Survey of digital image processing technology based on deep learning

Huaxiu Hong

Nanchang Normal University, Nanchang, Jiangxi, China

Keywords: neural network; image processing; deep learning; deep convolution.

Abstract: In the era of big data, the application domain of image information is extending, and the requirement of digital image processing technology is also improving. In recent years, with its powerful data recognition ability and high accuracy, in-depth learning has become a research hotspot in the field of digital image processing. This paper mainly introduces the technical framework of deep learning applied to digital image processing, introduces the deep convolution algorithm developed on the basis of machine learning neural network, and the application of deep learning image recognition, image HD and image understanding.

1. Introduction

With the advent of the era of big data, digital image, as an efficient information carrier, has been widely used in various fields of agriculture, industry, meteorology, medical and other national economy. On the one hand, with the continuous extension of the range of human activities, the application field of digital images is also expanding. On the other hand, in the era of large data, the production of massive data images makes people more and more efficient and performance requirements for digital image processing technology. One of the key technologies of digital image processing is image feature extraction and expression, and it is also one of the important technologies of efficient image processing. Deep learning forms a more dense high-level semantic Abstraction by combining low layer features, thus automatically discovering the feature representation of the data, and solves the problem of the need for artificial design features in traditional machine learning. In the field of image recognition, image classification and other digital image processing, breakthrough has been made. In recent years, deep learning has become an upsurge in the large data and artificial intelligence applications of the Internet. With its powerful data recognition ability and high accuracy, it has become a hot research field in the field of digital image processing.

2. The theoretical foundation of the development of deep learning technology

Deep learning is the further deepening of the field of artificial intelligence machine learning. As early as the Dartmouth conference in 1950s, the concept of artificial intelligence was put forward, the optimization of computer algorithm, the improvement of computer hardware ability, the production of large data, and the artificial intelligence entered a sTable and fast speed development stage. Artificial intelligence is a rich science. Among them, machine learning is very popular. Many tasks of AI processing data images first extract a suiTable feature set, and then provide these features to a simple machine learning algorithm. However, in daily life, many things are characterized by human intuition, and how do computers get these non-formal features? One of the ways to solve the problem is to use machine learning to discover the representation itself, and also to produce a series of machine learning algorithms, while deep learning is a specific type of machine learning. Figure 1 is the difference between machine learning and deep learning in the working principle of feature representation.

  DOI: 10.25236/mmmce.2019.012

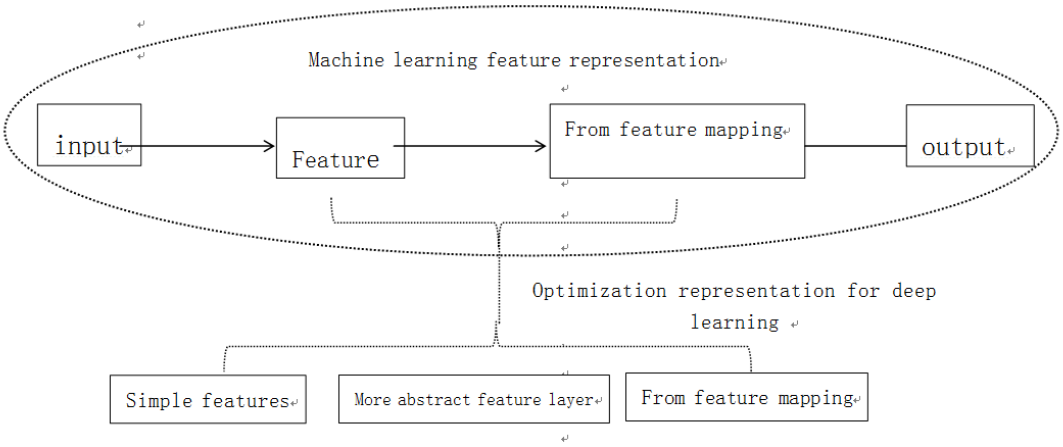

Figure 1: the principle of feature representation

(1) Foundation of machine learning

The combination of machine learning and image processing technology is an important driving force to promote the development of computer vision applications. It has been widely used in such aspects as license plate recognition, intelligent handwriting recognition, target recognition and so on.

The mathematical foundations of machine learning algorithms mainly involve linear algebra and probability theory. Linear algebra is one of the important tools to understand deep learning, and probability theory helps us understand the reasoning of deep learning system and analyze system behavior. In addition, information theory also helps us understand the concept of information quantity and analyze and quantify the information obtained. Machine learning often focuses on how to use computer to estimate complex functions statistically. Most machine learning algorithms belong to supervised learning or unsupervised learning algorithms.

Unsupervised learning algorithm The training dataset is not labeled, such as observing multiple samples of random vector x, trying to obtain probability distribution P (x). The typical algorithm is K- mean clustering, which can deal with image classification.

Supervised learning algorithm the training dataset contains many features. The dataset contains sample labels. For the given random vector x and its associated y, we try to predict y from X. The typical algorithms are support vector machines, decision trees and so on. In practical applications, sometimes the boundaries between unsupervised learning and supervised learning are not so strict. An important reason for the development of machine learning to deep learning is that these traditional machine learning algorithms have insufficient generalization ability in artificial intelligence.

(2) Deep learning algorithm

Deep learning is a new field of machine learning and belongs to the category of machine learning. The purpose of this study is to automatically extract multi level feature representation from massive data, and can extract feature data from the original image through complex nonlinear transformation, multilevel and multi angle, so as to obtain more accurate and enhanced expression and improve the efficiency of digital image processing. Therefore, the typical deep learning image processing model is a very deep neural network. By adding more layers and adding units into the layer, the depth network can represent the increasing complexity of the function and complete the mapping of the input vector to the output vector.

Deep learning puts forward higher requirements for mathematical algorithm and computer image processing capability based on machine learning.

3. Neural network

The neural network is a mathematical model or a calculation model for the approximate estimation of the function. It is calculated by a large number of artificial neurons, and the neural

network is a nonlinear statistical data modeling tool. Typical neural networks include structure, excitation function and learning rules. The neural network structure includes input layer, hidden layer and output layer (Figure 2).

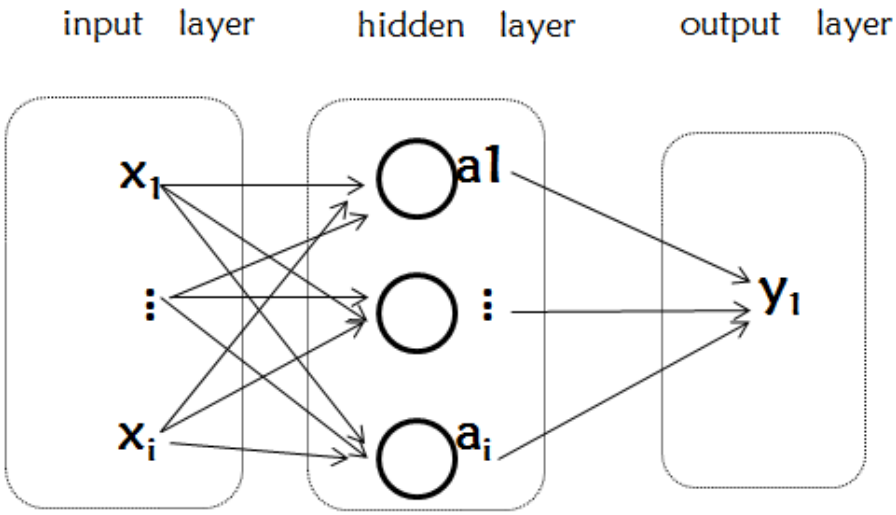

Figure 2 simple structure of neural network

4. Convolution neural network CNN

One of the problems that neural network may bring to image processing is “over fitting”. At present, the best way to solve over fitting is convolution, which can filter out unnecessary connections.

Convolution neural network has pushed neural network to deeper computation, and has played a great role in image recognition. By extracting the local areas of the image and fusing the information of these features into the subsequent processing stage, higher characteristics can be detected. The core of convolution neural network is convolution layer, pool layer and full connection layer.

The function of convolution layer is to perceive local and weight sharing. Convolution layer is composed of convolution core, convolution kernel is a two-dimensional matrix that can be smoothed and flipped on the image, and the same convolution layer is slid by multiple convolution kernel to extract data and feature operations. Example: suppose the picture of 5x5 is input, the convolution kernel is the matrix of 3X3, and every convolution kernel can be regarded as a neuron.

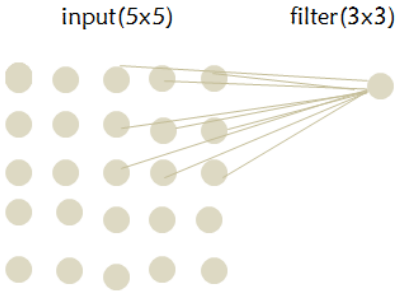

Figure 3: convolution kernel

Convolution operations mainly contain three important ideas: sparse interaction can reduce storage requirements, parameter sharing and equivariant representation. After the convolution layer is the excitation layer, it adds nonlinear mapping to the convoluted data. The pooling layer can further reduce dimension operations after the convolution operation. Finally, the full connection layer connects all the features and sends the output value to the classifier. A typical convolution neural network model has the following: VggNet,GoogleNET,Deep Residual Learning….

5. Application of deep learning in image processing

(1) Image recognition

Convolutional network was originally designed for image recognition and other problems, and now it is widely used in audio processing, text processing and other fields. In image recognition processing, the convolution neural network has become the research hotspot in the field of image recognition with its advantage of using three channels of color image directly as the input of the original data.

In 2012, Alex Krizhevsky proposed a deep convolution neural network model, AlexNet, for the first time using GPU acceleration operations, with 5 coiling layers, 3 maximum pool layers, 3 fully connected layers and 650 thousand neurons, winning the 2012 ILSVRC championship with absolute advantage. In 2014, the ILSVRC champion was GoogleNet, using 59 coiling layers, 16 pooling layers, and 2 fully connected layers, with more convolutions and deeper layers to get better structures. In the 2016 ImageNet Mass Vision Recognition Challenge Image Classification Contest, the accuracy rate based on depth learning algorithm exceeded 97%. It also establishes the dominance of deep convolution network in image recognition.

Image recognition technology is applied in various fields such as public safety, biology, industry, transportation, medical treatment and so on. For example, license plate recognition system, face recognition technology, fingerprint recognition technology, ECG recognition technology and so on. With the continuous development of computer technology, image recognition technology will be applied to more fields.

(2) Image high-definition

Improving image resolution and improving image quality is one of the hotspots in the field of image processing in recent years. The principle of image high-definition processing is: feature extraction of low-definition image, mapping low-definition feature to high-definition feature, and reconstruction of high-definition image. The model of deep convolution network is divided into three parts based on the traditional method: feature extraction, nonlinear mapping and feature reconstruction, and the three parts can be completed by convolution operation. In addition, we can calculate the loss function to evaluate the effect of image HD. Generating confrontation network (GAN) is a major change in machine learning recently. The typical applications are image high definition and image restoration. The main idea is to train two models, which are the generation of network G and the discrimination of network D. G is used to generate high definition image or image restoration task. After G generates the image, the generated image and the real image are put into D for classification and discrimination. In training process, the goal of generating network G is to generate real pictures as far as possible and deception to discriminate network D. The goal of D is to separate the G generated images from the real ones as far as possible. The discriminator D is a two classifier whose objective function is maximum likelihood estimation. The generation of G and D is a process of balancing each other's “game”. Based on the GAN framework, the generation of the network and the discriminating network can be defined, and the high-definition task of the image can be completed. GAN is a new and challenging field.

(3) image understanding

The core problems of image understanding are image classification, object detection, statement generation and image annotation, which are the results of convolution neural network and Recurrent neural network (RNN). RNN has an internal state in which the “state vector” is retained in the hidden unit, implicitly containing past input information about the sequence. When RNN accepts a new input, it will combine the hidden state vector with the new input, generate the output of the whole sequence, and cooperate with the CNN to form a more comprehensive and accurate understanding of the image. An important application is to first understand the original image through the convolution neural network and convert it into a distributed representation of the semantic, and then the recurrent neural network converts the advanced representation into a natural language.

6. Summary

The most successful of these years is the convolution neural network. Because of the characteristics of the image data, convolution and pooling are the two most important operations in the domain of image processing in the depth convolution network (Figure 4 is the technical framework for depth learning in the image processing field). In the current field of deep learning, convolutional neural network is the dominant depth algorithm in the present and future. With the further improvement of computer parallel computing ability, deep learning will appear in more and more fields.

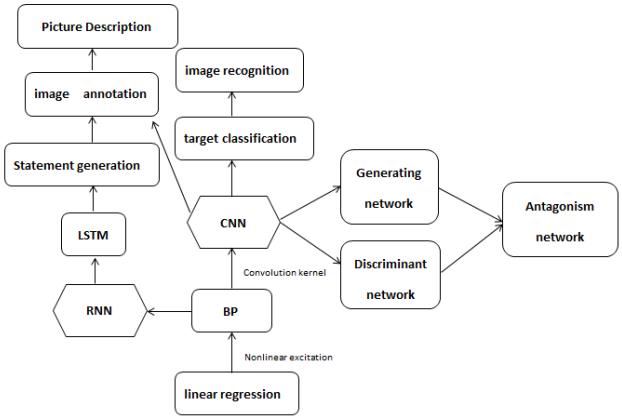

Figure 4 Technical framework about Deep Learning

References

[1] Wei Zhang,Shuo Yu. A Review of Digital Image Processing [J]. Communication World, 2015, (18): 258-259.

[2] Zhang Wei, Yu Shuo. Summary of Digital Image Processing [J]. Communication World, 2015, (18):258-259.