An Overview of Image Recognition Based on Deep Learning

Chen Changhui

Institute of General Aviation Industry, Fujian Chuanzheng Communications College, Fuzhou, 350007, China

Keywords: deep learning, image recognition, network structure, development and application

Abstract: Accurate image recognition is of great importance in academia. Image recognition technology plays an important role in many aspects such as medicine, aerospace, military, industry and agriculture. It is also an important branch of artificial intelligence research and one of the most widely used artificial intelligence technologies in the current world. Deep learning has made a breakthrough in applications such as image recognition and speech recognition, and has also shown unique advantages in the fields of face recognition and information retrieval. For instance, the X-ray images of bones show the changes of different levels of black, white and gray, and have the image characteristics of black-and-white contrast and level difference. Due to these advantages of deep learning in the image-recognition field, we organically combine it with bone age assessment research, aiming at providing basic data of the construction of forensic automatic bone age assessment system. In recent years, the application of deep-learning based image recognition has become a research hotspot in the field of image recognition, which has achieved good results and has broad research space.

1. Introduction

Image processes the most important position in all kinds of information. It contains a large amount of information and is most directly accepted by human. Compared with sound and word, image is more straightforward. Therefore, image processing has practical research significance and wide application prospect. Remote sensing data has been introduced into its application fields as basic data by the national, agriculture, forestry, construction, environment, ecology and other industries, and these applications have promoted the development of remote sensing technology to a certain extent [1]. In recent years, the emergence of high-resolution remote sensing images has made the spatial information of images more refined and much richer, and image data has accumulated tremendously. The high-resolution image data obtained worldwide exceeds 1.8 billion square kilometres annually, which is 12 times of the surface area of the earth. However, the remote sensing information that people can effectively use only accounts for less than 5% of the total information obtained. Bone X-ray image recognition belongs to the category of image recognition. Therefore, using deep learning to realize bone image recognition is a reliable way to realize artificial intelligence bone age evaluation system. This paper reviews the research progress of domestic and foreign academia using deep learning in image recognition and bone age evaluation in recent years [2].

The purpose of image recognition is for a well-trained computer to extract information from the image to be detected and identify. The reason why image recognition has been deeply researched and developed rapidly is its extensive and important application value. At present, the image recognition technology is mature and widely used in the recognition of faces, numbers and other objects. It is also used in agriculture, commerce, military and daily life [3]. At present, in the field of security, the intelligent functions that users pay attention to include moving target detection, target classification, target color recognition, pedestrian detection, face detection, face recognition, vehicle detection, vehicle type recognition, vehicle identification, license plate recognition traffic statistics, density detection, and etc. These intelligent modules are derived from image recognition is one of the core technologies of remote sensing application, and it is also the basis and premise for realizing various practical applications. How to classify large-scale logical images quickly and

accurately has become a hotspot in the field of remote sensing technology application research. Indepth learning adopts unsupervised learning, and the sample label is unknown in the training process. This learning method does not need manual participation. In the modern information age, Large samples and large data sets are becoming more and more popular. There is no doubt that deep learning will increasingly affect people's lives and become a research hotspot [4].

2. Deep Learning Technology

2.1. Deep Learning Concepts

Deep learning is a revolutionary achievement in the field of artificial intelligence research, which originates from the research of ANN by human beings. Based on the simulation of the information processing mode of human cerebral cortex and visual nervous system, by setting perceptrons with different number of hidden layers, the provided big data and obtain appropriate connection weights to extract the features of the original data can be trained from low level to high level step by step, such that the computer can realize the function of hierarchical processing and understanding data, and finally achieve the purpose of obtaining the required information. Deep learning was first applied in the field of image recognition, and has achieved great success in image classification, face recognition, general object detection, image segmentation, optical character recognition and other fields. The concept of deep learning originates from the research of computer science, ANN and artificial intelligence for more than 30 years [5]. These shallow learning methods rely on artificial experience to extract sample features. The obtained image features are the image representation features. The deep neural network model is an artificial neural network composed of several single-layer nonlinear network structures. It is an active, high-input understanding memory, which involves the strong transfer of high-order thinking, learning results, learning state and learning process. Figure 1 shows the general flow of image recognition.



Fig. 1 The general flow of image recognition

2.2. The Basic Idea of Deep Learning

Deep learning can fully and automatically learn abstract knowledge expression. In about ten years of its emergence, it has shown great application potential in the fields of speech recognition, image cognition, natural language processing and industry attention, such as fraud detection and recommendation system. The principle of deep learning is the mechanism of human brain's visual layered processing of information. The working principle of human brain is to recognize objects through the information processed by the aggregation and decomposition process of each layer of the brain. The earliest model of machine learning is shallow learning structure, followed by deep learning structure. Dividing machine learning into shallow learning and deep learning is determined by the hierarchical structure of learning network [6]. The network structure of shallow learning is not as many as that of deep learning. In addition, the relationship between levels is different. In the convolutional neural network, the image is sub-sampled according to the principle of local correlation in the spatial relationship of the image, and each neuron is connected to an adjacent local area of the previous layer, which can reduce the number of connections. The effect of optimizing the network structure. The more layers, the more the essence of objects can be abstracted. However, the increase of network layers causes the delay of training time. Furthermore, the increase of training layers does not guarantee the accuracy of the algorithm is improved. Figure 2 is a simplified convolutional neural network structure.



Fig. 2 Structure diagram of convolutional neural network

2.3. Deep Learning Network Structure

Feedforward deep network makes the data from the input layer one-way through one or more hidden layers to reach the output layer, including multilayer perceptrons, convolutional neural networks, etc., and its development and formation time is relatively early. The deep convolutional neural network requires that the input image to be classified is a standard rectangle of the same size, and the image objects obtained after segmentation are irregular polygons with different sizes, so the image objects need to be standardized before network training [7]. The size of the rectangular image enables the image object to be converted into a data input format that the network can accept, and minimizes the loss of shape information during the image stretching process. Mxnet is a fully functional, programmable and highly scalable deep learning framework developed by the distributed machine learning community. It has a few main advantages: 1) it has the ability to scale multiple GPUs; 2) it has superior programmability and development speed; 3) the portability between a wide range of devices and platforms is good. The important idea of deep learning is the "depth" of its model, that is, the network structure contains multiple hidden layers, not just from input to output, and the number of hidden layers in the structure model is generally greater than two. The feedback depth network is formed by superposition of multiple decoders to decode the input data, including deconvolution network, hierarchical sparse coding network and so on. Bidirectional depth network is composed of multiple encoders and decoders superimposed on each other, including depth confidence network, depth Boltzmann machine, trestle self encoder, and etc [8]. The structure of neural network is shown in Figure 3.



Fig. 3 Structure diagram of neural network

3. Development and Application of Image Recognition Technology

Graphic stimulation acts on sensory organs. People recognize that it is an experienced graphic, and can even perceive the change of distance or shape from the image. This process is called image recognition. In image recognition, there is not only the information that entered the senses at that

time, but also the information stored in the memory. The recognition of the image can only be achieved through the process of comparing the stored information with the current information [9]. This model does not indicate that the person is How to identify and process similar stimuli is also difficult to realize in computer programs. Therefore, someone proposed a more complex model, namely the "pan-magic" recognition model.

For the convolution layer, the convolution process of convolution neural network can be expressed as:

$$x_{j}^{l} = f\left(\sum_{i \in M_{j}} x_{i}^{l-1} * k_{ij}^{l} + b_{i}^{l}\right)$$
(1)

It can be seen that each feature map has a bias b, but since the convolution kernel is different, the output feature map is also different.

Given a characteristic diagram, we can find out its deviation gradient about bias, that is, sum:

$$\frac{\partial E}{\partial b_j} = \sum_{u,v} \left(\delta_j^i \right)_{uv} \tag{2}$$

The min max method is used to normalize the original image, and the normalization formula is:

$$DN_{nom} = \frac{DN}{DN_{max}}$$
(3)

Among them, DN is the pixel gray value of a certain band in the original image, DN_{nom} is the normalized value, and DN_{max} and DN_{max} are the maximum and minimum values of the corresponding bands in the original image, respectively.

The earliest research on image recognition in China was in the 1960s. In the early stage of the research, there was basically no great progress, and only preliminary research on number and letter recognition was carried out, not even the prototype. In the 1980s, the domestic recognition technology first stepped onto the international stage, and established its own key national laboratory of pattern recognition. Of course, pattern recognition is the parent class of image recognition. This period of time has greatly improved the Chinese people's own image recognition research ability. The extraction of basic image features includes the extraction of edges and corners. Generally, different feature extraction operators combined with corresponding thresholds are used to get these key points. Another method of feature extraction in frequency domain is to divide the image into different parts based on frequency by Fourier transform, so that the gray level change of the original image can be reflected in the frequency spectrum, and then the outline and edge of the image can be obtained [10].

At present, the popular basic algorithm of image recognition is deep learning method. Although deep learning has many advantages, it still has many shortcomings. First, because the deep learning model is a non convex function, its theoretical research is very difficult and lack of theoretical guarantee. When adjusting the data, it is still a simple "trial and error" and lacks theoretical support. At the same time, deep learning relies too much on data and computing resources. Learning a new concept often requires hundreds or even more marked samples. When the marked samples are difficult to obtain or the cost is over high, deep learning is not able to achieve good learning results. Moreover, deep learning requires a very expensive high-performance GPU, which makes deep learning difficult for civilians. At present, the speed of deep learning training is slow, which often takes several days or even a month. The model has poor scalability and lacks the ability to draw inferences from others. If the sample changes slightly, the system performance will drop rapidly. At present, deep learning is a static process and lacks interaction with the environment.

4. Conclusions

This paper studies deep learning technology and introduces the development and application of image recognition technology. Deep learning has brought a new type of solution to complex image recognition technologies such as face recognition, and some IT companies have achieved excellent "running scores" by applying deep learning. From traditional image recognition methods, to the rise of machine learning, to the rise of deep learning, the reason why image recognition technology can be updated like mushrooms is precisely because of the accumulation of theories and ideas. These theories and ideas have been proposed for several years, but due to the limitations of objective hardware, they have not been put into practice. On the basis of large-scale data and high-speed operation, these theories and ideas finally have a platform for them to show their strength, which makes image recognition develop rapidly in recent years, and even makes deep learning play an absolute leading role in computer vision. The result is also incredible. The performance of many neural networks has exceeded the human average.

Acknowledgements

Natural Science Foundation of Fujian Province project "Research on method and Theory of Emergency Plan Generation Based on Case Retrieval and Adjustment" (2021J01339)

References

[1] Hu Tinghong, Wan Lei, Liu Taiang, et al. The advantages and application prospects of deep learning in image recognition and bone age assessment. Journal of Forensic Medicine, vol. 33, no. 6, pp. 7, 2017.

[2] Li Bingzhen, Liu Ke, Gu Jiaojiao, et al. Research and application of deep learning in image recognition. Electronic Paradise, no. 3, pp. 1, 2021.

[3] Luo Xi. On the application status and advantages of deep learning in the field of image recognition. Science and Technology Information, vol. 18, no. 3, pp. 2, 2020.

[4] Ma Lele, Li Zhaoyang, Dong Jiarong, et al. UAV Gesture Control System Based on Computer Vision and Deep Learning. Computer Engineering and Science, vol. 40, no. 5, pp. 8, 2018.

[5] Liu Cuiming, Zhao Dongxu. Using deep learning drone images to quickly find illegal construction in bulldozing areas. Urban Survey, no. 4, pp. 4, 2020.

[6] Huang Xianfeng, Zhang Fan, Gao Yunlong, et al. Fusion Technology of Deep Learning and UAV Aerial Photogrammetry. China Disaster Reduction, no. 11, pp. 1, 2020.

[7] Li Chunxiao, Yin Zhenhong. Application Research of Deep Learning in Image Recognition. Satellite TV and Broadband Multimedia, no. 19, pp. 2, 2019.

[8] Li Chunxiao, Yin Zhenhong. Application Research of Deep Learning in Image Recognition. Satellite TV and Broadband Multimedia, no. 19, pp. 2, 2019.

[9] Liu Jie. Identifying Igneous Rocks and Fractures Using UAV and Deep Learning. Natural Gas Exploration and Development, vol. 42, no. 4, pp. 1, 2019.

[10] Cheng Jinfu, Zhang Wanhe. Research and Application of Deep Learning in Image Recognition. Electronic World, no. 19, pp. 2, 2020.