

Deep learning in image recognition

Yanxiang Huang¹

¹Department of International College of Chinese Studies, Fujian Normal University,
Fuzhou, Fujian, 350000, China

154032020049@student.fjnu.edu.cn

Abstract. With the rapid development of high-tech in China, the field of artificial intelligence has also been promoted. Image recognition is an important topic in the field of artificial intelligence, which mainly includes two modules: classification recognition and feature extraction; At the same time, as an important research direction of AI, deep learning has made rapid progress in recent years. It has been widely used in image recognition, speech recognition and other fields and has achieved great success. This paper analyzes the application of deep learning in image recognition, mainly from the aspects of face recognition, remote sensing image classification, etc. Its purpose is to provide help for relevant practitioners, so as to promote the development of image recognition in the tide of AI development.

Keywords: Deep learning, image recognition, application.

1. Introduction

Deep learning, which is the creation of deeper neural networks, is used to simulate the brain and is used for analysis and can also be used for understanding and comprehension. Image recognition, on the other hand, is the use of image processing, analysis and understanding of images for the purpose of recognising multiple image forms and objects [1]. Among the current image processing techniques, deep learning techniques have gained common usage in image processing. Therefore, the authors present an in-depth study of the basic connotations and constructive advantages of deep learning and its application to image recognition in order to be useful to practitioners in related fields.

2. Overview of traditional image recognition

In recent years, deep learning has become one of the hottest research directions in the world today due to the increasing socio-political and economic development of China. Advances in deep learning have also greatly influenced the field of computer vision and mechanical learning, and at this stage, new deep learning techniques are proliferating, which have a significant impact on people's working environment. This paper provides an in-depth discussion on how to effectively apply deep learning in image recognition.

There are two types of features, local features and global features. Global features are based on colour features and texture features, of which the most common ones include PCA-SIFT and SURF, SIFT and HOG. In general, a combination of local and global features can be used to better portray the image [2].

3. Overview of deep learning

3.1. The meaning of deep learning

In essence, deep learning is an emerging direction in machine learning. The ultimate goal of deep learning is to understand the internal rules and level of expression of a sample and to interpret the data in such a way that the computer is equipped to analyse and learn in the same way as a human being, and thus recognise data such as text, sound and images. Deep learning, as an advanced machine learning method, has promising applications in both speech and image recognition. The introduction of deep neural networks in image recognition can effectively overcome many complex pattern recognition problems, thus greatly contributing to the development and evolution of images [3]. Figure 1 shows a conceptual diagram of deep learning

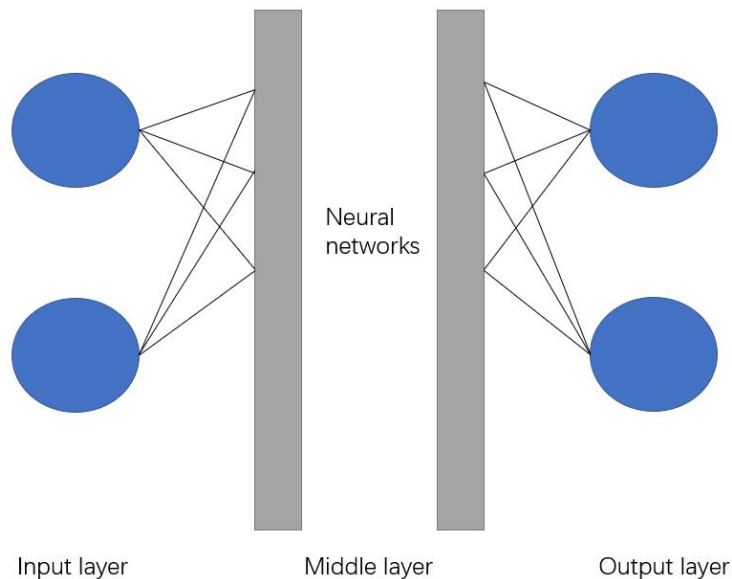


Figure 1. Concept map of deep learning.

3.2. Principles of deep learning

Deep learning is a machine learning method first proposed by Hinton and others, originating from the study and exploration of ANNs. In short, it is a simple merging of the basics of deep learning to form more complex, abstract and not clearly definable features. At the same time feature learning in deep learning is very similar to the principles of vision in the human brain [4]. The human brain's vision principle consists of the following: light acts on the visual organs through reflection, the pupil recognises pixels projected into the lens; the cerebral cortex carries out recognition and identification of colour and direction, after which the brain makes judgements about objects and eventually feeds back information. Deep learning in humans is a method of learning based on the description of information. The learning method can be simply explained by adjusting the hierarchy of the deep learning neural network to achieve maximum accuracy for training purposes. After gradient descent has been performed, learning has been carried out using statistical methods. As mankind continues to develop, it is being used more and more in areas such as automatic control, autonomous driving and computer vision.

3.3. Development of deep learning

At present, the development of machine learning techniques can be divided into two stages: a surface level research stage and a deep level research stage. Hierarchical learning has been studied in depth

because of its inability to deal effectively with the abstraction of structures and its inability to be formalised. The rapid development of deep learning techniques has had a positive effect on many aspects. For example, the use of convolutional neural networks can contribute well to the development of image recognition techniques. In addition, in speech recognition, the introduction of deep learning techniques has made a significant contribution to further improving the accuracy of speech. In addition, it has also achieved good results in the recognition and detection of road traffic signals, surpassing the human classification of road traffic signals. In the future, artificial intelligence will become one of the important research directions, and with the continuous improvement of science and technology, its development will become faster and faster. At the same time, artificial intelligence can also be combined with other disciplines to form a new body of knowledge, making it possible to give full play to all areas of artificial intelligence and to bring its unique advantages and value into play in all walks of life.

3.4. Prospects for deep learning

Currently, with the rapid development of deep research, its application in various industries and fields is also increasing. However, at present, researchers in depth cognition are still in a period of exploration and many problems need to be solved [5]. For example, in terms of model construction, the human brain is organised in three dimensions, taking into account both a planar hierarchy and a vertically oriented layout; however, at present, the network architecture used is still only a two-dimensional one. Therefore, a relative spatial structure model is necessary to explore deep learning problems in depth. In addition, the scientists involved would like to be able to maintain the speed of flight while improving training accuracy. However, because deep learning is complex, has many layered structures and is subject to vast amounts of information interference; their training, in all likelihood, will be extremely limited. Therefore, how to improve the efficiency of training while ensuring its effectiveness remains to be further explored and explored.

4. Application of deep learning in image recognition

4.1. Face recognition

At present, face recognition technology based on deep learning has been widely used in the recognition of images, and the biggest difficulty lies in its ability to distinguish different features from the many influencing factors. There are many factors that influence the recognisability of an image, including lighting, expression, identity, etc. However, the variations caused by these factors are also non-linear to a certain extent, and such non-linear variations are so complex that they cannot be analysed using conventional linear models. In the case of facial images, new properties can be obtained thanks to the use of multiple levels of non-linear transformation, which allows for better discrimination of the associations caused by various influences. Deepface is a face recognition algorithm based on DeepFace, which uses more than 200 neural structures to learn face features. The DeepID algorithm has excellent robustness and sparsity. Figure 2 shows the flow chart of face recognition.

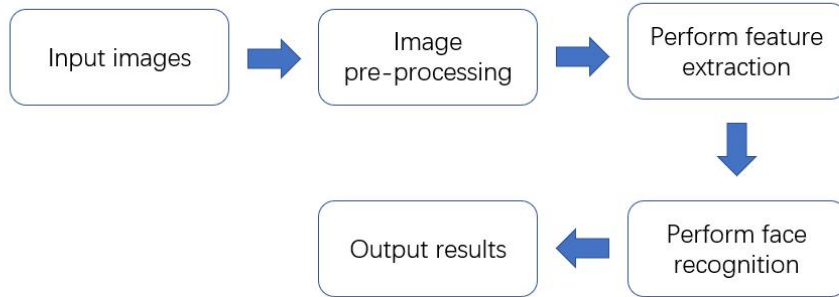


Figure 2.Flow chart of face recognition.

In face recognition, LFW data set was completed by a laboratory at the University of Massachusetts in 2007. It is a face recognition data set of unconstrained natural scenes, mainly applied to face recognition under unconstrained conditions. The LFW data set has 13323 face images of global celebrities, collected in different scenes. This dataset contains more than 5000 people's face images, of which 80% have collected one image, and the remaining 20% have collected more than two images. Each face image is marked with a different ID number. This dataset is the most commonly used method in training face recognition algorithm validation.



Figure 3. Sample LFW face image database.

VGGNet16 is a 16-layer planar convolutional network, including 13 layers of convolutional layer, 2 layers of full connection layer and 1 layer of Softmax layer. ResNet18 is a residual convolution network with an 18-layer structure, including 1 layer of convolution layer, 4 residual modules (each residual module consists of 2 layers of convolution layer) and 1 layer of Softmax layer. In order to compare the performance of the improved convolutional neural network with other networks, three network models are trained in the same experimental environment. Figure 4 shows the accuracy and loss function values of the three network models in the test set. At the same time, to calculate the time complexity of the model, this paper first counts the time required for model training, and then divides it by the number of iterations [6].

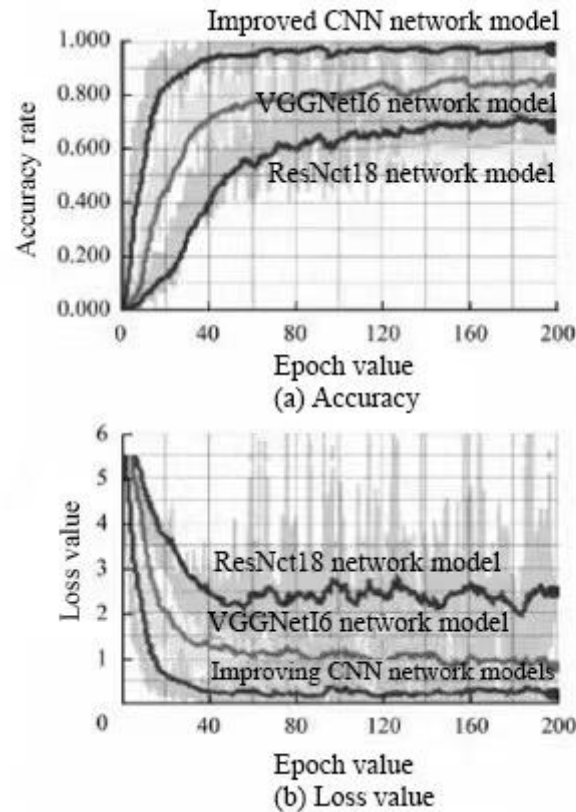


Figure 4. Prediction accuracy and loss value of network model [6].

4.2. Remote sensing image classification

Remote sensing images have important applications in a variety of fields because of the vast amount of information they contain. The redundancy caused by too much information due to the excessive amount of image data, and the fusion between various information due to the poor resolution of the images. Because of the difficulty in identifying remotely sensed images, it is difficult to segment meaningful image information correctly using conventional image classification techniques. In the classification of remotely sensed images, better classification results can be obtained by using appropriate deep learning models in conjunction with specific optimisation methods; this is the current direction of development of remotely sensed images.

4.3. ImageNet classification

Deep learning is also used extensively in ImageNet, where in many cases the results obtained by traditional computer vision techniques are highly biased and there is a large number of errors in the set of tests. Compared to traditional convolutional nets, AlexNet has the following advantages: firstly, AlexNet uses a Dropout approach to classification and classifies the neurons in it to zero, thus achieving a simulation of human nerve cells, although this process is slower, but results in a more robust network model [7]; secondly, AlexNet is able to reduce in a sense the amount of operations and can obtain neurons with sparse properties.

4.4. Traffic image recognition

In the rapid development of China, image recognition technology has been widely used in the transport industry. In the fields of vehicle number plate recognition, lane deviation warning, road recognition and road identification, vehicle recognition technology has brought great convenience to people's lives. The recognition technology of traffic images has also been widely used in areas such as

intelligent parking, toll management and traffic control. In recent years, some scholars have tried to introduce deep learning techniques into deeper road traffic signal recognition, for example, deep convolutional neural networks are used in the detection of road signals; the method is able to obtain fast and accurate recognition methods for road signals. On this basis, a new road traffic image detection technique is proposed. Figure 5 shows the flow chart for licence plate recognition

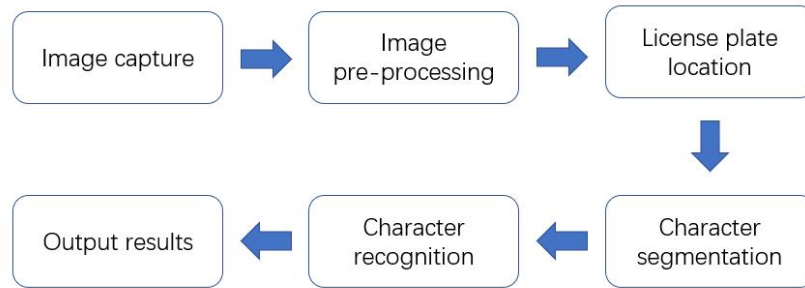


Figure 5. Flow chart of licence plate recognition.

4.5. Character image recognition

Character image recognition has been well used in many areas. However, early techniques for character image recognition mainly relied on manual pre-processing of characters, and the effectiveness and credibility of this method were significantly reduced. With the development of deep learning technology, it has been applied to the recognition of text images with better results, greatly reducing the error rate of text and narrowing the distance between computers and humans. It is also possible to improve the image recognition accuracy by feature extraction of the image, thus achieving good image recognition results.

4.6. Video image analysis

Although some research results are available, they are still in their infancy in terms of current practical applications. On ImageNet, we can use deep learning methods to portray still video images to obtain the corresponding deep learning patterns. In contrast, in deep learning, the inscription of motion characteristics of video has been a big challenge. Since in traditional view analysis, optical flow estimation and dynamic texture are usually used to portray dynamic characteristics, but the representation of such dynamic characteristics rarely reflects such dynamic characteristics [8]. The key to this problem is the failure to differentiate the spatial dimension from the temporal dimension, firstly by treating the image as a stereo image and then applying it to the convolutional network; the second approach is to use the spatial distribution of optical flow and other dynamical properties as the output channel of the convolutional network by pre-processing them; the third aspect is to use long and short term memory to obtain long time correlation information that allows a good to model complex motions. All three methods have their advantages and disadvantages, but they can all be used to process some specific types of data. For example, in the field of medical imaging, where there is a degree of uncertainty in information such as a patient's medical history, a classification algorithm based on a deep neural network structure is needed to enable better diagnosis.

5. Problems and future research directions

Despite the many advantages of using deep knowledge for image recognition, there are still a number of issues that need further research. The first reason is that deep learning mines large amounts of data and information, but the amount of effort and energy required to classify and label large amounts of data on a large scale limits the efficiency of learning. Secondly, as the environment in reality is uncontrollable and cannot be replicated, it is not possible to analyse the role of each element individually. Thirdly, due to the observations of the human eye, there is a high risk of error when tagging data information and therefore correct tagging cannot be made. In order to better address this issue and find a deep learning method with a strong scientific and theoretical approach, it is necessary to create a deep learning framework that can fuse different perceptual data together. The fourth point is that, with the development of computer technology, people have started to use a variety of intelligent devices to collect, process and store massive amounts of data, all of which make traditional machine learning methods face challenges, and how to use the powerful computing power of computers to improve efficiency has become an important research direction. The fifth point is that deep learning-based recognition methods mainly include convolutional neural networks, recurrent neural networks, etc. The most typical ones are BPNN algorithms and deep belief networks, but there is still a gap between BPNN and deep learning.

6. Conclusion

In conclusion, current deep research is widely used in various aspects and has been very effective in practice, especially in image recognition. Deep learning emphasises learning the internal rules and representational aspects of a sample, with the ultimate aim of giving machines similar learning and analysis to that of humans, accurately recognising information such as text, images and sound, while deep learning is a more advanced approach to machine learning. Therefore, image recognition techniques based on deep learning techniques are explored in depth, providing insight into their deeper benefits and scope of application, with a view to promoting the rapid development of image recognition techniques.

References

- [1] Xue Y, Wang LY, Zhang Y, et al. A GoogLeNet-based deep migration learning method for apple defect detection[J]. Journal of Agricultural Machinery, 2010(8):1-8.
- [2] Sadique M F , Biswas B K , Haque S M R . Unsupervised Content-Based Image Retrieval Technique Using Global and Local Features[C]// 2019 1st International Conference on Advances in Science, Engineering and Robotics Technology (ICASERT). 2019.
- [3] Peng X.Y., Li L. Application of deep learning neural network for image recognition system[J]. Fujian Computer, 2020, 36(1):65-67.
- [4] LeCun, Yann, Bengio, et al. Deep learning[J]. Nature, 2015.
- [5] Ma F. River management: deep learning-based image recognition and prediction of river floating materials[J]. Regional Governance, 2019(46):146-148.
- [6] Dong Luocheng, Chen Zhangping Face recognition algorithm based on improved convolutional neural network [J] Journal of Hangzhou University of Electronic Science and Technology: Natural Science Edition, 2018, 38 (5): 6
- [7] Fang Hao-Wen, Shi Hua-Jun. A deep learning-based classification method for satellite image recognition[J]. Computer System Applications, 2019, 28(10):27-34.
- [8] Wang J. A study of adversarial samples in image recognition problems based on deep learning[J]. Computer Knowledge and Technology, 2019, 15(28):222-223.