Transactions on Computational and Scientific Methods | Vo. 5, No. 2, 2025
ISSN: 2998-8780
https://pspress.org/index.php/tcsm
Pinnacle Science Press

Advancements and Challenges in Image Recognition: A Comparative Analysis of Traditional Neural Networks and Convolutional Neural Networks

Daxton Rivera[1], Maddox Ellsworth[2], Orson Winslow[3]
[1]School of Electrical and Computer Engineering, University of Iowa, USA
[2]School of Electrical and Computer Engineering, University of Iowa, USA
[3]School of Electrical and Computer Engineering, University of Iowa, USA
*Corresponding author: Daxton Rivera, daxton47@uiowa.edu

Abstract: Image recognition technology has evolved significantly since its inception in the 1940s, with notable advancements occurring in the 1990s through the integration of artificial neural networks and support vector machines. While traditional image recognition methods relied on shallow hierarchical structures and manual preprocessing, neural networks have revolutionized the field by enhancing accuracy and adaptability. This paper analyzes the advantages and limitations of both traditional neural networks and convolutional neural networks (CNNs), exploring their principles and real-world applications. With the increasing complexity of image recognition environments, standalone CNN techniques face challenges such as color similarity between objects and backgrounds, overlapping targets, and extreme occlusion. To address these limitations, integrating multiple CNN models and designing specialized neural network structures for different object types can further enhance recognition accuracy.

Keywords: Deep Learning; Convolutional Neural Network; Image Recognition.

1. Introduction

Image recognition technology originated in the 1940s. At that time, due to the lack of technology and imperfect hardware facilities, image recognition technology did not develop rapidly. Until the 1990s, the combination of artificial neural network and support vector machine promoted the development of image recognition technology. Image recognition technology has been widely used, such as license plate recognition, face recognition, object detection and so on. However, the traditional image recognition technology is mainly based on the shallow hierarchical structure model, which requires artificial image preprocessing, thus reducing the accuracy of image recognition. Neural network algorithm plays an important role in the field of image and character recognition, such as fraud detection, face recognition, the judgment of medical pathology, etc., can also be on stock prices, economic trends, etc on the generalized model for forecasting, it overcomes the traditional model can't predict the complex nonlinear model of faults, so more and more scholars engaged in this field.

2. The principle and development process of deep learning

The principle of deep learning is simply a kind of control principle based on multi-level structure; It's a tree-like manipulation and control mode; It's a principle of deep control and recognition. To be specific, we have A three-layer control structure: S, S1 and S2. The input image A passes through the recognition and extraction of S layer to transmit the result to S1 layer, and S1 layer transmits the result to S2 layer, and each layer passes through the recognition and extraction. And then we get B.

B is the final result extracted by A after feature recognition of each layer. The error between the result of each layer and the original image of A is very small, but after reaching B, the feature will be more specific, and the image B which is easy to judge the feature will finally be obtained. It does not mean that in-depth learning is just a simple three-layer method. In-depth learning is a judgment method that covers at least three layers of judgment, which can be divided into countless layers. The more layers, the higher the accuracy of recognition. This is the basic principle of deep learning. Deep learning, on the other hand, mimics the thinking patterns of the human brain to create a similar level of thinking. This method was first proposed in the 1980s, published in scientific journals in 2006, and achieved great success in image recognition in 2012. [1]

2.1 Limited Boltzmann machine

The limited Boltzmann machine is proposed by Smolensky on the basis of Boltzmann machine, which is composed of dominant variables and hidden variables. Among them, there is a mapping relationship between the dominant variables and the recessive variables, but there is no internal connection between them. Its training method is mainly a fast learning algorithm based on contrast divergence.

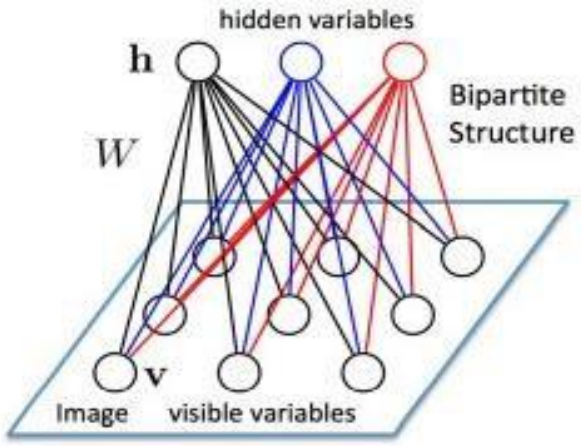

Figure 1. An overview of the restricted Boltzmann machine

2.2 Automatic encoder

Automatic encoder, proposed by Rumlhart in 1986, is mainly composed of an encoder and an interpreter [3]. The role of the encoder is to transmit the compressed representation of the input signal to the next layer of the network, and the role of the interpreter is to interpret the compressed reconstructed data signal of the encoder and transmit the output signal. These two parts are essentially the same thing, both of which apply some transformation to the input signal to restore the input signal.

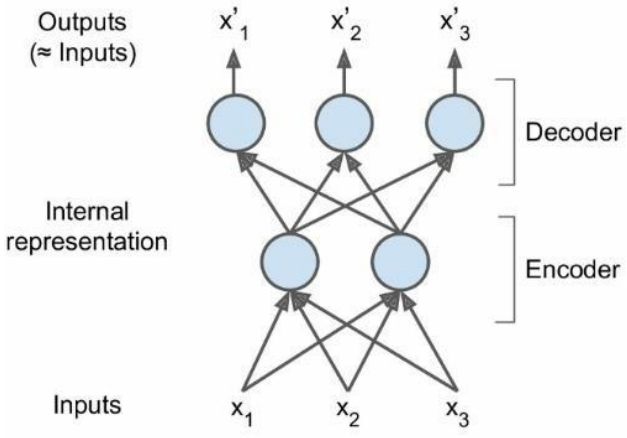

Figure 2. An overview of the automatic encoder

2.3 Convolutional neural network

A convolutional neural network usually consists of a convolutional layer, a downsampling layer, a full connection layer and an output layer. Its feature is that the original signal is directly used as the input of the network, which avoids the complex feature extraction and image reconstruction process in the traditional recognition algorithm. Among them, the function of convolutional layer is to extract features. For an image, a convolution layer contains multiple convolution kernels, every convolution kernels with input image convolution operation produce new image, new of each pixel in the image that represents the convolution kernels is covered inside the small area of the image of a trait, said with multiple convolution kernels respectively for convolution can be different kinds of feature extracting. As a deep feedforward neural network, convolutional neural network is most commonly used in supervised learning problems in the image field, such as image recognition and computer vision.

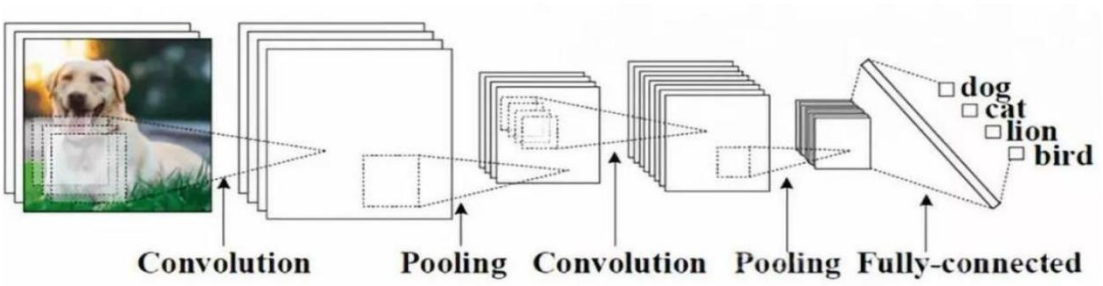

Figure 3. An overview of a convolutional neural network

3. Image recognition classification

The purpose of image recognition is to divide different types of images into different categories, and the convolutional neural network (CONVOLUtional neural network) is introduced to achieve the minimum classification error and the highest recognition rate. On the whole, image recognition problems [2] can be divided into cross-species semantic level classification, subclass fine-grained image classification and instance level image classification. Cross-species semantic level classification is to identify different kinds of objects. This kind of algorithm is characterized by small intra-class variance and large inter-class variance. Subclassification of subclass is realized in the same category. This kind of detection requires more classifier. On the basis of the first two categories of image classification, instance-level image classification also needs to recognize different individuals. At present, the image classification technology based on instance level is widely used, which can meet the requirements of precision and speed.

4. Image recognition algorithm

4.1 Traditional detection algorithm

Traditional image recognition algorithms include differential operator edge detection algorithm, Canny edge detection algorithm, corner detection algorithm and so on. This paper will elaborate on the latter two common algorithms

1) Canny edge detection algorithm generally includes four steps [3]: filtering, gradient amplitude and gradient direction calculation, non-maximum inhibition calculation, edge detection and connection. First by gaussian filtering function to remove the image noise, and to smooth the image, and then through a first order finite difference method, respectively, after filtering of the image pixels in the direction of the horizontal and vertical partial derivative model, then using the maximum suppression algorithm will be outside of the local maximum plus or minus gradient value is set to 0, at last, by two different threshold for the candidate edge pixel in image processing, keep two pixels within the

scope of threshold value, eventually detect objects. The traditional Canny edge detection algorithm has poor noise reduction ability. In view of this situation, literature [4] proposed to use four 5-order difference templates with anisotropy to detect pixel points in multiple directions at the same time, which can not only detect the gray weight value of the upper, lower, left and right fields, but also detect the value of the diagonal direction. In order to improve the adaptive ability of Canny algorithm, the literature [5] by using adaptive median filter and morphological closing operation to prevent more direction gradient amplitude to calculate the edge information is weakening, at the same time use the best separation point target and the background is the optimal gradient under the maximum and minimum variance between the variance within the class of this concept, to calculate upper and lower threshold of the Canny algorithm, in order to increase its adaptive abilities.

2) Corner detection algorithm [6] is to slide the image in any direction through a fixed pixel window and compare the gray value of pixels in the Windows before and after sliding. If there is a large change, it can be judged that there are corner points in the pixel. Corner detection algorithms fall into three categories [7]: corner detection based on binary image, corner detection based on gray image and edge contour

The detection of corner points. The traditional Harris corner detection algorithm has low accuracy and poor anti-noise performance. In literature [8], Sobel algorithm is combined with Harris algorithm. Sobel algorithm is first used to conduct corner selection, and the rectangle in the non-maximum suppression algorithm is reduced

In order to improve the detection accuracy, the circular template was used instead of the template. Finally, the adjacent point elimination method was used to improve the noise resistance of the algorithm. Literature [9] by comparing step edge, l-type corner, X, Y, or T corner corner and star corner strength change characteristics, puts forward using multi-scale anisotropic directional derivative gaussian filter, extracted from the input image grayscale change new method, the method can continuously extract the edge points in the image and angular point features.

4.2 Neural network algorithm

The neural network algorithm established by simulating the operation of human brain is a nonlinear dynamic system. The network system composed of a large number of input neurons can achieve the goals that can not be achieved by traditional algorithms. Many machine learning tasks that once relied heavily on hand-extracting features, such as target detection, machine translation and speech recognition, have been revolutionized by a variety of end-to-end neural network learning algorithms. Neural network learning algorithms [10] are mainly divided into five categories: convolutional neural network, attentional neural network, self-coding neural network, generative network and space-time network. Convolutional neural network is the basis of all other complex networks. However, due to its simple structure, it cannot achieve high-precision classification tasks in complex environments. For this reason, researchers have proposed some typical CNN frameworks [11-13]: LeNet, AlexNet, GooleNet, VGGNet and ResNet. In general, researchers will use it in combination to make it more effective.

5. Summary

In this paper, the advantages and disadvantages of traditional neural network and convolutional neural network are analyzed, and their principles and current applications are studied. Today has been widely used in image recognition convolution neural network technology, image recognition application in the future environment will be more complex, when used alone one kind of convolution neural network technology has certain limitation, in order to improve the recognition rate under special circumstances, such as the target color and background color is consistent, overlap multiple targets, identify target sheltered under extreme conditions, such as serious can be combining different convolution neural network technology, make full use of their respective advantages, at the same time can be used for different objects of special structure corresponding neural network, to further improve the recognition rate of object.

References

[1] D. Jin and Z. Liu, "A review of image recognition algorithms based on deep learning," Computer Products and Circulation, vol. 2018, no. 03, pp. 188, 2018.

[2] Q. Zhang, J. Li, and Zhuoli, "Overview of vehicle identification technology," Journal of Beijing University of Technology, vol. 44, no. 3, pp. 382-392, 2008.

[3] J. Shang, "Research on edge detection algorithm based on Canny operator," Information and Computers (Theoretical Edition), vol. 32, no. 2, pp. 39-41, 2020.

[4] Y. Wang, "Multi-directional anisotropic edge detection algorithm," Computer and Digital Engineering, vol. 48, no. 1, pp. 167-169, 2020.

[5] X. Zhang and F. Chen, "Lane line edge detection based on improved adaptive Canny algorithm," Proceedings of the Journal of Physics: Conference Series, vol. 1549, no. 2, pp. 022131, 2020.

[6] M. K. Mokhtar, F. Mohamed, M. S. Sunar, et al., "Image features detection and tracking for image-based target augmented reality application," Proceedings of the 2019 IEEE Conference on Graphics and Media (GAME), Erfurt, Germany, pp. 38-43, 2019.

[7] M. Xiao, Y. Li, X. Yan, M. Gao, and W. Wang, "Convolutional neural network classification of cancer cytopathology images: taking breast cancer as an example," Proceedings of the 2024 7th International Conference on Machine Vision and Applications, pp. 145–149, Singapore, 2024.

[8] L. Dong, Y. Peng, and L. Fu, "Sobel edge detection based circle Harris corner detection algorithm," Journal of Xi'an University of Science and Technology, vol. 39, no. 2, pp. 374-380, 2019.

[9] W. Zhang and C. Sun, "Corner detection using multi-directional structure tensor with multiple scales," International Journal of Computer Vision, vol. 128, no. 2, pp. 438-459, 2020.

[10] X. Yan, W. Wang, M. Xiao, Y. Li, and M. Gao, "Survival prediction across diverse cancer types using neural networks," Proceedings of the 2024 7th International Conference on Machine Vision and Applications, pp. 134-138, 2024.

[11] W. Wang, Y. Li, X. Yan, M. Xiao, and M. Gao, "Breast cancer image classification method based on deep transfer learning," Proceedings of the International Conference on Image Processing, Machine Learning and Pattern Recognition, pp. 190-197, 2024.

[12] X. Zhang, X. Zhou, M. Lin, et al., "Shufflenet: An extremely efficient convolutional neural network for mobile devices," Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, USA, pp. 6848-6856, 2018.

[13] S. Xie, R. Girshick, P. Dollar, et al., "Aggregated residual transformations for deep neural networks," Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, USA, pp. 1492-1500, 2017.