Artificial Neural Network and Its Application in Image Recognition

Abstract: Computer vision technology refers to the technical vision that uses cameras and computers to replace human eyes to recognize, track and measure objects. Image recognition (image object recognition and detection) is one of its basic tasks. The key to accomplish this task is the artificial neural network model and algorithm. With the continuous improvement of neural network model, image object recognition technology has also been steadily improved. In this paper, the origin, types and development history of artificial neural network are introduced, and its application in image recognition is reviewed.

Keywords: Computer vision technology, Image Recognition, Artificial neural network

1 Introduction

Artificial intelligence is to use computers to complete intelligent work that only talents can do, mainly including voice signal processing, intelligent search, reasoning, image recognition, machine learning, etc. These works are all related to artificial neural network (ANN). As an important branch of artificial intelligence, ANN can simulate the human brain to complete some tasks requiring high-intensity learning and computing, so as to solve some complex problems better and faster.

"Artificial neural network is a complex network structure composed of a large number of neurons connected with each other. It can be used in some aspects by simulating some mechanisms of the brain" [1,2]. "The unique structure and information processing method of neural network make it have obvious advantages in many aspects and wide application fields. The main application fields include intelligent driving [3-5], process control and optimization [6-9], voice processing [10-13], signal processing [14-17], medical and health care, target detection [18-20] and so on".

Target detection is involved in many fields, such as defect detection, face

recognition, dangerous area detection, medical pathology detection and vehicle target detection. The two key links of target detection are image preprocessing and image feature recognition. The main tasks of image preprocessing include image conversion, image smoothing and image enhancement [21]. On this basis, the iterative threshold image segmentation method is used to intercept the target image, and the target image features are extracted based on HOG features, so that they are not affected by the image state, and accurate image features are obtained and target detection or recognition is performed.

Image conversion is to convert the analog image collected by the camera into digital image, which is mainly divided into three steps: sampling, quantization and coding. Among them, sampling refers to the discretization of analog images according to fixed time or space intervals; Quantization refers to restricting the discrete points of an image to a fixed range of values, that is, realizing the level of discrete points to be limited; Coding is to convert the quantized discrete points of the image into the form of binary digital representation, so as to compress the amount of information [22].

Image smoothing is to eliminate all kinds of interference and noise in the process of image acquisition by the camera according to a certain filtering rule, and restore its original appearance. According to different noise characteristics of the image, different filtering methods are used for image smoothing. At present, there are three main methods: mean filtering, median filtering, high or low pass filtering. The mean filtering method uses the average value of each point in the neighborhood of a pixel to replace the original gray level of the pixel; The median filtering method uses the gray value of adjacent pixels of a pixel in the image to replace the pixel value; The low-pass (or high-pass) filtering law sets a cut-off frequency point to filter out all values lower (or higher) than the frequency of the digital image signal to eliminate high-frequency (or low-frequency) noise.

In the process of target detection, after the digital image is smoothed, the details of the image itself (such as boundary contour, line, etc.) will be blurred. Therefore, image enhancement processing is also required, which mainly includes two tasks: histogram equalization and edge sharpening, respectively strengthening the target object

information and image edge contour information in the image [23].

After image enhancement, target image capture and feature extraction can be performed by threshold method, that is, the target object in the image is separated from the background. Its basic principle is to divide image pixel points into several categories by comparing the difference between image pixel points and the set feature threshold [24]

At present, the target detection technology based on neural network algorithm has developed rapidly. This paper will briefly describe the origin and development of artificial neural network, and summarize the application research progress of artificial neural network in target detection.

2 Development history of artificial neural network

As early as the 1940s, the artificial neural network was officially born, marked by the M-P model proposed by McCulloch and Pitts [25]. Then Hebb [26] proposed a basic principle of neural network learning algorithm, namely Hebb learning rule; In 1960, Widrow [27] proposed the adaptive linear element model. These models and algorithms lay a solid foundation for the application of neural network theory. However, due to the difficulty of crossing limits of electronic circuits, the development of neural networks has stagnated for nearly 20 years. Until the 1980s, new theories such as ART network, cognitive machine network, Boltzmann machine theory, and parallel distributed processing have been proposed continuously, which solved two problems proposed by Minsky [28]. Starting from the Hopfield model proposed by Hopfield, artificial neural networks have entered a new era of development [29]. “In 1986, Rumelhart and Hinton [30] proposed a multilayer feedforward network (BP) algorithm. BP algorithm includes forward propagation of signal and back propagation of error. This two-way feedback structure can reduce the error signal to the minimum at that time”. “In 1998, Vapnik [31] proposed the concept of SVM algorithm, namely support vector machine”. Hinton [32-34] et al. alleviated “the local optimal solution problem of ANN by using the pre-training method”.

3 Types and research progress of artificial neural networks

Artificial neural network is composed of multiple neurons, generally including input layer X, hidden unit and output layer Y. The basic types of neural network include single-layer feedforward network, multi-layer feedforward network and feedback network.

"Since the advent of the Hopfield network model, the artificial neural network has derived hundreds of models to realize data analysis and utilization by simulating other industries, such as thermodynamics, mathematics, medicine, etc. The following will introduce the three most widely used neural networks: feedforward neural network (BP), radial basis function neural network (RBF) and convolution neural network (CNN)" [35].

3.1 Back Propagation (BP) Neural Network

BP neural network is a prediction model widely used in many fields. It obtains the relationship between input data and output data through simulation training, and then obtains the simulation prediction results by inputting new data. It is a new intelligent prediction method [36]. Its learning process mainly includes the following four parts: input mode forward propagation, output error reverse propagation, cyclic memory training, and learning result discrimination. Generally, the three-layer BP neural network is the most widely used in water demand prediction, which includes three parts: input layer, hidden layer and output layer. In the MATLAB running environment, the construction steps of three-layer BP neural network model are as follows

a) Data preprocessing. Before training the neural network, it is necessary to normalize the original data to map the data to [0,1] or [- 1,1] or smaller range. b) Determine the number of hidden layer nodes. The number of hidden layer nodes has a great impact on the training effect of BP neural network. If the number of hidden layer nodes is too large, the network training will be too slow and easy to fall into local optimization. If the number of nodes in the hidden layer is too small, the network learning ability will be insufficient and the mapping relationship between input and output variables cannot be realized. c) Create a new forward neural network. d) BP network training and simulation. The TRAINGDM algorithm is used to train the BP neural network, and the BP network is simulated.

"The calculation of error items needs to start from the output layer, calculate the error items of the hidden layer and the input layer in turn, and update all weights" [37]. The weight is adjusted once every time a sample is processed. After multiple iterations (that is, all training data are processed repeatedly for multiple rounds), the weight of the network model can be trained to achieve the objective function.

3.2 RBF neural network

"Moody and Darken proposed radial basis neural network (RBFNN), which is an abstraction and simplification of human brain neural network system" [38]. "RBFNN uses radial basis function as transfer function. Compared with BP neural network, there is only one hidden layer, so RBF neural network is superior to BP neural network in function approximation, classification ability and learning speed" [39].

3.3 Convolution neural network (CNN)

Vaillant et al. [40] first proposed the application of convolutional neural network to face detection. Until 2012, the convolutional neural network AlexNet [41] made a major breakthrough in image recognition. LeCun [42] proposed LeNet-5 with depth structure as a classifier for image recognition. Its basic structure mainly includes input layer, convolution layer, pooling layer, full connection layer and output layer. The features of the convolution layer are obtained from the local features of the previous layer through the weight shared by the convolution. The input image features are extracted through multiple convolution layers and pooling layers, and the low-level features are gradually changed into high-level features; The high-level features are then classified through the full connection layer and the output layer to generate a one-dimensional vector that represents the category of the current input image.

4 Research progress of application of artificial neural network in target detection

In recent years, with the continuous breakthrough of computer vision technology, the field of image recognition and target detection based on artificial neural network algorithm has been extended, and its accuracy has also been improved by leaps and bounds. The application and research progress of artificial neural network in defect detection, crop disease and pest control, auxiliary medical treatment, intelligent production line and other fields will be summarized below.

Image matching is one of the key technologies of image recognition. The significant geometric distortion in the image brings difficulties to dense matching. Shi et al. [43] proposed “an affine invariant dense wide baseline matching method, which includes two main steps: first, extract and match a sparse set of affine invariant features: seed points and their affine invariant regions; Then these initial matches are used as spatial prior to generate more sparse matches. In the second step, the sparse matching obtained is used to initialize the dense matching propagation process. Dense matching is carried out through region growth, in which the matching extends from the high discrimination region to the low discrimination region, and finally the recovery technology is completed by the low-rank matrix. The comparison experiment between the proposed method and the existing method shows that it has significant improvement in the presence of large affine deformation”.

Artificial intelligence and image recognition can be applied to the assembly operation of product production lines. Based on image recognition technology, Wang Bixian et al. [44] proposed a method of using cameras to assist operators to complete assembly in assembly engineering and accurately locate the assembly area, realizing assembly area recognition, assembly positioning and operation guidance, which can effectively reduce assembly error rate and improve work efficiency.

The detection of capsule defects can be intelligentized and mechanized by means of artificial neural network. Wang et al. [45] considered the fuzziness and diversity of capsule foreign body defects in the image, and applied BP neural network to identify capsule foreign body defects. Firstly, the capsule image is divided into three parts by the vertical Sobel operator, and each part of the image is processed to remove the noise through the median filter; Then the histogram features of all three parts of the image, namely smoothness, skewness, flatness, distortion, kurtosis and entropy, are extracted and used as the input of BP neural network. Aiming at the non-uniformity of input data, a normalization method based on clustering algorithm is proposed. The experimental results show that the method has high accuracy.

Multi-scale problem is one of the difficult problems in image recognition. Zhao et al. [46] proposed a feature pyramid bidirectional semantic feature information fusion

model to achieve bidirectional fusion of semantic feature information of images at different scales. Then, a new multi-scale image recognition method based on bidirectional semantic information fusion of feature pyramid is formed by embedding a deep neural network to improve the accuracy of object recognition at different scales. The experimental results show that the average accuracy of this method in PASCAL VOC dataset is at least 0.7% higher than that of other methods, which can effectively improve the accuracy of multiscale image recognition.

In order to realize the convenient recognition of wheat leaf disease image based on mobile terminal, Feng Xiao et al. [47] established a wheat leaf disease image recognition model based on lightweight convolutional neural network (CNN) and migration learning. The sample set is composed of three wheat leaf disease images of wheat powdery mildew, stripe rust and leaf rust. The image recognition model of wheat leaf disease is constructed using the deep learning framework Tensorflow 2.0 and MobileNetV2. The trained parameters on the ImageNet data set are used as the initial parameters of the model. The effects of migration learning method, sample size, adding Dropout layer, and initial learning rate on the performance of the model are discussed. The results show that the wheat leaf disease image recognition model with high recognition accuracy, strong generalization ability and suitable for mobile applications can be constructed based on MobileNetV2 and migration learning.

Jiang Yanxiang [48] and others proposed a computer vision target detection algorithm based on full convolution neural network to solve the problem of low accuracy of traditional computer vision target detection algorithm. The computer vision image is converted into a digital image by sampling, quantizing and coding. After a series of processing such as image smoothing, image enhancement, and target image interception, the target image feature extraction is completed to achieve image preprocessing. The computer vision image target detection is completed by using the full convolution neural network training. The Matlab simulation software is used as the experimental platform. The image set used in the test experiment is from the VisualGenome data set. The total number of image samples is 108249. 1000 static images and 1000 dynamic images are selected as the experimental test objects. The

results show that the accuracy and recall rate of this method are closer to 1, the detection accuracy is higher, and the detection time is less.

Wu Heng et al. [49] The network is composed of deep separable convolution and reverse residual, which greatly reduces the calculation amount of the model and maintains high recognition accuracy. The accuracy rate of image recognition reaches 91.9%, and the inference delay is less than 1.5 s. It is suitable for microcomputers.

The image recognition function of artificial neural network is also very important in auxiliary medical treatment. Kondo *et al.*[50] uses the modified GMDH (Group Method of Data Handling) neural network algorithm to identify the radial basis function network. The algorithm has a feedback loop, and the neural network identified by it is applied to the medical image recognition of the brain. It can automatically select useful input variables, the number of hidden layers, the number of neurons in the hidden layer and the optimal architecture of neurons in the hidden layer to minimize the error criterion defined as Akaike Information Criterion (AIC). The results show that the neural network algorithm is very useful in brain medical image recognition.

Moving C-arm X-ray images play a key role in minimally invasive spine surgery. However, automatic lumbar spine recognition is still a challenging task. To this end, Li et al. [51] proposed a new automatic recognition method of lumbar spine based on two-way short-term memory (LSTM) circulating neural network (RNN). They took the curvature features of 3D lumbar spine model common to 2D X-ray images as the input of the model, used bidirectional recurrent neural network to learn the correlation of curvature features of lumbar spine under different imaging angles, and used LSTM neurons to replace the notes of bidirectional RNN. The experimental results show that this method can recognize the lumbar spine more accurately.

By measuring the size of different structures in ultrasound (US) images, the fetal development can be non-invasive evaluated. However, due to many factors, such as low signal-to-noise ratio and small fetal size, automatic classification of anatomical planes in fetal ultrasound images is challenging. Kumar et al. [52] proposed a new general method for classifying anatomical planes in fetal ultrasound images. This method trains two convolutional neural networks to learn the best ultrasound (US) image and its

salient features. The fusion of these features overcomes the challenges related to ultrasound fetal imaging by emphasizing the salient features in the image that can best distinguish different planes. For 12 of the 13 different planes found in the clinical data set of fetal ultrasound images, the classification accuracy is higher than the most advanced baseline.

Saravanakumar *et al.* [53] proposed a method of hyperspectral remote sensing satellite scene segmentation. In this scenario, the mainstream algorithm fuzzy c-means is implemented. In addition, the algorithm is performed in inter-band and intra-band clustering and the band with the highest variance is selected from each cluster. The reduced frequency band is de-correlated and then segmented using the fuzzy algorithm.

5 Conclusion

This paper summarizes the typical applications of image recognition technology based on artificial neural network in the fields of auxiliary medical treatment, intelligent driving and fault recognition. It can be expected that the image recognition technology based on artificial neural network will be more widely used in industrial production, crop protection, medical care and other industries, bringing great convenience to people's production and life. With the further development of artificial intelligence, the accuracy of image recognition based on artificial neural network will be higher and higher, and the application range will be wider and wider.

Reference

[1] Zheng Z, Qi Y. Study on the Simulation Control of Neural Network Algorithm in Thermally Coupled Distillation. Asian Journal of Research in Computer Science, 2021, 10(3): 53-64.

[2] Mo R, Wang H. Application of Neural Network Algorithm in Optimal Control of Ethylene Distillation Tower, Asian Journal of Research in Computer Science, 2021, 9(2): 19-29.

[3] Kalal Z, Mikolajczyk K, Matas J. “Tracking-learning-detection,” IEEE Transactions on Pattern Analysis and Machine Intelligence, 2012, 34(7):1409–1422.

[4] Zheng Z H, Wang B. On-Road Vehicle Detection Based on Color Segmentation and Tracking Using Harris-SIFT[J], Advanced Materials Research, 2012, 433-440: 5334-5338.

[5] Wang W, Gee T, Price J, et al.. Real time multi-vehicle tracking and counting at intersections

from a fisheye camera[C], 2015 IEEE Winter Conference on Applications of Computer Vision, pp. 17-24, 2015.

[6] Li C, Wang C. Application of artificial neural network in distillation system: A critical review of recent progress. Asian Journal of Research in Computer Science, 2021, 11(1): 8-16.

[7] Wang H, Chen Y. Application of Artificial Neural Networks in Chemical Process Control. Asian Journal of Research in Computer Science, 2022, 14(1): 22-37..

[8] Zhao N, Lu J. Review of Neural Network Algorithm and its Application in Temperature Control of Distillation Tower, Journal of Engineering Research and Reports, 2021, 20(4): 50-61.

[9] Lu J, Zhao N. Application of Neural Network Algorithm in Propylene Distillation, Journal of Engineering Research and Reports, 2021, 20(12): 53-63.

[10] Graves A, Mohamed A, Hinton GE. Speech recognition with deep recurrent neural networks Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP 2013, Vancouver, BC, Canada, May 26–31, 2013, IEEE (2013), pp. 6645-6649.

[11] Gao J, Chakraborty D, Tembine H, Olaleye O. Nonparallel Emotional Speech Conversion, INTERSPEECH 2019, Graz, Austria, September 2019.

[12] Chen SH, Wang YR. Tone recognition of continuous Mandarin speech based on neural networks, Speech and Audio Processing IEEE Transactions on, 1995, 3(2): 146-150.

[13] Lotfidereshgi R, Gournay P. Speech Prediction Using an Adaptive Recurrent Neural Network with Application to Packet Loss Concealment, 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2018, pp. 5394-5398, doi: 10.1109/ICASSP.2018.8462185.

[14] Zoia G, Zhou R, Mlynek D. A multi-timbre chord/harmony analyzer based on signal processing and neural networks, IEEE 6th Workshop on Multimedia Signal Processing, 2004, pp. 219-222, doi: 10.1109/MMSP.2004.1436532.

[15] Hinton GE, Deng L, Yu D, *et al.*. Deep neural networks for acoustic modeling in speech recognition, Signal Process. Mag., 2012, 29 (6) : 82-97

[16] Gao J, Tembine H. Distributed mean-field filters, workshop on frontiers of networks: Theory and algorithms, seventeenth international symposium on mobile ad hoc networking and computing, Mobi Hoc. July 5-7, 2016, Paderborn, Germany.

[17] Vishnyakov S. Artificial Neural Networks Implementation in Digital Signal Processing

Courses," 2018 IV International Conference on Information Technologies in Engineering Education (Inforino), 2018, pp. 1-5, doi: 10.1109/INFORINO.2018.8581706.

[18] Naka M, Imai T, Shida T, Sato M, Ito R, Akamine I. Thermal image processing using neural network, Proceedings of 1993 International Conference on Neural Networks (IJCNN-93-Nagoya, Japan), 1993, pp. 2065-2068 vol.3, doi: 10.1109/IJCNN.1993.714129.

[19] Aizenberg I, Aizenberg N, Hiltner J. Cellular neural networks and computational intelligence in medical image processing [J]. Image & Vision Computing, 2001, 19(4):177-183.

[20] Gao J, Shi F. A Rotation and Scale Invariant Approach for Dense Wide Baseline Matching. Intelligent Computing Theory - 10th International Conference, ICIC (1) 2014: 345-356.

[21] Wang H, Qi X, Wu G. Research progress of object detection technology based on convolutional neural network in deep learning, computer science, 2018, 45(9): 11-19.

[22] Zhou L, Liu Y. Laser image processing based on adaptive genetic algorithm, Journal of Shenyang University of Technology, 2019, 41(02): 174-178.

[23] Wang W, Fu Y, Dong F, Li F. Infrared ship target detection method based on deep convolutional neural network, Acta Optica Sinica, 2018, 38(7): 160-166.

[24] Lu Y, Zhang S. Object Detection in Optical Remote Sensing Image with Convolutional Neural Networks, China Sciencepaper, 2017,12(14): 1583-1589.

[25] McCulloch WS, Pitts W. A logical calculus of the ideas immanent in nervous activity, Bulletin of Mathematical Biophysics, 1943, 5: 115-133.

[26] Hebb DO, The Organization of Behavior: A Neuropsychological Theory[M]. Lawrence Erlbaum Associates, New Jersey, 1949.

[27] Widrow B. Adaptive ' Adaline ' Neuron Using Chemical ' Memistors ', Stanford Electronics Laboratories Technical Report, No. 1553-2 (1960).

[28] Minsky M, Seymour AP. Perceptrons: An Introduction to Computational Geometry[M]. MIT Press, Cambridge, 1969.

[29] Hopfield JJ. Neurons with graded response have collective computational properties like those of two-state neurons. Proc Natl Acad Sci, 1984, 81(10): 3088–3092.

[30] Rumelhart DE, Hinton GE, Williams RJ. Learning representations by back-propagation errors, Naturem, 1986, 323: 53–536.

[31] Vapnik V. Statistical learning theory, 1998, Vol 3, Wiley, New York.

[32] Hinton GE, Salakhutdinov RR. Reducing the Dimensionality of Data with Neural Networks[J].

Science, 2006, 313(5786):504.

[33] Hinton GE, Krizhevsky A, Wang SD. Transforming auto-encoders, Proceedings of the Artificial Neural Networks and Machine Learning - ICANN 2011 - 21st International Conference on Artificial Neural Networks, Espoo, Finland, June 14–17, 2011, Proceedings, Part I, Springer (2011), pp. 44-51.

[34] Srivastava N, Hinton GE, Krizhevsky A, Sutskever I, Salakhutdinov R. Dropout: a simple way to prevent neural networks from overfitting, Journal of Machine Learning Research, 2014, 15 (1): 1929-1958.

[35] Bao Z, Yu J, Yang B. Intelligent optimization algorithm and its MATLAB example, 2nd. Beijing: Electronic Industry Press, 2018: 184.

[36] Zhang W, Zeng Z. Research Progress of Convolutional Neural Network and its Application in Object Detection, arXiv:2007.13284 [cs.CV], 2020.

[37] Han P, Zhou H, Zhou B. Research and implementation of BP neural network principle. Radio and TV Information, 2018; 10: 121-125. (in Chinese)

[38] Moody J, Darken CJ. Fast Learning in Networks of Locally-tuned Processing Units. ***Neural Computation***, 1989, 1(2): 281-294.

[39] Qi Y, Zheng Z. A review of neural network algorithms and their applications in supercritical extraction, arXiv:2011.05279 [cs.NE], 2021.

[40] Vaillant R, Monrocq C, Lecun Y. Original approach for the localization of objects in images. Vision, Image, and Signal Processing, 1994, 141, 245–250.

[41] Krizhevsky A, Sutskever I and Hinton G E. Imagenet Classification with Deep Convolutional Neural Networks. Advances in Neural Information Processing Systems (NIPS), 2012, 1097-1105.

[42] Lecun Y, Bottou L, Bengio Y, Haffner P. Gradient-based learning applied to document recognition. Proceedings of the IEEE, 1998, 86(11): 2278-2324.

[43] Shi F, Gao J, Huang X. An affine invariant approach for dense wide baseline image matching. International Journal of Distributed Sensor Networks (IJDSN), 2016, 12(12).

[44] Wang B, Lu P, Qian H. Research on Auxiliary Assembly System Based on Image Recognition, Changjiang Information & Communications, 2021, 5: 39-41

[45] Wang H, Liu X, Chen Y. Detection of capsule foreign matter defect based on BP neural

network," 2014 IEEE International Conference on Granular Computing (GrC), 2014, pp. 325-328, doi: 10.1109/GRC.2014.6982858.

[46] Zhao S, Zhao L. On Image Recognition Using Bidirectional Feature Pyramid and Deep Neural Network, Journal of Harbin Institute of Technology, 2021, 26(2): 44-50.

[47] Feng X, Li D, Wang W, et al.. Image Recognition of Wheat Leaf Diseases Based on Lightweight Convolutional Neural Network and Transfer Learning, Journal of Henan Agricultural Sciences, 2021, 50(4) : 174-180.

[48] Jiang YX. Computer Vision Target Detection Algorithm Based on Fully Convolutional Neural Network, Journal of Shenyang University of Technology, 2021, 43（5）: 1-6.

[49] Wu H, Dong Z. Vegetables and fruits Recognition based on microprocessor and neural network, Journal of Ningxia University (Natural Science Edition), 2021, 42（1）: 39-44.

[50] Kondo T, Pandya A S. Revised GMDH-type neural networks with a feedback loop and their application to the medical image recognition, Proceedings of the 9th International Conference on Neural Information Processing, 2002. ICONIP '02., 2002, vol.5 : 2274-2278.

[51] Li Y, Liang W, Tan J. Bi-Directional LSTM Recurrent Neural Network for Lumbar Vertebrae Identification in X-Ray Images, 2017 International Conference on Computer Systems, Electronics and Control (ICCSEC), 2017, pp. 1047-1051.

[52] Kumar A, Sridar P, Quinton A, et al.. Plane identification in fetal ultrasound images using saliency maps and convolutional neural networks," 2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI), 2016, pp. 791-794.

[53] Saravanakumar V, Kavitha M. Saravanan, V.V.S.S.S. Balaram, Anantha Sivaprakasam S. Fuzzy C-Means Technique for Band Reduction and Segmentation of Hyperspectral Satellite Image, International Journal of Fuzzy System Applications (IJFSA), 2021,10(4):79-100.