ISSN: 2664-9640

Application of Convolutional Neural Network in Face Recognition System

Qingyang Sheng [1], Yihong Gu [1], Chengcheng Han [1], Yunlan Zhou [1], Zhaohan Ding [2]

[1] School of Automation, Nanjing Institute of Technology, Nanjing 211167, China;

[2] School of Environmental Engineering, Nanjing Institute of Technology, Nanjing 211167, China.

Abstract

With the gradual maturity of neural network algorithm, the recognition technology represented by face recognition is gradually popularized. Face recognition system mainly includes four components: image acquisition and detection, preprocessing, feature extraction, and matching and recognition. By using the Convolutional Neural Network the image can be recognized and processed quickly. This paper mainly summarizes the basic structure and working principle of Convolutional Neural Network, and uses Convolutional Neural Network algorithm to realize face recognition. The Convolutional Neural Network (CNN) model is composed of multiple convolutional layers, and FD algorithm which is improved by fractional differential algorithm is adopted to improve the accuracy of face recognition and realize the purpose of optimizing face recognition system by simultaneously extracting local and global features of a face. At the same time, Sigmoid node processing technology is adopted to ensure that the network training set can quickly converge to a stable state, so as to improve the rapidity and stability of face recognition system.

Keywords

Convolutional Neural Network, face recognition, FD algorithm, Sigmoid node processing technology.

1. Introduction

In recent years, the neural network algorithm has gradually matured, Convolutional Neural Network has developed rapidly as a special kind of neural network. In the 1960s, Canadian neuroscientists David H. Hubel and Torsten Wisese discovered the neural network architecture for the first time and proposed the concept of "receptive field" of a single neuron. Fukushima adopted unsupervised learning through Convolutional Neural Network, and in 1980 proposed a receptive field based multi-layer neural network model – Neocognitron and simulated biological vision system, then proposed a hierarchical multi-layer artificial neural network -- "Neurocognitron", which is the predecessor of the present Convolutional Neural Network. In 1998, Lecun et al. proposed lenet-5, a Convolutional Neural Network structure with relatively simple structure. Its success in the field of handwritten character recognition has attracted the attention of the academic community to the Convolutional Neural Network. At the same time, the research of the Convolutional Neural Network in object detection, face recognition and other aspects has been gradually carried out. As an efficient deep learning model, Convolutional Neural Network was proposed in 2006, and was widely promoted by Krizhevsky et al in 2012. In 2016, Alphago successfully used Convolutional Neural Network to analyze go board information, leading to more in-depth research on Convolutional Neural Network. Among them, the application of Convolutional Neural Network in face recognition system is the focus of current research, which is of great significance both theoretically and technically.

ISSN: 2664-9640

Face recognition objects have quite complex changes in detail, manifested in the uncertainty of acquisition process, multidimensionality of image training, the diversity and variability of face patterns, face plastic randomness and uncertainty and other aspects, which are the challenges that must be overcome for the system to be able to make recognition quickly and accurately. With all the above problems in the process of face recognition being stacked together, the situation becomes more complex in the actual detection and recognition process. For example, under the Novel Coronavirus attack, a large number of face products in China cannot be scanned and identified while wearing masks. Therefore, in the face of uncertainties in the future, in-depth research is needed to get technological innovation and breakthrough achieved.

2. Overview of Convolutional Neural Networks

2.1. Basic Structure

The basic structure of Convolutional Neural Network (CNN) consists of input layer, convolutional layer, pooling layer (also known as sampling layer), full connection layer and output layer. The specific structure is shown in Figure 1.

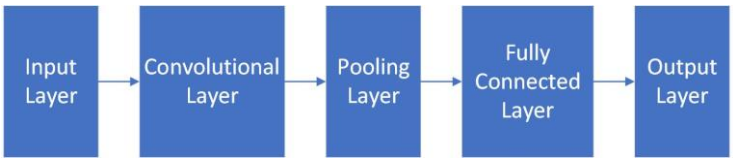

Figure 1: Basic structure diagram of Convolutional Neural Network

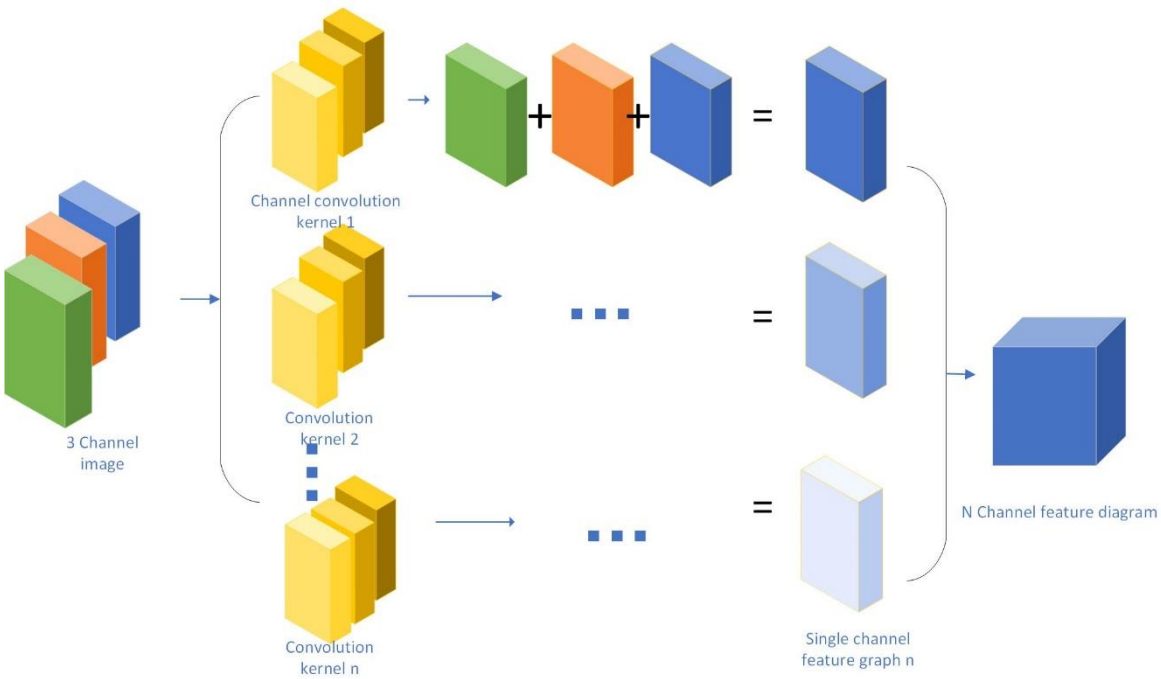

Figure 2: Calculation process diagram of filter extraction for each channel

In the convolution process, each neuron of the output feature plane in the convolution layer is locally connected with its input, and the input value of the neuron is obtained through the weighted sum of the corresponding connection weight and the local input plus the bias value, so as to complete the convolution. Convolutional Neural Network is mainly composed of input layer, output layer and multiple hidden layers. Based on each layer, it is divided into multiple

ISSN: 2664-9640

two-dimensional planes composed of numerous independent neurons. In comparison, traditional neural networks require far more parameters than Convolutional Neural Networks when applied to the full connection layer. In the process of operation, the idle area of the Convolutional Neural Network can share the same filter, whose depth is consistent with that of the input layer. Moreover, the filter of each channel performs channel operation independently, and the final weight is added to diminish the number of parameters of the Convolutional Neural Network, reduce its complexity and improves the operation efficiency.

The convolution layer is mainly used for feature extraction. Generally, the convolution layer and the convolution kernel are one-to-many modes. The convolution layer at the lower level receives the feature graph at the upper level and performs the convolution operation, and then adds the offset vector of this layer to obtain the output feature graph of this layer. N convolution kernels are matched with the feature graph, and the operation is repeated successively until it reaches the highest level.

The pooling layer aims at reducing the resolution of feature planes and acquiring features with spatial invariance. The output feature graph obtained through the calculation of the convolution layer is not much different from the original graph in dimension, so it needs to be compressed by the pooling layer. In the pooling layer, the nonlinear down-sampling method is adopted to reduce the dimension of the input feature graph, thus simplifying the computational complexity of the network. Subsequently, the pooling layer performs feature compression by extracting main features. In practical application, the above operations are repeated after feature sampling in the input pooling layer to reduce the size of the input feature graph.

Each neuron in the full connection layer is fully connected with all nodes in the upper layer, and end-to-end learning is realized by converting the two-digit feature graph of the convolution output into one-dimensional vector. The full connection layer can integrate the convolutional layer or the pooling layer, distinguish local information, and play the role of "classifier" in the whole Convolutional Neural Network. Compared with convolutional layer, pooling layer and activation function layer, which map original data to hidden layer feature space, full connection layer plays the role of mapping "distributed feature representation" to sample label space.

2.2. Working Principle of Convolutional Neural Network

The objects of Convolutional Neural Network include images and all the data that can be transformed into image-like structures. Compared with traditional algorithms and other types of neural network systems, Convolutional Neural Network has higher efficiency in processing two-dimensional local information of images, extracting image features and classifying images. The working principle diagram of Convolutional Neural Network is shown in Figure 3.

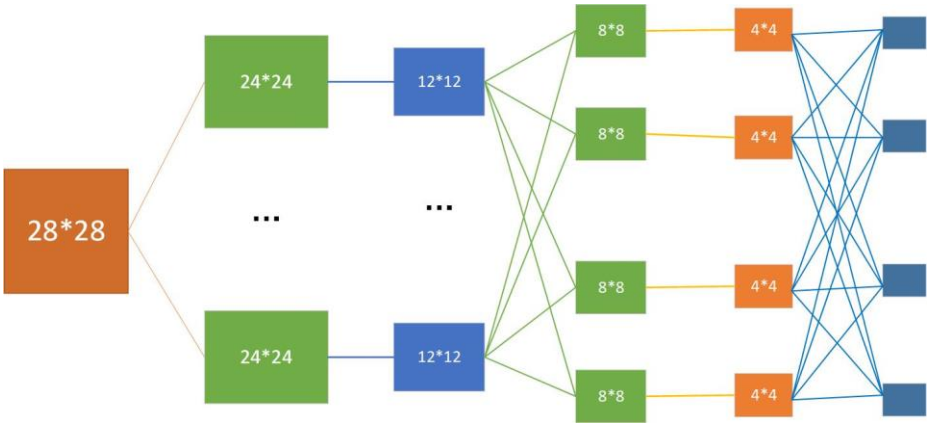

Figure 3: Working principle diagram of Convolutional Neural Network

As can be seen from Figure 3, assuming that the input size of the Convolutional Neural Network is 28×28, there are 6 5×5 convolutional kernels in the first convolutional layer, and 6 feature images with the size of 24×24 will be output through the convolution effect. Then, after the up-

ISSN: 2664-9640

sampling layer, the convolutional layer was averaged and transferred to the second convolutional layer, and six feature maps with a size of 12×12 were obtained. The second convolution layer is composed of 12 5×5 convolution kernels, each convolution kernel is convolved with the 6 feature graphs of the upper layer respectively, and the result is summated to form a new feature graph, that is, the layer generates 12 8×8 feature graphs, which are processed by sigmoid function and output to the lower sampling layer. In the lower sampling layer, 12 feature graphs with size of 4×4 are obtained, which are consistent with the number of feature graphs in the convolution layer. The output layer expands the feature graph of the lower sampling layer into a one-dimensional vector of length 192. After that, the vector is taken as input, and fully connected and classified with other layers below. In simple terms, two-dimensional images are compressed into one-dimensional vectors by dimensionality reduction, and fully linked with other network layers to simplify image processing.

3. Application of CNN in Face Recognition System

3.1. Basic process of face recognition

The purpose of face recognition system is to complete face recognition tasks through image feature extraction and data comparison, mainly including detection, characterization, recognition of three basic processes. Face detection is feature location and area ratio calculation for faces in different situations; Face representation is to extract facial features to determine the way of face description; Face recognition is the selection of matching algorithm to compare the detection object with the existing data in the database and get the results. The process of the system to complete the face recognition task is shown in Figure 4.

Figure 4: Flow chart of face recognition task

In the process of image acquisition, it is difficult to determine the neural network classifier for face recognition because the dimension of face image vector exceeds the number of training samples. Secondly, the interpolation algorithm of image compression is related to the calculation of the distortion degree of the image, so the selection of algorithm is closely related to the rapidity and accuracy of the system.

3.2. Face Recognition System

The whole system is composed of Convolutional Neural Network and face detection modules. Fractional differential is adopted to improve face recognition, and attention mechanism will be added to network model structure to enhance face feature extraction and integrate information of different channels and Spaces. Fractional differential was used to process Sigmoid node function, and convolution block was added to extract face details. Finally, loss function optimization model was constructed, and iterative training was carried out in the network to complete face recognition.

3.3. Convolutional Neural Network

Convolutional Neural Network is the core part of the system, which is used to extract face features and comprehensively process tasks such as feature extraction and network training and verification to obtain network model parameters of face features with good differentiation .

For network training and verification, the Convolutional Neural Network structure is constructed according to the configuration file, and network training based on error backpropagation algorithm is carried out. Through forward calculation and backward conduction, the update of network parameters can realize the tuning of parameters and greatly

improve network performance. In network training, the performance of the network should be verified periodically to judge the network convergence immediately.

After building the Convolutional Neural Network, feature extraction is performed on the preprocessed face images by fractional differential. Fractional differential is a differential form that extends the order of the traditional integer derivative to the field of fractions and complex numbers. Compared with the traditional neural network, feature extraction is more comprehensive and accurate, which can achieve the dimension reduction of any dimension of the image under sampling, and greatly improve the training efficiency of the network.

3.4. Face Detection

Based on ResNet residual network, FD algorithm mechanism is added to compress the feature graph, and then channel and spatial attention mechanism are added to gradually locate the feature information and obtain the results, which are presented in the form of probability feature vector. The specific FD-RESNET network structure is shown in Figure 5. Using the channel attention mechanism, the network compresses the spatial dimension of the input feature graph and sums and merges the elements one by one. After that, average pooling and maximum pooling are carried out through the pooling layer, and the feature map is generated and transmitted to the shared network. Spatial attention takes the generated channel attention as the input for secondary dimension reduction compression and outputs the feature map of spatial attention.

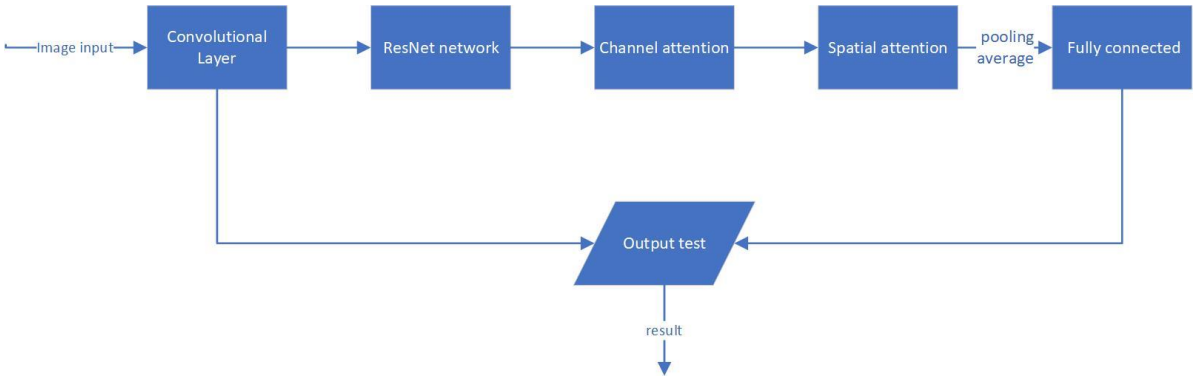

Figure 5: FD-RESNET network structure diagram

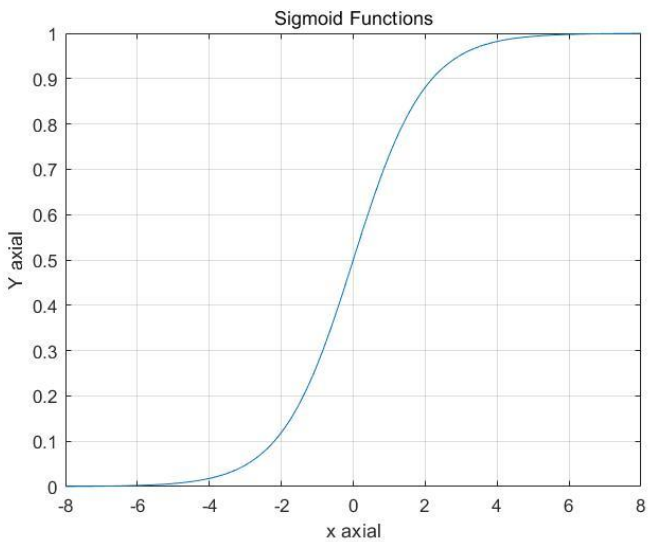

Figure 6: Sigmoid function diagram

In the attention module, the node function Sigmoid is processed by fractional order. Sigmoid function is often used as the activation function of neural network due to its properties of single

increment and differentiable everywhere. This function can map all the real numbers to the interval of (0,1) when the hidden layer neurons output, so as to achieve dichotomies. The Sigmoid function image is shown in Figure 6.

When the input is in an appropriate neighborhood near the equilibrium position, the function is basically linearly correlated with the independent variable, so the output is sensitive to the input. But outside the neighborhood, the input is insensitive to the output.

Based on ResNet residual network and FD algorithm mechanism, Convolutional Neural Network extracts features from channel and space dimensions, which greatly improves the speed of extracting facial image features and the feature capture ability of specific regions.

4. Conclusion

This paper summarizes the basic structure and working principle of Convolutional Neural Network and discusses the realization of face recognition using Convolutional Neural Network algorithm. Through the targeted analysis of fD-Resnet model structure and fractional-order Convolutional Neural Network model of residual network based on fractional differential, the FD algorithm mechanism based on ResNet network and the Sigmoid processing mechanism of fractional differential on node function are further summarized, comprehensively analyzed and described to improve the speed of model convergence. An effective method for model accuracy and face recognition accuracy. With the continuous development of artificial intelligence, face recognition technology with Convolutional Neural Network as the core is the key field of future research, which has practical significance for public security, criminal investigation and search.

References

[1] D H Hubel, T N Wiesel. Effects of monocular deprivation in kittens [J]. Naunyn-Schmiedebergs Archiv für experimentelle Pathologie und Pharmakologie, 1964, 248(6).

[2] K FUKUSHIMA. A hierarchical neural network model for associative memory [J]. Electronics and Communications in Japan (Part I: Communications), 1983, 66(12).

[3] K FUKUSHIMA. Analysis of the process of visual pattern recognition by the neocognitron [J]. Neural Networks, 1989, 2(6).

[4] Y LECUN, L BOTTOU, Y BENGIO, et al. Gradient-based learning applied to document recognition [J]. Proceedings of the IEEE, 1998, 86(11).

[5] N SRIVASTAVA, G E HINTON, A KRIZHEVSKY, et al. Dropout: a simple way to prevent neural networks from overfitting. [J]. Journal of Machine Learning Research, 2014, 15(1).

[6] M C FU. Simulation-Based Algorithms for Markov Decision Processes: Monte Carlo Tree Search from AlphaGo to AlphaZero [J]. Asia-Pacific Journal of Operational Research, 2019, 36(06).

[7] X LI, X Sun, L Wu, et al. The Survey of Methods and Algorithms for Computer Game Go; Proceedings of the 2017 Chinese Intelligent Automation Conference (CIAC 2017), Tianjin, F, 2017 [C].

[8] S P R BORRA, N V S S PRADEEP, N T S RAJU, et al. Face Recognition based on Convolutional Neural Network [J]. International Journal of Engineering and Advanced Technology (IJEAT), 2020, 9(4).

[9] Li Hui, Shi Bo. Face recognition algorithm based on convolutional Neural Network [J]. Software Guide, 2017, 16(03): 26-9.

[10] Forest in time. Research on Deep Convolutional Neural Network Based on Sequence Pooling [J]; Journal of Anhui Agricultural University (Natural Science edition); 2006-06 Zhengzhou University, 2017.

[11] Y D Chen, L M Wang. Face recognition Method based on convolutional Neural Network. Journal of Northeast Normal University (Natural Science), 2016, 48(02): 70-6.

[12] J Q Qin. Research on Image Classification and Weak Light Image Enhancement based on CNN [D]; Guilin University of Electronic Technology, 2019.

[13] Z H Zhang, R Wang. An improved Face recognition Method based on MobileFaceNet Network [J]. Journal of Beijing University of Aeronautics and Astronsutics, 2020, 46(09): 1756-62.

[14] Y Ge. Research on Remote Sensing Image Retrieval based on CNN Migration Feature Fusion and Pooling [J]; Journal of Anhui Agricultural Sciences; 2006-01 Nanchang University, 2019.

[15] R X Wu, Q K Xiao. International Journal of Electronic Measurement Technology, 2019, 38(05): 86-90.

[16] W J Chen, P Zhou, H Y Yang, et al. A saliency detection model for the mechanism of channel-spatial joint attention [J]. Computer Engineering and Applications, 2021, 57(19): 214-9.

[17] Y G Zhu, R M Liu, Q T Huang. Journal of Electronic Measurement and Instrument, 2020, 34(02): 115-22. (in Chinese)

[18] Yang Z, Wu B, LIAN W W, et al. Face recognition based on attention mechanism and depth identity mapping [J]. Sensors and Microsystems, 2020, 39(09): 150-3.

[19] Zhang Y N, QU L, Chen J W, et al. Multi-input Sigmoid excitation function neural network weight and structure determination method [J]. Application Research of Computers, 2012, 29(11): 4313-6 +51.

[20] X Zhang, X Huang, W H Zhong, et al. FPGA implementation of Sigmoid function and its derivative function [J]. Journal of Fujian Normal University (Natural Science Edition), 2011, 27(02): 62-5.