Computers and Electrical Engineering 92 (2021) 107128

Contents lists available at ScienceDirect

Computers and Electrical Engineering

journal homepage: www.elsevier.com/locate/compeleceng

Face recognition framework based on effective computing and adversarial neural network and its implementation in machine vision for social robots

Chenglin Yu [a,*], Hailong Pei [b]

[a] Key Laboratory of Autonomous Systems and Networked Control, Ministry of Education, Unmanned Aerial Vehicle Systems Engineering Technology Research Center of Guangdong, South China University of Technology, Guangzhou 510640, China

[b] Key Laboratory of Autonomous Systems and Networked Control, Ministry of Education, Unmanned Aerial Vehicle Systems Engineering Technology Research Center of Guangdong, School of Automation Scinece and Engineering, South China University of Technology, Guangzhou 510640, China

ARTICLE INFO

Keywords:
Computational model
Counter neural network
Deep learning
Face recognition
Machine vision

ABSTRACT

In recent years, with the continuous breakthrough of computer vision technology, the accuracy of object detection and target recognition has been improved by leaps and bounds. Face recognition is one of the important research directions in the field of computer vision, which is widely used in mobile payment, safe city, criminal investigation and other fields. Traditional face recognition methods need to extract face image features manually. The extracted features are greatly affected by subjective factors, and time-consuming and laborious. Deep learning is the most important technology in the field of computer vision at present. Compared with traditional face recognition methods, it can extract more essential features of face image without manual participation. In this paper, we build a face recognition system based on neural computing model and the principle of neural network. The experimental results show that the proposed method has high detection rate and short processing time.

1. Introduction

In 1950, Turing, the father of computer, put forward the idea of artificial intelligence in his paper. However, in recent years, with the deepening of the research on convolutional neural network, artificial intelligence has made a breakthrough. Compared with other biometric technologies (such as fingerprint recognition, speech recognition, iris recognition, etc.), face recognition technology has the advantages of non-invasive, strong interactivity, convenient and fast. Face recognition is a technology of identity recognition based on human face feature information. It analyzes the collected face image information by computer, extracts effective face feature information by certain feature extraction algorithm, and finally uses the extracted face feature information for identity recognition. Face recognition is a very important research direction in the field of pattern recognition and computer vision. Face recognition technology is widely used in mobile payment, safe city, criminal investigation, national defense security and other fields. It provides a safe, convenient and efficient solution for personal identity verification, and has a very broad application prospect and extremely important academic research value [1–4].

This paper is for special section VSI-ain. Reviews processed and recommended for publication by Guest Editor Dr. S. Smys.

* Corresponding author.
E-mail address: xueku97062@163.com (C. Yu).

https://doi.org/10.1016/j.compeleceng.2021.107128
Received 10 June 2020; Received in revised form 5 November 2020; Accepted 15 March 2021
Available online 30 March 2021

The development of artificial neural network has experienced three important periods. The first stage is from 1940s to 1960s. Neuroscientist MC CULLOCH and mathematician Pitts proposed to use neural network computing model to judge different types of input, and began to apply neural network to the research of artificial intelligence. However, due to the limitations of neural network perceptron and the lack of computer computing power, the research of neural network stagnated. The second stage is from 1980s to 1990s. Based on the experiment of cat's vision system, Yoshihiko Fukushima of Kyoto University put forward the concept of cognitive machine and receptive field. In 1989, Professor Yann Lecun of New York University realized the first convolution neural network combined with the back propagation algorithm which can be used to train neural network. In 1998, Lecun introduced a convolutional neural network called Lenet for handwritten digit recognition, which has been successfully applied in the commercial field. The third stage is since 2006, Professor Hinton first proposed the deep belief network (DBN), which has been extended to many different neural networks, greatly improving the performance of the network model. At the same time, the concept of deep learning has been formally proposed and promoted. By 2012, in the ilsvrc competition, the top-5 recognition rate of the convolutional neural network Alexnet reached an amazing 15.3%, while the second place at that time was only 26.2%. So far, deep learning is out of control, ilsvrc is refreshed by deep learning every year [5–8]. Feature extraction methods based on facial feature points mainly include geometric and texture features. Reference [5] mentions that the facial features that can be extracted include geometric features and appearance features. Geometric features can represent the shape and position of facial components. Literature [9] also mentioned feature points. Because feature point recognition algorithms are relatively mature, it is more intuitive and effective to directly detect such features. Literature [4] uses local texture features, global geometric features and hybrid features to propose a three-layer facial expression image classification framework of local layer-global layer-hybrid layer.

Most of the geometric feature vectors extracted from literatures are dominated by absolute distance, but the distance between feature points of different people may vary greatly, which has certain influence on the recognition. For example, in the surprise expression, different people have different degrees of opening their mouths, and the distance may vary greatly. According to the data, the distance is the absolute difference calculated by the pixels in the figure, and the range of variation may be relatively large. Compared with the commonly used biometrics for recognition, face recognition has become a hot technology in the field of identity recognition with its advantages of convenience, high intelligence, human-computer friendliness, security and stability, not suitable for forgery. With the increasing demand for security protection in various application fields, face recognition technology has undoubtedly a huge development prospect and market space. For traditional facial expression feature extraction methods, image preprocessing, such as eliminating the influence of light, face detection accuracy and face angle recovery, has a great impact on feature extraction. The existing methods will try to avoid the influence of light, angle and so on, and limit the environment of collecting data when collecting samples. The biggest difference with machine learning is that face recognition based on deep learning does not need artificial design features, but self-learning internal abstract features, which has high efficiency and universality. However, the neural network also has obvious defects: a large number of parameters, the lack of training resources and the computing power of CPU at that time lead to the inability to train the neural network with huge structure. Therefore, although the neural network was advanced in the assumption at that time, it did not perform well in the test effect, coupled with the difficulty of training, resulting in not being put into practical use. In essence, deep learning is very similar to human brain nerve in the way of solving problems. It abstracts and extracts features layer by layer through nonlinear model, which shows unprecedented advantages in expression ability and learning ability. As a classical research direction of artificial intelligence, face recognition has a deep foundation, but it has a great demand in performance and application.

Deep learning is a sub field of machine learning. In recent years, deep learning has shown great advantages in computer vision, speech recognition, image semantic analysis, recommendation system and other fields. It is the most popular technology at present. In recent years, computer hardware, especially GPU technology, has been changing rapidly. Google, Cambrian and other companies have launched AI chips specially for in-depth learning. Using computers to carry out large-scale parallel computing is no longer out of reach, which is also the most critical factor for the rise of in-depth learning. Feature extraction is the most important part of the whole recognition system. Deep learning can automatically extract more essential features from massive data, so deep learning can be far ahead of any previous shallow learning method. However, deep learning is still a variant of neural network in essence [9–12]. The common models of deep learning include the automatic coder, the extension of the automatic coder, and the convolutional neural network, which is composed of multiple restricted Boltzmann machines. In-depth learning these models are hierarchical, have many parameters and large enough capacity, so they can better represent the data characteristics. In view of the difficulty of image and speech recognition, deep learning can achieve good results. In addition, deep learning can combine the classifier and features in one frame, and use the training samples to learn the network parameters directly. This advantage is that it can reduce the artificial design intervention, greatly reduce the workload and improve the recognition rate. According to the above analysis, this paper constructs a face recognition system based on effective computing model and neural network.

The rest of the paper is organized as follows. In the Section 2, we introduce the neural computation and neural network basis. In the Section 3, we present the application of effective computing model in face recognition. In the Section 4, we discuss the experiment with the comarison analysis of the other state-of-the-art models. In the last section, we summarize the work and discuss the future work.

2. The backgorund and theoretical basis

2.1. Neural network structure

Neural network simulates the neural structure of human brain in order to solve some complex problems, which is essentially the category of traditional machine learning algorithm. The nervous system of human brain is extremely complex, with 100 billion neurons. Brain processing mechanism is to process information in a separate form through each neuron in turn. The artificial neural network imitates the neural mechanism of human brain and adds some "nodes" similar to neurons between the traditional computer instructions and outputs. When the data passes through these nodes in turn, it is calculated according to its specific coefficient. In a word, neural network can make the computer not only be a tool for calculation, but also have the learning ability of human like.

In the 1960s, Frank Rosenblatt, a psychologist at Cornell Aerospace Laboratory, put forward the first "artificial neuron" model, called "perceptron". Perceptron is essentially a 0 / 1 classification algorithm, and neural network is derived from it. The structure of the perceptron is very similar to that of the human nerve cell

$$output = f\left(\sum_{i=1}^{n} w_i x_i - b\right) \tag{1}$$

The perceptron model is shown in Fig. 1.

The activation function also increases the expression ability in order to add nonlinear factors, which is equivalent to human cell body. The most common and common activation function in the early days was sigmoid. The expression of sigmoid function is as follows:

$$f(x) = \frac{1}{e^{-x} + 1} \tag{2}$$

As shown in Fig. 2, the sigmoid function image is monotonically increasing, with a value range of 0–1, which is symmetric about (0,1 / 2). Because the derivation is simple, it is suitable for back propagation.

After the calculation of weight and bias, an output is generated by the neuron (activation function), which is similar to that the nerve endings feel the changes of various external environments, and finally the output is generated.

Back propagation algorithm is the basic algorithm of neural network. In the back propagation algorithm, the optimal training parameters are obtained by reducing the error. The error formula is:

$$E_{total} = \sum \frac{1}{2}(\text{target} - output)^2 \tag{3}$$

As shown in Fig. 3, the error of the output layer propagates back to the hidden layer:

The error can be transferred layer by layer through the weight matrix. Through the chain rule, the weight is updated by partial derivation of the corresponding weight with the overall error, as shown in Fig. 4:

The hidden layer does not directly receive and send signals, that is, it is in a hidden state for the outside world. As shown in Fig. 5:

From the development of neural network, we can see that the inspiration of neural network comes from neuroscience, and even sometimes the artificial neural network is used to help understand the brain mechanism. Therefore, many materials introducing neural network will appear the schematic diagram of biological neuron to help explain the principle behind neural network. And now, because we don't have enough information about the brain, neuroscience has less influence on deep learning. Modern deep learning draws inspiration from many fields, especially applied mathematics, including probability theory, linear algebra and information theory.

Two important reasons for the remarkable effect of deep neural network are the deep network model and the large scale of data, which lead to a large amount of computing resources consumed in its training. On the one hand, vectorization programming improves the operation speed of specific numerical calculation of a single CPU in the form of single instruction multi data stream programming. On the other hand, GPU uses its multicore architecture to parallelize vector operations. In addition, two distributed training methods, model parallelization and data parallelization, make the training of neural network not limited to a single machine. The large-scale distributed computing cluster further accelerates the training task of large-scale deep neural network by using the distributed storage and asynchronous parameter update mechanism of the model. At present, the mainstream deep learning framework supports the distributed training of network.

2.2. Neuron model

Neuron is the basic unit of neural network. Inspired by the information transmission mechanism in the biological nervous system,

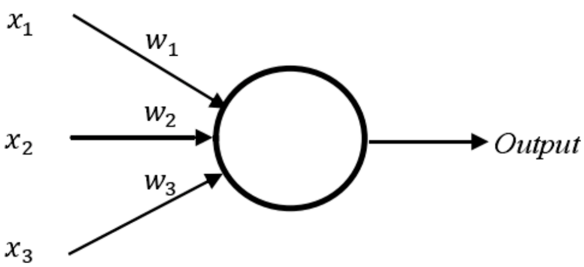

Fig. 1. Perceptron model diagram.

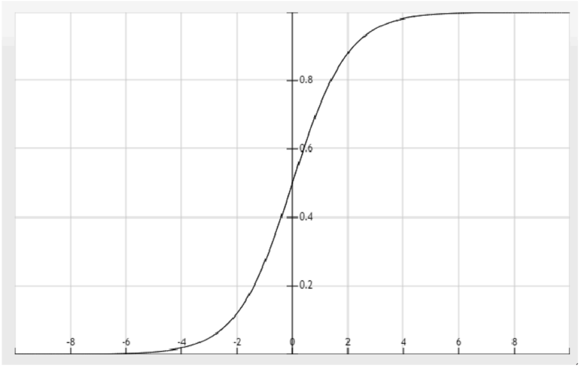

Fig. 2. Sigmoid function.

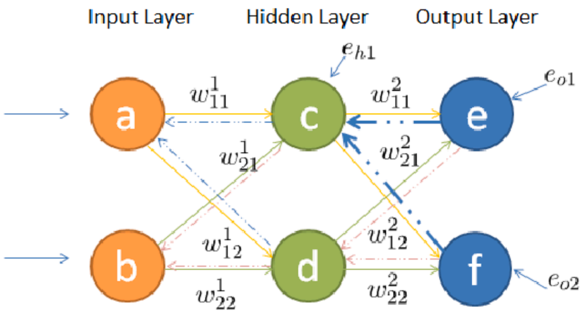

Fig. 3. Back propagation process.

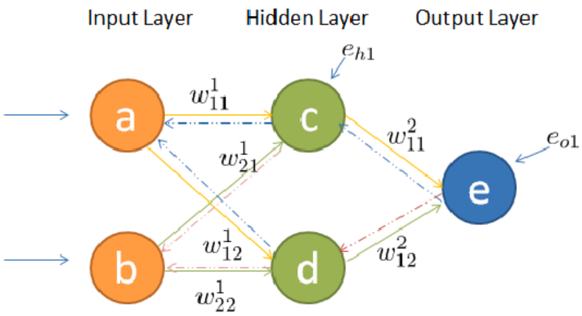

Fig. 4. Weight update process.

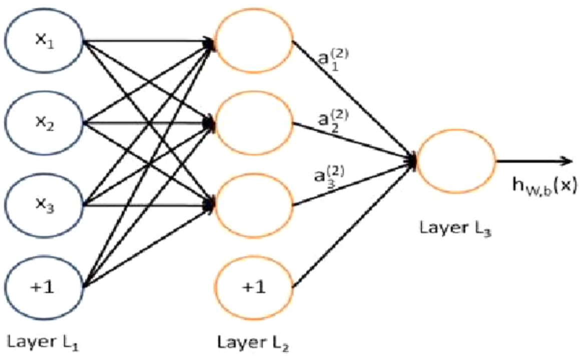

Fig. 5. Structure of feedforward neural network.

the earliest neuron model is proposed, as shown in Fig. 6 and Fig. 7.

Among them, x_i is the input of neurons, w_i is the connection weight, b is the bias term of neurons, the function is a nonlinear activation function, and h is the output of neurons. Its function expression is:

$$h_{w,b} = f\left(\sum_{i=1}^{n} W_i x_i + b\right) \tag{4}$$

Sigmoid function and tanh function are commonly used activation functions. The value range of the sigmoid function is (0, 1). Tanh function is a variation of sigmoid function, and its value range is (−1, 1).

Although fully connected network has advantages in classification, it is difficult to train because of too many parameters in modern large-resolution image. Human visual neurons sense local images rather than the whole. Finally, these neurons are combined in high latitude to obtain the perception of the whole image. This property is called image local connectivity. Inspired by the visual structure, in the neural network, each neuron only connects with the local area of the previous layer. This is equivalent to scanning the whole image with a local receptive field, and all neurons will be connected to a node in the next layer. This local connection mechanism can greatly reduce the number of network connections and weight parameters, and reduce the training difficulty.

A good parameter initialization method can make convolution neural network converge faster and faster. The common initialization methods of convolutional neural network are as follows:

1 Constant initialization: each parameter is initialized with a fixed constant. The constants commonly used for initialization are generally small, usually 0. The constant initialization method is usually used for offset terms.
2 Uniform distribution initialization: assume that the parameters are uniformly distributed in the interval [l, h], and then initialize the parameters. Xavier initialization method Xavier initialization method can make the output variance of each layer as equal as possible, so that the information in the network flows better.
3 Gabor initialization: the method of Gabor initialization is based on the parameters of Gabor filter directly as the parameters of neural network, which is generally used in the first layer of the network. Because of its natural image filtering characteristics, it can fix parameters in many cases. It can reduce the training time and burden.

2.3. Deep learning framework

Deep learning framework, simply speaking, is the library of deep learning and the tool to help deep learning. A single component represents each layer of the network or deep learning algorithm. Users can build according to the network structure to get the model they need. The advantage of using deep learning framework is that you don't have to write some components repeatedly. The model is equivalent to building blocks, which can be assembled directly, but different assembly methods, that is, different data sets depend on users.

Industry has played a decisive role in the open-source of deep learning framework: tensorflow opened by Google, torch on Facebook, paddy on Baidu, CNTK on Microsoft, MXNet on Amazon's AWS, and NL4J deep learning framework supporting Hadoop.

1 Caffe is, in a sense, the first widely popular deep learning framework. In 2013, Jia Yanqing of the University of California, Berkeley completed the development of Caffe during his doctoral study and opened it all through GitHub. Caffe provides a good implementation of CNN and a perfect Python and Matlab interface. It is simple to use, easy to expand the code, and fast to run. Therefore, after the open source, academia and industry are highly praised. Caffe has been used in a large number of papers. Today, Caffe is still the most mainstream framework in the field of deep learning, especially computer vision.
2 Tensorflow is a new generation of open-source deep learning framework of Google. Its predecessor is distbelief. Tensorflow is written in Python and accelerated by the C / C + + engine. The name of tensorflow comes from its unique processing method: tensor (vector), which represents data, and the runtime uses the data flow graph (flow) as the basic unit to represent the calculation

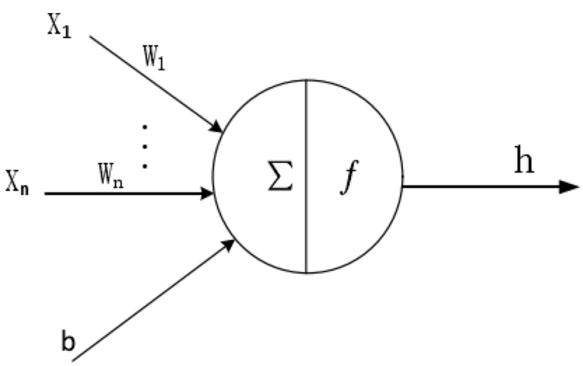

Fig. 6. Neuron model.

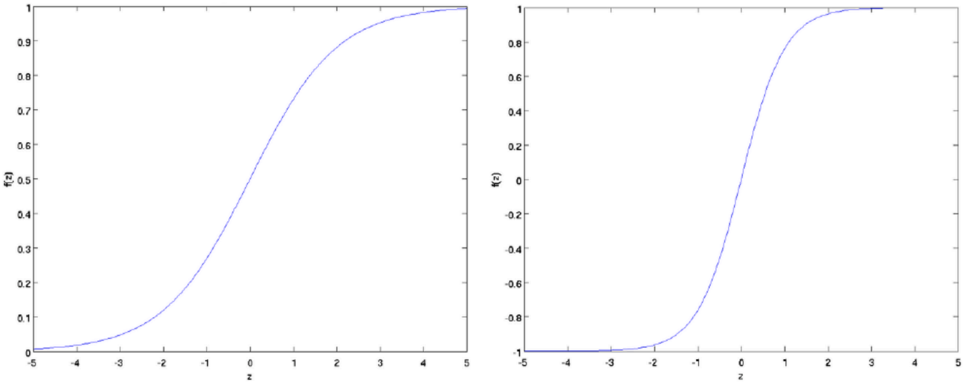

Fig. 7. Sigmoid function and tanh function.

task. This unique processing method makes it very simple to build a network through tensorflow, and the development efficiency is very high. At the same time, tensorflow supports the most popular network structures in the field of deep learning, such as CNN in the field of image, RNN in the field of language and LSTM in the field of natural language processing.

3 Torch is an open-source deep learning framework of Facebook. The underlying code is written in C / C + +, and the call interface of the middle layer is implemented in Lua language. In addition, pytorch, a Python based API, is also provided. Because Lua scripting language is easy to learn, it's easy to get started with torch. Torch provides a large number of modular components to build neural network, which is also a neural network itself. Flexible architecture, easy to write new network layer, can easily realize personalized algorithm based on modular components. Torch provides a large number of pre training models, which can be used in more complex applications.

4 Mxnet is a popular deep learning framework, and has become Amazon's AWS default deep learning engine for cloud services. Compared with other frameworks, mxnet has more convenient multi machine and multi card operation. Mxnet is easy to install and friendly for beginners.

2.4. The application of confrontation generation network in face recognition (GAN)

As early as 2014, people are very enthusiastic about the research of neural network technology. At the same time, some rational scientists think that the output judgment of neural network has a very high risk and the output has a very high instability. In order to get rid of this lie, they simply add some noise on the correct pictures judged by neural network. For human beings, they don't detect the changes of images at all, but in neural network, they completely judge it as another object. At the same time, scientists claim that such highly deceptive images are not accidental, but can be mass produced, for example, through confrontation to generate networks.

The principle of Gan: use the generator to generate a picture, and use the discriminator to judge whether the model generated by the generator is "real". If it is true, the image generated by the generator is almost the same as the real image; otherwise, the

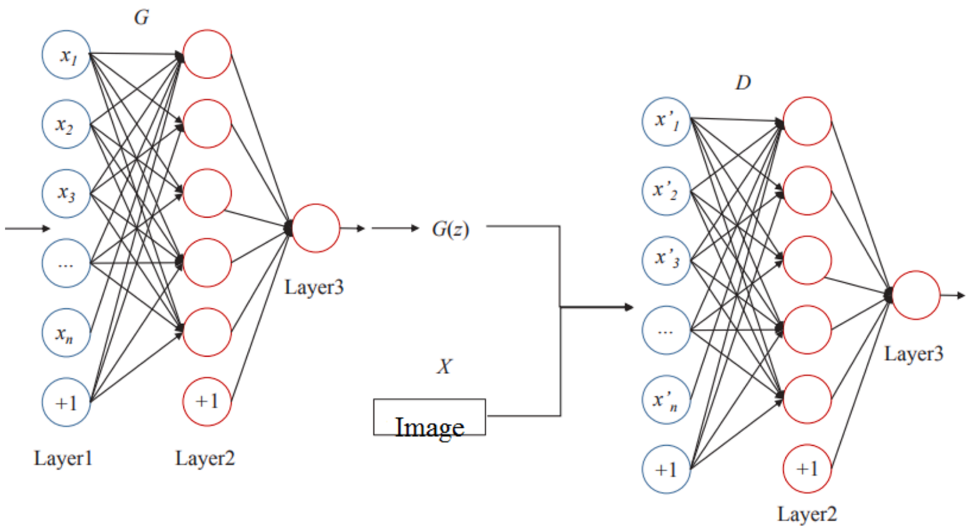

Fig. 8. The principle of antagonizing neural network training.

discriminator will judge that the image generated by the generator is not true. There is a "confrontation" relationship between the generator and the discriminator, and the confrontation is continuously strengthened with the increase of the picture library [13-16].

The training process of antagonizing neural network is shown in Fig. 8. It can be seen from Fig. 8 that the counter neural network trains two neural networks at the same time. The first neural network is a generator, whose input is generally the data generated by random distribution. The function of the generator is to generate false images, and then through the second neural network - discriminator, combined with the real images to distinguish the false images generated by the computer. The "real" scalar of false image is converted into probability by softmax function. If the probability is 0, it is false image, and the probability is 1, which means that the discriminator thinks the image is "real" and successfully "swindles" the discriminator.

In 2015, Radford et al. Proposed a deep convolution antagonism neural network, which is different from Gans in that both the generation model G(z) and the discrimination model D(x) are implemented by a convolution neural network. In DCGANs, the generative model is shown in Fig. 9.

When the size of the input image is too small, the generation model Gans is extended by interpolation between pixels through up sampling. When it is expanded to the set size and then convolution processing is carried out, the output size will be larger than the original size.

The discriminant model D(x) is also a deep convolution network. The input of the discriminant model is the output of the generated model. After a series of convolution operations, the final output is 1×1, which is probability P. If the probability p is close to 1, the image feature generated by the generated model is quite close to the feature of the real image; otherwise, there is a certain difference between the image feature generated by the generated model and the feature of the real image.

The judgment process of the discrimination model is shown in Fig. 10.

Figs. 11 and 12 present the flows.

3. The proposed methodology

3.1. An effective calculation model based on bilinear

Generally speaking, when implementing the Miller algorithm, the calculation efficiency can be improved from the following perspectives [17–20].

(1) Basic arithmetic in the base and extension fields. Obviously, any method to accelerate the basic arithmetic of the finite field in the bottom layer can speed up the calculation of bilinear pairs. Under the condition of security, we should choose the smallest base field as much as possible, and choose the irreducible polynomials which are propitious to multiplication and inversion in the extended field to construct the extended field.

(2) The base representation of prime number, which is beneficial to the operation of multiple points in elliptic curve group and the operation of multiple times of rational function assignment in Miller algorithm. For example, when calculating bilinear pairings in a hypersingular elliptic curve over a finite field characterized by 2 or 3, the prime number is represented by binary or ternary. In addition, the Hamming weight of R should be kept as small as possible. Obviously, the NaF representation of a number can reduce the addition of point addition and function assignment.

(3) Point addition and multiple point operation in elliptic curve group. In the whole calculation process of bilinear pairs, multiple point operation is needed. Therefore, the effective multiple point formula can accelerate the bilinear calculation.

(4) It can be seen from Miller algorithm that every cycle needs to calculate denominator and inverse. If the denominator value is in the extended domain and its inversion is quite complex, two intermediate variables F 1 and F 2 can be considered to replace one intermediate variable f until the last inversion.

(5) The number of cycles in the Miller algorithm. The less the number of cycles, the faster the calculation speed.

(6) The most direct acceleration idea is to use Frobenius mapping in finite field to accelerate the final power operation.

(7) Coordinate system selection of elliptic curve. Generally speaking, the projective coordinate system is better than the affine coordinate system when the time-consuming ratio of inversion and multiplication in the base domain is greater than 16. For special elliptic curves, special projective coordinates can be constructed to accelerate the calculation of bilinear pairs.

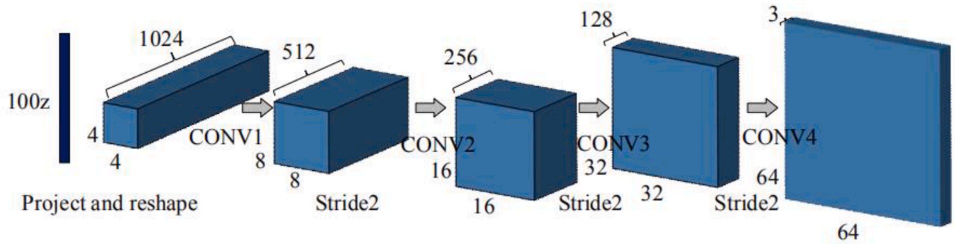

Fig. 9. Feature generation process of generation model.

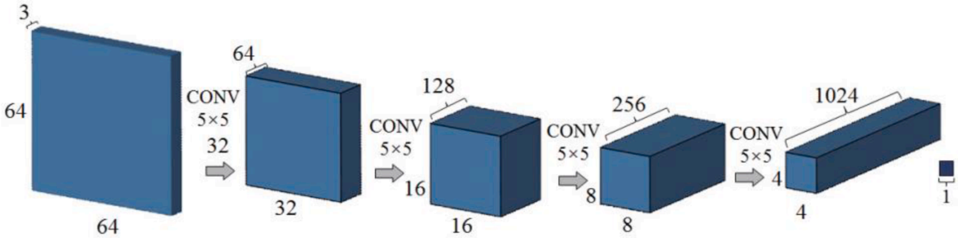

Fig. 10. Feature generation process of generation model.

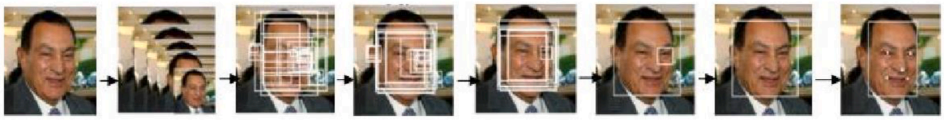

Fig. 11. MTCNN algorithm flow-chart.

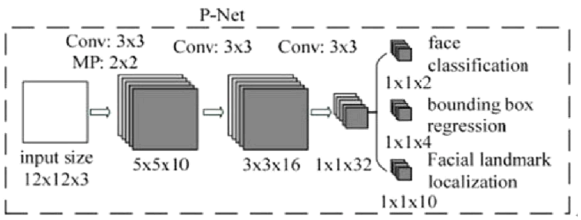

Fig. 12. P-net network structure.

3.2. System learning mode of face recognition

The process of face recognition includes: image acquisition, face detection, feature extraction, face image recognition.

Step one: face and face key point detection. Through the face detection algorithm, the face range and position are detected, and the face position coordinates and face key point coordinates are returned. In this paper, mtcnn based face and face key point detection algorithm is used to detect the location of face and face key point.
Step 2: face alignment. Face alignment is to calibrate the eyes, nose tip and mouth corner (five points in total) to the same position, which is the preprocessing of face recognition. The key point coordinates obtained in the previous process are used for affine transformation to complete the normalization of the face image.
The third step: face feature extraction. Through training the face feature extraction model based on deep learning, the face feature is extracted, which avoids the tedious feature selection by hand. In this paper, the feature extraction model based on Center loss is used to extract face features.
Step four: face feature comparison. Feature comparison is carried out by calculating cosine distance of feature vector to determine whether it is the same person.

This paper designs a face recognition system based on neural network. Its main functions are as follows:

(1) It can recognize faces in real time and efficiently, that is to find matching faces from the background database. If matching, the recognition results and matching degree are given. The function of face registration can be used to add face data and labels to the background database.
(2) Each module has high independence, the face detection module, the face preprocessing module and the face recognition module can be used separately, and the effect is good.
(3) The open interface is designed for each module, which can be compiled into dynamic connection library and called independently, which provides convenience for later program migration.

On the basis of ordinary camera recording image, the network camera encodes the video stream data according to the relevant protocol, uses the network transmission technology, and finally transmits the video image to the client. In order to achieve the purpose of real-time monitoring, a web-based operating system is also built in the computer. The client can access the camera according to the IP address assigned by the network camera, and can also use the relevant protocols to process and store the video data in real time. The network camera that supports multi angle shooting can also send instructions to the camera from the far end through the network, control the pan tilt and camera to rotate, and realize multi angle and omni-directional monitoring.

3.3. Face detection network: MTCNN

Similar to the traditional V-J cascaded AdaBoost method, mtcnn (multi task convolutional neural networks) adopts the idea of three class convolutional neural network (cascade CNN). In the form of multi task, the algorithm structure from coarse to fine is realized. Compared with the traditional face detection algorithm and single network structure, the speed is faster, the cost is smaller, and the detection is more accurate.

Data and processing: in order to solve the image scale invariance, the original image is scaled according to different scales, and the image pyramid is constructed as the input. The mtcnn consists of three network structures: p-net, R-Net and o-net. Through three cascaded networks in turn, boundary box regression and maximum suppression (NMS) are carried out, and finally the coordinates of face frame and face key points are obtained. The whole process is as follows:

The main function of p-net is to quickly generate candidate windows through a shallow CNN. As shown in Fig. 9, the network structure belongs to the full convolution structure, that is, it does not contain the full connection layer. Full convolution structure realizes the full connection layer in traditional CNN in the form of convolution layer: convolution core is 1×1 convolution layer. The biggest advantage of full convolution network is that it can input any size of image. The fixed size has the following problems:

(1) The selection of scale has certain subjectivity. For different targets, the most suitable size may be different.
(2) For images with different size and aspect ratio, forced transformation of fixed size will lose image information [21–23].
(3) The cropped image will destroy the integrity of the image, and the transformed image may lead to geometric deformation and distortion [24,25].

Full convolution network can improve the accuracy of face detection by accepting any scale of input. At the same time, it can reduce the problem of image size processing in training and testing.

4. Experiment and analysis

4.1. Experimental data analysis

In this experiment, the 0.12 version of Tensorflow is used as the platform. The face database chooses CASIA-Webface database for training. The database contains a total of 10,575 categories and 494,414 pictures, the size of which is next to Facebook's private database SCF. Some data of the database are shown in Fig. 13. Besides this, for the additional simulation, we verify the effectiveness of the algorithm in this article, the ORL face database and YALE face database were tested respectively. The ORL face database contains 400 images of 40 people, 10 images per person, and the size of the image is 112×92. YALE people as the face library contains 165 images of 15 people, 11 images for each person, the size of the image is 100×100. In order to avoid over-fitting, this experiment uses a five-fold cross-validation method. Each data set is randomly divided into five parts, four of which are used as training sets and the rest are used as test sets. This process was repeated five times, and the average of the five test results was used as the final recognition rate.

The network uses the same initialization parameters: the weight attenuation coefficient is 5e-5, the center error coefficient is 1e-4, the learning rate is initially set to 0.1, and then the attenuation is one tenth of every certain number of iterations. The training network iterates 100 times on all samples. When the center loss function is used as the classifier, the test results of the network in LFW dataset are shown in Fig. 14:

In the figure, the horizontal axis represents the number of iterations, the vertical axis represents the training accuracy of the depth neural network on the LFW database, and the final test accuracy of the network is 98.46%. When using the triple loss function as the classifier, the test performance of the network in LFW dataset is shown in Fig. 15:

Fig. 16 shows the visualization results based on Lenna head image and some features of CaffeNet:

It can be seen from the image that different convolution check source images extract different features, typical features can be roughly described as contour, texture and so on. With the deepening of the network, the information described by features is more abstract.

In order to facilitate the experiment, based on Caffe, this paper uses the open-source visualization tool deep visualization toolbox to operate. Through simple configuration, the visualization tool can observe the hidden layer output of convolution neural network and the original feature map contrast after deconvolution. The visualization results are shown in Fig. 17. The green marked out portion is the center of the clustering model. The Fig. 18–20. show the experimental result.

When the pose of the face changes, the system can detect the position of the face and mark it with the frame. The results are shown in the figure below.

In the process of recognition, if the picture uploaded by the user does not contain the face and so on, the steps of face detection fail, and the background will prompt the user to upload again. If the detection is successful, the neural network is used to extract the feature and calculate the feature distance with the face image in the database. Because the picture uploaded by the user does not exist in the database, a threshold value of 0.33 is set for the feature distance at this time. If the feature distance between the image to be recognized and all categories in the database is greater than the threshold value, the user will be prompted that it cannot be recognized, otherwise, the category name with the minimum feature distance will be returned. The experimental results are shown in the figure below.

Fig. 13. CASIA-Webface dataset.

k	Fisherfaces (UD)	CDI_Fisherfaces (UD)	Fisherfaces (OD)	CDI_Fisherfaces (OD)	DLDA	NLDA
2	75. 33	77. 77	79. 63	81. 24	37. 32	85. 32
3	86. 43	89. 05	91. 23	92. 20	85. 51	91. 64
4	90. 18	91. 93	94. 54	95. 03	92. 92	94. 69
5	91. 90	93. 38	95. 80	96. 32	94. 61	95. 98
6	93. 58	94. 65	97. 19	97. 58	97. 04	97. 02
7	94. 50	95. 19	97. 42	98. 06	97. 83	97. 53
8	95. 08	95. 71	97. 92	98. 17	98. 38	98. 04

Fig. 14. Center loss of LFW dataset.

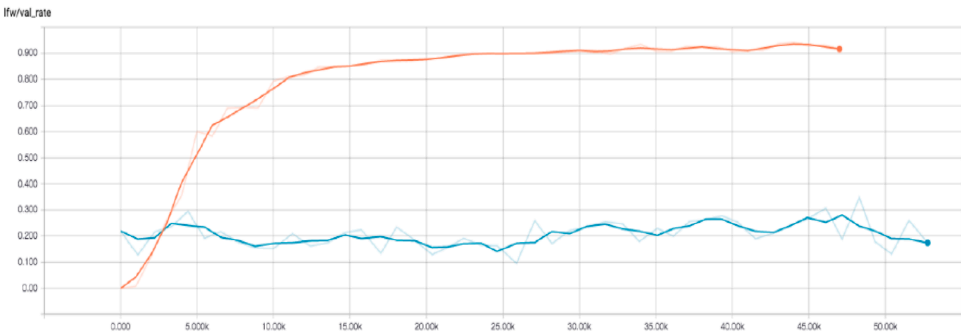

Fig. 15. Test precision of LFW dataset (triple loss).

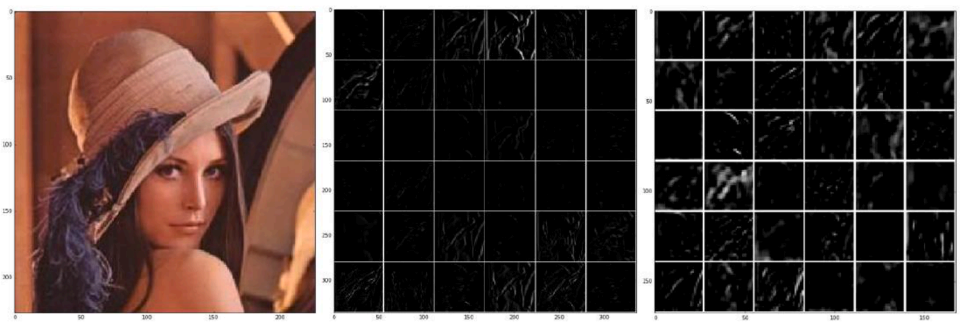

Fig. 16. Caffe visualization analysis results.

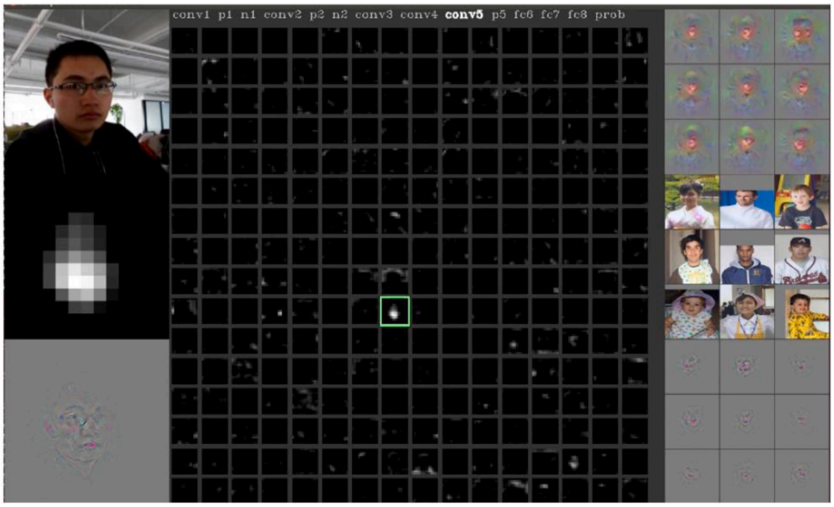

Fig. 17. Visual analysis results of neural network.

It can be seen that when there is no category corresponding to the face image uploaded by the user in the database, the feature distance calculated by the neural network is significantly greater than the threshold value. When there are face images uploaded by users in the database, neural network can effectively recognize the face categories in the images.

5. Conclusion

At present, deep learning plays an important role in the field of computer vision. Face recognition is one of the important research directions in the field of computer vision. The face recognition method based on deep learning has far exceeded the traditional method in recognition rate. In order to make computers think like people, machine learning develops a lot of image processing algorithms in the field of computer vision through feature extraction, which lays a solid theoretical foundation for the development of artificial intelligence. In this paper, the artificial neural network is briefly introduced, and some common activation functions and back propagation algorithms are described. Based on the above theory, this paper introduces the neural computing model and the counter neural network, and analyzes its application way. Face recognition based on deep learning does not need artificial design features, but self-learning internal abstract features, so deep learning method has high efficiency and universality. Based on this, this paper proposes a face recognition framework based on effective computation and counter neural network. The experimental results show that this method can effectively detect the face state. When the face pose changes, this method can also effectively recognize. In the future, we will test the proposed method on more data sets to validatethe robustness.

Author statement

Chenglin Yu is currently with the South China University of Technology, his research interests include image processing, data mining and pattern recognition.

Hailong Pei is currently with the South China University of Technology, his research interests include face analysis, patter analysis

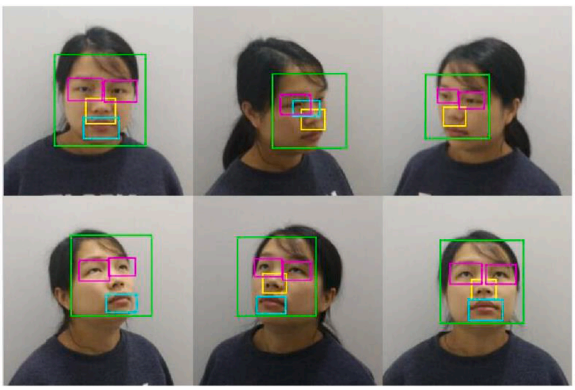

Fig. 18. Face detection results when posture changes.

Fig. 19. Classification results when no corresponding category exists in the database.

Fig. 20. Classification results when corresponding categories exist in the database.

and artificial intelligence.

Declaration of Competing Interest

The authors declare that they have no conflict of interest.

Acknowledgements

This research as supported by the following projects.

1. Supported by Scientific Instruments Development Program of NSFC [615278010]
2. Supported by Science and Technology Planning Project of Guangdong, China [2017B010116005].

References

[1] Wjun Li, Chongjun W, Wei Z, et al. A review of face recognition research. Pattern Recognit Artif Intell 2006;19(1):58–66.
[2] Zhao W, Chellappa R, Phillips PJ, et al. Face recognition: a literature survey. ACM Comput Surv 2003;35(4):399–458.
[3] Xiuping Z. A review of the development of biometric technology. Crim Technol 2011;(6):44–8.
[4] Taigman Y, Yang M, Marc, et al. DeepFace: closing the gap to human -level performance in face verification. Comput Vis Pattern Recognit IEEE 2014:1701–8.
[5] Pei W, Shang W, Liang C, Jiang X, Huang C, Yong Q. Using lignin as the precursor to synthesize Fe_3O_4@lignin composite for preparing electromagnetic wave absorbing lignin-phenol-formaldehyde adhesive. Ind Crops Prod 2020;154(2020):112638.
[6] Wan-hai X, Qian-nan Z, Wen-chen M, En-hao W. Response of two unequal-diameter flexible cylinders in a side-by-side arrangement: characteristics of FIV. China Ocean Eng 2020;34(4):475–87.
[7] Sutskever I, Martens J, Dahl G, et al. On the importance of initialization and momentum in deep learning. In: International conference on machine learning; 2013.
[8] Yosinski J, Clune J, Bengio Y, et al. How transferable are features in deep neural networks? Eprint Arxiv 2014;27:3320–8.
[9] Hinton GE, Salakhutdinov RR. Reducing the dimensionality of data with neural networks. Science 2006;313(5786):504–7.
[10] Kim SK, Park YJ, Toh KA, et al. SVM-based feature extraction for face recognition. Pattern Recognit 2010;43(8):2871–81.
[11] Bengio Y. Learning deep architectures for AI. Found Trends Mach Learn 2009;2(1):1–127.
[12] Wan-hai X, Qian-nan Z, Wen-chen M, En-hao W. Response of two unequal-diameter flexible cylinders in a side-by-side arrangement: characteristics of FIV. China Ocean Eng 2020;34(4):475–87.
[13] Bo Y, Yequan F. Liu min. image reconstruction algorithm based on deep convolution neural network. Comput Syst Appl 2018;27(9):170–5.
[14] Yan Ke, Xilong W, Yuhui Z. Deep convolution generation countermeasure network structure. Electron Technol Softw Eng 2018;(24):21–2.
[15] Bin L. Face detection research based on confrontation depth learning. Qingdao: Qingdao University of science and technology; 2018.
[16] Ju Li. Face grayscale image coloring based on generative antagonism network. Comput Knowl Technol 2018;14(11):185–7.
[17] Chen Z, Ahn H. Item response theory based ensemble in machine learning. 2019. arXiv preprint arXiv:1911.04616.
[18] Paterson KG. Cryptography from pairing-advances in elliptic curve cryptography. Cambridge: Cambridge University Press; 2005. p. 215–52.
[19] Stange KE. The Tate pairing via elliptic nets. In: Takagi T, editor. Proc. of the pairing 2007. Berlin, Heidelberg: Springer-Verlag; 2007. p. 329–48. LNCS 4575.

[20] Hess F, Smart P, Vercauteren F. The Eta pairing revisited. IEEE Trans. Inf Theory 2006;52(10):4595–602.
[21] Que S, Awuah-Offei K, Demirel A, Wang L, Demirel N, Chen Y. Comparative study of factors affecting public acceptance of mining projects: evidence from USA, China and Turkey. J Clean Prod 2019;237:117634.
[22] Hao Y, Wang N, Li J, Gao X. HSME: hypersphere manifold embedding for visible thermal person re-identification. Proc AAAI Conf Artif Intell 2019;33:8385–92.
[23] Chen Q, Zhang G, Yang X, Li S, Li Y, Wang HH. Single image shadow detection and removal based on feature fusion and multiple dictionary learning. Multimed Tools Appl 2018;77(14):18601–24.
[24] Pavlichin DS, Jiao J, Weissman T. Approximate profile maximum likelihood. J Mach Learn Res 2019;20(122):1–55.
[25] Wu J, Chen T, Wu H, Yang Z, Wang Q, Lin L. Concrete image captioning by integrating content sensitive and global discriminative objective. In: 2019 IEEE international conference on multimedia and expo (ICME). IEEE; 2019. p. 1306–11.

Chenglin Yu is with the Key Laboratory of Autonomous Systems and Networked Control,Ministry of Education, Unmanned Aerial Vehicle Systems Engineering Technology Research Center of Guangdong, South China University of Technology. His research focuses are machine learning, information theory and data mining.

Hailong Pei is with the Key Laboratory of Autonomous Systems and Networked Control,Ministry of Education,Unmanned Aerial Vehicle Systems Engineering Technology Research Center of Guangdong, School of Automation Scinece and Engineering, South China University of Technology. His research focuses are machine learning, information theory and data mining.