Research on Face Recognition Technology of Subway Automatic Ticketing System based on Neural Network and Deep Learning

Shuang Wu[1], Xin Lin[2], Tong Yao[3*]
Department of Transportation Engineering, Yangzhou Polytechnic Institute, Yangzhou, China[1, 2]
School of Economics and Management, Sichuan Tourism University, Chengdu, China[3]

***Abstract*—Face recognition technology is the core technology of the subway ticketing system, which is related to the efficiency of people's ticket purchase. In order to improve people's experience of taking public transport, it is necessary to improve the performance of face recognition technology. In this study, the Back Propagation (BP) algorithm is used to optimize the parameters of the SoftMax classifier of the convolutional neural network, and the branch structure is added to the structure of the SphereFace-36 convolutional neural network to extract the local features of the face. Based on the improved neural network, the face recognition system of the subway automatic ticketing system is established. The results show that the area under the ROC curve is the highest for validation and identification of the optimization model; The recognition accuracy of the optimized model in different data sets is 1.0%, 0.7%, 1.1%, 0.9% and 0.6% higher than that of SphereFace-36 respectively, and its specificity is higher than that of SphereFace-36, with the maximum difference of 9%; The average accuracy of global feature extraction and recognition of the optimized network model is 83.01%. In the simulation experiment, the optimized model can accurately recognize facial features, which has high practical value and can be applied to the automatic ticketing system.**

Keywords*—Automatic ticketing system; BP; CNN; deep learning; face recognition; SphereFac； SoftMax classifier

I. Introduction

Face authentication does not require physical contact with other systems, has a high false acceptance rate and false rejection rate, and can minimize errors [1]. Because of these unique characteristics of face recognition systems, it is considered one of the effective methods of data protection. Face recognition technology is widely used in many fields such as security, medical care, education, etc., and face recognition is required in various social places, which is very important for security, transportation, etc. [2]. Face recognition is mainly divided into appearance recognition and feature recognition. By extracting facial features, it is judged whether it matches the face information in the image or video stream, to identify the identity [3]. However, in the real environment, face recognition technology will be affected by posture, gender, lighting, and other conditions, so it is necessary to improve and optimize the method to improve face recognition's accuracy [4]. Changes in head posture will bring challenges to face recognition technology, and an unconstrained environment will increase the difficulty of face recognition in the field of human-computer interaction. Researchers designed a tool to collect key pose samples from face datasets and then identify changing head poses through cascaded convolutional neural networks [5]. With the continuous development of society, people have higher and higher requirements for travel quality. Face recognition technology can help people save time cost of transportation. In the subway automatic ticketing system, fast ticket purchases can be realized by scanning the face. The application of face recognition technology has significantly improved the service efficiency and quality of the subway ticketing system. Due to factors such as the new crown epidemic, wearing masks for self-protection in public places is necessary, which makes the face recognition technology of the automatic ticketing system of the subway more difficult. As an important factor, the occlusion problem can affect the accuracy of facial landmark detection in face recognition, and the occlusion of the face affects the application of face recognition [6]. To improve the face recognition technology of the subway automatic ticketing system, the neural network was optimized in this experiment, and the accuracy of the face recognition system was improved through deep learning technology, hoping to meet people's travel needs. To improve the face recognition technology of the subway ticketing system, the neural network was optimized in this experiment. The research content of this paper is mainly to optimize the parameters of the SoftMax classifier of the convolutional neural network using the Error Back Propagation (BP) algorithm, and add a branch structure on the basis of the structure of the SphereFace-36 convolutional neural network to extract the local features of the face. Then, based on the improved neural network, the face recognition system of the subway ticketing system is established. Finally, the performance of the optimized neural network and face recognition system is simulated and analyzed. It is hoped that the accuracy of face recognition system can be improved through deep learning technology to meet people's travel needs.

II. Related Work

In promoting the development of machine learning and deep learning, the application of CNN plays an important part and can reduce the complexity of feedback nerves. With in-depth research of scientific research, CNN has been continuously optimized and improved, and it has been effectively applied in various fields such as medicine, autonomous driving, and face recognition. In the technical research of heartbeat detection in the medical field, Chandra B

*Corresponding Author.

et al. applied a convolutional neural network for multimodal data analysis, and CNN can effectively improve the robustness and accuracy of heartbeat detection [7]. In disease classification, the bilateral CNN method can effectively distinguish dystrophic scoliosis from non-dystrophic scoliosis. This method has good classification accuracy, has strong application value, and provides a great deal for clinical diagnosis. Good technical support [8]. O Elharrouss and other scholars cascade the CNN technology, perform pose estimation, face segmentation, and recognition through three CNN models, and realize face recognition from the left, front and rear directions of the human face. The model training results show that the new model has high robustness [9]. As one of the ways of people’s daily communication, image data shows explosive growth with network technology’s continuous development. Machine vision is based on image recognition. CNN can perform feature extraction and recognition of image information, thereby extracting and analyzing different images. However, when recognizing face images, CNN may lose some image feature information during propagation and computation, and the multi-level information fusion method can further recover the information lost in CNN propagation and computation [10]. SoftMax classifier is the main technology of convolutional neural network to realize feature classification. The enhanced algorithm based on SoftMax classifier can effectively improve the accuracy and processing time of face recognition [11]. BP algorithm can be used to train artificial neural network to realize face recognition process. To improve the performance of face recognition system, artificial neural network is combined with the famous meta heuristic optimization algorithm, namely harmony search algorithm (HSA) [12]. Genetic algorithm, particle swarm optimization algorithm, etc. are heuristic algorithms, which can effectively solve the combination problems by examples. In recent years, researchers have obtained meta heuristic algorithm by combining random algorithm with local search algorithm. This algorithm can explore the solution space more thoroughly, so as to obtain a better solution. For example, Dwarf Mongoose Optimization Algorithm [13], the arithmetic optimization algorithm [14], Reptile Search Algorithm (RSA) [15], and other optimization algorithms are all improvements to heuristic algorithms, which can be independent of problems. Although meta heuristic algorithm has certain advantages, its intrinsic parameters still need to be adjusted and further studied in order to make the technology better used for problem solving.

Led by the rapid development of computer science and other fields, face recognition technology has become increasingly sophisticated and perfected. In various industries and fields, face recognition technology has been well applied. For example, based on face recognition technology such as neural networks and UAV, the online attendance system for students in the education industry is designed, and the research on the estimation of long-distance absolute distance and target tracking of face recognition technology is designed [16-17]. Some literature has conducted in-depth discussions on the lateralization of face recognition. The study believes that although the left hemisphere has achieved language evolution and is also the realization area of facial recognition, there is no evidence that the right hemisphere does not correlate with face recognition [18]. Image feature extraction is a key technology for face image classification in face recognition systems. Nejrs S et al. applied two-dimensional discrete Fourier transform technology to face image classification. The performance difference comparison method based on the correlation coefficient is effective. The classification accuracy of the target image reaches more than 90%, which proves that the classification result of the new model method is accurate [19]. Face recognition technology includes the recognition of facial expressions, and some scholars use machine learning algorithms to optimize face recognition technology. By using the gradient direction histogram to collect facial emotion data, and then applying the fast learning network algorithm to classify the extracted features for facial emotion recognition, the recognition of people’s facial expressions and emotions is realized [20]. CNN is an efficient and accurate technical method in face recognition method and has high robustness [21]. To improve the distinguishing degree of feature extraction in face recognition technology, researchers use the optimized loss function to train the existing neural network and use the multi-action form supervision force to supplement the origin loss, which promotes the feature learning of face recognition [22]. Face recognition technology is easily affected by factors such as illumination and environment, which reduces the accuracy of face recognition. Chen C et al. proposed to use the fuzzy discriminant analysis method to establish a face recognition system, which can reduce the influence of the environment on the face recognition technology to a certain extent. This method calculates the fuzzy membership degree of the sample, calculates the fuzzy mean value, and then establishes a new model to achieve face feature extraction [23]. Some scholars have established a face beauty prediction model with a simple structure and suitable for embedding in small devices. The model uses CNN technology to extract facial features and expands the face database to predict facial beauty [24].

Facial sign detection is an important step in face recognition technology. Under the influence of the COVID-19, masks have become a necessity for people's daily travel, and mask occlusion has brought challenges to facial sign detection. Based on the above research, it can be seen that CNN technology has a good application effect in face recognition technology. At the same time, in order to improve the efficiency of CNN face recognition, other methods can be introduced to optimize the parameters. In this experiment, BP neural network is introduced to optimize the weight matrix and bias parameter of SoftMax classifier in CNN algorithm, and branch structure is added to the structure of SphereFace-36 convolution neural network to extract local features of face. It is hoped that the optimization and improvement of the method can improve the efficiency of the subway ticketing system and provide convenience for people to purchase tickets.

III. Face Recognition Technology of Subway Automatic Ticketing based on Neural Network and Deep Learning

A. *Face Recognition Technology based on Neural Network and Deep Learning*

Due to the impact of the epidemic, people are required to wear masks when choosing public transportation, which puts forward higher requirements for the face recognition technology of the subway ticketing system. The face recognition system needs to achieve high accuracy of face recognition when ticket buyers wear masks and other facial masks. In the existing CNN face recognition systems, the extraction of individual facial sign features can improve the accuracy of face recognition and better classify features [25-28]. CNN is a feedforward neural network model, and convolutional neural networks can be used to detect targets in face recognition technology. CNN can repeatedly cut the target image through different sliding windows and then perform feature extraction and recognition. The CNN network layer includes three network layers: convolutional layer, downsampling layer, and fully connected layer. The weight matrix *W* and bias term *b* parameters in the convolutional layer need to be trained and learned, and formula (1) is the calculation formula of convolution.

$$y_{mm} = f(\sum_{j=0}^{J-1}\sum_{i=0}^{I-1} x_{m+i,n+j} W_{ij} + b) \quad (1)$$

i in formula (1) represents the *ith* parameter, *jth* parameter is represented by *j*, and the samples' total number is represented by *m*. The subsampling layer can use a nonlinear downsampling method to reduce the amount of input data, thereby the feature map's scale is reduced without affecting the feature extraction of the target. If the size of uniform sampling is $S_1 \times S_2$, then the convolution formula (2) can be obtained.

$$y_{mm} = \frac{1}{S_1 S_2}\sum_{j=0}^{S_2-1}\sum_{i=0}^{S_1-1} x_{m\times S_1+i, n\times S_2+j} \quad (2)$$

As a CNN classifier, the fully connected layer can be implemented by convolution operations in practical applications, and is widely used in deep learning network models, but there is also the problem of parameter redundancy. When performing multi-classification, the SoftMax multi-classification layer is added to the convolutional neural network. Assume that the training set sample $T = \{(x^{(1)}, y^{(1)}),...,(x^{(m)}, y^{(m)})\}$, which $x^{(i)}$ represents the sample *i*, $y^{(i)}$ is the label of the sample. For the training set *T*, the SoftMax classifier can determine the category of the samples in it. In the case of rule restrictions, the probability $P(y = j|x), (j = 1,....k)$ of converting to the *kth* sample is calculated by the formula, and then the sample category is calculated for face recognition, and the calculation formula of SoftMax can be obtained, see formula (3).

$$f(x^{(i)}|\theta) \begin{bmatrix} p(y^{(i)}=1|x^{(i)},\theta) \\ p(y^{(i)}=2|x^{(i)},\theta) \\ \vdots \\ p(y^{(i)}=k|x^{(i)},\theta) \end{bmatrix} = \frac{1}{\sum_{j=1}^{k} e^{\theta_j^T x^{(i)}}} \begin{bmatrix} e^{\theta_1^T x^{(i)}} \\ e^{\theta_2^T x^{(i)}} \\ \vdots \\ e^{\theta_k^T x^{(i)}} \end{bmatrix} \quad (3)$$

$\theta = [\theta_1^T \theta_2^T ... \theta_k^T]$ representation learning parameters in Eq. (3). The network model is continuously trained in the data set *T*, and the optimal one is obtained through iterative calculation and fitting θ. Formula (4) is the loss function used in the training.

$$J(\theta) = -\frac{1}{m}[\sum_{i=1}^{m}\sum_{j=1}^{k} 1\{y^{(i)} = j\} \log \frac{e^{\theta_j^T x^{(i)}}}{\sum_{i=1}^{k} e^{\theta_j^T x^{(i)}}}] \quad (4)$$

In the formula (3), $1\{y = j\}$ means that its value is 1 in case of $y = j$, and 0 in case of $y \neq j$. By calculating network prediction and real value's error, the loss function's value can be obtained and the optimization process of the neural network can make the loss function minimize. The direction of signal transmission is forward, and the propagation direction of the error is opposite to it, which are BP's characteristics. The parameters' reverse adjustment in the network layer can be achieved by BP. Fig. 1 shows the BP network's structure.

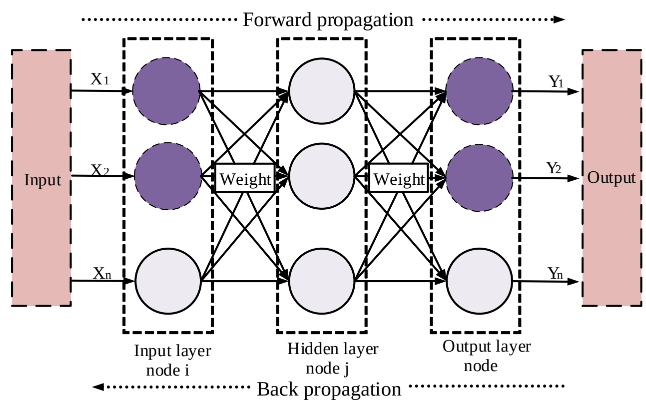


Fig. 1. Structure diagram of BP neural network

Through the continuous adjustment of the network weights in the BP neural network, the final network output close to the expected output can be obtained, that is, the error between the network prediction result and the given label is continuously reduced. Assuming that there are a total of *n* samples with a given label in the training set $\{(x_1, y_1), (x_2, y_2),...., (x_n, y_n)\}$, the neural network's training method is gradient descent. The variance loss function of the model is shown in formula (5) for the samples (x_i, y_i) in the training set.

$$L(w, b; x_i, y_i) = \frac{1}{2}\|h_{w,b}(x_i) - y_i\|^2 \quad (5)$$

In formula (5), w represents the weight and b represents the bias. The variance loss function can calculate the error, which can be used to evaluate the convergence situation. The loss function of n samples can be derived by formula (6).

$$\begin{aligned} L(w,b) &= \left[\frac{1}{n}\sum_{i=1}^{n} L(w,b;x_i,y_i)\right] + \frac{\lambda}{2}\sum_{l=1}^{n_l-1}\sum_{i=1}^{S_l}\sum_{j=1}^{S_l+1}(w_{ji}^{(l)})^2 \\ &= \left[\frac{1}{n}\sum_{i=1}^{n}\left\|h_{w,b}(x_i)-y_i\right\|^2\right] + \frac{\lambda}{2}\sum_{l=1}^{n_l-1}\sum_{i=1}^{S_l}\sum_{j=1}^{S_l+1}(w_{ji}^{(l)})^2 \end{aligned} \tag{6}$$

In formula (6), i represents the *ith* parameter, j represents the network layer unit of the previous layer, l represents the *lth* layer of the neural network, and S represents the variance. The parameters λ are introduced to adjust the loss value of the variance loss function and the proportion of the attenuation term. In the process of BP neural network training, it is necessary to continuously fine-tune the size of parameters w and b. The derivation formula of parameter w *is shown* in formulas (7) and (8).

$$w_{ij}^{(l)} = w_{ij}^{(l)} - \alpha\frac{\partial}{\partial w_{ij}^{(l)}}L(w,b) \tag{7}$$

In formula (7), α denotes the learning rate, and the step size of gradient descent can be controlled by adjusting the learning rate. The derivation formula of parameter b is shown in formula (8).

$$b_i^{(l)} = w_{ij}^{(l)} - \alpha\frac{\partial}{\partial b_i^{(l)}}L(w,b) \tag{8}$$

Through formulas (7) and (8), the partial derivatives of all sample parameters w and b can be deduced, and the formula for partial derivatives of all sample parameters w is shown in formula (9).

$$\frac{\partial}{\partial w_{ij}^{(l)}}L(w,b) = \left[\frac{1}{n}\sum_{i=1}^{n}\frac{\partial}{\partial w_{ij}^{(l)}}L(w,b;x_i,y_i)\right] + \lambda w_{ij}^{(l)} \tag{9}$$

The partial derivative formula of all sample parameters b is shown in formula (10).

$$\frac{\partial}{\partial b_{ij}^{(l)}}L(w,b) = \left[\frac{1}{n}\sum_{i=1}^{n}\frac{\partial}{\partial b_{ij}^{(l)}}L(w,b;x_i,y_i)\right] \tag{10}$$

In the BP neural network, according to the input sample signal of forward propagation, the two parameters of weight and bias are randomly initialized, and the loss value is calculated. The partial derivatives of the optimization function to the neuron weights in the neural network of each layer can be obtained in the process of backpropagation, to obtain the error signals of all neurons, as the basis for modifying the parameters w and b. Through the construction of a multi-layer neural network, the basic neural network can be continuously developed to form deep learning. In the field of image recognition, underlying features of the image can be detected by deep learning in the low-level neural network layer, then integrate and abstractly combine to generate the corresponding high-dimensional image features, and finally realize the classification of high-dimensional features at the top layer. When the deep neural network propagates forward, the corrected weight value and bias value of the neural network can be obtained by calculating the error, and then the weight and bias item are corrected by backpropagation.

B. *Face Recognition Technology for Subway Automatic Ticketing based on Neural Network and Deep Learning*

The application of face recognition technology in many scenarios is limited by many conditions, and the clarity of photos is one of the limiting factors. In the actual subway automatic ticketing face recognition scene, the face recognition technology will reduce the accuracy of face recognition due to the influence of masks, glasses, hats, and other obstructions, expressions, light, and other factors. To make the robustness of the face recognition technology for subway ticket vending improve, CNN was optimized in this experiment. A branch structure was added to the structure of SphereFace-36 to get the face's local features. Fig. 2 shows the network structure after the fusion of local features and global features. The resolution of facial details can be improved after local and global features are fused.

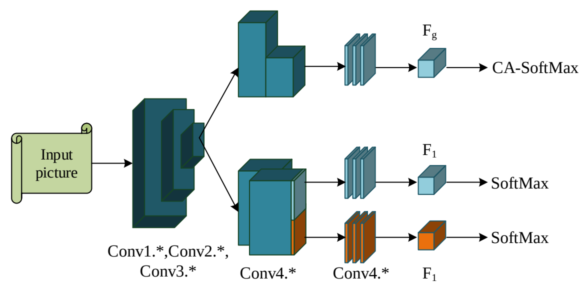


Fig. 2. Structure diagram of a network

One branch in the network structure diagram is the SphereFace-36, which can get the face's global features, and the other branch divides the feature image into upper and lower parts, which can get the face detailed features. The loss functions applied in these two branch network diagrams are CA-SoftMax (Center Angular SoftMax Loss) and SoftMax Loss respectively. In the original SoftMax Loss, it needs to be satisfied $w_1^T x > w_2^T x$, ie $\|w_1\|\|x\|\cos(\theta_1) > \|w_2\|\|x\|\cos(\theta_2)$. To achieve the effect of increasing the distance between classes, the variable m was added in this experiment to strictly limit the decision boundary of the classification, see formula (11).

$$\|w_1\|\|x\|\cos(m\theta_1) > \|w_2\|\|x\|\cos(\theta_2) \tag{11}$$

In the range of [0, π], since the *cos* function decreases monotonically, the larger the value of variable m is. Under the condition that and x are unchanged, when the classification is correct, the smaller the angle of θ_1, the smaller the angle between the sample feature vector and w_1, increase the distance between classes. The separability of the feature angle is a key part of Angular SoftMax Loss, and weight normalization can improve the separability of samples. After the parameter w is normalized by the weight, the value is 1, that is $\|w_i\| = 1$, the bias term b is 0, and formula (12) is the loss function.

$$L = \frac{1}{N}\sum_{i} -\log\left(\frac{e^{\|x_i\|\phi(\theta_{yi})}}{e^{\|x_i\|\phi(\theta_{yi})} + \sum_{j \neq y_i} e^{\|x_j\|\cos(\theta_j)}}\right) \quad (12)$$

$$\phi(\theta_{yi}) = (-1)^k \cos(m\theta_{yi}) - 2k, \theta_{yi} \in \left[\frac{k}{m}\pi, \frac{(k+1)}{m}\pi\right], k \in [0, m-1], m \geq 1$$

In Angular SoftMax Loss, the expression effect of facial features can be improved, but the convergence speed of the function is slow during training, and it is easy to fall into a local minimum. The experiment updated the weights in Angular SoftMax Loss, that is, the sample center, named CA-SoftMax Loss. In formula (13), at time t+1, the expected weight value is the mean value of all sample features and their weight values that conform to $y_i = j$ at time t.

$$w_j^{t+1} = \frac{w_j^t + \sum_{i=1}^{m} \delta(y_i = j)\cdot(x_i)}{1 + \sum_{i=1}^{m} \delta(y_i = j)} \quad (13)$$

The mean value of Euclidean distance is approximately equal to the mean value of angular distance after weight normalization. The updated value of the sample center can be obtained by formulas (14) and (15).

$$w_j^{t+1} - w_j^t = \frac{\sum_{i=1}^{m} \delta(y_i = j)\cdot(x_i) - \delta(y_i = j)\cdot w_j^t}{1 + \delta(y_i = j)} = \frac{\sum_{i=1}^{m} \delta(y_i = j)\cdot(x_i - w_j^t)}{1 + \delta(y_i = j)} \quad (14)$$

$$\Delta w_j = w_j^{t+1} - w_j^t = \frac{\sum_{i=1}^{m} \delta(y_i = j)\cdot(w_j^t - x_i)}{1 + \delta(y_i = j)} \quad (15)$$

To control the step size of the sample center update, a parameter is introduced during the update α, and it is set to 0.5, which achieves the purpose of preventing the jitter of the class center, see formula (16).

$$w_j^{t+1^*} = w_j^t - \alpha\Delta w_j \quad (16)$$

Based on neural network and deep learning, the face recognition system of subway automatic ticketing was designed and developed in the experiment. The process flow of the face recognition system is shown in Fig. 3. The overall framework of the system includes four modules: video image acquisition, face detection, normalization, and face recognition. The system needs to perform face registration first, and then extract the face features, and then save the acquired information and features into the database. When performing face recognition, it should extract the face information and features in the video and make it compared with the database's information and features to determine whether they belong to the same person.

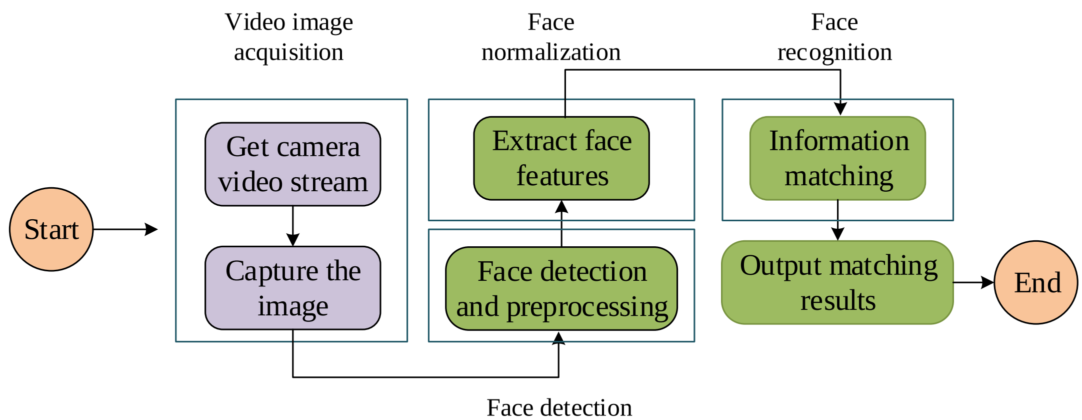
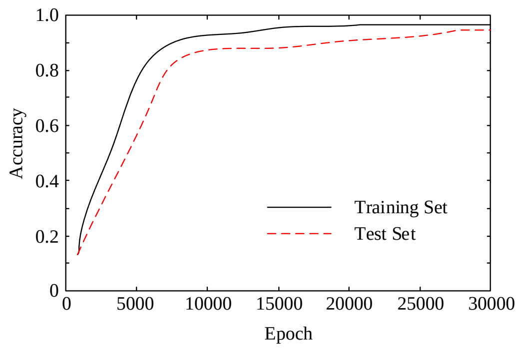


Fig. 3. Flow chart of face recognition

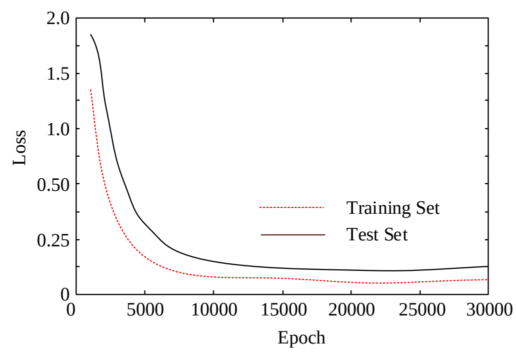
IV. Simulation Analysis of Face Recognition System for Subway Automatic Ticketing

To improve the classification performance of the improved CNN, the improved method is trained in the BLUFR general data set, and the training is completed in the training set and the test set respectively. It can be seen from Fig. 4 that the accuracy rate of the optimized CNN model tends to be stable when it is trained to about 8000 times. The final accuracy rate in the training set is 98.5%, and the final accuracy rate in the test set is 97.6%. The loss function value of the optimized CNN model tends to be stable when it is trained to about 5000 times. The loss function value in the training set is 0.09, and the loss function value in the test set is 0.18. It shows that the optimized CNN model can get higher accuracy and lower loss function value after training, and the algorithm has better performance.

In the process of neural network training, each weight value needs to be assigned in advance, that is, weight initialization. The method of weight assignment is very important to the convergence speed of the neural network model and the performance of the model. In this study, the original Angular SoftMax Loss has been updated with weights, and the form of the original function has not been changed. The weights can be updated faster and more efficiently after the weights are updated. In the data set WebFace, the loss function in the face recognition system is compared, and SphereFace-36 is selected as the test neural network, the parameter settings are consistent, and the initial learning rate is 0.01. The difference between the optimized CA-SoftMax Loss and the original loss function is shown in Fig. 5.



(a) Accuracy



(b) Loss

Fig. 4. Performance verification diagram of improved CNN model

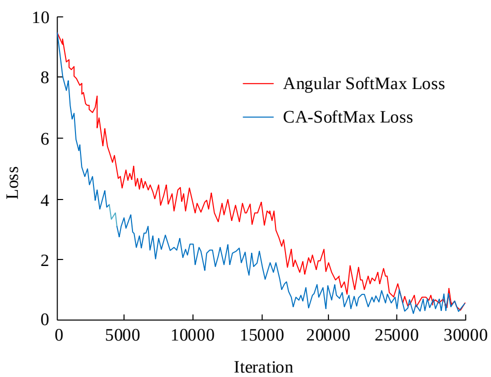
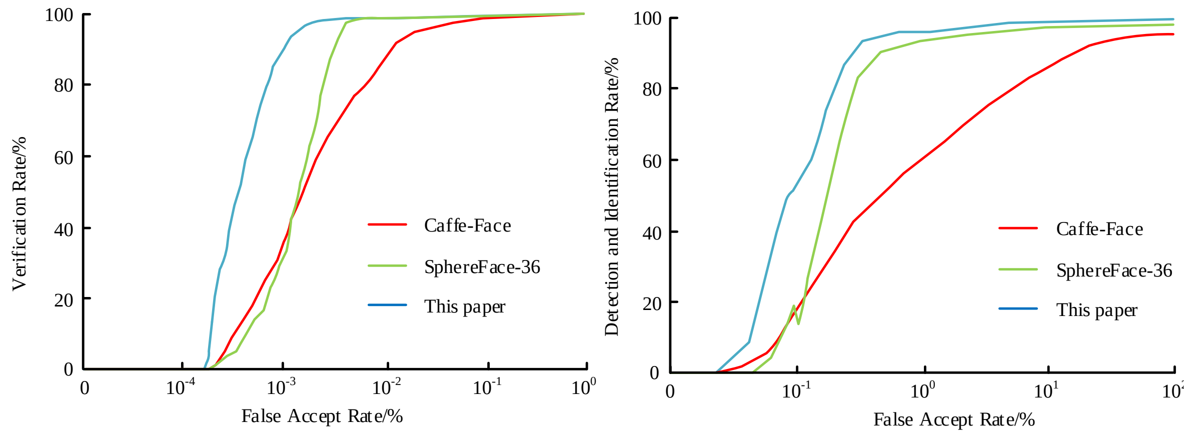


Fig. 5. Loss function curve

From Fig. 5, the initial Loss value of the optimized CA-SoftMax Loss and Angular SoftMax Loss is about 9.2. With the iterations number and pictures number increased, the Loss value decreases rapidly. The optimized CA-SoftMax Loss converges faster than the original Angular SoftMax Loss loss function, and the Loss value does not change or increase abnormally due to network gradient dissipation or gradient explosion during the descent process. In the experiment, we chose to test the model accuracy in the BLUFR general data set, LFW general data set, AgeDB-30 general data set, CFP-FP general data set and YTF general data set, and compared SphereFace-36, Caffe-Face, and optimization. The verification rate and recognition rate of the post-branch network structure are shown in Fig. 6.

Fig. 6 (a) and Fig. 6 (b) are the validation ROC curve and the identification ROC curve of the validation model, respectively. The area under the optimized branch network structure verification ROC curve and identification ROC curve is higher than that of the SphereFace-36 and Caffe-Face models, indicating that the optimized network structure proposed in this experiment has a better Good face recognition effect. In BLUFR general dataset, LFW general dataset, AgeDB-30 general dataset, CFP-FP general dataset and YTF general dataset, SphereFace-36, Center Loss, Angular SoftMax Loss, Caffe-Face. In this experiment, the optimization model is used to compare the algorithm recognition performance, and the results are shown in Table I.

The "-" in Table I indicates that the algorithm has not been tested for recognition performance in the corresponding dataset. It can be seen from the results in different datasets that the accuracy of the basic SphereFace-36 algorithm is lower than that of the optimized network model. In the BLUFR dataset, LFW dataset, AgeDB-30 dataset, CFP-FP dataset, and YTF dataset, the recognition accuracy of the optimized model is improved by 1.0%, 0.7%, 1.1%, 0.9%, and 0.6%, respectively. The evaluation indicators of the performance comparison in the experiment selected the true rate, false positive rate, prediction accuracy rate, and average accuracy rate, respectively in the BLUFR general data set, LFW general data set, AgeDB-30 general data set, CFP-FP general data set Compared with the YTF general data set. Table II shows the test results of SphereFace-36 and the optimized model in this experiment.



(a) Verification of ROC curve

(b) Detection and Identification of ROC curve

Fig. 6. ROC curve

TABLE I. ALGORITHM PERFORMANCE CONTRAST

Method	BLUFR	LFW	AgeDB	CFP	YTF
SphereFace-36	98.70%	93.25%	94.33%	92.01%	95.17%
Center Loss	98.31%	93.01%	-	92.57%	93.62%
Angular SoftMax Loss	98.12%	-	93.58%	-	94.83%
Caffe-Face	99.13%	92.87%	94.27%	92.33%	95.10%
This paper	99.70%	93.95%	95.43%	93.11%	95.77%

TABLE II. INDEX COMPARISON RESULTS

Data set	Method	True positive (%)	False positive (%)	Prediction accuracy (%)	Average accuracy (%)
BLUFR	SphereFace-36	84.84	9.79	90.59	87.87
	This paper	93.93	4.04	96.85	94.94
LFW	SphereFace-36	83.14	9.60	88.78	86.11
	This paper	92.05	3.95	94.92	93.04
AgeDB-30	SphereFace-36	78.98	9.12	84.34	81.80
	This paper	87.44	3.76	90.17	88.38
CFP-FP	SphereFace-36	86.88	10.03	92.78	89.98
	This paper	96.19	4.13	99.19	97.22
YTF	SphereFace-36	84.84	9.79	90.59	87.87
	This paper	93.93	4.04	96.85	94.94

From Table II, in BLUFR, LFW, AgeDB-30, CFP-FP, and YTF general data set, the true rate of optimizing the branch network structure is 93.93%, 92.05%, 87.44%, 96.19%, 93.93%. In the BLUFR general dataset, the LFW general dataset, the AgeDB-30 general dataset, the CFP-FP general dataset, and the YTF general dataset, the false positive rates of the optimized branch network structure are 4.04%, 3.95%, 3.76%, and 4.13%, 4.04%, respectively. In BLUFR, LFW, AgeDB-30, CFP-FP, and YTF general data set, the prediction accuracy of optimized branch network structure is 96.85%, 94.92%, 90.17%, 99.19%, 96.85%, respectively. In the BLUFR general dataset, the LFW general dataset, the AgeDB-30 general dataset, the CFP-FP general dataset, and the YTF general dataset, the average accuracies of optimizing the branch network structure are 94.94%, 93.04%, 88.38%, 97.22%, and 94.94%, respectively. In each general dataset, the true rate, prediction accuracy, and optimized network model's average accuracy are higher than those of the basic SphereFace-36; in each general dataset, the false positive rate of the optimized branch network is lower than that of the basic SphereFace-36. In Fig. 7, the sensitivity and specificity of the basic SphereFace-36 and optimized network models are tested on the BLUFR general dataset, the LFW general dataset, the AgeDB-30 general dataset, the CFP-FP general dataset, and the YTF general dataset.

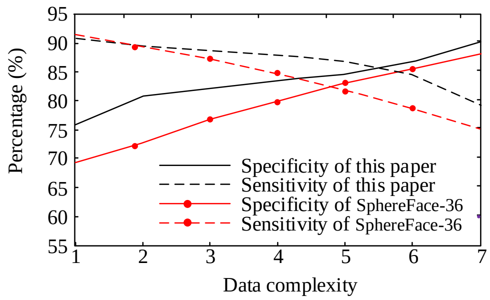


Fig. 7. Comparison results of sensitivity and specificity

Fig. 7 describes the comparison results of specificity and sensitivity between the basic SphereFace-36 and the optimized network model. In terms of specificity, the specificity of the optimized network model is higher than that of the basic SphereFace-36, and the difference can be up to 9 %. When the complexity of the data is low, the specificity of the optimized network model is significantly higher than that of the basic SphereFace-36; as the complexity of the data increases, the specificity gap between the basic SphereFace-36 and the optimized network model gradually narrows. In the sensitivity comparison of the methods, when the data complexity is low, the difference between the basic SphereFace-36 and the optimized network model is not significant, but as the data complexity increases, the optimized network model's sensitivity declines significantly less than the basic one. Fig. 8 shows the global feature recognition and extraction effects of different methods. Two sample images are chosen to optimize network model and the basic SphereFace-36 global feature's performance comparison.

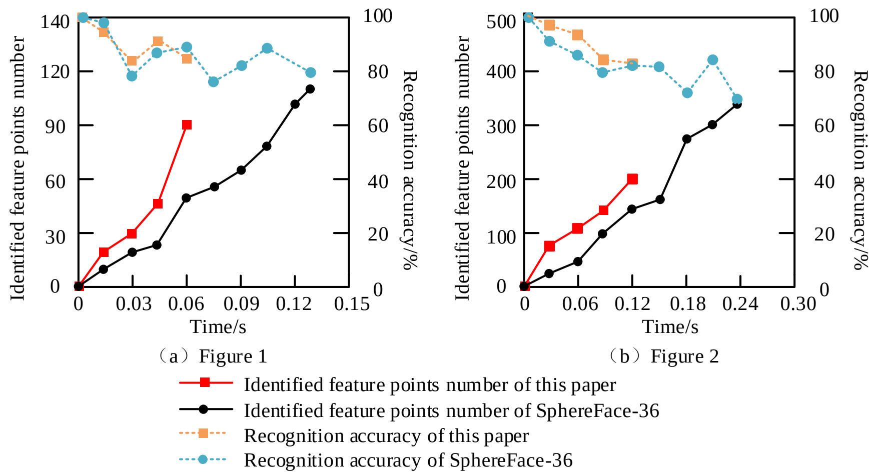


Fig. 8. Test results of global feature recognition and extraction

From Fig. 8, in the global feature extraction and recognition test of the two sample images, the optimized network model proposed in this experiment recognizes the number of feature points and the accuracy rate is higher than the basic SphereFace-36. Among them, the average accuracy rate of the optimized method is 83.01%, and basic SphereFace-36's average accuracy rate is 67.98%. The above results can show that the optimized network model has higher performance in global feature extraction and recognition of face images performance.

V. Conclusion

In the results of this research on face recognition technology, the area under the ROC curve for verification and recognition of the optimized model is the highest; The recognition accuracy of the optimized model in different data sets is 1.0%, 0.7%, 1.1%, 0.9% and 0.6% higher than that of SphereFace-36 respectively, and its specificity is higher than that of SphereFace-36, with the maximum difference of 9%. The average accuracy of global feature extraction and recognition of the optimized network model is 83.01%. The results show that the optimized method can accurately recognize facial features, and has high practical value, which can be applied to the automatic ticketing system. In a word, the improved method has good performance and plays a positive role in the establishment of face recognition system. Although the face recognition technology in this study has a good recognition effect, the running time of the algorithm is relatively long in the experimental process, which affects the running speed of the algorithm. In the later research, it is necessary to further improve the optimization of parameters. In addition, in actual application scenarios, face recognition may be affected by environmental factors, resulting in the inability to recognize low resolution faces. Therefore, in subsequent research, specific analysis of problems in the actual scene is required to improve the performance of face recognition technology.

References

[1] V. Sharma, A. M. Joshi, "VLSI Implementation of Reliable and Secure Face Recognition System," Wireless Personal Communications, vol. 122, no. 4, pp. 3485 - 3497, 2021.

[2] M. Kumar, S. Gupta, "2D - human face recognition using SIFT and SURF descriptors of face's feature regions." The Visual Computer, vol. 37, no 11, pp. 447 - 456, 2021.

[3] Y. Kortli, M. Jridi, A.A. Falou, M. Atri. "Face Recognition Systems: A Surve." Sensors (Basel, Switzerland), vol. 20, no. 2, pp. 1 - 34, 2020.

[4] F. Zhang, T. Zhang, Q. Mao, C. Xu. "A Unified Deep Model for Joint Facial Expression Recognition, Face Synthesis, and Face Alignment."

IEEE Transactions on Image Processing, vol. 29, no. 1, pp. 6574 - 6589, 2020.

[5] S. F. Baltanas, J. R. Ruizsarmiento, J. Gonzalezjimenez, "Improving the Head Pose Variation Problem in Face Recognition for Mobile Robots." Sensors, vol. 21, no. 2, pp. 659 - 676, 2021.

[6] M. Sadiq, D. Shi, J. Liang, "A robust occlusion - adaptive attention - based deep network for facial landmark detection." Applied Intelligence, vol. 52, no. 8, pp. 9320 - 9333, 2022.

[7] B. Chandra, S. Sastry, C.Jana, S. Robust, "Heartbeat Detection From Multimodal Data via CNN - Based Generalizable Information Fusion." IEEE Transactions on Biomedical Engineering, vol. 66, no. 3, pp. 710 - 717, 2019.

[8] H. Zhong, W. Yimu, Q. Xiaodong, Y. Rui, Q. Yong, H. Kelei, Z. Zezhang. "Classification of neurofibromatosis-related dystrophic or nondystrophic scoliosis based on image features using Bilateral CNN." Medical Physics, vol. 48, no. 4, pp. 1571 - 1583, 2021.

[9] O. Elharrouss, N. Almaadeed, S. Al - Maadeed, F. Khelifi. "Pose - invariant face recognition with multitask cascade networks." Neural Computing and Applications, vol. 34, no. 8, pp. 6039 - 6052, 2022.

[10] G. Lou, H. Shi, "Face image recognition based on convolutional neural network." China Communications, vol. 17, no. 2, pp. 117 - 124, 2020.

[11] R. Silwal, A. Alsadoon, P. Prasad, O.H. Alsadoon, A. Al-Qaraghuli. "A novel deep learning system for facial feature extraction by fusing CNN and MB-LBP and using enhanced loss function." Multimedia Tools and Applications, vol. 79. no. 41/42, pp. 31027-31047, 2020.

[12] M. M. Hussein, A. H. Mutlag, H. Shareef, "An Improved Artificial Neural Network Design for Face Recognition utilizing Harmony Search Algorithm." IOP Conference Series Materials Science and Engineering, vol. 745, pp. 012038, 2020.

[13] J. O. Agushaka, A. E. Ezugwu, L. Abualigah, "Dwarf Mongoose Optimization Algorithm." Computer Methods in Applied Mechanics and Engineering, vol. 391. no. 1, pp. 1-38, 2022.

[14] L. Abualigah, A. Diabat, S. Mirjalili, et al. "The Arithmetic Optimization Algorithm." Computer Methods in Applied Mechanics and Engineering, vol. 376. no.1, pp. 1-38, 2021.

[15] K. H. Almotairi, L. Abualigah, "Improved reptile search algorithm with novel mean transition mechanism for constrained industrial engineering problems." Neural Computing and Applications, vol. 34. no.20, pp.17257-17277, 2022.

[16] A.S. Bist, W. Febriani, C. Lukita, S. Kosasi, U. "Rahardja. Design of Face Recognition AttendX for Recording Student Attendance Data Based on Artificial Intelligence Technology." Solid State Technology, vol. 63, no. 2s, pp. 4505 - 4518, 2020.

[17] Y.H. Pu, P.S. Chiu, Y.S. Tsai, M.T. Liu, Y.Z. Hsieh, S.S. Lin. "Aerial face recognition and absolute distance estimation using drone and deep learning." The Journal of Supercomputing, vol. 78, no. 4, pp. 5285 - 5305, 2022.

[18] B. Rossion, A. Lochy, "Is human face recognition lateralized to the right hemisphere due to neural competition with left - lateralized visual word recognition?" A critical review. Brain Structure and Function, vol. 227, no. 2, pp. 599 - 629, 2021.

[19] S. Nejrs, M. S. Al – An, "Face Image Classification Based on Feature Extraction. Solid State Technology," vol. 63. no. 6, pp. 13515 – 13526, 2020.

[20] M.R.M. Alsemawi, M.H. Mutar, E.H. Ahmed, H.O. Hanoosh, F.T. Al - Dhief. "Emotions Recognition from Human Facial Images Using Machine Learning Technique." Solid State Technology, vol. 63, no. 5, pp. 8749 - 8761, 2020.

[21] H. B. Fredj, S. Bouguezzi, C. Souani, "Face recognition in unconstrained environment with CNN." The Visual Computer, vol. 37, no. 2, pp. 217 - 226, 2021.

[22] Z. Sun, L. Tian, Q. Du, et al. "Feature learning via multi - action forms supervising force for face recognition." Neural Computing and Applications, vol. 34, no. 6, pp. 4425 - 4436, 2022.

[23] C. Chen, X. Zhou, "Collaborative representation - based fuzzy discriminant analysis for Face recognition." The Visual Computer, vol. 38, no. 4, pp. 1383 - 1393, 2022.

[24] J. Gan, K. Jiang, H. Tan, G. He. "Facial Beauty Prediction Based on Lighted Deep Convolution Neural Network with Feature Extraction Strengthened." Chinese Journal of Electronics, vol. 29, no. 2, pp. 312 - 321, 2020.

[25] C. Han, S. Shan, M. Kan, S. Wu, X. Chen. "Personalized Convolution for Face Recognition." International Journal of Computer Vision, vol. 130, no. 2, pp. 344 - 362, 2022.

[26] Rizwan S A, Ghadi Y Y, Jalal A, Kim K. "Automated Facial Expression Recognition and Age Estimation Using Deep Learning." Computers, Materials, & Continua, vol. 6, pp. 5235-5252, 2022.

[27] Chen L, Li M, Lai X, Hirota K Pedrycz W. "CNN-based Broad Learning with Efficient Incremental Reconstruction Model for Facial Emotion Recognition." IFAC-PapersOnLine, vol. 53. no.2, pp.10236 - 10241, 2020.

[28] Mao L, Yan Y, Xue J H, Wang H. "Deep Multi-Task Multi-Label CNN for Effective Facial Attribute Classification." IEEE transactions on affective computing, 2022, vol. 13. no.2, pp. 818-828.