2021 20th International Symposium on Distributed Computing and Applications for Business Engineering and Science (DCABES)

Application of face recognition in port unrestricted scene

Song Qi
School of Computer Science and Technology,
Wuhan University of Technology
Wuhan, China
e-mail: 1192913039@qq.com

HanBing Yao
Wuhan University of Technology
Chongqing Research Institute
School of Computer Science and Technology,
Wuhan University of Technology
Wuhan, China
e-mail: 22876681@qq.com

WuWenBo
Imformation Technology Department,
Jiangjin Port Branch
Chongqing Gangjiu Co., LTD.
ChongQi China
e-mail:441479417@qq.com

***Abstract*—Traditional face recognition technology needs to collect data in a specified environment with stable facial features.However,due to the large scope and wide area of port operation,traditional identification methods have been unable to meet the requirements of port unrestricted scenarios. In this paper,the algorithms based on RetinaFace face detection and ArcFace face recognition are applied to port staff attendance in unrestricted scenarios.In order to obtain a more stable and accurate face representation method,this paper studies face detection, face image pretreatment,face recognition,loss function and other aspects,and verifies the feasibility of this model in port unrestricted scene attendance through the experimental results on PubFig dataset and the analysis of real-time face recognition effect.**

***Index Terms*—data sets;Face recognition;ArcFace;Loss functions;**

I. Introduction

Nowadays, with the trend of world economic globalization,the operation of port enterprises is also developing towards a more professional and information mode.A good attendance system not only enables enterprise leaders to better manage employees,but also enables employees to better restrain themselves,put more time and energy into work, and bring better development for the enterprise.Face recognition is a research hotspot in computer vision field.The traditional face recognition methods are only suitable for some specific scenes,such as the access board of public facilities.In general, only when the facial features reach a stable state, the traditional face recognition technology can accurately detect,and also require the detection of the face recognition equipment,the detection of active cooperation,such as removing masks,glasses and other objects that block the face recognition.The attendance method based on face recognition eliminates the disadvantages of traditional attendance,greatly facilitates personnel management,and makes attendance more reliable and efficient. Face recognition is usually composed of three key elements:face detection, face pretreatment and face recognition.The system designed in this paper uses face detection algorithm RetinaFace and face recognition part uses an additive angular edge loss (ArcFace) algorithm.However,the combination of Multi-task Cascaded Convolutional Neural Networks(MTCNN)face detection algorithm and Facenet face recognition algorithm had been better in the actual test environment.

II. FACE RECOGNITION

According to the different forms of data, face recognition can be divided into methods based on 2D images and methods based on 3D scanning,the two methods are very different in the development and application to a certain extent.In addition, with the rapid development of deep convolutional neural networks (DCNNs),deep learning-based methods have achieved significant performance improvements in including face recognition.In general, the depth face recognition system includes three basic elements:face detection, face pretreatment and face recognition.First,face detection locates the face region on the input image.Then face preprocessing is carried out to normalize the detected face into a standard layout. Finally,the face recognition algorithm will extract the recognition features from the preprocessed face.These features were used to calculate how similar they were to each other to determine whether faces belonged to the same identity.

A. Face detection

Face detection is the first step of face recognition system.Its main task is to lock the face region in the picture,for example, given an input picture,the task of face detection is to find all the faces in the picture, and give the boundary box coordinates with confidence.Face detection technology has made great progress in recent years,but there are still some difficulties and challenges,such as scale,resolution,occlusion,illumination and so on.Traditional face recognition technology designed manual features are mainly used to distinguish between face and background area.The method can be divided into single-stage and multi-stage methods by stage learning, and the anchored based and non-anchored methods can be divided according to the usage of anchor.However, a face detection method can be both single-phase and anchor-based.For the multistage approach,the multistage detector first generates many candidate boxes and

2473-3636/21/$31.00 ©2021 IEEE
DOI 10.1109/DCABES52998.2021.00029

then refines them in a coarse-to-fine or coarse-to-refinement strategy.

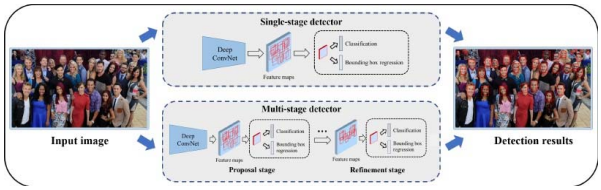

Fig. 1. Schematic diagram of single and multistage face detector

As shown in Figure 1,the single-stage detector directly completes face detection from the whole feature graph,while the multi-stage detector uses multi-stage to generate candidate images and refine candidate images through one or more stages. A classic single-stage detector,SSD, is the mainstream framework for target detection because it runs faster than the faster R-CNN while maintaining considerable accuracy. Many developers are using SSD face detection in the system. Although single-stage method has the advantage of high efficiency, its detection accuracy is lower than multi-stage method. This is partly due to the negative and positive imbalance caused by dense anchors.Due to the long development time and superior performance, most face detectors are based on anchor points. Dense anchor points are usually preset on the feature graph, and then these anchor points are classified and boundary box regressions are performed one or more times.

B. Face preprocessing

In the second stage,the purpose of face preprocessing is to calibrate the detected face to a typical view (mainly to improve the end-to-end performance of face recognition is an important step.Because the structure of the face is regular, and face parts (eyes, nose, mouth, etc.) of the arrangement is constant,so the arrangement of the face of the subsequent face recognition feature calculation has great benefits.In general, face alignment uses spatial transformation techniques to calibrate faces into a standardized layout.In the existing face alignment methods,face landmarks or so-called face keypoints are essential because they serve as a reference for similar transformation or affine transformation.Therefore, face landmark localization is the premise of face alignment.Face landmark localization methods based on deep convolutional neural network can be divided into three categories: coordinate regression based method,heat map regression based method and 3d model fitting method.RetinaFace puts the face image on the ordinate of the landmark based on the return target.The study of coordinate regression mainly adopts L1, L2 or smooth L1 loss function.

C. RetinaFace

RerinaFace is a single-stage target detection algorithm specialized in face detection.Based on multi-stage,Feature Pyramid Networks (FPN) are used to make full use of Feature information.The network structure is shown in Figure 2.RetinaFace uses multilevel feature images to form a feature pyramid,and a large number of anchor boxes with different sizes are designed on the multilevel feature images to achieve excellent detection performance.

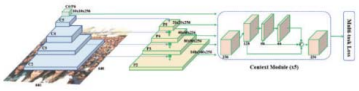

Fig. 2. RetinaFace algorithm network structure

D. Face recognition

After the face preprocessing,in the face recognition stage,the goal is to map the aligned face image to a feature space,where the features of the same identity are close,and the features of different identities are far apart.In practical application,face recognition mainly has two tasks:face verification and face recognition.But whether it is face verification or face recognition,are used to calculate the similarity of face image features,how to learn different face representation is the core goal of face recognition system.[1]Network architecture and training supervision play a key role in face recognition learning.

The supervised training methods of face recognition are divided into supervised mode,semi-supervised mode and unsupervised mode. In the supervision mode,it can be divided into three categories:classification,feature embedding and hybrid method. Deep face recognition learning based on classification method comes from general object classification task.Each class corresponds to an identity that contains multiple faces of the same person.Softmax training loss is the most widely used supervision method in classification tasks, which consists of full connection layer, Softmax and cross entropy loss.

E. ArcFace algorithm

SoftMax is often used in the classification process of various categories in deep learning.It allows the output values of multiple neurons to be mapped between 0 and 1 and their sum is 1.With similar probability,we take the node with the largest output value as the target category for prediction.Softmax cross entropy loss function:it is often used in the classification process of various categories in the past.[3]It allows the output values of multiple neurons to be mapped between 0 and 1 and their sum is 1.With similar probability,we take the node with the maximum output value as the target category for prediction.Softmax cross entropy loss function formula is shown in Formula 1:

$$L_1 = -\frac{1}{N}\sum_{i=1}^{N}\log\frac{e^{w_{yi}^T xi+by_i}}{\sum_{j=1}^{n} e^{w_j^T xi+bj}} \tag{1}$$

Xi is the sample feature,Yi is the category to which the sample belongs,Wj is the jth column of the weight matrix,bj is the paranoid amount,N is the batch processing size,and N is the number of categories.The unimproved SoftMax loss function can only classify them by output value,and cannot closely combine the same category through optimization,nor can it

distinguish different categories greatly,which is not suitable for face recognition training.ArcFace moves the additive Angle margin M to the interior of cos,and the loss function formula is shown in Formula 2:

$$L4 = -\frac{1}{N}\sum_{i=1}^{N}\log\frac{e^{s(\cos(\theta yi+m))}}{e^{s(\cos(\theta yi+m))}+\sum_{j\neq 1,j\neq yi}^{n}e^{s\cos\theta yi}} \quad (2)$$

Compared with other loss functions,ArcFace can better interpret Angle distance in geometric space,and it can correspond to geodetic lines in hyperplane and maximize decision boundary more intuitively.

III. EXPERIMENTATION

A. Experimental data set

The experimental Dataset PubFig is the Columbia University Open People Dataset.It is a large face dataset that covers 58,797 images of 200 people on the Internet.Unlike most existing face datasets,these images are taken with the subject completely out of control.Therefore,there are great differences in posture,illumination, expression,scene, camera,imaging conditions and parameters in different images. Through data set provides web crawl the data link,which the web page is inaccessible data cleaning,do not use data,not the camera captured data and light discomfort,mask,the image is not clear,such as data, lastly,everyone a few different positions (as shown in figure 3),facial expression, illumination and so on,Four images per person were used as validation sets,which contained 1 200 positive sample pairs and 398 000 negative sample pairs.Before training,face detection and clipping of the dataset with RetinaFace,image data preprocessing clipping size.See Figure 3

Fig. 3. Partial data set

The data set training in this paper is based on Ubantu20.04, Pytorch deep learning framework, and trained on two GeForce GTX 1080Ti GPU. The skeleton structure of the network is ResNet50 in deep convolutional neural Network. Arcface is used for the loss function, and stochastic gradient descent (SGD) is adopted. In order to make the network converge faster, cosine annealing attenuation method is used in the training. The initial learning rate is set to 0.1, and the maximum number of iterations is set to 100. After 100 batch training, the network reaches convergence state.Due to the small data set, the convergence speed of the loss function is fast, and the accuracy of the final retraining set and verification set is high. The final experiment also achieved the expected effect.

As can be seen from Table 1,after 10,000 batch training of the algorithm,ArcFace has the highest accuracy,SphereFace has the lowest accuracy than ArcFace,SoftMax performs poorly compared with other loss functions,CosFace and ArcFace have little difference.It can be seen that the performance of ArcFace loss function has a good performance in PubFig.The convergence of each loss function is good during the model training.Due to the small size of the data set and the high accuracy of training and verification,the loss function converges rapidly and achieves good results.It can be seen that the experiment on the PubFig public data set has achieved the expected goal.

TABLE I
ACCURACY OF DIFFERENT LOSS FUNCTIONS

loss function	Highest accuracy (percent)
CosFace	82.62
SphereFace	81.24
SoftMax	70.92
ArcFace	87.00

IV. PORT STAFF ATTENDANCE FACE RECOGNITION APPLICATION

A. Application Background

The common way of attendance is radio frequency card attendance and fingerprint attendance.These two methods of attendance are also widely used, but both card attendance and fingerprint attendance face the problem of whether the puncher is himself or not.Radio frequency card attendance can be handed to other people to clock in,while the fingerprint attendance mode seems to be unable to clock in,but now fingerprint film attendance technology has been very common,can not avoid the phenomenon of clock in. With the advent of the facial recognition system,the phenomenon of punching in has been solved,and it is not possible to use photos to punch in.As the facial recognition technology becomes more and more mature,the accuracy of facial recognition is gradually improved to achieve fast facial recognition.It has been a trend for companies to use face recognition attendance systems.

The working environment of ports is large,and the working time and location of employees are not uniform,which is quite different from that of ordinary enterprises.As a result,most port enterprises still use the most traditional paper and handwritten attendance method. On the one hand, port employees, this special group,are engaged in port operations, and a large number of port employees have incomplete fingerprint information,resulting in a large number of port enterprises in the use of fingerprint attendance is powerless; On the other hand,most domestic researches on ports focus on port intellectualization and informatization,and less attention is paid to port enterprise staff attendance management.Therefore, the mobile attendance mode of face recognition is also one of the attendance modes relatively suitable for this group.

B. The overall design of port face recognition system

The specific implementation steps of the system are as follows:

1) Login module : In order to ensure the security of employee information, the system has a login module, only authorized administrators can log in operations. This module is connected to the MySQL database. The administrator must enter the correct user name and password to log in successfully. After login successfully, the administrator can view the basic personal information and attendance information of employees in the system. After entering the correct user name and password, the administrator logs in to the main interface of the system. The interface mainly includes information collection module, face training module, face attendance module and attendance information query module.

2) Information collection module: The information collection module mainly completes the basic personal information and face information input of employees, including the basic personal information of the employee's name, number, phone number, gender, department, ID number and so on. Face information input includes two ways, one way is to call the camera to shoot 4 face photos, if the photo quality is qualified, saved in JPG format under the specified folder. If the photos do not meet the requirements, the staff will be prompted to take photos again; Another way is that employees can take good photos by themselves and send them to the administrator's email. The administrator can upload the photos directly on the computer.

3) Image preprocessing module: The image preprocessing module mainly processes the collected face images,because the port working environment is special,the face may be contaminated with dust and other impurities,causing a certain degree of influence on face recognition.In order to improve the accuracy of face recognition,the face is preprocessed before face training to reduce image noise pollution and enhance image features.And the processed photos were rotated, translated,mirrored and other operations to prepare for face training.

4) Face training module: After the pretreatment of face image,face training is needed.The system uses face detection algorithm RetinaFace and face recognition part uses an additive angular edge loss (ArcFace) algorithm.However, the combination of Multi-task Cascaded Convolutional Neural Networks(MTCNN) face detection algorithm and Facenet face recognition algorithm is better in the actual test environment.

5) Face attendance module: The face attendance module mainly completes the real-time attendance of employees. When the camera captures the face, it will collect the captured face and extract the corresponding face features, and then compare the collected face features with the face feature template in the database. If it is the face in the database, the second step is to judge the background environment taken by the staff when they punch in. If there are special characteristics of the port environment, the staff will be prompted to sign in successfully according to the location provided by the positioning system, and the sign in information will be written into the database. If it does not exist, prompt the staff to take a new picture and submit it; If the face information in the database cannot be matched, the employee will be reminded that the face cannot be matched.

6) Attendance information query module: The attendance information query module mainly completes the attendance information query of the employees.After the successful attendance,the attendance information of the employees is recorded,which mainly includes the employee's work number,name,check-in time,check-out time and attendance status.

SUMMARIZE

In this paper,the algorithms based on RetinaFace face detection and ArcFace face recognition are applied to port staff attendance in unrestricted scenarios.In order to obtain a more stable and accurate face representation method,this paper studies face detection,face image pretreatment, face recognition,loss function and other aspects.Through verifying ArcFace loss function has a good performance on PubFig data set.Using the camera to obtain face data information and face database can accurately identify whether the person is in the database.The feasibility of the model in port unrestricted scenarios is verified.

In the future,problems related to employee punch location,recognition rate,processing capacity and hardware equipment need to be further explored and realized.

ACKNOWLEDEGMENTS

The work described in this paper was supported by the Research Project of Wuhan University of Technology Chongqing Research Institute (YF2021-10).

REFERENCES

[1] LIU WY,WEN YD,YU ZD,et al.SphereFace:Deep Hypersphere Embedding for Face Recognition [C]//2017 IEEEConference on Computer Vision and Pattern Recognition(CVPR)IEEE,2017:6738-6746.

[2] WANG H,WANG YT,ZHOU Z,et al.Cosface:large margin cosine loss for deep facerecognition [C]//2018 IEEE CVF Conference on Computer Vision and Pattern Recognition (CVPR)IEEE,018:5265-5274.

[3] DENG JK,GUO J,XUE NN,et al.ArcFace:additive angular margin loss for deep face recognition [C]//2019 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)IEEE,2019:4690-4669

[4] Wang Mei,Zhang Yaobin,Deng Weihong. Meta Balanced Network for Fair Face Recognition.[J]. IEEE transactions on pattern analysis and machine intelligence,2021,PP:

[5] Qiu Haibo,Gong Dihong,Li Zhifeng,Liu Wei,Tao Dacheng. End2End Occluded Face Recognition by Masking Corrupted Features.[J]. IEEE transactions on pattern analysis and machine intelligence,2021,PP:

[6] Deng Jiankang,Guo Jia,Yang Jing,Xue Niannan,Cotsia Irene,Zafeiriou Stefanos P. ArcFace: Additive Angular Margin Loss for Deep Face Recognition.[J]. IEEE transactions on pattern analysis and machine intelligence,2021,PP:

[7] Liu Decheng,Gao Xinbo,Peng Chunlei,Wang Nannan,Li Jie. Heterogeneous Face Interpretable Disentangled Representation for Joint Face Recognition and Synthesis.[J]. IEEE transactions on neural networks and learning systems,2021,PP:

[8] Ming Xiang,Wei Fangyun,Zhang Ting,Chen Dong,Zheng Nanning,Wen Fang. Group Sampling for Scale Invariant Face Detection.[J]. IEEE transactions on pattern analysis and machine intelligence,2020,PP: