Research on Algorithms of Lateral Face Recognition Based on Data Generation

Zimin Zhang[1], Zhaohui Zhang[1,2(✉)], Xiaoyan Zhao[1,2], and Tianyao Zhang[1,2]

[1] School of Automation and Electrical Engineering, University of Science and Technology Beijing, 30# Xueyuan Road, Haidian District, Beijing 100083, China
zhangzhaohui@ustb.edu.cn

[2] Beijing Engineering Research Center of Industrial Spectrum Imaging, University of Science and Technology Beijing, 30# Xueyuan Road, Haidian District, Beijing 100083, China

Abstract. With the rapid progress of artificial intelligence, the methods of face recognition have also achieved considerable progress. However, when multiple head poses are present, the accuracy of face recognition decreases due to angular deviation. Therefore, how to realize the enhancement of accurate precision in two-dimensional multiple head poses is still a worthy research topic. In this paper, based on the ResNet50 residual network structure, data enhancement and attention mechanism are adopted, and the Public Figures Face Database public face database is used to partition the training and test sets. This hybrid approach effectively improved the accuracy of face recognition in lateral. However, the hybrid method has low accuracy for large angle deflection face recognition. Therefore, based on the hybrid method, this article proposes a side face recognition algorithm based on data generation, incorporating a data generation algorithm. The recognition accuracy of the front face of this algorithm is as high as 96.3%. This article presents a data-generated lateral face recognition algorithm that further improves the accuracy of large angle deflection face recognition.

Keywords: Face Recognition · Data Generation · Attention Mechanism

1 Introduction

With the rapid progress of science and technology, Chinese society is becoming more and more informatized, modernized and artificially intelligent. The drawbacks and shortcomings of the traditional identity authentication methods are gradually highlighted, and people are more and more hopeful that there will be new technologies to break the situation. Biometrics technology, which relies on computer vision technology to collect all kinds of characteristic information of living beings, has gradually become a key direction for experts and scholars to study and overcome by virtue of its own excellent characteristics, such as uniqueness and high efficiency, which are different from those of traditional identity identification. Compared with various biometric features extracted from the human body such as fingerprints, genes, voice, iris [1], etc., the human face has also become more and more a research hotspot in the field of deep learning and artificial

B. Xin et al. (Eds.): IWACIII 2023, CCIS 1932, pp. 182–194, 2024.
https://doi.org/10.1007/978-981-99-7593-8_17

intelligence due to its advantages such as easy to collect, high uniqueness, the ability to dynamically analyze the expression and gesture, broad application prospects, and the lack of easy to forge.

Face recognition technology originated in the 1950s for extracting and recognizing the contours of human faces, and then it has made rapid development as a research hotspot in the past decades. The following four methods are commonly used for face recognition: geometric feature-based methods, subspace analysis-based methods, local feature-based methods, and deep learning-based methods. Gu's team [2] analyzed the advantages, disadvantages, and typical cases of the above four methods. Li [3] compared and analyzed the current mainstream face recognition algorithms, and introduced in detail the two classification methods of face recognition technology, 2D and 3D face recognition process.

Face recognition technology is becoming more and more mature, and face recognition systems are appearing more and more in people's lives. However, under unconstrained natural conditions, face recognition technology is very susceptible to factors such as light intensity and posture changes, which in turn leads to a sharp decline in face recognition accuracy. When a person's head posture changes, his or her own facial features will be obscured by himself or herself, resulting in some key features of the face cannot be recognized. When the same person with different postures of the face image for comparison, it will be found that the difference between the images will change with the degree of change in posture, the greater the change in facial features, the greater the difference. Although current face recognition algorithms have achieved good recognition results, there is still room for improving the research on algorithms for lateral face images.

Currently researched algorithms for lateral face recognition can be categorized into two groups: First, feature extraction-based methods; Second, face rotation-based methods. Zhuang [4] categorized and summarized the face recognition methods based on geometric and algebraic features for people under natural conditions. The feature extraction-based methods actually extract feature information that is robust to the pose of the face. However, it is difficult to obtain a feature representation that is robust to pose changes through the feature extraction methods because of the unbalanced distribution of the number of large-pose face images. Zheng [5] provides a detailed introduction to the principles and implementation steps of two algorithms, LBP and Gabor, for extracting local features of faces. The face rotation-based methods are those that rotate the side image of the face into the frontal image of the face. Most of these methods are done using deep learning techniques. The methods can be further subdivided into two categories: methods based on convolutional neural networks and methods based on generative adversarial networks [6]. Generative adversarial networks with extremely strong generative capabilities have continuously achieved bright results in tasks such as face frontalization [7] and face repair [8] in recent years.

2 Lateral Face Recognition Algorithm Based on Data Generation

This section describes the lateral face recognition algorithm based on data generation, which incorporates Resnet50, attention mechanism and data augmentation in the data generation approach.

The data is generated by generating a 2D lateral face image from a 2D front face image. First, the parameter and texture information of the 3D image is obtained from the 2D image by the 3D Morphable Model method. Then, the 3D representation is rotated and rendered by using Neural Mesh Renderer which is an open source. After that the rendered image is mapped to the real image domain using Cycle GAN generator, Pix2PixHD multilayer discriminator and perceptual loss to generate the final multi-angle 2D image of the lateral face [9], the network structure is shown in Fig. 1.

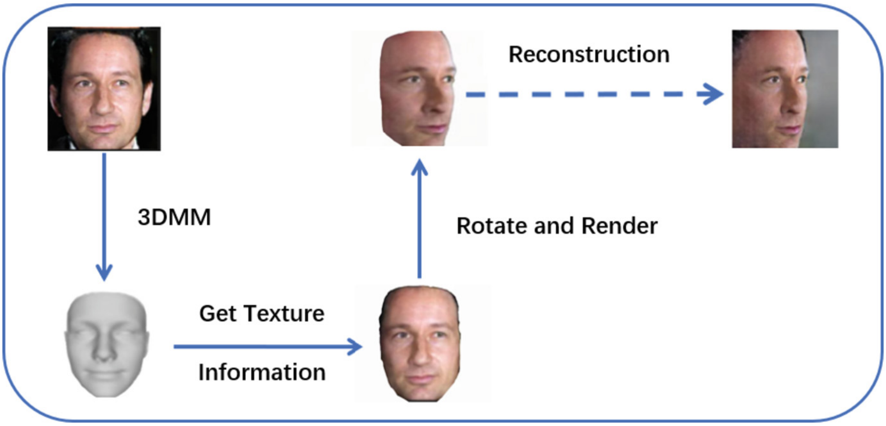

Fig. 1. Schematic diagram of the data generation methodology.

ResNet50 denoted as Base in Chapter 3 is an improved residual neural network that contains two basic blocks named Conv Block and Identity Block. One is the Conv Block: The dimensions of channels count and size for input and transformation processes are not the same, so it cannot be connected in a continuous series, which serves to change the dimensionality of the network. The other is Identity Block: The dimensions of channels count and size for input and transformation processes are the same, so it can be connected in series and used to deepen the network. The specific understanding can be seen in the following schematic diagram of the Resnet50 neural network, which is shown in Fig. 2.

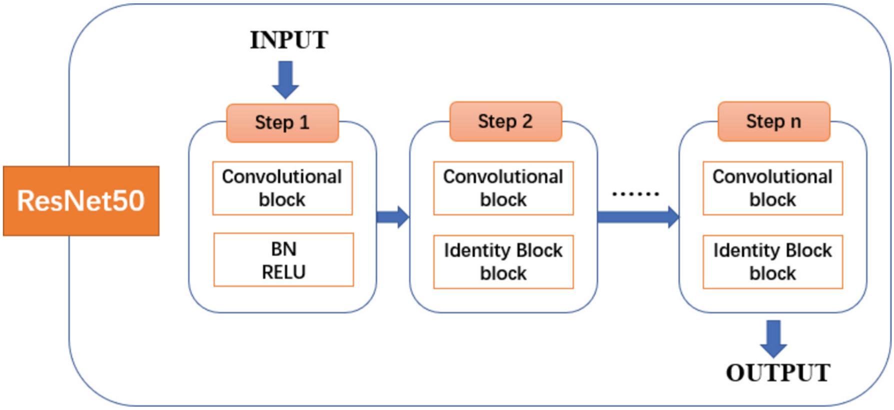

Fig. 2. Schematic diagram of Resnet50 neural network.

SE Net [10] is a method to add an attention mechanism to the channel dimension of a neural network, and its main operations are Squeeze and Excitation. A $1 * 1 * C$ weight matrix is obtained for the reconstruction of the original features by means of automatic learning. The different values after the reconstruction are used to measure the different levels of importance of the channels. This importance is then used to assign a weight value to each feature so that the neural network focuses on certain feature channels. The SE module is a plug-and-play module, which is added to the ResNet50 network structure in this paper which referred to as Attention in Chapter 3, and the algorithm structure is shown in Fig. 3. Structurally, it does not change the basic structure of the detection network, and at the same time, it does not increase the complexity, which is conducive to improving the convergence speed of model training.

Data augmentation [11] which often referred to as AUG in Chapter 3 is the process of processing more representations out of raw data without materially increasing the data. The purpose of the method is to increase the quantity and quality of the raw data so as to approach the value generated by a greater quantity of data. Data enhancement is often used with the IMGAUG toolkit, which is a more powerful data enhancement toolkit than TORCHVISION [12]. This toolkit is a packaged python library for image enhancement.

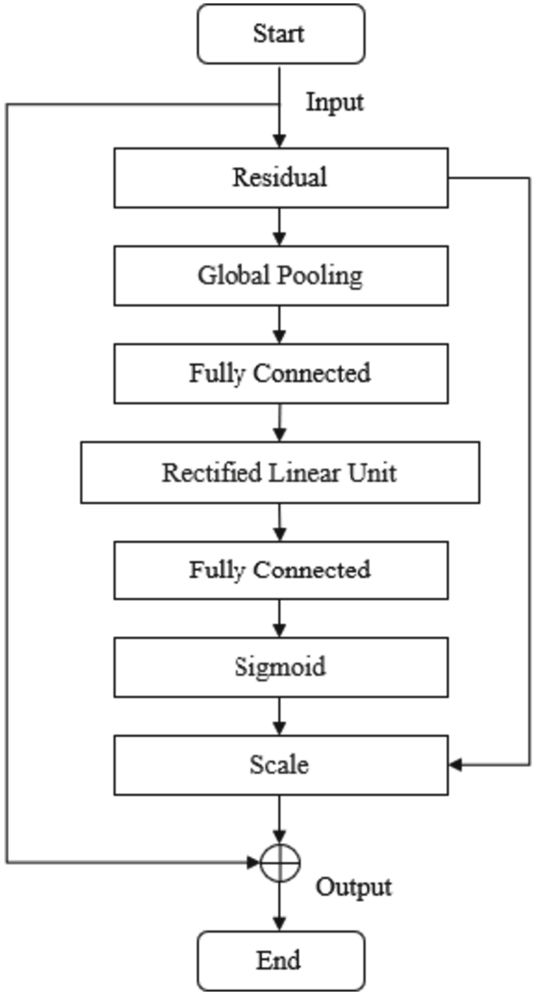

Fig. 3. Schematic of the ResNet50 network structure with the addition of the attention mechanism.

3 Experiments and Analysis

3.1 Environment Setup and Partial Parameterization

Since a large amount of image data processing is involved in this chapter and deep learning frameworks are required to be used for computation, there are high requirements on the hardware environment and software configuration of the computer. In terms of hardware, this section needs to use a highly configured computer with a large operating memory and a GPU graphics card for accelerated training so as to save time and cost, so the GPU rtx2080 is used as the graphics card in this section, and the CPU is chosen to be a core i7 processor. In terms of software, combined with the powerful compatibility of the Windows system and other advantages, choose Windows 10 as the main research and development platform of the operating system, this paper is based on python to build the system framework, using VS code as the project manager and editor, the training process using Adam optimizer to achieve the generation of adversarial network this deep learning model building, the learning rate is 1e-4, weight initialization using kaiming initialization method.

The data enhancement methods used in this chapter are Some Of Subitem Enhancer, Fliplr Horizontal Enhancer, Flipud Vertical Enhancer, Rotate Enhancer, Crop And Pad Enhancer, Crop Enhancer, Multiply Brightness Pixel Enhancer, Sharpen Enhancer, Contrast Normalization Enhancer, Gaussian Blur Enhancer, Motion Blur Enhancer, Gamma Contrast Enhancer. This section applies one to three of the above given enhancers, and the configuration of each enhancer is shown in Table 1 below.

Table 1. Enhancer configurations.

Enhancer	Value
Fliplr	0.5
Flipud	0.2
Rotate	(-45, 45)
Crop And Pad	Percent = (-0.25, 0.25)
Crop	Percent = (0, 0.1)
Multiply Brightness	(0.5, 1.5)
Sharpen	Alpha = 0 Lightness = 1 Name = None Deterministic = False Random _ State = None
Contrast Normalization	Alpha = 1.0 Per _ channel = False name = None Deterministic = False Random _ state = None
Gaussian Blur	Sigma = (0.0, 0.5)
Motion Blur	K = 3
Gamma Contrast	(0.5, 2.0)

3.2 Face Database Selection

Common face databases include CMU Multi-PIE [13], LEW [14], VGGFACE [15], CASIA-WEBFACE [16], and MS-Celeb-1M [17], etc., which can be broadly categorized into 2D face database and 3D face database.

This project was using Public Figures Face Database, as shown in Fig. 4. Public Figures Face Database is a large dataset of real faces, consisting of 797200 2D images of 60 public figures collected from the Internet. Unlike most other existing face datasets, these images were taken under completely uncontrolled conditions with uncooperative subjects [18].

As a result, there is a great deal of variation in pose, lighting, expression, scene, camera, imaging conditions and parameters. In this section, the database images were processed and the images of 60 people were screened and organized to generate the training and test sets in the ratio of 3:1 for the number of training and test sets, and it was ensured that both the training and test sets contained both front and side face images.

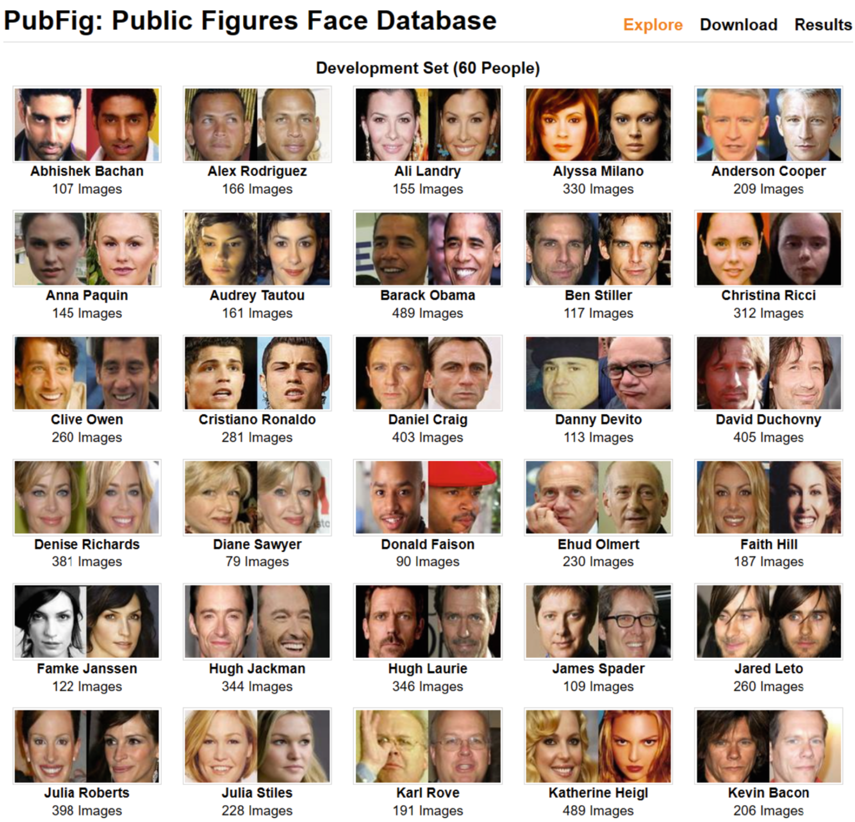

Fig. 4. Schematic of public face database [19].

3.3 Comparison of Face Recognition Methods and Analysis of Results

Experiment 1: Comparison of Recognition Accuracy of Four Models. The public face database is processed according to the proportion to complete the training set and test set classification, after which the face recognition results of the four models Base, Base + AUG, Base + Attention, Base + Attention + AUG are compared respectively.

The result curves are obtained as shown in the following Figs. 5, 6 and 7, and the table of collated key data is shown in Table 2.

Table 2. Comparison results of the accuracy of the four models.

Module	Accuracy	Recall
Base	0.528	0.721
Base + Aug	0.681	0.830
Base + Attention	0.723	0.91
Base + Attention + Aug	0.87	0.91

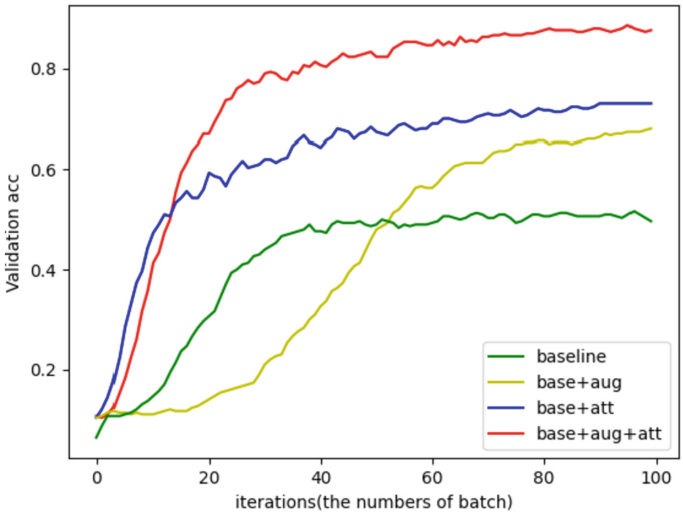

Fig. 5. Comparison of the accuracy of the four models.

According to the curve and the data in the table, it can be seen that after adding the attention mechanism or adding the data enhancement, the recognition accuracy is improved, and the attention mechanism has a relatively high accuracy due to its own characteristics, and the two hybrid methods reach 68.1% and 72.3% respectively. The hybrid method after adding the attention mechanism and data enhancement is better overall, with an accuracy rate of 87% and a recall rate of 91%, which is 34.2% and 18.9% higher than using only the most basic resnet50 network structure, respectively.

On this basis, the yaw angle in the face pose angle is classified, and seven angles of 0°, ± 30°, ± 45°, ± 60° are selected, and the training and test sets are reorganized, and the ratio of the number of images in the training set and the test set is still 3:1, and it is ensured that the proportion of the front face in the test set is 51%, and the proportion of the various angles in the side face is as follows: [0°-30°] accounts for 30%, [30°-45°]

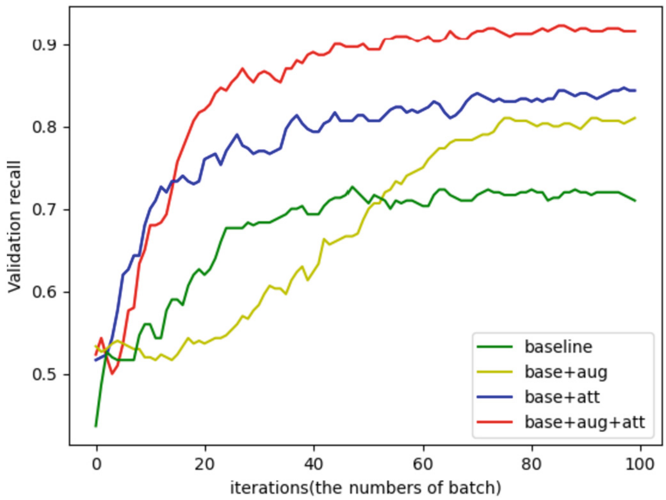

Fig. 6. Comparison of the recall of the four models.

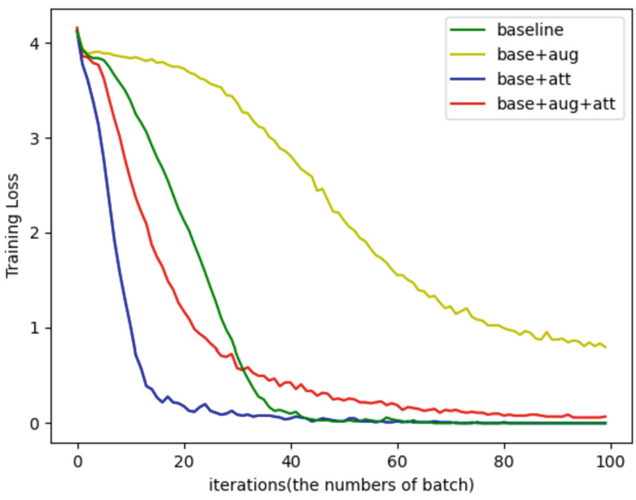

Fig. 7. Comparison of losses for the four models.

accounts for 15%, [45°- 60°] accounts for 4%. Compare the results of face recognition accuracy of four models, Base, Base + AUG, Base + Attention, Base + Attention + AUG, respectively, and organize the key data table in Table 3.

Further comparative analysis of the data results in the above table can be analyzed to obtain the following two points.

1) Side-by-side comparison: After adding the attention mechanism and data enhancement method, the accuracy rate of face recognition at each angle in the test set is

Table 3. Comparison of recognition accuracy results of four models at various angles.

Angle	Base	Base + Aug	Base + Att	Base + Att + Aug
[-60°, -45°]	0	0.1	0.23	0.35
[-45°, -30°]	0.312	0.462	0.47	0.5
[-30°, 0°]	0.34	0.621	0.764	0.892
0°	0.72	0.786	0.802	0.956
[0°, 30°]	0.35	0.65	0.769	0.87
[30°, 45°]	0.43	0.71	0.69	0.782
[45°, 60°]	0	0.13	0.21	0.48
Total accuracy	0.5264	0.6814	0.7348	0.8647

improved. Among them, the hybrid method that adds both data enhancement and attention mechanism had the highest accuracy rate at each angle and total accuracy rate, with the accuracy rate at 0° angle as high as 95.6%, which is 23.6% higher than the accuracy rate of the basic method, and the total accuracy rate reaches 86.5%, which is 33.83% higher than the accuracy rate of the basic method.

2) Vertical comparison: The recognition accuracy of all four methods decreases with the increase of face deflection angle. In large-angle deflection face recognition, the hybrid method that adds both data enhancement and attention mechanism perform better and has the highest recognition rate, but the difference with the total accuracy is still large.

Experiment 2: Comparison Results of this Paper's Method. The training set containing positive face images from the face database is used to generate 1200 images using the data generation method, in which 20 images are generated for each person containing nine angles of 0°, ± 15°, ± 25°, ± 30°, ± 45° are selected, and some of the results of the face generation are shown in Fig. 8 below.

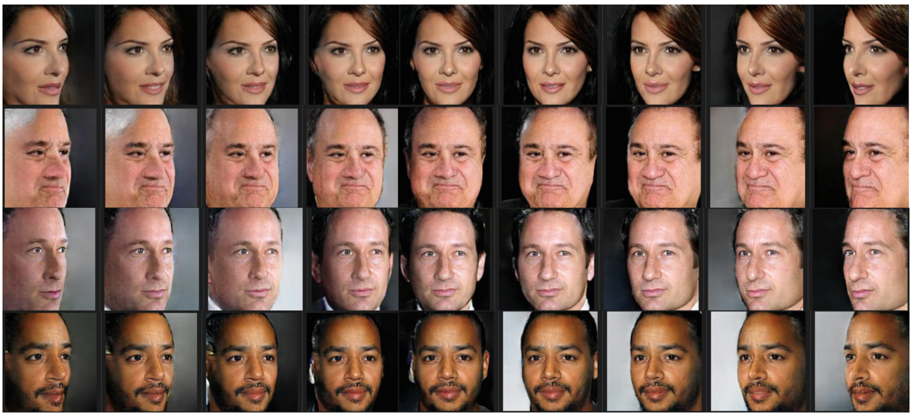

Fig. 8. Partial training set data generation results.

From the comparison of the experimental results in the previous section, it can be concluded that the hybrid method of Base + Attention + AUG, although the recognition accuracy is substantially improved, has a lower recognition accuracy in the case of lateral faces, especially in the case of large-angle face deflection. Next, the accuracy of the lateral face recognition algorithm based on data generation using this paper is compared with the Base + Attention + AUG method which has the best recognition effect in the previous section, and the results are obtained as shown in Table 4 below.

Table 4. Comparison of recognition accuracy results by angle after adding data generation.

Angle	Base + Att + Aug	Base + Att + Aug + Data generation
[-60°, -45°]	0.35	0.45
[-45°, -30°]	0.5	0.74
[-30°, 0°]	0.892	0.895
0°	0.956	0.963
[0°, 30°]	0.87	0.91
[30°, 45°]	0.782	0.85
[45°, 60°]	0.48	0.43
Total accuracy	0.8647	0.8988

Further comparative analysis of the data results in the above table can be analyzed to obtain the following two points.

1) Side-by-side comparison: First, after adding the data generation method, the face recognition accuracy improves at every other angle in the test set, except for a slight dip in face recognition accuracy at [45°, 60°]. Here, it is considered that more face features were lost due to extreme deflection angles, which affected the generation effect, and the symmetric loss function in the model will be improved subsequently. Second, by comparing the recognition accuracy of the same method at [-45°, -30°] versus [30°, 45°] and [-30°, -0°] versus [0°, 30°], it can be seen that the present algorithm has a higher recognition rate for forward deflection. This consideration is due to the fact that the left and right faces of each person are not perfectly symmetrical affecting the generation effect, which will be improved subsequently. Finally, the data-generated face recognition algorithm based on data generation has an accuracy of 96.3% at 0° angle and a total accuracy of 89.9%, which is the highest overall recognition rate.

2) Longitudinal comparison: The recognition accuracy of both methods decreases with the increase of face deflection angle. In large-angle deflection face recognition, the hybrid method that adds data generation performs better, and the recognition rate is significantly improved.

4 Conclusion

This article proposes a multi-posture face recognition algorithm based on data generation. By establishing a face recognition environment and utilizing public face database to organize and divide training and test sets, a series of comparative experiments among different methods are conducted. Data comparison confirms that the recognition accuracy of this multi-posture face recognition algorithm based on data generation is higher than that of four face recognition models: Base, Base + AUG, Base + Attention, and Base + Attention + AUG. The facial recognition algorithm generated from the data achieves an accuracy of 96.3% at 0° and an overall accuracy of 89.9%, effectively improving the accuracy of facial recognition for a variety of head postures.

In the future, the loss function of the data generation model will be improved to solve the problems of face feature loss due to extreme deflection angles and incomplete symmetry between a person's left and right faces.

Acknowledgements. This work was supported by the National Key Research and Development Project (2019YFB2101902).

References

1. Yansui, D.: Exploring 3D face recognition technology. Digital Technology & Applications **9**, 234–236 (2017)
2. Lichun, G., Ming, L.: Research on face recognition algorithm for multi-gesture. Comp. Knowl. Technol. **18**(21), 70–72 (2022)
3. Zhiyuan, L.: Overview of the current state of research on face recognition technology. Electr. Softw. Eng. **13**, 106–107 (2020)
4. Xinlin, Z.: A review of face recognition methods. Technol. Inno. Appl. **12**(2), 130–132 (2022)
5. Xueyuan, Z.: Optimization study of face recognition method based on local feature extraction. Automation Applications **64**(7), 54–56 and 60 (2023)
6. Goodfellow, I., Pouget-Abadie, J., Mirza, M., et al.: Generative adversarial networks. Commun. ACM **63**(11), 139–144 (2020)
7. Jie, C., Yibo, H., Hongwen, Z., et al.: Learning a high fidelity pose invariant model for high-resolution face frontalization. In: 32nd Conference on Neural Information Processing Systems, pp. 1–11. Montréal, Canada (2018)
8. Taherkhani, F., Talreja, V., Dawson, J., et al.: Pf-cpgan: profile to frontal coupled gan for face recognition in the wild. In: 2020 IEEE International Joint Conference on Biometrics (IJCB), pp. 1–10. IEEE, USA (2020)
9. Hang, Z., Jihao, L., Ziwei, L., Yu, L., Xiaogang, W.: Rotate-and-render: unsupervised photorealistic face rotation from single-view images. In: 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp. 5910–5919. USA (2020)
10. Jie, H., Li, S., Gang, S.: Squeeze-and-Excitation Networks. In: 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 7132–7141. USA (2018)
11. XerCis.: Images Data Enhancement Tools: Enhancers at a Glance, https://imgaug.readthedocs.io/en/latest/. Last accessed 21 June 2023
12. Xiang, W., Kai, W., Shiguo, L.: A survey on face data augmentation for the training of deep neural networks. Neural Comput. Appl. **32**(19), 15503–15531 (2020)

13. Gross, R., Matthews, I., Cohn, J., Kanade, T., Baker, S.: Multi-pie. Image Vis. Comput. **28**, 807–813 (2010)
14. Labeled Faces in the Wild, http://vis-www.cs.umass.edu/lfw/. Last accessed 21 June 2023
15. Parkhi, O.M., Vedaldi, A., Zisserman, A.: Deep face recognition. Proceedings of the British Machine Vision **1**(3), 6 (2015)
16. Yi, D., Lei, Z., Liao, S., Li, S.Z.: Learning face representation from scratch. Computer Science 1–9 (2014)
17. Yandong, G., Lei, Z., Yuxiao, H., Xiaodong, H., Jianfeng, G.: Ms-celeb-1m: a dataset and benchmark for large-scale face recognition. In: Computer Vision–ECCV 2016, pp. 87–102. Springer International Publishing, Amsterdam, The Netherlands (2016)
18. Kumar, N., Berg, A.C., Belhumeur, P.N. Nayar, S.K.: Attribute and simile classifiers for face verification. In: 2009 IEEE 12th International Conference on Computer Vision, pp. 365–372. Kyoto, Japan (2009)
19. PubFig: Public Figures Face Database, https://www.cs.columbia.edu/CAVE/databases/pubfig/explore/. Last accessed 21 June 2023