Studies Advanced in Robust Face Recognition under Complex Light Intensity

Zedong Fang[1,*] and Zhuoli Zhou[2]

[1] Alibaba Cloud Big Data Application College, Zhuhai College of Science and Technology, zhuhai, China, 519040
[2] School of computing, Zhuhai College of Science and Technology, zhuhai, China, 519040
*Corresponding author: zedongfang0829@stu.zcst.edu.cn

Abstract. Facial recognition tasks aim to automatically detect, recognize, and verify facial features through computer vision and pattern recognition technology. They have been widely used in various tasks, such as security monitoring and identity authentication. Thanks to the rapid development of machine learning and deep learning technology, breakthroughs have been made in the accuracy and speed of facial recognition. However, in complex scenes, especially when lighting conditions change, accurate facial recognition remains an unresolved issue. Focusing on the above issues, this article provides a detailed introduction to the latest research progress of facial recognition algorithms in dealing with complex lighting. Specifically, we introduced the representative work from three aspects: improving the algorithm to extract more features, selecting more appropriate data sets and more dimensional data. Secondly, we quantitatively compared the changes in facial recognition accuracy under different lighting conditions. Finally, we summarized the remaining issues in the field and discussed future development directions.

Keywords: Face recognition; light intensity; deep learning

1. Introduction

Face recognition task aims to automatically detect, recognize and verify human facial features through computer vision and pattern recognition technology. The commonly used facial features mainly include features such as eyes, mouth, and nose, as well as factors such as skin color, texture, and shape. Nowadays, facial recognition technology has been widely used in various scenarios, such as security monitoring, identity authentication, social media, and digital marketing.

Principal component analysis (PCA) and linear discriminant analysis (LDA), two feature-based techniques, are the most frequently used techniques in early face recognition technology. In order to recognize faces, these techniques essentially extract information from photos using mathematical operations and compare them with features from pre-existing facial databases. Face identification has been a particularly successful application of deep learning in computer vision since 2012. One of the most effective deep learning models currently is the convolutional

P. Kar et al. (eds.), *Proceedings of the 2023 International Conference on Image, Algorithms and Artificial Intelligence (ICIAAI 2023)*, Advances in Computer Science Research 108,
https://doi.org/10.2991/978-94-6463-300-9_101

network (CNN), which has made significant advances in the field of face recognition [1]. In 2014, Facebook's research team proposed DeepFace, the first facial recognition system to rival humans [2]. One of the greatest facial recognition systems currently available was provided by Google in 2015 with the introduction of FaceNet, an end-to-end system based on CNN [2].

The face recognition system first needs to perform face detection, then extract facial features, and finally use these features for training and matching. The quality of facial photos is crucial when extracting facial features [3]. There are several factors that can affect the quality of facial recognition photos, including posture changes, partial occlusion, lighting changes, and facial expressions. In the past decade, although various approaches such as LDA and LBP have been successfully applied to improve the accuracy of facial recognition algorithms, these methods did not achieve good results when faced with challenging conditions such as lighting and facial expression changes [4].

According to the process of facial recognition, in recent years, there have been three main directions for improving the accuracy and robustness of facial recognition algorithms in lighting problems: (1) improving algorithms to extract more features, (2) selecting more suitable datasets to obtain high-quality data, and (3) obtaining more dimensional data. The purpose of upgrading the algorithm is To enhance the number and quality of features that are extracted with the same dataset, thereby obtaining higher accuracy. In order to increase the amount and quality of features extracted while still utilizing the same technique, a more suitable dataset must be chosen. This will result in a dataset with higher accuracy. The purpose of obtaining more dimensional data is to improve accuracy by using data from other dimensions while maintaining the original algorithm and dataset.

Focusing on the above three frameworks, this paper aims to introduce the researches advanced in face recognition under complex light conditions. Specifically, the representative recognition methods will be detailed first, including their design ideas when dealing with light changes, their basic steps, advantages and disadvantages. Then, commonly used datasets are introduced and quantitative performance comparison and analysis of representative methods are reported. We finally summarize the challenging issues in the field of face recognition under complex light conditions and give a look out for its future development.

2. Face recognition for lights changes

2.1 Robust recognition through image enhancement

A time-tested technique for improving the contrast of digital images is histogram equalization (HE) [5-6]. Histogram h of digital image is defined as discrete function, which is given by equation (1).

$$h(x_k) = \mathrm{n_k} \tag{1}$$

Where x_k denotes the k-th intensity level in the range [0, L-1], and where $n_k$denotes the number of pixels in the input image.However, it has several drawbacks, such as losing some image features, making some local areas brighter than before, and being unable to maintain image brightness.Furthermore, there is another way to boost HE. Adaptive histogram equalization (AHE) works by dividing the input image into small "tiles." A local histogram will use a given probability density function to execute all regions occupying different grayscale ranges. By first computing the local histogram of the image and then dispersing the brightness, the AHE algorithm modifies the contrast of the image. The benefit of this technique is that it enhances the picture's local contrast while acquiring additional image information. However, if the local contrast is increased too much, the image will be distorted and the noise in the image will also be amplified.

Contrast limited adaptive histogram equalization (CLAHE) is an enhancement on adaptive histogram equalization , which primarily uses cutting histograms before calculating the value of Cumulative Distribution Function (CDF) to limit contrast magnification with a predefined value. The advantages of this method are faster, easier to implement, and fully automatic. Modified contrast limited adaptive histogram equalization (M-CLAHE) filtering was used to avoid the noise issue that CLAHE has after the enhancement procedure. However, image details will be lost due to low-pass filtering, which will result in information loss. Therefore, only the noisy areas will be subject to the discriminant low-pas$ filtering, while other areas will remain unchanged. For the M-CLAHE, the first stage is called pre-filtering stage, which uses Gaussian blur to smooth the enhanced image slightly. In the next stage, the intensity of the low-pass filtering operation can be carried out, and LP1 can do medium filtering. The original source low contrast input image is entered into the HE based enhancement block similarly to how AHE works. The discrimination filtering method eliminates the noise produced by the enhancement block.

Adjusting the double histogram while maintaining brightness Another version of contrast enhancement based on histogram equalization (HE) that does away with its drawbacks is maintaining double histogram equalization (BBHE). In order to preserve the average brightness of the image, the BBHE method divides the original image into two subimages based on the average value of the input image, and then equalizes the histogram of each subimage. The advantage is that the average brightness of the input image is maintained while the image contrast is enhanced, but when the histogram of the input image has peaks, histogram equalization will lead to excessive image enhancement.

2.2 Robust recognition through more dimensional data

As mentioned above, the quality of face photos is crucial, but in addition to directly processing images, we know that compared with two-dimensional face images, three-dimensional faces contain more geometric information and are not sensitive to pose and lighting changes [7]. Many research organizations have developed various 3D face databases in recent years to test and assess their methodologies, which can be recorded using the high-speed and high-precision structured light 3D measuring device. In addition, heterogeneous face recognition (HFR) is used to obtain more data through infrared scanning.

The task of face recognition across various racial, gender, age, expression, and other characteristics is known as heterogeneous face recognition. This corresponds to the traditional homogeneous face recognition, which refers to face recognition in the same race, gender, age and other aspects. Because of its greater difficulty and challenge, heterogeneous face recognition has attracted an increasing amount of attention and research in recent years. It can be used in many practical application scenarios, such as security monitoring, automatic access control system, authentication, etc. It can help solve the challenges and difficulties in face recognition across different races, genders, ages, expressions and other aspects, improve the accuracy and robustness of face recognition technology, and better serve society and people's lives. However, the disadvantage of this method is that the data generated by the 3D deformable model is lack details, which is quite different from the real face, and the actual application performance is poor. At the same time, the 3D face recognition system has high accuracy but also has shortcomings such as large computation, slow recognition speed, etc. The database is also scarce, lacks training samples, and requires special equipment such as 3D cameras, binocular cameras, etc. Such equipment is expensive. Therefore, from the perspective of real realization, it needs to invest higher costs and energy [8].

In dimly lit or completely dark conditions, heterogeneous near infrared visible light (NIR-VIS) facial recognition is achievable. Lezama et al. suggested using low rank embedding and cross spectral hallucination in order to generate heterogeneous photographs based on mosaic. Song et al. synthesized via images from NIR images using generative countermeasure network (GAN) and two-way model. Hot face images are used to create visible face images. Given the synthesized visible facial picture, any VIS facial recognition algorithm trained on VIS facial data can be used to match the synthesized image with the registered VIS image. This is a significant benefit of these synthesis methods [8].

2.3 Robust recognition through selecting more appropriate datasets

In addition to the above two methods, we also learned that we can increase the diversity of lighting by expanding the size of the database.The light condition during shooting can be obtained not only from the diffuse reflection of the skin, but also from the direction and range of the shadow cast, and the intensity and position of the highlight reflection. Inspired by this experience, the LeGendre et al. proposed a model that can achieve reverse lighting from a portrait, without any specific skin reflection model assumptions, and can also estimate omnidirectional high dynamic range lighting in the environment. This technology can obtain lighting information with higher frequency details, making more realistic portrait rendering and ARi visual effects possible. The advantage of this scheme is that it can improve the original quality of the image from the root, reduce the pressure of image enhancement in the later stage and improve the efficiency, but the disadvantage is that if the number of face images is increased blindly, it cannot effectively improve the diversity, but will increase data redundancy and increase the training burden [9].

The training model requires a large number of portrait photos with lighting labels, but it is almost impossible to collect such a large dataset in reality. Therefore, researchers have adopted a digital driven relighting technique based on images to synthesize portrait photos with lighting labels, and render realistic images by appropriately capturing complex light transmission phenomena.Under the theoretical framework of reflection field, people can obtain the subject image under re illumination by dot multiplication of reflection field and HDR ambient light.

In order to record the reflection field of characters, researchers used 331 LED lights installed in a spherical surface for photography. The reflection field was recorded through a series of reflective base images, with one LED light turned on each time to record an independent lighting result (One Light At a Time, OLAT). Six cameras were used to record the images of characters at different angles. The input image of the network is a normalized image of 256x256 pixels and corresponding LDR illumination. The encoder decoder architecture is used for training, and the final output is a 32×32 mirror sphere HDR image in logarithmic space to represent omnidirectional illumination.

Bendjillali et al. used face image enhancement technology to train Yale face database, which improve the accuracy by about 10%-15% in challenging lighting conditions [5]. Tian et al. conducted face identification and age estimation studies under three distances and ten dimming levels, and employed the Retinex image enhancement technique in YCbCr color space to improve the accuracy of face recognition and age estimation. The experimental results demonstrate that the face image created using the modified approach has a higher recognition rate when compared to the original image [10].

3. Experiment

In order to quantitatively analyze the changes in facial recognition accuracy under different lighting conditions, we introduce random lighting, extreme lighting and real lighting to construct the dataset following the paper [11]. We use the classical ResNet-50 to train the model. The hidden layer of the ResNet-50 network consists of four modules, each consisting of 3, 4, 6, and 3 residual structures. Each residual structure contains 3 layers of convolutional kernels with a kernel size of 1×1, 3×3, 1×1. Use the hard sharing to achieve parameter sharing between facial recognition tasks and lighting related tasks, where the first two modules share parameters while the last two modules learn the parameters of their respective tasks. The input of the network is facial images with different lighting, and the facial classification branch will pay more attention to the differences between different facial categories. The lighting processing branch will perform tasks related to the input image lighting, such as lighting parameter estimation and lighting classification. As the two branches of the network learn from specific tasks, the correlation between the obtained features will gradually decrease, which can achieve the effect of feature separation.

As shown in Table 1, the facial recognition network trained based on the base image can exhibit better recognition ability under extreme and abnormal lighting conditions such as the base image. With the increase of the number of base image samples, the recognition rate also increases. Under normal visual effects such as random and real lighting, the recognition rate of the base image and the network trained under random lighting is relatively similar. The experimental data also shows that blindly increasing the number of data does not necessarily improve the recognition rate. The amount of data that is added has a limited impact on enhancing the network's identification rate as the type of illumination increases to a certain point. The network's recognition rate can be increased by judiciously choosing the illumination sampling technique.

Table 1. Accuracy of ResNet with different sampling strategies

Train set	Extreme lighting	Base image	Random lighting	Real lighting	Average
Base image × 1	88.40%	75.50%	99.98%	97.20%	88.30%
Base image × 9	93.94%	93.23%	100.00%	99.73%	96.22%
Base image × 25	97.78%	97.15%	100.00%	99.87%	98.52%
Base image × 81	98.12%	97.68%	100.00%	99.87%	98.77%
Base image × 145	98.53%	98.65%	100.00%	99.87%	99.16%
Base image × 289	98.99%	98.65%	100.00%	99.87%	99.31%

4. Discussion

Face recognition technology is widely used in authentication applications because of its advantages of convenience and non-contact authentication. However, the existing face recognition technology is extremely vulnerable to artificial presentation attacks (PAs). For example, using photos, video playback, low-cost artificial masks and facial makeup attacks can easily deceive the face recognition algorithm. Face photos are now simple to get thanks to the growth of the Internet and social media, which raises the danger of face recognition systems being attacked. These problems have aroused widespread concern about the safety of face recognition technology [12].

In face recognition systems, an independent face anti-counterfeiting module is usually set in front of the face recognition module to ensure system security. Due to the variety of ways of face forgery attacks in real scenes, it is impossible to collect them comprehensively. This learning method, which is highly dependent on limited training sets, is difficult to predict all attack ways and cannot be effectively extended to unknown attacks. Different from the face anti-counterfeiting task other face-related tasks (such as face recognition and attribute editing) have strong generalization ability for applications in different scenes due to the training of a large number of real faces. Inspired by this this work proposes to replace the location of the face anti-counterfeiting module and the face-related tasks in the face system, and use the free prior knowledge obtained in the face-related tasks to serve the face representation attack detection task, so as to improve its generalization performance.

Face recognition has the greatest range of potential applications in the security sector, which not only gives the sector a fresh lease on life but also makes room for new development markets. The shortcomings of 2D projection are now being supplemented by 3D face recognition algorithms thanks to China's recent advancements in 3D measuring technology. They have also successfully overcome the conventional challenges, such as face rotation, occlusion, similarity, etc., which has also turned into one of the most significant development paths for face recognition technology.

Future development will primarily focus on integrating facial recognition technologies with smart homes.The embedded operating system and embedded hardware platform are combined to create the face recognition system in smart homes, which increases the integration of face recognition technology and smart home applications and has the qualities of an innovative concept and strong applicability.

Future big data-based technological development will focus heavily on face recognition. Big data has been adopted by public safety agencies today, which also addresses the shortcomings of conventional technologies. The ability to save and reuse this photo data using face recognition technology may significantly enhance the administration and general planning of public security information, which will turn face recognition into the key development trend in the next years.

5. Conclusion

In face recognition, lighting has always been a challenging problem. To solve this problem, researchers have tried many techniques and methods. There are three main directions, namely improving algorithms to extract more features, selecting more suitable datasets to obtain high-quality data, and obtaining more dimensional data. From this, it can be seen that future research directions on improving lighting issues are divided into three categories:

1. Develop more powerful algorithms: develop algorithms that can adapt to different lighting environments to improve the accuracy and robustness of face recognition. In the development of algorithms, light estimation methods based on physical models can be combined to restore 3D shape and lighting information more accurately. In addition, Generative adversarial network (GANs) are gradually blooming in this field.
2. Database construction: Establish more databases containing facial images under different lighting conditions, which will help train more accurate facial recognition models. Meanwhile, in recent years, research on heterogeneous facial recognition has become increasingly popular, and a database for HFR can also be established
3. Multi-sensor fusion: Combining different types of sensors (such as RGB, infrared, depth camera, etc.) to fuse data, using heterogeneous facial recognition to address the impact of lighting changes on facial recognition.

Acknowledgment

Zedong Fang is responsible for Abstract, Keywords, Introduction and Conclusion, Experimentation and Result. Zhuoli Zhou is responsible for Method, Discussion and References.

All the authors contributed equally and their names were listed in alphabetical order.

References

1. Guo, Guodong, and Na Zhang. "A survey on deep learning based face recognition." Computer vision and image understanding, 189: 102805 (2019).
2. Adjabi, Insaf, et al. "Past, present, and future of face recognition: A review." Electronics 9.8 : 1188 (2020).
3. Heinsohn, Daniel, et al. "Face recognition in low-quality images using adaptive sparse representations." Image and Vision Computing 85: 46-58 (2019).
4. Shepley, Andrew Jason. "Deep learning for face recognition: a critical analysis." arXiv preprint arXiv:1907.12739 (2019).
5. Bendjillali, Ridha Ilyas, et al. "Illumination-robust face recognition based on deep convolutional neural networks architectures." Indonesian Journal of Electrical Engineering and Computer Science 18.2: 1015-1027 (2020).
6. Singh, Gurjinder, and Amandeep Kaur. "Artificial Bee Colony Optimized Multi-Histogram Equalization for Contrast Enhancement and Brightness Preservation of Color Images." International Journal of Engineering and Manufacturing 13.1 (2023): 45.
7. Kortli, Yassin, et al. "Face recognition systems: A survey." Sensors 20.2 (2020): 342.
8. R. He, J. Cao, L. Song, Z. Sun and T. Tan, "Adversarial Cross-Spectral Face Completion for NIR-VIS Face Recognition," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 42, no. 5, pp. 1025-1037, 1 May, doi: 10.1109/TPAMI.2019.2961900 (2020).
9. LeGendre, Chloe, et al. "Learning illumination from diverse portraits." SIGGRAPH Asia 2020 Technical Communications. 2020. 1-4.
10. Huijuan Tian, Mingtian Qiao, and Minpeng Cai. "Face recognition and age estimation based on changing lighting conditions." Laser & Optoelectronics Progress 59.2: 0210019-0210019 (2022).
11. Wen Li, Yanli Liu, and Guanyu Xing. "Illumination Analysis for Deep Face Recognition." Journal of Computer-Aided Design & Computer Graphics/Jisuanji Fuzhu Sheji Yu Tuxingxue Xuebao 34.1 (2022).
12. Xu, Heng. "The application of deep learning-based face recognition system in public safety." International Conference on Cloud Computing, Performance Computing, and Deep Learning. Vol. 12287. SPIE, (2022).