

Research Advanced in the Face Recognition

Wenjie Liu^{1, a, †}, Xinyang Wang^{2, *, †}

¹Jingling Institute of Technology Nanjing, China

²University of Cincinnati Shandong, China

*Corresponding Author Email: wang5x8@mail.uc.edu, ^a1283244921@qq.com

[†]These authors contributed equally.

Abstract. Face recognition has always been a research hotspot in the field of computer vision, whose basic task is to recognize the identity of the face in the image. With the rapid development of computer technology, both the accuracy and speed of face recognition have achieved great breakthroughs, and face recognition technology has been widely used in the field such as cyber security, artificial intelligence and so on. In this paper, based on detailed literature research and analysis, we present a comprehensive review of the research work on face recognition technology. Specifically, we first analyze the difficulty of face recognition task from the human face feature extraction, which due to the variations within and between classes. We then introduce the main frameworks of face recognition from geometric feature-based method, template based method and model-based method. We further compare the performance of different algorithms on different common datasets. Finally, we give an outlook on the future research directions of face recognition.

Keywords: Face image recognition, Deep learning, Masked face, Convolutional neural network.

1. Introduction

Face recognition, as an advanced biometrics technology, is based on facial features to complete identity recognition, which has been used in many fields such as cyber security, artificial intelligence and so on. Face recognition can be traced back to the 1950s, whose technologies have mainly experienced three development stages with the development of computer technology. In the first stage, the human faces are described by the local features of facial organs and then recognized with personnel operation. In the second stage, face recognition method gradually evolved to rely on geometric feature parameters and multidimensional feature vectors to describe the face image information, and then use human-computer interaction to complete face recognition. In the 1990s, face recognition entered the third stage, where automatic face recognition with the help of artificial intelligence models [1]. On the basis of realizing face recognition, the recognition of occluded face images and 3D face recognition are further studied.

In recent decades, face recognition technology has been more and more popular. In 2014, the accuracy of face recognition based on computer vision reached 98.52%, surpassing the human eye recognition ability 97.53% for the first time. Early research on face image recognition focused on occlusion-free face recognition, focusing on resolving inter-class and intra-class differences in face images. Inter-class differences are generally learned to distinguish different individuals, while intra-class differences are often used to address the identification of the same individual in different scenarios. As the model continues to improve the accuracy of unoccluded face recognition, researchers gradually turn their attention to the task of occluded face recognition [2]. Especially in recent years, with the deteriorating situation of the new crown epidemic, wearing a mask has become a basic measure for daily protection. In this case, how to achieve accurate recognition of occluded faces has become an urgent problem to be solved. Face occlusion is a difficult scene of face, and the acquired face image information is often lost, which makes the learned face image inter-class difference features and intra-class similarity features often contain occlusion noise or cannot describe face information [3].

For occluded or unoccluded face images, the main difficulties in feature learning are mainly reflected in: (1) Since the structures and shapes of facial organs of different individuals are basically

similar, the learned face features are more conducive to face localization tasks; Differential features of face images are only reflected in parts of organs (such as slight differences in lip color or nose size), which makes the feature information that distinguishes faces often overwhelmed or ignored. (2) Face the face shape is very unstable, people's expression changes through the change of facial muscles, and from different angles, the vision of the face will also have very big differences. A face may be normal when viewed from the front, and angry when viewed from the side. And in this case face can also be influenced by light conditions, by coverings. (3) The inevitable interaction of the above two points makes face feature learning more difficult and further inhibits the accuracy of face recognition [4].

Focusing on the solutions of the face image recognition, in this paper, we conducted a detailed literature analysis and introduce the representative technical framework of masked face recognition, which mainly includes geometric feature-based methods, model-based methods and template-based methods. We further compare the performance of different algorithms on different common datasets. Finally, we give an outlook on the future research directions of face recognition.

2. Face Recognition

At present, we can see that there are three common different methods of face recognition in the market. They are geometric feature-based method, template based method and model-based method. The method based on geometric features is a very traditional method. It is also very old, and the accuracy of recognition is not very good. Usually, this method needs to be combined with other algorithms to get better results, that is, it depends on other algorithms. Second, template based method is the most commonly used framework in today's society, which is good for face recognition. It can be divided into correlation matching algorithm, feature face method, linear discriminant analysis method, singular value decomposition method, neural network method, dynamic connection matching method and so on. These representative algorithms will be introduced one by one in the following part of this paper.

2.1. Method based on Geometric feature

We are all part of the community of human destiny, so we should be very clear about human facial features, even if we don't know the knowledge of sociology and anthropology. Obviously, the face consists of eyes, nose, mouth, chin and other parts. There is no same leaf in the world. Similarly, there is no identical face or facial features in the world. It is precisely because of the various differences in the shape, size and structure of the five senses that every face in the world is very different. Therefore, the geometric description and size feature of the shape and structure relationship of these parts can be used as an important feature for face recognition and judgment. Firstly, we apply geometric features to the description and recognition of facial chin and side face contours. Then, multiple feature points are determined according to the 3D curve and contour of the face, and a set of feature measures for recognition, such as length, angle, etc., are derived from these feature points. Geometric features are usually used in frontal face recognition, because judging the face curve from different angles is a major problem in the current face recognition technology. The light and shadow at different times and the color of human face will affect the description of the curve. From the front, this technology can recognize faces by extracting the positions of important feature points such as eyes, mouth, nose, mouth, tongue and ears, as well as the geometric and three-dimensional features of important organs such as cheekbones as classification features. However, Professor Li Heping of Nanjing University is not optimistic about the accuracy of this feature extraction method. But a professor from Sichuan University improved the algorithm and the accuracy improved a little.

Nevertheless, the accuracy of this method in the field of face recognition is still too low. The main reason is that the second layer space of the features reflected by the main components of the face is too dense, which will greatly reduce the feature dimension during recognition, but this method is non local, which is not good. Then, the linear feature of the whole coordinate space is suddenly extended

to the core function in its deep level and core due to some security. At the same time, because its non topological essential characteristics have no reason, after being axially projected by the computer wide area network, the proximity of adjacent points has nothing to do with the proximity of the midpoint in the original image space, while the ideal features of pattern analysis and segmentation are the locality and topology of facial features and expressions. Although there are great differences between ideal situation and reality, there must be friction in reality. This seems in line with the universality of algorithms about neutral language handling to a certain extent in some ways. Therefore, it is very important to find an expression with this more reasonable function in the normal form. For this reason, Atick came up with a feature calculation method and recognition method for face recognition due to detailed features in the laboratory, which is slightly different from other methods. This method has good effect, more reasonable and high universality, and has made progress and good effect in practical application. It also becomes the basis of face recognition algorithm.

2.2. Template-based method

The deformation template method changed from the previous method is not a super novel method. It can be regarded as an effective and reasonable deformation and improvement of the geometric feature method, which must be better to some extent. However, the basic idea of this method is to design an organ change model with adjustable parameters (i.e. deformation template) to define and determine the energy function, and then minimize the energy function by adjusting the model parameters. In this step, the model parameters are regarded as the geometric features of the organ. Although the idea of this method is good, there are two problems. Firstly, the weight coefficients of various costs in the energy function can only be determined by experience, which is difficult to be extended to general algorithms, and this empirical state is difficult to change. Second, the optimization process of energy function is far more time-consuming than expected, so it is difficult to be applied in practice and can not produce objective economic value. Parameter based face representation, even if it is a mathematical formula or expression that can effectively describe the salient features of human face, requires a large number of complex irreducible preprocessing and fine parameter selection. At the same time, the use of general geometric features only describes the basic shape and structural relationship of components, which leads to the avoidance and omission of local subtle features and the loss of some information, so it is more suitable for rough classification. Moreover, the existing feature point detection technology is far from meeting the requirements of safety, intelligence, biology, chemistry and other interdisciplinary disciplines in accuracy, and the amount of calculation is amazing. Geometric features can be the shape of eyes, nose, mouth, etc. The geometric relationship between them is parallelogram, diamond or isosceles right triangle (such as the ratio of the distance between them to the corresponding length). Although the recognition speed of these algorithms is fast, the memory is too small and the recognition rate is low, which still can not meet the needs.

KL transform is an optimal face recognition method based on orthogonal transform. After KL transformation, a new set of orthogonal bases is obtained from high-dimensional image space, and the important orthogonal bases are retained and extended to low-dimensional linear space. If we assume that the projections of human face in these low dimensional linear spaces are separable, then these projections can be used as feature vectors for recognition, which is the basic idea of feature face method. These methods need more training samples and are completely based on the statistical characteristics of image gray. At present, there are some improved feature face methods.

The fish face was invented by Ronald Fisher, which must be the origin of the name of the fish face. The linear discriminant analysis (LDA) theory based on Fisher face is similar to the principal component analysis (PCA) of feature face. They are the methods of overall dimensionality reduction of original data and mapping to low dimensional space. LDA and PCA are based on overall data, which are different from LBP in extracting local texture features.

WeiWei proposed one Face recognition algorithm based on the ELM, which is called by human beings extreme learning machine. This method is similar to the fisherman facial recognition approach

described above. But the classifier used in the algorithm is ELM, which is called by human beings extreme learning magic machine. ELM can complete the application of regression and classification, it can randomly generate the hidden node parameters and determine the output weights of SLFNs instead of iterative tuning. So, ELM classification is fast and easy to implement. The algorithm reduces the dimension of the whole data through ELM. PCA and LDA are not optimal choices for dimensionality reduction. It is better to discriminate local alignment by DLA and classify by SVM, and ELM is a combination of these two methods. So ELM can realize face recognition faster and better.

The basic principle of PCA face recognition is to treat the face image as a feature vector. Then do the face image feature vector dimensionality reduction to find the key information for feature extraction. In the actual process of face recognition through PCA, PCA face recognition firstly preprocesses the face image, transforms the face into a row vector, and then obtains an average face through decentralization [5]. After that, its covariance matrix and its eigenvalues and eigenvectors are calculated, and then the eigenvalues are sorted and screened. The algorithm will decentralize to get an average face and calculate its covariance matrix. Then it will multiply the transpose of the different face matrix by the feature vector to get the featured face. Finally, it calculates Euclidean distance and projection vector so that finishes the face classification and face recognition. PCA can compound many features into a few features and reduce the loss of information during dimensionality reduction, this allows for more accurate recognition of faces with multiple features.

2.3. Model-based method

Machine learning originated from the origination of computer science, Professor Turing of Manchester university raised the question, can machine think? Machine learning is an interdisciplinary subject, which covers probability theory, statistics, approximation theory, dualism, convex analysis, data structure, complexity theory of algorithm, and mechanical model of physics. Deep and machine learning is aimed at how computers imitate or execute people's learning process to acquire new knowledge and skills, reconstruct existing knowledge and technology, and continuously improve their ability. DNN is a technology widely used in machine learning.

Convolutional neural network is an in-depth research widely used in face recognition. We can roughly divide it into three layers. The convolution layer extracts features, the physicochemical layer reduces the dimension of data, and the full connection layer classifies the latest data. All this is about data. Its biggest advantage is that it can use massive data for training and simulation, so it can learn and train robust facial expression. This method does not need to make special settings for various indoor temperature changes (such as lighting, posture, facial expression, psychological age, age, etc.), but can be understood through training materials. The biggest drawback of deep learning is its advantages. For example, in order to train, they must have a lot of data, which many people think is an advantage and many people think is a disadvantage. These data sets also need to be changed and calculated enough to expand them to unpredictable versions. Fortunately, some famous universities choose to contribute unconditionally, such as Zhejiang University and Shandong University. So many big datasets about our face contains those features, which are reduced and released by this algorithm, which is natural image of our face. These university datasets can omit the steps of searching data and then we found it is useful, which is good at training CNN model to nice ones. Then, we can use the neural network model to reduce the dimension and learn how to distinguish features at the same time., which can be trained as component classifier or measure the learning pan relationship using raised investment. It is considered that the specific method of training from one end to the other can be combined with other systems.

CNN face recognition model can be trained in very different ways. Whether supervised learning or unsupervised learning. There are many kinds of algorithms. One is to classify the problems one by one, and in the training set, each topic has a corresponding classification. After the training, we can remove the classification layer and happily use the value corresponding to the feature of the previous layer as a method to represent the face state. The model can be used to focus on identifying and

training topics that do not exist in this world. These features are often referred to as a feature in the literature related to deep learning, which is exactly a bottleneck feature. For example, the first segment has been completed by us. This model can also be further trained and optimized by other technology types, just like a tutorial class, so as to optimize the bottleneck characteristics of the target application (for example, the CNN model can be adjusted to varying degrees by using joint Bayesian or Jacobian determinants, cost functions, spherical harmonics and different loss correlation functions). Another commonly used perfect method of learning face representation is called directly learning bottleneck features by optimizing the distance measurement between pairs of faces or between three faces.

The outbreak of the epidemic at the beginning of last year quickly swept the world due to the strengthening of control measures. Now we basically have to wear masks when we go out. It is very important to recognize faces when wearing masks. Especially in the dense places such as shopping malls, it is unrealistic for you to ask customers to take off their masks in public places, and people will certainly refuse. From the beginning, apple wanted to solve this problem by letting customers wear watches and using the ecosystem of Apple system, but a complete product ecosystem means higher living costs. Many people don't have that much economic foundation, so this is not a perfect strategy. To solve this issue, a new masked face recognition framework is proposed, which is named as MTCNN. MTCNN, as its name implies, is a multi task convolutional neural network. The network model mainly frames the face through three cascaded networks, p-net, R-Net and o-net; For the model training of mtcnn, all the input images are positive samples (photos with masks), and the model input of negative samples is not used.

3. Experiments and discussion

3.1. Datasets

The research of any computer model and algorithm is inseparable from the help of the dataset. Dataset can teach our model to identify desired patterns or perform particular tasks. Based on the general feature of the dataset, the model and algorithm can learn to identify the range of work and the sequence of the task in the actual test. For face recognition, we also need existing datasets to train face recognition algorithms. Depending on the type of pictures in the dataset, it can determine the accuracy and method of face recognition by our algorithm. The dataset that is trained for unmasked face recognition requires the face recognition should be masked in various ways. Selecting the correct dataset enables the face recognition algorithm to operate normally and accurately. In this article we will show the training set used for face recognition that running in Table I.

Table 1. Commonly used datasets for face recognition

Name	Color Images	Image Size	People numbers	Pictures number each person	Introduction
MIT database	Yes	480 x 640	154; 82 Male, 74 Female	6	Data set includes 2 front, side, and 3/4 views At the same time, part of the data set includes photos of students making up
Color FERET	Yes	512×768	1199	N/A	The database contains 1564 sets of images In this data set totally contains 14,126 images. Data set contains 1199 individual images and 365 repeated images.
Olivetti - Att – ORL	No	92 x 112	40	10	All the image is frontal and slight tilt of the head
WIDER Face	Yes	N/A	32203	N/A	WIDER Face dataset contains 32,203 images and 393,703 faces bounding box annotations. All face images in this dataset contain different levels of posture and different degrees of movement

UCCS dataset	Yes	5184×3456	4160	N/A	this dataset is long-range surveillance photos that captured at University of Colorado Colorado Springs developed primarily for research and development of "face detection and recognition research towards surveillance applications
AR Face Database	Yes	576 x 768	126; 70 Male, 56 Female	26	Facial expressions in all positive photos include: angry, screaming, smiling, normal,etc. It also features a variety of shades (scarves and sunglasses).The photos were taken at CVC under strict control, with no restrictions on clothing, hair or makeup
CVL Database	Yes	640 x 480	114; 108 Male, 6 Female	7	The dataset contains photos of faces from different angles, including faces with full frontal smiles with teeth and smiles without teeth
The MUCT Face Database	Yes	480 x 640	276	10 to 15	The images contains frontal and 3/4 views, they all have different lighting sets and manual landmarks
The Yale Face Database	No	640 x 480	10	576	v9 poses x 64 illumination conditions

Table 2. The comparative result

Method	Dataset	Number of images used	Accuracy rate
Based on geometric features [6]	MIT database	The number of faces in each image ranged from 4 to 47, for a total of 620 faces	94.52%
Based on geometric features [7]	N/A	2570 pictures (1420 males and 1150 females)	86%
Based on geometric features [8]	Orl dataset and Caltech Faces	All	88.50%
Based on algebra features—PCA [9]		400 images of 40 people's faces	93.6%
Based on algebra features—PCA		One hundred and forty (140) training images generated average faces and feature faces	92.02%
Face recognition based on neural network [10]		270 images	92.2%
Local Face Analysis [11]	FERET	1002 quasi-positive images of 429 people/	94.7%
Local Face Analysis [12]			90.6%(dup1)
Based on algebra features—PCA	Yale Face Database	165 images of 15 people's faces	79.65%
Based on algebra features—fisherface classifier			87.44%
Face recognition based on neural network	CTK	All	86.7%
HFSRL-v [13]	UCCS dataset	Face images from 1,732 labeled persons are used	80.8%
HFSRL-m [13]			79.5
HFSRL-v [13]	NJU-ID dataset	dataset contains face samples from 256 persons	71.4 %

3.2. Performance comparison

In fact, the process of face recognition is similar to the process of target recognition, which is to recognize and classify objects in pictures. Therefore, the performance indicators of face recognition are also similar to those of target recognition. Index parameters include False Rejection Rate(FRR), False Acceptance Rate(FAR), Detection rate, False alarm rate, FPS, IOU. The accuracy of a face recognition algorithm refers to the proportion of successfully recognized samples to the total number of samples. This is the most direct indicator, through which we can directly see the recognition ability of the algorithm, but there are still some errors. So we need more another indicator to compare the performance of face recognition algorithms. FRR is the number of falsely rejected cases versus the total number of rejected cases. The higher the FRR, the higher the security performance of the algorithm. FAR is the ratio of falsely accepted cases to the total number of accepted cases. The false acceptance ratio refers to the probability that a stranger uses face recognition and is successfully recognized. The higher the false acceptance rate, the less secure the algorithm is. FPS is the frame rate, which refers to the speed and performance of the face detection algorithm to detect faces. IOU refers to the intersection of the face range detected by the algorithm and the actual face range. The larger the IOU, the more accurate the face recognition range.

There are many indicators that can judge the performance of face recognition, and these indicators are often used according to the different applications of the algorithm. Next, this article will compare the performance of each algorithm from the method used by the face recognition algorithm, the name of the training dataset used for algorithm training and testing, and the accuracy of target detection. The comparative result is shown in Table II.

The following conclusions can be drawn from Table II: For face recognition algorithms of various methods, using different datasets yields different accuracies. This may be related to the environment of the picture in the dataset, the appearance of the face, and the pose of the face. From the Table II, we can see that the recognition accuracy of the algorithm using ORL and FERET datasets is significantly higher than that of other datasets. For face recognition based on algebra features using OrL data sets, considering that the two algorithms use the same but different data sets and the datasets may be updated, the accuracy fluctuates, but the fluctuation range is within the error range. Therefore, it can be concluded that face recognition using PCR is relatively stable and accurate in the current development [8-9]. For Face recognition based on neural networks, their algorithms are the same, but because of the use of different data sets, the error range of the results is too large [10]. We can guess that the face images in the CTK dataset may be more complex than the OrL face data images, which leads to a larger bias in the training process of the neural network, which ultimately affects the accuracy of face recognition. For Hierarchical Deep CNN Feature Set-Based Representation Learning for Robust Cross-Resolution Face Recognition, they use different models, and the accuracy of target detection using the HFSRL-m model is 1.56% higher than that using the HFSRL-v model. The HFSRL-m model may slightly outperform the HFSRL-v model [13].

4. Discussion

The main function of face recognition application is authentication. Face recognition involves authentication from images and from real-time video. Their basic principle is to capture the features of the recognition target and match the features with the facial features in the face database to achieve identity recognition.

Biometrics Technology has been used in various fields, such as finance, e-commerce, military, enterprise security networks and so on. For example, The Business Department of United Bank of Texas once showed a real scene where a customer forgot to bring his Bank card and his password. But he did make a withdrawal from the ATM. Biometrics allows a camera to scan a customer's retina to quickly authenticate their identity and then conduct a business service. Many large commercial enterprises use the "Iris Recognition System" to authenticate employees, through this system can accurately identify staff attendance. In addition, After the 9.11 incident, national security becomes

more and more important, especially in airport security protection. With the help of the progress of face recognition technology, Visag's facial recognition technology can accurately identify the information of every passenger in the airport, and can timely identify and report to the police when illegal elements are found.

In the society theft, robbery occurs from time to time, and the traditional mechanical design security door has been unable to meet the needs of security. The application of face recognition technology in security door has greatly improved the safety of residents. This not only makes it more convenient for users to travel, but eliminates the need for the traditional means of opening doors with keys. At the same time, it also greatly protects the property and safety of customers.

Face recognition-based identity authentication in each industry, each product has played a crucial role. As face recognition algorithm continues to update and progress, its application is more mature and more extensive

Enterprise confidential assets security. This field contains confidential asset access management and Enterprise employees attendance system, anti-theft door and so on.

E-passport and ID card. Currently, China is promoting e-passports, In this program, face recognition has applied for verification of passenger entry and exit identity.

Public security and criminal investigation. for example, public security organ using face recognition system and network apply for the control monitor on everywhere to search, match and arrest fugitives.

Cyber security. for example, This field contains enterprise computer login, governmental network and e-commerce. In electronic commerce, all transactions are completed on online, At present, the authorization of transaction is still realized by password. But This transaction method is very risky. If the password is stolen, it not only make the transaction fail, but also the information of client and confidential data will be stolen by hackers. This will have a huge impact on e-commerce. However, through the method of double authentication, the use of password and face information recognition, which makes the electronic transaction security factor increased, greatly ensure the security of the transaction

5. Conclusion

In this thesis or article, we discuss issue of face recognition and presented a comprehensive review of the research work on masked and unmasked face recognition. We introduce the representative recognition algorithms according to the different technology frameworks, which mainly includes the Methods formed by geometric features, methods driven by modularity, and methods for model establishment and formation. We also introduce the common face recognition datasets and compare the performance of these recognition algorithms. Finally, we discuss the future direction of face recognition technology and predict their application scenarios.

References

- [1] Goldstein A J, Harmon L D, Lesk A B. Identification of human faces [J]. Proceedings of the IEEE, 1971, 59(5): 748-760.
- [2] Wang K, Li X, Wang W, et al. Application of kernel method on face feature extraction [C] //2007 International Conference on Mechatronics and Automation. IEEE, 2007: 3560-3564.
- [3] Zou, G., Fu, G., Li, H., Gao, M., & Wang, K. A survey of multi-pose face recognition. Moshie Shie Yu Rengong Zhineng/Pattern Recognition and Artificial Intelligence, 2015, 28(7), 613-625.
- [4] Li X X, Liang R H. A review for face recognition with occlusion: from subspace regression to deep learning[J]. Chinese Journal of Computers, 2018, 41(1): 177-207.
- [5] Owais M, Shaikh A, Jalal A A, et al. Human Face Recognition using PCA Eigenfaces[C]//2019 IEEE 6th International Conference on Engineering Technologies and Applied Sciences (ICETAS). IEEE, 2019: 1-6.

- [6] Yi Y C, Feng G. Face Detection Algorithm with Multi-stage Decision Based on Skin-color and Geometrical Features[C]//2018 12th IEEE International Conference on Anti-counterfeiting, Security, and Identification (ASID). IEEE, 2018: 39-42.
- [7] Xiao J, Guo Z, Cai C. Gender recognition based on face geometric features[C]//MIPPR 2013: Pattern Recognition and Computer Vision. International Society for Optics and Photonics, 2013, 8919: 89190R.
- [8] Ouarda W, Trichili H, Alimi A M, et al. Face recognition based on geometric features using Support Vector Machines[C]//2014 6th International Conference of Soft Computing and Pattern Recognition (SoCPaR). IEEE, 2014: 89-95.
- [9] Kwak K C, Pedrycz W. Face recognition using a fuzzy fisherface classifier[J]. Pattern recognition, 2005, 38(10): 1717-1732.
- [10] Mahmood S S H, Babaei P. Using SIFT Descriptors for Face Recognition Based on Neural Network and Kepenekci Approach[C]//2019 5th Conference on Knowledge Based Engineering and Innovation (KBEI). IEEE, 175-180.
- [11] Jiao F, Gao W, Chen X, et al. A face recognition method based on local feature analysis[C]//Proc. of the 5th Asian Conference on Computer Vision. 2002: 188-192.
- [12] Su Y, Shan S G, Chen X L, et al. Integration of global and local feature for face recognition[J]. Journal of Software, 2010, 21(8): 1849-1862.
- [13] Gao G, Yu Y, Yang J, et al. Hierarchical deep cnn feature set-based representation learning for robust cross-resolution face recognition[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2020.