The Development and Applications of Facial Recognition Technology

Yihan Xu[a]
College of Information Engineering, China Jiliang University, Hangzhou, Zhejiang, China

Keywords: Face Recognition, Deep Learning, Neural Network.

Abstract: The rapid development of face recognition technology has brought great convenience to society and has been widely used in many fields. With continuous improvement, the current face recognition technology has become more reliable. This article first introduces the role of pre-processing work in face recognition, focusing on how these pre-processing steps affect the final recognition results. Subsequently, this article lists three typical applications: emotion recognition, disease auxiliary diagnosis, and micro-expression lie detection. These applications demonstrate the development and application of face recognition technology in different fields. For these applications, different algorithms are used to achieve their respective recognition effects. Although existing algorithms have made significant progress in recognition efficiency and accuracy, they still face technical difficulties and security risks. In response to these problems, this article proposes strategies to solve them by increasing the amount of data, optimizing laws and regulations, and improving models. Finally, this article looks forward to the future development direction of face recognition technology and proposes the possibility of transforming static face recognition into more comprehensive and stable dynamic face recognition, which indicates the broad application prospects of dynamic face recognition technology.

1 INTRODUCTION

Face recognition is a kind of technology that processes and analyzes facial images through programs, extracts facial features for learning, and achieves computer recognition and verification of facial identity information. In the past few decades, with the development of computer technology and artificial intelligence, face recognition has been greatly refined (Chen, 2023). The development of face recognition can be traced back to the 1960s. At that time, face recognition mainly relied on manual methods for identification and extraction. However, due to various limitations, the accuracy and efficiency of manual and computer recognition were not high in the early days. Face recognition technology first entered the application stage in the late 1990s. Later, the rise of deep learning technology also improved the accuracy of predictions.

Deep learning-based facial recognition algorithms have become one of the most popular research topics. This method utilizes deep neural networks for feature extraction and classification, allowing for the learning of more abstract and high-level feature information to achieve accurate recognition of facial identity information (Wang, 2023; Shepley, 2019). Nevertheless, as society continues to progress and the internet becomes more widespread, these traditional recognition methods are facing severe challenges, and people are also proposing stricter requirements for recognition and authentication, such as higher accuracy and greater convenience.

This article introduces the development of facial recognition technology in three parts. The first part is the preparatory work. First, it introduces the two methods of data augmentation and normalization, which show that data preprocessing can improve the recognition accuracy of the system. Then the author summarizes the principles and effects of the two methods of coordinate regression and heat map regression. Finally paper introduces the advantages and disadvantages of the two algorithms, principal component analysis (PCA) and convolutional neural

[a] https://orcid.org/0009-0006-9755-0093

Xu, Y.
The Development and Applications of Facial Recognition Technology.
DOI: 10.5220/0013677200004670

In *Proceedings of the 2nd International Conference on Data Science and Engineering (ICDSE 2025)*, pages 11-15
ISBN: 978-989-758-765-8

network (CNN). The second part is the specific application of three types of face recognition: emotion recognition, disease-assisted diagnosis, and micro-expression polygraph. By summarizing the research results of several scholars, it proves the degree of development of face recognition technology in various fields. The last part is the existing limitations and future prospects. In this part, this paper points out that there is still a shortage of databases for face recognition technology, and the model has more room for development. At the same time, it also puts forward the goal of transforming static face recognition into a more comprehensive and stable dynamic face recognition in the future. The purpose of this paper is to summarize the development of face recognition technology and put forward the shortcomings and future development directions.

2 BASIC WORK BEFORE FACE RECOGNITION

2.1 Data Preprocessing

After the system collects images, it often undergoes multiple processing steps before learning their features. These steps include normalization and data augmentation. These have a very significant impact on the recognition accuracy of the overall system.

Normalization is divided into geometric normalization and grayscale normalization. Geometric normalization is to adjust the faces in the collected images to the same position through different methods such as cropping, scaling, and rotation, which can minimize the impact caused by differences in angle, size, and distance. CNN can construct a powerful face classifier with multiple processing layers, aligning the feature points marked by feature extraction with the model to achieve face alignment. Grayscale normalization, also known as grayscale conversion, is to convert the original photo into a grayscale photo within a specific range. This step is to improve the contrast between the face and the environment in the image, weaken the problems of different light intensities and light angles during the shooting stage, and also make up for the uneven image quality to improve image quality (Huang, 2024).

Data augmentation is a method to expand the data set. It uses techniques such as rotation, horizontal flipping, and scaling to increase the input sample library. Currently, the face data sets on the network are usually relatively small and unevenly distributed. Using data augmentation can handle the existing data sets, enrich the sample library, and balance the sample distribution.

In summary, these are the main steps of data set preprocessing, which have a very significant impact on the recognition accuracy of the system, especially when the image sizes are different and there are illumination changes.

2.2 Feature Detection and Learning

Feature point detection is the most crucial step before other steps. This step is not only the basis of data preprocessing but also the basis of subsequent feature extraction and feature learning. One method is coordinate regression, which uses the model to extract features and learn regression, and is suitable for some more complex tasks. Another method is to use the heatmap regression method. The heatmap regression method performs key point detection. Firstly, the input image is subjected to feature extraction, and then a heatmap is generated for each feature. The pixel value in the heatmap indicates the probability of the feature point appearing at that position. In the figure, the Gaussian distribution is used to simulate the position of the feature point, the distribution center corresponds to the true coordinate of the feature point, and the brightness indicates the distribution probability. Finally, the feature point coordinates positioning and coordinate mapping are performed. Feature learning is a technology set that converts raw data into a form that can be effectively developed by machine learning. There are many different kinds of methods for feature learning now, such as CNN and PCA.

PCA is a more traditional method, aiming to convert the processed face image sequence into several main technical indicators through the idea of dimensionality reduction. Then, the sequence data is centralized and its covariance matrix is calculated. After calculating the eigenvalues of the obtained covariance matrix, they are arranged in descending order, and the eigenvectors corresponding to the first k eigenvalues are selected to form the projection matrix A. This is the complete process of PCA learning features.

Another deep learning method is CNN, which uses activation functions to reflect the characteristics of neurons and uses convolutional layers, pooling layers, and fully connected layers for feature learning. The image undergoes convolution operations through multiple convolution kernels in the convolutional layer to analyze the features of each small area. These convolution kernels can automatically learn features such as edges and textures. Each convolution kernel generates a feature map, indicating how the features are extracted under this convolution kernel. Through multiple convolution kernels, multiple feature maps will be obtained. Then, the size of the feature map is reduced through the pooling layer to reduce the computational cost while retaining important features. After multiple convolution and pooling operations, the feature maps are usually flattened and converted into a one-dimensional vector. Finally, in the fully connected layer, the extracted features are mapped to the final output category and classified using the activation function. The entire process can be summarized as: the input image passes through multiple convolutional layers, activation functions, and pooling layers to gradually extract features, and finally is passed to the fully connected layer for classification. During this period, the weights are adjusted through the gradient calculation of the loss function to reduce the loss in the next forward propagation (Huang, 2024). CNN can automatically learn features at different levels, from low-level features such as edges and textures to high-level features such as face contours and the positions of facial organs. It can effectively learn complex nonlinear features without the need for manual feature extraction.

3 FACIAL RECOGNITION SPECIFIC APPLICATIONS

3.1 Emotional Recognition

Facial emotion recognition is a technology based on facial expression analysis, used to identify an individual's emotional state. By analyzing the changes in facial expressions, such as the movements of eyebrows, eyes, lips and other parts, the system can determine the person's emotional category, such as happiness, anger, sadness, and surprise. With the advancement of deep learning and computer vision technology, especially the application of CNN, the accuracy and practicality of facial emotion recognition have been significantly improved. Wang et al. proposed a feature extraction method that fuses the Complete Local Binary Pattern (CLBP) and geometric salient features. Using the Dlib library for feature point positioning, a feature ratio vector is constructed according to the significant regions of facial expression changes, and the fine-grained texture features extracted by fusing geometric salient features and CLBP are used as the input feature vector for expression classification. After the experiment, the performance of this algorithm on the CK+ database is that the accuracy rate is as high as 92.5% (Wang et al., 2020). Wang proposed a recognition method that combines Faster R-CNN in the process of facial recognition. Firstly, the Multi-Task Cascaded Convolutional Networks (MTCNN) is used to locate the facial key points of the image to generate a 3D reference model, and then the model is projected into the initial frontal face for comparison. Finally, the comparison data is stored in the database to complete the processing of the facial image. After that, the facial expression classification information is input into the Multi-Task Cascaded Convolutional Neural Network model to extract the facial expression features in an end-to-end manner. Then, after removing the redundant information, the generation of data labels for facial emotion recognition of the existing data is carried out. After the experiment, the results show that the recognition accuracy rate of using Faster R-CNN expressions is above 90% (Wang, 2023).

3.2 Disease Auxiliary Diagnosis

The application of face recognition technology in the aspect of disease auxiliary diagnosis is also continuously increasing. In this aspect, it mainly utilizes the accuracy of face recognition technology in recognizing regular features. The neural network will learn the facial expressions and facial features of each patient with different diseases, and then use face recognition technology to make a preliminary diagnosis of whether an unknown patient is ill. This diagnosis can assist doctors in evaluating the patient's condition. In the aspect of auxiliary diagnosis of depression, Li combines the Single Temporal Network (STNet) and the Full Temporal Network (FTNet). STNet is composed of a spatial convolution network, a contour capture network, and a temporal attention mechanism connecting the temporal backbone network. The spatial convolution network adopts the VGG 16 architecture and is composed of 5 spatio-temporal convolution blocks. The contour capture network is composed of 5 contour capture blocks, and the temporal backbone network can be served by the Long-Short Term Memory (LSTM) temporal model. The full temporal domain network is served by EfficientNet V2, with the first three layers connected by Fused-MBConv and the last three layers connected by MBConv. Then, the feature vectors of size 1000 generated by STNet and FTNet are concatenated into a feature vector of size 2000 and input into the fully connected network to make the final decision. Use the Cross Entropy Loss (CE Loss) as the loss function for training. The final result shows that the accuracy rate reaches 85.1% (Li, 2024). In the aspect of auxiliary diagnosis of Noonan syndrome, Noonan syndrome is a rare genetic syndrome caused by gene mutations that result in abnormalities in the RAS-MAPK pathway. Noonan syndrome has unique facial features, mostly manifested as a high forehead, wide eye spacing, epicanthus, ptosis, and horizontal or downward-sloping eye fissures. Using its more distinct features, the system can be recognized by convolutional neural networks such as AlexNet, Google Inception Net, VGGNet, and ResNet. However, so far, the accuracy rate of ResNet is the highest, with an error rate of only 3.75%, and it has a very good development space (Lin, 2022).

3.3 Micro-expression Lie Detection

Facial recognition also has relevant research in the police field. When people lie, they will have higher cognitive load and deliberate self-control and other psychological activities, which will lead to changes in the liar's facial micro-expressions, posture movements, and eye movements. This is one of the main principles of micro-expression lie detection.

Xiao Ziting proposed a multimodal lie detection method based on DG-MIFLD. That extracts the spatial features of eye fixation points, pupil diameter size, electroencephalogram signals and expressions in eye movement data through the spatial feature extraction module. Then it fuses local features into global features to extract features from the data and at the same time uses the temporal feature extraction module to extract the temporal features of each moment. Finally, the data is directly mapped to the final classification result, which can effectively learn the spatial and temporal information of the data and further improve the accuracy of lie recognition. The DG-MIFLD model using the Swish activation function can achieve the highest accuracy of 95.14% (Xiao, 2024).

Yu proposed a multi-label AU recognition model based on 3D-Net, using a two-channel convolutional neural network to extract the spatiotemporal features of the keyframes of micro-expressions. The system uses the method based on the optical flow method and LSTM to detect the micro-expression intervals in the video, and then it uses the multi-label AU recognition algorithm based on 3D-Net for micro-expression action unit recognition. The frequency of the action units in the video constitutes a feature vector, and a lie recognition algorithm based on micro-expression action units is proposed. A convolutional neural network is used to extract adjacent features and perform lie classification. Through experiments on the Real-lifeTrial Data dataset, the final accuracy is as high as 88.4% (Yu, 2023).

4 EXISTING LIMITATIONS AND FUTURE PROSPECTS

4.1 Disease Auxiliary Diagnosis

The facial data on the network is very scarce and difficult to collect. As facial data is the private data of each person, it is obviously unrealistic to conduct large-scale collection to expand the public database. After the collection is completed, it is necessary to pay a large cost for data protection to prevent the leakage of facial data and cause unnecessary troubles. For such problems, the data protection technology can be improved to enhance the data protection ability, so that the public can provide facial data with confidence. At the same time, the law should be strengthened and those who steal data should be severely punished. Strengthen publicity, call on the public to provide facial data, and set corresponding rewards.

Through improving the data enhancement technology. In the case of insufficient data, the data enhancement technology will be used to optimize and complete the incomplete and unbalanced data. However, the content of the current data enhancement technology is relatively scarce, often rotating, horizontally flipping, and zooming. In the future, more modes can be developed to enrich the database.

4.2 The Model is not Comprehensive

Facial recognition technology can be implemented through a variety of different neural networks, but these neural networks each have their own characteristics, but it is difficult to have both operating speed, accuracy and miniaturization. In future work, it should be attempted to propose new models or combine different existing models. Use a variety of different neural networks to handle a problem, so that each neural network is used in the most suitable position to make up for the deficiencies of another neural network. At the same time, it is also necessary to improve the existing neural networks and models to make these models more convenient and accurate.

4.3 Image Recognition Shifting to Video Recognition

Nowadays, compared with viewing a single image, watching a video is obviously more comprehensive, accurate, and stable in understanding emotions. However, this direction faces two problems. One is that currently, only a few datasets have dynamic sequences, and most are mainly static. The other is that how to fully utilize the dynamic expression sequences for facial expression recognition is also a difficulty in development. In this regard, more video face recognition algorithms need to be studied, and video face materials and databases need to be collected. Using deep learning, video face features are extracted for learning and classification to achieve the purpose of video recognition.

5 CONCLUSIONS

This article mainly summarizes the work required for facial recognition, the models and the results used by other scholars in three different fields: emotion recognition, disease-assisted diagnosis, and micro-expression polygraphy. The technical defects and safety problems of the current technology are analyzed, and the future development direction is proposed. The main existing problems also come mainly from the lack of databases and incomplete models. At the same time, this paper proposes that the lack of databases can be compensated for by enacting laws and improving data augmentation techniques. Proposing a new model and combining different existing models are also two ways to solve the problem of incomplete models. Finally, this paper proposes that static face recognition should be transformed into a more comprehensive and stable dynamic face recognition so that dynamic face recognition technology will also be widely used, which needs to be completed in the future.

REFERENCES

Chen, R. (2023). Review on the face recognition based on deep learning. *Applied and Computational Engineering, 22(1), 195‑199.

Wang, X. (2023). Analysis of face recognition technology based on deep learning. *Applied and Computational Engineering, 22, 258-264.

Shepley, A. J. (2019). Deep learning for face recognition: a critical analysis. *arXiv preprint arXiv:1907.12739*.

Huang, X. (2024). Research on facial expression recognition algorithm based on deep learning (Master's thesis, University of Electronic Science and Technology of China).

Wang, C., Huang, R., Sun, Y., Yang, B., & Sun, L. (2020). Fusion of CLBP and geometric significant features for facial emotion recognition. *Intelligent Computers and Applications* (05), 52-55.

Wang, X. (2023). Facial emotion recognition method based on Faster R-CNN. *Information and Computer (Theoretical Edition) (21), 148-150.

Li, X. (2024). Depression recognition based on facial expression changes in emotional stimulation experiments (Master's thesis, Qilu University of Technology).

Lin, M. (2022). Research on the construction of a facial recognition-assisted diagnosis model for Noonan syndrome based on integrated facial features (Master's thesis, Shantou University).

Xiao, Z. (2024). Research on lie detection technology based on multimodal information (Master's thesis, Jiangxi Normal University).

Yu, C. (2023). Lie detection based on micro-expressions and eye movements (Master's thesis, Southeast University).