Proceedings of the 3rd International Conference on Signal Processing and Machine Learning
DOI: 10.54254/2755-2721/6/20230894

Research of the occluded face recognition methods

Yunhao Cong[1], Shiqian Miao[2], Chaorong Zhang[3,4]

[1]School of information science and Engineering, Baoding University of Technology, Baoding, China
[2]College of Marine Electrical Engineering, Dalian Maritime University, Dalian, China
[3]School of Electronic and information Engineering, South China University of Technology, Guangzhou, China

[4]201830260361@mail.scut.edu.cn

Abstract. In face recognition, deep learning has always been a technology and problem that people need to improve. With the development of time, deep learning has the characteristics of self-learning ability, strong expression ability and better stability, and it is an effective solution to face recognition. However, in the sheltered environment, deep learning still faces many challenges. Different methods are used to expand the new deep learning methods, which can effectively solve the face recognition features in occlusion environment. This paper compares and classifies the related algorithms of deep learning and the traditional algorithms of deep learning. In particular, the network structure and construction principle of some classical algorithms are reviewed. Finally, discusses and summarizes their future development trends and directions. This study has certain reference significance for related scholars.

Keywords: face recognition, occluded, deep learning, feature extraction.

1. Introduction

Face recognition is a remained has important implications for human life and development of research. Now since the accuracy to the requirement of face recognition to both accurate and must meet certain quickening the development of face recognition is now on a sheltered side of the hottest research direction. Because due to the outbreak, people put on a mask, this makes it hard for some of the face recognition system in the case of face feature received keep out accurate analysis of facial features, the depth of the spawned a series of new learning algorithm to deal with the different shade environment, special shade environment includes the obstructions such as masks, the change of light and shadow, the change of the Angle, etc., can be in these environments accurately identify faces, for human production and living, and a series of activities is of important significance [1]. Including to authenticate by face recognition, due to wear some jewellery or lighting changes, will make testing environment full of shade, it needs to be through deep learning algorithm to extract the accurate facial feature, to face recognition to satisfy people's needs. At present due to the review material have not enough comprehensive correlation, this paper for face recognition on a sheltered side of the further research and generalization.

2. Traditional methods for face recognition

Face recognition is divided into two aspects: common face recognition and a certain shade of face recognition. Both have a common goal, that is in accurate detection of human face at the same time,

DOI: 10.54254/2755-2721/6/20230894

both satisfy the quickness and can satisfy the robustness. There are also essential differences between the two, that is, the integrity of the face sample features to be detected or whether there are occlusions. Ordinary human face recognition is through some for facial features extraction, comparison, analysis and so on process face to face to be detected and the data set model, calculated the accurate answer. In the early stage of traditional face recognition, mainly through edge detection algorithm, feature extraction and statistics to achieve the purpose, but due to the need for a large number of data sets and a long time of calculation, resulting in the early face recognition speed is slow, poor accuracy and other problems [2].

Traditional ways of face recognition have eigenfaces and fisher faces, etc. The Eigenfaces method is the PCA(Principal Component Analysis) supervised learning was applied to the face recognition, this method is through the use of the PCA supervised learning of the face image as a vector, and then by using the method of statistical characteristic, find out the characteristics of the covariance matrix of the human face image data vector, to represent the face image. The Fisher faces method is considered to be improvements to the eigenfaces method, the application of the LDA **(**linear discriminant analysis) supervised learning. The LDA supervised learning has in common with the PCA supervised learning, they are both on the matrix, but the LDA supervised learning has higher accuracy for large data sets, and the facial expression and the influence of illumination condition is very small, can improve the robustness [3].

Due to the outbreak of the COVID-19 in recent years, people are wearing masks and other facial coverings, this makes for a certain mask face recognition has become more and more popular. For the classification of occlusion there are light and shadow, occlusion, natural occlusion, random occlusion, etc., these occlusion will make the complete face features on some features of the loss or change. Therefore, for the obscured face recognition, it is necessary to extract the features of the face, enhance the features of the visible area of the face and the loss function training methods to carry out.

Tradition with a mask or other facial coverings on face recognition methods have FAN, CMS - RCNN, etc. The Facenity-Net algorithm uses the feature extraction method from the local to the whole. Firstly, feature extraction is carried out in the uncovering region, and then the extraction range is continuously expanded until the result is obtained. Its advantage is that when some unimportant positions are occluded, the recognition result can still be obtained very accurately. The FAN algorithm uses the method of feature enhancement and the Anchor mechanism to enhance the features of the face region including the area with occlusion and the area without occlusion at the same time, so as to compare with the data set (Figure 1). However, due to the possibility of enlarging the features of the occlusion. It would lead the identification results problematic.

DOI: 10.54254/2755-2721/6/20230894

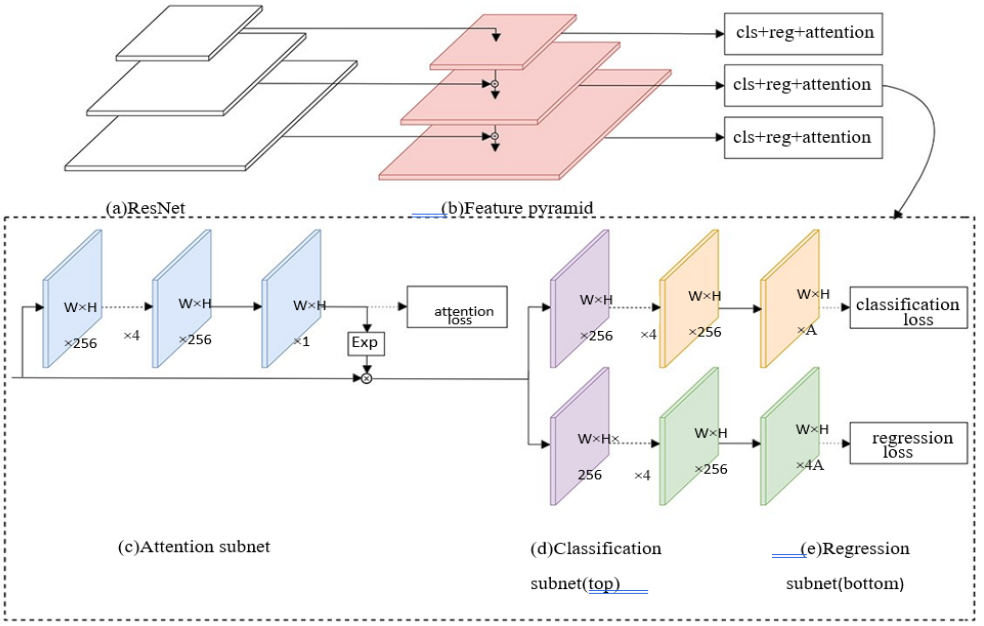

Figure 1. FANNet work model [3].

The CMS-RCNN (Contextual Multi-Scale Region-Based CNN) algorithm carries out feature analysis through the fusion of the context of face features, which improves the accuracy of the analysis, but the speed is greatly reduced. The SqueezeNet algorithm uses a modular convolution called the Fire module (Figure 2). The SqueezeNet algorithm by reducing the calculation parameters, so as to make the convolution layer decreases, and information available on the face of increased, did little model but the performance is not poor [4] .

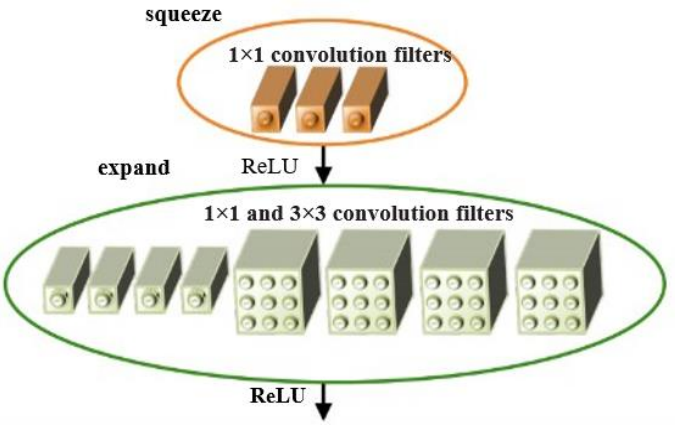

Figure 2. Fire module of SqueezeNet [4].

The MobileNet algorithm to further improve the operation speed, the algorithm by a 3 x3 standard convolution, the depth of separable convolution computation, such as pooling operation, won the convolution of a 1 x 1, to reduce the complexity of the model, the algorithm greatly accelerate the speed, accuracy but received a certain influence.

Face recognition algorithm on a sheltered side of the overall, for there are still focusing on the problem, the accuracy and efficiency but based on the deep learning algorithm for the covered face recognition compared to traditional methods still have certain advantages, due to the sudden appearance of all sorts of problems, a new shade scenarios and model is put forward, which has greatly affected the

DOI: 10.54254/2755-2721/6/20230894

depth of the existing learning algorithms, need for model and algorithm to optimize constantly to adapt to changing circumstances.

3. Deep learning-based methods for face recognition

In order to combine deep learning with traditional methods, the purpose is to improve robustness and extract effective features [5]. Among them, deep learning has the characteristics of self-learning ability, strong expression ability and better stability. At the same time, it can also integrate feature learning into the process of modeling, thus reducing the incompleteness caused by artificially designing features. Therefore, it is also a good method to solve the shallow features of traditional methods. Deep learning has gradually become the application of face recognition in occlusion environment.

With the development of the times, Faceness-Net method can detect faces with large posture changes, and has advantages in performance, running speed, recall rate, average accuracy and so on. However, when the occluded area is large or the face picture is not clear enough, it will affect the scoring and recognition effect. CMS-RCNN provides a method to integrate global and local context information, which can integrate multi-scale feature information into RPN proposals generation and RoI pooling plus detection, which is more conducive to small-scale face detection [6]. At the same time, we pay attention to the features of the face region and the face information, and fuse the features on the multi-layer feature map to form a long feature vector for subsequent classification. This kind of method has a high recognition accuracy, but the disadvantage is that the distribution and integration of feature weights of each part are difficult and inefficient, which will affect the stability of the model.

Deep learning requires a large number of complex parameters and neural network structures. Therefore, Chan et al. put forward a learning network method using PCANet, so that subspace can be extradited to deep learning (Figure 3). This method is prone to the problems of information loss and low recognition accuracy. Combining with tradition, a new method is proposed based on convolutional neural network (CNN) [7].

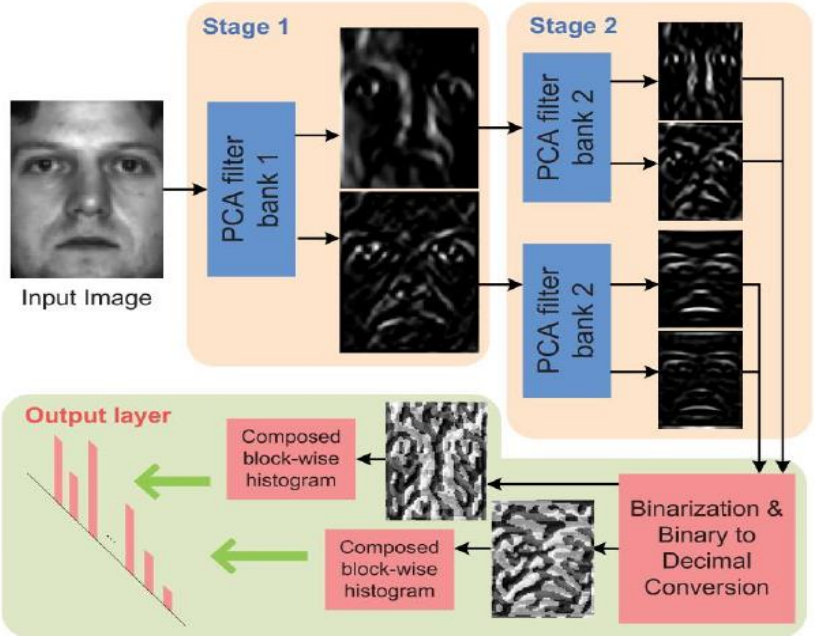

Figure 3. PCANet feature extraction process [7].

First of all, PCA and PCANet both extract information from two-dimensional images. Therefore, MPCANet, a face occlusion recognition algorithm based on multi-layer feature fusion, is extended by PCA and PCANet, which includes input, output, convolution and classification layers (Figure 4). Its advantage lies in that it can completely extract all features of occluded faces for recognition, which can save more cost and time. At the same time, MPCANet can use multi-layer features to fuse to extract the basic features of images, which is suitable for high-dimensional data. It makes use of PCANet's

DOI: 10.54254/2755-2721/6/20230894

redundancy and high classification efficiency, and combines the spatial pyramid pooling (SPP) to calculate a multi-scale method (MMPCANet) [8].

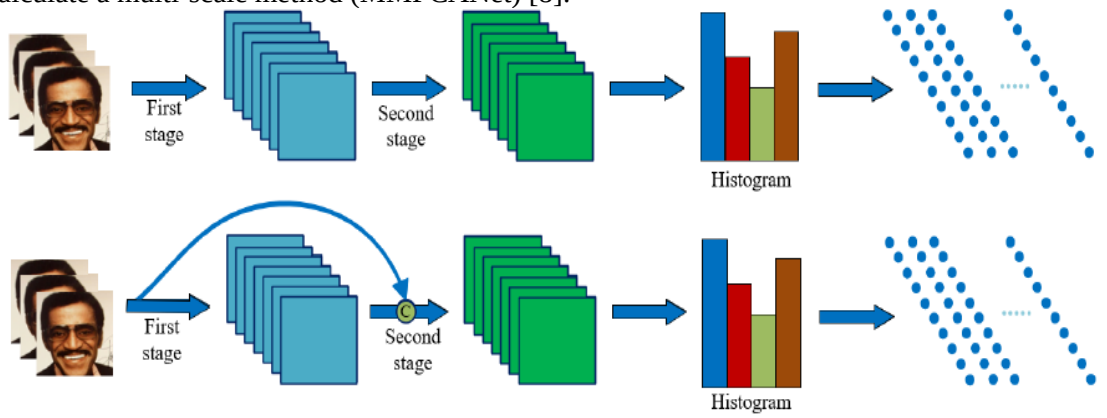

Figure 4 Simple process of PCANet and MPCANet model.

Secondly, the features of the original image are connected with the output features of the first layer, and then the connected result is used as the input of the second layer, which is convenient to use more image feature information. Use RF classifier for identification and classification. Finally, compared with the original PCANet, this algorithm can extract more low-level features of images, and does not need a large number of training samples, but it takes a long time, but it is beneficial to the occlusion recognition effect of human faces.

4. Evaluation metrics

In the process of performing face recognition, some common evaluation metrics are needed to assess how good a face recognition model is. Common evaluation metrics include recall rate, false recognition rate, accuracy, precision, and ROC curve. The task of face recognition is also essentially a binary classification task. In a binary classification task, samples are divided into positive and negative classes. Based on the actual class of the sample and the class predicted by the classifier, all the predictions can be further classified into four classes, called TP(True Positive), TN(True Negative), FP(False Positive), and FN(False Negative).

The Accuracy is the percentage of all samples that can be correctly classified. This metric is often used to indicate the global performance of a model. It is defined in equation 1.

$$Accuracy = (TP + TN)/(TP + TN + FP + FN) \tag{1}$$

The ROC Curve (Receiver Operating Characteristic Curve) is a plot with FPR as the horizontal coordinate and TPR as vertical coordinate, where the defining equations of FPR and TPR are shown in Eq. 2 and Eq. 3.

$$FPR = FP/(TN + FP) \tag{2}$$

$$TPR = TP/(TP + FN) \tag{3}$$

For classifiers, the better the performance, the closer the ROC curve of the classifier is to the upper left corner of the graph. In face recognition tasks, the ROC Curve is usually used to evaluate the false recognition level of the model.

AUC (Area under Curve) is defined as the area of the graph under the ROC curve. For classifiers, if the value of AUC is larger, the performance of the classifier will be better. Unlike ROC curves, AUC is usually a numerical value. Therefore, AUC can be used as a more intuitive judging metric.

5. Performance comparison

For the traditional methods of face recognition with occlusion, the more commonly used ones are PCA and SVM (Support Vector Machines). PCA is a subspace-based recognition method. It is less

DOI: 10.54254/2755-2721/6/20230894

computationally intensive, less complex, and can achieve more efficient classification of faces and non-faces. However, this method has low accuracy and no better robustness for recognizing faces with occlusions. And SVM, as a representative one of traditional face recognition methods, is widely used in the field of face recognition, but the robustness of face detection with occlusion still needs further improvement [9].

Deep learning methods enable researchers to improve recognition accuracy by recovering faces with occlusions and then passing the recovered faces through models for face recognition. As for the method of robust subspace regression proposed by the researchers for the problem of low robustness of the original recognition method, this method is more effective in recognizing faces with occlusions compared to the traditional methods [10]. However, for this method, it is still necessary to implement robust structured error coding for occluded faces to better solve the problem. Using deep learning to recognize obscured faces by suitable network construction can improve the accuracy of recognition. However, deep learning networks often have a large number of parameters and a complex network structure. This leads to an increase in the amount of computation required to perform training, which in turn makes training more difficult. In addition, the data set required for training still needs to be further supplemented.

We have summarized the advantages and disadvantages of the commonly used methods for face recognition with occlusion as well as the application areas, as shown in Table 1 below. Through comparison, it can be found that the deep learning-based occlusion face detection algorithm has better detection effect on the face detection problem of natural scenes compared with the traditional one, but there are also problems of training difficulty and poor stability in the training process.

Table 1. Comparison of different face recognition methods.

Recognition methods	Advantages	Disadvantages	Application scenarios
SVM	works well for small samples, nonlinear and high-dimensional problems	requires large storage space	expression and face recognition
Robust subspace regression	improves detection efficiency by representing test samples with partial samples	insufficient consideration of the structure of occlusion	partial occlusion, micro-expression change perception, face recognition
Faceness-Net	better recognition performance for face samples with large pose changes	recognition accuracy decreases when face clarity is low or when the occlusion area is large	face samples with high resolution
CMS-RCNN	the contextual information of the local face is utilized and fused with the global face information [11].	there are difficulties in assigning and integrating the weights of different features	partial occlusion
SqueezeNet	significantly reduces the number of parameters in the network model [12].	high requirement for balance of datasets	face recognition on mobile and embedded devices

DOI: 10.54254/2755-2721/6/20230894

6. Conclusion

By comparing traditional methods and deep learning methods in face recognition, we find that with the development of deep learning technology, it has a wider application prospect in face recognition applications with occlusion. The current traditional face recognition methods are already difficult to meet the increasingly diverse task scenarios in real life, but are only applicable to small sample face recognition tasks. It is easy to see that deep learning methods will become the mainstream of face recognition methods. Through the design and targeted modification of the model structure, researchers can make the model perform powerfully on specific tasks.

For future research directions, the following aspects can be taken into account: first, to improve the accuracy of detection of occlusion locations. At present, the computer still cannot know the exact location of the occlusion. Second, to produce data sets of different occlusions. The existing public dataset of occlusion is relatively single, and the collection method is also subject to human control, which still has a gap with the actual occlusion environment. Third, improve the detection capability of small samples. The current recognition of faces by deep learning methods often requires a large number of samples, and cannot rely on only a small number of images to summarize the overall characteristics of the obscured faces.

References

[1] Wang, M. , Deng, W. Deep face recognition: a survey. 2018 arXiv.
[2] Gao, X. Research on Robust Face Recognition with Occlusion and Complex Illumination. 2017 Dot. Dis., Shandong Univ. Sci. and Tech.
[3] Wei, D. Research and application of face recognition under variable illumination and occlusion. 2012 Dot. Dis., Nanjing Univ. Tech.
[4] Xu, X. , Liu, T. , Tian, G. , Yu, W. , Xiao, D. , & Liang, X. Review of Occlusion Face Recognition Methods. 2021 Com. Eng. Appl., 57(17), 15.
[5] Liu, R. , Xu, C. , Zhou T. Review of Face Recognition Methods under Local Occlusion. 2022 Jiangsu Ocean Univ.
[6] Pan, C. Research and Implementation of Local Occlusion Face Recognition Based on Deep Learning. 2022, Guizhou Univ.
[7] Xu X. ,Tao, J. ,Wu G. . Face Recognition with Masking Based on Convolutional Neural Network. 2019 J. Jianghan Univ.47(3):246-251.
[8] Wang, W. , Tang, Y. , Zhang, J. , etc. Face Recognition Algorithm Based on Convolution Neural Network Feature Fusion. 2020 Com. Dig. Eng., (1):88-92.
[9] Hu M. Face Recognition System Based on PCA and SVM. 2017, Comp. Era (12), 64-67+71.
[10] Wright, John, Yang, Allen, Y. , Ganesh, & Arvind, et al. Robust face recognition via sparse representation. 2009 IEEE Trans. Patt. Anal. Mach. Intel. 421(24):2436-2452
[11] Zhu, C. , Zheng, Y. , Luu, K. , & Savvides, M.. Cms-rcnn: contextual multi-scale region-based cnn for unconstrained face detection. 2016, 10.48550/arXiv.1606.05413
[12] Iandola, F. N. , Han, S., SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and <0.5MB model size. 2016 10.48550/arXiv.1602.07360.