A Review of Face Recognition based on Deep Learning

Run Zhang

Zhejiang University, China

Keywords: Deep Learning; Face Recognition; Convolutional Neural Network; Biometric Recognition.

Abstract: Face recognition based on deep learning, which has already become a hot research topic in the field of biometric recognition at present, was reviewed. Firstly, face recognition and the basic structure of deep learning were introduced. Then, the current international and domestic research status quo and application of the technology were summarized, such as face recognition method based on convolutional neural network(CNN), deep nonlinear face shape extraction method, robust modeling of face pose based on deep learning, fully automatic face recognition in the constrained environment, face recognition based on deep learning under video surveillance, low resolution face recognition based on deep learning, and other face information recognition based on deep learning. Finally, a general analysis was made on the existing problems and future development trend of face recognition in the application of deep learning.

1. Introduction

Face is highly non-rigid, and there are a lot of details reflecting individual differences. Face recognition is the process of finding the matching face by comparing the face images detected from static images or dynamic videos with the face images in the database. It is usually used for the purpose of identification [1]. It is a subject in the field of biometric recognition.

Face recognition research began in the 1950s. As an important biometric recognition technology, it has the advantages of direct, friendly, convenient and interactive, and has been concerned by researchers [2]. Early in 1966, Bledsoe *et.al.* studied on the human facial recognition based on the pattern recognition and created a technology named “man-machine facial recognition” as the preliminary development of the human facial recognition technology [3]. And there was an available facial detection system in 1997 that can find a certain face among the crowd by Malsburg *et.al.* [4]. In contemporary society, Face recognition technology has many practical applications in security and financial payment, such as video surveillance, intelligent payment, access control and so on. It is the most popular research direction in machine learning and pattern recognition.

At present, face recognition methods based on machine vision have achieved fruitful results. In this study, we need to consider the intra-class changes caused by facial expression, posture, age, location and occlusion, and the inter-class changes caused by different identities such as illumination and background [5-6]. The distribution of these two changes is highly complex and non-linear. Traditional face recognition methods based on shallow learning often fail to achieve the desired results for the complex distribution of intra-class and inter-class changes and the recognition of non-linear face data [7]. Deep learning simulates the cognitive learning of human visual perceptual nervous system, and can obtain more expressive high-level features, which can be used to solve the problem of intra-class and inter-class change distribution in face recognition.

This paper summarizes the face recognition technology based on deep learning, and lists the basic model structure of deep learning. This paper summarizes the research status of this technology at home and abroad, introduces the typical technology of deep learning for face recognition and the commonly used large-scale face recognition database, and finally analyses and prospects the problems and development trend of deep learning-based face recognition.

2. Deep Learning

In recent years, with the continuous development of machine learning, in-depth learning as a new research direction has attracted extensive attention in the field of artificial intelligence. Algorithmic speed, system performance and ease of use are the bottlenecks that restrict the popularization and application of machine learning. The structure of typical in-depth learning models can be roughly divided into three categories:

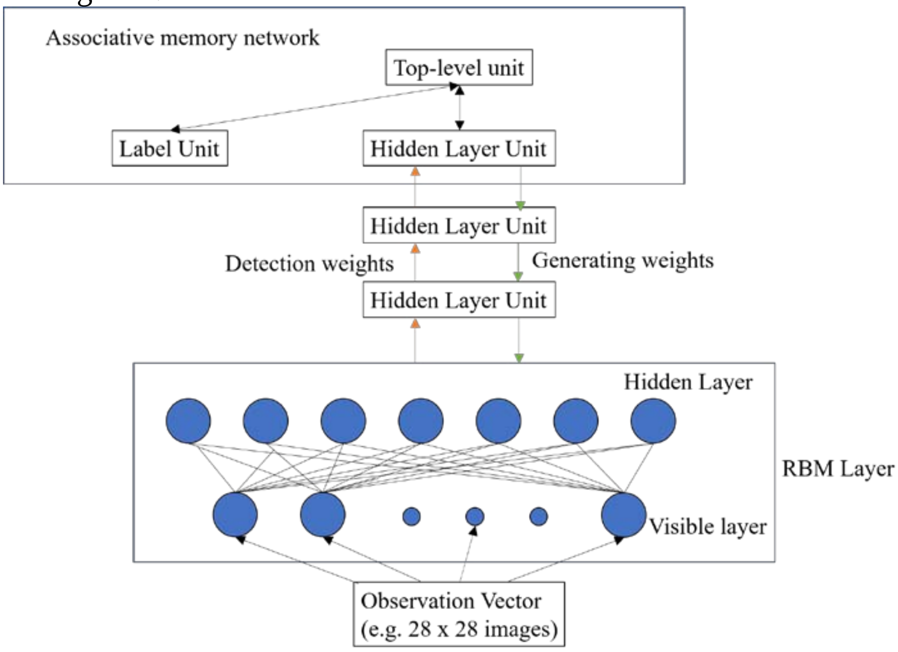

Fig.1 DBN framework

(1) Generative depth structure. Deep belief networks (DBN) [8] is a widely studied deep learning structure. As shown in Figure 1, DBN consists of a series of restricted Boltzmann machine (RBM) units. RBM is a random undirected graph model. Its visible layer and hidden layer units are interconnected with each other, and there is no connection in the layer. The hidden layer units can obtain high-order correlation of input visual units. In order to obtain generative weights, unsupervised greed layer by layer is adopted in pre-training. DBN can obtain the joint probability distribution of observation data and labels, which facilitates the estimation of prior probability and posterior probability, while the discriminant model can only estimate posterior probability [9].

DBN is widely used, flexible and easy to expand, but its input is a simple one-dimensional vectorization of image matrix, without considering the two-dimensional structure of the image. Convolutional deep belief networks (CDBN) [10-12] utilizes the spatial relationship of neighborhood pixels and transforms them into high-dimensional images through transformation invariance. In 2011, Liu *et.al.* [13] proposed a new semi-supervised learning algorithm, i.e. discriminant depth confidence network. Discriminative deep belief networks (DDBN), which uses a new deep architecture, combines the abstraction ability of DBN and the discriminatory ability of exponential loss function, and applies it to visual data analysis.

(2) Distinctive depth structure. Convolutional neural network (CNN) [14-16] is the first truly successful learning algorithm for training multi-layer network structure, which is designed and trained by BP algorithm. CNN is an effective learning method adapted to two-dimensional face image recognition scenarios. It has been widely used to solve face recognition problems in literature. It is mainly used to recognize displacement, scaling and other forms of distorted invariant two-dimensional images. Because the feature detection layer of CNN learns from training data, it avoids explicit feature extraction and implicitly learns from training data. Moreover, because the weights of the neurons on the same feature mapping surface are the same, the network can learn in parallel, which is also the advantage of CNN relative to the network where the neurons are connected to each other. Its layout is closer to the actual biological neural network. Weight sharing reduces the

complexity of the network. Especially the image of multi-dimensional input vectors can be directly input into the network, avoiding the complexity of data reconstruction in feature extraction and classification process. Figure 2 is an example of a two-layer convolution neural network for image classification.

(3) Mixed structure. The learning process of hybrid structure [17-18] includes generative part and discriminatory part, which are usually solved by optimization and discriminatory depth network model. This process of differentiated network optimization usually adds a top-level variable to estimate the parameters of any depth generation model or unsupervised depth network. BP algorithm is used to optimize the weights of DBN. The initial weights are obtained in the pre-training of RBM and DBN, not randomly generated. Generally, the performance of such a network is better than that of the network trained solely by BP algorithm, which speeds up the training and convergence time compared with the feedforward neural network.

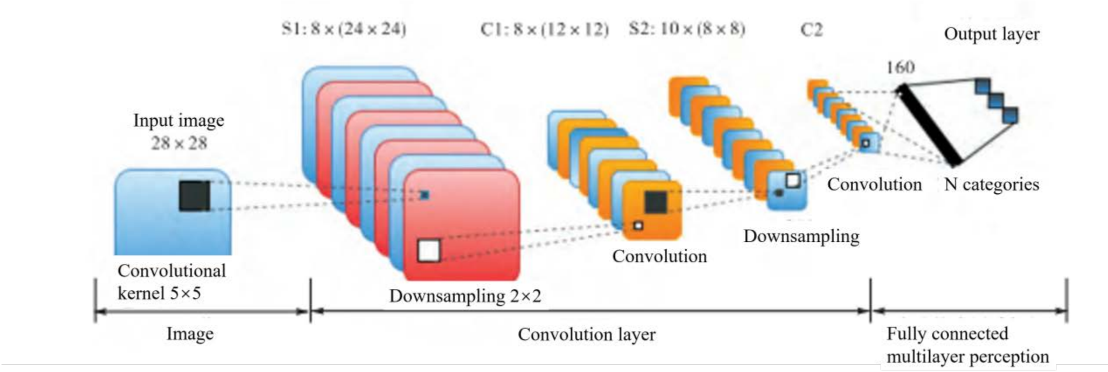

Fig.2 Two-layers CNN architecture of image classification

3. Face Recognition Technology based on Deep Learning

There are many traditional face recognition methods, such as active shape model.

(Active shape model, ASM) [19] and Active appearance models (AAM) [20]. Local-based methods such as local descriptor Gabor and local binary pattern (LBP) are used for recognition. There are also global-based methods, including classical face recognition algorithms, such as Eigenface [21], linear discriminant analysis (LDA) and local preserving projection (LPP) and other popular learning algorithms. 3D face recognition is also a new direction. However, due to the influence of illumination, gesture and expression changes, occlusion, massive data and other factors, the accuracy of traditional face recognition methods is limited due to their own limitations.

3.1 Overseas Research Status

In 2012, Leonid Miller's team took the lead in applying in-depth learning to face recognition in the LFW (labeled face in the wild) database. They adopted unsupervised feature learning method and achieved 87% recognition rate, which was still far behind the best face recognition algorithm at that time. The recognition rate of the classical face recognition algorithm Eigenface [22] in LFW is only 60%, while the recognition rate of the latest deep learning algorithm is 99.47%, which even exceeds the recognition rate of human eyes (99.25%). Among the above algorithms, FaceNet has the highest accuracy, which has reached 99.63% in LFW database, surpassing the recognition result of human eyes.

The LFW [23] database is compiled by the Computer Vision Laboratory of the University of Massachusetts, Amsterdam, USA. It is used to study face recognition in unconstrained situations. It has become a standard reference for academia to evaluate recognition performance. LFW project is the driving force of face recognition technology from the initial stage to the prototype system stage. The data from LFW test has become the reference standard for paper verification. This database greatly promotes the progress of face recognition technology in depth learning. Table 1 and Table 2 show the recognition rates of in-depth learning and other methods in LFW face database, respectively.

Tab.1 Recognition rate of deep learning in the LFW face database

Method	Key Points of Registration	Number of Training Samples	Recognition Rate%
Document [29]	3	Unsupervised	87
Document [23]	5	87,628	92.52
Deepface [30]	73	7 Million	97.35
Deepface [31]	73	500 Million	98.40
Deepid [32]	5	202,599	97.45
Deepid2 [33]	18	202,599	99.15
Deepid2+ [32]	18	450,000	99.47

Tab.2 Recognition rate of each method in the LFW face database

Method	Rank-1/%
COST-S1[34]	3.0
COST-S1+S2[34]	56.7
DeepFace [23]	66.5
DeepFace [30]	64.9
DeepID2 [33]	91.1
DeepID2+ [32]	950

Note: Rank-1 is the first time to find the correct face recognition rate.

3.2 Domestic Research Status

In recent years, domestic research institutes have also conducted a lot of in-depth research in the field of face recognition. At the International Conference on Computer Vision and Pattern Recognition in 2014, Taigman et al. from the Facebook Laboratory of Artificial Intelligence and a team from the Chinese University of Hong Kong, China. Under the condition of allowing the use of external label data and unrestricted test conditions, the average classification accuracy of 97.35% and 97.45% are obtained by using convolutional neural networks, respectively. The research team led by Tang Xiaoou and Wang Xiaogang of the Chinese University of Hong Kong has developed an automatic face recognition system with high accuracy. The recognition rate of the system in LFW database is as high as 99.15%.

On the most authoritative large-scale face recognition evaluation database FRVT 2006 (the face recognition vendor test, FRVT), the test completed by the research group of Tsinghua University such as Ding Xiaoqing reached the international leading level. Professor He Xiaofei's research group [24-25] of Zhejiang University has carried out a series of research on machine learning in the era of big data. Zhou Zhi of Nanjing University Professor Hua's research team has carried out machine learning and data mining long-term research, and put forward many new algorithms. In Xi'an University of Electronic Science and Technology, Professor Xue Chen Bo [26-27] has conducted in-depth research and application of deep learning algorithm based on hierarchical super-complete dictionary sparse representation. These studies show that domestic scholars have paid much attention to deep learning and face recognition related fields.

4. Typical Application of Deep Learning in Face Recognition

There are seven typical applications of deep learning in face recognition: face recognition based on convolutional neural network (CNN) and deep non-linear face shape extraction, robust modeling of face pose based on deep learning, automatic face recognition in constrained environment and face recognition under video surveillance based on deep learning, low-resolution face recognition based on deep learning and other face-related information recognition based on deep learning are shown in Figure 3.

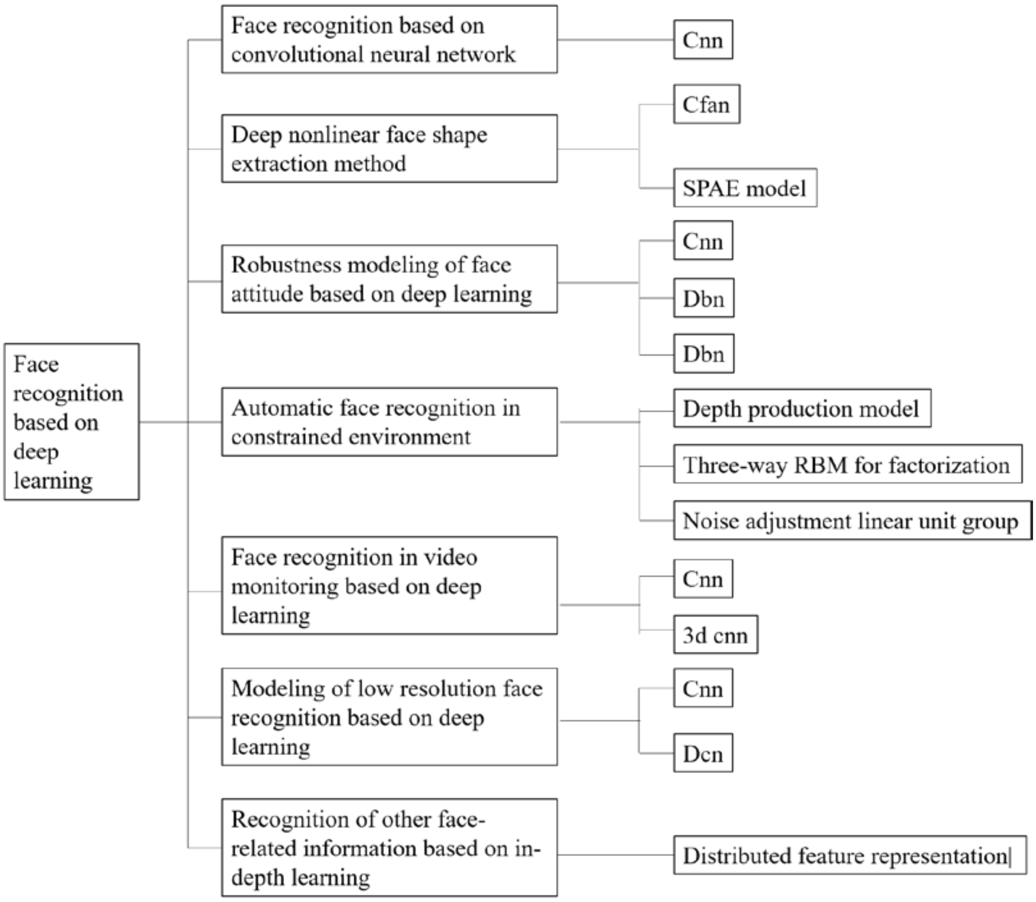

Fig.3 Face recognition methods based on deep learning

4.1 Face Recognition Method based on Convolutional Neural Network (CNN)

Convolutional neural network (CNN) is a machine learning model under deep supervised learning. It can mine local features of data, extract global training features and classify them. Its weight sharing structure network makes it more like biological neural network, and has been successfully applied in various fields of pattern recognition. CNN makes full use of the local features of data itself by combining local perception region, shared weights and downsampling in space or time to optimize the model structure and ensure certain displacement invariance [28].

Using CNN model, the correctness rates of the DeepID project of the Chinese University of Hong Kong [29] and the DeepFace project of Facebook on LFW database are 97.45% and 97.35%, respectively, which are slightly lower than that of human visual recognition 97.5% [30]. After a breakthrough, the DeepID2 project of the Chinese University of Hong Kong [31] raised the recognition rate to 99.15%. DeepID2 minimizes intra-class variations by learning non-linear feature transformations, while keeping the distance between face images of different identities constant, which surpasses the recognition rate of all leading deep learning and non-deep learning algorithms [32] on LFW database and human recognition rate in this database [33].

4.2 Depth Nonlinear Face Shape Extraction Method

Active Shape Model (ASM) [19] and Active Appearance Model (AAM) [20-21] are classical face alignment methods. They use linear principal component analysis to model the changes of face shape and texture, and optimize the model parameters to adapt to test face images. Due to the difficulty of describing complex facial shape and texture changes in linear model, the effect is not good under large posture, exaggerated expression, intense illumination changes and partial occlusion. The latest development of this problem is to predict face shape directly from face texture features by cascading multiple linear regression models.

The main difficulty in face recognition is the high complexity of face morphology and texture, which can further improve the non-linear regression ability of the algorithm to obtain robustness to shape and other changes. Zhang et al. [37] proposed a coarse-to-fine auto-encoders networks (CFAN) method for deep non-linear face shape extraction. CFAN cascades multiple non-linear regression models implemented by stack self-coding networks, each level depicts a partial non-linear mapping from face appearance to face shape. In the process of facial image recognition, deep learning method can not only extract useful facial texture features, but also obtain accurate information of face shape and geometric structure.

4.3 Robust Modeling of Face Attitude based on Deep Learning

The change of face appearance caused by pose change is a complex non-linear change. Although the method of generating virtual image by 3D model can solve the problem of non-linear change among different postures, it is very difficult to restore accurate 3D model from 2D image. In reference [38], a stack progressive self-coding (SPAE) neural network model is proposed to realize the non-linear modeling of attitude change in small-scale data. The change from side image to front image is very complicated, but it is slow and smooth. According to this feature, the modeling from side image to front image is divided into several sub-tasks. Each sub-task is only responsible for transforming the changed attitude to the changed attitude, not directly to the front attitude. Thus, the difficulty of each sub-problem is controlled. A shallow neural network can be used to model effectively, and then a deep neural network can be obtained by overlapping several shallow neural networks together to realize the smooth transformation from side image to front image. This idea of progressive learning divides the deep neural network into several shallow networks to match its model ability with limited data and avoid over-learning problems caused by small data scale. Literature [2] is to solve the problem of pose and resolution in face recognition by deep learning. The application of DBN in face pose processing, such as pose mapping and pose classification, is studied in detail. The change of face rotation caused by the change of pose in face imaging is not a simple linear method. Deep learning is a multi-layer complex neural network, which is a good non-linear generation model. DBN can learn a global mapping from side face image to front face image, but individual details are easily lost. Literature [53] uses local binary pattern to extract facial texture information, and classifies it as input of DBN to provide better recognition of facial features and obtain features of face images under unrestricted conditions. Document [54] adds a regression layer to the top of the RBM stack, which is used for feature extraction and classification under a unified framework of in-depth learning. Depth neural network has strong ability of non-linear modeling. It can be used to model the robustness of face posture by depth learning. These literatures provide a series of solutions. However, in-depth learning requires large-scale supervised multi-pose face image training, which is difficult to collect in practice.

4.4 Automatic Face Recognition in Constrained Environment

Document [34] uses a depth structure neural network composed of noise-adjusted linear units to apply depth learning to target recognition and face verification. DBN in document [35] is very good at dealing with occlusion in face image prediction expression categories, and can perform SIFT descriptors to distinguish different types of scene features. The method based on SIFT classification can get better recognition effect. The disadvantage is that the process of feature calculation is complex, and the process of point matching takes a long time, which has certain limitations. Document [36] applies the three-way RBM model of factorization to multi-face image matching, and the matching performance is better than that of the previous similar generation models. Literature [37] studies how to analyze facial components from locally occluded face images. Faces are divided into several overlapping blocks. Each block is only associated with some hidden nodes, and faces are detected at the block level. The training process is executed by DBN, and then the discriminant is adjusted by logistic regression, and the pixel-sensitive label mapping is calculated. Experiments on

2,239 images selected from three data sets of LFW, BiolD and CUFSF show that the method is effective. It not only has robustness to locally occluded face images, but also provides more abundant facial expression analysis for face synthesis and key point detection. The research shows that the method of deep learning can automatically learn face features in a constrained environment. Compared with the shallow method, it can make complex feature extraction easier and learn some implicit rules and rules in face images.

4.5 Face Recognition in Video Surveillance based on Deep Learning

In intelligent surveillance environment, the recognition of suspicious people is an important use of face recognition. Accurate and fast identification of human identity in video is very important for video search and video surveillance. Although many studies have been carried out, the results are not satisfactory. In recent years, there have been some data sets simulating similar scenarios, for example, the video face database published by NIST (Point and Shoot Face Recognition Challenge, PaSC) in the United States [38]. The first test was conducted at the International Joint Conference on Biometrics 2014 (IJCB'14). The algorithm is a feature probability elastic matching model. The correct recognition rate is only 26% when FAR is 1%.

In 2012, document [39] used continuous frames in video data as input data of convolutional neural network, together with information such as time dimension, to identify human actions. Document [40] integrates the continuous frames in video data, constructs data cube, adjusts RBM, and then fine-tunes the feedback of depth model to achieve the purpose of face recognition. In the re-evaluation of PaSC organized by the 11th International Conference on Automatic Face and Gesture Recognition (FG2015), Shan Shiguang led the team to adopt a deep learning-based technology solution with two core steps. The feature extraction method based on convolutional neural network for each frame of face in video and the ensemble modeling method integrating the features of the convolutional neural network for all video frames in video clips are presented. When FAR is 1%, the validation rate of this method reaches 58% and 59%, respectively, which is obviously higher than 48% and 38% of the second place. The essence of in-depth learning is to learn the multi-level non-linear functional relationship, which enables people to better model visual information, better understand images and videos, and better deal with complex problems such as video face object and behavior recognition. The above research results show that in-depth learning plays a certain role in accurate and fast video face recognition.

4.6 Other Face Information Recognition based on Deep Learning

Face is highly non-rigid, which forms a complex and non-linear face data in imaging. Deep learning is an effective and non-linear neural network model, which is also widely used in other features of the face. In document [41], the feature of convolutional neural network is applied to face expression recognition. Facial expression recognition used to use local features similar to face recognition (such as HOG, SIFT, LBP, LGBP, etc.), and on this basis, classifiers such as support vector machine were used to classify. Because some databases lack facial expression category tags, some models are not for facial expression recognition fine tuning, but for face recognition fine tuning. Literature [42] learns the deformable part to the 3D CNN framework, which can detect structural spatial constraints under specific facial motion parts, and obtains the representation of discriminant parts at the same time. Document [43] introduces depth belief network (DBN) to simulate the data distribution of input image, completes the hierarchical automatic extraction of fatigue features, and finally realizes the fatigue state recognition of video stream image based on time window. The basic idea of document [44] is to extract high-level features of face images through unsupervised greedy layer-by-layer training of deep belief network (DBN). Combining traditional image preprocessing and similarity measurement technology, the purpose of face verification is achieved. Literature [45] proposed a method of facial expression recognition based on depth belief network and multi-layer perception machine.

Deep learning provides distributed feature representation. In the highest hidden layer, each neuron represents an attribute classifier. It can also be used to identify and classify gender [46], race and hair color [47].

5. Analysis and Prospect

Compared with face recognition technology using other machine learning, in-depth learning has a key advantage. However, in-depth learning also has some shortcomings, such as the long time period of training model, the need for continuous iteration to optimize the model, cannot guarantee the global optimal solution, and so on. These also need to be explored in the future.

At present, there are four main problems to be solved in the theoretical aspect of in-depth learning. First, where is the theoretical limit of in-depth learning and whether there is a fixed number of layers. After reaching this level, computers can really realize artificial intelligence. Secondly, how to determine the number of layers of in-depth learning and the number of hidden layer nodes for a certain kind of problem. Thirdly, how to evaluate the characteristics of in-depth learning. Finally, how to improve the gradient descent method to achieve better local and even global optima

Because different recognition systems are suitable for different practical environments, different deep learning structures have different practical requirements. Facial recognition based on in-depth learning has made some progress in various aspects, and the following research directions are as follows:

(1) In feature learning of unlabeled data, the application of large-scale face search has attracted more and more attention from academia and high-tech enterprises. In addition to traditional face recognition based on the particularity of facial biological features, more attention should be paid to the research of automatic tagging technology for unlabeled data, to quickly and accurately search similar faces in large-scale face database.

(2) The resources of the database should be further enriched in both quantity of different faces and tags like expressions for further application. In the case of introducing external training data and unconstrained training protocol, since the face recognition rate of large-scale face database LFW has reached more than 95%, if we continue to expand the face database and design a more reasonable comprehensive evaluation protocol, it will be a very challenging task. As for the different expression detections, the database capacity is much less than that of traditional facial recognition, which need far more effort to spend on.

(3) Face recognition based on in-depth learning should be combined with other methods. When special equipment is allowed to be used, 3D model and depth information can be considered to improve the stability of the system. It may be necessary to redesign the current face design method for the apparent changes caused by posture change, occlusion and expression change, and use multiple local models instead of a whole model to represent [48]. In-depth learning can also be combined with some methods of facial posture or expression correction to carry out relevant exploration. According to the current development trend, in-depth learning will be the future development direction of face recognition. New methods, such as the combination of intensive learning and in-depth learning, have achieved great success [49]. Due to the high computational complexity of in-depth learning, parallel devices such as GPU are needed in the training process, and the training time is relatively long. In the process of testing, because of too many model parameters, it is not convenient to run directly on embedded devices. Using fast learning methods such as extreme learning [50-55] to construct in-depth learning units will be a research direction. Nowadays, deep learning-based algorithms are generally provided through cloud services, such as Face++, Microsoft Project Oxford and so on. These are also the future trends.

References

[1] Wu Jinyi and Zhou Delong. A Review of Face Recognition Methods [J]. Computer Applications, 2009, 26 (9): 3205-3209.

[2] Lin Miaozhen. Research on Face Recognition Based on Deep Learning [D]. Dalian: Dalian University of Technology, 2013.

[3] Boyer A.O., Boyer R.S. A Biographical Sketch of W. W. Bledsoe [J]. In: Boyer R.S. (eds) Automated Reasoning. Automated Reasoning Series, vol 1. Springer, Dordrecht, 1991.

[4] Christoph von der Malsburg. Mugspot Can Find A Face in The Crowd -- Face-Recognition Software Prepares to Go to Work in The Streets [EB/OL]. (1997-11-12) [2007-11-06]. https:// www. sciencedaily.com/releases/1997/11/971112070100.htm.

[5] Yael Adini, Yael Moses, Shimon Ullman. Face Recognition: The Problem of Compensating for Changes in Illumination Direction [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1997, 19 (7): 721-732.

[6] G. Yang, T.S. Huang, Human face detection in a complex background [J]. Pattern Recog., 1994 (27): 53-63.

[7] Belhumeur, P. N., Hespanha, J. P., & Kriegman, D. J. Eigenfaces vs. Fisherfaces: Recognition using class specific linear projection [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1997, 19 (7): 711-720.

[8] Hinton G, Osindero S, Teh Y.A fast learning algorithm for deep belief nets [J]. Neural Computation, 2006, 18 (7): 1527-1554.

[9] Sun Zhijun, Xue Lei, Xu Yangming, et al. Review of in-depth learning [J]. Computer Applied Research, 2012, 29 (8): 2806-2810.

[10] Lee H, Grosse R, Ranganath R. et al. Unsupervised learning of hierarchical representations with convolutional deep belief networks[J]. Communications of the ACM,2011,54(10): 95-103.

[11] Huang G B, Lee H, Miller E L. Learming hierarchical representations for face verification with convolution deep belief networks[c]/Proceedings of International Conference on Computer Vision and Pattern Recognition. 2012:223-226.

[12] Lee H, Grosse R, Ranganath R, et al. Convolutional deep belief networks for scalable unsupervised learning of hierarchical representations[C]/Proceedings of International Conference on Machine Learning.2009:21-30.

[13] Liu Y, Zhou SS, Chen Q C. Discriminative deep belief for visual data classification[J]. Patterm Recognition, 2011.44(10/11):2287-2296.

[14] Lecun Y, Boser L, Denker JS, et al. Backpropagation applied to handwritten zip code Recognition[J]. Neural Computation,1989,1(4):541-551.

[15] Lecun Y, Botou L. Bengio Y, et al. Gradient-based learning applied to document recognition[J]. 1998,86(11):2278-2324.

[16] Zhou s s, Chen Q C, Wang X L. Convolutional deep networks for visual data classification[J]. Neural Process Letters.2013.38(1):17-27.

[17] Ouyang W, Wang X. Joint deep leaming for pedestrian detection[C]/Proceedings of International Conference on Computer Vision. 2013:2056-2063.

[18] Ouyang W, Luo P. Zeng X, et al. DeeplD-Net: Multistage and deformable deep convolutional neural networks for object detection [EB/OL]. (2014-09-11) [2016-02-23]. http:/arxiv. Org /abs / 1409.3505v1.

[19] Cootes T F, Taylar CJ, Cooper D H, et al. Active shape models: Their training and application[J]. Computer Vision and Image Understanding,1995,61(1):38-59.

[20] Cootes T F, Taylor CJ. Statistical models of appearance for medical image analysis and computer vision[J]. Proceedings of SPIE,2004,4322(1):236-248.

[21] Turk M, Pentland A. Eigenfaces for recognition[J]. Journal of Cognitve Neuroscience, 1991, 3(1):71-86.

[22] Huang G B. Lee H, Leamed-Miller E. Learning hierarchical representations for face verification with convolutional deep belief networks [C]/Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. 2012:2518-2525.

[23] Sun Y, Wang X, Tang X. Hybrid deep learning for computing face Verification[C]//Proceedings of IEEE International Conference on Computer Vision. 2013:1489-1496.

[24] Taigman Y, Yang M, Ranzato M, et al. Web-scale training for face identification [EB/OL]. (2014-06-20) [2016-02-23]. http://arxiv. org/abs/1406.5266.

[25] Sun Y, Wang XD, Tang X. Deep learning face representation from predicting 10,000 classes[C]/Proceedings of IEEE International Conference on Computer Vision and Pattern Recognition. 2014:1891-1898.

[26] Sun Y, Wang XG, Tang X. Deeply learmed face representations are sparse, selective, and robust [EB/OL]. (2014-12-03) [2016-02-23]. hp://arxiv. org/abs/1412.1265.

[27] Sun Y, Wang X D, Tang X. Deep learning face representation by joint identification-verification[J]. Advances in Neural Information Processing Systems,2014,27:1988-1996.

[28] Best-Rowden L. Han H, Oto C, et al. Unconstrained face recognition: Identifying a person of interest from a media collection[J]. IEEE Transactions on Information Forensics and Security, 2014. 9 (12):2144-2157.

[29] Zhao JD, Lu K, He X F. Locality sensitive semisupervised feature selection[J]. Neurocomputing, 2008, 71(10):1842-1849.

[30] He X F, Sun C Y. Special issue on advances in intelligence science and intelligent data engineering [J]. Neurocomputing,2014,139(4860):1-2.

[31] Chen B, Polatkan G, Sapiro G, et al. Deep Leaming with hierarchical convolutional factor analysis[J]. IEEE Transactions on Patterm Analysis and Machine Intelligence,2013,35(8):1887-1901.

[32] Chen B, Polatkan G, Sapiro G, et al. The hierarchical beta process for convolutional factor analysis and deep learning [c]/Proceedings of the 28th International Conedings Ference on Machine Learning. 2011:361-368.

[33] License. Research on the application of convolutional neural network in image recognition [D]. Hangzhou: Zhejiang University, 2012.

[34] Kumar N, Berga A C, Belhumeur P N, et al. Atribute and simile classifiers for face verification [C]/Proceedings of the 12th International Conference on Computer Vision, 2009:365-372.

[35] Lu CC, Tang X. Surpass human-level face validation Performance on LFW with Gaussian Face [EB/OL]. (2014-04-15) [2016-02-23]. htp://arxiv.org/abs/1404.3840.

[36] Liang Shufen, Liu Yinhua and Li Lichen. Face Recognition Algorithms under Unconstrained Conditions Based on LBP and Deep Learning [J]. Journal of Communications, 2014, 35 (6): 154-160.

[37] Wang Ying, Fan Xin, Li Haojie, et al. Multimodal Face Recognition Based on Depth Network [J]. Computer Science, 2015, 42 (9): 61-65.

[38] Nair V, Hinton G E. Rectified linear units improve restricted Boltzmann machines [C]/ Proceedings of the 27th International Conference on Machine Learning. 2010:807-814.

[39] Ranzato M, Susskind J, MnihV, et al. On deep generative models with applications to recognition[C]/Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. 2011: 2857-2864.

[40] Susskind J. Hinton GE. Memisevic R. et al. Modeling the joint density of two images under a variety of transformations[C]/Proceedings of IEEE Conference on Computer Vision and Patern Recognition. 2011:2793-2800.

[41] Luo P, Wang XG, Tang X. Hierarchical face parsing via deep learning[C]/Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. 2012:2480-2487.

[42] BeveridgeJR. Phillips PJ, Bolme D S, et al. The challenge of face recognition from digital point-and-shoot cameras[C]//Proceedings of the 6th International Conference on Biometrics: Theory, Applications and Systems. 2013:1-8.

[43] JiS, Xu W, Yang M, et al.3D convolutional neural networks for human action recognition[J]. IEEE Transaction on Pattern Analysis and Machine Intelligence,2013, 35(1):221-231.

[44] Video Face Recognition Method Based on Deep Learning [D]. Harbin: Harbin University of Technology, 2012.

[45] Wang Jianyun, Li Xiaoxia. An expression recognition method based on in-depth learning [J]. Computer and Modernization, 2015 (1): 84-87.

[46] Liu MY, LiS x, Shan S G, et al. Deeply learning fefeFormable facial action parts model for dynamic expreAnalysis [M]/Computer Vision-ACCV 2014.Springer International Publishing, 2014: 143-157.

[47] Zhou Hui, Zhou Liang and Ding Qiulin. Fatigue state recognition based on in-depth learning Law [J]. Computer Science, 2015, 42 (3): 191-200.

[48] Zhang Yuanyuan, Huo Jing, Yang Wanqi, etc. Second Generation Identity of Deep Belief Network Verifying Heterogeneous Face Verification Algorithms [J]. Journal of Intelligent Systems, 2015,10(2): 193-200.

[49] Shi Xugan, Zhang Shiqing and Zhao Xiaoming. Integrating deep belief network and multi-layered sense Facial expression recognition with knowledger [J]. Minicomputer system, 2015, 36 (7): 1629-1632.

[50] LiS X. Xing J L, Niu Z H, et al. Shape drivern kemel adaptation in convolutional neural network for robust facial traits recognition [C]/Proceedings of the 28th IEEE Conference on Computer Vision and Pattem Recognition. 2015:222-230.

[51] Wang Dan, Shan Shiguang, Zhang Hongming, etc. A kind of hair cut from thick to thin METHODS [J]. Journal of Software, 2013, 24 (10): 2391-2404.

[52] Chen D, Cao X, Wen F. et al. Blessing of dimensionality: High-dimensional feature and its efficient compression for face verification[C]/proceedings of IEEE Conference on Computer Vision and Pattern Recognition. 2013:3025-3032.

[53] Mnih V. Kavukcuoglu K, Silver D, et al. Human-level control through deep reinforcement learning[J]. Nature.2015,518(7540):529-533.

[54] Wang Y Q, Xie ZG, Xu K, et al. An efficient and effective convolutional auto-encoder extreme learning machine network for 3d feature learning[J]. Neurocomputing,2016,174:988-998.

[55] Tisera M D, McDonnell M D. Deep extreme learning machines: Supervised autoencoding architecture for classification[J]. Neurocomputing,2016,174:42-49.