

# A novel face recognition algorithm based on the combination of LBP and CNN

Pengfei Ke, Maoguo Cai, Hanmo Wang, Jialong Chen

College of Information Engineering, Shenzhen University, Shenzhen, China

Email: kepengfei2016@email.szu.edu.cn, Caimg@szu.edu.cn, wanghanmo2017@email.szu.edu.cn, 188207157@qq.com

**Abstract**—In order to overcome the effects of posture, illumination, expression and other factors on face recognition, this paper proposes an algorithm which is based on local binary pattern (LBP) and convolutional neural network (CNN). LBP is a texture description method which describes the local texture features of an image. It has good robustness of illumination and posture. CNN can effectively extract the spatial features of images and reduce the dimensions of features. This paper combines the advantages of LBP and CNN to improve the accuracy of face recognition. The CNN in this algorithm has four convolution layers, two max-pooling layers, one activation layer, one fully connected layer, and one output layer. In order to optimize the network structure, batch normalization layer is added after the convolution layer. We get the local binary pattern coded images and put the images as the input of the CNN and train the network. Hence, we can use the well-trained CNN for classification and identification. The Experiments on the CMU-PIE face database show that our algorithm can effectively improve the rate of face recognition.

**Keywords**—face recognition; Local Binary Pattern; Convolutional Neural Network

## I. INTRODUCTION

Face recognition is a kind of biometrics that is widely used in information security, video surveillance, human-computer interaction and finance. It is one of the most challenging subjects in the fields of machine vision and pattern recognition. The research of face recognition has been more than 30 years when compared with some other biometrics, such as the iris, fingerprints, face recognition has the characteristics of convenience, directness and robustness. [1] proposes a method that uses skin color and Principal component analysis (PCA) for face recognition. Girish G N et al. make a comparison between the MS-LBP and PCA for the face recognition [2]. What's more, scholars also proposed face recognition methods based on different hand-crafted features. For example, Sobel filter is proposed in [3]. And Gabor filter is used to extract the feature for the face [4]. Later, the combination method based on different feature extraction operators was applied to face recognition and achieved better recognition results [5]. Bodla N et al. propose a deep heterogeneous feature fusion network to exploit the complementary information present in features generated by different deep convolutional neural networks (DCNNs) for template-based face recognition [6]. And a cascade CNN is proposed for face recognition in [7]. There are also many different deep learning approaches like Stacked Autoencoder [8], and Deep Belief Network [9]. Although there have made great achievements in the field of face recognition, some problems still exist [10]. For example, face recognition is affected by posture, illumination, occlusion and other factors. In order to eliminate the influence of pose, illumination and expression on the face recognition, we use the LBP descriptor to get the LBP coded image and put the processed image as an input to the convolutional neural network. In order to

Optimize the convolutional neural network, we developed an optimized CNN architecture by adding Batch Normalization.

Rest of the paper is organized as follows. We summarize the methods of face recognition in Section II, and then the proposed algorithm is introduced in Section III. Section IV presents the results of the experiments. Finally, we conclude the paper in Section V.

## II. RELATED WORK

There are many methods for face recognition which can be summarized as methods based on geometric features, methods based on statistical features and methods based on machine learning [11] [12].

### A. Methods Based on Geometric Features

This method is the earliest and most traditional, and it is simple and intuitive. The method is generally to extract the geometric features of eyes and nose. After that, the face can be represented by geometric features. The author uses the geometric feature for emotion recognition in [13]. And [14] proposed a method that combined the geometric feature and support vector machine(SVM). [15] improved performance in facial expression using 32 geometric features. Cheng S et al propose a 3D Constrained Local Model framework for deformable face alignment in depth image [16]. Their framework exploits the intrinsic 3D geometric information in depth data by utilizing robust histogram-based 3D geometric features that are based on normal vectors. Though it is easy to understand and achieve, it requires a high degree of the feature points. Therefore, the robustness of this method is poor.

### B. Methods Based on Statistical Features

This method treats the face image as a matrix and constructs a face pattern space for classification. PCA [17] is one kind of the methods based on statistical features. PCA transforms the image from high dimension to low dimension through K-L transformation to realize dimensionality reduction. Therefore, the image can be represented by a low dimensional vector and retains the important information. Linear discriminant analysis(LDA) [18] and hidden Markov model(HMM) belong to this method too. Now, there are also other method are proposed. In this paper [19], a novel recognition algorithm based on discriminant tensor subspace analysis and extreme learning machine is introduced. Wang H et al. put forward a new compact fisher vector descriptor for robust face recognition [20]. These improvements are designed to project face images into a subspace that is relatively easy to classify.

### C. Methods Based on Machine Learning

Multi-layer perceptron and back propagation were used for face recognition before 2006. And deep learning was proposed in 2006. It has made breakthrough result for

classification. Many different neural networks have proposed to identify the face, including CNN. And it includes convolutional layer, pooled layer, fully connected layer and classification layer. It has the advantages of adaptability and robustness, so it can improve the rate of face recognition [21][22][23]. Shaoxin Li et al. applied convolutional neural network to disentangling non-linear variations in the images and got accuracies of approximately 99% on the Multi-PIE Dataset [24]. [25] proposes a method that uses the CNN with the Logistic classifier to classify the different facial features from the image. [26] designed a four-layer CNN which improves performance of face recognition and yielded 85% accuracy on the FERET dataset. In order to increase the training speed and reduce the model parameters, Xiang Wu et al. proposed a Lighten CNN [27]. The model contains a total of about 4M parameter quantities. Compared with the VGG model, the calculation speed of the Lighten CNN is nine times.

### III. METHODOLOGY

The main work of this paper is studying the algorithm based on the combination of LBP and CNN.

#### A. LBP Feature Extraction

Local Binary Pattern is a texture description method which describes the local texture features of an image [28][29]. The method was first proposed in 1994. It has the advantages of rotation invariance and gray invariance, so it is very robustness in illumination and posture. It can reflect the relationship between the grayscale value of central pixel and the surrounding pixels.

The earliest LBP operator is defined in a window of 3\*3. The central pixel of the window is used as a threshold, and the adjacent eight pixels are compared with it in turn. As shown in Fig. 1, if the value of surrounding pixel is greater

44	118	192	0	1	1
32	83	204	0		1
61	174	250	0	1	1

→ (011111000)<sub>2</sub> = 124

Fig. 1. LBP operator

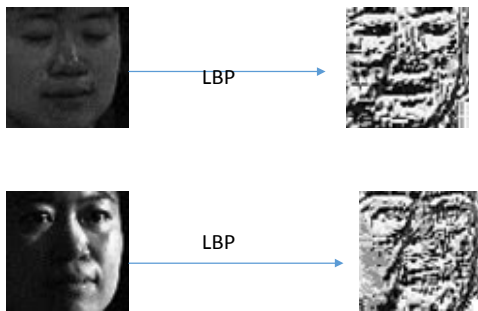


Fig. 2. Local binary pattern coded image

than the value of central pixel, the value is reset to 1, otherwise it is 0. In this way, 8-bit unsigned binary numbers can be generated and it will be converted to a decimal number. And

the value of the central pixel can be replaced by the decimal number. Thus, the LBP code of the window is obtained, and this value is used to represent the texture feature of the window. This method can be expressed by the following formula:

$$LBP_{P,R} = \sum_{i=0}^{n-1} S(g_i - g_c) \times 2^i \quad (1)$$

$$S(x) = \begin{cases} 1, & x > 0 \\ 0, & x \leq 0 \end{cases} \quad (2)$$

Where  $P$  denotes the number of the sample points,  $R$  is radius,  $g_c$  is the value of the central pixel and  $g_i$  is the value of the surrounding pixel.

From above analysis we can know that the extracted LBP operator can get an LBP code for each pixel in the image. After coding, a local binary pattern coded image is got from the original image. The result is shown in the Fig. 2.

#### B. Convolutional Neural Network

Convolutional neural network is a kind of artificial neural network, which has become a hotspot in the field of image recognition [30]. It can reduce the complexity of the network structure and reduce the number of weights. And this network structure is highly invariant to translation, scaling, tilting, or deformation.

The convolutional neural network (CNN) mainly includes the followings: convolution layer, pooling layer, fully connected layer and output layer.

##### 1) Convolution Layer

The convolution is actually an inner product. The size of the convolution kernel is generally smaller than the size of the input image, so the features extracted by the convolution will pay more attention to the partial image. Its weight-sharing can minimize the amount of calculations. (3) is the convolution formula. An example of convolution is shown in Fig.3.

1	1	1	0	0
0	1	1	1	0
0	0	1	1	1
0	0	1	1	0
0	1	1	0	0

Convolution

1	0	1
0	1	0
1	0	1

4	3	4
2	4	3
2	3	4

Fig. 3. Convolution

1	1	2	4
5	6	7	8
3	2	1	0
1	2	3	4

The kernel of max-pooling is 2\*2

6	8
3	4

Fig. 4. Max-pooling

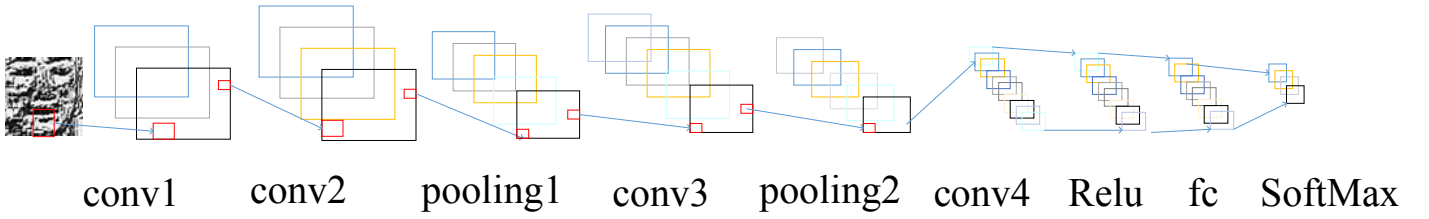


Fig. 5. Convolutional neural network structure

$$x_j^l = f(\sum_{i \in M_j} x_i^{l-1} \times k_{ij}^l + b_j^l) \quad (3)$$

Where  $f(\cdot)$  denotes the activation function,  $x_j^{l-1}$  is the pixel value of the input layer,  $x_j^l$  is the output of the function,  $k_{ij}^l$  is the weight of the convolution kernel,  $b_j^l$  is the bias.

#### 2) Pooling Layer

The pooling layer is also called the down-sampling layer. The pooling layer can reduce the dimension of the feature and decrease the number of parameters. It also has translation invariance. There are generally two kinds of pooling: max-pooling and mean-pooling. As shown in Fig. 4, the effect of the max-pooling is to reduce the original image to 1/4 and retain the maximum in each 2\*2 area as output. And the mean-pooling keeps the average of each 2\*2 area as output.

#### 3) Fully Connected Layer

Each node of the fully connected layer is connected to all the nodes of the previous layer. It is used to integrate the features extracted from the previous layer. The output of this layer is a highly purified feature, which is convenient for the classifier to judge.

#### 4) Output Layer

The output layer of the convolutional neural network is the classifier layer. We use the SOFTMAX regression classifier as the output layer. The SoftMax function is as follow:

$$p(C_j=1|z) = \frac{e^{z_j}}{\sum_{j=1}^k e^{z_j}} \quad (4)$$

Where  $z$  is the input data,  $p$  denotes the probability that the input is classified as the  $j^{th}$  class and  $k$  represents the total number of classes.

### C. Face Recognition Model

Fig.5 is the structure of the convolutional neural network proposed in this paper. It has nine layers in total, including four convolution layers, two max-pooling layers, one activation layer, one fully connected layer and one output layer. The experiment steps are as follows.

The size of the input image is 64\*64 pixels. First, we use LBP operator to get a local binary pattern coded image whose size is still 64\*64. Next, the data is normalized, and it is used as the input of the CNN.

The first convolution layer has a kernel of 3\*3. After convolution, the size of the image is unchanged, still 64\*64 pixels. Next, it is also a convolution layer with a kernel of 5\*5, and the size of the output is 60\*60. After the max-pooling

layer, the output is 30\*30. The fourth layer is still a convolution layer, and the size of the output image is 20 \*20, followed by the max-pooling layer. The output of the max-pooling layer is 10\*10. The sixth layer is convolutional layer and the size of the output is 1\*1. And the seventh layer is an activation function layer. Finally, the image passes through a fully connected layer and the last layer is the output layer.

## IV. EXPERIMENTAL RESULTS AND ANALYSIS

### A. Database Introduction

This experiment uses the CMU-PIE face database, created by Carnegie Mellon University of the United States. It contains a total of 337 volunteers with a total of 75,000 images. The images are collected under different conditions such as illumination and posture. It has gradually become an important test set for face recognition. In this paper, we selected 68 identities in the CMU-PIE database, 49 images of each identity, a total of 3332 images with different illumination and posture. Partial images in CMU-PIE database is shown in Fig.6.

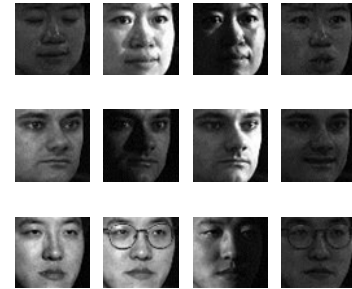


Fig. 6. Partial images in CMU-PIE database

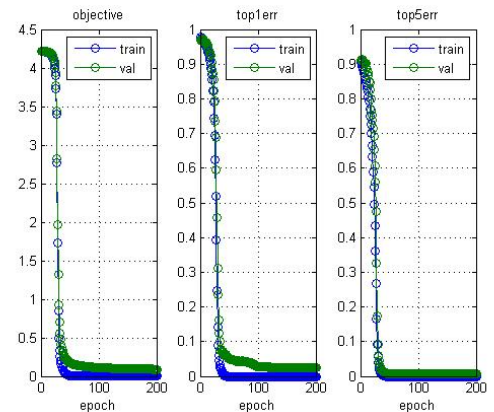
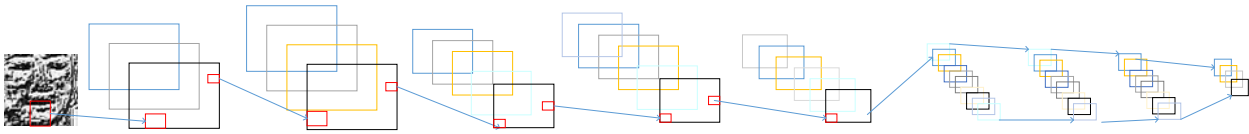


Fig. 7. The change when training the network.



conv1 conv2+bn pooling1 conv3+bn pooling2 conv4+bn Relu fc SoftMax

Fig. 8. Optimized network structure

## B. Experimental Results and Analysis

Fig.7 is the result of training the network. The objective represents the total loss of the function, the top1err is the error rate for the top class is not the same as the target label and the top5err is the error rate for the top 5 class is different from the target label.

### 1) Selection of parameters

Batch size will affect the direction of the gradient descent. At the same time, If the batch size is too small, it will increase the training time. And if the batch size is too large, it is very demanding on computer performance. Therefore, the appropriate batch size must be selected to achieve better results. The learning rate was set to 0.0001. 80% of each person's image are randomly selected as the training set, and the rest of the images are used as the test set. CNN is trained for 200 epochs.

From the TABLE I, the batch size has influence on the recognition rate. Take the training speed into consideration, the best batch size is 100. Therefore, we set the batch size to 100 in our experiment.

We keep the other parameters unchanged and change the proportion of the training set in the database for comparison. And the proportions of the training set were 90%, 80%, and 75%. The effect of the proportion of the training set on the recognition rate are shown in TABLE II. We can obtain the best accuracy when we use the 90% of the images in the database to train our model and set the batch size to 100.

### 2) Optimization

In order to achieve better results, batch normalization is added to the original CNN. When training the network, the update of the previous layer parameters will change the distribution of the input at each subsequent layer. This phenomenon is called internal covariance transfer [31]. For eliminating this phenomenon, a batch normalization layer is added to the network. The role of this layer is to normalize the input data to a mean of 0 and a variance of 1. Therefore, the network converges faster. The optimized network is shown in Fig.8. We compare the original network with the optimized network. In this part, we use 90% images in the database to train the models. The results are depicted in TABLE III. From the TABLE III, we can know that the optimized network can obtain a better result than the original network and we can reach the 100% accuracy on the CMU-PIE database with the optimized network.

TABLE I. THE EFFECT OF BATCH SIZE ON RECOGNITION RATE

Algorithm	Batch size	Recognition rate (%)
LBP+CNN	50	94.02
<b>LBP+CNN</b>	<b>100</b>	<b>94.12</b>
LBP+CNN	200	93.82
LBP+CNN	300	93.94

TABLE II. THE RECOGNITION RATE OF DIFFERENT ALGORITHMS

Algorithm	proportion	Recognition rate (%)	Time (second/epoch)
LBP+CNN	90%	97.65	31.49
LBP+CNN	80%	94.12	29.99
LBP+CNN	75%	95.48	27.12

TABLE III. THE EFFECT OF THE PROPORTION OF THE TRAINING SET ON THE RECOGNITION RATE

Algorithm	Batch size	Recognition rate (%)	Time (second/epoch)
<b>LBP+BN+CNN</b>	100	<b>100</b>	32.45
LBP+CNN	100	97.65	31.45
<b>LBP+BN+CNN</b>	200	<b>99.12</b>	33.76
LBP+CNN	200	97.65	32.44
<b>LBP+BN+CNN</b>	300	<b>97.65</b>	34.21
LBP+CNN	300	97.65	32.02

### 3) Comparison with different algorithms

To compare the performance of different methods, we implemented CNN and LDA algorithms and compared them with the proposed method. What is more, we compared our method with the other existing methods. TABLE IV shows the results comparison between the proposed method and other methods on the CMU-PIE database. Obviously, our method is superior to some other methods.

TABLE IV. COMPARISON WITH DIFFERENT ALGORITHMS

Algorithm	Recognition rate (%)
TWR+PCA [32]	99.20
ELDP [33]	98.29
VGG-Face [34]	93.16
VGG-Face +LDA [35]	99.80
CNN	95.02
LDA	94.61
<b>LBP+BN+CNN</b>	<b>100.00</b>

## V. CONCLUSION

A face recognition algorithm based on the combination of LBP and CNN was investigated in this paper. Firstly, LBP operator was used to get a local binary pattern coded image. Secondly, the images were used to train the CNN. The CMU-PIE face database was used as benchmark. Experiments showed that selection of the batch size could affect the recognition rate and the best batch size is 100 for this network. The original network could significantly improve the face recognition rate. The best accuracy of it is 97.65%. In order to increase the recognition rate, the network was optimized by adding three batch normalization layers. Compared with the original network, the improved network could obtain a better result. The best accuracy of the improved network is 100%. By comparing with some other methods, the algorithm proposed in this paper had good robustness of illumination and posture.

## REFERENCE

- [1] V. M. Praseetha, and S. Vadivel, "Face extraction using skin color and PCA face recognition in a mobile cloudlet environment". In Mobile Cloud Computing, Services, and Engineering (MobileCloud), 2016 4th IEEE International Conference on IEEE, March, 2016, pp. 41-45.
- [2] G. N. Girish, S. N. CL, and P. K. Das, "Face recognition using MB-LBP and PCA: A comparative study". In Computer Communication and Informatics (ICCCI), 2014 International Conference on IEEE, January 2014, pp. 1-6.
- [3] Y. Li, S. Gong, J. Sherrah, and H. Liddell, "Support vector machine based multi-view face detection and recognition". Image and Vision Computing, vol. 22, pp. 413-427, 2004.
- [4] L. A. Cament, F. J. Galdames, K. W. Bowyer, and C. A. Perez, "Face recognition under pose variation with active shape model to adjust Gabor filter kernels and to correct feature extraction location". In Automatic face and gesture recognition (FG), 2015 11th IEEE international conference and workshops on IEEE, May, 2015, pp. 1-6.
- [5] P. P. Das, T. Javid, and S. S. Mahamud, "Single image face recognition based on gabor, sobel and local ternary pattern". International Journal of Computer Applications, vol. 132, 2015.
- [6] N. Bodla, J. Zheng, H. Xu, J. C. Chen, C. Castillo, and R. Chellappa, "Deep heterogeneous feature fusion for template-based face recognition". In Applications of Computer Vision (WACV), 2017 IEEE Winter Conference on. IEEE, March, 2017, pp. 586-595.
- [7] A. Vinay, D. N. Reddy, A. C. Sharma, S. Daksha, N. S. Bhargav, M. K. Kiran, and S. Natrajan, (2017, March). "G-CNN and F-CNN: Two CNN based architectures for face recognition". In Big Data Analytics and Computational Intelligence (ICBDAC), 2017 International Conference on IEEE, March, 2017, pp. 23-28.
- [8] C. Ding, and D. Tao, "Robust face recognition via multimodal deep face representation". IEEE Transactions on Multimedia, vol. 17, pp. 2049-2058, 2015.
- [9] Y. Sun, Y. Chen, X. Wang, and X. Tang, "Deep learning face representation by joint identification-verification". In Advances in neural information processing systems, 2014, pp. 1988-1996.
- [10] G. M. Zafaruddin, and H. S. Fadewar, "Face Recognition: A holistic approach review". In Contemporary Computing and Informatics (IC3I), 2014 International Conference on . IEEE, November, 2014, pp. 175-178.
- [11] A. Özdil, and M. M. Özbilen, A survey on comparison of face recognition algorithms. In IEEE Conference Proceedings on Application of Information and Communication Technologies (AICT) 2014.
- [12] Y. Yan, "Improved face recognition algorithm using extended vector quantization histogram features". In Signal Processing (ICSP), 2016 IEEE 13th International Conference on IEEE, November, 2016, pp. 1046-1050.
- [13] A. Majumder, L. Behera, and V. K. Subramanian, "Emotion recognition from geometric facial features using self-organizing map. Pattern Recognition", vol. 47, pp. 1282-1293, 2014.
- [14] W. Ouarda, H. Trichili, A. M. Alimi, and B. Solaiman, "Face recognition based on geometric features using Support Vector Machines". In Soft Computing and Pattern Recognition (SoCPar), 2014 6th International Conference of IEEE, August, 2014, pp. 89-95.
- [15] G. Palestra, A. Pettinicchio, M. Del Coco, P. Carcagni, M. Leo, and C. Distanto, "Improved performance in facial expression recognition using 32 geometric features". In International Conference on Image Analysis and Processing, Springer, Cham, September 2015, pp.518-528.
- [16] S. Cheng, S.Zafeiriou, A. Athana, and M. Pantic, "3d facial geometric features for constrained local model". In Image Processing (ICIP), 2014 IEEE International Conference on IEEE, October, 2014, pp. 1425-1429.
- [17] J. Yang, D. Zhang, A. F. Frangi, and J. Y. Yang, "Two-dimensional PCA: a new approach to appearance-based face representation and recognition". IEEE transactions on pattern analysis and machine intelligence, vol. 26, pp. 131-137, 2004.
- [18] K. Z. Mao, F. Yang, and W. Tang, "Regularized linear discriminant analysis and its recursive implementation for gene subset selection". In Computational Intelligence in Bioinformatics and Computational Biology (CIBCB), 2011 IEEE Symposium on IEEE, April, 2011, pp. 1-7.
- [19] S. J. Wang, H. L. Chen, W. J. Yan, Y. H. Chen, and X. Fu, "Face recognition and micro-expression recognition based on discriminant tensor subspace analysis plus extreme learning machine". Neural processing letters, vol. 39, pp. 25-43, 2014.
- [20] H. Wang, J. Hu, W. Deng, "Compressing Fisher Vector for Robust Face Recognition". IEEE Access, vol. 5, pp. 23157-23165, 2017.
- [21] G. Liang, Z. Shu, and J. Jianguo, "Fusing deep convolutional network with SFM for 3D face reconstruction". In Signal Processing (ICSP), 2016 IEEE 13th International Conference on . IEEE, November, 2016, pp. 873-878.
- [22] M. Nakada, H. Wang, and D. Terzopoulos, "AcFR: active face recognition using convolutional neural networks". In Computer Vision and Pattern Recognition Workshops (CVPRW), 2017 IEEE Conference on IEEE, July, 2017, pp. 35-40.
- [23] R. Qian, Y. Yue, F. Coenen, and B. Zhang, "Visual attribute classification using feature selection and convolutional neural network". In Signal Processing (ICSP), 2016 IEEE 13th International Conference on IEEE, November, 2016, pp. 649-653.
- [24] S. Li, J. Xing, Z. Niu, S. Shan, and S. Yan, (2015). "Shape driven kernel adaptation in convolutional neural network for robust facial traits recognition". In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015, pp. 222-230.
- [25] H. Khalajzadeh, M. Mansouri, and M. Teshnehlab, "Hierarchical structure based convolutional neural network for face recognition". International Journal of Computational Intelligence and Applications, 12(03), 1350018, 2013.
- [26] A. R. Syafeeza, M. Khalil-Hani, S. S. Liew, and R. Bakhteri, "Convolutional neural network for face recognition with pose and illumination variation". International Journal of Engineering & Technology, vol. 6, pp. 0975-4024, 2014.
- [27] X. Wu, R. He, and Z. Sun, "A lightened cnn for deep face representation". In 2015 IEEE Conference on IEEE Computer Vision and Pattern Recognition (CVPR), Vol. 4, 2015.
- [28] S. C. Pei, M. S. Chen, Y. Yu, S. Tang, and C. Zhong, "Compact LBP and WLBP descriptor with magnitude and direction difference for face recognition". In Image Processing (ICIP), 2017 IEEE International Conference on IEEE, September, 2017, pp. 1067-1071.
- [29] D. Han, and Y. Ming, "Facial expression recognition with LBP and SLPP combined method". In Signal Processing (ICSP), 2014 12th International Conference on IEEE, October, 2014, pp. 1418-1422.
- [30] G. Liang, Z. Shu, and J. Jianguo, "Fusing deep convolutional network with SFM for 3D face reconstruction". In Signal Processing (ICSP), 2016 IEEE 13th International Conference on IEEE, November, 2016, pp. 873-878.
- [31] S. Ioffe, and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift". arXiv preprint arXiv:1502.03167, 2015.
- [32] X. Shi, Z. Guo, F. Nie, L. Yang, J. You, and D. Tao, "Two-dimensional whitening reconstruction for enhancing robustness of principal component analysis". IEEE transactions on pattern analysis and machine intelligence, vol. 38, pp. 2130-2136, 2016.
- [33] M. R. Faraji, and X. Qi, "Face recognition under illumination variations based on eight local directional patterns". IET Biometrics, vol. 4, pp. 10-17, 2015.
- [34] G. M. Mehdipour and E. H. Kemal, "A comprehensive analysis of deep learning based representation for face recognition". In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, 2016, pp. 34-41.
- [35] L. Wan, N. Liu, H. Huo, and T. Fang, "Face Recognition with Convolutional Neural Networks and subspace learning". In Image, Vision and Computing (ICIVC), 2017 2nd International Conference on IEEE, June, 2017, pp. 228-233.