Blind Source Separation for Face Image based on Deep Learning

Xiangning Zhao, Qijia Liu, Jianyu Li, Yina Guo*
Department of Electronics and Information Engineering, Taiyuan University of Science and Technology, China
E-mail: zulibest@163.com

Abstract——**Deep learning is a branch of machine learning, and it is also a particular type of machine learning. *JADE-MTCNN-FaceNet* algorithm is proposed by combining *JADE*, *MTCNN* and *FaceNe*t. The *MTCNN* deep learning convolution neural network exists the problem that the observation data has high dimension of feature information. Considering this problem, the dimension reduction of the feature information, face verification and experimental simulation on *LFW* database are realized by using *FaceNet* network. Experimental results show the algorithm can further improve the accuracy of the face detection and separation. The accuracy rate on *LFW* database is 0.98.**

Keywords—Multi-task convolutional networks, Joint approximation diagonalization eigenmatrices; FaceNET, Blind source separation for face image.

I. INTRODUCTION

Face detection and separation is a hotspot in the fields of machine language and computer vision. Face detection and separation is the phenomenon that the object occlusion and the repetition of the figure will occur in case of non-cooperation. Most people's faces are not frontal. They may be covered by hats, glasses, face towels, etc., which makes the collected face information incomplete as well as makes the face feature extraction difficult. The accuracy of the face detection and separation algorithm is significantly reduced. Therefore, the blind source separation of face image is an important subject. Face detection originated in 1970s. Francis Galton first proposed to represent face by matrix vector representation of geometric distance. The principle of face detection in this algorithm is to calculate the distance between face and each face in the database to match face. In 1997, Lanitis proposed the elastic model method, which is the elastic graph matching technology, by traversing the face image, then extracting the face feature points and coding the feature points into the parameter model. In 2010, You Yi improved the model objective function of neural network, and integrated wavelet transform into neural network. This transformation has a strong ability to overcome the external noise interference in face recognition process. Shi Wenfan improved the Gabor algorithm and combined the neural network model to increase the accuracy of face detection and recognition in 2013.

II. METHODS

A. Multi - task Convolutional Networks, MTCNN

Face detection module is used to separate and convolution cascade neural network algorithm is more tasks, multitasking convolution cascade neural network algorithm (Multi - task Convolutional Networks, MTCNN) is proposed by Qiao Yu group, the Hong Kong institute of face detection method in 2016. By using a well-designed three-level cascade mechanism, the position of the face can be predicted and the face can be detected. Get excellent accuracy while maintaining efficient time. The author first removed the pseudo region and then learned the remaining region through the complexity convolutional neural network (classifier), which is the advantage of the algorithm.

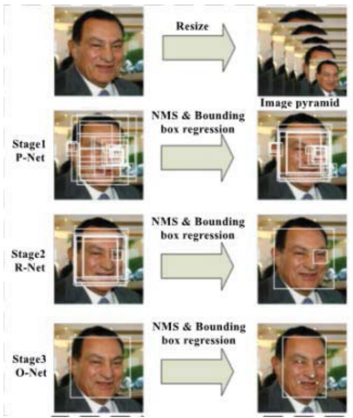

Fig 1. MTCNN frame diagram

Fig. 1 is the frame diagram of MTCNN. As can be seen from Fig. 1, this deep learning network is composed of three layers.

a) Proposleep-network (P-Net) is a regression vector for candidate Windows and bounding boxes of face areas. The bounding box is used for regression, the candidate window is calibrated, and then the highly overlapping candidate boxes are merged by non-maximum inhibition (NMS).

b) Refine Network (R-Net), which use bounding box regression and NMS to remove fake areas. Only because the network structure is different from the network structure of the previous layer, a full connection layer is added, so it can achieve better inhibition of the role of pseudoregions.

c) Output Network (O-Net). This layer has another convolutional layer than the second layer, so the result of processing will be more refined. The effect is the same as the second network effect. But the layer monitors the face area more, and it also detects facial features in five people. The detailed network structure is shown in Fig. 2.

Shenyang, China
November 16-18, 2018

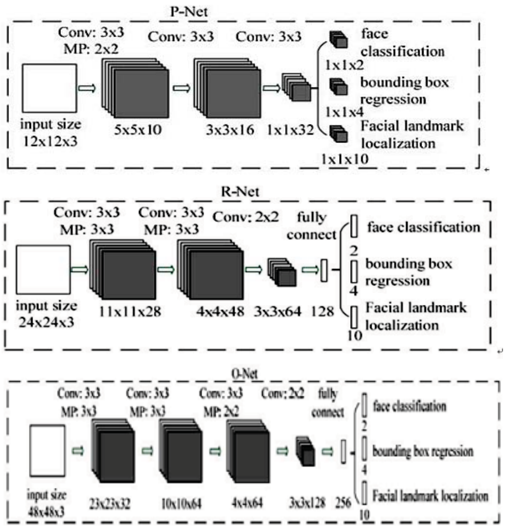

Fig 2. MTCNN network structure

B. JADE-MTCNN-FaceNet Algorithm

This paper mainly focuses on the blind source separation of face images. In recent years, there are many face detection and recognition algorithm models based on deep learning, but there are few researches on aliased face images.

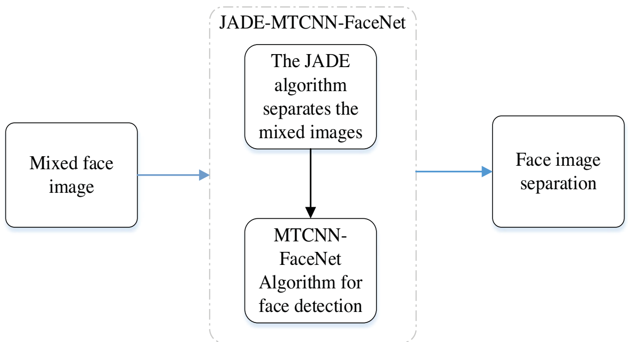

Fig 3. Blind source separation structure of face image

As shown in Fig. 3, JADE-MTCNN-FaceNet algorithm are proposed to realize blind source separation of aliased face images. Because the blind source separation module in the FastICA-MTCNN algorithm uses a batch algorithm and fixed-point iteration number theory to optimize the objective function, therefore, the non-correlated processing is performed for each batch separation matrix component, and thus redundancy is generated. Therefore, the JADE-MTCNN-FaceNet algorithm image separation module uses the JADE algorithm to establish and optimize the objective function by building a fourth-order cumulative energy matrix function. In the face detection and separation module, MTCNN has a large number of convolutional layers to convolve the feature information. Therefore, the feature information acquired by MTCNN is particularly high, so the depth learning FaceNet model is used to reduce the dimension and increase the accuracy of face detection.

The JADE-MTCNN-FaceNet algorithm involves the blind source separation of JADE algorithm and deep learning model MTCNN and FaceNet. Therefore, the model should be established for this algorithm, as shown in Fig. 4, which is the structure block diagram of JADE-MTCNN-FACENET algorithm.

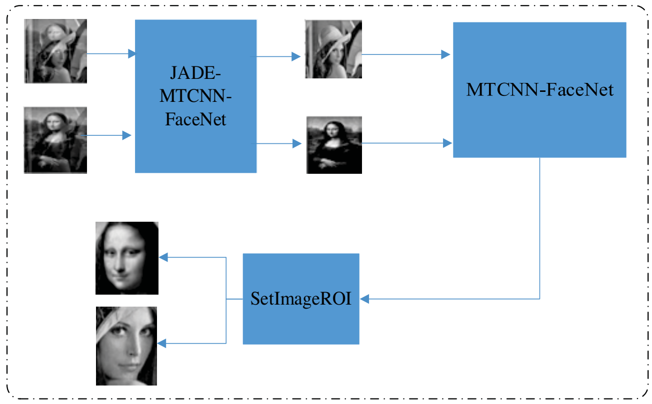

Fig 4. Structure block diagram of JADE-MTCNN-FaceNet

The JADE algorithm in the image separation module reduces the local redundancy of the observed signal and improves the precision of the image blind source separation module. In the face detection and separation module, MTCNN network model is used to detect the face, and FaceNet network model is used to verify the face. FaceNET is a deep learning model proposed by Google in the field of CV. This model mainly consists of three parts. The first part is a deep learning model, in which there is no software max output layer. In the second part, the extracted feature information is normalized by l_2 norm, which reduces the complexity and computational complexity of the model. The third part is to calculate the triple loss of feature information, which is the biggest feature of this model. It trains the classifier directly in Euclidean space, and integrates the function of feature information such as face clustering, face verification and so on.

The image separation module uses a joint diagonalization algorithm based on the fourth-order energy matrix and a Joint approximation diagonalization Eigen-matrices based on the fourth-order energy matrix, JADE is an algorithm proposed by Cardoso. It converts the observed signal into a fourth-order cumulative energy matrix, and then diagonalizes a group of matrices in its energy matrix to separate the matrices. The fourth-order cumulative energy matrix is defined as: Let Z be the N-channel observation vecto $z=[z_1,z_2\cdots z_n]^T$ after the observation signal is spheroidized, M is set to be the $\forall N\times N$ matrix, then the formula of the elements in the fourth-order cumulative energy matrix $Q_z(M)$ is Equation (1).

$$[Q_z(M)]_{ij} \overset{def}{=} \sum_{k=1}^{N}\sum_{l=1}^{N} K_{ijkl}(z)\bullet m_{kl} \quad i,j=1,2\cdots N \tag{1}$$

$K_{ijkl}(z)$ is the fourth-order cumulant in the fourth component of the observation vector. $N\times N$ matrix, m_{kl} is the elements in Row k, Column l.

For actual observation signal data, it can be assumed that,

a. The mean value of the source signals observing the relativity of the signals is zero, and the source signals are independent of each other.

b. There is at most one Gaussian signal.

c. The additive noise in practical application is zero and independent of the source signal.

The observed signal data is X, and there is a mixed matrix that makes Equation (2).

$$x = AS \tag{2}$$

In Equation (2), S is the source semaphore, and the matrix is a full rank matrix. Obviously, the vector is unknown, that is, the prior condition of the observation signal is not known. If you want to find the expression of the mixed matrix, you can spheroidize the observation signal, shown in Equation (3).

$$z = Wx \tag{3}$$

The elements of each row in the spheroidized matrix are orthogonal to each other, each row has equal energy and has a value of 1. Some data also show that spheroidization is whitened, shown in Equation (4).

$$z = WAS = VS \tag{4}$$

Equation (4) is obtained from the fourth-order accumulation of the spheroidized matrix. The matrix in Equation (5) is selected as a set of symmetric / antisymmetric matrices. $N \times N$ can be defined as M_{pq} , shown in Equation (5).

$$M_{pq} = \begin{cases} e_p e_q^T, p = q \\ \frac{1}{\sqrt{2}}\left[e_p e_q^T + e_p^T e_q\right], p < q \\ \frac{1}{\sqrt{2}}\left[e_p^T e_q + e_p^T e_q\right], p > q \end{cases} \tag{5}$$

III. SIMULATION

The face detection separation module in the algorithm FastICA-MTCNN uses a cascade convolutional neural network MTCNN in deep learning network, it is composed of a three-layer Network structure, Propossal Network (P-Net), Refine Network (R-Net) and Output Network (O-Net), respectively. Compared with the traditional CNN, the convolutional neural network structure of MTCNN has a great advantage in speed. The forward direction propagation time and accuracy of the three-layer network structure are shown in Table I.

TABLE I. MTCNN CONTRAST DIAGRAM

Group	CNN	Forward direction propagation time	Accuracy
1	12-Net(19)	0.038s	94.3%
	P-Net	0.031s	94.6%
2	24-Net(19)	0.738s	95.1%
	R-Net	0.458s	95.4%
3	48-Net(19)	3.577s	93.2%
	O-Net	1.347s	95.4%

Using a large number of data from LFW database, the experimental results of Blind Source Separation (BSS) for face images are satisfactory. The experimental results are analyzed by feature information, and the expected results are achieved.

Fig 5. JADE-MTCNN-FaceNet experimental result

TABLE II. EVALUATING INDICATOR

Image	A and F	B and E
IF	0.9876	0.9912
PSNR	29.007	28.258

Fig. 5 is the result of JADE-MTCNN-FaceNet experiment. A and B are two original images. After mixing matrix A, two mixed images, C and d, are formed. Graphs E and F are images separated by JADE algorithm. According to the theory of blind source separation, there will be uncertainties after blind source separation, that is, the order of the separated image will be different from the order of the source image, and this change will not affect the actual effect of the algorithm. Image G and H are face region images separated by MTCNN-FaceNet module. Then the global index PI of the source image and the separated image is calculated, and the peak signal to noise ratio PSNR of the image is calculated.

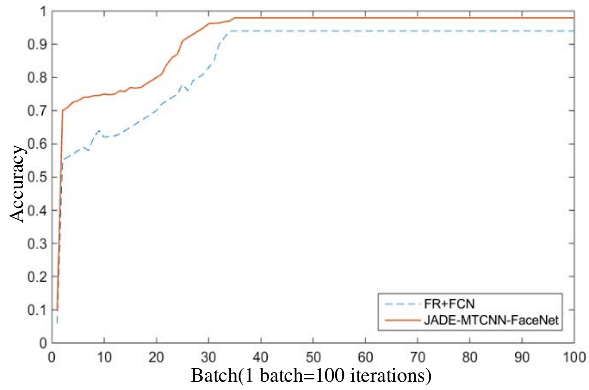

Fig 6. Accuracy chart

As can be seen from Table II, the PI value of the image fidelity is close to 1, so the similarity between the source image and the separated image is very high. From the results of PSNR peak signal to noise ratio, we can see that the results of image separation are more stable. Figure 6 shows the accuracy curve of the training depth learning module MTCNN + FaceNet. The abscissa is the number of times the test set is used for verification, and the ordinate is the accuracy. Compared with FR + FCN model, the accuracy of this algorithm is 0.98, and that of FR + FCN model is 0.94. Obviously, our algorithm has high accuracy.

Fig. 7 is the loss ratio chart of the algorithm and FR+FCN. The abscissa is the number of verification of the test set, and the ordinate is the loss rate. With the increase of the number of verification, the loss rate decreases obviously. It can be seen from the graph that the loss rate of JADE-MTCNN-FaceNet is obviously lower, that means, the learning effect is better.

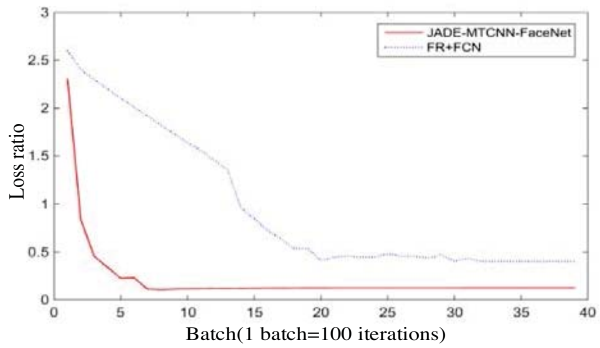

Fig 7. Loss rate diagram

IV. Conclusion

The JADE-MTCNN-FaceNet algorithm is an algorithm for blind source separation of face images. Firstly, the algorithm is modeled, and the blind source separation technology JADE involved in this algorithm is introduced in detail. The JADE algorithm in the face detection and separation module is reasoned, and the depth learning model MTCNN and FaceNET are trained. On the basis of MTCNN face detection, FaceNET model is used to verify the face to improve the accuracy of face detection. The SetImageROI function is then used to separate the detected face regions. A large number of experiments were carried out on the LFW database, and the accuracy rate was 0.98 on the standard set LFW.

Acknowledgement

This work was supported by National Natural Science Foundation of China (61301250), Outstanding Innovative Teams of Higher Learning Institutions of Shanxi ([2015]3), Shanxi Scholarship Council of China (2014-060), Doctoral Program of Taiyuan University of Science and Technology (20152003), Collaborative Innovation Center of Internet+3D Printing in Shanxi Province, Scientific and Technological Innovation Team of Shanxi Province (No. 201705D131025), and Key Innovation Team of Shanxi 1331 Project.

References

[1] [J. Herault, C. Jutten] Space or time adaptive signal processing by neural network model,Denker J. S(ed), Neural networks for computing: API conf.proceeding 151[J]. New York: American Institute for physics, 1986, 04: 13-16.

[2] [L. Tong, R. W. Liu, V. C. Soon, Y. E. Huang] Indeterminacy and identifiability of blind identification [J]. IEEE Trans. Circuits and Systems, 1991: 499-509.

[3] [P. Comon.] Independent component analysis,a new concept?[J]. Signal Processing, 1991: 224-238.

[4] [P. comon, C. Jutten, J. Herault] Blind separation of sources,Part2: Problems statement[J]. Signal Processing, 1991(24): 11-20.

[5] [A. J. Bell, I. J. Sejnowski] An information-maximization approach to blind separation and blind deconvolution[J]. Neural Computation, 1995(7): 1129-1159.

[6] [J. F. Cardoso, B. H. Laheld] Equivariant adaptive source separation[J]. IEEE Transactions Signal Processing, 1996, 44(12): 3017-30.

[7] [T. P. Chen, A. Cichocki.] Stability analysis of adaptive blind source separation[J]. Neural Networks, 1997, 10: 1345-1351.

[8] [A. Hyvarinen, E. Oja] Independent component Analysis:Algorithms and Applications[J]. Neural computation, 2000, 13(4-): 411-430.

[9] [A. Belouchrani, A. Cichocki] A robust whitening procedure in blind source separation context[J]. Electronics Letters, 2000(36): 2050-2051.

[10] [S. Amari,T. P. Chen, A. Cichocki] Non-holonomic Orthogonal learning algorithms for blind source separation[J]. Neural Computation, 2000, 6(12): 1463-1487.

[11] [S. Amari] A theory of adaptive pattern classifiers[J]. IEEE Trane, Electronic, 1967, 16.

[12] [S. Amari] Natural gradient works efficiently in learning[J]. Neural computation, 1998, 10.

[13] [G. S. Fu, P. Ronald, A. Matthew et al.] Blind source separation by entropy rate minimization[J]. IEEE Transactions on Signal Processing, 2014, 62(16): 4245-4255.

[14] [D. Lahat, J. Christian] An alternative proof for the identifiability of independent vector analysis using second order statistics[C]. ICASSP, Shanghai, 2016, 4363-4367.

[15] [N. G. Andrew] Machine learning yearning technical strategy for AI engineers in the era of deep learning [M]. 2016.

[16] [S. Shalev-Shwartz, S. Ben-David] Understanding machine learning from theory to algorithms[M]. Cambridge university press, 2014.

[17] [F. Galton] Personal identification and description[J]. Journal of anthropological institute of Great Britain and Ireland, 1889, 177-191.

[18] [F. Galton] Numeralised profiles for classification and recognition[J]. Nature, 1910, 83: 127-130.

[19] [M. Turk, A. Pentland] Eigenfaces for recognition[J]. Cognitive Neuroscience, 1991, 3(1).

[20] [M. Turk ,A. Pentland] Face recognition using eigenfaces[J]. Proceeding of IEEE Computer Society conference on computer Vision and Pattern Recognition, 1994: 586-591.

[21] [A. Lanitis] Automatic Interpretation and Coding of Face Images Using Flexible Models[J]. IEEE Trans. PAMI, 1997, 19: 743-756.

[22] [A. Lanitis] An automatic face identification system using flexible appearance models[J]. Image&Vision computing, 1995, 13: 393-401.

[23] [Garcia.Christophe, Delakis.Manolis] A neural architecture for fast and robust face detection[J]. IEEE in pattern Recognition, 2002, 2: 22-47.

[24] [Paul.Vioal, Jones.Michael]Robust real-time face detection[J]. International journal of computer vision, 2004, 57(2): 137-154.

[25] [J. Youyi, L. Xiao] Amethon for face recgnition based on wavelet neural network[J]. Intelligent systems(GCIS) 2010 Second WRI Global Congress on. Wuhan China, 2010, 3: 13