Case Studies on Neural Networks for Recognition in Biometric Identity Problem

Zhengwen Shen, Jun Wang*, Zaiyu Pan, Kai Yang

School of Information and Electrical Engineering
China University of Mining and Technology
Xuzhou, Jiangsu 221006, China
E-mails: szwfast@163.com, wj999lx@163.com, pzycumt@163.com, ts16060158a3@cumt.edu.cn

Corresponding author

Received: June 26, 2019 **Accepted: November 20, 2020**

Published: March 31, 2021

Abstract: *Hand-dorsa vein recognition using a convolutional neural network is presented. Our network contains five convolutional layers and three full connected layers, which have high recognition and more robust. The experimental results on the self-established database with the proposed CNN achieves 98.02% in training part and 97.65% in testing part, which demonstrates the effectiveness of the proposed CNN.*

Keywords*: Hand-dorsa vein recognition, High recognition, Convolutional neural network.*

Introduction

Biometrics is an automatic user authentication technology, which uses human physiological and/or behavioral characteristics with several desirable properties like universality, distinctiveness, permanence, and acceptability. In recent years, there has been an increasing interest in vein recognition, which is motivated by the advantages of live identification, non-invasive, and non-contact image capture, and high security over other biometric recognition techniques (e.g., fingerprint, face, iris, voice, gait, etc.) [11, 23, 24].

Commonly, the vein recognition technique involves image acquisition, vein image preprocessing, vein image feature extraction and representation, classifier design and vein recognition [10]. The most important and difficult part, in reality, is the feature extraction methods design, many efforts have been contributed to developing an effective feature extraction method towards vein recognition. The related work can be classified into three groups. The methods in the first group, which refers to the repeated line tracking method [1], the maximum curvature point method [3], the mean curvature [21] and the Gabor filter [22], are based on observation on the geometric shape or topological structure of the vein image. In these methods, the vein network is segmented firstly, and then the topological feature of the vein network is extracted for matching. However, the results of the geometrical models are usually unsatisfied due to the fact that segmentation results of low-quality images are often inaccurate. Vein texture descriptors based on the binary code are adopted in the second group, which covers the local binary pattern (LBP) [13], the local line binary pattern (LLBP) [8], the personalized best bit map (PBBM) [20], etc. These methods transform the image matrix into a 1-D or 2-D feature matrix. The characteristic of being sensitive to the translation and rotation of input image, however, would result in bad recognition performance. To overcome these problems, a multimodal biometric system is employed in the last group, which combines the evidence obtained from hand vein and other traits. Wang et al. [19] proposed the supervised local-preserving canonical correlation analysis method (SLPCCAM) to generate fingerprint-vein feature vectors in feature-level fusion. Finger vein and finger dorsal texture were fused using the proposed holistic fusion and non-linear fusion methods at the score level in [22]. Fan et al. [2] integrated finger vein and finger geometry by means of SVM-based core level fusion. However, along with the higher accuracy, a variety of limitations appear in multimodal biometric systems, such as the longer recognition time, the higher cost for multiple high-quality sensors and the more inconvenience to user. All the methods mentioned above, hand-crafted feature representation models, cannot ensure high recognition accuracy, well robust performance.

In recent years, deep learning has been successfully applied for computer vision, image processing and natural language processing. In the light of their powerful capacity for feature representation, some researchers brought them into biometrics. Several deep learning models such as in [14-16] have been built for face verification and have shown great success on the LFW face dataset. In [17] task-specific transfer learning has been applied to identification and gender recognition with hand-dorsa vein information. Inspired by this idea, we propose, in this paper, a convolutional neural network (CNN) model for hand-dorsa vein recognition. Our network contains five convolution layers, three full connected layer, which has high recognition and more robust. The experimental results on the lab-made hand-dorsa vein database with the proposed CNN achieves 98.02% in training part and 97.65% in testing part, which demonstrates the effectiveness of the proposed CNN.

Lab-made hand-dorsa vein database

To obtain a persuasive and satisfactory classification result, a comprehensive hand-dorsa vein database is built containing 40 volunteers whose ages vary from 19 to 62. For each sample, 20 hand-dorsal vein images were acquired in two specifically set sessions separated by a time interval of more than 10 days, and at each time, five images were acquired from each subject at the wavelength of 850 nm. To the fullest of the dorsal vein information, we set the size of the images as 460×680 with extremely high-quality.

Fig. 1 shows some samples of home-made database. The ROI (Region-Of-Interest) extraction [18] process specifically designed for this database is conducted followed by the grey and size normalization, which is as shown Fig. 2. Since the number of train data is insufficient for learning the many parameters and weights of five convolutional layers, three full connected layers and a hidden layer on CNN. The number of train data is increased through data augmentation for this research, which using existing data such as flip, translation or rotation to create more data, to make the neural network have better generalization effect. After data augmentation, we establish a hand-dorsa vein database which includes 50000 train sets and 10000 test sets.

Fig. 1 Samples of home-made database

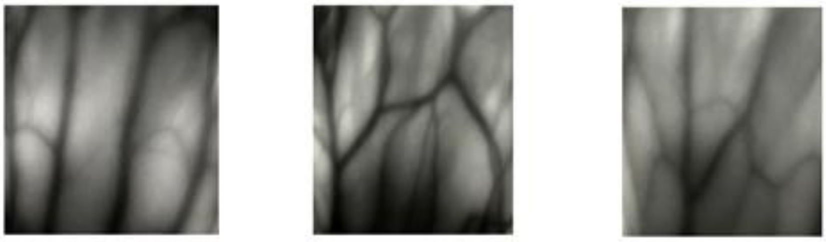


Fig. 2 Input samples after ROI extraction process and normalization

Proposed CNN architecture

As mentioned earlier, the proposed CNN architecture contains five convolutional layers, three full connected layers, as shown in Table 1. The input image size is set to 128×128, and in the 1st convolutional layer, 32 filers of 5×5 are used. Hence, the size of the feature map is 124×124×32 in the 1st convolutional layer, where 124 and 124 are the height and width of the feature map respectively. The rectified linear unit (Relu) layer can be expressed as follows [4, 12]:

$$Y = \max(0, x), \tag{1}$$

where x and y are the input and output value of the Relu function, respectively. The feature map obtained by passing the Relu layer (relu1) passes the 2nd convolutional layer and the Relu layer (relu2) before passing the max-pooling layer (pool1), as shown in Table 1.

Table 1. The architecture of CNN in our experiment

Layer type	Number of filter	Size of feature map	Size of kernel	Number of strides	Number of padding
Image input layer	128 (height) × 128 (width) × 3 (channel/RGB)				
Conv1 Relu1	32	124×124×32	5×5	1×1	0×0
Conv2 Relu2 Pool1	64	120×120×64 60×60×64	5×5	1×1	0×0
Conv3 Relu3 Pool2	128	50×50×128 25×25×128	11×11	1×1	0×0
Conv4 Relu4	256	15×15×256	11×11	1×1	0×0
Conv5 Relu5	512	15×15×512	15×15	1×1	0×0
Ip1 Relu6 Drop6	4096	1×1×4096	1×1	1×1	0×0
Ip2 Relu7 Drop7	4096	1×1×4096	1×1	1×1	0×0
Ip3	40	40			

Because 64 filers of 5×5 are used in 2nd convolutional layer, the feature map is 120×120×64 and the feature map in the max-pooling layer (pool1) is 60×60×64. The feature map passes the third convolutional layer, the relu3 layer and the pool2 layer. In third convolutional layer, 128 filers of 11×11 are used, therefore, their feature map is respectively 50×50×128, 50×50×128, 25×25×128. In fourth convolutional layer, 256 filers of 11×11 are used, and the feature map is 15×15×256. After passing the fifth convolutional layer, a feature map of 1×1×512 pixels is finally obtained.

After the input image of 128×128×3 pixels passes the 5 convolutional layers, 5 Relu layers and 2 pooling layers, a feature map of 1×1×512 pixels is finally obtained. In addition, it passes the three full connected layers (FCLs) The numbers of output nodes of the 1st, 2nd and 3rd FCLs are 4096, 4096, 40. For the 3rd FCL, the Softmax function is used, which is expressed as:

$$\sigma(p)_i = \frac{e^{pi}}{\sum_{n=1}^{R} e^{pn}}. \tag{2}$$

As shown in Eq. (2), given that the array of output neurons is set as *p*, the probability of neurons corresponding to the *i*-th class can be calculated by dividing the value of the *i*-th element by the summation of the values of all the elements.

Commonly, CNN has an over-fitting problem, which can lead to low recognition accuracy, to solve this problem, this paper uses dropout methods [7], which can reduce the effects of the over-fitting problem. In our experiment, we adopt the dropout probability of 50% to disconnect between the previous and the next layer.

Experiments and analysis

This paper presents a novel method for personal identification of hand-dorsa vein patterns using a CNN. The database used in our experiment is a lab-made database named CUMT-Hand-Dorsa Vein, which contains 50000 train sets and 10000 test sets by data augmentation.

Our network is trained using SGD in caffe, and the input image size is set to 128×128. We train using mini-batch size 32 and the maximum of iterations is 9000. The initial learning rate is set to 0.001, and is divided by 10 at 33%, and 66% of the total number of training iterations. We use a weight decay of 5×10^{-3} and a Nesterov momentum of 0.9 without dampening. The experimental results on the lab-made hand-dorsa vein database with the proposed CNN achieves 98.02% in the training part and 97.65% in the testing part.

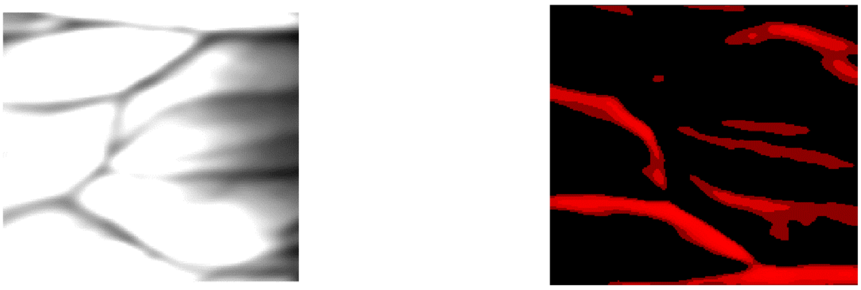
The paper also tries the classical hand-crafted feature model including a library for Support Vector Machines (LIBSVM) [5, 9], and the multi feature and multi classifier, a dorsal hand vein recognition algorithm based on establishing and fusing multi Hidden Markov Models (MFMC) [6] on CUMT-Hand-Dorsa Vein database to evaluate the performance of feature learning model together with hand-crafted model. The recognition result is shown in Table 2.

Table 2. Recognition rate comparison of different methods

Experiment method	Recognition rate, (%)
LIBSVM	88.62
MFMC	89.55
Proposed CNN	98.06

It can be concluded from Table 2 that the proposed CNN model could perform better than the traditional model. What is more, the proposed CNN model has more robust.

In our experiment, in order to demonstrate the effectiveness of the proposed method, we also visualize the feature map of the 1st convolution layer, which is shown in Fig. 3.



(a) original feature map (b) heat feature map

Fig. 3 A feature map of the 1st convolutional layer

It can be concluded from Fig. 3 that our proposed CNN could effectively learn hand-dorsa vein information, (a) is the original feature map, (b) is the convolution feature that the network learns, as shown in the red block of heat feature map, the red block represents the edge information of hand-dorsa vein features learned by the proposed method.

Conclusion

In this paper, a Hand-Dorsa Vein recognition method based on CNN is proposed. Our network contains five convolution layers, three full connected layers, which have high recognition and are more robust. The experimental results on the lab-made hand-dorsa vein database with the proposed CNN achieve 98.02% in the training part and 97.65% in testing part, which demonstrates the effectiveness of the proposed CNN. In the future work, we will enrich the current database to evaluate the efficiency of the proposed CNN model, and the proposed CNN model will be tested with finger-vein database and palm-vein database.

Acknowledgments

This work was supported by the National Natural Science Foundation of China under Grant 61379143 and the Provincial Innovative Program for Graduate Training under Grant SJLX15_0693.

References

1. Chatfield K., V. Lempitsky, A. Vedaldi, A. Zisserman (2011). The Devil Is in the Details: An Evaluation of Recent Feature Encoding Methods, British Machine Vision Conference, 45, 76.1-76.12.
2. Fan R. E., K. W. Chang, C. J. Hsieh, X. R. Wang, C. J. Lin (2008). Liblinear: A Library for Large Linear Classification, Journal of Machine Learning Research, 9(9), 1871-1874.
3. Grauman K. (2006). The Pyramid Match Kernal: Discriminative Classification with Sets of Image Features, Proceedings of the IEEE Int Conf on Computer Vision, 2, 1458-1465.
4. Heaton J. (2015). Artificial Intelligence for Humans, Vol. 3: Deep Learning and Neural Networks; Heaton Research, Inc., St. Louis, MO, USA.
5. http://www.csie.ntu.edu.tw/~cjlin/libsvm/ (Last access March 23, 2021).

6. Jia X., D. Y. Xue, J. J. Cui, J. Liu (2011). Vein Recognition Based on Fusing Multi HMMs with Contourlet Subband Energy Observations, Journal of Electronics & Information Technology, 33(8), 1877-1882.
7. Krizhevsky A., I. Sutskever, G. E. Hinton (2012). ImageNet Classification with Deep Convolutional Neural Networks, International Conference on Neural Information Processing Systems, Vol. 60, 1097-1105.
8. Lazebnik S., C. Schmid, J. Ponce (2006). Beyond Bags of Features: Spatial Pyramid Matching for Recognizing Natural Scene Categories, 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06), New York, NY, USA, 2169-2178, doi: 10.1109/CVPR.2006.68.
9. Li X., X. Liu, Z. Liu (2010). A Dorsal Hand Vein Pattern Recognition Algorithm, 3rd International Congress on Image and Signal Processing, Yantai, 1723-1726, doi: 10.1109/CISP.2010.5647776.
10. Mikolajczyk K., C. Schmid (2005). A Performance Evaluation of Local Descriptors, IEEE Trans Pattern Anal Machine Intell, 27(10), 1615-1630.
11. Moreno P., A. Bernardino (2007). A Comparative Study of Local Descriptors for Object Category Recognition: SIFT vs. HMAX, Iberian Conference on Pattern Recognition and Image Analysis, LNCS, 4477, 515-522.
12. Nair V., G. E. Hinton (2010). Rectified Linear Units Improve Restricted Boltzmann Machines, International Conference on International Conference on Machine Learning, 807-814.
13. Squire D. M., W. Müller, H. Müller, T. Pun (2000). Content-based Query of Image Databases: Inspirations from Text Retrieval, Pattern Recognition Letters, 21(13-14), 1193-1198.
14. Sun Y., D. Liang, X. Wang, X. Tang (2015). DeepID3: Face Recognition with Very Deep Neural Networks, arXiv:1502.00873.
15. Sun Y., X. Wang, X. Tang (2015). Deeply Learned Face Representations Are Sparse, Selective, and Robust, Computer Vision and Pattern Recognition, 2892-2900.
16. Taigman Y., M. Yang, M. Ranzato, L. Wolf (2014). DeepFace: Closing the Gap to Human-level Performance in Face Verification, 2014 IEEE Conference on Computer Vision and Pattern Recognition, Columbus, OH, 1701-1708, doi: 10.1109/CVPR.2014.220.
17. Wang G., J. Wang, Z. Pan (2017). Bimodal Vein Recognition Based on Task-specific Transfer Learning, IEICE Transactions on Information and Systems, E100.D(7), 1538-1541.
18. Wang J., G. Wang, M. Li, K. Wang, H. Tian (2014). Hand Vein Recognition Based on Improved Template Matching, International Journal Bioautomation, 18(4), 337-348.
19. Wang J., J. Yang, K. Yu, F. Lv, T. Huang, Y. Gong (2010). Locality-constrained Linear Coding for Image Classification, Computer Vision and Pattern Recognition, 119, 3360-3367.
20. Yang J., K. Yu, Y. Gong, T. Huang (2009). Linear Spatial Pyramid Matching Using Sparse Coding for Image Classification, 2009 IEEE Conference on Computer Vision and Pattern Recognition, Miami, FL, 1794-1801, doi: 10.1109/CVPR.2009.5206757.
21. Zhao J. Y., J. Gong, S. T. Ma, Z. M. Lu, S-C. Chu, J. Roddick (2019). Finger Vein Recognition Scheme Based on Convolutional Neural Network Using Curvature Gray Image, J Netw Intell, 4(3), 114-123.
22. Zhang J., J. Yang (2009). Finger-vein Image Enhancement Based on Combination of Gray-level Grouping and Circular Gabor Filter, 2009 International Conference on Information Engineering and Computer Science, Wuhan, China, 1-4.
23. Zhang L. (2018). Research on Human Body Movement Posture Based on Inertial Sensor, International Journal Bioautomation, 22(2), 179-186.

 doi: 10.7546/ijba.2021.25.1.000597

24. Zhang M. (2016). Snake Model Based on Improved Genetic Algorithm in Fingerprint Image Segmentation, International Journal Bioautomation, 20(4), 431-440.

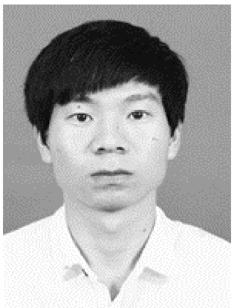
Zhengwen Shen, M.Sc. Student
E-mail: szwfast@163.com

Zhengwen Shen received the B.Sc. degree in College of Physical and Electronic Information from Luoyang Normal University, Henan, China, in 2017. He is currently pursuing the M.Sc. degree in Information and Communication Engineering from China University of Mining and Technology. His current research interests include image processing, machine learning, computer vision and pattern recognition, especially on vein recognition and other biometrical identification methods.

Assoc. Prof. Jun Wang, Ph.D
E-mail: wj999lx@163.com

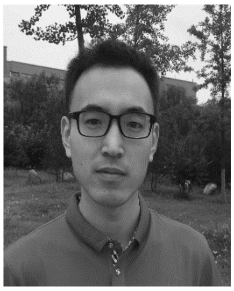
Jun Wang received the B.Sc., M.Sc. and Ph.D. degrees in Control Theory and Control Engineering from China University of Mining and Technology, Jiangsu, China, in 2003, 2009, and 2015, respectively. He has now been an Associate Professor with the School of Information and Electrical Engineering, China University of Mining and Technology. He also worked for the Jiangsu Province Laboratory of Electrical and Automation Engineering for Coal Mining. His current research interests include intelligent control system design, computer vision and pattern recognition, especially on vein recognition and other biometrical identification methods.

Zaiyu Pan, M.Sc. Student
E-mail: pzycumt@163.com



Zaiyu Pan received the B.Sc. degree in Electrical Engineering and Automation from Chaohu University, Anhui, China, in 2016. He is currently pursuing the M.Sc. degree in Control Science and Engineering from China University of Mining and Technology. His current research interests include image processing, machine learning, computer vision and pattern recognition, especially on vein recognition and other biometrical identification methods.

Kai Yang, M.Sc. Student
E-mail: ts16060158a3@cumt.edu.cn



Kai Yang received the B.Sc. degree in School of Mechanical and Electrical Engineering from Xuzhou University of Technology, Jiangsu, China, in 2015. He is currently pursuing the M.Sc. degree in Control Science and Engineering from China University of Mining and Technology. His current research interests include image processing, machine learning, computer vision and pattern recognition, especially on vein recognition and other biometrical identification methods.