



A Novel Method for Finger Vein Recognition

Junying Zeng, Yao Chen, Chuanbo Qin^(✉), Fan Wang, Junying Gan,
Yikui Zhai, and Boyuan Zhu

Faculty of Intelligent Manufacturing, Wuyi University,
Jiangmen 529020, Guangdong, China
zengjunying@126.com, sunnyshenyao@163.com,
tenround@163.com, victorwf1219@163.com,
junyinggan@163.com, yikuizhai@163.com,
zhuboyuan586@163.com

Abstract. Benefiting from CNN's strong feature expression ability, the finger vein recognition systems using the convolutional neural network (CNN) currently have shown a good performance. However, these systems usually adopt such large networks or complex step-by-step processes that they cannot be applied to the hardware platform with limited computing power and small memory. To address this limitation, this research proposes a finger vein recognition network based on difference image and 3C image for cascade fine-tuning. First, a difference image from the image pair for authentic matching or imposter matching is obtained by difference operation and a 2C image is acquired by regarding this image pair as a two-channel image; furthermore, a 3C image is gained with the channel connection of the difference image and 2C image. Then, the SqueezeNet (this network has been pre-trained on ImageNet) that receives the 3C image as input is fine-tuned and the best fine-tune manner is determined. Finally, a cascade fine-tune framework is designed to integrate the difference images and 3C image. In this paper, the size of SqueezeNet which is cascade fine-tuned on the basis of the pre-training weights is 5.63 MB, and the corresponding equal error rate(EER) acquired on the dataset MNCBNU_6000 and SDUMLA-HMT is 1.889% and 4.906% respectively. The experimental results fully prove that the proposed method achieves not only high recognition accuracy but also the simplification of network.

Keywords: Finger vein recognition · Convolutional neural network · Difference image · SqueezeNet

1 Introduction

At present, vein recognition, as a kind of biometric authentication technology with high anti-counterfeiting, has attracted much attention. Compared with traditional biometrics such as fingerprints, faces, and irises, vein recognition has two significant advantages:

This work is supported by NNSF (No. 61771347), Characteristic Innovation Project of Guangdong Province (No. 2017KTSCX181), Young innovative talents project of Guangdong Province (2017KQNCX206), Jiangmen science and technology project ([2017] No. 268), Youth Foundation of Wuyi University (No. 2015zk11).

internal features and liveness detection [1]. Finger vein recognition has been applied to ATM machines, access control systems, vending machines, and various login products in some countries or regions [2]. At the same time, it is difficult to capture high-quality finger vein images during the acquisition of finger vein images, which is still a great challenge for the subsequent recognition process. Due to unique advantages, broad application scenarios and existing challenges, finger vein recognition has attracted more and more researchers' attention.

Traditional finger vein recognition systems mostly use hand-craft features, which are generally sensitive to image quality and finger attitude changes; besides, the pre-processing process is too complicated and the final performance is still limited of the system. To overcome the shortcomings of this method, a few researchers have proposed the CNN-based finger vein recognition, which can automatically learn and extract features with stronger distinguishing ability from the original region of interest (ROI) image. There is no need to filter and enhance the ROI image and the pre-process of image is greatly simplified with CNN. Nevertheless, existing CNN-based finger vein recognition methods mostly adopt such large networks [3] or complex step-by-step processes [4] that they cannot be applied to the hardware platform with limited computing power and small memory.

Considering the weakness of existing CNN-based finger vein recognition methods, this paper introduces a finger vein verification system using cascade fine-tuning CNN, which achieves not only high recognition accuracy but also the simplification of network.

2 Related Works

In recent years, a few researchers have used CNN for finger vein recognition [4–11]. At present, CNN-based finger vein recognition can be roughly classified into the following categories: (1) Finger vein recognition is regarded as a multiclass classification problem [6–8]. In [6], they proposed a reduced complexity CNN with four convolutional layers for finger vein recognition. In [7], using CNN of the same structure as the one used in [6], they retrain the fully connected layer regardless of the pre-training weights. However, these types of methods use finger-vein images of the same class both for training and testing, and for this reason, they only work for 1:N recognition scene on some fixed small databases because they cannot recognize the finger-vein images of classes that have not been trained. (2) CNN is trained to extract the vein structure of a single image and template matching is applied to the obtained binary vein pattern image for verification [9]. This method can identify classes that have not been trained, but the vein pattern extraction and matching process are carried out step by step and the traditional algorithm is still used in the matching with this method. (3) A pair of homologous or heterogeneous images is treated as a single sample whose label is one or zero [4, 5, 10]. In [4], difference image of the image pair is taken as the input of CNN to fine-tune VGGNet-16 [12]. In [5], they design a two-channel network and a two-stream network with a two-channel image as the CNN's input, but in this paper, some images of the same finger both are used for training and others are used for testing in the experiment and there is no definite conclusion when finger images that

have not been trained are tested. In addition to the above, the literature [11] proposes a template image generation model of finger vein based on deep learning and random mapping according to the practical application. Of course, some of the CNN techniques in the above have also achieved good results in the fields of face recognition and image matching [13–15].

Therefore, how to train a finger vein recognition model with high precision and small model size is the focus of this paper, when there are only limited finger vein images. The main contributions of this paper are as follows:

1. This paper is the first to use 3C image as the input of CNN for finger vein verification, which not only make it possible to fine-tune the ImageNet pre-training model but also make full use of finger vein images. In addition, this paper compares the performances of the same network with difference image, 2C image and 3C image as the input, and makes a conclusion that the network performs best with 3C image as the input.
2. Considering that the finger vein recognition algorithm is often embedded in a small portable device, people maybe tend to the network with small size compared with that method using VGGNet-16. For this reason, this paper fine-tunes the light-weight SqueezeNet [16] with 3C image as the input and analyses the performance of SqueezeNet when we fine-tune it in different ways.
3. In order to make full use of the difference and self-information of the image pair to recognize and simultaneously fine-tune the light-weight SqueezeNet, this paper designs a cascade fine-tune framework to train difference images and 3C images hierarchically, further improving the performance of network.

The remainder of the paper is organized as follows: Sect. 3 introduces the proposed method, including how to obtain the 3C image, make use of SqueezeNet, design the cascade fine-tune framework for finger vein recognition. Section 4 describes our experiments as well as analysis on the experimental results. Finally we present conclusions in Sect. 5.

3 Proposed Method

3.1 Acquirement of the 3C Image

There are two methods of a CNN structure's identifying finger vein images of the untrained classes. (1) Using the trained CNN network to extract features from the original single finger vein image, Fig. 1(a) shows this flowchart. (2) Considering a pair of images to be verified as one sample. They use the difference image that is obtained with the difference operation between the input and enrolled image as the input to the CNN in [4], which is shown in Fig. 1(b). They use the 2C image that is obtained with the channel connection between the input and enrolled image as the input to the CNN in [5], which is shown in Fig. 1(c). If we try to fine-tune the ImageNet pre-training model with the difference image as the input, we have to copy the channel of the difference image. It is impossible to fine-tune the ImageNet pre-training model with the 2C image as the input because 2C image is a two-channel image. In addition, copying

channel of difference images is only the simple repetition of difference information and does not fully utilize the information about the image pair themselves. In order to fine-tune the ImageNet pre-training model and avoid copying the channel of the difference image, this paper combines the difference image with the 2C image in channel to acquire the 3C image, which is used as the input of CNN. This flowchart is shown in Fig. 1(d).

3.2 SqueezeNet-Based Finger Vein Recognition

In Fig. 2, the input of the network is a 3C image, and the output of the network is the category, that is, authentic matching or imposter matching. The SqueezeNet employed

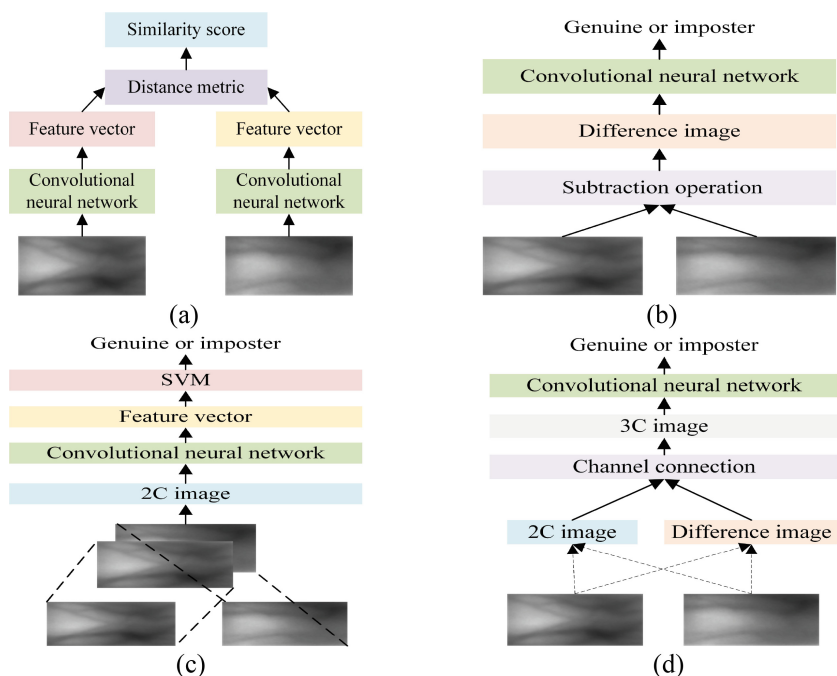


Fig. 1. The CNN structure with the finger vein image as the input

in this paper starts with the convolution layer (conv1), then uses 8 Fire Modules (fire2-fire9), and finally ends with the convolution layer (conv10). The number of filters in each Fire Module is gradually increased, and the layers conv1, fire3, and fire5 are all followed by the max-pooling with the step size of two. Like face recognition, finger vein recognition also consists of the 1:1 verification and the 1:N recognition. When the accuracy of 1:1 verification is high enough, the verification method will be applied to the 1:N recognition. Our method is aimed at solving the 1:1 verification problem that when there is a pair of images, the trained network can output the probability of this image pair's being authentic matching or imposter matching. (The probability of this

image pair’s being authentic matching is regarded as the similarity between them in this paper.)

3.3 Cascaded Fine-Tune Framework Based on Difference Image and 3C Image

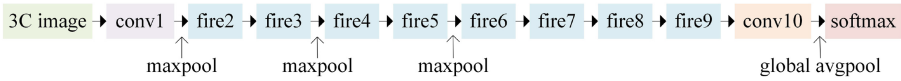


Fig. 2. Finger vein recognition network based on SqueezeNet

Considering a pair of images to be verified as one sample, and the difference image and 3C image can be seen as the various representations of the sample respectively: the difference image focus on describing the difference between the image pair corresponding to the sample; while, the 3C image not only retains the self-information of the image pair but also simply describes the difference between the image pair. In order to make full use of the difference between the image pair and self-information of them, and fine-tune the SqueezeNet at the same time, a cascade fine-tune framework based on the difference image and 3C image is proposed in this paper, which is shown in Fig. 3. Cascade fine-tune framework is a common fusion framework of network. In specific implementations, we fine-tune the pre-trained SqueezeNet with the difference image as the input for the first time, gaining the first optimized model, and then we fine-tune the first optimized model with the 3C image as the input for the second time, obtaining the second optimized model.

4 Experiments

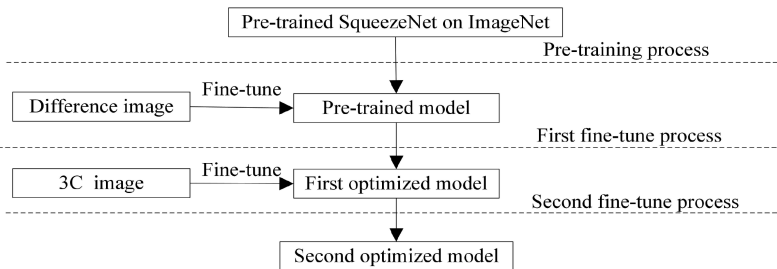


Fig. 3. Cascaded fine-tune framework for finger vein recognition

4.1 Experimental Environment and Data

The MMCBNU_6000 dataset has a total of 6000 images consisting of 100 people with 2 hands and 3 fingers, and 10 images per each finger. This dataset provides the original finger vein images whose resolution is 640×480 and corresponding ROI regional images whose resolution is 128×60 . In order to compare with other algorithms conveniently and fairly, we directly employ the ROI images of this dataset.

The SDUMLA-HMT dataset has a total of 3816 images consisting of 106 people with 2 hands and 3 fingers, and 6 images per each finger. All finger vein images in the dataset are gray-scale images with the resolution of 320×240 . Moreover, using the existing ROI extraction algorithm [18], we obtain the corresponding ROI images which are normalized to the size of 128×60 .

4.2 The Performance of Fine-Tuning Different Pre-Trained Models

The experimental results are shown in Table 1. C-1 is the method in reference [4] and the input of this method is replaced by the 3C image in this paper, namely, the C-2 method. It can be seen that the EER of C-2 method is lower than that of C-1 method on both datasets, which also shows that when fine-tuning the same network, inputting 3C image is better than inputting difference image. Then, this paper takes 3C image as input to fine-tune the SqueezeNet. In the process of transfer learning, there is no fixed choice that which layer weight of the pre-training model is used. In order to make the best transfer learning on two datasets, this paper further carries out experiments: the pre-training weights of SqueezeNet's pre-fire9, pre-fire8, pre-fire7 and pre-fire6 are taken as the initial weights respectively, and the weights of the other layers are randomly initialized. Then, the 3C images are used as input to fine-tune the SqueezeNet on the two datasets, respectively. The experimental results are shown in Table 2. On MMCBNU_6000, the EER of D-3 method is the lowest, while that of D-4 method is the lowest on SDUMLA-HMT. Thus, fine-tuning the same network is different on different datasets. Based on the previous experiments, the cascade fine-tune framework is tried in the experiment. The experimental results are shown in Table 3. This result is compared with those above. As shown in Fig. 4, the EER of this method is lower than that of the above on both datasets. So cascade fine-tuning the SqueezeNet based on difference images and 3C images will have the best performance.

Table 1. EER of fine-tuning VGGNet-16 with different inputs

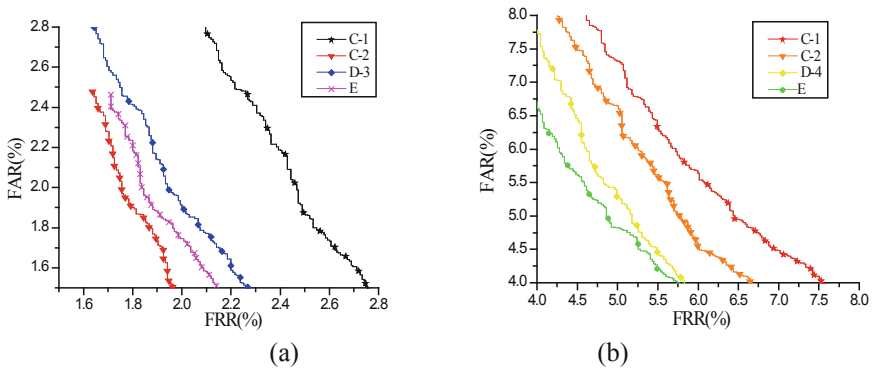
Method name	Input image	Layer for recognition	EER(%)	
			MMCBNU	SDUMLA
C-1(VGGNet-16)	Difference image	Output layer	2.333	5.825
C-2(VGGNet-16)	3C image	Output layer	1.845	5.535

Table 2. EER of fine-tuning SqueezeNet with different ways (Here fine-tune means that pre-trained weights of some layers are as initial weights)

Method name	Input image	Fine-tune	Layer for recognition	EER(%)	
				MMCBNU	SDUMLA
D-1(SqueezeNet)	3C image	Conv1-fire9	Output layer	2.304	5.406
D-2(SqueezeNet)		Conv1-fire8	Output layer	2.185	5.136
D-3(SqueezeNet)		Conv1-fire7	Output layer	1.967	5.281
D-4(SqueezeNet)		Conv1-fire6	Output layer	2.162	5.115

Table 3. EER based on cascaded fine-tune framework

Method name	Implementation process	Layer for recognition	EER(%)	
			MMCBNU	SDUMLA
E(cascaded fine-tune SqueezeNet)	See Part 3.3	Output layer	1.889	4.906

**Fig. 4.** EER of fine-tuning various pre-training models

4.3 Comparisons with Other Methods

Comparing the best E method with other existing algorithms, the results of comparing our E method with other finger vein recognition methods based on CNN are shown in Table 4. In Table 4, we can see that the proposed method has the best performance on MNCBNU_6000 and a little worse performance on SDUMLA-HMT than other methods. As mentioned above, image quality and ROI extraction process will affect subsequent recognition. In addition, the size of the model proposed in this paper is only 5.63 MB, which is much smaller than that of other methods. Considering the comprehensive performance and practical value, the proposed method is superior to other algorithms.

Table 4. Compare our method with other CNN methods in comprehensive performance

	EER(MMCBNU)	EER(SDUMLA)	Model Size
Hong et al. [3]	2.33	5.83	>500 MB
Kang et al. [4]	2.1	3.84	–
(proposed method)	1.89	4.91	5.63 MB

5 Conclusion

In this paper, 3C image is first proposed as CNN's input and we fine-tune the pre-trained SqueezeNet. Furthermore, a cascade fine-tune framework of network based on difference image and 3C image is proposed to improve the recognition accuracy of finger vein recognition. Firstly, the finger vein feature expression ability of three pre-training models is validated on two datasets. Secondly, the performance of the same network under different inputs is compared and the experimental results show that the 3C image has a good effect when it is used as the input of network. Finally, fine-tune is implemented in different ways based on SqueezeNet, and the performance of the model is further improved by using cascade fine-tune framework. Compared with other traditional methods, this method is obviously superior to the traditional feature extraction method. Compared with other CNN methods, this method has good comprehensive performance and high recognition accuracy. How to further improve the recognition accuracy on low-quality image datasets is still a challenging problem on the premise of smaller model. The next step is to consider further improving the recognition accuracy by data enhancement and more refined ROI extraction.

References

1. Yang, G., Xiao, R., Yin, Y., et al.: Finger vein recognition based on personalized weight maps. *Sensors* **13**(9), 12093–12112 (2013)
2. Yang, G., Xi, X., Yin, Y.: Finger vein recognition based on a personalized best bit map. *Sensors* **12**(12), 1738–1757 (2012)
3. Hong, H.G., Lee, M.B., Park, K.R.: Convolutional neural network-based finger-vein recognition using NIR image sensors. *Sensors* **17**(6), 1297 (2017)
4. Fang, Y., Wu, Q., Kang, W.: A novel finger vein verification system based on two-stream convolutional network learning. *Neurocomputing* **290**, 100–107 (2018)
5. Yang, L., Yang, G., Xi, X., et al.: Tri-branch vein structure assisted finger vein recognition. *IEEE Access* **5**, 21020–21028 (2017)
6. Radzi, S.A., Hani, M.K., Bakhteri, R.: Finger-vein biometric identification using convolutional neural network. *Turk. J. Electr. Eng. Comput. Sci.* **24**(3), 1863–1878 (2016)
7. Itqan, K.S., Syafeeza, A.R., Gong, F.G., et al.: User identification system based on finger-vein patterns using Convolutional Neural Network. *ARNP J. Eng. Appl. Sci.* **11**(5), 3316–3319 (2016)
8. Das, R., Piciucco, E., Maiorana, E., et al.: Convolutional neural network for finger-vein-based biometric identification. *IEEE Trans. Inf. Forensics Secur.* **14**(2), 360–373 (2019)

9. Qin, H., El-Yacoubi, M.A.: Deep representation-based feature extraction and recovering for finger-vein verification. *IEEE Trans. Inf. Forensics Secur.* **12**(8), 1816–1829 (2017)
10. Kim, W., Song, J., Park, K.: Multimodal biometric recognition based on convolutional neural network by the fusion of finger-vein and finger shape using near-infrared (NIR) camera sensor. *Sensors* **18**(7), 2296 (2018)
11. Liu, Y., Ling, J., Liu, Z., et al.: Finger vein secure biometric template generation based on deep learning. *Soft. Comput.* **22**(7), 2257–2265 (2018)
12. Simonyan, K., Zisserman, A.: Very deep convolutional networks for large-scale image recognition. *arXiv preprint [arXiv:1409.1556](https://arxiv.org/abs/1409.1556)* (2014)
13. Zagoruyko, S., Komodakis, N.: Learning to compare image patches via convolutional neural networks. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 4353–4361 (2015)
14. Han, X., Leung, T., Jia, Y., et al.: Matchnet: unifying feature and metric learning for patch-based matching. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 3279–3286 (2015)
15. Taigman, Y., Yang, M., Ranzato, M.A., et al.: Deepface: closing the gap to human-level performance in face verification. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1701–1708 (2014)
16. Iandola, F.N., Han, S., Moskewicz, M.W., et al.: Squeezenet: alexnet-level accuracy with 50x fewer parameters and <0.5 mb model size. *arXiv preprint [arXiv:1602.07360](https://arxiv.org/abs/1602.07360)* (2016)
17. Lu, Y., Xie, S.J., Yoon, S., et al.: An available database for the research of finger vein recognition. In: *2013 6th International Congress on Image and Signal Processing (CISP)*, vol. 1, pp. 410–415. IEEE (2013)
18. Kumar, A., Zhou, Y.: Human identification using finger images. *IEEE Trans. Image Process.* **21**(4), 2228–2244 (2012)