Journal of Information Hiding and Multimedia Signal Processing ©2018 ISSN 2073-4212
Ubiquitous International Volume **9**, Number **6**, November **2018**

Palm-print Recognition based on CNN against Rotation and Noise

Lijian Zhou, Zuowei Wang, Han Guo, Siyuan Hao, Zhao Zhuo

School of Communication and Electronic Engineering
Qingdao University of Technology
Qingdao, 266033, P. R. China
zhoulijian@qut.edu.cn, 735966414@qq.com,
364020539@qq.com, 609842513@qq.com, 187294993@qq.com

Received January, 2018; revised August, 2018
(Communicated by Lijian Zhou)

Abstract. *A palm-print recognition method based on Convolutional Neural Network (CNN) is proposed, which can improve the recognition effectiveness in rotation and noise condition. First, the training database with rotations and noises is created based on the original database considering the possible effects in palm-print image acquisition and the big samples set requirement of the CNN method. Second, the experimental CNN model construction approach is proposed, in which the number of hidden layers and convolution kernels, batchsize and epochs are determined by experiments. At last, the test palm-print image is recognized using the softmax method. The experimental results based on Institute of AutomationChinese Academy of Sciences (CASIA) palmprint database show that the proposed method is more effective than the Gabor, Local Binary Pattern (LBP), Multi-block Local Binary Patterns (MLBP) and Vanilla Neural Net (VNN) methods, and has robustness and stability for the images with different noises and rotation degrees.*
Keywords: Palm-print recognition, Database preparation, CNN, Model construction.

1. **Introduction.** Palm-print recognition is one of important biometrics, which is stable, reliable and unique. Main lines and wrinkles are its decisive characters, and the low resolution palm-print image can provide enough information to identify a person, so the palmprint is used widely.

The principal components, fisher features and textures are usually extracted by Principal Component Analysis (PCA) [1], Linear Discriminant Analysis (LDA) [2] and LBP [3] space domain features. The gabor [4], curvelet [5], multi-wavelet [6] are used to extract frequency features. Good recognition results can be obtained using these methods only when the images are registered relative precisely and with good quality. If the images are rotated and added some noises, the recognition rates will decrease obviously. Although the CNN [7, 8] has obtained great success in texture recognition [9], it only has some robustness for small rotations and noises, and usually need a big training database. Motivated by the properties of CNN methods, the expanded database preparation method is proposed to add some training samples with rotations and noises based on original palm-print database to overcome the influences of rotations and noises. Then, the CNN model construction is built referring to [10] and the characteristics of the palm-print images for improving the performance of the CNN method. Two parameters are set first (the 5×5

kernel size and 2×2 average window downsampling method in pooling layer). The experimental selection method of the number of hidden layers, the number of convolution kernels in every hidden layer, the batches and the epochs is proposed.

Combining the above analysis, the palm-print recognition method against rotation and noise is obtained. First, the training database is prepared. Second, the CNN model is built. At last the softmax method is used to recognize the palm-print images.

This paper is organized as follows. The database preparation, CNN construction method and the recognition method are proposed in section 2. The experiments and analysis are performed in section 3. The conclusion is given in section 4.

2. The Proposed Palm-print Recognition Approach.

For improving the robustness of the proposed method against rotation and noise, the database preparation and the CNN model construction are two keys, which are given as follows in detail.

2.1. The Proposed Preparation Scheme of Palm-print Image Database.

The palm-print image database is reconstructed for improving the recognition performance because palm-print images may be blurred, rotated or scaled in different palm-print acquisition processes on one hand. On the other hand, a large number of training samples is necessary in CNN training process.

2.1.1. *The Database Reconstruction Approach.* In this paper, CASIA palmprint database is focused as the original basic database, in which 310 sets of images from different people (8 right hand images for each person) are included. 8 images of a person are shown in figure 1. It can be seen that the images are obtained in different directions and different scales. In general, the Region of Interest (ROI) of a palmprint image is cropped according to the characteristics of fingers and wrists, as shown in figure 2. To improve the robustness of the proposed recognition method, the database is enlarged by adding noises to the images and rotating the images in different angles. Although this may cause the scale change or information loss, a similarity standard is provided for the underdetermined images. For convenience, the ROI is extracted for every palm-print image [11]. The detailed database reconstruction process is as follows.

(1) The rotated palm-print images

The obtained palm-print image in field is usually not a relatively complete hand image with fingers and wrist, so it is very difficult to crop the ROI and make it alignment correctly. Obviously, the recognition results will be better if the database has the images in different directions. Therefore, the basic ROI palm-print images are anticlockwise rotated with the angles 1, 2,$\cdots$, 359 respectively. The bilinear interpolating method [12] is applied to rotate the images. The first image in figure 2 is taken as an example. It is rotated 15 degrees anticlockwise using the bilinear interpolating method, as shown in figure 3 (a). The size of obtained image is larger than the original ROI image. Two kinds of image reserving methods are used for enhancing the robustness of the proposed method further. One is that the rotated image is resized to the size of the original image, as shown in figure 3(b). The other is that a part of information is cut to obtain the image with original size, as shown in figure 3(c). It can be seen that the rotated images not only include the main information in the center, but also provide the different scale information and local important information.

(2) The noised palm-print images

The noise is inevitable in image acquisition. As the gauss noise and PS noise are common, the original basic ROI images are added noises with 0 average value, different variances or intensities to obtain the noised images. The first image in figure 2 is still taken as an example. First, the gauss noise with 0 average value and 0.0095 variance is

added, as shown in figure 4(a). Next, the PS noise with 0 average value and 0.02 noise density is added, as shown in figure 4(b). The main line information can be seen clearly. Thus the images added noises can provide new samples to improve the recognition effects.

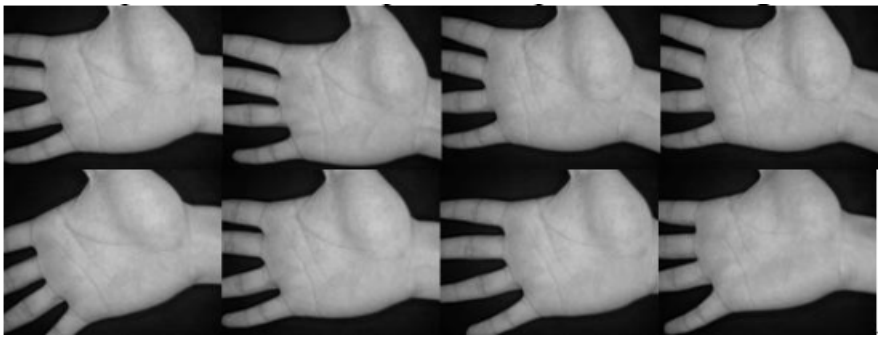

FIGURE 1. The original palmprint images (right hands of a person)

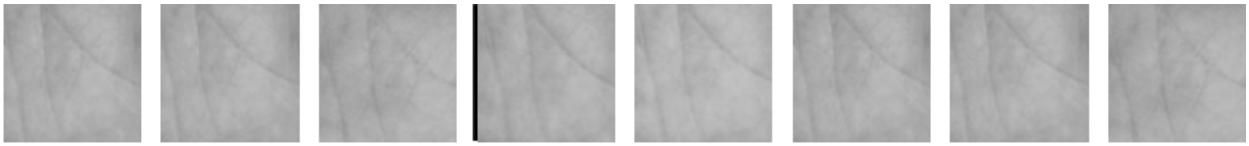

FIGURE 2. The ROI of the original palmprint images in figure 1

FIGURE 3. The images rotated with 15 degree anticlockwise

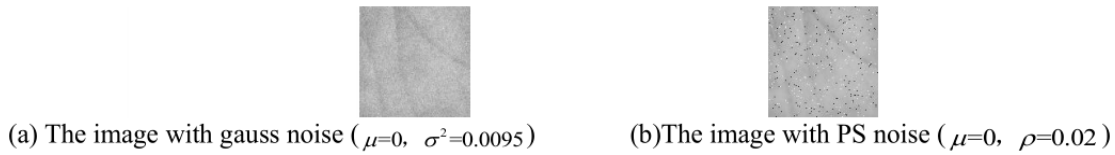

FIGURE 4. The images with gauss and PS noise

2.1.2. *Database Preparation.* To reduce the computation complexity, the size of ROI image is changed to 64×64 in this paper. According to the reconstructed palmprint database method in section 2.1.1, the detailed reconstructed database includes the following images. (1) The extracted ROI images. Total: $310 \times 8 = 2480$. (2) Two kinds of rotated images. Total: $310 \times 8 \times 359 \times 2 = 1780640$. (3) The images with gauss noise ($\mu = 0$), in

which variance changes in square relation because the noisy images with little variance is more common than the ones with big variance and the maximum pixel value is 255.

$$\sigma = x^2/255^2 \tag{1}$$

where the beginning value of x is set to 10, the end value of x is set to 25 and the step is 1. Thus, 16 noised images with gauss noise are obtained for an image.

(4) The images with PS noise ($\mu = 0$), in which density ρ changes in square relation like (3).

$$\rho = x^2/255^2 \tag{2}$$

where the beginning value of x should be italic is set to 36, the end value of x is set to 50 and the step is 1. Thus, 15 noised images with gauss noise are obtained for an image.

Overall, 749 processed images are obtained based on an original ROI image. Thus, there are 6000 images for each person. Considering the CNN training needs a big training database, 5000 images of a person are randomly chosen as the training samples and the rest 1000 images are taken as the testing samples in the experiments.

2.2. **The Proposed CNN Construction Approach for Palm-print.** The number of hidden layers, the number of convolution kernels, the epochs and the number of training samples for every batch (denoted as batchsize) can construct a CNN. It is very important to select the appropriate parameters in constructing the CNN for different objects because it can have a great effect on training processes and the recognition results. Considering the size of palm-print image and to decrease the computation complexity [13-15], the size of the kernel is set to 5×5, the 2×2 average window subsampling is used in pooling layer, a full connection layer (softmax layer) is set and the sigmoid function is chosen as the activation function in this paper.

2.2.1. *The Number of Hidden Layers Selection.* The distribution of the mail lines and wrinkles is the key in palm-print recognition, so the palm-print recognition can be still performed for the images with low resolution and quality. Therefore, the number of hidden layers should be relatively small to avoid over-fitting in the training. The number of hidden layers is selected from 1 to 6 to test, in which the number of kernels is chosen as 6, the batchsize is set to 50 and the epoch is set to 10. The recognition Error Rates (ER) are shown in table 1. For convenience, the number of hidden layers is denoted as NHL in table 1. It can be seen that the result is the best if the number of hidden layers is 2. As a result, the number of hidden layers is chosen as 2 in this paper.

Table 1. The recognition error results with different the number of hidden layers

NHL	1	**2**	3	4	5	6
ER	0.873	**0.192**	0.401	0.855	0.878	0.890

2.2.2. *The Number of Convolution Kernels Selection.* The input images are convolved with the convolution kernels of the hidden layers in the training process. Every convolution result can be seen as a projection component of original image. So the number of convolution kernels is also very important for recognition results. The number of convolution kernels is chosen from 1 to 9 in two hidden layers respectively in the experiments, denoted as KN_1 and KN_2, in which the batchsize is set to 50 and epochs is set to 20. The ERs are shown in table 2. The recognition results with the numbers of the 1^{st} and 2^{nd} hidden layers which are selected as (3, 5), (5, 3), (5, 5), (6, 6), (7, 6) are better than those with other parameters at the same number of hidden layers, batchsize and epochs.

Furthermore, the experiments are performed by selecting the different epochs (10, 20, 30, 40, 50) to find a best pair of kernel numbers. The recognition error results are shown in table 3. The result is the best if the kernel numbers is pair (6, 6). In this case, not only the better results can be obtained when the epochs are relatively small, but also the results are relatively more stable than that of other kernel numbers. Therefore, both kernel numbers are chosen as 6 in the 1^{st} and 2^{nd} hidden layers in this paper.

2.2.3. *The Batchsize Selection.* In general, it is very important to choose the appropriate batchsize. To begin with, the bigger the batchsize is, the faster the computation speed while processing the same data. However, the larger numbers of the epochs and training samples in a batch are needed, and the computation complexity is higher with increasing the batchsize if the same recognition accuracy is obtained. In addition, on one hand, the result may be poor if the batchsize is too big, on the other hand, the convergence of the CNN is very difficult to obtain even if the epochs is set to a big value. The tests are performed in order to obtain an optimum batchsize with epoch= 10 and 2 hidden layers (6 kernels). The results are shown in table 4. It can be seen that the result with batchsize=10 is better than the others. So the batchsize is set to 10 in this paper.

Table 2. The recognition error results with different the number of kernels

KN_1 \ KN_2	1	2	3	4	5	6	7	8	9
1	0.861	0.828	0.859	0.800	0.792	0.801	0.698	0.701	0.737
2	0.880	0.794	0.862	0.698	0.698	0.811	0.834	0.764	0.842
3	0.855	0.753	0.810	0.591	**0.271**	0.595	0.890	0.696	0.799
4	0.760	0.698	0.682	0.559	0.770	0.757	0.642	0.630	0.791
5	0.803	0.785	**0.489**	0.681	**0.238**	0.795	0.855	0.882	0.797
6	0.780	0.786	0.765	0.668	0.591	**0.180**	0.743	0.729	0.836
7	0.890	0.751	0.696	0.680	0.563	**0.440**	0.562	0.705	0.826
8	0.843	0.840	0.890	0.896	0.860	0.684	0.879	0.833	0.669
9	0.882	0.823	0.793	0.799	0.690	0.700	0.891	0.778	0.804

Table 3. The recognition error results with different the number of kernels and epochs

$KN_{1,2}$ \ Epochs	10	20	30	40	50
3,5	0.290	0.203	0.126	0.511	0.880
5,3	0.560	0.483	0.900	0.900	0.700
5,5	0.301	0.015	0.020	0.480	0.841
6,6	**0.192**	**0.008**	**0.007**	**0.590**	**0.780**
7,6	0.470	0.400	0.310	0.900	0.900

2.2.4. *The Epochs Selection.* The tests are performed to find optimum epochs with the parameters selected in the prior tests. The average recognition results of 10 runs are shown in table 5, in which the training and test samples are selected randomly. We can see that the recognition rates are better than 99.95% with epochs=28. Therefore, the epochs is set to 28 in this paper.

TABLE 4. The recognition error results and times with different batchsize

Batchsize	5	**10**	20	40	50	100	200	400
ER	0.419	**0.021**	0.037	0.126	0.192	0.871	0.880	0.900
Time (s)	2448	**2079**	1920	1982	1944	1989	1963	2096

TABLE 5. The recognition error results with different epochs

Epochs	10	11	12	13	14	15	16	17	18
ER	0.021	0.018	0.010	0.032	0.755	0.011	0.008	0.010	0.004
Epochs	19	20	21	22	23	24	25	26	27
ER	0.717	0.013	0.035	0.703	0.036	0.005	0.009	0.003	0.005
Epochs	**28**	29	30	31	32	33	34	35	36
ER	**0.0004**	0.008	0.007	0.001	0.005	0.005	0.004	0.006	0.004

The mean squared error between the desired output and the real output changes with training number is shown in figure 5 to test the effectiveness of the proposed CNN model using the prior design selection. It can be seen that the mean squared error the proposed CNN decreases relatively fast, which reflects the convergence property of the CNN is good. In addition, the recognition rate is more than or equal to 99.95% in 10 random experiments. Therefore, CNN for palm-print images is constructed considering training samples batch processing, computation complexity and over-fitting problem in training process, which includes input layer, two hidden layers (convolution, pooling, convolution, pooling), and the full connection layer, as shown in figure 6. Let take an input image (64×64) as an example to illustrate briefly the work process of an image in CNN model. First, the input image is convolved with 6 kernels (5×5) respectively, 6 images (60×60) are obtained. Second, the downsampling is performed for every convolved image using the window average method. Thus, 6 downsampling images (30×30) are obtained. Third, six convolutions between every downsampling image and 6 kernels are calculated and fused to an image (26×26) by adding them. Fourth, 6 fused images are downsampled like in the second step. Fifth, 6 downsampled images (13×13) is reshaped to a vector (1×1014), which are the input of the full connection layer. At last, the recognition is performed in full connection layer using the softmax algorithm, in which the number of nodes in the connection layer is set to the class number of the training samples.

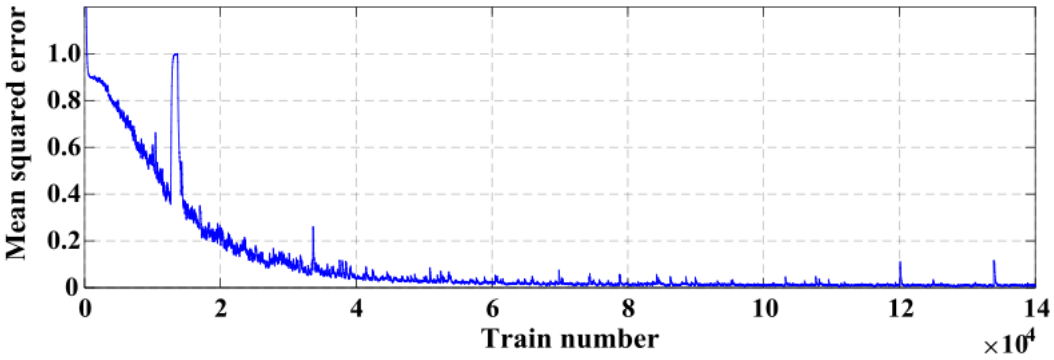

FIGURE 5. The mean squared error between desired and real output verses training number

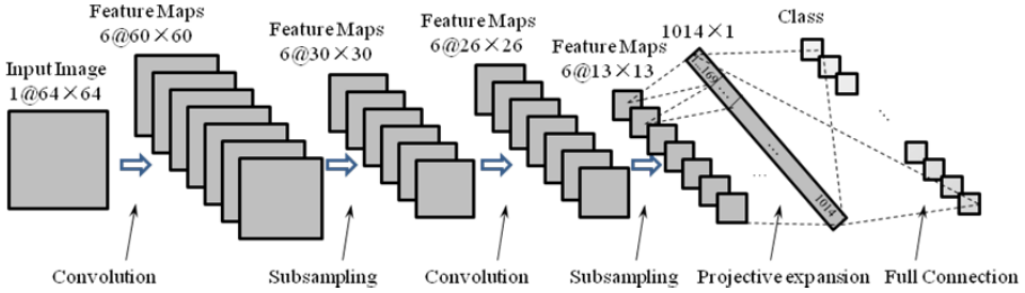

FIGURE 6. The proposed CNN model for palm-print

2.3. **The Process of Proposed Palm-print Recognition Approach.** Step 1: The training palm-print database is reconstructed based on the basic database according to the method proposed in section 2.1.

Step 2: The CNN model construction and training. The size of kernel is set to 5×5, the active function is set to the sigmoid function, the pooling is realized using average method with a 2×2 window and a full connection (softmax algorithm) is selected by considering the characteristics of the palm-print image and some experiences [13-15]. Other parameters are selected using the method in section 2.2, which includes the number of hidden layers, the number of convolution kernels, the batchsize and epochs. At last, the CNN model is trained using the supervised database (training samples).

Step 3: The test samples are recognized using the trained CNN.

3. **Experiments and Results.** Because the training and testing samples are from the reconstructed database in section 2.2 and the test results are effective ($\geq$99.95%), the tests are performed using new images with different rotation angles and noises, and the methods are compared to test the trained CNNs. The detailed test experiments are as follows.

3.1. **The Experiments with Different Rotation Angles and Noises.** The test images are obtained by rotating the original palm-print images with 1 interval from 0.5° to 330.5°. The experimental results are partly shown in table 6, which contains lossless and loss rotation respectively, denoted as ER and ER (loss). It can be seen that the proposed recognition method can effectively identify the palm-print even if the image is rotated.

3.2. **The Experiments with Different Noises.** The test images are obtained by adding the gauss noise (mean is 0) and PS noise to the original images respectively. The experimental results are shown in table 7. In addition, the mixed noises (gauss and PS) are added to test the trained CNN model. ERs are shown in table 8. It can be seen that the proposed recognition method is effective for the images polluted by noises.

Table Contents

TABLE 6. The error recognition results with different rotation angles

Angle	0.5	30.5	60.5	90.5	120.5	150.5	180.5	210.5	240.5	270.5	300.5	330.5
ER	**0**	**0**	**0.012**	**0.013**	**0.011**	**0.012**	**0.014**	**0.013**	**0**	**0**	**0**	**0**
ER(Loss)	**0**	**0**	**0**	**0.013**	**0.012**	**0.011**	**0.012**	**0.012**	**0**	**0**	**0**	**0**

TABLE 7. The error recognition results with different noises

Gauss variance	0.0041	0.0047	0.0052	0.0058	0.0064	0.007	0.008	0.009	0.01	0.011	0.012
ER	**0**	**0.003**	**0.009**	**0.012**	**0**	**0**	**0**	**0.013**	**0.025**	**0.012**	**0.007**
PS density	0.031	0.033	0.034	0.035	0.037	0.038	0.04	0.042	0.043	0.045	0.047
ER	**0.002**	**0**	**0.025**	**0**	**0.025**	**0.023**	**0.013**	**0.012**	**0.025**	**0.037**	**0.012**

TABLE 8. The error recognition results with mixed noises

Guass variance+PS density	0.0041+0.031	0.0047+0.033	0.0052+0.034	0.0058+0.035
ER	**0.027**	**0.025**	**0.025**	**0.029**

3.3. **The Comparison with Conventional Methods.** The experiments were performed by randomly choosing 10, 20, $\cdots$, 50 images as testing samples respectively from the rotated images (loss, lossless), the noised images (gauss, PS) used in two above experiments (1) and (2). The experimental results are shown in table 9, which are obtained by using the conventional Gabor+LDA, LBP+LDA, MLBP+LDA and the proposed methods. The detailed information of conventional methods can refer to [6]. According to the principles of the conventional methods, they should have robustness in small noise and rotation conditions to some extent and little robustness while big noises and rotating angles. From table 9, the recognition rates using the proposed method are better than the results using Gabor+LDA, LBP+LDA and MLBP+LDA methods. Especially, the proposed method is more stable than the other conventional ones.

TABLE 9. The recognition rates using different recognition methods (%)

N^*	Gabor+LDA	LBP+LDA	MLBP+LDA	Proposed approach
10	85.50	62.39	86.36	**95.64**
20	92.05	76.99	92.70	**97.32**
30	94.37	81.81	94.99	**99.88**
40	95.92	83.51	96.32	**99.69**
50	96.80	88.32	96.80	**99.97**

*N is the number of testing samples.

3.4. **The Comparison with Vanilla Neural Net Methods.** The samples are chosen as section 2.1.2, 5000 images are randomly taken as training samples, the rest 1000 images are testing samples. The batchsize is set to 10. Three layers architecture is taken in Vanilla Neural Net (VNN), in which the number of neurons in every layer is $N \times N$ (the size of input image), and the number of the class is 100. A normal VNN and a VNN with dropout are used in the experiments. Additionally, the tanh and sigmoid active functions are selected in different experiments. The results are shown in table 10. It can be seen that the recognition accuracy and stability using the proposed method is better than those using the VNN methods. The less epochs can obtain the better effectiveness using the proposed method (28) than the VNN methods (150).

TABLE 10. The recognition rates using VNN and the proposed approach(%)

Epochs	VNN+tanh	VNN+dropout+tanh	VNN+dropout+sigmoid	Proposed approach
28	67.36	82.05	85.47	**99.96**
60	70.34	88.49	82.18	**93.83**
90	69.55	90.13	83.83	**94.65**
120	70.15	81.02	86.01	**96.88**
150	75.28	84.69	90.44	**97.29**
180	70.01	83.95	89.13	**96.00**

4. **Conclusions.** For improving the palm-print recognition method against rotation and noise and conquering the training sample insufficient in using CNN method, the palm-print recognition method is proposed based on the expanded database and CNN. The experimental results based on CASIA palmprint database show that the proposed approach can effectively recognize the palm-print image. The database preparation method will be studied further based on frequency domain methods since the current methods only consider the space domain methods, such as Gabor or multi-wavelet transformation.

Acknowledgement. This work was supported by Shandong Province Natural Science Foundation, China: ZR2016FQ14. Qingdao Source Innovation Program (Special Project for Applied Research -Special Project for Young Scholars): 17-1-1-2-jch.

REFERENCES

[1] Cai P S, Yan L L. Application of Principal Component Analysis to Palmprint Images Recognition [J]. Computer Systems & Applications, 2010: 187-190.
[2] Yu P, Yu P, Xu D. Comparison of PCA, LDA and GDA for palmprint verification[C]//International Conference on Information NETWORKING and Automation. IEEE, 2010, 1: 148-152.
[3] Laxmi V. Palmprint Matching Using LBP[C]//International Conference on Computing Sciences, IEEE Computer Society. 2012: 110-115.
[4] Pan X, Ruan Q Q. Palmprint recognition using Gabor-based local invariant features [J]. Neurocomputing, 2009, 72(7): 2040-2045.
[5] Liu F, Zhou L, Lu Z M, et al. Palmprint feature extraction based on curvelet transform [J]. Journal of Information Hiding & Multimedia Signal Processing, 2015, 6(1): 131-139.
[6] Zhou L, Zhang C, Zhao K, et al. Palmprint feature extraction based on multi-wavelet and complex network [J]. Journal of Information Hiding & Multimedia Signal Processing, 2017, 8(3): 589-598.
[7] Jalali A, Mallipeddi R, Lee M. Deformation Invariant and Contactless Palmprint Recognition Using Convolutional Neural Network[C]//International Conference on Human-Agent Interaction, ACM. 2015: 209-212.
[8] Krizhevsky A, Sutskever I, Hinton G E. ImageNet classification with deep convolutional neural networks[C]//International Conference on Neural Information Processing Systems, Curran Associates Inc. 2012: 1097-1105.
[9] Cimpoi M, Maji S, Kokkinos I, et al. Deep Filter Banks for Texture Recognition, Description, and Segmentation [J]. International Journal of Computer Vision, 2016, 118(1): 65-94.
[10] Andrearczyk V, Whelan P F. Using filter banks in Convolutional Neural Networks for texture classification [J]. Pattern Recognition Letters, 2016, 84: 63-69.
[11] Zhang D, Kong W K, You J, et al. Online Palmprint Identification [J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2003, 25(9): 1041-1050.
[12] Gribbon K T, Bailey D G. A Novel Approach to Real-time Bilinear Interpolation[C]//IEEE International Workshop on Electronic Design, Test and Applications. IEEE Computer Society, 2004: 126.
[13] Bao X, Guo Z. Extracting region of interest for palmprint by convolutional neural networks[C]//International Conference on Image Processing Theory TOOLS and Applications. IEEE, 2017: 1-6.

[14] Huang L L, Li N. Palmprint Recognition Using Polynomial Neural Network, Berlin Heidelberg. Springer, 2010: 208-213.
[15] Lin H I, Hsu M H, Chen W K. Human hand gesture recognition using a convolution neural network[C]//IEEE International Conference on Automation Science and Engineering. IEEE, 2014: 1038-1043.