

A Novel Finger-Knuckle-Print Recognition Based on Batch-Normalized CNN

Yikui Zhai¹, He Cao¹, Lu Cao^{1(⊠)}, Hui Ma¹, Junyin Gan¹, Junying Zeng¹, Vincenzo Piuri³, Fabio Scotti³, Wenbo Deng¹, Yihang Zhi², and Jinxin Wang¹

¹ School of Information Engineering, Wuyi University, Dongchengstr. 22, Jiangmen 529020, China yikuizhai@163.com, caohell5@163.com, caolu20001742@163.com, mahuiwuyi@163.com, junyinggan@163.com, zengjunying@163.com, wenbodeng92@163.com, 18326410857@163.com ² School of Computer, Wuyi University, Dongchengstr. 22, Jiangmen 529020, China yihangzhi0@163.com
³ Dipartimento Di Informatica, Universita' degli Studi di Milano, Milan 26013, CR, Italy {vincenzo.piuri, fabio.scotti}@unimi.it

Abstract. Traditional feature extraction methods, such as Gabor filter and competitive coding, have been widely used in finger-knuckle-print (FKP) recognition. However, these methods focus on manually designed features which may not achieve satisfying results on FKP images. In order to solve this problem, a novel batch-normalized Convolutional Neural Network (CNN) architecture with data augmentation for FKP recognition is proposed. Firstly, a novel batch-normalized CNN is designed specifically for FKP recognition. Then, random histogram equalization is adopted as data augmentation here for training the CNN in FKP recognition. Meanwhile, batch-normalization is adopted to avoid overfitting during network training. Extensive experiments performed on the PolyU FKP database show that compared with traditional feature extraction method, the proposed method can not only extract more discriminative features, but also improve the accuracy of FKP recognition.

Keywords: Finger-knuckle-print · Batch-normalized · Data augmentation

1 Introduction

Due to the huge market demand of personal authentication, it has attracted much attention in the academic and industry fields. Biometric authentication [1-4] can provide higher security than normal computer passwords which are utilize in applications such as: bank security, computer security system and national ID card etc. Over the past few decades, researchers have focused on the use of biometric traits, like face, fingerprint, iris, palmprint, hand vein, voice, gait etc. In recent years, hand-based biometrics have attracted more attention compared to other biometrics identifiers.

© Springer Nature Switzerland AG 2018

J. Zhou et al. (Eds.): CCBR 2018, LNCS 10996, pp. 11–21, 2018. https://doi.org/10.1007/978-3-319-97909-0_2

Biometrics, such as palmprint [1], hand geometry [2], fingerprint [3] and hand vein [4], have been fully researched.

Owing to the uniqueness of FKP, FKP can be considered as a distinctive biometric identifier technique [5, 6]. The unique advantages of the FKP, compared with other biometrics are as follows: the surface of FKP is not easy to be abraded because people usually hold things with the inner side of their hands. Because of non-contact characteristic the collection of the FKP, the users usually have higher acceptance [7]. As such, FKP is considered one of the most promising personal identification technologies of the future.

However, to the best of our knowledge, these are no investigations about the application of convolution neural network for FKP recognition. Hence, we designed a novel batch-normalized CNN with deep learning method for FKP recognition to improve recognition accuracy. Compared with traditional methods, the proposed CNN could extract more distinctive features and achieve satisfying recognition performance. The main contributions of this paper are as follows: (i) A novel CNN architecture specialty for FKP recognition is designed. (ii) Histogram equalization method as a data augmentation method is adopted to get more training data for FKP recognition. (iii) Batch-normalization is utilized to prevent overfitting in CNN, respectively.

The rest of this paper is structured as follows: Sect. 2 introduces the existing FKP recognition methods. Section 3 presents the proposed batch-normalized CNN architecture. Section 4 describes the recognition process of FKP images. Section 5 shows and analyses the experimental results. Finally, Sect. 6 presents the conclusions.

2 Related Work

Many recent studies on FKP recognition attempt to generate distinctive and robust feature representation for FKP images. Woodard et al. [8, 9] built a 3D hand database and extract 3-D features from finger surface for authentication. Due to the high cost and time consuming of 3D acquisition equipment, the real-time performance of the biometrics system is affected. And there is no effective system for the extraction of the characteristics of the outer surface of the finger. Kumar et al. [10, 11] extracted fingerback surface image features to personal authentication by subspace analysis methods. Subspace analysis is widely used in face recognition task because it has the characteristics of strong distinctive, low computational cost, it is easy to realize and good separability; however, it cannot effectively extract the line features, such as FKP images. Zhang et al. [12] used Gabor filters to extract the feature and use the competitive coding (CompCode) to encode it, as the final feature representation for FKP images. Later, Zhang et al. [5] used Gabor filter to extract the orientation information and magnitude information of FKP images. In [13], the coefficient of the Fourier transform of the FKP images as features, and the similarity of the images is calculated by band-limited phase-only correlation technique. To obtain more FKP images, Morales et al. [14] used Gabor filter to enhance the FKP lines, then adopted Scale Invariant Feature Transform (SIFT) to extract features. Le et al. [15] proposed a robust feature presentation and matching method based on Speeded-Up Robust Features (SURF). This method is robust, which is invariant to the change of rotation, scale and

13

viewpoint. Barinath et al. [16] used a combination of SIFT and SURF to enhance the texture of FKP images. Owing to the complementarity of the two described approaches, the FKP recognition method made great progress. Li et al. [17] used a high-order steerable filter to extract continuous orientation feature maps, an Adaptive Steerable Orientation Coding Scheme (ASOC) is proposed. Yang et al. [18] proposed a Fisher discriminant analysis framework for FKP recognition. Zhang et al. [19] used RCode1 and RCode2 to code feature, improving the FKP recognition rate; however, these methods could not achieve desirable FKP recognition results.

3 Batch-Normalized CNN

A novel CNN architecture specifically for FKP recognition has been designed. The batch-normalized CNN architecture is shown in Fig. 1, which includes 4 convolution layers and 3 fully connected layers. 'C' denotes the convolution layer, the maxpooling layer and the full connection layer are represented by 'MP' and 'FC', respectively. During the training stage, the input of the CNN is a 220×110 grayscale image; all the images are cropped into 110×110 randomly as the input of the entire network. The parameters of each layer are optimized based on multiple experimental verification. Owing to the small FKP database, the solution to avoid overfitting is crucial. Hence, to prevent the training overfitting, a dropout layer is adopted in the proposed CNN and a batch of normalized layer is added after each convolution layer. Details of the network structure and parameters of batch-normalized CNN are shown in Table 1.



Fig. 1. The batch-normalized CNN architecture.

3.1 Batch-Normalized Convolutional Neural Networks

Due to the change of parameters of the previous network layers during the training process, the input of other layers will be affected. This leads to a large amount of computation from deep neural networks. This may reduce the speed of training, requiring lower learning rates and a careful parameter initialization. This makes training models notoriously difficult to saturating nonlinearities [21].

Batch normalization can be used for activation sets in a network. Here, we consider a affine transform followed by an element nonlinearity:

$$z = g(Wu + b) \tag{1}$$

where *W* and *b* represent the parameters that the model needs to optimize. Function $g(\cdot)$ represents a non-linearity function such as ReLU. Batch normalization is applied to the all convolutional layers and the full connected layers. And BN transform is added by normalizing x = Wu + b before non-linearity. The input layer *u* should be normalized, because *u* is the output of another non-linear layer whose distribution may change during the training process. This also limits the first and second phases of the input layer and mitigating the covariance shift. On the contrary, since Wu + b has symmetry and a non-sparse distribution, normalizing is more likely to produce a stable distribution of the excitation function.

Note that the subsequent mean subtraction can cancel the effects of bias b, so bias b can be ignored, as a result, z = g(Wu + b) can be replaced with:

$$z = g(BN(Wu)) \tag{2}$$

A pair of parameters $\gamma(k)$ and $\beta(k)$ are optimized at each layer, where BN transform is utilized.

For convolutional layers, we also require that the normalization should obey the convolutional property, so that the same feature maps at different locations using different elements are normalized in the same way. In order to achieve this, all the activations at all locations are normalized in a mini-batch. Given a feature map, all the activation of a given feature map adopts the same linear transformation.

Layer name	Alexnet [20]	Batch-normalized CNN	
Input	224	220 × 110	
Conv1	96,11 \times 11 kernels	96,7 \times 7 kernels	
Conv2	$256,5 \times 5$ kernels	$128,5 \times 5$ kernels	
Conv3	$384,3 \times 3$ kernels	$128,3 \times 3$ kernels	
Conv4	$384,3 \times 3$ kernels	$128,3 \times 3$ kernels	
IP1	4096	1024	
IP2	4096	512	
IP3	1000	165	
Loss	Softmax	Softmax	
Convolution kernels	1376	480	

Table 1. Comparison of network model architecture.

3.2 **ReLU Activation Function**

Compared with traditional saturated nonlinear activation functions, such as tanh and sigmoid, etc. ReLU (Rectified Linear Units), as non-saturated nonlinear activation function, has a faster network convergence speed. During forward propagation sigmoid and tanh function requires exponential calculation, while ReLU only needs to set a threshold value. In this paper, we utilized ReLU as the activation function of the CNN

architecture, which mimics the characteristics of a unilateral and a sparse activation of biological neurons. ReLU activation function can be expressed as:

$$f(x) = \max(0, x) \tag{3}$$

When the input is less than or equal to 0, the response is 0, otherwise the response directly equals its own value. Due to the characteristics of the ReLU function, the output has some sparsity, which can speed up the network convergence and make the CNN have a stronger classification ability.

$$f'(x) \begin{cases} 0, & x \le 0\\ 1, & x > 0 \end{cases}$$
(4)

By Eq. (4), the gradient is not saturated, only when x > 0, f'(x) = 1; therefore, in the process of backward propagation, the problem of gradient dispersion can be alleviated and the parameters of CNN can be updated quickly.

4 FKP Recognition Process

The recognition process of FKP images is divided into two stages: training and testing. Training stage includes three parts: ROI extraction, data augmentation and batchnormalized CNN training. The quality of the training data is critical to the training of the network, so it is necessary to ensure that the adopted database is effective. Firstly, we adopted a two-stage center point detection [22] to improve the positioning accuracy in skewed conditions. Secondly, histogram equalization method is utilized to obtain more data for the FKP recognition. Finally, the preprocessed training data is used as an input to train the proposed CNN model. The testing stage includes two parts: data preprocessing (ROI extraction) and identifying results. First, preprocessing testing data, and then the trained model is adopted to recognize FKP images.

4.1 ROI Extraction

The FKP images have a lot of background noise, which will do harm to the recognition rate. [23, 24] shows that the recognition performance of FKP image recognition highly depends on the accuracy of ROI extraction. This paper used a more effective two-stage center point detection method [22] to extract ROI of FKP images. The method has two stages: center point preliminary detection and center point secondary precise positioning. For more details, please refer to [22]. In Fig. 2, the first row are the original FKP images of four people in the PolyU FKP database [25] and images in the second row are the ROI images, respectively.



Fig. 2. Original FKP images and ROI images.



Fig. 3. (a) FKP ROI image, (b) FKP ROI image with histogram equalization.

4.2 Data Augmentation via Random Histogram Equalization

As shown in Fig. 3(a), the FKP ROI image has been curved with an nonuniform reflections, resulting in a low contrast; hence, we augmented the FKP ROI images by histogram equalization method to a better texture image distribution. Additionaly, PolyU database is small, which will bring limitations to the algorithm, making it appear overfitted in the learning process. Inspired by [26], histogram equalization was used with different levels to augment the FKP data, which significantly improve the classification results. In general, the FKP database after histogram equalization will satisfy the requirement of the data quantity and avoids the overfitting phenomenon.

Given an image, the gray level is distributed $0 \le r \le 1$ after normalization. For any r within [0,1] interval, it can be transformed as:

$$s = T(r) \tag{5}$$

The transformation function T(r) should satisfy two conditions: firstly, it increases monotonously in [0,1] interval to ensure the grayscale of the image changes from white to black orderly. Secondly, after the mapping transformation, the value of *s* must be guaranteed within the [0,1] interval. The discrete form of *r*'s probability density function is as follows:

$$P_r(r_k) = \frac{n_k}{n} \tag{6}$$

Among them, $P_r(r_k)$ represents the distribution of image gray level, r_k represents the discrete gray level, where $0 \le r_k \le 1$, k = 0, 1, 2..., n-1. *n* denotes the total number of

pixels, and n_k denote the number of r_k appearing in the image, respectively. Image histogram equalization can be expressed as:

$$S_r = T(r_i) = \sum_{i=0}^k \frac{n_i}{n} \tag{7}$$

where k is the gray level, $0 \le r_k \le 1$, k = 0, 1, 2..., n-1. The original FKP image and FKP image after histogram equalization is shown in Fig. 3(a) and (b).

By setting different thresholds (low-in, high-in, low-out and high-out) and using the transformation function (T(r)), the amount of the training images will be 360 times more than the original database. The values that are smaller than the low-in values will be mapped to low-out, while the values that bigger than high-in values will be mapped to high-out, and the values between low-in and high-in will be transformed into [low-out, high-out] by T(r).

5 Experimental Results

The experiment was configured with on a desktop computer with Intel Xeon E5-2620v2, 2.1 GHz CPUs, 80 GB RAM, a single NVIDIA Tesla K20c, on Windows 10 operation system. The training and testing of the proposed CNN model are based on the publicly available Caffe Library [29].

5.1 PolyU Database

Extensive experiment is performed on PolyU FKP database, which shows the performance of the proposed CNN architecture. The database's images were obtained by collecting finger images from 165 volunteers, with a ratio of nearly 3:1 for male and female. Among them, the proportion of people aged 20–30 and 30–50 years is close to 7:1. All images are collected in two sessions. In each session, the volunteer's left index, left middle, right index and right middle images were collected, each using 6 images, respectively. Overall, the database contained 7920 images of 660 different fingers from 165 volunteers. The average finger image collection time interval from the first to the second session was 25 days. Among them, the maximum and minimum collected time intervals between two sessions were 196 days and 14 days, respectively. After ROI extraction and data augmentation, the final image size was 220×110 . Here, we use the histogram equalization mentioned in Sect. 4.2 to augment the data.

5.2 Configuration of Training Parameters

Training parameters are set as follows:

(1) prepare data sets, and divide the PolyU database into 6:1:5 as training, validation, and test data sets, respectively. More specifically, the training set randomly selected 6 images from each subject. The same method is used to complete the preparation of the validation sets and test sets.

- (2) we set the momentum to 0.9, the weight decay of all the convolution layers and the first fully connected layer is 5e-4, and the fixed learning rate is 0.001. Randomly initialized the tunable network parameters and started the training of the network.
- (3) during the training process, when the verification accuracy is no longer trending upwards, then reducing the learning rate to continue training until the verification accuracy is no longer increasing.
- (4) the test results were obtained by selecting the model with the highest verification accuracy and the least loss.
- (5) we set the dropout rate to 0.5, which gets the best performance.

5.3 FKP Experimental Results

The recognition results are shown in Table 2, where LI denotes left index finger, LM denotes left middle finger, RI denotes right index finger and RM denotes right middle finger.

Recognition method	Finger			
	LI	LM	RI	RM
	TTR(%)	TTR(%)	TTR(%)	TTR(%)
BP [30]	92.6	92.5	93.2	93.0
Ordinal Code [31]	97.3	95.9	96.3	95.9
BOCV [32]	97.6	97.6	97.6	97.7
RLOC [33]	97.8	97.7	97.9	97.7
SRC [34]	98.7	98.7	98.4	98.0
Comp Code [35]	98.0	98.0	98.2	98.1
AlexNet	85.6	86.5	85.2	85.2
AlexNet with AD	88.3	90.2	86.8	89.9
Batch-normalized CNN	89.7	92.1	87.1	91.3
Batch-normalized CNN with AD	99.1	98.9	99.4	98.3

Table 2. Comparison results of proposed FKP recognition algorithms and other algorithms.

To verify the effect of the augmented data on the proposed CNN architecture, the contrast experiment was added. Experimental results show the effectiveness of the proposed CNN on augmentation data used in this paper. The training of network model is relying on the scale of training data; hence, increasing the training sample can improve network testing performance.

From the experimental results, we know that Comp Code accuracy is 98.0%, 98.0%, 98.2% and 98.1%, respectively. The proposed CNN on the proposed augmented data can obtain the state-of-the-art recognition accuracy of 99.1%, 98.9%, 99.4% and 98.3%, correspondingly. Compared with the state-of-the-art methods, such as AlexNet, BP [30], Ordinal Code [31], BOCV [32], RLOC [33], SRC [34] and Comp Code [35], the structure proposed in this paper has achieved the best results.

6 Conclusion

In this paper, we proposed a novel batch-normalized CNN for FKP recognition. A data augmentation method of random histogram equalization and a dropout layer were adopted to prevent overfitting during training in the proposed neural network architecture. Experimental results on finger-knuckle-print database established by the PolyU show that the batch-normalized CNN could achieved satisfying results in recognizing finger-knuckle-print.

Acknowledgments. This work is supported by National of Nature Science Foundation Grant (No. 61372193, No. 61771347), Guangdong Higher Education Outstanding Young Teachers Training Program Grant (No. SYQ2014001), Characteristic Innovation Project of Guangdong Province (No. 2015KTSCX 143, 2015KTSCX145, 2015KTSCX148), Youth Innovation Talent Project of Guangdong Province (No. 2015KQNCX172, No. 2016KQNCX171), Science and Technology Project of Jiangmen City (No. 201501003001556, No. 201601003002191), and China National Oversea Study Scholarship Fund.

References

- Fei, L., Wen, J., Zhang, Z., et al.: Local multiple directional pattern of palmprint image, 3013–3018. In: The 23rd International Conference on Pattern Recognition (ICPR) (2016)
- 2. Bapat, A., Kanhangad, V.: Segmentation of hand from cluttered backgrounds for hand geometry biometrics. In: IEEE Region 10 Symposium (TENSYMP), pp. 1–4 (2017)
- Chatterjee, A., Bhatia, V., Prakash, S.: Anti-spoof touchless 3D fingerprint recognition system using single shot fringe projection and biospeckle analysis. Opt. Lasers Eng. 95, 1–7 (2017)
- Huang, D., Zhang, R., Yin, Y., et al.: Local feature approach to dorsal hand vein recognition by centroid-based circular key-point grid and fine-grained matching. Image Vis. Comput. 58, 266–277 (2017)
- Zhang, L., Zhang, L., Zhang, D., et al.: Online finger-knuckle-print verification for personal authentication. Pattern Recogn. 43(7), 2560–2571 (2010)
- Zhang, L., Zhang, L., Zhang, D., et al.: Ensemble of local and global information for fingerknuckle-print recognition. Pattern Recogn. 44(9), 1990–1998 (2011)
- Kumar, A., Ravikanth, C.: Personal authentication using finger knuckle surface. IEEE Trans. Inf. Forensics Secur. 4(1), 98–110 (2009)
- Woodard, D.L, Flynn, P.J.: Personal identification utilizing finger surface features. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), vol. 2, pp. 1030–1036 (2005)
- Woodard, D.L., Flynn, P.J.: Finger surface as a biometric identifier. Comput. Vis. Image Underst. 100(3), 357–384 (2005)
- Ravikanth, C., Kumar, A.: Biometric authentication using finger-back surface. In: IEEE Conference on Computer Vision and Pattern Recognition, pp. 1–6 (2007)
- Kumar, A., Ravikanth, C.: Personal authentication using finger knuckle surface. IEEE Trans. Inf. Forensics Secur. 4(1), 98–110 (2009)
- Zhang, L., Zhang, L., Zhang, D.: Finger-knuckle-print: a new biometric identifier. In: IEEE International Conference on Image Processing (ICIP), pp. 1981–1984 (2009)

- Zhang, L., Zhang, L., Zhang, D.: Finger-knuckle-print verification based on band-limited phase-only correlation. In: Jiang, X., Petkov, N. (eds.) CAIP 2009. LNCS, vol. 5702, pp. 141–148. Springer, Heidelberg (2009). https://doi.org/10.1007/978-3-642-03767-2_17
- 14. Morales, A., Travieso, C.M., Ferrer, M.A., et al.: Improved finger-knuckle-print authentication based on orientation enhancement. Electron. Lett. **47**(6), 380–381 (2011)
- Le, Z.: Finger knuckle print recognition based on surf algorithm. In: The Eighth International Conference on Fuzzy Systems and Knowledge Discovery (FSKD), vol. 3, pp. 1879–1883 (2011)
- Badrinath, G.S., Nigam, A., Gupta, P.: An efficient finger-knuckle-print based recognition system fusing SIFT and SURF matching scores. In: Qing, S., Susilo, W., Wang, G., Liu, D. (eds.) ICICS 2011. LNCS, vol. 7043, pp. 374–387. Springer, Heidelberg (2011). https://doi. org/10.1007/978-3-642-25243-3_30
- Li, Z., Wang, K., Zuo, W.: Finger-knuckle-print recognition using local orientation feature based on steerable filter. In: Huang, D.-S., Gupta, P., Zhang, X., Premaratne, P. (eds.) ICIC 2012. CCIS, vol. 304, pp. 224–230. Springer, Heidelberg (2012). https://doi.org/10.1007/ 978-3-642-31837-5_33
- Yang, W., Sun, C., Zhang, L.: A multi-manifold discriminant analysis method for image feature extraction. Pattern Recogn. 44(8), 1649–1657 (2011)
- Zhang, L., Li, H.: Encoding local image patterns using riesz transforms: with applications to palmprint and finger-knuckle-print recognition. Image Vis. Comput. 30(12), 1043–1051 (2012)
- Krizhevsky, A., Sutskever, I., Hinton, G.E.: Imagenet classification with deep convolutional neural networks. In: Advances in Neural Information Processing Systems, pp. 1097–1105 (2012)
- Ioffe, S., Szegedy, C.: Batch normalization: accelerating deep network training by reducing internal covariate shift. In: International Conference on Machine Learning (ICML), pp. 448– 456 (2015)
- Yu, H., Yang, G., Wang, Z., et al.: A new finger-knuckle-print ROI extraction method based on two-stage center point detection. Int. J. Sig. Process. Image Proc. Pattern Recogn. 8(2), 185–200 (2015)
- Meraoumia, A., Chitroub, S., Bouridane, A.: Palmprint and finger-knuckle-print for efficient person recognition based on log-gabor filter response. Analog Integr. Circ. Sig. Process 69(1), 17–27 (2011)
- Xiong, M., Yang, W., Sun, C.: Finger-knuckle-print recognition using LGBP. In: Liu, D., Zhang, H., Polycarpou, M., Alippi, C., He, H. (eds.) ISNN 2011. LNCS, vol. 6676, pp. 270– 277. Springer, Heidelberg (2011). https://doi.org/10.1007/978-3-642-21090-7_32
- 25. PolyU Finger-Knuckle-Print Database: http://www.comp.polyu.edu.hk/~biometrics (2010)
- Kirthiga, R., Ramesh, G.: Efficient FKP based recognition system using k-mean clustering for security system. SSRG Int. J. Electron. Commun. Eng. 1–7 (2016)
- Konda, K., Bouthillier, X., Memisevic, R., et al.: Dropout as data augmentation. Computer. Science 1050, 29 (2015)
- Srivastava, N., Hinton, G., Krizhevsky, A., et al.: Dropout: a simple way to prevent neural networks from overfitting. J. Mach. Learn. Res. 15(1), 1929–1958 (2014)
- Jia, Y., Shelhamer, E., Donahue, J., et al.: Caffe: convolutional architecture for fast feature embedding. In: Proceedings of the 22nd ACM International Conference on Multimedia, pp. 675–678 (2014)
- Leung, H., Haykin, S.: The complex backpropagation algorithm. IEEE Trans. Signal Process. 39(9), 2101–2104 (1991)

- Sun, Z., Tan, T., Wang, Y., et al. Ordinal palmprint representation for personal identification. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), vol. 1, pp. 279– 284 (2005)
- Guo, Z., Zhang, D., Zhang, L., et al.: Palmprint verification using binary orientation cooccurrence vector. Pattern Recogn. Lett. 30(13), 1219–1227 (2009)
- Jia, W., Huang, D.S., Zhang, D.: Palmprint verification based on robust line orientation code. Pattern Recogn. 41(5), 1504–1513 (2008)
- 34. Wright, J., Yang, A.Y., Ganesh, A., et al.: Robust face recognition via sparse representation. IEEE Trans. Pattern Anal. Mach. Intell. **31**(2), 210–227 (2009)
- 35. Kong, A.W.K., Zhang, D.: Competitive coding scheme for palmprint verification. In: The 17th International Conference on Pattern Recognition (ICIR), vol. 1, pp. 520–523 (2004)