SCIENCE CHINA
Information Sciences

• LETTER •

January 2019, Vol. 62 019104:1–019104:3
https://doi.org/10.1007/s11432-018-9512-1

Learning discriminative and invariant representation for fingerprint retrieval

Dehua SONG*, Ruilin LI, Fandong ZHANG & Jufu FENG*

Key Laboratory of Machine Perception (Ministry of Education), Department of Machine Intelligence, School of Electronics Engineering and Computer Science, Peking University, Beijing 100871*, China*

Received 27 April 2018/Accepted 30 May 2018/Published online 19 December 2018

Citation Song D H, Li R L, Zhang F D, et al. Learning discriminative and invariant representation for fingerprint retrieval. Sci China Inf Sci, 2019, 62(1): 019104, https://doi.org/10.1007/s11432-018-9512-1

Dear editor,
Fingerprints are widely used in civil and forensic recognition systems owing to their uniqueness and immutability. However, the size of a modern gallery database has exceeded billions hitherto. It is challenging to identify a query fingerprint against a large-scale database [1]. In a fingerprint identification system, a query fingerprint needs to be compared with all the fingerprints in the gallery database by a one-to-one matching algorithm. The accuracy and efficiency of the fingerprint identification system deteriorate rapidly as the size of the gallery database increases. To reduce the total number of comparisons, the fingerprint identification system employs pre-filtering techniques. The most conventional technique is fingerprint retrieval, which characterizes each fingerprint with a multidimensional vector summarizing its primary features. In the search phase, the feature vector is utilized to pre-filter a large number of fingerprints with low similarity to the query fingerprint. Subsequently, the query fingerprint is identified from the remaining candidate hypotheses by an automated one-to-one matching algorithm.

Feature representation is the most important part of designing a fingerprint retrieval system, which greatly affects the accuracy and efficiency of the system. However, fingerprint distortion is a challenging problem for feature representation. When a fingerprint is captured, the nonuniform finger pressure applied by the subject and the inherent elasticity of the skin cause linear and nonlinear distortions in the ridge structure. To solve this problem, researchers have proposed various human-crafted features [2] and similarity measurement methods [3]. However, they could not achieve the satisfactory accuracy for fingerprint retrieval owing to the complexity of nonlinear distortion. The essential requirement of fingerprint retrieval is that the similarity between intra-class samples is greater than that between inter-class samples. Hence, to solve this problem, a natural idea is to employ metric learning and deep convolutional neural network (DCNN) to learn a discriminative and distortion invariant representation for fingerprint retrieval. DCNN possesses strong expression ability.

The most challenging problem in applying DCNN to fingerprint retrieval is the scarcity of training data, which limits the depth and width of the DCNN. By considering the scarcity of the training data and the fingerprint distortion, this study contributes in the following two aspects:

(1) A novel power activation function is proposed herein to learn a discriminative and distortion-invariant representation. It is nonlinear on the whole domain of the definition.

(2) A slight shear transformation and a distortion modeling method [1] are employed to augment the training data. This strategy ensures that the DCNN can learn the distortion variations of intra-class fingerprints.

To improve the retrieval accuracy, each fingerprint is represented by a single feature vector com-

* Corresponding author (email: songdehua@pku.edu.cn, fjf@cis.pku.edu.cn)

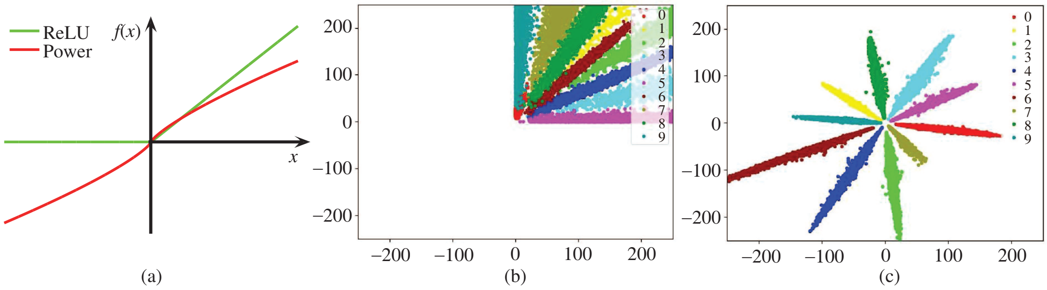

Figure 1 (Color online) The distribution of learned deep convolutional features $\boldsymbol{f}_c$ (without normalization) in fingerprint training data. A two-dimensional feature $\boldsymbol{f}_c$ is learned by the DCNN with ten fingerprint classes. (a) Activation functions; (b) the distribution of features with rectified linear unit (ReLU); (c) the distribution of features with power function.

bined with multiple patch features. The single feature vector is named the multiple patches convolutional (MPC) feature. Extensive experimental results reveal that the MPC method achieves better performance than other prominent approaches.

Method. To improve the accuracy and efficiency of fingerprint retrieval, we propose a novel MPC feature, which combines six different patch features to capture more detailed information of the fingerprint. Each patch feature is learned by the DCNN with metric loss to ensure some degree of translation, rotation, and distortion invariance.

(1) Preprocessing. The ridge structures of the fingerprint provide discriminatory information for retrieval. However, many reasons can result in a poor quality image that is marked by low contrast and ill-defined boundaries between the ridges, such as bruises, dry fingers, and sweat. Poor quality fingerprint images may prevent the DCNN from learning the ridge structures of the fingerprints. Therefore, FingerNet [4] is utilized to enhance the ridge structure and filter out noise.

(2) Architecture of our DCNN. Our DCNN contains ten learned layers: eight convolutional layers, a fully connected layer, and a softmax layer. Six convolutional layers exist that are followed by a max pooling layer. Unlike the typical DCNN, the size of the fully connected layer is designed as small as possible to ensure that the learned representation is compact. The activation values of the fully connected layer are the learned feature, which is viewed as deep convolutional feature $\boldsymbol{f}_c$. Subsequently, mean subtraction and l_2-normalization are employed on the deep convolutional feature $\boldsymbol{f}_c$. The normalized deep convolutional (NDC) feature $\boldsymbol{f}_N$ is defined as

$$\boldsymbol{f}_N = \frac{\boldsymbol{f}_c - \bar{\boldsymbol{f}}_c}{\|\boldsymbol{f}_c - \bar{\boldsymbol{f}}_c\|_2}, \tag{1}$$

where $\bar{\boldsymbol{f}}_c$ denotes the mean of feature $\boldsymbol{f}_c$.

(3) Activation function. Efficiency is crucial to fingerprint retrieval, in which a compact feature is required. Additionally, the scarcity of training data limits the DCNN depth. However, we hope that the learned feature should be discriminative and invariant to distortion. Hence, we designed a novel power activation function to accomplish the representation learning task. The nonsaturated ReLU [5], $f(x) = \max(0, x)$, is widely utilized in the DCNN. Unfortunately, for feature representation, the ReLU weakens the discriminability of the feature, because it truncates the negative part of the feature vector to zero. This is done by assuming that the dimension of the feature is n, and the features that have not passed through the ReLU are distributed in the whole n-dimensional space. However, the ReLU compresses the distribution space of the features to the space enclosed by the positive parts of all the coordinate axes. Consequently, the distribution space of the features is decreased to $\frac{1}{2^n}$ of the whole n-dimensional space. It causes all the fingerprint classes to distribute more densely, which complicates the identification of different fingerprints. Figure 1 illustrates this phenomenon. Despite the parametric rectified linear units (PReLU) [6], $f(x) = \max(0, x) + a\min(0, x)$, can solve this problem to some extent by retaining the negative part, the nonlinearity of the function is only on the origin of the coordinates. The nonlinearity of the activation function is crucial to neural networks fitted for intra-class variations. Hence, this study proposes the power function defined in Eq. (2) to enhance the expression ability of the DCNN. The power activation function has three admirable properties:

- Central symmetry. Power function does not only retain the negative part of the input but also is central symmetry. It enables the DCNN to distribute the centroids of the fingerprint classes uniformly in the whole space, which contributes to learning a discriminative representation.

- Nonlinearity on the whole domain of definition. It is helpful to learn an invariant representation for nonlinear distortion.

- Nonsaturated function. It assists the DCNN to alleviate the vanishing gradient problem.

$$f(x) = x^{\frac{p}{q}}, \quad q > p > 0, \tag{2}$$

where x is the input of the activation function. p and q are fixed odd numbers that ensure the power function is odd. The exponential number p/q is searched in the range $(0, 1)$. The power activation function is applied on the fully connected layer to enhance the feature discriminability. To accelerate the convergence speed, the activation function of the convolutional layer is set to the PReLU.

(4) Training. We employed four different forms of data augmentation: cropping, rotation, slight shear transformation and nonlinear distortion [1]. To ensure the distortion variance in the training data, we utilized a slight shear transformation to simulate the linear distortion and the distortion model method to simulate the nonlinear distortion. To learn the discriminative and invariant representation, we combined the softmax loss and cosine loss [7] as the metric loss. It utilizes the softmax loss to constrain the inter-class discrimination and cosine center loss to constrain the intra-class invariance. The whole loss function is defined in Eq. (3). Each finger is considered as a class. To accelerate the convergence speed, the DCNN is first pretrained with the PReLU on all layers. Subsequently, we fine-tuned the DCNN with the power function on the fully connected layer.

$$\mathcal{L} = \mathcal{L}_{\mathcal{S}} + \lambda \mathcal{L}_{\mathcal{C}}, \tag{3}$$

where $\mathcal{L}_{\mathcal{S}}$ denotes the softmax loss. λ is the coefficient of the cosine center loss $\mathcal{L}_{\mathcal{C}}$.

(5) Retrieval system. Each fingerprint is represented by six patch features $\boldsymbol{f}_N$. For retrieval efficiency, these normalized features are orderly concatenated with parameter η to construct a fixed-length feature vector $\boldsymbol{f}_M$, named the MPC feature, which is defined in Eq. (4). The similarity between query fingerprint feature and gallery fingerprint features is calculated in accordance with Eq. (5). Finally, the top N_C-ranked fingerprints are selected as the candidates by sorting the gallery fingerprints according to the similarity.

$$\boldsymbol{f}_M = [\eta_1 \boldsymbol{f}_{N_1}{}^{\mathrm{T}}, \eta_2 \boldsymbol{f}_{N_2}{}^{\mathrm{T}}, \ldots, \eta_P \boldsymbol{f}_{N_P}{}^{\mathrm{T}}]^{\mathrm{T}}, \tag{4}$$

$$\mathrm{Sim}(\boldsymbol{f}_{M_1}, \boldsymbol{f}_{M_2}) = \langle \boldsymbol{f}_{M_1}, \boldsymbol{f}_{M_2} \rangle, \tag{5}$$

where $\sum_{p=1}^{P} \eta_p^2 = 1$. P denotes the number of patches in one image. $\boldsymbol{f}_{M_1}$ and $\boldsymbol{f}_{M_2}$ denote two MPC feature vectors. $\langle \cdot, \cdot \rangle$ denotes scalar product.

Performance evaluation. Comprehensive experiments were conducted to evaluate the retrieval performance of the proposed method on four benchmark databases. The experimental results in Appendix A reveal that the proposed MPC approach achieves a lower error rate than other prominent methods [8]. This is primarily because the power activation function enables the DCNN to learn a more discriminative and invariant representation of distortion. The details of the experiments are provided in the supporting information.

Conclusion. Herein, a novel fingerprint retrieval approach based on representation learning is proposed to improve the accuracy and efficiency of retrieval in large databases. By considering the scarcity of training data and fingerprint distortion, this study proposes a novel power activation function that enables the DCNN to learn a discriminative and invariant representation. Additionally, we employ distortion generation methods to augment the training data. Our experimental results reveal that the proposed MPC method achieves better performance than other prominent approaches on four benchmark databases.

Acknowledgements This work was supported by National Natural Science Foundation of China (Grant No. 61333015).

Supporting information Appendixes A–D. The supporting information is available online at info.scichina.com and link.springer.com. The supporting materials are published as submitted, without typesetting or editing. The responsibility for scientific accuracy and content remains entirely with the authors.

References

1 Maltoni D, Maio D, Jain A K, et al. Handbook of Fingerprint Recognition. Berlin: Springer, 2003
2 Gago-Alonso A, Hernández-Palancar J, Rodríguez-Reina E, et al. Indexing and retrieving in fingerprint databases under structural distortions. Expert Syst Appl, 2013, 40: 2858–2871
3 Chen X J, Tian J, Yang X. A new algorithm for distorted fingerprints matching based on normalized fuzzy similarity measure. IEEE Trans Image Process, 2006, 15: 767–776
4 Tang Y, Gao F, Feng J, et al. FingerNet: an unified deep network for fingerprint minutiae extraction. In: Proceedings of IEEE International Joint Conference on Biometrics, 2017. 108–116
5 Krizhevsky A, Sutskever I, Hinton G E. ImageNet classification with deep convolutional neural networks. In: Proceedings of Annual Conference on Neural Information Processing Systems, 2012. 1097–1105
6 He K, Zhang X, Ren S, et al. Delving deep into rectifiers: surpassing human-level performance on imagenet classification. In: Proceedings of the IEEE International Conference on Computer Vision, 2015. 1026–1034
7 Wang J, Li Y, Miao Z, et al. Learning deep discriminative features based on cosine loss function. Electron Lett, 2017, 53: 918–920
8 Song D, Feng J. Fingerprint indexing based on pyramid deep convolutional feature. In: Proceedings of IEEE International Joint Conference on Biometrics, 2017. 200–207