CROSS-DOMAIN PALMPRINT RECOGNITION BASED ON TRANSFER CONVOLUTIONAL AUTOENCODER

Huikai Shao[1], *Dexing Zhong*[1], *Xuefeng Du*[1]

[1]School of Electronic and Information Engineering, Xi'an Jiaotong University,
28 West Xianning Road, Xi'an, Shaanxi 710049, P.R. China

ABSTRACT

Recently, excellent palmprint recognition algorithms have emerged and achieved satisfactory performance. However, cross-domain palmprint recognition is rarely considered. In this paper, we proposed transfer autoencoder for cross-domain palmprint recognition. Convolutional autoencoders were firstly used to extract low-dimensional features. A discriminator was then introduced to reduce the gap of two domains. The autoencoders and discriminator were alternately trained, and finally the features with the same distribution were extracted. The databases collected from different environments are defined as source and target domains, respectively. Based on the labels in source domain, unsupervised identification of target domain can be achieved. The experiments were performed on 24 cross-domain pairs composed by multispectral database and our self-built uncontrolled databases. The results show that transfer autoencoder can greatly improve cross-domain recognition accuracy, up to 23.26%. At the same time, the accuracy in a single domain can reach over 99% in controlled database.

Index Terms— palmprint recognition, cross-domain, transfer autoencoder, uncontrolled database

1. INTRODUCTION

Nowadays, information security has received extensive attention. Biometrics-based identification is superior to traditional identification technologies such as keys and passwords. Palmprint recognition, a contactless biometrics, has the advantages of high security, convenience, and user friendliness. It consists of four steps: image acquisition, image preprocessing, feature extraction, and matching [1]. The image acquisition mainly captures palmprint images through the camera, and many popular public databases have emerged, such as PolyU [2] and IITD [3]. The image preprocessing includes filtering and extracting regions of interest (ROI) [4]. The extraction of ROI refers to obtaining the main palmprint area without background. The feature extraction mainly extracts texture and directional features of palmprint [5]. In feature matching, Euclidean distance [6] or Hamming distance [7] between the features are calculated to identify palmprint.

However, the current palmprint recognition algorithms are mainly based on a single domain, *i.e.* both the training set and the test set belong to the same database. When a model trained on a unique database is used directly for identification in another different database, the performance will be greatly degraded. This brings a lot of inconvenience to palmprint recognition. For example, in practical applications, it is likely that samples of registration and identification will be collected by different devices. One way to solve this problem is to adopt transfer learning approaches. Transfer learning takes the knowledge learned in one domain into another domain to implement domain adaptation [8].

In this paper, we proposed transfer autoencoder for cross-domain palmprint recognition, and the flow chart is shown in Fig. 1. Two different palmprint databases are selected as the source domain and target domain, where the source domain has labels and the target domain has no labels. Firstly, convolutional autoencoders were adopted to extract the low-dimensional palmprint features in source and target domains, which were used for identification. Then a discriminator was introduced to reduce the gap between two domains. The autoencoders and discriminator were trained in turn. In fact, the reason that the model trained in source domain cannot be used for target domain is the feature distributions of two domains are different. Using adversarial adaptive model, *i.e.* transfer autoencoder, the distance between the empirical source and target mapping distributions is minimized [9]. Therefore, the unsupervised identification in target domain can be achieved using the information of source domain.

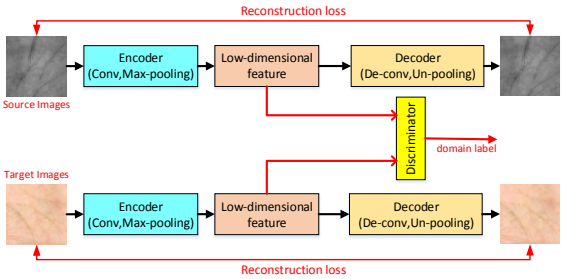


Fig. 1. An overview of our proposed transfer autoencoder. The images in source domain and target domain are respectively input to convolutional autoencoders, and both discriminative features are obtained by optimizing the reconstruction loss. The discriminator is then adopted to reduce the gap between two domains. Finally, the autoencoder and discriminator are trained in turn.

The contributions can be summarized as follows:

1） Transfer autoencoder was proposed for cross-domain palmprint recognition. The autoencoder can extract discriminative low-dimensional features, while the gap between two domains was reduced by the discriminator.

2） We built two databases using mobile phones, which were divided into four independent sub-databases. They were collected in uncontrolled environments, which was more relevant to the actual scenario.

3） There were 24 cross-domain pairs composed by several popular databases and our self-built databases, and cross-domain experiments were performed. The results showed the high efficiency of transfer autoencoder, and the cross-domain recognition accuracy increased by up to 23.26%. At the same time, the recognition accuracy in a single domain was also very high.

The paper consists of five sections. Section 2 introduces some related works. Section 3 explains our methods in details. The experiments and results are presented in section 4. Section 5 concludes the paper.

2. RELATED WORK

Palmprint recognition mainly includes encoding, photo, and learning-based approaches [4]. Zhang *et al.* [10] adopted a Gabor filter to extract a certain orientation feature of palmprint images, named palmcode. After that, orientation-based coding methods are successfully developed, such as competitive code [11] and robust line orientation code (RLOC) [12]. Zhong *et al*. [13] proposed deep hashing network for palmrpint recognition. Izadpanahkakhk *et al.* [14] extracted ROI and discriminative features for palmprint recognition using a neural network and transfer learning fusion method. Recent progress in cross-device palmprint recognition only focused on algorithms under each independent application environment. Jia *et.al*. [15] collected palmprint images using one digital camera and two smart-phones and proposed a method to calculate the palm width for scale normalization. They only tested separate recognition performance using images from each device. However, the transferring recognition task was not solved.

Nowadays transfer learning has thrived in the area of deep learning. The first type is fine tuning [16] that only trains the last couple of layers of a neural network while keeping other parameters fixed when facing a new situation. The others are feature-based transferring methods, which aim to decrease the difference between source and target domains in order to learn a good representation in the target domain. Long *et.al*. [17] proposed a framework that embeds the deep features of all task-specific layers to reproducing kernel Hilbert spaces (RKHSs). Cycle-Consistent Adversarial Domain Adaptation (CyCADA) was proposed by Hoffman *et.al.* [18], which adapts representations at both the pixel-level and feature-level, enforces cycle-consistency while leveraging a task loss, and does not require aligned pairs. Based on a simple linear autoencoder, Xiao *et.al.* [19] used Maximum Mean Discrepancy (MMD) to derive the latent representations of both domains. Our method combines autoencoder and adversarial domain adaptation, which decreases the difference of feature distribution in the latent space. As a result, based on source domain, unsupervised palmprint recognition in the target domain can be fulfilled.

3. METHOD

Two palmprint databases collected from different devices and conditions are referred to as the source domain and target domain, respectively. The samples from the source domain are defined as $X_S = \left\{x_{S_1}, \dots, x_{S_{n_1}}\right\} \in \mathbb{R}^{d \times n_1}$, whose labels are known and defined as $Y_S = \left\{y_{S_1}, \dots, y_{S_{n_1}}\right\} \in \mathbb{R}^{n_1}$. The samples from the target domain are defined as $X_T = \left\{x_{T_1}, \dots, x_{T_{n_1}}\right\} \in \mathbb{R}^{d \times n_2}$, which have no labels. Here, d is the dimension of each image, while n_1 and n_2 are the number of samples in the source and target domains respectively. Our goal is to predict the labels for target domain based on X_S and Y_S. Transfer autoencoder consists of two parts: convolutional autoencoders extracting palmprint features and adversarial domain adaptation using a discriminator to reduce the gaps between two domains. The detailed structures are shown in Fig. 2.

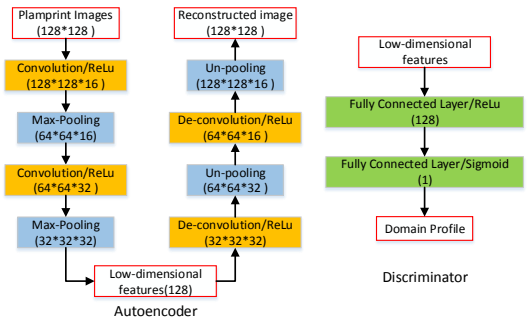


Fig. 2. Architecture of autoencoder and discriminator. The output of each layer is represented in the box. ReLu means the activation function is Rectified Linear Unit.

3.1. Convolutional Autoencoder

The convolutional autoencoder contains encoder and decoder with low information loss and simple structure. The encoder is used to convert the input image $x \in \{X_S, X_T\}$ into low-dimensional features $F(x)$, which results in a loss of information. The decoder projects the feature back to the original space to obtain a reconstructed image $x' \in \{X'_S, X'_T\}$. By constraining difference between the reconstructed image and the original image, the lost information can be minimized and the discriminative features can be extracted. Here, the reconstruction loss of source domain and target domain can be expressed as:

$$\mathcal{L}_{r_S} = \|x'_S - x_S\|^2 \quad (1)$$

$$\mathcal{L}_{r_T} = \|x'_T - x_T\|^2 \quad (2)$$

Compared with the traditional linear autoencoder, the convolutional autoencoder retains its original property, *i.e.* unsupervised learning, but it combines the convolution and pooling operations to fully extract the unique features of palmprint images. In this paper, the encoder consists of two convolutional layers and two maximum pooling layers; the decoder consists of two deconvolution layers and two un-pooling layers. Finally, 128-dimensional features were obtained for recognition.

3.2. Adversarial Domain Adaptation

The gap between the source and target domains results in different distributions of extracted features. In order to reduce the domain gap, a discriminator, D, was introduced. As shown in Fig. 1, in training, low-dimensional feature codes $F(x)$ from the source and target domains were respectively input into the discriminator. Then, the discriminator was used to identify which domain the input features came from. The autoencoder and discriminator were alternately trained. Based on the idea of Generative Adversarial Nets (GAN) [20], the convolutional autoencoder attempts to optimize the extracted features so that the discriminator cannot distinguish which domain the features belong to, while the discriminator is trying to accurately distinguish which domain the input comes from. Finally, the autoencoders in the two domains extract features that are indistinguishable by the discriminator, which proves that the feature distributions of two domains become consistent. As a result, the classifier trained by X_S and Y_S can be directly used for the target domain. Thus adversarial domain adaptation corresponds to the following unconstrained optimization.

$$\min_{F_S,F_T} \mathcal{L}_F(D, X_S, X_T) = -\mathbb{E}_{x_T \sim X_T}\left[\log\left(1 - D\left(F_T(x_T)\right)\right)\right] \quad (3)$$

$$\min_{D} \mathcal{L}_D(F_S, F_T, X_S, X_T) = -\mathbb{E}_{x_S \sim X_S}\left[\log\left(D\left(F_S(x_S)\right)\right)\right] - \mathbb{E}_{x_T \sim X_T}\left[\log\left(D\left(1 - F_T(x_T)\right)\right)\right] \quad (4)$$

To sum up, the loss function for training the autoencoders is given in the following Equation:

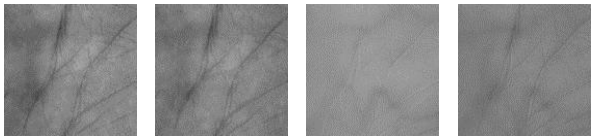
$$\mathcal{L}_{ae} = \alpha\mathcal{L}_F(D, X_S, X_T) + \beta\mathcal{L}_{r_S} + \gamma\mathcal{L}_{r_T} \quad (5)$$

Where α, β, and γ are trade-off parameters

4. EXPERIMENTS AND RESULTS

4.1. Database

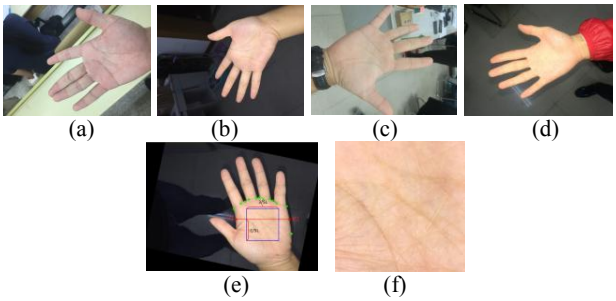
Multispectral palmprint database: Multispectral palmprint database was established under blue, red, green, and near-infrared (NIR) illuminations [21]. Under each illumination, palmprint images of 250 volunteers were collected in two sessions. In a session, each individual was asked to provide 6 images from a palm. According to different illuminations, the multispectral database can be divided into four spectrum databases, denoted as Blue (B), Green (G), NIR (N), and Red (R), which can be used as four different domains. Therefore, in each sub-database, a total of 6,000 images were divided into 500 categories, and some samples are shown in Fig. 3. In this paper, two different domains from B, G, N, and R were randomly selected to form 12 cross-domain pairs, e.g., B→N, B→G, B→R, … , R→G.



(a) (b) (c) (d)

Fig. 3. Samples of multispectral palmprint database. (a) is in Blue, (b) is in Green, (c) is in NIR, and (d) is in Red.

Uncontrolled database: The database mentioned above was collected from a controlled environment. It is necessary to limit the postures of palms in a closed environment and add extra illuminations. It reduces the friendliness of palmprint recognition in practical applications. So we used iPhone6s and HUAWEI Mate8 to create new databases in uncontrolled environments. Indoors, volunteers hold mobile phones to determine the posture and angle of their palms according to their own wishes, and selected different scenes as the backgrounds of images. Two kinds of illuminations were adopted respectively, one is the natural light in the room, and the other is the flash of mobile phones. Therefore, there are a total of four databases, which can also be defined as four different domains, called iN (iPhone6s under Natural illumination), iF (iPhone6s under Flash illumination), HN (HUAWEI Mate8 under Natural illumination), and HF (HUAWEI Mate8 under Flash illumination). Using two mobile phones, each of total 98 volunteers was asked to capture 10 images of each palm under two illuminations. In summary, there are 1960 images from 196 categories in each independent database. ROIs of images collected from uncontrolled environments are difficult to extract due to the complex background. We manually marked 14 feature points, including the roots of the fingers, valleys, *etc.*, available to the academic community, as shown in Fig. 4.



(a) (b) (c) (d) (e) (f)

Fig. 4. Samples of uncontrolled palmprint database. (a) is in iN, (b) is in iF, (c) is in HN, and (d) is in HF. (e) is the setting of ROI extraction, (f) is a ROI extracted.

In this paper, we selected two points on the edge of palm (the red point in Fig. 4 (e)) as the reference points, and a square area, whose side length was three-fifths of the length between the two points, was extracted as ROI, as shown in Fig. 4 (f). Similarly, two databases were randomly selected to form cross-domain pairs, with a total of 12 pairs.

Table 1: recognition results on different databases

Database	Autoencoder (baseline)	Adversarial Autoencoder	Accuracy	
			Source	Target
B→N	80.63%	92.23%	99.87%	**99.87%**
B→G	85.93%	90.17%	99.47%	99.70%
B→R	78.50%	90.80%	99.87%	99.83%
N→B	85.57%	92.83%	99.10%	99.23%
N→G	88.17%	95.87%	99.73%	99.70%
N→R	79.47%	94.00%	99.86%	99.83%
G→B	80.27%	90.03%	99.83%	99.70%
G→N	89.30%	92.23%	99.63%	99.50%
G→R	82.00%	87.63%	**99.90%**	99.83%
R→B	89.87%	96.83%	99.30%	99.57%
R→N	**80.83%**	**96.63%**	99.87%	**99.87%**
R→G	77.07%	86.23%	99.67%	99.80%
Average	83.13%	92.12%	99.68%	99.70%
IF→IN	62.24%	71.02%	99.08%	92.55%
IF→HF	62.04%	78.78%	96.94%	97.96%
IF→HN	56.43%	62.55%	**99.29%**	89.49%
IN→IF	64.08%	74.29%	85.92%	97.86%
IN→HF	**59.39%**	**82.65%**	90.41%	98.06%
IN→HN	68.98%	73.67%	90.82%	84.39%
HF→IF	73.21%	82.52%	82.52%	92.23%
HF→IN	51.23%	66.77%	96.01%	91.62%
HF→HN	48.16%	68.61%	95.60%	87.53%
HN→IF	56.43%	62.27%	88.75%	96.22%
HN→IN	58.90%	68.40%	85.37%	91.10%
HN→HF	57.96%	64.69%	89.39%	**98.47%**
Average	59.92%	71.35%	91.67%	93.12%

4.2. Implementation details

Transfer autoencoder is used for cross-platform palmprint recognition. By optimizing, autoencoder can get 128-dimensional features for palmprint recognition. In training, half of each category in the source and target domains were used as training sets. In testing, the remaining half of samples were used as test sets to evaluate the performance. The test methods fall into two categories. Firstly, for cross-domain palmprint recognition, following the settings in [22], 1-NN was chosen as the base classifier. The source and target domains were as the training set and test set respectively. Secondly, palmprint recognition experiment was performed in each domain. For each image in the test set, the image that was most similar to it was found in the remaining images. If they were from the same class, the identification was accurate, otherwise, it was fail. The ratio of the number of correctly recognized images to the total number of images in the test set was calculated as the recognition accuracy. The model that used autoencoder for cross-domain recognition was adopted as a baseline. The entire experiments were performed using TensorFlow on a NVIDIA GPU GTX1080 with 6G memory power. Learning rate was set to 0.00001, and Adam optimization algorithm was adopted. The weights of different loss functions are three hyper parameters, specifically set to $\alpha = 0.5$, $\beta = 1$, and $\gamma = 1$.

4.3. Results and analysis

The experimental results in the multispectral database and uncontrolled databases are shown in Table 1. First of all, it can be seen from the results that transfer autoencoder can greatly improve the accuracy of cross-domain palmprint recognition compared to the autoencoder (baseline). In the multispectral database, in the best case, the accuracy rate was increased by 15.8% in R→N, and the average increase reached 8.99%. Similarly, for the uncontrolled databases, in the best case, the accuracy rate is increased by 23.26% in IN→HF, and even in the worst case, the accuracy is increased by 4.69% in IN→HN. It shows that the transfer autoencoder proposed can significantly reduce the gap between different domains and improve the accuracy of cross-domain palmprint recognition. Secondly, while the cross-domain recognition accuracy is greatly improved, the recognition accuracy in each domain can also reach a high level. In the Green of multispectral database, the recognition rate reached 99.90%. In the uncontrolled databases, due to the complexity of images, the recognition rate is relatively low, however the average accuracy is also over 91.67%.

5. CONCLUSION

In this paper, we proposed transfer autoencoder for cross-domain palmprint recognition. Transfer autoencoder contains two convolutional autoencoders and one discriminator. Convolutional autoencoders introduces convolutional layers and pooling layers on the basis of linear autoencoders. By optimizing the reconstruction loss, discriminative low-dimensional features are extracted and the loss of information is minimized. The discriminator is adopted to reduce the gap between source and target domains, so that the same feature distribution can be extracted. We designed 24 cross-domain pairs on controlled multispectral database and our self-built uncontrolled database. The experimental results show that, on the one hand, transfer autoencoder proposed can significantly improve the accuracy of cross-domain identification, with a maximum increase of 23.26%. On the other hand, transfer autoencoder can also achieve more than 99% recognition accuracy in a single domain, which fully demonstrates its efficiency in biometrics.

ACKNOWLEDGMENT

This work is supported by the National Natural Science Foundation of China (No. 61105021), and Natural Science Foundation of Zhejiang Province (No. LGF19F030002).

6. REFERENCES

[1] A. Kong, D. Zhang, and M. Kamel, "A survey of palmprint recognition," *Pattern Recognition,* Review vol. 42, no. 7, pp. 1408-1418, Jul 2009.

[2] K. Zhang, D. Huang, B. Zhang, and D. Zhang, "Improving texture analysis performance in biometrics by adjusting image sharpness," *Pattern Recognition,* Article vol. 66, pp. 16-25, Jun 2017.

[3] A. Kumar and S. Shekhar, "Personal Identification Using Multibiometrics Rank-Level Fusion," *IEEE Transactions on Systems Man and Cybernetics Part C-Applications and Reviews,* Article vol. 41, no. 5, pp. 743-752, Sep 2011.

[4] D. Zhong, X. Du, and K. Zhong, "Decade progress of palmprint recognition: A brief survey," *Neurocomputing,* vol. 1, no. 328, pp. 16-28, 2018.

[5] P. Dubey, T. Kanumuri, and R. Vyas, "Sequency codes for palmprint recognition," *Signal Image and Video Processing,* Article vol. 12, no. 4, pp. 677-684, May 2018.

[6] Y. Xu, Z. Fan, M. Qiu, D. Zhang, and J. Yang, "A sparse representation method of bimodal biometrics and palmprint recognition experiments," *Neurocomputing,* Article vol. 103, pp. 164-171, Mar 2013.

[7] Z. Guo, D. Zhang, L. Zhang, W. Zuo, and G. Lu, "Empirical study of light source selection for palmprint recognition," *Pattern Recognition Letters,* Article vol. 32, no. 2, pp. 120-126, Jan 2011.

[8] D. Cook, K. D. Feuz, and N. C. Krishnan, "Transfer learning for activity recognition: a survey," *Knowledge and Information Systems,* Article vol. 36, no. 3, pp. 537-556, Sep 2013.

[9] E. Tzeng, J. Hoffman, K. Saenko, T. Darrell, and Ieee, "Adversarial Discriminative Domain Adaptation," in *30th IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, HI, 2017, pp. 2962-2971.

[10] D. Zhang, W. Kong, J. You, and M. Wong, "Online palmprint identification," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* Article vol. 25, no. 9, pp. 1041-1050, Sep 2003.

[11] A. W. K. Kong and D. Zhang, "Competitive coding scheme for palmprint verification," in *17th International Conference on Pattern Recognition (ICPR)*, Cambridge, ENGLAND, 2004, pp. 520-523.

[12] W. Jia, D. Huang, and D. Zhang, "Palmprint verification based on robust line orientation code," *Pattern Recognition,* Article vol. 41, no. 5, pp. 1504-1513, May 2008.

[13] D. Zhong, H. Shao, and X. Du, "A Hand-based Multi-biometrics via Deep Hashing Network and Biometric Graph Matching," *IEEE Transactions on Information Forensics and Security,* 2019. DOI: 10.1109/TIFS.2019.2912552

[14] M. Izadpanahkakhk, S. M. Razavi, M. Taghipour-Gorjikolaie, S. H. Zahiri, and A. Uncini, "Deep Region of Interest and Feature Extraction Models for Palmprint Verification Using Convolutional Neural Networks Transfer Learning," *Applied Sciences-Basel,* Article vol. 8, no. 7, pp. 1-20, Jul 2018, Art. no. 1210.

[15] W. Jia, R. Hu, J. Gui, Y. Zhao, and X. Ren, "Palmprint Recognition across Different Devices," *Sensors,* vol. 12, no. 6, pp. 7938-7964, 2012.

[16] M. Wang and W. Deng, "Deep visual domain adaptation: A survey," *Neurocomputing,* vol. 312, pp. 135-153, OCT 2018.

[17] M. Long, Y. Cao, Z. Cao, J. Wang, and M. I. Jordan, "Transferable Representation Learning with Deep Adaptation Networks," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* pp. 1-14, 2018.

[18] J. Hoffman *et al.*, "CyCADA: Cycle-Consistent Adversarial Domain Adaptation," in *36th International Conference on Machine Learning (ICML 2018)*, California, USA, 2018, pp. 1994-2003.

[19] P. Xiao, B. Du, J. Wu, L. Zhang, R. Hu, and X. Li, "TLR: Transfer Latent Representation for Unsupervised Domain Adaptation," in *2018 IEEE International Conference on Multimedia and Expo (ICME)*, San Diego, CA, USA, 2018, pp. 1-6.

[20] I. J. Goodfellow *et al.*, "Generative Adversarial Nets," in *28th Conference on Neural Information Processing Systems (NIPS)*, Montreal, CANADA, 2014, pp. 1-9.

[21] D. Zhang, Z. Guo, G. Lu, L. Zhang, and W. Zuo, "An Online System of Multispectral Palmprint Verification," *IEEE Transactions on Instrumentation and Measurement,* Article vol. 59, no. 2, pp. 480-490, Feb 2010.

[22] B. Gong, Y. Shi, F. Sha, and K. Grauman, "Geodesic Flow Kernel for Unsupervised Domain Adaptation," in *25th IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Providence, RI, 2012, pp. 2066-2073.