Efficient Deep Palmprint Recognition via Distilled Hashing Coding

Huikai Shao[1], Dexing Zhong *[1,2], and Xuefeng Du[1]
[1] School of Electronic and Information Engineering, Xi'an Jiaotong University, Xi'an, Shaanxi, 710049, P.R. China
[2] Research Institute of Xi'an Jiaotong University, Zhejiang, Hangzhou, Zhejiang, 311215, P.R. China
shaohuikai@stu.xjtu.edu.cn dxfsxl@163.com bell@xjtu.edu.cn

Abstract

Efficient deep palmprint recognition has become an urgent issue for the demand of personal identification on mobile/wearable devices. Compared to other biometrics, palmprint recognition has many unique advantages, e.g. richness of features, high user-friendliness, suitability for private security, etc. Existing deep learning based methods are computationally exhaustive in feature representation and learning, which are not suitable for large-scale deployment in portable authentication systems. In this paper, we combine hash coding and knowledge distillation to explore efficient deep palmprint recognition. Based on deep hashing network, palmprint images were converted to binary codes to save storage space and speed up matching. Combining hashing coding with knowledge distillation can further compress deep model to achieve an efficient recognition by light networks. Unlike previous palmprint recognition on datasets collected by dedicated devices in a controlled environment, we establish a novel database for unconstrained palmprint recognition, which consists of more than 30,000 images collected by 5 different mobile phones. Moreover, we manually labeled 14 key points on each image for region of interest (ROI) extraction. Comprehensive experiments were conducted on this palmprint database. The results indicate the feasibility of our database and the potential of palmprint recognition to be used as an efficient biometrics for deployment on consumer devices.

1. Introduction

Recently, information security has become especially important. Traditional methods for identification, such as keys and passwords, have many drawbacks [1]. Biometrics is a technique that uses physiological or behavioral properties of human body for identification [12]. Two types of biometric recognition are mainly considered in literature: biometric verification and biometric identification [11]. Biometric verification refers to the comparison between the tester and her own biometric template in the database to determine whether the claim is true or not, which is one-to-one comparison. Biometric identification refers to matching the tester with all the samples in the database to determine the identity, which is one-to-many comparison. As one of the unique biometric characteristics, palmprint has received wide attention from researchers [40]. Many excellent palmprint recognition algorithms are emerging, such as robust line orientation code (RLOC) [14], linear programming (LP) formulation [25], and discriminative and robust competitive code (DRCC) [29]. Satisfactory results have been obtained on some public palmprint databases, such as PolyU multi-spectral dataset [32]. However, the current palmprint recognition system still has many shortcomings. Firstly, dedicated palmprint acquisition devices limit the efficient usage of palmprint recognition. In order to avoid external illumination interference, additional lighting needs to be added to a enclosed space for acquisition. In addition, volunteers need to limit their palms during acquisition, which reduces user-friendliness. Secondly, the recognition accuracies of traditional palmprint algorithms are relatively low, such as method based on Scale Invariant Feature Transform (SIFT) [39] and method based on Histogram of Oriented Gradients (HOG) [23]. Though the deep learning-based methods can greatly improve the performance, they could have high model complexity and low computational efficiency, which cannot meet the requirement of its practical application. For biometric recognition, recognizing efficiently is as critical as recognizing accurately. Proper acquisition of data is the basis of efficient identification. Feature representation and learning approaches are the key to efficient recognition.

In response to the above problems, in this paper, we carried out a comprehensive study on efficient deep palmprint recognition. Using 5 different mobile phones, we captured

*The corresponding author is Dexing Zhong, email: bell@xjtu.edu.cn.

about 30,000 palmprint images of 100 people in different environments. Based on this database, we adopted several algorithms to perform efficient palmprint recognition. Firstly, we tried to perform region of interest (ROI) extraction task. In order to extract stable ROIs, in each palmprint image, we manually marked 14 key points. We achieve automatic and precise positioning for the key points, and then extract the central area of palm as ROI. Secondly, based on the deep hashing network (DHN), we performed palmprint verification and identification tasks. DHN converts the palmprint images into binary codes. Due to the efficiency of hashing, we only need to obtain the Hamming distance between codes through XOR operation to realize identity recognition. Furthermore, to reduce the complexity of model, we implement efficient palmprint recognition tasks based on knowledge distillation (KD). The DHN, based on a pretrained VGG-16, is regarded as a teacher network. A light model with only two convolutional layers and three fully connected layers is selected as the student network. Through the guidance of teacher network, the student network can achieve higher recognition accuracy and efficiency. To the best of our knowledge, our proposed dataset is the most comprehensive palmprint database until now, allowing for multiple research tasks in addition to the above. The contributions can be summarized as follows:

1. Based on deep hashing network and knowledge distillation, we have realized efficient palmprint verification and identification. DHN converted images into binary codes, which reduced storage space and comparison time. Based on knowledge distillation, light student networks can achieve efficient performance under the guidance of teacher network.

2. We introduced a comprehensive palmprints database. It was collected from 5 mobile phones under unconstrained environments, containing more than 30,000 images and 10 domains, called Xi'an Jiaotong University Unconstrained Palmprint database (XJTU-UP) [1]. In order to obtain ROIs, 14 key points were manually marked.

This paper consists of 5 sections. Section 2 introduces related works. Section 3 explains our proposed efficient palmprint recognition methods in detail. The experiments and results are presented in section 4. Section 5 concludes the paper.

2. Related Works

2.1. Palmprint Recognition

Palmprint recognition has enjoyed great research popularity for identity authentication these years [16, 21]. Typically, a palmprint recognition pipeline consists of image acquisition, preprocessing, feature extraction and matching [40]. For image acquisition, there are several popular benchmark datasets until now [32, 31, 7, 17, 35] . Besides, there are also some databases acquired by mobile devices. Chora and Kozik *et al.* [3] used a smartphone to collect 252 palmprint images from 84 samplers. Jia *et al.* [13] used 2 smartphone cameras and a compact camera to collect 12,000 palmprints under two different illumination conditions for cross device recognition. Even though there are many mobile phone-based palmprint datasets, most of them are not open to public. Only Adrian-Stefan Ungureanu *et al.* [27] publicized their NUIG_Palm1 dataset collected by 5 smartphone cameras. After image acquisition and preprocessing, feature extraction will be conducted for identifying different profiles in the registered dataset. Existing palmprint recognition methods can be divided into several categories, such as structure-based and texture-based methods [33]. Guo *et al.* [6] proposed binary orientation co-occurrence vector (BOCV) to represent multiple orientations for a local region. After that, Zhang *et al.* [34] incorporated fragile bits information in the code maps from different palms and proposed extended binary orientation co-occurrence vector (E-BOCV). Dai *et al.* [4] designed a ridge-based palmprint matching system. Currently, deep learning-based methods have been introduced to recognize palmprints [43, 22, 44, 41]. Zhao *et al.* [37] trained the parameters of a deep belief net and obtained higher accuracy compared with traditional methods. Izadpanahkakhk *et al.* [10] extracted ROI and discriminative features from palmprints using transfer learning fusion.

2.2. Deep Hashing Network

Deep hashing network recently thrived with the development of deep learning [45, 2]. Due to the storage and retrieval efficiency, Zhu *et.al.* [45] firstly proposed supervised hashing to approximate nearest neighbor search for large-scale multimedia retrieval, which improved the quality of hash coding by exploiting the semantic similarity on data pairs. Cao *et.al.* [2] proposed Deep Visual-Semantic Quantization (DVSQ) to learn deep quantization models from labeled image data as well as the semantic information underlying general text domains. For effective palm vein verification, Zhong *et al.* [42] adapted DHN for extracting features in the palm vein configuration and matching, which reached an Equal Error Rate (EER) close to 0%. Currently, DHN is mainly based on convolutional neural networks (CNN) and hashing algorithms. There are several types of DHN: network in network hashing (NINH) [18], deep semantic ranking based hashing (DSRH) [38], deep regularized similarity comparison hashing (DRSCH) [36] and so on.

2.3. Knowledge distillation

Knowledge distillation is a technique for transferring information from a complex deep model to a light model [19]. KD has been widely used for model compression and ac-

[1]http://gr.xjtu.edu.cn/web/bell/resource

celeration to improve the performance of fast and light networks [5]. Hinton *et al.* [8] first distilled knowledge from an ensemble of pre-trained models to improve a small target net via high-temperature softmax training. Then, FitNet developed KD using the pre-trained wide and shallow teachers hint layer to assist thin and deep students guided layer [24]. Li *et al.* [20] proposed a unified distillation framework to use a small clean dataset and label relations in knowledge graph to guide the distillation process. Yim *et al.* [30] proposed a method of transferring the distilled knowledge as the flow between two layers by computing the inner product between features, and the student model outperformed the original model that was trained from scratch. Recently, KD is also successfully used for pedestrian detection [9] and face recognition [26]. Wei *et al.* [28] quantized a large network and then mimiced a quantized small network for object detection. The model improved the performance of a student network by transferring knowledge from a teacher network.

3. Method

3.1. Deep hashing network

DHN combines the high accuracy of deep learning and the efficiency of hash coding. In this paper, it converts palmprint images into 128-bit binary codes. On the one hand, the storage space of binary codes is greatly compressed compared to the original image, so that the storage efficiency is greatly improved. On the other hand, a simple XOR operation can be used to obtain the Hamming distance between codes, and their similarity can be quickly obtained to improve the efficiency of palmprint retrieval. In this paper, DHN based on VGG-16 is used for efficient palmprint recognition. VGG-16 has 5 batches of convolutional layers and 3 fully connected layers. We use the weights from batch-1 to batch-4 which are trained on ImageNet, as shown in Figure 1. The batch-5 and fully connected layers are trained with our palmprints. In fact, DHN transforms the softmax layer of VGG-16 into a coding layer, where *sgn* function is used as activation function. In order to obtain hash codes, the activation function of the last fully connected layer is set to tanh. The optimization goal consists of two parts: hash loss and quantization loss.

Hash loss: hash loss is based on contrastive loss, mainly to close the distance between similar samples and to extend the distance between heterogeneous images. Suppose for two images i and j, the features extracted by DHN are h_i and h_j, respectively, and the hash loss between them is defined as:

$$L_{h_{i,j}} = \frac{1}{2} S_{i,j} D\left(h_i, h_j\right) + \frac{1}{2}\left(1 - S_{i,j}\right) \max\left(t - D\left(h_i, h_j\right), 0\right), \quad (1)$$

where $S_{i,j}$ represents their relationship label. If images i and j are from the same class, $S_{i,j} = 1$, otherwise, $S_{i,j} = 0$. t is a threshold used to balance the distance between heterogeneous images.

Quantization loss: the quantization loss is mainly due to the fact that the feature is directly converted to binary code by the *sgn* function, which is defined as:

$$L_{q_i} = \|h_i - b_i\|_2 = \||h_i| - 1\|_2, \quad (2)$$

where b_i is the code of image i. Assuming we have N images in total, the optimization goal is:

$$\min L = \sum_{i=1}^{N} \sum_{j=i+1}^{N} L_{h_{i,j}} + w \times \sum_{i=1}^{N} L_{q_i} = L_h + wL_q, \quad (3)$$

where w is used to balance the weight between hash loss and quantization loss.

3.2. Knowledge distillation based DHN

The above DHN-based palmprint recognition can obtain potential results. However, the model based on VGG-16 is quite complex, which costs a lot of training time. In order to solve this problem, a KD model is applied here to realize efficient recognition under a light network. The KD model includes a teacher and a student network. In general, the teacher network is a complex network with strong feature extraction capabilities to to obtain distinguishable features. The student network is a light network with limited recognition capabilities but high recognition efficiency. During training, the feature learned by the teacher network is transferred to the student network as knowledge, so that the student network can achieve the same performance as the teacher network. In this paper, we introduce a KD-based DHN for both palmprint verification and classification. The teacher network is the above-mentioned VGG-16-based DHN. The student network consists of only three fully connected layers and two convolutional layers. The schematic of the model are shown in Figure 2, and related parameters are in Table 1.

Different optimization goals are selected to train the student network for different recognition tasks. For palmprint classification, a common classification model with cross entropy loss is adopted. In order to transfer the knowledge from the teacher to the student network, we refer to Hinton's method and introduced the soft labels. In general, the neural network converts the output of the last layer into a class probability through softmax layer to achieve the purpose of classification.

$$q_i = \frac{\exp\left(z_i\right)}{\sum_{z_i} \exp\left(z_i\right)}, \quad (4)$$

where z_i is the output of the $i-$th neuron in the last layer, corresponding to the $i-$th category. q_i is the probability that the input image belongs to the $i-$th category. A temperature

Student Network	
Conv1	3×3×16, stride 4, padding 0, ReLU
Pool1	max-pool
Conv2	5×5×32, stride 2, padding 0, ReLU
Pool2	max-pool
Full3	FC-512, ReLU
Full4	FC-128, tanh
Full5	FC-128, sgn

Table 1: Parameters of student network. In the convolutional layers, the first column specifies the number of filters and channels; the second column is convolution stride and spatial padding; the third column is activation function. Full3 and full4 specify the number of neurons and activation function.

T is introduced in Eq. (5) to provide a soft label for the student network, which improves the discrimination ability and generalizability of a model.

$$q_i^T = \frac{\exp\left(z_i/T\right)}{\sum_{z_i}\exp\left(z_i/T\right)}, \tag{5}$$

where q_i^T is the soft probability for the teacher network. When T is equal to 1, Eq. (4) and Eq. (5) are equivalent. Using a higher value of T produces a softer probability distribution over classes.

Wei *et al.*[28] trained very tiny CNN through the quantization method and the knowledge transfer-based mimic method. Inspired by it, for palmprint verification, DHN can be seen as a simple quantitative method. The realization of KD is mainly to find a way to contact the teacher network and the student network. Suppose for image i, the code obtained by the teacher network is b_i^T, and the code obtained by the student network is b_i^S. The task of KD is to make b_i^T and b_i^S as similar as possible. Here L2 loss is used to constrain the difference between them.

$$L_{b_i} = \left\|b_i^T - b_i^S\right\|_2, \tag{6}$$

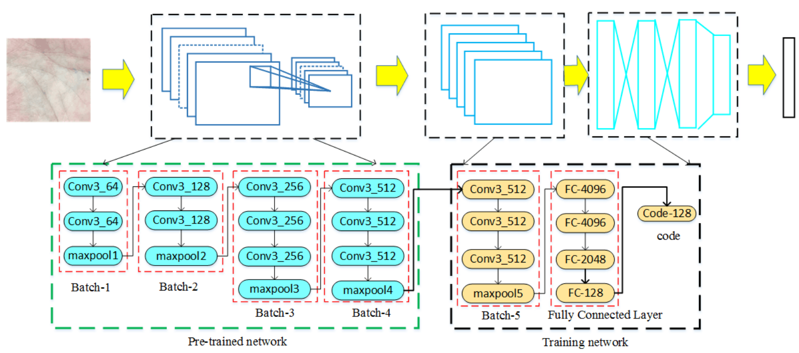


Figure 1: The schematic diagram of DHN

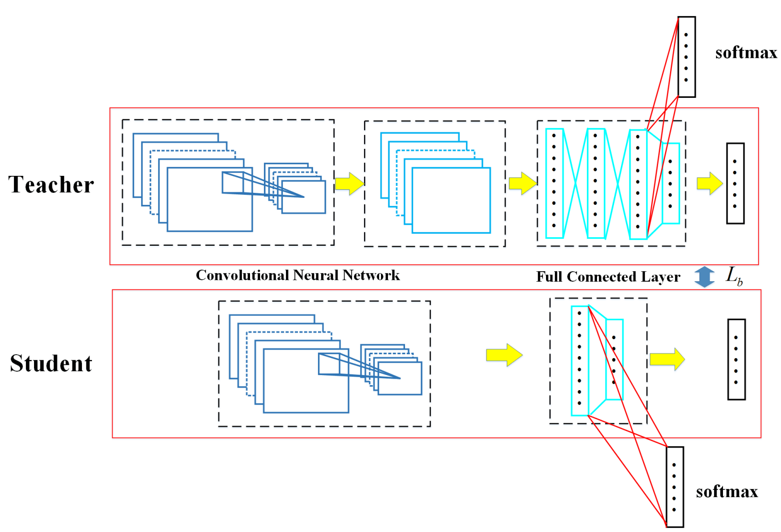


Figure 2: The schematic of KD model proposed based on DHN

Device	iPhone 6S	HUAWEI Mate8	LG G4	Galaxy Note5	MI8
Pixels	12 million	16 million	8 million	16 million	12 million
Image Size	3264×2448	3456×4608	5312×2988	5312×2988	4032×3024

Table 2: Details of acquisition devices

The teacher network is firstly trained in a supervised manner where T is set to 20. L is used to train the DHN and the cross entropy loss with T is adopted to train the classification network. In order to combine KD and DHN to achieve efficient palmprint verification and classification in a single network, their loss functions need to be added as the final optimization target.

Therefore, we formulate the loss of palmprint classification as follows:

$$L_c = mL_s\left(p_i, q_i\right) + nL_T\left(p_i, q_i^T\right), \tag{7}$$

where L_s is the normal cross entropy loss for student network, and L_T is soft cross entropy after adding temperature T which is set to 5.

Similarly, the loss of palmprint verification is:

$$\begin{aligned} L_v &= \sum_{i=1}^{N}\sum_{j=i+1}^{N} L_{h_{i,j}} + w \times \sum_{i=1}^{N} L_{q_i} + \alpha \times \sum_{i=1}^{N} L_{b_i} + \beta \times \sum_{i=1}^{N} L_T\left(p_i, q_i^T\right) \\ &= L_h + wL_q + \alpha L_b + \beta L_T. \end{aligned} \tag{8}$$

L_v is used to update the network parameters of DHN. L_c is only used to update the parameters of classification network.

4. Experiments and result analysis

4.1. Dataset

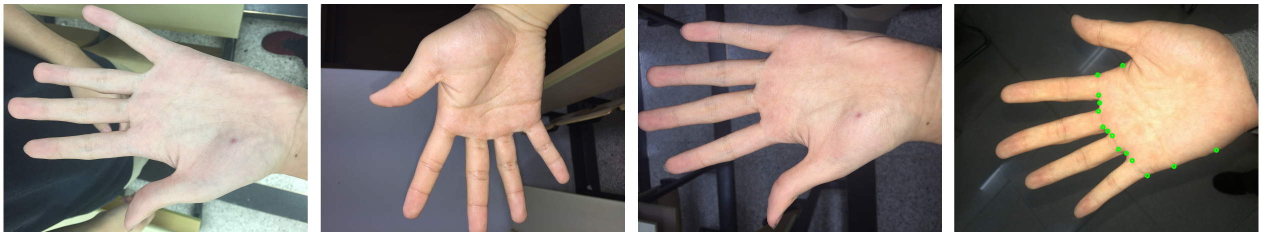
Our palmprint database is collected under unconstrained environments. It reduces the constraints of acquisition compared to other public databases. First of all, the collection devices are 5 most commonly used smartphones, *i.e.* iPhone 6S, HUAWEI Mate8, LG G4, Samsung Galaxy Note5, and MI8. The details related to the acquisition devices are shown in Table 2.

The volunteers captured palmprint images indoors without other controlled equipment. They chose the angles and backgrounds of shooting as they wish, as long as the entire palm was captured. In order to increase the diversity of samples, the postures and backgrounds of palms are continuously changed in each acqusition scenario. Two kinds of illumination are taken, one is the indoor natural illumination, and the other is flash lighting of mobile phone. 100 volunteers, which are 19 to 35 years old, provided their palmprint images. For different mobile phones, each volunteer was asked to capture about 10 to 15 images of the left and right hands under different illuminations. So for every tester, there are at least 5 (devices)$\times$2(illuminations)$\times$10$\times$2(palms) = 200 images with a total of 5 (devices)$\times$2 (illuminations) = 10 domains. The size of each RGB image is shown in Table 2. The exemplary images are shown in Figure 3. In the images acquired from the unconstrained environments, the palm area is difficult to extract, especially under natural illumination, which has a great influence on the ROI extraction methods based on contour and valley points. In order to extract stable ROIs, 14 key points are manually marked in every image, as shown in Figure 3(d). These points include 3 valley points between fingers, 8 points at the bottom of fingers, and 3 points on either side of palm. Compared to other points, the points on the sides of palm are less affected by the posture of hand. A ROI extraction method based on these feature points will be given in subsequent sections. According to the acquisition devices and illuminations, 10 sub-databases were named as IN (iPhone 6s under Natural illumination), IF (iPhone 6s under Flash illumination), HN (HUAWEI Mate8 under Natural illumination), HF (HUAWEI Mate8 under Flash illumination), LN (LG G4 under Natural illumination), LF (LG G4 under Flash illumination) SN (Samsung Galaxy Note5 under Natural illumination), SF (Samsung Galaxy Note5 under Flash illumination), MN (MI8 under Natural illumination), and MF (MI8 under Flash illumination).

To extract the ROIs in our dataset, we try to find the points based on the regression trees algorithm and then extract the ROI based on distance since it is difficult to extract the palm and its contours in the practical scenarios. The method mainly includes three steps: palm detection, key point positioning, and ROI extraction, as shown in Figure 4.

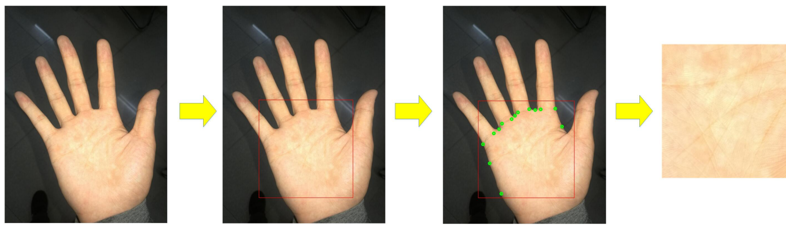
Step 1: Palm detection. In order to improve the accuracy and efficiency of key point positioning, the palm is firstly positioned. The target recognition algorithm based on HOG is mainly used. The HOG features of palm image are extracted. Then, in several scaled images of the original image, several candidate positions of target palm are detected using sliding window algorithm based on the pretrained Support Vector Machine (SVM). Finally, the palm position is selected from these candidate positions using non-maximum suppression algorithm, as shown in Figure 4(b).

Step 2: Key point positioning. After detecting the palm area, it is necessary to locate the key points in the area. Inspired by [15], regression tree algorithm is adopted. Each



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

Figure 3: Some typical examples in XJTU-UP. (a)-(c) are original images and (d) is the image with 14 key points



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

Figure 4: Schematic diagram of ROI extraction, (a) is original image, (b) is palm detection, (c) is key point positioning, and (d) is the ROI extracted

key point is located by a regression tree, and each regression tree has multiple sub-trees. The parameters of each tree are updated by the difference between the current shape predicted and the manually labeled ground truth of each key point.

Step 3: ROI extraction. After obtaining the precise key point positions, stable ROIs can be extracted based on different methods. Here, we give a distance-based approach to extract square ROI. The valley between the index finger and the middle finger, P_3, and the valley between the ring finger and the little finger, P_9, are selected to determine the direction of ROI. The edge points on both palm sides, P_0 and P_{12}, which are less affected by the postures of palm, are selected to determine the length of ROI. The valley between the middle finger and the ring finger, P_6, is used to determine the center point, P_o, of ROI. Suppose the distance between P_0 and P_{12} is L, therefor the side length of ROI is set to $3/5L$, and the distance between P_o and P_6 is $2/5L$.The details are shown in the Figure 5.

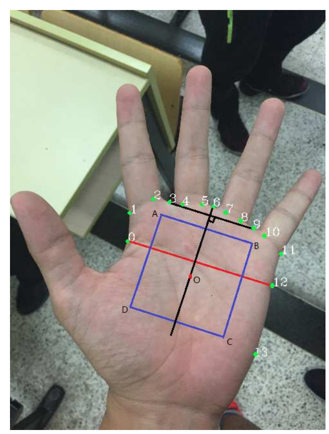


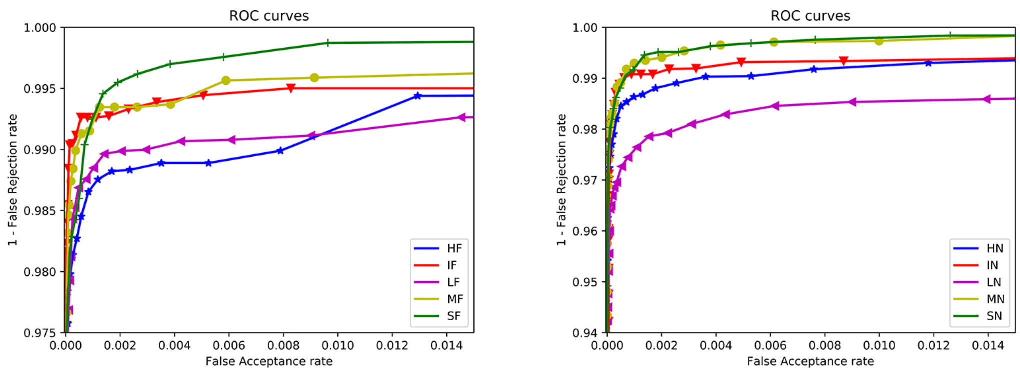
Figure 5: The details of ROI extraction

4.2. Palmprint verification and identification

In this paper, DHN is mainly adopted for efficient palmprint verification and identification. For each sub-database, the first five images of each category were used to train network parameters and the remaining images were used for testing. The teacher network, student network, and KD model are trained separately. For palmprint verification, after obtaining the code of each image, Hamming distance of any two codes was calculated. According to prior knowledge, different distance thresholds were set to get False Acceptance Rate (FAR) and False Rejection Rate (FRR). FAR is the ratio of the number of genuine matches, whose distance is greater than the threshold, to the number of all genuine matches. FRR is the ratio of the number of imposter matches, whose distance distance is less than the threshold, to the number of all imposter matches. When FAR is equal to FRR, the EER can be obtained. The Receiver Operating Characteristic (ROC) curve can be obtained according to FAR and FRR. In the palmprint identification, the category of image is obtained directly through the network, and then

Database	Teacher		Student		KD	
	EER (%)	Accuracy (%)	EER (%)	Accuracy (%)	EER (%)	Accuracy (%)
IF	0.50	98.57	5.91	69.23	3.96	90.31
IN	0.64	98.37	7.29	40.51	4.75	84.57
HF	0.70	97.37	6.47	46.16	5.44	88.38
HN	0.72	97.06	7.76	43.25	7.65	79.90
LF	0.79	97.10	6.36	39.07	5.11	87.25
LN	1.35	95.70	10.45	29.80	9.69	80.20
MF	0.43	97.84	5.29	49.90	5.15	88.45
MN	0.32	97.40	7.75	39.06	6.28	85.10
SF	0.29	97.5	4.14	52.19	4.00	88.75
SN	0.33	97.95	6.49	43.18	6.27	85.54
Average	0.607	97.49	7.47	45.235	5.83	85.845

Table 3: The results of palmprint recognition based on DHN and KD



(a) (b)

Figure 6: The ROC curves of palmprint recognition based on teacher DHN

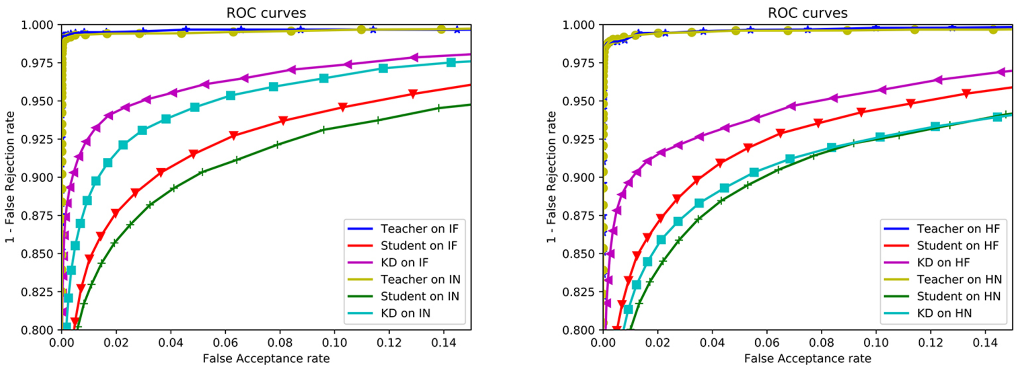
the accuracy is calculated. Some hyperparameters are set as: t=180, w=0.5, α=5, β=10, m=0.2, and n=0.8. Table 3 presents the EERs of 10 databases on different models, the ROC curves of teacher network are shown in Figure 6. Correspondingly, the ROC curves of KD and student network are shown in Figure 7.

It can be seen from the experiment results that the performance of the student network is significantly worse than the teacher network, especially for the palmprint classification task. This is because the student network is too simple and its ability to extract distinguishable features is weaker. In addition, for each sub-database, the illuminations, angles, and postures of images change greatly due to the acquisition environments, which greatly increases the difficulty of recognition. The limited number of images in every sub-database is also one of the reasons.

When using teacher network for palmprint verification, the best result is in SF database where EER is equal to 0.29%. The worst result is in LN database where EER equals to 1.35%. For palmprint identification, the best result is in IF where the recognition accuracy is 98.75%, and the worst result is in LN where recognition accuracy is 95.70%. From the experiment results, the average EER is 0.607%, and the average recognition accuracy is 97.49%, which proves that the databases can be used for palmprint recognition with high accuracy.

Before transferring the knowledge, the best performance of student network is in SF with an EER of 4.14%. The best classification accuracy is 69.23% in IF, where the ultimate goal for both tasks are not well achieved. After introducing the knowledge gained from the teacher network, there has been a potential improvement in recognition performance, especially for palmprint classification tasks. In IN, the EER decreases by up to 2.54%; the classification accuracy in the dataset of LN increases by up to 50.40%.

From the presented results, the performance of palmprint images acquired under the flash illumination is better than that under natural illumination for both palmprint verification and identification. The flash lights are equivalent to extra filling light, so the illumination is more uniform, which



(a) (b)

Figure 7: The ROC curves of palmprint recognition based on KD. (a) is on IF and IN, (b) is on HF and HN. Note that we omit the results on LF, LN, SF SN, MF and MN for simple illustration.

reduces the influence of unbalanced illumination and the difficulty of recognition. However, when considering the improvement of recognition accuracy, the experiments under the natural illumination is greater than those under the flash lights, which proves that the KD model has obtained useful knowledge from the teacher network.

Compared with teacher networks, light student networks perform relatively poorly. However, the KD-based student network achieves efficient palmprint recognition with little reduction in accuracy. In Table 4, we compare the performance of these three models. All the experiments are implemented using TensorFlow on a NVIDIA GPU GTX1080 with 8G memory. The model size and average feature extraction time for each image are obtained for comparison. The model size and complexity of teacher network are the largest, which is much larger than of student network. At the same time, the teacher network is the slowest with respect to the time spent on feature extraction, which is about 4 times the student network. For KD model, although it is necessary to obtain the trained teacher network before training the student network, the performance has been significantly improved. For palmprint classification, the accuracy is more than two times of the student network, but the speed is not reduced. This is very important for the palmprint recognition system on the consumer device, because the complex model cannot work well with limited computational resources.

	Teacher	Student	KD
Model size	93 M	0.51 M	0.51 M
Total parameters	156 M	48 M	48 M
Feature extraction time	12.30 ms	3.33 ms	3.32 ms
Iterations	10,000	10,000	10,000

Table 4: Computational cost of three different models

5. Conclusion

In this paper, we conduct a comprehensive analysis for efficient deep palmprint recognition via distilled hashing coding. We built an unconstrained palmprint image database using 5 mobile phones. Compared to other public databases, this database has more palmprint images and modalities with fewer restrictions on acquisition. There are 14 key points manually labeled on each image for ROI extraction, which will be publicly available to the research community. Based on the regression tree algorithm, we achieved the positioning of key points and extracted relatively stable ROIs. With Deep hashing network, we performed efficient palmprint verification and identification with an average accuracy of 97.49% and average EER of 0.607%. Knowledge distillation algorithm is used to improve recognition performance on light networks. The results show that the average classification accuracy is increased by 40.61% and the average EER is reduced by 1.64%. Experiment results show that the database we proposed has many advantages and can be successfully used in the research of multiple tasks. Further research would focus on deep learning-based models for more efficient representation learning in order to deploy the palmprint recognition systems in real world applications.

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

References

[1] X. Bai, N. Gao, Z. Zhang, and D. Zhang. 3D palmprint identification combining blocked ST and PCA. *Pattern Recognition Letters*, 100:89–95, 2017.

[2] Y. Cao, M. Long, J. Wang, and S. Liu. Deep visual-semantic quantization for efficient image retrieval. In *2017 IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017, July 21-26, 2017*, pages 916–925, 2017.

[3] M. Choras and R. Kozik. Contactless palmprint and knuckle biometrics for mobile devices. *Pattern Anal. Appl.*, 15(1):73–85, 2012.

[4] J. Dai, J. Feng, and J. Zhou. Robust and efficient ridge-based palmprint matching. *IEEE Trans. Pattern Anal. Mach. Intell.*, 34(8):1618–1632, 2012.

[5] Ç. Gülçehre and Y. Bengio. Knowledge matters: Importance of prior information for optimization. *Journal of Machine Learning Research*, 17:8:1–8:32, 2016.

[6] Z. Guo, D. Zhang, L. Zhang, and W. Zuo. Palmprint verification using binary orientation co-occurrence vector. *Pattern Recognition Letters*, 30(13):1219–1227, 2009.

[7] Y. Hao, Z. Sun, T. Tan, and C. Ren. Multispectral palm image fusion for accurate contact-free palmprint recognition. In *International Conference on Image Processing, ICIP 2008, October 12-15, 2008*, pages 281–284, 2008.

[8] G. Hinton, O. Vinyals, and J. Dean. Distilling the knowledge in a neural network. *CoRR*, abs/1503.02531, 2015.

[9] Q. Hu, P. Wang, C. Shen, A. Hengel, and F. Porikli. Pushing the limits of deep cnns for pedestrian detection. *IEEE Trans. Circuits Syst. Video Techn.*, 28(6):1358–1368, 2018.

[10] M. Izadpanahkakhk, S. Razavi, M. Gorjikolaie, S. Zahiri, and A. Uncini. Deep region of interest and feature extraction models for palmprint verification using convolutional neural networks transfer learning. *Applied Sciences*, 8(7), 2018.

[11] A. K. Jain, A. Ross, and S. Prabhakar. An introduction to biometric recognition. *IEEE Trans. Circuits Syst. Video Techn.*, 14(1):4–20, 2004.

[12] A. K. Jain, B. Ruud, and P. Sharath. Introduction to biometrics. *A Handbook of Biometrics*, 148(1):89–111, 2008.

[13] W. Jia, R. Hu, J. Gui, Y. Zhao, and X. Ren. Palmprint recognition across different devices. *Sensors*, 12(6):7938–7964, 2012.

[14] W. Jia, D. Huang, and D. Zhang. Palmprint verification based on robust line orientation code. *Pattern Recognition*, 41(5):1504–1513, 2008.

[15] V. Kazemi and J. Sullivan. One millisecond face alignment with an ensemble of regression trees. In *2014 IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2014, June 23-28, 2014*, pages 1867–1874, 2014.

[16] A. Kong, D. Zhang, and M. Kamel. A survey of palmprint recognition. *Pattern Recognition*, 42(7):1408–1418, 2009.

[17] A. Kumar and S. Shekhar. Personal identification using multibiometrics rank-level fusion. *IEEE Trans. Systems, Man, and Cybernetics, Part C*, 41(5):743–752, 2011.

[18] H. Lai, Y. Pan, Y. Liu, and S. Yan. Simultaneous feature learning and hash coding with deep neural networks. In *IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2015, June 7-12, 2015*, pages 3270–3278, 2015.

[19] S. Lee, D. Kim, and B. Song. Self-supervised knowledge distillation using singular value decomposition. In *15th European Conference on Computer Vision, ECCV 2018, September 8-14, 2018*, pages 339–354, 2018.

[20] Y. Li, J. Yang, Y. Song, L. Cao, J. Luo, and L. Li. Learning from noisy labels with distillation. In *IEEE International Conference on Computer Vision, ICCV 2017, October 22-29, 2017*, pages 1928–1936, 2017.

[21] G. Lu, D. Zhang, and K. Wang. Palmprint recognition using eigenpalms features. *Pattern Recognition Letters*, 24(9-10):1463–1467, 2003.

[22] S. Minaee and Y. Wang. Palmprint recognition using deep scattering convolutional network. *CoRR*, abs/1603.09027, 2016.

[23] A. Pariharand, A. Kumar, O. Verma, A. Gupta, P. Mukherjee, and D. Vatsa. Point Based Features for Contact-less Palmprint Images. In *IEEE International Conference on Technologies for Homeland Security*, pages 165–170, 2013.

[24] A. Romero, N. Ballas, S. Kahou, A. Chassang, C. Gatta, and Y. Bengio. Fitnets: Hints for thin deep nets. *CoRR*, abs/1412.6550, 2014.

[25] Z. Sun, L. Wang, and T. Tan. Ordinal feature selection for iris and palmprint recognition. *IEEE Trans. Image Processing*, 23(9):3922–3934, 2014.

[26] Y. Tai, J. Yang, Y. Zhang, L. Luo, J. Qian, and Y. Chen. Face recognition with pose variations and misalignment via orthogonal procrustes regression. *IEEE Trans. Image Processing*, 25(6):2673–2683, 2016.

[27] A. Ungureanu, S. Thavalengal, T. Cognard, C. Costache, and P. Corcoran. Unconstrained palmprint as a smartphone biometric. *IEEE Trans. Consumer Electronics*, 63(3):334–342, 2017.

[28] Y. Wei, X. Pan, H. Qin, W. Ouyang, and J. Yan. Quantization mimic: Towards very tiny CNN for object detection. In *15th European Conference on Computer Vision, ECCV 2018, September 8-14, 2018*, pages 274–290, 2018.

[29] Y. Xu, L. Fei, J. Wen, and D. Zhang. Discriminative and robust competitive code for palmprint recognition. *IEEE Trans. Systems, Man, and Cybernetics: Systems*, 48(2):232–241, 2018.

[30] J. Yim, D. Joo, J. Bae, and J. Kim. A gift from knowledge distillation: Fast optimization, network minimization and transfer learning. In *2017 IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017, July 21-26, 2017*, pages 7130–7138, 2017.

[31] J. You, D. Zhang, M. Wong, and W. Kong. Online palmprint identification. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25:1041–1050, 09 2003.

[32] D. Zhang, Z. Guo, G. Lu, L. Zhang, and W. Zuo. An online system of multispectral palmprint verification. *IEEE Trans. Instrumentation and Measurement*, 59(2):480–490, 2010.

[33] D. Zhang, W. Zuo, and F. Yue. A comparative study of palmprint recognition algorithms. *ACM Comput. Surv.*, 44(1):2:1–2:37, 2012.

[34] L. Zhang, H. Li, and J. Niu. Fragile bits in palmprint recognition. *IEEE Signal Process. Lett.*, 19(10):663–666, 2012.

[35] L. Zhang, L. Li, A. Yang, Y. Shen, and M. Yang. Towards contactless palmprint recognition: A novel device, a new benchmark, and a collaborative representation based identification approach. *Pattern Recognition*, 69:199–212, 2017.

[36] R. Zhang, L. Lin, R. Zhang, W. Zuo, and L. Zhang. Bit-scalable deep hashing with regularized similarity learning for image retrieval and person re-identification. *IEEE Trans. Image Processing*, 24(12):4766–4779, 2015.

[37] D. Zhao, X. Pan, X. Luo, and X. Gao. Palmprint recognition based on deep learning. In *6th International Conference on Wireless, Mobile and Multi-Media (ICWMMN 2015)*, pages 214–216, Nov 2015.

[38] F. Zhao, Y. Huang, L. Wang, and T. Tan. Deep semantic ranking based hashing for multi-label image retrieval. In *IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2015, June 7-12, 2015*, pages 1556–1564, 2015.

[39] Q. Zhao, W. Bu, and X. Wu. Sift-based image alignment for contactless palmprint verification. In *International Conference on Biometrics, ICB 2013, June 4-7, 2013, Madrid, Spain*, pages 1–6, 2013.

[40] D. Zhong, X. Du, and K. Zhong. Decade progress of palmprint recognition: A brief survey. *Neurocomputing*, 328:16–28, 2019.

[41] D. Zhong, M. Li, H. Shao, and S. Liu. Palmprint and dorsal hand vein dualmodal biometrics. In *2018 IEEE International Conference on Multimedia Expo Workshops (ICMEW)*, pages 1–6, July 2018.

[42] D. Zhong, S. Liu, W. Wang, and X. Du. Palm vein recognition with deep hashing network. In *First Chinese Conference on Pattern Recognition and Computer Vision, PRCV 2018, November 23-26, 2018*, pages 38–49, 2018.

[43] D. Zhong, Y. Yang, and X. Du. Palmprint recognition using siamese network. In *13th Chinese Conference on Biometric Recognition, CCBR 2018, August 11-12, 2018*, pages 48–55, 2018.

[44] D. Zhong and J. Zhu. Centralized large margin cosine loss for open-set deep palmprint recognition. *IEEE Transactions on Circuits and Systems for Video Technology*, 2019.

[45] H. Zhu, M. Long, J. Wang, and Y. Cao. Deep hashing network for efficient similarity retrieval. In *Thirtieth AAAI Conference on Artificial Intelligence, February 12-17, 2016*, pages 2415–2421, 2016.