IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received August 23, 2021, accepted September 11, 2021, date of publication September 13, 2021, date of current version October 11, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3112513

Palmprint-Palmvein Fusion Recognition Based on Deep Hashing Network

TENGFEI WU[1,2], LU LENG[1,2,3], (Member, IEEE), MUHAMMAD KHURRAM KHAN[4], (Senior Member, IEEE), AND FARRUKH ASLAM KHAN[4], (Senior Member, IEEE)

[1]School of Software, Nanchang Hangkong University, Nanchang 330063, China
[2]Key Laboratory of Jiangxi Province for Image Processing and Pattern Recognition, Nanchang Hangkong University, Nanchang 330063, China
[3]School of Electrical and Electronic Engineering, College of Engineering, Yonsei University, Seoul 120749, Republic of Korea
[4]Center of Excellence in Information Assurance, King Saud University, Riyadh 11653, Saudi Arabia

Corresponding author: Lu Leng (leng@nchu.edu.cn)

This Project was funded by the National Plan for Science, Technology and Innovation (MAARIFAH), King Abdulaziz City for Science and Technology, Kingdom of Saudi Arabia, Award Number (3-17-09-001-0008).

ABSTRACT Palmprint has attracted increasing attention due to its several advantages in the biometrics field. Deep learning has achieved remarkable performance in the computer vision area, so a large number of deep-learning-based methods have been proposed by the research community for palmprint recognition. The outputs of a deep hashing network (DHN) can be represented as a binary bit string, so DHN can reduce the storage and accelerate the matching/retrieval speed. In this paper, DHN is employed to extract the binary template for palmprint and palmvein verification. Spatial transformer network is used to overcome the rotation and dislocation. Palmprint and palmvein can be acquired from visible-light spectrums, including red (R), green (G), blue (B), and near infrared (NIR) spectrum, respectively. Since the features in different spectrums are different, their complementary advantages can be exploited to the full by fusion. Image-level fusion and score-level fusion are developed for palmprint-palmvein fusion recognition. The experiments demonstrate that score-level fusion can improve the accuracy efficiently.

INDEX TERMS Biometric recognition, palmprint verification, palmvein verification, fusion recognition, deep hashing network.

I. INTRODUCTION

Recently, biometric recognition has been applied in numerous security applications [1]. Both palmprint and palmvein have high discrimination, robustness and privacy, and have attracted increasing attention in recent years. For different palmprint modalities, different methods have been proposed, e.g., 2D palmprint recognition [2], [3], 3D palmprint recognition [4], palmprint and vein recognition [5], [6], etc. The existing methods can be briefly divided into the subspace-based [7]–[9], statistic-based [10]–[12], deep-learning-based [13], [14] and coding-based methods [15], [16].

In texture coding-based methods, some filters are typically employed to extract and encode the texture features, and generate the binary code as the template. The storage cost is low. In addition, the dissimilarity between two binary templates can be measured by Hamming distance, so the XOR computation in matching is fast.

Coding-based methods are suitable for ideal environments, such as contact acquisition. However, coding-based methods are not robust to some interferences in contactless acquisition, such as deformation due to hand gesture, placement movement, rotation, and tilt due to imperfect segmentation and localization.

Deep learning has achieved remarkable performance in many computer vision tasks. Deep learning models can solve the aforementioned interferences. Deep hashing network (DHN) is a state-of-the-art deep learning model, which inherits the advantages of deep learning and the strengths of coding-based methods, including strong robustness to interference, low storage cost, and fast matching speed. The size of the binary template of DHN is much smaller than the traditional coding-based methods, so this paper employs DHN for palmprint/palmvein recognition to further reduce the storage

The associate editor coordinating the review of this manuscript and approving it for publication was Yizhang Jiang.

cost, accelerate the speed, but without accuracy degradation even in severe environments.

The spectrums for palmprint acquisition include red (R), green (G), blue (B), and their combinations. The spectrums for the palmvein acquisition include Infrared (IR) and Near Infrared (NIR). Palmprint and palmvein can be acquired simultaneously, and they can use the same recognition algorithms, so they are suitable for fusion recognition. Accordingly, the main contributions of this paper are as follows:

(1) Spatial transformer network (STN) is embedded into DHN to align the images. Since STN overcomes the problems of dislocation and rotation, the verification accuracy is improved.

(2) Two image-level fusion modes, namely spatial concatenation and channel concatenation, are developed to improve the accuracy of palmprint-palmvein fusion recognition. Different combinations of the spectrums in the fusions are compared and analyzed to confirm the optimal scheme.

(3) Score-level fusion is performed for the fusion recognition of palmprint and palmvein. Compared with the recognition of a single modality, the recognition accuracy is greatly improved.

The rest of this paper is organized as follows: Section II reviews the related work. Section III describes the proposed method, and Section IV presents the experimental results. Finally, the conclusions are drawn in Section V.

II. RELATED WORK

Deep learning has achieved remarkable performance in many computer vision tasks. A limited number of deep-learning-based methods have been proposed for palmprint and palmvein recognition, which can be briefly categorized into four classes as follows:

A. DIRECT USAGE OF PRE-TRAINED NETWORKS

The pre-trained deep neural networks (DNNs) can be directly used for palmprint/palmvein recognition. Tarawneh *et al.* [17] compared different convolutional neural network (CNN) models for contactless palmprint recognition, including AlexNet [18], VGG-16 [19] and VGG-19. Their experiments demonstrate that VGG-16 and VGG-19 outperform AlexNet. Similarly, Ramachandra *et al.* [20] used AlexNet and SVM for the newborns' contactless palmprint recognition. Since the pre-trained networks are directly used, they are not trained specially for palmprint/palmvein recognition, so their accuracies are limited.

B. TRAINING NETWORKS FOR PALMPRINT/PALMVEIN RECOGNITION

Svoboda *et al.* [21] trained AlexNet to enlarge the separability between the genuine and imposter distributions, but it requires supervision training. Wen *et al.* [22] proposed a novel loss function, which aimed to increase the inter-class distances. Inspired by this work, Zhong and Zhu [23] proposed centralized large margin cosine loss on the benchmark structure in [24], which enhanced the intra-class compactness. Matkowski *et al.* [25] proposed a CNN framework for palmprint recognition in an uncontrollable environment, which included two sub-networks for segmenting region of interest (ROI) and extracting features, respectively. Chai *et al.* [26] pre-trained a network with gender soft biometric, and then trained the network for palmprint classification. Du *et al.* [27] proposed a CNN-based regularized adversarial domain model for cross-domain recognition. Liu *et al.* [28] used fully convolutional network (FCN) to develop soft-shifted triplet loss function to learn the discriminate palmprint features.

C. COMBINATION OF SPECIAL FILTERS AND DEEP LEARNING

Minaee and Wang [29] used deep scattering network and a bank of special filters to extract palmprint features. SVM was selected as the classifier. Genovese *et al.* [30] proposed PalmNet, which combined CNN, Gabor filters and principal component analysis (PCA). The significant advantage of PalmNet is that it can be trained in an unsupervised mode.

D. HASHING-BASED NETWORK

The classical CNN frameworks are suitable for classification (identification), but they are not good at verification (authentication). DHN was specially proposed for verification to enlarge the inter-class variance and reduce the intra-class variance. The outputs of DHN can be represented as a binary bit string, so DHN can reduce the storage and accelerate matching/retrieval speed.

Chen *et al.* [31] proposed discriminant spectral hashing for compact palmprint representation. Cheng *et al.* [32] combined supervised hashing with deep convolutional features for palmprint recognition. Zhong *et al.* [33] recognized hand-based multi-biometrics by using DHN and biometric graph matching to separately extract the features of palmprint and dorsal hand veins. A trained SVM was used as the classifier. Zhong *et al.* [34] employed deep hashing network for palmvein recognition; their method was named as deep hashing palmvein network (DHPN). The equal error rate of DHPN was close to 0% on NIR. Li *et al.* [35] used softmax classification loss and improved ternary losses to learn the hashing code to maintain the consistency with high-dimensional features. Liu *et al.* [36] used deep self-taught hashing to generate pseudo labels and then used DHN to generate hashing code.

The deep-learning-based palmprint/palmvein recognition methods are summarized in Table 1.

III. METHODLOGY

A. SPATIAL TRANSFORMER NETWORK

CNN encounters some restrictions due to the lack of spatial invariance. Jaderberg *et al.* [37] first proposed STN that can be used in spatial transform and data alignment according to specific tasks. STN is a differential module that can be embedded into the existing convolutional structures to shift,

TABLE 1. Deep-learning-based palmprint/palmvein recognition methods.

Ref.	Year	Network	Class	Description
[17]	2018	AlexNet, VGG-16, VGG-19	Direct usage of pre-trained DNNs	DNN for feature extraction, SVM for classification
[20]	2018	AlexNet		
[21]	2016	AlexNet	Trained for recognition	Enlarge the separability between the genuine and imposter distributions
[23]	2020	C-LMCL		Centralized large margin cosine loss (C-LMCL) for intra-class compactness enhancement
[25]	2019	Two sub-networks		Two sub-networks for ROI segmentation and feature extraction.
[26]	2019	CNN		Gender soft biometric recognition.
[27]	2020	CNN		CNN-based regularized adversarial domain model for cross-domain recognition
[28]	2020	FCN		Soft-shifted triplet loss function to learn the discriminate palmprint features
[29]	2017	Deep scattering network	Combination of special filters and deep learning	Deep scattering network+special filters
[30]	2019	PalmNet		CNN+Gabor filters+PCA
[31]	2013	DNN	Hashing-based network	Discriminant spectral hashing
[32]	2017	DNN		Supervised hashing+DNN
[33]	2019	DHN		DHN+biometric graph matching
[34]	2018	DHPN		Palmvein recognition
[35]	2019	DHN		Softmax classification loss+ improved ternary losses
[36]	2019	DHN		Deep self-taught hashing to pseudo labels

scale, rotate, and distort the feature maps while without any additional training or optimization. STN can generate the models that are invariant to translation, zoom, rotation, and deformation [38]. STN consists of three parts, namely localization network, grid generator, and sampler, as shown in Figure 1. The localization network is a regression network, which aims to generate the parameter θ. The input feature map passes several layers of convolution operations and fully connected layer regression to output θ. The size of θ depends on the transformation type, e.g. θ is 6-dimension in affine transformation. The grid generator uses the coordinates of the target feature map to calculate the corresponding coordinates of the target feature map T in the source feature map.

$$\begin{pmatrix} x_i^s \\ y_i^s \end{pmatrix} = T_\theta(G_i) = A_\theta \begin{pmatrix} x_i^t \\ y_i^t \\ 1 \end{pmatrix} = \begin{bmatrix} \theta_{11} & \theta_{12} & \theta_{13} \\ \theta_{21} & \theta_{22} & \theta_{23} \end{bmatrix} \begin{pmatrix} x_i^t \\ y_i^t \\ 1 \end{pmatrix} \quad (1)$$

where (x_i^s, y_i^s) are the source coordinates in the input feature map, (x_i^t, y_i^t) are the target coordinates, and θ is the transformation parameter. The sampler samples the original feature map according to the coordinates in $T(G)$, and copies the pixels in the source feature map S to the target feature map.

B. NETWORK ARCHITECTURE

CNN-F in [34] is selected as the backbone. The architecture of the modified CNN-F is shown in Figure 2. The network configuration is shown in Table 2. Compared with the network in [34], the network structure in this paper adds STN after the input layer, and the modified CNN-F uses 5 layers of convolutional layers instead of 4 layers, and uses PReLU as the activation function after each convolutional layer. In the last fully connected layer, three lengths, 64, 128, and 256, are tested for experiments. Compared with the original CNN-F, the modified CNN-F uses fewer channels and further reduces the amount of calculation.

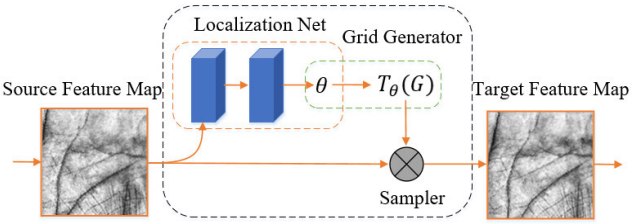


FIGURE 1. Architecture of spatial transformer network.

TABLE 2. Modified CNN-F parameters.

Type	Configuration	Output size (C×H×W)
Input		1×128×128
STN		1×128×128
Conv1	16×3×3, st.4, pad 0, BN, PReLU	16×32×32
Max_pool	2×2, st.1, pad 0	16×31×31
Conv2	32×5×5, st.2, pad 2, BN, PReLU	32×16×16
Max_pool	2×2, st.1, pad 0	32×15×15
Conv3	64×3×3, st.1, pad 1, PReLU	64×15×15
Conv4	64×3×3, st.1, pad 1, PReLU	64×15×15
Conv5	128×3×3, st.1, pad 1, PReLU	128×15×15
Max_pool	2×2, st.1, pad 0	128×14×14
Full6	2048	2048
Full7	2048	2048
Full8	64/128/256 tanh and sgn	64/128/256

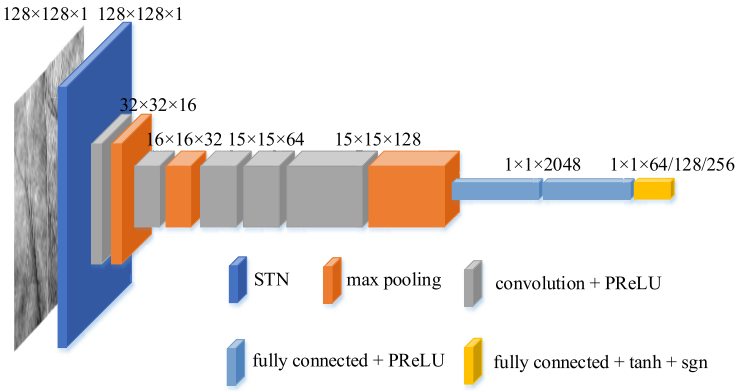


FIGURE 2. Architecture of modified CNN-F.

STN module is embedded at the back of the network input layer. STN module does not need to design a loss function separately, its parameter learning can be realized by using the loss function of the network. In order to avoid the problem of the gradient disappearance when the convolution result is negative, PReLU activation function is used following each convolution layer. Tanh function is used as the activation function of the output layer. Sign function is used to quantize the output. Finally, the hashing code with a binary value of -1 or 1 is obtained.

C. LOSS FUNCTION

Controlling quantization error and cross entropy loss can effectively improve network performance. Therefore, the proposed loss function is based on two parts, namely distance loss and quantization loss.

1) DISTANCE LOSS

The purpose of the distance loss is to enlarge inter-class variance and reduce intra-class variance. The distance loss is defined as:

$$L_S\left(x_i, x_j, l_{ij}\right) = \sum_{i=1}^{M}\sum_{j=1}^{M}\left[S\left(x_i, x_j\right) + \bar{S}\left(x_i, x_j\right)\right] \quad (2)$$

$$S\left(x_i, x_j\right) = \frac{1}{2} l_{ij} D\left(f_i, f_j\right) \quad (3)$$

$$\bar{S}\left(x_i, x_j\right) = \frac{1}{2}\left(1 - l_{ij}\right) max\left(T - D\left(f_i, f_j\right), 0\right) \quad (4)$$

M is the number of palmprint images in the training set, x_i and x_j are two input samples, f_i and f_j are their binary hashing codes, and $D(f_i, f_j)$ is the Hamming distance between two hashing codes. If x_i and x_j are of the same class, l_{ij} =1; else l_{ij} =0. T is the threshold. If x_i and x_j are of different classes and $D(f_i, f_j) > T$, there is no need to further enlarge the distance between the two samples. The value of T is determined by the length of the hashing code output from the network.

2) QUANTIZATION LOSS

In the output of DNN, if the output of the last layer is randomly distributed, binarization by the tanh and sgn functions will inevitably lead to a large quantization error. In order to reduce the quantization error, the quantization loss should make each entry in the output closer to 1 or -1. The quantization error is defined as:

$$L_Q = \sum_{i=1}^{M} \frac{1}{2}\left(\|1 - |f_i|\|_2\right) \quad (5)$$

$|f_i|$ is the absolute value of f_i, $\|\cdot\|$ denotes the $L2$ norm. Then the optimization is:

$$minL = \alpha L_S + L_Q \quad (6)$$

α is the scaling factor, whose empirical value is 0.5 according to experiences.

D. IMAGE-LEVEL FUSION

Image-level fusion includes two modes, namely spatial concatenation and channel concatenation.

The ROI size of palmprint and palmvein is 128 × 128. Assume $\mathbf{I}_1$, $\mathbf{I}_2$, $\mathbf{I}_3$ and $\mathbf{I}_4$ are the ROI images of R, G, B, NIR spectrums of the same class. The images can be concatenated, i.e., they are fused at image level.

1) SPATIAL CONCATENATION

Two spectral ROI images, $\mathbf{I}_i$, $\mathbf{I}_j$, of the same class, can be concatenated as $\mathbf{I}$=[$\mathbf{I}_i$, $\mathbf{I}_j$] with the size of 128×(2 × 128)=128 × 256;

Three spectral ROI images, $\mathbf{I}_i$, $\mathbf{I}_j$, $\mathbf{I}_k$, of the same class, can be concatenated as $\mathbf{I}$=[$\mathbf{I}_i$, $\mathbf{I}_j$, $\mathbf{I}_k$] with the size of 128×(3 × 128)=128 × 384;

Four spectral ROI images, $\mathbf{I}_i$, $\mathbf{I}_j$, $\mathbf{I}_k$, $\mathbf{I}_l$ of the same class, can be concatenated as $\mathbf{I}$=[$\mathbf{I}_i$, $\mathbf{I}_j$, $\mathbf{I}_k$, $\mathbf{I}_l$] with the size of (2 × 128)×(2 × 128)=256 × 256.

2) CHANNEL CONCATENATION

Each spectral image can be regarded as a single channel in channel concatenation.

TABLE 3. Different databases.

Database	Acquisition	Number of categories	Number of samples per category	Total number of samples
PolyU [39]	Contact	378	20	7560
Red [39]	Contact	500	12	6000
Green [39]	Contact	500	12	6000
Blue [39]	Contact	500	12	6000
NIR [39]	Contact	500	12	6000
IITD [40]	Contactless	460	5	2300
Tongji [41]	Contactless	600	20	12000
Tongji Palmvein [41]	Contactless	600	20	12000

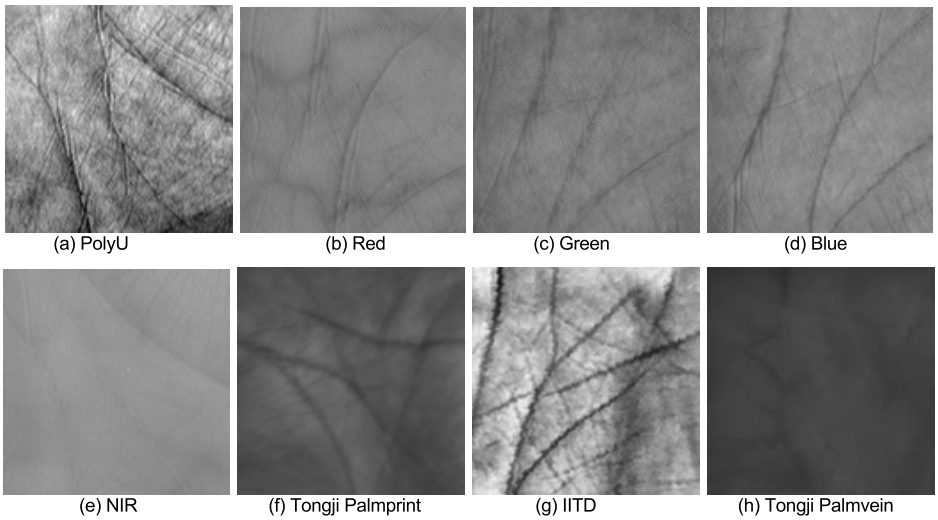


FIGURE 3. Samples in different datasets.

TABLE 4. EERs(%) of different methods on different databases.

	PolyU	Blue	Green	Red	NIR	IITD	Tongji	Tongji Palmvein
PalmCode [43]	0.4082	0.3524	0.2774	0.2376	0.2230	6.2690	0.1255	2.5980
OrdinalCode [44]	0.2455	0.1716	0.1542	0.0723	0.1193	5.5144	0.1089	1.0815
FusionCode [45]	0.2915	0.2810	0.2147	0.1264	0.1700	6.5042	0.0753	1.1743
CompCode [46]	0.1530	0.0845	0.0952	0.0345	0.0562	6.3912	0.0151	0.8945
RLOC [47]	0.1543	0.0892	0.1020	0.0414	0.0727	5.7972	0.0248	0.7823
HOC [48]	0.1935	0.1837	0.1518	0.0113	0.0952	7.1563	0.1038	2.7471
DOC [49]	0.1839	0.1476	0.1175	0.0649	0.0598	7.5081	0.0416	1.1304
DCC [50]	0.1699	0.1048	0.0979	0.0447	0.0637	5.8401	0.0554	1.4377
DRCC [50]	0.1580	0.1211	0.1052	0.0678	0.0555	6.1882	0.0323	1.4352
BOCV [51]	0.1082	0.0526	0.0604	0.0242	0.0365	5.0312	0.0068	0.7754
DHPN [34]	0.0302	0.0213	0.0352	0.0369	0.0020	3.7316	0.0694	0.6569
PalmNet [30]	0.1110	0.0178	0.0087	0.0366	0.0871	4.2041	0.0332	0.7363
Ours	**0.0227**	**0**	**0**	**0**	**0**	**3.1183**	**0.0001**	**0.1723**

Three spectral ROI images, $\mathbf{I}_i$, $\mathbf{I}_j$, $\mathbf{I}_k$, of the same class, can be concatenated as $\mathbf{I}$=[[$\mathbf{I}_i$], [$\mathbf{I}_j$], [$\mathbf{I}_k$]] with the size of 128 × 128×3.

The concatenated images can be input to DHN for training, which contain the multi-spectral information with image-level fusion, so the accuracy can be improved.

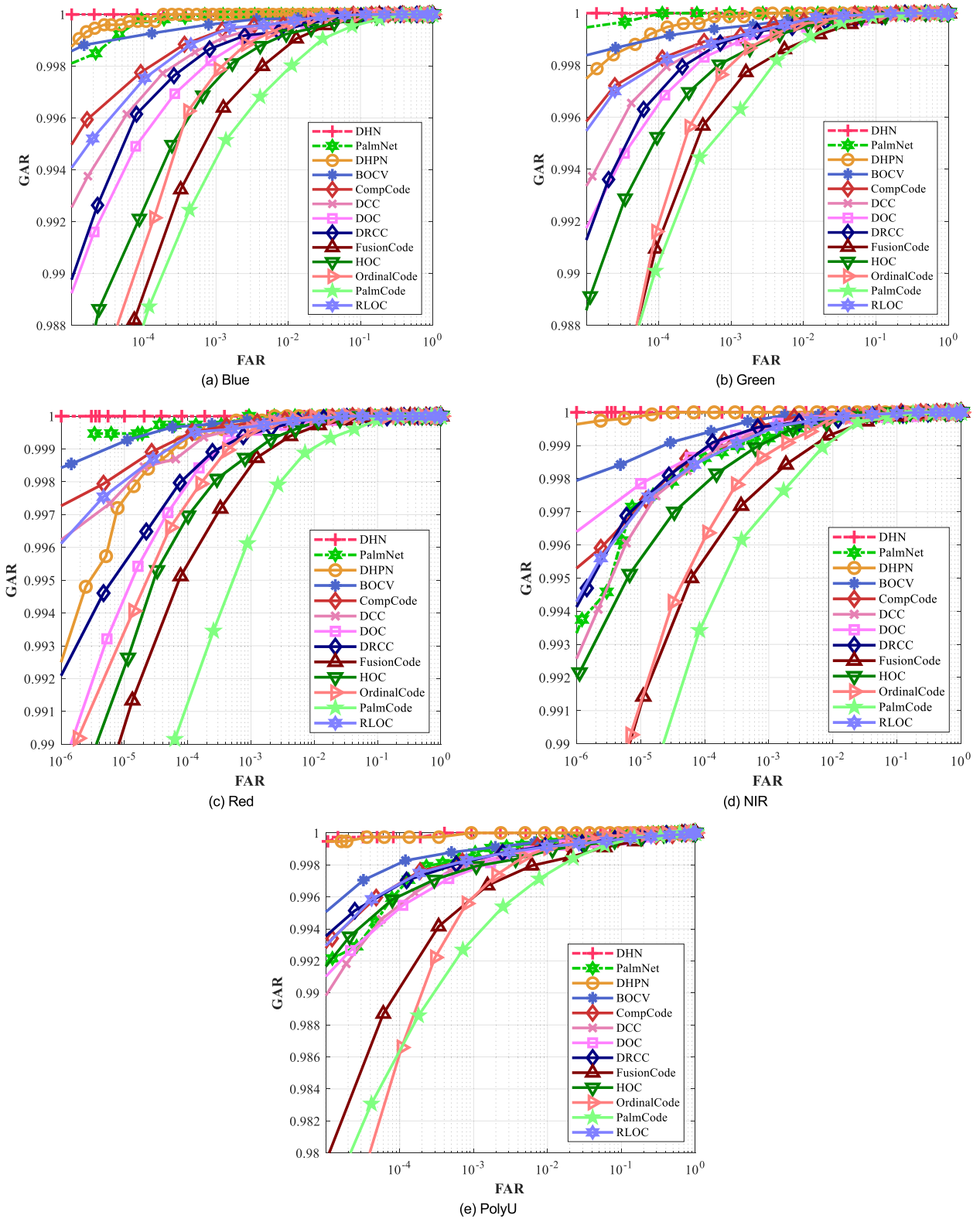


FIGURE 4. ROC curves on different contact palmprint databases. (a)-(e) represent Blue,Green, Red, NIR and PolyU, respectively.

E. SCORE-LEVEL FUSION

The output of DHN is two-valued, −1 and 1, and the output can be converted to binary, 0 and 1. The Hamming distance can be fast computed by XOR operation on each bit. The proposed model is trained on each spectrum sample independently. There are four spectrums

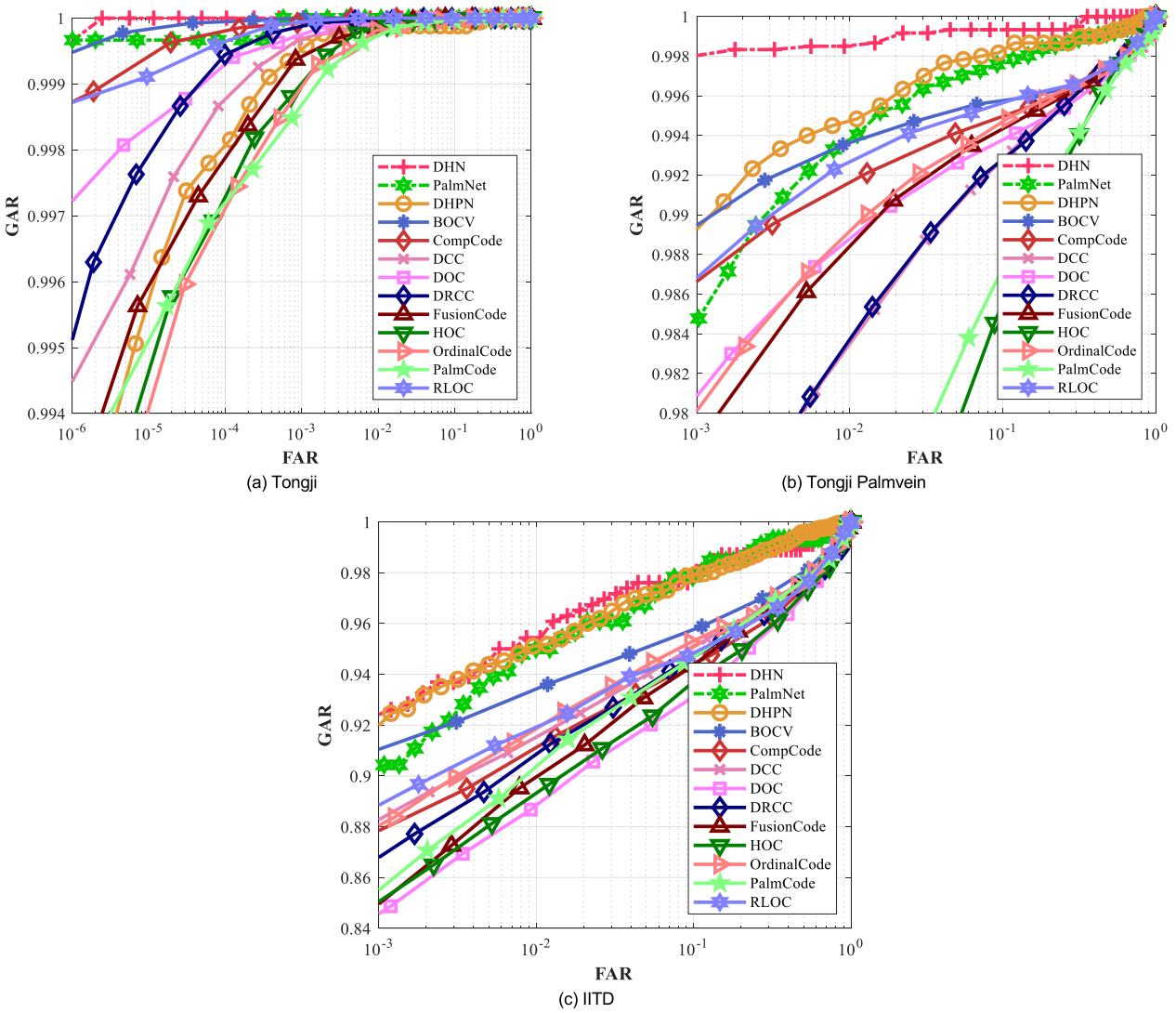


FIGURE 5. ROC curves on different contactless palmprint(palmvein) databases. (a)-(c) represent Tongji-print, Tongji-vein and IITD, respectively.

in multi-spectral database, so four proposed models are trained.

The dissimilarity between two deep hashing codes is measured by Hamming distance. Assume the Hamming distance of one spectrum is h_i, $1 \leq i \leq 4$. In score-level fusion, the weighted sum of Hamming distances is:

$$H_W = \sum_{i=1}^{n} w_i h_i, \quad \sum_{i=1}^{n} w_i = 1 \tag{7}$$

n is the number of spectrums in the fusion at score level, $n = 2$, 3 or 4. If all the weights are equal, $w_i = 1/n$.

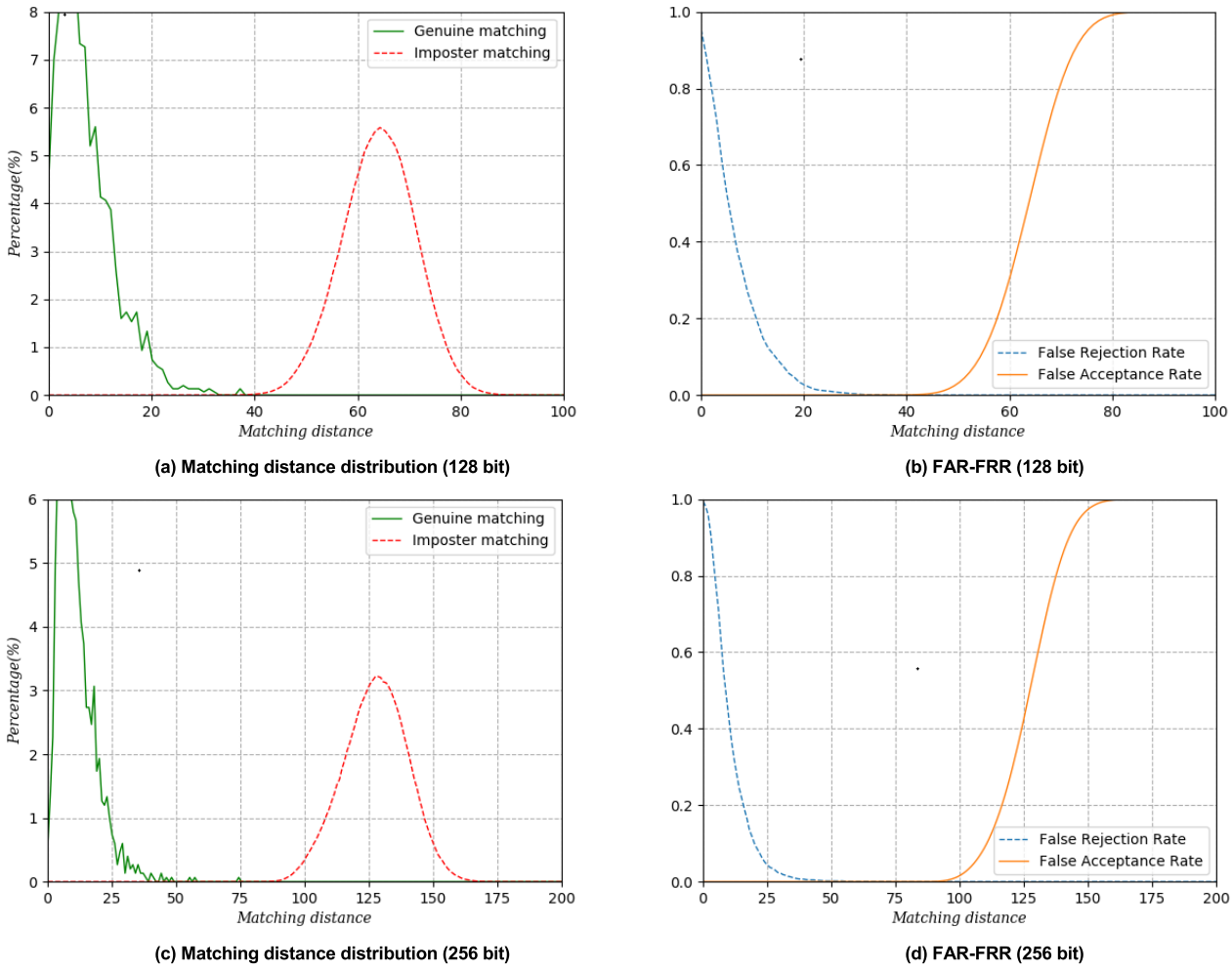
IV. EXPERIMENTAL RESULTS AND DISCUSSION

A. DATABASES

In order to confirm that the methods in this paper have satisfactory performance, the experiments are conducted on four public databases, which are PolyU [39], multi-spectral [39], IITD [40] and Tongji [41], [42]. The left and right hands of one person are considered as two different classes. The samples of some classes are removed so that each class has the same sample number. Table 3 lists the information of the databases. Figure 3 shows the image samples of these databases. The palmprint and palmvein look dissimilar because they are acquired with different devices.

B. VERIFICATION

When false rejection rate (FRR) is equal to false acceptance rate (FAR), their values are equal to error rate (EER). The proposed method, namely DHN+STN, is compared with coding-based methods and deep-learning-based methods. Coding-based methods include PalmCode [43], OrdinalCode [44], FusionCode [45], Competitive Code (CompCode) [46], Robust Line Orientation Code (RLOC) [47], Half-orientation Code (HOC) [48], Double Orientation Code (DOC) [49], Discriminative Competitive Code (DCC) [50], Discriminative Robust Competitive Code (DRCC) [50],



(a) Matching distance distribution (128 bit)

(b) FAR-FRR (128 bit)

(c) Matching distance distribution (256 bit)

(d) FAR-FRR (256 bit)

FIGURE 6. Matching distance distributions and FAR-FRR curvers using 128-bit and 256-bit encoding on Red.

and Binary Orientation Co-occurrence Vector (BOCV) [51]. Deep-learning-based methods include DHPN [34] and PalmNet [30]. For all the deep-learning-based methods, the ratio between the sample numbers in training set and testing set is 3:1. The hashing code length is set to 128.

Table 4 compares the EERs of different methods on different databases. Our method yields the best results on all databases. The accuracies on contact databases are better than those on contactless databases since the acquisition conditions can be controlled satisfactorily on contact databases.

The receiver operating characteristic (ROC) curves of the proposed method and the state-of-the-art methods on the contact and contactless databases are shown in Figure 4 and Figure 5, respectively. The ROC curves of the proposed method are almost always the highest.

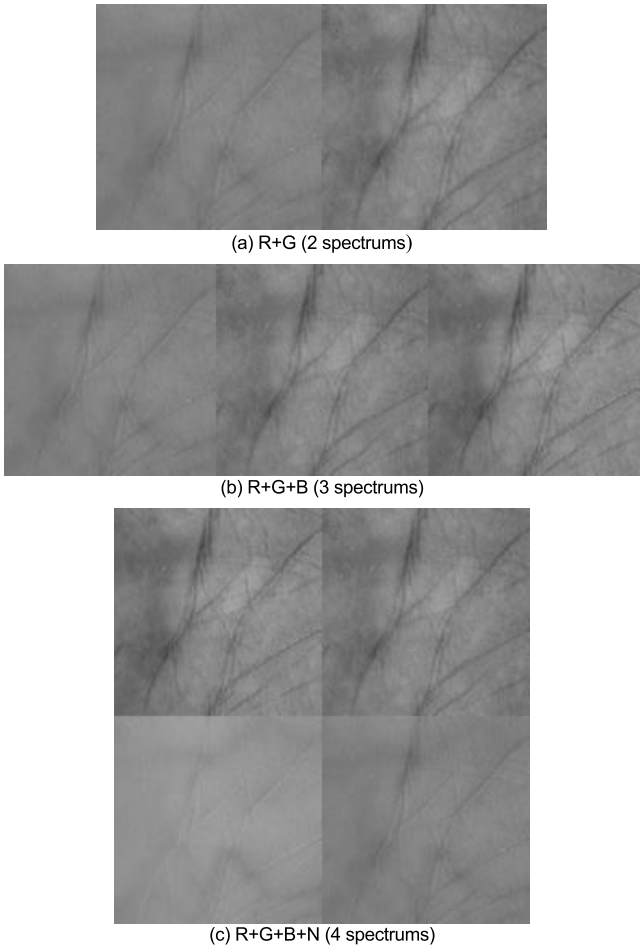
To confirm that our method can yield zero EER, Figure 6 shows the results on Red, including the intra-class and inter-class matching distance distribution curves as well as FAR-FRR curves using 128-bit and 256-bit encoding. The ratio between the training set and the test set is 3:1. In Figure 6(a), when the distance threshold is about 40, the two curves are not overlapped, at the same time, FAR=FRR=0 in Figure 6(b). The same conclusions can be drawn from Figure 6(c) and (d). The failure of other methods to get EER=0 is possibly caused by the number of test samples and the difficulty of network training. The test set of the coding-based methods is the entire data set, while the method in this paper and the deep-learning-based methods are tested only on the test set. Obviously, the latter test set is a subset of the former test set. The method in this paper takes full advantages of fusion of multiple spectrums to enhance the discrimination, so it can yield zero EER.

The data sizes of the outputs of the methods are compared in Table 5. The size of all the outputs is measured by bit number. The number of filter orientations is typically 6, so the orientation index, an integer within the range of [0,5], can be represented by 3 bits. Our method and DHPN require much less storage than the other methods.

Table 6 compares the EERs of our method with different ratios between the training and testing sample numbers. Since the sample numbers per class are different in different databases, the ratios are different in the databases. The longer

TABLE 5. Data sizes of the outputs.

PalmCode [43]	OrdinalCode [44]	FusionCode [45]
32×32×2=2048	32×32×3=3072	32×32×2=2048
HOC [48]	DOC [49]	CompCode [46]
32×32×2=2048	32×32×2=2048	32×32×3=3072
DCC [50]	DRCC [50]	BOCV [51]
32×32×2=2048	32×32×2=2048	32×32×6=6144
DHPN [34]	PalmNet [30]	Ours
128	15×25×2^15=12,288,000	128



(a) R+G (2 spectrums)

(b) R+G+B (3 spectrums)

(c) R+G+B+N (4 spectrums)

FIGURE 7. Spatial-concatenated image with different spectrum numbers.

the hashing code length is, the lower the EER will be. The EER can be 0 in some databases with good acquisition conditions.

C. IMAGE-LEVEL FUSION

Image-level fusion includes two modes, namely spatial concatenation and channel concatenation. Figure 7 shows the spatial-concatenated image with different spectrum numbers.

Table 7 shows the EERs of spatial concatenation. The ratio between the training and testing sample numbers is 1:1, i.e., 6 samples of each class are for training; while the other 6 samples of each class are for testing. The hashing code length is set to 128 and 64. The best results on one spectrum and two-, three-, four-spectrum spatial concatenation are bold. For single spectrum, N outperforms the other three spectrums. Spatial concatenation cannot remarkably improve the accuracy.

TABLE 6. EERs(%) of our method with different ratios between the training and testing sample numbers.

Database	Ratios		
PolyU	1:1	3:2	3:1
256bits	0.2960	0.0345	0.0031
128bits	0.3310	0.0815	0.0227
64bits	1.060	0.3370	0.1981
Red	1:1	2:1	3:1
256bits	0.0545	0	0
128bits	0.1297	0.0016	0
64bits	0.4277	0.0414	0.0005
Green	1:1	2:1	3:1
256bits	0.1328	0	0
128bits	0.1959	0.0304	0
64bits	0.4410	0.0554	0
Blue	1:1	2:1	3:1
256bits	0.0568	0	0
128bits	0.1320	0.0019	0
64bits	0.5970	0.0117	0.0010
NIR	1:1	2:1	3:1
256bits	0.0122	0	0
128bits	0.0601	0	0
64bits	0.2864	0.0056	0
Tongji	1:1	3:2	3:1
256bits	0.1691	0.0021	0.00001
128bits	0.3991	0.0879	0.0001
64bits	1.5181	0.8182	0.4689
Tongji Palmvein	1:1	3:2	3:1
256bits	0.4875	0.2971	0.1458
128bits	0.7265	0.3005	0.1723
64bits	1.9472	1.5362	1.1422

Table 8 shows the EERs of channel concatenation. The ratio between the training and testing sample numbers is 1:1. The hashing code length is set to 128 and 64. Compared with single spectrum, channel concatenation improves the accuracy. In addition, channel concatenation is better than spacial concatenation according to Tables 7 and 8.

TABLE 7. EERs(%) of spatial concatenation.

Spectrum	128bit	64bit
B	0.1320	0.6970
G	0.1959	0.4410
N	**0.0601**	**0.2864**
R	0.1297	0.4276
B+N	**0.0574**	0.7339
B+G	0.1716	1.3670
B+R	0.0884	**0.5280**
G+N	0.1359	2.7196
R+G	0.1231	1.1396
N+R	0.0579	0.5389
R+G+B	0.1978	0.6701
G+B+N	0.1448	1.0175
R+B+N	**0.1169**	**0.5952**
G+N+R	0.1692	1.7160
G+R+N+B	**0.1899**	**2.6921**

TABLE 8. EERs(%) of channel concatenation.

Spectrum	128bit	64bit
B	0.1320	0.6970
G	0.1959	0.4410
N	**0.0601**	**0.2864**
R	0.1297	0.4276
R+G+B	0.1006	0.3303
G+B+N	0.0625	0.2546
R+B+N	**0.0560**	**0.1608**
G+N+R	0.0732	0.3382

D. SCORE-LEVEL FUSION

Table 9 shows the EERs of multi-spectral fusion at score level. The ratio between the training and testing sample numbers is 1, i.e., 6 samples of each class are for training; while the other 6 samples of each class are for testing. The hashing code length is set to 128 and 64. n is the number of spectrums in the fusion at score level, $n = 2$, 3 or 4. All the weights are equal. $n = 1$ means single spectrum is used without fusion.

(1) All the results when $2 \leq n \leq 4$ are better than those when $n = 1$, i.e., any score-level fusion can yield better result than that of any single spectrum.

(2) The best results appear when $n = 3$, but the results when $n = 3$ are unstable, the results are different when different spectrums are fused. Thus, it is not easy to select the spectrums when $n = 3$.

(3) The results when $n = 4$ are highly close to the best results.

TABLE 9. EERs (%) of multi-spectral fusion at score level.

Spectrum	128bit	64bit
B	0.1320	0.5970
G	0.1960	0.4410
N	**0.0601**	**0.2864**
R	0.1297	0.4277
B+N	0.0198	0.0936
B+G	0.0583	0.1264
B+R	0.0213	0.0799
G+N	0.0135	0.0782
G+R	0.0294	**0.0387**
N+R	**0.0096**	0.0613
B+G+R	0.0260	0.0492
B+G+N	0.0016	0.0657
B+N+R	**0.0013**	0.0550
G+N+R	0.0023	**0.0248**
G+R+N+B	**0.0013**	**0.0253**

TABLE 10. EERs (%) of Tongji-print and Tongji-vein fusion at score level.

	128bit	64bit
Tongji	0.3991	1.5181
Tongji Palmvein	0.7265	1.9472
Tongji+Tongji Palmvein	**0.1484**	**0.5063**

Table 10 shows the EERs of Tongji palmprint and Tongji palmvein fusion at score level. The ratio between the training and testing sample numbers is 1:1, i.e., 10 samples of each class are for training; while the other 10 samples of each class are for testing. The hashing code length is set to 128 and 64. The weights of palmprint and palmvein are set to 0.5. From the experimental results, the score-level fusion improves the accuracy significantly.

V. CONCLUSION AND FUTURE WORK

In this paper, DHN is employed to extract the binary template for palmprint and palmvein verification. Spatial Transformer Network is used to overcome the rotation and dislocation, and accordingly improve the accuracy. The spectrums include R, G, B and NIR. Palmprint and palmvein can be acquired from visible-light spectrums and NIR spectrums, respectively. Since the features in different spectrums have different information, their complementary advantages can be exploited to the full by fusion. According to our investigation, image-level fusion of channel concatenation and score-level fusion can improve the accuracy. In our future work, we will try to select the discriminative bits from the hashing code to reduce the data size while without degrading the accuracy, and will explore a deep hash network recognition framework for open data sets.

REFERENCES

[1] A.-S. Ungureanu, S. Salahuddin, and P. Corcoran, "Toward unconstrained palmprint recognition on consumer devices: A literature review," *IEEE Access*, vol. 8, pp. 86130–86148, 2020.

[2] X. Zhou, K. Zhou, and L. Shen, "Rotation and translation invariant palmprint recognition with biologically inspired transform," *IEEE Access*, vol. 8, pp. 80097–80119, 2020.

[3] Y. Hao, Z. Sun, T. Tan, and R. Chao, "Multispectral palm image fusion for accurate contact-free palmprint recognition," in *Proc. 15th IEEE Int. Conf. Image Process.*, Oct. 2008, pp. 281–284.

[4] A. Iula and M. Micucci, "A feasible 3D ultrasound palmprint recognition system for secure access control applications," *IEEE Access*, vol. 9, pp. 39746–39756, 2021.

[5] S. Sun, X. Cong, P. Zhang, B. Sun, and X. Guo, "Palm vein recognition based on NPE and KELM," *IEEE Access*, vol. 9, pp. 71778–71783, 2021.

[6] D. Zhong, M. Li, H. Shao, and S. Liu, "Palmprint and dorsal hand vein dualmodal biometrics," in *Proc. IEEE Int. Conf. Multimedia Expo Workshops (ICMEW)*, Jul. 2018, pp. 1–6.

[7] X. Bai, N. Gao, Z. Zhang, and D. Zhang, "3D palmprint identification combining blocked ST and PCA," *Pattern Recognit. Lett.*, vol. 100, pp. 89–95, Dec. 2017.

[8] L. Leng, J. Zhang, M. K. Khan, X. Chen, and K. Alghathbar, "Dynamic weighted discrimination power analysis: A novel approach for face and palmprint recognition in DCT domain," *Int. J. Theor. Phys.*, vol. 5, no. 17, pp. 2543–2554, Dec. 2020.

[9] T. Connie, A. Teoh, M. Goh, and D. Ngo, "Palmprint recognition with PCA and ICA," in *Proc. Image Vis. Comput.*, Nov. 2003, pp. 227–232.

[10] L. Fei, J. Wen, Z. Zhang, K. Yan, and Z. Zhong, "Local multiple directional pattern of palmprint image," in *Proc. 23rd Int. Conf. Pattern Recognit. (ICPR)*, Dec. 2016, pp. 3013–3018.

[11] L. Leng, M. Li, C. Kim, and X. Bi, "Dual-source discrimination power analysis for multi-instance contactless palmprint recognition," *Multimedia Tools Appl.*, vol. 76, no. 1, pp. 333–354, Jan. 2017.

[12] A. Bruno, P. Carminetti, V. Gentile, M. La Cascia, and E. Mancino, "Palmprint principal lines extraction," in *Proc. IEEE Workshop Biometric Meas. Syst. Secur. Med. Appl. (BIOMS) Proc.*, Oct. 2014, pp. 50–56.

[13] S. Zhao and B. Zhang, "Deep discriminative representation for generic palmprint recognition," *Pattern Recognit.*, vol. 98, Feb. 2020, Art. no. 107071.

[14] J.-H. Kim, B.-G. Kim, P. P. Roy, and D.-M. Jeong, "Efficient facial expression recognition algorithm based on hierarchical deep neural network structure," *IEEE Access*, vol. 7, pp. 41273–41285, 2019.

[15] Z. Yang, L. Leng, and W. Min, "Extreme downsampling and joint feature for coding-based palmprint recognition," *IEEE Trans. Instrum. Meas.*, vol. 70, pp. 1–12, 2021.

[16] L. Leng, Z. Yang, and W. Min, "Democratic voting downsampling for coding-based palmprint recognition," *IET Biometrics*, vol. 9, no. 6, pp. 290–296, Nov. 2020.

[17] A. S. Tarawneh, D. Chetverikov, and A. B. Hassanat, "Pilot comparative study of different deep features for palmprint identification in low-quality images," 2018, *arXiv:1804.04602*. [Online]. Available: http://arxiv.org/abs/1804.04602

[18] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in *Proc. 25th Int. Conf. Neural Inf. Process. Syst. (NIPS)*, 2012, pp. 1097–1105.

[19] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," 2014, *arXiv:1409.1556*. [Online]. Available: http://arxiv.org/abs/1409.1556

[20] R. Ramachandra, K. B. Raja, S. Venkatesh, S. Hegde, S. D. Dandappanavar, and C. Busch, "Verifying the newborns without infection risks using contactless palmprints," in *Proc. Int. Conf. Biometrics (ICB)*, Feb. 2018, pp. 209–216.

[21] J. Svoboda, J. Masci, and M. M. Bronstein, "Palmprint recognition via discriminative index learning," in *Proc. 23rd Int. Conf. Pattern Recognit. (ICPR)*, Dec. 2016, pp. 4232–4237.

[22] Y. Wen, K. Zhang, Z. Li, and Y. Qiao, "A discriminative feature learning approach for deep face recognition," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, Sep. 2016, pp. 499–515.

[23] D. Zhong and J. Zhu, "Centralized large margin cosine loss for open-set deep palmprint recognition," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 30, no. 6, pp. 1559–1568, Jun. 2020.

[24] D. Palma, P. L. Montessoro, G. Giordano, and F. Blanchini, "Biometric palmprint verification: A dynamical system approach," *IEEE Trans. Syst., Man, Cybern. Syst.*, vol. 49, no. 12, pp. 2676–2687, Dec. 2019.

[25] W. M. Matkowski, T. Chai, and A. W. K. Kong, "Palmprint recognition in uncontrolled and uncooperative environment," *IEEE Trans. Inf. Forensics Security*, vol. 15, pp. 1601–1615, Oct. 2020.

[26] T. Chai, S. Prasad, and S. Wang, "Boosting palmprint identification with gender information using DeepNet," *Future Gener. Comput. Syst.*, vol. 99, pp. 41–53, Oct. 2019.

[27] X. Du, D. Zhong, and H. Shao, "Cross-domain palmprint recognition via regularized adversarial domain adaptive hashing," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 31, no. 6, pp. 2372–2385, Jun. 2021.

[28] Y. Liu and A. Kumar, "Contactless palmprint identification using deeply learned residual features," *IEEE Trans. Biometrics, Behav., Identity Sci.*, vol. 2, no. 2, pp. 172–181, Apr. 2020.

[29] S. Minaee and Y. Wang, "Palmprint recognition using deep scattering network," in *Proc. IEEE Int. Symp. Circuits Syst.*, May 2017, pp. 1–4.

[30] A. Genovese, V. Piuri, F. Scotti, and K. N. Plataniotis, "PalmNet: Gabor-PCA convolutional networks for touchless palmprint recognition," *IEEE Trans. Inf. Forensics Security*, vol. 14, no. 12, pp. 3160–3174, Dec. 2019.

[31] Y. C. Chen, M. H. Lim, P. C. Yuen, and J. H. Lai, "Discriminant spectral hashing for compact palmprint representation," in *Proc. Chin. Conf. Biometric Recognit.*, Nov. 2013, pp. 225–232.

[32] J. Cheng, Q. Sun, J. Zhang, and Q. Zhang, "Supervised hashing with deep convolutional features for palmprint recognition," in *Proc. 12th Chin. Conf. Biometric Recognit.*, Oct. 2017, pp. 259–268.

[33] D. Zhong, H. Shao, and X. Du, "A hand-based multi-biometrics via deep hashing network and biometric graph matching," *IEEE Trans. Inf. Forensics Security*, vol. 14, no. 12, pp. 3140–3150, Dec. 2019.

[34] D. Zhong, S. Liu, W. Wang, and X. Du, "Palm vein recognition with deep hashing network," in *Proc. Chin. Conf. Pattern Recognit. Comput. Vis. (PRCV)*. Cham, Switzerland: Springer, 2018, pp. 38–49.

[35] D. Li, Y. Gong, D. Cheng, W. Shi, X. Tao, and X. Chang, "Consistency-preserving deep hashing for fast person re-identification," *Pattern Recognit.*, vol. 94, pp. 207–217, Oct. 2019.

[36] Y. Liu, J. Song, K. Zhou, L. Yan, L. Liu, F. Zou, and L. Shao, "Deep self-taught hashing for image retrieval," *IEEE Trans. Cybern.*, vol. 49, no. 6, pp. 2229–2241, Jun. 2019.

[37] M. Jaderberg, K. Simonyan, A. Zisserman, and K. Kavukcuoglu, "Spatial transformer networks," 2015, *arXiv:1506.02025*. [Online]. Available: http://arxiv.org/abs/1506.02025

[38] N. Xu, Q. Zhu, X. Xu, and D. Zhang, "An effective recognition approach for contactless palmprint," *Vis. Comput.*, vol. 37, no. 4, pp. 695–705, Apr. 2021.

[39] *PolyU Palmprint Database (Version 2.0)*. [Online]. Available: https://www.comp.polyu.edu.hk/~biometrics

[40] *IITD Touchless Palmprint Database (Version 1.0)*. [Online]. Available: http://www4.comp.polyu.edu.hk/?csajaykr/IITD/Database_Palm.htm

[41] *Tongji Palmprint Image Database*. [Online]. Available: https://cslinzhang.github.io/ContactlessPalm/

[42] L. Zhang, Z. Cheng, Y. Shen, and D. Wang, "Palmprint and palmvein recognition based on DCNN and a new large-scale contactless palmvein dataset," *Symmetry*, vol. 10, no. 4, p. 78, 2018.

[43] D. Zhang, W. K. Kong, J. You, and M. Wong, "Online palmpint identification," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 25, no. 9, pp. 1041–1050, Sep. 2003.

[44] Z. Sun, T. Tan, Y. Wang, and S. Z. Li, "Ordinal palmprint represention for personal identification," in *Proc. Int. Conf. Comput. Vis. Pattern Recognit.*, Jun. 2005, pp. 279–284.

[45] A. Kong, D. Zhang, and M. Kamel, "Palmprint identification using feature-level fusion," *Pattern Recognit.*, vol. 39, no. 3, pp. 478–487, Mar. 2006.

[46] A. W.-K. Kong and D. Zhang, "Competitive coding scheme for palmprint verification," in *Proc. 17th Int. Conf. Pattern Recognit. (ICPR)*, vol. 1, Aug. 2004, pp. 520–523.

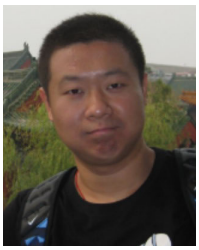
[47] W. Jia, D.-S. Huang, and D. Zhang, "Palmprint verification based on robust line orientation code," *Pattern Recognit.*, vol. 41, no. 5, pp. 1504–1513, 2008.

[48] L. Fei, Y. Xu, and D. Zhang, "Half-orientation extraction of palmprint features," *Pattern Recognit. Lett.*, vol. 69, pp. 35–41, Sep. 2016.

[49] L. Fei, Y. Xu, W. Tang, and D. Zhang, "Double-orientation code and nonlinear matching scheme for palmprint recognition," *Pattern Recognit.*, vol. 49, pp. 89–101, Jan. 2016.
[50] Y. Xu, L. Fei, J. Wen, and D. Zhang, "Discriminative and robust competitive code for palmprint recognition," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 48, no. 2, pp. 232–241, Feb. 2018.
[51] Z. Guo, D. Zhang, L. Zhang, and W. Zuo, "Palmprint verification using binary orientation co-occurrence vector," *Pattern Recognit. Lett.*, vol. 30, no. 13, pp. 1219–1227, 2009.

TENGFEI WU received the B.S. degree from Hubei Engineering University, Wuhan, China, in 2018. He is currently pursuing the M.S. degree with Nanchang Hangkong University, Nanchang, China.

His research interests include pattern recognition, image processing, and deep learning.



LU LENG (Member, IEEE) received the Ph.D. degree from Southwest Jiaotong University, Chengdu, China, in 2012.

He did his postdoctoral research at Yonsei University, Seoul, Republic of Korea, and Nanjing University of Aeronautics and Astronautics, Nanjing, China. From 2015 to 2016, he was a Visiting Scholar with West Virginia University, USA, and Yonsei University, from 2019 to 2020. He is currently an Associate Professor with Nanchang Hangkong University. He has published more than 90 international journal articles and conference papers, and been granted several scholarships and funding projects in his academic research. His research interests include computer vision, biometric template protection, and biometric recognition.

Dr. Leng is a member of the Association for Computing Machinery (ACM), China Society of Image and Graphics (CSIG), China Computer Federation (CCF), and China Association of Artificial Intelligence (CAAI). He is a reviewer of several international journals and conferences.

MUHAMMAD KHURRAM KHAN (Senior Member, IEEE) is currently working as a Professor of cybersecurity with the Center of Excellence in Information Assurance, King Saud University, Saudi Arabia. He is the Founder and the CEO of Global Foundation for Cyber Studies and Research (http://www.gfcyber.org), an independent and non-partisan cybersecurity think-tank in Washington, DC, USA. He is the Editor-in-Chief of *Telecommunication Systems* (Springer-Nature) with its recent impact factor of 2.314 (JCR 2021). He is also the Editor-in-Chief of *Cyber Insights Magazine* (http://www.cyber-insights.org). He is on the editorial board of several journals, including IEEE Communications Surveys and Tutorials, *IEEE Communications Magazine*, IEEE Internet of Things Journal, IEEE Transactions on Consumer Electronics, *Journal of Network and Computer Applications* (Elsevier), IEEE Access, *IEEE Consumer Electronics Magazine*, *PLOS ONE*, and *Electronic Commerce Research*. He has published more than 400 papers in the journals and conferences of international repute. In addition, he is an Inventor of ten U.S./PCT patents. He has edited ten books/proceedings published by Springer-Verlag, Taylor & Francis, and IEEE. His research interests include cybersecurity, digital authentication, the IoT security, biometrics, multimedia security, cloud computing security, cyber policy, and technological innovation management. He is a fellow of the IET (U.K.), the BCS (U.K.), and the FTRA (South Korea). More details can be found at: http://www.professorkhurram.com.

FARRUKH ASLAM KHAN (Senior Member, IEEE) received the M.S. degree in computer system engineering from the GIK Institute of Engineering Sciences and Technology, Pakistan, in 2003, and the Ph.D. degree in computer engineering from Jeju National University, South Korea, in 2007. He also received professional training from Massachusetts Institute of Technology, New York University, IBM, and other professional institutions. He is currently a Professor of cybersecurity with the Center of Excellence in Information Assurance, King Saud University, Riyadh, Saudi Arabia. He is also the Founding Director of the Wireless Networking and Security Research Group at the National University of Computer and Emerging Sciences, Islamabad, Pakistan. He has over 120 publications in refereed international journals and conferences. He has co-organized several international conferences and workshops. He has successfully supervised/co-supervised five Ph.D. students and 18 M.S. thesis students. Several M.S. and Ph.D. students are currently working under his supervision. His research interests include cybersecurity, body sensor networks and e-health, bio-inspired and evolutionary computation, and the Internet of Things. He is on the panel of reviewers of over 40 reputed international journals and numerous international conferences. He serves/served as an Associate Editor for prestigious international journals, including IEEE Access, *PLOS ONE*, *Neurocomputing* (Elsevier), *Ad Hoc and Sensor Wireless Networks*, *KSII Transactions on Internet and Information Systems*, *Human-Centric Computing and Information Sciences* (Springer), *PeerJ Computer Science*, and *Complex and Intelligent Systems* (Springer). He is a fellow of the British Computer Society (BCS).