Contactless Palmprint Image Recognition Across Smartphones With Self-Paced CycleGAN

Qi Zhu, Guangnan Xin, Lunke Fei, *Senior Member, IEEE*, Dong Liang, Zheng Zhang, *Senior Member, IEEE*, Daoqiang Zhang, *Senior Member, IEEE*, and David Zhang, *Life Fellow, IEEE*

***Abstract*— Contactless palmprint recognition, an emerging biometric technology, has attracted increasing attention due to its noninvasive and high practicability characteristics. Although it is naturally suitable for mobile application scenarios, the following two challenges severely limit its recognition performance: 1) the inconsistency in acquisition devices used in training and testing, and 2) many subjects are unable to be imaged on each device, resulting in incomplete data problems. To address these issues, we propose a self-paced CycleGAN with self-attention modules, which simultaneously synthesizes missing data and alleviates the influence of different imaging devices. Specifically, we develop CycleGAN with self-attention modules to generate missing training data by effectively mining the structural correlation among samples while capturing the cross-domain features. Furthermore, a self-paced learning strategy, which is a human cognitive-driven learning mechanism, is used to guide learning the robust cross-domain feature representation and recognition model, by which the relatively easy learning samples are gradually involved in the training process. To verify the effectiveness of the proposed method, we conduct experiments on contactless palmprint datasets collected using different smartphones. The results show that our approach outperforms state-of-the-art methods in classifying contactless palmprint images.**

***Index Terms*— Biometric, contactless palmprint, cross-domain recognition, image-to-image translation, self-paced learning, self-attention.**

Manuscript received 2 May 2022; revised 6 November 2022 and 21 May 2023; accepted 30 July 2023. Date of publication 3 August 2023; date of current version 14 August 2023. This work was supported in part by the National Natural Science Foundation of China under Grant 62076129, Grant 61501230, Grant 61732006, Grant 61876082, Grant 61861130366, and Grant 62006115; in part by the Key Research and Development Plan of Jiangsu Province under Grant BE2022842; and in part by the National Key Research and Development Programme of China under Grant 2018YFC2001600 and Grant 2018YFC2001602. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Luisa Verdoliva. *(Corresponding author: Daoqiang Zhang.)*

Qi Zhu, Guangnan Xin, Dong Liang, and Daoqiang Zhang are with the College of Computer Science and Technology, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China, and also with the Key Laboratory of Brain-Machine Intelligence Technology, Ministry of Education, Nanjing University of Aeronautics and Astronautics, Nanjing 211106, China (e-mail: zhuqi@nuaa.edu.cn; xin_gn@nuaa.edu.cn; liangdong@nuaa.edu.cn; dqzhang@nuaa.edu.cn).

Lunke Fei is with the School of Computer Science and Technology, Guangdong University of Technology, Guangzhou 510006, China (e-mail: flksxm@126.com).

Zheng Zhang is with the Shenzhen & Peng Cheng Laboratory, Harbin Institute of Technology, Nanshan, Shenzhen 518055, China (e-mail: darrenzz219@gmail.com).

David Zhang is with the School of Data Science, The Chinese University of Hong Kong (Shenzhen), Shenzhen 518172, China (e-mail: davidzhang@cuhk.edu.cn).

Digital Object Identifier 10.1109/TIFS.2023.3301729

I. Introduction

COMPARED with traditional palmprint recognition technology, contactless palmprint recognition has better user acceptance and ease of use [1] because of fewer hand position restrictions and more flexible image acquisition scenarios. Based on the above inherent advantages, contactless palmprint recognition technology has wide application prospects in identity authentication [2], [3].

In practice, contactless palmprint recognition easily suffers from noise and varies in illumination translations, scales, and rotations. Most of the previous palmprint recognition approaches [4], [5], [6] were originally derived for contact palmprint images and cannot achieve satisfactory performance on contactless palmprint recognition tasks [7]. Recently, some contactless palmprint feature extraction and recognition algorithms have been proposed. An iterative RANSAC (I-RANSAC) algorithm was introduced in [8] to refine matched scale-invariant feature transform (SIFT) points. Xu et al. [9] applied a spatial transformation network to align images and exploited the residual network for classification problems. Ajay Kumar et al. [10] developed a deformation alignment and matching approach to achieve feature alignment of the local regions to alleviate the effect of deformations on classification performance. A subspace method [11] was developed to simultaneously capture the similarity and the low-rank structural data relationships. In [12], a fusion model with a quaternion matrix was proposed for multi-domain palmprint image feature extraction.

Notably, most of the existing contactless palmprint recognition methods are based on the assumption that training and test data are collected from the same imaging device, which implies that the samples have the same data domain. However, the above condition is difficult to guarantee in practical contactless palmprint applications. Often, the palmprint acquisition equipment used for testing is different from the training data acquisition equipment. When using the smartphone as the imaging acquisition device for palmprint recognition applications, the users can collect palmprint images conveniently without the need for equipping the explicitly designed imaging devices, thereby expanding the application scenarios of palmprint recognition. For example, in many distributed systems, palmprint authentications can be performed on remote authentication servers, which requires collecting palmprint images from different devices, such as smartphones. In mobile payment applications, the user can upload palmprint images

captured by his or her smartphone to remote servers for enrolment. When checking out, the salesperson can ask the user to show his or her palm for payment. The salesperson captures the palmprint image with the smartphone and uploads the palmprint images to the remote server for authentication. Many people change their smartphones after using them for a period, so they need to use palmprint images collected by different devices for authentication. With the development of contactless palmprint recognition, cross-device palmprint recognition application scenarios will be increasingly common. Therefore, data heterogeneity and distribution differences are challenging for contactless palmprint recognition across smartphones.

The incomplete data problem is widespread in contactless palmprint applications across devices such as smartphones. There are many kinds of imaging devices in real-world applications, especially for distributed identity authentication systems, which makes it difficult to guarantee that the palmprint images of all users will be collected from each device. Therefore, during both the training and testing phases, it is a usual situation that palmprint images of some users may be collected by partial devices, while the images of these users from other devices are missing, as shown in Fig.1(a), which leads to poor recognition results. Moreover, the impact of noise and pose variations in contactless palmprint images also lead challenging problem in both data complement and recognition.

In this paper, to solve the incomplete multi-domain contactless palmprint recognition problem, we propose a cycled self-paced generative adversarial network (SPCGAN), which is a cognitive-driven multi-view data synthesis and fusion method with a self-attention mechanism [13]. First, a double-channel generative adversarial network is employed to fuse the user identity information from one data domain and the image pattern from other data domains by two inverse generators and cycle-consistency loss. Fake missing images can be generated using this model, where the data distribution of generated images is similar to the missing data while the identity information remains the same. Second, since traditional convolutional generative adversarial networks are limited by the size of convolution kernels, we introduce self-attention modules to model long-range dependencies, which can facilitate maintaining the internal correlations with a small computational cost. Finally, we introduce a self-paced learning (SPL) strategy, which is an easy-to-hard human learning mechanism, by ordering the training images in a meaningful way. This strategy can alleviate the impacts of noise and pose variations in contactless palmprint recognition and improve the stability of the generative adversarial module. We define significant quality scores to dynamically measure the learning ease of training samples, by which relatively easy-simple learning images are gradually involved in training progress to learn the robust recognition model.

The contributions of this article can be summarized as follows:

1) To our knowledge, this is the first work to investigate incomplete contactless palmprint recognition across data acquisition equipment.
2) We propose a self-paced CycleGAN, which is called SPCGAN, to simultaneously synthesize missing cross-device contactless palmprint images and fuse these data by mining the complementary information among data domains.
3) To alleviate the influence of noise and deformation from contactless palmprint images and improve the stability of the generative adversarial module, we introduce the SPL-based cognitive-driven mechanism to the model to guide the gradual involvement of relatively reliable training samples in the training process.
4) The experimental results on two contactless palmprint databases imaged by different smartphones show that the proposed method outperforms several state-of-the-art methods.

The remainder of this paper is structured as follows. Section II presents a detailed review of the related relevant techniques for palmprint recognition and image-to-image translation. Section III introduces the proposed self-paced CycleGAN method. Section IV describes the experimental results. Finally, Section V concludes the work.

II. Related Work

A. Palmprint Recognition

In recent years, feature learning-based approaches have been the most widely used in palmprint recognition [7]. Convolutional neural networks (CNNs), including AlexNet [14], VGG-16 [15], ResNet-50 [16], and Inception-V3 [17], are applied for feature extraction and classification because of their great success in large-scale image recognition. Moreover, feature learning-based methods are also applied for contactless palmprint recognition. In [18], AlexNet with an eight-layer architecture was introduced for contactless palmprint feature extraction. Similarly, a fine-tuned pretrained AlexNet [19] achieved promising verification accuracy on contactless palmprint images of newborn babies. In addition, CNNs are used as fusion approaches in multi-modal fusion with contactless palmprints and other biometric traits. For example, PCANet, a CNN trained using an unsupervised procedure based on principal component analysis (PCA), was used for fusing palmprint and inner finger texture in [20]. In [21], a transfer autoencoder was developed for cross-domain palmprint recognition. Features with the same distribution were extracted after alternating autoencoder and discriminator training. Shao et al. proposed a two-stage alignment, called joint pixel and feature alignment [22], to reduce the dataset gaps at the pixel level and align the distributions at the feature level.

However, most existing cross-domain contactless palmprint recognition approaches are trained on complete multi-domain data. It cannot be guaranted in the real application of contactless palmprint applications with multiple acquisition devices, which leads to poor classification accuracy and an inability to predict samples with missing data in some domains.

B. Image-to-Image Translation

Image-to-image translation learns the mappings among images in different domains and can be used to recover the missing samples by cross-domain feature representation. Generative adversarial networks (GANs) have achieved remarkable

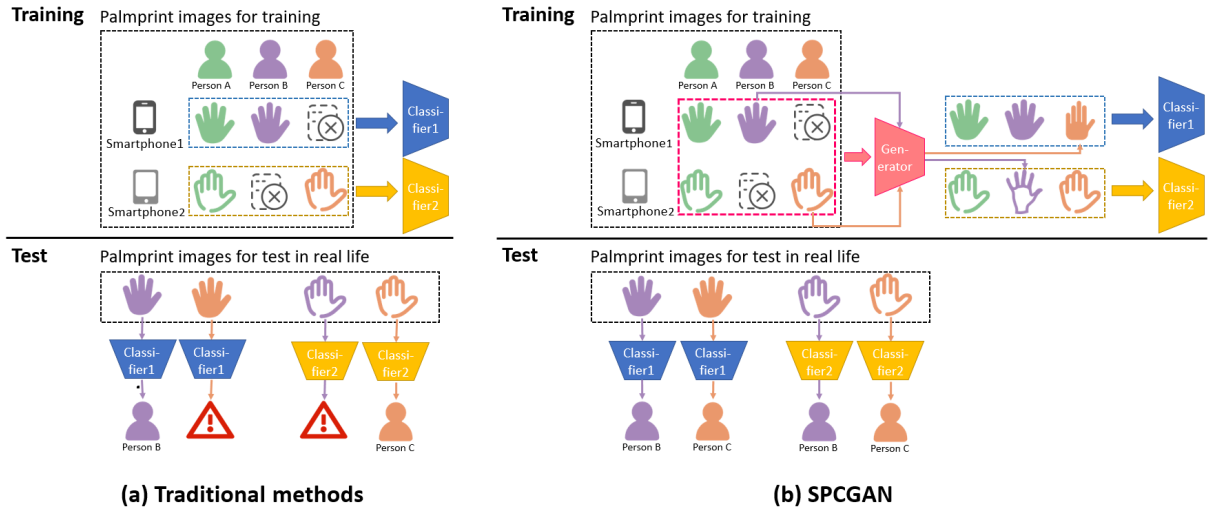


Fig. 1. An example of the missing data problem in real contactless palmprint recognition across smartphones. (a) Traditional methods: in the training process, the palmprint images of some subjects may be missing in partial domains, while existing in other domains which leads to false recognition results of those subjects that are not involved in the training process. (b) SPCGAN: our method simultaneously synthesizes missing data and alleviates the influence of different imaging devices, by which the missing subjects are expected to be correctly classified.

performance in image translation tasks [23], [24], [25] due to their great feature representation learning capacity. GANs mainly contain two modules, a generator and a discriminator. The generator learns to generate fake data, and the discriminator attempts to distinguish between real and fake data. In [26], a general-purpose model called pix2pix was proposed for supervised image-to-image translation tasks. A U-Net-based architecture is applied as the generator in pix2pix, and the discriminator is designed as PatchGAN. Although pix2pix has demonstrated excellent performance, such as [27], [28], and [29], it requires well-aligned paired training images, which are not available in many contactless palmprint recognition applications. Zhu et al. proposed CycleGAN [30] for image translation without paired training samples. In CycleGAN, the generator is designed as a cycle architecture for unpaired data and cycle-consistency loss is introduced to keep the key information between the input and output of the cycle architecture. In [30], Zhu et al. found that the CycleGAN model is robust with colour and texture changes but sensitive to geometric changes in the image-to-image translation task.

C. Self-Paced Learning

SPL was proposed to simulate the human learning process [31]. It progressively arranges learning tasks and gradually obtains a robust learning model. Recently, SPL has shown efficient results on many learning tasks, in which training samples are gradually involved from easy to hard according to the learning task. By defining a computable measure for learning hardness of samples, SPL can avoid falling into local minima and improve the generalization of the learning model. The SPL strategy is adopted with a regularization term in the objective function to evaluate the hardness of training samples. Jiang et al. [32] considered diversity and proposed a more generalized approach called self-paced learning with diversity (SPLD). The SPLD regularization term is independent of the objective function. In [33], a multi-objective SPL method was introduced, which treats the SPL problem as a multi-objective issue. Many researchers have adopted SPL in their tasks, such as multi-modal few-shot learning [34], long-term tracking [35], and co-saliency detection [36]. In this paper, we exploit SPL to reduce the learning difficulties of image-to-image translation tasks on incomplete multi-domain contactless palmprint images.

III. Proposed Method

A. Architecture of Proposed SPCGAN

To synthesize the missing samples in the target domain from the source domain, we transform the image data distribution in the source domain to be similar to the target domain by CycleGAN armed with self-attention modules [13].

Given domains A and B and training samples $\{a_i\}_{i=1}^{M_A} \in A$ and $\{b_j\}_{j=1}^{M_B} \in B$, where M_A and M_B are the sample numbers in domains A and B, respectively. There are two mapping functions $G : A \rightarrow B$ and $F : B \rightarrow A$ in this model. In addition, there are two associated adversarial discriminators, D_A and D_B. D_A encourages F to translate B into outputs indistinguishable from A. In the same way, D_B aims to distinguish between $\{b\}$ and $\{G(a)\}$. The generator and discriminator architectures are shown in Fig. 2. Considering that geometrical characteristics are one of the most significant contactless palmprint image features [7], a convolutional generator may not achieve significant performance for the multi-domain contactless palmprint image translation task, because long-range dependencies cannot be captured only by several convolution operators [37]. To address this issue, we introduce self-attention modules as a complement to convolution operators to model long-range dependencies efficiently, as shown in Fig. 3. Convolution feature maps $x \in \mathbb{R}^{(C \times N)}$, where C is the number of channels and N is the number of features, are transformed into three feature spaces q, k, v, where $q(x) = W_q x, k(x) = W_k x, v(x) = W_v x$. The attention map $Attn$ and the output of the self-attention layers O can be calculated as:

$$Attn = softmax(qk^{\mathrm{T}}), \tag{1}$$

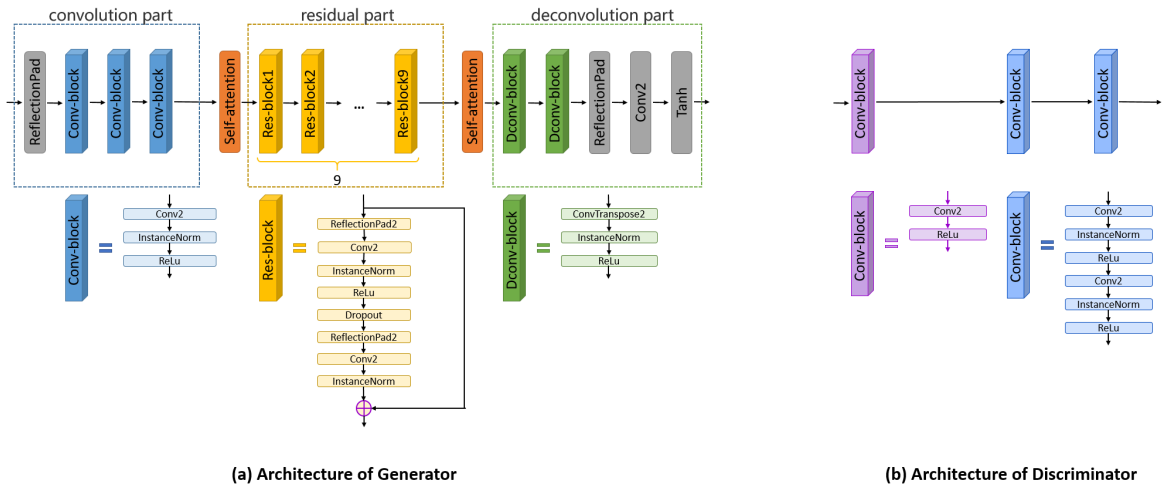


Fig. 2. Architecture of the proposed model for contactless palmprint image recognition across domains.

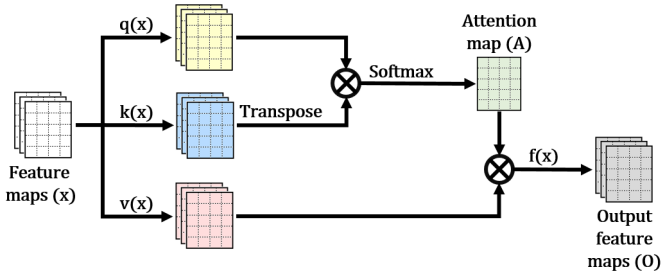


Fig. 3. The framework of self-attention model, where ⊗ denotes matrix multiplication.

$$O = f(Attnv),\ f(x) = W_f x, \tag{2}$$

where W_f is the learned weight matrix. In addition, a learning rate γ is introduced to modify the self-attention layer weight. The learnable scalar γ is initialized as 0 and learns to assign a larger weight gradually. Therefore, the final output of the self-attention module is given by

$$y = \gamma O + x. \tag{3}$$

We fuse two self-attention modules in the generator before and after the residual part, as shown in Fig. 2.

The typical overall loss function of CycleGAN is:

$$\begin{aligned}\mathcal{L}(G, F, D_A, D_B) = &\mathcal{L}_{GAN}(G, D_B, A, B)\\ &+ \mathcal{L}_{GAN}(F, D_A, B, A)\\ &+ \lambda\mathcal{L}_{cyc}(G, F),\end{aligned} \tag{4}$$

where $\mathcal{L}_{GAN}$ and $\mathcal{L}_{cyc}$ represent adversarial loss and cycle-consistency loss, respectively. λ controls the relative importance of cycle-consistency loss. In particular, we introduce identity loss to ensure identity invariance:

$$\begin{aligned}\mathcal{L}_{idt}(G, F) = &\mathbb{E}_{a\sim p_{data}(A)}[\parallel G(a) - a \parallel_1]\\ &+ \mathbb{E}_{b\sim p_{data}(B)}[\parallel F(b) - b \parallel_1].\end{aligned} \tag{5}$$

Finally, the optimizing objective can be obtained by:

$$\begin{aligned}\mathcal{L}_{SPCGAN}(G, F, D_A, D_B) = &\mathcal{L}_{GAN}(G, D_B, A, B)\\ &+ \mathcal{L}_{GAN}(F, D_A, B, A)\\ &+ \lambda\mathcal{L}_{cyc}(G, F)\\ &+ \mu\mathcal{L}_{idt}(G, F),\end{aligned} \tag{6}$$

where μ controls the relative importance of identity loss.

B. Self-Paced Learning Strategy

Compared with contact-based palmprint images, in practice, contactless palmprint images have more deformation and noise. In addition, despite GAN's excellent success, it still has obstacles to stable training due to the Nash equilibrium, vanishing gradient, and mode collapse. To alleviate the influence of the above aspects, we introduce SPL into the proposed model. The core of SPL is to iteratively sort training samples via human learning [31]. Motivated by [38], we developed an SPL strategy based on the quality score to induce the samples to participate in the training process. In our method, the quality score is defined to quantify the learning difficulties of generated contactless palmprint images. This measure is composed of two parts: the consistency score and the identity score. The consistency score is defined to judge the gap between the input images and the output images after two transformations ($A - B - A$), and the identity score measures the identity invariance. For domain B, the score is as follows:

$$\begin{aligned}Score_b = &\mathbb{E}_{b\sim p_{data}(B)}[\parallel F(G(b)) - b \parallel_1]\\ &+ \mathbb{E}_{b\sim p_{data}(B)}[\parallel F(b) - b \parallel_1].\end{aligned} \tag{7}$$

A similar quality score is defined in the reverse direction for domain A. We conducted several experiments to verify the effectiveness of adversarial loss, consistency loss, and identity loss in quality scores in Section IV.

We developed a self-paced training scheme based on quality scores, as shown in Fig. 4, by which reliable images are gradually involved in training progress. Assume that dataset $X_{all} = X_{train} \cup X_{rest}$, $X_{train} \cap X_{rest} = \emptyset$, where X_{all} denotes all training data, X_{train} denotes data involved in training and X_{rest} denotes the remaining training data. In the first iteration, X_{all} is randomly divided into two parts: X^1_{train} and X^1_{rest},

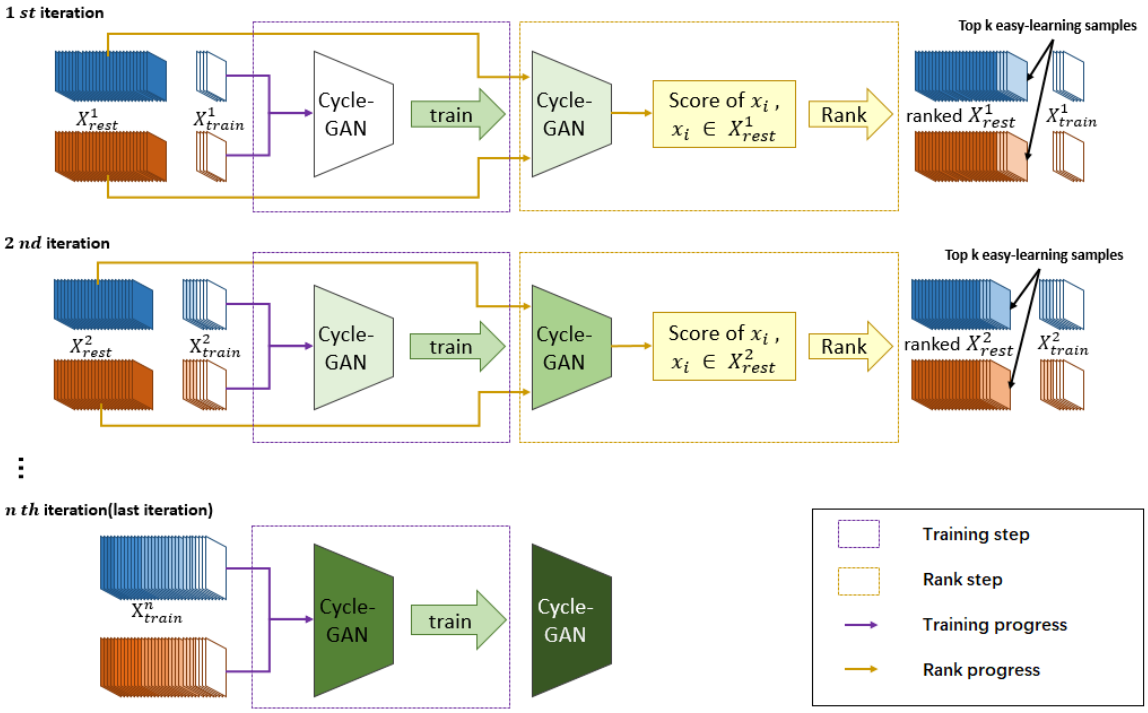


Fig. 4. The proposed cross-domain fusion model with the SPL strategy framework. A quality score is proposed to measure sample learning ease. Contactless palmprint images are divided into a training group and a rest group. In the first iteration, the training samples X^1_{train} are selected randomly. When the loss function converges, the quality scores of the remaining samples X^1_{rest} are calculated by the trained model. Then, the top k easy learning samples are iteratively selected into the training group to learn a more robust model. In the last iteration, all of the training data are involved in the training process.

where the proportion of X^1_{train} in X_{all} is r_{ini}. In each iteration, there are two steps: a training step and a ranking step. In the training step, SPCGAN is trained with X_{train} until the loss function converges. Then, in the ranking step, the trained model is used to calculate the quality scores of X_{rest}. The samples in X_{rest} are ranked from easy to hard according to the quality scores. The top k samples are selected as training data for the next iteration. k specifies a series of maximum numbers of the selected samples in X_{rest} for the next iteration and is designed to be related to the sample number m_{rest} of X_{rest}, where $k = r_k m_{rest}$. In the nth iteration, which is the last iteration, all of the training data are involved in the training process. The pseudocode of the proposed method is given in Algorithm 1.

Algorithm 1 Self-Paced Learning Strategy.
CALSCORE( ) Returns the Quality Scores.
RANK( ) Is a Function That Ranks the Images Based on Quality Scores.
TOP( ) Returns the Lowest-Ranked Images.

Require: training set X_{all}
Ensure: mapping functions G, F
1: Randomly divide training set X_{all} into X_{train} and X_{rest}, where $X_{all} = X_{train} \cup X_{rest}$, $X_{train} \cap X_{rest} = \emptyset$, and the proportion of X_{train} is r_{ini}.
2: For i = 1,2...n-1 do
3: While not converged do
4: Update G, F, D_A, D_B with X_{train}
5: Update $k : k = r_k m_{rest}$
6: $Score$ = CALSCORE(X_{rest}, G, F)
7: X_{sorted} = RANK (X_{rest}, $Score$)
8: X_{topk} = TOP(X_{sorted}, k)
9: Update $X_{train} : X_{train} = X_{train} \cup X_{topk}$
10: Update $X_{rest} : X_{rest} = X_{rest} - X_{topk}$
11: While not converged do
12: Update G, F, D_A, D_B with X_{train}

IV. Experiments

A. Databases of Contactless Palmprint Across Smartphones

To evaluate the recognition accuracy of the proposed method, we considered two contactless palmprint databases imaged by smartphones. Because of palmprint pose variation and the complicated backgrounds and illumination conditions in uncontrolled environments, extracting regions of interest (ROIs) from contactless palmprint images is challenging. ROI extraction and classification are both important topics in contactless palmprint recognition. This work focuses on cross-device classification issues, while [3] and [39] proposed ROI extraction approaches for contactless palmprints.

Mobile Palmprint Database (MPD) [3] has 16,000 palmprint images from 200 subjects. Two smartphones with different cameras, Huawei and Xiaomi, were used to collect palm images. Each volunteer was asked to provide 10 images of each hand with each camera. In addition, there were two collection rounds in two sessions, which avoided the time influence. A local coordinate system with "double-finger-gap" and "palm-center" in the training contactless palmprint image

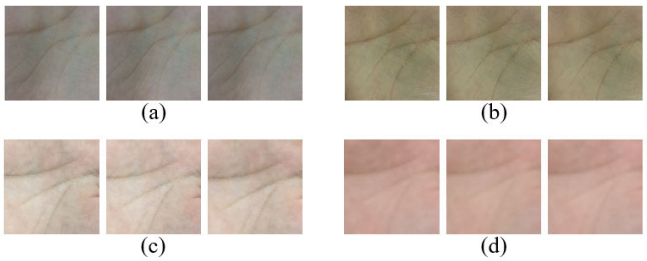


Fig. 5. Typical palmprint ROI images. (a) and (b) are ROI images of the same sample collected by Xiaomi and Huawei, respectively, in the MPD database. (c) and (d) are selected from Huawei and iPhone images, respectively, in the XJTU-UP database, which belong to the same sample.

was established according to the finger-gap points and was used to obtain the ground truth for each training image. Then, the Tiny-YOLOV3 detector was trained to obtain the key points in test contactless palmprints, which helped to construct the local coordinate system for extracting the ROI. More details about ROI extraction on the MPD database can be found in [3]. All of the images were cropped as ROIs with a size of 1896 × 1896 pixels. In this paper, the images were divided into two datasets according to the acquisition smartphones, denoted as HW (Huawei) and XM (Xiaomi). There were 8,000 images belonging to 400 categories in each dataset.

Xi'an Jiaotong University Unconstrained Palmprint Database (XJTU-UP) [39] is an unconstrained palmprint database collected by smartphones, i.e., iPhone 6S, Huawei Mate8. One hundred volunteers were asked to take 10 photos of each hand under indoor natural light and a flashlight, holding the smartphone and choosing the hand angles as desired. There were 2,000 palmprint images belonging to 200 categories in total. A target recognition algorithm based on the histogram of oriented gradients (HOG) was applied for palm detection, and key points were identified in the detected palm area by regression tree algorithms. Based on key points, ROIs can be extracted through a distance-based approach. More details about the ROI extraction on the XJTU-UP database can be found in [39]. All of the images were extracted as ROIs with a size of 280 × 280 pixels. We chose two datasets, denoted as HWN (Huawei under natural illumination) and IPN (iPhone under natural illumination).

B. Implementation Details

Our method is trained with the Adam optimizer, the learning rate of 0.00001, 200 epochs in each iteration, and the batch size of 32. Contactless palmprint ROI images are resized to 64 × 64 pixels. For each experiment, we randomly divide the dataset into 10 folds. Each fold maintains the same category proportion as the original dataset. For example, there are 2 images from each subject in each fold on the MPD dataset, and there is 1 image for each subject from each fold on the XJTU-UP dataset. Afterwards, 1-fold is taken as the test set each time. Among the remaining 9 folds, 8 folds are used for the training set, and 1-fold is treated as the validation set to determine the value of the parameter with grid search. In the training set, we randomly remove 10%, 20%, and 30% of subjects from each domain, and the removed subjects in different domains do not overlap to ensure that all subjects exist in at least one domain.

C. Comparisons With Different Feature Learning-Based Approaches

We compared the performance of our method and different feature learning-based methods on each dataset with different missing rates. ResNet-50, AlexNet, and VGG-16 are applied as the classifiers and trained with complete data, where missing subjects in the missing domain are copied from the corresponding subjects of the other domain. Meanwhile, the same classifiers are trained with the dataset, which was completed using our method. ResNet-50 trained with complete ground-truth data is regarded as the baseline. The experimental conditions of these three classifiers (ResNet-50, AlexNet, and VGG-16) are the same. We apply the cross-entropy loss function, ADAM optimizer, learning rate of 0.0002, 200 epochs, and batch size of 32.

The top-1 classification accuracies on the MPD database are listed in Table I, where the baselines, trained with complete ground-truth data, on XM and HW. It is obvious that the proposed method achieved significant performance in every situation. In particular, when 10% of subjects were missing, the best top-1 accuracies were 100% for Xiaomi and Huawei on our method with ResNet-50, which was the same as the baseline and over 5% higher than the mixture with the same classifier. In addition, the top-1 recognition accuracy of our method showed a relatively slight decline as the missing rate rose. When 30% of subjects were missing, the best top-1 accuracies were 99.25% and 98.5% for Xiaomi and Huawei, respectively, on our method with ResNet-50 or VGG16. Compared with other methods, the top-1 recognition rate of our method was the closest to the baselines trained with complete ground-truth data.

The top-1 classification accuracies on the XJTU-UP database are listed in Table II, where the baselines on HWN and IPN are 98.96% and 99.48%, respectively. Our method outperformed the mixture methods in most situations. In particular, the best accuracies were 86.98% and 87.50% for HWN and IPN, respectively, on our method with ResNet-50 with a missing rate of 30%.

Fig. 6 and Fig. 7 show the cumulative match characteristic (CMC) curves on the MPD and XJTU-UP databases, respectively. The CMC curves confirmed our method outperformed the comparison approaches. It can also be observed in Fig. 6 and Fig. 7 that our method showed a significant improvement in the rank-5 recognition rate, which is the probability that the top five results with the highest confidence contain the correct result.

D. Comparisons With Other Image-to-Image Models

We conducted several experiments to compare the state-of-the-art image-to-image translation approaches, including pix2pix [26], CycleGAN [30], RAU-net [40], and AttentionGAN [41].

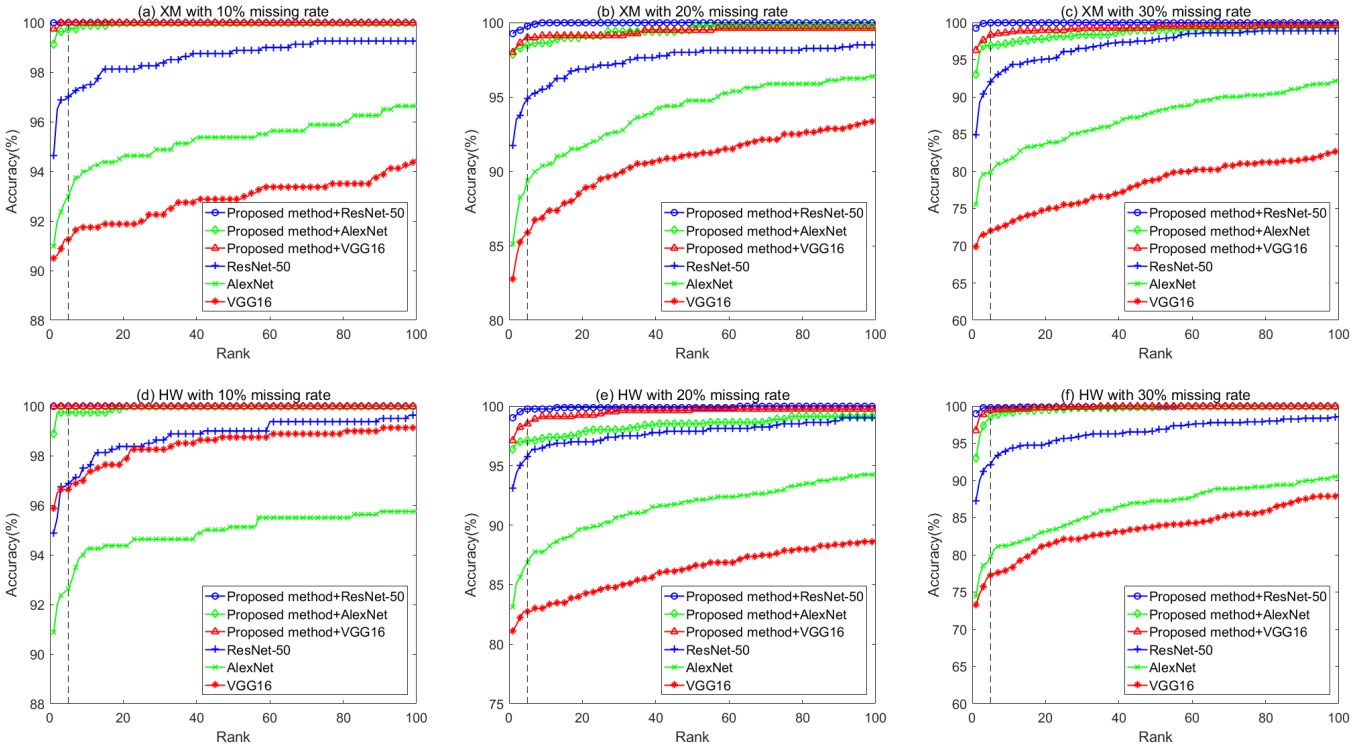


Fig. 6. CMC curves on MPD databases: (a-c) are the CMC curves on the XM dataset with missing rates of 10%, 20%, and 30%, respectively. (d-f) are the CMC curves on the HW dataset with missing rates of 10%, 20%, and 30%, respectively. (The dotted line represents rank=5.)

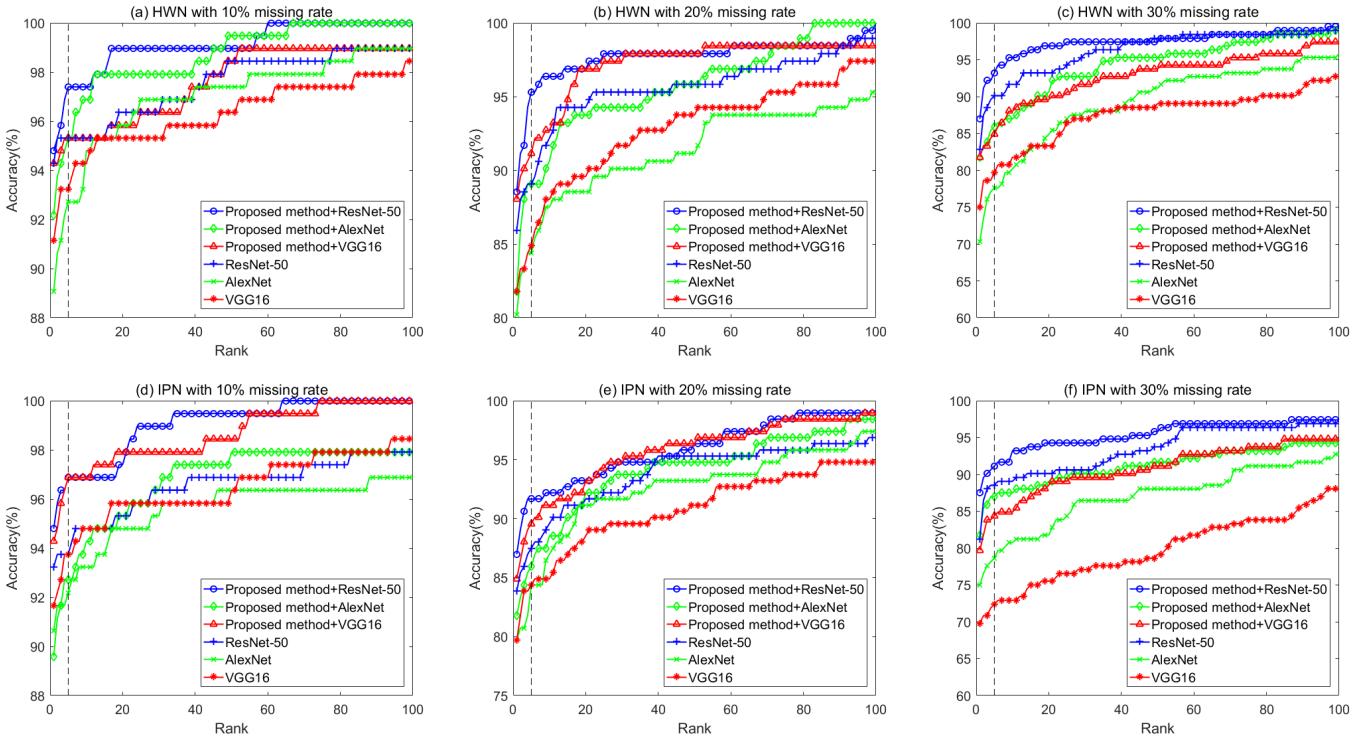


Fig. 7. CMC curves on the XJTU-UP databases: (a-c) are the CMC curves on the HWN dataset with missing rates of 10%, 20%, and 30%, respectively. (d-f) are the CMC curves on the IPN dataset with missing rates of 10%, 20%, and 30%, respectively. (The dotted line represents rank=5.)

1) Pix2pix is a general-purpose model for image-to-image supervised translation tasks.
2) CycleGAN can translate images from a source domain to a target domain using the cycle-consistency loss.
3) RAU-net is an improvement of pix2pix, which changes the generator from a traditional U-Net into a residual attention U-Net.
4) AttentionGAN can identify the most discriminative foreground objects and minimize the change in the

TABLE I

THE TOP-1 ACCURACY (%) OF DIFFERENT RECOGNITION METHODS ON THE MPD DATABASE

Missing Rate	10%		20%		30%	
Dataset	XM	HW	XM	HW	XM	HW
ResNet-50 [16]	94.625	94.875	91.750	93.125	84.875	73.750
Proposed method+ResNet50	100.00	100.00	99.250	99.000	99.250	99.000
AlexNet [14]	91.000	90.875	85.125	83.125	75.500	73.625
Proposed method+AlexNet	99.125	98.875	97.875	96.375	93.000	93.000
VGG-16 [15]	90.500	95.875	82.750	81.125	69.875	84.500
Proposed method+VGG16	99.750	100.00	98.000	97.125	96.750	99.000

NOTE: The baselines trained with complete ground-truth data on XM and HW are both 100.00%.

TABLE II

THE TOP-1 ACCURACY (%) OF DIFFERENT RECOGNITION METHODS ON THE XJTU-UP DATABASE

Missing Rate	10%		20%		30%	
Dataset	HWN	IPN	HWN	IPN	HWN	IPN
ResNet-50 [16]	94.27	93.23	86.98	85.42	82.81	81.25
Proposed method+ResNet50	94.79	94.79	88.54	86.98	86.98	87.50
AlexNet [14]	89.06	90.63	80.21	79.17	70.31	75.00
Proposed method+AlexNet	92.19	89.58	81.77	79.67	81.77	81.77
VGG-16 [15]	91.15	91.67	81.77	79.69	75.00	69.79
Proposed method+VGG16	94.27	94.27	88.02	84.90	81.77	79.67

NOTE: The baselines trained with complete ground-truth data on HWN and IPN are 98.96% and 99.48%, respectively.

TABLE III

THE TOP-1 ACCURACY (%) OF RESNET-50 WITH DIFFERENT DATA SYNTHESIS METHODS ON THE MPD BATABASE

Missing Rate	10%		20%		30%	
Dataset	XM	HW	XM	HW	XM	HW
Pix2pix [26]	88.875	89.000	78.625	76.125	69.625	69.250
CycleGAN [30]	93.000	94.000	87.125	85.500	81.625	82.625
RAU-net [40]	92.500	90.500	83.625	84.125	78.500	73.625
AttentionGAN [41]	98.250	97.250	97.875	97.625	97.375	96.625
Ours	100.00	100.00	99.250	99.000	99.250	99.000

TABLE IV

THE TOP-1 ACCURACY (%) OF RESNET-50 WITH DIFFERENT DATA SYNTHESIS METHODS ON THE XJTU-UP DATABASE

Missing Rate	10%		20%		30%	
Dataset	HWN	IPN	HWN	IPN	HWN	IPN
Pix2Pix [26]	90.10	90.10	80.20	79.16	70.31	70.31
CycleGAN [30]	88.02	88.02	82.81	81.77	78.64	77.60
RAU-net [40]	92.19	89.06	79.69	79.69	75.00	76.04
AttentionGAN [41]	92.19	93.23	86.98	88.54	80.73	82.29
Ours	94.79	94.79	88.54	86.98	86.98	87.50

background. The attention-guided generators in AttentionGAN can produce attention masks and then fuse the generated output with the attention masks to obtain high-quality target images.

Note that all these methods used ResNet-50 as the classifier using similar hyperparameters with a slight difference in each model to achieve the best performance. The experimental settings, such as the training and testing data split and the incomplete data simulation method, were consistent.

The classification results of the methods on the MPD database and XJTU-UP database are shown in Tables III and IV, respectively. We can observe the significant improvement in our method in these tables, and the recognition accuracies of our method showed a relatively slight decline as the missing rate increased. Specifically, all the test samples were correctly classified by our method on the MPD dataset with a missing rate of 10%. This dataset has superior image quality compared to the XJTU-UP dataset, with clearer palm lines and more detailed texture information as demonstrated in Figure 5. Additionally, the low rate of missing data allows for the use of a larger number of real palmprint images during model training, which helps to improve the generalization ability of the model. As can be seen from Tables I to IV, the accuracies of the methods tend to decrease as the rate of missing data increases. On the MPD database, our method outperformed all the other image-to-image translation approaches. On the XJTU-UP database, our method achieved the same classification rate as AttentionGAN when 10% of the subjects were missing. Under the missing rate of 30%, the best accuracies were 86.98% for HWN and 87.50% for IPN by our method. Fig. 8 show some images generated by our method and several state-of-the-art methods. As shown in Fig. 8, the palmprint synthesized by our method best approximated the corresponding image style among all the methods. Specifically, compared with Pix2Pix and RAU-net, the images generated by our method are much clear in palm lines. Compared with CycleGAN and AttentionGAN, our method preserves more

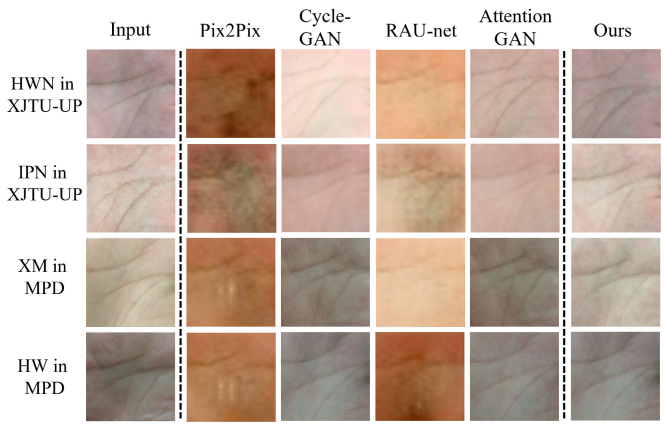


Fig. 8. The synthesized images by different methods. From left to right: input, pix2pix, CycleGAN, RAU-net, AttentionGAN, and ours.

TABLE V
THE ABLATION STUDY RESULT (TOP-1 ACCURACY (%)) ON THE XJTU-UP DATABASE

CycleGAN	Self-attention	SPL	HWN	IPN
✓			88.02	88.02
✓	✓		89.06	89.58
✓		✓	92.19	92.70
✓	✓	✓	94.79	94.79

texture details. The examples in Fig. 8 shows that our method not only can achieve the transfer of image style across different palm imaging devices, but also maintains more discriminative information of the palm images.

E. Ablation Experiments

We performed an ablation study on the XJTU-UP database with a missing rate of 10% to show the effect of the self-attention layer and SPL strategy. The results are listed in Table V. When self-attention modules were adopted, the performance increased by over 2% in both datasets. When further combined with the SPL strategy, our method achieved the highest accuracy.

F. Parameter Tuning and Sensitivity Analysis

1) Analysis of Scalars of Consistency Loss and Identity Loss: In this experiment, we demonstrated the performance of our method under different parameter combinations. We tried several different combinations of λ and μ on the MPD database with a missing rate of 30%. As shown in TableVI, the classification results were relatively stable, when the two parameters were slightly tuned. In particular, the proposed method achieved better recognition performance when λ both and μ are 10.

2) Analysis of Quality Score Component: To assess the influence of adversarial loss, consistency loss, and identity loss on recognition results, we performed several experiments on different quality score schemes on the MPD database with a missing rate of 30%. As shown in Table VII, the accuracies achieved by combining consistency loss and identity loss were higher than those of the methods with a single-loss quality score. The recognition rate dropped when the adversarial loss was introduced as the quality score.

TABLE VI
THE TOP-1 ACCURACY (%) OF OUR METHOD FOR DIFFERENT COMBINATIONS OF PARAMETERS λ AND μ ON MPD WITH A MISSING RATE OF 30%

	Datasets	$\lambda = 5$	$\lambda = 10$	$\lambda = 15$
$\mu = 5$	XM	98.625	98.625	98.875
	HW	99.125	98.375	98.625
$\mu = 10$	XM	98.375	99.125	98.500
	HW	98.750	99.000	98.625
$\mu = 15$	XM	99.125	99.25.	99.000
	HW	98.750	99.00	99.250

TABLE VII
THE TOP-1 ACCURACY (%) OF RESNET-50 WITH DIFFERENT DATA SYNTHESIS METHODS ON MPD DATABASE

Score	XM	HW
Adversarial loss	97.75	97.625
Consistency loss	98.125	98.50
Identity loss	98.25	98.625
Adversarial loss + Consistency loss + Identity loss	98.50	98.375
Ours	99.25	99.00

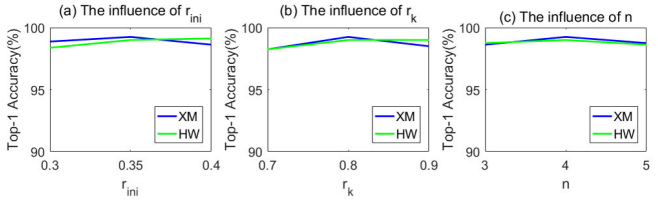


Fig. 9. The influence of r_{ini}, r_k and n on the MPD database with 30% missing data, where r_k=0.8 and n=4 in (a), r_{ini}=0.35 and n=4 in (b), and r_{ini}=0.35 and r_k=0.8 in (c).

3) Analysis of Hyperparameters in the SPL Strategy: We investigated the influence of r_{ini}, r_k, n on the MPD database with 30% missing data. It can be observed in Fig. 9 that the optimal settings of r_{ini}, r_k, n are 0.35, 0.8, and 4, respectively. The recognition results showed relatively slight variations when r_{ini} and n were slightly changed and showed a decrease of 0.75% when r_k was tuned.

V. CONCLUSION

In this paper, we proposed a self-paced CycleGAN with self-attention modules to solve multi-domain contactless palmprint recognition with missing data issues. We developed a novel deep network with CycleGAN to synthesize missing training data by learning the mapping between different domains, and self-attention modules were introduced as a supplement for the convolutional generator to model the long-range dependencies. Simultaneously, the SPL strategy, motivated by a human learning mechanism, was adopted in our model to alleviate the influence of noise in contactless palmprint images and avoid training instability in the generative adversarial module. The experimental results showed that our methods can effectively relieve the decline in accuracies caused by missing inputs, and our method outperformed several state-of-the-art approaches in contactless palmprint recognition.

However, our method has several limitations in practice. We assumed that there were only two known devices for image acquisition. Many kinds of smartphones can be used for palmprint image collection. In addition, palmprints from unknown devices can be used in real-world applications. To explore this issue further, in future work, we can cluster the contactless palmprint images from unknown devices through unsupervised methods and then investigate the transfer learning-based recognition model established in the data domains from known devices to unknown devices.

REFERENCES

[1] A. Genovese, V. Piuri, and F. Scotti, *Touchless Palmprint Recognition Systems*, vol. 60. Cham, Switzerland: Springer, 2014.

[2] Y. Zhang, L. Zhang, X. Liu, S. Zhao, Y. Shen, and Y. Yang, "Pay by showing your palm: A study of palmprint verification on mobile platforms," in *Proc. IEEE Int. Conf. Multimedia Expo (ICME)*, Jul. 2019, pp. 862–867.

[3] Y. Zhang, L. Zhang, R. Zhang, S. Li, J. Li, and F. Huang, "Towards palmprint verification on smartphones," 2020, *arXiv:2003.13266*.

[4] 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.

[5] W. Zuo, Z. Lin, Z. Guo, and D. Zhang, "The multiscale competitive code via sparse representation for palmprint verification," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, Jun. 2010, pp. 2265–2272.

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

[7] L. Fei, G. Lu, W. Jia, S. Teng, and D. Zhang, "Feature extraction methods for palmprint recognition: A survey and evaluation," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 49, no. 2, pp. 346–363, Feb. 2019.

[8] X. Wu, Q. Zhao, and W. Bu, "A SIFT-based contactless palmprint verification approach using iterative RANSAC and local palmprint descriptors," *Pattern Recognit.*, vol. 47, no. 10, pp. 3314–3326, Oct. 2014.

[9] 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.

[10] A. Kumar, "Toward more accurate matching of contactless palmprint images under less constrained environments," *IEEE Trans. Inf. Forensics Security*, vol. 14, no. 1, pp. 34–47, Jan. 2019.

[11] Q. Zhu, N. Xu, Z. Zhang, D. Guan, R. Wang, and D. Zhang, "Cross-spectral palmprint recognition with low-rank canonical correlation analysis," *Multimedia Tools Appl.*, vol. 79, nos. 45–46, pp. 33771–33792, Dec. 2020.

[12] X. Xu, Z. Guo, C. Song, and Y. Li, "Multispectral palmprint recognition using a quaternion matrix," *Sensors*, vol. 12, no. 4, pp. 4633–4647, Apr. 2012.

[13] A. Vaswani et al., "Attention is all you need," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 30, 2017, pp. 1–11.

[14] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 25, 2012, pp. 1–9.

[15] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," 2014, *arXiv:1409.1556*.

[16] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 2818–2826.

[17] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 770–778.

[18] L. Dian and S. Dongmei, "Contactless palmprint recognition based on convolutional neural network," in *Proc. IEEE 13th Int. Conf. Signal Process. (ICSP)*, Nov. 2016, pp. 1363–1367.

[19] 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.

[20] A. Genovese, V. Piuri, F. Scotti, and S. Vishwakarma, "Touchless palmprint and finger texture recognition: A deep learning fusion approach," in *Proc. IEEE Int. Conf. Comput. Intell. Virtual Environments Meas. Syst. Appl. (CIVEMSA)*, Jun. 2019, pp. 1–6.

[21] H. Shao, D. Zhong, and X. Du, "Cross-domain palmprint recognition based on transfer convolutional autoencoder," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Sep. 2019, pp. 1153–1157.

[22] H. Shao and D. Zhong, "Towards cross-dataset palmprint recognition via joint pixel and feature alignment," *IEEE Trans. Image Process.*, vol. 30, pp. 3764–3777, 2021.

[23] A. Gonzalez-Garcia, J. Van De Weijer, and Y. Bengio, "Image-to-image translation for cross-domain disentanglement," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 31, 2018, pp. 1–12.

[24] T. Kim, M. Cha, H. Kim, J. K. Lee, and J. Kim, "Learning to discover cross-domain relations with generative adversarial networks," in *Proc. Int. Conf. Mach. Learn.*, 2017, pp. 1857–1865.

[25] H. Tang, W. Wang, D. Xu, Y. Yan, and N. Sebe, "GestureGAN for hand gesture-to-gesture translation in the wild," in *Proc. 26th ACM Int. Conf. Multimedia*, Oct. 2018, pp. 774–782.

[26] P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros, "Image-to-image translation with conditional adversarial networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2017, pp. 1125–1134.

[27] S. Chen, S. Chen, Z. Guo, and Y. Zuo, "Low-resolution palmprint image denoising by generative adversarial networks," *Neurocomputing*, vol. 358, pp. 275–284, Sep. 2019.

[28] W. Yifan, F. Serena, S. Wu, C. Öztireli, and O. Sorkine-Hornung, "Differentiable surface splatting for point-based geometry processing," *ACM Trans. Graph.*, vol. 38, no. 6, pp. 1–14, Dec. 2019.

[29] P. Salehi and A. Chalechale, "Pix2Pix-based stain-to-stain translation: A solution for robust stain normalization in histopathology images analysis," in *Proc. Int. Conf. Mach. Vis. Image Process. (MVIP)*, Feb. 2020, pp. 1–7.

[30] J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros, "Unpaired image-to-image translation using cycle-consistent adversarial networks," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Oct. 2017, pp. 2242–2251.

[31] M. Kumar, B. Packer, and D. Koller, "Self-paced learning for latent variable models," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 23, 2010, pp. 1–9.

[32] L. Jiang, D. Meng, S.-I. Yu, Z. Lan, S. Shan, and A. Hauptmann, "Self-paced learning with diversity," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 27, 2014, pp. 1–10.

[33] H. Li, M. Gong, D. Meng, and Q. Miao, "Multi-objective self-paced learning," in *Proc. 13th AAAI Conf. Artif. Intell.*, 2016, pp. 1–7.

[34] F. Pahde, O. Ostapenko, P. J. Hnichen, T. Klein, and M. Nabi, "Self-paced adversarial training for multimodal few-shot learning," in *Proc. IEEE Winter Conf. Appl. Comput. Vis. (WACV)*, Jan. 2019, pp. 218–226.

[35] J. S. Supancic III and D. Ramanan, "Self-paced learning for long-term tracking," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2013, pp. 2379–2386.

[36] D. Zhang, D. Meng, C. Li, L. Jiang, Q. Zhao, and J. Han, "A self-paced multiple-instance learning framework for co-saliency detection," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Dec. 2015, pp. 594–602.

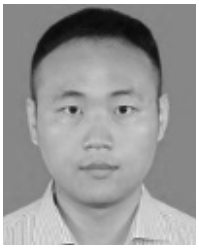
[37] H. Zhang, I. Goodfellow, D. Metaxas, and A. Odena, "Self-attention generative adversarial networks," in *Proc. Int. Conf. Mach. Learn.*, 2019, pp. 7354–7363.

[38] Y. Ge, F. Zhu, D. Chen, and R. Zhao, "Self-paced contrastive learning with hybrid memory for domain adaptive object re-ID," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 33, 2020, pp. 11309–11321.

[39] H. Shao, D. Zhong, and X. Du, "Efficient deep palmprint recognition via distilled hashing coding," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW)*, Jun. 2019, pp. 714–723.

[40] J. Guo, W. Zeng, S. Yu, and J. Xiao, "RAU-Net: U-net model based on residual and attention for kidney and kidney tumor segmentation," in *Proc. IEEE Int. Conf. Consum. Electron. Comput. Eng. (ICCECE)*, Jan. 2021, pp. 353–356.

[41] H. Tang, H. Liu, D. Xu, P. H. S. Torr, and N. Sebe, "AttentionGAN: Unpaired image-to-image translation using attention-guided generative adversarial networks," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 34, no. 4, pp. 1972–1987, Apr. 2023.

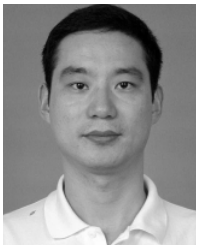


Qi Zhu received the B.S., M.S., and Ph.D. degrees from the Harbin Institute of Technology in 2007, 2010, and 2014, respectively. He is currently a Professor with the College of Computer Science and Technology, Nanjing University of Aeronautics and Astronautics. He has published over 40 scientific articles in refereed international journals and conference proceedings. His research interests include pattern recognition, feature extraction, and medical image analysis.

Guangnan Xin received the B.E. degree in digital media technology from Shandong University in July 2014. She is currently pursuing the M.S. degree with the College of Computer Science and Technology, Nanjing University of Aeronautics and Astronautics. Her current research interests include feature extraction and pattern recognition.

Lunke Fei (Senior Member, IEEE) received the Ph.D. degree in computer science and technology from the Harbin Institute of Technology, Harbin, China, in 2016. Since 2017, he has been with the School of Computer Science and Technology, Guangdong University of Technology, Guangzhou, China. His research interests include pattern recognition, biometrics, image processing, and machine learning.

Dong Liang received the B.S. degree in telecommunication engineering and the M.S. degree in circuits and systems from Lanzhou University, China, in 2008 and 2011, respectively, and the Ph.D. degree from the Graduate School of IST, Hokkaido University, Japan, in 2015. He is currently an Associate Professor with the College of Computer Science and Technology, Nanjing University of Aeronautics and Astronautics. He has published several research papers, including *Pattern Recognition*, IEEE TRANSACTIONS ON IMAGE PROCESSING, IEEE TRANSACTIONS ON NEURAL NETWORKS AND LEARNING SYSTEMS, IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING, IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY, and AAAI. His research interests include model communication in pattern recognition and image processing. He was received the Excellence Research Award from Hokkaido University in 2013.



Zheng Zhang (Senior Member, IEEE) received the M.S. degree in computer science and the Ph.D. degree in computer applied technology from the Harbin Institute of Technology, Shenzhen, China, in 2014 and 2018, respectively. He was a Post-Doctoral Research Fellow with The University of Queensland, Australia. He is currently an Associate Professor with the Harbin Institute of Technology. He has published over 50 technical papers at prestigious international journals and conferences, including the IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, IEEE TRANSACTIONS ON NEURAL NETWORKS AND LEARNING SYSTEMS, IEEE TRANSACTIONS ON IMAGE PROCESSING, IEEE TRANSACTIONS ON CYBERNETICS, IEEE CVPR, ECCV, AAAI, IJCAI, SIGIR, and ACMM.

Daoqiang Zhang (Senior Member, IEEE) received the B.Sc. and Ph.D. degrees in computer science from the Nanjing University of Aeronautics and Astronautics, Nanjing, China, in 1999 and 2004, respectively. He is currently a Professor with the Department of Computer Science and Engineering, Nanjing University of Aeronautics and Astronautics. His current research interests include machine learning, pattern recognition, and biomedical image analysis. In these areas, he has authored or coauthored more than 100 technical papers in refereed international journals and conference proceedings.

David Zhang (Life Fellow, IEEE) received the Graduate degree in computer science from Peking University, Beijing, China, the M.Sc. degree in computer science and the Ph.D. degree from the Harbin Institute of Technology (HIT), Harbin, China, in 1982 and 1985, respectively, and the Ph.D. degree in electrical and computer engineering from the University of Waterloo, Waterloo, ON, Canada, in 1994. From 1986 to 1988, he was a Post-Doctoral Fellow with Tsinghua University, Beijing, and then an Associate Professor with the Academia Sinca, Beijing. He has been a Chair Professor with The Hong Kong Polytechnic University, where he has been the Founding Director of the Biometrics Research Centre (UGC/CRC) supported by the Hong Kong SAR Government, since 1998. He is currently the Presidential Chair Professor with The Chinese University of Hong Kong (Shenzhen). He is also a Visiting Chair Professor with Tsinghua University and an Adjunct Professor with Peking University, Shanghai Jiao Tong University, Shanghai, China, HIT, and the University of Waterloo. He has authored more than ten books and 200 journal articles. He is also a Croucher Senior Research Fellow, a Distinguished Speaker of the IEEE Computer Society, and a fellow of International Association for Pattern Recognition (IAPR). He is an Organizer of the International Conference on Biometrics Authentication. He was a Book Editor of the *International Series on Biometrics* (Springer) and an Associate Editor of more than ten international journals, including the IEEE TRANSACTIONS AND PATTERN RECOGNITION. He is the Founder and the Editor-in-Chief of the *International Journal of Image and Graphics*.