RPG-Palm: Realistic Pseudo-data Generation for Palmprint Recognition

Lei Shen*
Youtu Lab, Tencent
Shanghai, China
shenlei1996@gmail.com

Jianlong Jin*
Hefei University of Technology
Hefei, China
2022111048@mail.hfut.edu.cn

Ruixin Zhang
Youtu Lab, Tencent
Shanghai, China
ruixinzhang@tencent.com

Huaen Li
Hefei University of Technology
Hefei, China
huaenli@mail.hfut.edu.cn

Kai Zhao
UCLA
California, USA
kz@kaizhao.net

Yingyi Zhang
Youtu Lab, Tencent
Shanghai, China
sherlyzhang@tencent.com

Jingyun Zhang
WeChat Pay Lab33, Tencent
Shenzhen, China
naskyzhang@tencent.com

Shouhong Ding†
Youtu Lab, Tencent
Shanghai, China
ericshding@tencent.com

Yang Zhao†
Hefei University of Technology
Hefei, China
YZhao@hfut.edu.cn

Wei Jia†
Hefei University of Technology
Hefei, China
jiawei@hfut.edu.cn

Abstract

Palmprint recently shows great potential in recognition applications as it is a privacy-friendly and stable biometric. However, the lack of large-scale public palmprint datasets limits further research and development of palmprint recognition. In this paper, we propose a novel realistic pseudo-palmprint generation (RPG) model to synthesize palmprints with massive identities. We first introduce a conditional modulation generator to improve the intra-class diversity. Then an identity-aware loss is proposed to ensure identity consistency against unpaired training. We further improve the Bézier palm creases generation strategy to guarantee identity independence. Extensive experimental results demonstrate that synthetic pretraining significantly boosts the recognition model performance. For example, our model improves the state-of-the-art BézierPalm by more than 5% *and* 14% *in terms of TAR@FAR=1e-6 under the* $1:1$ *and* $1:3$ *Open-set protocol. When accessing only* 10% *of the real training data, our method still outperforms ArcFace with* 100% *real training data, indicating that we are closer to real-data-free palmprint recognition.*

*Equal contribution.
†Corresponding authors.

1. Introduction

Palmprint is an excellent biometric in terms of privacy, willingness, convenience, and security. Since the palmprint locates on the inner side of the palm, obtaining the palmprint without one's permission is nearly impossible, which is more privacy-friendliness over the face. Compared to iris and fingerprints, palmprint has fewer usage restrictions, which makes palmprint more user-friendly. Furthermore, Palmprints and palm veins can be collected simultaneously to form a highly secure dual-modal system.

Due to these advantages, big companies such as Amazon [1] and Tencent [69] begin to apply palmprint recognition in their payment services. However, for the same reason, palmprints are rare in public and expensive to collect. To our best knowledge, publicly available datasets [2, 21, 29, 30, 36, 41, 52, 57, 68] only contain thousands of identities and tens of thousands of images in total. Meanwhile, face recognition has several million-level publicly available datasets [4, 9, 27, 38, 44, 62] that contain tens to hundreds of thousands of identities. The lack of large-scale public dataset seriously inhibits the research on palmprint recognition.

To solve this problem, one way is to collect a large-scale palmprint dataset. However, this way is very time-

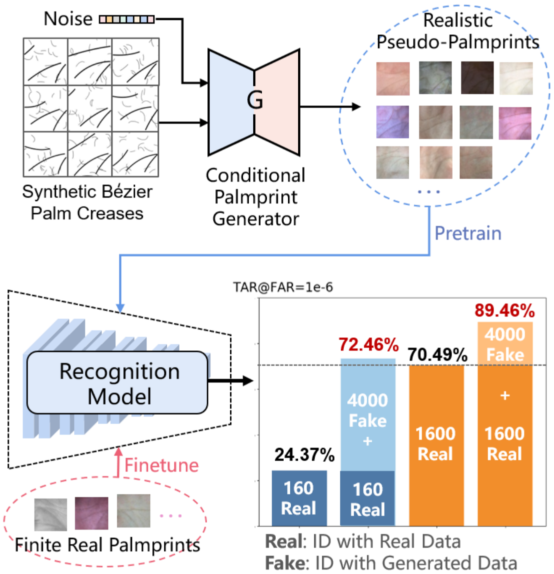

Figure 1: Synthetic Bézier palm creases are used as the identity control condition to guide the generator to produce diverse synthetic data. Recognition models are pretrained with pseudo-data and then finetuned on real data, achieving significant performance boosts. Besides, even if only accessing 10% of the real training data for generation and recognition model, our method still gets comparable performance to 100% of real data direct training.

consuming and expensive. In addition, collecting biometric data causes more and more privacy and ethical concerns [7] and is strictly restricted by legislation in many countries. Another way is to augment the training dataset with synthetic palmprints. To better benefit palmprint recognition applications, the synthetic method should support generating massive identities with inter-class and intra-class diversity, and the gap between synthetic palmprints and real palmprints should be small.

Recently, BézierPalm [69] synthesizes fake palmprint images by generating palmprint creases with several parameterized Bézier curves for recognition model pretraining. BézierPalm first shows the ability to output massive new identities without using real palmprints and significantly improves the performance of the palmprint recognition models. However, BézierPalm has some problems unsolved. Firstly, synthetic palmprints have a large gap to real ones, resulting in non-neglectable finetune data requirements in real-world applications. Secondly, the curve parameter difference of each identity cannot ensure inter-class independence when generating massive data. Thirdly, the intra-class diversity only contains small curve deformation in BézierPalm, while real palmprints have diverse textures, lighting and et al.

In this paper, we propose a realistic pseudo-palmprint generation (RPG) model. As shown in Fig.1, the RPG model takes synthetic Bézier palm creases as an identity (ID) condition and outputs realistic pseudo-palmprints with a bidirectional mapping from the Gaussian noise domain to the palmprints domain. Since there is no correspondence between synthetic Bézier palm creases and real palmprints, the generation model can only use unpaired data and will lose intra-class identity consistency easily during training. To solve this problem, we introduce an identity-aware loss that restrains the identity consistency between palmprints generated from the same Bézier palm creases. In addition, we design a conditional modulation generator to generate diversified intra-class textures and lighting conditions using a latent control vector encoded from random noises. To further reduce the distribution gap between synthetic palm creases and real-world palmprints, we refined the synthetic palmprint creases generation strategy with a more reasonable parameter design and identity independence check.

The contributions of this paper are as follows:

- We propose a realistic pseudo-palmprint generation (RPG) model with a conditional modulation generator to improve the intra-class diversity and an ID-aware loss to help the RPG model ensure identity consistency under unpaired training.
- We improve the Bézier palm creases synthetic method to get more reasonable palm creases and independent identities.
- Extensive experimental results on 13 public datasets demonstrate that recognition models pretrained with our synthetic pseudo-palmprints achieve state-of-the-art recognition accuracy.
- With our RPG pretraining, even if accessing only 10% of the real data, the recognition model performance still outperforms 100% real data direct training. Showing we are closer to real-data-free palmprint recognition.

2. Related Work

2.1. Palmprint Recognition Methods

Palmprint recognition methods can be divided into two categories: traditional-based and deep-learning based [18]. Traditional methods extract various kinds of local or global features to make different palmprints more discriminative. Local-based methods [19, 28, 34, 40, 43, 57, 70] manually design effective local features for recognition. Global-based methods [3, 20, 32, 42, 51, 59] extract low-dimensional features from the whole image to distinguish different palmprints. Deep Learning based methods [16, 25, 35, 58] train modified neural networks to extract features with classification or pair-wise loss [14, 55, 71]. However, the lack of large-scale public palmprint datasets limits the potential of existing palmprint recognition methods.

2.2. Generation Model for Image Transformation

With the development of image generation models, such as generative adversarial network (GAN)-based models [26, 37] and diffusion-based models [47, 48], image-to-image generation/translation methods have achieved impressive performance [11, 12, 33]. However, many typical models, e.g., conditional generation model pix2pixHD [61], multi-domain transformation model StarGAN [12], dual-domain mapping model BicycleGAN [74] and recent conditional diffusion model [50, 60], rely on paired training data, which is unavailable in many applications. Therefore, some unpaired image-to-image translation models have been proposed [39, 54, 64, 72], which usually adopt cycle consistency loss to train the models without paired data. However, these models are not designed for recognition tasks, and thus ignore the ID preservation in the generation process. Overall, there is still a lack of research on generating diversified and realistic palmprints with ID consistency from limited unpaired samples.

2.3. Data generation for Recognition Tasks

In order to improve the recognition performance, data generation can be used to expand the depth (diversity of each identity) and width (total number of identities) of training data. For example, in face recognition field, several facial image generation methods [15, 24, 45, 46, 65, 22, 23] have been proposed to generate facial samples based on the priors of facial attributes, facial structures and 3D faces [6]. Also, in field of fingerprint recognition, some methods based on hand-crafted or learning-based approaches [5, 17, 63] have been proposed to generate high-fidelity fingerprint images.

For palm recognition tasks, BézierPalm [69] use Bézier curves to synthesize fake palm creases by changing the Bézier control parameters. PalmGAN [53] improves the cross-domain recognition performances by using CycleGAN to transfer the styles between normal and multispectral palmprints. However, PalmGAN doesn't create virtual identities, and the BézierPalm suffers from the domain gap between palmprint images and geometry curves. By introducing the new generation model and ID-aware loss, our method can expand the inter-class diversity and intra-class diversity of palmprint dataset simultaneously.

3. Method

3.1. Overall Framework

Fig.2 illustrates the whole framework of the proposed realistic pseudo-palmprint generation (RPG) model, which includes a training stage and a forward palmprint generation stage. In the training phase, palmprint B is first mapped from the palmprint image domain to the latent vector $Q(z|B)$ in the Gaussian noise domain through the encoder E. Then, the generator G utilizes the unpaired palm creases A as condition, and remaps the encoded noise vector $Q(z|B)$ back to the palmprint image domain. In order to increase the diversity and randomness of the generated palmprints, a conditional modulation structure is designed for G, which uses the input noise vector to control the modulation of intermediate features.

In order to constrain the ID consistency of the generated pseudo palmprints, an ID-aware loss is presented to enforce the generator to maintain the ID information of input palm creases. As shown in Fig.2 (a), a siamese generator G is used to produce another palmprint B'' with the same palm creases A. Then, a pretrained palmprint recognition discriminator is used to measure the ID consistency between B' and B''.

In the forward pseudo-palmprint generation stage, synthetic Bézier creases are input into the generator G to obtain corresponding palmprint images for virtual identities. Motivated by BézierPalm [69], we improve the random Bézier curves generation strategy to obtain a more reasonable layout of principal lines and wrinkle lines. Then, a classic palmprint recognition method RLOC [34] is applied to ensure the inter-class difference.

In the following, we will introduce the details of generator G, encoder E, ID-aware loss, and the improved Bézier palm creases synthesis strategy.

3.2. Conditional Modulation Palmprint Generator

The proposed generator G takes Gaussian noise vector as input, palm creases image A as condition, and reproduces the pseudo-palmprint image B'. As a typical image generation task, the common UNet [49] architecture is adopted. The detailed structure of G is illustrated in Fig.3 (a). To generate diversified results, a conditional adaptive instance normalization module (CAdaIN) is introduced to modulate the generated details in each down-block and up-block.

In the CAdaIN module, the input noise vector $N(z)$ is first encoded into a latent control vector $w(z)$ through two fully-connected (FC) layers. Note that the parameters of these two FC layers are shared for the whole generator G, so that the generated style can be consistent by means of the same $w(z)$. Then, two other FC layers are used to modulate the mean and variance of intermediate feature maps

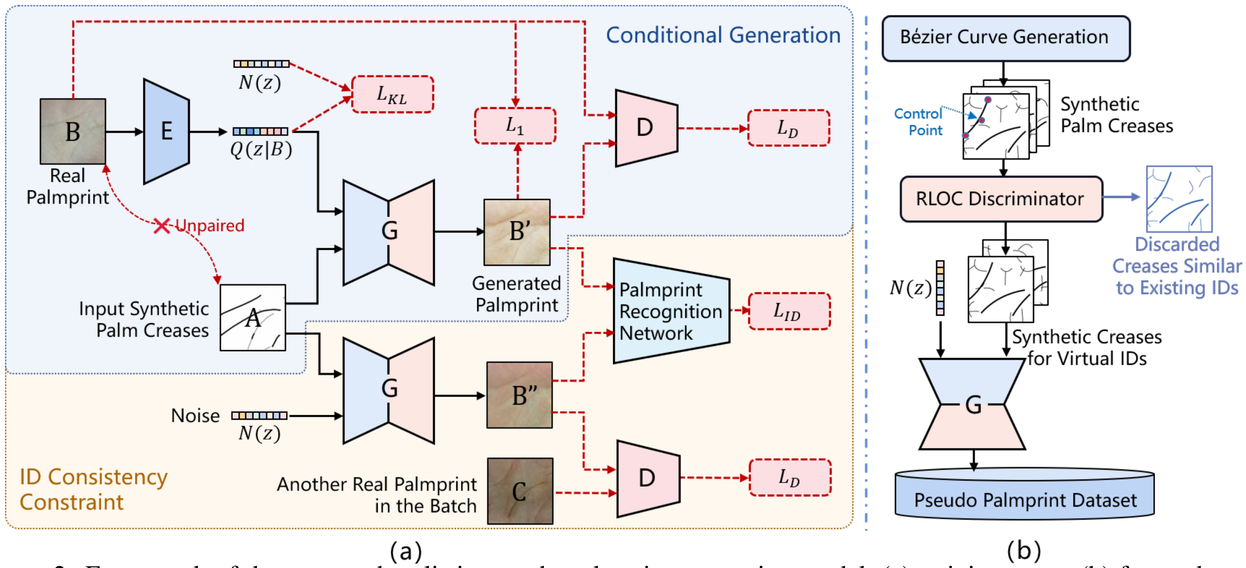

Figure 2: Framework of the proposed realistic pseudo-palmprint generation model, (a) training stage, (b) forward pseudo-palmprint generation stage.

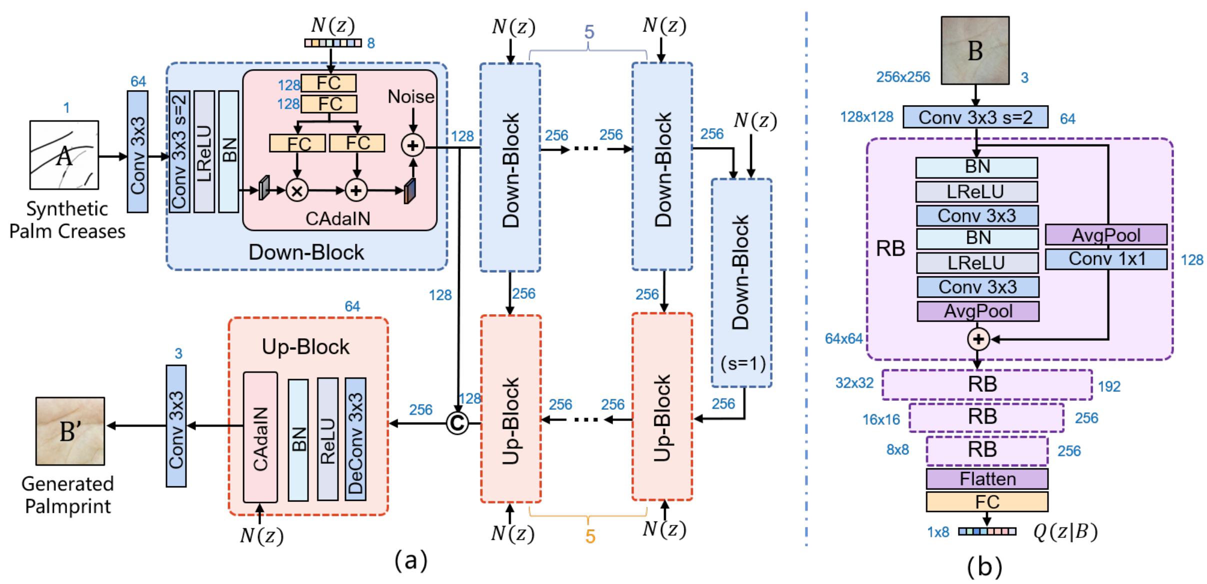

Figure 3: The network structure of the proposed method, (a) structure of generator G, (b) structure of encoder E.

respectively, which can be calculated as,

$$\boldsymbol{X}_o = f_{c1}(w(z))\boldsymbol{X}_i + f_{c2}(w(z)) + \boldsymbol{n}_0, \qquad (1)$$

where $\boldsymbol{X}_i$ and $\boldsymbol{X}_o$ denote the input and output features of CAdaIN, and f_{c1}, f_{c2} represent two FC layers. By using a latent control vector encoded from random noise, the CAdaIN can modulate the features to produce different distributions, resulting in diversified palmprint images. In addition, in order to further improve the diversity, random noise $\boldsymbol{n}_0$ with the same spatial resolution as the feature map is added to the modulated features to inject more randomness.

For producing realistic results, the loss $\mathcal{L}_1$ and adversarial loss $\mathcal{L}_D$ are used to restrain the learning of G, as follows,

$$\mathcal{L}_G = \lambda_1 \mathcal{L}_1(B, B') + \lambda_2 \mathcal{L}_D(B, B'), \qquad (2)$$

where λ_1 and λ_2 are weights. The $\mathcal{L}_1$ is a commonly used pixel-wise loss, which can ensure the numerical similarity between the generated image and real palmprint. But note that the palm creases image A is not matched with palmprint B, so that too strong pixel-wise constraint may cause wrong overfitting to the details in B. Hence, adversarial loss $\mathcal{L}_D$ is also used to relax the constraint and restrain the semantic similarity between B and B'. For the discriminator in $\mathcal{L}_D$, we adopt the PatchGAN [13], which determines patch-wise authenticity by mapping the image to a 70×70 grid.

3.3. Palmprint Encoder

The encoder E is used to map the palmprint image to the Gaussian noise domain. Its structure is shown in Fig.3 (b), which is a straightforward ResNet structure. The original size of the feature map is 256×256, and it gradually decreases to 16×16 through several residual blocks (RB). Then, a FC layer is used to estimate the target mean μ_Q and variance σ_Q^2 from flattened features. Finally, the noise vector $Q(z|B)$ is sampled from Gaussian space $\mathcal{N}(\mu_Q, \sigma_Q^2)$ via reparameterization trick [74].

In the training stage, we constrain the KL divergence between $Q(z|B) \sim \mathcal{N}(\mu_Q, \sigma_Q^2)$ and noise vector $N(z) \sim \mathcal{N}(0,1)$ sampled from Gaussian space, so that the target domain of E can keep approximating to the standard normal distribution. This loss can be computed as,

$$\mathcal{L}_{KL} = -\frac{1}{2}(1 + \log \sigma_Q^2 - \sigma_Q^2 - \mu_Q^2). \qquad (3)$$

In this paper, E and G are combined to learn a domain-to-domain mapping from unpaired training data, instead of using a fully supervised manner. Firstly, this structure can avoid the dependence on manual labeling of paired data. Secondly, conditional bidirectional domain-to-domain mapping can reproduce random and realistic images and keep the information of palm creases condition as well.

3.4. ID-aware Loss

For training recognition models, the generated palmprints not only need to be diversified, but also have to preserve ID information. For a synthetic palm creases image A, the generated palmprints should have the same ID information. Therefore, an ID-aware loss $\mathcal{L}_{ID}$ is added to restrain the generator. As shown in Fig.2 (a), a siamese generator G is used to produce another palmprint image B'' with the same creases A and random noise vector $N(z)$. Then the $\mathcal{L}_{ID}$ restricts the ID consistency of two generated results of B' and B'', as follows,

$$\mathcal{L}_{ID} = 1 - \frac{D_{MB}(B') \cdot D_{MB}(B'')}{||D_{MB}(B')|| \times ||D_{MB}(B'')||}, \qquad (4)$$

where "$\cdot$" denotes the vector dot product operation, D_{MB} is a pretrained palmprint recognition model using MobileFaceNet [10] and extracts the 512×1 feature of B' and B''. That is, $\mathcal{L}_{ID}$ calculates cosine similarity between the extracted features of two generated images for the same ID. Owing to the ID-aware loss, increasing the randomness and diversity of the generator will not destroy the intra-class ID consistency.

The total loss of entire model, as follows,

$$\mathcal{L}_{total} = \lambda_D \mathcal{L}_D + \lambda_1 \mathcal{L}_1 + \lambda_{KL} \mathcal{L}_{KL} + \lambda_{ID} \mathcal{L}_{ID}, \qquad (5)$$

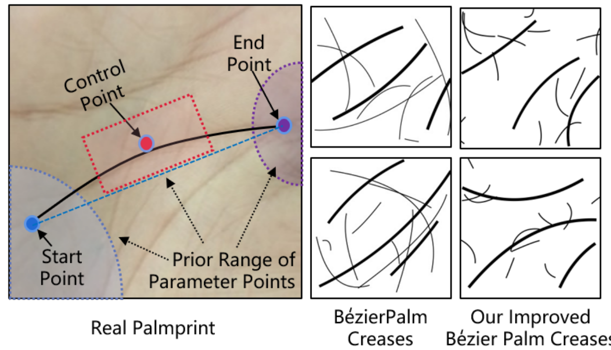

Figure 4: Examples of some synthetic palm creases in BézierPalm [69] and our method.

where $\mathcal{L}_D$ denotes GAN loss, $\mathcal{L}_1$ denotes absolute error loss, $\mathcal{L}_{KL}$ denotes KL divergence loss, and $\mathcal{L}_{ID}$ denotes ID-aware loss. We will add the total loss in final version.

3.5. Improved Bézier Palm Creases Synthesis

In the forward stage, only the generator G is used to produce palmprint images from synthetic palm creases, as shown in Fig.2 (b). Motivated by BézierPalm [69], we also adopt the two-level Bézier curves to synthesize palm creases image that contains three principal lines and random wrinkle lines. The parametric control points of the Bézier curve can be adjusted randomly within a preset range to obtain a large number of fake palm creases for virtual ID.

However, we have observed that randomly adjusted Bézier curves may lead to some unreasonable results. As shown in Fig.4, the layout of some randomly generated creases is quite different from that of the real palmprints.

As a result, this paper improves the Bézier curves synthesis strategy based on real palmprint prior from the following three aspects. Firstly, we observe and adjust the rough range of the parameter points (i.e., start point, control point, and end point) of three principal lines according to real palmprints. So that the layout of synthetic principal lines becomes more similar to that of real palms. Secondly, we adjust the synthesis rules to generate more wrinkle lines with moderate length and more uniform distribution. Thirdly, a similarity constraint based on RLOC [34] is added, which makes the principal lines of different IDs sufficiently distinguishable. RLOC is a classic palmprint recognition method that can measure the similarity of two creases images. In order to avoid very similar principal lines being generated for different IDs, we filter out some creases images which exceed the RLOC inter-class similarity threshold. We ablate the threshold of RLOC with 0.1, 0.5, 0.9, 0.95 and experimentally set it as 0.9. More details of improved Bézier palm creases can be found in supplementary materials.

Finally, the improved random Bézier curves are input into the generator G to obtain corresponding palmprint images. This generation process can be repeated until a large-scale realistic pseudo-palmprint dataset is established.

4. Experimental Settings

We adopt the same experimental datasets and Open-set evalutation protocol following BézierPalm [69]. The identities of the training and test set are isolated. TAR(True Acceptance Rate) @FAR(False Accept Rate) is used to evaluate model performance. Details about Open-set and evaluation protocol can be found in supplementary materials.

4.1. Datasets

We adopt 13 public datasets in our experiments as in BézierPalm [69], including $3,268$ IDs and $59,162$ palmprint images. Detailed descriptions of these datasets are shown in Tab.1. We follow the detect-then-crop protocol in [67] to extract the Region Of Interests(ROIs).

Name	#IDs	#Images	Device
MPD [68]	400	16,000	Phone
TCD [68]	600	12,000	Contactless
IITD [41]	460	2,601	Contactless
CASIA [57]	620	5,502	Contactless
CASIA-MS [29]	200	7,200	Contactless
COEP [2]	167	1,344	Digital camera
MOHI [30]	200	3,000	Phone
WEHI [30]	200	3,000	Web cam
GPDS [21]	100	2,000	Web cam
XJTU_UP [52]	200	30,000	Phone
XJTU_A [52]	114	1,130	CMOS camera
PolyU-MS [66]	500	24,000	Contactless
PolyU(2D+3D) [36]	400	8,000	Web cam

Table 1: Details of the 13 public palmprint datasets.

4.2. Implementation Details

Generation Model Training. We generate 4000 identities and 100 samples for each identity by default following BézierPalm [69]. For training the generation model, the weights of $\mathcal{L}_1$, $\mathcal{L}_D$ and $\mathcal{L}_{KL}$ are set as 10.0 1.0 and 0.01 according to [74]. We ablate the weight of $\mathcal{L}_{ID}$ with 1.0, 5.0, 10.0 and experimentally set it as 5.0. The resolution of input and output images is 256×256. The learning rate is 0.0002 in the first 30 epochs and linearly decays to 1e-8 in the last 30 epochs. The generation model is trained using Adam optimizer with parameters $(0.5, 0.99)$. For comparative experiment, we use the source codes of pix2pixHD [61], CycleGAN [73] and BicycleGAN [74] in their original papers.

Recognition Model Training. For our recognition model, we use ResNet50 [31] and MobileFaceNet [10] as the backbone with the input resolution of 224×224. The model is first pretrained on synthesized data for 25 epochs and then finetuned on real datasets for 50 epochs. For the baseline model, we train the model on real datasets for 50 epochs. The feature extractor in ID-aware Loss uses the same training setting as the baseline. We use Arcface [14] with margin $m = 0.5$ and scale factor $s = 48$ for the pretraining, finetuning and baseline training supervision. We use the cosine learning rate schedule with a warmup start for one epoch. The maximum learning rate is 1e-2 and the minimum learning rate is 1e-6 for pretraining and finetuning. All recognition models are trained with a mini-batch SGD optimizer. We use 4 NVIDIA Tesla V100 GPUs for training with total batch size of 128.

It should be emphasized that the generation model, the feature extractor in ID-aware loss, and the recognition model all ensure that the training and test sets are completely isolated.

5. Experimental Results

5.1. Open-set Palmprint Recognition

We first test our method under the open-set protocol with two different training and test ratios 1:1 and 1:3 (trainIDs:testIDs=1634:1632, 818:2448). Details about the "Open-set" protocol can be found in supplementary materials. The quantitative results are shown in Tab.2. The TAR $v.s.$ FAR curves of the 1:1 setting are in Fig.5. Our method outperforms BézierPalm by 5.09% and 14.73% under 1:1 and 1:3 settings @FAR=1e-6 using 'MB', which establishes a new state-of-the-art. Our method achieves more significant improvement under 1:3 setting than 1:1 setting, which demonstrates the effectiveness of our method in scenarios with only a small amount of real training data.

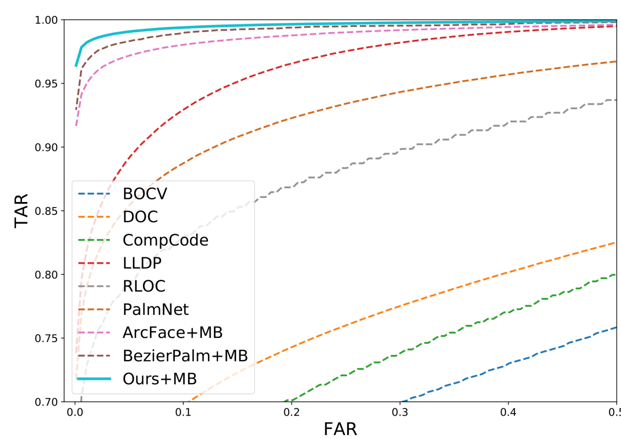

Figure 5: TAR@FAR curves of different methods under the open-set 1:1 setting. The backbone of ArcFace, BézierPalm and our method is MobileFaceNet.

5.2. Palmprint Recognition with Limited Identities

To further verify the performance of our method with limited real training IDs, we test our method with different sizes of real training IDs and fix the test set under the 1:1

Method	Backbone	train : test = 1 : 1				train : test = 1 : 3			
		TAR@ 1e-3	TAR@ 1e-4	TAR@ 1e-5	TAR@ 1e-6	TAR@ 1e-3	TAR@ 1e-4	TAR@ 1e-5	TAR@ 1e-6
CompCode [40]	N/A	0.4800	0.4292	0.3625	0.2103	0.4501	0.3932	0.3494	0.2648
LLDP [43]	N/A	0.7382	0.6762	0.5222	0.1247	0.7372	0.6785	0.6171	0.2108
BOCV [28]	N/A	0.4930	0.4515	0.3956	0.2103	0.4527	0.3975	0.3527	0.2422
RLOC [34]	N/A	0.6490	0.5884	0.4475	0.1443	0.6482	0.5840	0.5224	0.3366
DOC [19]	N/A	0.4975	0.4409	0.3712	0.1667	0.4886	0.4329	0.3889	0.2007
PalmNet [25]	N/A	0.7174	0.6661	0.5992	0.1069	0.7217	0.6699	0.6155	0.2877
C-LMCL [71]	MB	0.9290	0.8554	0.7732	0.6239	0.8509	0.7554	0.7435	0.5932
ArcFace [14]	MB	0.9292	0.8568	0.7812	0.7049	0.8516	0.7531	0.6608	0.5825
BézierPalm [69]	MB	0.9640	0.9438	0.9102	0.8437	0.9407	0.8861	0.7934	0.7012
Ours	MB	**0.9802**	**0.9714**	**0.9486**	**0.8946**	**0.9496**	**0.9267**	**0.8969**	**0.8485**
C-LMCL [71]	R50	0.9545	0.9027	0.8317	0.7534	0.8601	0.7701	0.6821	0.6254
ArcFace [14]	R50	0.9467	0.8925	0.8252	0.7462	0.8709	0.7884	0.7156	0.6580
BézierPalm [69]	R50	0.9671	0.9521	0.9274	0.8956	0.9424	0.8950	0.8217	0.7649
Ours	R50	**0.9821**	**0.9732**	**0.9569**	**0.9347**	**0.9533**	**0.9319**	**0.9016**	**0.8698**

Table 2: Quantitative performances under the open-set protocol where the performances are evaluated in terms of TAR@FAR. 'MB' represents the MobileFaceNet [10] backbone and 'R50' is resnet50 [31] backbone.

open-set protocol. We synthesize 4000 identities and 100 samples for each identity in all experiments. As shown in Tab.3, the performance of ArcFace and BézierPalm drops quickly with the reduction of real data, while our method still performs well with few real training IDs. Note that our method with 160(10%) real IDs even outperforms Arcface with 1600 real IDs. Besides, our method significantly improves the TAR by 20.81%@FAR=1e-6 against BézierPalm [69] with 80 training IDs, which shows the superiority of our method under few real training IDs.

5.3. Palmprint Recognition at Million Scale

In order to verify the effectiveness of our method on large-scale real-world datasets, we test our method on our internal dataset with millions of palmprint images. Our dataset contains $19,286$ training identities with $2,871,073$ images and $1,000$ test identities with $182,732$ images. For a fair comparison with BézierPalm, we generate $20,000$ IDs with 100 samples in each ID for pretraining. Quantitative results are shown in Tab.4. Our method outperforms ArcFace and BézierPalm with a clear margin, showing its practical application value on large-scale datasets.

Method	#ID	TAR@FAR= 1e-3	1e-4	1e-5	1e-6
ArcFace	1,600	0.9292	0.8568	0.7812	0.7049
BézierPalm	1,600	0.9640	0.9438	0.9102	0.8437
Ours	1,600	**0.9802**	**0.9714**	**0.9486**	**0.8946**
ArcFace	800	0.8934	0.7432	0.7104	0.6437
BézierPalm	800	0.9534	0.9390	0.9025	0.8164
Ours	800	**0.9783**	**0.9687**	**0.9356**	**0.8741**
ArcFace	400	0.8102	0.7050	0.6668	0.3320
BézierPalm	400	0.9189	0.8497	0.7542	0.6899
Ours	400	**0.9573**	**0.9324**	**0.8836**	**0.8162**
ArcFace	160	0.6761	0.5294	0.4783	0.2437
BézierPalm	160	0.8179	0.6998	0.5826	0.4832
Ours	160	**0.9356**	**0.8641**	**0.8063**	**0.7246**
ArcFace	80	0.5384	0.4682	0.3249	0.1173
BézierPalm	80	0.6547	0.5511	0.4490	0.3743
Ours	80	**0.8974**	**0.8092**	**0.6947**	**0.5824**

Table 3: Performance under different real training identities. The generation model, feature extractor in ID-aware loss, and the recognition model access the same number of real training identities. The backbone is MobileFaceNet.

6. Ablation Study

In this section, we conduct ablation studies to verify different components and design choices of our method. The MobileFaceNet is used as the backbone for all experiments under the same Open-set protocol.

6.1. Components and design choices

The main components and design choices of our method are ID-aware loss, conditional modulation generator and improved strategy for Bézier palm creases synthesis. Tab.5 shows the results of recognition models with or without these components. 'I', 'C', 'S' represent the three components respectively. For baseline without 'C', the normal UNet [49] as in [74] is used to take place of 'C'. ID-aware loss achieves the greatest improvement by 11.72%@FAR=1e-6 against the baseline, which reflects its advantage for preserving intra-class consistency. The con-

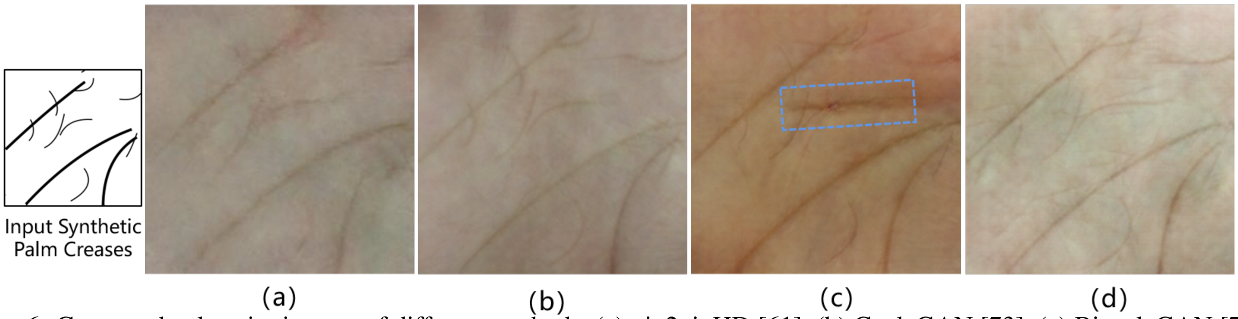

Figure 6: Generated palmprint images of different methods, (a) pix2pixHD [61], (b) CycleGAN [73], (c) BicycleGAN [74], (d) the proposed method.

Method	Backbone	TAR@FAR= 1e-6	1e-7	1e-8	1e-9
ArcFace	MB	0.9770	0.9550	0.9251	0.8833
BézierPalm		0.9803	0.9605	0.9301	0.9015
Ours		**0.9871**	**0.9684**	**0.9416**	**0.9194**
ArcFace	R50	0.9986	0.9964	0.9931	0.9879
BézierPalm		0.9996	0.9975	0.9943	0.9911
Ours		**0.9998**	**0.9983**	**0.9972**	**0.9954**

Table 4: Palmprint recognition performance on million scale dataset.

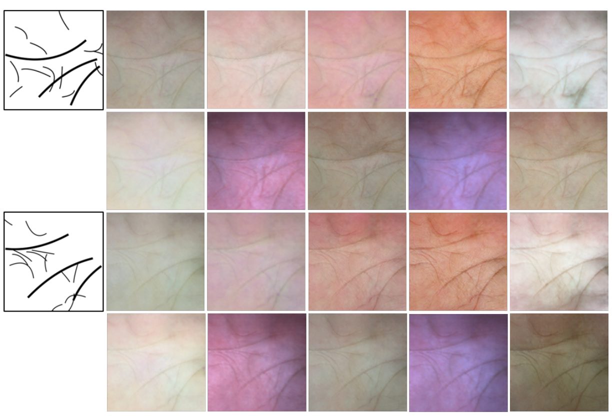

Figure 7: Diverse synthetic palmprints can be generated by adjusting the input noise vector $N(z)$.

ditional modulation generator also significantly improves the performance by 4.49%@FAR=1e-6, which reflects its advantage for improving the intra-class diversity. The improved Bézier palm creases bring an improvement by 1.15%@FAR=1e-6.

6.2. Comparison of Generation Methods

Three generation methods pix2pixHD [61], CycleGAN [73], BicycleGAN [74] are used for comparison, and all of them are retrained on the same training set with unpaired data. Fig.6 illustrates the generated palmprints of different generation methods. It can be found that pix2pixHD

I	C	S	TAR@FAR= 1e-3	1e-4	1e-5	1e-6
✗	(UNet)	✗	0.9399	0.8905	0.8164	0.7210
✓	(UNet)	✗	0.9687	0.9571	0.9076	0.8382
✓	✓	✗	0.9796	0.9689	0.9441	0.8831
✓	✓	✓	**0.9802**	**0.9714**	**0.9486**	**0.8946**

Table 5: Ablation of different components in our method. 'I', 'C' and 'S' denote ID-aware loss, conditional modulation generator and improved synthetic creases, respectively.

Method	FID↓	train:test=1:1 TAR↑ @1e-5	train:test=1:1 TAR↑ @1e-6	train:test=1:3 TAR↑ @1e-5	train:test=1:3 TAR↑ @1e-6
ArcFace [14]	- -	0.7812	0.7049	0.6608	0.5825
BézierPalm [69]	- -	0.9102	0.8437	0.7934	0.7012
pix2pixHD [61]	97.5801	0.9156	0.8734	0.8052	0.7244
CycleGAN [73]	50.7704	0.9136	0.8703	0.7863	0.7189
BicycleGAN [74]	35.2801	0.8173	0.7254	0.6783	0.5929
Ours	**16.4762**	**0.9486**	**0.8946**	**0.8969**	**0.8485**

Table 6: Quantitative recognition results using different generation methods under the open-set protocol.

and CycleGAN tend to synthesize blurred results. As marked in blue rectangle, BicycleGAN may generate some creases that are inconsistent with the input Bézier curves. Our method can synthesize clearer and sharper principal lines and faithfully preserve the ID information of the input Bézier curves. For intra-class diversity, as shown in Fig.7, our method is able to randomly generate high-fidelity palmprints with diverse lighting and skin types by adjusting the input noise vector $N(z)$.

Quantitative results are shown in Tab.6. Our method substantially outperforms the existing methods in terms of TAR@FAR. Also, we use FID [8] to evaluate the quality of generated results and our method also effectively decreases the FID score by more than 50% against other methods. Besides, we reproduced some other generation methods [54, 56] for unpaired image-to-image transfer, including diffusion-based models, but the generated results are unsatisfying. Related details and more subjective results can be found in supplementary material.

6.3. Number of Synthesized Identities and Images

In this ablation, we investigate the influence of the number of synthesized identities and images. Specifically, we generate 4000 identities and 100 samples by default, and fix one number and vary the other under the open-set protocol(train:test=1:1). The results are shown in Fig.8. With the increase of "width" and "depth" of synthetic data, our method can continuously improve the performance of the recognition model and achieve a higher upper bound than BézierPalm. The performance reaches the upper bound with $80k$ synthetic identities and 160 samples per identity.

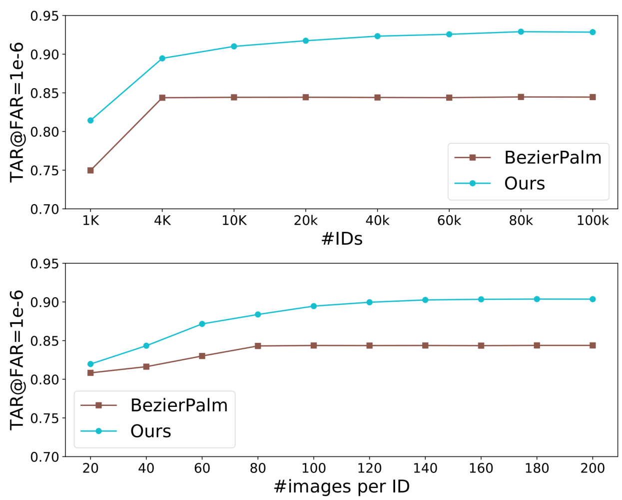

Figure 8: TAR@FAR=1e-6 of recognition models pre-trained with different numbers of synthetic identities and samples. The backbone is MobileFaceNet.

7. Conclusion

This paper proposed an ID-aware conditional modulation generation model which can produce realistic and diversified palmprint images. Specifically, a conditional modulation generator was designed, which adopted synthetic palm creases as condition, and used encoded Gaussian noise vector to modulate the generated diversity. An ID-aware loss was proposed to preserve the identity information of input palm creases during the unpaired training process. In the forward pseudo-palmprint generation stage, we improved the Bézier curves generation strategy to produce more realistic synthetic palm creases. From experimental results, we can obtain the following findings. Firstly, the generated pseudo-palmprint samples can effectively improve the performance of palmprint recognition models. Secondly, by using the synthetic palmprints, our method can effectively reduce the dependence of real data by 90%. In future work, we hope to implement complete real-data-free palmprint recognition with synthetic data.

8. Acknowledgements

This work is partly supported by the grants of the National Natural Science Foundation of China, Nos.62076086, 62272142 and 61972129.

References

[1] Amazon one. `https://one.amazon.com/`.

[2] The ceop dataset. `https://www.coep.org.in/resources/coeppalmprintdatabase`.

[3] Denise Almeida, Konstantin Shmarko, and Elizabeth Lomas. The ethics of facial recognition technologies, surveillance, and accountability in an age of artificial intelligence: a comparative analysis of us, eu, and uk regulatory frameworks. *AI and Ethics*, 2(3):377–387, 2022.

[4] Xiang An, Xuhan Zhu, Yuan Gao, Yang Xiao, Yongle Zhao, Ziyong Feng, Lan Wu, Bin Qin, Ming Zhang, Debing Zhang, et al. Partial fc: Training 10 million identities on a single machine. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 1445–1449, 2021.

[5] Keivan Bahmani, Richard Plesh, Peter Johnson, Stephanie Schuckers, and Timothy Swyka. High fidelity fingerprint generation: Quality, uniqueness, and privacy. In *2021 IEEE International Conference on Image Processing (ICIP)*, pages 3018–3022. IEEE, 2021.

[6] Volker Blanz and Thomas Vetter. A morphable model for the synthesis of 3d faces. In *Proceedings of the 26th annual conference on Computer graphics and interactive techniques*, pages 187–194, 1999.

[7] Fadi Boutros, Marco Huber, Patrick Siebke, Tim Rieber, and Naser Damer. Sface: Privacy-friendly and accurate face recognition using synthetic data. In *2022 IEEE International Joint Conference on Biometrics (IJCB)*, pages 1–11. IEEE, 2022.

[8] Naresh Babu Bynagari. Gans trained by a two time-scale update rule converge to a local nash equilibrium. *Asian Journal of Applied Science and Engineering*, 8:25–34, 2019.

[9] Qiong Cao, Li Shen, Weidi Xie, Omkar M Parkhi, and Andrew Zisserman. Vggface2: A dataset for recognising faces across pose and age. In *FG 2018*, pages 67–74. IEEE, 2018.

[10] Sheng Chen, Yang Liu, Xiang Gao, and Zhen Han. Mobilefacenets: Efficient cnns for accurate real-time face verification on mobile devices. In *Chinese Conference on Biometric Recognition*, pages 428–438. Springer, 2018.

[11] Yuan Chen, Yang Zhao, Wei Jia, Li Cao, and Xiaoping Liu. Adversarial-learning-based image-to-image transformation: A survey. *Neurocomputing*, 411:468–486, 2020.

[12] Yunjey Choi, Minje Choi, Munyoung Kim, Jung-Woo Ha, Sunghun Kim, and Jaegul Choo. Stargan: Unified generative adversarial networks for multi-domain image-to-image translation. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 8789–8797, 2018.

[13] Ugur Demir and Gozde Unal. Patch-based image inpainting with generative adversarial networks. *arXiv preprint arXiv:1803.07422*, 2018.

[14] Jiankang Deng, Jia Guo, Niannan Xue, and Stefanos Zafeiriou. Arcface: Additive angular margin loss for deep face recognition. In *CVPR*, pages 4690–4699, 2019.

[15] Yu Deng, Jiaolong Yang, Dong Chen, Fang Wen, and Xin Tong. Disentangled and controllable face image generation via 3d imitative-contrastive learning. In *IEEE Computer Vision and Pattern Recognition*, 2020.

[16] Liu Dian and Sun Dongmei. Contactless palmprint recognition based on convolutional neural network. In *IEEE ICSP*, pages 1363–1367. IEEE, 2016.

[17] Joshua J Engelsma, Steven A Grosz, and Anil K Jain. Printsgan: Synthetic fingerprint generator. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2022.

[18] Lunke Fei, Guangming Lu, Wei Jia, Shaohua Teng, and David Zhang. Feature extraction methods for palmprint recognition: A survey and evaluation. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 49(2):346–363, 2018.

[19] Lunke Fei, Yong Xu, Wenliang Tang, and David Zhang. Double-orientation code and nonlinear matching scheme for palmprint recognition. *PR*, 49:89–101, 2016.

[20] Guiyu Feng, Dewen Hu, David Zhang, and Zongtan Zhou. An alternative formulation of kernel lpp with application to image recognition. *Neurocomputing*, 69(13-15):1733–1738, 2006.

[21] Miguel A Ferrer, Francisco Vargas, and Aythami Morales. Bispectral contactless hand based biometric system. In *CONATEL 2011*, pages 1–6. IEEE, 2011.

[22] Chaoyou Fu, Xiang Wu, Yibo Hu, Huaibo Huang, and Ran He. Dual variational generation for low shot heterogeneous face recognition. *Advances in neural information processing systems*, 32, 2019.

[23] Chaoyou Fu, Xiang Wu, Yibo Hu, Huaibo Huang, and Ran He. Dvg-face: Dual variational generation for heterogeneous face recognition. *IEEE transactions on pattern analysis and machine intelligence*, 44(6):2938–2952, 2021.

[24] Zhenglin Geng, Chen Cao, and Sergey Tulyakov. 3d guided fine-grained face manipulation. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 9821–9830, 2019.

[25] Angelo Genovese, Vincenzo Piuri, Konstantinos N Plataniotis, and Fabio Scotti. Palmnet: Gabor-pca convolutional networks for touchless palmprint recognition. *IEEE TIFS*, 14(12):3160–3174, 2019.

[26] Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative adversarial networks. *Communications of the ACM*, 63(11):139–144, 2020.

[27] Yandong Guo, Lei Zhang, Yuxiao Hu, Xiaodong He, and Jianfeng Gao. Ms-celeb-1m: A dataset and benchmark for large-scale face recognition. In *ECCV*, pages 87–102. Springer, 2016.

[28] Zhenhua Guo, David Zhang, Lei Zhang, and Wangmeng Zuo. Palmprint verification using binary orientation co-occurrence vector. *PRL*, 30(13):1219–1227, 2009.

[29] Ying Hao, Zhenan Sun, Tieniu Tan, and Chao Ren. Multispectral palm image fusion for accurate contact-free palmprint recognition. In *ICIP*, pages 281–284. IEEE, 2008.

[30] Ahmad Hassanat, Mouhammd Al-Awadi, Eman Btoush, Amani Al-Btoush, Esra'a Alhasanat, and Ghada Altarawneh. New mobile phone and webcam hand images databases for personal authentication and identification. *Procedia Manufacturing*, 3:4060–4067, 2015.

[31] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *CVPR*, pages 770–778, 2016.

[32] Dewen Hu, Guiyu Feng, and Zongtan Zhou. Two-dimensional locality preserving projections (2dlpp) with its application to palmprint recognition. *PR*, 40(1):339–342, 2007.

[33] Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, and Alexei A Efros. Image-to-image translation with conditional adversarial networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1125–1134, 2017.

[34] Wei Jia, De-Shuang Huang, and David Zhang. Palmprint verification based on robust line orientation code. *PR*, 41(5):1504–1513, 2008.

[35] Wei Jia, Qiang Ren, Yang Zhao, Shujie Li, Hai Min, and Yanxiang Chen. Eepnet: An efficient and effective convolutional neural network for palmprint recognition. *Pattern Recognition Letters*, 159:140–149, 2022.

[36] Vivek Kanhangad, Ajay Kumar, and David Zhang. Contactless and pose invariant biometric identification using hand surface. *IEEE TIP*, 20(5):1415–1424, 2010.

[37] Tero Karras, Samuli Laine, and Timo Aila. A style-based generator architecture for generative adversarial networks. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 4401–4410, 2019.

[38] Ira Kemelmacher-Shlizerman, Steven M Seitz, Daniel Miller, and Evan Brossard. The megaface benchmark: 1 million faces for recognition at scale. In *CVPR*, pages 4873–4882, 2016.

[39] Taeksoo Kim, Moonsu Cha, Hyunsoo Kim, Jung Kwon Lee, and Jiwon Kim. Learning to discover cross-domain relations with generative adversarial networks. In *International conference on machine learning*, pages 1857–1865. PMLR, 2017.

[40] AW-K Kong and David Zhang. Competitive coding scheme for palmprint verification. In *ICPR*, volume 1, pages 520–523. IEEE, 2004.

[41] Ajay Kumar. Incorporating cohort information for reliable palmprint authentication. In *2008 Sixth Indian conference on computer vision, graphics & image processing*, pages 583–590. IEEE, 2008.

[42] Guangming Lu, David Zhang, and Kuanquan Wang. Palmprint recognition using eigenpalms features. *PRL*, 24(9-10):1463–1467, 2003.

[43] Yue-Tong Luo, Lan-Ying Zhao, Bob Zhang, Wei Jia, Feng Xue, Jing-Ting Lu, Yi-Hai Zhu, and Bing-Qing Xu. Local line directional pattern for palmprint recognition. *PR*, 50:26–44, 2016.

[44] Brianna Maze, Jocelyn Adams, James A Duncan, Nathan Kalka, Tim Miller, Charles Otto, Anil K Jain, W Tyler Niggel, Janet Anderson, Jordan Cheney, et al. Iarpa janus benchmark-c: Face dataset and protocol. In *ICB*, pages 158–165. IEEE, 2018.

[45] Thu Nguyen-Phuoc, Chuan Li, Lucas Theis, Christian Richardt, and Yong-Liang Yang. Hologan: Unsupervised learning of 3d representations from natural images. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 7588–7597, 2019.

[46] Haibo Qiu, Baosheng Yu, Dihong Gong, Zhifeng Li, Wei Liu, and Dacheng Tao. Synface: Face recognition with synthetic data. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 10880–10890, 2021.

[47] Aditya Ramesh, Mikhail Pavlov, Gabriel Goh, Scott Gray, Chelsea Voss, Alec Radford, Mark Chen, and Ilya Sutskever. Zero-shot text-to-image generation. In *International Conference on Machine Learning*, pages 8821–8831. PMLR, 2021.

[48] Robin Rombach, Andreas Blattmann, Dominik Lorenz, Patrick Esser, and Björn Ommer. High-resolution image synthesis with latent diffusion models. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 10684–10695, 2022.

[49] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III 18*, pages 234–241. Springer, 2015.

[50] Chitwan Saharia, Jonathan Ho, William Chan, Tim Salimans, David J Fleet, and Mohammad Norouzi. Image super-resolution via iterative refinement. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2022.

[51] Haifeng Sang, Weiqi Yuan, and Zhijia Zhang. Research of palmprint recognition based on 2dpca. In *International Symposium on Neural Networks*, pages 831–838. Springer, 2009.

[52] Huikai Shao, Dexing Zhong, and Xuefeng Du. Effective deep ensemble hashing for open-set palmprint recognition. *Journal of Electronic Imaging*, 29(1):013018, 2020.

[53] Huikai Shao, Dexing Zhong, and Yuhan Li. Palmgan for cross-domain palmprint recognition. In *2019 IEEE International Conference on Multimedia and Expo (ICME)*, pages 1390–1395. IEEE, 2019.

[54] Xuning Shao and Weidong Zhang. Spatchgan: A statistical feature based discriminator for unsupervised image-to-image translation. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 6546–6555, 2021.

[55] Lei Shen, Yingyi Zhang, Kai Zhao, Ruixin Zhang, and Wei Shen. Distribution alignment for cross-device palmprint recognition. *Pattern Recognition*, 132:108942, 2022.

[56] Xuan Su, Jiaming Song, Chenlin Meng, and Stefano Ermon. Dual diffusion implicit bridges for image-to-image translation. In *International Conference on Learning Representations*, 2023.

[57] Zhenan Sun, Tieniu Tan, Yunhong Wang, and Stan Z Li. Ordinal palmprint represention for personal identification [represention read representation]. In *CVPR*, volume 1, pages 279–284. IEEE, 2005.

[58] Jan Svoboda, Jonathan Masci, and Michael M Bronstein. Palmprint recognition via discriminative index learning. In *ICPR*, pages 4232–4237. IEEE, 2016.

[59] Meng Wang and Qiuqi Ruan. Palmprint recognition based on two-dimensional methods. In *ICSP*, volume 4. IEEE, 2006.

[60] Tengfei Wang, Ting Zhang, Bo Zhang, Hao Ouyang, Dong Chen, Qifeng Chen, and Fang Wen. Pretraining is all you need for image-to-image translation. *arXiv preprint arXiv:2205.12952*, 2022.

[61] Ting-Chun Wang, Ming-Yu Liu, Jun-Yan Zhu, Andrew Tao, Jan Kautz, and Bryan Catanzaro. High-resolution image synthesis and semantic manipulation with conditional gans. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2018.

[62] Cameron Whitelam, Emma Taborsky, Austin Blanton, Brianna Maze, Jocelyn Adams, Tim Miller, Nathan Kalka, Anil K Jain, James A Duncan, Kristen Allen, et al. Iarpa janus benchmark-b face dataset. In *CVPRW*, pages 90–98, 2017.

[63] André Brasil Vieira Wyzykowski, Mauricio Pamplona Segundo, and Rubisley de Paula Lemes. Level three synthetic fingerprint generation. In *2020 25th International Conference on Pattern Recognition (ICPR)*, pages 9250–9257. IEEE, 2021.

[64] Zili Yi, Hao Zhang, Ping Tan, and Minglun Gong. Dualgan: Unsupervised dual learning for image-to-image translation. In *Proceedings of the IEEE international conference on computer vision*, pages 2849–2857, 2017.

[65] Xi Yin, Xiang Yu, Kihyuk Sohn, Xiaoming Liu, and Manmohan Chandraker. Towards large-pose face frontalization in the wild. In *In Proceeding of International Conference on Computer Vision*, Venice, Italy, October 2017.

[66] David Zhang, Zhenhua Guo, Guangming Lu, Lei Zhang, and Wangmeng Zuo. An online system of multispectral palmprint verification. *IEEE TIM*, 59(2):480–490, 2009.

[67] Yingyi Zhang, Lin Zhang, Xiao Liu, Shengjie Zhao, Ying Shen, and Yukai Yang. Pay by showing your palm: A study of palmprint verification on mobile platforms. In *2019 IEEE International Conference on Multimedia and Expo (ICME)*, pages 862–867. IEEE, 2019.

[68] Yingyi Zhang, Lin Zhang, Ruixin Zhang, Shaoxin Li, Jilin Li, and Feiyue Huang. Towards palmprint verification on smartphones. *arXiv preprint arXiv:2003.13266*, 2020.

[69] Kai Zhao, Lei Shen, Yingyi Zhang, Chuhan Zhou, Tao Wang, Ruixin Zhang, Shouhong Ding, Wei Jia, and Wei Shen. Bézierpalm: A free lunch for palmprint recognition. In *Computer Vision–ECCV 2022: 17th European Conference, Tel Aviv, Israel, October 23–27, 2022, Proceedings, Part XIII*, pages 19–36. Springer, 2022.

[70] Qian Zheng, Ajay Kumar, and Gang Pan. Suspecting less and doing better: New insights on palmprint identification for faster and more accurate matching. *IEEE TIFS*, 11(3):633–641, 2015.

[71] Dexing Zhong and Jinsong Zhu. Centralized large margin cosine loss for open-set deep palmprint recognition. *IEEE TCSVT*, 2019.

[72] Jun-Yan Zhu, Taesung Park, Phillip Isola, and Alexei A Efros. Unpaired image-to-image translation using cycle-consistent adversarial networks. In *Proceedings of the IEEE international conference on computer vision*, pages 2223–2232, 2017.

[73] Jun-Yan Zhu, Taesung Park, Phillip Isola, and Alexei A Efros. Unpaired image-to-image translation using cycle-consistent adversarial networkss. In *Computer Vision (ICCV), 2017 IEEE International Conference on*, 2017.

[74] Jun-Yan Zhu, Richard Zhang, Deepak Pathak, Trevor Darrell, Alexei A Efros, Oliver Wang, and Eli Shechtman. Toward multimodal image-to-image translation. In *Advances in Neural Information Processing Systems*, 2017.