Supplementary Material For RPG-Palm: Realistic Pseudo-data Generation for Palmprint Recognition

Appendix

This supplementary includes:

A: More Implementation Details: Open-set Protocol, Evaluation protocol, Improved Bézier Plam Creases Synthesis.

B: More Palmprint Recognition Results: Closed-set Validation, Cross-dataset Validation.

C: More Ablation Studies: RLOC Threshold, Weight of $\mathcal{L}_{ID}$, Palmprint Recognition model in ID-aware Loss, Number of Synthesized Identities on Million scale dataset.

D: More Subjective Results and Comparisons: More Comparisons of Generated Palmprints, Diversity and Identity Consistency of Our Method, Comparison with Recent Generation Models.

A. More Implementation Details

A.1. Open-set Protocol

Dataset split setting. Under the Open-set protocol, we randomly select part of identities from each public dataset and combine them as the training set, and the other identities are merged as the test set following BézierPalm [16]. Example ROIs of used public datasets are shown in Fig.1.

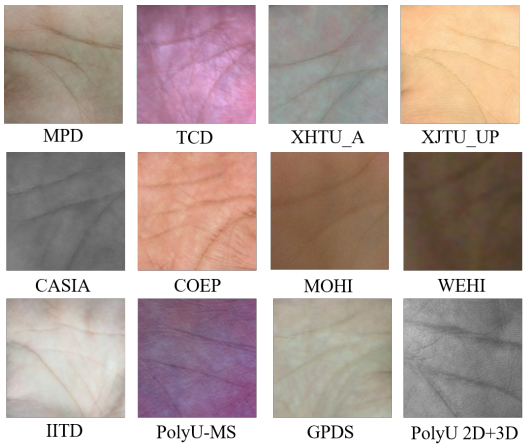

Figure 1: Example ROIs of different public datasets.

We conduct two split settings according to BézierPalm. In the first setting, 1/2 of the identities are used for training and the left 1/2 identities are used for testing. In the second setting, 1/4 of the identities are used for training and the left 3/4 identities are used for testing. The number of images and identities under two split settings are listed in Tab.1.

Split setting	Mode	#IDs	#Images
train:test = 1:1	train	1,634	29,347
	test	1,632	29,815
train:test = 1:3	train	818	14,765
	test	2,448	44,397

Table 1: Split settings under open-set protocol.

A.2. Evaluation protocol

We use TAR (True Acceptance Rate) @FAR (False Acceptance Rate) to evaluate the model performance following BézierPalm [16], which is widely used for Open-set recognition tasks, e.g. face recognition. Specifically, given several test images, we randomly sample several positive pairs where the two samples share the same identity, and negative pairs whose samples are from distinct identities. Let sim_{pos}, sim_{neg} be the cosine similarities of positive and negative pairs. We fix the FAR and calculate the corresponding similarity threshold t from N_{neg} negative pairs, then we compute TAR using the threshold t on the N_{pos} positive pairs. The threshold t@different FAR can be searched by the following metric, where sim^i_{neg} is the cosine similarity of the i-th negative pair, N_{neg} is the number of negative pairs (N_{pos}=28,183 N_{neg}=166,094,362 for 1:1 setting, N_{pos}=41,949 N_{neg}=210,205,085 for 1:3 setting):

$$\text{FAR} = \frac{1}{N_{neg}} \sum_{i=1}^{N_{neg}} (sim^i_{neg} > t) \tag{1}$$

The TAR is calculated with the threshold t as follows, where sim^j_{pos} is the cosine similarity of the j-th positive pair, N_{pos} is the number of positive pairs:

$$\text{TAR} = \frac{1}{N_{pos}} \sum_{j=1}^{N_{pos}} (sim^j_{pos} > t) \tag{2}$$

A.3. Improved Bézier Plam Creases Synthesis

The different preset ranges of parameter points of BézierPalm [16](based on the official released code) and

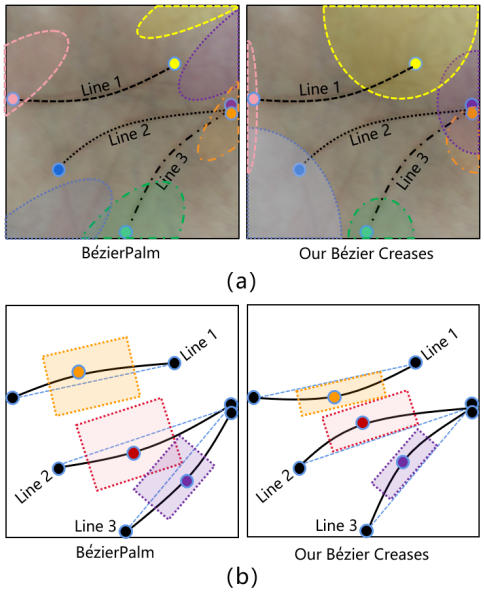

Figure 2: Preset ranges of parameter points in BezierPalm [16] and our method, (a) ranges of start points and end points, (b) ranges of control points.

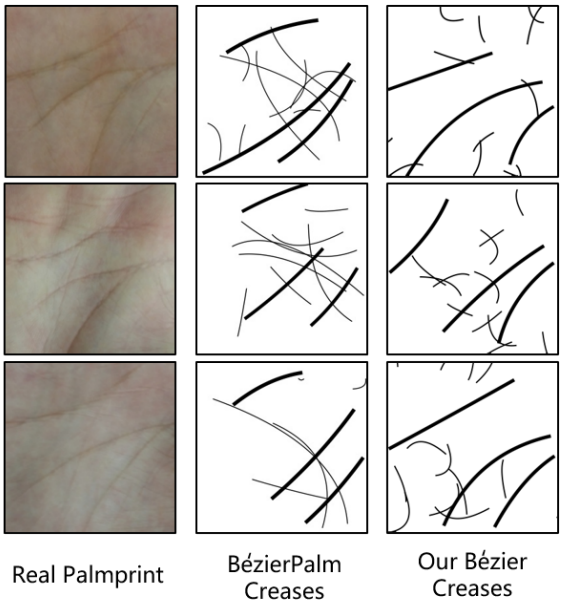

Figure 3: Comparisons of synthetic creases in BézierPalm [16] and our method.

our method are shown in Fig.2. As illustrated in Fig.2(a), BézierPalm [16] adopts preset random positions for the start points and end points, but ignores the real distributions of these points. By roughly counting the start points and end points on real palmprints, we observe that the ranges of these points are different. For instance, the start point of line 1 is usually near the left edge of the palmprint. But some other points may be distributed in a much larger range, such as the end point of line 1 and the start point of line 2. In addition, the uppermost principal line (line 1) is basically convex downwards, while the other two principal lines (line 2 and line 3) are convex upwards. Therefore, we also adjust the location of control points to produce more reasonable convex direction, as shown in Fig.2(b).

As shown in Fig.3, by comparing with real palmprints, the synthetic palm creases in our method are more reasonable than that of BézierPalm. The detailed parameters of ranges are provided in our open-sourced codes. Also note that we only illustrate the right palm, and the situation of the left palm is a simple mirror image of the right palm.

B. More Palmprint Recognition Results

B.1. Closed-set Validation

We perform experiments on five datasets: CASIA, IITD, PloyU, TCD, MPD under Closed-set protocol, where the training and test set share the same identities but the samples are different. We conduct 5-fold cross-validation on each dataset and report the average top-1 accuacy and EER(Equal Error Rate) result. We randomly select one sample from each identity as the register image and the left samples are query images. Top-1 accuracy is the number of successfully matched queries divided by the total number of queries. The EER is the intersection of FAR(False Acceptance Rate) and FRR(False Rejection Rate). We generate 4,000 identities and 100 samples by default. The experiments are conducted individually on the five datasets.

The quantitative results can be found in Tab.2. Our method achieves saturated results on all five datasets, where top-1 results are all nearly 100% and EER results are all nearly 0.000.

B.2. Cross-dataset Validation

In real-world applications, the training and testing dataset may not be the same. To evaluate the cross-dataset generalization of the proposed method, we add 5 different cross-dataset experiments using MPD, PolyU, TCD and IITD datasets. We train our generation and recognition model on one dataset and test the performance on the other dataset. We generate 4,000 identities and 100 samples by default.

TAR@FAR, top-1 and EER are used to evaluate the model generalization performance. As shown in Tab.3, our method surpasses BézierPalm [16] remarkably on 4 settings and reaches a saturated result on the M→I setting, which verifies the cross-dataset generalization of our method.

Method	CASIA	IITD	PolyU	TCD	MPD
CompCode [8]	79.27 / 1.08	77.79 / 1.39	99.21 / 0.68	- / -	- / -
Ordinal Code [14]	73.32 / 1.75	73.26 / 2.09	99.55 / 0.23	- / -	- / -
DoN [18]	99.30 / 0.53	99.15 / 0.68	100.0 / 0.22	- / -	- / -
PalmNet [4]	97.17 / 3.21	97.31 / 3.83	99.95 / 0.39	99.89 / 0.40	91.88 / 6.22
FERNet [10]	97.65 / 0.73	99.61 / 0.76	99.77 / 0.15	98.63 / -	- / -
DDBC [3]	96.41 / -	96.44 / -	-	98.73 / -	- / -
RFN [9]	- / -	99.20 / 0.60	- / -	- / -	- / -
C-LMCL [19]	- / -	- / -	100.0 / 0.13	99.93 / 0.26	- / -
JCLSR [17]	98.94 / -	98.17 / -	- / -	- / -	- / -
ArcFace [2] + MB	97.92 / 0.009	98.73 / 0.012	98.58 / 0.014	98.83 / 0.008	96.12 / 0.022
BézierPalm [16] + MB	99.75 / 0.004	**100.0 / 0.000**	**100.0 / 0.000**	**100.0 / 0.000**	99.96 / 0.001
Ours + MB	**99.96 / 0.001**	**100.0 / 0.000**	**100.0 / 0.000**	**100.0 / 0.000**	**100.0 / 0.000**

Table 2: Top-1 accuracy and EER under the 'closed-set' protocol.

Datasets	Method	TAR@FAR= 1e-4	1e-5	Top-1	EER
M→P	AF	0.9499	0.9210	99.93	0.007
	BézierPalm	0.9766	0.9622	**100.0**	**0.002**
	Ours	**0.9907**	**0.9863**	**100.0**	**0.002**
T→P	AF	0.8981	0.8509	98.22	0.018
	BézierPalm	0.9748	0.9591	**100.0**	**0.003**
	Ours	**0.9826**	**0.9753**	**100.0**	**0.003**
I→P	AF	0.9001	0.8020	97.67	0.019
	BézierPalm	0.9224	0.8728	99.04	0.009
	Ours	**0.9436**	**0.9047**	**99.37**	**0.005**
T→I	AF	0.7872	0.7306	97.47	0.033
	BézierPalm	0.9864	0.9745	98.85	0.007
	Ours	**0.9916**	**0.9871**	**99.14**	**0.003**
M→I	AF	0.9846	0.9717	99.76	0.004
	BézierPalm	**1.0000**	**1.0000**	**100.0**	**0.000**
	Ours	**1.0000**	**1.0000**	**100.0**	0.001

Table 3: Cross-dataset validation. 'M', 'P', 'T' and 'I' represent MPD, PolyU, TCD, and IITD datasets, respectively. M→P indicates the model is trained on M and evaluated on P.

RLOC threshold	Pass Rate	TAR@FAR= 1e-3	1e-4	1e-5	1e-6
0.1	0.0%	–	–	–	–
0.3	0.0%	–	–	–	–
0.5	0.0%	–	–	–	–
0.7	2.0%	–	–	–	–
0.8	38.8%	0.9773	0.9662	0.9413	0.8794
0.9	99.4%	**0.9802**	**0.9714**	**0.9486**	**0.8946**
0.95	100.0%	0.9796	0.9689	0.9441	0.8831

Table 4: Ablation of different RLOC thresholds under Open-set 1:1 setting. The backbone is MobileFaceNet.

weight of $\mathcal{L}_{ID}$	TAR@FAR= 1e-3	1e-4	1e-5	1e-6
1.0	0.9791	0.9653	0.9211	0.8636
3.0	0.9796	0.9687	0.9332	0.8725
5.0	**0.9802**	**0.9714**	**0.9486**	**0.8946**
7.0	0.9738	0.9642	0.9251	0.8673
10.0	0.9614	0.9493	0.9138	0.8492

Table 5: Ablation of different weight of $\mathcal{L}_{ID}$ under Open-set 1:1 setting. The backbone is MobileFaceNet.

C. More Ablation Studies

C.1. RLOC Threshold

We ablate the threshold of RLOC [6] with 0.1, 0.3, 0.5, 0.7, 0.8, 0.9, 0.95 under the Open-set 1:1 setting. First, we calculate the passing rate of generating 100,000 synthesized identities with different thresholds. As shown in Tab.4, when the threshold is less than or equal to 0.7, most of the identities are filtered out. Then we uniformly generated 4000 identities with 100 samples for each identity under different thresholds for recongnition model pretraining. As shown in Tab.4, the best recognition result is achieved when the threshold is set as 0.9.

C.2. Weight of $\mathcal{L}_{ID}$

We ablate the weight of $\mathcal{L}_{ID}$ with 1.0, 3.0, 5.0, 7.0, 10.0 under the Open-set 1:1 setting. As shown in Tab.5, the best recognition result is achieved when the weight is set as 5.0. As shown in Fig.4, the clarity of the generated images will be reduced when the weight is set too large.

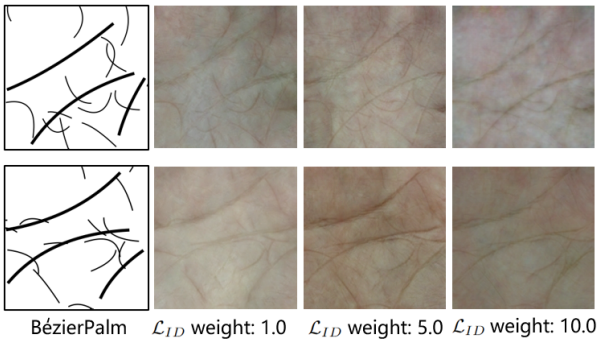

Figure 4: Generated palmprints of different weight of $\mathcal{L}_{ID}$.

C.3. Palmprint Recognition model in ID-aware Loss

We compare the recognition result of using different pretrained palm recognition model in ID-aware loss under the Open-set 1:1 setting. The weight of $\mathcal{L}_{ID}$ is set as 5.0. The quantitative results are shown in Tab.6, the recognition performance of using MoileFaceNet [1] as the discriminator is better than ResNet50 [5]. As shown in Fig.5, the clarity of the generated images also decreases slightly when using ResNet50 as the discriminator, indicating that a too strong discriminator will affect the quality of the generated images.

Discriminator	TAR@FAR=			
	1e-3	1e-4	1e-5	1e-6
MobileFaceNet [1]	**0.9802**	**0.9714**	**0.9486**	**0.8946**
ResNet50 [5]	0.9763	0.9672	0.9321	0.8737

Table 6: Ablation of palmprint recognition model in ID-aware loss under Open-set 1:1 setting.

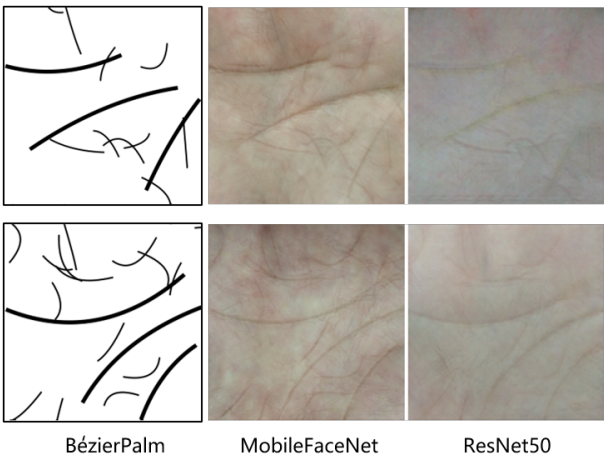

Figure 5: Generated palmprints of different palm recognition model in ID-aware Loss.

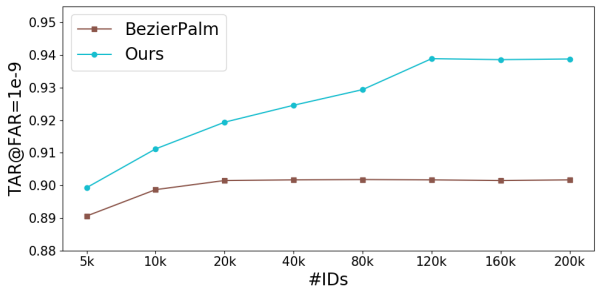

Figure 6: TAR@FAR=1e-9 of recognition models pretrained with different numbers of synthetic identities on million-scale dataset. The backbone is MobileFaceNet.

C.4. Number of Synthesized Identities on Million scale dataset

We also investigate the influence of the number of synthesized identities on million scale dataset. Specifically, we generate 20,000 identities and 100 samples by default, and fix the number of samples and vary the identities under the Open-set protocol(train:test=1:1). The results are shown in Fig.6. With the increase of synthetic identities, our method can continuously improve the performance of the recognition model on large-scale palmprint dataset. The performance of our method reaches the upper bound with 120k synthetic identities and surpasses BézierPalm [16] with 3.71% in terms of TAR@FAR=1e-9 (90.18%→93.89%).

D. More Subjective Results and Comparisons

D.1. More Comparisons of Generated Palmprints

More comparisons of generated palmprints are shown in Fig.7. It can be found that the pix2pixHD [15] and CycleGAN [20] lead to blurred results, and BicycleGAN [21] may generate redundant lines marked in blue rectangle. Our method can produce clearer and more vivid palmprints. Note that our method can even restore some **realistic light and shadow effects**, such as the shadow in the root area of thumb.

D.2. Diversity and Identity Consistency of Our Method

Fig.8 shows more results of our method for the same input creases. First, it is obvious that our method can produce various and diversified palmprints. Second, these different pseudo-palmprints strictly preserve the same identity information according to the input palm creases. More generated results of different creases are provided in Fig.9, which also demonstrates our method faithfully preserve the information of input palm creases. Generally speaking, our method can yield a high degree of diversity and identity consistency at the same time.

D.3. Comparison with Recent Generation Models

We also implement some other models for palmprint generation, i.e., UGATIT [7], SpatchGAN [12], and DDIB [13] under the Open-set(train:test=1:1) protocol. Although these recent models achieve state-of-the-art natural image generation performance, these large models require abundant training samples and perform unsatisfactory for limited and fixed type palmprints. As shown in Fig.10, UGATIT and SpatchGAN lead to significant artifacts, and diffusion-based model DDIB severely changes the identity information. We think that these models have strong learning ability, but they are not suitable for current palmprint data. Therefore, we did not use these methods for detailed comparisons.

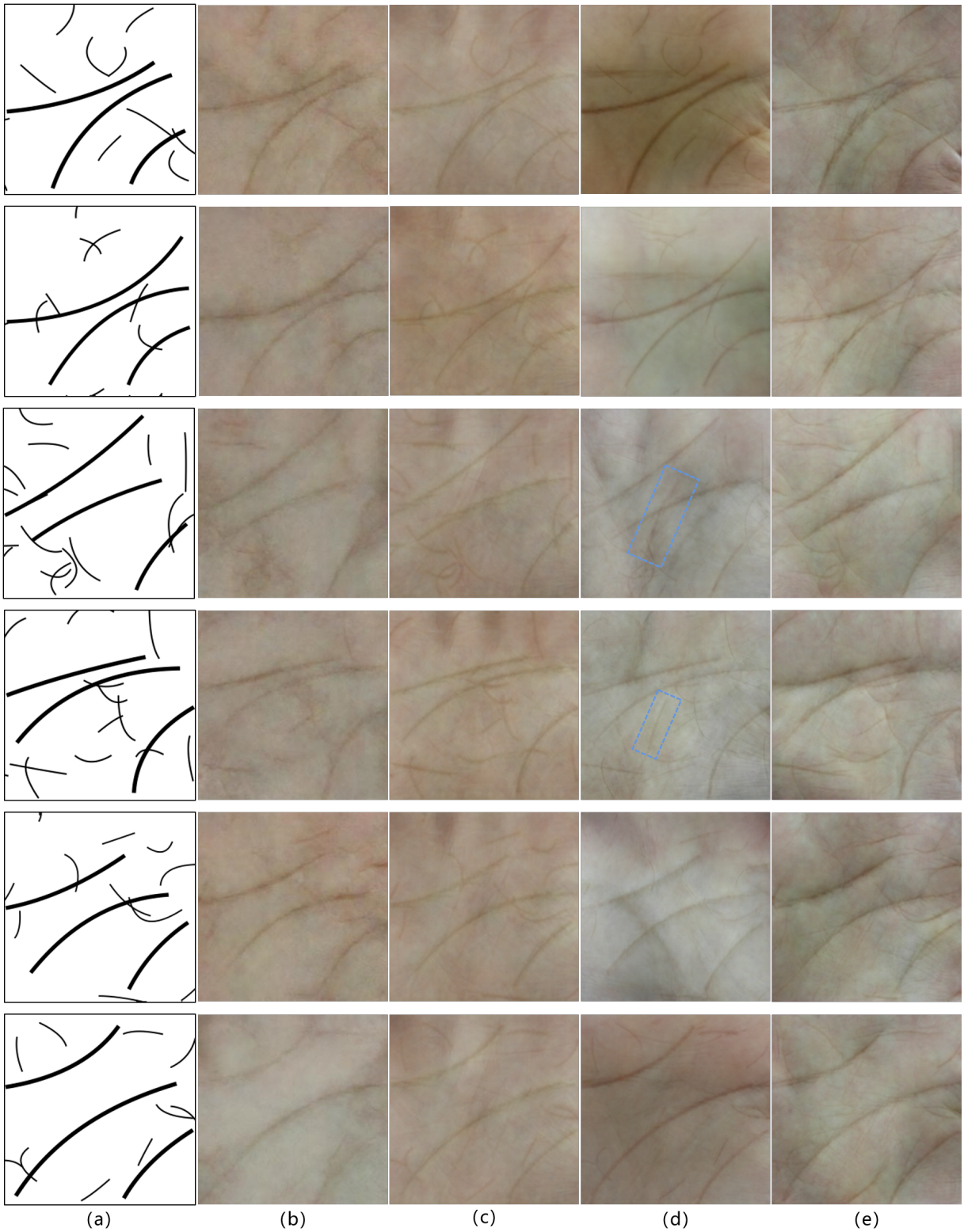

Figure 7: Generated palmprint images of different methods, (a) synthetic Bézier palm creases, (b) pix2pixHD, (c) CycleGAN, (d) BicycleGAN, (e) our method. Redundant lines are marked in blue rectangle

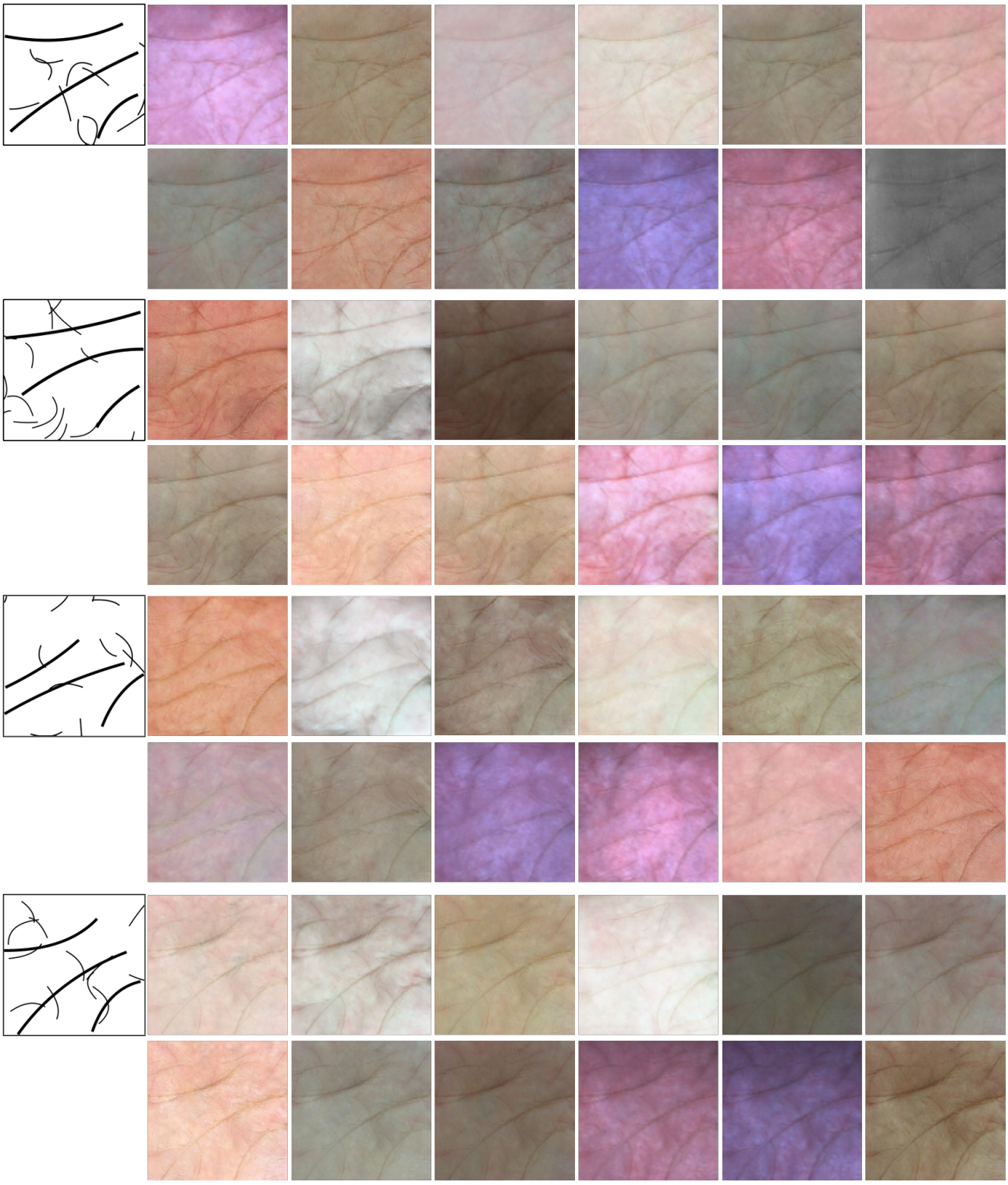

Figure 8: Our method can generate diversified palmprints for the same palm creases by using different noise vectors.

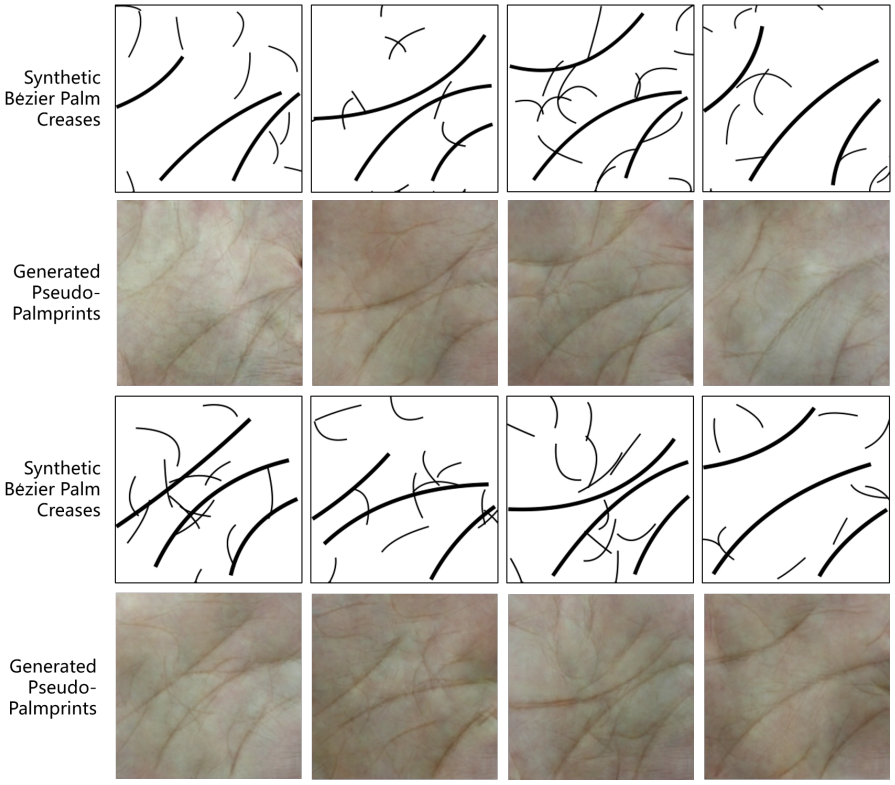

Figure 9: Generated pseudo-palmprints with different input Bézier palm creases in our method.

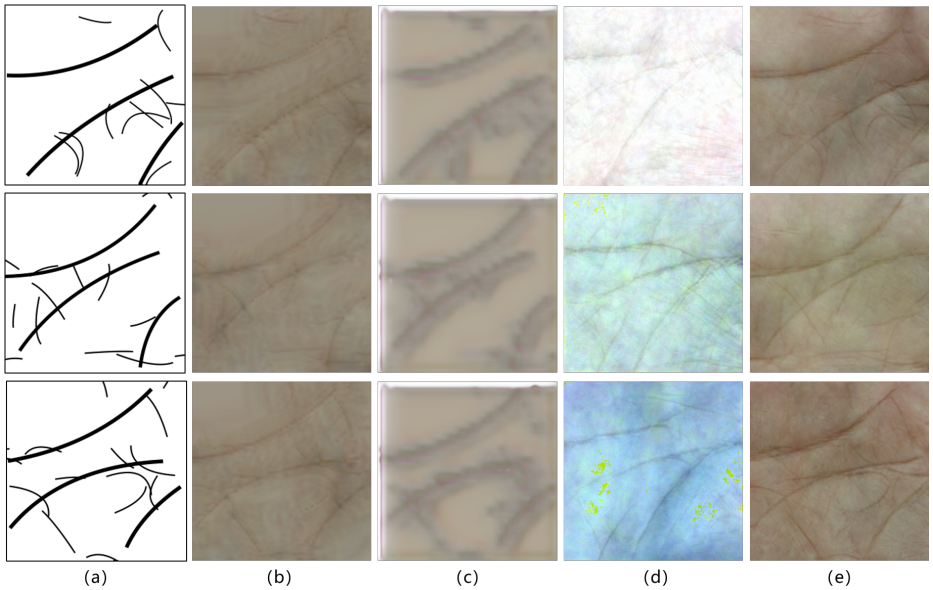

Figure 10: More comparison with recent large generation models, (a) synthetic Bézier palm creases, (b) UGATIT [7], (c) SpatchGAN [11], (d) DDIB [13], (e) our method.

References

[1] 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.

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

[3] Lunke Fei, Bob Zhang, Yong Xu, Zhenhua Guo, Jie Wen, and Wei Jia. Learning discriminant direction binary palmprint descriptor. *IEEE TIP*, 28(8):3808–3820, 2019.

[4] 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.

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

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

[7] Junho Kim, Minjae Kim, Hyeonwoo Kang, and Kwanghee Lee. U-gat-it: Unsupervised generative attentional networks with adaptive layer-instance normalization for image-to-image translation. *arXiv preprint arXiv:1907.10830*, 2019.

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

[9] Yang Liu and Ajay Kumar. Contactless palmprint identification using deeply learned residual features. *IEEE TBBIS*, 2(2):172–181, 2020.

[10] Wojciech Michal Matkowski, Tingting Chai, and Adams Wai Kin Kong. Palmprint recognition in uncontrolled and uncooperative environment. *IEEE Transactions on Information Forensics and Security*, 15:1601–1615, 2019.

[11] 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.

[12] 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.

[13] 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.

[14] 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.

[15] 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.

[16] 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.

[17] Shuping Zhao and Bob Zhang. Joint constrained least-square regression with deep convolutional feature for palmprint recognition. *IEEE TSMC*, 2020.

[18] Qian Zheng, Ajay Kumar, and Gang Pan. A 3d feature descriptor recovered from a single 2d palmprint image. *IEEE TPAMI*, 38(6):1272–1279, 2016.

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

[20] 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.

[21] Jun-Yan Zhu, Richard Zhang, Deepak Pathak, Trevor Darrell, Alexei A Efros, Oliver Wang, and Eli Shechtman. Toward multimodal image-to-image translation. *Advances in neural information processing systems*, 30, 2017.