PatchNet: A Simple Face Anti-Spoofing Framework via Fine-Grained Patch Recognition (Supplementary Material)

Chien-Yi Wang[1] Yu-Ding Lu[2] Shang-Ta Yang[1] Shang-Hong Lai[1]
[1]Microsoft AI R&D Center, Taiwan [2]HTC
{chiwa, shanya, shlai}@microsoft.com jonlu.citi@gmail.com

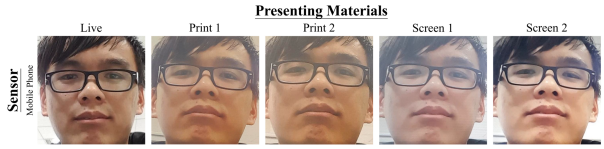

Figure 1. Fine-grained patch classes samples of Oulu-NPU [1] dataset.

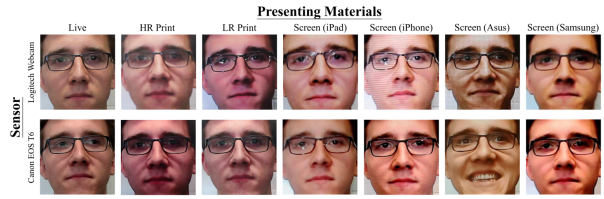

Figure 2. Fine-grained patch classes samples of SiW [6] dataset.

Figure 3. Fine-grained patch classes samples of MSU-MFSD [11] dataset.

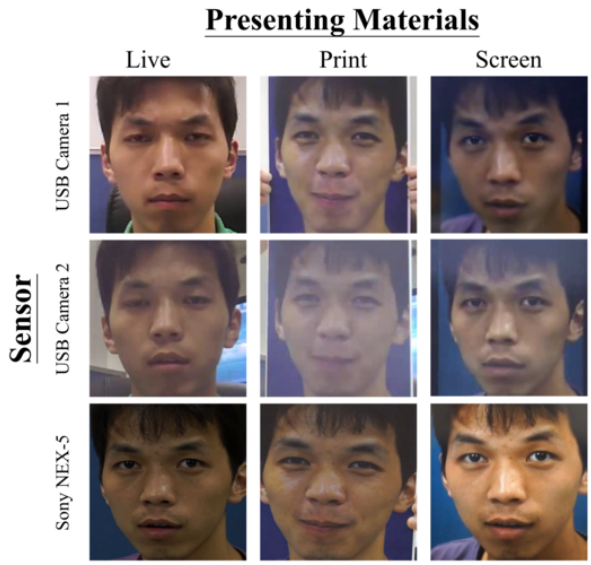

Figure 4. Fine-grained patch classes samples of CASIA-FASD [12] dataset.

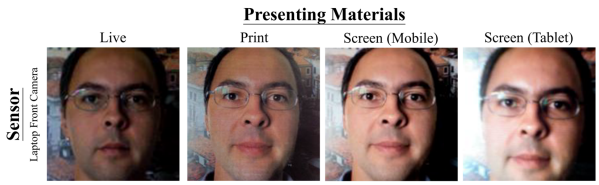

Figure 5. Fine-grained patch classes samples of ReplayAttack [2] dataset.

1. Dataset

The samples of fine-grained patch classes from O, S, M, C, and I dataset are shown in Fig. 1, 2, 3, 4, and 5, respectively. Different patch classes exhibit different capture characteristics to be modelled in the patch embedding space.

2. Ablation Study

Network Architecture. We have experimented the proposed PatchNet with different network architectures, and the results on Oulu-NPU protocol 1 are shown in Tab. 1. The model can reach competitive and SOTA performance with similar backbones like VGG11 [9], MobileNet-V2 [8], and ResNet18 [4]. However, we do not observe performance improvement with larger capacity backbones like ResNet34 [4] and ResNet50 [4]. Moreover, small backbones like AlexNet [5] and ShuffleNetV2 [7] do not have enough capacity to learn robust patch features for FAS tasks. We adopt ResNet18 to build PatchNet in all the other experiments.

Backbone	ImageNet top-5(%)	ACER(%) (Crop Patch 96)	Crop160 ACER(%) (Crop Patch 160)
AlexNet [5]	79.07	6.46	2.5
ShuffleNetV2(0.5) [7]	81.75	5.21	2.1
VGG11 [9]	89.81	1.04	0.0
MobileNetV2 [8]	90.28	1.25	0.0
ResNet18 [4]	89.08	1.04	0.0
ResNet34 [4]	91.42	1.04	0.2
ResNet50 [4]	92.86	1.04	0.0

Table 1. The ablation study on the network backbone architecture. The second column shows the top-5 accuracy on the ImageNet [3] classification task. The third and fourth columns show the ACER(%) on Oulu-NPU protocol 1 with patch crop size 96 and 160, respectively.

Method	M&C&O to I	I&C&M to O
PatchNet	95.96	95.07
PatchNet w/ 5-shot	93.29	95.67
PatchNet w/ 10-shot	94.40	95.79

Table 2. Few-shot live reference testing on **M&C&O to I** and **I&C&M to O** protocols. The testing score is averaged by 10 experiment runs. AUC(%) score is reported.

Crop Patch Num	ACER(%)
2 x 2	0.2
3 x 3	0.0
4 x 4	0.0
Center Crop	0.2

Table 3. Influence of patch numbers during testing.

Influence of Testing Strategy. As shown in Tab. 3, the cropped patch number during testing does not have much impact on the performance. Notably, even with only one center crop patch, the performance is still competitive with state-of-the-art methods. It implies that the model can make robust decisions only from local patches cropped anywhere inside the face region.

3. Few-Shot Reference Anti-Spoofing

As described in Sec. 4.7 in the main paper, we can leverage the learned patch type embedding space to perform few-shot FAS with live features as the reference. We report the experiments on M&C&O to I and I&C&M to O protocols in Tab. 2. We can achieve improved performance on I&C&M to O, both with 5-shot and 10-shot live samples as the reference. However, for M&C&O to I protocol, the few-shot performance is inferior to the original one. We suspect that it is due to the low sensor resolution of I dataset, which results in the difficulty for modeling the intrinsic cues from the captures.

4. Visualizations

The t-SNE [10] visualizations of domain generalization protocols O&C&I to M, I&C&M to O, and M&C&O to I are shown in Fig. 6, 7, and 8.

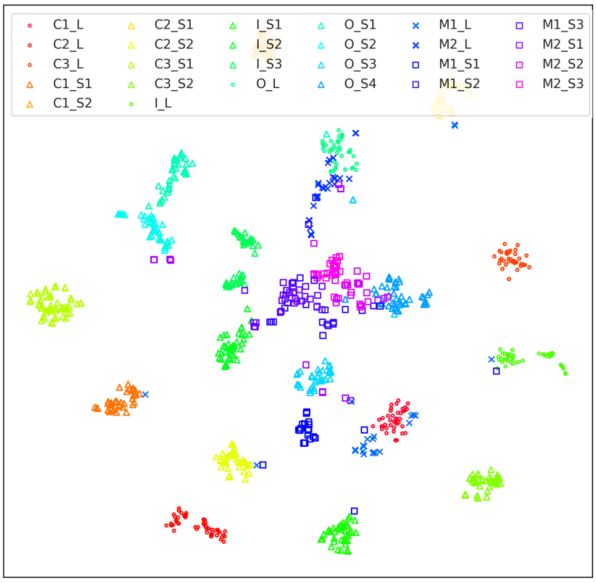

Figure 6. The t-SNE visualizations of normalized patch features in cross-dataset protocol O&C&I to M. The fine-grained patch class is denoted by (Dataset)(SensorID)_(Liveness)(MediumID).

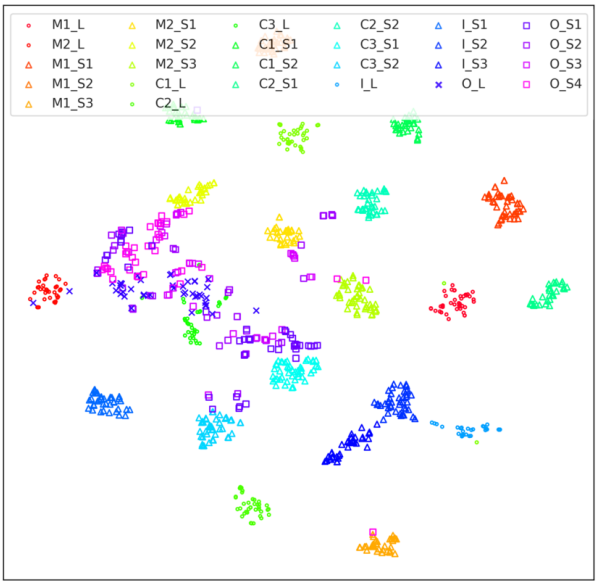

Figure 7. The t-SNE visualizations of normalized patch features in cross-dataset protocol I&C&M to O. The fine-grained patch class is denoted by (Dataset)(SensorID)_(Liveness)(MediumID).

References

[1] Z. Boulkenafet, J. Komulainen, L. Li, X. Feng, and A. Hadid. Oulu-npu: A mobile face presentation attack database with real-world variations. In *IEEE International Conference on Automatic Face Gesture Recognition (FG)*, 2017. 1

[2] Ivana Chingovska, André Anjos, and Sébastien Marcel. On the effectiveness of local binary patterns in face anti-spoofing. In *BIOSIG-proceedings of the international conference of biometrics special interest group (BIOSIG)*. IEEE, 2012. 1

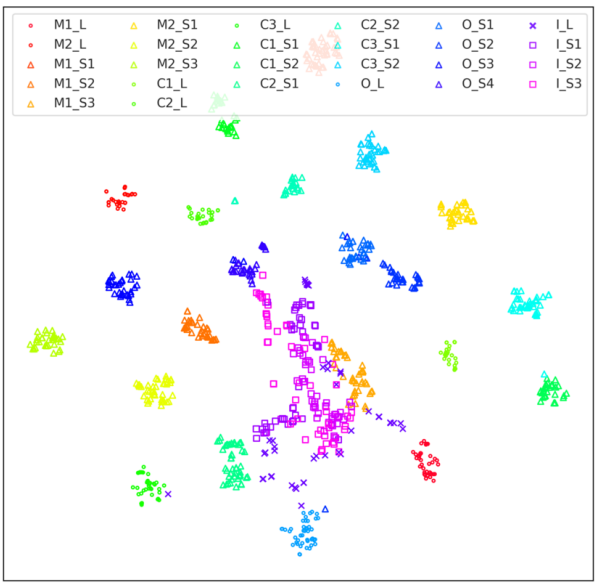

Figure 8. The t-SNE visualizations of normalized patch features in cross-dataset protocol M&C&O to I. The fine-grained patch class is denoted by (Dataset)(SensorID)_(Liveness)(MediumID).

[3] Jia Deng, Wei Dong, Richard Socher, Li-Jia Li, Kai Li, and Li Fei-Fei. Imagenet: A large-scale hierarchical image database. In *CVPR*, 2009. 2

[4] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *CVPR*, 2016. 1, 2

[5] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. *NIPS*, 2012. 1, 2

[6] Yaojie Liu, Amin Jourabloo, and Xiaoming Liu. Learning deep models for face anti-spoofing: Binary or auxiliary supervision. In *CVPR*, 2018. 1

[7] Ningning Ma, Xiangyu Zhang, Hai-Tao Zheng, and Jian Sun. Shufflenet v2: Practical guidelines for efficient cnn architecture design. In *ECCV*, 2018. 1, 2

[8] Mark Sandler, Andrew Howard, Menglong Zhu, Andrey Zhmoginov, and Liang-Chieh Chen. Mobilenetv2: Inverted residuals and linear bottlenecks. In *CVPR*, 2018. 1, 2

[9] Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*, 2014. 1, 2

[10] Laurens Van der Maaten and Geoffrey Hinton. Visualizing data using t-sne. *Journal of machine learning research*, 2008. 2

[11] D. Wen, H. Han, and A. K. Jain. Face spoof detection with image distortion analysis. *IEEE Transactions on Information Forensics and Security*, 2015. 1

[12] Z. Zhang, J. Yan, S. Liu, Z. Lei, D. Yi, and S. Z. Li. A face antispoofing database with diverse attacks. In *IAPR International Conference on Biometrics (ICB)*, 2012. 1