Supplementary Material for Regularized Fine-grained Meta Face Anti-spoofing

Rui Shao, Xiangyuan Lan, Pong C. Yuen
Department of Computer Science, Hong Kong Baptist University, Hong Kong
{ruishao, pcyuen}@comp.hkbu.edu.hk, xiangyuanlan@life.hkbu.edu.hk

Datasets

Table 1: Comparison of four experimental datasets.

Dataset	Light variation	Complex background	Attack type	Display devices
C	No	Yes	Printed photo Cut photo Replayed video	iPad
I	Yes	Yes	Printed photo Display photo Replayed video	iPhone 3GS iPad
M	No	Yes	Printed photo Replayed video	iPad Air iPhone 5S
O	Yes	No	Printed photo Display photo Replayed video	Dell 1905FP Macbook Retina

The evaluation of our method is conducted on four public face anti-spoofing datasets that contain both print and video replay attacks: Oulu-NPU (Boulkenafet and et al 2017) (O for short), CASIA-MFSD (Zhang and et al 2012) (C for short), Idiap Replay-Attack (Chingovska, Anjos, and Marcel 2012) (I for short), and MSU-MFSD (Wen, Han, and Jain 2015) (M for short). From Table 1 and Fig. 1, it can be seen that many kinds of variations, due to the differences on materials, illumination, background, resolution and so on, exist across these four datasets. Therefore, significant domain shift exists among these datasets.

Network Structure

The detailed structure of the proposed network is illustrated in Table 2. To be specific, each convolutional layer in the feature extractor, meta learner and depth estimator is followed by a batch normalization layer and a rectified linear unit (ReLU) activation function, and all convolutional kernel size is 3×3. The size of input image is $256 \times 256 \times 6$, where we extract the RGB and HSV channels of each input image. Inspired by the residual network (He et al. 2016), we use a short-cut connection, which is concatenating the responses of pool1-1, pool1-2 and pool1-3, and sending them to conv3-1 for depth estimation. This operation helps to ease the training procedure.

References

Boulkenafet, Z., and et al. 2017. Oulu-npu: A mobile face presentation attack database with real-world variations. In *FG*.

Chingovska, I.; Anjos, A.; and Marcel, S. 2012. On the effectiveness of local binary patterns in face anti-spoofing. In *BIOSIG*.

He, K.; Zhang, X.; Ren, S.; and Sun, J. 2016. Deep residual learning for image recognition. In *CVPR*.

Wen, D.; Han, H.; and Jain, A. K. 2015. Face spoof detection with image distortion analysis. In *IEEE Trans. Inf. Forens. Security, 10(4): 746-761*.

Zhang, Z., and et al. 2012. A face antispoofing database with diverse attacks. In *ICB*.

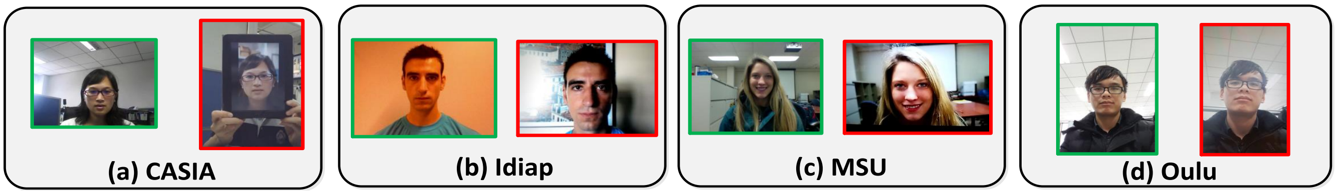

Figure 1: Sample frames from CASIA-MFSD (Zhang and et al 2012), Idiap Replay-Attack (Chingovska, Anjos, and Marcel 2012), MSU-MFSD (Wen, Han, and Jain 2015), and Oulu-NPU (Boulkenafet and et al 2017) datasets. The figures with green border represent the real faces, while the ones with red border represent the video replay attacks. From these examples, it can be seen that large cross-dataset variations due to the differences on materials, illumination, background, resolution and so on, cause significant domain shift among these datasets.

Table 2: The structure details of all components of the proposed network.

Feature Extractor			**Meta Learner**			**Depth Estimator**		
Layer	Chan./Stri.	Out.Size	Layer	Chan./Stri.	Outp.Size	Layer	Chan./Stri.	Outp.Size
Input image			Input pool1-3			Input pool1-1+pool1-2+pool1-3		
conv1-1	64/1	256	conv2-1	128/1	32	conv3-1	128/1	32
conv1-2	128/1	256	pool2-1	-/2	16	conv3-2	64/1	32
conv1-3	196/1	256	conv2-2	256/1	16	conv3-3	1/1	32
conv1-4	128/1	256	pool2-2	-/2	8			
pool1-1	-/2	128	conv2-2	512/1	8			
conv1-5	128/1	128	Average pooling					
conv1-6	196/1	128	fc2-1	1/1	1			
conv1-7	128/1	128						
pool1-2	-/2	64						
conv1-8	128/1	64						
conv1-9	196/1	64						
conv1-10	128/1	64						
pool1-3	-/2	32						