A Cross Domain Multi-modal Dataset for Robust Face Anti-spoofing

Qiaobin Ji†,Shugong Xu†,Xudong Chen†,Shunqing Zhang†, Shan Cao†
† Shanghai Institute for Advanced Communication and Data Science,
Shanghai University, Shanghai, 200444, China
Email:{qiaobin, shugong, xudongchen, shunqing, cshan}@shu.edu.cn

Speaker: Xudong Chen

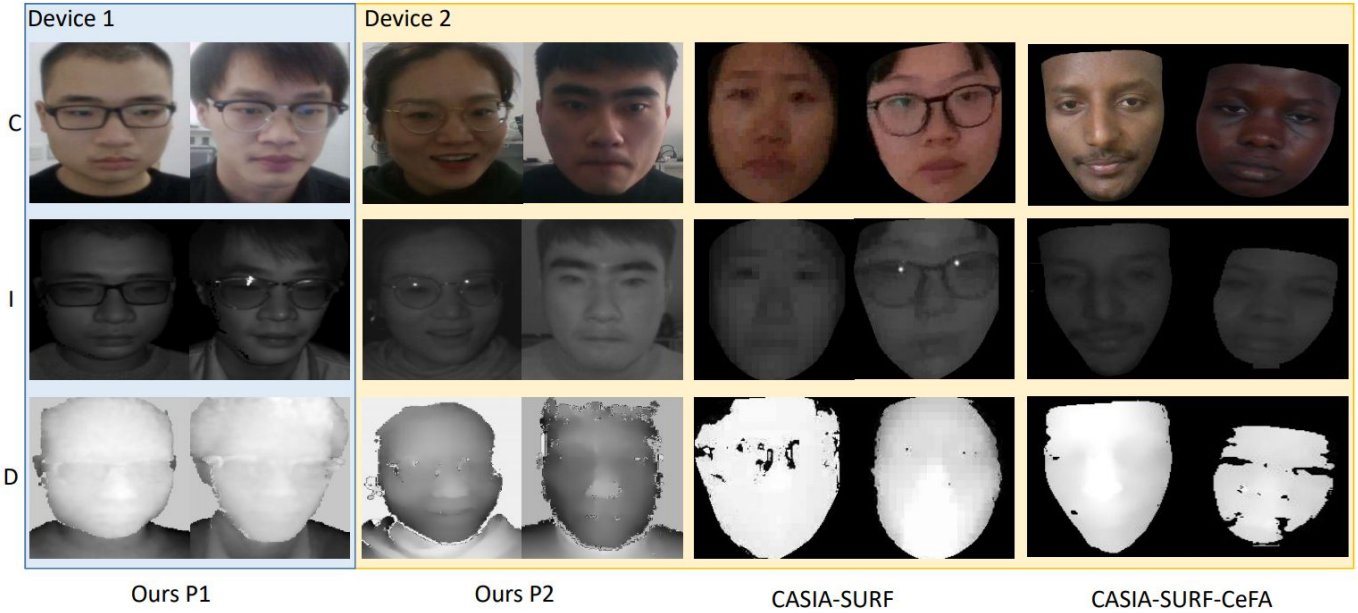
Problems

1. Multi-modal FAS datasets rarely pay attention to the cross domain problems.
2. Previous multi-modal dataset preprocessing method requires a lot of computing resources and causes the grid effect and depth holes.

Contribution

1. A cross-device face anti-spoofing dataset is proposed.
2. Two depth map preprocessing and normalization methods are proposed.
3. A novel unified multi-head end-to-end convolutional neural network architecture for face anti-spoofing is proposed.

Dataset



1. The dataset has 4 human species, 4 attack types, and the age range is from 20s -50s, two different depth sensor principle multi-modal cameras are included.
2. Our preprocessing only uses the face detector. It can recover the grid effect and noise caused holes, and can keep more details in the quantification process.

Model

A combination of position-wise attention and residual block is used.

Loss function

$$L_s = -\sum_{i=1}^{m} log \frac{e^{W_{y_i}^T xi + b_{y_i}}}{\sum_{j=1}^{n} e^{W_j^T xi + b_j}}, \quad (1)$$

$$L_c = \frac{1}{2} \sum_{i=1}^{m} \| x_i - c_{y_i} \|_2^2, \quad (2)$$

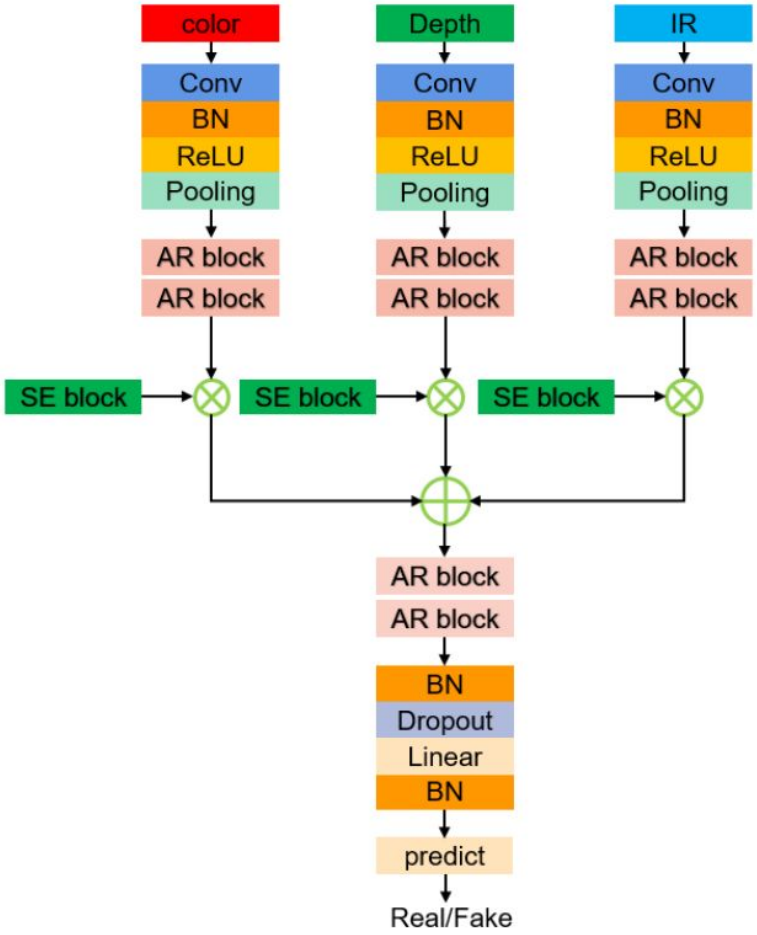


Fig. 4. Attention based anti-spoofing network(AASN) architecture

Experiments

TABLE III
PERFORMANCE OF AASN ON GREAT-FASD'S THREE DIFFERENT PROTOCOLS.

Method	simple	medium	hard
Res18-SE [28]	98.08%	93.23%	67.17%
Res18 [28]	97.75%	91.92%	63.41%
FaceBagNet [21]	98.00%	94.31%	67.02%
VisionLabs [17]	**98.69%**	95.56%	69.13%
ours	98.06%	**98.39%**	**72.50%**

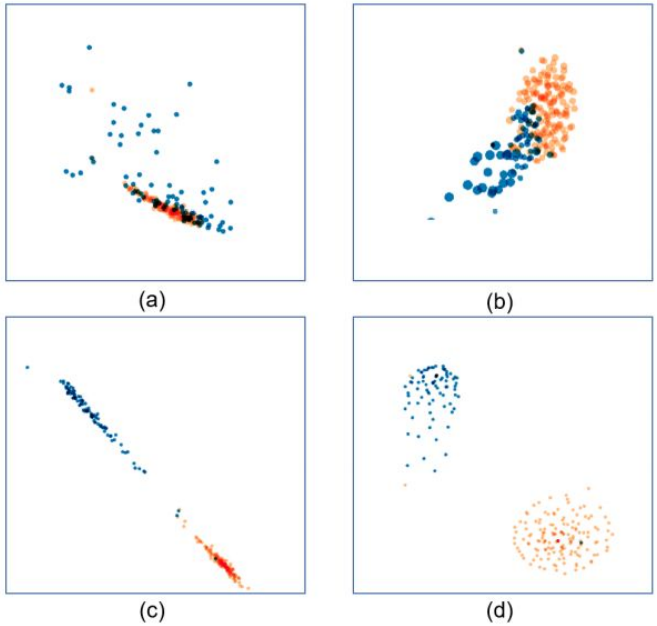


Fig. 6. The 2D visualization of spoof faces and real faces on baseline model and ours AASN. The first row is dimension-reduced 2D face features extracted by baseline and second row is extracted by AASN. The first column and second column is dimension-reduced by PCA and t-SNE [14] respectively. Please zoom in to check details. Best viewed in color.

Experiments

Model	attenion	shifting	center loss	accuracy	BPCER	APCER	ACER
baseline [28]				93.23%	0.0102	0.1297	0.0699
+a	✓			96.15%	**0.0014**	0.0765	0.0389
+a+msa	✓	✓		97.04%	0.0192	0.0547	0.0369
+a+c	✓		✓	97.35%	0.0089	0.0157	0.0124
Ours	✓	✓	✓	**98.39%**	0.0062	**0.0024**	**0.0043**

TABLE II

PERFORMANCE OF THE BASELINE MODEL TRAINED WITH AND WITHOUT ATTENTION RESIDUAL BLOCK, CENTER LOSS AND MODAL SHIFTING DATA AUGMENTATION ON THE GREAT-FASD DATASET UNDER MEDIUM PROTOCOL. "A" DENOTES ATTENTION,"C" DENOTES CENTER LOSS EQ(2), AND "MSA" MEANS MODAL SHIFTING AUGMENTATION.

During the training process, there is a 20% probability that two pictures of the same modality in the minibatch will be randomly exchanged.

Thanks for listening.