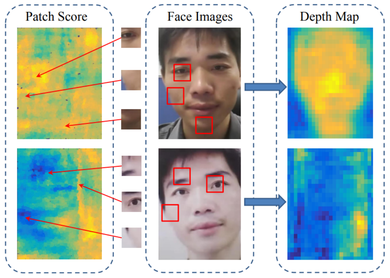
Face Anti-Spoofing Using Patch and Depth-Based CNNs

Yousef Atoum*, Yaojie Liu*, Amin Jourabloo*, Xiaoming Liu
Department of Computer Science and Engineering
Michigan State University

Computer Vision Lab

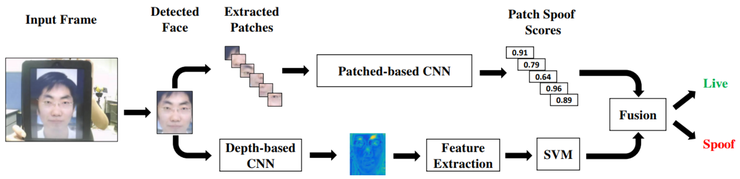
Introduction

- Face image is the most accessible biometric modality.
- Face anti-spoofing is a very critical step before face recognition.
- We utilize:
  - Local feature → Independent of spatial face areas
  - Holistic feature → Face-like depth



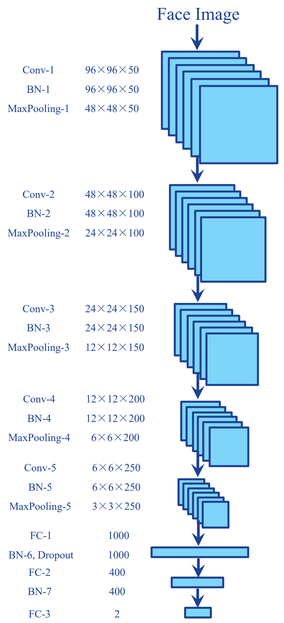
Proposed Methods

- The architecture of two-stream CNN-based face anti-spoofing method.
- Each stream is trained end-to-end independently.
- The fusion of scores provides the final decision.

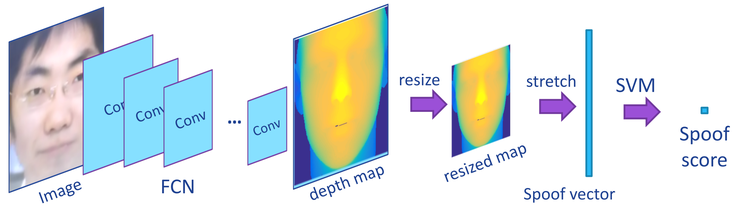


Patch-Based CNN

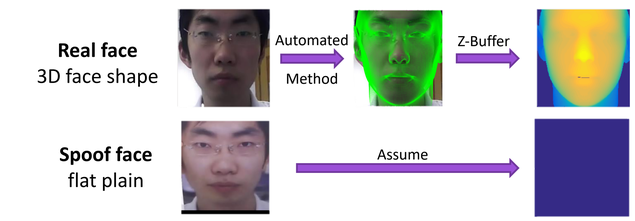
- Feed the CNN with extracted patches
- Independent of their spatial location
- Advantages of patch based approach:
  - Increase number of training samples
  - Maintain the native resolution
  - Discover spoof-specific information
- Size of extracted patches are 96×96
- Select 10 patches randomly from face area
- Each convolutional layer is followed by batch normalization and RELU layers.
- Different types of features can be input to the network (HSV+YCbCr).



Depth-Based CNN



- Depth labels for real and spoof faces:



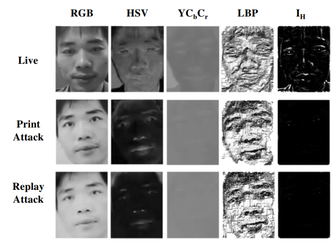
- The loss function is pixel-level Euclidean loss:

$$\arg\min_{\Theta} J = \|f(\mathbf{I};\Theta) - \mathbf{M}\|_F^2$$

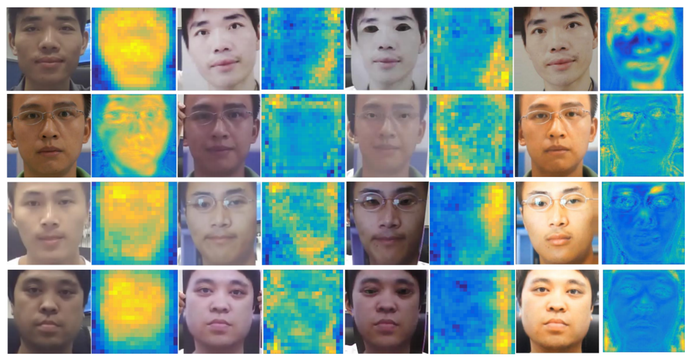
Ablation study

- We explore different input feature maps to train the patch-based CNN.
- The HSV+YCbCr achieves the highest performance.

Feature	EER (%)	HTER (%)
YC_bC_r	4.82	3.95
$YC_bC_r + HSV$	**4.44**	**3.78**
$YC_bC_r + HSV + LBP$	7.72	6.09
$(YC_bC_r + HSV)_H$	9.58	5.57



- The depth estimation on CASIA-FASD testing subjects.



Live image | Print attack | Cut print attack | Video replay attack

Experimental Comparison

- Replay-Attack dataset:

Contains replay, print and cut-print attacks.

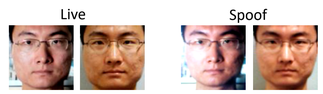
Method	EER (%)	HTER (%)
DPCNN	2.90	6.10
Yang *et al.*	2.14	-
Boulkenafet *et al.*	**0.10**	2.20
Moire pattern	-	3.30
Our patch-based CNN	2.50	1.25
Our depth-based CNN	0.86	0.75
Our fusion	0.79	**0.72**

- MSU-USSA dataset:

Contains eight types of print attacks.

Method	EER (%)	HTER (%)
Patel *et al.*	3.84	-
Our patch-based CNN	0.55 ± 0.26	0.41 ± 0.32
Our depth-based CNN	2.62 ± 0.73	2.22 ± 0.66
Our fusion	$\mathbf{0.35} \pm 0.19$	$\mathbf{0.21} \pm 0.21$

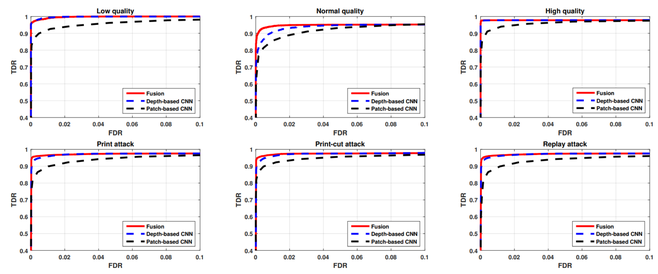
- CASIA-FASD dataset:



Contains two types of replay attacks.

Method	EER (%)	HTER (%)
DPCNN	4.50	-
LSTM-CNN	5.17	5.93
Boulkenafet *et al.*	2.80	-
Moire pattern	-	**0**
Our patch-based CNN	4.44	3.78
Our depth-based CNN	2.85	2.52
Our fusion	**2.67**	2.27

- Frame-based ROC curves on CASIA-FASD:



Conclusions

- Introduce a novel face anti-spoofing method based on fusing two CNN streams.
- We leverage both the full face image and patches extracted from the same face to distinguish the spoof from live faces.

Acknowledgment

- This research is based upon work supported by the Office of the Director of National Intelligence (ODNI), Intelligence Advanced Research Projects Activity (IARPA), via IARPA R&D Contract No. 2014-14071600012.

*denotes equal contribution by the authors.