

# Face Recognition System Security: Anti-spoofing and Template Protection

Pong C Yuen  
Professor  
Department of Computer Science  
Hong Kong Baptist University

# Thanks...

- Collaborators:
  - Anil Jain, Michigan State University
  - G Y Zhao, University of Oulu
  - M H Lim, former Postdoc
  - Y C Feng, former PhD student
  - G C Mai, PhD student
  - S Q Liu, PhD student
  - R Shao, PhD student
- Funding Agents
  - Hong Kong Research Grant Council
  - Hong Kong Baptist University

# Outline

1. Background and Motivations
2. Face Anti-spoofing
3. Face Template Protection
4. Conclusions

# Biometrics

## ■ Deployed practical applications



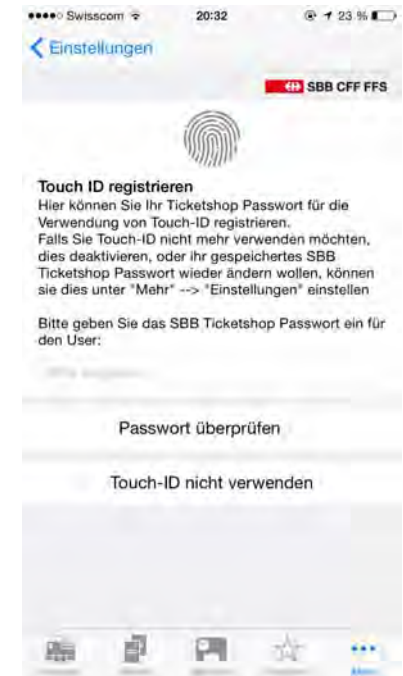
Border Control




Door Access Control



Touch ID (iPhone)



SBB for buying ticket 

# Face Biometrics

## ■ Face Recognition Technology

### Jack Ma's first unmanned supermarket

Today, on a street in Hangzhou (Zhejiang province), Jack Ma's first unmanned supermarket officially opened for business. Because there are no costs for manpower, the expenses for running the unmanned supermarket only add up to about a quarter of those of traditional supermarkets. The shop owner just needs to replenish the inventories every morning - nothing else needs to be done.



Entrance to the unmanned supermarket



face-recognition payment Alipay



World's first' facial recognition ATM unveiled in China

Monday, 31 May 2015, 6:38am  
Monday, 01 June, 2015, 11:31am

COMMENTS: 2



Source: china.com and iomniscient.com

# Face Biometrics



Photo: Bloomberg

Face ID is enabled by the TrueDepth camera and is simple to set up. It projects and analyzes more than 30,000 invisible dots to create a precise depth map of your face.

## FaceID in iPhone X

Announced on 12 September 2017

“With a simple glance, Face ID securely unlocks your iPhone X. You can use it to **authorize purchases from the iTunes Store, App Store, iBooks Store, and payments with Apple Pay.** Developers can also allow you to use Face ID to sign into their apps. ....”

## 3D Face Recognition:


Employed Structured-light 3D technology

Your face is your  
secure password.



With Face ID, iPhone X unlocks only when you're looking at it. It's designed to resist spoofing by photos or masks. Your facial map is encrypted and protected by the Secure Enclave. And authentication happens instantly on the device, not in the cloud.



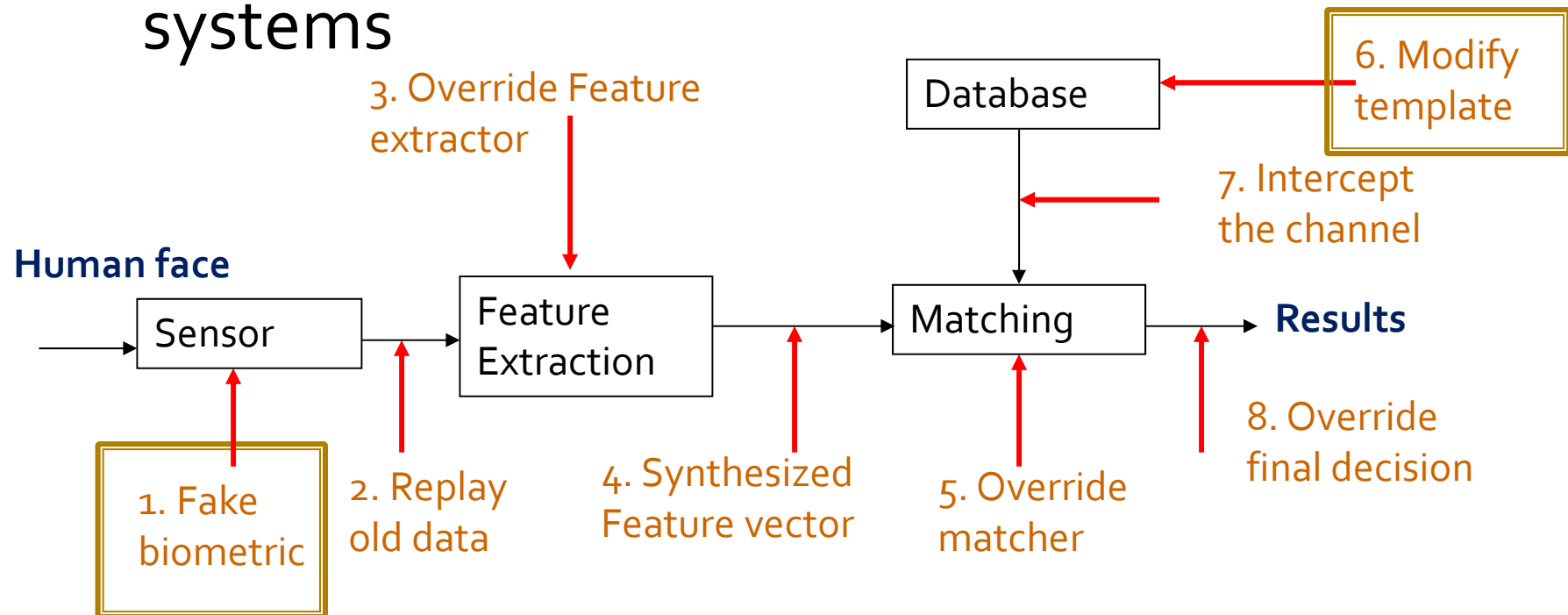


**What happens if  
a face recognition system is NOT secure?**



# Background and Motivations

- Vulnerabilities: Ratha *et al.* [IBM Sys J 2001] pointed out eight possible attacks on biometric systems





# Part I: Face Anti-Spoofing



## Mission Impossible - Rogue Nation (2015): Biometric Spoofing

# Outline: Face Anti-spoofing

1. Background and Motivations
2. Related Work
3. rPPG Approach
4. Deep Learning Approach
5. Conclusions

# Background and Motivations

- Face Spoofing Attack
  - With rapid development of social network such as Facebook and Twitter, face information can be easily acquired (facebook, twitter) and abused



✓ Real Face



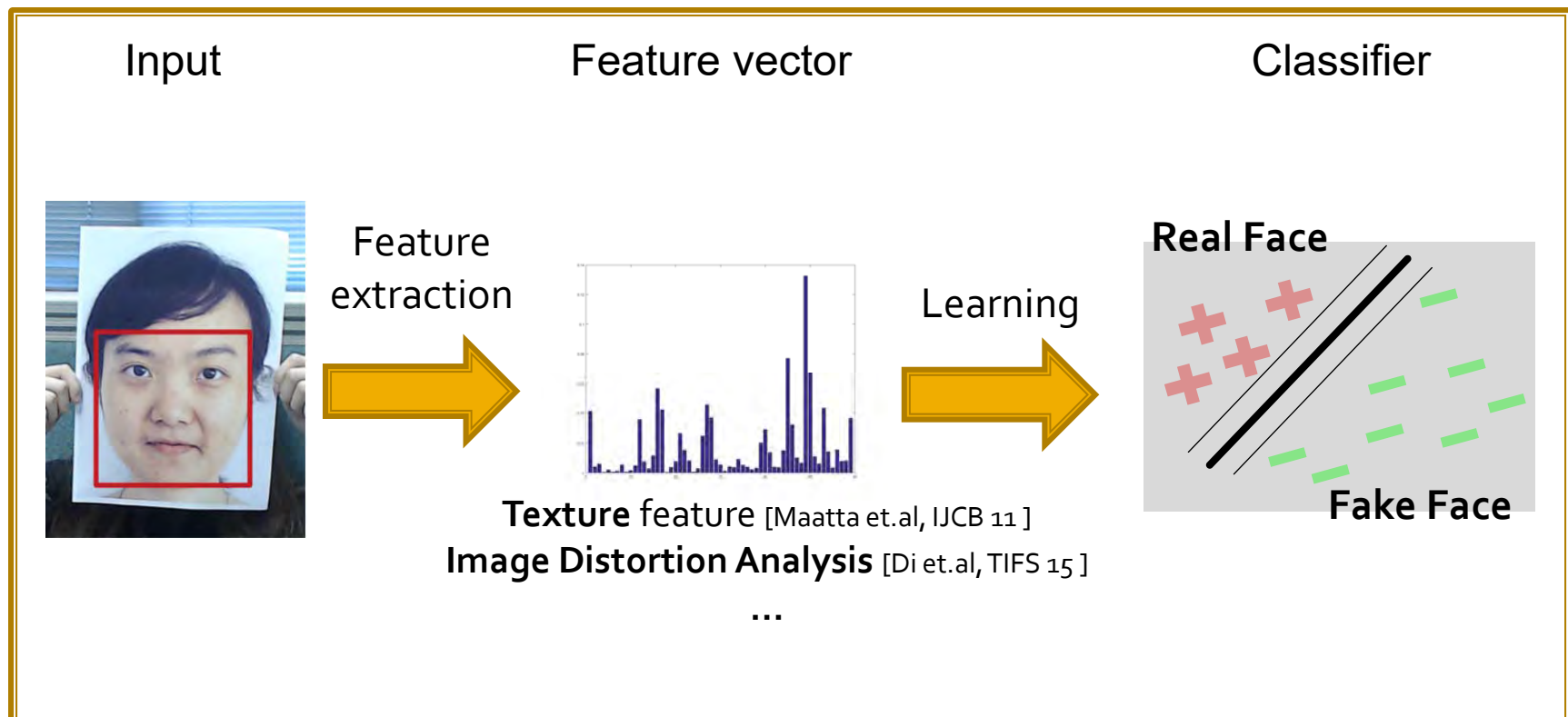
✗ Prints Attack



✗ Replay Attack

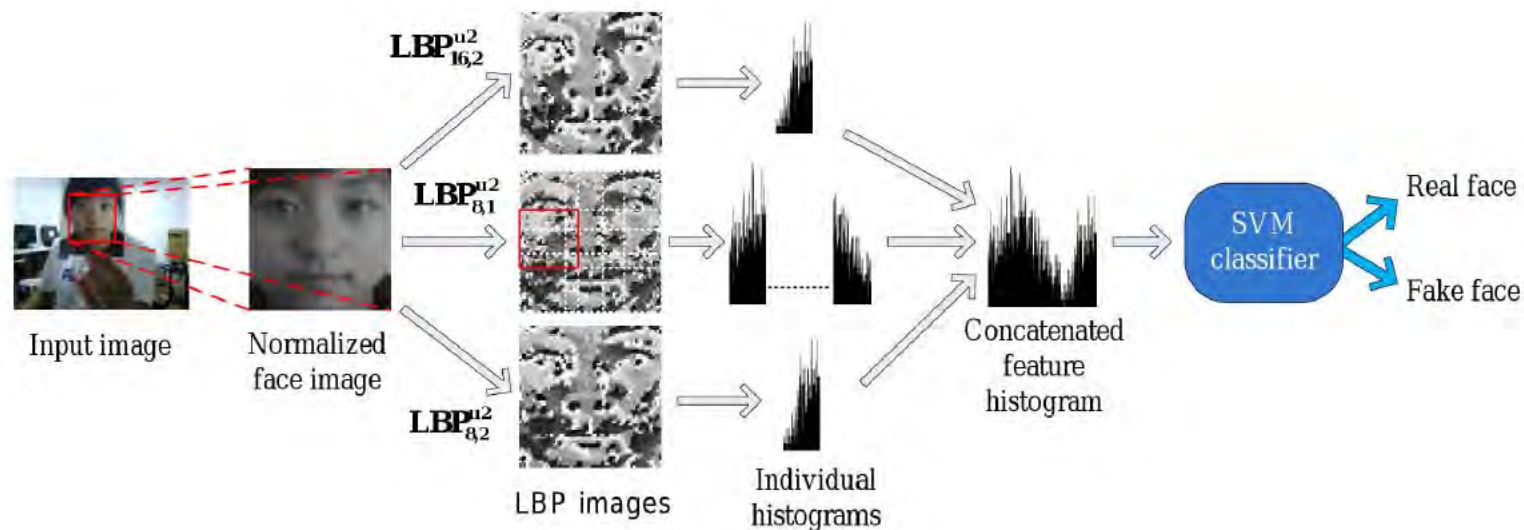
# Background and Motivations

- Anti-spoofing approach: Appearance-based
  - Spoof media (print and screen) and genuine face has different appearance



# Background and Motivations

- Anti-spoofing approach: Appearance-based
  - Spoof media (Prints and screen) has different texture, comparing with genuine face

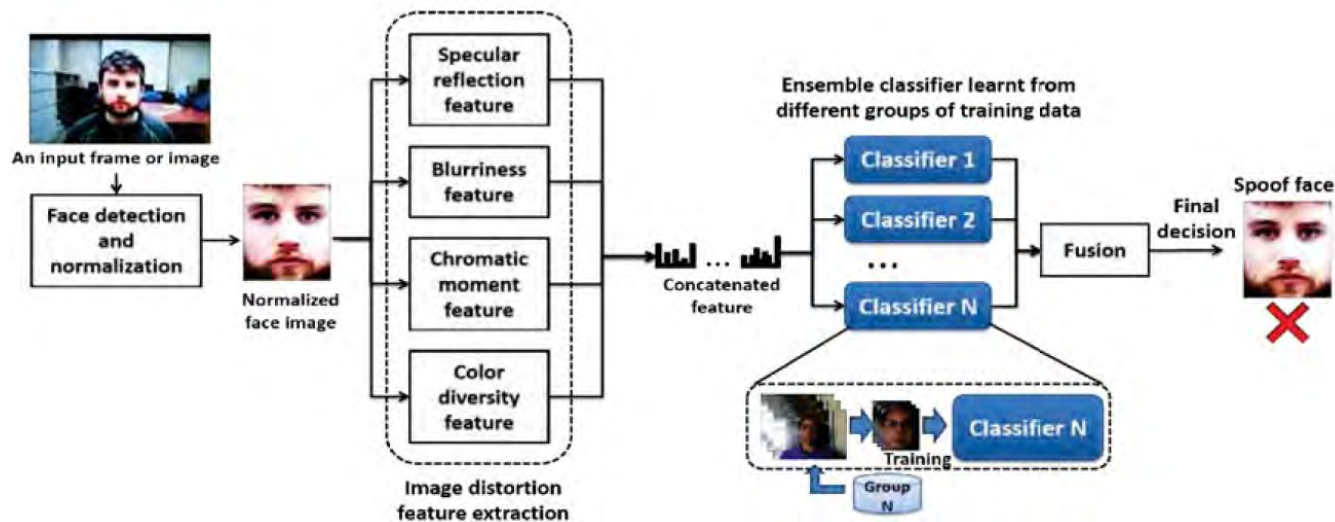


**Source:** Jukka Maatta, Abdenour Hadid, Matti Pietikainen, "Face Spoofing Detection From Single Images Using Micro-Texture Analysis", *IJCB* 2011



# Background and Motivations

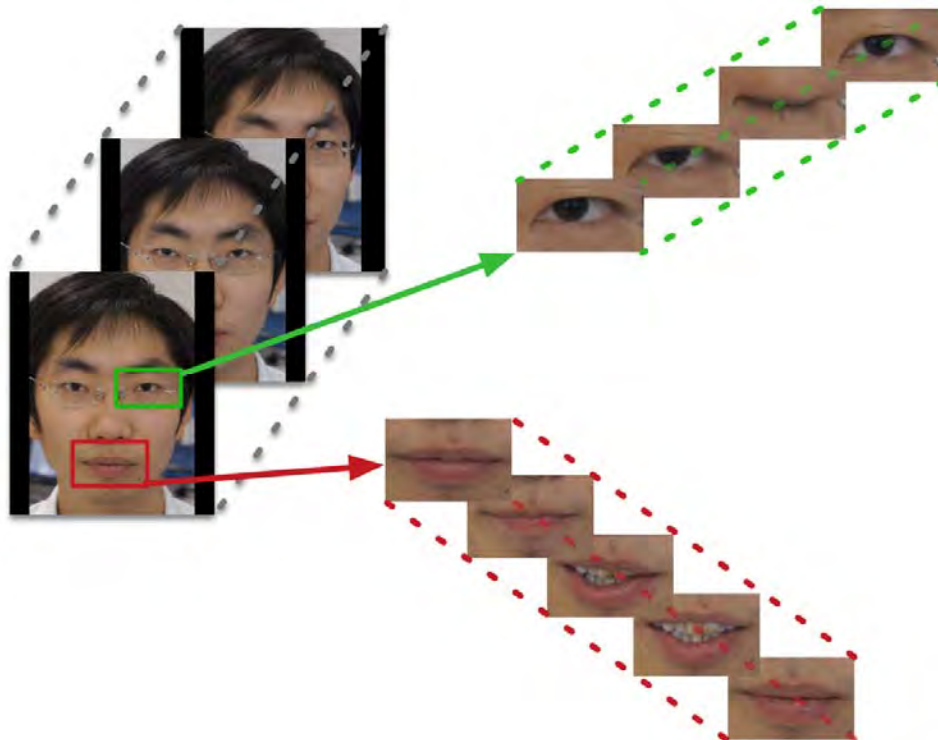
- Anti-spoofing approach: Appearance-based
  - Spoof media (prints and screen) has specific quality defects



**Source:** Di Wen, Hu Han, Anil K. Jain, "Face Spoof Detection with Image Distortion Analysis", *TIFS* 2015

# Background and Motivations

- Anti-spoofing approach: Motion-based
  - 2D spoofing medium cannot move, or has different motion pattern compare with real face





# Background and Motivations

- Anti-spoofing approach: Motion-based
  - **Eyeblick-based** anti-spoofing in face recognition from a generic web-camera (G.Pan et al., ICCV'07)
  - Real-time face detection and **motion analysis** with application in liveness assessment. (K. Kollreider et al., TIFS'07)
  - A liveness detection method for face recognition based on **optical flow field** (W. Bao et al., IASP'09)
  - Face liveness detection using **dynamic texture** (Pereira et al., JIVP'14)
  - Detection of face spoofing using **visual dynamics** (S. Tirunagari et al., TIFS'15)

# Background and Motivations

- Performance on traditional face spoofing attack

<i>Pipelines</i>	<b>Replay Attack</b>		<b>Print attack</b>	
	<i>Dev</i>	<i>Test</i>	<i>Dev</i>	<i>Test</i>
DMD+SVM (face region)	8.50	7.50	0.00	0.00
DMD+LBP+SVM (face region)	5.33	3.75	0.00	0.00
PCA+SVM (face region)	20.00	21.50	16.25	15.11
PCA+LBP (face region)	11.67	17.50	9.50	5.11
DMD+LBP+SVM (entire video)	0.50	0.00	0.00	0.00
PCA+LBP+SVM (entire video)	21.75	20.50	11.50	9.50

[S. Tirunagari et al., TIFS'15]

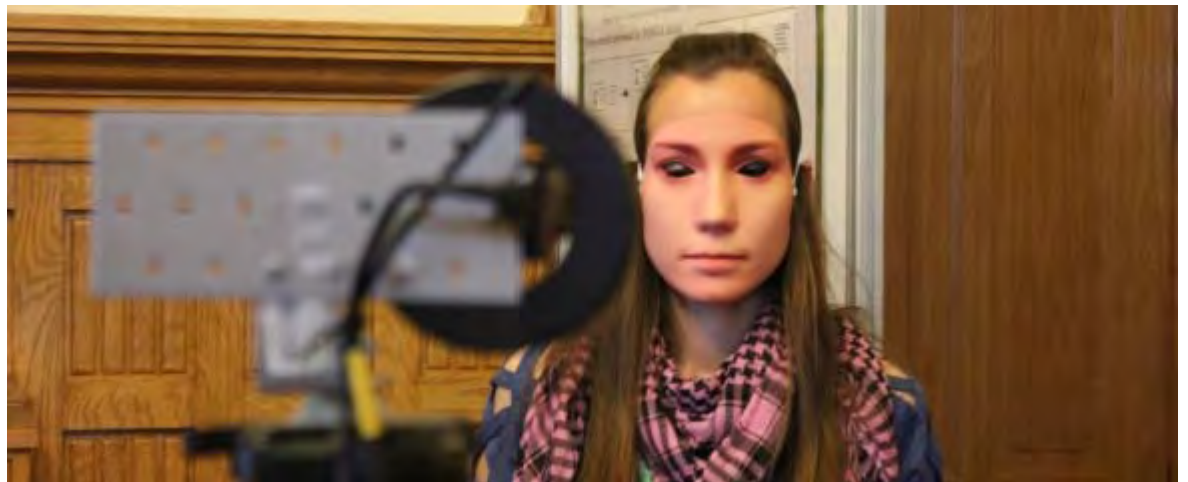
Promising results are achieved on tradition face spoofing attack

# Background and Motivations

**Problem solved?**

# Background and Motivations

- **New Challenge:** 3D Mask Attack
  - With the advanced development on 3D reconstruction and 3D printing technology, 3D face model can easily be constructed and used to spoof recognition systems



Source: [idiap.ch](http://idiap.ch)

Mask is made from [ThatsMyFace.com](http://ThatsMyFace.com)

# Background and Motivations

- **New Challenge:** Super-realistic 3D Mask
  - 3D mask can be so real that we can hardly differentiate them from appearance



(a)

Life face

(b)

Real-F hyper real mask



# Background and Motivations

## Hong Kong airport security fooled by these hyper-real silicon masks

Masks like the one that transformed a Chinese kid into a U.S. grandpa are now available online.  
By Joe Li (// author/see-hong-kong-edited) 8 November, 2010

Suspicious old folks: the Elder Mask from SPFX Masks is so real.



Before...

That Chinese guy who disguised himself as an old white man to slip by Hong Kong airport security and board an Air Canada flight might have ordered his old man mask from [SPFX Masks](http://www.spfxmasks.com/) (<http://www.spfxmasks.com/>).

This is the stuff that entered popular imagination with the [Mission Impossible television series](http://www.youtube.com/watch?v=b6gqllhLE3c&NR=1) (<http://www.youtube.com/watch?v=b6gqllhLE3c&NR=1>) and is used by the CIA (<http://abcnews.go.com/Health/Cosmetic/story?id=1354130>) and as prosthetics for medical conditions (<http://www.prosthetics.com/services.html>).

Now we can order our own so-real-its-creepy mask online.

Silicon masks from SPFX adhere to facial features such that the mask is able to move with the musculature of the wearer, like a second skin. The mask is attached to a neck flap and some come with silicon gloves to disguise the hands and forearms as well.

is attached to a neck flap and some come with silicon gloves to disguise the hands and forearms as well.

Check out the video above of a demonstration of the [Elder Mask](http://www.spfxmasks.com/maskelder.html) (<http://www.spfxmasks.com/maskelder.html>) from SPFX, which resembles the one that Chinese stowaway was caught with in Canada. Priced at US\$689, the mask is aimed at Halloween revelers and haunted house actors.



... and after.

But the passenger who breached Hong Kong airport security on October 29 used his mask to smuggle himself into Canada.

The Chinese man who appeared to be in his early 20s disguised himself as an elderly Caucasian man, obtained a boarding pass from a U.S. citizen while in transit in Hong Kong, and boarded the Air Canada flight using an Aeroplan card for identification.

Read more details about the case from the confidential alert obtained by CNN ([http://articles.cnn.com/2010-11-04/world/canada.disguised.passenger.1\\_flight\\_crew\\_hong\\_kong\\_regional-communications-officer?\\_s=PM:WORLD](http://articles.cnn.com/2010-11-04/world/canada.disguised.passenger.1_flight_crew_hong_kong_regional-communications-officer?_s=PM:WORLD)).

Source: <http://travel.cnn.com/hong-kong/visit/hong-kong-airport-security-fooled-these-hyper-real-silicon-masks-743923/>

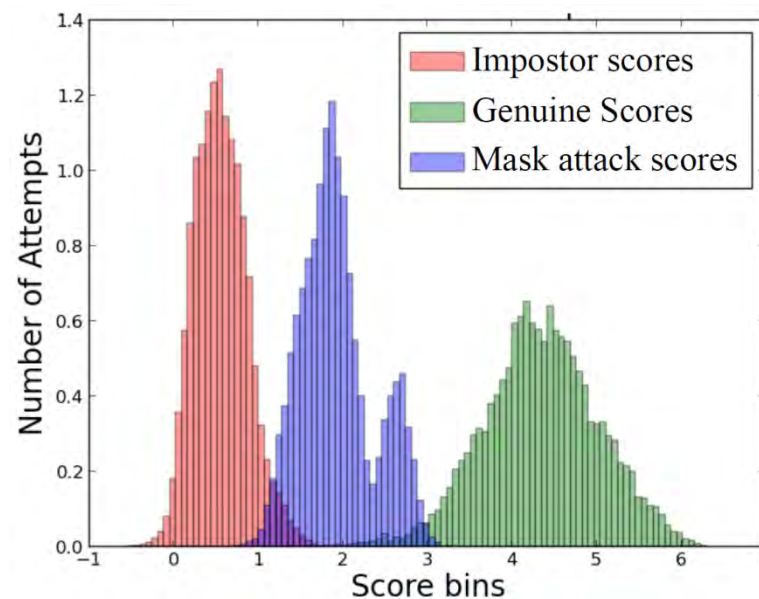
# Related Work

- Existing works on 3D Mask Spoofing Attack
  - The 3DMAD dataset [Erdogmus et al., BTAS'13]
  - LBP-based solution [Erdogmus et al., TIFS'14]



# Related Work

- The 3DMAD dataset
  - Score distributions of genuine, impostor, and mask attack scores of 3DMAD using ISV for 2D face verification

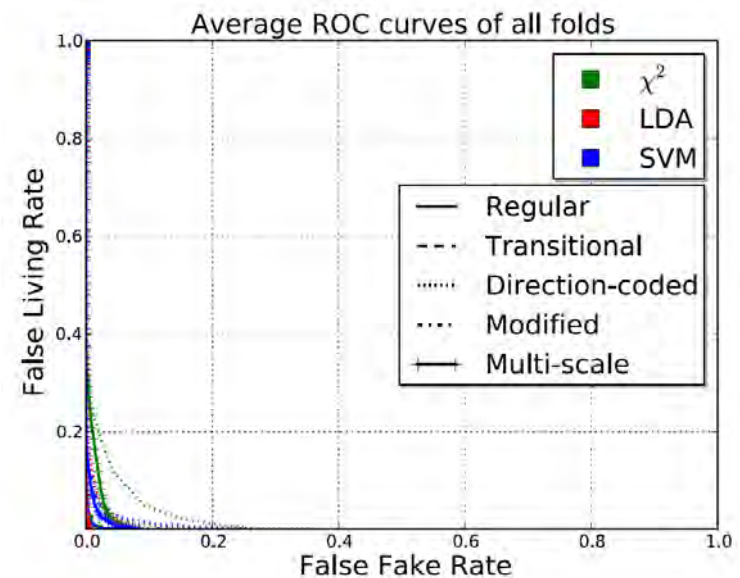
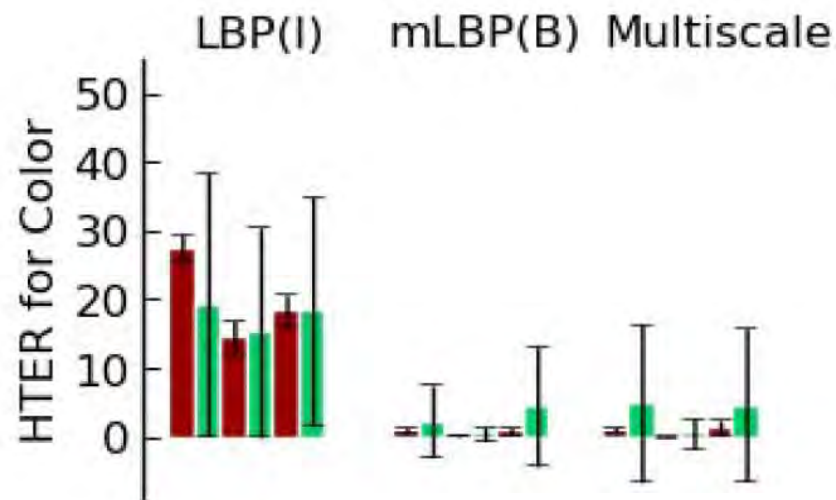


[Erdogmus et al., BTAS'13]



# Related Work

- LBP-based solution
  - The multi-scale LBP features yield to very good results on 3DMAD [Erdogmus et al., TIFS'14]



[Erdogmus et al., TIFS'14]

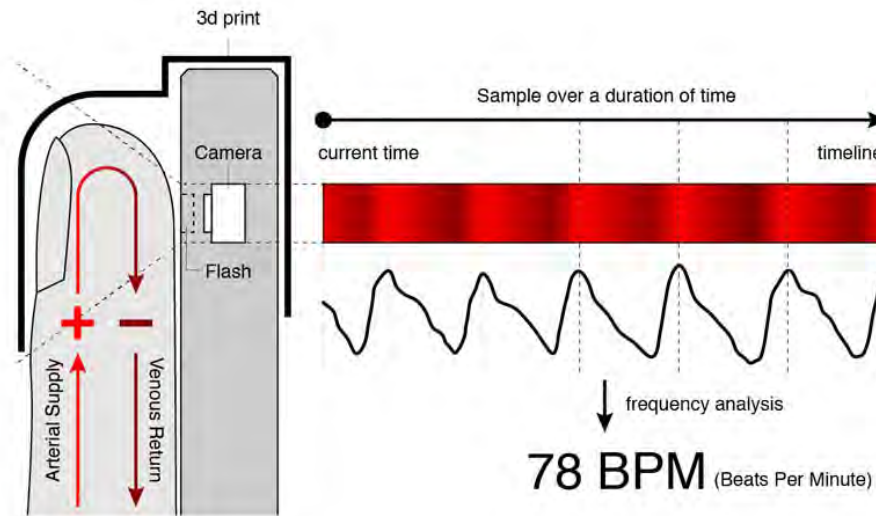
# Analysis of Existing Methods

- Pros and Cons
  - + Achieve high performance in 3DMAD dataset
  - Hyperreal 3D mask may not have quality defects
  - LBP-based solution may not have good generalization ability across databases

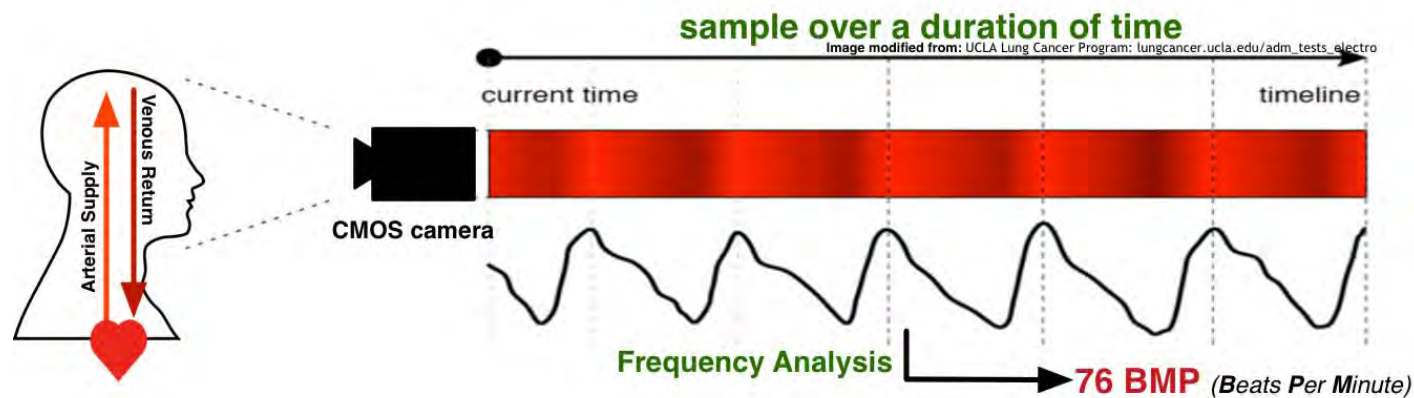


# **rPPG Approach for 3D Face Anti-spoofing**

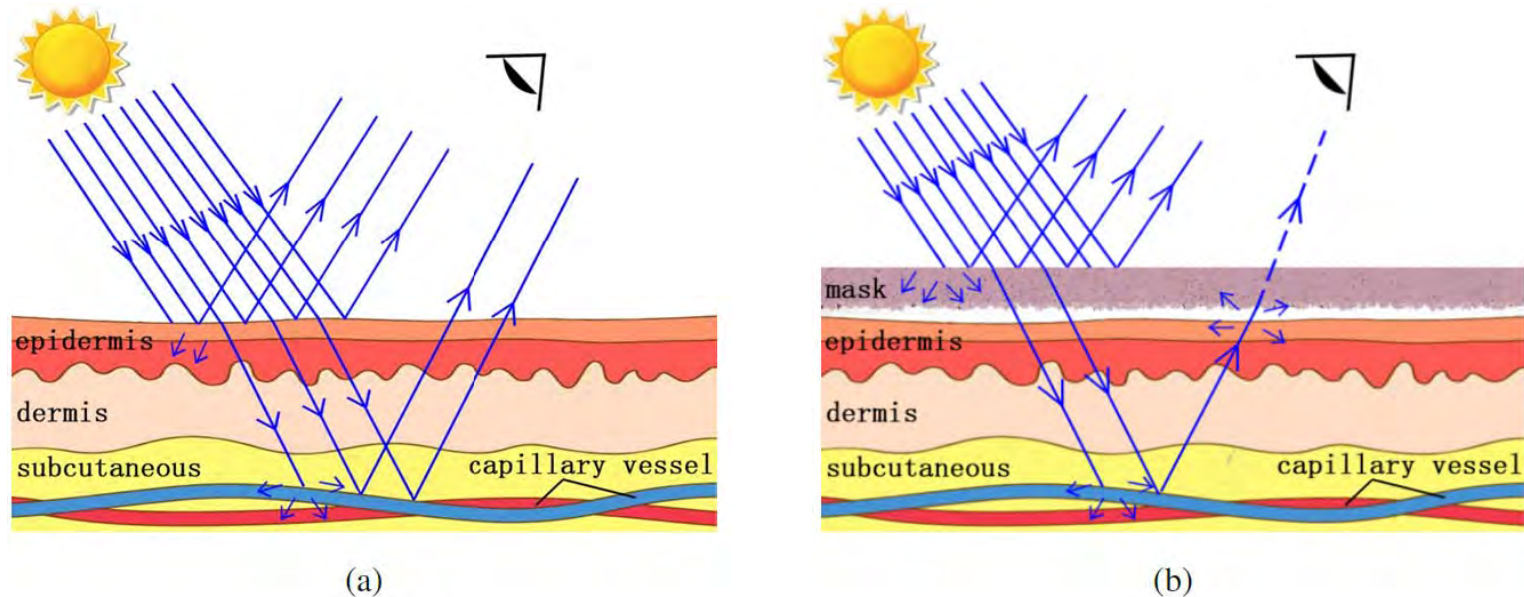
# PhotoPlethysmoGraphy (PPG)



# remote PhotoPlethysmography (rPPG)



# Principle of rPPG Based Face Anti-Spoofing



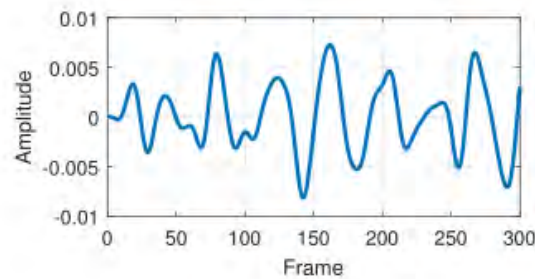
- (a) rPPG signal can be extracted from genuine face skin.
- (b) rPPG signals will be **too weak** to be detected from a masked face.
- light source needs to penetrate the mask before interacting with the blood vessel.
  - rPPG signal need to penetrate the mask before capturing by camera

# Principle of rPPG Based Face Anti-Spoofing

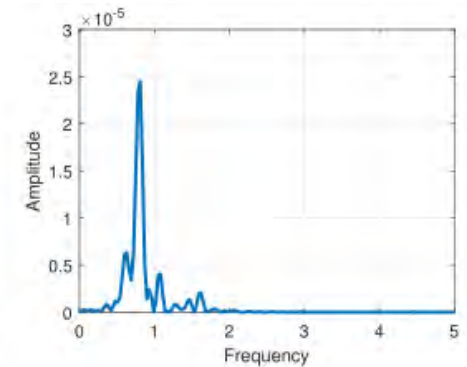
genuine face



(a)



(b)

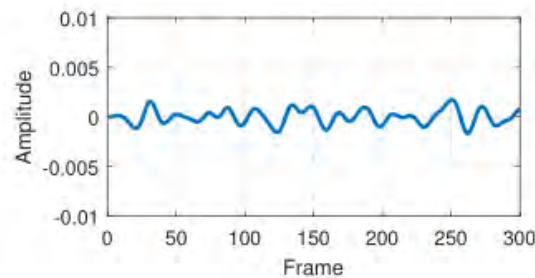


(c)

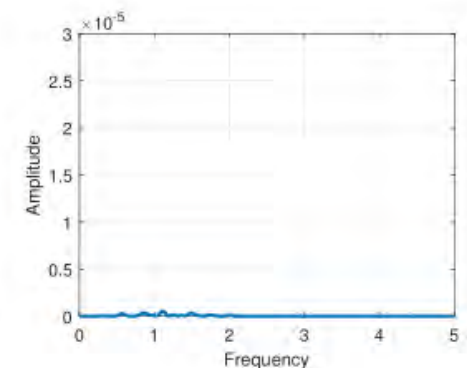
masked face



(d)



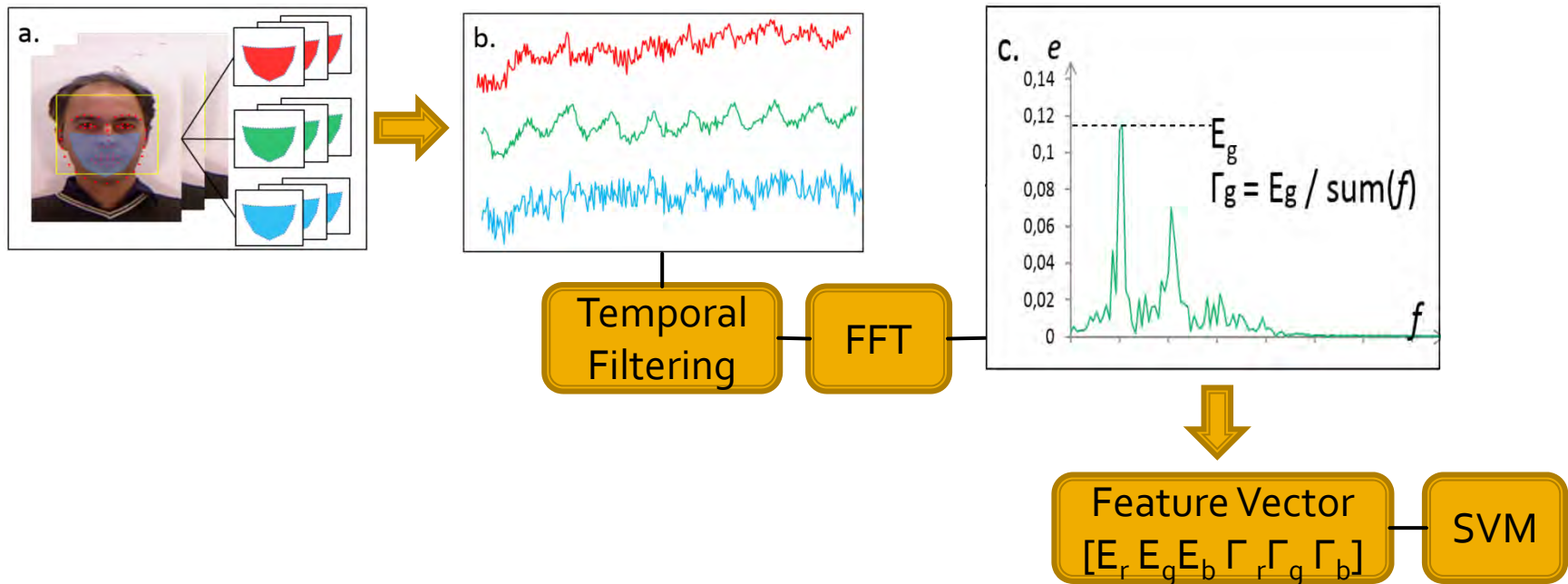
(e)



(f)



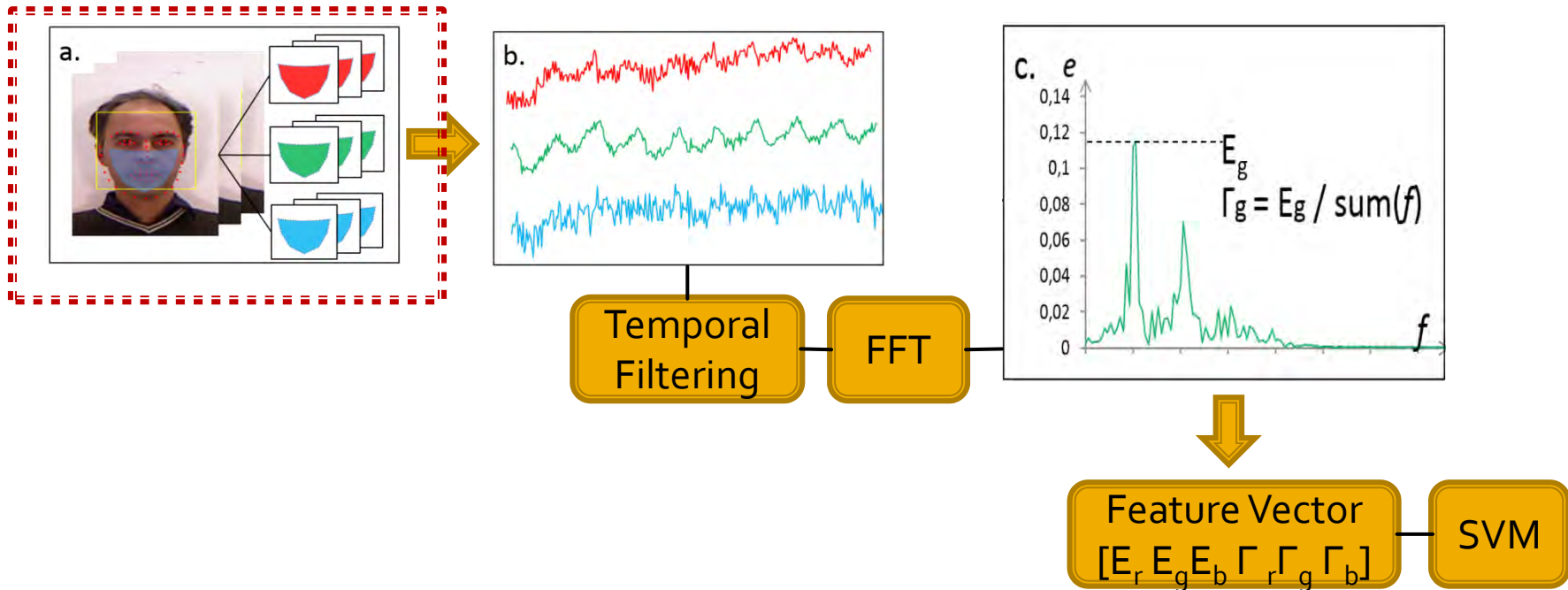
# Global rPPG-based Face Anti-Spoofing [ICPR 2016]



X Li, J Komulainen, G Zhao, P C Yuen and M Pietikainen,  
"Generalized face anti-spoofing by detecting pulse from face videos"  
ICPR 2016



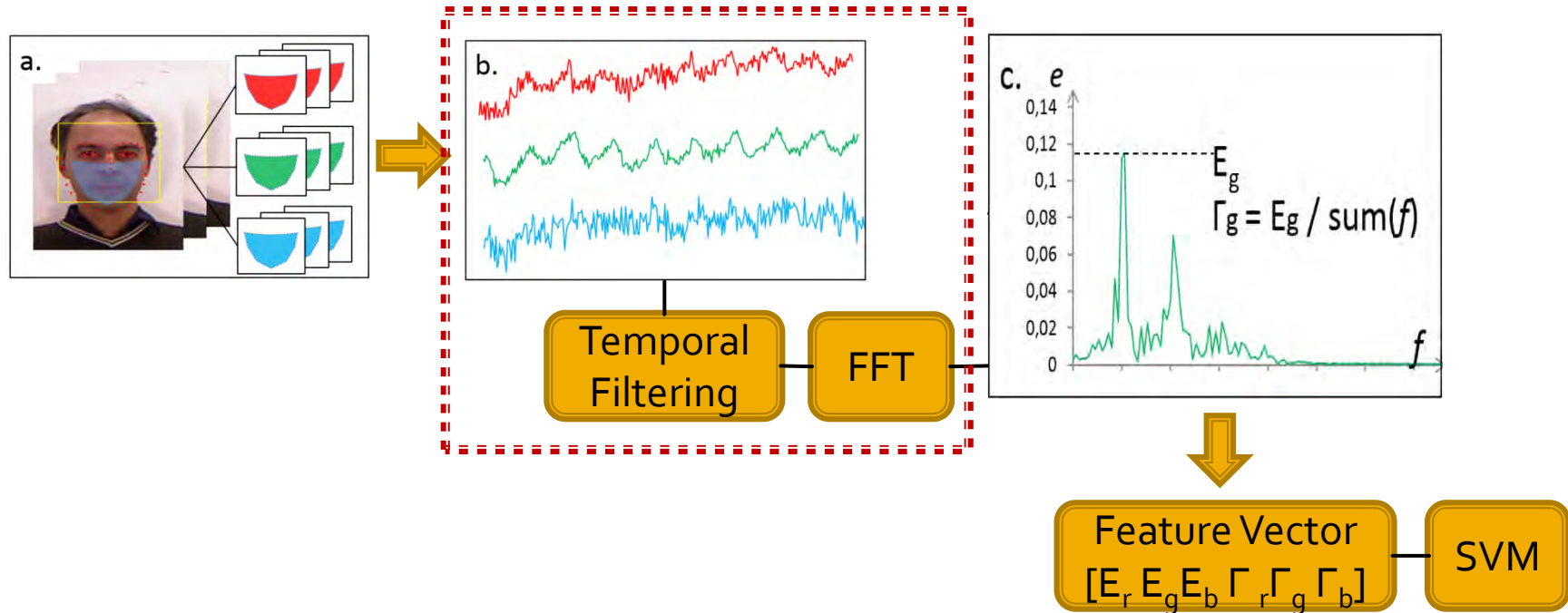
# Global rPPG-based Face Anti-Spoofing



## a. Face Detection and ROI tracking

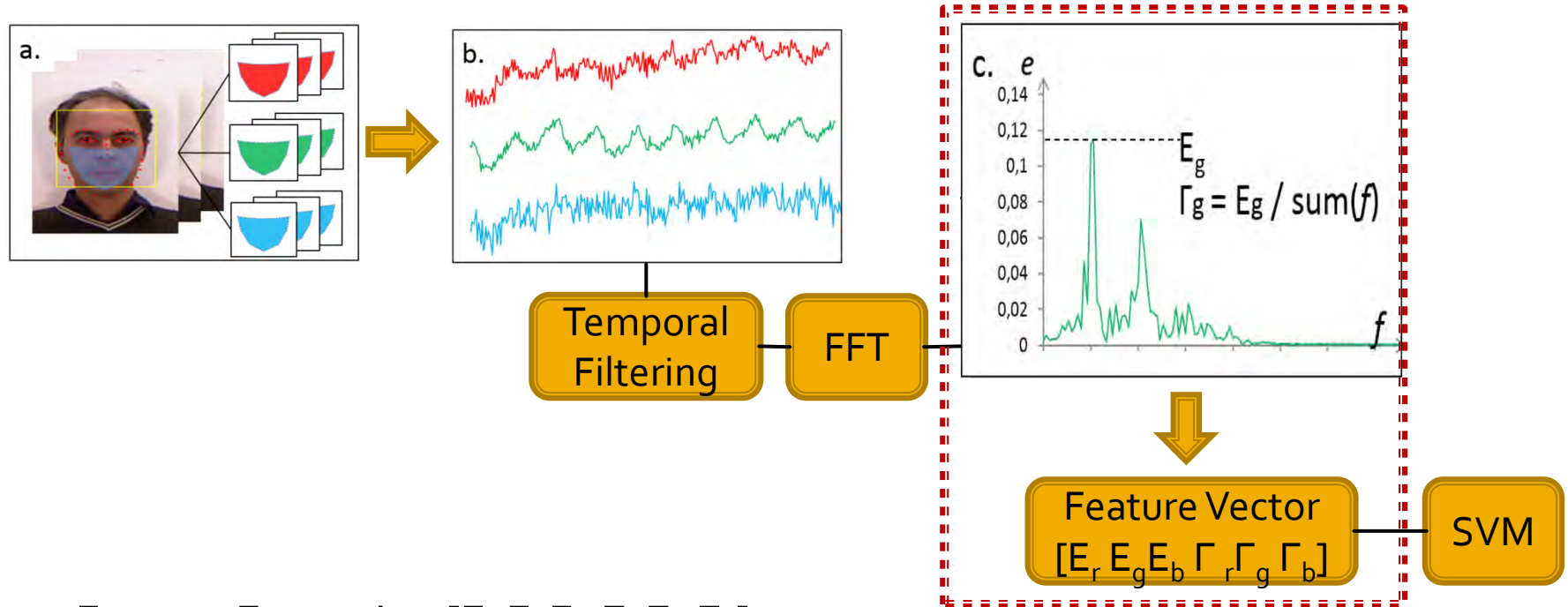
- Use Viola-Jones face detector on the first frame
- Find 66 facial landmarks [CVPR'13 Asthana et.al] within the face bounding box. Use 9 of them to define the ROI
- ROI is tracked through all frames using KLT

# Global rPPG-based Face Anti-Spoofing



- b. Three raw pulse signals  $r_{raw}$ ,  $g_{raw}$  and  $b_{raw}$  are computed, respectively.
- FIR bandpass filter with a cutoff frequency range of  $[0.7; 4]$  Hz ( $[42; 240]$  beat-per-minute)
  - Use fast Fourier transform (FFT) to convert the pulse signals into frequency domain  $\rightarrow$  PSD curve:  $f$

# Global rPPG-based Face Anti-Spoofing



c. Feature Extraction  $[E_r, E_g, E_b, \Gamma_r, \Gamma_g, \Gamma_b]$

- $E = \max(e(f))$
- $\Gamma = \frac{E}{\sum_{\forall f \in [0.7, 4]} e(f)}$

# Experimental Results

- Data:
  - 3DMAD [Erdogmus et.al TIFS'14]
    - 255 videos recorded from 17 subjects
    - Masks made from *ThatsMyFace.com*
  - 2 REAL-F Masks
    - 24 videos recorded from 2 subjects
    - Hyper real masks from *REAL-F*



# Experimental Results

- Results on 3DMAD
  - LOOCV protocol [Erdogmus *et.al* TIFS'14 ]

	3DMAD-dev	3DMAD-test	
Method	EER(%)	HTER(%)	EER(%)
<b>Pulse (ours)</b>	<b>2.31</b>	<b>7.94</b>	<b>4.17</b>
LBP-blk	0	0	0
LBP-blk-color	0	0	0
LBP-ms	0	0	0
LBP-ms-color	0	0	0

Note:

**LBP-blk:**  $LBP_{8,1}$  extracted from 33 blocks of a gray-scale face

**LBP-blk-color:** LBP-blk extracted separately from each RGB color channel

**LBP-ms:** multi-scale LBP extracted from a whole gray-scale face image combining  $LBP_{8,1}$ ,  $LBP_{8,2}$ ,  $LBP_{8,3}$ ,  $LBP_{8,4}$ , and  $LBP_{16,2}$

**LBP-ms-color:** LBP-ms extracted separately from each RGB color channel

# Experimental Results

## ■ Results on REAL-F

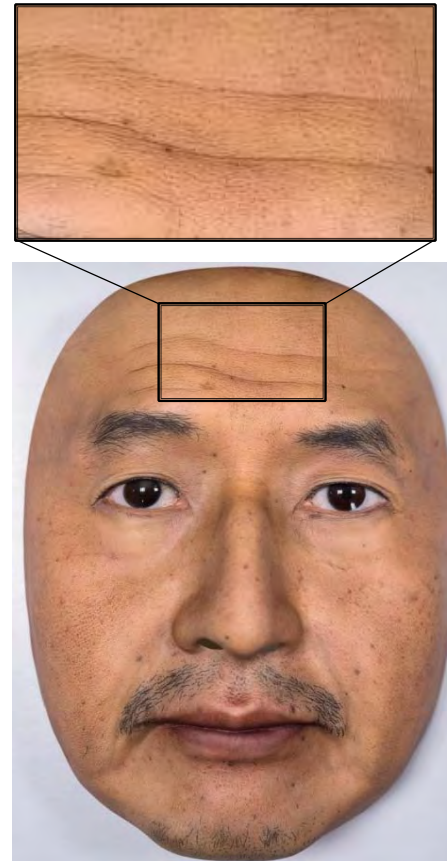
- Randomly select 8 subjects from 3DMAD for training and the other 8 subjects as the development set

	REAL-F			
Method	HTER(%)	EER(%)	FPR (@FNR=0.1%)	FPR (@FNR=0.01%)
<b>Pulse (ours)</b>	<b>4.29</b>	<b>1.58</b>	<b>0.25</b>	<b>3.83</b>
LBP-blk	26.3	25.08	37.92	48.25
LBP-blk-color	25.92	20.42	31.5	48.67
LBP-ms	39.87	46.5	59.83	73.17
LBP-ms-color	47.38	46.08	86.5	95.08

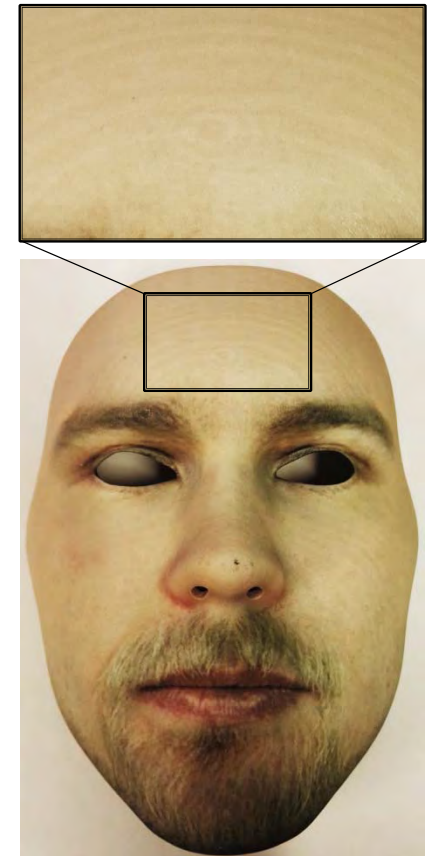


# Analysis of Results

- Observations:
  - LBP-based texture method gives *zero error* for *3DMAD* dataset but *very large error* in *REAL-F*
  - Global rPPG method (pulse) provides *very small errors* in both *3DMAD* and *REAL-F* datasets



REAL-F



3DMAD

# Limitations on Global rPPG method

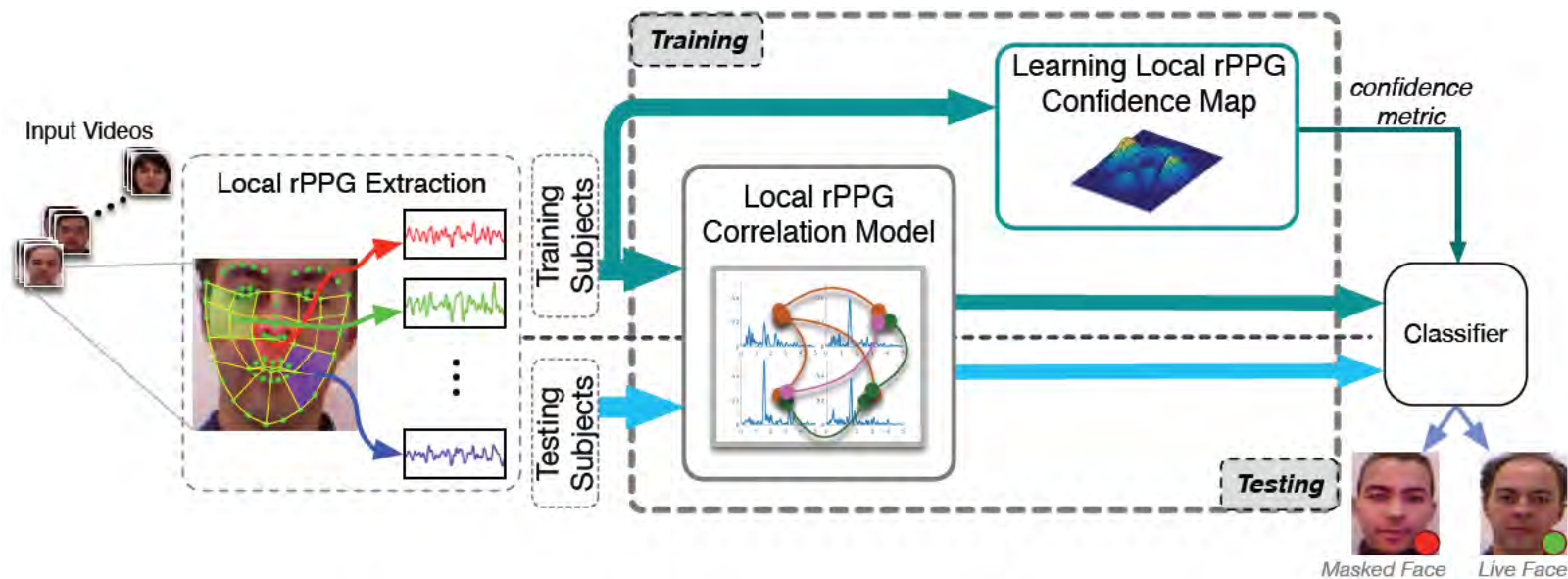
- Global rPPG signal is sensitive to certain variations such as illuminations, head motion and video quality
- rPPG signal strength may vary with different subjects





**How to increase the robustness of  
rPPG-based Face Anti-spoofing?**

# Local rPPG based Face Anti-Spoofing Method [ECCV 2016]



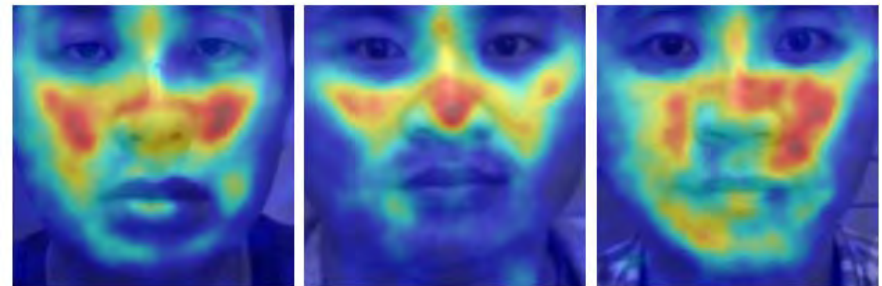
S Q Liu, P C Yuen, S P Zhang and G Y Zhao  
**3D Mask Face Anti-spoofing with Remote  
Photoplethysmography**  
*ECCV 2016*

# Rationale

- For different subjects, the patterns of facial blood vessels are similar.

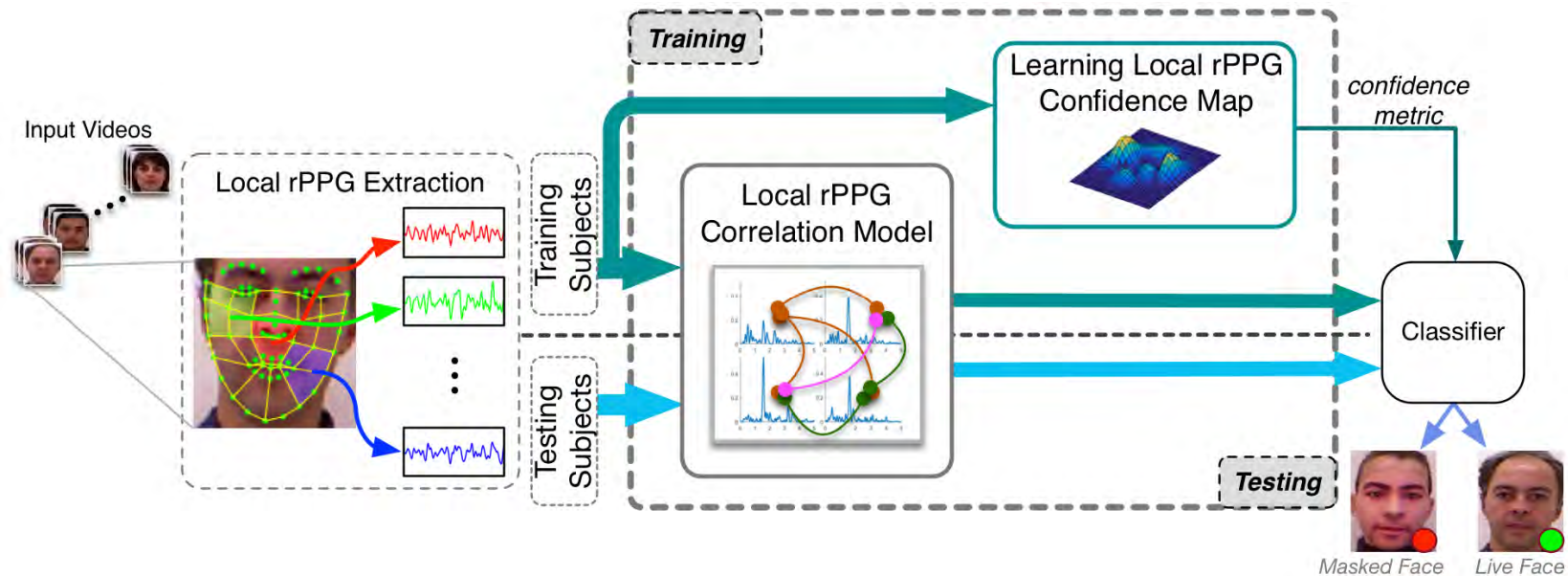


Generic map of blood vessels on the face



SNR map of local rPPG signals for different subjects

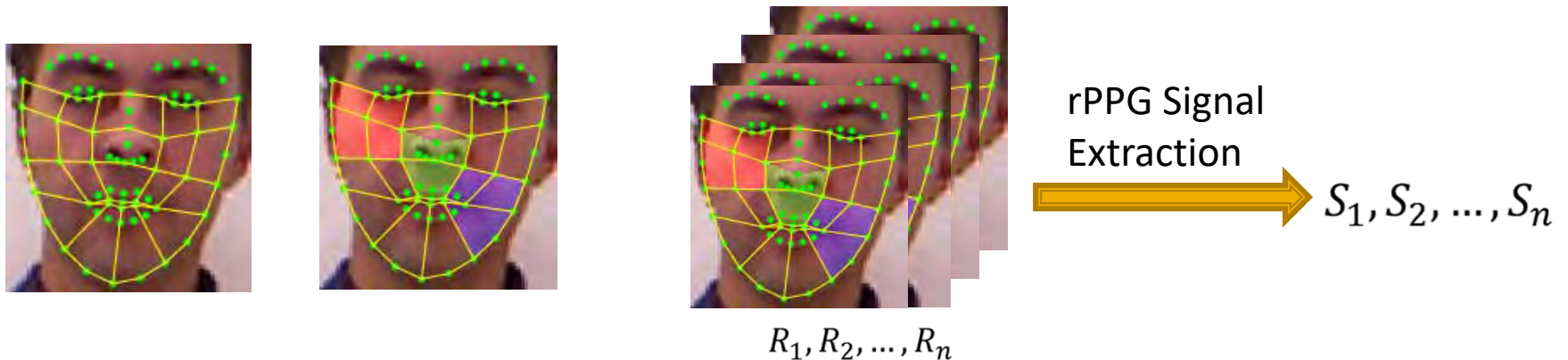
# Local rPPG based Face Anti-Spoofing Method



- (a) Local ROIs are pre-defined based on the facial landmarks. Local rPPG signals are extracted from these local face regions.
- (b) Extract Local rPPG patterns through the proposed **local rPPG correlation model**.
- (c) Training stage: local rPPG confidence map is learned, and then transformed into distance metric for classification.
- (d) Classifier: SVM

# 1. Local rPPG Signal Extraction

- (i) ROI detection and tracking
  - Landmark detection and tracking
  - Local ROIs are pre-defined based on the facial landmarks
- (ii) rPPG Signal Extraction
  - We adopt (Haan et.al., TBE, 2013) method to extract rPPG signals.



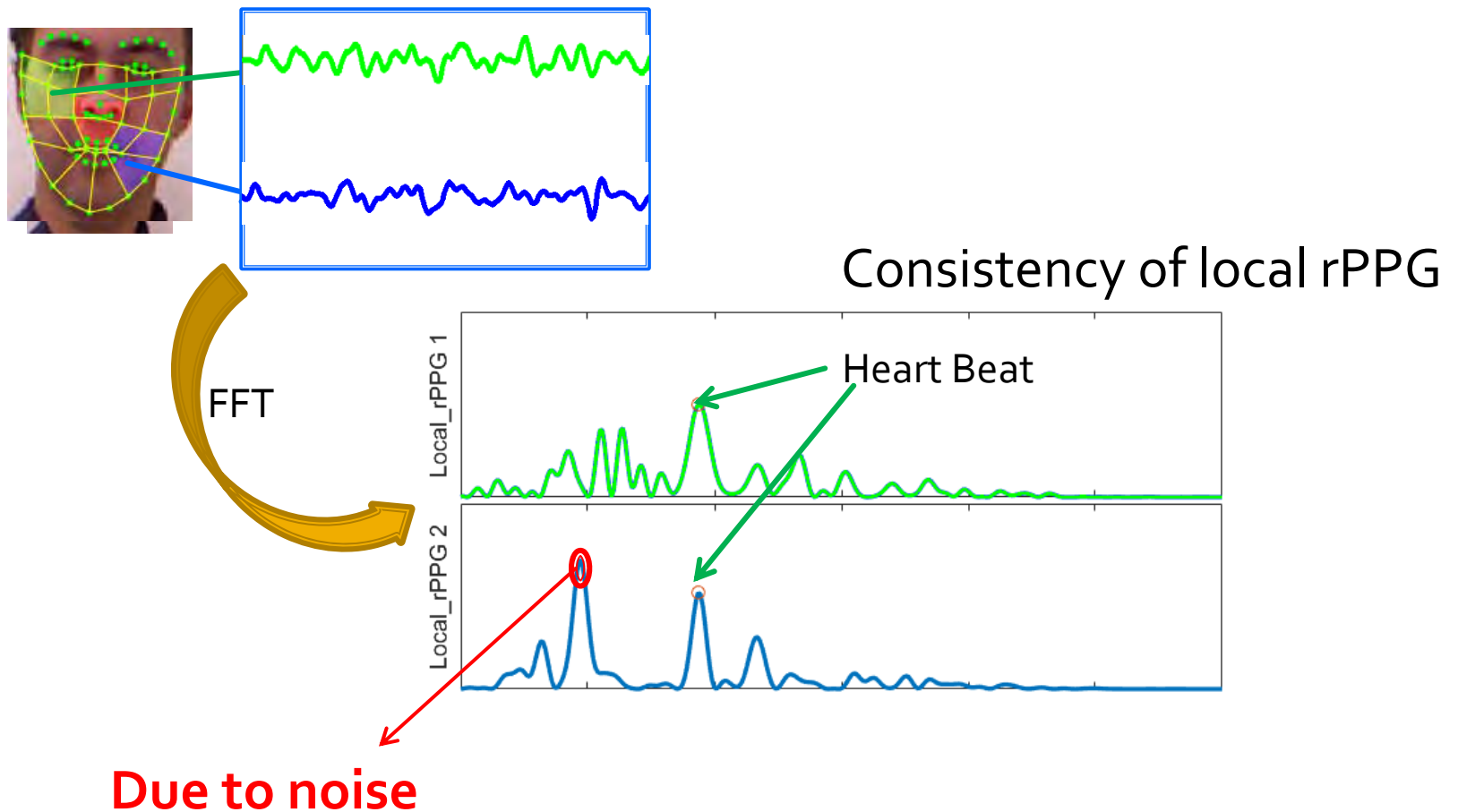
## 2. Local rPPG Correlation Model

- To handle noise introduced in rPPG signal due to different variations, such as illuminations, head motion, ...
- For genuine face, local rPPG signals should have high consistency
- For masked face, local rPPG signals should have a small frequency similarity and periodicity

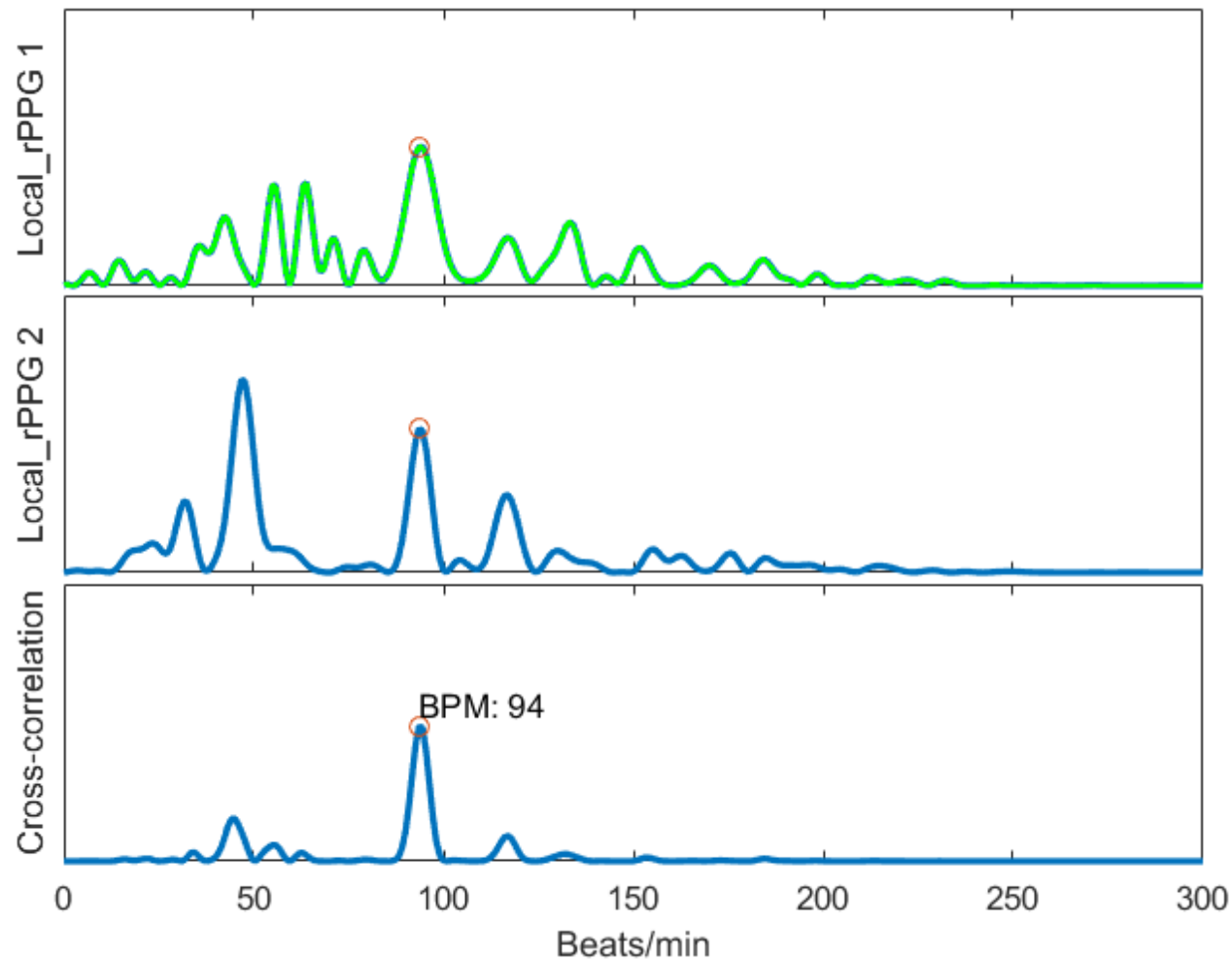


## 2. Local rPPG Correlation Model

- Local rPPG on genuine face

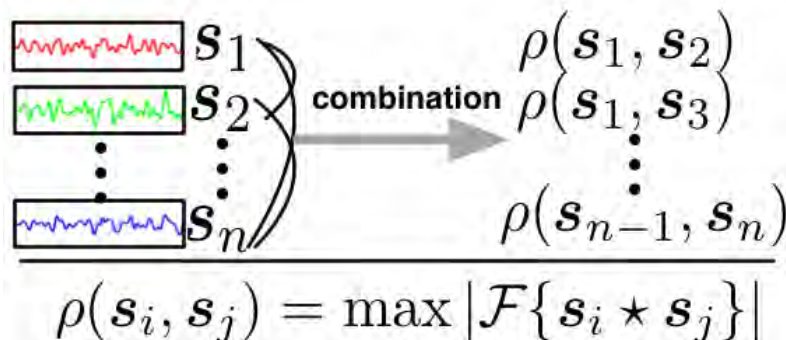


## 2. Local rPPG Correlation Model



## 2. Local rPPG Correlation Model

Similarity of all possible combinations of local rPPG signals

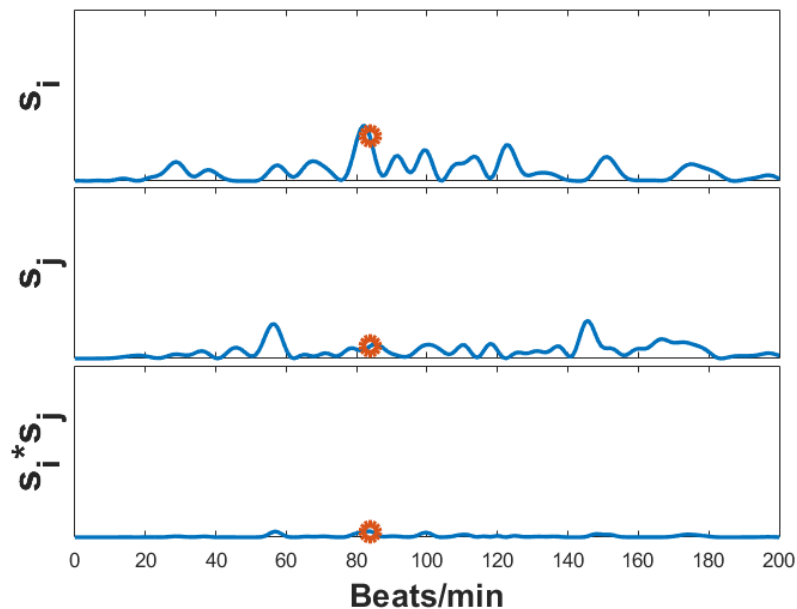


Through the cross-correlation operation, we could **reduce the effect of noise due to illumination and motion**.

Moreover, **noise from local masked face regions will be suppressed using cross correlation**, because random noise do not share the same periodic frequency.

### 3. Learning Local rPPG Confidence Map

- Local rPPG correlation pattern may not be sufficient to handle noise in some cases
  - rPPG signals may be too weak in low quality video



### 3. Learning Local rPPG Confidence Map

- Local rPPG correlation pattern may not be sufficient to handle noise in some cases
  - rPPG signals may be too weak in low quality video and concealed by noise
- rPPG signal strength varies with different local face regions

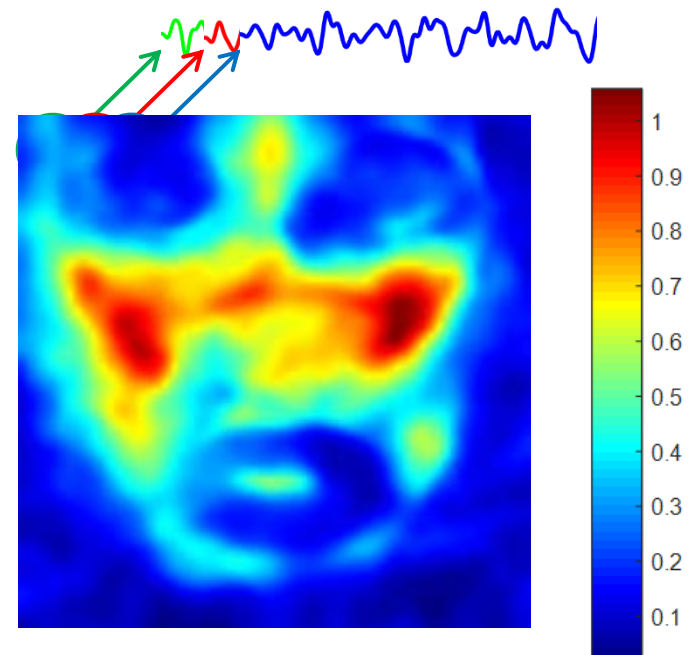
*We propose to learn a **local rPPG confidence map***

1. emphasizing the region with strong HR signal, and
2. weaken the unreliable region with pale HR signal.

### 3. Learning Local rPPG Confidence Map



Generic map of blood vessels on the face



**The distribution of local rPPG signals should be considered**



### 3. Learning Local rPPG Confidence Map

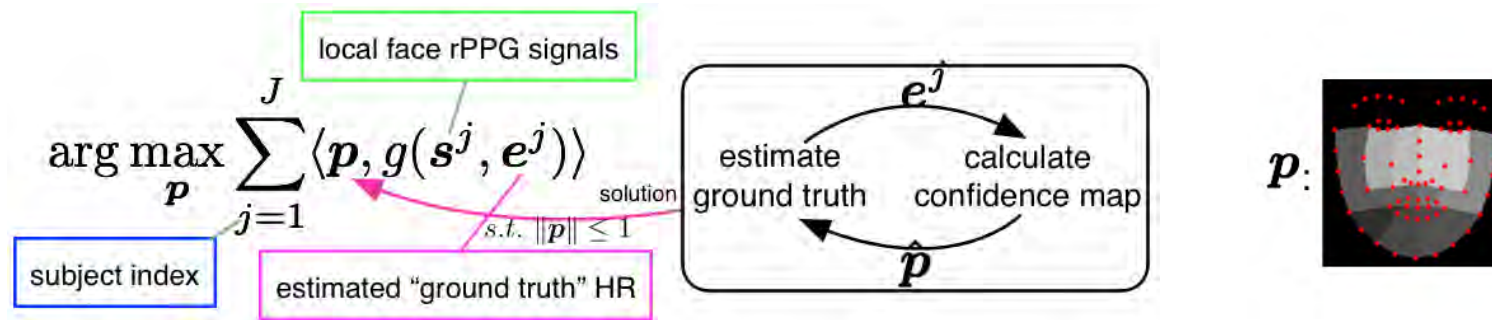
- How to measure the strength of HR signal?
- Signal to Noise Ratio (SNR)

$$\frac{\sum_{f_{HR-r}}^{f_{HR+r}} \hat{\mathbf{s}}^j(f)}{\sum \hat{\mathbf{s}}^j(f) - \sum_{f_{HR-r}}^{f_{HR+r}} \hat{\mathbf{s}}^j(f)}$$

- How to find the estimated ground truth HR signal?

### 3. Learning Local rPPG Confidence Map

- An iterative algorithm: Given  $J$  training subjects, learn the local rPPG confidence map  $\mathbf{p}$  which reflects the reliability of local face regions:

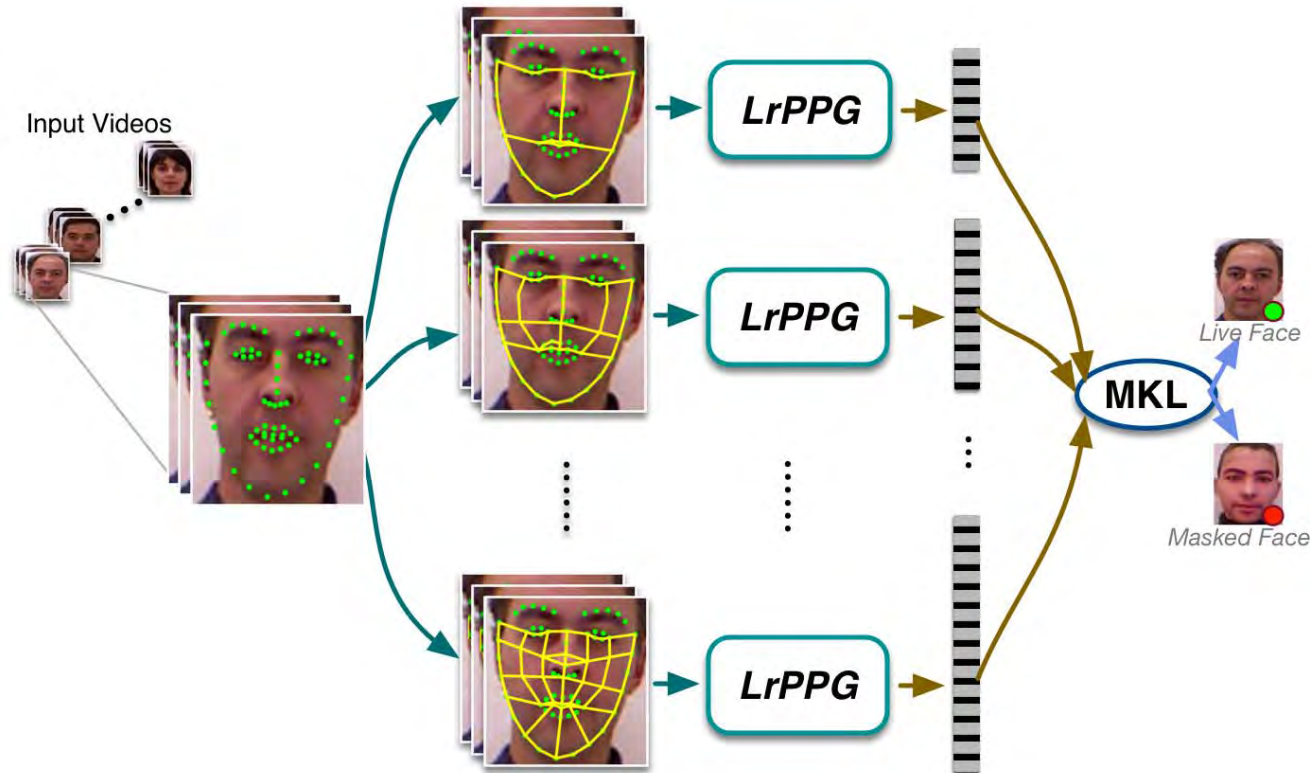


- Using local rPPG confidence map  $\mathbf{p}$  to weight the distance metric in classifier

# Limitations on Local rPPG Method

- rPPG quality (Discriminability) highly depends on the local regions size:
  - Smaller region: Signal quality ↓, spatial information ↑
  - Larger region: Signal quality ↑, spatial information ↓
- Large variations (lighting condition & camera settings)
- Multi-scale ROI strategy can better adapt different application environment in practice

# Multi-Scale Local rPPG Method



$$K_{MS}(x, x') = \sum_{m \in M} d_m K(x_m, x'_m)$$

$$s.t., d_m \geq 0, \sum_{m \in M} d_m = 1$$

$M$  : number of scales  
 $x_m$ : LrPPG feature extracted from of scale  $m$   
 $d_m$ : weight of scale  $m$

# Experimental Results

## ■ Datasets

- 3DMAD [TIFS'14 Erdogmus et.al]
  - 255 videos recorded from 17 subjects
  - Masks made from ThatsMyFace.com



- HKBU MARs V2 Dataset:
  - 2 Mask types and **12** subjects: ThatsMyFace (6), REAL-F (6)
  - Captured by WebCam Logitech C920 (1280\*720 RGB)



More details can be found: <http://rds.comp.hkbu.edu.hk/mars/>

# Experimental Results

- Intra-Database Experiment (*LOOCV*)
  - 3DMAD, HKBU MARs V2, and Combined Dataset (3DMAD+HKBU MARs V2)
- Cross-Database Experiment
  - Training on 3DMAD, Test on HKBU MARs V2 dataset
  - Training on HKBU MARs V2 , Test on 3DMAD dataset
- Cross-Mask Experiment
  - Training and test using different mask types



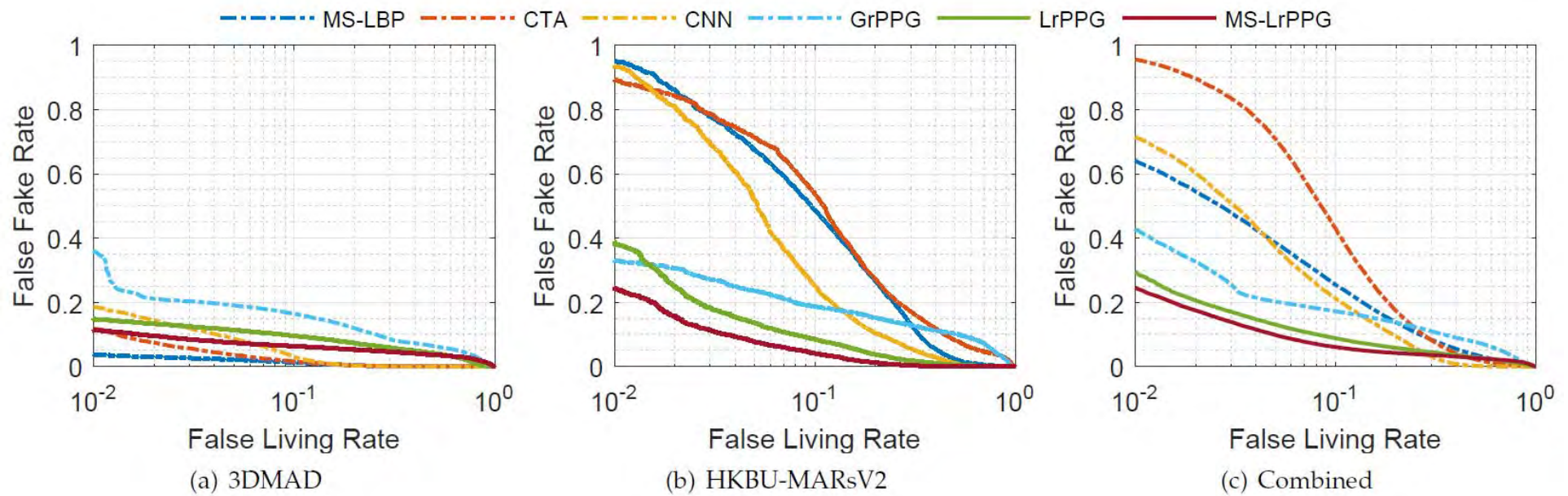
# Experimental Results

- Intra-database experiments (*LOOCV*)

Combined		HTER_dev	HTER_test	EER	AUC	FFR@ FLR=0.1	FFR@ FLR=0.01
	MS-LBP[2]	15.7 $\pm$ 4.2	16.2 $\pm$ 22.6	16.6	91.0	25.4	64.2
	CTA [1]	18.4 $\pm$ 5.8	19.5 $\pm$ 21.5	18.9	87.7	42.9	95.7
	CNN [6]	13.5 $\pm$ 5.9	14.6 $\pm$ 20.6	14.5	93.5	21.2	71.5
	GrPPG	15.3 $\pm$ 2.9	15.5 $\pm$ 18.5	15.2	91.1	17.2	42.8
	<b>LrPPG [5]</b>	8.69 $\pm$ 1.5	9.16 $\pm$ 11.9	9.21	95.7	8.79	29.4
	<b>MS-LrPPG</b>	<b>6.93 <math>\pm</math> 1.2</b>	<b>6.92 <math>\pm</math> 11.1</b>	<b>7.41</b>	<b>96.4</b>	<b>6.07</b>	<b>24.6</b>
HKBU MARs V2		HTER_dev	HTER_test	EER	AUC	FFR@ FLR=0.1	FFR@ FLR=0.01
	MS-LBP[2]	20.5 $\pm$ 8.9	24.0 $\pm$ 25.6	22.5	85.8	48.6	95.1
	CTA [1]	22.4 $\pm$ 10.4	23.4 $\pm$ 20.5	23.0	82.3	53.7	89.2
	CNN [6]	13.7 $\pm$ 10.8	14.8 $\pm$ 22.2	15.2	91.4	25.1	93.5
	GrPPG	15.4 $\pm$ 6.7	16.1 $\pm$ 20.5	16.4	89.4	18.6	32.9
	<b>LrPPG [5]</b>	8.43 $\pm$ 2.9	8.67 $\pm$ 8.8	9.07	97.0	8.51	38.9
	<b>MS-LrPPG</b>	<b>6.07 <math>\pm</math> 2.6</b>	<b>6.44 <math>\pm</math> 7.6</b>	<b>6.38</b>	<b>98.5</b>	<b>4.08</b>	<b>24.5</b>

1. Z. Boulkenafet, J. Komulainen, and A. Hadid, “Face spoofing detection using colour texture analysis”, *TIFS*, 2016
2. N. Erdogmus and S. Marcel, “Spoofing face recognition with 3d masks”, *TIFS*, 2014
5. S. Liu, P.C. Yuen, S. Zhang, and G. Zhao, “3D Mask Face Anti-spoofing with Remote Photoplethysmography”, *ECCV*, 2016.
6. J. Yang, Z. Lei, and S. Z. Li, “Learn convolutional neural network for face anti-spoofing”, *arXiv*, 2014.

# Experimental Results: Intra-database



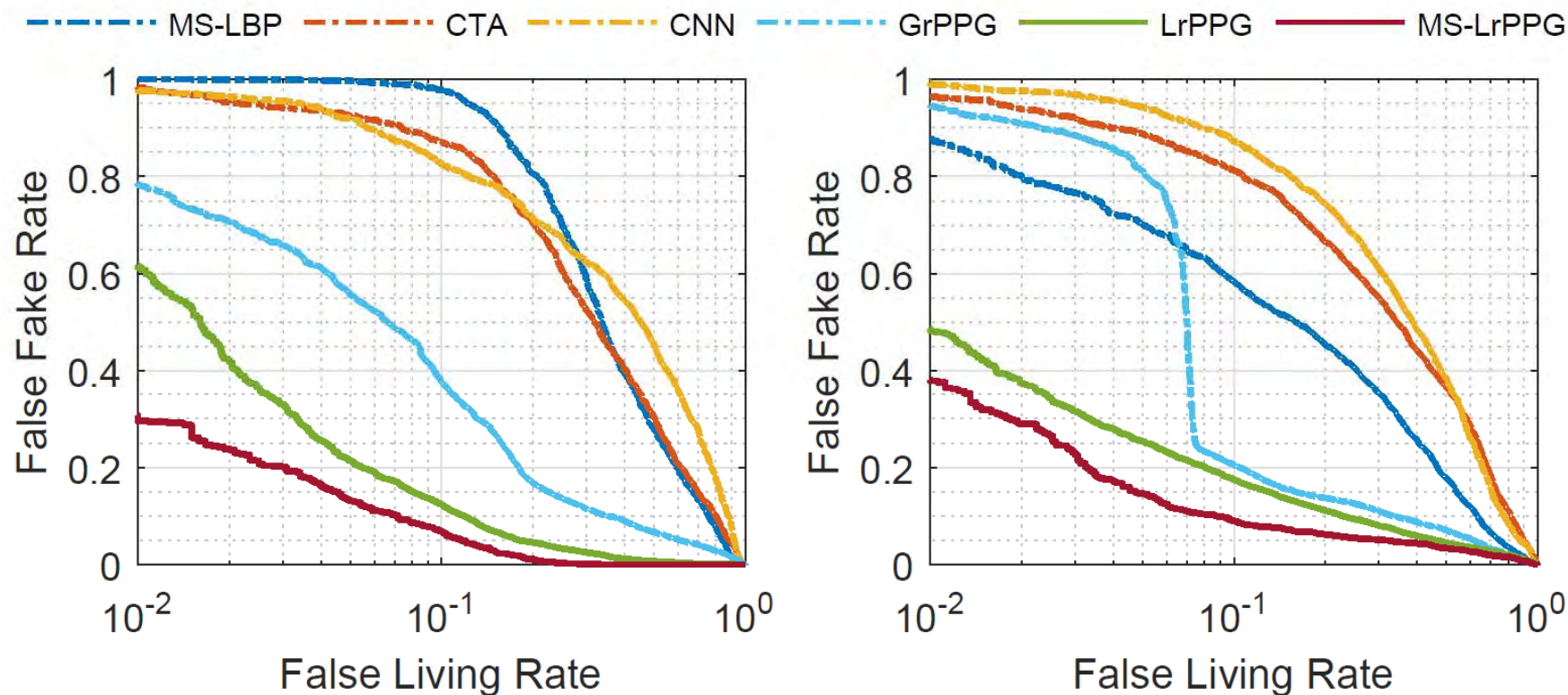
# Experimental Results

## ■ Cross-database experiments

	3DMAD→HKBU-MARsV2					HKBU-MARsV2→3DMAD				
	HTER(%)	EER(%)	AUC(%)	FFR@ FLR=0.1	FFR@ FLR=0.01	HTER(%)	EER(%)	AUC(%)	FFR@ FLR=0.1	FFR@ FLR=0.01
MS-LBP[2]	53.0 ± 3.6	39.8	60.4	97.8	100.0	32.8 ± 11.5	32.5	75.3	58.5	87.8
CTA [1]	40.1 ± 7.8	40.2	62.1	87.1	98.3	47.7 ± 5.4	42.5	60.5	81.2	96.5
CNN [6]	50.0 ± 0.0	47.8	54.6	82.6	97.9	50.0 ± 0.0	44.3	58.6	87.3	99.3
GrPPG	29.2 ± 9.7	20.4	87.7	34.8	62.8	18.4 ± 8.3	16.8	89.9	27.1	53.9
<b>LrPPG [5]</b>	16.8 ± 5.0	10.9	95.6	12.4	61.7	17.4 ± 4.4	14.0	92.3	17.4	48.7
<b>MS-LrPPG</b>	<b>13.2 ± 4.8</b>	<b>8.35</b>	<b>98.0</b>	<b>6.83</b>	<b>30.6</b>	<b>11.0 ± 2.0</b>	<b>9.59</b>	<b>95.0</b>	<b>9.00</b>	<b>38.2</b>

- Our proposed method is robust in cross-database scenario

# Experimental Results: Cross-database

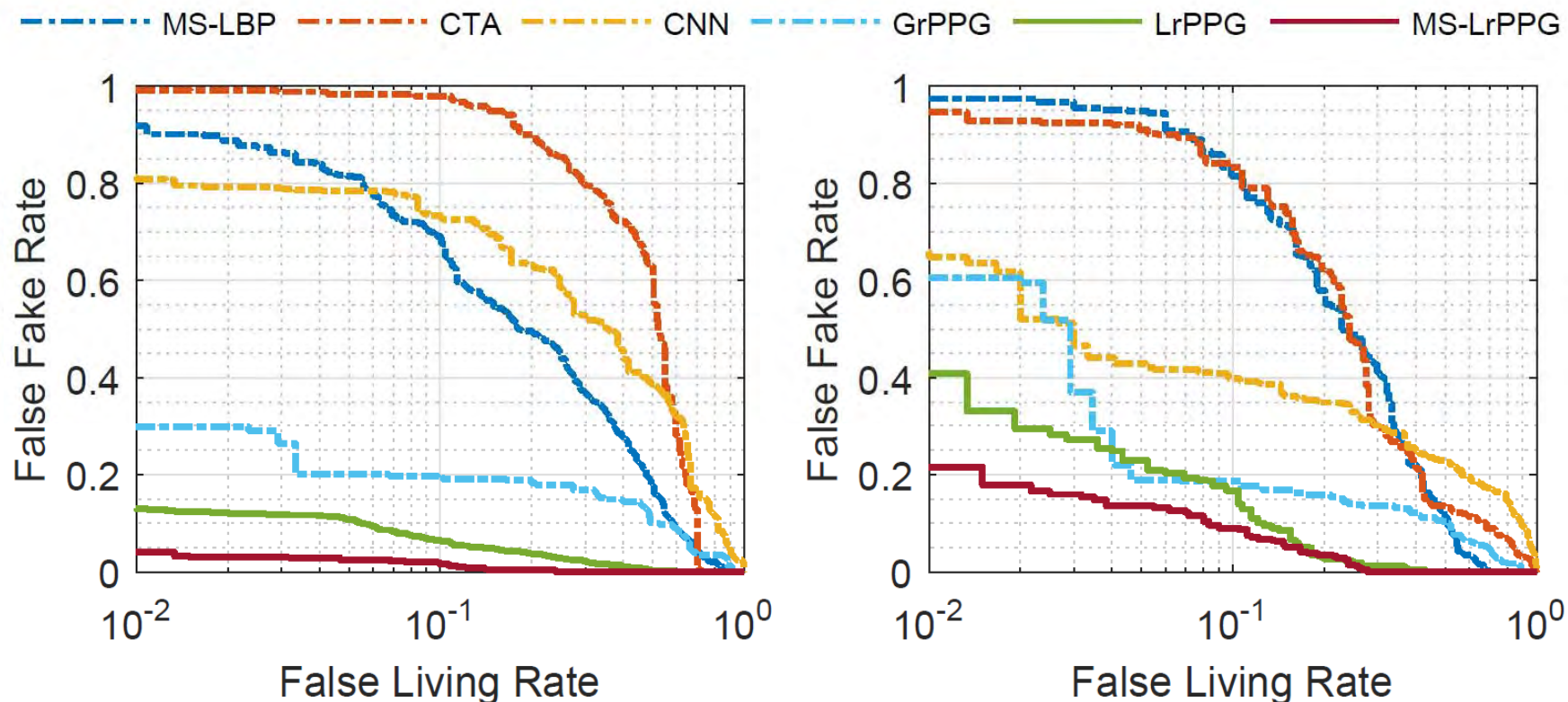


(a) 3DMAD  $\rightarrow$  HKBU-MARsV2

(b) HKBU-MARsV2  $\rightarrow$  3DMAD



# Experimental Results: Cross-mask



(a) TMF to RF (HKBU-MARsV2)

(b) RF to TMF (HKBU-MARsV2)

# Section Conclusion

- rPPG is a promising approach for face anti-spoofing against 3D mask
- We proposed a global and a local rPPG models for 3D mask face anti-spoofing and the results are encouraging.
- Still, a lot work needs to be done on improving the robustness on the variations, such as head motion, illuminations, video quality, pose and occlusion.

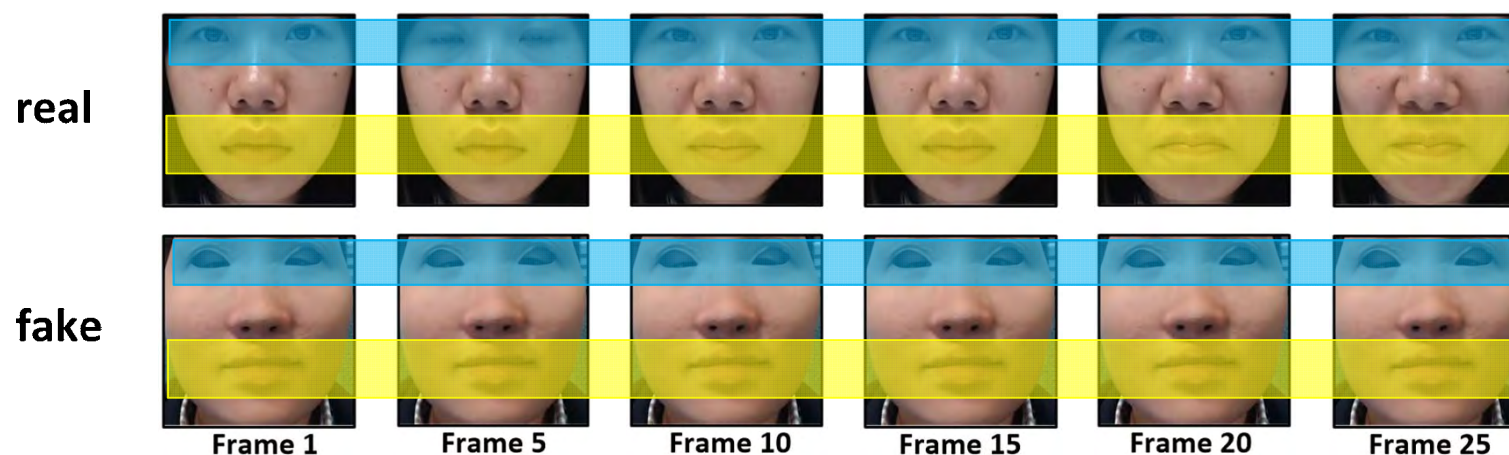




# Deep Learning Approach

# Joint Discriminative Learning of Deep Dynamic Textures

## ▪ Basic Idea



- Eye blinking
- Lip movements
- Some other facial components movements

} Captured by **dynamic textures**

Reference:

1. R Shao\*, X Y Lan\* and P C Yuen, "Deep Convolutional Dynamic Texture Learning with Adaptive Channel-discriminability for 3D Mask Face Anti-spoofing", *IAPR/IEEE International Joint Conference on Biometrics (IJCB)*, Oct 2017
2. R Shao\*, X Y Lan\* and P C Yuen, "Joint Discriminative Learning of Deep Dynamic Textures for 3D Mask Face Anti-spoofing", Submitted to Journal, 2017.

# Joint Discriminative Learning of Deep Dynamic Textures

## ■ Challenge

- A large portion of these facial movements are **subtle**
- Hand-crafted features are not fine-grained and descriptive enough to capture these **subtle** dynamic texture differences between real faces and 3D masks.

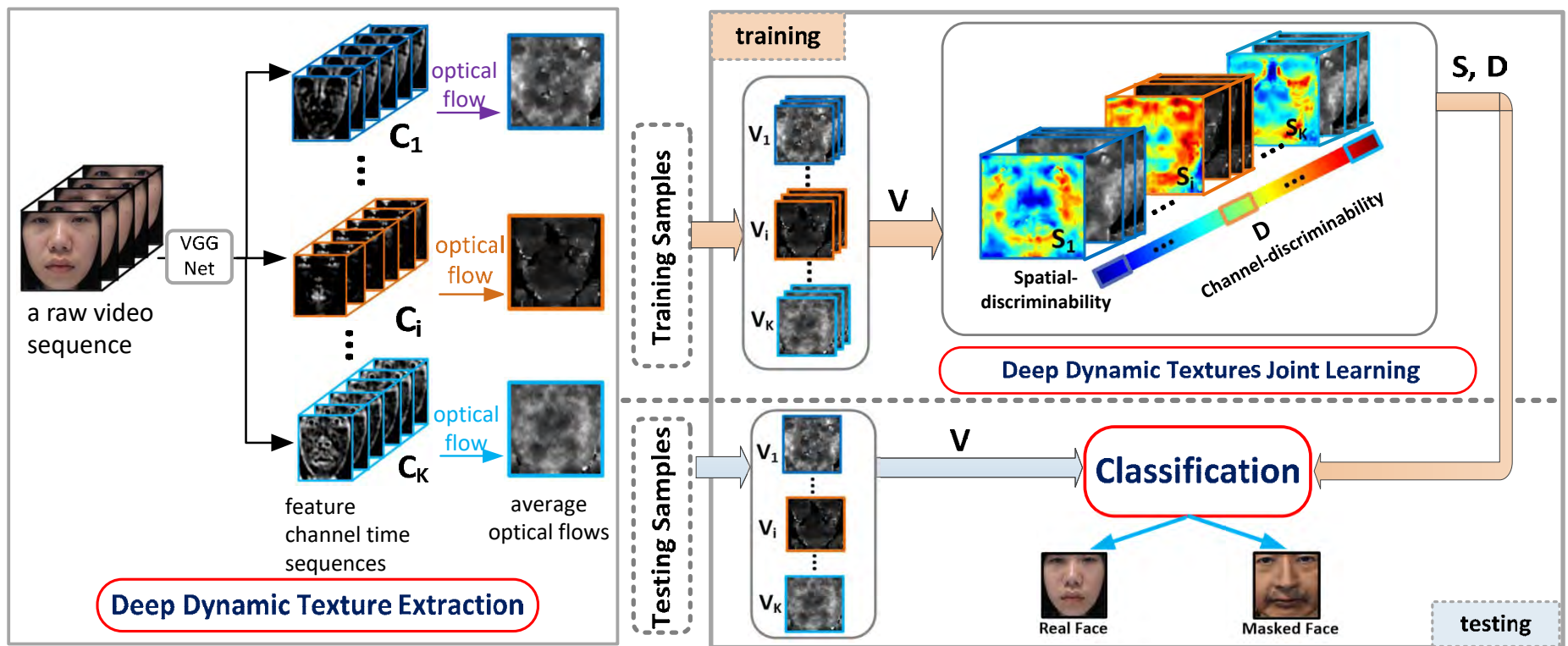
# Joint Discriminative Learning of Deep Dynamic Textures

## ■ LBPTOP vs Deep Dynamic Features

- LBPTOP [1]: Extract hand-crafted dynamic textures for anti-spoofing
- Deep Dynamic Features: Extract the dynamic information from texture features of the convolutional layer of CNN
  - Deep textures of the lower convolutional layer have strong description ability.
  - The dynamic feature estimated from these descriptive deep textures is more able to differentiate subtle facial motion differences than hand-crafted dynamic textures.

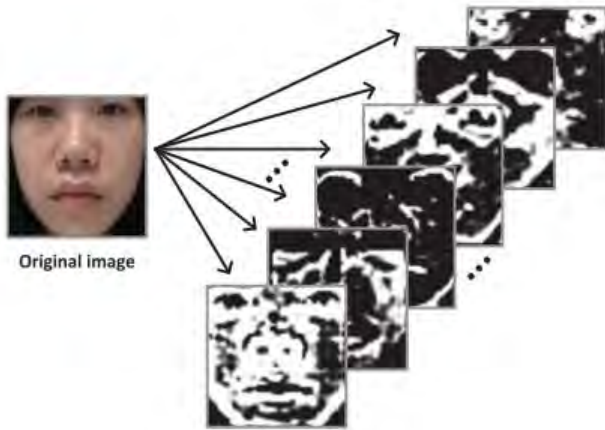
# Joint Discriminative Learning of Deep Dynamic Textures

## Framework



# Joint Discriminative Learning of Deep Dynamic Textures

## ▪ Deep Dynamic Texture Extraction



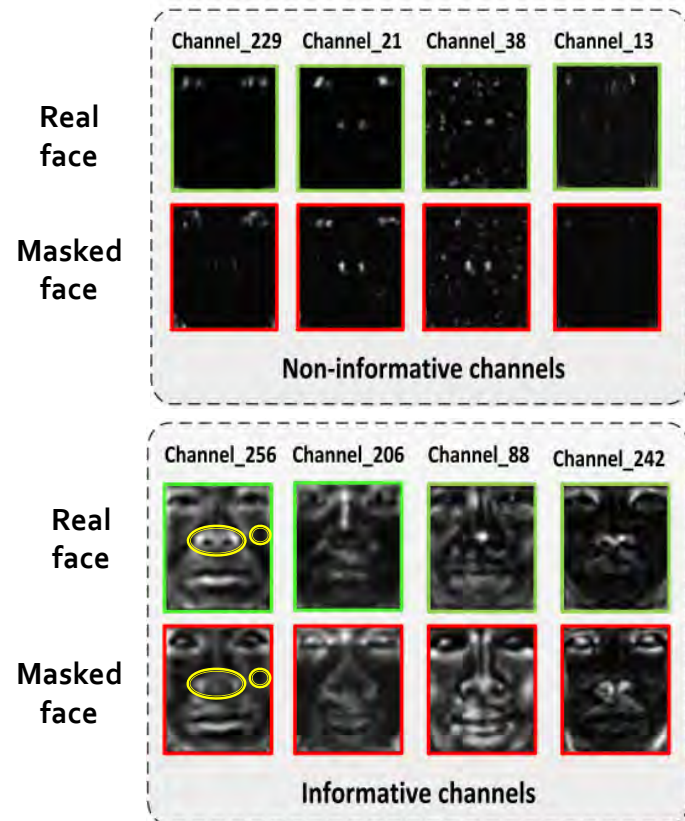
( The responses in feature channels of a lower convolutional layer of a sample )

- An original image of a sample can be decoupled into various texture responses in feature channels of a convolutional layer
  - Various fine-grained deep textures
- Every facial local region can be described by various fine-grained deep textures
  - Motion information of every facial local region can be described by the proposed visual cues of multiple deep dynamic textures
  - Differentiate the various subtle motion differences between the real face and masked face.



# Joint Discriminative Learning of Deep Dynamic Textures

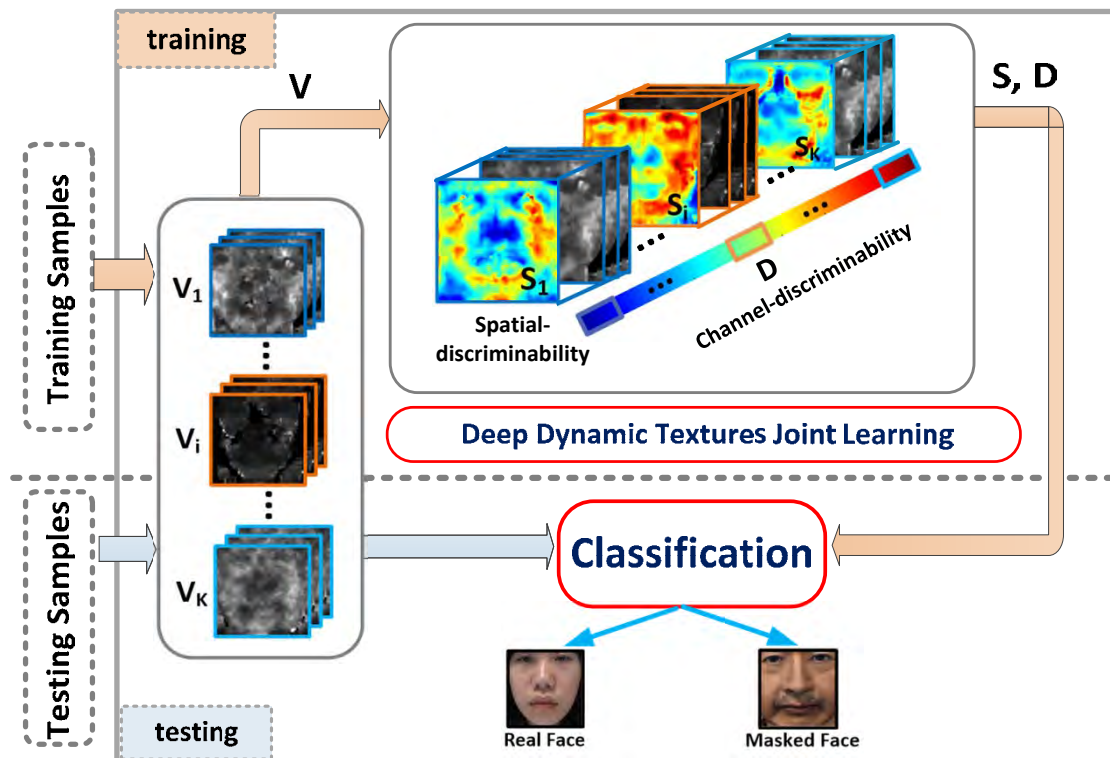
## ■ Deep Dynamic Texture Joint Learning



- Not all deep dynamic textures are useful for our task
  - Weighting the importance of different deep dynamic textures based on their discriminability

# Joint Discriminative Learning of Deep Dynamic Textures

## Deep Dynamic Texture Joint Learning



- Channel-discriminability  $D_i$
- Spatial-discriminability  $S_i$

### Discriminative learning model:

To capture the difference in channel- and spatial-discriminability for feature learning which enables more discriminative features to play more important role in face/mask decision

# Joint Discriminative Learning of Deep Dynamic Textures

## ▪ Dataset

- 3DMAD [TIFS'14 Erdogmus et.al]
  - 255 videos recorded from 17 subjects
  - Masks made from ThatsMyFace.com
- Supplementary (SUP) Dataset:
  - 120 videos recorded from 8 subjects
  - 2 Mask type: 8 subjects: ThatsMyFace (6), REAL-F (2)



# Joint Discriminative Learning of Deep Dynamic Textures

## ■ Protocols

- Intra-database Experiment (LOOCV) [TIFS'14 Erdogmus et.al]
  - 3DMAD Dataset
  - Supplementary (SUP) Dataset
- Cross-database Experiment:
  - Train on 3DMAD, Test on SUP dataset
  - Train on SUP, Test on 3DMAD dataset
- Evaluation metrics:
  - False Fake Rate (FFR)
  - False Liveness Rate (FLR)
  - Area Under Curve (AUC)
  - Equal Error Rates (EER)

# Experiments

## ■ Features for Comparison

- Appearance-based features:
  - Multi-scale LBP (MS LBP for short) [1]
  - Color Texture (CT for short) [2]
  - Deep features from last fully connected layer of CNN (fc CNN for short) [3]
  - Image distortion analysis features (IDA for short) [4]
- Motion-based features:
  - LBPTOP features [5]
  - Multifeature videolet aggregation (Videolet for short)[6]
  - Optical flow field (OFF for short) [7]
  - Optical flows on Gabor features [8](OF Gabor for short)
  - Optical flows on raw images (OF raw for short)
- Liveness-based features:
  - rPPG [9]

[1] Spoofing face recognition with 3D masks., 2014. IEEE transactions on information forensics and security, 2014

[2] Face spoofing detection using color texture analysis., 2016. IEEE Transactions on Information Forensics and Security, 2016

[3] Learn convolutional neural network for face anti-spoofing., 2014. arXiv

[4] Face spoof detection with image distortion analysis., 2015. IEEE Transactions on Information Forensics and Security

[5] Face liveness detection using dynamic texture., 2014. EURASIP Journal on Image and Video Processing, 2014

[6] Face anti-spoofing with multifeature videolet aggregation., 2016. Pattern Recognition, 2016 23rd International Conference on.

[7] A liveness detection method for face recognition based on optical flow field., 2009. International Conference on Image Analysis and Signal Processing.

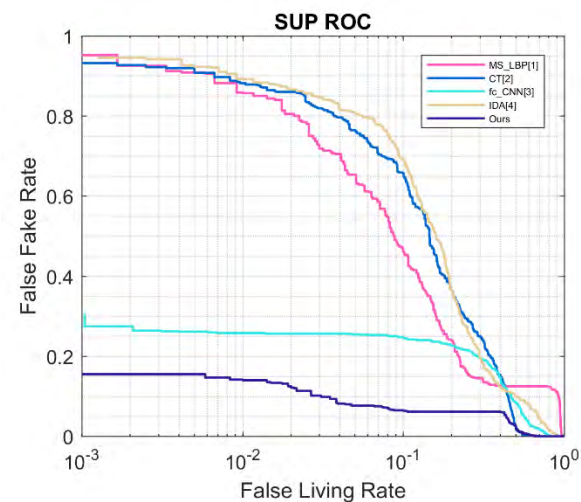
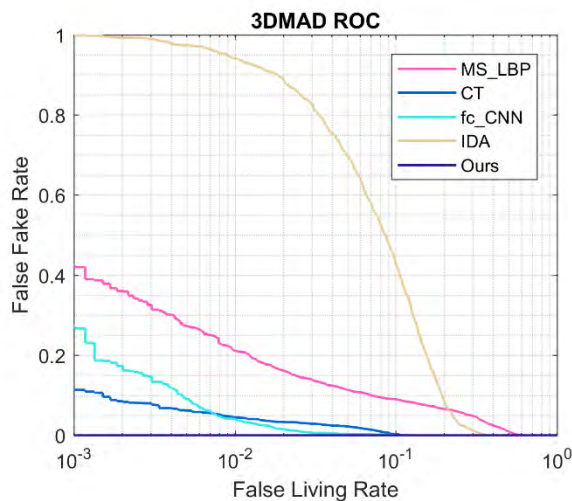
[8] Nonlinear operator for oriented texture., 1999. IEEE Transactions on image processing.

[9] 3D mask face anti-spoofing with remote photoplethysmography., 2016. European Conference on Computer Vision. Springer International Publishing

# Experiments

## ■ Intra-dataset Evaluation

Appearance-based features comparison:



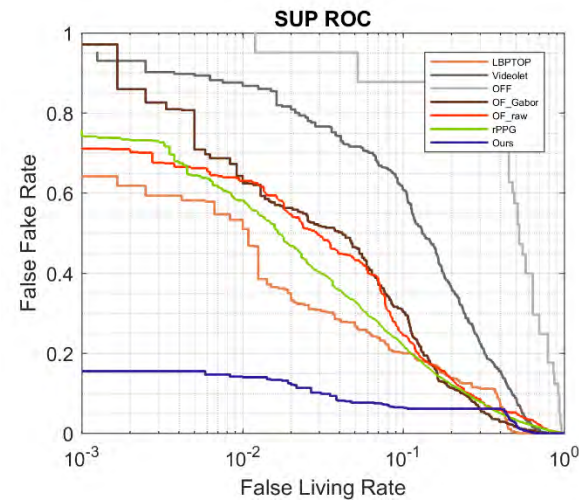
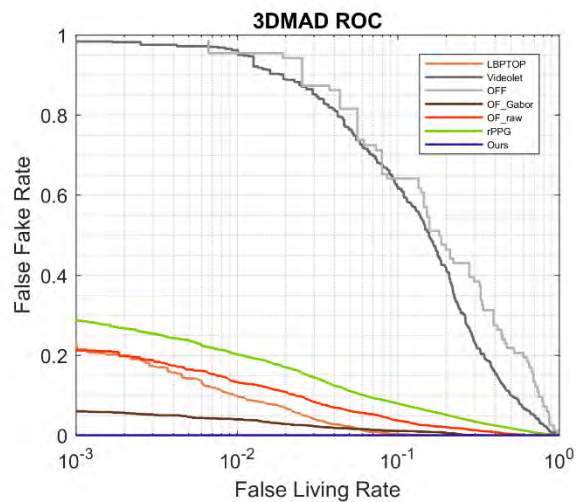
- Hand-crafted features and deep features of last fully-connected layer of CNN can achieve good results in 3DMAD dataset which are comparable with our method, but the performance of these methods drop a lot in supplementary dataset  
-> **Appearance-based features are not discriminative enough to capture subtle texture differences when facing masks with good appearance quality**



# Experiments

## ■ Intra-dataset Evaluation

Motion-based features and other cues-based features comparison:



- Similarly, the intrinsic limitation of the hand-crafted feature leads to the same degraded performance of motion-based features and other cues-based features
  - > Motion-based features and other cues-based features are not descriptive enough for **subtle** motion differentiation

# Experiments

## ■ Intra-dataset Evaluation

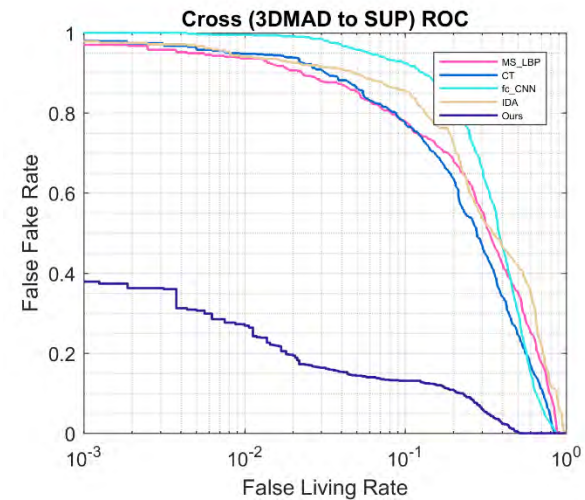
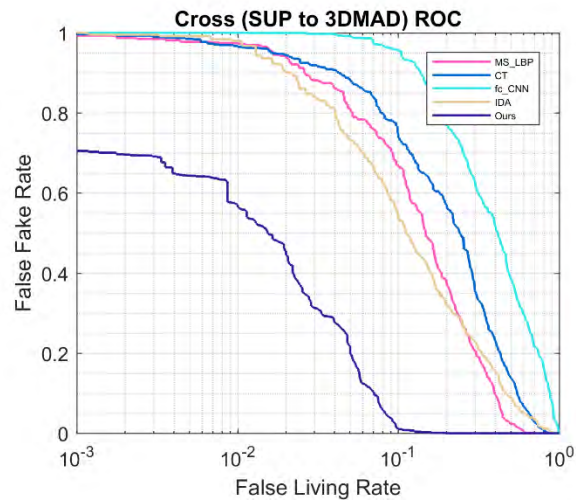
TABLE II: Experimental results of False Fake Rate at chosen False Living Rate on 3DMAD and SUP under intra-dataset test protocol.

Method	3DMAD dataset			SUP dataset		
	FFR @FLR=0.001	FFR @FLR=0.01	FFR @FLR=0.1	FFR @FLR=0.001	FFR @FLR=0.01	FFR @FLR=0.1
MS_LBP[14]	42.11	21.08	9.01	94.60	85.66	46.00
CT[6]	11.37	4.45	0.24	93.05	87.93	65.44
fc_CNN[38]	26.77	3.98	0.03	27.53	25.79	24.51
IDA[36]	99.83	94.09	42.45	94.61	89.16	69.15
LBPTOP[12]	21.52	9.64	0.30	63.52	51.00	19.95
Videolet[30]	98.30	95.09	61.72	93.64	86.75	60.73
OFF[3]	97.89	95.18	64.13	99.34	95.47	87.69
OF_Gabor	5.98	3.96	1.07	94.77	62.41	30.35
OF_raw	21.19	13.31	3.69	71.05	63.00	24.40
rPPG[24]	28.76	20.15	7.97	75.73	58.13	22.08
DTAC[29]	28.34	7.86	0.20	63.06	39.96	14.16
<b>Ours</b>	<b>0</b>	<b>0</b>	<b>0</b>	<b>15.50</b>	<b>14.00</b>	<b>6.50</b>

# Experiments

## ■ Cross-dataset Evaluation

Appearance-based features comparison:

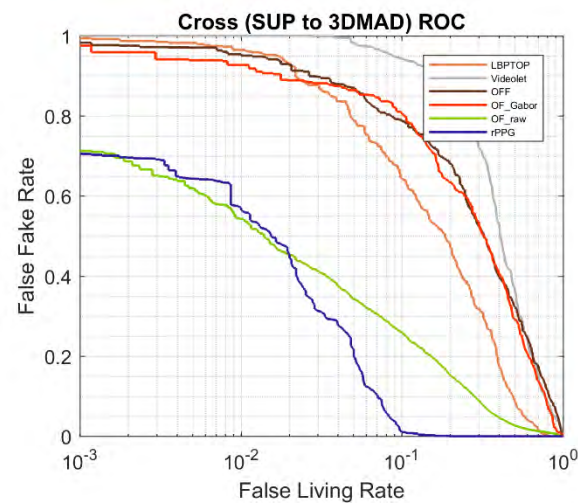
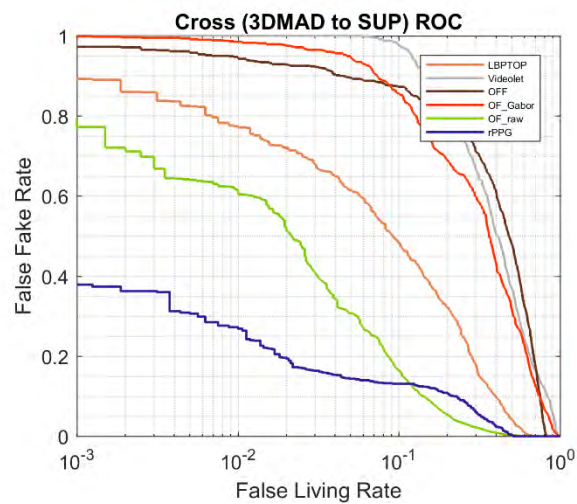


- The proposed method generalizes well between different masks
- The existing appearance-based methods have limited generalization ability

# Experiments

## ■ Cross-dataset Evaluation

Motion-based features and other cue-based features comparison:



- The proposed method is more able to find invariant features cross the datasets than motion-based features and other cues-based features

# Experiments

## ■ Cross-dataset Evaluation

TABLE III: Experimental results of False Fake Rate at chosen False Living Rate on 3DMAD and SUP under cross-dataset test protocol.

Method	3DMAD to SUP dataset			SUP to 3DMAD dataset		
	FFR @FLR=0.001	FFR @FLR=0.01	FFR @FLR=0.1	FFR @FLR=0.001	FFR @FLR=0.01	FFR @FLR=0.1
MS_LBP[14]	96.95	93.62	77.87	99.29	97.29	66.82
CT[6]	97.96	94.75	77.25	99.64	96.47	73.76
fc_CNN[38]	100	99.50	92.37	100	100	95.64
IDA[36]	97.61	94.00	85.62	99.82	97.94	54.64
LBPTOP[12]	89.44	77.12	48.12	99.31	96.47	64.23
Videolet[30]	100	100	97.62	100	100	94.35
OF_Gabor	97.24	94.25	87.25	98.06	95.05	78.70
OF_raw	99.79	98.50	85.50	96.94	92.70	80.47
rPPG[24]	79.32	61.62	16.47	72.94	<b>54.31</b>	25.97
DTAC[29]	70.77	55.37	18.37	83.63	74.76	36.64
<b>Ours</b>	<b>37.35</b>	<b>26.87</b>	<b>13.12</b>	<b>70.46</b>	56.43	<b>1.05</b>



# Experiments

## ■ Intra-dataset and cross-dataset Evaluation

TABLE I: Experimental results of AUC curve and EER data on 3DMAD and SUP under intra-dataset and cross-dataset test protocol.

Method	3DMAD dataset		Supplementary dataset		3DMAD to SUP dataset		SUP to 3DMAD dataset	
	EER(%)	AUC(%)	EER(%)	AUC(%)	EER(%)	AUC(%)	EER(%)	AUC(%)
MS_LBP[14]	9.14	96.71	21.17	80.29	41.00	62.35	26.12	81.67
CT[6]	2.92	99.74	27.00	81.40	37.16	67.59	32.65	73.59
fc_CNN[38]	1.77	99.82	21.98	89.17	42.63	60.89	45.25	54.69
IDA[36]	16.57	90.25	25.67	79.27	44.38	57.82	25.82	80.05
LBPTOP[12]	3.46	99.60	16.50	92.71	24.97	84.99	31.06	77.67
Videolet[30]	27.19	78.69	26.81	81.31	44.88	55.83	43.58	55.60
OFF <sup>1</sup> [3]	33.43	70.72	52.31	46.54	—	—	—	—
OF_Gabor	2.55	99.67	15.98	91.69	49.00	55.30	39.88	62.27
OF_raw	5.68	98.61	16.26	90.65	41.00	61.65	40.51	64.18
rPPG[24]	8.59	96.81	15.38	92.03	12.25	94.89	17.67	91.83
DTAC[29]	2.66	99.69	12.54	95.94	13.03	94.98	18.00	90.21
<b>Ours</b>	<b>0</b>	<b>100</b>	<b>7.33</b>	<b>96.68</b>	<b>12.75</b>	<b>95.64</b>	<b>7.60</b>	<b>97.44</b>

<sup>1</sup> The method of OFF [3] does not need training process and thus the cross-dataset test is not necessary for this method.



# Section Conclusion

- We propose to estimate multiple deep dynamic textures from fine-grained textures in different feature channels of a convolutional layer of CNN for 3D mask face anti-spoofing, which are able to fully capture the subtle facial motion differences and have a strong generalization ability.
- We propose a joint discriminative learning model of the multiple deep dynamic textures to weight the discriminability of deep dynamic textures in different spatial regions and different feature channels.
- Experimental results on intra- and inter- dataset protocol shows that the proposed method has a good generalization ability.

# New dataset: HKBU-MARs

- new dataset: **HKBU-MARs**
- <http://rds.comp.hkbu.edu.hk/mars>



(a) Room light

(b) Low light

(c) Bright light



(d) Warm light

(e) side light

(f) Up side light



(a)

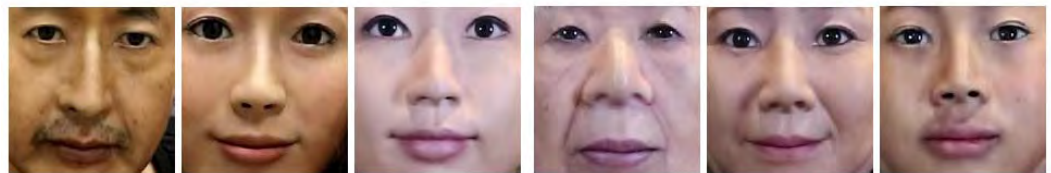
(b)

(c)

(d)

(e)

(f)



(g)

(h)

(i)

(j)

(k)

(l)

# **Part II:**

# **Face Template Protection**

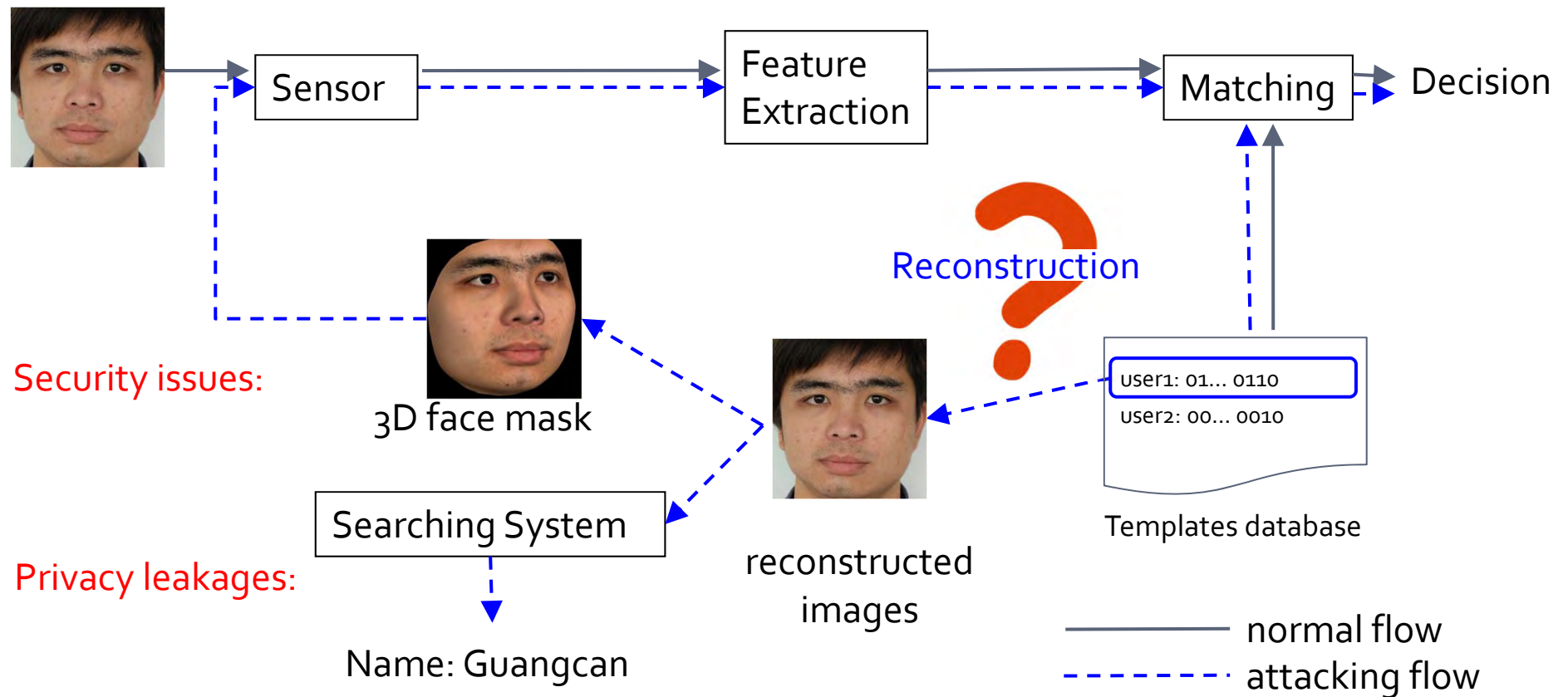
# Outline: Face Template Protection

1. Face Image Reconstruction Attack
2. Face Template Protection
  - a. Hybrid approach
  - b. Binary Discriminative Analysis for binary template generation
  - c. Binary template fusion for multi-biometric cryptosystems
3. Other Security Research Issues
  - a. Entropy Measurement for Biometric Verification Systems

# Image Reconstruction Attack

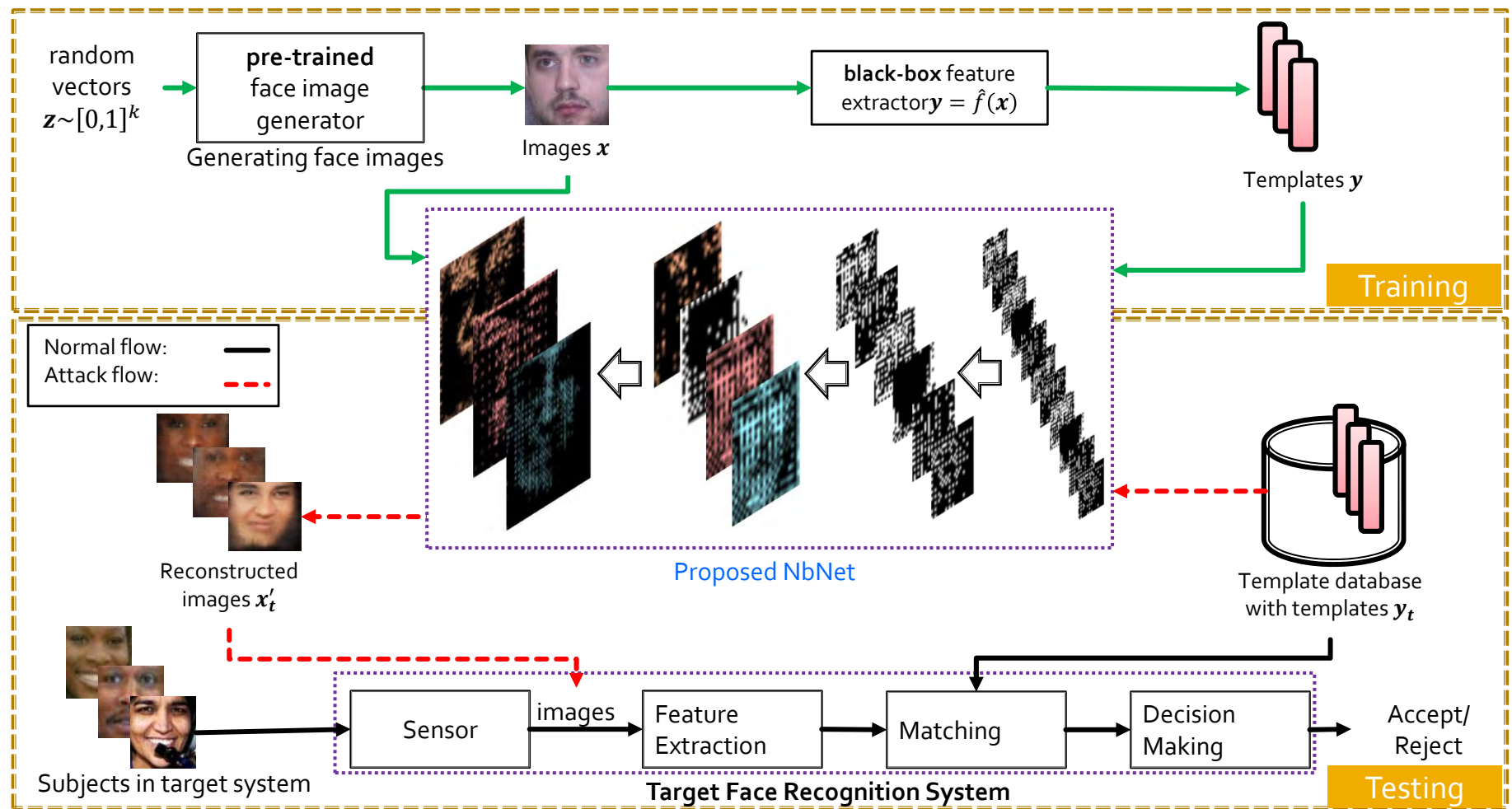
Can we reconstruct a fake face  
from templates?

# Image Reconstruction Attack

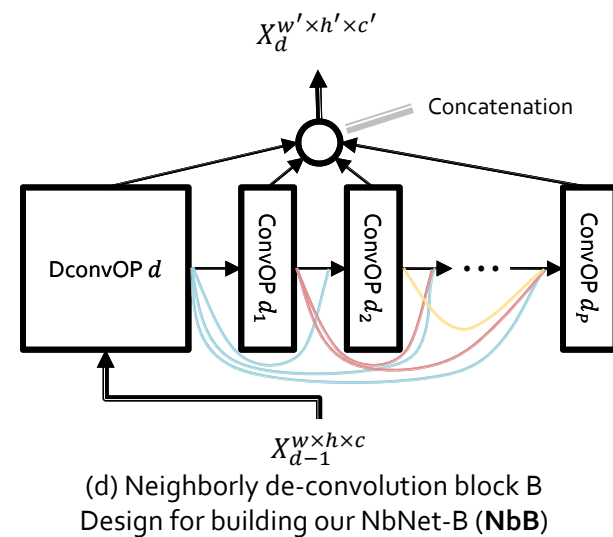
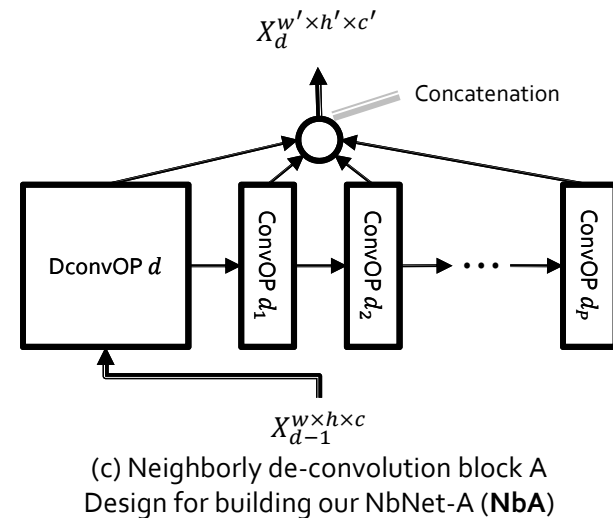
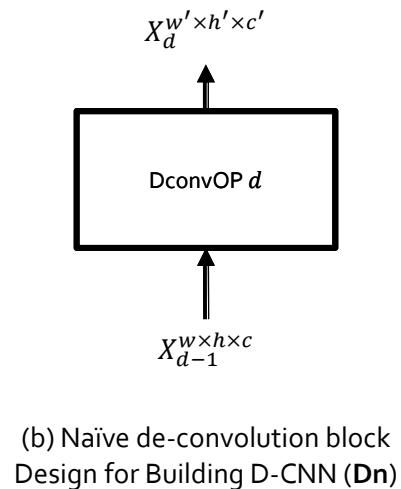
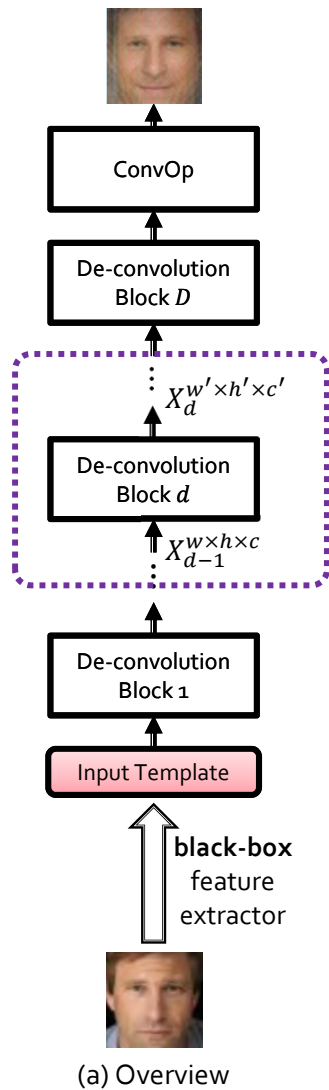




# From Deep Template to Face Image



# Proposed NbNet for Reconstruction



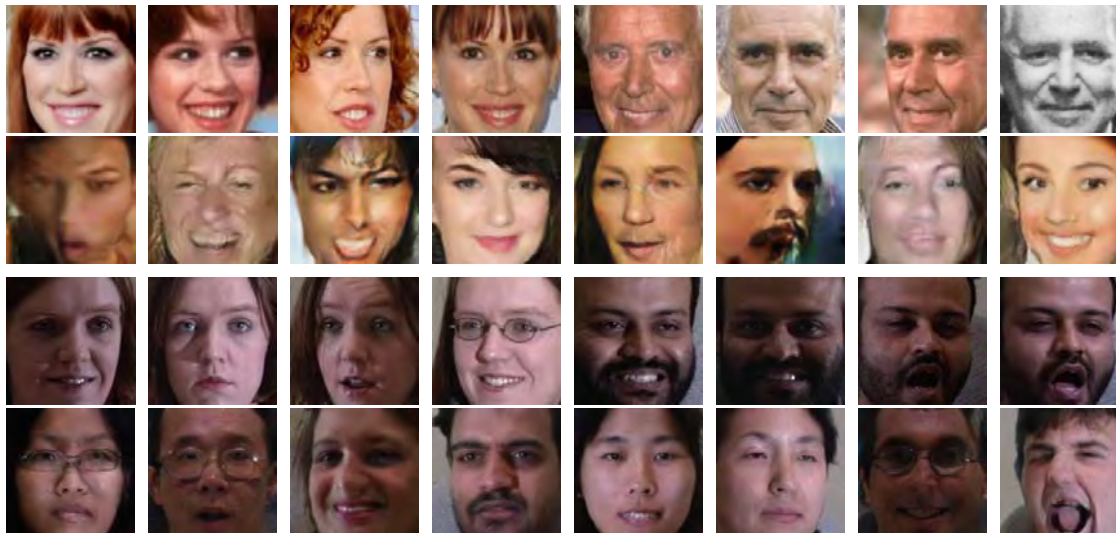
# Network Details

Layer name	Output size ( $c \times w \times h$ )	D-CNN	NbNet-A, NbNet-B
input layer	$128 \times 1 \times 1$		
De-convolution Block (1)	$512 \times 5 \times 5$	$[5 \times 5, 512]$ DconvOP, stride 2	$[5 \times 5, 256]$ DconvOP, stride 2 $\{[3 \times 3, 8]$ ConvOP, stride 1 $\} \times 32$
De-convolution Block (2)	$256 \times 10 \times 10$	$[3 \times 3, 256]$ DconvOP, stride 2	$[3 \times 3, 128]$ DconvOP, stride 2 $\{[3 \times 3, 8]$ ConvOP, stride 1 $\} \times 16$
De-convolution Block (3)	$128 \times 20 \times 20$	$[3 \times 3, 128]$ DconvOP, stride 2	$[3 \times 3, 64]$ DconvOP, stride 2 $\{[3 \times 3, 8]$ ConvOP, stride 1 $\} \times 8$
De-convolution Block (4)	$64 \times 40 \times 40$	$[3 \times 3, 64]$ DconvOP, stride 2	$[3 \times 3, 32]$ DconvOP, stride 2 $\{[3 \times 3, 8]$ ConvOP, stride 1 $\} \times 4$
De-convolution Block (5)	$32 \times 80 \times 80$	$[3 \times 3, 32]$ DconvOP, stride 2	$[3 \times 3, 16]$ DconvOP, stride 2 $\{[3 \times 3, 8]$ ConvOP, stride 1 $\} \times 2$
De-convolution Block (6)	$16 \times 160 \times 160$	$[3 \times 3, 16]$ DconvOP, stride 2	$[3 \times 3, 8]$ DconvOP, stride 2 $\{[3 \times 3, 8]$ ConvOP, stride 1 $\} \times 1$
ConvOP	$3 \times 160 \times 160$	$[3 \times 3, 3]$ ConvOP, stride 1	
Loss layer	$3 \times 160 \times 160$	Pixel difference or perceptual loss [36]	

$[k_1 \times k_2, c]$  DconvOP (ConvOP), stride  $s$  denotes cascade of a de-convolution (convolution) layer with  $c$  channels, kernel size  $(k_1 \times k_2)$ , and stride  $s$ , batch normalization and ReLU (tanh for the bottom ConvOP) activation layer

# Experiments - Training

- **Feature extractor** [1], an implementation of FaceNet [2]
- **Network architecture**: D-CNN (Dn), NbNet-A (NbA), NbNet-B (NbB)
- **Training approach**: Generated images & Raw images (r)
- **Loss function**: Pixel difference (M) & Perceptual Loss [3] (P)
- **Training datasets**:



VGG Raw Images: 1.94 M

VGG Gen Images: 19.2 M

Multi-PIE Raw Images: 150,760

Multi-PIE Gen Images: 19.2 M

1. <https://github.com/davidsandberg/facenet> (model: 20170512-110547)
2. Schroff, Florian et al. "Facenet: A unified embedding for face recognition and clustering." *CVPR2015*
3. Johnson et. al., "Perceptual losses for real-time style transfer and super-resolution", *ECCV2016*

# Experiments

- Verification (protocol: BLUFR[1], comparison: RBF [2])
  - Type-I attack: match the reconstructed image against the same one from which representation was extracted
  - Type-II attack: match the reconstructed image against a different one of the same subject
- Identification (Rank-one identification rate)
  - Type-I attack: identify the images reconstructed from the gallery set
  - Type-II attack: identify the images reconstructed from the probe set
- Testing datasets



(a) LFW (Verification)



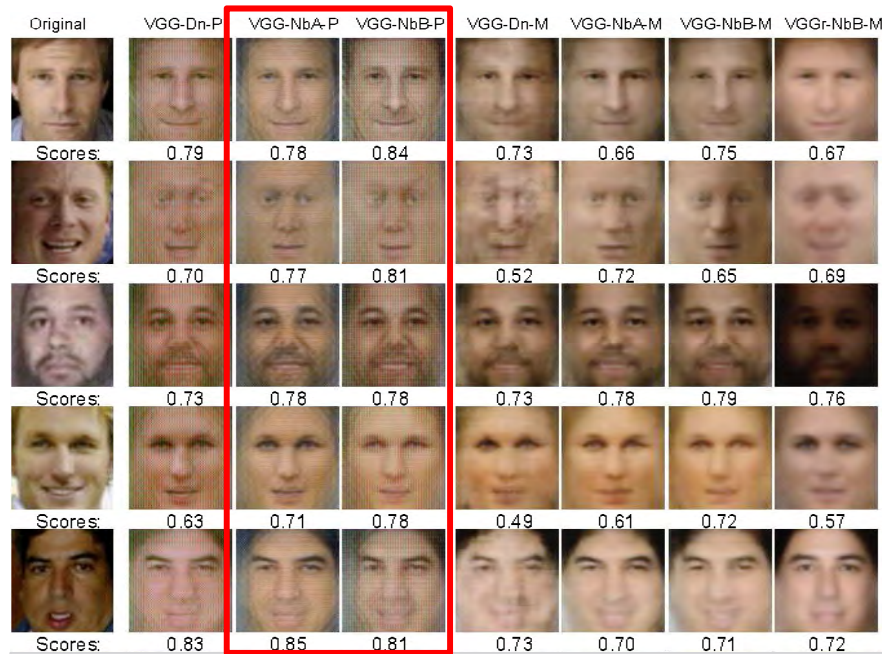
(b) FRGC V2.0 (Verification)



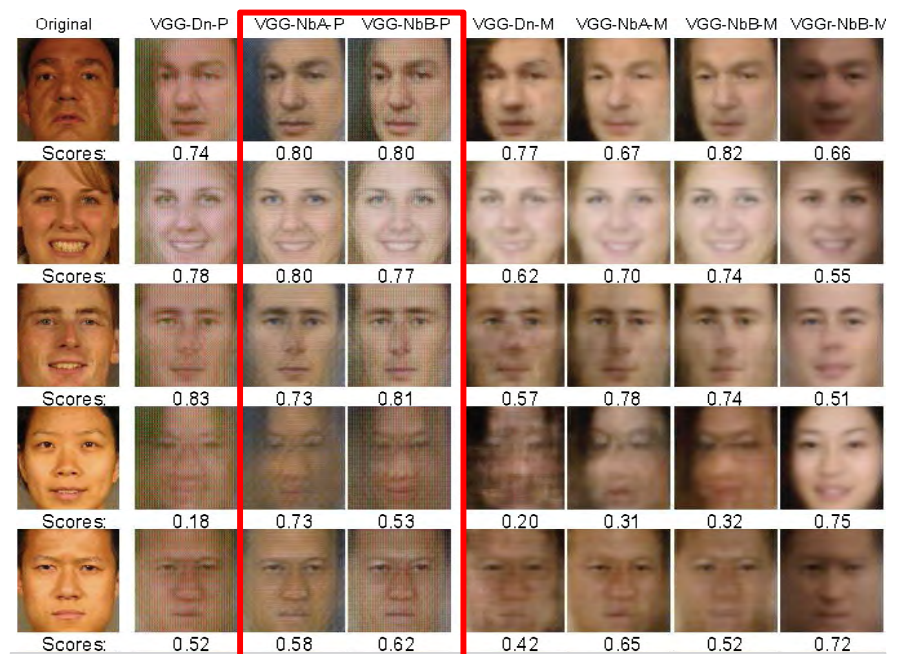
(c) Color FERET (Identification)



# Reconstruction of First 5 Subjects\*



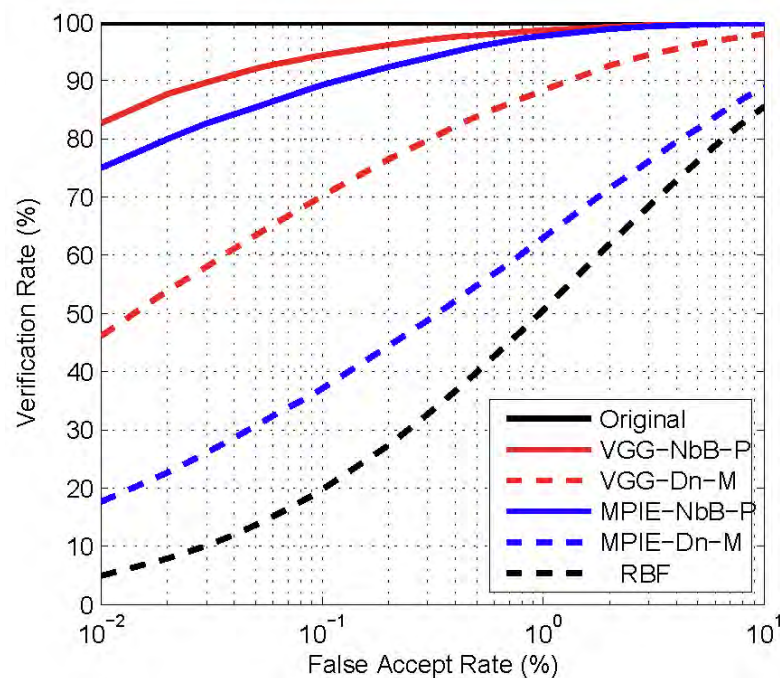
LFW Thresholds:  
 @(FAR=0.1%) = 0.51 @ (FAR=1%) = 0.38



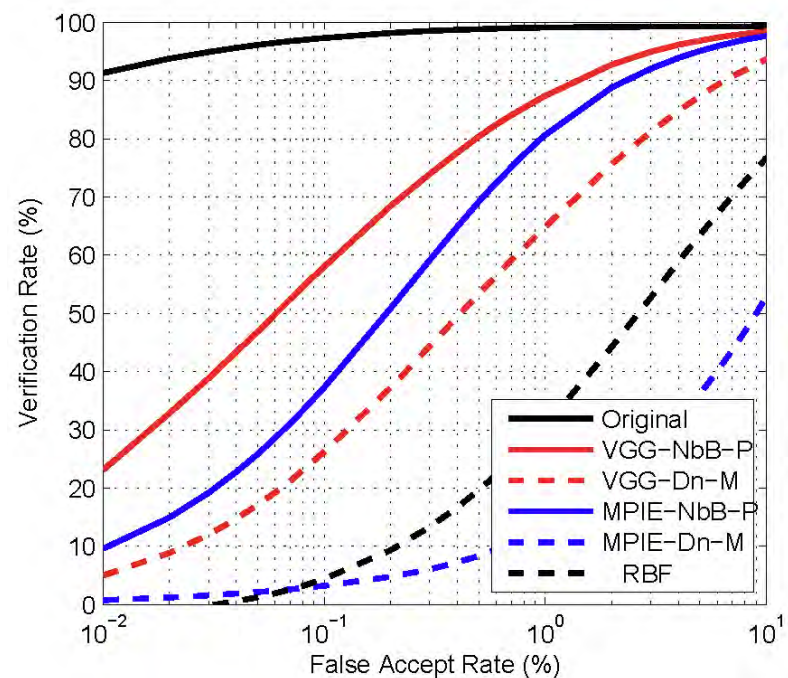
FRGC v2.0 Thresholds:  
 @(FAR=0.1%) = 0.80 @ (FAR=1%) = 0.64

- ❖ \* As specified in the image list of the BLUFR protocol [1]
- ❖ 'VGG-', 'MPIE-' denotes the face image generator is pretrained by the VGG-Face (2.6M) and MultiPIE (frontal images, 150K)
- ❖ 'VGGr-' denotes the NbNet directly trained by the raw images in VGG-Face, no face image generator is used.
- ❖ '-Dn-', '-NbA-', '-NbB-' denote the network architecture, i.e., D-CNN, NbNet-A and NbNet-B
- ❖ '-P' trained with perceptual loss
- ❖ '-M' trained with pixel-wise mean absolute error

# Experiments – Verification on LFW



(a) Type-I attack on LFW

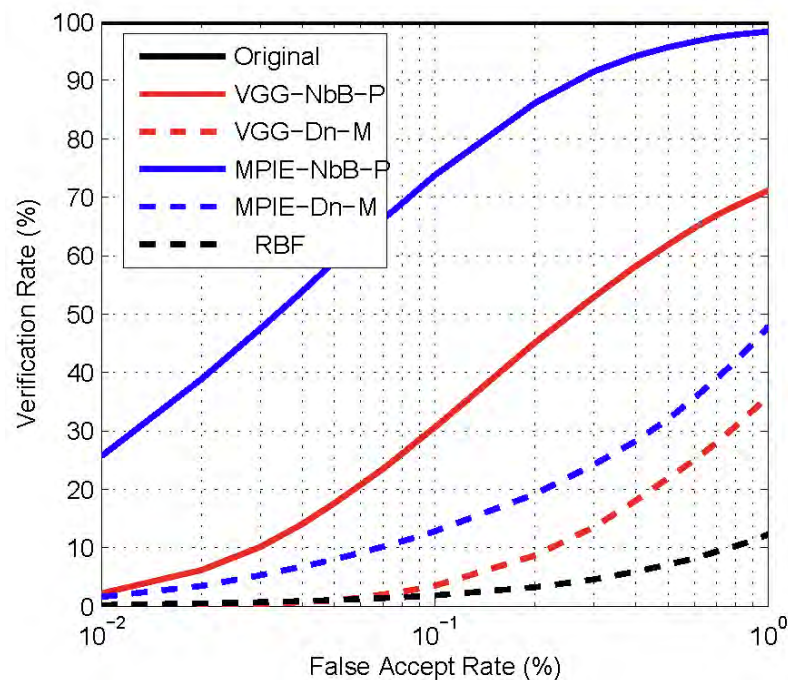


(b) Type-II attack on LFW

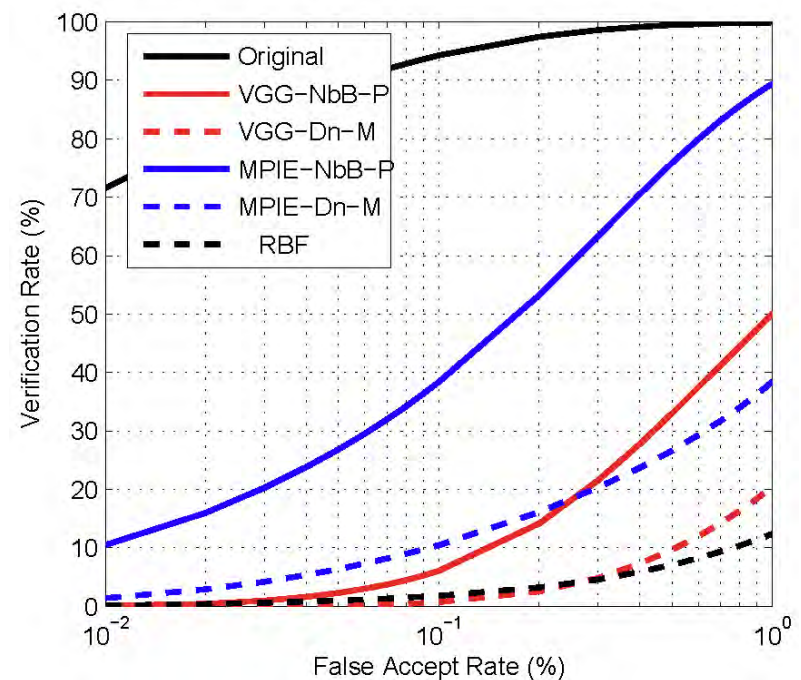
- Type-I attack: match the reconstructed image against the same one from which representation was extracted
- Type-II attack: match the reconstructed image against a different one of the same subject



# Experiments – Verification on FRGC



(a) Type-I attack on FRGC



(b) Type-II attack on FRGC

- Type-I attack: match the reconstructed image against the same one from which representation was extracted
- Type-II attack: match the reconstructed image against a different one of the same subject

# Identification with Reconstructed Images on Color FERET

Probe (partition specified by protocol)	Type-I	Type-II		
	fa	fb	dup1	dup2
Original	100.00	98.89	97.96	99.12
VGG-Dn-P	89.03	86.59	76.77	78.51
VGG-NbA-P	94.87	90.93	80.30	81.58
VGG-NbB-P	95.57	<b>92.84</b>	<u>84.78</u>	84.65
VGG-Dn-M	80.68	74.40	62.91	65.35
VGG-NbA-M	86.62	80.44	64.95	66.67
VGG-NbB-M	92.15	87.00	75	75.44
VGGr-NbB-M	81.09	74.29	61.28	62.28
MPIE-Dn-P	<u>96.07</u>	91.73	84.38	<u>85.53</u>
MPIE-NbA-P	93.86	90.22	79.89	79.82
MPIE-NbB-P	<b>96.58</b>	<b>92.84</b>	<b>86.01</b>	<b>87.72</b>
MPIE-Dn-M	73.54	64.11	53.26	49.12
MPIE-NbA-M	72.23	64.01	51.09	44.74
MPIE-NbB-M	85.61	78.22	71.06	68.42
MPIEr-NbB-M	63.88	54.54	44.57	35.96
Mixedr-NbB-M	82.19	76.11	62.09	58.77

Best:

**boldface**

Second best:

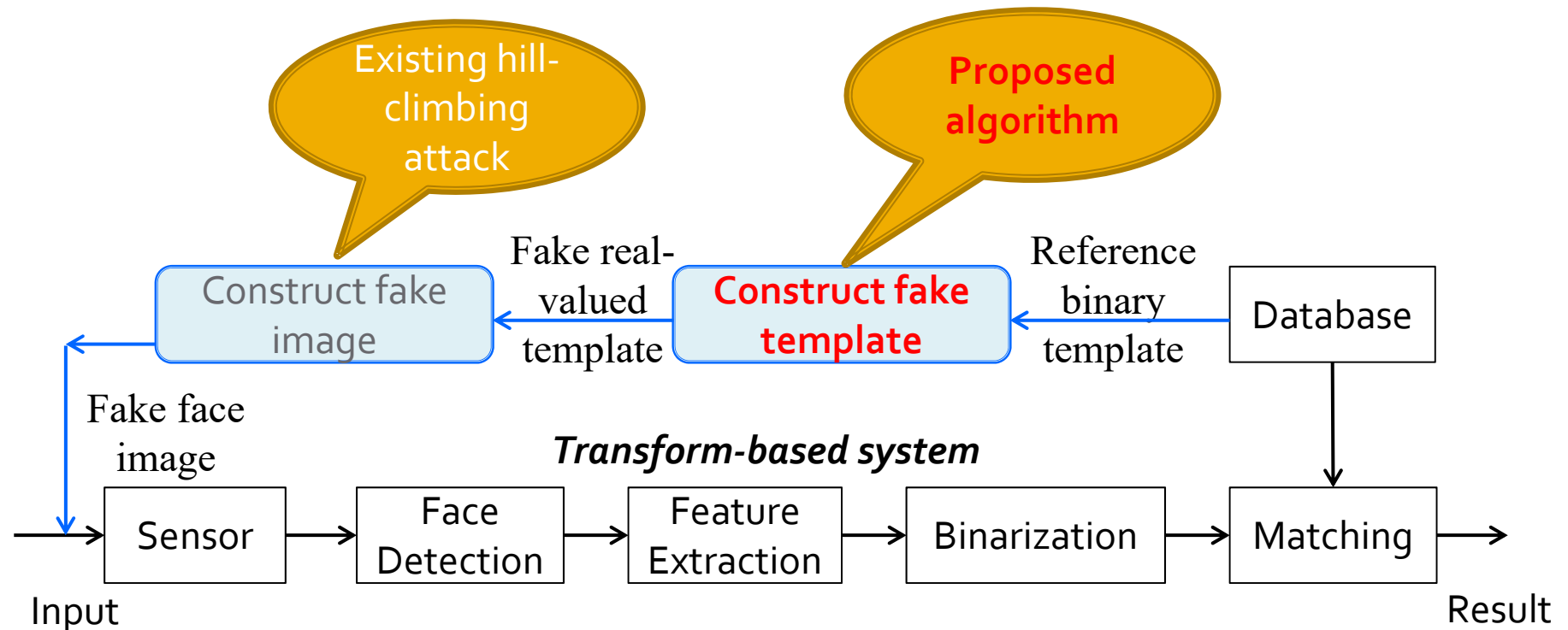
underline

Type-I attack:  
identify the images  
reconstructed from the  
gallery set (partition *fa*)

Type-II attack:  
identify the images  
reconstructed from the  
images which not used  
in the gallery set  
(partition *fb*, *dup1*, *dup2*)

# From Binary Template to Face Image

- Attack a transform-based system (partially protected, will be introduced later) with two steps



# From Binary Templates to Faces

- Consider two scenarios
  - The binarization scheme is understood by the attacker
  - The binarization scheme is unknown to the attacker
- Assumptions

The reference  
binary templates  
stored in database  
are exposed

*Masquerade attack  
assumptions*

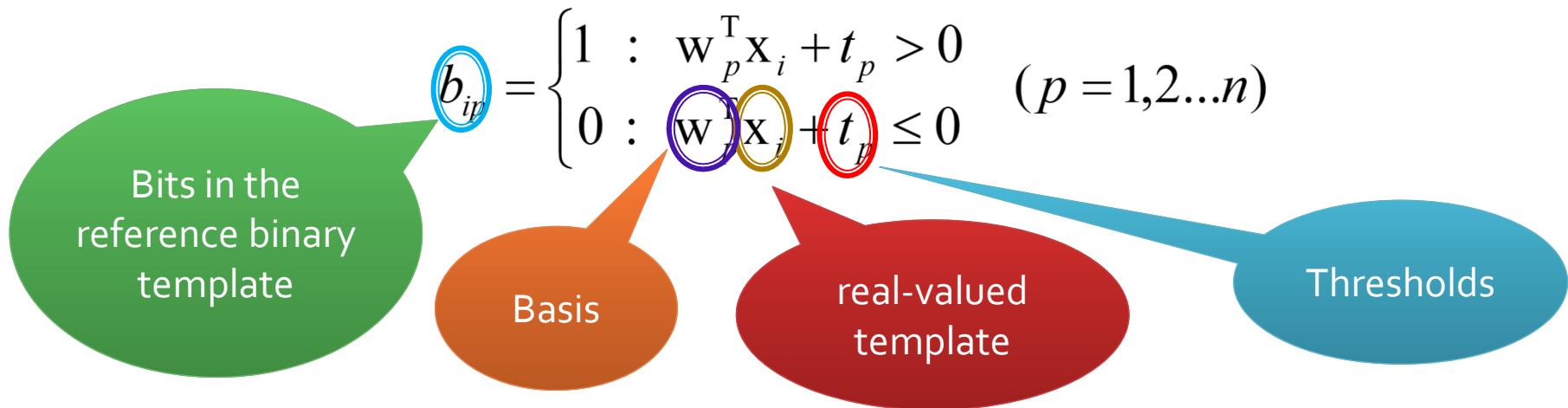
Attacker knows  
the feature  
extraction  
algorithm

Matching scores  
can be accessed

*Hill-climbing attack  
assumptions*

# Scenario One

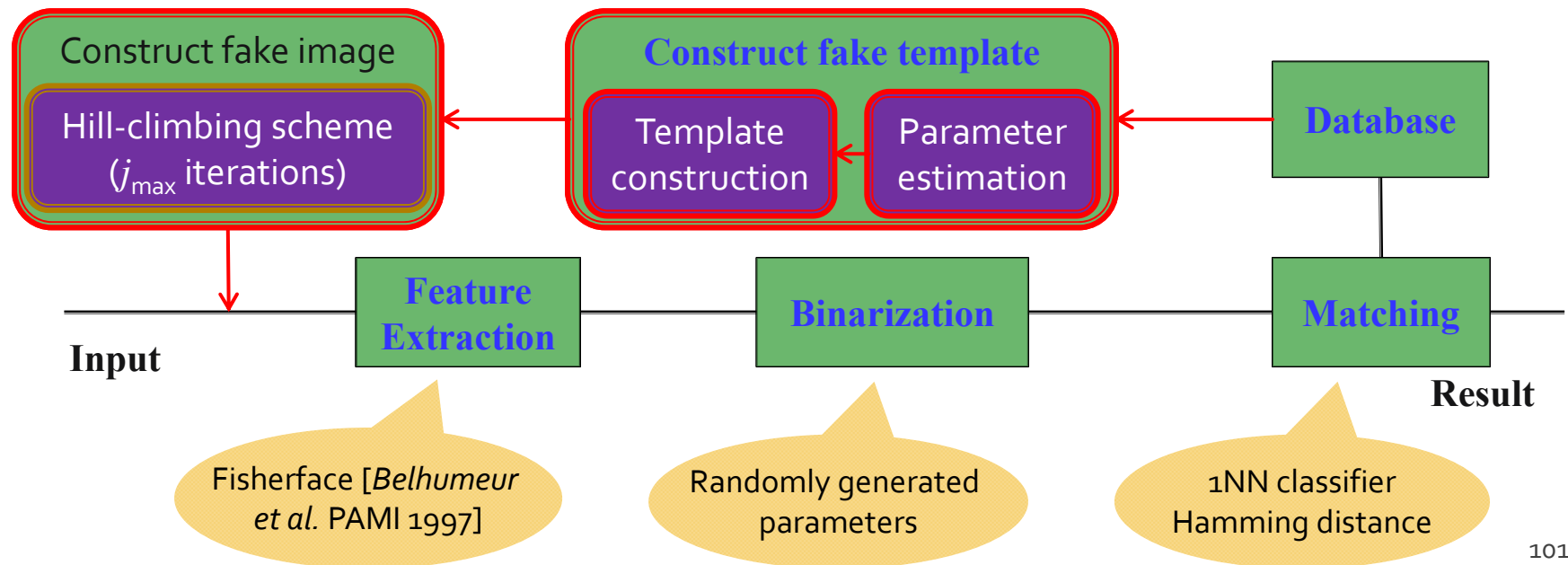
- Understand the binarization scheme
  - Most schemes follow “projection + thresholding” approach



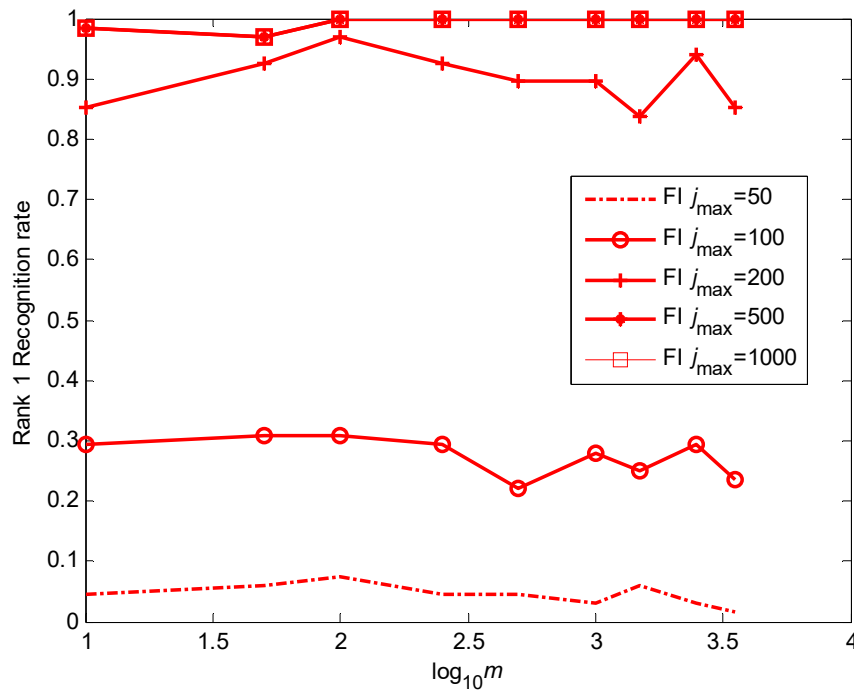
- Two steps to construct fake template
  - Binarization parameters estimation
  - Construct fake template with estimated parameters

# Scenario One

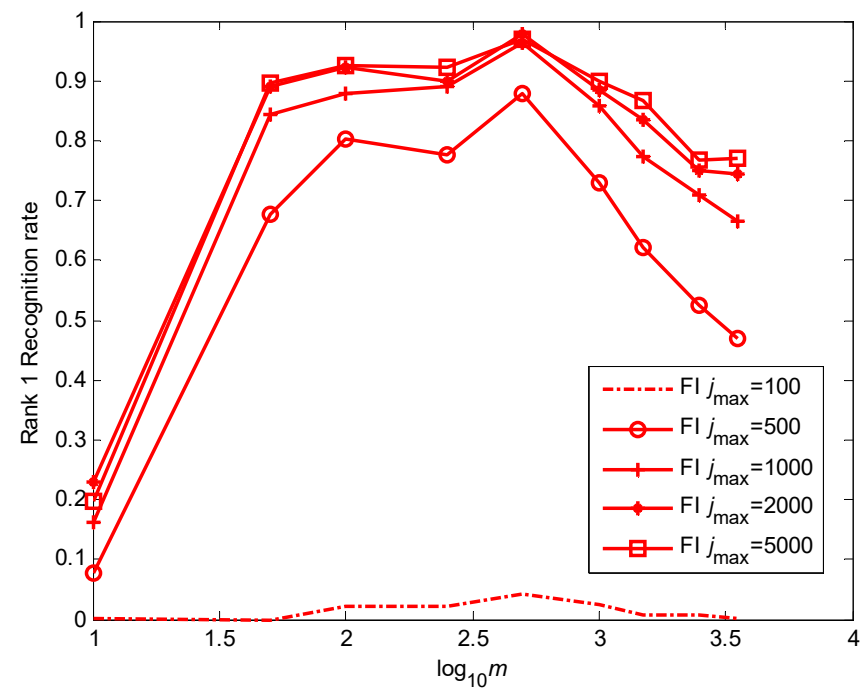
- Experimental results
  - Experiment settings
    - CMU PIE & FRGC databases employed
    - Choose different  $m$  (No. of local faces) in testing



# Scenario One – Results



CMU PIE



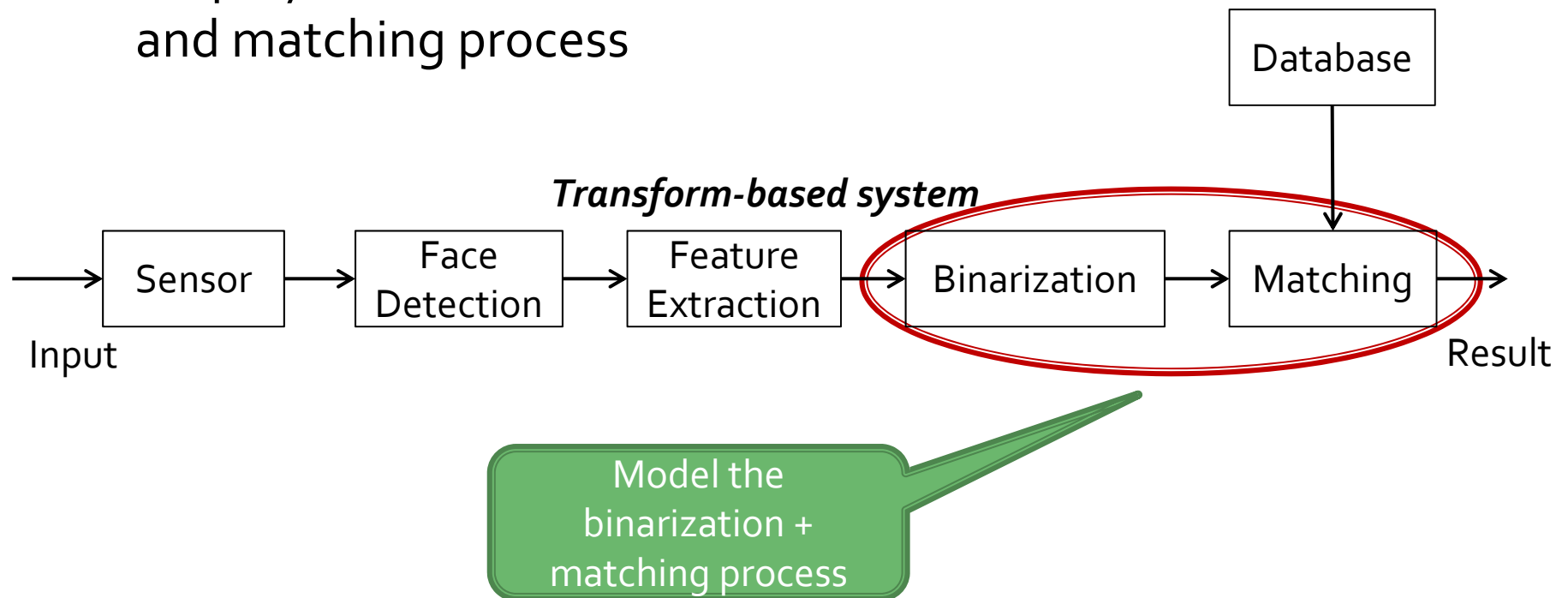
FRGC

Our proposed algorithm achieves high performance in both databases



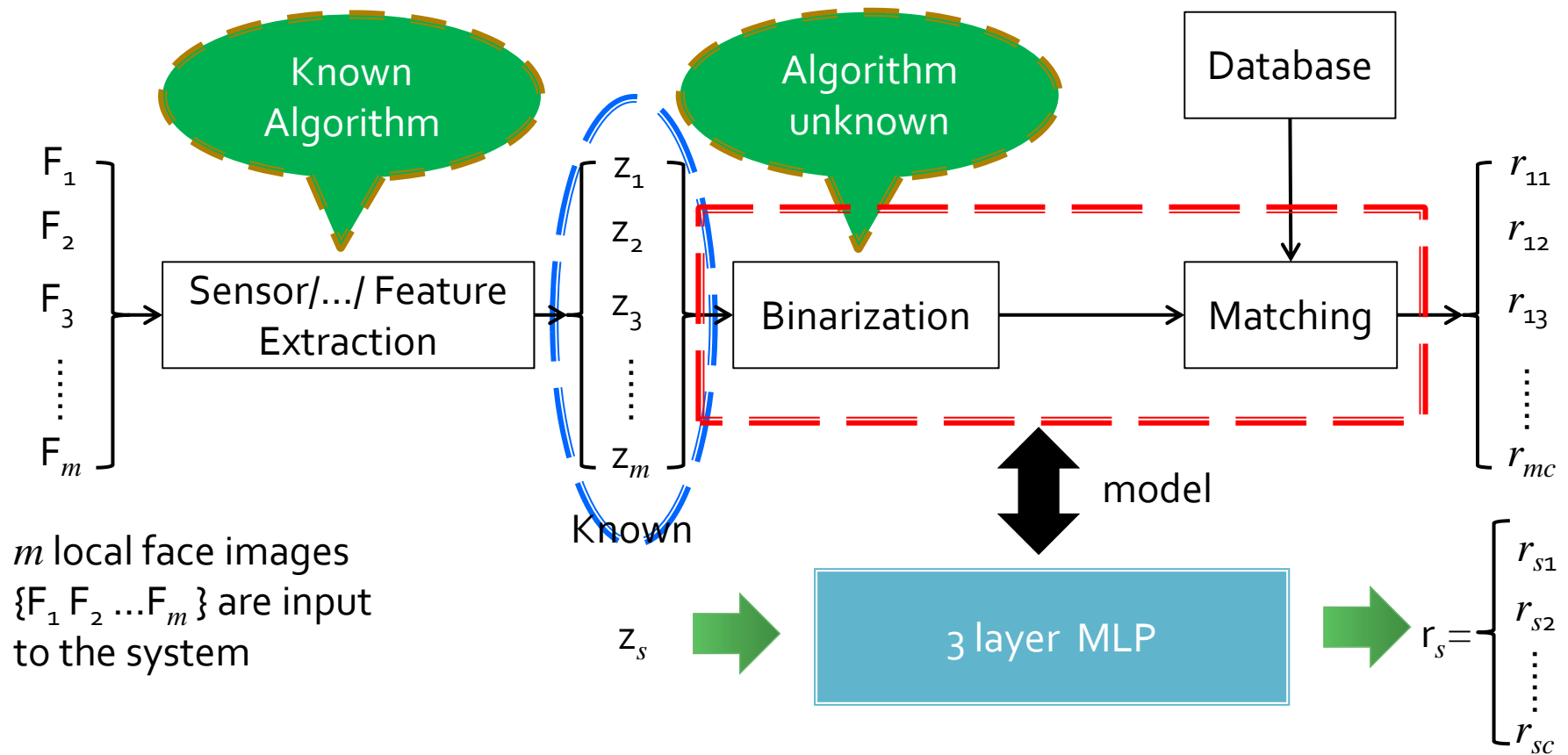
# Scenario Two

- Since the attacker does not understand the binarization algorithm, the binarization process needs to be modeled.
- Employ artificial neural networks to model the binarization and matching process



# Scenario Two

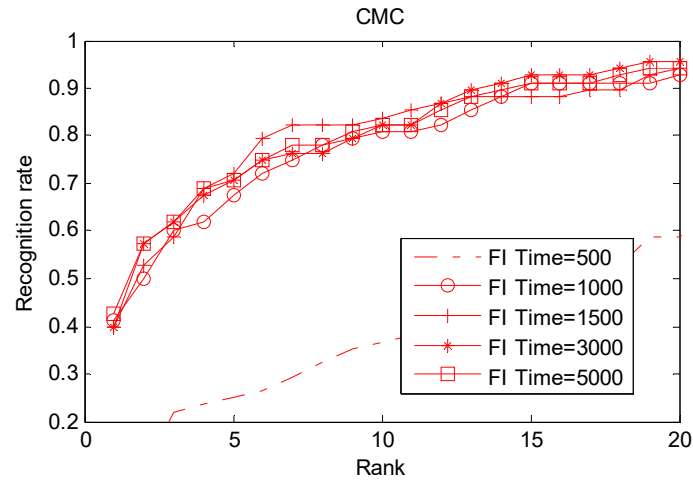
- Use local faces for modeling



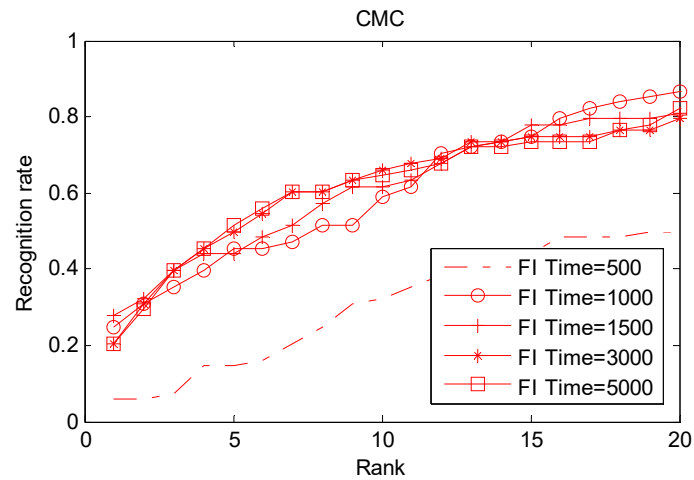
# Scenario Two

- Experimental results
  - Experiment settings
    - Follow the settings in scenario one
    - Implement the proposed attack in different binarization schemes
      - Biohashing (BH)
      - Multi-stage biohashing (MBH)
      - Feature binarization (FB)
      - Discriminability-preserving transform (DP)

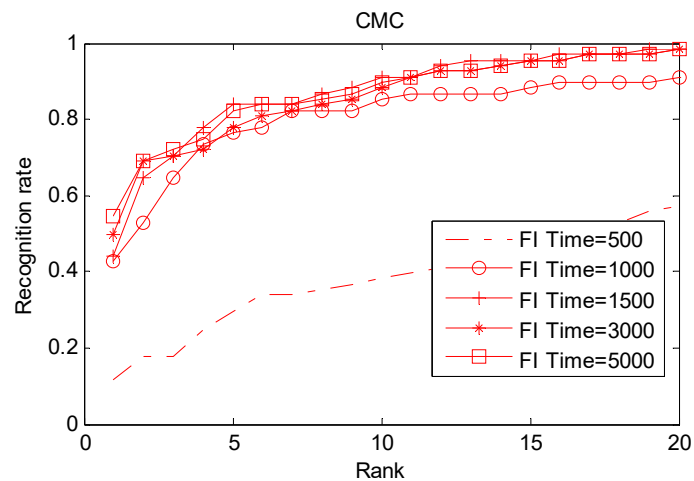
# Scenario Two – CMU-PIE Results



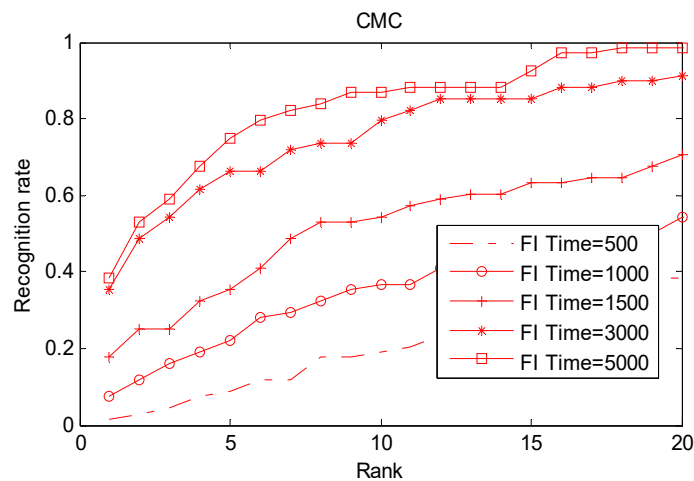
BH



MBH

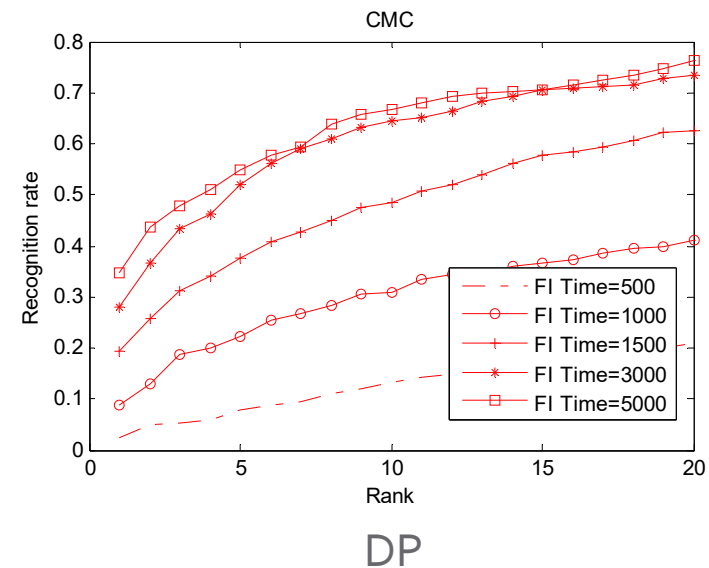
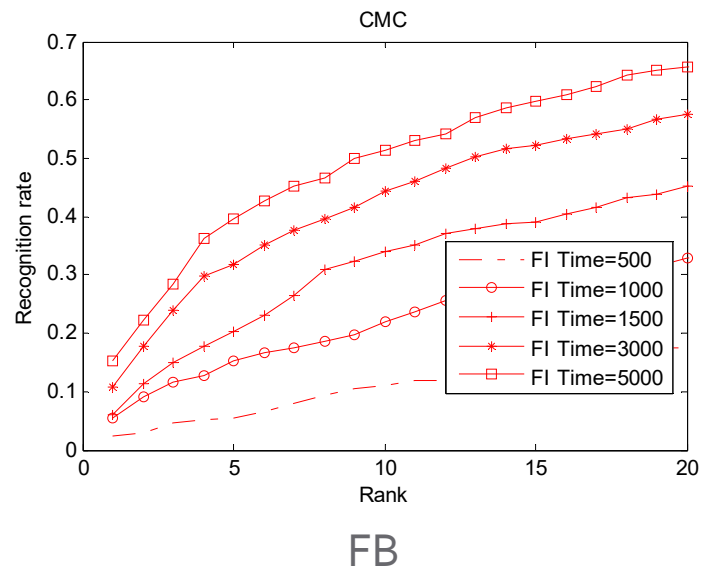
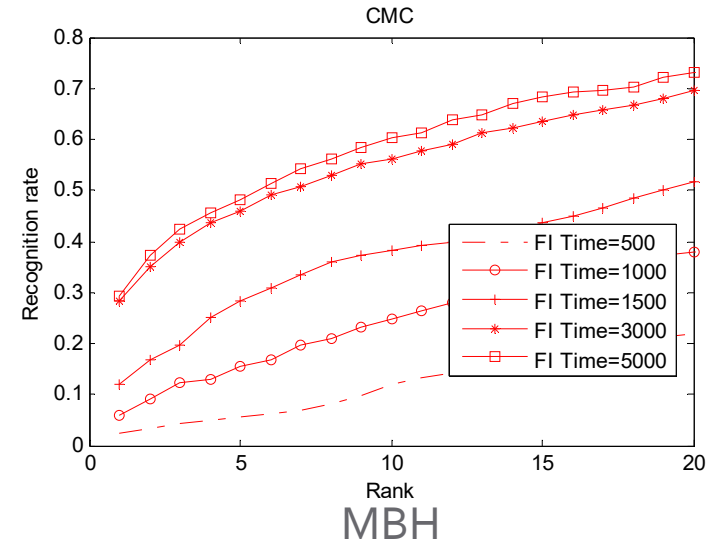
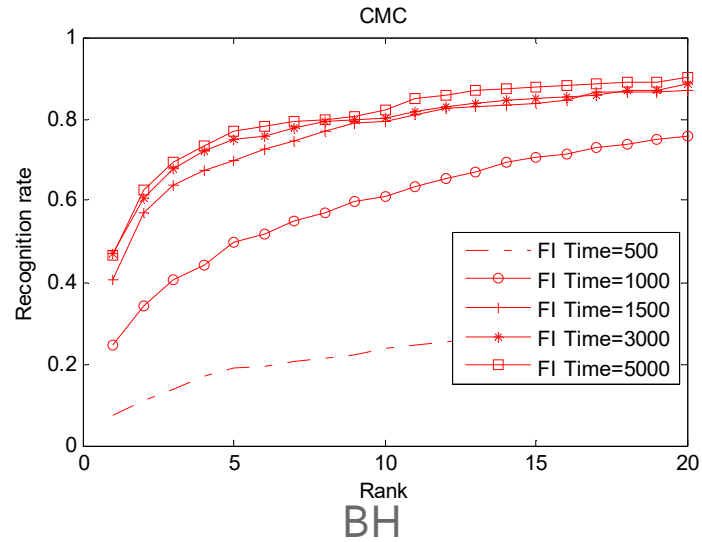


FB



DP

# Scenario Two – FRGC Results



# Section Conclusion

Both deep template and binary template are not save!

# Security and Privacy Issues

- If face template stored in database is compromised, it may cause security and privacy problems.

<i>Modification/ Replacing</i>	<ul style="list-style-type: none"><li>• Modify/replace the templates to the ones preferred by attackers.</li></ul>
Uniqueness	<ul style="list-style-type: none"><li>• Since biometric is “unique” feature for individual, it is hard to reset or re-issue</li></ul>
<i>Information Leakage</i>	<ul style="list-style-type: none"><li>• May cause fake face attacks</li></ul>
<i>Cross-platform matching</i>	<ul style="list-style-type: none"><li>• Templates stolen from one system may be used to attack another system</li></ul>





# Face Template Protection

# Requirements

## Security

- Computationally hard to reconstruct the original template from the secure template.

## Discriminability

- The discriminative power of the secure template should be as good as that of the original face template so that system performance will not be affected.

## Cancelability

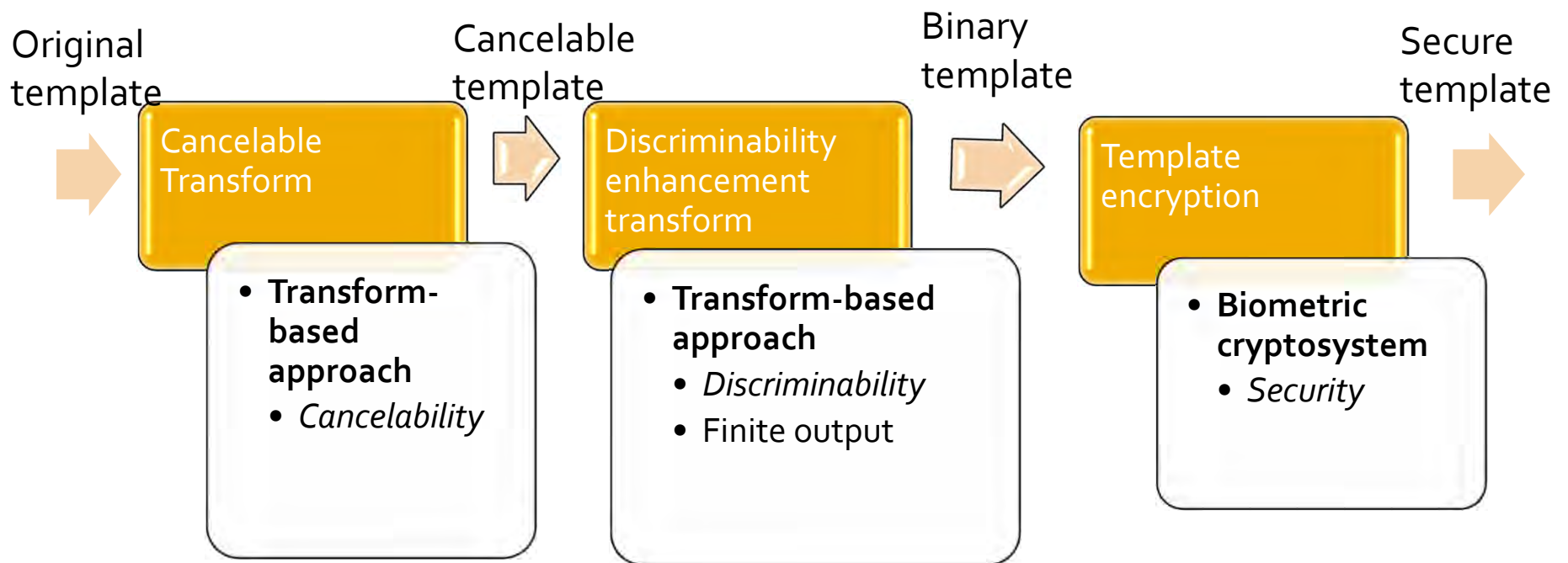
- The secure template can be canceled and re-issued from original template if it is stolen or lost.

# Basic Idea

- General approach: *Never* store the original raw biometric template
- Straightforward method: Protection with traditional encryption/hashing methods (e.g. DES, MD5)
  - Small change in input cause large change in output
  - Intra-class variations => not good for matching
  - Not feasible
- The template protection schemes can be classified into two categories (Jain *et al.* *EURASIP JASP* 2008).
  - *Biometric Cryptosystem*
  - *Transformation-based*

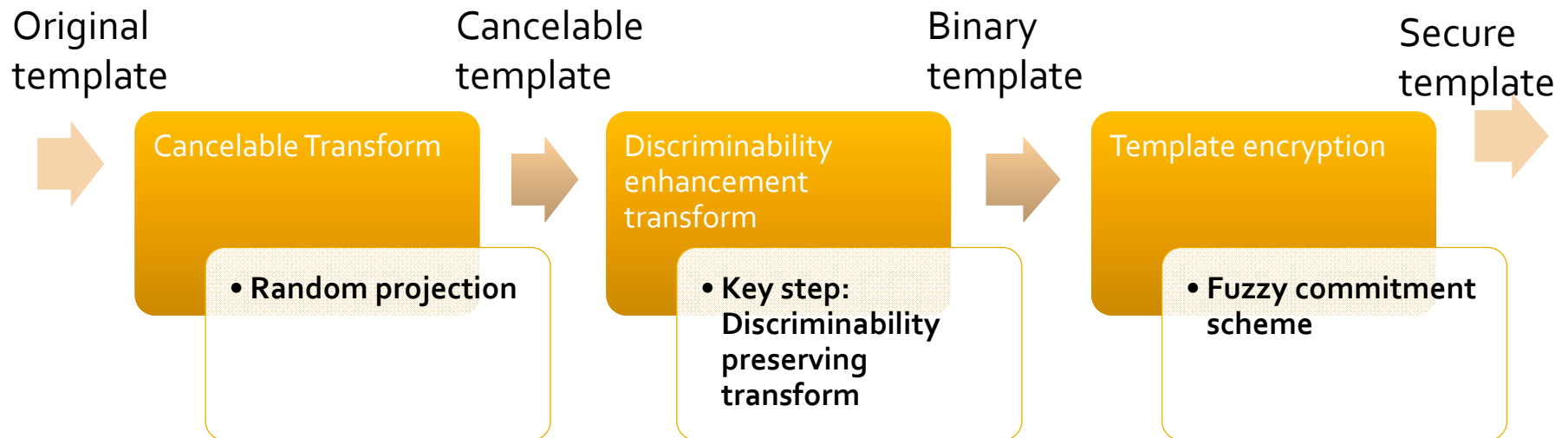
# Proposed Hybrid Framework [TIFS 2010]

- One single approach cannot achieve all security, discriminability and cancelability requirements
- A three-step hybrid approach: transformation-based biometric cryptosystem



# 3-step Algorithm

- The three-step hybrid algorithm



- The discriminability preserving transform should
  - Convert the cancelable template into binary template
  - Preserve the discriminability via transform.

# Experimental Results

- Experiment settings:

- Database:

$c$  : No. of individuals.

$m$ : No. of samples for each individual.

$q$  : No. of training samples per individual



CMU PIE



FERET



FRGC

Database	$c$	$m$	$q$	Variations
CMU PIE	68	105	10	Illumination, pose, expression
FERET	250	4	2	Mild expression, illumination
FRGC	350	40	5	expression, illumination, mild pose

# Experimental Results

- Experiment settings
  - Fisherface [Belhumeur *et al.* PAMI 1997] applied for feature extraction
  - Experiments
    - Template discriminability
    - Recognition accuracy
    - Cancelability

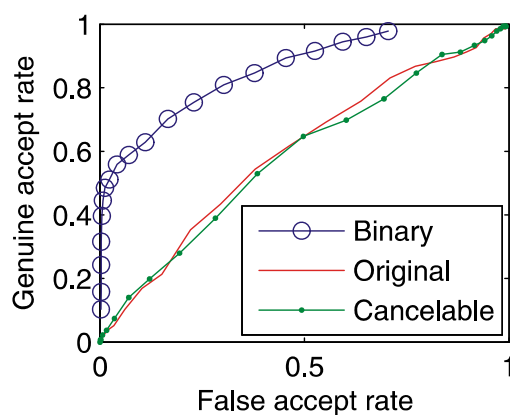


# Template Discriminability

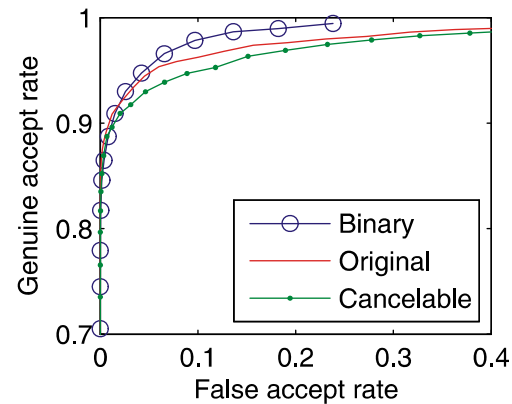
- Experimental settings
  - Choose three subsets from the CMU PIE database for experiments.
  - $kr$ : length of the cancelable templates
  - $kc$ : length of the binary templates

Database	$c$	$m$	$q$	$kr$	$kc$	Variations
CMU PIE-1	68	4	2	40	56	Pose
CMU PIE-2	250	21	4	40	84	Illumination
CMU PIE-3	350	105	10	40	210	Pose & illumination

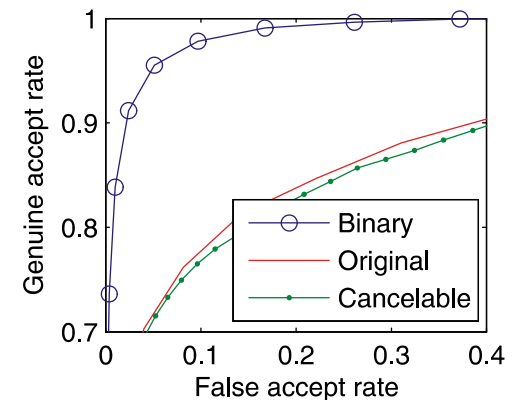
# Template Discriminability



(a) Pose



(b) Illumination



(c) Pose & Illumination

## ■ Observations

- Overlapping rate increased: Cancelable templates lightly degrade some discriminability
- Overlapping rate significantly decreased: binary templates enhance discriminability.
- The recognition performance conforms it.

# Recognition Accuracy

- Experimental settings
  - CMU PIE, FERET, FRGC databases used.

Database	$c$	$m$	$q$	$kr$	$kc$
CMU PIE	68	105	10	40	120, 150, 180, 210
FERET	250	4	2	150	120, 150, 180, 210
FRGC	350	40	5	250	150, 200, 250, 350

- Implement authentication with different  $kc$ . And comparing the performance with the
  - Original fisherface algorithm ("Original")
  - Random multispace quantization scheme ("RMQ-S") [Teoh *et al.* PAMI 2006]

# Recognition Accuracy

- In the transformed-based scheme (random projection), keys can be issued in two ways.
- Experiments are done in two scenarios
  - Common key scenario ("SRC")
  - User-specified key scenario ("DRC")

# Common-key Scenario

- Observation
  - The proposed hybrid algorithm outperforms the original fisherface and the RMQ algorithm

EER(%)	Fisherface	<i>kc-1</i>	<i>kc-2</i>	<i>kc-3</i>	<i>kc-4</i>	RMQ
CMU PIE	18.18	7.61	7.30	6.95	6.81	11.93
FERET	12.58	9.52	8.86	8.61	8.55	12.83
FRGC	31.75	17.93	17.40	16.70	16.68	21.87

# User-specified Key Scenario

- Observation
  - The proposed hybrid algorithm outperforms the original fisherface and the RMQ algorithm

EER(%)	Fisherface	<i>kc-1</i>	<i>kc-2</i>	<i>kc-3</i>	<i>kc-4</i>	RMQ
CMU PIE	18.18	9.41	8.41	8.70	8.26	11.68
FERET	21.66	3.38	3.36	3.34	3.62	4.49
FRGC	31.75	9.03	9.18	9.08	9.13	11.03

# Binary Template Generation [TIFS 2012]

- The discriminability of the binary templates receives little attention

Insufficient  
information carried  
in binary templates

Binarization  
process may cause  
distortion

Matching  
performed in  
Hamming space

Ad-hoc  
thresholding/  
quantization

Feng [1] optimizes  
discriminability, but  
with user-specific  
parameters

1. Y C Feng and P C Yuen, "Binary Discriminant Analysis for Generating Binary Face Template," *IEEE Transactions on Information Forensics and Security*, vol. 7, no. 2, pp.613-624, 2012.



# Rationale

- Existing schemes lack of discriminability evaluations of the binary templates
- Traditional discriminability optimization methods are not effective
  - Employ differentiation
  - Differentiation is not feasible in Hamming space
- Propose a **binary discriminant analysis (BDA)** to optimize the discriminability of the binary templates

# Rationale

- Use a series of **linear discriminant functions (LDF)** to form a binary template  $b=(b_1, b_2 \dots b_i \dots b_k)$  from input sample  $x$ .

$$b_i(x) = \begin{cases} 0 & \text{if } w_i^T x + t_i > 0 \\ 1 & \text{if else} \end{cases}$$

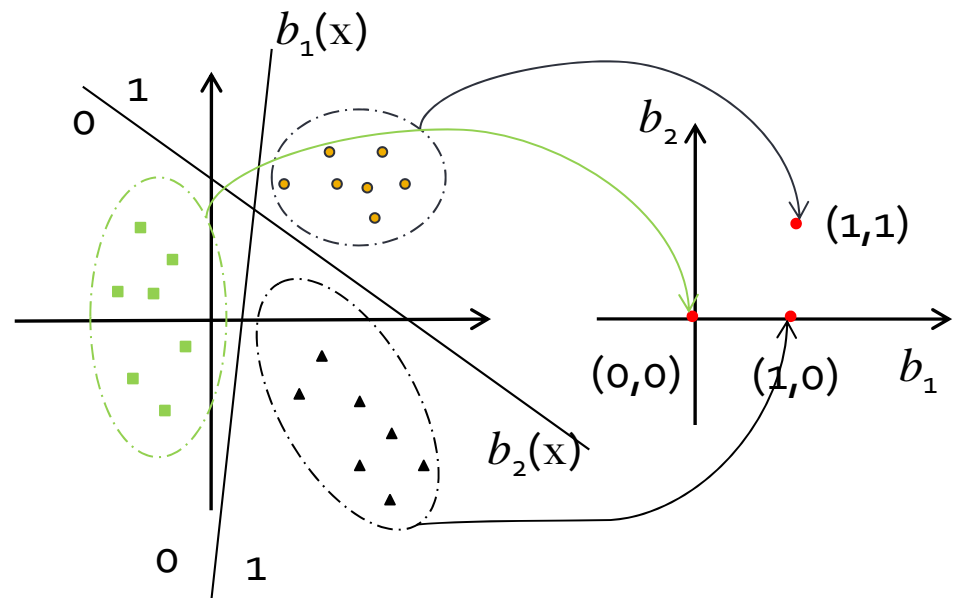
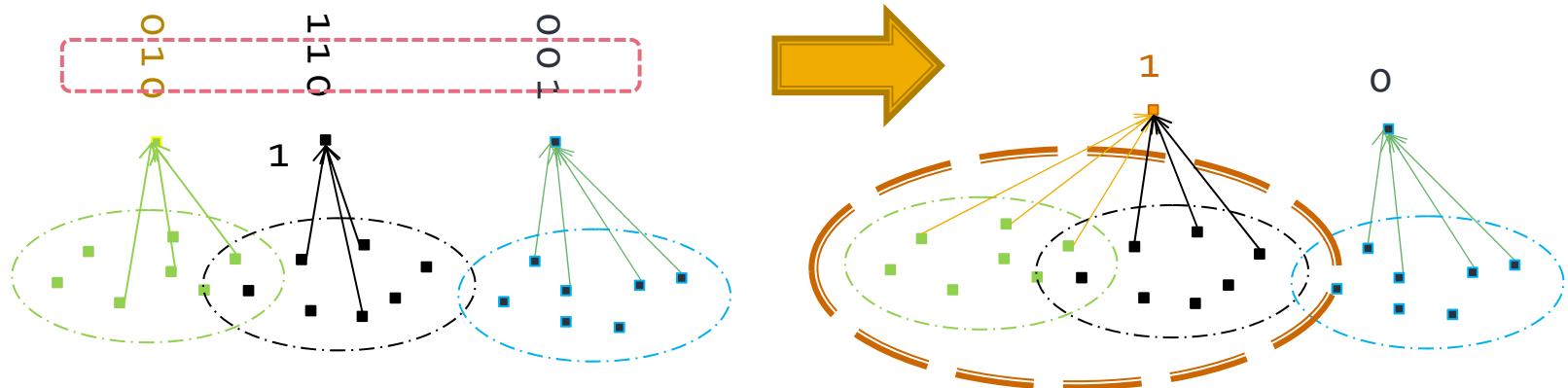
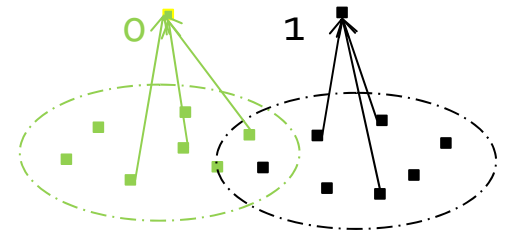


Illustration in 2-D space

# Rationale

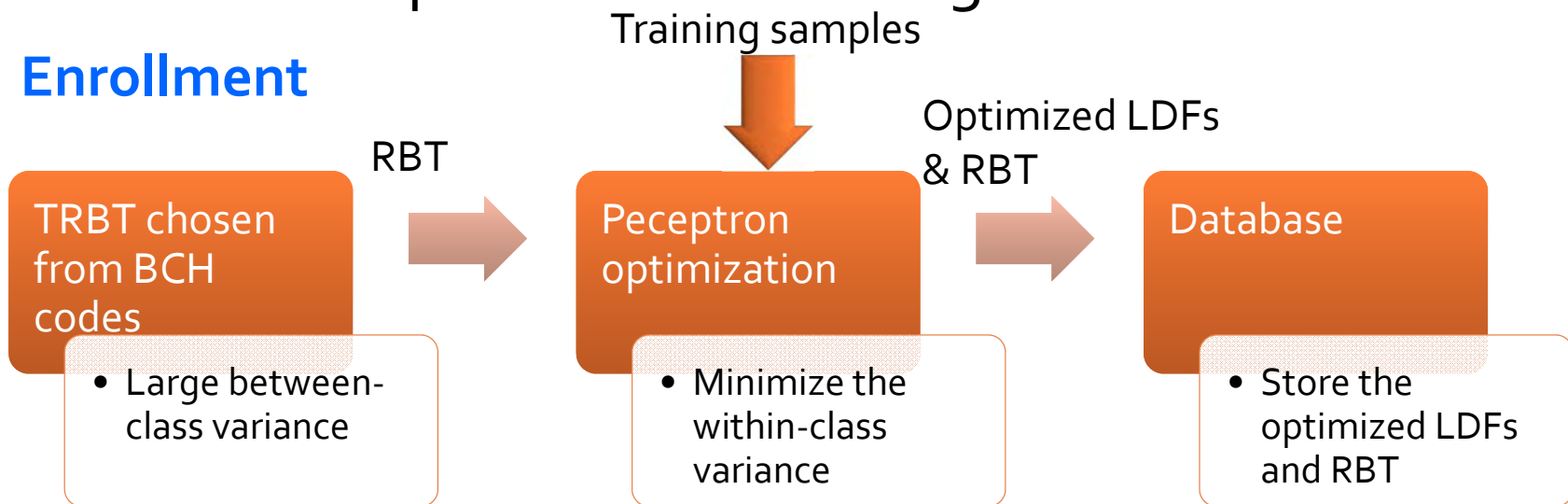
- Inspired by perceptron
  - Can find a LDF to classify two classes
  - Construct a continuous perceptron criteria function to find optimal  $(w, t)$ 
    - Can be extended to multiple classes with labels of multiple bits, just like binarization



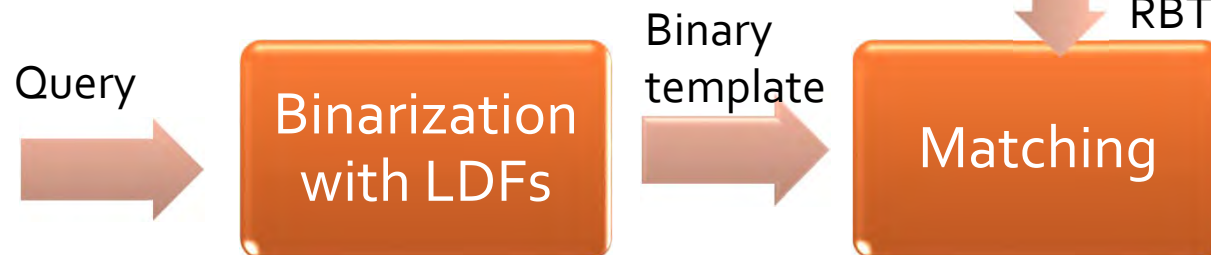
# Detailed Algorithm

- The whole procedure of the algorithm

## Enrollment



## Query



# Experimental Results

- Experiment settings



CMU PIE



FRGC

$c$  : No. of individuals.

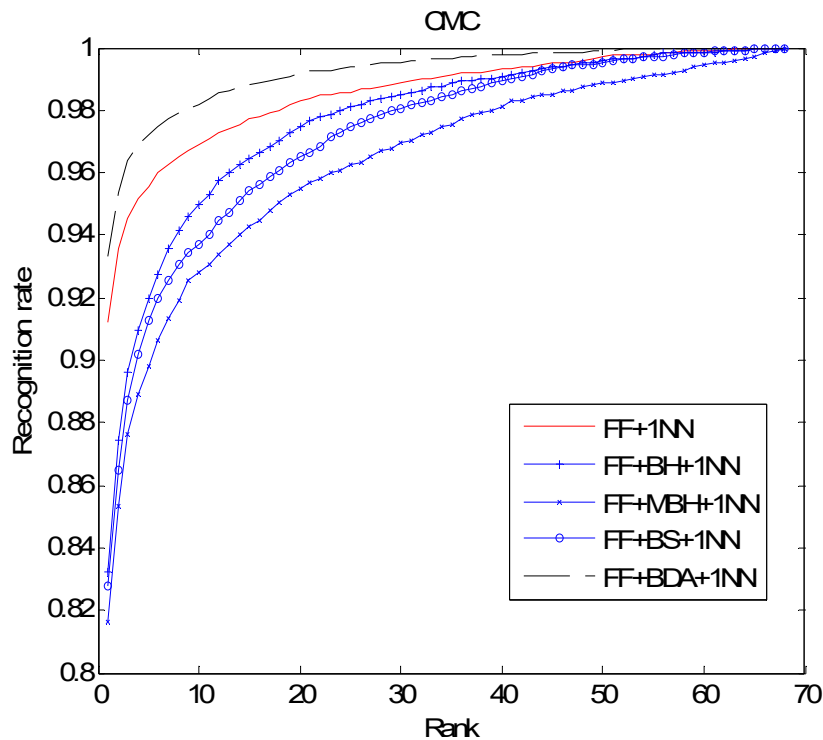
$N_p$ : No. of samples for each individual.

$N_t$  : No. of training samples per individual

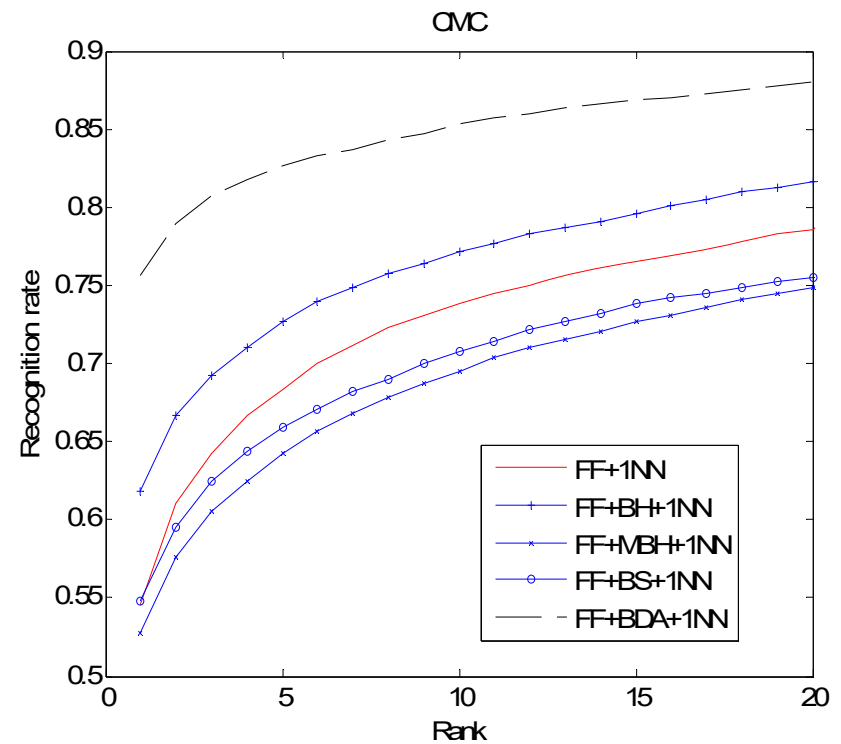
Database	$c$	$N_p$	$N_t$	Variations
CMU PIE	68	105	10	Illumination, pose, expression
FRGC	350	40	5	expression, illumination, mild pose

# Experimental Results

## ■ Experiment settings

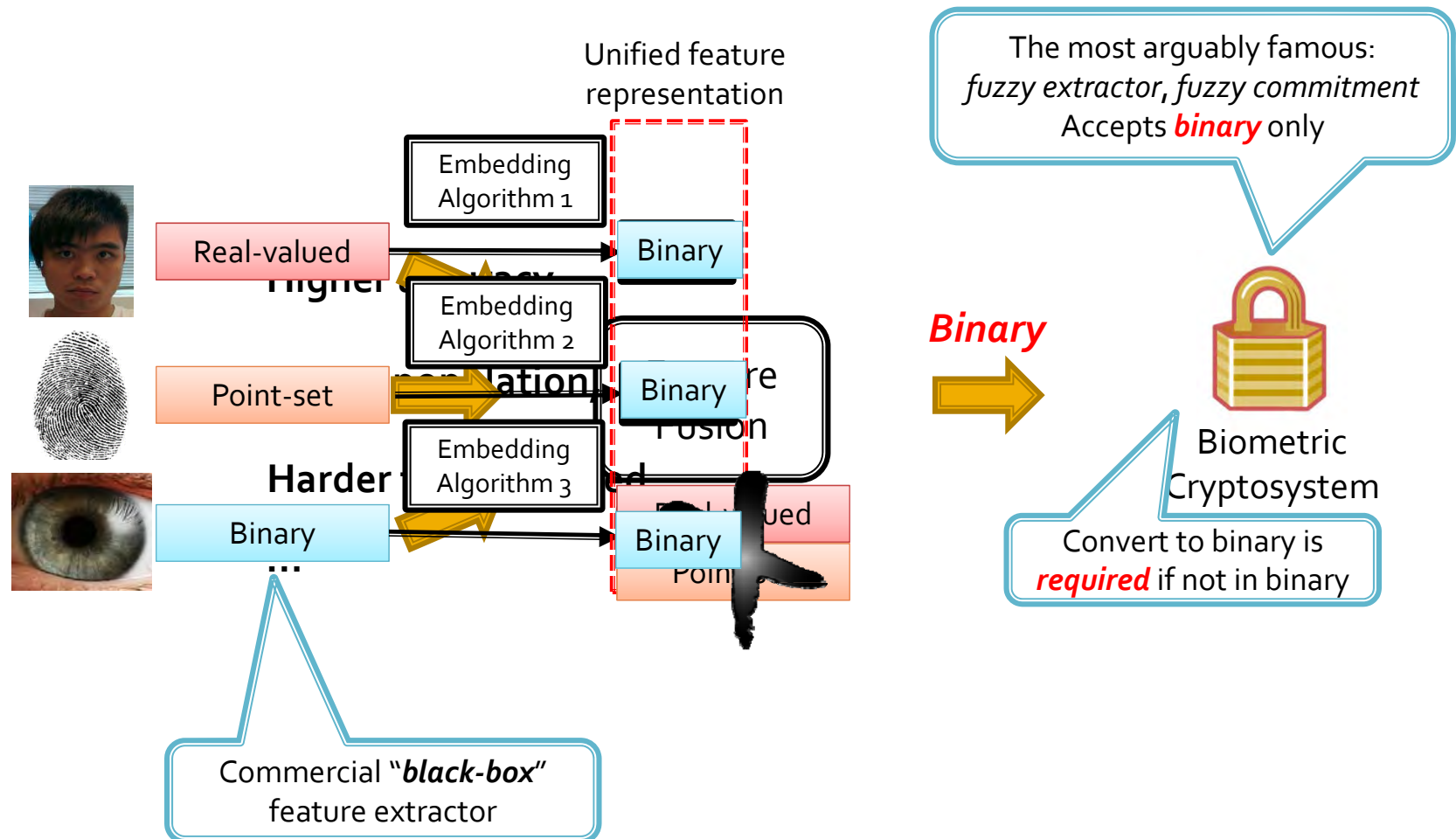


(a) CMU PIE



(b) FRGC

# Binary Template Fusion for Multi-biometric Cryptosystem [IVC 2017]



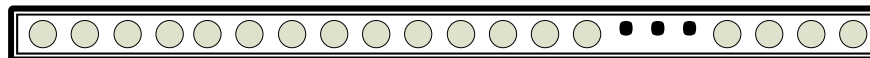


# Criteria for Binary Template Fusion

- Discriminability
  - Small intra-user variations *of* feature bits
  - Large inter-user variations *of* feature bits
- Security (high-entropy)
  - Low dependency *among* bits
  - High uniformity *of* feature bits
- Privacy
  - No information leakage from helper data

# Proposed Binary Template Fusion

- Stage one: dependency-reductive bit grouping
  - Dependency among bits (**security**)
- Stage two: discriminative within-group fusion
  - Bit-uniformity (**security**), intra-user variations (**discriminability**), inter-user variations (**discriminability**)



dependency reductive bit-grouping

*Grouping information*



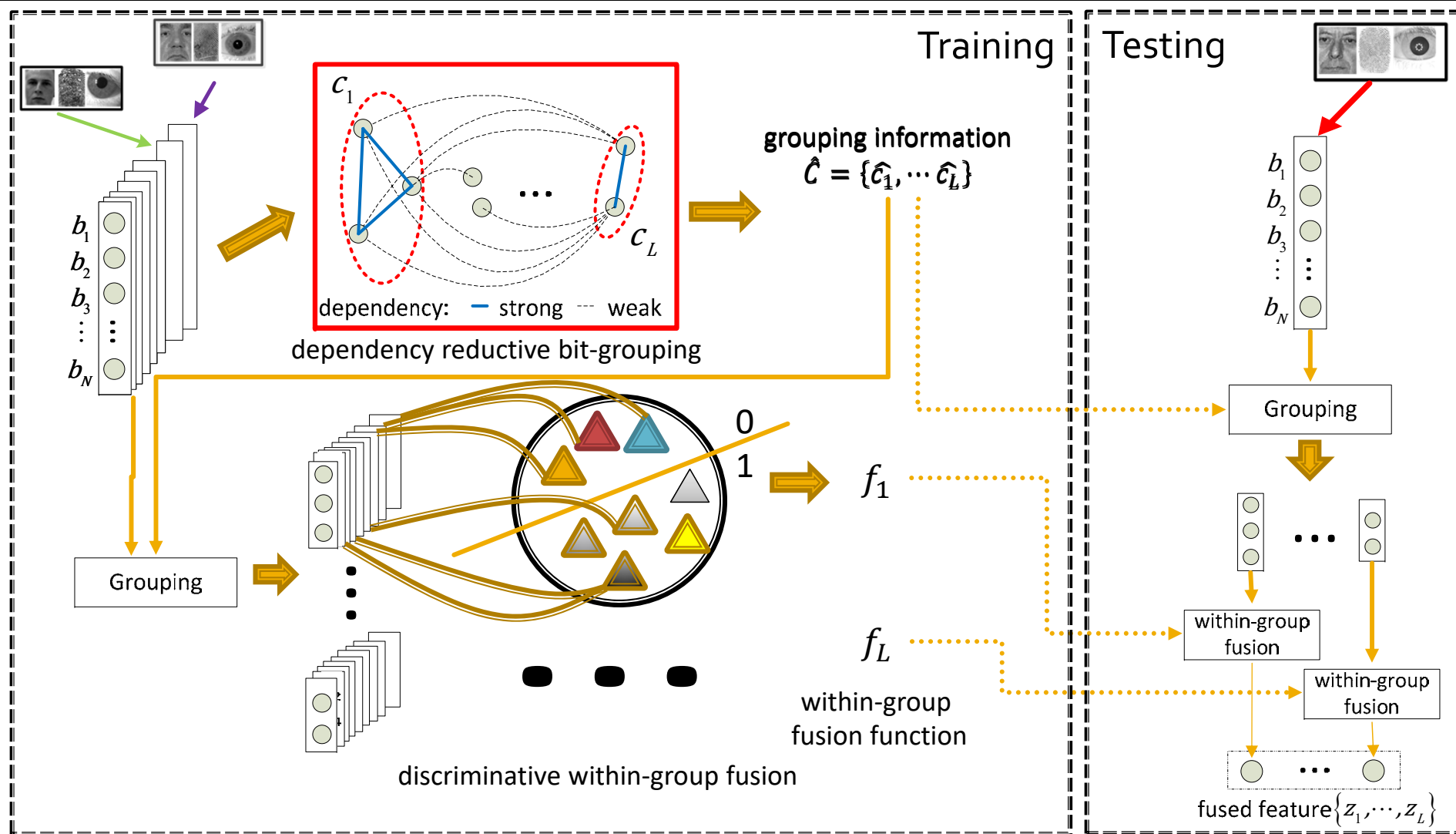
discriminative within-group fusion

*Fusion function*

Discriminative  
binary feature  
with high entropy



# Proposed Binary Template Fusion

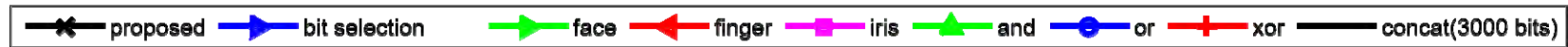


# Experiments

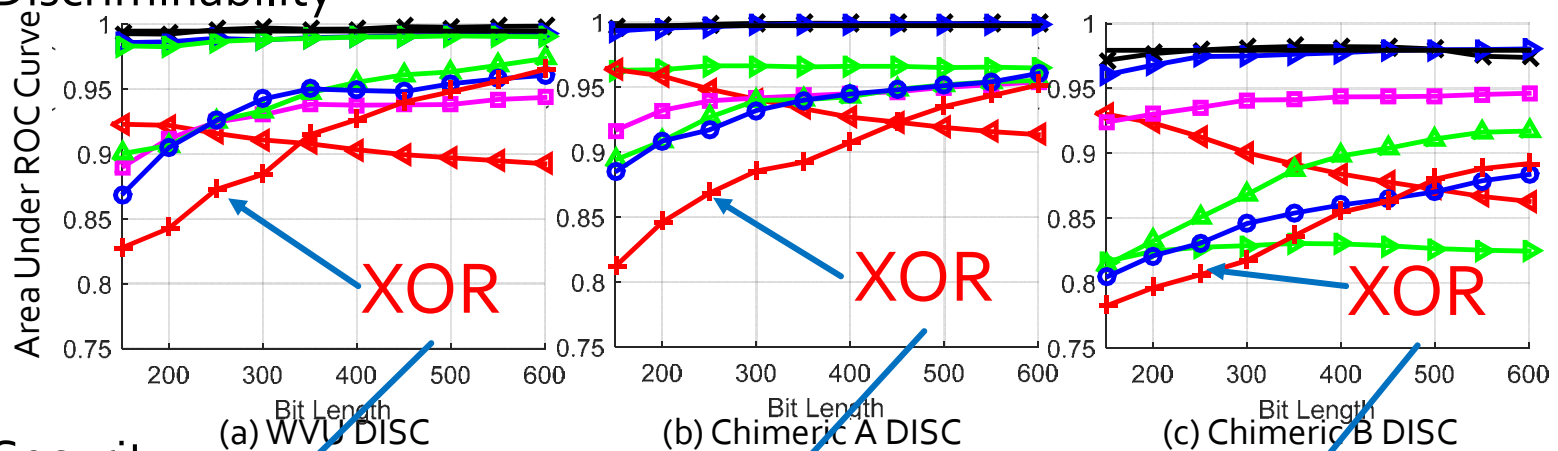
- Evaluation
  - Discriminability ( Area under ROC curve)
  - Security (average Renyi entropy, Hidano et al. BIOSIG2012)
- Experimental setting

Multimodal Database	WVU	Chimeric A (FVC2000DB2 + FERET + CASIA)	Chimeric B (FVC2002DB2 + FRGC + ICE2006)
Subjects	106	100	100
Training Sample	3	4	4
Testing Sample	2	4	4

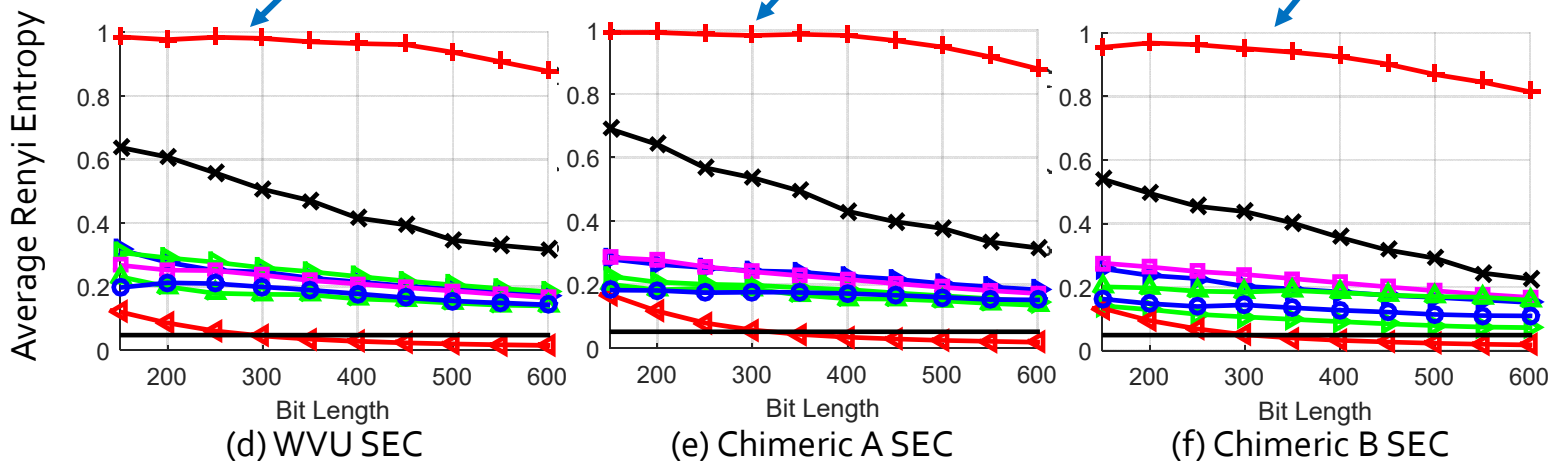
# Experimental Results



## Discriminability



## Security



**Thank you!**

# References (Face Template Protection)

- G. MAI, M H Lim and P C Yuen, Binary Feature Fusion for Discriminative and Secure Multi-biometric Cryptosystems, *Image and Vision Computing*, 2017
- M H Lim and P C Yuen, Entropy Measurement for Biometric Verification Systems, *IEEE Transactions on Cybernetics*, 2016
- M H Lim, S Verma, G C Mai and P C Yuen, "Learning discriminability-preserving histogram representation from unordered features for multibiometric feature-fused template protection", *Pattern Recognition*, 2016
- Y C Feng, M H Lim and P C Yuen, Masquerade attack on transform-based binary-template protection based on perceptron learning, *Pattern Recognition*, 2014
- YC Feng & P C Yuen, Binary discriminant analysis for generating binary face template, *IEEE Transactions on Information Forensics and Security*, 2012
- YC Feng, P C Yuen, AK Jain, A hybrid approach for generating secure and discriminating face template, *IEEE Transactions on Information Forensics and Security*, 2010



# References (Face Anti-spoofing)

1. N. Erdogmus and S. Marcel, "Spoofing face recognition with 3d masks", *TIFS*, 2014
2. J. Maatta, A. Hadid, and M. Pietikainen. "Face spoofing detection from single images using micro-texture analysis", *IJCB*, 2011.
3. D. Wen, H. Han, and A. K. Jain, "Face spoof detection with image distortion analysis", *TIFS*, 2015.
4. G. Pan, L. Sun, Z. Wu, and S. Lao. "Eyeblick-based antispoofing in face recognition from a generic webcam", *ICCV*, 2007.
5. T. de Freitas Pereira, J. Komulainen, A. Anjos, J. M. De Martino, A. Hadid, M. Pietikainen, and S. Marcel, "Face liveness detection using dynamic texture.", *EURASIP JIVP*, 2014.
6. X. Li, J. Komulainen, G Y Zhao, P C Yuen, and M. Pietikainen, "Generalized face anti-spoofing by detecting pulse from face videos", *ICPR*, 2016.
7. S. Liu, P C. Yuen, S. Zhang, and G Y Zhao, "3D Mask Face Anti-spoofing with Remote Photoplethysmography", *ECCV*, 2016.
8. S. Liu, B Y Yang, P C Yuen, G Y Zhao, "A 3D Mask Face Anti-spoofing Database with RealWorld Variations", *CVPRW*, 2016.
9. R Shao, X Y Lan and P C Yuen, "Deep Convolutional Dynamic Texture Learning with Adaptive Channel-discriminability for 3D Mask Face Anti-spoofing", *IAPR/IEEE International Joint Conference on Biometrics (IJCB)*, Oct 2017