Remote Photoplethysmography Based 3D Facial Mask Presentation Attack Detection (a.k.a Face Anti-spoofing)

P C Yuen

Department of Computer Science
Hong Kong Baptist University

Emai: pcyuen@comp.hkbu.edu.hk

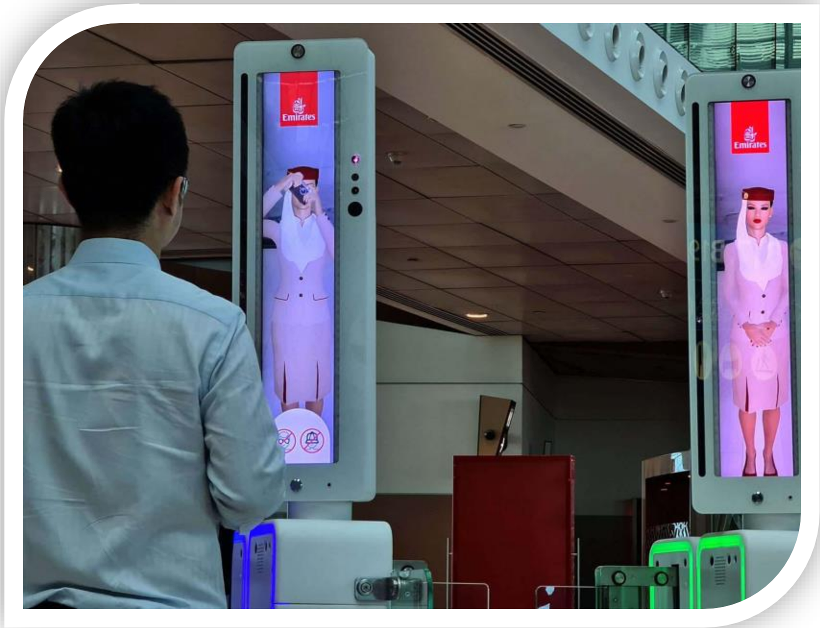
Outline

1. Background and Motivations
2. Basic principle for remote photoplethysmography (rPPG) for Face Presentation Attack Detection
3. rPPG based Face Presentation Attack Detection Methods
4. Conclusions

Background and Motivations

- Extensive deployed biometrics practical applications

Door Access Control



Iris recognition at Dubai's airport

Palmprint recognition – Amazon One

Border Control

Background and Motivations

- Face Recognition Technology

2022 – The year that facial recognition will lead the fintech industry

Contactless e-channel in HK


'World's first' facial recognition ATM unveiled in China

PUBLISHED : Sunday, 31 May, 2015, 6:38am
UPDATED : Monday, 01 June, 2015, 11:51am

COMMENTS:

ent.com

MIT Technology Review:
10 breakthrough technologies 2017

https://www.financedigest.com/2022-the-year-that-facial-recognition-will-lead-the-fintech-industry.html

Background and Motivations

Is Face Recognition Technology Secure?

Primary students spoof the face recognition system of auto courier cabinet with a printed photo

"A few days ago, the Science Team of Class 402 of Xiuzhou Foreign Language School of Shanghai International Studies University discovered in an extracurricular scientific experiment that as long as a printed photo can be used instead of a real person to scan their face, it can fool the Fengchao smart cabinet in the community and take out parents' personal information. shipment. is this real?"

刷脸取件被小学生用照片破解，丰巢快递柜紧急下线相关功能

2019-10-17 14:05

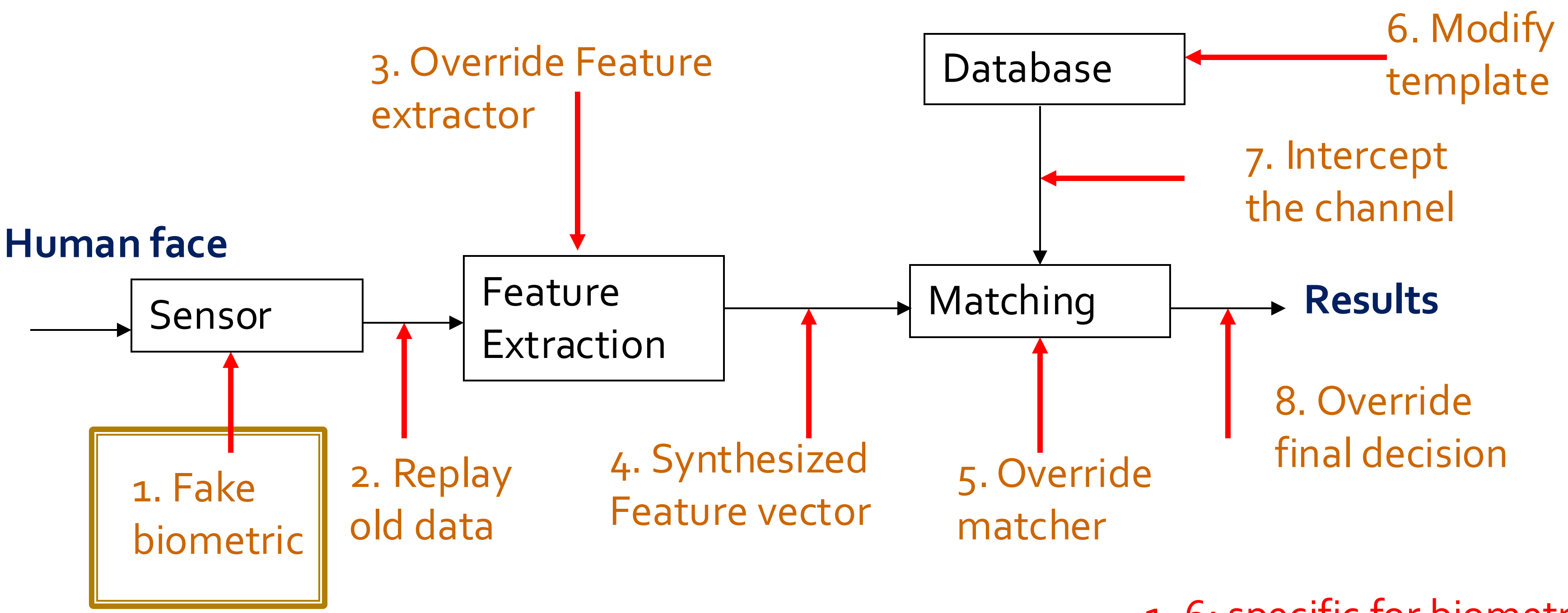
本文转载自微信公众号"央视新闻"（ID：cctvnewscenter）

前些天，上海外国语大学秀洲外国语学校402班科学小队在一次课外科学实验中发现：只要用一张打印照片就能代替真人刷脸、骗过小区里的丰巢智能柜，取出父母们的货件。这是真的吗？

News from https://www.sohu.com/a/347612078_115479

Background and Motivations

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



1, 6: specific for biometric systems

Background and Motivations

- Face Presentation Attack Detection (PAD)
  - Face information can be easily acquired (facebook, twitter) and abused
  - 3 popular attacks: Print (image), Replay (video), and 3D mask

Low Cost

High cost, but hard to detect

Image and Video Face PAD

- A straightforward approach: a two-class classification problem

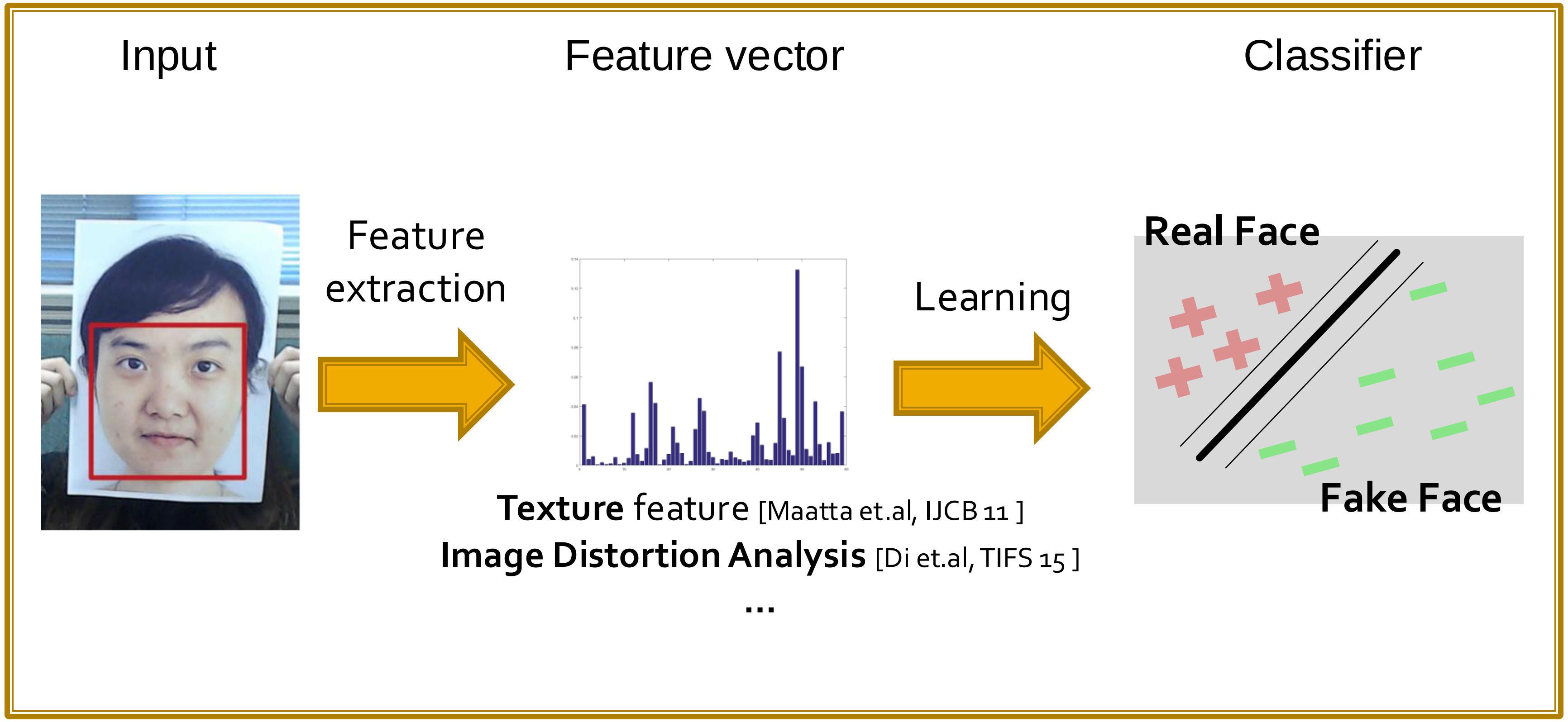
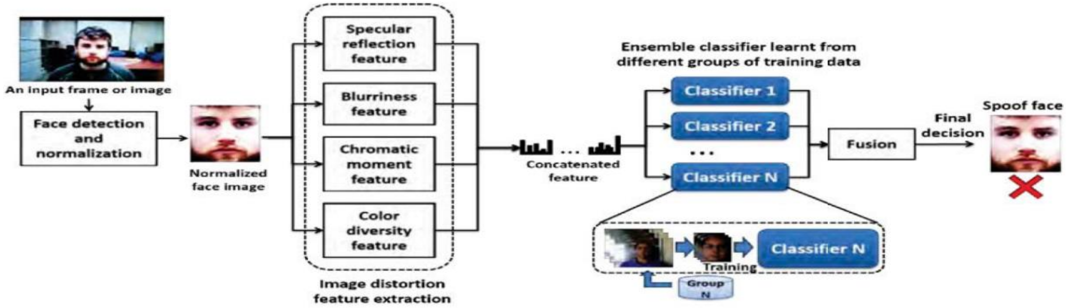
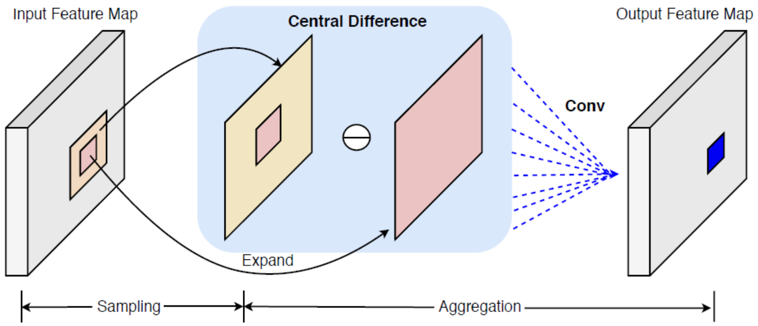


Image and Video Face PAD

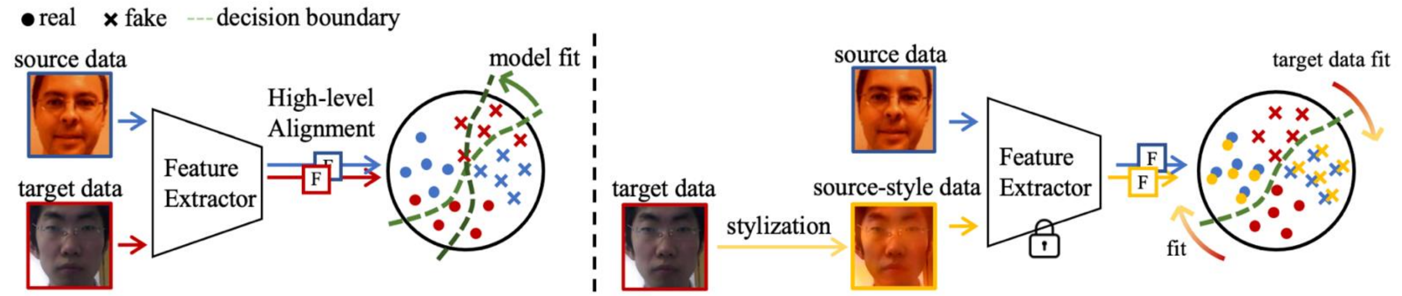
- Many methods have been proposed in the past decade



Appearance-based

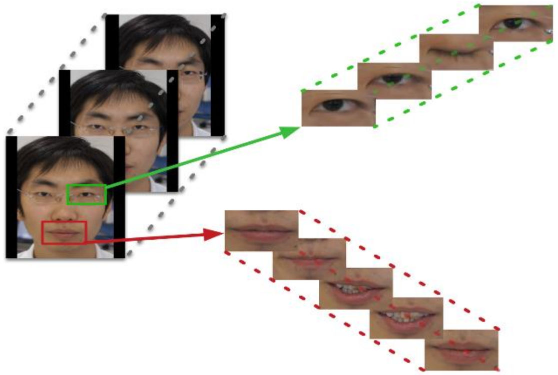


Deep Representation Learning



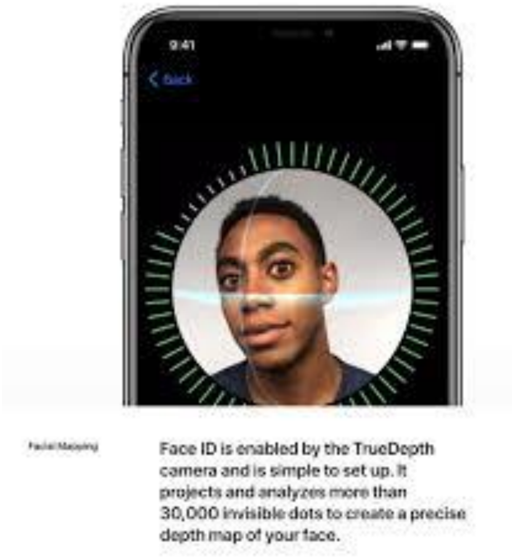
(a) Previous UDA in FAS: Model fit to Target data (b) UDA in FAS: Target data fit to Model

Domain Adaptation and Generalization



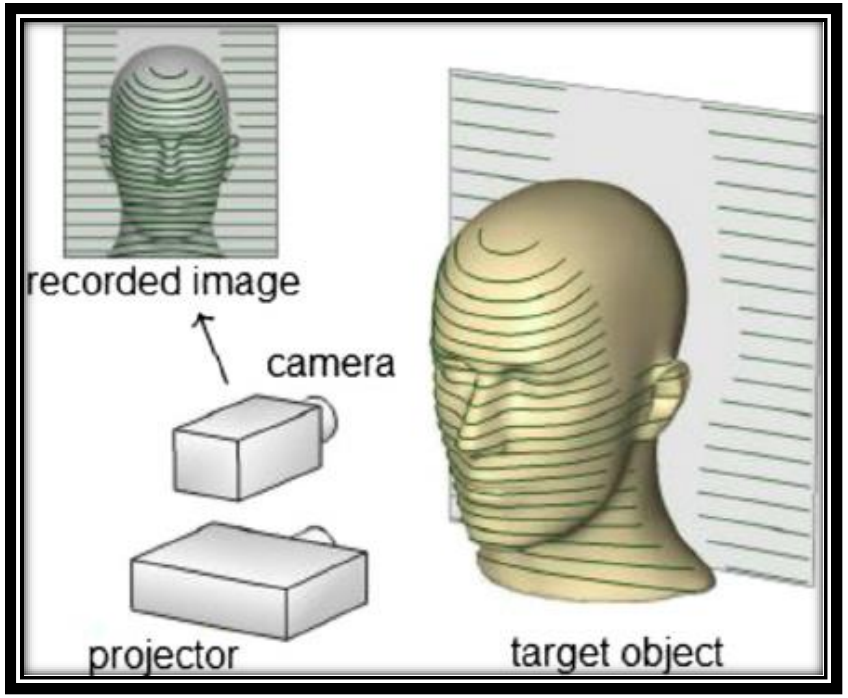
Motion-based

3D Face Recognition



FaceID in iPhone X

Announced on 12 September 2017



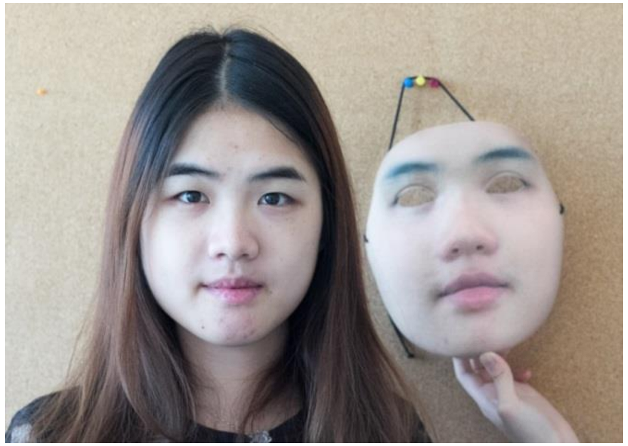
3D Face Recognition:

Employed Structured-light 3D technology

https://matterandform.net/blog/how-do-3d-scanners-work

3D Mask Face PAD

- 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

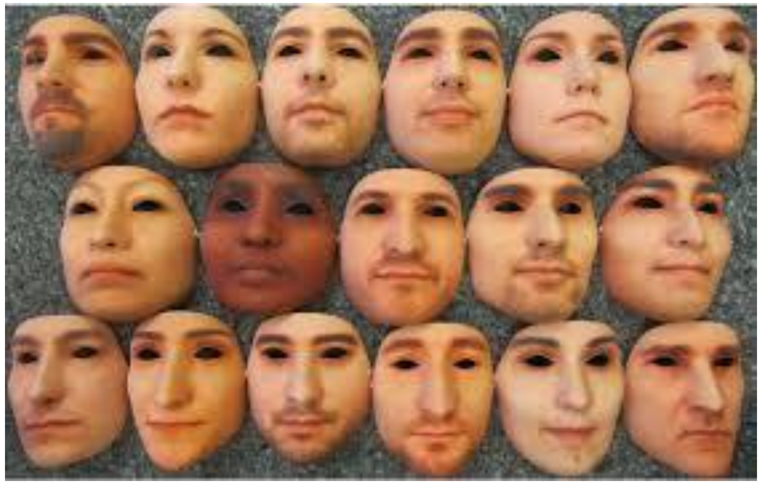


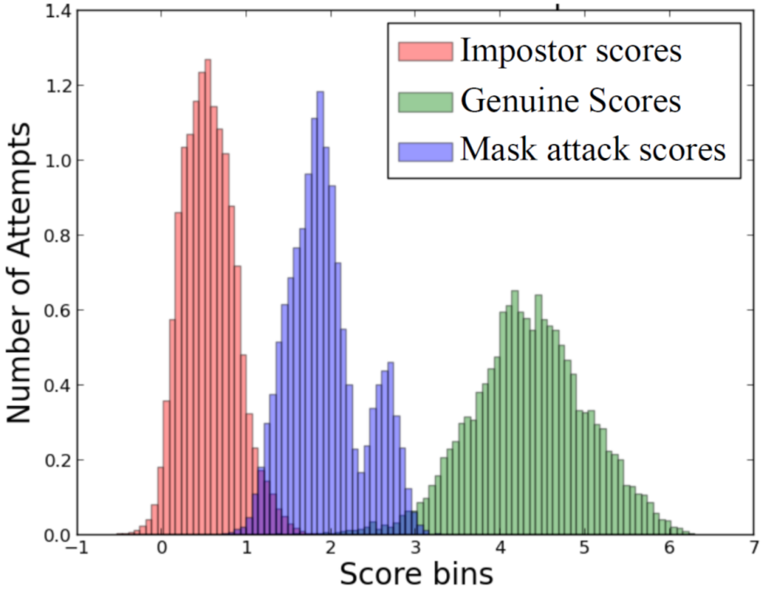
Source: idiap.ch

Mask is made from ThatsMyFace.com

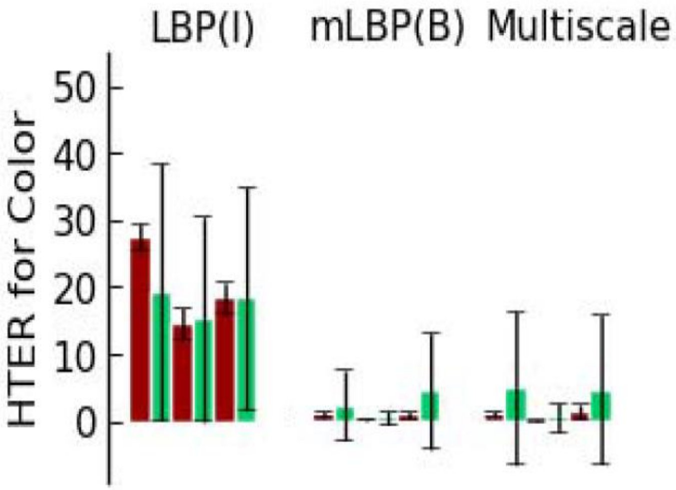
3D Mask Face PAD

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





[Erdogmus et al., TIFS14]



3D Mask Face PAD

- Super-realistic 3D Mask

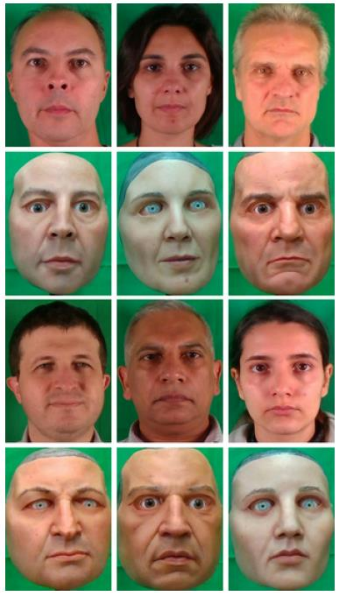
(a) Life face

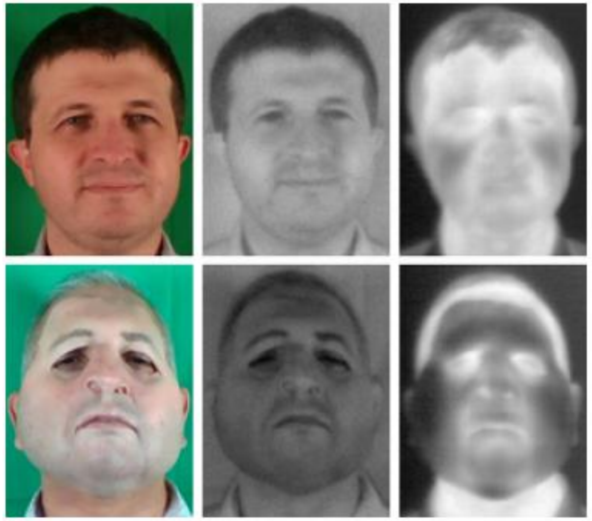
(b) Real-F hyper real mask

Source: real-f.jp

3D Mask Face PAD

- Custom Silicone Masks Datasets
  - Consider PAs performed using custom-made flexible silicone masks..
  - A new dataset based on six custom silicone masks





Bhattacharjee S, Mohammadi A, Marcel S. Spoofing deep face recognition with custom silicone masks, BTAS 2018

BBC NEWS

Brazil drug dealer dresses up as daughter in bungled jail escape

05 August 2019 | Latin America & Caribbean

Airport and Payment Facial Recognition Systems Fooled by Masks and Photos, Raising Security Concerns

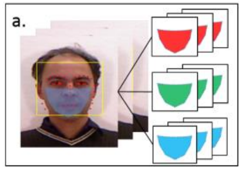
By Jeff John Roberts December 12, 2019

The test, by artificial intelligence company Kneron, involved visiting public locations and tricking facial recognition terminals into allowing payment or access. For example, in stores in Asia—where facial recognition technology is deployed widely—the Kneron team used high quality 3-D masks to deceive AliPay and WeChat payment systems in order to make purchases.

More alarming were the tests deployed at transportation hubs. At the self-boarding terminal in Schiphol Airport, the Netherlands' largest airport, the Kneron team tricked the sensor with just a photo on a phone screen. The team also says it was able to gain access in this way to rail stations in China where commuters use facial recognition to pay their fare and board trains.

Today Journey on PhotoPlethysmoGraphy based Face PAD Methods for 3D Mask Attack

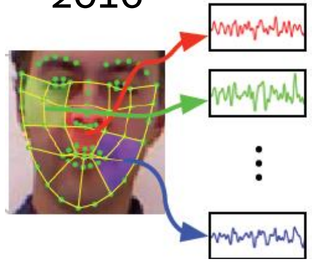
Global rPPG (ICPR16)



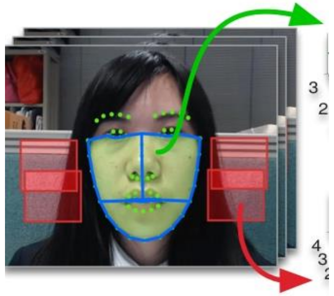
2016

Local rPPG (ECCV 16)

2016



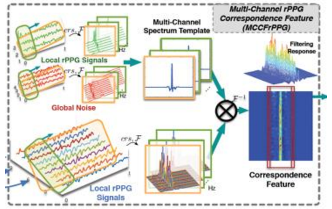
Local rPPG with background information for noise removal (ECCV18)



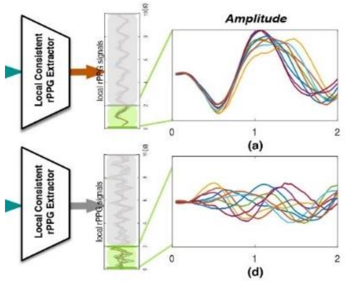
2018

Multi-Channel rPPG (TIFS 21)

2021



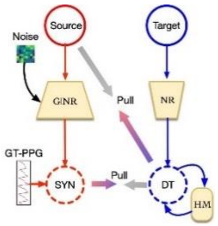
Fast rPPG estimation using Temporal Similarity (TIFS 22)



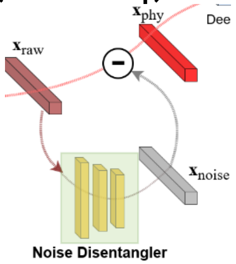
2022

Robust and generalised rPPG estimation (CVPR 23)

2023

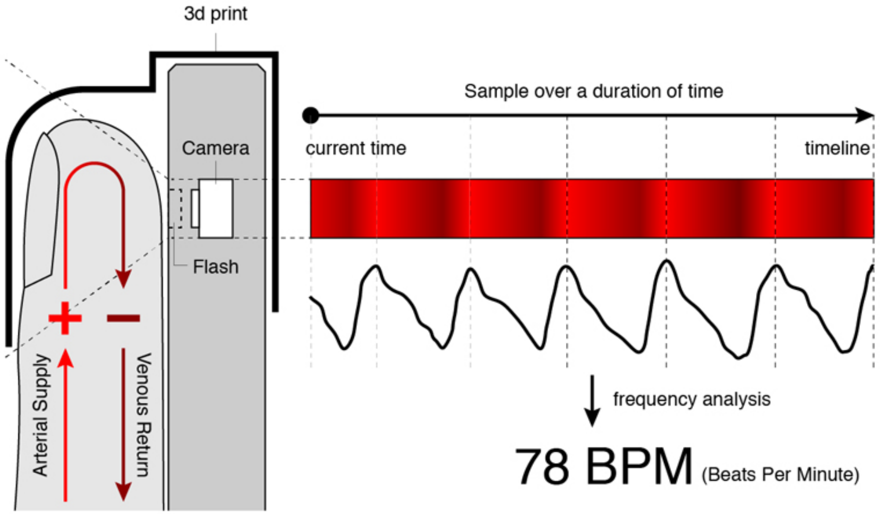


Noise disentangled rPPG (TIFS 24)



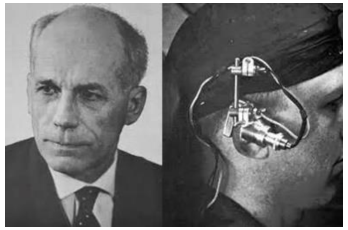
2024

PhotoPlethysmoGraphy (PPG)



Pic. from UCLA Lung Cancer Program http://lungcancer.ucla.edu/adm_tests_electro.html

History of PhotoPlethysmoGraphy



1935

1977

2001
The MIT ring with PPG, Rhee et al.

2010
Magnetic ear clip, Poh et al.

2016-2019
Oura PPG sensing ring

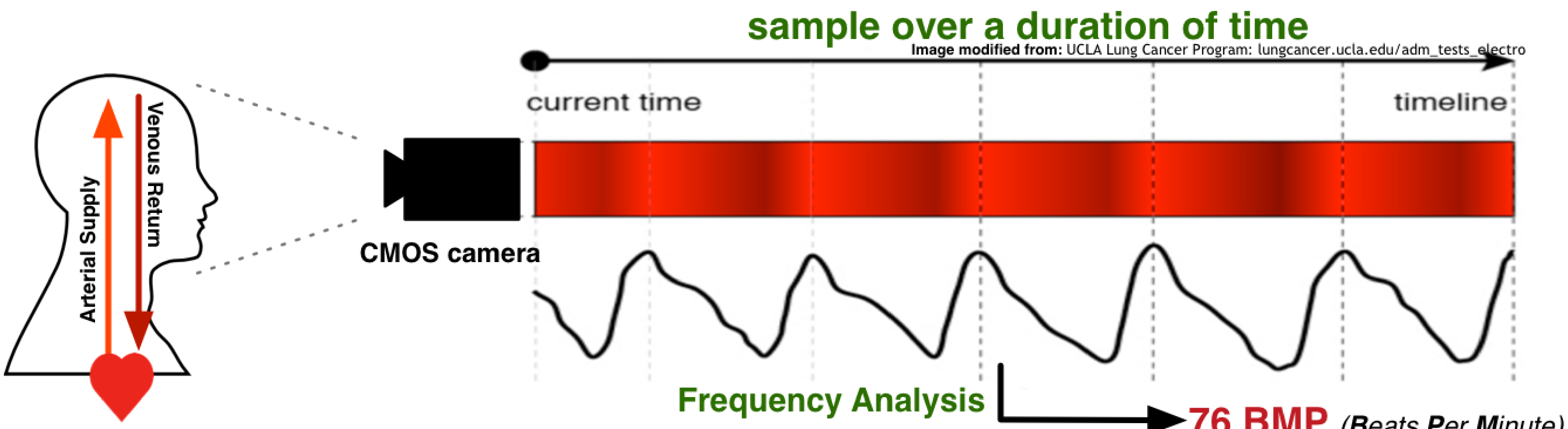
Limitation:

The contact approach is too restrictive and limits the applications

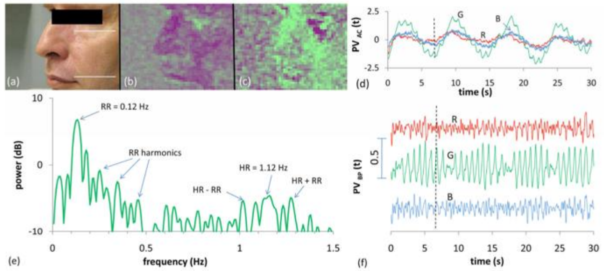
Applications: physiological monitoring, Blood oxygen saturation, Blood pressure, Cardiac output, Heart rate, ...

F. Wolling, S. Heimes and K. Van Laerhoven, "Unity in Diversity: Sampling Strategies in Wearable Photoplethysmography," in *IEEE Pervasive Computing*, 2019

remote PhotoPlethysmoGraphy (rPPG)



Brief History of rPPG

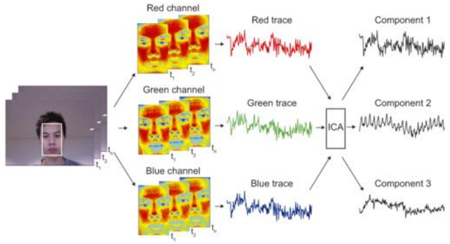


2008

Verkruysse et al ***first*** *verified the validity of rPPG for HR estimation from facial videos*

2010 | BSS-based

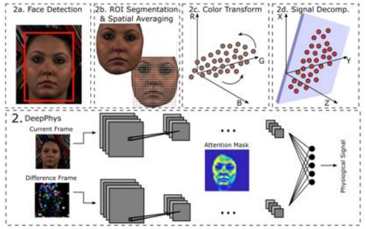
Poh et al. used video imaging and Blind Source Separation for Cardiac pulse measurement



2018 | ICA-based

MRICA applies Independent Component Analysis (ICA) to enhance iPPG signal

2018 | Deep Learning in rPPG

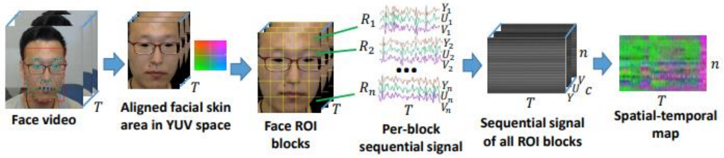


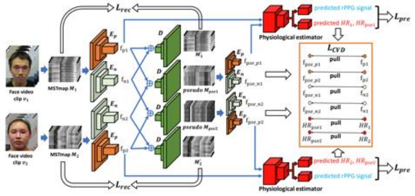
2018 | 2D CNN

DeepPhys, the first end-to-end system for video-based measurement using a deep convolutional network.

2019 | Spatial-Temporal Representation

Rhythmnet uses spatial-temporal representation to encode the HR signals;



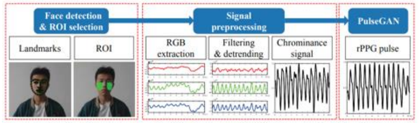


2020 | Feature Disentangling

Niu et al. disentangle the physiological features with non-physiological representations

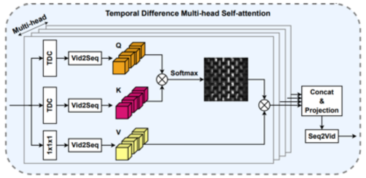
2021 | Adversarial learning

PulseGAN generates realistic rPPG pulse signals through denoising the chrominance signals



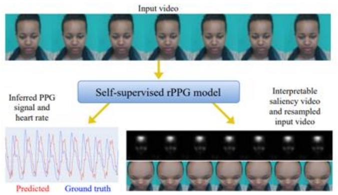
2022 | Transformer-based

First time to explore the long-range spatio-temporal relationship, by *Yu et al.*



2022 | Unsupervised Learning

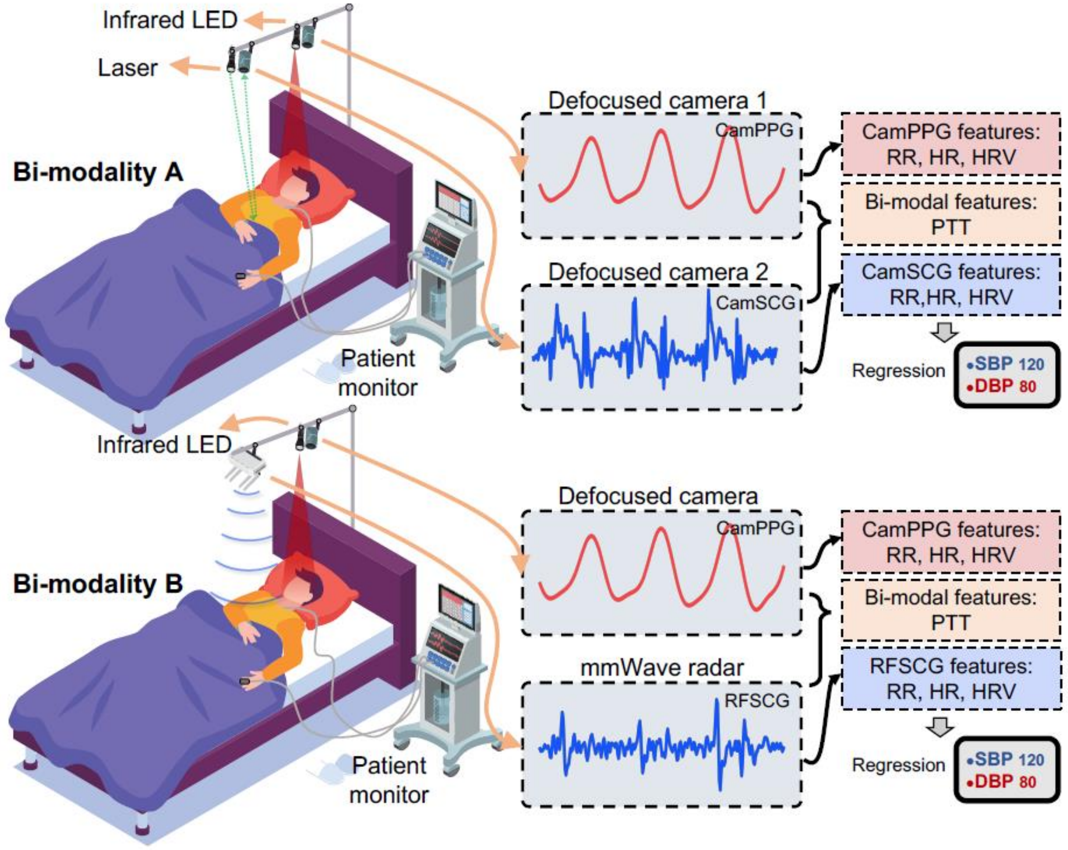
Fully self-supervised training approach by *Gideon et al.*



2024 | Data Augmentation

Paruchuri et al. explore motion transfer as a form of data augmentation to introduce motion variation

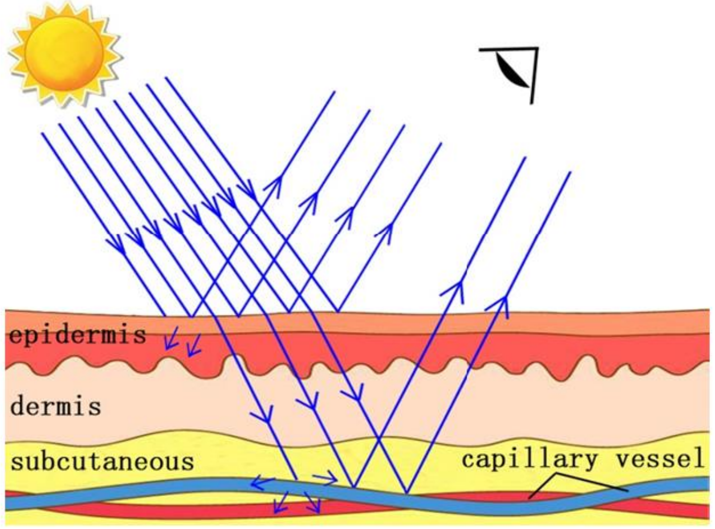
Privacy-protected Contactless Sleep Disorder Monitoring and Detection



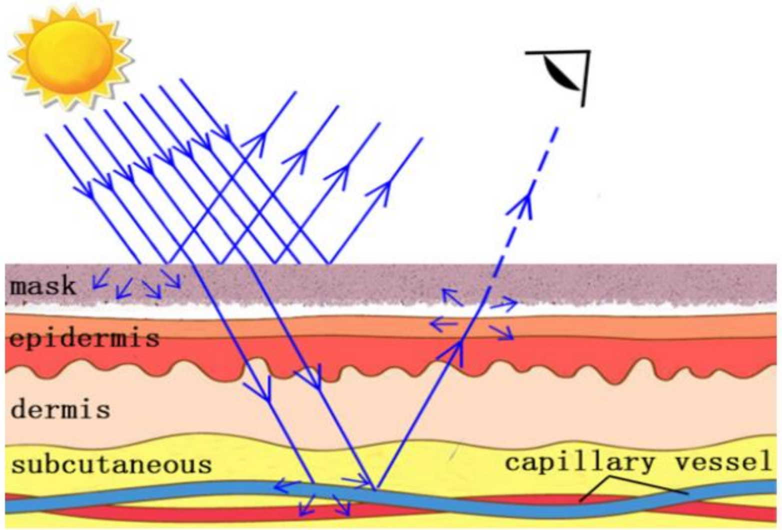
- Sleep disorder Statistics:
  - 62% of adults around the world say they don't sleep as well as they'd like (Philips Global Sleep Survey, 2019).
  - As many as 67% of adults report sleep disturbances at least once every night (Philips Global Sleep Survey, 2019).
  - 44% of adults around the world say that the quality of their sleep has gotten worse over the past five years (Philips Global Sleep Survey, 2019).
- Polysomnography (PSG) is the gold standard for sleep monitoring in clinical settings
- A new camera-based rPPG and Seismocardiography (SCG) system

Y Zhu1, Y Ge, Q Wei, Y Huang, D Huang, P C Yuen, X Ji, F Xia, and W Wang, "Camera-based Bi-modal PPG-SCG: Sleep Privacy-protected Contactless Vital Signs Monitoring", Submitted to Journal under review, May 2024

Principle of rPPG Based Face PAD



(a)



(b)

(a) rPPG signal can be extracted from genuine face skin.

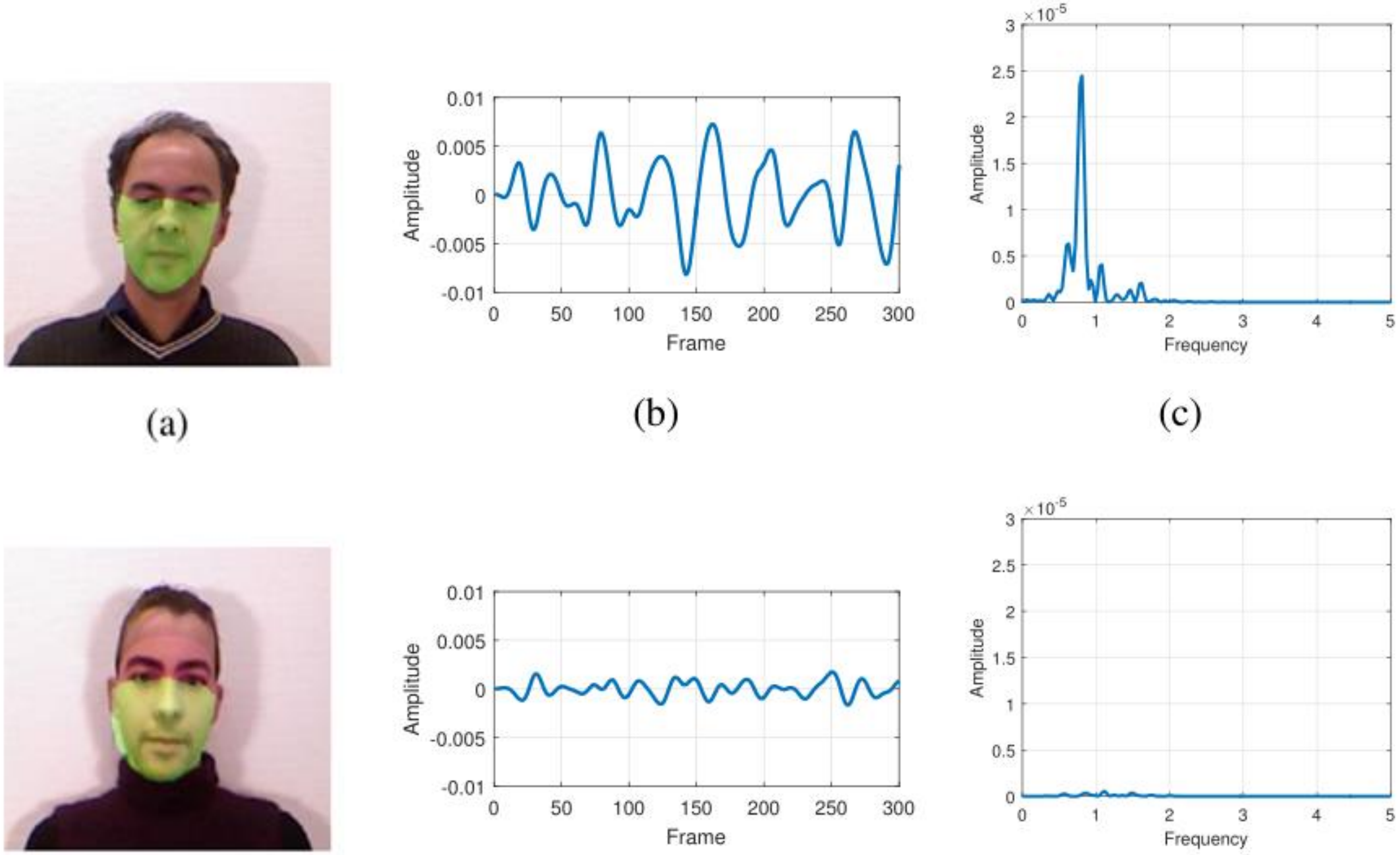
(b) rPPG signals will be **too weak** to be detected from a masked face.

- 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 PAD

genuine face

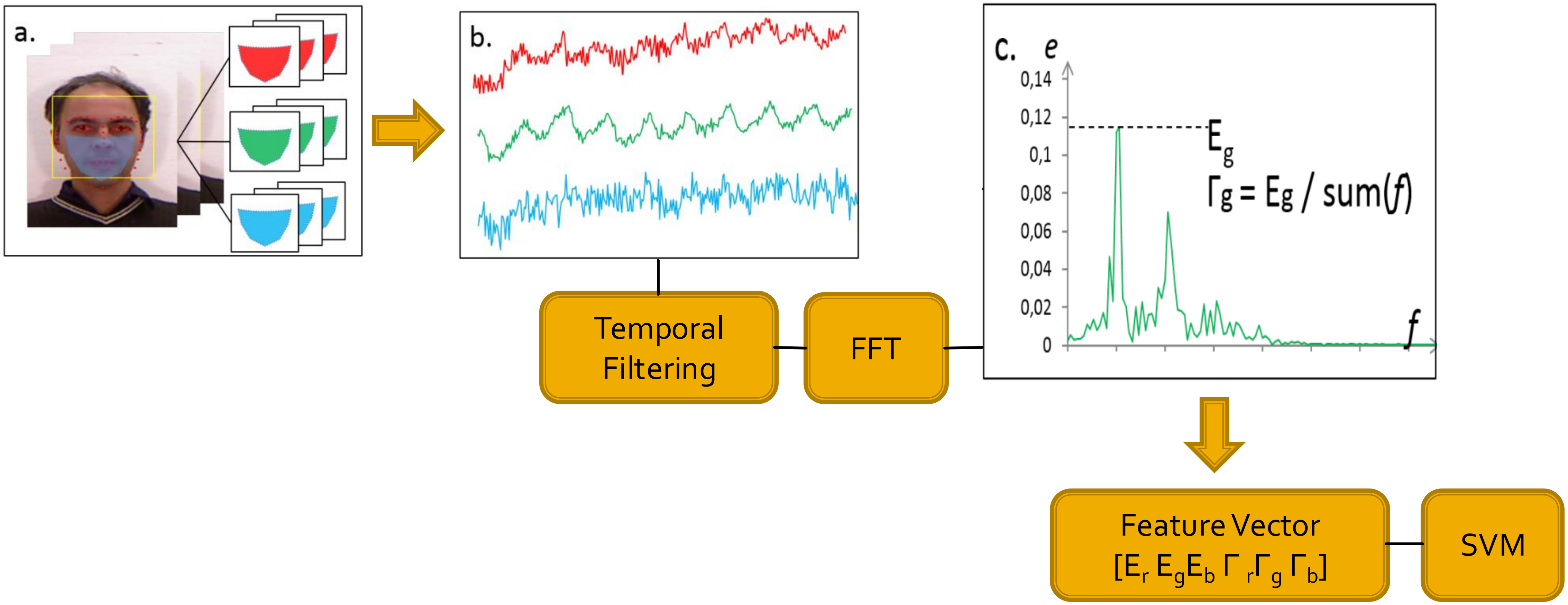
masked face



(a) (b) (c)

(d) (e) (f)

Global rPPG-based Face PAD [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

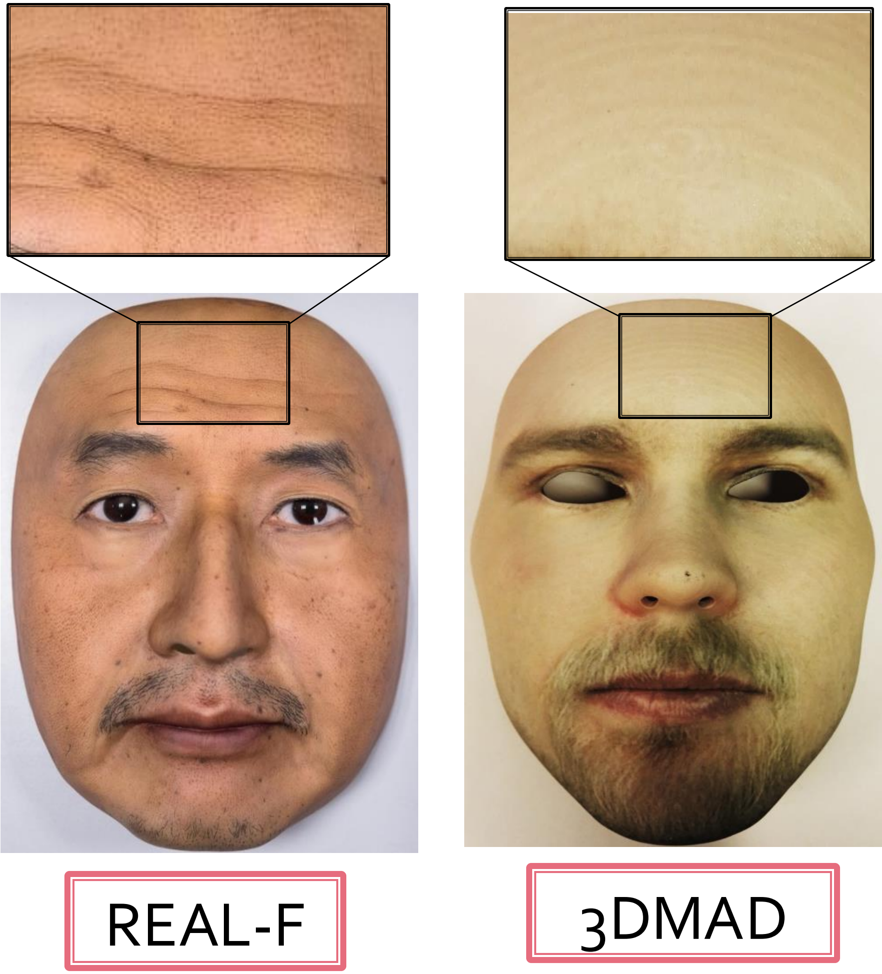
Experimental Results

- 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 REAL-F (cross dataset)
  - Randomly select 8 subjects from 3DMAD for training and the other 8 subjects as the development set

Method	HTER(%)	EER(%)	FPR @FNR=0.1%	FPR @FNR=0.01%
Pulse (ours)	**4.29**	**1.58**	**0.25**	**3.83**
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



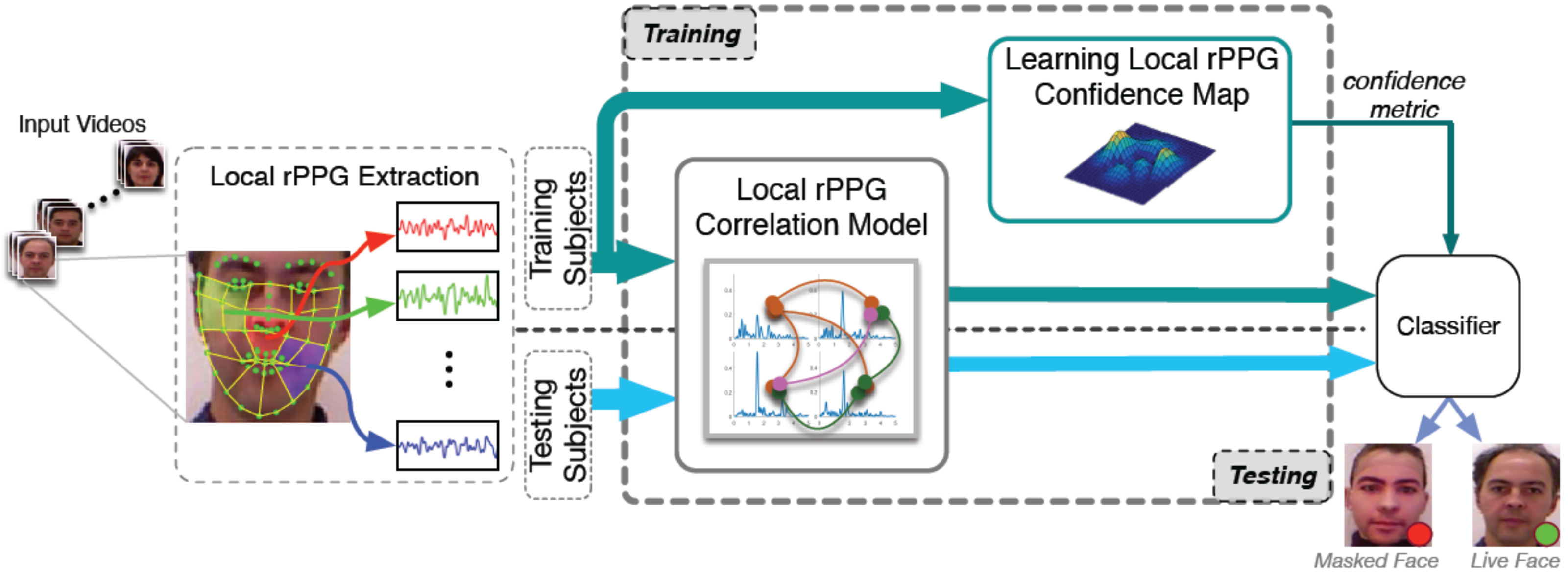
Limitations on Global rPPG method

- Global rPPG signal is sensitive to certain variations such as illuminations, head motion and video quality

- rPPG signal strength may vary with different subjects

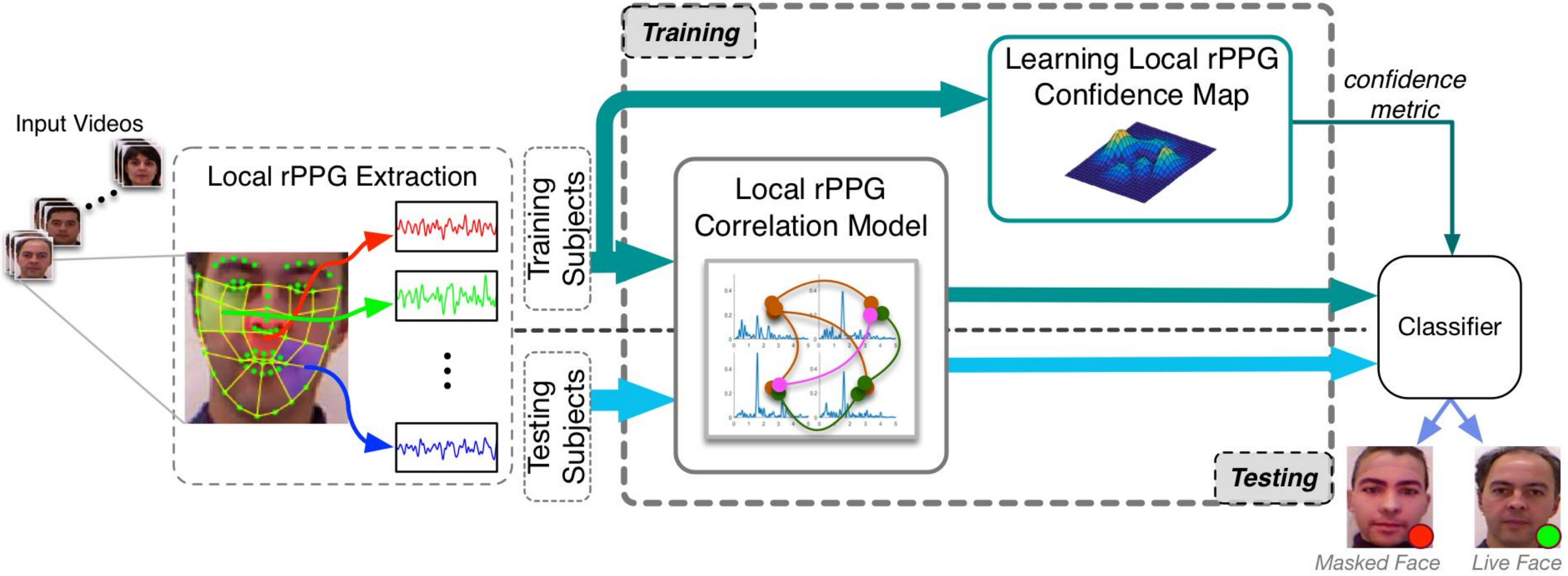
How to increase the robustness of rPPG-based face presentation attack detection?

Local rPPG based Face PAD 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*

Local rPPG based Face PAD Method



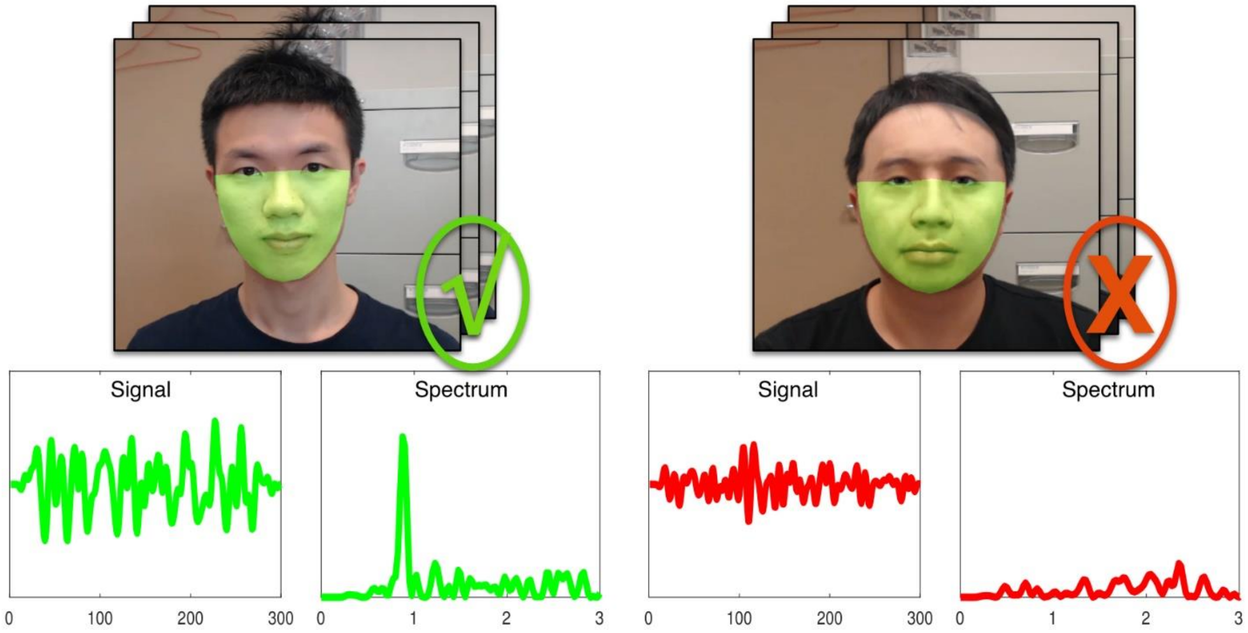
(a) Local ROIs are pre-defined based on the facial landmarks. Local rPPG signals are extracted from these local face regions.

(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

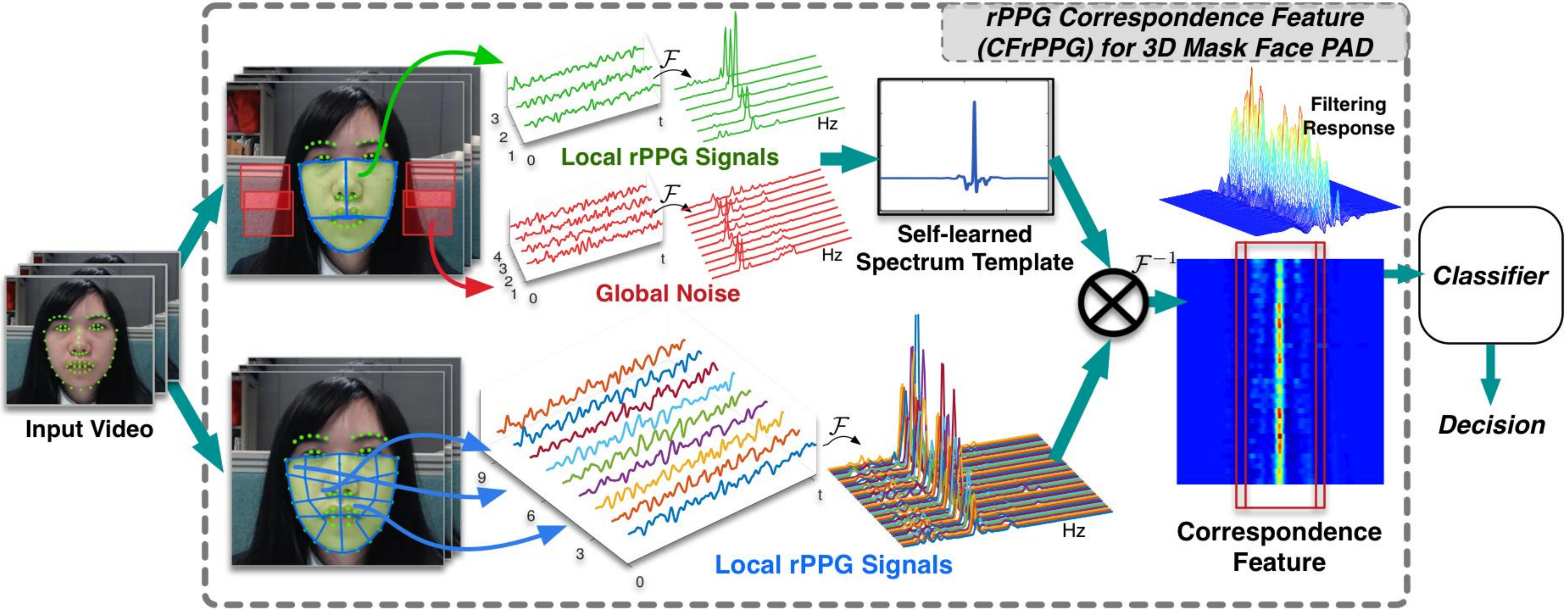
Limitation on Local rPPG Approach



How to **accurately obtain the liveness evidence** from the observed noisy rPPG signals**?**

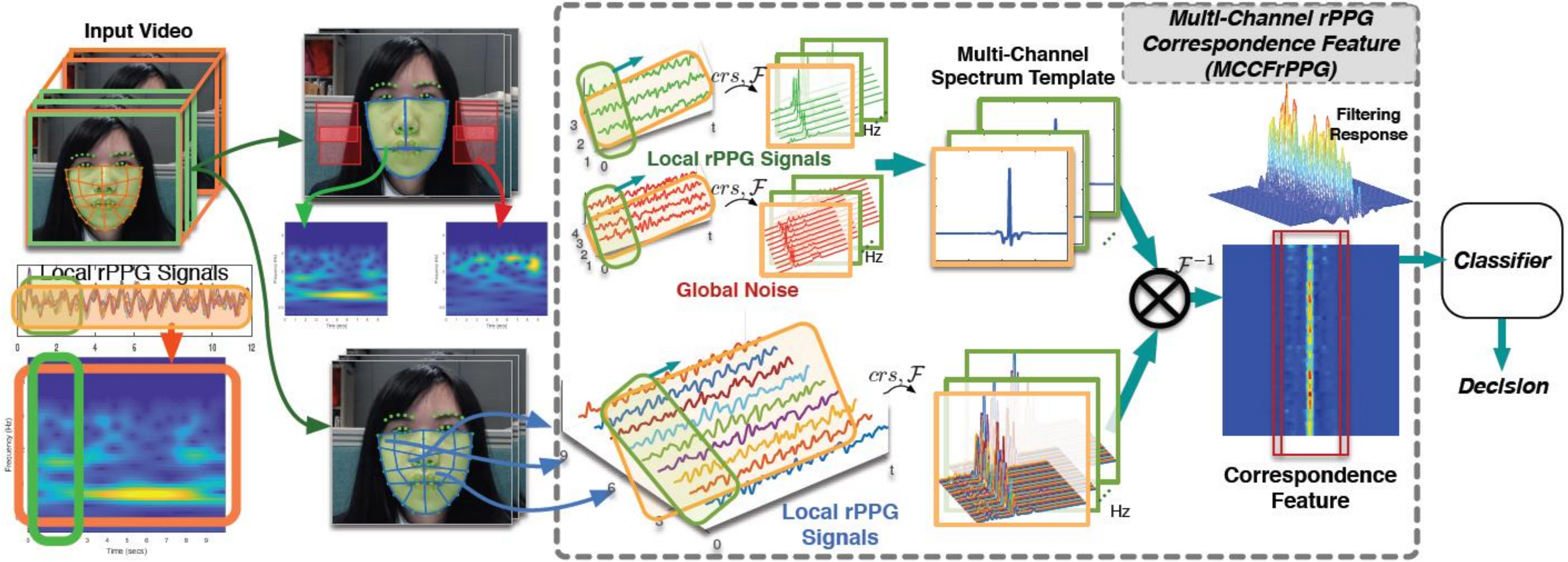
Improved Method: rPPG Correspondence Feature

[ECCV 2018]



1. S Q Liu, X Y Lan and P C Yuen, "Remote Photoplethysmography Correspondence Feature for 3D Mask Face Presentation Attack Detection", *ECCV* 2018

Improved Method: Multi-channel rPPG Correspondence Feature [TIFS 2021]



1. S Q Liu, X Y Lan and P C Yuen, "Multi-Channel Remote Photoplethysmography Correspondence Feature for 3D Mask Face Presentation Attack Detection", *IEEE Transactions on Information Forensics and Security (TIFS)*, 2021.

Experimental Results

➢ Dataset

- 3DMAD [TIFS'14 Erdogmus et.al]
- HKBU MARs V1+
- Custom silicone mask attack dataset (CSMAD)
- HKBU MARs V2+

(a) ThatsMyface (b) REAL-f (c) Silicone

VARIATION SUMMARY OF 3D MASK ATTACK DATASETS USED IN THE EXPERIMENT

	#Subject/Mask	#Video	Mask Type	Lighting Condition	Camera	Face (pixel) Resolution	Compression
3DMAD [11]	17 17	255	TMF	1(Studio)	Kinect	80×80	Motion JPEG
HKBU-MARsV1+ [14]	12 12	180	TMF+RF	1(Room)	Logitech C920	200×200	H.264
CSMAD [33]	14 6	246	Silicon	4	RealSense SR300	350×350	H.264
HKBU-MARsV2+	16 16	1048	TMF+RF	6	3(C920, M3, MV-U3B)	200×200	both
Summary	59 39	1729	3	12	6	4	2

Experimental Results

INTRA DATASET EVALUATION RESULTS(%) ON 3DMAD

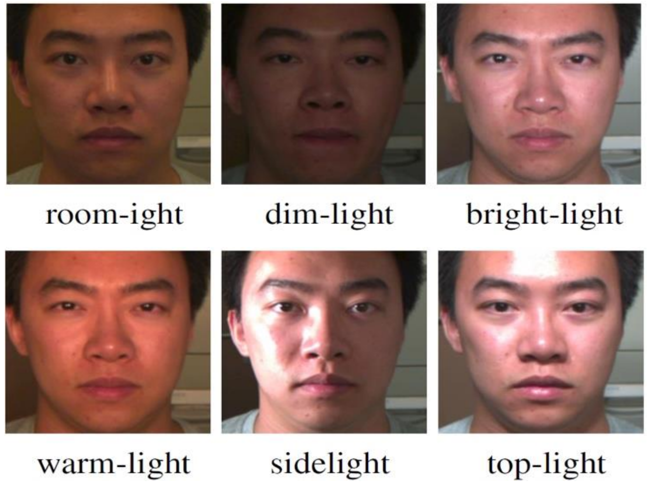
	HTER_dev	HTER_test	EER	AUC	BPCER@ APCER=0.1	BPCER@ APCER=0.01
MS-LBP [7]	**1.25 ± 1.9**	4.22 ± 10.3	2.66	99.6	1.50	**4.00**
CTA [20]	2.78 ± 3.6	4.40 ± 9.7	4.24	99.3	1.32	12.8
CNN	1.58 ± 1.6	**1.93 ± 3.4**	**2.07**	**99.7**	**0.38**	4.26
FBNet-RGB [48]	3.91 ± 2.4	5.66 ± 9.7	5.54	98.6	2.21	19.9
GrPPG [12]	13.4 ± 4.2	13.2 ± 13.2	13.9	92.6	15.4	36.2
PPGSec [36]	15.2 ± 4.4	15.9 ± 14.6	15.8	90.8	20.5	35.9
CFrPPG-crs	9.06 ± 4.4	8.57 ± 13.3	8.88	96.0	8.41	14.1
CFrPPG	5.95 ± 3.3	6.82 ± 12.1	6.94	97.1	5.85	11.6
MCCFrPPG	4.42 ± 2.3	5.60 ± 8.8	5.01	98.7	3.76	8.24

INTRA DATASET EVALUATION RESULTS(%) ON HKBU-MARSV2+

	HTER_dev	HTER_test	EER	AUC	BPCER@ APCER=0.1	BPCER@ APCER=0.01
MS-LBP [7]	12.4 ± 5.3	12.9 ± 14.4	12.8	94.2	16.6	59.6
CTA [20]	13.1 ± 4.6	14.0 ± 13.8	13.9	93.5	18.8	57.4
CNN	12.3 ± 3.9	13.3 ± 12.1	13.3	93.8	17.1	64.2
FBNet-RGB [48]	29.4 ± 3.0	29.7 ± 8.9	29.7	77.8	57.3	89.7
GrPPG [12]	31.3 ± 2.3	31.3 ± 7.6	32.1	74.3	67.5	94.6
PPGSec [36]	14.4 ± 3.0	15.0 ± 10.9	15.0	91.6	19.1	42.7
CFrPPG-crs	9.11 ± 1.7	9.53 ± 6.1	9.55	96.3	9.23	36.7
CFrPPG	3.84 ± 1.1	3.91 ± 2.7	3.92	99.2	2.19	6.59
MCCFrPPG	**2.88 ± 0.9**	**3.12 ± 3.2**	**3.17**	**99.6**	**1.23**	**4.84**

(a) ThatsMyFace

(b) REAL-f



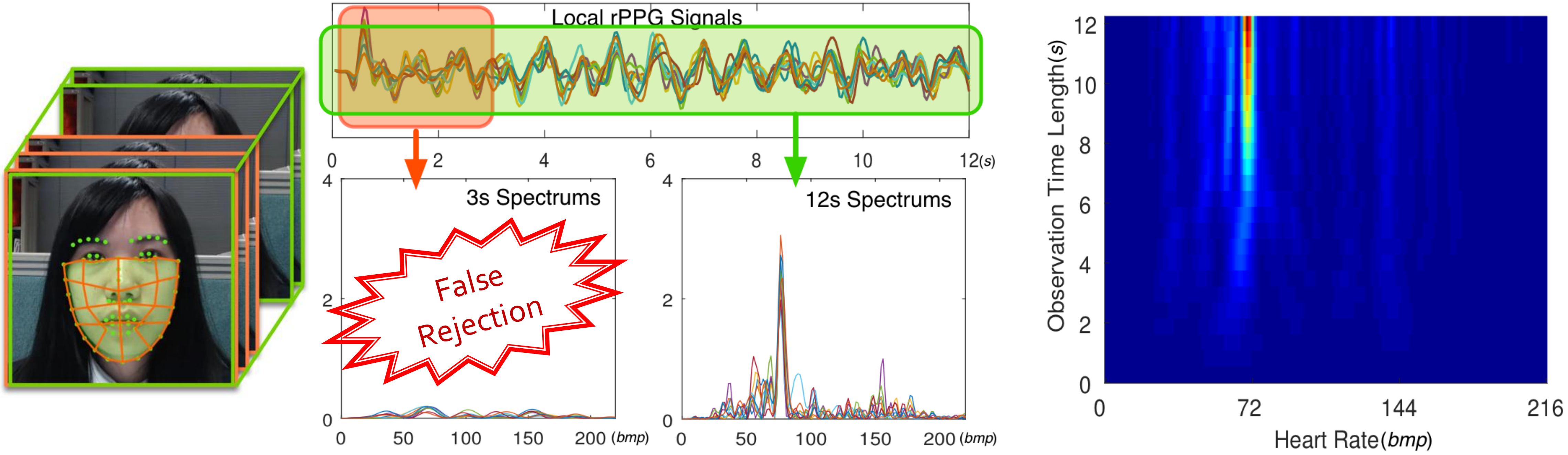
APCER: Attack presentation classification error rate" (false accept rate FAR

BPCER: "bona-fide presentation classification error rate" (false reject rate FRR)

CROSS-DATASET EVALUATION RESULTS (%) BETWEEN 3DMAD, HKBU-MARSV1+, HKBU-MARSV2+, AND CSMAD. A⇔B INDICATES THE EVALUATION ACROSS DATASETS A AND B, WHERE THE LEFT COLUMN IS A→B AND RIGHT ONE IS B→A. HTER STANDARD DEVIATION IS IN BRACKET

	Methods	3DMAD⇔MARsV1+		3DMAD⇔MARsV2+		3DMAD⇔CSMAD		MARsV1+⇔MARsV2+		MARsV1+⇔CSMAD		MARsV2+⇔CSMAD	
HTER	MS-LBP [7]	36.8 (2.9)	41.3 (14.0)	47.7 (7.0)	43.2 (7.3)	50.6 (5.6)	42.7 (6.4)	45.2 (3.9)	24.6 (5.6)	42.3 (3.2)	45.0 (5.8)	34.4 (3.8)	39.9 (2.2)
	CTA [20]	71.8 (2.1)	55.7 (8.7)	51.5 (2.4)	68.2 (7.7)	48.9 (5.8)	58.4 (7.8)	50.7 (4.8)	20.8 (5.4)	53.6 (5.0)	37.8 (4.8)	35.7 (3.3)	41.3 (3.7)
	CNN	49.4 (1.7)	62.5 (7.4)	50.8 (1.6)	46.5 (4.7)	45.6 (3.2)	46.5 (4.0)	31.3 (5.1)	33.8 (17.0)	45.9 (4.3)	42.6 (5.6)	45.7 (3.5)	42.3 (2.8)
	FBNet-RGB [48]	34.0 (1.4)	12.3 (10.6)	44.5 (0.3)	26.4 (21.5)	46.3 (2.3)	50.2 (18.1)	43.2 (1.4)	36.5 (5.4)	41.6 (3.7)	40.9 (6.8)	43.6 (3.4)	46.1 (3.4)
	GrPPG [12]	35.9 (4.5)	36.5 (6.8)	50.5 (0.2)	49.5 (4.0)	43.6 (3.7)	50.0 (0.0)	50.3 (0.2)	50.3 (3.3)	54.0 (11.4)	50.0 (0.0)	44.1 (3.2)	50.6 (0.3)
	PPGSec [36]	14.4 (1.4)	19.1 (2.3)	33.5 (0.5)	14.0 (2.0)	43.6 (1.5)	24.8 (11.9)	31.5 (1.6)	9.06 (1.4)	52.2 (2.2)	37.6 (3.9)	41.4 (3.8)	54.2 (4.6)
	CFrPPG-crs	4.46 (0.9)	8.46 (0.3)	31.4 (1.0)	8.44 (0.6)	40.5 (2.6)	17.0 (7.2)	27.3 (1.5)	5.02 (1.7)	40.4 (2.9)	13.8 (8.0)	36.0 (4.8)	31.7 (2.8)
	CFrPPG	4.23 (0.3)	4.81 (0.4)	11.0 (0.3)	6.71 (1.1)	22.7 (0.6)	6.37 (1.0)	11.0 (0.2)	3.21 (1.0)	22.5 (0.7)	**2.58 (0.8)**	22.7 (1.4)	10.4 (0.4)
	MCCFrPPG	**3.46 (0.6)**	**4.78 (0.8)**	**3.76 (0.2)**	**3.46 (0.6)**	**9.98 (0.4)**	**3.71 (0.8)**	**3.99 (0.2)**	**1.21 (0.6)**	**10.8 (0.5)**	2.67 (0.9)	**10.5 (0.7)**	**4.08 (0.4)**
AUC	MS-LBP [7]	60.7	62.2	52.4	58.8	49.5	58.2	53.4	75.3	52.3	54.8	68.8	64.1
	CTA [20]	45.9	48.6	48.9	40.1	50.7	46.5	52.8	84.5	48.8	61.5	67.1	62.6
	CNN	72.1	50.4	52.7	86.1	78.2	75.5	76.9	88.4	62.0	83.7	81.1	67.9
	FBNet-RGB [48]	73.6	89.6	56.2	74.5	56.6	52.5	59.1	69.3	57.0	56.1	58.1	56.0
	GrPPG [12]	67.2	66.5	49.9	49.9	52.7	50.0	49.8	49.8	48.9	50.0	59.9	50.0
	PPGSec [36]	91.8	87.2	73.5	91.8	60.7	77.2	76.4	96.5	52.1	53.3	61.7	58.7
	CFrPPG-crs	98.9	95.3	77.3	95.8	67.0	84.7	82.5	98.9	65.4	88.6	66.4	80.0
	CFrPPG	99.0	98.1	95.0	95.7	82.6	96.3	95.0	98.5	84.0	99.3	83.9	95.6
	MCCFrPPG	**99.6**	**98.5**	**99.1**	**97.1**	**95.7**	**98.6**	**99.3**	**99.8**	**95.3**	**99.7**	**95.1**	**99.3**
BPCER@ APCER=0.1	MS-LBP [7]	87.5	89.2	86.4	87.5	89.9	83.9	84.9	64.1	85.4	87.1	78.2	78.3
	CTA [20]	96.8	89.9	90.5	94.7	88.6	93.6	83.2	46.3	90.3	80.2	67.3	80.6
	CNN	86.4	90.5	90.7	35.4	49.3	69.1	61.3	30.2	84.9	49.8	65.1	75.5
	FBNet-RGB [48]	65.7	26.8	86.2	95.0	80.8	87.5	84.2	78.0	87.1	86.0	80.4	87.0
	GrPPG [12]	75.8	86.3	89.9	90.0	85.5	90.0	89.8	90.0	88.7	90.0	76.7	90.0
	PPGSec [36]	16.9	26.2	79.6	17.1	87.4	46.5	76.6	9.42	91.6	83.7	76.3	79.2
	CFrPPG-crs	1.33	8.79	80.4	8.03	63.5	46.3	61.6	2.25	65.1	39.2	56.7	71.6
	CFrPPG	2.83	4.44	12.9	9.88	51.1	8.26	12.4	4.13	47.8	1.29	41.7	10.9
	MCCFrPPG	**0.25**	**4.00**	**2.57**	**6.47**	**10.9**	**3.47**	**2.38**	**0.62**	**12.0**	**0.75**	**11.5**	**2.11**
BPCER@ APCER=0.01	MS-LBP [7]	97.0	99.5	97.6	99.2	98.1	98.9	96.4	98.8	97.9	99.0	96.6	95.8
	CTA [20]	99.3	97.4	98.8	99.9	98.1	99.3	95.9	93.0	98.7	96.2	88.7	97.4
	CNN	99.1	98.7	99.2	71.6	94.0	93.0	90.8	72.6	98.2	87.7	96.6	96.3
	FBNet-RGB [48]	97.8	66.8	97.3	99.9	99.0	98.5	96.8	96.2	99.1	97.2	94.3	98.4
	GrPPG [12]	97.6	98.6	99.4	99.7	96.5	99.0	100.1	100.5	98.6	99.0	85.6	98.6
	PPGSec [36]	25.8	45.0	94.8	36.1	98.6	64.6	94.8	17.3	99.9	96.0	96.2	89.5
	CFrPPG-crs	31.2	17.0	93.2	41.9	82.3	99.9	88.4	31.1	81.5	98.1	75.3	94.6
	CFrPPG	19.9	14.3	68.7	17.9	89.1	16.7	68.4	13.2	85.8	17.3	76.5	66.8
	MCCFrPPG	**7.63**	**8.59**	**10.8**	**10.7**	**46.9**	**8.09**	**7.92**	**2.79**	**40.9**	**5.50**	**38.6**	**7.72**

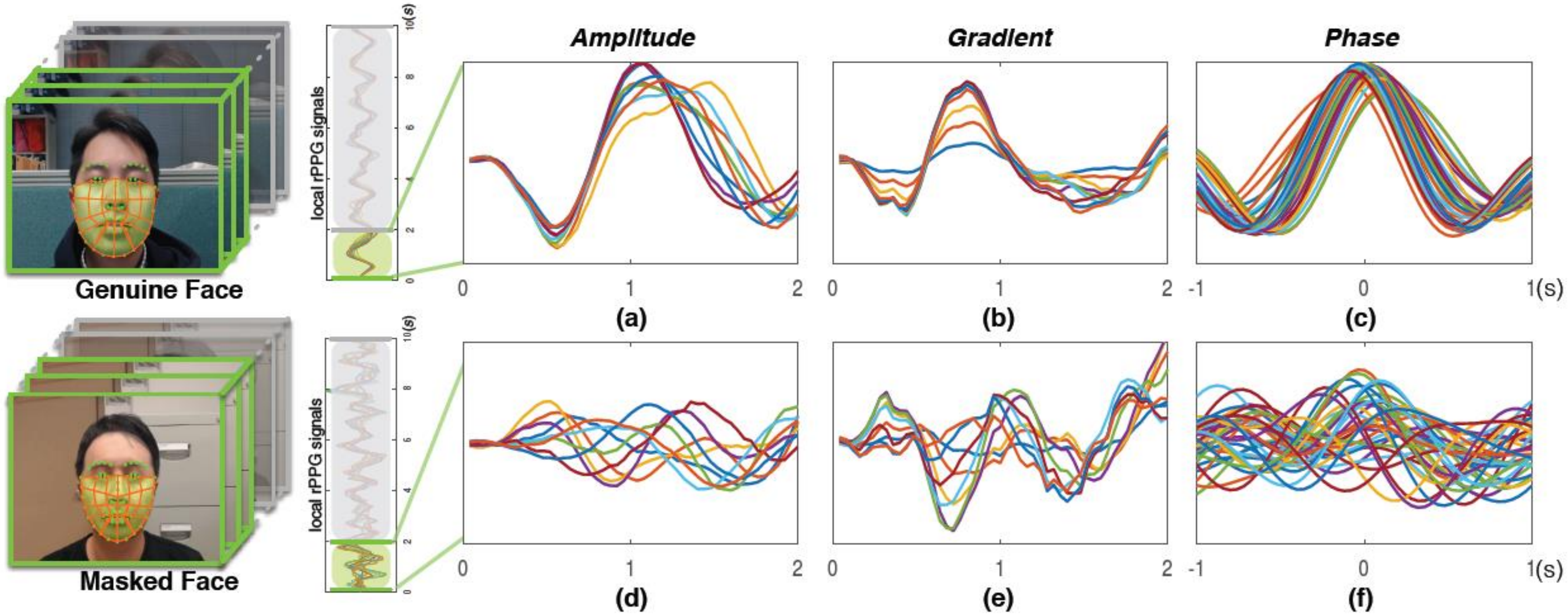
Limitations on existing rPPG Methods



Existing rPPG-based 3D mask PAD methods are based on spectrum analysis

→ Require long observation time (8-10 seconds) to identify heartbeat information

Learnable Temporal Similarity Analysis of rPPG (TSrPPG) for Fast 3D Mask Face PAD

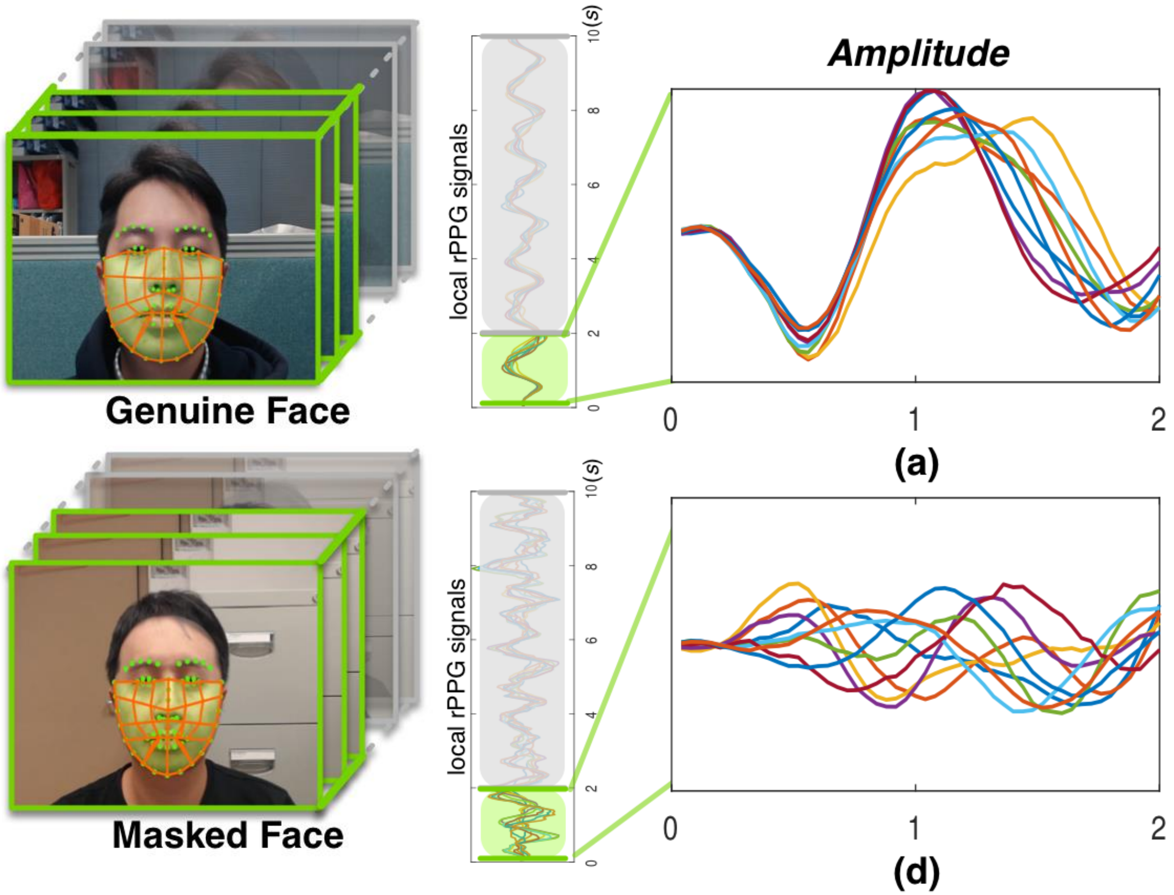


Reference:

1. S Q Liu, XY Lan, and P C Yuen, "Temporal Similarity Analysis of Remote Photoplethysmography (TSrPPG) for Fast 3D Mask Face Presentation Attack Detection", *WACV*, 2020.

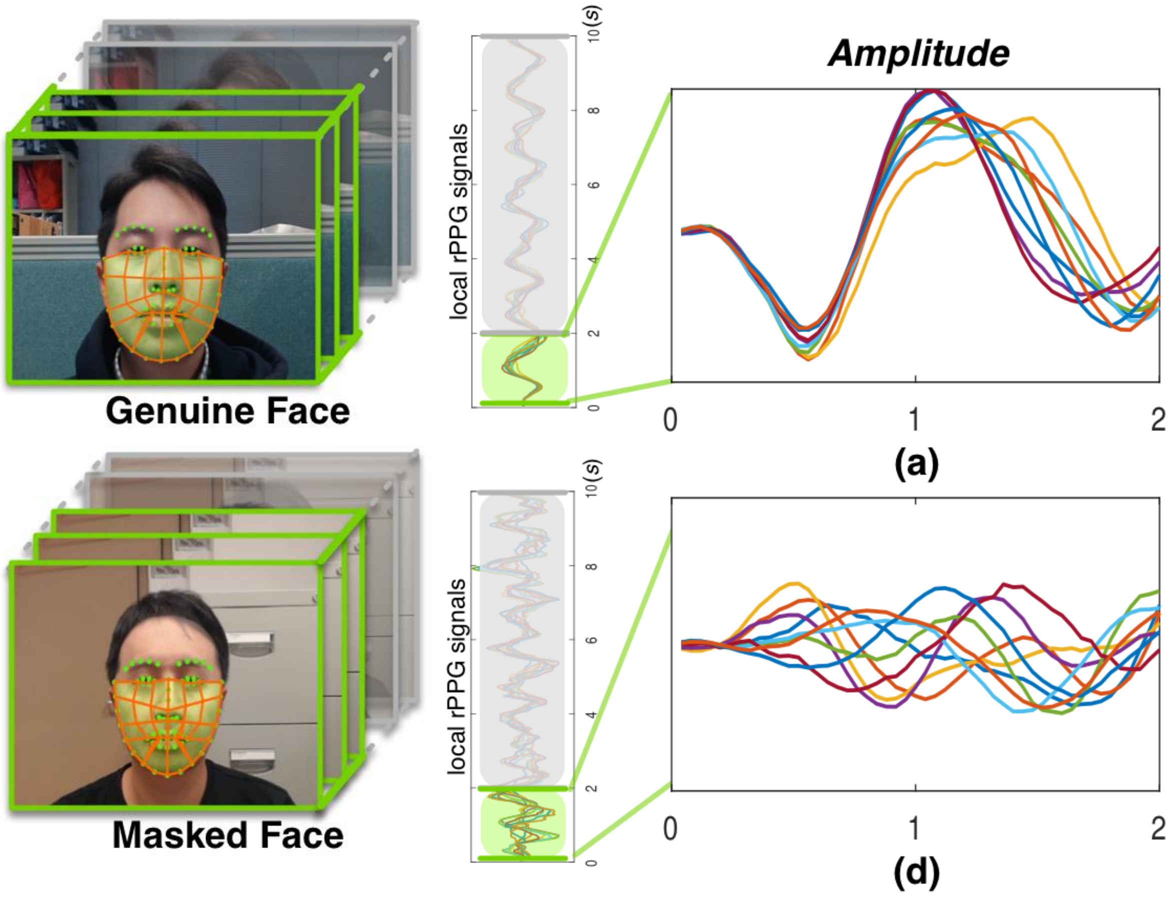
2. S Q Liu, X Y Lan and P C Yuen, "Learning Temporal Similarity of Remote Photoplethysmography for Fast 3D Mask Face Presentation Attack Detection", *IEEE Transactions on Information Forensics and Security (TIFS)*,2022.

The proposed TSrPPG

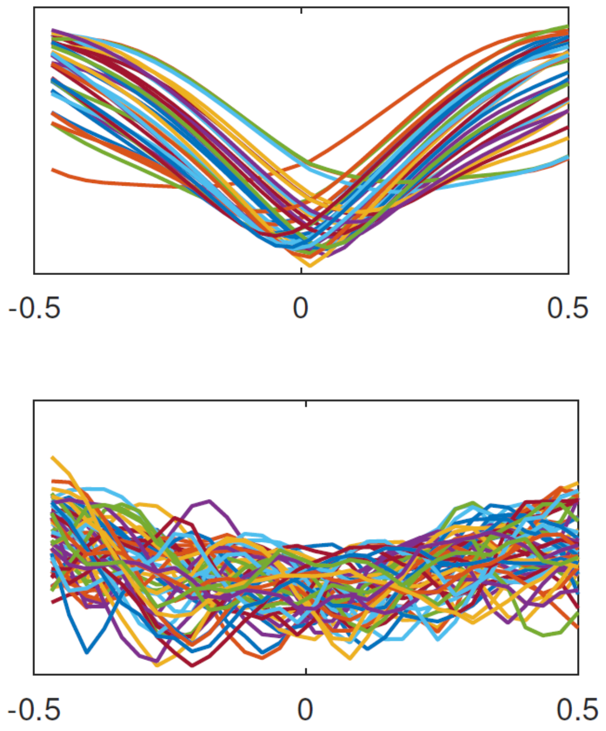


- Rationale
  - The periodicity information is not available within short observation time.
    - Hard to adopt spectrum analysis
  - Correlation of local rPPG signals on genuine faces is higher compared with those on masked faces.
  - Design liveness feature in temporal space

The proposed TSrPPG



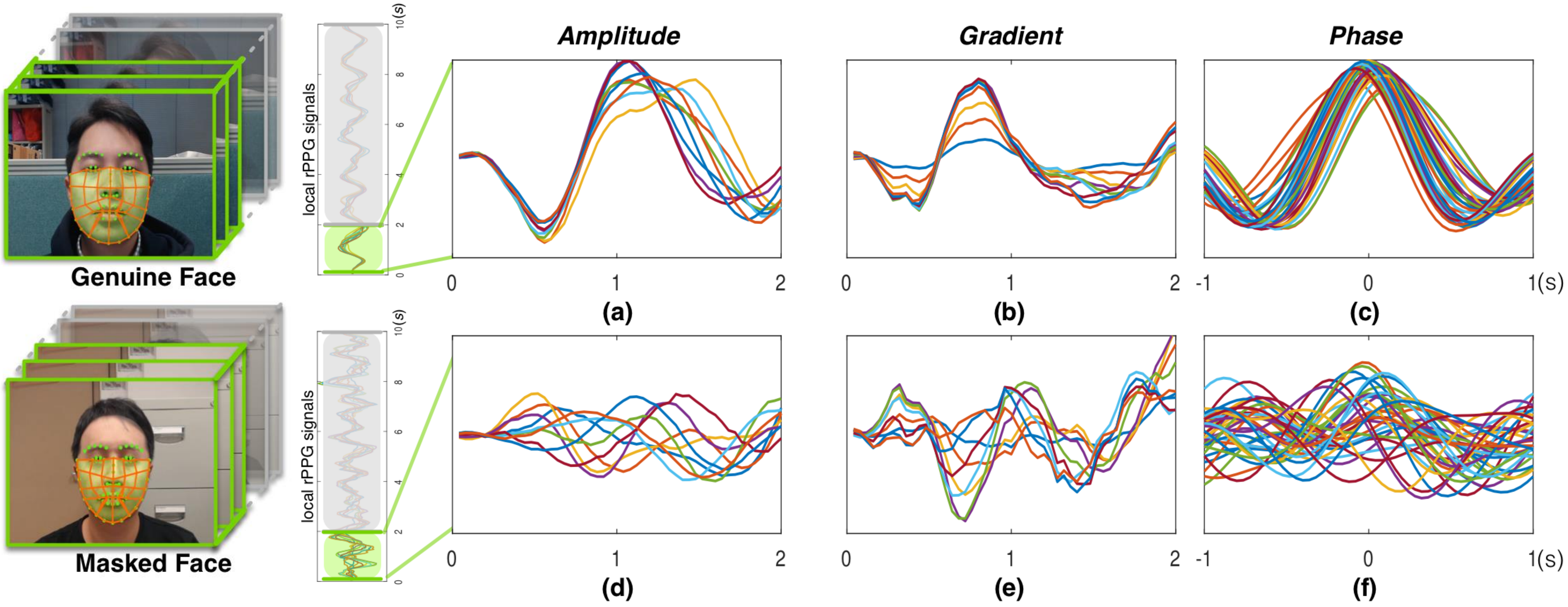
$$\boldsymbol{TSrPPG}_{i,j}[m] = \int_{-\infty}^{+\infty} \mathcal{D}(\boldsymbol{s}_i[t], \boldsymbol{s}_j[t+m])dt$$



Extract features on the result pattern
→ Min, Mean, Std (… etc.)

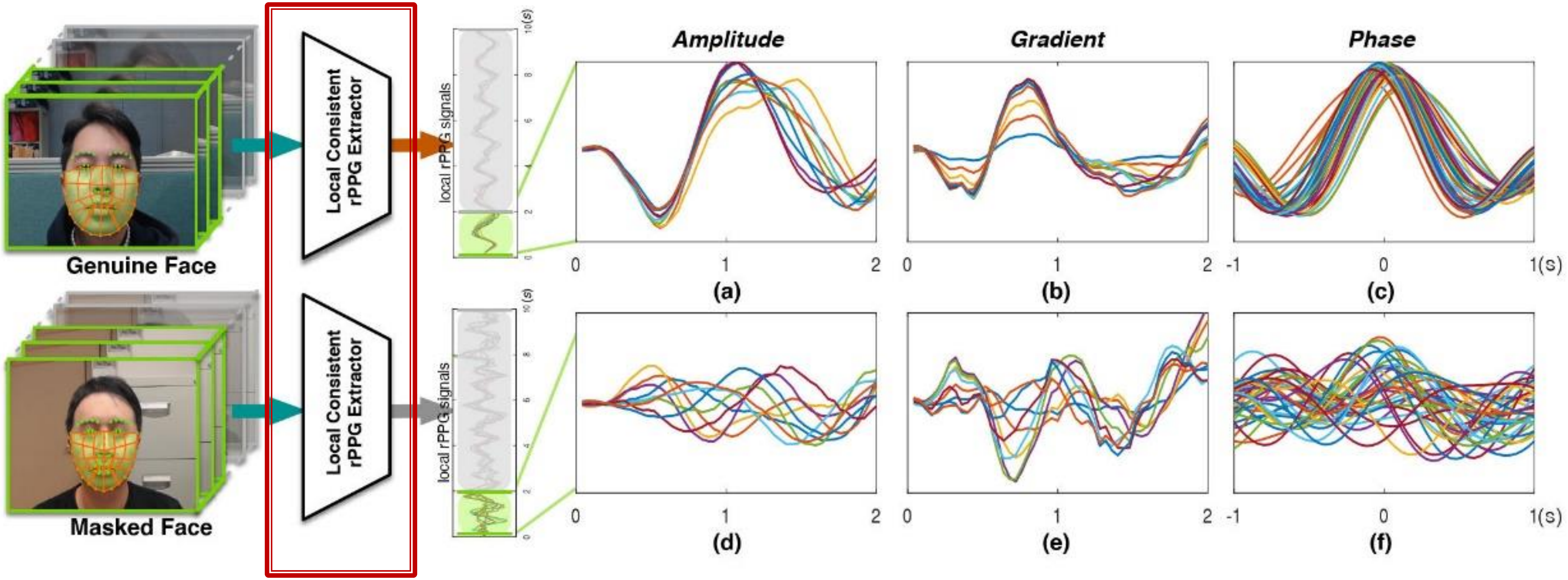
The proposed TSrPPG

$$\boldsymbol{TSrPPG_{i,j}}[m] = \int_{-\infty}^{+\infty} \mathcal{D}(\boldsymbol{s}_i[t], \boldsymbol{s}_j[t+m])dt$$



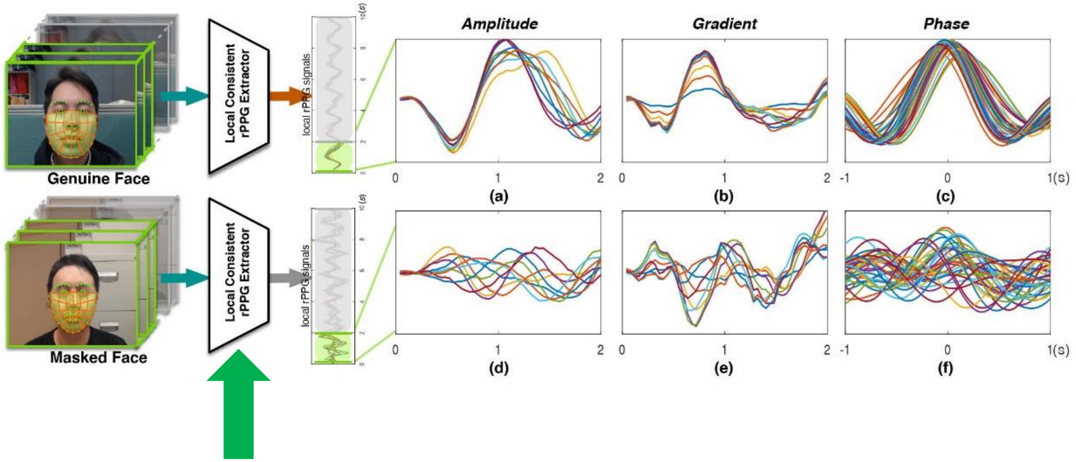
Final result is obtained through score-level-fusion

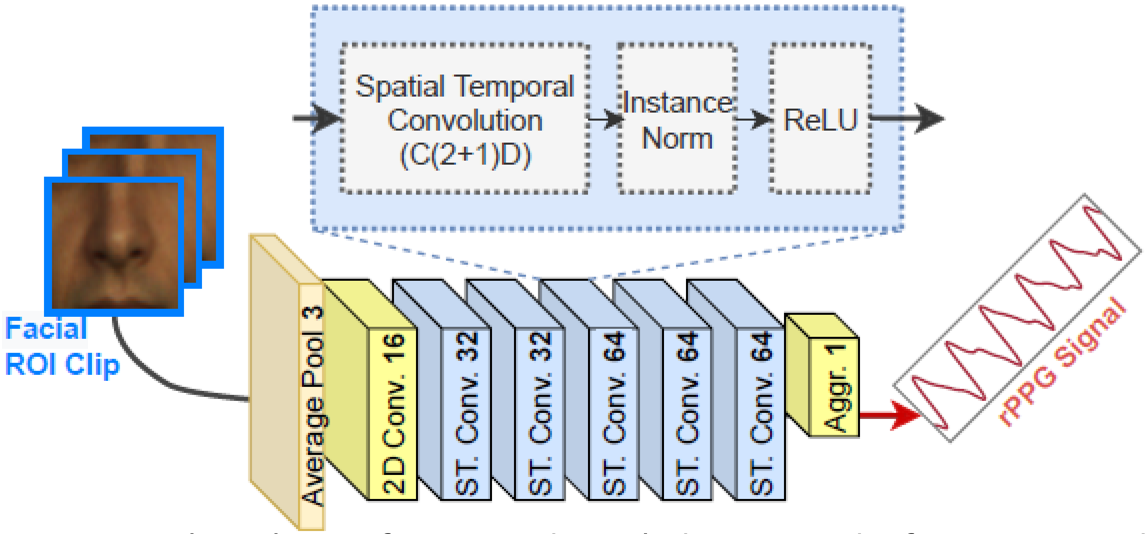
LeTSrPPG: Learnable rPPG to enhance Temporal Similarity of TSrPPG



LeTSrPPG: Learnable rPPG to enhance temporal similarity of TSrPPG

- Learnable rPPG estimator:
  - Learn robust rPPG feature through 3D convolution

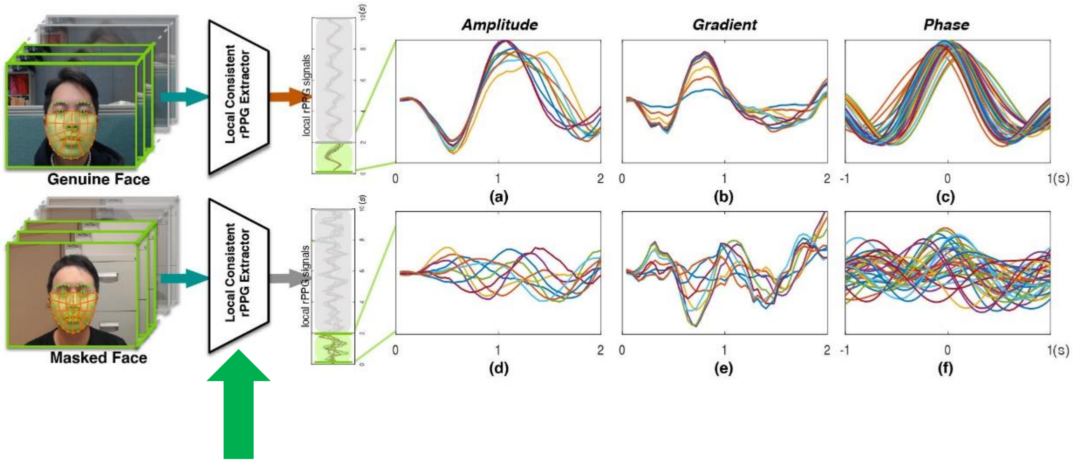


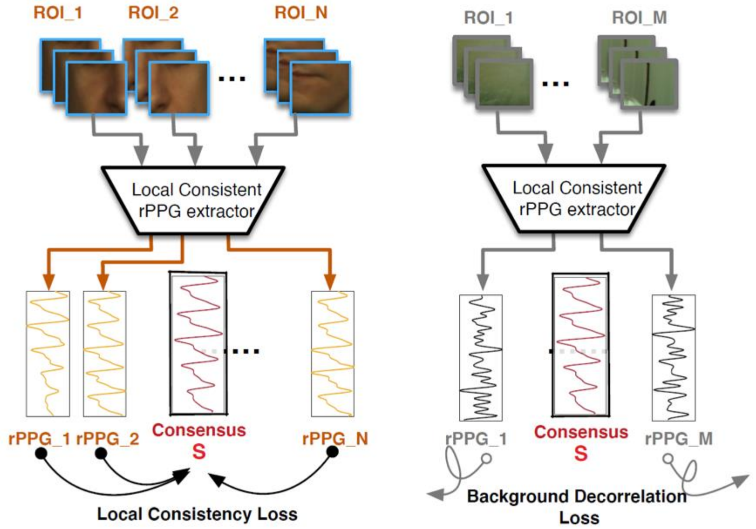


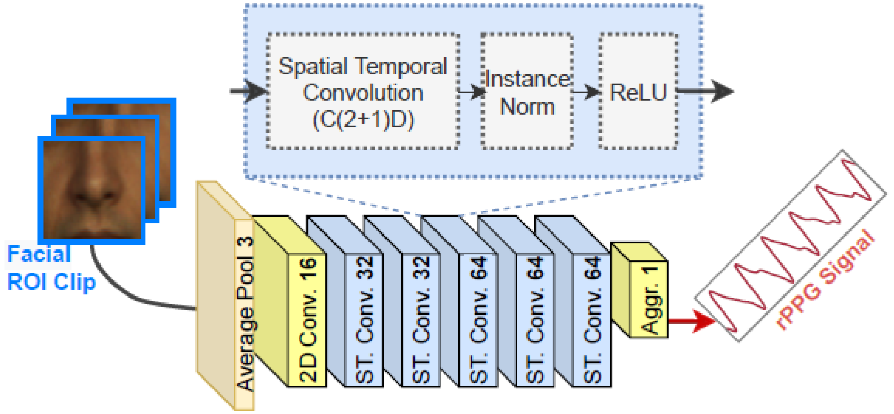
S Q Liu, X Y Lan and P C Yuen, "Learning Temporal Similarity of Remote Photoplethysmography for Fast 3D Mask Face Presentation Attack Detection", *IEEE Transactions on Information Forensics and Security (TIFS)*, 2022.

LeTSrPPG: Learnable rPPG to enhance temporal similarity of TSrPPG

- Learnable rPPG estimator:
  - Learn robust rPPG feature through 3D convolution
  - Boost the discriminability of TSrPPG using local consistency loss
    - Genuine face: Enhance the temporal similarity
    - Fake face: Reduce the temporal similarity



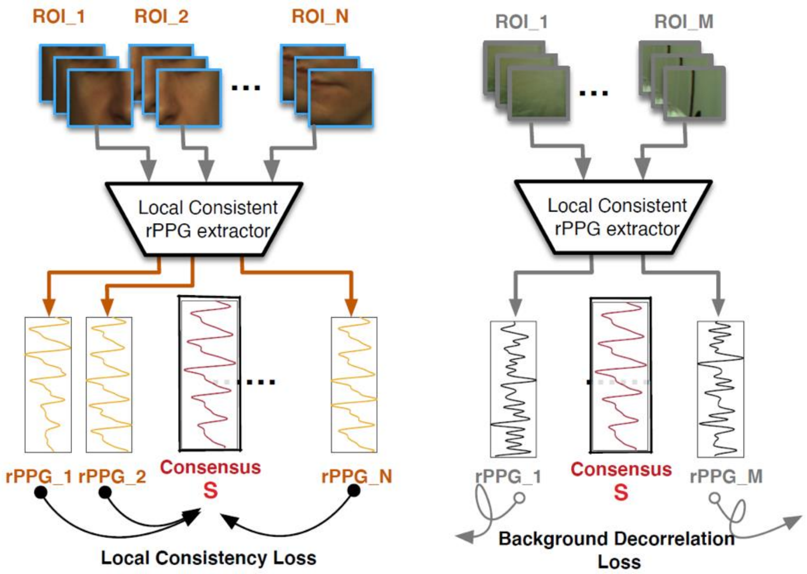


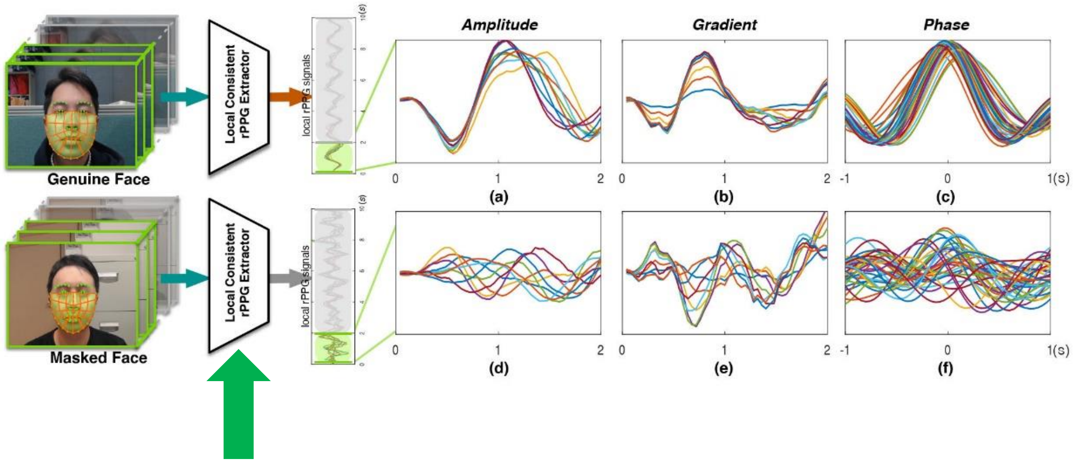


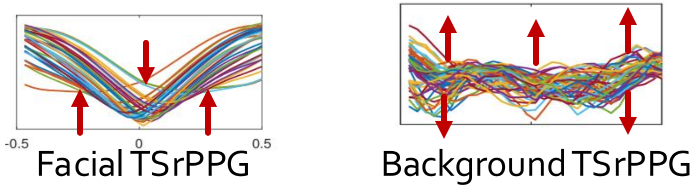
S Q Liu, X Y Lan and P C Yuen, "Learning Temporal Similarity of Remote Photoplethysmography for Fast 3D Mask Face Presentation Attack Detection", *IEEE Transactions on Information Forensics and Security (TIFS)*, 2022.

LeTSrPPG: Learnable rPPG to enhance temporal similarity of TSrPPG

- Learnable rPPG estimator:
  - Learn robust rPPG feature through 3D convolution
  - Further boost the discriminability of TSrPPG
    - Genuine face: Enhance the temporal similarity
    - Fake face: Reduce the temporal similarity







- Improve TSrPPG in rPPG extraction stage
  - Enhance the consistency of local rPPG signals
  - Reduce the correlation between background rPPG and facial rPPG

S Q Liu, X Y Lan and P C Yuen, "Learning Temporal Similarity of Remote Photoplethysmography for Fast 3D Mask Face Presentation Attack Detection", *IEEE Transactions on Information Forensics and Security (TIFS)*, 2022.

LeTSrPPG: Learnable rPPG to enhance temporal similarity of TSrPPG

- Experimental Setting:

	#Subjects/Masks	#Video Slots	Mask Type	Lighting Condition	Camera	Face (pixel) Resolution	Compression
3DMAD [13]	17 17	2550	TMF	1(Studio)	Kinect	80×80	Motion JPEG
HKBU-MARsV1+ [15]	12 12	2160	TMF+RF	1(Room)	Logitech C920	200×200	H.264
CSMAD [30]	14 6	1582	Silicon	4	RealSense SR300	350×350	H.264
HKBU-MARsV2+	16 16	12480	TMF+RF	6	3	3	2
Summary	59 39	18772	3	12	6	5*	2

(a) ThatsMyface (b) REAL-f (c) Silicone

- Evaluation Protocols:
  - Intra-dataset evaluation
    - Leave one subject out cross validation (**LOOCV**)
  - Cross-dataset evaluation
    - Train and test on different datasets

LeTSrPPG: Learnable rPPG to enhance Temporal Similarity of TSrPPG

- Intra dataset evaluation with short observation time (1 second) :

	HTER_dvlp	HTER_test	EER	AUC
GrPPG	34.1 ± 5.7	33.7 ± 11.6	38.3	65.9
PPGSec	33.3 ± 3.1	33.0 ± 8.1	34.8	69.4
LrPPG	45.2 ± 3.2	44.8 ± 8.8	45.3	55.7
CFrPPG	32.8 ± 1.7	32.7 ± 7.4	32.5	70.8
TransrPPG	20.7 ± 2.2	20.6 ± 8.3	20.8	84.5
TSrPPG	13.1 ± 3.0	13.4 ± 11.2	13.3	93.8
LeTSrPPG	**11.5 ± 2.7**	**11.8 ± 8.6**	**11.9**	**94.4**

3DMAD

	HTER_dvlp	HTER_test	EER	AUC
GrPPG	29.2 ± 4.7	29.1 ± 9.7	33.8	72.0
PPGSec	42.4 ± 2.1	42.9 ± 5.8	43.0	59.3
LrPPG	45.3 ± 3.7	45.1 ± 12.0	45.3	56.2
CFrPPG	41.6 ± 3.3	42.1 ± 5.6	42.0	60.8
TransrPPG	32.9 ± 2.8	32.7 ± 6.4	33.1	72.0
TSrPPG	21.5 ± 2.6	22.3 ± 8.8	22.0	85.2
LeTSrPPG	**15.3 ± 2.2**	**15.8 ± 6.5**	**15.7**	**91.5**

HKBU-MARsV1+

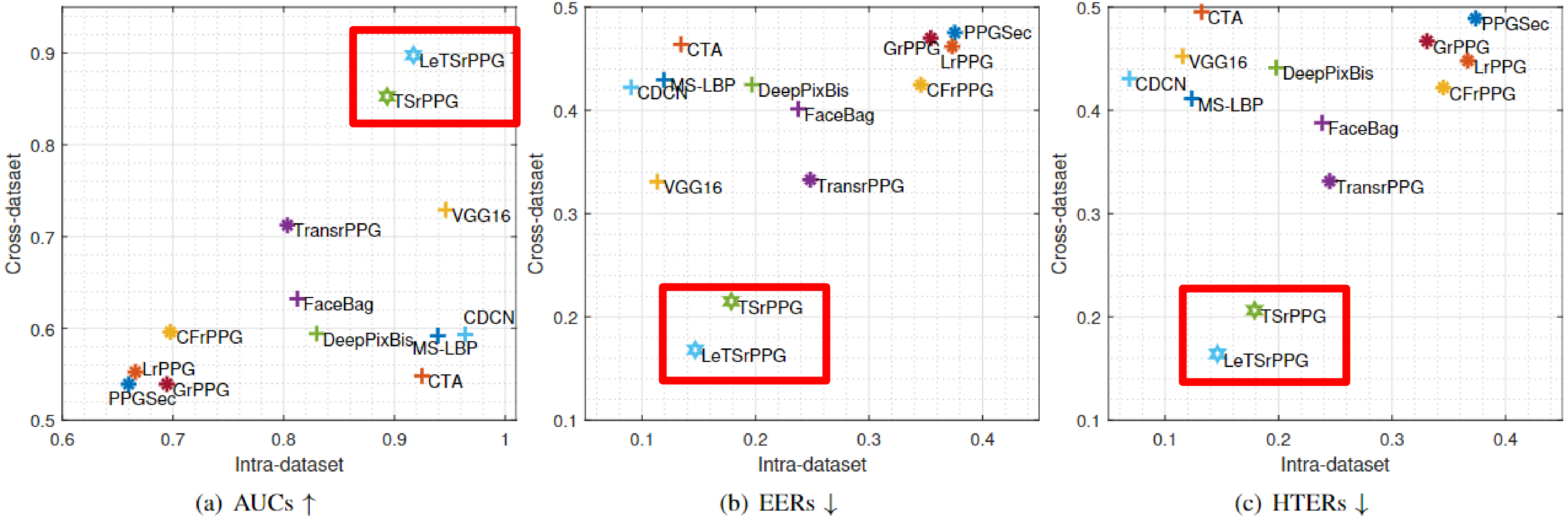
	3DMAD				HKBUMARsV1+			
	1s	2s	3s	4s	1s	2s	3s	4s
GrPPG [14]	65.9	79.1	84.6	87.7	72.0	79.2	80.3	82.3
LrPPG [13]	69.4	84.1	89.3	92.0	59.3	71.5	78.8	84.5
PPGSec [40]	55.7	68.3	74.5	80.0	56.2	74.4	76.7	79.8
CFrPPG [15]	70.8	88.1	93.1	94.4	60.8	78.6	85.8	89.0
TransrPPG [41]	84.5	87.3	89.4	88.1	72.0	76.8	77.6	79.6
TSrPPG	93.8	97.0	97.7	98.4	85.2	89.0	89.9	90.3
LeTSrPPG	**94.4**	**97.1**	**98.0**	**98.6**	**91.5**	**96.0**	**97.3**	**98.0**

Performance (AUC) with different length of observation

S Q Liu, X Y Lan and P C Yuen, "Learning Temporal Similarity of Remote Photoplethysmography for Fast 3D Mask Face Presentation Attack Detection", *IEEE Transactions on Information Forensics and Security (TIFS)*, 2022.

LeTSrPPG: Learnable rPPG to enhance temporal similarity of TSrPPG

- Overall comparison with state of the arts on both intra and cross dataset evaluation (1 second)
  - TSrPPG and LeTSrPPG achieve the best robustness and top-level discriminability



S Q Liu, X Y Lan and P C Yuen, "Learning Temporal Similarity of Remote Photoplethysmography for Fast 3D Mask Face Presentation Attack Detection", *IEEE Transactions on Information Forensics and Security (TIFS)*, 2022.

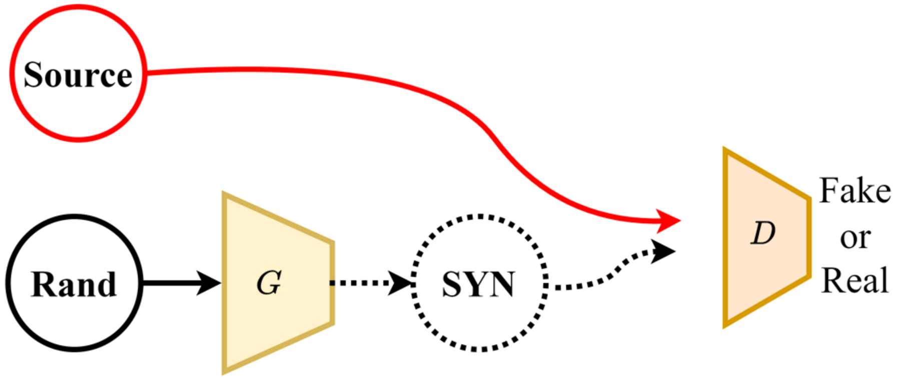
How to improve the robustness and generalization of rPPG estimation (1)?

Existing Approaches for Cross Domain Estimation

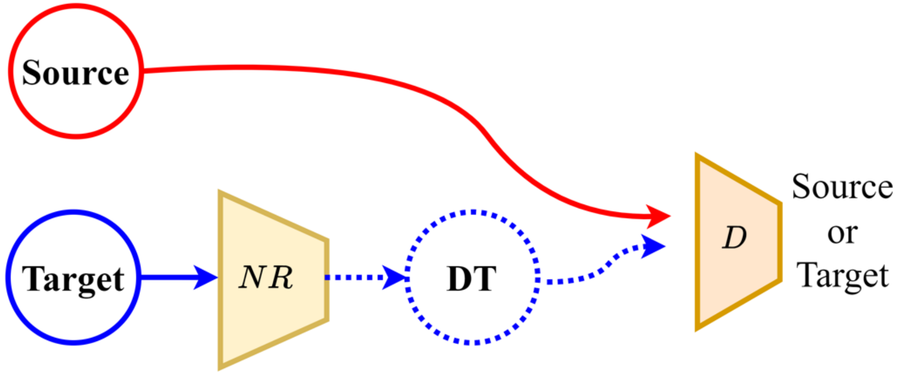
Problems:

- Robustness of rPPG estimation
- Generalisation to unseen interference

- Solution 1: GAN-based
  - Perform well under intra-dataset evaluation
  - Not aim to handle unseen scenarios

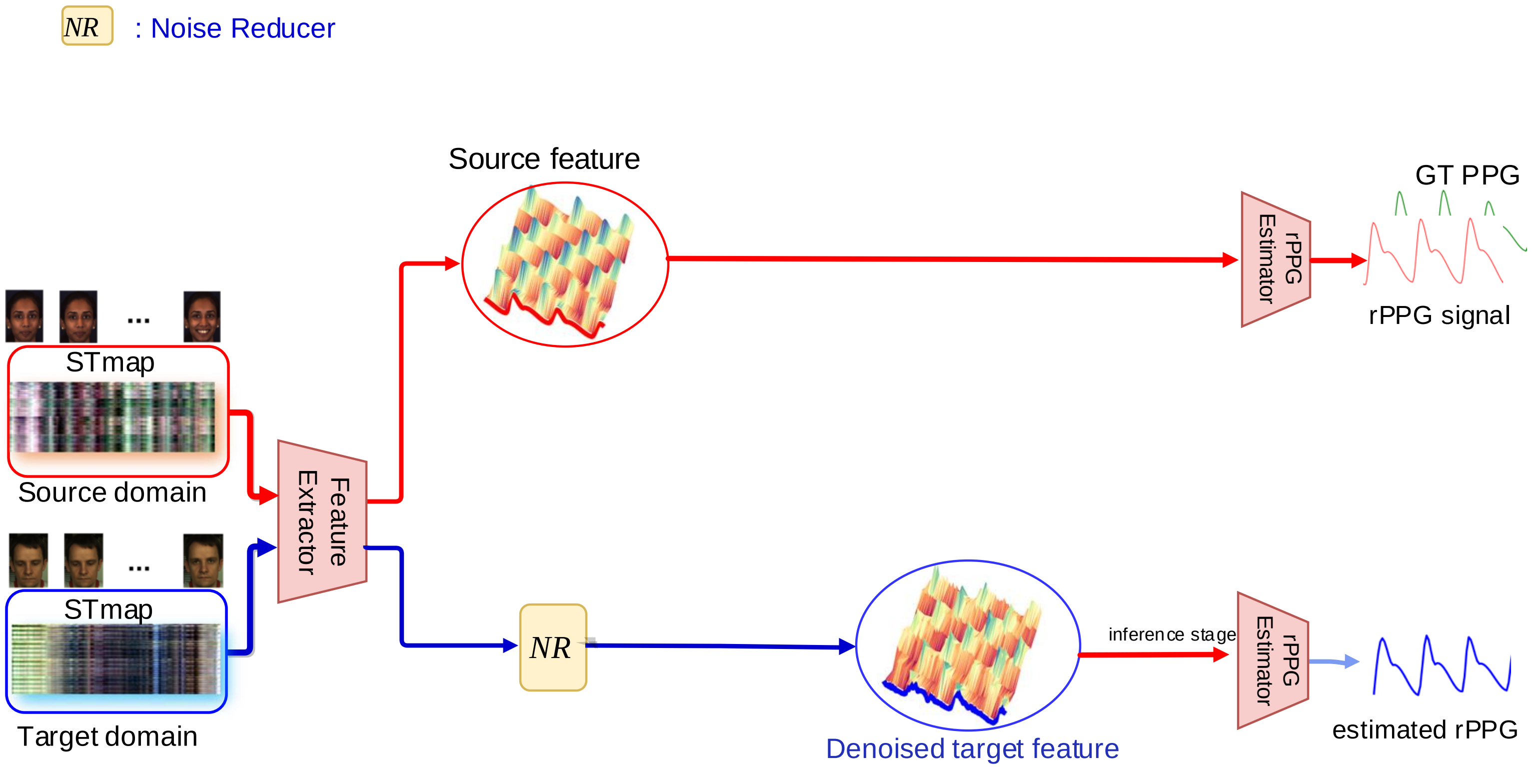


- Solution 2: Unsupervised domain adaptation
  - Denoise -> domain invariant feature
  - Successful experience in natural image tasks
  - Domain classification may not give sufficient information in rPPG regression task



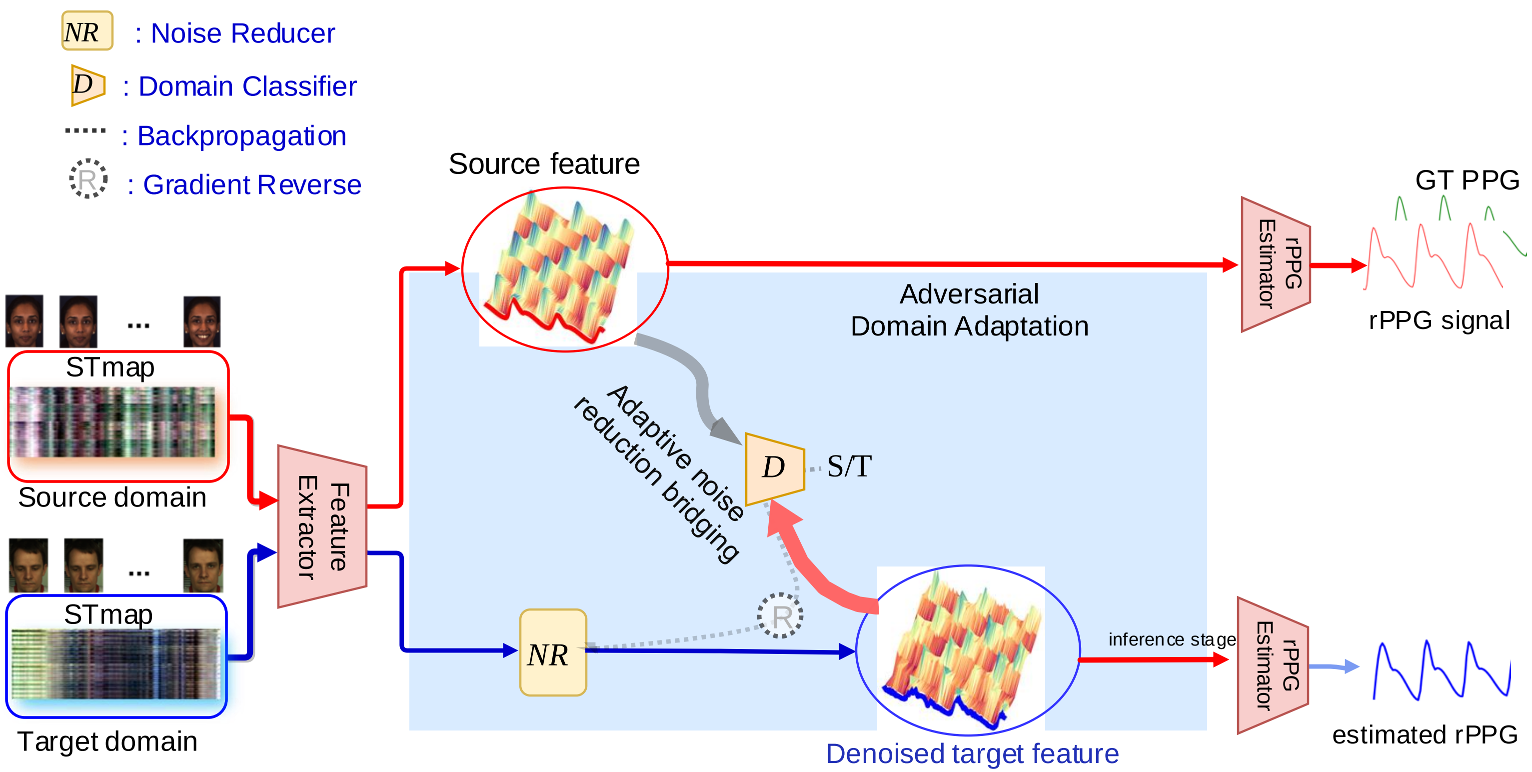
[1] Ganin, Yaroslav, et al. "Domain-adversarial training of neural networks." *The journal of machine learning research*, 2016.
[2] G. Wei, C. Lan, W. Zeng, Z. Zhang, and Z. Chen, “Toalign: Task-oriented alignment for unsupervised domain adaptation,” Advances in Neural Information Processing Systems, vol. 34, 2021

Dual-bridging with Adversarial Noise Generation for Domain Adaptive rPPG Estimation [CVPR23]



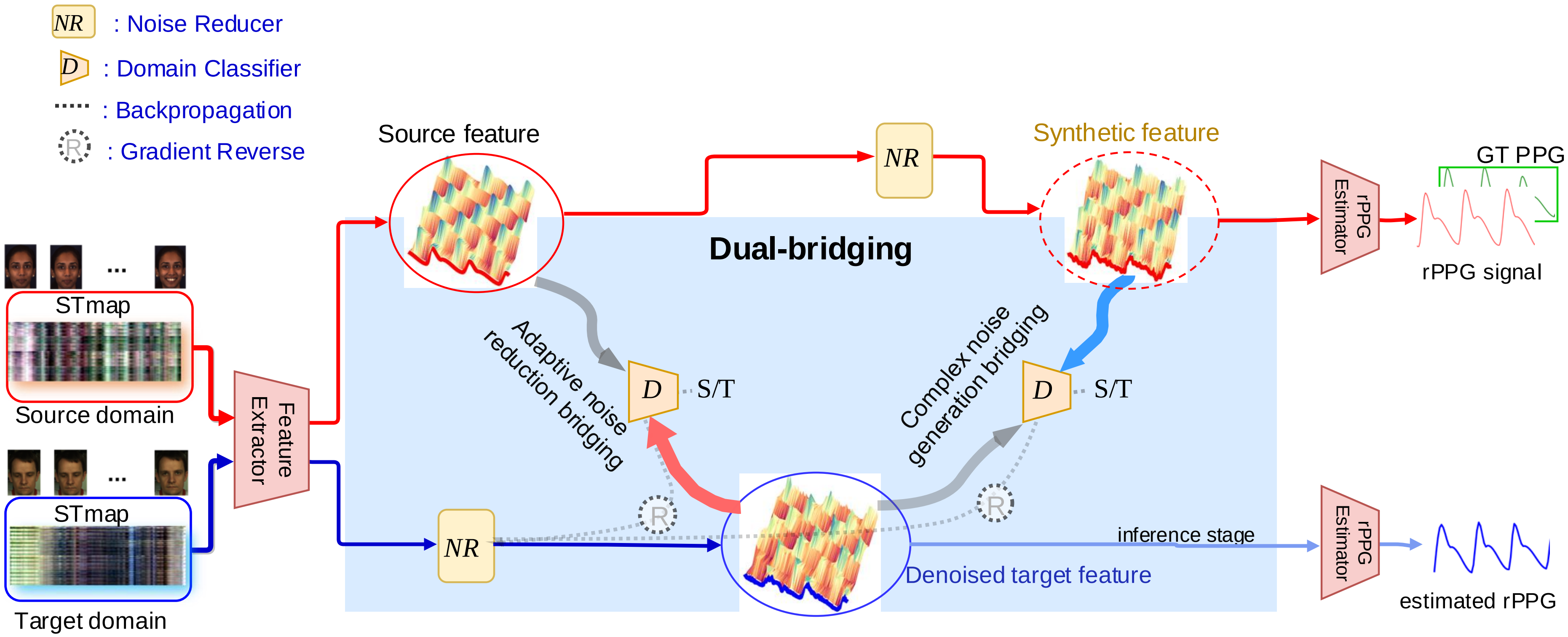
J Du, S Liu, B Zhang, P C Yuen, "Dual-bridging with Adversarial Noise Generation for Domain Adaptive rPPG Estimation", CVPR 2023

Dual-bridging with Adversarial Noise Generation for Domain Adaptive rPPG Estimation



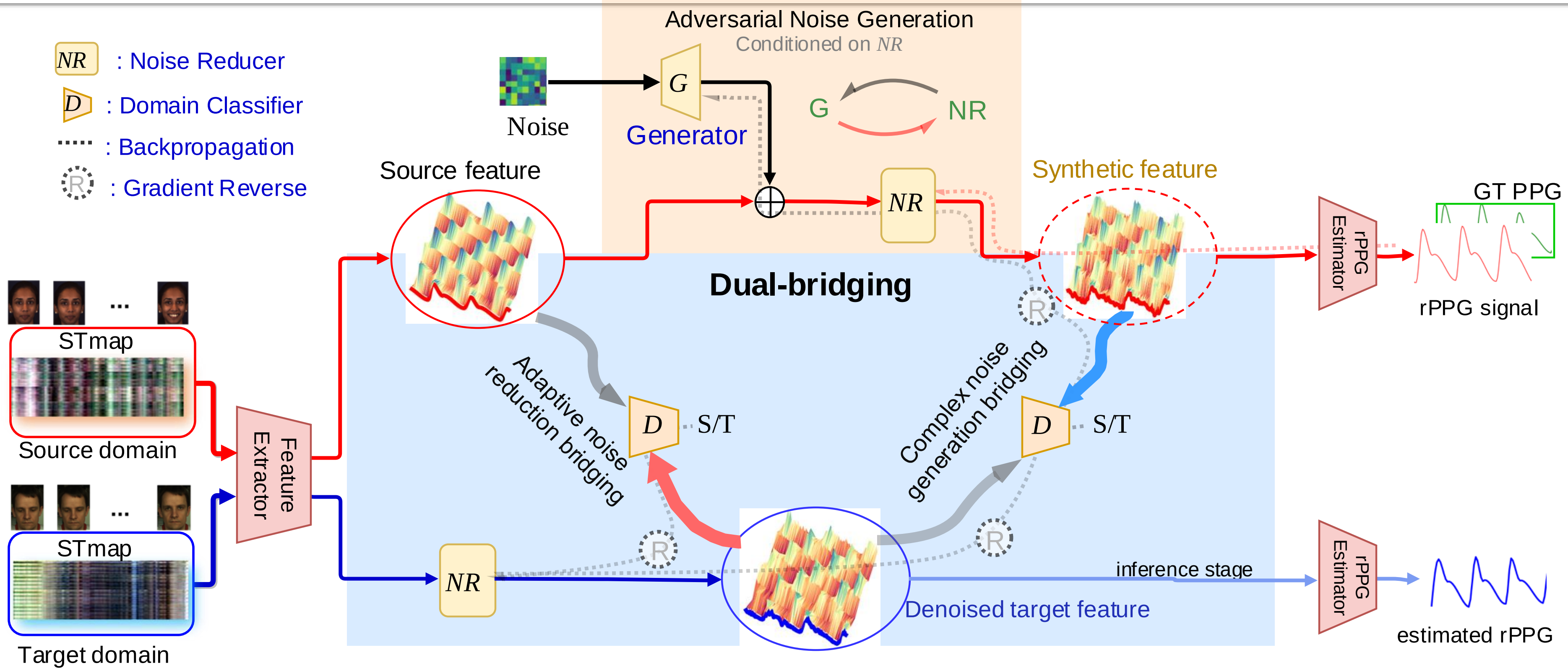
J Du, S Liu, B Zhang, P C Yuen, "Dual-bridging with Adversarial Noise Generation for Domain Adaptive rPPG Estimation", CVPR 2023

Dual-bridging with Adversarial Noise Generation for Domain Adaptive rPPG Estimation



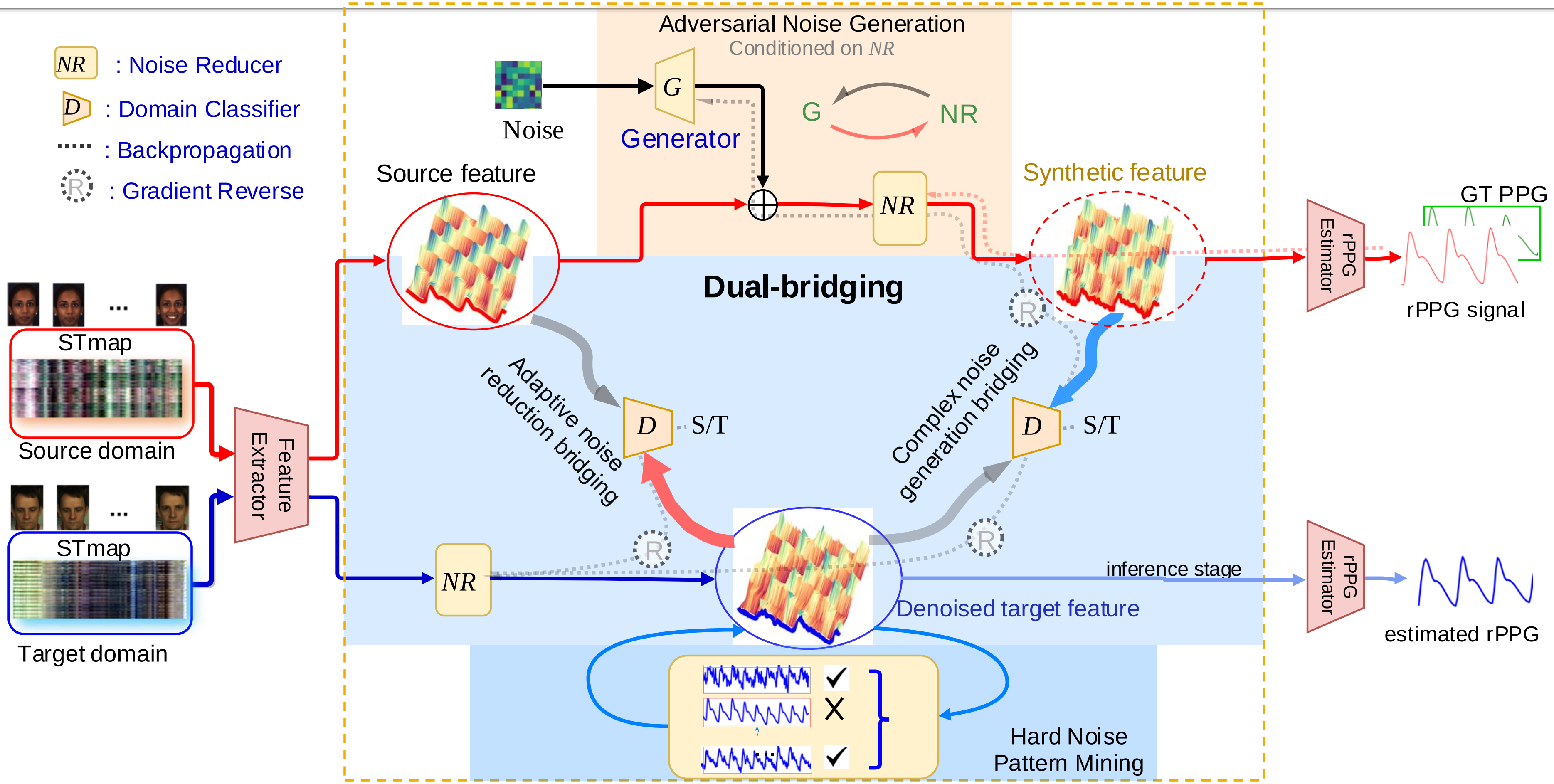
J Du, S Liu, B Zhang, P C Yuen, "Dual-bridging with Adversarial Noise Generation for Domain Adaptive rPPG Estimation", CVPR 2023

Dual-bridging with Adversarial Noise Generation for Domain Adaptive rPPG Estimation



J Du, S Liu, B Zhang, P C Yuen, "Dual-bridging with Adversarial Noise Generation for Domain Adaptive rPPG Estimation", CVPR 2023

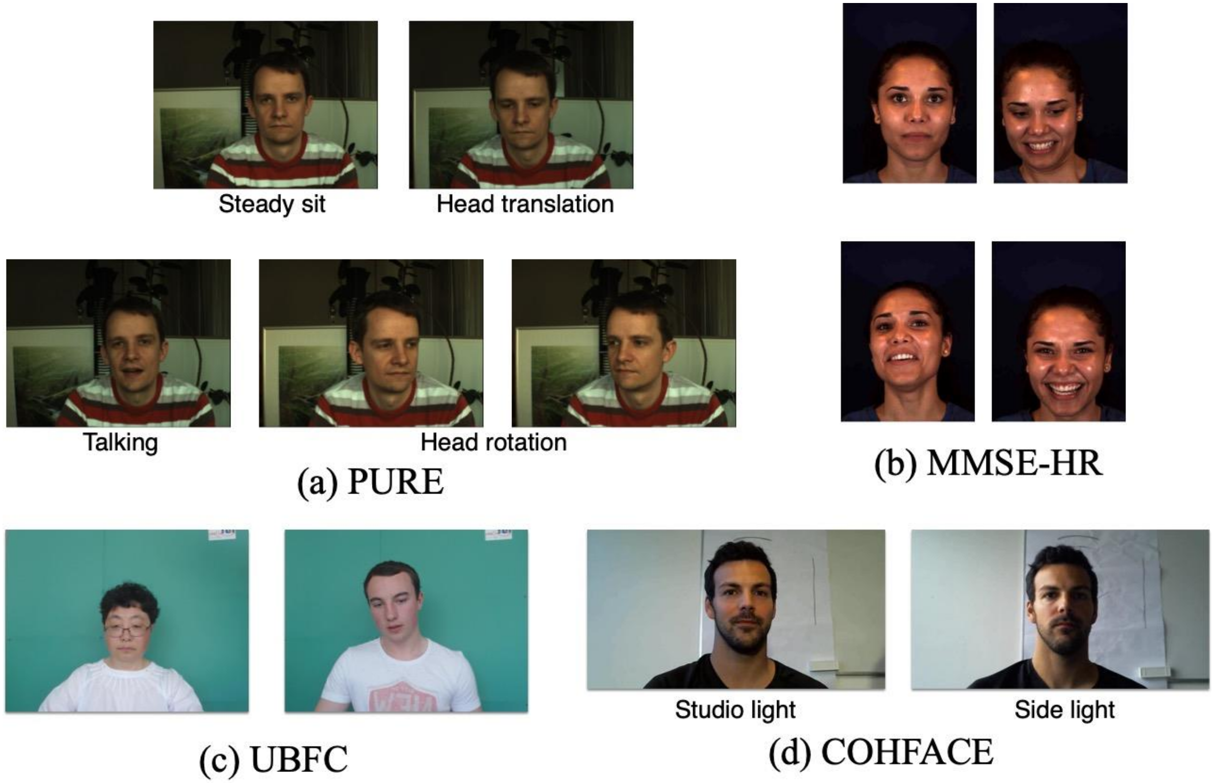
Dual-bridging with Adversarial Noise Generation for Domain Adaptive rPPG Estimation



J Du, S Liu, B Zhang, P C Yuen, "Dual-bridging with Adversarial Noise Generation for Domain Adaptive rPPG Estimation", CVPR 2023

Experiments

- 4 datasets
  - PURE, MMSE-HR, UBFC, COHFACE
- Variations:
  - Illumination
  - Facial motion and expression,
  - Camera and video compression
  - Skin tone
  - Heartbeat ranges



(a) PURE
(b) MMSE-HR
(c) UBFC
(d) COHFACE

Dual-bridging with Adversarial Noise Generation for Domain Adaptive rPPG Estimation

Task-independent evaluation on MMSE-HR dataset

Method	MAE	RMSE	r
Li2014 [16]	-	19.95	0.38
CHROM [5]	-	13.97	0.55
Tulyakov2016 [39]	-	11.37	0.71
ST-Attention* [29]	-	10.10	0.64
RhythmNet [27]	-	5.03	0.86
CVD* [28]	-	6.04	0.84
PhysNet [47]	-	13.25	0.44
DeepPhys [3]	4.43	9.98	0.80
TS-CAN [21]	3.85	7.21	0.86
AutoHR [45]	-	5.87	0.89
BVPNet [4]	-	7.47	0.79
Federated2022 [23]	2.99	2.42	0.79
EfficientPhys-C [22]	2.91	5.43	0.92
EfficientPhys-T1 [22]	3.48	7.21	0.86
PhysFormer* [49]	2.84	5.36	0.92
ERM [12]	1.30	2.58	**0.99**
DANN [7]	1.24	2.71	**0.99**
CST [17]	1.20	2.42	**0.99**
Ours	**0.85**	**2.05**	**0.99**

* Trained on VIPL-HR datasets due to the large model-scale

Method	MAE	RMSE	r
GREEN [40]	4.47	11.6	0.842
ICA [32]	3.51	8.64	0.908
CHROM [5]	3.44	4.61	0.968
POS [41]	2.44	6.61	0.936
CK [35]	2.29	3.80	0.981
Frédéric [2]	5.45	8.64	-
HeartTrack [31]	2.41	3.37	0.983
ETA-rPPGNet [10]	1.46	3.97	0.93
DAE [34]	1.48	2.49	0.97
PulseGAN [34]	1.19	2.10	0.98
Meta-rPPG [13]	5.97	7.42	0.53
CVD [28]	2.19	3.12	0.99
Gideon2021 [8]	3.6	4.6	0.95
Federated2022 [23]	2.00	4.38	0.93
Dual-GAN [24]	0.44	0.67	0.99
ContrastPhys [37]	0.64	1.00	0.99
ERM [12]	0.75	1.84	**0.99**
DANN [7]	0.58	1.19	**0.99**
CST [17]	0.41	1.04	**0.99**
Ours	**0.16**	**0.57**	**0.99**

Participant-independent evaluation on UBFC-rPPG dataset

Dual-bridging with Adversarial Noise Generation for Domain Adaptive rPPG Estimation

	MMSE-HR $\rightarrow$ PURE			PURE $\rightarrow$ MMSE-HR		
Method	MAE	RMSE	r	MAE	RMSE	r
CHROM [5]	3.25	12.92	0.84	5.72	12.69	0.58
POS [41]	2.83	12.49	0.85	4.98	13.11	0.53
CVD [28]	2.75	3.98	0.98	4.08	7.03	0.84
ERM [12]	2.49	8.48	0.93	2.59	5.44	0.96
DANN [7]	2.69	6.97	0.95	2.84	7.65	0.93
CST [17]	1.27	2.96	**0.99**	2.32	5.97	0.96
EfficientT1 [22]	-	-	-	3.04	5.91	0.92
PhysFormer [49]	-	-	-	2.84	5.36	0.92
Synthetic [25]	-	-	-	2.26	3.70	0.97
Ours	**1.10**	**1.67**	**0.99**	**1.71**	**3.72**	**0.98**

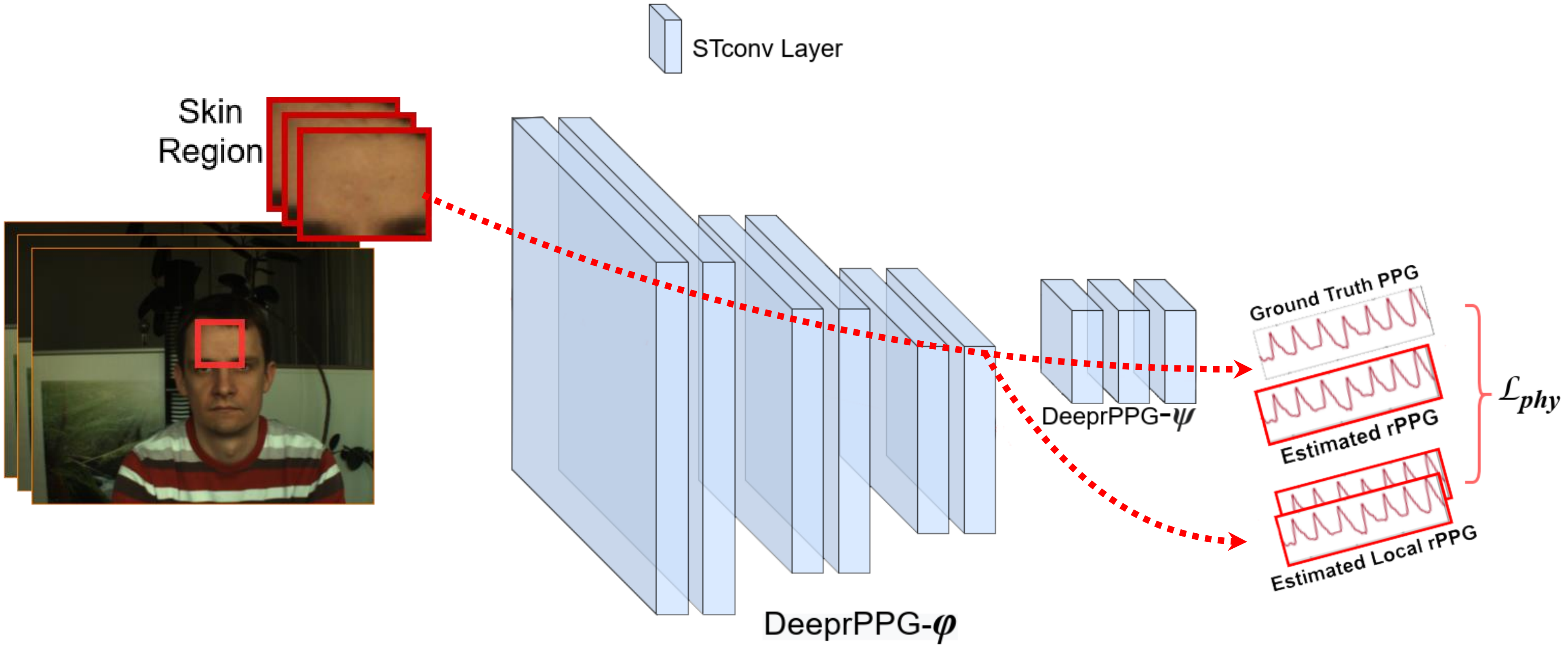
Cross-datasets

How to further improve the noise robustness and generalization of rPPG estimation (2)?

Noise Disentangled DeeprPPG Framework

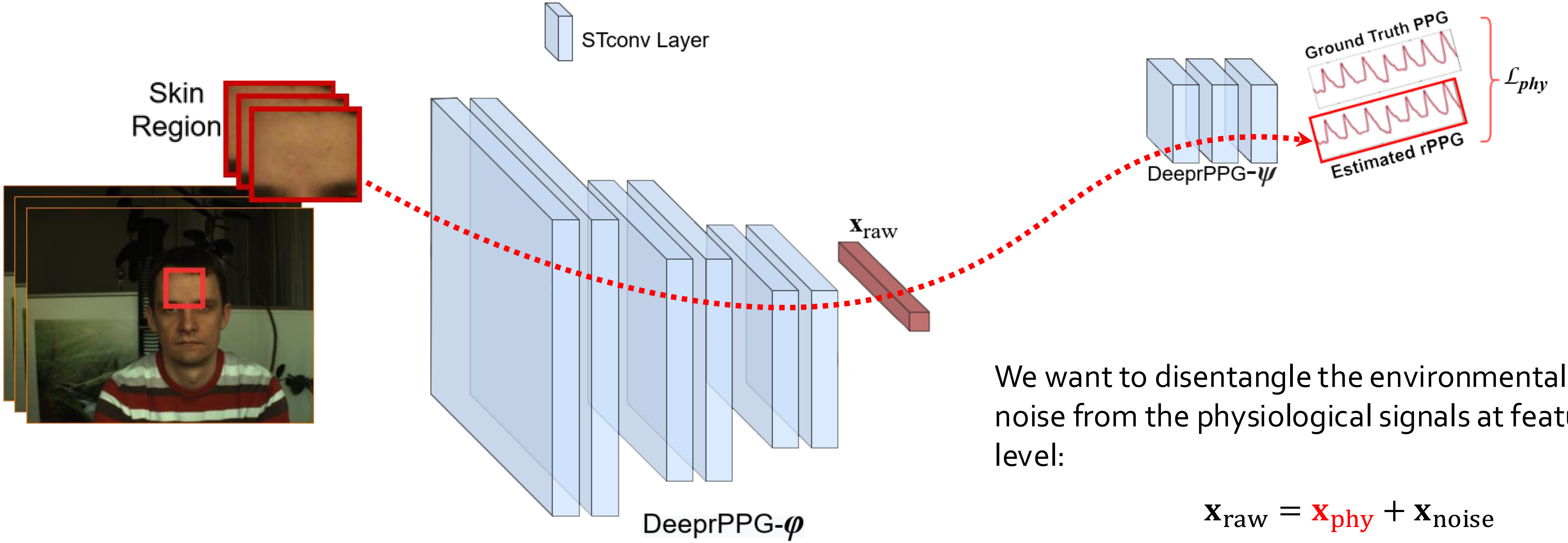
[TIP 24]

➢ DeeprPPG backbone



1. SQ. Liu and PC. Yuen, "A General Remote Photoplethysmography Estimator with Spatiotemporal Convolutional Network," FG, 2020.
2. S Q Liu and P C Yuen, "Robust Remote Photoplethysmography Estimation With Environmental Noise Disentanglement" TIP, In Press, 2024

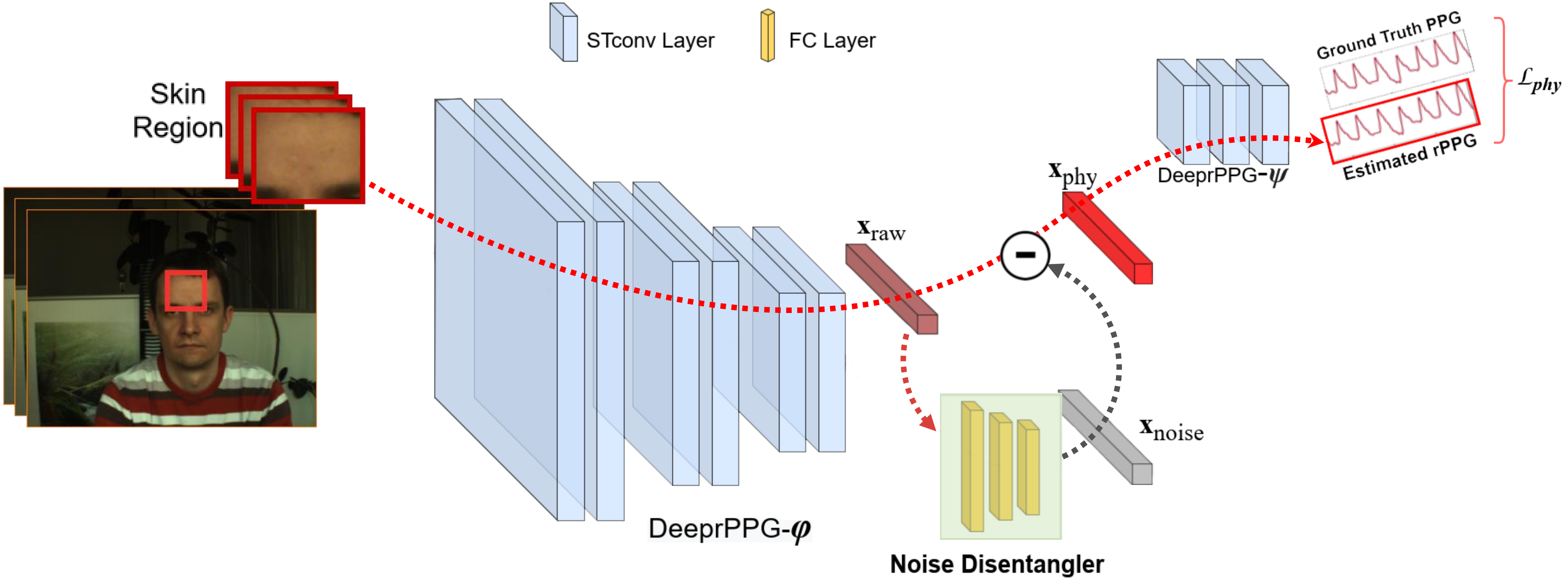
Noise Disentangled DeeprPPG Framework



We want to disentangle the environmental noise from the physiological signals at feature level:

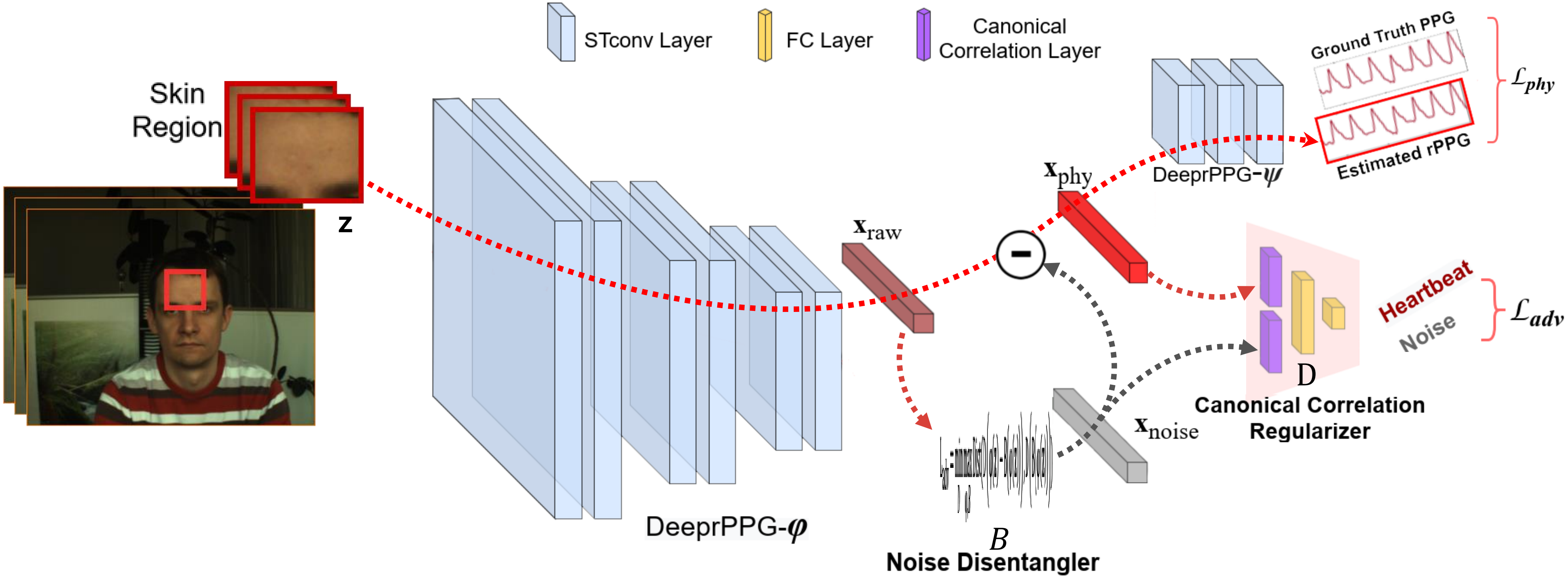
$$\mathbf{x}_{\mathrm{raw}} = \mathbf{x}_{\mathrm{phy}} + \mathbf{x}_{\mathrm{noise}}$$
$$\mathbf{x}_{\mathrm{phy}} = \mathbf{x}_{\mathrm{raw}} - \mathbf{x}_{\mathrm{noise}}$$

Noise Disentangled DeeprPPG Framework



Disentangle the environmental noise $\mathbf{x}_{noise}$ from the observed raw signal feature $\mathbf{x}$

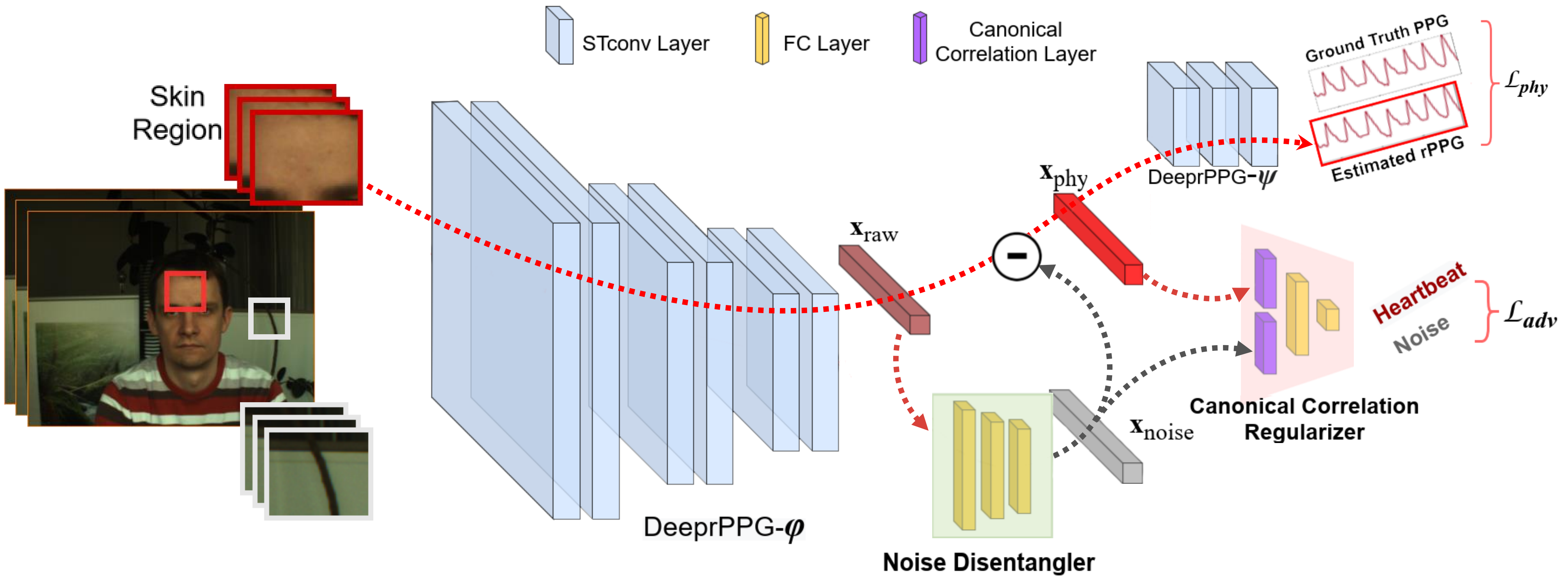
Noise Disentangled DeeprPPG Framework



Adversarial training: two-player min-max game:

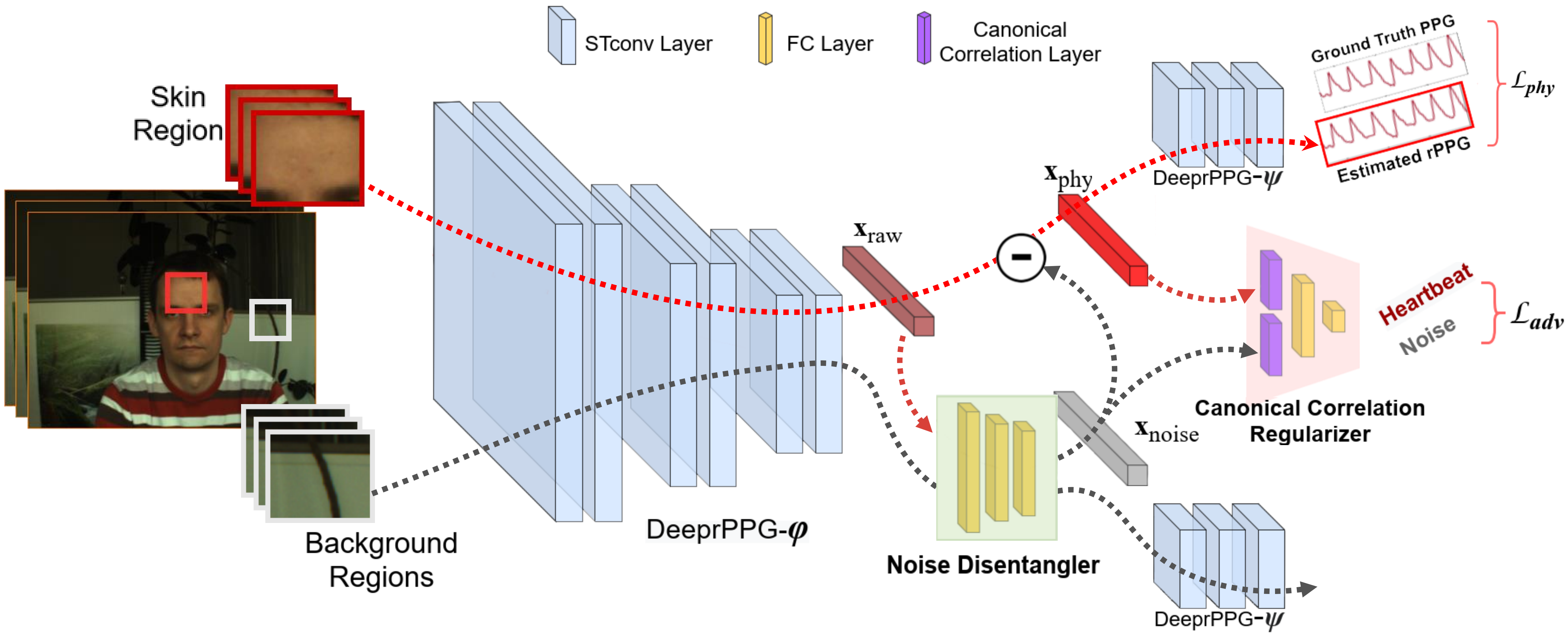
$$L_{adv} = \min_{D} \max_{\varphi,B} \mathrm{Dist}(D\left(\varphi(\mathbf{z}) - B(\varphi(\mathbf{z}))\right), D\left(B(\varphi(\mathbf{z}))\right))$$

Noise Disentangled DeeprPPG Framework

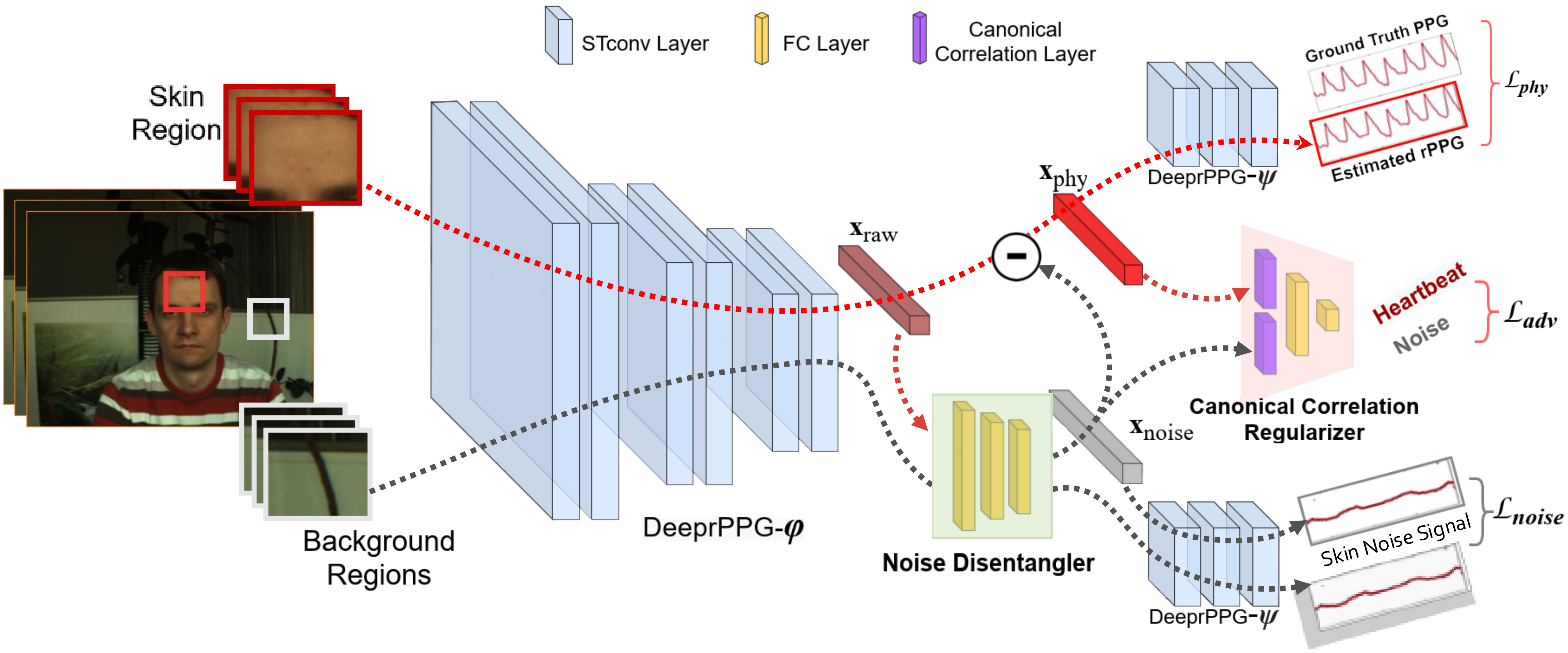


Use background regions as reference to train the Noise Disentangler *B(x)*

Noise Disentangled DeeprPPG Framework

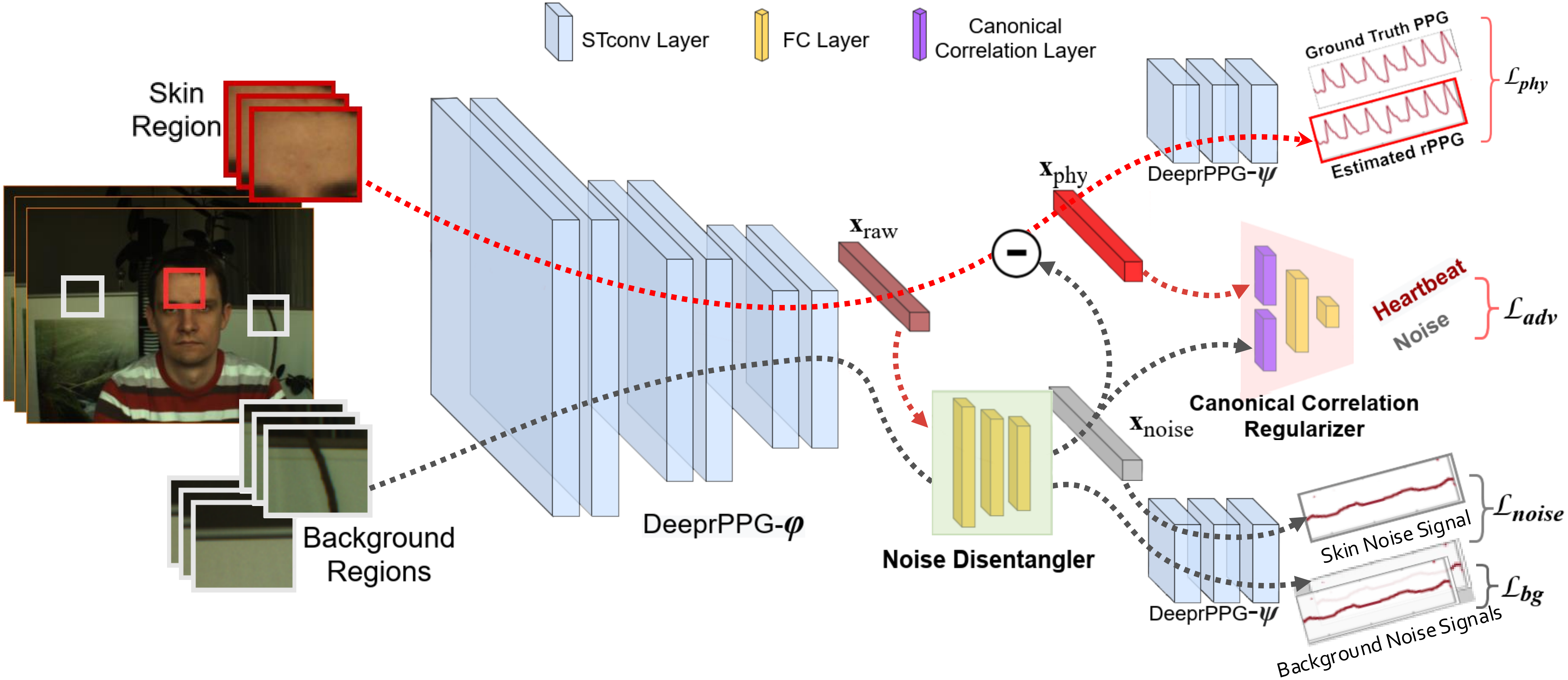


Noise Disentangled DeeprPPG Framework



The noise rPPG signals extracted from skin region and background region should be similar

Noise Disentangled DeeprPPG Framework



The noise rPPG signals extracted different background regions should be similar to each other

Experiments

- Datasets
  - PURE
    - 6 statuses: steady sitting, talking, slow head translation, fast head translation, small head rotation, and medium head rotation
    - Industrial camera, uncompressed video storage.
  - COHFACE
    - 2 lighting conditions: (a) studio light (b) natural side
    - Web camera, compressed video storage.
  - UBFC
    - Small head movement (Subjects are asked to play a time sensitive mathematical game)
    - Larger heart rate variation (80-120 bpm)
    - Logitech C920, uncompressed video storage
  - MMSE-HR
    - Spontaneous larger facial expressions and head motions
    - dark skin tones
  - MAHNOB-HCI
    - subjects are stimulated with emotion-eliciting clips and behave with corresponding facial expressions and head motions

Experiments

- Evaluation of average HR on PURE, COHFACE, UBFC
  - Performance metrics:
    - RMSE (root mean square error)
    - MAE (mean absolute error)
    - Pearson correlation R

Experiments

	MAE (bpm)	RMSE (bpm)	R
2SR	2.44	3.06	0.98
CHROM	2.07	2.50	0.99
LiCVPR	28.22	30.96	-0.38
HR-CNN	1.84	2.37	0.98
CVD	27.0	28.5	0.11
DeeprPPG	0.28	0.43	0.999
ND-DeeprPPG	0.18	0.41	0.999

Evaluation results on PURE

	MAE (bpm)	RMSE (bpm)	R
ICA [1]	8.16	13.99	0.36
MRICA	5.87	11.26	0.55
iBCG	8.20	14.09	0.39
CHROM[2]	8.44	13.74	0.34
LiCVPR	19.98	25.59	-0.44
2SR	20.98	25.84	-0.32
POS [4]	6.58	11.90	0.49
HR-CNN[3]	8.10	10.78	0.29
Two-stream CNN [12]	8.09	9.96	0.40
MOMBAT[11]	5.89	-	0.62
EEMD-MCCA[10]	2.08	4.80	0.91
CVD [9]	14.2	17.7	0.01
DeeprPPG	3.07	7.06	0.86
ND-DeeprPPG	0.64	1.89	0.98

Evaluation results on COHFACE

	MAE (bpm)	RMSE (bpm)	R
GREEN	4.47	11.6	0.842
ICA [1]	3.51	8.64	0.908
CHROM[2]	3.44	4.61	0.968
POS[4]	2.44	6.61	0.936
CK [5]	2.29	3.80	0.981
Frédéric [6]	5.45	8.64	-
Meta-rPPG[8]	5.97	7.42	0.53
HeartTrack[7]	2.41	3.37	0.983
CVD[9]	18.8	23.9	0.10
DeeprPPG	0.67	1.70	0.995
ND-DeeprPPG	0.31	0.98	0.999

Evaluation results on UBFC

[1] Poh et.al., "Non-contact, automated cardiac pulse measurements using video imaging and blind source separation.". Optical Society of America, 2010
[2] G. de Haan et al., "Robust pulse rate from chrominance-based rppg", TBE, 2013
[3] R. Spetlik et al., "Visual heart rate estimation with convolutional neural network", BMVC, 2018
[4] Wang et.al., "Algorithmic principles of remote PPG", TBE, 2015
[5] Song et.al., "New insights on super-high resolution for video-based heart rate estimation with a semi-blind source separation method", Computers in Biology and Medicine, 2020
[6] Frederic et.al., "3d convolutional neural networks for remote pulse rate measurement and mapping from facial video", Applied Sciences, 2019
[7] Olga et.al., "HeartTrack: Convolutional neural network for remote video-based heart rate monitoring", CVPRW, 2020
[8] Lee et.al., "Meta-rppg: Remote heart rate estimation using a transductive meta-learner", ECCV, 2020
[9] Niu et.al., "Video-based Remote Physiological Measurement via Cross-verified Feature Disentangling", ECCV,2020
[10] Song et.al., "Remote Photoplethysmography with an EEMD-MCCA Method Robust Against Spatially Uneven Illuminations", Sensors Journal, 2021
[11] P.Gupta et.al., "Mombat: Heart Rate Monitoring from Face Video using Pulse Modeling and Bayesian Tracking", Computers in biology and medicine, 2020
[12] Wang et.al., "Vision-Based Heart Rate Estimation via a Two-Stream CNN", ICIP, 2019

Qualitative comparisons

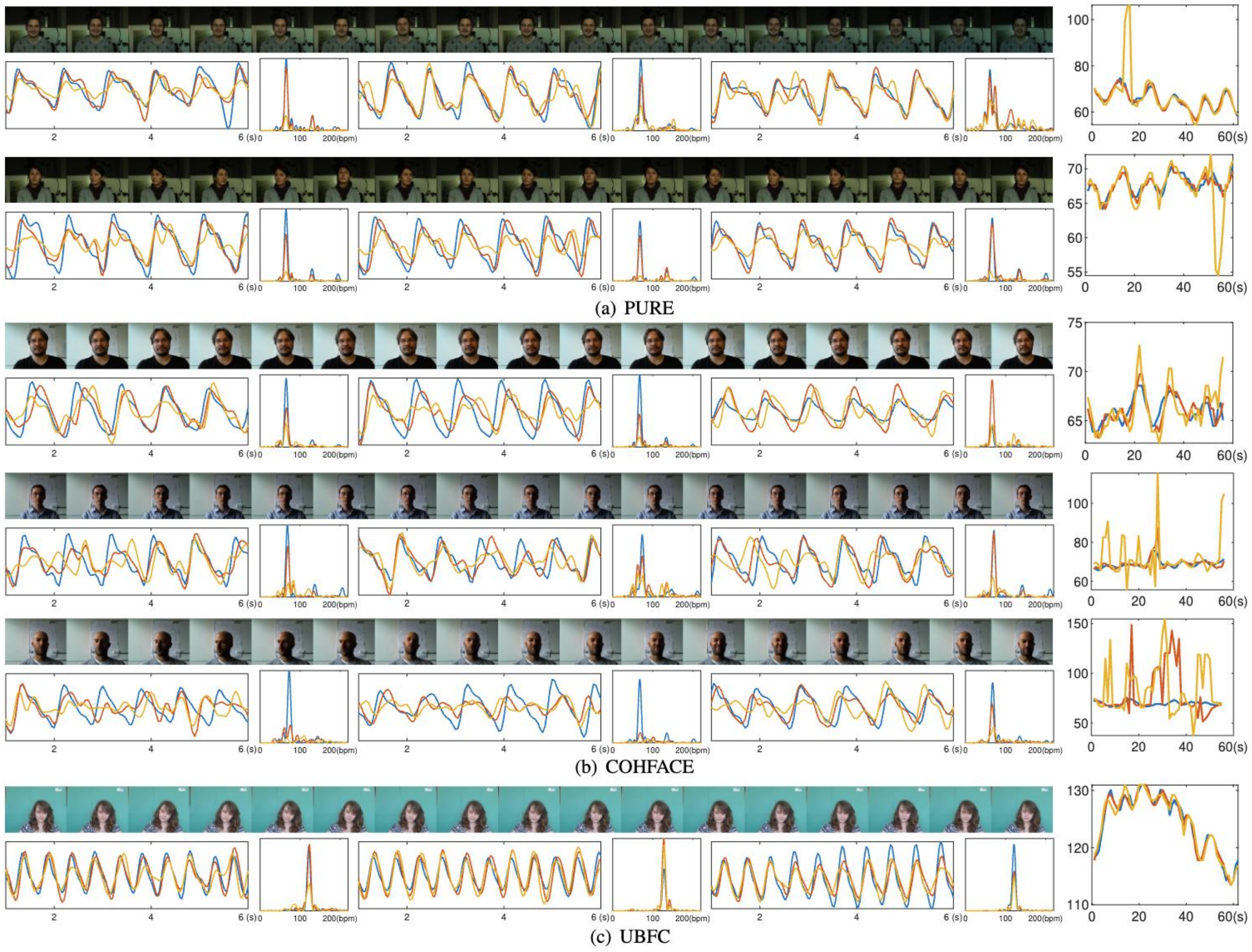


Fig. 9. Qualitative comparisons of DeeprPPG, ND-DeeprPPG, and ground-truth PPG on PURE, COHFACE, and UBFC dataset. rPPG signal slots and corresponding spectrums are visualized. Right column shows the HR trace where each point of HR (bmp) is obtained from 6 secs. signal. with $stride = 1sec$

Experiments

➢ Ablation Study

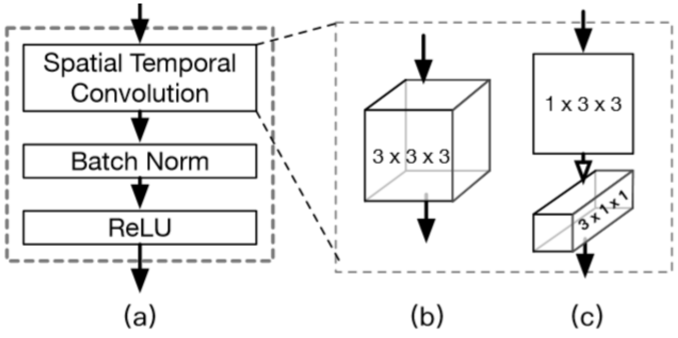
Components				COHFACE		
$\mathcal{L}_{phy}$	$\mathcal{L}_{adv}$	$\mathcal{L}_{noise}$	$\mathcal{L}_{bg}$	MAE (bpm)	RMSE (bpm)	R
✓				3.07	7.06	0.86
✓	✓			1.67	5.85	0.94
✓	✓	✓		1.11	3.90	0.94
✓	✓	✓	✓	0.64	1.89	0.98

Ablation study of the four components of ND-DeeprPPG on COHFACE dataset

➢ Different Backbone

	MAE (bpm)	RMSE (bpm)	R
DeeprPPG((2+1)D)	3.07	7.06	0.86
DeeprPPG(3D)	2.05	6.80	0.82
ND-DeeprPPG((2+1)D)	0.95	2.84	0.98
ND-DeeprPPG(3D)	0.84	2.86	0.97

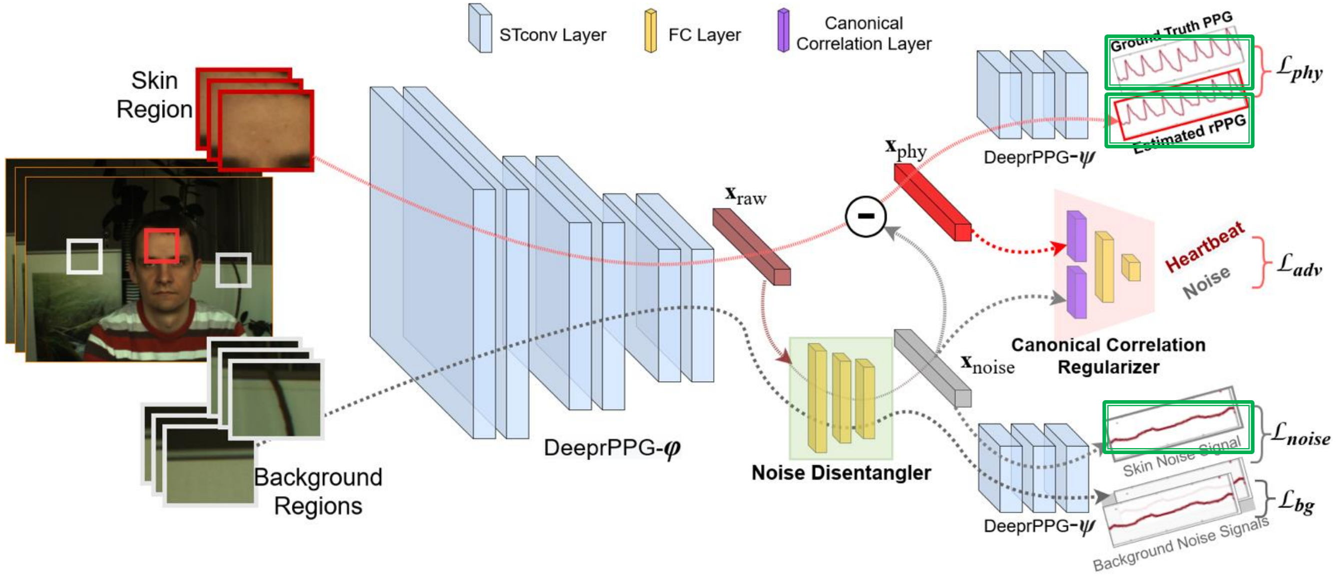
Evaluation the effectiveness of ND-DeeprPPG using different spatiotemporal convolutions on COHFACE dataset

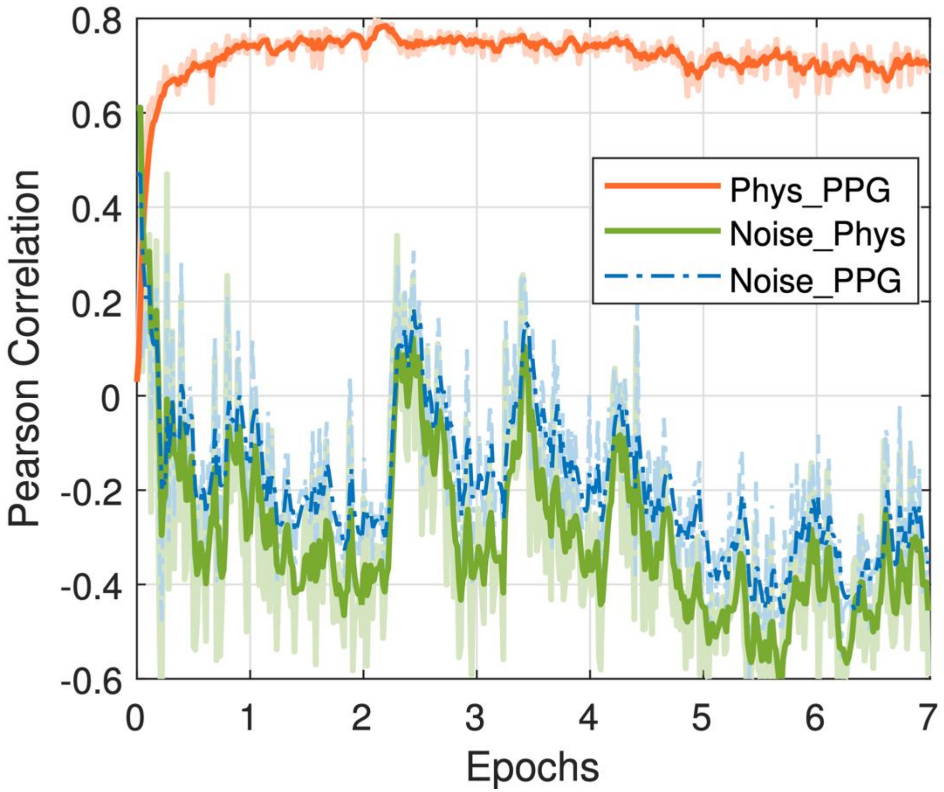


3D (2+1)D

Visualization of the Disentangling Process

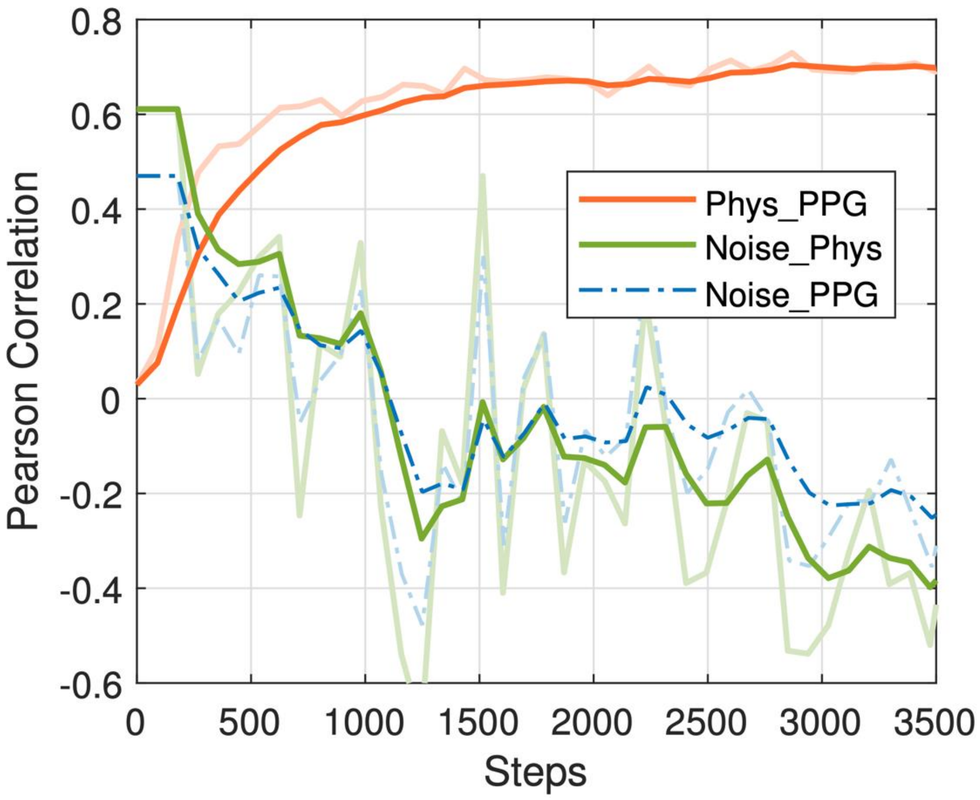
- Visualize the correlation of rPPG signals of $\boldsymbol{\psi}(\mathbf{x}_{\text{phy}})$ and $\boldsymbol{\psi}(\mathbf{x}_{\text{noise}})$:

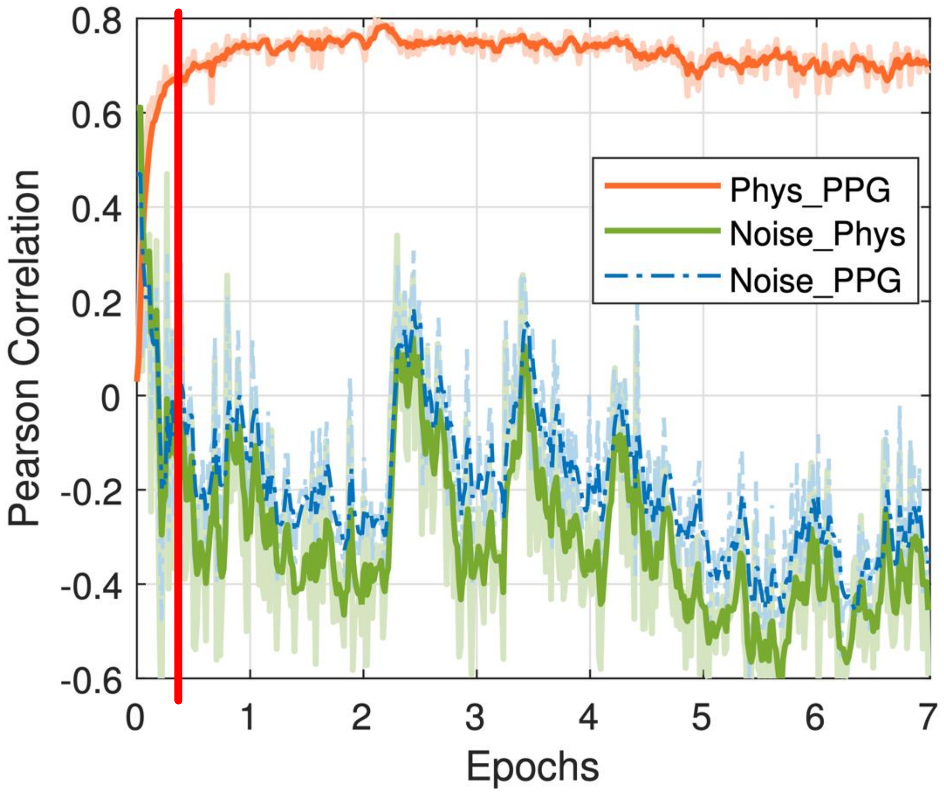




Visualization of the Disentangling Process

- Visualize the correlation of rPPG signals of $\boldsymbol{\psi}(\mathbf{x}_{\text{phy}})$ and $\boldsymbol{\psi}(\mathbf{x}_{\text{noise}})$:





Experiments

- Cross-dataset evaluation between PURE, COHFACE, and UBFC

	PURE→COHFACE			COHFACE→PURE		
	MAE (bpm)	RMSE (bpm)	R	MAE (bpm)	RMSE (bpm)	R
HR-CNN*	-	-	-	8.72	11.0	0.70
Two-stream CNN*	-	-	-	9.81	11.81	0.42
DeeprPPG	7.66	13.35	0.46	6.55	20.83	0.54
ND-DeeprPPG	3.04	7.10	0.78	0.29	0.62	0.997

	UBFC→COHFACE			COHFACE→UBFC		
	MAE (bpm)	RMSE (bpm)	R	MAE (bpm)	RMSE (bpm)	R
DeeprPPG	4.0	10.6	0.70	4.52	9.69	0.86
ND-DeeprPPG	2.39	6.65	0.84	0.49	1.12	0.998

	PURE→UBFC			UBFC→PURE		
	MAE↓	RMSE↓	R↑	MAE↓	RMSE↓	R↑
DAE	2.70	5.17	0.96	-	-	-
PulseGAN	2.09	4.42	0.97	-	-	-
Dual-GAN	0.74	1.02	0.997	-	-	-
DeepPhys	1.02	2.53	0.99	5.80	17.1	0.71
PhysNet	1.99	4.49	0.97	8.39	19.2	0.71
TS-CAN	0.99	2.41	0.99	5.75	16.3	0.74
DeeprPPG	2.30	4.15	0.97	0.29	0.63	0.997
ND-DeeprPPG	0.34	0.98	0.999	0.17	0.35	0.999

Experiments

- DeeprPPG for 3D mask face PAD
  - Extract local rPPG signals from forehead, cheek and low-face region
  - Apply LrPPG on the extracted 3 local rPPG signals

	3DMAD				HKBU-MARsV1+			
	HTER_dev(%)	HTER_test(%)	EER(%)	AUC(%)	HTER_dev(%)	HTER_test(%)	EER	AUC(%)
CHROM	10.82	11.65	11.68	94.69	8.83	10.10	9.81	96.48
DeeprPPG	16.46	17.15	17.04	89.68	37.67	38.83	38.28	64.90
ND-DeeprPPG	8.42	8.81	8.51	94.77	1.88	2.67	2.19	99.14

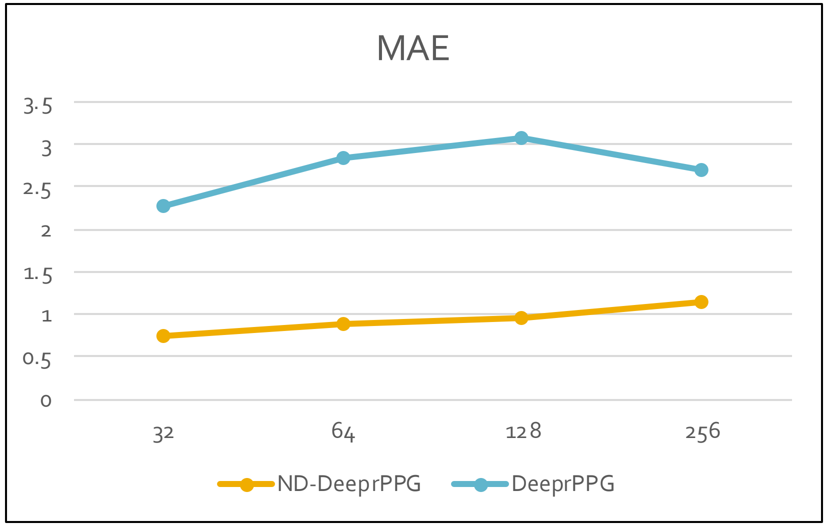
Intra-dataset evaluation on 3DMAD and HKBU MARs V1+ using LrPPG with different rPPG extractor

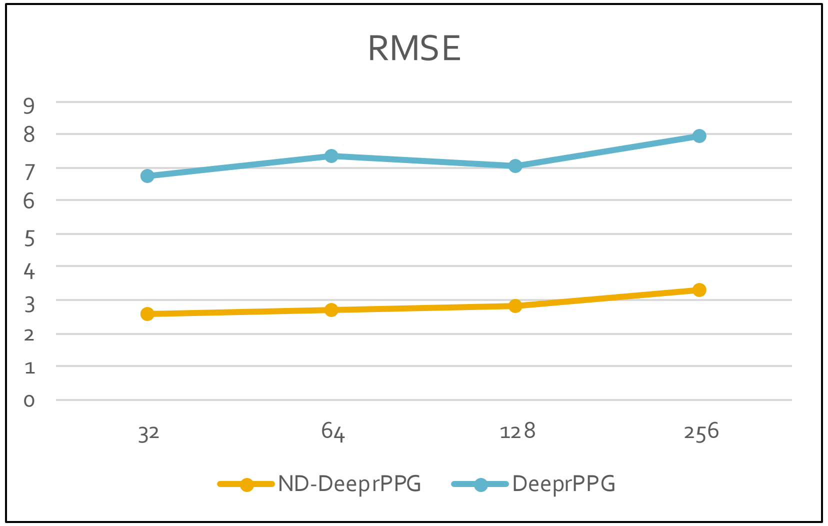
	HKBU-MARsV1+→3DMAD			3DMAD→HKBU-MARsV1+		
	HTER_test(%)	EER(%)	AUC(%)	HTER_test(%)	EER	AUC(%)
CHROM	12.47	12.47	93.97	11.23	10.90	94.88
DeeprPPG	48.25	51.03	50.03	40.83	47.42	54.39
ND-DeeprPPG	7.24	7.76	95.76	2.81	3.42	99.12

Cross-dataset evaluation between 3DMAD and HKBU MARs V1+ using LrPPG with different rPPG extractor

Experiments

- Performance comparison with various training video clip lengths: T=32, 64, 128, 256
  - Demonstrate the effectiveness of ND-DeeprPPG using different video clip length settings

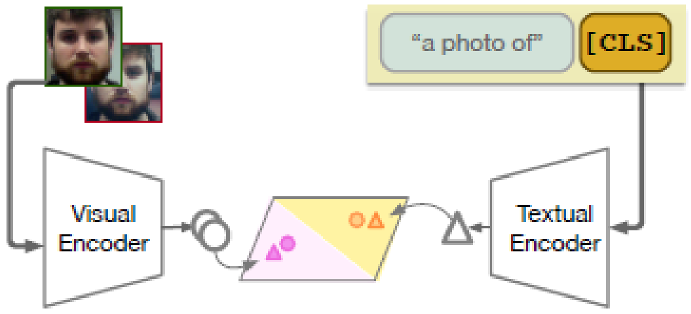




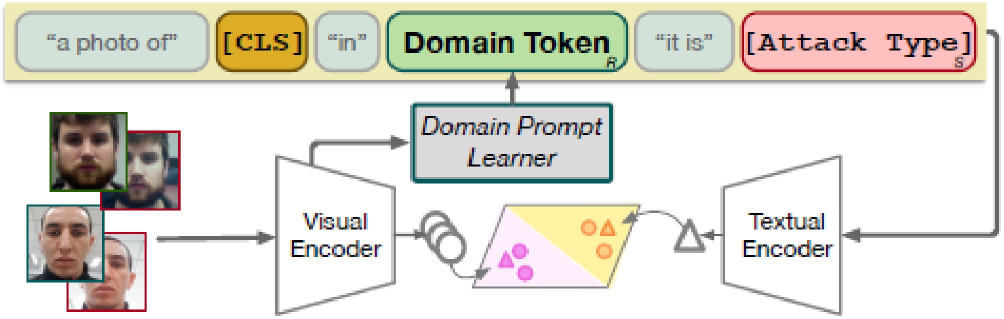
COHFACE:

Vision-Language Model for FAS

- Aligning the image representation with an ensemble of class text prompts.



(a) Intuitive solution of CLIP for Face Anti-Spoofing



(b) *Bottom-Up* Domain-Generalized Prompt Tuning for FAS (ours)

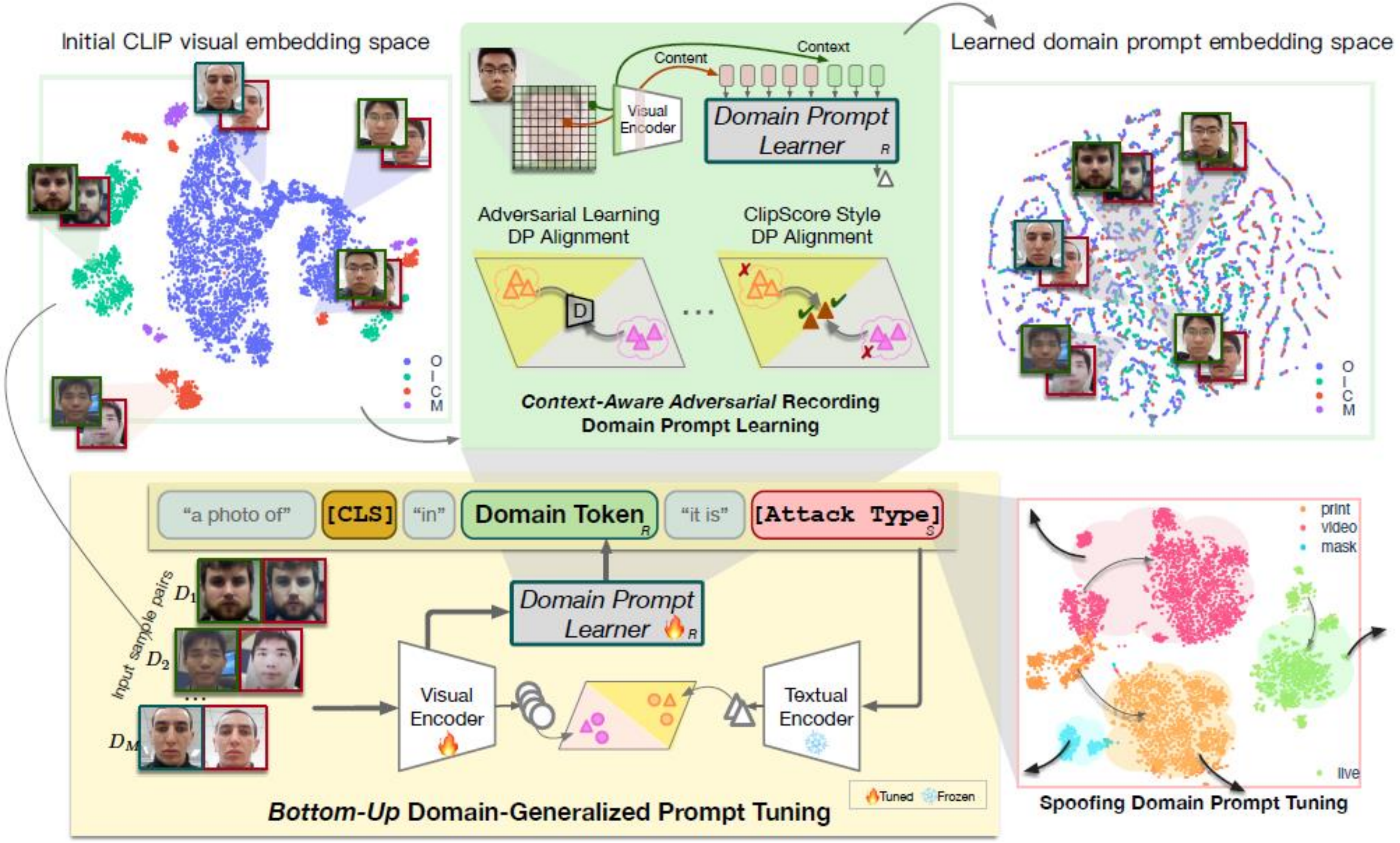
- Question: How to further improve the domain generalizability of Vision-Language-based FAS model?
- ✓ Solution: Introduce domain information with Prompt Tuning.

`{a photo of [CLS] captured under [Recording Setting]}`

- Issues: How to define domain?
  How to ensure domain information helps learn more generalized FAS feature?

Srivatsan et al. (2023). "Flip: Cross-domain face anti-spoofing with language guidance.", *ICCV,* 2023

Bottom-up Domain Prompt Tuning [ECCV2024]



BUDoPT: Covering different levels of domain variance to improve generalizability

S Q Liu, B Q Wang and P C Yuen, "Bottom-Up Domain Prompt Tuning for Generalized Face Anti-Spoofing", *ECCV* 2024.

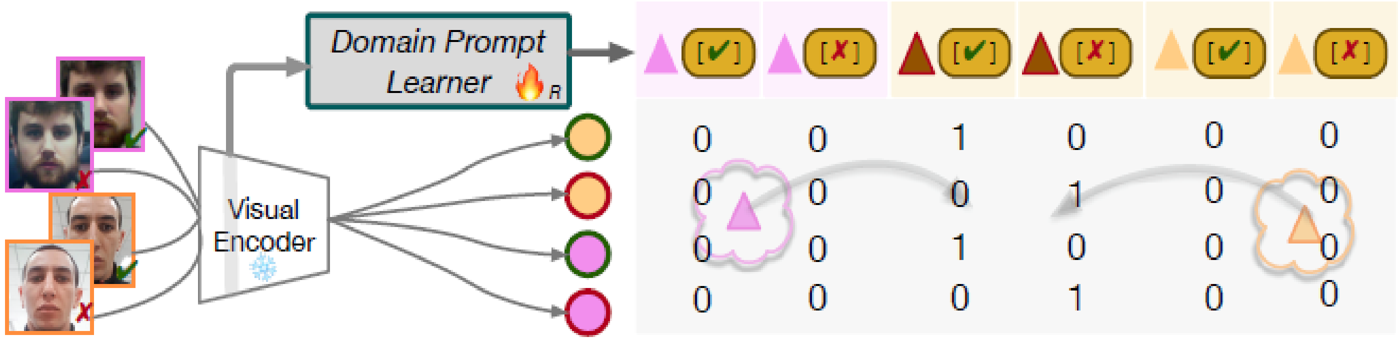
Bottom-up Domain Prompt Tuning [ECCV2024]

➢ Recording Domain Generalized Prompt Learning

✓ Employ Prompt Learner PL to learn the generalized domain prompt $\boldsymbol{p}_{GD}$ from vision domain features.

✓ **CLIP score-based DG learning strategy**: For image embedding, push away domain-specific prompt $\boldsymbol{p}_{D_m}$ (▲, ▲) and pull with generalized domain prompt $\boldsymbol{p}_{GD}$ (▲).

$$\mathcal{L}_{ce} = -\sum_{z=1}^{2\times(M+1)} y \log\left(\frac{\exp(\langle f_v(\boldsymbol{x}), f_t(\boldsymbol{p}_z)\rangle)}{\sum_{k=1}^{2\times(M+1)} \exp(\langle f_v(\boldsymbol{x}), f_t(\boldsymbol{p}_k)\rangle)}\right)$$



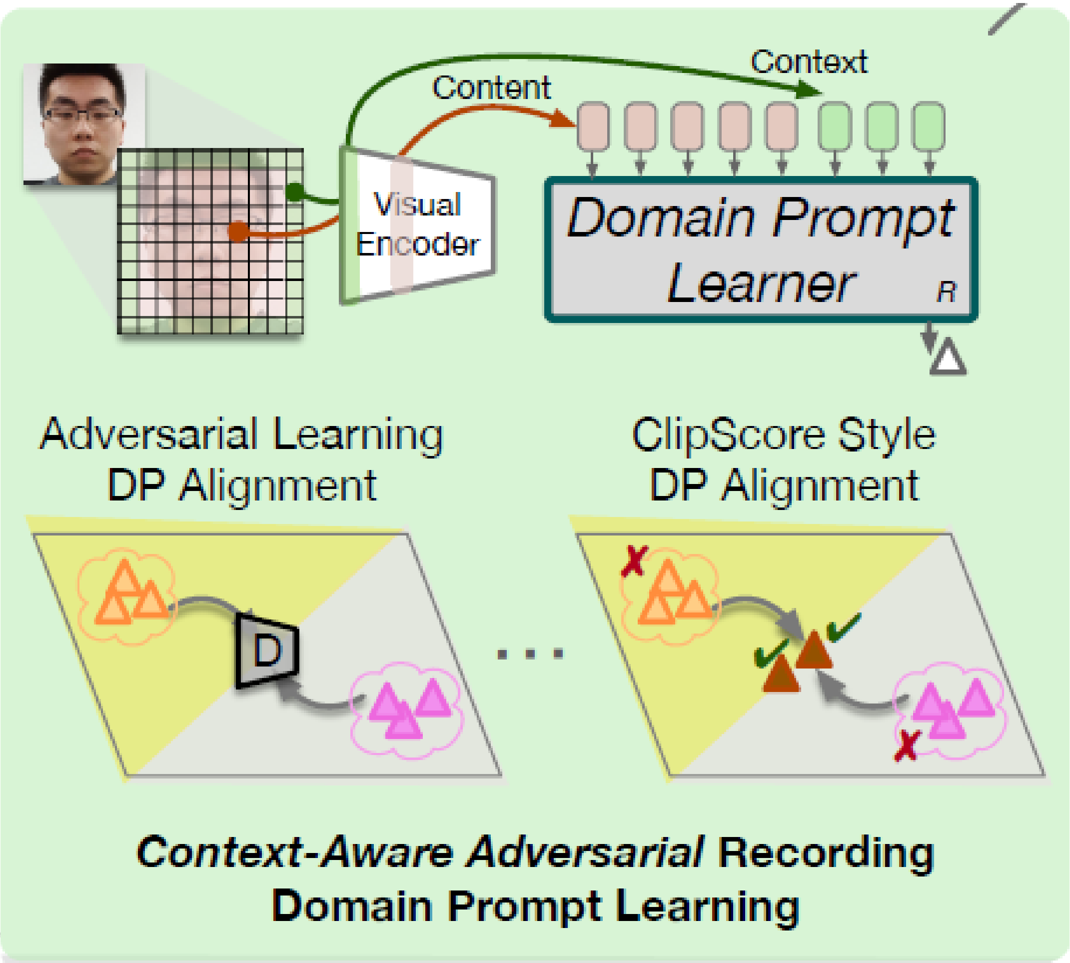
➢ Domain Partition Strategies: dataset, capturing devices, environments, sessions ···

S Q Liu, B Q Wang and P C Yuen, "Bottom-Up Domain Prompt Tuning for Generalized Face Anti-Spoofing", *ECCV* 2024.

Bottom-up Domain Prompt Tuning [ECCV2024]

- **Context-aware Domain Clue**
  - Highlight the shallow layer's context patches with purer domain information.

- **Adversarial Domain Prompt Learning**
  - Minimize the domain-specific trace in prompts



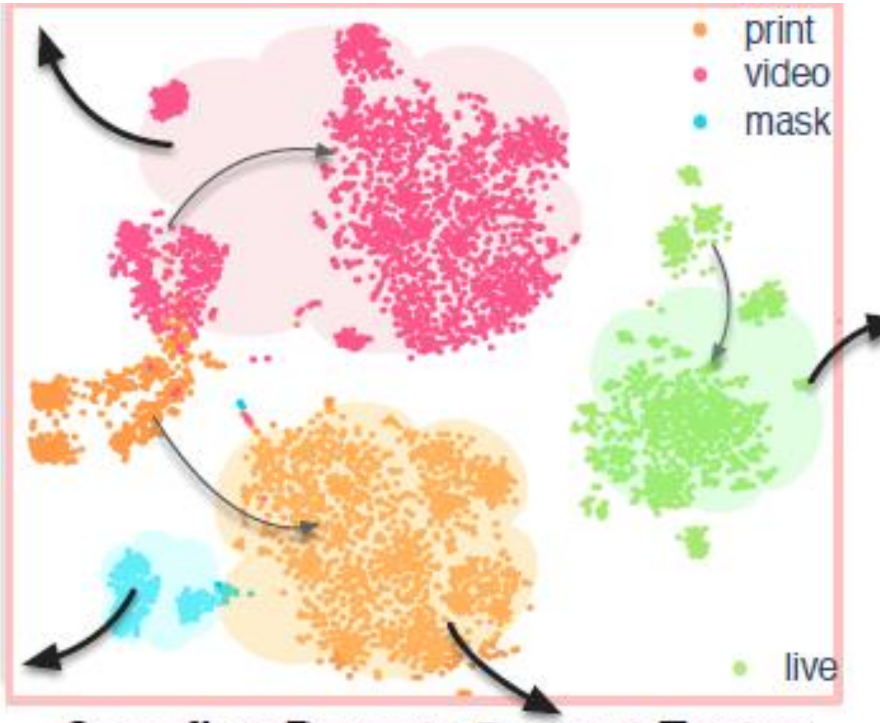
S Q Liu, B Q Wang and P C Yuen, "Bottom-Up Domain Prompt Tuning for Generalized Face Anti-Spoofing", *ECCV* 2024.

Bottom-up Domain Prompt Tuning [ECCV2024]

➢ Spoofing Domain Discriminative Prompt Tuning

- Attack types can be clearly described with high-level textual prompts
- Regularize the visual embedding with finer grained attack types text

$$\mathcal{L}_t = -\sum_{o=1}^{N*M} o \log\left(\frac{\exp(\langle f_v(\boldsymbol{x}), f_t(\boldsymbol{pa}_o)\rangle)}{\sum_{w=1}^{N*M} \exp(\langle f_v(\boldsymbol{x}), f_t(\boldsymbol{pa}_w)\rangle)}\right)$$



Spoofing Domain Prompt Tuning

S Q Liu, B Q Wang and P C Yuen, "Bottom-Up Domain Prompt Tuning for Generalized Face Anti-Spoofing", *ECCV* 2024.

Experimental Results

Table 1: Evaluation on **I**,**M**,**C**,**O** with coarse domain partition. † employs CelebA-Spoof as supplementary training dataset.

Methods	OCI→ M			OMI→ C			OCM→ I			ICM→ O			Avg.
	HTER	AUC	TPR@ FPR=1%	HTER	AUC	TPR@ FPR=1%	HTER	AUC	TPR@ FPR=1%	HTER	AUC	TPR@ FPR=1%	
MADDG	17.69	88.06	-	24.50	84.51	-	22.19	84.99	-	27.98	80.02	-	23.09
MDDR	17.02	90.10	-	19.68	87.43	-	20.87	86.72	-	25.025	81.47	-	20.64
NAS-FAS	16.85	90.42	-	15.21	92.64	-	11.63	96.98	-	13.16	94.18	-	14.21
RFMeta	13.89	93.98	-	20.27	88.16	-	17.30	90.48	-	16.45	91.16	-	16.97
D^2AM	12.70	95.6	-	20.98	85.58	-	15.43	91.22	-	15.27	90.87	-	16.09
DRDG	12.43	95.81	-	19.05	88.79	-	15.56	91.79	-	15.63	91.75	-	15.66
Self-DA	15.40	91.80	-	24.50	84.40	-	15.60	90.10	-	23.10	84.30	-	19.65
ANRL	10.83	96.75	-	17.85	89.26	-	16.03	91.04	-	15.67	91.90	-	15.09
FGHV	9.17	96.92	-	12.47	93.47	-	16.29	90.11	-	13.58	93.55	-	12.87
SSDG-R	7.38	97.17	-	10.44	95.94	-	11.71	96.59	-	15.61	91.54	-	11.28
SSAB-R	6.67	98.75	-	10.00	96.67	-	8.88	96.79	-	13.72	93.65	-	9.80
PatchNet	7.10	98.46	-	11.33	94.58	-	13.40	95.67	-	11.82	95.07	-	10.90
GDA	9.20	98.00	-	12.20	93.00	-	10.00	96.00	-	14.40	92.60	-	11.45
LDCformer	6.43	98.39	-	8.11	96.67	-	8.57	97.09	-	11.17	95.58	-	8.57
TTDG-V	4.16	98.48	-	7.59	98.18	-	9.62	98.18	-	10	96.15	-	7.84
BUDoPT(ours)	**0.95**	**99.70**	**87.12**	**2.85**	**98.03**	**55.00**	**4.4**	**98.54**	**86.46**	**2.26**	**98.78**	**55.25**	**2.62**
ViT†	1.58	99.68	96.67	5.70	98.97	88.57	9.25	97.15	51.54	7.47	98.42	69.30	6.00
FLIP-V†	3.79	99.31	87.99	1.27	99.75	95.85	4.71	98.80	75.84	4.15	98.76	66.47	3.48
FLIP-IT†	5.27	98.42	79.33	0.44	99.98	99.86	2.94	99.42	84.62	3.61	99.15	84.76	3.06
FLIP-MCL†	4.95	98.11	74.67	0.54	**99.98**	**100.00**	4.25	99.07	84.62	2.31	**99.63**	92.28	3.01
BUDoPT-CoOp†	1.98	99.54	80.63	1.93	99.58	88.06	3.18	99.3	90.4	3.33	98.91	89.29	2.61
BUDoPT-CoCoOp†	1.51	99.43	75.55	1.55	99.56	84.74	3.53	99.2	92.53	3.51	98.98	87.98	2.53
BUDoPT(ours)†	**0.40**	**99.99**	**99.67**	**0.26**	99.96	99.92	**1.38**	**99.69**	**98.1**	**1.60**	99.51	**97.18**	**0.91**

- **Datasets**: MSU-MFSD(**M**),
  CASIA-FASD(**C**),
  Idiap Replay attack(**I**),
  OULU-NPU(**O**),
  CelebA-Spoof

- **Domain gaps**:
  - Recording devices
  - Attack types
  - Illumination
  - Environment
  - ......

Experimental Results

One-to-one cross-dataset evaluation

Table 2: One-to-one cross-dataset evaluation on **I,M,C,O** with fine-grained domain partition. Comparing with domain adaptation models (Unseen=No).

Methods	Unseen	C → I	C → M	C → O	I → C	I → M	I → O	M → C	M → I	M → O	O → C	O → I	O → M	Avg.
ADDA	No	41.8	36.6	-	49.8	35.1	-	39.0	35.2	-	-	-	-	39.6
DRCN	No	44.4	27.4	-	48.9	42.0	-	28.9	36.8	-	-	-	-	38.1
DupGAN	No	42.1	33.4	-	46.5	36.2	-	27.1	35.4	-	-	-	-	36.8
KSA	No	39.3	15.1	-	12.3	33.3	-	9.1	34.9	-	-	-	-	24.0
DR-UDA	No	15.6	9.0	28.7	34.2	29.0	**3.5**	16.8	3.0	30.2	19.5	25.4	27.4	23.1
ADA	No	17.5	9.3	29.1	41.5	30.5	39.6	17.7	5.1	31.2	19.8	26.8	31.5	25.0
USDAN-Un	No	16.0	9.2	-	30.2	25.8	-	13.3	3.4	-	-	-	-	16.3
GDA	No	15.10	5.8	-	29.7	20.8	-	12.2	2.5	-	-	-	-	14.4
CDFTN-L	No	**1.7**	8.1	29.9	11.9	9.6	29.9	8.8	**1.3**	25.6	19.1	5.8	6.3	13.2
BUDoPT(ours)	Yes	9.65	**4.37**	**6.76**	**8.26**	**5.64**	7.73	**6.07**	2.9	**6.16**	**4.81**	**2.97**	**2.94**	**5.69**
FLIP-V†	Yes	15.08	13.73	12.34	4.30	9.68	7.87	0.56	3.96	4.79	2.09	5.01	6.00	7.12
FLIP-IT†	Yes	12.33	15.18	7.98	1.12	8.37	6.98	0.19	5.21	4.96	**0.16**	4.27	5.63	6.03
FLIP-MCL†	Yes	10.57	7.15	**3.91**	0.68	7.22	4.22	0.19	5.88	3.95	0.19	5.69	8.40	4.84
BUDoPT(ours)†	Yes	**4.33**	**2.62**	4.98	**0.48**	**1.83**	**4.14**	**0**	**2.45**	**1.87**	0.44	**2.53**	**1.43**	**2.46**

Unseen attack type evaluation

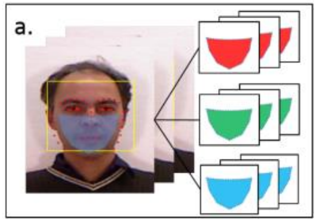
Table 3: Evaluation on WMCA with unseen 3D mask attack.

Methods	IMCO → WMCA	
	HTER	AUC
DiVT-M	22.36	86.82
DGUA-FAS	20.62	88.07
Method	C → WMCA	
BUDoPT(ours)	**9.20**	**95.93**

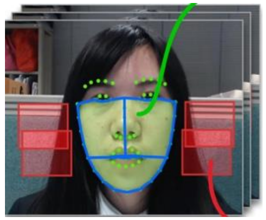
Implementation of our rPPG-based Face PAD Method

My Journey: Face PAD

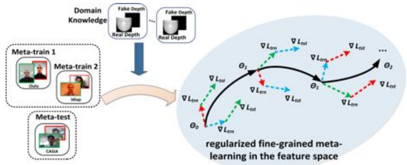
Global rPPG (ICPR16)



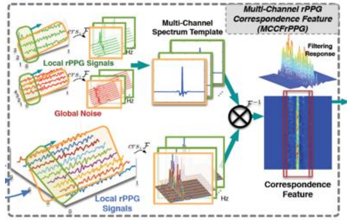
Local rPPG with background information for noise removal (ECCV18)



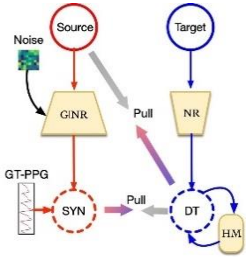
Meta-learning (AAAI 20)



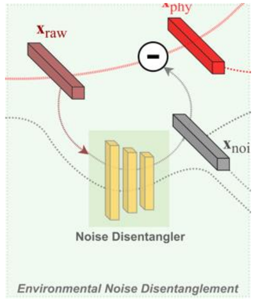
Multi-Channel rPPG (TIFS2021)



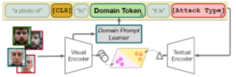
Robust and generalized rPPG estimation (CVPR23)



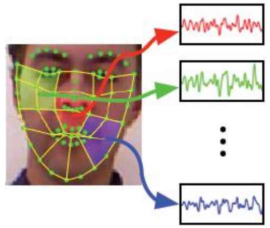
ND-DeeprPPG (TIFS2024)



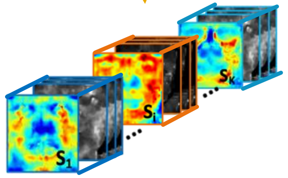
2016 2017 2018 2019 2020 2021 2022 2023 2024



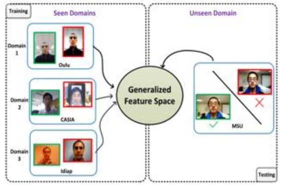
Local rPPG (ECCV 16)



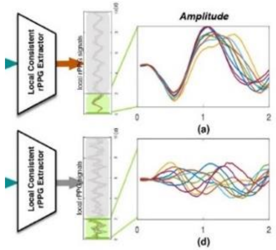
Deep Dynamic Textures (TIFS 19)



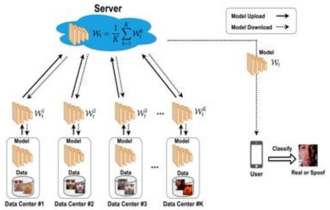
Deep domain generalization (CVPR 19)



Fast rPPG estimation using Temporal Similarity (TIFS2022)



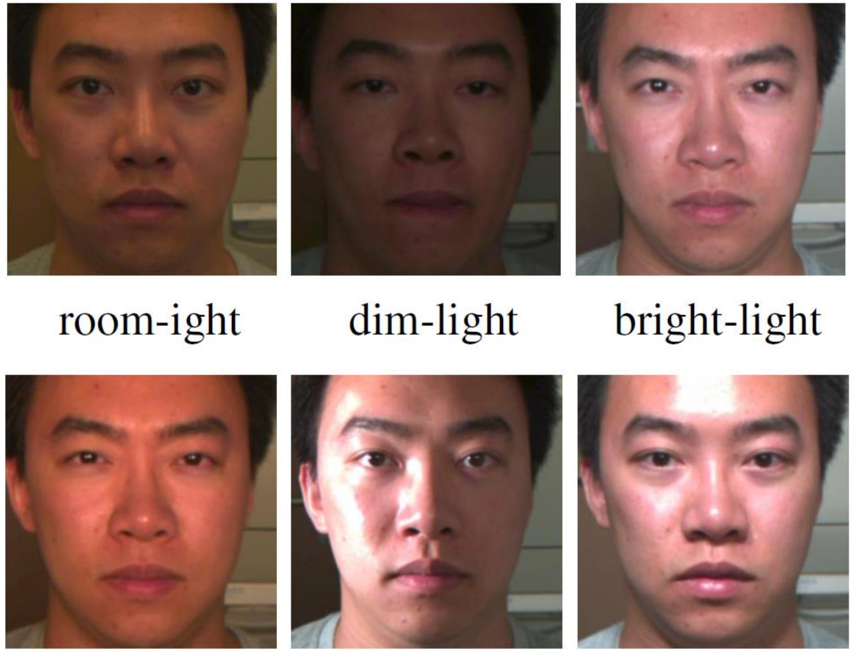
Federated Generalisation face PAD [TNNLS 22] and test-time adaptation [FG21]



Bottom-Up Domain Prompt Tuning for Generalized Face PAD (ECCV 24)

Our dataset: HKBU-MARs

- http://rds.comp.hkbu.edu.hk/mars



room-ight dim-light bright-light

warm-light sidelight top-light

(a) ThatsMyFace

(b) REAL-f

Conclusions

- PAD is an important and un-solved issue in biometric systems
- Rapid progress in the past 5 years, still a lot issues needed to be solved
- Face PAD has high academic and commercial values
- rPPG offers very good generalisation ability for face PAD, in particular 3D mask attack. Performance can be further improved by integrating other PAD methods
- rPPG is also a powerful tool in healthcare domain

Special Thanks …

Collaborators:
- Prof. GY Zhao, The University of Oulu
- Prof. V Patel, Johns Hopkins University

Current/Former PhD Students:
- Dr. Siqi Liu, CUHK Shenzhen
- Dr. Rui Shao, HIT Shenzhen
- Dr. X Lan, Pengcheng Lab Shenzhen
- Mr. J Du

Funding:
- Hong Kong Research Grant Council

Prof. GY Zhao

Prof. V Patel

Dr. Siqi Liu

Dr. Rui Shao

Dr. X Lan

Mr. J Du

Welcome Submission to TBIOM

Browse Journals & Magazines > IEEE Transactions on Biometric... ?

IEEE Transactions on Biometrics, Behavior, and Identity Science

IEEE Biometrics Council

Submit Manuscript | Add Title To My Alerts | Add to My Favorites

Home | Popular | Early Access | Current Issue | All Issues | About Journal

3.9
CiteScore
Powered by Scopus

The articles in this journal are peer reviewed in accordance with the requirements set forth in the IEEE PSPB Operations Manual (sections 8.2.1.C & 8.2.2.A). Each published article was reviewed by a minimum of two independent reviewers using a single-blind peer review process, where the identities of the reviewers are not known to the authors, but the reviewers know the identities of the authors. Articles will be screened for plagiarism before acceptance.

Corresponding authors from low-income countries are eligible for waived or reduced open access APCs. Read Full Aims & Scope

Author Resources

Submission Guidelines

Submit Manuscript

Popular Articles | Latest Published Articles

Fun Selfie Filters In Face Recognition: Impact Assessment And Removal
Cristian Botezatu; Mathias Ibsen; Christian Rathgeb; Christoph Busch

Feedback

Thank you!