Face Presentation Attack Detection

aka Face Anti-Spoofing

Prof Sébastien Marcel (www.idiap.ch/∼marcel)

January 12, 2025

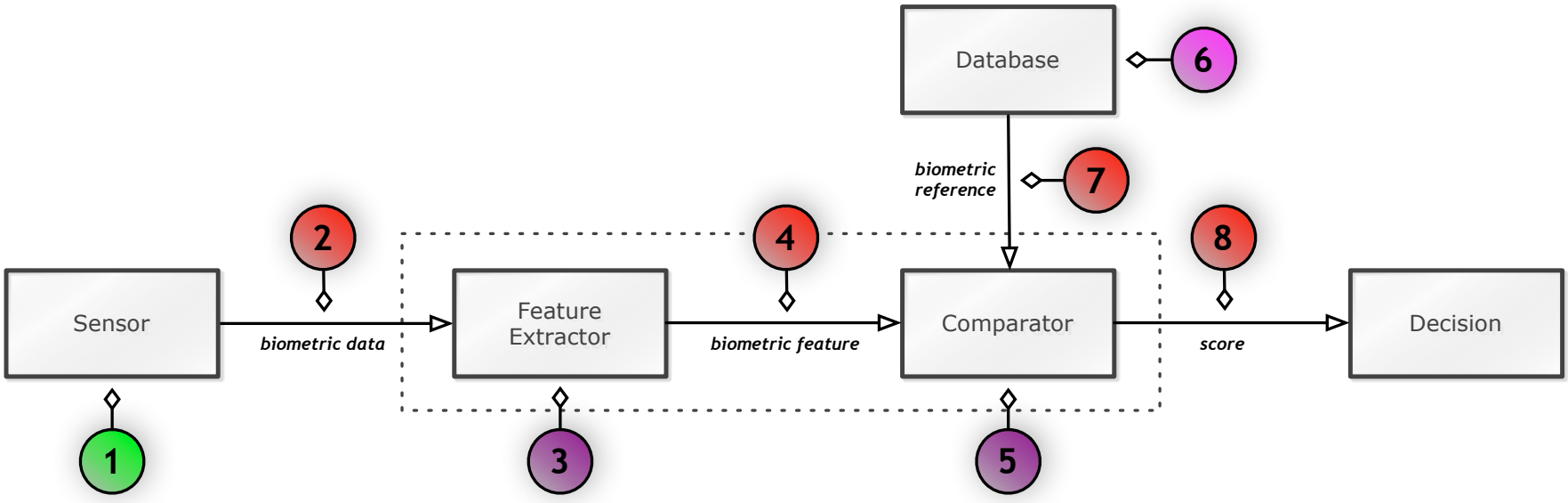
Idiap Research Institute

idiap RESEARCH INSTITUTE

Outline

Outline

Presentation Attacks



In this talk we will focus:

- on direct attacks to the sensor (1), referred to as **Presentation Attacks** (PA),
- on biometric systems using face (aka face recognition),
- on methods to detect face PAs i.e. face **Presentation Attack Detection** (PAD).

Definitions

Presentation Attack (PA)

- An attempt to **fool** the biometric recognition system by presenting **fake** biometric data to the sensor, e.g.,
  - A **replica** of an enrolled user's biometric features (if the goal is to **impersonate** that user), or
  - **Generic** biometric features (if the goal is to **avoid recognition**)
- PAs are also commonly called **spoofing attacks**, and the fake biometric data is referred to as a **spoof**

Presentation Attack Detection (PAD)

- The **determination** of a PA (i.e., "the presented biometric data is/is not a spoof")
- Also commonly referred to as **anti-spoofing**

Definitions

Presentation Attack Instrument (PAI)

- The **biometric characteristic** or **object** used to launch a PA
- *Examples:* Face mask, gummy fingerprint, dead body parts, etc.

Bona Fide Presentation

- **Normal** (intended) interaction of the subject with the biometric system's sensor
- Basically, anything which is **not** a PA

Note: See *Biometric presentation attack detection – part 1*, ISO/IEC 30107-1:2016 (2016) for formal (standardised) definitions.

Importance

PAs pose a **major threat** to biometric recognition systems:

- because the attack is **external** to the system (i.e., at the sensor), so the attacker does **not** need to have any **knowledge** about the **internal workings** of the system,
- PAs can be launched by basically **anyone**, often using very **basic tools**.

Growing field of research[1]:

- novel methods for innovative PAIs and PAs,
- novel techniques and algorithms for PAD (e.g. sensors, signal processing, machine learning, generalisation to unseen attacks),
- not only for face biometrics but also fingerprint, iris, voice, vein, ...

[1] S. Marcel *et al.*, "Handbook of Biometric Anti-Spoofing", Third Edition, *Springer*, 2023 (10.1007/978-981-19-5288-3)

Outline

Face PA in the reality

Locker unlock (2019)

A group of primary school children in China showed that lockers secured by face recognition technology could be spoofed using a photograph of the locker owner's face[2]

[2] http://www.sixthtone.com/news/1004698/facial-recognition-smart-lockers-hacked-by-fourth-graders

Face PA in the reality

Robbery (2010)

Conrad Zdzierak used a silicone face mask to pass himself off as a black character "SPFX The Player" during bank robberies[3]

[3] http://www.telegraph.co.uk/news/worldnews/northamerica/usa/8193185/US-criminals-using-film-quality-masks-during-bank-robberies.html

Face PA in the reality

Immigration (2011)

A young Asian man disguised himself as an old Caucasian man using a silicone face mask, boarded a plane in Hong Kong, then removed the disguise mid-flight and asked for refugee status upon arriving in Canada[4]

[4] http://www.dailymail.co.uk/news/article-1326885/Man-boards-plane-disguised-old-man-arrested-arrival-Canada.html

Face PA in the reality

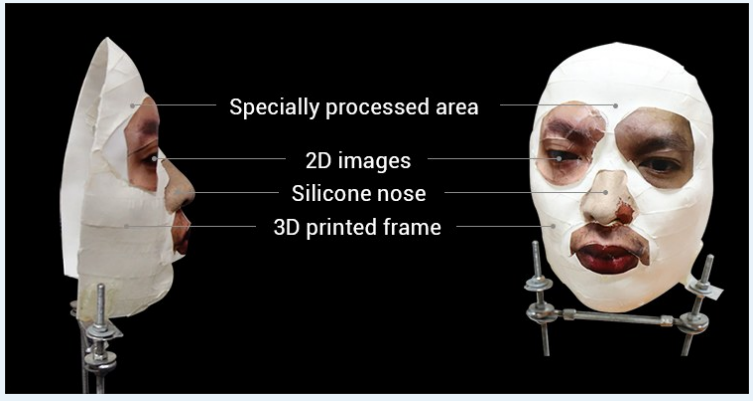
Smartphone unlock (2011)

The Face Unlock feature on Galaxy Nexus, running Android 4.0, was spoofed by a face photograph[5]

[5] http://www.geek.com/android/android-face-lock-feature-spoofed-by-photograph-1440953

Face PA in the reality

Smartphone unlock (2017)



iPhone X's Face ID was spoofed by a specially crafted face mask[6], despite claims that it is robust to mask attacks

[6] https://www.youtube.com/watch?v=i4YQRLQVixM

Outline

Face PAI

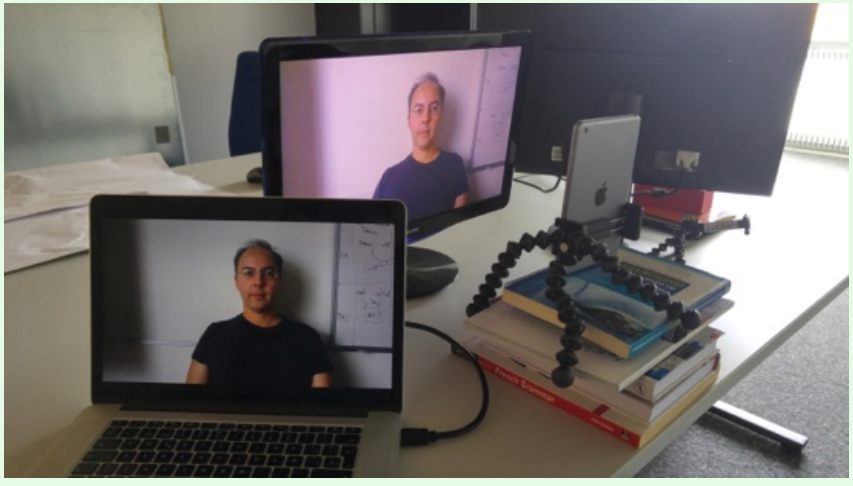
Printed face image[7]

1. Print an image of the target's face
2. Present the face image to the face recognition system

[7] A. Anjos and S. Marcel, "Counter-Measures to Photo Attacks in Face Recognition: a public database and a baseline", *IEEE IJCB* 2011 (10.1109/IJCB.2011.6117503)

Face PAI

Digital face image or video[8]

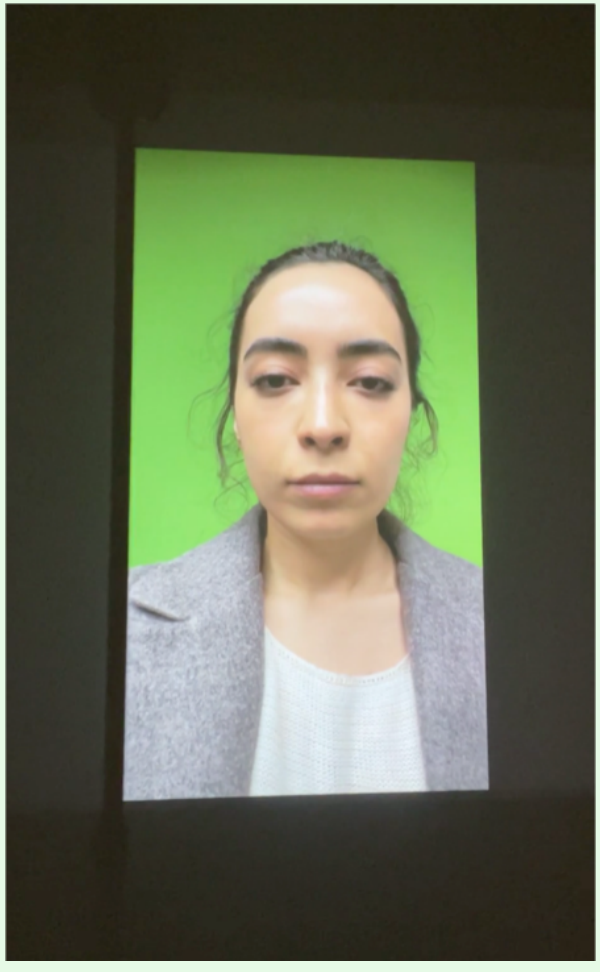


1. Capture a digital image or record a video of the target's face (e.g., using a smartphone or tablet)
2. Present the image or video (e.g. **deepfake**) to the face recognition system

[8] I. Chingovska, A. Anjos and S. Marcel, "Biometrics Evaluation Under Spoofing Attacks", *IEEE TIFS* 2014 (10.1109/TIFS.2014.2349158)

Face PAI

Video projection of an image or a video on any surface[9]



[9] N. Ramoly and al., "A Novel and Responsible Dataset for Face Presentation Attack Detection on Mobile Devices", *IEEE IJCB* 2024

Face PAI

Printed face on a t-shirt[10]

[10] M. Ibsen and al., "Attacking Face Recognition with T-shirts: Database, Vulnerability Assessment and Detection", *IEEE Access* 2024 (10.1109/ACCESS.2023.0322000)

Quizz time !

Which images are Bona Fide and which are PA?

Quizz time !

Which images are Bona Fide and which are PA?

All are PAs!

- *Left:* Printed images
- *Middle:* iPhone (digital) images
- *Right:* iPad (digital) images

Face PAI

Hard (resin composite) face mask[11]

1. 3D print a model of the target's face (plain or eye holes)
2. Present the corresponding hard face mask (made of a resin composite) to the face recognition system

[11] N. Erdogmus and S. Marcel, "Spoofing Face Recognition with 3D Masks", *IEEE TIFS*, 9(7), pp. 1084–1097, 2014 (10.1109/TIFS.2014.2322255)

Face PAI

Hyper-realistic face masks [12]

Same as previous example but hyper-realistic plastic masks from HiRes pictures

[12] K. Kotwal *et al.*, "Domain-Specific Adaptation of CNN for Detecting Face Presentation Attacks in NIR", *TBIOM* 2022 (10.1109/TBIOM.2022.3143569)

Face PAI

Silicone face mask – generic

A generic silicone face mask could be used to obfuscate an attacker's identity, but it does not correspond to any specific target face

Face PAI

Silicone face mask – customised[13]

1. Manufacturer a custom 3D silicone mask,
2. Present the mask to the face recognition system

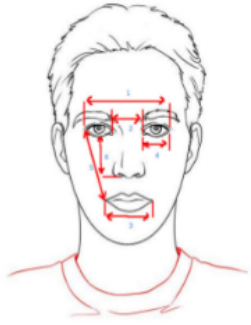
[13] K. Kotwal *et al.* "Multispectral Deep Embeddings As a Countermeasure To Custom Silicone Mask Presentation Attacks", *IEEE TBIOM*, 4(1), pp. 238–251, 2019 (10.1109/TBIOM.2019.2939421)

Face PAI

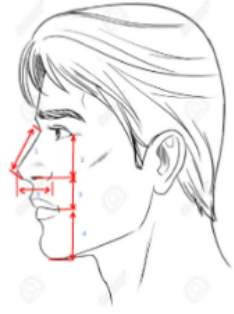
Silicone face mask – customised

- The method
  1. Acquire a 3D scan, measurements, and multiple 2D colour images of the target's face

(a) 3D-scan



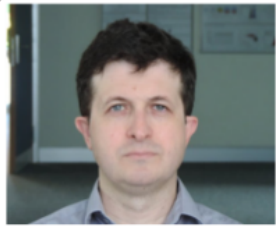
(b) Frontal face measurements (6)



(c) Profile measurements (5)

(d) Right 90°

(e) Right 45°



(f) Frontal

(g) Left 45°

(h) Left 90°

Face PAI

Silicone face mask – customised

2. Send the information to a manufacturer (e.g., Nimba Creations[14]), who will generate a customised 3D silicone mask, including manual application of facial features (e.g., skin colour, eyebrows, etc.), for ≈ 4,000 USD

Raw mask

Intermediate mask

Final mask

3. Present the mask to the face recognition system

[14] https://www.nimbacreations.com/

Face PAI

Silicone face mask – customised

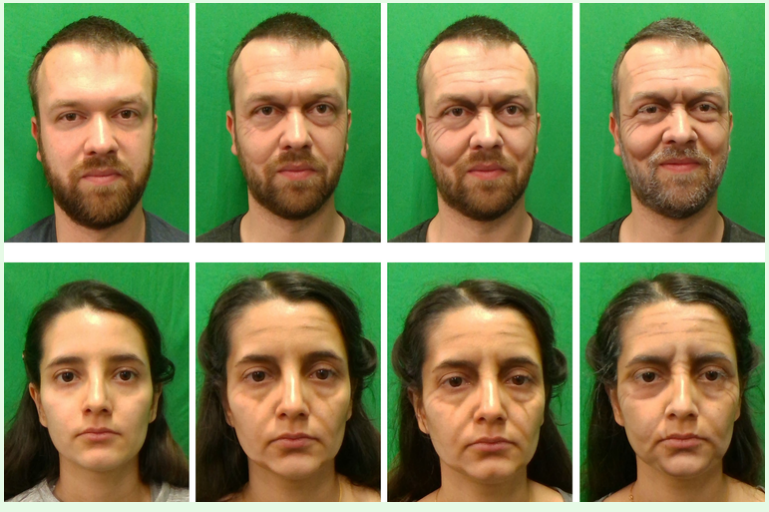
- The customised silicone face masks are quite life-like and they allow for some flexibility in facial movement
- Effective for launching PAs against face recognition systems[15]

[15] Ramachandra, R. *et al.* "Custom silicone Face Masks: Vulnerability of Commercial Face Recognition Systems & Presentation Attack Detection", *IEEE IWBF*, pp. 1–6 (2019)

Face PAI

More methods under investigation

- make-up: apply make-up to the attacker's face to impersonate an enrolled user of a face recognition system or to obfuscate by simulating aging[16]:



[16] K. Kotwal *et al.*, "Detection of Age-Induced Makeup Attacks on Face Recognition Systems Using Multi-Layer Deep Features", *IEEE TBIOM*, 2019 (10.1109/TBIOM.2019.2946175)

Face PAI

More methods under investigation

- digital morphing[17], deepfake face-swaps[18], biometric template inversion[19], ...

[17] E. Sarkar *et al.*, "Are GAN-based morphs threatening face recognition?", *IEEE ICASSP*, 2022 (10.1109/ICASSP43922.2022.9746477)

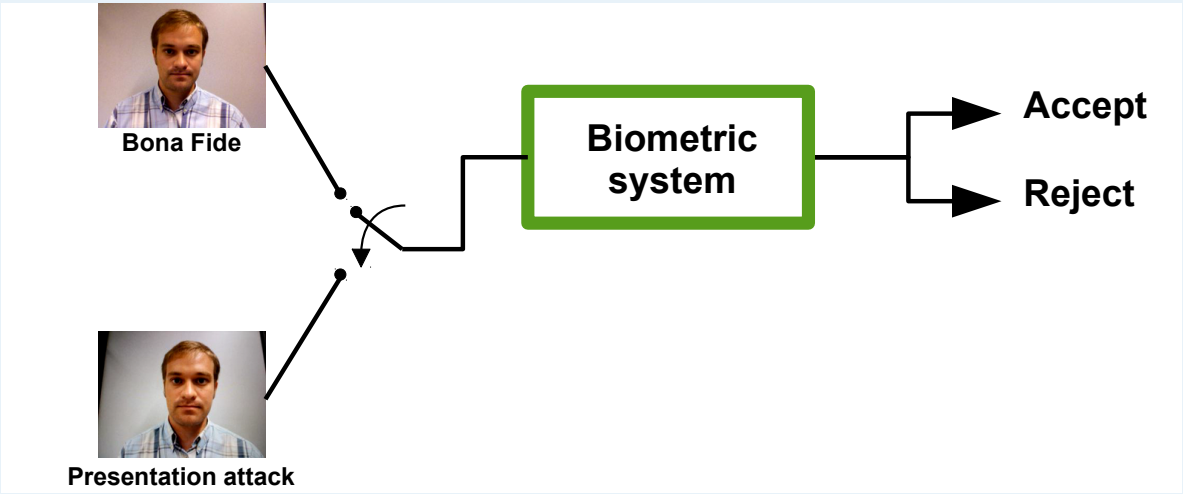
[18] P. Korshunov and S. Marcel, "Subjective and Objective Evaluation of Deepfake Videos", *IEEE ICASSP* 2021 (10.1109/ICASSP39728.2021.9414258)

[19] H. Otroshi Shahreza, V. Krivokuca Hahn and S. Marcel, "Face Reconstruction from Deep Facial Embeddings using a Convolutional Neural Network", *IEEE ICIP* 2021 (10.1109/ICIP46576.2022.9897535)

Outline

Biometrics and PAD

Biometric sub-system: a binary classifier

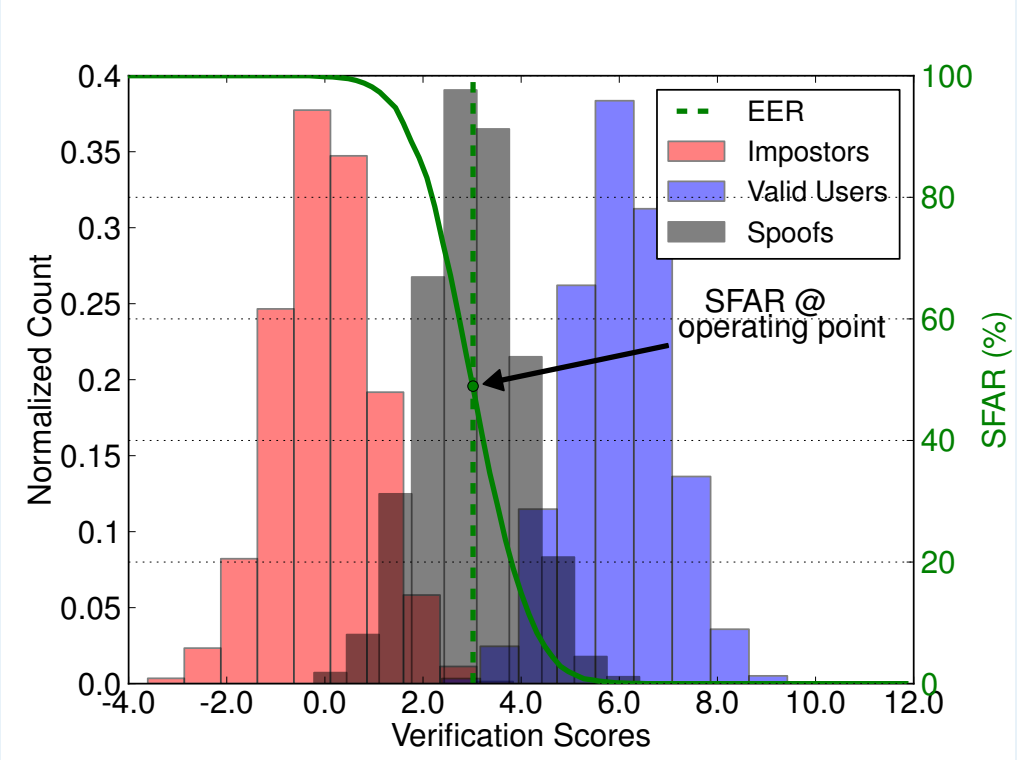


We measure the recognition accuracy :

- **False Match Rate (FMR)** or **False Accept Rate (FAR):** Proportion of bona fide zero-effort impostors that are *accepted* (i.e., classified as bona fide genuine presentations)
- **False Non-Match Rate (FNMR)** or **False Reject Rate (FRR):** Proportion of bona fide genuine presentations that are *rejected* (i.e., classified as either bona fide zero-effort impostors or PAs)

Biometrics and PAD

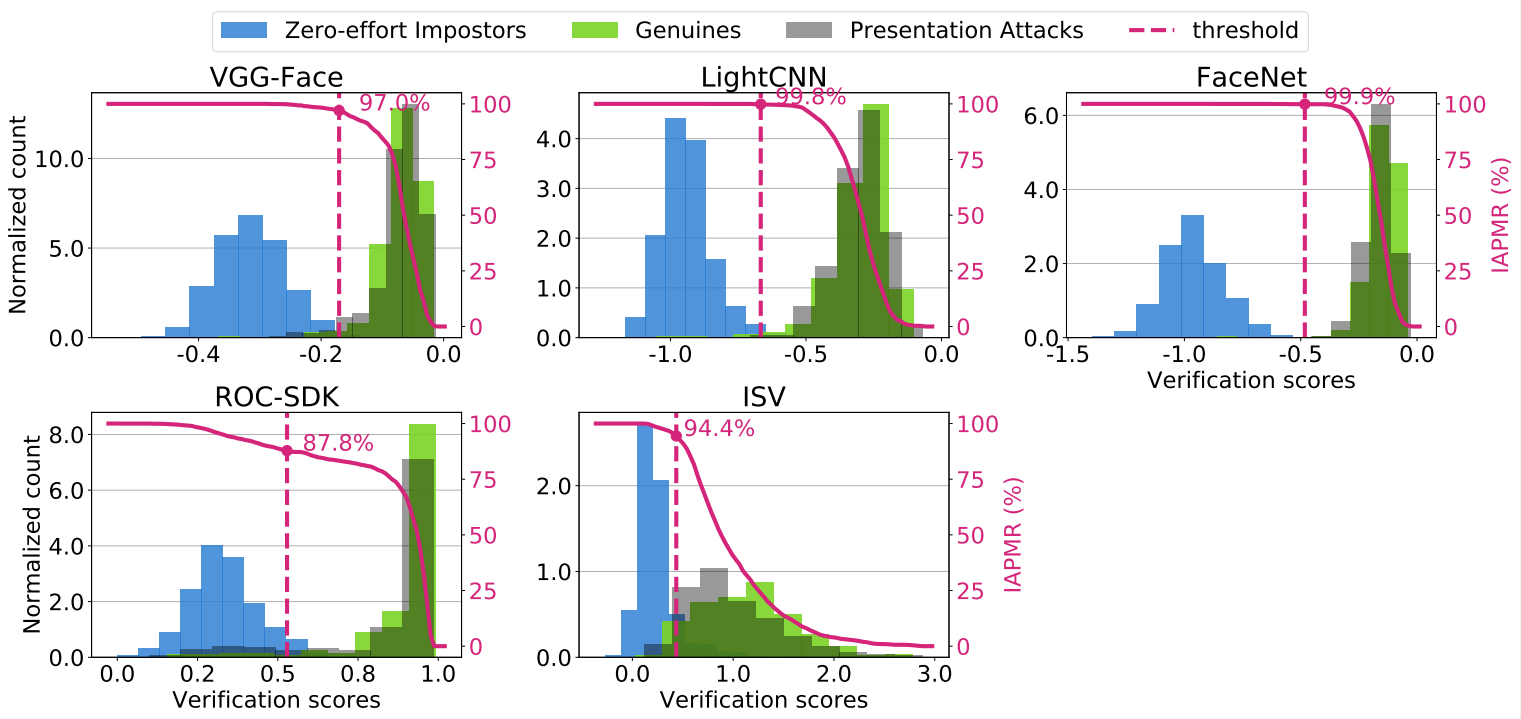
Biometric sub-system: a binary classifier



We can also measure the vulnerability as:
Impostor Attack Presentation Accept Rate (IAPAR) or
Impostor Attack Presentation Match Rate (IAPMR) or
Spoofing False Accept Rate (SFAR): Proportion of PAs that are *accepted* (i.e., classified as bona fide genuine presentations)

Biometrics and PAD

Vulnerability of Deep Face Recognition [20]

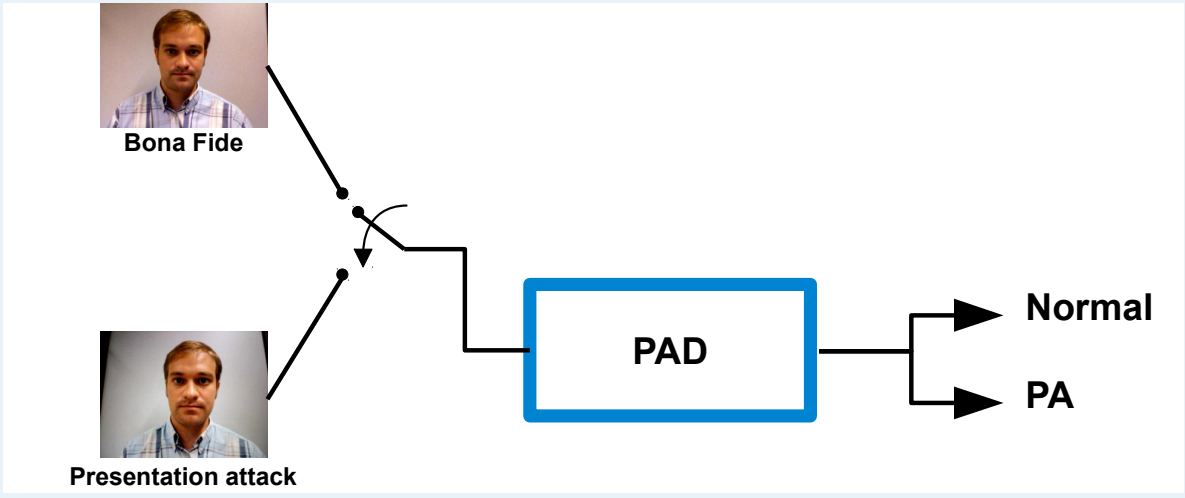


- FR systems using CNN are very vulnerable (up to 99% IAPMR)
- Improved FR accuracy translates into improved vulnerability

[20] A. Mohammadi et al., "Deeply vulnerable: a study of the robustness of face recognition to presentation attacks", IET Biometrics, 2017 (10.1049/iet-bmt.2017.0079)

Biometrics and PAD

PAD sub-system: a binary classifier



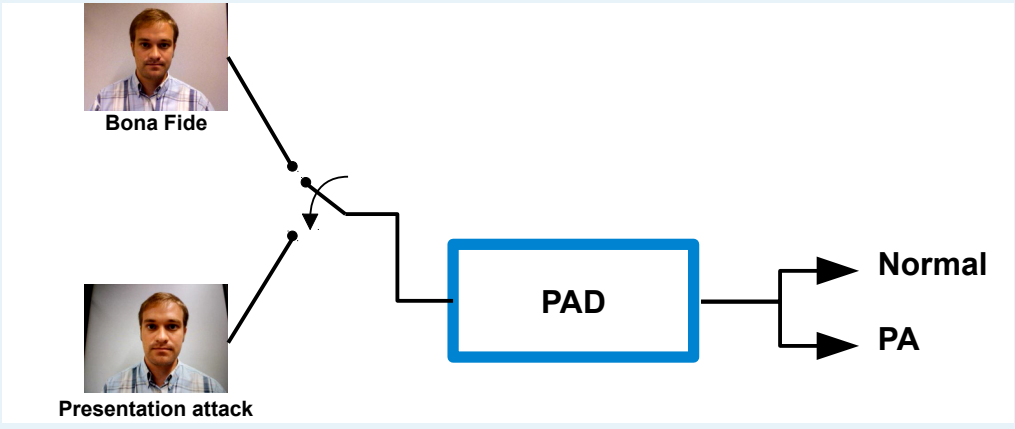
We measure 2 errors:

- **Attack Presentation Classification Error Rate (APCER):** Proportion of PAs *incorrectly* classified as bona fide presentations
- **Bona fide Presentation Classification Error Rate (BPCER):** Proportion of bona fide presentations *incorrectly* classified as PAs

Biometrics and PAD

PAD methods

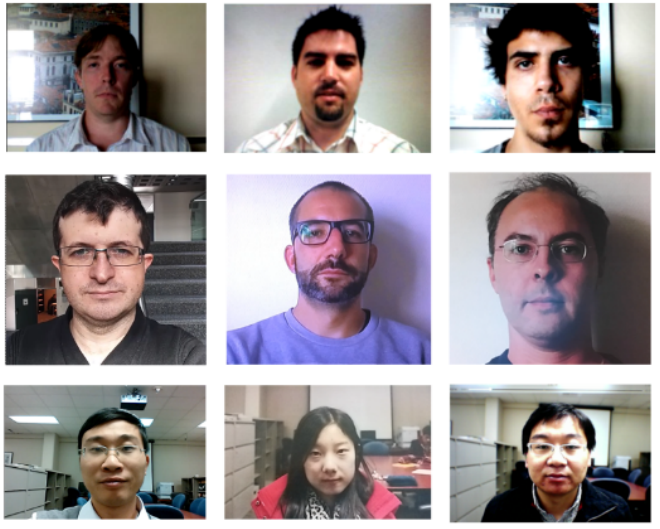
- software-based (**SW**): biometric data from the sensor is analysed to discriminate bona fide vs PA (eg. motion, texture)
- hardware-based (**HW**): an additional sensor (eg. multi-spectra) is used and its data analysed to discriminate bona fide vs PA (eg. temperature, pulse)
- challenge-response: the user interacts with the system (eg. prompted text in face/speaker recognition)



Outline

SW-based Face PAD

RGB only



Replay-Attack,Mobile, MSU-MFSD

WMCA

Waxface

SW-based Face PAD

From handcrafted classifiers to deep learning

- Motion analysis: Optical flow correlation and MLP to detect static PAIs [21]
- Texture analysis: Local Binary Patterns (LBP) and LDA/SVM to detect static/dynamic PAIs [22]
- Image quality: general image quality measures (IQM) and LDA to detect static/dynamic PAIs [23]

---

[21] A. Anjos and S. Marcel, "Motion-Based Counter-Measures to Photo Attacks in Face Recognition", *IET Biometrics*, 3(3), pp. 147–158, 2013 (10.1049/iet-bmt.2012.0071)

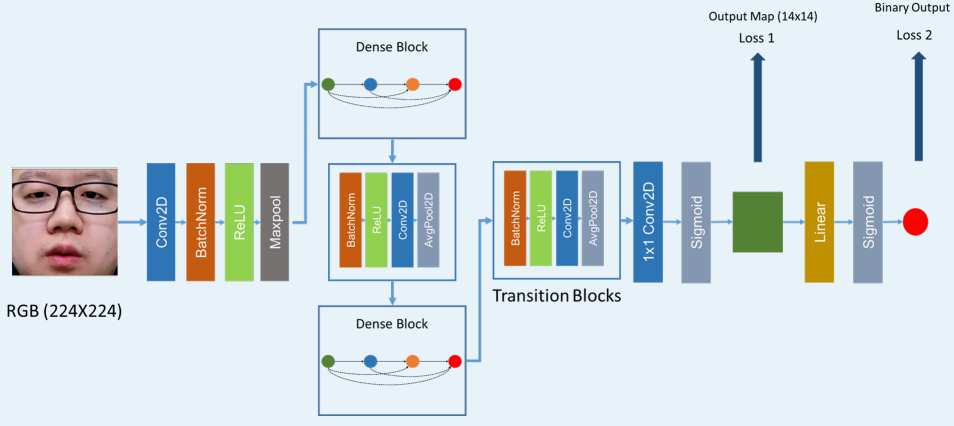
[22] I. Chingovska *et al.*, "On the Effectiveness of Local Binary Patterns in Face Anti-spoofing", *IEEE BIOSIG*, 2012

[23] J. Galbally, S. Marcel and J. Fierrez, "Image Quality Assessment for Fake Biometric Detection: Application to Iris, Fingerprint, and Face Recognition", *IEEE Transactions on Image Processing*, 2013 (10.1109/TIP.2013.2292332)

SW-based Face PAD

From handcrafted classifiers to deep learning

- Convolutional Neural Networks: DenseNet-based pixel-wise binary supervision [24] **outperformed LBP and IQM**



- Vision Transformers (ViTran): Fine-tuned ViTran[25] **outperformed CNNs on unseen attacks**

---

[24] A. George and S. Marcel, "Deep pixel-wise binary supervision for face presentation attack detection" *IEEE ICB* 2019 (10.1109/ICB45273.2019.8987370)

[25] A. George and S. Marcel, "On the Effectiveness of Vision Transformers for Zero-shot Face Anti-Spoofing" *IEEE IJCB* 2021 (10.1109/IJCB52358.2021.9484333)

Outline

HW-based Face PAD

RGB, Depth, NIR, SWIR and Thermal[26]

Home-made multi-spectra sensing station to capture:

- HQ VIS, HQ stereo NIR ($_{660,735,850,940\ nm}$)
- Depth from Intel SR435 (stereo) and HQ Thermal
- HQ SWIR ($_{1050,1200,1300,1450,1550,1650\ nm}$)

[26] https://www.idiap.ch/en/dataset/hq-wmca

HW-based Face PAD

PAD across spectrum

different channels provide complementary information from different sources and hence more robust PAD

- NIR offers several advantages for face PAD, especially for 2D attacks
- Prints and masks colored with non-metallic inks should be barely visible in NIR spectrum
- Wavelengths around 850–950 *nm* should provide better discrimination between human skin and other materials

HW-based Face PAD

PAD across spectrum

different channels provide complementary information from different sources and hence more robust PAD

- Thermal images[27] should make it easier to detect 2D and 3D mask attacks using the temperature distribution
- More precise information about the distribution of temperature should be needed to identify more sophisticated attacks such as make-up [28]
- SWIR channel should make it easy to identify skin easily due to the specific nature of reflectance spectra[29]

---

[27] Bhattacharjee, S. and Marcel, S. "What you can't see can help you – extended-range imaging for 3D-mask presentation attack detection", *IEEE BIOSIG* (2017)

[28] Kotwal, K. *et al.* "Detection of Age-Induced Makeup Attacks on Face Recognition Systems Using Multi-Layer Deep Features", *IEEE T-BIOM*, 2(1), pp. 15–25 (2020)

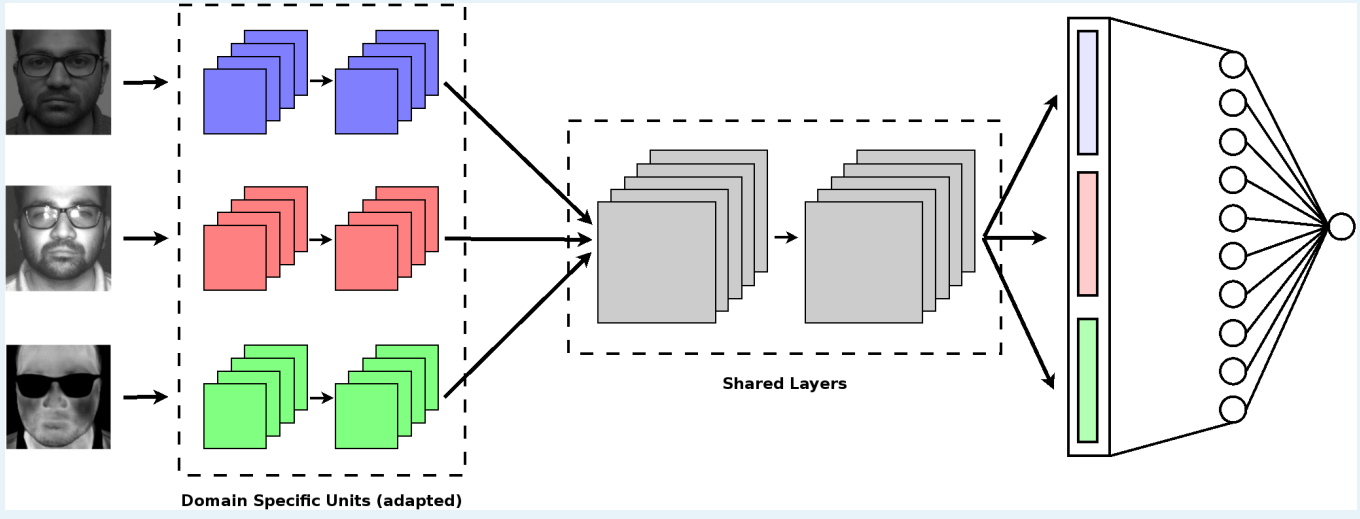
[29] Kotwal, K. *et al.* "Multispectral Deep Embeddings As a Countermeasure To Custom Silicone Mask Presentation Attacks", *IEEE T-BIOM*, 4(1), pp. 238–251 (2019)

HW-based Face PAD

Deep Learning (DL) PAD across spectrum

DL-based methods can be explored to detect a large range of PAIs:

- multi-channel (**RGB+NIR+SWIR**) CNN-based approaches [30] [31]



[30] G. Heusch *et al.*, "Deep Models and Shortwave Infrared Information to Detect Face Presentation Attacks", *IEEE TBIOM*, 2020 (10.1109/TBIOM.2020.3010312)

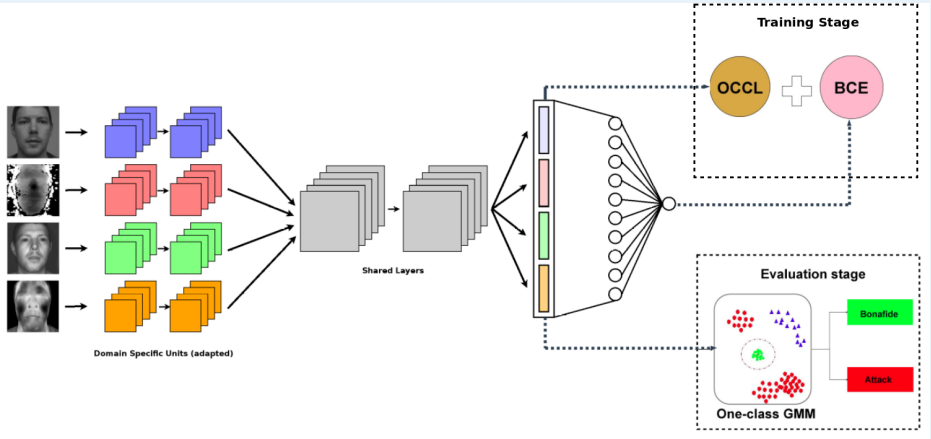
[31] A. George *et al.*, "Biometric face presentation attack detection with multi-channel convolutional neural network", *IEEE TIFS*, 2019 (10.1109/TIFS.2019.2916652)

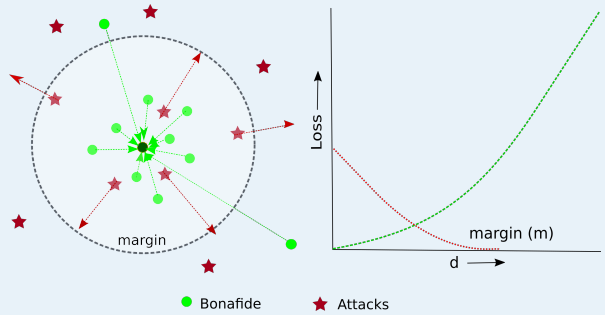
HW-based Face PAD

Deep Learning (DL) PAD across spectrum

DL-based methods can be explored to detect a large range of PAIs:

- a one class classifier framework: CNN embedding + One Class Constrastive Loss + GMMs [32]





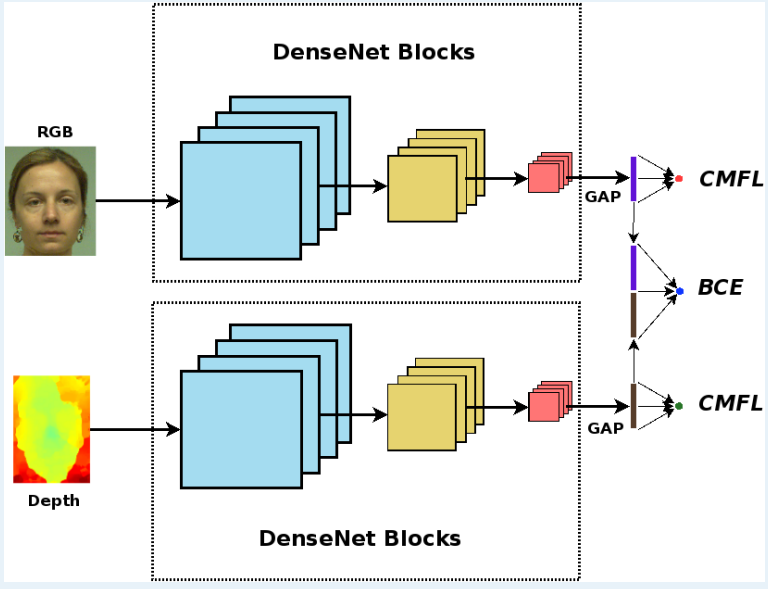
[32] A. George and S. Marcel, "Learning one class representations for face presentation attack detection using multi-channel convolutional neural networks", *IEEE TIFS*, 2020 (10.1109/TIFS.2020.3013214)

HW-based Face PAD

Deep Learning (DL) PAD across spectrum

DL-based methods can be explored to detect a large range of PAIs:

- a cross-modal (**RGB+Depth**) loss [33]



[33] A. George and S. Marcel, "Cross Modal Focal Loss for RGBD Face Anti-Spoofing", *IEEE CVPR* 2021 (10.1109/CVPR46437.2021.00779)

HW-based Face PAD

PAD across spectrum: some conclusions

- Even simple average temperature of face regions is effective against simpler 2D/3D attacks
- Majority of impersonation/obfuscation attacks can be identified using SWIR or RGB+NIR
- Wavelengths around 1450 *nm* provide good separation between skin and other objects
- Even low resolution SWIR sensors improves the PAD performance greatly

Outline

To conclude

Is PAD a solved problem?

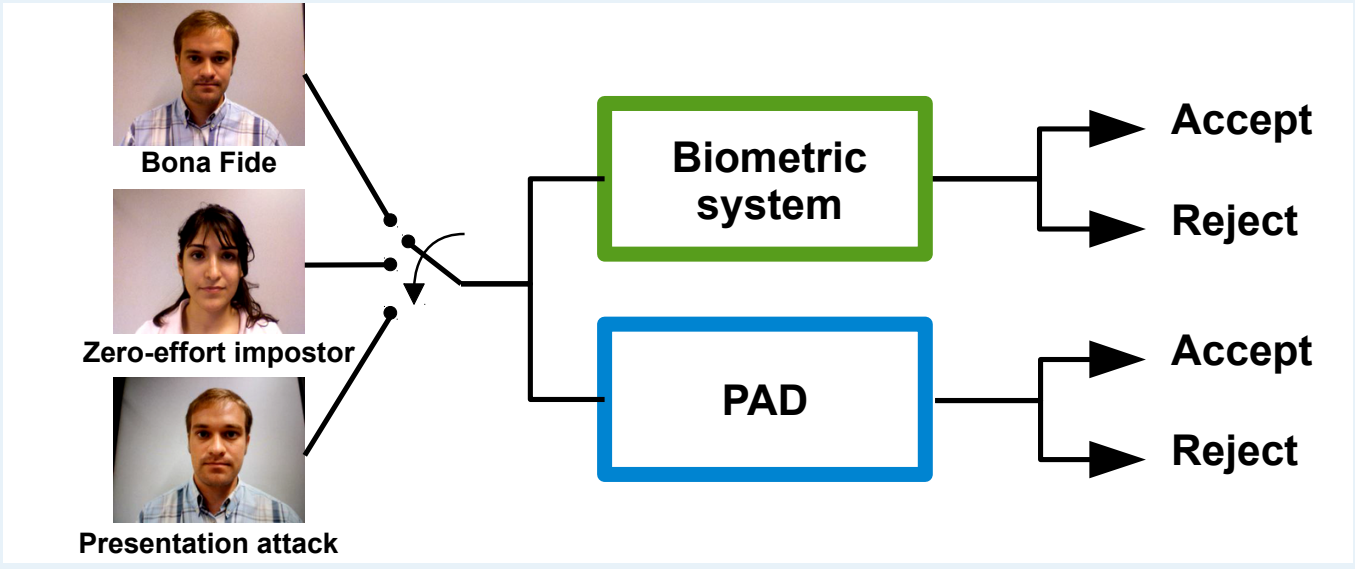
- biometrics is more prevalent hence **incentives** for launching PAs are multiplying
- active PAD research but generalisation (to unseen attacks) is challenging – **arms race**
- PAD is not a solved problem, it continues to be an **important field of research**

References

- S. Marcel *et al.*, "Handbook of Biometric Anti-Spoofing", Third Edition, *Springer*, 2023 (10.1007/978-981-19-5288-3)
- N. Evans, S. Marcel, A. Ross and A. Teoh, "Biometrics Security and Privacy Protection", *IEEE Signal Processing Magazine*, 2015 (10.1109/MSP.2015.2443271)
- Z. Yu *et al.*, "Deep Learning for Face Anti-Spoofing: A Survey", *IEEE TPAMI*, 2022 (10.1109/TPAMI.2022.3215850)

Outline

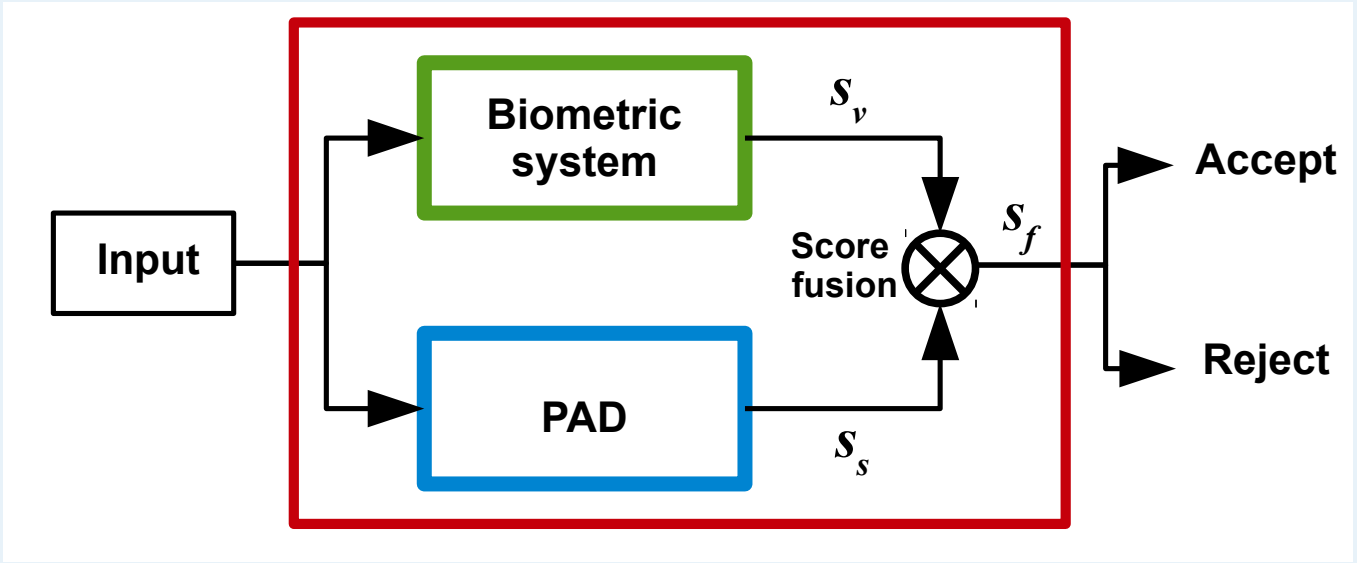
Two separate components



How to combine these two components ?

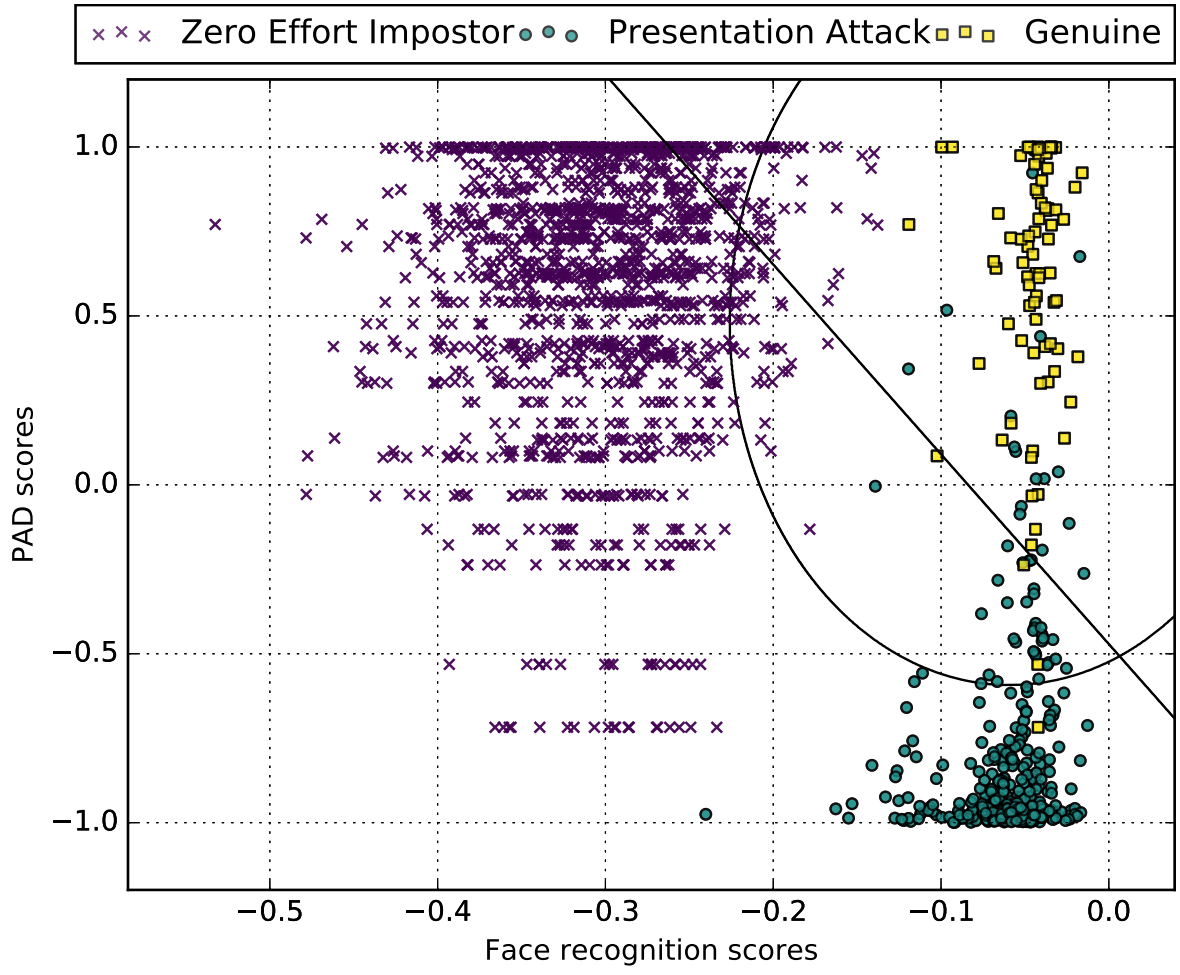
[34] *Biometrics Evaluation under Spoofing Attacks*, I. Chingovska and al., IEEE TIFS, 2014.

Fusion scheme

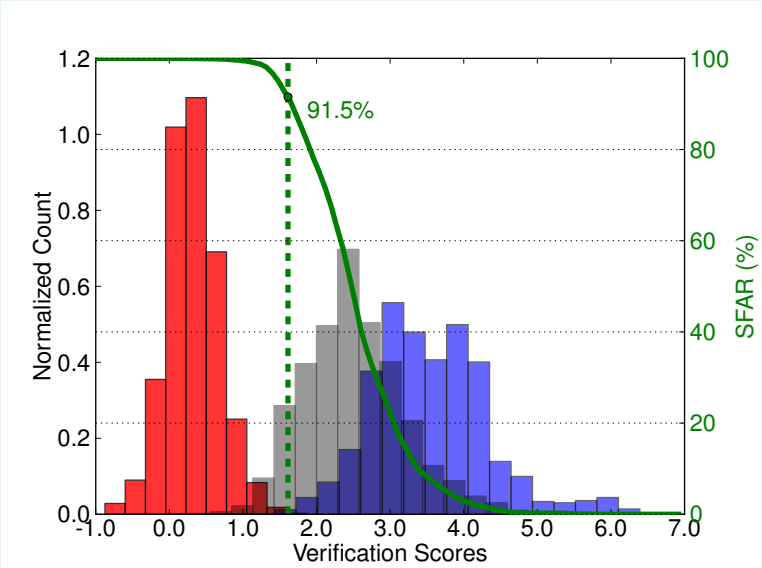


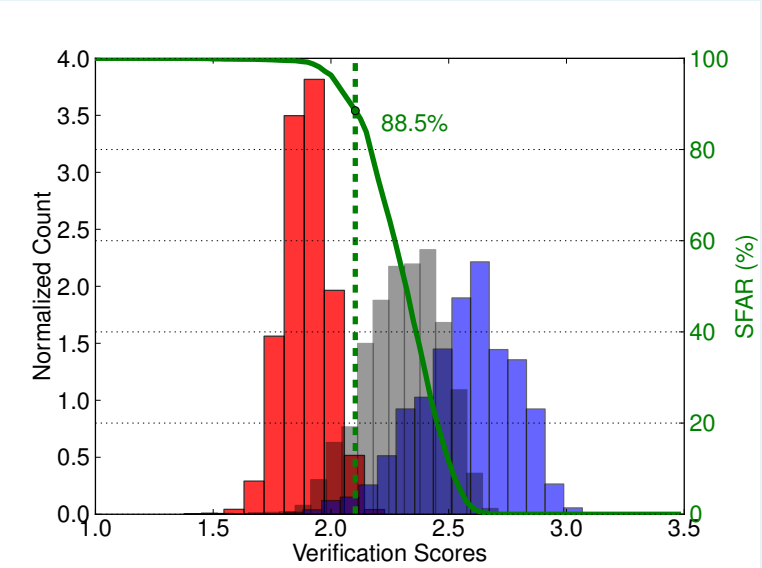
One unique threshold to be determined

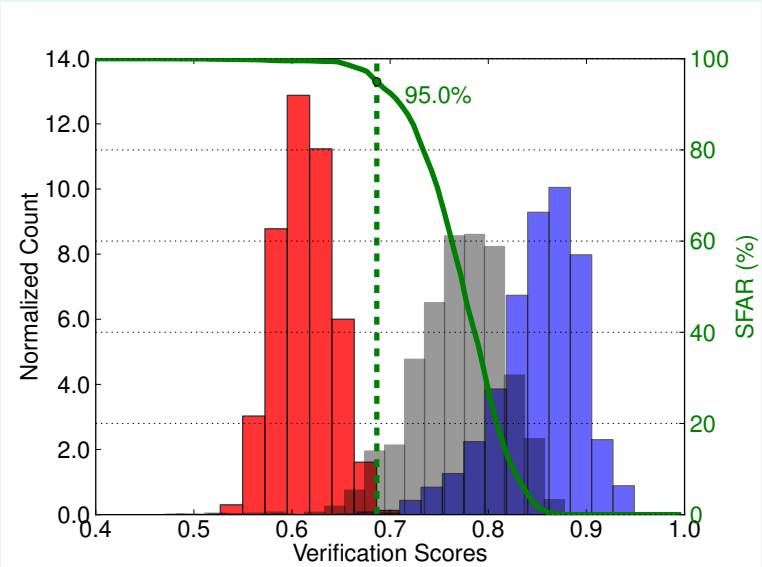
Fusion scheme

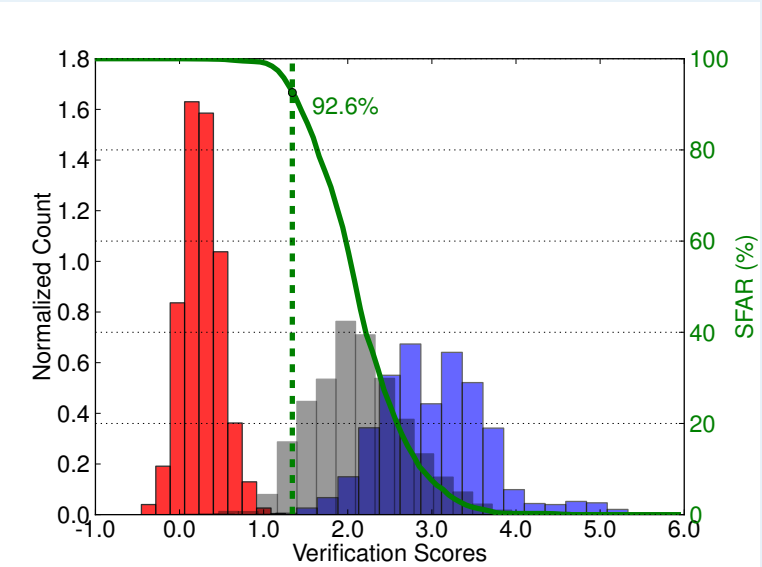


Biometric systems without PAD (no fusion)

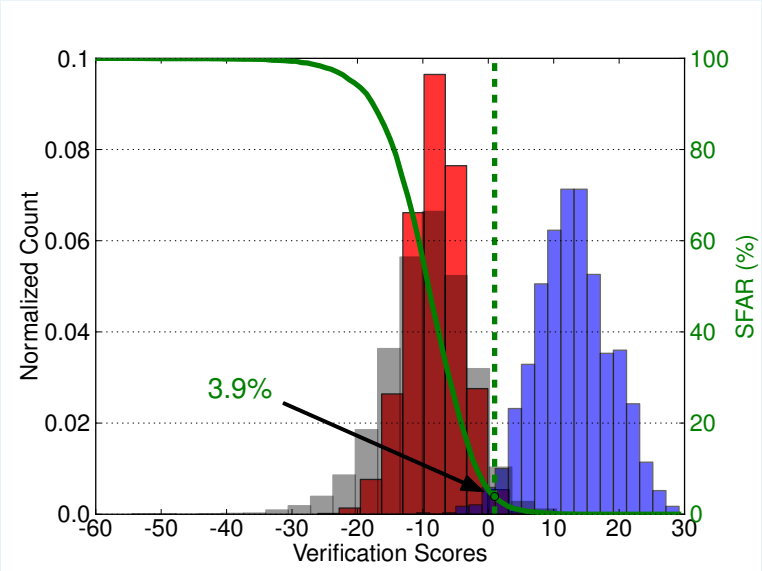


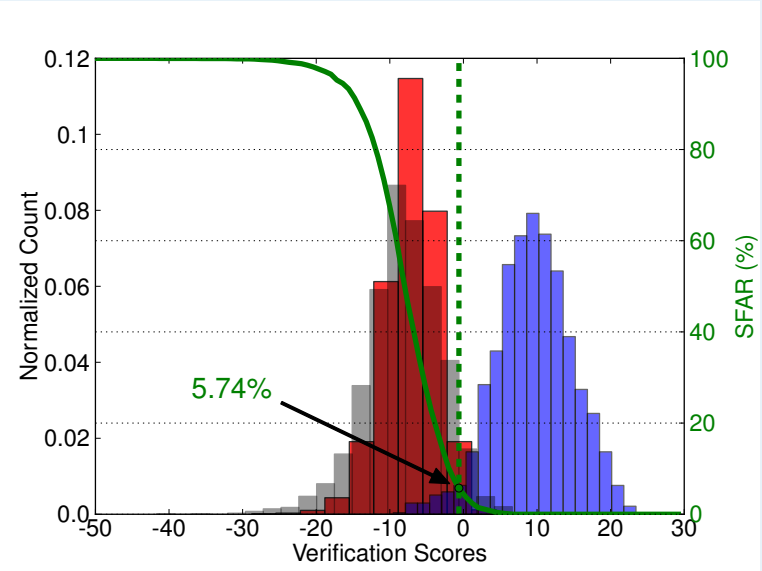


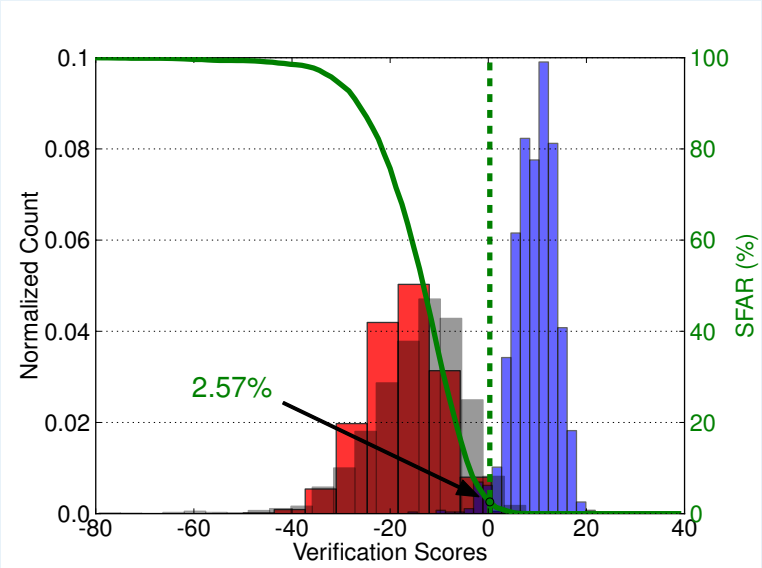


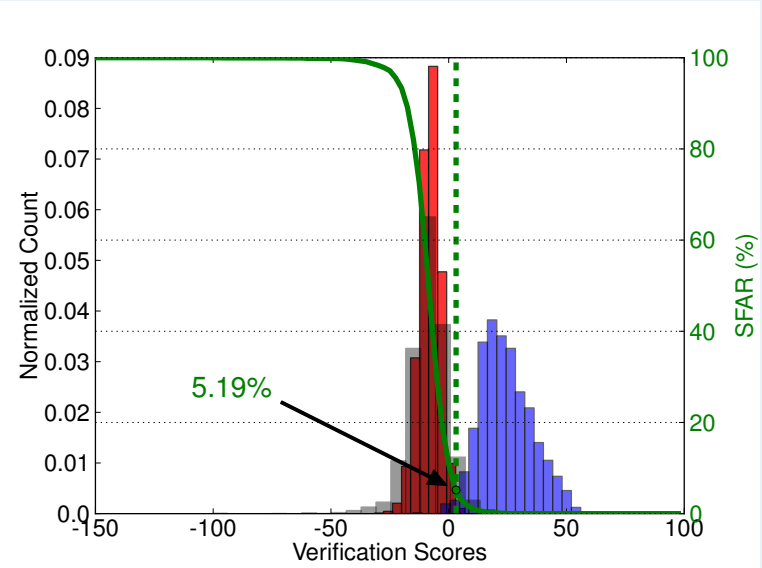


Biometric systems + PAD (fusion)

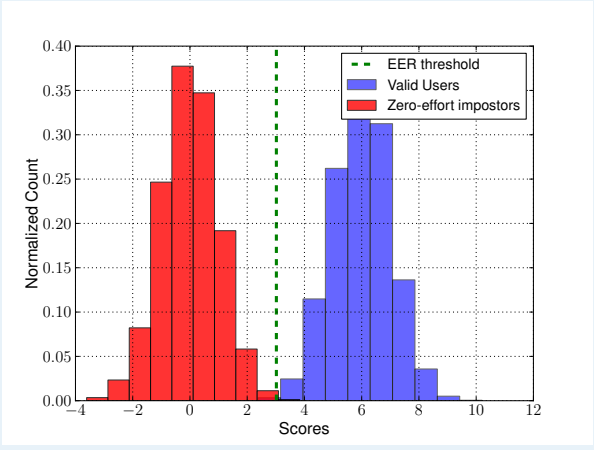


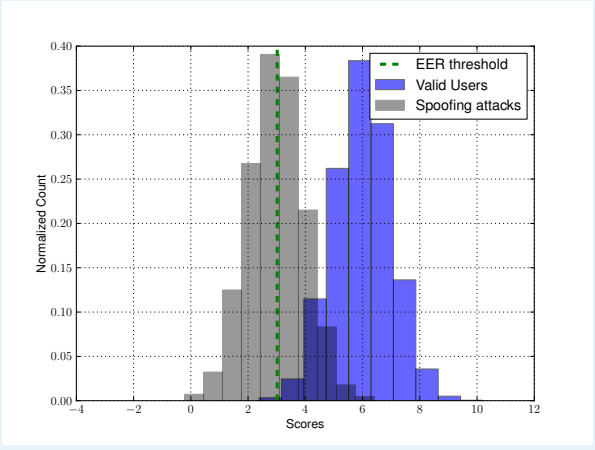






Measuring the performance





We still measure 3 errors:

- False Rejection Rate (FRR): % of genuine users falsely rejected
- False Acceptance Rate (FRR): % of zero-effort impostors falsely accepted
- Spoof False Acceptance Rate (SFAR): % of presentation attacks falsely accepted

FAR_ω (development set)

Weighted error rate for the two negative classes (zero-effort impostors and presentation attacks):

$$\text{FAR}_\omega = (1-\omega) \cdot \text{FAR} + \omega \cdot \text{SFAR}$$

Determine τ^*_ω to minimize the difference between FAR_ω and FRR on the development set:

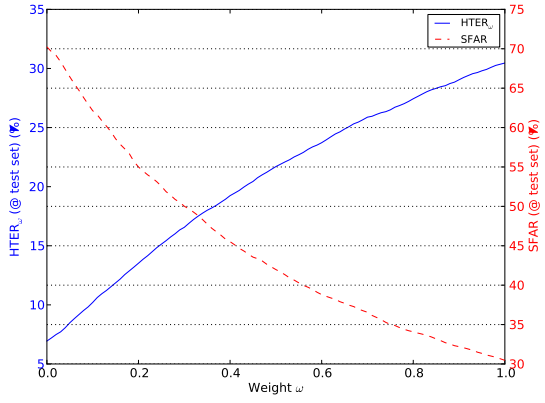
$$\tau^*_\omega = \arg \min_\tau |\text{FAR}_\omega(\tau, \mathcal{D}_{dev}) - \text{FRR}(\tau, \mathcal{D}_{dev})|$$

HTER$_\omega$ (test set)

Measuring both the verification performance and the spoofability of the system

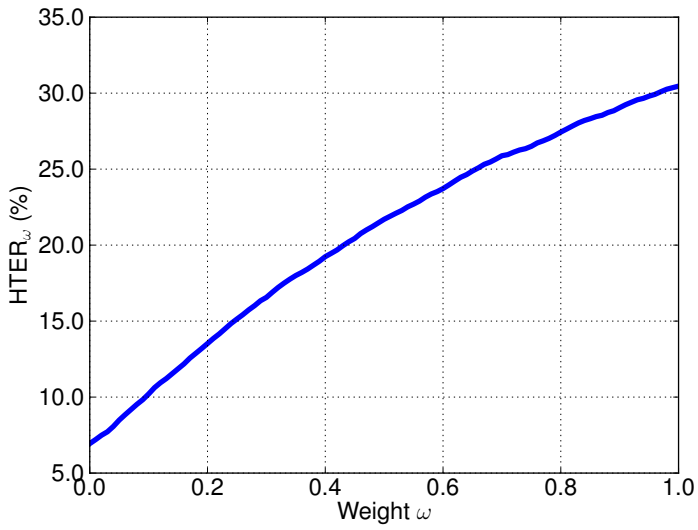
$$\mathrm{HTER}_\omega(\tau^*_\omega, \mathcal{D}_{test}) = \frac{\mathrm{FAR}_\omega(\tau^*_\omega, \mathcal{D}_{test}) + \mathrm{FRR}(\tau^*_\omega, \mathcal{D}_{test})}{2}$$

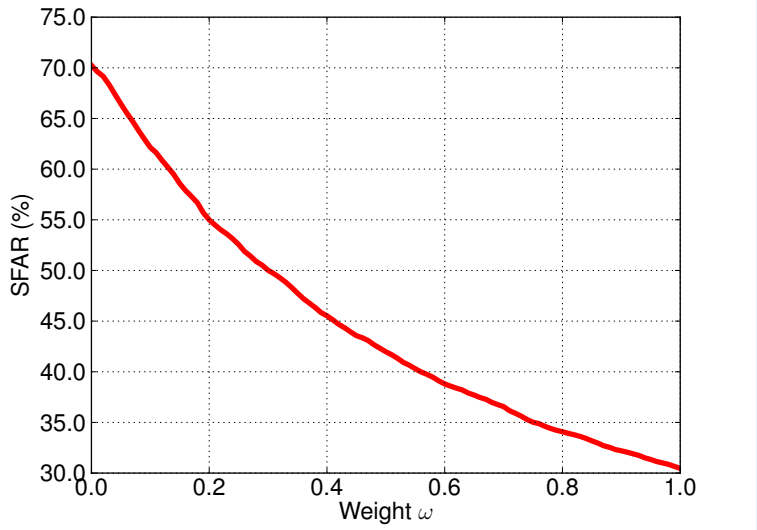
Ploting HTER$_\omega$ or SFAR



EPSC in action

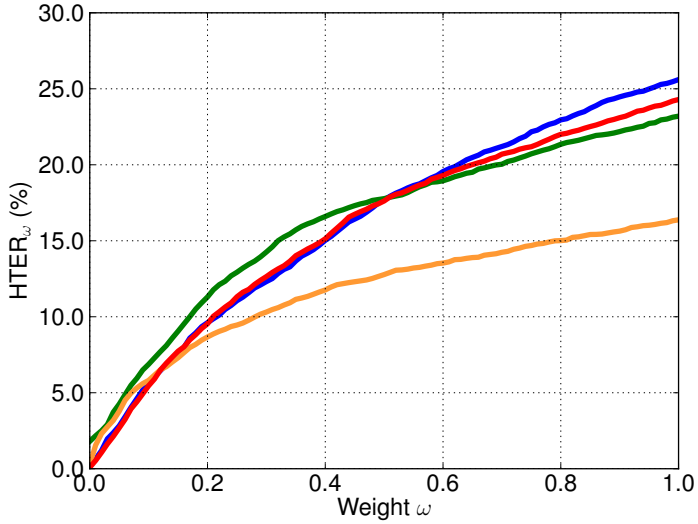
EPSC: HTER$_\omega$ and SFAR

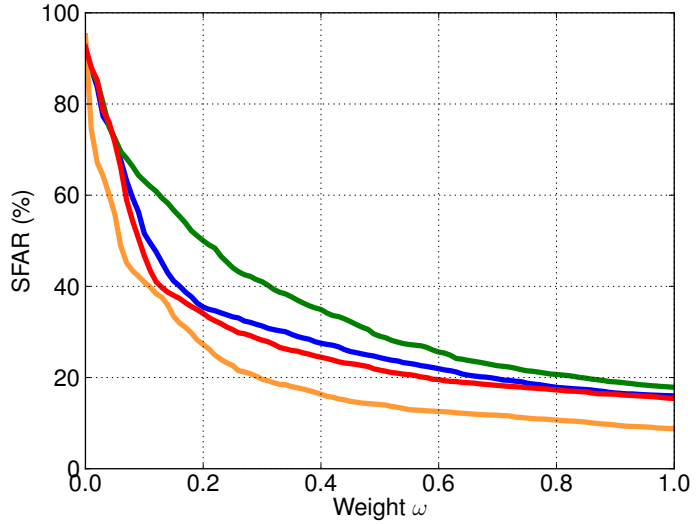




EPSC Examples

EPSC to compare biometric systems only

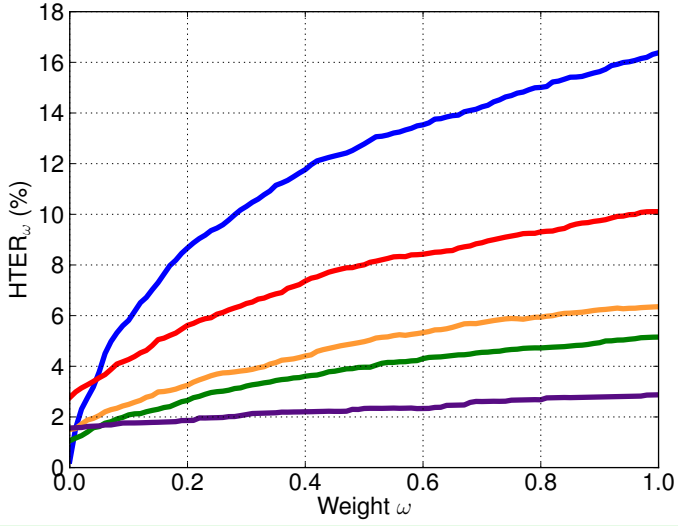


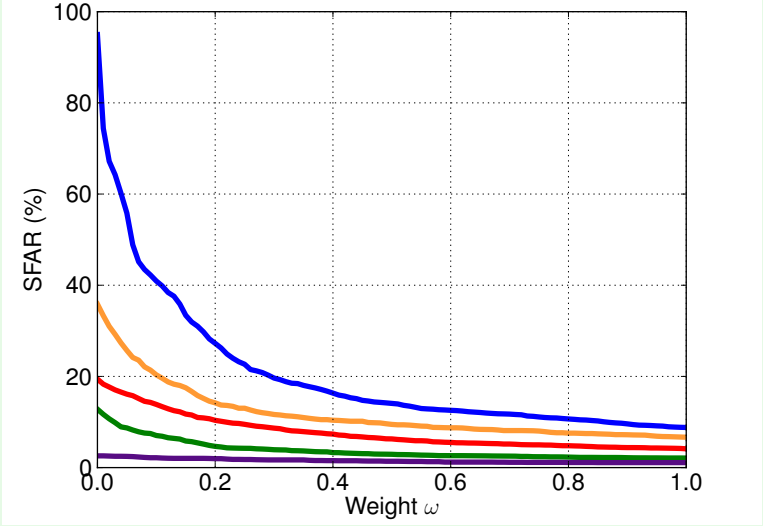


4 biometric systems (no PAD): using the orange subsequently

EPSC Examples

EPSC to compare PAD

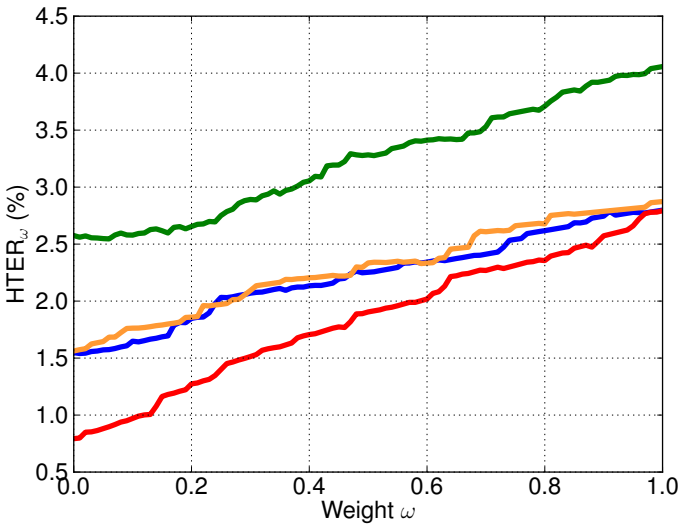


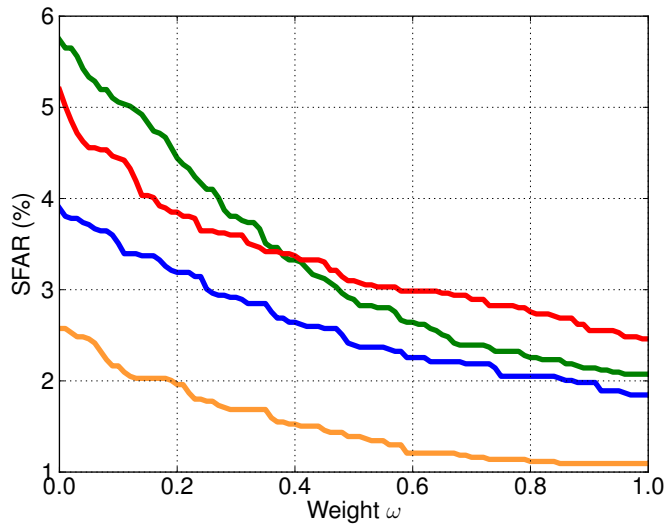


1 biometric system (blue), same system + 3 PADs (red, orange,green), same system + all PADs (purple)

EPSC Examples

EPSC to compare biometric systems fused with ALL PADs





4 biometric system + all PADs

Thank you for your attention!

Prof Sébastien Marcel (www.idiap.ch/∼marcel)

Idiap Research Institute, Martigny, Switzerland