Presentation Attack Detection and Unknown Attacks

Marta Gomez-Barrero

Hochschule Ansbach

IFPC'20, 28/10/2020

Outline

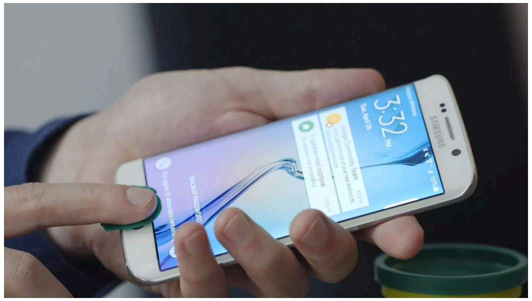
HOCHSCHULE ANSBACH

- Introduction
- Unknown PAD
- Train on PA and BF
- Anomaly detection
- Conclusions

HOCHSCHULE ANSBACH

What are Presentation Attacks? [ISO/IEC IS 30107-1 on PAD]

- Presentation attack: presentation to the biometric capture subsystem with the goal of interfering with the operation of the biometric system
  - **Impostor**: the attacker attempts to being matched to someone else's biometric reference
  - **Identity concealer**: the attacker attempts to avoid being matched to their own biometric reference (i.e., to avoid a black-list)



Introduction

Current situation

- There is a clear need to develop Presentation Attack Detection (PAD) methods, especially for non-supervised access control scenarios
- Good news, there is a lot of work done, reporting really good results and error rates close to 0%

😃

J. Galbally, S. Marcel, J. Fierrez, “Biometric antispoofing methods: A survey in face recognition”, IEEE Access, 2014

R. Ramachandra, C. Busch, “Presentation attack detection methods for face recognition systems: A comprehensive survey”, ACM CSUR, 2017

E. Marasco, A. Ross, “A survey on antispoofing schemes for fingerprint recognition systems” ACM CSUR, 2014

C. Sousedik, C. Busch, “Presentation attack detection methods for fingerprint recognition systems: a survey”, IET Biometrics, 2014

J. Galbally, M. Gomez-Barrero, , “Presentation attack detection in iris recognition”, in Iris and Periocular Biometrics, IET, 2017

A. Czajka, K. Bowyer, “Presentation attack detection for iris recognition: An assessment of the state-of-the-art”, ACM CSUR, 2018

Introduction

Current situation

- But what about unknown attacks? Attacks not seen during training

- State of the art PAD methods, based on image quality measures and Support Vector Machines (SVMs), present poor generalisation capabilities to detect unknown attacks
  - Error rates are multiplied by up to 6!
  - Even worse, the performance of the system drops significantly for lower APCER values – probably the most interesting operating points

😟

T. de Freitas Pereira, A. Anjos, J. D. Martino, S. Marcel, “Can face anti-spoofing countermeasures work in a real world scenario?”, in Proc. ICB, 2013

O. Nikisins, A. Mohammadi, A. Anjos, S. Marcel, “On effectiveness of anomaly detection approaches against unseen presentation attacks in face anti-spoofing”, in Proc. ICB, 2018

HOCHSCHULE ANSBACH

Evaluation Metrics – Let‘s keep it standard

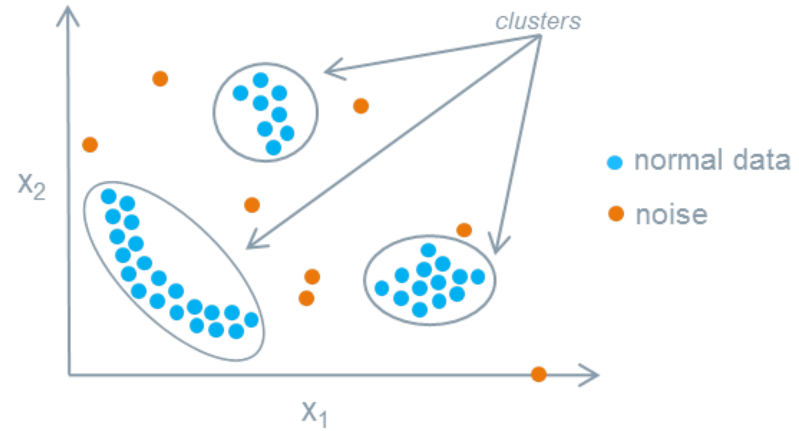
- In compliance with the ISO/IEC IS 30107-3 on biometric PAD – Part 3: Testing and reporting, the following metrics should be used:
  - **Attack Presentation Classification Error Rate (APCER):** percentage of attack presentations wrongly classified as bona fide presentations.
  - **Bona Fide Presentation Classification Error Rate (BPCER):** percentage of bona fide presentations wrongly classified as attack presentation.
  - **Detection Equal Error Rate (D-EER)**: operation point where APCER = BPCER.

HOCHSCHULE ANSBACH

What can we do?

- Profit from ALL the data we have and simulate an unknown attack scenario, using different Presentation Attack Instrument (PAI) species for train and test
  - Leave-One-Out (LOO) protocol
  - Deep learning! Generative Adversarial Networks (GANs)

- Employ anomaly detection techniques:
  - Bona fides = “normal“ data
  - Presentation attacks = outliers, noise



HOCHSCHULE ANSBACH

Fisher vector representation

- Add generalisation capabilities to existent PAD approaches
- For instance: project the features used by the classifiers into a new feature space, where new attacks will look similar to known attacks
- Good results with Fisher vectors for face, fingerprints, and speech!
  - Deep learning is not the only way to good results and generalisation 😉

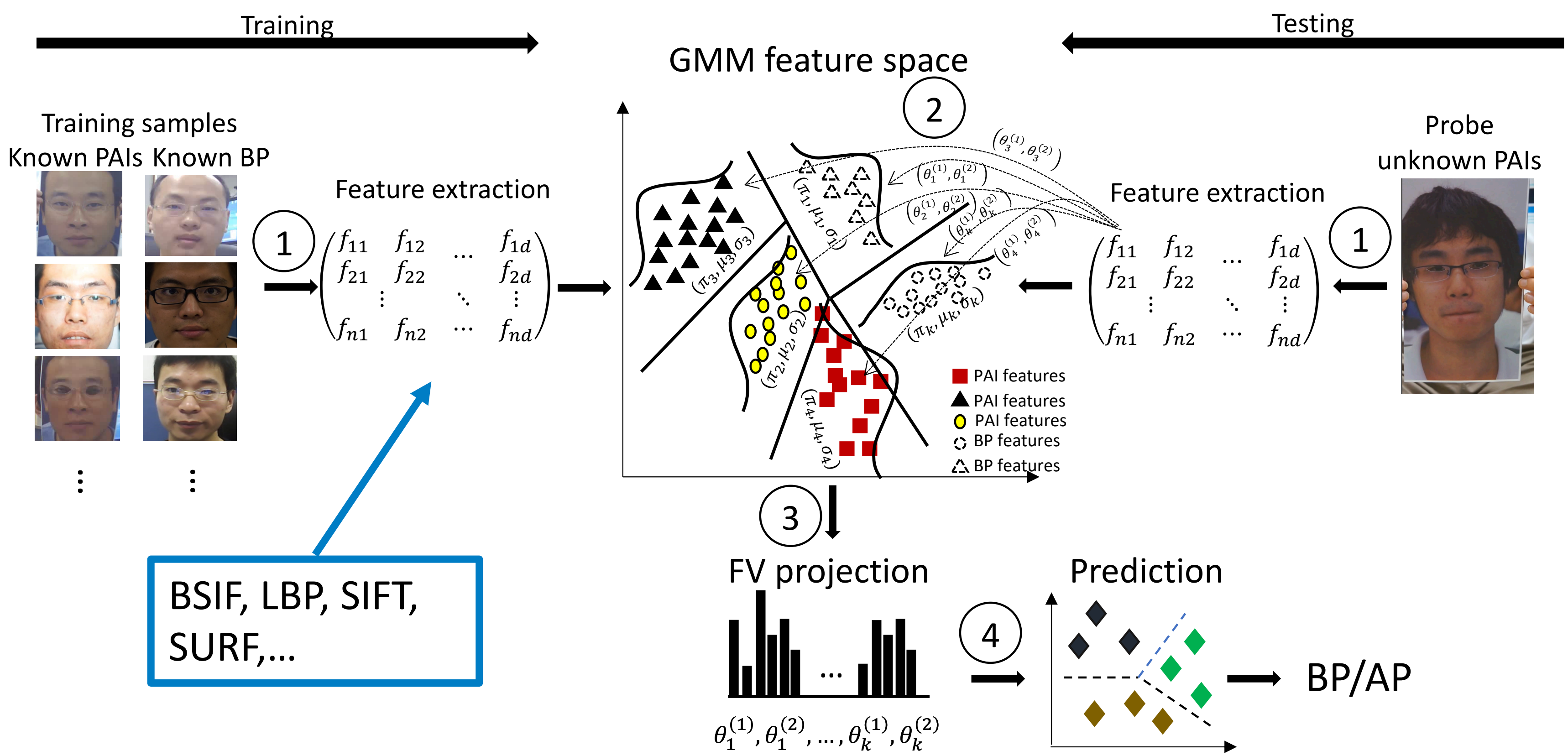
L. J. Gonzalez-Soler, M. Gomez-Barrero, C. Busch, “Fisher Vector Encoding of Dense-BSIF Features for Unknown Face Presentation Attack Detection”, in Proc. BIOSIG, 2020

L. J. Gonzalez-Soler, J. Patino, M. Gomez-Barrero, M. Todisco, C. Busch, N. Evans, “Texture-based Presentation Attack Detection for Automatic Speaker Verification”, in Proc. WIFS, 2020 (ArXiV)

L. J. Gonzalez-Soler, M. Gomez-Barrero, L. Chang, A. Perez-Suarez, C. Busch, „Fingerprint Presentation Attack Detection Based on Local Features Encoding for Unknown Attacks“, in arXiv:1908.10163, 2019

HOCHSCHULE ANSBACH

Fisher vector representation



L. J. Gonzalez-Soler, M. Gomez-Barrero, C. Busch, "Fisher Vector Encoding of Dense-BSIF Features for Unknown Face Presentation Attack Detection", in Proc. BIOSIG, 2020

HOCHSCHULE ANSBACH

Fisher vector representation

- For known attacks (CASIA database):
  - BSIF + SVM → D-ERR = 10.21%
  - State of the art (no deep learning) → D-ERR = 4.60%
  - BSIF + Fisher Vectors + SVM → D-EER = 1.79%

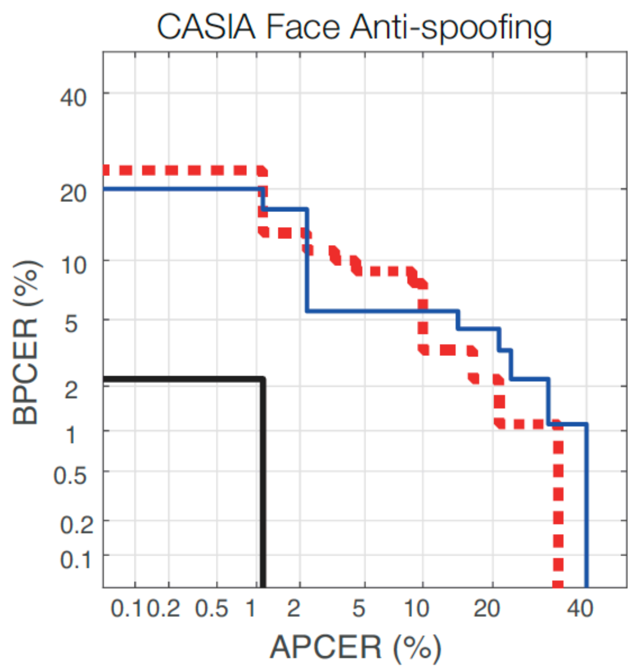
- For unknown attacks:

	CASIA			REPLAY-ATTACK			REPLAY-MOBILE		
	Cut	Warped	Video	Digital	Printed	Video	Digital	Printed	Video
AUC	99.6	97.9	99.9	100	99.9	100	100	100	100
D-EER	4.11	6.15	1.37	0.00	1.35	0.00	0.00	0.34	0.02

L. J. Gonzalez-Soler, M. Gomez-Barrero, C. Busch, “Fisher Vector Encoding of Dense-BSIF Features for Unknown Face Presentation Attack Detection”, in Proc. BIOSIG, 2020

Train on PA and BF

Fisher vector representation



Train vs test	BPCER10	BPCER20	BPCER100
warped + video vs cut	5.56	5.56	16.7
video + cut vs warped	3.33	8.89	13.3
cut + warped vs video	0.00	0.00	0,00

- Depending on the PAI species / training set, very different results!
- And this is not exclusive to face, but has also been reported for other characteristics (e.g., fingerprint)

🤔

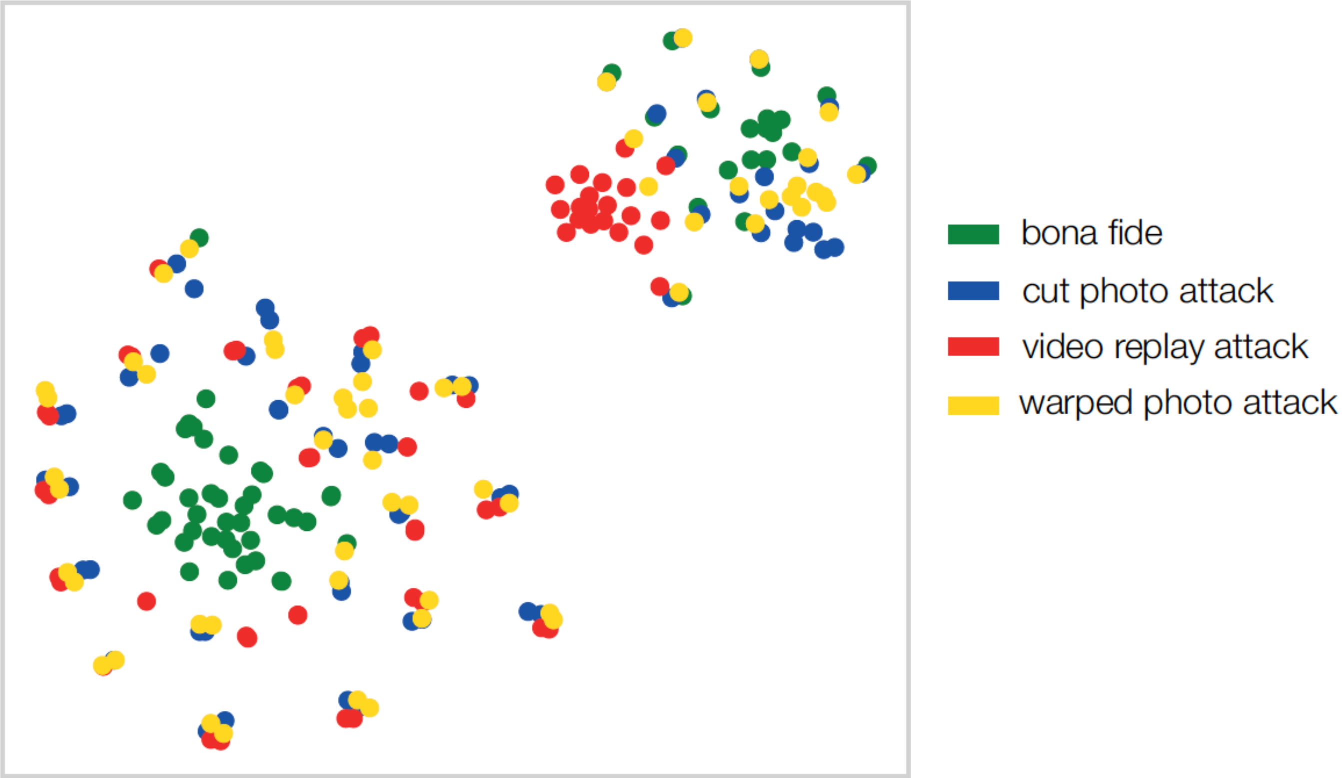
L. J. Gonzalez-Soler, M. Gomez-Barrero, L. Chang, A. Perez-Suarez, C. Busch, “On the Impact of Different Fabrication Materials on Fingerprint Presentation Attack Detection”, in Proc. ICB, 2019

T. Chugh, A. K. Jain, “Fingerprint presentation attack detection: Generalization and efficiency”, in Proc. ICB, 2019

HOCHSCHULE ANSBACH

Fisher vector representation

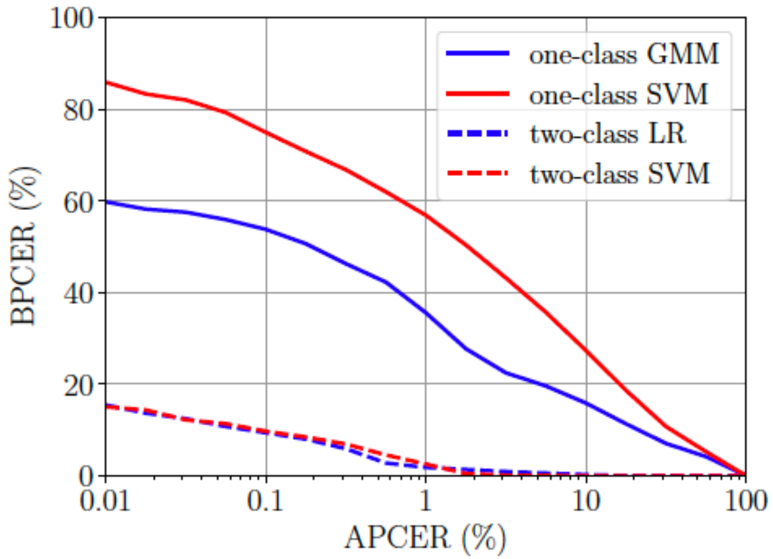
t-SNE visualisation of the FV
common feature space for CASIA



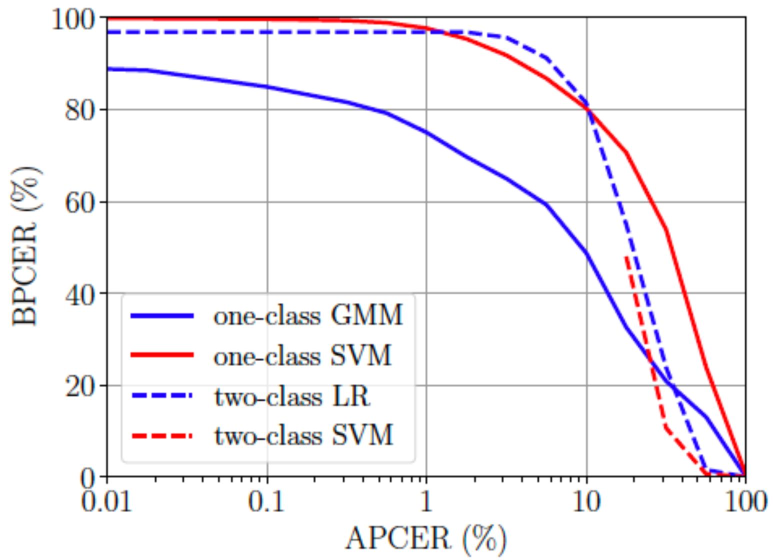
One-Class Classifiers

- Adapt classifiers such as SVMs or Gaussian Mixture Models (GMMs) to be trained only on bona fide data

Development



Evaluation

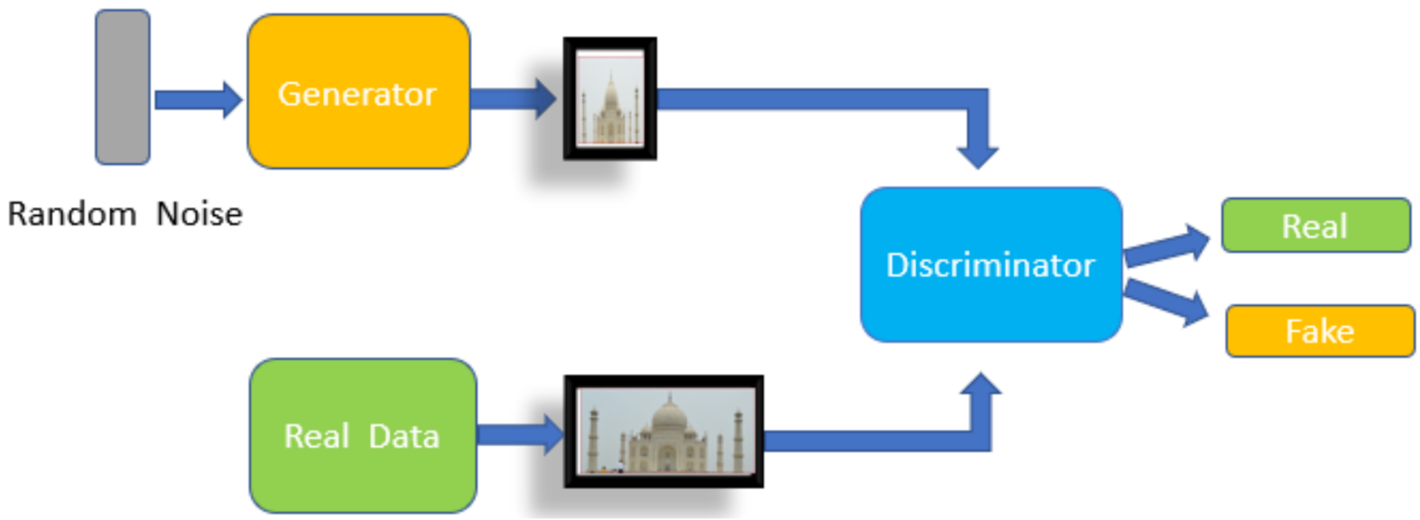


O. Nikisins, A. Mohammadi, A. Anjos, S. Marcel, “On effectiveness of anomaly detection approaches against unseen presentation attacks in face anti-spoofing”, in Proc. ICB, 2018

HOCHSCHULE ANSBACH

Generative Adversarial Networks (GANs)

- A GAN is made of two parts:
  - Generator network: takes as input a random vector and decodes it into a synthetic image
  - Discriminator network (adversary): takes as input an image (real or synthetic) and predicts its type



HOCHSCHULE ANSBACH

Generative Adversarial Networks (GANs)

- We can use a trained discriminator for PAD:
  - Synthetic = presentation attack
  - Real = bona fide

- For fingerprint PAD, APCER @ BPCER = 0.2%:

	Gelatin	Pigmented	Playdoh	Woodglue	Transpa-rency	Gold finger
Binary CNN	32.3%	**70.6%**	99.4%	44.3%	66.0%	88.2%
1-class GAN	**26.4%**	77.7%	**3.7%**	**14.8%**	**6.0%**	**60.8%**

	Dragonskin	Ecoflex	Monster Latex	Crayola Magic	Body Latex	2D paper
Binary CNN	**51.0%**	**60.7%**	**45.7%**	21.9%	**87.9%**	53.9%
1-class GAN	97.9%	95.2%	61.5%	**16.4%**	99.7%	**43.2%**

J. J. Engelsma, A. K. Jain, "Generalizing fingerprint spoof detector: Learning a one-class classifier" in Proc. ICB, 2019

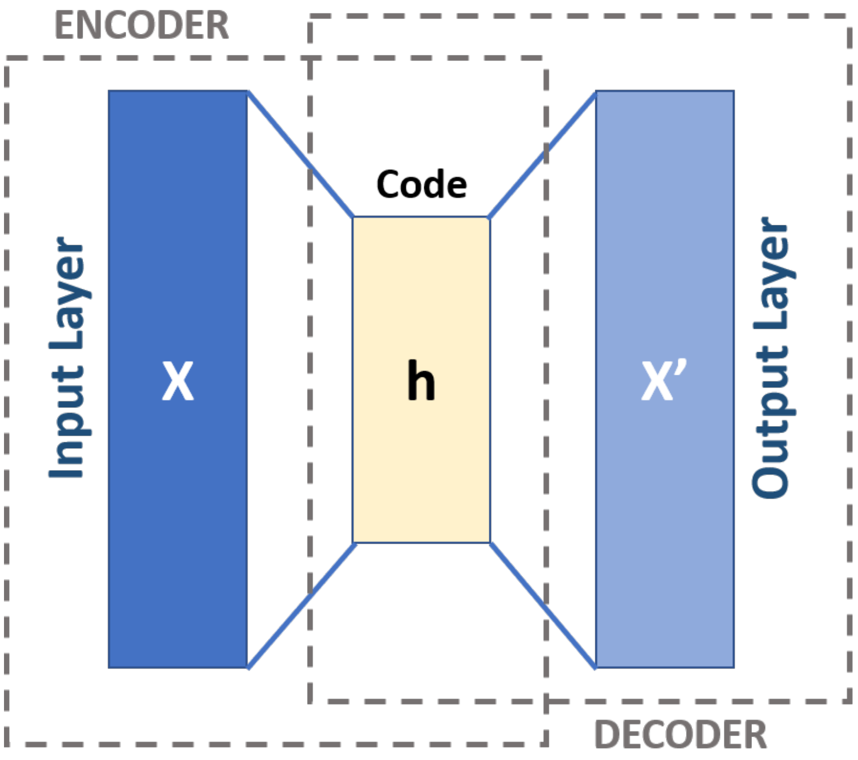
HOCHSCHULE ANSBACH

Autoencoders (AEs)

- A classical image autoencoder takes an image, maps it to a latent vector space, and decodes it back

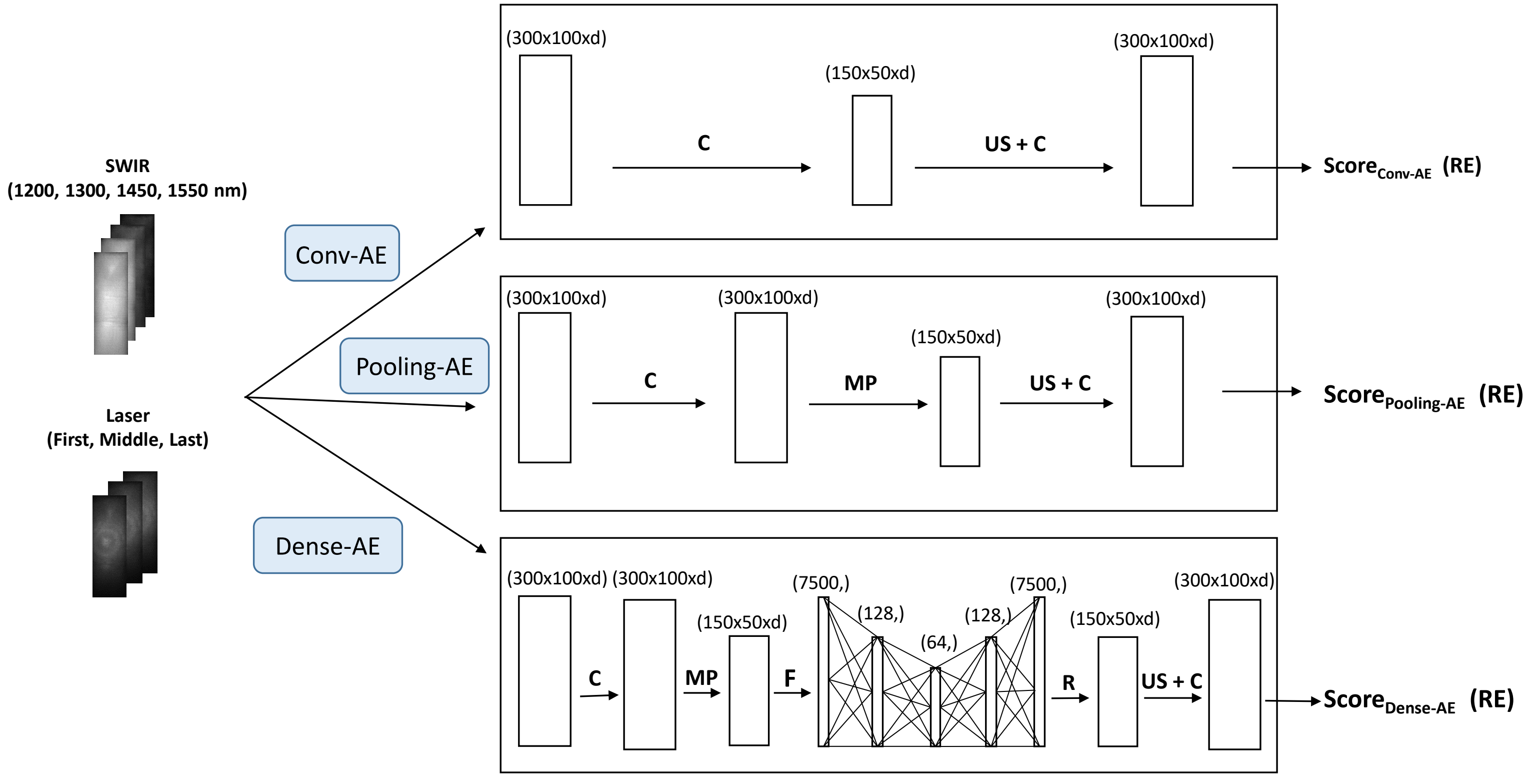
- The autoencoder learns to reconstruct the original inputs

- If we only train it on bona fide data, it will produce “weird” results for presentation attacks on the test set



HOCHSCHULE ANSBACH

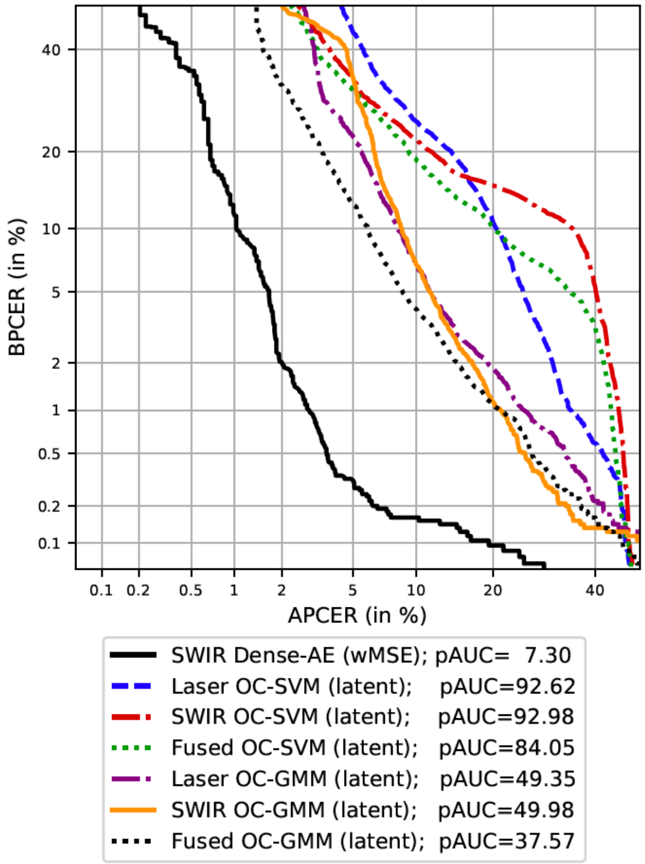
Autoencoders

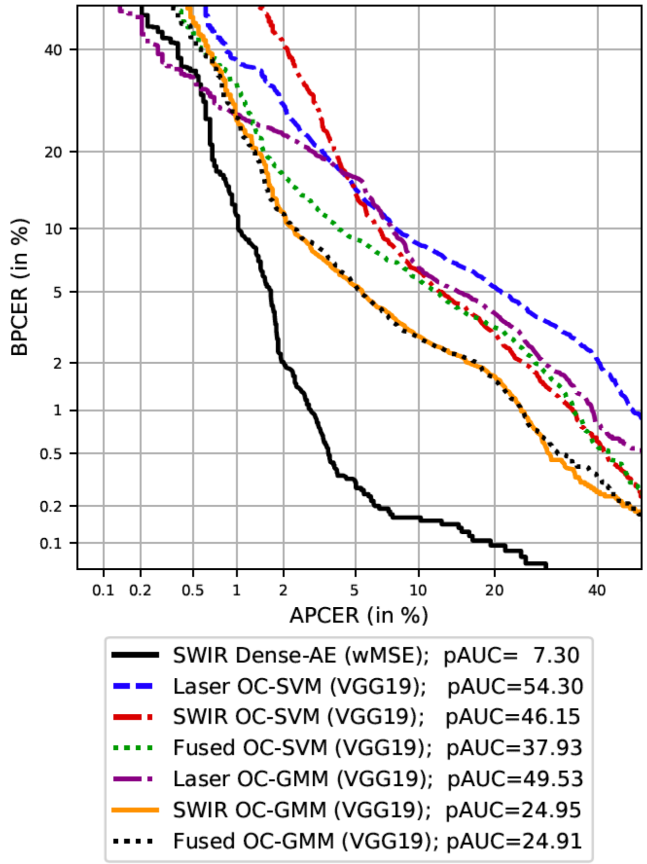


J. Kolberg, M. Grimmer, M. Gomez-Barrero, C. Busch, “Anomaly Detection with Convolutional Autoencoders for Fingerprint Presentation Attack Detection”, in ArXiV:2008.07989, 2020.

Autoencoders

- 19,711 bona fide and 4,339 PA images, 45 different PAI species

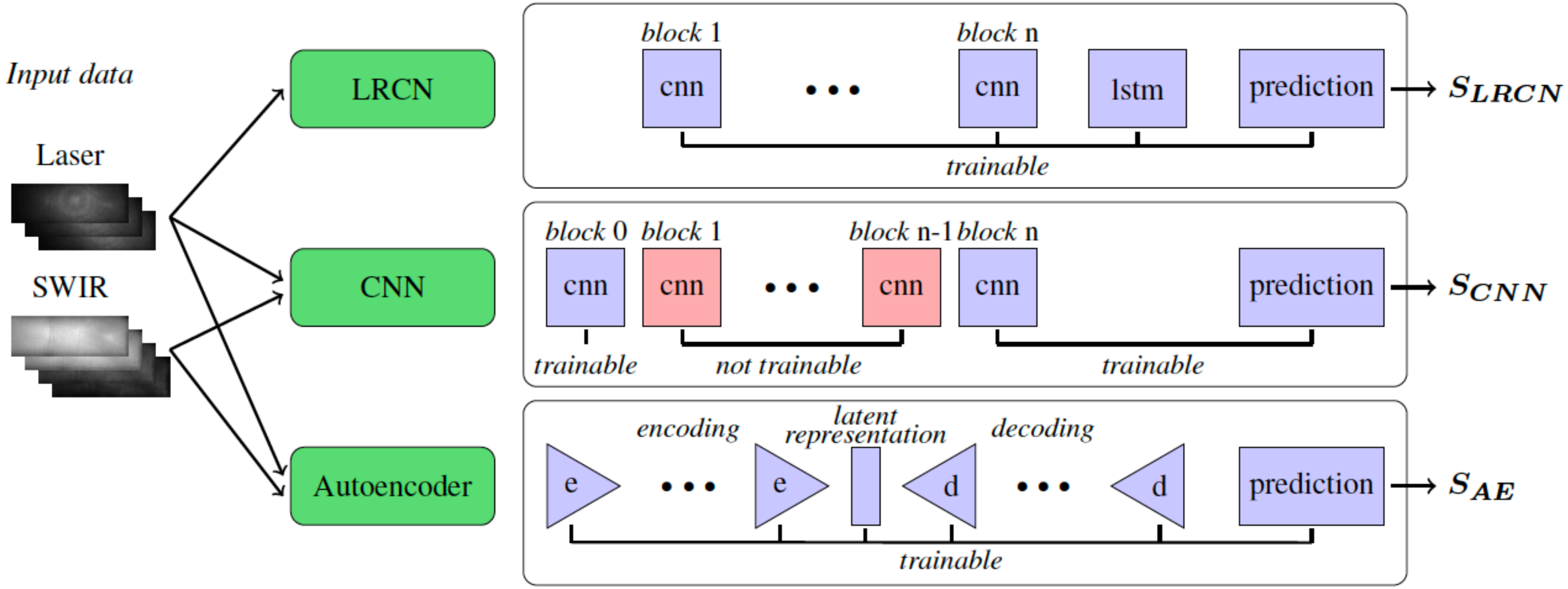




HOCHSCHULE ANSBACH

Autoencoders vs. Two-class CNNs

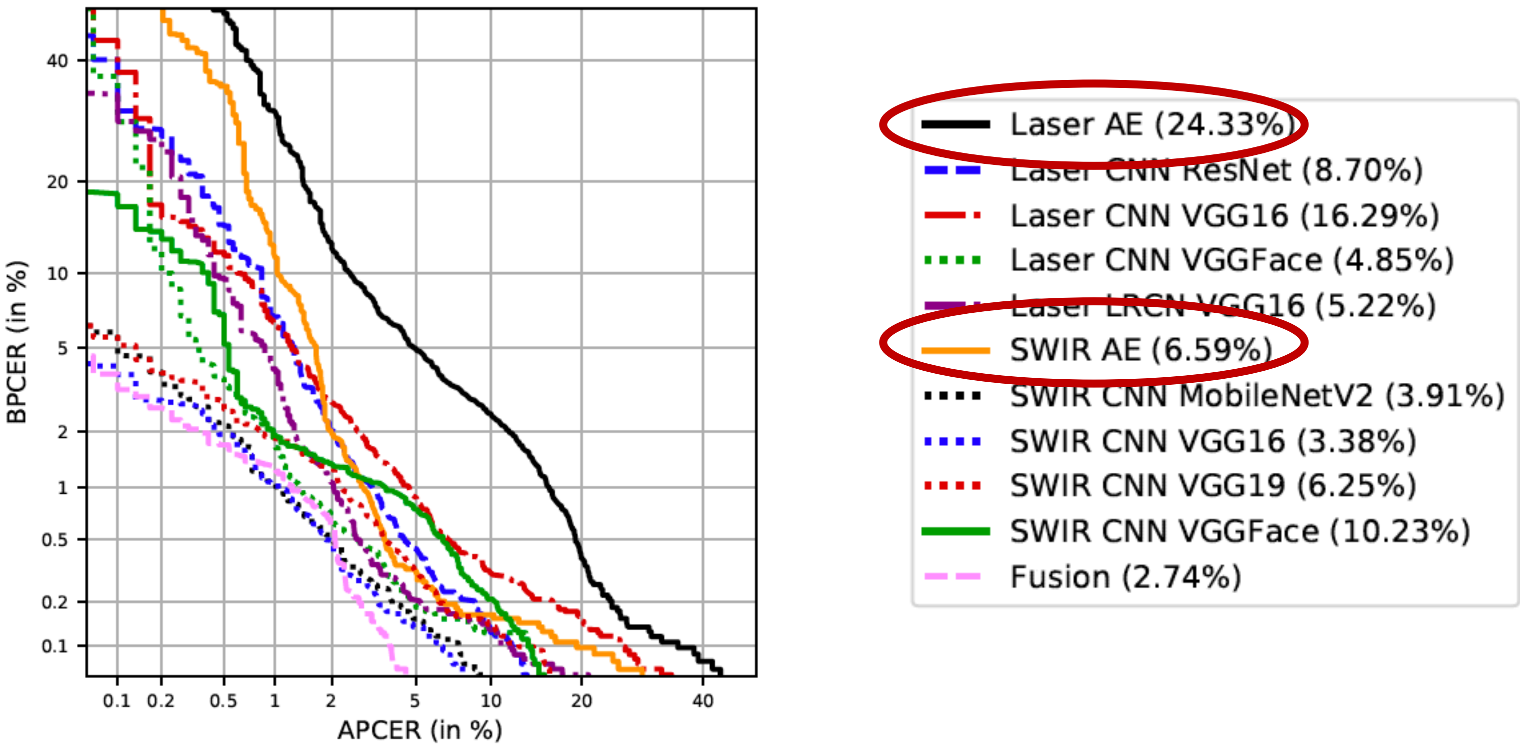
- What happens when we benchmark Autoencoders with two-class classifiers evaluated on a leave-one-out (LOO) protocol?



J. Kolberg, M. Gomez-Barrero, C. Busch, "On the Generalisation Capabilities of Fingerprint Presentation Attack Detection Methods in the Short Wave Infrared Domain", in ArXiX:2010.09566, 2020.

Autoencoders vs. Two-class CNNs

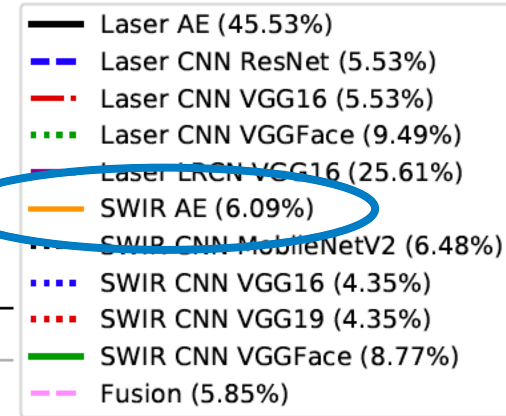
- Baseline training on BFs + PAs (APCER @ BPCER = 0.2%)

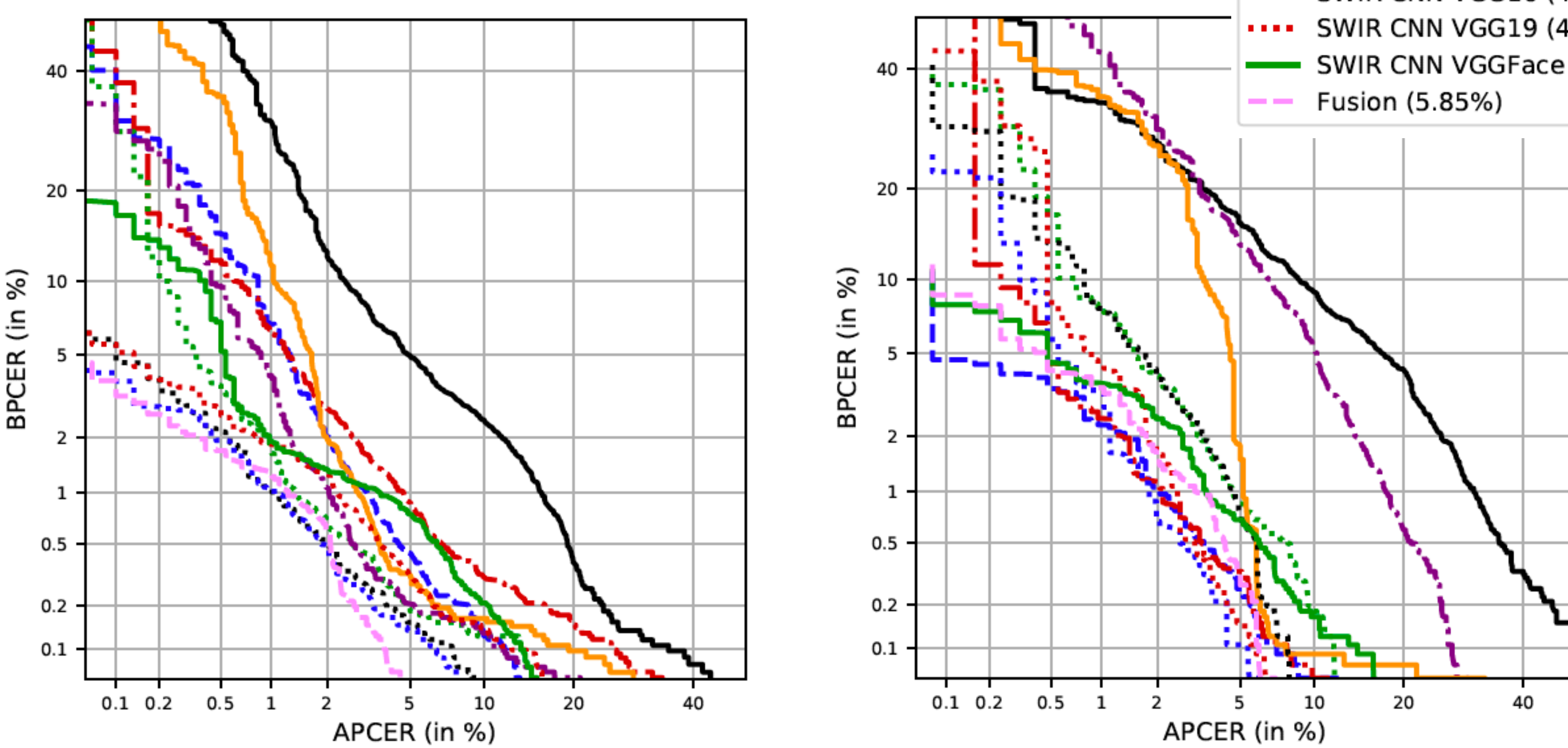


HOCHSCHULE ANSBACH

Autoencoders vs. Two-class CNNs

- LOO per groups: fake fingers



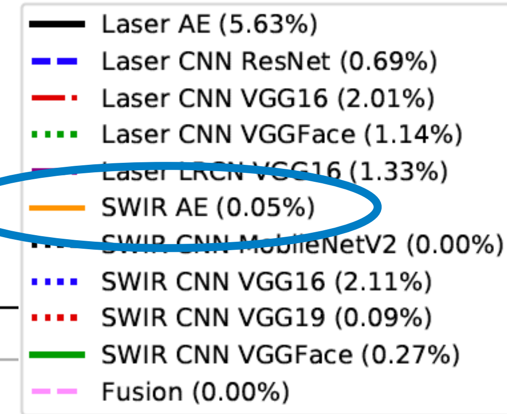


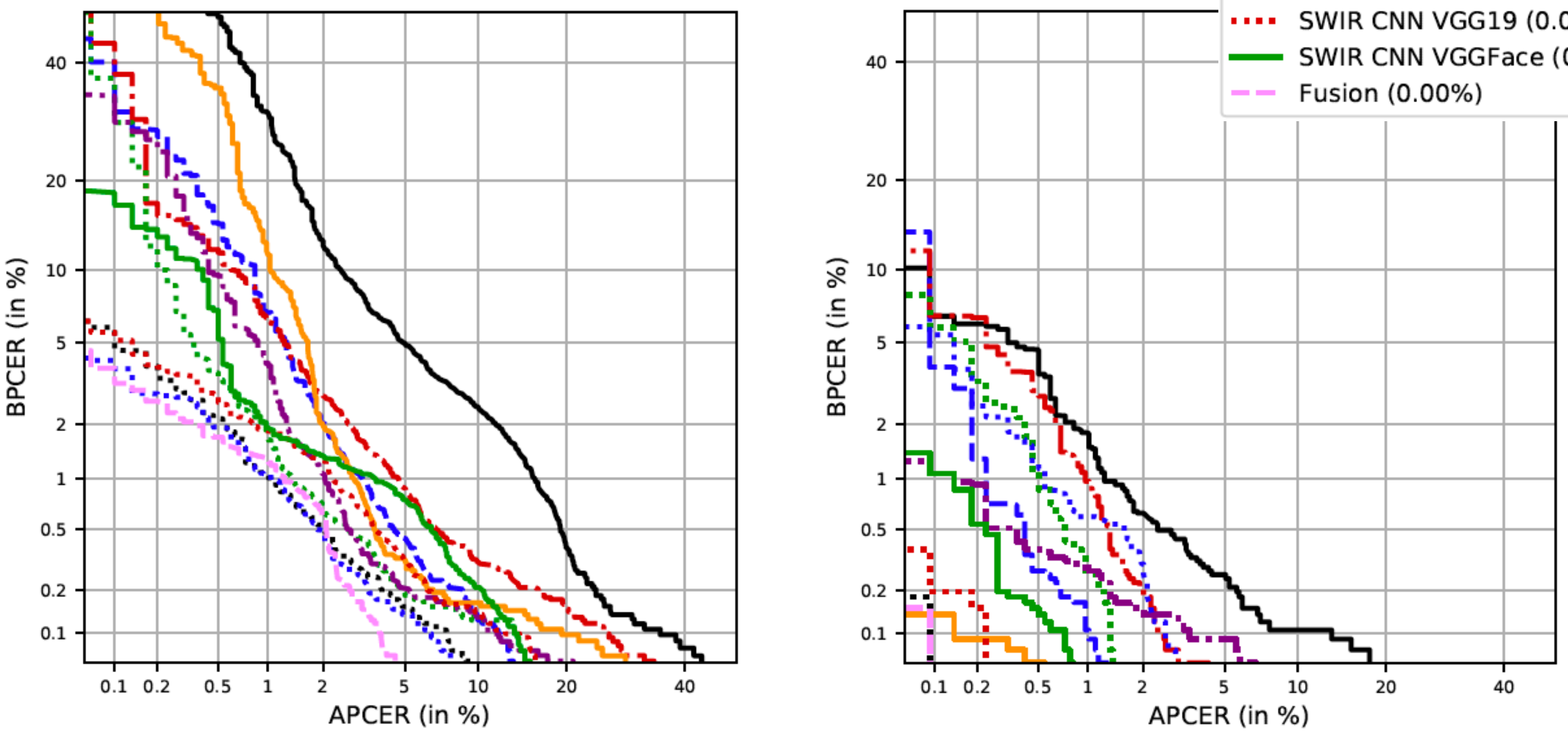
Anomaly detection

HOCHSCHULE ANSBACH

Autoencoders vs. Two-class CNNs

➢ LOO per groups: opaque overlays

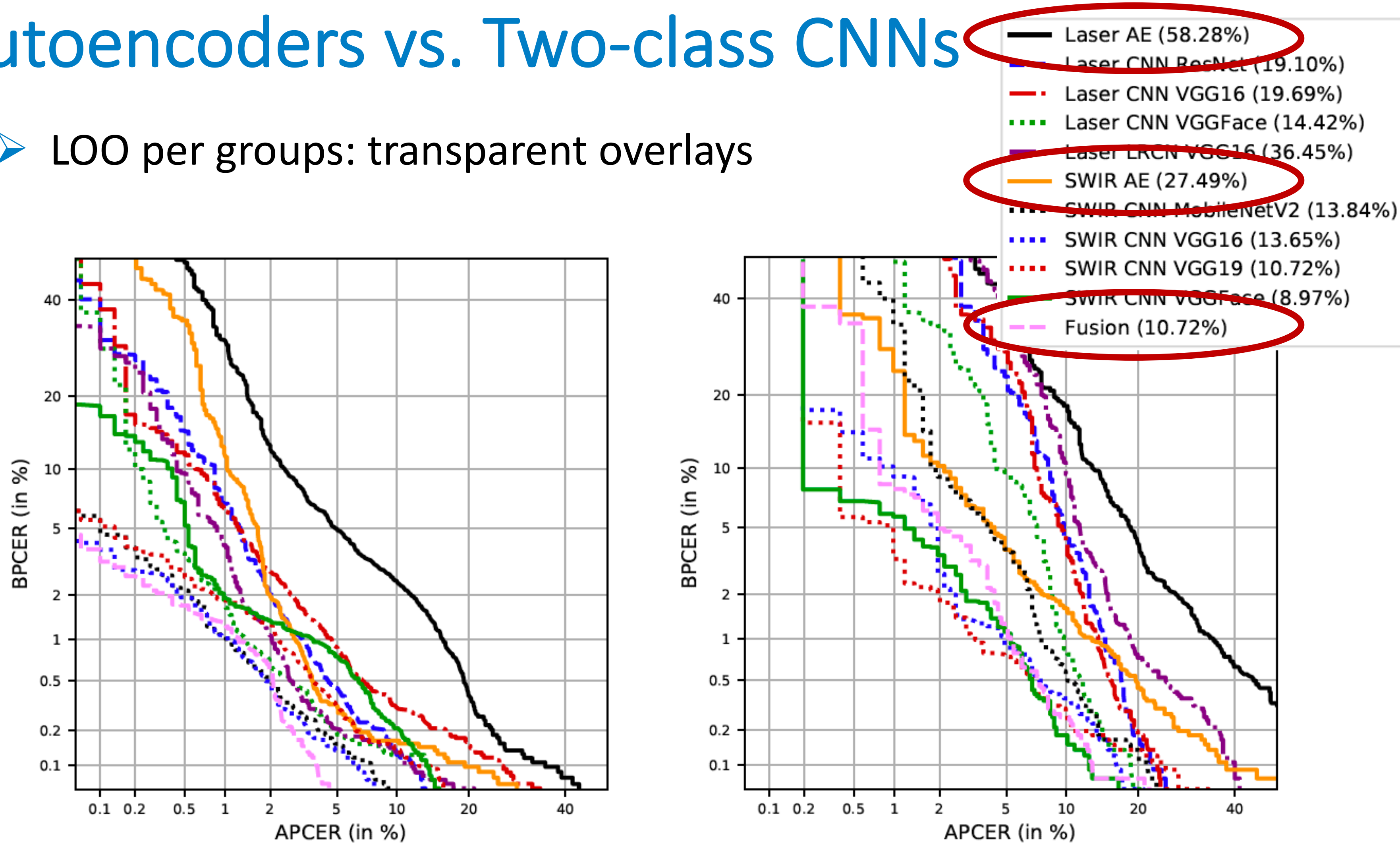




HOCHSCHULE ANSBACH

Autoencoders vs. Two-class CNNs

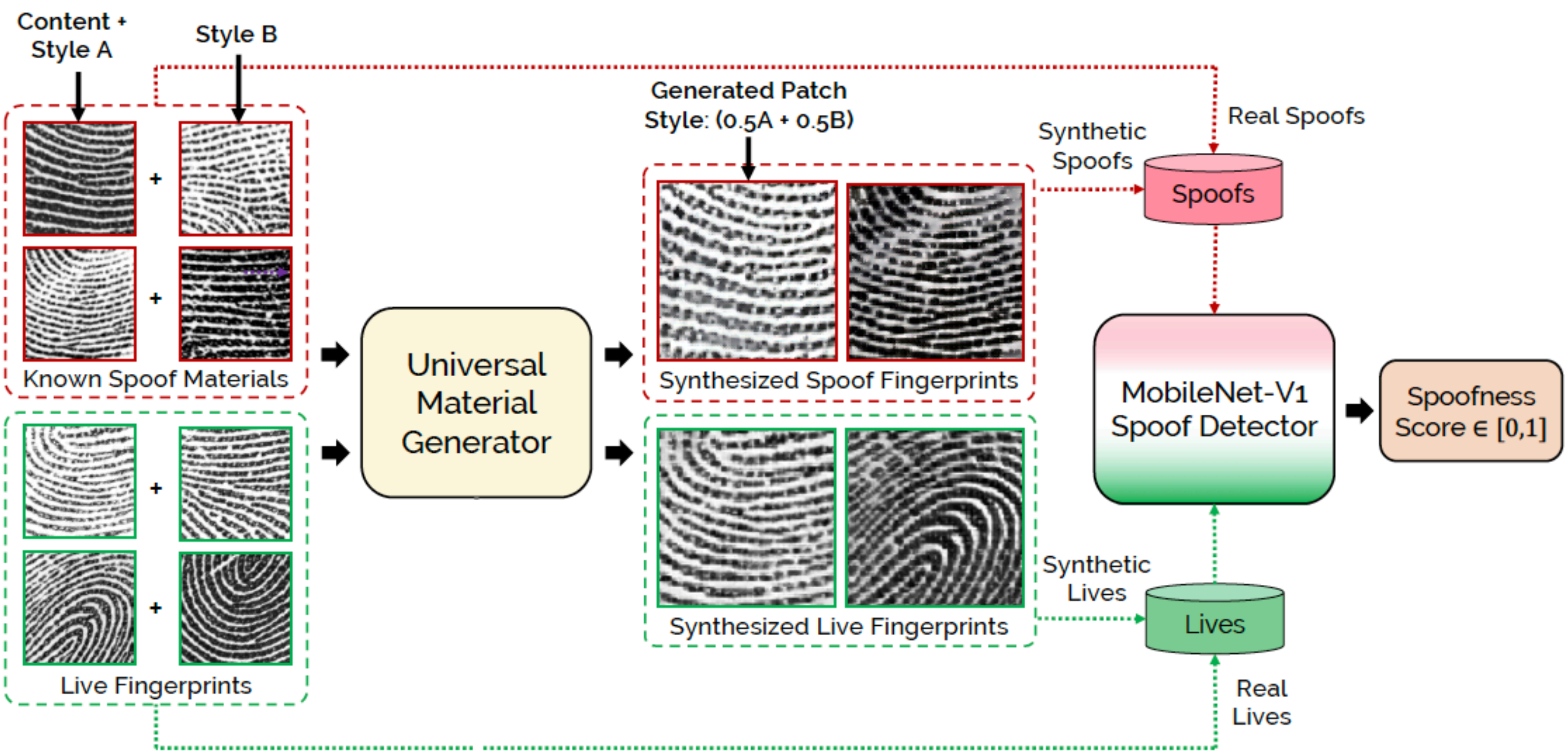
- LOO per groups: transparent overlays



HOCHSCHULE ANSBACH

Style transfer with deep learning

- Goal: generate synthetic PA images of unknown materials, by transferring the style (texture) characteristics from known materials

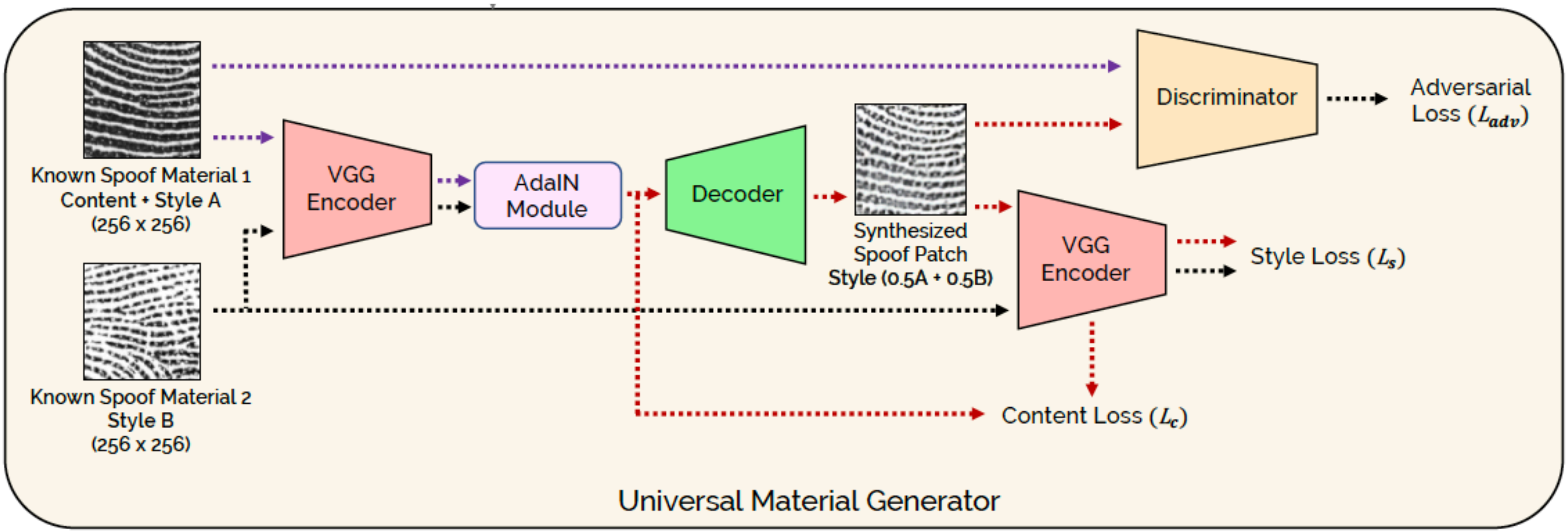


T. Chugh, A. K. Jain, "Fingerprint Spoof Generalization", ArXiV:1912.02710, 2019

HOCHSCHULE ANSBACH

Style transfer with GANs and AEs

- Style transfer is done with an AE construction
- A GAN construction enhances the appearance of the synthetic samples



T. Chugh, A. K. Jain, "Fingerprint Spoof Generalization", ArXiV:1912.02710, 2019

HOCHSCHULE ANSBACH

Style transfer with deep learning

- For fingerprint PAD, APCER @ BPCER = 0.2% on a LOO partition:

	Gelatin	Silicone	Playdoh	Woodglue	Transpa-rency	Gold Finger
CNN	45.05%	**32.28%**	41.58%	13.62%	4.17%	11.78%
CNN + UMG	**2.04%**	**1.36%**	**27.64%**	**1.03%**	**0.00%**	**11.41%**

	Dragonskin	Conductive Ink + Paper	Monster Latex	3D Univ. Targets	Body Latex	2D paper
CNN	2.52%	10.00%	5.23%	5.00%	23.65%	44.56%
CNN + UMG	**0.00%**	**0.00%**	**3.76%**	**0.00%**	**10.28%**	**19.78%**

T. Chugh, A. K. Jain, "Fingerprint Spoof Generalization", ArXiV:1912.02710, 2019

Conclusions

- Good performing PAD methods are still not robust to unknown attacks

- Two approaches to tackle the problem:
  - Use a limited number of PAI species for training in order to simulate an unknown attacks scenario
  - Use anomaly detection approaches

- Main findings:
  - Traditional one-class classifiers cannot reach a good performance in the state of the art (so far)
  - Alternatives based on feature projection, autoencoders, or style transfer do offer good results, but highly dependent on the PAI species

- There is still work to be done!

Marta Gomez-Barrero
(marta.gomez-barrero@hs-ansbach.de)

HOCHSCHULE
ANSBACH