PICUS lab

PATTERN ANALYSIS AND INTELLIGENT COMPUTATION FOR MULTIMEDIA SYSTEMS

Artificial Intelligence and Intelligent Systems
cini National Lab

# Fingerprint Presentation Attacks: Tackling the Ongoing Arms Race in Biometric Authentication

Roberto Casula[2], Antonio Galli[1], Michela Gravina[1], Stefano Marrone[1], Domenico Mattiello[1], Marco Micheletto[2], Giulia Orrù[2], Gian Luca Marcialis[2] and Carlo Sansone[1]

[1] University of Naples Federico II
[2] University of Cagliari

Ital-IA 2023: 3rd National Conference on Artificial Intelligence May 29-31, 2023, Pisa, Italy
Workshop AI per Cybersecurity

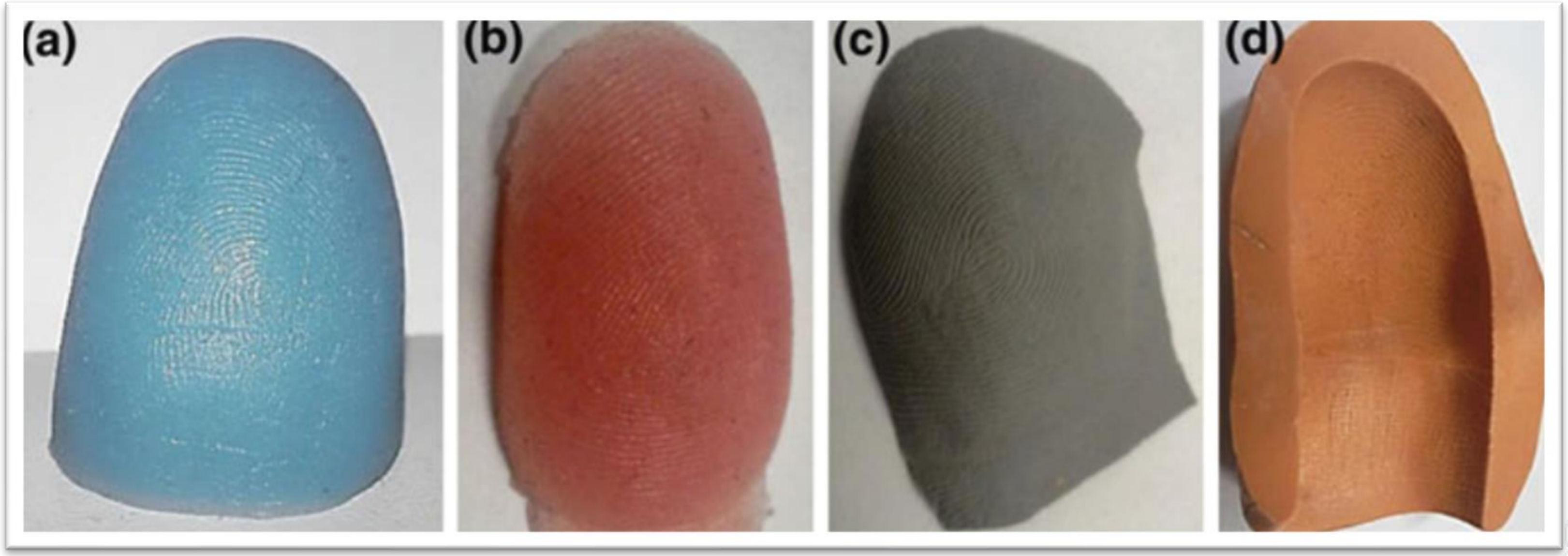
# Biometric Based Authentication Systems

- Biometric-based authentication systems allows to recognize subjects according to “what they are” rather then “what they use”
- Among all, fingerprints represent one of the easiest to implement and the most accepted by end-users

# Presentation Attacks

- Subject authentication based on fingerprints is widely adopted in public security systems (e.g. banks) as well as on personal devices

- Presentation attacks: procedures aimed at bypassing an Automated Fingerprint Identification System (AFIS) by using an artificial fingerprint replica
  - Artificial fingerprints can be crafted using different materials



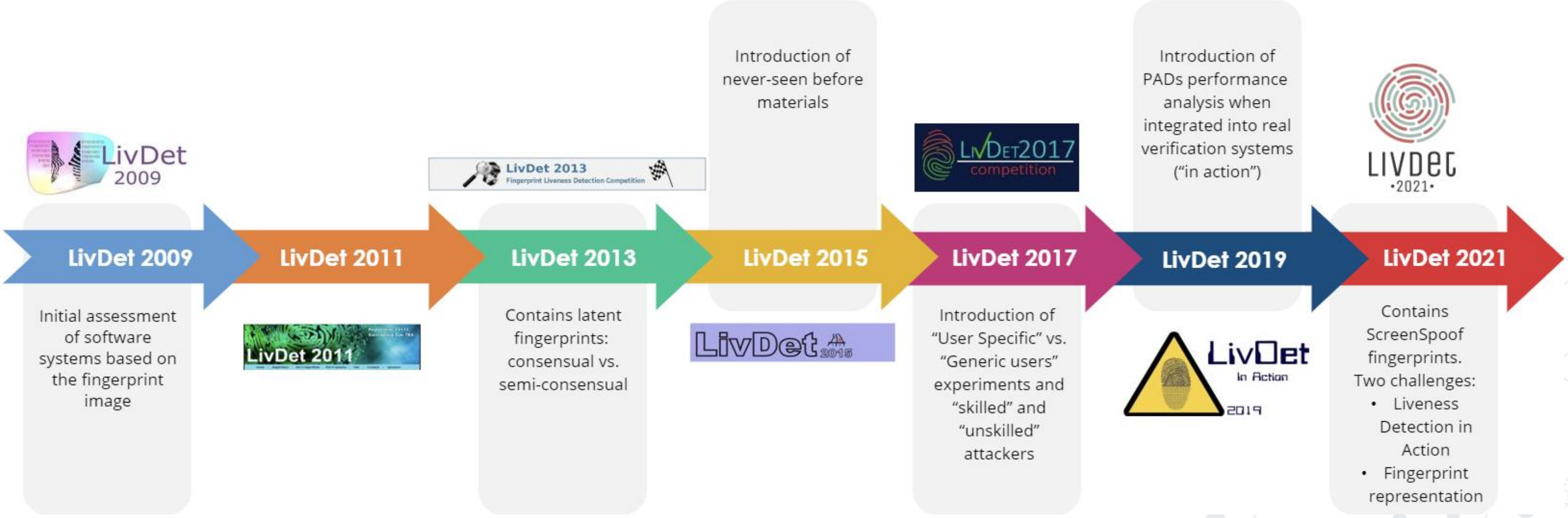
Artificial finger replicas made using GLS (a), Ecoflex (b), Liquid Ecoflex (c) and Modasil (d). Images taken from Bhanu et al.[1]

[1] Bhanu et al. “Deep learning for biometrics”, Springer (2017)

# Fingerprint Liveness Detection Competition

- The Liveness Detection (LivDet) Competition is a biennial competition in which participants are challenged to identify spoofs from live samples

- Each edition has its own distinctive set of challenges that competitors must overcome:
  - ✓ the presence of different materials for the training and test sets
  - ✓ the integration of FPADs into AFIS.



LivDet 2009

Initial assessment of software systems based on the fingerprint image

LivDet 2011

LivDet 2013

Contains latent fingerprints: consensual vs. semi-consensual

LivDet 2015

Introduction of never-seen before materials

LivDet 2017

Introduction of "User Specific" vs. "Generic users" experiments and "skilled" and "unskilled" attackers

LivDet 2019

Introduction of PADs performance analysis when integrated into real verification systems ("in action")

LivDet 2021

Contains ScreenSpoof fingerprints.
Two challenges:
- Liveness Detection in Action
- Fingerprint representation

# Challenges to be Faced

- The introduction of a new spoof fabrication technique (ScreenSpoof) highlighted the ongoing vulnerability of modern FPADs to never-seen attacks:
  - attacks unknown in the training phase of the classification model

- Another key point in the design of a reliable FPAD is considering its integration with an AFIS:
  - the evaluation of integrated systems has also been introduced

- The LivDet competition is therefore based on the concept that to design a robust and efficient FPAD system, both the defender's and attacker's perspectives must be considered:

😈

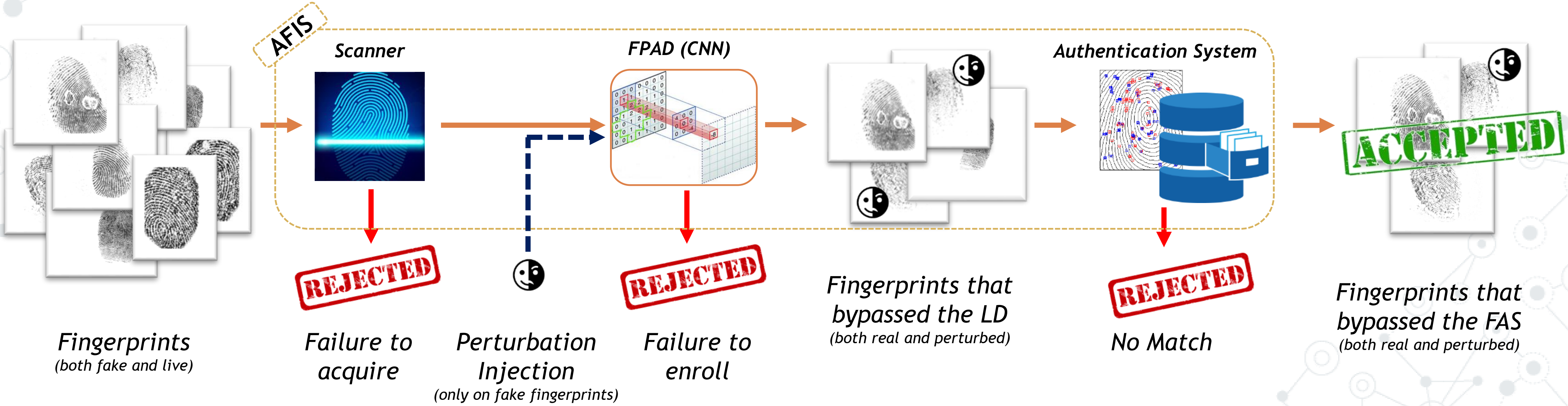
- the organizers put themselves in the shoes of the attackers, by simulating real-world attacker scenarios.

😇

- Participants are expected to develop robust systems

# Fingerprint Liveness Detector

- AFIS are often equipped with a Fingerprint Liveness Detection (FLD) module, often also referred to as Fingerprint Presentation Attack Detection (PAD or FPAD), to recognize real fingerprints from spoof replicas
- FPAD can both rely on external sensors or be based only on image processing techniques
    - In the latter case, as for several other computer vision domains, Artificial Intelligence (AI) based approaches represent the current state-of-the-art



Fingerprints
(both fake and live)

Failure to acquire

Perturbation Injection
(only on fake fingerprints)

Failure to enroll

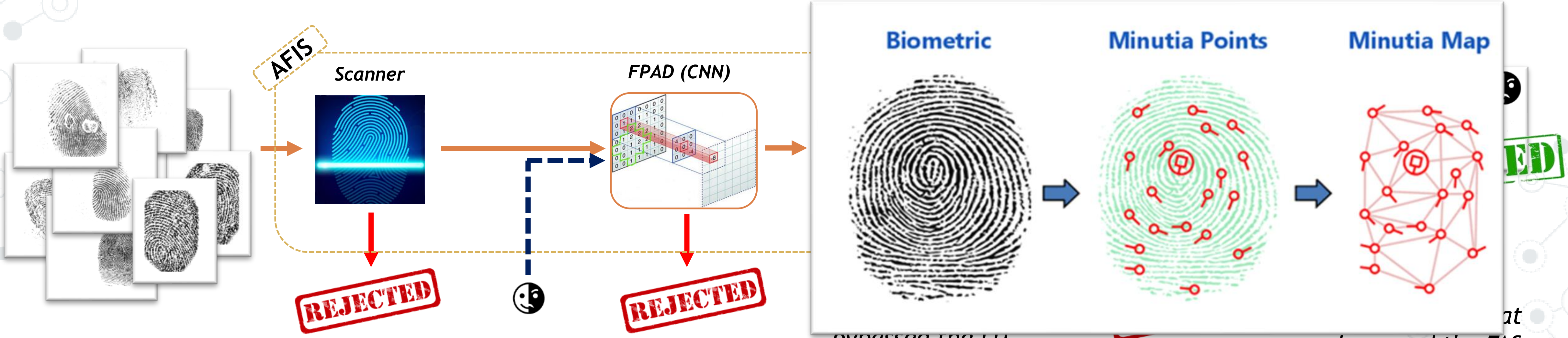
Fingerprints that bypassed the LD
(both real and perturbed)

No Match

Fingerprints that bypassed the FAS
(both real and perturbed)

# Fingerprint Liveness Detector

- AFIS are often equipped with a Fingerprint Liveness Detection (FLD) module, often also referred to as Fingerprint Presentation Attack Detection (PAD or FPAD), to recognize real fingerprints from spoof replicas
- FPAD can both rely on external sensors or be based only on image processing techniques
    - In the latter case, as for several other computer vision domains, Artificial Intelligence (AI) based approaches represent the current state-of-the-art



*Fingerprints*
*(both fake and live)*

*Failure to acquire*

*Perturbation Injection*
*(only on fake fingerprints)*

*Failure to enroll*

*bypassed the LD*
*(both real and perturbed)*

*No Match*

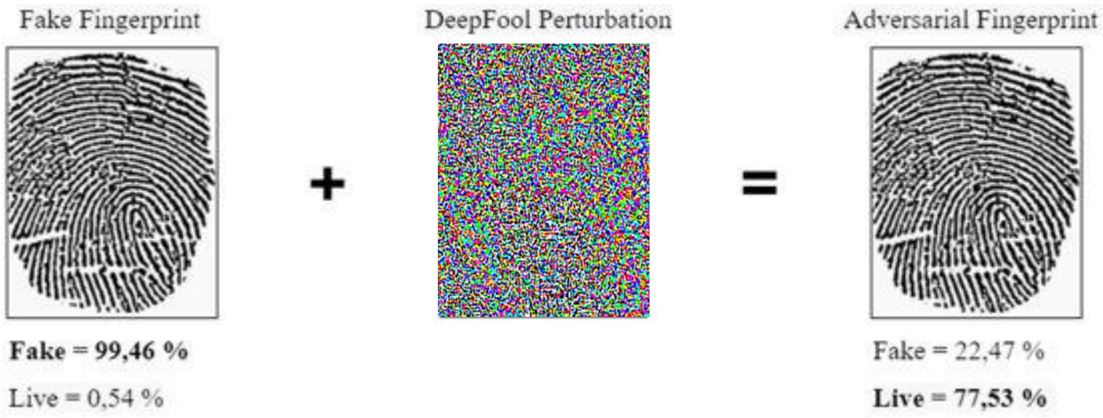
*at*
*bypassed the FAS*
*(both real and perturbed)*

# The two Sides of AI in FPAD

- FPAD has seen the rise and establishment of Convolutional Neural Networks (CNNs) as an effective approach to the problem:
  - ✓ CNNs obtain state-of-the-art performance in identifying a large number of fake samples

- The development of counter-anti-spoofing techniques:
  - ✓ approaches aimed at bypassing an AFIS despite being protected by an FPAD

- Adversarial fingerprints:
  - ✓ set of attacks designed to circumvent an FPAD by exploiting adversarial perturbations
  - ✓ Adversarial perturbation: a set of algorithms designed to mislead a target CNN by means of a specifically crafted noise



# Adversarial Liveness Detector (ALD)

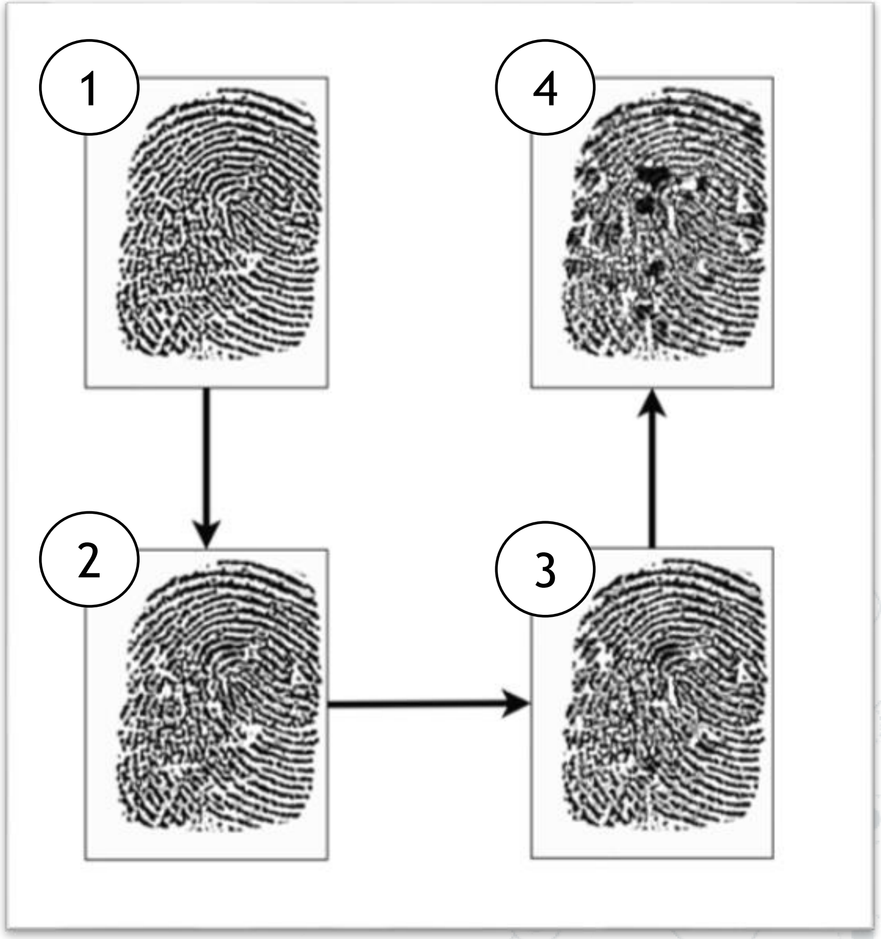
- ALD represents a deep learning-based fingerprint liveness detection whose aim is to exploit the experience matured as attackers to design an ad-hoc adversarial data augmentation strategy intended to increase the effectiveness of CNN-based presentation attack detection

😈 + 😇 = ALD

ALD has been submitted to the LivDet 2021 competition and obtained the first place out of 23 participants in the “Liveness Detection in Action track”.

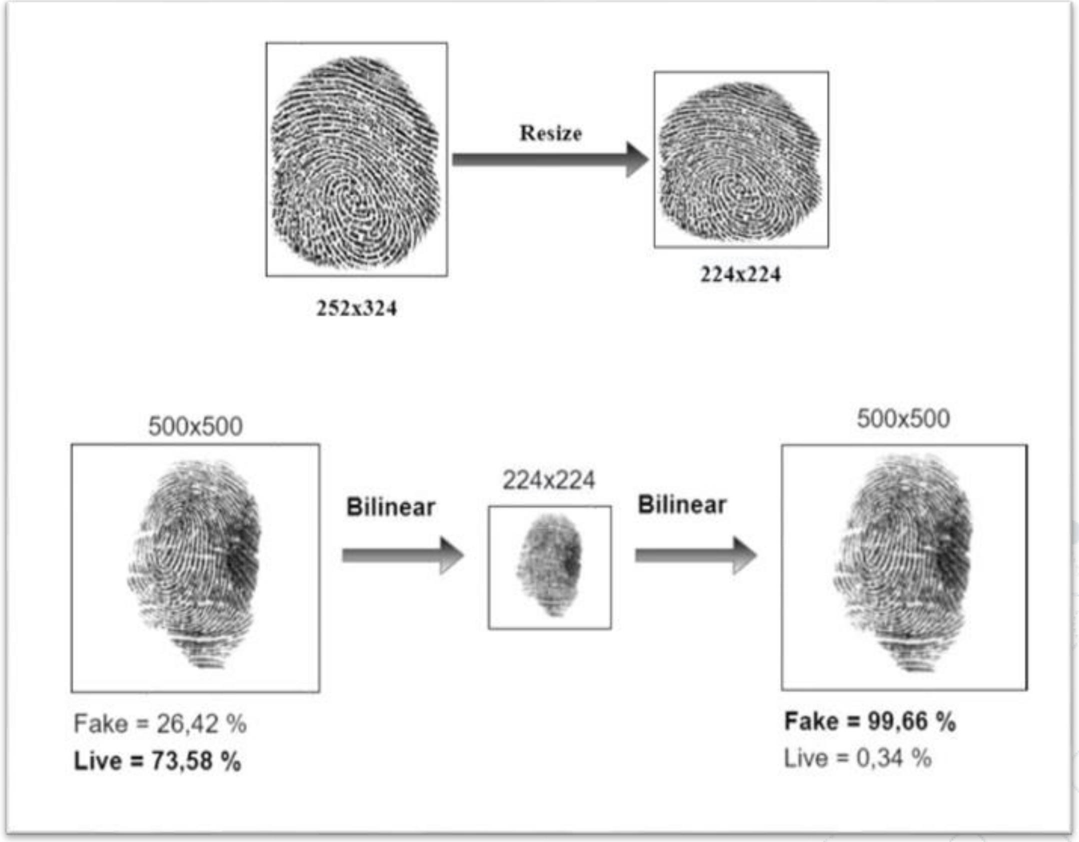
# Effects of Adversarial Samples for Training

- The idea is to realize an iterative training procedure consisting of three main steps:
    - Train a CNN-based Liveness Detector (LD) on the available data
    - Generate adversarial fingerprints (DeepFool) using the trained LD as target CNN
    - Add these new fingerprints to the pool of training data
- Despite this “adversarial data augmentation” schema sounds legit, fingerprints are different from natural images:
    - It turned out that the adversarial attack success rate quickly decreases as the described training iteration increases
    - As a consequence, only the first iteration allows the adversarially trained CNN to increase its generalization ability
- The reason is that the so trained LD tends to be very robust against the attack, causing the adversarial algorithm to “destroy” the fingerprints



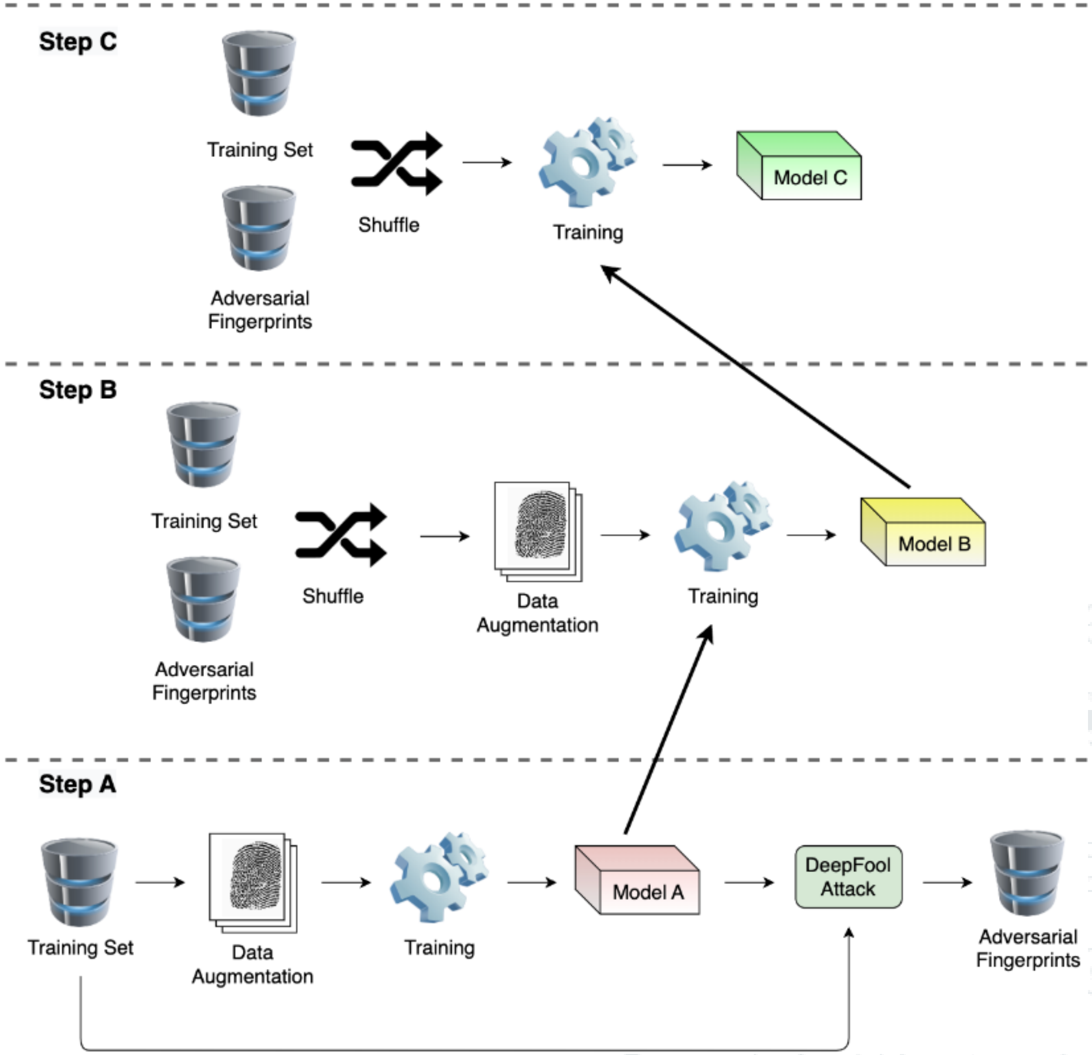
# Effects of Image Resize

- Using a CNN also makes harder to deal with the size of the acquired fingerprint:
  - Different scanners acquire fingerprint at a different resolution, not necessarily resulting in “square” images
  - CNNs (almost always) expect a square image as input, having sizes (commonly) smaller than the acquired fingerprint
  - Image resize is not an option, as it tends to destroy details in the acquired fingerprint



# Result Training Schema

- The result is a three-steps training schema, where each step fine-tunes the model obtained in the previous one:
  - In the first step (A), the model (pre-trained on ImageNet) is fine-tuned on the challenge data. Data augmentation is used, limited to algorithms operating only on pixels intensity values (saturation, shading, etc.). Once the model is trained, the DeepFool algorithm is used to create a new dataset of adversarially perturbed fingerprint
  - In the second step (B), the model is further fine-tuned by using the new dataset consisting of both original and perturbed fingerprints. The same set of data augmentation algorithms has been used
  - In the third step (C), the model is fine-tuned for the last time by using the new dataset consisting of both original and perturbed fingerprints

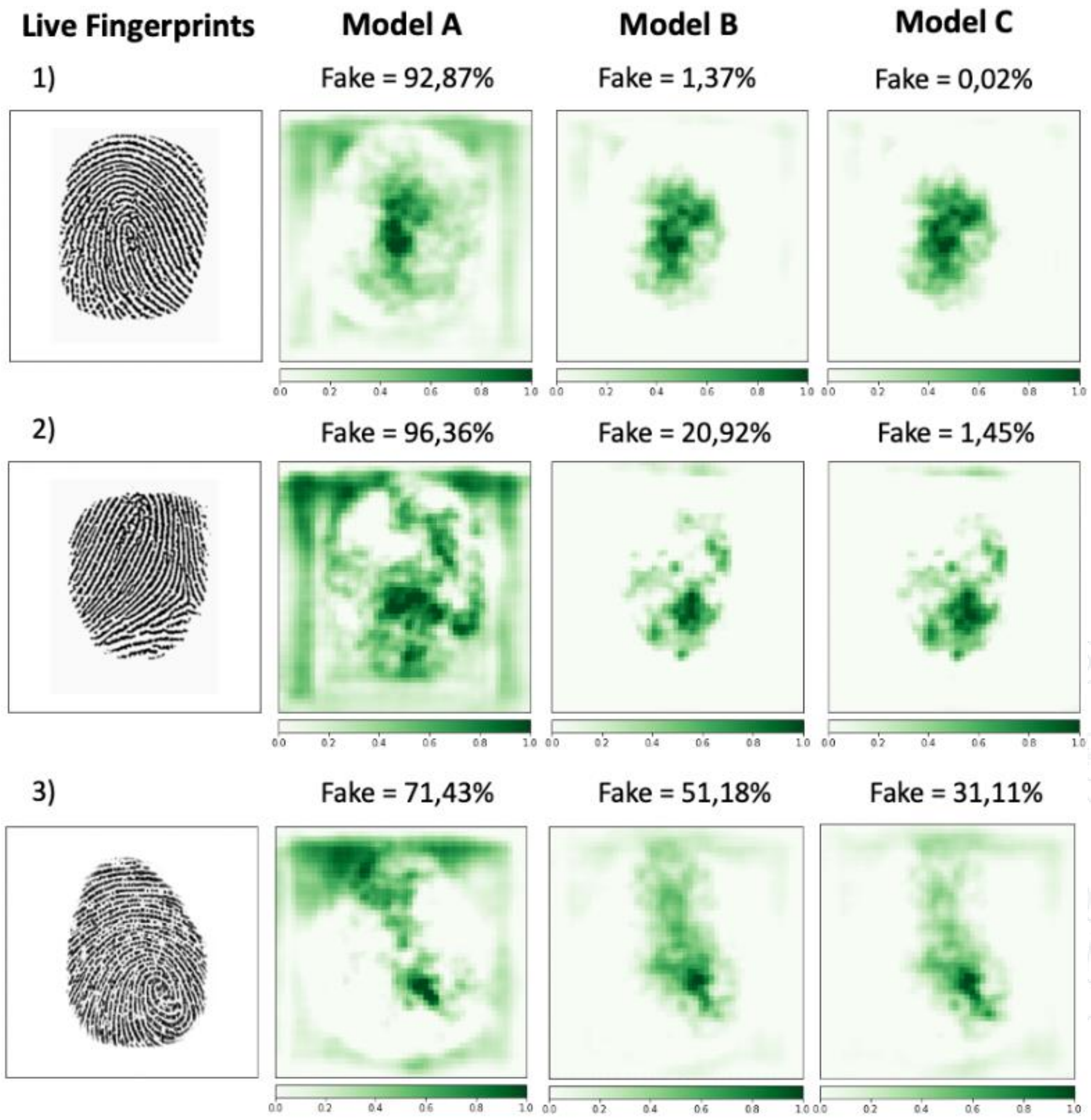


# Explainability Analysis

- Finally, with the aim of further analyzing the effects of the proposed approach, we performed an explainability analysis for each of the three steps

- The Occlusion method is used to highlight the most portion of the image that causes the liveness detector to classify the fingerprint as fake

- It can be noted that using adversarial fingerprints (models B and C) causes the model to mostly focus on the fingerprint inner region, almost ignoring borders and background

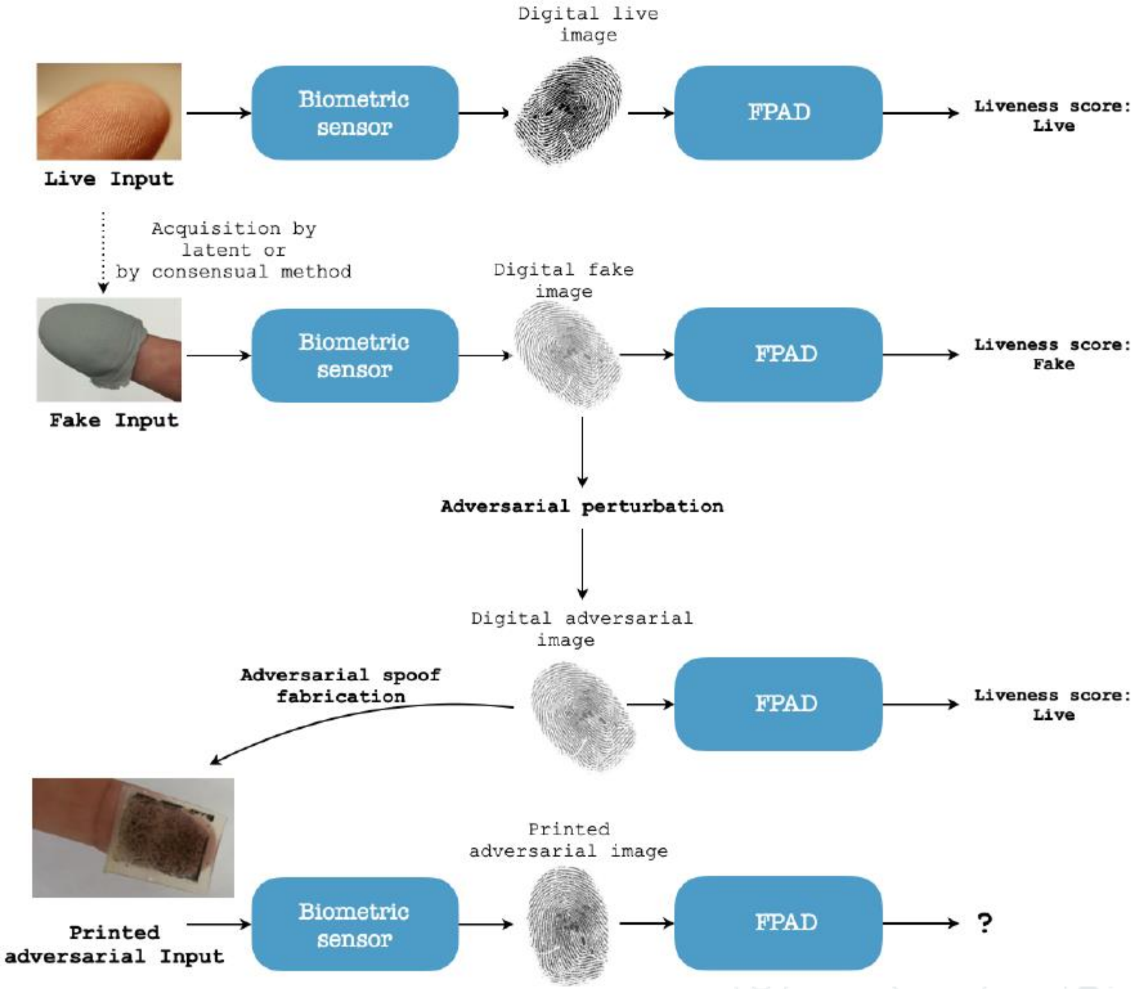


# Fingerprint Adversarial Presentation Attack

- Digital adversarial attacks have proven effective against modern AFIS using adversarial perturbations:
  - small changes added to the fingerprint image designed to mislead the system

- We explored the threat level of physical adversarial attacks in a realistic scenario where attackers must create a physical replica:
  - Starting from the image of a fake fingerprint it is possible to inject noise to obtain an adversarial image considered live by a classifier
  - The presentation attack obtained by printing the digital adversarial image using a standard laser

- We evaluated the percentage of successful fingerprint adversarial presentation attacks on both white-box and black-box systems



# Fingerprint Adversarial Presentation Attack

- Digital adv
  against mo
  perturbatio
  - small c
    designe
- We explore
  attacks in a
  create a phy
  - Starting
    possible
    conside
  - The pre
    digital a
- We evaluat
  fingerprint
  white-box a

Liveness score:
Live

Liveness score:
Fake

Liveness score:
Live

?

# Conclusions

- The availability of a public available dataset for fingerprint liveness detection (LivDet), serving as a known and shared benchmark evaluated under the same experimental conditions, is strongly pushing the researcher in exploring new approaches and solutions for FPAD
- ALD highlighted the significant contribution of adversarial perturbation techniques to the generalization capacity of the CNNs considered as FPAD
- At the same time, adversarial perturbations have proved to be an effective strategy for the generation of fake fingerprint in the real world
- The experience in the international competition LivDet as organizers (University of Cagliari), and as participants (University of Naples Federico II) has allowed to highlight that the dual approach that considers the two points of view during the design of FPADs and their integration into AFIS is crucial

# Questions?