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, Naples, Italy*
[2]*University of Cagliari, Cagliari, Italy*

Abstract
The widespread use of Automated Fingerprint Identification Systems (AFIS) in consumer electronics opens for the development of advanced presentation attacks, i.e. procedures designed to bypass an AFIS using a forged fingerprint. As a consequence, AFIS are often equipped with a fingerprint presentation attack detection (FPAD) module, to recognize live fingerprints from fake replicas, in order to both minimize the risk of unauthorized access and avoid pointless computations. The ongoing arms race between attackers and detector designers demands a comprehensive understanding of both the defender's and attacker's perspectives to develop robust and efficient FPAD systems. This paper proposes a dual-perspective approach to FPAD, which encompasses the presentation of a new technique for carrying out presentation attacks starting from perturbed samples with adversarial techniques and the presentation of a new detection technique based on an adversarial data augmentation strategy. In this case, attack and defence are based on the same assumptions demonstrating that this dual research approach can be exploited to enhance the overall security of fingerprint recognition systems against spoofing attacks.

Keywords
Fingerprint, Presentation Attack Detection, Convolutional Neural Networks, Adversarial Perturbation, Data Augmentation

1. Introduction

Recent consumer electronics, such as smartphones, laptops, etc., are more and more focused on protecting users' privacy by enforcing access only to authorized subjects. This is resulting in the development and integration of authentication devices, often exploiting users' biometrics in order to identify them based on who they are rather than on what they carry. Among all, subject authentication based on fingerprints is widely adopted in public security systems (e.g. banks) as well as on personal devices, thanks to its low invasiveness and high precision. The flip side of this massive and long-time usage is in the broader range of opportunities an attacker has to develop more ingenious presentation attacks, namely procedures aimed at bypassing an Automated Fingerprint Identification System (AFIS) by using an artificial fingerprint replica (spoof or presentation attack)[1].

It is extremely important to identify a fake fingerprint as early as possible in the AFIS processing pipeline in order to both minimize the risk of unauthorized access and to avoid pointless computations [2]. To this aim, in recent years, 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 (i.e., coming from a "live" finger) from spoof replicas. Artificial fingerprints can be crafted using different materials, including very common ones, such as latex and wood glue. Liveness detection can exploit external sensors (e.g. temperature) or rely only on the acquired image. The latter solution is often preferred, as an image processing approach suits the FPAD for several devices and scenarios.

As for other image processing tasks, presentation attack detection has seen the rise and establishment of Machine Learning (ML) as an effective approach to the problem, with deep Convolutional Neural Networks (CNNs) obtaining state-of-the-art performance in identifying a large number of fake samples [3, 4]. An interesting side effect resulting from the availability of CNN-based FPAD and of a large number of fake and real samples is the development of counter-anti-spoofing techniques, namely approaches aimed at bypassing an AFIS despite being protected by an FPAD. On this line, in previous works [5, 6], we introduced the concept of *adversarial fingerprints*, i.e., the first time ever set of attacks designed to circumvent an FPAD by exploiting adversarial perturbations, namely

Ital-IA 2023: 3rd National Conference on Artificial Intelligence, organized by CINI, May 29–31, 2023, Pisa, Italy
*Corresponding author.
roberto.casula@unica.it (R. Casula); antonio.galli@unina.it (A. Galli); michela.gravina@unina.it (M. Gravina); stefano.marrone@unina.it (S. Marrone); do.mattiello@studenti.unina.it (D. Mattiello); marco.micheletto@unica.it (M. Micheletto); giulia.orru@unica.it (G. Orrù); marcialis@unica.it (G. L. Marcialis); carlo.sansone@unina.it (C. Sansone)
0000-0003-3810-5935 (R. Casula); 0000-0001-9911-1517 (A. Galli); 0000-0001-5033-9617 (M. Gravina); 0000-0001-6852-0377 (S. Marrone); 0000-0003-1196-7869 (M. Micheletto); 0000-0002-7802-2483 (G. Orrù); 0000-0002-8719-9643 (G. L. Marcialis); 0000-0002-8176-6950 (C. Sansone)

CEUR Workshop Proceedings (CEUR-WS.org)

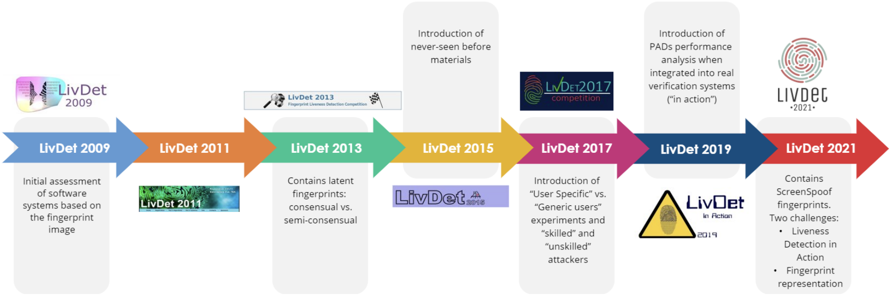
LivDet 2009
Initial assessment of software systems based on the fingerprint image
LivDet 2011
LivDet 2013
Fingerprint Liveness Detection Competition
Contains latent fingerprints: consensual vs. semi-consensual
LivDet 2015
Introduction of never-seen before materials
LivDet 2017
competition
Introduction 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 In Action 2019
LivDet 2021
Contains ScreenSpoof fingerprints.
Two challenges:
• Liveness Detection in Action
• Fingerprint representation

Figure 1: Challenges introduced in each edition of the international competition LivDet.

a set of algorithms designed to mislead a target CNN by means of a specifically crafted noise. Using the same principle, in [7], we introduced *ALD* (Adversarial Liveness Detector), whose core idea is to exploit adversarial fingerprint [5, 6] as a way to perform data augmentation in order to increase the model generalization ability.

2. Fingerprint Liveness Detection Competition

To support the research and development of increasingly sophisticated presentation attack detectors on a common experimental protocol, in 2009 the first Liveness Detection (LivDet) Competition[1] [8] was started through the collaboration of the University of Cagliari and Clarkson University. LivDet is a biennial competition in which participants from both academia and industry are challenged to identify spoofs from live samples [9]. Each edition has its own distinctive set of challenges that competitors must overcome, such as the presence of different materials for the training and test sets (never-seen-before materials) and the integration of FPADs into AFIS. These challenges have highlighted the arms race nature of fingerprint presentation attack detection. For example, a new spoof fabrication technique, called ScreenSpoof [10], was introduced in LivDet2021, which highlighted the ongoing vulnerability of modern FPADs to never-before-seen-before attacks, i.e. attacks unknown in the training phase of the classification model (Figure 2). 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, allowing the competitors to assess the effectiveness of the presented algorithms by simulating real-world attacker scenarios. Another key point in the design of a reliable FPAD is considering its integration with an AFIS.

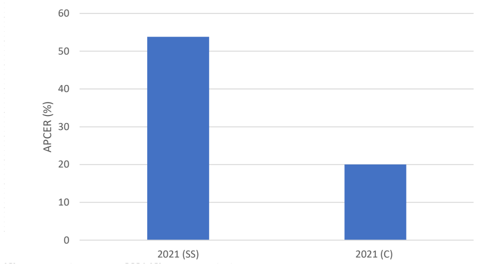

Figure 2: Comparison between mean APCER on the consensual test set (C) and ScreenSpoof test set (SS) for LivDet2021.

For this purpose, starting from LivDet 2019, the evaluation of integrated systems has also been introduced. This is a critical step since anti-spoofing algorithms are not expected to work independently, and the integration may significantly influence the recognition system's performance[2] [11]. In this respect, the LivDet competition is crucial in identifying different algorithms' strengths and weaknesses and guiding the development of more robust and efficient integrated systems. Designers can then use the knowledge resulting from each edition to improve their solutions.

3. Fingerprint Adversarial Presentation Attack in the Physical Domain

Digital adversarial attacks have proven effective against modern AFISs, even when protected with an FPAD module. In particular, this type aims to deceive the AFIS/FPAD module using adversarial perturbations, i.e. small

[1] https://livdet.diee.unica.it/

[2] https://livdet.pythonanywhere.com/

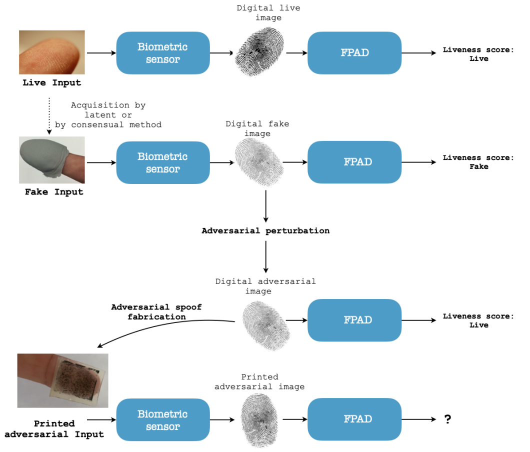

Figure 3: Adversarial presentation attack schema.

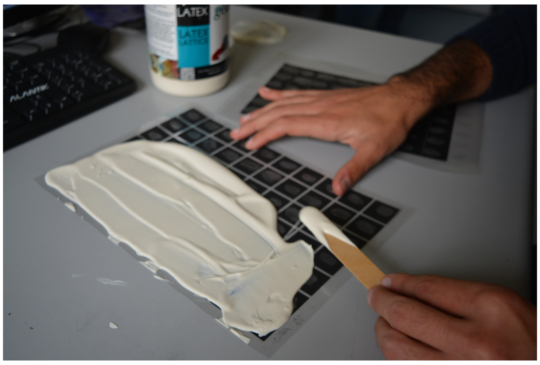

Figure 4: Example of the adversarial fingerprint physical spoof realisation. The expert is depositing a layer of latex over the printed adversarial fingerprints.

changes added to the fingerprint image designed to mislead the system without being visually noticeable to a human observer. However, these digital attacks typically assume access to the internal modules of the system, making them unrealistic. For this reason, we explored the threat level of physical adversarial attacks in a realistic scenario where attackers cannot directly feed a digitally perturbed image to the FPAD and they have to create a physical replica to breach the system through the sensor [5]. Figure 3 shows the process of creating the adversarial presentation attack. 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 PA is obtained by printing the digital adversarial image negative on a translucent sheet using a standard laser printer and casting a silicone material on top of it (Figure 4). We evaluated the percentage of successful fingerprint adversarial presentation attacks on both white-box and black-box systems, with white-box systems referring to AFISs and FPADs in which the attacker has complete knowledge of the system architecture and parameters and black-box systems referring to those in which the attacker has no prior knowledge of the underlying system. These experiments have highlighted the feasibility and danger of the attack.

4. Adversarial Liveness Detector

ALD represents the deep learning-based fingerprint liveness detection proposed in [7], 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. To test the effectiveness of the proposed approach, we took part in the LivDet 2021 [12] competition, submitting a methodology to recognize counterfeit biometry from live ones and obtaining first place out of 23 participants in the "Liveness Detection in Action track". In particular, the idea of the proposed solution is to leverage adversarial fingerprints as a way to force the designed CNN-based liveness detector to focus only on the most important portions of the fingerprint, with the aim of reducing the chance of it being misled by minor details. Indeed, adversarial perturbations are very suited for this goal, as they tend to highlight such minor details that tend to mislead the pad. To maximize this effect, it is convenient to use a gradient-based adversarial perturbation algorithm, in order to exploit the gradient with respect to the input of the used CNN-based FPAD. Among all the fingerprints adversarial algorithms, we made use of the modified version of DeepFool [13] based on an efficient iterative approach exploiting the network gradient of a locally linearized version of the loss. More in detail, we further modified the perturbation strategy by not interrupting the attack as soon as the target liveness detector recognized a fake fingerprint as live with a probability $\geq 70\%$ and by amplifying the perturbation at each iteration by a magnification factor of 10^3. As a result, we obtained an attack success rate $\geq 99\%$, with every single fingerprint able to fool the FPAD with a confidence of at least 70%.

However, using adversarial fingerprints as an ad-hoc data augmentation strategy is not trivial, as a target CNN-based FPAD is needed to craft the adversarial fingerprints and it is important to not cause the final model to be overfitted to the adversarial samples. In ALD, we designed an iterative training procedure consisting of three main steps: we first train a CNN-based FPAD on the clean (i.e. non adversarially perturbed) fingerprint data, we then generate the adversarial fingerprints as described above by using the target CNN FPAD as LD, and repeat the training by also adding the adversarial fingerprint to the training data with their original label (i.e. the pre-perturbation class), as we want to make the FPAD more robust. The result is an adversarial data augmentation schema, summarized in Figure 5, where adversarial attacks are exploited to improve the network generalization ability.

The performance obtained in the LivDet 2021 [12] international competition proved the effectiveness of the methodology proposed in [7] and highlighted the significant contribution of adversarial perturbation techniques to the generalization capacity of the CNNs considered as FPAD. In future works we will further investigate the use of adversarial fingerprints in the context of both liveness detection and subject matching, trying to understand whether this experience can be used also to support or against impersonification attacks.

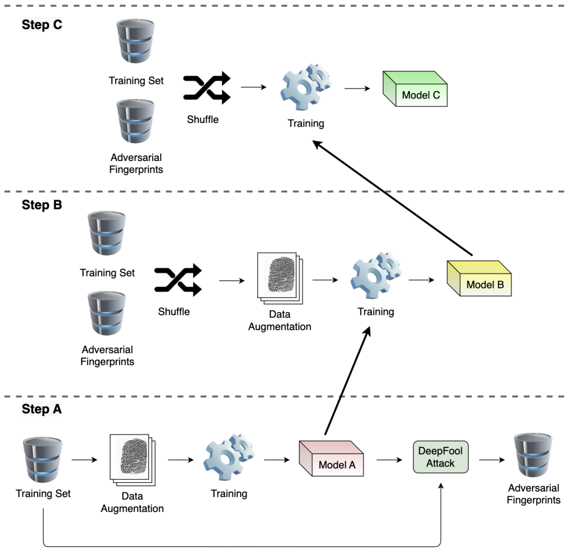

Figure 5: The proposed three-stage training schema: In the first step (A), the model (pre-trained on ImageNet) is fine-tuned on the clean challenge data. We also adopt classical data augmentation, limited to algorithms operating only on pixel intensity values (saturation, shading, etc.). Once the model is trained, the DeepFool algorithm is used to create a new dataset of adversarially perturbed fingerprints (both live and fake fingerprints are perturbed). 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 classical data augmentation algorithms is also 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, but without using any other data augmentation operation.

5. Conclusions

Fingerprint Presentation Attack Detection (FPAD) is considered an arms race problem due to the continuous and dynamic struggle between attackers who develop novel techniques to deceive fingerprint recognition systems and defenders who design and improve FPAD methods to counter these threats. For this reason, the dual approach that considers the two points of view during the design of FPADs and their integration into AFIS is crucial to discover unknown vulnerabilities and fix them. Our experience in the international competition LivDet as organizers, for the University of Cagliari, and as participants, for the University of Naples Federico II, has allowed us to highlight this aspect. Moreover, in this paper, we have presented a case of dual approach in the FPAD related to the exploitation of spoofs obtained with adversarial processes: we have shown that it is possible to start from the analysis of the danger deriving from a new attack technique, in this case the adversarial presentation attack, a defence technique can be designed.

References

[1] S. Marcel, M. S. Nixon, J. Fiérrez, N. W. D. Evans (Eds.), Handbook of Biometric Anti-Spoofing - Presentation Attack Detection, Second Edition, Advances in Computer Vision and Pattern Recognition, Springer, 2019. doi:10.1007/978-3-319-92627-8.

[2] Z. Akhtar, C. Micheloni, G. L. Foresti, Biometric liveness detection: Challenges and research opportunities, IEEE Security & Privacy 13 (2015) 63–72.

[3] L. Ghiani, D. A. Yambay, V. Mura, G. L. Marcialis, F. Roli, S. A. Schuckers, Review of the fingerprint liveness detection (livdet) competition series: 2009 to 2015, Image and Vision Computing 58 (2017) 110–128.

[4] T. Chugh, K. Cao, A. K. Jain, Fingerprint spoof buster: Use of minutiae-centered patches, IEEE Transactions on Information Forensics and Security 13 (2018) 2190–2202.

[5] S. Marrone, C. Sansone, Adversarial perturbations against fingerprint based authentication systems, in: 2019 International Conference on Biometrics (ICB), IEEE, 2019, pp. 1–6.

[6] S. Marrone, C. Sansone, On the transferability of adversarial perturbation attacks against fingerprint based authentication systems, Pattern Recognition Letters 152 (2021) 253–259.

[7] A. Galli, M. Gravina, S. Marrone, D. Mattiello, C. Sansone, Adversarial liveness detector: Leveraging adversarial perturbations in fingerprint liveness detection, IET Biometrics (2023).

[8] G. L. Marcialis, F. Roli, Liveness detection competition 2009, Biometric Technology Today 17 (2009) 7–9.

[9] M. Micheletto, G. Orrù, R. Casula, D. Yambay, G. L. Marcialis, S. Schuckers, Review of the Fingerprint Liveness Detection (LivDet) Competition Series: From 2009 to 2021, Springer Nature Singapore, Singapore, 2023, pp. 57–76. URL: https://doi.org/10.1007/978-981-19-5288-3_3. doi:10.1007/978-981-19-5288-3_3.

[10] R. Casula, M. Micheletto, G. Orrú, G. L. Marcialis, F. Roli, Towards realistic fingerprint presentation attacks: The screenspoof method, Pattern Recognition Letters (2022). URL: https://www.sciencedirect.com/science/article/pii/S0167865522002653. doi:https://doi.org/10.1016/j.patrec.2022.09.002.

[11] M. Micheletto, G. L. Marcialis, G. Orrù, F. Roli, Fingerprint recognition with embedded presentation attacks detection: are we ready?, IEEE Transactions on Information Forensics and Security 16 (2021) 5338–5351.

[12] R. Casula, M. Micheletto, G. Orrù, R. Delussu, S. Concas, A. Panzino, G. L. Marcialis, Livdet 2021 fingerprint liveness detection competition-into the unknown, in: 2021 IEEE International Joint Conference on Biometrics (IJCB), IEEE, 2021, pp. 1–6.

[13] S.-M. Moosavi-Dezfooli, A. Fawzi, P. Frossard, Deepfool: a simple and accurate method to fool deep neural networks, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016, pp. 2574–2582.