Balancing Accuracy and Error Rates in Fingerprint Verification Systems Under Presentation Attacks With Sequential Fusion

Marco Micheletto, *Member, IEEE*, and Gian Luca Marcialis, *Senior Member, IEEE*

***Abstract*—The assessment of the fingerprint PADs embedded into a comparison system represents an emerging topic in biometric recognition. Providing models and methods for this aim helps scientists, technologists, and companies to simulate multiple scenarios and have a realistic view of the process's consequences on the recognition system. The most recent models aimed at deriving the overall system performance, especially in the sequential assessment of the fingerprint liveness and comparison pointed out a significant decrease in Genuine Acceptance Rate (GAR). In particular, our previous studies showed that PAD contributes predominantly to this drop, regardless of the comparison system used. This paper's goal is to establish a systematic approach for the "trade-off" computation between the gain in Impostor Attack Presentation Accept Rate (IAPAR) and the loss in GAR mentioned above. We propose a formal "trade-off" definition to measure the balance between tackling presentation attacks and the performance drop on genuine users. Experimental simulations and theoretical expectations confirm that an appropriate "trade-off" definition allows a complete view of the sequential embedding potentials.**

***Index Terms*—Biometrics, fingerprint recognition, presentation attack detection.**

I. Introduction

THE PRESENTATION attack detection ability is explicitly required for current fingerprint-based personal verification systems in many security applications. As a matter of fact, the design of any biometric verification system cannot ignore the vulnerability to spoofing or presentation attacks (PA), which must be addressed by effective countermeasures from the beginning. The increasing attention to the presentation attack detection (PAD) [1], also called anti-spoofing or liveness detection, has led to substantial advancements in fingerprint-based security solutions, achieving excellent results [2], [3], [4], [5].

Manuscript received 12 June 2023; revised 12 October 2023 and 31 March 2024; accepted 16 May 2024. Date of publication 27 May 2024; date of current version 19 July 2024. This work is supported by the European Union - NextGenerationEU within the PRIN 2022 PNRR - BullyBuster 2 - the ongoing fight against bullying and cyberbullying with the help of artificial intelligence for the human wellbeing (CUP: P2022K39K8, prj: F53C2200074007). This article was recommended for publication by Associate Editor H. Han upon evaluation of the reviewers' comments. *(Corresponding author: Marco Micheletto.)*

The authors are with the Department of Electrical and Electronic Engineering, University of Cagliari, 09123 Cagliari, Italy (e-mail: marco.micheletto@unica.it; marcialis@unica.it).

Digital Object Identifier 10.1109/TBIOM.2024.3405554

Nevertheless, a major limitation of the current research is that most studies design and evaluate PADs as independent systems without considering their integration with recognition systems in real-world applications. This is a crucial factor to consider, as the performance of the recognition system may be significantly impacted by the integration of PAD algorithms [6], [7]. The most-reported problem is that the decrease of attacks classification error rate (APCER) [8], [9], [10], also generates a lower genuine acceptance rate (GAR) than the baseline system. This issue is commonly considered an intrinsic consequence of combining the two systems [11], but no further explanation on the extent or predictability of the GAR decrease is given. Accordingly, the impact of embedding a specific PAD in a fingerprint verification system implies the adoption of a "trial and error" process and choosing the parameters that most fit the requirements demanded by the design constraints. Moreover, the possible amount of errors can only be evaluated *a posteriori*, namely, after the whole system has been implemented, resulting in time and resource waste.

What the literature is missing seems effective strategies that optimize the performance of both the recognition system and PADs to ensure the overall system's security and reliability. Up to date, it is difficult to draw the effect on several possible scenarios; in other words, having a clear idea of the conditions for which this embedding may lead to a real gain, that is, an advantage with respect to neglect the presentation attack problem. To fill this gap, we presented in [12] a novel simulation approach based on the probability modeling of receiver operating characteristic (ROC) curves of PADs and verification systems. The goal was to simulate the performance of a sequentially integrated system. This led to the Bio-WISE system, which was made publicly available.[1] Reference [12] exemplified the performance in two possible operational points of state-of-the-art fingerprint PADs and analyzed the performance of a benchmark and a top-level comparator when embedding such PADs. We noticed that the PAD significantly impacts the performance of the integrated system, regardless of the comparison algorithm used. Bio-WISE allowed also exploring the performance as function of the attack probability. However, the current version of Bio-WISE and, in general, the findings of [12] and related previous works [6], [7], [8], [9], [10], [13], [14] do not allow assessing for which PAD's operational points the overall GAR

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

degradation can be still acceptable, with the advantage of handling presentation attacks. The key question is: what is the best way to embed a PAD into a recognition system so that the final product is robust to spoofing (IAPAR low *enough*), without suffering from significantly reduced genuine recognition accuracy (GAR still *acceptable*)? To properly answer, especially considering PADs with different characteristics, we introduce a formal definition of "*trade-off*", a term that is used when referring to "a balancing of factors all of which are not attainable at the same time."[2] We carried out a new set of simulations using LivDet 2019 and 2021 data sets, specifically oriented to derive, from the proposed formulation of trade-off, the extent to which the PAD can be integrated without significantly degrading the whole performance.

From our investigations, we have developed an approach that surpasses the foundational concepts of Bio-WISE. Though it might superficially appear as an extension, our emphasis on formalizing the trade-off problem demonstrates that it is not merely an addition of a parameter. Instead, our methodology is deeply rooted in established theory, addressing dimensions previously unexplored by Bio-WISE. The final result is a novel analytical framework that delivers specific operational ranges for the integrated system, where the performance agrees with expectations and constraints.

The paper is organized as follows. Section II makes the point about the current research on fingerprint comparison and PAD fusion, with specific reference to other biometrics where this problem was put on the table. Section III firstly summarizes the basic model fully described in [12]. Then, it introduces the concept of "trade-off", which is well known in many engineering fields, and in particular how it has been used so far in pattern recognition applications. This allows us to motivate our formal definition of "trade-off" measurement, the information it adds, and how it can be used in practice. A set of experiments, which supports both the theoretical findings and also shows the main guidelines for the designer, is reported in Section IV. The paper concludes with a discussion of the advances and limitations of the present work.

II. Embedding PAD and Comparison Systems: State-of-the-Art Update

Despite the increasing demand for secure and accurate biometric systems, the research on embedding PAD and comparator systems is still far from being mature. While various attacks on biometric systems have been analyzed, also across different modalities (e.g., face, iris or speech [4]), there has been limited exploration on integrating spoof detection techniques with verification systems. The trend in the literature has been to evaluate the spoof detection task independently, leading to a lack of understanding of how to effectively combine the outputs of the PAD and comparator systems to make a final authentication decision [7]. The first works on the topic focused on investigating the performance differences between various frameworks for the fusion of the two modules [6], in order to identify the most effective approach. The authors also designed a Bayesian Belief Network (BBN) to model the relationship between liveness and comparison scores explicitly. Similarly, [10] presents a Bayesian framework for fusing comparison, quality, and liveness measures while also considering the influence of the sensor. Furthermore, Ding et al. [13] extensively investigated diverse BBN architectures to capture the influence of PAD on comparator scores and vice versa.

A noteworthy contribution to the field has been made by the last editions of LivDet, aimed at promoting the development of integrated systems by providing a common platform for researchers to evaluate and compare the performance of their algorithms. The solutions presented employ score-level fusion to generate a unified metric score [14]. In general, all these works follow a similar approach, utilizing two independent architectures to carry out the presentation attack detection and recognition task and implementing a fusion method at the output level. On the other hand, recent studies have shown promising results in developing a unified system model able to perform both tasks simultaneously [15], [16], eliminating the need for separate modules and reducing processing time and computational complexity. However, the major limitation of all the aforementioned works lies in their exclusive emphasis on performance evaluation of the proposed integrated system, overlooking the crucial aspect of system design. There is a noticeable absence of substantial discussions on quantifying the trade-offs between the verification system and PAD performance, specifically in determining the integrated system's optimal operational point for a given application.

A notable exception can be found in the domain of speech biometrics. In this context, the authors of [17], [18], starting from the definition of the detection cost function (DCF) [19], proposed an extension called the tandem detection cost function (t-DCF) that specifically addresses the evaluation of combined automatic speaker verification (ASV) and spoofing countermeasure systems. Their final goal is to provide a comprehensive evaluation metric that accurately reflects the impact of PAD on verification decisions and enables a reliable ranking of ASV systems in the presence of spoofing attacks, regardless of the approach used for the fusion. Nevertheless, despite the significant step forward made in this work, the issue of threshold setting was not addressed. The author focused primarily on developing a protocol for posterior evaluation, conducted after the system had already been constructed and without knowledge of the integration process. This approach differs from simulation-based methods, which involve exploring various scenarios before implementation [12]. As a result, the question of how threshold calibration should be defined for biometric integrated systems remains unresolved. The threshold plays a critical role in striking a balance between false acceptances (accepting impostors) and false rejections (rejecting genuine users), as well as in determining the effectiveness of the PAD system in rejecting presentation attacks. To address this challenge, it is necessary to consider different scenarios and optimize the integration process proactively. By pursuing this goal, it becomes feasible to establish an appropriate threshold(s) set-up that ensures the biometric system's security and usability. The definition of an appropriate method for such proactive assessment is the topic of the next Section.

[2]https://www.merriam-webster.com/dictionary/trade-off

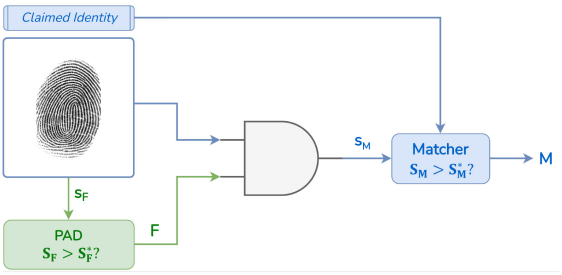

Fig. 1. Example of a possible integrated system configuration. The scheme also represents the boolean events F and M, driven respectively by the PAD and the comparator modules.

III. The "Trade-Off": From a Qualitative to a Quantitative Definition

A. The Bio-WISE Theoretical Model

In this section, we summarize the main assumptions of the Bio-WISE simulator developed in [12]. The sequentially integrated system's structure allowed us to conveniently relate the actual performance to the probabilistic relationships of the two modules, under appropriate working hypotheses.

In particular, to adequately describe the comparator and the presentation attack detector's acceptance rate, we introduced two boolean events driven by their outcomes, namely the comparison and liveness scores. The Match event is defined as the evidence that the comparison score s_M between the input trait and the claimed identity's template is over a given acceptance threshold s_M^*. In other words: $M = s_M > s_M^*$. Similarly, the PAD classes a specific input sample as alive or fake when the liveness score s_F is greater than a certain liveness threshold s_F^*. Therefore, we defined the following event: $F = s_F > s_F^*$.

The sequential nature of the embedding represented in Fig. 1 sets up the final decision to be an AND-like boolean one, that is, the pattern is finally accepted when both F and M events are *True*. This assumption has allowed us to consider the PAD-comparator fusion as a particular case of AND fusion system, where even the error rate evaluation can be treated similarly [12].

Thus, we proved that the expressions governing the performance[3] of a sequential system, regardless of whether the comparator precedes or follows the PAD, can be approximated as:

$$\begin{aligned} GAR_S &= GAR(M) \cdot (1 - BPCER(F)) \\ FMR_S &= FMR(M) \cdot (1 - BPCER(F)) \\ IAPAR_S &= IAPAR(M) \cdot APCER(F) \end{aligned} \tag{1}$$

where the three significant indices of the sequential system are always the simple product of the error rates of the individual systems. This simulator allowed us to assess the impact of sensors, spoof materials and PADs on the integrated system. Our analysis was referred to two fixed operational points of the PAD, namely $BPCER = 1\%$ and $APCER = 1\%$, and the probability of a presentation attack w (see also the term ω in [11]), without lack of generalization. Indeed, the following parameter was added, in agreement with what reported in [11]:

$$GFMR_S = FMR_S \cdot (1 - w) + IAPAR_S \cdot w \tag{2}$$

Eq. (2) allows representing the probability that an attack, whatever is, succeded. Reference [12] proved that w is really the prior probability of a presentation attack.

The model defined by Eqs. (1)–(2) pointed out one crucial aspect: the GAR will unavoidably decrease in an integrated system, in inverse proportion to the PAD's BPCER. The less the BPCER, the more the GAR degradation.

The PAD's impact was highlighted in [12], in agreement with previous evidences,[4] but not in-depth, since the [12]'s analysis was aimed at inspecting the performance of the comparator, keeping the PAD operational point constant.

The primary objective of this work is to introduce a well-grounded tool that can effectively quantify the loss of performance in terms of GAR and the related gain in terms of GFMR (Eq. (2)). Our methodology is tailored to rigorously evaluate the optimal operational points of the presentation attack detection module when the operational points of the comparator are already defined. We can discern the best trade-off based on the intended application by employing this tool across a range of PADs and comparators. Using this tool with multiple PADs and comparators, we can assess the best trade-off based on the target application, thus excluding, for example, the out-of-the-trade-off PAD modules from the following design steps or being aware of the scenarios[5] where they cannot work according to the overall system's requirements.

Although it is built upon the foundation laid by the Bio-WISE model, it is pivotal to delineate how our proposed model differentiates itself and addresses the existing gaps. While BIO-WISE was adept at setting the PAD operating point and demonstrating how performance modulates with the abstract parameter w (representing the attack probability), our framework provides a detailed analysis by systematically evaluating each PAD operational point and its subsequent impact on GAR. This methodological advancement enables our model to not only identify shifts in GAR but also to predict and address them proactively. This addresses a notable limitation observed in the Bio-WISE approach. Additionally, the Bio-WISE model's reliance on the parameter w, grounded in probabilistic reasoning, can sometimes render its practical interpretation elusive. This may create ambiguities for professionals aiming to derive direct applications from the model. In contrast, our current approach provides distinct metrics and concrete insights, facilitating a more straightforward decision-making process in system integration. The proposed methodology allows for a more direct assessment of system performance against real-world threats without the dependency on probabilistic parameters that may be difficult to estimate accurately in practice. This refinement ensures the model's

[3] According to the ISO definition released in https://www.iso.org/obp/ui/#iso:std:iso-iec:30107:-3:ed-2:v1:en, we adopted the term IAPAR instead of previous ones IAMPR [12] and S-FAR [8].

[4] https://cordis.europa.eu/project/id/257289/it

[5] In terms of probability of presentation attack, or typology of presentation attack instrument.

applicability is more precise and better aligned with real-world scenarios.

B. Performance "Trade-Off": A Formal Definition

Before introducing our contribution, it is essential to emphasize that although the "trade-off" may be a well-known concept in general, the specific meaning and application can vary widely depending on the context. In engineering, "trade-off" may refer to balancing competing objectives or factors, such as cost, quality, and time [20], [21], or trading off different aspects of performance, such as accuracy, robustness and reliability [22], [23]. Researchers and practitioners may use the term differently based on their specific goals and constraints. For instance, [20] formalizes the process of making trade-off decisions by modeling a hybrid approach that combines the preferences of the design and performance parameters. The design problem, in this case, is to identify the set of design parameters that maximizes overall preference.[6]

The design strategy described above was also used in [21] to optimize operations in a fabrication company by conceiving a production layout that would increase productivity. The process and production layout were then analyzed according to five design criteria: economic, health and safety, ergonomics, environmental impact, and productivity. In general, such quantitative trade-off approaches are widespread in the context of optimization, decision support and design space exploration. In the latter case, trade-space analysis is a standard methodology to evaluate design options by comparing them across various performance dimensions such as functionality, efficiency, safety, or reliability. This allows decision-makers to explore the "trade-space" of design alternatives, that is, the range of possible design options that meet the required criteria. The evaluation process may employ models or simulations to assess the performance of different options [22], or it may depend on expert judgment.

In Pattern Recognition, to our knowledge, this concept has been defined clearly and precisely through two significant contributions, namely, [24], [25]. It is closely linked to the mechanism of the "reject" option.

The reject option is a feature that allows the system to withhold the automatic classification of a particular input pattern when it is likely to be misclassified. The rejected patterns are flagged for further review or processing, such as manual inspection, or fed to a more accurate classifier. This approach helps to reduce the misclassification rate and improve the system's overall accuracy, especially when dealing with complex or noisy input patterns. However, since the error and rejection rates are inversely related (i.e., the lower the error rate, the higher the rejection rate), it is necessary to find the optimal "trade-off" that maximizes the system's overall accuracy based on the specific requirements of the target application.

[6]In the design of imprecision [32], the "preference" is a value associated with a certain parameter's value to indicate the designer's uncertainty about that parameter's value. If the designer does not trust that parameter's value, the preference is ranked low, near zero, for example.

Chow [24] proposed the optimal classification rule with the reject option based on the minimum risk theory, considering the costs of different types of errors and their probabilities. Specifically, the rule rejects a pattern if its maximum class posterior probability is lower than a given threshold. However, the rule's optimality depends on the exact knowledge of posterior probabilities, which are usually unknown in practical applications [26]. For these reasons, Fumera and Roli [25] extended the framework developed by [27], which assessed the increase in error probability when *estimates* of class posterior probabilities are used. By incorporating the reject option, they evaluated and compared the performances of individual and combined classifiers under different assumptions about the distribution of the estimation errors. Finally, they presented guidelines for determining whether a linear combination of classifiers can improve the error-reject trade-off through simple or weighted averaging of their outputs. In particular, they conclude that, for classifiers exhibiting the same accuracy, the simple average can be expected to provide the best error-reject trade-off.

In both works, the meaning of trade-off was related to avoiding classifying patterns that were likely to be misclassified, thus, the measurement of trade-off was designed accordingly, in agreement with the theoretical findings.

However, in our case, we must design a trade-off measurement to find the PADs operational points, if any, such that the overall genuine users' acceptance rate does not degrade and the performance on presentation attacks is still under control.

Therefore, our trade-off rule represents a novelty in the literature. Contrary to previous works, it is not limited to the simple fusion of features of individual patterns. Still, it is designed for the sequential fusion of the two characteristics of fingerprints: proximity to the reference template and authenticity in a strict sense. Developed for the generic comparator, it enriches with new definitions the existing predictive model reported in [12]. Indeed, Eqs. (1) are not expressive enough to evaluate whether the loss of accepted genuine samples introduced by the Presentation Attack Detection (PAD) embedding can be kept within a given tolerance range. On this basis, we finally introduce our definition of *trade-off*, expressed as the ratio between the fraction of attackers and genuines accepted for a given match threshold value s_M^*. Since the attackers can be classified into two groups, impostors (zero-effort) and PAs, we have two trade-off values:

$$T_{ZE}^{M} = \frac{FMR(M)}{GAR(M)} \tag{3}$$

$$T_{PA}^{M} = \frac{IAPAR(M)}{GAR(M)} \tag{4}$$

where the abbreviations ZE and PA stand respectively for "Zero-Effort" and "Presentation Attack". It is worth noting that the formal definition above quantifies the balance between the error rate and the (opposite of) rejection rate, fitting perfectly the thread established by [24], [25]. Thus, it is reasonable to refer to Eqs. (3)–(4) as representatives of the term "trade-off" when a comparator must deal with genuine users and attacks "at the same time".

Due to the cumulative nature of the error curves and since $IAPAR \geq FMR$ [28], the relation $T_{ZE} \leq T_{PA}$ is always valid, whatever the comparator threshold value found. Additionally, these metrics can be successfully employed to assess the worst-case performance scenarios without using a PAD, as, by definition, the verification system cannot counter a presentation attack.

Since we want to evaluate the improvement achievable by a sequential integrated system, we may express the relative trade-off values by recalling Eqs. (1):

$$T^S_{ZE} = \frac{FMR_S}{GAR_S} = \frac{FMR(M)\cdot(1-BPCER(F))}{GAR(M)\cdot(1-BPCER(F))} = T^M_{ZE} \quad (5)$$

$$T^S_{PA} = \frac{IAPAR_S}{GAR_S} = \frac{IAPAR(M)\cdot(APCER(F))}{GAR(M)\cdot(1-BPCER(F))} = \\ = T^M_{PA}\cdot\tau_F \quad (6)$$

From these formulations, we can mainly highlight the following aspects:

- The trade-off values relating to zero-effort attacks are independent of the liveness threshold. In other words, the original relationship between *FMR* and *GAR* cannot be changed by any PAD.
- The performance ratio, denoted as τ_F, is always less than one since $APCER_F \leq (1 - BPCER_F)$ for any liveness threshold.
- The PAD inclusion reduces the maximum error obtainable by the verification system alone, namely T^M_{PA}, in proportion to the τ_F parameter. For the same liveness operating point, the more efficient the liveness detector, the better the improvement.
- As for the original simulator of [12] and the [25]'s findings, it is not required the exact knowledge of the individual systems' operational points values.

To further study the role of the trade-off in the systems embedding, we focus on determining whether an operational point of the presentation attack detector exists, such as to keep the loss of GAR within a specific *tolerance* margin.

C. A Case Study: The Equal Error Rate (EER)

We report here an example of a case study obtained by selecting, for the sake of simplicity, the Equal Error Rate (EER), which can be considered the comparator operational point par excellence. However, our findings can be extended to any other operational point.

In this instance, Eqs. (3)–(4) assume the following constant values:

$$T^{EER}_{ZE} = \frac{EER}{1-EER} \quad (7)$$

$$T^{EER}_{PA} = \frac{EER+\Delta}{1-EER} = T^{EER}_{ZE}\left(1+\frac{\Delta}{EER}\right) = \\ = T^{EER}_{ZE}(1+\Delta_{EER}) \quad (8)$$

where the term Δ_{EER} in Eq. (8) expresses the fraction deviation (also representable in percentage) from T^{EER}_{ZE} and depends on the relative performance difference Δ between the percentage of impostors (FMR) and presentation attacks accepted (IAPAR) at the EER. It is worth remarking that these quantities are estimated from the ROC of the verification system.

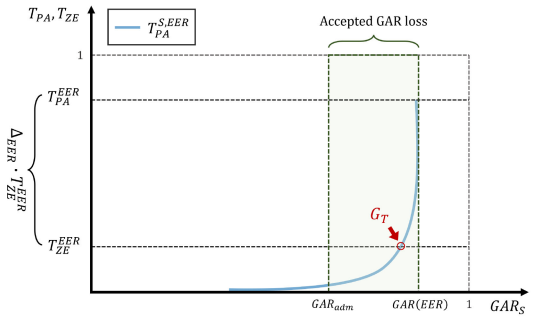

Fig. 2. Toy graph displaying the relationship between T^{EER}_{ZE}, T^{EER}_{PA} and $T^{S,EER}_{PA}$ in an integrated system.

Similarly, we can easily define from Eq. (6) the trade-off for presentation attacks relating to the serial system, as it is the only one subject to the PAD influence:

$$T^{S,EER}_{PA} = T^{EER}_{PA}\cdot\tau_F = T^{EER}_{ZE}(1+\Delta_{EER})\cdot\tau_F \quad (9)$$

In order to show how the trade-off values can be exploited to select the most appropriate PAD operating point, we provide in Figure 2 a possible trend of the trade-off curves defined by Eqs. (7)–(9) when plotted against the GAR of the sequential system. We remember that our model (Eqs. (1)) may simulate the integrated system's performance parameters without actually implementing it overall.

First of all, we observe that the $T^{S,EER}_{PA}$ curve (blue line) is included within the operational points of $zero-APCER$ of the PAD ($\tau_F = 0$) and the first liveness threshold value for which $\tau_F = 1$. At this point, the serial system equals the comparator's performance in detecting spoofs, thus cancelling all PAD advantages.

This means that, through this curve, we can estimate the GAR loss associated to each operational point of the liveness detector. Among these, what is the working point that may guarantee the most appropriate balance? Ideally, the best possible compromise would allow keeping the performance of the integrated system stable on zero-effort and at the same time improve that relating to PAs.

In Fig. 2, this point corresponds to the intersection of the $T^{S,EER}_{PA}$ curve with the T^{EER}_{ZE} straight line, in which the ratio between impostors/genuine is equivalent to the false/genuine one, namely the integrated system detects fakes with the same "efficiency" with which the comparison system alone blocks the impostors. For the sake of clarity, we have marked this point in Fig. 2 as G_T.

As previously stated, this returns a GAR loss, which is proportional to the performance of the liveness detector. Accordingly, the tolerance margin, within which to accept the genuines' loss to improve the fakes' detection, can be defined as follows:

$$\rho = GAR(EER) - GAR_{adm} \quad (10)$$

where GAR_{adm} is the minimum *adm*issible GAR of the integrated system, which is still compatible with the simulated

TABLE I
LIVDET 2019 AND 2021 DEVICE FEATURES

Scanner	Model	Res.[dpi]	Img Size	Format	Type
Green Bit	DactyScan84C	500	500x500	BMP	Optical
Digital Persona	U.are.U 5160	500	252x324	PNG	Optical
Orcanthus	Certis2 Image	500	300xN	PNG	Thermal swipe
Dermalog	LF10	500	500x500	PNG	Optical

scenario's constraints. The lower its value, the greater the tolerance for GAR loss. Accordingly, ρ indicates the maximum percentage deviation from the nominal GAR value of the comparator alone at the EER. In our plots (starting from Fig. 2), it is represented by the green area.

Once the region has been delimited, we can derive the following guideline from Fig. 2:

$$GAR_{adm} \leq G_T \tag{11}$$

This means that when the accepted GAR loss (GAR_{adm}) falls on or to the left of the intersection point G_T, the most advisable decision is to set the corresponding working point for the PAD at that value. Consequently, the loss of GAR is within the fixed range. Otherwise, it is possible to evaluate any intermediate point that generates a satisfactory advance compared to the verification system's case. Whether such a point does not exist, the PAD under consideration *does not fit* the scenario's constraints and can be discarded.

Finally, the example also suggests that the trade-off on presentation attacks can be made even better than the T_{ZE}^{EER} value by choosing any point at the left of G_T. However, significant attention must be paid to the fact that the GAR of the sequential system degrades rapidly.

In summary, once the comparator operational point has been set, our trade-off definition allows to accurately assess under which conditions a presentation attack detector can be integrated without significantly degrading the overall performance in terms of GAR. The following section shows how to apply the outlined guidelines to a real-case study.

IV. EXPERIMENTAL SIMULATIONS

A. Datasets and Protocol

The proposed experimental analysis was performed on LivDet 2019 and 2021 datasets [29]. Both datasets contain high-resolution fingerprint images, including live and spoof samples. LivDet 2019 consists of three datasets (Table I rows 1-3), obtained from different acquisition sensors but common spoof materials. LivDet 2021, on the other hand, includes two sub-datasets obtained by two sensors (Table I rows 1 and 4), different materials and covers a more comprehensive range of presentation attacks by including a new spoof fabrication technique called ScreenSpoof [30]. We employed all the algorithms submitted to both competitions for preliminary analysis, followed by a more focused investigation of the top-two performing algorithms of LivDet 2019 and 2021 on the datasets of the same competition; their behavior allows us to summarize that of the other algorithms and cover diverse simulation scenarios. These PADs can be considered among the best at the state of the art. Then, for each dataset, we followed a four-step procedure: (1) we computed the liveness scores using the aforementioned PADs, whose details are reported in Table II; (2) we computed the match score employing the standard NIST Bozorth3 and the top-level VeriFinger 12 comparator; (3) we derived individual acceptance rates for the comparison system, namely GAR, FMR and IAPAR, and the error rates for the liveness detector, that is, BPCER and APCER and subsequently we applied Eqs. (1) for computing the acceptance rates of the integrated system; (4) we computed the trade-off values by setting the operating point of the comparator at EER (Eqs. (7)–(9)).

TABLE II
PARTICIPANTS ALGORITHMS NAME OF THE TOP-TWO WINNERS OF LIVDET 2019 AND 2021

Partecipant	Algorithm name	Year	Type
CENATAV	PADUnkFv	2019	Hand-crafted
Hangzhou Jinglianwen Technology Co.,Ltd	JLW_LivDet	2019	Deep-learning
MEGVII (BEIJING) TECHNOLOGY CO, LTD	megvii_ensemble	2021	Deep-learning
Dermalog	LivDet21_Dob_C2	2021	Deep-learning

This analysis demonstrated that our novel instrument may be employed not only in the meta-design process to determine the optimum PAD operating point, but also as a comparator of current PAD technology when applied to a specific comparator and sensor combination.

B. Results

In order to guarantee a correct evaluation of the data and graphs, we first report in Table III-IV the values of T_{ZE}, T_{PA} and GAR calculated at the EER working point of the comparator for the analyzed datasets. The significant difference between the two trade-off values of ZE and PA testify to the danger of spoofing if not correctly contrasted. This is particularly apparent for the Verifinger 12 comparator, which although it provides a benefit to zero-effort attacks detection, presents a much higher T_{PA}^{EER} than Bozorth3, resulting, for specific datasets, being utterly vulnerable to presentation attacks. We can hypothesize that this dissimilarity is due to the different nature of the two comparators: in contrast to Bozorth3, VeriFinger leverages deep neural networks combined with exclusive algorithmic solutions that amplify the system's effectiveness and reliability. It is plausible that the use of deep neural networks in VeriFinger allows for more effective feature extraction from fake fingerprints, leading to better detection of minutiae and improved comparison performance. However, the exact mechanisms by which VeriFinger handles fingerprints are proprietary and not publicly disclosed.

In light of this, integrating a PAD with the verification system is crucial to ensure robust and reliable detection

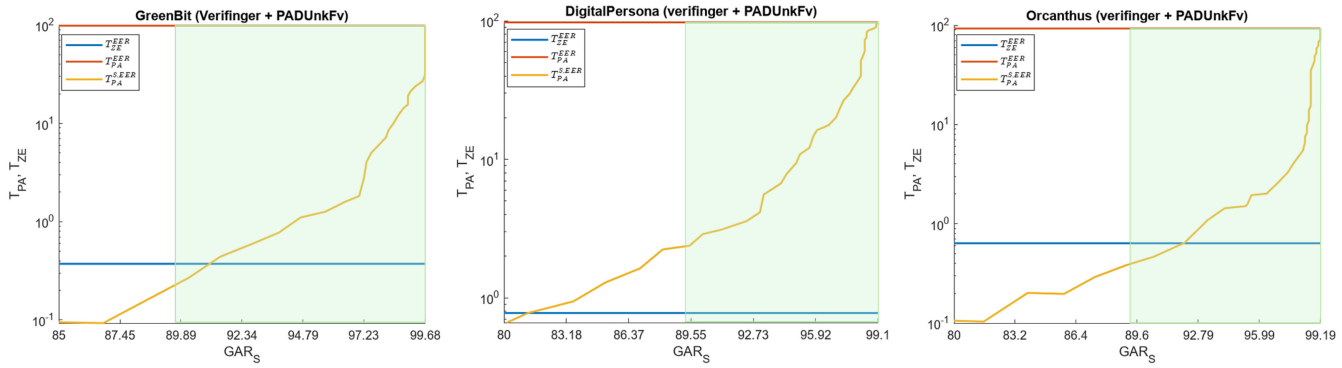

Fig. 3. $T_{PA}^{S,EER}$ trend for GreenBit (a), DigitalPersona (b) and Orcanthus (c) sensors equipped with Verifinger 12 comparison system at the EER operating point and "PADUnkFv" liveness detector. The tolerance margin ρ (green area) is set to 10%. The y-axis is in logarithmic scale.

TABLE III
GAR AND TRADE-OFF VALUES (IN PERCENTAGE) AT THE EER OPERATIONAL POINT FOR LIVDET 2019 DATASETS EQUIPPED WITH BOZORTH3 AND VERIFINGER 12 COMPARATOR: GREENBIT (GB), DIGITALPERSONA (DP) AND ORCANTHUS (OR) SENSORS

		Bozorth3	Verifinger 12
GB	GAR	98.77	99.68
	T_{ZE}	1.13	0.37
	T_{PA}	69.43	98.92
DP	GAR	94.73	99.10
	T_{ZE}	4.66	0.78
	T_{PA}	58.48	97.49
OR	GAR	95.30	99.20
	T_{ZE}	5.93	0.64
	T_{PA}	52.79	92.72

TABLE IV
GAR AND TRADE-OFF VALUES (IN PERCENTAGE) AT THE EER OPERATIONAL POINT FOR LIVDET 2021 DATASETS EQUIPPED WITH BOZORTH3 AND VERIFINGER 12 COMPARATOR: GREENBIT (GB) CONSENSUAL (CC) AND SCREENSPOOF (SS) AND DERMALOG (DL) CONSENSUAL (CC) AND SCREENSPOOF (SS). NOTE THAT *GAR@EER* AND T_{ZE} VALUES ARE THE SAME FOR CC AND SS DATASETS SINCE THEY SHARE THE SAME LIVE FINGERPRINTS

		Bozorth3	Verifinger 12
GB CC	GAR	97.40	98.02
	T_{ZE}	1.88	1.76
	T_{PA}	27.76	75.35
GB SS	GAR	97.40	98.02
	T_{ZE}	1.88	1.76
	T_{PA}	43.55	61.22
DL CC	GAR	96.33	97.54
	T_{ZE}	2.77	2.64
	T_{PA}	46.52	86.69
DL SS	GAR	96.33	97.54
	T_{ZE}	2.77	2.64
	T_{PA}	51.78	66.93

of presentation attacks and prevent unauthorized access to sensitive information. At the same time, it is essential to thoroughly evaluate the impact of the PAD on the overall system's performance. Below, we present some use-case examples obtained from the examined datasets and generated using our tool to illustrate its effectiveness in evaluating such impact.

Let us consider, for instance, the GreenBit sensor equipped with the comparator VeriFinger and the best detector of LivDet 2019, namely "PADUnkFv" (Figure 3a). The last value indicated in the x-axis corresponds to the GAR value at the EER, and the y-axis is presented in logarithmic scale. The acceptance area is obtained by setting $\rho = 10\%$. To evaluate the PAD's effectiveness in detecting spoofs, the T_{PA}^{S} (yellow curve) trade-off curve should be observed. We note that this curve crosses the straight line T_{ZE} for a value of $GAR \simeq 92\%$. It means that to bring the liveness detection rate (trade-off on PAs) to the same level of the verification system's accuracy on impostors (trade-off on zero-effort attacks), we should accept a loss of GAR of approximately eight per cent. The crossing point is located within the green area of tolerance, therefore this could be a case of a feasible integrated system, as it can block presentation attacks with high efficiency, keeping the associated GAR loss within the performance constraints. Another example of a suitable embedded system is depicted in Figure 3c), obtained by applying the same configuration of PAD/comparator to the Orcanthus sensor.

However, it is important to point out that the intersection point is only a *possible* choice. As a matter of fact, our simulator allows to clearly view the integrated system's behaviour for each PAD operational point and, accordingly, choose the one that best suits the final application context. In both cases (Fig. 3{a,c}), the T_{PA}^{S} curve decreases rapidly at first and then more slowly until it crosses the T_{ZE} line; therefore, we could select an intermediate point shortly before the gradient becomes too small, achieving a good compromise between rejected PAs and GAR loss.

This also applies when the crossing point is not located within the green area. Figure 3b shows a case of this kind, related to the DigitalPersona sensor. Here, the high performance guaranteed by the comparator VeriFinger 12 ($EER < 1\%$) generates a not practicable trade-off point due to the high loss of accepted genuines ($GAR \simeq 83\%$). Nevertheless, the simulation shows us that it is still possible to consistently improve the detection of fakes of over 90% than in the case of the recognition system alone, by choosing, for instance, the point corresponding to the maximum accepted loss value as the PAD's threshold or any other value within the green area.

Another advantage of our trade-off definition is the ability to compare several PADs simultaneously, study their

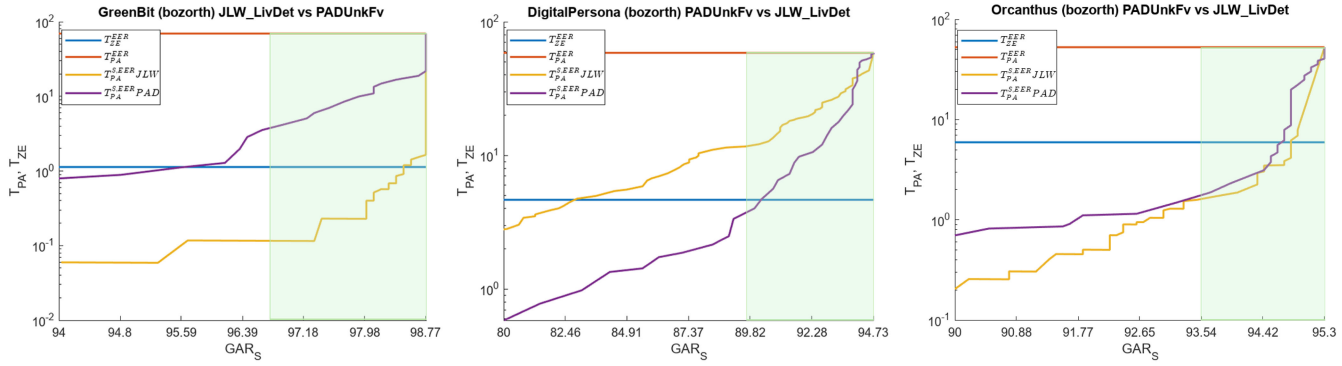

Fig. 4. Comparison between $T_{PA}^{S,EER}$ of the top two PADs of LivDet 2019 integrated with Bozorth3 comparator at the EER operating point on the GreenBit (a), DigitalPersona (b) and Orcanthus (c) sensor. The tolerance margin ρ (green area) is set to 2% (a) and 5% (b,c). The y-axis is in logarithmic scale.

behaviour and consequently choose the one that provides better performance. For this purpose, we show the comparison between the two PADs under examination when they are integrated with the Bozorth3 comparator on all the investigated sensors (Figure 4). This analysis shows three different scenarios, exemplified by each sensor:

1) Figure 4a: the *PADUnkFV* algorithm (green curve) does not meet the required specifications since not only its intersection point is outside the acceptance area defined by the tolerance parameter ($\rho = 2\%$) but also has lower accuracy under the same liveness threshold. On the other hand, the *JLW_LivDet* algorithm (yellow curve) fits the GreenBit characteristics perfectly, achieving a trade-off on PAs comparable to the trade-off on ZE attacks with only 1% loss of GAR.
2) Figure 4b: in this case, the *JLW_LivDet* algorithm (yellow curve) does not match the needed parameters. The tolerance area is defined setting $\rho = 5\%$ and the *PADUnkFV*, albeilt borderline, respects the performance constraints.
3) Figure 4c shows instead a situation of equality among the two PADs since the two curves are nearly superimposed. Therefore, both could be valid choices in an application scenario.

It is worth recalling that we focused only on the top-two winners algorithms of LivDet 2019, nevertheless, the proposed tool can be easily employed to compare several PAD and assess which solution is the most accurate for a given task or simply to evaluate the performance in terms of genuine loss. For this purpose, we present in Figure 5 a comparison of seven different PADs submitted to LivDet 2019 embedded with Bozorth3 on DigitalPersona sensor. For the sake of clarity, we did not draw the green area. However, we can immediately notice that the best PAD is the "PAD 6", with a GAR loss of approximately four percent at the intersection point. This means that the integrated system can work at the EER operational point of the verification system by improving its spoof detection by over 90% with a relatively small cost.

Regarding the Livdet 2021 datasets, they provide crucial data for evaluating integrated fingerprint systems. The four

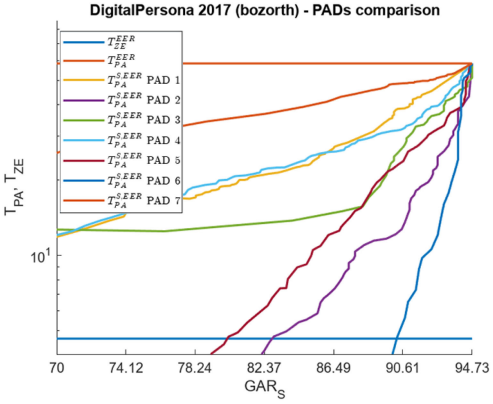

Fig. 5. Comparison between $T_{PA}^{S,EER}$ of eight different PADs of LivDet 2017 integrated with Bozorth3 comparator at the EER operating point on the DigitalPersona sensor. The y-axis is in logarithmic scale.

test sets of the competition were generated from two different sensors and two methods of spoof fabrication: the traditional method and the semi-consensual ScreenSpoof technique [30], by sharing the same live fingerprints. This unique characteristic allows us to establish a relationship between the integrated system's GAR loss and the type of attack. Through the trade-off analysis, we can, for instance, visually identify the optimal operating point for the integrated system that ensures protection against both attack types. This offers a more comprehensive understanding than a mere liveness accuracy evaluation. Let us focus on the GreenBit sensor. Based on the data presented in [31], the average accuracy on the Consensual and ScreenSpoof datasets was 95.52% for Megvii and 91.82% for *Dob_C*2. Assuming the goal is to achieve absolute security (i.e., T_{PA}^{S} at EER equal to zero for both CC and SS attacks), the performance of *Dob_C*2 is more efficient compared to Megvii. As evidence, the two T_{PA}^{S} curves for Dob_C2 approach zero earlier, incurring only a GAR_S loss slightly greater than 2%, against the approximately 7% of Megvii (Figure 6). This is valid for both comparators. This difference can be attributed to *Dob_C*2's higher effectiveness in detecting attacks performed through the ScreenSpoof technique. However, under a less stringent security constraint, that is, considering the point

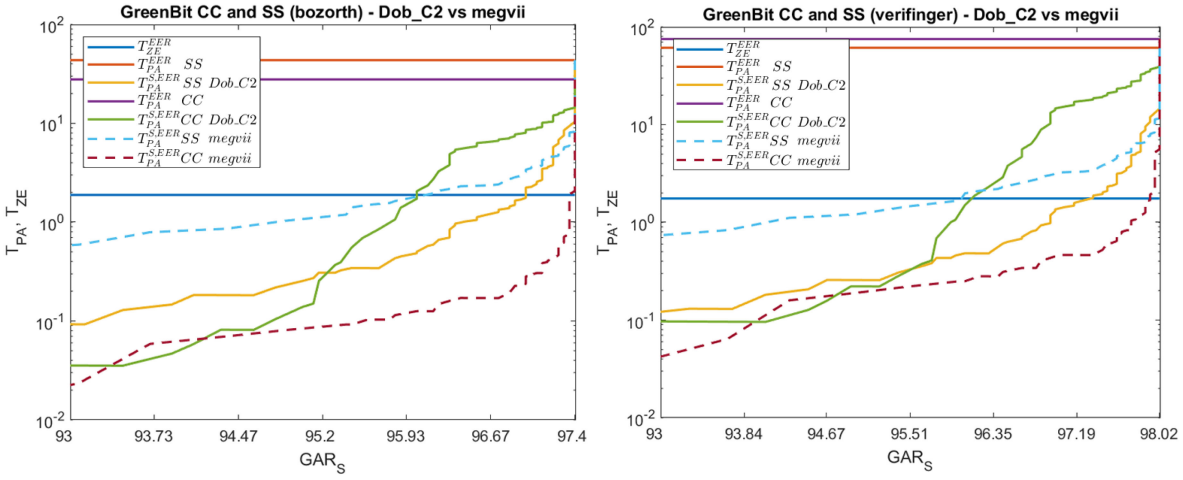

Fig. 6. GreenBit dataset from LivDet2021: Comparison between $T_{PA}^{S,EER}$ of the two subsets Consensual (CC) and ScreenSpoof (SS). *Dob_C*2 (solid line) and *megvii* (dashed line) PADs are integrated with Bozorth3 (a) and Verifinger 12 (b) comparator at the EER operating point. The tolerance margin is omitted. Note that the T_{ZE} value is the same for CC and SS datasets since they share the same live fingerprints. The y-axis is in logarithmic scale.

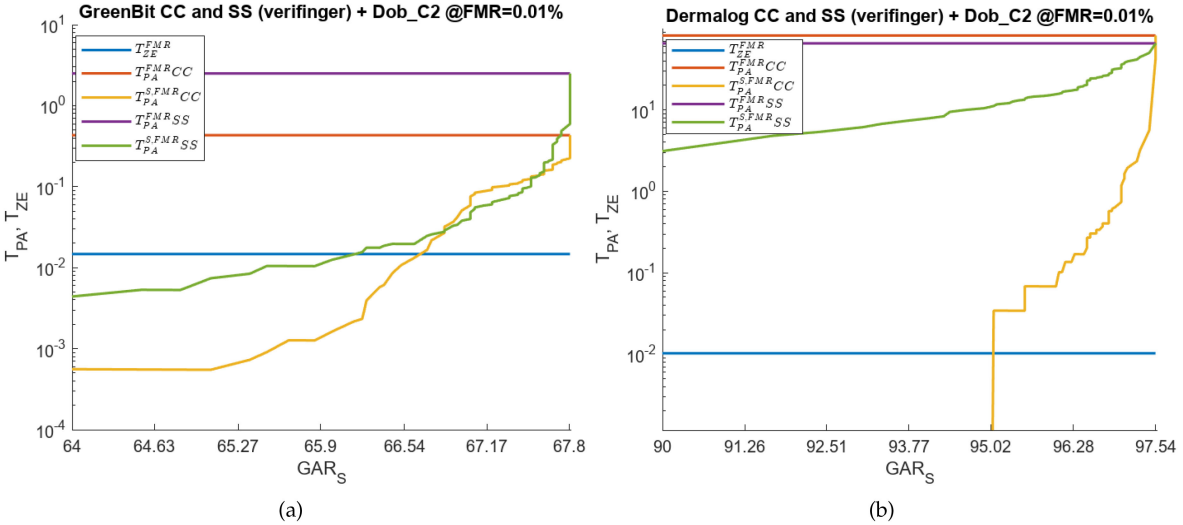

Fig. 7. GreenBit (a) and Dermalog (b) dataset from LivDet2021: Comparison between $T_{PA}^{S,FMR}$ of the two subsets Consensual (CC) and ScreenSpoof (SS). *Dob_C*2 PAD is integrated with Verifinger 12 comparator at the FMR=0.01% operating point. The tolerance margin is omitted. Note that the T_{ZE} value is the same for CC and SS datasets since they share the same live fingerprints. The y-axis is in logarithmic scale.

where both T_{PA}^{S} curves are less than or equal to T_{ZE}, the two solutions display similar performance. This is evident as the G_T value for megvii's T_{PA}^{S} on the SS dataset and *Dob_C*2's T_{PA}^{S} on the CC dataset (respectively, cyan dashed and green solid curve in Fig. 6) is almost the same. As a result, *Dob_C*2 may be considered the optimal choice from an integration perspective when the goal is to face multiple methods of spoof fabrication.

As a concluding analysis of this study, we investigated the system's performance at a stringent operating point of $FMR = 0.01\%$, a threshold aligned with real-world security requirements. Leveraging our trade-off model, analogous to the EER scenario, we observed the inherent system dynamics upon PAD integration. In particular, a pronounced decrease in GAR, T_{ZE} and T_{PA} values became evident (Figure 7a). This immediate visibility into system behaviour at different thresholds stands as a prime advantage of our approach. However, this decrease is not a universally observed trend but is contingent on the dataset. For the GreenBit dataset (Fig. 7a), the FMR curve's more gradual approach to zero establishes a comparatively higher threshold, amplifying the observed reduction in GAR. This effect becomes particularly conspicuous in the system's sustained detection of numerous imposter attempts. In contrast, analysis using the Dermalog dataset (Fig. 7b) demonstrated a performance trajectory more congruent with EER-based evaluations. Thus, analogous insights and considerations can be gleaned from an integration perspective, underscoring the dataset-dependent aspects our model unveils. Further, they reaffirm its utility in offering nuanced insights, essential for researchers and industry practitioners in tailoring biometric systems to address real-world challenges effectively.

V. Conclusion

This paper introduces the "trade-off" measurement, a novel simulation-based technique for evaluating integrated fingerprint systems. While drawing insights from the Bio-WISE model, our methodology transcends its scope by addressing the aspect of threshold calibration, which has also been overlooked in previous studies. By linking the GAR decrease and PAD working points, our method enables the selection of an optimal PAD setting, ensuring a practical operational point for the integrated system. To validate the efficacy of our approach, extensive simulations were conducted on two benchmark datasets, LivDet2019 and LivDet 2021, employing state-of-the-art comparators and PADs. The results demonstrate its efficacy in evaluating whether error rates are within an acceptable range, facilitating the designer's decision-making process in selecting the optimal working point of the PAD.

However, it is important to acknowledge the inherent limitation of our framework, which is its lack of generalizability to fusion techniques beyond sequential fusion and to systems that do not adhere to this specific architecture. Although sequential fusion remains the most straightforward and widely adopted approach, there is a growing trend towards developing unified models that integrate both recognition and presentation attack detection modules. These unified models offer significant advantages regarding reduced parameters and latency in applications with limited storage space and low-performance hardware. Future research should therefore aim to extend our methodology to encompass these diverse fusion techniques, enabling a more exhaustive evaluation of integrated biometric systems that go beyond the constraints of sequential fusion.

Additionally, exploring the applicability of our model as an *a posteriori* method, as exemplified in Kinnunen et al.'s works [17], [18], would be valuable. While grounded in theoretical considerations and simulations, our *a priori* approach is centred on the phase before the full-scale integration of the biometric system. It is important to note that in this stage, single components like the PAD and comparator are evaluated individually. However, the real essence of the "a priori" term captures the evaluations before the complete assembly and operationalization of the integrated system. In contrast, the *a posteriori* method involves assessing the entire system after its implementation, offering a comprehensive view of its real-world effectiveness. Combining insights from both these evaluation methods could pave the way for a richer understanding of the system's performance throughout its lifecycle.

By considering different scenarios, we can enhance the overall evaluation process and improve the security and usability of biometric systems. This opens up exciting avenues for future research in the field of biometric security and evaluation.

References

[1] *Information Technology—Biometric Presentation Attack Detection—Part 3: Testing and Reporting*, ISO/IEC Standard 30107-3, 2023.

[2] E. Marasco and A. Ross, "A survey on antispoofing schemes for fingerprint recognition systems," *ACM Comput. Surv.*, vol. 47, no. 2, pp. 1–36, 2014. [Online]. Available: https://doi.org/10.1145/2617756

[3] C. Sousedik and C. Busch, "Presentation attack detection methods for fingerprint recognition systems: A survey," *IET Biom.*, vol. 3, no. 4, pp. 219–233, 2014. [Online]. Available: https://doi.org/10.1049/iet-bmt.2013.0020

[4] S. Marcel, J. Fierrez, and N. Evans, "Handbook of biometric anti-spoofing," in *Advances in Computer Vision and Pattern Recognition*. Singapore: Springer Nat., 2023. [Online]. Available: https://doi.org/10.1007/978-981-19-5288-3

[5] T. Chugh, K. Cao, and A. K. Jain, "Fingerprint spoof buster: Use of minutiae-centered patches," *IEEE Trans. Inf. Forensics Security*, vol. 13, no. 9, pp. 2190–2202, Sep. 2018, doi: 10.1109/tifs.2018.2812193.

[6] E. Marasco and C. Sansone, "Combining perspiration- and morphology-based static features for fingerprint liveness detection," *Pattern Recognit. Lett.*, vol. 33, no. 9, pp. 1148–1156, 2012. [Online]. Available: https://doi.org/10.1016/j.patrec.2012.01.009

[7] I. Chingovska, A. Anjos, and S. Marcel, "Anti-spoofing in action: Joint operation with a verification system," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. Workshops*, 2013, pp. 98–104. [Online]. Available: https://doi.org/10.1109/cvprw.2013.22

[8] A. Abhyankar and S. Schuckers, "Integrating a wavelet based perspiration liveness check with fingerprint recognition," *Pattern Recognit.*, vol. 42, no. 3, pp. 452–464, 2009. [Online]. Available: https://doi.org/10.1016/j.patcog.2008.06.012

[9] A. Rattani and N. Poh, "Biometric system design under zero and non-zero effort attacks," in *Proc. Int. Conf. Biom. (ICB)*, 2013, pp. 1–8. [Online]. Available: https://doi.org/10.1109/icb.2013.6612999

[10] A. Rattani, N. Poh, and A. Ross, "A bayesian approach for modeling sensor influence on quality, liveness and match score values in fingerprint verification," in *Proc. IEEE Int. Workshop Inf. Forensics Security (WIFS)*, 2013, pp. 37–42. [Online]. Available: https://doi.org/10.1109/wifs.2013.6707791

[11] I. Chingovska, A. Mohammadi, A. Anjos, and S. Marcel, "Evaluation methodologies for biometric presentation attack detection," in *Handbook of Biometric Anti-Spoofing*. Cham, Switzerland: Springer Int. Publ., 2019, pp. 457–480. [Online]. Available: https://doi.org/10.1007/978-3-319-92627-8_20

[12] M. Micheletto, G. L. Marcialis, G. Orrù, and F. Roli, "Fingerprint recognition with embedded presentation attacks detection: Are we ready?" *IEEE Trans. Inf. Forensics Security*, vol. 16, pp. 5338–5351, 2021. [Online]. Available: https://doi.org/10.1109/tifs.2021.3121201

[13] Y. Ding, A. Rattani, and A. Ross, "Bayesian Belief models for integrating match scores with liveness and quality measures in a fingerprint verification system," in *Proc. Int. Conf. Biom. (ICB)*. 2016, pp. 1–8. [Online]. Available: https://doi.org/10.1109/icb.2016.7550095

[14] Y. Zhang, C. Gao, S. Pan, Z. Li, Y. Xu, and H. Qiu, "A score-level fusion of fingerprint matching with fingerprint liveness detection," *IEEE Access*, vol. 8, pp. 183391–183400, 2020. [Online]. Available: https://doi.org/10.1109/access.2020.3027846

[15] A. Popli, S. Tandon, J. J. Engelsma, and A. Namboodiri, "A unified model for fingerprint authentication and presentation attack detection," in *Handbook of Biometric Anti-Spoofing*. Singapore: Springer Nat., 2023, pp. 77–99. [Online]. Available: https://doi.org/10.1007/978-981-19-5288-3_4

[16] S. A. Grosz, K. P. Wijewardena, and A. K. Jain, "ViT unified: Joint fingerprint recognition and presentation attack detection," 2023, *arXiv:2305.07602*.

[17] T. Kinnunen et al., "t-DCF: A detection cost function for the tandem assessment of spoofing countermeasures and automatic speaker verification," in *Proc. Speaker Lang. Recognit. Workshop*, 2018, pp. 1–8. [Online]. Available: https://doi.org/10.21437/odyssey.2018-44

[18] T. Kinnunen et al., "Tandem assessment of spoofing countermeasures and automatic speaker verification: Fundamentals," *IEEE/ACM Trans. Audio, Speech, Lang. Process.*, vol. 28, pp. 2195–2210, Jul. 2020. [Online]. Available: https://doi.org/10.1109/taslp.2020.3009494

[19] G. R. Doddington, M. A. Przybocki, A. F. Martin, and D. A. Reynolds, "The NIST speaker recognition evaluation—Overview, methodology, systems, results, perspective," *Speech Commun.*, vol. 31, nos. 2–3, pp. 225–254, 2000. [Online]. Available: https://doi.org/10.1016/S0167-6393(99)00080-1

[20] K. N. Otto and E. K. Antonsson, "Trade-off strategies in engineering design," *Res. Eng. Design*, vol. 3, no. 2, pp. 87–103, 1991. [Online]. Available: https://doi.org/10.1007/bf01581342

[21] M. M. Navarro and B. B. Navarro, "Engineering trade-off strategies design evaluation for fabrication company in the Philippines," in *Proc. Int. Conf. Ind. Eng. Oper. Manag.*, 2016, pp. 23–25. [Online]. Available: http://ieomsociety.org/ieomdetroit/pdfs/167.pdf

[22] N. Nassar and M. Austin, "Model-based systems engineering design and trade-off analysis with RDF graphs," *Procedia Comput. Sci.*, vol. 16, pp. 216–225, Jan. 2013 [Online]. Available: https://doi.org/10.1016/j.procs.2013.01.023
[23] N. S. Sigurdarson, T. Eifler, and M. Ebro, "Functional trade-offs in the mechanical design of integrated products—Impact on robustness and optimisability," *Proc. Design Soc., Int. Conf. End. Design*, vol. 1, no. 1, pp. 3491–3500, 2019. [Online]. Available: https://doi.org/10.1017/dsi.2019.356
[24] C. Chow, "On optimum recognition error and reject tradeoff," in *IEEE Trans. Inf. Theory*, vol. 16, no. 1, pp. 41–46, Jan. 1970. [Online]. Available: https://doi.org/10.1109/tit.1970.1054406
[25] G. Fumera and F. Roli, "Analysis of error-reject trade-off in linearly combined multiple classifiers," *Pattern Recognit.*, vol. 37, no. 6, pp. 1245–1265, 2004. [Online]. Available: https://doi.org/10.1016/j.patcog.2003.12.005
[26] G. Fumera, F. Roli, and G. Giacinto, "Reject option with multiple thresholds," *Pattern Recognit.*, vol. 33, no. 12, pp. 2099–2101, 2000. [Online]. Available: https://doi.org/10.1016/S0031-3203(00)00059-5
[27] K. Tumer and J. Ghosh, "Analysis of decision boundaries in linearly combined neural classifiers," *Pattern Recognit.*, vol. 29, no. 2, pp. 341–348, 1996. [Online]. Available: https://doi.org/10.1016/0031-3203(95)00085-2
[28] B. Biggio, G. Fumera, G. L. Marcialis and F. Roli, "Statistical meta-analysis of presentation attacks for secure multibiometric systems," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 3, pp. 561–575, Mar. 2017. [Online]. Available: https://doi.org/10.1109/tpami.2016.2558154
[29] M. Micheletto, G. Orrù, R. Casula, D. Yambay, G. L. Marcialis, and S. Schuckers, "Review of the fingerprint liveness detection (LivDet) competition series: From 2009 to 2021," in *Handbook of Biometric Anti-Spoofing*. Singapore: Springer Nat., 2023, pp. 57–76. [Online]. Available: https://doi.org/10.1007/978-981-19-5288-3_3
[30] R. Casula, M. Micheletto, G. Orrú, G. L. Marcialis, and F. Roli, "Towards realistic fingerprint presentation attacks: The screenspoof method," *Pattern Recognit. Lett.*, vol. 171, pp. 192–200, Jul. 2023. [Online]. Available: https://doi.org/10.1016/j.patrec.2022.09.002
[31] R. Casula et al., "LivDet 2021 fingerprint liveness detection competition—Into the unknown," in *Proc. IEEE Int. Joint Conf. Biom. (IJCB)*, 2021, pp. 1–6. [Online]. Available: https://doi.org/10.1109/ijcb52358.2021.9484399
[32] E. K. Antonsson and K. N. Otto, "Imprecision in engineering design," *J. Vib. Acoust.*, vol. 117, pp. 25–32, Jun. 1995. [Online]. Available: https://doi.org/10.1115/1.2838671.

Marco Micheletto (Member, IEEE) is currently an assistant professor with the Pattern Recognition and Applications Laboratory (PRA Lab), University of Cagliari. His research interests include integration of fingerprint comparison systems with presentation attack detector, fingerprint liveness detection, electroencephalography signal processing for biometric purposes and deepfake detection.

Gian Luca Marcialis (Senior Member, IEEE) is currently an Associate Professor of Computer Engineering with the University of Cagliari, Italy, with national habilitation to full professorship. He is the Head of the Biometric Unit with the Pattern Recognition and Applications Laboratory leaded by Prof. F. Roli. His research interests include biometric-based personal recognition; in particular, fingerprint and face recognition, multimodal fusion, self update algorithms, EEG-based features, and behavioral detection in crowds. He acts as a referee for the main international journals and conferences on pattern recognition, biometrics, and image processing. He also acts as an external project referee for public and private institutions. He is the Chair of the International Fingerprint Liveness Detection Competition, aimed at assessing biennially the state of the art on fingerprint presentation attack detection. He is an IAPR Member.

Open Access funding provided by 'Universitá degli Studi di Cagliari' within the CRUI CARE Agreement