

Biometric Presentation Attack Detection: Beyond the Visible Spectrum

Ruben Tolosana[✉], Marta Gomez-Barrero[✉], Christoph Busch[✉], and Javier Ortega-Garcia, *Fellow, IEEE*

Abstract—The increased need for unattended authentication in multiple scenarios has motivated a wide deployment of biometric systems in the last few years. This has in turn led to the disclosure of security concerns specifically related to biometric systems. Among them, presentation attacks (PAs, i.e., attempts to log into the system with a fake biometric characteristic or presentation attack instrument) pose a severe threat to the security of the system: any person could eventually fabricate or order a gummy finger or face mask to impersonate someone else. In this context, we present a novel fingerprint presentation attack detection (PAD) scheme based on *i)* a new capture device able to acquire images within the short wave infrared (SWIR) spectrum, and *ii)* an in-depth analysis of several state-of-the-art techniques based on both handcrafted and deep learning features. The approach is evaluated on a database comprising over 4700 samples, stemming from 562 different subjects and 35 different presentation attack instrument (PAI) species. The results show the soundness of the proposed approach with a detection equal error rate (D-EER) as low as 1.35% even in a realistic scenario where five different PAI species are considered only for testing purposes (i.e., unknown attacks).

Index Terms—Biometrics, presentation attack detection, deep learning, CNN, SWIR, fingerprint.

I. INTRODUCTION

THERE is an increasing demand in the current society for automatic and reliable authentication of individuals in a wide number of scenarios. To address this need, biometric recognition systems based on the individuals' biological (e.g., iris or fingerprint) or behavioural (e.g., signature or voice) characteristics have been consolidated as a reliable paradigm in the last decades. Their advantages over traditional authentication methods (e.g., no need to carry tokens or memorise

passwords, they are harder to circumvent and provide at the same time a stronger link between the subject and the action or event), have allowed a wide deployment of biometric systems, including large-scale national and international initiatives such as the Unique ID program of the Indian government [1] or the Smart Border project of the European Commission [2].

In spite of their numerous advantages, biometric systems are vulnerable to external attacks as any other security-related technology. Among all possible attack points defined in [3]–[5], the biometric capture device is probably the most exposed one: an eventual attacker requires no knowledge about the inner operating of the system in order to break the system. Instead, one can simply present the capture device with a *presentation attack instrument* (PAI), such as a gummy finger or a fingerprint overlay, in order to interfere with its intended behaviour. The main goal might be to impersonate someone else (i.e., active impostor) or to avoid being recognised (i.e., identity concealer). These attacks are known in the ISO/IEC 30107 [5] as *presentation attacks* (PAs).

Given the severe security threat posed by such PAs, the development of automatic techniques which are able to distinguish between bona fide (i.e., real or live) presentations and access attempts carried out by means of PAIs has become of the utmost importance [6], [7]. Referred to as *presentation attack detection* (PAD) methods, research in this area has been recently funded by several international projects like the European Tabula Rasa [8] and BEAT [9], or the more recent US ODIN research program [10]. Together with the organisation of the LivDet – liveness detection competition series on iris and fingerprint [11], [12], where the number of participants has been increasing year after year (up to 17 algorithms submitted in 2017), these initiatives have fostered a considerable number of publications on PAD for different biometric characteristics, including iris [13], fingerprint [14], [15], face [16], or handwritten signature [17].

The initial approaches to PAD were based on the so-called handcrafted features, such as texture descriptors or motion analysis [6], [18]. However, deep learning (DL) has become a thriving topic in the last years [19]–[21], and biometric recognition in general, and PAD in particular, are not an exception. DL allows systems to learn from experience and understand the world in terms of a hierarchy of simpler units, thereby enabling significant advances in complex domains. The main reasons to understand its high deployment lie on the increasing amount of available data and the evolution of graphical processing units (GPU), which in turn allows the

Manuscript received March 21, 2019; revised June 24, 2019 and August 1, 2019; accepted August 6, 2019. Date of publication August 12, 2019; date of current version December 11, 2019. This work supported in part by the Office of the Director of National Intelligence (ODNI), Intelligence Advanced Research Projects Activity (IARPA), under Contract 2017-17020200005, in part by the German Federal Ministry of Education and Research (BMBF), in part by the Hessen State Ministry for Higher Education, Research and the Arts (HMWK) within the National Research Centre for Applied Cybersecurity (CRISP), in part by the project BIBECA (MINECO/FEDER) under Grant RTI2018-101248-B-I00, and in part by the Bio-Guard (Ayudas Fundacion BBVA a Equipos de Investigacion Cientifica 2017). The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Clinton Fooke. (Corresponding author: Ruben Tolosana.)

R. Tolosana and J. Ortega-Garcia are with the Biometrics and Data Pattern Analytics (BiDA) Lab, Universidad Autonoma de Madrid, 28049 Madrid, Spain (e-mail: ruben.tolosana@uam.es; javier.ortega@uam.es).

M. Gomez-Barrero and C. Busch are with the da/sec—Biometrics and Internet Security Research Group, Hochschule Darmstadt, 64295 Darmstadt, Germany (e-mail: marta.gomez-barrero@h-da.de; christoph.busch@h-da.de).

Digital Object Identifier 10.1109/TIFS.2019.2934867

successful training of deep architectures. However, the belief that DL schemes can be only used in tasks with massive amounts of available data has changed in the last years thanks to the combination of both pre-trained models and transfer learning techniques. These consist in network models which are first trained for a given task with large available databases, including any kind of images and not only those expected for the problem at hand. These pre-trained models are subsequently retrained (a.k.a. fine-tuned, adapted) for a different task for which data are usually scarce [22], [23].

All the aforementioned advances have allowed the deployment of DL architectures in many different fields, including biometric recognition [24], [25]. More specifically, convolutional neural networks (CNNs) and deep belief networks (DBNs) have been used for fingerprint PAD purposes, based either on the complete fingerprint samples [26]–[28] or on a patch-wise manner [29]–[31].

As it will be described in more detail in Sect. III, DL based PAD approaches have boosted the performance over common PAD benchmarks from the LivDet competitions, achieving detection rates over 90%. Such high accuracy rates indicate the valuable contributions of the existing approaches. However, the LivDet databases comprise altogether up to 11 different materials for the fabrication of PAIs, even though the choice for the attacker is much wider based on commercial products readily available even online. As a consequence, other databases, comprising a larger number of materials for the fabrication of the PAIs, should be explored. Very few studies have considered this issue, including a database comprising over twelve different PAI species in [31], and 21 materials in [32]. We address this issue with the acquisition of a database including 35 different PAI species, within the US ODIN research program [10].

In addition, there is one question that remains mostly unanswered in the literature: Once a deep neural network is trained on a large number of PAI species, will future unknown attacks also be detected? Some previous studies have evaluated this challenging scenario using handcrafted feature approaches. In [33], the authors evaluated this scenario using standard databases taken from LivDet 2009 competition. The results achieved showed a high degradation of the system performance when unknown attacks were presented to the system. In [34], the authors designed a scheme for automatic adaptation of PAD systems in order to detect novel unknown attacks. The experiments conducted on the LivDet 2011 database suggested that up to 46% improvement can be achieved considering their proposed adaptive approach. Later on, the same authors proposed in [35] the use of Weibull-calibrated SVM (W-SVM) algorithm in order to improve the detection of unknown attacks. The experiments conducted over LivDet 2011 database achieved up to 44% improvement. Another interesting evaluation was carried out in [36]. In that work the authors compared the performance of supervised and semi-supervised approaches that rely solely on the bona fide samples. The results obtained remarked the true vulnerability of the biometric systems. Finally, novel PAI species were considered in [32], noting that the error rates were multiplied by a factor of six when unknown PAI species were tested, with respect

to the detection accuracy reached on known attacks. This challenging scenario has also been studied in other biometric characteristics such as face [37] and iris [38]. Therefore, we can conclude that additional research efforts are needed in this area. To further tackle these issues, some researchers have considered other sources of information different from traditional capture devices [13], [15]. More specifically, the use of multi-spectral near infrared (NIR) technologies has been studied for face [39], [40] and fingerprint [41], [42].

In this new context, a recent trend for both biometric PAD and face recognition enhancement is based on skin detection. On the one hand, non-skin materials (e.g., a mask or a scarf) can be masked for recognition purposes. On the other hand, such materials can be considered a PA attempt. This will be the fundamental idea followed in this article: PAD is regarded as the problem of discriminating skin vs. non-skin materials. In order to overcome one of the main challenges of skin detection, namely, the plurality of different skin colours [43], we choose the short wave infrared (SWIR) band as a promising information source. It has been proved that human skin shows characteristic remission properties for multi-spectral SWIR wavelengths, which are independent of the capture subject's age, gender or skin type [44]. In fact, several approaches have been proposed for face recognition in the infrared domain [45], [46]. In particular, for surveillance purposes, the SWIR range has been analysed by several research groups, either as solely source of information or in combination with visible light images [47]–[49]. The advantages of SWIR are mostly its robustness in challenging environmental conditions (e.g., with fog or at night time). In addition, the benefits of multi-spectral hand based recognition within the SWIR bands were studied in [50], outperforming state-of-the-art recognition approaches.

For the particular task of PAD, the characteristic remission properties of the human skin observed in the multi-spectral SWIR band were exploited in [40] for facial PAD, achieving a 99% detection accuracy. A similar approach was analysed in [51] over a small fingerprint database, comprising 60 samples. It was shown that the method was able to detect all 12 PAIs except for one. In addition, a preliminary DL approach based on a pre-trained CNN model was tested on the same database in [52], achieving perfect detection rates over the small preliminary database.

Keeping these thoughts in mind, the main contributions of this work compared with the state-of-the-art can be summarised as follows:

- We present a novel fingerprint PAD scheme based on *i)* a new capture device able to acquire images within the SWIR spectrum, and *ii)* an in-depth analysis of several state-of-the-art techniques based on both handcrafted and deep learning features. A final fusion of both handcrafted and deep learning features is carried out for completeness, as depicted in Fig. 1.
- We study multiple state-of-the-art CNN architectures for fingerprint PAD purposes. Both networks trained from scratch (i.e., a residual network [53]) and also pre-trained models (i.e., MobileNet [54] and VGG19 [55]) are analysed. In addition, two different approaches are

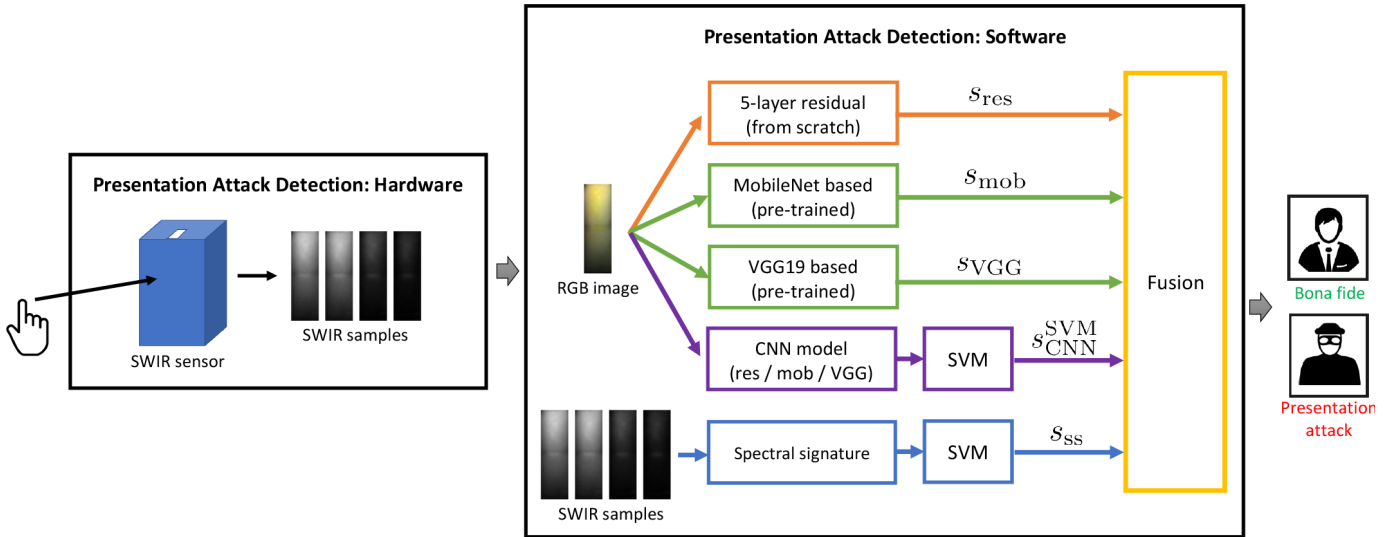


Fig. 1. General diagram of the proposed PAD method. On the left, the capture device acquires the samples at four different wavelengths within the SWIR spectrum. On the right, several software approaches have been proposed, namely: *i*) Three different state-of-the-art CNN architectures have been tested as an end-to-end solution, *ii*) the features output by the CNN models have been used to feed an SVM, *iii*) handcrafted features (i.e., spectral signatures) have been extracted, and *iv*) a final fusion of the aforementioned algorithms has been evaluated for completeness.

considered: *i*) using the CNNs as an end-to-end solution, and *ii*) utilising the CNNs as a feature extractor and carrying out classification with support vector machines (SVMs).

- We evaluate our proposed PAD approach on a large database comprising over 4700 samples, stemming from 562 different subjects and 35 different PAIs. We include a benchmark of deep learning approaches with high-performing handcrafted features [51]. Our proposed approach has achieved a final 1.35% detection equal error rate (D-EER), outperforming the state-of-the-art.
- We also evaluate the robustness of our proposed PAD approach against new PAIs not used during the development of the system (i.e., unknown attacks). Our final fused system is able to correctly detect all unknown attacks, proving its high generalisation capacity to new PAI species that can appear in the future.
- We review the state-of-the-art on fingerprint PAD based on either *i*) non-conventional sensors, or *ii*) conventional sensors in combination with deep learning approaches.

Finally, it should be noted that, being a skin detection based method, the proposed PAD technique can be applied not only to fingerprints but also to other biometric characteristics, such as the face, the hand, or the periocular regions.

The rest of the article is organised as follows. Sect. II presents the main terms which will be used in the remainder of the article. Related works on fingerprint PAD are summarised in Sect. III. Sects. IV and Sect. V describe the proposed approach. The evaluation framework is presented in Sect. VI, and the results discussed in Sect. VII. Final conclusions are drawn in Sect. VIII.

II. DEFINITIONS

In the following, we include the main definitions stated within the ISO/IEC 30107-3 standard on biometric

presentation attack detection - part 3: testing and reporting [56], which will be used throughout the article:

Bona fide presentation: “interaction of the biometric capture subject and the biometric data capture subsystem in the fashion intended by the policy of the biometric system”. That is, a normal or genuine presentation.

Presentation attack (PA): “presentation to the biometric data capture subsystem with the goal of interfering with the operation of the biometric system”. That is, an attack carried out on the capture device to either conceal your identity or impersonate someone else.

Presentation attack instrument (PAI): “biometric characteristic or object used in a presentation attack”. For instance, a silicone 3D mask or an ecoflex fingerprint overlay.

PAI species: “class of presentation attack instruments created using a common production method and based on different biometric characteristics”.

In order to evaluate the vulnerabilities of biometric systems to PAs, the following metrics should be used:

Attack Presentation Classification Error Rate (APCER): “proportion of attack presentations using the same PAI species incorrectly classified as bona fide presentations in a specific scenario”.

Bona fide Presentation Classification Error Rate (BPCER): “proportion of bona fide presentations incorrectly classified as presentation attacks in a specific scenario”.

Derived from the aforementioned metrics, the detection equal error rate (D-ERR) is defined as the error rate at the operating point where $APCER = BPCER$.

III. RELATED WORKS

In this section we summarise the key works on fingerprint PAD for both non-conventional optical or capacitive sensors (see Sect. III-A and Table I) and conventional sensors in combination with DL approaches (see Sect. III-B and Table II).

TABLE I
SUMMARY OF THE MOST RELEVANT METHODOLOGIES FOR FINGERPRINT PAD BASED ON NON-CONVENTIONAL SENSORS

Year	Spectrum	Ref.	Description	Performance	Database (# PAIs)
2008	430 – 630 nm	[57]	Wavelet transform	APCER = 0.9% BPCER = 0.5%	Unavailable DB (49)
2011	400 – 1630 nm	[58]	Spectroscopic properties	-	Unavailable DB (0)
	OCT 400 – 850 nm	[42]	-	-	Unavailable DB (-)
2018	1200 – 1550 nm	[51]	Multi-spectral signatures	APCER = 5.7% BPCER = 0.0%	Unavailable DB (12)
		[52]	Pre-trained VGG19 model	APCER = 0.0% BPCER = 0.0%	Unavailable DB (12)
	1310 nm (LSCI)	[59]	Texture descriptors	APCER = 10.97% BPCER = 0.84%	Self-acquired DB (32)
2019	Finger vein 940 nm	[60]	Pyramid Local Binary Patterns (PLBP)	APCER \approx 10% BPCER \approx 1%	Self-acquired DB (32)
	1200 – 1550 nm		Proposed Approach	APCER \approx 2% BPCER = 0.2%	Self-acquired DB (35)

For further details on fingerprint PAD, the reader is referred to [14], [15].

It should be noted that, in addition to the metrics defined in Sect. II, two different metrics are used in the LivDet competitions [11], [12]. The Average Classification Error Rate (ACER) is defined as the average of the APCER and the BPCER for a pre-defined decision threshold δ :

$$\text{ACER}(\delta) = \frac{\text{APCER}(\delta) + \text{BPCER}(\delta)}{2} \quad (1)$$

It should be noted that averaging APCER and BPCER has been deprecated in ISO/IEC 30107-3. The ACER is reported here for the only purpose to relate our results to the LivDet competition, where ACER has been used.

The detection accuracy (Acc.) refers to the rate of correctly classified bona fide and PAs at $\delta = 0.5$:

$$\text{Acc}(\delta) = \frac{1}{\# \text{ samples}} \cdot \left\{ (1 - \text{APCER}(\delta)) \cdot \{\# \text{ PA samples}\} + (1 - \text{BPCER}(\delta)) \cdot \{\# \text{ BF samples}\} \right\} \quad (2)$$

These metrics will be used in Table II where needed.

A. Non-Conventional Fingerprint Sensors

To the best of our knowledge, the pioneering work on fingerprint multi-spectral PAD with non-conventional capacitive or optical sensors was carried out by Rowe *et al.* in [57]. The presented, and now widely used Lumidigm sensor, captures multi-spectral images in four different wavelengths (i.e., 430, 530, and 630 nm, as well as white light). In their work, the authors studied the PAD capabilities of the combined images using absolute magnitudes of the responses of each image to dual-tree complex wavelets. In a self-acquired database including 49 PAI species, they obtained an APCER of 0.9% for a BPCER of 0.5%. Even if these results are remarkable, the PAD methods used are not described and not many details about the acquired database or the experimental

protocol are available. Therefore, it is difficult to establish a fair benchmark.

Three years later, Hengfoss *et al.* analysed extensively the spectroscopic properties of living against the cadaver fingers using four wavelengths between 400 nm and 1630 nm [58]. However, no PAIs were analysed in their work. Later that year, Chang *et al.* studied in [42] the complex properties of the skin, which differentiate it from PAIs, using optical coherence tomography (OCT) and nine different wavelengths between 400 nm and 850 nm. A single volunteer provided the bona fide and PA samples, and not many details about the algorithms used were reported.

More recently, in 2018, some preliminary PAD studies were carried out in [51], [52] on a small database, comprising a total of 60 samples and 12 different PAI species, which was acquired at University of South California within the BATL project [61]. Gomez-Barrero *et al.* extracted multi-spectral signatures from four different wavelengths in SWIR spectrum, achieving an APCER = 5.7% and a BPCER = 0%. In this case, all classification errors stem from a single PAI made with orange playdoh. In a subsequent work on the same database, Tolosana *et al.* used a pre-trained VGG19 CNN model [55] for PAD purposes. In this case, all 60 samples were correctly classified (i.e., APCER = BPCER = 0%).

Keilbach *et al.* also analysed in [59] the PAD capabilities of laser speckle contrast images (LSCI) over a larger database, also acquired within the BATL project and comprising 32 PAIs and more than 750 samples. In this case, several descriptors were extracted from the LSCI sequences, including the well-known local binary patterns (LBP) or the histogram of oriented gradients (HOG). The final cascaded score level fusion yielded an APCER = 10.97% for a BPCER = 0.84%.

Finally, Kolberg *et al.* studied in [60] the feasibility of detecting fingerprint PAs with finger vein images over the same database analysed in [59]. To that end, texture information was acquired at different resolution levels using Gaussian pyramids and LBP (PLBP). Combining the information of up to 16 levels with SVMs, an APCER \approx 10% can be achieved

TABLE II
SUMMARY OF THE MOST RELEVANT METHODOLOGIES FOR FINGERPRINT PAD BASED ON CONVENTIONAL SENSORS AND DL

Category	Year	Ref.	Description	Performance	Database (# PAIs)
Full Sample	2015	[62]	CNN optimisation	Acc. = 98.97%	LivDet 2013 (7)
		[26]	Pre-trained CNNs (Best: VGG)	ACER = 2.90%	LivDet 2009-13 (8)
	2016	[28]	DBN with RBMs	1 - ACER = 97.10%	LivDet 2013 (7)
		[63]	Pre-trained CNNs and Siamese networks (Best: GoogLeNet)	APCER = 4.3% BPCER = 2.5%	LivDet 2011-13 (8)
ROI	2017	[64]	CNNs + ROI and PCA optimization and SVM classification	ACER = 4.57% (2011) ACER = 7.25% (2013)	LivDet 2011-13 (8)
Patch-wise	2015	[65]	DCNN (CiFar10-Net + FingerNet)	ACER = 0.88% (2011) ACER = 0.90% (2013)	LivDet 2011-13 (8)
	2016	[66]	CNN trained from scratch	ACER = 3.42%	LivDet 2009 (Identix, 3)
		[27]	Contrast enhancement + Ad hoc CNN	ACER = 0.20%	ATVS FP (2)
	2017	[30]	Deep Boltzmann Machine	Acc. = 85.96%	LivDet 2013 (7)
		[29]	Pre-trained AlexNet + Data augmentation and log-likelihood	ACER = 4.63% (2011) ACER = 1.90% (2013)	LivDet 2011-13 (8)
		[67]	Deep triplet embedding	ACER = 1.74%	LivDet 2009-13 (8)
	2018	[31]	Pre-trained MobileNet + Minutiae patches	ACER = 0.96% ACER = 2.00%	LivDet 2011-15 (11) Own DB (12)
		[68]	Fully CNN (SqueezeNet) + Data augmentation	ACER = 1.43%	LivDet 2011-15 (11)
Deep Fusion	2017	[69]	Texture based features and DNN fusion	ACER \approx 1.70%	LivDet 2009-13 (8)

for a BPCER \approx 1%, thereby showing the main limitation of these studies: both LSCI and finger vein samples acquire information below the skin and are thus unable to detect thin transparent overlays.

B. Deep Learning for Conventional Sensors

The DL based fingerprint PAD approaches proposed in the literature can be widely classified depending on the input of the networks into: *i*) using the full samples as input to the network, *ii*) cropping the region of interest (ROI) and feeding it to the network, and *iii*) extracting patches from the ROI as input. Moreover, *iv*) some articles use the network for feature level fusion of handcrafted descriptors. In the following, we summarise the main studies of each category.

1) *Full Samples*: To the best of our knowledge, the first work on fingerprint PAD based on deep learning algorithms was presented in 2015 by Menotti *et al.* [62]. The authors proposed two different CNN optimisation approaches for the particular purpose of PAD. On the one hand, the architecture was optimised with feedforward convolutional operations and hyperparameter optimisation. On the other hand, the inner weights of the network were optimised via back-propagation. Both techniques were tested on iris, face and fingerprint benchmarks, thus proving the generalisation capabilities of the proposal. Their best fingerprint results achieved an average detection accuracy, Acc., across the four fingerprint sensors of LivDet 2013 of 98.97%.

A year later, three different approaches were proposed. Nogueira *et al.* [26] tested three different CNNs, namely: *i*) the pre-trained VGG [55], *ii*) the pre-trained Alexnet [70], and *iii*) a CNN with randomly initialised weights and trained from scratch. They compared the ACER obtained with the networks over the LivDet 2009, 2011 and 2013 databases to a classical state-of-the-art algorithm based on LBP. In the evaluation, the best detection performance was achieved using a VGG pre-trained model and data augmentation (average ACER = 2.9%), with a clear improvement with respect to LBP (average ACER = 9.6%). It should also be noted that the ACER decreased between 25% and 50% (relative decrease) for all three networks tested when data augmentation was used.

Then, Kim *et al.* analysed the use of deep belief networks based on superimposed restricted Boltzmann machines (RBMs) [28]. The global network was trained in a two-stage manner with layer-wise greedy training and fine-tuning with labelled inputs. On LivDet 2013, they achieved a detection accuracy Acc. of 97.10%, noting again the considerable improvement achieved with data augmentation.

Finally, Marasco *et al.* explored in [63] two different pre-trained CNNs: *i*) CaffeNet [70], and *ii*) GoogLeNet [71]. The performance of such networks was compared with a Siamese network, which optimised a metric distance to yield high bona fide - PA distances and low bona fide - bona fide distances. In a thorough evaluation on LivDet 2011 and 2013, a detection accuracy over 96% was achieved for GoogLeNet, closely followed by the other networks. The authors showed

an accuracy decrease when dealing with either unknown attack or a cross-sensor scenario.

2) *ROI*: In 2017, Yuan *et al.* followed a different approach to optimise the performance of CNN models [64]. First, only the ROI was fed to the network. Then, principal component analysis (PCA) was introduced for each convolutional and pooling operation in order to discard non-relevant information. Finally, the output was classified with SVMs. This way, no data augmentation was required to achieve a 4.57% ACER over LivDet 2013, thereby outperforming other existing approaches.

3) *Patch-Wise*: In 2015, a different two-step approach was proposed by Wang *et al.* [65]. First, the ROI of the fingerprint was segmented. Then, two deep CNNs (DCCNs) were used in a patch-wise manner: *i*) the CiFar10-Net [72], and *ii*) the self-developed Finger-Net, yielding an ACER under 1% over LivDet 2011 and 2013.

In 2016, Park *et al.* extracted random patches from the fingerprint samples and trained a CNN from scratch in [66], achieving an ACER = 3.4% over the Identix subset of LivDet 2009.

In 2017, Jang *et al.* proposed contrast enhancement and block-wise processing of the fingerprint to improve the state-of-the-art results achieved with DL [27]. The blocks were then combined with a majority voting rule. They also designed a CNN from scratch inspired in the VGG19 model, and evaluated the proposed approach over the ATVS fake fingerprint DB [73]. An ACER of 0.2% was reported.

Souza *et al.* analysed again in [30] the use of Boltzmann machines, this time in a patch-wise manner and using a majority voting rule. In particular, they used deep Boltzmann machines (DBMs), which can learn more complex and internal representations from a small number of labelled samples. The accuracy obtained over LivDet 2013 was 85.96%.

Following this patch-wise trend, Toosi *et al.* tested in [29] the accuracy of AlexNet with data augmentation. For classification, the scores were calibrated using log-likelihood ratios. The average ACER on LivDet 2011 and 2013 is 3.26%.

Similarly, Pala *et al.* tested the feasibility of using deep triplet embedding for PAD purposes [67]. In contrast to Siamese networks, this method requires no enrolment database, since the triplets are selected from patches within the input sample. Over LivDet 2009 to 2013, an ACER of 1.74% was reported. The robustness to unknown attacks was also evaluated on LivDet 2013, achieving ACERs much lower than other approaches (e.g., 0.7% vs. 1.4% for Siamese networks for the modasil PAIs).

In 2018, Chugh *et al.* presented in [31] a different way to extract fingerprint patches: around the minutiae. The idea behind this patch computation is the fact that PAIs can present spurious minutiae, which can be surrounded by a distinct texture. Therefore, these patches were fed to the MobileNet pre-trained network [54]. The detection performance was evaluated on LivDet 2011 to 2015, achieving a remarkable ACER of 0.96% on average. However, the ACER increased to 2.0% for a self-acquired database, comprising a larger number of PAIs (12).

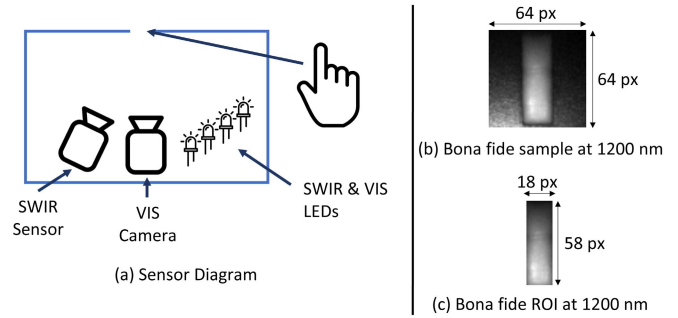


Fig. 2. Finger sensor diagram. **Left**: a diagram of the inner components: two different sensors for the SWIR images and the visible (VIS) light images, and the corresponding LEDs. **Right**: Original sample and the corresponding ROI for a bona fide at 1200 nm.

In the same year, Park *et al.* developed in [68] a fully CNN based on the Fire module of SqueezeNet [74]. They analysed different patch sizes and compared the common voting method to an optimal thresholding approach, which yielded a better performance: an ACER of 1.43% over LivDet 2011 to 2015.

4) *Deep Fusion*: Toosi *et al.* proposed in [69] a completely different approach to use DL for fingerprint PAD. Instead of using deep networks for feature extraction, ten different handcrafted descriptors, including the well-known local phase quantization (LPQ), binary statistical features (BSIF) or scale invariant feature transform (SIFT) were fed to a self-developed deep network (Spidernet) for final fusion and classification. The performance was compared with classical fusion approaches such as SVMs and AdaBoost, achieving ACERs around 1.6-1.8% for LivDet 2009 to 2013.

IV. PRESENTATION ATTACK DETECTION METHODOLOGY: HARDWARE

The finger SWIR capture device used in the present work was developed within the BATL project [61] in cooperation with our project partners. A general diagram of its inner components is included in Fig. 2 (a). As it may be observed, the camera and lens are placed inside a closed box, which includes an open slot on the top, about 30 cm away from the cameras. When the finger is placed there, all ambient light is blocked and therefore only the desired wavelengths are considered during the acquisition. In particular, we have used a Hamamatsu G11097-0606S InGaAs area image sensor, which captures 64×64 px images, with a 25 mm fixed focal length lens optimised for wavelengths within 900 – 1700 nm. More specifically, the following SWIR wavelengths are selected for PAD purposes: $\lambda_1 = 1200$ nm, $\lambda_2 = 1300$ nm, $\lambda_3 = 1450$ nm, and $\lambda_4 = 1550$ nm. These are similar to the wavelengths considered in [40] for the skin vs. non-skin facial classification.

An example of the acquired images for a bona fide sample is shown in Fig. 2 (b) for the 1200 nm wavelength. As it may be observed, before applying any PAD algorithm, the region of interest (ROI) (i.e., the central finger-slot region corresponding to the open slot where the finger is placed) needs to be extracted from the background. Given that the finger is always placed over the fixed open slot, and the camera does not move, a simple fixed size cropping can be applied. The ROI

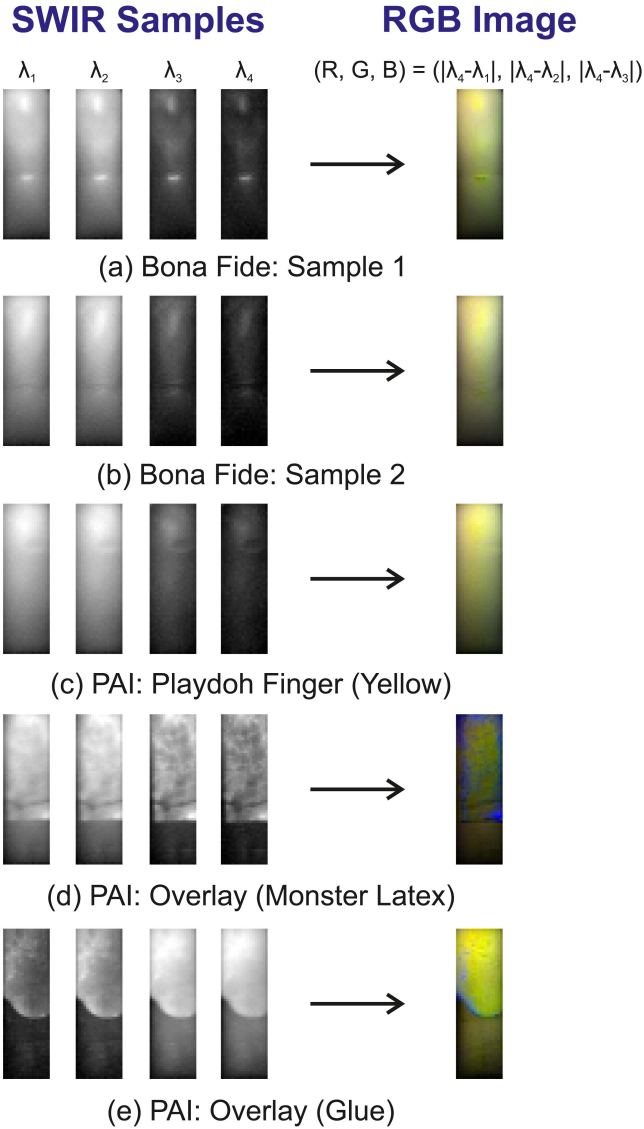


Fig. 3. Examples of bona fides and PAs acquired by the SWIR sensor and the RGB image created for the input of the deep neural network systems (see Eq. 6).

corresponding to Fig. 2 (b) with a size of 18×58 px is depicted in Fig. 2 (c).

Finally, the four wavelength samples acquired from two bona fides Fig. 3 (a, b), and three PAIs Fig. 3 (c to e) fabricated with different materials are included in Fig. 3. As it may be observed, the playdoh finger shows some similarities with respect to the bona fide presentations (i.e., a similar change of intensity across wavelengths), which will make the PAD task harder. However, the change trend is completely different for the other two PAIs, thereby making it easier to discriminate them from bona fide presentations.

In addition to the SWIR images captured by the device, fingerprint verification can be carried out with contactless finger photos acquired in the visible spectrum with a 1.3 MP camera and a 35 mm VIS-NIR lens, which are placed next to the SWIR sensor within the closed box (see Fig. 2 (a)). As shown in [60], commercial off-the-shelf systems can extract minutiae

correctly from these samples, thereby granting compatibility with conventional fingerprint sensors.

V. PRESENTATION ATTACK DETECTION METHODOLOGY: SOFTWARE

This section describes the state-of-the-art software methods proposed in order to detect fingerprint PAs, as summarised in Fig. 1. Two different approaches are studied: *i*) handcrafted features, and *ii*) deep learning features. For both approaches, the information provided by the sensor described in Sect. IV is used as input.

In general, it should be noted that each individual score s_i generated by the individual PAD algorithms needs to be transformed into a single range to allow the final fusion and a fair benchmark. In compliance with the ISO/IEC 30107-2 standard on biometric presentation attack detection – Part 2: data formats [75], we define $s_i \in [0, 100]$, where low values close to 0 will represent bona fide presentations and high values close to 100 will denote presentation attacks.

A. Handcrafted Features

As it was firstly proposed in [51], this method builds upon the raw spectral signatures of the pixels across all four acquired wavelengths, in order to capture the different properties attributed to skin (i.e., bona fide presentation) and non-skin (i.e., PAI) materials. In particular, for each pixel (x, y) , the SWIR sensor provides the raw spectral signature ss as follows:

$$ss(x, y) = \{i_1(x, y), \dots, i_N(x, y)\} \quad (3)$$

where $i_n(x, y)$ represents the intensity value of the pixel for the n -th wavelength λ_n . In our particular case study, $N = 4$.

However, this original representation of the sensor is vulnerable to illumination changes. Even if they have been minimised in the sensor due to only having the finger slot open to the outer world, thinner fingers for instance can let some tiny amounts of light through. In addition, since our final goal is to capture the distinct trends across the different wavelengths shown in Fig. 3, only the differences among wavelengths will be used as our set of handcrafted features. Therefore, for each pixel, the final normalised difference feature vector $\mathbf{d}(x, y)$ is computed as follows:

$$\mathbf{d}(x, y) = \{d[i_a, i_b](x, y)\}_{1 \leq a < b \leq N} \quad (4)$$

$$d[i_a, i_b](x, y) = \frac{i_a(x, y) - i_b(x, y)}{i_a(x, y) + i_b(x, y)} \quad (5)$$

with $-1 \leq d[i_a, i_b](x, y) \leq 1$. In other words, the normalised differences among all possible wavelength combinations are computed. For our case study with $N = 4$, a total number of six differences are calculated per pixel. Finally, these normalised difference feature vectors $\mathbf{d}(x, y)$ will be used to classify skin vs. non-skin pixels using an SVM.

The procedure described so far performs a pixel-wise classification. Hence, the final score s_{ss} returned by the PAD method will be the proportion of non-skin pixels of the sample ROI in a range of 0 to 100.

B. Deep Learning Features

CNNs have been one of the most successful deep neural network architectures in the last years. Some of their key design principles were drawn from the findings of the Neurophysiologists Nobel Prize awardees David Hubel and Torsten Wiesel in the field of human vision [19]. Traditional (a.k.a. plain) CNN based systems are mainly composed of convolutional and pooling layers. The former extracts patterns from the images through the application of several convolutions in parallel to local regions of the images. These convolutional operations are carried out by different kernels that are adapted by the learning algorithm, assigning a different weight to each pixel of the local region of the image depending on the type of patterns to be extracted. Therefore, each kernel of one convolutional layer is focused on extracting different patterns, such as horizontal or vertical edges, over image patches whose size is determined by the dimension of the layer. The output of these operations produces a set of linear activations (a.k.a. feature map), which serve as input to nonlinear activations, such as the rectified linear activation function (ReLU). Finally, it is common to use pooling layers to make the representation invariant to small translations of the input. The pooling function replaces the output of the network at a certain region with a statistical summary of the nearby outputs, and facilitates the learning convergence. For instance, the max-pooling function selects the maximum value of the region.

As it was summarised in Fig. 1, in this study we explore the potential of deep learning features in comparison to handcrafted features by means of two different strategies: *i*) using CNNs as an end-to-end approach (i.e., for both feature extraction and classification), and *ii*) using CNNs as feature extractors in combination with SVMs for classification. In addition, two different training scenarios have been analysed, namely: *i*) training CNN models from scratch, and *ii*) adapting CNN pre-trained models.

For the input of the networks, and in order to consider the information provided by the four wavelengths captured by the sensor, we need to build a single RGB image. To that end, each dimension or channel of the RGB space will comprise information stemming from different SWIR wavelengths or combinations thereof. To maximise the discriminative power of the input images, we analysed which wavelengths provided a higher inter-class (i.e., between bona fide and PA presentations) and a lower intra-class (i.e., within the bona fide presentation samples) variation in terms of the heatmaps of the differences between samples. That is, to estimate the inter-class variability we computed the pixel wise difference of bona fide and PA samples, and for the intra-class variability, the differences between bona fide samples. The former should have high intensity values and the latter low values. After an exhaustive analysis of the different possible combinations, we defined the three dimensions as follows:

$$\text{image}(R, G, B) = (|i_{\lambda_4} - i_{\lambda_1}|, |i_{\lambda_4} - i_{\lambda_2}|, |i_{\lambda_4} - i_{\lambda_3}|) \quad (6)$$

Fig. 3 shows examples of bona fides and PAIs acquired by the SWIR sensor and the RGB image created following Eq. 6. Finally, RGB images are resized to the corresponding input size of the pre-trained networks using the nearest-neighbour

interpolation. All strategies have been implemented under the Keras framework using Tensorflow as back-end, with a NVIDIA GeForce GTX 1080 GPU. Adam optimizer is considered with a learning rate value of 0.0001 and a loss function based on binary cross-entropy. We now describe the details of each deep learning strategy studied in this work.

1) Training CNN Models From Scratch: The first approach is focused on training **residual CNNs** [53] from scratch. These networks have outperformed traditional (a.k.a. plain) networks in many different datasets such as ImageNet 2012 [76], CIFAR-10 [77], PASCAL VOC 2007/2012 [78] and COCO [79] for both image classification and object detection tasks. The peculiarity of this network is the insertion of shortcut connections every few stacked layers, converting the plain network into its residual version. The residual connections allow the use of deeper neural network architectures and at the same time decrease their training time significantly [53], [80].

Our proposed residual CNN is depicted in Fig. 4 (left). Batch normalization (BN) is applied right after each convolution and before the activation as described in [81]. All activation functions are based on ReLU except from the Sigmoid activation used in the last fully-connected layer, which provides output values between 0 and 1. This value is finally multiplied by 100 in order to obtain scores between 0 and 100.

2) Adapting Pre-Trained CNN Models: The second approach evaluates the potential of state-of-the-art pre-trained models for fingerprint PAD. In order to adapt the pre-trained models to our task, we replace and retrain the classifier (i.e., the fully-connected layers), and adapt the weights of the last convolutional layers to the fingerprint PAD task. The reason for adapting only the last convolutional layers lies on the fact that the first layers of the CNN extract more general features related to directional edges and colours, whereas the last layers of the network are in charge of extracting more abstract features related to the specific task. We propose to use both MobileNet and VGG19 network architectures pre-trained using the ImageNet database [54], [55]. This database contains more than one million images from 1000 different classes, thereby allowing the extraction of very robust features in the first layers [76].

Fig. 4 (middle) shows the architecture of our adapted **MobileNet** network. This architecture has been modified compared to the original version by removing some of the last convolutional layers in order to reduce the complexity of the features extracted. Furthermore, the fully-connected layers designed for the ImageNet classification task have been also removed. This network is based on depthwise separable convolutions, which factorise a standard convolution into: *i*) a depthwise convolution, and *ii*) a 1×1 convolution called pointwise convolution. Therefore, the depthwise convolution applies a single filter to each input channel, and the pointwise convolution subsequently applies a 1×1 convolution to combine the outputs of the depthwise convolution [54]. Downsampling is directly applied by the convolutional layers that have a stride of 2 (represented by $/2$ in the convolutional layers of Fig. 4). This network architecture allows to reduce both model

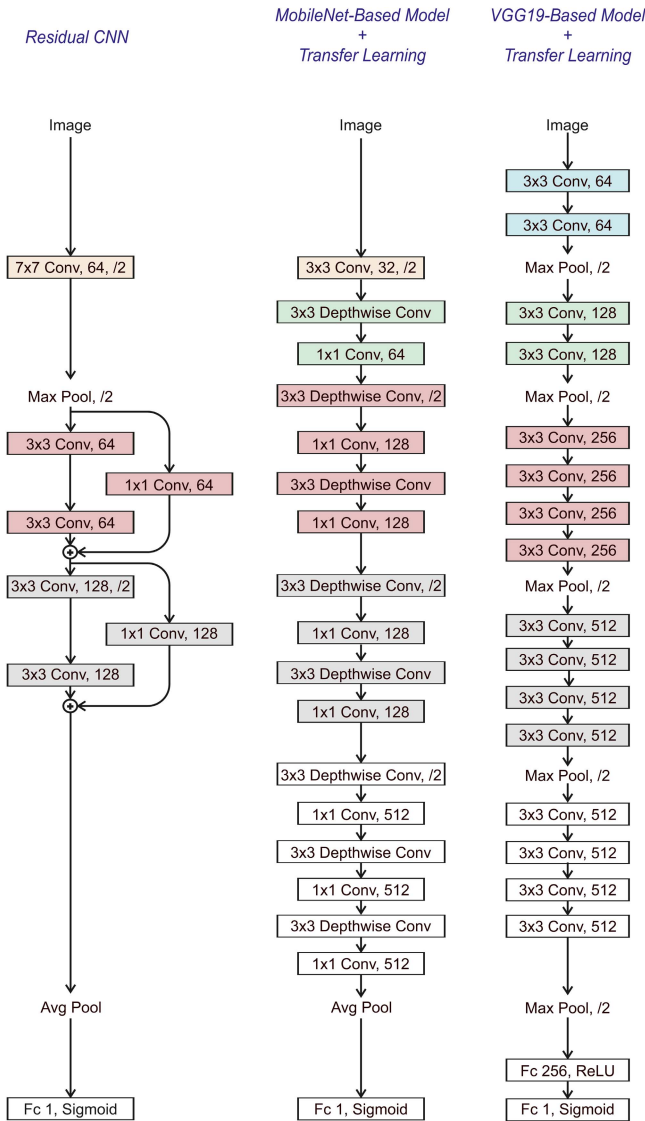


Fig. 4. Proposed network architectures. **Left:** the residual CNN trained from scratch using only the SWIR fingerprint database (319,937 parameters). **Middle:** the pre-trained MobileNet-based model (815,809 parameters). **Right:** the pre-trained VGG19-based model (20,155,969 parameters). Both middle and right networks are adapted using transfer learning techniques over the last white-background layers.

size and training/testing times, thus being a good solution for mobile and embedded vision applications. It has been tested in different datasets such as ImageNet [76], PlaNet [82], and COCO [79] with state-of-the-art results.

Finally, Fig. 4 (right) shows the architecture of the adapted **VGG19** network [55]. This architecture has also been modified replacing the last 3 fully-connected layers with 2 fully-connected layers (with a final sigmoid activation). This network architecture belongs to the family of traditional or plain networks and appeared before the residual and MobileNet configurations. Despite of that, and due to its simplicity, it is one of the most popular network architectures nowadays, providing very good results in many different competitions.

3) *Using CNNs as Feature Extractors:* In addition to the end-to-end approaches described in Sect. V-B.1 and V-B.2,

we also analyse the potential of adapting and using all the aforementioned CNNs (i.e., the residual CNN trained from scratch, and the adapted MobileNet and VGG19 models) as feature extractors. For this strategy, we consider the same network architectures described in Fig. 4, but removing the last fully-connected layers in order to use only the features provided by the last convolutional layer (after the Average or Max pool layers, respectively). Then, these features are transformed to the range [0, 1] and subsequently used to train an SVM for final classification purposes.

C. Fused Approach

Finally, we analyse to which extent the proposed algorithms complement each other to enhance the final fingerprint PAD decisions. To that end, the algorithms are fused with a weighted sum of the individual PAD scores as follows:

$$s = (1 - \alpha) \cdot s_1 + \alpha \cdot s_2 \quad (7)$$

where s_1, s_2 with $1, 2 \in \{\text{ss}, \text{res}, \text{mob}, \text{VGG}\}$ represent the individual scores output by the approaches described above, α the fusion weight value selected using the validation dataset, and s the final fusion score.

VI. EXPERIMENTAL FRAMEWORK

A. Database

The database considered in the experimental evaluation was acquired within the BATL research project [61] in collaboration with our partners at the University of South California (USC). The project is financed by IARPA ODIN program [10]. Data were collected in two different stages and comprise both bona fide and PA samples. Note that the project sponsor IARPA has indicated that they will make the SWIR finger database available in the near future such that research results presented in this article can be reproduced.

For the bona fide samples, a total of 163 subjects participated during the first stage. For each of them, all 5 fingers of the right hand were captured. For the second stage, there were a total of 399 subjects. Index, middle and ring fingers of both hands were captured from each subject. It is important to highlight that people from different gender, ethnicity, and age were considered during the acquisition in order to evaluate the systems and algorithms in realistic conditions.

For the PA samples, the selection of the PAI fabrication materials was based on the requirements of IARPA ODIN program evaluation, covering the most challenging PAIs [14], [15]. There are a total of 35 different PAI species, which can be further categorised into eight main groups, namely: dragon skin, latex, overlay, playdoh, printed fingers, silicone, silly putty and wax. All details are included in Table III. It should be noted that each material combination specified in Table III is unique, and thus constitutes a different PAI species.

Finally, all captured samples were manually reviewed in order to remove all samples with operational errors (e.g., finger movement) or hardware failures, ending up with a total of 4,290 and 443 bona fide and PA samples, respectively.

TABLE III

PAI SPECIES INCLUDED IN THE EXPERIMENTAL WORK OF THIS STUDY. PAI SPECIES USED ONLY FOR TESTING AND NOT FOR TRAINING (I.E., UNKNOWN ATTACKS) HAVE BEEN UNDERLINED

PAI Group	PAI Species
Dragon Skin	Finger, conductive, conductive nanotips white, <u>graphite</u>
Latex	Finger
Overlay	Conductive silicone, monster latex, glue, silicone, <u>urethane</u> , wax, dragon skin
Playdoh	Black, blue, green, orange, pink, purple, red, teal, <u>yellow</u>
Printed	2D photograph/matte paper, 3D normal/Ag paint,
Silicone	Barepaint coating, finger flesh/yellow, graphite, normal, <u>coating</u>
Silly Putty	Glow in the dark, normal, <u>metallic</u>
Wax	Finger

B. Experimental Protocol

The main goal behind the experimental protocol design is to analyse and prove the soundness of our proposed fingerprint PAD approach in a realistic scenario. Therefore, the database described in the previous section is split into non-overlapping training, validation, and test datasets following the same procedure considered in previous studies [53], [55]. All details are shown in Table IV. In order to provide a fair comparison among the approaches described in Sect. V, the same partitions will be used in all experiments.

For the development of our proposed fingerprint PAD methods, both training and validation datasets are used in order to train the weights of the systems and select the optimal network architectures. For the training dataset, we consider a total of 130 samples for each class (i.e., bona fide and PA), whereas for the validation dataset the number of samples is reduced to 90 per class. It is important to highlight that we consider the same number of samples per class during the development of the systems in order to avoid bias towards one class.

For the final evaluation, the test dataset comprises the remaining bona fide (4070) and PA (223) samples not used during the development of the systems, thereby allowing a fair performance analysis.

Moreover, it is important to remark that the test dataset includes 5 unknown PAI species, which were not considered during the development stage (i.e., they are not present either in the train or in the validation datasets). This way, we can also evaluate the robustness of our proposed methods to unknown attacks, thereby modelling realistic scenarios. These unknown attacks are underlined in Table III.

Based on these partitions, three different sets of experiments are carried out:

A. Exp. 1 - Handcrafted features: first, the performance of the handcrafted features described in Sect. V-A is evaluated.

B. Exp. 2 - Deep learning features: then, we evaluate the performance of each deep learning approach described in Sect. V-B (i.e., end-to-end and feature-extraction + SVM classification, CNNs trained from scratch and transfer learning), and establish a fair benchmark by following the same experimental protocol.

C. Exp. 3 - Fused system: in the last set of experiments, the score level fusion (see Sect. V-C) of the aforementioned systems will be evaluated in order to determine the best performing configuration and assess the complementarity of the individual algorithms.

TABLE IV

PARTITION OF TRAINING, VALIDATION AND TEST DATASETS

	# Samples	# PA Samples	# BF Samples
Training set	260	130	130
Validation set	180	90	90
Test set	4293	223	4070

VII. EXPERIMENTAL RESULTS

A. Exp. 1 - Handcrafted Features

Fig. 5a shows the DET curves of each of the individual methods proposed in this study. As it may be observed, the spectral signature pixel-wise approach has achieved a 12.61% D-EER. Compared with the results first reported in [51] (APCER = 5.6% and BPCER = 0%), there is a clear decrease in the detection performance. This is due to the preliminary character of the first study, over a small database comprising only 60 samples and 12 different PAI species. In this work, the more thorough evaluation unveils the main drawbacks of the approach: it is not possible to get an APCER $\leq 2\%$, and for APCER $\approx 5\%$, the BPCER is over 20% (i.e., the system is not convenient any more).

B. Exp. 2 - Deep Learning Features

Deep learning strategies have considerably improved the results achieved using handcrafted features (see Fig. 5a for a comparison). In general, the features extracted by the neural network models provide a higher discriminative power and generalisation to new samples (note that during the development of the systems, all strategies were able to achieve loss values very close to zero for both training and validation datasets). This is due to the fact that, in contrast to the handcrafted features, not only the pixel-wise differences across wavelengths are taken into account, but also: *i)* global information from the complete images, and *ii)* further non-linear transformations of these differences (i.e., RGB images constructed following Eq. 6) through the convolutional layers of the networks (see Fig. 4).

For the case of training end-to-end residual CNN models from scratch, the best result obtained is a 2.25% D-EER. This result outperforms the handcrafted feature approach by an 82% relative improvement. Furthermore, low APCERs below 1% can be achieved for BPCERs below 8%, thereby overcoming the main drawback of the handcrafted feature approach. Similarly, for high convenient systems with BPCERs under 1%, the APCER ranges between 4% and 15%. These facts highlight the potential of incorporating residual connections to plain CNNs, being able to easily train neural network models without the necessity of having thousands of labelled images for each class, but only 130 (see Table IV).

Very good results have been also obtained for the use of pre-trained CNN models. In particular, the proposed MobileNet- and VGG19-based models have obtained state-of-the-art results with final values of 1.80% and 1.35% D-EER, respectively. These error rates have further improved

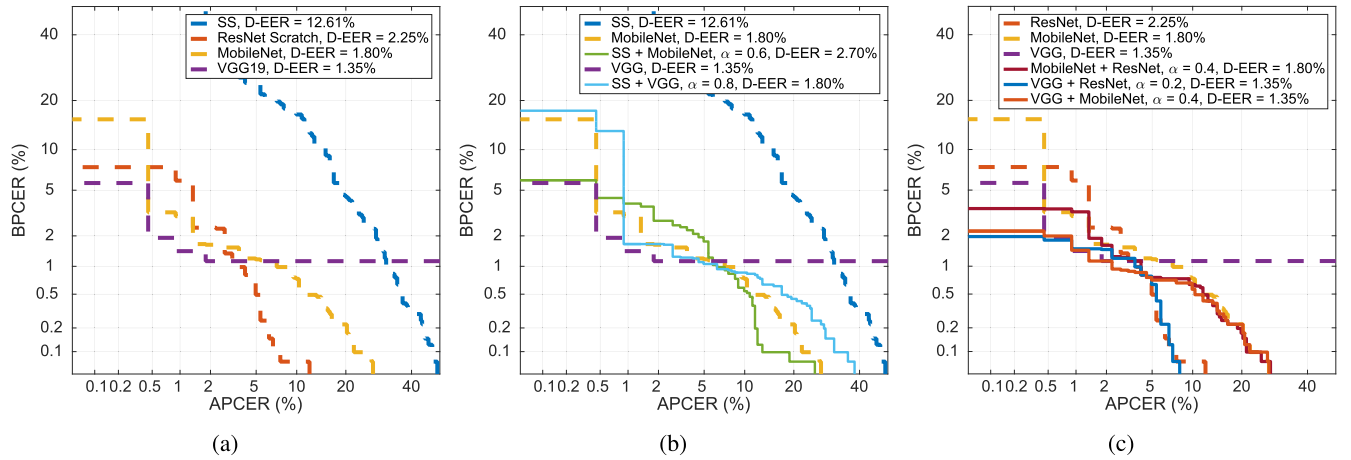


Fig. 5. Performance evaluation of: (a) All individual systems, (b) fusion of handcrafted features and end-to-end deep learning approaches (MobileNet and VGG19), and (c) fusion of end-to-end deep learning approaches (ResNet, MobileNet, and VGG19).

the ones obtained using handcrafted features, achieving an average relative improvement of 86% and 89%, respectively.

In addition, it is important to note that, even though an improvement at the D-EER operating point is achieved using these pre-trained models in combination with transfer learning techniques, this does not hold for all operating points. If we take a closer look at Fig. 5a, we can observe that for low BPCERs (i.e., high convenience), the best performing approach is the residual CNN trained from scratch. On the contrary, the lowest BPCER for an $APCER \leq 2\%$ (i.e., high security) is achieved by the VGG19 pre-trained model. However, it should be noted that the VGG19 based system cannot reach BPCERs under 1%, which can be done using the pre-trained MobileNet model. Therefore, depending on the final application, some CNN approaches might be more suitable than others.

For completeness, we also analyse the potential of using CNNs as feature extractors in combination with SVM classifiers. This way, we can also analyse the improvement achieved using deep learning features compared to the handcrafted features, which were also classified using SVMs. The performance in terms of APCER and BPCER is summarised in Table V (note that a single binary decision in terms of the thresholded distance is considered for the SVMs trained on the CNN features). As it may be observed, the operating points are always contained within the DET curves reported in Fig. 5a, which means that no further improvement has been achieved using the SVM for classification with respect to the last fully-connected sigmoid activation layer of the end-to-end CNNs. Therefore, these results further confirm the advantages of the learned features with respect to the handcrafted approach. Thus, in the remaining experiments only the end-to-end CNNs will be considered.

All these results show the potential of using CNNs in combination with SWIR images for fingerprint PAD purposes, and the robustness of the features extracted. Fig. 6 shows some examples of the features extracted in the first convolutional layer (64 filters) of the VGG19-based model for bona fide and PA samples. In general, very different features are extracted for bona fide and PA samples. This fact can be easily observed

TABLE V
PERFORMANCE EVALUATION OF THE DEEP LEARNING FEATURE EXTRACTORS IN COMBINATION WITH THE SVM CLASSIFIERS

	BPCER (%)	APCER (%)
Residual CNN	3.37	1.35
MobileNet-Based Model	5.33	0.45
VGG19-Based Model	1.89	0.90

when considering overlays based on monster latex and glue, Fig. 6 (d) and (e), respectively. However, for other materials such as the yellow playdoh, the features extracted by the network are more similar to the bona fide samples (Fig. 6 (a) vs. Fig. 6 (c)), indicating the difficulty of the task.

C. Exp. 2 - Deep Learning: Robustness to Unknown Attacks

Finally, we have also studied the robustness and generalisation capacity of the deep learning methods to new PAIs (a.k.a. unknown attacks). In order to do that, 30 samples acquired from 5 out of the 35 total PAIs available in the database (see Table III) were considered only for testing the systems (i.e., none of those PAI species were included in the training and validation datasets). The reason behind this particular PAI selection is twofold. On the one hand, we choose one PAI species from each PAI category of Table III, to increase the variability also in the unknown attacks. On the other hand, we select the PAI species with the smallest number of samples available, in order to maximise the number of training samples and hence the detection performance.

In general, very good results have been achieved for all methods. At the D-EER operating point, only one sample from a yellow playdoh finger has been misclassified by the residual CNN and MobileNet-based models, whereas for the case of using the VGG19-based model, all 30 samples stemming from the unknown attacks have been correctly classified. This proves the robustness of the proposed PAD approach against unknown attacks that may appear in the future.

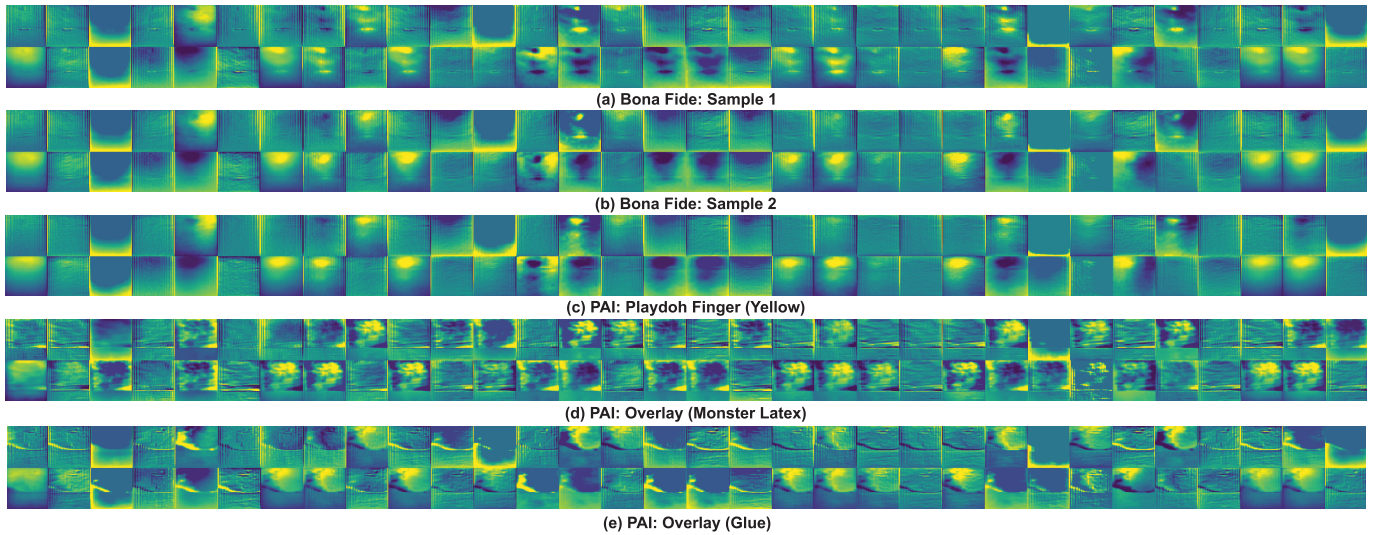


Fig. 6. Examples of the features extracted in the first convolutional layer (64 filters) of the VGG19-based model from the samples depicted in Fig. 3.

D. Exp. 3 - Fused Systems

In order to further enhance the results achieved by individual methods, and analyse to which degree the systems complement each other, we study in this last set of experiments the fusion of multiple systems at score level. In all cases, the performance is optimised using the validation dataset in terms of the D-EER for values of $\alpha \in [0, 1]$ (see Eq. 7), where this α weight corresponds to the second system referred to in the legend.

First, the fusion of handcrafted and deep learning features is evaluated in Fig. 5b. Only the MobileNet and VGG19 on are depicted, since no improvement was achieved for the fusion of the residual network and the spectral signatures with respect to the individual CNN. As it could be expected given the high performance gap between the spectral signatures and the deep learning counterparts, the score level fusion yields a minimum improvement with respect to the CNNs only in two cases: *i*) for either low BPCER $\leq 0.5\%$ or low APCER $\leq 0.5\%$ for the MobileNet approach (dashed yellow vs solid green curves), and for *ii*) BPCER $\leq 1\%$ for the VGG19 network (dashed purple vs solid light blue curves).

Afterwards, the three CNN approaches have been fused in a two-by-two basis (the fusion of all three systems showed no further improvement), and the best performing fusions are depicted in Fig. 5c. As it may be observed, no further improvements have been achieved for the operating point around the D-EER. However, for APCER $\leq 0.5\%$, the corresponding BPCER values for the fused systems (solid lines) are significantly lower than those of the individual networks (dashed lines): close to 2% for the fusions with VGG19 instead of between 5% and 15% (i.e., close to a 90% relative improvement). That yields convenient systems (i.e., low BPCER) even for highly secure (i.e., very low APCER) scenarios. On the other hand, for low BPCER $\leq 1\%$, the best APCER ($\leq 10\%$) is achieved for either the residual CNN alone (dashed orange) or its fusion with the VGG19 (solid dark blue). Taking a closer look at the individual PAD scores, we can see that both networks complement each other. Lastly, if we compare

Figs. 5b and 5c, we observe a superior performance in the latter case, thereby further supporting the fact that CNNs can perform better than the baseline handcrafted fusion in this task.

All in all, we can conclude that a remarkable performance can be achieved for fingerprint PAD using SWIR images and the fusion of two CNN models: a residual CNN trained from scratch and a pre-trained VGG19 CNN. We now compare our proposed approach with *i*) non-conventional sensors (Table I), and *ii*) conventional sensors in combination with deep learning approaches (Table II). Note that most algorithms and experimental conditions vary between the listed works, e.g., the database and PAI species considered for training and testing. Therefore, Tables I and II should be mainly interpreted in general terms to compare how different scenarios of use based on conventional and non-conventional sensors, as well as machine learning algorithms, are able to detect PAs. Our proposed fingerprint PAD system has achieved a final 1.35% D-EER. Furthermore, other operating points yield a BPCER of 2% for APCER $\leq 0.5\%$, and an APCER $\approx 7\%$ for BPCER = 0.1%. These results have outperformed the most similar and recent studies based on conventional sensors with deep learning approaches (Table II) even when increasing the variety of PAIs (35). In [31], the authors achieved a final 2.00% ACER over an own-acquired database composed of 12 PAIs. The selection of the PAIs was also based on the requirements of IARPA ODIN program, allowing therefore a fair comparison with the results achieved in the present study. In [68], the authors achieved a final 1.43% ACER on LivDet 2011-15 composed of 11 PAIs. Our proposed fingerprint PAD system has also outperformed in large margins our preliminary handcrafted feature approaches evaluated on similar conditions [59], [60]. Finally, it is important to analyse the challenging unknown attack scenario. Different approaches have been proposed in the last years [35], [36], all of them vulnerable against unknown attacks. Our proposed fused system has been able to correctly detect all unknown attacks, proving its high generalisation capacity to new PAIs that can appear in the future.

VIII. CONCLUSION

In this article we have presented a fingerprint PAD scheme based on *i*) a new capture device for the acquisition of finger samples in the SWIR spectrum, and *ii*) state-of-the-art deep learning techniques. An in depth analysis of several networks, either trained from scratch or using transfer learning over pre-trained models, and either as end-to-end solutions or as feature extractors in combination with SVMs for classification, has revealed the soundness of the proposed approach.

Three different CNN architectures have been tested: a residual CNN trained from scratch [53], [80], and the adaptation of the final layers of the VGG19 [55] and the MobileNet [54] pre-trained models. In addition, the performance of the proposed DL approaches has been benchmarked against the only handcrafted approach for fingerprint PAD based on SWIR images available in the literature [51]. The performance of all the individual algorithms has been tested over a database comprising more than 4700 samples, stemming from 562 different subjects and 35 different PAI species. Furthermore, several score level fusion schemes have been evaluated. The experimental protocol was designed to simulate a real life scenario: only 260 samples were used for training, and 30 samples acquired from 5 PAI species were excluded from the development stage and considered only for testing (i.e., unknown attack scenario).

In the aforementioned conditions, the best performance was reached for the fusion of two end-to-end CNNs: the residual CNN trained from scratch and the adapted VGG19 pre-trained model. A D-EER of 1.35% was obtained. Moreover, this system can be used for different applications. First, if high user convenience is preferred, an APCER around 7% can be achieved for a BPCER of 0.1% (i.e., only 1 in 1000 bona fide samples will be rejected). Also, for highly secure scenarios, a BPCER of 2% can be achieved for any APCER under 0.5%. These results clearly outperform those achieved with the handcrafted features, which yielded a D-EER over 12% and had trouble reaching APCERs under 2%.

We may thus conclude, that the use of SWIR images in combination with state-of-the-art CNNs offers a reliable and efficient solution to the threat posed by presentation attacks. However, the development of new countermeasures usually brings the corresponding development of new attacks, in this case, new PAI species. To tackle them, we plan to fuse the techniques developed in this work, which analyse the surface of the finger within the SWIR spectrum, with other approaches analysing bona fide properties below the skin [59], [60].

ACKNOWLEDGMENT

The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies, either expressed or implied, of ODNI, IARPA, or the U.S. Government. The U.S. Government is authorized to reproduce and distribute reprints for governmental purposes notwithstanding any copyright annotation therein. This work was carried out during an internship of R. Tolosana at da/sec.

REFERENCES

- [1] Government of India. (2012). *Unique Identification Authority of India*. [Online]. Available: <https://uidai.gov.in/>
- [2] European Commission. (2013). *Smart Borders*. [Online]. Available: http://ec.europa.eu/dgs/home-affairs/what-we-do/policies/borders-and-visas/smart-borders/index_en.htm
- [3] A. Zwiesele, A. Munde, C. Busch, and H. Daum, "BioIS study: Comparative study of biometric identification systems," in *Proc. IEEE 34th Annu. Int. Carnahan Conf. Secur. Technol.*, Oct. 2000, pp. 60–63.
- [4] N. K. Ratha, J. H. Connell, and R. M. Bolle, "Enhancing security and privacy in biometrics-based authentication systems," *IBM Syst. J.*, vol. 40, no. 3, pp. 614–634, 2001.
- [5] *Information Technology—Biometric Presentation Attack Detection*, document ISO/IEC 30107-1 ISO/IEC JTC1 SC37 Biometrics, International Organization for Standardization, 2016.
- [6] S. Marcel, M. S. Nixon, and S. Z. Li, Eds., *Handbook of Biometric Anti-Spoofing: Trusted Biometrics Under Spoofing Attacks*. Berlin, Germany: Springer, 2014.
- [7] A. Hadid, N. Evans, S. Marcel, and J. Fierrez, "Biometrics systems under spoofing attack: An evaluation methodology and lessons learned," *IEEE Signal Process. Mag.*, vol. 32, no. 5, pp. 20–30, Sep. 2015.
- [8] Tabularasa. (2010). *Trusted Biometrics Under Spoofing Attacks*. [Online]. Available: <http://www.tabularasa-euproject.org/>
- [9] BEAT. (2012). *Biometrics Evaluation and Testing*. [Online]. Available: <http://www.beat-eu.org/>
- [10] ODNI and IARPA. (2016). *IARPA-BAA-16-04 (Thor)*. [Online]. Available: <https://www.iarpa.gov/index.php/research-programs/odin/odin-baa>
- [11] L. Ghiani, D. A. Yambay, V. Mura, G. L. Marcialis, F. Roli, and S. A. Schuckers, "Review of the fingerprint liveness detection (LivDet) competition series: 2009 to 2015," *Image Vis. Comput.*, vol. 58, pp. 110–128, Feb. 2016.
- [12] V. Mura *et al.*, "LivDet 2017 fingerprint liveness detection competition," in *Proc. Int. Conf. Biometrics (ICB)*, Feb. 2018, pp. 297–302.
- [13] J. Galbally and M. Gomez-Barrero, "Presentation attack detection in iris recognition," in *Iris and Periocular Biometrics*, C. Busch and C. Rathgeb, Eds. Edison, NJ, USA: IET, Aug. 2017, pp. 235–263.
- [14] E. Marasco and A. Ross, "A survey on antispooing schemes for fingerprint recognition systems," *ACM Comput. Surv.*, vol. 47, no. 2, p. 28, 2015.
- [15] C. Sousedik and C. Busch, "Presentation attack detection methods for fingerprint recognition systems: A survey," *IET Biometrics*, vol. 3, no. 4, pp. 219–233, Dec. 2014.
- [16] J. Galbally, S. Marcel, and J. Fierrez, "Biometric antispooing methods: A survey in face recognition," *IEEE Access*, vol. 2, pp. 1530–1552, 2014.
- [17] R. Tolosana, R. Vera-Rodriguez, J. Fierrez, and J. Ortega-Garcia, "Handbook biometric anti-spoofing," in *Presentation Attacks in Signature Biometrics: Types and Introduction to Attack Detection*, 2nd ed. Berlin, Germany: Springer, 2018.
- [18] R. Raghavendra, M. Avinash, S. Marcel, and C. Busch, "Finger vein liveness detection using motion magnification," in *Proc. IEEE 7th Int. Conf. Biometrics Theory, Appl. Syst. (BTAS)*, Sep. 2015, pp. 1–7.
- [19] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. Cambridge, MA, USA: MIT Press, 2016.
- [20] I. Sutskever, O. Vinyals, and Q. V. Le, "Sequence to sequence learning with neural networks," in *Proc. Adv. Neural Inf. Process. Syst.*, 2014, pp. 3104–3112.
- [21] B. Zhou, A. Khosla, A. Lapedriza, A. Oliva, and A. Torralba, "Learning deep features for discriminative localization," in *Proc. Conf. Comput. Vis. Pattern Recognit.*, Jun. 2016, pp. 2921–2929.
- [22] S. J. Pan and Q. Yang, "A survey on transfer learning," *IEEE Trans. Knowl. Data Eng.*, vol. 22, no. 10, pp. 1345–1359, Oct. 2010.
- [23] J. Hu, J. Lu, and Y.-P. Tan, "Deep transfer metric learning," in *Proc. Conf. Comput. Vis. Pattern Recognit.*, Jun. 2015, pp. 325–333.
- [24] A. Rattani and R. Derakhshani, "On fine-tuning convolutional neural networks for smartphone based ocular recognition," in *Proc. Int. Joint Conf. Biometrics*, Oct. 2017, pp. 762–767.
- [25] R. Tolosana, R. Vera-Rodriguez, J. Fierrez, and J. Ortega-Garcia, "Exploring recurrent neural networks for on-line handwritten signature biometrics," *IEEE Access*, vol. 6, pp. 5128–5138, 2018.
- [26] R. F. Nogueira, R. de Alencar Lotufo, and R. C. Machado, "Fingerprint liveness detection using convolutional neural networks," *IEEE Trans. Inf. Forensics Security*, vol. 11, no. 6, pp. 1206–1213, Jun. 2016.
- [27] H.-U. Jang, H.-Y. Choi, D. Kim, J. Son, and H.-K. Lee, "Fingerprint spoof detection using contrast enhancement and convolutional neural networks," in *Proc. Int. Conf. Inf. Sci. Appl.*, 2017, pp. 331–338.

- [28] S. Kim, B. Park, B. S. Song, and S. Yang, "Deep belief network based statistical feature learning for fingerprint liveness detection," *Pattern Recognit. Lett.*, vol. 77, pp. 58–65, Jul. 2016.
- [29] A. Toosi, S. Cumani, and A. Bottino, "CNN patch-based voting for fingerprint liveness detection," in *Proc. Int. Joint Conf. Comput. Intell.*, 2017, pp. 158–165.
- [30] G. B. Souza, D. Santos, R. G. Pires, A. N. Marana, and J. P. Papa, "Deep boltzmann machines for robust fingerprint spoofing attack detection," in *Proc. Int. Joint Conf. Neural Netw. (IJCNN)*, May 2017, pp. 1863–1870.
- [31] T. Chugh, K. Cao, and K. Anil Jain, "Fingerprint spoof buster: Use of minutiae-centered patches," *IEEE Trans. Inf. Forensics Security*, vol. 13, no. 9, pp. 2190–2202, Sep. 2018.
- [32] O. Kanich, M. Drahanský, and M. Mézl, "Use of creative materials for fingerprint spoofs," in *Proc. Int. Workshop Biometrics Forensics*, Jun. 2018, pp. 1–8.
- [33] E. Marasco and C. Sansone, "On the robustness of fingerprint liveness detection algorithms against new materials used for spoofing," in *Proc. Int. Conf. Bio-Inspired Syst. Signal*, vol. 8, 2011, pp. 553–558.
- [34] A. Rattani and A. Ross, "Automatic adaptation of fingerprint liveness detector to new spoof materials," in *Proc. IEEE Int. Joint Conf. Biometrics*, Sep. 2014, pp. 1–8.
- [35] A. Rattani, W. J. Scheirer, and A. Ross, "Open set fingerprint spoof detection across novel fabrication materials," *IEEE Trans. Inf. Forensics Security*, vol. 10, no. 11, pp. 2447–2460, Nov. 2015.
- [36] A. F. Sequeira and J. S. Cardoso, "Fingerprint liveness detection in the presence of capable intruders," *Sensors*, vol. 15, no. 6, pp. 14615–14638, Jun. 2015.
- [37] S. R. Arashloo and J. Kittler, "An anomaly detection approach to face spoofing detection: A new formulation and evaluation protocol," in *Proc. IEEE Int. Joint Conf. Biometrics*, Oct. 2017, pp. 80–89.
- [38] A. F. Sequeira, S. Thavalengal, J. Ferryman, P. Corcoran, and J. S. Cardoso, "A realistic evaluation of iris presentation attack detection," in *Proc. 39th Int. Conf. Telecommun. Signal Process. (TSP)*, Jun. 2016, pp. 660–664.
- [39] Y. Wang, X. Hao, Y. Hou, and C. Guo, "A new multispectral method for face liveness detection," in *Proc. 2nd IAPR Asian Conf. Pattern Recognit.*, Nov. 2013, pp. 922–926.
- [40] H. Steiner, S. Sporrer, A. Kolb, and N. Jung, "Design of an active multispectral SWIR camera system for skin detection and face verification," *J. Sensors*, vol. 2016, Mar. 2015, Art. no. 9682453.
- [41] R. K. Rowe, K. A. Nixon, and P. W. Butler, *Multispectral Fingerprint Image Acquisition*. London, U.K.: Springer, 2008, pp. 3–23.
- [42] S. Chang, K. Larin, Y. Mao, W. Almuhtadi, and C. Fluerau, "Fingerprint spoof detection by NIR optical analysis," in *State Art Biometrics*. Rijeka, Croatia: InTech, 2011, pp. 57–84.
- [43] A. Lumini and L. Nanni, "Fair comparison of skin detection approaches on publicly available datasets," 2018, *arXiv:1802.02531*. [Online]. Available: <https://arxiv.org/abs/1802.02531>
- [44] J. A. Jacquez, J. Huss, W. McKeehan, J. M. Dimitroff, and H. F. Kuppenheim, "Spectral reflectance of human skin in the region 0.7–2.6 μ ," *J. Appl. Physiol.*, vol. 8, no. 3, pp. 297–299, Nov. 1955.
- [45] R. S. Ghiass, O. Arandjelović, A. Bendada, and X. Maldague, "Infrared face recognition: A comprehensive review of methodologies and databases," *Pattern Recognit.*, vol. 47, no. 9, pp. 2807–2824, Sep. 2014.
- [46] T. Bourlai, *Face Recognition Across the Imaging Spectrum*. Berlin, Germany: Springer, 2016.
- [47] T. Bourlai, N. Kalka, A. Ross, B. Cukic, and L. Hornak, "Cross-spectral face verification in the short wave infrared (SWIR) band," in *Proc. Int. Conf. Pattern Recognit.*, Aug. 2010, pp. 1343–1347.
- [48] F. Nicolo and N. A. Schmid, "Long range cross-spectral face recognition: Matching SWIR against visible light images," *IEEE Trans. Inf. Forensics Security*, vol. 7, no. 6, pp. 1717–1726, Dec. 2012.
- [49] N. Narang and T. Bourlai, "Face recognition in the SWIR band when using single sensor multi-wavelength imaging systems," *Image Vis. Comput.*, vol. 33, pp. 26–43, Jan. 2015.
- [50] M. A. Ferrer, A. Morales, and A. Díaz, "An approach to SWIR hyperspectral hand biometrics," *Inf. Sci.*, vol. 268, pp. 3–19, Jun. 2014.
- [51] M. Gomez-Barrero, J. Kolberg, and C. Busch, "Towards fingerprint presentation attack detection based on short wave infrared imaging and spectral signatures," in *Proc. Norwegian Inf. Secur. Conf.*, Sep. 2018, pp. 1–5.
- [52] R. Tolosana, M. Gomez-Barrero, J. Kolberg, A. Morales, C. Busch, and J. Ortega, "Towards fingerprint presentation attack detection based on convolutional neural networks and short wave infrared imaging," in *Proc. Int. Conf. Biometrics Special Interest Group (BIOSIG)*, Sep. 2018, pp. 1–5.
- [53] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, 2016, pp. 770–778.
- [54] A. Howard *et al.*, "MobileNets: Efficient convolutional neural networks for mobile vision applications," pp. 1–9, 2017, *arXiv:1704.04861*. [Online]. Available: https://arxiv.org/abs/1704.04861?source=post_page
- [55] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," in *Proc. Int. Conf. Learn. Representations*, 2015, pp. 1–14.
- [56] *Information Technology—Biometric presentation attack detection—Part 3: Testing and Reporting*, Documents ISO/IEC JTC1 SC37 Biometrics, ISO/IEC FIDS 30107-3, International Organization for Standardization, 2017.
- [57] R. K. Rowe, K. A. Nixon, and P. W. Butler, *Multispectral Fingerprint Image Acquisition*. London, U.K.: Springer, 2008, pp. 3–23.
- [58] C. Hengfoss, A. Kulcke, G. Mull, C. Edler, K. Püschel, and E. Jopp, "Dynamic liveness and forgeries detection of the finger surface on the basis of spectroscopy in the 400–1650 nm region," *Forensic Sci. Int.*, vol. 212, nos. 1–3, pp. 61–68, Oct. 2011.
- [59] P. Keilbach, J. Kolberg, M. Gomez-Barrero, C. Busch, and H. Langweg, "Fingerprint presentation attack detection using laser speckle contrast imaging," in *Proc. Int. Conf. Biometrics Special Interest Group (BIOSIG)*, Sep. 2018, pp. 1–6.
- [60] J. Kolberg, M. Gomez-Barrero, S. Venkatesh, R. Raghavendra, and C. Busch, "Presentation attack detection for finger recognition," in *Handbook of Vascular Biometrics*, S. Marcel, A. Uhl, R. Veldhuis, and C. Busch, Eds. Springer, 2019.
- [61] *Biometric Authentication With A Timeless Learner*, BATL, New York, NY, USA, 2017.
- [62] D. Menotti *et al.*, "Deep representations for iris, face, and fingerprint spoofing detection," *IEEE Trans. Inf. Forensics Security*, vol. 10, no. 4, pp. 864–879, Apr. 2015.
- [63] E. Marasco, P. Wild, and B. Cukic, "Robust and interoperable fingerprint spoof detection via convolutional neural networks," in *Proc. Int. Conf. Technol. Homeland Secur.*, May 2016, pp. 1–6.
- [64] C. Yuan, X. Li, Q. Wu, J. Li, and X. Sun, "Fingerprint liveness detection from different fingerprint materials using convolutional neural network and principal component analysis," *Comput., Mater. Continua*, vol. 53, no. 4, pp. 357–372, 2017.
- [65] C. Wang, K. Li, Z. Wu, and Q. Zhao, "A DCNN based fingerprint liveness detection algorithm with voting strategy," in *Proc. Chin. Conf. Biometric Recognit.*, 2015, pp. 241–249.
- [66] E. Park, W. Kim, Q. Li, J. Kim, and H. Kim, "Fingerprint liveness detection using CNN features of random sample patches," in *Proc. Int. Conf. Biometrics Special Interest Group (BIOSIG)*, Sep. 2016, pp. 1–4.
- [67] F. Pala and B. Bhanu, "On the accuracy and robustness of deep triplet embedding for fingerprint liveness detection," in *Proc. Int. Conf. Image Process. (ICIP)*, Sep. 2017, pp. 116–120.
- [68] E. Park, X. Cui, W. Kim, J. Liu, and H. Kim, "Patch-based fake fingerprint detection using a fully convolutional neural network with a small number of parameters and an optimal threshold," Mar. 2018, *arXiv:1803.07817*. [Online]. Available: <https://arxiv.org/abs/1803.07817>
- [69] A. Toosi, A. Bottino, S. Cumani, P. Negri, and P. L. Sottile, "Feature fusion for fingerprint liveness detection: A comparative study," *IEEE Access*, vol. 5, p. 23695–23709, 2017.
- [70] A. Krizhevsky, I. Sutskever, and G. E. Geoffrey, "ImageNet classification with deep convolutional neural networks," in *Proc. Adv. Neural Inf. Process. Syst.*, 2012, pp. 1097–1105.
- [71] C. Szegedy *et al.*, "Going deeper with convolutions," in *Proc. Conf. Comput. Vis. Pattern Recognit.*, Jun. 2015, pp. 1–9.
- [72] L. Wan, M. Zeiler, S. Zhang, Y. L. Cun, and R. Fergus, "Regularization of neural networks using dropconnect," in *Proc. Int. Conf. Mach. Learn.*, Feb. 2013, pp. 1058–1066.
- [73] J. Galbally, J. Fierrez, F. Alonso-Fernandez, and M. Martinez-Diaz, "Evaluation of direct attacks to fingerprint verification systems," *Telecommun. Syst.*, vol. 47, nos. 3–4, pp. 243–254, 2011.
- [74] F. N. Iandola, S. Han, M. W. Moskewicz, and K. Ashraf, "SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and <0.5MB model size," 2016, *arXiv:1602.07360*. [Online]. Available: <https://arxiv.org/abs/1602.07360>
- [75] *Information Technology—Biometric Presentation Attack Detection—Part 2: Data Formats*, Documents ISO/IEC JTC1 SC37 Biometrics, ISO/IEC DIS 30107-2, International Organization for Standardization, 2017.
- [76] O. Russakovsky *et al.*, "Imagenet large scale visual recognition challenge," *Int. J. Comput. Vis.*, vol. 115, no. 3, pp. 211–252, 2015.

- [77] A. Krizhevsky and G. Hinton, "Learning multiple layers of features from tiny images," Univ. Toronto, Canada, Tech. Rep., 2009, vol. 1, no. 4.
- [78] M. Everingham, L. Van Gool, C. K. I. Williams, J. Winn, and A. Zisserman, "The Pascal visual object classes (VOC) challenge," *Int. J. Comput. Vis.*, vol. 88, no. 2, pp. 303–338, Sep. 2009.
- [79] T.-Y. Lin *et al.*, "Microsoft COCO: Common objects in context," in *Proc. Eur. Conf. Comput. Vis.*, 2014, pp. 740–755.
- [80] C. Szegedy, S. Ioffe, and V. Vanhoucke, "Inception-v4, inception-resnet and the impact of residual connections on learning," in *Proc. 31st AAAI Conf. Artif. Intell.*, 2017, pp. 1–7.
- [81] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," pp. 1–11, 2015, *arXiv:1502.03167*. [Online]. Available: <https://arxiv.org/abs/1502.03167>
- [82] T. Weyand, I. Kostrikov, and J. Philbin, "Planet-photo geolocation with convolutional neural networks," in *Proc. Eur. Conf. Comput. Vis.*, 2016, pp. 37–55.



Ruben Tolosana received the M.Sc. degree in telecommunication engineering, and the Ph.D. degree in computer and telecommunication engineering from the Universidad Autonoma de Madrid, in 2014 and 2019, respectively. In April 2014, he joined the Biometrics and Data Pattern Analytics (BiDA) Lab, Universidad Autonoma de Madrid, where he is currently collaborating as a Post-Doctoral Researcher. Since then, he has been granted with several awards, such as the FPU Research Fellowship from Spanish MEC (2015) and the

European Biometrics Industry Award (2018). His research interests are mainly focused on signal and image processing, pattern recognition, deep learning, and biometrics, particularly in the areas of handwriting and handwritten signature. He is the author of several publications and also collaborates as a reviewer in many different high-impact conferences, including the ICDAR, ICB, BTAS, and EUSIPCO, and journals such as the IEEE TRANSACTIONS ON INFORMATION FORENSICS AND SECURITY, the IEEE TRANSACTIONS ON CYBERNETICS, the IEEE TRANSACTIONS ON IMAGE PROCESSING, and the *ACM Computing Surveys*. Finally, he has participated in several National and European projects focused on the deployment of biometric security throughout the world.



Marta Gomez-Barrero received the M.Sc. degree in computer science and mathematics and the Ph.D. degree in electrical engineering from the Universidad Autonoma de Madrid in 2011 and 2016, respectively. Since 2016, she has been a Postdoctoral Researcher with the National Research Centre for Applied Cybersecurity (CRISP), Germany. Her current research focuses on the development of privacy-enhancing biometric technologies and presentation attack detection methods within the wider fields of pattern recognition and machine learning.

She has been actively involved in international projects dealing with vulnerability evaluation of biometric systems, including the EU FP7 projects Tabula Rasa and BEAT or the BATL Project within the U.S. IARPA Odin Program. She was a recipient of a number of distinctions, including the EAB European Biometric Industry Award 2015, the Best Ph.D. Dissertation Award by the Universidad Autonoma de Madrid (2015–2016), the Siew-Ngiem Best Paper Award at ICB 2015, the Archimedes Award for young researches from the Spanish Ministry of Education in 2013, and the Best Poster Award at ICB 2013.



Christoph Busch received the Diploma degree from the Technical University of Darmstadt (TUD), Darmstadt, Germany, and the Ph.D. degree in computer graphics from TUD, in 1997. He joined the Fraunhofer Institute for Computer Graphics, Darmstadt, in 1997. He is currently a member of the Norwegian Biometrics Laboratory, Norwegian University of Science and Technology, Norway. He holds a joint appointment with the Faculty of Computer Science, Hochschule Darmstadt. He lectures a course on biometric systems with DTU, Copenhagen, since 2007. He has coauthored over 400 technical papers. His research interests include pattern recognition, multimodal and mobile biometrics, and privacy enhancing technologies for biometric systems. He is also the Co-Founder of the European Association for Biometrics and convener of WG3 in ISO/IEC JTC1 SC37 on Biometrics. He has been a speaker at international conferences.



Javier Ortega-Garcia received the M.Sc. and Ph.D. (*cum laude*) degrees in electrical engineering from the Universidad Politécnica de Madrid, Spain, in 1989 and 1996, respectively. He is currently a Full Professor with the Signal Processing Chair, Universidad Autónoma de Madrid, Spain, where he holds courses on biometric recognition and digital signal processing. He is also the Founder and the Director of the Biometrics and Data Pattern Analytics (BiDA) Lab, Biometrics and Data Pattern Analytics Group. He has authored over 300 international contribu-

tions, including book chapters, refereed journal, and conference papers. His research interests are focused on biometric pattern recognition (online signature verification, speaker recognition, and human-device interaction) for security, e-health, and user profiling applications. He has chaired Odyssey-04, the Speaker Recognition Workshop, ICB-2013, the 6th IAPR International Conference on Biometrics, the ICCST 2017, and the 51st IEEE International Carnahan Conference on Security Technology.