The Rise of Data-driven Models in Presentation Attack Detection

Luis A. M. Pereira, Allan Pinto, Fernanda A. Andaló, Alexandre M. Ferreira, Bahram Lavi, Aurea Soriano-Vargas, Marcos V. M. Cirne, and Anderson Rocha

Abstract Biometric systems are prevalent in access control but are vulnerable to frauds. A typical attempt of violating them is through presentation attacks, in which synthetic data is directly presented to an acquisition sensor to deceive these systems. A well-designed biometric system should have a presentation attack detection (PAD) module. A fruitful way to perform PAD is to model properties of peculiar traits (artifacts) in synthetic data. Studies have been advocating for approaches that seek to model the artifacts automatically from data (data-driven), achieving state-of-the-art results in PAD. However, the following questions arise from this literature: Which approaches are state of the art? When do these approaches fail? How can such approaches complement the proposed ones based on human knowledge on PAD? How robust are these approaches under cross-dataset scenarios? Are these approaches robust against new attack types (e.g., face morphing)? Do these methods provide other ways to perform PAD, for example, using open-set classifiers rather than the classical binary formulation? Are these methods applicable to the multi-biometric setting? In this chapter, we address these questions through a literature review, focusing on three biometric modalities: face, fingerprint, and iris.

1 Introduction

In the contemporary society, oftentimes people and corporations manipulate information taking advantage of the increased adoption of digital systems — smartphones, bank and airport control systems, the Internet, and so on. To prevent such

Luis A. M. Pereira, Allan Pinto, Fernanda A. Andaló, Alexandre M. Ferreira, Bahram Lavi, Aurea Soriano-Vargas, Marcos V. M. Cirne, and Anderson Rocha,
Reasoning for Complex Data Lab. Institute of Computing, University of Campinas, Av. Albert Einstein, 1251, Campinas, SP, Brazil. CEP 13.083-852,
e-mail: {luis.pereira, allan.pinto, feandalo, melloferreira, bahram.lavi, aurea.soriano, marcos.cirne, anderson.rocha}@ic.unicamp.br

manipulations, much of the generated data, such as photos, conversational histories, transactions, and bank statements, should be protected from indiscriminate access, so people have control over their data and can maintain their right to privacy.

Traditionally, data protection methods rely on the use of external knowledge (e.g., passwords and secret questions) or tokens (e.g., smartcards), which may not be secure, as they can be forgotten, lost, stolen, or manipulated with ease. In addition, by using knowledge- or token-based solutions, the user is not required to claim an identity [18], which allows the use of multiple identities by a single person.

To overcome the disadvantages of traditional security methods, biometric systems use biological or behavioral traits pertaining to a user — face, iris, fingerprint, voice, gait, among others — to automatically recognize her/him, therefore granting access to private data. A biometric system gathers traits from an individual through a sensor, extracts features from such traits, and compares them with feature templates in a database [18], enabling the recognition of particular individuals.

However, biometric traits cannot be considered as completely private information, as we inevitably expose them in our everyday life [12]: our faces in social media, our fingerprints where we touch, our gait when we are recorded by surveillance cameras, among many other examples. This leads to the biggest drawback of pure biometric systems. In practice, our traces can be captured and offered to the system by an adversary, in order to circumvent security mechanisms.

The attack to a biometric system, which occurs whenever an adversary offers a counterfeit biometric trace to the acquisition sensor, is called a ***presentation attack*** (or *spoofing attack* in earliest literature). It is considered as the most dangerous type of attack to a biometric system [12], as the attacker primarily needs only access to the victim's traits, which are often plenty, and replay them to the biometric sensor.

In the earliest biometric methods and systems put into operation, there was little to no concern in providing countermeasures to presentation attacks, mainly because of the assumption that the counterfeiting of biometric traces, such as fingerprints and faces, was difficult to achieve. However, it was not long before we began to receive news of biometric systems hacked by the use of false traits. One of the most prominent examples took place in 2013,[1] when a Brazilian physician fooled the biometric employee attendance device, by using prosthetic fingers bearing the fingerprints of co-workers (Figure 1(b)). The investigation of this fraud scheme revealed that 300 public employees had been receiving pay without going to work.

To deal with the urging threat, academia and industry have been researching and applying automated methods to counterattack presentation attacks, a field of research known as presentation attack detection (PAD). These methods perform the task of differentiating between genuine (or *bona fide*) trait samples from attack ones, referred to as ***presentation attacks***. For some years, researched methods relied on human knowledge to automatically look for specific characteristics expected on genuine trait biometric samples, such as shape, texture, or liveness signs; or on attack samples, such as artifacts and noise. Recently, data-driven methods have been increasingly employed to learn relevant characteristics automatically from training

[1] BBC News: https://www.bbc.com/news/world-latin-america-21756709.

data, yielding state-of-the-art results for PAD without relying on knowledge-based algorithms to extract specific characteristics from samples.

The clear advances in the area after the dissemination of data-driven approaches have resulted in the possibility of generating models more robust to variations and nuances, which are captured during training, and capable of extracting and analyzing relevant sets of characteristics without the need of aggregating prior human knowledge. Nevertheless, as the models generated by data-driven methods are well fit to the training data, sometimes they do not generalize well when applied in a cross-dataset scenario, i.e., when testing data have characteristics not present/seen in the training data (such as different sensor noise, different geometric and photometric variations and distortions, etc.). This may indicate that methods based on prior knowledge conceivably complement data-driven ones. Although extremely important, these and other aspects are not often explored and discussed when new methods are published.

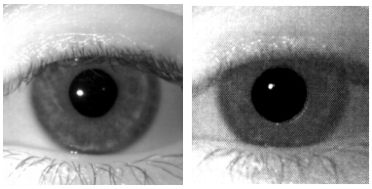
In this chapter, we cover the relevant literature on data-driven PAD methods, presenting a critical analysis of open, often overlooked, issues and challenges, in order to shed light on the problem. We answer and provide insights to important questions surrounding PAD research and applications. In which scenarios do these methods fail? How can such methods complement the ones based on human knowledge? How robust are these methods under the cross-dataset scenario? Are these methods robust against new attack types? Do these methods provide other ways to model the PAD problem besides the classical binary decision? And finally, are these methods applicable to multi-biometric settings?

To survey the relevant literature, we examine three widely used biometric modalities: face (Figure 1(a)), fingerprint (Figure 1(b)), and iris (Figure 1(c)). Faces are the most common biometric characteristic used by humans to recognize others [9], and they are often considered in biometric systems due to their non-intrusiveness and simplicity of acquisition (by any current camera). Systems based on face biometrics can be attacked by the presentation of a photograph, video, or 3D model of the user's face. Fingerprints are patterns of ridges and furrows located on the tip of each finger and can be captured by sensors that provide digital images of these patterns [9]. Fingerprint biometric systems are often spoofed by the presentation of fingerprints printed on paper or by 3D finger casts. Irises contain many distinctive features [9], such as ligaments, ridges, rings, and others, which favors their use as a biometric indicator. A fake iris sample can be created from an artificial eyeball, textured contact lens, and iris patterns printed on paper [43].

2 Benchmarks to Evaluate PAD Solutions

In this section, we describe the most adopted benchmarks in the three biometric authentication modalities explored in this paper. For each modality, there is a specific table to favor the comparison. These tables present some characteristics for each benchmark, such as dataset name, the total number of samples (videos or photos), the division between *bona fide* (BF) samples and PA (presentation attack) and, if

(a) (b)



(c)

Fig. 1 Presentation attack modalities. (a) Face presentation attack, CASIA-FAS dataset [57] examples: genuine face, warped paper attack, cut paper attack, and video attack. (b) Fingerprint presentation attack: six silicone fingers used to fool the biometric employee attendance device at a hospital in Brazil. Source: BBC [1]. (c) Iris presentation attack, LivDet-Iris Warsaw 2017 dataset [53]: genuine iris and printed attack.

the data is already split into training and testing, the train/test sets ratio. This last aspect is important to check whether the amount of data is enough to be used in a data-driven approach and to ensure that there are no significant unbalanced sets either between BF/PA or train/test. The number of subjects indicates diversity and average resolution is important to check the sample sizes (height $\times$ width) coming from different acquisition sensors. The setup column shows information regarding the capturing environment, and the following column describes which type of attack is considered. Acquisition and used devices for attacks give some information about the dataset representativeness, as increasing the number and adopting realist devices are pivotal for reducing the likelihood of specific artifacts that can be explored instead of specific characteristics. Some cells of these tables are marked as N/A, indicating "Not Applicable" and N/R indicating "Not Reported." A short discussion about the aspects of these benchmarks is also provided. To facilitate the link between the benchmarks and the described research papers, Table 4 (Appendix A) shows all references that have used each benchmark per modality.

Face PAD:

Face PAD benchmarks (Table 1) are usually composed of screen- and print-based forms of attacks. However, some of them have also tried to use 3D masks. In some of the benchmarks, the small number of subjects can be problematic as it shows less variability. For instance, 3DMAD and NUAA contain faces from less than 20 subjects. In turn, MSU USSA, although with more subjects, is composed of

celebrity faces acquired from the Internet, which may not retain detailed acquisition information. The main drawback of CASIA is having only one device for the attacks, lacking variety on acquisition artifacts. This is not the case for UVAD dataset, which contains videos recorded with six different devices and considered seven different devices when performing presentation attacks.

Iris PAD:

Iris PAD benchmarks (Table 2) are, in part, similar to Face PAD ones from the attack interface perspective (screen- and print-based). Some of them, however, have also explored specific aspects, such as contact lenses and artificial eyeballs (ND CLD 2015). IIITD Iris Spoofing dataset is unique in providing a combined attack setting comprising paper printout of eyes and contact lenses. Yet, the most common attacks are static, in which printed iris photos are adopted. The CASA Iris Fake provides a considerable number of subjects and a consistent ratio between live and PA samples, although training and testing sets are not specified. Not having a fixed split between training and testing data may harden comparisons of different solutions and may increase the likelihood of data leakage. An important aspect of IIITD WVU is its naturally cross-dataset setting, once its training set incorporates 3,000 samples from IIITD Iris Spoofing dataset.

Fingerprint PAD:

Fingerprint PAD benchmarks (Table 3) may consider different physical materials, such as ecoflex and gelatin. PBSKD is the benchmark with most spoofing materials (ten in total), creating one hundred spoof specimens that are acquired five times using two different fingerprint scanners. However, the most important data-source available nowadays is generated by the LivDet fingerprint competitions (years 2009, 2011, 2013, 2015, and 2017), which provide us with intra-sensor, cross-material, cross-sensor, and cross-dataset validation protocols. In the two last editions, testing samples were created using materials that are not seen in the training set, naturally enabling (and pushing for) open-set experimentation. The table does not show the numbers of each competition internal dataset, but it gives a holistic view of them concerning size, the ratio between live and PA samples, and the ratio between training and testing samples. The information about the number of subjects and resolution is omitted in the table as they differ among LivDet internal datasets given that their capturing scanners have different purposes (border control, mobile device authentication, among others).

Table 1: Face PAD benchmarks.

Dataset	Size	BF/PA	Train/Test	Subj.	Resolution	Setup	Types of attacks	Acquisition	Devices for attacks
Replay-Attack [7]	1,300 photos and videos	140/700	360/480	50	752×544 photos 720p videos	Lighting and holding	Screen and print	Canon PowerShot SX150	4x iPhone 3GS, 4x iPad first generation and 2x Triumph-Adler DCC 2520 color laser
CSMAD [4]	263 photos and videos	104/159	N/R	14	640 × 480	Four lighting conditions	6 silicone masks and 2 wearing ways	RealSense SR300 and Compact Pro (thermal)	Nikon Coolpix P520
CASIA-FASD [57]	600 videos	150/450	Yes	50	1226 × 813	Seven evaluation scenarios and three image quality	Three modes: warped photo, cut photo and video playback	Sony NEX-5, new and old USB camera	iPad
3DMAD [11]	76,500 frames	51k/25,500	N/R	17	640 × 480	Three sessions (2 weeks interval), having 5 videos of 300 frames from each session	3D mask	ThatsMyFace.com	Kinect
MSU USSA [32]	10,260 images	1,140/9,120	N/R	1,140	705 × 865	Uncontrolled	Screen and printed photo	Celebrities from internet and Nexus 5	Cannon PowerShot 550D, iPhone 5S and HP Color Laserjet CP6015
OULU-NPU [5]	4,950 videos	720/2,880	1,800/1,800	55	N/R	Lighting and background in three sections	Screen and print	6 mobile devices front camera	2 printers and 2 display
NUAA [47]	12,614	5,105/7,509	Yes	15	640 × 480	N/R	Printed photos	Webcam	N/R
UVAD [34]	17,076 videos	808/16,268	N/R	404	1366 × 768	Lighting, background and places in two sections	Screen	6 cameras	Seven displays

Table 2: Iris PAD benchmarks.

Dataset	**Size**	**Live/PA**	**Train/Test**	**Subj.**	**Resolution**	**PA types**	**Acquisition**	**PA devices**
Clarkson17 [53]	6,749	3,954/2,795	3,591/3,158	25	640×480	Printed images, patterned contact lenses, and printouts of patterned contact lenses	LG IrisAccess EOU2200	iPhone 5 and printer
IIITD Iris Spoofing [15]	4,848	0/4,848	N/R	101	640×480	Printed images	Cogent CIS 202, VistaFA2E and HP flatbed optical scanner	HP Color LaserJet 2025 printer
IIITD Delhi [22]	1,120	N/R	N/A	224	320×240	Textured and soft contact lens	JIRIS, JPC1000, digital CMOS camera	N/R
ND CLD 2015 [10]	7,300	4,800/2,500	6,000/1,200	278	640×480	Textured contact lens	IrisAccess LG 4000 and IrisGuard AD100	JJ, Ciba, UCL and ClearLab
Warsaw17 [53]	12,013	5,168/6,845	4,513/2,990	186	640×480	Printed images	IrisGuard AD100 and Sony EX-View CCD	Panasonic ET100
IIITD WVU [53]	7,459	2,952/4,507	3,250+3,000 /4,209	N/R	Diverse	Textured lens and printed images	CLI plus IIITD datasets and IrisShield sensor	Ciba and Aryan for the lens and HP P3015 for the printouts
CASIA Iris Fake [44]	10,240	6,000/4,120	N/A	500	640×480	Print, Contact, Plastic and Synth	LG-H100	Fuji Xerox C1110 printer, contact lens, replayed video and artificial eyeballs

Table 3: Fingerprint PAD datasets.

Dataset	**Size**	**Live/PA**	**Train/Test**	**Scanner**	**Model**	**PA devices**
LivDet 2009 [26]	11,000	5,500/5,500	2,750/8,250	Crossmatch, Identix and Biometrika	Verifier 300 LC, DFR2100 and FX2000	Gelatin, silicone, play-doh
LivDet 2011 [54]	16,000	8,000/8,000	8,000/8,000	Biometrika, Digital Persona, ItalData and Sagem	FX2000, 4000B, ET10 and MSO300	Gelatin, latex, ecoflex, Play-doh, silicone and wood glue
LivDet 2013 [13]	20,590	11,740/8,850	10,350/10,240	Biometrika, Crossmatch, ItalData and Swipe	FX2000, L Scan Guardian, ET10 and Swipe	Gelatin, body double, latex, play-doh, ecoflex, modasil, and wood glue
PBSKD [8]	1,800	900/900	1,000/1,000	CrossMatch and Lumidigm	Guardian 200 and Venus 302	Ecoflex, galatin, latex body paint, ecoflex with silver colloidal ink coating, ecoflex with BarePaint coating, ecoflex with nanotips coating, Crayola model magic, wood glue, monster liquid latex, 2D printed on office paper
Bogus [45]	16,000	8,000/8,000	8,000/8,000	Biometrika, Digital, Italdata and Sagem	FX2000, 4000B, ET10 and MSO300	Gelatin, Playdoh, silicone, wood glue, ecoflex and silgum

3 Data-Driven Methods for Presentation Attack Detection

In this section, we present the state-of-the-art methods for the three modalities considered in this chapter (face, fingerprint, and iris). We present methods for each modality, separately, and methods proposed for the multi-biometric scenario. We also discuss hybrid methods designed from data-driven and handcrafted methodologies. Recently, some methods proposed in the literature showed how to harmoniously mix these two approaches in order to take advantage from both.

3.1 Face PAD

Face recognition systems are one of the least intrusive biometric approaches and can be performed with low-cost sensors (e.g., smartphone cameras). The intrinsic nature of such systems, however, makes them the most vulnerable ones. An impostor can perform illegal access in such systems by presenting a synthetic sample to the acquisition sensor to impersonate a genuine user. An attack can be carried out through presenting, to the acquisition sensor, 2D-printed photos, electronic display of facial photos or videos, or 3D face masks. Mask-based attacks, although more sophisticated than the other forms, are increasingly easy to produce. The process of presenting synthetic samples to an acquisition sensor, however, inevitably includes noise information and telltales, which are added to the biometric signal and can be used to identify attempted attacks.

Despite being widely used for face recognition, data-driven models have just a recent history in the Face PAD problem and have been showing their potential to detect this kind of attack. Existing solutions are distinct, but a slight tendency can be perceived for the ones based on neural networks. Usually, pre-trained Convolution Neural Network (CNN) architectures are used as feature extractors, and these features are then used to train a classifier (e.g., SVM). Ito *et al.* [17], for instance, have investigated two different CNN architectures for Face PAD: CIFAR-10 and AlexNet. Instead of using cropped images of faces (as in traditional face recognition literature), the authors used the whole image as input to their method. The rationale behind this approach is that by exploiting the whole image, more information about the artifacts present in synthetic samples can be acquired. Although the proposed method overcame some baselines, experiments were performed only on one dataset. Thus, no general conclusion can be drawn about the robustness of the presented methods.

Due to the nature of data-driven approaches, it is not always possible to decode the artifacts in attacks that are being exploited by the model. The model is entirely in charge of extracting features from the data (images or videos) that maximize the learning process. This aspect, however, can be partially controlled by extracting dynamic and static features. In this context, Wu *et al.* [52] proposed a well-engineered data-driven method. The idea is that, by combining the movements of a person in a video (dynamic features) with texture features from the frames (static features), complementary telltales of an attack can be assessed. The method extracts static

features frame-by-frame using a CNN. For dynamic features, however, it employs the horizontal and vertical optical flow by using the Lucas-Kanade Pyramid method to extract dynamic maps from the frames followed by the CNN on the dynamic maps. Both static and dynamic features are combined through concatenation and used as input to a binary SVM classifier for decision making.

Yang *et al.* [55] have also explored static and dynamic features. An initial step is based on the use of Local Binary Patterns (LBP) descriptors to extract more generalized and discriminative low-level features of face images. LBP features are successfully applied in an intra-dataset protocol, but the performance may degrade severely in a more realistic scenario, i.e., inter- or cross-dataset protocol, due to factors such as abnormal shadings, specular highlights, and device noise [6]. For that reason, the authors encoded these low-level features into high-level features via deep learning and proposed a sparse auto-encoder (SAE) to tackle the aforementioned issues. SAE consists on the application of a sparse penalty in a traditional auto-encoder, which is an axisymmetric single hidden-layer neural network, to strengthen the generalization ability of the model. It has a significant advantage when addressing complex problems: extract characteristics that reflect the adhesion state. Finally, a binary SVM classifier is used to distinguish genuine from synthetic face samples. Nonetheless, the training was only performed with an intra-dataset protocol.

In the last years, dictionary learning, a well-known machine learning concept, has been introduced as a candidate to build deep architectures, creating a new branch called Deep Dictionary Learning (DDL). The idea consists of stacking up dictionary learning layers to form a DDL [48] structure. Manjani *et al.* [25] developed a DDL approach in which layers of single-level dictionaries are stacked one after another, yielding a sparse representation of features. The main advantages are the mitigation of the requirement of large training datasets, promising intra-dataset results, and the discernment between different types of attacks, even unknown ones. However, the main concerns about this representation are the difficulty of extracting fine-grained features to deal with real mask attacks, and the lack of generalization of the method.

3.2 Iris PAD

Despite the better accuracy of iris authentication methods in comparison with methods based on face and fingerprint traits, the use of this technology was limited to protect only high-restrict systems and places due to mainly the costs associated to implementation of this technology. Nowadays, iris authentication methods permeate our daily life due to research efforts toward making processes and sensors cheaper and smaller, as we can found in modern computing devices, such as smartphones.

However, even the high-accuracy and advances in iris biometrics, the current iris-based authentication systems still suffer from vulnerabilities to presentation attacks. Currently, the most effective PA methods to bypass an iris authentication system consist of showing to an acquisition sensor printed photos containing iris patterns of legitimate users, textured and transparent contact lenses used for impersonating a

legitimate user or for concealing the real identity of an attacker. From this perspective, we categorized the existing works on iris presentation attack detection into two non-disjoint groups based on their ability to detect the following attack types: print-based attacks, performed with printout iris images; and methods aiming at detecting attempted attacks performed with contact lenses.

The first CNN-based approach proposed to detect iris presentation attacks was presented to the community by Menotti *et al.* [27]. In this study, the authors presented a unified framework to perform architecture- and filter-level optimization for three biometric modalities (iris, face, and fingerprint). The proposed framework was developed based on a shallow CNN architecture, the SpoofNet, with two convolutional layers tailored to detect PAs in different modalities. This framework was convenient to the community due to the small size of the datasets freely available at the time to build iris PA systems. A drawback of the study, regarding the iris modality, relies on the fact that the SpoofNet was evaluated only for print-based attempted attacks.

Similarly, Silva *et al.* [39] presented a study that evaluated the SpoofNet in other attack scenarios different from those proposed in [27]. In that work, the authors proposed some training methodologies considering the Notre Dame and IIIT-Delhi datasets, which are composed by Near Infrared (NIR) iris images that represent *bona fide* presentations and presentation attacks performed with textured (colored) contact lenses and soft (transparent) contacted lenses. The authors also adapted the SpoofNet to detect three type of images: non attack, textured, and transparent contact lenses. The proposed method outperformed existing methods for the Notre Dame dataset achieving an overall accuracy of 82.80%. The reported results suggested that SpoofNet was able to detect transparent contact lenses better than textured contact lenses and *bona fide* presentations (iris images without any contact lenses for this particular dataset). The obtained results considering the IIIT-Delhi dataset reveal some limitations of the SpoofNet architecture to distinguish these different kinds of presentations, especially the confusion of bone fide samples with transparent contact lenses attacks.

Toward overcoming the SpoofNet limitation, Raghavendra *et al.* [36] proposed a novel CNN architecture more robust to distinguish *bona fide* presentations, textured, and transparent contact lenses. Similarly to Silva *et al.* [39], the authors proposed a multi-class CNN designed to classify an input image into *bona fide* and PA performed with textured and transparent contact lenses. However, similar to He *et al.*, the authors used normalized iris image patches as input to the CNN, while Silva *et al.* fed their CNN with raw images and also used a six-layer CNN and dropout, a mechanism to reduce over-fitting of the network.

In [16], He *et al.* proposed a multi-patch CNN capable of detecting both attack types, print- and contact lens-based attempted attacks, by training a CNN with normalized iris image patches of size 80×80 pixels, which can significantly reduce the trainable parameters of a deep network and, therefore, prevent possible generalization problems such as over-fitting. The authors reported near-perfect classification results for both attack types. The comparison among the proposed method and other hand-craft methods in prior art shows the superiority of CNN-based approaches over feature engineering-based methods.

Pala and Bhanu [29] developed a deep learning approach based on triplet convolutional neural networks, whereby three networks map iris image patches into a representation space. This is done by either taking two real examples and a fake one or two fakes and a real one, yielding intra-class and inter-class distances. The goal is to learn a distance function so that two examples taken from the same class are closer together than two examples taken from different classes. Two samples of the same class are as close as possible, according to the learned distance function. The method was evaluated in three different datasets containing print- and contact-lens attack and compared with descriptor-based methods and a CNN approach from [27], achieving the lowest average classification error rates for all datasets. The main advantage of this framework relies on its small architecture, being easy to implement on hardware, with reduced computational complexity. However, there is no guarantee that the framework performs well in a cross-dataset scenario as no tests in this regard were discussed in their work.

Kuehlkamp *et al.* [21] combined handcrafted and data-driven features to generate multiple transformations on the input data looking for more appropriate input-space representations. Handcrafted features are obtained by extracting multiple viewpoints of binarized statistical image features (BSIF), which are then used to train lightweight CNNs. After that, a meta-analysis technique is used for selecting the most important and discriminative set of classifiers, performing meta-fusion from selected viewpoints to build a final classification model that performs well not only under cross-domain constraints, but also under intra- and cross-dataset setups. As an advantage, this approach offers an iris PAD algorithm that better generalizes to unknown attack types, also outperforming state-of-the-art methods in this regard.

3.3 Fingerprint PAD

Fingerprints are one of the most present biometric traits nowadays, being widely adopted in security systems and sensitive environments. In some cases, a biometric system could be potentially defenseless against fake fingerprints. However, research has been made to mitigate the risk of attacks by proposing software- or hardware-based solutions. Among the hardware-based solutions, fingerprint liveness detection has been considered by most of the recent works. For software-based solutions, deep learning approaches play a crucial role, yielding state-of-the-art results.

Pala et al. [30] proposed a patch-based triplet siamese network for fingerprint PAD. Under a classical binary classification formulation (live/fake), the network comprises a deep metric learning framework that can generate representative features of real and artificial fingerprints. The proposed method evaluates liveness by comparing with the same fingerprint set of patches used for training, instead of requiring an enrollment database. It also tackles the limitations of current deep learning approaches regarding computational cost, thus allowing mobile and off-line implementation.

In [28], three well-known deep learning networks were utilized in the form of a binary classification problem. The networks were fine-tuned using the weights of

a pre-trained network originally trained on the ImageNet dataset [20], rather than training them from scratch for each network. The authors have shown the effect of data augmentation techniques not only in the case of deep learning framework but also when a shallow technique such as LBP was utilized. Additionally, they followed an experimental protocol taking cross-dataset validation into consideration and made a significant comparison among the methods in their approach. Sundaran *et al.* [45] showed how training a single CNN-based classifier using different available datasets can aid generalization and boost performance.

An analytical study has been made in [49] for feature fusion by taking into account different features and methods. A two-stage deep neural network that starts from general image descriptors was adopted. In the first stage, the method is capable of simultaneously learning a transformation of different features into a common latent space used for classification in a second stage. Nogueira *et al.* [28] compared four deep learning techniques for liveness detection. They studied the effect of using pre-trained weights, concluding that using a pre-trained CNN could yield good results without considering modifying neither the architectures nor their hyperparameters.

Deep Boltzmann Machines (DBMs) are another type of neural network that consist of learning stochastic energy-based on complex patterns. DBMs were considered in [42] and [41] for liveness detection in fingerprint data. In [42], a fingerprint spoofing detection method was proposed based on DBMs. After the network was trained on fingerprint data samples, a SVM classifier was trained in order to classify the high-level features generated by the DBMs. In [41], the authors proposed a DBM-based method which had a final layer added at the top of the network with two softmax units, forming an MLP network, to identify normal and attack patterns.

Toosi *et al.* [50] proposed a patch-based approach for liveness detection. The method extracts a set of patches and then a classifier is applied for each patch by utilizing the AlexNet architecture as a feature extractor. The final class label is computed based on the probability scores of patches. Another patch-based method [31] attempted to detect liveness based on a voting strategy on patches classified by a CNN.

Wang *et al.* [51] developed FingerNet, a DNN for fingerprint liveness detection with a voting strategy at the end for decision making. Its architecture was inspired by another DNN called CIFAR-10, with the difference that the convolutional and pooling layers are more complex, besides the inclusion of an extra inner product layer. The training process, however, is not done directly on the images from these datasets. Instead, each image is cut into patches of 32×32 pixels, followed by a segmentation step that excludes patches depicting background content, leaving the remaining ones for training. The test process is also performed on image patches, and the voting strategy is then applied by computing labels of each patch within a fingerprint image. The label with the highest vote is chosen as the image label. Experiments were performed with LivDet2011 and LivDet2013 datasets, with FingerNet outperforming CIFAR-10.

Zhang *et al.* [56] improved the Inception [46] architecture and built a lightweight CNN for fake fingerprint detection. In the proposed architecture, the original fully-connected layer was replaced by a global average pooling layer to reduce overfitting

and enhance robustness to spatial translations. For the experiments, the authors created an in-house 2D dataset with fingerprints made from different materials (along with some live examples). The reported results were expressed regarding a weighted average rate of correctly classified live fingerprints and fake fingerprints, outperforming not only the original Inception architecture but also other methods from the literature based on CNNs.

4 Countermeasures for Face, Iris, and Fingerprint Presentation Attack Detection

In this section, we discuss two approaches to deploy countermeasures based on data-driven models for detecting presentation attacks, which were successfully used to detect face-, iris-, and fingerprint-based presentation attacks in prior art.

4.1 Architecture and Filter Optimization

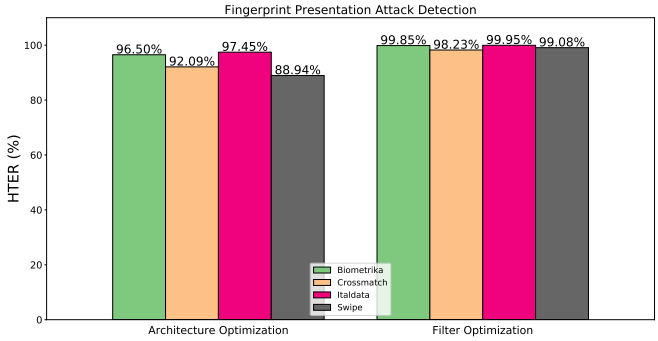
Finding suitable architectures for a given application is a challenge and time-consuming task due to the high dimensionality of the hyper-parameter space. Moreover, the absence of enough data makes the hyper-parameter search harder for some applications. For this reason, several techniques have been proposed in the literature toward mitigating these problems by proposing heuristics to find suitable architectures faster.

The Tree-structured Parzen Estimator (TPE) is a sequential model-based optimization approach [3], which can construct models to approximate the performance of hyper-parameters based on historical observations. In summary, the TPE algorithm models probabilities $P(x|y)$ and $P(y)$, where x represents hyper-parameters and y represents the loss function's value associated with x. This modeling is performed by using historical observations to estimate non-parametric density distributions for the hyper-parameters, which is used to predict good values for them [2].

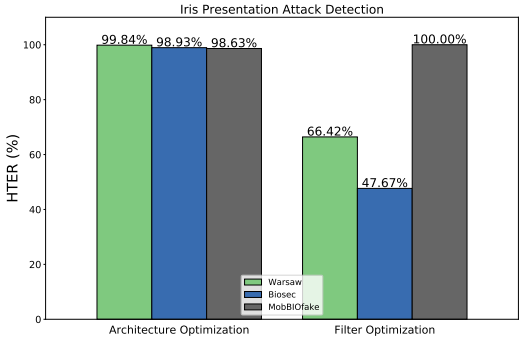
Another approach to explore the hyper-parameter space toward finding suitable architectures for a given problem is the random search algorithm. In general, this approach is more efficient than manual search and the grid search algorithm [35]. The random search algorithm explores the hyper-parameter space by identifying valid hyper-parameter assignments, which defines a valid configuration space for the problem. Finally, hyper-parameters are randomly selected, considering a uniform distribution. The main advantage of this strategy is that it is simpler to implement it in non-parallel and parallel versions.

The search for good filter weights in a deep learning architecture is also a challenging task. The huge number of parameters and the small size of the datasets available for PAD in the literature may prevent optimization algorithms to find an optimal or even reasonable solution.

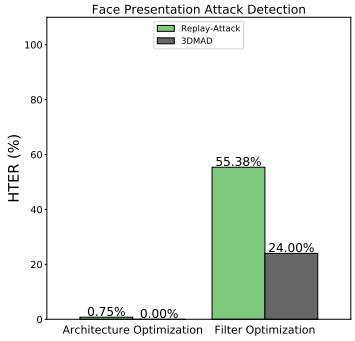
According to Menotti *et al.* [27], architecture and filter optimization strategies are effective toward deploying suitable deep learning models. Fig. 2 shows a comparison study of both strategies. The architecture optimization was performed considering shallow architectures with up to three layers and filter kernel with random weights, while the filter optimization was performed considering pre-defined architectures to the problem and Stochastic Gradient Descent (SGD) for optimizing the filter weights, which were initialized randomly.



(a) Fingerprint modality.



(b) Iris modality.



(c) Face modality.

Fig. 2 Filter and architecture optimization results for face, fingerprint and iris presentation attack detection considering different available datasets.

As we can observe, the architecture optimization approach was able to find architectures with near-perfect classification results for face PAD problem (Fig. 2(c)), with a Half Total Error Rate of $\approx 0.0\%$. For the iris PAD problem, this approach also found an architecture with near-perfect classification results, whose accuracy was superior to the filter optimization strategy in two of the three datasets (i.e., Warsaw and Biosec) evaluated for this modality and competitive results for the MobBio fake dataset, as illustrated in Fig. 2(b). In contrast, for the fingerprint PAD problem, the filter optimization presented better results than the architecture optimization

approach, obtaining a near-perfect classification result for all datasets evaluated in this modality (Fig. 2(a)). For this reason, the interplay of these two approaches is recommended, in cases that architecture optimization is not enough to find good solutions. The two options are recommended specially when there is not enough training data to represent the PAD problem of interest.

4.2 Fine-tuning of Existing Architectures

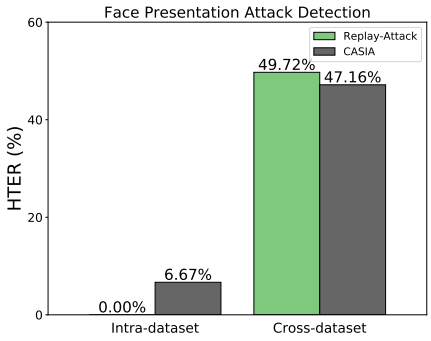
A second trend in the literature consists of building models using transfer learning techniques to adapt pre-existing deep learning solutions pretrained with thousand of hundreds of images to PAD problem. Depending on the modality, the transfer-learning process could take advantage of pretrained architectures whose source problem is related to the target problem, for instance, a pretrained architecture for face recognition (source) for optimizing a deep learning architecture for the face PAD problem (target).

With this in mind, Pinto [33] adapted the VGG network architecture [40], which was originally proposed for object recognition by the Visual Geometry Group, by transferring the knowledge obtained from the training process conducted with a huge dataset, the ImageNet [20] dataset. The fine-tuned architectures were evaluated considering the original protocols of the datasets used by the authors, as well as cross-dataset protocol, whose training and testing subsets come from different sources. Fig. 3 shows results for these two evaluation protocols. For all modalities, the cross-dataset evaluation protocol presented poor results in comparison with the intra-dataset evaluation protocol, with exception to ATVS Iris dataset. In several cases, the performance of PAD solutions ranges from near-perfect classification results to worse than random performance, showing us a clear need of focus for new techniques considering the cross-dataset setup as well as robustness to unseen attacks (open-set scenarios).

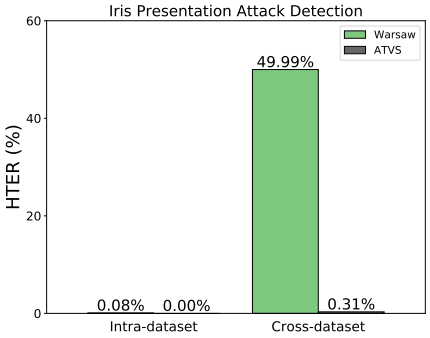
5 Challenges, Open Questions and Outlook

This section presents the main existing limitations of current methods and shed some light on aspects that further research paths could undertake in order to tackle PAD problems.

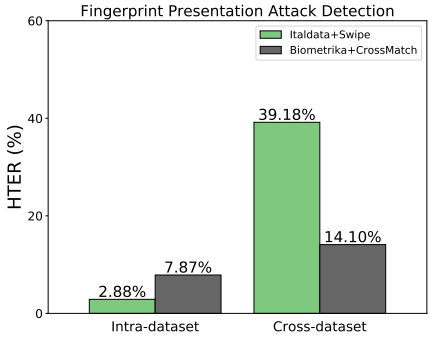
First of all, one of the main aspects preventing the introduction of robust methods is the lack of representative public datasets. Oftentimes the existing datasets lack generalization features and are normally small when we consider the era of big data. This limitation leads to another one, which is the difficulty of performing cross-dataset and domain-adaptation validations. In the absence of data, it is hard to learn the changing aspects from one dataset to another or from one condition to another.



(a) Face modality.



(b) Iris modality.



(c) Fingerprint modality.

Fig. 3 Intra- and cross-data validation results considering different datasets and biometric modalities. Note how performance degrades when we move from the controlled setting of intra-dataset validation to the wild settings of cross-dataset validation.

Although we can see some reasonable effort from researchers in designing techniques for iris and fingerprint presentation attack detection, the same is still not true for faces. Most works in this modality still rely on handcrafted features. Probably this fact is related to the lack of good and representative datasets but we believe the most likely reason is that faces vary much more often than iris and fingerprints (shadows, lighting, pose, scaling) and deep-learning methods would need very robust functions to deal with such transformations without incurring in overfitting. Here is probably a very nice area of research.

Existing work in iris PAD using deep learning seems to be ahead of fingerprint and face modalities. A recent method proposed by [21] seems to represent an important step toward cross-dataset validation. The metafusion of different views of the input data has yielded the best results and represents a good tackle on handling variations of different datasets. Despite this, faces and fingerprints are still far from a reasonable path toward solving the problem. For fingerprints, there is a sensible change in performance when changing acquisition sensors. For face, although ac-

quisition sensors play an important role in rendering the problem difficult to solve, old problems known to researchers since the development of the first face recognition methods are still present such as shadow and illumination changes. Even in the iris case, oftentimes existing methods fail under domain shift conditions caused by cross-sensors and also for detecting transparent contact lenses. Clear efforts from the community are needed in this regard as well.

Recently, Li et al. [23] have shed some light on learning generalized features for face presentation attack detection. The authors proposed training their solution with augmented facial samples based on cross-entropy loss and further enhanced the training with a specifically designed generalization loss, which coherently serves as a regularization term. The training samples from different domains can seamlessly work together for learning the generalized feature representation by manipulating their feature distribution distances. When learning domain-shift conditions, it seems that proposed robust loss functions, as well as ways of implicitly learning data variations, is a promising path to solve this hard problem.

With respect to open-set validation conditions and robustness against unseen types of attacks, the literature is far behind. One of the first works considering this aspect was presented in the context of fingerprint PAD [37], but the authors show that much is still to be done in this regard. For faces and irises, the story is even worse in this regard. This is surely one of the hottest topics in PAD nowadays along with domain shift robustness as non-existing work shows reasonable performance for open-set conditions in any of face, fingerprint or iris modalities.

Thus far, most of the existing prior art in PAD relying on deep learning methods have only scratched the surface of their potential. Virtually all methods in the prior art only adopt existing networks and tweak them somehow. Faces are a bit different in this regard, as some authors have investigated the effects of dictionary learning and stacking of different solutions. Nonetheless, few authors have endeavored to propose significant modeling changes to such solutions, and this surely represents an open avenue of opportunities.

Finally, it is worth mentioning that although we surveyed some 60 papers in this chapter, the literature still presents a series of open problems in face presentation attack detection ranging from proposing representative datasets to robust methods under the cross-dataset and open-set validation setups to representative methods robust to domain shifts. Calling the attention of the community for such problems was the main motivation that led us to write this chapter, and we sincerely hope researchers will consider these aspects in their future investigations for their PAD problems.

Acknowledgements The authors thank the financial support of the European Union through the Horizon 2020 Identity Project as well as the São Paulo Research Foundation – Fapesp, through the grant #2017/12646-3 (DéjàVu), and the Brazilian Coordination for the Improvement of Higher Education Personnel – Capes, through the DeepEyes grant.

Appendix 1

Datasets and research work.

Table 4: Datasets per modality and prior work relying on them.

Modality	Dataset	References
Face PAD	Replay-Attack	[25, 38, 17, 52, 6]
	SMAD	[25]
	CASIA-FASD	[25, 38, 55, 27, 6]
	3DMAD	[25, 24, 27]
	MSU MFSD	[24, 6]
	MSU USSA	
	NUAA	[55, 6]
	UVAD	[25]
Iris PAD	Clarkson17	[21]
	IIITD Iris Spoofing	[19]
	IIITD Delhi	[19, 39, 36]
	ND CLD 2015	[21, 14, 16, 39, 36]
	Warsaw17	[21, 29, 16, 27]
	IIITD WVU	[21]
	CASIA Iris Fake	[16]
Fingerprint PAD	LivDet 2009	[30, 28, 49, 31]
	LivDet 2011	[30, 28, 49, 50, 51]
	LivDet 2013	[30, 28, 49, 41, 50, 51, 42, 27]
	PBSKD	
	Bogus	[45]

Appendix 2

List of acronyms.

BSIF Binary Statistical Image Features
CNN Convolutional Neural Network
DBM Deep Boltzmann Machine
DCNN Deep Convolutional Neural Network
DNN Deep Neural Network
HTER Half Total Error Rate
LBP Local Binary Pattern
MLP Multilayer Perceptron
PA Presentation Attack
PAD Presentation Attack Detection
SAE Sparse Auto-Encoder
SVM Support Vector Machines

References

1. BBC News: Doctor 'used silicone fingers' to sign in for colleagues. BBC News, https://www.bbc.com/news/world-latin-america-21756709 (2013). Accessed: 2019-01-13
2. Bergstra, J., Bardenet, R., Bengio, Y., Kégl, B.: Algorithms for hyper-parameter optimization. In: Advances in neural information processing systems, pp. 2546–2554 (2011)
3. Bergstra, J., Yamins, D., Cox, D.D.: Making a science of model search: Hyperparameter optimization in hundreds of dimensions for vision architectures. In: International Conference on Machine Learning, pp. I–115–I–123 (2013)
4. Bhattacharjee, S., Mohammadi, A., Marcel, S.: Spoofing deep face recognition with custom silicone masks. In: IEEE International Conference on Biometrics: Theory, Applications and Systems (2018)
5. Boulkenafet, Z., Komulainen, J., Li, L., Feng, X., Hadid, A.: OULU-NPU: A mobile face presentation attack database with real-world variations. In: IEEE International Conference on Automatic Face and Gesture Recognition (2017)
6. Cai, G., Su, S., Leng, C., Wu, J., Wu, Y., Li, S.: Cover patches: A general feature extraction strategy for spoofing detection. Concurrency and Computation: Practice and Experience p. e4641 (2018)
7. Chingovska, I., Anjos, A., Marcel, S.: On the effectiveness of local binary patterns in face anti-spoofing. In: International Conference of the Biometrics Special Interest Group (2012)
8. Chugh, T., Cao, K., Jain, A.K.: Fingerprint spoof buster: Use of minutiae-centered patches. IEEE Transactions on Information Forensics and Security **13**(9), 2190–2202 (2018)
9. Delac, K., Grgic, M.: A survey of biometric recognition methods. In: International Symposium Electronics in Marine, vol. 46, pp. 16–18 (2004)
10. Doyle, J.S., Bowyer, K.W.: Robust detection of textured contact lenses in iris recognition using BSIF. IEEE Access **3**, 1672–1683 (2015)
11. Erdogmus, N., Marcel, S.: Spoofing in 2D face recognition with 3D masks and anti-spoofing with Kinect. In: IEEE International Conference on Biometrics: Theory, Applications and Systems (2013)
12. Erdoğmuş, N., Marcel, S.: Introduction. In: Handbook of Biometric Anti-Spoofing: Trusted Biometrics under Spoofing Attacks, pp. 1–11. Springer London, London (2014)
13. Ghiani, L., Yambay, D., Mura, V., Tocco, S., Marcialis, G.L., Roli, F., Schuckcrs, S.: LivDet 2013 fingerprint liveness detection competition 2013. In: IAPR International Conference on Biometrics, pp. 1–6 (2013)
14. Gragnaniello, D., Sansone, C., Poggi, G., Verdoliva, L.: Biometric spoofing detection by a domain-aware convolutional neural network. In: International Conference on Signal Image Technology and Internet-Based Systems, pp. 193–198 (2017)
15. Gupta, P., Behera, S., Vatsa, M., Singh, R.: On iris spoofing using print attack. In: International Conference on Pattern Recognition, pp. 1681–1686 (2014)
16. He, L., Li, H., Liu, F., Liu, N., Sun, Z., He, Z.: Multi-patch convolution neural network for iris liveness detection. In: IEEE International Conference on Biometrics: Theory, Applications and Systems, pp. 1–7 (2016)
17. Ito, K., Okano, T., Aoki, T.: Recent advances in biometrie security: A case study of liveness detection in face recognition. In: Asia-Pacific Signal and Information Processing Association Annual Summit and Conference, pp. 220–227 (2017)
18. Jain, A.K., Ross, A., Prabhakar, S.: An introduction to biometric recognition. IEEE Transactions on Circuits and Systems for Video Technology **14**(1), 4–20 (2004)
19. Kohli, N., Yadav, D., Vatsa, M., Singh, R., Noore, A.: Synthetic iris presentation attack using iDCGAN. In: IEEE International Joint Conference on Biometrics, pp. 674–680 (2018)
20. Krizhevsky, A., Sutskever, I., Hinton, G.: ImageNet classification with deep convolutional neural networks. In: Advances in neural information processing systems, pp. 1097–1105 (2012)
21. Kuehlkamp, A., Pinto, A., Rocha, A., Bowyer, K.W., Czajka, A.: Ensemble of multi-view learning classifiers for cross-domain iris presentation attack detection. IEEE Transactions on Information Forensics and Security **4**(6) (2019)

22. Kumar, A., Passi, A.: Comparison and combination of iris matchers for reliable personal authentication. Pattern Recognition **43**(3), 1016–1026 (2010)
23. Li, H., He, P., Wang, S., Rocha, A., Jiang, X., Kot, A.C.: Learning generalized deep feature representation for face anti-spoofing. IEEE Transactions on Information Forensics and Security **13**(10), 2639–2652 (2018)
24. Li, X., Komulainen, J., Zhao, G., Yuen, P.C., Pietikäinen, M.: Generalized face anti-spoofing by detecting pulse from face videos. In: International Conference on Pattern Recognition, pp. 4244–4249 (2016)
25. Manjani, I., Tariyal, S., Vatsa, M., Singh, R., Majumdar, A.: Detecting silicone mask-based presentation attack via deep dictionary learning. IEEE Transactions on Information Forensics and Security **12**(7), 1713–1723 (2017)
26. Marcialis, G.L., Lewicke, A., Tan, B., Coli, P., Grimberg, D., Congiu, A., Tidu, A., Roli, F., Schuckers, S.: First international fingerprint liveness detection competition – LivDet 2009. In: International Conference on Image Analysis and Processing, pp. 12–23 (2009)
27. Menotti, D., Chiachia, G., Pinto, A., Schwartz, W.R., Pedrini, H., Falcao, A.X., Rocha, A.: Deep representations for iris, face, and fingerprint spoofing detection. IEEE Transactions on Information Forensics and Security **10**(4), 864–879 (2015)
28. Nogueira, R.F., de Alencar Lotufo, R., Machado, R.C.: Fingerprint liveness detection using convolutional neural networks. IEEE Transactions on Information Forensics and Security **11**(6), 1206–1213 (2016)
29. Pala, F., Bhanu, B.: Iris liveness detection by relative distance comparisons. In: IEEE Conference on Computer Vision and Pattern Recognition Workshops (2017)
30. Pala, F., Bhanu, B.: On the accuracy and robustness of deep triplet embedding for fingerprint liveness detection. In: IEEE International Conference on Image Processing, pp. 116–120 (2017)
31. Park, E., Kim, W., Li, Q., Kim, J., Kim, H.: Fingerprint liveness detection using cnn features of random sample patches. In: International Conference of the Biometrics Special Interest Group, pp. 1–4 (2016)
32. Patel, K., Han, H., Jain, A.K.: Secure face unlock: Spoof detection on smartphones. IEEE Transactions on Information Forensics and Security **11**(10), 2268–2283 (2016)
33. Pinto, A., Pedrini, H., Krumdick, M., Becker, B., Czajka, A., Bowyer, K.W., Rocha, A.: Deep Learning in Biometrics, chap. Counteracting Presentation Attacks in Face Fingerprint and Iris Recognition. Taylor and Francis (2018)
34. Pinto, A., Schwartz, W.R., Pedrini, H., de Rezende Rocha, A.: Using visual rhythms for detecting video-based facial spoof attacks. IEEE Transactions on Information Forensics and Security **10**(5), 1025–1038 (2015)
35. Pinto, N., Doukhan, D., DiCarlo, J., Cox, D.: A high-throughput screening approach to discovering good forms of biologically-inspired visual representation. PLoS Computational Biology **5**(11), e1000579 (2009)
36. Raghavendra, R., Raja, K.B., Busch, C.: ContlensNet: Robust iris contact lens detection using deep convolutional neural networks. In: IEEE Winter Conference on Applications of Computer Vision, pp. 1160–1167 (2017)
37. Rattani, A., Scheirer, W., Ross, A.: Open set fingerprint spoof detection across novel fabrication materials. IEEE Transactions on Information Forensics and Security **10**(11), 2447–2460 (2015)
38. Rehman, Y.A.U., Po, L.M., Liu, M.: LiveNet : Improving features generalization for face liveness detection using convolution neural networks. Expert Systems with Applications **108**, 159–169 (2018)
39. Silva, P., Luz, E., Baeta, R., Pedrini, H., Falcao, A., Menotti, D.: An approach to iris contact lens detection based on deep image representations. In: Conference on Graphics, Patterns and Images, pp. 157–164 (2015)
40. Simonyan, K., Zisserman, A.: Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556 (2014)
41. Souza, G.B., Santos, D.F., Pires, R.G., Marana, A.N., Papa, J.P.: Deep boltzmann machines for robust fingerprint spoofing attack detection. In: International Joint Conference on Neural Networks, pp. 1863–1870 (2017)

42. de Souza, G.B., da Silva Santos, D.F., Pires, R.G., Marana, A.N., Papa, J.P.: Deep features extraction for robust fingerprint spoofing attack detection. Journal of Artificial Intelligence and Soft Computing Research **9**(1), 41–49 (2019)
43. Sun, Z., Tan, T.: Iris anti-spoofing. In: Handbook of Biometric Anti-Spoofing: Trusted Biometrics under Spoofing Attacks, pp. 103–123. Springer London, London (2014)
44. Sun, Z., Zhang, H., Tan, T., Wang, J.: Iris image classification based on hierarchical visual codebook. IEEE Transactions on Pattern Analysis and Machine Intelligence **36**(6), 1120–1133 (2014)
45. Sundaran, S., Antony, J.K., Vipin, K.: Biometrie liveness authentication detection. In: International Conference on Innovations in Information, Embedded and Communication Systems, pp. 1–3 (2017)
46. Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., Rabinovich, A.: Going deeper with convolutions. In: IEEE Conference on Computer Vision and Pattern Recognition, pp. 1–9 (2015)
47. Tan, X., Li, Y., Liu, J., Jiang, L.: Face liveness detection from a single image with sparse low rank bilinear discriminative model. In: European Conference on Computer Vision, pp. 504–51 (2010)
48. Tariyal, S., Majumdar, A., Singh, R., Vatsa, M.: Deep dictionary learning. IEEE Access **4**, 10096–10109 (2016)
49. Toosi, A., Bottino, A., Cumani, S., Negri, P., Sottile, P.L.: Feature fusion for fingerprint liveness detection: a comparative study. IEEE Access **5**, 23695–23709 (2017)
50. Toosi, A., Cumani, S., Bottino, A.: CNN patch-based voting for fingerprint liveness detection. In: International Joint Conference on Computational Intelligence, pp. 158–165 (2017)
51. Wang, C., Li, K., Wu, Z., Zhao, Q.: A DCNN based fingerprint liveness detection algorithm with voting strategy. In: Chinese Conference on Biometric Recognition, pp. 241–249 (2015)
52. Wu, L., Xu, Y., Xu, X., Qi, W., Jian, M.: A face liveness detection scheme to combining static and dynamic features. In: Chinese Conference on Biometric Recognition, pp. 628–636 (2016)
53. Yambay, D., Becker, B., Kohli, N., Yadav, D., Czajka, A., Bowyer, K.W., Schuckers, S., Singh, R., Vatsa, M., Noore, A., et al.: LivDet iris 2017 – iris liveness detection competition 2017. In: IEEE International Joint Conference on Biometrics, pp. 733–741 (2017)
54. Yambay, D., Ghiani, L., Denti, P., Marcialis, G.L., Roli, F., Schuckers, S.: LivDet 2011 – fingerprint liveness detection competition 2011. In: IAPR International Conference on Biometrics, pp. 208–215 (2012)
55. Yang, D., Lai, J., Mei, L.: Deep representations based on sparse auto-encoder networks for face spoofing detection. In: Chinese Conference on Biometric Recognition, pp. 620–627 (2016)
56. Zhang, Y., Zhou, B., Qiu, X., Wu, H., Zhan, X.: 2D fake fingerprint detection for portable devices using improved light convolutional neural networks. In: Chinese Conference on Biometric Recognition, pp. 353–360 (2017)
57. Zhang, Z., Yan, J., Liu, S., Lei, Z., Yi, D., Li, S.Z.: A face antispoofing database with diverse attacks. In: IAPR International Conference on Biometrics, pp. 26–31 (2012)