Recent Advances in Face Presentation Attack Detection

Sushil Bhattacharjee, Amir Mohammadi, André Anjos, and Sébastien Marcel

Abstract The undeniable convenience of face-recognition (FR) based biometrics has made it an attractive tool for access control in various applications, from immigration-control to remote banking. Widespread adoption of face biometrics, however, depends on the how secure such systems are perceived to be. One particular vulnerability of FR systems comes from presentation attacks (PA), where a subject **A** attempts to impersonate another subject **B**, by presenting, for example, a photograph of **B** to the biometric sensor (*i.e.*, the camera). PAs are the most likely forms of attacks on face biometric systems, as the camera is the only component of the biometric system that is exposed to the outside world. Robust presentation attack detection (PAD) methods are necessary to construct secure FR based access control systems. The first edition of the Handbook of Biometric Anti-spoofing included two chapters on face-PAD. In this chapter we present the significant advances in face-PAD research since the publication of the first edition of this book. In addition to PAD methods designed to work with color images, we also discuss advances in face-PAD methods using other imaging modalities, namely, near-infrared (NIR) and thermal imaging. This chapter also presents a number of recently published datasets for face-PAD experiments.

Sushil Bhattacharjee
IDIAP Research Institute, Centre du Parc, Rue Marconi 19, PO Box 592, CH - 1920 Martigny Switzerland, e-mail: sushil.bhattacharjee@idiap.ch

Amir Mohammadi
IDIAP Research Institute, Centre du Parc, Rue Marconi 19, PO Box 592, CH - 1920 Martigny Switzerland, e-mail: amir.mohammadi@idiap.ch

André Anjos
IDIAP Research Institute, Centre du Parc, Rue Marconi 19, PO Box 592, CH - 1920 Martigny Switzerland, e-mail: andre.anjos@idiap.ch

Sebastien Marcel
IDIAP Research Institute, Centre du Parc, Rue Marconi 19, PO Box 592, CH - 1920 Martigny Switzerland, e-mail: sebastien.marcel@idiap.ch

1 Introduction

As pointed out by Ratha et al. [48] and many other researchers, biometrics based access-control systems can be attacked in several ways. Most kinds of attacks on a biometric system require privileged access to the various components of the system. The biometric sensor in the system is the easiest to attack, as it is the most exposed component in the system. By definition, privileged access is not necessary to interact with the sensor. Attacks on the biometric sensor are called *presentation attacks* (PA).

The ISO standard for biometric presentation attack detection[1] defines a PA as "*a presentation to the biometric data capture subsystem with the goal of interfering with the operation of the biometric system*."

An attacker, **A**, mounts a PA on a previously enrolled identity, **B**, using a *presentation attack instrument* (PAI). For FR systems, common PAIs are images, videos, or even 3D masks depicting the victim **B**. Such attacks fall into the category of *impersonation* attacks. It is important to note that the ISO standard also includes *obfuscation* as a kind of PA. An obfuscation attack is said to occur when the attacker attempts to spoof the biometric sensor in order to avoid being correctly recognized. Classic examples of obfuscation in face biometrics are the use of clothing or facial makeup, or a mask to avoid identification by a FR system.

Presentation attack detection (PAD) is an essential component in any secure biometric system. The first edition of this handbook included a comprehensive chapter describing the approaches face-PAD. In this chapter we review advances in face-PAD research since the publication of the first edition. Specifically, we review significant works in face-PAD published since the year 2015. Besides discussing the significant face-PAD methods proposed in the past three years, we also describe recently published datasets useful for research on this topic.

1.1 Standardization Efforts

One of the most significant developments in PAD has been the formal adoption of ISO standards[2]1. Among other things, the standard defines several metrics for reporting experimental results. The metrics relevant to this chapter are listed below:

- IAPMR: the *Impostor Attack Presentation Match Rate* quantifies the vulnerability of a biometric system, and is given as the proportion of impostor attack presentations that are incorrectly accepted by the biometric security system,
- APCER: *Attack Presentation Classification Error Rate* gives the proportion of PAs that is accepted by the system in question, and,
- BPCER: *Bona fide Presentation Classification Error Rate* specifies the proportion of *bona fide* presentations that are incorrectly rejected by the system as PA.

[1] ISO/IEC 30107-1:2016 Part 1

Note that the IAPMR is computed in the licit scenario (the scenario where PAs are not expected, and every presentation is considered *bona fide*), whereas APCER and BPCER are computed in the PA scenario. There is a further subtlety to be taken into account when computing the APCER in a given experiment, namely, that APCER values should be computed separately for each PAI. In other words, for a FR system, separate APCER values should be determined for print-attacks, video-replay attacks, 3D-mask attacks, and so on. If an experiment includes attacks based on different PAIs, that is, if a certain test dataset contains PAs involving different kinds of PAIs, then the APCER corresponding to the PAI that is the most expensive (in terms of cost, as well as manufacturing effort) should be specified as the overall APCER achieved in the experiment. It is often more practical to report the BPCER when the APCER is no greater than a preset value, for example BPCER @ APCER=10% (sometimes abbreviated as BPCER10).

1.2 Structure of the Chapter

The remainder of the chapter is organized in four sections. In Section 2 we discuss some recent studies on the vulnerability of FR systems to PAs. This section highlights the importance of continuing research and development of face-PAD technology. Following the discussion on vulnerability, a range of recent research publications relevant to face-PAD are summarized in Section 3. To facilitate comparison with the state of the art, most research publications on face-PAD include results on publicly available datasets. As technology for mounting PAs improves, new datasets are needed to evaluate the performance of face-PAD algorithms. Section 4 presents a number of recent public datasets for face-PAD experiments. We end the chapter with concluding remarks in Section 5.

2 Vulnerability of FR Systems to PA

FR systems are explicitly trained to handle session-variability, that is, variability due to changes in scale, orientation, illumination, facial expressions, and to some extent even make-up, facial grooming, and so on. This capacity to deal with session-variability also opens the door to presentation attacks. In 2016, a wide-ranging European project (TABULA RASA[3]) hypothesized that the higher the efficacy of a FR system in distinguishing between genuine and zero-effort-impostor (ZEI) presentations, the more vulnerable the system is to PAs. Several studies investigating the vulnerability to PAs of various FR systems, under different scenarios, have provided quantitative evidence that most FR schemes are very vulnerable in this respect.

[3] http://www.tabularasa-euproject.org/

Hadid [22] analyses the vulnerability of a FR system that uses a parts-based Gaussian mixture-model (GMM). His experiments show that when the false rejection rate (FRR) is constrained to 0.1%, the presence of spoof attacks causes the false acceptance rate (FAR) of the trained GMM is 80%. In standardized metric terms, for this GMM-FR system, the IAPMR @ FAR=0.1% is 80%.

Raghavendra *et al.* [47] report on the vulnerability of a FR system relying on presentations in different spectral ranges. Their study is based on the Sparse Representation based Classifier (SRC) [57]. They capture 2D color-print PAIs (color face-images printed on two types of printers: laser, and ink-jet) in several wavelength bands, ranging from visible light (RGB) to near-infrared (NIR) (specifically, at the following seven wavelengths: 425nm, 475nm, 525nm, 570nm, 625nm, 680nm and 930nm). Evaluating the vulnerability in individual bands separately, they show that in almost all cases the chosen FR system shows very high vulnerability (IAPMR in the range of 95% – 100%). Only in one case, namely, laser-printed PAIs captured in the 930nm wavelength, does the IAPMR drop to acceptable levels (IAPMR = 1.25%). This experimental result is consistent with the finding that the reflectance of facial skin dips sharply in a narrow spectral-band around 970nm [25].

Deep learning based FR systems are now considered the state of the art. In the current decade convolutional neural networks (CNN) based FR systems have achieved near-perfect FR performance [40, 58, 50] on highly unconstrained datasets, such as the well known Labeled Faces in the Wild (LFW) dataset [24]. Mohammadi *et al.* [36] have studied the vulnerability of several CNN-FR systems. Their study, based on several publicly available PAD datasets, shows that CNN-FR systems are in fact more vulnerable (IAPMR up to 100%) to PAs than older FR methods.

One class of PAs not often considered is the morphed-image attack [18, 49]. Here, face images of two different subjects, say, **A** and **B**, are morphed into a single image. The morphed image is constructed to resemble both subjects sufficiently closely to pass a quick visual inspection. Then, if, say, subject **A** wishes to avoid detection at an international border, he may alter his passport using such a morphed-image to impersonate **B**. Raghavendra *et al.* [45] have shown, using a commercial off-the-shelf (COTS) FR system, that vulnerability of FR systems to morphed-image attacks may be as high as 100%.

3 Recent Approaches to Face PAD

It is not straightforward to impose a neat taxonomy on existing face-PAD approaches. Chingovska *et al.* [13] group face-PAD methods into three categories: motion based, texture based, and image-quality based. Other works [16] have considered image-quality based face-PAD methods as a subclass of texture-based methods. Ramachandra and Büsch [43] offer a hierarchical organization of face-PAD methods, with most general groups: hardware-based and software-based.

Here, it is not our aim to propose any specific taxonomy of face-PAD methods. To provide some order to our discussion, however, we have organized our survey of

recent face-PAD methods in several sections: methods that operate on visible-light imagery, methods the rely on inputs captured in wavelengths outside the visible-range of light, and a separate category of methods designed to detect 3D-mask based attacks. In the following discussion, the term *extended-range (ER) imagery* refers to data captured in wavelengths outside the visible-range of light.

3.1 Visible-Light Based Approaches

A majority of studies on face-PAD so far have relied exclusively on visible-light imagery (commonly called color imagery) as input. The term *visible light* here refers to the range of the electromagnetic spectrum – approximately from 380 to 750 nm – that is typically perceptible by the human visual system. One reason for the use of color-imagery is that the proliferation of high-quality and low-cost color cameras has made digital color-imagery widely accessible. Another reason is the need for face-PAD on mobile devices such as laptops, smartphones and tablet devices. With the sharp increase in the use of mobile devices in sensitive applications such as remote-banking and online education, secure identity-verification on such devices has become a critical issue. Although recently some companies have introduced products that include NIR cameras, a large majority of mobile devices still come with only color cameras. It is, therefore, important to continue developing face-PAD methods that can function with only color imagery as input.

Successful application of histograms of local binary pattern (LBP) coeffients to the problem of face-PAD [7, 13, 33] has made LBP and its various variants a mainstay for face-PAD. Initial LBP based methods for face-PAD relied on gray-level images. Boulkenafet *et al.* [8, 9] have used LBP features to characterize color-texture. For a given color image in RGB color-space, they first generate the YC_bC_r as well as HSV representations of the image. Uniform LBP histograms are then computed on the Y, C_b, C_r, H, S, and V components and concatenated together to generate the final feature-vector representing the input color image. These color-texture feature-vectors may be classified using support vector machines (SVM). Boulkenafet *et al.* have shown that color-texture features outperform gray-level LBP features in the face-PAD task [8]. In a separate work [9], they have also shown that this color-texture representation leads to significantly better generalization to unknown attacks, compared to other hand-crafted face-PAD features. Indeed, in a recent face-PAD competition [10], the winning entry also combined motion-information with color-texture information using LBP histograms.

Notwithstanding the success of LBP based methods, in the past three years researchers have also explored other approaches for face-PAD. Prominent recent works using color imagery have focussed on a variety of features characterizing local motion, local texture and more generally, image-quality. Wen *et al.* [56] propose several features for image distortion analysis (IDA) to tackle the problem of face-PAD for 2D (print and video-replay) attacks. Their features characterize the color-diversity, image-sharpness and the presence of specular regions in the input

images. The IDA features are computed only over the face-region (*i.e.*, on the output of the face-detection step), and are classified using a two-class SVM classifier. The authors present results on several public datasets, including a new dataset (MSU-MFSD, see Section 4) introduced in this paper. In intra-database experiments the IDA features perform competitively to other face-PAD approaches. Cross-dataset experiments [56] indicate that these features show better generalization properties than previous approaches, notably when compared to LBP+SVM (*i.e.*, LBP features classified using a SVM).

The IDA features [56] complement the image quality measures (IQM) proposed earlier by Galbally *et al.* [19]. The IQM features are all computed on gray-level images. The IDA features provide a way of additionally capturing information relevant to face-PAD available in the color domain.

Costa-Pazo *et al.* [16] have proposed a face-PAD approach using a set of Gabor features, which characterize the image-texture over the face-region. This work represents the first use of Gabor features for face-PAD. Their experiments show that the Gabor features perform better than the IQM features [19] in detecting PAs. Texture information, captured using shearlets, has also been exploited in the method proposed by Li *et al.* [29].

Certain face-PA cues are not as consistent as others. For example, the set of IDA feature-set includes several features characterizing the amount of specularity in a image. The underlying expectation is that the presence of large specular regions indicates that the input is a PA. There are, however, many instances of PAs that do not include significant specularity. Similarly, although the presence of Moiré patterns is also a strong indicator of PAs [20, 41], the absence of Moiré patterns does not rule of a PA.

Tirunagari *et al.* [54] exploit motion cues to detect face liveness. Specifically, they detect micro-motions, such as slight head movements, lip movements, and eye-blinks, to identify *bona fide* presentations. Unlike the work of Anjos *et al.* [3] – where motion information derived from optical flow computation is directly used to identify PAs – here the video is treated a three-dimensional data, and apply dynamic mode decomposition (DMD) to this 3D data. The result of the DMD procedure is an image where regions of high local micro-motion are marked with brighter pixels. The micro-texture information in the resulting image is characterized using LBP histograms, which are subsequently classified using a SVM.

In the past few years several specific research directions have attracted attention in the context of face-PAD. Unsurprisingly, the application of deep learning methods for face-PAD has become a popular research track. The idea of personalized face-PAD, where client information is incorporated into the PAD process, has also been explored. Several works have been published on the subject of detecting obfuscation attacks. Finally, as the question of detecting previously unseen kinds of PAs becomes important, several researchers have posed face-PAD as an anomaly-detection problem. In the following sections we discuss publications on each of these topics separately.

3.1.1 Deep Learning Approaches To PAD

Following the success of deep learning based approaches for face recognition, there has been a proliferation in CNN based approaches for face-PAD. One reason why researchers are looking into the use of deep networks for face-PAD is that as the quality of PAIs improves, it is becoming increasing difficult to design explicit hand-crafted features able to distinguish PAs from *bona fide* presentation. Here, we highlight a few representative works, to provide readers with a general idea about current research activities on this topic.

In one of the first works in this area, Yang *et al.* [60][4] have proposed a CNN with the same architecture as ImageNet [27], but with the output layer configured for only two outputs: *bona fide* or PA. In this work the authors augment the training data by using input images at multiple scales and also multiple frames of video. The trained CNN is used to extract a feature-vector (from the penultimate fully-connected layer, fc7, of the network) for each input test image. The feature-vector is then classified using a two-class SVM.

More recent works on the use of CNNs for face-PAD have focussed on newer CNN architectures. Lucena *et al.* have proposed FASNet[5] [32], a deep network for face-anti-spoofing. They start with the VGGNet16 (16-layer VGGNet [51]) and modify only the top fully-connected section of the network by removing one fc-layer, and changing the sizes of the subsequent two fc-layers to 256 units and 1 unit, respectively. FASNet shows a small improvement over SpoofNet [35] on the two datasets, 3DMAD and REPLAY-ATTACK, used in both works.

Nagpal and Dubey [37] compare the performances of three different CNN architectures: the Inception-v3 [53] and two versions of ResNet [23], namely ResNet50 (a 50-layer ResNet) and ResNet152 (the 152-layer version). For each architecture, they have conducted six experiments, by training the networks with different parameter-settings. Their study is based on the the MSU-MSFD dataset (see Section 4), which is a relatively small dataset. The authors augment their training data by using flipped versions of each frame in the training-set as well. The best result achieved in this work is an accuracy of 97.52%, produced by the ResNet152 initialized with weights taken from the ImageNet, and where only the final densely connected layers have been re-trained using the MSU-MSFD data. Their experiments also seem to indicate that using lower learning-rates may lead to better discrimination in face-PAD tasks.

Li *et al.* have used a hybrid CNN [28] to model *bona fide* and attack presentations in a parts-based fashion. The face-region is divided into rectangular sub-regions, and a separate two-class CNN (VGG-Face network [40]) is trained for each sub-region. Given a test image, a feature-vector is constructed by concatenating the output vectors from the last fully connected layer of each CNN. This feature-vector is then classified using a SVM.

---

[4] Open source implementation available on `https://github.com/mnikitin/Learn-Convolutional-Neural-Network-for-Face-Anti-Spoofing`

[5] Open-source implementation of FASNet is available on `https://github.com/OeslleLucena/FASNet`

Nguyen *et al.* [38] have explored the idea of combining hand-crafted features with deep learning based features. They train a 19-layer VGGNet [51] (with only two output classes), and take the output of the fc7 layer as a descriptor for the input test image. The descriptors from the CNN are concatenated with a multi-level LBP (MLBP) histogram, a set of hand-crafted features, to construct a combined feature-vector. Principal Component Analysis (PCA) is used as a dimensionality-reduction step to reduce the combined feature-vector to a much shorter feature-vector (reduced from 7828-D to between 90-D and 530-D depending on the dataset). Finally, the reduced feature-vectors are classified using a two-class SVM classifier.

Xu *et al.* [59] combine a long short-term memory (LSTM) network with a CNN to extract features that encode both temporal as well as spatial information. The input to the LSTM-CNN network is a short video, instead of individual frames. The LSTM is plugged on top of the CNN, to model the temporal information in the video. The authors show that this network can outperform straight-forward CNNs, as well as various hand-crafted features.

Liu *et al.* [31] combine a CNN and a LSTM network for face-PAD. In this architecture, the CNN is trained on individual video-frames (images) to extract image-feature-maps as well as depth-maps of the face-region. The LSTM network takes the feature-map produced by the CNN, and is trained to extract a rPPG signal from the video. They present results on the OULU-NPU dataset (see Section 4. A new dataset, named Spoof in the Wild (SiW, discussed in Section 4) is also introduced in this paper.

In general, current datasets for face-PAD are too small to train CNNs from scratch. Most works involving CNNs for face-PAD so far have adapted existing FR CNNs for face-PAD applications, using transfer-learning.

3.1.2 Client-Specific Face-PAD

In real world applications PAD systems are not expected to function in isolation – a PAD system is usually deployed in conjunction with a biometric-verification system. The client-identity information available to the verification system may also be incorporated into the PAD process to improve the PAD performance. This approach to PAD has been explored in various other biometric modalities (such as for fingerprint PAD).

Chingovska and Anjos [14] have proposed client-specific face-PAD methods using both discriminative as well as generative approaches. In both cases, essentially, a separate classifier is constructed for each enrolled client. In the discriminative scheme, for each client, they train a two-class SVM in a one-versus-all configuration. In the generative approach, GMMs are trained for each client using a cohorts-based approach to compensate for the lack of adequate numbers of PAs for each client.

Although the idea of a client-specific approach to face-PAD sounds attractive, one severely limiting factor is the cost of constructing a sufficient variety and number of PAs for every enrolled client. Indeed, the cost may quickly become prohibitive

when PAs based on custom silicone 3D-masks are considered. Yang *et al.* [61] have also proposed a face-PAD method that incorporates client-specific information. Again, they train a separate classifier for each enrolled client. They propose an innovative solution to the problem of lack of sufficient PA samples to train classifiers for newly enrolled clients. Their solution is to use domain-adaptation to generate virtual PA samples to train the client-specific classifiers. The domain-adaptation model learns the relationship between the *bona fide* and attack presentations from the training partition of a dataset. Thereafter, the trained adaptation model is used to generate PA samples for clients in the test partition.

3.1.3 Obfuscation Attacks

An obfuscation attack is said to occur if the attacker actively attempts to alter one's appearance to the extent that FR systems may fail to recognize the subject. Obfuscation attacks may take the form of the use of extreme facial makeup, the use of clothing, or simple medical masks, to occlude significant portions of the face, or even the use of facial masks (mask that resemble faces) made of various materials.

In case of severe occlusion, even localizing the face region in the image (face detection) is a significant challenge. Ge *et al.* [21] have proposed a LLE-CNN – combining CNN based feature-extraction with locally linear embedding (LLE) – to detect the face-region even in the presence of extensive occlusion. For subjects wearing makeup, Wang and Fu [55] have proposed a method for reconstructing makeup-free face images, using local low-rank dictionary learning. Kose *et al.* [26] use a combination of LGBP (LBP histograms computed over a set of Gabor-filtered images) and HOG (histogram of gradients) to classify face-images as containing makeup or not. Agarwal *et al.* [2] tackle the problem of detecting obfuscation using 3D flexible masks, that is, detecting whether the subject in the presentation is wearing a mask, using multispectral imagery. Specifically, they capture images in visible, NIR and thermal wavelength-ranges of the spectrum. Their experiments, based on a variety of local texture descriptors, show that thermal imagery is the best suited for detecting masks reliably. (The use of multispectral data for face-PAD is discussed in more detail in Section 3.2.)

The morphed-image attacks mentioned in Section 2 may be seen as a kind of obfuscation attack. Raghavendra *et al.* [45] have demonstrated the superiority of binarized statistical image features (BSIF) over LBP histograms in detecting morphed-image attacks.

3.1.4 One-Class Classification for PAD

Most researchers approach PAD as a two-class problem. That is, data is collected for both *bona fide* and attack presentations, and, using suitable feature-descriptors, a two-class classifier is trained to discriminate between *bona fide* presentations and attacks. The greatest disadvantage of this general scheme is poor generalization to

unknown attacks. A recent face-PAD competition [11] showed that the performance of all entries deteriorated in the test-protocol involving unknown attacks, relative to their respective performances in test-protocols involving known attacks. Most published face-PAD methods have performed relatively poorly in cross dataset tests (see, for example [19, 56]). The reason is that different datasets include attacks of different kinds (different PAIs, or even just different devices used for performing the attacks). Consequently, the attacks in a given dataset are very likely to be unknown to the classifier that has been trained on a different dataset. This issue – generalization to unknown attacks – has emerged as the most significant challenge in face-PAD.

Indeed, when implementing countermeasures to PAs, the goal is simply to detect PAs, and not necessarily to identify the class of the PA. The problem of PAD may therefore be formulated as one of anomaly detection, where only the *bona fide* class is modelled using a one-class classifier (OCC). In general OOCs may be grouped under two categories: generative and non-generative. A GMM modelling only the *bona fide* class is an example of a generative OCC. A one-class SVM, on the other hand, is a non-generative OCC. Arashloo and Kittler [4] have investigated the use of both kinds of OCCs for the purpose of face-PAD. They report results using a SVM as the non-generative classifier, and a SRC [57] as the generative classifier. The authors compare the performances of two-class GMM and two-class SVM with one-class GMM and one-class SVM respectively, for face-PAD. In total they have considered 20 different scenarios, that is 20 different combinations of classifiers and features. From their experiments, performed with three publicly available datasets, the authors conclude that the OCC based outlier-detection approach can perform comparably to a two-class system. More importantly, the OCC results are better than their two-class counterparts in tests involving unknown PAs (*i.e.*, tests where certain PAs are not represented in the training dataset).

Nikisins *et al.* [39] have also studied the use of OCCs for face-PAD. They base their work on an aggregate dataset composed of three publicly available datasets: REPLAY-ATTACK, REPLAY-MOBILE, and MSU-MFSD (discussed in Section 4). The difference between this work and that of Arashloo and Kittler [4] is that Nikisins *et al.* [39] train their classifiers using the *bona fide* presentations from all three component datasets at once, where as Arashloo and Kittler use *bona fide* presentations of only one dataset at a time in a given experiment. Nikisins *et al.* [39] use a one-class GMM (a generative OCC) to model the distribution of *bona fide* presentations in the aggregated dataset, using a set of image-quality features [19, 56]. Their experiments also show that although two-class classifiers perform better than their one-class counterparts for known attacks (*i.e.*, the case where samples of the attack-types have been included in the training set), their performance deteriorates sharply when presented with unknown attacks, that is PAIs that were not included in the training set. By contrast, the one-class GMM appears to generalize better to unknown classes of PAs [39].

The advantage of using a one-class system is that only data for *bona fide* presentations is necessary. Although experimental test datasets usually include a variety of

attack presentations, in real scenarios it is quite difficult to collect sufficient data for all the various possible kinds of attacks.

3.2 Approaches Based on Extended-Range Imagery

Broadly speaking, visible-light based approaches rely on identifying subtle qualitative differences between *bona fide* and attack presentations. As the quality (color-fidelity, resolution, and so on) of PA devices improves, distinctions between the two kinds of presentations are becoming increasing narrower. That is, progress in PAI quality impacts the performance of existing face-PAD methods. This phenomenon is concretely illustrated by Costa-Pazo *et al.* [16]. They apply the same face-PAD method – SVM classification using a set of image-quality measures – to two datasets. Their experiment shows that the performance of the chosen face-PAD method is significantly worse on the newer dataset (REPLAY-MOBILE [16]) than on the older (REPLAY-ATTACK [13]) dataset. The reason is that as technology (cameras, electronic screens, printers, etc.) improves, the quality of PAs in visible-light is also approaching that of *bona fide* presentations, and therefore it is becoming increasingly difficult to separate the two classes.

A new approach to face-PAD involves the use of ER imagery. Both active- as well as passive-sensing approaches have been considered in recent works. In active ER imagery, the subject is illuminated under a chosen wavelength-band, for example, with NIR and SWIR illumination, and the biometric-sensor (camera) is equipped with appropriate filters, to be able to capture data only in the chosen wavelength band. In passive sensing no specific illumination is used, and the camera is designed to capture radiation in a given wavelength band. One example of passive sensing is the use of thermal cameras to capture the heat radiated by human subjects.

When using active ER imagery for face-PAD, the general idea is to model the reflectance properties of human skin at different wavelengths. Steiner *et al.* [52] have proposed the design of a multi-spectral SWIR camera for face-PAD applications. The camera captures images at four narrow wavelength bands, namely, 935nm, 1060nm, 1300nm, and 1550nm. The image-sensor is sensitive in the range 900-1700nm. The camera is equipped with a ring-illuminator consisting of LEDs emitting NIR in the four wavelength-bands of interest. During image-acquisition the camera cycles through the illumination in the different bands one by one, and synchronizes the image-capture to the duration of illumination at a given wavelength. Thus, the camera captures at multispectral-stack of four images at each time interval. This camera can capture 20 stacks, or frames per second (FPS) – a significant improvement on a previous design of a SWIR camera proposed by Bourlai [12], which was able to capture image at an average rate of 8.3 FPS. Using this camera, human skin can be reliably distinguished from other materials. Steiner *et al.* show results demonstrating the efficacy of face-PAD using data acquired with this camera.

Raghavendra *et al.* [46] have used 7-dimensional multispectral imagery for face-PAD, captured using a SpectroCam^{TM} multispectral camera. This device cap-

tures presentations in narrow bands centered at the following wavelengths: 425nm, 475nm, 525nm, 570nm, 625nm, 680nm and 930nm. The authors propose two face-PAD approaches based on:

- image fusion, where the 7 images in a given multispectral stack are fused into a single image, and a PAD algorithm processes the fused image, and
- score fusion, where the individual images in the multispectral stack are classified separately, and the 7 scores are then fused to generate the final classification score.

Quantitative results [46] show that the score-fusion approach performs significantly better than the image-fusion approach.

Bhattacharjee and Marcel [5] have also investigated the use of ER imagery for face-PAD. They demonstrate that a large class of 2D attacks, namely, video-replay attacks, can be easily detected using NIR imagery. In live presentations under NIR illumination the human face is clearly discernible. However, electronic display monitors appear almost uniformly dark under NIR illumination. Therefore, using NIR imagery, it is possible to design simple statistical measures to distinguish between *bona fide* presentations and attacks. This approach may also be applied to detect print-based attacks. It may fail, however, if the PAIs are printed using metallic inks. The authors also demonstrate that NIR imagery is not particularly useful in detecting 3D mask based attacks. They go on to show that thermal (LWIR) imagery can be used to easily distinguish *bona fide* presentations from mask-based attacks. This is because, in a *bona fide* presentation, the heat emanating from the subject's face renders it very brightly in the thermal image. In contrast, in a mask-attack, the mask appears very dark in the image, because it has a much lower temperature than the subject's body.

This direction of research is still in its infancy. One reason why research in ER imagery has not yet been widely explored is the high cost of IR and thermal cameras. In recent years, however, low-cost options such as the Microsoft Kinect, Intel's RealSense range of sensors, and inexpensive thermal cameras such as from FlirOne and SeekThermal have become widely available. Availability of affordable hardware will be a key factor in advancing research in this direction.

3.3 Detection of 3D Mask Attacks

Good quality 3D masks present clear threats in both impersonation as well as obfuscation categories. As custom 3D masks become increasingly affordable, research on PAD for 3D masks is also gaining critical importance. Bhattacharjee *et al.* [6] have recently demonstrated empirically, that several state-of-the-art FR CNNs are significantly vulnerable to attacks based on custom silicone 3D masks (IAPMR is at least 10 times greater than FNMR).

Initial research was directed towards detecting custom rigid masks, typically made of sandstone powder and resin, with hand-painted facial features. Publicly

available datasets 3DMAD [17] and HKBU-MARs [30] contain data pertaining to custom rigid masks. More recent face-PAD research has focussed on detecting attacks based on hyper-realistic flexible custom masks, usually made of silicone. Although custom silicone masks are still fairly expensive to manufacture, in the coming years the cost of creating such masks is expected to drop to affordable levels.

Another strand of research involving 3D masks is to detect obfuscation attacks mounted using readily available, generic latex masks. Agarwal *et al.* [1] have used texture cues characterized using a set of features computed over co-occurence matrices (so called Haralick-features) to detect rigid-mask attacks in the 3DMAD dataset [17]. Liu *et al.* [30] have published the more recent HKBU-MARs dataset containing images of 3D-mask based PAs. They have proposed a remote photoplethysmography (rPPG) based approach to detecting 3D-mask PAs.

Manjani *et al.* [34] present an observational study into obfuscation attacks using 3D-masks. They describe PAD experiments based on the SMAD dataset (see Section 4), which consists of public-domain videos collected from the World-wide Web. Although observational studies such as this may indicate association between variables (in this case between the true labels of the test videos and the classifier-score), the influence of other confounding variables here cannot be ruled out. To demonstrate the efficacy of a method for detecting 3D-mask based PAs, it is important to design a controlled experiment to highlight exclusively the causal effect of 3D-masks on the resulting classifier-score.

4 New Datasets for Face PAD Experiments

One significant reason for rapid advances in face PAD research is the availability of publicly shared datasets, which facilitates comparison of the performance of new PAD algorithms with existing baseline results. As the quality of devices used to mount attacks improves, the older datasets tend to become less relevant. It is, therefore, important for the research community to continually collect new datasets, representing attacks created using state of the art technology.

Table 1 lists some recently published face-PA datasets. The MSU-MFSD, UVAD, REPLAY-MOBILE, MSU-USSA, OULU-NPU and SiW datasets contain 2D attacks captured under the visible-light illumination. The other datasets include data representing 3D attacks (HKBU-MARs and SMAD) or 2D attacks captured under non-standard illumination, such as extended-range (multispectral) imagery (MS-Face, EMSPAD and MLFP), or light-field imagery (GUC-LiFFAD). Brief descriptions of these datasets follow:

- MSU-MFSD: The public version of the MSU-MFSD dataset [56] includes real-access and attack videos for 35 subjects. Real-access videos ($\tilde{1}2$ sec. long) have been captured using two devices: a 13 MacBook Air (using its built-in camera), and a Google Nexus 5 (Android 4.4.2) phone. Videos captured using the laptop camera have a resolution of 640 $\times$ 480 pixels, and those captured using the Android camera have a resolution of 720 $\times$ 480 pixels. The dataset also includes

Dataset Name	Year	PAIs	Comment
MSU-MFSD [56]	2015	2D attacks: print and replay	70 *bona fide* and 210 PA videos representing 35 subjects, collected using laptop and smartphone.
GUC-LiFFAD [44]	2015	2D attacks: print and replay	Light-field imagery collected from 80 subjects, using a Lytro camera. For each presentation several images are collected, each at a different depth-of-focus.
UVAD [42]	2015	2D attacks: video-replay	17076 videos corresponding to 404 identities.
REPLAY-MOBILE [16]	2016	2D attacks: print, replay	1200 *bona fide* and attack videos representing 40 subjects, captured using only smartphone and tablet.
MSU-USSA [41]	2016	2D attacks: print, replay	1000 *bona fide* presentations and 8000 PAs representing 1000 subjects, captured using only smartphone and tablet.
MS-Face [15]	2016	2D attacks: print (visible and NIR)	Based on 21 subjects print PAIs. Data captured using hi-res. CMOS sensor.
HKBU-MARs [30]	2016	3D attacks: rigid masks	1008 videos corresponding to 12 subjects and their masks.
SMAD [34]	2017	3D attacks: silicone masks	The dataset contains 130 presentations: 65 *bona fide*, and 65 mask attacks.
EMSPAD [47]	2017	2D attacks: print (laser & inkjet)	7-band multispectral data for 50 subjects
OULU-NPU [10]	2017	2D attacks: print, video-replay	5940 videos corresponding to 55 subjects using 6 different smartphones, captured in 3 different environments
MLFP [2]	2017	3D attacks: obfuscation with latex masks	1350 videos based on 10 subjects in visible, NIR and thermal bands, captured in indoor and outdoor environments
SiW [31]	2018	2D attacks: 2 print- and 4 replay-attacks	4620 videos based on 165 subjects, captured in various head-poses and environments. Replay-attacks captured using 4 different PAIs.

Table 1 Recently published datasets for face-PAD experiments.

PA videos representing printed photo attacks, and mobile video replay-attacks where video captured on an iPhone 5s is played back on an iPhone 5s, and high-definition (HD) (1920 $\times$ 1080) video-replays (captured on a Canon 550D SLR, and played back on an iPad Air).

- GUC-LiFFAD: The GUC Light Field Face Artefact Database (GUC-LiFFAD) has been created for face-PAD experiments based on light-field imagery. Specifically, the biometric-sensor used in this dataset is a Lytro[6] camera, which, for every presentation, captures several images, each at a different depth-of-focus. Data corresponding to 80 subjects is included in this dataset. Only print-attacks, based on high-quality photographs (captured using a Canon EOS 550D DSLR

[6] `www.lytro.com`

camera, at 18 mega-pixel resolution, and printed on both laser and inkjet printers) are represented in this dataset.

- UVAD: The Unicamp Visual Attack Database (UVAD) consists of 17076 *bona fide* and attack presentation videos corresponding to 404 identities. All videos have been recorded at full-HD resolution, but subsequently cropped to a size of 1366×768. The dataset includes *bona fide* videos collected using six different cameras. Two videos have been captured for each subject, both using the same camera but under different ambient conditions. PA videos corresponding a given subject have also been captured using the same camera as that used for the *bona fide* videos of the subject in question. The PAs have been generated using seven different electronic monitors, and all PA videos have also been cropped to the same shape as the *bona fide* videos.
- REPLAY-MOBILE: This dataset contains short (10 sec. long) full-HD resolution (720×1280) videos corresponding to 40 identities, recorded using two mobile devices: an iPad Mini 2 tablet and a LG-G4 smartphone. The videos have been collected under six different lighting conditions, involving artificial as well as natural illumination. Four kinds of PAs are represented in this database have been constructed using two PAIs: matte-paper for print-attacks, and matte-screen monitor for digital-replay attacks. For each PAI, two kinds of attacks have been recorded: one where the user holds the recording device in hand, and the second where the recording device is stably supported on a tripod.
- MSU-USSA: The Unconstrained Smartphone Spoof Attack dataset from MSU (MSU-USSA) aggregates *bona fide* presentations from a variety of Internet-accessible sources. In total 1000 *bona fide* presentations of celebrities have been included in this dataset. Two cameras (front and rear camera of a Google Nexus 5 smartphone) have been used to collect 2D attacks using four different PAIs (laptop, tablet, smartphone and printed-photographs), resulting in a total of 8000 PAs.
- HKBU-MARs: This dataset is designed to test countermeasures for 3D rigid-mask based attacks. The second version (V2) of this dataset contains data corresponding to 12 subjects. Rigid masks created by two different manufacturers have been used to construct this dataset. Presentations have been captured using seven different cameras (including mobile devices), under six different illumination conditions.
- MS-Face: This is the first public dataset to explore the use of NIR imagery for face-PAD. Specifically, data is collected under two kinds of illumination: visible-light and 800nm (NIR) wavelengths. The dataset contains data captured from 21 subjects. *Bona fide* presentations in this dataset have been collected under five different conditions. Only print-attacks have been considered in this dataset. For PAs under visible-light, high-quality color prints have been used, whereas PAs under NIR illumination have been created using gray-level images printed at 600 dpi.
- SMAD: the Silicone Mask Attack Database (SMAD) consists of videos collected from the Internet. The authors [34] have collected 65 videos of celebrities (which form the *bona fide* presentations) as well as 65 videos of actors wearing a variety

of flexible masks. Although the authors refer to the masks as silicone masks, some of the masks in the dataset appear to be constructed from latex, instead of silicone. Some of the original videos collected for this dataset may be rather long. For the purposes of experiments, long videos have been trimmed, so that all videos in the dataset are between 3 and 10 sec. long.

- EMSPAD: the Extended Multispectral Presentation Attack Database (EMSPAD) contains images captured using a Pixelteq SpectroCam^{TM} camera. This camera captures multispectral images using a set of filters mounted on a continuously rotating wheel. The dataset contains 7D multispectral stacks per time-instant, that is, for each frame, 7 images have been captured in narrow wavelength bands centered at the following values: 425nm, 475nm, 525nm, 570nm, 625nm, 680nm and 930nm. *Bona fide* and attack presentations for 50 subjects comprise this dataset. *Bona fide* presentations have been collected in two sessions, and in each session, five frames (*i.e.*, 5×7 images) have been collected for each subject. This dataset includes only one kind of PAI, namely, 2D color-print attacks. To construct the attacks, high quality color photographs of each subject have been printed on two kinds of printers – a color laser printer, and a color inkjet printer – at 600dpi resolution, and multispectral images of these printed photographs have been captured using the SpectroCam camera.
- OULU-NPU: This dataset includes data corresponding to 55 subjects. Front cameras of 6 different mobile devices have been used to capture the images included in this dataset. The images have been collected under three separate conditions, each corresponding to a different combination of illumination and background. PAs include print-attacks created using two printers, as well as video-replay attacks using two different displays. In total, 4950 *bona fide* and attack videos comprise the dataset.
- MLFP: The Multispectral Latex Mask based Video Face Prepresentation Attack (MLFP) dataset has been prepared for experiments in detecting obfuscation attacks using flexible latex masks. The dataset consists of 150 *bona fide* and 1200 attack videos, corresponding to 10 subjects. In fact the attacks have been performed using seven latex masks and three paper masks. Data has been collected in both indoor and outdoor environments.
- SiW: The Spoof in the Wild dataset consists of 1320 *bona fide* videos captured from 165 subjects, and 3300 attack videos. Liu *et al.* [31] mention that the dataset encaspulates greater racial diversity than previous datasets. Varying ambient conditions, as well as different facial expressions and head-poses are also represented in the SiW dataset. Two kinds of print attacks and four kinds of video replay-attacks have been included in this dataset. Replay-attacks have been created using four PAIs: two smartphones, a tablet device, and a laptop-monitor screen.

For detailed descriptions of the datasets, such as the experimental protocols as well as how to access the datasets, the reader is referred to the respective references cited in Table 1.

5 Conclusion

As several studies have quantitatively demonstrated, modern face-recognition (FR) methods are highly susceptible to presentation attacks (PA). This vulnerability is a consequence of the desired ability of FR methods to handle inter-session variability. In order to have secure face-verification systems, the underlying FR methods need to be augmented with appropriate presentation attack detection (PAD) methods. Consequently, face-PAD has become a topic of intense research in recent years. In this chapter we have attempted to summarize several prominent research directions in this field.

A large majority of face-PAD methods operate on color-imagery. Several new kinds of features characterizing local motion information, image-quality, as well as texture information have been proposed in the recent scientific literature. Deep-learning based methods for face-PAD have also been widely explored. Most works involving deep-learning methods have started with a CNN designed for FR, and have adapted the network for face-PAD using transfer-learning. The reason for this approach is that current face-PAD datasets are still too small to train really deep networks from scratch. Given this constraint on the size of available training data, perhaps researchers should investigate the use of relatively smaller networks for face-PAD.

In addition to well studied categories of 2D attacks, namely, print attacks and video-replay attacks, several research groups are now developing methods to detect attacks performed using hyper-realistic custom-made masks. Attacks based on both rigid and flexible masks have been considered. In the past this category of attacks did not receive much attention as constructing custom-masks was prohibitively expensive. Although, even today the cost of manufacturing high-quality custom masks remains high, the costs have come down significantly, and we may expect PAs based on such masks to be highly likely in the near future. The research community would benefit from a concerted effort to produce large and significantly diverse datasets based on a variety of custom-made masks.

Extended-range (ER) imagery, that is, imagery in wavelengths outside the visible light spectrum, is proving to be a valuable tool in tackling both 2D and 3D PAs. Given the availability of low-cost infrared and thermal cameras, this is a promising direction of research in face-PAD.

Besides impersonation attacks, the recently adopted ISO standard for PAD also considers obfuscation attacks as PAs. Specifically, there is a need to detect presentations where makeup or a mask is used to hide one's identity. This category of PA has not received the same amount of attention as impersonation attacks. The availability of carefully constructed datasets representing obfuscation attacks is key to the progress of research on this topic.

We note, in general, that most recent papers on face-PAD still report results on relatively old datasets, such as CASIA and REPLAY-ATTACK – datasets that are more than five years old now. With ever-improving technology for constructing PAs, older datasets become increasingly irrelevant. In order to have the true snapshot

of the state of the art, besides publishing new datasets at a steady rate, it is also important that face-PAD researchers report results on recent datasets.

Although most state-of-the-art face-PAD methods seem to perform well in intra-dataset tests, generalization in cross-dataset scenarios remains a significant challenge. Cross-dataset generalization is an important goal, because it indicates the ability of a given PAD method to tackle previously unseen attacks. In this context the use of one-class classifiers (OCC) have been shown to be a step in the right direction.

There is a growing interest in developing face-PAD methods for scenarios involving previously unseen attacks. We expect this trend to grow in the coming years. Another research direction with great potential is the use of ER imagery to tackle various kinds of PAs. So far, deep learning based methods for face-PAD have been shown to be roughly as accurate as state-of-the-art methods relying on hand-crafted features. As mentioned earlier, current efforts involving deep learning start with well understood deep networks designed for object recognition or FR. Further research is required in this area, perhaps involving bespoke deep architectures for face-PAD.

Acknowledgements This work has been supported by the European H2020-ICT project TeSLA[7] (grant agreement no. 688520), the project on Secure Access Control over Wide Area Networks (SWAN) funded by the Research Council of Norway (grant no. IKTPLUSS 248030/O70), and by the Swiss Center for Biometrics Research and Testing.

References

1. Agarwal, A., Singh, R., Vatsa, M.: Face Anti-Spoofng Using Haralick Features. In: Proceedings of the IEEE International conference on Biometrics: Theory, Applications, and Systems (BTAS), pp. 1 – 6. Niagara Falls, NY, USA (2016)
2. Agarwal, A., Yadav, D., Kohli, N., Singh, R., Vatsa, M., Noore, A.: Face Presentation Attack with Latex Masks in Multispectral Videos. In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), pp. 275 – 283 (2017). DOI 10.1109/CVPRW.2017.40
3. Anjos, A., Chakka, M.M., Marcel, S.: Motion-Based Counter-Measures to Photo Attacks in Face Recognition. IET Biometrics **3**(3), 147 – 158 (2014). DOI 10.1049/iet-bmt.2012.0071
4. Arashloo, S.R., Kittler, J.: An Anomaly Detection Approach to Face Spoofing Detection: A New Formulation and Evaluation Protocol. In: Proceedings of the IEEE International Joint Conference on Biometrics (IJCB), pp. 80 – 89 (2017)
5. Bhattacharjee, S., Marcel, S.: What You Can't See Can Help You – Extended Range Imaging for 3d-Mask Presentation Attack Detection. In: Proceedings of the 16th International Conference of the Biometrics Special Interest Group (BIOSIG). Darmstadt, Germany (2017)
6. Bhattacharjee, S., Mohammadi, A., Marcel, S.: Spoofing Deep Face Recognition With Custom Silicone Masks. In: Proceedings of the IEEE International Conference on Biometrics: Theory, Applications and Systems (BTAS). Los Angeles, USA (2018)
7. Boulkenafet, Z., Komulainen, J., Hadid, A.: Face Anti-Spoofing Based on Color Texture Analysis. In: IEEE International Conference on Image Processing (ICIP), pp. 2636 – 2640 (2015)

[7] www.tesla-project.eu

8. Boulkenafet, Z., Komulainen, J., Hadid, A.: Face Spoofing Detection Using Colour Texture Analysis. IEEE Transactions on Information Forensics and Security **11**(8), 1818 – 1830 (2016). DOI 10.1109/TIFS.2016.2555286
9. Boulkenafet, Z., Komulainen, J., Hadid, A.: On the generalization of color texture-based face anti-spoofing. Image and Vision Computing (2018). Accepted at the time of writing.
10. Boulkenafet, Z., Komulainen, J., Li, L., Feng, X., Hadid, A.: OULU-NPU: A Mobile Face Presentation Attack Database with Real-World Variations. In: Proceedings of 12th IEEE International Conference on Automatic Face and Gesture Recognition (FG 2017) (2017)
11. Boulkenafet, Z., et al.: A Competition on Generalized Software-Based Face Presentation Attack Detection in Mobile Scenarios. In: Proceedings of IEEE International Joint Conference on Biometrics (IJCB), pp. 688 – 696 (2017). DOI 10.1109/BTAS.2017.8272758
12. Bourlai, T., Narang, N., Cukic, B., Hornak, L.: On Designing a SWIR Multi-Wavelength Facial-Based Acquisition System. In: Proceedings of SPIE: Infrared Technology and Applications, vol. 8353 (2012)
13. Chingovska, I., Anjos, A., Marcel, S.: On The Effectiveness of Local Binary Patterns in Face Anti-Spoofing. In: Proceedings of the International Conference of Biometrics Special Interest Group (BIOSIG) (2012)
14. Chingovska, I., dos Anjos, A.R.: On the Use of Client Identity Information for Face Antispoofing. IEEE Transactions on Information Forensics and Security **10**(4), 787 – 796 (2015). DOI 10.1109/TIFS.2015.2400392
15. Chingovska, I., Erdogmus, N., Anjos, A., Marcel, S.: Face Recognition Systems Under Spoofing Attacks. In: T. Bourlai (ed.) Face Recognition Across the Imaging Spectrum, pp. 165 – 194. Springer (2016)
16. Costa-Pazo, A., Bhattacharjee, S., Vazquez-Fernandez, E., Marcel, S.: The Replay-Mobile Face Presentation-Attack Database. In: Proceedings of International Conference of the Biometrics Special Interest Group (BIOSIG) (2016). DOI 10.1109/BIOSIG.2016.7736936
17. Erdogmus, N., Marcel, S.: Spoofing in 2D Face Recognition With 3D Masks and Anti-Spoofing With Kinect. In: Proceedings of the IEEE International Conference on Biometrics: Theory, Applications and Systems (BTAS) (2013)
18. Ferrara, M., Franco, A., Maltoni, D.: The Magic Passport. In: Proceedings of IEEE International Joint Conference on Biometrics (IJCB) (2014). DOI 10.1109/BTAS.2014.6996240
19. Galbally, J. and Marcel, S. and Fierrez, J.: Image Quality Assessment for Fake Biometric Detection: Application to Iris, Fingerprint, and Face Recognition. IEEE Transactions on Image Processing **23**(2), 710 – 724 (2014). DOI 10.1109/TIP.2013.2292332
20. Garcia, D.C., de Queiroz, R.L.: Face-Spoofing 2D-Detection Based on Moiré-Pattern Analysis. IEEE Transactions on Information Forensics and Security **10**(4), 778 – 786 (2015). DOI 10.1109/TIFS.2015.2411394
21. Ge, S., Li, J., Ye, Q., Luo, Z.: Detecting Masked Faces in the Wild with LLE-CNNs. In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 426 – 434 (2017). DOI 10.1109/CVPR.2017.53
22. Hadid, A.: Face Biometrics Under Spoofing Attacks: Vulnerabilities, Countermeasures, Open Issues, and Research Directions. In: Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), pp. 113 – 118 (2014)
23. He, K., Zhang, X., Ren, S., Sun, J.: Deep Residual Learning for Image Recognition. In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 770 – 778 (2016). DOI 10.1109/CVPR.2016.90
24. Huang, G.B., Ramesh, M., Berg, T., Learned-Miller, E.: Labeled Faces in the Wild: A Database for Studying Face Recognition in Unconstrained Environments. Technical report 07-49, University of Massachusetts, Amherst (MA), USA (2007)
25. Kanzawa, Y., Kimura, Y., Naito, T.: Human Skin Detection by Visible and Near-Infrared Imaging. In: Proceedings of the 12th IAPR Conference on Machine Vision Applications, MVA 2011. Nara, Japan (2011)
26. Kose, N., Apvrille, L., Dugelay, J.L.: Facial makeup detection technique based on texture and shape analysis. In: Proceedings of 11th IEEE International Conference on Automatic Face

and Gesture Recognition, May 4-8, 2015, Ljubljana, Slovenia (FG). Ljubljana, SLOVENIA (2015). DOI http://dx.doi.org/10.1109/FG.2015.7163104. URL http://www.eurecom.fr/publication/4494

27. Krizhevsky, A., Sutskever, I., Hinton, G.E.: ImageNet Classification with Deep Convolutional Neural Networks. In: Advances in Neural Information Processing Systems, vol. 25 (2012)
28. Li, L., Xia, Z., Li, L., Jiang, X., Feng, X., Roli, F.: Face Anti-Spoofing Via Hybrid Convolutional Neural Network. In: Proceedings of International Conference on the Frontiers and Advances in Data Science (FADS), pp. 120 – 124 (2017). DOI 10.1109/FADS.2017.8253209
29. Li, Y., Po, L.M., Xu, X., Feng, L., Yuan, F.: Face Liveness Detection and Recognition Using Shearlet Based Feature Descriptors. In: Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 874 – 877 (2016)
30. Liu, S., Yang, B., Yuen, P.C., Zhao, G.: A 3D Mask Face Anti-Spoofing Database With Real World Variations. In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), pp. 1551 – 1557 (2016). DOI 10.1109/CVPRW.2016.193
31. Liu, Y., Jourabloo, A., Liu, X.: Learning Deep Models for Face Anti-Spoofing: Binary or Auxiliary Supervision. In: In Proceeding of IEEE Computer Vision and Pattern Recognition. Salt Lake City, USA (2018)
32. Lucena, O., Junior, A., Hugo G. M., V., Souza, R., Valle, E., De Alencar Lotufo, R.: Transfer Learning Using Convolutional Neural Networks for Face Anti-Spoofing. In: F. Karray, A. Campilho, F. Cheriet (eds.) Proceedings of International Conference on Image Analysis and Recognition (ICIAR), pp. 27 – 34. Springer International Publishing, Cham (2017)
33. Määttä, J., Hadid, A., Pietikäinen, M.: Face Spoofing Detection From Single Images Using Micro-Texture Analysis. In: Proceedings of International Joint Conference on Biometrics (IJCB) (2011). DOI 10.1109/IJCB.2011.6117510
34. 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). DOI 10.1109/TIFS.2017.2676720
35. Menotti, D., Chiachia, G., Pinto, A., Schwartz, W.R., Pedrini H. Falco, 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). DOI 10.1109/TIFS.2015.2398817
36. Mohammadi, A., Bhattacharjee, S., Marcel, S.: Deeply Vulnerable: A Study of the Robustness of Face Recognition to Presentation Attacks. IET Biometrics **7**(1), 15 – 26 (2018). DOI 10.1049/iet-bmt.2017.0079
37. Nagpal, C., Dubey, S.R.: A Performance Evaluation of Convolutional Neural Networks for Face Anti Spoofing. CoRR **abs/1805.04176** (2018). URL https://arxiv.org/abs/1805.04176
38. Nguyen, T.D., Pham, T.D., Baek, N.R., Park, K.R.: Combining Deep and Handcrafted Image Features for Presentation Attack Detection in Face Recognition Systems Using Visible-Light Camera Sensors. Journal of Sensors **18**(3), 699 – 727 (2018). URL https://doi.org/10.3390/s18030699
39. Nikisins, O., Mohammadi, A., Anjos, A., Marcel, S.: On Effectiveness of Anomaly Detection Approaches against Unseen Presentation Attacks in Face Anti-Spoofing. In: Proceedings of International Conference on Biometrics (ICB) (2018). DOI 10.1109/ICB2018.2018.00022
40. Parkhi, O.M., Vedaldi, A., Zisserman, A.: Deep Face Recognition. In: British Machine Vision Conference (2015)
41. Patel, K., Han, H., Jain, A.: Secure Face Unlock: Spoof Detection on Smartphones. IEEE Transactions on Information Forensics and Security **11**(10), 2268 – 2283 (2016). DOI 10.1109/TIFS.2016.2578288
42. Pinto, A., Schwartz, W.R., Pedrini, H., Rocha, A.D.R.: Using Visual Rhythms for Detecting Video-Based Facial Spoof Attacks. IEEE Transactions on Information Forensics and Security **10**(5), 1025 – 1038 (2015). DOI 10.1109/TIFS.2015.2395139
43. Ramachandra, R., Büsch, C.: Presentation Attack Detection Methods for Face Recognition Systems - A Comprehensive Survey. ACM Computing Surveys **50** (2017)
44. Ramachandra, R., Raja, K., Büsch, C.: Presentation Attack Detection for Face Recognition using Light Field Camera. IEEE Transactions on Image Processing **24**(3), 1 – 16 (2015)

45. Ramachandra, R., Raja, K.B., Büsch, C.: Detecting Morphed Face Images. In: Proceedings of IEEE 8th International Conference on Biometrics Theory, Applications and Systems (BTAS), pp. 1 – 7 (2016). DOI 10.1109/BTAS.2016.7791169
46. Ramachandra, R., Raja, K.B., Venkatesh, S., Büsch, C.: Extended Multispectral Face Presentation Attack Detection: An Approach Based on Fusing Information From Individual Spectral Bands. In: Proceedings of 20th International Conference on Information Fusion (Fusion) (2017). DOI 10.23919/ICIF.2017.8009749
47. Ramachandra, R., Raja, K.B., Venkatesh, S., Cheikh, F.A., Büsch, C.: On the Vulnerability of Extended Multispectral Face Recognition Systems Towards Presentation Attacks. In: Proceedings of IEEE International Conference on Identity, Security and Behavior Analysis (ISBA), pp. 1 – 8 (2017). DOI 10.1109/ISBA.2017.7947698
48. Ratha, N.K., Connell, J.H., Bolle, R.M.: An Analysis of Minutiae Matching Strength. In: J. Bigun, F. Smeraldi (eds.) Audio- and Video-Based Biometric Person Authentication, pp. 223 – 228. Springer, Berlin, Heidelberg (2001)
49. Scherhag, U., Nautsch, A., Rathgeb, C., Gomez-Barrero, M., Veldhuis, R.N.J., Spreeuwers, L., Schils, M., Maltoni, D., Grother, F., Marcel, S., Breithaupt, R., Ramachandra, R., Büsch, C.: Biometric Systems under Morphing Attacks: Assessment of Morphing Techniques and Vulnerability Reporting. In: Proceedings of International Conference of the Biometrics Special Interest Group (BIOSIG) (2017). DOI 10.23919/BIOSIG.2017.8053499
50. Schroff, F., Kalenichenko, D., Philbin, J.: FaceNet: A Unified Embedding for Face Recognition and Clustering. CoRR **abs/1503.03832** (2015). URL `http://arxiv.org/abs/1503.03832`
51. Simonyan, K., Zisserman, A.: Very Deep Convolutional Networks for Large-Scale Image Recognition. CoRR **abs/1409.1556** (2014). URL `http://arxiv.org/abs/1409.1556`
52. Steiner, H., Sporrer, S., Kolb, A., Jung, N.: Design of an Active Multispectral SWIR Camera System for Skin Detection and Face Verification. Journal of Sensors **2016**(1), 1 – 8 (2016). Article ID 9682453, Special Issue on Multispectral, Hyperspectral, and Polarimetric Imaging Technology
53. Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., Wojna, Z.: Rethinking the Inception Architecture for Computer Vision. In: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 2818 – 2826 (2016). DOI 10.1109/CVPR.2016.308
54. Tirunagari, S., Poh, N., Windridge, D., Iorliam, A., Suki, N., Ho, A.T.S.: Detection of Face Spoofing Using Visual Dynamics. IEEE Transactions on Information Forensics and Security **10**(4), 762 – 777 (2015). DOI 10.1109/TIFS.2015.2406533
55. Wang, S., Yun Fu, Y.: Face Behind Makeup. In: Proceedings of the Thirtieth Conference of the Association for the Advancement of Artificial Intelligence, pp. 58 – 64 (2016)
56. Wen, D., Han, H., Jain, A.K.: Face Spoof Detection with Image Distortion Analysis. IEEE Transactions on Information Forensics and Security **10**(4), 746 – 761 (2015)
57. Wright, J., Yang, A.Y., Ganesh, A., Sastry, S.S., Ma, Y.: Robust Face Recognition via Sparse Representation. IEEE Transactions on Pattern Analysis and Machine Intelligence **31**(2), 210 – 227 (2009). DOI 10.1109/TPAMI.2008.79
58. Wu, X., He, R., Sun, Z.: A Lightened CNN for Deep Face Representation. CoRR **abs/1511.02683** (2015). URL `http://arxiv.org/abs/1511.02683`
59. Xu, Z., Li, S., Deng, W.: Learning Temporal Features Using LSTM-CNN Architecture for Face Anti-Spoofing. In: Proceedings of 3rd IAPR Asian Conference on Pattern Recognition (ACPR), pp. 141 – 145 (2015). DOI 10.1109/ACPR.2015.7486482
60. Yang, J., Lei, Z., Li, S.Z.: Learn Convolutional Neural Network for Face Anti-Spoofing. CoRR **abs/1408.5601** (2014). URL `http://arxiv.org/abs/1408.5601`
61. Yang, J. and Lei, Z. and Yi, D. and Li, S. Z.: Person-Specific Face Antispoofing With Subject Domain Adaptation. IEEE Transactions on Information Forensics and Security **10**(4), 797–809 (2015). DOI 10.1109/TIFS.2015.2403306