

# Fingerprint Presentation Attack Detection: A Sensor and Material Agnostic Approach

Steven A. Grosz  
Michigan State University  
East Lansing, MI, 48824  
groszste@msu.edu

Tarang Chugh  
Michigan State University  
East Lansing, MI, 48824  
chughtar@msu.edu

Anil K. Jain  
Michigan State University  
East Lansing, MI, 48824  
jain@msu.edu

## Abstract

The vulnerability of automated fingerprint recognition systems to presentation attacks (PA), i.e., spoof or altered fingers, has been a growing concern, warranting the development of accurate and efficient presentation attack detection (PAD) methods. However, one major limitation of the existing PAD solutions is their poor generalization to new PA materials and fingerprint sensors, not used in training. In this study, we propose a robust PAD solution with improved cross-material and cross-sensor generalization. Specifically, we build on top of any CNN-based architecture trained for fingerprint spoof detection combined with cross-material spoof generalization using a style transfer network wrapper. We also incorporate adversarial representation learning (ARL) in deep neural networks (DNN) to learn sensor and material invariant representations for PAD. Experimental results on LivDet 2015 and 2017 public domain datasets exhibit the effectiveness of the proposed approach.

## 1. Introduction

Fingerprints are considered one of the most reliable biometric traits due to their inherent uniqueness and persistence, which has led to their widespread adoption in secure authentication systems [27]. However, it has been demonstrated that these systems are vulnerable to presentation attacks by adversaries trying to gain access to the system [14, 37]. A presentation attack (PA) as defined by the ISO standard IEC 30107-1:2016(E) [24] is a “presentation to the biometric data capture subsystem with the goal of interfering with the operation of the biometric system.” These attacks often involve a fingerprint cast from a mold using common household materials (gelatin, silicone, wood glue, etc) and aim to mimic the ridge-valley structure of an enrolled user’s fingerprint [33, 3, 11, 48, 28].

The vulnerability of these systems to presentation attacks

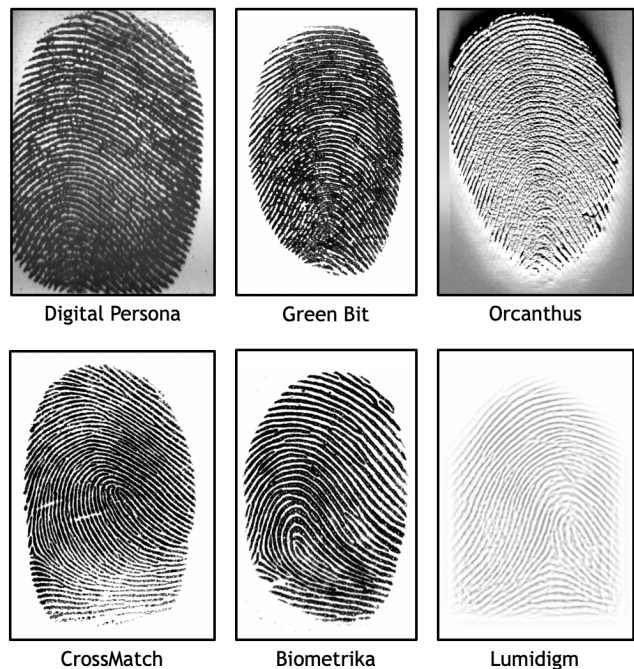


Figure 1: Illustration of the differences in textural appearance of live fingerprints captured on six different fingerprint readers. Images from LivDet 2015 [34], LivDet 2017 [35], and MSU-FPAD datasets [4].

led to a series of standard assessments of fingerprint presentation attack detection (PAD) methods developed by different organizations<sup>1</sup>. The First International Fingerprint Liveness Detection Competition debuted in 2009 [31] with subsequent competitions every two years, the most recent being 2019 [47, 19, 34, 35, 38].

There are numerous published approaches to liveness detection, which can be classified as hardware-based, software-based or a combination of both. Hardware based methods use a number of additional sensors to gain further insight into the liveness of the presented fingerprint [1, 25, 12]. Similarly, a few sensing technologies are inherently

<sup>1</sup>In the literature, presentation attack detection (PAD) is also commonly referred to as spoof detection and liveness detection. In this work, we use these terms interchangeably.

Table 1: Summary of Published Fingerprint Cross-Material Generalization Studies.

Study	Approach	Database	Performance
Rattani et al. [40]	Weibull-calibrated SVM	LivDet 2011	EER = 19.70 %
Ding & Ross [9]	Ensemble of multiple one-class SVMs	LivDet 2011	EER = 17.06 %
Chugh & Jain [4]	MobileNet-v1 trained on minutiae-centered local patches	LivDet 2011-2015	ACE = 1.48 % (LivDet 2015), 2.93 % (LivDet 2011, 2013)
Chugh & Jain [5]	Identify a representative set of spoof materials to cover the deep feature space	MSU-FPAD v2.0, 12 spoof materials	TDR = 75.24 % @ FDR = 0.2 %
Engelsma & Jain [13]	Ensemble of generative adversarial networks (GANs)	Custom database with live and 12 spoof materials	TDR = 49.80 % @ FDR = 0.2 %
Gonzalez-Soler et al. [20]	Feature encoding of dense-SIFT features	LivDet 2011-2015	TDR = 7.03 % @ FDR = 1.0 % (LivDet 2015), ACE = 1.01 % (LivDet 2011, 2013)
Tolosana et al. [43]	Fusion of two CNN architectures trained on SWIR images	Custom database with live and 8 spoof materials	EER = 1.35 %
Gajawada et al. [15]	Style transfer from spoof to live images with a few samples of target material	LivDet 2015, CrossMatch sensor	TDR = 78.04 % @ FDR = 0.1 %
Chugh & Jain [6]	Style transfer between known spoof materials to improve generalizability against unknown materials	MSU-FPAD v2.0, 12 spoof materials & LivDet 2017	TDR = 91.78 % @ FDR = 0.2 % (MSU-FPAD v2.0); Avg. Accuracy = 95.88 % (LivDet 2017)
<b>Proposed Approach</b>	Style transfer with a few samples of target sensor fingerprint images + ARL	LivDet 2015	TDR = 87.86 % @ FDR = 0.2 % cross-sensor & cross-material

well suited for liveness detection and have been used for fingerprint PAD, such as the multispectral Lumidigm sensor or OCT based sensors [7]. On the other hand, software-based solutions use only the information contained in the captured fingerprint image (or a sequence of images) to classify a fingerprint as bonafide or PA [32, 30, 18, 17, 36, 39, 4]. Of the existing software solutions, convolutional neural network (CNN) approaches have shown the best performance on the respective genuine vs. PA benchmark datasets. However, it has been shown that the spoof detection error rates of these approaches suffer up to a three fold increase when applied to datasets containing spoof materials not seen during training, denoted as *cross-material generalization* [29, 42].

Some published studies aimed at reducing the performance gap due to cross material evaluations are summarized in Table 1. A similar performance gap exists for *cross-sensor generalization*, in which presentation attack algorithms are applied to fingerprint images captured on new fingerprint sensor devices that were not seen during training. One explanation for the challenge of cross-sensor generalization is the different textural characteristics in the fingerprint images from different sensors (Figure 1). This discrepancy in the representation performance between the *seen source domain* and the *unseen target domain* has been referred to as the “domain gap” in the growing literature of deep neural network models applied for representational

learning [2]. The cross-sensor evaluation can be considered as two separate cases: (i) all sensors in the evaluation employ the same sensing technology, e.g., all optical FTIR, and (ii) the sensors may vary in the underlying sensing mechanisms used, e.g., optical direct-view vs. capacitive.

In this work, we aim to improve the fingerprint presentation attack detection generalization across novel spoof materials and fingerprint sensing devices<sup>2</sup>. Our approach builds off any existing CNN-based architecture trained for fingerprint liveness detection combined with cross material spoof generalization using a style transfer network wrapper. We also incorporate adversarial representation learning (ARL) in deep neural networks (DNN) to learn sensor and material invariant representations for presentation attack detection.

The main contributions of this study are enumerated below:

1. A robust PAD solution with improved cross-material and cross-sensor generalization performance.
2. Our solution can be built on top of any CNN-based fingerprint PAD solution for cross-sensor and cross-

<sup>2</sup>Generally, fingerprint sensor refers to the fingerprint sensing mechanism (e.g., camera and prism for FTIR optical, direct-view camera, thermal measurement device, etc.) and fingerprint reader refers to the entire process of converting a physical fingerprint into a digital image. In this work, similar to the literature, we use these two terms interchangeably.

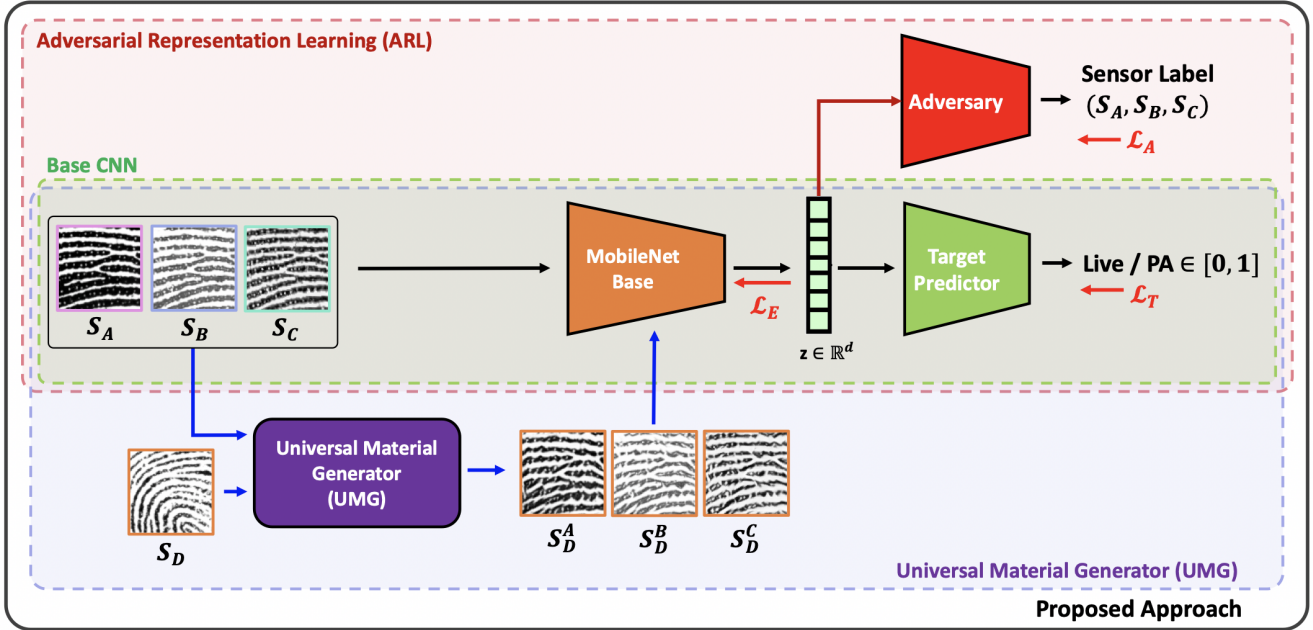


Figure 2: Overview of the network architecture for the proposed UMG + ARL approach for live vs. presentation attack (PA) detection.  $S_A$ ,  $S_B$ ,  $S_C$ , and  $S_D$  represent fingerprint images from four different fingerprint sensors.  $\mathcal{L}_T$  denotes a cross-entropy loss on the target prediction,  $\mathcal{L}_A$  denotes a cross-entropy loss on the sensor label prediction, and  $\mathcal{L}_E$  denotes the loss propagated to the encoder.

material spoof generalization using adversarial representational learning.

3. Experimental evaluation of the proposed approach on publicly available datasets LivDet 2015, LivDet 2017, and MSU-FPAD. Our approach is shown to improve the cross-sensor (cross-material) generalization performance from a TDR of 88.36% (78.76%) to a TDR of 92.94% (87.86%) at a FDR of 0.2%.
4. Feature space analysis of cross-sensor domain separation of the embedded representations prior to and following adversarial representation learning.
5. Detailed discussion of the challenges and techniques involved in applying deep-adversarial representation learning for spoof detection.

## 2. Related Work

In this section we briefly discuss the preliminaries of domain adaptation and domain generalization in the context of machine learning. Csurka provides a more in-depth review of domain adaptation [8]. Similarly, Wang and Deng provide a specific survey of the recent deep domain adaptation methods [45]. We also describe adversarial representation learning (ARL) as it is applied to the tasks of domain adaptation and domain generalization.

### 2.1. Domain Adaptation and Domain Generalization

A domain refers to a probability distribution over which data examples are drawn from. In this context, domain adaptation and domain generalization are approaches to machine learning aimed at minimizing the performance gap between training data examples from a seen “source” domain and testing data from a related, but different “target” domain. Therefore, domain adaptation and domain generalization are applied to situations in which the training and testing data points are not both independently and identically sampled from the same distribution. While domain adaptation involves training on labeled examples from the source domain and unlabeled data from the target domain, domain generalization assumes no access to labeled or unlabeled data examples from the target domain.

### 2.2. Adversarial Representation Learning (ARL)

Adversarial representation learning is a machine learning technique that can be applied to both domain adaptation and domain generalization. Adversarial representation learning has been applied in DNN architectures to extract discriminative representations for a given target prediction task (say face recognition), while obfuscating some undesired attributes present in the data (say gender information) [10, 16, 44, 49].

The general setup of ARL involves (i) an encoder network, (ii) a target prediction network, and (iii) an adversary network. The encoder network aims to extract a latent rep-

resentation ( $\mathbf{z}$ ) that is not only informative for the target prediction task ( $t$ ), but also does not leak any information for the sensitive task ( $s$ ). Meanwhile, the adversary network is tasked with extracting the sensitive information from the encoded latent representation. The entire network is trained in a minimax game similar to the generative adversarial networks introduced by Goodfellow et al. [21].

In Xie et al., the parameters of the adversary network are optimized to maximize the likelihood of the sensitive label prediction, whereas the encoder is trained to maximize the likelihood of the target task, while minimizing the likelihood of the sensitive task [46]. In contrast, our proposed work is more aligned with the approach proposed by Roy and Bodetti [41], where the adversary network is optimized to maximize the likelihood of the sensitive label prediction from the latent representation and the encoder is trained to maximize the entropy of the sensitive label prediction. In this manner, the base network is encouraged to encode a representation that aims to confuse the sensitive label prediction such that the adversary predicts equal probabilities (maximum entropy) for all classes of the sensitive label.

### 3. Proposed Approach

Our proposed approach is multifaceted and combines ideas from style transfer, which was previously applied for spoof detection, and adversarial representation learning to improve the generalization performance of PAD across different fingerprint sensing devices. An overview of the approach which highlights each of the individual components is shown in Figure 2. Here we introduce each individual component and later discuss the generalization performance improvement achieved with the incorporation of each technique leading up to the final approach.

#### 3.1. Base CNN

What we refer to as the *base CNN* approach is a convolutional neural network (CNN) trained on 96 x 96 aligned minutiae-centered patches for classifying a given fingerprint impression as live or spoof. As was shown by Chugh and Jain [4], utilizing minutiae patches, as opposed to whole images, overcomes the difficulty in processing fingerprint images of different sizes, provides large amounts of training examples suitable to training deep CNN architectures, and encourages the network to learn local textural cues to robustly separate bonafide from fake fingerprints. This base CNN approach is illustrated in Figure 2 as the box enclosed by the green line.

The specific architecture of the CNN model employed is the MobileNet-v1 model [23] (the same as in [4])<sup>3</sup>, where the final 1000-unit softmax layer is replaced with a 2-unit softmax layer suitable for the two-class problem of live vs.

<sup>3</sup>Any other CNN-based approach other than [4] can be used instead.

spoof. The network is trained from scratch with an RMSPprop optimizer at a batch size of 64. During training, data augmentation tools of random distorted cropping, horizontal flipping and random brightness were employed to encourage robustness to overfitting to minute variations of the input images.

#### 3.2. Adversarial Representational Learning (ARL)

ARL is an approach to domain generalization that does not require any knowledge of the unseen target domain, yet aims to learn a generalized and robust feature representation for both source and target domains. The goal of the ARL approach is to encourage an encoding network to learn a representation that is invariant to which sensor generated the input fingerprint images (sensitive label), while accurately predicting live vs. PA (target label).

In this setup, the encoder network is represented as a deterministic function,  $\mathbf{z} = E(\mathbf{x}; \theta_E)$ , the target prediction network estimates the conditional distribution  $p(t|\mathbf{x})$  through  $q_T(t|\mathbf{z}; \theta_T)$ , and the adversary network estimates the conditional distribution  $p(s|\mathbf{x})$  through  $q_A(s|\mathbf{z}; \theta_A)$ ; where  $\mathbf{x}$  denotes the input fingerprint image,  $p(t|\mathbf{x})$  and  $p(s|\mathbf{x})$  represent the probabilities of ground truth target and sensitive labels  $t$  and  $s$ , respectively.

To learn this sensor-invariant representation, the adversary network is trained to maximize the likelihood of predicting which sensor generated the input fingerprint image from the encoded representation. The parameters,  $\theta_A$ , of the adversary network are updated to minimize the loss defined in equation 1. The output of the adversary network is used to encourage the encoder to produce a representation that obfuscates the sensitive class labels by penalizing the parameters of the encoder,  $\theta_E$ , to minimize the loss in equation 2, where  $\alpha$  is a hyper-parameter that allows for a trade-off between obfuscation of the sensitive label and prediction of the target label. Meanwhile, to accurately predict live vs. PA, the parameters of target prediction network,  $\theta_T$ , are optimized to minimize the loss in equation 3. The ARL approach is shown in Figure 2 by the box enclosed by the red line.

$$\mathcal{L}_A = \mathbb{E}_{\mathbf{x}, s}[-\log q_A(s|E(\mathbf{x}; \theta_E); \theta_A)] \quad (1)$$

$$\begin{aligned} \mathcal{L}_E = \mathbb{E}_{\mathbf{x}, t}[-\log q_T(t|E(\mathbf{x}; \theta_E); \theta_T)] \\ + \alpha \mathbb{E}_{\mathbf{x}} \left[ \sum_{i=1}^m q_A(s_i|E(\mathbf{x}; \theta_E); \theta_A) \log q_A(s_i|E(\mathbf{x}; \theta_E); \theta_A) \right] \end{aligned} \quad (2)$$

$$\mathcal{L}_T = \mathbb{E}_{\mathbf{x}, t}[-\log q_T(t|E(\mathbf{x}; \theta_E); \theta_T)] \quad (3)$$

### 3.3. Naïve

A simple approach to cross-sensor generalization is one in which we assume access to a limited number of training examples (100 live and PA fingerprint images) from the target sensor that we include during training, which doesn't require collecting extensive amounts of data from the target domain. This is a reasonable assumption in the case of cross-sensor generalization, where we have access to the sensing device on which the system will be deployed. This is in contrast to generalization to unknown spoof materials, where we cannot assume any prior knowledge of the unknown target materials. We denote this method as the *naïve* approach to cross-sensor spoof detection as it does not require any changes to the system architecture.

### 3.4. Naïve + ARL

We combine the naïve approach with ARL to take advantage of the benefits of each separate approach. By exposing the network to the textural characteristics inherent to the small number of target sensor images during training, the goal is that the network will better learn a mapping from images to representations for each sensor domain. Furthermore, by incorporating the adversary during training to learn a sensor-invariant representation, we aim to overcome the apparent imbalance in the number of training examples from source and target sensors.

### 3.5. Universal Material Generator (UMG)

The final technique that we incorporate is a style transfer approach, coupled on top of the naïve approach, to augment the training data from the target sensor. The specific style transfer network we use is the Universal Material Generator (UMG) proposed in [6] that inputs source and target domain minutiae patches and produces a large amount of synthetic training images in the target sensor domain. UMG achieves this by learning a mapping from the style of the source domain image patches to the style of the target domain image patches. Concretely, the UMG separates the content information, i.e., the fingerprint ridge structure, and the style, i.e., textural information, of a given fingerprint minutiae patch and produces a synthetic image that has the content of the source domain and the style of the target domain. An overview of the *UMG* approach is shown as the box enclosed by the blue line in Figure 2.

### 3.6. UMG + ARL (Proposed Approach)

The proposed approach applies ARL with the UMG style transfer wrapper to further improve generalization performance. An illustration of the *ARL + UMG* approach is illustrated in Figure 2 as everything enclosed by the box formed by the solid, black line. Like the naïve approach, this method inherently assumes knowledge of a limited set of examples from the target domain sensor. Specifically, we assume 100 live and 100 PA images from the target sensor. From this small set of images from the target sensor, we produce a much larger set of synthetic images in the target domain using the UMG wrapper to transfer the style of the target domain to the content of the source domain training images.

The advantage of this approach is that we leverage the ability of the UMG wrapper to ensure a balanced dataset from all sensors (source and target), which we combine with ARL that forces the

network to learn a sensor-invariant representation. In the following section, we demonstrate the performance gains over the previous approaches and show that the UMG coupled with ARL achieves the new state-of-the-art in cross-sensor and cross-material generalization of fingerprint PAD.

## 4. Evaluation Procedure

In this section we describe the experimental protocol of the various experiments carried out in this study, the datasets involved in each experiment, and the implementation details of the UMG + ARL approach.

### 4.1. Experimental Protocol

To evaluate cross-sensor PAD performance, we adopt the leave-one-out protocol where one sensor is set aside for testing and the network is trained on data from all remaining sensors. To analyze separately the cross-sensor performance and the cross-material performance, we segment our evaluation to include the case where all materials during testing were included during training (cross-sensor only) and the case where no materials during training were seen in testing (cross-sensor and cross-material).

### 4.2. Datasets

The data used in the experiments for this paper are from the LivDet 2015, LivDet 2017, and MSU-FPAD datasets, which are summarized in Tables 2 and 3. The LivDet 2015 dataset consists of four sensors: Biometrika, CrossMatch, Digital Persona, and Green Bit. These sensors are all FTIR optical image capturing devices. We utilize this dataset to evaluate the generalization performance across different fingerprint readers with the same sensing technology. To further evaluate our approach on fingerprint readers with different sensing mechanisms, we experiment on fingerprint data from the Lumidigm sensor of the MSU-FPAD dataset. This sensor uses different sensing technology from the four seen in the LivDet 2015 as it is a multi-spectral, direct-view capture device. Finally, we incorporate a third dataset, LivDet 2017, which consists of three sensors: Digital Persona, Green Bit, and Orcanthus, where Orcanthus uses thermal-based imaging.

### 4.3. Implementation Details

The architecture of the encoder in the proposed approach is MobileNet-v1 with the final 1000-unit softmax layer removed, which is used to encode a latent representation  $\mathbf{z} \in \mathbb{R}^d$ . In our implementation,  $d = 1024$ . The target predictor is a single fully connected layer of 2-dimensions (for predicting live vs. PA) with a softmax activation. The adversary network consists of a fully connected layer with a softmax activation of output dimension equal to the number of source sensors in the training dataset, e.g., 3 in the leave-one-out protocol on the LivDet 2015 dataset.

Training adversarial losses is notoriously difficult and often requires extensive hyper-parameter tuning. For example, it was found advantageous during training to update the parameters,  $\theta_A$ , of the adversary network five times per every update of the encoder and target predictor. We also explored adjusting the number of hidden layers in the adversary network, but no significant improvements over a single layer network were observed. A grid

Table 2: Summary of the 2015 and 2017 Liveness Detection (LivDet) Datasets.

Dataset	LivDet 2015				LivDet 2017		
	Green Bit	Biometrika	Digital Persona	CrossMatch	Green Bit	Orcanthus	Digital Persona
Fingerprint Reader	DactyScan26	HiScan-PRO	U.are.U 5160	L Scan Guardian	Dacty Scan 84C	Cerits2 Image	U.are.U 5160
Image Size	500 x 500	1000 x 1000	252 x 324	640 x 480	500 x 500	300 x $n^\dagger$	252 x 324
Resolution (dpi)	500	1000	500	500	569	500	500
#Live Images Train / Test	1000 / 1000	1000 / 1000	1000 / 1000	1510 / 1500	1000 / 1700	1000 / 1700	1000 / 1692
#Spoof Images Train / Test	1000 / 1500	1000 / 1500	1000 / 1500	1473 / 1448	1200 / 2040	1180* / 2018	1199 / 2028
Spoof Materials	Ecoflex, Gelatine, Latex, Wood Glue, Liquid Ecoflex, RTV			Body Double, Ecoflex, Play-Doh, OOMOO, Gelatin	Wood Glue, Ecoflex, Body Double, Gelatine, Latex, Liquid Ecoflex		

<sup>†</sup> Fingerprint images captured by Orcanthus have a variable height (350 - 450 pixels) depending on the friction ridge content.

\* A Set of 20 Latex spoof fingerprints were present in the training data of Orcanthus; which were excluded in our experiments because only Wood Glue, Ecoflex, and Body Double are expected to be in the training dataset.

Table 3: Summary of the MSU-FPAD Dataset.

Dataset	MSU-FPAD	
	CrossMatch	Lumidigm
Fingerprint Reader	Guardian 200	Venus 302
Model	Guardian 200	Venus 302
Image Size	750 x 800	400 x 272
Resolution (dpi)	500	500
#Live Images Train / Test	2250 / 2250	2250 / 2250
#Spoof Images Train / Test	3000 / 3000	2250 / 2250
Spoof Materials	Ecoflex, PlayDoh, 2D Print (Matte Paper), 2D Print (Transparency)	

search was performed over the value of  $\alpha$  for selecting the influence of the adversary on updating the parameters,  $\theta_E$ , of the encoder, and the optimal parameter value of  $\alpha = 0.1$  was selected (See Eq. 2).

## 5. Experimental Results

Here we present the results of each experiment to evaluate the cross-sensor and cross-material generalization performance of the proposed approach. This section is divided into several parts to facilitate an in-depth analysis of the generalization performance of the algorithm to each of the following cases: cross-sensor, cross-material, and cross-sensing technology. A discussion on the effect of varying the number of assumed target domain images is included in section 5.4. We conclude this section with an analysis of the deep feature space prior to and following the application of the proposed methodology for fingerprint spoof generalization. The feature space analysis is conducted utilizing a 2-dimensional t-Distributed Stochastic Neighbor Embedding (t-SNE) visualization [26].

There has not been much prior work aimed specifically at improving cross-sensor generalization of fingerprint PAD; nonetheless, there are a few cross-sensor performance results reported in the literature. Chugh and Jain report the cross-sensor performance of Fingerprint Spoof Buster, which shares the same architecture of our base encoder model [4]. Therefore, in the following sections we compare our performance against that of Fingerprint Spoof Buster as the Base CNN model. Furthermore, Chugh and Jain report cross-sensor results in their work toward improving cross-

material generalization with the introduction of their UMG network wrapper [6]. For comparison with this approach, we refer to their work as the UMG approach in Tables 4 and 5 of this section.

### 5.1. Cross-Sensor Performance

To evaluate cross-sensor generalization we utilize the LivDet 2015 dataset which consists of four different FTIR optical fingerprint imaging devices and we apply a leave-one-out strategy where the algorithm is trained on only three of the four sensors at a time. We then compare the performance on a test set of data from these three sensors included in the training to the performance on a test set consisting of data from the remaining sensor. We repeat this procedure for all four combinations of sensors and report the results in Table 4.

To separate out the cross-sensor generalization performance from the related task of cross-material generalization, we first remove all the non-overlapping materials between the testing dataset of the target sensor and the training datasets of the three source sensors. For this experiment, Liquid Ecoflex and RTV materials were excluded from the testing sets when Green Bit, Biometrika, and Digital Persona were the target sensors; whereas, Body Double, Playdoh, and OOMOO were excluded from the testing set with CrossMatch as the target sensor.

As shown in Table 4, the proposed approach of UMG + ARL increases the average cross-sensor generalization in terms of True Detection Rate (TDR) at a False Detection Rate (FDR) of 0.2%<sup>4</sup> from 88.36% to 92.94% over the UMG only method. The proposed approach also maintains higher performance (TDR = 90.13%) on the source domain sensors compared to the UMG only approach (TDR = 86.98%). Lastly, we note that the standard deviation (s.d.) across the four experiments of cross-sensor generalization on the LivDet 2015 dataset is significantly reduced for the UMG + ARL method (11.27% to 4.09%), in comparison UMG only, indicating the robustness of the proposed approach.

For completeness, we include an evaluation of using an additional CNN architecture, Resnet-v1-50<sup>5</sup> [22], as the base encoder

<sup>4</sup>We consider this metric to be more representative of actual use cases as opposed to EER and ACE. Space limitation does not allow us to show the full ROC curve.

<sup>5</sup>Resnet-v1-50 was chosen since the authors of other SOTA fingerprint PAD algorithms were not willing to share their code and we found the

Table 4: Cross-Sensor Generalization Performance (TDR (%) @ FDR = 0.2 %)† with Leave-One-Out Method on LivDet 2015 Dataset with Materials Common to Training and Testing, i.e., Excluding Cross-Materials‡. Bio = Biometrika, CM = CrossMatch, DP = Digital Persona, and GB = GreenBit.

	Source* CM, DP, GB	Target* Bio	Source Bio, DP, GB	Target CM	Source Bio, CM, GB	Target DP	Source Bio, CM, DP	Target GB	Source Mean ± s.d.	Target Mean ± s.d.
Base CNN [4]	90.34	75.16	88.20	3.33	98.40	10.76	92.82	70.74	92.44 ± 4.40	40.00 ± 38.21
ARL	93.44	80.51	91.03	2.11	98.73	11.74	92.04	64.74	<b>93.81 ± 3.43</b>	39.78 ± 38.67
Naïve	87.74	84.80	88.23	97.37	96.96	59.13	88.08	90.68	90.25 ± 4.48	83.00 ± 16.72
UMG [6]	89.10	94.33	84.28	90.70	96.39	71.85	78.14	96.57	86.98 ± 7.71	88.36 ± 11.27
Naïve + ARL	90.18	91.86	87.87	98.95	94.21	52.07	89.15	83.92	90.35 ± 2.74	81.70 ± 20.69
UMG + ARL	88.98	92.83	88.48	97.54	96.18	87.61	86.88	93.78	90.13 ± 4.13	<b>92.94 ± 4.09</b>

† We use FDR = 0.2 % because this is the stringent metric being used by the IARPA Odin program. Due to space limits, it is challenging to show the complete Receiver Operating Curve (ROC) or Detection Error Tradeoff (DET) curve.

‡ Liquid Ecoflex and RTV materials were excluded from the testing sets of Green Bit, Biometrika, and Digital Persona. Body Double, Playdoh, and OOMOO were excluded from the testing set of CrossMatch.

\* Sensors included in the training set (source)

\* Sensors included in the test set (target)

Table 5: Cross-Sensor and Cross-Material Generalization Performance (TDR (%) @ FDR = 0.2 %) with Leave-One-Out Method on LivDet 2015 Dataset with Materials Exclusive to the Testing Datasets, i.e., Cross-Material Only. Bio = Biometrika, CM = CrossMatch, DP = Digital Persona, and GB = GreenBit.

	Source CM, DP, GB	Target Bio	Source Bio, DP, GB	Target CM	Source Bio, CM, GB	Target DP	Source Bio, CM, DP	Target GB	Source Mean ± s.d.	Target Mean ± s.d.
Base CNN [4]	90.34	63.92	88.20	4.46	98.40	11.39	92.82	72.39	92.44 ± 4.40	38.04 ± 35.06
ARL	92.78	72.58	91.03	6.06	98.73	13.08	92.04	49.69	<b>93.65 ± 3.47</b>	35.35 ± 31.33
Naïve	87.74	77.11	88.23	96.80	96.96	42.62	88.08	85.69	90.25 ± 4.48	75.56 ± 23.39
UMG [6]	89.10	87.01	84.28	81.37	96.39	54.43	78.14	92.23	86.98 ± 7.71	78.76 ± 16.82
Naïve + ARL	90.18	86.19	87.87	97.45	94.21	35.65	82.51	65.44	88.69 ± 4.88	71.18 ± 27.15
UMG + ARL	89.31	89.07	88.48	92.69	96.18	78.69	86.88	91.00	90.21 ± 4.10	<b>87.86 ± 6.29</b>

Table 6: Cross-Sensor Generalization Performance (TDR (%) @ FDR = 0.2 %) on Leave-Out Biometrika (LivDet 2015) using Resnet-v1-50 as the Base CNN Model. Bio = Biometrika, CM = CrossMatch, DP = Digital Persona, and GB = GreenBit.

	Source CM, DP, GB	Target Bio
Base CNN [22]	65.29	76.02
ARL	72.72	72.27
Naïve	73.55	90.79
UMG [6]	72.76	91.76
Naïve + ARL	73.05	92.18
UMG + ARL	<b>75.94</b>	<b>92.83</b>

to demonstrate the generality of the proposed approach. In Table 6, we report the performance with ResNet-v1-50 as the Base CNN model on LivDet 2015 with leaving Biometrika out as the target sensor. We see that the performance improvement is consistent for both Base CNN models, supporting the generality of the approach to any existing CNN architecture trained for fingerprint spoof detection. In the remaining experiments, we continue to report results for only Spoof Buster as the Base CNN model.

## 5.2. Cross-Sensor and Cross-Material Performance

We now compare the performance of each solution on the cross-sensor and cross-material experiment by following the same procedure as the cross-sensor experiment, while including only materials exclusive to the test datasets of LivDet 2015. Even

details of their reported implementations insufficient for reproducing for a fair evaluation.

though our system was trained to adversarially learn a sensor-invariant representation, we report the results of including unseen materials to evaluate whether we automatically obtain the added benefit of cross-material generalization (Table 5).

The results of Table 5 agree with the results of the cross-sensor only experiment shown previously; however, we note small performance declines due to the evaluation on only unknown spoof materials. Specifically, the average TDR at a FDR of 0.2% of the proposed approach decreased from 92.38% for cross-sensor only to 87.86% for cross-sensor and cross-material generalization on the target sensor. However, we notice that the performance degradation of the UMG + ARL method is less than the drop in performance of the UMG only approach, which further demonstrates the generalization benefits of incorporating ARL for fingerprint PAD. It seems that learning an invariance to the textural differences between different sensors also encourages an invariance to the textural differences between different spoof materials.

## 5.3. Cross-Sensing Technology Performance

In this section, we expand our analysis to include generalization across different fingerprint sensing mechanisms, where the sensing technology of the source fingerprint readers during training is different from the target test reader. For the first experiment we incorporate the data from the Lumidigm multispectral sensor of the MSU-FPAD database as the test sensor and the four FTIR optical sensors of LivDet 2015 as our training sensors. In this experiment we do not control for unknown materials between training and test sets, thus we could consider the evaluation as a combination of cross-sensor, cross-material, and cross-sensing technology. The results show that UMG + ARL achieves the highest general-

Table 7: Cross-Sensing Technology Generalization Performance (TDR (%) @ FDR = 0.2 %) with Four Sensors of LivDet 2015 Dataset Included During Training and Lumidigm from the MSU-FPAD Dataset Left Out For Testing. Bio = Biometrika, CM = CrossMatch, DP = Digital Persona, GB = GreenBit, and Lum = Lumidigm.

	Source Bio, CM, DP, GB	Target Lum
Base CNN [4]	<b>90.40</b>	0.60
ARL	87.41	3.00
Naïve	63.54	61.27
UMG [6]	88.24	80.60
Naïve + ARL	87.22	84.93
UMG + ARL	88.45	<b>88.60</b>

Table 8: Cross-Sensing Technology Generalization Performance (TDR (%) @ FDR = 0.2 %) on LivDet 2017 Dataset.

	Source Mean $\pm$ s.d.	Target Mean $\pm$ s.d.
Base CNN [4]	41.43 $\pm$ 5.83	4.63 $\pm$ 8.71
ARL	38.92 $\pm$ 6.64	7.35 $\pm$ 12.27
Naïve	43.90 $\pm$ 7.26	27.30 $\pm$ 6.82
UMG [6]	39.02 $\pm$ 14.71	34.80 $\pm$ 4.96
Naïve + ARL	<b>44.63 <math>\pm</math> 15.52</b>	30.30 $\pm$ 11.97
UMG + ARL	38.50 $\pm$ 14.63	<b>36.47 <math>\pm</math> 9.86</b>

ization TDR of 88.60% on the target domain sensor (Figure 7).

To further evaluate the generalization performance of the proposed UMG + ARL approach, we repeat the experiments on a third dataset, LivDet 2017, which consists of data from three different sensors: Green Bit (optical FTIR), Digital Persona (optical FTIR), and Orcanthus (thermal). With the inclusion of the Orcanthus sensor as a thermal based technology, we can evaluate cross-sensing technology performance where the underlying imaging technology between the sensors is substantially different. Further, we do not remove unseen material types between the training and testing datasets of LivDet 2017 for this experiment. As shown in Table 8, the generalization performance (TDR @ FDR = 0.2%) on LivDet 2017 improves over the state-of-the-art from 34.80% to 36.47%.

#### 5.4. Varying Number of Target Domain Images

To study the effect of varying the number of assumed target domain images available during training, we repeat the experiments in the leave-out Biometrika (LivDet 2015) scenario. Specifically, we run experiments on 50 and 250 live and PA training images from the target domain. As shown in Table 9, increasing the number of target domain images greatly benefits the naïve approach, but only marginally affects the UMG + ARL method. Therefore, the benefit of UMG + ARL is most pronounced in cases with limited target domain training examples. In the trade-off between time spent for data collection and performance, the proposed method can significantly help reduce the burden of expensive data collection.

#### 5.5. Feature Space Analysis

To explore the benefits of incorporating ARL on top of the UMG only approach, we extract 2-dimensional t-SNE feature embeddings of the live and spoof fingerprint minutiae patches from the final 1024-unit layer of the MobileNet-v1 encoder network,

Table 9: Cross-Sensor Generalization Performance (TDR (%) @ FDR = 0.2 %) on Leave-Out Biometrika (LivDet 2015) with Varying Number of Target Sensor Training Images.

	50 Images		250 Images	
	Source	Target	Source	Target
Naïve	91.21	90.15	91.04	95.29
UMG [6]	<b>93.19</b>	90.47	91.00	89.19
Naïve + ARL	85.64	91.43	<b>95.50</b>	<b>95.40</b>
UMG + ARL	90.76	<b>93.25</b>	90.71	93.04

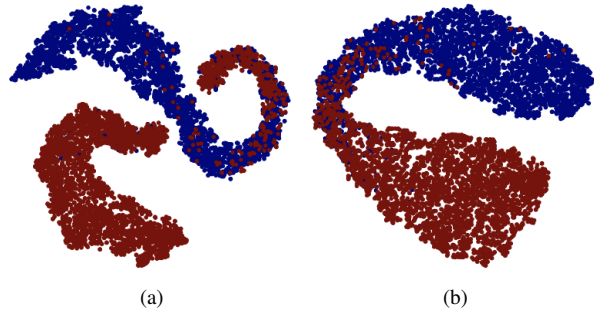


Figure 3: 2-dimensional t-SNE feature embeddings of the target sensor fingerprint minutiae patches for the (a) UMG only and (b) UMG + ARL models trained on the LivDet 2015 dataset with Biometrika, Green Bit, and Digital Persona as the source sensors and CrossMatch as the target sensor. The blue and red dots represent live and spoof minutiae patches of fingerprint impressions captured on the target sensor (CrossMatch), respectively.

prior to the softmax non-linearity, from the UMG only network and the UMG + ARL network. For brevity, we just show the results of the leave-one-out protocol on the LivDet 2015 dataset with Biometrika, Green Bit, and Digital Persona as the source sensors and CrossMatch as the target sensor. In Figure 3, we plot these embeddings to analyze the effect of adversarially enforcing the learning of a sensor-invariant representation. Figure 3 (a) shows the separation between live and spoof fingerprint minutiae patch embeddings of the UMG only network for minutiae patches from the target sensor, i.e., CrossMatch, whereas (b) shows the separation of the embeddings produced by the UMG + ARL approach. We can see that the proposed method provides noticeably better separation between the live and fingerprint spoof patches, resulting in the improved PAD performance.

## 6. Conclusion

Diverse and sophisticated presentation attacks pose a threat to the effectiveness of fingerprint recognition systems for reliable authentication and security. Previous PAD algorithms have demonstrated success in scenarios for which significant training data of bonafide and spoof fingerprint images are available, but are not robust to generalize well to novel spoof materials unseen during training. Additionally, previous fingerprint PAD solutions are not generalizable across different fingerprint readers, meaning that a PAD algorithm trained on a specific fingerprint reader will not perform well when applied to different fingerprint sensing devices.

The proposed approach towards fingerprint PAD demonstrates an improvement over the state-of-the-art, in terms of true detection rate (TDR) at a false detection rate (FDR) of 0.2%, on cross-sensor and cross-material generalization. In particular, incorporat-



ing adversarial representation learning with the Universal Material Generator (UMG) improves the cross-sensor generalization performance from a TDR of  $88.36 \pm 11.27\%$  to  $92.94 \pm 4.09\%$  on the LivDet 2015 dataset, while maintaining higher performance on the sensors seen during training. Further, including cross-materials with the cross-sensor evaluation leads to an improvement of  $78.76 \pm 16.82\%$  to  $87.86 \pm 6.29\%$ . Lastly, experiments involving cross-sensor, cross-material, and cross-sensing technology show average improvements of 80.60% to 88.60% and 34.80% to 36.47% with the proposed approach over state-of-the-art, on the MSU-FPAD and LivDet 2017 datasets, respectively.

## 7. Acknowledgment

This research is based upon work supported in part by the Office of the Director of National Intelligence (ODNI), Intelligence Advanced Research Projects Activity (IARPA), via IARPA R&D Contract No. 2017 - 17020200004. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies, either expressed or implied, of ODNI, IARPA, or the U.S. Government. The U.S. Government is authorized to reproduce and distribute reprints for governmental purposes notwithstanding any copyright annotation therein.

## References

- [1] D. Baldisserra, A. Franco, D. Maio, and D. Maltoni. Fake fingerprint detection by odor analysis. In *International Conference on Biometrics*, pages 265–272. Springer, 2006. 1
- [2] Y. Bengio, A. Courville, and P. Vincent. Representation learning: A review and new perspectives. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35(8):1798–1828, 2013. 2
- [3] K. Cao and A. K. Jain. Hacking mobile phones using 2d printed fingerprints. 2016. [https://www.youtube.com/watch?v=fZJI\\_BrMZXU&feature=youtu.be](https://www.youtube.com/watch?v=fZJI_BrMZXU&feature=youtu.be). 1
- [4] T. Chugh, K. Cao, and A. K. Jain. Fingerprint spoof buster: Use of minutiae-centered patches. *IEEE Transactions on Information Forensics and Security*, 13(9):2190–2202, 2018. 1, 2, 4, 6, 7, 8
- [5] T. Chugh and A. K. Jain. Fingerprint presentation attack detection: Generalization and efficiency. *arXiv preprint arXiv:1812.11574*, 2018. 2
- [6] T. Chugh and A. K. Jain. Fingerprint spoof generalization. *arXiv preprint arXiv:1912.02710*, 2019. 2, 5, 6, 7, 8
- [7] T. Chugh and A. K. Jain. OCT fingerprints: Resilience to presentation attacks. *arXiv preprint arXiv:1908.00102*, 2019. 2
- [8] G. Csurka. Domain adaptation for visual applications: A comprehensive survey. *arXiv preprint arXiv:1702.05374*, 2017. 3
- [9] Y. Ding and A. Ross. An ensemble of one-class svms for fingerprint spoof detection across different fabrication materials. In *2016 IEEE International Workshop on Information Forensics and Security (WIFS)*, pages 1–6, 2016. 2
- [10] H. Edwards and A. Storkey. Censoring representations with an adversary. *arXiv preprint arXiv:1511.05897*, 2015. 3
- [11] J. J. Engelsma, S. S. Arora, A. K. Jain, and N. G. Paulter. Universal 3d wearable fingerprint targets: advancing fingerprint reader evaluations. *IEEE Transactions on Information Forensics and Security*, 13(6):1564–1578, 2018. 1
- [12] J. J. Engelsma, K. Cao, and A. K. Jain. Raspireader: Open source fingerprint reader. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 41(10):2511–2524, 2018. 1
- [13] J. J. Engelsma and A. K. Jain. Generalizing fingerprint spoof detector: Learning a one-class classifier. *2019 International Conference on Biometrics (ICB)*, 2019. 2
- [14] N. Evans. *Handbook of Biometric Anti-spoofing: Presentation Attack Detection*. Springer, 2019. 1
- [15] R. Gajawada, A. Popli, T. Chugh, A. Namboodiri, and A. K. Jain. Universal material translator: Towards spoof fingerprint generalization. *2019 International Conference on Biometrics (ICB)*, 2019. 2
- [16] Y. Ganin, E. Ustinova, H. Ajakan, P. Germain, H. Larochelle, F. Laviolette, M. Marchand, and V. Lempitsky. Domain-adversarial training of neural networks. *The Journal of Machine Learning Research*, 17(1):2096–2030, 2016. 3
- [17] L. Ghiani, A. Hadid, G. L. Marcialis, and F. Roli. Fingerprint liveness detection using binarized statistical image features. In *2013 IEEE Sixth International Conference on Biometrics: Theory, Applications and Systems (BTAS)*, pages 1–6, 2013. 2
- [18] L. Ghiani, G. L. Marcialis, and F. Roli. Fingerprint liveness detection by local phase quantization. In *Proceedings of the 21st International Conference on Pattern Recognition*, pages 537–540, 2012. 2
- [19] L. Ghiani, D. Yambay, V. Mura, S. Tocco, G. L. Marcialis, F. Roli, and S. Schuckers. Livdet 2013 fingerprint liveness detection competition 2013. In *2013 International Conference on Biometrics (ICB)*, pages 1–6, 2013. 1
- [20] L. J. González-Soler, M. Gomez-Barrero, L. Chang, A. Pérez-Suárez, and C. Busch. Fingerprint presentation attack detection based on local features encoding for unknown attacks. *arXiv preprint arXiv:1908.10163*, 2019. 2
- [21] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio. Generative adversarial nets. In *Advances in Neural Information Processing Systems*, pages 2672–2680, 2014. 4
- [22] K. He, X. Zhang, S. Ren, and J. Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 770–778, 2016. 6, 7
- [23] A. G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto, and H. Adam. Mobilenets: Efficient convolutional neural networks for mobile vision applications. *arXiv preprint arXiv:1704.04861*, 2017. 4
- [24] International Standards Organization, “iso/iec 30107-1:2016, Information Technology - Biometric Presentation Attack Detection - Part 1: Framework”. <https://www.iso.org/standard/53227.html>. 1

- [25] P. D. Lapsley, J. A. Lee, D. F. Pare Jr, and N. Hoffman. Anti-fraud biometric scanner that accurately detects blood flow, Apr. 7 1998. US Patent 5,737,439. **1**
- [26] L. v. d. Maaten and G. Hinton. Visualizing data using t-sne. *Journal of Machine Learning Research*, 9(Nov):2579–2605, 2008. **6**
- [27] D. Maltoni, D. Maio, A. K. Jain, and S. Prabhakar. *Handbook of Fingerprint Recognition*. Springer Science & Business Media, 2nd edition, 2009. **1**
- [28] E. Marasco and A. Ross. A survey on antispoofing schemes for fingerprint recognition systems. *ACM Computing Surveys*, 47(2):1–36, 2014. **1**
- [29] E. Marasco and C. Sansone. On the robustness of fingerprint liveness detection algorithms against new materials used for spoofing. In *BIOSIGNALS*, volume 8, pages 553–555, 2011. **2**
- [30] E. Marasco and C. Sansone. Combining perspiration-and morphology-based static features for fingerprint liveness detection. *Pattern Recognition Letters*, 33(9):1148–1156, 2012. **2**
- [31] G. L. Marcialis, A. Lewicke, B. Tan, P. Coli, D. Grimberg, A. Congiu, A. Tidu, F. Roli, and S. Schuckers. First International Fingerprint Liveness Detection Competition Livdet 2009. In *International Conference on Image Analysis and Processing*, pages 12–23. Springer, 2009. **1**
- [32] G. L. Marcialis, F. Roli, and A. Tidu. Analysis of fingerprint pores for vitality detection. In *2010 20th International Conference on Pattern Recognition*, pages 1289–1292, 2010. **2**
- [33] T. Matsumoto, H. Matsumoto, K. Yamada, and S. Hoshino. Impact of artificial” gummy” fingers on fingerprint systems. In *Optical Security and Counterfeit Deterrence Techniques IV*, volume 4677, pages 275–289, 2002. **1**
- [34] V. Mura, L. Ghiani, G. L. Marcialis, F. Roli, D. A. Yambay, and S. A. Schuckers. Livdet 2015 fingerprint liveness detection competition 2015. In *2015 International Conference on Biometrics Theory, Applications, and Systems*, 2015. **1**
- [35] V. Mura, G. Orrù, R. Casula, A. Sibiriu, G. Loi, P. Tuveri, L. Ghiani, and G. L. Marcialis. Livdet 2017 fingerprint liveness detection competition 2017. In *2018 International Conference on Biometrics (ICB)*, pages 297–302, 2018. **1**
- [36] R. F. Nogueira, R. de Alencar Lotufo, and R. C. Machado. Fingerprint liveness detection using convolutional neural networks. *IEEE Transactions on Information Forensics and Security*, 11(6):1206–1213, 2016. **2**
- [37] ODNI, IARPA, “IARPA-BAA-16-04”. <https://www.iarpa.gov/index.php/research-programs/odin/odin-baa>. **1**
- [38] G. Orr, R. Casula, P. Tuveri, C. Bazzoni, G. Dessalvi, M. Micheletto, L. Ghiani, and G. L. Marcialis. Livdet in action - fingerprint liveness detection competition 2019, 2019. **1**
- [39] F. Pala and B. Bhanu. Deep triplet embedding representations for liveness detection. In *Deep Learning for Biometrics*, pages 287–307. Springer, 2017. **2**
- [40] A. Rattani, W. J. Scheirer, and A. Ross. Open set fingerprint spoof detection across novel fabrication materials. *IEEE Transactions on Information Forensics and Security*, 10(11):2447–2460, 2015. **2**
- [41] P. C. Roy and V. N. Boddeti. Mitigating information leakage in image representations: A maximum entropy approach. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 2586–2594, 2019. **4**
- [42] B. Tan, A. Lewicke, D. Yambay, and S. Schuckers. The effect of environmental conditions and novel spoofing methods on fingerprint anti-spoofing algorithms. In *2010 IEEE International Workshop on Information Forensics and Security*, pages 1–6, 2010. **2**
- [43] R. Tolosana, M. Gomez-Barrero, C. Busch, and J. Ortega-Garcia. Biometric presentation attack detection: Beyond the visible spectrum. *IEEE Transactions on Information Forensics and Security*, 15:1261–1275, 2019. **2**
- [44] E. Tzeng, J. Hoffman, K. Saenko, and T. Darrell. Adversarial discriminative domain adaptation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 7167–7176, 2017. **3**
- [45] M. Wang and W. Deng. Deep visual domain adaptation: A survey. *Neurocomputing*, 312:135–153, 2018. **3**
- [46] Q. Xie, Z. Dai, Y. Du, E. Hovy, and G. Neubig. Controllable invariance through adversarial feature learning. In *Advances in Neural Information Processing Systems*, pages 585–596, 2017. **4**
- [47] D. Yambay, L. Ghiani, P. Denti, G. L. Marcialis, F. Roli, and S. Schuckers. Livdet 2011 fingerprint liveness detection competition 2011. In *2012 5th IAPR international conference on biometrics (ICB)*, pages 208–215, 2012. **1**
- [48] S. Yoon, J. Feng, and A. K. Jain. Altered fingerprints: Analysis and detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 34(3):451–464, 2012. **1**
- [49] B. H. Zhang, B. Lemoine, and M. Mitchell. Mitigating unwanted biases with adversarial learning. In *Proceedings of the 2018 AAAI/ACM Conference on AI, Ethics, and Society*, pages 335–340, 2018. **3**