Fingerprint Spoof Detection Using Minutiae-based Local Patches

Tarang Chugh, Kai Cao and Anil K. Jain
Department of Computer Science and Engineering
Michigan State University, East Lansing, Michigan 48824
{chughtar, kaicao, jain}@cse.msu.edu

Abstract

The individuality of fingerprints is being leveraged for a plethora of day-to-day applications, ranging from unlocking a smartphone to international border security. While the primary purpose of a fingerprint recognition system is to ensure a reliable and accurate user authentication, the security of the recognition system itself can be jeopardized by spoof attacks. This study addresses the problem of developing accurate and generalizable algorithms for detecting fingerprint spoof attacks. We propose a deep convolutional neural network based approach utilizing local patches extracted around fingerprint minutiae. Experimental results on three public-domain LivDet datasets (2011, 2013, and 2015) show that the proposed approach provides state of the art accuracies in fingerprint spoof detection for intra-sensor, cross-material, cross-sensor, as well as cross-dataset testing scenarios. For example, the proposed approach achieves a 69% reduction in average classification error for spoof detection under both known material and cross-material scenarios on LivDet 2015 datasets.

1. Introduction

With the ubiquitous deployment of fingerprint recognition systems in many day-to-day applications, such as financial transactions, international border security, unlocking a smartphone, etc., the vulnerability of the system security to *presentation attacks* is of growing concern [1, 2]. The ISO standard *IEC 30107-1:2016(E)* [3] defines presentation attacks as the *"presentation to the biometric data capture subsystem with the goal of interfering with the operation of the biometric system"*. These attacks can be realized through a number of methods including, but not limited to, use of (i) *gummy fingers* [4], *i.e.* fabricated finger-like objects with accurate imitation of another individual's fingerprint ridge-valley structures, (ii) *2D or 3D printed fingerprint targets* [5, 6, 7], (iii) *altered fingerprints* [8], *i.e.* intentionally tampered or damaged real fingerprint patterns to avoid identification, and (iv) *cadaver fingers* [9] (see Fig. 1).

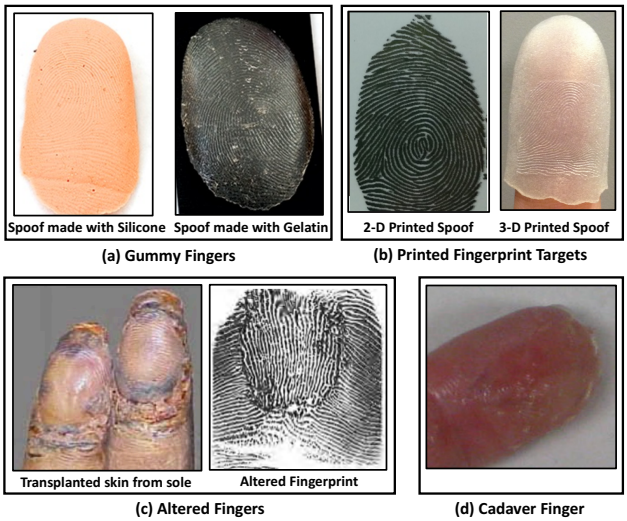

Figure 1: Fingerprint presentation attacks can be realized using (a) gummy fingers [4, 11], (b) 2D or 3D printed fingerprint targets [5, 6, 7], (c) altered fingers [8], or (d) cadaver fingers [9].

Among these, fingerprint spoof attacks (*i.e.* gummy fingers and printed targets) are the most common form of presentation attacks, with a multitude of fabrication processes ranging from basic *molding and casting* to utilizing sophisticated 2D and 3D printing techniques [4, 5, 6, 7]. Commonly available materials, such as gelatin, silicone, playdoh, etc., have been utilized to generate fingerprint spoofs (see Fig. 2), capable of circumventing a fingerprint recognition system security with a reported success rate of more than 70% [10]. For instance, in 2013, a Brazilian doctor was arrested for using spoof fingers made of silicone to fool the biometric attendance system at a hospital in Sao Paulo[1]. Cao and Jain [5] demonstrated a simple hack to bypass the biometric security of two smartphones by scanning and printing fingerprints using silver conductive ink on a special AgIC (silver ink circuit) paper.

Fingerprint spoof detection methods are urgently needed

[1] http://www.bbc.com/news/world-latin-america-21756709

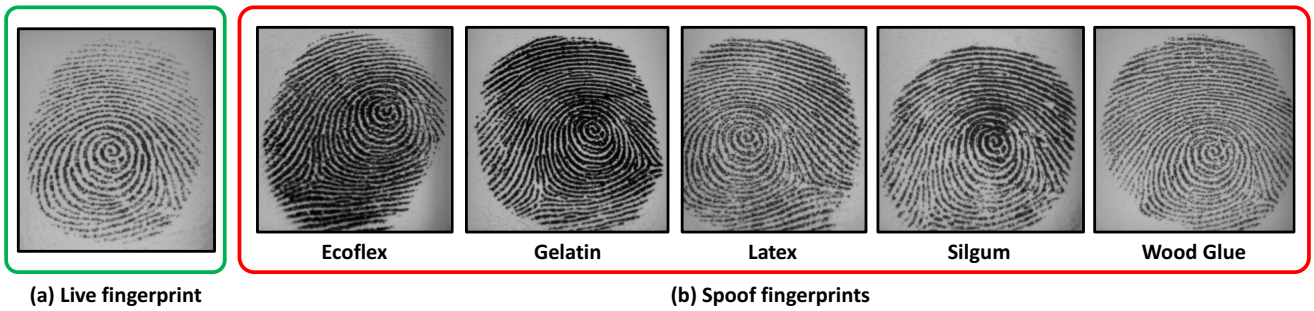

Figure 2: Visual comparison between (a) a live Fingerprint, and (b) the corresponding spoofs (of the same live finger) made with different materials. Images are taken from LivDet-2011 dataset (Biometrika sensor) [11].

to thwart such attacks on fingerprint authentication systems, thereby increasing user confidence in such systems. The various anti-spoofing approaches proposed in the literature can be broadly classified into hardware-based and software-based solutions [1, 9, 12]. The hardware-based solutions typically require the fingerprint reader to be augmented with additional sensor(s) to detect the characteristics of vitality, such as blood flow [13], skin distortion [14], odor [15] and so on. There are also special types of fingerprint sensors, such as Lumidigm's multispectral scanner [16], that capture sub-dermal ridge pattern in the finger. Software-based solutions, on the other hand, extract features from the presented fingerprint image (or a sequence of frames) acquired by the fingerprint sensors, without incurring any additional hardware cost, to differentiate between live and spoof fingers. The software-based solutions published in the literature typically utilize one of the following approaches: (i) anatomical features (e.g. pore locations and their distribution [17]), (ii) physiological features (e.g. perspiration [18]), or (iii) texture-based features (e.g. Local Phase Quantization (LPQ) [19], Binarized Statistical Image Features (BSIF) [20], and Weber Local Descriptor [21]). Gragniello et al. [22] proposed a 2D local contrast-phase descriptor (LCPD), utilizing both spatial and frequency domain information. In contrast to the custom-tailored anti-spoof features, Menotti et al. [23] and Nogueira et al. [24] have proposed convolutional neural network (CNN) based solutions whose performances were shown to surpass many published spoof detection algorithms.

One of the limitations of many of the published anti-spoof methods is their inability to generalize across spoof materials. Studies in [24, 25, 26] have shown that when a spoof detector is evaluated on spoofs fabricated using materials that were not seen during training, there can be up to a three-fold increase in the spoof detection error rates. To generalize an algorithm's effectiveness across fabrication materials, called cross-material performance, some studies have approached spoof detection as an *open-set problem*[2]. Rattani et al. [26] applied the Weibull-calibrated SVM (W-SVM), a variant of SVM based on properties of statistical extreme value theory, to detect spoofs made of new materials. Ding and Ross [27] trained an ensemble of multiple One-Class SVMs using textural features extracted from only live fingerprint images.

A series of fingerprint Liveness Detection (LivDet) competitions have been held since 2009 to advance state-of-the-art and benchmark the proposed anti-spoofing solutions [28]. The best performer in LivDet 2015 [12], Nogueira et al. [24], utilized transfer learning, where pre-trained deep CNNs originally designed for object recognition were fine-tuned on fingerprint images to differentiate between live and spoof fingerprints. In their approach, the networks were trained on whole fingerprint images resized to 227×227 pixels for VGG [29] and 224×224 pixels for AlexNet [30] as required by these networks. However, there are three disadvantages of using this approach. First, fingerprint images from some sensors, such as Crossmatch L Scan Guardian (640×480), have a large blank area ($\geq 50\%$) surrounding the friction ridge region. Directly resizing these images, from 640×480 to 227×227, eventually results in the friction ridge area occupying less than 10% of the original image size. Secondly, resizing a rectangular image, say $w \times h$, to a square image, say $p \times p$, leads to different amounts of information retained in the two spatial image dimensions. Lastly, since different sensors capture images of varying size, such sensor specific characteristics learned by the networks can adversely affect the cross-sensor performance.

The randomness involved in the spoof fabrication process itself, can generate some artifacts, such as missing friction ridge regions, cracks, etc. The primary consequence of such noise is the creation of spurious minutiae in the fingerprint (see Fig. 3). The local regions around these spurious minutiae can provide salient cues to differentiate a spoof fingerprint from live fingerprint. We utilize this observation to train a two-class CNN using local patches ex-

[2]Open-set problems address the possibility of unknown classes during testing, compared to closed-set problems where all possible classes (spoof materials) that will be encountered are known during training.

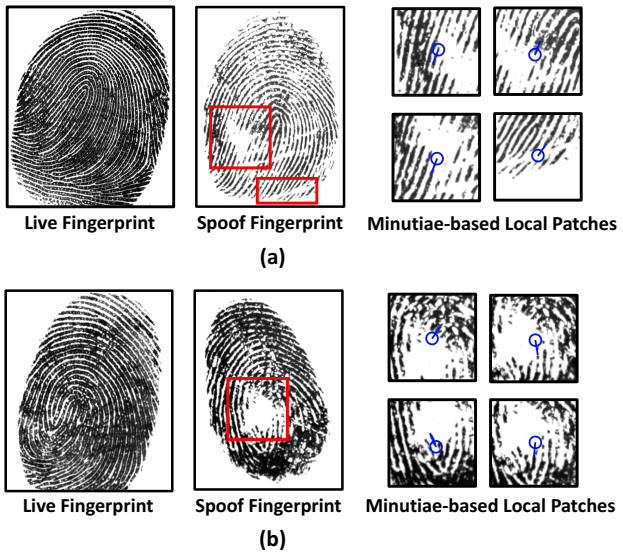

Figure 3: Examples of two fingerprint images (a) and (b) from live fingers and corresponding spoof fingers, with the artifacts introduced in the spoofs highlighted in red. The minutiae based local patches extracted around the artifacts are also presented. The images are taken from LivDet 2015 - Biometrika Sensor and the spoof material used is RTV.

tracted around minutiae, as opposed to the whole fingerprint image, to design a fingerprint spoof detector. We show that this approach is more robust to novel fabrication materials and cross-sensor scenarios than earlier approaches that utilize the whole image [24]. The proposed approach, utilizing $p \times p$ local patches ($p = 96$), (i) addresses the drawbacks of using the whole fingerprint image to train the CNN, (ii) provides large amount of data (an average of 48 patches/fingerprint image) to train the CNN, and (iii) captures salient information from local regions, required to differentiate between spoof and live fingerprints. The output of the CNN is a probability score in the range $[0 - 1]$, defined as *Spoofness Score*; higher the spoofness score, more likely the patch is extracted from a spoof fingerprint. For a given image, the spoofness scores corresponding to the local patches are averaged to give the global spoofness score for the whole image. Furthermore, a fusion of CNN models trained on multi-scale patches, centered on minutiae, is shown to further boost the spoof detection performance. The main contributions of this study are as follows:

- Utilized fingerprint domain-knowledge to design a robust fingerprint spoof detector, where local patches centered around fingerprint minutiae are used for training a CNN model.
- Utilized multi-scale local patches to improve the spoof detection performance. The proposed approach is shown to be robust to the different image sizes output by different sensors.
- Experimental results on LivDet 2011, LivDet 2013, and LivDet 2015 datasets show that the proposed minutiae-based fingerprint spoof detector outperforms the best results published on these three datasets. For example, in LivDet 2015, our algorithm achieves 98.61% average accuracy compared to 95.51% achieved by the LivDet 2015 winner [12].

2. Proposed Spoof Detection Approach

The proposed approach includes two stages, an offline training stage and an online testing stage. The offline training stage involves (i) detecting fingerprint minutiae, (ii) extracting local patches centered on minutiae locations, and (iii) training CNN models on the local patches. During the testing stage, the final spoof detection decision is made based on the average of spoofness scores output from the CNN model. An overview of the proposed approach is presented in Fig. 4.

2.1. Minutiae Detection

The minutiae were extracted using the algorithm in [31]. The different LivDet datasets used in this study comprise of fingerprint images captured at varying resolutions, ranging from 500 dpi to 1000 dpi. Since the minutiae detector in [31] was designed for 500 dpi images, all fingerprint images are resized to ensure a standard resolution of 500 dpi. The average number of minutiae detected for the LivDet datasets were 46 per live image (s.d. = 6.2) and 50 per spoof image (s.d. = 6.9).

2.2. Local Patch Extraction

Given a fingerprint image I and a set of k detected minutiae points $M = \{m_1, m_2, \ldots, m_k\}$, a corresponding set of k local patches $L = \{l_1, l_2, \ldots, l_k\}$, each of size $[p \times p]$ where $(p = 96)$, are extracted. Each local patch (l_i) is centered at the corresponding minutia point (m_i). In case the detected minutiae is close to the image boundary, i.e. some region of the local patch lies outside the image region, then the patch region is shifted inwards such that it is completely embedded within the fingerprint region, ensuring the size of each patch to be $[p \times p]$. Figs. 4 and 5 present examples of real and spoof fingerprint images and the corresponding local patches centered around minutiae points. All the local patches are resized[3] to 299×299 pixels as required by the Inception-v3 model.

2.3. Convolutional Neural Network

A Convolutional Neural Network (CNN) is a cascade of multiple layers consisting of linear and non-linear pro-

[3]TensorFlow's resize utility with bilinear interpolation was used; available at `https://www.tensorflow.org/api_docs/python/tf/image/resize_images`

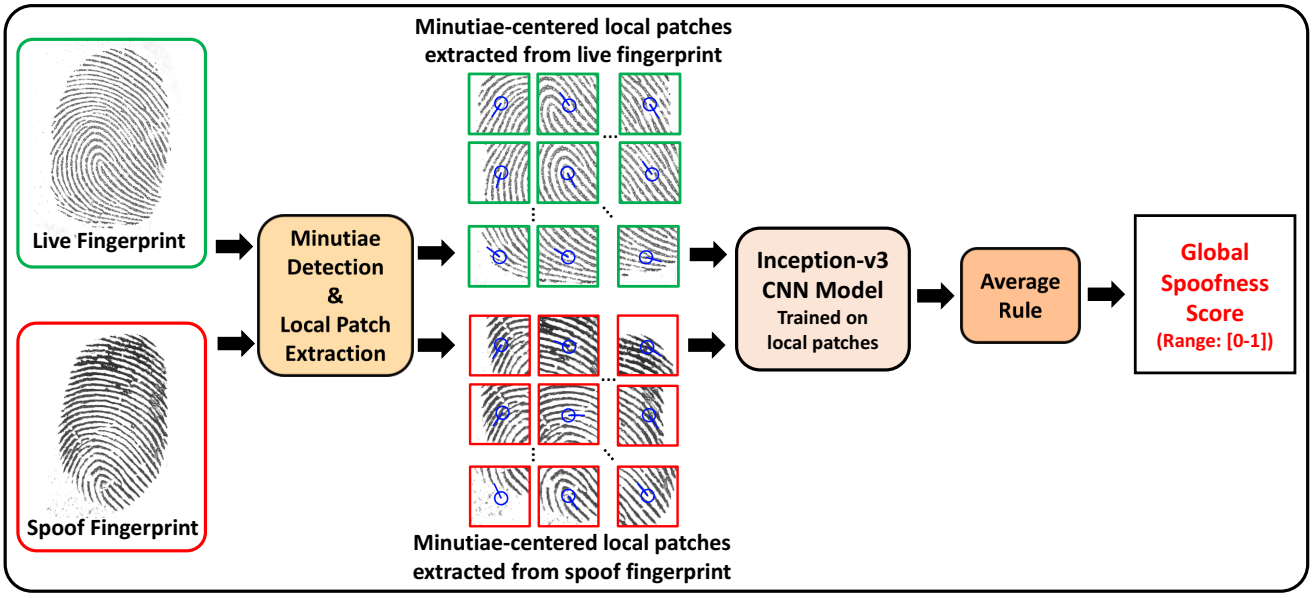

Figure 4: An overview of the proposed approach for fingerprint spoof detection using convolutional neural networks trained on local patches based on minutiae locations.

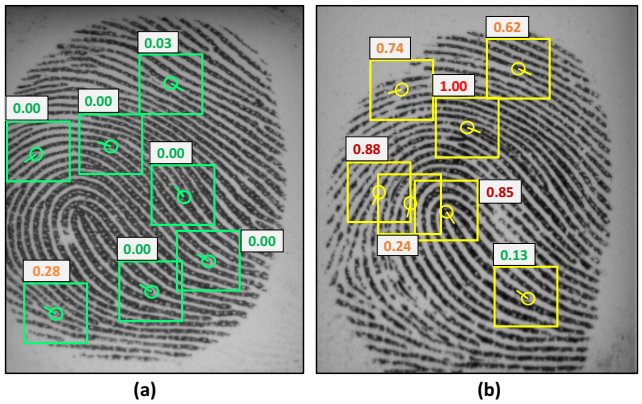

Figure 5: Local patches extracted around the fingerprint minutiae for (a) real fingerprint, and (b) spoof fingerprint (gelatin). The spoofness score for each patch is in the range $[0-1]$; higher the score, more likely the patch is extracted from a spoof fingerprint.

cessing units. These layers, when stacked together, can learn a complex multi-level representation of the input data corresponding to different levels of abstraction. Since the success of AlexNet [30] in ILSVRC-2012 [32], different deep CNN architectures have been proposed in literature, such as VGG, GoogleNet (Inception), Inception v2-v4, and ResNets. Nogueira et al. [24] utilized a pre-trained VGG architecture [29] to achieve the best performance in LivDet 2015 [12]. In this study, we utilize the Inception-v3 architecture [33] because it offers the following advantages: (i) Inception-v3 achieved a reduction of 40% top-5 error rate over VGG on the ILSVRC 2012 classification benchmark, (ii) the number of model parameters to be trained in Inception-v3 is significantly smaller than the number of model parameters in VGG, and (iii) it allows a larger input image size of 299×299 pixels, compared to 227×227 pixels for VGG.

We utilized the TF-Slim library[4] implementation of the Inception-v3 architecture. The last layer of the architecture, a 1000-unit softmax layer (originally designed to predict the $1,000$ classes of ImageNet dataset) was replaced with a 2-unit softmax layer for the two-class problem, *i.e.* Live vs. Spoof. The optimizer used to train the network was RMSProp, with a batch size of 32, and an adaptive learning rate with exponential decay, starting at 0.01 and ending at 0.0001. Data augmentation techniques, such as random cropping, brightness adjustment, horizontal and vertical flipping, are employed to ensure the trained model is robust to the possible variations in fingerprint images.

2.4. Spoofness Score

The output from the softmax layer of the trained Inception-v3 model is in the range $[0-1]$, defined as *Spoofness Score*. The larger the Spoofness Score, the higher the likelihood that the input local patch belongs to the Spoof class (see Fig. 5). For an input test image I, the spoofness scores $s^I_{i\in\{1,2,\ldots,k\}}$ corresponding to the k minutiae-based local patches, extracted from the input image, are averaged to give a global Spoofness Score S^I. An adaptive threshold that minimizes the average classification error on training dataset is utilized as the classification threshold. An image with a Spoofness Score below the threshold is classified as

[4] https://github.com/tensorflow/models/tree/master/slim

Table 1: Summary of the Liveness Detection (LivDet) datasets utilized in this study.

Dataset	**LivDet 2011 [11]**				**LivDet 2013 [34]**		**LivDet 2015 [12]**			
Sensor	Biometrika	ItalData	Digital Persona	Sagem	Biometrika	ItalData	GreenBit	Biometrika	Digital Persona	Crossmatch
Model	FX2000	ET10	4000B	MSO300	FX2000	ET10	DactyScan26	HiScan-PRO	U.are.U 5160	L Scan Guardian
Image Size	315×372	640×480	355×391	352×384	315×372	640×480	500×500	1000×1000	252×324	640×480
Resolution (dpi)	500	500	500	500	569	500	500	1000	500	500
#Live Images Train / Test	1000 / 1000	1000/1000	1000/1000	1000/1000	1000/1000	1000/1000	1000/1000	1000/1000	1000/1000	1510/1500
#Spoof Images Train / Test	1000 / 1000	1000/1000	1000/1000	1000/1000	1000/1000	1000/1000	1000/1500	1000/1500	1000/1500	1473/1448
Cooperative*	Yes	Yes	Yes	Yes	No	No	Yes	Yes	Yes	Yes
Spoof Materials	Ecoflex, Gelatine, Latex, Silgum, Wood Glue		Gelatine, Latex, Play Doh, Silicone, Wood Glue		Ecoflex, Gelatine, Latex, Modasil, Wood Glue		Ecoflex, Gelatine, Latex, Wood Glue, Liquid Ecoflex, RTV			Body Double, Ecoflex, Play Doh, OOMOO, Gelatin

*In the cooperative method, a subject willingly provides a negative impression of the fingerprint as a mold, while in the non-cooperative method, the fingerprint mold is created by using the latent fingerprint lifted off a surface touched by the subject.

Table 2: Performance (Average Classification Error [%]) comparison of software-based spoof detection studies, most of them compiled from [24, 35].

Study	**Approach**	**LivDet 2011**	**LivDet 2013***	**LivDet 2015**
Ghiani et al. [19]	Local Phase Quantization (LPQ)	11.1	3.0	N/A
Gragniello et al. [21]	Weber Local Descriptor (WLD)	7.9	N/A	N/A
Ghiani et al. [20]	Binarized Statistical Image Features (BSIF)	7.2	2.1	N/A
Gragniello et al. [22]	Local Contrast-Phase Descriptor (LCPD)	5.7	1.3	N/A
Nogueira et al. [24]	Transfer Learning + CNN-VGG + Whole Images	4.5	1.1	4.5
Proposed Approach	CNN-Inception v3 + Minutiae-based local patches	**2.6**	**0.5**	**1.4**

*LivDet 2013 includes results for Biometrika and Italdata sensors.

live, otherwise as spoof. The adaptive threshold performed slightly better than selecting a pre-defined threshold of 0.5.

3. Experimental Results

3.1. Datasets

In order to evaluate performance of the proposed approach, we utilized LivDet 2011 [11] and LivDet 2015 [12] datasets. Each of these datasets contain over $16,000$ fingerprint images, acquired from four different sensors, with equal numbers of live and spoof fingerprints that are equally split between training and testing sets. However, all the spoof fingerprints are fabricated using the *cooperative method i.e.* with user cooperation. To analyze the performance of the proposed approach on spoofs fabricated using non-cooperative method, fingerprint images from Biometrika and Italdata sensors from LivDet 2013 dataset [34] are also used. In LivDet 2015, the testing set included spoofs fabricated using new materials, that were not known in training set. These new materials included liquid ecoflex and RTV for Biometrika, Digital Persona, and Green Bit sensors, and OOMOO and gelatin for Crossmatch sensor. Table 1 presents a summary of the datasets used in this study, and Table 2 presents a performance comparison between software-based spoof detection solutions utilizing these datasets.

3.2. Performance Evaluation Metrics

The performance of the proposed approach is evaluated following the metrics used in LivDet [28].

- $Ferrlive$: Percentage of misclassified live fingerprints.
- $Ferrfake$[5]: Percentage of misclassified spoof fingerprints.

The average classification error (ACE) is defined as:

$$ACE = \frac{F_{errlive} + F_{errfake}}{2} \tag{1}$$

Additionally, we also report the $Ferrfake$ @ $Ferrlive = 1\%$ for each of the experiments as reported by [28]. This value represents the percentage of spoofs able to breach the biometric system security when the rate of legitimate users that are rejected is no more than 1%.

3.3. Results

The proposed approach is evaluated under the following four scenarios of fingerprint spoof detection, which reflect an algorithm's robustness against new spoof materials, use of different sensors and/or different environments.

[5]When all the spoof fabrication materials are known during the training, this metric is referred to as $Ferrfake\_known$, and in case all the spoof fabrication materials to be encountered during testing are not known during training, this metric is referred to as $Ferrfake\_unknown$.

Table 3: Performance comparison between the proposed approach (bottom) and state-of-the-art (top) reported on LivDet 2015 dataset [12]. Separate networks are trained on the training images captured by each of the four sensors. *Ferrfake known* and *Ferrfake unknown* correspond to Known Spoof Materials and Cross-Material scenarios, respectively.

	LivDet 2015	**Ferrlive (%)**	**Ferrfake† (%)**	**Ferrfake known (%)**	**Ferrfake unknown* (%)**	**ACE (%)**	**Ferrfake (%) @ Ferrlive= 1% [28]**
State-of-the-Art [12]	GreenBit	3.50	5.33	4.30	7.40	4.60	17.90
	Biometrika	8.50	3.73	2.70	5.80	5.64	15.20
	Digital Persona	8.10	5.07	4.60	6.00	6.28	19.10
	Crossmatch	0.93	2.90	2.12	4.02	1.90	2.66
	Average	4.78	4.27	3.48	5.72	4.49	13.24
Proposed Approach	**LivDet 2015**	**Ferrlive (%)**	**Ferrfake† (%)**	**Ferrfake known (%)**	**Ferrfake unknown* (%)**	**ACE (%)**	**Ferrfake (%) @ Ferrlive = 1%**
	GreenBit	1.20	2.53	2.80	2.00	2.00	3.07
	Biometrika	2.20	1.47	1.80	0.80	1.76	3.80
	Digital Persona	1.80	0.60	0.60	0.60	1.08	1.47
	Crossmatch	0.00	1.66	2.82	0.00	0.81	0.28
	Average	**1.16**	**1.56**	**1.97**	**0.81**	**1.39**	**2.17**

† Ferrfake includes spoofs fabricated using both known and previously unseen materials. It is an average of Ferrfake-known and Ferrfake-unknown, weighted by the number of samples in each category.
*The unknown spoof materials in LivDet 2015 test dataset include Liquid Ecoflex and RTV for Green Bit, Biometrika, and Digital Persona sensors, and OOMOO and Gelatin for Crossmatch sensor.

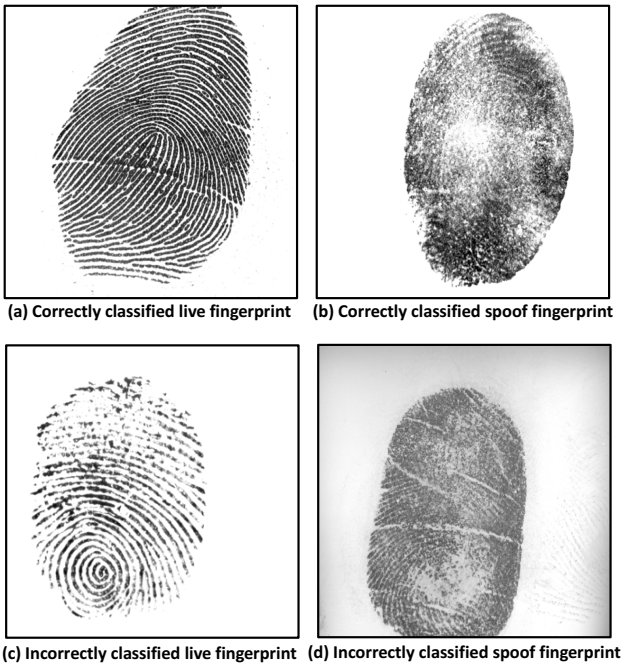

Figure 6: Example fingerprints for Biometrika sensor from LivDet 2015 dataset, correctly and incorrectly classified by our proposed approach.

3.3.1 Intra-Sensor, Known Spoof Materials

In this setting, all the training and testing images are captured using the same sensor, and all spoof fabrication materials utilized in the test set are known a priori. Our experimental results show that training the Inception-v3 model from scratch, using minutiae-based local patches, performs better than utilizing a pre-trained network, as reported in [24]. The large amount of available data, in the form of local fingerprint patches, is sufficient to train the deep architecture of Inception-v3 model without overfitting. Also, a score level fusion of the proposed Inception-v3 model trained on local patches, with an independent Inception-v3 model trained on whole fingerprint images does not offer any advantage in terms of performance improvement. It was reported in [28] that most of the algorithms submitted to LivDet 2015 did not perform well on Digital Persona sensor due to the small image size. Our approach based on local patches does not suffer from this limitation. Tables 3 and 4 present the performance comparison between the proposed approach and the state-of-the-art results for the LivDet datasets utilized in this study. Independent Inception-v3 networks are trained for each evaluation. Note that in LivDet 2015 (Table 3), this scenario is

Table 4: Performance comparison between the proposed approach and state-of-the-art results reported on LivDet 2011 and LivDet 2013 datasets for intra-sensor, known materials scenario, in terms of Average Classification Error (ACE), and Ferrfake @ Ferrlive = 1%.

Dataset	**State-of-the-Art**	**Proposed Approach**	
LivDet 2013	**ACE (%)**	**ACE (%)**	**Ferrfake @ Ferrlive = 1%**
Biometrika	1.10 [20]	0.60	0.10
ItalData	0.40 [24]	0.40	0.30
Average	0.75	**0.50**	**0.20**
LivDet 2011			
Biometrika	4.90 [22]	2.60	6.30
Digital Persona	2.00 [35]	2.70	5.50
ItalData	8.00 [24]	3.25	8.00
Sagem	1.70 [24]	1.80	18.92
Average	4.15	**2.59**	**9.68**

Table 5: Performance comparison between the proposed approach and state-of-the-art results [24] reported on LivDet 2011 and LivDet 2013 datasets for cross-material experiments, in terms of Average Classification Error (ACE), and Ferrfake @ Ferrlive = 1%.

Dataset	**Spoof Materials**		**Nogueira et al. [24]**	**Proposed Approach**	
	Materials - Training	**Materials - Testing**	**ACE (%)**	**ACE (%)**	**Ferrfake @ Ferrlive= 1%**
Biometrika 2011	EcoFlex, Gelatine, Latex	Silgum, WoodGlue	10.1	4.4	6.5
Biometrika 2013	Modasil, WoodGlue	EcoFlex, Gelatine, Latex	4.9	3.1	4.2
ItalData 2011	EcoFlex, Gelatine, Latex	Silgum, WoodGlue, Other	22.1	5.7	7.8
ItalData 2013	Modasil, WoodGlue	EcoFlex, Gelatine, Latex	6.3	0.8	0.9
Average			10.9	**3.5**	**4.9**

represented by the *Ferrfake known*. For LivDet 2011 and 2013 datasets (Table 4), all spoof materials in the test set were known during training. Fig. 6 presents example fingerprint images for Biometrika sensor from LivDet 2015 dataset that were correctly and incorrectly classified by the proposed approach.

In order to evaluate the significance of utilizing minutiae locations for extracting local patches, we trained independent Inception-v3 models on a similar number of local patches, extracted in a raster scan mode from LivDet 2015 datasets. It was observed that the models trained on minutiae-centered local patches achieved a significantly higher reduction (69%) in average classification error, compared to the reduction (17%) achieved by the models trained on raster scan local patches.

We also evaluate the impact of local patch size on the performance of the proposed approach, by comparing the performance of three CNN models trained on minutiae-centered local patches of size $[p \times p]$ where $p = \{64, 96, 128\}$, extracted from the fingerprint images captured by Biometrika sensor for LivDet 2011 dataset. Among these three models, the one trained on local patches of size $[96 \times 96]$ performed the best. However, a score-level fusion, using average-rule, of the three models reduced the average classification error (ACE) from 2.6% to 2.2%, and Ferrfake from 6.3% to 5.5% @ Ferrlive $= 1\%$. Similar performance gains were observed for other sensors, but there is a trade off between the performance gain and the computational requirements for spoof detector.

3.3.2 Intra-Sensor, Cross-Material

In this setting, the same sensor is used to capture all training and testing images, but the spoof images in the testing set are fabricated using new materials that were not seen during training. For the first set of cross-material experiments, we utilize the LivDet 2015 dataset which contains two new spoof materials in the testing set for each sensor, *i.e.* Liquid Ecoflex and RTV for Green Bit, Biometrika, and Digital Persona sensors, and OOMOO and Gelatin for Crossmatch sensor. The performance of the proposed approach on cross-material experiments for LivDet 2015 dataset is presented in Table 3 (column $Ferrfake\_unknown$) and is compared with the state-of-the-art performance reported in [12]. A significant reduction in the error rate is achieved by the proposed method. For better generalizability, a second set of cross-material experiments are performed on LivDet 2011 and LivDet 2013 datasets, following the protocol adopted by the winner of LivDet 2015 [24]. Table 5 presents the achieved error rates on these experiments, along with the spoof fabrication materials used in training and testing sets.

3.3.3 Cross-Sensor Evaluation

In this evaluation, the training and the testing images are obtained from two different sensors but from the same database. This setting reflects the algorithm's strength in learning the common characteristics used to distinguish live and spoof fingerprints across fingerprint acquisition devices. For instance, using LivDet 2011 dataset, images from Biometrika sensor are used for training, and the images from ItalData sensor are used for testing. We follow the protocol for selection of training and testing sets for cross-sensor and cross-dataset experiments as adopted by Nogueira et al. [24]. Table 6 compares the average classification error and Ferrfake @ Ferrlive = 1% for the proposed approach with the results obtained by [24] on cross-sensor experiments.

Table 6: Performance comparison between the proposed approach and state-of-the-art results [24] reported on LivDet 2011 and LivDet 2013 datasets for cross-sensor experiments, in terms of Average Classification Error (ACE), and Ferrfake @ Ferrlive = 1%.

Training Dataset, Testing Dataset	**Nogueira et al. [24]**	**Proposed Approach**	
	ACE (%)	**ACE (%)**	**Ferrfake (%) @ Ferrlive = 1%**
Biometrika 2011, ItalData 2011	37.2	29.5	76.4
ItalData 2011, Biometrika 2011	31.0	24.9	72.4
Biometrika 2013, ItalData 2013	8.8	6.7	41.6
ItalData 2013, Biometrika 2013	2.3	5.1	74.5
Average	19.8	**16.6**	**66.2**

Table 7: Performance comparison between the proposed approach and state-of-the-art results [24] reported on LivDet 2011 and LivDet 2013 datasets for cross-dataset experiments, in terms of Average Classification Error (ACE) and Ferrfake @ Ferrlive = 1%.

Training Dataset, Testing Dataset	**Nogueira et al. [24]**	**Proposed Approach**	
	ACE (%)	**ACE (%)**	**Ferrfake (%) @ Ferrlive = 1%**
Biometrika 2011, Biometrika 2013	15.5	7.9	90.6
Biometrika 2013, Biometrika 2011	46.8	34.4	90.1
ItalData 2011, ItalData 2013	14.6	3.3	4.3
ItalData 2013, ItalData 2011	46.0	29.9	92.4
Average	30.7	**18.9**	**69.4**

3.3.4 Cross-Dataset Evaluation

In this scenario, the training and the testing images are obtained using the same sensor, but from two different databases. For instance, training images are acquired using Biometrika sensor from LivDet 2011 dataset and the testing images are acquired using Biometrika sensor from LivDet 2013. This set of experiments captures the algorithm's invariance to the changes in environment for data collection. Table 7 presents the average classification error and Ferrfake @ Ferrlive = 1%. Results in Table 7 show that the proposed local patch based approach achieves a reduction of 38% in the average classification error from 30.7% in [24] to 18.9% in our approach. However, the average Ferrfake @ Ferrlive = 1% that we report is 66.2% and 69.4% for cross-sensor and cross-dataset scenarios respectively, indicating the challenges, especially in applications where a high level of spoof detection accuracy is needed.

3.3.5 Processing Times

The Inception-v3 CNN model takes around 4-6 hours to converge using a single Nvidia GTX Titan GPU utilizing approximately $96,000$ local patches for a training set with $2,000$ fingerprint images ($2,000$ images $\times$ 48 patches/fingerprint image). The average classification time for a single input image, including minutiae detection, local patch extraction, inference of Spoofness Scores for local patches, and producing the final decision, on a single Nvidia GTX Titan GPU is 800ms.

4. Conclusions

A robust and accurate method for fingerprint spoof detection is critical to ensure the reliability and security of the fingerprint authentication systems. In this study, we have utilized fingerprint domain knowledge by extracting local patches centered on minutiae locations in the input fingerprint image for training an Inception-v3 CNN model. The local patch based approach provides salient cues to differentiate spoof fingerprints from live fingerprints. The proposed approach is able to achieve a significant reduction in the error rates for intra-sensor (55%), cross-material (78%), cross-sensor (17%) as well as cross-dataset scenarios (38%) compared to state-of-the-art on public domain LivDet databases.

5. 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] S. Marcel, M. S. Nixon, and S. Z. Li, *Handbook of Biometric Anti-Spoofing*. Springer, 2014.

[2] ODNI, IARPA, "IARPA-BAA-16-04 (Thor)." https://www.iarpa.gov/index.php/research-programs/odin/odin-baa, 2016.

[3] International Standards Organization, "ISO/IEC 30107-1:2016, Information Technology—Biometric Presentation Attack Detection—Part 1: Framework." https://www.iso.org/standard/53227.html, 2016.

[4] T. Matsumoto, H. Matsumoto, K. Yamada, and S. Hoshino, "Impact of artificial gummy fingers on fingerprint systems," in *SPIE*, vol. 4677, pp. 275–289, 2012.

[5] K. Cao and A. K. Jain, "Hacking mobile phones using 2D Printed Fingerprints," *MSU Technical report, MSU-CSE-16-2*, 2016.

[6] S. S. Arora, K. Cao, A. K. Jain, and N. G. Paulter, "Design and Fabrication of 3D Fingerprint Targets," *IEEE Transactions on Information Forensics and Security*, vol. 11, no. 10, pp. 2284–2297, 2016.

[7] S. S. Arora, A. K. Jain, and N. G. Paulter, "Gold Fingers: 3D Targets for Evaluating Capacitive Readers," *to appear in the IEEE Transactions on Information Forensics and Security*, 2017.

[8] S. Yoon, J. Feng, and A. K. Jain, "Altered fingerprints: Analysis and detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 34, no. 3, pp. 451–464, 2012.

[9] E. Marasco and A. Ross, "A survey on antispoofing schemes for fingerprint recognition systems," *ACM Computing Surveys*, vol. 47, no. 2, p. 28, 2015.

[10] B. Biggio, Z. Akhtar, G. Fumera, G. L. Marcialis, and F. Roli, "Security evaluation of biometric authentication systems under real spoofing attacks," *IET Biometrics*, vol. 1, no. 1, pp. 11–24, 2012.

[11] D. Yambay, L. Ghiani, P. Denti, G. L. Marcialis, F. Roli, and S. Schuckers, "LivDet 2011-Fingerprint liveness detection competition 2011," in *5th IAPR International Conference on Biometrics*, pp. 208–215, 2012.

[12] V. Mura, L. Ghiani, G. L. Marcialis, F. Roli, D. A. Yambay, and S. A. Schuckers, "LivDet 2015 - Fingerprint liveness detection competition 2015," in *IEEE 7th International Conference on Biometrics Theory, Applications and Systems*, pp. 1–6, 2015.

[13] P. D. Lapsley, J. A. Lee, D. F. Pare Jr, and N. Hoffman, "Anti-fraud biometric scanner that accurately detects blood flow." US Patent 5,737,439, 1998.

[14] A. Antonelli, R. Cappelli, D. Maio, and D. Maltoni, "Fake finger detection by skin distortion analysis," *IEEE Transactions on Information Forensics and Security*, vol. 1, no. 3, pp. 360–373, 2006.

[15] D. Baldisserra, A. Franco, D. Maio, and D. Maltoni, "Fake fingerprint detection by odor analysis," in *International Conference on Biometrics*, pp. 265–272, Springer, 2006.

[16] R. K. Rowe and D. P. Sidlauskas, "Multispectral biometric sensor," Dec. 12 2006. US Patent 7,147,153.

[17] G. L. Marcialis, F. Roli, and A. Tidu, "Analysis of fingerprint pores for vitality detection," in *20th International Conference on Pattern Recognition*, pp. 1289–1292, 2010.

[18] E. Marasco and C. Sansone, "Combining perspiration-and morphology-based static features for fingerprint liveness detection," *Pattern Recognition Letters*, vol. 33, no. 9, pp. 1148–1156, 2012.

[19] L. Ghiani, G. L. Marcialis, and F. Roli, "Fingerprint liveness detection by local phase quantization," in *21st International Conference on Pattern Recognition*, pp. 537–540, 2012.

[20] L. Ghiani, A. Hadid, G. L. Marcialis, and F. Roli, "Fingerprint liveness detection using Binarized Statistical Image Features," in *IEEE Sixth International Conference on Biometrics: Theory, Applications and Systems*, pp. 1–6, 2013.

[21] D. Gragnaniello, G. Poggi, C. Sansone, and L. Verdoliva, "Fingerprint liveness detection based on Weber Local Image Descriptor," in *IEEE Workshop on Biometric Measurements and Systems for Security and Medical Applications*, pp. 46–50, 2013.

[22] D. Gragnaniello, G. Poggi, C. Sansone, and L. Verdoliva, "Local contrast phase descriptor for fingerprint liveness detection," *Pattern Recognition*, vol. 48, no. 4, pp. 1050–1058, 2015.

[23] D. Menotti, G. Chiachia, A. Pinto, W. R. Schwartz, H. Pedrini, A. X. Falcao, and A. Rocha, "Deep representations for iris, face, and fingerprint spoofing detection," *IEEE Transactions on Information Forensics and Security*, vol. 10, no. 4, pp. 864–879, 2015.

[24] 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*, vol. 11, no. 6, pp. 1206–1213, 2016.

[25] E. Marasco and C. Sansone, "On the Robustness of Fingerprint Liveness Detection Algorithms against New Materials used for Spoofing," in *International Conference on Bio-Inspired Systems and Signal Processing*, pp. 553–558, 2011.

[26] A. Rattani, W. J. Scheirer, and A. Ross, "Open set fingerprint spoof detection across novel fabrication materials," *IEEE Transactions on Information Forensics and Security*, vol. 10, no. 11, pp. 2447–2460, 2015.

[27] Y. Ding and A. Ross, "An ensemble of one-class SVMs for fingerprint spoof detection across different fabrication materials," in *IEEE International Workshop on Information Forensics and Security*, pp. 1–6, 2016.

[28] L. Ghiani, D. A. Yambay, V. Mura, G. L. Marcialis, F. Roli, and S. A. Schuckers, "Review of the Fingerprint Liveness Detection (LivDet) competition series: 2009 to 2015," *Image and Vision Computing*, vol. 58, pp. 110–128, 2017.

[29] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556*, 2014.

[30] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in *Advances in Neural Information Processing Systems*, pp. 1097–1105, 2012.

[31] K. Cao, E. Liu, L. Pang, J. Liang, and J. Tian, "Fingerprint matching by incorporating minutiae discriminability," in *International Joint Conference on Biometrics*, pp. 1–6, 2011.

[32] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, *et al.*, "Imagenet large scale visual recognition challenge," *International Journal of Computer Vision*, vol. 115, no. 3, pp. 211–252, 2015.

[33] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the Inception Architecture for Computer Vision," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2818–2826, 2016.

[34] L. Ghiani, D. Yambay, V. Mura, S. Tocco, G. L. Marcialis, F. Roli, and S. Schuckcrs, "LivDet 2013 Fingerprint Liveness Detection Competition 2013," in *International Conference on Biometrics*, pp. 1–6, 2013.

[35] D. Gragnaniello, G. Poggi, C. Sansone, and L. Verdoliva, "An investigation of local descriptors for biometric spoofing detection," *IEEE Transactions on Information Forensics and Security*, vol. 10, no. 4, pp. 849–863, 2015.