Y. Ding and A. Ross, "An Ensemble of One-Class SVMs for Fingerprint Spoof Detection Across Different Fabrication Materials," Proc. of 8th IEEE International Workshop on Information Forensics and Security (WIFS), (Abu Dhabi, UAE), December 2016

An Ensemble of One-Class SVMs for Fingerprint Spoof Detection Across Different Fabrication Materials

Yaohui Ding* and Arun Ross†
Michigan State University, East Lansing, MI 48824, USA
Email: *dingyaoh@msu.edu, †rossarun@cse.msu.edu

***Abstract*—A fingerprint recognition system is vulnerable to spoof attacks, where a fake fingerprint is used to circumvent the system. To counter such attacks, an automated spoof detector is used to distinguish images of fake fingerprints from those of real live fingerprints. Most spoof detectors adopt a machine learning approach, where a classifier is trained to distinguish between "spoof" and "live" samples. Such approaches require training samples from both classes. However, there are two fundamental concerns. Firstly, the number of spoof samples available during the training stage is typically much less than the number of live samples, resulting in imbalanced training sets. Secondly, the spoof detector may encounter spoofs fabricated using materials that were not "seen" in the training set, thereby failing to detect them. In order to alleviate some of these concerns, we adopt a One Class Support Vector Machine (OC-SVM) approach that predominantly uses training samples from only a single class, i.e., the live class, to generate a hypersphere that encompasses most of the live samples. The goal is to learn the concept of a "live" fingerprint. The boundary of the hypersphere is refined using a small number of spoof samples. The proposed method uses an ensemble of such OC-SVMs based on different feature sets. Experimental results on the LivDet2011 database show the advantages of the proposed ensemble of OC-SVMs for detecting spoofs generated from previously "unseen" materials.**

I. Introduction

Recent research has highlighted the vulnerability of biometric systems to "spoof attacks", commonly realized by presenting a falsified or altered biometric trait to the sensor [1], [2]. For instance, it has been shown that some fingerprint systems can be fooled by using a finger-like object made of easily available materials such as gelatine, woodglue or silicone that has the fingerprint ridges of another person inscribed on it [3].

In order to detect or deflect fingerprint spoof attacks, a number of sensor-based and image-based anti-spoofing solutions have been proposed [4], [5]. Image-based solutions, in particular, have received plenty of attention in the literature since they do not require the use of additional hardware and are based only on the images that are subsequently used by the fingerprint matcher. Such algorithms typically extract texture-based features [6], [7], anatomical features [8], [9] or physiological features [10], [11] from a fingerprint image (or sequence of images), and then train a binary classifier (such as a Support Vector Machine) that distinguishes the features of "Live" and "Spoof" samples.

However, there are some concerns associated with the use of binary classifiers in the context of spoof detection. In practice, it is easy to obtain training samples pertaining to the "Live" class but difficult to obtain samples for the "Spoof" class, thereby leading to imbalanced training sets where the latter has substantially fewer training samples. Further, the training set for the spoof class may not have data corresponding to all possible types of fabrication materials. This makes it difficult for the classifier to reliably learn the concept of a spoof. In fact, it has been shown that spoof detection accuracy degrades sharply, when the test set has fake samples fabricated using materials that were previously "unseen" in the training set (as reported in [12], [13]). As spoof attacks evolve, it is likely that new and more sophisticated materials will be used to create fake fingerprints thereby undermining existing learning-based spoof detectors.

To generalize the effectiveness of spoof detectors across fabrication materials - even those that are not encountered during training - recent work has formulated spoof detection as an open-set problem [14], [15]. Others utilize quality-based measures to minimize the impact of fabrication materials [16], [17]. While such methods have demonstrated success, they still require a large number of training samples from the spoof class. Menotti et al. [18] proposed a convolutional neural network (CNN) whose performance exceeded that of many fingerprint spoof detection benchmarks. However, just like other CNN-based methods, it requires a large number of training samples. Further, its robustness across fabrication materials was not evaluated.

The aforementioned concerns (related to interoperability across fabrication materials and limited spoof training samples) motivated us to consider approaching spoof detection as a *one-class* problem. The one-class classification paradigm differs from the multi-class paradigm in that only data from a single class (e.g., the live class) is used for training the classifier [19]. The task in one-class classification is to derive a decision boundary around samples of the live class that accepts as many samples as possible from that class while excluding other samples. Take the one-class support vector machine classifier as an example. The idea is to minimize the volume of the decision hypersphere containing the training data from a single class. However, this makes the problem harder than two-class classification because it is difficult to determine the

Y. Ding and A. Ross, "An Ensemble of One-Class SVMs for Fingerprint Spoof Detection Across Different Fabrication Materials,"
Proc. of 8th IEEE International Workshop on Information Forensics and Security (WIFS), (Abu Dhabi, UAE), December 2016

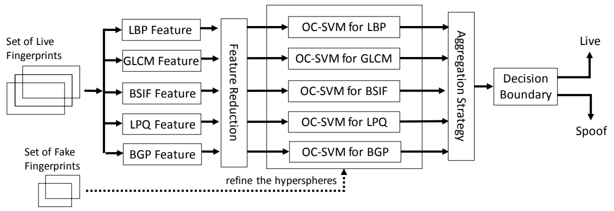


Fig. 1. Schematic of the proposed ensemble framework that uses multiple One Class Support Vector Machines (OC-SVMs). Each OC-SVM utilizes a different set of features. While spoof fingerprints are not necessary for training the OC-SVMs, they are used to refine the decision boundary in the validation phase.

tightness of the hypersphere enclosing the training samples. Moreover, it is difficult to determine what type of features extracted from a sample would effectively model the samples from the "Live" class.

In this paper, we propose a novel ensemble of multiple One-Class SVM (OC-SVM) classifiers to address the problem of spoof detection. Each OC-SVM uses a different feature set to find the smallest possible hypersphere around the majority of training samples pertaining to the "Live" class. The boundaries of the individual hyperspheres are further refined using a small number of spoof samples (as shown in Figure 1). Then, a Least Square Estimation (LSE) based weighting algorithm aggregates those independently trained OC-SVMs via a weighted linear combination scheme.

The experimental results show three significant advantages of the proposed ensemble of OC-SVMs: i) the detection accuracy is comparable with state-of-art spoof detection algorithms; however, only a smaller number of spoof samples is required for training; ii) the spoof detection accuracy remains consistent, regardless of what fabrication material is used, and iii) the detection accuracy can be further increased by utilizing a larger number of spoof fingerprint samples, without suffering from the imbalanced class problem encountered by conventional binary SVM (B-SVM) classifiers. In short, the classifier learns the concept of a "real live" fingerprint and uses this learned concept to reject spoofs.

This paper is organized as follows: Section II lists the different feature sets that are used in this work. Section III reviews the theory of a conventional one-class SVM classifier and Section IV proposes renovations that can effectively combine multiple OC-SVM classifiers to achieve an optimal decision boundary. Section V presents the experimental protocol and analyzes the results based on commonly used performance metrics. Section VI summarizes the findings of this work.

II. Features for Spoof Detection

As stated earlier, an image-based fingerprint spoof detector is a pattern classifier that is used to distinguish a live finger from a fake (spoof) artifact. Thus far, four fingerprint liveness detection competitions (LivDet) have been conducted between 2009 and 2015. From the results reported in these competitions [10], local texture-based features have been shown to outperform other competing techniques based on anatomical (such as pore detection [9]) or perspiration [20] features. Hence, the experiments in this work are conducted using local textural features, viz., Gray Level Co-occurrence Matrix (GLCM), Local Phase Quantization (LPQ), Binarized Statistical Image Features (BSIF), Local Binary Pattern (LBP) and Binary Gabor Pattern (BGP).

A detailed description of the above local textual features can be found in [17]. Briefly, Grey Level Co-occurrence Matrix (GLCM) characterizes the texture of an image by calculating the frequency of occurrence of pairs of pixels with specific values and in a specified spatial relationship; statistical measures are then extracted from this matrix [6]. Local Phase Quantization (LPQ) utilizes phase information computed locally in a window [7]. The phases of the four low-frequency coefficients are decorrelated and uniformly quantized. Binary Statistical Image Features (BSIF) encode texture information as a binary code for each pixel by linearly projecting local image patches onto a subspace, whose basis vectors are from natural images [21]. The Local Binary Pattern (LBP) operator compares a pixel with its neighbours, thresholds the ensuing results into a decimal value, and converts this value into a binary code [22]. Binary Gabor Patterns (BGP) encode textual information by convolving the image with Gabor filters and binarizing the responses [23]. Since each of the above texture descriptors are expected to capture different attributes of live and spoof samples, there is a need for fusing these descriptors and adapting them to generalize over multiple spoof materials.

III. One Class SVM (OC-SVM)

The one-class paradigm, also known as single-class classification or anomaly/novelty detection, is a learning scheme developed by Schölkopf et al [19]. One-class paradigm allows for the modeling of just a single class of patterns (e.g., real live fingerprints), and distinguishing them from all other possible patterns (e.g., spoof fingerprints fabricated by different materials). Tax and Duin [24] constructed a hypersphere with radius $R > 0$ and center a around the positive class data, which encompasses almost all points in the data set, while allowing for some samples to be excluded as outliers. This method is called Support Vector Data Description (SVDD), and the hypersphere formulation involves solving the following quadratic programming optimization problem:

$$\arg\min_{\mathbf{a},R,\xi} \left\{ R^2 + \frac{1}{N\nu}\sum_i \xi_i \right\}, \qquad (1)$$
$$\text{subject to} \quad ||\phi(\mathbf{x_i}) - \mathbf{a}||^2 \leq R^2 + \xi_i \quad \xi_i \geq 0.$$

Here, the training set is denoted as $\{\mathbf{x_i}\}, i = 1 \ldots N$, where $\mathbf{x_i}$ are column vectors. The term $\phi(\mathbf{x_i})$ is a non-linear mapping function that maps each input feature vector to a higher dimensional space. ν is a predefined regularisation parameter that governs the trade-off between the size of the hypersphere and the fraction of data points falling outside the hypersphere, i.e., the fraction of training examples that can be classified as

Y. Ding and A. Ross, "An Ensemble of One-Class SVMs for Fingerprint Spoof Detection Across Different Fabrication Materials,"
Proc. of 8th IEEE International Workshop on Information Forensics and Security (WIFS), (Abu Dhabi, UAE), December 2016

outliers. The ξ_i terms are the slack variables that allow some of the data points to lie outside the hypersphere. The Lagrange multipliers $\alpha_i \geq 0$ and $\gamma_i \geq 0$ are used to solve Eqn. (1):

$$L(\mathbf{a}, R, \xi, \alpha_i, \gamma_i) = R^2 + \frac{1}{N\nu}\sum_i \xi_i - \sum_i \alpha_i\{R^2 + \xi_i - (||\phi(\mathbf{x_i})||^2 - 2\mathbf{a}\cdot\phi(\mathbf{x_i}) + ||\mathbf{a}||^2)\} - \sum_i \gamma_i\xi_i.$$

L should be minimized with respect to $\mathbf{a}, R$ and ξ, and maximized with respect to α_i and γ_i. When L's partial derivatives w.r.t $\mathbf{a}$ and ξ_i are set to zero, it results in the following constraints:

$$\begin{aligned} \frac{\partial L}{\partial \mathbf{a}} &: \quad \mathbf{a} = \frac{\sum_i \alpha_i \mathbf{x_i}}{\sum_i \alpha_i} = \sum_i \alpha_i \mathbf{x_i}. \\ \frac{\partial L}{\partial \xi_i} &: \quad \frac{1}{N\nu} - \alpha_i - \gamma_i = 0. \end{aligned} \tag{2}$$

Eqn. (2) suggests that the center of the hypersphere is a linear combination of the input vectors. Further, because $\alpha_i \geq 0$ and $\gamma_i \geq 0$, the Lagrange multiplier γ_i can be removed when we require that $0 \leq a_i \leq \frac{1}{N\nu}$. As a result, the dual problem for Eqn. (1) can be written as:

$$\begin{gathered} \arg\max_{\alpha_i}\left\{\sum_i \alpha_i(\mathbf{x_i}\cdot\mathbf{x_i}) - \sum_{i,j}\alpha_i\alpha_j(\mathbf{x_i}\cdot\mathbf{x_j})\right\}, \\ \text{subject to} \quad 0 \leq \alpha_i \leq \frac{1}{N\nu}. \end{gathered} \tag{3}$$

When a training sample $\mathbf{x_i}$ satisfies the inequality $||\phi(\mathbf{x_i}) - \mathbf{a}||^2 < R^2 + \xi_i$, the constraint in Eqn. (3) is satisfied and the corresponding Lagrange multiplier α_i will be zero. For training samples that satisfy the equality $||\phi(\mathbf{x_i}) - \mathbf{a}||^2 = R^2 + \xi_i$, the constraints have to be enforced and the Lagrange multiplier will become greater than zero. This can be summarized as:

$$\begin{aligned} ||\phi(\mathbf{x_i}) - \mathbf{a}||^2 < R^2 + \xi_i \rightarrow &\quad \alpha_i = 0 \quad \text{(inlier)} \\ ||\phi(\mathbf{x_i}) - \mathbf{a}||^2 = R^2 + \xi_i \rightarrow &\quad 0 < \alpha_i < \tfrac{1}{N\nu} \quad \text{(border SVs)} \\ ||\phi(\mathbf{x_i}) - \mathbf{a}||^2 > R^2 + \xi_i \rightarrow &\quad \alpha_i = \tfrac{1}{N\nu} \quad \text{(outlier).} \end{aligned}$$

After the center $\mathbf{a}$ and the radius R of the hypersphere are deduced, a test sample $\mathbf{z}$ can be detected as an outlier, i.e., assigned to the spoof class, if its distance to the center of the hypersphere is greater than the radius:

$$||\phi(\mathbf{z})-\mathbf{a}||^2 = (\mathbf{z}\cdot\mathbf{z}) - 2\sum_i \alpha_i(\mathbf{z}\cdot\mathbf{x_i}) + \sum_{i,j}\alpha_i\alpha_j(\mathbf{x_i}\cdot\mathbf{x_j}) > R^2.$$

In this work, the LIBSVM package [25] (ver 3.18) was used to solve the above optimization problem.

IV. Ensemble of OC-SVM Classifiers

In the context of spoof detection, if the training data resides in a single feature space (e.g., LPQ feature space), the use of a single OC-SVM classifier can easily lead to overfitting problems. This is because the hypersphere attempts to tightly encompass live fingerprints and so a single feature space may not adequately capture the concept of a "live" class. To overcome this drawback, diversity is intuitively induced by combining several OC-SVMs that are based on descriptions of live fingerprint patterns in different feature spaces. Two different combination methods, the majority voting and the LSE-based weighting approach, are used here for combining the outputs of multiple OC-SVMs.

Majority voting is the simplest method for combining multiple classifiers. Multiple OC-SVMs, pertaining to different feature sets but derived from the same training samples, will result in multiple hyperspheres as decision functions, $f_j(\mathbf{x})$, $j = 1 \ldots L$. Here, L is the number of feature sets (OC-SVMs) considered. Let $\mathbf{y_i}$ denote the class label. While $\mathbf{y_i}$ is always $+1$ for the training data (i.e., the live class), $\hat{\mathbf{y}}_\mathbf{i}$, which denotes the output label of an OC-SVM classifier, could be -1 (i.e., the spoof class) or $+1$. Let $N_k(\mathbf{x}) = \sum_{j=1}^{L} \ (\hat{\mathbf{y}} = k|f_j(\mathbf{x}))$ where $k \in \{+1, -1\}$, denote the number of OC-SVMs that assign the input sample to the live or spoof class. Then the final decision of the OC-SVM ensemble via majority voting, $f_{MV}(\mathbf{z})$, for a test sample, $\mathbf{z}$, is determined by:

$$f_{MV}(\mathbf{z}) = \arg\max_k(N_k(\mathbf{z})) \quad k \in \{+1, -1\}. \tag{4}$$

An alternative to majority voting is the LSE-based weighting approach. The LSE-based weighting technique assigns different weights to individual OC-SVMs based on their classification accuracy. In the training phase, the weight vector $\mathbf{w}$ is estimated as $\hat{\mathbf{w}} = \mathbf{A}^{-1}\mathbf{y}$, where $\mathbf{A} = (f_j(\mathbf{x}_i))_{N\times L}$ consists of the estimated class label of each OC-SVM on training samples, and $\mathbf{y} = (\mathbf{y_i})_{N\times 1}$. The final decision of the OC-SVM ensemble for a given input sample $\mathbf{z}$ due to the LSE-based weighting is determined by:

$$f_{LSE}(\mathbf{z}) = sign\{\hat{\mathbf{w}}\cdot(f_j(\mathbf{z}))_{L\times 1}\}. \tag{5}$$

Since the performance of the LSE-based weighting approach was consistently better than the majority voting approach, only results from the LSE-based weighting are reported.

As stated earlier, one of the challenges in one-class classification is to determine how tightly the boundary should fit the training data. We propose two adjustments to the proposed ensemble OC-SVM scheme to address this concern. Firstly, the global regularisation parameter ν, that governs the trade-off between the radius of each hypersphere and the fraction of training data falling outside of the hypersphere, is gradually adjusted in the interval $[0.1\%, 10\%]$ in increments of 0.001. The LSE-based weights are also adjusted to optimize the detection accuracy during the training phase. In order to evaluate the detection accuracy, the Correct Detection Rate (CDR) on live fingers is defined as follows:

- CDR of "Live" fingers (CDR$_L$): the proportion of live samples that are correctly classified as "Live".

The rationale behind the adjustment is for the decision hypersphere to better fit the training data in every feature space rather than on a single feature space.

Secondly, the hypersphere is further refined by using a relatively small number of spoof fingerprints in a **validation**

Y. Ding and A. Ross, "An Ensemble of One-Class SVMs for Fingerprint Spoof Detection Across Different Fabrication Materials," Proc. of 8th IEEE International Workshop on Information Forensics and Security (WIFS), (Abu Dhabi, UAE), December 2016

phase. As discussed by Tax and Duin [24], the rationale behind the validation phase is to adjust the decision hypersphere to better classify the points that are in the vicinity of the hypersphere of any one of the L OC-SVMs by utilizing negative examples (spoof fingerprints). The available negative examples are labelled as outliers. Hence, they decrease the fraction of positive training samples that are classified as outliers, which leads to a readjustment of the global regularisation parameter ν. Hence, the following performance metric is defined to validate the detection accuracy on spoof fingerprints.

- CDR of "Spoof" fingers (CDR_S): the proportion of fake samples that are correctly classified as "Spoof".

V. Experimental Methodology

Datasets and Experimental Protocol: We used the LivDet2011 dataset for performance assessment of the proposed ensemble of one-class SVMs [12]. This database comprises images from 4 different sensors. Corresponding to each sensor, there are 1,000 live and 1,000 fake fingerprint samples in the training set, and the same number of samples, but from different subjects, in the test set. The fake fingerprints were fabricated using five materials, and for each material, 200 fake fingerprints were fabricated from 20 fingers. Five different kinds of texture descriptors were used in this work, and their dimensionalities were 40, 516, 256, 54 and 216 for GLCM, BSIF, LPQ, LBP and BGP, respectively. Since similar trends were observed across all 4 sensors, only results from the Biometrika sensor are reported in this paper.

In order to compare the proposed method against state-of-the-art spoof detection algorithms that exhibit interoperability across fabrication materials, the experimental protocol described in [12] is carefully followed in this work. The only difference is that the fake fingerprints in the training set were *not* used during the training phase of the proposed OC-SVMs. In order to analyze the impact of fabrication materials on spoof detection, Rattani and Ross [14] divided the test set of LivDet2011 dataset into two non-overlapping subsets according to the fabrication materials used. Each subset consists of 500 live samples and 500 fake fingerprints, where 200 fake fingerprints correspond to two fabrication materials that are used during the training stage (these are the "known" materials) and 300 fake fingerprints correspond to the other three fabrication materials that are *not* used during the training stage (these are the "novel" materials). As noted, there are ten possible combinations of known materials (and ten combinations of novel materials as well). In order to be consistent with [14], only nine of ten combinations have been evaluated in this work, using the exact same test sets as described. To show the advantage of the proposed ensemble OC-SVM on the detection of novel fabrication materials, CDR_S is intuitively divided into two parts:

- CDR of "Known" fake samples (CDR_K): the proportion of fake samples generated using known materials (i.e., materials encountered in the training set) that are correctly classified as "Spoof";

TABLE I
Establishing the **baseline performance** using conventional OC-SVM and B-SVM.

	Training Materials Only For B-SVM	Conventional B-SVM (OC-SVM) LBP	BGP	GLCM
1	Latex+EcoFlex	53.5 (28.5)	58.2 (40.4)	47.4 (40.2)
2	WoodGlue+Latex	58.2 (30.2)	60.0 (37.5)	55.0 (38.0)
3	Gelatine+Latex	53.5 (28.3)	55.0 (40.4)	55.7 (42.2)
4	Silgum+Latex	47.4 (27.5)	53.4 (33.3)	50.2 (30.2)
5	EcoFlex+Silgum	49.7 (33.9)	55.0 (33.3)	50.2 (28.9)
6	Gelatine+EcoFlex	40.0 (33.3)	50.2 (40.2)	47.0 (37.9)
7	Silgum+Gelatine	47.0 (38.0)	53.5 (37.5)	53.9 (40.4)
8	WoodGlue+Silgum	40.9 (33.3)	47.4 (38.0)	47.4 (42.2)
9	Gelatine+WoodGlue	47.9 (31.2)	47.9 (37.2)	50.2 (31.2)
	Average CDR_N	48.7 (31.6)	53.4 (37.5)	50.8 (36.8)

- CDR of "Novel" fake samples (CDR_N): the proportion of fake samples generated using novel materials (i.e., materials *not* encountered in the training set) that are correctly classified as "Spoof".

Experiment #1: Spoof detection accuracy of the conventional B-SVM and OC-SVM: This section evaluates the performance of conventional binary SVM (B-SVM) and conventional one-class SVM classifiers. This provides a baseline for the experiments in the subsequent sections. The training set used in this experiment consists of 400 live samples and 400 fake fingerprints made using two fabrication materials. Note that only B-SVM classifiers use fake fingerprints for training. In this experiment, no validation phase for the OC-SVM is implemented and the parameters of both classifiers are tuned following a conventional estimation procedure. Table I shows the correct detection rates on "novel" fake samples (CDR_N) using conventional B-SVM and conventional OC-SVM (in parentheses). Note that all the accuracy rates reported here are carried out on the exact same test set that was stated earlier.

It can be seen that both the conventional classifiers *do not* provide an acceptable correct detection accuracy on the fake samples manufactured using novel materials. The conventional OC-SVM classifier performed worse than the conventional binary SVM. However, we observed that the conventional OC-SVM provides higher correct detection rates on live fingers (CDR_L) than conventional B-SVM in some cases (results not shown here). These results are not surprising because the conventional OC-SVM is unable to find a tight enough decision boundary when using only the live fingerprints for training, leading to a higher CDR_L but a much lower CDR_N compared to B-SVM.

Experiment #2: Performance of proposed ensemble OC-SVM classifier: This section evaluates the performance of the ensemble OC-SVM classifier, especially on novel materials. To achieve a fair comparison, two variations of the conventional B-SVM were used as baselines:

- A feature-level fusion of B-SVM (referred to as B-SVM-F): The feature sets are concatenated into a single feature vector and the concatenated feature vector is used to train the conventional B-SVM and generate the binary outputs.

Y. Ding and A. Ross, "An Ensemble of One-Class SVMs for Fingerprint Spoof Detection Across Different Fabrication Materials," Proc. of 8th IEEE International Workshop on Information Forensics and Security (WIFS), (Abu Dhabi, UAE), December 2016

TABLE II
PERFORMANCE OF THE PROPOSED ENSEMBLE OF OC-SVMS COMPARED TO TWO B-SVM CLASSIFIERS AND THE AUTOMATIC ADAPTATION APPROACH IN [14]. THE CDRS BEFORE AND AFTER THE VALIDATION PHASE ARE REPORTED. SEE TEXT FOR EXPLANATION.

	Training Materials only for B-SVM	Proposed Ensemble of OC-SVMs			B-SVM-F			B-SVM-D			Automatic Adaptation	
		CDR_L	CDR_K	CDR_N	CDR_L	CDR_K	CDR_N	CDR_L	CDR_K	CDR_N	CDR_K	CDR_N
1	Latex+EcoFlex	90.4 - 0.6	83.1 + 2.4	82.7 + 2.4	76.5	77.3	63.8	74.3	71.8	77.2	92.9	83.6
2	WoodGlue+Latex	91.5 - 0.8	84.4 + 2.4	83.2 + 2.4	77.2	78.0	65.0	75.0	74.2	79.7	86.9	83.5
3	Gelatine+Latex	89.0 - 0.6	89.2 + 2.6	84.7 + 2.5	75.9	75.7	61.8	73.7	71.7	77.1	90.8	81.4
4	Silgum+Latex	83.9 - 0.4	79.4 + 2.2	77.9 + 2.3	69.1	69.8	61.6	66.8	69.3	74.5	91.7	81.5
5	EcoFlex+Silgum	90.8 - 0.4	86.6 + 2.3	84.6 + 2.3	77.0	77.8	67.7	74.8	76.1	81.9	85.8	75.5
6	Gelatine+EcoFlex	91.4 - 0.2	81.6 + 2.2	80.2 + 2.2	79.0	77.8	66.1	76.8	74.3	79.9	83.6	76.2
7	Silgum+Gelatine	90.5 - 0.4	84.0 + 1.9	84.0 + 2.1	76.7	77.5	66.2	71.5	75.6	81.3	86.3	82.0
8	WoodGlue+Silgum	91.4 - 0.4	84.2 + 2.4	84.2 + 2.3	77.1	77.9	65.9	74.9	74.2	79.7	84.8	78.4
9	Gelatine+WoodGlue	88.7 - 0.3	81.9 + 2.3	80.4 + 2.3	73.9	72.7	64.0	69.6	72.0	77.4	83.9	80.9
	Average CDR	89.7 - 0.5	$\mathbf{83.8 + 2.3}$	$\mathbf{82.4 + 2.3}$	75.8	76.1	64.7	73.0	73.2	78.7	**87.4**	**80.3**

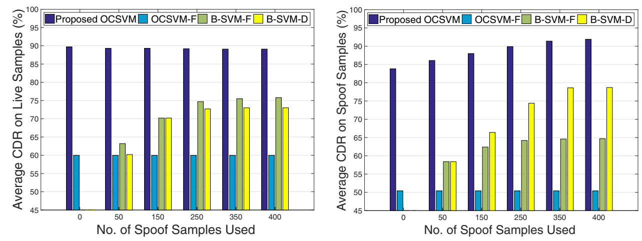
- A decision-level fusion of B-SVM (referred to as B-SVM-D): Several B-SVM classifiers are trained, and each of them is trained on a different feature set to generate binary outputs, then those outputs are combined using the majority vote rule.

Table II reports the performance of the proposed ensemble OC-SVM compared to an adaptive approach (referred to as Adaptive Detection) proposed earlier by Rattani and Ross [14], which was shown to significantly increase the correct detection rate on novel spoof materials (CDR_N).

As described earlier, the proposed ensemble OC-SVM utilizes the live fingerprint samples in the training set to generate the decision hypersphere. Although the spoof samples are not used by the learning procedure, they are used to readjust the decision boundary. In order to demonstrate the impact of this readjustment, the table reports the CDRs before and after the validation phase in the same cell. For example, the average CDR_N of the proposed OC-SVM is reported as $83.8 + 2.4\%$; this means the correct detection rate before the validation phase was 83.8%, and it increased by 2.4% after the validation. It must be noted that the number of fake samples used for validation is relatively small (50 spoof samples) compared to the larger training set (400 spoof samples) used by other approaches.

From Table II, it can be seen that the ensemble OC-SVM provides significantly higher correct detection rates than the other two SVM-based fusion schemes. One possible reason for the poor performance of the feature-level B-SVM (B-SVM-F) is the curse of dimensionality. A similar feature-level fusion was implemented for the conventional one-class SVM (OCSVM-F) as well. However, the poor performance (as shown in Figure 2) on both live samples (60.0%) and spoof samples (50.4%) indicates that multiple feature sets need to be aggregated more carefully to avoid potential issues such like the curse of dimensionality. It must be noted that the decision-level fusion of B-SVM results in an improvement in accuracy for detecting novel materials (as evidenced by the CDR_N for B-SVM-D). This result substantiates our previous conjecture that the use of different feature sets can better characterize the concept of "live" fingerprints to some extent.

The proposed ensemble OC-SVM is comparable with the best reported algorithm in LivDet2011 (89% CDR_L and 81% CDR_S on the Biometrika sensor as shown in [12]). However, it does not exceed the performance of the automatic adaptation approach [14], which has the best performance on the same database thus far. Further, while the accuracy of spoof detection was significantly improved (as evidenced by CDR_N and CDR_K), the use of the validation phase decreased the accuracy on detecting live samples (CDR_L reduced by 0.5%). We address both issues in the next experiment.



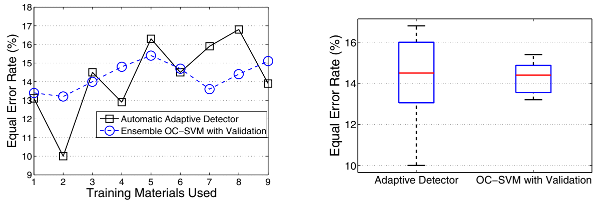
(a) (b)

Fig. 2. Performance of the ensemble OC-SVM after increasing the number of fake samples used in the validation phase.

Experiment #3: Improved performance of the ensemble OC-SVM using fake samples: This section evaluates the performance of the proposed ensemble OC-SVM by increasing the number of fake samples used in the validation phase. Although fake samples are not required for training the classifier, they can be used to improve the overall accuracy by tuning the decision hypersphere (i.e., the global regularisation parameter ν). Figure 2 presents two bar plots of the average CDR on live and spoof samples under this experimental design.

Figure 2(b) indicates that when increasing the number of fake samples in the validation phase (from 0 to 400), the proposed ensemble OC-SVM provides consistently higher CDRs on spoof samples than the binary SVM classifier with feature level fusion (B-SVM-F) and decision-level fusion (B-SVM-D). Figure 3(a) suggests that the proposed ensemble OC-SVM can provide similar detection rates as the state-of-art detector in [14], although the former only needs half the number of spoof samples as the latter (200/400). Moreover, the detection rates of the proposed method are more stable with a smaller standard deviation across different fabricated materials. This

Y. Ding and A. Ross, "An Ensemble of One-Class SVMs for Fingerprint Spoof Detection Across Different Fabrication Materials," Proc. of 8th IEEE International Workshop on Information Forensics and Security (WIFS), (Abu Dhabi, UAE), December 2016



(a) (b)

Fig. 3. Comparing the proposed method with the method in [14]. Equal Error Rate of the ensemble OC-SVM when 200 spoof samples are used in validation phase. Training materials used here are as same as in Table I and II.

suggests that the proposed method is not unduly impacted by the choice of fabrication material used for generating the spoof fingerprint. The average CDRs on live samples are presented in Figure 2(a). Similar to the results in Table II, the CDR$_L$ marginally decreased by 0.5% to 0.8% when the number of fake samples is increased during validation. This demonstrates the trade-off between the misclassification of live samples and the size of the decision hypersphere. However, compared to the performance gain on spoof detection (an increase from 83.0% to 89.7%), the modest degradation in CDR$_L$ is acceptable.

VI. Summary

In this work, the problem of spoof detection is posed as a one-class problem where the classifier learns the concept of a "live" fingerprint sample and uses this to reject spoof samples. It was shown that the accuracy of a conventional one-class SVM (OC-SVM) could be significantly improved by fusing multiple kinds of features and optimizing the decision functions across these features. Experimental analysis conducted on the LivDet2011 database show that the proposed ensemble OC-SVM outperforms Binary SVMs, and its performance is comparable with state-of-art spoof detection algorithms that are interoperable across fabrication materials. However, the proposed method requires much fewer spoof training samples than competing techniques. Further, the performance of the proposed method is observed to be stable across different fabrication materials. Thus, the proposed approach successfully mitigates some of the concerns associated with the issue of "imbalanced training sets" and "insufficient spoof samples" encountered by conventional spoof detection algorithms.

Acknowledgment

This project was supported by the NSF Center for Identification Technology Research (CITeR).

References

[1] S. Schuckers, "Spoofing and anti-spoofing measures," *Information Security Technical Report*, vol. 7, no. 4, pp. 56–62, 2002.

[2] B. Toth, "Biometric liveness detection," *Information Security Bulletin*, vol. 10, no. 8, pp. 291–297, 2005.

[3] T. Matsumoto, H. Matsumoto, K. Yamada, and S. Hoshino, "Impact of artificial gummy fingers on fingerprint systems," *Proceedings of Society of Photographic Instrumentation Engineers (SPIE)*, Jan. 2002.

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

[5] V. Mura, L. Ghiani, G. L. Marcialis, F. Roli, D. A. Yambay, and S. A. Schuckers, "Livdet 2015 fingerprint liveness detection competition," in *IEEE 7th International Conference on Biometrics: Theory, Applications and Systems (BTAS)*, 2015, pp. 1–6.

[6] S. Nikam and S. Aggarwal, "Wavelet energy signature and glcm features-based fingerprint anti-spoofing," in *IEEE Int. Conf. On Wavelet Analysis and Pattern Recognition*, Hong Kong, 2008, pp. 717–723.

[7] L. Ghiani, G. Marcialis, and F. Roli, "Fingerprint liveness detection by local phase quantization," in *21st Int. Conf. On Pattern Recognition (ICPR)*, 2012, pp. 537–540.

[8] M. Espinoza and C. Champod, "Using the number of pores on fingerprint images to detect spoofing attacks," in *Intl. Conf. on Hand-based Biometrics*, Hong Kong, China, 2011, pp. 1–5.

[9] G. Marcialis, F. Roli, and A. Tidu, "Analysis of fingerprint pores for vitality detection," in *Proc. of Intl. Conf. on Pattern Recognition (ICPR)*, Aug 2010, pp. 1289–1292.

[10] L. Ghiani, P. Denti, and G. L. Marcialis, "Experimental results on fingerprint liveness detection," in *IEEE Intl. Conf. on Articulated Motion and Deformable Objects*, Spain, 2012, pp. 210–218.

[11] E. Marasco and C. Sansone, "Combining perspiration and morphology based static features for fingerprint liveness detection," *Pattern Recognition Letters*, vol. 33, pp. 1148–1156, 2012.

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

[13] L. Ghiani, D. Yambay, V. Mura, S. Tocco, G. L. Marcialis, F. Roli, and S. Schuckcrs, "Livdet 2013 fingerprint liveness detection competition," in *International Conference on Biometrics (ICB)*, 2013, pp. 1–6.

[14] A. Rattani and A. Ross, "Automatic adaptation of fingerprint liveness detector to new spoof materials," in *International Joint Conference on Biometrics (IJCB)*, 2014, pp. 1–8.

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

[16] J. Galbally, F. Alonso-Fernandez, J. Fierrez, and J. Ortega-Garcia, "A high performance fingerprint liveness detection method based on quality related features," *Future Generation Computer Systems*, vol. 28, no. 1, pp. 311–321, 2012.

[17] Y. Ding, A. Rattani, and A. Ross, "Bayesian belief models for integrating match scores with liveness and quality measures in a fingerprint verification system," in *International Conference on Biometrics (ICB)*, 2016, pp. 1–8.

[18] 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.

[19] B. Schölkopf, R. C. Williamson, A. J. Smola, J. Shawe-Taylor, J. C. Platt *et al.*, "Support vector method for novelty detection," in *The Annual Conference on Neural Information Processing Systems (NIPS)*, vol. 12, 1999, pp. 582–588.

[20] A. Abhyankar and S. Schuckers, "Integrating a wavelet based perspiration liveness check with fingerprint recognition," *Pattern Recognition*, vol. 42, pp. 452–464, 2009.

[21] L. Ghiani, A. Hadid, G. Marcialis, and F. Roli, "Fingerprint liveness detection using binarized statistical image features," in *IEEE 6th International Conference on Biometrics: Theory, Applications and Systems (BTAS)*, 2013, pp. 1–6.

[22] S. Nikam and S. Aggarwal, "Local binary pattern and wavelet-based spoof fingerprint detection," *Intl. Journal of Biometrics*, vol. 1, no. 2, pp. 141–159, 2008.

[23] L. Zhang, Z. Zhou, and H. Li, "Binary gabor pattern: An efficient and robust descriptor for texture classification," in *19th International Conference on Image Processing (ICIP)*, 2012, pp. 81–84.

[24] D. M. Tax and R. P. Duin, "Support vector data description," *Machine learning*, vol. 54, no. 1, pp. 45–66, 2004.

[25] C. C. Chang and C. J. Lin, "LIBSVM: A library for support vector machines," *ACM Transactions on Intelligent Systems and Technology*, vol. 2, pp. 1–27, 2011, software available at http://www.csie.ntu.edu.tw/cjlin/libsvm.