Custom silicone Face Masks: Vulnerability of Commercial Face Recognition Systems & Presentation Attack Detection

Raghavendra Ramachandra†; Sushma Venkatesh†; Kiran B. Raja†;
Sushil Bhattacharjee‡; Pankaj Wasnik†; Sebastien Marcel‡, Christoph Busch†
† Norwegian Biometrics Laboratory, NTNU-Norway
‡ Idiap Research Institute, Martigny, Switzerland

***Abstract*—The vulnerability of face recognition systems towards evolving presentation attacks has drawn significant interest in the last decade. In this paper, we present an empirical study on both vulnerability analysis and presentation attack detection for commercial face recognition systems (FRS) using custom 3D silicone face masks corresponding to real subjects. To this end, a new database is collected consisting of 8 custom 3D silicone masks together with *bona fide* presentations of the corresponding subjects using three different devices (smartphones). The vulnerability of FRS for 3D face silicone face masks is effectively evaluated using two well-known commercial-off-the-shelf (COTS) FRS (Verilook from Neurotechnology, and the Cognitec Face-VACS). Further, extensive experiments are carried out to evaluate the effectiveness of five state-of-the-art presentation attack detection (PAD) techniques for detecting such masks. Key insights on silicone mask PAD are provided along with a discussion on the accuracy achieved in our experiments.**

I. Introduction

Face presentation attack detection (PAD) has received significant interest from the biometrics community because of the vulnerability of existing Face Recognition Systems (FRS) in differentiating these attacks from bona fide presentations. The ease of creating face Presentation Attack artefacts has been well demonstrated in the literature and further, the use of simple Presentation Attack Instruments (PAI) such as printed attacks, display and wrap photo attacks has shown remarkable attack potential on the FRS. To effectively address this problem several Presentation Attack Detection (PAD) algorithms have been developed that are largely based on differentiating the texture information [13]. However, creation of realistic and textured masks have challenged existing FRS and the PAD schemes thereby as shown in earlier works [3], [13].

Among the different types of Presentation Attack Instruments (PAIs), the 3D face mask based attacks are highly sophisticated due to close-to-real appearance[3]. Arguably, the detection challenge increases proportionally with very good quality 3D face masks. Early work in this direction has addressed the detection of PAs based on 3D rigid masks collected using the Kinect device [6]. It is demonstrated in [12] that, the use of the depth information together with the texture extraction methods can successfully detect 3D rigid mask PAs. Agarwal et al. [1] have proposed a method for detecting PAs based on generic latex face masks, using

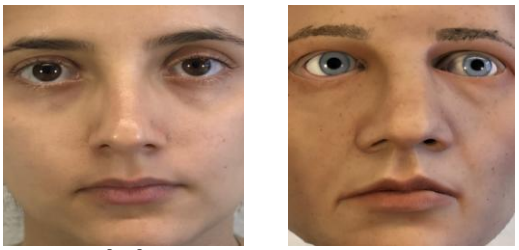


Fig. 1: Illustration of the example 3D face silicone mask image from the newly collected dataset (a) *Bona fide* presentation (b) Corresponding 3D silicone face mask attack presentation

texture-based approaches on the multi-spectral imagery [15]. Note that their work relies on only three bands, visible-light (VIS), Thermal and near-infrared (NIR). More specifically, (i) it does not establish the vulnerability of FRS to latex mask based impersonation attacks, and (ii) it does not address the challenges from custom masks (i.e., corresponding masks of actual/real data subjects). Another study on detecting silicone face masks [10] relies on data collected from different Youtube videos. This study also does not address the vulnerability of FRS to silicone mask attacks, as the masks used in these videos have not been custom-made, to impersonate specific identities. A preliminary study on the use of extended-range imagery for detecting custom 3D-mask based attacks [2] showed that thermal (long-wave infrared (LWIR)) imagery could be effective in detecting custom-mask based attacks. A recent work [9] proposes a Convolutional Neural Network (CNN) based PAD method, to detect full head 3D mask PAs, using bi-spectral imagery (VIS and NIR). Unlike previous studies, this study [9] employs only a single mask, presented by different subjects. We note again, that this work does not include any vulnerability analysis of FRS, as *bona fide* presentations of the data subject corresponding to the mask have not been collected.

The first systematic study on the threat posed by custom 3D silicone masks using deep learning (CNN) based FRS is presented in [3]. The work uses a dataset based on six custom

silicone masks, with presentations captured using two different cameras: (1) Realsense SR300 for VIS and NIR images, and (2) Seek Thermal Compact-Pro camera to capture the thermal (LWIR) images. Using the VIS images, vulnerability analysis is reported for three different CNN based FRS, namely, VGG-Face [11], FaceNet [14], and LightCNN [17]. It is demonstrated that all three FRS are vulnerable to custom 3D silicone face mask based PAs. Further, [4] also proposes a silicone mask PAD method using thermal images. The authors report an *a posteriori* equal-error rate (D-EER) of 7.5%, using the mean thermal intensity in the face-region as the discriminating feature between the two classes of presentations. This method, however, may not be very robust, because, as noted in [2], the temperature of a silicone mask can rise significantly when it is worn by the attacker for a substantial period of time. In summary, we note that the work reported in [4] is limited to discussion of the vulnerability of deep CNN FRS, which are purely academic solutions. *To the best of our knowledge, there exists no published study on the vulnerability of widely deployed commercial FRS to such attacks. Moreover, there exists no baseline evaluation of the current PAD techniques to provide the benchmark on the detection accuracy for the custom 3D silicone face mask.*

A. Contributions of our work

In this paper we present a study of the vulnerability of two well known and widely deployed FRS, namely, *Neurotechnology Verilook 10.0* and *Cognitec Face-VACS 9.1.4*, to custom 3D silicone face mask based PAs. Further, we benchmark the performance of five popular PAD techniques. The main contributions of this work are as follows.

1) A new dataset involving custom silicone face masks for eight data subjects. *Bona fide* presentations for the corresponding subjects, as well as PAs have been collected using three different smartphones, iPhone X, Samsung S7 and Samsung S8. To the best of our knowledge, this is the largest custom silicone mask dataset (compared to earlier works) collected so far.
2) Vulnerability analysis of two well known commercial FRS to custom silicone face masks, using the new dataset.
3) Benchmarks the performance of five commonly used PAD algorithms on the newly collected dataset.

The rest of the paper is organised as follows: Section II discusses the data collection procedure and the statistics of the newly collected dataset. Section III presents both quantitative and qualitative experiments on the collected dataset, and Section IV summarizes the conclusions drawn from the experiments discussed here.

II. Silicone Mask dataset

In this work, we have created a new database of images captured from custom 3D face silicone masks corresponding to eight subjects and captured using three different smartphones. We refer this database as **C**ustom **S**ilicon **M**ask **a**ttack **d**atabase - **M**obile (CSMad-Mobile) dataset [1]. The masks, each costing about USD 4000, have been manufactured by a professional special-effects company. The process of silicone mask generation is the same as described in [4]. For the *bona fide* presentations of the same eight subjects, each data subject is asked to pose in a manner compliant to standard portrait capture. The data is captured indoors, with adequate artificial lighting. Silicone mask presentations have been captured under similar conditions, by placing the masks on their bespoke support provided by the manufacturer, with prosthetic eyes and silicone eye sockets.

Both *bona fide* presentations and mask PAs have been captured using the rear cameras of three smartphones: (1) iPhone X, (2) Samsung S7 and (3) Samsung S8. All three cameras have a resolution of 12 Mega-pixels. *The use of smartphones in this work is motivated by the real-life application of the PAs in unsupervised applications such as banking applications where such mask based attacks can be carried out easily.* Using each smartphone, we have collected 108 *bona fide* samples and 155 mask PA samples, resulting in a dataset with $(108 + 155) \times 3 = 789$ samples. Figure 2 shows examples of *bona fide* and PAs collected in this work using three different smartphones for all eight subjects.

To evaluate PAD methods we divide the dataset into two disjoint partitions for each smartphone. One partition is used as the training set, and the other is used as the testing set. Table I summarizes the statistics of training and testing partitions.

III. Experiments and results

In this section, we present both vulnerability analysis on two different commercial FRS and the performance of the PAD techniques on the newly collected custom silicone mask attack dataset.

A. Vulnerability analysis

First we evaluate the threat posed by custom silicone masks to two commercial FRS – Neurotechnology Verilook 10.0 and Cognitec FaceVACS 9.1.4 [2]. Both these FRS are already deployed in numerous real-life face recognition applications and are also tested by the Face Recognition Vendor Test of the National Institute of Standards and Technology (NIST) [7]. The objective of the vulnerability analysis is to determine the probability of a custom silicone mask to be accepted as the corresponding genuine subject. According to [8], the vulnerability (or attack success rate) of the biometrics systems under attacks can be quantified using the metric 'Imposter Attack Presentation Match Rate (IAPMR)', defined as *the proportion of the impostor attack presentations using the same Attack Instrument species in which the target reference is matched in a full-system evaluation of a verification system* [8]. Higher values of IAPMR indicate a more vulnerable FRS.

While evaluating the commercial face SDKs (software development kits), we enrol one image corresponding to the *bona*

[1]The dataset can be availed www.idiap.ch/dataset/csmad-mobile

[2]The experiments are conducted on Cognitec FaceVACS SDK directly and our results does not necessarily constitute Cognitec's best effort results.

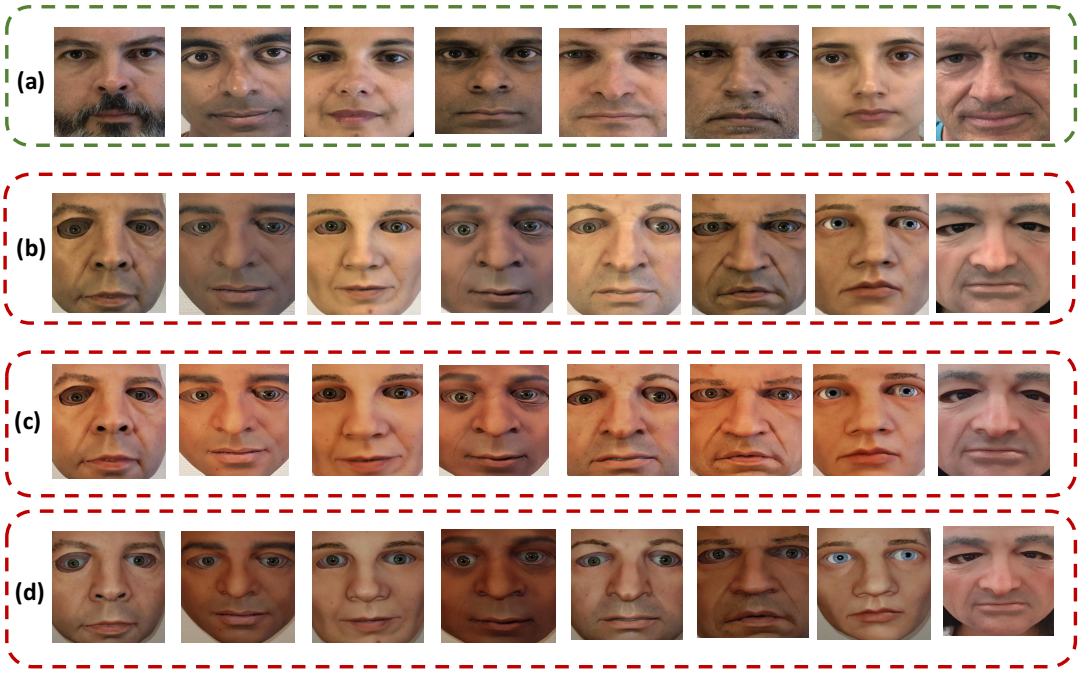


Fig. 2: Illustration of the sample data from the eight subjects used in this work. (a) *Bona fide* presentations. Corresponding 3D silicone face mask attack presentations captured using (b) iPhone X, (c) Samsung S8, and (d) Samsung S7.

TABLE I: Statistics of the collected dataset and protocol to evaluate PAD techniques.

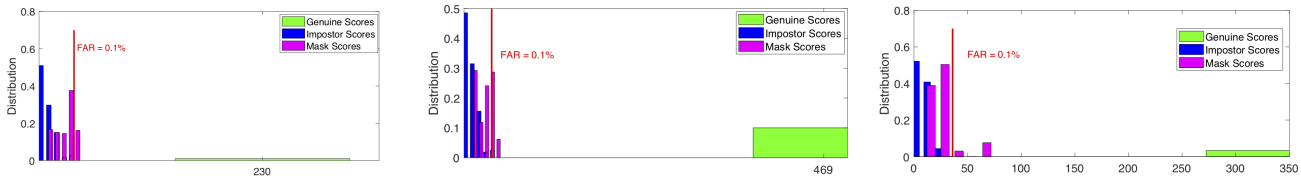
Devices	Training images		Testing images	
	No. of *Bona fide*	No. of PAs	No. of *Bona fide*	No. of artefact
iPhone	50	50	58	105
Samsung S7	50	50	58	105
Samsung S8	50	50	58	105

fide to obtain the comparison scores. The data is captured from three different smartphones in separate sessions. Under this setting, we have obtained 99 genuine scores, 749 zero-effort impostor (ZEI) scores and 247 scores for silicone mask PAs for iPhone X. In the case of the Samsung S7 dataset, there are 88 genuine scores, 679 ZEI scores and 181 mask PA scores. For the Samsung $S8$ dataset, there are 100 genuine scores, 756 ZEI scores and 195 PA scores.

Figure 3 shows the verification score distributions produced by the Neurotechnology Verilook FRS corresponding to iPhone X (Fig. 3a), Samsung S8 (Fig. 3b) and Samsung S7 (Fig. 3c). For each device, the threshold corresponding to (false-accept rate) FAR = 0.1% is also indicated in the corresponding plot in Figure 3 by a red vertical straight-line. Note that, for all three devices, (1) the scores for genuine and ZEI classes are well separated, (2) the mask PA score-distribution (magenta) significantly overlaps the ZEI score distribution (depicted in blue). Table II summarizes the vulnerability of the two commercial FRS in quantitative terms (IAPMR) for two different thresholds, set at FAR = 0.1% and FAR = 0.01% [3]. These score-threshold values have been selected based on the recommendations from the respective manufacturers of the two commercial FRS. The main observations from Table II are:

- In general, commercial FRS are vulnerable to 3D silicone mask PAs.
- The vulnerability is noted only at the higher FAR thresholds. By setting the threshold with lower values of FAR, for example, at $FAR = 0.01\%$ the FRS is not vulnerable to the 3D silicone mask PAs.
- Comparing the two commercial FRS, the Neurotechnology Verilook 10.0 FRS indicates higher IAPMR values, that is, it is more vulnerable than Cognitec's Face-VACS 9.1.4.
- The highest vulnerability is noted for data from the Samsung S8 phone with Neurotechnology Verilook FRS resulting in IAPMR = 28.20 @ FAR = 0.1%, while the lowest vulnerability is indicated by the Cognitec Face-VACS 9.1.4 on Samsung S7 images, with IAPMR =

[3]Threshold obtained from respective vendors without fine-tuning on CSMad-Mobile dataset.



(a) iPhone X dataset (b) Samsung S8 dataset (c) Samsung S7 dataset

Fig. 3: Verfication-score distributions from Neurotechnology Verilook 10.0 FRS (comparison scores on x-axis)

TABLE II: Quantitative results of vulnerability analysis for two commercial face recognition systems (FRS).

Commercial FRS	Devices	Threshold @ FAR = 0.1%		Threshold @ FAR = 0.01%	
		FRR (%)	IAPMR(%)	FRR (%)	IAPMR(%)
Neurotech	iPhone	0	10.12	0	0
	Samsung S7	0	12.21	0	0
	Samsung S8	0	28.20	0	0
Cognitec	iPhone	0	19.02	0	0
	Samsung S7	0	3.31	0	0
	Samsung S8	0	20.51	0	0

3.31% @ FAR = 0.1%.

- Despite the moderate number of attack sample size (593 attempts from different subjects), we note that FRS are quite vulnerable to such PAs.

B. Evaluation of baseline PAD techniques

In this section, we present the performance of five state-of-the-art PAD algorithms on the 3D silicone face mask dataset. Each PAD technique involves feature-extraction followed by classification based on a supervised two-class classifier. Here, we have used the two-class support-vector machine (SVM) classifier for all our experiments. The state-of-the-art feature-extraction algorithms evaluated in this study, all characterize image-texture information using: Local Binary Patterns (LBP) [6], Binarized Statistical Image features (BSIF) [13], Local Phase Quantization (LPQ) [13], Image Distortion Analysis (IDA) [16] and Colour texture [5].

The performance of the PAD techniques are presented following the IS0/IEC 30107-3 [8] metrics with Bona fide *Presentation Classification Error Rate (BPCER) and Attack Presentation Classification Error Rate (APCER)*. **BPCER** is defined as the proportion of *bona fide* presentations incorrectly classified as PAs whereas **APCER** is defined as the proportion of PAs incorrectly classified as *bona fide* presentations. In particular, for each method, we report the BPCER while fixing the APCER at 5%, and at 10%, following the IS0/IEC 30107-3 [8] recommendations. In addition, we also present the results in terms of Detection-Equal Error Rate (D-EER) (%) obtained on the testing set.

In this work we perform two different experiments to evaluate the effectiveness of the chosen PAD techniques. The first experiment (Experiment-I) measures the performance of the PAD techniques when data from the same device is used for both training and testing. The second experiment (Experiment-II) evaluates the performance of the PAD methods in cross-device scenarios, where a PAD technique is trained using the data from one device and tested with data from the remaining two devices.

1) Experiment I: Table III shows the quantitative results of the five different state-of-the-art methods. The key observations are:

- State-of-the-art features like LBP, IDA and color textures perform well, showing lower values of D-EER.
- Among the five techniques, the LBP-SVM configuration shows the best results with D-EER = 0% and BPCER = 0 @APCER = 5% & 10% respectively across all three devices.
- In general, the five PAD techniques all lead to similar PAD rates across the three devices. However, closer observation reveals that attacks in the iPhone X images are more difficult to detect than attacks on Samsung S8 and Samsung S7 phones.
- It is also interesting to observe that the use of color texture based PAD techniques [5] has indicated the second best performance on both iPhone X and Samsung S8 images. However, for data from the Samsung S7 device, the performance is the same as that of LBP-SVM PAD technique.
- We attribute the deviations in IAPMR obtained for different devices mainly to variations in data-capture conditions and in the optics of the smartphones.

Based on the limited analysis provided above, it can be asserted that the 3D silicone mask attacks with artificial eyes and mounted on bespoke support can be effectively detected using LBP-SVM method.

2) Experiment-II: In this section we discuss the performance of the state-of-the-art PAD techniques in cross-device tests. Here, the PAD techniques are trained using images captured from one device and tested using images captured

TABLE III: Experiment I: Quantitative performance of the PAD techniques.

Devices	Algorithms	D-EER(%)	BPCER @ APCER =	
		Testing data	5%	10%
iPhoneX	**LBP-SVM**	**0**	**0**	**0**
	LPQ-SVM	20.34	24.13	22.41
	BSIF-SVM	13.56	17.24	15.51
	IDA-SVM	6.78	12.06	3.44
	Color Textures-SVM	3.62	3.44	1.72
Samsung-S8	**LBP-SVM**	**0**	**0**	**0**
	LPQ-SVM	12.22	12.06	12.06
	BSIF-SVM	12.22	13.79	12.06
	IDA-SVM	6.78	29.31	0
	Color Textures-SVM	6.78	77.58	0
Samsung-S7	**LBP-SVM**	**0**	**0**	**0**
	LPQ-SVM	13.56	13.79	13.79
	BSIF-SVM	13.56	13.79	13.79
	IDA-SVM	**0**	**0**	**0**
	Color Textures-SVM	4.96	5.17	5.17

using remaining two devices. This experiment evaluates the generalization of the PAD techniques across different capture devices.

Table IV shows the quantitative performance of the five PAD techniques. From the obtained results, following observations can be made:

- In general, the performance of the five PAD techniques degrades in the cross-device scenario, compared to protocol of Experiment-I (See Section III-B1). The drop in performance can be attributed to the variations in image statistics among the three different devices.
- Among the three different evaluation scenarios summarized in the Table IV, the scenario where data from Samsung S7 is used for training, and the test data comes from the remaining two devices produces the worst performance.
- In general, the poor performance in cross-device tests indicates a challenge in detecting silicone mask based PAs with state-of-the-art PAD methods when the sources of training and probe images are different.

We note that there is a necessity for generalizability studies of such LBP-SVM method for detecting the attacks in cross-database scenario.

3) Discussion and future work: The threat of PAs using high quality silicone masks is demonstrated in this work with the vulnerability analysis of two commercial FRS. The FRS included in this study appear to be less vulnerable to PAs based on custom silicone masks than the academic FRS discussed in a previous work [4]. The commercial off-the-shelf FRS even exhibit stronger resistance to attacks at lower thresholds of FAR. Although the experiments present in this work indicate that these two FRS have moderate vulnerability to silicone mask PAs, a more detailed analysis on varying conditions for presentations, such as attack by wearing the masks, attacks by concealing the discontinuities around eye region and so on, need to be evaluated to establish the true robustness of the FRS against mask based PAs. Further, the experiments with the dataset collected indicate high PAD accuracy in the intra-device scenario (where training and testing data come from the same device). The performance degrades when the training and testing data varies or corresponds to different capture devices. Although our preliminary conjecture is that the reliance on image-texture information plays a role in lowering the performance in cross-device experiments, more detailed investigation in this direction needs to carried out.

IV. Conclusion

The vulnerability analysis of two different commercial FRS towards custom 3D silicone face masks is presented in this work. For this purpose we have collected a new dataset based on eight subjects. For each subject a custom silicone mask has been created. Data for mask-presentations as well as *bona fide* presentations of the eight subjects has been captured using three different smartphones. Our experiments indicate that the two commercial FRS are vulnerable to PAs based on custom 3D silicone face masks, especially when operating threshold corresponds to higher values of FAR. When the threshold is set at the lower values of FAR (*e.g.*, FAR = 0.01%), both commercial FRS are not vulnerable to the custom silicone mask PAs.

We have also presented an extensive evaluation of five different state-of-the-art techniques to benchmark the silicone face mask presentation attack detection. Our experiments indicate outstanding detection accuracy of the LBP-SVM method with D-EER = 0%, BPCER = 0% @ APCER = 10% and 5%, when training and testing images come from

TABLE IV: Experiment II: Quantitative performance of the PAD techniques

Train data	Development and testing data	Algorithms	D-EER(%)	BPCER @ APCER =	
			Testing Data	5%	10%
iPhoneX	Samsung S8 & Samsung S4	LBP-SVM	24.66	51.71	43.34
		LPQ-SVM	28.68	53.23	45.62
		BSIF-SVM	25.43	53.99	45.24
		IDA-SVM	45.89	61.21	59.31
		Color Textures-SVM	26.01	50.92	43.51
Samsung S8	iPhone X & Samsung S4	LBP-SVM	28.1	39.16	36.88
		LPQ-SVM	29.63	39.54	37.26
		BSIF-SVM	35.75	49.8	49.04
		IDA-SVM	23.51	41.44	28.13
		Color Textures-SVM	28.31	91.66	74.53
Samsung S7	iPhone X & Samsung S8	LBP-SVM	40.68	94.29	81.36
		LPQ-SVM	38.47	75.66	61.97
		BSIF-SVM	42.39	77.94	66.53
		IDA-SVM	32.69	76.42	64.25
		Color Textures-SVM	40.85	95.37	93.51

the same device (same smartphone). However, when the PAD systems are trained with the images from one device and subsequently tested with the images captured with other devices, the experimental results have indicated a significant drop in detection performance for all five state-of-the-art PAD techniques studied in this work. This highlights the challenge in detecting PAs based on 3D silicone face masks when the source of image capture is not known to the PAD system. Future work in this direction will focus on developing new PAD algorithms, especially for the cross-dataset (or cross-device) scenarios. A dataset with larger set of PAs needs to be created for further validation of initial observations made in this work.

ACKNOWLEDGEMENT

We acknowledge the support of the Research Council of Norway under Grant No. IKTPLUSS 248030/O70 (SWAN project), the European H2020 ICT project TeSLA (grant agreement no. 688520), and of the Swiss Center for Biometrics Research and Testing, for making this work possible.

REFERENCES

[1] A. Agarwal, D. Yadav, N. Kohli, R. Singh, M. Vatsa, and A. Noore. Face presentation attack with latex masks in multispectral videos. In *IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW-2017)*, pages 275–283, 2017.

[2] S. Bhattacharjee and S. Marcel. What you can't see can help you – extended-range imaging for 3d-mask presentation attack detection. In *Proceedings of the 16th International Conference on Biometrics Special Interest Group*. Gesellschaft fuer Informatik e.V. (GI), 2017.

[3] S. Bhattacharjee, A. Mohammadi, and S. Marcel. Spoofing deep face recognition with custom silicone masks. In *IEEE Ninth International Conference on Biometrics: Theory, Applications and Systems (BTAS)*, pages 1–8, 2018.

[4] S. Bhattacharjee, A. Mohammadi, and S. Marcel. Spoofing deep face recognition with custom silicone masks. In *Proceedings of BTAS-2018*, 2018.

[5] Z. Boulkenafet, J. Komulainen, and A. Hadid. Face spoofing detection using colour texture analysis. *IEEE Transactions on Information Forensics and Security*, 11(8):1818–1830, 2016.

[6] N. Erdogmus and S. Marcel. Spoofing attacks to 2d face recognition systems with 3d masks. In *International Conference of the Biometrics Special Interest Group*, 2013.

[7] P. J. Grother, M. L. Ngan, and K. K. Hanaoka. Ongoing face recognition vendor test (frvt) part 2: Identification. Technical report, 2018.

[8] ISO/IEC JTC1 SC37 Biometrics. *ISO/IEC 30107-3. Information Technology - Biometric presentation attack detection - Part 3: Testing and Reporting*. International Organization for Standardization, 2017.

[9] J. Liu and A. Kumar. Detecting presentation attacks from 3d face masks under multispectral imaging. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, June 2018.

[10] I. Manjani, S. Tariyal, M. Vatsa, R. Singh, and A. Majumdar. Detecting silicone mask-based presentation attack via deep dictionary learning. *IEEE Transactions on Information Forensics and Security*, 12(7):1713–1723, 2017.

[11] O. M. Parkhi, A. Vedaldi, and A. Zisserman. Deep face recognition. In *British Machine Vision Conference*, 2015.

[12] R. Raghavendra and C. Busch. Novel presentation attack detection algorithm for face recognition system: Application to 3d face mask attack. In *IEEE International Conference on Image Processing (ICIP), Paris, France*, pages 323–327, Oct 2014.

[13] R. Raghavendra and C. Busch. Presentation attack detection methods for face recognition systems: A comprehensive survey. *ACM Computing Survey*, 50(1):8:1–8:37, Mar. 2017.

[14] F. Schroff, D. Kalenichenko, and J. Philbin. FaceNet: A Unified Embedding for Face Recognition and Clustering. In *Proceedings of the IEEE Intl. Conf. on Computer Vision and Pattern Recognition (CVPR)*, pages 815 – 823, 2015.

[15] H. Steiner, A. Kolb, and N. Jung. Reliable face anti-spoofing using multispectral swir imaging. In *2016 International Conference on Biometrics (ICB)*, pages 1–8. IEEE, 2016.

[16] D. Wen, H. Han, and A. Jain. Face spoof detection with image distortion analysis. *IEEE Transactions on Information Forensics and Security*, 10(99):1–16, 2015.

[17] X. Wu, R. He, and Z. Sun. A Lightened CNN for Deep Face Representation. *CoRR*, abs/1511.02683, 2015.