Presentation Attack Detection for Mobile Device-based Iris Recognition

Ewelina Bartuzi[1][0000−0001−6245−2908] and
Mateusz Trokielewicz[1][0000−0002−7363−8385]

Research and Academic Computer Network (NASK)
Kolska 12, Warsaw 01-045, Poland
{ewelina.bartuzi, mateusz.trokielewicz}@nask.pl
https://eng.nask.pl

Abstract. Apart from ensuring high recognition accuracy, one of the main challenges associated with mobile iris recognition is reliable Presentation Attack Detection (PAD). This paper proposes a method of detecting presentation attacks when the iris image is collected in visible light using mobile devices. We extended the existing database of 909 bonafide iris images acquired with a mobile phone by collecting additional 900 images of irises presented on a color screen. We explore different image channels in both RGB and HSV color spaces, deep learning-based and geometric model-based image segmentation, and use Local Binary Patterns (LBP) along with the selected statistical images features classified by the Support Vector Machine to propose an iris PAD algorithm suitable for mobile iris recognition setups. We found that the red channel in the RGB color space offers the best-quality input samples from the PAD point of view. In subject-disjoint experiments, this method was able to detect 99.78% of screen presentations, and did not reject any live sample.

Keywords: biometrics · iris recognition · presentation attack detection · mobile devices

1 Introduction

Iris biometrics is popular in high-security applications, since it provides a decent level of recognition accuracy compared to other biometric traits. Due to the ubiquity of mobile devices, with a 10 year increase in the average number of devices per 100 people from 50.6 to 103.5 [2], currently one of the fastest-growing branches of biometrics is the mobile-based authentication. In addition to high recognition accuracy, these solutions must also be resilient to attacks. This paper proposes an open-source Presentation Attack Detection for biometric systems using iris images collected in a mobile environment with non-specialized sensors, which broadens the possible applications beyond those to which the device manufacturer restricts a given biometric implementation. The method, utilizing local binary patterns and statistical image features selected with PCA and classified with an SVM, is able to reach close-to-perfect performance on a subject-disjoint testing subset consisting of live and screen iris presentations.

The study advances the research in iris and periocular biometrics by offering the **following contributions to the state of the art:**

- an open-source code for two PAD methods suitable for detecting color irises presented on a screen: one based on LBP texture descriptor, which detects areas with specific frequencies, the second employing statistical features of the image; this can serve as a useful benchmark method for visible light iris PAD,[1].
- extension of the existing database, consisting of 900 attack iris presentations obtained by imaging an iris sample displayed on a color screen,[2],
- analysis of the optimal color representation of samples collected in visible light employing different color spaces and their individual channels, with respect to both the PAD component as well as the overall recognition accuracy.

2 Related work review

The importance of equipping a biometric system with a reliable Presentation Attack Detection component is already well recognized throughout the biometrics community. Researchers have proposed numerous methods for detection of attack irises, including paper printouts, textured contact lenses, or prosthetic eyes. These techniques include employing image texture descriptors, such as Local Binary Patterns (LBP) [14], Binarized Statistical Image Features (BSIF) [16], or Local Phase Quantization [22], keypoint detectors and descriptors such as Scale Invariant Feature Transform (SIFT) [18], as well as calculating image quality metrics [8]. Deep-learning-based PAD techniques have also recently emerged, *e.g.*, [20]. Thavalengal *et al.* proposed a multi-spectral analysis of the iris [23]. Czajka, on the other hand, exploited biological properties of the eye's reaction to light stimuli, introducing a method based on pupil dynamics [12]. Recently, Trokielewicz *et al.* proposed a deep-learning-based PAD component for detecting cadaver iris presentations [4]. In the mobile domain, methods such as exploiting the properties of a light field camera have been proposed [19], or employing magnified phase information[21]. Recently published review paper by Czajka and Bowyer presents a systematic summary of PAD for iris recognition [13].

3 Experimental data

A part of the existing multimodal biometric database, created by these authors in the past, containing eye and periocular images was used in this paper [10]. Photographs have been collected in visible light using Huawei Mate S (13 Mpx, f/2.0). Data acquisition from 53 people (20 women and 33 men), aged from 14 to

[1] Available for download at `http://zbum.ia.pw.edu.pl/EN/node/46`.

[2] This dataset of attack iris samples is available to researchers at `http://zbum.ia.pw.edu.pl/EN/node/46`.

71 years, has been divided into two measurement sessions. Example photographs from this database are shown in Fig. 1. This dataset contains both the high quality (referred to as HQ) and low quality (LQ) images, with HQ images being taken with flash, and LQ images being taken without flash illumination.

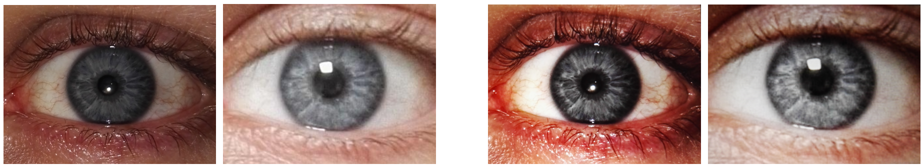


Fig. 1. From left to right: Example same-eye images from the smartphone camera acquired with flash (HQ image), without flash (LQ image), attack screen presentations for the same samples (with and without flash), taken with the same smartphone camera as original images.

Part of this study was to extend the live iris dataset with a complementary set of attack iris samples. For this purpose, a popular (in visible light iris recognition) way of creating artificial data was used: displaying photos on the screen of a phone and then taking photographs of the screen. The device used for taking pictures of artifacts was the same device as the one used to collect real samples in the original study. A total of 900 attack representations were created.

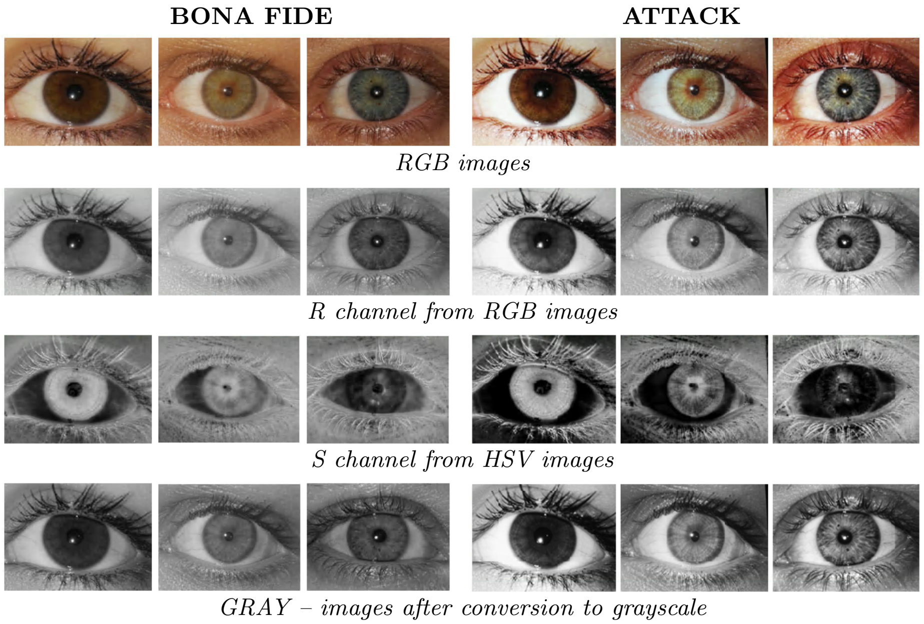
The iris images were cropped in accordance with the ISO/EIC 19794-6:2011, up to a resolution of 640×480 pixels. Color preprocessing was applied to come up with five different representations, the first two including **RGB**, unmodified images straight from the sensor (Fig. 2, top row), **R** images created by extracting the red channel of the RGB color space (Fig. 2, second row), **S** images created by extracting the saturation component from the HSV representation (Fig. 2, third row), and **GRAY** images (Fig. 2, bottom row), grayscale created from the RGB image.

4 Methodology

4.1 Presentation Attack Detection methods

For the purpose of mitigating screen presentation attacks, we have implemented two PAD methods, both requiring only a single, static iris image, and both not requiring any exhaustive computations. The first method relies on the statistical features of an image, whereas the second one utilizes an LBP-based texture descriptor.

Statistical image features: In this method, seven statistical characteristics of the image were calculated: average image intensity (μ), variance of the pixels value (σ), skewness ($\frac{\sum_{i=1}^{M}\sum_{j=1}^{N}(I(i,j)-\mu)^3}{M\cdot N\cdot\sigma^{3/2}}$), kurtosis ($\frac{\sum_{i=1}^{M}\sum_{j=1}^{N}(I(i,j)-\mu)^4}{M\cdot N\cdot\sigma^{2}}$), the 10th, 50th, and 90th percentile of the pixels value.



RGB images

R channel from RGB images

S channel from HSV images

GRAY – images after conversion to grayscale

Fig. 2. Examples of real (**left**) and fake (**right**) irises of brown/hazel, green and blue/gray eyes.

Local Binary Patterns: LBP is one of the most popular texture descriptors, which involves the analysis of a pixel in relation to its surroundings [3]. The value of each pixel is compared to the value of the central pixel, and the binary code created in this way is converted into a number in the decimal system: $LBP_{N,R}(I_C) = \sum_{n=1}^{N} s(I_n - I_C)2^{n-1}$ where N, R are the number of surrounding neighbors and the radius, respectively, I_C denotes the central pixel, I_N denotes the n-th pixel from neighborhood of central pixel, and

$$s(I_n - I_C) = \begin{cases} 1 & \text{if } I_n - I_C \geq 0 \\ 0 & \text{otherwise} \end{cases} \tag{1}$$

The obtained values are represented as histograms: one for the entire image, and individual histograms created for each piece of the image divided into 100 parts. In addition, we implemented version of the LBP algorithm resistant to rotation (uniform LBP code) [3]. The best results were obtained for the surroundings of eight neighbors analyzing the whole picture. In this way, the feature vectors counted 59 elements.

Features selection and classification: Features obtained from each method were then sorted by relevance using principal component analysis (PCA), the influence of subsequent features on attack detection accuracy was examined, and an optimal number of features is determined, Fig. 3. Table 1 summarizes the optimal number of LBP features for each type of image. The binary classification

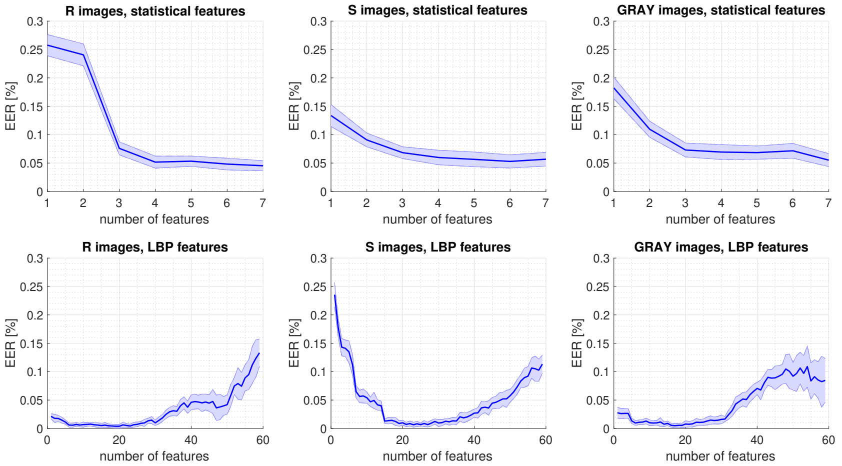


Fig. 3. EER as a function of the number of successive statistical (**top row**) and LBP (**bottom row**) features ordered according to the PCA. Average (solid dark blue lines) and standard deviation obtained from 20 training/testing data splits (light blue shades) are shown.

was carried out using a support vector machine (SVM) classifier with a radial basis function kernel. The data were divided into subject-disjoint training and testing subsets in a ratio of 80 : 20.

4.2 Iris recognition: OSIRIS

For the purpose of iris recognition, the Open Source for IRIS (OSIRIS) is employed [5]. The academic-based software was developed as a part of the BioSecure project and implements the original Daugman concept, including segmentation of the iris and its normalization by transformation from cartesian to polar coordinates using the *rubber sheet model*. The encoding of the iris features is performed using phase quantization of a response of Gabor filtering outcomes, and then comparing the binary codes using the XOR operation to obtain the normalized Hamming distance. Values close to zero should indicate data from the same iris, whereas typical results for different irises comparisons oscillate around 0.5 (usually they are concentrated in the range of 0.4-0.45 because of shifting of the iris codes).

OSIRIS segmentation: The first stage of iris image processing is location and segmentation, with the exception of image noise, among others in the form of eyelids, eyelashes, reflections, shadows. The result is a binary mask, which determines which pixels belong to the iris. In the original OSIRIS, this is performed by a circular Hough transform used to roughly approximate the circles representing the edges of the iris, and then employing active contours algorithm to exclude noisy regions. Examples of segmentation results can be seen in Fig. 4,

Method	*Optimal no. of features*	*Image type*	*APCER [%]*	*BPCER [%]*	*EER_{PAD} [%]*
Statistical features	7	R	1.89	4.22	5.00
	6	S	5.94	4.84	4.95
	7	GRAY	3.92	4.53	4.97
LBP	20	R	0.22	0.00	0.10
	23	S	0.11	0.44	0.28
	17	GRAY	0.44	0.00	0.22

Table 1. APCER, BPCER, EER$_{PAD}$ for two PAD methods for the optimal no. of parameters.

top row.

DCNN-based segmentation: Due to the type of photos that is different from NIR images for which OSIRIS was originally built, a second segmentation method is also used, namely a model based on the SegNet architecture retrained with iris images taken in infrared light, as well as images representing only the R channel of the RGB color scheme, using the implementation from [6].

5 Experiments and results

Stage 1: Presentation Attack Detection

The LBP descriptor and seven statistical image characteristics were used as a PAD component put in front of the iris recognition pipeline. In both cases, the features were sorted from the most important, and then combined in the order given and treated as a feature vector, which was then classified using a SVM (Fig. 3). The data were divided with 20 subject-disjoint, 80/20 training and test sets. In this part of the analysis, we used iris representations in form of R, S and GRAY images. Tab. 1 presents the optimal number of parameters for each image type and the results of the binary classification in the form of the error metrics as recommended by the ISO/IEC standard on presentation attack detection.

Both the methods based on texture analysis and statistical image features obtained good results, with LBP allowing for almost perfect discrimination between bona-fide and attack samples. In the case of the R channel image representation, the EER_{PAD} was as low as 0.11%. Since the EER_{PAD} was considered as a minimization target for parameter choice, the APCER here is larger than BPCER, however, this can be tuned by moving the acceptance threshold in the desired direction. Analysis of the statistical features of the image gave higher errors, but a downward trend can be seen here, therefore adding other image features could also reduce these errors. In each case the most discriminative information seems to be represented in the red channel R. **The APCER and BPCER errors concern LQ images, for the HQ both errors are 0%, and both the detection rate of attack and bona fide samples, is perfect.**

Stage 2: Iris recognition

In this Section we evaluate the impact that attack samples can have on recognition accuracy. The first stage of the OSIRIS recognition pipeline consists of image segmentation. The original OSIRIS segmentation repeatedly encountered problems with correctly segmenting the images, Fig. 4.

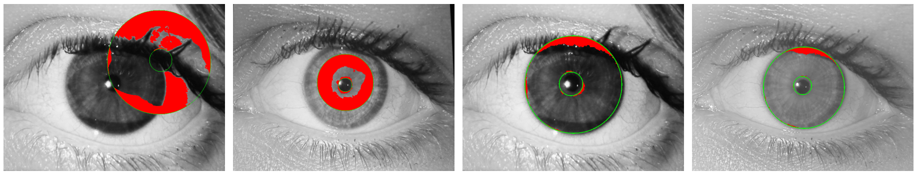


Fig. 4. Examples of incorrect segmentation from the OSIRIS algorithm (left) and corresponding results for DCNN-based approach (right) for the same samples. The iris should ideally be located within the two green circles, and the red regions should denote non-iris portions that lay within these circles.

Since most of these errors will likely lead to erratic iris verification, the segmentation phase has to be altered or replaced to be able to test the effectiveness of the iris feature representation, encoding, and matching of OSIRIS for iris images taken in visible light. The original segmentation was therefore replaced with a solution based on convolutional neural network and the Hough transform, cf. Sec. 4.2, which allowed for close-to-perfect segmentation of most samples, cf. Fig. 4, bottom row.

Image Type	*R*	*S*	*GRAY*
Without PAD component			
HQ	29.15 (±0.66)	27.55 (±0.54)	6.86 (±0.29)
LQ	38.71 (±0.48)	46.28 (±0.30)	37.14 (±0.58)
HQ:LQ	40.40 (±0.48)	47.52 (±0.24)	53.45 (±0.53)
Including PAD component			
HQ	7.68 (±0.58)	26.02 (±0.47)	6.57 (±0.29)
LQ	18.09 (±0.41)	43.77 (±0.33)	21.36 (±0.57)
HQ:LQ	20.24 (±0.42)	47.48 (±0.23)	47.83 (±0.53)

Table 2. Equal Error Rates and their standard deviations obtained using OSIRIS for three different image representation: R, S, and GRAY, for higher- and lower quality images and between different qualities, both without and with the PAD component.

After the segmentation stage, the OSIRIS method was employed to calculate all possible comparison scores between within-class image pairs (*genuine*), between-class image pairs (*impostors*), and between real samples and their corresponding attack representations (*real vs fake*). This was done for all three image types (R, S, and GRAY), as well as for high quality images only, low quality images only, and between images of different quality (denoted as HQ, LQ, and HQ:LQ, respectively), leading to a total of 9 experiments without, and 9 experiments with the PAD component included. Comparison scores distributions are presented in Fig. 5, whereas the obtained average Equal Error Rates for all 18 experiments are summarized in Tab. 2. In the experiments run without the PAD method, all comparisons with *attack* samples are considered as impostor comparisons.

Conclusions drawn from comparison score distributions plotted in Fig. 5:

1) the best separation between genuine and impostor comparisons can be found when matching high quality (HQ) images representing the R channel and the grayscale (GRAY) conversion of the RGB image,
2) both the low quality (LQ) and mixed quality (HQ:LQ) comparisons do not offer distribution separation that would enable reasonably accurate iris recognition,
3) using the R channel, ***bonafide-vs-attack*** **scores overlap with those obtained from** ***genuine comparisons*****, thus making the PAD component crucial** (including it improves the EER from almost 30% to 7.68%),
4) but, surprisingly, for grayscale (GRAY) images the same scores overlap with the *impostor* scores distribution (and including PAD in this case improves the EER from 6.86% to 6.57%).

The best results EER-wise were obtained for high-quality images, in the representations of the R channel and grayscale conversion of the RGB color space, which gave EER=7.68% and EER=6.57%, respectively. For lower-quality images, R images yielded better results, giving (still unacceptable) EER of 18.13%, compared against 21.36% obtained for the GRAY images and EER=43.77% for S images. For mixed quality database, only the R channel allows for a non-random recognition accuracy with EER=20.24%, whereas S and GRAY images yield EERs close to 50%.

6 Conclusions

This paper offers an open-source presentation attack detection method designed to detect attack representations of iris samples collected with a mobile phone in visible light. LBP-derived features allow for a nearly perfect attack detection accuracy with APCER=0.11% and BPCER=0% with a dataset of attack iris representations consisting of irises displayed on a screen, which was created for this study.

By testing the proposed PAD method coupled with the OSIRIS recognition pipeline with DCNN-based image segmentation stage, we show that by employing the R channel of the RGB color space, recognition accuracy of 7.68% EER can be achieved, compared to EER=30% obtained prior to the inclusion of a

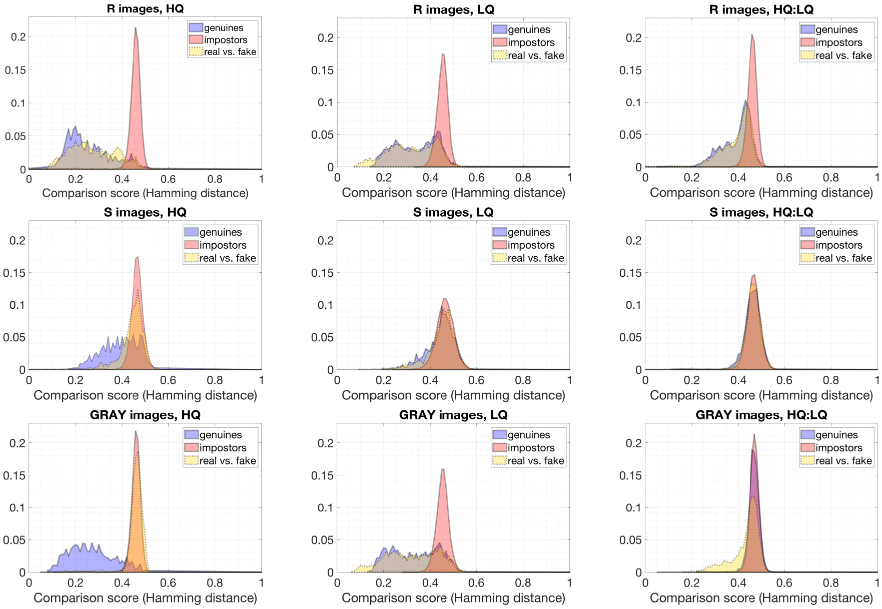


Fig. 5. Score distributions obtained for HQ images only (**left**), for LQ images only (**middle**), and between HQ and LQ images (**right**). Genuine scores (blue), impostor scores (red), and scores between real samples and their fake representations (yellow) are shown.

PAD component. Surprisingly, we have found the grayscale representation of the RGB image color space to offer some kind of resilience to this particular attack, as comparison scores obtained from matching real iris samples and their fake counterparts were similar to the scores obtained from matching impostor image pairs. Here, the proposed PAD allowed for a moderate reduction of EER from 6.86% to 6.57%.

This paper follows the guidelines for research reproducibility by making the dataset of attack iris representations, as well as source codes for the PAD methods proposed, open-sourced to serve as a benchmark for visible light iris presentation attack detection, especially with respect to mobile applications.

Acknowledgments

Project CYBERSECIDENT/382354/II/NCBR/2018 financed by the National Centre for Research and Development in the framework of CyberSecIdent programme.

References

1. Quinn G. W., Matey J. R., Grother P., "IREX IX Part One: Performance of Iris Recognition Algorithms", NIST Interagency Report 8207, https://doi.org/10.6028/NIST.IR.8207, 2018
2. International Telecommunication Union (ITU), "Global and regional ICT Data: Mobile-cellular subscriptions", `https://www.itu.int/en/itud/statistics/pages/stat/default.aspx`," accessed on June, 2018.
3. Ojala T., Pietikainen M., and Maenpaa T., "Multiresolution Gray Scale and Rotation Invariant Texture Classification With Local Binary Patterns." IEEE Transactions on Pattern Analysis and Machine Intelligence. vol. 24, issue 7, July 2002, pp. 971-987, DOI: 10.1109/TPAMI.2002.1017623.
4. Trokielewicz M., Czajka A., Maciejewicz P., "Presentation Attack Detection for Cadaver Iris", 9th IEEE International Conference on Biometrics: Theory, Applications and Systems (BTAS 2018), Los Angeles, California, October 22-25, 2018
5. Othman N., Dorizzi B., Garcia-Salicetti S., OSIRIS: An open source iris recognition software, https://doi.org/10.1016/j.patrec.2015.09.002
6. Trokielewicz M., Czajka A., Maciejewicz P., "Post-mortem Iris Recognition with Deep-Learning-based Image Segmentation", http://arxiv.org/abs/1901.01708, 2019
7. ISO/IEC 30107-1:2016, "Information technology – Biometric presentation attack detection – Part 1: Framework", 2016
8. J. Galbally, S. Marcel, and J. Fierrez. Image Quality Assessment for Fake Biometric Detection: Application to Iris, Fingerprint, and Face Recognition. *IEEE Transactions on Image Processing*, 23(2):710–724, Feb 2014.
9. Ren J., Jiang X., Yuan J., "Noise-Resistant Local Binary Pattern with an Embedded Error-Correction Mechanism", IEEE Transactions on Image Processing, vol. 22, no. 10, DOI: 10.1109/TIP.2013.2268976, 2013
10. Bartuzi E., Roszczewska K., Trokielewicz M., Białobrzeski R., "MobiBits: Multimodal Mobile Biometric Database", BIOSIG 2018: 17th International Conference of the Biometrics Special Interest Group, Sep. 26 - 29, 2018, Darmstadt, Germany, DOI: 10.23919/BIOSIG.2018.8553108
11. ISO/IEC JTC 1/SC 37, "Information technology – Vocabulary – Part 37: Biometrics (FDIS)," October 2016.
12. A. Czajka. Pupil dynamics for iris liveness detection. *IEEE Trans. Inf. Forens. Security*, 10(4):726–735, April 2015.
13. A. Czajka and K. Bowyer. Presentation Attack Detection for Iris Recognition: An Assessment of the State of the Art. *ACM Computing Surveys (in review), https://arxiv.org/abs/1804.00194*, 2018.
14. J. S. Doyle, P. J. Flynn, and K. W. Bowyer. Automated classification of contact lens type in iris images. In *IEEE Int. Conference on Biometrics (ICB)*, pages 1–6, June 2013.
15. ISO/IEC 30107-1:2016. Information technology – Biometric presentation attack detection – Part 1: Framework, 2016.
16. J. Komulainen, A. Hadid, and M. Pietikäinen. Generalized textured contact lens detection by extracting bsif description from cartesian iris images. In *IEEE Int. Joint Conference on Biometrics (IJCB)*, pages 1–7, Sept 2014.
17. D. Menotti, G. Chiachia, A. Pinto, W. R. Schwartz, H. Pedrini, A. X. Falcão, and A. Rocha. Deep representations for iris, face, and fingerprint spoofing detection. *IEEE Transactions on Information Forensics and Security*, 10(4):864–879, April 2015.

18. F. Pala and B. Bhanu. Iris liveness detection by relative distance comparisons. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, July 2017.
19. R. Raghavendra and C. Busch. Presentation attack detection on visible spectrum iris recognition by exploring inherent characteristics of light field camera. In *IEEE International Joint Conference on Biometrics*, pages 1–8, Sep. 2014.
20. R. Raghavendra, K. B. Raja, and C. Busch. Contlensnet: Robust iris contact lens detection using deep convolutional neural networks. In *IEEE Winter Conference on Applications of Computer Vision (WACV)*, pages 1160–1167, March 2017.
21. K. B. Raja, R. Raghavendra, and C. Busch. Video presentation attack detection in visible spectrum iris recognition using magnified phase information. *IEEE Transactions on Information Forensics and Security*, 10(10):2048–2056, Oct 2015.
22. A. F. Sequeira, S. Thavalengal, J. Ferryman, P. Corcoran, and J. S. Cardoso. A realistic evaluation of iris presentation attack detection. In *Int. Conference on Telecommunications and Signal Processing (TSP)*, pages 660–664, June 2016.
23. S. Thavalengal, T. Nedelcu, P. Bigioi, and P. Corcoran. Iris liveness detection for next generation smartphones. *IEEE Trans. Consumer Electronics*, 62(2):95–102, May 2016.