LivDet-Iris 2013 – Iris Liveness Detection Competition 2013

David Yambay
Dept. of Electrical and Computer Eng.
Clarkson University
yambayda@clarkson.edu

James S. Doyle
Dept. of Computer Science and Eng.
University of Notre Dame
jdoyle6@nd.edu

Kevin W. Bowyer
Dept. of Computer Science and Eng.
University of Notre Dame
kwb@nd.edu

Adam Czajka
Research and Academic Computer Network (NASK)
Warsaw University of Technology
aczajka@elka.pw.edu.pl

Stephanie Schuckers
Dept. of Electrical and Computer Eng.
Clarkson University
sschucke@clarkson.edu

Abstract

The use of an artificial replica of a biometric characteristic in an attempt to circumvent a system is an example of a biometric presentation attack. Liveness detection is one of the proposed countermeasures, and has been widely implemented in fingerprint and iris recognition systems in recent years to reduce the consequences of spoof attacks. The goal for the Liveness Detection (LivDet) competitions is to compare software-based iris liveness detection methodologies using a standardized testing protocol and large quantities of spoof and live images. Three submissions were received for the competition Part 1; Biometric Recognition Group de Universidad Autonoma de Madrid, University of Naples Federico II, and Faculdade de Engenharia de Universidade do Porto. The best results from across all three datasets was from Federico with a rate of falsely rejected live samples of 28.6% and the rate of falsely accepted fake samples of 5.7%.

1. Introduction

Results have shown that the iris biometric modality can be spoofed by obfuscating the natural iris pattern with an opaque contact lens or through the use of printed images of irises. There have been numerous techniques proposed to answer the problem of this susceptibility of iris systems. One primary countermeasure to spoofing attacks is called "liveness detection." Liveness detection is based on the principle that additional information can be garnered above and beyond the data procured by a standard verification system, and this additional data can be used to verify if an image is authentic. Liveness detection uses either a hardware-based system or software-based system coupled with the authentication algorithm to provide additional security. Hardware-based systems use additional sensors to gain measurements outside of the iris image itself to detect liveness. Software-based systems use image processing algorithms to gather information directly from the collected iris image in order to detect liveness. These systems classify images as either live or fake.

In 2009, in order to assess the main achievements of the state of the art in fingerprint liveness detection, University of Cagliari and Clarkson University, organized the first Fingerprint Liveness Detection Competition. In 2013, the LivDet competitions expanded into the iris biometrics with collaborations between Clarkson University, Warsaw University of Technology, and University of Notre Dame.

The First International Fingerprint Liveness Detection Competition LivDet 2009 [9], provided an initial assessment of software systems based on the fingerprint image only. The second Liveness Detection Competition (LivDet 2011 [15]) was created in order to ascertain the current state of the art in liveness detection, and also included integrated system testing. The third Liveness Detection Competition (LivDet 2013 [6]) expanded on previous competitions with the inclusion of a iris component. LivDet 2013 was split into two separate competitions, LivDet-Fingerprint 2013 and LivDet-Iris 2013. LivDet-Iris 2013 was held in con-

David Yambay, James Doyle, Kevin Bowyer, Adam Czajka, Stephanie Schuckers, "LivDet-Iris 2013 - Iris Liveness Detection Competition 2013",
International Joint Conference on Biometrics - IJCB 2014, 29 September - 2 October 2014, Clearwater, Florida, USA

junction with BTAS 2013.

The results were originally presented only as a poster. In this paper, the competition design and results of the submitted algorithms for LivDet-Iris 2013 are summarized. Section 2 describes the background of liveness detection in iris recognition. Section 3 explains the methods and protocol used to evaluate the submitted algorithms. Section 4 discusses the results of the competition. Section 5 covers the conclusions that are garnered through the analysis of this competition.

2. Background

The concept of spoofing has existed for some time now. Research into spoofing can be seen beginning in 1998 from research conducted by D. Willis and M. Lee where six different biometric fingerprint devices were tested against fake fingers and it was found that four of the six were susceptible to spoofing attacks [14]. A year later Daugman proposed an idea how to spoof and, simultaneously, how to parry an attack based on iris printouts [4]. This idea was lately incorporated into the first security evaluation of commercial iris recognition systems by Thalheim *et al.* [11], which had shown that the iris could be printed, presented and positively verified by a commercial system. This pioneer research stimulated others presenting their own security evaluation of additional hardware. Matsumoto presented in 2004 alarming lack of effective countermeasures in the commercial equipment not used by Thalheim *et al.*. Pacut and Czajka examined the severe lack of anti-spoofing measures in additional iris systems, and made a survey of types of eye forgery attacks as well as proposed multiple solutions to these forms of attacks in 2006 [10]. Z. Wei *et al*. produced results on detection of counterfeit irises in 2008 using three different anti-spoofing iris measures [13]. Recently Czajka proposed a frequency-based method dedicated to detection of printed irises [3]. Currently we may observe more frequent implementations of liveness detection, with either hardware-based or software-based systems. Usually the result of this analysis is a score used to classify images as either live or fake.

There are two main attack methods for iris systems, the spoofing and the obscuring of natural patterns, Fig. 1. In the spoofing method, an image of a subject's iris is obtained and printed out onto paper. This iris image is then presented to the iris system. Obscuring of natural iris patterns is accomplished through the use of wearing a patterned contact lens. This pattern covers over the natural iris pattern of a subject to hide their identity. LivDet 2009 created a benchmark for measuring liveness detection algorithms. It provided results that showed the current state of the art at that time [9]. The objective of this competition is to expand on LivDet 2009 and 2011 by evaluating the performance of algorithms for the iris liveness detection operating in two mentioned scenarios: impersonation using paper printouts, and obfuscation employing printed contact lens. The value in the addition of the iris biometrics is to present the current state of the art for this modality related to software-based liveness routines.

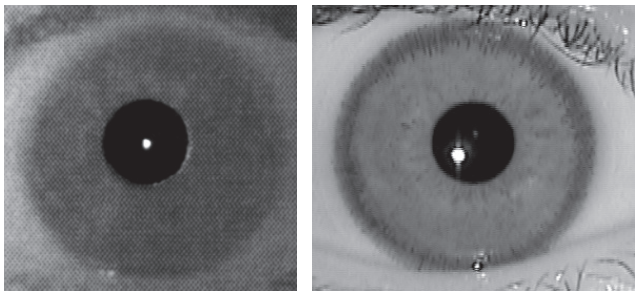


Figure 1: Examples of impersonation and obscuring of the iris individual patterns. The former may be done by printing a given pattern on a paper (left), and for the latter we may use a contact lens (right) with any pattern printed on it, and interfering with an actual texture of the iris tissue.

3. Experimental Protocol and Evaluation

The competition for LivDet-Iris 2013 was focused on algorithms. The design of the experiment will be discussed in detail in this section also outlining the constraints placed on entrants. To request a copy of the dataset, please visit `http://people.clarkson.edu/projects/biosal/iris/registration.php`.

3.1. Participants

The competition is open to all academic and industrial institutions. Upon registration, each participant was required to sign a database release agreement detailing the proper usage of data made available through the competition. Participants were then given a database access letter with a user name and password to access the server to download the training data. In Table 1 are presented the participants names and the correspondent algorithms names as they are used in this paper. Three groups submitted algorithms for the competition.

3.2. Part 1: Algorithm – Data Sets

The dataset contains images from three different datasets. Spoof images were collected using two different presentation attack types. The first being patterned contact lenses which are used to obscure the natural iris pattern. (Prepared by Clarkson University and University of Notre Dame.) The second attack type is printed iris spoofs which aim to identify as another person (prepared by Warsaw University of Technology). Table 2 illustrates number of im-

Participant name	Algorithm name
ATVS - Biometric Recognition Group, Universidad Autonoma de Madrid	ATVS
University of Naples Federico II	Federico
Faculdade de engenharia de Universidade do Porto	Porto

Table 1: Name of the participants and the submitted iris algorithms.

ages made available for training and used in testing. The next subsections present details of each dataset.

	Training		Testing	
	Life	Spoof	Live	Spoof
Notre Dame	2000	1000	800	400
Warsaw	228	203	624	612
Clarkson	270	400	246	440

Table 2: Number of images in training and testing iris datasets.

3.2.1 University of Notre Dame Subset

The ND Contact Lens Detection 2013 (NDCLD'13) dataset [12] [5] consists of 4200 images (3000 used for classifier training and 1200 reserved for testing) acquired with an LG 4000 [8] iris camera. Both the training set and the verification set are divided equally into three classes: (1) no contact lenses, (2) soft, non-textured contact lenses, and (3) textured contact lenses. Classes (1) and (2) are both "live" images and class (3) constitutes "fake" irises. Classes (1) and (2) are additionally balanced between male and female, and represent a variety of ethnicities. Category (3) images are predominantly from Caucasian males. All images were acquired in a laboratory with constant lighting and under the observation of a trained acquisition operator.

All textured contact lenses in this dataset came from three major suppliers: Johnson&Johnson [7], CIBA Vision [1], and Cooper Vision [2]. Subjects in the database belong to four different ethnic categories (Caucasian, Asian, Black, and Other). Multiple colors of contact lenses were selected for each manufacturer. Some were also "toric" lenses, meaning that they are designed to correct for astigmatism. Toric lenses are designed to maintain a preferred orientation around the optical axis. As such, they may present different artifacts than non-toric lenses but also may have less variation in the position on the eye.

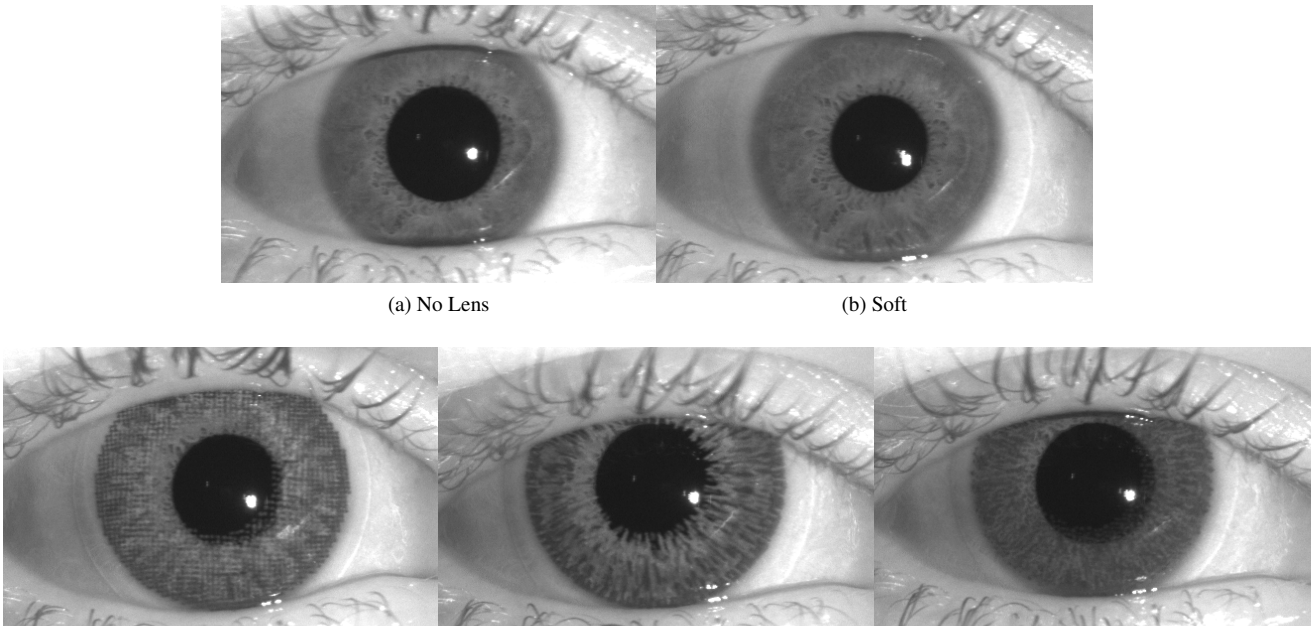
Sample images of the three classes and three manufacturers of cosmetic contact lenses represented in the NDCLD'13 can be found in Figure 2. All images are from the same subject eye but are for visual comparison only, no single subject is represented in all three classes in the NDCLD'13 dataset. The images are cropped to allow the texture in the cosmetic lens to be more apparent.

3.2.2 Warsaw subset

Warsaw group followed earlier experiments [10] when preparing this subset suggesting a laser printing on a matt paper as an optimal process to produce artificial paper irises eagerly accepted by some commercial systems. Additional gimmick of making a hole instead of a pupil was applied, as cameras typically search for a specular reflection from a cornea when detecting the iris presence behind the camera.

Preparing fake irises in a form of paper printouts gives an opportunity to impersonate subjects, as printing a given (rather than any) iris pattern on a carrier can be done at the same cost. Hence, in the Warsaw subset authentic samples have their fake counterparts. In this approach, level of imitation sophistication cannot be accidental, though the precise rules how to prepare good printouts are difficult to be formally defined. To find this borderline all the prepared printouts were presented to an example commercial camera (Panasonic ET-100, purchased in 2003). Only printouts that were accepted by this system (i.e. the iris pattern read from artifacts matched the corresponding iris templates based on authentic, living eyes) were then photographed by a separate commercial iris capture camera with convenient iris capture capability (IrisGuard AD100, purchased in 2009 and with **liveness detection intentionally set off**). Each printout presented to Panasonic ET-100 and IrisGuard AD00 had an iris image printed with 1:1 scale, and was cut keeping a few mm of white background. This allowed to keep the printout very closely to the subjects eye, who looked through the hole (cut instead of the pupil) to interact with the camera in a fashion it expects when a living eye is captured. There were many different subjects presenting the printouts, yet the presentation manner was stable across the volunteers. Subject presenting the printout took care to present it with no rotation and 3D distortions. Such an approach of printouts preparation increases the value of this subset, as it correctly predicts reliability of fake samples expected in real attack scenarios.

Two different printers were used to build Warsaw subset: a) HP LaserJet 1320 and b) Lexmark c534dn. The former represents a standard black and white laser printer, delivering iris images of "low resolution" (not exceeding 600 dpi), and was intentionally selected as an example of a low-cost printing device. The latter device is semi-professional color laser printer allowing to produce printouts of "high resolu-



(a) No Lens (b) Soft

(c) Ciba (d) Cooper (e) J&J

Figure 2: Sample images of University of Notre Dame subset showing the variety of cosmetic lens textures that are available from different manufacturers. All images are from the same subject eye.

tion" (though not exceeding 1200 dpi), Fig. 3. The Warsaw subset comprises 852 images of 284 distinct eyes, and 815 biometric images of paper printouts for 276 distinct eyes. We divided this set into training and testing subsets, and composed the training subset with high resolution printouts only. In turn, we intentionally mixed high and low resolution printouts in the testing subset to favor algorithms which are agnostic to the printout resolution. Table 2 details number of images used in both subsets.

3.2.3 Clarkson subset

The Clarkson University Dataset was created using patterned contacts lens in an attempt to obscure the subject's true identity from the system. The dataset sought to add in an element of difficulty through the use of slight blur of the images. Images were collected through the use of video capture of 100 frames at 25 frames/sec using a Dalsa camera. The sequence began out of focus and was moved through the focus range across full focus and back to out of focus. This allowed images to be taken at multiple different levels of blur. Initial images were taken at 0% blur as well as at 5% blur on each side of the 0% frame and finally 10% blur on each side of the 0% frame. The training dataset contained images at these 5 blur levels, however it was found that the 10% blur was more than would be accepted by a system and the testing dataset only contained 0% and 5% blur images. The dataset also contained two illuminations of images. the first illumination was with using a lamp along the camera with illumination just enough to fully illuminate the subject's face to simulate a night environment. The second illumination was with normal lights in the room turned on as well as the addition of the lamp with the camera.

The live images were taken from a total of 64 eyes with up to 5 images supplied from each eye at each illumination level for the training set. With the removal of the 10% blur from the testing set, there were only 3 images per eye for each illumination level.

Spoof images were collected from 6 subjects that had participated in the live collection. Since patterned contact lenses obstruct the underlying iris image, it was seen that varying the subject wearing the contact lens did not alter the final iris image as what was seen is the pattern on the lenses. This brought the idea that a wider array of contacts was more important than a wide array of subjects for patterned lenses.

The spoof dataset consisted of images from the 6 subjects each wearing 19 different patterned contact lenses. The types of lenses are listed in Table 3.

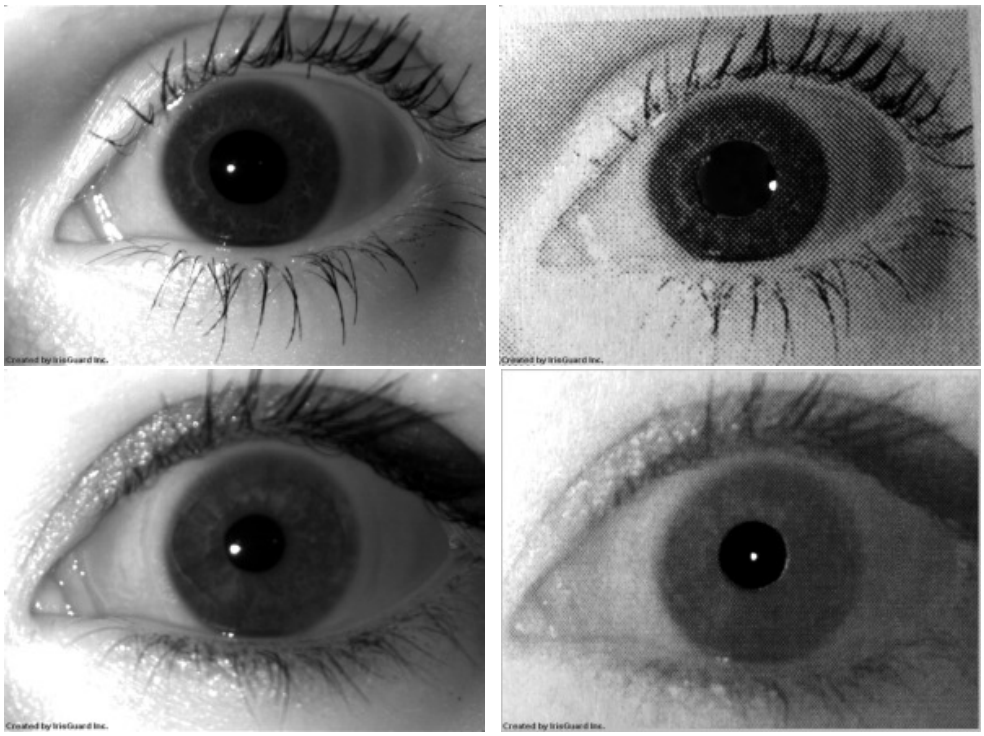


Figure 3: Sample images of Warsaw subset. Images of the authentic eyes are shown in the left column, and their fake counterparts are shown in the right column. Low and high resolution printouts are presented in the upper and lower rows, respectively.

3.3. Algorithm Submission

The algorithm submission for LivDet 2013 used the same structure as LivDet 2011, and each submitted algorithm had to be a Win32 console application `LIVENESS_XYZ.exe` (where `XYZ` is the identification number of the participant) with the following list of parameters:

> `ndataset`: identification number of the data set to analyze, i.e. 1=Clarkson , 2=Notre Dame, 3=Warsaw
>
> `inputfile`: TXT file with the list of images to analyze. Each image will be in the same format as the training data.
>
> `outputfile`: TXT file with the output of each processed image with a newline between each output, in the same order of inputfile. There was one output file for each input file. In the case that the algorithm could not be able to process the image, the correspondent output was -1000 (failure to enroll).

Each user had a chance to configure his algorithm by the training set available after registration. Participants could also choose to publish a description and/or source code of their algorithm, but this was not mandatory.

3.4. Performance Evaluation

Each of the algorithms returned an integer value representing a posterior probability of the live class given the image or a degree of "liveness" normalized in the range 0 and 100 (100 is the maximum degree of liveness, 0 means that the image is fake). The threshold value for determining liveness was set at 50. This threshold was then used to calculate *Ferrfake* and *FerrLive* error estimators, where:

- *Ferrlive* is the rate of misclassified live iris images (live called fake),
- *Ferrfake* is the rate of misclassified fake iris images (fake called live).

Both *Ferrlive* and *Ferrfake* were calculated for each data subset separately, as well as the average values across all subsets were presented. To select a winner the average of *Ferrlive* and *Ferrfake* was calculated for each participant and each data subset.

4. Results and Discussion

Three algorithm submissions successfully completed the competition at the time of submission of this paper. Re-

Figure 4: Sample images from Clarkson Dataset. Live Images (top) and spoof images (bottom)

Contact Number	Patterned Contact Type	Color
1	FreshLook Dimensions	Pacific Blue
2	FreshLook Dimensions	Sea Green
3	FreshLook ColorBlends	Green
4	FreshLook ColorBlends	Blue
5	FreshLook ColorBlends	Brown
6	FreshLook Colors	Hazel
7	FreshLook Colors	Green
8	FreshLook Colors	Blue
9	Phantasee Natural	Turquoise
10	Phantasee Natural	Green
11	Phantasee Vivid	Green
12	Phantasee Vivid	Blue
13	Phantasee Diva	Black
14	Phantasee Diva	Brown
15	ColorVue BiggerEyes	Cool Blue
16	ColorVue BiggerEyes	Sweet Honey
17	ColorVue 3 Tone	Green
18	ColorVue Elegance	Aqua
19	ColorVue Elegance	Brown

Table 3: Clarkson Patterned Contact Types

sults of this competition are not reflective of performance for spoof attacks not included in this study.

Results from the Iris Part 1: Algorithms are shown in Figures 5 and 6. ATVS and Porto declined to participate against the Notre Dame dataset and only Federico was tested against all three subsets. Since Federico was the only submission to test against the Notre Dame dataset, they are the only group that has a 3 dataset average error rate.

Across the two datasets that all three algorithms submitted against, Porto had the lowest average rates with a 12.2% FerrLive and 10.0% FerrFake. Federico tested a lower FerrFake than the other algorithms, but their higher FerrLive created a higher average error rate.

Error rates were lower for spoof images on the Warsaw dataset than any other, which demonstrates that printed irises are easier to detect than printed contact lenses. The Clarkson University dataset had the overall worst error rates.

The error rates for each patterned contact type for the Clarkson Dataset are given in Table 4. The FreshLooks Dimensions brand had the highest error rates among the contact types.

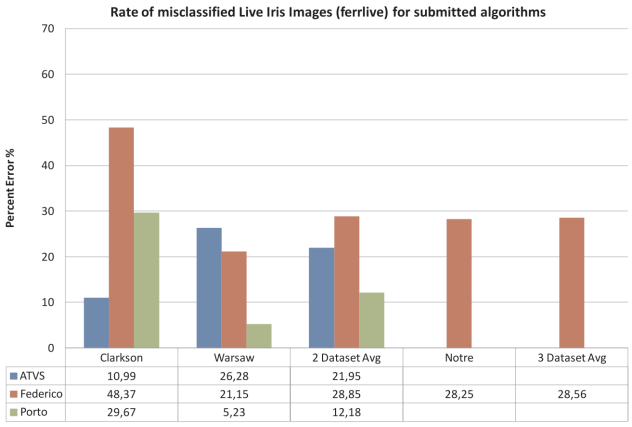


Figure 5: Rate of misclassified live iris images for submitted algorithms.

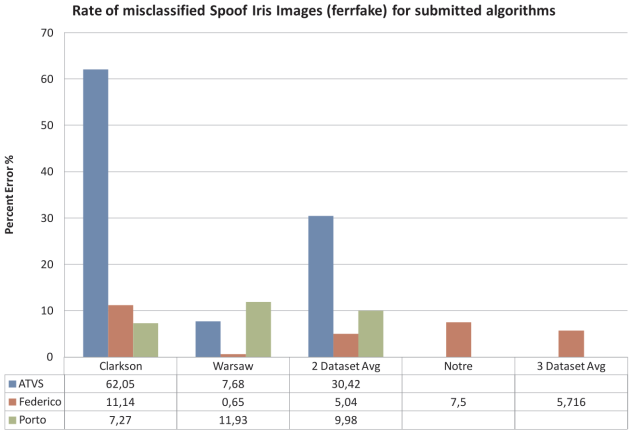


Figure 6: Rate of misclassified spoof iris images for submitted algorithms.

5. Conclusions

LivDet-Iris 2013 was the first public assessment of algorithm-based iris liveness detection. As the first competition it shows similar promise as the original LivDet 2009 with the enthusiasm for the competition. Although only three competitors submitted algorithms for the competition, the dataset used in the competitions is available by request and there are many institutions which have applied to receive the dataset. The best results were seen by Porto in Part 1, but Federico overall across the 3 datasets. Upon testing the submitted algorithms and compiling all results for submitted algorithms, it was shown that printed iris images were much easier to detect than patterned contact lens.

It is hoped that this competition will be continued in order to continually understand and promote the state of the art in liveness detection. Creating effective liveness detection methodologies is an important step in minimizing the vulnerability of spoof attacks.

Contact Number	ATVS	Federico	Porto
1	29.41	52.94	47.06
2	46.15	19.23	11.54
3	50.00	0.00	0.00
4	47.06	5.88	0.00
5	66.67	5.56	16.67
6	70.00	0.00	0.00
7	77.78	0.00	11.11
8	29.41	5.88	5.88
9	55.56	22.22	11.11
10	65.00	10.00	5.00
11	77.78	0.00	0.00
12	66.67	11.11	11.11
13	41.67	0.00	8.33
14	30.30	3.03	9.09
15	74.42	4.65	0.00
16	69.77	2.33	9.30
17	78.57	16.67	0.00
18	80.95	14.29	2.38
19	73.17	19.51	7.32

Table 4: Percent Error by Pattern Type for Clarkson Patterned Contacts

6. Acknowledgements

This material is based upon work supported by the National Science Foundation under Grant No. #1068055 and the Center for Identification Technology research. We would also like acknowledge the funding support from Research and Academic Computer Network (NASK), Warsaw, Poland.

References

[1] CibaVision. Freshlook colorblends, April 2013. `http://www.freshlookcontacts.com`.

[2] Cooper Vision. Expressions colors, April 2013. `http://coopervision.com/contact-lenses/expressions-color-contacts`.

[3] A. Czajka. Database of iris printouts and its application: Development of liveness detection method for iris recognition. In *Methods and Models in Automation and Robotics (MMAR), 2013 18th International Conference on*, pages 28–33, Aug 2013.

[4] J. Daugman. Countermeasures against subterfuge. In Jain, Bolle, and Pankanti, editors, *Biometrics: Personal Identification in Networked Society*, pages 103–121. Amsterdam: Kluwer, 1999.

[5] J. Doyle, K. Bowyer, and P. Flynn. Variation in accuracy of textured contact lens detection. In *Proceedings of the 6th*

IEEE International Conference of Biometrics: Technology, Applications, and Systems (BTAS'13). IEEE, 2013.

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

[7] Johnson&Johnson. Acuvue2 colours, April 2013. `http://www.acuvue.com/products-acuvue-2-colours`.

[8] LG. Lg 4000 camera, October 2011. `http://www.lgiris.com`.

[9] G. L. Marcialis, A. Lewicke, B. Tan, P. Coli, F. Roli, S. Schuckers, D. Grimberg, A. Congiu, and A. Tidu. First international fingerprint liveness detection competition - livdet 2009, 2009.

[10] A. Pacut and A. Czajka. Aliveness detection for iris biometrics. In *Carnahan Conferences Security Technology, Proceedings 2006 40th Annual IEEE International*, pages 122–129, Oct 2006.

[11] L. Thalheim, J. Krissler, and P.-M. Ziegler. Biometric Access Protection Devices and their Programs Put to the Test. In *c't Magazine 11/2002, p. 114*, November 2002.

[12] University of Notre Dame Computer Vision Research Lab. University of Notre Dame Contact Lenses Detection 2013 (NDCLD'13) Dataset, December 2013. `http://www.nd.edu/~cvrl/CVRL/Data_Sets.html`.

[13] Z. Wei, X. Qiu, Z. Sun, and T. Tan. Counterfeit iris detection based on texture analysis. In *Pattern Recognition, 2008. ICPR 2008. 19th International Conference on*, pages 1–4, Dec 2008.

[14] D. Willis and M. Lee. Six biometric devices point the finger at security. biometrics under our thumb. *Network Computing*, June 1998.

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