LivDet-Iris 2015 – Iris Liveness Detection Competition 2015

David Yambay
Clarkson University
Potsdam, NY USA
yambayda@clarkson.edu

Brian Walczak
Clarkson University
Potsdam, NY USA
walczabg@clarkson.edu

Stephanie Schuckers
Clarkson University
Potsdam, NY USA
sschucke@clarkson.edu

Adam Czajka
NASK and WUT
Warsaw, Poland
aczajka@elka.pw.edu.

Abstract

Presentation attacks such as printed iris images or patterned contact lenses can be used to circumvent an iris recognition system. Different solutions have been proposed to counteract this vulnerability with Presentation Attack Detection (commonly called liveness detection) being used to detect the presence of an attack, yet independent evaluations and comparisons are rare. To fill this gap we have launched the first international iris liveness competition in 2013. This paper presents detailed results of its second edition, organized in 2015 (LivDet-Iris 2015). Four software-based approaches to Presentation Attack Detection were submitted. Results were tallied across three different iris datasets using a standardized testing protocol and large quantities of live and spoof iris images. The Federico Algorithm received the best results with a rate of rejected live samples of 1.68% and rate of accepted spoof samples of 5.48%. This shows that simple static attacks based on paper printouts and printed contact lenses are still challenging to be recognized purely by software-based approaches. Similar to the 2013 edition, printed iris images were easier to be differentiated from live images in comparison to patterned contact lenses.

1. Introduction

Iris recognition has been shown to be susceptible to presentation attacks in the form of printed images of the iris or the obscuring of the natural iris pattern through wearing patterned contact lenses. Presentation Attack Detection (PAD) has been proposed as solution to these vulnerabilities. Presentation Attack Detection (commonly called liveness detection) is based on the principle that additional information can be garnered above and beyond the data procured by a standard verification system to verify if an image is authentic.

These schemes are split into two categories, hardware-based and software-based implementations. Hardware-based systems make use of additional sensors to take measurements to detect a presentation attack. Software-based systems make use of a variety of image processing algorithms to take additional measurements from collected iris images to detect presentation attacks. Both of these systems classify the input images as either live or fake images.

The First International Fingerprint Liveness Detection Competition LivDet 2009 [1] provided an initial assessment of **s**oftware systems based on the fingerprint image only. The second Liveness Detection Competition (LivDet 2011 [2]) also included integrated system testing. The third Liveness Detection Competition (LivDet 2013) expanded on previous competitions with the inclusion of an iris component. LivDet 2013 was split into two separate competitions, LivDet 2013 - Fingerprint and LivDet 2013 – Iris [3,4]. The first half of LivDet 2015, LivDet-Finger 2015 is completed and showed the growth improvement in PAD performance and increased participation [5].

The competition design and results for the submitted algorithms for LivDet-Iris 2015 are summarized in this paper. Section 2 delves into the background of iris presentation attack detection. Section 3 discusses the methods and the protocol used to evaluate the algorithms submitted for testing. Section 4 explains the results of the competition. Section 5 discusses conclusions from the algorithms and future thoughts on the LivDet competitions.

2. Background

Vulnerabilities in iris have been known to exist for over a decade. Daugman proposed an iris liveness technique based on the 2-D Fourier Transform in 2003 that showed patterned contact lenses had points in the Fourier spectrum that were not in the natural iris [6]. In 2006, Pacut and Czajka examined the weakness of iris systems to spoof attacks through a survey of different types of forgery attacks as well as examining solutions to these forms of attacks [7]. Z. Wei in 2008 examined three different anti-spoofing iris measures which gave new results on the detection of counterfeit irises [8]. Czajka later released an article on a database of iris printouts and their applications [9]. More recent work includes that by Galbally who combined frequency analysis with other quality features to successfully detect printed iris images [10]. Sequeira et al.

employed a similar method of combining frequency analysis and quality features for both printed iris and patterned contact lenses [11]. Additionally, Czajka proposed to use pupil dynamics as a liveness indicator [12].

The two attack scenarios employed in previous research are the main attack methods for iris systems. Sample images can be seen in Figure 1. In the printed iris case, an image of the subject's iris is printed onto paper and presented to the iris system as though it were a genuine iris image. Patterned contacts act to obscure one's natural iris pattern through the use of patterns printed onto a contact lens that will cover one's natural iris image and block it from the system.

Algorithms and hardware are both developing over time to better differentiate if an iris image presented to a system is from the genuine user who is not attempting to interfere with the function of the biometric system.

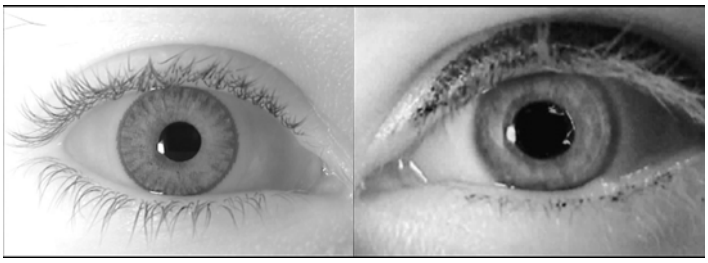

Figure 1: Example iris attack methods. A patterned contact obscuring a natural iris patterns (left), and a printed iris image (right).

3. Experimental Protocol and Evaluation

The competition for LivDet-Iris 2015 continued to focus on iris algorithms. The protocol for LivDet-Iris 2015 competition will be outlined in this section.

3.1 Participants

The competition was open to all academic and industrial institutions. All participants are required to sign a database release agreement that outlines the usage limitations of data made available. Participants then download the training datasets to create their algorithms. Participants are allowed to submit as an anonymous submission and not have their organization's name in the publication. Table 1 displays the participant names and the corresponding algorithm names as they are used throughout the paper. A total of four algorithms were submitted to LivDet 2015-Iris.

Participant Name	Algorithm Name
Anonymous 0	Anon0
Anonymous1	Anon1
Anonymous2	Anon2
University of Naples Federico II	Federico

Table 1: Name of participants and submitted iris algorithms

3.2. Datasets

The database for LivDet-Iris 2015 consisted of three datasets. Similar to 2013, presentation attacks are represented with patterned contact lenses and printed iris images. Clarkson University prepared images using both presentation attacks and Warsaw University of Technology prepared a dataset with only printed iris images. Patterned contact lenses from the LG and Dalsa datasets consisted of 20 different patterned contacts. 15 of these patterns were present only in the training set and were "known" spoof attack types, whereas the remaining 5 patterned contact types were only present in the testing datasets as "unknown" spoof attacks. Table 2 displays the patterned contact types present in the LG and Dalsa datasets.

Contact Number	Patterned Contact Type	Color
1	Expressions Colors	Brown
2	Expressions Colors	Jade
3	Expressions Colors	Blue
4	Expressions Colors	Hazel
5	Air Optix Colors	Brown
6	Air Optix Colors	Green
7	**Air Optix Colors**	**Blue**
8	Freshlook Colorblends	Brilliant Blue
9	Freshlook Colorblends	Brown
10	Freshlook Colorblends	Honey
11	Freshlook Colorblends	Green
12	Freshlook Colorblends	Sterling Gray
13	Freshlook One-Day	Green
14	Freshlook One-Day	Pure Hazel
15	Freshlook One-Day	Gray
16	Freshlook One-Day	Blue
17	**Air Optix Colors**	**Gray**

18	**Air Optix Colors**	**Honey**
19	**Expressions Colors**	**Blue Topaz**
20	**Expressions Colors**	**Green**

Table 2: Patterned Contact Types from LG and Dalsa datasets. **Unknown patterns in Bold**

3.2.1 Clarkson LG Dataset

The first subset within Clarkson LivDet 2015 uses an LG IrisAccess EOU2200 camera for capture of the irises. There are a total of 828 live images over 45 subjects, 1152 patterned contact lens images from 20 contact types and 7 subjects, and 1746 printed images. The training subset contains 450 live images, 576 patterned contact lens images, and 846 printed images. The testing subset contains 378 live images, 576 patterned contact lens images, and 900 printed images. The printed images use a variety of configurations including 1200 dpi versus 2400 dpi printouts, contrast adjustment versus raw images, pupil hole versus no pupil hole, and glossy paper versus matte paper. The printouts came from images collected from both the LG iris camera and from the Dalsa camera. Figure 2 shows sample images from the LG dataset.

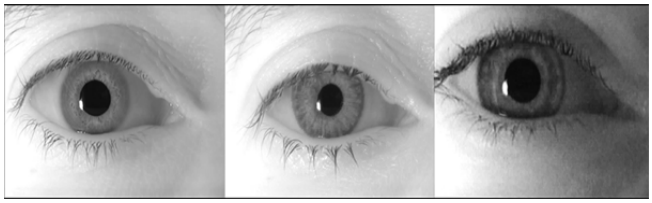

Figure 2: Images from LG Dataset. Left to right; Live, Patterned, Printed

3.2.2 Clarkson Dalsa Dataset

The second subset within Clarkson LivDet 2015 uses a Dalsa camera for iris capture. The camera is modified to capture in the NIR spectrum similar to commercial iris cameras. It captures a section of the face of each subject that includes both eyes. The eyes are then cropped out of the images to create the subset. There are a total of 1078 live images, 1431 patterned contact lens images, and 1746 printed images are in this dataset. The training subset contains 700 live, 873 patterned contact lens, and 846 printed images. The testing subset contains 378 live, 558 patterned contact lens, and 900 printed images. Figure 3 gives examples images from the Dalsa dataset.

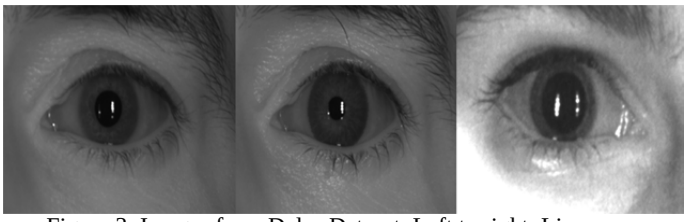

Figure 3: Images from Dalsa Dataset. Left to right; Live, Patterned, Printed

3.2.3 Warsaw IrisGuard Dataset

The Warsaw group has followed up on their dataset created in 2013, with a larger scale dataset. This new dataset is based on live irises captured with a commercial iris system, IrisGuard AD100, with the liveness detection functionality intentionally turned off. Each live eye that is captured also has its printed counterpart. This utilizes printed images which are created using a Lexmark 534dn printer. This device is a semi-professional color laser printer which creates printouts of a resolution up to 1200 dpi. The printer is set to use two different modes for the printouts: color printing and black and white printing. In addition, pupil holes are added in order to have a live user presented behind the printouts. This is to counter a camera that searches for specular reflection of a live cornea. With this dataset being a larger scale than in 2013, the testing set uses 100 unique irises for 2002 live samples, and 100 unique irises for 3890 printed samples. The training set also contains 852 live iris images as well as 815 printed samples. Figure 4 showcases images from the Warsaw database.

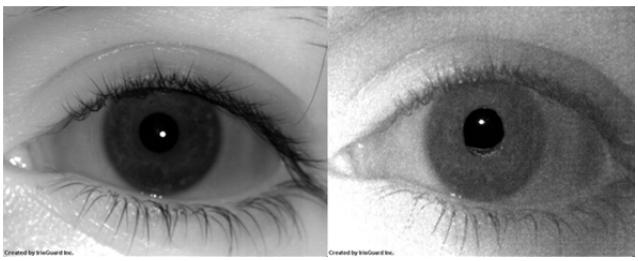

Figure 4: Images from Warsaw Dataset. Left to right; Live, Printed

	Training			Testing		
	Live	Patterned Contacts	Printed	Live	Patterned Contacts	Printed
Clarkson LG	450	576	846	378	576	900
Clarkson Dalsa	700	873	846	378	558	900
Warsaw	852	N/A	815	2002	N/A	3890

Table 3: Dataset composition for LivDet-Iris 2015 competition

3.3. Algorithm Submission

The algorithm submission for LivDet 2015 is the same as all previous competitions. Each submitted algorithm is given as a Win32 console application LIVENESS_XYZ.exe, unless otherwise arranged with the testing committee prior to submission, with the following list of parameters:

Ndataset: Identification Number of dataset to be analyzed, i.e., 1=LG, 2=Dalsa, 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 "return" 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 their algorithm by the training set made available to them. Participants could also choose to publish a description and/or source code of the algorithm, but this was not mandatory.

3.4. Performance Evaluation

Each of the algorithms returned an integer value representing a percentage of posterior probability of the live class given the image or a degree of "liveness" normalized in the range 0 to 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 is used to calculate Attack Presentation Classification Error Rate (APCER) and Normal Presentation Classification Error Rate (NPCER) error estimators, where

- APCER is the rate of misclassified spoof images (spoof called live)
- NPCER is the rate of misclassified live images (live called spoof)

Both APCER and NPCER are calculated for each dataset separately, as well as the average values across all datasets. To select a winner the average of APCER and NPCER was calculated for each participant across datasets. The weight of importance between APCER to NPCER will change based on use case scenario, such as low NPCER is more important for low security implementations such as unlocking phones, however low APCER is more important for high security implementations. Due to this treating APCER and NPCER as equal is used in this competition.

4. Results and Discussion

Four algorithms were successfully submitted and completed the competition at the time of submission of this paper. Table 4 below details the error rates of iris algorithms.

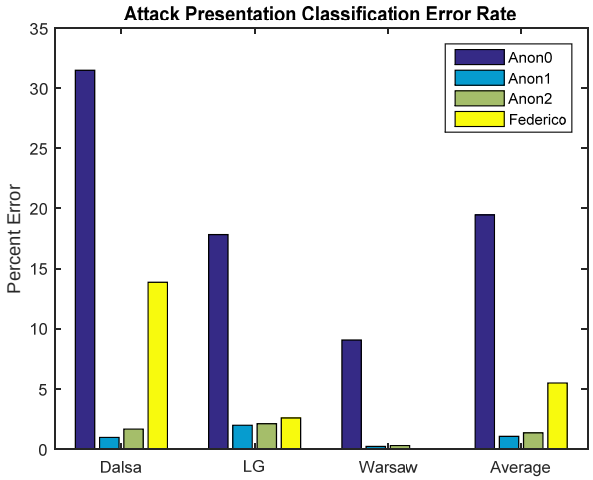

Figure 5: Rate of Misclassified Spoof Images

Figures 5 and 6 show the results from the iris algorithms. Among the four algorithms, Federico performed the best with an average APCER of 5.48% and a NPCER of 1.68%. Anon2 performed at a close rate to the Federico algorithm with an average APCER of 1.34% and a NPCER of 7.53%. Figure 7 shows the combined APCER and NPCER from each algorithm as an average of the APCER and NPCER.

In general as with LivDet 2013, error rates were lower on the Warsaw dataset. The printed iris images tend to be easier for the algorithms to differentiate from live images in comparison to patterned contact lenses. The error rates for printed images are near 0 whereas the error rates for

Algorithm	Dalsa		LG		Warsaw		Average	
	APCER	NPCER	APCER	NPCER	APCER	NPCER	APCER	NPCER
Anon0	31.48	13.23	17.82	11.64	9.05	3.25	19.45	9.37
Anon1	0.96	11.9	1.97	13.23	0.21	2.35	1.05	9.16
Anon2	1.65	10.85	2.1	10.85	0.28	0.9	1.34	7.53
Federico	13.85	3.18	2.58	1.85	0	0	5.48	1.68

Table 4: Error Rates by Dataset

patterned lenses are much larger in comparison. The Federico algorithm on the LG dataset had an APCER of 6.25% against patterned irises and an APCER of 0.22% against printed iris images.

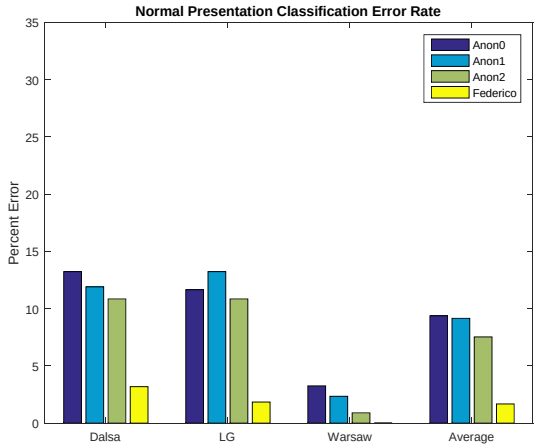

Figure 6: Rate of Misclassified Live Images

Figure 7: Combined Average APCER and NPCER

Examining the known vs. unknown patterned contact lenses produced a curious output. While all competitors had a large number of known patterned contact types to train their algorithms, error rates were almost universally better on the unknown patterned contact lenses than the known. The only case in which the unknown patterned contact lenses had a higher error rate than the known is for Federico on the LG dataset. Further investigation is warranted as to why the unknown patterned contacts were more easily identified than the known.

Overall the results have shown tremendous improvement over those seen in LivDet 2013- Iris. Examining the best performing algorithm for the Warsaw dataset in 2013 to the results from 2015, it can be seen that error rates have declined heavily among algorithms. Given that the Warsaw dataset uses the same data collection protocol among both LivDet competitions just with additional data for 2015, results are directly compared. The best result for NPCER in Warsaw was 5.23%. In LivDet 2015, all 4 submitted algorithms have significantly lower error rates of 3.25%, 2.35%, 0.9%, and 0%. The best results for APCER for Warsaw in 2013 was 0.65%. In LivDet 2015, three of the four submitted algorithms have lower error rates of 0.21%, 0.28%, and 0%.

Examining the overall average error rates across all datasets there is also distinct decreases in error rates. The overall average APCER for the best algorithm in 2013 was 5.716%. Three of the algorithms submitted for LivDet 2015 have lower error rates. The overall average NPCER for the best algorithm in 2013 was 28.56%. All four algorithms in LivDet 2015 presented a lower number of errors. These results are shown in figure 8.

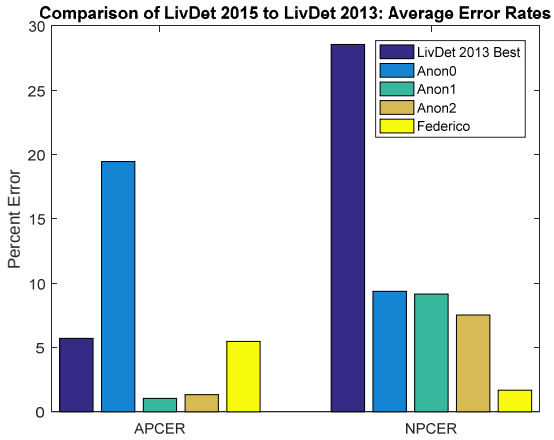

Figure 8: Average Error Rate for Best Algorithm in 2013 vs Average Error Rates for All Algorithms in 2015

5. Conclusions

LivDet-Iris 2015 is the second public assessment of algorithm-based iris presentation attack detection. This competition has shown growth from LivDet-Iris 2013 with the addition of a fourth algorithm. The datasets are available upon request and are reaching beyond the competition as large numbers of requests for data have been received even once the competition is complete.

The best results from the competition were shown by the Federico algorithm and these error rates have shown tremendous improvement over the span of two years.

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] Marcialis, Gian Luca, et al. "First international fingerprint liveness detection competition—livdet 2009." *Image Analysis and Processing–ICIAP 2009*. Springer Berlin Heidelberg, 2009. 12-23.

[2] Yambay, David, et al. "LivDet 2011—Fingerprint liveness detection competition 2011." *Biometrics (ICB), 2012 5th IAPR International Conference on*. IEEE, 2012.

[3] Ghiani, Luca, et al. "Livdet 2013 fingerprint liveness detection competition 2013." *Biometrics (ICB), 2013 International Conference on*. IEEE, 2013.

[4] Yambay, David, et al. "Livdet-iris 2013-iris liveness detection competition 2013." *Biometrics (IJCB), 2014 IEEE International Joint Conference on*. IEEE, 2014.

[5] Mura, Valerio, et al. "LivDet 2015 Fingerprint Liveness Detection Competition 2015." *Biometrics Theory, Applications and Systems (BTAS), 2015 IEEE 7th International Conference on*. IEEE, 2015.

[6] Daugman, John. "Demodulation by complex-valued wavelets for stochastic pattern recognition." *International Journal of Wavelets, Multiresolution and Information Processing* 1.01 (2003): 1-17.

[7] Pacut, Andrzej, and Adam Czajka. "Aliveness detection for iris biometrics."*Carnahan Conferences Security Technology, Proceedings 2006 40th Annual IEEE International*. IEEE, 2006.

[8] Wei, Zhuoshi, et al. "Counterfeit iris detection based on texture analysis."*Pattern Recognition, 2008. ICPR 2008. 19th International Conference on*. IEEE, 2008.

[9] Czajka, Adam. "Database of iris printouts and its application: Development of liveness detection method for iris recognition." *Methods and Models in Automation and Robotics (MMAR), 2013 18th International Conference on*. IEEE, 2013.

[10] Galbally, Javier, et al. "Iris liveness detection based on quality related features." *Biometrics (ICB), 2012 5th IAPR International Conference on*. IEEE, 2012.

[11] Sequeira, Ana F., Juliano Murari, and Jaime S. Cardoso. "Iris liveness detection methods in the mobile biometrics scenario." *Neural Networks (IJCNN), 2014 International Joint Conference on*. IEEE, 2014.

[12] Czajka, Adam. "Pupil dynamics for iris liveness detection." *IEEE Transactions on Information Forensics and Security* 10.4 (2015): 726-735.