LivDet 2011 – Fingerprint Liveness Detection Competition 2011

David Yambay[1], Luca Ghiani[2], Paolo Denti[2], Gian Luca Marcialis[2], Fabio Roli[2], S Schuckers[1]
[1] Clarkson University - Department of Electrical and Computer Engineering – USA,
{yambayda, sschucke}@clarkson.edu
[2] University of Cagliari - Department of Electrical and Electronic Engineering – Italy,
l.ghiani@tiscali.it, paolode@hotmail.com, (roli@diee.unica.it marcialis@diee.unica.it

Abstract

"Liveness detection", a technique used to determine the vitality of a submitted biometric, has been implemented in fingerprint scanners in recent years. The goal for the LivDet 2011 competition is to compare software-based fingerprint liveness detection methodologies (Part 1), as well as fingerprint systems which incorporate liveness detection capabilities (Part 2), using a standardized testing protocol and large quantities of spoof and live fingerprint images. This competition was open to all academic and industrial institutions which have a solution for either software-based or system-based fingerprint vitality detection problem. Five submissions across the two parts of the competition resulted in successful completion. These submissions were: Chinese Academy of Sciences Institute of Automation (CASIA), Federico II University (Federico) and Dermalog Identification SystemsGmbH (Dermalog) for Part 1: Algorithms, and GreenBit and Dermalog for Part 2: Systems. Part 1 was evaluated using four different datasets. The best results were from Federico on the Digital Persona dataset with error for live and spoof detection of 6.2% and 11.61% respectively. The best overall results for Part 1 were Dermalog with 34.05 FerrFake and 11.825% FerrLive. Part 2 was evaluated using live subjects and spoof finger casts. The best results were from Dermalog with an error for live and spoof of 42.5% and 0.8%, respectively.

1. Introduction

Biometrics has been an expanding industry in recent years and provides security through identifying a person based on physiological or behavior characteristics. However, it has been shown that biometric systems are vulnerable to spoof attacks by artificial fingerprint casts made of materials such as PlayDoh, silicone, or latex.

There have been numerous methods proposed to solve the susceptibility of fingerprint devices to attacks by spoof fingers. 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 program to provide additional security. Hardware-based systems use additional sensors to gain measurements outside of the fingerprint image itself to detect liveness. Software-based systems use image processing algorithms to gather information directly from the collected fingerprint to detect liveness. These systems classify images as either live or fake

A standard assessment of fingerprint liveness detection methods has been lacking in the industry by which different organizations can be compared. The First International Fingerprint Liveness Detection Competition – LivDet 2009, provided an initial assessment of software systems based on the fingerprint image only, but did not address into integrated systems. The second Liveness Detection Competition 2011 (LivDet 2011) was created in order to ascertain the current state of the art in liveness detection, including integrated system testing. LivDet 2011 was open to all academic and industrial institutions and contained two parts: evaluation of software-based systems in Part 1: Algorithms, and evaluation of integrated systems in Part 2: Systems.

In this paper, the competition design and results of the submitted algorithms and systems are summarized. Section 2 of this paper describes the background of spoofing and liveness detection. Section 3 delves into the protocol behind the evaluation of algorithms and systems. Section 4

discusses the results of the competition and conclusions that are drawn from the results.

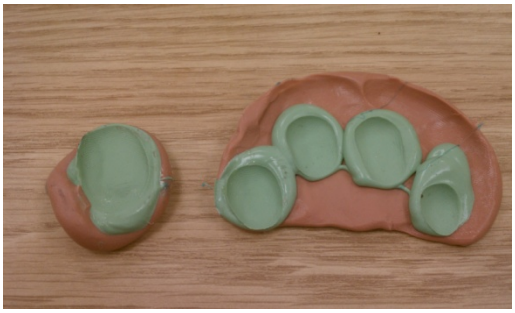

Figure 1. Negative impression of five fingers using consensual method.

Figure 2. Latex spoof on finger

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 [1]. This research was approached again in 2000-2002 by multiple institutions including; Putte and Kuening as well as Matsumoto (et al.) [2-3]. The research presented by these researchers looked at the vulnerability of spoofing. In 2001, Kallo (et al.) looked at a hardware solution to Liveness Detection while in 2002; Schuckers delved into using software approaches for Liveness Detection [4-5].

There are two general forms of creating artificial fingers, the cooperative method and non-cooperative method. In the cooperative method the subject pushes their finger into a malleable material such as dental impression material, plastic, or wax creating a negative impression of the fingerprint as a mold, see Figure 1. The mold is then filled with a material, such as gelatin, PlayDoh or silicone. This cast can be used to represent a finger from a live subject, see Figure 2.

The non-cooperative method involves enhancing a latent fingerprint left on a surface, digitizing it through the use of a photograph, and finally printing the negative image on a transparency sheet. This printed image can then be made into a mold, for example, by etching the image onto a printed circuit board which can be used to create the spoof cast.

The challenge to fingerprint recognition systems is the ability to detect if a presented fingerprint is from a live person or an artificial finger. Systems are being upgraded to incorporate liveness detection solutions that will be able to detect if the submitted probe is a spoof or live finger.

Liveness detection can be incorporated into a system through the addition of hardware components to the capture device that can search for traits in the fingerprint through the use of blood pressure, electrocardiogram, temperature or other methods. Liveness detection can also be implemented through the use of algorithms that are added to the system. This method looks to see if there are features inside of the fingerprint images to determine liveness.

There are many solutions that have been proposed to solve the vulnerability of spoofing [6-8]. 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 by evaluating the performance of both algorithms and systems using a standardized experimental protocol. The value in the addition of system testing is to assess the performance of liveness detection approaches tuned to a specific commercial device.

3. Experimental Protocol and Evaluation

The competition features two distinct parts; Part 1: Algorithms and Part 2: Systems, with separate protocols designed for each part. Each part contains their own constraints necessary to eliminate the variability that may be present across algorithms or systems. The design of the experiment will be discussed in detail in this section also outlining the constraints placed on entrants for each part.

3.1 Participants

The competition is open to all academic and industrial institutions. Upon registration, each participant is required

to sign a database release agreement detailing the proper usage of data made available through the competition. Participants are then given a database access letter with a username and password to access the server to download the training data. The fingerprint data and its release to participants were approved by the Institutional Review Board at Clarkson University.

3.2 Part 1: Algorithm Data Set

The dataset for Part 1: Algorithms consists of images from four different devices; Biometrika, Digital Persona, Italdata and Sagem. There are 4000 images for each of these devices, 2000 live images and 2000 spoof images (400 of each of 5 spoof materials). The spoof materials used for this experiment were gelatine, latex, PlayDoh, silicone and wood glue for Digital Persona and Sagem (400 each) and gelatine, latex, ecoflex (platinum-catalysed silicone), silicone and wood glue for Biometrika and ItalData (400 each)The dataset of 4000 images per scanner were divided into two equal datasets, training and testing. Details are described in Table 1 and Table 2. Live images came from 400 fingers from 50 people for Biometrika and ItalData datasets, 200 fingers representing 100 people for, Digital Persona dataset, and 112 fingers from 56 people for Sagem dataset. Spoof images come from approximately 100 fingers representing 50 people for the Digital Persona and Sagem Datasets and 81 fingers representing 22 subjects for the Biometrika and ItalData datasets.

The spoof images were collected using the consensual method that was described earlier. After the competition is completed, the entire dataset will be made available to those who sign the proper data release agreement. Figure 3 below shows examples of images used in the experiment.

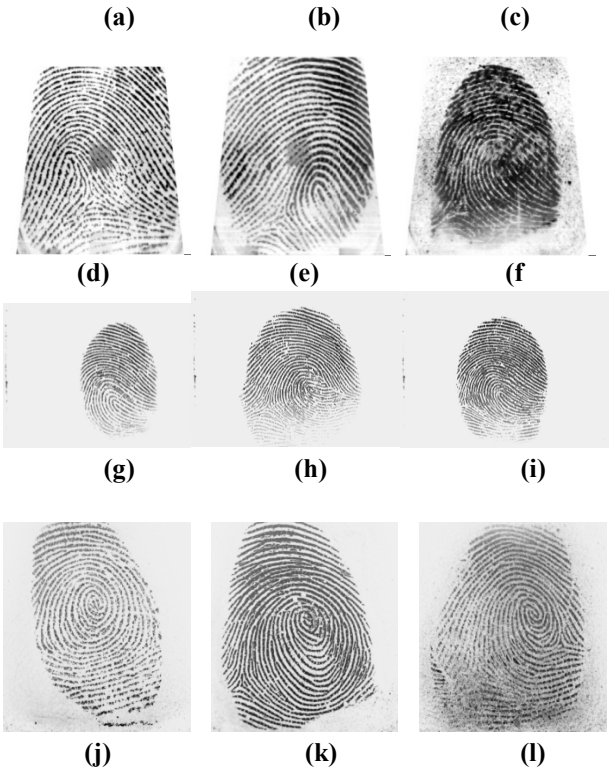

Figure 3. Examples of fake fingerprint images, from Biometrika (a) latex, (b) gelatin, (c) silicone; from Digital Persona: (d) latex, (e) gelatin, (f) silicone; from Italdata: (g) latex (h) gelatin (i) silicone; from Sagem: (j) latex (k) gelatin (l) silicone.

Table 1. Device characteristics for Part 1 datasets

Dataset	Sensor	Model No.	Resolution (dpi)	Image Size
#1	Biometrika	FX2000	500	315x372
#2	Digital Persona	4000B	500	355x391
#3	ItalData	ET10	500	640x480
#4	Sagem	MSO300	500	352x384

Table 2. Training and test set for Part 1 datasets

Dataset	Sensor	Live Training Samples	Live Testing Samples	Number of Fingers
#1	Biometrika	1000	1000	200
#2	Digital Persona	1000	1000	200
#3	ItalData	1000	1000	200
#4	Sagem	1000	1000	112

3.3 Algorithm Submission

The algorithm submission for LivDet 2011 uses the same structure as LivDet 2009. As stated for LivDet 2009 each submitted algorithm must be a Win32 console application with the following list of parameters:
LIVENESS_XYZ.exe [ndataset] [inputfile] [outputfile]

Each parameter, specified in Table 3, and related to the data set configuration must be set before submission. Only Win32 console applications with the above characteristics will be accepted for the competition [9].

Table 3. Formats of submission requirements

Arguments	Description
LIVENESS_XYZ.exe	It is the executable name, where XYZ is the identification number of the participant. LIVENESS_XYZ.exe Format : Win32 console application (.exe)
[ndataset]	It is the identification number of the data set to analyse. Legend: 1=Biometrika, 2=Digital Persona, 3=Italdata, 4=Sagem
[inputfile]	Txt file with the List of images to analyse. Each image is in RAW format (ASCII)
[outputfile]	Txt file with the output of each processed image, in the same order of inputfile. The output is 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). In the case that the algorithm has not been able to process the image, the correspondent output must be -1000 (failure to enroll).

3.4 Part 2: Systems Submission

In this component, participants were asked to ship a fingerprint system which captures a fingerprint image as well as outputs a liveness detection score. Three *spoof recipe and methods were provided* upon registration (Play-doh, Gelatin, Silicone). These were used for testing the system. Two additional unspecified methods were also used in testing (Body Double, Latex). The requirements for installation are that the system will be run on a Windows XP 32-bit system, that the file will be an ".exe" or similar executable and that the system will use either a USB or Firewire connection.

The system is required to output the collected image and a liveness score normalized in the range of 0 and 100 (100 is the maximum degree of liveness, 0 spoof). In the case that the algorithm has not been able to process the image, the correspondent output must be -1000 (failure to enroll).

Laboratory staff systematically attempted to spoof the system and collect corresponding live data. For each submitted system, 500 live attempts from 50 people (5 images each of R1 and R2 fingers) were performed, as well as 750 spoof attempts for five different materials (Play-doh, Gelatin, Silicone, Body Double, and Latex). Gelatin, PlayDoh and Silicone recipes were supplied to the participants. Body Double and Latex were recipes unknown to the participants. This experiment used molds from 25 subjects with 2 fingers each and 3 images per finger per spoof. The same physical spoof fingers were placed on both scanners. A spoof image was collected on one scanner and then the next scanner alternating between the two scanners.

3.5 Performance Evaluation

The parameters adopted for the performance evaluation will be the following:

Evaluation per sensor/system

- *Frej_n*: Rate of failure to enroll for the sub-set *n*.
- *Fcorrlive_n*: Rate of correctly classified live fingerprints for sub-set *n*.
- *Fcorrfake_n*: Rate of correctly classified fake fingerprints for sub-set *n*.
- *Ferrlive_n*: Rate of misclassified live fingerprints for sub-set *n*.
- *Ferrfake_n*: Rate of misclassified fake fingerprints for sub-set *n*.
- *ET*: Average processing time per image

.

4 Results and Discussion

Three algorithm submissions and two system submissions successfully completed the competition at the time of submission of this paper: Dermalog, Federico and CASIA for Part 1, Dermalog and GreenBit for Part 2.

4.1 Part 1: Algorithms

The rate at which each algorithm produced a false rejection of a live subject (*Ferrlive*) is shown in Figure 4 and summarized in Table 4. The rate at which each algorithm produced a false acceptance of a spoof image is shown in Figure 5 and summarized in Table 4 (*FerrFake*).

The threshold value for determining liveness was set at 50. This threshold is used for the values given for *Ferrfake* and *FerrLive*. The liveness values were relatively varied and thus in many cases, changing the threshold would have an effect on the false accept and false reject rates. Figure 6 shows the change in error rates as a function of the threshold for all algorithms across all datasets and Figure 7 for the Sagem Dataset.

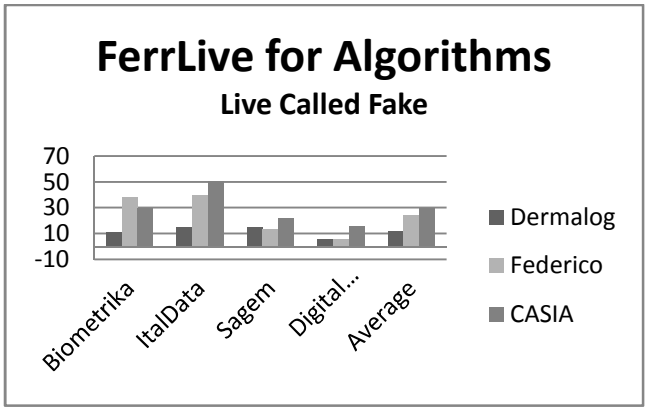

Figure 4: *FerrLive* (%) for each dataset, (left to right) for two submitted algorithms (Dermalog, Federico).

The best balance of spoof and live detection was the Federico algorithm on the Digital Persona dataset at only 6.2% *FerrLive* and 11.61% *FerrFake*. The Sagem dataset had a low *FerrLive* for all three algorithms. The overall classification rate is 27.2% for CASIA, 25.5% for Federico and 22.9% for Dermalog.

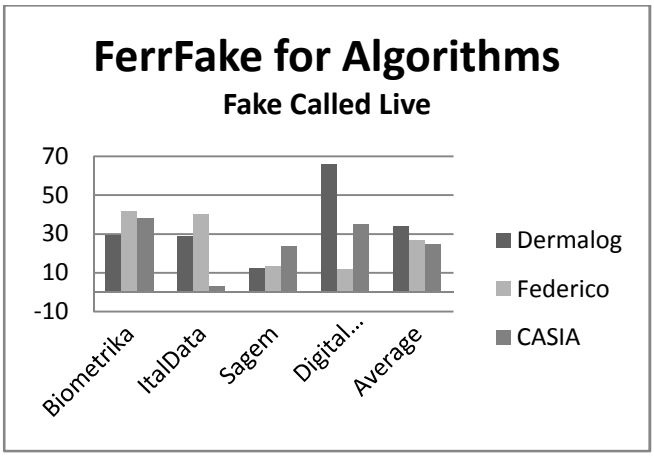

Figure 5. *FerrFake (%)* for each dataset (left to right) for two submitted algorithms (Dermalog, Frederico)

Algorithm (columns)	Dermalog	Federico	CASIA
Datasets (rows)	*FerrLive*	*FerrLive*	*FerrLive*
Biometrika	11%	38%	29.7%
ItalData	15.10%	39.90%	50.6%
Sagem	15.10%	13.80%	22.1%
Digital Persona	66%	6.20%	16.1%
Average	35.30%	26.60%	29.625%

Algorithm (columns)	Dermalog	Federico	CASIA
Datasets (rows)	*FerrFake*	*FerrFake*	*FerrFake*
Biometrika	29%	42%	38.1%
ItalData	28.50%	40.10%	2.8%
Sagem	12.50%	13.10%	23.6%
Digital Persona	6.20%	11.60%	34.7%
Average	17.60%	24.50%	24.8%

Table 4: *FerrFake* and *FerrLive* for submitted algorithms

Biometrika and ItalData both had high *FerrFake* rates for Dermalog and Federico algorithms. *FerrLive* rate does not vary for Dermalog algorithm, whilst it gets worse for Federico algorithm. This can be due to several reasons, which should be better investigated in a future paper, maybe related to the characteristics of these data sets.

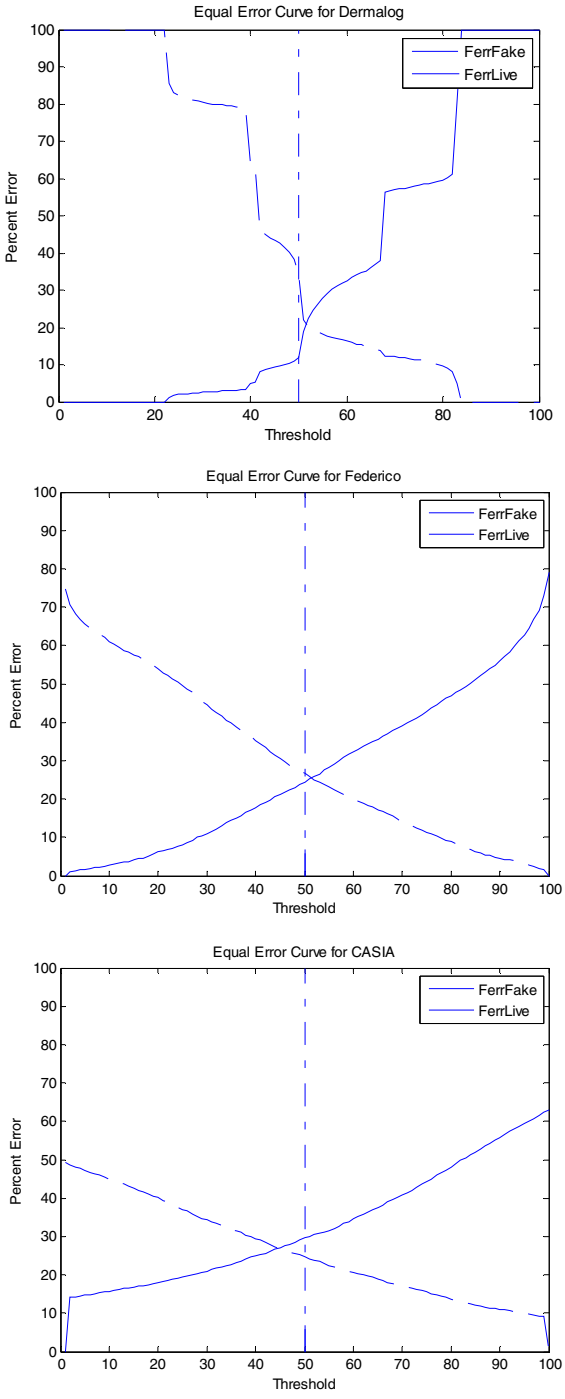

Figure 6. Change in error rate: *FerrFake* (dashed) and *FerrLive* (solid) compared to thresholds for both algorithms (Dermalog top, Frederico middle, CASIA bottom) across all datasets

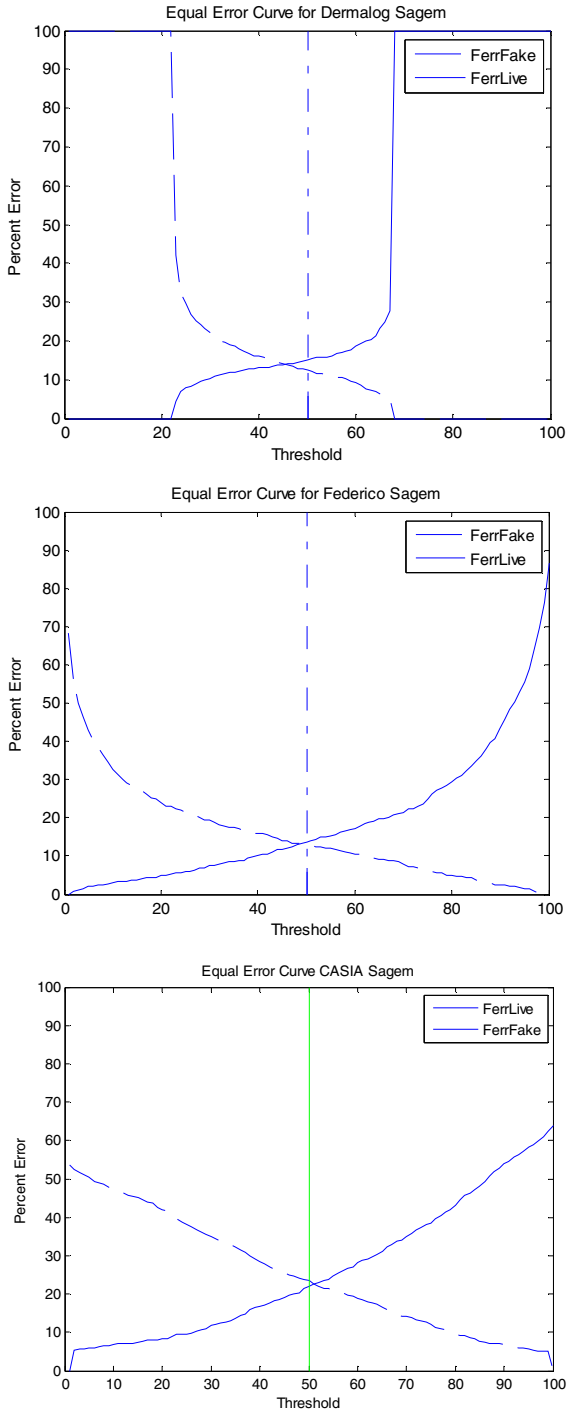

Figure 7. Change in error rate: *FerrLive* (solid) and *FerrFake* (dashed) compared to thresholds for both algorithms (Dermalog top, Federico middle, CASIA bottom) on Sagem Dataset

The failure to enroll rate was 0% for all algorithms on all datasets. The average processing time per image was about ten times lower for Dermalog as Federico and about two times lower than CASIA. Dermalog processed images at an

average elapsed time of 0.28 seconds per image. Federico processed images at an average elapsed time of approximately 3 seconds per image. CASIA processed images at an average elapsed time of approximately 0.6 seconds.

4.2 Part 2: Systems

FerrLive and *FerrFake* for the two submitted systems can be found in Figure 8 below with the data summarized in Table 5. Dermalog performed at a *FerrLive* of 42.5% and a *FerrFake* of 0.8%. GreenBit performed at a *FerrLive* of 38.8% and a *FerrFake* of 39.47%. Both systems had high *FerrLive* scores.

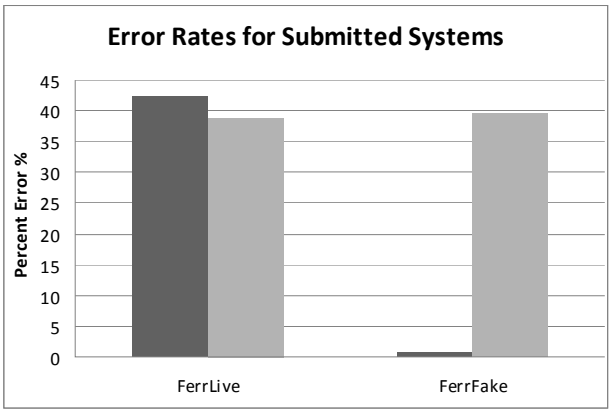

Figure 8. *FerrLive* and *FerrFake* for submitted systems for Dermalog, GreenBit

Table 5. *FerrLive* and *FerrFake* for submitted systems

Submitted Systems	*FerrFake*	*FerrLive*	*FerrFake Known*	*FerrFake Unknown*
Dermalog	0.8%	42.5%	0.4%	1.3%
Greenbit	39.5%	38.8%	19.1%	70%

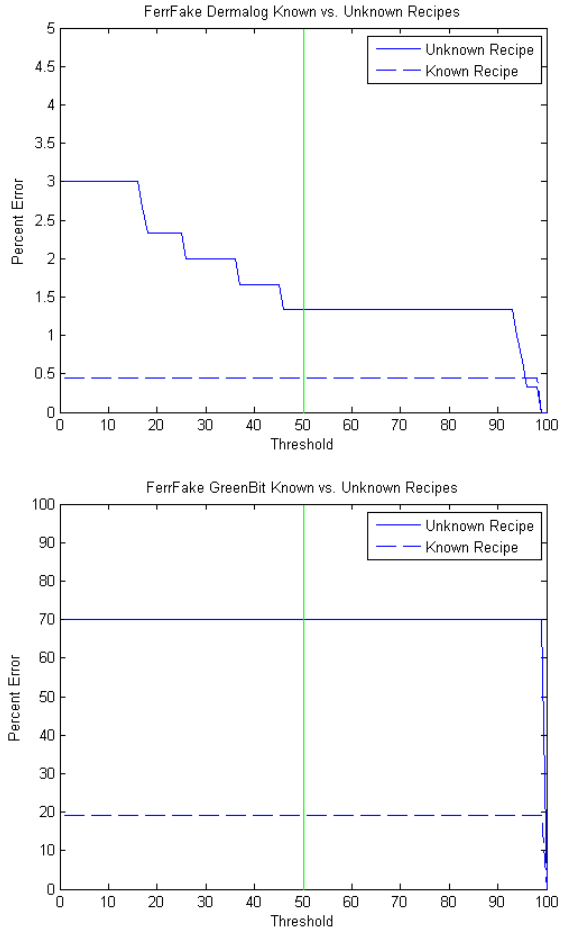

Figure 9: *FerrFake* for unknown vs. known recipes for Dermalog (top) and Greenbit (bottom)

Figure 9 shows the *FerrFake* rate for both systems for the unknown recipes and known recipes. It can be seen both systems performed better against known recipes.

The software for each system had built-in auto-detect that would allow the system to keep attempting to detect an image until either an image was detected or the scan manually stopped. The frej_n was 0% for both systems because eventually all images were detected.

5 Conclusions

In summary, LivDet 2011 is the second international public competition for software-based fingerprint liveness detection and first public assessment of system-based fingerprint liveness detection. Entries were submitted from a total of five participants demonstrating the state-of-the art in fingerprint liveness. LivDet 2011 evaluates (1) software

approaches applicable to four fingerprint sensors represented in the training data as well (2) embedded hardware/software systems for liveness detection specialized to a specific fingerprint sensor. The best results shown were by Dermalog in Part 1 and Dermalog again in Part 2. It is hoped that this competition will be continued in order to continually understand and promote the state of the art in liveness detection. Results of this competition are not reflective of performance for spoof attacks not included in this study. Creating effective liveness detection methodologies is an important step in minimizing the vulnerability of spoof attacks.

Acknowledgements

We would like acknowledge the funding support from Center for Identification Technology Research (CITeR), from TABULA RASA project, 7th Framework Research Programme of the European Union (EU), grant agreement number: 257289, and from PRIN 2008 project ``Biometric Guards - Electronic guards for protection and security of biometric systems" funded by the Italian Ministry of University and Scientific Research (MIUR).

References

[1] D. Willis, M. Lee, *Six Biometric Devices Point The Finger At Security.* Biometrics Under Our Thumb, Network Computing, June 1998

[2] van der Putte, T. and Keuning, J.: Biometrical Fingerprint Recognition: Don't Get Your Fingers Burned, *SMART CARD RESERCH AND ADVANCED APPLICATIONS*, IFIP TC8/WG8.8 Fourth Working Conference on Smart Card Research and Advanced Applications, pp. 289-303 (2001).

[3] T. Matsumoto, H. Matsumoto, K. Yamada, and S. Hoshino, Impact of artificial gummy fingers on fingerprint systems, In Proceedings of SPIE, 4677, Optical Security and Counterfeit Deterence Techniques IV, Yokohama, Japan.

[4] Peter Kallo, Imre Kiss, Andras Podmaniczky, all of Budapest, Janos Talosi, Negykanizsa, All of (HU). "Detector for Recognizing the Living Character of a Finger in a Fingerprint Recognizing Apparatus" Patent US 6,175641, Jan. 16, 2001

[5] Schuckers SAC. *Spoofing and Anti-Spoofing Measures*. Information Security Technical Report, Vol 7. No. 4, pages 56-62, 2002.

[6] Bozhao Tan, Stephanie Schuckers, Spoofing protection for fingerprint scanner by fusing ridge signal and valley noise, Pattern Recognition, Volume 43, Issue 8, August 2010, Pages 2845-2857, ISSN 0031-3203, DOI: 10.1016/j.patcog.2010.01.023.

[7] Martin Drahansky, Dana Hejtmankova, New Experiments with Optical Liveness Testing Methods, Journal of Information Hiding and Multimedia Signal Processing, Volume 1, Number 4, October 2010

[8] P. Coli, G.L. Marcialis, and F. Roli, Fingerprint silicon replicas: static and dynamic features for vitality detection using an optical capture device, International Journal of Image and Graphics, World Scientific, 8 (4) 495-512, 2008.

[9] Gian Luca Marcialis, Aaron Lewicke, Bozhao Tan, Pietro Coli, Fabio Roli, Stephanie Schuckers, Dominic Grimberg, Alberto Congiu, Alessandra Tidu, First International Fingerprint Liveness Detection Competition – LivDet 2009.