LivDet 2013 Fingerprint Liveness Detection Competition 2013

Luca Ghiani, Valerio Mura, Simona Tocco, Gian Luca Marcialis, Fabio Roli
University of Cagliari - Department of Electrical and Electronic Engineering Italy
{luca.ghiani, marcialis, roli}@diee.unica.it, {valeriomura, tocco.simona@}gmail.com

David Yambay, Stephanie Schuckers
Clarkson University - Department of Electrical and Computer Engineering USA
{yambayda, sschucke}@clarkson.edu

Abstract

A spoof or fake is a counterfeit biometric that is used in an attempt to circumvent a biometric sensor. Liveness detection distinguishes between live and fake biometric traits. 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 a biometric measure is authentic.

The Fingerprint Liveness Detection Competition (LivDet) goal is to compare both software-based (Part 1) and hardware-based (Part 2) fingerprint liveness detection methodologies and is open to all academic and industrial institutions. Submissions for the third edition were much more than in the previous editions of LivDet demonstrating a growing interest in the area. We had nine participants (with eleven algorithms) for Part 1 and two submissions for Part 2.

1. Introduction

Among biometric systems, fingerprints systems are probably the best-known and widespread because of the fingerprint properties: universality, durability and individuality. Unfortunately it has been shown that fingerprint scanners are vulnerable to spoof attacks, i.e. it is possible to deceive a fingerprint system with an artificial replica of a fingertip. Therefore, it is important to develop countermeasures to those attacks.

Liveness detection, with either hardware-based or software-based systems, is used to check if a presented fingerprint originates from a live person or an artificial finger. Usually the result of this analysis is a score used to classify images as either live or fake.

To detect liveness, hardware-based systems employ a combination of additional sensors and software and may include measurements outside of the fingerprint image itself while the software-based only use image processing algorithms to gather information directly from the collected fingerprint.

Since 2009, in order to assess the main achievements of the state of the art in fingerprint liveness detection, the Department of Electrical and Electronic Engineering of the University of Cagliari, and the Department of Electrical and Computer Engineering of the Clarkson University, have organized the Fingerprint Liveness Detection Competition.

The First International Fingerprint Liveness Detection Competition LivDet 2009 [1], provided an initial assessment of software systems based on the fingerprint image only. The second and third Liveness Detection Competition (LivDet 2011 [2] and 2013) were created in order to ascertain the current state of the art in liveness detection, including integrated system testing. LivDet 2011 and 2013 were both 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, we describe the LivDet 2013 competition characteristics and we summarize the results achieved from the participants. Section 2 describes some of the fingerprint spoofing techniques and the liveness detection countermeasures. In section 3 the evaluation protocols of the algorithms and the systems are examined in depth. Section 4 presents the competition results and section 5 concludes the paper.

2. Fingerprint liveness detection

There are two methods to create an artificial fingertip, the cooperative method and the non-cooperative method. In the cooperative method the subject pushes the finger into a plasticine-like material creating a negative impression of the fingerprint as a mold.

The mold is then filled with a material, such as gelatin, PlayDoh or silicone that will reproduce the fingerprint char-

acteristics. In the non-cooperative method a latent fingerprint left on a surface is enhanced, digitized through the use of a photograph, and, finally, the negative image is printed 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.

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. In Table 1 are presented the participants names and the correspondent algorithms names as they're used in this paper. Three out of nine preferred to remain anonymous and the University of Naples Federico II submitted three different algorithms.

3.2. Part 1: Algorithm Data Set

The dataset for *Part 1: Algorithms* consists of images from four different devices; Biometrika, Crossmatch, Italdata and Swipe. There are 4000 or more images for each of these devices as detailed in Table 3 and 4. The spoof materials used for this experiment were Body Double, latex, Play-Doh and wood glue for Crossmatch and Swipe and gelatine, latex, ecoflex (platinum-catalysed silicone), modasil and wood glue for Biometrika and Italdata. The images were divided into two equal datasets, training and testing. Details are described in Table 2 and Table 3 and 4. Live images came from 300 fingers from 30 subjects for Biometrika and Italdata datasets, 440 fingers representing 44 subjects for Crossmatch dataset, and 250 fingers from 50 subjects for Swipe dataset. Spoof images come from approximately 100 fingers representing 20 people for the Crossmatch and Swipe Datasets and 100 fingers representing 15 subjects for the Biometrika and Italdata datasets.

The spoof images of two of the LivDet 2013 datasets (Crossmatch and Swipe) were collected using the cooperative method that was described earlier. The other two datasets (Biometrika and Italdata) were created using the non-cooperative method. This is the reason of the great difference between the error rates: a fake created from a latent fingerprint (therefore without cooperation) may, in some cases, be less similar to the original one then one created with cooperation.

After the competition is completed, the entire dataset will be made available to those who sign the proper data release agreement. Figure 1 below shows examples of images used in the experiments.

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

Participants	**Algorithm names**
Dermalog Identification Systems GmbH	Dermalog
First anonymous participant	Anonym1
Universidad Autonoma de Madrid	ATVS
Second anonymous participant	Anonym2
University of Naples Federico II (first algorithm)	UniNap1
University of Naples Federico II (second algorithm)	UniNap2
University of Naples Federico II (third algorithm)	UniNap3
Third anonymous participant	Anonym3
HangZhou JLW Technology Co Ltd	HZ-JLW
Federal University of Pernambuco	Itautec
Chinese Academy of Sciences	CAoS
Safran Morpho	Morpho

Table 2: Device characteristics for Part 1 datasets

Data set	Sensor	Model No.	Resolution(dpi)	Image size
#1	Biometrika	FX2000	569	315x372
#2	Italdata	ET10	500	640x480
#3	Crossmatch	L SCAN GUARDIAN	500	800x750
#4	Swipe		96	208x1500

Table 3: Training and test set characteristics for Part 1 datasets (live samples)

Data set	Sensor	Live training samples	Live testing samples
#1	Biometrika	1000/200	1000/200
#2	Italdata	1000/200	1000/200
#3	Crossmatch	1250/500	1250/440
#4	Swipe	1250/500	1250/500

Table 4: Training and test set characteristics for Part 1 datasets (fake samples)

Data set	Sensor	fake training samples	fake testing samples
#1	Biometrika	1000/100	1000/100
#2	Italdata	1000/100	1000/100
#3	Crossmatch	1000/125	1000/100
#4	Swipe	1000/125	1000/100

3.3. Algorithm Submission

The algorithm submission for LivDet 2013 uses the same structure as LivDet 2009 and 2011. As stated for the other two LivDet editions each submitted algorithm must be a Win32 console application with the following list of parameters:

LIVENESS_XYZ.exe [ndataset] [inputfile] [outputfile]

In Part 2: Systems, participants were asked to ship a fingerprint system which captures a fingerprint image as well as outputs a liveness detection score. Three spoof recipes and methods were made available to the participants upon their registration. These materials were Playdoh, Gelatin and Ecoflex. In addition the systems were tested with two unspecified methods. These materials were Modasil and Latex. The requirements for installation are that the system will be run on either a Windows XP 32-bit or 64-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 if the image is considered a live image and a liveness score normalized in the range of 0 and 100 (100 is the maximum degree of liveness, 0 spoof) for images detected as either live of fake. In the case that the algorithm has not been able to process the image it is considered an error for a live image and a success for a spoof image. Laboratory staff systematically collected live data from participating subjects and attempted to spoof the system with casts made from the varying materials. Each submitted system was given 2000 test attempts. This corresponds to 1000 live attempts from 50 people (2 images each of all 10 fingers) were performed, as well as 1000 spoof attempts for the five different materials listed above. The spoof attempts were conducted with 2 images per finger from the 5 right hand fingers of 20 subjects for each spoof material. The same physical spoof fingers were placed on both scanners. A spoof image was collected on one scanner and then the next scanner alternating which scanner was first for a new spoof finger. In order to ensure quality of the spoof images, 15 spoof fingers were collected on a third party system to examine the quality of that batch

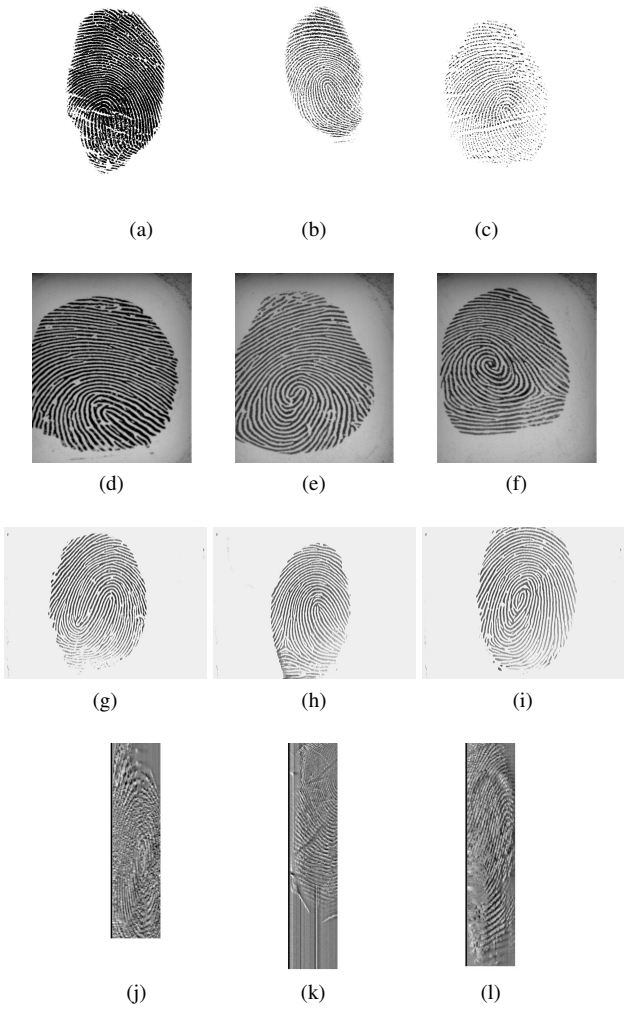


Figure 1: Examples of fake fingerprints acquired with the 4 sensors. From Crossmatch (a) body double, (b) latex, (c) wood glue, from Biometrika (d) gelatine, (e) latex, (f) wood glue, from Italdata (g) gelatine, (h) latex, (i) wood glue, from Swipe (j) body double, (k) latex, (l) wood glue

of each spoof recipe.

3.4. 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.

Overall evaluation:

- *Frej_n*: Rate of failure to enroll.
- *Fcorrlive_n*: Rate of correctly classified live fingerprints.
- *Fcorrfake_n*: Rate of correctly classified fake fingerprints.
- *Ferrlive_n*: Rate of misclassified live fingerprints.
- *Ferrfake_n*: Rate of misclassified fake fingerprints.

4. Results and Discussion

Eleven algorithms and two systems successfully completed the competition at the time of submission of this paper: Dermalog, ATVS, the University of Naples Federico II (three different algorithms), HZ-JLW, Itautec, CAoS and the three Anonymus for Part 1; Dermalog and Morpho for Part 2.

4.1. Part 1: Algorithms

The competition results are presented in the following tables, for the sake of space we only show the rate of misclassified live fingerprints (ferrlive) in Table 5, the rate of misclassified fake fingerprints (ferrfake) in Table 6 and the accuracy rate in Table 7.

The threshold value for determining liveness was set at 50. This threshold is used to calculate Ferrfake and FerrLive.

Since, as stated earlier, the Biometrika and Italdata datasets were created using the non-cooperative method, the fakes were easier to detect therefore the error rates are much lower than those obtained with the Crossmatch and Swipe datasets.

The live images collected with the Crossmatch sensor turned to be especially difficult to recognize for most of the algorithms.

Table 5: Rate of misclassified live fingerprints (ferrlive) for submitted algorithms

	Biometr.	Italdata	Crossm.	Swipe	Average
Dermalog	3.30	0.50	99.84	3.82	26.86
Anonym1	1.50	0.50	86.96	N.A.	N.A.
ATVS	4.60	0.00	90.40	0.00	23.75
Anonim2	2.30	0.20	98.40	2.52	25.85
UniNap1	30	2.10	31.28	11.45	11.96
UniNap2	1.80	5.00	55.20	33.22	23.80
UniNap3	1.80	2.10	55.20	11.45	17.64
Anonum3	3.30	1.00	95.52	2.69	25.63
HZ-JLW	65.30	26.10	100.00	25.33	54.18
Itautec	1.10	1.30	64.96	N.A.	N.A.
CAoS	5.50	21.10	41.92	N.A.	N.A.

The best performances were those of Dermalog with a 98.3% accuracy for the Biometrika dataset, of Anonym2 with a 99.4% for the Italdata dataset, of the first Federico II algorithm with a 68.8% for the Crossmatch dataset and of Dermalog again with a 96.47% for the Swipe dataset. Finally, despite the great Dermalog and Anonym2 performances, the best average results were those of the first Federico II algorithm with a 86.63%of accuracy and with an average of 11.96% FerrLive and 14.62% FerrFake.

The Anonym1, Itautec and CAoS algorithms were not able to process the Swipe images and for this reason they did not compete for the final win despite the good performances of Anonym1 on the Biometrika and Italdata datasets.

The failure to enroll rate was zero on all datasets for all algorithms except the Anonymous1 one which has not been able to analyze one image from both the Biometrika and Italdata datasets and two images from the Crossmatch dataset.

4.2. Part 2: Systems

FerrLive and FerrFake for the two submitted systems can be found in Figure 2 with the data summarized in Table 8. Dermalog performed at a FerrLive of 11.8% and a FerrFake of 0.6%. Morpho performed at a FerrLive of 1.4% and a FerrFake of 0%. Both systems had low FerrFake rates. Morpho received a perfect score of 0% error, successfully determining every spoof finger presented as a spoof.

Figure 3 and 4 above shows the equal error rate curves for the Dermalog and Morpho systems. Figure 5 shows

Table 6: Rate of misclassified fake fingerprints (ferrfake) for submitted algorithms

	Biometr.	Italdata	Crossm.	Swipe	Average
Dermalog	0.10	1.10	0.00	3.20	1.1
Anonym1	2.40	1.70	2.40	N.A.	N.A.
ATVS	5.50	100.00	10.30	100.00	53.95
Anonim2	1.30	1.00	0.30	9.60	3.05
UniNap1	6.40	4.90	31.10	16.10	14.62
UniNap2	11.30	13.90	48.30	19.50	23.25
UniNap3	11.30	4.90	48.30	16.10	20.15
Anonym3	8.10	4.60	0.10	8.20	5.25
HZ-JLW	0.60	0.20	0.00	3.50	1.07
Itautec	16.90	6.50	13.90	N.A.	N.A.
CAoS	3.70	70.70	54.20	N.A.	N.A.

Table 7: Rate of accuracy for submitted algorithms

	Biometr.	Italdata	Crossm.	Swipe	Average
Dermalog	98.30%	99.20%	44.53%	96.47%	84.63%
Anonym1	98.00%	98.85%	50.53%	N.A.	N.A.
ATVS	94.95%	50.00%	45.20%	53.55%	60.93%
Anonym2	98.20%	99.40%	45.20%	94.19%	84.25%
UniNap1	95.30%	96.50%	68.80%	85.93%	86.63%
UniNap2	93.45%	90.55%	47.87%	73.15%	76.26%
UniNap3	93.45%	96.50%	47.87%	85.93%	80.94%
Anonym3	94.30%	97.20%	46.89%	94.75%	83.29%
HZ-JLW	67.05%	86.85%	44.44%	84.81%	70.79%
Itautec	91.00%	96.10%	57.73%	N.A.	N.A.
CAoS	95.40%	54.10%	52.62%	N.A.	N.A.

Table 8: FerrLive and FerrFake for submitted systems

Submitted Systems	FerrFake	FerrLive	FerrFake Known	FerrFake Unknown
Dermalog	0.6%	11.8%	0.3%	1%
Morpho	0%	1.4%	0%	0%

Figure 2: FerrLive and FerrFake for submitted systems for Dermalog, Morpho.

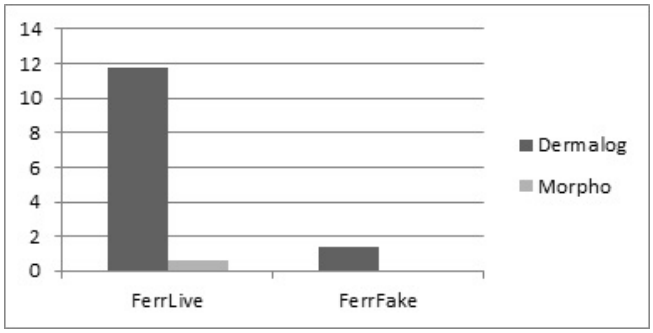
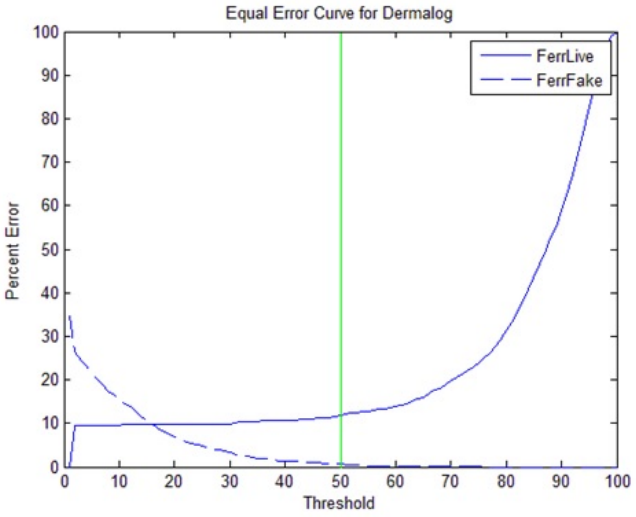


Figure 3: Equal Error Rate Curves for Dermalog.



the FerrFake rate for the Dermalog system for the unknown recipes and known recipes. Since Morpho had a 0% overall FerrFake, it did not require a plot. Dermalog performed considerably better against the known recipes than the unknown recipes. The results from Dermalog continue to show the importance of learning different methods for creating spoof fingers as the knowledge of recipes helps to significantly lower the FerrFake rate. Figure 4 shows examples of accepted and rejected live images on Dermalog and Figure 5 shows examples of accepted and rejected spoof images on Dermalog. Morpho did not record images in a format able to be viewed and thus images are not available.

Both systems had software built-in to auto-detect fingers and collect when a finger was found. All live fingers were able to be detected by both systems, however not all spoof fingers were detected. Dermalog had a 23% Fake Non-Response Rate. The Morpho device had a 49% Fake Non-

Figure 4: Equal Error Rate Curves for Morpho.

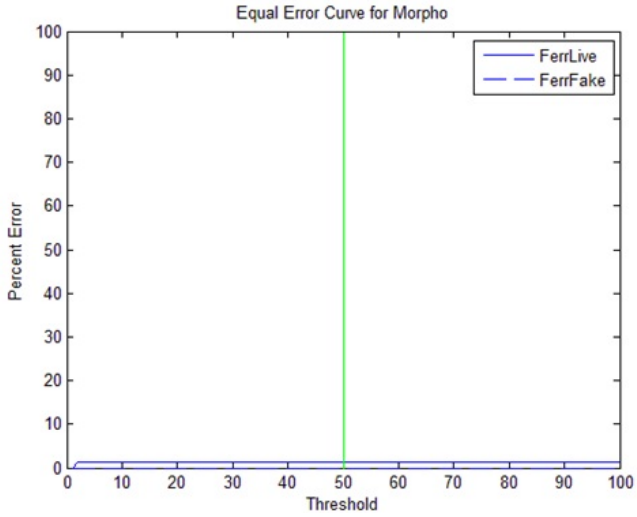
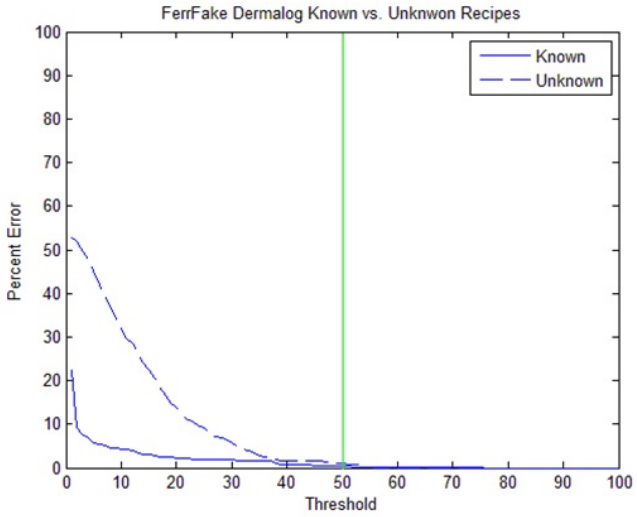


Figure 5: FerrFake for unknown vs. known recipes for Dermalog.



Response Rate. It should be noted that a non-response is recorded as a successful detection of a spoof attempt.

Morphos high Fake Non-Response rate can partially be attributed to the lack of moisture in some of the recipes. The system was unable to detect any Playdoh or Ecoflex images.

In addition to the testing reported here, a small subset of 50 images per material was conducted with the inclusion of a brushed on saline solution. The Morpho system was able to detect the previously undetected spoof fingers. However, even though the spoof was detected as a finger, none of the spoofs were detected as a live finger, i.e., there was no increase in FerrFake using the saline solution.

5. Conclusions

LivDet 2013, with the third international public competition for software-based fingerprint liveness detection and the second public assessment of system-based fingerprint liveness detection, proved to be an important competition. The number of participants, from both academic and industrial institutions, is growing with respect to previous editions. As a matter of fact entries were submitted from a total of ten participants demonstrating the state-of-the art in fingerprint liveness detection. Since an effective liveness detection algorithms is a key component to minimize the vulnerability of fingerprint systems to spoof attacks, we hope that this competition success continues to increase such that further improvement in algorithm performance is encouraged.

6. Acknowledgements

This work was partly supported by the Tabula Rasa project, 7th FP of the European Union (EU), grant agreement number: 257289 and the Center for Identification Technology Research (CITeR).

References

[1] G.L. Marcialis, et al., First International Fingerprint Liveness Detection Competition, 14th Int. Conf. on Image Analysis and Processing (ICIAP 2009), Springer LNCS 5716, pp. 12-23.

[2] D. Yambay, et al., LivDet 2011 - Fingerprint Liveness Detection Competition 2011, 5th IAPR/IEEE Int. Conf. on Biometrics (ICB 2012), New Delhi (India), March, 29th, April, 1st, 2012.

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

[4] P. Kallo, et al., 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] P. Coli, et al., 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.