
LIVDET2023 - FINGERPRINT LIVENESS DETECTION COMPETITION: ADVANCING GENERALIZATION

**Marco Micheletto¹, Roberto Casula¹, Giulia Orrù¹, Simone Carta¹, Sara Concas¹,
Simone Maurizio La Cava¹, Julian Fierrez², Gian Luca Marcialis¹**

¹ University of Cagliari, DIEE, Cagliari, Italy

² Universidad Autonoma de Madrid, BiDA Lab, Madrid, Spain

{marco.micheletto, roberto.casula, giulia.orrù, marcialis}@unica.it, julian.fierrez@uam.es

ABSTRACT

The International Fingerprint Liveness Detection Competition (LivDet) is a biennial event that invites academic and industry participants to prove their advancements in Fingerprint Presentation Attack Detection (PAD). This edition, LivDet2023, proposed two challenges, “Liveness Detection in Action” and “Fingerprint Representation”, to evaluate the efficacy of PAD embedded in verification systems and the effectiveness and compactness of feature sets. A third, “hidden” challenge is the inclusion of two subsets in the training set whose sensor information is unknown, testing participants’ ability to generalize their models. Only *bona fide* fingerprint samples were provided to participants, and the competition reports and assesses the performance of their algorithms suffering from this limitation in data availability.

1 Introduction

Due to their convenience and security, fingerprint-based authentication systems have garnered significant attention in numerous applications, ranging from financial transactions to healthcare management [14]. However, these systems are vulnerable to various attacks, including spoofing or presentation attacks [13], where attackers use artificial replicas of live fingers to deceive the sensors [10]. Such attacks can lead to severe consequences, such as unauthorized access, identity theft, and financial fraud. To address this vulnerability, automated presentation attack detection (PAD) systems that utilize either hardware or software have been developed over the last few decades [20, 11]. Software-based methods, in particular, have seen significant advancements [12] thanks to pattern recognition research innovations and larger datasets’ availability [3]. Among other initiatives, the International Fingerprint Liveness Detection Competition (LivDet)* [18], now in its eighth edition, has significantly promoted research and development in this area since its inception in 2009 and has become a well-known benchmark for assessing the effectiveness of PAD techniques.

LivDet2023 presents two familiar challenges from the previous edition, “Liveness Detection in Action” and “Fingerprint Representation” [18], while introducing new evaluation criteria and datasets. Challenge 2 now evaluates system speed to encourage efficient and practical PAD systems that can operate in time-critical real-world scenarios. The retention of previous challenges ensures continuity and comparability with past results.

Furthermore, LivDet2023 has set out to tackle a new challenge often overlooked in previous competition editions: the issue of generalization. Generalization [6] refers to the ability of systems to detect the authenticity of fingerprints in a wide variety of sensor technologies, Presentation Attack Instruments (PAIs) and attack scenarios rather than being limited to specific, predefined conditions. Despite the improvements in the field, developing generalized PADs remains challenging for several reasons. Firstly, creating different PAIs for training is a non-trivial, expensive, and time-consuming task requiring skilled operators. Secondly, even with a comprehensive training dataset, existing

*<https://sites.unica.it/livdet/>

methods often struggle to detect PAs captured using new acquisition methods or materials [15, 17]. Previous editions of LivDet have provided strong evidence of this phenomenon [18].

Recent works have focused on addressing the lack of interoperability across fabrication materials by using one-class classification [8, 9]. Unlike the traditional multi-class paradigm, one-class classification utilizes data from a single class, typically the bona fide class, for training the classifier. The ultimate objective of this approach is to establish a decision boundary around bona fide class samples that can accommodate as many samples as possible from that class while rejecting samples from other classes. Driven by the potential of this approach, we have added a “hidden” challenge next to the previous ones. In particular, we included two subsets in the LivDet 2023 training set whose sensor information is unknown and contains only bona fide fingerprint samples. This challenge presents a valuable opportunity for participants to assess the effectiveness of their algorithms on unknown data, an essential aspect of real-world deployment. We hope this will increase LivDet2023’s rigour and relevance while promoting the advancement of research and development in PADs.

2 LivDet2023

As in the previous edition, two distinct challenges characterize the LivDet2023 competition:

- Challenge 1, *Liveness Detection in Action* [5]: Competitors were asked to submit a complete algorithm capable of producing both the “score”, that is, the probability of being a bona fide sample and the “integrated score”, which combines the previous score with the probability of belonging to the claimed user. Participants in this challenge can choose whether to use the related “user-specific” information [19].
- Challenge 2, *Fingerprint representation*: Compactness and discriminability of feature vectors are critical in modern authentication systems to ensure high performance in terms of accuracy and speed. With the aim of evaluating speed and compactness, we asked competitors to submit PADs that return the feature vector corresponding to the input image in addition to the score.

Furthermore, to evaluate the participating PADs’ ability to generalize, an additional challenge, called *Unknown sensors*, is introduced: in the training set, two sensors were unknown, the name and brand of the sensor were not declared, and only bona fide fingerprint samples were provided.

2.1 Datasets and participants

Although the number of competitors was lower than in the previous edition, the competition showcased a diverse range of algorithms from each participant. It is worth noting that many competitors submitted multiple algorithms, which highlights their dedication to finding innovative solutions. Table 3 provides further details on each competitor, including the name of their presented algorithm, the type of solution adopted, and the challenge(s) in which they participated. In addition, we considered the quantity of data utilized by each participant in the training phase. In fact, some competitors have generated for each test set a model trained on data from the specific sensor (single, in Table 3); others have generated a single model trained with data from multiple sensors suitable for multiple test sets (multiple, in Table 3). Moreover, although we strongly advised utilizing only the LivDet 2023 dataset to maintain consistency in the results, some participants have employed additional data, which could have given them an advantage. Conversely, some competitors opted to use fewer data, typically omitting unknown sensors during the training phase. These instances will be designated by a plus (+) or minus (-) sign, respectively.

The LivDet 2023 training set and test set comprise four sub-sets containing fingerprint images from four different capture devices: GreenBit DactyScan 84C, Dermalog LF10, Jenetric LiveTouch Quattro and Integrated Biometrics Watson Mini.

The sensors can be grouped into two categories: known and unknown sensors. GreenBit and Dermalog are known sensors; therefore, we provided competitors with these devices’ names, brands, and technical details. The training set for these sensors included 25 users, for a total of 2750 images, subdivided into 1250 bona fide and 1500 PAs collected with the classic consensual method. On the other hand, Jenetric and Integrated Biometrics were unknown sensors: we did not declare any information about these devices to the competitors. The training set included only bona fide fingerprint samples, totalling 1250 samples. To ensure the accuracy and dependability of the algorithms, our test sets were carefully designed to facilitate cross-material and cross-method experiments. In order to introduce more significant variability, we included synthetic fingerprints fabricated using materials different from those used in the training set. Furthermore, we incorporated presentation attacks generated with our semi-consensual ScreenSpooft technique [4], which is known for its ability to produce highly realistic forgeries. The test set is four times larger than the training set, comprising 2500 bona fide and 6000 attack presentations (including both consensual and ScreenSpooft-generated PAIs)

Table 1: Device characteristics for LivDet 2023 datasets.

Scanner	Model	Res. [dpi]	Img Size	Format	Type
Green Bit	DactyScan84C	500	500x500	JPEG	Optical
Dermalog	LF10	500	500x500	JPEG	Optical
Jenetric	Livetouch Quattro	500	500x500	JPEG	Optical
Integrated Biometrics	Watson Mini	500	500x500	JPEG	Hybrid

Table 2: Number of samples for each scanner and each part of the dataset.

Dataset	Training			Test Consensual/ScreenSpoof					
	Bona fide	Latex_v2	RPro10	Bona fide	BodyDouble_new	ElmersGlue	R15	GLS	Mix3
Green Bit	1250	750	750	2500	1000/1000	1000/1000	1000/1000	-	-
Dermalog	1250	750	750	2500	1000/1000	-	1000/1000	1000/1000	-
Jenetric	1250	-	-	2500	1500	1500	-	-	-
Int. Biometrics	1250	-	-	2500	-	1500	-	-	1500

for known sensors. However, we deliberately included only 3000 ScreenSpoof-generated PAs for unknown sensors to create a more challenging scenario for the algorithms to detect and classify presentation attacks.

2.2 Algorithms submission

Algorithms for Challenge 1 must be submitted as console programs with the following parameters:

[nameOfAlgorithm] [ndataset] [templateimagesfile] [probeimagesfile] [livenessoutputfile] [IMSoutputfile].

The parameter *ndataset* is an identification number for the dataset used. The file *templateimagesfile* contains a list of absolute paths to every template image stored in the system, while the file *probeimagesfile* contains a list of absolute paths to each probe image that the algorithm will test. The algorithm outputs are saved to the paths specified by the last two parameters. The file *livenessoutputfile* contains the degree of “liveness” for each processed image, normalized between 0 and 100, where 100 indicates the highest degree of liveness, and 0 denotes a fake image. Fingerprint images with scores [0,50) are classified as “presentation attack”, while those with scores [50,100] are classified as “bona fide”. The file *IMSoutputfile* lists, for each probe image, the normalized probability of a fingerprint belonging to the declared identity and being authentic. Scores [0,50) classify the probe as “presentation attack” or the probe-template comparison as no-mated comparison, while scores [50,100] classify the comparison as bona-fide and mated. The evaluation threshold is set to 50. If the algorithm is unable to process an image, the corresponding value in both outputs is set to -1000.

The submission process for Challenge 2 in LivDet 2023 is the same as in LivDet 2021. In addition to the parameters *nameOfAlgorithm*, *ndataset*, *probeimagesfile*, and *livenessoutputfile*, Challenge 2 applications require an additional parameter called *embeddingsfile*, representing the file of feature vectors for each processed image.

2.3 Performance Evaluation

In both challenges, the performance of the PADs will be evaluated using the standard PAD ISO metrics [2, 1]:

- PAD Accuracy: percentages of fingerprint images correctly classified by the PAD.
- BPCER (Bona fide Presentation Classification Error Rate): Rate of misclassified bona fide images.
- APCER (Attack Presentation Classification Error Rate): Rate of misclassified fake images.

In Challenge 1, to evaluate the performance of the integrated system, we employed the following metrics:

- FNMR (False Non-Match Rate): Rate of mated comparisons that result in rejection.
- FMR (False Match Rate): Rate of non-mated comparisons that result in acceptance.
- IAPAR (Impostor Attack Presentation Accept Rate): rate of presentation attacks that result in acceptance.
- Integrated Matching (IM) Accuracy: percentages of samples correctly classified by the integrated system.

To simulate real-world scenarios, we conducted comparisons using templates derived from bona fide fingerprints. The testing involved matching a fingerprint template to a fingerprint image from the same finger and user (mated), a fake fingerprint image from the same finger and user (presentation attack), or a bona fide fingerprint image from a different user (no-mated). Overall, we performed 5000 mated, 10000 no-mated and 10000 presentation attack comparisons.

Table 3: Participants name and submitted algorithms, alongside information about their training approach. The terms 'Single' and 'Multiple' denote whether the model training used data from a specific or multiple sensors respectively. The symbols (+) and (-) indicate if participants used additional or fewer data for training their respective models.

Participant	Algorithm name	Acronym	Challenge	Type	Training data
UNESP	Contreras_1_ch11/2[7]	contr1	1,2	Hand-crafted	Single
	Contreras_2_ch11/2[7]	contr2	1,2		Single
Peking University	CIS_PAD_F	CIS_F	1	Deep-learning	Multiple(-)
	CIS_PAD_W	CIS_W	1		Multiple(-)
	CIS_PAD_W_ensemble	CIS_Wens	1		Multiple(-)
	CIS_PAD_F_v2	CIS_F2	1		Multiple(-)
	PAD_Supcon_cls	S_cls	1		Multiple(-)
	PAD_Supcon_knn	S_knn	1		Multiple
Hanbat National University	HNU_AIM	HNU	1		Multiple
Università degli Studi di Napoli Federico II	mod1	unina1	1		Single/Multiple(-)
	mod2	unina2	1,2		Single/Multiple(-)
	mod3	unina3	1,2		Single/Multiple(-)
	unina_grbt	unina4	1		Multiple
	unina_derm	unina5	1		Multiple
	unina_grbt_derm	unina6	1		Multiple
JIIOV Technology	run	jiiov	1,2		Multiple
	run_all_data	jiiov_all	1,2	Multiple(+)	

Challenge 2 aimed to evaluate the compactness and the discriminability of feature vectors generated by various algorithms. We considered both the speed and size of the feature vectors to be essential parameters for this edition. To ensure fairness in the evaluation, we specified two machines where the algorithms were tested: a Desktop-PC Linux 18.04.1 Ubuntu or Windows 10 Pro system with an Intel® Core™ i9 9900K @ 3.60GHz processor, 64 GB DDR4 2.933 MHz RAM, and dual NVIDIA® GeForce® RTX 2080 Ti (11GB each) graphics cards. The final ranking was determined based on the speed of the algorithms in generating and comparing the feature vectors, their size, and the accuracy achieved on the specific dataset. The final score was obtained by combining the contributions of speed, compactness and PAD accuracy, normalized and averaged.

3 Results

This section examines the results of the algorithms submitted to LivDet2023. The global results of the two challenges are shown in Tables 4 and 5.

Seventeen algorithms were submitted to Challenge 1, which evaluates the integration between PAD and comparator; the results are shown in Table 6 for the known, that is GreenBit and Dermalog, consensual test datasets, in Table 7 for the known ScreenSpooF test datasets and in Table 8 for the two unknown sensors, Jenetric and Int. Biometrics.

Analyzing the results as a whole, it is evident that the test sets acquired with unknown sensors reported higher average errors. In particular, while on average, competitors achieved about 88% IMS on data acquired on known sensors, this value dropped to 84% for data from unknown sensors. Although, on average, this drop does not seem particularly significant, if we analyze the APCERs of the single methods, we can deduce that the rate of erroneously classified unknown presentation attacks is very high (14.91% for known sensors vs. 39.58% for unknown sensors). This aspect shows that the interoperability problem in fingerprint presentation attack detection is still open.

In this challenge, the CIS_W/Wens model emerges as the undisputed winner in terms of PAD/IM accuracy. This model, which is sensor-interoperable, has been designed to combine metric learning with the spoof detection task, aiming to encode more PAD-related information while minimizing sensor-related interference. Interestingly, despite the model's training being conducted with a smaller volume of data than what was fully available - specifically on the Dermalog and Greenbit sub-datasets - it does not appear that this necessarily led to superior performance. This suggests that the relationship between the volume of training data and the model's performance may not be directly proportional [17].

Nevertheless, a notable observation from the data is the high FNMR, much more significant than typical verification systems without PAD [14]. This distinctive characteristic applies across all participating algorithms and confirms what has been reported in [16], namely, the integration of a PAD algorithm has a substantial impact on the performance of the recognition system.

Table 4: Challenge 1: Integrated and PAD overall results. The jiiiov_all method is not considered in the final ranking as it uses additional data.

Algorithm	Overall PAD Accuracy [%]	Overall IM Accuracy [%]
Contr1	92,47	53,69
Contr2	87,42	42,09
CIS_F	88,75	91,55
CIS_W	96,22	95,99
CIS_Wens	97,54	96,35
CIS_F_v2	89,22	91,2
S_cls	93,76	94,11
S_knn	94,05	94,11
HNU_AIM	75,8	86,29
unina1	86,24	93,14
unina2	87,42	92,86
unina3	88,66	93,47
unina4	87,23	84,87
unina5	86,65	93,07
unina6	88,3	93,06
jiiiov	90,12	94,68
jiiiov_all	87,11	94,33

Table 5: Challenge 2: overall results.

Algorithm	Overall Time/im [ms]	Feat. vect. size	Accuracy [%]	Score
Contr1	1302.97	800	87.65	0.57
Contr2	4511.78	800	79.03	0.00
unina2	93.80	32	79.80	0.69
unina3	94.10	32	80.70	0.73
jiiiov	46.89	192	84.29	0.80
jiiiov_all	47.42	192	80.55	0.66

Compared to the previous year, there has been a shift in trend concerning the detection of consensual and ScreenSpooof attacks for the Greenbit dataset. In fact, the IAPAR is higher for this sensor in consensual scenarios. The reasons behind this phenomenon will be the subject of future research.

It is important to highlight that the handcrafted algorithms, Contr1 and Contr2, report a low IM accuracy due to the very high FMR. However, this behaviour is strictly linked to the choice of the comparator since PAD performances are in line with the other detectors.

As we shift focus to deep learning methods, the underperformer is hnu_aim. Despite its singular distinction as the quickest method in this edition (20 ms for probe/template comparison), it fails to demonstrate competitive potential in real-world applications due to its low accuracy.

The only algorithm that used additional data jiiiov_all has not demonstrated substantial effectiveness, particularly concerning PAD performance. It exhibited an unacceptably high error margin in the APCER metric, specifically when detecting ScreenSpooof-fabricated PAs.

For Challenge 2, six algorithms were submitted. The goal of this challenge is to encourage the development of algorithms that strike a balance between accuracy, speed, and compactness. These are crucial factors for ensuring high-performance fingerprint recognition.

Tables 9 and 10 present the results for known and unknown sensors. The overall evaluation considering processing time and feature vector size is shown in Table 5. While the handcrafted methods exhibit the highest accuracy, they are also characterized by larger size and longer computational time. For example, the Contr2 method exceeds an average processing time of 4 seconds per image, which is impractical for real-world scenarios.

Considering these aspects, the algorithm that offers the best compromise is jiiiov. By leveraging the learning capabilities of a CNN, this algorithm effectively identifies patterns and distinctive characteristics in fingerprints, enabling fast image processing. In terms of compactness, the top-ranking algorithm is unina, which employs an autoencoder-based approach to achieve a condensed representation of relevant information.

Table 6: Challenge 1 Integrated and PAD Consensual results - Known scanners.

	Algorithms	FNMR [%]	FMR [%]	IAPAR [%]	IM Acc. [%]	BPCER [%]	APCER [%]	PAD Acc. [%]
GreenBit CC	Contr1	1.42	98.5	22.82	51.19	1.36	22.9	90.03
	Contr2	0.94	97.6	43.54	43.36	0.41	43.76	82.25
	CIS_F	15.32	0.1	8.98	93.30	0.16	11.45	95.33
	CIS_W	15.24	0.12	5.53	94.69	0.05	7.22	97.08
	CIS_Wens	15.24	0.12	5.53	94.69	0.05	7.22	97.08
	CIS_F_v2	15.26	0.10	8.36	93.56	0.10	10.82	95.61
	S_cls	15.22	0.10	14.78	91.00	0.06	19.33	92.24
	S_knn	15.22	0.10	14.77	91.00	0.06	19.33	92.24
	HNU_AIM	83.94	16.01	15.31	70.68	38.81	60.72	52.62
	unina1	19.02	1.55	12.69	90.50	0.00	32.86	86.86
	unina2	23.32	1.51	10.38	90.58	5.35	26.51	86.19
	unina3	21.46	1.54	11.37	90.54	2.61	30.67	86.16
	unina4	20.92	3.54	13.23	89.11	2.01	32.75	85.70
	unina5	19.02	1.55	12.90	90.42	0.00	33.07	86.77
	unina6	19.02	1.55	12.90	90.42	0.00	33.07	86.77
	jiiov	18.06	0.09	8.75	92.85	2.75	28.56	86.93
jiiov_all	16.38	0.08	5.55	94.47	1.08	16.83	92.62	
Dermatlog CC	Contr1	3.00	97.44	3.57	59.00	1.58	3.57	97.62
	Contr2	1.86	97.38	6.04	58.26	1.83	6.04	96.48
	CIS_F	14.86	0.01	2.52	96.02	0.14	3.65	98.46
	CIS_W	14.90	0.03	0.02	97.00	0.20	0.06	99.86
	CIS_Wens	14.90	0.03	0.02	97.00	0.20	0.06	99.86
	CIS_F_v2	14.82	0.01	1.56	96.41	0.06	0.29	99.08
	S_cls	14.80	0.00	0.18	96.97	0.06	0.29	99.85
	S_knn	14.80	0.00	0.18	96.97	0.06	0.29	99.85
	HNU_AIM	37.58	7.87	0.27	89.23	0.39	3.09	98.53
	unina1	18.78	2.71	2.39	94.20	0.03	5.06	97.96
	unina2	21.08	2.59	5.49	92.55	3.34	9.99	94.00
	unina3	19.34	2.69	1.08	94.62	0.70	2.10	98.74
	unina4	20.06	4.74	3.52	92.68	1.92	5.20	96.77
	unina5	18.78	2.71	2.53	94.15	0.03	5.20	97.90
	unina6	18.78	2.71	2.53	94.15	0.03	5.20	97.90
	jiiov	18.32	0.12	0.34	96.15	3.19	1.59	97.45
jiiov_all	16.10	0.17	0.18	96.64	0.95	1.31	98.90	

However, it is important to note that the algorithms generally demonstrate a limited ability to handle the hidden challenge effectively. Despite reporting an acceptable BPCER, an average APCER close to 50% is observed, implying that the PAs classification is akin to a coin toss. This result emphasizes the critical need to develop more sophisticated algorithms that can accurately identify and differentiate bona fide samples from presentation attacks, even without PA examples during training.

4 Discussion and conclusions

The eighth edition of the Fingerprint Liveness Detection Competition allowed for the evaluation of the degree of interoperability of current PADs, in addition to the impact of integrating a PAD system with an AFIS and the level of compactness, speed, and representativeness. To simulate a worst-case scenario for an AFIS designer, we only provided competitors with bona-fide samples for two of the sensors used. The competitors have faced the challenge in the most different ways: someone has trained with only the data of the other sensors to carry out a sort of transfer domain; others have used data from multiple sensors to increase the PAD’s ability to generalize; others have used information from the training data to generate unknown PAs synthetically. No one reported using a one-class classifier. Although the proposed solutions were very different, the APCER on the unknown sensors is still high, especially on one of the two sensors. This shows that the interoperability problem is still open but that solutions to solve it are under development and have potential. A comparison with past LivDet editions reveals a pause in the rise of accuracy typical of earlier editions, with some fluctuations due to the diverse challenges and materials involved.

Acknowledgements

We would like to thank Dirk Morgeneier and his company Jenetric for sponsoring the LivDet 2023 competition. J.F. is supported by project BBforTAI (PID2021-127641OB-I00MICINN/FEDER).

Table 7: Challenge 1 Integrated and PAD ScreenSpooF results - Known scanners.

	Algorithms	FNMR [%]	FMR [%]	IAPAR [%]	IM Acc. [%]	BPCER [%]	APCER [%]	PAD Acc. [%]
GreenBit SS	Contr1	1.42	98.79	1.52	59.59	1.17	1.52	98.69
	Contr2	0.94	97.45	3.55	59.41	0.45	6.39	97.18
	CIS_F	15.32	0.03	6.91	94.16	0.12	14.45	94.15
	CIS_W	15.24	0.07	0.18	96.85	0.03	0.48	99.79
	CIS_Wens	15.24	0.07	0.18	96.85	0.03	0.48	99.79
	CIS_F_v2	15.26	0.03	9.10	94.90	0.09	15.65	93.68
	S_cls	15.22	0.03	13.59	91.51	0.03	28.69	88.50
	S_knn	15.22	0.03	13.59	91.51	0.03	28.69	88.50
	HNU_AIM	28.40	14.20	6.07	82.21	0.23	65.85	73.52
	unina1	19.02	1.73	4.62	93.66	0.00	40.40	83.84
	unina2	23.32	1.69	0.04	94.64	5.23	0.03	96.85
	unina3	21.46	1.71	0.32	94.90	2.59	7.39	95.49
	unina4	26.04	9.33	39.11	75.42	7.69	28.17	84.12
	unina5	19.02	1.73	4.93	93.53	0.00	40.71	83.72
	unina6	19.02	1.73	4.93	93.53	0.00	40.71	83.72
	jiiov	18.06	0.11	1.02	95.94	2.68	4.81	96.47
jiiov_all	16.38	0.11	7.42	96.24	0.99	38.39	84.05	
Dermalog SS	Contr1	3.00	97.32	1.66	59.81	1.61	1.66	98.37
	Contr2	1.86	97.11	3.12	39.91	2.09	3.40	97.39
	CIS_F	14.86	0.07	3.03	95.79	0.18	6.39	97.34
	CIS_W	14.9	0.05	0.42	96.83	0.23	0.64	99.61
	CIS_Wens	14.9	0.05	0.42	96.83	0.23	0.64	99.61
	CIS_F_v2	14.82	0.07	16.49	90.41	0.16	25.63	89.65
	S_cls	14.80	0.07	6.35	94.47	0.08	12.19	95.08
	S_knn	14.80	0.07	6.36	94.47	0.08	12.19	95.08
	HNU_AIM	37.58	7.68	2.03	88.60	0.46	27.30	88.84
	unina1	18.78	2.91	1.38	94.53	0.06	8.09	96.73
	unina2	21.08	2.85	0.05	94.62	3.35	0.62	97.74
	unina3	19.34	2.90	0.18	94.90	0.69	1.01	99.18
	unina4	18.98	3.16	4.83	93.01	0.31	8.18	96.54
	unina5	18.78	2.94	1.64	94.41	0.06	8.35	96.62
	unina6	18.78	2.94	1.64	94.41	0.06	8.35	96.62
	jiiov	18.32	0.06	1.18	95.84	3.38	3.48	96.58
jiiov_all	16.10	0.13	8.02	93.52	1.04	31.34	86.84	

References

- [1] ISO/IEC 2382-37:2022 Information technology — Vocabulary — Part 37: Biometrics, 2022.
- [2] ISO/IEC 30107-3:2023 Information technology — Biometric presentation attack detection — Part 3: Testing and reporting, 2023.
- [3] F. Alonso-Fernandez and J. Fierrez. *Fingerprint Databases and Evaluation*, pages 599–606. Springer, 2015.
- [4] R. Casula, M. Micheletto, G. Orrú, G. L. Marcialis, and F. Roli. Towards realistic fingerprint presentation attacks: The screenspooF method. *Pattern Recognition Letters*, 2022.
- [5] I. Chingovska, A. Anjos, and S. Marcel. Anti-spoofing in action: Joint operation with a verification system. In *2013 IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pages 98–104, 2013.
- [6] T. Chugh and A. K. Jain. Fingerprint presentation attack detection: Generalization and efficiency. In *2019 International Conference on Biometrics (ICB)*, pages 1–8. IEEE, 2019.
- [7] R. C. Contreras, L. G. Nonato, M. Boaventura, I. A. G. Boaventura, F. L. D. Santos, R. B. Zanin, and M. S. Viana. A new multi-filter framework for texture image representation improvement using set of pattern descriptors to fingerprint liveness detection. *IEEE Access*, 10:117681–117706, 2022.
- [8] Y. Ding and A. Ross. An ensemble of one-class svms for fingerprint spooF detection across different fabrication materials. In *2016 IEEE International Workshop on Information Forensics and Security (WIFS)*, pages 1–6. IEEE, 2016.
- [9] J. J. Engelsma and A. K. Jain. Generalizing fingerprint spooF detector: Learning a one-class classifier. In *2019 International Conference on Biometrics (ICB)*, pages 1–8. IEEE, 2019.
- [10] J. Galbally, R. Cappelli, et al. An evaluation of direct attacks using fake fingers generated from iso templates. *Pattern Recognition Letters*, 31(8):725–732, June 2010.
- [11] J. Galbally, J. Fierrez, R. Cappelli, and G. L. Marcialis. *Introduction to Presentation Attack Detection in Fingerprint Biometrics*. Springer, 2023. 3rd Ed.
- [12] J. Galbally, S. Marcel, and J. Fierrez. Image quality assessment for fake biometric detection: Application to iris, fingerprint and face recognition. *IEEE Trans. on Image Processing*, 23(2):710–724, February 2014.
- [13] A. Hadid, N. Evans, S. Marcel, and J. Fierrez. Biometrics systems under spoofing attack: an evaluation methodology and lessons learned. *IEEE Signal Processing Magazine*, 32(5):20–30, September 2015.
- [14] D. Maltoni, D. Maio, A. K. Jain, and J. Feng. Fingerprint matching. In *Handbook of Fingerprint Recognition*, pages 217–297. Springer, 2022.

Table 8: Challenge 1 Integrated and PAD results - Unknown scanners.

	Algorithms	FNMR [%]	FMR [%]	IAPAR [%]	IM Acc. [%]	BPCER [%]	APCER [%]	PAD Acc. [%]
Jenetric SS	Contr1	1.38	98.45	21.99	51.55	0.79	22.17	90.66
	Contr2	0.00	99.99	50.55	39.78	0.00	51.21	79.52
	CIS_F	21.76	0.04	7.06	92.81	4.75	16.71	90.47
	CIS_W	26.48	0.05	0.00	94.68	10.41	0.00	93.75
	CIS_Wens	20.76	0.05	0.75	95.53	3.39	1.56	97.34
	CIS_F_v2	19.92	0.04	18.34	88.66	2.37	32.46	85.60
	S_cls	20.46	0.04	6.38	93.34	3.28	13.80	92.51
	S_knn	18.86	0.04	10.44	92.04	1.55	16.71	92.38
	HNU_AIM	18.66	9.54	1.60	91.81	0.11	59.89	75.98
	unina1	24.24	2.14	1.84	93.56	5.71	11.37	92.03
	unina2	23.58	2.03	5.31	92.35	4.53	18.82	89.75
	unina3	21.36	2.18	3.47	93.47	2.00	16.11	92.36
	unina4	30.50	12.66	40.40	72.68	11.25	34.39	79.49
	unina5	20.40	2.33	3.56	93.56	1.01	19.18	91.72
	unina6	20.04	2.24	6.40	92.54	0.04	35.59	85.74
jiiov	19.92	0.09	12.06	91.16	0.32	52.09	78.97	
jiiov_all	20.92	0.11	6.90	93.01	2.17	30.47	86.51	
Int. Biometrics SS	Contr1	3.04	97.09	48.87	41.01	1.45	49.16	79.46
	Contr2	0.00	99.75	70.71	11.82	0.00	70.77	71.69
	CIS_F	16.72	0.01	48.62	77.20	17.03	82.64	56.73
	CIS_W	19.38	0.03	0.52	95.90	20.31	1.54	87.20
	CIS_Wens	10.24	0.03	1.87	97.19	11.13	4.38	91.57
	CIS_F_v2	5.26	0.02	35.17	84.87	5.00	63.27	71.69
	S_cls	3.84	0.02	4.68	97.35	3.70	8.50	94.38
	S_knn	5.70	0.02	0.45	98.67	5.68	0.85	96.25
	HNU_AIM	3.98	16.13	3.86	91.21	0.00	86.23	65.51
	unina1	3.90	3.32	13.68	92.42	0.00	100.00	60.00
	unina2	3.94	3.32	13.68	92.41	0.05	99.96	59.99
	unina3	3.94	3.32	13.68	92.41	0.00	99.96	60.02
	unina4	11.04	8.29	20.42	86.31	8.70	35.13	80.73
	unina5	4.34	3.25	13.72	92.34	0.72	90.95	63.19
	unina6	6.92	3.41	9.84	93.32	4.21	46.09	79.04
jiiov	1.94	0.15	8.46	96.17	1.08	37.55	84.33	
jiiov_all	1.30	0.15	12.59	94.64	1.11	36.06	73.76	

Table 9: Challenge 2 PAD accuracy of the algorithms on the test sets. For each known dataset the rate of misclassified bona fide and fake fingerprints are reported. The last column is relative to the average of the total accuracy on the four known datasets.

Algorithm	Green Bit					Dermalog					Overall PAD Acc. [%]
	BPCER [%]	CC		SS		BPCER [%]	CC		SS		
		APCER [%]	PAD Acc. [%]	APCER [%]	PAD Acc. [%]		APCER [%]	PAD Acc. [%]	APCER [%]	PAD Acc. [%]	
Contr1	1.20	23.13	86.84	1.57	98.60	1.64	3.43	97.38	1.70	98.33	95.29
Contr2	0.44	43.83	75.89	6.13	96.45	2.12	5.87	95.84	3.53	97.11	91.32
unina2	4.96	26.57	83.25	0.03	97.73	3.44	9.97	93.00	3.44	98.07	93.01
unina3	2.32	31.23	81.91	7.60	94.80	0.68	2.17	98.51	1.07	99.11	93.58
jiiov	2.68	28.50	83.24	4.63	96.25	3.60	1.63	97.47	3.47	96.47	93.36
jiiov_all	0.96	17.03	90.27	38.47	78.58	1.12	1.30	98.78	31.00	82.58	87.55

- [15] E. Marasco and C. Sansone. On the robustness of fingerprint liveness detection algorithms against new materials used for spoofing. In *Multivariable Processing for Biometric Systems*, volume 2, pages 553–558. SCITEPRESS, 2011.
- [16] M. Micheletto, G. L. Marcialis, G. Orrù, and F. Roli. Fingerprint recognition with embedded presentation attacks detection: are we ready? *IEEE Transactions on Information Forensics and Security*, 16:5338–5351, 2021.
- [17] M. Micheletto, G. Orrù, R. Casula, and G. L. Marcialis. Mitigating sensor and acquisition method-dependence of fingerprint presentation attack detection systems by exploiting data from multiple devices. *Applied Sciences*, 12(19):9941, 2022.
- [18] M. Micheletto, G. Orrù, R. Casula, D. Yambay, G. L. Marcialis, and S. Schuckers. Review of the fingerprint liveness detection (livdet) competition series: from 2009 to 2021. *Handbook of Biometric Anti-Spoofing: Presentation Attack Detection and Vulnerability Assessment*, pages 57–76, 2023.
- [19] G. Orrù, P. Tuveri, L. Ghiani, and G. L. Marcialis. Analysis of “user-specific effect” and impact of operator skills on fingerprint pad systems. In *ICIAP 2019*, pages 48–56, Cham, 2019. Springer.
- [20] C. Sousedik and C. Busch. Presentation attack detection methods for fingerprint recognition systems: a survey. *Iet Biometrics*, 3(4):219–233, 2014.

Table 10: Challenge 2 PAD accuracy of the algorithms on the test sets. For each unknown dataset the rate of misclassified bona fide and fake fingerprints are reported. The last column is relative to the average of the total accuracy on the four unknown datasets.

Algorithm	Jenetric			Integrated Biometrics			Overall PAD
	<i>BPCER [%]</i>	<i>APCER [%]</i>	<i>PAD Acc. [%]</i>	<i>BPCER [%]</i>	<i>APCER [%]</i>	<i>Acc. [%]</i>	Acc. [%]
Contr1	0.84	22.33	87.44	1.48	49.07	72.56	80.00
Contr2	0.00	51.23	72.05	0.00	70.77	61.40	66.73
unina2	4.36	18.90	87.71	0.04	99.97	45.45	66.58
unina3	1.88	16.50	90.15	0.00	99.97	45.47	67.81
jiiiov	0.36	52.13	71.40	1.08	37.53	79.03	75.22
jiiiov_all	2.40	30.13	82.47	1.08	63.83	64.60	73.54