Multi-challenge database for active liveness

Bruno Kamarowski*, Raul Almeida*, Bernardo Biesseck*, Roger Granada†, Gustavo Führ†, David Menotti*
* Federal University of Paraná, Curitiba, PR, Brazil {bhkamarowski, rgpalmeida, bernardo, menotti}@inf.ufpr.br
‡unico - idTech, Brazil {roger.granada, gustavo.fuhr}@unico.io

***Abstract*—Facial authentication on mobile devices has been widely applied in various scenarios. The field of Face Liveness (or Face Anti-Spoofing, FAS) focuses on methods and tools for detecting attacks (spoofs) where a malicious user tries to impersonate someone else or obfuscate their own identity. The specific problem of Active Liveness consists of analyzing the input signal and also the user behavior while performing some required challenge to determine whether the presented face is real or not. Despite the large amount of public datasets in FAS, very few contemplate active liveness, a phenomenon that typically results in new solutions being developed and evaluated with in-house data that cannot be shared due to its sensitive nature. This disjoint character in evaluations leads to irreproducible works and greatly hinders the mutual contribution inside the scientific community. In this paper, we describe an approach for creating a new database for active liveness detection. In this database, volunteers use a mobile application (under alpha testing for Android and iOS devices) to record themselves executing three distinct interactions (challenges) with their heads, namely face close up, head orientation, and flashes. Users will be encouraged to acquire videos in different environments and times of day, which will contribute to high variance in the dataset. After all acquisitions, we will split the dataset into training and testing protocols.**

I. Introduction

As virtual interfaces become more popular, an issue emerges: ensuring user authenticity in the system [1]. Some verification methods use facial biometrics from images and not only identify the user but distinguish real faces (indicating legitimacy) from fake ones (suggesting a potential impersonation attack) [2], [3]. This is essential to avoid attackers maliciously impersonate some victim by using a presentation attack instrument (PAI), like a photo or video. The field related to the ability of a system to determine whether the presented face image belongs to a genuine person, as opposed to a fake or artificial representation is called facial liveness, face anti-spoofing, or face presentation attack detection (Face PAD) [4].

To counteract malicious liveness activities, face anti-spoofing systems are developed, known as fake face detectors [4]. There are two primary branches of face anti-spoofing: active and passive. Active methods involve prompting the user to perform specific actions or gestures to verify if a living person is in front of the camera during the authentication process. Nevertheless, passive methods rely on the system's ability to detect signs of vitality in the presented facial image or video without requiring any deliberate user interaction. Within the realm of active liveness, a wide array of user interactions can be used, including involuntary signals or physiological reflexes, such as spontaneous head or eye movements, blinking, or pupil dilation. They can also be induced by displaying light or dark patterns on the user's screen [5]–[7]. Another approach involves injecting information during capture and analyzing its response, such as sound or light pattern, to determine whether it interacted with a real face or a PAI [8]–[11]. A more intrusive approach involves requesting the user to solve a challenge, such as smiling, nodding, or blinking voluntarily [12], [13]. Such methods depend on the successful completion of the proposed challenges, while also leveraging their dynamic aspects to extract additional information. For example, the 3D facial properties can be analyzed as the user's face gets closer and moves away from the camera [14].

In recent years, passive FAS methods have shifted from handcrafted features to learned feature maps, which require a substantial amount of data for training. While there are numerous public databases available for development and testing of passive methods, scientific works on active approaches typically depend on using in-house databases and report their results within them, making it challenging to achieve a fair comparison between similar approaches and impossible to reproduce experiments. To overcome such issues, this work proposes the creation of a new active liveness database containing facial videos acquired by volunteer subjects with their own smartphones. The proposed collection methodology ensures high variance in camera model, background, ethnicity, gender, and illumination.

There is no consensus in the literature on the most efficient user interaction approach. We therefore incorporate three different approaches: two based on head movement and a third involving emitting light patterns from the device screen.

II. Related Work

In this section, we will present studies related to FAS. Subsection II-A discusses important datasets in the field of passive liveness, Subsection II-B summarizes studied passive face anti-spoofing methods, and Subsection II-C presents studied active face anti-spoofing methods.

A. Datasets for Passive Facial liveness

With the growth in the field of facial anti-spoofing research, there have been various approaches to dataset collection and determining the most relevant information for liveness assessment. Currently, there is a focus on sample abundance and variability in acquisition conditions (such as user movement, camera quality, and ambient lighting), user characteristics (gender, ethnicity, etc.), and attack methods. Table I summarizes the studied datasets and includes our own, with estimates w.r.t. data volume.

TABLE I
STUDIED DATASETS' MAIN CHARACTERISTICS.

Dataset	Samples	Subjects	Attack types	User interaction
NUAA [15]	5105 real, 7509 fake	15	1	Passive
PRINT-ATTACK [16]	200 real, 200 fake	50	1	Passive
CASIA [17]	150 real, 450 fake	50	3	Passive
Replay-Attack [18]	200 real, 1000 fake	50	3	Passive
MSU-MFSD [19]	110 real, 330 fake	55	3	Passive
MSU-USSA [20]	1140 real, 9120 fake	1140	2	Passive
MLFP [21]	150 real, 1200 fake	10	2	Passive
Oulu-NPU [22]	990 real, 3960 fake	55	4	Passive
SiW [23]	1320 real, 3300 fake	165	6	Multiple angles, face expressions and the subjects move
SiW-M [24]	660 real, 968 fake	493	13	Passive
HQ-WMCA [25]	555 real, 2349 fake	51	10	Passive
DMAD [26]	900 real, 1800 fake	300	6	Passive
Celeb A-Spoof [27]	156,384 real, 469,153 fake	10,177	6	Passive
WFAS [28]	529,571 real, 853,729 fake	469,920	18	Passive
Ours (work in progress)	1200 to 6000 real, 750 to 3000 fake	200 to 1000	at least 3	Close up, face orientation and flashes

It is noteworthy, however, that all presented datasets are primarily intended for passive liveness studies, as subjects were not exposed to external stimuli or instructed to perform specific actions to interact with the system at the time of sample capture. The only listed work where users perform non-trivial activities is SiW [23], where the authors' intention was not to simulate the active liveness scenario found in many recent applications, but to enhance variety in data. To the best of our knowledge, no previous public dataset simulates active interactions in the scope presented in this work.

We expect our new database to contain at least 200 subjects, aiming to reach 1000. Each subject will record 2 sessions of the proposed 3 challenges, producing 6 videos of 20 seconds in length on average, with 20 frames per second (fps). This will result in at least 1200 videos of real faces. Estimating the number of 1000 volunteers, the dataset will contain 6000 videos of real faces.

B. Passive Facial liveness Mechanisms

The studied methods for FAS will be presented in this section. The strategies used by each method vary based on the approach and even the technology used, but they can be grouped into several categories.

Throughout the evolution of liveness studies, strategies for FAS have transitioned from detecting constructed features to learned feature maps. Among the first type of strategies, Boulkenafet et al. [29] describe facial appearance by applying Fisher vector encoding to features extracted from different color spaces for FAS. Wang et al. [30] also combine handcrafted 3D features with image properties for liveness detection. More recently, Liu et al. [31] fuse low and high-level features from different modalities to enhance sample representation.

While more traditional FAS methods rely on handcrafted features, such as those from Local Binary Patterns [18], recent approaches depend heavily on deep learning, with a clear division between traditional deep learning applications (binary cross-entropy- [31]–[35] or pixel-wise supervised [36]–[41]) and generalized deep learning methods [42]–[44], which aim to train models that generalize either to unseen domains or unknown attack types. Examples of pixel-wise supervision and domain generalization are, respectively, the DC-CDN network [36], which produces a face depth map as output, and the IADG method [42], which whitens instance-specific features to avoid domain bias.

C. Active Facial Liveness Mechanisms

As mentioned earlier, active methods depend on user interaction for liveness detection. They can be classified into three main lines: based on involuntary interaction, based on voluntary interaction, and injected information.

Face anti-spoofing based on involuntary interaction typically employs features from natural physiological movements. Some works extract specific features [7], [45], classifying a sample as real or spoof based on blinking patterns and lip movement patterns. Pupil movement is also used as a cue for liveness detection [6], and it has been experimented with combining more of such cues [5] (namely blinking, mouth movements, face-background consistency and other aspects of samples) for facial liveness detection. In [46], [47] remote photoplethysmography (rPPG) is used to analyze heart rate-related information in videos, resulting in a proposed Convolutional Attention Network to detect attacks.

In systems relying on user cooperation, also known as challenge-response systems, the user is instructed to perform simple actions. For instance, [48], [49] require the user to follow a displayed pattern with their eyes or to point their eyes at a designated point on the screen. If the user fails the challenge or completes it with suspicious patterns, they are classified as spoofed. [50]–[52] asks the user to pronounce a randomized sequence of words, and based on the consistency between mouth and face movement with the audio sample, liveness is detected. [14] locates facial landmarks in frames where the user is close to and far from the recording device, calculates the distances between these points in each frame, and uses these distances as input for a classifier. [12], [13] prompts the user to display a random sequence of emotions and classifies the images of each emotion as real or spoof using a CNN. Additionally, there are approaches based on

head movements to verify facial three-dimensionality through projective invariants [53].

In strategies based on injected data, additional information is introduced during media capture. Studied works following this trend include [8] and [9], which emit a light pattern and use a CNN for depth map recovery and liveness classification, along with a regression branch for light CAPTCHA checking to search for the injected pattern in the user's face and eyes. In [10], it is also emitted light during acquisition but enhances 3D features. In [11], it is emitted sound signals while the user is engaged in a simple task, analyzes the recovered signal (the echo of the emitted signal) to extract 3D facial geometry properties, and feeds these properties to an SVM classifier.

III. Methodology

The approach for the new active FAS database involves collecting videos from volunteer users who will participate in two recording sessions at different times. Each session consists of recording three videos while solving liveness challenges and uploading them to a server. Since there are no previous public active liveness databases to base our work on, there is no specific standard or protocol to active liveness data collection.

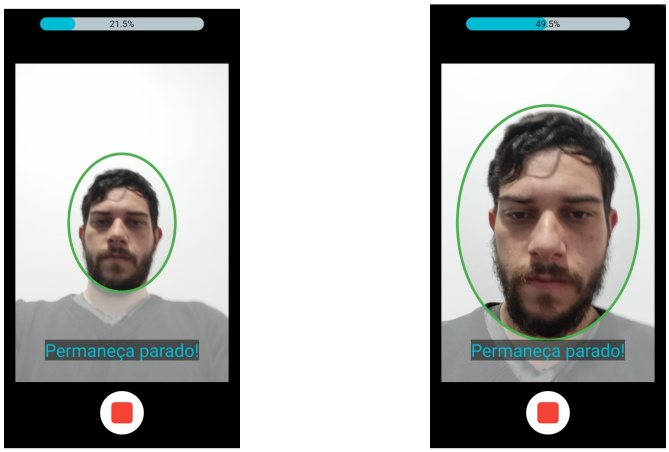
To adhere to research and data collection standards, an application was developed for Android and iOS systems using the Dart language, Flutter framework, and implementations of public packages for the mentioned framework. This application provides a simple interface for volunteers to register, receive challenge instructions, record a video while solving the challenge, and upload the generated content to the server. The mentioned functionalities will be described below.

The application implements a user registration system, collecting only identification information such as gender, age, and agreement with the consent terms. These data will be used later for diversity and balancing indicators in the database.

Additionally, the application provides the infrastructure for users to record videos while solving each of the challenges in each session and, at a later time, send the videos to a server that stores these files. During a session, the user is invited to solve the three proposed liveness challenges: close-up focus on the face, face orientation, and flashes. After submitting the files for a session, the user must wait eight hours before being able to conduct another session. This measure was adopted to introduce more variability into the data, as it is expected that the same user may be in different environments, lighting conditions, and appearances in each session, enhancing the quality of the database.

The application offers users real-time feedback on the progression of challenges. To achieve this, a facial landmark detector and a facial angle estimator were employed. These detectors identify the facial position and angles in each frame. They assess whether the current face position is centered, determine the direction the face is pointing during head movement challenges, and confirm if the face aligns with the expected region as indicated on the screen. However, it has been noted that there is a trade-off between resolution and frame rate. When frames have resolutions higher than 1080p, the application becomes less responsive, impacting its usability. Consequently, a decision was made to establish 720p as the default resolution for the dataset.

The "close up" challenge has two stages: In the first one, the user must align their face within a small region indicated on the screen so that when aligned, the user is distant from the device. Once aligned, the user must maintain this position for a brief moment, and then the second stage will start. In the second stage, the user is asked to align their face once again within a region indicated on the screen, but this time the region is large, causing the user to approach the screen. This adds a dynamic aspect to the challenge, making simpler attack schemes more difficult and enabling the extraction of three-dimensional information from the displayed face. Figure 1 presents the two stages of this challenge.



(a) (b)

Fig. 1. 1a - Distant alignment step; 1a - Close alignment step.

In the first stage, the user must align their face with the center of the screen, similar to the close-up challenge. The face orientation challenge involves aligning the face and positioning the head toward four random directions. In the initial stage, the user is required to center their face on the screen, similar to the close-up challenge. Subsequently, in each subsequent stage, the following sequence unfolds: A random face orientation is designated, and the user is prompted to adjust their head to match the specified orientation. The user needs to maintain this pose momentarily before proceeding to the next step. The potential face orientations encompass facing upwards, downwards, leftwards, and rightwards, as illustrated in Figure 2. This challenge introduces a heightened level of complexity for attacks, as resolving it necessitates the attacker's emulation of the victim's face in various orientations.

The flashes challenge requires the user to stay still for a few seconds as the application's screen emits six patterns of varying lights. These lights are presented in a variety of colors and are distributed across the screen's four quadrants. The patterns are generated at random, with each pattern assigned

(a) (b) (c) (d)

Fig. 2. 2a - Upward pose; 2b - Downward pose; 2c - Rightward pose; 2d - Leftward pose

to a specific quadrant and colored according to blue, yellow, green, or red. This challenge adds extra information during capture, specifically in the form of screen lighting patterns. This is based on the assumption that such added information will interact uniquely with real and fake faces, leading to distinguishable differences in the captured light signals. To see an example of solving this challenge, refer to Figure 3.

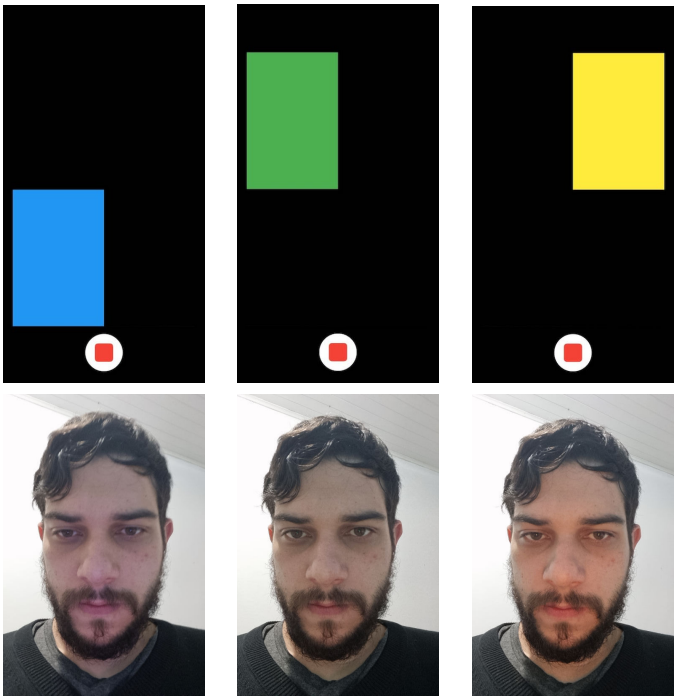


Fig. 3. In the first line, three flash patterns are shown, and in the line below, the reflection effect caused on the user's face by the respective flashes is displayed.

A public celebrity face database [54] will be utilized to select a subset of 200 to 1000 individuals (depending on the number of real subjects in our final database). This subset, in conjunction with the outlined application, will be employed to generate the spoof samples for each of the three challenges. We choose to use these images instead of the database participants due to the sensitive nature of producing spoofs. We aim to execute fundamental attack types, encompassing photo-based methods involving printed media and digital devices. Moreover, video-based attacks will also be carried out. The extent of attacks will be adjusted in accordance with the available resources, using 3D face projection viewers to create head movements such as in [55], [56].

Once the database is created, training and testing protocols will be developed, along with establishing metrics to evaluate experiments on the database.

The database will be made available to the scientific community for model training and evaluation in a closed manner (by us), allowing experiments and ensuring reproducibility without exposing sensitive data. We will also publish baseline results on popular state-of-the-art models and in a detector of our own, which will be presented in the same work. These models are yet to be defined.

IV. Conclusion

The contributions of the project are primarily two-fold. First, it involves mitigating the challenge of creating a new database for scientific studies on active liveness and, thereby, facilitating future work in this area by providing a database dedicated for active liveness, which is often scarce in academic circles. Secondly, by establishing training protocols and metrics, it enables a fair comparison of active Face Anti-Spoofing (FAS) methods that utilize any of the three interaction approaches covered by the database.

We plan on documenting the challenges in producing spoofs for each type of user interaction, which are themselves obstacles to malicious users. Furthermore, experiments on this dataset will enable a better understanding of how different user interactions make face liveness detection easier. We expect these findings to guide researchers and system designers in balancing trade-offs between cost and effectiveness of implementing each user challenge.

References

[1] D. Gollmann, "Computer security," *Wiley Interdisciplinary Reviews: Computational Statistics*, vol. 2, no. 5, pp. 544–554, 2010.

[2] S. Z. S. Idrus, E. Cherrier, C. Rosenberger, and J.-J. Schwartzmann, "A review on authentication methods," *Australian Journal of Basic and Applied Sciences*, vol. 7, no. 5, pp. 95–107, 2013.

[3] L. Li, P. L. Correia, and A. Hadid, "Face recognition under spoofing attacks: countermeasures and research directions," *IET Biometrics*, vol. 7, no. 1, pp. 3–14, 2018.

[4] Z. Yu, Y. Qin, X. Li, C. Zhao, Z. Lei, and G. Zhao, "Deep learning for face anti-spoofing: A survey," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 45, no. 5, pp. 5609–5631, 2022.

[5] J. Yan, Z. Zhang, Z. Lei, D. Yi, and S. Z. Li, "Face liveness detection by exploring multiple scenic clues," *2012 12th Int. Conf. on Control Automation Robotics & Vision (ICARCV)*, pp. 188–193, 2012.

[6] M. Killioğlu, M. Taşkiran, and N. Kahraman, "Anti-spoofing in face recognition with liveness detection using pupil tracking," in *2017 IEEE 15th International Symposium on Applied Machine Intelligence and Informatics (SAMI)*, 2017, pp. 000 087–000 092.

[7] M. Singh and A. Arora, "A robust anti-spoofing technique for face liveness detection with morphological operations," *Optik*, vol. 139, pp. 347–354, 2017. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0030402617303935

[8] M. Mohzary, K. J. Almalki, B.-Y. Choi, and S. Song, "Your eyes show what your eyes see (y-eyes): Challenge-response anti-spoofing method for mobile security using corneal specular reflections," in *1st Workshop on Security and Privacy for Mobile AI*, ser. MAISP'21. New York, NY, USA: Association for Computing Machinery, 2021, p. 25–30. [Online]. Available: https://doi.org/10.1145/3469261.3469408

[9] Y. Liu, Y. Tai, J. Li, S. Ding, C. Wang, F. Huang, D. Li, W. Qi, and R. Ji, "Aurora guard: Real-time face anti-spoofing via light reflection," *CoRR*, vol. abs/1902.10311, 2019. [Online]. Available: http://arxiv.org/abs/1902.10311

[10] J. M. D. Martino, Q. Qiu, T. Nagenalli, and G. Sapiro, "Liveness detection using implicit 3D features," *CoRR*, vol. abs/1804.06702, 2018. [Online]. Available: http://arxiv.org/abs/1804.06702

[11] W. Xu, J. Liu, S. Zhang, Y. Zheng, F. Lin, J. Han, F. Xiao, and K. Ren, "Rface: Anti-spoofing facial authentication using cots rfid," in *IEEE INFOCOM 2021 - IEEE Conference on Computer Communications*, 2021, pp. 1–10.

[12] M. Ezz, M. Ayman, and A. Elshenawy Elsefy, "Challenge-response emotion authentication algorithm using modified horizontal deep learning," *Intelligent Automation and Soft Computing*, vol. 35, pp. 3659–3675, 09 2022.

[13] Z. Ming, J. Chazalon, M. M. Luqman, M. Visani, and J.-C. Burie, "Facelivenet: End-to-end networks combining face verification with interactive facial expression-based liveness detection," in *2018 24th Int. Conf. on Pattern Recognition (ICPR)*. IEEE, 2018, pp. 3507–3512.

[14] Y. Li, Z. Wang, Y. Li, R. Deng, B. Chen, W. Meng, and H. Li, "A closer look tells more: A facial distortion based liveness detection for face authentication," in *Asia CCS '19: ACM Asia Conference on Computer and Communications Security*, 07 2019, pp. 241–246.

[15] X. Tan, Y. Li, J. Liu, and L. Jiang, "Face liveness detection from a single image with sparse low rank bilinear discriminative model," in *Computer Vision – ECCV 2010*, K. Daniilidis, P. Maragos, and N. Paragios, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2010, pp. 504–517.

[16] A. Anjos and S. Marcel, "Countermeasures to photo attacks in face recognition: A public database and a baseline," in *2011 International Joint Conference on Biometrics (IJCB)*, 2011, pp. 1–7.

[17] Z. Zhang, J. Yan, S. Liu, Z. Lei, D. Yi, and S. Z. Li, "A face antispoofing database with diverse attacks," in *2012 5th IAPR Int. Conf. on Biometrics (ICB)*, 2012, pp. 26–31.

[18] I. Chingovska, A. Anjos, and S. Marcel, "On the effectiveness of local binary patterns in face anti-spoofing," in *2012 BIOSIG - Int. Conf. of Biometrics Special Interest Group (BIOSIG)*, 2012, pp. 1–7.

[19] D. Wen, H. Han, and A. K. Jain, "Face spoof detection with image distortion analysis," *IEEE Transactions on Information Forensics and Security*, vol. 10, no. 4, pp. 746–761, 2015.

[20] K. Patel, H. Han, and A. K. Jain, "Secure face unlock: Spoof detection on smartphones," *IEEE Transactions on Information Forensics and Security*, vol. 11, no. 10, pp. 2268–2283, 2016.

[21] A. Agarwal, D. Yadav, N. Kohli, R. Singh, M. Vatsa, and A. Noore, "Face presentation attack with latex masks in multispectral videos," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, July 2017.

[22] Z. Boulkenafet, J. Komulainen, L. Li, X. Feng, and A. Hadid, "OULU-NPU: A mobile face presentation attack database with real-world variations," in *2017 12th IEEE Int. Conf. on Automatic Face & Gesture Recognition (FG 2017)*, 2017, pp. 612–618.

[23] Y. Liu, A. Jourabloo, and X. Liu, "Learning deep models for face anti-spoofing: Binary or auxiliary supervision," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2018.

[24] Y. Liu, J. Stehouwer, A. Jourabloo, and X. Liu, "Deep tree learning for zero-shot face anti-spoofing," in *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2019, pp. 4675–4684.

[25] Z. Mostaani, A. George, G. Heusch, D. Geissbühler, and S. Marcel, "The high-quality wide multi-channel attack (HQ-WMCA) database," *CoRR*, vol. abs/2009.09703, 2020. [Online]. Available: https://arxiv.org/abs/2009.09703

[26] Z. Wang, Z. Yu, C. Zhao, X. Zhu, Y. Qin, Q. Zhou, F. Zhou, and Z. Lei, "Deep spatial gradient and temporal depth learning for face anti-spoofing," in *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2020, pp. 5041–5050.

[27] Y. Zhang, Z. Yin, Y. Li, G. Yin, J. Yan, J. Shao, and Z. Liu, "CelebA-spoof: Large-scale face anti-spoofing dataset with rich annotations," in *Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part XII 16*. Springer, 2020, pp. 70–85.

[28] D. Wang, J. Guo, Q. Shao, H. He, Z. Chen, C. Xiao, A. Liu, S. Escalera, H. J. Escalante, Z. Lei *et al.*, "Wild face anti-spoofing challenge 2023: Benchmark and results," in *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2023, pp. 6379–6390.

[29] Z. Boulkenafet, J. Komulainen, and A. Hadid, "Face antispoofing using speeded-up robust features and fisher vector encoding," *IEEE Signal Processing Letters*, vol. 24, no. 2, pp. 141–145, 2017.

[30] T. Wang, J. Yang, Z. Lei, S. Liao, and S. Z. Li, "Face liveness detection using 3D structure recovered from a single camera," in *2013 Int. Conf. on Biometrics (ICB)*, 2013, pp. 1–6.

[31] W. Liu, X. Wei, T. Lei, X. Wang, H. Meng, and A. K. Nandi, "Data fusion based two-stage cascade framework for multi-modality face anti-spoofing," *IEEE Transactions on Cognitive and Developmental Systems*, pp. 1–1, 2021.

[32] S. Garg, S. Mittal, P. Kumar, and V. Anant Athavale, "DeBNet: Multi-layer deep network for liveness detection in face recognition system," in *2020 7th Int. Conf. on Signal Processing and Integrated Networks (SPIN)*, 2020, pp. 1136–1141.

[33] X. Yang, W. Luo, L. Bao, Y. Gao, D. Gong, S. Zheng, Z. Li, and W. Liu, "Face anti-spoofing: Model matters, so does data," in *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2019, pp. 3502–3511.

[34] S. Luo, M. Kan, S. Wu, X. Chen, and S. Shan, "Face anti-spoofing with multi-scale information," in *2018 24th Int. Conf. on Pattern Recognition (ICPR)*, 2018, pp. 3402–3407.

[35] Z. Yu, C. Zhao, Z. Wang, Y. Qin, Z. Su, X. Li, F. Zhou, and G. Zhao, "Searching central difference convolutional networks for face anti-spoofing," in *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2020, pp. 5294–5304.

[36] Z. Yu, Y. Qin, H. Zhao, X. Li, and G. Zhao, "Dual-cross central difference network for face anti-spoofing," *CoRR*, vol. abs/2105.01290, 2021. [Online]. Available: https://arxiv.org/abs/2105.01290

[37] W. Zheng, M. Yue, S. Zhao, and S. Liu, "Attention-based spatial-temporal multi-scale network for face anti-spoofing," *IEEE Transactions on Biometrics, Behavior, and Identity Science*, vol. 3, no. 3, pp. 296–307, 2021.

[38] Y. Wang, X. Song, T. Xu, Z. Feng, and X.-J. Wu, "From RGB to depth: Domain transfer network for face anti-spoofing," *IEEE Transactions on Information Forensics and Security*, vol. 16, pp. 4280–4290, 2021.

[39] K.-Y. Zhang, T. Yao, J. Zhang, Y. Tai, S. Ding, J. Li, F. Huang, H. Song, and L. Ma, "Face anti-spoofing via disentangled representation learning," in *European Conference on Computer Vision*. Springer, 2020, pp. 641–657.

[40] Z. Wang, Z. Yu, C. Zhao, X. Zhu, Y. Qin, Q. Zhou, F. Zhou, and Z. Lei, "Deep spatial gradient and temporal depth learning for face anti-spoofing," 2020. [Online]. Available: https://arxiv.org/abs/2003.08061

[41] Z. Yu, C. Zhao, Z. Wang, Y. Qin, Z. Su, X. Li, F. Zhou, and G. Zhao, "Searching central difference convolutional networks for face anti-spoofing," *CoRR*, vol. abs/2003.04092, 2020. [Online]. Available: https://arxiv.org/abs/2003.04092

[42] Q. Zhou, K.-Y. Zhang, T. Yao, X. Lu, R. Yi, S. Ding, and L. Ma, "Instance-aware domain generalization for face anti-spoofing," 2023.

[43] Y. Sun, Y. Liu, X. Liu, Y. Li, and W.-S. Chu, "Rethinking domain generalization for face anti-spoofing: Separability and alignment," 2023.

[44] Z. Wang, Z. Wang, Z. Yu, W. Deng, J. Li, T. Gao, and Z. Wang, "Domain generalization via shuffled style assembly for face anti-spoofing," 2022. [Online]. Available: https://arxiv.org/abs/2203.05340

[45] M. Singh and A. S. Arora, "A novel face liveness detection algorithm with multiple liveness indicators," *Wireless Personal Communications*, vol. 100, no. 4, pp. 1677–1687, Jun 2018. [Online]. Available: https://doi.org/10.1007/s11277-018-5661-1

[46] J. Hernandez-Ortega, J. Fierrez, A. Morales, and D. Díaz, "A comparative evaluation of heart rate estimation methods using face videos," *2020 IEEE 44th Annual Computers, Software, and Applications Conference (COMPSAC)*, pp. 1438–1443, 2020.

[47] J. Hernandez-Ortega, R. Tolosana, J. Fierrez, and A. Morales, "Deepfakeson-phys: Deepfakes detection based on heart rate estimation," *ArXiv*, vol. abs/2010.00400, 2020.

[48] I. Sluganovic, M. Roeschlin, K. Rasmussen, and I. Martinovic, "Using reflexive eye movements for fast challenge-response authentication," in *CCS '16: ACM SIGSAC Conference on Computer and Communications Security*, 10 2016, pp. 1056–1067.

[49] M. Shen, Z. Liao, L. Zhu, R. Mijumbi, X. Du, and J. Hu, "Iritrack: Liveness detection using irises tracking for preventing face spoofing attacks," 2018. [Online]. Available: https://arxiv.org/abs/1810.03323
[50] P. McShane and D. Stewart, "Challenge based visual speech recognition using deep learning," in *2017 12th Int. Conf. for Internet Technology and Secured Transactions (ICITST)*, 2017, pp. 405–410.
[51] E. Uzun, S. Chung, I. Essa, and W. Lee, "rtCaptcha: A real-time captcha based liveness detection system," in *Conference: The Network and Distributed System Security Symposium (NDSS)*, 02 2018.
[52] C.-L. Chou, "Presentation attack detection based on score level fusion and challenge-response technique," *The Journal of Supercomputing*, vol. 77, 05 2021.
[53] M. De Marsico, M. Nappi, D. Riccio, and J.-L. Dugelay, "Moving face spoofing detection via 3D projective invariants," in *2012 5th IAPR Int. Conf. on Biometrics (ICB)*. IEEE, 2012, pp. 73–78.
[54] Z. Liu, P. Luo, X. Wang, and X. Tang, "Large-scale celebfaces attributes (celeba) dataset," *Retrieved August*, vol. 15, no. 2018, p. 11, 2018.
[55] "Spark AR," https://spark.meta.com/learn/articles/people-tracking/face-mesh#face-mesh-properties, accessed on: August 21, 2023.
[56] "Reface," https://reface.ai/, accessed on: August 21, 2023.