

Face Liveness Detection Competition (LivDet-Face) - 2026

Lambert Igene^{1†}, Mohammad Zahir Uddin Chowdhury¹, Mst Rumana Sumi¹, Jahangir Khondkar¹,
Stephanie Schuckers¹, ¹Clarkson University, USA, ¹ University of North Carolina @ Charlotte,

[†]igenela@clarkson.edu

Juan E. Tapia^{*2} Lázaro Janier González-Soler^{*2} Christoph Busch^{*2} Vasilije Pantić^{*3}

Aleksandr Andrusenko^{*3} Efim Boieru^{*3} Alejandro Hinke^{*4}

Raúl Manchón-Alcaraz^{*4} Mario Nieto-Hidalgo^{*4} Juan M. Espín^{*4} Gabriel García^{*4}

João C. Monteiro^{*5} João Ribeiro Pinto^{*5} Pedro Couto Pereira^{*5}

Mariana Xavier^{*5} Luciano Tarouco^{*5} César Gouveia^{*5}

João Ferreira^{*5} Liepiao Zhang^{*6} Hui Huang^{*6}

Wenxiong Kang^{*6} Keivan Bahmani^{*7} Becky Gross^{*7} Mikhail Pozigun^{*7}

Anton Sinkov^{*8} Andrey Parasich^{*8} Anant Gupta^{*9} Gourav Gupta^{*9}

²Hochschule Darmstadt (HDA), Darmstadt, Germany. ³Incode Technologies, Inc, CA, USA.

⁶South China University of Technology, Guangzhou, China.

⁷Mitek Systems, CA, USA. ⁸ 3DiVi Inc, Kazakhstan, Astana. ⁹ArogyaPandit Research, India

^{*}Competitors

Abstract

Face Presentation Attack Detection (PAD) remains a critical challenge in biometric security due to the increasing sophistication of spoofing attacks, including replay attacks, printed photos, projection attacks, and three-dimensional masks. LivDet-Face 2026 continues the LivDet series by providing an international benchmark for evaluating state-of-the-art face liveness detection algorithms and systems under standardized testing protocols. Organized in conjunction with the International Joint Conference on Biometrics (IJCB), the competition offers an independent and comprehensive evaluation framework designed to measure the robustness and generalization capabilities of modern face PAD solutions against diverse and previously unseen Presentation Attack Instruments (PAIs). No official training dataset was provided to participants, and all evaluation data remained sequestered throughout the competition, ensuring an unbiased assessment of algorithm generalization to unseen attack scenarios. The competition consisted of two primary categories: algorithms and systems. A total of nineteen algorithms and one system were evaluated across image and video-based subcategories using a diverse dataset containing bona fide presentations and multiple attack types collected under realistic acquisition conditions. Mitek's submission achieved the best perfor-

mance in the image subcategory with an ACER of 0.56% (APCER = 1.12%, BPCER = 0%), while Incode submissions achieved the best performance in the video subcategory with an ACER of 4.74% (APCER = 0.98%, BPCER = 8.5%). In the systems category, team anonymous achieved an ACER of 13.64% (APCER = 20.6%, BPCER = 6.67%). The results demonstrate significant progress in face PAD performance while also highlighting the continuing challenges posed by advanced and unseen presentation attacks in real-world biometric deployments.

1. Introduction

Face biometric recognition systems have become an essential component of modern identity management and authentication applications due to their convenience, usability, and high recognition accuracy. These systems are now widely deployed across border control, mobile authentication, financial services, healthcare, and access control applications. Despite these advantages, face recognition systems remain vulnerable to presentation attacks, including printed photos, replay attacks, projection attacks, and three-dimensional masks, which can compromise system security and reliability [20, 2, 7, 15]. Such attacks are designed to impersonate legitimate users, conceal identities, or fraudulently gain access to protected systems [10].

To mitigate these threats, robust Presentation Attack Detection (PAD) mechanisms are required to distinguish bona fide presentations from attack presentations. Face PAD has therefore become a critical research area in biometric security, attracting substantial attention from both academia and industry [1, 8, 17]. Although recent deep learning-based PAD approaches have demonstrated strong performance on benchmark datasets [5, 4], many methods still face challenges in real-world deployments due to variations in sensors, environments, illumination conditions, acquisition quality, and previously unseen Presentation Attack Instruments (PAIs) [20, 3]. Improving the robustness and generalization capability of face PAD systems against diverse and unknown attacks therefore remains an open research challenge.

Consequently, large-scale and standardized evaluation platforms are essential for benchmarking the effectiveness and generalization capability of modern PAD algorithms. In parallel, several face biometric datasets have been introduced to evaluate recognition, detection, and security performance under realistic acquisition conditions, including CelebA-Spoof [23], CASIA-FASD [5], and Replay-Attack [4]. These datasets highlight the challenges posed by presentation attacks, facial occlusions, and operational variability, further emphasizing the need for robust and generalizable PAD solutions capable of maintaining reliable performance under real-world deployment scenarios.

To address these challenges, LivDet-Face 2026 continues the LivDet competition series by providing a standardized international benchmark for evaluating state-of-the-art face liveness detection algorithms and systems under realistic and diverse attack scenarios. The competition enables independent evaluation of submitted solutions using unseen test data, multiple sensors, and a broad range of attack instruments designed to simulate real-world presentation attack conditions.

The LivDet-Face 2026 competition includes two evaluation tracks: algorithms and systems. The algorithm track evaluates software-only PAD solutions submitted by participants and tested independently by the organizers using sequestered image and video datasets. In contrast, the system track evaluates complete biometric systems consisting of both hardware and software components operating in an end-to-end manner. A total of nineteen algorithms and one system were evaluated under standardized protocols using ISO-compliant PAD metrics. The competition dataset contains a diverse set of bona fide presentations and presentation attacks collected across multiple sensors, environments, and attack categories.

Key contributions of LivDet-Face 2026 include:

- **Comprehensive State-of-the-Art Evaluation:** Independent evaluation of nineteen face PAD algorithms

and one integrated PAD system across image and video-based scenarios.

- **Standardized PAD Evaluation Protocol:** Evaluation protocols developed following Fast Identity Online (FIDO) biometric requirements [16] and ISO/IEC PAD evaluation guidelines [18].
- **Sequestered Evaluation Protocol:** No official training dataset was provided to participants, and all evaluation data remained sequestered throughout the competition, ensuring an independent and unbiased assessment of algorithm generalization to unseen attacks and acquisition conditions.
- **Diverse and Realistic PAI Categories:** Evaluation using a broad spectrum of presentation attack instruments, including replay attacks, print attacks, projection attacks, and three-dimensional mask attacks acquired under realistic conditions.
- **Evaluation on Unseen Attacks:** Assessment of algorithm robustness and generalization capability against previously unseen PAIs and acquisition conditions.
- **Introduction of Novel PAIs:** A novel PAI, namely makeup skin-tone 3D masks of bona fide subjects, was introduced.

1.1. Performance Evaluation Metrics

For LivDet-Face 2026, we maintain the evaluation metrics used in LivDet-Face 2024 [12]. APCER and BPCER, defined below, are the metrics specified by ISO/IEC 30107-3 [18]:

Attack Presentation Classification Error Rate (APCER): Measures the proportion of attack presentations (PAIs) incorrectly classified as genuine, indicating the system’s susceptibility to accepting fraudulent inputs. **Bona Fide Presentation Classification Error Rate (BPCER):** Quantifies the proportion of genuine presentations incorrectly classified as attacks, highlighting the system’s ability to reject legitimate inputs.

Overall performance is summarized by $APCER_{average}$, the *unweighted* mean of APCER across the ten PAI categories, so that each PAI type contributes equally regardless of its sample count and a large category cannot mask a small but security-critical one. For ranking, we additionally report ACER, the average of $APCER_{average}$ and BPCER. Unlike APCER and BPCER, ACER is not defined by ISO/IEC 30107-3 and has been deprecated in recent guidance; we retain it as the single scalar needed to rank submissions and determine a winner, and because participants were informed in advance of the fixed operating threshold (a score of 50) at which their algorithms would be evaluated. It also provides longitudinal comparability with LivDet-Face 2024. The

ISO/IEC 30107-3-compliant metrics, per-category APCER and BPCER, are reported in full in Tables 3–5 and constitute the primary results.

2. Comparison with Recent Face PAD Benchmarks and Challenges

Recent years have witnessed substantial progress in face PAD, driven by large-scale datasets such as CelebA-Spoof [23], CASIA-FASD [5], Face projection [11] and Replay-Attack [4], by public challenges including the CVPR Face Anti-Spoofing Challenges [22] and the Wild Face Anti-Spoofing Challenge [19], and by the LivDet-Face competition series [12]. Table 1 compares LivDet-Face 2026 against representative prior datasets in terms of scale, sensor diversity, PAI coverage, and evaluation protocol, and summarizes how it extends each of them. Notably, several widely used benchmarks release their data publicly for training and evaluation, whereas LivDet-Face keeps its evaluation data sequestered to test generalization to genuinely unseen attacks.

Compared with LivDet-Face 2024, the 2026 edition uses a substantially smaller and less sensor-diverse test set (two sensors versus nine), a deliberate trade-off favoring a broader range of realistic mask-based PAI types, including the newly introduced makeup skin-tone mask, rather than scaling sensor and environmental diversity. We discuss the implications of this trade-off for statistical reliability in Section 4.3.

3. Competition Design

3.1. LivDet-Face 2026

LivDet-Face 2026, co-organized by Clarkson University (USA) and UNC Charlotte, is the third edition of the LivDet competition focused on face PAD [14]. The competition evaluates state-of-the-art face PAD algorithms and systems against both conventional and newly introduced PAIs. Similar to previous editions, the competition consists of two primary categories: Algorithm and System, where the Algorithm category is further divided into Image and Video subcategories. The Algorithm category evaluates software-only PAD solutions, submitted as standalone executables or Docker containers, which the organizers run offline against sequestered image or video files. The System category, by contrast, evaluates complete end-to-end biometric capture systems, combining the competitor’s own sensor hardware with its PAD software, against physical presentation attack instruments presented live to the sensor rather than pre-recorded files.

The competition utilized ten PAI types for both subcategories: Laptop_Photo, Cloth_mask, Normal_Paper, Paper_Mask, Photo_Paper, HQ_3D_mask, LQ_3D_mask, Makeup_skintone_mask, Skintone_3D_Mask_with_Hair,

and Skintone_Filament_Mask, along with bona fide live presentations. The overall test samples and attack categories were not disclosed to the competitors in order to evaluate the robustness and generalization capability of submitted algorithms under unseen attack scenarios.

Algorithm performance was measured using an output score ranging from 0 to 100, with a threshold set at 50. Scores below 50 were classified as PAIs, while scores of 50 and above were classified as bona fide presentations. A score of -1000 was used to indicate undetected samples or other processing failures, as specified in the competition instructions [13]. During PAI evaluation, a score of -1000 was considered a correct rejection and therefore did not contribute to the Attack Presentation Classification Error Rate (APCER). However, for bona fide samples, a score of -1000 was considered an incorrect classification and was included in the Bona Fide Presentation Classification Error Rate (BPCER) calculation.

All evaluations were conducted independently by the competition organizers to ensure unbiased and standardized performance assessment.

4. Experimental Protocol

The LivDet-Face 2026 competition welcomed participation from international academic and industrial organizations, both anonymously and non-anonymously. Non-anonymous competitors were included as co-authors in the publication. A total of eleven teams registered globally, submitting nineteen algorithms for the image category, eleven algorithms for the video category, and one system submission. All submitted algorithms and systems were successfully evaluated under the competition protocol.

4.1. Training Dataset

No official training dataset was provided for LivDet-Face 2026. Participants were encouraged to use any available data, whether public or proprietary, to train their algorithms. To aid familiarization, the organizers shared a few examples of known PAIs. However, the remaining samples of the disclosed PAI types were kept unknown to competitors.

4.2. Test Dataset

The LivDet-Face 2026 competition utilized a diverse dataset collected at Clarkson University and UNC Charlotte, comprising 1,866 images (95 bona fide and 1,771 PAI samples) and 375 videos (94 bona fide and 281 PAI samples). The data was collected using two sensors: a DSLR camera and a Samsung Galaxy S24 Ultra smartphone, from 15 subjects under realistic acquisition conditions. The dataset includes both bona fide presentations and multiple PAI categories for image and video-based evaluation. Video samples were recorded under varied environmental and acquisition conditions to simulate real-world biometric

Table 1. Comparison of LivDet-Face 2026 with representative prior face PAD datasets. Image/Video columns report dataset scale; “Seq.” indicates whether evaluation data is sequestered (withheld from participants) rather than publicly released for training; the final column notes how LivDet-Face 2026 extends each.

Benchmark	Images	Videos	Sensors	PAIs	Seq.	How LivDet-Face 2026 extends it
CelebA-Spoof [23]	625,537	–	multiple (crowd-sourced)	10	No	Adds video evaluation and realistic 3D/skin-tone and makeup-enhanced mask attacks beyond controlled image-only capture.
CASIA-FASD [5]	–	600	3 (low/normal/high quality)	3	No	Adds image-based evaluation, a far broader set of physical PAIs, and a sequestered protocol with no released training data.
Replay-Attack [4]	–	1,300	2 (webcam, mobile)	3	No	Introduces realistic 3D and makeup-enhanced mask attacks well beyond print and replay, under a sequestered protocol.
LivDet-Face 2024 [12]	17,512	16,291	9	8	Yes	Introduces the novel makeup skin-tone 3D mask and grows participation (9→19 image, 7→11 video algorithms).
LivDet-Face 2026 (ours)	1,866	375	2	10	Yes	Image, video, and system tracks with sequestered evaluation and a new makeup skin-tone 3D mask PAI.

authentication scenarios. A summary of the dataset is provided in Table 2.

A subset of the evaluation data was curated by selecting captures spanning acquisition distances from approximately 1 m to 10 m from the larger MORSE-Face collection [6], in order to introduce realistic variation in subject-to-sensor distance. The MORSE-Face collection was not publicly available during the competition and was distributed only through controlled access; no competition participant was granted access to it. It will be made available to the research community following the conclusion of the competition. The specific images and videos used in LivDet-Face 2026 were independently selected for this competition and remained fully sequestered throughout, and no evaluation samples were shared with participants prior to or during evaluation. Consequently, all evaluation samples were effectively unseen by participants, ensuring that the reported results reflect algorithm robustness and generalization rather than familiarity with previously available data.

4.2.1 Test Dataset Details

Bona Fide Presentations: The dataset contains 95 bona fide face images and 94 bona fide face videos collected from 15 subjects using DSLR and smartphone sensors under realistic acquisition conditions.

Laptop Display Attacks: Laptop display attacks were created by displaying face images and videos on laptop screens and recapturing them using DSLR and smartphone cameras. The dataset contains 330 image samples and 15 video samples.

Normal Paper Attacks: Printed face images on standard paper were used to generate normal paper attacks. The dataset includes 299 image samples and 16 video samples collected under varied lighting and acquisition conditions.

Paper Mask Attacks: Paper mask attacks were created using printed facial photographs with eye-region cutouts to simulate realistic presentation attacks. The dataset contains 37 image samples and 37 video samples.

Photo Paper Attacks: High-quality face photographs printed on glossy photo paper were used to generate spoof presentations. The dataset includes 383 image samples and 15 video samples.

Cloth Mask Attacks: Colored cloth masks covering facial regions were used to simulate occlusion-based spoofing attacks. The dataset contains 17 image samples and 17 video samples.

3D Mask Attacks: The competition includes multiple types of three-dimensional mask attacks generated using 3D printed facial models reconstructed from subject face images. The dataset contains 10 low-quality masks and 8 high-quality masks captured under both image and video acquisition settings. The image dataset contains 70 HQ_3D_mask samples and 36 LQ_3D_mask samples, while the video dataset contains 69 HQ_3D_mask samples and 36 LQ_3D_mask samples.

Skin-tone 3D Mask Attacks: In addition to conventional 3D masks, the competition introduced realistic skin-tone masks and filament-based masks to evaluate algorithm robustness against advanced appearance-based spoofing attacks. The SkinTone_3D_Mask_with_Hair category contains 145 image samples and 31 video samples, while the SkinTone_Filament_Mask category contains 152 image samples and 30 video samples.

Makeup Skin-tone Mask Attacks: A novel makeup skin-tone mask category was introduced in LivDet-Face 2026. These attacks were generated using white resin 3D masks followed by manual hand painting and cosmetic makeup processing to simulate realistic human facial appearance and skin texture. The dataset contains 302 image

samples and 15 video samples.

This comprehensive dataset, with representative examples of the presentation attack instruments shown in Figure 1, provides a realistic and challenging evaluation environment for benchmarking the robustness and generalization capability of modern face PAD systems under real-world biometric deployment scenarios.

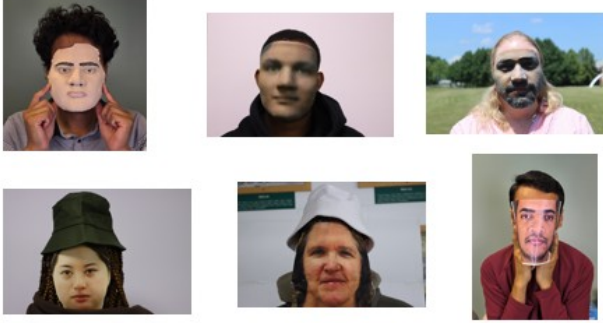


Figure 1. Example images of presentation attack types present in the LivDet-Face 2026 test dataset. Top (left to right): Makeup Skin-tone Mask, LQ 3D Mask, HQ 3D Mask, Paper Mask, Cloth Mask and Photo Paper.

Table 2. Test Dataset Summary

Class	Types of PAIs	Total Images	Total Videos
Bona fide	-	95	94
PAI	Paper Mask (PM)	37	37
PAI	Normal Paper (NP)	299	16
PAI	Photo Paper (PP)	383	15
PAI	Laptop Display (LP)	330	15
PAI	Cloth Mask (CM)	17	17
PAI	Makeup Skin-tone Mask (MM)	302	15
PAI	Skin-tone 3D Mask with Hair (SM-hair)	145	31
PAI	HQ 3D Mask (HQ)	70	69
PAI	LQ 3D Mask (LQ)	36	36
PAI	Skin-tone Filament Mask (SM)	152	30

4.3. Dataset Limitations, Ethics, and Demographics

Design considerations and limitations. As a sequestered PAD benchmark, LivDet-Face 2026 prioritizes breadth and realism of attack instruments over raw dataset scale: the evaluation set concentrates on ten diverse PAI categories, including newly introduced realistic 3D and makeup-enhanced masks, rather than maximizing image count or sensor variety. This design yields an attack-heavy test set, with bona fide-to-attack ratios of roughly 1:18 for images and 1:3 for videos. Because APCER and BPCER are computed independently per class, this composition does not bias either metric directly; we note, however, that BPCER is estimated from a smaller bona fide set and is therefore best interpreted with correspondingly wider uncertainty. A few PAI categories are also represented by relatively few samples (e.g., Cloth Mask, 17; LQ 3D Mask, 36), and the single System-category submission was scored on per-PAI subsets of 2–9 samples, so per-category rates in those cases are sensitive to individual decisions. We report exact per-class counts in Table 2 for full transparency, and

we view expanding sensor diversity and per-category sample sizes as natural directions for future editions.

Subjects and demographics. The dataset was collected from 15 subjects spanning multiple races; a detailed demographic breakdown of the broader collection is provided in the MORSE-Face paper [6].

Ethics, consent, and data handling. Data collection for LivDet-Face 2026 was conducted under an approved Institutional Review Board (IRB) protocol. Written informed consent was obtained from all participants, including for the mask-fabrication process, which involves capturing detailed facial photographs. All collected and sequestered data was stored on a secure Clarkson University server, with access restricted to the competition organizers.

Data availability. The evaluation dataset is made available on request through <https://livedet.org> for non-competition research use, while the sequestered evaluation protocol is maintained for future LivDet-Face editions.

5. LivDet-Face 2026 Competition Algorithms

A total of eleven teams from multiple countries participated in the LivDet-Face 2026 competition, submitting nineteen algorithms for the image category, eleven algorithms for the video category, and one submission for the system category. All competitors were given the option to participate anonymously, and non-anonymous teams were included as co-authors. The following subsections summarize the submissions for which the participating teams provided a method description; the remaining submissions are reported by their results only.

5.1. Mitek

Mitek’s Presentation Attack Detection (PAD) solutions are passive, requiring only a single facial image for inference, and are available as both standalone Dockerized services and SDK integrations. The underlying PAD framework is based on an ensemble architecture that combines multiple liveness pipelines to improve robustness against a wide range of presentation attacks. The processing pipeline begins with a proprietary face detector, which localizes the facial region and estimates image-quality metrics. The extracted facial image is then analyzed by multiple deep neural network-based discriminator models. The outputs of these models are then combined using a linear score-fusion mechanism to generate the final liveness score. Training is performed on proprietary Mitek datasets collected across diverse capture devices and representative demographic groups. The datasets include more than 20 presentation attack instruments (PAIs), encompassing print attacks, replay/display attacks, face-mask attacks, and other spoofing scenarios, thereby supporting robust generalization across diverse operational environments. Mitek PAD-R1 and PAD-R2 are two production-ready variants which

have been optimized for commercial deployment. In addition, PAD-P3 serves as an experimental research build.

5.2. SCUT_BIPLab_LivDetFace

Our method is based on MoGA-ETA [21], a language-guided face anti-spoofing framework for cross-domain generalization. In-context text prompt enhancement builds fine-grained spoof descriptions covering global anomalies (illumination artifacts, recapture) and local cues (moiré patterns, screen boundaries, 3D mask artifacts). Mixture-of-gradient adapters are introduced into the visual encoder during post-training for task-specific adaptation, and hard positive/negative mining strengthens cross-modal alignment. Final predictions combine adapted visual representations with enhanced textual semantics.

5.3. InvestigAI

InvestigAI uses an ensemble of CNNs, each trained to detect specific attack types on overlapping but non-identical data partitions with varied architectures, encouraging complementary feature learning from RGB images. Face alignment/cropping optimizes the region of interest and data augmentation reduces overfitting. Individual predictions are combined via minimum-based score fusion into a score in [0,1]. For video, frames are sampled and independently scored, and per-frame scores are aggregated by the median before thresholding.

5.4. HDA-Team

5.4.1 HDA-Team (Alg1)

HDA-Team (Alg1) proposes a face PAD method based on the DINOv2 foundation model. Freezing the backbone and adding a lightweight three-layer feed-forward head enables efficient training with minimal overfitting. Discriminative features from diverse attack types are fused into a unified embedding space, and a compact classifier distinguishes bona fide presentations (BPs) from attack presentations (APs), generalizing across device, material, and lighting variations.

5.4.2 HDA-Team (Alg2)

HDA-Team (Alg2) [9] presents a zero-shot face PAD algorithm using a single pretrained vision transformer. Without PAD-specific training, the model extracts semantic and fine-grained visual features, with lightweight adaptation modules inserted for parameter-efficient tuning while preserving the frozen backbone. This supports cross-domain generalization under unseen attack types and heterogeneous acquisition conditions with low adaptation complexity.

5.4.3 HDA-Team (Alg3)

HDA-Team (Alg3) combines multiple foundation models via parameter-efficient fine-tuning. Parallel vision transformers capture complementary semantic, geometric, and texture-based representations, with Low-Rank Adaptation (LoRA) modules applied to attention and feed-forward layers while backbones remain frozen. Features are fused into a unified embedding space and classified by a lightweight head.

5.5. CosmicFace

CosmicFace is an ensemble PAD pipeline for fully offline execution. The largest face is localized using a detector with fallback strategies, then processed by multiple liveness classifiers with complementary preprocessing and test-time augmentation. Handcrafted quality cues (texture strength, color statistics, sharpness, spectral characteristics, and replay/print artifacts) are fused with learned scores into a final 0–100 score; for video, frame-level scores are aggregated into a sequence-level prediction. It targets print, replay, and mask attacks, packaged as a self-contained Dockerized offline solution.

5.6. Anonymous0.0.2

This passive single-image system uses two complementary CNNs on different spatial contexts: one on a tight face crop for local texture and fine-grained artifacts, and one on an extended crop for contextual evidence such as medium borders, reflections, and background inconsistencies. Both are trained over 50 epochs with SGD and softmax on proprietary data covering printed, cutout, replay, 3D mask, and fabric mask attacks, with checkpoints selected by minimum APCER at fixed BPCER. At inference the two outputs are fused into a single liveness score.

5.7. APHI

5.7.1 APHI(Alg1)

APHI is a passive PAD algorithm handling both video and single-image inputs, treating images as one-frame videos. It uses a security-oriented ensemble of three deep models, each trained on a distinct subset of PAIs for attack-specific feature learning: convolutional backbones capture local evidence (texture, printing artifacts, replay patterns) while a Vision Transformer models global structure and longer-range inconsistencies. A frame is accepted as bona fide only when all models agree, and per-frame scores are aggregated into a video-level prediction.

5.7.2 APHI(Alg2) and APHI(Alg3)

Alg2 and Alg3 are variants of the Alg1 ensemble described above, evaluated with a stricter unanimity threshold across

the constituent models; this stricter setting drives their near-zero average APCER (0% and 0%, image and video tracks) at the cost of substantially higher BPCER (43.1%/80.8% for Alg2 and 62.1%/80.8% for Alg3, image/video).

5.8. Incode

Incode’s image and video entries are production PAD systems. The image solution fuses a single-frame appearance model with a monocular depth-estimation model, combining texture evidence with geometric cues that separate genuine 3D faces from flat or masked presentations. The video solution samples frames via a production-grade face detector with quality-aware selection favoring sharpness and frontal pose, fused with a temporal branch built on an internal foundation-model backbone specialized for physical spoof detection. Both branches expose tunable severity levels swept to balance APCER and BPCER, and are delivered as self-contained Docker images.

5.9. Remaining Submissions

The remaining competitors, including the anonymous submissions, elected not to share a method description, as permitted under the competition rules. Their results are nonetheless reported in full in Tables 3 and 4.

6. Results and Discussion

This section discusses the performance of the competing algorithms and systems, evaluated using APCER for each PAI category and BPCER for bona fide samples at a threshold of 50, as specified in the competition protocol. A summary of error rates for the image, video, and system categories is provided in Table 3 and Table 4, and Table 5, respectively, while performance comparisons for the algorithm track are illustrated in Figure 2a,b.

LivDet-Face 2026 Image Category: Mitek(pad-1) emerged as the winner of the image category with the lowest ACER of 0.56%, followed by Mitek (pad-2) with 0.78% and Mitek(p-3) with 0.84%. The winning algorithm achieved a BPCER of 0% and an average APCER of 1.12%, demonstrating excellent discrimination between bona fide and attack presentations. The top-performing algorithms consistently achieved very low APCER values across nearly all attack categories, including print attacks, display attacks, and three-dimensional mask attacks.

The top-ranked image algorithms substantially improve on prior editions: the winning ACER fell from 4.93% in LivDet-Face 2024 to 0.56% in 2026, reflecting progress in modern deep-learning PAD architectures and training strategies.

Some algorithms (e.g., SCUT_BIPLab_LivDetFace and HDA-team(Alg3)) achieved very low average APCER but high BPCER (73.6% and 40%), effectively rejecting at-

tacks while frequently misclassifying bona fide presentations. High APCER concentrated in advanced categories like skin-tone masks, makeup skin-tone masks, and realistic 3D masks, with CosmicFace elevated across many categories, indicating that balancing attack detection against reliable bona fide classification remains a key challenge.

LivDet-Face 2026 Video Category: Incode achieved the best overall performance in the video category with an ACER of 4.74%, followed by InvestigAI with 8.94% and APhi(Alg1) with 10.17%. The winning algorithm achieved an average APCER of 0.98% while maintaining a relatively low BPCER of 8.5%, demonstrating strong performance across both attack and bona fide presentations. This is a modest regression relative to LivDet-Face 2024, where the winning video algorithm achieved an ACER of 4.13% (APCER = 6.11%, BPCER = 2.16%) [12]; unlike the image track, the video track did not show clear improvement over the previous edition, which we attribute to the more challenging and previously unseen 3D mask categories introduced in 2026. Compared to previous LivDet-Face editions, the results otherwise indicate continued progress in video-based face PAD systems, particularly against conventional print and display attacks.

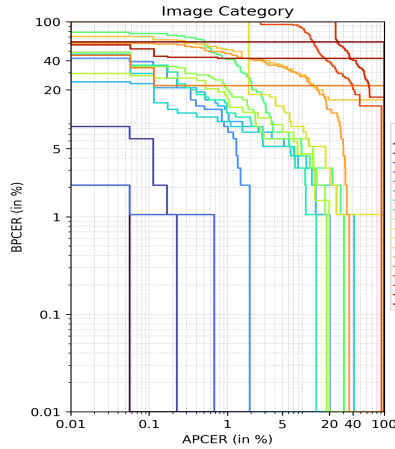
Several algorithms demonstrated very low APCER values but experienced substantial degradation in bona fide performance. For example, APhi (Alg2) and APhi (Alg3) achieved an average APCER of 0% but recorded BPCER values of 80.8%, indicating a tendency to reject a large proportion of genuine presentations. Similarly, HDA-team (Alg3) achieved an average APCER of only 0.32% but exhibited a BPCER of 55.3%.

The most challenging categories for video PAD were realistic 3D masks (high- and low-quality) and skin-tone masks with hair; HDA-team (Alg2) and arogyapandit_research showed high APCER across several of these. Thus, while video systems reliably detect conventional print and display attacks, realistic mask attacks remain a significant open challenge.

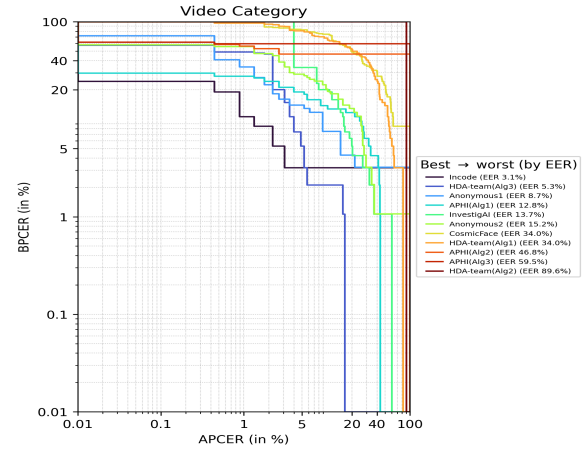
LivDet-Face 2026 System Category: As in LivDet-Face 2024, the system category attracted a single submission. Team Anonymous achieved an ACER of 13.64%, with an average APCER of 20.6% and a BPCER of 6.67%, as shown in Table 5, a regression relative to the sole 2024 system submission (Team Dermalog), which achieved an ACER of 3.12% (APCER = 6.25%, BPCER = 0%) [12]. We note, as in Section 4.3, that both editions’ system results are each based on a single submission evaluated on small per-PAI sample counts, so neither result should be read as representative of system-based PAD broadly. The system demonstrated excellent performance against several attack categories, including Paper Mask, Normal Paper, Cloth Mask, Silicone 3D Mask, and White 3D Mask attacks, achieving 0% APCER for these presentation attack instruments.

Table 3. Face PAD Competition Summary: Image Category. All values are reported as percentages (%)

Algorithms	Presentation Attack Instruments Level Types (APCER %)										Overall Performance		
	Level A				Level B	Level C					APCER _{avg} %	BPCER %	ACER %
	Print Attack				Display Attacks	3D Face Masks							
	CM	NP	PM	PP	LP	HQ	LQ	MM	SM-hair	SM			
Mitek(pad-1)	0	0	0	0	0	5.7	5.5	0	0	0	1.12	0	0.56
Mitek(pad-2)	0	0	0	0	0	5.7	0	0	0	0	0.57	1	0.78
Mitek(pad-3)	0	0	0	0	0	14	2.7	0	0	0	1.69	0	0.84
Anonymous0.0.2	0	1.3	0	1.8	0	20	11	0	0	0	3.42	8.4	5.91
Anonymous0.0.3	29.4	4	5.4	0.5	0	39	2.7	0	0	0	8.05	7.3	7.67
APHI(Alg1)	11.7	0	0	0.3	0	21	25	0	0	0	5.83	10.5	8.16
Anonymous3	0	0	0	0	0	8.6	0	0	0	0	0.86	16.84	8.85
Anonymous0.0.1	30.7	10.7	2.7	11	0	64	50	0	0	0	16.91	4.4	10.65
HDA-team(Alg2)	5.8	86.6	2.7	3.1	96.9	10	8.3	0	0	0	21.34	1	11.17
InvestigAI	17.6	6	16.2	0	0	46	61	0.3	53.1	22.3	22.23	5.2	13.71
Incode	0	0	0	0	0	1.4	2.7	0	0	0	0.41	29.4	14.90
Anonymous1	0	20.4	5.4	34	11.8	13	17	0.3	0	0	10.09	20.2	15.14
HDA-team(Alg3)	0	0	0	0	0	1.4	0	0	0	0	0.14	40	20.07
APHI(Alg2)	0	0	0	0	0	2.8	5.5	0	0	0	0.83	43.1	21.96
Anonymous2	0	0	0	0	0	29	11	0	0	0	4.03	41.4	22.71
APHI(Alg3)	0	0	0	0	0	0	0	0	0	0	0	62.1	31.05
HDA-team(Alg1)	11.7	0	5.4	0	0	37	50	59.6	0	0	16.38	55.7	36.04
SCUT_BIPLab_LivDetFace	0	0	0	0	0	4.2	0	0	0	0	0.42	73.6	37.01
CosmicFace	29.4	30.1	70.2	48	22.1	61	78	86.7	89.6	61.1	57.60	32.6	45.10



(a) Image Category



(b) Video Category

Figure 2. DET curves for 18 of the 19 image algorithms (Anonymous2 excluded, as only attack scores were available for it) and all 11 video algorithms in the competition, plotted on logarithmic APCER/BPCER axes per ISO/IEC 30107-3 convention, pooling samples across all ten PAI types.

Table 4. Face PAD Competition Summary: Video Category. All values are reported as percentages (%)

Competitor Name	Presentation Attack Instruments Level Types (APCER %)										Overall Performance		
	Level A				Level B	Level C					APCER _{avg} %	BPCER%	ACER%
	Print Attack				Display Attacks	3D Face Masks							
	CM	NP	PM	PP	LP	HQ	LQ	MM	SM-hair	SM			
Incode	0	0	0	0	0	4.3	5.5	0	0	0	0.98	8.5	4.74
InvestigAI	5.8	6.2	8.1	0	0	23.1	41.6	0	38.7	13.3	13.68	4.2	8.94
APHI(Alg1)	5.8	0	0	0	0	31.8	38.8	0	0	0	7.64	12.7	10.17
Anonymous2	0	0	0	0	0	45.5	48.5	0	6.6	0	10.06	10.7	10.38
Anonymous1	0	0	0	0	0	1.4	0	0	0	0	0.14	41.9	21.02
HDA-team(Alg3)	0	0	0	0	0	0	0	0	3.2	0	0.32	55.3	27.81
HDA-team(Alg1)	5.8	0	8.1	0	0	30.4	52.7	40	0	0	13.70	50.0	31.85
CosmicFace	52.9	0	59.4	20	26.6	63.7	72.2	53.3	6.4	10	36.45	30.8	33.62
APHI(Alg2)	0	0	0	0	0	0	0	0	0	0	0	80.8	40.40
APHI(Alg3)	0	0	0	0	0	0	0	0	0	0	0	80.8	40.40
HDA-team(Alg2)	100	100	100	100	80	100	100	100	64.5	70	91.45	0	45.72

However, the system struggled with Makeup Skin-tone Masks, High-Quality 3D Masks, and Low-Quality 3D Masks (APCER 66.6%, 42.8%, and 55.5%). Realistic

3D and appearance-based attacks – especially makeup-enhanced skin-tone masks that closely resemble bona fide faces – thus remain the hardest instruments for integrated

Table 5. Face PAD Competition Summary: System Category

Attack Type	Samples	Errors	APCER (%)
Paper Mask	24	0	0
Normal Paper	3	0	0
Cloth Mask	4	0	0
Makeup Skin-tone Mask	3	2	66.6
Silicone 3D Mask	2	0	0
HQ 3D Mask	7	3	42.8
LQ 3D Mask	9	5	55.5
White 3D Mask	5	0	0
Average APCER (%)			20.6
BPCER (%)			6.67
ACER (%)			13.64

PAD systems and the main target for improvement.

7. Conclusion

The LivDet-Face 2026 competition introduced a novel makeup skin-tone 3D mask attack and evaluated state-of-the-art face PAD algorithms across image, video, and system categories. The winning image algorithm achieved an ACER of 0.56% (APCER = 1.12%, BPCER = 0%), while the winning video algorithm achieved an ACER of 4.74% (APCER = 0.98%, BPCER = 8.5%). The sole system submission achieved an ACER of 13.64%. A total of nineteen image algorithms, eleven video algorithms, and one integrated PAD system were evaluated using a diverse dataset containing bona fide presentations and ten presentation attack instrument categories collected under realistic acquisition conditions. Compared with LivDet-Face 2024, the results demonstrate substantial progress in image-based face PAD performance, with the best image algorithm reducing ACER from 4.93% to 0.56%. Key observations indicate that most algorithms successfully detected conventional print and display attacks, achieving low APCER values. However, highly realistic attack types, including high-quality 3D masks, skin-tone masks, and makeup-enhanced masks, remained challenging for several competitors. In addition, a number of algorithms achieved excellent attack detection performance but exhibited high BPCER values, highlighting ongoing difficulties in correctly classifying bona fide presentations. Several factors contributed to performance degradation, including the presence of unseen attack types, limited training data for emerging PAIs, variations in acquisition conditions and sensor characteristics, and failures during intermediate processing stages such as face detection and quality assessment. These findings demonstrate that face PAD remains an active and challenging research area where robustness and generalization continue to be critical concerns. The LivDet-Face 2026 competition and benchmark dataset provides a valuable platform for measuring progress in face presentation attack detection and will support the development of more secure, reliable, and generalizable biometric systems.

8. Acknowledgment

The organizers thank all teams that participated in LivDet-Face 2026. This material is based upon work supported in part by the National Science Foundation under Grant No. 1650503, the Center for Identification Technology and Research (CITeR).

References

- [1] Y. Atoum, Y. Liu, A. Jourabloo, and X. Liu. Face anti-spoofing using patch and depth-based cnns. In *2017 IEEE international joint conference on biometrics (IJCB)*, pages 319–328. IEEE, 2017. 2
- [2] J. Bigun, H. Fronthaler, and K. Kollreider. Assuring liveness in biometric identity authentication by real-time face tracking. In *Proceedings of the 2004 IEEE International Conference on Computational Intelligence for Homeland Security and Personal Safety, 2004. CIHSPS 2004.*, pages 104–111. IEEE, 2004. 1
- [3] Z. Boulkenafet, J. Komulainen, and A. Hadid. Face anti-spoofing based on color texture analysis. In *2015 IEEE international conference on image processing (ICIP)*, pages 2636–2640. IEEE, 2015. 2
- [4] M. M. Chakka, A. Anjos, S. Marcel, R. Tronci, D. Muntoni, G. Fadda, M. Pili, N. Sirena, G. Murgia, M. Ristori, et al. Competition on counter measures to 2-d facial spoofing attacks. In *2011 international joint conference on biometrics (IJCB)*, pages 1–6. IEEE, 2011. 2, 3, 4
- [5] I. Chingovska, J. Yang, Z. Lei, D. Yi, S. Z. Li, O. Kahm, C. Glaser, N. Damer, A. Kuijper, A. Nouak, et al. The 2nd competition on counter measures to 2d face spoofing attacks. In *2013 international conference on biometrics (ICB)*, pages 1–6. IEEE, 2013. 2, 3, 4
- [6] M. Z. U. Chowdhury and L. Igene. Morse-face: A multi-distance, occlusion-rich face dataset for recognition and presentation attack detection. In *Proceedings of the IEEE International Joint Conference on Biometrics (IJCB)*, 2026. To appear. 4, 5
- [7] T. de Freitas Pereira, A. Anjos, J. M. De Martino, and S. Marcel. Can face anti-spoofing countermeasures work in a real world scenario? In *2013 international conference on biometrics (ICB)*, pages 1–8. IEEE, 2013. 1
- [8] J. Galbally, S. Marcel, and J. Fierrez. Biometric antispoofing methods: A survey in face recognition. *Ieee Access*, 2:1530–1552, 2014. 2
- [9] J. Gonzalez-Soler, J. Tapia, and C. Busch. Are foundation models all you need for zero-shot face presentation attack detection? In *IEEE Intl. Conf. on Automatic Face and Gesture Recognition (FG)*, pages 1–10. IEEE, 2025. 6
- [10] A. Husseis, J. Liu-Jimenez, I. Goicoechea-Telleria, and R. Sanchez-Reillo. A survey in presentation attack and presentation attack detection. In *2019 International Carnahan Conference on Security Technology (ICCST)*, pages 1–13. IEEE, 2019. 1
- [11] L. Igene, A. Hossain, R. Plesh, S. Purnapatra, S. Singh, M. Imtiaz, and S. Schuckers. Development of novel 3d face masks for assessment of face recognition systems. In *2024*

International Conference of the Biometrics Special Interest Group (BIOSIG), pages 1–7. IEEE, 2024. [3](#)

- [12] L. Igene, A. Hossain, M. Z. Uddin Chowdhury, H. Rezaie, A. Rollins, J. Dykes, R. Vijaykumar, A. Komaty, S. Marcel, S. Schuckers, J. E. Tapia, C. Aravena, D. Schulz, B. Adami, N. Karimian, D. Nunes, J. Marcos, N. Gonçalves, L. Sikošek, B. Batagelj, A. Alenin, A. Alkhaddour, A. Pimenov, A. Tregubov, I. Avdonin, M. Kazantsev, M. Pozigun, V. Pryadchenko, N. Schei, D. Pabon, and M. Tiedemann. Face liveness detection competition (livdet-face) - 2024. In *2024 IEEE International Joint Conference on Biometrics (IJCB)*, pages 1–9, 2024. [2](#), [3](#), [4](#), [7](#)
- [13] LivDet-Face 2026 . competition’s instructions. <https://face2026.livdet.org/>. [3](#)
- [14] S. Purnapatra, N. Smalt, K. Bahmani, P. Das, D. Yambay, A. Mohammadi, A. George, T. Bourlai, S. Marcel, S. Schuckers, et al. Face liveness detection competition (livdet-face)-2021. In *2021 IEEE International Joint Conference on Biometrics (IJCB)*, pages 1–10. IEEE, 2021. [3](#)
- [15] R. Ramachandra and C. Busch. Presentation attack detection methods for face recognition systems: A comprehensive survey. *ACM Computing Surveys (CSUR)*, 50(1):1–37, 2017. [1](#)
- [16] S. Schuckers, G. Cannon, E. Tabassi, M. Karlsson, and E. Newton. Fido biometrics requirements. *Population*, 5(2-1):2–3, 2019. [2](#)
- [17] S. A. Schuckers. Spoofing and anti-spoofing measures. *Information Security technical report*, 7(4):56–62, 2002. [2](#)
- [18] I. G. Tellería. Evaluation of presentation attack detection under the context of common criteria. *Universidad Carlos III de Madrid*, 2019. [2](#)
- [19] 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 *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 6380–6391, 2023. [3](#)
- [20] D. Wen, H. Han, and A. K. Jain. Face spoof detection with image distortion analysis. *IEEE Transactions on Information Forensics and Security*, 10(4):746–761, 2015. [1](#), [2](#)
- [21] F. Zhang and Others. Moga-eta: A language-guided face anti-spoofing framework for cross-domain generalization. *Journal Name*, 2026. To be updated with correct details. [6](#)
- [22] S. Zhang, X. Wang, A. Liu, C. Zhao, J. Wan, S. Escalera, H. Shi, Z. Wang, and S. Z. Li. A dataset and benchmark for large-scale multi-modal face anti-spoofing. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2019. [3](#)
- [23] Y. Zhang, Z. Yin, J. Shao, Z. Liu, S. Yang, Y. Xiong, W. Xia, Y. Xu, M. Luo, J. Liu, et al. Celeba-spoof challenge 2020 on face anti-spoofing: Methods and results. *arXiv preprint arXiv:2102.12642*, 2021. [2](#), [3](#), [4](#)