To whom it may concern

The Swiss Center for Biometrics Research and Testing (SCBRT) is part of the Biometrics Security and Privacy group at the Foundation of the Idiap Research Institute[1]. The SCBRT is also a FIDO Alliance Accredited Biometrics Laboratory (ABL).[2] Under contract with Mobai AS[3] (referred to as *the vendor* hereafter), the SCBRT has evaluated the efficacy of the vendor's face biometrics solution. Our evaluation complies with ISO/IEC 30107 recommendations. Additionally, in accordance with the ISO/IEC 19989 standard, in this evaluation we have also verified a set of performance-claims made by the vendor. This letter describes the evaluation protocol, and summarizes the main outcomes.
Target Of Evaluation (TOE): Client version 2.1.1 and Server version 2.4.0.
Test Harness:

- two smartphones: iPhone 14, running iOS 16, and Samsung S21, running Android 11;
- a laptop computer, running Ubuntu version 24.04; and
- a wireless access-point.

The client app, running on each phone, captures face biometric samples. It transmits each sample to the server (running on the laptop, for the purposes of this evaluation) via a wireless connection. The server processes the biometric sample and transmits the decision (*accept* or *reject*) back to the client, which displays the result on the phone-screen.Test subjects and the testers interact only with the client app on each phone.

Scope of Evaluation

Two aspects of the TOE have been studied in this evaluation:

- its face verification performance, and
- its presentation attack detection (PAD) capability (for Level A and Level B attacks only).

Face verification performance of the TOE has been evaluated using a test crew of 20 subjects (10 men and 10 women). First, all the subjects enrolled themselves on each of the two phones. Then, each subject made at least 10 probes on each phone. Each probe sample was compared with all reference samples of all subjects (two reference templates per subject). From these comparisons, we compute the false match rates (FMR) and false non-match rates (FNMR) of the TOE.
The Presentation Attack Instruments (PAI) have been created in accordance with ISO standards used to assess and triage the attack potential of presentation attacks (PA).

- Level A PAIs are attack instruments that can be created within a day, requiring neither sophisticated equipment nor a high degree of expertise.
- Level B PAIs are attack instruments that may take one to three days to create, requiring some degree of technical expertise.

By prior agreement with the vendor, we have used PAIs of six Level A species and six Level B species. The PAI species used in this evaluation are described in Table 1. For each species, 10 PAIs have been created – five PAIs have been used to perform PAs on the iPhone 14, and the remaining five PAIs have been used to attack the TOE on the Samsung S21 device. The testers have performed 10 PAs using each PAI. Thus, in total, 100 PAs have been performed for each PAI species included in this evaluation. The PAD scores generated for the *bona fide* presentations and PAs are used to estimate the Bona-Fide Presentation Classification Error Rate (BPCER), and the Attack Presentation Classification Error Rate (APCER) of the TOE. The face-verification scores and PAD scores for the PAs are also used to determine the Impostor Attack Presentation Accept Rate (IAPAR), which quantifies the vulnerability of the TOE to PAs.

[1]Website: `www.idiap.ch`
[2]The evaluation discussed in this letter is not related to FIDO certification.
[3]Company website: `www.mobai.bio`

The performance metrics estimated for the TOE in this evaluation are listed in Table 2. The metrics have been computed in accordance with their definitions in the ISO/IEC 30107 standards. In particular, the metrics have been estimated on *transaction basis*, where a transaction consists of multiple *attempts* performed by the user within a predefined time-period. For this evaluation, the TOE implements a transaction has having up to five attempts, that must be completed within 60 seconds. A transaction is terminated when (a) a given attempt is accepted by the TOE, (b) the maximum number of attempts has been reached, or (c) the time-out limit has been reached. If a transaction ends with no recorded attempt within the time-out period, then we say that a *failure-to-acquire* (FTA) event has occured. The total number of transactions and attempts for both *bona fide* and attack presentations are shown in Table 3.

Species	Description	Count
Level A		**(60)**
A1	Photo printed on glossy paper using inkjet printer	10
A2	Photo printed on glossy paper using laser printer	10
A3	Photo printed on matte paper using inkjet printer	10
A4	Photo printed on matte paper using laser printer	10
A5	Photo displayed on smartphone (replay attack)	10
A6	Still photo displayed on iPad Pro (replay attack)	10
Level B		**(60)**
B1	Enhanced photo printed on glossy paper using inkjet printer	10
B2	Enhanced photo printed on glossy paper using laser printer	10
B3	Enhanced photo printed on matte paper using inkjet printer	10
B4	Enhanced photo printed on matte paper using laser printer	10
B5	Video replayed on iPad Pro (video replay attack)	10
B6	Paper cut-out masks	10

Table 1: The PAI species used in this evaluation. Ten PAIs have been created for each PAI species. Numbers in parentheses indicate the total number of PAIs used for the corresponding Level.

Term	Definition
APAR	*Attack Presentation Acquisition Rate*: Proportion of attack presentations using the same PAI species from which the data capture subsystem acquires a biometric sample of sufficient quality.
APCER	*Attack Presentation Classification Error Rate*: the proportion of attack presentations that have been misclassified as *bona fide* by the by the presentation-attack detection (PAD) subsystem of the TOE.
APNRR	*Attack Presentation Non-Response Rate*: the proportion of attack presentations using the same PAI species that cause no response at the PAD subsystem or data capture subsystem.
BPCER	*Bona-fide Presentation Classification Error Rate*: the proportion of *bona fide* presentations that have been misclassified as presentation attacks (PA) by the PAD subsystem of the TOE.
BPNRR	*Bona-fide Presentation Non-Response Rate*: the proportion of bona-fide presentations that cause no response at the PAD subsystem or data capture subsystem.
FMR	*False Match Rate*: the proportion of zero-effort-impostor (ZEI) transactions that are falsely declared to match a template of a subject.
FNMR	*False Non-Match Rate*: the proportion of genuine transactions falsely reported as non-match for a template of the subject in question.
IAPAR	*Impostor Attack Presentation Accept Rate*: the proportion of impostor attack presentations using the same presentation attack instrument (PAI) species that result in *accept* decision by the TOE.
RIAPAR	*Relative Imposter Presentation Accept Rate*: the sum of IAPAR and FNMR at a fixed decision threshold.

Table 2: Definitions of metrics estimated used in this evaluation. These definitions conform to the ISO/IEC 30107 standards.

Presentation Category	Number of Transactions	Number of Attempts
Bona fide	407	419
PA	1200	4583

Table 3: Statistics of the transactions and attempts performed in this evaluation.

Evaluation Results

1. **Face Verification Performance:**
   Based on a total of 814 *genuine* comparisons and 15,466 zero-effort impostor (ZEI) comparisons, the TOE performed perfectly, achieving a FMR of 0% as well as FNMR of 0%. This result is based on both *same device* comparisons (where the reference and probe samples are captured using the same device) and *cross-device* comparisons (where the reference and probe samples are recorded using different devices).
   As no FTA event occured for *bona fide* presentations in this evaluation, the BPNRR of the TOE is 0%.
2. **Vulnerability to Presentation Attacks:**
   For PAIs of Level A and Level B, no PA transaction was successful in attacking the TOE. All transactions using PAIs of these two Levels were correctly classified as PAs. Therefore, the IAPAR of the TOE is 0% for both Level A and Level B. As required by ISO/IEC 30107:2023, we report here the RIAPAR, also as 0% for Level A, as well as for Level B.
3. **Presentation Attack Detection:**
   In this evaluation the TOE exhibited perfect APCER (0%) for all PAI species of Level A and Level B (see Table 4). The BPCER of the TOE was estimated at 0.246% in this evaluation, corresponding to one *bona fide* transaction misclassified as a PA.
   Other metrics related to FTA events for PAs in this study are also shown in Table 4. APAR (%) is computed as $[100 - APNRR]$. According to the ISO/IEC 30107 standard, to determine the APCER of the TOE for a given Level, we take the APCER for the PAI species for which the TOE shows the worst performance among all the species of that Level. However, this rule does not extend to the computation of APAR. Therefore, here we have computed the APNRR and APAR of each Level as simple percentages, considering all PAs of the Level in question. In this evaluation, the APNRR for Level A PAs came to 3.33%, and the corresponding APAR was 96.67%. For Level B PAs, the APNRR was 5%, and therefore the APAR was 95%.

ISO/IEC 19989-3 Conformance

The vendor has mandated us (the SCBRT) to perform this evaluation in conformance with the ISO/IEC 19989-3 standard for biometrics PAD evaluation. To this end, the vendor provided us with a list of performance claims for the three PAD metrics: APCER, BPCER and IAPAR. The claims are of the form: 'For PAI species x, the metric y shall not exceed the value z.' These claims are indicated in Table 5. For example, the vendor claims that, 'for the PAI species A1, the APCER shall not exceed 10%, the BPCER shall not exceed 15% and the IAPAR shall not exceed 10%.' The vendor provided us the list of claims before the commencement of this evaluation. The ISO/IEC 19989-3 standard requires that we verify whether the TOE conforms to these claims. *As indicated in Table 5, all the claims have been validated in this evaluation.*

PAI Species	FTA	APNRR (%)	APAR (%)	APCER (%)
Level A		**3.33**	**96.67**	**0**
A1	0 of 100	0	100	0
A2	0 of 100	0	100	0
A3	0 of 100	0	100	0
A4	0 of 100	0	100	0
A5	10 of 100	10	90	0
A6	10 of 100	10	90	0
Level B		**5.0**	**95.0**	**0**
B1	10 of 100	10	90	0
B2	0 of 100	0	100	0
B3	10 of 100	10	90	0
B4	10 of 100	10	90	0
B5	0 of 100	0	100	0
B6	0 of 100	0	100	0

Table 4: PAD performance metrics for the TOE. APCER values are indicated for each PAI species, as well as the aggregate value for each Level. The aggregate APCER of a given Level is determined by selecting the APCER of the PAI species for which the TOE shows the worst performance in that Level. The metrics APNRR and APAR summarize the FTA events seen for PAs. For each Level, we express these two metrics as simple percentages, considering all presentations using all PAIs of the Level. For every PAI species, 10 PA transactions have been performed for each of 10 PAIs, resulting in a total of 100 PA transactions per PAI species.

PAI Species	APCER		BPCER		IAPAR	
	Claimed (Max. (%))	Claim Validated	Claimed (Max. (%))	Claim Validated	Claimed (Max. (%))	Claim Validated
Level A	10	✓	15	✓	10	✓
Level B	10	✓	15	✓	10	✓

Table 5: Validation of claimed PAD performance, in fulfilment of ISO/IEC 19989-3 requirements. Before the start of the evaluation, the vendor provided us with 'Claimed Performance Limits' for three PAD metrics – APCER, BPCER, and IAPAR. Complying with the ISO/IEC 19989-3 standard for biometrics PAD evaluation, we formally verify whether the PAD performance of the TOE meets these claims. For each of the three metrics, the value obtained in this evaluation for each PAI species did not exceed the upper limit claimed by the vendor.

Sincerely yours,

Institut de Recherche Idiap
Centre du Parc
Rue Marconi 19, 1920 Martigny
Tél. 027 721 77 11 - Fax 027 721 77 12
info@idiap.ch - www.idiap.ch

Prof. Sébastien Marcel
Director, SCBRT

Prof. Andrea Cavallaro
Director, IDIAP