iBeta
QUALITY ASSURANCE

Innovatrics
DOT Passive Liveness Application - V2.1.0(1) iOS/ V3.0.0 Android
PAD Level 1 Test Report

Prepared for

Innovatrics
Jarosova 1 832 03 Bratislava
Slovak Republic

16 Sept 2020
Report #200908-iBetaCTR-v2.0

Trace to Standards
ISO 30107-3
Sections 1.3, 6.0, 7.1, 7.2, 8.1, 10.2, 11.3, 11.5, 13.1, 13.2, 13.3 and 13.4
NIST Handbook 150-25
Sections 4.1.5, 5.10.1 through 5.10.4

Test Results in this report apply to the biometrics system configuration tested. Testing of biometric systems that have been modified may or may not produce the same test results. This report shall not be reproduced, except in full. iBeta Quality Assurance is accredited for Biometric System Testing:

NVLAP®
TESTING
NVLAP LAB CODE 200962-0

2675 S. Abilene Street, Suite 300, Aurora, Colorado, 80014

Version History				
Ver #	**Description of Change**	**Author**	**Approved by**	**Date**
V1.0	Initial release of Innovatrics report	*Ricky Brown Jr*	*Karen Wilson-Winfrey*	08-September-2020
V2.0	Final release based on Innovatrics review	*Ricky Brown Jr*	*Gail Audette*	16-September-2020

TABLE OF CONTENTS

1 Executive Summary

iBeta conducted testing in compliance with the requirements of ISO/IEC 30107-1 and ISO/IEC 30107-3 with the Innovatrics Digital Onboarding Toolkit (DOT) Passive Liveness Application Version 2.1.0(1) for iOS and V3.0.0 for Android facial recognition biometric system on the device from 21 August through 4 September 2020. The testing was conducted on two smartphones loaded with the application. The application uses passive liveness detection.

Conformance testing was performed in compliance with the requirements of Level 1 testing which was conducted in accordance with the contract for a level of spoofing technique that only utilized simple, readily available methods to create an artefact of the genuine biometric for use in the presentation attack. The test record included all test executions and reviews. All test executions and reviews included the record of requirements that were satisfactorily and unsatisfactorily completed, deficiencies noted, reports to Innovatrics, software and manufacturing resolutions, validations of resolutions and documentation of incorporation of resolutions into the biometric system. This test report bearing the NVLAP symbol must not be used by the client to claim product certification, approval, or endorsement by NVLAP, NIST, or any agency of the U.S. Government.

The application was tested on two smartphones (Google Pixel 2 with Android 8.1.0 and iPhone 8 with iOS 13.6). The liveness testing applied 1 bona fide subject presentation alternated with 3 presentations for each species, such that 150 Presentation Attacks (PAs) and 50 bona fide presentations were applied to the device per species on each device.

This corresponds to over 1800 presentation attacks over the entire test effort on the Google Pixel 2 and iPhone 8.

During testing on both the Google Pixel 2 and iPhone 8, iBeta was unable to gain a liveness classification with 150 presentation attacks (PAs) with each of the 6 species of attacks. With 900 transaction attempts on each device, the Presentation Attack (PA) success rate is 0%.

The overall combined Attack Presentation Classification Error Rate (APCER) equates to an overall PA success rate of 0%. The Bona Fide Presentation Classification Error Rate (BPCER) is the proportion of bona fide presentations incorrectly classified as presentation attacks during the testing that produced an overall BPCER of 0.0% on the Google Pixel 2 and 0.7% on the iPhone 8. The BPCER represents the percentage of genuine, live subjects whose liveness could not be determined. The summary of testing is provided below in Table 1.

Table 1 Summary of Test Results

	Test Species	Android v3.0.0			iOS v2.1.0(1)		
		PAs	APCE	APCER	PAs	APCE	APCER
1.	2D photo on matte paper with edges cut	150	0 of 150	0%	150	0 of 150	0%
2.	2D photo on matte paper presented on a curved surface	150	0 of 150	0%	150	0 of 150	0%
3.	2D photo (as above) with eyes cut out	150	0 of 150	0%	150	0 of 150	0%
4.	3D Layered paper photo	150	0 of 150	0%	150	0 of 150	0%
5.	Video displayed on laptop	150	0 of 150	0%	150	0 of 150	0%
6.	Video displayed on smartphone	150	0 of 150	0%	150	0 of 150	0%
Total per species			0 of 150	0%		0 of 150	0%
Total for all species			0 of 900	0%		0 of 900	0%

1.1 Background

iBeta is nationally accredited as a test lab by the National Voluntary Lab Accreditation Program (NVLAP Testing Lab Code 200962) to the requirements of ISO/IEC 17025:2017 (General requirements for the competence of testing and calibration laboratories). In 2011, iBeta was accredited by NIST under the National Voluntary Laboratory Accreditation Program (NVLAP) for Biometric Testing under NIST handbook 150-25 and has become an expert in the field of biometrics. In addition, iBeta procedures against the ISO 30107-3 Presentation Attached Detection (PAD) standard were audited by our accrediting body and iBeta's Scope of Accreditation was increased to include conformance testing to the ISO 30107-3 standard in April 2018.

The terms and definitions within this report are directly from the ISO 30107-3 standard.

1.2 Internal Documentation

The documents identified below are iBeta internal documents used in conformance testing.

Table 2 Internal Documents

Version #	Title	Abbreviation	Date	Author (Org.)
iBeta	Contractual Documents			
V01	Agreement for Presentation Attack Detection ISO 30107-3 Testing Services v01	SOW	31 March 2019	iBeta Quality Assurance
V01	Change Order		23 July 2020	iBeta Quality Assurance
	Mutual Confidential Disclosure Agreement	NDA		iBeta Quality Assurance
iBeta	PAD Procedures			
1.0	Biometric Deliverable Receipt Procedure		6/1/11	iBeta Quality Assurance
3.0	Biometric Security Procedure		5/20/13	iBeta Quality Assurance
1.0	Biometrics Configuration Management Procedure		6/9/11	iBeta Quality Assurance
1.0	PAD Certification Test Procedure		1/24/18	iBeta Quality Assurance
1.0	Biometric Training and Training Records Procedure		6/1/11	iBeta Quality Assurance
B	Biometric Certification Report Template		1/24/18	iBeta Quality Assurance
iBeta	Project Documents			
	PAD Level 1 Test Case - Innovatrics.xlsx		09/08/20	iBeta Quality Assurance

1.3 External Documentation

The documents identified below are external resources used in conformance testing.

Table 3 External Documents

Version #	Title	Abbreviation	Date	Author (Org.)
NIST Handbook 150 2006 Edition	NVLAP System Testing	NIST 150	February 2006	National Voluntary Lab Accreditation Program
NIST Handbook 150-25	NVLAP Biometric System Testing	NIST 150-25		National Voluntary Lab Accreditation Program
2010	International Standard: Conformity assessment – General requirements for proficiency testing	ISO/IEC 17043:2010	2010-02-01	ISO/IEC
2017-09	ISO/IEC 30107-3 Information technology — Biometric presentation attack detection — Part 3: Testing and reporting	ISO 30107-3	September 2017	ISO/IEC
2016-01-15	ISO/IEC 30107-1 Information technology — Biometric presentation attack detection — Part 1: Framework	ISO 30107-1	January 2015	ISO/IEC
2012-12-15	ISO/IEC 2382-37, Information technology — Vocabulary — Part 37: Biometrics		December 2012	ISO/IEC
2016	Presentations and attacks, and spoofs, oh my." Image and Vision Computing 55 (2016): 26-30	Schuckers(2016)	2/3/2016	Schuckers, Stephanie, Clarkson University

1.4 Technical Documents

The Technical Documents submitted for this conformance test effort are listed in Section 3 System Identification.

1.5 Test Report Contents

The contents of this Test Report include:

- Section 1 The Introduction identifies the scope of testing.
- Section 2 The Test Background identifies the process for testing.
- Section 3 The Biometrics System Identification identifies the system configuration including hardware, software and the technical documentation.
- Section 4 The Biometrics System Overview identifies the overall design and functionality of biometrics system.
- Section 5 The Review and Test Results are the methods and results of the testing effort.
- Section 6 The conformance statement of the biometrics system. Test Operations, Findings and Data Analysis are in the appendices.
- Appendix A: Test Results for PAD Level 1 (conforming to the applicable standard).

2 Conformance Test Background

The testing performed was completed per ISO-IEC 30107-3, which does not have specific pass/fail criteria or target APCER. Instead, the results of the testing presented in this report serve as meeting conformance standards that the system as described was tested to provide the reported results. As such, Innovatrics may use the terms compliance or conformance with the ISO 30107-3 standard when discussing or communicating the completion of this testing

As part of their application for Conformance testing Innovatrics submitted their implementation statement for the Innovatrics DOT Passive Liveness facial recognition application for PAD Level 1 testing.

The Systems under Test (SUT) are Facial Recognition biometric systems developed by Innovatrics. iBeta was also informed by ISO 30107-4 for mobile device based application testing. iBeta follows the Levels of Testing as defined below in Table 4 that closely relates to the Levels A, B, and C as defined as the Level of Effort of PAD Artefact Generation from Schuckers, Stephanie. "Presentations and attacks, and spoofs, oh my." Image and Vision Computing 55 (2016): 26-30.

Table 4 iBeta Levels of PAD Testing

Level	Time	Expertise	Artefact source	Limit
1	8 hours per subject	None - minimal	Cooperative subject and equipment is readily available in a normal home or office environment	0% penetration or match rate allowed
2	2-4 days per subject	Moderate – participated in at least 1 other PAD test with the target modality	Cooperative subject and equipment is more expensive (such as a 3D printer)	1% penetration or match rate allowed
3*	3 weeks per subject	Significant – has dedicated at least 16 hours to research of presentation attacks of the target modality and has participated in at least 2 other PAD tests with the target modality	Cooperative Subject and latent sources for subject data. Equipment is extensive e.g., special order contact lenses, facial masks, and 3D printed spoofs	5% penetration or match rate allowed
*Currently, iBeta does not offer testing to Level 3 as a service.				

2.1 Terms and Definitions

The Terms and Definitions identified below are used in this test report.

Table 5 Terms and Definitions

Term	Abbreviation	Definition
attack potential		Measure of the capability to attack an IUT (TOE) given the attacker’s knowledge, proficiency, resources and motivation
attack type		Element and characteristic of a presentation attack, including PAI species, concealer or impostor attack, degree of supervision, and method of interaction with the capture device
attack presentation classification error rate	APCER	Proportion of attack presentations using the same PAI species incorrectly classified as bona fide presentations in a specific scenario

Term	Abbreviation	Definition
attack presentation non-response rate	APNRR	Proportion of attack presentations using the same PAI species that cause no response at the PAD subsystem or data capture subsystem.
bona fide presentation		Interaction of the biometric capture subject and the biometric data capture subsystem in the fashion intended by the policy of the biometric system
bona fide presentation classification error rate	BPCER	Proportion of bona fide presentations incorrectly classified as presentation attacks in a specific scenario
Bona fide presentation non-response rate	BPNRR	Proportion of bona fide presentations that cause no response at the PAD subsystem or data capture subsystem.
Failure to acquire	FTA	The system fails to capture a sample from the subject. This is normally reported as a rate based on the number of subjects x attempts that the system attempted to acquire.
Failure to enroll	FTE	The system fails to enroll the subject. This is normally reported as a rate based on the number of subjects whom the system attempted to enroll.
Full-system evaluations		Full-system evaluations add a comparison subsystem to the IUT, generating a comparison score or candidate list. This situation is illustrated in ISO/IEC 30107-1:2016, Figure 3.
impostor attack presentation match rate	IAPMR	Proportion of impostor attack presentations using the same PAI species in which the target reference is matched
presentation attack	PA	Presentation to the biometric data capture subsystem with the goal of interfering with the operation of the biometric system
presentation attack detection	PAD	Automated determination of a presentation attack
presentation attack instrument	PAI	Object used in a presentation attack
PAI species		Class of presentation attack instruments created using a common production method and based on different biometric characteristics
PAI series		Presentation attack instruments based on a common medium and production method and a single biometric characteristic source
Implementation under test	IUT	That which implements the standard(s) being tested
Subject		The person from whom the biometric enrolment was taken. The target of the attack.
System under test	SUT	The computer system of hardware and software on which the implementation under test operates.

Term	Abbreviation	Definition
Target of evaluation	TOE	Within Common Criteria, the IT product that is the subject of the evaluation. Note: The TOE in Common Criteria evaluations is the equivalent of IUT in biometric evaluations.
Test approach		Totality of considerations and factors involved in PAD evaluation
Vendor		Biometrics system manufacturer

2.2 Presentation-Attack-Detection Conformance Testing

As described above, the results in this report serve as a conformance. No target values for these results exist.

2.2.1 *Definition of Test Criteria*

The test criteria determined the configuration and test cases were performed. The DOT passive liveness application conformance checklist was provided by the vendor during contracting.

Evaluations of PAD mechanisms are classifiable as one of three general types – concealer, verification, or identification. This report is limited to:

- Liveness detection
- Enrollment attacks – such as when an actor attempts to enroll a non-live face for purposes of subverting the system for some reason.
- Application-focused evaluations of PAD mechanisms in which the set/range of attack types is selected to be appropriate to the application, such as those discussed in Clause 11 of ISO 30107-3;
- In particular, this report covers only Level 1 or Level A types of attacks. Such attacks are performed with cooperation by subjects providing authentic biometric samples to create the artefacts, using manufactured materials, and produced and tested in an 8 hour period per subject.

The evaluation did not cover:

- Verification or authentication presentation attacks
- Concealer attacks – such as when an actor attempts to subvert the system by concealing that their biometric is enrolled in a given system.
- Identification attacks – such as when an actor is attempting to be identified in a one-to-many search of a database.

The following metrics were measured and reported here.

APCER – Attack Presentation Classification Error Rate

Overall APCER – is the largest APCER reported for all species

$$APCER = \frac{Number\ of\ Imposter\ Classification\ Errors}{Number\ of\ Imposter\ Attempts}$$

BPCER – Bona Fide Presentation Classification Error Rate

$$BPCER = \frac{Number\ of\ Bona\ fide\ Classification\ Errors}{Number\ of\ Imposter\ Attempts}$$

APNRR – Attack Presentation Non-response Rate

$$APNRR = \frac{Number\ of\ Attack\ Presentation\ Failures\ to\ Match\ or\ to\ Acquire}{Number\ of\ Imposter\ Attempts}$$

BPNRR – Bona Fide Presentation Non-Response Rate
Overall BPNRR – is the largest BNPRR reported for all species

$$BPNRR = \frac{Number\ of\ Bona\ Fide\ Failures\ to\ Match\ or\ Acquire}{Number\ of\ Bona\ Fide\ Attempts}$$

2.2.1.1 *Levels of Evaluation*

Evaluation of PAD could occur at various levels within the biometric system. For example:

- The PAD subsystem may return a classification of the attack or non-attack as live or non-live
- The data capture subsystem may return a classification of the attack or non-attack as live or non-live
- The full system may report the above, or it may only report match/no-match result for a given verification attempt.

Evaluation of PAD for this report consisted of the following:

- The PAD subsystem was tested as it returned a classification of the attack or non-attack as live or non-live when presented with a bona fide subject or a PAD species.

The attack potential of PAD evaluation for this study was performed somewhat similar to Level A of Schuckers (2016), which corresponds to iBeta defined levels as provided in Table 4.

Table 6 Industry Accepted Levels of Attack

Level	Attack Potential	Examples
Level A = iBeta Level 1	Time: short (iBeta 8 hours per subject) Expertise: none Equipment: readily available	Paper printout of face, fingerprint, and palm image and mobile phone display of face and palm photo. iBeta also included: video (from mobile phone), display of face (with movement and blinking), and dusted finger and palm prints.
Level B = iBeta Level 2	Time: >3 days Expertise: moderate skill and practice needed Equipment: available but requires planning	Paper masks, video display of face (with movement and blinking), and fingerprint and palm casting.
Level C = iBeta Level 3	Time: >10 days Expertise: extensive skill and practice needed Equipment: specialized and not readily available	Silicon and theatrical masks.

2.2.2 Test Environment Setup

The test environment consisted of natural lighting as well as regular indoor lighting.

The test platform consisted of two smartphones provided from the iBeta inventory, which included a Google Pixel 2 with Android 8.1.0 and an iPhone 8 with iOS 13.6.

2.2.2.1 Bona-fide population

iBeta utilized 2 testers to provide bona-fide/authentic samples which is not the standard iBeta process. Typically between 3 and 5 testers are available to provide their live faces as the bona fide source; however, due to the Covid-19 pandemic and the requirements to work remotely, only 2 bona fide subjects were utilized. The 5 subjects that provided samples of the facial recognition biometric for the creation of the artefacts were able to be captured by a single tester in the lab. For diversity in the testing within the artefacts, subjects were recruited across age, gender, and ethnic backgrounds such that 40% of the subjects were female, representation was provided from each age group (2 subjects were between the ages of 18-35, 2 subjects were between the ages of 36-53, and 1 subject was between the ages of 54-70), and 1 non-Caucasian subject presented. Subjects were cooperative.

The testing process evaluated liveness only on the test devices. The testers, as the bona fide presentations, applied three imposter samples and then a bona fide sample. This was repeated until 150 PAs were submitted for each of the 6 species yielding 50 bona fide presentations per species.

2.2.2.2 Artefact Generation

For biometric impostor attacks where the subject intends to be recognized as a specific, targeted individual known to the system, it was necessary to create artefacts with three properties:

- Property 1. The sample appears as a natural biometric characteristic to any PAD mechanisms in place.
- Property 2. The sample appears as a natural biometric characteristic to any biometric data quality checks in place.
- Property 3. A sample acquired by a capture device from the artefact contains extractable features that match against the targeted individual's reference.

Artefacts for the testing consisted of six species:

1. 2D photo on matte paper with edges cut
2. 2D photo on matte paper presented on a curved surface
3. 2D photo (as above) with eyes cut out
4. 3D handmade paper mask
5. Video displayed on laptop
6. Video displayed on smartphone

As the subjects were cooperative, each species appeared as a natural face duplication (meeting the requirements of Property 1 and 2). All of the facial features captured in the artefacts contained extractable features as they were acquired from the genuine subject (meeting the requirement of Property 3).

Artefact generation for this system did not rely on white-box or gray-box analysis of the SUT. Iterative techniques were not used during this test effort.

Based on the modality and type of PAD testing being performed, artefact generation was chosen to be captured on smartphones (photos) and from a mid-level digital camera as these are devices that a novice or Level 1 attacker would have available. Similarly, the images were printed either at FedEx or on office printers that iBeta determined would be accessible to a novice or Level 1 attacker. The videos were taken on a Samsung Galaxy S8 or Samsung Galaxy Note 5 and uploaded to a Windows 10 laptop for display.

Per the statement of work for this Level 1 test effort, iBeta performed the testing using cooperative subjects. For example, photos of the test subject’s faces were obtained in office lighting conditions and later used as the PAIs for the testing.

The artefacts were created with minimum effort by the testers in that the creation of the artefacts and presentation of the artefacts were completed in an 8 hour day for each of the 5 subjects. The testers had no specific knowledge of the application functionality and had not habituated to the application prior to testing. The source of the biometric artefacts was access to the cooperative subject. The testers conducted previous facial recognition spoofing projects.

2.2.2.3 Artefact Usage

The tester was provided with the species and artefacts but the decision to use the normal household items within the lab or at home, as well as the lighting levels was not dictated. The tester was allowed to use the items from the lab and items at their workstation, or homes during Covid-19 mandates, with no limits applied. Artefacts were also attached to different backgrounds in some presentations.

Sufficient artefacts were printed so that the photographs could be cut out as the tester determined. The artefacts are durable and may be used repeatedly. The tester kept track of progression of the project and checked in with the Director of Biometrics.

2.2.2.4 Iterative Approaches to Artefact Design

No iterative approaches were used to generate and use artefacts.

2.2.2.5 Test Design

The test design and test case development was conducted for the liveness detection process only.

Innovatrics provided a test application that was ready for testing upon delivery. A successful message that stated “Status: Passed” for the live person or a failure message that stated “Status: Rejected” was displayed for the bona fide and non-live person.

2.2.3 Test Execution

Final test execution was conducted from 21 August through 04 September 2020 and the results are listed in Appendix A. Two software deliveries (for retest) of the application were provided for both devices by Innovatrics.

The subject demographics are provided below in Table 7.

Table 7 Subject Demographics

Subject	Age	Gender	Self-declared ethnicity
1	45	Male	Caucasian
2	26	Female	Caucasian
3	36	Female	Caucasian
4	62	Male	Caucasian
5	30	Male	African American

In summary, the testing was conducted on the Google Pixel 2 and iPhone 8 using the application for the liveness detection solution as follows:

1. The tester(s) then applied Presentation Attack Instrument Species (PAIS) three times each until the application provided results of “Passed” or “Rejected”. All photos were captured with a digital camera/camera phone in Quad HiDef (2560 x 1440). The species were:
    a. 2D photo on matte paper with edges cut
    b. 2D photo on matte paper presented on a curved surface
    c. 2D photo (as above) with eyes cut out
    d. 3D handmade paper mask
    e. Video displayed on laptop
    f. Video displayed on smartphone

2. The sequence was to present 1 bona fide and then 3 PAIs. This was alternated until 150 PAs of each species and 50 bona fides for each subject were presented on the device.
3. All results were recorded.

For each subject, 4 photos were taken with the digital camera and 4 photos were taken with the test smartphone(s). The tester determined how many printouts to use and anywhere between 4 and 8 were utilized for a subject test.

The number of subjects selected and the number of times each species was presented was documented within the contract scope of work. This number and presentation was limited by this being a Level 1 PAD test effort which, by definition, only allowed a tester 8 hours per subject.

Performance metrics discussed in ISO 30107-3 Clause 13 can fail to achieve statistical significance due to limitations in sample size. iBeta determined the metrics that would be recorded and reported during test case development as:

$$APCER_{PAIS} = 1 - \left(\frac{1}{N_{PAIS}}\right)\sum_{i=1}^{N_{PAIS}} Res_i \text{ (1)}$$

Where

N_{PAIS} is the number of attack presentations for the given PAI species;

Res_i takes the value of 1 if the i^{th} presentation is classified as an attack and value 0 if classified as a bona fide

$$BPCER = \frac{\sum_{i=1}^{N_{BF}} Res_i}{N_{BF}}.$$

$$BPNRR = \left(\frac{1}{N_{BF}}\right)\sum_{i=1}^{N_{BF}} Res_i \text{ (2)}$$

Where

N_{BF} is the number of bona fide presentations;

Res_i takes value 1 if the i^{th} presentation produces a non-response or failure to match and value 0 if the bona fide subject matches.

2.2.3.1 Deviations and Exclusions

This report certifies only the following Presentation Attack Detection Testing was performed. ISO 30107-3 covers a number of attack types, system operational types, and evaluation techniques. This report certifies only the following items tested:

- A mobile device authentication system using Innovatrics DOT passive liveness application
- Attacks involving photos and videos
- Evaluation of the PAD classification subsystem.

There were no deviations or omissions from the standard.

3 Biometrics System Identification

The System Identification stipulates the Innovatrics facial recognition biometric application submitted for testing and the hardware, software, and the documentation used in testing.

3.1 Submitted Biometrics System Identification

Table 8 Biometrics System Name and Version

Biometric System Name	Version
DOT passive liveness application	V2.1.0(1) iOS
DOT passive liveness application	V3.0.0 Android

This Biometrics System includes the following:

Table 9 Biometrics System Software

Software Applications	Version	Function Description
Innovatrics DOT passive liveness application	Version 2.1.0(1)	System Under Test on iPhone 8
Innovatrics DOT passive liveness application	Version 3.0.0	System Under Test on Google Pixel 2

3.2 Biometrics System Test Environment

The Biometrics System Test Environment identifies the specific hardware that was used in the test environment. For this test effort, iBeta located all equipment in the biometrics lab or at home due to Covid-19 restrictions.

Table 10 Biometrics System Test Hardware

Hardware	OS or Version	Manufacturer	Description
Google Pixel 2	Android 8.1.0	Samsung	Utilized for testing Model: Pixel 2 Baseband version: g8998-00202-1802061358 Kernel Version: 4.4.88-g3acf2d53921d
iPhone 8	iOS 13.6	Apple	Utilized for testing Model Number: MWLC2LL/A Serial: C8PWL5GRJC6C IMEI: 35 489809 133033 1

Table 11 Other Software, Hardware and Materials

Material	Material Description	Use in the Biometrics System
Other		
Canon EOS Rebel T1	SLR Digital Camera color DS126231	Used to acquire color 2D facial images as attack species.
Samsung Galaxy S8	Model number SM-G950U Serial number RF8JA1T8H4y	Used to acquire video and also to present video on the cell-phone species
Samsung Galaxy Note 5	Model number SM-N920V Serial number R38GA13CGAJ	Used to acquire photo and/or video and also to present photos and video on the cell-phone species
Dell Inspiron 15	Model 3542 Intel Pentium 3542 Windows 7 Home Premium SP1 64-bit	Presentation of attack videos.
Multiple desktop and laptop PCs	A variety of PCs running Microsoft operating systems	Supplied by iBeta: Preparation, management and recording of test plans, test cases, reviews, results and reports

Material	Material Description	Use in the Biometrics System
Microsoft Office 2013	Excel and Word software and document templates	Supplied by iBeta: The software used to create and record test plans, test cases, reviews and results
SharePoint 2010	TDP and test documentation repository	Supplied by iBeta: Vendor document and test documentation repository and configuration management tool
Other standard business application software	Internet browsers, PDF viewers, and email	Supplied by iBeta: Industry standard tools to support testing, business and project implementation
Visual Studio 2013 v.12.0.40629.0 Update 5 (Microsoft)	Build and source code Integrated Development Environment	Supplied by iBeta: View source code
Beyond Compare 4 v.4.1.9 (Scooter Software)	Comparison utility	Supplied by iBeta: used to compare file/folder differences
WinDiff 5.1 (Microsoft)	Comparison utility	Supplied by iBeta: used to compare file/folder differences

No documents from Innovatrics were delivered for this test effort.

4 Biometrics System Overview

The application consists of a biometric face liveness detection system. The application is a facial liveness product that incorporates built-in presentation attack detection.

Innovatrics DOT passive liveness applications (versions 2.1.0[1] iOS and 3.0.0 Android) were tested on two smartphone test platforms using the front facing (selfie) camera. Enrollment was conducted in accordance to the instructions within the application.

5 Conformance Review and Test Results

The results and evaluations of the tests are identified below. Detailed data regarding the Acceptance/Rejection criteria, reviews and tests are found in the appendices.

- Appendix A identifies all test results for Conformance Testing

5.1 Limitations

The results and conclusions of this report are limited to the specific IUT/SUT applications and versions described below.

It is the responsibility of the vendor to provide the laboratory with systems and devices which are representative of those systems and devices produced for the consumer.

These results represent usage of falsification testing methodology. Testing can only demonstrate non-conformity, i.e., if errors are found, non-conformance of the SUT shall be proven, but the absence of errors does not necessarily imply the converse. These results are intended to provide a reasonable level of confidence and practical assurance that the SUT conforms to the standard. Use of these results will not guarantee conformity of an implementation to the standard; that normally would require exhaustive testing, which is impractical for both technical and economic reasons.

As described elsewhere, this report covers only Level 1 or relatively low level PAD species for the biometric system under test.

IBeta did attempt to differentiate classification errors from non-responses. All results are reported as the subject or attack species was either classified as live or non-live. Innovatrics has indicated that the system responses do not normally provide classification responses to mitigate hill-climbing attacks against the system.

5.2 PAD Testing Results

The application provided by Innovatrics did operate sufficiently during PAD Testing.

5.2.1 Innovatrics DOT Passive Liveness Version 3.0.0 Android Results

As stated above in Section 2.2.3, bona fide presentations were alternated with presentation attacks.

BPCER on the Pixel 2 was 0%. In total, there were 300 bona fide attempts with 300 successes. As stated previously, 1 tester acted as the bona fide subject due to the work-from-home order.

Both the BPNRR and the APNRR on the Android application was 0%. There were no instances where the presentation of the bona fide did not receive a message from the application. For the APNRR, the tester set a 30 second time limit before declaring a non-response when an artefact was presented but this limit was not met so no APNRs were recorded.

For APCER, iBeta considered a single result from the PAs if the application accepted the artefact as alive. The Artefacts were presented approximately 150 times each to yield an APCER of 0 of 150 for each species and 0 of 900 presentations overall.

Table 12 Android Results

	Test Species	**Android Innovatrics**		
		PAs	**APCE**	**APCER**
1.	2D Photo on matte paper with edges cut out	150 per subject	0 of 150 per subject	0%
2.	2D Photo with eyes cut out	150 per subject	0 of 150 per subject	0%
3.	2D photo on matte paper presented on a curved surface	150 per subject	0 of 150 per subject	0%
4.	3D Layered paper photo	150 per subject	0 of 150 per subject	0%
5.	Video displayed on laptop	150 per subject	0 of 150 per subject	0%
6.	Video displayed on smartphone	150 per subject	0 of 150 per subject	0%
Total per species			0 of 150	0%
Total for all species			0 of 900	0%

5.2.2 Innovatrics DOT Passive Liveness Version 2.1.0(1) iOS Results

As stated above in Section 2.2.3, bona fide presentations were alternated with presentation attacks.

BPCER on the iPhone 8 was 0.7%. In total, there were 302 bona fide attempts with 300 successes. As stated previously, 2 testers acted as the bona fide subjects due to the work-from-home order.

Both the BPNRR and the APNRR on the iOS application were 0%. There were no instances where the presentation of the bona fide did not receive a message from the application. For the APNRR, the tester set a 30 second time limit before declaring a non-response when an artefact was presented but this limit was not met so no APNRs were recorded.

For APCER, iBeta considered a single result from the PAs if the application accepted the artefact as alive. The Artefacts were presented approximately 150 times each to yield an APCER of 0 of 150 for each species and 0 of 900 presentations overall.

Table 13 iOS Results

	Test Species	iOS Innovatrics		
		PAs	APCE	APCER
1.	2D Photo on matte paper with edges cut out	150 per subject	0 of 150 per subject	0%
2.	2D Photo with eyes cut out	150 per subject	0 of 150 per subject	0%
3.	2D photo on matte paper presented on a curved surface	150 per subject	0 of 150 per subject	0%
4.	3D Layered paper photo	150 per subject	0 of 150 per subject	0%
5.	Video displayed on laptop	150 per subject	0 of 150 per subject	0%
6.	Video displayed on smartphone	150 per subject	0 of 150 per subject	0%
Total per species			0 of 150	0%
Total for all species			0 of 900	0%

5.2.3 Exclusions

When interpreting the performance of a PAD subsystem, it is important to recognize that there may be presentation attack types, PAI species and factors which have not been tested. Therefore, the reported performance of a PAD subsystem does not provide any information regarding its effectiveness in detecting presentation attacks which have not been tested.

6 Opinions & Recommendations

6.1 Recommendations

iBeta Quality Assurance has completed the Level 1 PAD testing of Innovatrics DOT passive liveness applciations -Versions 2.1.0(1) for iOS and 3.0.0 for Android. The purpose of this report is to describe the testing performed and the metrics obtained for that testing. Conformance to any criteria was not tested.

Based on the test results of Section 5, the overall system design and construction of the application meets all of the normative requirements with the ISO/IEC 30107-3 for Level 1 testing.

iBeta Quality Assurance confirms that Innovatrics DOT passive liveness applications - Version 2.1.0(1) for iOS and Version 3.0.0 for Android meet the Level 1 criteria for Presentation Attack Detection.

6.1.1 Limitations

As described in section 5.1 Limitations, iBeta has tested what it believes to be a representative sample of the commercially available system and used the appropriate test methods to test conformance to the standards.

As stated also in Section 2.0, this report does not contain a certification per se, but only results of testing per a certified procedure. There are no ISO 30107-3 requirements stating specific levels of passing or failing values for example of BPCER and APCER.

The results reported here were obtained during PAD testing of the Innovatrics DOT passive liveness applications - Versions 2.1.0(1) iOS and 3.0.0 Android provided.

6.1.2 Exceptions

There were no exceptions to the test method. The data supporting this review is found in Appendix A.

6.2 Opinions

iBeta has no other remarks or opinions not reflected in the above report.

Gail Audette

Gail Audette
iBeta Quality Assurance Director of Biometrics
GAudette@ibeta.com
303.627.1110 extension 182

APPENDICES: TEST OPERATION, FINDINGS & DATA ANALYSIS

A.1 Appendix A: PAD – Test Case 1

		Test Case 1 - Presentation Attack Detection (PAD)						
		Execute PAD Artefact testing	**Test Objective**					
		PAD artefacts have been designed PAD artefacts have been generated. Subjects have been identified.	**Test Prerequisite** **• Pass (P): the expected result is observed** **• Fail (F): the expected result of the test case is not observed** **• Not Testable (NT): rejection of a previous test step prevents execution of this step.** **• Not Applicable (NA): not applicable to test scope**					
		ISO/IEC 30107-3 informed by ISO/IEC 30107-1	**Record Standards or non-standard test methods**					
		Passive Liveness 2 - v2.1.0(1) iOS/ v3.0.0 Android	**SUT**					
		08/21/20-09/04/20	**Dates**	78.8%	0%	0%	21.2%	
		Gail Audette, Ricky Brown, Karen Wilson-Winfrey, Ryan Borgstrom	**Validator (s)**	**26**	**0**	**0**	**7**	
Req.	**Step**		**Expected Result**	**Pass**	**Fail**	**NT**	**NA**	**Notes**
		Getting Started: - Fill out or have the customer fill out Pre-engagement Checklist and use that to fill out the Info Tab. This is a required table in the report - If necessary for Security, install TrueCrypt or VeraCrypt on the Test System Create a P:\ drive encrypted system with TrueCrypt to contain the transactions Make the file at least 4 GB, use the assigned password - Mount the TrueCrypt volume to P:\ - Collect the Biometric PII or artefact data		**X**				

		Test Steps						
		PAD Type						
9	1	Type of presentation: Concealer, Imposter	Liveness detection	**X**				
7.1	2	Evaluations of PAD mechanisms and resulting reports shall specify the applicable evaluation level, whether PAD subsystem, data capture subsystem, or full system. The resulting reports should discuss how the evaluation level influenced PAD testing.	PAD subsystem on Android and iOS device	**X**				
11.3-11.4	2a	If a Comparison subsystem, record verification or identification	Liveness detection only				**X**	Not a comparison subsystem
11.3-11.4	2b	If identification subsystem, record concealer or imposter	N/A				**X**	Not an identification subsystem
11.3-11.4	2c	Record the PAD Certification Test Procedure: Summary of Required Metrics for the test to be performed as described in step 3 and the Req. Per Subsystem tab for required metrics	Presentation Attack Detection Certification Test Procedure v1.0 dated 1/24/18	**X**				
	2d	Record the number of species	The species is a class of presentation attack instruments created using a common production method and based on different biometric characteristics. For this Level 1 pad testing, there will be 6 species.	**X**				
		For each PAD species						
	3	Describe the PAD mechanism	Cooperative subject photo laser printed color, cooperative subject video.	**X**				
7	4	Describe how or why the species is expected to meet: For biometric impostor attacks in which the subject intends to be recognized as a specific, targeted individual known to the system, it will be necessary to create an artefact with three properties: — Property 1. The sample appears as a natural biometric characteristic to any PAD mechanisms in place. — Property 2. The sample appears as	Species 1: The cooperative subject photo appears in a life like position as normal person would, the photo has similar dimensions and cut out to remove traces of it being a photo. Species 2: Similar to Species 1, this is displayed on a curved surface, against a paper towel roll, for example. Species 3: Similar to Species 1,	**X**				

		a natural biometric characteristic to any biometric data quality checks in place. — Property 3. A sample acquired by a capture device from the artefact contains extractable features that match against the targeted individual's reference.	this has eyes cut out. Species 4: This species uses video displayed and the subject's liveness as well as the laptop's larger size to simulate a head's natural size. Species 5: The 3D paper mask uses layering to display depth Species 6: This species uses video displayed on Android Device and the size can be adjusted and the positioning changed.					
8.1	5	Describe how or why the species is expected to meet: Artefacts created for the biometric concealer attack are meant to appear as a natural biometric characteristic to any PAD mechanisms and any biometric quality checks in place. Such artefacts should contain extractable features that can be compared to stored references. In addition to Properties 1 and 2, artefacts in biometric concealer attacks should also have the following property (continuing the list of properties from 8.1): — Property 4. The extractable features should not match any stored references.					X	Not a concealer test.
8.2	6	Describe in the Species-x or alternate tab the reasoning for artefact creation and preparation	This is a Level 1 test and per the contract, each subject is given 8 hours to produce the artefacts with no expertise and with only equipment readily available in a normal home or office environment. The subject is cooperative meaning that the biometric characteristics are capture directly from the individual with assistance such as a photo or video for facial recognition systems.	X				
10.2	6a	Peer-review the PAI species and PAI series	Previously used on facial recognition PAD level 1 and peer	X				

			reviewed by the Director of Biometrics.					
	7	Record the number of specimens in the series	The series are presentation attack instruments based on a common medium and production method and a single biometric characteristic source. For this test effort, the 5 species will be: 1. 2D photo on matte paper with edges cut 2. 2D photo on matte paper presented on a curved surface 3. 2D photo (as above) with eyes cut out 4. Video displayed on a laptop 5. 3D handmade paper mask 6. Video displayed on an Android device	**X**				
11.2-11.4	8	Record the number of bona fide subjects	5 Subjects will provide artefacts	**X**				
		PAD mechanism subsystem						
	9	Evaluations of PAD mechanisms and resulting reports shall describe whether evaluation design considered enrolment, identification, and/or verification processes, or alternatively whether evaluation design considered a generic biometric sub-system independent of a specific process.	Liveness detection only on a PAD subsystem.	**X**				
11.2	10	Evaluations of PAD mechanisms and resulting reports that apply to enrollment processes shall describe the following: — use of enrollment-specific quality thresholds or presentation policy; — parameters of the enrolment transaction, including number and duration of presentations; — level of operator oversight present in the process; — manner in which operator functions were applied or emulated in the evaluation.	Liveness detection being tested as enrollment process. APCER will be classified as any PAI that can be successfully recognized by the application. BPCER will be classified as any bona fide presentation that is unable to be recognized on the application. APNRR will be classified as any PAI that is unable to produce the green circle and 'Passed" result or red X and "Rejected" result. This will be the same for BPNRR	**X**				

11.3	10a	Evaluations of PAD mechanisms and resulting reports that apply to verification processes shall describe the following: — use of quality thresholds and presentation policy; — parameters of the verification transaction, including the number and duration of presentations; — level of operator oversight present in the process; — manner in which operator functions were applied or emulated in the evaluation.					**X**	Verification not being tested.
11.4	10b	Evaluations of PAD mechanisms and resulting reports that apply to identification processes shall describe the following: — use of quality thresholds and presentation policy; — parameters of the identification transaction, including the number and duration of presentations; — configuration of system to perform negative or positive identification; — whether capture subjects were enrolled in the databases against which identification took place; — level of operator oversight present in the process; — whether and how an operator adjudicates candidate identities returned by the system; — manner in which operator functions were applied or emulated in the evaluation.					**X**	Identification not being tested.
11.5	10c	Reports that evaluate offline PAD mechanisms shall describe their implementation in the overall processing scheme.					**X**	Authentication is not being conducted off-line.
		As applicable. The next three categories are mutually exclusive a) PAD subsystem,						

		b) data capture subsystem, or c) full system						
13.1	11	Evaluations of PAD mechanisms shall report the following:	PAD subsystem	**X**				
	a	— number of presentation attack instruments, PAI species, and PAI series used in the evaluation;	There will be 6 species all presentation (manually). There will be no duplicate artefacts in the PAI series.	**X**				
	b	— number of test subjects involved in the testing, including those unable to utilize artefacts or present non-conformant characteristics;	5 subjects will provide artefacts.	**X**				
	c	— number of artefacts created per test subject for each material tested;	There will be multiple copies of each of the 6 artefacts.	**X**				
	d	— number of sources from which artefact characteristics were derived;	All artefacts are derived from cooperative subjects.	**X**				
	e	— number of tested materials;	Test materials include photographs, cameras, cell phones, videos, photo paper, etc.	**X**				
	f	— description of output information available from PAD mechanism;	Center your face, Move closer, Move back, Stay still, and Move towards light were the messages displayed.	**X**				
	g	— ordering of subject presentations with and without PAI, and whether subjects were reused;	Subject order will follow the data sheet, and no subjects will be reused.	**X**				
	h	— ordering of subject presentations to the PAD enabled and disabled system, and whether subjects were reused.	PAD system enabled at all times	**X**				
13.2.2	12	Evaluations of PAD mechanisms shall report the number of artefact presentations correctly and incorrectly classified: total, by PAI species, by PAI series, by capture subject, and by source.	This will be documented in the final report.	**X**				
13.2.2	13	Evaluations of PAD mechanisms shall report the number of bona fide presentations correctly and incorrectly classified – total.	This will be documented in the final report.	**X**				
		b) Data Capture Subsystem						
13.3.2	14	In data capture subsystem evaluations, performance metrics for presentation attacks shall be calculated and reported as APCER and BPCER.	This will be documented in the final report.	**X**				

13.3.3	15	The evaluator shall report non-response rates of the data capture subsystem using the following metrics: — for each PAI species, APNRR and the sample size on which the computed rate is based; — BPNRR and the sample size on which the computed rate is based.	This will be documented in the final report.	**X**				
13.3.3	16	The evaluator shall report capture rates of the data capture subsystem using the following metrics: — for each PAI species, attack presentation acquisition rate (APAR) and the sample size on which the computed rate is based; — for bona fide capture subjects erroneously rejected by capture or quality sub-systems, FTA and/or FTE as defined in ISO/IEC 19795-1 and the sample size on which the computed rate is based.	Capture rates of data capture subsystem will not be recorded				**X**	
		c) Full system (For verification or identification subsystems)						
13.4.2	17	For verification systems, for each PAI species, at least one of the following shall be reported: — IAPMR and the sample size on which this computed rate is based; — CAPNMR and the sample size on which this computed rate is based.	N/A - Verifications will not be recorded.				**X**	Not a verification system
13.4.2.2	18	For positive identification systems, for each PAI species, impostor attack presentation identification rate (IAPIR) and the sample size on which the computed rate is based shall be reported.	N/A				**X**	Not a positive identification system.
13.4.2.3	19	For negative identification systems, for each PAI species, concealer attack presentation non-identification rate (CAPNIR) and the sample size on which the computed rate is based shall be reported.	N/A				**X**	Not a negative identification system.

A.2 Appendix A: PAD Testing - Test Case

		Test Case – PAD Testing						
		To document and perform the PAD testing (Level 1) per the contract	Test Objective					
		On device application PAD Artefacts have been designed. PAD Artefacts have been generated. Subjects have been identified.	Test Prerequisites					
		ISO 30107-3	Record the relevant regulations.					
		Passive Liveness 2 - v2.1.0(1) iOS and v3.0.0 Android	Verify and Record any CTS Name and Version					
		08/21/20-09/04/20	Validation Date	100 %	0%	0%	0%	
		Gail Audette, Ricky Brown, Ryan Borgstrom, and Karen Wilson-Winfrey	Validator(s)	9	0	0	0	
Reference	Test Step		Expected Result	Pass	Fail	NT	NA	Notes
		Getting Started: Complete the prerequisites; Verify the environment & installation of the CTS; and record the testers & date Record the filename of the test data	The test environment matches any vendor documentation.	X				
		Configuration of the CTS and Test Environment - include any pertinent configuration information of the CTS . archive configuration files if they exist . make note of any special settings - include . OS, Service Pack, . hardware description . Network ID and iBeta tag number . the name of the Assessment Spreadsheet summarizing this test case	CTS and Test environment configuration. On device via a delivered application. The application verified the licenses via the in CTS and Test environment configuration <u>iPhone 8:</u> iOS: 13.5.1 Model: MQ722LL/A Serial Number: C8PWL5GRJC6C <u>Google Pixel 2:</u> Android: 8.1.0 Model: Pixel 2 Baseband version: g8998-00202-	X				

			1802061358 Kernel Version: 4.4.88-g3acf2d53921d					
		In the event that there are deviations or exclusions to the test method, the test lead shall document, technically justify, and notify the project lead and the project lead shall document and notify the vendor and obtain vendor approval prior to performing the testing. Insert a row for each such deviation or exclusion here or at the appropriate spot in the test case.	Technical Justification ref: Vendor notification ref:	X				
		Test Steps/Setup						
	1	Acquire and record the final version of the applications:		X				
	2	Passive Liveness 2 for Android	https://drive.google.com/drive/folders/1zBPrOjMPfwBsTjiQjaOo6dnUt2sdSYvY	X				
	3	Passive App for iOS	Redeemed in TestFlight (code was in email from Innovatrics and was one-time use). Passive Liveness App - 1.3(1) 6/19/20	X				
		PAD Testing – iPhone 8						
	4	Record iOS device information	iPhone 8 - iOS Version 13.6 Model Number: MWLC2LL/A Serial:FK1ZCQQUN72N IMEI: 35 398510 157086 4	X				
	5	Install the application on the iPhone 8	Open TestFlight to and redeem access code (delivered via email).	X				
	6	Select Passive Liveness Check		X				
	7	Position face inside of the oval		X				
	8	Record messages during capture	Center your face, Move closer, Move back, Stay still, and Move towards light were the messages displayed.	X				
	9	Define a non-response	The capture process does not time out. A non-response occurs after 30 seconds and 3 attempts.	X				
	10	Record output messages	Status: Rejected or Passed	X				
		PAD Testing - Google Pixel 2						
	11	Record Android device information	Google Pixel 2 - Android Version Android: 8.1.0	X				

			Model: Pixel 2 Baseband version: g8998-00202-1802061358 Kernel Version: 4.4.88-g3acf2d53921d					
	12	Install the Android application on the Google Pixel 2	Link delivered via email.	**X**				
	13	Follow steps 6-10		**X**				