Цели и задачи алгоритмов Liveness

Цель

Противодействие биометрическому спуфингу*

Стандарты и рекомендации

- **ISO/IEC WD 30107-3** (Link);
- **FIDO Biometric Requirements** (Link)

Способы захвата изображения

- **RGB-камера**

- **ИК-камера**

- **Depth-камера (карта глубины)**

Задачи

Распознавание атак (Presentation Atack Detection, PAD) следующих типов:

- **Printed Photo Attack** – атака с помощью одной или нескольких фото человека);
- **Printed Mask Attack** – атака с помощью вырезанной фото (бумажной маски), часто с вырезанными отверстиями для глаз и рта;
- **Video Replay Attack** – атака с помощью видеозаписи с человеком;
- **3D Mask Attack** – атака с помощью маски (силиконовой, керамической или пластмассовой).

Методы

- **Активные** (кооперативные) – методы проверки Liveness, при которых объект участвует напрямую в процессе (необходимо посмотреть в камеру или совершить действие – моргание, улыбка, поворот головы и проч.)
- **Пассивные** (некооперативные) – методы проверки Liveness, при которых объект проверки может не взаимодействовать с системой проверки Liveness и даже не знать о её существовании.

* Биометрический спуфинг – атака на лицевую биометрическую систему с использованием изображения человека (например, фотографии) с целью обмана процесса аутентификации

AGSR
АГЕНТСТВО СЕРВИСИЗАЦИИ И РЕИНЖИНИРИНГА

Примеры атак

Real
Реальный пример / Живой человек

Spoofing Attack

	Printed Бумажные фейки	**Printed mask** Бумажные маски	**Replay** С экрана устройств	**3D Mask** 3D-маски
Статичное изображение				
В движении				

Уровень A* **Уровень B*** **Уровень C***

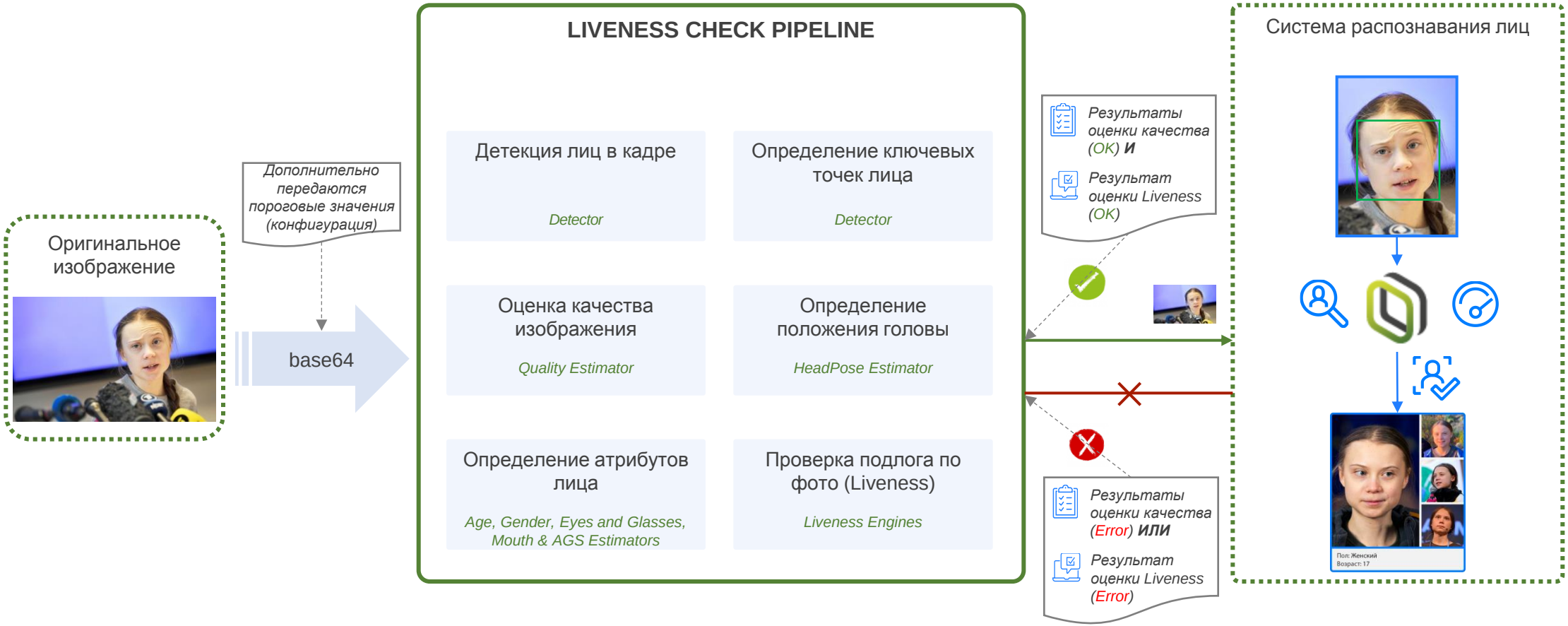
* Подробно описаны в приложении

AGSR
АГЕНТСТВО СЕРВИСИЗАЦИИ И РЕИНЖИНИРИНГА

Уровни атак по классификации FIDO*

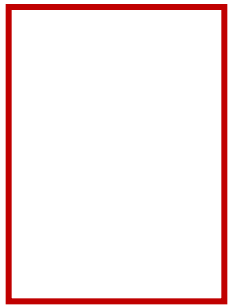
LVL	ОПИСАНИЕ	ЭКСПЕРТИЗА	ВРЕМЯ	ОБОРУДОВАНИЕ	ИСТОЧНИК ДАННЫХ
A	Атаки заключаются в получении и использовании фотографии человека, подвергающегося нападению. Атаки этого уровня не предъявляют высоких требований к оборудованию и навыкам. Примеры видов PAI** уровня A: • набор изображений лиц, напечатанных на струйных/лазерных принтерах; • фотографий, напечатанных в фотолаборатории; • фотографий, отображаемых на мобильных устройствах и мониторах (подвиды – разные модели мобильных телефонов, планшетов, мониторов и т.д.) Кроме того, может быть выполнена предобработка для улучшения фотографии; вырезаны отверстия для глаз, носа, рта или контура лица. Любые изменения, подобные этому, будут отнесены к разным видам и подвидам*** PAI.	Дилетант	< 1 дня	Бумажная распечатка, экран телефона	Фото из соц. сетей
B	Атаки схожи с атаками уровня A, за исключением того, что требуется не фотография лица, а видео субъекта. Кроме того, с изображением лица в высоком разрешении можно создать бумажную маску человека. Примеры видов PAI уровня B: • показ видео на электронных устройствах (в т.ч. различные кооперативные видео с морганием, поворотами головы и т.д.), • бумажные маски с добавлением трёхмерности, полумаски.	Профессионал	< 7 дней	бумажная маска, видео лица с движением	Фотография высокого качества, видео
C	Атаки включают использование более сложных масок, которые сделаны не из бумаги, а из специализированных материалов. Создание этих масок занимает больше времени, они более дороги и требуют фотографии с высоким разрешением и / или 3D-информации. Трехмерная информация также может быть получена из двумерной фотографии с использованием сложных методов компьютерного зрения. 3D маски могут быть жесткими с глазными отверстиями и без них, либо гибкими силиконовые маски и 3D-отпечатанные цветные копии лица. Примеры видов PAI уровня C: • керамические или силиконовые маски	Эксперт	> 7 дней	3D маска	Изображение / серия изображений высокого качества, 3D информация о лице

Описание процесса

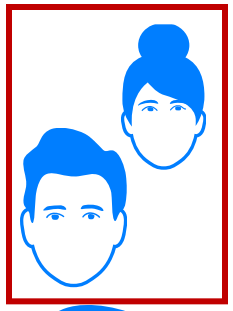


AGSR
АГЕНТСТВО СЕРВИСИЗАЦИИ И РЕИНЖИНИРИНГА

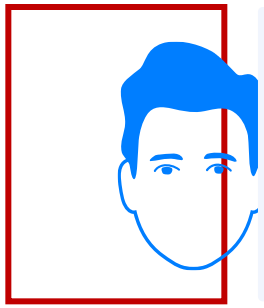
Типовые ошибки при проверке качества



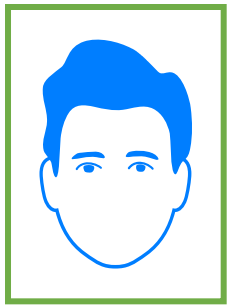
Отсутствие лица на изображении

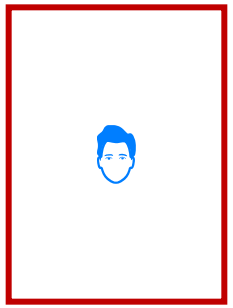


Несколько лиц на изображении

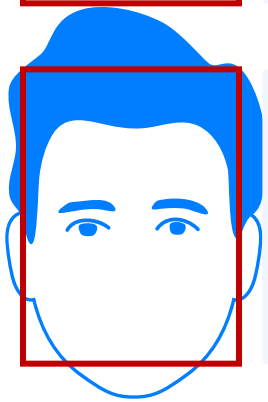


Часть лица вне кадра





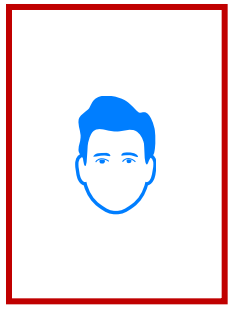
Лицо на изображении меньше заданного



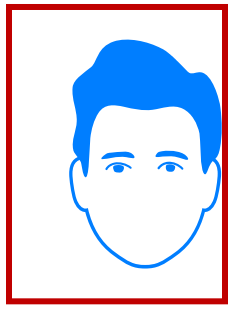
Лицо на изображении больше заданного

Перекрытия лица

OK

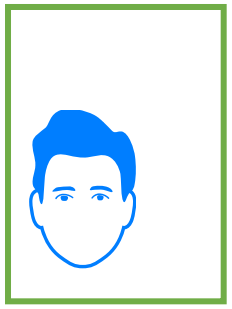


Межзрачковое расстояние меньше заданного



Отступ от края меньше заданного

Углы поворота головы больше заданных



AGSR
АГЕНТСТВО СЕРВИСИЗАЦИИ И РЕИНЖИНИРИНГА

Используемые метрики

False Accept Rate (FAR)

Ошибки 2 рода. Доля ошибочных подтверждений верификации фотографий разных людей. Применяется в контексте модуля распознавания лиц

$$FAR\ (\%) = N_{\text{False accepted face recognition transactions}} / N_{\text{Total number of face recognition transactions}} * 100$$

Attack Presentation Classification Error Rate (APCER)

Отношение успешно пройденных атак к общему количеству попыток атак. Применяется в контексте модуля liveness.

$$APCER\ (\%) = N_{\text{Accepted liveness transactions}} / N_{\text{Total number of spoof transactions}} * 100$$

False Reject Rate (FRR)

Ошибки 1 рода. Доля ошибочных отказов верификации фотографий разных людей. Применяется в контексте модуля распознавания лиц

$$FRR\ (\%) = N_{\text{False rejected face recognition transactions}} / N_{\text{Total number of face recognition transactions}} * 100$$

Bona fide Presentation Classification Error Rate (BPCER)

Отношение верно классифицированных добросовестных прохождений liveness к общему количеству добросовестных попыток. Применяется в контексте модуля liveness.

$$BPCER\ (\%) = (1-N_{\text{accepted liveness transactions}} / N_{\text{Total number of bonafide transactions}}) * 100$$

Impostor Attack Presentation Match Rate (IAPMR)

Отношение успешно пройденных сессий атак к общему количеству сессий атак. Сессия считается пройденной, если пройдена хотя бы одна из K попыток пройти

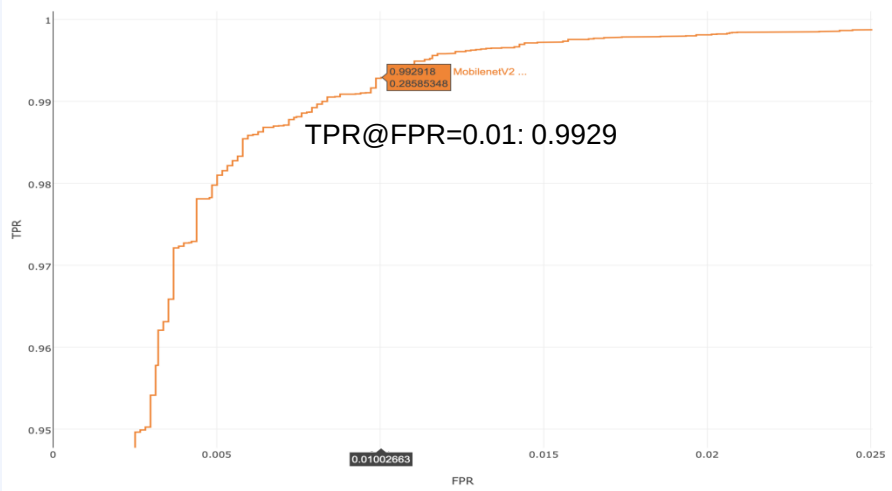
$$IAPMR\ (\%) = N_{\text{Accepted sessions}} / N_{\text{Total number of attack sessions}} * 100$$

Bona fide Presentation Error Rate (BFPER)

Отношение верно классифицированных добросовестных прохождений сессий к общему количеству добросовестных сессий. Сессия считается пройденной, если пройдена хотя бы одна из K попыток пройти

$$BFPER\ (\%) = (1-N_{\text{Accepted sessions}})/ N_{\text{Total number of bona fide sessions}} * 100$$

Пример: ROC-кривая, полученная при тестировании IR Liveness v2 на level A testset



	TPR@FPR=0.1	TPR@FPR=0.01
Level A	**0.9998**	**0.9929**
Level B	**0.9995**	**0.9472**

Примечание: Дополнительно могут использоваться метрики:
HTER (Half-Total Error Rate) = (FAR + FRR) / 2; TPR (True positive rate) = 1 – FRR; FPR (False positive rate) = FAR

Positive class: real; Negative class: fakes

AGSR
АГЕНТСТВО СЕРВИСИЗАЦИИ И РЕИНЖИНИРИНГА

Документирование тестирования алгоритмов

Протокол тестирования

- Фиксирование кейса использования системы
- Описание разнообразия условий использования системы
- Фиксация необходимой точности системы
- Набор тестовой группы и сбор тестовых данных в реальных условиях максимально приближенных к боевым
- Фиксация устойчивого кейса взлома при обнаружении (необходима воспроизводимость взлома)
- Запуск системы в тестовом режиме (silent mode), сбор статистики ошибок второго рода
- Запуск системы в боевом режиме

Правила сравнения алгоритмов / вендоров

- Фиксирование кейса использования системы
- Фильтрация вендоров по требованиям системы liveness
- Сбор общих тестовых данных: сбор под каждую систему отдельно и объединение
- Первый отбор на основе тестовых данных
- Тестирование в боевом режиме основных вендоров на реальных кейсах с увеличением тестовой выборки и времени проведения теста

Отчёт о проведении тестирования

- Описание тестовой среды;
- Описание тестируемой платформы;
- Количество биометрических шаблонов в базе распознавания лиц;
- Количество и описание типов атак с привязкой к их уровню по FIDO;
- Количество испытуемых, участвующих в тестировании;
- Количество атакующих приспособлений (PAI) для каждого испытуемого;
- Таблицы результатов по участникам, атакам (см. Приложение);
- Полученные IAPMR (общий и для каждой атаки), BFPMR. Для более детальных результатов - полученные APCER, BPCER;
- Заключение (общее впечатление/сложные кейсы/вердикт о готовности системы).

Примечание:
Тестирование должно проводиться по аналогии с методикой FIDO. Количество участников тестирования должно быть не менее 10.
Цель тестирования – оценка точности противодействия системы атаке самозванца (Impostor Attack Presentation Match Rate, IAPMR) и свободному пропусканию добросовестных людей (Bona fide Presentation Match Rate, BPMR).

AGSR
АГЕНТСТВО СЕРВИСИЗАЦИИ И РЕИНЖИНИРИНГА

IR Liveness (Passive)

Описание

Алгоритм основан на изображении ближнего инфракрасного диапазона.
Преимущества: Изображения с телефонов/мониторов не отображаются в данном диапазоне, устойчив к атакам с бумаги.
Недостатки: Может понадобиться дообучение под новые камеры. Плохая работа при прямых солнечных лучах.
Рекомендации по использованию:
Требуется камера с инфракрасным сенсором. Важна подсветка. Качественная работа только в статичных (неменяющихся) условиях.

Совместимость

Камеры:
- RealSense SR300;
- RealSense D415;
- RealSense D435;
- Mouse;
- IR-модуль Devices

Сценарии применения

Любой в помещении

Оценка устойчивости к атакам

A B C
★★☆

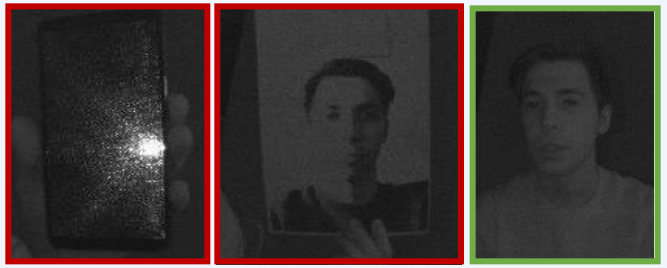
Наличие в продуктах

- Windows & Linux

Примеры

RGB-изображение

Изображение в IR



AGSR
АГЕНТСТВО СЕРВИСИЗАЦИИ И РЕИНЖИНИРИНГА

Depth Liveness (Passive)

Описание

Алгоритм основан на паре изображений RGB (для детекции лица) и карты глубины (для проверки liveness).

Преимущества: устойчив к самым популярным атакам с бумаги и с телефонов

Недостатки: Слабоустойчив к 3D маскам под конкретного человека, дорогие камеры

Рекомендации по использованию:
Качественная работа только в статичных условиях IR засветки (исключить прямой солнечный свет).

Совместимость

Камеры:
- RealSense D415;
- RealSense D435;
- IR-модуль Devices

Сценарии применения

Любой в помещении (СКУД, СУО, АТМ и т.д.)

Оценка устойчивости к атакам

A B C
★ ★ ☆

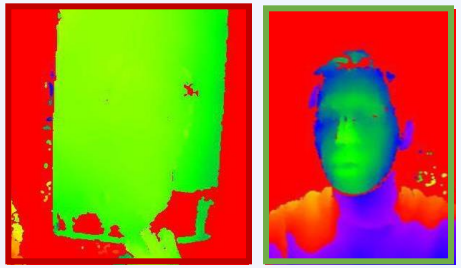
Наличие в продуктах

- Windows & Linux

Примеры

RGB-изображение

Карта глубины



СКУД Liveness (Passive)

Описание

Алгоритм основан на получении дополнительной информации с фиксированного состояния камеры и постоянной работы системы.

Преимущества: Устойчив к самым популярным атакам с бумаги и с телефонов

Недостатки: Слабоустойчив к маскам под конкретного человека - камера должна быть установлена с выдаваемой спецификацией.

Рекомендации по использованию:

Камера должна быть установлена неподвижно (с момента инициализации фон не должен радикально меняться)

Совместимость

IP-камеры (например, ACTi E38, AXIS Q1615, Bosch NBN-50022-V3, Dahua IPC-HDW5231RP-ZE, Hikvision DS-2CD2822F, Samsung XNV-8040RP, Vivotek IB9367-EHT и другие)

Сценарии применения

Системы контроля и управления доступом (СКУД)

Оценка устойчивости к атакам

A B C

★ ★ ☆

Наличие в продуктах

- Windows & Linux;
- FaceStream

Примеры

Кадры с камер, установленных на турникетах

Качественные изображения (условия соблюдены)

Некачественные изображения (условия не соблюдены)

AGSR
АГЕНТСТВО СЕРВИСИЗАЦИИ И РЕИНЖИНИРИНГА

Devices Liveness (Passive)

Описание

Алгоритм основан на анализе RGB изображения и изображения ближнего инфракрасного диапазона, сделанные в один момент времени.

Преимущества: Изображения с телефонов/мониторов не отображаются в IR, устойчив к атакам с бумаги.
Недостатки: Плохая работа при прямых солнечных лучах.

Совместимость

Сценарии применения

Любой в помещении (СКУД, СУО, АТМ и т.д.)

Оценка устойчивости к атакам

A B C

★ ★ ⯪

Наличие в продуктах

- Devices

Примеры

Изображение с устройства

RGB: IR:

AGSR
АГЕНТСТВО СЕРВИСИЗАЦИИ И РЕИНЖИНИРИНГА

FPR PC Liveness (Passive)

Описание

Алгоритм включает комплекс проверок:

- Replay Liveness - определяет, артефактов видеозаписи;
- Phone Liveness - определяет, есть ли в расширенном bbox телефон;
- FlyingFaces Liveness - определяет распечатанные фотографии и маски.

Преимущества: устойчив к самым популярным атакам с бумаги и с телефонов

Недостатки: Слабоустойчив к 3D маскам под конкретного человека

Рекомендации по использованию:

Расстояние до камеры не более 80 см (на расстоянии вытянутой руки)

Совместимость

Web-камеры, IP-камеры с качеством не ниже 3.7MP, 720p (иначе шумы убивают высокочастотные признаки)

Сценарии применения

Двухфакторная аутентификация и иные кооперативные/некооперативные сценарии

Оценка устойчивости к атакам

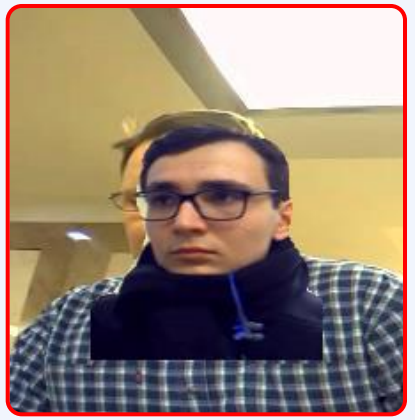
A B C

★ ★ ☆

Наличие в продуктах

- Windows & Linux

Примеры



AGSR
АГЕНТСТВО СЕРВИСИЗАЦИИ И РЕИНЖИНИРИНГА

Active Mobile Liveness

Описание

Алгоритм основан на получении пространственной информации за счет **кооперативного** взаимодействия с устройством (приближение-отдаление устройства, повороты головы, улыбка, моргание и иные действия).

Преимущества: Устойчив к самым популярным атакам с бумаги и телефонов.

Недостатки: Поддержка огромного разнообразия устройств, из-за чего слабое тестирование на всех возможных условиях. Невысокая скорость работы на старых процессорах.

Рекомендации по использованию:
Использовать вместе с пассивным Liveness

Совместимость

Камеры мобильных устройств (телефоны, планшеты и т.д. с ОС iOS 10.0+ или Android 7.0+)

Сценарии применения

Мобильная аутентификация и иные кооперативные сценарии

Оценка устойчивости к атакам

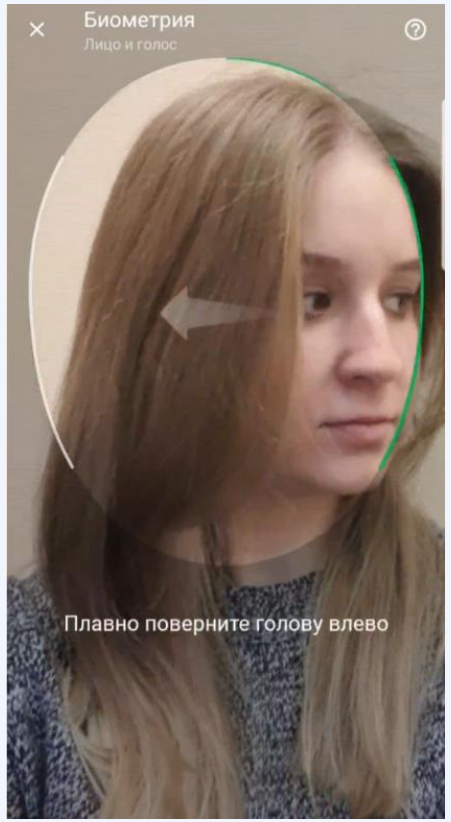
A B C

★ ★ ☆

Наличие в продуктах

- Mobile iOS & Android

Примеры



АЛГОРИТМЫ

Active PC Liveness

Описание

Алгоритм основан на получении пространственной информации за счет **кооперативного** взаимодействия (повороты головы, улыбка, моргание, движения бровями) с приложением.

Преимущества: устойчив к самым популярным атакам с бумаги и с телефонов
Недостатки: Слабоустойчив к 3D маскам под конкретного человека

Рекомендации по использованию:
Использовать вместе с пассивным Liveness

Совместимость

Web-камеры, IP-камеры с качеством не ниже 3.7MP, 720p (иначе шумы убивают высокочастотные признаки)

Сценарии применения

Двухфакторная аутентификация и иные кооперативные сценарии

Оценка устойчивости к атакам

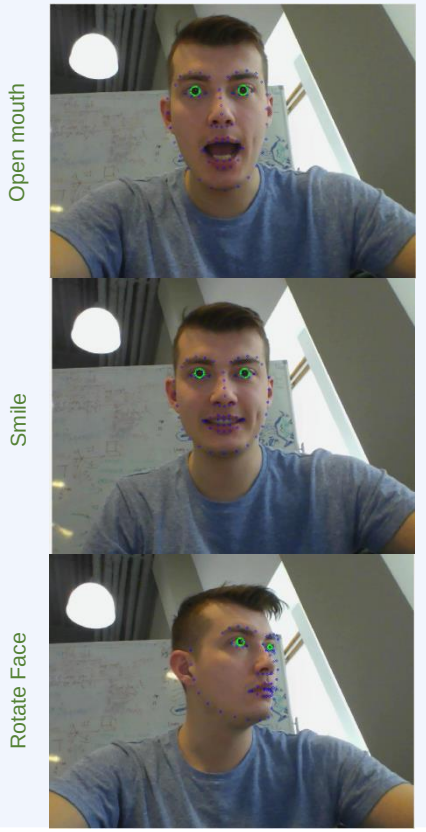
A B C

★ ★ ☆

Наличие в продуктах

- SDK Windows & Linux

Примеры



AGSR
АГЕНТСТВО СЕРВИСИЗАЦИИ И РЕИНЖИНИРИНГА

Виды атак уровня А*

КОД	НАЗВАНИЕ	ОПИСАНИЕ
P1	Printed square paper with large face	Распечатанное изображение (например, на А4) без модификаций. Края бумаги видны в кадре.
P2	Printed paper with full body, camera don't capture margins	Распечатанное изображение (например, на А4) без модификаций. Края бумаги не видны в кадре. То есть фейк поднесем близко к камере.
P3	Printed flat head mask	Распечатанное изображение лица, вырезанное по контуру. Без изменений на изображении лица, как в P4.
P1/P2/P3 ext	Series of P1/P2/P3 with different cooperative movements	Несколько артефактов из P1/P2/P3 - нейтральное лицо, улыбающиеся лицо и т.д.
P1/P2/P3 bw	P1/P2/P3 printed in BW mode	Артефакт из P1/P2/P3, распечатанный в ч/б
P4	Printed flat mask with cropped eye etc.	Распечатанное изображение лица, вырезанное по контуру + вырезанные глаза или рот для выполнения активных выражений, как моргнуть или улыбнуться.
D1	Phone Screen	Статическое изображения лица (как из социальных сетей), выведенное на экран телефона. Границы телефона находятся в кадре.
D2	Tablet/PC Screen, camera don't campure margins	Статическое изображения лица (как из социальных сетей), выведенное на экран планшета / ноутбука / монитора. Границы экрана не видны в кадре.
D1/D2 ext	Series of D1/D2	Несколько кадров из D1/D2 с фиксированными кооперативными выражениями лица
D3	Random video of target on device	Случайное видео с атакуемым (что можно добыть в соц. сетях) с движениями губ, либо с морганием, либо с сменой выражения лица.

ПРИЛОЖЕНИЕ

Виды атак уровня B*

КОД	НАЗВАНИЕ	ОПИСАНИЕ
P1/P2/P3 ir	P1/P2/P3 printed in IR mode	Артефакт из P1/P2/P3, распечатанный в IR
P5	Partial paper mask	Вырезанный бумажный участок лица, для накладывание на реальное лицо, например, область вокруг глаз + щеки. Перед тестированием подобной атаки против liveness необходимо убедиться в том, что этой информации достаточно для распознавания лиц с ожидаемым ответом. Для выделения областей, которые влияют на распознавание, необходима экспертиза.
D4	Video of target with defined movement (blink/headpose/...)	Видео с атакуемым повторяющего близкие движения к тому, что требуется в системе, например, кивок.
M1	Papercraft mask	Объемная бумажная маска атакуемого. Для изготовления необходима экспертиза.

ПРИЛОЖЕНИЕ

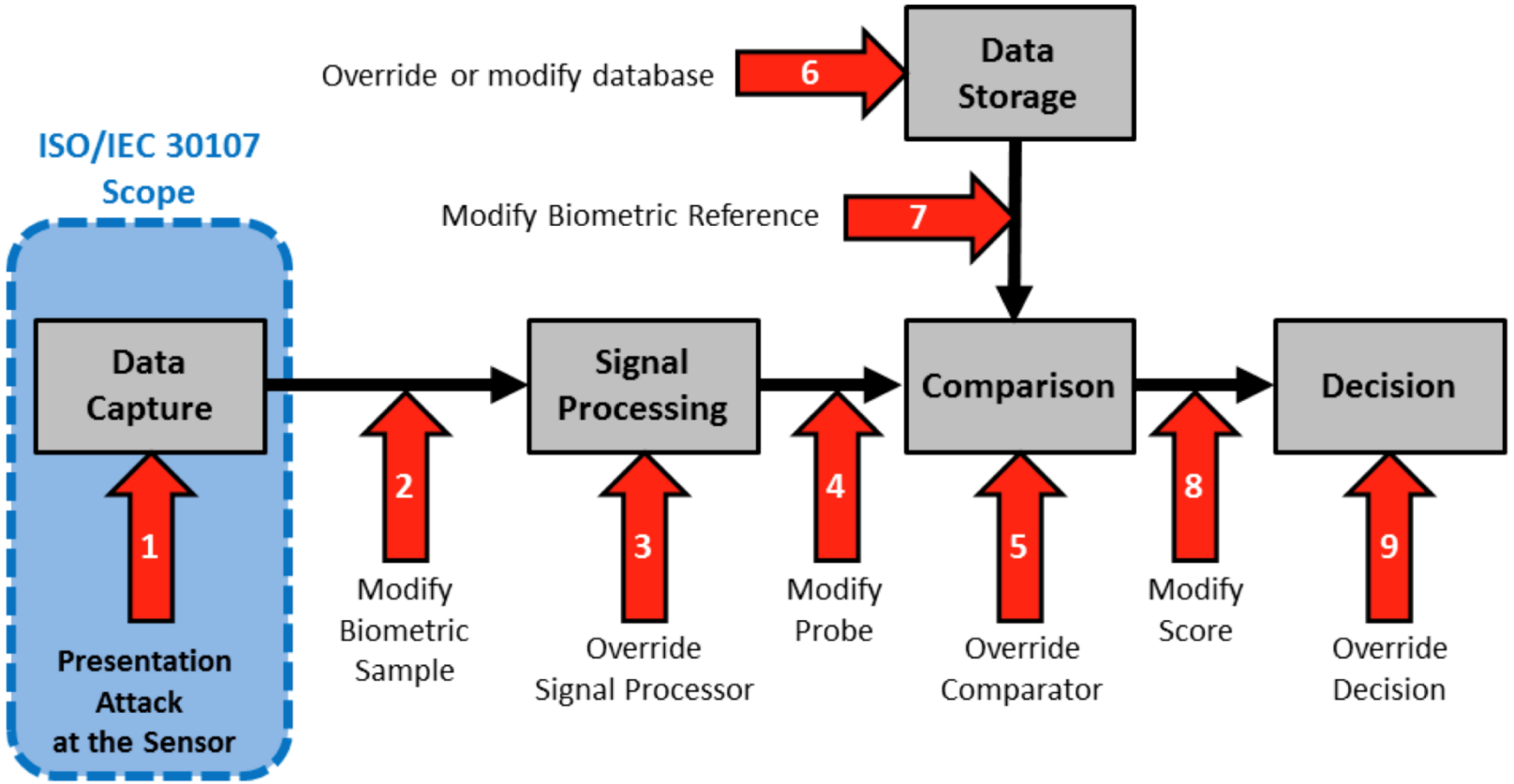
AGSR
АГЕНТСТВО СЕРВИСИЗАЦИИ И РЕИНЖИНИРИНГА

Виды атак уровня C*

КОД	НАЗВАНИЕ	ОПИСАНИЕ
D5	Skype with target	Видеозвонок с попыткой удаленного доступа
M2	Silicon simple mask	Точная силиконовая маска атакуемого без возможности выполнения активного действия (вырезанные глаза, рот и так далее).
M3	Silicon mask with cropped eyes/mouth	Точная силиконовая маска атакуемого с возможностью выполнения активного действия.
M4	Ceramic mask	Точная керамическая маска без возможности выполнения активного действия.

ПРИЛОЖЕНИЕ

Pipeline (ISO/IEC 30107)



Source: ISO/IEC 30107-1
Inspired by N.K. Ratha, J.H. Connell, R.M. Bolle, “Enhancing security and privacy in biometrics-based authentication systems,” IBM Systems Journal, Vol 40. NO 3, 2001.