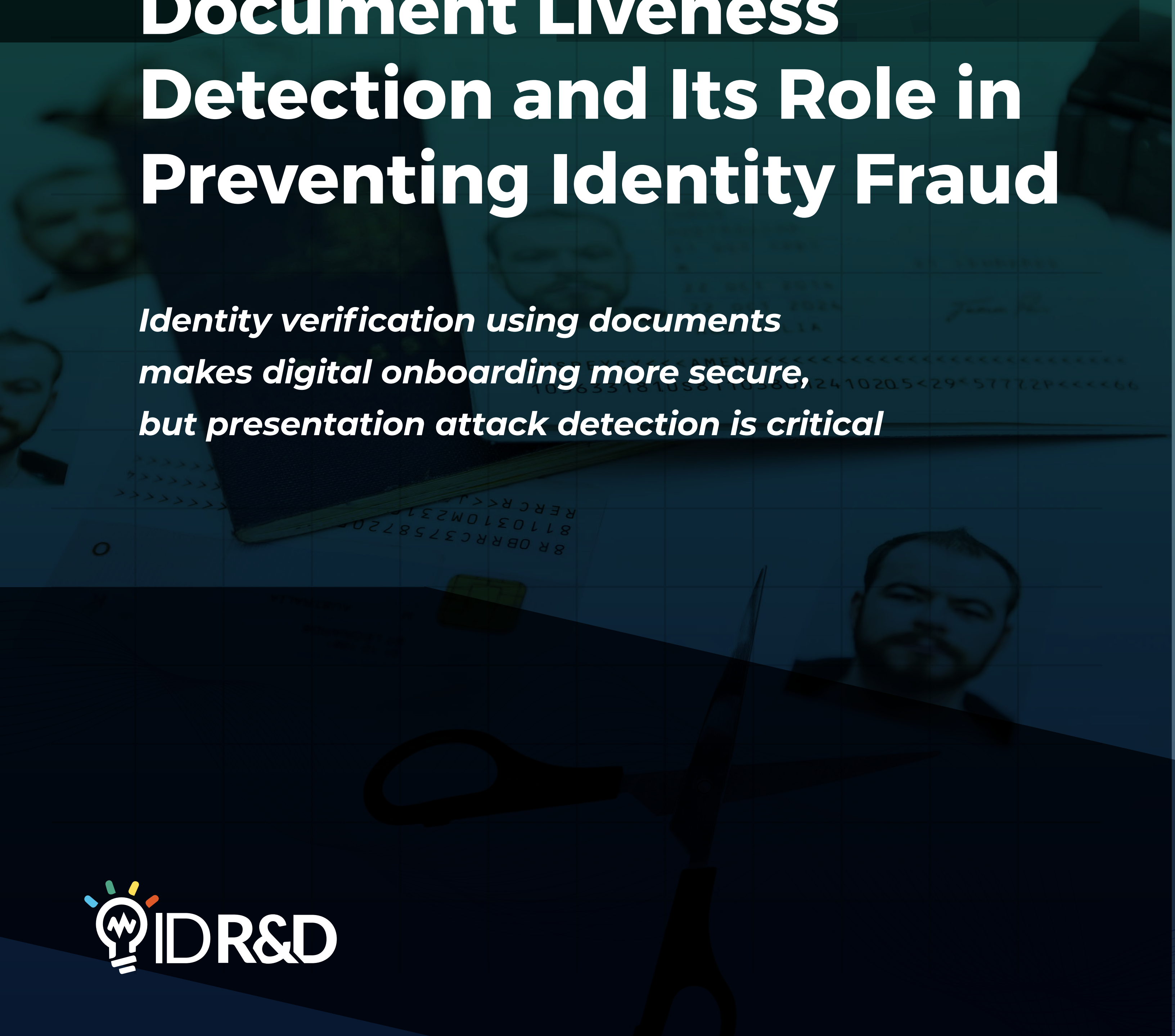
WHITE PAPER

Document Liveness Detection and Its Role in Preventing Identity Fraud

Identity verification using documents makes digital onboarding more secure, but presentation attack detection is critical



ID R&D

Digital identity verification can vastly improve onboarding efficiency, increase customer growth rates, and provide convenient, equitable access to services for new customers. Leveraging government-issued identity documents for the process is extremely effective, if not essential to making it secure. But an unsupervised process creates vulnerabilities to fraud, where bad actors can use counterfeits, forgeries, and reproductions to misrepresent their identity.

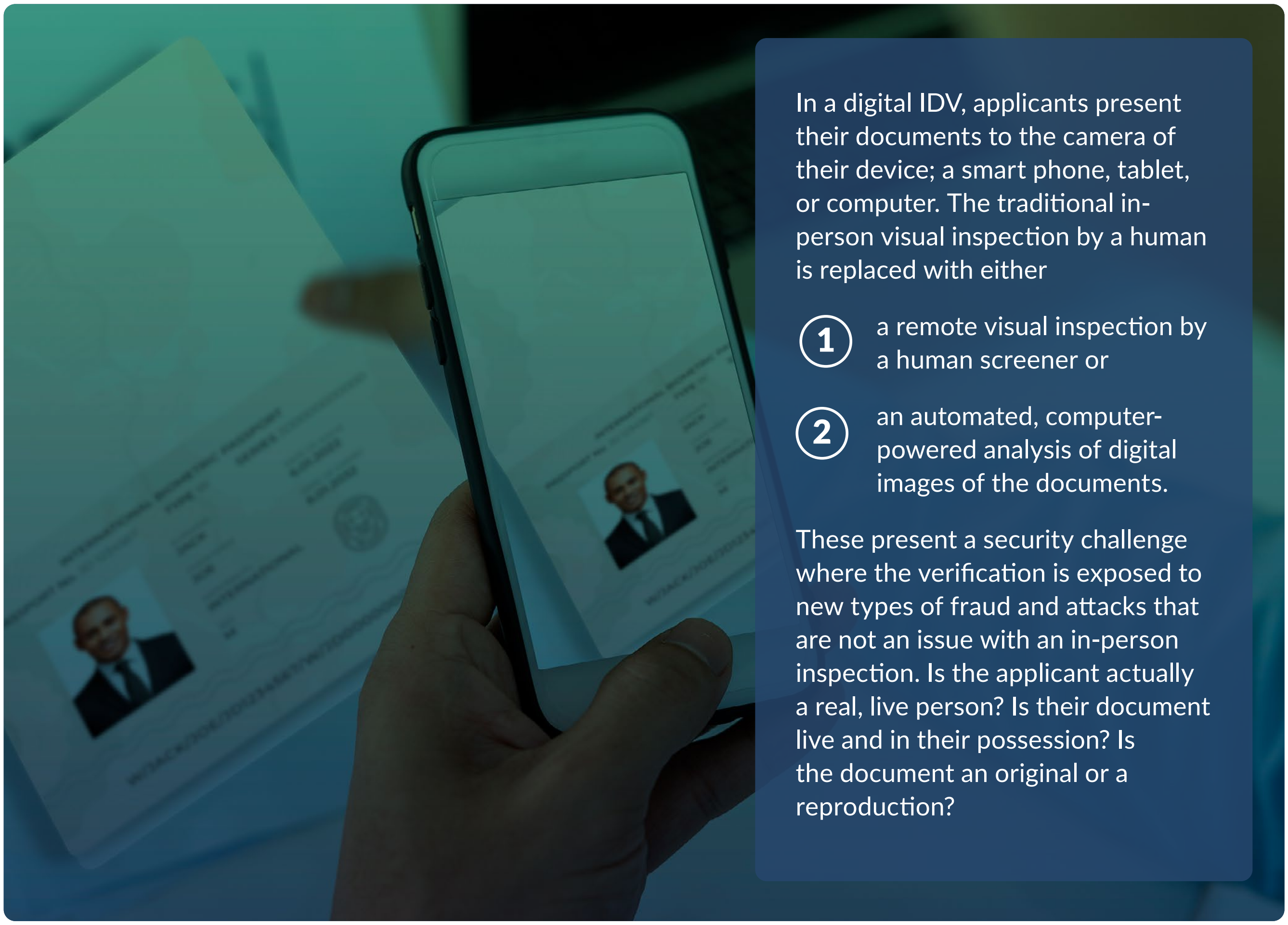
Presentation attacks are when a fraudster uses an identity document comprised either fully or partially of a reproduction. A reproduction can be physically printed or digitally displayed on a screen. The three most prevalent presentation attack methods are screen replays, printed copies, and portrait substitutions. Recent ID R&D research shared herein shows that presentation attacks make up 90% of document-based attacks, with screen replays making up a majority of this type of attack.

Where documents are used for digital onboarding, It is essential to use accurate and effective document liveness detection.It also should be easy to implement into legacy systems and easy to use; applicants should not be impeded by friction in the form of complex instructions, failed image captures, and false-positives.

Introduction - the value of digital identity verification

Identity verification (IDV) is the process of confirming the claimed identity of an individual in order to establish their eligibility for access to a service or resource, and is essential to a secure onboarding process. Organizations need to protect onboarding from bad actors looking to establish fraudulent accounts or to illegally access an account. Digital IDV makes is possible to perform onboarding remotely ("digitally") without supervision by using the applicant's device. Avoiding an in-person visit makes onboarding more convenient, particularly for applicants with less ability to participate in person. The reduced friction that digital IDV enables can increase customer growth rates while making services more accessible and equitable.

Government-issued identity documents such as passports and driver's licenses are clearly a fundamental tool for verifying identity. They leverage the substantial investment made in identity proofing that is performed as part of their issuance. They can and should be used for digital IDV, but precautions are needed to avoid allowing bad actors through the process.



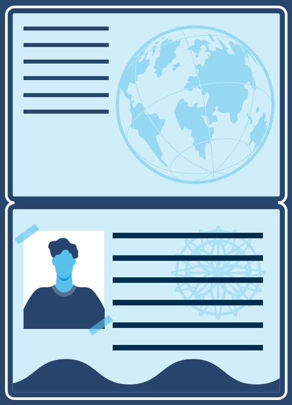
In a digital IDV, applicants present their documents to the camera of their device; a smart phone, tablet, or computer. The traditional in-person visual inspection by a human is replaced with either

1. a remote visual inspection by a human screener or
2. an automated, computer-powered analysis of digital images of the documents.

These present a security challenge where the verification is exposed to new types of fraud and attacks that are not an issue with an in-person inspection. Is the applicant actually a real, live person? Is their document live and in their possession? Is the document an original or a reproduction?

Two categories of fraudulent documents: counterfeits and forgeries

With document-based identity verification in place, bad actors attempting to commit fraud need to present fraudulent or stolen documents to conceal their true identity. These fraudulent documents fall into two categories: counterfeits and forgeries. Following are descriptions and examples.

Types of Fraudulent Identity Documents	Examples
COUNTERFEIT A physical or digital fabrication of a document designed to resemble either a genuine document or an invented document 	• A New York driver's license physically fabricated to resemble a genuine New York driver's license • A digitally created and displayed or printed US passport based on online examples • A passport issued by Utopia • A New York driver's license that is of an invented design
FORGERY A genuine document that has been physically altered or tampered, or a reproduction of a forged document 	• A US passport with the facial image physically altered, such as with an photo overlay • A photocopy of a tampered New York driver's license, cut to proper size and shape

There are a wide variety of alterations that a fraudster can make to documents, including the following:

Text and font manipulation. This involves altering the text or font of a document in order to change or falsify the information contained in it.

Signature falsification. The signature on a document is altered.

Portrait substitution. A type of document forgery in which the photograph or image on a document is replaced with a different image.

Reproductions as a form of fraud: photocopies, scans, and screen replays

Every identity document contains unique information and is designed to be difficult to reproduce in order to prevent copies being used by people other than those to whom they were issued. Photocopies and scans are physical reproductions created using a copier or scanner, respectively. A screen replay is a digital reproduction using a device with a digital display, such as a smart phone, tablet, or computer monitor.

A reproduction of a genuine, untampered document is not considered fraudulent in and of itself when in the possession of its rightful owner; it's actually a good practice to travel with a copy of a genuine passport to use in the event that it is lost or stolen. Using a reproduction of a counterfeit or forged document can make those fraudulent features more difficult to detect. This is why prohibiting and detecting the attempted use of reproductions in an unsupervised IDV process is essential to preventing fraud.

Document presentation attacks: identity fraud in action

A document presentation attack is when a bad actor tries to misrepresent their identity by present identity documents that are either fully or partially reproduced. The reproductions can be either physical (full or partial) or digital, and can include counterfeits or forgeries.

ID R&D research indicates that 90% of document-based attacks are presentation attacks (see figure).

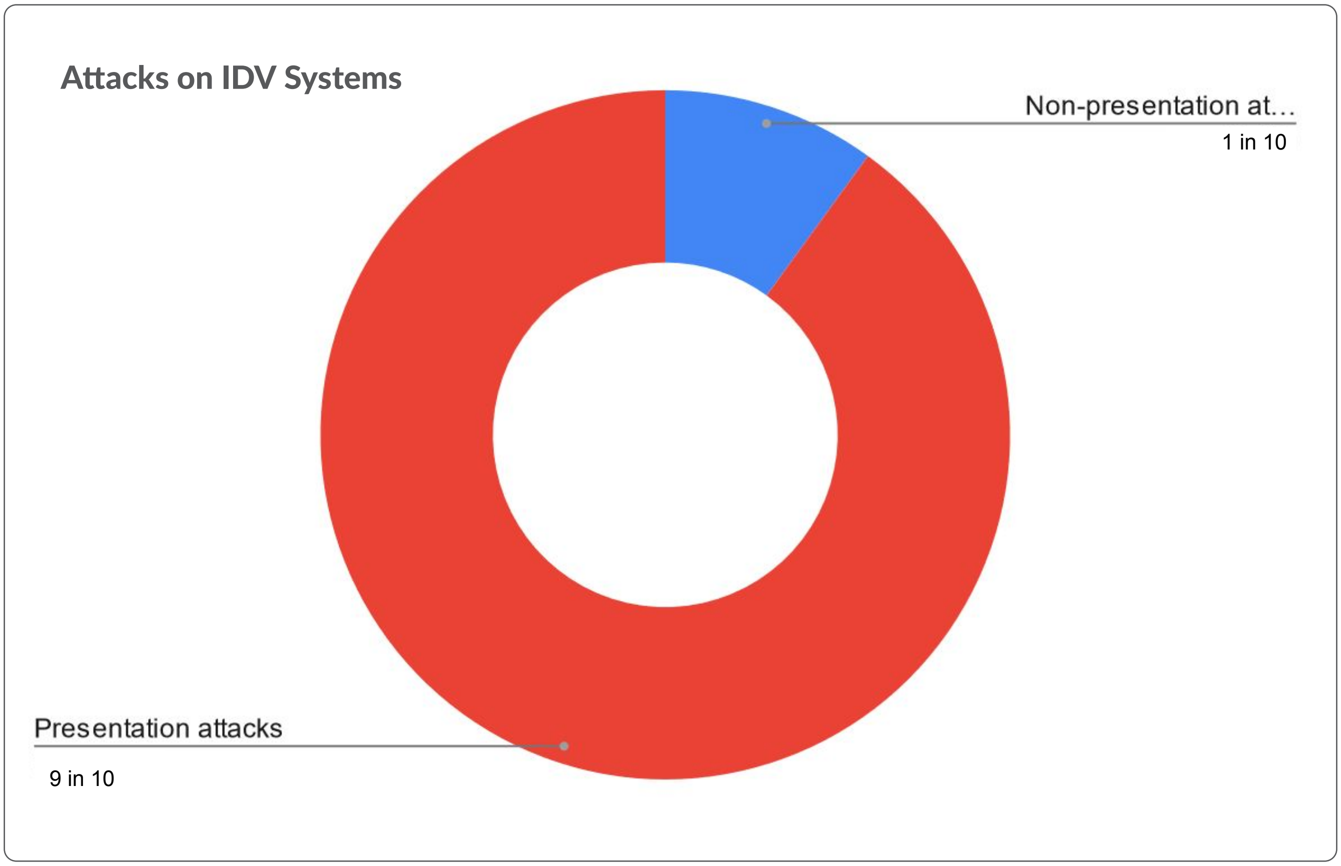


Figure: A large majority of document-based attacks on digital IDV systems are presentation attacks

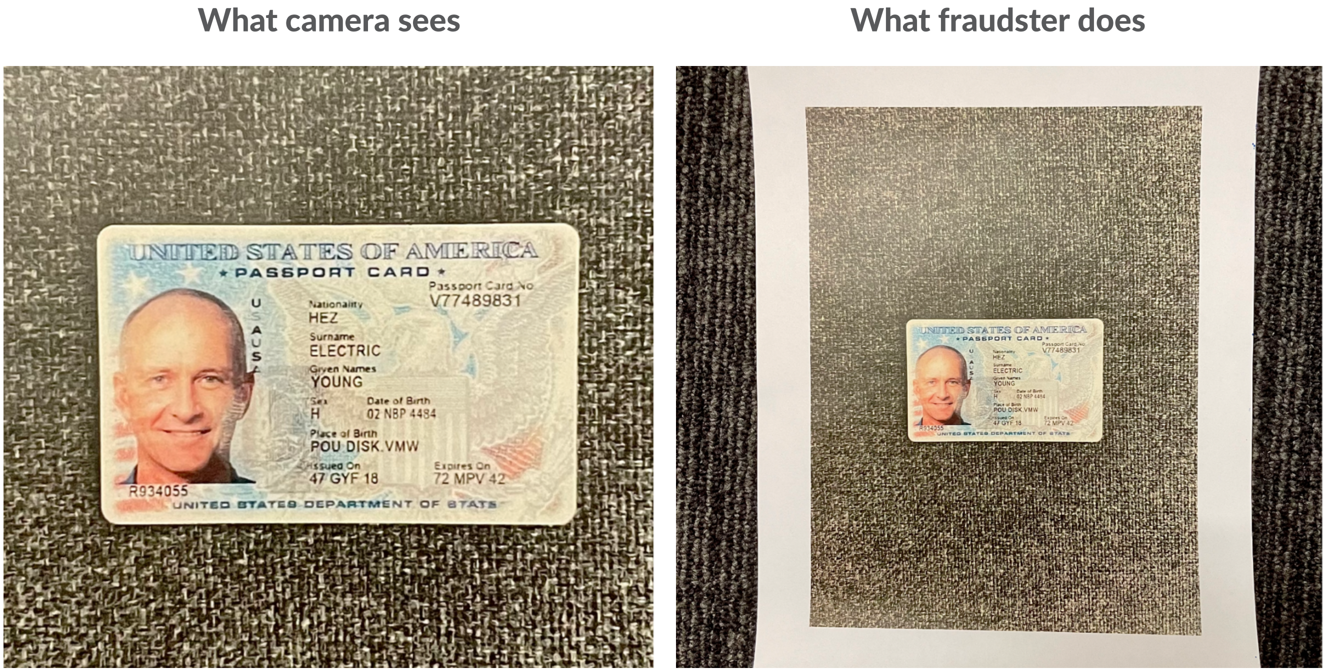
The majority of attack vectors are presentation attacks. While most KYC platforms focus on confirming the authenticity of the document (correctness of structure, text, bar codes), presentation attacks are usually not detected.

Following are examples of attack methods that use counterfeits and forgeries in conjunction with full or partial reproductions to conduct presentation attacks:

PRINTED COPY

The attacker presents a physical copy of the document; a colored printed copy or black-and-white Xerox copy of a document in front of the camera. The source of a document could be a digital image found on the internet. The attacker can use different techniques to bypass the liveness detection (in order from simple to hard):

- Printed cutouts without lamination
- Printed cutouts with lamination imitation
- Laminated cutouts
- Plastic printed forgery


What camera sees

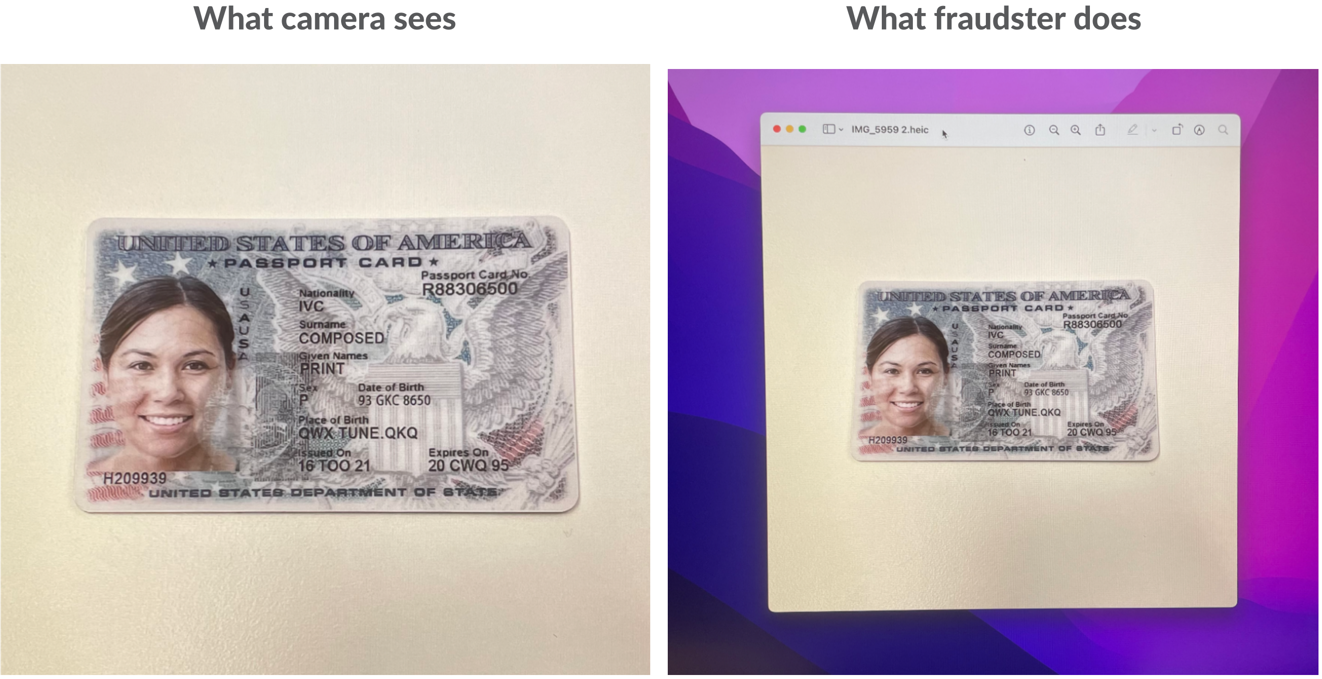
What fraudster does

Figure: Printed copy

Presentation attacks include an attack when the original document is genuine, but the user presents a reproduction of athe document instead of the original. An example is a fraudster copying a digital image of an ID availble on the dark web and using it to misrepresent their identity. In this case, the security checks of document structure, text, and bar codes will likely pass successfully and the copy will go undetected.

SCREEN REPLAY

The attacker presents a digital reproduction of a counterfeit, forged, or stolen document using a device with a screen, such as a laptop, tablet, smartphone. This type of attack refers to the use of a pre-recorded video or image of a document in order to bypass identity verification systems. The attacker records the video of the legitimate or generated document and then demonstrates it from a display of smartphone/laptop or another device during the liveness check in order to appear as if the document is physically presented in front of the camera.



What camera sees

What fraudster does

Figure: Screen replay

PORTRAIT SUBSTITUTION (PORTRAIT OVERLAY)

The face image on a document is replaced with a different image in an attempt to deceive someone into believing that the document is genuine. One way that a portrait substitution attack could be organized is by physically altering the document. This could involve cutting out the original face image and pasting in a new one, or using other methods to alter the appearance of the document. This type of attack may be more difficult to carry out, as it requires access to the physical document.

Figure: Portrait substitution

The following figure shows the breakdown of document-based attacks on IDV systems according to ID R&D data. Note that a significant majority of document-based attacks on IDV systems are conducted using screen replays, making these attacks particularly important to detect.

Sample Case Study: Breakdown of attacks on IDV Systems

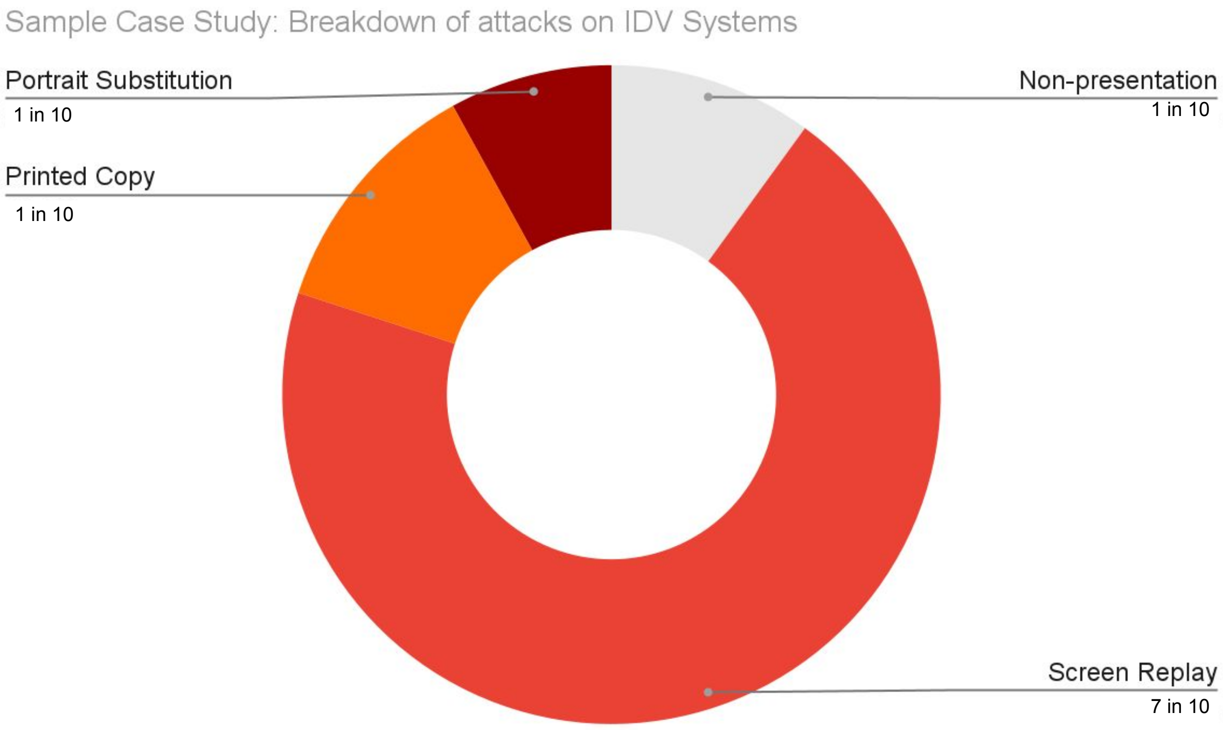


Figure: Typical Use Case: Breakdown of document attacks on IDV system

Mitigating the risk: document liveness detection

Document liveness detection products such as IDLive Doc from ID R&D are designed to detect and prevent presentation attacks including screen replays, printed copies, and other kinds of attacks as described herein. Following are features that are desirable in a solution.

Ability to detect the presentation of printed copies, screen replays, digital snapshots, and portrait substitutions

Detect attacks in near real-time without adding friction to the user experience or tipping off fraudsters

Stop fraud that humans cannot detect

Work universally across all types of identity documents from around the globe

Works independently of user experience

Integrate into existing Know Your Customer (KYC) processes without disrupting other core components

Can be added to any existing remote onboarding system in less than a day.

Windows or Linux server compatible