情報セキュリティ・セミナー
「生体認証におけるディープフェイクの脅威と対策」

HITACHI
Inspire the Next

ディープフェイクの脅威およびeKYCへの攻撃と対策

2022/11/22
日立製作所　研究開発グループ
川名のん

0. 目次

HITACHI
Inspire the Next

1. Deepfakeの概要
2. eKYCへの攻撃実験の内容
3. 実験結果と対策
4. まとめ

1. Deepfakeの概要

HITACHI
Inspire the Next

1. Deepfakeの概要
2. eKYCへの攻撃実験の内容
3. 実験結果と対策
4. まとめ

1-1. Deepfakeとは

DeepfakeとはDeep learningとFakeをあわせた造語で、機械学習を用いて人物画像を合成する技術の総称である。

Deepfake技術は大きく分けて4種類ある。
・Face Synthesis：何もないところから架空の顔を生成する

・Identity Swap：画像内の人物の顔(目や口などのパーツ)を別の人物のものに入れ替える

・Attribute Manipulation：顔の特徴を変更する。髪型の変更や、装飾品の追加、皺の削除など

・Expression Swap：画像内の人物の表情を別の人物の表情にする

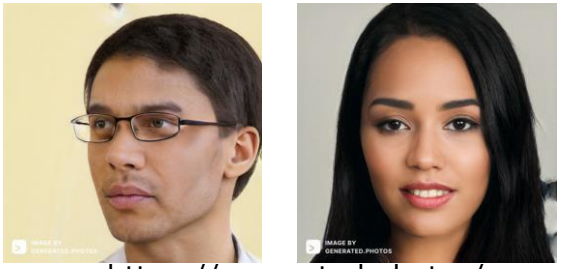
https://arxiv.org/pdf/2001.00179.pdf

1-2. Deepfakeツール

HITACHI
Inspire the Next

DeepfakeのアルゴリズムやソースコードはGitHubなどに公開されており、
誰でもその技術やノウハウを習得できる。
※2022年5月 Google ColabでのDeepfake作成を禁止に

・GENERATED PHOTOS
自動で存在しない人物の顔を生成できるツール
ブラウザで閲覧可能



https://generated.photos/

・FaceSwap(https://faceswap.dev/)
動画内の人物の顔を他の人物に変えることができるOSS
かなりの高スペックPCが必要

・FaceApp
顔を幼くしたり、性別を変えたりできるスマホアプリ
数年前にSNSで流行った

1-3. Deepfakeの用途

HITACHI
Inspire the Next

本来の利用方法

・美容手術の術後イメージ画像
・映画撮影での代役

悪質な利用方法

・著名人の顔を用いた動画作成による人権侵害、名誉毀損
・著名人になりすました政治利用や詐欺
　ーオバマ前大統領がトランプ政権を批判する(2018/04)
https://www.youtube.com/watch?v=cQ54GDm1eL0

　ーゼレンスキー大統領が人々に降伏を呼びかける(2022/03)

　ー暗号通貨取引所のCCOになりすまして上場詐欺(2022/08)
https://news.bitcoin.com/hackers-used-deepfake-of-binance-cco-to-perform-exchange-listing-scams/

1-4. Deepfakeに関する動向

HITACHI
Inspire the Next

■Deepfake検出コンテストが各地で開催

・MetaやMicrosoftなどがディープフェイクの検知ツール開発コンペ
「Deepfake Detection Challenge(DFDC)」を開催→成功率は最高で83%

https://antispoofing.org/Deepfake_Detection_Competitions

Deepfakeデータ：**Celeb-DF (v2): A New Dataset for DeepFake Forensics**

2. eKYCへの攻撃実験の内容

HITACHI
Inspire the Next

1. Deepfakeの概要
2. eKYCへの攻撃実験の内容
3. 実験結果と対策
4. まとめ

2-1. KYCについて

HITACHI
Inspire the Next

■KYC(Know Your Customer)とは

金融機関における口座開設やクレジットカードの作成の際に求められる、
本人確認の手続き(本人確認書類の金融機関への提示)のこと。
マネーロンダリングやテロ組織への資金流入を防ぎ、犯罪やテロ活動の防止を図るためにある。

https://www.rakuten-bank.co.jp/loan/cardloan/entry/application/

この手続きは、「犯罪による収益の移転防止に関する法律」(犯収法)で、本人特定事項の確認として定められている。
従来、オンラインの際は「本人限定受け取り郵便」での住所確認をしていたが、
オンラインでも本人確認が完結できるように、2018年に法改正が行われた。

2-1. KYCについて

HITACHI
Inspire the Next

■eKYC (electronic Know Your Customer)とは

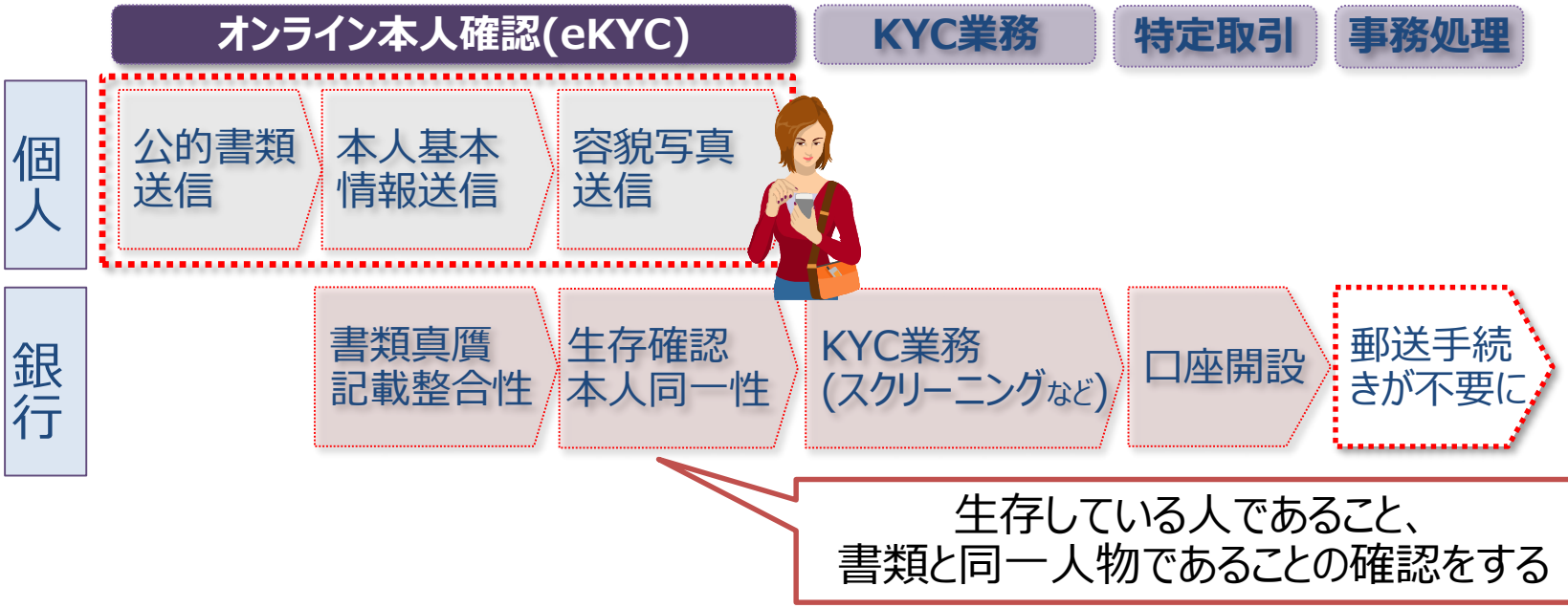
オンラインで完結する本人確認手続きのこと。

・セルフィー写真の撮影(容貌写真)

・事業者による生存確認と本人同一性確認(公的書類の顔写真)

を行う。

改正犯収法(6条1項1号ホ)によるオンライン確認機能(eKYC)

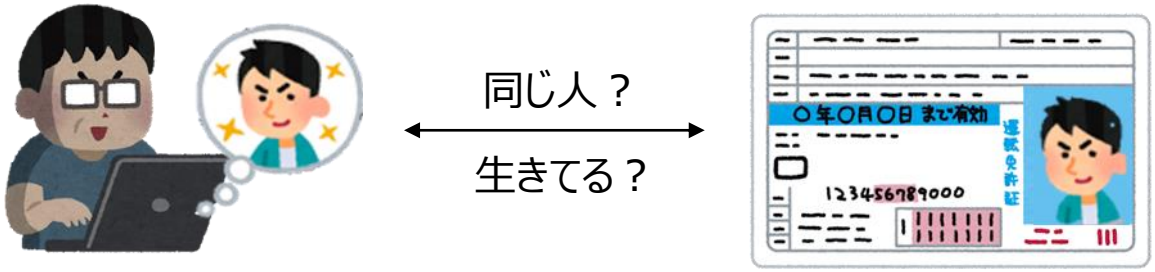


2-2. なりすまし攻撃の実験概要

HITACHI
Inspire the Next

■実験内容

他人の顔写真を用いてDeepfakeでなりすましをして、オンライン本人確認eKYCを突破できるか(本人確認書類と同じ人、生存している人と判定されるか)確かめる



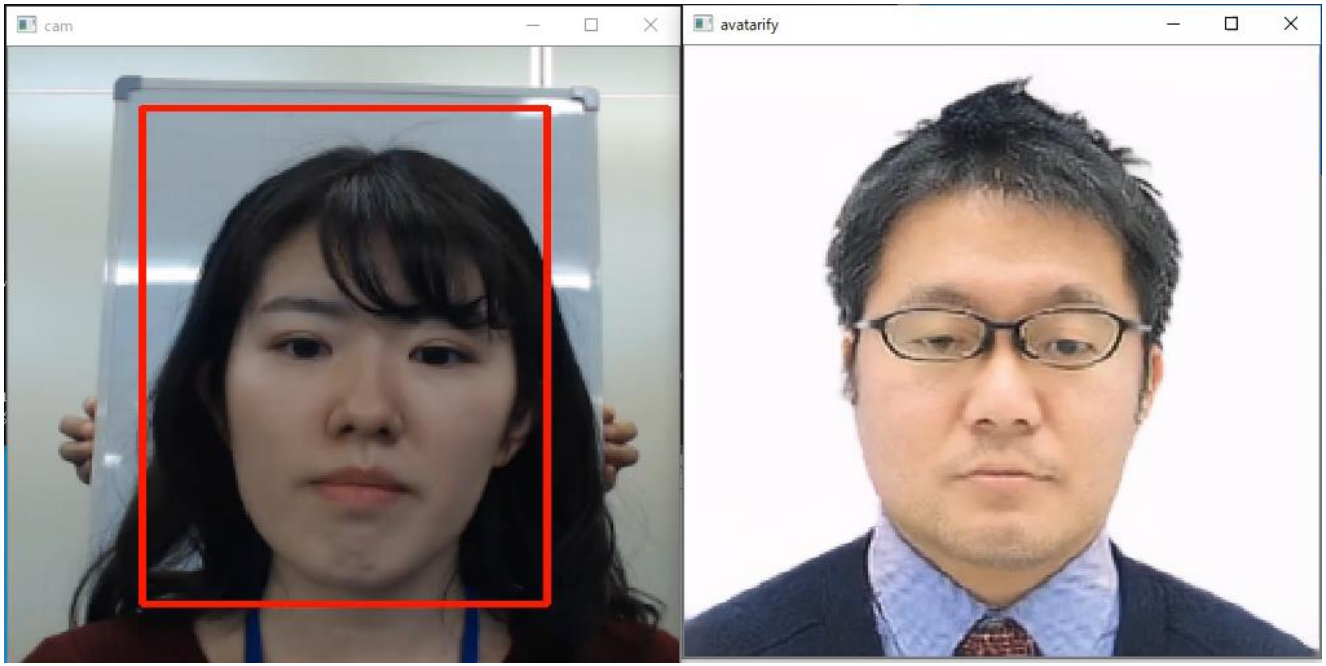
■実験環境

・Deepfakeツール：Avatarify
・グラフィックボード：GeForce RTX 2080 SUPER GAMING X TRIO, MSI
・webカメラ：C922n, ロジクール

・攻撃対象：独自に作成した模擬的なeKYCシステム

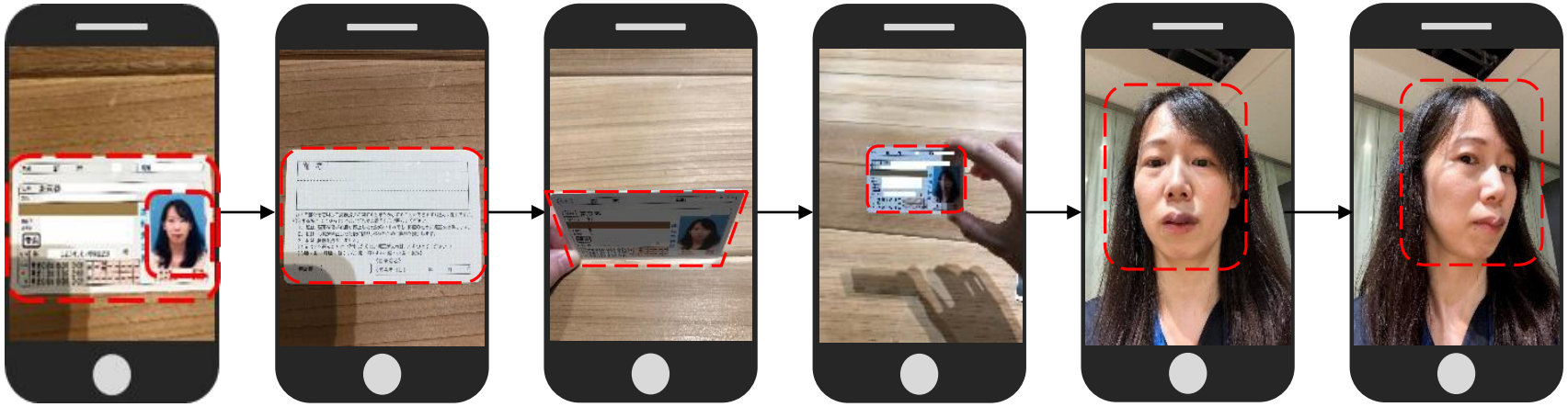
2-2. なりすまし攻撃の実験環境

■**Avatarifyとは** (https://github.com/alievk/avatarify)
リアルタイムでDeepfake動画を生成するOSS。
事前のトレーニングが必要なく、ターゲットの顔写真1枚からリアルタイムで表情の自動生成が可能。
今回は男性の顔写真1枚を用いてなりすましを行った。



webカメラで撮影している顔　　Avatarifyで生成している顔

2-3. 攻撃対象の独自システム



①運転免許証の表・裏・厚みの撮影

②正面撮影

③ランダムな指示に従って顔を動かして撮影

※顔が検出されたまま3秒待機
3秒後の顔が撮影される

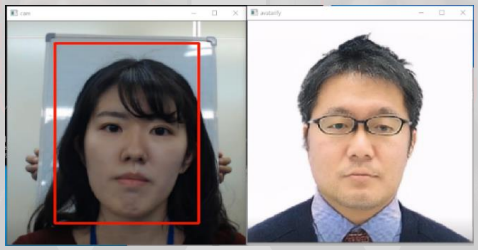
④運転免許証と撮影した顔が一致するか照合

2-4. なりすまし攻撃の実験方法

HITACHI
Inspire the Next

なりすましたい男性の運転免許証を使用

攻撃者が男性になりすまして撮影



①運転免許証の表・裏・厚みの撮影

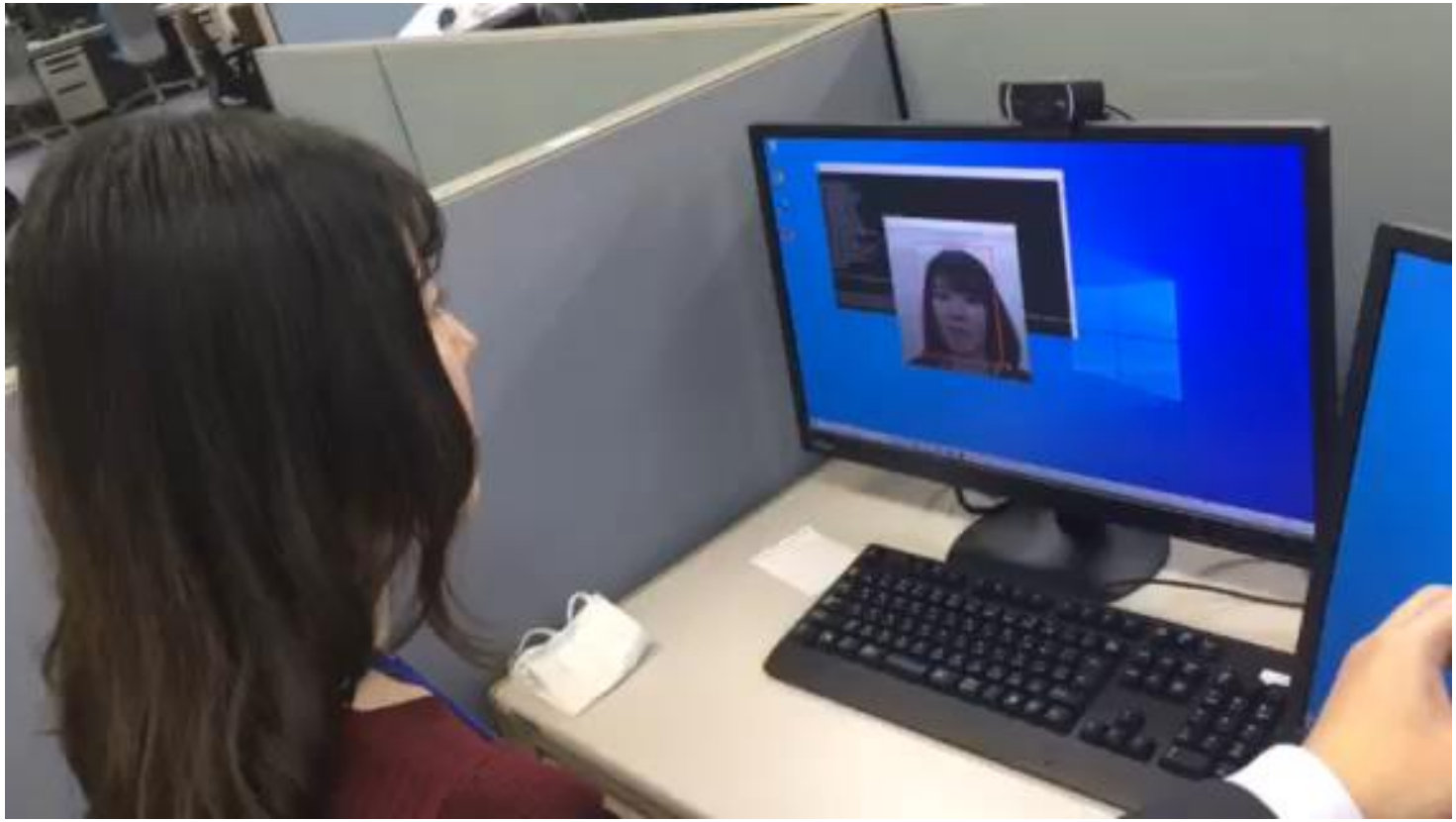
②正面撮影　③動かして撮影

※顔が検出されたまま3秒待機
3秒後の顔画像が撮影される

④運転免許証と撮影した顔が一致するか照合
→**一致したら攻撃成功**

2-4. 実験の様子

スマートフォンのインカメラで顔の正面の撮影をしている様子。
図の左側のモニターに表示している顔がWEBカメラでキャプチャしたもので、
右側のモニターに表示している顔がAvatarifyで生成されたものである。



3. 実験結果と対策

HITACHI
Inspire the Next

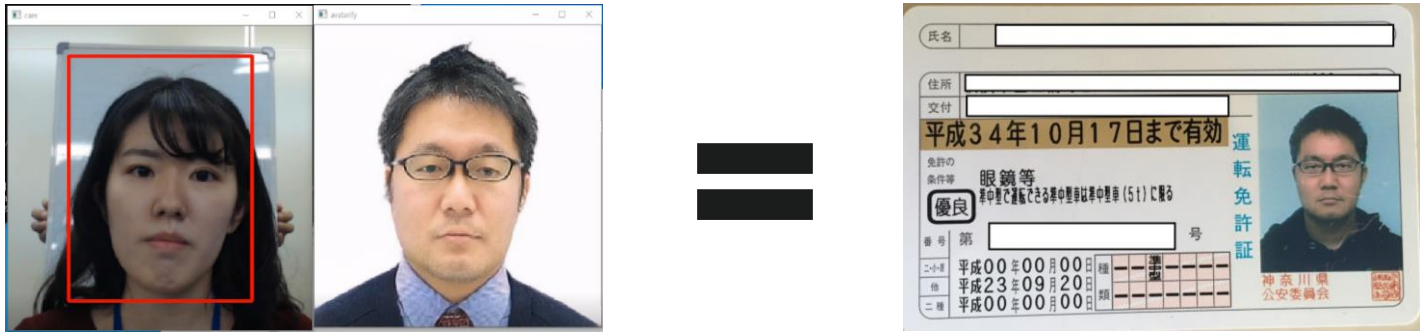
1. DeepfakeとE-KYCの概要
2. eKYCへの攻撃実験の内容
3. 実験結果と対策
4. まとめ

3-1. 実験結果

HITACHI
Inspire the Next

■実験結果

Deepfakeでなりすました顔が運転免許証と同一人物で、生存している人という結果が出力された。



これにより、Deepfakeによるe-KYCへのなりすまし攻撃、およびこの攻撃による不正な口座作成が現実的な脅威であることが判明した。

3-2. 対策技術案

HITACHI
Inspire the Next

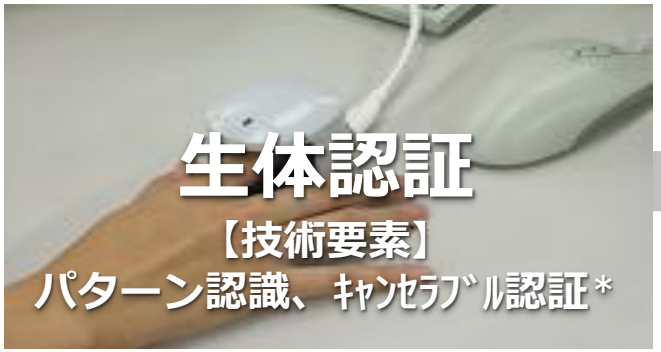
■eKYC向けDeepfake対策案

- Deepfake検知技術の実装
- スマホからフラッシュして、画面色変化に伴う顔の明度変化で検知
- 撮影時のモニター検知
- ランダムアクションの高度化
  （腕で顔を隠すなどのDeepfake範囲外のアクション）
- スマホ指静脈認証との連携（例：日立PBI認証技術）
- 顔画像の使いまわし、SIMカード、端末の使いまわしを検知

参考. 日立PBI技術概要

HITACHI
Inspire the Next

PBI（Public Biometrics Infrastructure）は、日立が世界で初めて実用化した、生体認証の技術とPKIの技術を融合した公開型生体認証基盤です。



＊2010年に日立が世界に先駆けて実用化した認証方式
（生体情報に特殊な暗号化を施し復号せず照合）

世界初日立特許取得

＊公開鍵暗号基盤
（公開鍵と秘密鍵のキーからなる公開鍵暗号技術）

公開型生体認証基盤（PBI）＝ 生体情報から「秘密鍵」を生成する公開鍵暗号基盤

PBIの特長

◆ **機微情報にあたる生体情報の管理が不要のため安全にセンタ保管が可能**
- 生体情報を直接利用せず、元の情報に復元できない形にして利用可能
- プライバシを高度に保護。破棄更新も可能

◆ **ICカードや暗証番号を使わず、便利で安全なPKIを実現**
-「忘れない」「無くさない」「手ぶら」で認証・署名が可能
- 生体情報に基づく確実な本人確認と、PKIの高セキュリティを両立

4. まとめ

HITACHI
Inspire the Next

1. DeepfakeとE-KYCの概要
2. eKYCへの攻撃実験の内容
3. 実験結果と対策
4. まとめ

4. まとめ

HITACHI
Inspire the Next

■Deepfakeとは

Deepfakeは人物画像を合成する技術の総称
さまざまな用途があるが悪用される危険性が存在する
実際に被害者や逮捕者が出ていて、サイバー攻撃にも利用されている

■Deepfakeの危険性

Deepfakeでなりすました顔と免許証の顔が同一人物で、生存している人
という結果が出た
→DeepfakeによるeKYCへのなりすまし攻撃が現実的な脅威である

■Deepfakeへの対策

・スマホ指静脈認証との連携
・ランダムアクションの高度化
・顔画像の使いまわし、SIMカード、端末の使いまわしを検知
・Deepfake検知技術の実装


HITACHI
Inspire the Next