様　式　Ｃ－１９、Ｆ－１９－１、Ｚ－１９　（共通）

科学研究費助成事業　　研究成果報告書

令和　５ 年　 ６ 月 ２２ 日現在

機関番号：１３８０１
研究種目：基盤研究(C)（一般）
研究期間：2018～2022
課題番号：１８Ｋ１１２９４

研究課題名（和文）未知の攻撃を検知するUnspoofable Biometricsの研究

研究課題名（英文）Research on Unspoofable Biometrics to detect unknown presentation attacks

研究代表者
大木　哲史（OHKI, Tetsushi）

静岡大学・情報学部・准教授

研究者番号：８０５３７４０７

交付決定額（研究期間全体）：（直接経費）　　3,300,000円

研究成果の概要（和文）：未知のなりすまし攻撃を検知し，あらゆるなりすまし攻撃に対して安全な生体認証方式，Unspoofable Biometricsの開発を目指した研究を実施した．この目的を達成するために，(1)未知のなりすまし攻撃対策，および(2)未知のなりすまし攻撃，の2つの観点から取り組みを進め，(1)に関して高精度かつ省モデルサイズのなりすまし検知が可能であることに成功した．さらに，(2) に関して，Adversarial Examples等を応用した攻撃手法の検討を行うことで，本手法の「未知の攻撃に対する安全性」を実験的な評価を実施した．

研究成果の学術的意義や社会的意義
なりすまし攻撃の脅威は日々多様化しており，とりわけ未知の攻撃に対する対応能力が求められているが，特定の偽造物の検知を目的とした従来型の検知手法は充分な防御策になり得ていない．本研究では，高次元の複雑な構造を持つ生体情報を確率分布でモデル化可能することで，幅広い未知の攻撃を検知可能な生体認証方式を実現・実証した．人工知能が普及するネットワーク社会では，端末利用者が本人自身である，という真正性を保証することは必須の要件となる.ゆえに，本研究のなりすまし不能性を保証に関する成果は，今後のAI社会インフラ構築に向けた大きな一歩となる．

研究成果の概要（英文）：We have conducted research aimed at developing Unspoofable Biometrics, a secure biometric verification method against all forms of spoofing attacks, including unknown ones. Our pursuit of this goal has been approached from two main perspectives: (1) measures against unknown spoofing attacks, and (2) dealing with unknown spoofing attacks themselves.

With respect to (1), we have succeeded in creating an efficient spoofing detection model with high accuracy while maintaining a minimal model size. Regarding the second aspect, we have considered different attack methods, including the use of model inversion attacks, which allowed us to experimentally evaluate the 'security against unknown attacks' of our method. Our research efforts represent a significant leap in strengthening the security measures of biometric verification systems.

研究分野：情報セキュリティ

キーワード：生体認証　バイオメトリクス　なりすまし　セキュリティ

様　式　Ｃ－１９、Ｆ－１９－１、Ｚ－１９（共通）

１．研究開始当初の背景

人工知能や暗号技術の発達に伴い，人がコンピュータに単純なタスクのみならず，その意思決定を委託する，という社会が近づいている．そのような社会の安心安全を保つために，委託行為が誰の意思で行われたかが保証され，かつその偽造や不当な改ざんを正しく検知する技術が必要である．この要求に応えるためには，利用端末自体に加え，端末利用者自身の真正性を保証する技術が必要である．このための技術として生体認証は必須の技術と言えるが，近年では本物に限りなく近いが本物ではないメディア（顔，音声，指紋等）の流通によるなりすましが生体認証への大きな脅威となっており，この対策なしに端末利用者の真正性を保証することはできないと言える．日々多様化するこれらの脅威に対しては，とりわけ未知の攻撃に対する対応能力が求められており，特定の偽造物を検知する従来型の検知手法は充分な防御策になり得ていない．本研究ではそのためのキー技術として，未知のなりすまし攻撃に対して安全な生体認証方式，Unspoofable Biometrics の研究を行う．

２．研究の目的

本研究では，なりすまし攻撃の種類を一切仮定しない，未知のなりすまし攻撃に対して安全な，Unspoofable Biometrics 方式を開発する．Unspoofable Biometrics は，なりすまし攻撃に対する事前の知識なしに構築が可能であり(可用性要件)かつ，未知のなりすまし攻撃に対する安全性を保証し(安全性要件)，さらに，既存の生体認証方式の認証精度に影響を与えない（有用性要件）新たななりすまし攻撃検知技術である．これにより，なりすまし攻撃の多様化に伴う攻撃検知の困難化という根本的な問題を解決し，学術面と実用面の両方で大きなインパクトを与える．従来行われてきた特定の偽造物に対するなりすまし攻撃検知対策は，「真の生体情報」か「なりすまし攻撃」かの二値分類を解く識別器を複数組み合わせることにより複数のなりすまし攻撃を検知可能としてきた．しかし，この手法には，(1)生体情報だけでなく対象のなりすまし攻撃に関しても大量のサンプルが必要，(2)事前に学習したなりすまし攻撃以外の攻撃が二値のどちらに分類されるかは保証されない，といった問題が存在する．そこで本研究では，従来の二値分類に基づく識別器の考えに対し，教師あり異常値検知の考えを導入し，かつ敵対的な攻撃に安全なメカニズムを応用することで，未知の攻撃を行う攻撃者に対しても安全ななりすまし攻撃検知を実現可能であることを実証する．

３．研究の方法

未知のなりすまし攻撃を検知し，あらゆるなりすまし攻撃に対して安全な生体認証方式，Unspoofable Biometrics の実現を目指した研究を行う．まず，（１）生体情報の取りうる分布が個人に依らず同じであり，かつセンサや生体情報取得環境に起因するノイズ（以下，環境ノイズ）の分布も特定の確率分布で表せる，という単純な設定で本技術を開発し，評価実験で有効性を示す．なお，実機検討も考慮し，モバイル端末に搭載されたセンサによる取得が容易である顔および音声を対象として検討を進める．この技術は，「生体情報モデルを求めるメカニズム」と「生体情報モデルと，入力された情報とを比較し，入力された情報がなりすまし攻撃であるかを判定するメカニズム」とに分けられる．ここで「生体情報モデル」とは，全人類の生体情報を表す１つの確率分布であり，より具体的には，任意の生体情報に対する尤度の期待値を最大化する確率分布を求めることが目標となる．実際には，全人類の生体情報を収集することは不可能であるが，ここでは敵対的学習を用いた生成型モデルの学習手法(Generative Adversarial Network)等を

応用する．敵対的学習では，攻撃側のニューラルネットがより効果的な攻撃方法を次々と獲得していくことによって，守備側のニューラルネットが未知の攻撃に対する安全性を獲得していく．この特性を応用し，限られた生体情報から未知の情報の判定に有効な真の生体情報モデルを効率的に推定し，推定したモデルに基づき未知の攻撃を判定するメカニズムについて検討を行う．判定メカニズムの検討にあたっては，未知のなりすまし攻撃を検知できる確率（安全性要件）と生体をなりすましと誤検知する確率（有用性要件）の 2 つの観点を評価指標とし，これらを最適化する判定メカニズムを求める．

４．研究成果

本研究期間においては，未知のなりすまし攻撃を検知し，あらゆるなりすまし攻撃に対して安全な生体認証方式，Unspoofable Biometrics の実現するための第一歩として，敵対的生成ネットワークを応用したなりすまし検知システムを提案し，手のひら画像認証システムを対象としてその有用性を検証した．

（１）手法の概要

生体認証に対するなりすまし検知（以下，PAD）のこれまでのアプローチでは，印刷写真の周波数スペクトル，顔の 3 次元性，あるいは動画像における動き特徴検出など，特定のなりすまし攻撃を想定した検知手法，あるいはそれらの組み合わせにより安全性を保証する PAD 特徴の抽出手法が議論されてきた．しかし，日々多様化しているなりすまし攻撃をすべて検出できる万能の PAD 特徴を事前に学習することは困難である．この問題に対して本研究では，図 1 に示すように，敵対的生成ネットワークを利用した異常検知アルゴリズムを応用することで，偽サンプルと真サンプルを区別する 1 クラスシステムのニューラルネットワークを構成する．これにより，高い精度かつ多様な攻撃に対する攻撃耐性を備えたなりすまし攻撃対策システムを構築することができる．

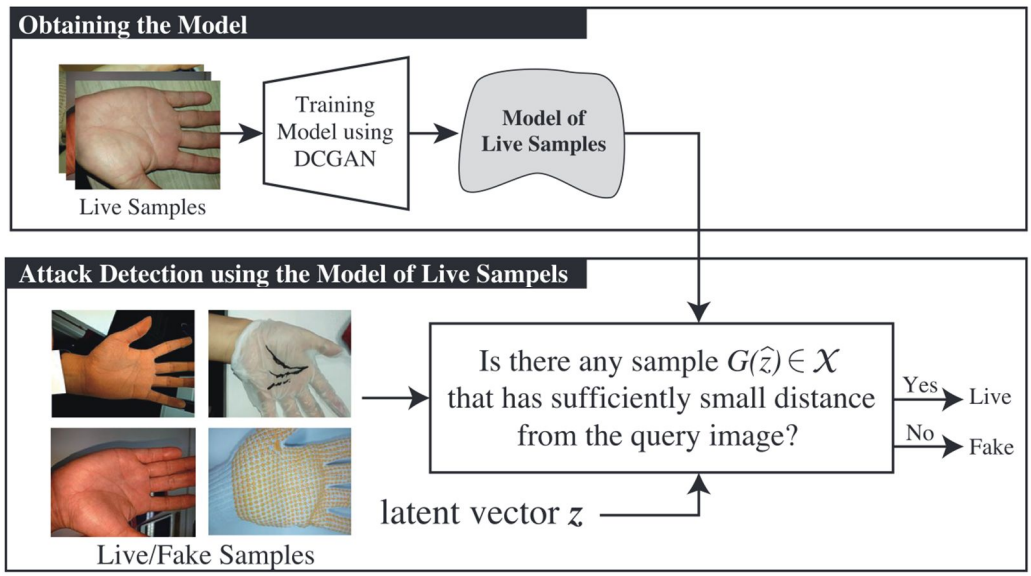


図 1：敵対的生成ネットワークを用いたなりすまし検知手法の概要

より具体的には，真の（偽画像を含まない）多数の手のひら画像のみを用いて教師なしの敵対的生成ネットワーク（Generative Adversarial Network：GAN）G を学習する．これにより画像空間上の手のひら画像多様体 X を学習することが可能となる（なお，ここでは画像に適した GAN である DCGAN を使用する）．G では，潜在空間 Z 上の潜在ベクトル z を入力することで，手のひら

多様体上の画像 $G(z) \in X$ を得ることが可能である．検知にあたっては，ランダムに選択した初期ベクトルｚに基づき得られる $G(z) \in X$ と入力画像 $\tilde{x}$ との損失を最小化するように誤差逆伝播を用いながら$\hat{z} \leftarrow z$と潜在ベクトルを更新していく．一定の回数α回のうちに，入力画像$\tilde{x}$との距離に十分近い画像 $G(\hat{z}) \in X$ が見つからなければ偽画像，見付かれば正規画像と判定する方式を構成した．

（２）実験方法

従来のなりすまし検知に関する研究では，Replay-Attack Database や Unconstrained Smartphone Spoof Attack (USSA) Database など，公開された生体／偽造物データセットを利用した研究が多く行われている．しかし，このデータベースには，特定の種類の偽の写真や動画サンプルしか含まれていないため，異常サンプルの点では不十分である．そこで，本実験では，手袋をはめた手のひら，ビニール手袋をはめた手のひらなど，手とは直接関係のない想定外の入力があった場合でも，システムが生と偽のサンプルを明確に区別可能なデータベースを独自に構築した．データベースは，10 種類の携帯カメラで約 2000 人から直接撮影した画像解像度 160x120 ピクセル，8748 点のライブサンプルおよび 6648 点の偽サンプル手のひら画像で構成された．撮影画像のサンプルを図 2 に示す．

本データセットを学習および評価セットに分割して実験を行った．学習セットは，ランダムに選択された 8000 の生体サンプルで構成し，評価セットは，7396 個のサンプルから構成し，そのうち 748 個を生体サンプル，6648 個を学習セットに含まれないケースからの偽造手のひらサンプルとした．

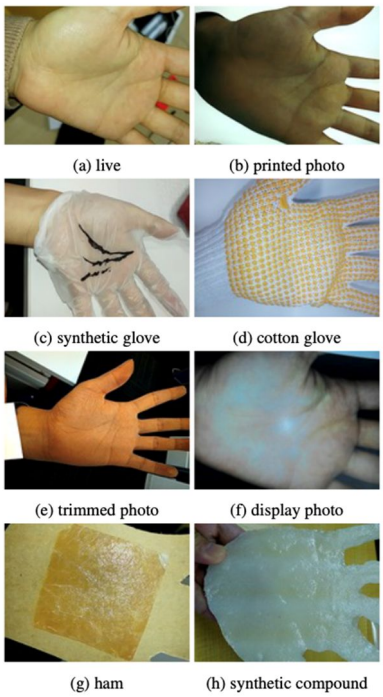


図 2：データセットに含まれる画像の例　(b)印刷した写真，(c)合成手袋をはめた手，(d)軍手をはめた手，(e)手の境界部分をカットし，2 次元的に手を再現した印刷写真，(f)iPad や Web カメラなどのデジタルデバイスから撮影した写真，(g，h)ゼラチンやハムなど手とは直接関係ないが皮膚を模した入力

（３）実験結果

図 3 に提案手法の性能評価結果を示す．横軸を False Positive Rate(FPR)，縦軸を True Positive Rate(TPR)とし，モデルの学習を 25 および 50 エポック実施した場合の結果を示している．ここで FPR は「未知のなりすまし攻撃に対する安全性を保証し(安全性要件)」に，TPR は「既存の生

体認証方式の認証精度に影響を与えない（有用性要件）」の評価結果にそれぞれ対応する．学習を進めることで AUC=96.8%と高い精度を達成できることを示した．また，表 1 に比較として代表的な機械学習による異常検知手法である One-Class SVM（OCSVM）との比較を実施した結果を示す．One-Class SVM については入力を無加工および Local Binary Pattern (LBP)画像の 2 種類としたもので評価を実施した．表 1 からもわかるように，いずれの評価においても提案手法の AUC が OCSVM の結果を上回っており，提案手法の優位性が確認できた．

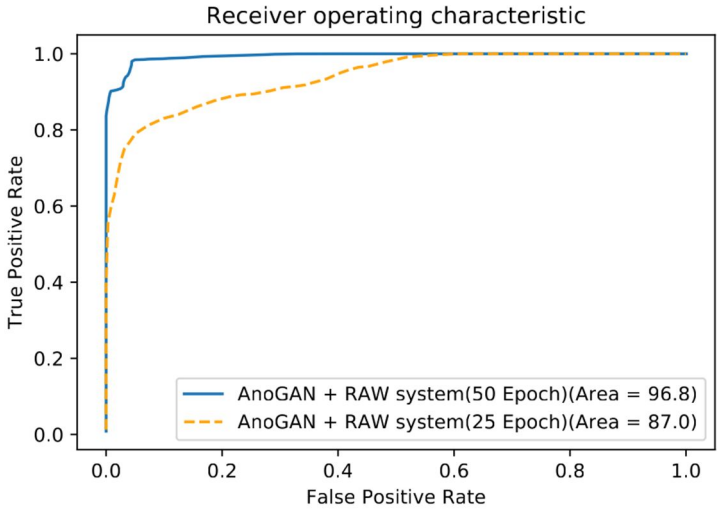
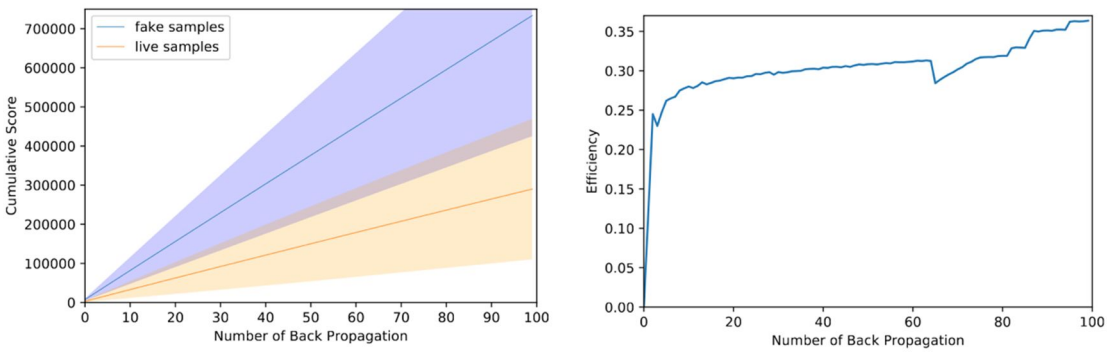


図 3：提案手法の性能評価

表 1: 提案手法と他の機械学習手法との比較

System	AUC(%)
Proposed (50 Epoch)	96.8
OCSVM + RAW	34.3
OCSVM + LBP	83.5

提案手法は誤差逆伝播を一定数繰り返すことで生体か偽造物であるかを判定する．このためα=100 等と設定した場合に判定にかかる時間が長くなることが懸念される．そこで，アルゴリズムの高速化を目的として，偽造サンプルを示す Anomaly Score の累積値が所定の値を超えた時点で入力を偽造サンプルと判定し，アルゴリズムを早期終了する手法を提案した．



(a)：50 エポック学習後の提案モデルの累積スコア計算結果　(b)：累積スコアに基づいた計算量評価結果

図 4：アルゴリズムの早期終了による効率化

図 4(a)は提案モデルの累積スコアの計算結果である．明らかに偽画像において累積スコアの伸びが速いことから，早期段階での終了が有効であることが直感的にもわかる．また，図 4(b)に横軸をαとし，縦軸を累積スコア導入により何割の計算量を削除できたかを示した（Efficiency=0.2 であれば 8 割の計算量で計算ができることを示す）．図 4(b)からわかるように，αが大きくなるほど累積スコアによる計算量削減の効果が高くなることがわかった．

５．主な発表論文等

〔雑誌論文〕　計4件（うち査読付論文　4件／うち国際共著　0件／うちオープンアクセス　0件）

１．著者名 塩見 祐哉、大内 結雲、藤田 真浩、眞野 勇人、大木 哲史、西垣 正勝	４．巻 62
２．論文標題 プレゼンテーション攻撃検知とQRコードの導入によるマイクロ爪認証の改良	５．発行年 2021年
３．雑誌名 情報処理学会論文誌	６．最初と最後の頁 1996～2010
掲載論文のDOI（デジタルオブジェクト識別子） 10.20729/00214242	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 －

１．著者名 郷間 愛美, 大木 哲史, 吉浦 裕, 市野 将嗣	４．巻 61(12)
２．論文標題 スマートフォンの通話に着目した音声と耳介による個人照合	５．発行年 2020年
３．雑誌名 情報処理学会論文誌	６．最初と最後の頁 1881-1891
掲載論文のDOI（デジタルオブジェクト識別子） なし	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 －

１．著者名 Vishu Gupta, Masakatsu Nishigaki, Tetsushi Ohki	４．巻 11
２．論文標題 Unsupervised Biometric Anti-spoofing using Generative Adversarial Networks	５．発行年 2019年
３．雑誌名 International Journal of Informatics Society	６．最初と最後の頁 45-53
掲載論文のDOI（デジタルオブジェクト識別子） なし	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 －

１．著者名 高橋洋介，遠藤将，松野宏昭，村松弘明，大木哲史，西垣正勝	４．巻 60
２．論文標題 眼球-頭部協調運動における生体反射型反応に基づく生体認証方式に関する検討	５．発行年 2019年
３．雑誌名 情報処理学会論文誌	６．最初と最後の頁 1-12
掲載論文のDOI（デジタルオブジェクト識別子） なし	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 －

〔学会発表〕　計17件（うち招待講演　4件／うち国際学会　6件）

１．発表者名
Vo Ngoc Khoi Nguyen, Takamichi Terada, Masakatsu Nishigaki, Tetsushi Ohki

２．発表標題
Examining of Shallow Autoencoder on Black-box Attack against Face Recognition

３．学会等名
APSIPA ASC 2021（国際学会）

４．発表年
2021年

１．発表者名
井田 天星, 竹内 廉, ヴォ ゴック コイ グエン, 西垣 正勝, 大木 哲史

２．発表標題
ブラックボックス型モデル反転攻撃におけるユーザ類似性を考慮した生成モデルの検討

３．学会等名
暗号と情報セキュリティシンポジウム2022

４．発表年
2022年

１．発表者名
土屋 純, 西垣 正勝, 大木 哲史

２．発表標題
Deep Master Voiceによる話者照合システムへのウルフ攻撃可能性の検証

３．学会等名
電子情報通信学会BioX研究会

４．発表年
2022年

１．発表者名
Yumo Ouchi, Ryosuke Okudera, Yuya Shiomi, Kota Uehara, Ayaka Sugimoto, Tetsushi Ohki and Masakatsu Nishigaki

２．発表標題
Study on Possibility of Estimating Smartphone Inputs from Tap Sounds

３．学会等名
2020 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference（国際学会）

４．発表年
2020年

１．発表者名
Yuya Shiomi, Genki Sugimoto, Ayaka Sugimoto, Kota Uehara, Masahiro Fujita,Yuto Mano, Tetsushi Ohki and Masakatsu Nishigaki

２．発表標題
Micro Biometric Authentication Using Fingernail Surfaces: A Study of Practical Use

３．学会等名
Advanced Information Networking and Applications（国際学会）

４．発表年
2020年

１．発表者名
大内結雲，奥寺瞭介，塩見祐哉，大木哲史，西垣正勝

２．発表標題
スマートフォンのタップ音からの入力内容推測可能性に関する研究(その2)

３．学会等名
暗号と情報セキュリティシンポジウム2021

４．発表年
2021年

１．発表者名
藤垣成汰朗，成田惇，菅沼弥生，西垣正勝，大木哲史

２．発表標題
ディープフェイク画像からの個人再識別化に関する検討

３．学会等名
暗号と情報セキュリティシンポジウム2021

４．発表年
2021年

１．発表者名
塩見祐哉，大内結雲，奥寺瞭介，藤田真浩，眞野勇人，大木哲史，西垣正勝

２．発表標題
生体検知とQRコードの導入によるマイクロ爪認証の改良にむけての一検討

３．学会等名
コンピュータセキュリティシンポジウム2020

４．発表年
2020年

１．発表者名
Vo Ngoc Khoi Nguyen, 西垣 正勝, 大木 哲史

２．発表標題
生体認証を回避する物理的なAdversarial Exampleの検討

３．学会等名
第82回情報処理学会全国大会

４．発表年
2020年

１．発表者名
Tetsushi Ohki, Vishu Gupta, Masakatsu Nishigaki

２．発表標題
Efficient Spoofing Attack Detection against Unknown Sample using End-to-End Anomaly Detection

３．学会等名
2019 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC)（国際学会）

４．発表年
2019年

１．発表者名
Ayaka Sugimoto, Yuya Shiomi, Akira Baba, Norihiro Okui, Tetsushi Ohki, Yutaka Miyake, Masakatsu Nishigaki

２．発表標題
A Liveness Detection Method for Palmprint Authentication

３．学会等名
2020 International Conference on Intelligent Human Systems Integration: Integrating People and Intelligent Systems（国際学会）

４．発表年
2020年

１．発表者名
Vishu Gupta, Masakatsu Nishigaki, and Tetsushi Ohki

２．発表標題
Effective countermeasure against biometric spoofing attacks using unsupervised one-class learning

３．学会等名
International Workshop on Informatics（国際学会）

４．発表年
2019年

１．発表者名
大木 哲史

２．発表標題
話者照合システムの脆弱性とウルフなりすまし攻撃

３．学会等名
日本音響学会 秋季研究発表会（招待講演）

４．発表年
2018年

１．発表者名
大木 哲史

２．発表標題
生体認証に対するなりすまし攻撃とその対策

３．学会等名
情報処理学会 音声言語情報処理研究会（招待講演）

４．発表年
2018年

１．発表者名
大木 哲史

２．発表標題
目に見えるものが真実とは限らない：なりすましとの戦い

３．学会等名
情報処理学会 IPSJ-ONE（招待講演）

４．発表年
2019年

１．発表者名
Tetsushi Ohki

２．発表標題
Presentation Attacks and its Countermeasures on Biometric Authentication Systems

３．学会等名
BioPro A+ Workshop（招待講演）

４．発表年
2019年

１．発表者名 寺田崇倫，大木哲史，西垣正勝
２．発表標題 Proof of Human-work実現に向けたCAPTCHAの検討
３．学会等名 情報処理学会 第81回全国大会
４．発表年 2019年

〔図書〕　計0件

〔産業財産権〕

〔その他〕

―

６．研究組織

	氏名 （ローマ字氏名） （研究者番号）	所属研究機関・部局・職 （機関番号）	備考
研究分担者	西垣　正勝 (Nishigaki Masakatsu) (20283335)	静岡大学・情報学部・教授 (13801)	
研究分担者	大塚　玲 (Otsuka Akira) (50415650)	情報セキュリティ大学院大学・その他の研究科・教授 (32721)	

７．科研費を使用して開催した国際研究集会

〔国際研究集会〕　計0件

８．本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関