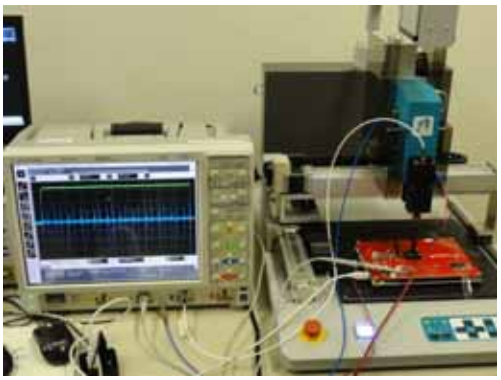
RITSUMEIKAN NTSOC LAB.

ネットワークLSIシステム研究室紹介



電子情報工学応用演習

2022.10.17

教授　藤野　毅

助教　吉田　康太

RITSUMEIKAN

ネットワークLSIシステム研究室紹介

RITSUMEIKAN NTSOC LAB.

- 設立：2003年4月
- 研究テーマ
  - 暗号回路やPUFの実装などのハードウェアセキュリティ技術を使ってネットワーク＆セキュリティーシステムへ応用する研究をしています
  - 最近のメインテーマは「IoTとAIセキュリティ」，
    - サイドチャネル攻撃耐性のある暗号回路の実装やPUFの実装
    - 自動車や監視カメラなどをターゲットにしたAIの応用研究
    - AIの誤動作を誘発するような攻撃手法の研究と対策
- 研究室メンバー
  - 教授：藤野　毅（1986年～2003年まで三菱電機でLSIの製造技術や設計に従事）
  - 助教：吉田　康太（2022年3月当研究室で博士学位を取得，大倉研と兼務）
  - 構成員
    - (D1)：2名，(M2)：4名，(M1)：6名
    - 学部学生：9名（6名がM1に進学予定）
    - 秘書：松田　詩織
  - 客員教授：白畑　正芳，客員協力研究員：中井綱人（三菱電機）

今日の紹介内容

1. Society 5.0とエッジAI
2. ハードウェアセキュリティ技術
   1. 消費電力・電磁波を用いたサイドチャネル攻撃
   2. 複製不可能デバイスPUF
3. AI技術のセキュリティ
   1. AIと信頼性
   2. AIを狙った攻撃
   3. エッジAIのハードウェアセキュリティAI技術（深層学習）
4. AI技術の応用
   1. 車載カメラにおけるセンサーフュージョン（RGB-FIRカメラ）
5. 研究の進め方
   1. 応用演習
   2. 卒業研究以降

藤野

吉田

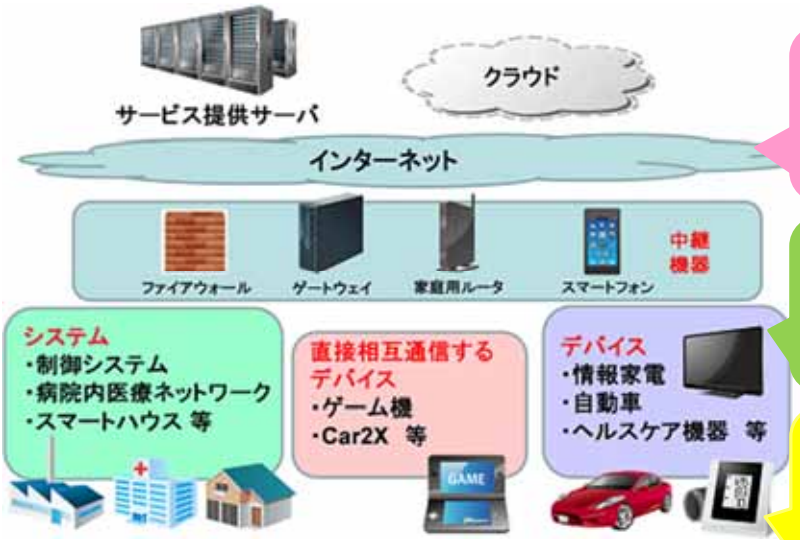
藤野

RITSUMEIKAN NTSOC LAB.

1. SOCIETY5.0とエッジAI

IoT機器とは？

- 「IoT」とは「Internet of Things」の頭文字を取った単語で，「モノのインターネット」
- 「身の周りのあらゆるモノがインターネットにつながる」仕組みのこと．
- 最近では，クラウド（インターネット上の計算資源）上のAIと通信



農業応用
農地に取り付けたセンサーで読み取った日射量や土壌の状況をもとに、水や肥料の量を制御

医療応用
着用型ウェアラブルデバイスによる健康状態の記録・管理、医師との共有

自動運転システム
自動車が自動車同士または周囲のインフラや歩行者と通信して，安全走行

スマートスピーカー

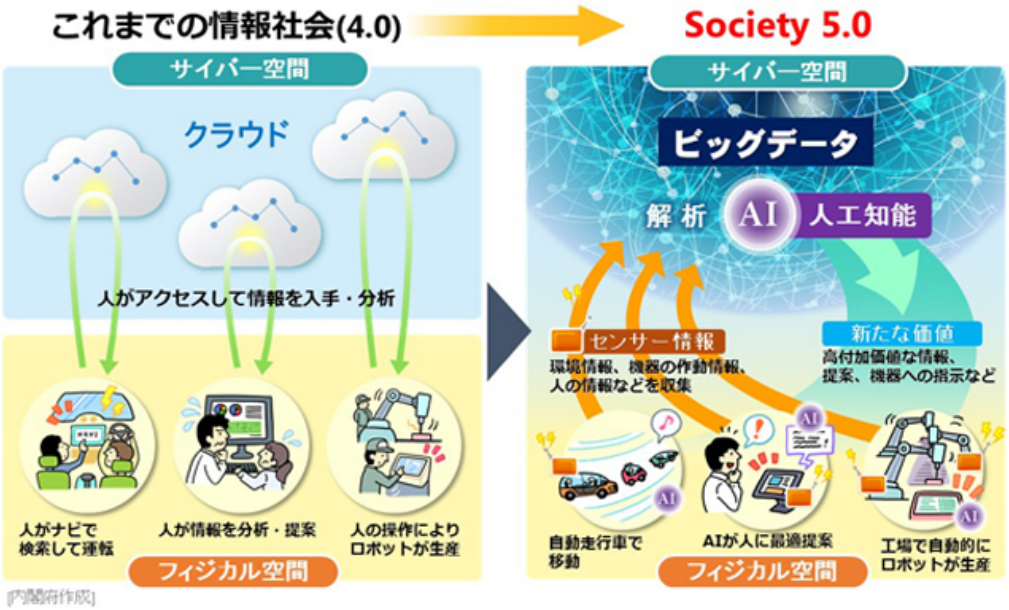
監視カメラ

コネクティッドカー

https://www.ipa.go.jp/security/iot/iotguide.html より引用

Society 5.0（スマート社会）

- サイバー空間（仮想空間）とフィジカル空間（現実空間）を高度に融合させたシステムにより、経済発展と社会的課題の解決を両立する、人間中心の社会（Society）
- 狩猟社会（Society 1.0），農耕社会（Society 2.0），工業社会（Society 3.0），情報社会（Society 4.0）



[内閣府作成]

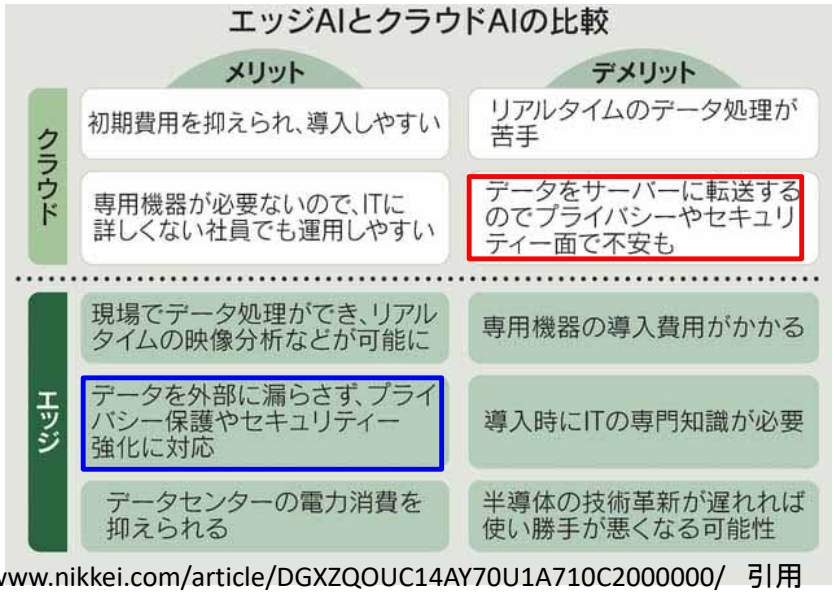
ドローン宅配，AI冷蔵庫，AIスピーカー，遠隔診療，介護ロボット，見守りサービス，無人トラクター，清掃ロット，ICT栽培，会計クラウド，旅館クラウド，商品の生産・販売管理，無人走行バス，隊列トラック

https://www.youtube.com/watch?v=BoEl9K2v2B0

Society5.0を実現するためのエッジAI

■ 今後のAI活用はエッジAIが重要に＠日経新聞記事（2021.7.29）

- クラウドからエッジへ――。人工知能（AI）の活用の舞台が利用者に近い端末（エッジ）に広がってきた。
- エッジAIは初期費用がかかるなど導入のハードルは高い。だが、データを外部に出さない解析が可能なため、プライバシー保護などで利点がある。

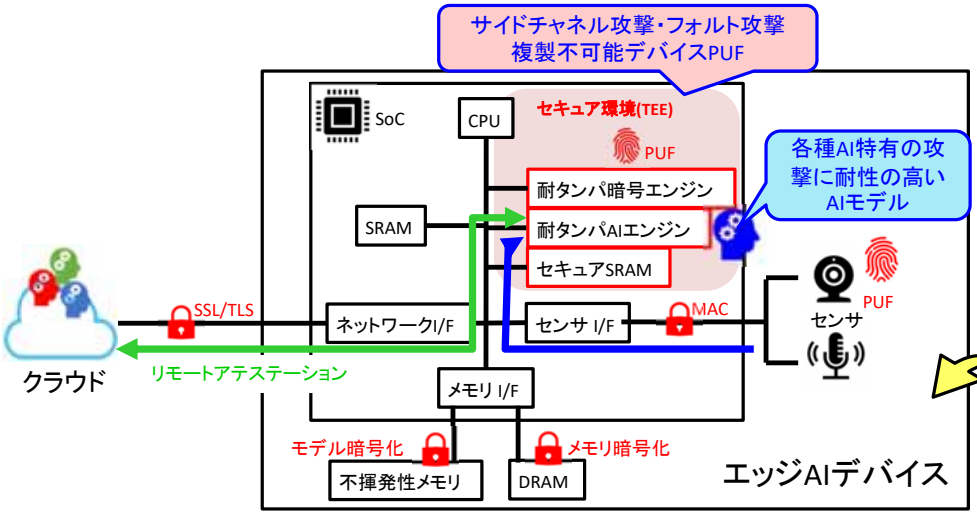


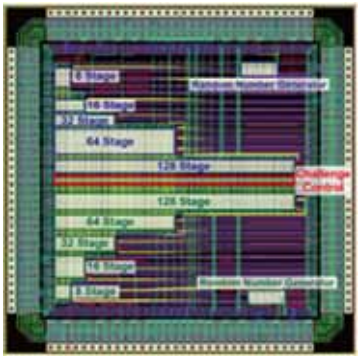
https://www.nikkei.com/article/DGXZQOUC14AY70U1A710C2000000/　引用

当研究室の目標：セキュアなエッジAIシステムハードウェア

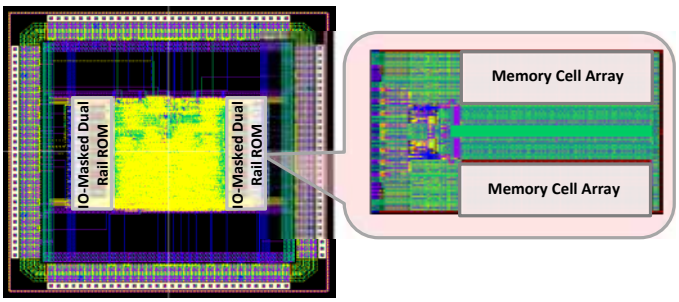
■ エッジAIを搭載したIoT機器の動作要件

1. 各種センサで正しいデータを取得する
2. 搭載されるエッジAIで正しい判断を行う
3. IoT機器が搭載される機器の知的財産を守る
4. 収集されるデータのプライバシーを守る





立命館大学　試作
アービターPUFテスト回路
by CREST Proj.



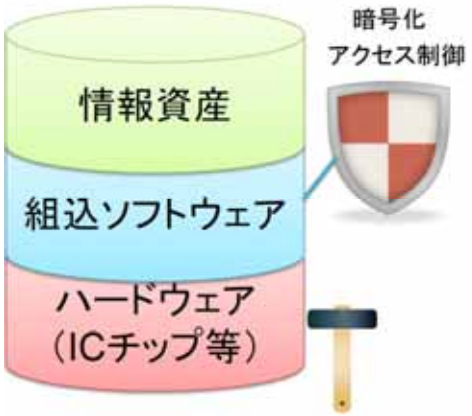
立命館大学　試作
サイドチャネル攻撃に耐性のあるAES暗号回路
by CREST Proj.

2.　ハードウェアセキュリティ技術

- サイドチャネル攻撃
- 複製不可能デバイスPUF

ハードウェアセキュリティ

- 情報セキュリティ
  - 電子的な手段を利用した情報のやり取りに関する安全性や信頼性の確保のこと
- サイバーセキュリティ
  - コンピューターネットワークに接続された機器で安全性や信頼性を確保すること
- ハードウェアセキュリティ
  - Root of Trust （信頼の基点）
  - たとえば，情報セキュリティを確保し，サイバー攻撃を防止するために，暗号技術を用いることが多いが，秘密にしなければならない暗号鍵が攻撃者に知られると無力
    ⇒ハードウェア（一般にはLSI）で暗号鍵を守る



https://www.ipa.go.jp/files/000048318.pdf より引用↑

ソフトウェアで実装するセキュリティ機能は【正しく動く】ハードウェアでの動作が大前提

↓

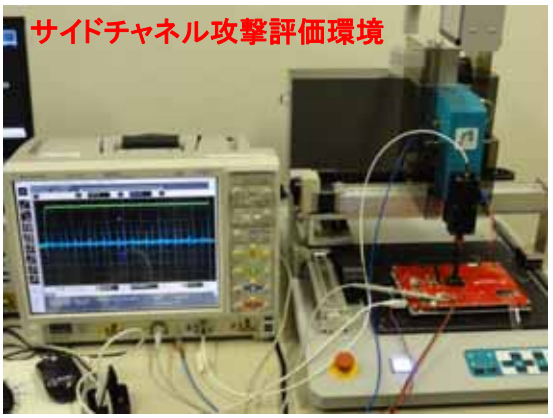
しかし、ハードウェアが攻撃を受ける可能性がある環境ではその大前提がなりたない

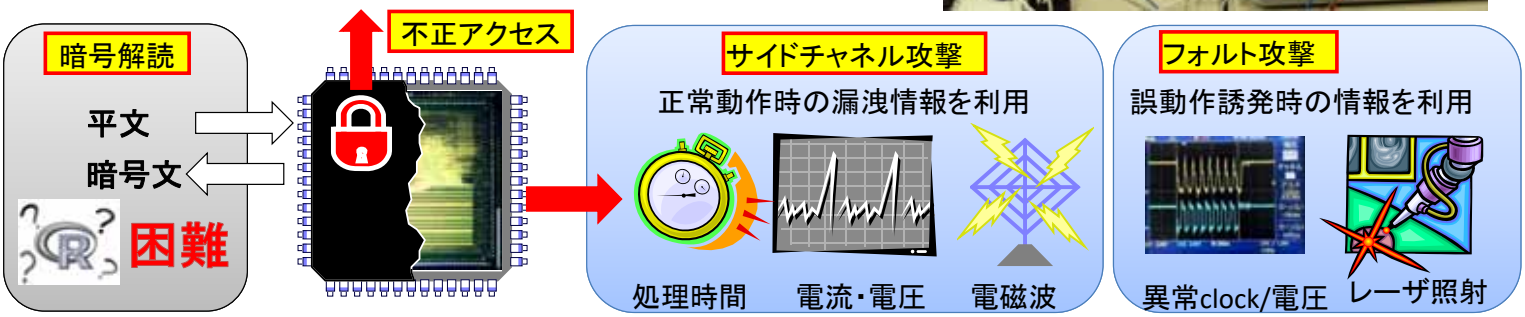
↓

【正しく動かない】ようハードウェアを攻撃された場合には、ソフトウェアで実装したセキュリティ機能も当然【正しく動かない】

サイドチャネル攻撃（SCA）とフォルト攻撃

- 暗号通信は万能ではない，鍵データが窃取されれば無効化される
  - ECUの鍵データ保管領域へのデバッグモードなどを利用した不正アクセスで鍵データを窃取可能
  - 暗号回路動作時の電力や電磁波（サイドチャネル情報）を用いて秘密鍵を窃取可能（サイドチャネル攻撃）
  - 暗号回路に対して，電圧やレーザーを使って誤動作させ秘密鍵を窃取可能（フォルト攻撃）

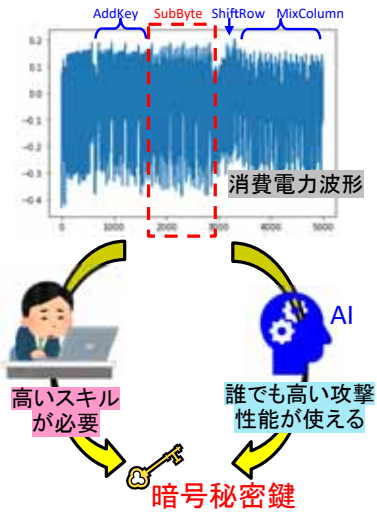


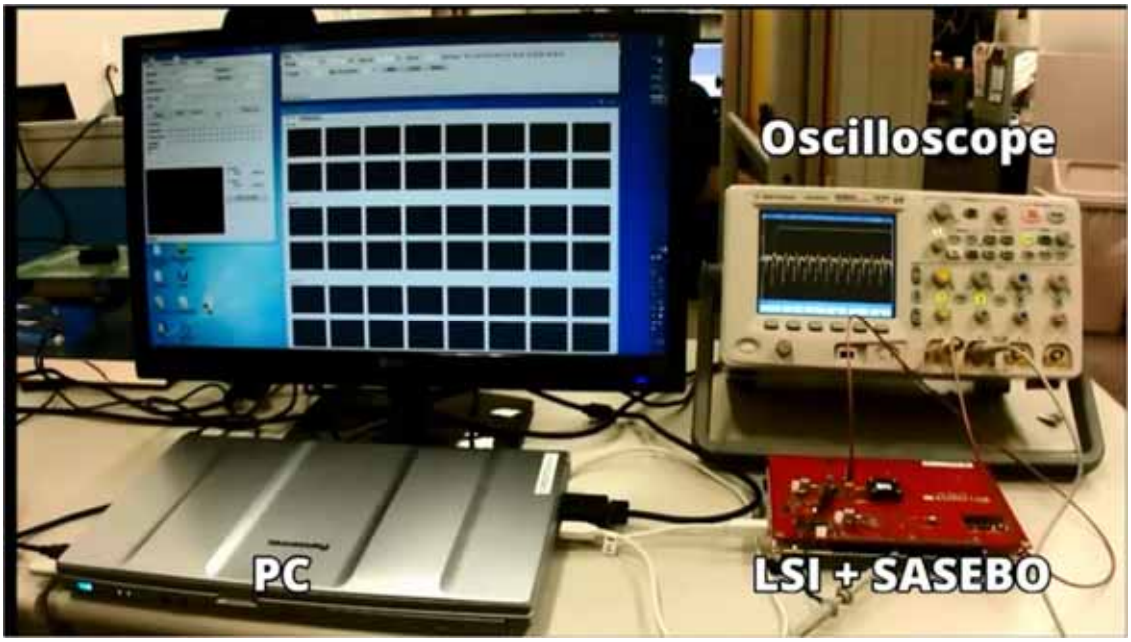


電力を利用したサイドチャネル攻撃実験デモ

- 当研究室で開発した，サイドチャネル攻撃に強いAES暗号回路を使っている
- 最近では深層学習技術を使ったサイドチャネル攻撃技術の研究を行っている

学会でもホットな話題
セキュリティとAIを両方
研究できる





最近使用しているSCA評価ボードChipWhisperer

- NewAETechnology社のボードでPythonを使って攻撃
  - ChipWhisperer-LiteまたはChipWhisperer-Proを使用（オシロも接続可能）
  - マイコン上に実装したソフトウェアの攻撃用ボード　または　HDLを使って実装したハードウェア評価用のFPGAボードを使って暗号処理時の波形取得
  - 取得波形をPythonを使って統計解析or/and深層学習で暗号鍵を導出

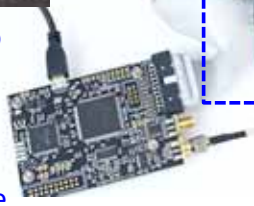
ハードウェア(HDL)評価用

この本でも本ボードを使って消費電力波形からパスワードを窃取する簡単攻撃手法が詳細されています

ChipWhisperer-Pro

UFO ターゲットボード
（MPUごとに交換する

CW308 ベースボード
ソフトウェア評価用



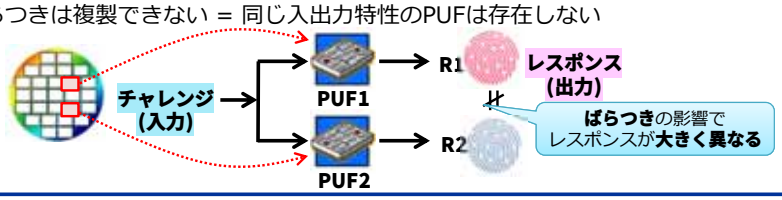
ChipWhisperer-Lite
(PCと接続して波形取得制御)

https://www.newae.com/chipwhisperer　より引用

PUF(Physically Unclonable Function)複製不可能デバイス

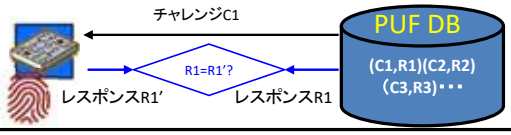
- 機器認証（偽造品の検知）や，暗号鍵の生成・保管に活用できる

- **PUF** ≒ ICチップの「指紋」
  - 設計データ上は同じ回路だが、入力に対して半導体の「**ばらつき**」によって異なる値を出力
  - ばらつきは複製できない = 同じ入出力特性のPUFは存在しない



PUFを使ってホンモノを見分ける

・PUFは同じチャレンジ（C）を与えても個体ごとに異なるレスポンス（R）が得られるので，事前にC-Rペアをデータベース（DB）に登録しておけば容易に真贋判定が可能



PUFを使った暗号鍵の生成と秘密情報の保管

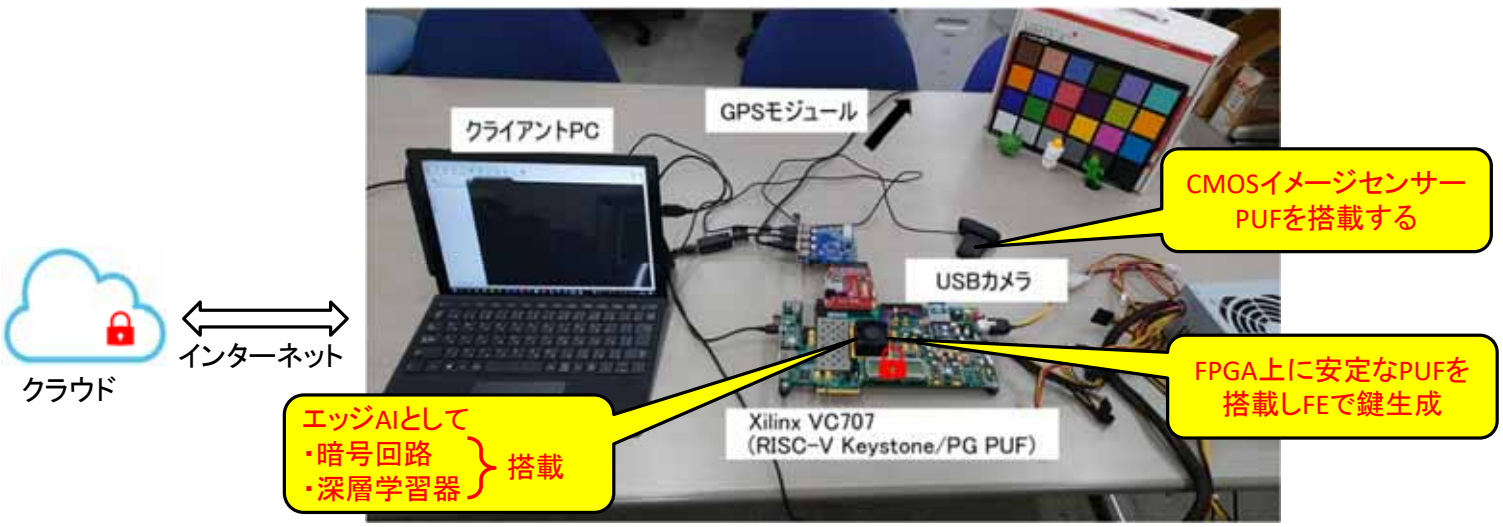
・PUFの秘密の指紋データを使って機密データや暗号鍵を暗号化して保管すると，そのデータ（鍵）はPUFを搭載したデバイス内だけで使用できるデータ（鍵）となる．

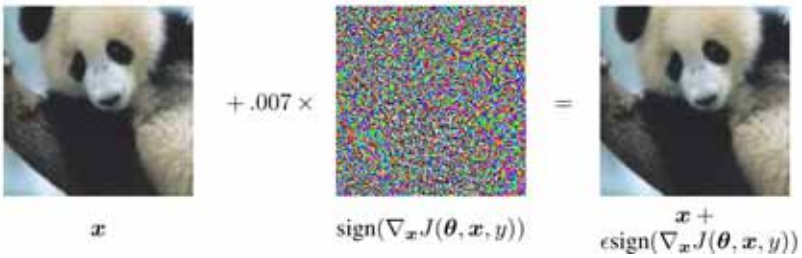
PUFを使った耐タンパキーレスエントリーデモビデオ

- 鍵データをPUF_IDで暗号化することで安全性を高める

PUFを使った暗号鍵生成に向けて

- PUFの出力は環境（温度電圧など）変動により不安定
  - 誤りを訂正する回路（FE:Fuzzy Extractor）で安定な鍵を生成
  - 暗号回路ではなくFEの動作消費電力を使ったサイドチャネル攻撃対策





3. AI技術のセキュリティ

- AIと信頼性
- AIを狙った攻撃
- エッジAIのハードウェアセキュリティ

RITSUMEIKAN

AIと信頼性の話(2)

✓ AIシステムの普及により，ハッカーにとって魅力的な攻撃目標に対策をしなければ 金銭的・人的被害がより大きくなる恐れも

◆ Amazon Go*にて決済せず退店できる事例 (どちらも2018年)

“私は万引きしたと思う？ Amazon Goは私にヨーグルトを請求しなかった”

https://twitter.com/dee_bosa/status/955459507403997184/photo/1

- Youtuberによる万引き “私はAmazon Goから非常に簡単に盗めることを証明した”

https://www.youtube.com/watch?v=JbyjL9tazxQ

◆ 顔認証システムへの攻撃による不正アクセス

http://www.itmedia.co.jp/news/articles/1711/13/news081.html

* Amazon Goとは，Amazonが展開する無人コンビニチェーン．レジがなく，店舗中に設置されたカメラで人物と商品を認識し，店から出ると自動的に請求と引き落としが行われる

AIと信頼性の話(2)

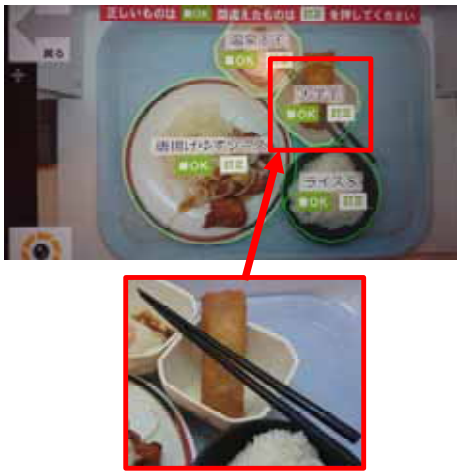
✓ 立命館大学生協(ユニオンスクエア)に一時期導入された自動レジシステム

カメラでメニューを認識
生協カードから引き落とし

写真のライスSを認識

箸を乗せる前

春巻　　55円
大学芋　88円

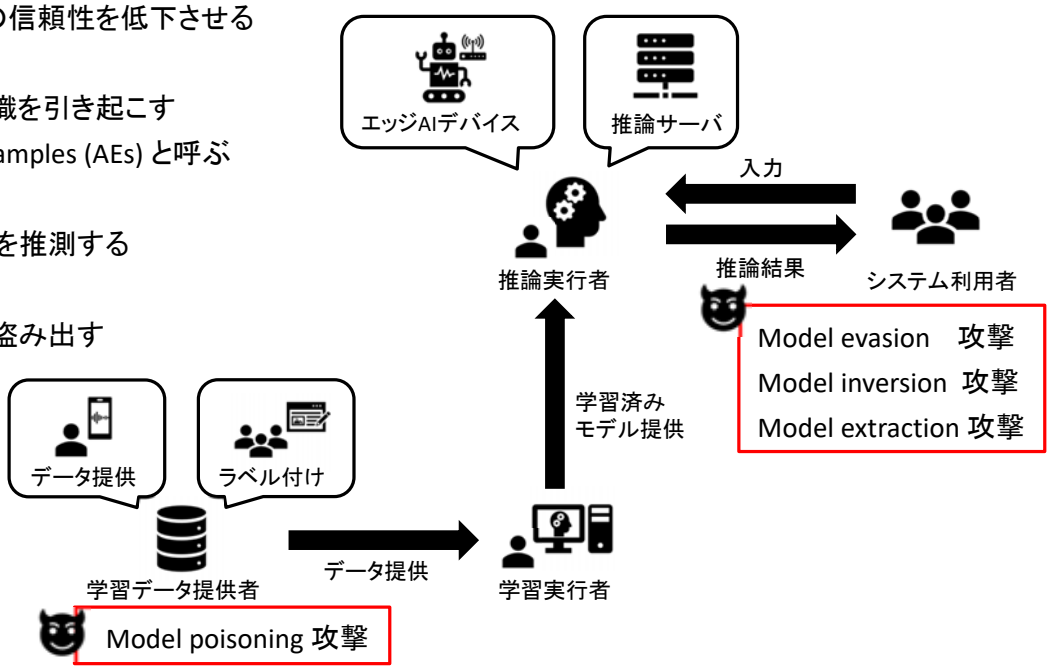


春巻きに 箸を乗せると 芋になる

深層学習におけるセキュリティ

✓ 深層学習アルゴリズムにおけるセキュリティ課題 **(AISec)**

- **Model poisoning 攻撃**
  学習データに細工し，AIモデルの信頼性を低下させる
- **Model evasion 攻撃**
  入力に細工し，AIモデルの誤認識を引き起こす
  細工された画像をAdversarial examples (AEs) と呼ぶ
- **Model inversion攻撃**
  入力と推論結果から学習データを推測する
- **Model extraction攻撃**
  入力と推論結果からAIモデルを盗み出す



Model Poisoning 攻撃

✓ バックドア攻撃 (Model poisoning攻撃)

データセットに悪意あるデータを混入させる攻撃

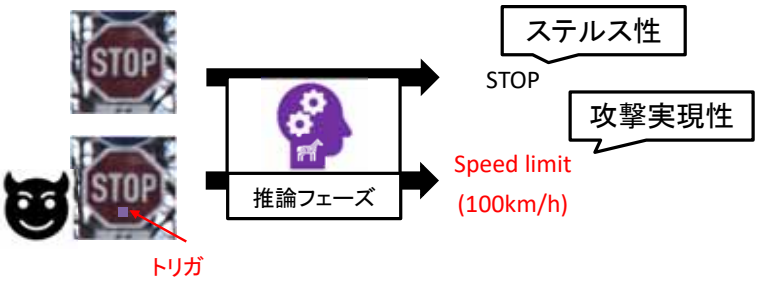
AIモデルの学習データセットの巨大化・ラベル付け作業の外注化を背景とする

バックドアモデルのステルス性と攻撃実現性が高く，通常のAI学習フローでは検知・対策が困難

✓ 学習フェーズ

➢ 学習データにトリガつき毒データを混入

✓ 推論フェーズ



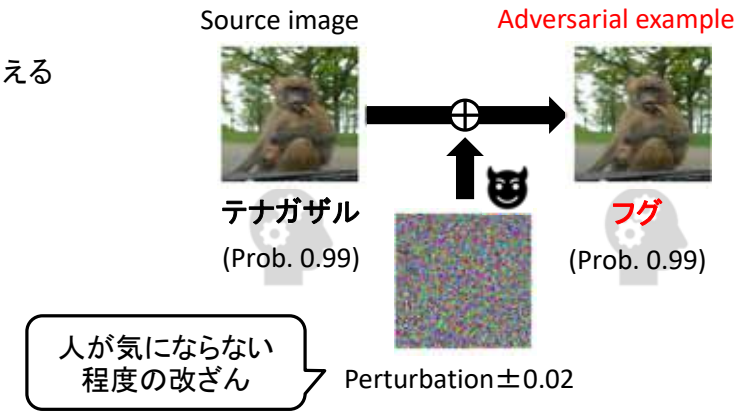
➢ 攻撃者の設定したトリガに反応して誤認識が発生

Model evasion攻撃

✓ 敵対的サンプル生成攻撃

人が気にならない程度の小さな改ざん(摂動)を画像に加える

摂動によって誤った認識が誘発される



✓ 道路標識認識システムに対する攻撃

“止まれ”の標識を“速度制限(30km/h)”に誤るような摂動を計算

現実世界でも高い効果が得られるよう，Unity(3Dゲームエンジン)を用いて様々な環境を模擬し，より強力な摂動を計算

距離や角度など様々な条件で評価

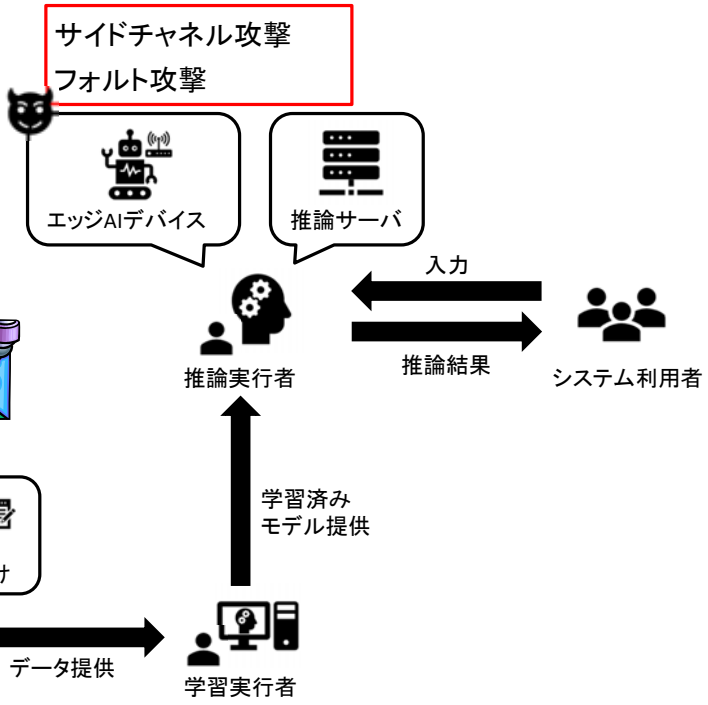
攻撃成功率100%　攻撃成功率83.3%

深層学習ハードウェアのセキュリティ

✓ 深層学習ハードウェアにおけるセキュリティ課題 **(AIHWS)**

- **サイドチャネル攻撃**
  サイドチャネル情報; e.g., 処理時間, **消費電力**
  推論結果・入力データ・**AIモデルを窃取**する
  AEsを生成する

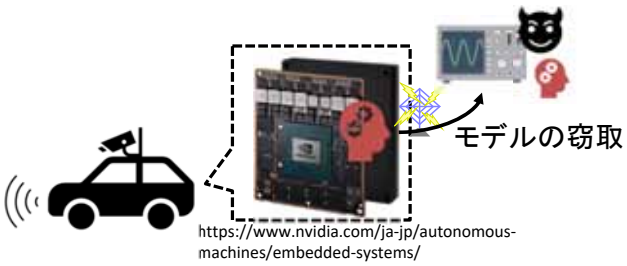
- **フォルト攻撃**
  フォルト; e.g., 電源・クロックグリッチ, レーザ照射
  推論結果の改ざん・誤認識の誘発



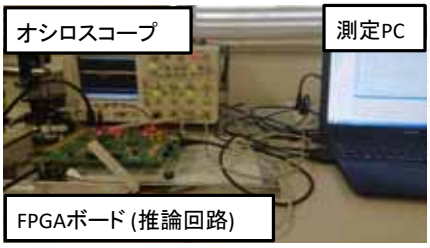
サイドチャネル攻撃によるパラメータの窃取

✓ 推論アクセラレータに対するサイドチャネル攻撃
FPGA上に深層学習モデルの推論(行列演算)アクセラレータを実装
推論回路の消費電力から深層学習モデルを摂取する攻撃を実証

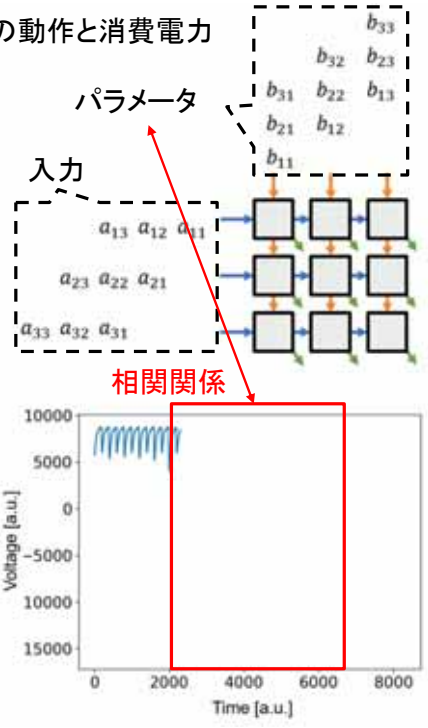
- 演算時に消費される電力とモデルパラメータの相関性を発見
- 商用にも利用される推論回路構造の一種に対して攻撃できることを報告



https://www.nvidia.com/ja-jp/autonomous-machines/embedded-systems/



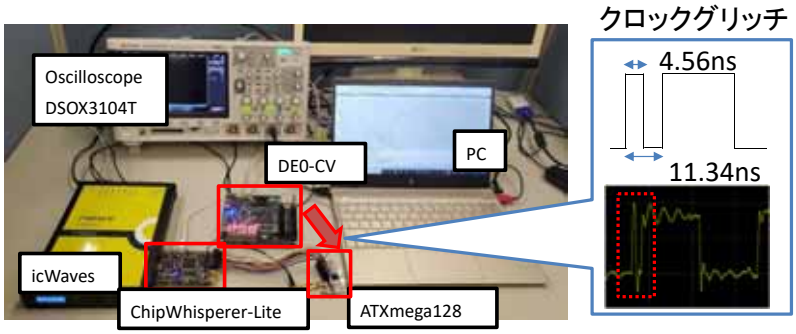
✓ 推論アクセラレータ回路の動作と消費電力

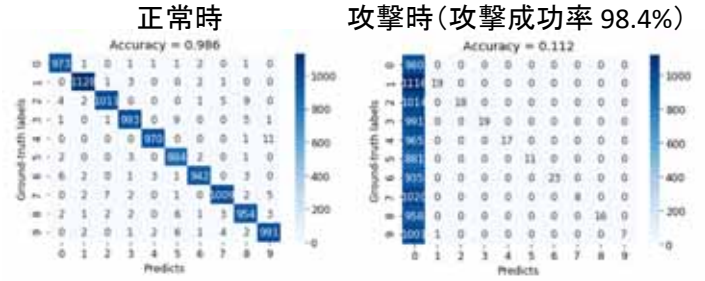


クロックグリッチによる誤認識の誘発

- ✓ マイコン上で実行されているCNNの推論処理に対してクロックグリッチを注入
- ✓ Softmax関数の演算処理の途中で強制的に処理を中断させることで推論結果を改ざん

- ✓ MNISTを用いた実験
  10,000枚のテスト画像を全て“0”に誤認識させることに成功





イメージセンサインターフェースへのフォルト注入による Adversarial Examples/Model Poisoning攻撃と対策

■ 攻撃手法

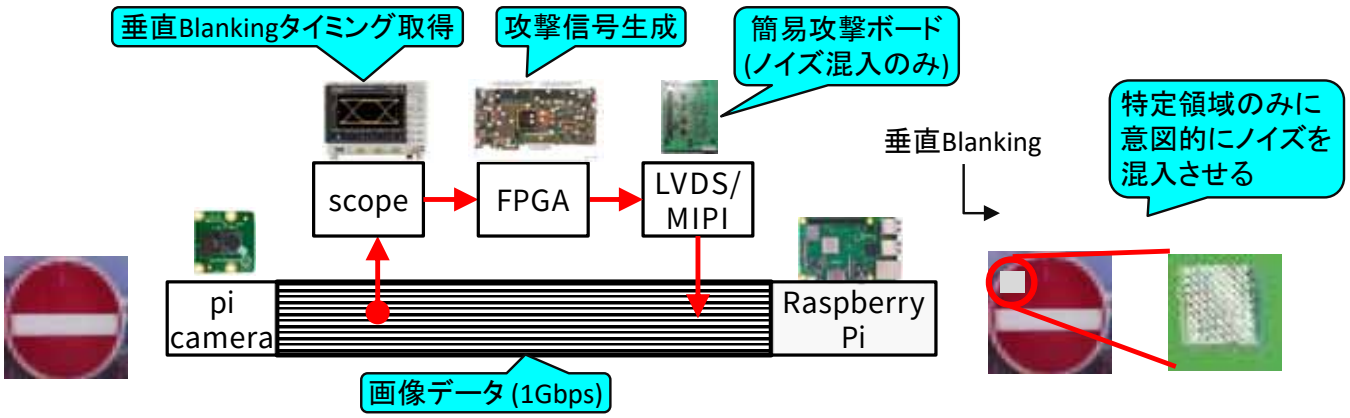
- ➢ イメージセンサと後段チップのMIPI通信路に改ざんデータを注入し，AIの誤動作を誘発させるAdversarial Examples(AE)攻撃

■ 対策手法

- ➢ イメージセンサで画像に低コストでMACを付与

■ 進捗

raspberry piカメラを用いてストリーム撮影中に、Poisoning攻撃に成功！



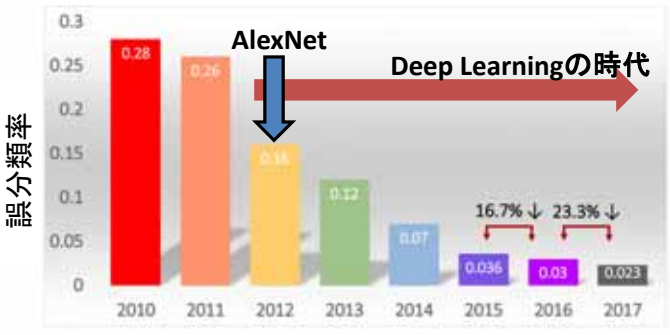
4. AI技術の応用
車載カメラにおけるセンサーフュージョン（RGB-FIRカメラ）

深層学習 (Deep Learning)と自動運転

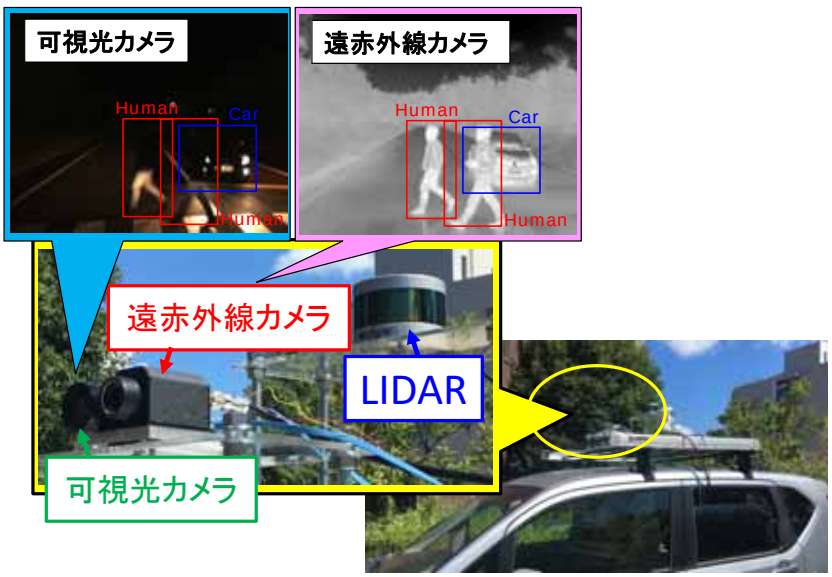
- 第三次AIブームの火付け役
- 画像認識チャレンジILSVRC（クラス分類：1000種類の画像データ画像を正しく分類できるか）において，CNN*の登場により精度が大幅に改善
- これを皮切りに，音声認識など様々な分野で深層学習技術の応用が活発化
- 藤野研でもセンサーフュージョンに活用

* 畳み込みニューラルネットワーク

ILSVRC 画像分類タスク



http://image-net.org/challenges/talks_2017/ILSVRC2017_overview.pdf



RGB-FIRカメラによる物体認識の研究

- RGBカメラとFIRカメラの併用により，環境の変化に頑健な物体認識を実現
- RGBカメラは環境光の変化によって，FIRカメラは環境温度の変化によって画像の写りが変化することを考慮する必要がある
- RGBカメラとFIRカメラの設置位置のずれを補正する必要がある

夏の日中

冬の夜間

✓ YOLOv3を用いた検出例

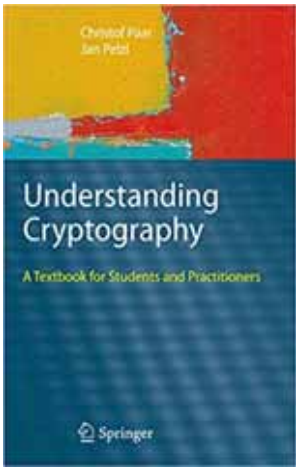
RITSUMEIKAN NTSOC LAB.

5. 研究の進め方
・応用演習
・卒業研究以降

RITSUMEIKAN

ゼミの進め方

- 3回生応用演習
  - 14inchノートPCを一人ずつ配布（卒業まで占有使用可能）
  - テキスト:「ディープラーニングのしくみがわかる数学入門」
    - Jupyter Notebookを導入してPythonプログラミング
    - Numpyとmatplotlibなどを使ってPythonになれる
    - 機械学習の基礎から深層学習へ
- 4回生ゼミ
  - 毎週約2時間半の卒研ゼミ（参加必須）
    - セキュリティの基本（暗号技術）を英語で学ぶ
    - テキスト:「Understanding Cryptography」
  - 先輩からの研究基礎ゼミ（春学期，参加必須）
  - 同じ研究グループの学生ゼミ（参加必須）
  - 教員参加進捗ゼミ（毎週１回開催，希望制）
  - ７月と１２月に研究室での成果発表



大学院・就職について

- **研究の主体は大学院生**
  - 4回生で学んだ研究のやり方を活用して，自主的に研究を進める能力を育成する.
- **学会発表**
  - **国内学会（M1,M2）**
    春（5月）「システムLSIワークショップ」
    冬（1月）「暗号と情報セキュリティーシンポジウム」
    ハードウェアセキュリティ研究会（3月，5月，7月，11月）
  - **M2は海外発表・論文誌投稿が目標**
    春（3月）NCSP（Nonlinear Circuit and Signal Processing）
    春（11月）ASHES（Attacks and Solutions in Hardware Security），
    AIsec(ARTIFICIAL INTELLIGENCE AND SECURITY)
    春（12月） asianHOST（Asian Hardware Oriented Security and Trust Symposium）
- **対外研究交流**
  - **耐タンパLSI設計,PUF，AIのセキュリティ技術：三菱電機，産総研**
- **過年度（2006-2022）の就職先（含内定）**
  - **東芝（キオクシア）10，ローム7，日立6，デンソー5，村田製作所4，アイシン精機4，パナソニック4，**
  - **トヨタ自動車3，イシダ3，富士通3，三菱電機2，SUBARU2，京セラ2，キャノン2，ルネサス2，**
  - **NTT西日本2，アドビックス2，ネットワンシステムズ2 他**
  - **博士課程進学4**

研究室訪問・見学

- 相談・質問のためのアポイントメント受付
  - fujino@se.ritsumei.ac.jp にメールしてください.
- 研究室配属見学（研究室HPおよびmanabaで確認してください）
  - 10/31（月）16:20頃～19:20頃
  - （16:20頃第1陣，17:20頃第2陣，18:20頃第3陣の交代制を予定）
    - ローム記念館の3Fの藤野第1研究室に先輩学生が待機しています.
    - また，教員もその隣の個人研究室にいますので相談があれば来てください.

    先輩学生からも以下のような説明がある予定です.

    (1)AIを使ったセキュリティ技術

    (2)AIのためのセキュリティ技術

    (3)サイドチャネル攻撃技術

    (4)PUF技術
  - 11/1（火）18:00～19:00@Zoom

    ・アドレスはmanabaで見てください

最後に

- ダーウィンの言葉？
  - 最も強い者が生き残るのではなく、
  - 最も賢い者が生き延びるのでもない。
  - 唯一生き残るのは、変化できる者である。
- 常に新しい技術にアンテナをはり，研究者・エンジニアとして第一線で活躍しよう
- 本研究室で扱う，活躍するための武器
  - （ハードウェア）セキュリティ技術
  - AI処理（深層学習）技術
  - 暗号回路やAIなどのセキュアなLSI設計技術

AI
×
セキュリティ

RITSUMEIKAN NTSOC LAB.

ご清聴ありがとうございました

立命館大学　ネットワークLSIシステム研究室
藤野　毅：fujino@se.ritsumei.ac.jp
吉田　康太：y0sh1d4@fc.ritsumei.ac.jp