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）

SCIS 2021 2021 Symposium on
Cryptography and Information Security
On-line, Jan. 19 – 22, 2021
The Institute of Electronics,
Information and Communication Engineers

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）
Study on Possibility of Estimating Smartphone Inputs from Tap Sounds (part2)

大内 結雲* 奥寺 瞭介* 塩見 祐哉*
大木 哲史* 西垣 正勝*
Yumo Ouchi* Ryosuke Okudera* Yuya Shiomi*
Tetsushi Ohki* Masakatsu Nishigaki*

あらまし スマートフォンのキー入力盗取に関するサイドチャネル攻撃として，タップ音からその入力内容が推定可能であるという脅威に関する研究が複数提案されている．しかし既存の手法は，正規ユーザへの積極的な干渉によりタップ音を入手することを前提としており，能動的な攻撃シナリオといえる．これに対し著者らは,「正規ユーザがスマートフォンにキー入力をする際のタップ音を, 攻撃者が外部のマイクで単純に盗聴する」という受動的な攻撃シナリオを想定し，その脅威に関する基礎検討を SCIS2020 で行った．その結果，高い精度でキー入力の識別が可能であること，ならびに，単純に攻撃者から距離を取るだけでは，攻撃を防ぎきれない可能性が残ることが判明した．ただし SCIS2020 の実験では，攻撃モデルに関する多くの制約が存在しており，１名の実験協力者による限定的な評価となっていた．そこで本稿では，実験条件の制約を一部緩和させるとともに，複数名のタップ音を収集し，著者らの SCIS2020 の実験で得た結果の再現性を確認する．さらに，収集した複数名のデータを用いることで，既知のユーザのタップ音で学習した識別器を用いて未知のユーザの入力内容を推定可能であるかを検証する．具体的には，まず，実験環境を防音室から環境音下に移し，7 名の実験協力者のタップ音を収集する．そして，各実験協力者に対して，本人を除く実験協力者 6 名のタップ音で識別器を学習し，その識別器での本人のキー入力に対する推定精度を算出する．また，騒音環境下における識別器の性能を評価することにより，この攻撃に対する防御策の検討に資する．タップ音に対して周辺騒音によるノイズ付加を行い，キー入力推定精度を十分に低下させることのできる騒音環境を探る．

キーワード サイドチャネル攻撃，情報漏えい対策

1 はじめに

近年，スマートフォンの普及やキャッシュレス決済サービスの普及により, スマートフォンを用いて個人情報やパスワード等のセンシティブな情報を入力する機会が増加している. そのような秘密情報を盗む攻撃手法の１つとしてサイドチャネル攻撃という攻撃が存在する[1]．サイドチャネル攻撃は, 暗号モジュールが搭載されている機器を外部から観察し, 得られる副次的な情報を元に暗号解析を行う攻撃である. サイドチャネル攻撃はログに残らないために, 攻撃の証拠が残りにくいという特徴がある．サイドチャネル攻撃の 1 つにテンペスト攻撃が存在する[2]．テンペスト攻撃はディスプレイやケーブルから漏洩する微弱な電磁波や音を検知することで, ディスプレイに表示された情報や入力された文字列等を取得する攻撃である. このようなテンペスト攻撃の一手法とし[1]て，スマートフォンやタブレット端末への入力操作に伴い発生する音響を利用して入力内容を推測する攻撃手法が提案されている[3][4][5]．しかし，既存の手法は正規ユーザの端末に対して, 積極的な干渉が必要な攻撃シナリオが想定されており, 現実的な脅威にはなり得ない.

著者らは文献[6]にて,「正規ユーザがスマートフォンにキー入力をする際のタップ音を, 攻撃者が外部のマイクで単純に盗聴する」という受動的な攻撃シナリオを想定し, 正規ユーザのキー入力が攻撃者にどの程度漏れるのか検証した．文献[6]の評価は攻撃モデルに関する多くの制約が存在しており，１名の実験協力者による限定的な予備実験ではあったが, タップ音から高い精度でキー入力の識別

* 静岡大学, Shizuoka University

が可能であることが判明した．また，攻撃対象のスマートフォンから攻撃者側の録音デバイス（タブレット端末）までの距離を 10cm～70cm の範囲内で変化させて評価したところ，キー入力推測の識別率は大きく変化しなかった．この結果から，「攻撃者からの距離を取る」という単純な対策では，攻撃を防ぎきれない可能性が残ることが判明した．

そこで本稿では，実験条件を一部緩和させるとともに，複数のユーザのタップ音を収集し，文献[6]の実験で得られた結果の再現性を確認する．さらに，既知のユーザのタップ音で学習した識別器を用いて未知のユーザの入力内容を推定可能であるかを検証する．また，この攻撃に対する防御策として，タップ音に対するノイズ付加を検討する．騒音を重畳したタップ音に対する識別器のキー入力精度を評価することによって，推定精度を低下させることのできる騒音環境を探る．

2 関連研究

スマートフォンのキー入力盗取に関するサイドチャネル攻撃として，タップ音からその入力内容が傍受可能であるという脅威が存在する．

Shumailo らは正規ユーザのスマートフォンやタブレットの内蔵マイクとタップ音によって入力内容を推測する手法を提案している[3]．正規ユーザの端末に複数のマイクが内蔵されている場合，タップした際に発生する音響は，上部に設置されているマイクと下部のマイクで受信する時間に差が生じる．この音響の到達時間の差から画面上のどこをタップしたときの音であるのかを計算し，正規ユーザのキー入力を 61%の精度で推測可能であることを報告している．しかしこの攻撃手法は，タップ音を盗聴する録音デバイスが正規ユーザの端末に内蔵されているマイクであり，事前侵入が必要という点で妥当性を欠く攻撃シナリオとなっている．攻撃対象のスマートフォンに侵入できたのならば，不正者はキーロガー等を用いて正規ユーザのキー入力を直接取得できる．

Lu らは攻撃者のスマートフォンの内蔵スピーカから攻撃対象のスマートフォンに向けてソナー音を放射し，タップ入力の際の正規ユーザの指からの反射波を分析することによって，正規ユーザのキー入力を 90%の精度で推測可能であることを報告している[4]．しかしこの攻撃手法は，攻撃者が正規ユーザの端末に対して積極的に干渉するタイプのサイドチャネル攻撃となっており，能動的な攻撃シナリオであると言える．

Zhuang らは PC の物理キーボードの打鍵音から入力内容を推測する攻撃を提案している[5]．正規ユーザの打鍵音を近辺に設置されているマイクから盗聴し，得られた音声データに対してケプストラム分析で特徴量抽出を行い，クラスタリングを行うことによって，正規ユーザのキー入力を 96%の精度で推測可能であることを報告している．この事実は，スマートフォンにおいても，タップ入力音を外部マイクによって盗聴するだけで，正規ユーザのスマートフォンへの入力を推測できる可能性があることを意味している．

そこで本研究では，「正規ユーザがスマートフォンにキー入力をする際のタップ音を，攻撃者が外部のマイクで単純に盗聴する」という受動的な攻撃シナリオを想定し，その脅威の深刻度を評価するとともに防御策を検討する．

3 攻撃手法

正規ユーザがスマートフォンにキー入力をする際のタップ音を，攻撃者が外部のマイクで単純に盗聴する受動的な攻撃シナリオにおいて，正規ユーザのキー入力に関する情報が攻撃者にどの程度漏れるのかを検証する．本研究の現段階では，キー入力を PIN 入力に限定して調査を行っている．

攻撃者は攻撃対象のスマートフォンのタップ音を，その近辺に設置した録音デバイスを用いて盗聴する．収集した音声データを畳み込みニューラルネットワーク（CNN: Convolutional Neural Network）に学習させて識別器を作成する．CNN の入力は，音声データのメル周波数スペクトログラムをヒートマップ化した画像である．CNN の出力は，正規ユーザが入力したキー情報である． CNN は深層学習の一種で，画像識別で広く用いられている．文献[6]では，メル周波数ケプストラム係数（MFCC: Mel-Frequency Cepstrum Coefficients）を時系列に表示したメル周波数ケプストログラム画像を CNN に入力していたが， 文献[7]によると，MFCC 算出時に行う離散コサイン変換は，学習時に必要な情報を除去してしまうため，深層学習には離散コサイン変換を行わないメル周波数スペクトルの方が適切だとされている．そのため，今回の実験ではメル周波数スペクトルを時系列順に表示したメル周波数スペクトログラム画像を CNN の入力とする．

正規ユーザのタップ音は静寂な会議室内で収録する．この静音下タップ音音源に騒音音源を重畳することによって，騒音環境下タップ音音源を作成する．静音下タップ音音源を用いて文献[6]の実験の追試を行う．騒音下タップ音音源を用いて攻撃の脅威範囲と防御策を検討する．

4 実験方法

4.1 実験環境

実験に使用した機器の諸元を表 1 に示す．実験は静岡大学浜松キャンパス情報学部棟 1 号館内の会議室で行い，会議室に設置されている机の上に，攻撃対象のスマートフォンと攻撃者の録音デバイス（タブレット端末）を設置した．文献[6]の実験より，タブレット端末とスマートフォンの距離が異なっていてもキー入力推定の精度は大きく変化しなかったことから，今回はタブレット端末とスマートフォンの距離は 10cm に固定した．実験環境を図 1 に示す．実験時の会議室は静寂であり，タブレット端末の横に

普通騒音計（リオン株式会社製 NL-42）を設置し，周波数重みづけ：A 特性，時間重みづけ：Fast 特性のモードで暗騒音を測定したところ，騒音レベルは 35-45dB であった．正規ユーザ役の実験協力者には，椅子に座って机上のスマートフォンにキーを入力してもらった．非利き手はスマートフォンには触れずに，利き手の人差し指のみでキー入力を行うよう指示した．

表 1 実験機器

種類	名称
スマートフォン	iPhone6
タブレット	iPad Pro
分析用 CPU	2.3 GHz クアッドコア Intel Core i7
OS(PC)	macOS Catalina, macOS Big Sur
音声編集ソフト	Audacity
言語	Python 3.7
音声処理ライブラリ	librosa 0.7.0
深層学習ライブラリ	Keras 2.3.1
	TensorFlow 1.14.0

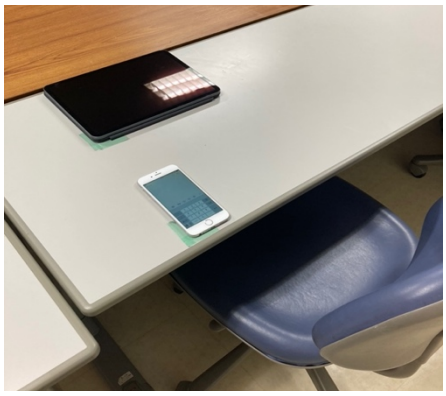

図 1　実験環境

4.2　音響データ収集

攻撃対象のスマートフォンの画面に日本語用ソフトウェアキーボードの PIN 入力インタフェースを表示させた．イヤホンを装着した実験協力者 7 名に，爪が画面に当たるように 10 種類の数字キーを「1」，「2」，・・・，「9」，「0」の順番でタップしてもらった．音声データの分析を簡易にするために，タップの際には，100bpm のメトロノームの音声をイヤホンから流し，実験協力者はそのリズムに合わせてタップを行うようにした．「1」から「0」のタップ入力を 1 セットとし，今回は各実験協力者に 100 セット分の入力を繰り返してもらい，合計 1,000 回のタップ音を収集した．ただし，タップミスの混入に備え，実験協力者には各自の判断で1~10セット余分にタップするよう指示した．タブレット端末に内蔵されているマイクと録音アプリで，タップ音を録音した．録音する音響データの形式は m4a である．収録された音響データを「静音下タップ音音源」と呼ぶこととする．

4.3　音響データ処理

収録後，各数字の音声データを m4a 形式から wav 形式に変換した．音声編集ソフトウェア Audacity [8]を用いて，1 つの音声データが 1 タップ分になるように音声データ全体を約 0.35sec ごとに時分割した．図 2 に音声データの時分割の例を示す．静音化音響音源から生成した音響データを「静音下データセット」と呼ぶこととする．5.1 節および 5.2 節の実験（文献[6]の実験の追試）は，この静音下データセットを用いて実施する．

また，静音下タップ音音源に騒音音源を重畳することによって，騒音下タップ音音源を作成した．今回は，騒音音源は電子協騒音データベース[9]内に収録されている「計算機室（中型）」の音源を用い，タップ音音源と騒音音源の信号対雑音比（SNR：Signal-to-Noise Ratio）が，0dB，-5dB，-10dB となるように重畳した．静音下タップ音音源から静音下データセットの生成と同じ手順を用いて，騒音下タップ音音源から騒音下データセットを生成した．5.3 節の実験（攻撃の脅威範囲と防御策の検討）は，この騒音下データセットを用いて実施する．

音響データ収集時に実験協力者のタップミスが見られた場合は，当該時間区間のタップ音をデータセットから除外した．また，今回の実験では全実験協力者のほとんどのタップ音の振幅（強度）が-0.4~0.4 の範囲にあることが確認できた．そのため，振幅が-0.4~0.4 を超えたタップ音もデータセットから除外した．タップ音の除外が生じた際には余分に録音しておいたタップ音を補填し，各実験協力者の静音下データセット，騒音下データセットがすべて「各 PIN のタップ音×100 セット＝1000 個のタップ音」になるようにした．

静音下データセット，騒音下データセットのすべてのタップ音に対し，Python の音声処理ライブラリ librosa を利用して，各 1 タップ分の音声データのメル周波数スペクトログラム画像を作成した．作成した画像例を図 3 に示す．横軸が時間，縦軸が周波数のヒートマップ画像である．実際の学習／識別においては，カラーバー，x 軸，y 軸，ラベルの表示は削除し，640×480 ピクセルの画像情報として CNN に入力した．メル周波数スペクトログラム画像はカラー画像のため，RGB の 3 チャネルの画像情報として CNN に入力される．

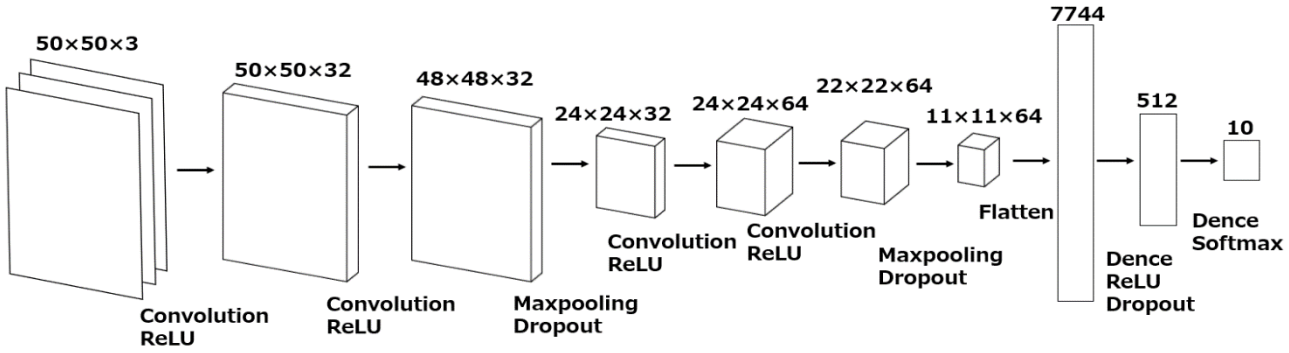

図 4　CNN モデル

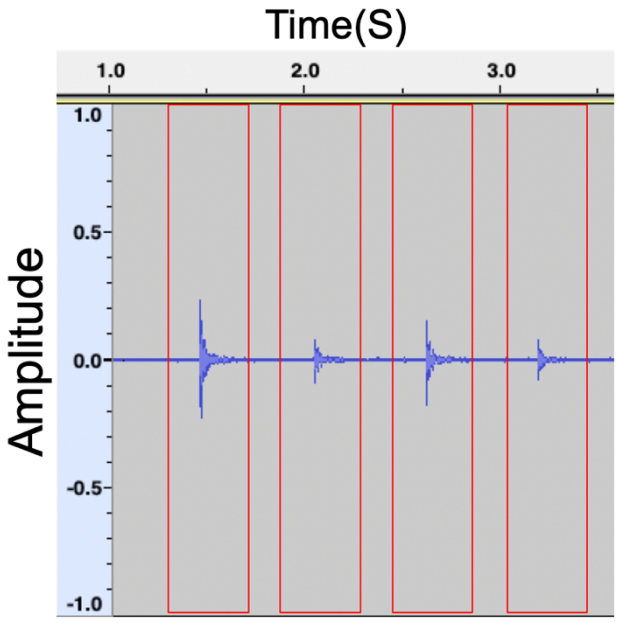

図 2　音響データの時分割

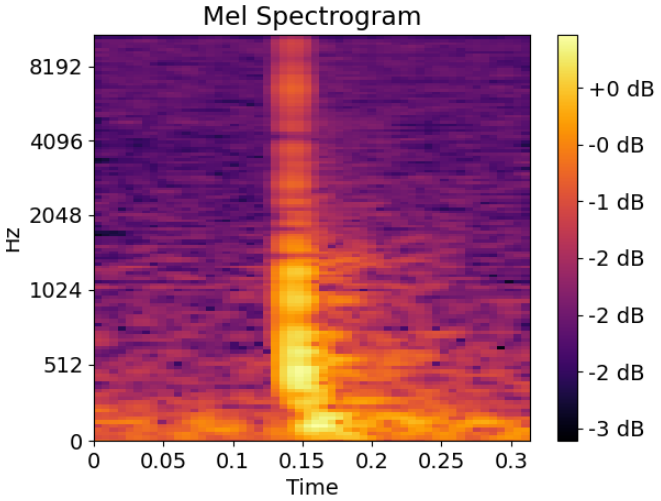

図 3　メル周波数スペクトログラム画像

4.4　機械学習

Python の深層学習ライブラリ Keras，TensorFlow を用いて CNN の学習および識別を行った．CNN のネットワーク構成は文献[10]を参考にした．今回使用した CNN モデルを図 4 に示す．

4.3 節で作成したメル周波数スペクトログラム画像（RGB の 3 チャネルの画像）を CNN の入力として与える．CNN は，まず，これを 50×50 ピクセルに圧縮した上で，3×3 のフィルタを用いて 2 連続で畳み込みを行い，32 枚の特徴量マップを得る．次に，Max プーリングにより画像サイズを半分に縮小する．今回は Max プーリングを適用する際の小領域サイズは 2×2 とした．更に，畳み込みを 2 連続で行い，Max プーリングを行った．この結果得られた 3 次元の配列を 1 次元に平滑化し，全結合層につなげた．今回は PIN 入力の推測（10 クラス分類）が目的であるため，最後に 10 個のノードを持つ全結合層につなげた．活性化関数は出力層ではソフトマックス関数，その他の層ではランプ関数を用いた．

5　実験結果

学習を終えた CNN に対して評価用データを入力し，識別率を評価した．

5.1　静音環境下の本人間識別精度

静音環境下での本人-本人間の識別精度を評価するため，実験協力者ごとに，4.3 節で作成した静音下データセットのメル周波数スペクトログラム画像を用いて実験を行った．各実験協力者の全データセットを，それぞれ，テスト用と訓練用に 2:8 の割合で分割し，更に訓練用データを検証用と学習用に 2:8 の割合で分割して使用した．識別率の算出には 5-fold Cross Validation を用い，5 回分の評価で得られた識別率の平均を用いて評価を行った．実験協力者ごとに，本人の訓練用データを学習させた識別器を作成し，その識別器によって本人のテスト用データの識別を試みた．その結果，93.1%の識別率が得られた．実験協力者ごとの識別率を表に示す．表 2 から，静音環境下においては，本攻撃手法を用いた PIN 入力の識別は高い精度で可能であることが確認できた．

表 3　本人-他人間の識別率

テストデータ＼訓練データ	ユーザ 1	ユーザ 2	ユーザ 3	ユーザ 4	ユーザ 5	ユーザ 6	ユーザ 7
ユーザ 1		6.4%	18.8%	8.0%	35.3%	13.9%	29.1%
ユーザ 2	17.1%		17.4%	12.3%	19.7%	39.1%	52.3%
ユーザ 3	23.2%	18.0%		7.9%	21.4%	17.1%	17.4%
ユーザ 4	20.5%	8.0%	12.9%		25.1%	10.0%	20.1%
ユーザ 5	45.8%	17.2%	24.3%	9.3%		35.8%	46.3%
ユーザ 6	31.6%	45.1%	23.6%	22.9%	43.3%		48.8%
ユーザ 7	18.5%	28.3%	19.4%	13.1%	32.6%	21.8%	

表 2　本人間の識別率

実験協力者	識別率(%)
ユーザ 1	98.1
ユーザ 2	98.6
ユーザ 3	98.1
ユーザ 4	91.4
ユーザ 5	97.6
ユーザ 6	99.0
ユーザ 7	69.4

5.2　静音環境下の他人間識別精度

実際の攻撃においては, 攻撃者が事前に攻撃対象のユーザのスマートフォン入力のタップ音を取得できるというような状況にはないことが通常である. すなわち攻撃者は, 自分自身あるいは攻撃者の仲間内のタップ音を使って識別器を学習することになるだろう. そこで, 静音環境下での本人-他人間の識別精度を評価する. 実験協力者ごとに, 4.3 節で作成した静音下データセットのメル周波数スペクトログラム画像を用いて実験を行った.

まず, 攻撃者が自分自身のタップ音を使って識別器を学習する場合を想定しての実験を行った. 5.1 節の実験と同様, 各実験協力者の全データセットを, それぞれ, テスト用と訓練用に 2:8 の割合で分割し, 更に訓練用データを検証用と学習用に 2:8 の割合で分割して使用している. 識別率の算出方法も, 5.1 節の実験と同じである. 実験協力者ごとに, 本人の訓練用データを学習させた識別器を作成し, その識別器によって他人（本人以外の 6 名の実験協力者）のテスト用データの識別を試みた. その結果を表 3 に示す. 表 3 から, 既知のユーザ 1 名分のタップ音を学習した場合, 未知のユーザのキー入力を識別できる精度は最も高い場合で 52.3%, 最も低い場合で 6.4%という結果となった.

次に, 攻撃者が自分自身のタップ音と攻撃者仲間のタップ音を使って識別器を学習する場合を想定した実験を行った. 5.1 節, 5.2 節の実験と同様, 各実験協力者の全データセットを, それぞれ, テスト用と訓練用に 2:8 の割合で分割し, 更に訓練用データを検証用と学習用に 2:8 の割合で分割して使用している. 識別率の算出方法も, 5.1 節, 5.2 節の実験と同じである. 実験協力者ごとに, 他人（当該実験協力者以外の 6 名の実験協力者）の訓練用データのすべてを学習させた識別器を作成し, その識別器によって本人のテスト用データの識別を試みた. 全員のデータを用いて識別器を作成し, その識別器によって本人（当該実験協力者）のテスト用データの識別を試みた. その結果を表 4 に示す. 表 4 から, 既知のユーザ 6 名分のタップ音を学習した場合, 未知のユーザのキー入力を識別できる精度は最も高い場合で 71.6%, 最も低い場合で 24.3%という結果となった.

表 4　本人以外の他人データを学習した場合の識別率

テストデータ	ユーザ 7 を含めた場合の識別率(%)	ユーザ 7 を除いた場合の識別率(%)
ユーザ 1	54.5	35.1
ユーザ 2	56.3	30.0
ユーザ 3	31.7	27.9
ユーザ 4	24.3	17.8
ユーザ 5	68.5	34.2
ユーザ 6	71.6	61.2
ユーザ 7	41.2	NA

表 3, 表 4 から, 既知のユーザのタップ音を学習することで, 未知のユーザのキー入力についてもある程度の精度で識別が可能であることがわかった. また, ほとんどの実験協力者において, 既知のユーザ 1 名分のタップ音を用いて識別器を学習させた場合よりも, 既知のユーザ 6 名分のタップ音を用いて識別器を学習させた場合のほうが, 未知のユーザのキー入力推定に対する識別率が高まるという結果が得られた.

5.1 節の実験では, ユーザ 7 の本人間識別率が他の実験協力者の識別率よりも低い結果となった. すなわち, ユーザ 7 のデータセットは他の実験協力者と比べて「性質が悪

いデータ」であった可能性がある（実際，当該実験協力者は，実験終了後に「単調なキー入力操作が続いたため，知らず知らずの中にタップの強さが大きくなったり，小さくなったりしてしまった」という感想を漏らしていた）．そこで，ユーザ 7 のデータセットを除いた上で，本人-他人間の実験の再実施を試みた．表 4 にはその結果も併記してある．その結果から，ユーザ 7 のデータを学習に加えた場合のほうが，高い識別率が得られることがわかった．このことから，本人-本人間の識別精度が低い場合のデータであっても，そのデータは他人を識別するにあたっての学習データとしての効果を持つことが示唆される．

5.3 ノイズ付加による精度評価

騒音環境下での本人-本人間の識別精度を評価するため，4.3 節で作成した騒音下データセットのメル周波数スペクトログラム画像を用いて，5.1 節と同じ実験を行った．その結果を表 5 に示す．

表 5　騒音環境下の本人間の識別率

実験協力者	識別率(%)		
	SNR = 0	SNR = -5	SNR = -10
ユーザ 1	98.5	97.5	10.7
ユーザ 2	63.8	54.2	27.7
ユーザ 3	95.7	55.2	25.6
ユーザ 4	95.8	7.9	7.9
ユーザ 5	95.7	94.0	7.9
ユーザ 6	97.3	97.5	98.0
ユーザ 7	7.8	7.4	7.6

表 5 から，1 名を除いては，騒音環境下ではキー入力識別精度が下がることが明らかとなった．SNR＝-10 の際の識別率の低下が顕著なことから，スマートフォン操作時にタップ音がかき消される程度の音楽などをスマートフォンで再生してやるような方法が，今回の攻撃に対する防御策となり得ると考えられる．

6 制限

5 章の実験からは，スマートフォンのタップ音からキー入力を推定するという攻撃が，CNN を用いることによって高い確率で実行可能であるという結果が得られたが，本稿で行った評価にはまだ多くの制限がある．

第一にタップの仕方である．今回の実験では爪を立ててスマートフォンにタップした際の音声を収集したため，正規ユーザの行動シナリオが限定的である．したがって，今後は爪を立てずに指の腹のみでタップした際の識別可能性を検証する必要がある．

第二に端末依存性である．今回の実験では特定のスマートフォン端末と録音デバイス（タブレット端末）の組み合わせに対して評価を行った．しかし，タップ音の特性は正規ユーザが使用しているスマートフォン，攻撃者が使用する録音デバイスに依存する可能性がある．今後は多様なスマートフォン，録音デバイスに対する評価を検討する．

第三に攻撃対象である．今回は，本研究の基礎検討の段階であったため，キー入力を PIN に限定している．今後は，フリック入力型 50 音キーボードや QWERTY キーボードに対する識別精度ついても検討していく必要がある．

7 研究倫理

本研究は提案シナリオにおけるタップ音識別の脅威を調査することを目的とし，実験を通じて相応の危険性が確認される結果となった．しかし，今回はまだ多くの実験条件の制約を前提としており，「実際の製品に対する現実的な脅威」レベルとは大きな隔たりがある．本研究の目的は「タップ入力」自体の安全性に関する一般的な性質を調査することにあり，特定の機種に対する攻撃を狙ったものではない．今後は調査結果の進展によって，脅威レベルに応じて関係者と連携して対応を進めていく予定である．

8 まとめ

本研究では，「正規ユーザがスマートフォンにキー入力を行う際のタップ音を，攻撃者が外部のマイクで単純に盗聴する」という受動的な攻撃シナリオにおいて，正規ユーザの入力情報が攻撃者にどの程度漏れるのかについて検証を行った．本稿では，本研究の第一段階として，キー入力を PIN 入力に限定して調査を行った．タップ音の音声データのメル周波数スペクトログラム画像を入力データとして，畳み込みニューラルネットワークを用いてキー入力の識別を行った結果，静音環境下での本人-本人間の識別の場合は 90%以上の精度で PIN 入力の識別が可能であることが判明した．また，本人-他人間の識別の場合（複数名の既知のユーザのタップ音で学習した識別器を用いて，未知のユーザの PIN 入力を識別する場合）も，時として静音環境下で 50%程度の精度で PIN 入力の識別が可能であることが明らかとなった．そして，騒音環境を模擬した実験からは，タップ音をかき消すようなノイズ音の不可が攻撃の防御に貢献することが示唆される結果が得られた．

今後は，実験協力者を増やしながら，現実的な状況を考慮した実験環境での実験を繰り返し，タップ音によるキー入力の識別精度を更に精査していく．また，今回の結果を踏まえ，ノイズ付加に関する具体的な防御策についても検討を深めていく．

参考文献

[1] 本間尚文，青木孝文：知っておきたいキーワード サドチャネル攻撃，映像情報メディア学会誌，Vol.64，No.11，pp.1576-pp.1576，2010

[2] National Securiy Agency: TEMPEST fundamentals, NACSIM 5000, Feb 1982

[3] Ilia Shumailov, Laurent Simon, Jeff Yan and

RossAnderson: Hearing your touch: A new acoustic sidechannel on smartphones, arXiv : 1903.11137,2019.
[4] Li Lu, Jiadi Yu, Yingying Chen, Yanmin Zhu, Xiangyu Xu, Guangto Xue and Minglu: KeyListener: Inferring Keystrokes on QWERTY Keyboard of Touch Screen through Acoustic Signals, IEEE INFOCOM 2019.
[5] Li Zhuang, Feng Zhou and J. D. Tygar: Keyboard Acoustic Emanations Revisited, ACM Conference on Computer and Communications Security, November 2005, pp. 373-382.
[6] 大内結雲, 奥寺瞭介, 塩見祐哉, 上原航汰, 杉本彩歌, 大木哲史, 西垣正勝 : 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, 暗号と情報セキュリティシンポジウム 2020（SCIS2020)予稿集, 1E2-4, (2020)
[7] Hendrik Purwins, Bo Li, Tuomas Virtanen, Jan Schluter, Shuo-yiin Chang and Tara Sainath: Deep Learning for Audio Signal Processing, JOURNAL OF SELECTED TOPICS OF SIGNAL PROCESSING, VOL 13, NO.2, May 2019, pp. 206-219.
[8] Audacity: The Free, Cross-Platform Sound Editor, available from <http://audacity.sourceforge.net> (accessed 2020-12-6)
[9]音声資源コンソーシアム: 電子協 騒音データベース (JEIDA-NOISE),available from<http://research.nii.ac.jp/src/JEIDA-NOISE.html>(accessed 2020-12-9)
[10] Keras Documentation: Train a simple deep CNN on the CIFAR10 small images dataset(online), available from<https://github.com/keras-team/keras/blob/master/examples/cifar10_cnn.py>(accessed 2019-12-12)