M-035

機械学習を用いた スマートフォンのモーションセンサによる個人認証方式の提案

A proposal of personal authentication method by motion sensor of smartphone using machine learning

播磨尚希 † 平松耕輔 † 小林孝史 ‡
Naoki Harima Kosuke Hiramatsu Takashi Kobayashi

1. はじめに

近年普及しているスマートフォンなどのスマートデバイスの利用用途は多岐にわたり，スマートデバイスは従来型携帯電話と比べ，個人情報や鍵情報などのより重要な情報を多く保有している．そのため，従来型携帯電話に比べてより強固なセキュリティの担保が求められる．中でも利用者の個人認証は，端末内の情報の機密性を担保する上で重要な要素である．従来のスマートフォンの個人認証では，パスコード認証やパターンロック認証など，スマートフォンを利用する度にユーザが意識的に鍵情報を入力する必要があるものがほとんどであった．しかし，鍵入力が煩雑であるとユーザビリティが低下し，これを解消するために強度の低い認証コードを利用するユーザが生まれる可能性があり，Li らの研究によって携帯電話ユーザの 30%以上が PIN コード認証すら利用していないことが判明している [1]．そのため，ユーザに負担をかけない簡便かつ強力な認証方法が求められ，近年は指紋認証や顔画像認証などのユーザの負担が少ない認証手法が利用される機会が増加した．しかし，強度の高い指紋認証や顔画像認証には特殊なハードウェアを利用する必要があるため，実装コストに課題があった．一方，加速度センサと角速度センサは市場に出回っているほとんどのスマートデバイスに搭載されていることから，近年，加速度センサや角速度センサのデータを用いた個人識別手法の研究が進んでいる [2][3][4]．

また，近年スマートデバイスの個人認証に用いられているようになった生体認証は身体的特徴をもとにした生体認証（以下，身体的生体認証）に分類されるが，身体的生体認証に用いられる鍵情報は攻撃対象に接触せずとも得られる外見などの情報からの複製が容易であったり，意識を失っている状態でも認証を突破できてしまうなどの問題があった．これに対し，ユーザの行動をもとに個人識別を行う生体認証（以下，行動的生体認証）では，攻撃対象に接触せずに外見などの情報から複製することは困難であり，認証の主体が意識を保っている状態でなければ認証が行えないなどの利点がある．

本研究では，市場に出回っている殆どのスマートフォンに搭載されている加速度センサと角速度センサによって得られるユーザのロック解除に至るまでの動作データ（以下，モーションデータ）をもとに機械学習にて個人分類器を作成することにより，明示的な鍵入力プロセスを必要としない強固な行動的生体認証手法を提案する．また，畳み込みニューラルネットワーク（以下，CNN）と Long short-term memory（以下，LSTM）を組み合わせたネットワーク（以下，CNN+LSTM）による分類器を実装し，サポートベクターマシン（以下，SVM），ランダムフォレスト（以下，RF），多層パーセプトロン（以下，MLP）を用いた分類器との性能比較実験を行った．

†関西大学大学院 総合情報学研究科
‡関西大学総合情報学部

2. 従来型個人認証手法

現在スマートデバイスで利用されている個人認証手法にはパスコード認証，パターンロック認証のような非生体認証と，指紋認証，虹彩認証，顔画像認証などの身体的生体認証が存在する．パスコード認証やパターンロック認証には個人識別を可能とするのに十分なパターン数の認証キーが存在するが，実際に多くのユーザに利用される認証キーの種類は多くない．これを確かめた研究としては，実際に利用されているパターンロック認証の傾向を調査し，パターンロック認証は 3 桁の PIN コード認証ほどの複雑さしかないことを示した Aviv らの研究 [5] が挙げられる．また，パスコード認証やパターンロック認証には覗き見による認証情報漏えいの可能性がある．これらの問題を解決するため，生体認証への期待が高まっている．

現在スマートデバイス市場に投入されている生体認証手法としては，光学式指紋センサによる指紋認証，静電容量式指紋センサによる指紋認証，超音波式指紋センサによる指紋認証，虹彩認証，2 次元顔画像認証，3 次元顔画像認証などの身体的生体認証が挙げられる．新たな方式の指紋センサの導入によりその強度を高める指紋認証であるが，指紋情報が攻撃者に渡ってしまうと多くの場合で認証を突破できる [6][8]．また，従来攻撃者は攻撃対象の残留指紋などを得ることで指紋を複製していたが，starbug らは市販のデジタルカメラで撮影した指画像から指紋を複製できることを示した [7]．更に，虹彩認証や二次元顔画像認証も同様に第三者が不正に認証を突破することが示されている [9][10]．3 次元顔画像認証では，3 次元顔画像を容易に複製する術は未だ確立されていないため一般的に認証突破を危惧する必要はまだないが，一卵性双生児による認証実験では第三者によって認証が突破された [11]．

この様に複製が容易な認証情報を用いた認証や，先天的な特徴をもとにした認証では不正に認証を突破されてしまう恐れがある．加えて，これらの認証は一貫して認証の主体の意識がない状態でも認証可能である．このため，睡眠中などの意識を失っている状態や盗撮によって認証が突破される恐れがある．これらの問題を避けるためには，ユーザ行動をもとに個人識別を行う行動的生体認証が利用できる．機微なユーザ行動は攻撃者に接触すること無く外見などの情報から複製することが困難であり，後天的に特徴を変化させることに加え，認証の主体が意識を保っていなくては取得できない．また，身体的生体認証に存在した，手袋やマスク，サングラスなどによって認証に必要な部位が隠された状態では認証が行えないという問題も行動的生体認証によって回避できる．

3. 先行研究

高坂の研究 [12] では，パスコード認証方式における認証作業の煩わしさと鍵情報の自由度が制限される課題点を軽減することを目的として，スマートフォンに搭載された加速度センサと角速度センサを利用し，端末を振るモーションで個人認証を行う手法を提案した．この手法では，鍵情報と認証

モーションのパターンマッチには，それぞれのモーション間のコサイン類似度を利用した．

高坂の手法では，認証のたびに同じ動作を繰り返し入力するため，第三者がユーザの動きを見てモーションを覚えることが容易く，なりすましによる認証のリスクが高い．また，端末を利用するたびに意識的にモーションを入力する必要があるため，指紋認証や，内蔵カメラによる顔画像認証などの認証手法と比較すると認証にかかる手間が大きいことも課題であった．

4. 本研究のシステム

本研究では，強固なセキュリティと高い利便性を合わせ持つ個人認証方式の実現を目的とし，スマートフォンに搭載された加速度センサと角速度センサから取得したモーションデータをもとに学習した分類器を用いた個人認証手法を提案する．

4.1. 学習データ

学習に利用するモーションデータは，ユーザのロック解除試行からさかのぼって取り出した過去一定時間のモーションデータである．1 秒間に記録されるデータの件数を n ，認証に使用するモーションの秒数を t としたとき，1 件のデータには 3 軸の加速度と 3 軸の角速度の計 6 個の数値データが含まれるため，1 件のモーションデータは合計 $6nt$ 個の数値データで構成される．

また，得られたモーションデータは分類器が対象とするユーザ（以下，本人）とそれ以外のユーザ（以下，他人）にラベル分けして利用するが，他人のデータ件数に対して本人のデータ件数が著しく少なくなる．これによる分類機の性能低下を防ぐため，本人のデータに対して重複を許してランダムにサンプリングするランダムオーバーサンプリングを行うことで，本人データの件数と他人データの件数を同数にした．

また，分類器の性能を向上させるため，次の二種類のデータオーギュメンテーションを行った．(1) モーションデータの抽出位置を時系列方向にずらす．(2) 加速度データの大きさを増減させる．前者の手法により，データの時系列方向へのずれに対する頑健性の高い学習が期待できる．また，後者の手法により，動きの大きさの大小に対する頑健性の高い学習が期待できる．

4.2. 分類器

本手法での個人認証に用いる分類器は，本人ならば 1 を，それ以外ならば 0 を出力する一対他分類器であり，ユーザごとに個別の分類器を学習する．提案手法には順伝播型人工ディープニューラルネットワークの一種である CNN と人工回帰型ニューラルネットワークの一種である LSTM を組み合わせたネットワークアーキテクチャを用いる．CNN は画像認識分野において顕著な成功を収めているアーキテクチャであるが，複数の加速度センサから得られたデータをもとにして人間の活動状態を識別する手法を提案する Jian Bo Yang らの研究 [13] によって，多チャンネルの時系列データを取り扱うための機械学習モデルとして特徴抽出能力と識別力において SVM，k 近傍法などの手法よりも優れていると示された．また，LSTM は時系列データにもとづく分類，処理，予測によく適したアーキテクチャである．

学習に利用するモーションデータは 3 次元空間での方向を表す軸と時間を表す軸からなる 2 次元データであるとみなすことができる．この 2 次元データを CNN を用いて時間軸方向に畳み込むことで，時間軸方向にモーションデータが集約し，時間軸方向のデータ 1 件に含まれる情報量が増える．このデータを用いて LSTM の学習を行うことで，学習のステップ数が削減され，学習が高速化することが期待できる．

提案手法のネットワークは 5 層の畳み込み層と 3 層の LSTM からなるディープニューラルネットワークである．入力層に一番近い畳み込み層への入力は，加速度データと角速度データを別チャンネルに分離した $2@1800 \times 3$ という形状のデータとする．LSTM への入力は畳み込み層の第 3 層の出力を用いるようになっており，畳み込み層の第 5 層の出力データと LSTM の出力データの大きさを揃えるため，畳み込み層の第 5 層の出力を全結合層へ通し，128 個のデータに変換する．最終層の全結合層への入力は，CNN と LSTM の出力を結合した 256 個のデータとし，最終層の出力を分類器全体の出力とする．活性化関数には ReLU を使用し，最適化手法には Adam を使用した．

4.3. 利用の流れ

個人認証を行うユーザはまず，分類器を作成するためのモーションデータの登録を行う．スマートデバイスのロック解除を幾度か試行し，その際のモーションデータを記録する．その後．それらのモーションデータが学習用サーバに送られ，他ユーザのロック解除モーションデータとともに学習用サーバ上で学習に用いられ，各ユーザ毎に固有の分類器を学習する．その後，学習済みモデルをユーザに送信し，ユーザ端末上に学習済みモデルが保存される．

個人認証を利用する場面では，ユーザ端末上で常にモーションデータを記録し，ロック解除試行をきっかけにその時点から過去一定時間のモーションデータを取り出し，学習済みモデルによって本人によるロック解除モーションであるかを識別する．その結果本人であると識別されれば，認証し，ロックを解除する．なお，学習サーバにモーションデータが追加された際には，オンライン学習を行い，モデルを再配布する．

5. 評価実験

評価実験に用いるモーションデータは，端末がロックされた状態において常にモーションデータを取得，記録するアプリを 8 名の協力者にインストールして日常生活を送ってもらうことで収集したデータから，ロック解除を試行したタイミングからさかのぼって 30 秒間のモーションデータを抽出することで取得した．その結果，協力者によってデータ件数にばらつきはあるものの，合計 11817 件のモーションデータが集まった．なお，モーションデータ中には極めて短いモーションデータなども含まれているが，30 秒に満たないモーションデータは前方を 0 で埋めることによって長さを揃えた．

評価実験では，より軽量なモデルである SVM，RF，MLP による分類器との比較実験を行った．なお，比較に用いた各分類器では，グリットサーチを行い最も性能の良かったハイパーパラメータを利用した．SVM では誤分類の許容値を決定する正則化パラメータ $C = 10$，モデルの複雑さを決定するカーネルパラメータ $\gamma = 0.001$，RF では決定木の数 $n\_estimators = 30$，各決定木に使用する説明変数の数 $max\_features = 1000$，各決定木の深さ $max\_depth = None$，MLP では入力層，中間層，出力層の 3 層構造のネットワークを用い，中間層のニューロン数 $hidden\_layer\_sizes = 1000$，エポック数 $max\_iter = 10$，ミニバッチのサイズ $batch\_size = 100$，CNN+LSTM ではエポック数 $max\_iter = 30$，ミニバッチのサイズ $batch\_size = 100$ とし，その他のパラメータに関しては人手で試行した中で最も良い性能を示したものを用いた．CNN+LSTM の詳細なネットワーク図を図 1 に示す．なお，MLP，CNN+LSTM では活性化関数には ReLU，最適化手法には Adam を使用

した．

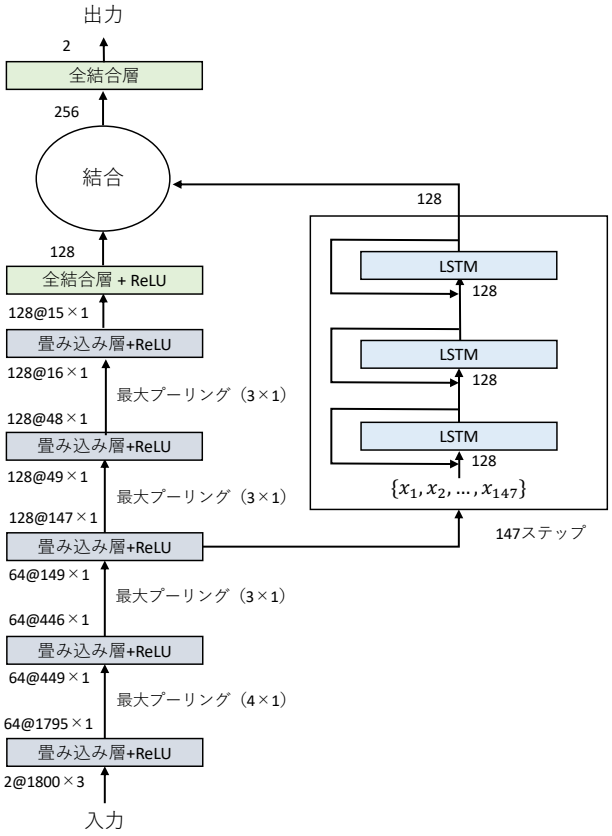

図 1: CNN+LSTM のネットワーク構造

また，実際の利用では各学習済みモデルを各ユーザの端末に配布しユーザ端末上で認証する事を想定しているが，検証においては，学習サーバ上で収集したデータセットに対するモデルの性能評価を行った．

5.1. 分類精度

評価指標として，本人を誤って拒否した割合である本人拒否率（FRR）と，他人を誤って受け入れた割合である他人受入率（FAR）の二つの指標と，その平均値である Balanced Error Rate（BER）の指標を用いる．

データセットをそれぞれ 4 分割し，分割交差検証法を用いた実験を行った．全協力者の結果の平均値を表 1 に示す．また，学習データ数と性能の関係示すグラフを図 2 に示す．

全協力者の結果の平均値に着目すると，CNN+LSTM，RF，SVM，MLP の順に分類精度が高く，提案手法である CNN+LSTM での BER は約 7.9%であり，約 92%の精度で個人認証が成功している．また，どの協力者にでも提案手法の精度が高いことがわかる．各協力者毎の結果に注目すると，最もデータ件数が多かった協力者では CNN+LSTM の BER が約 4.8%であり，平均を上回る約 95%という精度で分類ができている一方，最もデータ件数が少なかった協力者では CNN+LSTM の BER が約 13%であり，平均を下回る約 87%という分類精度となっている．このことや図 2 から概ねデータ件数と分類精度の間には相関関係があることが見て取れるので，学習データを増やすことで分類精度が向上する可能性がある．

また，学習に利用したモーションデータ中には極端に短く識別に適さないものも含まれていた．そのため，学習データを精査することで分類精度は向上すると考えられる．

表 1: 全協力者の結果の平均値

手法	FRR	FAR	BER
SVM	0.155918	0.500799	0.328359
RF	0.307450	0.189845	0.248648
MLP	0.336662	0.358135	0.347398
CNN+LSTM	**0.079885**	**0.077978**	**0.078932**

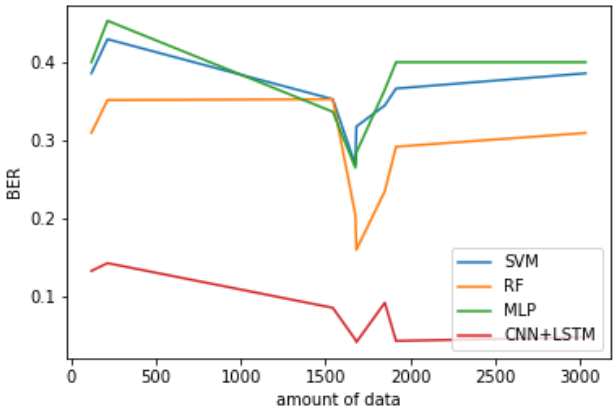

図 2: 学習データ件数とエラー率

5.2. 学習時間

性能評価を行った 4 種類の機械学習手法について，モデルの学習にかかる時間の比較を行った．表 2 に，各手法で 8 名の協力者の分類器を作成するのに掛かった時間の平均値を示す．なお，学習サーバには GPU（GTX1080）が 1 枚取り付けられており，多層パーセプトロンと CNN+LSTM の学習には chainer の GPGPU 機能を用いた．

表 2: 学習時間の平均値

Architecture	SVM	RF	MLP	CNN+LSTM
時間（秒）	39	56	115	445

ニューラルネットワークを使用した多層パーセプトロンと CNN+LSTM は他の手法と比べて学習にかかる時間が長く，特に CNN+LSTM は多層パーセプトロンの 4 倍程の学習時間がかかることが分かった．実際の運用ではニューラルネットワークモデルは逐次オンライン学習を行うことになるため，計算リソースをある程度分散することができるが，多人数の同時利用を考えると多量の計算リソースが必要になることが予想される．なお，本手法が提案する CNN+LSTM モデルはハイパーパラメータのチューニングが不十分である可能性が高く，より軽量なモデルを構築することで計算リソースの問題は緩和できる可能性がある．

6. モーションデータとモデルの可視化

モデルの学習結果を理解するため，学習に利用したモーションデータの可視化と，学習済み CNN+LSTM モデルの判断根拠を可視化する．

評価実験に用いたモーションデータ中から二名の協力者のデータをそれぞれ 1 件ずつサンプリングし，速度変化を時系列順に 3 次元空間にプロットしたグラフを図 3 に示す．図中の原点を初期速度とし，そこからの相対速度の変化が線として示されることで端末の 3 次元空間での動きを視覚的に解釈することができる．

また，協力者 A を識別するように学習した CNN+LSTM モデルにモーションの可視化に用いた二つのデータを分類させた際の判断根拠を Smooth Grad によって可視化したもの

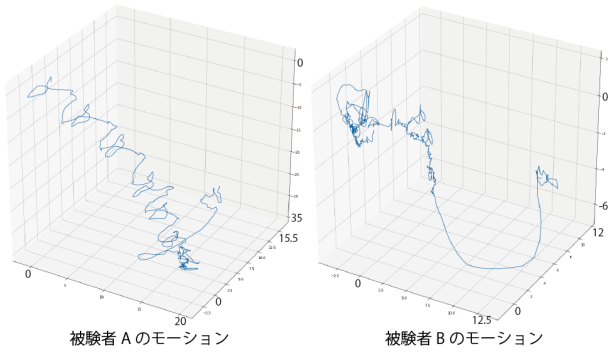

図 3: モーションの可視化

を図 4 に示す．

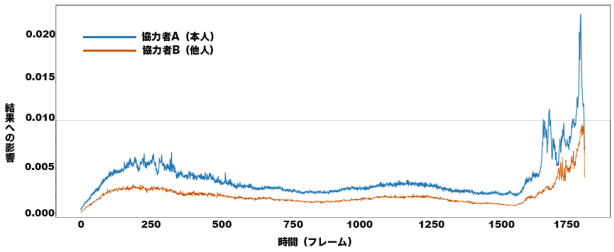

図 4: Smooth Grad による判断根拠の可視化

横軸は時間（フレーム）を示しており，縦軸はその瞬間のモーションがラベル付けに与えた影響の強さを表している．どちらのデータにおいても，モデルが強く注目しているのは末尾から数秒間となっているが，それ以外の部分も小さいながら予測結果へ影響を与えていることが読み取れる．この可視化結果を活用することで，入力データの長さの最適化といったモデルのチューニングを効率的に行うことができる可能性がある．

7. おわりに

高度な機密性と利便性が同時に求められるスマートデバイスの個人認証において，加速度センサと角速度センサの値をもとにした行動的生体認証を用いることで明示的な鍵入力プロセスを必要としない個人認証が実現可能であることを示し，SVM，RF，MLP と比較することで，精度において CNN+LSTM が優れていることを示した．また，本提案手法では既存の殆どのスマートデバイスに搭載されている加速度センサと角速度センサを用いるため，追加で特殊なハードウェアを搭載すること無く利用可能である．

実利用に際しては，分類精度と学習時間の面で課題が残るが，データ数と分類精度や学習時間の関係，学習データの精査，モデルのチューニングなどによって解決可能な可能性がある．また，実機における盗み見耐性の評価や，認証対象数に応じた分類精度の評価などが課題として残される．また，モデルの精度を向上するためにはオンライン学習によって逐次学習を続けることが必要であると考えられるが，提案手法によって認証が失敗した直後の他の手法での認証の成否をもとにしたモデルの再学習や，他人データが追加される状況でのオンライン学習についての検証も必要である．

参考文献

[1] L. Li, X. Zhao, G. Xue, ”Unobservable re-authentication for smartphones”, Proc. 20th Netw. Distrib. Syst. Secur. Symp. (NDSS), vol. 13, pp. 1-16, 2013.

[2] Jennifer R. Kwapisz , Gary M. Weiss , Samuel A. Moore, “Cell phone-based biometric identification”, 2010 Fourth IEEE International Conference on Biometrics: Theory, Applications and Systems (BTAS), 2010.

[3] Claudia Nickel , Tobias Wirtl , Christoph Busch, “Authentication of Smartphone Users Based on the Way They Walk Using k-NN Algorithm”, 2012 Eighth International Conference on Intelligent Information Hiding and Multimedia Signal Processing, 2012.

[4] 佐藤 悠祐, 神山 剛, 福田 晃, 小口 正人, 山口 実靖 , “スマートフォンに搭載されている加速度センサー情報を用いた 2 クラスの分類による身長推定”, 研究報告コンシューマ・デバイス＆システム（CDS）2018-CDS-22 巻 10 号 pp.1-6，2018

[5] Adam J Aviv, Devon Budzitwoski, Ravi Kuber, “Is bigger better? comparing user-generated passwords on 3x3 vs. 4x4 grid sizes for android ’s pattern unlock.”, Proceedings of the 31th Annual Computer Security Applications Conference, ACSAC ’15, 2015..

[6] starbug, “iPhone 5s Touch ID hack in detail”, https://www.heise.de/multimediadatei/iPhone-5s-Touch-ID-hack-in-detail-1965628.html，2019 年 6 月参照.

[7] starbug, “CCC — Fingerprint Biometrics hacked again”, https://www.ccc.de/en/updates/2014/ursel，2019 年 6 月参照.

[8] darkshark, “I attempted to fool the new Samsung Galaxy S10’s ultrasonic fingerprint scanner by using 3d printing. I succeeded. - Album on Imgur”, https://imgur.com/gallery/8aGqsSu#WlRksn6，2019 年 6 月参照.

[9] iDeviceHelp, “Galaxy S8 Facial recognition can be bypassed With a Photo DEMO”, https://www.youtube.com/watch?time_continue=2&v=uS1NmvJvHNk，2019 年 6 月参照.

[10] starbug, “media.ccc.de - Die Sendung mit dem Chaos - Iris-Scanner im Samsung Galaxy S8”, https://media.ccc.de/v/biometrie-s8-iris-fun，2019 年 6 月参照.

[11] Wall Street Journal, “iPhone X Review: Testing (and Tricking) FaceID”, https://www.youtube.com/watch?v=FhbMLmsCax0，2019 年 6 月参照.

[12] 高坂賢佑, 平松耕輔, 小林孝史, “スマートフォンのモーションセンサを利用した個人認証アプリケーションの開発”, FIT2016 第 4 分冊 pp.149-150，2016.

[13] Jian Bo Yang, Minh Nhut Nguyen, Phyo Phyo San, Xiao Li Li, Shonali Krishnaswamy, “Deep Convolutional Neural Networks On Multichannel Time Series For Human Activity Recognition”, IJCAI 2015 pp.3995-4001，2015.