深層学習を用いたキーボード入力とマウス操作情報による個人識別

木村 悠生[1,a)]　猪俣 敦夫[2,b)]　上原 哲太郎[3,c)]

概要：近年，情報通信機器を用いたサービスが広がるにつれ，その利用等の場面で操作者個人の識別が必要な場面が増加している．これらのサービスにおいては，クレデンシャルの利用によるユーザ識別が広く行われているが，不正ログインによるなりすましや，故意の認証情報譲渡による代理ログインなどの偽装行為により，ログイン情報のみでは個人識別の信頼が保てない場合がある．オンラインにおける資格試験や各種学校におけるリモート試験の実施など，サービスの性質によってはより厳密な個人識別のニーズがあることを踏まえ，本研究ではパッシブ認証の一種として，情報通信機器の操作特徴による個人識別を採用する．本論文では，日本語入力時の PC 操作情報を深層学習を用いて既知のユーザと照合し，操作ユーザの特定を行うための手法を提案する．10 人の被験者の操作状況を用いて実験した結果，従来手法より少ない操作時間で操作者を特定することが可能になった．

キーワード：動的生体識別，深層学習，個人識別

Personal Identification by Keyboard Input and Mouse Operation Information Using Deep Learning

KIMURA Yuuki[1,a)]　INOMATA Atsuo[2,b)]　UEHARA Tetsutaro[3,c)]

Abstract: In recent years, as services using information and communication devices expand, the need for identification of individual operators has increased in situations such as use of such services. Although credential-based user identification is widely used for these services, there are cases in which login information alone cannot be trusted for personal identification due to impersonation by unauthorized login or proxy login by intentional transfer of authentication information. In addition, there is a need for more strict personal identification depending on the nature of the service, such as online qualification examinations and remote examinations in various schools. In this work, we focus on personal identification based on operational characteristics of information communication devices as a type of passive authentication. In this paper, we propose a method to identify an operating user by matching PC operation information during Japanese input with known users using deep learning. Experimental results using the operating conditions of 10 subjects showed that it was possible to identify the operator in less operating time than with conventional methods.

Keywords: Dynamic Biometrics，Deep Lerning，Individual Identification

1. はじめに

様々なサービスにおいて，利用者の認証・識別手段として，ユーザが自発的に認証・識別情報を提供し，提供された情報からユーザを認証・識別するアクティブ認証が行わ

[1] 立命館大学大学院情報理工学研究科
Graduate School of Infomation Science and Engineering, Ritsumeikan University
[2] 立命館大学総合科学技術研究機構・大阪大学
Research Organization of Science and Technology, Ritsumeikan University/Osaka University
[3] 立命館大学情報理工学部
College of Information Science and Engineering, Ritsumeikan University
[a)] ykimura@cysec.cs.ritsumei.ac.jp
[b)] inomata.atsuo.cysec@osaka-u.ac.jp
[c)] t-uehara@fc.ritsumei.ac.jp

れている．現状では，利用者 ID とパスワードによる記憶による認証・識別や，指紋や顔による生体認証・識別，またワンタイムパスワードを用いた所有による認証などが行われている．しかし，ユーザが自発的に認証・識別情報を提供するアクティブ認証では，利益を得ることを企図した利用者が悪意を持って ID やパスワードといったクレデンシャルを他人に流出させたり，虚偽の識別情報を提供した際に正確に利用者を識別できない場合がある．一方で，しばしばアクティブ認証と比較されるパッシブ認証では，普段の利用者の行動パターンなどをプロファイルとして収集することで利用者を識別するため，ユーザの自発的な情報提供を必要とせず，流出や偽装も難しい．また，所有による認証は適正な管理を行わなければ不正利用が可能であり，生体認証は多くの場合専用の入力装置が必要であるという問題点もある．しかしパッシブ認証では，行動に着目する限り，厳重な管理も不要であり，専用の入力装置も不要である．

そのため本研究では，サービスにおけるパッシブ認証の重要性を評価し，利用者自身が悪意を持ってクレデンシャルを流出させ，または偽装した場合の対策として，偽装及び複製が難しい行動的特徴を用いた生体認証技術に着目した個人識別を試みる．特に，専用の入力機器に依存しないように，マウス及びキーボードの操作特徴の利用という，多くの PC に搭載されている入力装置の利用による識別方法を採用する．さらに，実験ではこれらの操作特徴を，事前に収集した既知のユーザのプロファイルと照合し，操作ユーザの識別を行う．

人のキーボード及びマウス操作の時系列データには個人差があり，この個人固有のパターンに基づいて個人識別をする手法はキーストロークダイナミクス及びマウスダイナミクスと呼ばれる．キーストロークダイナミクス及びマウスダイナミクスより得られる情報は，PC を使用する際にマウスやキーボードを利用することは必須であることから，確実に収集可能な識別情報であると考えられる．更に，キーストロークダイナミクス及びマウスダイナミクスによる識別は，インターネットプロトコルに関わる情報以外を用いる識別方式である．したがって，インターネットを介す必要がないだけでなく同じデバイスを利用していても利用者によって操作の癖が出るため，デバイス自体が不正利用されても本識別方法を活用することが可能である．

本論文の構成について述べる．2 章では，研究に至る経緯を述べるとともに，先行研究について，3 章では，提案手法について，4 章では，実験手法及び実験結果について，5 章では，実験結果に考察を示す．最後に 6 章では，本論文の総括と今後の展望を述べる．

2. 研究背景

2.1 研究目的

様々なサービスの利用に際して，ユーザの識別で頻繁に用いられる要素として，識別子 (ID) とパスワードがある．しかし，識別子による識別は，悪意あるユーザによる盗用や，ユーザ本人による他人への譲渡の増加によって確実とは言えなくなっている．また，近年では ID とパスワードの他に，指紋や虹彩などの生体情報や，ワンタイムパスワードなどを活用した多要素認証も導入が進んでいる．しかし生体認証では一般に，専用の入力機器が必要であり，ワンタイムパスワードも盗用のリスクに晒されているため，利便性や安全性の観点から不十分である．本研究の目的は，悪意ある利用者によって識別や認証が回避または偽造された場合に，悪意ある利用者の能動的な識別情報の提供による識別に依存することなく，かつ専用の入力機器を用いることなく利用者の識別を行うことにある．本研究では，先行研究において長時間の操作が必要であった操作について，深層学習を活用することによって，短い操作時間で利用者識別を可能にすることを目標とする．

2.2 アクティブ認証

アクティブ認証とは，ユーザ自身による能動的な識別情報の提供によってユーザを認証・識別する手法である．従来の ID・パスワードによる認証や指紋認証，ワンタイムパスワードを用いた認証など，サービスに直接関係ない情報をユーザが入力する必要がある認証・識別手法がアクティブ認証に該当する．ユーザに悪意がない場合本人拒否率や他人受入率は極めて低い一方で，複製可能である場合が多いため，他人への譲渡や偽装が行われるリスクが問題点として挙げられる．

2.3 パッシブ認証

パッシブ認証とは，ユーザ自身による能動的な識別情報の提供を必要としない認証・識別手法である．一般に，IP アドレスやブラウザ情報，サービスへのログイン後の行動特性を用いてプロファイルを作成し，プロファイルから逸脱しないユーザを本人として認証・識別する．複製や偽造が難しく，ユーザの負担なく収集できることがメリットである一方，プロファイルの作成手法及び閾値によっては本人拒否率及び他人受入率が高くなることが課題である．

2.4 生体識別

生体識別とは，顔や指紋，歩き方など人間の生体に関わる情報を用いてユーザを識別する手法である．なりすましや紛失及び盗難のリスクが少ない反面，生体情報の経年的変化による識別精度の低下や，流出した際の変更の難しさ

が弱点として挙げられる．生体識別の中でも，行動的特徴を用いるものは行動生体識別と呼ばれる．

行動生体識別とは，筆跡情報や歩行姿勢及びタイピング速度などの行動的特徴の情報を用いてユーザを識別する手法である．他の生体識別手法に比べてなりすましの難易度は高い反面，人間の行動的特徴は一定でないため，識別精度の低下による本人拒否率や他人受入率の上昇が課題である．多くの場合，機械学習を利用してユーザの行動的特徴と予め登録されている行動的特徴を突合し，一致度を一定の閾値で区切ることでユーザの識別を行っている．

2.5 関連研究

北條ら [1] は，プロファイルを用いたパッシブ認証による識別手法として，ブラウザフィンガープリンティングによる端末識別を提案した．利用者の自発的な情報提供を必要とせずに端末を高精度で分類できる画期的な成果を示したが，Computer Based Testing(CBT) 等の同一ないしは同様の端末を複数のユーザが操作するシーンにおいて識別が難しい．

粕川ら [2] は，コンピュータログイン時に行うクレデンシャル入力について，事前に任意のキーが打たれてから次のキーが打たれるまでの時間を打鍵間時間として測定し平均を算出し，それらとクレデンシャル入力時の打鍵間時間との差が一定の範囲内であれば本人であると認証する手法を提案した．アクティブ認証とパッシブ認証を併用する認証手法の提案であり，かつ明解な判断基準を提案した手法であるが，本人拒否率が 35.5%となり，識別手法としての実用性には乏しい．

山田 [3] は，被験者として 47 人の学生から 270 分間あるいは 360 分間のキーボード及びマウスの操作ログデータを収集し，平均 1 分未満で筆者自身の操作か否かを判断する手法を提案した．サービス利用開始時に認証を行った後も継続的に認証を行うために，継続認証アルゴリズムとして提案された手法であるが，筆者自身の操作が極めて特異である可能性が排除できない．また，47 人の被験者の操作特徴は「コンピュータリテラシー」及び「Java プログラミング」の授業中に取得されているが，山田の操作特徴は別日に 1 週間分の日常業務の操作ログデータを収集して作成されたものであり，操作内容の差異によって分類された可能性もある．そのため本研究では被験者の操作内容を統一し，同一操作での識別を図る．

佐村ら [4] は，112 人の被験者に 5 分間の web 上でのタイピング試験を課してキーボード操作ログデータを収集し，1 つまたは連続する 2 つの打鍵の押下・離上時間を用いてユーザを識別する手法を提案した．実験の結果，打鍵速度によってばらつきはあるものの，90%以上の高い精度での識別が可能であったと述べている．

櫻井ら [5] は粕川らや佐村らの研究を応用し，中間層が一層のニューラルネットワークを用いてユーザごとに学習する手法を提案した．これにより，本人拒否率が 0.17%，他人受け入れ率が 2.38%と，佐村らの提案手法よりさらに高精度での識別が可能であることを示した．

伊藤ら [6] は，スマートフォンのフリック入力動作に着目したプロファイリングを行い，ユーザの操作特徴による継続認証手法を提案した．実験の結果，9 割程度の精度で識別が可能であることを示した．

しかし，佐村ら及び櫻井ら，また伊藤らの研究はいずれも 1 回の入力にあたって日本語 300 字から 500 文字程度の入力を課しており，短答式の CBT 等，入力文字数が少ない操作で識別を行う場合，識別されるべきユーザ全員から規定文字数以上の入力を収集できない場合がある．実際に，CBT に拘らず一般的な短答式試験で課される設問として，独立行政法人情報処理推進機構が主催する応用情報技術者試験では，令和 4 年度春期試験の午後試験 [7] で短文での解答を求める設問の約 80%が漢字を含め 15 文字から 40 文字を字数上限とするものであった．そのため，15 文字の入力から個人を識別することを示すことができれば，短答式の CBT でもユーザの識別が可能になると仮定し，本研究では 1 回の入力あたりの日本語文字数を 15 文字程度に抑え実験を行う．

3. 提案手法

本研究では，キーボード及びマウスの操作ログデータを深層学習を用いて学習することでユーザを識別可能な学習器を構築する手法を提案する．

3.1 使用する特徴点

キーボード操作ログからの特徴抽出は，佐村らの先行研究が一定の成果を得ている，1 つまたは 2 つの打鍵についての特徴点とローマ字表記の違いによる入力の差異を特徴として採用する．ローマ字表記の違いとは，表音文字を用いる日本語における訓令式とヘボン式による入力方法の違いであり，たとえば “し”という文字は訓令式では “si”だがヘボン式だと “shi”であるし，“つ”という文字は訓令式では “tu”だがヘボン式だと “tsu”である．大戸の研究 [8] で指摘されている通り，ユーザが訓令式を用いるかヘボン式を用いるかは個人差が大きく，また通常一つの文字に対して訓令式とヘボン式を併用する場合は少ない．そのため，本研究ではローマ字表記の違い及び，訓令式とヘボン式で差異のあるもののうち，日本語入力で頻出である “し”の入力方法について特徴点として採用した．一般に機械学習では，次元数が多い場合に過学習や学習効率の低下が問題として挙げられる．そこで今回は，一つの打鍵についての特徴量では，頻出である母音と “n”，及び変換の際に用いる “space”が押されてから離されるまでの時間を用いる．また同様に，二つの打鍵についての特徴量では，佐村らの先

行研究で採用されていた “ka”，“no”，及び本研究で用いる入力で頻出である “ar” を採用する．特徴点の一覧をキーボード特徴点と呼び，表 1 に示す．また，抽出された特徴量をキーボードログと呼ぶ．

表 1 キーボード操作ログから収集する特徴点
Table 1 Feature Points Collected from Keyboard Operation Logs

表記	説明
ht	1 つ目のキーが押されてから離されるまでの平均時間． 頻出である a，i，u，e，o，n，space で取得する
pp	1 つ目のキーが押されてから 2 つ目のキーが押されるまでの平均時間． 頻出である ar，no，ka で取得する
rp	1 つ目のキーが離されてから 2 つ目のキーが押されるまでの平均時間． 頻出である ar，no，ka で取得する
rr	1 つ目のキーが離されてから 2 つ目のキーが離されるまでの平均時間． 頻出である ar，no，ka で取得する
pr	1 つ目のキーが押されてから 2 つ目のキーが離されるまでの平均時間． 頻出である ar，no，ka で取得する
romaji	ローマ字入力が訓令式かヘボン式かを記録する 頻出である 「し」(shi/si) で取得する

マウス操作ログからの特徴量抽出は，マウスカーソルの画面上の絶対座標の変化の絶対値を 10 ミリ秒ごとに取得し，時系列データとして抽出する．本研究では抽出された時系列データをマウスログと呼ぶ．また，キーボードログとマウスログを総称して，入力ログと呼ぶ．

3.2 データセットの作成

本研究では，キーボード及びマウスログを取得する際に，データ不足による過学習を防ぐため，同一ユーザ間で任意の 2 つの入力ログを組合わせ，サンプルを作成する．任意の 2 つのサンプルを結合した組を用意し，データセットとして利用する．二つの入力ログを組み合わせてサンプルを作成することで，n 個の入力ログから ${}_nC_2$ 個のサンプルを作成することが可能となる．データセット作成時も同様に，m 個のサンプルから ${}_mC_2$ 個のデータセットを作成することが可能となる．これにより，少ない入力ログからより多くのデータセットを作成することが可能になる．一般に深層学習の精度はデータ数が増加するほど向上するため，少ない入力ログをそのまま学習する場合と比較して，深層学習の精度向上が期待できる．

データセットに付与する教師データは，結合した 2 つのサンプルが同一ユーザのものであれば正解ラベルを，異なるユーザのものであれば不正解ラベルを付与する．

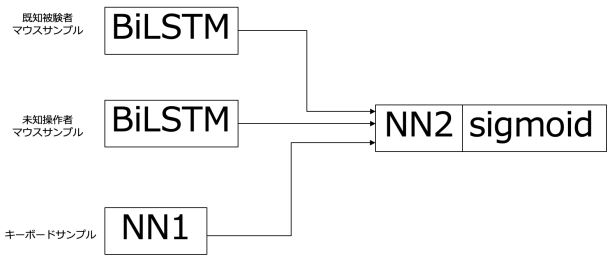

図 1 学習器の構成
Fig. 1 Structure of the Learner

3.3 学習器の構成

データセットの入力から処理及び結果の結合の流れを図 1 に示す．時系列データと統計データを効率よく処理するため，複数の構造のニューラルネットワークを組み合わせ，予備実験による比較検討の結果，図 1 の構造を採用した．

データセットの中で組み合わされているサンプルのうち，マウスサンプルはそれぞれ別の Bidirectional LSTM(BiLSTM) を用いて学習を行う．キーボードサンプルは，Neural Network 1(NN1) を用いて学習を行う．それぞれの結果を Neural Network 2(NN2) を用いて結合し，sigmoid 関数の出力を 0.5 を閾値として二値化を行う．各層は全結合層であり，損失関数は交差エントロピー関数，最適化関数には Adam を使用した．学習に関しては誤差逆伝播法を用いた．学習器のハイパーパラメータを表 2 に示す．

表 2 ハイパーパラメータ
Table 2 The Hyperparameter

	BiLSTM	NN1	NN2
隠れ層の数	1	2	1
隠れ層の次元	256	80，40	4
ドロップアウト	0.6	0.7	0

4. 実証実験

本章では実験手法とその実装，そして実験結果について述べる．

4.1 実験手法

4.1.1 実験 1

実験 1 では，無作為に抽出した 2 つのサンプルについて，操作者が同一か否かを判別することを目的とする．操作者が既知のデータのみを用いてデータセットを作成し，学習器に学習させる．教師データはデータセット作成時の通り，組み合わせた 2 つのサンプルの操作者が同一か否かとする．データセットは学習精度向上のため，正解ラベルと不正解ラベルの数が均等になるように調整し作成する．学習の際は過学習を防ぐため，作成したデータセットのうち 20%を検証用として隔離し，データセットの 80%を利用し

表 3　利用諺リスト
Table 3　Proverb List

百聞は一見に如かず
二兎追う者は一兎も得ず
塵も積もれば山となる
開いた口が塞がらない
後は野となれ山となれ
犬も歩けば棒に当たる
井の中の蛙大海を知らず
溺れる者は藁をも掴む
風が吹けば桶屋が儲かる
壁に耳あり障子に目あり
昨日の敵は今日の味方
清水の舞台から飛び降りる
結構毛だらけ猫灰だらけ
朱に交われば赤くなる
心頭滅却すれば火もまた涼し
捨てる神あれば拾う神あり
飛んで火に入る夏の虫
二度あることは三度ある
腹が減っては戦はできぬ
火のないところに煙は立たない
下手な鉄砲も数打ちゃ当たる
盆と正月が一緒に来たよう
無理が通れば道理が引っ込む
安物買いの銭失い
笑う門には福来る

て学習を行い，検証用データセットで精度の算出を行う．

4.1.2　実験 2

実験 2 では，実験 1 で作成した学習器を用いて，操作者が不明のサンプルが，いずれの既知操作者の操作であるかを判別することを目的とする．サンプルを結合しデータセットを作成する際に操作者が既知のサンプルと操作者が不明のサンプルを結合し，各既知操作者のサンプルとの一致度の高低で操作者を特定することを試みる．

4.2　実装

本章では操作ログ収集のための実験環境について述べる．

4.2.1　操作ログ収集システムの実装

本節では，操作ログ収集システムについて説明する．本実験では，デバイス間のインターフェースの差異による収集情報への影響を防ぐため，同一端末を用いた実験システムを用いる．

CBT 等での利用シーンを想定した少ない文字数での識別を図るため，入力文書は，日本の諺のうち 10 音以上 20 音以下のものを 25 句用い，被験者は漢字への変換も含め完全に一致する文章を入力する．利用した諺を表 3 に示す．変換には MicrosoftIME を利用し，MicrosoftIME の学習機能は無効に設定する．被験者は，同一の 25 句を 3 日以上間隔を空けて 2 回入力する．

4.3　実験結果

4.3.1　識別精度算出に使った指標

学習結果を，学習器が出力した予測値と教師データに基づいて表 4 のように分類する．

表 4　データの分類
Table 4　Data Classification

	操作者が同一と予測	操作者が異なると予測
操作者が同一	TP	FN
操作者が異なる	FP	TN

識別精度では，学習器の予測値と教師データに基づき，Precision, Recall, Accuracy, F_1 値, 本人拒否率 (FRR), 他人受入率 (FAR), Balanced Error Rate(BER) を評価指標として使用する．各評価指標の算出式は以下の通りである．

$$\text{Precison} = \frac{\text{TP}}{\text{TP}+\text{FP}}$$
$$\text{Recall} = \frac{\text{TP}}{\text{TP}+\text{FN}}$$
$$\text{Accuracy} = \frac{\text{TP}+\text{TN}}{\text{TP}+\text{FP}+\text{TN}+\text{FN}}$$
$$\text{F}_1 = \frac{2\times\text{Precison}\times\text{Recall}}{\text{Precison}+\text{Recall}}$$
$$\text{FRR} = \frac{\text{FN}}{\text{FN}+\text{TP}}$$
$$\text{FAR} = \frac{\text{FP}}{\text{TN}+\text{FP}}$$
$$\text{BER} = \frac{1}{2}(\text{FRR}+\text{FAR})$$

作成した学習器の予測値とはデータセットによる判定結果，教師データは結合した 2 つのサンプルの操作者の一致有無とする．また，実験 2 で評価に用いる一致率とは，学習器が任意の 2 者を同一操作者だと判断した割合であり，被験者 1 から被験者 10 までの各被験者と未知の操作者 X のサンプルを結合した際の結果を以下の算出式で算出する．

$$\text{一致率} = \frac{\text{TP}+\text{FP}}{\text{TP}+\text{FP}+\text{TN}+\text{FN}}$$

4.3.2　実験 1 結果

実験 1 では，提案手法の学習器を用いて学習を行った．結果を表 5 に示す．

表 5 より，Precision・Recall とも 0.92 以上であり，FRR・FAR とも 0.09 以下であった．この結果から，任意の 2 つのサンプルについて，深層学習を用いて分類を行うことで，操作者が同じであるか否かを良好に判定できると言える．

4.3.3　実験 2 結果

実験 1 で作成した学習器を用いて実験を行った．既知操作者との一致率の高低で，未知の操作者 X がどの既知操作者と一致するかを検討する．

表 5　実験 1 結果
Table 5　Results of the Experiment 1

	値
Precison	0.9209
Recall	0.9363
Accuracy	0.9277
F_1	0.9285
FRR	0.06375
FAR	0.0808
BER	0.0723

実験結果は表 6 の通り．便宜上被験者に通し番号を振って識別する．また，最も既知被験者との一致率が高い数値を太字にする．

表 6 より，すべての場合で，操作者が不明の場合でもどの被験者の操作かを同定することが出来ているということがわかる．

5. 考察

5.1 実験 1

実験 1 では，表 5 で示したとおり，任意の二操作者が同一であるか否かを，90%以上の精度で分類できた．この要因をデータセットの各特徴点の重要度から考察する．

重要度は，Permutation Importance を利用して識別に影響のある特徴点を調べる．Permutation Importance は，Fisher ら [9] によって 2019 年に提案された，機械学習モデルの各特徴量の効果を測定する手法である．通常の学習結果に対し，各特徴量をランダムにシャッフルしたときに，どれだけ分類精度が悪化するかを算出することで各特徴量の重要度を測定する．

本手法を用いて，重要であるとされた特徴点上位 5 位までを表 7 に示す．

これらの特徴点が識別に影響している要因として，出現頻度が突出して高いことが考えられる．特に “space” と “a” についてはその他のキーと比較しても倍以上の出現回数であった．それにより，概ねどのサンプルでも収集できた特徴点であったことから，識別に大きく寄与したと考えられる．また時系列情報であるマウスサンプル以外の標準偏差のうち，大きかったもの上位 5 点は表 8 の通りである．

以上より，識別に大きく関わっているのは主に ht であり，2 連続打鍵及びマウスの特徴点は識別に大きく関わってはいないものの，精度向上に寄与していると考えられる．

5.2 実験 2

実験 2 では，不明な操作者の操作特徴がどの既知の操作者の特徴と一致するかについて，一致率が最も高い既知操作者が未知操作者の正体であるとすることで，操作者を特定することが出来た．既存手法より短時間のログ収集で識別が可能になった要因として，複数ログの組み合わせにより学習データが増加したことを挙げる．未知の被験者を各既知被験者と突合し，一致率の高低でどの既知被験者の操作であるかを判断するという実験 2 において，10 名すべてが本人との一致率が一番高かったという結果により，本研究の目的である，既知操作者の短文入力時の操作ログを用いた未知操作者の特定について，達成することができた．

一方で本人ではない既知操作者との一致率に目を向けると，30%近くなっている組み合わせもあり，多くのユーザで高くなっていることがわかる．また，最も一致率の高い既知操作者に目を向けても，特に未知の操作者として入力した際の被験者 6, 7 で顕著であるが，一致率が低くなっていることがわかる．これは，操作特徴の個人間差を，個人の各回の操作ごとの軽微な差異が上回っているからであると考えられる．また実験 1 の学習の際に，正解ラベルと不正解ラベルの数を均等にするため不正解ラベルを持つデータを大きく削減したことも，本人ではない既知操作者との一致率の上昇の一因として考えられる．

6. まとめ

本研究では深層学習を用いて，キーボード及びマウス操作のログデータを分析し，操作者が不明であるログデータから，操作者を特定する手法の提案及び検証を行った．データセットを工夫して作成し，適切な前処理を行ってから深層学習を用いたことにより，少ない特徴量によって識別を行うことが可能になった．不明な操作者を既知の操作者と突合し，操作者を同定する手法の有効性を評価するために実施した実験 2 の結果より，提案手法で操作者を同定することが可能であることが判明した．このことから，日本語 10 音程度の文字列でも，5 分ほど収集を続ければ過去に登録されたキーボード及びマウス操作のログデータと照合し，操作者を同定することが可能であることを示した．

一方で，実験 2 で見られるような他人との一致率の高さが課題として残る．特徴点を減らすことで高速かつ簡単な識別が可能となるが，その分操作者間の個人差よりも，操作時の誤差のほうが大きくなってしまうことが考えられる．10 人程度であれば個人間の特徴の有意な差を機械学習によって見出すことが可能であるが，実際の運用を想定すれば更に大人数の中から識別しなければいけない状況もあり得る．今後の課題として，学習データを増加させるなどして，より識別精度を向上させるべきであろう．

謝辞　実験への協力を快諾して頂いた立命館大学学友会中央事務局の同期や後輩の皆様，数多くの助言やご指摘を頂いたサイバーセキュリティ研究室の皆様に心から感謝します．

参考文献

[1]　北條大和，細谷竜平，齋藤祐太，齋藤孝道：DNN を用いたパッシブフィンガープリンティング手法の提案と実

表 6　実験 2 結果

Table 6　Results of Experiment 2

既知 \ 未知	被験者 1	被験者 2	被験者 3	被験者 4	被験者 5	被験者 6	被験者 7	被験者 8	被験者 9	被験者 10
被験者 1	**68.84**	0.75	0	1.03	10.13	1.22	7.06	0	0.15	16.20
被験者 2	1.81	**68.60**	0.02	0.01	0.27	0.02	7.19	0.02	4.60	2.18
被験者 3	0.09	2.05	**64.89**	1.60	0	9.55	5.57	0.02	26.45	1.06
被験者 4	3.36	0	3.48	**96.82**	0.60	17.94	9.82	0	0.05	0.34
被験者 5	8.19	0.10	0	0.25	**86.24**	1.03	4.19	0	0	4.26
被験者 6	0	0	9.10	0.60	0	**59.03**	0.85	0	0.25	0
被験者 7	0	2.75	5.39	0.27	0	4.50	**59.74**	0.01	0.15	1.70
被験者 8	1.25	1.25	0.06	0.08	0	0	0	**81.38**	0.9	18.04
被験者 9	1.75	17.10	2.08	0.38	0.01	0.17	3.65	0.03	**66.25**	3.88
被験者 10	4.26	1.10	0	0	0	0	0	28.15	0.80	**75.08**

表 7　識別において重要な特徴点

Table 7　Important Features in Identification

rank	特徴点
1	「space」の ht
2	「a」の ht
3	「i」の ht
4	「o」の ht
5	「e」の ht

表 8　標準偏差が大きい特徴点

Table 8　Feature Points with Large Standard Deviation

rank	特徴点
1	「space」の ht
2	「し」の入力方法
3	「u」の ht
4	「i」の ht
5	「a」の ht

装, 第 81 回全国大会講演論文集 Vol. 2019, pp. 445-446, (2019).

[2] 粕川正充，角田博保，森裕子：アルペジオ打鍵列を利用した個人認証手法の提案．情報処理学会論文誌 Vol. 34, pp. 1198-1205，(1993).

[3] 山田 猛矢，行動的特徴を用いた継続認証アルゴリズム DPTM の有用性．第一工業大学研究報告，第 30 号，pp. 15-20，(2018).

[4] 佐村敏治，西村治彦，非定型な日本語文入力におけるキーストロークダイナミクス識別，システム制御情報学会論文誌，Vol. 22, No. 4, pp. 145–153, (2009).

[5] 櫻井啓志，宮本貴朗，青木茂樹，岩田基，汐崎陽：ニューラルネットワークを用いたキーストローク特徴によるユーザ認証．電子情報通信学会技術研究報告．WBS，ワイドバンドシステム：IEICE technical report 110(444)，pp. 213-220，(2011).

[6] 伊藤駿吾，白石陽：スマートフォンのフリック入力方式の特徴に着目した継続認証手法の提案，第 25 回マルチメディア通信と分散処理ワークショップ論文集 Vol. 2017 pp. 1-8, (2017).

[7] 令和 4 年度春期応用情報技術者試験午後問題，入手先〈https://www.jitec.ipa.go.jp/1_04hanni_sukiru/mondai_kaitou_2022r04_1/2022r04h_ap_pm_qs.pdf〉，(参照 2022-8-8)

[8] 大戸あや香．ローマ字の規範意識と実態．日本文学ノート，Vol. 55，pp1-21，(2020).

[9] Aaron，F., Cynthia，R., Francesca，D.：All Models are Wrong, but Many are Useful: Learning a Variable's Importance by Studying an Entire Class of Prediction Models Simultaneously，Journal of Machine Learning Research 20 (177), 1-81, (2019)