Vol.2019-MBL-92 No.3
Vol.2019-CDS-26 No.3
2019/8/29

加速度センサを用いる運転者識別と詐称者検出手法

森 優稀[†1,a)]　内藤 克浩[†2,b)]

概要：近年，自動車の盗難が大きな問題となっている．盗難問題への対策として運転者の認識が挙げられる．指紋に代表される生体情報を用いた認識技術は，高い精度で運転者を認識し盗難防止システムとして適している．しかし一度認証を行った後に発生する盗難への対応は不可能である．認証後に発生する盗難に対応するためには，継続的に運転者を識別し，詐称者を検知する必要がある．本稿では，運転者の行動を監視する連続識別手法と詐称者の検知手法を提案する．提案手法は，加速度データによる運転者の行動を識別する分類器に長短記憶（LSTM : Long Short Term Memory）を使用し，運転者を分類する．また，詐称者の検知は，運転者識別の分類結果の統計情報を特徴量とし，登録済み運転者と詐称者の二値分類をすることで行う．実験結果は，実車の加速度センサは実時間運転者の挙動に従って 15 人の運転者を分類するのに十分であることを示した．さらに提案手法は，実際の実験データの検証を通して詐称者を検出できることを示した．

1. はじめに

車両の技術が進歩すると共に，盗難の技術も進歩する．FBI の調査によると，2017 年にアメリカでは約 80 万台の車両が盗難されている．そのため，自動車の盗難防止技術はより重要になる．特に新たなタイプの盗難攻撃に対しては，継続的な監視や追跡などの高度な盗難防止策が必要となる．

いくつかの研究で，盗難防止のための運転者認証や認識システムが提案されている．車両の施錠に生体認証を用いるシステムや指紋認証を用いた運転者の認証システムが挙げられる [1], [2]．これらのシステムに用いる生体情報や指紋情報は偽装することが難しため，盗難防止に適したシステムであると言える．しかし，運転者の認証プロセスははじめの 1 度のみ行われ，認証後に発生する盗難に対応することは難しい．そのためより安全な盗難防止システムは，運転中に運転者を継続的に認識するシステムを必要とする．

継続的な運転者の認識は，運転行動信号を監視し続けることで行う．運転行動信号はアクセルペダルやブレーキペダル，ハンドルの操作を用いる．従来の研究は，これらの制御要素を用いて個々の運転者を識別することが可能であることを示している [3], [4], [5]．さらに，運転者の行動信号の偽装は困難なため，運転行動信号の監視による盗難防止技術が役に立つことが知られている．従って，多くの研究は運転行動信号による運転者の識別において，高い精度で識別が可能であることが示されている．従来の手法は，運転者の推定確率を出力する分類器を使用し，推定確率の閾値に基づいて運転者の識別を行う [6] ．また，システムに登録されていない詐称者の検知システムについても研究が行われている [7] ．このシステムは，ニューラルネットワークの出力ラベルに詐称者を追加することにより詐称者の検知を実現している．しかし，登録済みの運転者の数が増加すると精度が低下する問題が指摘されている．

従来のシステムでは，運転行動信号を取得するために追加のデバイスを必要としている．一般的な車両は Controller Area Network (CAN) バス信号を取得するために自己診断機能インターフェイス (OBD2 : On-Board Diagnostics 2) を備えている．従来のシステムは OBD2 端子に追加のデバイスを接続し，データの取得を行っている．しかし近年の研究では，運転行動信号を CAN バス信号に変わり，スマートフォンやウェアラブル端末により取得する方法が提案されている [8], [9], [10]．スマートフォンやウェアラブル端末には加速度センサが搭載されているため，デバイスから取得される加速度データを運転行動信号として用いる [11]．しかし加速度データを用いた提案システムは，1 日の各時間毎における行動で分類器を作成するため，リア

†1 現在，愛知工業大学 経営情報科学研究科, 〒 464-0807 愛知県名古屋市千種区東山通 1-38-1
†2 現在，愛知工業大学 情報科学部 情報科学科, 〒 470-0392 愛知県豊田市八草町八千草 1247
a) yuki7291@pluslab.org
b) naito@pluslab.org

ルタイムでの識別は困難である．盗難防止技術には，リアルタイムで運転者を識別可能なシステムが必要になる．

現在，いくつかの研究では運転者の識別のためにニューラルネットワークによる分類器を使用している．ニューラルネットワークは与えられたデータから特徴を抽出し分類を行う．しかし一般的なニューラルネットワークは加速度データのような時系列データには適さない．そのため時系列データを扱う分類では，一般的に再帰型ニューラルネットワーク (RNN : Recurrent Neural Network) を用いる．例として，運転者の感情分類タスクにおいて，通常のニューラルネットワークに比べ RNN は効果的であることが示されている [12]．

本稿では，連続的に運転者を識別する手法と詐称者の検知手法を提案する．運転行動信号として加速度データを使用し，RNN による分類器を作成し運転者を識別する．また識別結果を元にした詐称者の検知も行う．実験結果は，実際の車両に搭載された加速度センサを使用し，15 人の運転者を高い精度で識別可能であることを示した．さらに提案手法は，識別結果の統計情報を使用し，詐称者の検知が可能であることも示した．

2. 提案手法

本稿では，運転行動信号に基づく運転者の識別手法及び詐称者の検知手法を提案する．運転行動信号は加速度データを用いる．加速度データはアクセルペダルやブレーキペダル，ハンドリングなどの運転行動の特徴を含んでいる．加速度データは時系列データであるため，提案手法では長短期記憶 (LSTM : Long Short Term Memory) ユニットを含む再帰型ニューラルネットワークによる分類器を用いて運転者の識別を行う．また，詐称者の検知は，登録済み運転者と詐称者を分類する二値分類器を用いて行う．従って運転者の識別と詐称者の検知は 2 種類の分類器によって行われる．

提案手法の概要を図 1 に示す．本手法の運転者の識別では特徴量選択によって抽出した特徴に対し，ウィンドウ分割による分析を行う．スライディングウィンドウ方式によって分割された各ウィンドウを分類器へ入力する．分類器は RNN 分類器を用いて行う．運転行動信号は時系列データであるため，データの前後の関係性が重要となる．RNN 分類器は前後関係を学習することが可能であるため，時系列データを入力として受け取ることが可能である．RNN 分類器の出力は学習した各運転者である確率ベクトルを出力する．最終的な運転者の識別結果は，分類器の出力に適切な処理を行った確率ベクトルである．

詐称者の検知は，運転者の識別結果である確率ベクトルを特徴量として用いる．確率ベクトルから各運転者に対する確率の平均や分散などの統計情報に変換し，二値分類器へ入力する．二値分類器は入力された統計情報を元に登録

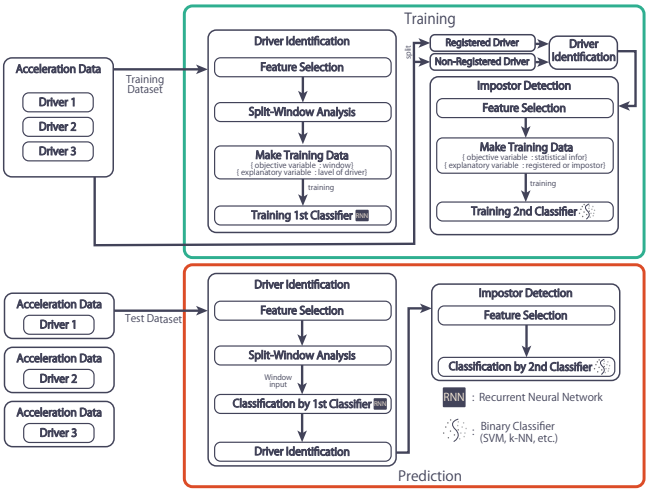

図 1 提案手法の概要

済み運転者か詐称者かを分類する．

本手法は複数の分類器により構成されるため，学習段階と推論段階から構成される．学習段階では予め用意されたデータセットを用いて分類器の学習を行う．推論段階では，入力されたデータと学習済みの分類器を用いて推論を行う．この推論は運転者の識別と詐称者の検知が含まれる．

2.1 学習段階

学習段階では，運転者識別に用いる RNN 分類器と詐称者の検知に用いる二値分類器の学習を行う．はじめに RNN 分類器の学習を行い，その後に学習済み RNN 分類器を用いて二値分類器の学習を行う．

RNN 分類器の学習では，はじめにデータセットから教師データの生成を行う．データセットは，一定距離の走行データに対し運転者のラベルが付けられている．走行データは 3 軸の加速度データで構成されている．そこで，3 軸データからアクセル・ブレーキペダルの操作を反映する前後加速度データと，ハンドル操作を反映する左右加速度データの 2 軸のデータを特徴量として抽出し，学習に用いる．抽出したデータが一定のサンプリングレートで記録されていない場合があるため，一定のサンプリングレートに統一する必要がある．そこで，走行データは 20Hz にダウンサンプリング処理を行う．リサンプリング処理されたデータは，複数のウィンドウに分割し，各ウィンドウ毎に運転者のラベル付けを行う．ウィンドウとウィンドウの間のデータの関係性は分割により失われてしまう．そこでウィンドウ分割はスライディングウィンドウによって行う．これによりウィンドウ間のデータは別のウィンドウデータに保持される．ラベル付けされたウィンドウデータを教師データとして RNN 分類器の学習を行う．

RNN 分類器は教師データを用いて学習する．しかし RNN 分類器の学習を行うためには，いくつかのハイパーパラメータを設定する必要がある．図 2 は RNN 分類器の

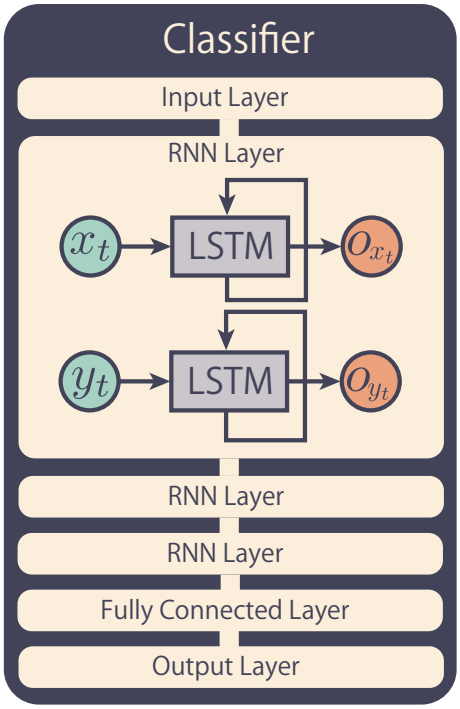

図 2　RNN 分類器の構造

構造を示している．RNN 分類器は入出力層や複数の RNN 層，全結合層から構成される．RNN 層は LSTM ユニットを内包している．そこでここでは，RNN 分類器の層の数や LSTM の隠れユニットの数を変化させた複数のモデルを設計し，正解率を比較検討し用いるモデルを決定する．

詐称者の検知は二値分類器によって行う．二値分類器は登録済み運転者と詐称者を分類する．分類器の学習のための教師データは運転者識別の推論結果を用いる．

RNN 分類器の学習に含まれない運転者のデータに対して運転者識別を行った場合，登録済み運転者に対するスコアベクトルを出力し，登録済み運転者から尤もらしい運転者を推定する．しかし出力されるスコアベクトルの様子は，登録済み運転者のデータに対して識別を行った場合のスコアベクトルと異なる．登録済み運転者の場合は，入力された運転者に対するスコアが大きく，その他の運転者に対するスコアは低くなる傾向がある．一方で，登録されていない運転者のデータである詐称者のデータを入力した場合は，各運転者に対するスコアの差は少なく，すべての運転者において一様なスコアを出力する．そこで，登録済み運転者を入力したときの出力と詐称者を入力したときの出力にラベル付けを行い，これを分類する分類器を設計する．

学習はデータセットを登録済み運転者と詐称者のサブセットに分割し，登録済み運転者のデータのみで RNN 分類器の学習を行う．学習を行った RNN 分類器を用いた運転者識別の推論を登録済み運転者と詐称者のデータで行う．その結果に対しラベル付けを行い教師データを作成する．推論結果は 2.2 章に示すように，各運転者に対するスコアベクトルである．特徴量は以下のスコアベクトルの統計情報を用いる．

- 最大値
- 平均値
- 中央値
- 分散
- 標準偏差
- 共分散

統計情報を含む特徴量を二値分類器に入力し学習を行う．分類器のアルゴリズムはランダムフォレストやサポートベクターマシンを想定し，実験により最も高い精度のアルゴリズムを選定し，これを用いる．

2.2　推論段階

推論段階では，学習段階で学習を行った RNN 分類器や二値分類器を用いて運転者の識別や詐称者の検知を行う．入力されたデータを元に，運転者の識別を行い，その後詐称者の検知を行う．運転者の識別では，登録済み運転者のうち最もらしい運転者を推論する．その後推論結果の運転者が登録済み運転者か詐称者かを二値分類器によって判断する．入力されたデータが詐称者のデータであっても，運転者識別では登録済み運転者のうちのいずれかの運転者として識別されてしまう．二値分類器は，識別結果が正しいか，詐称者のデータであったかを判断する．登録済み運転者の場合はそのまま認識され，詐称者の場合は識別結果は無効とし，詐称者として検知する．

運転者識別における推論では，入力されたデータに対して学習段階と同様の処理を行い分類器へ入力する．入力されるデータは一定のサンプリングレートではない場合があるため，20Hz にダウンサンプリングされる．ダウンサンプリングされたデータは，スライディングウィンドウにより複数のウィンドウに分割される．分割されたウィンドウを学習済みの RNN 分類器へ入力される．ここで，登録済み運転者の数を $|D|$ とすると，第 i 番目のウィンドウを入力したときの分類器の出力は次のような確率ベクトルで表される．

$$\boldsymbol{o}_i = [s_1 \ s_2 \ \cdots \ s_{|D|}] \tag{1}$$

最終的な運転者識別の推論を行うため，以下の式を用いてスコアベクトル $\boldsymbol{S}$ を算出する．

$$\boldsymbol{S} = \sum_i \boldsymbol{o}_i \tag{2}$$

算出されたスコアベクトルで，最も高いスコアの運転者を識別すべき運転者とする．

詐称者の検知は，運転者識別の推論で算出されたスコアベクトル $\boldsymbol{S}$ を用いて行う．スコアベクトルは各運転者に対するスコアを内包している．このスコアベクトルを平均値や中央値などの統計情報に変換し，二値分類器へ入力する．二値分類器は入力されたデータに対して，登録済み運

表 1　データセットのパラメータ

Number of Driver	15 people
Number of Round per Driver	15 rounds
Test data	1, 2, 4, 5, 7, 8, 10, 11, 13, 14th rounds
Training Data	3, 6, 9, 12, 15th rounds
Distance of Course	400m
Sampling Rate	200Hz
Resampling Rate	20Hz
Window size	128
Sliding size	16

転者か詐称者かを判断する．登録済み運転者と判断された場合は，入力されたデータは識別された運転者であると判断される．詐称者と判断された場合は，入力されたデータは識別された運転者ではなく，登録済み運転者に属さない運転者である詐称者と判断される．

3. 実験及び検証

3.1 実験

手法の検証に用いるデータセットは，実際の車両による実験によって収集された．車両は，加速度センサと Raspberry Pi を搭載した日産マーチを使用した．加速度センサは LIS3DH を用いた．Raspberry Pi と加速度センサは I2C 通信によりデータを送受信する．Raspberry Pi は加速度センサから送られてくるデータを約 200Hz で記録する．記録の際のサンプリングレートはソフトウェアの仕様のため，一定ではない．記録されたデータは後に検証のためのコンピュータに移される．

実験は 15 人の被験者により行われた．被験者は予め定められたコースを走行する．コースは 1 週約 400m であり，スタート地点が設けられている．被験者はスタート地点から走行を開始し，再びスタート地点まで戻り停止する．走行開始から停止までを 1 週と定義し，1 人の被験者につき 15 週のデータを収集した．

収集したデータは各分類器の学習のための教師データと検証に用いるテストデータの 2 種類に分割される．被験者の走行特性は，車両やコースに対する慣れにより変化することが予想される．そこで教師データは各走行のうちの 3，6，9，12，15 週とし，残りの走行データをテストデータとした．表 1 にデータセットの各パラメータを示す．

3.2 運転者識別

運転者の識別を行う分類器は再帰型ニューラルネットワーク（RNN）を用いる．運転者識別の検証では，入力された 1 週の走行データに対して，正しい運転者を識別できたかを検証する．検証では，RNN 分類器の RNN 層の数や各層に含まれる LSTM ユニットの数を変化させた複数のモデルを設計し，モデルごとの分類精度を比較する．設計したモデルの各パラメータを表 2 に示す．

表 2　RNN 分類器モデルのハイパーパラメータ

Hyper Parameters	Model 1	**Model 2**	Model 3	Model 4
Window size	128	**128**	128	128
Sliding size	16	**16**	16	16
1st LSTM Units	100	**200**	200	100
2nd LSTM Units	100	**200**	100	50
3rd LSTM Units	-	**-**	50	100
Classifier Accuracy	54.2	**59.0**	54.2	55.9
Identification Accuracy	98.0	**98.7**	98.0	97.3
Rank	2	**1**	3	4

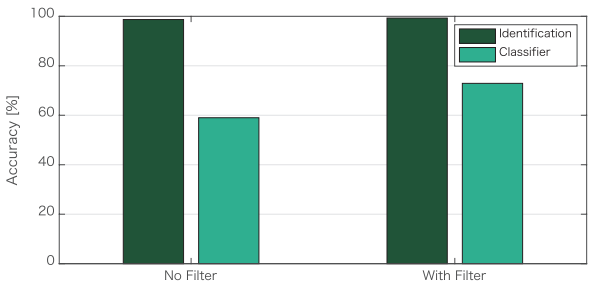

図 3　フィルタを適応したモデルの正解率

表 3　追加の RNN 分類器モデルのハイパーパラメータ

Hyper Parameters	Model 2 (No Filter)	Model 2	**Model 5**
Window size	128	128	**128**
Sliding size	16	16	**16**
1st LSTM Units	200	200	**200**
2nd LSTM Units	200	200	**200**
3rd LSTM Units	-	-	**200**
Classifier Acc	59.0	72.9	**74.2**
Accuracy	98.7	99.3	**99.3**
Rank	3	2	**1**

運転者識別の正解率は全てのモデルにおいて 90%以上を示した．しかし RNN 分類器の正解率は 60%以下に留まった．これは分類器に入力されたデータにエイリアシングが含まれていることに起因すると考えられる．初めに入力されたデータは 200Hz から 20Hz にダウンサンプリングされる．ダウンサンプリングの際にエイリアシングが発生していると考えられる．そこでダウンサンプリングの際にアンチエイリアシングフィルタを適応したデータを用いて識別を行う．図 3 は最も高い正解率を示したモデル 2 を用いてフィルタの有効性を検証した結果である．RNN 分類器の正解率はフィルタを適応した場合が適応しなかった場合に比べ，約 10%高いことが示された．

各層の LSTM ユニット数を変化させたモデルは，変化させないモデルに比べ正解率が低下する傾向が確認された．そこでモデル 2 と同数の LSTM ユニットを用いて RNN 層だけ変化させたモデルを設計し，モデル 2 と比較した．入力するデータにはフィルタを適応したデータを用いた．比較結果を表 3 に示す．RNN 分類器の正解率において，少量の改善が確認された．よって以降の詐称者検知の検証ではモデル 5 を用いる．

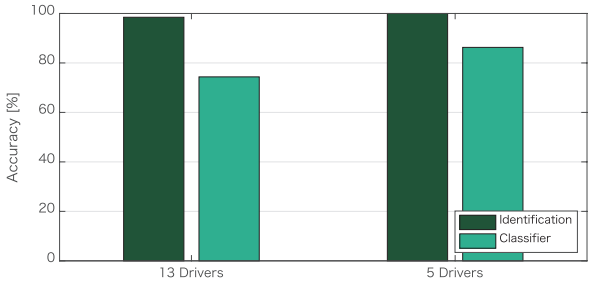

図 4　詐称者検知の正解率

表 4　二値分類器のアルゴリズム比較

Algorithm	Accuracy (%)
Logistic Regression	90.7
Cubic SVM	92.5
Weighted k-NN	93.8
Bagged Trees	92.3

表 5　5 人サブデータセットによる詐称者検知の混同行列

	Impostor	Registered Driver
Impostor	766	34
Registered Driver	41	359

3.3　詐称者の検知

詐称者検知の検証では，初めに登録済み運転者のみのデータを用いて RNN 分類器の学習を行う必要がある．そこで，15 人のデータから登録済み運転者が 13 人のサブデータセットと，5 人のサブデータセットを作成し比較する．13 人のサブデータセットには 2 人の詐称者データが含まれ，5 人のサブデータセットには 10 人の詐称者データが含まれる．各サブデータセットによる運転者識別の正解率を図 4 に示す．運転者の識別にはモデル 5 を用いて行った．5 人のサブデータセットによる識別では，分類器の正解率が 80%以上であることが示された．

そこで 5 人のサブデータセットを用いて二値分類器のアルゴリズムの比較検討を行う．分類器に用いるアルゴリズムの比較には，MATLAB の分類器 APP を用いて行った．正解率の検証には 5 交差検証を用いた．その結果を表 4 に示す．全てのアルゴリズムにおいて 90%以上の高い正解率を確認した．また最も高い正解率を示したアルゴリズムは k 近傍法であった．k 近傍法による分類結果を表した混同行列を表 5 に示す．

表 5 が示すように，登録済み運転者と詐称者のデータ数には偏りがある．偏りがあるデータでの評価には，一般的に再現率や適合率，F 値が用いられる．各評価指標は真陽性を TP，偽陽性を FP，偽陰性を FN とすると，次のように表される．

$$precision = \frac{TP}{FP + TP} \tag{3}$$

$$recall = \frac{TP}{FN + TP} \tag{4}$$

表 6　5 人サブデータセットによる詐称者検知の評価

	precision	recall	F-measure
Impostor	94.9%	95.8%	95.3%
Registered Driver	91.3%	89.8%	90.5%

表 7　13 人サブデータセットによる詐称者検知の評価

	precision	recall	F-measure
Impostor	89.2%	86.7%	87.9%
Registered Driver	96.9%	97.6%	97.3%

表 8　13 人サブデータセットによる詐称者検知の評価

	precision	recall	F-measure
Impostor	89.2%	86.7%	87.9%
Registered Driver	96.9%	97.6%	97.3%

$$F\text{-}measure = 2 \times \frac{presicion \times recall}{precision + recall} \tag{5}$$

登録済み運転者と詐称者に対する各評価は表 6 の通りである．登録済み運転者と詐称者共に F 値は 90%以上である．これは k 近傍法が適したアルゴリズムであることを示している．

そこで 13 人のサブデータセットを用いて同様の検証を行った．用いたアルゴリズムは k 近傍法である．その結果を表 8 に示す．詐称者の F 値は 5 人のサブデータセットに比べ低下し，登録済み運転者は増加した．これは詐称者のデータ数が大幅に減少したためであると考えられる．また RNN 分類器の正解率が低下していることも挙げられる．

4.　結論

本稿では，運転者の識別手法と詐称者の検知手法の提案を行った．運転者識別では，時系列データを扱う再帰型ニューラルネットワークが有効的であることを示した．15 人の運転者を 98%以上の正解率で識別できることを示した．詐称者の検知では，k 近傍法を用いて登録済み運転者と詐称者を 90%以上の正解率で分類することが可能であることを示した．今回の検証では，定められたコースを走行したデータを用いて行われた．今後の課題として，走行コースの一般化が挙げられる．運転者の識別と詐称者検知は任意のコースや場所で行える必要がある．そこで今後は，様々なコースのデータを収集し，未知のコースにおいての検証を行っていく．

参考文献

[1] N. N. Nagamma, M. V. Lakshmaiah, and T. Narmada, “Raspberry pi based biometric authentication vehicle door locking system,” in *2017 IEEE International Conference on Power, Control, Signals and Instrumentation Engineering (ICPCSI)*, pp. 2348–2351, September 2017.

[2] J.-D. Wu and S.-H. Ye, “Driver identification using finger-vein patterns with radon transform and neural network,” *Expert Systems with Applications*, vol. 36, no.

3, Part 2, pp. 5793–5799, April 2009.

[3] H. Qian, Y. Ou, X. Wu, X. Meng, and Y. Xu, "Support vector machine for behavior-based driver identification system," *Journal of Robotics*, Vol. 2010, Article ID 397865, 11 pages, March 2010.

[4] M. V. Martìnez, I. D. Campo, J. Echanobe, and K. Basterretxea, "Driving behavior signals and machine learning: A personalized driver assistance system," in *2015 IEEE 18th International Conference on Intelligent Transportation Systems*, pp. 2933–2940, September 2015.

[5] S. Jafarnejad, G. Castignani, and T. Engel, "Towards a real-time driver identification mechanism based on driving sensing data," in *2017 IEEE 20th International Conference on Intelligent Transportation Systems (ITSC)*, pp. 1–7, October 2017.

[6] B. I. Kwak, J. Woo, and H. K. Kim, "Know your master: Driver profiling-based anti-theft method," in *2016 14th Annual Conference on Privacy, Security and Trust (PST)*, pp. 211–218, December 2016.

[7] M. V. Martìnez, J. Echanobe, and I. del Campo, "Driver identification and impostor detection based on driving behavior signals," in *2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC)*, pp. 372–378, November 2016.

[8] C. Yang, D. Liang, and C. Chang, "A novel driver identification method using wearables," in *2016 13th IEEE Annual Consumer Communications Networking Conference (CCNC)*, pp. 1–5, January 2016.

[9] P. Phumphuang, P. Wuttidittachotti, and C. Saiprasert, "Driver identification using variance of the acceleration data," in *2015 International Computer Science and Engineering Conference (ICSEC)*, pp. 1–6, November 2015.

[10] Z. Chen, J. Yu, Y. Zhu, Y. Chen, and M. Li, "D3: Abnormal driving behaviors detection and identification using smartphone sensors," in *2015 12th Annual IEEE International Conference on Sensing, Communication, and Networking (SECON)*, pp. 524–532, June 2015.

[11] N. Virojboonkiate, P. Vateekul, and K. Rojviboonchai, "Driver identification using histogram and neural network from acceleration data," in *2017 IEEE 17th International Conference on Communication Technology (ICCT)*, pp. 1560–1564, October 2017.

[12] U. E. Manawadu, T. Kawano, S. Murata, M. Kamezaki, J. Muramatsu, and S. Sugano, "Multiclass classification of driver perceived workload using long short-term memory based recurrent neural network," in *2018 IEEE Intelligent Vehicles Symposium (IV)*, pp. 1–6, June 2018.