The Car is Safe: A Fast and Accurate Pressure-based Authentication System for Identifying Car Drivers

Mohsen Ali Alawami, Dahyun Jung, Yewon Park, Yoonseo Ku, Gyeonghwan Choi, and Ki-Woong Park

Department of Information Security, Sejong University, Seoul, South Korea
{mohsencomm, woongbak}@sejong.ac.kr
{dahyunj, sebong0412}@sju.ac.kr
{sksmsdbsej, jacksimon86}@gmail.com

Abstract

Although authorized drivers gain access to their vehicles via tokens such as ignition keys, fingerprints, or RFID tags, the security of vehicles still has essential vulnerabilities that unauthorized drivers (*e.g.,* thieves or attackers) can access and misuse for criminal purposes. Therefore, driver authentication is an important rule in which the identity of a person who driving the car can be identified. Many previous studies have proposed solutions using behavioral characteristics based on sensory data collected from car devices on the controller area network (CAN) bus, GPS data, or face recognition. However, by exploring these technologies, we found that they have shortcomings such as cybersecurity attacks on the CAN buses, violating the privacy of users via GPS and face techniques, and consuming a long time (a couple of minutes) to authenticate drivers. Our research tackles the problems mentioned above by investigating the feasibility of authenticating a driver quickly in a few seconds (using a few data samples) and accurately using only a single sensor without hurting the user's privacy. In this paper, we develop a fast and accurate authentication scheme that depends on driver's pressure data collected from sensors attached to the driver's seat and belt. To do this, we conducted real-world experiments and collected large-scale pressure datasets from 12 users under different settings. To evaluate the effectiveness of our system, we implemented three different evaluation scenarios for seat-only, belt-only, and fusion seat and belt datasets through various train/test split ratios using two machine learning algorithms. Our experimental results of pressure datasets show that the system can authenticate the driver with an F1 score of 99.7% for belt-only (in 0.5 seconds), an F1 score of 94% for seat-only (in 1.3 seconds), and an F1 score of 96% for fusion module (in 2.3 seconds) for the best case.

Keywords: Car, Driver, Pressure, Authentication, Machine Learning.

1 Introduction

Many advanced technologies are applied to various automobiles (*e.g.,* machine learning, deep learning, supplemental restraint systems, automatic collision avoidance system (ACAS), and millimeter wave radar) to increase the passengers' facilities and improve their safety. Specifically, these technologies provide various benefits to automatically detect vehicles, avoid them from colliding with obstacles, and determine nearby pedestrians [1]. However, they can not eliminate vehicle security issues such as driver authentication in which an unauthorized person (*e.g.,* thief) drives the car and misuses it for criminal purposes. Although legitimate drivers gain

authorization to access the vehicles via tokens (*e.g.,* ignition keys, fingerprints, and RFID tags), these approaches can not secure all instances of vehicle misuse because they can be stolen or forged by attackers (or thieves). Therefore, driver authentication for vehicles is one of the key issues and important research topics nowadays. Many applications and cases in special-purpose vehicles, public transportation (bus/taxi/metro), and private cars are vulnerable to driver authentication in which drivers need to be authorized. For example, administrators of public transportation need to only authorize certain persons to drive their vehicles (*i.e.,* bus/taxi/metro), insurance companies need to identify drivers to better understand their habits and behaviors for accurate insurance pricing determination, military vehicles need to accept special soldiers to drive to ensure the safety of weapons and equipment, and parents sometimes want to prevent their children or strangers from driving cars. A most important case of driver authentication is to protect the car from unauthorized access such as theft who steal it and misuse it for criminal purposes.

Many driver authentication solutions are proposed to identify driver identity using various methods such as fingerprinting driving habits and monitoring behavioral characteristics based on sensory data collected from controller area network (CAN) bus [2]. However, there are still many limitations and security issues that need to be addressed to secure vehicles. First, some proposals require a large amount of data on the CAN bus from various controllers (*e.g.,* on-board diagnostics-II OBD-II, steering wheel, brake pedal, accelerator, engine speed, and vehicle speed) and through external devices that use the same networks such as smartphones, which makes the approaches vulnerable to cyber-attacks and needs a long time (about 8 to 15 minutes) for authentication [3, 4, 5, 6, 7, 8]. Second, other solutions require collecting sensitive information such as GPS information (locations and time) and accelerometer readings by utilizing that the driver almost drives at the same speed through the same specific routes and destinations everyday [9, 10]. Besides that these approaches contain private location and time synchronization data, which are problematic from a privacy perspective if shared, they are also vulnerable to location and time spoofing attacks. Third, many techniques use the camera to record the driver's face and build an in-car face database for applying face-based and eye-based driver recognition systems with the aid of deep learning [11, 12, 13]. However, camera-based driver authentication solutions still suffer from many challenges reducing the authentication performance such as covering the face, extreme illumination change, different head pose, camera location, multiple face appearances, and wearing glasses. Fourth, many other authentication studies use the concept of wireless sensing such as RF radio signals (*e.g.,* Wi-Fi or Bluetooth) or channel state information (CSI) as well as require external devices with various sensor data from devices carried by the driver such as smartphones [14, 15, 16, 17, 18]. However, extracting biometric information recorded in wireless signals suffers from multipath propagation issues. Also, involving external devices increases the system cost and adds security threats that attackers can utilize. Finally, most driver authentication studies have performed experiments using data generated from computer simulators rather than collecting actual data from drivers on cars [4, 15, 17, 19, 20, 20, 21] – which limits its performance in real-world implementation and usage.

Therefore, based on the aforementioned shortcomings, we target developing a novel driver authentication system that has the following design goals. First, instead of relying on multi-sensor data from the CAN bus, we propose a single-sensor-based authentication approach using only pressure measurements. Second, our system does not require any external devices such as CSI information or smartphones – which makes the system cost-less and applicable for real usage. Third, we aim to make the system fast and accurate in that a driver can be authenticated within a few seconds (on average from 5 to 10 seconds). Fourth, our implementations are

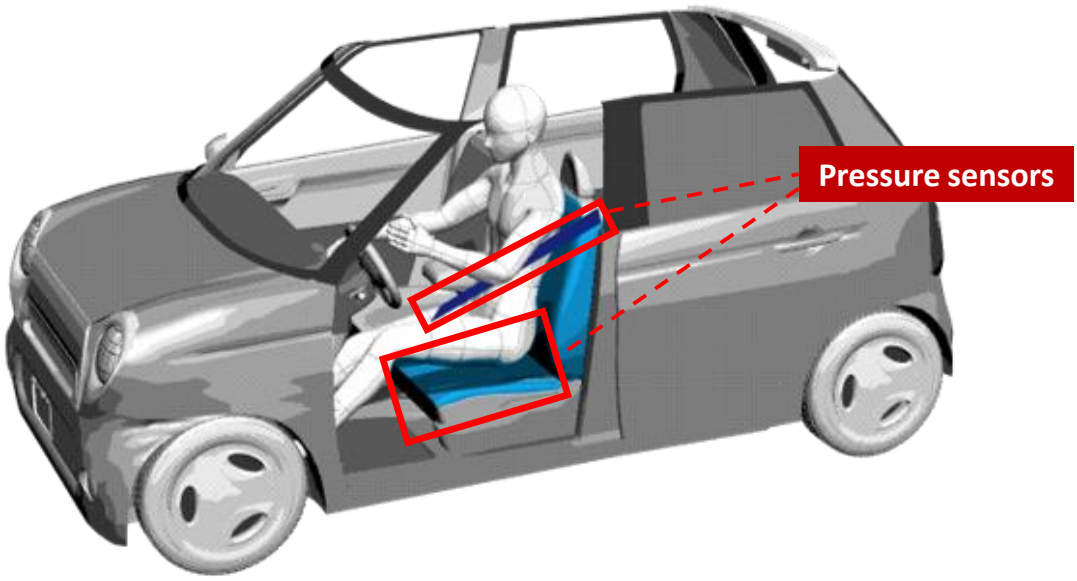


Figure 1: The general illustration of pressure-based driver authentication system.

realistic so that the datasets are collected from real experiments rather than from computer simulators. Finally, our system supports data privacy-preserving since it does not rely on recording drivers' faces or their location information.

The objective of this work is to evaluate the feasibility of authenticating the driver's identity using a few samples of data collected from pressure sensors installed on the belt and seat of the car driver. To do this, we conducted real experiments and collected actually pressure datasets from 12 users under several iterations and settings. We also implemented machine-learning models to achieve high authentication accuracy as well as fast authentication time in which a driver can be correctly classified using only a few samples of data that are collected within a few seconds.

Here, we summarize our work's contributions as follows.

1. We propose a novel, fast, and accurate driver authentication system using user-specific data acquired from only pressure sensors equipped in the driver's belt and seat.

2. We conducted real experiments, rather than using computer simulators, by installing 60 pressure sensors on the driver's seat (30 on the belt and 30 on the seat) and collected large-scale and real-world datasets from 12 users, each of them repeated for 10 iterations under different settings.

3. We constructed two machine-learning algorithms and evaluated the system using three evaluation settings: four train/test split ratios of the users' data and three scenarios (Belt-only, Seat-only, Belt+Seat); as well as computed the train/test time consumption for authentication.

4. Our experimental results show the effectiveness of the proposed system by achieving the best Belt-only performance of an F1 score equal to 99.7% and an accuracy of 99.61% in 0.54 seconds. For the Seat-only setting, the system achieved the best performance of an

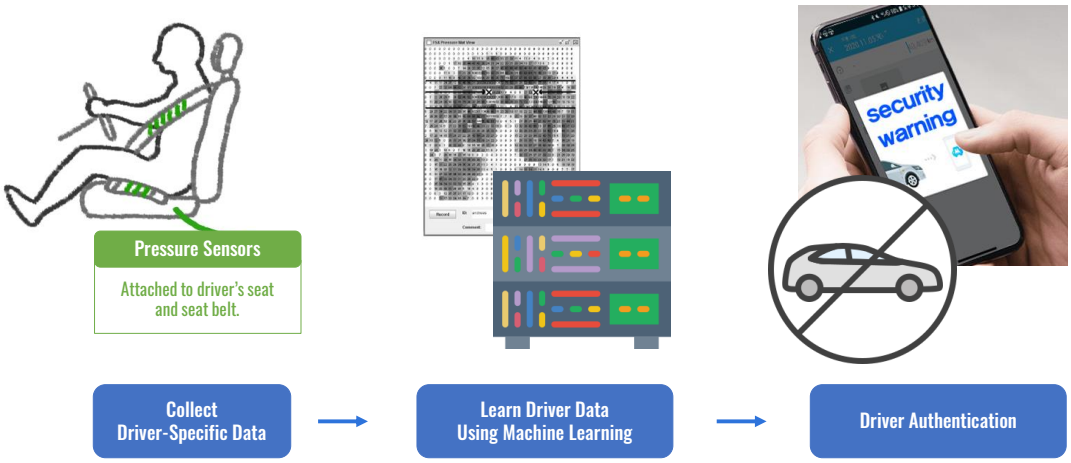


Figure 2: The pipeline of pressure data collection, preprocessing, model training, and authentication.

F1 score equal to 94% and an accuracy of 94.04% in 1.34 seconds. When fusion both belt and seat data samples, the system achieved the best performance of an F1 score equal to 96% and an accuracy of 96.08% in 2.32 seconds.

2 System Design and Methodology

Here, we explain the system overview design and experiments that we conducted for the dataset collection process.

2.1 System Overview

In this research, we were inspired by the shortcomings of the existing studies of driver authentication mentioned in the previous section to initiate the work's idea. We put the following research question: *Given data collected from only a single sensor from the car, can we quickly and accurately authenticate the driver within a few seconds using a few samples of data without requiring external devices?'*. The answer to this question can solve several limitations of previous studies that involve huge data types from multiple devices on CAN bus, requiring additional devices, needing a long time (reach to 15 minutes) of authentication, or disclosing private information. After deep thinking, we found that attaching pressure sensors to the driver's belt and seat can be a valid trick to identify who is setting and driving the car. In other words, pressure sensors can be implicitly mounted by car manufacturers one time during the manufacturing in which only the driver's seat will be equipped with sensors to collect pressure data whenever a driver seats and drives. Therefore, the objective of this research is to investigate whether we can develop a lightweight (only using one sensor), fast (within a few seconds), and accurate driver authentication system. Note that, using pressure data, any user (legitimate or attacker) must first sit and drive for at least a few seconds to steal the car. This is entirely consistent with our key idea that pressure sensors immediately generate pressure data when a user starts to sit down and drive.

(a) (b) (c) (d)

Figure 3: Illustrations show the testbed of data collection experiments: (a) Driver seat attached with belt and seat pressure sensors. Figures [(b), (c), and (d)] show participants under different postures, clothes, and gender settings.

To do this, we decided to investigate the feasibility of classifying the pressure data of multiple users and check the performance of how much we can distinguish one user's pressure data from others for the purpose of driver authentication. Figure 1 shows the general illustration of our experimental thoughts in which the pressure sensors can be installed in the belt and seat of the car driver. Our system has three main stages shown in Figure 2. The first stage is about collecting pressure data from the sensors attached to the driver belt and seat. After that, the data will be immediately sent to a database in a trusted server (the server is under our control). During the offline learning, we execute data preprocessing, learn driver data using machine-learning models, and save the models' files (.pickle) inside the server. This stage is only conducted for the legitimate users who are authorized to drive the car to build their pressure-based templates or profiles in the database. Finally, it is the stage for online prediction that is whenever a user (legitimate or attacker) sits and drives the car, the sensors collect his/her pressure data for a few seconds and then send it to the server for authenticating him based on the models' predictions and outputs. Then, a warning message can be sent to the owner's smartphone in case the models predict that the tested sensory data does not belong to the legitimate user's template. Note that, in this work, we only focus on investigating the feasibility of whether pressure-based authentication for drivers is effective enough and promising. We leave the entire development of the system including sending warning messages to the owner for future work.

2.2 Experiments of Dataset collection process

Here, we explain in detail about the dataset collection process and its demography. We asked 12 volunteers to participate in the experiments of data collection. They are all undergraduate students (males and females), who have a background in computer science and applications, and they are informed about the experiment's purpose. We collected pressure datasets from the belt (top and bottom) and the seat. Figure 3a shows the way we installed the pressure sensors in three positions of the driver seat: (1) the top of the belt (15 sensors), (2) the bottom of the belt (15 sensors), and (3) the base of the seat (30 sensors). Specifically, we intentionally asked participants to wear different heavy and light clothes as well as different setting postures as shown in Figure 3b, Figure 3c, and Figure 3d. These settings are important for studying many impacts that may directly affect the pressure measurements and then change the authentication

performance. The participant's data consists of the pressure measurements from the belt that were collected every 5 seconds immediately after the belt was worn. Also, the participant's data consists of the pressure measurements from the seat that were collected for 10 seconds immediately after seating and wearing the belt.

The sensor values were measured at 0.1-second intervals. This means that for each experiment, for the belt-only pressure data, we collected 50 samples (*i.e.,* 50 rows) where each sample has 30 sensory values (*i.e.,* 30 columns). Similarly, for the seat-only pressure data, we collected 100 samples (*i.e.,* 100 rows) where each sample has 30 sensory values (*i.e.,* 30 columns). In short, for each experiment, we create two matrices of pressure dataset for each participant. The first matrix is called "belt-only" and has dimensions of 50 rows $\times$ 30 columns. The second matrix is called "seat-only" and has dimensions of 100 rows $\times$ 30 columns. Note that, we intentionally repeated each experiment for ten iterations for each participant to make sure that the pressure values were consistent under different settings and through all iterations. The data of each case is saved in a separate (.csv) file. In total, our pressure dataset contains 240 (.csv) files: 12 users $\times$ two data types (Belt, Seat) $\times$ ten iterations. In terms of the number of samples, we collected about 500 samples for belt-only data and 1000 samples for seat-only data from each user. Therefore, we collected about 6000 samples and 12000 samples of belt-only and seat-only datasets from all 12 users. Each sample has 30 values (i*i.e.,* 30 columns) representing the 30 attached sensors for either belt or seat. All dataset files are then saved into a database on our server for later learning processes.

2.3 Implementation and evaluation metrics

Here, we explain the preprocessing steps applied to the dataset, types of used machine-learning algorithms, evaluation scenarios, and evaluation metrics that we used for assessing the performance of our system.

Preprocessing: We first explored the whole dataset and found some missing values in some data samples (*i.e.,* values of some sensors are missed from the rows) for both belt and seat dataframes. For example, some samples that were collected during the 5 seconds of belt-only or 10 seconds of seat-only dataframes have less than 30 values (usually ranging from 20 to less than 30 data points). Since we found that the counted incomplete samples are few (usually less than 5 samples per user), we decided to remove these incomplete samples from the datafreames. To input the data into the machine-learning model, we vertically contacted dataframes from the ten iterations for each user so that the model is sufficiently trained with enough samples to better understand the user's behaviors and create an offline template that will be used later for online predictions. Then, we applied the "max-min" normalization method for the whole dataset to re-scale the pressure measurements to be within 0 and 1 values.

Driver-pressure learning process: We have tried several machine-learning algorithms and found that two of them can be efficiently used for our application since they have achieved good performance results. In short, we used the "RandomForest" and "LogisticRegression" machine learning algorithms to evaluate the accuracy of our system. Random forest classifier is a meta estimator that fits a number of decision tree classifiers on various sub-samples of the dataset and uses averaging to improve the predictive accuracy and control over-fitting. Random forest classifier has two main hyperparameters that can affect the classification performance as follows. The first is the "n_estimators", which controls the number of trees in the forest by changing the number of estimators. The second is "max_depth" of the tree, which refers to the depth that nodes are expanded until all leaves are pure. In contrast, logistic regression is used as a type of statistical model for classification and predictive analytics to estimate

the probability of an event occurring, such as driver authentication based on given pressure samples of the belt and seat dataset of each user. The output of the model is a probability, which is a dependent variable bounded between 0 and 1. Specifically, in our work, we used "RandomForest" and "LogisticRegression" machine learning algorithms supported by the scikit-learn 1.3.0 library with the multi-class classification of the 12 users in which the algorithms use the one-vs-rest (OvR) scheme. Note that, for "sklearn.ensemble.RandomForestClassifier", we set hyperparameters (n_estimators = 1000) and (max_depth = 100), while we used the default settings of hyperparameters for the "sklearn.linear_model.LogisticRegression".

For every user's pressure dataset, we split the data samples into various training and testing parts. Then, for every evaluation scenario, we used the training amount of samples to learn the "RandomForest" and "LogisticRegression" classifiers and create two (.pickle) files of the two models. We save them in the user's directory on the server for later testing and output predictions.

Evaluation scenarios: We considered three evaluation scenarios in this work to assess the performance of the proposed system as follows. First, we evaluate two modules separately, which are Seat-only and Belt-only. We report their detailed confusion matrices regarding the number of pressure samples to show how many samples are correctly predicted for each user corresponding to the actual tested samples. For each module, we created four different confusion matrices using four different training and testing split ratios of the number of samples; (1) Train: 60%, Test: 40%; (2) Train: 70%, Test: 30%; (3) Train: 80%, Test: 20%; (4) Train: 90%, Test: 10%. In this evaluation scenario, we only used the "RandomForest" machine learning classifier for both Seat-only and Belt-only modules. The purpose of this evaluation is to investigate how much pressure samples of a user can be misclassified to other users as well as what the best train/test split ratio is valid (*i.e.,* what is the minimum number of samples that we can use to achieve a reasonable classification accuracy of users' pressure data). Second, we consider the evaluation of several aspects such as considering another machine learning algorithm (*i.e.,* "LogisticRegression") for the purpose of performance and time trade-off by comparing performance (accuracy, Precision, recall, and F1 score) versus time consumption for authentication. In other words, utilizing another classifier can let us see what classifier is suitable in terms of both achieving good accuracy and consuming reasonable time for training and testing data. Similarly, as in scenario 1, we did this evaluation scenario for the seat-only and belt-only modules separately. Third, from the about of the above two evaluation scenarios, we used the best machine learning classifier to evaluate the performance of the system using the fusion of both seat and belt pressure datasets and computed the train and test time required for authentication.

Evaluation metrics: To evaluate the effectiveness of the proposed system, we use the following metrics. *True negative (TN)*: The pressure fingerprints from legitimate users authorized to access the car (*e.g.,* owners) are correctly identified by the system as legitimate data and hence the system grants car access. *True positive (TP)*: The pressure fingerprints from attackers who haven't access to the car are correctly identified as unauthorized data and hence the system rejects car's access and sends a warning message to the owner's smartphone. *False positive (FP)*: The pressure fingerprints from legitimate users authorized to access the car (*e.g.,* owners) are incorrectly rejected by the system. *False negative (FN)*: The pressure fingerprints from attackers are incorrectly identified as authorized data and hence granted by the system to get car access. *Precision* is defined as in Equ. 1, which measures the portion of true positives divided by the summation of true positives and false positives. *Recall* is defined as in Equ. 2, which measures the portion of true positives divided by the summation of true positives and false negatives. The formula for the F1 score is shown in Equ. 3, which is a weighted average

metric emphasizing the model's performance regarding false positives and false negatives. We also computed the accuracy metric to show how many true samples are correctly classified from all tested samples for every user. Note that, our dataset is balanced for all users so that the macro accuracy is weighted too. Finally, we computed the confusion matrix to show how much the system is effective in terms of the detailed number of samples for actual and predicted users.

$$Precision = \frac{TP}{TP + FP} \tag{1}$$

$$Recall = \frac{TP}{TP + FN} \tag{2}$$

$$F1 = \frac{2 \times precision \times recall}{precision + recall} \tag{3}$$

3 Performance and Evaluation Results

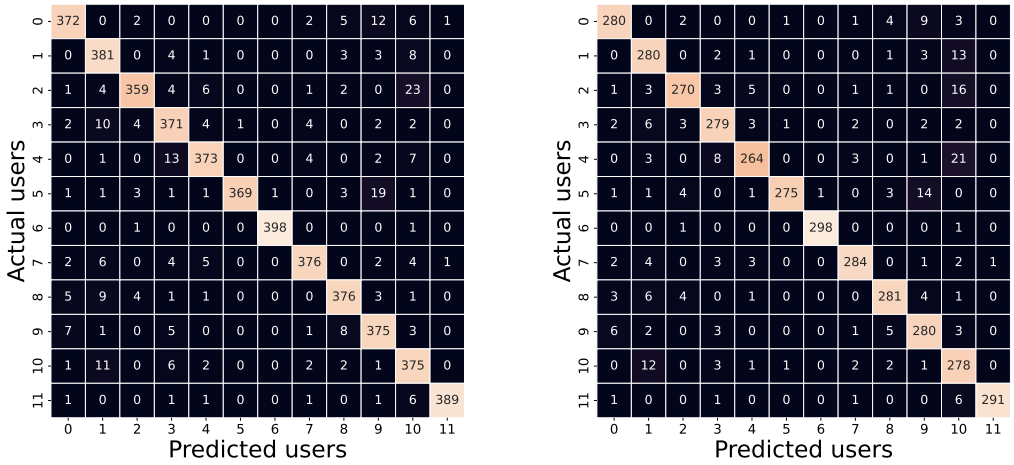
In this section, we demonstrate in detail the performance results under the aforementioned three evaluation scenarios to show the feasibility and effectiveness of our work.

3.1 Evaluation Scenario 1: Confusion matrix of users' samples

Our evaluation results in this scenario show the confusion matrix for seat-only and belt-only modules in terms of the number of samples that were correctly predicted by the RandomForest classifier under various train and test split ratios of pressure data. Specifically, as mentioned in section 2.2, we collected a total amount of 5 seconds and 10 seconds of pressure data for each user for seat-only and belt-only modules. The data collection was then repeated ten times for each user. When evaluating each user's data, we used multi-class classification with four different train and test split ratios of the data. Our goal is to investigate what optimal amount of samples that are enough for the classification of 12 users and how it changes from one user to another and from the belt-only module to the seat-only module. After conducting the experiments, we report the confusion matrices in Figure 4 and Figure 5.

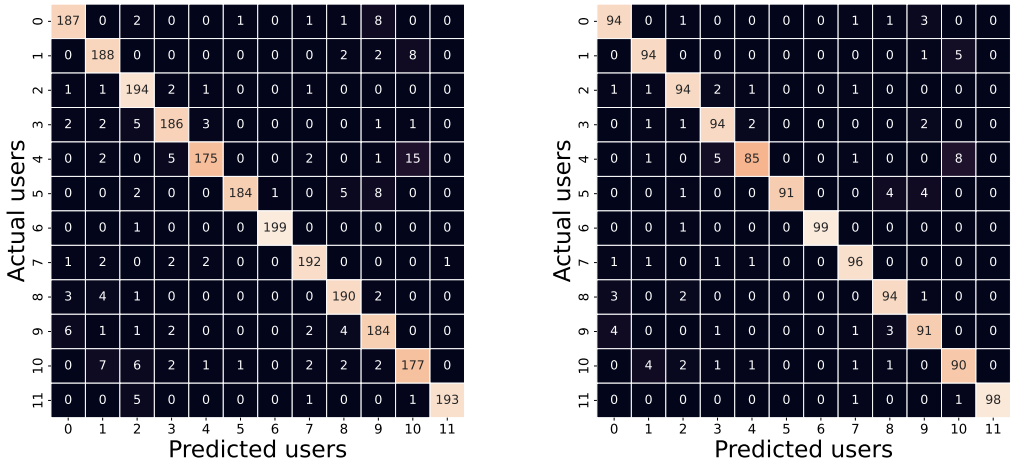
Note that, the number of samples used to train and test each user differs based on the train/test split ratio and the type of module (*i.e.,* Seat-only or Belt-only). In detail, the number of samples used for the seat-only module is 900/100, 800/200, 700/300, and 600/400 using train/test split ratios of 90/10, 80/20, 70/30, and 60/40 respectively. Similarly, the number of samples used for the belt-only module is 441/49, 392/98, 343/147, and 294/196 using train/test split ratios of 90/10, 80/20, 70/30, and 60/40 respectively. Note that, in the preprocessing steps, we found that belt pressure data have some incomplete samples. We removed the missing and incomplete samples from the dataset. All classification tasks in this evaluation scenario were done using the RandomForest classifier with hyperparameters mentioned in section 2.3.

Our results of the Seat-only module are presented in Figure 4 in which we plotted four fusion matrices based on the type of train/test split ratio. The index of the actual 12 users (indexed from 0 to 11) is set on the y-axis while the label of predicted users is set on the x-axis of the confusion matrix. The number of samples shown in the diagonal of each matrix refers to the amount of correctly classified samples for each user out of the total tested samples based on the train/test split ratio. For example, in Figure 4a, (train: 60%, test: 40%), the total



(a) Train: 60%, Test: 40%

(b) Train: 70%, Test: 30%



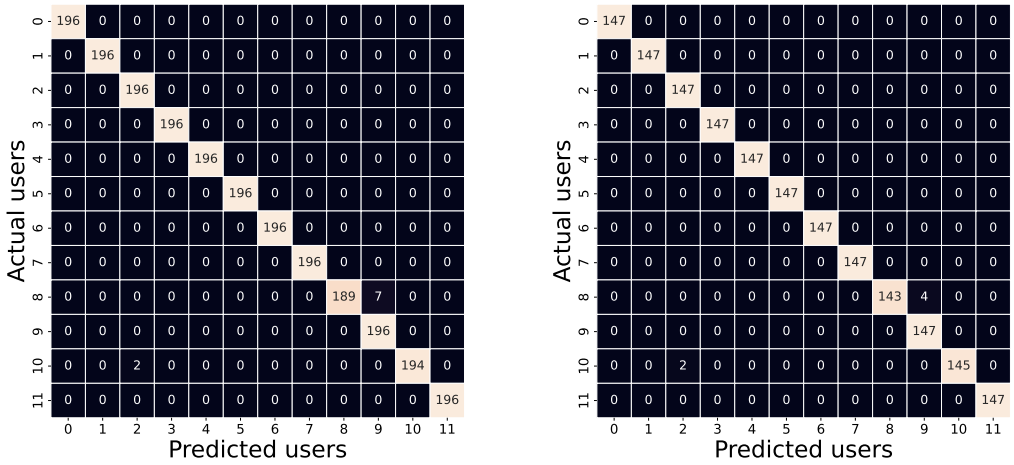
(c) Train: 80%, Test: 20%

(d) Train: 90%, Test: 10%

Figure 4: Confusion matrix shows the performance of 12 users of driver's authentication using pressure datasets collected from Seat-based sensors.

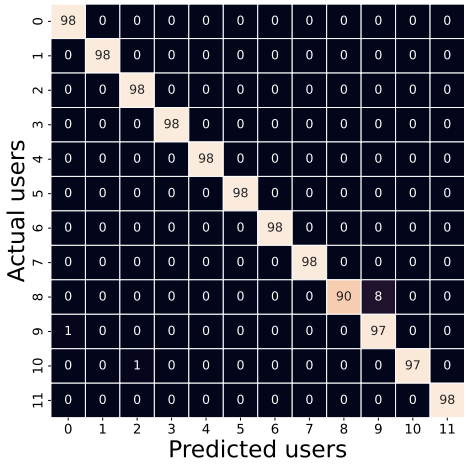
tested samples were 400 for each user. However, the confusion matrix reports that there are some samples of each user are classified wrongly as other user's samples. In detail, the best classification results in Figure 4a are reported for user 6 where 398 out of 400 samples are correctly classified as user 6. However, the worst classification results were reported at user 2 where only 359 samples out of 400 are correctly classified as user 2, while 23 samples are classified as user 10 and 6 samples as user 4, *etc.*

By looking at the confusion matrices in Figure 4b,(train: 70%, test: 30%), we can see that

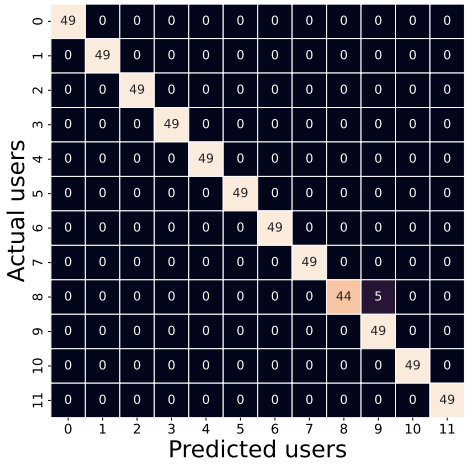


(a) Train: 60%, Test: 40%

(b) Train: 70%, Test: 30%



(c) Train: 80%, Test: 20%



(d) Train: 90%, Test: 10%

Figure 5: Confusion matrix shows the performance of 12 users of driver's authentication using pressure datasets collected from Belt-based sensors.

using a less number of tested samples (*i.e.,* 300), report distribution of results is almost the same where user 6 has the best amount of correctly detected samples and user 2 has the worst. However, in Figure 4c, (train: 80%, test: 20%), the classification results using 200 samples are a bit different where user 4 is now providing the worst performance than user 2 with about 25 misclassified samples. Lastly, Figure 4d confirms the same result distribution as Figure 4c which indicates that we can reduce the amount of test samples and get almost similar performance. Note that 20% (*i.e.,* 200 samples) and 10% (*i.e.,* 100 samples) of test samples for each user

are collected in 2 seconds and 1 second respectively. This definitely shows the ability of our system to provide fast and accurate authentication decisions by predicting user identity using only 20% or 10% test samples.

By looking at the confusion matrices of the belt-only module shown in Figure 5, we can see in general that the overall performance is better than the seat-only module. In detail, number of misclassified samples per user is less than those in the seat-only module. For example, in Figure 5a, (train: 60%, test: 40%), the total tested samples were 196 for each user. However, the confusion matrix reports that there are only a few samples of some users are classified wrongly as other user's samples. Specifically, the best classification results in Figure 5a are reported that most users have no misclassified samples. However, the worst classification results were reported at user 8 where 189 samples out of 196 are correctly classified as user 8, while 7 samples are classified as user 9. By looking at the confusion matrix in Figure 5b, (train: 70%, test: 30%), we can see there is no change in the distribution results across all users even when using fewer test samples (*i.e.,* 147 samples per user). In addition, when changing the train/test split ratios to 80/20 in Figure 5c, we observe a change in the result distribution in terms of the increasing number of users that get the worst performance and have misclassified samples. Lastly, when reducing test samples to 10% as shown in Figure 5d, we still get accurate performance for most users except user 8 where 5 samples are misclassified as user 9. From this evaluation scenario, we can conclude that belt-only performance is higher than seat-only even across various train/test data split amounts. However, in real-world usage, we can not guarantee that the attacker will wear the driver belt when conducting the attack. Therefore, including seat-based pressure data in our system is essential because the attacker must have to sit and drive during the car attack.

3.2 Evaluation Scenario 2: Authentication performance and time trade-off

In this evaluation scenario, we conducted more experiments to check the validity of our system when considering authentication accuracy versus time trade-off. In other words, we tried to answer the question that is: *How fast does our system authenticate a user while maintaining accurate performance?*. To do this, we implemented two machine learning algorithms which are "RandomForest" and "LogisticRegression" to compare their performance and time trade-off using various train/test split ratios. In addition, we measured the performance using four metrics which are accuracy, precision, recall, and F1-score. Also, we plotted train and test time consumption for each classifier for each belt-only and seat-only pressure dataset.

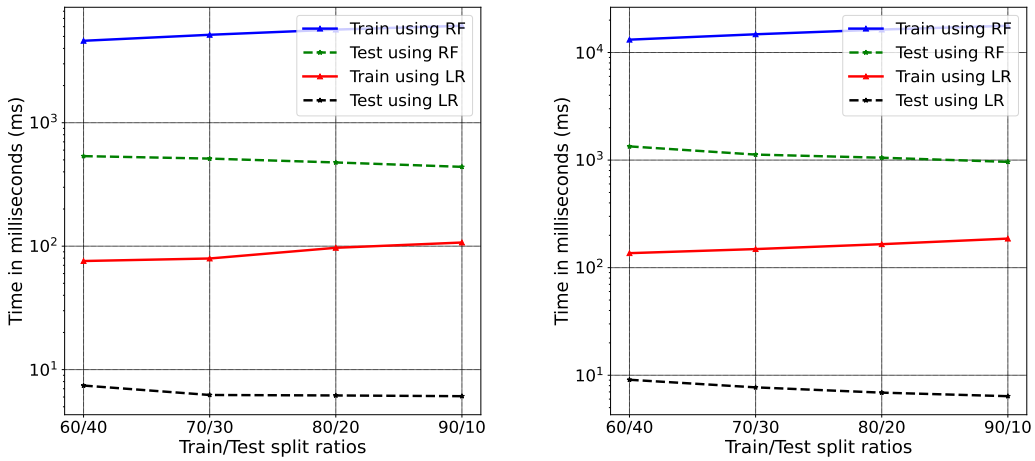
Table 1: Evaluation results of the 12 users' pressure datasets using driver's belt.

Data type	Tr/Ts ratio [%]	Samples /user	Total tested samples	Model	Accuracy [%]	Precision [%]	Recall [%]	F1-score [%]
Belt	90/10	49	588	LR	95.06	95	95	95
	80/20	98	1176		96.14	96	96	96
	70/30	147	1764		95.83	96	96	96
	60/40	196	2352		96.25	96	96	96
	90/10	49	588	RF	99.15	99	99	99
	80/20	98	1176		99.15	99	99	99
	70/30	147	1764		99.65	100	99	99.49
	60/40	196	2352		99.61	99.5	100	99.75

Table 2: Evaluation results of the 12 users' pressure datasets using driver's seat.

Data type	Tr/Ts ratio [%]	Samples /user	Total tested samples	Model	Accuracy [%]	Precision [%]	Recall [%]	F1-score [%]
Seat	90/10	100	1200	LR	82.67	86	83	83
	80/20	200	2400		84.67	88	85	85
	70/30	300	3600		84.55	88	85	85
	60/40	400	4800		84.62	88	85	86
	90/10	100	1200	RF	93.33	93	93	93
	80/20	200	2400		93.7	94	94	94
	70/30	300	3600		93.33	94	93	93
	60/40	400	4800		94.04	94	94	94

We reported our experimental results of this evaluation scenario in Table1 and Table 2, in which each table shows the number of tested samples per user, the total number of tested samples for the 12 users, the type of used machine learning model, and the four performance measurement metrics. By looking at the tables, we found two observations: First, the classifier "RandomForest" provides overall performance results better than the "LogisticRegression" classifier for both belt-only and seat-only datasets. Second, the performance results of the belt-only dataset are better than those of the seat-only dataset for both RF and LR classifiers and across all train/test data split ratios. For example, the average RF classifier results of the Belt-only pressure for 12 users reached above 99% F1 score and above 99% accuracy using 80/20 and 90/10 train/test split ratios. However, the LR classifier achieved average results of only 96% F1 score and 96.25% accuracy using the same conditions. Similarly, The average results of the RF classifier for seat-only achieved 94% F1 score and accuracy, which are better performance than those of the LR classifier of only 86% F1 score and 84.62%. This emphasizes that the RF classifier provides better authentication performance for our work. However, by looking at the time consumption aspect, we found that LogisticRegression is faster than RandomForest as shown in Figure 6. In detail, we also measured the time consumption required for training and testing the two classifiers when conducting the experiments mentioned in the above tables. For every user, we measured train time and test time across the four train/test split ratios using the two classifiers. Then we averaged the total train time and test time for the 12 users and plotted them in Figure 6a and in Figure 6b for RF and LR classifiers. As expected, we found that for both belt-only and seat-only datasets, the RF classifier consumes more time than the LR classifier across four train/test split ratios. For example, the LR classifier only consumes up to 107 ms for training (when Tr/Ts: 90/10) and 7.4 ms for testing (when Tr/Ts: 60/40), while the RF classifier consumes up to 6 seconds for training (when Tr/Ts: 90/10) and 0.5 seconds for testing (when Tr/Ts: 60/40) for the belt-only dataset. For the Seat-only module, the number of samples is more than the belt-only module which increases time consumption using the two classifiers as shown in Figure 6b– however, the LR is still much lower and faster than the RF classifier. At the end of this evaluation scenario, we conclude that even though the RF classifier provides better performance accuracy, it consumes much more time because of the hyperparameter "n_estimators = 1000" used for decision trees for classification. Although the time consumption of the RF classifier is higher, it is still within a few seconds for training (which occurs only one time when creating the model) and less than a second for testing – which can be acceptable and fast for the purpose of driver authentication. Therefore, to keep the good performance, we decided to select the RF model for the rest evaluations in this paper.



(a) Belt-based train/test time

(b) Seat-based train/test time

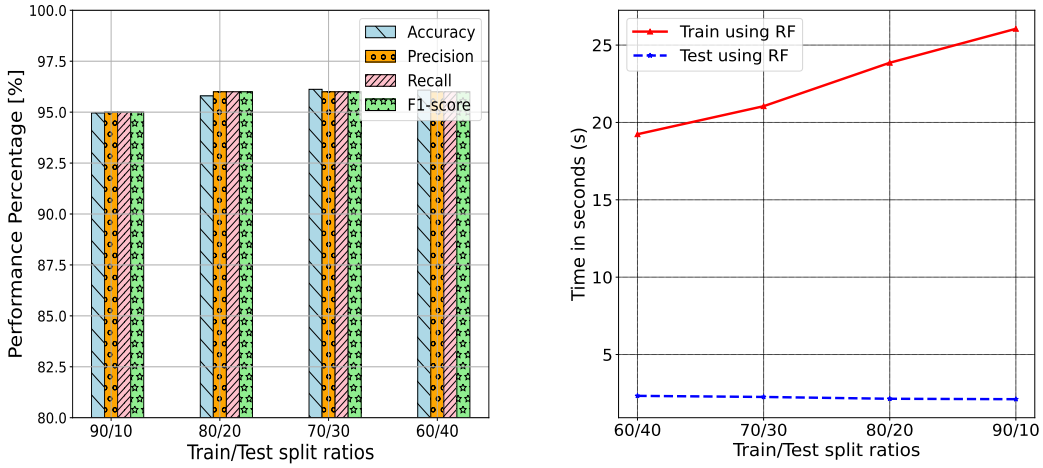
Figure 6: Average timing distributions required for training/testing models of Belt-based and Seat-based datasets using two different machine learning algorithms.

3.3 Evaluation Scenario 3: Performance of fusion datasets

In the previous two evaluation scenarios, we conducted experiments to check the performance of separate belt-only and seat-only modules. Here, we evaluate the fusion dataset from the belt and seat. Since the number of pressure sensors attached to the belt and seat are equal (*i.e.,* 30 sensors for each), we vertically concatenated all dataframes of belt and seat datasets for each user to build a fusion dataset. After that, we trained the RF classifier and tested it using four train/test split ratios similar to the implementation of the previous sections. Our fusion experiment results are reported in Figure 7. In detail, we reported accuracy, precision, recall, and F1 score for four train/test split ratios as well as computed the corresponding train/test consumption times. We found that 70/30 data size provides the best performance with 96.12 % accuracy and 96% F1 score; while consuming about 19.2 seconds for training and 2.3 seconds for testing. Note that, even though the performance of the fusion module is lower than the belt-only module (which is 99.7%), it is better than the seat-only module (which is 94%). In practice, the belt-only data may be difficult to collect in real life since the attacker (or the unauthorized users) may not wear the belt in order to avoid detection. Thus, considering the fusion dataset of both seat and belt is important for the real applications of driver authentication. In the end, our findings and results of the three evaluation scenarios show the feasibility and the effectiveness of the work for supporting research in driver authentication.

4 Related work

In this section, we explain the related works published for performing driver authentication. Many previous studies have focused on collecting sensory data from the controller area network (CAN) from various devices and circuits in the car. Xun et al. [22] proposed an automobile fingerprinting scheme by studying the behavioral characteristics of the drivers using data



(a) Seat and Belt fusion performance

(b) Seat and Belt fusion time consumption

Figure 7: Average performance and time consumption results of the 12 users using the fusion of seat and belt datasets.

collected from the CAN bus in two vehicles. They implemented two models using a convolution neural network (CNN) and support vector machines (SVM) to evaluate the system performance. Banerjee et al. [9] presented an approach to authenticating drivers using features extracted from GPS-only data, and achieved an overall mean area under the receiver operating characteristic curve (AUC) of 0.9. Regani et al. [14] exploited the unique radio-based biometric signals in the channel state information (CSI) for recognizing driver identity. They started by addressing the in-car change environment to overcome the problem of existing wireless sensing-based identification systems, and they achieved an accuracy of up to 99.13% for the best case. Kwak et al. [18] investigated the auto-theft of cars caused by the increasing number of computerized electric circuits. The author analyzed the measurements collected from a sensor in the vehicle and selected the most significant features to develop a driver verification method with less time cost and good detection accuracy. Burton et al. [17] addressed the problem of carjacking that may occur in the middle of a driving session. The author focused on collecting sensory data on driver habits such as steering wheel and pedal pressure to extract features based on individuals' driving behaviors for authentication purposes. Recently, Xun et al. [23] addressed the problem of a fixed number of drivers based on their biometrics behaviors and proposed a growable scheme for adding a new authorized driver to the system. The idea was based on performing incremental learning in dynamic scenarios where the number of drivers can be increased during the authentication. Xun et al. [24] established a multi-task learning network for driver authentication by extracting behavioral characteristics from the CAN bus data to conduct three tasks which are illegal driver detection, legal driver identification, and driving behavior evaluation.

Another direction of driver authentication research is to exploit the face fingerprinting and eye movements of the driver by using camera recording. Taha et al. [11] proposed a biometric-based continuous authentication by developing end-to-end learning of eye movement profiles and producing embeddings for identifying drivers. Borghi et al. [25] investigated that monitoring

the upper body and head pose of the driver is the key task for driver authentication. The author developed a framework using deep-learning for computing the accuracy of identifying depth images based on head localization and pose estimation. Derman et al. [12] introduced a continuous driver authentication using face recognition and its features extracted from pre-trained convolution neural networks of 52 different subjects. Gupta [26] addressed the problem of the reliability of drivers for on-demand rides and ride-sharing services such as Uber and Lyft companies. In fact, these services are based on a client-server infrastructure where the smartphone is the client and companies' servers manage the registration of drivers and customers. The author proposed a multi-model risk-based system based on the biometrics (*e.g.,* swipe, text-independent, voice, and face images) of drivers and customers to make the on-demand ride and ride-sharing services more safer and secure. Also, many other proposals utilize facial features and use face recognition technology for developing solutions for detecting several issues of drivers while driving the cars such as driver fatigue detection, Distracted driving detection, and Driver drowsiness detection [27, 28, 29, 27].

On the basis of previous research, we explore their shortcomings such as involving sensory data from many devices on the CAN bus in the car which are prone to cyber-threats and need a long time (a couple of minutes) to produce outputs and authenticate drivers, requiring sensitive data that disclose the privacy of driver's locations and time such as GPS, or exploiting wireless sensing of the radio signal that suffer from severe multipath propagation [30, 31, 32]. Also, we found that most of the previous studies were based on simulation datasets collected from simulation tools on computers. All the above issues may limit the practical employment of these schemes for real-world applications. Our work is different since it uses a practical dataset collected from a single pressure sensor that can be attached to the driver's seat and belt, and can authenticate the driver in a few seconds with high accuracy as well as preserve the privacy aspects of users.

5 Conclusion and Future work

We designed a novel driver authentication scheme that achieves fast and accurate performance using pressure data attached to the driver's seat and belt. We conducted real experiments and collected real pressure datasets (rather than simulation data) of 12 users and built up a database in our server for later learning and classification processes. We conducted extensive evaluation implementations to prove the validity of our work under various evaluation settings and different pressure data sizes. Our results show the promising aspects of the pressure-based system in a fast and accurate way.

For future work, we plan to extend the work as follows. First, continue gathering pressure datasets while driving the car from a larger and more diverse pool of users. This can help enhance the system's adaptability to various driving conditions and driver profiles. Second, extend the research to focus on mid-session driver authentication. This will be valuable for scenarios where driver changes occur during a journey or for additional security measures during longer trips. Third, we will address situations where the driver may board under various conditions different from the training period. For example, changes in body weight, injuries, or seat adjustments. Lastly, we plan to conduct a full application design of a client-server architecture to send a warning alarm to the owner's smartphone in case of detecting a car's unauthorized access.

Acknowledgement

This work was supported by the Institute of Information & Communications Technology Planning & Evaluation (IITP),(Project No. RS-2022-00165794, Development of a Multi-Faceted Collection-Analysis-Response Platform for Proactive Response to Ransomware Incidents, 30%), (Project No.2022-0-00701, 10%, Project No.RS-2023-00228996, 10%), the ICT R&D Program of MSIT/IITP, (Project No. 2021-0-01816, A Research on Core Technology of Autonomous Twins for Metaverse, South Korea, 10%), and a National Research Foundation of Korea (NRF), South Korea grant funded by the Korean government (Project No. RS-2023-00208460, 40%).

References

[1] Jiajia Liu, Shubin Zhang, Wen Sun, and Yongpeng Shi. In-vehicle network attacks and countermeasures: Challenges and future directions. *IEEE Network*, 31(5):50–58, 2017.

[2] Bing Zhu, Zhipeng Liu, Jian Zhao, Yizhou Chen, and Weiwen Deng. Driver behavior characteristics identification strategies based on bionic intelligent algorithms. *IEEE Transactions on Human-Machine Systems*, 48(6):572–581, 2018.

[3] Miro Enev, Alex Takakuwa, Karl Koscher, and Tadayoshi Kohno. Automobile driver fingerprinting. *Proc. Priv. Enhancing Technol.*, 2016(1):34–50, 2016.

[4] Chiyomi Miyajima, Yoshihiro Nishiwaki, Koji Ozawa, Toshihiro Wakita, Katsunobu Itou, Kazuya Takeda, and Fumitada Itakura. Driver modeling based on driving behavior and its evaluation in driver identification. *Proceedings of the IEEE*, 95(2):427–437, 2007.

[5] Sahil Garg, Kuljeet Kaur, Neeraj Kumar, and Joel JPC Rodrigues. Hybrid deep-learning-based anomaly detection scheme for suspicious flow detection in sdn: A social multimedia perspective. *IEEE Transactions on Multimedia*, 21(3):566–578, 2019.

[6] Saad Ezzini, Ismail Berrada, and Mounir Ghogho. Who is behind the wheel? driver identification and fingerprinting. *Journal of Big Data*, 5(1):1–15, 2018.

[7] Cheng Zhang, Mitesh Patel, Senaka Buthpitiya, Kent Lyons, Beverly Harrison, and Gregory D Abowd. Driver classification based on driving behaviors. In *Proceedings of the 21st International Conference on Intelligent User Interfaces*, pages 80–84, 2016.

[8] Yijie Xun, Yuanyuan Sun, and Jiajia Liu. An experimental study towards driver identification for intelligent and connected vehicles. In *ICC 2019-2019 IEEE International Conference on Communications (ICC)*, pages 1–6. IEEE, 2019.

[9] Tanushree Banerjee, Arijit Chowdhury, Tapas Chakravarty, and Avik Ghose. Driver authentication by quantifying driving style using gps only. In *2020 IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops)*, pages 1–6. IEEE, 2020.

[10] Sebastian Bittl, Arturo A Gonzalez, Matthias Myrtus, Hanno Beckmann, Stefan Sailer, and Bernd Eissfeller. Emerging attacks on vanet security based on gps time spoofing. In *2015 IEEE Conference on Communications and Network Security (CNS)*, pages 344–352. IEEE, 2015.

[11] Bilal Taha, Sherif Nagib Abbas Seha, Dae Yon Hwang, and Dimitrios Hatzinakos. Eyedrive: A deep learning model for continuous driver authentication. *IEEE Journal of Selected Topics in Signal Processing*, 2023.

[12] Ekberjan Derman and Albert Ali Salah. Continuous real-time vehicle driver authentication using convolutional neural network based face recognition. In *2018 13th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2018)*, pages 577–584. IEEE, 2018.

[13] Mohamed Aly. Real time detection of lane markers in urban streets. In *2008 IEEE intelligent vehicles symposium*, pages 7–12. IEEE, 2008.

[14] Sai Deepika Regani, Qinyi Xu, Beibei Wang, Min Wu, and KJ Ray Liu. Driver authentication for smart car using wireless sensing. *IEEE Internet of Things Journal*, 7(3):2235–2246, 2019.

[15] Shihong Duan, Tianqing Yu, and Jie He. Widriver: Driver activity recognition system based on wifi csi. *International Journal of Wireless Information Networks*, 25:146–156, 2018.

[16] Mohsen Ali Alawami, Aishwarya Ram Vinay, and Hyoungshick Kim. Locid: A secure and usable location-based smartphone unlocking scheme using wi-fi signals and light intensity. *IEEE Internet of Things Journal*, 9(23):24357–24372, 2022.

[17] Angela Burton, Tapan Parikh, Shannon Mascarenhas, Jue Zhang, Jonathan Voris, N Sertac Artan, and Wenjia Li. Driver identification and authentication with active behavior modeling. In *2016 12th International Conference on Network and Service Management (CNSM)*, pages 388–393. IEEE, 2016.

[18] Byung Il Kwak, JiYoung Woo, and Huy Kang Kim. Know your master: Driver profiling-based anti-theft method. In *2016 14th Annual Conference on Privacy, Security and Trust (PST)*, pages 211–218. IEEE, 2016.

[19] Minh Van Ly, Sujitha Martin, and Mohan M Trivedi. Driver classification and driving style recognition using inertial sensors. In *2013 IEEE Intelligent Vehicles Symposium (IV)*, pages 1040–1045. IEEE, 2013.

[20] Toshihiro Wakita, Koji Ozawa, Chiyomi Miyajima, Kei Igarashi, Katunobu Itou, Kazuya Takeda, and Fumitada Itakura. Driver identification using driving behavior signals. *IEICE TRANSACTIONS on Information and Systems*, 89(3):1188–1194, 2006.

[21] Xiaoning Meng, Ka Keung Lee, and Yangsheng Xu. Human driving behavior recognition based on hidden markov models. In *2006 IEEE International Conference on Robotics and Biomimetics*, pages 274–279. IEEE, 2006.

[22] Yijie Xun, Jiajia Liu, Nei Kato, Yongqiang Fang, and Yanning Zhang. Automobile driver fingerprinting: A new machine learning based authentication scheme. *IEEE Transactions on Industrial Informatics*, 16(2):1417–1426, 2019.

[23] Yijie Xun, Wei Guo, and Jiajia Liu. G-driveraut: A growable driver authentication scheme based on incremental learning. *IEEE Transactions on Vehicular Technology*, 2023.

[24] Yijie Xun, Jiajia Liu, and Zhenjiang Shi. Multitask learning assisted driver identity authentication and driving behavior evaluation. *IEEE Transactions on Industrial Informatics*, 17(10):7093–7102, 2020.

[25] Guido Borghi, Marco Venturelli, Roberto Vezzani, and Rita Cucchiara. Poseidon: Face-from-depth for driver pose estimation. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 4661–4670, 2017.

[26] Sandeep Gupta, Attaullah Buriro, and Bruno Crispo. Driverauth: A risk-based multi-modal biometric-based driver authentication scheme for ride-sharing platforms. *Computers & Security*, 83:122–139, 2019.

[27] Suhandi Junaedi and Habibullah Akbar. Driver drowsiness detection based on face feature and perclos. In *Journal of Physics: Conference Series*, volume 1090, page 012037. IOP Publishing, 2018.

[28] Jianju Xing, Guoxin Fang, Juping Zhong, and Jun Li. Application of face recognition based on cnn in fatigue driving detection. In *Proceedings of the 2019 International Conference on Artificial Intelligence and Advanced Manufacturing*, pages 1–5, 2019.

[29] Belhassen Akrout and Walid Mahdi. A novel approach for driver fatigue detection based on visual characteristics analysis. *Journal of Ambient Intelligence and Humanized Computing*, 14(1):527–552, 2023.

[30] Sai Deepika Regani, Qinyi Xu, Beibei Wang, Min Wu, and K. J. Ray Liu. Driver authentication for smart car using wireless sensing. *IEEE Internet of Things Journal*, 7(3):2235–2246, 2020.

[31] Yongsen Ma, Gang Zhou, and Shuangquan Wang. Wifi sensing with channel state information: A survey. *ACM Comput. Surv.*, 52(3), jun 2019.

[32] Mohsen A Alawami and Hyoungshick Kim. Locauth: A fine-grained indoor location-based authentication system using wireless networks characteristics. *Computers & Security*, 89:101683,

2020.