Advances in Reliability, Safety and Security, Part 3 – Kołowrocki, Dąbrowska (eds.)
© 2024 Polish Safety and Reliability Association, Gdynia, ISBN 978-83-68136-15-9 (printed), ISBN 978-83-68136-02-9 (electronic)

Advances in Reliability, Safety and Security

ESREL 2024
Monograph Book Series

Utilization Of AI-Based Driver Style Detection Using Vehicle Dynamics Measurements

Marcin Hinz, Matthias Becker, Peter E. Pfeffer

Munich University of Applied Sciences, Munich, Germany

Abstract

This paper presents a comprehensive study on the application of artificial intelligence (AI) for the detection of driving styles, utilizing data derived from vehicle dynamics measurements. The research was conducted using an advanced driving simulator setup, which combined physical components of a vehicle's operating system with a virtual driving scenario. This approach facilitated the collection of detailed data from various proband groups, encompassing both human drivers and automated driving models. The experimental setup included a multifunction steering wheel, control levers, accelerator and brake pedals, and a force feedback actuator, all integrated into a simplified vehicle cockpit mock-up. Visual, acoustic, and kinesthetic feedback mechanisms were employed to provide a realistic driving experience. Data acquisition focused on driver interface signals, vehicle state signals, and traffic object signals, captured with high precision to reflect driving behaviors. The core of the study involved the extraction and analysis of features from the collected data and the employment of the TSFresh library and supervised learning algorithms, including Decision Trees, Random Forest, and Support Vector Machines (SVM). Hyperparameter tuning was a critical aspect of this process, ensuring the optimization of each algorithm despite the constraints posed by the limited size of the dataset. Our results demonstrate the efficacy of the employed algorithms in accurately classifying driver types, with a considerable level of consistency across different models. However, the study also revealed challenges, particularly the misclassification of specific driving patterns, underscoring the need for more extensive and diverse datasets for training and validation.

Keywords: car dynamics, driver recognition, artifitial intelligence, supervised learning

1. Introduction

In this paper, we present an extensive study into the detection of driving styles using artificial intelligence (AI), focusing on the analysis of vehicle dynamics measurements. Our study is set apart by the use of a driving simulator to collect diverse and comprehensive data from different proband groups, allowing for a detailed analysis of driver behavior in a controlled environment. This setup comprised both physical elements of a vehicle's operating system and a virtual driving scenario, providing a holistic approach to data acquisition.

The experimental setup involved a multifunction steering wheel, control levers, and pedals within a simplified vehicle cockpit mock-up, offering realistic driving input elements. Feedback to the driver was given through visual, acoustic, and kinesthetic channels, including a curved wide screen for environmental visualization and a high-performance force feedback actuator for realistic steering feel. Our setup did not include vehicle motion in terms of translational and rotational accelerations, focusing instead on low lateral dynamics scenarios.

Two computing units processed the inputs and outputs of the simulator, with a real-time capable industrial PC handling sensor inputs and executing a two-degree-of-freedom steering system model. The second unit ran the vehicle dynamics simulation and visualization. This comprehensive setup allowed us to create a realistic driving environment and collect high-fidelity data on driver behavior and vehicle dynamics.

Our study employed an innovative approach to data manipulation and feature extraction, utilizing the TSFresh library to generate a substantial number of features from the time domain data. The focus was on extracting meaningful patterns from driver interface signals, vehicle state signals, and traffic object signals, each meticulously recorded during the simulation runs. This approach allowed us to capture the nuanced aspects of driving behavior, which are crucial for developing an accurate AI model for driver style detection.

In the learning process with supervised learning algorithms, we applied Decision Trees (Gareth, 2013), Random Forests (Breiman Leo, 2001), and Support Vector Machines (SVM) (Mammone, 2009), to analyze the extracted features. The limited size of our dataset, comprising data from 20 human drivers and 32 automated driver model runs, necessitated a focus on these specific algorithms and emphasized the importance of hyperparameter tuning for optimization. Despite the dataset size constraints, our models achieved high accuracy levels in classifying driver types, demonstrating the effectiveness of our approach. However, the occasional misclassification of one subject as an algorithmic driver highlights the need for larger and more diverse datasets to refine the model's predictive capabilities.

The diverse studies in the realm of driver behavior analysis illustrate various methodologies and focuses. Dimitar Filev's (2009) research concentrates on real-time characterization of driver behavior using vehicle dynamics and driver actuations, primarily for identifying aggressive and cautious behaviors. Kaustubh V. Sakhare (2020) shifts the focus to the use of vision systems in autonomous vehicles, utilizing computer vision techniques for dynamic vehicle detection. Derick A. Johnson (2011) introduces a novel approach using smartphone-based sensor fusion and Dynamic Time Warping to detect aggressive driving behaviors.

Lei's (2016) paper takes a predictive modeling approach, classifying driving styles into categories like sports and economical, based on historical data analysis. Joshi's (2017) research proposes an algorithm for evaluating driver performance using real-time vehicle state data, focusing on steering and pedal inputs. Lastly, Patrick Brombacher (2017) employs artificial neural networks for classifying driving styles, based on driving maneuvers identified during test trips.

Contrasting with these approaches, our study harnesses AI-based methods, specifically supervised learning algorithms, to analyze driver styles using vehicle dynamics measurements. Our methodology is distinct in its rigorous feature extraction from vehicle dynamics data and the application of various machine learning algorithms, including Decision Trees, Random Forest, and SVM. This choice is driven by the need to address the challenges posed by a smaller dataset, making our study unique in its approach to optimizing these algorithms through hyperparameter tuning and achieving consistent results across different models.

This study contributes to the field of AI-based driver style detection by providing a comprehensive analysis of vehicle dynamics measurements, enhanced by our novel experimental setup and methodical approach to data processing and machine learning application.

2. Experimental setup

In order to acquire driving style sample data from different proband groups, a driving simulator setup was utilized. The setup consisted of physical parts of a vehicle's operating elements exchanging control signals with a virtual driving scenario. The specific input and feedback elements to carry out the driving task are described as follows. Regarding the input elements, a multifunction steering wheel, control levers attached to the steering column as well as accelerator and brake pedals were installed in a simplified vehicle cockpit mock-up, which is illustrated in Figure 1.

The feedback elements served the drivers' visual, acoustic, and kinaesthetic channels of perception. For the visual feedback of the virtual environment, two displays came to use. A curved wide screen display visualized the present scenario including the road, scenery, rear view mirrors, and traffic. A second display was placed above the steering column acting as an instrument cluster providing mainly vehicle speed information. Driving sound outputs, such as noise from rolling tires, engine, airstream and passing vehicles were provided through loudspeakers. Lastly, a force feedback actuator was attached to the steering column which generated validated steering feel corresponding to the dynamic vehicle state (Joerg et al, 2023). The aspect of vehicle motion in terms of translational and rotational accelerations acting on the driver himself was not implemented in the static experimental setup and therefore limited the experimental scenarios to driving with low lateral dynamics.

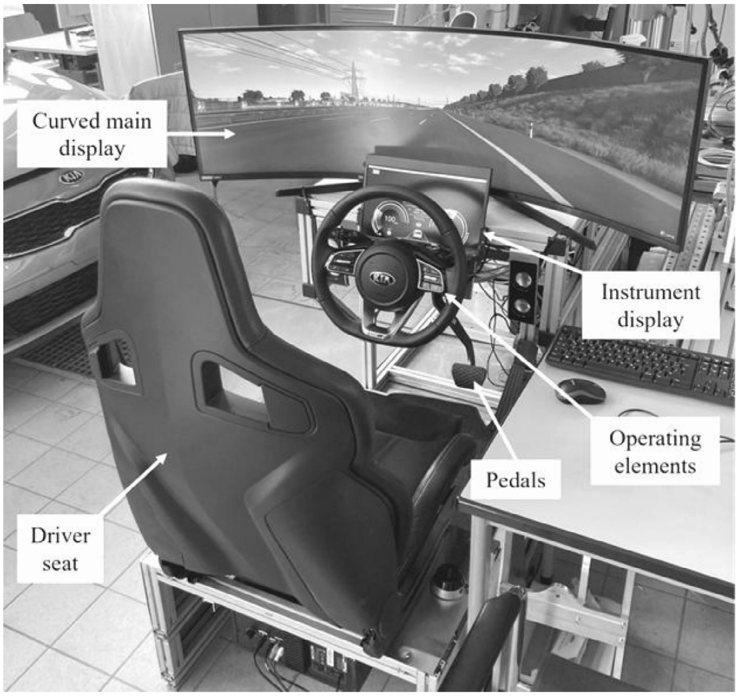

Fig. 1. Experimental static driving simulator setup.

The processing of the physical driving simulator input and output interfaces and calculation of simulation models was distributed on two computing units. A real-time capable industrial PC collected the sensor inputs coming from the steering feedback actuator (steering wheel angle, velocity, and acceleration), from the pedals (accelerator and brake pedal value) and from all buttons and levers. This PC performs pre-processing of the input signals and executes a two degree of freedom steering system model. It passes the virtual tie rod displacement, which represents the steering interface to the vehicle model, together with the remaining input signals to the second computing unit. On this second unit runs the vehicle dynamics simulation in soft real time as well as the maneuver configuration and rendering of the visualization. The feedback forces for the steering system are calculated inside the vehicle dynamics model and sent back to the real-time PC, where a corresponding steering wheel torque is calculated and controlled by the feedback actuator. A schematic topology of the setup components and interfaces is shown in Figure 2.

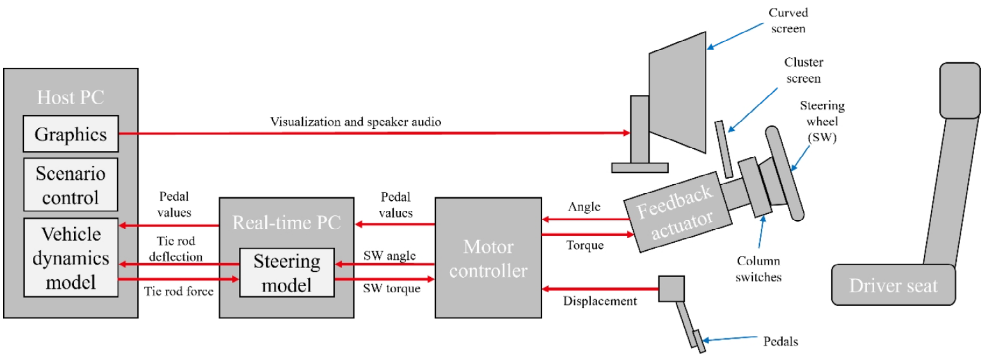

Fig. 2. Driving simulator setup topology.

The simulated scenario used to identify driving styles consisted of driving along a three-lane highway section of 6.6 km length modeled from real road data. It contained road markings, lane barriers, traffic signs (directions and speed limits), highway access and exit lanes as well as multiple traffic objects as shown in Figure 3. The initial conditions of all traffic objects like driving lane, starting position, time of appearance and initial speed, were kept identical in every run. However, the traffic objects were configured to interact with other traffic objects including the ego vehicle, which resulted in varying traffic situations depending on the proband's driving pattern.

Fig. 3. Road and environment visualization from the driver's perspective.

2.1. Description of the signals

For the objective of methodically characterizing driving styles, several driver- and vehicle related measurement quantities were acquired in the form of time domain data during each experiment run.

Table 1. Driver interface signals.

Signal label	Signal description	Unit
VC.Gas	Accelerator pedal value (between 0 and 1)	1
VC.Brake	Brake pedal value (between 0 and 1)	1
Steer.WhlAng	Steering wheel angle	rad
Steer.WhlTrq	Steering wheel torque	Nm

The data was grouped into three groups: driver interface signals, vehicle state signals and traffic object signals. Table 1 to 3 list the specific measurement signals of each group. A data acquisition sample time of 0.001 seconds was chosen to sufficiently cover the bandwidth of driver, vehicle, and traffic dynamics.

Table 2. Vehicle state signals.

Signal label	Signal description	Unit
Car.tx, .ty, .tz	Global x-, y- and z-position	m
Car.vx, .vy, .vz	Velocity in x-, y- and z- direction	m/s
Car.v	Absolute velocity	m/s
Car.ax, .ay, .az	Acceleration in x-, y- and z- direction	m/s²
Car.Roll, .Pitch, .Yaw	Roll, pitch and yaw angle	rad
Car.RollVel, .PitchVel, .YawVel	Roll, pitch and yaw rate	rad/s
Sensor.ObjByLane.OB00.LaneDiff, .tPOI, .tPath	Lateral offset with respect to original driving lane	m
Sensor.Road.RD01.Path.DevAng	Heading angle with respect to lane direction	rad
Vhcl.Distance	Driven distance from simulation start	m

The traffic object detection was carried out in the current driving lane of the ego vehicle as well as in the left and right neighbouring lane. In case no neighbouring lane existed because the ego vehicle was driving in the leftmost or rightmost lane, detection for the corresponding neighbouring lane was inactive. For each lane, relative signals always refer to the closest vehicle in front of the ego vehicle in the direction of the corresponding lane. A maximum detection range of traffic objects in front was set to 200 m. Outside the detection range, signals were marked invalid.

Table 3. Traffic object signals, the asterisk* is placeholder for lane index left (L), right (R) and ego/center (C).

Signal label	Signal description	Unit
Sensor.ObjByLane.OB00.Lanes*.0.ObjF.0.VelLong	Absolute velocity of traffic object in left, right or ego lane	m/s
Sensor.ObjByLane.OB00.Lanes*.0.ObjF.0.sMax, .sMin	Maximum and minimum relative longitudinal distance	m
Sensor.ObjByLane.OB00.Lanes*.0.ObjF.0.tMax, .tMin	Maximum and minimum relative lateral distance	m
Sensor.ObjByLane.OB00.Lanes*.0.nObjF	Number of objects in left, right, and ego lane	1

2.2. Driver description

Following the introduced approach, two experimental datasets were generated. The first dataset contains recordings from 20 individual human drivers. The drivers' ages ranged between 20 and 30 years while the driving experience ranged from little to moderate. After the drivers were introduced to the experimental driving simulator setup, they were instructed to drive one scenario run for settling in and one subsequent run which was eventually recorded. The driving task during experimental runs of this proband group was to follow the predefined road while maintaining their personal driving preferences, especially regarding respect of traffic rules, driving speed, lane choice and handling of takeover situations.

The second dataset was created by running 32 variations of an automated parametrizable driver model. The variations were formed by permuting four different driver model attributes specified in Table 4.

Table 4. Driver model attribute variations.

Driver model attribute	Variation values
Driving style preset (influencing maximum applied longitudinal and lateral accelerations and decelerations)	normal defensive
Minimum time gap to traffic objects ahead	1.5 s 1.0 s
Minimum differential velocity for initiating a takeover maneuver	10 km/h 5 km/h
Desired driving velocity	100 km/h 130 km/h 160 km/h 200 km/h

Since the scenario length was fixed to 6.6 km, different average driving speeds resulted in different durations of recorded data.

3. Data manipulation and feature extraction

Our study focuses on the utilization of artificial intelligence (AI) in detecting driver styles using vehicle dynamics measurements. This comprehensive analysis entails a meticulous process of data ingestion, graphical analysis, consistency checks, feature extraction, and selection, ensuring the integrity and applicability of the data for developing a robust AI model.

Initially, each dataset was individually ingested, providing a detailed base for our analytical procedures. The graphical plotting of these datasets was not just a preliminary step but a crucial one, as it allowed us to visually inspect the data for underlying patterns, anomalies, or any form of irregularities that might skew our analysis. The use of advanced plotting techniques facilitated a deeper understanding of the multidimensional nature of vehicle dynamics data, which includes parameters like acceleration, braking patterns, steering angle, and vehicle speed.

Following the graphical analysis, a critical phase was the inspection for inconsistencies within the data. This step was crucial in ensuring the reliability of our dataset for AI processing. The inspection involved checking for NaN (Not a Number) values, constant value signals, and any other forms of data anomalies. Remarkably, none of the datasets required cleaning, a testament to the meticulous data collection process. This rarity in data quality meant that each dataset was viable for feature extraction without the need for further preprocessing.

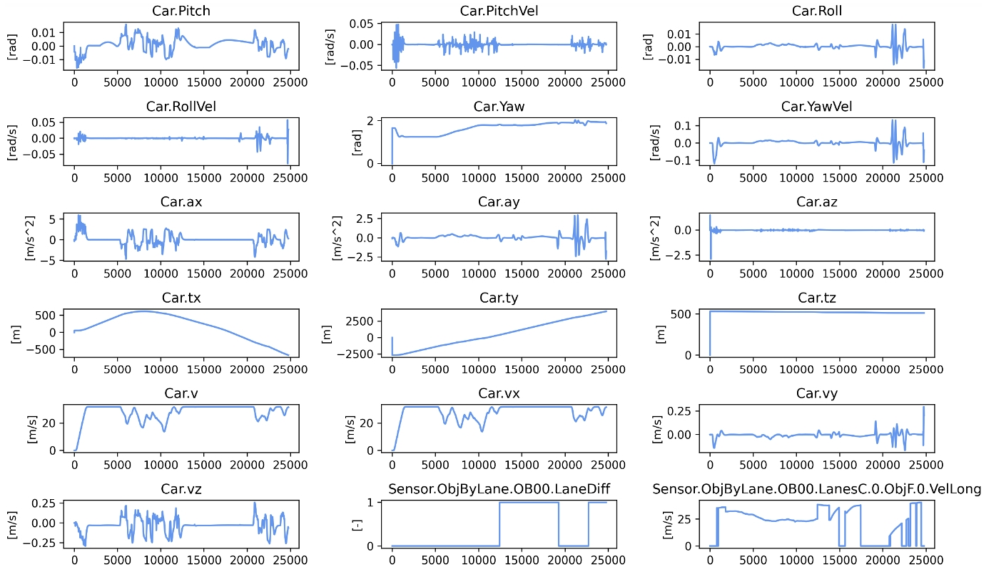

Fig. 4. Excerpt of the graphical representation of the recorded signals (18 out of 45).

A segment of the signals analyzed is exemplified in Figure 4 (excerpt). This illustration not only serves as a representative sample of the data but also highlights the complexity and diversity of the vehicle dynamics measurements. The signals, varying in amplitude, trend, and dynamic behavior, underscore the need for a sophisticated approach to feature extraction.

For the extraction of features, we employed the TSFresh library, a powerful tool for time series data processing. The EfficientFCParameters setting was chosen for this purpose. This choice was pivotal in generating a comprehensive set of features - a total of 35,235, to be precise. An initial trial was also conducted using the MinimumFCParameters setting; however, given the complexity of our data, EfficientFCParameters was deemed more suitable with an acceptable time and resource effort. It offered a wider range of features capturing the intricate aspects of driver behavior as reflected in the vehicle dynamics data.

To manage the vast number of features generated and to identify the most significant ones, we calculated feature importance using the Random Forest algorithm. This method allowed us to sift through the plethora of features and focus on those most relevant to driver style detection (according to the given target variable, as described in the following section). Furthermore, to assess the significance of the extracted features, we introduced an additional randomized variable into the feature matrix. This approach served as a benchmark - any feature less significant than this random variable was deemed non-essential and subsequently eliminated. This rigorous selection process resulted in approximately 12,000 features being retained. These selected features, now refined and concentrated, form the foundation for our subsequent analysis involving supervised learning techniques. The random forest was trained with various numbers of trees ending up by the number of 50,000. The according train test split for the analysis of the feature importance was set to 80-20%.

In addition to the comprehensive analysis detailed earlier, a significant aspect of our study is depicted in Figure 5, which presents a visual representation of the feature importance as determined by the Random Forest algorithm. This figure specifically illustrates a bar plot of the top 20 features, offering a clear insight into the features that most significantly influence the driver style detection model.

What stands out in this visualization, as well as a further analysis of the most important features not present in this figure, is the recurrent appearance of certain sensor data, notably those labeled "Sensor.ObjByLane", "Steer.WhlTrq", and "Vhcl.Distance". The frequent occurrence of these specific sensors among the top-rated features is particularly noteworthy. It indicates that these sensors play a pivotal role in the assessment of vehicle dynamics and, consequently, in the determination of driver style for this dataset and the drivers in question.

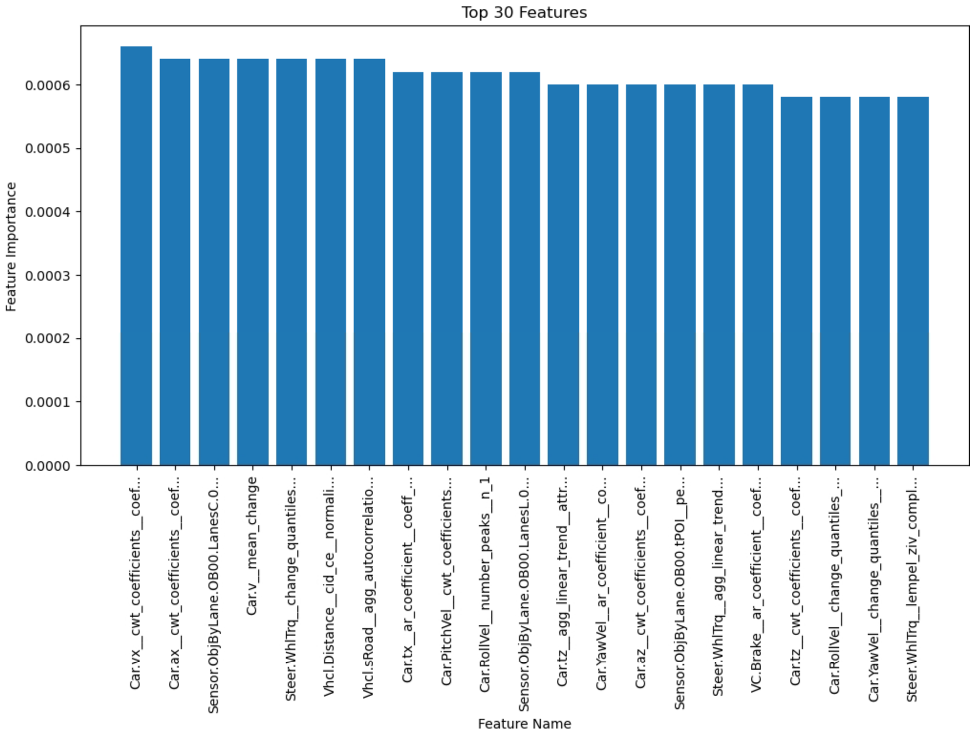

Fig. 5. Top 20 features identified by Random Forest algorithm.

The predominance of these sensors in the top features of our model highlights their collective importance in the accurate assessment of driving dynamics. It suggests that a driver's interaction with their environment, their steering behavior, and their acceleration patterns are among the most defining elements of their driving style. This insight is invaluable in further refining our AI model, ensuring it captures the most relevant and impactful aspects of driver behavior.

The meticulous process of data preparation, from graphical analysis to feature extraction and selection, ensures the integrity and applicability of our dataset. This rigorous approach lays the groundwork for developing an AI model capable of accurately detecting and categorizing driver styles based on vehicle dynamics measurements.

4. Learning process with supervised learning algorithms

In the quest to advance the field of driver style detection using artificial intelligence, our study embarked on an exploration of supervised learning algorithms to analyze extracted and reduced features from vehicle dynamics measurements. The essential part of our research involved the meticulous application of three distinct machine learning algorithms: Decision Trees, Random Forest, and Support Vector Machines (SVM). The choice of these algorithms was primarily driven by the limited scope of our dataset, which comprised data from 20 human drivers and 32 automated driver model runs. This dataset size, while significant for initial studies, is relatively small compared to the vast data requirements of deep learning models, thereby necessitating the use of more traditional machine learning approaches.

A pivotal aspect of our methodology was the optimization of each selected algorithm through rigorous hyperparameter tuning. This process is fundamental in machine learning, as it involves fine-tuning various parameters within each algorithm to identify the most effective configuration for the specific characteristics of our dataset. The tuning process is both an art and a science, requiring a deep understanding of the algorithms and a methodical approach to testing different parameter combinations.

Despite the inherent differences in the theoretical foundations and operational mechanisms of Decision Trees, Random Forest, and SVM, our study revealed a remarkable consistency in the results obtained from these algorithms. This consistency not only reinforces the robustness and reliability of our feature selection process but also highlights the versatility of the features in adapting to different machine learning techniques. Such adaptability is crucial in the field of AI, where the choice of algorithm can significantly impact the model's performance and accuracy.

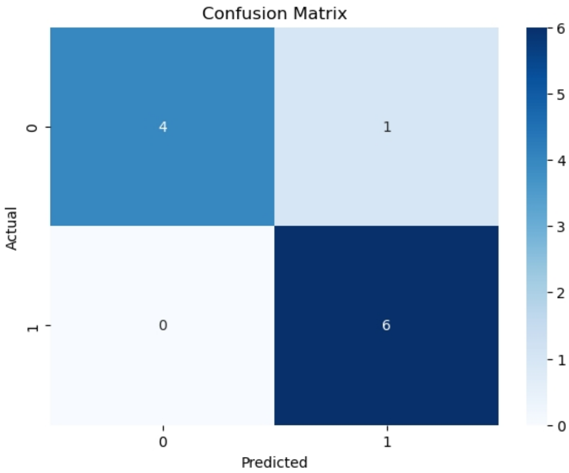

Fig. 5. Confusion Matrix with the averaged results across all models.

A key element in evaluating the effectiveness of our classification models was the use of confusion matrices. An illustrative example, represented in Figure 6, provides a detailed breakdown of the successes and failures of the model in predicting the target variable – the type of driver, classified as either human or algorithmic. The confusion matrix is an invaluable tool in machine learning, offering insights into the precision and recall of the model, and helping identify areas where the model may be misclassifying data.

The consistent application of an 80-20 train-test split across all algorithms ensured the robustness of our testing process. This standard split, where 80% of the data is used for training the model and 20% for testing its predictive power, is a widely accepted practice in machine learning, providing a balanced approach to model training and validation.

Table 5. Classification report.

Class	Precision	Recall	F1-Score	Support
1.0	1.00	0.80	0.89	5
2.0	0.86	1.00	0.92	6
accuracy			0.91	11
macro avg	0.93	0.90	0.91	11
weighted avg	0.92	0.91	0.91	11

Interestingly, while our models achieved high accuracy levels, consistently classifying the majority of drivers correctly, there was a recurring anomaly (as shown in Table 5). One particular subject was consistently misclassified as an algorithmic driver in the test dataset. This consistent misclassification raises intriguing questions about the unique aspects of this subject's driving style, which the models consistently interpreted as non-human. Such findings highlight the nuances and complexities involved in driver style detection and underscore the need for models to be adaptable and sensitive to a wide range of driving behaviors.

In conclusion, our study demonstrates that while the employed machine learning algorithms are highly effective in classifying driver types based on vehicle dynamics data, the limited size of the dataset poses a constraint on the overall validity of the model. The occasional misclassification, particularly of one subject as an algorithmic driver, indicates that our model, although robust, requires a larger and more diverse dataset to enhance its accuracy and reliability. Future research should focus on expanding the dataset, incorporating a broader range of driving styles and scenarios, to refine the model's predictive capabilities further and ensure its applicability across a wider spectrum of driving behaviors.

5. Summary and outlook

Our research has provided valuable insights into the utilization of artificial intelligence for detecting driving styles using vehicle dynamics measurements. The study used a state-of-the-art driving simulator to gather data from a variety of proband groups. This setup enabled the collection of rich data sets encompassing a range of driver interface signals, vehicle state signals, and traffic object signals. The experimental approach and subsequent analysis demonstrated the potential of AI, specifically supervised learning algorithms like Decision Trees, Random Forest, and SVM, in accurately classifying driver types.

The feature extraction process, conducted with the TSFresh library, was pivotal in identifying relevant driving behavior patterns. Despite the limited size of the dataset, the study achieved high accuracy in classifying driver types, although it also revealed challenges such as the misclassification of specific driving patterns. These findings underscore the complexity of driver style detection and the critical role of comprehensive data in developing effective AI models.

A primary focus of future studies should be the expansion of the human data set. Recording data from a larger and more diverse group of human drivers, encompassing different age groups, driving experiences, and various backgrounds, is essential. This expansion will not only enhance the robustness of the AI models but also ensure their applicability and accuracy across a broader spectrum of the driving population. A further focus relies on the repeatability of the same drives performed by certain drivers to achieve a dataset with consistent driving characteristics.

In conclusion, while our study has made significant strides in the field of AI-based driver style detection, there is a clear path forward for further research. Expanding the dataset is the crucial step in refining and enhancing the accuracy and applicability of AI models in understanding and predicting driver behavior.

References

Breiman, L. 2001. Random Forests. In: Machine Learning 45(1), 5–32.

Brombacher, P., Masino, J., Frey, M., Gauterin, F. 2017. Driving event detection and driving style classification using artificial neural networks. 2017 IEEE International Conference on Industrial Technology (ICIT) Toronto, ON: IEEE, 997–1002.

James, G., Witten, D., Hastie, T., Tibshirani, R. 2013. An introduction to statistical learning. With applications in R. New York, Heidelberg, Dordrecht, London: Springer (Springer texts in statistics).

Filev, D., Lu, J., Prakah-Asante, K., Tseng, F. 2009. Real-time driving behavior identification based on driver-in-the-loop vehicle dynamics and control. 2009 IEEE International Conference on Systems, Man and Cybernetics, San Antonio, TX, USA, IEEE, 2020–2025.

Johnson, D. A., Trivedi, M. M. 2011. Driving style recognition using a smartphone as a sensor platform. 2011 14th International IEEE Conference on Intelligent Transportation Systems (ITSC). Washington, DC, USA, IEEE, 1609–1615.

Joerg, K., Becker, M., Thaler, K., Pfeffer, P. E. 2023. Steering Feel Simulation with a High Performance Force Feedback Steering Wheel. Proceedings of the 28th Symposium of the International Association for Vehicle System Dynamics (IAVSD 2023).

Joshi, S. S., Maas, N., Schramm, D. 2017. A vehicle dynamics based algorithm for driver evaluation. 2017 11th International Conference on Intelligent Systems and Control (ISCO). Coimbatore, India, IEEE, 40–44.

Lei, Y., Liu, K., Fu, Y., Li, X., Liu, Z., Sun, S. 2016. Research on driving style recognition method based on driver's dynamic demand. Advances in Mechanical Engineering 8(9).

Mammone, A., Turchi, M., Cristianini, N. 2009. Support vector machines. WIREs Computational Stats 1(3), 283–289.

Sakhare, K. V. Tewari, T., Vyas, V. 2020. Review of Vehicle Detection Systems in Advanced Driver Assistant Systems. Arch Computat Methods Eng 27(2), 591–610.