robotics

MDPI

Article

How AI from Automated Driving Systems Can Contribute to the Assessment of Human Driving Behavior

Tom Driessen [1], Olger Siebinga [1], Thomas de Boer [1], Dimitra Dodou [1], Dick de Waard [2] and Joost de Winter [1,*]

[1] Faculty of Mechanical Engineering, Delft University of Technology, 2628 CD Delft, The Netherlands; t.driessen@tudelft.nl (T.D.); o.siebinga@tudelft.nl (O.S.); thomas.boer.rc@gmail.com (T.d.B.); d.dodou@tudelft.nl (D.D.)

[2] Faculty of Behavioural and Social Sciences, University of Groningen, 9712 TS Groningen, The Netherlands; d.de.waard@rug.nl

* Correspondence: j.c.f.dewinter@tudelft.nl

Abstract: This paper proposes a novel approach to measuring human driving performance by using the AI capabilities of automated driving systems, illustrated through three example scenarios. Traditionally, the assessment of human driving has followed a bottom-up methodology, where raw data are compared to fixed thresholds, yielding indicators such as the number of hard braking events. However, acceleration threshold exceedances are often heavily influenced by the driving context. We propose a top-down context-aware approach to driving assessments, in which recordings of human-driven vehicles are analyzed by an automated driving system. By comparing the human driver's speed to the AI's recommended speed, we derive a level of disagreement that can be used to distinguish between hard braking caused by aggressive driving and emergency braking in response to a critical event. The proposed method may serve as an alternative to the metrics currently used by some insurance companies and may serve as a template for future AI-based driver assessment.

Keywords: artificial intelligence; context-aware driving assessment; hard braking; emergency braking; aggressive driving style

check for updates

Citation: Driessen, T.; Siebinga, O.; De Boer, T.; Dodou, D.; De Waard, D.; De Winter, J. How AI from Automated Driving Systems Can Contribute to the Assessment of Human Driving Behavior. *Robotics* **2024**, *13*, 169. https://doi.org/10.3390/robotics13120169

Academic Editor: Girijesh Prasad

Received: 2 October 2024
Revised: 7 November 2024
Accepted: 18 November 2024
Published: 21 November 2024

Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Human drivers are increasingly being evaluated by algorithms, particularly in terms of fuel efficiency and safety. Safe driving is of interest to insurance companies, fleet owners, and licensing organizations. For example, some insurance companies now offer incentives for defensive driving based on acceleration-based metrics such as hard braking (exceeding longitudinal acceleration thresholds), sharp cornering (exceeding lateral acceleration thresholds), and speeding [1–6]. Similarly, Tesla's Safety Score Beta calculates a score based on factors like hard braking, speeding, and following distance [7].

While studies have found that acceleration-based metrics can be used to predict drivers' likelihood of being involved in crashes and damage incidents (e.g., [8–11]), there are challenges in holding drivers accountable for exceeding acceleration thresholds. One issue is that the need for high acceleration levels can depend on the road type. For example, urban areas typically require more frequent (hard) braking than rural roads. One solution is to make the driver's assessment dependent on the road type based on GPS (e.g., [12–14]). However, such location-based assessments lack the ability to account for local road geometry.

A second concern is that traffic conditions can play a key role. For example, an unexpected event, such as a pedestrian crossing the road or a lead vehicle suddenly decelerating, may force the driver to brake hard. When evaluating drivers based on deceleration events or offering discounts and rewards for not showing such events, it is important to consider the traffic conditions in which these actions occur. One approach to include contextual information is to compare the speed of the vehicle with the speed of

other traffic or to take weather conditions into consideration (e.g., [11,15,16]). However, such global statistical measures will only provide limited context.

A third concern is the risk of unintended consequences. If drivers are penalized for hard braking, they may hesitate to brake hard when needed, leading to dangerous situations. In summary, simply counting "dangerous events" based on predefined acceleration thresholds may not be a reliable indicator of an individual's driving behavior.

In recent years, computational resources in vehicles and the sophistication of algorithms have increased to the point that vehicles have become advanced enough to perform most of the driving tasks themselves. These developments can be seen in systems like those of Waymo [17,18], Lyft [19,20], Tesla's "Full Self-Driving" (FSD), and AI/data companies such as NVIDIA [21] and comma.ai [22,23]. A dilemma, however, arises from the fact that these systems are not yet perfect. They either have to operate in limited regions or still require human attention and intervention. For example, current Tesla FSD systems require drivers to keep their hands on the wheel or remain attentive to the road, as monitored by a cabin camera [24], whereas the autonomous fleets of Waymo and Zoox rely on remote operators to solve difficult situations [25,26] (see also [27]).

Given that vehicles are becoming increasingly capable of driving automatically in a human-like manner, yet still cannot drive wholly automatically, we propose the concept of using automation to assist in measuring human driving behavior. The idea we explore is to evaluate human driving by comparing it to AI-generated driving behavior. This concept can be traced back over three decades to the GIDS (Generic Intelligent Driver Support) project in the late 20th century [28], where a reference driving behavior was generated through computer-simulated driving. However, at the time, this idea was considered "(too) far ahead of its time" [29] and did not gain traction. Today, however, it has become a realistic possibility.

Comparing human driving behavior to an AI's intended plan offers advantages over the aforementioned assessment strategies. For example, in the case of a detected hard-braking event, an AI system could be used to retrospectively differentiate between reckless driving and a sudden, unavoidable event requiring an immediate response. A high level of disagreement between the driver and the AI may suggest reckless driving, where both the AI and the human would have had the time to respond to a visible obstruction. On the other hand, a high level of agreement could indicate a justified reaction, where neither the driver nor the AI could anticipate the need to brake sooner. An algorithm based on fixed acceleration thresholds, however, would flag both scenarios as instances of unsafe driving.

It is important to note that our proposed approach does not assume the AI is infallible. The idea is that even an imperfect AI can be effective for assessing human driving behavior, as long-term data collection may reveal that certain drivers consistently deviate more from AI-referenced behavior than others.

Aim

This study presents three driving scenarios conducted in a driving simulator. After recording, the scenarios were analyzed using Openpilot, an open-source platform [30] that can be described as SAE level 2 vehicle automation [31]. Online videos [32,33] demonstrate that Openpilot is capable of driving vehicles both on highways and in urban environments. We fed the recordings of our driving scenarios to the Openpilot system and analyzed the internal metrics of Openpilot to understand how it interprets each situation.

This study aims to present a technical demonstration of how AI can be used to assess human driving behavior. Using three driving scenarios in a simulator environment, we demonstrate the feasibility of comparing human driving behavior with driving behavior of an existing automated driving system.

Our scenarios include hard braking, which is practically relevant for applications like insurance assessments. However, the technical framework we demonstrate is adaptable to other scenarios due to the generalizable nature of the automated driving system. In the discussion section, we illustrate this broader potential impact with specific examples.

Additionally, we contribute a publicly accessible adaptation of Openpilot's tools, which enables the replay of externally recorded driving data to support future research.

2. Method

2.1. Setup

Three demonstration driving scenarios were driven by the first author of this paper. The video and vehicle data were recorded in a virtual world using JOAN [34], a Python software package developed to enable human-in-the-loop experiments in the CARLA driving simulator [35]. JOAN provides the possibility to connect USB steering wheels to CARLA vehicles (in this case, the Logitech Driving Force G923 steering wheel), record vehicle data, and create reproducible experiments with other traffic following predefined trajectories. The simulation ran on a PC running Windows 11. The repository contains the experimental setup files that were used within JOAN.

2.2. Scenarios

The three scenarios were chosen to be simple to explain and feasible for simulator implementation without the need for complex hard-coded choreographies. Moreover, the scenarios were chosen to make our specific point: that AI-based assessments can be used to help classify a hard-braking event as either unnecessary or necessary. In other words, while traditional assessment methods, such as those used by insurance companies, would classify every hard-braking event as undesirable and contributing to the driver's risk profile, we illustrate that hard braking in a surprise emergency condition does not need to be marked as undesirable but rather as desirable and necessary.

The scenarios involved two situations where an obstruction was clearly visible and the driver of the ego-vehicle reacted either aggressively (i.e., braked too late) or calmly (i.e., braked in time), and a third scenario where a surprise event justified hard braking by the driver. These scenarios are illustrated in Figure 1, where orange represents the obstructions encountered in the *Aggressive* and *Calm* scenarios, and pink indicates the moving obstruction in the *Surprise* scenario. Figure 2 shows a view during the *Surprise* scenario, and Figure 3 shows the buses from the *Aggressive* and *Calm* scenarios. Video files of the three scenarios can be found in the repository.

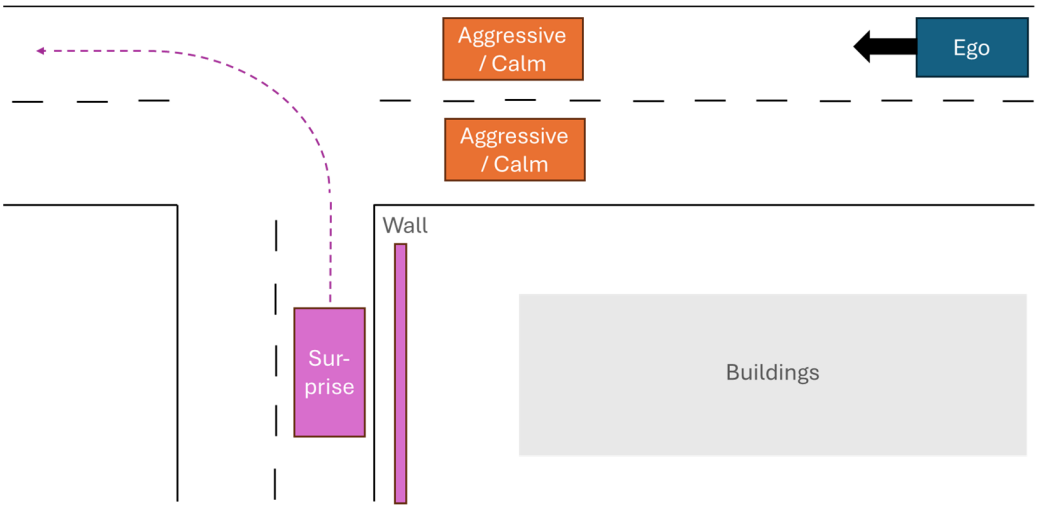


Figure 1. Demonstration scenarios. *Aggressive* and *Calm*: The ego-vehicle (blue) approached stationary buses (orange) at a T-junction requiring a full stop, which the driver of the ego-vehicle executed aggressively or calmly. *Surprise*: The ego-vehicle drove along an empty road until a bus (pink) emerged from behind a wall, necessitating an emergency brake to avoid a collision.



Figure 2. Video stills from the *Surprise* scenario.

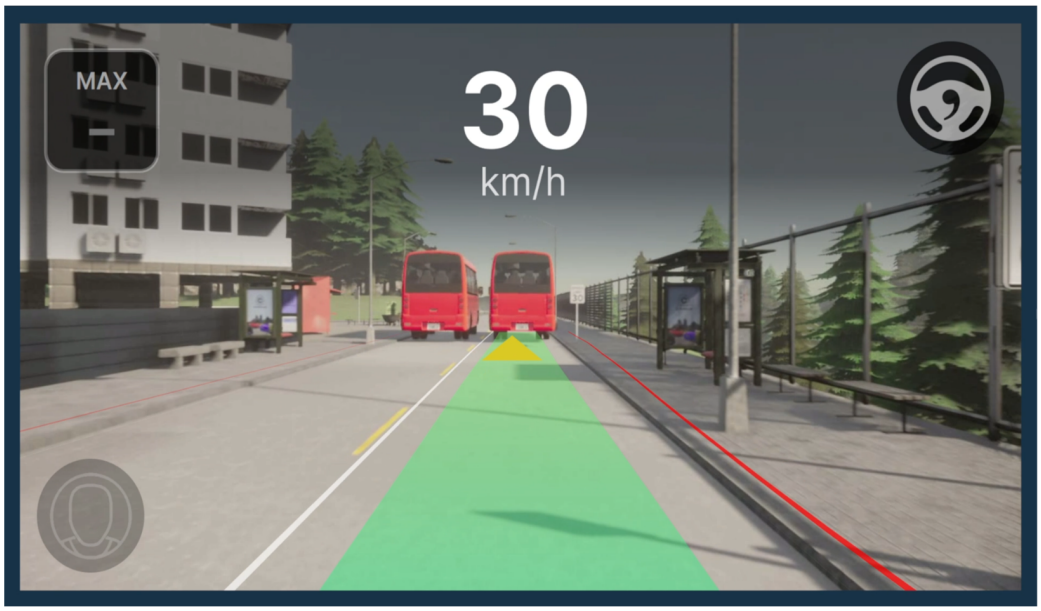


Figure 3. Openpilot user interface during the *Calm* scenario. The *Aggressive* scenario uses the same setup. The green band displays the predicted trajectory, and the lines provide information about the lane line estimates. The yellow triangle indicates a detected lead vehicle.

In all scenarios, the ego-vehicle began from the same starting point and followed a straight road leading to a T-junction (Figure 1). The driver controlling the vehicle aimed at a target speed of 50 km/h.

2.2.1. Calm Scenario (Driver of Ego-Vehicle Did Not Brake Hard)

In the *Calm* scenario, two stationary buses were positioned on the road, forcing the ego-vehicle to stop behind them. The experimenter achieved this by braking moderately and gradually letting the vehicle come to a stop.

2.2.2. Aggressive Scenario (Driver of Ego-Vehicle Braked Hard, but Hard Braking Was Avoidable)

The *Aggressive* scenario was identical to the *Calm* scenario, except for the behavior of the ego-vehicle. The experimenter maintained the target speed until approaching the buses and then applied a high braking input at the last moment, to demonstrate an aggressive driving style.

2.2.3. Surprise Scenario (Driver of Ego-Vehicle Braked Hard, and Hard Braking Was Unavoidable)

The *Surprise* scenario contained an unforeseeable event. The scenario started with the ego-vehicle in the same position as in the previous two scenarios. Unlike the other scenarios, no stationary vehicles obstructed the road, allowing the ego-vehicle to proceed

straight through the T-junction. However, a bus (depicted in pink in Figure 1) approached the junction at the same time as the ego-vehicle. The bus was hidden from the ego-vehicle's view before entering the junction due to a high wall along the sidewalk. The bus followed a predefined path, turning left at the junction without yielding to the ego-vehicle (Figure 2). The timing of the bus's appearance forced the ego-vehicle to brake abruptly to avoid a collision.

2.3. Analysis

The Openpilot software was adapted to work with pre-recorded driving data to enable a comparison between human driving and AI-generated predictions. This adaptation involved modifying Openpilot's existing simulation module, originally designed to interface with the CARLA driving simulator for development purposes. Driving in the simulator with Openpilot running in real time could also have been a suitable setup for this demonstration. However, we first recorded the data and replayed it with Openpilot, as this also enables the module to work with other sources of data, such as dashcam videos supplemented with a CSV file containing speed and acceleration information (e.g., logged using a phone), allowing for the post-hoc analysis of other existing driving datasets.

The adapted module allows two primary inputs:

1. A 1928 × 1208 pixel forward-facing driving video recorded at 20 frames per second, simulating the visual input an autonomous system would receive.
2. A corresponding CSV file with a row for each video frame containing vehicle state data, including speed, bearing, steering angle, brake, and throttle inputs.

Thus, the video file provided the visual context, while the CSV file supplied data about the vehicle state. The module processed these inputs sequentially, mimicking the real-time data flow that Openpilot would experience in a live driving situation. Note that the module still works when omitting some of this data, such as brake and throttle input. This allows for some flexibility when using other datasets that may not feature all vehicle state data. Our limited experimentation with other sources suggests that reasonable estimates can often still be obtained. Figure 3 shows a screenshot of the Openpilot user interface during the replay of one of our recordings.

The software was run on an Ubuntu 20.04 desktop PC. As the input data were processed by Openpilot, the system's generated predictions and plans were logged to a CSV file for later analysis. For the current demonstration, we used the desired speed originating from Openpilot's longitudinal planner module. Other variables logged by the module that are suitable for comparison to human driving execution but not used are discussed in Appendix A. The full implementation details can be found in the project repository.

3. Results

Figure 4 compares Openpilot's desired speed with the actual speed at which the human drove. A positive difference (shown in green) can be interpreted as a desire by the model to obtain a higher speed (a desire to accelerate), while a negative difference (shown in red) indicates a desire to decelerate.

3.1. Calm Scenario

In the *Calm* scenario, the driver drove slower than the AI deemed appropriate for the current situation (Figure 4 left, annotation A). As the two stationary buses became closer, the driver decelerated, and the AI indicated a similar desire. At some point, the model suggested a speed lower than the current speed (Figure 4 left, B), indicating the AI preferred to slow down faster for the oncoming obstacles. The human driver also decelerated around the same time, indicated by the steeper slope, coming to a full stop.

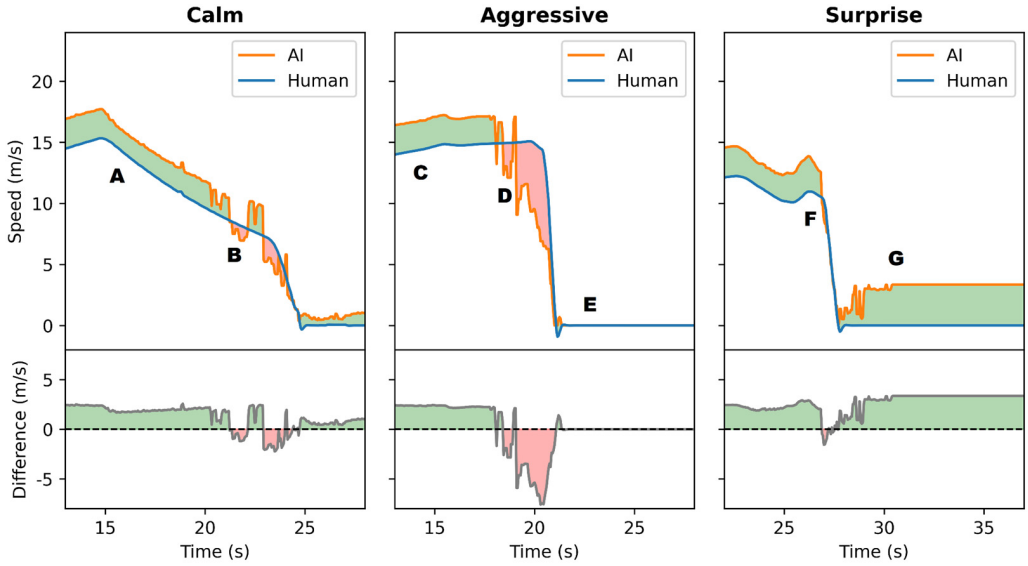


Figure 4. Measured speed (blue), AI's desired speed (orange), and the difference between the two, for all scenarios. To reduce noise in the visualization, a median filter with a window size of 3 samples (0.15 s) was applied to the prediction signal. The letters denote specific moments in the scenarios that are discussed in the text.

3.2. Aggressive Scenario

In the *Aggressive* scenario, we see a similar start, with Openpilot suggesting a higher speed for the current empty road (Figure 4 middle, C). As the car approached the obstacles, the driver kept his speed, while the AI suggested deceleration (Figure 4 middle, D). Only at the last moment, the driver decided to brake, and the speed dropped to 0 (Figure 4 middle, E).

When comparing the *Calm* and the *Aggressive* scenarios, we see that, in both graphs, some level of disagreement was present. However, the graph of the *Aggressive* scenario clearly shows a larger negative difference.

3.3. Surprise Scenario

The *Surprise* scenario started similarly until the surprise event (a bus suddenly appearing and cutting off the path of the ego-vehicle) happened. The model only suggested deceleration after the surprise event, around the same time the human driver was decelerating (Figure 4 right, F). Further note that, after the full stop, the model suggested increasing speed again (Figure 4 right, G), since the bus had continued on its way and the ego-vehicle was alone again.

The *Aggressive* and *Surprise* scenarios would both likely have been flagged in traditional assessment methods. When presented with the *Aggressive* scenario, Openpilot suggested braking earlier, as judged by the difference graph at the bottom. In other words, it saw reason to decelerate before the human decided to decelerate. In the *Surprise* scenario, however, the model suggested deceleration at about the same time the human started decelerating. There was no large negative peak visible in the difference graph at the bottom, as the bus was a surprise for Openpilot as well. This means that this stop was executed in a manner that was more similar to the way the AI would have executed it than in the case of the *Aggressive* scenario. In other words, the AI's interpretation allows us to distinguish between an unnecessary hard brake and a necessary hard brake.

In the above comparisons, the height of the difference peak was highest in the *Aggressive* scenario, which means it could be identified as the braking event where there was the most disagreement between the AI and the human execution. This means that the current method provides a way to judge whether a hard brake event was justified after this brake event happened. A possible implementation could be, for every hard brake event (when a certain static threshold was exceeded), to let the AI assess whether the brake event was

justified, in order to improve the validity of the assessments of driver behavior. Other metrics that could be considered are the total area between the two curves or the time since the first model suggested deceleration until a full stop.

4. Discussion

This study demonstrated the potential of using AI from an automated driving system to assess human driving behavior. The method compared the state of a human-driven vehicle with the AI's desired state. This approach allowed discrimination between necessary and unnecessary hard braking by using a variable that indicates the level of disagreement of the human driver's actions relative to the AI's recommendations.

Traditional methods for assessing driver behavior, such as the predefined thresholds used by insurance companies (e.g., for speed, acceleration, or braking), are situation-agnostic. In contrast, the AI-assisted method provides a context-aware evaluation by comparing the human driver's actions with model predictions. Driving examiners previously identified the lack of a holistic perspective as a key barrier to effective data-driven driver assessment [36].

We achieved our results using Openpilot, a system that relies on a vision feed and CAN-bus data to control the vehicle automatically in many situations. One may argue that using Openpilot is overkill for the current demonstration purposes and that simpler, more widely used radar-based solutions, like those found in forward collision warning systems (FCWS) or adaptive cruise control (ACC), could have sufficed. An equivalent argument could be made that potential field methods, which examine the extent to which other objects and road users intrude in this field, could be used to quantify the level of risk in driving (e.g., [37–39]). These existing concepts, however, are still limited in their ability to interpret the overall driving context. Our approach relies on a visual understanding of the scenario, including, e.g., traffic signs. It uses a neural network [40] that provides an estimate of the appropriate speed given the entire driving situation, rather than providing an estimate of risk based on relatively simple indicators such as time to collision (TTC) or potential field intrusion. The system's ability to process visual information and make context-aware decisions mirrors the decision-making process of human drivers.

4.1. Limitations

The system that we used still has limitations. Like all current automated driving systems, it can make errors in perception, prediction, and decision-making. For example, it may misclassify objects, fail to anticipate complex traffic situations, or make wrong speed recommendations. However, our method's effectiveness does not require the AI to be infallible. Rather than relying on individual instances of disagreement, our approach is envisioned to identify patterns in driving behavior over extended periods. This aggregation helps mitigate the impact of occasional AI errors. Even with imperfect AI, the additional context helps distinguish between necessary and unnecessary hard braking events better than approaches relying on kinematic thresholds only.

As explained in the last paragraph of the Results section, the calculation and accompanying threshold values that distinguish a "high level of disagreement" from a "low level of disagreement" warrant further development and research with real datasets. Research will also be needed into "gray areas", where, for example, the AI detects a precursor to a hazard and considers braking intervention necessary while a human driver may barely recognize it. To this end, for different traffic situations, the driving behavior or recommendations of expert drivers could be compared to those of AI. This, in turn, raises fundamental questions about who should ultimately be the arbiter in defining successful driving performance: an expert driver or an AI agent. It also brings up questions about the required quality of cameras and the level of intelligence an AI-based automated driving system should or could possess.

The demonstrated scenarios represent only a fraction of potential applications. Future research could examine disagreements in lateral driving behavior, i.e., identify discrep-

ancies between a driver's chosen maneuver and the vehicle's suggested alternative. It should be noted here that, in some emergency situations, where the braking distance is too long to avoid a collision, an evasive steering maneuver may be the only way to prevent an accident [41]. Additionally, the method could assess whether higher-level strategic decisions (e.g., Michon's model of driving behavior; [42]) can be evaluated using multiple agents as reference points. As AI continues to improve, its ability to assess human driving behavior will likely expand. This is especially relevant because improved AI systems may not always result in safer automated driving. As AI takes more control, human drivers may be kept out of the loop for longer periods, potentially leading to slower response times when manual intervention is required.

Finally, it is important to note that this study serves as a proof of concept only; it represents a single demonstration conducted by the author in a virtual environment. Further research is needed to validate the approach, which should include testing the system across a broader range of real-world driving scenarios with more human participants.

4.2. Recommendations

In the insurance industry, our proposed method could allow for fairer assessments of driving incidents. This could lead to more precise risk profiling and personalized insurance premiums based on individual driving styles. For fleet operators, the ability to detect risky driving behaviors and offer targeted coaching could help reduce accidents, improve driver performance, and lower operating costs. Additionally, our algorithmic assessment concept could benefit driver education and testing, as well as provide concurrent feedback on driving behavior after obtaining a license.

In academia, our work opens up possibilities in several ways. First, existing datasets with human driving videos and car state variables could be augmented with AI-generated predictions, removing the need for costly new data annotation or data collection efforts. Discrepancies could be used to extract critical situations from the datasets, for example. Secondly, future research could explore the use of multi-agent systems, where disagreements between two or more concurrently observing AI models could provide a more robust assessment of driver behavior. Using multiple AI systems would essentially provide the driver with a "committee" of reference drivers, reducing the potential for bias present in individual models. An approach in which multiple agents arbitrate their perspectives on current driving events was previously suggested by Fridman et al. [43].

The current method assumes that hard braking events are evaluated after they occurred by observing whether the AI detected a need to decelerate earlier for these specific events. Future work could explore whether more continuous measures of disagreement can be developed. We propose an overall "deviation score", determined by the total level of disagreement between the driver and one (or multiple) AI(s) regarding variables such as speed, acceleration, steering input, and braking input.

The software was run on a desktop PC, with the data analyzed after the scenario recordings. However, future versions could allow real-time data collection in a real vehicle. Openpilot's plug-and-play design offers advantages over traditional onboard systems. It can be installed in different vehicle models [22]. In contrast, manufacturer-installed automation systems are typically integrated into onboard computers.

Though the model weights remain the private property of comma.ai, the software itself has been made open-source. The broader automotive industry still lags behind in embracing open-source frameworks. This closed nature may hinder innovation and restrict improvements in safety. We argue that manufacturers should, at a minimum, consider offering accessible APIs to allow developers and researchers to interact with vehicle data. This may be counterintuitive to traditional automotive manufacturers, where innovation is typically patented and kept private. However, some examples indicate that the industry is increasingly seeing the value in open-sourcing data and code. Waymo has released its Open Dataset [44], Tesla occasionally publishes repositories that may prove useful to outsiders [45], and the Automotive Grade Linux [46] initiative is another example. By

inviting public collaboration, manufacturers could benefit from faster development cycles, while the broader community gains access to tools for creating safer vehicles. When done right, such openness advances the company's technology and also strengthens the entire ecosystem.

Author Contributions: Conceptualization, T.D., O.S. and J.d.W.; methodology, T.D. and O.S.; software, T.D., O.S. and T.d.B.; validation, T.D. and J.d.W.; formal analysis, T.D.; investigation, T.D. and O.S.; resources, J.d.W.; data curation, T.D.; writing—original draft preparation, T.D. and J.d.W.; writing—review and editing, T.D., O.S., D.D., D.d.W. and J.d.W.; visualization, T.D.; supervision, D.D., D.d.W. and J.d.W.; funding acquisition, D.d.W. and J.d.W. All authors have read and agreed to the published version of the manuscript.

Funding: This research is supported by a Transitions and Behaviour grant [403.19.243] ("Towards Safe Mobility for All: A Data-Driven Approach"), provided by the Dutch Research Council (NWO).

Data Availability Statement: The demonstration recordings, analysis, and the Openpilot fork used in this project can be accessed via https://github.com/tomdries/AI-driving-assessment (accessed on 17 November 2024).

Conflicts of Interest: The authors declare no conflicts of interest.

Appendix A.

This appendix provides an overview of how Openpilot works, what parts of the software were modified, and what other variables are suited for similar analyses where human and AI plans are compared. For implementation details, we refer to the GitHub repository of this paper.

Appendix A.1. Openpilot Overview and Modifications

Openpilot's intended usage is in a live vehicle, where it is run on a comma device [22]. On the device, the vehicle's CAN-bus data are processed and combined with data from the device's sensors and two forward-facing cameras. These data are fed to a neural network, along with a data buffer of previous predictions, providing the model with a temporal context of 5 s. The model returns the location of road features and lead vehicles and creates a plan for the vehicle's coming states (i.e., where the vehicle wants to be). In addition, a second model can monitor the driver (e.g., distraction, face pose, phone usage, etc.); this module was disabled in the current work. Figure A1 gives an overview of the way the car, the comma device, and Openpilot interact under normal operating conditions.

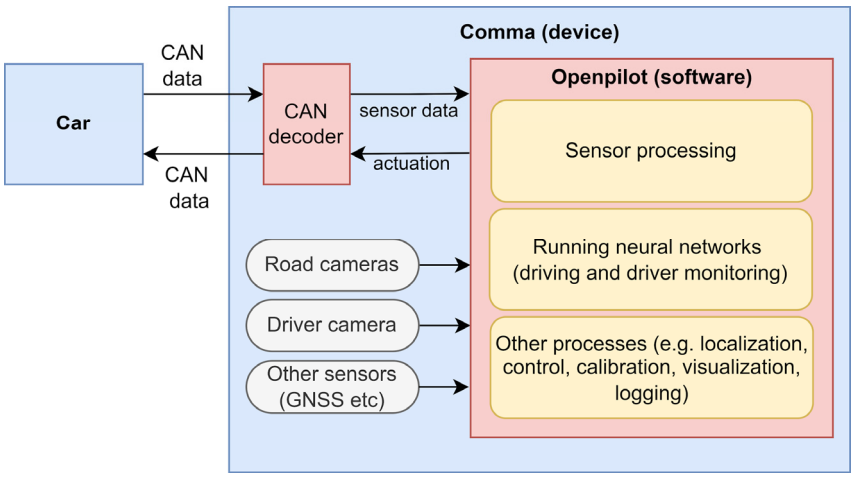


Figure A1. Diagram of Openpilot's intended use. Openpilot ingests vehicle data and data from the comma device and feeds this data to the neural networks for driving and driver monitoring. When Openpilot is in control of the vehicle, it provides steering, throttle, and brake commands to the car via the CAN interface.

The Openpilot repository contains developer tools that allow bridging Openpilot with a CARLA [47] driving simulator. However, this module requires a live simulator and does not allow replays of pre-recorded rides. While driving in the simulator in real time would have been a suitable setup for the current demonstration, we opted for an approach where we first recorded our data and then replayed the recording, with Openpilot observing the recordings in the background.

To achieve this, we modified the existing *sim* module so that it can ingest a forward-facing driving video (20 Hz, 1928 × 1208 pixels), along with an input CSV file that contains, for each video frame, car state data such as speed and bearing, as well as steering, brake, and throttle inputs. Note that the module still works when omitting some of these data, such as brake and throttle input. This allows for some flexibility when using other datasets that may not feature all vehicle state data. However, it comes at the expense of the model prediction and planning quality (from our limited experimentation with other sources, reasonable estimates can often still be obtained).

Appendix A.2. Other Variables

Although not directly used in the current demonstration, other variables are being logged in the current implementation of the module. These variables could be relevant for the evaluation of human driving. Noteworthy are variables contained in the *lateralPlan* and the *metaPredictions* structures of the model. The lateral plan contains the desired curvature (rad/s) and curvature rate (rad/s^2) predictions. These values could be compared to human execution of trajectories, for example, when merging or overtaking. Moreover, there are variables that represent the model's planned "lateral desire", which can take on the values *none*, *turnLeft*, *turnRight*, *laneChangeLeft*, *laneChangeRight*, *keepLeft*, or *keepRight*. The meta-prediction data structure contains several probabilistic measures, such as the probability that a hard brake will be executed within the upcoming seconds or the probability of executing one of the maneuvers mentioned before.

The desired predictions can be considered higher-level decisions (in comparison to the predicted dynamics values, such as future speeds or accelerations). Comparison to human execution data becomes more complex in this area. However, the data do look promising: If the current paradigm of using AI as a reference can be extended to these decision-level measures, this opens up possibilities for even more holistic assessments of human driving.

Appendix A.3. Further Notes

The performance of the current module is sufficient for demonstration, but we still had some issues with high-frequency noise. Possible causes could be the implementation of the bridge model, incorrect calibration of vehicle data, mismatch in frame rates between Openpilot and the video, or bugs in the implementation of the simulator bridge.

References

1. Admiral. Black Box Insurance. 2024. Available online: https://www.admiral.com/black-box-insurance (accessed on 6 November 2024).
2. Allianz. BonusDrive. 2024. Available online: https://www.allianz.de/auto/kfz-versicherung/telematik-versicherung (accessed on 6 November 2024).
3. Allstate. Drivewise. 2024. Available online: https://www.allstate.com/drivewise (accessed on 6 November 2024).
4. ANWB. Veilig Rijden [Safe Driving]. 2024. Available online: https://www.anwb.nl/verzekeringen/autoverzekering/veilig-rijden/hoe-werkt-het (accessed on 6 November 2024).
5. Direct Assurance. YouDrive. 2024. Available online: https://www.direct-assurance.fr/nos-assurances/assurance-auto-connectee (accessed on 6 November 2024).
6. Nationwide. SmartRide. 2024. Available online: https://www.nationwide.com/personal/insurance/auto/discounts/smartride (accessed on 6 November 2024).
7. Tesla. Safety Score Beta: Version 2.1. Tesla Support. 2024. Available online: https://www.tesla.com/support/insurance/safety-score#version-2.1 (accessed on 6 November 2024).
8. Cai, M.; Yazdi, M.A.A.; Mehdizadeh, A.; Hu, Q.; Vinel, A.; Davis, K.; Xian, H.; Megahed, F.M.; Rigdon, S.E. The association between crashes and safety-critical events: Synthesized evidence from crash reports and naturalistic driving data among commercial truck drivers. *Transp. Res. Part C Emerg. Technol.* **2021**, *126*, 103016. [CrossRef]

9. Driessen, T.; Dodou, D.; De Waard, D.; De Winter, J.C.F. Predicting damage incidents, fines, and fuel consumption from truck driver data: A study from the Netherlands. *Transp. Res. Rec.* **2024**, *2678*, 1026–1042. [CrossRef]
10. Hunter, M.; Saldivar-Carranza, E.; Desai, J.; Mathew, J.K.; Li, H.; Bullock, D.M. A proactive approach to evaluating intersection safety using hard-braking data. *J. Big Data Anal. Transp.* **2021**, *3*, 81–94. [CrossRef]
11. Ma, Y.-L.; Zhu, X.; Hu, X.; Chiu, Y.-C. The use of context-sensitive insurance telematics data in auto insurance rate making. *Transp. Res. Part A Policy Pract.* **2018**, *113*, 243–258. [CrossRef]
12. Guillen, M.; Pérez-Marín, A.M.; Nielsen, J.P. Pricing weekly motor insurance drivers' with behavioral and contextual telematics data. *Heliyon* **2024**, *10*, e36501. [CrossRef] [PubMed]
13. Melman, T.; Abbink, D.; Mouton, X.; Tapus, A.; De Winter, J. Multivariate and location-specific correlates of fuel consumption: A test track study. *Transp. Res. Part D Transp. Environ.* **2021**, *92*, 102627. [CrossRef]
14. Moosavi, S.; Ramnath, R. Context-aware driver risk prediction with telematics data. *Accid. Anal. Prev.* **2023**, *192*, 107269. [CrossRef] [PubMed]
15. Masello, L.; Castignani, G.; Sheehan, B.; Guillen, M.; Murphy, F. Using contextual data to predict risky driving events: A novel methodology from explainable artificial intelligence. *Accid. Anal. Prev.* **2023**, *184*, 106997. [CrossRef] [PubMed]
16. Reig Torra, J.; Guillen, M.; Pérez-Marín, A.M.; Rey Gámez, L.; Aguer, G. Weather conditions and telematics panel data in monthly motor insurance claim frequency models. *Risks* **2023**, *11*, 57. [CrossRef]
17. Waymo. *Waymo One*. 2024. Available online: https://waymo.com (accessed on 6 November 2024).
18. Hu, X.; Zheng, Z.; Chen, D.; Sun, J. Autonomous vehicle's impact on traffic: Empirical evidence from Waymo open dataset and implications from modelling. *IEEE Trans. Intell. Transp. Syst.* **2023**, *24*, 6711–6724. [CrossRef]
19. Lyft. *Lyft*. 2024. Available online: https://www.lyft.com (accessed on 6 November 2024).
20. Li, T.; Han, X.; Ma, J.; Ramos, M.; Lee, C. Operational safety of automated and human driving in mixed traffic environments: A perspective of car-following behavior. *Proc. Inst. Mech. Eng. Part O J. Risk Reliab.* **2023**, *237*, 355–366. [CrossRef]
21. NVIDIA. Self-Driving Vehicles. 2024. Available online: https://www.nvidia.com/en-us/self-driving-cars (accessed on 6 November 2024).
22. comma.ai. Openpilot. 2024. Available online: https://comma.ai/openpilot (accessed on 6 November 2024).
23. Dorr, B. Prius Sets 'Autonomous' Cannonball Run Record with AI Driving Assistant. 2024. Available online: https://www.yahoo.com/tech/prius-sets-autonomous-cannonball-run-180347033.html (accessed on 6 November 2024).
24. Tesla. Full Self-Driving (Supervised). 2024. Available online: https://www.tesla.com/ownersmanual/modely/en_us/GUID-2CB60804-9CEA-4F4B-8B04-09B991368DC5.html (accessed on 6 November 2024).
25. Waymo. Fleet Response: Lending a Helpful Hand to Waymo's Autonomously Driven Vehicles. 2024. Available online: https://waymo.com/blog/2024/05/fleet-response (accessed on 6 November 2024).
26. Zoox. How Zoox Uses Teleguidance to Provide Remote Human Assistance to Its Autonomous Vehicles [Video]. YouTube. 2020. Available online: https://www.youtube.com/watch?v=NKQHuutVx78 (accessed on 6 November 2024).
27. Lu, S.; Shi, W. Teleoperation in vehicle computing. In *Vehicle Computing: From Traditional Transportation to Computing on Wheels*; Lu, S., Shi, W., Eds.; Springer: Cham, Switzerland, 2024; pp. 181–209. [CrossRef]
28. Michon, J.A. (Ed.) *Generic Intelligent Driver Support*; Taylor Francis Ltd.: London, UK, 1993.
29. Michon, J.A. GIDS: Generic Intelligent Driver Support. 1993. Available online: https://www.jamichon.nl/jam_writings/1993_car_driver_support.pdf (accessed on 6 November 2024).
30. comma.ai. Openpilot: An Operating System for Robotics. GitHub. 2024. Available online: https://github.com/commaai/openpilot (accessed on 6 November 2024).
31. Chen, L.; Tang, T.; Cai, Z.; Li, Y.; Wu, P.; Li, H.; Shi, J.; Yan, J.; Qiao, Y. Level 2 autonomous driving on a single device: Diving into the devils of Openpilot. *arXiv* **2022**. [CrossRef]
32. comma. *A Drive to Taco Bell* [Video]. YouTube. 2022. Available online: https://www.youtube.com/watch?v=SUIZYzxtMQs (accessed on 6 November 2024).
33. Greer Viau. I Turned My Toyota Corolla into a Self Driving Car [Video]. YouTube. 2021. Available online: https://www.youtube.com/watch?v=NmBfgOanCyk (accessed on 6 November 2024).
34. Beckers, N.; Siebinga, O.; Giltay, J.; Van der Kraan, A. JOAN: A framework for human-automated vehicle interaction experiments in a virtual reality driving simulator. *J. Open Source Softw.* **2023**, *8*, 4250. [CrossRef]
35. Dosovitskiy, A.; Ros, G.; Codevilla, F.; Lopez, A.; Koltun, V. CARLA: An open urban driving simulator. In Proceedings of the 1st Annual Conference on Robot Learning, Mountain View, CA, USA, 13–15 November 2017; Available online: https://proceedings.mlr.press/v78/dosovitskiy17a.html (accessed on 6 November 2024).
36. Driessen, T.; Picco, A.; Dodou, D.; De Waard, D.; De Winter, J.C.F. Driving examiners' views on data-driven assessment of test candidates: An interview study. *Transp. Res. Part F Traffic Psychol. Behav.* **2021**, *83*, 60–79. [CrossRef]
37. Hennessey, M.P.; Shankwitz, C.; Donath, M. Sensor-based virtual bumpers for collision avoidance: Configuration issues. In Proceedings of the PHOTONICS EAST '95, Philadelphia, PA, USA, 22–26 October 1995; Volume 2592, pp. 48–59. [CrossRef]
38. Li, L.; Gan, J.; Ji, X.; Qu, X.; Ran, B. Dynamic driving risk potential field model under the connected and automated vehicles environment and its application in car-following modeling. *IEEE Trans. Intell. Transp. Syst.* **2020**, *23*, 122–141. [CrossRef]
39. Kolekar, S.; Petermeijer, B.; Boer, E.; De Winter, J.; Abbink, D. A risk field-based metric correlates with driver's perceived risk in manual and automated driving: A test-track study. *Transp. Res. Part C Emerg. Technol.* **2021**, *133*, 103428. [CrossRef]

40. comma.ai. *Development Speed Over Everything* [Blog]. 11 October 2022. Available online: https://blog.comma.ai/dev-speed (accessed on 6 November 2024).
41. Allen, R.W.; Rosenthal, T.J.; Aponso, B.L. Measurement of behavior and performance in driving simulation. In Proceedings of the Driving Simulation Conference North America, 2005, Orlando, FL, USA, 30 November–2 December 2005; pp. 240–250. Available online: https://www.nads-sc.uiowa.edu/dscna/2005/papers/Measurement_Behavior_Performance_Driving_Simulation.pdf (accessed on 6 November 2024).
42. Michon, J.A. A critical view of driver behavior models: What do we know, what should we do? In *Human Behavior and Traffic Safety*; Evans, L., Schwing, R.C., Eds.; Springer: Boston, MA, USA, 1985; pp. 485–524. [CrossRef]
43. Fridman, L.; Ding, L.; Jenik, B.; Reimer, B. Arguing machines: Human supervision of black box AI systems that make life-critical decisions. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, Long Beach, CA, USA, 16–17 June 2019. [CrossRef]
44. Waymo-Research. Waymo-Open-Dataset. GitHub. 2024. Available online: https://github.com/waymo-research/waymo-open-dataset (accessed on 6 November 2024).
45. Tesla. Teslamotors. GitHub. 2024. Available online: https://github.com/teslamotors (accessed on 6 November 2024).
46. Sivakumar, P.; Neeraja Lakshmi, A.; Angamuthu, A.; Sandhya Devi, R.S.; Vinoth Kumar, B.; Studener, S. Automotive Grade Linux. An open-source architecture for connected cars. In *Software Engineering for Automotive Systems*; CRC Press: Boca Raton, FL, USA, 2022; pp. 91–110. [CrossRef]
47. comma.ai. Openpilot in Simulator. GitHub. 2024. Available online: https://github.com/commaai/openpilot/tree/master/tools/sim (accessed on 6 November 2024).

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.