Inferring Driver Behavior Profiles Using Digital Twins in Simulated Environments

Matheus Andrade†, Morsinaldo Medeiros†, Thaís Medeiros†, Marianne Silva†‡, Ivanovitch Silva†
†Federal University of Rio Grande do Norte, Natal, Brazil
‡Federal University of Alagoas, Penedo, Brazil
E-mails: {matheus.diniz.122, morsinaldo.medeiros.075, thais.araujo.707}@ufrn.edu.br,
marianne.silva@penedo.ufal.br and ivanovitch.silva@ufrn.br

***Abstract*—Understanding driver behavior is crucial for enhancing road safety and improving fuel efficiency. However, collecting data on these behaviors in real-world settings is challenging due to vehicle complexity and high instrumentation costs. Driving simulators offer a viable alternative by creating environments that replicate real-life situations. This study explores the use of digital twins in simulated environments, specifically in Euro Truck Simulator 2 (ETS2), to infer driver behavior using virtual sensors. A case study was conducted where a driver simulated routes under two driving conditions: cautious and aggressive. The telemetry data collected during these simulations were analyzed to identify behavioral patterns and assess fuel consumption efficiency. The results demonstrated that digital twins enable real-time capture of driver behavior information, revealing significant differences between driving styles. Analysis of the accumulated data and the radar area soft sensor indicated that cautious driving practices are associated with greater fuel efficiency.**

***Index Terms*—driver behavior analysis, digital twins, simulated environments, virtual sensors, and fuel efficiency.**

I. Introduction

The automotive sector has undergone significant transformations driven by technological advancements in areas such as the Internet of Things (IoT), artificial intelligence (AI), and, more recently, the concept of digital twins [1]. Digital twins, which are dynamic virtual representations of physical systems, enable real-time monitoring, simulation, and prediction of the behavior of objects or processes [2]. In the automotive context, they have shown great promise for vehicle optimization, safety enhancement, and detailed driver behavior analysis [3].

Understanding driver behavior is crucial for road safety, fuel efficiency, and vehicle wear [4]. For example, aggressive acceleration, abrupt braking, and inconsistent speed can indicate reckless or inefficient driving [5]. However, collecting behavioral data in real-world scenarios is challenging due to system complexity and high costs [6]. Driving simulators offer a practical alternative by creating virtual environments that replicate realistic scenarios. These simulations allow for the precise recording of interactions, facilitating detailed studies without the risks and expenses of real-world testing.

In this context, soft sensors—virtual sensors—are tools for estimating variables that are not easily measurable directly [7]. Combining data from physical sensors with computational models and machine learning algorithms can infer latent variables such as the driver's emotional state or propensity for risky behaviors [8]. In simulated environments, these sensors can monitor driving style in real-time, offering analyses of factors like stress, fatigue, and reaction patterns to different traffic scenarios [9]. Therefore, integrating soft sensors with simulated data enables the inference of driver profiles and enhances the understanding of how variables such as road conditions and vehicle characteristics affect the driver's interaction with the environment [10].

To facilitate these analyses, the Euro Truck Simulator 2 (ETS2)[1] stands out as a simulation platform that allows for the collection of detailed real-time driving data [11]. These data, processed and analyzed using Python, serve as a foundation for building models capable of inferring behavioral patterns.

In this article, we explore the use of digital twins in simulated environments to infer driver behavior using soft sensors within ETS2. The collected data enabled the identification of driving profiles based on two distinct behavioral patterns. The results demonstrate that it is feasible not only to infer driver behavior but also to highlight the benefits of using simulators as tools for data collection and improving safety and efficiency in transportation. Moreover, the simulated environment offers an effective method for testing and validating technologies aimed at vehicle data analysis.

II. Related Works

Several studies have explored using simulators to collect and analyze vehicle data, shedding light on recent advancements and highlighting specific gaps that warrant further investigation.

In the work by [12], the challenges faced by commercial drivers are examined, particularly the need to maintain high levels of attention and quick reactions under adverse conditions like fatigue and sleep deprivation. Utilizing Euro Truck Simulator 2 alongside the Varjo VR system equipped with eye-tracking technology, the study analyzes how drivers' attention spans fluctuate over time in fatigue-inducing situations. The results indicate that as fatigue intensifies, drivers are more likely to avert their gaze from the road and neglect the vehicle's instruments, thereby increasing the risk of incidents. However, this research does not investigate how these attention-related behaviors correlate with different driving styles, such as aggressive or moderate driving.

[1]https://eurotrucksimulator2.com/

Expanding on simulation for vehicular studies, [13] uses the Simulation of Urban Mobility (SUMO) to model urban traffic, highlighting its effectiveness for testing vehicular data architectures while reducing real-world risks. However, it does not develop virtual sensors for detailed vehicle dynamics.

Similarly, [14] employs SUMO to analyze vehicle emissions from fuels like ethanol and gasoline, offering a controlled environment for consistent results. Yet, this study also lacks focus on data collection integration and real-time predictive models. These studies underscore SUMO's potential in vehicular research while pointing to areas needing further exploration.

While these works significantly contribute to the field by demonstrating the utility of simulations in vehicular data analysis, specific gaps remain unaddressed. Notably, there is a lack of integration between driving simulators and data collection platforms to infer driver behavior through soft sensors. The present work aims to bridge these gaps by integrating Euro Truck Simulator 2 with vehicular data collection platforms and developing virtual sensors to infer driver profiles. This approach enhances the understanding of driver behavior and advances the application of simulated environments in automotive research.

III. Proposed Approach

This section presents the proposed methodology for collecting and inferring driver behavior within a simulation environment. As illustrated in Figure 1, the approach comprises five stages, each elaborated upon in the subsequent subsections.

Fig. 1. Overview of proposed Approach.

A. Truck Data Collection

To collect data on driving behavior, we utilized the Euro Truck Simulator 2 (ETS2), which accurately replicates truck driving conditions, including acceleration, braking, fuel consumption, and interactions with traffic. The simulator provides telemetry data such as speed, revolutions per minute (RPM), and acceleration. This rich dataset enables assessing the driver's efficiency in maintaining appropriate speeds and evaluating the impact on fuel consumption.

B. Telemetry Web Server

Extracting data from the simulator required the installation of the ETS2 Telemetry Web Server[2]. This open-source telemetry server is developed in C# and leverages WebSockets and REST APIs.

[2]https://github.com/Funbit/ets2-telemetry-server

The ETS2 Telemetry Web Server facilitates data collection from the simulator by decoding information and making it accessible via a REST API. This functionality ensures seamless integration between the simulator and applications requiring telemetry data, streamlining the data acquisition process.

C. Python Code to Request Data

We developed a custom application to consume the data generated and stored by the Telemetry Web Server. Python was chosen for this purpose due to its versatility and extensive library support for data manipulation, API integration, and statistical analysis.

The Python application sends requests to the telemetry server's REST API at one-second intervals, extracting real-time data. This information enables detailed analyses of driver behavior and facilitates the creation of graphical visualizations and reports, aiding in interpreting the collected data.

D. Computing Radar Area

Complementing the Python code, we calculated the radar area using speed, RPM, throttle position, and fuel consumption variables. This metric serves as an indicator to infer driver behavior under various driving conditions.

In the study by [7], sensors for speed, RPM, throttle position, and engine load were utilized to calculate the radar area. However, since ETS2 does not provide engine load data, we selected fuel consumption as the fourth axis.

This substitution is appropriate because fuel consumption is directly related to the driver's driving style; more aggressive driving results in higher consumption, thereby increasing the radar area. Consequently, fuel consumption becomes a valuable indicator for identifying behaviors such as harsh acceleration and improper use of pedals, directly reflecting the driver behavior metric.

We calculated the corresponding radar area after receiving telemetry data from the simulator via the REST API. Normalization was necessary to ensure that the variables are on comparable scales. For the RPM variable, we used the engine's maximum RPM value available in the simulator. For speed, we implemented a limiter of 90 km/h to prevent the truck from exceeding this maximum speed, aligning with common regulatory limits. The simulator provides the throttle position as a value between 0 and 1, requiring no additional normalization.

Finally, Equation 1 illustrates how the radar area is calculated:

$$Area = 0.5 \left| \sum_{i=1}^{n} x_i \cdot x_{i+1} \cdot \sin\left(\frac{2\pi}{n}\right) \right| \tag{1}$$

Given that the number of variables n is 4, when $i = 4$, the term $x_4 \cdot x_5$ is undefined since only four sensors are used. To address this, we employed the `roll` operation, allowing x_5 to be the same as x_1. This adjustment enables the calculation to proceed and effectively closes the polygon formed by the normalized variables.

E. Data Analysis

In the final data analysis stage, we accessed server-stored data to infer driving behaviors by extracting and interpreting telemetry information to identify patterns and trends. We conducted a comparative analysis of driver performance across scenarios using radar area values obtained every second, providing valuable insights for enhancing driving practices and optimizing vehicle performance.

IV. Case Study

The present case study aims to evaluate the proposed methodology, addressing the following research questions:(1) Can it identify driver behavior patterns in different driving styles? (2) How can the driving style inform practical recommendations that optimize fuel consumption?

A. Preparation

To ensure the consistency and validity of the collected data, several preparatory steps were established: **Truck Selection**: a standard truck within the simulator was chosen to maintain data consistency across all tests; **Route Definition**: a predetermined route was established, which would be used in both driving conditions to ensure comparability; **Driver Consistency**: the same driver was responsible for both driving scenarios to minimize the effects of human variability. Figure 2 shows the external view of the truck within the simulation environment, depicting the rear of the vehicle and the surrounding area.

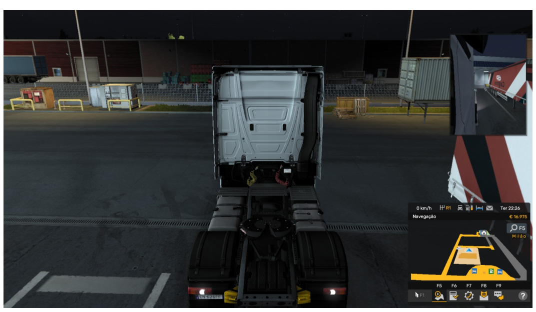

Fig. 2. External view of the truck within the ETS2.

B. Execution

For this study, the predefined route was traversed under two distinct driving conditions, allowing for a direct comparison between cautious and aggressive driving styles, both of which significantly impact fuel consumption and vehicle performance. **Cautious Driving - Scenario 1:** The driver focused on fuel economy, adhering to speed limits, avoiding abrupt accelerations, and employing smooth techniques to minimize fuel consumption and reduce vehicle wear. **Aggressive Driving - Scenario 2:** The driver repeated the route with an aggressive style, characterized by rapid accelerations and frequent speed limit violations, leading to increased fuel consumption and stress on the vehicle.

V. Results and Discussion

The results address the research questions by revealing distinct patterns in driver behavior and their implications for fuel consumption.

A. Identification of Driver Behavior Patterns

Figure 3 illustrates the radar area time series for both driving scenarios. In the cautious driving scenario (A), the radar area exhibits smooth variations, indicating a stable, controlled driving style prioritizing safety. In contrast, the aggressive driving scenario (B) shows sharp peaks and troughs, reflecting sudden maneuvers and less predictable driving patterns. This contrast confirms that the radar area metric can identify different driving styles.

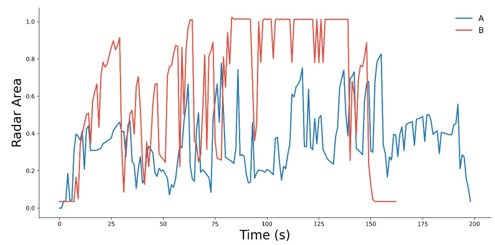

Fig. 3. Time series of radar area for cautious (A) and aggressive (B) driving scenarios.

Figure 4 presents the density distribution of the radar area. Scenario A is characterized by a higher density at lower radar area values, indicating that the driver spent most of the time in safer conditions. Conversely, scenario B shows a broader distribution with significant occurrences at higher radar area values, suggesting more time spent in riskier situations. These findings demonstrate the radar area's effectiveness in distinguishing between different driver behaviors.

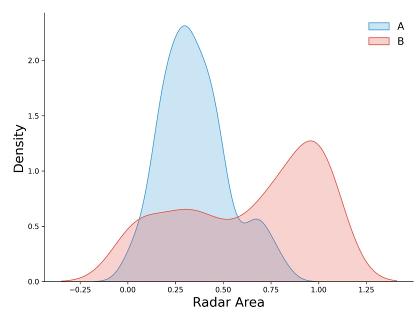

Fig. 4. Density distribution of radar area for both driving scenarios.

B. Recommendations for Optimizing Fuel Consumption

Figure 5 compares the average radar area between the two scenarios. The aggressive driving scenario (B) exhibits a significantly larger radar area, indicating a more intense driving style with frequent adjustments and reactive maneuvers. Although this increased activity is typically associated with higher fuel consumption due to frequent accelerations and decelerations, further analysis may be needed to confirm the exact impact on fuel efficiency, as the data does not consistently reflect this correlation.

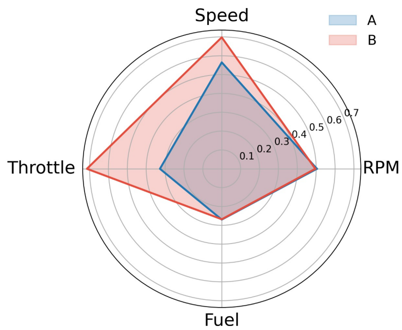

Fig. 5. Average radar area for cautious (A) and aggressive (B) driving scenarios.

Figure 6 shows the cumulative sum of the radar area over time. In scenario A, the cumulative radar area increases gradually, reflecting consistent and efficient driving practices that optimize fuel usage. In contrast, scenario B exhibits a steeper cumulative increase, indicative of driving behaviors that lead to higher fuel consumption and increased risk.

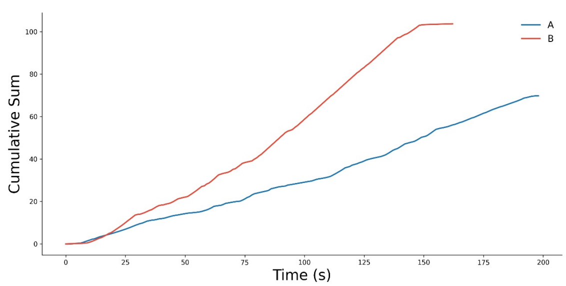

Fig. 6. Cumulative sum of radar area over time for both driving scenarios.

These results suggest that adopting a cautious driving style that maintains constant speeds and avoids abrupt maneuvers can optimize fuel consumption and enhance safety. The radar area metric effectively captures these differences, providing a valuable tool for driver behavior analysis.

VI. Conclusion

This study explored the use of digital twins in simulated environments, specifically utilizing Euro Truck Simulator 2 (ETS2) to infer driver behavior through virtual sensors. The application of digital twins demonstrated significant potential in providing real-time information on driving patterns without the need for physical vehicles. The radar area metric effectively distinguished between cautious and aggressive driving styles, though it revealed that ETS2's fuel consumption model might lack sensitivity to key engine demand parameters, suggesting a possible simplification in the simulation.

The case study results addressed the proposed research questions and offered insights into how different driving styles impact efficiency and safety. Although increased activity is typically linked to higher fuel consumption, the consistent fuel consumption values suggest that the impact of driving behavior on efficiency may need further investigation. Future work will focus on developing machine learning algorithms to analyze driving data and predict behaviors in real time and evaluating the proposed approach across various types of vehicles and routes to generalize the findings.

Acknowledgment

The research presented in this paper was financially supported by the Brazilian National Council for Scientific and Technological Development (CNPq) under Grant No. 405531/2022-2, the Research Development Foundation—FUNDEP/Mover-VI under Grant No. 29271.01.01/2023.03-00.

References

[1] O. Vermesan, R. John, P. Pype, G. Daalderop, K. Kriegel, G. Mitic, V. Lorentz, R. Bahr, H. Sand, S. Bockrath *et al.*, "Automotive intelligence embedded in electric connected autonomous and shared vehicles technology for sustainable green mobility. front. future transp. 2, 688482 (2021)," 2021.

[2] M. S. Dihan, A. I. Akash, Z. Tasneem, P. Das, S. K. Das, M. R. Islam, M. M. Islam, F. R. Badal, M. F. Ali, M. H. Ahmed *et al.*, "Digital twin: Data exploration, architecture, implementation and future," *Heliyon*, 2024.

[3] S. Deng, L. Ling, C. Zhang, C. Li, T. Zeng, K. Zhang, and G. Guo, "A systematic review on the current research of digital twin in automotive application," *Internet of Things and Cyber-Physical Systems*, vol. 3, pp. 180–191, 2023.

[4] T. Mecheva, R. Furnadzhiev, and N. Kakanakov, "Modeling driver behavior in road traffic simulation," *Sensors*, vol. 22, no. 24, p. 9801, 2022.

[5] F. Alonso, M. Faus, J. V. Riera, M. Fernandez-Marin, and S. A. Useche, "Effectiveness of driving simulators for drivers' training: a systematic review," *Applied Sciences*, vol. 13, no. 9, p. 5266, 2023.

[6] N. Fouladinejad, M. K. A. Jalil, and J. M. Taib, "Reduction of computational cost in driving simulation subsystems using approximation techniques," in *2014 International Conference on Industrial Automation, Information and Communications Technology*. IEEE, 2014, pp. 111–117.

[7] M. Medeiros, T. Flores, M. Silva, and I. Silva, "A multi-layered methodology for driver behavior analysis using tinyml and edge computing," in *2024 IEEE International Conference on Evolving and Adaptive Intelligent Systems (EAIS)*. IEEE, 2024, pp. 1–8.

[8] S. Tomažič, "Intelligent soft sensors," p. 6895, 2023.

[9] A. A. Al-Atawi, S. Alyahyan, M. N. Alatawi, T. Sadad, T. Manzoor, M. Farooq-i Azam, and Z. H. Khan, "Stress monitoring using machine learning, iot and wearable sensors," *Sensors*, vol. 23, no. 21, p. 8875, 2023.

[10] T. Flores, M. Silva, P. Andrade, J. Silva, I. Silva, E. Sisinni, P. Ferrari, and S. Rinaldi, "A tinyml soft-sensor for the internet of intelligent vehicles," in *2022 IEEE International Workshop on Metrology for Automotive (MetroAutomotive)*. IEEE, 2022, pp. 18–23.

[11] H. Shu, "Safety prompt advanced driver-assistance system," 2024.

[12] S. Stavrev and D. Ginchev, "Detection and analysis of commercial drivers' focus and attention using sensors and simulators," in *AIP Conference Proceedings*, vol. 3064, no. 1. AIP Publishing, 2024.

[13] M. Azevedo, M. Andrade, M. Medeiros, T. Medeiros, M. Silva, I. Silva, E. Sisinni, and P. Ferrari, "Optimizing vehicle iot systems: Sumo-digital twin performance analysis," in *2024 IEEE International Workshop on Metrology for Industry 4.0 & IoT (MetroInd4. 0 & IoT)*. IEEE, 2024, pp. 204–209.

[14] M. Andrade, M. Medeiros, T. Medeiros, M. Azevedo, M. Silva, D. G. Costa, and I. Silva, "On the use of biofuels for cleaner cities: Assessing vehicular pollution through digital twins and machine learning algorithms," *Sustainability*, vol. 16, no. 2, p. 708, 2024.