A Cloud-Edge Digital Twin Platform for Real-Time Driver Behaviour Analysis

Enrico Landolfi*, Giuseppe Loseto†, Tommaso Barrasso*, Giovanni Carpino†, Ferdinando Potenza†, Demetrio Alessandro Trunfio†, Filippo Gramegna†

* NetCom Engineering S.p.A. SB – Via Nuova Poggioreale, Centro Polifunzionale T7, I-80143, Napoli
{e.landolfi,to.barrasso}@netcomgroup.eu
† Department of Engineering, LUM "Giuseppe Degennaro" University – S.S. 100 km 18, I-70010, Casamassima, Bari
{loseto,gramegna}@lum.it, {g.carpino,f.potenza,d.trunfio}@studenti.lum.it,

***Abstract*—The paper presents a Digital Twin platform for in-vehicle monitoring, integrating OBD-II vehicle telemetry and custom driver-monitoring sensors. A cloud–edge approach is adopted to enable low-latency data processing using Machine Learning models for driving style classification and edge-based event detection. An early validation has been conducted to demonstrate the feasibility of the proposed approach and its suitability for real-time edge deployment.**

***Index Terms*—Digital Twin, Smart Mobility, Cloud-Edge Computing, Machine Learning**

I. Introduction and Motivation

The evolution of in-vehicle sensing devices has significantly increased the interest in advanced driver monitoring solutions aimed at improving road safety, reducing accident risks, and enabling intelligent mobility services [1]. Modern vehicles generate large volumes of real-time data through onboard Electronic Control Units (ECUs), providing valuable information about vehicle dynamics, driver interactions, and environmental conditions. Within this context, driving behaviour analysis has become a critical research topic. In particular, the identification of aggressive driving patterns or altered driver states caused by substances such as alcohol represents a major challenge for intelligent vehicular systems. According to recent road safety studies [2], impaired driving remains one of the leading causes of severe traffic accidents worldwide and the capability to continuously monitor driver behaviour is essential for the development of proactive safety mechanisms.

Traditional monitoring systems are often based on isolated data acquisition modules and centralized architectures. Vehicle telemetry is typically obtained through On-Board Diagnostics (OBD) interfaces [3], while additional driver-related information may be collected through external sensors, wearable devices, or environmental monitoring systems resulting in fragmented data processing pipelines and limited integration between vehicle-centric and driver-centric information. Furthermore, time-sensitive applications such as impaired driving detection require rapid local processing capabilities in order to identify potentially dangerous conditions and generate immediate alerts. For this reasons, Digital Twin (DT) technology represents a promising approach for creating dynamic virtual representations of physical systems and supporting the integration of vehicle telemetry, driver-related information, and contextual data into a unified intelligent ecosystem.

This research proposes the design and implementation of a cloud-edge DT platform for driver behaviour analysis and in-vehicle monitoring. The primary objective of the proposed system is the identification and classification of driving styles together with the detection of possible driver status alterations, including abnormal behaviours potentially associated with alcohol consumption or other impairment conditions. The proposed architecture combines cloud computing resources with edge-based processing capabilities in order to support real-time monitoring and low-latency analytics. The platform also integrates Machine Learning (ML) models deployed at the edge layer to enable real-time event detection and behavioural analysis directly within the vehicular environment.

The remainder of this paper is organized as follows. Section II presents the proposed DT framework architecture and its main technological components, Section III describes the early validation activities conducted using two reference datasets, whereas Section IV concludes the paper and discusses future research directions.

II. Framework Architecture

The reference platform, depicted in Figure 1, acquires vehicular telemetry data through an OBD-II module connected to the vehicle diagnostic interface, enabling access to parameters such as vehicle speed, engine revolutions per minute, throttle position, acceleration patterns, and steering behaviour. In addition, custom sensors dedicated to driver monitoring are integrated into the system in order to collect supplementary behavioural and physiological information capable of supporting the identification of altered driving conditions. The edge layer performs local data acquisition, preprocessing, filtering, and event classification within the vehicle environment. Meanwhile, the cloud layer provides advanced computational capabilities for long-term storage, training of the ML model for behavioural classification, and dashboarding. This approach enables rapid detection of critical behavioural anomalies and reduces dependency on continuous cloud communication.

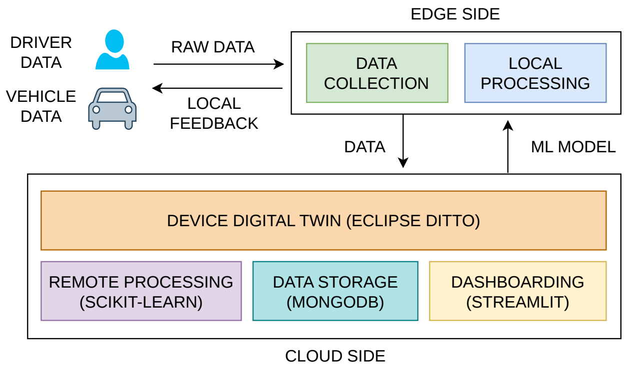

Fig. 1. Cloud-Edge Digital Twin Platform

At the core of the proposed architecture, the DT implementation has been developed using *Eclipse Ditto* (https://eclipse.dev/ditto/), an open-source framework specifically designed for the representation of physical entities through virtual models based on standardized JSON structures and communication interfaces. Each monitored vehicle is associated with a DT instance capable of continuously receiving telemetry data and driver-related information from the edge layer. The acquired data streams are locally preprocessed and transmitted to the DT infrastructure through an event-driven approach. For event detection, the platform integrates ML models developed using *scikit-learn* (https://scikit-learn.org/), a python-based library adopted for implementing classification algorithms. Data persistence and historical storage functionalities are implemented through *MongoDB* (https://www.mongodb.com/), a NoSQL document-oriented database particularly suitable for managing heterogeneous and semi-structured vehicular data. The choice of MongoDB is motivated by its document-store architecture, which allows efficient storage and retrieval of JavaScript Object Notation (JSON)-based telemetry records and DT state information. Since Ditto internally manages device representations using JSON structures, MongoDB provides a natural solution for storing synchronized vehicle and driver-related datasets while maintaining high flexibility for schema evolution and future platform extensions. In addition, to support remote monitoring, the platform integrates a visualization layer developed using *streamlit* (https://streamlit.io/). The web dashboard provides operators and researchers with a comprehensive overview of driving sessions, detected anomalies, and system status, facilitating both real-time supervision and offline analysis activities.

III. Platform Validation

An initial validation of the proposed platform has been conducted using two publicly available datasets related to alcohol detection [4] and driving behaviour analysis [5]. The objective of this preliminary evaluation was to assess the feasibility of the proposed architecture in terms of heterogeneous data integration and ML-based event detection.

The first dataset focuses on driver status monitoring and alcohol detection through the acquisition of multimodal physiological, biological, and visual information. In particular, the dataset includes measurements collected from three different sensing modalities: a sensor for detecting alcohol concentration in the surrounding environment, a thermal sensor capable of measuring temperature variations at predefined regions of the driver's face, and a visual monitoring component designed to identify and estimate pupil thickness variations. The second dataset contains vehicle telemetry data including several vehicular parameters related to speed, acceleration, engine dynamics, and driving manoeuvres. In addition, the dataset provides labeled behavioural classes associated with different driving styles, specifically distinguishing between moderate and aggressive driving behaviours.

The validation has been conducted by evaluating and comparing the performance of several ML classification models, including Logistic Regression, Decision Tree, Random Forest, Support Vector Machine, and K-Nearest Neighbors. The experimental results show that the Random Forest model consistently outperforms the other approaches, achieving an overall accuracy greater than 97% across both reference datasets. The model also demonstrates inference times of only a few milliseconds, making it particularly suitable for edge-side event detection and to support the generation of immediate alerts and feedback.

IV. Conclusion

This paper presented a DT platform for in-vehicle monitoring, integrating OBD-II telemetry, custom driver-monitoring sensors, and classification models. Future work will focus on extending the platform towards extensive real-world evaluations involving additional on-board sensing devices, in order to further improve robustness and generalization under diverse driving conditions. Moreover, the architecture will be enhanced to incorporate federated learning techniques, enabling collaborative model training across multiple vehicles while preserving data privacy and reducing communication overhead.

References

[1] E. T. Michailidis, A. Panagiotopoulou, and A. Papadakis, "A Review of OBD-II-Based Machine Learning Applications for Sustainable, Efficient, Secure, and Safe Vehicle Driving," *Sensors*, vol. 25, no. 13, p. 4057, 2025.

[2] E. Beaulieu, R. B. Naumann, G. Deveaux, L. Wang, E. J. Stringfellow, K. H. Lich, and M. S. Jalali, "Impacts of alcohol and opioid polysubstance use on road safety: Systematic review," *Accident Analysis & Prevention*, vol. 173, p. 106713, 2022.

[3] M. Ruta, F. Scioscia, G. Loseto, A. Pinto, and E. Di Sciascio, "Machine learning in the Internet of Things: A semantic-enhanced approach," *Semantic Web*, vol. 10, no. 1, pp. 183–204, 2018.

[4] P. D. Rosero-Montalvo and D. H. Peluffo-Ordóñez, "Alcohol Detection in Drivers by Sensors and Computer Vision," 2025, accessed: 2026-05-12. [Online]. Available: https://ieee-dataport.org/open-access/alcohol-detection-drivers-sensors-and-computer-vision

[5] I. G. Zamora, "OBD-II & CAN-Based Driving Behavior Dataset," 2024, accessed: 2026-05-12. [Online]. Available: https://www.kaggle.com/datasets/isaygerardozamora/obd-ii-and-can-based-driving-behavior-dataset