Date of publication xxxx 00, 0000, date of current version xxxx 00, 0000.

Digital Object Identifier 10.1109/ACCESS.2024.0429000

Anomaly Detection for Vehicle Diagnostics based on OBD Snapshots with Cause Investigation

Veljko Vučinić[1], Luca Seidel[1], Nikola Lukežić[1], Frank Hantschel[2], Thomas Kotschenreuther[2], Dragan Aleksendrić[3], and Eric Sax[1]

[1]Institute for Information Processing Technologies, Karlsruhe Institute of Technology, 76131 Karlsruhe, Germany
[2]RA Consulting GmbH, 76646 Bruchsal, Germany
[3]University of Belgrade Faculty of Mechanical Engineering, Automotive department, 11120 Belgrade, Serbia

Corresponding author: Veljko Vučinić (e-mail: veljko.vucinic@kit.edu).

ABSTRACT Vehicle diagnostic systems are crucial for the normal operation of vehicles and their propulsion-related systems. Undetected unusual behaviour of such systems makes the vehicle diagnostic system unreliable. Current diagnostic systems, such as On-Board Diagnostics (OBD), are limited to monitoring only specific systems in order to make fault decision. However, various anomalies, including drastic performance drops, vehicle tampering, and changes in the driving environment, often go undetected during OBD system testing, validation, and inspection. This research presents a novel explainable OBD anomaly detection pipeline that is able to detect anomalies based only on OBD data snapshots during processes of OBD validation and inspection. The novel approach is implemented using combined dimension reduction and data clustering methodologies. First, the data is transformed into a latent space using t-distributed Stochastic Neighbor Embedding (t-SNE), where the general structure of the anomaly in the data can be exploited. Subsequently, clustering using Density-Based Spatial Clustering of Applications with Noise (DBSCAN) is applied to group similar normal data and identify anomalous patterns. The novelty of the solution is further extended with a feedback loop that suggests the root cause of OBD signals for individual anomalies using explainable AI (XAI) methodology, in this case Shapley additive explanations (SHAP). The proposed concept was verified and evaluated using real OBD snapshots with synthetically generated anomalies in two scenarios with different engine status, engine off and on, with an achieved accuracy of 92.89% and 96.45% for anomaly detection, respectively. The majority of anomaly causes in the form of specific OBD signals from propulsion- and emission-related systems were successfully explained using SHAP.

INDEX TERMS anomaly detection, dimension reduction, clustering, explainable AI, OBD, validation, vehicle diagnostics

LIST OF ABBREVIATIONS

CTC	Clean Truck Check
DBSCAN	Density-Based Spatial Clustering of Applications with Noise
ECM	Engine Control Module
ECU	Electronic Control Unit
DTC	Diagnostic Trouble Code
FN	False Negative
FP	False Positive
ICE	Internal Combustion Engine
IF	Isolation Forest
KL	Kullback–Leibler
LOF	Local Outlier Factor
OBD	On-Board Diagnostic
PCA	Principal Component Analysis
PID	Parameter Identification
PVE	Production Vehicle Evaluation
RF	Random Forest
SHAP	Shapley Additive Explanations
t-SNE	t-distributed Stochastic Neighbor Embedding
TN	True Negative
TP	True Positive
TPR	True Positive Rate
UMAP	Uniform Manifold Approximation and Projection
XAI	Explainable Artificial Intelligence

I. INTRODUCTION

THE purpose of vehicle diagnostic systems is to monitor the vehicle emissions and the operation of propulsion-related systems. The On-Board Diagnostic (OBD) system has the intended function of making a fault decision, including fault detection, isolation, and identification [1]. In this context, the diagnostic system must answer the question of when a fault occurred, while also pinpointing the faulty components with the fault effect magnitude. All this makes vehicle diagnostics essential for the reliability of vehicles and their systems, especially considering the importance of maintaining the normal operation of critical systems, like active safety [2]. However, modern vehicle diagnostic applications, such as OBD, have various limitations due to the conventional approach of diagnostics [3]. Firstly, they monitor only specific propulsion- and emission-related systems, and not all critical systems on the vehicle. Secondly, they can make a fault decisions only for a finite number of faults, excluding all malfunction possibilities [4]. Lastly, they do not cover the detection of vehicles' and diagnostic systems' unusual behaviours, such as tampering, drastic system performance drops, changes in driving environment, etc. These limitations often make various anomalies undetected by classical OBD systems, which can lead to critical consequences [5]. More general anomaly detection has proven to be a valuable approach for ensuring the safety and security [6]. In addition to these limitations, the design space of modern vehicle platforms contributes to an increasingly high-dimensional build space due to a multitude of software and hardware configurations across vehicle variants [7]. This leads to significant challenges in validating diagnostic behavior across all possible variants. With this in mind, vehicle diagnostic validation engineers, testers, and inspectors cannot fully rely on the output of current OBD systems when checking for problems and status.

This research presents an innovative pipeline of anomaly detection with root detection based on OBD snapshot embedded into the processes of OBD testing, validation, and inspection. The pipeline contains anomaly detection based on OBD signals, able to detect anomalies based only on snapshots of OBD data. A snapshot is a single measurement or observation of a system's state at a specific moment. The proposed solution does not require labeled data, but can be applied to various datasets with an unknown number and types of anomalies. This makes it applicable for investigating specific in-vehicle system problems, but also for general use, such as vehicle systems or data tampering detection. The proposed anomaly detection employs a two-stage OBD snapshot processing pipeline combining dimension reduction and data clustering techniques. First, the high-dimensional OBD data is transformed into a latent space where the general structure of the anomaly in the data can be exploited. One general example of dimension reduction used on real OBD vehicle snapshots data is represented in Figure 1. Subsequently, dimension reduction keeps the grouped structural patterns of normal and unexpected data in low latent spaces, and makes them visualizable and deducible using clustering techniques. The methodology is extended with the root cause suggestion of OBD signals for individual anomalies using the explainable AI approach, mainly Shapley Additive Explanations (SHAP) technique [8]. The overall methodology benefits validation and workshop engineers as an innovation tool during system validation after production and in the aftermarket vehicle use case.

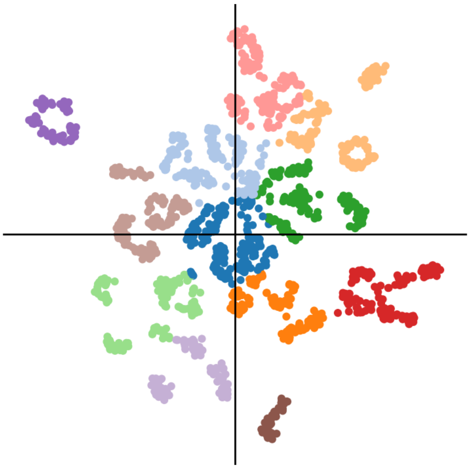


FIGURE 1: Representation of OBD snapshot data after dimension reduction and general clustering.

In the context of vehicle diagnostic validation, novel data processing applications have the potential to enhance the reliability and durability of vehicle systems in testing phases [9], by marking potential weak points in design and control strategies. For example, the SAE J1699-3 is a diagnostic compliance test procedure done as part of Production Vehicle Evaluation (PVE), and is used to validate the operation and communication of diagnostic systems and main propulsion-related controllers [10]. The methodology proposed in this paper for anomaly detection could extend the standardized SAE J1699-3 test by introducing snapshot-based anomaly detection with root detection. This holds a potential to identify anomalous behaviour and its roots before a vehicle enters the market, significantly improving diagnostic and propulsion-related systems' reliability and overall vehicle safety.

Beyond pre-market validation, similar anomaly detection techniques also play an important role for in-use compliance and emissions monitoring. Tampering with the OBD systems of diesel trucks is becoming a critical issue worldwide, as it directly leads to a significant increase in emissions [11], [12]. Another practical application of OBD snapshot-based anomaly detection could be to aid California's Clean Truck Check (CTC). The initiative is motivated by minimizing the pollution effect from heavy-duty vehicles, and deals with active monitoring of the emission-related systems (covered by the OBD system) for on-road trucks [13]. As part of the CTC initiative, truck drivers and fleet managers are obligated to provide periodical OBD snapshots to ensure their trucks comply with the emission regulations. The solution proposed in this paper can help catch anomalies such as system, scan tool, or data tampering on a large scale and further ensure the emission minimization effect in the area.

The paper is organized in a way that section II discusses the research gap, state of the art, and principal paper contributions. Section III contains the background about the vehicle and OBD anomalies with their generation for the later concept verification and evaluation. The following section IV proposes a novel concept of anomaly detection with cause investigation using OBD snapshots. Experiments with evaluation scenarios are described in the section V, where results are visualized and later discussed in section VI. The conclusions and future work are stated in section VII.

II. RELATED WORKS

The development of OBD was primarily motivated by the need to reduce carbon emissions and to standardize diagnostics across all vehicle manufacturers. The functionality of OBD for diagnostics is straightforward and rule-based. For example, if a signal exceeds a predefined threshold, a fault is detected and a Diagnostic Trouble Code (DTC) is raised. With this in mind, rule-based diagnostics fail to capture the complexity of various real-world vehicle anomaly patterns. This makes them incapable of handling variations for different vehicle types and environmental situations [14]. Together with that, existing OBD approaches cover only a predefined number of finite faults for which the OBD monitors the thresholds. This set of detectable DTCs with regular OBD is defined with standard SAE J2012, which is updated periodically [15], [16]. Even if OBD detected a problem, multiple DTCs are often triggered, making it challenging to pinpoint the exact root cause. There are no indicators which signals were out of order, and only a handful of signal values are recorded when a fault is detected.

Anomaly detection has been widely explored in many fields to catch problems and has been deployed in, but not limited to healthcare [6], finance [17], and cybersecurity [18]. In the automotive sector, anomaly detection methods focus on emission control, driving behavior, energy consumption, and cybersecurity. Table 1 summarizes a comparison of anomaly detection methods for fault detection in vehicles. Data is mainly based on CAN data, while lesser research focuses on OBD-specific data [3]. Anomaly detection methods focused on emission control and prediction can also identify malfunctions and manipulations. Although several successful approaches have been implemented and tested, most methods operate online and rely on historical driving data [19]–[21]. Anomaly detection methods have also been widely explored in the field of vehicle cybersecurity, where standardized OBD data is used as a base for intrusion detection [22]. However, such approaches often rely on time series data and cannot be applied to snapshot data [23]–[25]. The majority of them use supervised machine learning methods that are unsuitable for a wide range of vehicle types and unknown problems outside of the training dataset [24]–[26].

The authors of [27] propose an anomaly detection approach in vehicles using Principal Component Analysis (PCA) for dimensionality reduction. Anomalies are identified using One Class Support Vector Machines (SVM), Isolation Forest, and Local Outlier Factor, while the best results were achieved by One Class SVM, reaching an accuracy as high as 96.12%. The approach processes the timeseries data and their trends, limiting its deployment to snapshot data.

The authors of [24] developed a method for anomaly detection on CAN sensor timeseries data. The method compares Long Short Term Memory-, Gated recurrent unit-, and Convolutional neural networks-based anomaly detectors. The best results were achieved with Long Short Term Memory-anomaly detectors with a successful detection of 86%. The system has its limitations as it requires labeled timeseries data for model training, meaning it cannot cope with snapshot data, nor anomalies outside of the training data range. The authors of [28] introduced a KL divergence-based method for detecting lateral deviation anomalies in autonomous mining vehicles by segmenting road intervals and applying threshold determination. While this method could be adopted for snapshot data, it would rely on historical data.

Recent advancements showed that related works integrated machine learning with Explainable Artificial Intelligence (XAI) in vehicle diagnostics [29]–[32]. It enables the transition from simple problem detection to comprehensive root cause analysis. For instance, authors of [29] proposed a framework for heavy-duty diesel vehicles that identifies high-emission trucks and utilizes XGBoost alongside model interpretation methods like Partial Dependence Plots to trace the underlying causes of specific emissions. Similarly, the work [30] developed a model-based diagnostic framework using Random Forest and SHAP to pinpoint the specific OBD parameters responsible for low in-use monitor performance ratio output. Furthermore, authors of [31] established a hybrid CNN-LSTM-Transformer architecture that integrates XAI with Local Interpretable Model-agnostic Explanations (LIME) and SHAP to provide interpretability for maintenance technicians. Another application of XAI used SHAP and Individual Conditional Expectation to analyze diesel particulate filter regeneration states in urban bus fleets to optimize maintenance management [32]. Despite these OBD root detection innovations with XAI, existing research predominantly relies on continuous time-series data and large historical datasets.

Various vehicle and diagnostic anomaly detection have been implemented and studied so far. While the majority of related works focus on known problems (supervised learning) and timeseries data, the research gap remains for general explainable anomaly detection based on only OBD snapshots. This is crucial for the use cases of OBD testing and inspection. Furthermore, the related works showed no published research that aimed to optimize the OBD testing, validation, or inspection concerning explainable anomaly detection. Primary contributions of the current paper can be summarized as follows:

- Novel pipeline for snapshot-based OBD anomaly detection during processes of OBD validation and inspection

TABLE 1: Comparison of existing anomaly detection methods for fault detection in vehicles.

Approach	Method	Dataset	Performance	Limitations
Cherdo et al. (2023) [24]	Long Short Term Memory	CAN data from Alpine Renault car	86% accuracy	Timeseries data and non standardized signals
Aloqaily et al. (2025) [25]	Random Forest and LightGBM	DoS, Fuzzy, Gear and RPM OBD intrusion datasets	99.9% accuracy	Known anomalies only (supervised learning)
Van Wyk et al. (2024) [26]	Convolutional Neural Network and Kalman filtering	N.A	85.4-95.3% accuracy	Limited to known DTC anomalies
Jain et al. (2023) [27]	Principal Component Analysis and One-Class Support Vector Machine	OBD-II Data from electric off-road vehicles	96.12% accuracy	Classification of anomalies based on DTC
Zhang et al. (2024) [28]	Kullback-Leibler divergence	Autonomous vehicle data in mining areas	91.4% accuracy	Detection based on vehicles driving on the same road intervals

using a combination of dimension reduction and clustering methodology embedded in the mentioned processes.

- OBD system anomaly detection based on discrete data snapshots collection, avoiding the need for continuous timeseries data.
- Cause detection for anomalies in vehicles and OBD systems using Explainable AI methodology based on data snapshots.

III. BACKGROUND

A. VEHICLE AND OBD SYSTEMS ANOMALIES

The normal operational behavior of a vehicle system can be defined as the expected range of system states and signal patterns that occur under standard usage conditions, defined by manufacturer specifications. It is typically marked by consistent statistical properties and repeatable parameter relationships. Anomalies in vehicle systems represent deviations from the normal operational behavior of those systems that can originate from various sources, such as physical component malfunctions, software failures, or cyberattacks targeting the vehicle's communication networks [33]. In the sense of vehicle diagnostics, all faults are considered anomalies, but not all anomalies are faults. The fault is defined by the hypothesized cause of the system failure, where failure represents an event that occurs when the delivered system service deviates from the service implementing the system function [34]. Vehicle systems and diagnostic anomalies manifest as unexpected events or data patterns that deviate from established baselines and usually can be detected through the in-vehicle network or by specific processing of the sensors, actuators, and Electronic Control Unit (ECU) data. Anomaly detection can be represented as a task of detecting these anomalous pieces of data that exhibit different patterns from normal data [35]. There are various categories of anomalies in sensor data [36], and such anomalies can be described as one or a successive series of signal vectors that deviate from the current data and therefore represent a singularity [37].

In the course of this paper, the anomalies will be considered as all unexpected measured system behaviours that do not cause direct system failure (i.e., system performance drop), and the system failure causes (faults) that are out of the scope of the OBD II system (i.e,. vehicle system, data, external scan tool tampering). The focus of this work are anomalies of hardware and software components of passenger and commercial vehicle propulsion- and emission-related systems, that are detectable within vehicle OBD II data. Propulsion- and emission-related systems are a group of original equipment systems, components, and parts whose failure will directly impact the ability to propel the vehicle or raise direct vehicle emissions above the legal limit. Such systems are involved in refueling/recharging the vehicle, storing and transferring fuel/energy, the combustion process, and propelling or decelerating the vehicle, together with systems in charge of delivering torque to the wheel. Included in that group are components used to control or thermally manage such systems and their emissions. The novel concept will be verified using Internal Combustion Engine (ICE) passenger vehicles data that have high intersimilarity due to the standardization of OBD systems among them. Anomalies defined as such are undetectable using state of the art vehicle diagnostic system, therefore will be the contribution of this paper.

Detecting vehicle systems and OBD anomalies is pivotal before the vehicle enters the market, but also in the vehicle exploitation. In the phases of vehicle testing and validation during production, OEMs can still actively fix the root problems, ensuring higher vehicle system reliability. This is the purpose for which PVE is introduced for the OBD system before entering the market. PVE is a requirement introduced as a worldwide standard to ensure the functionality of OBD in production vehicles. It comprises three stages: J1, J2, and J3 tests, where all stages are used in the USA and China markets, while only J1 is used for vehicles produced in Europe. The J1 test verifies the conformity of OBD communication in accordance with SAE J1699-3. It ensures that the vehicle's OBD interface supports the required protocols (e.g., CAN, UDS, ISO9141, ISO15765) and transmits diagnostic data in accordance with SAE J1979. In the J2 test, all DTCs that trigger the malfunction indicator lamp (MIL) are checked. Faults must be simulated using hardware-based methods and checked over several driving cycles. The test confirms the generation of pending, confirmed, and permanent fault codes and the activation of the MIL. The J3 test evaluates the in-use performance of OBD monitoring by collecting in-use performance ratio data from vehicles in the field over a period of 6 to 12 months. To ensure representative results, the vehicle sample must reflect normal use and be selected statistically from national data sets [38].

Due to the conventional way of fault detection, not all anomalies are being detected with OBD. One example of this represents tampering with the vehicle, data, network, OBD system, external tools, etc. This is being done for engine tuning, ECU enhancements, upgrading infotainment systems, or simply misrepresentation of the real vehicle systems and subsystems status, such as emissions. A popular case of tampering is the Dieselgate scandal, including one of the biggest worldwide OEMs [39]. Other examples of anomalies can be specific vehicle systems problems that are outside the monitoring and analysis of OBD, such as drastic vehicle performance drops that affect fuel consumption and emissions, or changes in the driving environment. Such anomalies do not raise any particular faults, but affect the vehicle's compliance, driving, or overall safety.

The current state of the vehicle E/E architecture supports a wide variety of in-vehicle signals from sensors, actuators, and controllers. A subset of those signals is made available through the OBD system. A standardized OBD system has 10 different diagnostic services 01-0A (hexadecimal) with various purposes, where services 01 and 09 give current powertrain data and in-use performance metrics and parameters, respectively [40]. According to the standard SAE J1979 digital annex, OBD services 01 and 09 support more than 250 signals [41], whereas in reality, a modern vehicle supports around 50 [9]. Support of the OBD signals from mode 1, Parameter Identification (PID)s, depends upon the vehicle model, fuel type, model year, manufacturer, etc. Quantity, quality, and representability of such signals are a solid ground base for anomaly detection. Data available from OBD mode 01 will be used in the course of this work for designing and implementing a novel anomaly detection method.

B. GENERATION OF SPECIFIC OBD ANOMALIES

In order to evaluate and verify the novel anomaly detection solution proposed in the following section, specific cases of anomalies are generated. The anomaly generation was done using real OBD data snapshots that represent normal cases. Specific or random parameters were adjusted to generate different use cases of anomalies with OBD data, depending on the anomaly types replicated. As it is customary in the literature [35], [36], [42]–[44], we use this strategy to test the performance boundaries of the anomaly detection method proposed. Anomaly detection will be covered in four specific cases of OBD and vehicle system anomalies in this work:

1) Vehicle system operating performance drop
2) Engine coolant system problem
3) Fuel system problem
4) Engine ECU tampering

The first type of anomaly addressed within this work is the vehicle system operating performance drop. Since performance drops with the vehicle systems can be very generic, it cannot be expected that a specific set of parameters is problematic. It represents the anomaly where data patterns and correlations are out of scope from the normal data, but for a random set of OBD PID signals. One example is the battery energy storage capacity drop, impacting the ability to charge the battery to its original manufacturer-specified capacity [45]. For this purpose, snapshots representing this anomaly type are generated on random OBD PIDs x_i^{obd} using the Lorenz Attraction model, also known as Chaos theory [46], [47]. This anomaly generation technique utilizes a set of differential equations (1a)-(1c), where small input differences allow recreation of general anomalies, fitting the use case of vehicle system performance drop. In the course of this work, the equations of the Lorenz model (1a)-(1c) are solved for parameters $\sigma = 10$, $\beta = 2.65$, $\rho = 28$, and $\mathrm{d}t = 0.01$. The solutions of the differential equations (α, θ, and γ) are normalized (α_{norm}, θ_{norm}, and γ_{norm}), and later injected in the random PIDs $x_{i,orig}^{\mathrm{obd}}$ of snapshots according to Equation (2), emulating performance drop anomaly PIDs $x_{i,anom}^{\mathrm{obd}}$.

$$\frac{\mathrm{d}\alpha}{\mathrm{d}t} = \sigma(\theta - \alpha) \tag{1a}$$

$$\frac{\mathrm{d}\theta}{\mathrm{d}t} = \alpha(\rho - \gamma) - \theta \tag{1b}$$

$$\frac{\mathrm{d}\gamma}{\mathrm{d}t} = \alpha\theta - \beta\gamma \tag{1c}$$

$$x_{i,anom}^{\mathrm{obd}} = x_{i,orig}^{\mathrm{obd}} * (1 + (\alpha_{norm} + \theta_{norm} + \gamma_{norm})) \tag{2}$$

Other types of anomalies are generated using a Gaussian perturbation to the selected PIDs for each case, depending on the anomaly type. The Gaussian perturbation is done by multiplying the Gaussian noise with the original signals from normal data, representing statistical deviation or anomaly in a data sense. The Gaussian noise magnitude is scaled differently in order to achieve a range of anomaly severity levels, giving more realistic cases of variable severity of anomalies found in real OBD cases. In contrast to the chaotic perturbations described earlier, this approach aims to generate more localized outliers and sensor level anomalies by perturbing selected signals toward the statistical borders of their normal distribution. For a given snapshot, a subset of OBD PID signals x_i^{obd} is selected, and their original values $x_{i,orig}^{obd}$ are modified using a multiplicative Gaussian deviation.

The selection of OBD PID signals for each anomaly type created using Gaussian perturbation is shown in Table 2. Since the generation of each instance of vehicle operation performance drop anomaly (using the Lorenz attraction model) is random, it is excluded from the mentioned table. In the case of an engine coolant system anomaly, the outputs of engine coolant temperature sensors are tuned. One example of such an anomaly can be the air trapped in the coolant system, causing air pockets due to improper bleeding after coolant refill. The impact of such an anomaly is reduced heat transfer efficiency or localized temperature irregularities. This anomaly will not cause a system failure and is not recognized with the OBD system, but the long-term damage risk increases.

Furthermore, fuel system problems are generated using fuel-related parameters, such as various temperature and pressure sensor parameters. An example of such an anomaly would be delayed or noisy fuel pump priming, causing the longer time to build the required pressure in the vehicle engine off, ignition on state. It can lead to the extended cranking time and inconsistent cold starts, again not being detected within the OBD system. Lastly, anomalies related to engine ECU tampering are created using dynamic propulsion system sensors and actuators, such as engine speed, torque, and throttle positions. For example, the engine parameters could be forcefully remapped outside of official workshops for engine tuning. This anomaly can lead to a failure of the propulsion system, since it starts behaving outside of predetermined manufacturer specifications, but is not detected within OBD. In this way, only the specific physically correlated parameters are manipulated, generating relevant real-life problem scenarios where exceptional, out-of-order patterns are exhibited. This is important to test the feedback part of the proposed anomaly detection solution, used for identifying the primary causes for various anomalies.

The idea of using Gaussian perturbation is to replicate anomalies that fall outside the usual operating range but remain physically plausible, such as various inconsistencies found in real OBD data. To emulate different severity levels, an anomaly is constructed by sampling a Gaussian random variable centered at zero and scaled by a deviation factor $d_a \in \{0.5, 1.5, 3\}$, depending on the anomaly magnitude. Anomaly values are generated using the equation (3):

$$x_{i,\text{anom}}^{\text{obd}} = x_{i,\text{orig}}^{\text{obd}} \cdot [1 + \mathcal{N}(0,\ d_a)] \tag{3}$$

where $\mathcal{N}(0, d_a)$ represents a Gaussian distribution with zero mean and standard deviation d_a. After perturbation, the resulting values are bounded within predefined physical limits to prevent impossible or unrealistic OBD signals (e.g., negative RPM or temperatures below hardware thresholds). This Gaussian-based injection method ensures the controllable and reproducible creation of severity-graded anomalies across multiple PID inputs.

The similarity between the generated synthetic anomalies and the real anomalies found in actual J1699 log data was checked using Kullback–Leibler (KL) divergence. This approach is a common measure for quantifying the difference between two probability distributions [48]. In this context, a lower KL divergence, measured in nats for multivariate distributions, indicates a higher statistical resemblance between the datasets. It should be taken into consideration that the anomaly types found in the actual J1699 log files are much more diverse than those generated in this work. The KL divergence analysis revealed that the generated anomalies emulating system performance drops, created using a Lorenz attraction model, showed the closest alignment with the real data, with a total KL divergence of 18.77 nats. This low divergence suggests that the chaotic perturbations effectively capture the structure of a general anomaly type, such as system performance drops.

On the other hand, anomalies generated using Gaussian perturbation, designed to reflect component-specific issues such as engine coolant faults and fuel system problems, exhibited higher KL divergence values of 25.22 and 33.70 nats, respectively. While still being reasonably aligned with real-world trends, these results indicate a growing deviation in statistical behavior. The anomalies intended to represent ECU tampering produced the highest KL divergence at 47.55 nats, showing a significant shift from the distribution of tampering-like behavior in the J1699 logs. This is expected, since the original anomalies do not specifically contain this type of anomaly. This concludes that the generated anomalies provide a meaningful approximation of real vehicle systems and OBD anomalies found in actual J1699 log files.

IV. METHODOLOGY

A primary challenge in implementing heuristic and expert knowledge-based systems for anomaly detection in OBD data lies in the high dimensionality of OBD datasets. Conventional statistical analyses, such as correlation studies [49], as well as data visualization, face significant limitations when applied to such high-dimensional spaces, especially for holistic data interpretation. Removing some OBD PIDs temporarily solves this problem, but removes potentially relevant information from the dataset. The generalization of the data processing results in that case is lost. The challenge grows when considering only OBD snapshots as data input, since less amount of data per vehicle is present. As a first step of the methodology, a variety of OBD data snapshots from different vehicles need to be acquired through testing, measurements, validation, and other means through the OBD port inside the vehicles. This will create a database of OBD data snapshots ready in the backend for further processing, visualization, and archiving purposes. The database is managed by the engineers who are in charge of the process of testing, validation, monitoring, etc. A database that contains a collection of OBD Mode 01 snapshots can be defined with the relation (4), where X_{obd} is $m \times n$ OBD snapshot matrix, n represents the number of available PIDs, and m indicates the total number of snapshots in the database. Each OBD PID x_i^{obd} represents a column in X_{obd}. Practical dimensionality revolves around $n \approx 50$ (up to 250).

$$\begin{gathered} X_{\text{obd}} \in \mathbb{R}^{m \times n}, \quad x_i^{\text{obd}} \in \mathbb{R}^{m}, \quad x_i^{\text{obd}} = X_{\text{obd}}[:, i], \\ \text{for } 1 \leq i \leq n \end{gathered} \tag{4}$$

Data analyses and processing tend to use lower-dimensional data for better and more transparent results [50]. In order to achieve lower dimensions and keep the data structure consistent, the first step of the anomaly detection for the use case of OBD snapshots proposed in this paper is dimension reduction. Reduced data loses the physical meaning of dimensions, but gains potential to analyse and process data in lower dimensions, revealing the topological structure of the

TABLE 2: List of OBD parameters that are used as a basis for the generation of specific anomaly types using Gaussian perturbation.

Anomaly root	PID	PID abbreviation	PID description
Engine coolant system	05	ECT	Engine coolant temperature
	67	ECT_1	Engine coolant temperature 1
	67	ECT_2	Engine coolant temperature 2
Fuel system	0B	MAP	Manifold absolute pressure
	0F	IAT	Intake air temperature
	10	MAF	Airflow rate
	23	FRP	Fuel rail pressure
	68	IAT_11	Intake air temperature 1
	68	IAT_12	Intake air temperature 2
Engine ECU tampering	04	LOAD_PCT	Engine torque percentage
	0C	RPM	Engine RPM
	11	TP	Absolute throttle position
	43	LOAD_ABS	Engine torque value
	45	TP_R	Relative throttle position
	47	TP_B	Absolute throttle position B
	5C	EOT	Engine Oil Temperature

data inside a dataset. After dimension reduction is applied to the initial data (4), the resulting relation (5) defines Y_{obd} as the OBD snapshot data matrix with k reduced dimensions, where each column y_j corresponds to a reduced feature. Dimension reduction mapping function ϕ maps the n-dimensional data points to k-dimensional target points, see relation (6) [51]. Using lower dimensions, such as $k = 2$ or $k = 3$, better visualization and clustering of data are possible, making the results and the data structure intuitive for systematic anomaly detection and further heuristic reasoning of the data.

$$Y_{\text{obd}} = \phi(X_{\text{obd}}), \quad Y_{\text{obd}} \in \mathbb{R}^{m \times k}, \quad y_j = Y_{\text{obd}}[:, j], \\ \text{for } 1 \leq j \leq k, \quad k < n \tag{5}$$

$$\phi : \mathbb{R}^n \rightarrow \mathbb{R}^k, \quad x_i \rightarrow y_j, \text{ for } 1 \leq j \leq k \tag{6}$$

The central premise of this work is formalized in Hypothesis 1. The authors propose that dimension reduction techniques can reveal the disputancy in data structure between normal and abnormal (anomaly) patterns required for anomaly detection. The hypothesis revolves around the notion that by mapping OBD data into a lower-dimensional space, it becomes feasible to identify abnormal patterns through clustering and outlier analysis. The proof of such a hypothesis would enable effective anomaly detection in automotive applications, specifically crucial for diagnostic systems.

Hypothesis 1. *Anomalies in vehicle operation and diagnostic systems are detectable within lower-dimensional representations of OBD data snapshots, perceptible by their individual distances and inherent structural patterns.*

The architecture of anomaly detection based on OBD data snapshots proposed in this work is visualized in Figure 2. The figure shows the collection of OBD snapshots in the database, forming $m \times n$-dimensional matrix X_{obd}, described by the Equation (4). The input is preprocessed using min-max data normalization to ensure that all features contribute uniformly [52], and later processed by the dimension reduction algorithm. The output of dimension reduction is an OBD snapshot dataset with reduced dimensions Y_{obd}, according to Equation (5). Such data is further processed with a clustering technique to identify groups of normal data points and potential outliers outside of those cluster zones. Each anomaly is individually inspected in the backend to determine the cause of its labeling as a potential anomaly. This inspection results in a selected group of OBD PIDs (i.e. $x_u^{\text{obd}}, x_v^{\text{obd}}, x_m^{\text{obd}}$) that exhibit the strongest influence on the anomaly compared to normally clustered data. This concept aims to flag potential anomaly OBD snapshots and provide a focused list of the specific causes for the observed abnormal behaviour.

Fig. 3 illustrates a typical utilization of the proposed snapshot-based anomaly detection. First, the vehicle test is performed in the sense of PVE J1 (SAE J1699) test, CTC implementation, or simply by checking OBD vehicle compliance (described in Sections I and III). The OBD data snapshot is derived from the first phase, since all mentioned technical procedures include it. Secondly, the tester stores the snapshot in the database, collecting groups of snapshots from various vehicles. This generates a considerable snapshots database that is ready for further processing in the backend, which is managed by the engineers in charge of the tests, validations, checks, and other programs. At this point, the anomaly detection pipeline is initiated in the backend by an application engineer who has access to the snapshots database and selects the desired set of snapshots for the analysis. This is not done in real-time, but independently of the vehicle in the backend, after data collection inside the vehicle is done. The goal of the application engineer here is to provide final checks for the vehicle tests by analysing the data found in the database. This is important since the OBD system that does not show any faults or other problems does not guarantee compliance or normal behaviour, as discussed in previous sections. An anomaly detector, the blue-highlighted process in Figure 3, is designed specifically to fill this existing gap. The anomaly detector process consists of the following steps: dimension reduction, cluster identification, outlier extraction, and root

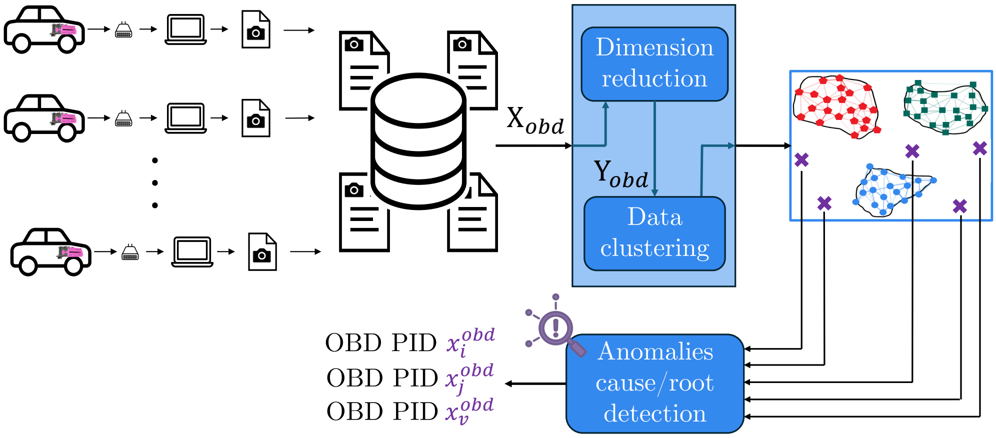


FIGURE 2: Architecture of proposed vehicle diagnostic system anomaly detection based on OBD snapshot.

cause investigation. These steps are explained in greater detail further in this section. The anomaly detector finalizes with the creation of a technical report for the application engineer, including a visualized representation of the database in reduced dimensions, problematic snapshots, their data, and potential root causes. The engineer either acknowledges for each vehicle separately that it is normal, or provides additional actions for anomaly-labelled vehicle snapshots. Such actions can be repair instructions, issuing fines, reinitiating the tests, or further investigations.

As a first step, the t-distributed Stochastic Neighbor Embedding (t-SNE) algorithm is used for dimension reduction, as it outperformed other approaches. t-SNE is a dimension reduction algorithm that maps data points to a k-dimensional space. It is one of the most popular methods used for dimension reduction and is widely used in machine learning and data visualization. In the following, the basics of the algorithm are explained. $\mathbf{P}$ is a similarity matrix of the OBD snapshot X_{obd}, while $\mathbf{Q}$ is the similarity matrix of the resulting dimension-reduced data Y_{obd}. The exact definitions of these similarity matrices can be found in the originally proposed algorithm [53]. t-SNE aims to find y_j that minimizes the KL divergence between $\mathbf{P}$ and $\mathbf{Q}$, that is described by Equation (7).

$$\begin{aligned}(y_1,\dots,y_k) &= \arg\min_{y_1,\dots,y_k} D_{KL}(\mathbf{P},\mathbf{Q})\\ &= \arg\min_{y_1,\dots,y_k} \sum_{\substack{i,j\in\{1,2,\dots,n\}\\ i\neq j}} p_{ij}\log\frac{p_{ij}}{q_{ij}} \qquad (7)\end{aligned}$$

Many algorithms have been proposed to solve this equation, and the most common is a variant of the gradient descent algorithm, with an updating equation [54]. While other dimension reduction algorithms have been evaluated in this work, such as Uniform Manifold Approximation and Projection (UMAP) [55], the best results were obtained with t-SNE. The results are represented as labelless, dimension-reduced OBD snapshot data points scattered across the latent space.

The following step is applying a clustering algorithm on the dimension-reduced data matrix Y_{obd}. The Density-Based Spatial Clustering of Applications with Noise (DBSCAN) clustering approach is chosen for this use case, as better performance was obtained in comparison with other methods such as k-means and Local Outlier Factor (LOF). DBSCAN is an unsupervised learning method and belongs to the class of density-based clustering algorithms. It identifies clusters as regions of high density in the data or latent space, which are separated by areas of lower density. In contrast to partitioning clustering algorithms such as k-means, density-based methods allow for the identification of clusters with arbitrary shapes in n-dimensional space. This is especially favorable because the latent representations of data, which are generated through dimensionality reduction, frequently involve intricate structures that are not adequately described by spherical boundaries. By connecting points with locally high density, dense regions are formed that can be interpreted as clusters. The local density of a data point q is defined by

$$N_\epsilon(q) = \{p \in D | dist(p,q) \leq \epsilon\}, \qquad (8)$$

where ϵ describes the radius of the neighborhood of the data point q. A core object is a point q that satisfies $|N_\epsilon| \geq MinPts$, which means that a sufficient number of neighboring points are located within its density region. A point p is said to be directly density-reachable from another q if $p \in N_\epsilon(q)$ and q is a core object held. If a point p is reachable from q via a point o, and both p and q are density-reachable from o, then p and q are considered density-connected. A dense region thus comprises all points that are mutually density-connected. The set of densely clustered points from Y_{obd} can be partitioned into clusters $\{C_1,...,C_w\}$, such that: $C_i \subseteq Y_{\text{obd}}, C_i \cap C_j = \emptyset$. Points that do not belong to any of these clusters form the residual set $Y_{\text{obd}} \setminus \{C_1,...,C_w\}$ and are referred to as outliers. These points lie in low-density regions and, with respect to the parameters ϵ and $MinPts$, cannot be assigned to any cluster [56]. It is assumed that such points are generated by a different process and can therefore be interpreted as anomalies. In other

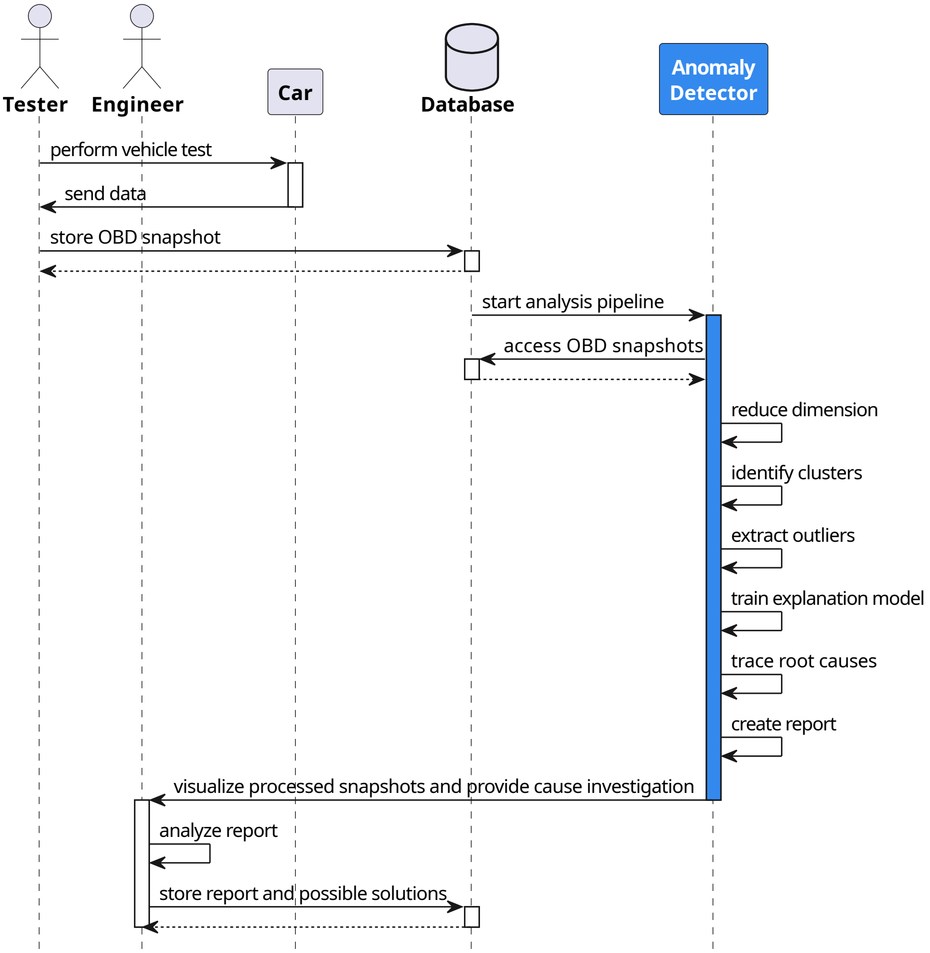


FIGURE 3: Illustration of the interaction between the stakeholders of the anomaly detection pipeline in OBD vehicle diagnostics.

words, DBSCAN labels each OBD snapshot by assigning it a cluster $\in \{C_1, ..., C_w\}$, or marks potential anomalies by assigning them outside of all clusters $\notin \{C_1, ..., C_w\}$.

Finally, the causes for anomalies are investigated using SHAP methodology. SHAP provides a unified framework for interpreting the output of various data processing models by quantifying the contribution of each input feature to a given label prediction [57]. The methodology is often referred to as a benchmark for XAI, the solution of transforming systems from black-box models into white-box ones. It aims to achieve transparent, interpretable, explainable, and dependable systems [58], overall very valuable in the context of anomaly detection in automotive. SHAP methodology outputs a single SHAP value for each feature, and can be calculated for a set of OBD snapshots using the equation below:

$$SHAP_{x_i^{\text{obd}}} = \sum_{S \subseteq N\{x_i^{\text{obd}}\}} \frac{|S|!(n-|S|-1)!}{n!}[\nu(S \cup \{x_i^{\text{obd}}\}) - \nu(S)], \tag{9}$$

where $SHAP_i$ is the SHAP value of each feature x_i^{obd}, N represents a set of all features $[x_1^{\text{obd}}, ...x_i^{\text{obd}}, ...x_n^{\text{obd}}]$, n is the number of OBD PIDs, set S is the subset of N which contains feature x_i^{obd}, and finaly ν is the base value of the predicted outcome for each feature x_i^{obd} in N [59].

The explainability using SHAP is done on a trained Random Forest (RF) model that has a goal of anomaly classification. After the dimension reduction and clustering, RF needs to be trained based on the output labels from previous steps. Based on trained RF model using the OBD snapshots and labels, SHAP values enable a quantified understanding of which specific OBD PIDs (i.e. x_u^{obd}, x_v^{obd}, x_m^{obd}) strongly influence the DBSCAN anomaly detection decisions. The SHAP values are ordered based on the magnitude, and the highest ones are considered to be the causes for the anomalies. The SHAP methodology is done on the preprocessed snapshot input data X_{obd} (see Equation (4)), after DBSCAN labels them as anomaly snapshots. By attributing anomalies to particular OBD features, SHAP facilitates the identification of underlying systems or components that are likely responsible for abnormal behavior. This interpretability not only supports targeted troubleshooting and maintenance but also highlights potential design flaws or areas where system improvements are necessary, giving it high importance in vehicle testing phases.

In the case of the explainable anomaly detection approach depicted in the Figure 2, the anomalies are firstly detected in lower dimensions, as previously described. Each snapshot is therefore labeled normal or anomaly. For the labeled snapshots, the next step for tracing root causes of anomalies is performed by training an RF classifier on combined labeled data in the original dimensions. The combined labeled data consists of all normal snapshots and each snapshot that is labeled as an anomaly separately. The RF model is chosen here for its compatibility with the TreeExplainer method in the SHAP framework, which efficiently computes SHAP values for tree-based models. After training, the TreeExplainer generates SHAP values for each anomaly individually to quantify the contribution of OBD PID features to the classification decision, or the root cause of the anomaly in this case. These features are then ranked by the mean absolute SHAP value over all anomaly samples.

V. EXPERIMENTAL STUDY

A. DATASETS AND EVALUATION METRICS

The evaluation of the anomaly detection concept proposed in the previous section is done with two scenarios, the first one with the engine off, and the second with the engine on data snapshots. Both scenarios have separate sets of regular 1057 snapshots, expected normal, from different ICE vehicles. A high variety of different ICE vehicles in exploitation were included, both spark and compression ignition engine types, model years from 2014 to 2024. On top of that, for each scenario, 140 anomaly snapshots were generated, accounting for 13.2% of total OBD snapshots. Anomaly snapshots include 50 snapshots of vehicle operating performance drop anomalies, 30 engine coolant system problems, 30 snapshots of fuel system problems, and 30 snapshots of engine ECU tampering, again for each evaluated scenario. The last three anomalies that were generated using Gaussian perturbation, each severity level $d_a \in \{0.5, 1.5, 3\}$ had 10 anomaly snapshots (see Section III-B and Equation (3)).

Evaluation of proposed anomaly detection performance for both cases is done with the model confusion matrix and its derivatives, accuracy, precision, recall, and F1 score. The confusion matrix consists of four basic characteristics that are used to define the measurement metrics of the classifier, in this case, anomaly or normal OBD snapshot. These four characteristics are: True Positive (TP) that represents the percentage of data points that have been properly classified as anomalies; True Negative (TN) the percentage of correctly classified snapshots that are normal; False Positive (FP) the percentage of misclassified snapshots with the anomaly but they are clasified as normal; False Negative (FN) the percentage of snapshots misclassified as normal but actually are anomalies [60]. Accuracy, precision, recall, and F1 score are calculated from the values of TP, TN, FP, and FN. The proposed anomaly detection approach is evaluated against the Isolation Forest (IF). The IF is considered to be a benchmark for the general anomaly detection in the literature due to its ability to isolate anomalies effectively by recursively partitioning the data [61]. It identifies outliers as points that require fewer splits to isolate in random trees.

B. ENGINE OFF SCENARIO

1) Scenario Description

Engine off represents the state of the vehicle where the engine is not active, but the ignition is on. This can happen before engine cranking or during a short stop in the driving cycle (i.e., during a traffic light) for start-stop system engine types. In this state, the main propulsion-related systems and controllers are powered on and are in the stage of preparing to turn on the engine. Usually, the Engine Control Module (ECM), engine ECU, coordinates the state of the vehicle with component boot order and monitors their early behaviour. A total of 55 PIDs x_i^{obd} are available from the engine off OBD snapshots, the complete list is in the Table 5. A lot of irregularities in the engine operation, emission regulation, fuel system, and others could be detected in this state using anomaly detection. The challenging task in the engine-off scenario arises from the data being more uniformly distributed, resulting in lower Shannon data entropy for many parameters during this stage (in the case of our data, 29%). In this case, anomalies representing engine ECU tampering are disregarded since the engine is off and the vehicle is not driving. The visualization of key PIDs for anomalies (see Table 2) in normal and anomaly snapshots is shown in Figure 10.

2) Dimension Reduction

The scenario including engine off data with true labels (normal and anomaly types) after dimension reduction using t-SNE gives a result represented in Figure 4. Each data point in the 2-D plot represents one vehicle OBD snapshot, in the engine off state. The figure can be interpreted as being divided into two parts, left and right, from the t-SNE dimension 1 value 0. The t-SNE created larger line-shaped normal snapshot clusters (blue points), indicating their data-similarity closeness. The figure shows variable separation of anomaly (orange/red/purple points) from these normal snapshot clusters. Some anomalies are obviously separated, while others are merged among the normal points (blue). The best separation gave the general system performance drop anomalies (orange), while other types of anomalies are more mixed with the normal data in the latent space. A greater distance between anomalies and normal data points makes it possible to detect anomalies using DBSCAN. The different anomaly types are usually kept in separate smaller groups (2-6 snapshots), around the aforementioned line-shaped normal data clusters. The t-SNE model with a perplexity of 200 and $k = 2$ reduced dimensions shows the best results for dimensionality reduction. Quantitatively, t-SNE outperformed the other considered dimension reduction technique UMAP, as concluded from the Silhouette score (s) for each approach: $s_{\text{off}}^{\text{t-SNE}} = 0.0657$; $s_{\text{off}}^{\text{UMAP}} = 0.0002$. The Silhouette score represents a widely used metric that measures how well data

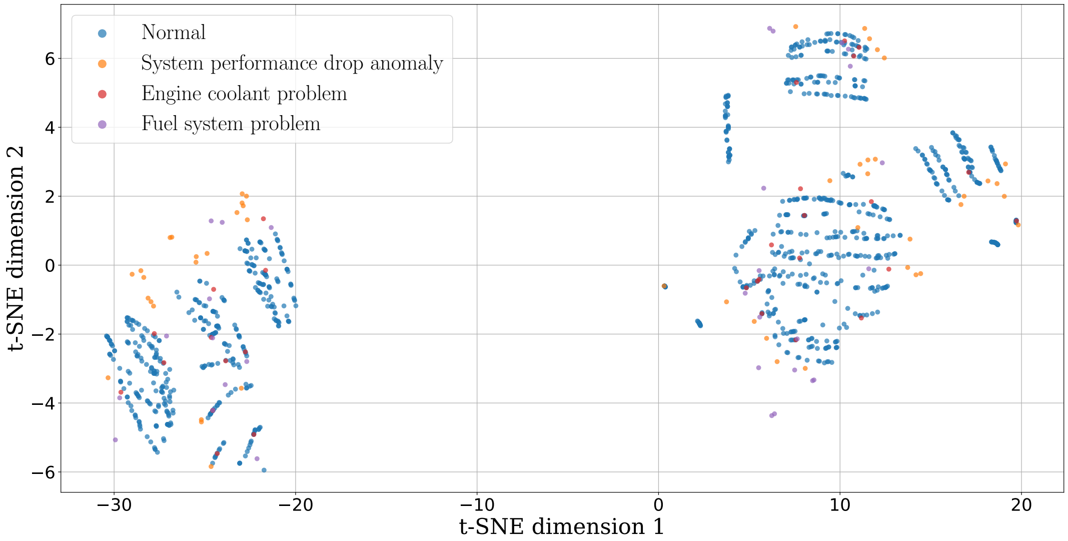


FIGURE 4: Results of dimension reduction using t-SNE for engine off data snapshots scenario. The data is reduced to two dimensions with a hyperparameter perplexity of 200.

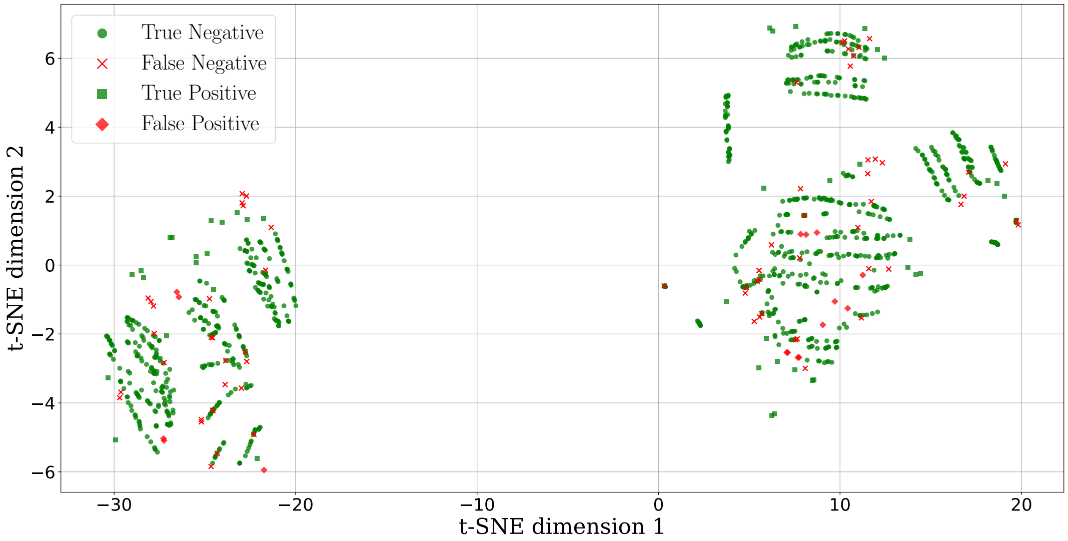


FIGURE 5: Results of t-SNE and DBSCAN in anomaly detection for engine off OBD snapshot. The results are shown labeled with confusion matrix results, and the model in this case showed an accuracy of 92.89%.

points with different labels (normal/anomaly) are separated in the embedding.

3) Anomaly Detection

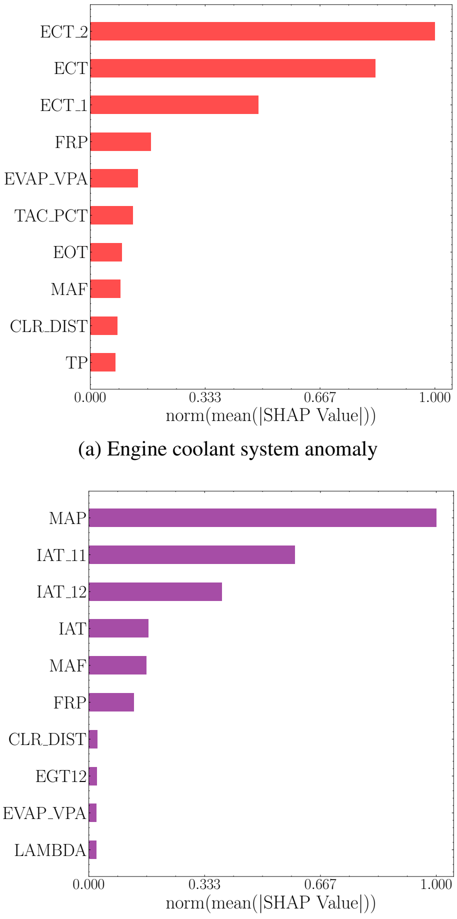
Further anomaly classification in the engine off scenario is done using DBSCAN. The normal/anomaly labels are originally unknown to the DBSCAN model, they are derived from the formed dense regions and isolated points after DBSCAN processing. Various DBSCAN hyperparameters were evaluated with the dimension-reduced data, and the best result provided a hyperparameter combination of maximum distance between two neighbor points $\epsilon = 0.5$, minimum number of samples within ϵ to form a cluster min_dist $= 3$, and distance metric type *Manhattan*. The model in this scenario showed overall anomaly detection accuracy of 92.89%, precision of 0.729, recall of 0.391, and F1 score of 0.509. The overall and per-anomaly type normalized confusion matrices are displayed in the Table 3. The confusion matrices are normalized based on the actual label for better visual inspection in a way that $\mathrm{TP} + \mathrm{FN} = 100\%$ and $\mathrm{TN} + \mathrm{FP} = 100\%$. The results of anomaly classification using this DBSCAN model are shown in Figure 5, where each snapshot is labelled according to the confusion matrix. As expected, the DBSCAN anomaly classification performed well for those anomaly data points that were clearly separated from the clusters of normal data in the latent space (i.e., TP data points located on the left side above the normal clusters). Anomaly snapshot data points that overlapped with normal data after dimension reduction were not identified as anomalies (i.e., FN data point groups on the right side). The analysis of per-anomaly type confusion matrices is done against normal data. Each anomaly type excludes other types of anomalies in evaluation for more interpretable results. The results of such analysis show that the performance drop anomalies had the highest rate of positive detections, compared to the other two types of anomalies. The reason for this can be found in the dimension reduction analysis done in the subsection V-B2, where the performance drop anomaly type had the largest overall separation from the normal datapoints. The solid separation had the fuel system anomalies, and the much poorer separation was done with the engine coolant system anomaly. This directly reflects the results of DBSCAN (see Table 3). This concludes that the anomaly detection in this case highly depends on the dimension reduction result in terms of the algorithm and hyperparameters. Compared to the baseline benchmark, the results can be seen in Appendix III, Table 6. IF performs better on data points with clear separation, as shown in V-B2. In the more challenging anomaly patterns, such as the engine coolant and fuel system, it becomes apparent that the proposed outlier detection pipeline outperforms the pure IF approach due to its dimension-reducing preprocessing. The proposed approach detects approximately 30% more fuel system anomalies in this scenario, while it is outperformed in the performance drop anomaly case by 27%.

TABLE 3: Confusion matrices results for overall and per-anomaly type of anomaly detection using DBSCAN for engine off scenario.

Anomaly group	TP	TN	FP	FN
All	39.09%	98.49%	1.51%	60.91%
Performance drop	54.00%	98.49%	1.51%	46.00%
Engine coolant	6.67%	98.49%	1.51%	93.33%
Fuel system	46.67%	98.49%	1.51%	53.33%

4) Cause Analysis

SHAP analysis of the engine off snapshots scenario is done to identify the causes of individual anomalies. The SHAP was implemented by training an RF classifier in the original data space after the snapshots were binarly classified to be normal or anomaly. The binary classification RF model was trained using 100 decision trees and is initialized with a fixed random seed to ensure reproducibility of results. After training, TreeExplainer with an RF model was used to assign the SHAP value for each OBD PID of anomaly-labelled snapshots. In order to evaluate the explainability and root cause identification of the proposed anomaly detection approach, the anomalies of the same type are grouped, and their mean SHAP values for each PID (mean$|SHAP_{x_i^{\text{obd}}}|$, for $1 \leq i \leq n$) are calculated. Furthermore, the mean SHAP values are normalized, since only the relative ratio between the SHAP values of different PIDs is relevant. The outcome in the cases of engine coolant systems and fuel system problems is shown in Figure 6. The SHAP analyzed each of 55 PIDs in the engine off scenario and assigned a SHAP value, while the 10 highest are shown in the Figure. In both cases, the SHAP method showed success in finding the root causes of specific anomalies by giving the specific PIDs the highest magnitudes of SHAP values. In the case of engine coolant temperature, root causes are coolant temperature sensors (ECT, ECT_1, ECT_2), as described previously in Table 2. Using heuristic investigation, the problematic system can be pinpointed using marked signals, in this case engine coolant system. The SHAP in the second anomaly case of the fuel system problem marked all 6 root cause PIDs from Table 2. The final cause of the second anomaly can be pinpointed using the 6 detected signals, leading to the fuel system anomaly. The SHAP analysis in the anomaly case of general system performance drop is skipped, since the anomalies are generated on random sets of PIDs for each anomaly of this type, making the results impossible to validate. Overall, SHAP analysis showed success in detecting the root causes for individual anomalies in the case of the engine off snapshots scenario.



(a) Engine coolant system anomaly

(b) Fuel system anomaly

FIGURE 6: Results of the SHAP analysis for the engine off snapshots scenario, including the top 10 most influential parameters with normalized mean SHAP values for different anomalies.

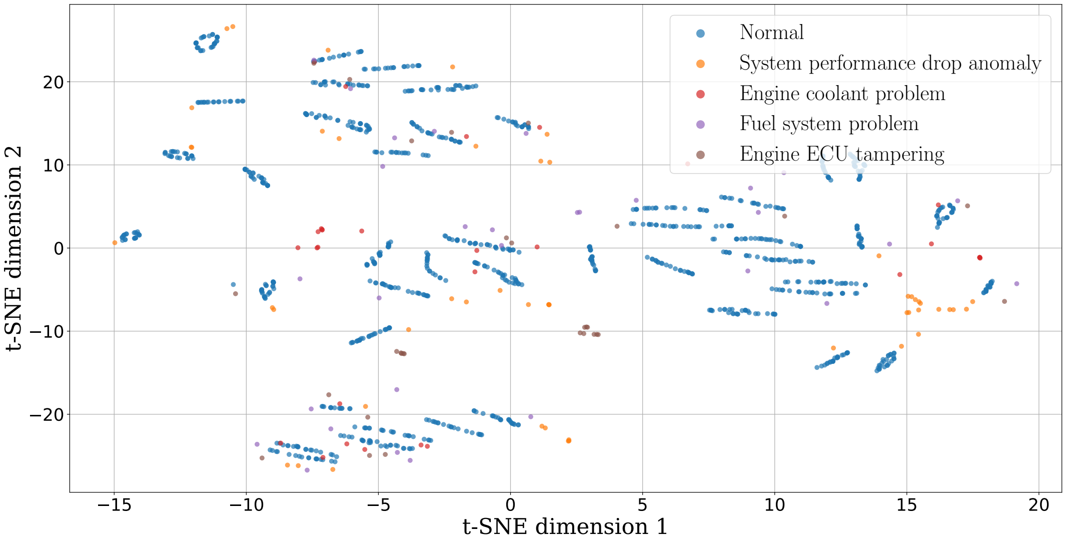


FIGURE 7: Results of dimension reduction using t-SNE for the engine on data snapshots scenario. The data is reduced to two dimensions with a hyperparameter perplexity of 100.

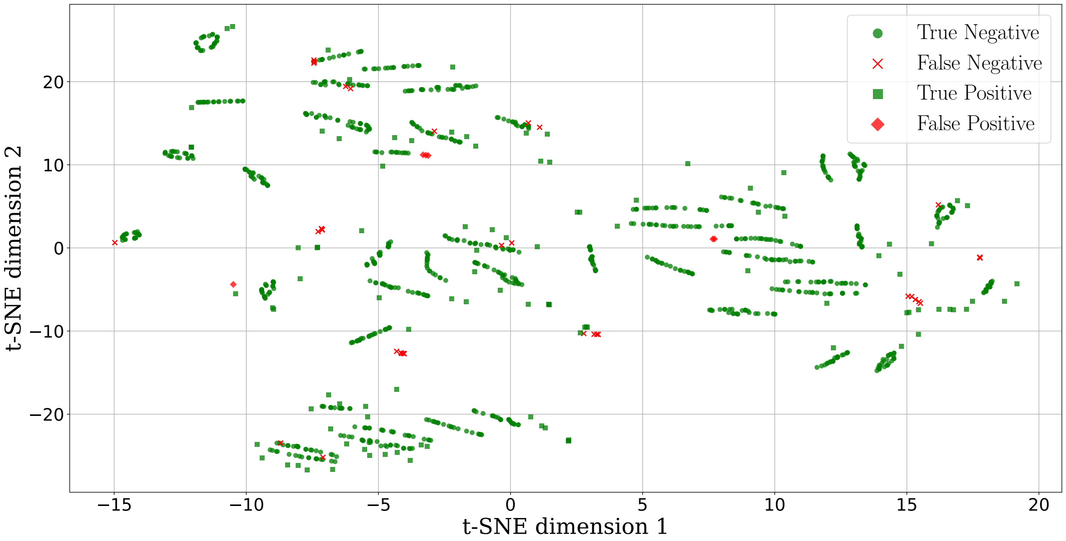


FIGURE 8: Results of t-SNE and DBSCAN in anomaly detection for engine on OBD snapshot. The results are shown labeled with confusion matrix results, and the model in this case showed an accuracy of 96.45%.

C. ENGINE ON SCENARIO

1) Scenario Description

The engine on data snapshots represents the vehicle state in driving or parking mode, where the engine has been running for some period. This is important to avoid the potential false positive anomalies at the specific moment of cranking the engine or during preparation, warm-up cycles. For example, this is managed in the J1699 test by forcing the tester to wait 30 seconds after the engine is turned on. All anomaly types from Table 2 are accounted for in this scenario, and the total number of PIDs x_i^{obd} is 57. The total list of PID parameters included here is shown in Table 5. Some parts of propulsion-related systems are still not active or not responding to ECM in the engine off scenario, therefore more PIDs are found in the engine on state. The representation of normal and anomaly snapshot values for PIDs used to create anomalies in the engine on scenario (see Table 2), is shown in the Figure 11.

2) Dimension Reduction

The t-SNE for dimension reduction in this scenario results in Figure 7 that shows snapshots with their true labels. The axes in the figure represent the abstract t-SNE output variables, y_j for $1 \leq i \leq k$, after dimension reduction to a latent space. Dimension-reduced normal snapshots, marked with blue dots, in this case form more distinct cluster groups, compared to the engine off case. This creates a better ground base for the separation of anomalies (orange, red, purple, brown points), crucial for their later detection. Contrary to the other scenario, the reduced dataset cannot be divided into two parts, but represents relatively equally distant clusters. The normal data keeps the line-shaped clusters for most parts of the dataset, out of which the majority is horizontally oriented (in the direction of the constant values of t-SNE dimension 2 axis). The engine on scenario contains an additional anomaly compared to the engine off scenario - an engine ECU tampering. The anomalies are reduced in relative proximity to the normal data clusters, but far enough to be detectable as anomalies in latent space. In this case, all types of anomalies are well separated from normal clusters, but they create larger groups (2-12 datapoints) than in other scenario. The larger groups are a direct cause of the majority of false negatives in later anomaly detection. The dimension reduction was done using the t-SNE model with perplexity 100, and $k = 2$ reduced dimensions. Once more, the t-SNE outperformed UMAP in the engine on scenario, as depicted with Silhouette scores: $s_{\text{on}}^{\text{t-SNE}} = 0.3238$; $s_{\text{on}}^{\text{UMAP}} = 0.2934$.

3) Anomaly Detection

The DBSCAN hyperparameters combination that gave the best results in the engine on case is: maximum distance between two neighbor points $\epsilon = 0.5$, minimum number of samples within ϵ min_dist = 4, with distance type *euclidean*. After visually better separation of normal and anomaly snapshots with t-SNE than in other scenario, evaluation metrics in the engine on scenario show an increase in achieved results with accuracy of 96.45%, precision of 0.945, recall of 0.743, and F1 score of 0.832. Normalized confusion matrices in the engine on scenario for anomaly detection using DBSCAN overall and per anomaly type are shown in the Table 4. The majority of misslabeled anomaly snapshots in this case (False Negatives) are the tight groups of the same anomaly labels that form a cluster and are hard to detect. For example, this is the case for grouped tampering anomalies in the middle (brown points in the Figure 7, red × in the Figure 8) or system performance drop anomalies in the lower right side (orange points in the Figure 7, red × in the Figure 8). This proves once more that the dimension reduction step is crucial for precise anomaly detection using DBSCAN. The per-anomaly type confusion matrices show the best detection of the performance drop anomalies. Furthermore, other types of anomalies performed much better than in the engine off scenario. This is mainly due to the increased dynamics of PIDs in the engine on scenario. In the case of the engine on scenario, the dataset has a broader range of signals. Consequently, interpretability decreases for high-dimensional spaces (see Section V-C2). This is also evident in the application of the IF. Here, the presented anomaly detection pipeline outperforms the IF applied to this scenario. In each anomaly case, fewer anomalies are detected with the benchmark solution (see Table 7). Furthermore, the IF showed a lower rate of TN compared to the proposed anomaly detection.

TABLE 4: Confusion matrices results for overall and per-anomaly type of anomaly detection using DBSCAN for the engine on scenario.

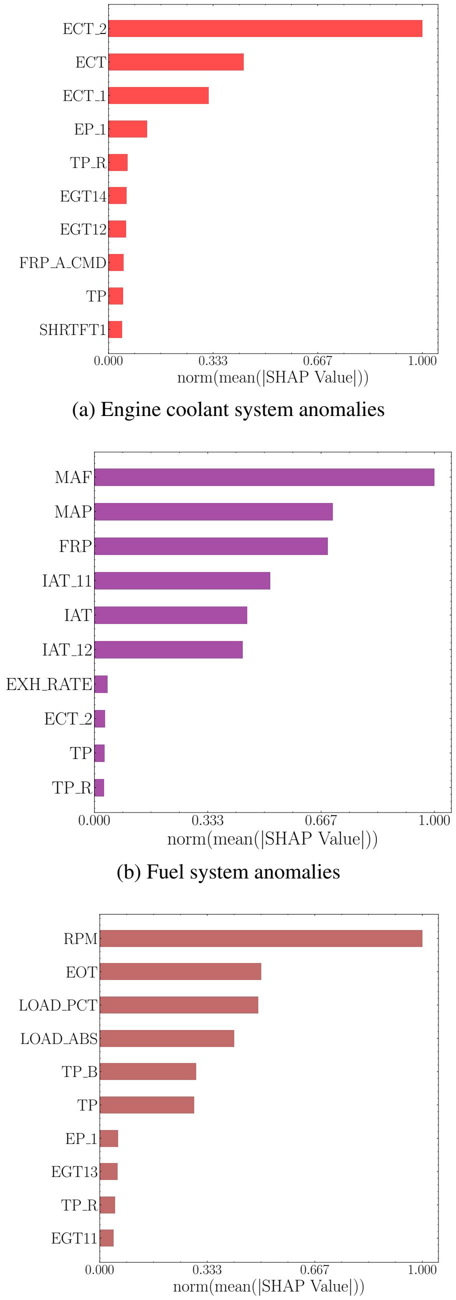
Anomaly group	TP	TN	FP	FN
All	74.29%	99.43%	0.57%	25.71%
Performance drop	88.00%	99.43%	0.57%	12.00%
Engine coolant	53.33%	99.43%	0.57%	46.67%
Fuel system	86.67%	99.43%	0.57%	13.33%
Tampering	60.00%	99.43%	0.57%	40.00%

4) Cause Analysis

SHAP analysis further investigates the specific causes of individual anomalies in the engine on evaluation scenario. More PIDs are available in this scenario, and their expected values in snapshots should be more structured and dynamic. Again, the SHAP was implemented by training the RF classifier after DBSCAN with labeled snapshots. The RF model used the same hyperparameters as in the last case, 100 decision trees and a fixed random seed. TreeExplainer assigned the SHAP value to each PID of snapshots that are labelled as anomalies. The results are grouped according to the anomaly type, and the normalized mean SHAP values are ranked according to magnitude. The 10 highest values for the cases of anomalies in the engine on scenario are shown in Figure 9. Three anomaly types were evaluated using SHAP: engine coolant system problems, fuel system problems, and engine ECU tampering. For the cases of engine coolant problems and fuel system problems, all influencing PIDs were detected. Three PIDs in case of coolant system problem and all six PIDs in the case of fuel system problem (see Figures 9a, 9b and Table 2). The output of their main subsystems showed the highest magnitude of normalized mean SHAP values, therefore the anomalies can be seamlessly pinpointed to the respective root systems. Finally, the SHAP managed to detect 6 out of 7 relevant signal causes in the case of tampering. The relative throttle position signal (TP_R) was not included in the top 10 signals of average SHAP value (see Figure 9c, and Table 2). The main reason for this can be found in the anomaly generation part, since the values of this PID for generated anomalies are almost identical to the normal snapshots (see blue and brown datapoints in the TP_R plot inside the Figure 11). With one signal missing, the problematic component (e.g., engine ECU, engine speed, and throttle sensors) could still be focused, since the other two throttle sensors were detected as potentially problematic. With this precision of the cause detection in the engine on scenario, it can be concluded that the SHAP methodology showed great success and practicality in the case of OBD data. Combining with the results of the engine off evaluation scenario, it proves the potential for cause detection with vehicle known and unknown problems using OBD data.

VI. DISCUSSION

An observation can be made from the distance metrics perspective of the t-SNE dimension reduction algorithm for different scenarios of OBD anomaly detection based on snap-



(a) Engine coolant system anomalies

(b) Fuel system anomalies

(c) Engine ECU tampering anomalies

FIGURE 9: Results of the SHAP analysis for the engine on snapshots scenario, including the top 10 most influential parameters with the normalized mean SHAP values for different anomalies.

shots. The *Manhattan* distance metric achieved the best separation between true anomalies and normal data, while the *Euclidean* metric in the engine on scenario performed better. This is likely due to the unique inherent characteristics of the data distributions in each scenario. This is emphasized with the lower Shannon data entropy for 29% of the PID parameters in the case of engine off, which is not the case with the engine on scenario. When the ignition is on and the engine is off, the OBD data tends to be more discrete, sparse, or less smoothly varying, which can result in higher-dimensional data with localized clusters. The *Manhattan* distance metric is often more effective at preserving meaningful neighborhood structures during dimensionality reduction in this case. Contrary to the engine in the engine on scenario, the engine system is active and generating interrelated measurements across a broader range of sensors, leading to a denser and more smoothly varying dataset. Therefore, the *Euclidean* distance is better suited here to capture the global geometric relationships among snapshots. Thus, the difference in optimal distance metrics for t-SNE in the two scenarios reflects the underlying difference in the structure and variability of the OBD data in the scenarios.

A sensitivity analysis of the DBSCAN hyperparameters, namely ϵ and *MinPts* (see Section IV), is also considered. A parameter sweep over ϵ and *MinPts* was conducted to evaluate how the sensitivity of the model depends on the chosen distance metric. The sensitivity of the True Positive Rate (TPR) is used as the evaluation criterion. In the engine off scenario, the choice of hyperparameters has a significantly stronger effect on model performance. From $\epsilon = 0.5$ onwards, no reliable predictions can be achieved (TPR $< 70\%$). In contrast, the spatial density in the engine on scenario is lower, which is reflected in the reduced influence of the neighborhood radius. A noticeable degradation in prediction performance occurs only from $\epsilon = 0.9$ for the *Euclidean* metric and from $\epsilon = 1.1$ for the *Manhattan* metric. If ϵ is chosen too large, the neighborhood around each core object becomes excessively wide. As a result, individual clusters and noise points can no longer be separated, making the identification of anomalies impossible. The choice of *MinPts*, however, has only a negligible effect. Since the points lie sufficiently densely in the projected feature space, the formation of clusters with too few points does not occur. Furthermore, the influence of distance metric shows that in both scenarios the *Manhattan* distance yields a more robust metric in terms of the TPR (see Figures 12, 13). The superior sensitivity of the *Manhattan* distance may be explained by the geometry of the low-dimensional feature space. While the *Euclidean* metric defines a circular (or spherical) neighborhood around each core point, the *Manhattan* distance forms a diamond-shaped region. This leads to a different notion of locality and thus influences the clustering result. However, as already demonstrated in Section V-C2, the spherical approximation yielded better overall results. In summary, the sensitivity of the proposed pipeline depends primarily on the preceding dimensionality reduction step, as it implicitly determines an appropriate choice of the ϵ parameter. Based on the distribution obtained from the parameter study (see fig. 12a, 12b, 13a, 13b), it can be determined that a suitable initial value for the ϵ parameter lies in the range of 0.1 to 0.3. In contrast, the result is largely independent of the choice of *minPts*, provided that this parameter is not selected too small (*minPts* ≥ 4). In the presented DBSCAN use case, the Silhouette Coefficient in combination with elbow point detection could potentially be employed to estimate a suitable number of clusters and, consequently, to determine the associated exclusion of noise points [62].

The consistent performance across both scenarios demonstrates the robustness of the approach and suggests its applicability with OBD data. Higher precision of 96.45% showed the engine on scenario (contrary to other scenario precision of 92.89%), which is more feasible and reasonable to be used for this type of application. Dimension reduction was a critical step in the overall process of anomaly detection and had a major influence on the anomaly detection output. The overall result provides empirical evidence supporting and validating the Hypothesis 1, where latent spaces after dimension reduction of the original data can effectively be used for anomaly detection of OBD. Limitations of the proposed solution represent a relatively high percentage of false negative classifications in both scenarios (60.91% and 25.71% for engine off and on, respectively). This is acceptable for the use case of engine on scenario, and even common in the vehicle anomaly detection due to the wide range of anomaly variations [63]. For the engine off scenario, it is suggested to use benchmark solutions, such as IF. Furthermore, the analysis of FN rate per anomaly type reveals that the highest number comes from engine coolant anomaly type (93.33% and 46.67% for engine off and on, respectively). This is a direct result of a poor separation of this anomaly type from normal data in the latent space after t-SNE dimension reduction, observed with red and blue datapoints in Figures 4 and 7).

In addition to anomaly detection, the methodology successfully identified root causes for various anomaly types using the SHAP interpretability method, by pinpointing almost all causing PIDs for individual anomalies. More specifically, all 9 PID signals were identified correctly as causes in the engine off scenario, and 15 out of 16 in the engine on scenario. Despite being evaluated using synthetic anomalies, the proposed solution is expected to hold practical usefulness with groundthrough data. Real anomaly datasets have a larger variation of anomalies, but the anomaly quantity is reflected in the paper. With this in mind, a slight variation of precision for t-SNE/DBSCAN combined anomaly detection is expected. The retuning of hyperparameters is almost certain for the optimal results with different OBD snapshot datasets. Due to the higher anomaly variability in real cases, it is expected that the SHAP method performs less precisely in general, but remains useful in the root detection for the majority of anomaly cases. It can be concluded that the contribution keeps the practical use despite the limitations of using synthetic data.

VII. CONCLUSIONS AND FUTURE WORK

This paper presents a novel pipeline for explainable anomaly detection in the case of vehicle diagnostics testing, validation, and inspection. The concept was evaluated using OBD data snapshots from ICE vehicles. The proposed approach combines t-SNE for dimensionality reduction and DBSCAN for clustering and anomaly detection. Furthermore, the solution supports anomaly cause investigation using SHAP to determine potential roots of the individual anomalies. Four different OBD-relevant anomaly types are used for verification of the concept. The paper presented an integration into the real technical applications for vehicle systems validation and aftermarket checks, such as PVE OBD compliance tests (SAE J1699-3) and CTC emission regulation tests. This provides qualitative enhancements to the mentioned technical procedures by extending the range of detectable out-of-order vehicle systems behaviours that modern OBD systems are not capable of. In addition, the root anomaly causes in the sense of problematic systems would be traceable using the XAI proposed solution for OBD PID signals.

The presented results provide empirical support for the initial hypothesis that anomalies in vehicle operation and diagnostic systems are detectable within lower-dimensional representations of OBD data. The aim is to find the most distinct separation between normal and anomaly snapshots, in this case achieved with t-SNE. This validates the hypothesis that latent spaces derived from the high-dimensional OBD data can serve as an effective basis for anomaly detection, and indicates that the dimension reduction is a critical step. This is proven in the evaluation for both scenarios, where poor separation of anomalies from normal data in lower dimensions made such anomalies undetectable using DBSCAN. Within the success of t-SNE dimension reduction, an observation is made that the *Manhattan* distance metric has better results for the lower Shannon entropy case of engine off high-dimensional OBD data. Conversely, in the engine on case of denser varying OBD snapshot data, *Euclidean* distance dominated. Overall, the solution proved better results in the engine on scenario, and showed significant application potential for real-world vehicle diagnostics and compliance testing.

The future work shall cover the cases of hybrid and electric vehicles. Besides the architecture and operation of propulsion-related systems, the main difference is the snapshot input OBD data dimensionality. HEVs support, on average, more than 100 PID, while EVs support around only 20 so far. This would bring necessary changes to the initial part of the pipeline, more specifically the dimension reduction and anomaly detection. Furthermore, while hybrid and ICE vehicles are falling under the regulation of the SAE J1699-3 test, the EVs shall in the future use a different test procedure SAE J1699-5. This subsequently leads to the additional modification of the pipeline for its integration into the test procedure. A suitable extension of the proposed pipeline is an iterative process for the automated detection of appropriate parameters for the selected anomaly detection method. Moreover, the future work will include more detailed analysis and comparison of state of the art root detection approaches, including LIME, Deep Learning Important FeaTures (DeepLIFT), and Layerwise Relevance Propagation (LRP).

REFERENCES

[1] Uwe Kiencke and Lars Nielsen. Automotive control systems: For engine, driveline, and vehicle. *Measurement Science and Technology*, 11(12):1828, dec 2000.

[2] Veljko Vučinić and Dragan Aleksendrić. Neuro-fuzzy control of commercial vehicles braking. *Soft Computing*, 2025.

[3] Emmanouel T. Michailidis, Antigoni Panagiotopoulou, and Andreas Papadakis. A Review of OBD-II-Based Machine Learning Applications for Sustainable, Efficient, Secure, and Safe Vehicle Driving. *Sensors*, 25(13), June 2025.

[4] Sandra Bickelhaupt, Michael Hahn, Nikolai Nuding, Andrey Morozov, and Michael Weyrich. Comprehensive evaluation of logging frameworks for future vehicle diagnostics. *SAE International Journal of Advances and Current Practices in Mobility*, 6(2):1061–1072, 2023. Publisher: SAE International.

[5] Shichun Yang, Hanchao Cheng, Mingyue Wang, Meng Lyu, Xinlei Gao, Zhengjie Zhang, Rui Cao, Shen Li, Jiayuan Lin, Yang Hua, Xiaoyu Yan, and Xinhua Liu. Multi-scale battery modeling method for fault diagnosis. *Automotive Innovation*, 5(4):400–414, 2022.

[6] Andreas Puder, Moritz Zink, Luca Seidel, and Eric Sax. Hybrid anomaly detection in time series by combining kalman filters and machine learning models. *Sensors*, 24(9):2895, 2024.

[7] Luca Seidel, Houssem Guissouma, Andreas Schmid, and Eric Sax. Variant-aware reconfiguration of automotive virtual test environments. In *Vol.9 - Driving Simulation Conference Europe 2024 VR (DSC 2024), Driving Simulation and Virtual Reality Conference and Exhibition, Straßburg, 18th-20th September 2024*, pages 221 – 226, 2024.

[8] Scott Lundberg and Su-In Lee. A unified approach to interpreting model predictions, 2017.

[9] Veljko Vučinić, Frank Hantschel, and Thomas Kotschenreuther. Synthetic on-board diagnostics data generation and evaluation for vehicle diagnostic testing. *SAE Technical Paper*, 2025.

[10] SAE International. J1699-3: Vehicle OBD II Compliance Test Cases, 2021. https://www.sae.org/standards/j16993_202104-vehicle-obd-ii-compliance-test-cases.

[11] Piroska Haller, Béla Genge, Fabrizio Forloni, Gianmarco Baldini, Massimo Carriero, and Georgios Fontaras. VetaDetect: Vehicle tampering detection with closed-loop model ensemble. *International Journal of Critical Infrastructure Protection*, 37:100525, 2022.

[12] Barouch Giechaskiel, Fabrizio Forloni, Massimo Carriero, Gianmarco Baldini, Paolo Castellano, Robin Vermeulen, Dimitrios Kontses, Pavlos Fragkiadoulakis, Zissis Samaras, and Georgios Fontaras. Effect of tampering on on-road and off-road diesel vehicle emissions. *Sustainability*, 14(10):6065, 2022.

[13] Tom Cackette and Zhenying Shao. California's clean diesel program. *The International Council on Clean Transportation, Washington, USA*, 2021.

[14] Yashashree Mahale, Shrikrishna Kolhar, and Anjali S. More. Enhancing predictive maintenance in automotive industry: addressing class imbalance using advanced machine learning techniques. *Discover Applied Sciences*, 7(4), April 2025.

[15] SAE International. J2012: Diagnostic Trouble Code Definitions, 2025. https://www.sae.org/standards/j2012_202509-diagnostic-trouble-code-definitions.

[16] SAE International. Digital Annex of Diagnostic Trouble Code Definitions and Failure Type Byte Definitions, 2024. https://www.sae.org/standards/j2012da_202403-digital-annex-diagnostic-trouble-code-definitions-failure-type-byte-definitions.

[17] Jaydeep Taralkar. FinAI: Deep learning for real-time anomaly detection in financial transactions. *World Journal of Advanced Engineering Technology and Sciences*, 15(2):454–461, May 2025.

[18] Vignes V. M., Sri Harini M. P., Rahul Satheesh, Vipin Das, and Sanjeevikumar Padmanaban. AI-driven cybersecurity framework for anomaly detection in power systems. *Scientific Reports*, 15(1):35506, October 2025.

[19] Weixia Li, Zhurong Dong, Ling Miao, Guoyuan Wu, Zhijun Deng, Jianfeng Zhao, and Wenwei Huang. On-road evaluation and regulatory recommendations for NOx and particle number emissions of China VI heavy-duty diesel trucks: A case study in Shenzhen. *Science of The Total Environment*, 928:172427, 2024.

[20] Zhenyi Xu, Ruibin Wang, Kai Pan, Jiaren Li, and Qilai Wu. Two-stream networks for copert correction model with time-frequency features fusion. *Atmosphere*, 14(12), 2023.

[21] Zhenyi Zhao, Yang Cao, Zhenyi Xu, and Yu Kang. A seq2seq learning method for microscopic emission estimation of on-road vehicles. *Neural Computing and Applications*, 36(15):8565–8576, May 2024.

[22] Emad E. Abdallah, Ahmad Aloqaily, and Hiba Fayez. Identifying Intrusion Attempts on Connected and Autonomous Vehicles: A Survey. *Procedia Computer Science*, 220:307–314, January 2023.

[23] Pedro Andrade, Marianne Silva, Morsinaldo Medeiros, Daniel G. Costa, and Ivanovitch Silva. Teda-rls: A tinyml incremental learning approach for outlier detection and correction. *IEEE Sensors Journal*, 24(22):38165–38173, 2024.

[24] Yann Cherdo, Benoit Miramond, Alain Pegatoquet, and Alain Vallauri. Unsupervised anomaly detection for cars can sensors time series using small recurrent and convolutional neural networks. *Sensors*, 23(11), 2023.

[25] Ahmad Aloqaily, Emad E. Abdallah, Hiba Abuzaid, Alaa E. Abdallah, and Malak Al-hassan. Supervised Machine Learning for Real-Time Intrusion Attack Detection in Connected and Autonomous Vehicles: A Security Paradigm Shift. *Informatics*, 12(1):4, March 2025.

[26] Franco van Wyk, Yiyang Wang, Anahita Khojandi, and Neda Masoud. Real-time sensor anomaly detection and identification in automated vehicles. *IEEE Transactions on Intelligent Transportation Systems*, 21(3):1264–1276, 2020.

[27] Aditya Jain and Piyush Tarey. Anomaly Detection for Early Failure Identification on Automotive Field Data. *International Journal of Prognostics and Health Management*, 14(3), February 2023.

[28] Yifang Zhang, Guizhen Yu, Han Li, Chaoqi Zhang, Lecong Li, and Chuanying Zhang. Anomaly Detection and Fault Diagnosis Method for Autonomous Transport Vehicles on Unstructured Roads. In *2024 IEEE 22nd International Conference on Industrial Informatics (INDIN)*, pages 1–7, August 2024. ISSN: 2378-363X.

[29] Zeping Cao, Kai Shi, Hao Qin, Zhou Xu, Xiaoyang Zhao, Jiawei Yin, Zhenyu Jia, Yanjie Zhang, Hailiang Liu, Qijun Zhang, and Hongjun Mao. A comprehensive OBD data analysis framework: Identification and factor analysis of high-emission heavy-duty vehicles. *Environmental Pollution*, 368:125751, March 2025.

[30] Elifnur Tafralı, Sena Yazan, Gizem Çakırbaş, and Cemhan Demirci. Machine Learning Based Root Cause Analysis of IUMPR Performance in Automotive OBD Systems. In *ELECO 2025 International Conference on Electrical and Electronics Engineering*, 2025.

[31] Chien-Yu Lu, Hong-Yi Hsu, Bo-Siang Chen, Wei-Lun Huang, and Wei-Sho Ho. Development and Validation of an Explainable Hybrid Deep Learning Model for Multiple-Fault Diagnosis in Intelligent Automotive Electronic Systems. *Electronics*, 14(22):4488, January 2025.

[32] Bernardo Tormos, Benjamín Pla, Ramón Sánchez-Márquez, and Jose Luis Carballo. Explainable AI Using On-Board Diagnostics Data for Urban Buses Maintenance Management: A Study Case. *Information*, 16(2):74, February 2025.

[33] Övgü Özdemir, Tuğberk İşyapar, Pınar Karagöz, Klaus Werner Schmidt, Demet Demir, and N. Alpay Karagöz. A survey of anomaly detection in in-vehicle networks, 2024.

[34] A. Avizienis, J.-C. Laprie, B. Randell, and C. Landwehr. Basic concepts and taxonomy of dependable and secure computing. *IEEE Transactions on Dependable and Secure Computing*, 1(1):11–33, 2004.

[35] Seungho Jeon, Kijong Koo, Daesung Moon, and Jung Taek Seo. Mutation-based multivariate time-series anomaly generation on latent space with an attention-based variational recurrent neural network for robust anomaly detection in an industrial control system. *Applied Sciences*, 14(17):7714, 2024. Number: 17 Publisher: Multidisciplinary Digital Publishing Institute.

[36] Thien-Binh Dang, Duc-Tai Le, Moonseong Kim, and Hyunseung Choo. A general model for long-short term anomaly generation in sensory data. *2022 16th International Conference on Ubiquitous Information Management and Communication (IMCOM)*, pages 1–5, 2022.

[37] Marc Weber. *Untersuchungen zur Anomalieerkennung in automotive Steuergeräten durch verteilte Observer mit Fokus auf die Plausibilisierung von Kommunikationssignalen*. PhD thesis, Karlsruher Institut für Technologie (KIT), 2019. Translated title: Investigations into anomaly detection in automotive control units using distributed observers with a focus on the plausibility of communication signals.

[38] Lihui Wang, Xionghui Zou, Hongyu Qin, and Peilin Geng. Design of OBD function test on production vehicle (PVE). *E3S Web of Conferences*, 268:01047, 2021.

[39] Marco Frigessi Di Rattalma. *The dieselgate: a legal perspective*. Springer International Publishing, 2017.

[40] SAE International. J1979: E/E diagnostic test modes, 2025. https://www.sae.org/standards/j1979_202505-e-e-diagnostic-test-modes.

[41] SAE Internationa. J1979-DA: Digital annex of E/E diagnostic test modes, 2025. https://www.sae.org/standards/j1979da-j1979-da-digital-annex-e-e-diagnostic-test-modes.

[42] Masoud Pourreza, Bahram Mohammadi, Mostafa Khaki, Samir Bouindour, Hichem Snoussi, and Mohammad Sabokrou. G2d: Generate to detect anomaly. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, pages 2003–2012, 2021.

[43] Milad Salem, Shayan Taheri, and Jiann Shiun Yuan. Anomaly generation using generative adversarial networks in host-based intrusion detection. In *2018 9th IEEE Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON)*, pages 683–687, 2018.

[44] Chengyu Wang, Kui Wu, Tongqing Zhou, Guang Yu, and Zhiping Cai. TSAGen: Synthetic time series generation for KPI anomaly detection. *IEEE Transactions on Network and Service Management*, 19(1):130–145, 2022.

[45] Mohamed Ahmeid, Musbahu Muhammad, Simon Lambert, Pierrot S. Attidekou, and Zoran Milojevic. A rapid capacity evaluation of retired electric vehicle battery modules using partial discharge test. *Journal of Energy Storage*, 50:104562, 2022.

[46] A. Rauh, L. Hannibal, and N. B. Abraham. Global stability properties of the complex lorenz model. *Physica D: Nonlinear Phenomena*, 99(1):45–58, 1996.

[47] Bo Wen Shen. A review of lorenz's models from 1960 to 2008. *International Journal of Bifurcation and Chaos*, 33(10):2330024, 2023.

[48] John R. Hershey and Peder A. Olsen. Approximating the Kullback Leibler divergence between Gaussian mixture models. In *2007 IEEE International Conference on Acoustics, Speech and Signal Processing - ICASSP '07*, volume 4, pages IV–317–IV–320, 2007.

[49] Thiago B. Murari, Roberto C. da Costa, Hernane B. de B. Pereira, Roberto L. S. Monteiro, and Marcelo A. Moret. Early detection of failing lead-acid automotive batteries using the detrended cross-correlation analysis coefficient. *Applied System Innovation*, 8(2):29, 2025.

[50] Kiran Maharana, Surajit Mondal, and Bhushankumar Nemade. A review: Data pre-processing and data augmentation techniques. *Global Transitions Proceedings*, 3(1):91–99, 2022.

[51] Daniel Engel, Lars Hüttenberger, and Bernd Hamann. A survey of dimension reduction methods for high-dimensional data analysis and visualization. *OASIcs, Volume 27, VLUDS 2011*, 27:135–149, 2013.

[52] Kelsy Cabello-Solorzano, Isabela Ortigosa de Araujo, Marco Peña, Luís Correia, and Antonio J. Tallón-Ballesteros. The impact of data normalization on the accuracy of machine learning algorithms: A comparative analysis. In *18th International Conference on Soft Computing Models in Industrial and Environmental Applications (SOCO 2023)*, pages 344–353, Cham, 2023. Springer Nature Switzerland.

[53] Laurens van der Maaten and Geoffrey Hinton. Visualizing data using t-SNE. *Journal of Machine Learning Research*, 9(86):2579–2605, 2008.

[54] T Tony Cai and Rong Ma. Theoretical Foundations of t-SNE for Visualizing High-Dimensional Clustered Data. *Journal of Machine Learning Research*, 23(301):1–54, 2022.

[55] Leland McInnes, John Healy, and James Melville. Umap: Uniform manifold approximation and projection for dimension reduction, 2020.

[56] Martin Ester, Hans-Peter Kriegel, Jörg Sander, and Xiaowei Xu. A density-based algorithm for discovering clusters in large spatial databases with noise. In *Proceedings of the Second International Conference on Knowledge Discovery and Data Mining*, KDD'96, page 226–231. AAAI Press, 1996.

[57] Guy Van den Broeck, Anton Lykov, Maximilian Schleich, and Dan Suciu. On the tractability of SHAP explanations. *Journal of Artificial Intelligence Research*, 74:851–886, 2022.

[58] Kashif Alam, Kashif Kifayat, Gabriel Avelino Sampedro, Vincent Karovič, and Tariq Naeem. SXAD: Shapely eXplainable AI-based anomaly detection using log data. *IEEE Access*, 12:95659–95672, 2024.

[59] Abhirup Dikshit and Biswajeet Pradhan. Interpretable and explainable AI (XAI) model for spatial drought prediction. *Science of The Total Environment*, 801:149797, 2021.

[60] Pushpa Singh, Narendra Singh, Krishna Kant Singh, and Akansha Singh. *Chapter 5 - Diagnosing of disease using machine learning*. Academic Press, 2021.

[61] Fei Tony Liu, Kai Ming Ting, and Zhi-Hua Zhou. Isolation forest. In *2008 Eighth IEEE International Conference on Data Mining*, pages 413–422, 2008.

[62] Peter J. Rousseeuw. Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics*, 20:53–65, 1987.

[63] Haojie Ji, Liyong Wang, Hongmao Qin, Yinghui Wang, Junjie Zhang, and Biao Chen. In-vehicle network injection attacks detection based on feature selection and classification. *Automotive Innovation*, 7(1):138–149, 2024.

VELJKO VUČINIĆ is currently a Research Associate at the Institute for Information Processing Technology (ITIV), Karlsruhe Institute for Technology (KIT), since November 2022, and a Software Engineer in the Research and Development department in the company RA Consulting GmbH since October 2022. He holds a Master's and a Bachelor's degree in Mechanical Engineering with a specialization in Control Engineering from the University of Belgrade, which he completed in 2022 and 2020, respectively. His master's thesis covered the topic of intelligent control of braking systems within the context of commercial vehicles under the guidance of Prof. Dr.-Ing. Dragan Aleksendrić. Veljko is a PhD student at KIT, pursuing his research interests of system engineering, diagnostic systems, and AI applications for electric vehicles, under the mentorship of Prof. Dr.-Ing. Eric Sax.

LUCA SEIDEL is a Research Associate at the Institut für Technik der Informationsverarbeitung (ITIV), Karlsruhe Institute of Technology (KIT), where he has been working since February 2023. He holds a Bachelor's and a Master's degree in Electrical Engineering and Information Technology with a specialization in System Engineering from KIT, completed in 2020 and 2022, respectively. His master's thesis focused on the relative positioning of traffic participants in the context of urban platooning, under the supervision of Prof. Dr.-Ing. Eric Sax. Luca is a PhD student at KIT, pursuing his research interests in context-based uncertainty monitoring to enable self-healing processes in the domain of automated driving, also under the mentorship of Prof. Dr.-Ing. Eric Sax.

NIKOLA LUKEŽIĆ received the B.Eng. in Electrical Engineering from the Hochschule Karlsruhe in 2019 and M.S in Electrical Engineering and Information Technology from the Karlsruhe Institute of Technology in 2022, where he is currently pursuing the Ph.D. degree with the Institute for Information Processing Technologies in the area of vehicle to everything communication.

FRANK HANTSCHEL is currently team leader in the Research and Development department in the company RA Consulting GmbH. He holds a Phd in Physics, which was achieved in the field of Quantum Physics at the Institute of Theoretical Physics of the Ruprecht-Karls University of Heidelberg in 2015, following his diploma in Physics, which was awarded in 2010 at the same place.

Dr. Frank Hantschel started to work at RA Consulting GmbH in 2015 as a Software Engineer, implementing a tool for the measurement and calibration of vehicles. In 2018, he joined the Research and Development department as a Project Manager, organising research projects in the field of autonomous driving, like KIsSME and RepliCar, besides developing automotive standards by being part of ASAM standardization groups. Since 2020, he works as a Team Leader in addition to his other tasks.

THOMAS KOTSCHENREUTHER is leading the department for Research and Development at RA Consulting GmbH, Bruchsal. After finishing his Diploma in Computer Science in 2001 at the University of Karlsruhe (now KIT), he started as a Research Assistant at the FZI Forschungszentrum Informatik in the field of embedded Systems and model-based development (ESM). In 2008, he started working for RA Consulting GmbH as a software developer with a focus on embedded development and research projects, which developed into its own department at RA Consulting in 2018. He also participates in standardisation groups of ASAM e.V. and contributed to EU or national research projects, like e.g. ASIMOV and Real Driving Validation (RDV).

DRAGAN ALEKSENDRIĆ is a Full Professor at the Automotive Department, University of Belgrade Faculty of Mechanical Engineering and Head of Laboratory for motor vehicles and trailers safety - LaBMV. He is also a Leading Expert of the Republic of Serbia for braking systems and running gears. He received his Dipl. Ing., Master's, and Dr.-Ing. degrees from the University of Belgrade Faculty of Mechanical Engineering, in 1996, 2000, and 2007, respectively. His research interests include braking systems, friction materials in brakes, system engineering, vehicle design and maintenance, intelligent systems, and machine and deep learning.

ERIC SAX is the Head of the Institute for Information Processing Technology (ITIV) and Dean of the KIT Department of Electrical Engineering and Information Technology (ETIT) at Karlsruhe Institute of Technology (KIT). He is also a member of the KIT Senate, spokesperson for the “Zentrum Mobilitätssysteme”, Director of the “Innovations-Campus Mobilität”, Director at the Forschungszentrum Informatik, and Director of the International Department for the business school of Hector School. A tight link to industry derives from his previous roles as Head of E/E at Daimler Buses (2009–2014) and Head of test engineering at the MBtech Group (2002-2009). He earned his Diplom and Ph.D. degrees at the University of Karlsruhe in 1993 and 1999, respectively. His research interests include processes, methods, and tools in systems engineering, data-driven and service-oriented architectures, and the application of machine learning.

APPENDIX I. LIST OF OBD SIGNALS INCLUDED IN THE VERIFICATION

TABLE 5: List of OBD PIDs considered in the Engine Off and Engine on scenarios.

PID	Name	Description	Eng. Off	Eng. On
01	MIL	Malfunction Indicator Lamp Status	✓	✓
04	LOAD_PCT	Calculated LOAD Value	✓	✓
05	ECT	Engine Coolant Temperature	✓	✓
06	SHRTFT1	Short Term Fuel Trim - Bank 1	✓	✓
07	LONGFT1	Long Term Fuel Trim – Bank 1	✓	✓
0B	MAP	Intake Manifold Absolute Pressure	✓	✓
0C	RPM	Engine RPM	✓	✓
OD	VSS	Vehicle Speed Sensor	✓	✓
OE	SPARKADV	Ignition Timing Advance for #1 Cylinder	✓	✓
0F	IAT	Intake Air Temperature	✓	✓
10	MAF	Air Flow Rate	✓	✓
11	TP	Absolute Throttle Position	✓	✓
15	O2Sv12	Oxygen Sensor Output Voltage	✓	✓
15	SHRTFT12	Oxygen Sensor 2 Short term fuel trim	✓	✓
1C	OBDSUP	OBD requirements of vehicle	✓	✓
1F	RUNTM	Time Since Engine Start	✓	✓
21	MIL_DIST	Distance Traveled While MIL is Activated	✓	✓
23	FRP	Fuel Rail Pressure	✓	✓
24	O2SV11	Oxygen Sensor Voltage - Bank 1, Sensor 1	✓	✓
2E	EVAP_PCT	Commanded Evaporative Purge	✓	✓
2F	FLI	Fuel Level Input	✓	✓
30	WARM_UPS	Number of warm-ups since DTCs cleared	✓	✓
31	CLR_DIST	Distance traveled since DTCs cleared	✓	✓
33	BARO	Barometric Pressure	✓	✓
34	LAMBDA11	Equivalence Ration - Bank 1, Sensor 1	✓	✓
34	O2Sc11	Oxygen Sensor Current - Bank 1, Sensor 1	✓	✓
3C	CATEMP11	Catalyst temperature Bank 1 Sensor 1	✓	✓
42	VPWR	Control module voltage	✓	✓
43	LOAD_ABS	Absolute Load Value	✓	✓
44	LAMBDA	Fuel/Air Commanded Equivalence Ratio	✓	✓
45	TP_R	Relative Throttle Position	✓	✓
46	AAT	Ambient air temperature	✓	✓
47	TP_B	Absolute Throttle Position B	✓	✓
49	APP_D	Accelerator Pedal Position D	✓	✓
4A	APP_E	Accelerator Pedal Position E	✓	✓
4C	TAC_PCT	Commanded Throttle Actuator Control	✓	✓
53	EVAP_VPA	Absolute Evap System Vapor Pressure	✓	✓
56	LGSO2FT1	Long Term Secondary O2 Sensor Fuel Trim	✓	✓
5C	EOT	Engine Oil Temperature	✓	✓
5E	FUEL_RATE	Engine Fuel Rate	✓	✓
62	TQ_ACT	Actual Engine - Percent Torque	✓	✓
63	TQ_REF	Engine Reference Torque	✓	✓
67	ECT_1	Engine Coolant Temperature 1	✓	✓
67	ECT_2	Engine Coolant Temperature 2	✓	✓
68	IAT_11	Intake Air Temperature - Bank 1, Sensor 1	✓	✓
68	IAT_12	Intake Air Temperature - Bank 1, Sensor 2	✓	✓
73	EP_1	Exhaust Pressure Sensor Bank 1	✓	✓
78	EGT11	Exhaust Gas Temperature - Bank 1, Sensor 1	✓	✓
78	EGT12	Exhaust Gas Temperature - Bank 1, Sensor 2	✓	✓
78	EGT13	Exhaust Gas Temperature - Bank 1, Sensor 3	✓	**x**
78	EGT14	Exhaust Gas Temperature - Bank 1, Sensor 4	✓	**x**
8E	TQ_FR	Engine Friction - Percent Torque	✓	✓
9D	ENG_FUEL_RATE	Engine Fuel Rate	✓	✓
9D	VEH_FUEL_RATE	Vehicle Fuel Rate	✓	✓
9E	EXH_RATE	Engine Exhaust Flow Rate	✓	✓
A4	GEAR_ACT	Actual Transmission Gear	✓	✓
A6	ODO	Odometer	✓	✓

APPENDIX II. VISUALIZATION OF THE OBD SIGNALS USED FOR ANOMALIES.

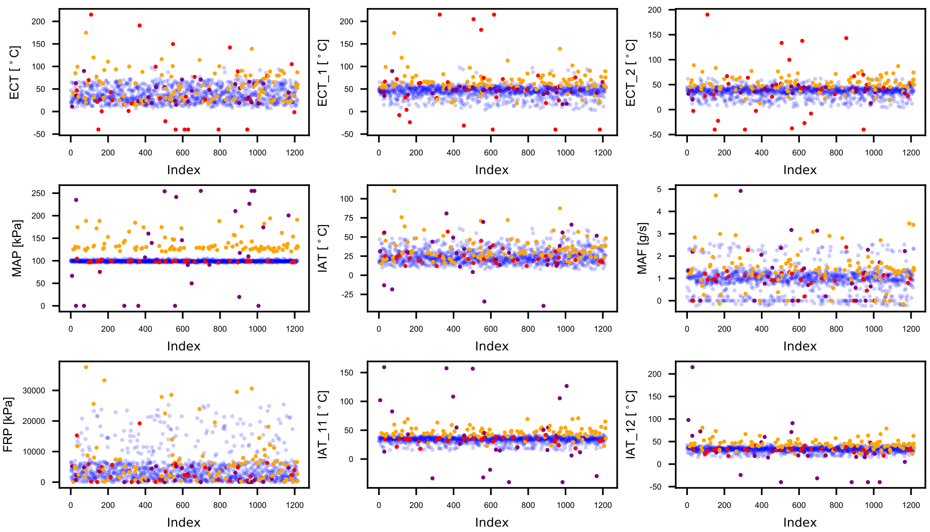


FIGURE 10: Values of key PID signals of normal and anomaly snapshots in the Engine Off scenario, as described in Table 2. Color code is as follows: transparent blue - normal data, orange - system performance drop anomaly, red - engine coolant system anomaly, purple - fuel system anomaly.

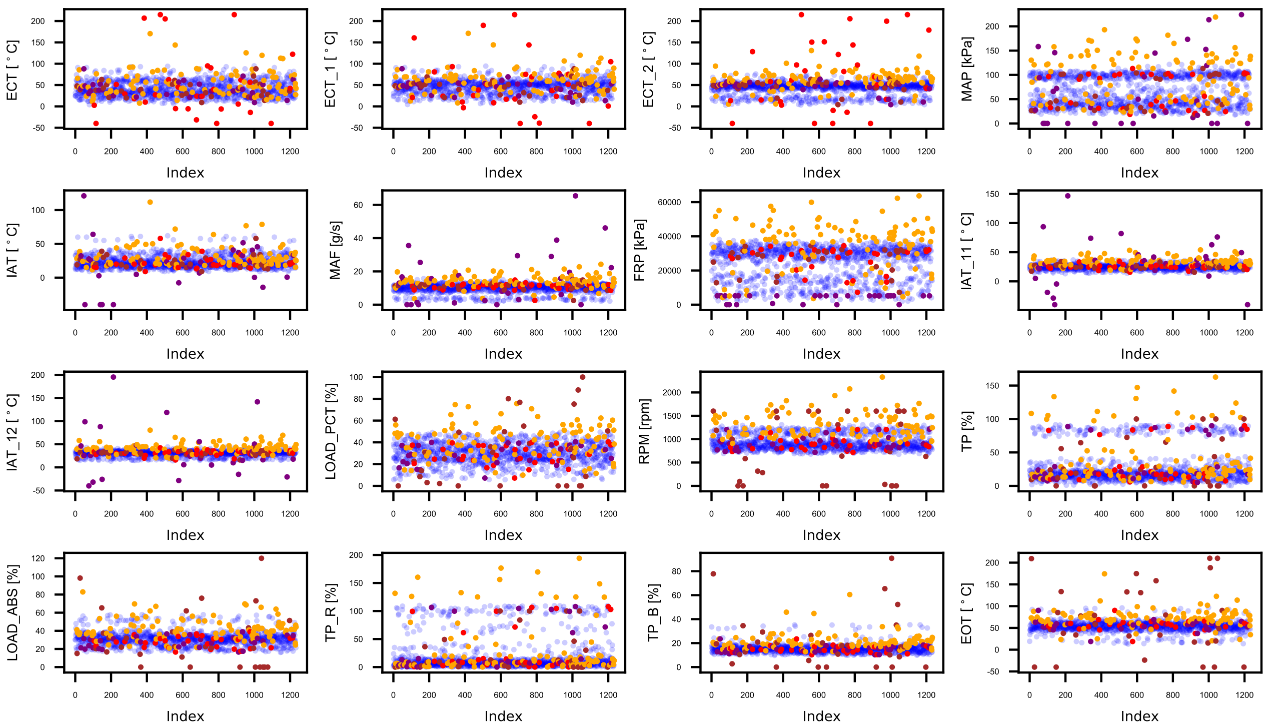


FIGURE 11: Values of key PID signals of normal and anomaly snapshots in the Engine On scenario, as described in Table 2. Color code is as follows: transparent blue - normal data, orange - system performance drop anomaly, red - engine coolant system anomaly, purple - fuel system anomaly, brown - engine ECU tampering.

APPENDIX III. EXPERIMENTAL STUDY WITH ISOLATION FOREST (IF)

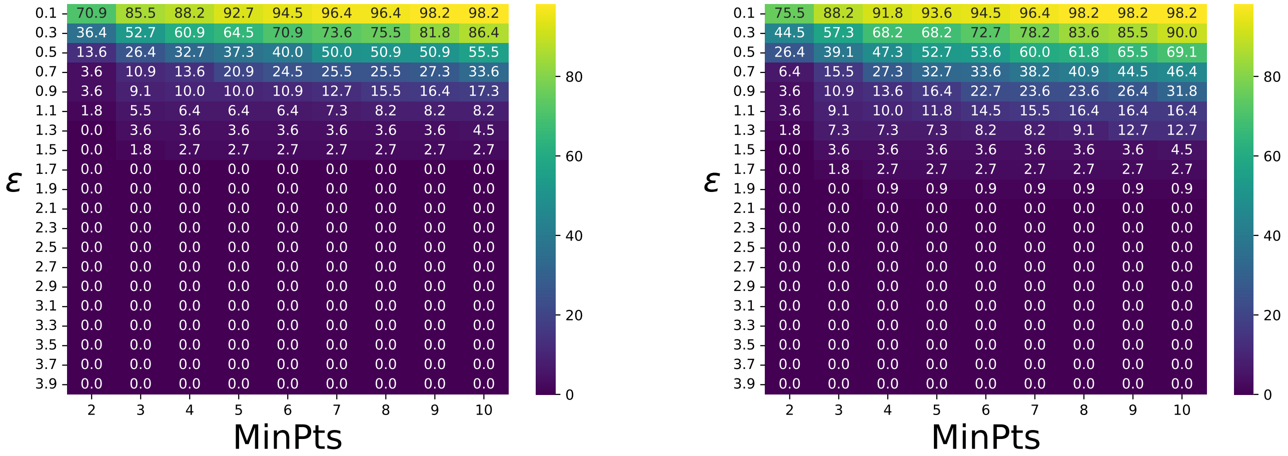
TABLE 6: Confusion matrices results for overall and per-anomaly type of anomaly detection using IF for engine off scenario. The IF is created with the following parameters: number of trees - 100, features per tree - all, data points per tree - 256.

Anomaly group	**TP**	**TN**	**FP**	**FN**
All	55.00%	98.55%	1.45%	45.00%
Performance drop	81.00%	98.55%	1.45%	19.00%
Engine coolant	6.67%	98.55%	1.45%	92.33%
Fuel system	16.67%	98.55%	1.45%	83.33%

TABLE 7: Confusion matrices results for overall and per-anomaly type of anomaly detection using IF for engine on scenario. The IF is created with the following parameters: number of trees - 100, features per tree - all, data points per tree - 256.

Anomaly group	**TP**	**TN**	**FP**	**FN**
All	46.32%	98.08%	1.92%	53.68%
Performance drop	69.00%	98.08%	1.92%	31.00%
Engine coolant	3.33%	98.08%	1.92%	96.67%
Fuel system	33.33%	98.08%	1.92%	66.67%
Tampering	26.67%	98.08%	1.92%	73.33%

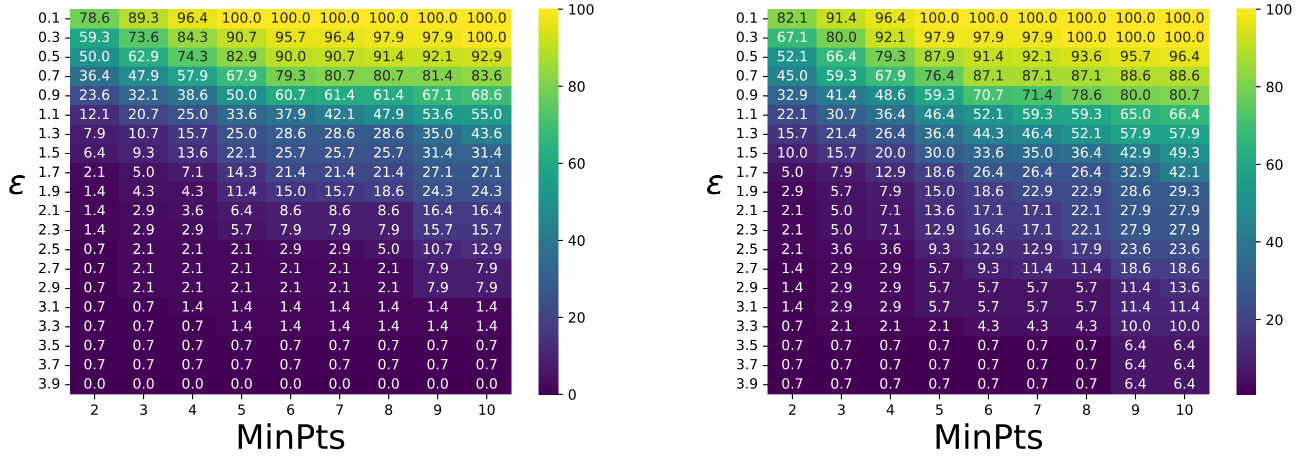
APPENDIX IV. SENSITIVITY ANALYSIS OF THE DBSCAN HYPERPARAMETERS



(a) Influence of ϵ and *MinPts* on sensitivity for Euclidean distance

(b) Influence of ϵ and *MinPts* on sensitivity for Manhattan distance

FIGURE 12: Comparison of sensitivity behavior with variable ϵ and *MinPts* for different distance metrics in the engine off application scenario.



(a) Influence of ϵ and *MinPts* on sensitivity for Euclidean distance

(b) Influence of ϵ and *MinPts* on sensitivity for Manhattan distance

FIGURE 13: Comparison of sensitivity behavior with variable ϵ and *MinPts* for different distance metrics in the engine on application scenario.