NIST Internal Report
NIST IR 8534

Feature Description for Assessing Autonomous Vehicle Performance

Hadhoum Hajjaj
Thoshitha T. Gamage
Edward R. Griffor
Thomas P. Roth
Wenqi W. Guo

This publication is available free of charge from:
https://doi.org/10.6028/NIST.IR.8534

NATIONAL INSTITUTE OF STANDARDS AND TECHNOLOGY
U.S. DEPARTMENT OF COMMERCE

NIST Internal Report
NIST IR 8534

Feature Description for Assessing Autonomous Vehicle Performance

Hadhoum Hajjaj
Thoshitha T. Gamage
Edward R. Griffor
Thomas P. Roth
Wenqi W. Guo
Smart Connected Systems Division
Communication Technology Laboratory

This publication is available free of charge from:
https://doi.org/10.6028/NIST.IR.8534

September 2024

U.S. Department of Commerce
Gina M. Raimondo, Secretary

National Institute of Standards and Technology
Laurie E. Locascio, NIST Director and Under Secretary of Commerce for Standards and Technology

Publication History
Approved by the NIST Editorial Review Board on 2024-08-22

How to cite this NIST Technical Series Publication:
Hajjaj H, Gamage TT, Griffor ER, Roth TP, Guo WW (2024) Feature Description for Assessing Autonomous Vehicle Performance. (National Institute of Standards and Technology, Gaithersburg, MD), NIST IR 8534. https://doi.org/10.6028/NIST.IR.8534

Author ORCID iDs
Hadhoum Hajjaj: 0009-0001-4237-4874
Thoshitha T. Gamage: 0009-0002-2797-2929
Edward R. Griffor: 0000-0001-5241-7551
Thomas P. Roth: 0000-0002-9986-7784
Wenqi W. Guo: 0000-0002-9712-8363

Contact Information
hadhoum.hajjaj@nist.gov
100 Bureau Dr, Gaithersburg, Maryland, 20899

Abstract

This technical report presents a structured methodology for describing and assessing automated vehicle (AV) features, with a focus on ensuring consistent and reproducible testing across various scenarios. The document delineates a systematic approach to integrating operational design domains (ODD), levels of automation, and vehicle behaviors into AV feature descriptions. Through the development of a feature description framework and a corresponding testing template, the report provides guidance on how to measure and validate the performance of AV features under controlled and variable conditions. Key elements of the methodology include detailed behavior specifications, performance metrics, and scenario-based testing, which are essential for evaluating the effectiveness and safety of AV technologies. The report also explores the application of simulation-based testing to assess AV feature behaviors in a dynamic and virtual environment, providing a critical link between theoretical frameworks and practical implementation. The outcomes aim to enhance communication among manufacturers, regulators, and other stakeholders, promoting a unified understanding of AV capabilities and compliance.

Keywords

Autonomous Vehicle (AV); Operational Design Domain (ODD); Operating Envelope Specification (OES); behavior specification; feature description; feature specification; scenario-based testing; feature performance assessment; co-simulation.

Table of Contents

List of Tables

List of Figures

Author Contributions

Hadhoum Hajjaj: Conceptualization, Methodology Development, and Writing. **Thoshitha T. Gamage**: Simulation Execution, Writing, Reviewing, and Editing. **Edward R. Griffor**: Co-development of Methodology, Reviewing, and Validating Terminology. **Thomas P. Roth**: Supervision, Reviewing, and Editing. **Wenqi W. Guo**: Reviewing, Editing.

Executive Summary

The Autonomous Vehicle (AV) technology landscape is rapidly evolving, presenting a persistent challenge in assessing the performance of components that automate the driving tasks traditionally performed by human operators. These automated tasks, referred to as AV 'features', are frequently described with inconsistent technical documentation that lacks the standardization crucial for effective assessment of task performance under real-world conditions. Feature performance is assessed by how well a feature executes its designated function, whether it improves road safety, vehicle reliability, or occupant comfort, and whether the feature performs consistently under various operating conditions.

Establishing a consistent approach to feature description is a prerequisite for developing a comprehensive methodology for assessing AV feature performance. While features are prominent in existing AV literature, they are often discussed in plain language and lack comprehensive and mathematical descriptions that list the factors relevant to feature performance. Features such as adaptive cruise control (ACC) and automatic emergency braking (AEB) are typically defined by their functions; however, the parameters describing their operational effectiveness, vehicle sub-systems involved, behaviors and other parameters are not uniformly detailed. The lack of a robust and uniform framework for feature description could lead to variability in performance and potentially uneven safety outcomes across models and brands. The need for detailed, standardized feature descriptions is therefore not just a concern of regulatory compliance, but a fundamental prerequisite for advancing AV technology in a manner that assures reliability and safety.

Recognizing these gaps in the feature description, this technical report introduces a feature description framework and associated test plan template designed for AV feature testing. This template acts as a nexus, systematically integrating the technical parameters of features into a format usable for feature-level performance testing. The proposed framework aims to facilitate a more consistent and reliable assessment of AV technologies and provide manufacturers, regulatory authorities, and the research community with a foundation for a common understanding of feature performance requirements.

1. Introduction and Background

1.1. Background

Autonomous Vehicles (AVs) hold the potential to significantly transform the future of automotive mobility, providing benefits such as enhanced fuel efficiency, reduced traffic accidents, reduced road congestion, and improved accessibility [2]. Starting with innovations like cruise control, the automotive industry has seen significant milestones including the advent of electronic fuel injection systems and GPS (Global Positioning System) navigation integration. These technological advancements have not only improved vehicle control accuracy but also introduced vehicle connectivity and data-driven navigation. Particularly, Automated Driving Assistance Systems (ADAS) such as automatic emergency braking and lane-departure warning have significantly advanced vehicle automation leading to a decrease in vehicle crash rates [3, 4]. These innovations continue to drive automotive research and development, pushing the boundaries of what automated systems can achieve in modern transportation.

As AV technologies evolve, there is a growing need to understand how to best integrate these new technologies with existing road traffic and infrastructure while ensuring the safety and maintaining the trust of people in the driving environment. AV-related legislation has gained momentum across the United States since 2012, with legislation considered in at least 41 states. In 2018 alone, 15 states enacted 18 AV-related bills, demonstrating the growing need for comprehensive regulatory frameworks and a proactive approach to integrating AVs into public roads and transport systems [5].

SAE International categorizes AV features from Level 0 to Level 5 based on their functionality and degree of driving automation as shown in Fig. 1 [1]. Each level of automation has specific requirements that are crucial for the technical design and regulatory compliance of AV features. Regulatory frameworks often align with these levels, influencing how vehicles are certified and brought to market [6]. As the automotive industry progresses toward features with higher levels of automation, there is an increasing need for unified methodologies in describing, testing, and validating AV features. This demand is not only driven by the aim for safety, reliability, and market readiness but also by the requirement to establish universal approaches that guide both industry and regulatory bodies.

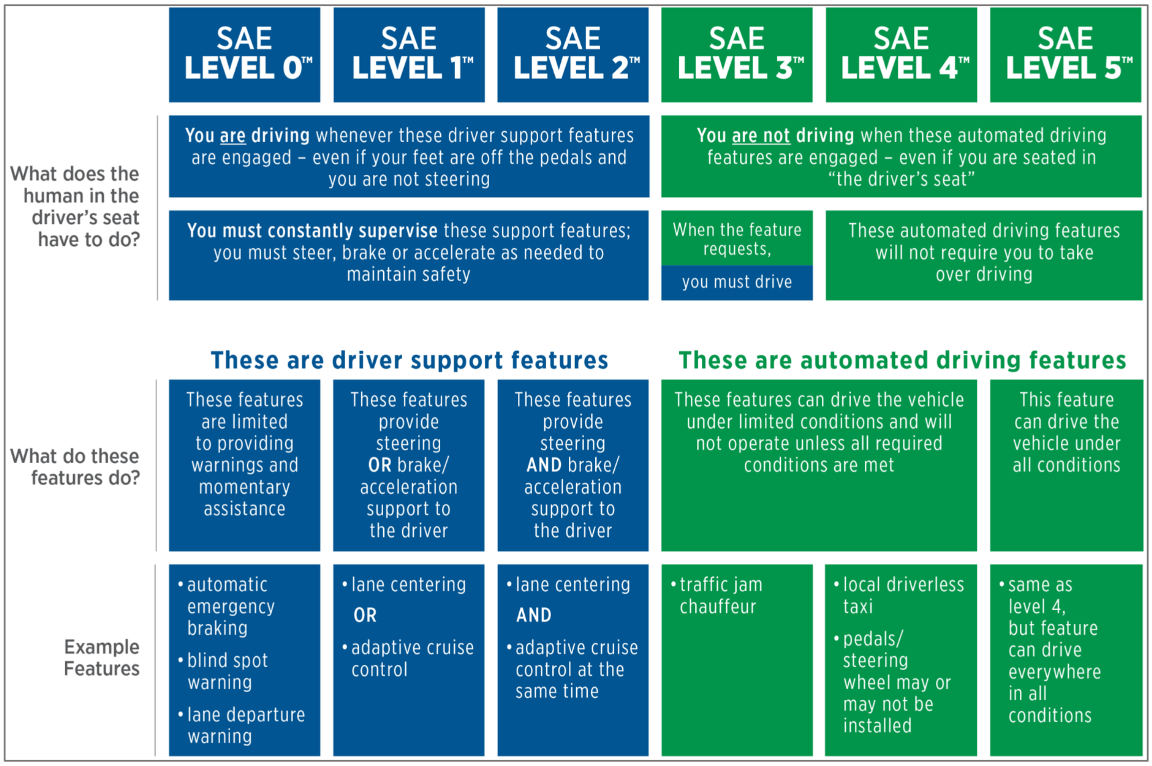

Fig. 1. SAE Levels of Driving Automation [1]

1.2. Terms and Definitions

The AV field has a variety of terms related to features. This terminology comes from various sources, including government institutions, Standards Development Organizations (SDOs), and the automotive industry. There is some, but not complete, consensus on their usage. Table 1 summarizes the terms and definitions adopted in this report. While the term feature from this table is capitalized in the context of SAE standards, this report will use a lowercase feature for readability.

As an example of using these terms, a vehicle's adaptive cruise control maintains a preset distance to vehicles ahead while striving to maintain a preset speed. Adaptive cruise control is a **feature** that is considered Level 1 automation as the driver still needs to monitor other aspects of driving, like steering and braking. This feature operates within a defined **ODD**, which for most adaptive cruise systems includes highways or well-marked roads with clear lane markings. The **OES** for adaptive cruise control might specify that it can function properly on lanes at least 10 ft wide in clear conditions with less than 10 mm/h of rainfall. The **behaviors** of adaptive cruise control include maintaining a speed set by the driver, responding to other vehicles, and maintaining a lane. This relates to the **DDT**, where the system takes over longitudinal vehicle motion control to maintain its speed and lateral vehicle motion control to maintain its lane.

Table 1. Terms and Definitions

Term	Definition
Dynamic Driving Task (DDT) [7]	All the real-time operational and tactical functions required to operate a vehicle in on-road traffic, excluding the strategic functions such as trip scheduling and selection of destinations and waypoints, and including without limitation: lateral and longitudinal vehicle motion control, and monitoring the driving environment.
Operational Design Domain (ODD) [1]	The operating conditions under which a given driving automation system or feature thereof is specifically designed to function, including, but not limited to, environmental, geographical, and time-of-day restrictions, and/or the requisite presence or absence of certain traffic or roadway characteristics.
Operating Envelope Specification (OES) [8]	A structured description of the operating environment for driving, usable for formal reasoning about that environment in testing and certification applications and in real-time driving conditions. The OES provides measurement for ODD factors and other elements of the operating environment.
Autonomous Vehicle (AV)	A vehicle that has the capability to perform the DDT under some or all conditions, elaborated in the ODD of the automated driving feature. Some refer to AVs as Automated Vehicles or Automated Driving Systems Dedicated Vehicles (ADS-DV).
Feature [1]	A level 1 through 5 driving automation system's design-specific functionality at a given level of driving automation within a particular ODD, if applicable.
Behavior [7]	A specific goal-oriented action directed by an engaged Automated Driving System (ADS) in the process of completing the DDT or DDT fallback within the ODD (if applicable) at a variety of timescales.
Scenario	A description of the temporal development through several consecutive events. See [9–12], for alternative definitions.
Event	A specific occurrence, action, or state within a scenario. This could include changes in the environment, actions by other road users, or vehicle responses.

A **scenario** involving adaptive cruise control might include two vehicles on a highway traveling at a steady speed. When the lead vehicle begins to slow down, the following vehicle, equipped with adaptive cruise control, must automatically adjust its speed to maintain a safe distance. Within this scenario, the moment the leading vehicle decelerates acts as a triggering **event**, prompting the adaptive cruise control system to respond accordingly.

1.3. Methodology and Section Outline

To develop the AV feature description, the following steps were conducted:

- Identify the key elements relevant to documenting an AV feature
- Identify the principal factors relevant for testing an AV feature
- Develop the AV feature description and testing framework, incorporating the identified elements and factors into a structured template
- Develop a simulated case study to demonstrate the usage of the framework to evaluate feature performance

The remainder of the report is structured as follows:

- **Section 2** presents existing methodologies for assessing the performance of AV features. It explores various approaches and highlights their implications in the context of AV systems.
- **Section 3** describes the architecture of the AV feature description framework, detailing each component and its role within the framework. It uses a case study on the Automatic Emergency Braking (AEB) feature as an illustrative example.
- **Section 4** implements a use case in simulation to demonstrate the application of the framework. It illustrates the practical deployment of the test plan and provides insights into the performance assessment of AV features within a simulation environment.

2. AV Performance Assessment

As the level of automation in vehicles increases, evaluating the performance of AVs becomes more complex and challenging. This section highlights existing approaches to AV performance evaluation.

One such approach is Naturalistic Field Operational Tests (N-FOTs) [13], which involve deploying multiple vehicles to drive under natural conditions. This method allows for direct observation of AVs in natural settings but requires significant time and budget [14, 15], and offers low exposure to safety-critical scenarios [16].

The European New Car Assessment Program (Euro NCAP) is widely recognized for assessing vehicle performance through safety reports based on performance in crash tests [17]. To keep up with advancements in vehicle safety technology, Euro NCAP has introduced a roadmap for 2025 [18]. This roadmap expands the scope of assessments to include Advanced Driver Assistance Systems (ADAS), addresses driver monitoring to mitigate issues such as distraction and fatigue, and transitions to a more scenario-based assessment approach. Additionally, it incorporates the use of virtual testing for a more robust evaluation and includes advancements in technologies such as AEB and Autonomous Emergency Steering to address various crash scenarios and improve overall vehicle safety.

Another common method is the use of test matrix scenarios, which involve defining a suite of scenarios to test each case systematically. This approach is advantageous due to the low cost, high controllability, and repeatability of scenarios. However, vehicles may perform well in these predetermined scenarios but not necessarily under real-world conditions. These scenarios often derive from collision databases, thus limiting the range of influencing factors considered [14].

Worst-case scenario testing is another method utilized to evaluate the performance of AV features [19]. Typically, these scenarios are generated from crash databases to identify critical system failures that can be addressed to enhance driving safety [20]. Researchers also employ techniques such as defining criticality metrics and filtering accidents derived from naturalistic driving data [21] to identify the worst-case scenarios. However, this method does not guarantee comprehensive coverage of all potential worst cases nor does it reflect normal operational conditions. Additionally, it is usually difficult to reconstruct the context of traffic accidents to create these scenarios.

Critical-scenario testing [22] involves generating relevant scenarios. One of the methods used to identify critical scenarios is proposed in [23] and is called Testing Scenario Library Generation (TSLG). This method includes the definition of decision variables, the design of performance metrics, and the generation of a testing scenario library. Critical scenarios are identified by evaluating scenario criticality, which combines maneuver challenges and exposure frequency derived from naturalistic driving data. Optimization methods are then employed to efficiently identify these critical scenarios. However, despite generating relevant scenarios, critical-scenario testing shares the limitations of worst-case methodologies in terms of coverage and completeness [24].

Another methodology for evaluating the performance of autonomous vehicles integrates both scenario-based and functionality-based testing methods. Scenario-based testing involves evaluating the vehicle's performance in various driving environments and situations, ensuring it can handle real-world conditions. Functionality-based testing, on the other hand, focuses on assessing specific system capabilities, such as object detection, decision-making, and vehicle control functions. To address the limitations of each method when used in isolation, a hybrid approach has been proposed [25].

While scenario-based testing is crucial for assessing performance under diverse conditions, it is impractical to anticipate every possible scenario an AV might encounter. Relying solely on specific edge-case, critical-case, or worst-case scenarios for feature testing and performance assessment may not provide a comprehensive view of an AV feature's capabilities and limitations across typical operating conditions. While edge cases are essential for pushing the feature to its limits, they may offer limited insights into the overall objectives of the feature. This does not reduce the importance of edge cases but suggests a balanced approach where structured feature descriptions guide the selection of the most relevant and impactful scenarios for the feature.

AVs can be conceptualized as a constellation of features, each representing a specific functionality that collectively defines the vehicle's automated driving capabilities. Describing these features accurately and testing and evaluating their behaviors within targeted scenarios can significantly enhance the approach to assessing the functionality of AVs. This approach is modular and not only aims to assure the reliability of individual features but also emphasizes the overall functional integration and testing of automated systems.

Overall, the core contribution of this report is the proposal of a feature description framework for AV performance evaluation, distinguishing it from existing methodologies that predominantly rely on scenario generation.

3. AV Feature Description for Performance Assessment

This section introduces the AV feature description framework, starting with an overview of the framework content and then providing a detailed explanation of the individual fields. An AEB use case is used as an example to demonstrate how the framework can be populated to describe a specific AV feature.

3.1. Overview of the AV Feature Description Framework

Figure 2 depicts the proposed AV feature description and testing framework template, which provides a visual representation of the entire process. This template draws from the corresponding ODD and OES to systematically and comprehensively describe the feature, its associated systems and system technologies, the behaviors, and the performance. Moreover, the scenarios, derived from these foundational components, serve as practical applications to test the feature.

Four key components have been identified that constitute the feature description framework. These categories are shown in Fig. 3 and include the Feature Definition, the Feature Behaviors, the Feature Performance, and the Relevant Scenarios.

The proposed methodology integrates established elements such as ODD, behavior, and levels of automation. It also includes measurable elements such as OES, behavioral specifications, and influences to give features a tangible dimension and aid in performance

<table>
<tr><th colspan="4">1. Feature</th></tr>
<tr><td colspan="4"></td></tr>
<tr><th colspan="4">2. SAE Automation Level</th></tr>
<tr><td colspan="4"></td></tr>
<tr><th colspan="2">3. Operational Design Domain (ODD)</th><th colspan="2">4. Operating Envelope Specification (OES)</th></tr>
<tr><td colspan="2"></td><td colspan="2"></td></tr>
<tr><th colspan="4">5. Feature Description</th></tr>
<tr><td colspan="4"></td></tr>
<tr><th colspan="4">6. Systems and System Technologies</th></tr>
<tr><th colspan="2">6.1. Systems</th><th colspan="2">6.2. System Technologies</th></tr>
<tr><td colspan="2"></td><td colspan="2"></td></tr>
<tr><th colspan="4">7. Behaviors Assessment</th></tr>
<tr><th colspan="2">7.1. Behavior Name</th><th>7.2. Behavior Description</th><th>7.3. Behavior Specification</th></tr>
<tr><td colspan="2"></td><td></td><td></td></tr>
<tr><th colspan="4">8. Feature Performance</th></tr>
<tr><th colspan="2">8.1. Feature Specification</th><th colspan="2">8.2. Performance Metric</th></tr>
<tr><td colspan="2"></td><td colspan="2"></td></tr>
<tr><th colspan="4">9. Scenarios</th></tr>
<tr><th rowspan="6">9.1 Scenario 1: Scenario Name</th><th colspan="3">9.1.1. Scenario and Context</th></tr>
<tr><th>9.1.1.1. Relevant ODD Factors</th><th>9.1.1.2. Relevant OES Measurements</th><th>9.1.1.3. Events</th></tr>
<tr><td></td><td></td><td></td></tr>
<tr><th colspan="3">9.1.2. Influences</th></tr>
<tr><th>9.1.2.1. Influence Name</th><th>9.1.2.2. Description</th><th>9.1.2.3. Influence Specification</th></tr>
<tr><td></td><td></td><td></td></tr>
<tr><td colspan="4">⋮</td></tr>
<tr><th rowspan="6">9.X Scenario X: Scenario Name</th><th colspan="3">9.X.1. Scenario and Context</th></tr>
<tr><th>9.X.1.1. Relevant ODD Factors</th><th>9.X.1.2. Relevant OES Measurements</th><th>9.X.1.3. Events</th></tr>
<tr><td></td><td></td><td></td></tr>
<tr><th colspan="3">9.X.2. Influences</th></tr>
<tr><th>9.X.2.1. Influence Name</th><th>9.X.2.2. Description</th><th>9.X.2.3. Influence Specification</th></tr>
<tr><td></td><td></td><td></td></tr>
</table>

Fig. 2. The AV Feature Description and Testing Framework Template

assessment. The template is available as an Excel spreadsheet, which can be accessed at https://doi.org/10.18434/mds2-3270. Each instance of the template describes a single AV feature, with multiple scenarios detailed that capture the varied conditions under which the feature's performance is to be evaluated. This structured approach allows for the systematic examination of each feature, while also providing the flexibility to assess a comprehensive range of scenarios relevant to that feature.

3.2. Key Components of AV Feature Description and Testing Framework

This subsection describes each element of the proposed framework highlighted in Fig. 3 using an AEB use case as a practical example.

An AEB feature is designed to automatically stop a vehicle before it collides with an object in its path, primarily using sensor data from radar and cameras. For this example, the

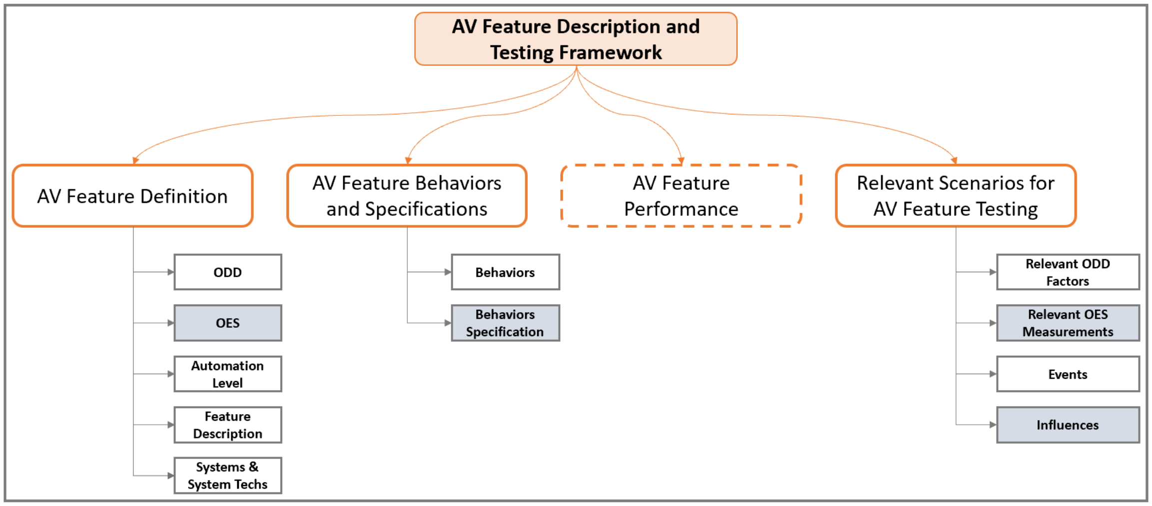

Fig. 3. AV Feature Description and Testing Framework Components

AEB feature has been enhanced to include vehicle-to-vehicle (V2V) communications that enable it to detect obstacles using external communications in addition to its onboard sensors. The same AEB feature is implemented in simulation in Sec. 4.

3.2.1. AV Feature Definition

The feature description framework begins with an "AV Feature Definition". This includes the feature name, its level of automation according to SAE standards, and a feature description that describes the purpose of the feature, its functional objectives, and the conditions that trigger its activation.

Following SAE terminology, a feature is defined not just by its intended function but also through its ODD, which specifies the environment in which the feature is designed to operate. However, there is no standardized method to describe an ODD, and most are loosely defined in plain language such as "on a highway in clear weather". The OES, introduced by NIST, is crucial in this context as it quantifies the operational concept represented by the ODD, offering a structured and measurable framework for understanding it.

Lastly, the AV Feature Definition lists the systems and system technologies required to implement the feature. In this context, "systems" refers to the integrated set of components and functions that enable a vehicle to perceive its environment, make decisions, communicate with the surrounding environment, and control its movement, whether autonomously or with human intervention. And "system technology" refers to the fundamental technological domain that supports these systems, such as communication and perception, as discussed in [26] and [27] respectively. These system technologies are interdependent and collectively contribute to the feature's performance, with each AV system potentially integrating one or more of these technologies.

Example: AEB Feature Definition – Thorn et al. [6] provides a taxonomy to describe the ODD of the AEB feature. This description is intended to be descriptive rather than normative. It organizes the ODD elements into a hierarchical structure of categories and subcategories. The OES for AEB quantifies these ODD elements, establishing measurable parameters as detailed in [8]. The other fields for the AEB definition are presented in Fig. 4.

1. Feature	
Automatic Emergency Braking (AEB)	
2. SAE Automation Level	
Level 1	
3. Operational Design Domain (ODD)	**4. Operating Envelope Specification (OES)**
DOT HS 812 623	https://doi.org/10.6028/NIST.SP.1900-301
5. Feature Description	
The Automated Emergency Braking (AEB) Feature is designed to avoid collisions by automatically applying the brakes when obstacles are detected ahead, leveraging a combination of sensor data and vehicle-to-vehicle (V2V) communication.	
6. Systems and System Technologies	
6.1. Systems	**6.2. System Technologies**
AV Controller for Brake Monitoring	Perception
V2V Communication	Communication
Object Detection	Perception
Braking System	Mechanical and Control

Fig. 4. AEB Feature Definition

3.2.2. AV Feature Behaviors and Specifications

The second category in the AV feature description framework is "Behaviors and Specifications" which lists all behaviors expected of the AV feature. A partial list of vehicle behaviors has been published by the SAE Automated Vehicle Safety Consortium (AVSC) [7]. The relevant behaviors should be identified based on the AV Feature Definition and the different objectives outlined in the feature description field. However, similar to the ODD, behaviors are described loosely in plain language and NIST recommends a formal, mathematical specification be associated with each behavior to support feature testing and validation.

Each behavior should have one or more specifications that set a measurable and quantifiable standard, defining the exact parameters within which the feature must operate to fulfill the behavior. These specifications set clear performance criteria with thresholds that can be used to evaluate whether a feature achieves its intended functions. The thresholds are derived from the OES, maintaining consistency and alignment with the foundational elements of the feature description.

Example: AEB Feature Behaviors and Specifications – The AEB feature includes two primary behaviors: i) responding to lane obstructions and obstacles, and ii) responding to external communication signals. The inclusion of a response to external communication signals extends beyond [7], aiming to fill a documented gap concerning V2V communication. This adaptation ensures alignment with the AEB feature's description and addresses essential aspects of vehicle interaction not thoroughly explored in existing standards.

Specifications for the AEB behaviors are quantitatively defined to ensure precise evaluation. For the first behavior, the distance between the ego vehicle and the obstacle, *veh_to_obst_distance*, when the vehicle comes to a full stop must be greater than or equal to the minimum safe distance, d_{min}, to be considered sufficient. For the second behavior, the time required to process communication signals, *time_to_process*, must not exceed the maximum allowable time, t_{max}, for effective response. These behaviors and specifications for AEB are summarized in Fig. 5.

> The ego vehicle refers to the primary vehicle under study or control within a given scenario. This vehicle is equipped with autonomous driving capabilities and is the focus of the simulation, where its behaviors, decision-making processes, and interactions with the surrounding environment and other road users are analyzed and tested.

7. Behaviors Assessment		
7.1. Behavior Name	**7.2. Behavior Description**	**7.3. Behavior Specification**
Responding to Lane Obstructions and Obstacles	The vehicle stops with sufficient stopping distance before the obstacle.	$veh\_to\_obst\_distance \geq d_{min}$
Responding to External Communication Signals	The vehicle appropriately responds to messages transmitted from another vehicle through V2V Communication.	$time\_to\_process \leq t_{max}$

Fig. 5. AEB Feature Behaviors and Specifications

3.2.3. AV Feature Performance

The purpose of a test plan for AV features is to evaluate their performance. While this report primarily focuses on describing the features and creating scenarios for testing, the subsequent step of assessing how these features perform is crucial. Performance assessment determines whether each feature conforms to its intended design and functions effectively within the AV system. This document includes a section for performance evaluation, with a comprehensive methodology of feature performance assessment planned for future work. A subsequent effort will present a full examination of performance, providing data on how well each feature meets its targets.

Example: AEB Feature Performance – The evaluation process for AEB feature performance has two primary components: specifications and performance metrics. Specifications comprise one or more criteria, or constraints, that the AEB performance metrics must meet. Performance metrics are the quantifiable measures that assess how well the AEB feature meets these specifications, such as the activation rate of the AEB in relevant scenarios. The adequacy of AEB feature performance is determined by whether these metrics meet or exceed the constraints set by the specifications and is established through rigorous testing and compliance to industry standards, ensuring the system performs reliably and effectively. For an example, see Fig. 6.

8. Feature Performance	
8.1. Feature Specification	**8.2. Performance Metric**
Predefined criteria and constraints that the AEB system must meet.	*Quantifiable measure of performance (e.g., the AEB activation rate). The specification for success could be that this metric must achieve a 95% activation rate when an obstacle is detected within a critical distance or when a message is received through the communication channel.*

Fig. 6. AEB Feature Performance

3.2.4. Relevant Scenarios for AV Feature Testing

The feature description framework emphasizes the generation of meaningful scenarios as pivotal for effectively testing and validating AV features. Factors from the ODD and OES are selectively chosen to directly challenge the behaviors outlined for the AV feature. This targeted approach ensures that the AV feature is assessed under conditions related to its behaviors, allowing for precise testing of its operational capacity. Focusing on scenarios that demonstrate or challenge the feature behavior specifications avoids the creation of numerous irrelevant scenarios.

Furthermore, this element of the framework includes influences that may introduce greater variability into system performance. In this context, influences refer to either external or internal factors or conditions that can impact the performance of the feature. These influences may be included in the AVs ODD, but are not generally characterized quantitatively by the ODD. Influences are incorporated into the test plan, and quantitatively characterized, to assess the sensitivity and resilience of AV feature performance to them. The selection of these influences is deliberate and aligned with the behaviors defined in the previous element of the approach.

Example: A Sample of Scenarios for AEB Feature Testing – The scenarios should be designed to assess the identified behavior specifications and test whether the AEB feature satisfies the objectives introduced in its feature description, which is intended to stop the vehicle using either its perception system or its communication systems. Consequently, two scenarios are designed for the AEB feature: one employing the vehicle's perception system for braking and the other utilizing its communication system.

Effective testing of these scenarios involves selecting ODD factors and OES measurements that stringently challenge the AEB's behaviors. Both scenarios consider ODD factors such as weather conditions, highway terrain, highway speed, and illumination. These factors are varied for both scenarios to methodically evaluate the system's response and effectiveness. These factors are chosen because they represent critical conditions that significantly challenge the behaviors and impact the performance and safety of AEB feature.

Weather conditions, such as rain, can significantly impact sensor accuracy and vehicle control, which is crucial for the AEB system to correctly detect and respond to lane obstructions and obstacles. Variations in highway terrain can challenge the vehicle's stability, ensuring that the AEB feature can effectively operate and apply braking on different surfaces. Highway speed influences the reaction time and braking distance required for safe operation, which is vital for both responding to lane obstructions and accurately interpreting external communication signals. Illumination, including different levels of daylight, affects the visibility and detection capabilities of the vehicle's sensors. These ODD factors are critical in determining how well the AEB feature detects obstacles and interprets communication signals under varying conditions. It is important to note that these are just examples from the entire ODD and can be expanded to include other factors that challenge the feature's behaviors, such as traffic density.

Both scenarios consider two vehicles on a highway, with the ego vehicle equipped with AEB following a lead vehicle. In the first scenario, the lead vehicle suddenly applies heavy braking. This allows for an evaluation of the AEB's response to lane obstructions and obstacles under realistic highway conditions. In the second scenario, the lead vehicle also suddenly applies heavy braking but simultaneously broadcasts a Basic Safety Message (BSM) [28]. This scenario assesses the AEB's ability to respond to the received communication signal.

Regarding the influences applicable to the scenarios, the first scenario incorporates the distance between the two vehicles as a variable influence. This is relevant because the distance between vehicles can significantly affect the AEB's ability to detect and react to sudden braking by the lead vehicle. For example, if the distance is too short between the two vehicles, the ego vehicle may not be able to stop before colliding with the obstacle detected. In the second scenario, both the inter-vehicle distance and the delay in the communication signal (BSM) are varied. The inter-vehicle distance is varied to evaluate how different following distances impact the AEB's response to communication inputs. The communication system may be more efficient in cases of short distances between vehicles, providing quicker response times compared to relying solely on the perception system. The delay in the BSM message is relevant because if there is a delay in communication due to external and/or internal factors, this can impact the response of the ego vehicle. The BSM delay is crucial for assessing the system's robustness in handling communication latency, which is vital for the timely and accurate response to external communication signals. Figure 7 shows how these scenarios are represented in the template for AEB testing.

These scenarios illustrate the process of how to populate the scenario field in the template; a complete test plan would need to consider a variety of ODD factors and influences relevant to the feature behaviors. Figure 7 shows how these scenarios are represented in the template for AEB testing.

<table>
<tr><th colspan="4">9. Scenarios</th></tr>
<tr><td rowspan="9">9.1 Highway Rear-End Collision (w/o comms)</td><td colspan="3">9.1.1. Scenario and Context: Ego vehicle travelling at the speed limit on a highway behind another vehicle. The lead vehicle suddenly brakes heavily.</td></tr>
<tr><td>9.1.1.1. Relevant ODD Factors</td><td>9.1.1.2. Relevant OES Measurements</td><td>9.1.1.3. Events</td></tr>
<tr><td>Weather Condition</td><td>Rain: {2 - 50} mm/h</td><td rowspan="4">AV starts traveling
AV maintains lane
Front vehicle suddenly brakes
AV reacts to stop safely</td></tr>
<tr><td>Highway Terrain</td><td>{incline, flat, decline}</td></tr>
<tr><td>Highway Speed</td><td>{55, 60, 65, 70 mph}</td></tr>
<tr><td>Illumination</td><td>{0.1 - 30,000} lux</td></tr>
<tr><td colspan="3">9.1.2. Influences</td></tr>
<tr><td>9.1.2.1. Influence Name</td><td>9.1.2.2. Description</td><td>9.1.2.3. Influence Specification</td></tr>
<tr><td>Inter-vehicle distance</td><td>Change the steady-state distance between the two vehicles</td><td>follow_distance: {2, 5, 10} m</td></tr>
<tr><td rowspan="11">9.2 Highway Rear-End Collision (with comms)</td><td colspan="3">9.2.1. Scenario and Context: Ego vehicle travelling at the speed limit on a highway behind another vehicle. The lead vehicle suddenly brakes heavily and broadcast a BSM.</td></tr>
<tr><td>9.2.1.1. Relevant ODD Factors</td><td>9.2.1.2. Relevant OES Measurements</td><td>9.2.1.3. Events</td></tr>
<tr><td>Weather Condition</td><td>Rain: {2 - 50} mm/h</td><td rowspan="5">AV starts traveling
AV maintains lane
Front vehicle suddenly brakes and broadcast a BSM
AV reacts to stop safely</td></tr>
<tr><td>Highway Terrain</td><td>{incline, flat, decline}</td></tr>
<tr><td>Highway Speed</td><td>{55, 60, 65, 70 mph}</td></tr>
<tr><td>Illumination</td><td>{0.1 - 30,000} lux</td></tr>
<tr><td></td><td></td></tr>
<tr><td colspan="3">9.2.2. Influences</td></tr>
<tr><td>9.2.2.1. Influence Name</td><td>9.2.2.2. Description</td><td>9.2.2.3. Influence Specification</td></tr>
<tr><td>Inter-vehicle distance</td><td>Change the steady-state distance between the two vehicles</td><td>follow_distance: {2, 5, 10} m</td></tr>
<tr><td>BSM delay</td><td>Change the network delay for the BSM</td><td>bsm_delay: {100 - 1000} ms</td></tr>
</table>

Fig. 7. A Sample of Relevant Scenarios for AEB Feature Testing

4. Feature Assessment using Simulation

AV feature performance measurement is a data-driven process. Enumerating a broad spectrum of scenarios over applicable ODD factors and OES variations using real vehicles, even on dedicated closed-circuit test tracks, is prohibitively time-consuming and resource-intensive. This is further the case for scenarios with a risk of potential harm to humans due to testing failure modes, which may lead to collisions or near-misses. High-fidelity physics-based simulations offer a practical way to safely and efficiently generate measurement data required for comprehensive evaluation. Simulations provide a controlled environment that closely replicates real-world conditions, enabling data collection and feature behavior assessment across a broad spectrum of scenarios. This facilitates early identification of potential safety risks and edge cases, leading to informed adjustments in a safe environment before real-world AV testing and deployment.

4.1. Simulation Software Architecture

Co-simulation principles can be leveraged to develop a high-fidelity simulation environment for AV feature testing. Key components of a typical architecture are as follows:

- Physics-based Simulation Engine: Provides a realistic virtual environment encompassing physics modeling, sensor simulation, and rendering capabilities. Example tools include `CARLA` https://carla.org/, `AWSIM` https://tier4.github.io/AWSIM/, and `CarSim` https://www.carsim.com/.
- Scenario/Traffic Management: Used for designing, configuring, and executing test scenarios covering a wide range of ODD/OES conditions, both environmental and network-related. These tools also can provide vehicle control automation and traffic management capabilities. Examples include `Scenario Runner` https://github.com/carla-simulator/scenario_runner and `Eclipse Sumo` https://eclipse.dev/sumo/.
- Communication Middleware: Provides real-time data exchange between the simulators, scenario management tools, and the AV software under test. Examples include `ROS2` (which leverages Data Distribution Services (DDS) https://docs.ros.org/en/humble/index.html, pure DDS implementations (e.g. `RTI Connext DDS`, `Eclipse Cyclone DDS`), and potentially custom middleware solutions.
- (optional) Additional Simulator(s): Dedicated tools to simulate other specific aspects. For example, network behavior such as packet delays, bandwidth limitations, and potential network-bound disturbances can be simulated using a dedicated network simulator. Popular options include `ns-3` https://www.nsnam.org/ and `OMNeT++` https://omnetpp.org/.

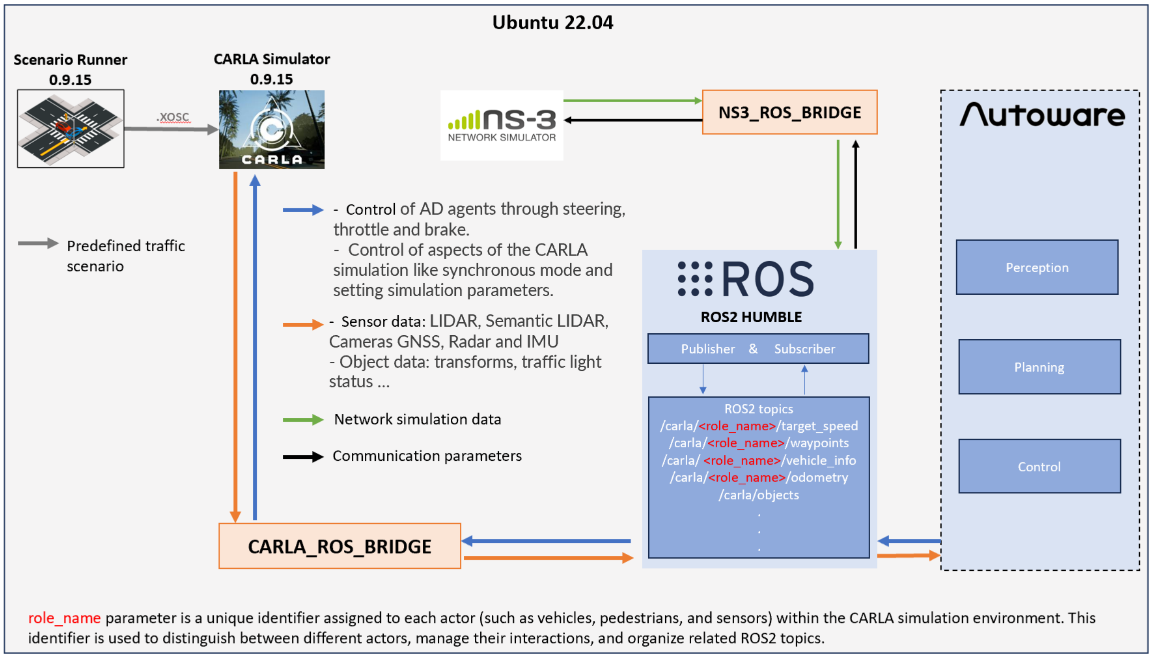

Fig. 8. An Illustrative Simulation Software Architecture

Figure 8 depicts an illustrative simulation software architecture utilizing CARLA, Scenario Runner, ROS2, and ns-3 for simulation engine, scenario management, communication middleware, and network simulator respectively. CARLA operates on a client-server model. The server handles the core simulation, including physics, sensor rendering, and the *"simulation world"* environment. Clients connect to the server, issuing commands to control vehicle actors and retrieve relevant sensor data. Typically, these are direct TCP connections. For example, Scenario Runner is a CARLA client that issues commands to the CARLA server to spawn vehicle actors and their route control. CARLA provides rich APIs with C++ and Python support, offering developers extensive control over simulation elements. This facilitates the customization and development of new functionalities tailored to specific AV testing scenarios.

The CARLA_ROS_BRIDGE (see Fig. 8) functions as a software bridge, introducing ROS2 middleware as an intermediary between the CARLA server and its clients. This bridge directly interacts with the CARLA server, centralizing all communication. Other simulation components interact with the server indirectly through the bridge, improving clock management and time synchronization. Additionally, the CARLA_ROS_BRIDGE enables the integration of other simulators (e.g., ns-3).

Direct access to the CARLA server allows the CARLA_ROS_BRIDGE to expose runtime configurable API elements as either ROS2 *topics* or ROS2 *services*. This design leverages ROS2's publish/subscribe paradigm to enable flexible, real-time access to sensor data and other simulation world information. Controlling weather conditions in the simulation is a good example. A customized copy of the CARLA_ROS_BRIDGE can be found at https://github.com/usnistgov/cav-cosim.git.

Other noteworthy open-source and/or research-driven AV co-simulation efforts elsewhere include CARMA https://github.com/usdot-fhwa-stol/carma-platform, Eclipse MOSAIC https://eclipse.dev/mosaic/, InterACT https://www.interact-roadautomation.eu/. NVIDIA DRIVE Sim https://developer.nvidia.com/drive/simulation and Virtual Test Drive https://hexagon.com/products/virtual-test-drive are several examples for commercial AV co-simulation platforms.

4.2. Simulation Setup for Feature Assessment and Testing

This section describes the process of configuring simulation software to evaluate the AEB feature introduced in Sec. 3.2 to demonstrate the potential application of simulation in assessing AV features. While all four key components in Fig. 3 are important, the most crucial for simulation is the definition of relevant scenarios. This work focuses on the scenario setup process, while also outlining how the other components contribute to the overall simulation framework. Figure 9 depicts the initial scene for AEB feature testing from a third-person view. For improved clarity and readability, the complete AEB feature testing template is depicted in Fig. 10.

Fig. 9. Initial Scene for the AEB Scenario

1. Feature	
Automatic Emergency Braking (AEB)	
2. SAE Automation Level	
Level 1	
3. Operational Design Domain (ODD)	**4. Operating Envelope Specification (OES)**
DOT HS 812 623	https://doi.org/10.6028/NIST.SP.1900-301
5. Feature Description	
The Automated Emergency Braking (AEB) Feature is designed to avoid collisions by automatically applying the brakes when obstacles are detected ahead, leveraging a combination of sensor data and vehicle-to-vehicle (V2V) communication.	
6. Systems and System Technologies	
6.1. Systems	**6.2. System Technologies**
AV Controller for Brake Monitoring	Perception
V2V Communication	Communication
Object Detection	Perception
Braking System	Mechanical and Control

7. Behaviors Assessment		
7.1. Behavior Name	**7.2. Behavior Description**	**7.3. Behavior Specification**
Responding to Lane Obstructions and Obstacles	The vehicle stops with sufficient stopping distance before the obstacle.	$veh\_to\_obst\_distance \geq d_{min}$
Responding to External Communication Signals	The vehicle appropriately responds to messages transmitted from another vehicle through V2V Communication.	$time\_to\_process \leq t_{max}$

8. Feature Performance	
8.1. Feature Specification	**8.2. Performance Metric**
Predefined criteria and constraints that the AEB system must meet.	*Quantifiable measure of performance (e.g., the AEB activation rate). The specification for success could be that this metric must achieve a 95% activation rate when an obstacle is detected within a critical distance or when a message is received through the communication channel.*

9. Scenarios

9.1 Highway Rear-End Collision (w/o comms)		
9.1.1. Scenario and Context: Ego vehicle travelling at the speed limit on a highway behind another vehicle. The lead vehicle suddenly brakes heavily.		
9.1.1.1. Relevant ODD Factors	**9.1.1.2. Relevant OES Measurements**	**9.1.1.3. Events**
Weather Condition	Rain: {2 - 50} mm/h	AV starts traveling
Highway Terrain	{incline, flat, decline}	AV maintains lane
Highway Speed	{55, 60, 65, 70 mph}	Front vehicle suddenly brakes
Illumination	{0.1 - 30,000} lux	AV reacts to stop safely
9.1.2. Influences		
9.1.2.1. Influence Name	**9.1.2.2. Description**	**9.1.2.3. Influence Specification**
Inter-vehicle distance	Change the steady-state distance between the two vehicles	follow_distance: {2, 5, 10} m

9.2 Highway Rear-End Collision (with comms)		
9.2.1. Scenario and Context: Ego vehicle travelling at the speed limit on a highway behind another vehicle. The lead vehicle suddenly brakes heavily and broadcast a BSM.		
9.2.1.1. Relevant ODD Factors	**9.2.1.2. Relevant OES Measurements**	**9.2.1.3. Events**
Weather Condition	Rain: {2 - 50} mm/h	AV starts traveling
Highway Terrain	{incline, flat, decline}	AV maintains lane
Highway Speed	{55, 60, 65, 70 mph}	Front vehicle suddenly brakes and broadcast a BSM
Illumination	{0.1 - 30,000} lux	AV reacts to stop safely
9.2.2. Influences		
9.2.2.1. Influence Name	**9.2.2.2. Description**	**9.2.2.3. Influence Specification**
Inter-vehicle distance	Change the steady-state distance between the two vehicles	follow_distance: {2, 5, 10} m
BSM delay	Change the network delay for the BSM	bsm_delay: {100 - 1000} ms

Fig. 10. AEB Feature Description and Testing Framework Example

4.2.1. Mapping Feature Definition to Simulation

The relevant elements from the Feature Definition that map to the simulation are Systems (Fig. 10, item 6.1) and System Technologies (Fig. 10, item 6.2), and the feature itself and its description (Fig. 10, item 1, and item 5). The underlying physics engine (like CARLA) typically fulfills many of the systems and system technologies requirements. This includes object detection and core vehicle control functionality. The AEB feature must also be explicitly mapped to the simulation, which involves implementing the AEB's decision-making algorithms, simulating its interaction with vehicle sensors, and defining its response to detected obstacles and emergencies. Capabilities not intrinsically supported by the engine may require custom programming or the integration of dedicated simulators (e.g., `ns-3` for V2V communication). For more advanced object detection, the simulation setup can be enhanced by integrating specialized AV software stacks like `Autoware` (https://autoware.org/).

4.2.2. Mapping Behaviors Assessment and Feature Performance to Simulation

Behavior Specifications (Fig. 10, item 7.3), Feature Specifications (Fig. 10, item 8.1), and Performance Metric (Fig. 10, item 8.2) are key elements for evaluating AV behavior and feature performance through simulation. These elements directly correspond to measurements and observations gathered within the simulated environment.

Behaviors exhibited by an ego vehicle relate to performing subtasks of the DDT, and the ability to measure and customize these depends directly on the capabilities offered by the test setup, including the capabilities of a simulation engine API for virtual testing. Simulation engines offer native support for core DDT subtasks like lateral and longitudinal vehicle control *(operational)*, along with perception capabilities for object and event detection, recognition, and classification *(tactical)*. These operational and tactical functionalities can be further customized through `CARLA` API (see `carla.VehiclePhysicsControl`, `carla.Sensor`, `carla.VehicleControl`, etc., from the `CARLA` Python API https://carla.readthedocs.io/en/0.9.15/python_api/) to align with specific simulation requirements. For behaviors not natively supported, custom programming is necessary; for example, responding to V2V signals requires setting up appropriate `ROS2` *nodes*, and their corresponding publishers and subscribers.

Access to operational, tactical, and environmental data from the simulated world is crucial for accurate measurement of feature behavior. Real-time access and exchange of this information can be achieved through `ROS2` *topics*. For example, assessing `veh_to_obst_distance` (see Sec. 3.2.2) relies on continuously updated information on the ego vehicle's location.

Features are rarely found natively within simulation engines. Even if some features exist within a simulator, it's essential to remember that their assessments would reflect the performance of the simulator's feature, not the actual AV feature being evaluated. The accuracy of the Feature Performance assessment is primarily determined by how well the simulation scenarios are designed and executed. Additionally, the responsibility for accurately implementing the feature within the simulation lies with the test conductor. This implementation can be carried out by integrating the feature as hardware-in-the-loop (HIL) or by incorporating external software via another simulator. A comprehensive and well-structured test plan is essential for ensuring the reliability of the feature performance evaluation.

4.2.3. Scenario Setup for Simulation

Relevant ODD Factors, Relevant OES Measurements, and Events (refer to Fig. 10, items 9.1.1.1-3 and 9.2.1.1-3) are all simulation parameters from the scenario simulation perspective. The process of defining scenarios for simulation involves setting values for these parameters, often in a structured way to facilitate comprehensive assessment. Simulation parameters can broadly be categorized as either fixed/static, which remain constant across a set of scenarios, or variable parameters, which are adjusted between simulation runs to assess their impact on feature performance. ODD factors tend to have a longer lifespan with adjustments between simulation sets, while OES measurements exhibit more frequent fluctuations.

Fixed/static parameters provide a controlled backdrop for assessing the AV feature in question. These parameters, including behaviors and routes of non-ego vehicles, initial vehicle separation, vehicle models, initial spawn point (or specific road segment), simulation maps/towns, and more, can be pre-configured to streamline the simulation process. They are defined in the OpenSCENARIO (https://www.asam.net/standards/detail/openscenario-xml/) files, which allow for a precise definition of maneuvers for scenario actors through their storyboard elements. This approach is exemplified in the excerpt (Fig. 11). Scenario Runner's support for OpenSCENARIO simplifies the definition and execution of simulation scenarios leveraging these pre-configured elements.

Variable parameters like ODD factors need to be carefully set up for a systematic and comprehensive assessment. The best approach is to initially focus on a single ODD factor at a time over a range of realistic values that reflect the full scope of conditions the feature is designed to handle. For example, testing under different rainy weather conditions might involve simulating a range from light drizzle to heavy downpour to see how the feature performs. During actual operation, the AV will encounter even greater variability within its ODD, reflected in real-time OES measurements (e.g., the precisely measured rainfall in mm/h at a given moment).

Figure 12 showcases the impact of varying a single ODD factor on the simulation environment. Here, the ODD factor under investigation is rain intensity. Three separate simulation runs are depicted as screenshots, each experiencing a different level of rainfall: light rain (Fig. 12a), moderate rain (Fig. 12b), and heavy rain (Fig. 12c). Variations in rain intensity are achieved through mechanisms such as `ROS2` *services* that can modify weather conditions at simulation runtime to avoid restarting the simulation each time an OES measurement is modified. The screenshots depict the simulated environment from the ego vehicle's front-facing camera perspective.

This dynamic adjustment of simulation parameters contrasts with the use of fixed parameters often defined within scenario files (like OpenSCENARIO). The ability to alter parameters at runtime allows for greater flexibility and the exploration of a wider range of scenarios during the assessment process. This same systematic approach to runtime parameter

```xml
<?xml version="1.0" encoding="UTF-8"?>
<OpenSCENARIO>
    <ParameterDeclarations>
        <ParameterDeclaration name="InitSeperation" parameterType="double" value="40.0" />
        <ParameterDeclaration name="LeadCarSpeed" parameterType="double" value="19.44" />
    </ParameterDeclarations>
    <RoadNetwork>
        <LogicFile filepath="Town05"/>
    </RoadNetwork>
    <Entities>
        <ScenarioObject name="hero"><Vehicle name="vehicle.tesla.model3" vehicleCategory="car"> ⋯
        </ScenarioObject>
        <ScenarioObject name="LeadCar"><Vehicle name="vehicle.lincoln.mkz_2020" vehicleCategory="car">⋯
        </ScenarioObject>
    </Entities>
    <Storyboard>
        <Init>⋯
        </Init>
        <Story name="ReadEnd">
            <Act name="Act1">
                <ManeuverGroup name="LeadCar_Behavior" maximumExecutionCount="1">
                    <Maneuver name="LeadCar_Actions">
                        <Event name="LeadCar_Start" priority="parallel" maximumExecutionCount="1">
                            <Action name="LeadCar_Start">
                                <PrivateAction>
                                    <LongitudinalAction>
                                        <SpeedAction>
                                            <SpeedActionDynamics dynamicsShape="step" value="0" dynamicsDimension="time"/>
                                            <SpeedActionTarget>
                                                <AbsoluteTargetSpeed value="$LeadCarSpeed"/> <!-- 19.44 m/s = 70mph -->
                                            </SpeedActionTarget>
                                        </SpeedAction>
                                    </LongitudinalAction>
                                </PrivateAction>
```

Fig. 11. An Excerpt of an OpenSCENARIO file Showing Pre-Configured, Fixed Simulation Parameters

variations can also be applied to assess the impact of other ODD factors (Fig. 10, items 9.1.1.1, 9.2.1.1).

After individual ODD factor evaluations are completed, different ODD factors can be combined in assessment to gain a more comprehensive understanding of the feature's robustness and limitations. For instance, simulations might combine different levels of rain intensity (encompassing both design specifications and potentially more extreme conditions) with variations in highway terrain (flat vs. incline). This approach enables a thorough assessment of how the AV feature behaves across a wide range of conditions, both within and beyond its intended ODD. Figure 13 illustrates this idea, with contrasting scenarios like clear weather on flat terrain (Fig. 13a) and rainy conditions on an incline terrain (Fig. 13b). These simulations yield valuable data for analyzing feature performance under diverse operational conditions.

To customize the simulation environment with the specificity required by the testing template, `CARLA` provides in-depth control over OES measurements and other relevant parameters. APIs such as `carla.WeatherParameters` enable fine-grained adjustments to weather conditions, including variables like fog density, precipitation intensity, and cloudiness. Additionally, the `carla.WheelPhysicsControl` API allows for the modification of

(a) Light Rain **(b)** Mid Rain

(c) Heavy Rain

Fig. 12. Single ODD Factor Variation (Rain Intensity)

(a) Clear Weather and Flat Highway Terrain **(b)** Rainy Weather and Incline Highway Terrain

Fig. 13. Multiple ODD Factor Combinations: Weather and Highway Terrain

elements like tire friction and suspension damping rates, influencing vehicle dynamics and road interactions. The CARLA API offers control over many aspects of the simulation from the client side. For even more advanced customization, modifications can be made directly to the CARLA server source code prior to compilation.

NOTE: It's important to recognize that, however, any sort of simulation-based assessment is directly influenced by the fidelity of the underlying physics engine. Unfortunately, limitations exist in vehicle simulators. Specifically the CARLA physics engine (as of this writing) doesn't automatically adjust the vehicle slippage based on factors like rain intensity, which stands in contrast to how vehicle dynamics are affected by weather conditions in the real world. Similarly, CARLA doesn't easily replicate the impact of real-world elements like tire thread patterns, tire wear, and road surface material on a vehicle's stopping distance. These nuances, highlight the challenges of fully replicating all real-world conditions assessing AV behavior within a simulation. Careful planning by researchers requires a thorough understanding of such limitations, underscoring their impact on meaningful simulations and accurate interpretations of results.

4.3. Simulations and Feature Performance

The simulation setup described in this section is designed to meticulously gather data for Feature Performance assessment. By systematically varying ODD factors both individually and in combinations, the simulation environment captures the full range of conditions the feature is designed to handle and potentially challenging scenarios beyond the specified ODD. Dynamically updated OES measurements further enrich these simulations with real-time environmental variations.

This comprehensive approach facilitates a thorough analysis of the feature's behavior across a wide range of operational conditions. The insights generated can directly inform the definition of robust system behavior, ensuring that the feature maintains stability and safety across a variety of scenarios. Additionally, the evaluation process can assess the feature's adaptability by examining its ability to adjust performance parameters in response to changes in the environment or operating conditions. This ensures that the feature's behavior aligns with safe and reliable operation in complex real-world environments.

5. Conclusion

This report outlines a comprehensive framework for the description and assessment of AV features, emphasizing the need for a systematic approach to enhance communication among manufacturers, regulators, and other stakeholders. The framework integrates established elements such as ODD, levels of automation, and behaviors while introducing measurable components like OES, behavior specifications, and feature metrics. This provides a comprehensive approach to defining and testing AV features where each feature is assessed against specific, quantifiable parameters.

An example AEB use case is used to demonstrate the use of the framework and its implementation in simulation. Simulations are pivotal for gathering data on the measurable parameters outlined in the framework, allowing for both consideration of key scenarios that cannot practically be executed using real vehicles and iterative development of the feature description based on preliminary simulation results.

Overall, the core contribution of this report is the proposal of a feature description framework for AV performance evaluation, distinguishing it from existing methodologies that predominantly rely on scenario generation. Future work will consider extending the behavior specifications with constraints to develop a comprehensive framework to assess the performance of these features and identify ways to improve them.

References

[1] Committee ORADO (2021) *Taxonomy and definitions for terms related to driving automation systems for on-road motor vehicles* (SAE international).
[2] Klaver F (2020) The economic and social impacts of fully autonomous vehicles. *Retrieved February* 18:2022.
[3] Hellman II, Lindman M (2023) Estimating the crash reducing effect of advanced driver assistance systems (adas) for vulnerable road users. *Traffic Safety Research* 4:000036–000036.
[4] Validi A, Ludwig T, Olaverri-Monreal C (2017) Analyzing the effects of v2v and adas-acc penetration rates on the level of road safety in intersections: Evaluating simulation platforms sumo and scene suite. *2017 IEEE International Conference on Vehicular Electronics and Safety (ICVES)* (IEEE), pp 38–43.
[5] Ackerman RP (2019) Regulating autonomous vehicles in the united states .
[6] Thorn E, Kimmel SC, Chaka M, Hamilton BA, et al. (2018) A framework for automated driving system testable cases and scenarios (United States. Department of Transportation. National Highway Traffic Safety ...),
[7] (2021) AVSC Best Practice for Evaluation of Behavioral Competencies for Automated Driving System Dedicated Vehicles (ADS-DVs) (SAE International), Technical report AVSC00008.
[8] Griffor E, Wollman D, Greer C (2021) Automated driving system safety measurement part I: Operating envelope specification. *NIST Special Publication* 1900:301.
[9] Consortium AVS, et al. (2021) Best practice for metrics and methods for assessing safety performance of automated driving systems (ads). *SAE Industry Technologies Consortium* .
[10] De Gelder E, Paardekooper JP, Saberi AK, Elrofai H, den Camp OO, Kraines S, Ploeg J, De Schutter B (2022) Towards an ontology for scenario definition for the assessment of automated vehicles: An object-oriented framework. *IEEE Transactions on Intelligent Vehicles* 7(2):300–314.
[11] Ko W, Park S, Yun J, Park S, Yun I (2022) Development of a framework for generating driving safety assessment scenarios for automated vehicles. *Sensors* 22(16):6031.
[12] Weber H, Bock J, Klimke J, Roesener C, Hiller J, Krajewski R, Zlocki A, Eckstein L (2019) A framework for definition of logical scenarios for safety assurance of automated driving. *Traffic injury prevention* 20(sup1):S65–S70.
[13] Connected Automated Driving (2021) Methodology - festa. [Accessed: 08-July-2024] Available at https://www.connectedautomateddriving.eu/methodology/festa/.
[14] Alghodhaifi H, Lakshmanan S (2021) Autonomous vehicle evaluation: A comprehensive survey on modeling and simulation approaches. *IEEE Access* 9:151531–151566.
[15] Zhang S, Peng H, Zhao D, Tseng HE (2018) Accelerated evaluation of autonomous vehicles in the lane change scenario based on subset simulation technique. *2018 21st International Conference on Intelligent Transportation Systems (ITSC)*, pp 3935–3940. https://doi.org/10.1109/ITSC.2018.8569800

[16] Zhao D, Peng H, Lam H, Bao S, Nobukawa K, LeBlanc DJ, Pan CS (2015) Accelerated evaluation of automated vehicles in lane change scenarios. *Dynamic Systems and Control Conference* (American Society of Mechanical Engineers), Vol. 57243, p V001T17A002.
[17] Takács A, Drexler DA, Galambos P, Rudas IJ, Haidegger T (2018) Assessment and standardization of autonomous vehicles. *2018 IEEE 22nd International Conference on Intelligent Engineering Systems (INES)*, pp 000185–000192. https://doi.org/10.1109/INES.2018.8523899
[18] ZERO IPOV (2017) Euro ncap 2025 roadmap .
[19] Nilsson J, Ödblom AC, Fredriksson J (2015) Worst-case analysis of automotive collision avoidance systems. *IEEE Transactions on Vehicular Technology* 65(4):1899–1911.
[20] Li S, Li W, Li P, Ma P, Yang M (2022) Novel test scenario generation technology for performance evaluation of automated vehicle. *International Journal of Automotive Technology* 23(5):1295–1312.
[21] Junietz P, Bonakdar F, Klamann B, Winner H (2018) Criticality metric for the safety validation of automated driving using model predictive trajectory optimization. *2018 21st International Conference on Intelligent Transportation Systems (ITSC)* (IEEE), pp 60–65.
[22] Hallerbach S, Xia Y, Eberle U, Koester F (2018) Simulation-based identification of critical scenarios for cooperative and automated vehicles. *SAE International Journal of Connected and Automated Vehicles* 1(2018-01-1066):93–106.
[23] Feng S, Feng Y, Yu C, Zhang Y, Liu HX (2021) Testing scenario library generation for connected and automated vehicles, part i: Methodology. *IEEE Transactions on Intelligent Transportation Systems* 22(3):1573–1582. https://doi.org/10.1109/TITS.2020.2972211
[24] Alnaser AJ, Akbas MI, Sargolzaei A, Razdan R (2019) Autonomous vehicles scenario testing framework and model of computation. *SAE International Journal of Connected and Automated Vehicles* 2(4).
[25] Li L, Huang WL, Liu Y, Zheng NN, Wang FY (2016) Intelligence testing for autonomous vehicles: A new approach. *IEEE Transactions on Intelligent Vehicles* 1(2):158–166.
[26] Mehr G, Ghorai P, Zhang C, Nayak A, Patel D, Sivashangaran S, Eskandarian A (2022) X-car: An experimental vehicle platform for connected autonomy research. *IEEE Intelligent Transportation Systems Magazine* 15(2):41–57.
[27] Van Brummelen J, O'brien M, Gruyer D, Najjaran H (2018) Autonomous vehicle perception: The technology of today and tomorrow. *Transportation research part C: emerging technologies* 89:384–406.
[28] Ahmed-Zaid F (2019) Sae standard overview of basic safety message (bsm).