Conceptual Approaches to Identify the Hazardous Scenarios in Safety Analysis for Automated Driving Systems

Marzana Khatun[1][a], Florence Wagner[2][b], Rolf Jung[2] and Michael Glass[3]

[1]*Kempten University of Applied Sciences, Bahnhofstrasse 61, Kempten, Germany*

[2]*Institute for Driver Assistance and Connected Mobility, Benningen, Germany*

[3]*Embedded Systems/ Real-Time Systems, University of Ulm, Ulm, Germany*

{*marzana.khatun, florence.wagner, rolf.jung*}*@hs-kempten.de, michael.glass@uni-ulm.de*

Keywords: Functional Safety, Safety of the Intended Functionality, Hazard Analysis and Risk Assessment, Automated Driving System, Machine Learning.

Abstract: To ensure safety of the road users is one of the major challenges in highly automated driving. The technologies applied in semi or fully-automated vehicles that are safer than human drivers compromise functionalities and human comfort. A comprehensive understanding of the use of complex driving systems and the Operational Design Domain (ODD) is essential for the effective deployment and safe operation of Automated Driving Systems (ADSs). Hazard analysis is a foundation of various safety engineering methods, which include Functional Safety (FuSa) and Safety Of The Intended Functionality (SOTIF). The scenario-based analysis offers significant advantages in the safety analysis of automated vehicles but poses inherent difficulties in identifying unknown-hazardous scenarios. The work presented in this paper deals with the conceptual approaches of hazard scenario identification. Moreover, discusses the incorporation of Machine Learning (ML) in Hazard Analysis and Risk Assessment (HARA) for vehicles equipped with ADSs. Furthermore, this paper can serve as foundation support for research inquiries related to ADSs validation and safety assessment.

1 INTRODUCTION

The complexity in Automated Driving Systems (ADSs) has brought safety aspects to the forefront of research efforts (Tu and Sun, 2023). According to Society of Automotive Engineers (SAE) J3016:2021 standard, automated or autonomous vehicles are categorized into six automation levels from level 0 (No Driving Automation) to level 5 (Full Driving Automation). First three driving automation levels (0-2) require a human driver who is at all times supervising the support features and handling the driving for safety reasons. Level 3 higher are taking over the driving task while level 3 can detect the limits of its application range and request the driver be in control, level 4 and 5 are not requiring someone to take over driving. This paper focuses on driving automation level 4 (High Driving Automation) and/or level 5 (SAE, 2021). Current Functional Safety (FuSa) and Safety Of The Intended Functionality (SOTIF) related standards are focusing on the automation level 2 where humans are responsible to ensure driving safety. Standards such as ISO 26262:2018 (ISO26262, 2018) for road vehicle FuSa and ISO 21448:2022 (ISO21448, 2022) for road vehicle SOTIF consider performance insufficiencies of the complex components used in an ADS. Moreover, the upcoming ISO/PAS 8800 standard is a valuable addition that enhances safety in Artificial Intelligence (AI)/Machine Learning (ML), providing essential guidelines as well (ISO/PAS8800, 2024).

Fully-automated vehicles are intended to drive efficiently from one point to another. According to a study, the uses of automated driver-assistance systems in Europe could reduce the number of accident by about 15% by 2030 (Deichmann and Steiner, 2023). ADS-equipped vehicles can add additional value together with human comfort for automotive industry and could generate between $300 billion and $400 billion in the passenger car market by 2035, according to McKinsey analysis (Deichmann and Steiner, 2023). According to (Feldmann, 2023), Feldmann mentions that the adoption of safety measures is hindered by slow progress, mainly due to lack of resources. From this motivation, the research questions addressing in

[a] https://orcid.org/0000-0002-3839-1575

[b] https://orcid.org/0009-0009-2515-1116

Khatun, M., Wagner, F., Jung, R. and Glass, M.
Conceptual Approaches to Identify the Hazardous Scenarios in Safety Analysis for Automated Driving Systems.
DOI: 10.5220/0013250900003890

In *Proceedings of the 17th International Conference on Agents and Artificial Intelligence (ICAART 2025) - Volume 1*, pages 549-555
ISBN: 978-989-758-737-5; ISSN: 2184-433X

this paper with respect to safety analysis of ADS-equipped vehicles are:

1. What methodologies can be employed to identify unknown-hazardous scenarios and subsequently reduce unknown-hazardous concerns in the context of SOTIF-related scenario evaluation?
2. What steps can be involved to perform scenario-based hazard analysis, including scenario simulation?
3. Can continuous improvement be achieved in hazard analysis and risk assessment to effectively support the verification and validation of automated driving systems?
4. What strategies or methodologies can be employed to ensure ongoing enhancement in this critical aspect of automated vehicle development?

This paper focuses on the scenario-based safety analysis in terms of hazardous scenario identification and support to reduce the unknown-hazardous scenarios. Unknown-hazardous scenarios are described as Area 3 based on the scenario category division from SOTIF standard. Here, the scenarios are classified as known-not hazardous (Area 1), known-hazardous (Area 2), unknown-hazardous (Area 3) and unknown-not hazardous (Area 4). In relation to ADS safety, the scenarios assigned to Area 3 are considered to be the most critical. Scenario modeling and simulation play a vital role in investigating and understanding the critical aspects of ADS safety. The contributions of this paper are listed below:

- Propose conceptual approaches to incorporate a new technology in hazardous scenario identification.
- Description of safety-related scenarios to support scenario-based hazard analysis and risk assessment for automated vehicles.
- Illustrate incorporation of ML and deterministic approaches in scenario-based Hazard Analysis and Risk Assessment (HARA) applicable in automation level 4 or higher. Looking at level 4 and 5 of driving automation the driving features will be in control and will not require a human driver to take over.

The outline of this paper is as follows, Section 2 introduces normative definitions of terms used in safety and cybersecurity domain. Section 3 presents the proposed approaches to identify hazardous scenarios and risk assessment. Finally, Section 4 provides a redundant summary of the discussion and offers an outlook on the future perspectives of the study.

2 LITERATURE STUDY

2.1 Normative Terms and Definitions

According to SAE J3016:2021 (SAE, 2021), ADS is the ”hardware and software that are collectively capable of performing the entire dynamic driving task on a sustained basis, regardless of whether it is limited to a specific ODD; this term is used specifically to describe a Level 3, 4, or 5 driving automation system.” The ODD is described as the ”specific conditions under which a given driving automation system is designed to function” (ISO21448, 2022). The ODD is not limited to, environmental, geographical, and time-of-day restrictions, and/or the requisite presence or absence of certain traffic or roadway characteristics (SAE, 2021). According to ISO 26262-1:2018 (ISO26262, 2018), SAFETY is the ”absence of unreasonable risk”. FUNCTIONAL SAFETY is defined as an ”absence of unreasonable risk due to hazards caused by malfunctioning behavior of electrical and/or electronic systems”. A HAZARD is defined as a ”potential source of harm caused by malfunctioning behavior of the item”, and a HAZARDOUS EVENT is defined as a ”combination of a hazard and an operational situation”. The HARM indicates the ”physical injury or damage to the health of persons” (ISO26262, 2018).

According to ISO 21448:2022, a SCENARIO is described as ”a temporal relationship between several scenes in a sequence of scenes, with goals and values within a specified situation, influenced by actions and events” (ISO21448, 2022). The HAZARDOUS SCENARIO is a ”scenario when harm occurs unless prevented by an entity other than the ADS” (ISO34502, 2022).

2.2 Methods and Technologies

As stated in 26262-1:2018 (ISO26262, 2018), a FUNCTIONAL CONCEPT is described as ”specification of the intended functions and their interactions necessary to achieve the desired behavior. The functional concept developed during the concept phase describes a set of specification of the functional safety requirements, with associated information, their allocation to elements within the architecture, and their interaction necessary to achieve the safety goals”. HARA is a ”method to identify and categorize hazardous events of items and to specify safety goals and automotive safety integrity levels related to the prevention or mitigation of the associated hazards in order to avoid unreasonable risk” (ISO26262, 2018).

The growing need for large-scale coverage of scenarios in autonomous vehicles is driving the adoption of technologies such as ML models. These technologies are used to support scenario-based analysis, identifying safety-critical parameters, scenario simulation and reduction in testing and optimize time efficiency. In (Grigorescu and Macesanu, 2020), GRIGORESCU, TRASNEA, COCIAS AND MACESANU present a survey of the current state-of-the-art on deep learning technologies used in autonomous driving. Simulation-based HARA is suggested by (Oakes, 2021) but focused on fault injection not in hazardous scenario identification. Furthermore, a concept of hazard scenario identification is presented in (Khatun and Glaß, 2023) based on simulated scenario but risk assessment was not present in detail.

Scenario-based HARA supports to understand and realize the hazardous events and is widely applied in automated vehicle's safety assessment (Madala and Solmaz, 2023; Khatun and Jung, 2020). To investigate the hazard and risk methods like, System-Theoretic Process Analysis (STPA) and dynamic hazard approach are applied in ADS-equipped vehicles (Chen and Zhao, 2020; Schwalb, 2021). A priori safety risk assessments of ADS-equipped vehicles road tests are of great importance as they allow quantification of the level of risk in different road scenarios and provide guidance for such vehicles road tests (Tu and Sun, 2023),(TR5469, 2024). AI technologies can not only help to recognize, research and to define hazardous scenarios, but they can also accelerate functional safety analyses. Although validating AI system safety poses significant challenges, it can be managed effectively through rigorous confirmation measures.

A scenario-based HARA approch is presented here with the support of ML model to define the inputs that are the key to establish the base scenario framework in HARA, applicable for L3 and/or higher level of automation (L4/L5.)

3 PROPOSED HAZARD ANALYSIS APPROACHES

3.1 Scenario-Based Hazard Analysis

While scenario-based analysis has many advantages by means of detail description and reducing misunderstanding in autonomous safety analysis, it also has some shortcomings. The time and effort required for scenario-based HARA is very high. One of the biggest challenges is the number of scenarios that need to be considered for risk assessment. A further obstacle arises from the limited applicability of real-world scenarios to automation level 3 and/or higher, leading to the emergence of novel hazard scenarios that cannot be effectively derived from real-world scenarios. In such cases, the inclusion of simulation-based scenarios is essential for comprehensive safety analyses.

Simulation-based approaches have gained significant traction for identifying unknown-hazardous scenarios (Zhang and Felbinger, 2023). ISO 21448:2022 introduces scenario categories divided into areas 1 to 4, based on the attributes known/unknown and hazardous/non-hazardous. These categories are used to facilitate the preparation of the safety concept and the definition of hazardous events. The hazard analysis begins with the known hazardous scenarios (Area 2). Simulation-based and knowledge-based investigations based on safety-related parameters can be used to reduce the scope of Area 3, as shown by the arrows in Fig. 1. In cases where simulation-based scenario analysis begins with known-hazardous scenarios, the utilization of parameter variation techniques aids in the identification of unknown-hazardous scenarios and thus leads to a reduction in the scope of Area 3 (See Fig. 1).

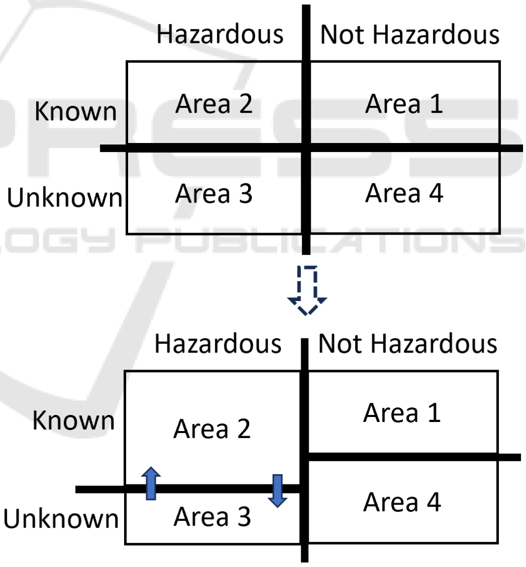


Figure 1: Scenario Categories Visualization with Areas.

3.2 Proposed Concepts

The conceptual approaches proposed in this take into account the different representations of real-world crash scenarios, virtual scenarios, and pre-crash scenarios. To generate scenarios methods like, tree diagram, formal rule-based and model-based approaches can be applied for safety analyses (Khan and Vijay, 2023; Faysal, 2022). For scenario-simulation tools like, CarMaker (IPG-Automotive,), MATLAB (Matlab,), Simulation of Urban Mobility (SUMO) (Li, 2023), Unity3D (Gang, 2016), CARLA Simulator (Horel, 2022) are utilized in various stages of verification and validation. To generate and identify critical scenarios, SOTIF-related hazard analysis and risk assessment (HARA) includes several areas to consider (Koné and Géronimi, 2023; Yang et al., 2023; Sana and Raahemifar, 2023; Kramer et al., 2020). These include ODD, environmental conditions, and complex system features leading to limitations in human-performed HARA due to the heightened volume of scenarios.

The different representations of these scenarios cause difficulties in analyzing of hazards for automated driving systems. To aid in mitigating challenges, a generic flow diagram is depicted in Fig. 2. The flow diagram in Fig. 2 shows the possibilities how a scenario database can be created and simulated for identifying unknown-hazardous scenarios and continuously update the scenario database.

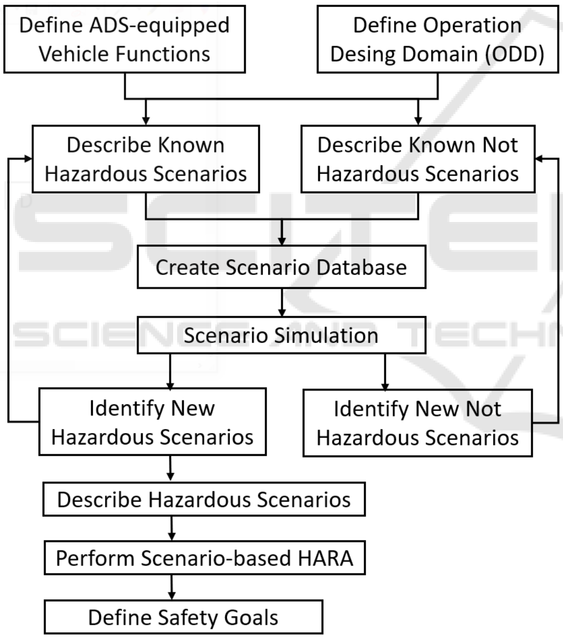

Define ADS-equipped Vehicle Functions
Define Operation Desing Domain (ODD)
Describe Known Hazardous Scenarios
Describe Known Not Hazardous Scenarios
Create Scenario Database
Scenario Simulation
Identify New Hazardous Scenarios
Identify New Not Hazardous Scenarios
Describe Hazardous Scenarios
Perform Scenario-based HARA
Define Safety Goals

Figure 2: Proposed Flow Diagram for Scenario-based HARA.

In the following two concepts, both deterministic (yellow color) and ML model (light blue color) approaches are integrated within the context of hazardous scenario identification and risk assessment for ADS-equipped vehicles. The concepts include the use of Hazard and Operability Analysis (HAZOP) keywords displayed as dark green color (see in Fig. 3 and Fig. 4) and characteristics that are used in describing hazardous events (light green color), as shown in Fig. 3 and Fig. 5.

Concept 1. The proposed concept 1 is divided into two segments. In the first segment, the hazardous events are compiled based on the characteristics of the scenario descriptions. Finite number of Hazardous Events (HE) are presented in Fig. 3 from 1 to n (HE1, HE2, ..., HEn). Then, hazardous events are combined with the predefined HAZOP keywords by applying an ML technique. The characteristics mentioned in Fig. 3 are operating mode, operational situation and vehicle functions. The deterministic approach is marked by a yellow block and can be carried out by a knowledge-based method for describing hazard events. The number of hazard events can range from a finite number between 1 and n (see Fig. 3). The hazardous events database can be employed to identify previously unknown-hazardous scenarios, thereby contributing to the reduction of Area 3.

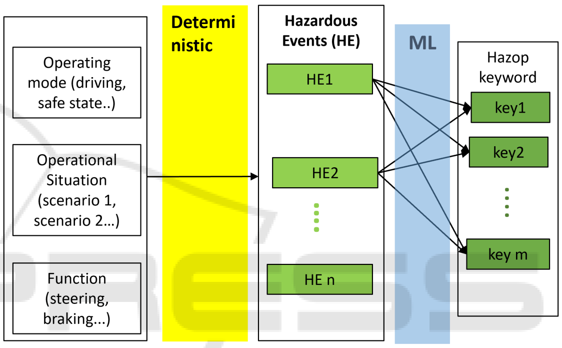

Operating mode (driving, safe state..)
Operational Situation (scenario 1, scenario 2...)
Function (steering, braking...)
Determi nistic
Hazardous Events (HE)
HE1
HE2
HE n
ML
Hazop keyword
key1
key2
key m

Figure 3: Hazardous Scenario Identification - Concept 1 (Segment 1).

In the second segment of concept 1, hazard events combined with HAZOP keywords are used to formulate Accident Scenarios (ACs). Each HAZOP keyword is associated with individual hazardous events, to define multiple sets of accident scenarios (red color) as shown in Fig. 4. A ML model can be applied to determine the Severity (S), Exposure (E) and Controllability (C). Based on these three parameters, Automotive Safety Integrity Level (ASIL) can be defined using risk graph or risk matrix for risk assessment of ADS-equipped vehicles.

An exemplary case: The operating mode is driving an ADS-equipped vehicle in an operating scenario in which the vehicle performs a lane change from the right to the left lane on a straight highway in sunny weather. To perform the lane change, the ADS-equipped vehicle uses the steering, acceleration, and braking functions. There are other road users on the road that are passed by an ADS-equipped vehicle while it is performing a lane change. A hazardous event can be described as follows: The camera sensor fails to detect the lane marking while an ADS-equipped vehicle is performing a lane change.

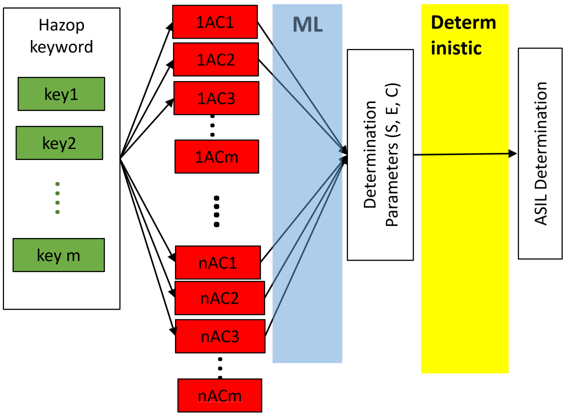


Figure 4: Hazardous Scenario Risk Assessment - Concept 1 (Segment 2).

Some of the traditional HAZOP keywords are no or not, late, after. Based on the HAZOP keywords, ACs can be described as follows: A rear-end collision occurs. Accident database can be used as mixed datasets. Mixed datasets can be used to identify the parameters (S, E, C) using deep neural networks, such as convolutional neural networks and recurrent neural networks. Finally, a deterministic method can be applied to determine ASIL based on ISO 26262:2018.

Concept 2. Concept 2 is divided into two segments. First, hazardous events are considered as database for ML. Hazardous events are classified into several sets as shown in Fig. 5 The HE classification sets are presented in Fig. 5 from 1 to k (set 1, set 2, ..., set k).

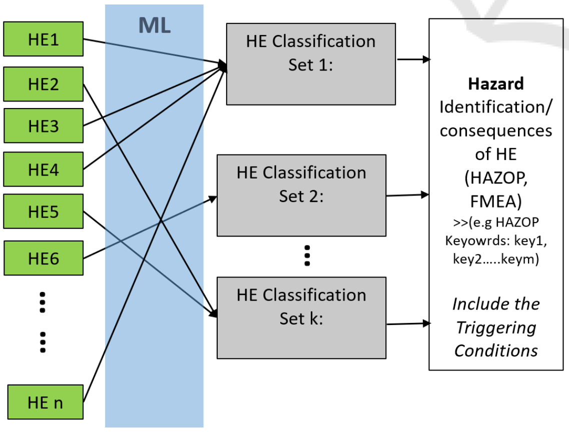
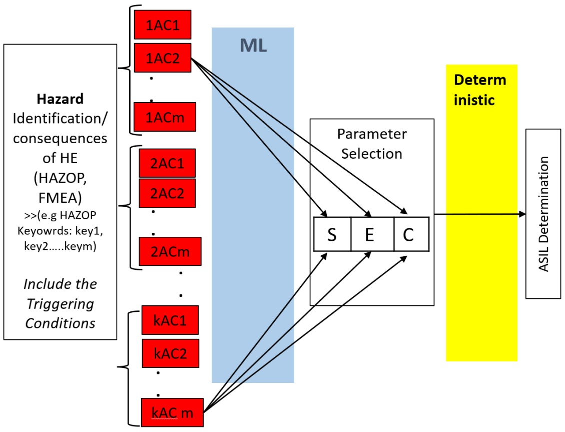

HE1
HE2
HE3
HE4
HE5
HE6
HE n
ML
HE Classification Set 1:
HE Classification Set 2:
HE Classification Set k:
Hazard Identification/ consequences of HE (HAZOP, FMEA)
>>(e.g HAZOP Keyowrds: key1, key2.....keym)
Include the Triggering Conditions

Figure 5: Hazardous Scenario Identification - Concept 2 (Segment 1).

Afterward, the sets of hazardous events are combined with the HAZOP keywords to describe the ACs. Several sets of finite ACs are shown in Fig. 6 that are considered based on the hazardous events with HAZOP keywords. The accident database can be considered as a mixed dataset, containing both numerical and textual values.


Hazard Identification/ consequences of HE (HAZOP, FMEA)
>>(e.g HAZOP Keyowrds: key1, key2.....keym)
Include the Triggering Conditions
1AC1
1AC2
1ACm
2AC1
2AC2
2ACm
kAC1
kAC2
kAC m
ML
Parameter Selection
S
E
C
Determ inistic
ASIL Determination

*xACy: x=HE Set1 with HAZOP y=Key2 lead to an accident- 1AC2

Figure 6: Hazardous Scenario Risk Assessment - Concept 2 (Segment 2).

An exemplary case: publicly available hazardous events from real-world scenarios are used together with simulated scenarios and considered as the base for ML models, which can be used to classify hazardous events. Each of these classification blocks can be associated with HAZOP keywords to create ACs. An AC can be expressed as follows: Accident on the highway, as vehicles equipped with ADS show an unintended behavior and injure a human. The ML model should provide the parameters (S, E, C), and later ASIL can be determined based on the deterministic approach.

4 CONCLUSION

The work showcases the conceptual approaches how ML models can be applied for the scenario-based HARA with respect to FuSa and SOTIF aspects. Two conceptual approaches are presented which gives the fundamental of incorporating ML in safety analyses by means of HARA. This paper presents the basics of HARA through the interplay of ML and deterministic approaches. This integration provides valuable insights for safety mechanisms in hazardous scenarios. The proposed approach not only improves the depth of risk assessment, but also enables real-time adjustments and scenario updates as new data becomes available, making the risk management process more dynamic and responsive. Ultimately, the use of AI in scenario-based HARA supports a more proactive security strategy that responds to emerging threats in complex systems such as autonomous vehicles and other AI-driven applications.

In future, simulation and publicly available real-world scenario database will be considered for safety analysis. To perform verification of the scenario-based HARA for ADS-equipped vehicle, safe human intervention needs to be ensured. The safety processes should include all possible safety-critical human interventions and provide safety measures to reduce the risk. To support safety analysis, ML techniques can be applied in HARA with acceptable verification and validation processes that include new test strategies and methods. To validate the scenario-based HARA for ADS-equipped vehicles, it is essential to incorporate a human safety intervention process. This process is aimed at evaluating the HARA results in view of the ML applications to maintain the integrity of the assessment.

ACKNOWLEDGEMENTS

This research work is funded by the Institute for Driver Assistance and Connected Mobility (IFM), Junkersstraße 1A, 87734 Benningen, Germany.

REFERENCES

Chen, L., J. J. and Zhao, T. (2020). A novel hazard analysis and risk assessment approach for road vehicle functional safety through integrating stpa with fmea. *Applied Sciences 10, no. 21: 7400*.

Deichmann, J., E. E. H.-K. H. R.-K. M. and Steiner, F. (2023). *Autonomous driving's future: Convenient and connected*. McKinsey Center for Future Mobility, McKinsey & Company.

Faysal, J. A. (2022). Formal rule-based scenarios for the design of safe autonomous vehicles. modeling and simulation. *HAL open Science, Universié Côte d'Azur*.

Feldmann, A. (2023). *The industry adoption of innovative safety-analysis approaches for autonomous vehicles*. KTH Industrial Engineering and Management, master thesis edition.

Gang, S.-M., C. H.-W. K. D.-R. C. Y.-J. (2016). A study on the construction of the unity 3d engine based on the webgis system for the hydrological and water hazard information display. *Procedia Engineering, vol 154, PP. 138–145*.

Grigorescu, S., T. B. C.-T. and Macesanu, G. (2020). *A Survey of Deep Learning Techniques for Autonomous Driving*. Elektrobit Automotive and the Robotics Vision and Control Laboratory (ROVIS Lab), Department of Automation and Information Technology, Transilvania University of Brasov, Romania.

Horel, J.-B., L. C.-M. L. M.-R. M. L.-e. a. (2022). Using formal conformance testing to generate scenarios for autonomous vehicles. In *Automation and Test in Europe - Autonomous Systems Design*.

IPG-Automotive. Carmaker: Software tool. https://ipg-automotive.com/de/anwendungen/autonomes-fahren/.

ISO21448 (2022). *ISO 21448: Road vehicles - Safety of the intended functionality*. ISO Standard, 1 edition.

ISO26262 (2018). *ISO 26262: Road vehicles - Functional Safety: Part 1 to Part 13*. ISO Standard, 2 edition.

ISO34502 (2022). *ISO 34502: Road vehicles — Test scenarios for automated driving systems — Scenario based safety evaluation framework*. ISO Standard, 1 edition.

ISO/PAS8800 (2024). *ISO PAS 8800: Road vehicles - Safety and artificial intelligence*. ISO Standard/Under development, 1 edition.

Khan, J. and Vijay, C. (2023). 27th international technical conference on the enhanced safety of vehicles (esv), yokohama, japan. In *alidation of safety of the intended functionality for autonomous driving systems*.

Khatun, M., G. M. and Jung, R. (2020). Scenario-based extended hara incorporating functional safety & sotif for autonomous driving. In *30th European Safety and Reliability Conference*.

Khatun, M., J. R. and Glaß, M. (2023). Scenario-based collision detection using machine learning for highly automated driving systems. In *In Journal Systems Science & Control Engineering, vol. 11, no. 1*. Taylor & Francis.

Koné, T. F., B. E.-L. E. M.-F. and Géronimi, S. (2023). An approach to guide the search for potentially hazardous scenarios for autonomous vehicle safety validation. *Applied Sciences, vol. 13, no. 11, p. 6717*.

Kramer, B., Neurohr, C., Büker, M., Böde, E., Fränzle, M., and Damm, W. (2020). Identification and quantification of hazardous scenarios for automated driving. In *Model-Based Safety and Assessment*.

Li, X., T. S. L.-B. D. X.-N. X. W.-F.-Y. (2023). Ieee transactions on intelligent vehicles, vol. 8, no. 5. In *Letters: Advanced Scenario Generation for Calibration and Verification of Autonomous Vehicles*.

Madala, K. and Solmaz, M. (2023). Scenario-based risk quantification approach for assuring safety in autonomous vehicles. *SAE Technical Paper 2023-01-0584*.

Matlab. Matlab: Software tool, automatic scenario generation. https://www.mathworks.com/help/driving/ug/automatic-scenario-generation.html.

Oakes, B. J., M. M.-M. S. V.-V. H. D.-J. (2021). Machine learning-based fault injection for hazard analysis and risk assessment. In *Computer Safety, Reliability, and Security*. Springer International Publishing.

SAE (2021). *SAE J3016: Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles*. SAE Standards.

Sana, F., A. N. L. and Raahemifar, K. (2023). Autonomous vehicle decision-making and control in complex and unconventional scenarios—a review. *Machines, vol. 11, no. 7, p. 676*.

Schwalb, E. (2021). Analysis of hazards for autonomous driving. *ASME Journal of Autonomous Vehicles Systems,vol.1, no. 2: 021003*.

TR5469, I. (2024). Iso/iec tr 5469: Artificial intelligence — functional safety and ai systems. Technical report, ISO/IEC.

Tu, H., W. M. L.-H. and Sun, L. (2023). Safety risk assessment for autonomous vehicle road testing. In *Traffic Injury Prevention*. Taylor & Francis.

Yang, Y., Kujanpää, K., Babadi, I. A., Pajarinen, J., and Ilin, A. (2023). Suicidal pedestrian: Generation of safety-critical scenarios for autonomous vehicles. In *2023 IEEE 26th International Conference on Intelligent Transportation Systems (ITSC)*, pages 1983–1988.

Zhang, X., T. J. T.-K. T. M.-J. S. M. G.-R. M. T. X. G. M. W. F. F. M. N. N. M. and Felbinger, H. (2023). Ieee trans. softw. eng., vol. 46, no. 3, ieee press, pp. 991–1026. In *Finding Critical Scenarios for Automated Driving Systems: A Systematic Mapping Study*.