Critical Scenario Techniques for Automated Vehicles: Literature Review

Umut Durak[1], Siddhartha Gupta[2], Kaan Akgün[3]

[1]DLR Institute of Flight Systems, Lilienthalplatz 7, 38108 Braunschweig, Germany; umut.durak@dlr.de

[2]Institut für Informatik, Technische Universität Clausthal, 38678, Clausthal-Zellerfeld, Germany; siddhartha.gupta@tu-clausthal.de

[3]Institut für Informatik, Technische Universität Clausthal, 38678, Clausthal-Zellerfeld, Germany; kaan.akguen@tu-clausthal.de

Abstract. As Automated Driving Systems (ADS) become more widely used, there is rising worry regarding their safety and security. Traditional verification techniques, such as evaluating the vehicle's performance over a certain number of miles or kilometres, are insufficient to assess ADS risk. Instead, alternate approaches, such as scenario-based testing, which involves assessing the vehicle's performance in simulated situations that match real-world settings, are required to determine the system's safety and performance. This article presents a literature review on the importance of critical scenario analysis in ensuring the safety and reliability of ADS. Critical scenario analysis is a complete approach for identifying, quantifying, prioritising, selecting, and validating the most critical situations for ADS development and testing. ADS can reduce the frequency of accidents and fatalities caused by human error. Through this literature review, readers will understand the importance and benefits of critical scenario analysis methods and approaches, allowing them to systematically evaluate risks and opportunities and better comprehend the potential outcomes of future scenarios.

Introduction

The increasing number of vehicle accidents and fatalities is a global concern, with the World Health Organization reporting that 1.3 million deaths occur yearly [1]. The introduction of Automated Driving Systems (ADS) has the potential to drastically reduce this number, as it aims to remove human error – the leading cause of such accidents. This is because ADSs are meant to remove human mistakes, which are the primary cause of automobile accidents. The Society of Automobile Engineers (SAE) has classified different levels of automation for ADSs, ranging from no automation (level 0) to full automation (level 5) [2]. ADSs with an SAE level of 3 or above have a higher degree of automation and can accomplish many driving activities without human involvement [2]. As the level of automation increases, so does the level of safety concern, including the potential for system malfunction or failure [3]. To ensure the reliability and security of these systems, ADS must undergo comprehensive security evaluations before being deployed to the public[3],[4].

The most common procedures for validating ADS involve evaluating the vehicle's performance over a certain number of miles or kilometres and assessing the system's safety and dependability [6]. This is done by measuring the miles or kilometres a vehicle can travel without incident. The assumption is that the further a vehicle can go without incident, the safer it is. However, these techniques are not considered sufficient for evaluating the risk of ADS as they do not consider the various operational scenarios that a vehicle may encounter in the real world [5]. As an alternative, scenario-based testing may be used to more accurately assess the system's safety and performance [6]. Scenario-based testing involves evaluating the vehicle's performance in simulated scenarios replicating real-world conditions, such as variable weather, traffic patterns, and road conditions. The process involves deliberately modifying the conditions and parameters of the simulated scenario to test the vehicle's ability to operate safely and effectively. This method aims to evaluate the safety and reliability of the ADS by exposing it to a wide range of potential conditions and situations it may encounter in real-world operations. Virtual testing is scenario-based testing that uses computer simulations to replicate these scenarios instead of physical testing. This approach allows for a more efficient and cost-effective ADS testing method [29].

Critical scenarios refer to specific driving situations with a high probability of causing safety issues or problems in the ADS [10]. These scenarios may present a significant risk to the safety of the autonomous vehicle, its passengers, and other road users. They may pose a challenge to the proper functioning of the ADS. These scenarios can include but are not limited to, situations

ARGESIM Report 21 (ISBN 978-3-903347-61-8), p87-94, DOI: 10.11128/arep.21.a2124

such as heavy traffic, poor weather conditions, complex road layouts, and unexpected obstacles. Identifying and testing for these critical scenarios is an essential step in ensuring the safety and reliability of ADS before they are released to the public [10].

The literature review focuses on specific methods proposed in previous research for identifying critical scenarios likely to present significant risks or challenges to the safety or operation of ADS. These methods can include techniques for identifying and quantifying the potential impact of different scenarios, prioritising them based on their likelihood or potential impact, selecting the most critical ones, and evaluating them can increase performance to focus on during the development of the ADS. These methods aim to help ensure that the ADS is thoroughly tested and validated against a wide range of scenarios likely to be encountered in real-world operations before deployment to improve its safety and reliability. First, section 1 introduces fundamental concepts related to scenario-based safety validation for ADS. Then, Section 2 describes the literature review, which is used to identify relevant information on critical scenario analysis. This process involves searching through academic journals, conference proceedings, and other relevant sources to identify articles, papers, and studies that address critical scenario analysis in the context of ADS.

This paper presents a comprehensive overview of various techniques in a cohesive manner, allowing the reader to gain a holistic understanding of the subject matter. Bringing together a diverse range of concepts and approaches facilitates a deeper understanding of the interconnectedness of these different aspects and how they can be applied in practice.

1. KEY CONCEPTS

ADS are vehicles designed to operate and navigate without human intervention. These systems use a combination of sensors, cameras, and other technologies to sense their surroundings and decide how to drive and operate the vehicle. SAE has defined many levels of autonomy for ADS, ranging from Level 0, which requires human intervention for all driving activities, to Level 5, which can do all driving functions without human intervention. Many levels of autonomy for ADS, ranging from Level 0, which requires human intervention for all driving activities, to Level 5, which can do all driving functions without human intervention. The criteria for each level are clearly defined, including the tasks the system can perform and the situations in which it can perform them without human intervention. ADS aims to increase safety by using advanced technologies to perceive and analyse the driving environment, make decisions, and control the vehicle. By relying on these technologies rather than human drivers, ADS can potentially reduce the frequency of accidents and fatalities caused by human error, such as distracted or impaired driving [11],[12].

ISO/PAS 21448:2019, as cited by Geyer et al. and Ulbrich et al., defines a scenario as a sequence of events occurring in a particular order and in different locations or settings (referred to as "scenes"). Scenarios help describe the temporal series of images portrayed by scenes and actions, and occurrences that can enhance them[13]-[15].

Thorn et al. employed a method of testing called scenario-based testing to evaluate the performance of ADS. This method involves creating simulated scenarios replicating real-world conditions and situations that the ADS may encounter during operation. For example, scenario-based testing can include testing the vehicle's performance in simulated scenarios replicating real-world conditions such as variable weather, traffic patterns, or road conditions. These scenarios evaluate the ADS' ability to safely and effectively navigate different driving conditions [16].

Menzel et al. introduced the concept of using different levels of abstraction to represent and analyse potential scenarios that an autonomous vehicle may encounter. Functional scenarios are described using semantic or linguistic notations and are defined at a high level of abstraction. Logical scenarios are represented using state-space level with parameter ranges and are defined at a medium level of abstraction. Finally, concrete scenarios are described using concrete parameter values and are defined at a low level of abstraction. Using these different levels of abstraction allows researchers to better understand a vehicle's capabilities and limitations and develop strategies for improving safety and reliability [17].

Although not a topic for discussion in this paper, it should be noted that Operational Design Domain (ODD) is a concept mainly used in ADS to define the operating conditions and limitations of the vehicle to ensure safe operation. ODD specifies the range of environmental conditions within which the ADS is designed to operate. In reality, ADS is the combination of hardware and software that collectively performs the entire dynamic driving task (DDT) on a sustained basis, regardless of whether it is limited to a specific ODD. Specifying the ODD helps to set the boundaries of safe operation for the vehicle. At the same time, the ADS is responsible for navigating the car and reacting to the environment and traffic based on the ODD-defined parameters. Also

included in ADS is the concept of ontology, which is a formal representation of knowledge that describes concepts and relationships within a specific domain [22]. In the context of ADS, an ontology can be used to represent knowledge about the driving scene, including information about the road layout, traffic signs, and the behaviour of other vehicles. The ontology can also be used to represent the relationships between concepts, such as how traffic signs constrain the behavior of vehicles. Using an ontology to represent this knowledge allows for automation in creating traffic scenarios for testing and simulation, which is a more efficient process than manually creating these scenarios by experts [9]. This can be useful to identify critical scenarios in the design of ADS and to evaluate the performance of the ADS in a wide range of conditions.
Additionally, ontology metamodels can provide a standardised, structured way of representing and organising knowledge about the environment and the vehicle's capabilities [21]. Therefore, ODD and ontology metamodels may be considered vital components in ensuring the safety and performance of ADS. However, noting that ODD is not always required for scenario generation [7],[8],[9],[21],[22].

Critical scenarios generally refer to specific driving situations that present a high risk or danger to the safety and operation of ADS. These scenarios may include potential collisions, traffic violations, or other dangerous situations. Analysing critical scenarios is essential for developing strategies to enhance the safety and performance of ADS [18].

According to Zhang et al., critical scenarios are defined as difficult circumstances under which an autonomous vehicle can operate safely. The authors specify the specific parameters, such as weather and road conditions, and external circumstances, such as traffic density and presence of pedestrians, necessary for the vehicle to run safely. A comprehensive set of essential scenarios should be defined to ensure that the vehicle operates effectively in various locations and scenarios. Examining critical scenarios aids in understanding and developing strategies for improving the vehicle's performance and safety limits [19].

2. CRITICAL SCENARIO ANALYSIS

Critical scenario analysis is a comprehensive process used to develop and test ADS. This process aims to ensure the safety and reliability of ADS by thoroughly testing their capabilities under a range of scenarios based on real-world data and expert opinions. The following sections will present a literature review and explain the critical scenario analysis process. The process starts with critical scenario identification and quantification, followed by critical scenario prioritisation, validation, selection, and evaluation.

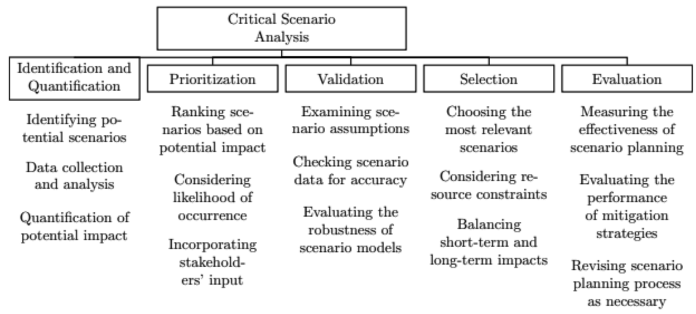

Figure 1: Overview of Critical Scenario Analysis[97]

2.1 Critical Scenario Identification & Quantification

Critical scenario identification and quantification is identifying possible scenario types, grouping them into categories, and giving numerical numbers to each scenario to compare and rank them. The purpose of scenario identification is to develop a well-defined set of scenarios that may be utilised to assess a specific problem or circumstance, identify potential risks and opportunities, and develop strategies for managing them. Scenario quantification is an essential step in the scenario management process. It helps to identify the most critical scenarios and concentrate on the issues most likely to have a significant impact.

Articles by Kramer et al. and Papa and Ferreira present solutions for meeting critical scenario identification and quantification requirements. Kramer et al. propose combining existing methodologies for hazard analysis and risk assessment (HARA) with further advancements and recommend using an integrated HARA strategy early in the development phase of ADS so that the results may be included during the design and testing phases. They also propose a method for identifying and specifying hazards related to ADS. The technique consists of modelling the system design architecture and intended system design, identifying hazards, performing a causal chain analysis, deriving the triggering conditions for hazards, and quantifying and assessing the associated risks. Papa and Ferreira use a mix of exploratory and back-casting approaches to scenario building, allowing flexibility in the process and customisation for the specific situation under investigation. They define scenarios as hypothetical futures highlighting the policy quandaries and behavioural conflicts that may arise as ADSs transition from speculation to reality. They provide two alternatives, one optimistic and the other pessimistic, and identify critical decisions related to the deployment and governance of ADS. Additionally, Song et al. propose a new research strategy for selecting critical ADS test scenarios, which includes studying the system requirements and implementation of the autonomous driving functionalities, defining the system's essential parameters, describing the parameter's ratio and dispersion, and establishing feasible object operations and severity levels for the ADS, and using the modeFrontier tool to develop an optimisation technique for identifying critical test scenarios for ADS. These are the main approaches for critical scenario identification and quantification being utilised today [22],[23],[24].

2.2 Critical Scenario Prioritisation

Critical scenario prioritisation is used in ADS to prioritise tests for ADS based on diversity and cost. Critical scenario prioritisation requirements include minimising the data needed for autonomous vehicle testing while maintaining accuracy and detecting almost all inserted defects.

Several solutions have been proposed to meet the requirements of critical scenario prioritisation. Deng et al. offer STRaP, a method that enables the alignment of messages from different channels in the recording and the conversion of each frame of the aligned recording into a vector based on a driving situation data structure. Additionally, the STRaP method allows for slicing the aligned and vectorised recording into components based on the correlation of successive vectors and the prioritisation of the classes based on their coverage of driving scene aspects and rarity. Birchler et al. present two evolutionary techniques for selecting ADS experiments in virtual settings based on diversity measures derived from static road characteristics. They demonstrate that their strategy enhances the detection of security issues and that multi-objective meta-heuristics better single-objective procedures for promoting SDC tests. The solutions provided by this article, such as the Singular-Objective Genetic Algorithm (SO-SDC-Prioritizer), Multi-Objective Genetic Algorithm (MO-SDC-Prioritizer), and the Black-box Greedy Algorithm, all aim to minimise the amount of data needed for autonomous vehicle testing while maintaining accuracy and detecting almost all inserted defects [26],[27].

2.3 Critical Scenario Selection

Critical scenario selection is choosing the most appropriate scenario for a given situation using the results of scenario evaluation, quantification, and prioritisation. The goal of scenario selection is to select the scenario that offers the best balance of benefits, risks, and uncertainties based on the specific goals and objectives of the situation.

Several strategies have been presented to address the essential scenario selection requirements. Wang et al. recommended choosing the best scenario for a given condition. Riedmaier et al. distinguish between techniques based on testing and falsification in their literature review. Testing-based approaches to safety assessment involve evaluating the safety of a system by testing it under a range of scenarios selected to assess the system's safety under various conditions. Examples include N-wise sampling, Interactive Design of Experiments (DoE), automated method for regression testing, Satisfiability Modulo Theories (SMT), Signal Temporal Logic (STL), Randomization procedures, and Developing logical scenarios. These methods all ensure that the selected scenarios represent real-world conditions and can identify potential issues in the ADS. Falsification is a method of testing that seeks out instances that break safety criteria for ADSs. Examples include using real-world accident data, increasing the criticality of specific scenarios, identifying critical scenarios within established parameter ranges, using simulation-based falsification, using reinforcement-based adaptive stress testing learning, using optimisation methods such as Particle Swarm Optimization, Differential Evolution Genetic Optimization, Simulated Annealing and using machine learning models such as random forest for scenario selection [28 -85]

2.4 Critical Scenario Validation

Critical scenario validation is the process of checking scenarios to ensure that they are accurate and realistic. This can involve comparing the scenarios with real-world data and expert opinions and adjusting them as needed. The goal of scenario validation is to ensure that the scenarios being considered are based on the most precise and up-to-date data and accurately reflect the real-world situations that ADS may encounter.

Various solutions have been proposed to meet the requirements of critical scenario validation. Wang et al. propose VAAFO, an ADS validation approach that combines the benefits of test drives with virtual assessment. This approach includes a route analysis unit, a reality design adjustment unit, an evaluation unit, and a scenario record unit. Second, Elrofai et al. propose a method for efficiently and systematically evaluating the safety of advanced ADS by integrating deterministic algorithms with data science to uncover events concealed in massive volumes of the dataset. Finally, Weng et al. propose a total and effective probabilistic solution to the security assessment challenge. This approach is based on set invariance, divides the given data into separate groups, and validates each group independently. These solutions aim to improve the accuracy and reliability of critical scenario validation by providing methods for comparing scenarios to real-world data, adjusting them as needed, and evaluating the safety and security of ADS [86],[87],[88].

2.5 Critical Scenario Evaluation

Critical scenario evaluation assesses the likelihood and potential impact of different scenarios to identify the most critical ones and focus on the issues most likely to have a significant effect. Requirements for critical scenario evaluation include the ability to analyse decision-making in uncertain settings, identify the best possible actions to take in each situation, and evaluate the performance of the autonomous system in critical scenarios.

The articles provided describe various solutions for meeting these requirements. Huang et al., for example, present scenario analysis as a method for identifying critical scenarios. Van Der Pol suggests using Markov Decision Process (MDP) to analyse decision-making in uncertain settings. Sutton et al. present reinforcement learning (RL) as a machine learning approach that uses MDPs to identify the best possible actions to take in each situation. Levine et al. present REINFORCE, a specific type of RL algorithm that uses a policy gradient method called Monte Carlo sampling to find the optimal policy. Ren and colleagues present Neural architecture search (NAS) as a parameter optimisation strategy for constructing a learning algorithm for a given task. Shalev-Shwartz et al. present the Responsible and Safe Social Scenario (RSS) method, which guarantees that an operator cannot be the cause of an incident. Dosovitskiy et al. present Carla, an open-source simulator for ADS that aims to provide realistic settings and a ROS bridge to test and evaluate the system's performance in critical scenarios. These are the most commonly used approaches for critical scenario evaluation [89 - 96].

3. CONCLUSION AND FUTURE WORK

The scenario-based methodology presented in this work helps discover, measure, and prioritise relevant situations of critical scenarios to assist ADS development. As a result, researchers may better understand the capabilities and limits of ADS and devise ways to improve their safety and performance by methodically creating and analysing demanding scenarios. Future studies on this issue might include developing and testing the scenario-based approach to critical scenario analysis. This may also require conducting more complete assessments of the technique using more extensive and more diverse case datasets, as well as investigating the impact of different aspects, such as the ADSs level and the complexity of the environment, on the approach's performance. Furthermore, it may be worthwhile investigating possible applications of the technique outside of the creation of ADS, such as designing other complex systems or decision-making in dynamic situations. Overall, it is imperative to study more in this field to understand better crucial circumstances and their importance in developing sophisticated technology.

References

[1] World Health Organization. Number of Road Traffic Deaths. [Online] . Available: https://www.who.int/data/gho/data/themes/topics/topic-details/GHO/road-traffic-mortality

[2] SAE International, ``Taxonomy and definitions for terms related to driving automation systems for on-road motor vehicles," SAE Tech. Paper J3016, 2018.

[3] P. Koopman and M. Wagner, ``Challenges in autonomous vehicle testing and validation," SAE Int. J. Transp. Saf., vol. 4, no. 1, pp. 1524, 2016.

[4] P. Koopman and M. Wagner, ``Toward a framework for highly automated vehicle safety validation," in Proc. WCX World Congr. Exper., 2018, pp. 113, DOI: 10.4271/2018-01-1071.

[5] A. Shetty, M. Yu, A. Kurzhanskiy, O. Grembek, H. Tavafoghi, and P. Varaiya, "Safety challenges for autonomous vehicles in the absence of connectivity," Transportation Research Part C: Emerging Technologies, vol. 128, p. 103133, 2021.

[6] N. Kalra and S. M. Paddock, "Driving to safety: How many miles of driving would it take to demonstrate autonomous vehicle reliability?" Transportation Research Part A: Policy and Practice, vol. 94, pp. 182–193, 2016.

[7] IEEE Transactions on Intelligent Transportation Systems (2017). A survey of autonomous driving system design and development.

[8]Czarnecki, K. (2018). Operational Design Domain for Automated Driving Systems. Taxonomy of Basic Terms ", Waterloo Intelligent Systems Engineering (WISE) Lab, University of Waterloo, Canada.

[9]Klück, F., Li, Y., Nica, M., Tao, J., & Wotawa, F. (2018, October). Using ontologies for test suite generation for automated and autonomous driving functions. In 2018 IEEE International symposium on software reliability engineering workshops (ISSREW) (pp. 118-123). IEEE.

[10]Khastgir, S., Brewerton, S., Thomas, J., & Jennings, P. (2021). Systems approach to creating test scenarios for automated driving systems—reliability engineering & system safety, 215, 107610.

[11]Society of Automotive Engineers International. (2019). Autonomous driving. Retrieved from http://www.sae.org/standards/content/j3016_201901/

[12]IEEE Transactions on Intelligent Transportation Systems. (2019). Autonomous driving: Challenges and opportunities. 20(4), 1718-1731.

[13] ISO/PAS 21448:2019(E) (2019). "Road vehicles - Safety of the intended functionality ."ISO

[14] S. Geyer et al. (2014). "Concept and development of a unified ontology for generating test and use-case catalogs for

assisted and automated vehicle guidance ."In: IET Intelligent Transportation Systems, Vol. 8, Iss. 3, pp.193-189.
[15] S. Ulbrich et al. (2015). "Defining and Substantiating the Terms Scene, Situation, and Scenario for Automated Driving ."IEEE 18th International Conference on Intelligent Transportation Systems, Las Palmas, pp. 982-988.
[16]Thorn, E., Kimmel, S. C., Chaka, M., & Hamilton, B. A. (2018). A framework for automated driving system testable cases and scenarios (No. DOT HS 812 623). United States. Department of Transportation. National Highway Traffic Safety Administration.
[17]T. Menzel, G. Bagschik, and M. Maurer. Scenarios for Development, Test, and Validation of Automated Vehicles. arXiv preprint:1801.08598, pages 1–7, 2018.
[18]Khastgir, S., Brewerton, S., Thomas, J., & Jennings, P. (2021). Systems approach to creating test scenarios for automated driving systems. Reliability engineering & system safety, 215, 107610.
[19]Zhang, X., Tao, J., Tan, K., Törngren, M., Sánchez, J. M. G., Ramli, M. R., ... & Felbinger, H. (2021). Finding critical scenarios for automated driving systems: A systematic literature review. arXiv preprint arXiv:2110.08664.
[20]Colwell, I., Phan, B., Saleem, S., Salay, R., & Czarnecki, K. (2018, June). An automated vehicle safety concept based on runtime restriction of the operational design domain. In 2018 IEEE Intelligent Vehicles Symposium (IV) (pp. 1910-1917). IEEE.
[21]Aßmann, U., Zschaler, S., & Wagner, G. (2006). Ontologies, metamodels, and the model-driven paradigm. In Ontologies for software engineering and technology (pp. 249-273). Springer, Berlin, Heidelberg.
[22] Kramer, B., Neurohr, C., Büker, M., Böde, E., Fränzle, M., & Damm, W. (2020, September). Identification and quantification of hazardous scenarios for automated driving. In International Symposium on Model-Based Safety and Assessment (pp. 163-178). Springer, Cham.
[23]Papa, E., & Ferreira, A. (2018). Sustainable accessibility and the implementation of automated vehicles: Identifying critical decisions. Urban Science, 2(1), 5.
[24]Song, Q., Tan, K., Runeson, P., & Persson, S. (2022). Critical Scenario Identification for Realistic Testing of Autonomous Driving Systems.
[25]Deng, Y., Zheng, X., Zhang, M., Lou, G., & Zhang, T. (2022, November). Scenario-based test reduction and prioritisation for multi-module autonomous driving systems. In Proceedings of the 30th ACM Joint European Software Engineering Conference and Symposium on the Foundations of Software Engineering (pp. 82-93).
[26]Birchler, C., Khatiri, S., Derakhshanfar, P., Panichella, S., & Panichella, A. (2022). Single and Multi-objective Test Cases Prioritisation for Self-driving Cars in Virtual Environments. Proceedings of the ACM on Measurement and Analysis of Computing Systems.
[27]Wang, J., Pun, A., Tu, J., Manivasagam, S., Sadat, A., Casas, S., ... & Urtasun, R. (2021). Advise: Generating safety-critical scenarios for self-driving vehicles. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 9909-9918).
[28]Riedmaier, S., Ponn, T., Ludwig, D., Schick, B., & Diermeyer, F. (2020). Survey on the scenario-based safety assessment of automated vehicles. IEEE Access, 8, 87456-87477.
Just 18 deleted
[29] T. Ponn, D. Fratzke, C. Grandt, and M. Lienkamp, ``Towards certification of autonomous driving: Systematic test case generation for a comprehensive but economically-feasible assessment of lane keeping assist algorithms," in Proc. 5th Int. Conf. Vehicle Technol. Intell. Transp. Syst., 2019, pp. 333342.
[30] H. Beglerovic, A. Ravi, N. Wikström, H.-M. Koegeler, A. Leitner, and J. Holzinger, ``Model-based safety validation of the automated driving function highway pilot," in Proc. 8th Int. Munich Chassis Symp. (Proceedings), P. P. E. Pfeffer, Ed. Wiesbaden, Germany: Springer, 2017, pp. 309329.
[31] E. Rocklage, H. Kraft, A. Karatas, and J. Seewig, ``Automated scenario generation for regression testing of autonomous vehicles," in Proc. IEEE 20th Int. Conf. Intell. Transp. Syst. (ITSC), Oct. 2017, pp. 476483.
[32] B. Kim, A. Jarandikar, J. Shum, S. Shiraishi, and M. Yamaura, ``The SMT-based automatic road network generation in vehicle simulation environment," in Proc. 13th Int. Conf. Embedded Software. (EMSOFT), Oct. 2016, pp. 110.
[33] B. Kim, T. Masuda, and S. Shiraishi, ``Test specification and generation for connected and autonomous vehicle in virtual environments," ACM Trans. Cyber-Phys. Syst., vol. 4, no. 1, pp. 126, Jan. 2020.
[34] I. Majzik, O. Semerath, C. Hajdu, K. Marussy, Z. Szatmari, Z. Micskei, A. Voros, A. A. Babikian, and D. Varro, ``Towards system-level testing with coverage guarantees for autonomous vehicles," in Proc. ACM/IEEE 22nd Int. Conf. Model Driven Eng. Lang. Syst. (MODELS), Sep. 2019, pp. 8994.
[35] S. Khastgir, G. Dhadyalla, S. Birrell, S. Redmond, R. Addinall, and P. Jennings, ``Test scenario generation for driving simulators using constrained randomisation technique," in Proc. SAE World Congr. Exper. Warrendale, PA, USA, 2017.
[36] L. Huang, Q. Xia, F. Xie, H.-L. Xiu, and H. Shu, ``Study on the test scenarios of level 2 automated vehicles," in Proc. IEEE Intell. Vehicles Symp. (IV), Jun. 2018, pp. 4954.
[37] M. Althoff and J. M. Dolan, ``Reachability computation of low-order models for the safety verification of high-order road vehicle models," in Proc. Amer. Control Conf. (ACC), Jun. 2012, pp. 35593566.
[38] M. O'Kelly, A. Sinha, H. Namkoong, R. Tedrake, and J. C. Duchi, ``Scalable end-to-end autonomous vehicle testing via rare-event simulation," in Proc. Adv. Neural Inf. Process. Syst., 2018, pp. 98279838.
[39] D. Åsljung, J. Nilsson, and J. Fredriksson, ``Comparing collision threat measures for verification of autonomous vehicles using extreme value theory," IFAC-PapersOnLine, vol. 49, no. 15, pp. 5762, 2016. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S2405896316309855
[40] D. Asljung, J. Nilsson, and J. Fredriksson, ``Using extreme value theory for vehicle level safety validation and implications for autonomous vehicles," IEEE Trans. Intell. Vehicles, vol. 2, no. 4, pp. 288297, Dec. 2017.
[41] D. Zhao, ``Accelerated evaluation of automated vehicles," Ph.D. dissertation, Univ. Michigan, Ann Arbor, MI, USA, 2016.
[42] D. Zhao, H. Lam, H. Peng, S. Bao, D. J. LeBlanc, K. Nobukawa, and C. S. Pan, ``Accelerated evaluation of

automated vehicles safety in lane change scenarios based on importance sampling techniques," IEEE Trans. Intell. Transp. Syst., vol. 18, no. 3, pp. 595607, Mar. 2017.
[43] D. Zhao, X. Huang, H. Peng, H. Lam, and D. J. LeBlanc, ``Accelerated evaluation of automated vehicles in car-following maneuvers," IEEE Trans. Intell. Transp. Syst., vol. 19, no. 3, pp. 733744, Mar. 2018.
[44] Z. Huang, D. Zhao, H. Lam, D. J. LeBlanc, and H. Peng, ``Evaluation of automated vehicles in the frontal cut-in scenario An enhanced approach using piecewise mixture models," in Proc. IEEE Int. Conf. Robot. Automat. (ICRA), May/Jun. 2017, pp. 197202.
[45] Z. Huang, H. Lam, and D. Zhao, ``Sequential experimentation to efficiently test automated vehicles," in Proc. Winter Simulation Conf. (WSC), Dec. 2017, pp. 30783089.
[46] Z. Huang, H. Lam, and D. Zhao, ``Towards affordable on-track testing for autonomous vehicle Kriging-based statistical approach," in Proc. IEEE 20th Int. Conf. Intell. Transp. Syst. (ITSC), Oct. 2017, pp. 16.
[47] Z. Huang, Y. Guo, M. Arief, H. Lam, and D. Zhao, ``A versatile approach to evaluating and testing automated vehicles based on kernel methods," in Proc. Annu. Amer. Control Conf. (ACC), Jun. 2018, pp. 47964802.
[48] Z. Huang, H. Lam, D. J. LeBlanc, and D. Zhao, ``Accelerated evaluation of automated vehicles using piecewise mixture models," IEEE Trans. Intell. Transp. Syst., vol. 19, no. 9, pp. 28452855, Sep. 2018.
[49] Z. Huang, H. Lam, and D. Zhao, ``An accelerated testing approach for automated vehicles with background traffic described by joint distributions," in Proc. IEEE 20th Int. Conf. Intell. Transp. Syst. (ITSC), Oct. 2017, pp. 933938.
[50] Z. Huang, M. Arief, H. Lam, and D. Zhao, ``Synthesis of different autonomous vehicles test approaches," in Proc. 21st Int. Conf. Intell. Transp. Syst. (ITSC), Nov. 2018, pp. 20002005.
[51] Z. Huang, M. Arief, H. Lam, and D. Zhao, ``Evaluation uncertainty in data-driven self-driving testing," in Proc. IEEE Intell. Transp. Syst. Conf. (ITSC), Oct. 2019, pp. 19021907.
[52] S. Zhang, H. Peng, D. Zhao, and H. E. Tseng, ``Accelerated evaluation of autonomous vehicles in the lane change scenario based on subset simulation technique," in Proc. 21st Int. Conf. Intell. Transp. Syst. (ITSC), Nov. 2018, pp. 39353940.
[53] M. Arief, P. Glynn, and D. Zhao, ``An accelerated approach to safely and efficiently test pre-production autonomous vehicles on public streets," in Proc. 21st Int. Conf. Intell. Transp. Syst. (ITSC), Nov. 2018, pp. 20062011.
[54] M. O'Kelly, A. Sinha, H. Namkoong, R. Tedrake, and J. C. Duchi, ``Scalable end-to-end autonomous vehicle testing via rare-event simulation," in Proc. Adv. Neural Inf. Process. Syst., 2018, pp. 98279838.
[55] X. Wang, H. Peng, and D. Zhao, ``Combining reachability analysis and importance sampling for accelerated evaluation of highly automated vehicles at pedestrian crossing," in Proc. Dyn. Syst. Control Conf., 2019.
[56] S. P. Olivares, N. Rebernik, A. Eichberger, and E. Stadlober, ``Virtual stochastic testing of advanced driver assistance systems," in Advanced Microsystems for Automotive Applications. Cham, Switzerland: Springer, 2016, pp. 2535.
[57] Y. Akagi, R. Kato, S. Kitajima, J. Antona-Makoshi, and N. Uchida, ``A risk-index base sampling method to generate scenarios for the evaluation of automated driving vehicle safety*," in Proc. IEEE Intell. Transp. Syst. Conf. (ITSC), Oct. 2019, pp. 667672.
[58] D. Asljung, M. Westlund, and J. Fredriksson, ``A probabilistic framework for collision probability estimation and an analysis of the discretisation precision," in Proc. IEEE Intell. Vehicles Symp. (IV), Jun. 2019, pp. 5257.
[59] L. Stark, M. Düring, S. Schoenawa, J. E. Maschke, and C. M. Do, ``Quantifying vision zero: Crash avoidance in rural and motorway accident scenarios by combination of ACC, AEB, and LKS projected to German accident occurrence," Traffic Injury Prevention, vol. 20, no. 1, pp. S126S132, Jun. 2019.
[60] G. E. Mullins, ``Adaptive sampling methods for testing autonomous systems," Ph.D. dissertation, Univ. Maryland, College Park, MD, USA, 2018.
[61] Stark, L., Kühn, A., & Kühn, A. (2017). Deriving test cases from accident data for the safety assessment of automated vehicles. In Transportation Research Part C: Emerging Technologies (Vol. 80, pp. 134-144).
[62] Stark, L., Kühn, A., & Kühn, A. (2018). Test case selection from an accident database for the safety assessment of automated vehicles. In Transportation Research Part C: Emerging Technologies (Vol. 94, pp. 1-11).
[63] Fahrenkrog, F., Kluge, B., & Kühn, A. (2018). Deriving safety test cases from real-world accident data for the assessment of automated vehicles. In Transportation Research Part C: Emerging Technologies (Vol. 93, pp. 314-326).
[64] Geiger, M., Kühn, A., & Kühn, A. (2019). Deriving representative test cases for the safety assessment of automated vehicles from an extensive accident database. In Transportation Research Part C: Emerging Technologies (Vol. 101, pp. 199-213).
[65] A. Pierson, W. Schwarting, S. Karaman, and D. Rus, ``Learning risk level set parameters from data sets for safer driving," in Proc. IEEE Intell. Vehicles Symp. (IV), Jun. 2019, pp. 273280.
[66] M. Althoff and S. Lutz, ``Automatic generation of safety-critical test scenarios for collision avoidance of road vehicles," in Proc. IEEE Intell. Vehicles Symp. (IV), Jun. 2018, pp. 13261333.
[67] M. Klischat and M. Althoff, ``Generating critical test scenarios for automated vehicles with evolutionary algorithms," in Proc. IEEE Intell. Vehicles Symp. (IV), Jun. 2019, pp. 23522358.
[68] F. Gao, J. Duan, Y. He, and Z. Wang, ``A test scenario automatic generation strategy for intelligent driving systems," Math. Problems Eng., vol. 2019, pp. 110, Jan. 2019.
[69] Q. Xia, J. Duan, F. Gao, T. Chen, and C. Yang, ``Automatic generation method of test scenario for ADAS based on complexity," SAE Tech. Paper 2017-01-1992, 2017, DOI: 10.4271/2017-01-1992.
[70] Q. Xia, J. Duan, F. Gao, Q. Hu, and Y. He, ``Test scenario design for intelligent driving system ensuring coverage and effectiveness," Int. J. Automot. Technol., vol. 19, no. 4, pp. 751758, Aug. 2018.
[71] J. Wang, C. Zhang, Y. Liu, and Q. Zhang, ``Traffic sensory data classification by quantifying scenario complexity," in Proc. IEEE Intell. Vehicles Symp. (IV), Jun. 2018, pp. 15431548.

[72] C. Zhang, Y. Liu, Q. Zhang, and L. Wang, ``A graded offline evaluation framework for intelligent vehicle's cognitive ability," in Proc. IEEE Intell. Vehicles Symp. (IV), Jun. 2018, pp. 320325.
[73] Y. Qi, Y. Luo, K. Li, W. Kong, and Y. Wang, ``A trajectory-based method for scenario analysis and test effort reduction for highly automated vehicle," in Proc. SAE Tech. Paper Series, Warrendale, PA, USA, 2019.
[74] T. Ponn, C. Gnandt, and F. Diermeyer, ``An optimisation-based method to identify relevant scenarios for type approval of automated vehicles," in Proc. 26th Int. Tech. Conf. Enhanced Saf. Vehicles (ESV). Washington, DC, USA: National Highway Traffic Safety Administration, 2019.
[75] T. Ponn, A. Schwab, F. Diermeyer, C. Gnandt, and J. Záhorský, ``A method for the selection of challenging driving scenarios for automated vehicles based on an objective characterisation of the driving behavior," in 9. Tagung Automatisiertes Fahren. Munich, Germany: Technical Univ. Of Munich, 2019.
[76] P. Koopman and F. Fratrik, ``How many operational design domains, objects, and events?" in Proc. SafeAI, 2019, pp. 14.
[77] A. Corso, P. Du, K. Driggs-Campbell, and M. J. Kochenderfer, ``Adaptivestress testing with reward augmentation for autonomous vehicle validation," in Proc. IEEE Intell. Transp. Syst. Conf. (ITSC), Oct. 2019, pp. 163168.
[78] H. Beglerovic, M. Stolz, and M. Horn, ``Testing of autonomous vehicles using surrogate models and stochastic optimisation," in Proc. IEEE 20th Int. Conf. Intell. Transp. Syst. (ITSC), Oct. 2017, pp. 16.
[79] G. E. Mullins, ``Adaptive sampling methods for testing autonomous systems," Ph.D. dissertation, Univ. Maryland, College Park, MD, USA, 2018.
[80] C. E. Tuncali, T. P. Pavlic, and G. Fainekos, ``Utilising S-TaLiRo as an automatic test generation framework for autonomous vehicles," in Proc. IEEE 19th Int. Conf. Intell. Transp. Syst. (ITSC), Nov. 2016, pp. 14701475.
[81] C. E. Tuncali, G. Fainekos, H. Ito, and J. Kapinski, ``Simulation-based adversarial test generation for autonomous vehicles with machine learning components," in Proc. IEEE Intell. Vehicles Symp. (IV), Jun. 2018.
[82] C. E. Tuncali, G. Fainekos, D. Prokhorov, H. Ito, and J. Kapinski, ``Requirements-driven test generation for autonomous vehicles with machine learning components," IEEE Trans. Intell. Vehicles, early access, Nov. 25, 2019, DOI: 10.1109/TIV.2019.2955903.
[83] B. Gangopadhyay, S. Khastgir, S. Dey, P. Dasgupta, G. Montana, and P. Jennings, ``Identification of test cases for automated driving systems using Bayesian optimisation," in Proc. IEEE Intell. Transp. Syst. Conf. (ITSC), Oct. 2019, pp. 19611967.
[84] M. Nabhan, M. Schoenauer, Y. Tourbier, and H. Hage, ``Optimising coverage of simulated driving scenarios for the autonomous vehicle," in Proc. IEEE Int. Conf. Connected Vehicles Expo (ICCVE), Nov. 2019.
[85] Hungar, H. (2020, October). A concept of scenario space exploration with criticality coverage guarantees. In International Symposium on Leveraging Applications of Formal Methods (pp. 293-306). Springer, Cham.
[86] Wang, C., & Winner, H. (2019, October). Overcoming challenges of validation automated driving and identification of critical scenarios. In 2019 IEEE Intelligent Transportation Systems Conference (ITSC) (pp. 2639-2644). IEEE.
[87] Elrofai, H., Worm, D., & Op den Camp, O. (2016). Scenario identification for validation of automated driving functions. In Advanced Microsystems for Automotive Applications 2016 (pp. 153-163). Springer, Cham.
[88] Weng, B., Capito, L., Ozguner, U., & Redmill, K. (2021). Towards guaranteed safety assurance of automated driving systems with scenario sampling: An invariant set perspective. IEEE Transactions on Intelligent Vehicles, 7(3), 638-651.
[89] Huang, W., Wang, K., Lv, Y., & Zhu, F. (2016, November). Autonomous vehicles testing methods review. In 2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC) (pp. 163-168). IEEE.
[90] Van Der Pol, E. (2016). Deep reinforcement learning for coordination in traffic light control. Master's thesis, University of Amsterdam.
[91] R. S. Sutton and A. G. Barto, Reinforcement learning: An introduction. MIT Press, 2018.
[92] S. Levine, "Cs 294: Deep reinforcement learning," 2017.
[93] Ren, P., Xiao, Y., Chang, X., Huang, P. Y., Li, Z., Chen, X., & Wang, X. (2021). A comprehensive survey of neural architecture search: Challenges and solutions. ACM Computing Surveys (CSUR), 54(4), 1-34.
[94] S. Shalev-Shwartz, S. Shammah, and A. Shashua, "On a formal model of safe and scalable self-driving cars," CoRR, vol. abs/1708.06374, 2017. [Online]. Available: http://arxiv.org/abs/1708.06374
[95] A. Dosovitskiy, G. Ros, F. Codevilla, A. Lopez, and V. Koltun, "CARLA: An open urban driving simulator," in Proceedings of the 1st Annual Conference on Robot Learning, 2017, pp. 1–16.
[96] Karunakaran, D., Worrall, S., & Nebot, E. (2020). Efficient falsification approach for autonomous vehicle validation using a parameter optimisation technique based on reinforcement learning. arXiv preprint arXiv:2011.07699.
[97] Figure 1