Autonomous Vehicle Evaluation: A Comprehensive Survey on Modeling and Simulation Approaches

HESHAM ALGHODHAIFI, (Student Member, IEEE), SRIDHAR LAKSHMANAN
Department of Electrical and Computer Engineering, University of Michigan-Dearborn, Dearborn, MI 48128 USA
(e-mails: halghodh & lakshman@umich.edu)
Corresponding author: Hesham Alghodhaifi (e-mail: halghodh@umich.edu).

ABSTRACT In recent years, autonomous vehicles (AVs), which observe the driving environment and lead a few or all of the driving duties, have garnered tremendous success. The field of AVs has been developing rapidly and has found many applications. As a safety requirement by policymakers, these vehicles must be evaluated before their deployment. The evaluation process for AVs is challenging because crashes are rare events, and AVs can escape passing predefined test scenarios. Therefore, capturing crashes and creating real test scenarios should be considered in order to have an evaluation approach that represents the real-world scenarios. One evaluation approach is based on the naturalistic field operational test (N-FOT), in which prototype AVs are driven by volunteers or test engineers on the roads. Unfortunately, this approach is time-consuming and costly because one needs to drive thousands of miles to experience a police-reported collision and nearly millions of miles for a fatal crash. Another approach is the Accelerated Evaluation method. The core idea of the Accelerated Evaluation approach is to modify the statistics of naturalistic driving so that safety-critical events are emphasized.

This paper presents a brief survey of the advances that have occurred in the area of the evaluation of partly or fully AVs, starting with naturalistic field operational tests (N-FOTs). The review goes on to cover test matrix evaluation, worst-case scenario evaluation (WCSE), Monte Carlo simulations, and accelerated evaluation (AE). We also present all the simulation-based and agent-based modeling approaches that do not follow any evaluation protocol listed above. This study provides a scientific analysis of each of the evaluation techniques, focusing on their advantages/disadvantages, inherent restrictions, practicability, and optimality. The results reveal that the accelerated evaluation approach outperforms naturalistic field operational tests (N-FOTs), test matrix evaluation, worst-case scenario evaluation (WCSE), Monte Carlo simulations methods in some of the car-following, and lane-change studies when using specific models. Moreover, the agent-based model and augmented and virtual reality approaches show promising results in AVs evaluation. Furthermore, integrating machine and deep learning into the available AV evaluation methods can improve its performance and generate encouraging outcomes.

INDEX TERMS Accelerated Evaluation, Agent-based Modeling, Autonomous Vehicles, Evaluation, Modeling, Monte Carlo Simulations, N-FOTs, Safety, Simulation, Simulation-based Model, Test Matrix, Testing, Worst-Case Scenario, Augmented Reality, Virtual Reality.

I. INTRODUCTION

DECADES-long mobile robot navigation and more recent artificial intelligence (AI) and wireless communication advances have created the technological possibilities to make the semi-autonomous road vehicles of today possible and have brought the fully autonomous intelligent transportation systems (ITS) of tomorrow within reach. This body of research in AI also offers excellent potential to increase the efficiency and safety of future transportation substantially. Autonomous vehicles (AVs) can help to save fuel, decrease traffic crashes, reduce traffic congestion, and provide better transportation services to older people and people with disabilities [1]. There are many legal challenges in developing AVs, and an effort is being made to remove
these challenges. Since 2012, at least 41 states and D.C. have considered legislation related to AVs [2]. Twenty-nine states have proceeded with laws authorizing the testing of AVs on public roads [2]. In Arizona, Delaware, Hawaii, Idaho, Illinois, Maine, Massachusetts, Minnesota, Ohio, Washington, and Wisconsin, executive orders have been issued regarding AVs [2], [3], [4], as shown in Figure 1.

In Europe, the United Kingdom allowed the testing of AVs on public streets beginning in January 2015 [5]. On February 6, 2019, a formal statement issued by the Department for Transport (DfT) said that the UK is "on track to meet its commitment to have fully self-driving vehicles on UK roads by 2021" [6]. Table 1 shows the AVs Readiness Index for 30 countries [7], [294]. This index shows the level of preparedness for AVs [7]. It is a compound index that combines 28 individual measures from various sources into a single score [7].

The 30 countries have been evaluated on 28 measures collected into four pillars, namely policy and legislation, technology and innovation, infrastructure, and consumer acceptance [294]. Four measures are used to define the final index score for each country by KPMG International and the ESI ThoughtLab [7]. They use public data such as media reports, press releases, and other materials [7]. One of these measures utilizes a consumer survey conducted by Branded Research in every country on the list [7].

All variables are given equivalent weight except the mobile connection speed and broadband measures in the infrastructure pillar [294]. These two measures have half the weight of other measures. The collected data is normalized first before being combined using the min-max approach because every variable has a different measurement unit. The normalization method changes the variables into a range between zero and one [294]. As a result, the top-ranked country gets one and the bottom country zero. Thus, each pillar has an equivalent weight in the final score for each country [294]. Figure 2 shows these pillars with the associated variables based on [294].

As the pace of AV innovation picks up, cities have become the proving grounds of choice. Tech giants, automakers, and startups alike are focused on cities because that is where future customers live and work [8]-[13].

Many major car companies have begun research and development programs for AVs. Table 2 presents the reported AV production schemes [14], [15]. Three features are considered, namely traffic jam assist (TJA), autonomous parking assist (APA), and automated highway driving (AHD). On October 14, 2015, Tesla operated the autopilot function on the Model S through an over-the-air software update [16], enabling features such as adaptive cruise control, lane-keeping, auto lane change, autopark, and automatic emergency steering [17].

After 2016, a large number of companies entered the AVs industry such as Amazon, Apple, Microsoft, Nvidia, Tesla, Toyota, Uber, Volvo, Huawei, and many more [18], [19]. Recent developments in the area of advanced driver assistance systems (ADAS) show vast improvements toward the accessibility of autonomous driving. Many companies have raised the level of autonomy over the last few years. Several projects are targeting SAE level 4 or higher. A list of the definition of SAE levels of AVs is explained in [20] and is shown in figure 3. Advancement in autonomous driving requires high-level algorithms that are efficient enough to solve complicated scenarios, especially urban scenarios, such as intersections with multiple pedestrians, pedestrians with unknown intent, traffic lights, cars, and bicycles, which are a real challenge to predict. These high-level algorithms include pattern recognition (classification) [232], [233], [297]-[299], clustering [236, 237, 238], decision matrix algorithms [239, 240, 296], pedestrian intention prediction [234, 235] and many more algorithms.

Driving in urban environments has been both a potential and a hot area of research due to the high density of vehicles and many obstacles that must be avoided. There have been
TABLE 1: AVs Readiness Index Results [7]

| Rank | 2020 | 2019 | 2018 | Country       | 2020 Scores | 2019 Scores | 2018 Scores |
|------|------|------|------|---------------|-------------|-------------|-------------|
| 1    | 2    | 2    | Singapore | 25.45         | 24.32       | 26.08       |
| 2    | 1    | 1    | The Netherlands | 25.22         | 25.05       | 27.73       |
| 3    | 3    | n/a  | Norway     | 24.25         | 23.75       | n/a         |
| 4    | 4    | 3    | United States | 23.99         | 22.58       | 24.75       |
| 5    | 6    | 4    | Finland    | 23.58         | 22.28       | n/a         |
| 6    | 5    | 4    | Sweden     | 23.17         | 22.48       | 24.73       |
| 7    | 13   | 10   | South Korea | 22.71         | 19.79       | 20.71       |
| 8    | 9    | 8    | United Arab Emirates | 22.23 | 20.69 | 20.89 |
| 9    | 7    | 5    | United Kingdom | 21.36         | 21.58       | 23.99       |
| 10   | n/a  | n/a  | Denmark    | 21.21         | n/a         | n/a         |
| 11   | 10   | 11   | Japan      | 20.88         | 20.53       | 20.28       |
| 12   | 12   | 7    | Canada     | 20.68         | 19.80       | 22.61       |
| 13   | n/a  | n/a  | Taiwan     | 19.97         | n/a         | n/a         |
| 14   | 8    | 6    | Germany    | 19.88         | 21.15       | 22.74       |
| 15   | 15   | 14   | Australia  | 19.70         | 19.01       | 19.40       |
| 16   | 14   | n/a  | Israel     | 19.40         | 19.60       | n/a         |
| 17   | 11   | 9    | New Zealand | 19.19         | 19.87       | 20.75       |
| 18   | 16   | 12   | Austria    | 19.16         | 18.85       | 20.00       |
| 19   | 17   | 13   | France     | 18.59         | 18.46       | 19.44       |
| 20   | 20   | 16   | China      | 16.42         | 14.41       | 13.94       |
| 21   | n/a  | n/a  | Belgium    | 16.23         | n/a         | n/a         |
| 22   | 18   | 15   | Spain      | 16.15         | 15.50       | 14.58       |
| 23   | 19   | n/a  | Czech Republic | 13.99         | 14.46       | n/a         |
| 24   | n/a  | n/a  | Italy      | 12.70         | n/a         | n/a         |
| 25   | 21   | n/a  | Hungary    | 11.66         | 11.99       | n/a         |
| 26   | 22   | 18   | Russia     | 11.45         | 8.55        | 7.09        |
| 27   | n/a  | n/a  | Chile      | 11.28         | n/a         | n/a         |
| 28   | 23   | 19   | Mexico     | 7.42          | 7.73        | 6.51        |
| 29   | 24   | 20   | India      | 7.95          | 6.87        | 6.14        |
| 30   | 25   | 17   | Brazil     | 5.49          | 6.41        | 7.17        |

TABLE 2: Announced autonomous vehicle technologies until 2016 [14], [15]

| OEM         | Automation Level | Launch Year | First Model | TJA | APA | AHD | Region of Introduction                  |
|-------------|------------------|-------------|-------------|-----|-----|-----|-----------------------------------------|
| BMW         | Semi-Highly      | 2014–2017   | X5 - 5/7-series | Yes | Yes | Yes | Europe, North America                  |
| Mercedes Benz| Semi-Highly     | 2014–2017   | E-Class S-Class | Yes | Yes | Yes | Europe, North America, Australia        |
| Audi        | Semi-Highly      | 2016        | A8          | Yes | No  | Yes | Europe & North America                 |
| Volkswagen  | Semi-Highly      | 2017        | Passat      | Yes | No  | Yes | Europe & North America                 |
| General Motors (Cadillac) | Semi-Highly | 2018 >2020 | SRX, Cadillac CT6 ATS & XTS | No  | Yes | Yes | North America                           |
| Ford        | Semi-Highly      | 2017 - 2020 | Fusion Explorer | Yes | Yes | Yes | Europe & North America                 |
| Lexus       | Highly           | 2020        | LS          | Yes | Yes | Yes | Japan, Europe & North America           |
| Nissan      | Highly           | 2020        | Leaf        | Yes | Yes | Yes | Europe & North America                 |
| Volvo       | Semi            | 2015        | XC90        | Yes | Yes | Yes | Europe & North America                 |
| Google      | Fully            | 2018        | Model/OEM agnostic | Yes | Yes | Yes | North America                           |

Several in-depth efforts to study this problem, such as the DARPA Urban Challenge [21], the V-Charge Project [22], and at least three US military applications—urban operations (UO), manned-unmanned teaming (MUM-T), and AGR [23]. The challenge of driving in an urban environment is complex because it considers increasing the speed of AVs and environmental complexity [24].

By increasing the level of automation, the evaluation process becomes challenging because the AV system will become more complex. An AV may have 100 million lines of code, while the Boeing 787 has only 6.5 million [28] (Figure 4). It is a real challenge for a company and also for the evaluation authorities, such as the National Highway Traffic Safety Administration (NHTSA) [30], to check every line of code. Many problems may be uncovered after the product release, which could cost the company a lot of money [31], [32]. It is, therefore, necessary to evaluate the AV system...
during the design process. In this paper, we focus on the evaluation of AVs in level 3 to level 5.

The term "autonomous vehicle" is used in this article instead of "automated vehicle." We have chosen to utilize "autonomous" because it is a common term, and the general public is familiar with it. However, the term "automated" implies control or operation by a machine and refers to connected vehicles, while "autonomous" have more intelligence than the term "automated" and suggests the vehicle is acting independently [295]. The following sections show several AV evaluation methods.

II. NATURALISTIC FIELD OPERATIONAL TESTS

Naturalistic field operational tests (N-FOTs) [33] have been used to evaluate AVs. In this test, several equipped vehicles are deployed on the road and are driven in naturalistic conditions [34]. During the driving time, the data are collected for evaluation purposes. An advantage of naturalistic field operational tests (N-FOTs) observation is that it allows the investigators to directly observe the CAVs and AVs in a natural setting. A naturalistic driving study of 100 vehicles was implemented by Virginia Polytechnic Institute and Virginia Tech to investigate the major contributing components of crashes. The collected data had been used to inspect many elements, such as driver performance, surrounding environment, driving conditions, and other components, that are related to critical incidents, near collisions, and collisions [35]–[39]. A list of large-scale N-FOT projects carried out in the U.S. is shown in table 3.

Some companies, such as Waymo (formerly the Google self-driving car project), have designed several SAE Level 4 AVs [41] and evaluated the whole autonomous system on actual roads since 2012. Up to Jan 2020, the Waymo AVs have logged nearly 20 million miles of self-driven operation on public roads in 25 cities and tens of billions of miles through computer simulations, with thousands of scenarios and different individual test tracks [42], [304]. In the N-FOT test, the drivers are trained and know where to drive. Thus, the evaluation process involves non-intrusive driving conditions. The N-FOT test approach has many restrictions, such as the time required to conduct this test. In addition, the probability to expect critical events under naturalistic conditions is very low. For example, in the U.S., the vehicles should travel a total of 0.53 million miles for a police-reported collision and 99 million miles for a fatal collision [43]. Therefore, the N-FOT projects require many vehicles, a lot of time, and large budgets. In [44], it is noted that an N-FOT “cannot be conducted with less than $10,000,000.” An efficient test approach for AV evaluation is required.

III. TEST MATRIX EVALUATION

A test matrix evaluation is defined as a series of test scenarios that are defined at the start of the process. Then, the autonomous vehicles go through each test and are assessed objectively or subjectively [40]. Figure 5 shows an example of a test matrix evaluation process. In figure 5, the design cycle starts with using specific scenarios or cases. Then, the functional and technical specifications are constructed from these cases. The final design is then verified at a component and function level. In the evaluation cycle, the function description is established using the functional and technical specifications. Then, three types of tests are applied to the constructed function. These tests are potential safety tests, human factors, and technical tests. At the end of the process, a validation test and a safety impact analysis are implemented.

Test matrix evaluation scenarios can be applied in field tests, hardware-in-the-loop (HIL) tests, driving simulator tests, and computer simulation [40]. Field tests are utilized by all certification authorities [40]. Driving simulator tests and computer simulations have also been used to decrease the cost and time. The Test Matrix evaluations are mainly based on crash databases. The pre-crash scenarios are investigated in many studies [241]–[245]. Figure 6 shows the General Motors 44-crashes typology. The United States Department of Transportation designed the pre-crash typology based on the NASS crash databases GES (General Estimates System) [246] and CDS (Crashworthiness Data System) [247]. Volpe National Transportation Systems Center integrated these two typologies to create the 37 pre-crash scenarios and capture the vehicle movements and dynamics in real-world and pre-crash critical scenarios. The top five scenario groups are namely, car-following, lane change, left turn, crossing, and opposite direction, were generated by Volpe using the GES, NMVCCS (National Motor Vehicle Crash Causation Survey) [248], and EDR (Event Data Recorder) [249] databases.
Table 3: Major N-FOT projects in the U.S. [40][304]

| Name                          | Conductor | Period       | Mileage [mile] | Vehicle | Sensor                  | Drivers                           | Research Topic                                      |
|-------------------------------|-----------|--------------|----------------|---------|-------------------------|-----------------------------------|----------------------------------------------------|
| 100-Car Naturalistic Driving Study | VT        | 2001–2009    | 2,000,000      | 100 sedans | Camera                  | 109 primary drivers, 100 secondary drivers | Rear end collision                                |
| ACAS                          | UM        | 2004–2005    | 137,000        | 11 sedans | Camera, Radar           | 96 drivers                        | Forward collision warning                        |
| RDCW                          | UM        | 2005–2006    | 83,000         | 11 sedans | Camera, Radar           | 11 drivers                        | Lane departure warning                           |
| SeMiFOT                       | UM        | 2008–2009    | 106,528        | 10 sedans, 4 trucks | Camera, Radars | 39 drivers                        | Forward collision warning, lane departure warning, blind spot information systems, electronic stability control, and impairment warning |
| IVBSS                         | UM        | 2010–2011    | N/A            | sedans: 213,309; trucks: 601,944 | Camera, DSRC | 108 drivers for sedans, 19 professional truck drivers | Integrated warning                                |
| SPMD                          | UM        | 2012–2014    | Over 34 millions | 2,800 vehicles of various types | Camera, Radar, DSRC | 2,700 volunteer drivers and several professional bus and truck drivers | Connected vehicle                                 |
| Waymo-(Google driverless car) | Google    | 2009–present | Over 20 millions | At least 50 sedans and SUVs | Lidar, Camera, Radar | Google technicians and volunteers | Fully self-driven vehicle                         |

FIGURE 6: Pre-crash scenarios defined by NHTSA [244].

Table 4 presents the major crash databases in the USA and Europe.

For more details about crash analysis, reference [250] presents extensive reviews about it. Test matrix forms scenarios from the data acquired from the NFOTS and acquired by technical document analysis [251], [252], [253]. Several programs and research projects have been started to develop evaluation policies using the test matrix technique, such as the collision scenarios designed in Crash Avoidance Metrics Partnership (CAMP) [252], the critical scenarios created through the classification tree method for ADAS [253], and the scenarios constructed based on ontologies [254]. Table 5 shows all the test matrix projects. The significant advantages of the test matrix technique are that the determined test policy is repeatable, well-grounded, and quick to finish [46]. Nevertheless, several challenges come to the table, such as all the test scenarios being predetermined. Thus, the AV control system can attain excellent results in these test scenarios, but the results under real-life scenarios may not be satisfying. In an analogy, “Having a standard test is akin to holding an SAT exam for students with all problems pre-announced. Students do well in the test, but the score may tell very little about how much they learn” [45].

Besides, the test matrix scenarios are usually chosen from collision databases in which most of the collisions were caused by human-controlled vehicles (HVs). Therefore, the test scenarios and the evaluation processes applied to AVs may not accurately capture the safety-critical events of AVs [40]. Moreover, according to CAMP, ADAS, and ontologies projects, the results indicated that the test matrix is more appropriate in autonomous driving system evaluation with the availability of low cost and high controllability of scenarios. However, the generation of test scenarios using the traditional test matrix approach is usually based on few influence factors. These factors are generally integrated simultaneously to generate an ultimate scenario [252]. The influence factors are usually taken from the NFOTS database, WCS database, test standards, and many more. The influence factors can be then divided into surrounding environment parameters, AV parameters, and road users’ parameters [131], [132], [40], [40].
TABLE 4: Crash Databases in the USA and Europe – Databases with Larger 5000 Crashes [40]

| Database                                           | Country     | Crashes       | Data Years     | Owner                      |
|----------------------------------------------------|-------------|---------------|----------------|----------------------------|
| National Automotive Sampling System - Crashworthiness Data System (NASS-CDS [247]) | USA         | 3300-5000 per year | 1988-present | NHTSA                      |
| National Automotive Sampling System - General Estimated System (NASS-GES [246]) | USA         | 50,000 per year   | 1988-present | NHTSA                      |
| Fatal Accidents Recording System (FARS [256])        | USA         | 33,000 per year   | 1975-present | NHTSA                      |
| National Motor Vehicle Crash Causation Survey (NMVCCS [257]) | USA         | 5,470           | 2005-2007     | NHTSA                      |
| Community Road Accident Database (CARE [258])        | EU & CH, IS, NO | 1,000,000+   | 1991-present | European Commission.       |
| German In-Depth Accident Study (GIDAS [259])        | DE          | 22000          | 1999-present | BASf and several manufacturers. |
| ADAC Accident Investigation Study (ADAC [260])       | DE          | 11456          | 2015-present | ADAC.                     |
| Co-operative Crash Injury Study (CCIS [261])        | UK          | 15000          | 1983-2010    | Department for Transport.  |
| Vehicle Occupant Infrastructure Etudes de la Sécurité des Usagers de la Route (VOIESUR [262]) | FR          | 9000           | 2011          |                            |

TABLE 5: Projects Studying the Test Matrix Technique [40]

| Project                                           | Institute                                      | Year | Test scenario                                                                 | Evaluation approach                  | Basis                                      |
|----------------------------------------------------|-----------------------------------------------|------|-------------------------------------------------------------------------------|--------------------------------------|--------------------------------------------|
| CAMP                                               | Ford/GM/NHTSA                                  | 1999 | 26 tests to evaluate FCW system performance.                                 | Test track.                          | GM “44 Crashes.” Real-world scenarios.     |
| HASTE                                              | HASTE consortium                               | 2005 | Three levels of road complexities.                                           | Simulator test track.                | HMI conflict scenarios in previous research. |
| AIDE                                               | AIDE consortium                                | 2005 | Three levels of road complexities.                                           | Simulator.                           | National statistics combined with in-depth crash data. |
| TRACE                                              | LMS, LAB, INRETS, VW, TNO, ALLIANZ             | 2006 | Scenarios from crash database.                                                | Simulator test track, computer simulation. | SAVE-U COMPOSE PReVENT INVENT project GIDAS data. |
| Pre-Crash Scenario Typology for Crash Avoidance Research | Volpe                                        | 2007 | 37 pre-crash scenarios for light vehicles.                                   | Crash data analysis.                 | GES (General Estimates System).            |
| APROSYS                                            | APROSYS consortium                             | 2007 | Scenarios for cars, heavy trucks, motorcyclists, pedestrians, and pedal cyclists. | Simulator Test track.                |                                          |
| CICAS                                              | U of Minnesota, PATH, VTech                   | 2008 | Scenarios for intersection conflict.                                          | Test track.                          |                                          |
| ASSESS                                             | ASSESS consortium                              | 2010 | Stopped lead vehicle, slow lead vehicle with constant speed, braking lead vehicle with constant deceleration. | Test track.                          | GIDAS.                                    |
| interactIVe                                        | interactIVe consortium                        | 2011 | Multiple scenarios for active intervention systems.                          | Hardware-in-the-loop testing, simulator test track. | Previous studies based on real traffic crashes. |
| Accident data study in support of development of Autonomous Emergency Braking (AEB) test procedures | Loughborough University                      | 2012 | EURO NCAP AEB test procedure.                                                 | Test track.                          | STARS 19 (2008) OTS (2000-2009).          |
| Depiction of priority light-vehicle pre-crash scenarios for safety applications based on vehicle-to-vehicle communications | Volpe                                        | 2013 | 37 pre-crash scenarios for light vehicle impact.                             | Crash data analysis.                 | GES NMVCCS EDR.                           |
| Development of Performance Evaluation Procedures for Active Safety Systems | UMTRI                                         | 2013 | 11 scenarios to assess DBS and CIB.                                           | Test track.                          | Real traffic accidents.                    |

An accident analysis report by Tesla revealed that the faults of a critical system might be generated by mixtures of particular values of some factors. Moreover, test scenarios [255]. Thus, adding additional factors will show a geometric growth in the number of scenarios and, as a result, increase the test cost.
that integrate these mixtures of values can help in the evaluation process by revealing new problems [255]. Therefore, using the traditional test matrix approach that is based on the exhaustive testing of all influence factors is redundant and ineffective.

Furthermore, specific test scenarios should be generated using certain influence factors (elements) to evaluate AV systems. For example, the generated test scenarios to evaluate the Lane Departure Warning (LDW) system take into consideration the traffic environment parameters, subject vehicle driver’s behavior, and traffic participants’ state [263]. The traffic environment parameters are subdivided into lighting environment (weather, time, and rapid changes in light), lane lines parameters (lane lines clarity, lane line integrity, lane line number, lane line color, etc.), and road parameters (curvature, lane number, lane marks, slope, and roadside facilities). Every parameter, such as weather condition, has many values such as sunny, cloudy, rainy, and foggy. To design test scenarios according to ISO 17361, the total test scenarios are only eight [8], which are not enough to find the system failure. According to the test matrix method using the exhaustive testing approach, the total test scenarios are 497,664,000 [264]. Assuming an average running time of 30 s for each scenario, a total time of 473 years is required to finish all scenarios [263]. It is undoubtedly an inefficient and unacceptable testing approach.

Authors in [263] proposed a Combinatorial Testing Scenario Generation Method Based on Complexity (CTBC) to improve the effectiveness of the traditional test matrix technique. The presented method considers decreasing the number of test scenarios and improving the overall complexity of the scenarios. The results revealed that scenarios with high complexity are effective in finding system failures. Moreover, the CTBC method reduces the number of test scenarios and generates more complex scenarios than the traditional test matrix methods. On the other hand, each AV system has unique influence elements, and the coupling relationship between these elements from different systems is not investigated yet. Therefore, the system’s defects under the coupling relationship condition are not explored yet by the CTBC method. Thus, many AV systems, subsystems, and advanced features are not yet tested and evaluated by the CTBC method and traditional test matrix methods.

**IV. WORST-CASE SCENARIO EVALUATION**

The worst-case scenario evaluation (WCSE) technique is suggested to recognize the highly challenging scenarios for any cars, with or without active control systems [40]. In [47], [48], Ma tries WCSE on rollover (overturning of a vehicle) and jackknifing of articulated cars using a dynamic game theory. The term Jackknifing refers to the folding of an articulated vehicle so that it resembles the acute angle of a folding pocket knife. This approach suggests that the control inputs and disturbance inputs take part in a two-player game condition. In [49], Ungoren proposes another approach as a one-player game by considering the car and its control system as a joint dynamic structure. Then, for solving the WCSE problem numerically, the iterative dynamic programming technique is applied [40]. This technique is used in [50] to assess the integrated chassis control (ICC) system. Therefore, a mathematical model of the vehicle is established, and the WCSE is defined as a horizon optimization problem to resolve for a trajectory (e.g., a sequence of steering inputs) that minimizes or maximizes the cost function (e.g., rollover index) [50]. A solution for two different systems is conducted. For a linear system (SISO time-invariant and MIMO systems), the worst bounded inputs are acquired from the convolution of impulse responses [51]. For a nonlinear system (nonlinear dynamical system for a control problem), the solution of the Hamilton-Jacobi-Bellman equations is obtained by the calculus of variations to resolve the optimal trajectory task [52].

Even though the WCSE approach can recognize the weakness of a car and a car control system, it does not take into account the occurrence probability of such worst-case scenarios [53], [54]. Therefore, the WCSE results do not provide sufficient data about critical real-world situations. Furthermore, there are some limitations when using complicated control algorithms or when the control algorithms are not in numerical form. As a result, the WCSE techniques may either face difficulties finding the worst scenarios or be time-consuming.

**V. MONTE CARLO SIMULATIONS**

Monte Carlo simulation or Monte Carlo method is a mathematical procedure utilized to predict the likely results of an unpredictable event. The Monte Carlo simulation creates a model of potential outcomes by leveraging a probability distribution (uniform or normal distribution etc.) for any variable with uncertainty [300]. In this approach of AV evaluation, the N-FOTs data are used to construct stochastic models, and Monte Carlo simulation is applied to assess partly or fully AVs. Table 6 presents a list of all the papers related to this method, with the objectives, techniques & models, and scenarios.

In addition to table 6, table 7 presents all the references with the associated AV tasks.

In [55], collision avoidance systems are evaluated by establishing an "errorable" driver model to mimic human distraction based on road-departure crash-warning (RDCW) FOT and intelligent cruise control (ICC) FOT naturalistic driving databases. In [56], heavy trucks’ collision warning and collision mitigation braking technologies are assessed by building 1.5 million forward-collision test scenarios from naturalistic driving data conflicts. The main advantage of this approach is that naturalistic driving data are used to create all the scenarios/models. Therefore, these scenarios/models represent real-world scenarios. As a result, Monte Carlo simulation models may decrease the assessment cost compared to the field tests. Moreover, this method evaluates data collected from human driving databases without any actual crashes [55], [56]. Therefore, using Monte Carlo simulations
TABLE 6: Monte Carlo simulations studies: objectives, techniques and models, and scenarios

| Ref # | Objectives | Techniques and Models | Scenarios |
|-------|------------|----------------------|-----------|
| [55]  | Evaluate the performance of several Collision warning/collision avoidance (CW/CA) algorithms using an errable driver model. | Errable driver model: perceptual limitation, time delay, and distraction. Stochastic driver model, and a forward multi-step prediction. | Car-following scenario. |
| [56]  | Evaluate the safety performance of Commercial Vehicle Forward Collision Avoidance and Mitigation Systems (F-CAM). | Collision warning/collision mitigation braking technologies. Naturalistic driving data conflicts to build test scenarios. | Rear-end crashes using five crash threat scenarios: lead vehicle moving slower at a constant speed, lead vehicle decelerating, lane change (cut-in/out), lead vehicle is not moving when the ego vehicle’s radar detects the lead vehicle. |
| [265] | Compare the use of Markov Chain Abstraction and Monte Carlo Simulation for the safety evaluation of fully AVs. | Markov chain abstraction & Monte Carlo simulation. | Car-following scenarios. |
| [266] | Extend the adaptive stress evaluation technique to test the AVs and compare Markov decision process (MDP) and Monte Carlo tree search (MCTS). | Markov decision process (MDP) and Monte Carlo tree search (MCTS). | An AV with noisy sensors approaching a pedestrian crosswalk. |
| [268] | Investigate the behavior of eight tracking controllers under extreme situations, uncertain parameters, and sensor noises. | Monte Carlo simulations & Rapid Exploring Random Trees (RRTs). | Single and double lane change scenarios. |
| [269] | Compare the univariate Gaussian Probability Density Functions to Monte Carlo simulations in estimating the future discrete state transitions. | Univariate Gaussian Probability Density & Monte Carlo simulations. Uncertain mental model. | A vehicle crossing the path of an AV in a roundabout. An autonomous UAS approaching a busy airfield with the intention of landing. |
| [270] | Utilize model predictive control and its optimization function to find a smoother trajectory. Use Monte Carlo simulations as a safety assessment. | Model predictive control and Monte Carlo simulations. | Lane-change on the straight road and turning at the intersection. |
| [277] | Analyze the safety of complex road scenes using a reasoning framework and Monte Carlo integration. | Reasoning framework, Monte Carlo integration, and Runga-Kutta-Fehlberg algorithm. | A road scene with a parked car and a moving bicycle. Vehicle approaching an obstacle at a high velocity. Two vehicles are moving, with one approaching the other from behind with a higher velocity. |
| [278] | Present a threat evaluation algorithm for general road scenes. | A dynamic driver model, Monte Carlo sampling, and an extended Kalman filter. | Lane change scenario. |
| [279] | Present an algorithm to predict the driver intention of other vehicles using a random-forests classifier. Compute possible future trajectories with a sequential Monte Carlo method. | A sequential Monte Carlo method. | Race road with four sharp curves with and without obstacles. |
| [284] | Propose a reinforcement learning-based deep-MCTS algorithm for vision-based autonomous driving control. | Deep-MCTS. | A highway scenario consisting of six agents (five are part of the environment). |
| [285] | Propose a risk-based framework to evaluate AV systems and subsystems as a black box. | Multilevel splitting method and adaptive importance-sampling methods. | A highway scenario. |
| [286] | Implement a simulation framework to evaluate a fully AV system using adaptive importance-sampling methods. | Adaptive importance-sampling methods. | A Mixed Urban Route. |
| [287] | Proposes an Enhanced Driver Model (EDM) that predicts the driver action in an urban environment. Using the EDM model, a Monte Carlo simulation is used to identify the statistical distribution of fuel consumption and travel time. | EDM. | A race track at Halmstad University Sports Center. |
| [293] | Implement and validate an autonomous agent based on a Monte Carlo Tree Search | MCTS. | |

directly may result in an inefficient simulation model because of the dominance of the non-safety-critical portions of the naturalistic driving data.

In [55], the model coefficients are usually acquired by fitting a collection of driving data. Then, the errable model can be improved in real-time to achieve a higher level of false positives and false negatives for better modification to the driver. However, according to [55] and [56], tuning the models under open-loop usually show inefficient results with human or hardware-in-the-loop. Therefore, when the hardware or human is in the loop, this method may not speed up the evaluation procedure.

The security problem in the autonomous driving domain, especially trajectory planning, is investigated heavily in the literature reviews. It is essential to estimate the surrounding road users and then predict the probabilistic occupancy of...
each road user to identify any future risk [271]. Achieving accurate estimations help the AV navigates to the final destination with the least crash probability. According to some literature reviews, it is commonly proposed that all road users keep their initial motion state [272], [273]. Therefore, the actual motion of the road user might be different from the estimated motion because of the uncertainty of the road user detection and future intent. In literature reviews, there are many estimation techniques based on kinematic models or dynamic models [274] – [276]. The kinematic models have some limitations in neglecting the forces that influence the road user movement [274]. At the same time, the dynamic models consider the variety of forces that affect driving, such as tire forces and air friction. Using dynamic models is excessively complex. Moreover, it requires to have a different model for each vehicle [270]. Therefore, Monte Carlo simulation [55], [56], [265], [266], [268], [270], [277], [279], Gaussian distribution [280], [281], and Markov chain abstraction [265], [282], [283] are usually used to tackle the above issue.

In [270], the motion prediction of road users is studied. The authors focus on trajectory planning in two typical lane-change scenarios (lane change on a straight road and turning at intersections). The Monte Carlo simulation is used as a safety assessment to estimate the probabilistic path planning of the road users and then produce a map from probability statistics to actual scenarios. Furthermore, Monte Carlo simulations have a limitation because of the probabilistic errors from the random sampling [265]. According to literature reviews, it is required to use more samples to achieve accurate outcomes by Monte Carlo simulations. However, using more samples will accumulate more errors, and the results will not be accurate [265], [270]. The outcomes in [270] revealed that Monte Carlo simulation is inefficient in real-time computation.

Authors in [265] compared Markov Chain Abstraction and Monte Carlo Simulation for the safety evaluation of fully AVs. The two methods have common differences in terms of their error sources. The significant errors in the Markov chain approach are only the systematic errors from the discretization of the state and input space [265]. Moreover, the Markov chain has no probabilistic errors since no random sampling is implemented [265]. In the Monte Carlo method, the main errors are probabilistic errors from sampling the initial states and the input sequences [265].

Furthermore, the Monte Carlo simulation has no systematic errors because each simulation uses the system’s main dynamical equations [265]. The outcomes show that the Markov Chain resulting probability distributions outperformed the Monte Carlo simulation approach in terms of accuracy and simulation speed [265]. On the other hand, the Monte Carlo simulation produces superior outcomes than the Markov Chain method when computing the crash probabilities [265]. Thus, the outstanding performance of Monte Carlo simulations in crash probabilities is due to the absence of systematic errors in the Monte Carlo simulation.

Koren et al. extended the adaptive stress evaluation technique that was used to evaluate the aircraft collision avoidance system to test the AVs [266]. Adaptive stress testing (AST) is an approach to find the critical scenarios using a Markov decision process (MDP) [266]. The original AST approach utilizes a Monte Carlo tree search (MCTS) with double progressive widening (DPW) to explore any failure in the system [267]. The test scenario in [266] includes an AV with noisy sensors approaching a pedestrian crosswalk. The deep reinforcement learning solver is used to improve the efficiency of AST instead of using the Monte Carlo tree search (MCTS). Authors claimed that the deep reinforcement learning approach is more efficient and can discover more critical events than the Monte Carlo tree search [266]. The results reveal that both methods can recognize the failure trajectories in an AV-pedestrian conflict. The two solvers produce many events where the AV hits the pedestrian. A major problem with the MCTS is that MCTS has non-zero noise that adds up over time. As a result, the MCTS does not reduce this noise to zero, leading to a considerable amount of probability error with time. Thus, the AV will not detect and predict the pedestrian and result in a critical collision in the

| Reference | Collision Avoidance | Safety Assessment | Trajectory & Prediction | Driver Behavior | Control | End-to-End Testing | Fuel Consumption |
|-----------|---------------------|-------------------|--------------------------|-----------------|--------|-------------------|-----------------|
| [55]      | ✓                   | -                 | -                        | -               | -      | -                 | -               |
| [56]      | ✓                   | -                 | -                        | -               | -      | -                 | -               |
| [265]     | -                   | ✓                 | -                        | -               | -      | -                 | -               |
| [266]     | -                   | ✓                 | -                        | -               | -      | -                 | -               |
| [268]     | -                   | ✓                 | ✓                        | -               | -      | -                 | -               |
| [269]     | -                   | -                 | ✓                        | -               | -      | -                 | -               |
| [270]     | -                   | -                 | ✓                        | -               | -      | -                 | -               |
| [277]     | -                   | -                 | ✓                        | -               | -      | -                 | -               |
| [278]     | -                   | ✓                 | ✓                        | -               | -      | -                 | -               |
| [279]     | -                   | -                 | ✓                        | -               | -      | -                 | -               |
| [284]     | -                   | ✓                 | -                        | ✓               | -      | -                 | -               |
| [285]     | -                   | -                 | ✓                        | -               | -      | -                 | -               |
| [286]     | -                   | -                 | ✓                        | -               | -      | -                 | -               |
| [287]     | -                   | ✓                 | -                        | ✓               | -      | -                 | ✓               |
| [293]     | -                   | ✓                 | -                        | -               | -      | -                 | -               |

TABLE 7: Summary of the Main Monte Carlo Studies with the Associated Tasks.
generated scenarios. Table 8 shows the numerical results of the two solvers of adaptive stress testing. In table 8, the number of calls to STEP for MCTS is the needed number of calls to find a critical accident in the AV-pedestrian conflict. In other words, it refers to the algorithm’s computational capability to find a critical conflict. Moreover, this number of calls represents the required number of calls to trust the results. The presented approach in [266] consists of three scenarios shown in figure 7. In scenario 1, the generated event by MCTS and DRL send one pedestrian into the scene towards the AV to establish a conflict. DRL approach produces a straightforward path for a pedestrian towards the vehicle, which is better than the MCTS method. In short, according to the results presented in [266], the DRL solver for AST outperformed the MCTS solver, especially in higher-dimensional scenarios.

![Generated scenarios for MCTS and DRL](image)

**FIGURE 7:** The generated scenarios for MCTS and DRL. The collision point is the point of conflict between the AV and the pedestrian. In scenario 3, pedestrian 1 trajectory is far from the collision point [266].

Jönsson et al. implemented and validated an autonomous agent based on a Monte Carlo Tree Search [293]. Three action generators and two reward functions are compared. The outcomes revealed that the MCTS performs well and converges to a driving agent in static conditions. Moreover, the results showed that the MCTS only succeeds at low speeds in real-time driving [293].

Reference [268] investigates the behavior of eight tracking controllers under extreme situations, uncertain parameters, and sensor noises. Different tests from the single and double lane change scenarios are generated to evaluate the tracking controllers. Monte Carlo simulations and Rapid Exploring Random Trees (RRTs) are utilized to assess the controllers’ average and worst-case performance [268]. The authors state that most controllers’ properties (e.g., stability, noise rejection, robustness to model variations) are not strongly compromised during the turning phase [268]. Moreover, the authors concluded that the obtained outcomes by Monte Carlo simulations and Rapid Exploring Random Trees (RRTs) may not be perfect but can help choose the suitable controllers.

In [269], the univariate Gaussian Probability Density Functions are used to estimate future discrete state transitions such as the beginning of a turn by other agents. Then, the outcomes are compared to the Monte Carlo simulations. The results showed a remarkable correlation between the proposed prediction distributions and the Monte Carlo simulations, especially over long prediction horizons [269]. Although the outcomes revealed an excellent correlation between the two methods, more investigations and validations are required for this model.

Reference [278] presents a risk evaluation algorithm for public road scenes:

1) The driver behavior is modeled as a probabilistic prior.
2) A Monte Carlo sampling is used to approximate the distribution of future scenarios.
3) Different safety measures are computed based on the distribution of future scenarios.
4) A variety of techniques are implemented to increase the performance of this algorithm.

The results showed that the algorithm was tested on simulated data and sensor data and was able to differentiate between the safe and non-safe road scenes. However, the dataset used in this paper is not enough to achieve optimal risk assessment because the data does not have a lot of variation. Moreover, this algorithm removes samples with conflicts with other objects and replaces them with non-conflict samples. Thus, a real-time assessment is required to judge the efficiency of this algorithm.

Reference [279] captures the driver behavior of other vehicles using a random-forests classifier. Then, the likely future trajectories are computed with a sequential Monte Carlo simulation followed by the possible risk assessment. This method is tested by conducting numerical simulations. The simulation results exposed that the algorithm was able to recognize the driver’s behavior. However, a limitation to this method is that no real-time deployment and evaluation is conducted.

Reference [277] presents a reasoning framework for the future movement of multiple road users. The probability distribution of every vehicle’s future motion is generated by Monte Carlo planning. The Synthetic data that is based on a real-world scenario is used to test this approach. The suggested approach shows excellent outcomes but requires more improvements and validation to handle more V2V and V2P complex scenarios.

Reference [284] develops a reinforcement learning-based Monte Carlo Tree Search (deep-MCTS) control technique for an AV vision-based system. Two deep neural networks (DNNs) are utilized to predict action probabilities and then fed to deep-MCTS to reconstruct multiple future trajectories. The deep-MCTS method outperforms existing methods and shows 50.0%, 66.30%, and 59.06% improvement in training efficiency, steering control stability, and driving trajectory stability, respectively [284]. The deep-MCTS is evaluated using the US and the Torcs simulators.

In [285], the authors present a simulation testing platform to evaluate the whole AV system as a black box. The multilevel splitting method and adaptive importance-sampling
methods are used to address the shortcomings of naïve Monte Carlo simulations in estimating rare event probabilities. The approach in [285] outperforms naïve Monte Carlo for events with a probability lower than $10^{-3}$. Moreover, the variance of the failure probability is decreased by up to 10x [285]. In contrast, naïve Monte Carlo performs well comparing to the above method in predicting non-rare events [285].

In [286], O’Kelly et al. implemented an end-to-end AV testing framework using adaptive importance-sampling methods to speed up rare-event probability validation. As a result, the system validation is accelerated by 2-20 times compared to naïve Monte Carlo methods and 10-300p times (where P is the number of processors) over real-world evaluation [286].

Reference [287] proposes an Enhanced Driver Model (EDM) that predicts the driver action in an urban environment. The effects of Signal Phasing and Timing (SPaT) are considered by presenting the concept of Line-of-Sight (LOS). Signal Phasing and Timing (SPaT) data gives information on signal states by motion [288]. Detailed studies about SPaT and LOS are presented in [288]-[292]. The EDM model is then validated against data collected from equipped vehicles with different drivers. Using the EDM model, a Monte Carlo simulation is used to identify the statistical distribution of fuel consumption and travel time under other states such as traffic conditions, SPaT, and driver behavior. This study evaluates the influence of uncertainties related to real-world driving in fuel consumption in connected vehicles [287].

In short, based on the literature reviews presented in this section, Monte Carlo simulations perform well in evaluating AV in some scenarios and fail to produce outstanding outcomes in other scenarios compared to other methods. Thus, more improvements are required to this method using new techniques such as deep learning and reinforcement learning.

### VI. ACCELERATED EVALUATION

Zhao in [40] proposed the accelerated evaluation test. The main objective of this test is to establish a method that can speed up the AV evaluation course of action. Moreover, this method can precisely show AVs’ real-life safety benefits. The main idea of the accelerated evaluation approach is to reduce the evaluation process time and eliminate the safe parts of daily driving by skewing the statistics of the surrounding vehicles. This process consists of many steps, as shown below:

- Collect a massive amount of real-world driving data.
- Take out events that have possible conflicts between AVs and surrounding human-driven vehicles.
- Model the surrounding human-driven vehicle behaviors as the main distraction to AV. Then, a modeling of the randomness as random variables vector $x$ with probabilistic distribution $f(x)$ is conducted.
- Skew the disturbance statistics to minimize the safe part of daily driving by replacing $f(x)$ with the accelerated distribution $f^*(x)$.
- Run Monte Carlo tests with the accelerated probability density function $f^*(x)$. The results will provide more intense interactions/collisions between AVs and human-driven vehicles.
- “Skew back” the outcomes of the accelerated tests to understand the performance of AVs under real-life driving scenarios using the statistical analysis.

Figure 8 shows the accelerated evaluation procedure.

![Figure 8: Concept of the Accelerated Evaluation technique](image)

The suggested method can be used in computer simulations, human-in-the-loop tests with driving simulators, hardware-in-the-loop tests, or vehicle tests. Four procedures were established in this approach to shape the foundation of the accelerated evaluation idea. The first technique depends on a likelihood analysis of naturalistic driving. A probabilistic model approach based on time series driving data was used to build the test scenarios. The assessment policy is sped up by decreasing the relatively safe events that are highly likely to occur. The second technique gives a mathematical base for the "skewing back" mechanism depending on the importance sampling theory, such that the statistical equivalence between the accelerated tests and the naturalistic driving tests can be rigorously demonstrated. The third technique is adaptive accelerated evaluation. This technique shows a policy to recursively discover the best way to skew...
the probabilistic density functions of human-driven vehicles to maximally decrease the evaluation duration. Finally, the accelerated evaluation method for analyzing the dynamic interactions between AVs and human-driven vehicles was established based on stochastic optimization procedures.

In [59], three indicators – crash, injury, and conflict rates – are calculated to test the accelerated evaluation’s accuracy and performance. The crash and conflict cases are binary events, and the injury event is modeled as a probabilistic function [59]. Two types of simulations are performed: accelerated evaluation and naturalistic driving simulations (non-accelerated, based on Monte Carlo). Table 9 shows the accelerated rates of crash, injury, and conflict events in car-following scenarios [59]. The $N_{\text{nature}}$, $N_{\text{acc}}$, $r_{\text{acc}}$ in table 9 represent the number of naturalistic driving simulations, the number of accelerated tests, and the accelerated rate, respectively. In the crash and injury events, the accelerated approach expedites the evaluation by five orders of magnitude [59]. In the conflict event, the acceleration rate is lower by 300 times [59]. Table 10 shows the accelerated rates of the crash, injury, and conflict events in lane change scenarios [57]. The $D_{\text{nature}}$, $D_{\text{acc}}$, and $r_{\text{acc}}$ in table 10 represents the driving distance needed in naturalistic test, the driving distance required in accelerated test, and the accelerated rate, respectively.

The simulation outcomes in car-following and lane change scenarios in [59] and [57] revealed that the accelerated tests could decrease the assessment time of the collision, injury, or conflict events by 300 to 100,000 times. Otherwise stated, driving for 1,000 miles can show the AVs challenging scenarios that would take 300 thousand to 100 million miles to encounter in the real world [57], [59]. As a result, the development and validation time for AVs will be reduced.

Table 11 presents a list of all the papers related to this method, with the objectives, techniques & models, and scenarios.

Even though the accelerated evaluation methods can produce excellent results and reduce the duration of the evaluation, they do not take into account the following tasks:

- The AVs to AVs and human-driven vehicles to human-driven vehicles interactions are not studied and are used only as a benchmark.
- The AV sensors and controls are suggested to work accurately. Thus, the measurements and perception errors, and control are assumed to be accurate.
- AVs are assumed to look like human-driven vehicles. Therefore, human drivers’ reactions to AVs are the same as to the other human-driven vehicles.
- The secondary impacts of crashes are not considered in these methods.
- The human-driven vehicles model is not accurate enough to mimic real-world scenarios.
- Many real-world scenarios are not investigated and developed in this technique, such as sensing/detection fail scenarios (e.g., fog, snow, low light), perception fails scenarios (e.g., hand gesture, eye contact, blinking lights), vehicles/pedestrians/pedal-cyclists conflicts scenarios (e.g., running, red light, cut-in, jaywalk), making-decisions scenarios (e.g., low confidence, multiple threats), and so on.

Therefore, more improvements for this method are required to have a well-rounded and efficient technique to evaluate AVs.

VII. SIMULATION-BASED MODEL APPROACH

The goal of autonomous driving is mainly to decrease the number of deadly accidents in a highly uncertain environment as well as to provide a high quality of comfort and efficiency and create unprecedented intelligent transportation for individuals within cities. In the interest of getting the AV to navigate safely and dependably in uncertain environments, many challenges need to be considered. Modeling AVs is one of these challenges and is regarded as an essential step toward accurately validating AVs in highly uncertain environments. Then, the interaction between AVs and the surrounding vehicles or vulnerable road users should be investigated and validated in various real-world scenarios. A well-established validation approach is required to fill many gaps in the AVs evaluation process during the design, pre-deployment, and deployment stages. However, based on the previous AV validation techniques, the real-world data are limited, and many safety-critical scenarios are hard to capture in real life. Therefore, the simulation-based model approach is introduced to tackle these challenges.

In 1934, Greenshield introduced the first traffic model [307]. Since 1934, three major model classes have been presented: microscopic, macroscopic, and mesoscopic models depending upon the level of details needed for network analysis [305]. The microscopic models study the behavior and interaction of individual vehicles based on car-following, lane-changing, and gap-acceptance algorithms [305] [306]. The microscopic models are used to model the sophisticated urban street networks, intersections, vulnerable road users (VRU) movements, traffic lights, multi-model systems, and many more. The macroscopic models represent traffic as a continuous sequence and are based on the relationships of the flow, speed, and density of the traffic stream [306]. The macroscopic models focus on modeling large-scale traffic networks such as freeways, corridors, surface-street grid networks, and rural highways on a section-by-section basis instead of following individual vehicles [306]. Moreover, the macro simulation-based model approach requires a traffic assignment policy, which can be implemented by utilizing activity-based models [70]–[73] or modified traditional four-step models [74]. Finally, the mesoscopic models integrate the properties of micro and macroscopic simulation models and allow less fidelity than the micro models for individual vehicles [306]. An example of mesoscopic simulation studies can be found in [310]–[312].

The focus of this section is mainly on the connected and AVs’ micro simulation-based model studies that consider longitudinal and lateral dynamics. The micro-simulation-
TABLE 9: Accelerated Rates of Crash, Conflict and Injury Events – Car-Following Scenarios [59].

| Method                  | Crash        | Injury       | Conflict     |
|-------------------------|--------------|--------------|--------------|
| $N_{nature}$            | $4.30 \times 10^8$ | $4.20 \times 10^8$ | $1.07 \times 10^6$ |
| $N_{acc}$               | $3.84 \times 10^3$  | $3.1 \times 10^3$  | $3.26 \times 10^3$  |
| $r_{acc} = N_{nature}/N_{acc}$ | $1.12 \times 10^5$  | $1.35 \times 10^5$  | $3.28 \times 10^2$  |

TABLE 10: Accelerated Rates of Crash, Conflict and Injury Events – Lane Change Scenarios [57].

| Method                  | Crash        | Injury       | Conflict     |
|-------------------------|--------------|--------------|--------------|
| $D_{nature}[\text{mile}]$ | $4.53 \times 10^4$ | $4.71 \times 10^5$ | $4.70 \times 10^7$ |
| $D_{acc}[\text{mile}]$  | 16.4         | $4.02 \times 10^3$ | $2.53 \times 10^3$ |
| $r_{acc} = D_{nature}/D_{acc}$ | $2.77 \times 10^3$ | $1.17 \times 10^4$ | $1.86 \times 10^4$ |

TABLE 11: Accelerated evaluation studies: objectives, techniques and models, and scenarios

| Ref # | Objectives                                                                 | Techniques and Models                                                                 | Scenarios                                                                 |
|-------|---------------------------------------------------------------------------|--------------------------------------------------------------------------------------|--------------------------------------------------------------------------|
| [40]  | Accelerated evaluation framework for highly A V s.                        | Likelihood analysis, importance sampling techniques, adaptive accelerated evaluation, multi-step stochastic sampling. | Lane change scenarios and car-following scenarios.                       |
| [57], [58] | Accelerated evaluation of AV safety.                                     | Importance sampling techniques.                                                      | Lane changes scenarios. The main crash type is the frontal collision due to unsafe cut-ins. |
| [59], [60] | Accelerated evaluation of AVs safety.                                   | Extracted naturalistic driving data are used to build the statistics of the motion of the primary other vehicles (POV). | Car-following scenarios.                                                 |
| [61]  | Accelerated evaluation of AVs safety.                                    | Piecewise mixture models.                                                            | Lane - change scenarios.                                                 |
| [62]  | Accelerated evaluation of AVs safety.                                    | Piecewise mixture distribution models.                                              | Lane - change scenarios.                                                 |
| [63]  | An Accelerated testing approach for AVs.                                 | Joint statistical models, accelerated distributions for Gaussian mixture models using importance sampling techniques. | Lane - change scenarios.                                                 |
| [64]  | Accelerated evaluation and validation of complicated control systems within AVs. | Kernel methods.                                                                     | Lane - change scenarios.                                                 |
| [65]  | Accelerated evaluation of AVs: rare-event simulation.                    | Learning-based approach.                                                            | Unknown rare-event sets.                                                 |
| [66]  | Accelerated evaluation of AVs: rare-event.                               | Statistical learning models.                                                        | Rare-event sets.                                                         |
| [67]  | Evaluation uncertainty in data-driven self-driving testing.              | classical bootstrap method with likelihood ratio-based scheme.                     | Lane change test scenario.                                               |
| [68]  | On-track testing for AV.                                                 | Kriging-based statistical approach.                                                 | Lane change scenario.                                                   |
| [69]  | An accurate, affordable, and safe way to evaluate a design of an AV.     | Co-Kriging-based statistical model.                                                 | Lane change test scenario.                                               |
| [302] | A rare-event simulation for neural network and random forest predictors is presented. | Importance Sampling, Sequential Mixed Integer Programming, Random Forest, Neural Network, Deep-Learning-Based Importance Sampling, Lazy-learner Importance Sampling, Cross entropy method (CE), adaptive multilevel splitting (AMS), and naive Monte Carlo (NMC). | Rare-event sets.                                                        |
| [303] | Present a Deep Probabilistic Accelerated Evaluation (Deep-PrAE) to estimate rare-event probabilities in safety-critical applications. |                                                                                       | Rare-events: car-following scenarios.                                    |

based model produces valuable data for AVs’ future development based on the level of details in the model. Moreover, the interactions of AVs with human-driven vehicles and vulnerable road users (VRU) are presented and discussed. In addition, the agent-based simulation modeling of AVs is presented.

A. MICRO-SIMULATION-BASED MODELS

In microscopic simulation models, the AV’s behaviors can be modeled by adopting the available human drivers’ models or by inventing new intelligent models that consider V2X capabilities [110, 313]. Moreover, the car-following model is required in collaboration with the driver models or any new innovative model for V2X communications to represent how the simulated vehicles interact with the vehicle ahead [110]. Therefore, the car-following model is an essential part of modeling the behavior of human-driven vehicles (HVs), connected autonomous vehicles (CAVs), and AVs in micro-simulation modeling [110, 313]. A car-following model can be designed based on these assumptions by keeping a safe distance between the lead and host vehicles and controlling the vehicles’ speed and accelerations [135].

This section presents all the available car-following models for CAVs and AVs in micro-simulation modeling in lateral and longitudinal directions. Then, the AV simulation platforms are taken into the spotlight. Finally, the agent-based modeling studies are covered.
1) Car-Following Models for HVs, CAVs, and AVs

The car-following, lane change and distance headway are the three significant behaviors of any vehicle in micro-simulation modeling. Figure 9 shows an illustration of these major behaviors. In addition, the interaction between vehicles on the road is determined by many factors such as lateral vehicle maneuver, driver behavior, and surrounding vehicles behavior [317]. This section discusses the longitudinal and lateral dynamics for human-driven vehicles (HV), CAVs, and AVs. The following vehicle will follow the lead vehicle with proper distance headway, speed, and intended acceleration or deceleration in the longitudinal scenario. In the lateral scenario, a lane-change maneuver is performed, as shown in figure 9. This chapter covers all the significant contributions in microscopic analysis of traffic flow and safety evaluation and how the diverse traffic flow modeling has been presented and developed from homogeneous microscopic modeling to mimic the real-world environment.

![Figure 9: Major Behaviors of Any Vehicle in Micro-simulation Modeling.](image)

In micro-simulation modeling, all vehicle parameters such as the max, min, intended acceleration or deceleration, and desired speed values are defined using many statistical distributions and functions [315]. In 1945, researchers introduced vehicle trajectories for car-following modeling instead of using speeds and distance between the two vehicles [314]. They proposed the safe driving distance between the lead and following vehicles that the following vehicle driver must maintain. The car-following models can then be applied after determining the safe gap between the lead and the following vehicle in every scenario. The popular car-following models are GHR models, safe distance models, Intelligent Driver Model (IDM), ACC CACC Models, optimal velocity models, psychophysical models, fuzzy logic models, and cellular automata (CA) models. The car-following models are classified into types or categories depending on the utilized logic [316]. For example, in [318], the authors presented five different classes while other researchers, such as [319], suggested three more types to include the 21st-century models. Figure 10 illustrates the available car-following models.

The following subsections will describe each type of car-following model.

1.1 GHR Model

In 1958, Chandler introduced the first version of the GHR model to find the relative velocity between two vehicles, known as a stimulus [324]. In 1961, General Motors (GM) presented a car-following model using the acceleration/deceleration values as a stimulus [328,329]. The GM model used the speed of the leader and follower vehicles to estimate the acceleration/deceleration values. Then, the estimated acceleration/deceleration values are used to calculate the driver’s reaction time. The GM model is a simple linear car-following model with a constant sensitivity parameter, and the acceleration of the following vehicle can be approximated [324, 330]. In this model, the gap between the leader and follower vehicles affects the stimulus, making the GM linear model impractical [315]. Moreover, this model does not consider the driver’s acceleration and deceleration in the lane-change maneuvers [325]. Furthermore, the diversity of vehicles on the road is neglected in the reaction time calculation [326, 327]. Therefore, the GM and GHR models are failed to address the drivers and vehicles’ diverse conditions [321]. Other restrictions in these models are the non-availability of the acceleration/deceleration limits [320]. Recently, many researchers suggested a wide range of solutions to expand the original models and overcome some GM model limitations. For example, in [322], the following driver is permitted to accelerate if the relative speed of the lead vehicle is increased. Moreover, the authors also introduced various acceleration and deceleration parameters to improve the original models in making an efficient decision. The authors in [323] also proposed an extension to speed up the driver reaction under deceleration scenarios as compared to the acceleration cases. Furthermore, the authors in [331] suggested a critical headway value to estimate the state of driving behavior. Finally, another extension to the GM model is presented in [329] to consider the non-linear behavior in terms of relative speed and distance between the lead and following vehicles. Despite all the extensions, the GM model has its behavioral limitations, such as the drivers’ reactions to random changes in the stimulus, and the actions of the leader vehicle keep impacting the relative speed and the driver of the following vehicles even when the gap between these vehicles
are high [315]. Table 12 shows an outline of the GHR model studies.

1.2 Safety-Distance or Collision Avoidance Model

The safety-distance model keeps a safe distance between the lead and follower vehicle. This model is based on the fundamental motion equation [318]. Pipes explains the term of a safety distance in 1953 [112] as “a good rule for following another vehicle at a safe distance is to allow yourself at least the length of a car (about 15 ft) between you and the vehicle ahead for every 10 miles of hour speed at which you are traveling”. In 1981, Gipps presented the first acceleration model that documents car-following and non-car-following maneuvers [339]. However, a limitation to the Gipps model is that the model requires keeping the safe-distance headway and not exceeding the desired speed. As a result, many researchers performed extensive extensions, modifications, and calibrations studies to the Gipps model [340-342]. Table 13 shows a list of these studies. Nowadays, the Gipps model is widely used in micro-simulations modeling because of its basic calibration assumptions about human driving behavior [315].

1.3 Psychophysical Models

The psychophysical or action point model [318] utilizes space headway and relative velocity for the following vehicle as a threshold. An effort is taken by the drivers when the threshold values are reached. The threshold values of the spacing or relative velocity should be reached to see a reaction from the drivers. The psychophysical model is also helping to record if the drivers are paying attention to the small spacing and the associated effect when there is a large spacing on the following behavior [330]. In [343], the psychophysical model is implemented to a simulation platform using a framework called “MISSION.” Moreover, the interaction between the lead and follower vehicles is investigated by defining four thresholds and regimes, as shown in figure 11.

![FIGURE 11: Wiedemann Model - Thresholds and Regimes.](image-url)

The drivers’ behavior in the psychophysical model is suggested to be naturally distributed and can be represented as normal distributions [343]. That means that each driver has unique driving skills for perception, reaction, and prediction of the surrounding environment. The authors in [343] also proposed that the vehicles have different abilities to perform simple techniques such as max velocity and max acceleration/deceleration values. The Wiedemann model in [343] presented various ranges of other random parameters to be used to calculate threshold values and driving functions. An example of these parameters are, namely, the desired distance (AX), the desired minimum following distance (ABX), the maximum following distance (SDX), the perception threshold (SDX), and the decreasing and increasing speed differences (CLDV, OPDV) [315]. More improvements are required to this model because the calibration of its parameters is challenging [315, 321]. Examples of simulators using psychophysical models include VISSIM and PARAMICS. The VISSIM platform is widely used to model heterogeneous traffic conditions. However, using VISSIM to simulate 2D traffic conditions is not recommended because it produces inefficient outcomes [321].

1.4 Optimal Velocity Model (OVM)

The optimal velocity model (OVM) is proposed to model the traffic flow instabilities that cause congestion in public roads. The OVM model is developed to describe the dynamical behavior of traffic flow using the motion equation of each vehicle. The model is based on the relative distance to the lead vehicle. Moreover, the acceleration of the following vehicle is controlled in a way that the final velocity is modified according to the trajectory of the lead vehicle [321]. Several modifications and extensions are proposed to overcome some of the limitations of the OVM approach for modeling the mixed traffic conditions. The significant improvement to the OVM approach is the two-velocity difference model (TVDM) [345]. The TVDM approach is developed to integrate the intelligent transport system (ITS) with the OVM approach. The integration of the two models provides a comprehensive car-following model that incorporates multiple leading vehicles [321]. Table 14 shows a list of all the OVM modification and extension studies.

1.5 Intelligent Driver Model (IDM), Adaptive Cruise Control (ACC), and Cooperative Adaptive Cruise Control (CACC) Models

These days, with the advancement of the intelligent driving assistance system (IDAS), vehicles have become more intelligent and are expected to perform many driving tasks. The cruise control (CC) system is an early step in connected AVs’ intelligent driving assistance system. More improvements are added to the cruise control (CC) system to have more advanced methods such as the adaptive cruise control (ACC) and cooperative adaptive cruise control (CACC) systems. Intelligent driving assistance systems such as CC, ACC, and CACC are essential for supporting acceleration control for longitudinal motions based on the gap distance and the speed difference between the lead and host vehicles. Therefore, many researchers utilize the simulation-based model approach to evaluate and study the impacts of connected and AVs. Furthermore, the simulation-based model analysis provides the flexibility to build safety-critical scenarios and validate the AVs during their development to avoid mistakes before public road deployment. Many of the micro-simulation papers in this section established their own ACC, CACC, AV, or CAV car-following models. Furthermore, each article has implemented a unique method and produced distinct performance indicators. Figure 12 presents the four...
TABLE 12: GHR Model studies.

| Ref # | Objective | Vehicle Types | Year | Parameters | Main Results |
|-------|-----------|---------------|------|------------|--------------|
| [324] | Investigate car-following scenarios on a highway (without passing). | Cars | 1958 | Relative velocity, and acceleration. | The acceleration of a car that is attempting to follow a leader is proportional to the difference in velocity of the two cars. |
| [332] | Include vehicle mix in stimulus-response car-following models. | Cars, trucks | 2016 | Acceleration/deceleration, driver response time lag, movement state, and vehicle type for different pairs of leader-followers (car-truck, truck-car, and car-car). | The speed, relative speed, space headway parameters are estimated. |
| [333] | Present a latent leader acceleration model. | LMV, HMV, 2W. | 2016 | Acceleration/deceleration, relative speed, headway. | Predicting acceleration behavior in presence of weak lane discipline where the subject driver may have multiple vehicles in its front and there may not be a distinct leader vehicle. |

*Table legend: LMV(Private car, Microbus, Human Hauler, CNG Auto-rickshaw), HMV(SUV, Bus, Truck), 2W(Motorcycle, Bicycle, Cycle Rickshaw).*

TABLE 13: Safety-Distance Model studies.

| Ref # | Objective | Vehicle Types | Year | Parameters | Simulator | Main Results |
|-------|-----------|---------------|------|------------|-----------|--------------|
| [334] | Incorporate lateral discomfort in car-following modeling. | Cars | 2007 | Acceleration, deceleration, vehicle widths/lengths, desired speeds, lateral separation, frictional clearance, and veering distance. | C++ | Estimating speed, flow, density diagrams, space-time trajectories. |
| [335] | Develop a simulation model with various types of vehicular interactions, including lane-change, and perform certain validation tests. | Cars | 2009 | Lateral separation, frictional clearance, veering distance, maximum escape speed. | C++ | Estimating speed, flow, density diagrams, space-time trajectories. |
| [336] | Propose a full velocity difference model for traffic flow, including driver’s reaction time delay. | Cars | 2015 | Driver’s reaction time delay, headway, velocity, velocity difference, and traffic flow. | Matlab | Estimating the classification of traffic jams and space-time diagrams. |
| [337] | Extend the widely used Gipps’s car-following model to include vehicle-type dependent parameters. | 3W, trucks, buses. | 2011 | Driver reaction time, acceleration, desired speed of follower vehicle, size of leader vehicle, vehicle type. | C | Estimating the follower behavior and suggest the need of incorporating vehicle-type combination specific parameters into traffic simulation models. Providing a robust behavioral model to simulate Asian traffic conditions. Estimating the capacity and jam density. |
| [338] | Developed a new mixed-flow model for the micro simulator of Aimsun. | 3W, trucks, buses. | 2015 | Driver reaction time, acceleration, desired speed of follower vehicle, size of leader vehicle, vehicle type. | AIMSUN | Estimating the classification of traffic jams and space-time diagrams. |

main intelligent vehicle types and their definitions.

The ACC system extends the existing CC system to incorporate a headway sensor that observes the distance between the host vehicle and the vehicle in front of it. The essential function of the CC system is to keep a constant vehicle speed adjusted by the driver. In contrast, the ACC system’s principal purpose is to control vehicles’ acceleration based on a distance gap and a speed difference between lead and host vehicles. Moreover, the ACC system can accelerate or decelerate based on the lead vehicles’ speed changes. Figure 13 shows the ACC scenario.

FIGURE 12: AVs Definition and their Categorizations [101].

FIGURE 13: Adaptive Cruise Control [103].
TABLE 14: Optimal Velocity Model (OVM) Studies.

| Ref# | Objective | Vehicle Types | Year | Parameters | Simulator | Main Results |
|------|-----------|---------------|------|------------|-----------|--------------|
| [344] | Presents a non-lane-based car following model by incorporating the effects of the lane width in traffic. | Cars | 2010 | Lane width, lateral separation, critical headway. | Numerical simulation | Estimating the evolution of traffic congestion. Lane width effects on the car-following model can lower critical headway. The lateral separation effects greatly enhance the realism of car-following models. |
| [346] | Evaluate the vehicle type-dependent car-following heterogeneity from micro- and macro-aspects by using Next Generation Simulation (NGSIM) trajectory data. | Cars | 2019 | Time headway (TH), time to collision (TTC), and safety margin (SM)). | Numerical simulation | During the close car-following process with the same speed, the driver generally keeps a higher gap distance with the leader when following or driving a larger vehicle. |
| [347] | Develop a new staggered car-following model taking into consideration lateral separation effects. | Cars | 2012 | Time-to-collision, visual angle variables, lateral separation distance. | Numerical simulation | Incorporating lateral separation effects into the car-following model leads to the suppression of traffic jams and greatly enhances the realism of the model. |
| [348] | Propose a new car-following model that considers the effects of two-sided lateral gaps on a road without lane discipline. | Cars | 2015 | Lateral gaps, Two-sided lateral separation. | Numerical simulation | Considering the lateral separation and overtaking expectation in the model can better simulate car-following behavior, especially in some complex driving conditions. |
| [349] | Present a new model that considers the effects of lateral separation and overtaking expectation based on OVM. | Cars | 2016 | Lateral headway, escape corridor. | Numerical simulation | |

Furthermore, communication capabilities are added to the ACC system. The modified ACC system with V2V and V2I communications is called the CACC system. The CACC system shares the acceleration, deceleration, braking capability, and vehicle positions using V2V and V2I communications [75], [102]. The communications capabilities of the CACC provide shorter headway time compared to the ACC. Figure 14 shows the CACC scenario setup.

FIGURE 14: Cooperative Adaptive Cruise Control [102].

The ultimate goal of the intelligent driving assistance system (IDAS) is to control vehicles fully. Connected AVs have all the AV functions along with V2V and V2X communications. Figure 15 shows the CAV and AV scenarios. One significant difference between CACC and CAV is the automated lateral movement. The standard car-following motions established for human-driven vehicles are old-fashioned compared to the CC, ACC, CACC, and CAV. Thus, the literature studies’ related terms and approaches are slightly different and are not profoundly classified.

Table 15 presents the simulation-based modeling studies of ACC and CACC and their validation analysis. This table provides the objective of each review, the base model, modeling and validation scenarios, vehicle types, assessment basis, and primary outcomes.

FIGURE 15: CAVs & AVs [308, 350].

According to literature reviews, the IDM and MIXIC models are used widely as benchmark car-following models. Some modifications to the IDM [106, 78, 79, 95] and MIXIC [94, 116, 133] models are added to understand AVs’ longitudinal motion fully. The car-following model is an essential model that is used widely in simulating AVs. Research efforts have been conducted to establish AV car-following models by improving the traditional car-following models (IDM and MIXIC). The intelligent driver model (IDM) [134] is one of the uncomplicated and safe models that produce practical outcomes [135]. The main goal for developing the intelligent driver model (IDM) is to tackle the modeling of the mixed traffic conditions. The initial development of the IDM was done by Treiber et al. [134] for a single-lane scenario. Furthermore, the acceleration is described as a function of
### TABLE 15: Simulation-based AVs studies: ACC and CACC studies [101]

| Ref # | Objectives | Base Model(s) | Scenarios | Vehicle Types | Evaluation Criteria | Main Results |
|-------|------------|---------------|-----------|---------------|--------------------|--------------|
| [77]  | Investigate the ACC vehicles’ lane-changing effects compared to manual vehicles. | Pipes [112], ACC [113], and comprehensive modal emissions model (CMEM). | Position of ACC vehicles and Market-penetration rate of ACC (5%, 10%, 15% and 30%). | Manual vehicle, ACC, and CACC | Fuel consumption and emission (CO, HC, NOx) | The smooth response of the ACC vehicles has a beneficial effect on the environment. These benefits vary with the levels of the disturbance, the position of the ACC vehicle in the string of manually driven vehicles, and the ACC vehicle penetration. |
| [78]  | Present the new ACC car-following model with its impact analysis. | IDM with constant-acceleration heuristic (CAH). | Market-penetration rate of ACC (10%, 20%, 30%, 40%, and 50%). | Manual vehicle and ACC | Throughput | 1% more ACC vehicles will lead to an increase in the road capacities of about 0.3%. |
| [79]  | Present the ACC-based traffic assistance system intended to improve traffic flow and road capacity. | IDM | Market-penetration rate of ACC (0%, 5%, 15% and 25%). | Manual vehicle and ACC | Throughput | A small proportion (5%) of ACC vehicles can improve the traffic flow. An increasing proportion of ACC vehicles reduces traffic congestion. |
| [80]  | Establish the ACC and CACC car-following models and estimate their impact. | An error-based control law for ACC and CACC. The lane change is under human control. | A 100% market-penetration rate of each vehicle type | Manual vehicle, ACC, and CACC | Throughput | Throughput of the manual, ACC, and CACC vehicles was, respectively, 2,050, 2,200, and 4,550 vehicles/h. |
| [82]  | Approximate the effect on highway capacity of varying market penetrations of vehicles with ACC and the CACC. | NSGIM model [111]. ACCs: Proprietary to Nissan. CACCs: Car-following behavior[83]. | The ACC and CACC vehicles 10% increase proportion. | Manual vehicle, ACC, and Here-I-Am (HIA) vehicle | Highway throughput | The use of ACC was unlikely to change lane capacity significantly. The CACC can increase capacity greatly after its market penetration reaches moderate to high percentages. The capacity benefits of CACC can be accelerated at somewhat lower market penetrations if the rest of the vehicles are equipped with VADs. |
| [86]  | Investigate the impact of the CACC vehicle string operation on the capacity of multi-lane highway with merging bottlenecks. | ACC and CACC car-following models [83]. | Market-penetration rate of CACC (0%, 20%, 40%, 60%, 80% and 100%). | Manual vehicle and CACC | Bottleneck capacity | The freeway capacity increases quadratically as the CACC increases, with a maximum of 3080 vehicles/hour/lane at 100% CACC penetration. A rapid increase in bottleneck capacity from 80% to 100% CACC penetration, especially with high on-ramp inputs. |
| [87]  | Extend the CACC to incorporate new algorithms describing the interactions between CACC and manual vehicles. | The CACC model reported in [114]. The anticipatory lane change (ALC) for lane changing. | Market-penetration rate of CACC (0%, 20%, 40%, 60%, 80% and 100%). | Manual vehicle and CACC | Throughput | Freeway capacity is 90% higher at a 100% CACC penetration compared to 0%. The capacity increase is insignificant under low to medium CACC market penetrations (e.g., 20–60%) in the absence of additional management strategies. |
| [105] | Inspect the CACC vehicles’ impacts on traffic flow characteristics. | IDM. | Arrival rate scenarios: 7,000/veh (moderate), 8,000/veh (saturated), 9,000/veh (oversaturated), 10,000/veh (oversaturated). | Manual vehicle and CACC | Throughput | At a low-to-moderate penetration rate of CACC, the CACC impact is not statistically significant. A very large improvement is noticed at a high penetration rate of CACC, especially in high traffic conditions. |
| [106] | Present an ACC-based traffic assistance system. | IDM. | Market-penetration rate of ACC (0%, 5%, 15% and 25%). | Manual vehicle and ACC | Throughput | ACC vehicles improve the traffic stability and the road capacity. A penetration rate of 25% ACC eliminates traffic congestion during simulation. |
| [110] | Examine the impact of the CACC on traffic-flow characteristics. | MIXIC model. | Market-penetration rate of CACC (0%, 20%, 40%, 60%, 80% and 100%). | Manual vehicle and CACC | Throughput | CACC can improve traffic-flow characteristics. A low market-penetration rate of CACC (< 40%) would not have an impact on the throughput. |
| [119] | Present an up-to-date algorithm for CACC systems for collaborative driving. | Effective CACC (ECACC) algorithm | Market-penetration rate of CACC (0%, 20%, 40%, 60%, 80% and 100%). | Manual vehicle and CACC | Throughput | The congestion reduction is higher when the market-penetration rate of the CACC-equipped vehicles increases. At a low penetration rate, the effect of the CACC on traffic dynamics is not significant. |
| [119] | Present an up-to-date algorithm for CACC systems for collaborative driving. | Effective CACC (ECACC) algorithm | Market-penetration rate of CACC (0%, 20%, 40%, 60%, 80% and 100%). | Manual vehicle and CACC | Throughput | The IDM controller in the experimental test vehicles does not perceptibly follow the speed changes of the preceding vehicle. Strings of consecutive ACC vehicles are unstable, amplifying the speed variations of preceding vehicles. Strings of consecutive CACC vehicles overcome these limitations, providing smooth and stable car-following responses. |
the gap, velocity, and space difference between the lead and following vehicles [321]. The maximum acceleration and minimum headway are considered to maintain the minimum gap and acquire the required velocity. Additional extensions to the model are required to cover the multilane traffic modeling and consider the potential risk elements [135]. The IDM and linear approaches do not support the modeling and validation of 2D traffic scenarios [321]. The modified IDM can be used as the ACC model or as a human-driven vehicle model. Furthermore, the improved version of the IDM can be utilized to simulate connected autonomous vehicles (CAVs) [95].

To fully understand and estimate the impact of AVs using a simulation-based model, autonomous modeling should be able to examine the performance of the AV under highly uncertain conditions. Moreover, the model should evaluate AV safety, fuel consumption and emissions, noise emission, and traffic performance. The MICroscopic model for simulation of intelligent cruise control (MIXIC) is then suggested and developed as a stochastic simulation model to overcome these challenges. The MIXIC is implemented widely for cooperative AVs simulations because this model uses V2V communication and can optimize traffic capacity under real-world conditions. The MIXIC allows interaction between the lead and host vehicles to share the actual speed, acceleration, maximum potential braking, and warnings. In [94], the CAV is established based on the MIXIC model. This technique utilizes smart-micro automotive radar (UMRR-00 Type 30) (90 m ± 2.5% detection range and ± 35 horizontal FOV) as an input for the MIXIC model. The sensor update rate is 50 ms and can track up to 64 objects. The AV speed should be low enough to apply a complete stop when the lead vehicle is detected and reached a full stop. Using the maximum deceleration of the AV (host) and lead vehicle, the AV’s maximum safe speed and acceleration and the minimum safe distance can be calculated. In [89], the authors presented a hardware-in-the-loop (HIL) testing system for the CAV applications. The results showed the effectiveness of the CACC in absorbing certain disturbances and oscillations of speeds. Moreover, speed oscillation decreases as vehicle position in the string increases. In addition, perfect communication/radar contributes to string stability.

1.6 Cellular Automata Model

In 1992, the cellular automata (CA) model was presented by Negal and Schreckenberg [351]. The road segment in this model is divided into cells with the same size of almost 7.5 meters long [351]. Each cell can fit a single vehicle or can be unoccupied. The longitudinal dynamics of the vehicles are integrated into the CA model by including the acceleration, braking time, and randomization of vehicle types [321]. The CA model is then extended to include the two-lane traffic conditions [352]. As a result, large-scale traffic dynamic modeling is achieved easily using this model. A limitation to this model is the loss of information due to the discretization of cells [321]. Due to the discretization of cells that have the same size, vehicles are required to update their parameters, such as velocity and acceleration/deceleration in multiple cells. Another limitation is the difficulty of representing all vehicle types within cells. The cell size is an essential factor in the CA model. The representation of all vehicle types cannot be achieved if a large cell is utilized. Moreover, the computational workload might be increased when using a small cell. Furthermore, this model cannot validate the changeability in headway distance between the lead and follower vehicles regarding vehicle velocity due to the same cell size assumption. Therefore, a wide range of modifications and extensions to the CA model was performed by Incorporating many parameters such as other vehicle types, vehicle size, mechanical properties, lateral arrangement, lateral gaps between vehicles, flow, velocity, occupancy of a vehicle in a cell, cell size, and acceleration&vehicle type [353-362]. As a result, the improved CA model utilizes a cell size of 0.5 m in length, a safe gap at the front and back of vehicles, the relationship area of occupancy, interaction rate, and structure of vehicles [361, 362]. However, according to outcomes presented in [353-362], further validation and investigation of the model in various traffic conditions and lateral and longitudinal interactions is required to expand its application.

1.7 Fuzzy Logic Model

In 1992, Kikuchi et al. incorporated the relative headway distribution, velocity, and acceleration into the fuzzy model [363]. The developments of the fuzzy logic model continued with time to include the car-following model [364-367]. The generated outcomes show several issues due to the inadequate establishment of the drivers’ perception. Moreover, the mixed traffic flow in car-following behavior in this model is not included [321]. An advantage of the fuzzy logic model in car-following maneuvers is the ability to determine the lane shifting behavior of the vehicles [368-370].

2) Other micro-simulation studies

Lane-changing models are essential elements for modeling HVs, CAVs, and AVs in traffic micro-simulation tools. In 1978, Sparmanns proposed a lane-change model to classify lane-change behavior as slower-to-faster and faster-to-slower lanes based on driver need [321]. In 1986, Gipps presented the well-known lane-change model for urban driving that considers the effects of elements such as traffic signals, obstructions, and heavy vehicles in traffic flow [371]. The main focus of the Gipps model is to investigate the critical interaction between vehicle to vehicle, vehicle to obstructions, and other real-world driving behaviors [371]. More studies about lane-change modeling in micro-simulation can be found in [372-373].

The intersection is one of the challenging environments for AVs because of the unpredicted interactions among pedestrians, bicycles, and vehicles and the intersection users’ highly complicated design and behavior. As a result, many studies cover a wide range of research related to intersection scenarios. For example, in [123], a turning vehicle is modeled, and its surrogate safety indicators are investigated at mixed-flow intersections. Authors in [124] validated the automated...
intersection traffic management applications using a vehicle-in-the-loop (VIL) verification environment. In [127], Mohd et al. evaluated CAVs and AVs applications and fuel consumption and emission using a hardware-in-the-loop testbed. The outcomes showed fast data transfer at every 200 ms, and the optimized engine operating points and the desired vehicle speed are tracked precisely. In [126], Shao et al. assessed CAVs and AVs using a hardware-in-the-loop testbed and a living lab focusing on fuel consumption and emissions. The outcomes revealed that the error between the virtual vehicle and the actual testing vehicle is 1%. Thus, the results support the use of the HIL testbed to evaluate CAVs in real-world scenarios. Furthermore, in [128], Li et al. presented an advanced intersection control system to support CAVs and AVs trajectories and validate their safety and performance at intersections. Table 16 shows a list of studies related to the modeling and validation of AVs in different scenarios.

With the advancements toward fully AVs, human drivers will not control and understand the surrounding environment. Therefore, AVs should have a social understanding of the interaction between their control systems and road users to ensure a safe driving environment [137]. The meaning of interaction in driving involves many tasks such as identifications, behavior analysis, future action prediction, and so on, and taking the right actions to avoid any severe collisions. Behavioral psychology studies have investigated the social aspects of driving and have shown what factors can significantly impact road users’ decisions [138]–[140]. These factors are pedestrian demographics [140], road conditions [139], social factors [139], and traffic characteristics [142]. Thus, a deep understanding of pedestrian crossing behavior, the extent of these factors, and how they are connected is required.

In the case of autonomous driving, intent prediction algorithms have been established to estimate the next moves of pedestrians [143] and drivers [144]. A wide range of technologies has been developed to assist AVs in communicating with road users, such as V2V [145] and V2P [146] communications. Moreover, visual intent interfaces such as LED lights [147] or projectors [148] are used. The problem with all of these studies is that they consider the technologies a rigid active thing rather than a social interaction [149]. Pedestrian behavior studies are classified into two categories, classical studies and AV conflicts studies. The traditional methods focus on studying pedestrian behavior when interacting with human drivers. A wide range of data-collection methods is used in classical pedestrian behavior studies, such as observation, police reports, video recording, photography, simulation, scripted observation, questionnaires, literature surveys, and interviews. The focus of this section is on the simulation-based method. A study conducted by Caird and Hancock [151], which involves 48 men and women, shows that the road users misjudge the vehicle arrival time as the size of the vehicle increases. In [152], Sun et al. study the relationship between pedestrian waiting time before crossing and gap acceptance. The outcomes show that a long pedestrian wait time results in a low acceptance gap. Another study investigates the impact of vehicle size on pedestrian behavior and shows that pedestrians are more careful when interacting with a larger vehicle [153]. Wiedemann [154] shows that pedestrian flow and pedestrian speed have a linear relationship with no interaction between pedestrians. Rasouli and Tsotsos [150] classify the factors that impact pedestrian behavior into two groups: pedestrian and environmental factors. Figure 16 shows a list of these factors and how they are connected under classical studies.

In contrast, figure 17 presents a list of factors that impact pedestrian behavior when facing AVs and how they are connected. Various methods are used to collect data that are used in pedestrian behavior studies involving AVs, namely, observation, video recording, photography, simulation, scripted observation, questionnaires, literature surveys, interviews, Wizard of Oz (A research experiment in which subjects interact with a computer framework that subjects believe to be autonomous but actually run or partially run by a hidden human being.).

The simulation data collection methods used in pedestrian intention studies involving AVs are discussed briefly in this section. Beggiato et al. [155] investigate the indirect forms of communication between vehicle and pedestrian, such as vehicle speed and distance. The authors claim that many factors impact the interpretation of the signal, such as vehicle speed, road users’ age, and time of day.

Jayaraman et al. [156] investigated how the availability of traffic signals at crosswalks slightly influences pedestrian crossing behavior while the AV’s driving decisions significantly impact such behavior. In [157], Chang et al. suggest a method for intent display by placing moving eyes at the
TABLE 16: Additional Simulation-based AVs studies [101]

| Ref # | Objectives | Base Model(s) | Scenarios | Vehicle Types | Evaluation Criteria | Main Results |
|-------|-------------|---------------|-----------|---------------|---------------------|--------------|
| [76]  | Suggest a pedestrian collision avoidance system based on V2X and another one based on V2P information. | Pedestrian collision avoidance model. | A pedestrian is crossing in front of two vehicles. The host vehicle has three different speeds and is in a different lane from the vehicle in the front. | CAVs & AVs | Pedestrian safety, V2X and V2P technologies | Vehicle emergency braking method is tested. A vehicle that is using V2V and V2P technologies will engage in a severe conflict with a pedestrian when the delay or packet loss rate increases to a certain value. The communication delay and packet loss rate should be reduced to avoid deadly conflicts with pedestrians. |
| [104] | Examine the safety impacts of leader-follower AVs | Bicycle mode. It comes with a linear tire model. | Connected autonomous semi-trucks under different weather conditions. | ACC & platooning | Fuel consumption, safety margin | Connected-autonomous semi-trucks that are operating in a leader-follower configuration under different weather conditions. |
| [121] | Inspect roundabout safety level in circumstances where different numbers of AVs are mixed with manual vehicles. | Surrogate safety assessment model (SSAM), Wiedemann 74 & VISSIM. | Market-penetration rate of AVs (0%, 10%, 25%, and 50%). | Manual vehicle, heavy commercial & HGV & AVs | Average travel speed | An increase of travel speed and decrease of average stop delay with an increase of the percentage of AVs. Increases in estimated crash numbers at roundabouts in terms of rear-end conflict when the AVs percentage is increased. |
| [122] | Create a decision-making CAV control algorithm in VISSIM for safety assessments. | Safety impact: SSAM, CAV. External driver model API written in C++. Manual vehicles: Wiedemann 99. | Market-penetration rate of CAVs (0%, 25%, 50%, 75%, and 100%). Daily based estimation, Monday to Friday. | Manual vehicle and CAV | Conflicts based on the threshold values of TTC (1.5 s) and PET (5 s). | CAVs bring about compelling benefit to road safety as traffic conflicts significantly reduce even at relatively low market-penetration rates (12–47%, 50–80%, 82–92% and 90–94%) for 25%, 50%, 75%, and 100% CAV penetration rates, respectively. |
| [123] | Model the turning vehicles at mixed-flow intersections and investigate their surrogate safety measures. | Car-following, yielding & SSM models. | Turning vehicles at mixed-flow intersections. | Regular vehicle: autonomous vehicles in the future. | Evaluate turning vehicles at intersections | Great capability of reproducing four features of real turning trajectories. The travel time of vehicles using simulation does not show significant difference from that in the field. The PET for vehicles and VRU does not show significant difference from the real data. |
| [128] | Propose an advanced intersection control system to coordinate vehicle trajectories and verify safety. | Temporal-spatial trajectory model. Two signal-control algorithms. | The effects of rain and pedestrian interaction on AVs. | AV | Evaluate safety and performance efficiency at intersections | The total traffic delays at intersections are reduced. The intersection capacity and performance efficiency are increased. |
| [129] | Model AVs and validate their safety and performance in a highly uncertain environment. | American Center for Mobility (ACM): No vehicle cut-in (Radar only & Radar and DSRC). Vehicle cut-in (Radar only & Radar and DSRC). | Pedestrian classification time to collision (PCT) | AV | Throughput | With an increase in rainfall, the detection and classification of lanes and obstacles in the environment takes a longer time to complete. The time-to-collision decreases with increasing rainfall. |
| [130] | Study different unsafe scenarios for leader-follower to understand the critical driving conditions of AV platoons. | ACC, DSRC. | AV | Evaluate unsafe scenarios | Having DSRC communications between two trucks in a convoy is an efficient way of decreasing the amount of “blind time,” where the following vehicle is operating without awareness of where the lead vehicle is. |
| [131] | Model the SSM of pedestrian-AV conflicts. | Pedestrian protection system (PPS) model. Bicycle model. | Pedestrian, child walking and running, and bicycle crossing in front of the AVs. | AV | AV-pedestrian evaluation & SSM | Pedestrian Classification Time to Collision (PCT), Total Braking Time to Collision, and Total Minimum Time to Collision can be used as safety margins for AV-pedestrian conflicts. |
| [132] | Investigate the SSM of AV-VRU conflict-weather conditions and sensor failure. | Pedestrian protection system (PPS) model. Bicycle model. | Two test scenarios: Pedestrian walking and running in front of the AVs. | AV-VRU conflicts evaluation & SSM | Pedestrian classification time to collision (PCT) and post encroachment (PET) can be utilized as safety indicators for AV-VRU conflicts. |
| [160] | Present a new technique for developing a pedestrian CIB simulation model. | Pedestrian protection system (PPS) | AV (speed varies in a straight road with a pedestrian crossing (diff weather&light)). | AVs | Throughput | The stop at distance is very sensitive to the error of warning starting TTC, braking starting TTC, and average deceleration. |
| [161] | Present a V2P-PAEB system that permits the vehicles to share the data of pedestrians observed by the PAEB system in the V2V network. | Pedestrian automatic emergency braking system (PAEB), bicycle model. | Intersection with 5 vehicles and a pedestrian. Vehicle 5 didn’t view the pedestrian while crossing because the view is blocked by vehicle 2. | Pedestrian safety at intersection evaluation using PAEB and V2V | Two cases are studied: Vehicle 5 uses PAEB only, and vehicle 5 uses PAEB and V2V from vehicle 1 and 2. When vehicle 5 uses a V2P-PAEB system, the vehicle can detect pedestrians much earlier than in case 1. A V2P-PAEB system can react quickly to avoid a collision. |
Autonomous vehicle developments. V-model is one of the primary methodologies for the design and validation of autonomous or non-autonomous systems. Simulation and modeling platforms are well-developed tools for testing and validating these systems. One approach to validating autonomous vehicle systems is the vehicle-hardware-in-the-loop (VHIL) approach. The vehicle test simulation approach is based on the V-model criteria. Several factors must be considered when choosing a simulator for autonomous vehicles, such as the availability and compatibility of models, subsystems that can be tested, and the availability of real-time simulation communication protocols, and compliance with ISO 26262 [163].

Pedestrian factors are critical when designing intent displays for autonomous driving. A deep understanding and consideration of pedestrian factors are essential for developing safe autonomous systems. These factors include the group size, pedestrian speed, and street width, among others. Moreover, some elements from classical studies, such as road structure, gap acceptance, and social norms factors, are considerations when selecting a simulator for autonomous vehicles. For instance, a study by Pillai [158] suggests that pedestrians' crossing decisions depend on the erratic behavior of the vehicle (speed and distance) and that under specific weather conditions with low visibility, the use of intent display will be helpful. Finally, Pillai concluded that culture is an essential factor that should be considered when designing any intent displays. According to literature reviews [150], pedestrian behavior under autonomous driving conditions needs more focus to include signal, location, road structure, gap acceptance, and social norms factors. Moreover, some elements from classical studies, such as group size, pedestrian speed, and street width, should be evaluated under autonomous driving circumstances. These factors are essential for understanding pedestrian intention to cross the road. A deep understanding and consideration of these factors will result in safe autonomous driving. In short, the V2V, V2P, and V2I communications can provide a safe environment for autonomous driving and road users. However, although using these technologies is advantageous, several questions are raised regarding sharing pedestrians' data via these technologies [159].

3) Autonomous Vehicle Simulation Platforms

Simulation and modeling platforms are well-developed tools for the design and validation of autonomous or non-autonomous vehicle developments. V-model is one of the most popular simulation methods used to cover the testing and design of the entire AV development process [164]. In the development process of autonomous vehicles, virtual simulation methods are applied at different stages, and various testing setups are achieved, such as model-in-the-loop (MIL), software-in-the-loop (SIL), and hardware-in-the-loop (HIL). The ISO 26262 is based on V-model and does not match an agile development process. As a result, there is a wide range of simulation platforms. In [163], Rosique et al. classify the simulation platforms into four different approaches that can be considered when selecting a simulator for autonomous vehicles, namely, vehicle test simulation, games, and physics engines for simulation, robotics simulators, and perception simulators.

Autonomous vehicle development is based on V-model development. Model V has several phases of development and testing, such as the model-in-the-loop (MIL) [165] approach, software-in-the-loop (SIL) [166] validation, hardware-in-the-loop (HIL) [167], and the vehicle-hardware-in-the-loop (Ve-HIL) approach. The vehicle test simulation approach is based on the V-model criteria. Several factors must be considered when choosing an autonomous vehicle simulator, such as the availability and compatibility of models, subsystems that can be tested, availability of real-time simulation communication protocols, and compliance with ISO 26262 [163].

Table 17 shows some of the simulation platforms used to validate and test autonomous vehicles based on the vehicle test simulation approach.

Another ordinary simulator is the use of an available game engine. A game engine is defined as the software part of a computer game that has a rendering engine, a physics engine, collision detection, and response, sound, scripting, animation, artificial intelligence, networking, streaming, memory control, threading, localization support, and scene graph. In addition, it might incorporate video support for cinematic and virtual reality (VR) simulation [163]. Game engines provide some features that are advantageous for autonomous vehicles and robotics researchers, such as physical fidelity, distributed architecture, cutting-edge graphics, and scriptable environments [163]. The main game engines that are used widely in the development of autonomous vehicle systems or subsystems are Unity 3D [174], Unreal Engine [175], Blender [176], and Cry Engine [177]. The physics engine is an essential component when simulating an autonomous vehicle perception system. This engine provides less fidelity and works according to the detection of collisions. Examples of the high-performance physics engines that are used in AV simulation are Open Dynamics Engine (ODE) [178], Bullet physics [179], NVidia PhysX [180], and PreScan [231].

Robotics simulation platforms are also used in autonomous vehicle simulation. Models of all sensors and actuators should be provided for the effective use of this type of simulation [181-183]. Moreover, a realistic environment for testing and validating all kinds of algorithms and subsystems should be provided as well. Many features should be considered when choosing a robotics simulator, such as 3D...
rendering, license, external agent support, sensor noise, parallelism/distribution, level of maturity, fault-tolerance, realistic scenario simulation, and HIL simulation techniques [163]. Examples of current robotics simulators that incorporate data simulation sensors are Gazebo [184], V-REP [185], Webots [186], and Microsoft Robotics Developer Studio (MRDS). In the robotics domain, USARSim [187], BlenSor [188], and MORSE [189] are three leading simulators that are used extensively for research. For example, MORSE was used by Ford Motor Company to test the 2021 Ford Mustang Mach-E [190]. Table 18 presents a broader list of some robotics simulators that integrate data simulation data. Table 19 provides a comparison between sensors that are simulated using robotics simulators.

Simulation platforms should mimic real-world environments to model and validate any perception algorithm. Therefore, the available simulation platforms tend to have these features: fast prototyping, physics engines for realistic motions, realistic 3d rendering, and dynamics with scripting. Tables 20 and 21 show a list of perception algorithms simulation platforms and their features.

### B. AGENT-BASED MODELS

Agent-based models integrate activity-based demand generation and dynamic traffic assignments [204]. This approach covers all macroscopic four-step procedures, namely, demand generation, demand distribution, model choice, and traffic assignment [203]. Agent-based models (ABMs) also utilize independent agents with a bottom-up technique to simulate a highly complex system [202]. This type of modeling and simulation is considered a superior simulation approach compared to other methods in terms of flexibility, hierarchy, intuition, and dealing with complex systems. For example, an AV operating on public roads while interacting with human-driven cars, vulnerable road users, and road networks is highly problematic. Within the independent vehicle system, all subsystems are interconnected and worked simultaneously. Moreover, a wide range of elements, such as the diverse behavior of agents (people and vehicles), is integrated within agent-based modeling. With high-end computers, the agent-based modeling approach is used to build challenging models with more realistic scenarios.

The autonomous-vehicles agent-based modeling studies are diverse. They include the travel and environmental impacts of autonomous vehicles [70],[206], the parking requirements with the arrival of autonomous vehicles [207],[208], the traffic congestion caused by autonomous cars [209], the system performance of the self-driving vehicle [210],[211],[208],[212],[213], and the autonomous vehicles’ modal share and travel modes [214],[72],[215]. There are many critical variables that can impact the system performance of the AV, namely, fleet size, demand, strategy, ride-sharing, pricing schemes, configurations of stations, travel mode, vehicle capacity, service area, refuel/recharge time, maximum waiting time, and cruising time [221]. Many researchers consider these variables in the sensitivity analysis and modeling of various simulation scenarios. For example, in the literature review, there are 27 papers related to the fleet-size research area. In the fleet-size studies, the regular vehicles are replaced by autonomous vehicles (AV) [71], autonomous taxis (ataxi) [216], autonomous mobility on demand (AMOD) [217], autonomous transit on demand (ATOD) [218], shared autonomous vehicles (SAV) [70], or shared autonomous electric vehicles (SAEV) [219].

The fleet size or replacement rate is considered one of the significant outcomes of the agent-based simulation. The replacement rate is used as an indicator to show the efficiency of an autonomous vehicle system. Fagnant et al. [220] argue that the travel demand, average speeds, and average trip distances impact AV system performance. Moreover, the replacement rate of the autonomous vehicle is investigated with a case study in Austin. Fagnant et al. [70] argue that one autonomous vehicle can replace ten human-driven cars. In [221], the outcomes show that the replacement rate in [70] is 1:11 with link-level travel time and is 1:9 with constant speed in [220]. In [210], Marczuk et al. show that fleet size relies on many factors, such as service area, average demand, level of service (based on average waiting time, service and reject rate), routing scheme, relocation plan, and design of the facility. Moreover, in [206], the fleet size can be minimized by ride-sharing. Additionally, the average trip distances data in the survey papers are not precise enough. Furthermore, environmental scenarios, such as an urban area or highway, can impact the travel distance. In [221], ride-sharing is considered as an indicator for the routing scheme. In short, many major factors can affect the fleet size or replacement rate, namely, service data, average demand, average speed, average waiting time, service and reject rate, ride-sharing, relocation plan, and design of the facility. The replacement rate in [70],[220],[216] is the same, considering that one
TABLE 18: Summary of the main features of robotic simulator platforms for AVs [163]

| Simulator | License                      | Simulation engine | Graphical engine | External Agent |
|-----------|------------------------------|-------------------|------------------|----------------|
| Gazebo    | GPL/Open Source              | ODE, Bullet, Simbody Art | Ogre3D          | Yes            |
| V-Rep     | GPL/Open Source, Commercial  | ODE, Bullet, Vortex | OpenGL          | Yes            |
| Webots    | Commercial                   | ODE               | -                | Yes            |
| MRDS      | Commercial                   | PhysX             | DirectX          | No             |
| USARSim   | GPL                          | Unreal Engine     | -                | Yes            |
| BlenSor   | GPL/Open Source              | Blender, Bullet   | OpenGL          | No             |
| MORSE     | GPL/Open Source              | -                 | OpenGL          | Yes            |

TABLE 19: Summary of the main sensors simulated by robotic simulator platforms for AVs [163]

| Simulator | GPS | IMU | LIDAR | Ultrasonic | Radar | Infrared | Stereo Camera | ToF Camera |
|-----------|-----|-----|-------|------------|-------|----------|---------------|-----------|
| Gazebo    | x   | x   | x     | x          | x     | x        | x             | x         |
| V-Rep     | x   | x   | x     | x          | x     | x        | n/a           | n/a       |
| Webots    | x   | n/a | x     | -          | x     | x        | -             | -         |
| MRDS      | x   | n/a | x     | n/a        | x     | n/a      | n/a           | n/a       |
| USARSim   | x   | x   | x     | n/a        | x     | n/a      | n/a           | n/a       |
| BlenSor   | x   | x   | x     | n/a        | n/a   | x        | x             | x         |
| MORSE     | x   | x   | x     | -          | x     | x        | n/a           | n/a       |

*Table legend: x—Yes, n/a—Unknown or could not be determined, - —No.

TABLE 20: Summary of the features of specific perception simulation platforms for AVs [163].

| Simulator         | License            | Simulation engine | Graphical engine | Scripting Language | External Agent | Notes               |
|-------------------|--------------------|-------------------|------------------|--------------------|----------------|---------------------|
| CARLA [191]       | GPL/Open Source    | Unreal Engine     | GPU              | Python             | Yes            | Driving             |
| AirSim [200]      | GPL/Open Source    | Unreal Engine     | n/a              | C++, Python, C, Java | Yes           | Driving/HIL, SIL    |
| DeepDrive [192]   | GPL/Open Source    | Unreal Engine     | n/a              | C++, Python        | Yes            | Driving             |
| Udacity * [193]   | GPL/Open Source    | Unity             | n/a              | C++, Python        | n/a            | Driving             |
| Constellation [194]| Restricted         | PhysX/CUDA        | GPU              | C/C++, Python      | Yes            | Cloud, HIL, VR      |
| Carcraf/Waymo [195]| Restricted         | n/a               | n/a              | n/a                | Yes            | Driving             |
| SIMLidar [196]    | GPL/Open Source    | n/a               | n/a              | C++                | n/a            | LiDAR               |
| Helios [197]      | GPL/Open Source    | JMonkey Engine    | OpenGL           | Java               | n/a            | LiDAR               |
| GLIDAR [198]      | GPL/Open Source    | OpenGL            | C++              | n/a                | LiDAR          |                    |
| RADSim [199]      | Commercial         | n/a               | n/a              | MATLAB             | n/a            | RADAR               |
| SIMSonic [201]    | GPL/Open Source    | n/a               | n/a              | R                  | n/a            | Ultrasonic          |
| PreScan [231]     | Commercial         | Unreal Engine     | GPU              | Matlab, C++        | n/a            | Driving/HIL, SIL    |

*Table legend: n/a—Unknown or could not be determined.

autonomous taxi can replace ten regular vehicles, excluding relocation and travel demand. The outcomes show that the average waiting time is approximately 2.28 minutes, which is considered too large.

In [162], the authors present an autonomous intersection management algorithm called AIM-ped, which considers vehicles and pedestrians. The total optimal throughput is created when incorporated with max pressure control. Moreover, the conflict region model conducts stability analysis of the autonomous intersection management system’s properties. The AIM-ped algorithm is implemented by integrating the max-pressure control with an existing trajectory optimization algorithm to obtain the optimal vehicle trajectories. The result is that the AIM-ped algorithm can trigger vehicle movements when there is a change in pedestrian demand. The simulation outcomes show that pedestrians and vehicle delays are negatively correlated. In [125], the sequential movements of vehicles in intersections are modeled as a multi-agent Markov decision process (MAMDPS). The outcomes show that the optimal sequential decision from DCL-AIM outperforms all the other control policies. In [205], the authors present a model to develop the interaction dynamics
between drivers and pedestrians in dense traffic areas where pedestrians and/or drivers do not obey the traffic laws and regulations. This approach can be used in control systems of AVs and drivers’ onboard alert systems [205]. In [92], the performance of many SAV fleets and vehicle sizes serving travelers across France’s Rouen Normandie metropolitan area is evaluated. Moreover, the effect of ride-sharing and rebalancing strategies on service is studied. This study emphasizes that the performance of SAV is strongly correlated with the fleet size and the shared rides.

Table 22 presents a summary of the features of specific agent-based simulation platforms. Most agent-based simulation research papers use the MATSim simulation platform to model all the autonomous vehicle system operations. In the past, MATSim was used to simulate regular vehicles and not autonomous vehicles. With the need to validate AVs, [209] and [222] establish the AV toolkit for MATSim. In the agent-based simulation, the autonomous vehicles are likewise simulated, and the difference between all the simulated vehicles is the data source. For example, for an autonomous taxi, taxi data are used, and travel surveys are taken into account for other car-sharing services. Private AVs simulation in the literature surveys is limited. In addition to agent-based modeling, the augmented and virtual reality methods have a great potential to be essential methods for AV evaluation. Recent studies about these methods are presented in [374-377]. In short, the agent-based simulation approach for AVs is in its infancy stage. More focus on agent-based simulation of private autonomous vehicles is required for this approach to compete with other AV validation methods.

VIII. DISCUSSION AND FUTURE RESEARCH DIRECTIONS

Two main tests are currently used to evaluate CAVs and AVs, namely Naturalistic field operational tests (N-FOTs) and virtual tests. The virtual tests include test matrix evaluation, worst-case scenario evaluation (WCSE), Monte Carlo simulations, accelerated evaluation (AE), simulation-based and agent-based modeling approaches. In some cases, both N-FOTs and virtual tests are combined simultaneously to evaluate CAVs and AVs. In N-FOTs, vehicles are equipped with the required sensors and are driven in naturalistic conditions, which is not the case in virtual tests. The N-FOTs allow the investigators to observe CAVs and AVs in a natural setting. The data collected from N-FOTs are utilized to investigate many elements, such as driver performance, surrounding environment, driving conditions, and other components related to critical incidents, near collisions, and collisions. However, the N-FOTs have many restrictions, such as the time required to conduct the test, trained drivers, the low probability of critical events. Moreover, the test requires many vehicles, a lot of time, and large budgets. Therefore, virtual tests are used as an efficient alternative approach to model and validate CAVs and AVs. Many questions are raised regarding virtual tests and how these tests can be reliable and replace the Naturalistic field operational tests. For example, urban scenes are essential in virtual tests, which inevitably involve pedestrians, vehicles, cyclists, motorcyclists, etc. Simulating traffic congestion, lane-change scenarios, car-following scenarios, pedestrian-vehicle conflict, vehicle-vehicle conflict, pedestrian behavior, driver behavior, human-driven vehicle behavior, weather conditions, and so many scenarios in a large-scale traffic scenario is a complicated multi-layer task. Usually, the resulting movements of each object in the simulation rarely follow the physical laws. Moreover, accessing vehicle trajectories and including them in virtual tests or applications in real-time is challenging. Furthermore, road networks generation and representation is also a fundamental task in traffic simulation and modeling. Several simulation tools support road networks, but the outcomes do not resemble real-world traffic at the street level. Therefore, in virtual testing methods, model verification in terms of the similarity between the simulated traffic model and real-world scenarios is always a concern. In this review paper, we presented the pros and cons of each evaluation method. The review papers presented in this survey show a clear gap in the research area of CAVs and AVs evaluation. Every method has its strengths and weaknesses. For example, many techniques focus on car-following and lane-change modeling and evaluation and neglect the remaining traffic conditions and the modeling of severe weather conditions. Furthermore, the V2V, V2P, and V2I technologies are still under investigation and require
more validation.

Moreover, the CAVs and AVs modeling and evaluation is a task that requires integrating many models simultaneously with a wide range of parameters and variables. Choosing the suitable models will produce satisfying outcomes. Based on our findings, different models related to behavior, car-following, lane-change, vehicle dynamics, etc. being used in every research paper. Therefore, establishing a comparison study is a very challenging task. For example, some car-following models outperform other models. Using the superior car-following model in a research paper with a specific evaluation method will produce promising outcomes for this method. Moreover, several architecture models have been developed from completely modular to fully end-to-end, each with its limitations. The optimal algorithms for localization, mapping, and perception still lack accuracy and efficiency. In short, for safe autonomous driving, a high-fidelity driving simulator, which includes realistic traffic streams and complicated traffic conditions, is necessary. Such a simulator can construct critical training environments in an efficient and reproducible manner. New evaluation methods need to be developed for more scenarios to give a thorough validation of AVs. The community has not fully understood the full failure modes of AVs to design a complete list of test scenarios, but possible elements incorporate:

1) Challenges in sensing/detection under severe weather conditions such as heavy snow, rain, fog, etc.
2) Aggression of surrounding vehicles/vulnerable road users such as running a red light, cut-in, jaywalk, etc.
3) Challenges in making decisions such as under low confidence, multiple threats at a time, and so on.
4) Challenges due to road types and vehicle types

Moreover, these simulators and evaluation methods should provide clear answers to the following questions:

1) What are scalable driving policies to control many AVs in mixed traffic comprised of human-driven vehicles (HVs), CAVs, AVs, vulnerable road users, etc.?
2) How do we estimate human driver behaviors, pedestrian behaviors, surrounded vehicles (HVs, CAVs, and AVs)?
3) How to make sure that the behaviors of drivers and pedestrians are accurate and capture the real-world behaviors?
4) How should the driving behavior of HVs, CAVs, AVs be modeled in the environment?
5) How are the interactions between human-driven vehicles (HVs) and AVs characterized?
6) How are the interactions between CAVs and AVs characterized?
7) How are the interactions between pedestrians and AVs characterized?
8) How are the interactions between other vulnerable road users (VRU) and AVs characterized?
9) How should pedestrian behavior be modeled in the environment?
10) How should the severe weather conditions be modeled in the environment?

Many methods showed promising outcomes but did not provide answers to all of these questions.

### IX. CONCLUSION

It is critical to evaluate AVs thoroughly before their release and deployment to the general public. However, because most trips are not safety-critical in naturalistic driving, testing AVs on public roads is hugely time-consuming, inefficient, and expensive. In this paper, we surveyed all the evaluation methods. These methods are naturalistic field operational tests, test matrix evaluation, worst-case scenario evaluation, Monte Carlo simulations, accelerated evaluation, and simulation-based model approach. This survey has shown that there is a clear gap in the field of AVs evaluation. Many factors affect our judgment on what is the best approach to evaluate the AVs. These factors include:

1) The AVs to AVs and HVs to HVs interactions are not studied and used only as a benchmark.
2) The AV sensors and controls are suggested to work accurately in many papers, and the measurements are presumed to be accurate.
3) The drivers’ reactions to AVs are assumed to be the same as to HVs.
The accelerated evaluation approach outperforms naturalistic field operational tests (N-FOTs), test matrix evaluation, worst-case scenario evaluation (WCSE), Monte Carlo simulations methods in some of the car-following, and lane-change studies when using specific models in terms of the assessment time of the collision, injury, or conflict event. In addition, some studies show that integrating machine and deep learning techniques to test matrix evaluation, Monte Carlo simulations, and accelerated evaluation can reveal significant improvements. In the simulation-based model approach, the agent-based modeling approach was investigated and shown to be advantageous in AV modeling and validation. However, more works are needed to implement the agent-based modeling approach to cover a wide range of self-driving vehicle research. Another promising approach for AV evaluation is the augmented and virtual reality methods. The development of AVs depends on advancements in scientific disciplines and new technologies. Therefore, the AV research development has a high impact on AV driving technology by overcoming the weaknesses of the available evaluation methods and by inventing new evaluation methods.

REFERENCES

[1] A. Alessandri, A. Campagna, P. D. Site, F. Filippi, and L. Persia, “Automated Vehicles and the Rethinking of Mobility and Cities,” Transportation Research Procedia, vol. 5, pp. 145–160, 2015, doi: 10.1016/j.trpro.2015.01.002.

[2] “Automonomous Vehicles | Self-Driving Vehicles Enacted Legislation,” https://www.ncsl.org/research/transportation/autonomous-vehicles-self-driving-vehicles-enacted-legislation.aspx (accessed Jun. 6, 2020).

[3] J. M. Anderson, N. Kalra, K. D. Stanley, P. Sorenson, C. Samaras, and T. A. Oluwotola, “Autonomous Vehicle Technology: A Guide for Policymakers,” Mar. 2016. Accessed: Oct. 26, 2020. [Online]. Available: https://www.rand.org/pubs/research_reports/RR443-2.html.

[4] S. Beiker and R. Calo, “Legal Aspects of Autonomous Driving,” SSRN Journal, 2010, doi: 10.2139/ssrn.2767899.

[5] “UK government paves way for driverless cars,” BBC News, Dec. 05, 2013. https://www.bbc.com/news/technology-25230483. (accessed Aug. 22, 2019).

[6] L. K. CNN, “Driverless cars will be on UK roads by 2021, says government,” CNN. https://www.cnn.com/2019/02/06/uk/driverless-cars-selfdriving-intl/index.html (accessed Dec. 6, 2019).

[7] “2019 Autonomous Vehicles Readiness Index - KPMG Belgium,” KPMG, Feb. 26, 2020. https://home.kpmg/be/en/home/insights/2019/02/2019-autonomous-vehicles-readiness-index.html (accessed Nov. 5, 2019).

[8] “Bloomberg Aspen Initiative on Cities and Autonomous Vehicles,” Global Atlas of AVs in Cities. https://avinicsities.bloomberg.org/ (accessed Jan. 28, 2020).

[9] C. B. Oct. 30 and 2017, “Automontous vehicles hitting Greenville County, S.C. roads,” NACO. https://www.naco.org/articles/autonomous-vehicles-hitting-greenville-county-sc-roads (accessed Jan. 29, 2020).

[10] “WeRide kicks off a robotaxi pilot program in Guangzhou - TechNode,” TechNode, Nov. 28, 2019. https://technode.com/2019/11/28/wereide-robotaxi-launch-guangzhou/ (accessed Jan. 29, 2020).

[11] “Walmart partners with self-driving startup Nuro to test autonomous grocery delivery in Houston,” TechCrunch. https://social.techcrunch.com/2019/12/10/walmart-partners-with-self-driving-startup-nuro-to-test-autonomous-grocery-delivery-in-houston/ (accessed Jan. 29, 2020).

[12] “Houston’s First Driverless Shuttle Now Rolling at TSU,” https://www.gototech.com/tsu/automation/houstons-first-driverless-shuttle-now-rolling-at-tsu.html (accessed Mar. 12, 2020).

[13] “Waymo’s self-driving Jaguar I-Pace vehicles are now testing on public roads,” TechCrunch. https://social.techcrunch.com/2019/06/17/waymos-self-driving-jaguar-i-pace-vehicles-are-now-testing-on-public-roads/ (accessed Jan. 15, 2020).

[14] “Autonomous Driving Connected Mobility,” Frost Sullivan. https://ws2.frost.com/research/industry/mobility-automotive-transportation/autonomous-driving-connected-mobility/ (accessed Oct. 27, 2020).

[15] “Strategic Analysis of the European and North American Market for Automated Driving,” English. https://store.frost.com/strategic-analysis-of-the-european-and-north-american-market-for-automated-driving.html (accessed Dec. 3, 2019).

[16] “Your Autopilot has arrived,” Oct. 14, 2015. https://www.tesla.com/blog/your-autopilot-has-arrived (accessed Jun. 10, 2020).

[17] “Tesla Version 7.0 Software Update Activates First Autopilot Features,” Motor Authority. https://www.motorauthority.com/news/1100476_tesla-version-7-0-software-update-activates-first-autopilot-features (accessed Jun. 10, 2020).

[18] “The autonomous car takes to the road,” Valeo. Jul. 28, 2015. https://www.valeo.com/en/the-autonomous-takes-to-the-road/ (accessed Mar. 12, 2020).

[19] “Autonomous Vehicles Car Companies 1 CB Insights,” CB Insights Research. Mar. 04, 2020. https://www.cbinsights.com/research/autonomous-driverless-vehicles-companies-list/ (accessed Mar. 8, 2020).

[20] “Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles (J3016 Ground Vehicle Standard) - SAE Mobiltus.” https://saemobiltus.sae.org/content/j3016_201806 (accessed Nov. 22, 2020).

[21] M. Buehler, K. Iagnemma, and S. Singh, Eds., The DARPA Urban Challenge: Autonomous Vehicles in City Traffic. Berlin Heidelberg: Springer-Verlag, 2009.

[22] “CVG @ ETHZ - V-Charge: Automated Valet Parking and Charging for e-Mobility,” https://www.cvg.ethz.ch/research/v-charge/ (accessed Jun. 11, 2020).

[23] P. Furgale et al., “Toward automated driving in cities using close-to-market sensors: An overview of the V-Charge Project,” in 2013 IEEE Intelligent Vehicles Symposium (IV), Jun. 2013, pp. 809–816, doi: 10.1109/IVS.2013.6628566.

[24] S. D. Pendleton et al., “Perception, Planning, Control, and Coordination for Autonomous Vehicles,” Machines, vol. 5, no. 1, Art. no. 1, Mar. 2017, doi: 10.3390/machines5010006.

[25] M. K. Mills, “Army’s ST Investment in Ground Vehicle Robotics,” B. S. Jahromi, “Levels of Automation for Self-Driving Cars,” Medium, Aug. 16, 2019. https://medium.com/@BabakShah/levels-of-automation-for-self-driving-cars-64104f467987 (accessed Dec. 3, 2019).

[26] “The public are ready for autonomous vehicles, but our cities aren’t.” https://www.aurecongroup.com/markets/transport/future-of-transport/australia/autonomous-transport (accessed Feb. 15, 2020).

[27] I. is Beautiful, “Million Lines of Code,” Information is Beautiful. https://informationisbeautiful.net/visualizations/million-lines-of-code/ (accessed Sep. 5, 2019).

[28] U. S. G. Office, “Vehicle Cybersecurity: DOT and Industry Have Efforts Under Way, but DOT Needs to Define Its Role in Responding to a Real-world Attack,” no. GAO-16-350, Apr. 2016, Accessed: Nov. 22, 2019. [Online]. Available: https://www.gao.gov/products/GAO-16-350. (accessed Sep. 5, 2019).

[29] “2009–11 Toyota vehicle recalls,” Wikipedia. Nov. 06, 2020. Accessed: Nov. 22, 2020. [Online]. Available: https://en.wikipedia.org/wiki?title=2009%E2%80%9311_Toyota_vehicle_recalls&oldid=987346472.

[30] S. Lassaire et al., “FESTA Support Action Field OpErational teSt supporT Action,” Jan. 2009.

[31] M. L. Aust, “Improving the Evaluation Process for Active Safety Functions,” p. 61.
D. Zhao et al., “Accelerated Evaluation of Automated Vehicles Safety,” IEEE Transactions on Intelligent Transportation Systems, vol. 18, no. 3, pp. 595–607, Mar. 2017, doi: 10.1109/TITS.2016.2582208.

D. Zhao et al., “Accelerated Evaluation of Automated Vehicles in Lane Change Scenarios,” presented at the ASME 2015 Dynamic Systems and Control Conference, Jan. 2016, doi: 10.1115/DSCC2015-9718.

D. Zhao, X. Huang, H. Peng, H. Lam, and D. J. LeBlanc, “Accelerated Evaluation of Automated Vehicles in Car-Following Maneuvers,” IEEE Transactions on Intelligent Transportation Systems, vol. 19, no. 3, pp. 733–744, Mar. 2018, doi: 10.1109/TITS.2017.2701846.

D. Zhao, H. Peng, S. Bao, K. Nobukawa, D. LeBlanc, and C. Pan, “Accelerated evaluation of automated vehicles using extracted naturalistic driving data,” 2016, pp. 287–296.

Z. Huang, H. Lam, D. J. LeBlanc, and D. Zhao, “Accelerated Evaluation of Automated Vehicles Using Piecewise Mixtures Models,” IEEE Transactions on Intelligent Transportation Systems, vol. 19, no. 9, pp. 2845–2855, Sep. 2018, doi: 10.1109/TITS.2017.2766172.

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 2017 IEEE International Conference on Robotics and Automation (ICRA), May 2017, pp. 197–202, doi: 10.1109/ICRA.2017.7989024.

Z. Huang, H. Lam, and D. Zhao, “An accelerated testing approach for automated vehicles with background traffic described by joint distribution,” in 2017 IEEE 21st International Conference on Intelligent Transportation Systems (ITSC), Oct. 2017, pp. 933–938, doi: 10.1109/ITSC.2017.8317940.

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 2018 Annual American Control Conference (ACC), Jun. 2018, pp. 4800–4802, doi: 10.23919/ACC.2018.8431590.

Z. Huang, H. Lam, and D. Zhao, “RARE-EVENT SIMULATION WITHOUT STRUCTURAL INFORMATION: A LEARNING-BASED APPROACH,” in 2018 Winter Simulation Conference (WSC), Dec. 2018, pp. 1826–1837, doi: 10.1109/WSC.2018.8632355.

Z. Huang, H. Lam, and D. Zhao, “DESIGNING IMPORTANCE SAMPLERS TO SIMULATE MACHINE LEARNING PREDICTORS VIA OPTIMIZATION,” in 2018 Winter Simulation Conference (WSC), Dec. 2018, pp. 1730–1741, doi: 10.1109/WSC.2018.8632173.

Z. Huang, M. Arief, H. Lam, and D. Zhao, “Evaluation Uncertainty in Data-Driven Self-Driving Testing,” arXiv:1904.09306 [cs, stat], Jul. 2019, Accessed: Nov. 12, 2020. [Online]. Available: http://arxiv.org/abs/1904.09306.

Z. Huang, H. Lam, and D. Zhao, “Towards affordable on-track testing for autonomous vehicle — A Kriging-based statistical approach,” in 2017 IEEE 20th International Conference on Intelligent Transportation Systems (ITSC), Oct. 2017, pp. 1–6, doi: 10.1109/ITSC.2017.8317656.

Z. Huang, M. Arief, H. Lam, and D. Zhao, “Synthesis of Different Autonomous Vehicles Test Approaches,” in 2018 21st International Conference on Intelligent Transportation Systems (ITSC), Nov. 2018, pp. 2000–2005, doi: 10.1109/ITSC.2018.8569672.

Z. Huang, D. Zhao, and M. Arief, “Anticipating the Regional Impacts of Connected and Automated Vehicle Travel in Austin, Texas,” Journal of Urban Planning and Development, vol. 144, no. 4, pp. 04180832, Dec. 2018, doi: 10.1061/(ASCE)UP.1943-5449.0000463.

Z. Huang, M. Arief, H. Lam, and D. Zhao, “Synthesis of Different Autonomous Vehicles Test Approaches,” in 2018 21st International Conference on Intelligent Transportation Systems (ITSC), Nov. 2018, pp. 2000–2005, doi: 10.1109/ITSC.2018.8569672.

Z. Huang, H. Lam, and D. Zhao, “Anticipation of the Regional Impacts of Connected and Automated Vehicle Travel in Austin, Texas,” Journal of Urban Planning and Development, vol. 144, no. 4, pp. 04180832, Dec. 2018, doi: 10.1061/(ASCE)UP.1943-5449.0000463.

Z. Huang, H. Lam, and D. Zhao, “Anticipating the Regional Impacts of Connected and Automated Vehicle Travel in Austin, Texas,” Journal of Urban Planning and Development, vol. 144, no. 4, pp. 04180832, Dec. 2018, doi: 10.1061/(ASCE)UP.1943-5449.0000463.

P. M. Boesch, F. Cari, and K. W. Axhausen, “Autonomous Vehicle Fleet Sizes Required to Serve Different Levels of Demand,” Transportation Research Part C, Jan. 2016, doi: 10.3141/2524-13.

T. D. Chen and K. M. Kockelman, “Management of a Shared Autonomous Electric Vehicle Fleet: Implications of Pricing Schemes,” Transportation Research Record, Jan. 2016, doi: 10.3141/2572-05.

T. D. Chen and K. M. Kockelman, “Management of a Shared Autonomous Electric Vehicle Fleet: Implications of Pricing Schemes,” Transportation Research Record, Jan. 2016, doi: 10.3141/2572-05.

T. D. Chen and K. M. Kockelman, “Management of a Shared Autonomous Electric Vehicle Fleet: Implications of Pricing Schemes,” Transportation Research Record, Jan. 2016, doi: 10.3141/2572-05.

T. D. Chen and K. M. Kockelman, “Management of a Shared Autonomous Electric Vehicle Fleet: Implications of Pricing Schemes,” Transportation Research Record, Jan. 2016, doi: 10.3141/2572-05.

T. D. Chen and K. M. Kockelman, “Management of a Shared Autonomous Electric Vehicle Fleet: Implications of Pricing Schemes,” Transportation Research Record, Jan. 2016, doi: 10.3141/2572-05.

T. D. Chen and K. M. Kockelman, “Management of a Shared Autonomous Electric Vehicle Fleet: Implications of Pricing Schemes,” Transportation Research Record, Jan. 2016, doi: 10.3141/2572-05.
This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/
tions with Airborne, Ground-Based Mobile and Stationary Platforms,” ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences, vol. 3, pp. 161–168, Jun. 2016, doi: 10.5194/isprs-annals-III-3-161-2016.

[198] J. O. Woods and J. A. Christian, “Glider: An OpenGL-based, Real-Time, and Open Source 3D Sensor Simulator for Testing Computer Vision Algorithms,” Journal of Imaging, vol. 2, no. 1, Art. no. 1, Mar. 2016, doi: 10.3390/jimag20100005.

[199] “Radar Simulator (RADSim) | Riverside Research.” https://www.riversideresearch.org/what-we-do/tools/radar-simulator-standalone (accessed Aug. 6, 2020).

[200] “Home - AirSim,” https://microsoft.github.io/AirSim/ (accessed Aug. 6, 2020).

[201] “SimSonic: free fdtd software for the simulation of ultrasonic waves propagation.” http://www.simsonic.fr/ (accessed Aug. 7, 2020).

[202] Y. B. Moon, “Simulation modelling for sustainability: a review of the literature,” International Journal of Sustainable Engineering, vol. 10, no. 1, pp. 2–19, Jan. 2017, doi: 10.1080/19397038.2016.1220990.

[203] J. Ortúza and L. Willumsen, Modelling Transport, 4th Edition | Wiley, 4th Edition. Wiley, 2011.

[204] M. Rieser, K. Nagel, U. Beuck, M. Balmer, and J. Rümenapp, “Fleet performance and cost evaluation of shared autonomous vehicles: A case study for Austin, Texas,” Transportation Research Part A: Policy and Practice, vol. 121, pp. 374–385, Mar. 2019, doi: 10.1016/j.tra.2019.01.025.

[205] J. Kamel, R. Vosoghi, J. Puchinger, F. Ksontini, and G. Sirin, “Exploring the Impact of User Preferences on Shared Autonomous Vehicle Modal Split: A Multi-Agent Simulation Approach,” Transportation Research Procedia, vol. 26, pp. 13–21, Jan. 2017, doi: 10.1016/j.trpro.2017.07.004.

[206] M. Maciejewski and J. Bischoff, “Congestion effects of autonomous shared taxis and shared autonomous taxis: A case study of NYC,” Transportation Research Part C: Emerging Technologies, vol. 97, pp. 45–60, Dec. 2018, doi: 10.1016/j.trc.2018.10.007.

[207] M. Lokhandwala and H. Cai, “Dynamic ride sharing using traditional taxis and shared autonomous taxis: A case study of NYCC,” Transportation Research Part C: Emerging Technologies, vol. 97, pp. 78–96, Jan. 2018, doi: 10.1016/j.trc.2017.10.024.

[208] W. Zhang, S. Guhatakarta, J. Fang, and G. Zhang, “Exploring the impact of shared autonomous vehicles on urban parking demand: An agent-based simulation approach,” Sustainable Cities and Society, vol. 19, pp. 34–45, Dec. 2015, doi: 10.1016/j.scs.2015.07.006.

[209] H. Dia and F. Javanshour, “Autonomous Shared Mobility-On-Demand: Melbourne Pilot Simulation Study,” Transportation Research Procedia, vol. 22, pp. 285–296, Jan. 2017, doi: 10.1016/j.trpro.2017.03.035.

[210] M. Maciejewski and J. Bischoff, “Congestion effects of autonomous taxi fleets,” Transportation, vol. 33, no. 4, Art. no. 4, Dec. 2018, doi: 10.3846/16484142.2017.1347827.

[211] K. A. Marczuk et al., “Autonomous mobility on demand in SimMobility: Case study of the central business district in Singapore,” in 2015 IEEE 7th International Conference on Cybernetics and Intelligent Systems (CIS) and IEEE Conference on Robotics, Automation and Mechatronics (RAM), Jul. 2015, pp. 1673–1678, doi: 10.1109/CIS-RAM.2015.7274567.

[212] Z. Liu, T. Moulin, W. Zeng, and T. Morikawa, “An Agent-Based Simulation Model for Shared Autonomous Taxi System,” Asian Transport Studies, vol. 5, no. 1, pp. 1–13, 2018, doi: 10.11175/aisasts.5.1.

[213] W. Shen and C. Lopes, “Managing Autonomous Mobility on Demand Systems for Better Passenger Experience,” arXiv:1507.02563 [cs], Apr. 2015, doi: 10.1007/978-3-319-25524-8_2.

[214] F. Javanshour, H. Dia, and G. Duncan, “Exploring the performance of autonomous mobility on-demand systems under demand uncertainty,” Transportmetrica A: Transport Science, vol. 15, no. 2, pp. 698–721, Nov. 2019, doi: 10.1080/23329915.2018.1528485.

[215] M. Heilig, T. Hilgert, N. Mallig, M. Kagerbauer, and P. Vortisch, “mobiTopp – A Modular Agent-based Travel Demand Modelling Framework,” Procedia Computer Science, vol. 130, pp. 797–802, Jan. 2020.

[216] B. Paden, M. ˇCáp, S. Z. Yong, D. Yershov, and E. Frazzoli, “A Survey of Motion Planning and Control Techniques for Self-Driving Urban Vehicles,” IEEE Transactions on Intelligent Vehicles, vol. 1, no. 1, pp. 33–55, Mar. 2016, doi: 10.1109/TIV.2016.2578706.
[240] R. Alagibri and M.-T. Chot, “Deep-Learning-Based Indoor Human Following of Mobile Robot Using Color Feature,” Sensors, vol. 20, no. 9, Art. no. 3699, Jan. 9, 2020, doi: 10.3390/s20093699.

[241] C. J. Wielick and W. G. Najim, “Driver/Vehicle Characteristics in Rear-End Precrash Scenarios Based on the General Estimates System (GES),” Mar. 1999, pp. 1999-01-0817, doi: 10.4271/1999-01-0817.

[242] W. G. Najm, J. D. Smith, and M. Yanagisawa, “Pre-Crash Scenario Typology for Crash Avoidance Research,” Art. no. DOT-VNTSC-NHTSA-06-02, Apr. 2007, Accessed: Jan. 13, 2021. [Online]. Available: https://trid.trb.org/view/1086065.

[243] W. Najim, and J. D. Smith “Crash Imminent Test Scenarios for Integrated Vehicle-Based Safety Systems,” DOT HS 810 757, 2007.

[244] S. Toma, “Depiction of Priority Light Vehicle Pre-Crash Scenarios For Safety Applications Based OnVehicle-to-Vehicle Communications,” DOT HS 811 732, 2013.

[245] W. G. Najim et al., “Description of Light-Vehicle Pre-Crash Scenarios for Safety Applications Based on Vehicle-to-Vehicle Communications,” DOT HS 811 731, 2013.

[246] “NASS General Estimates System,” NHTSA, Nov. 14, 2016, https://www.nhtsa.gov/national-automotive-sampling-system-nass-general-estimates-system (accessed Jan. 13, 2021).

[247] “Crashworthiness Data System,” NHTSA, Nov. 14, 2016, https://www.nhtsa.gov/national-automotive-sampling-system-crashworthiness-data-system (accessed Jan. 13, 2021).

[248] “NCSA I Tools, Publications, and Data,” https://cdan.nhtsa.gov/ (accessed Jan. 14, 2021).

[249] “Event Data Recorders (EDRs),” NHTSA, Aug. 02, 2016, https://www.nhtsa.gov/fmvss/event-data-recorders-eds (accessed Jan. 13, 2021).

[250] C. Flanagan et al., “Mutual Recognition Methodology Development,” 2014. Accessed: Jan. 14, 2021. [Online]. Available: https://www.dotbeer.org/pdf/16-0837.pdf.

[251] T. Unger, J. L. Haslbeck, and S.-A. Schneider, “ADAC accident research accident analysis based simulation of the most dangerous scenarios,” in Proceedings of the 34th International Congress on Road Safety, Munich, Germany, 2011, pp. 319–324. doi: 10.1109/IVS.2011.2157342.

[252] M. Tatar, “Enhancing ADAS Test and Validation with Automated Search for Critical Situations,” presented at the Driving Simulation Conference Exhibition, Tübingen, Germany, Sep. 2015.

[253] G. Bagschik, T. Menzel, and M. Maurer, “Ontology based Scenario Creation for the Development of Automated Vehicles,” in 2018 IEEE Intelligent Vehicles Symposium, vol. 5, no. 1, pp. 42–57, Spring 2013, doi: 10.1109/MIVS.2012.2217995.

[254] M. Shulman and R. K. Deering, “Third Annual Report of the Crash Avoidance Metrics Partnership,” National Highway Traffic Safety Administration Office of Advanced Safety Research Advanced Technology Division, DOT HS 809 837, Jan. 2005.

[255] M. Tatar, “Enhancing ADAS Test and Validation with Automated Search for Critical Situations,” presented at the Driving Simulation Conference Exhibition, Tübingen, Germany, Sep. 2015.

[256] J. Karl, G. Berg, F. Ruger, and B. Farber, “Driving Behavior and Simulator Sickness While Driving the Vehicle in the Loop: Validation of Longitudinal Driving Behavior,” IEEE Intelligent Transportation Systems Magazine, vol. 5, no. 1, pp. 42–57, Spring 2013, doi: 10.1109/MIVS.2012.2217995.

[257] D. Meissner, S. Reuter, E. Strigel, and K. Dietmayer, “Intersection-Based Preceding Vehicle Tracking Using a Classifying Multiple-Model PHD Filter,” IEEE Robotics and Automation Letters, vol. 4, no. 3, pp. 1373–1379, Jul. 2019.

[258] A. Eidehall and L. Petersson, “Statistical Threat Assessment for General Situation Awareness by autonomous vehicles,” Research in progress supported by the EPSRC, grant numbers EP/L024942/1 and EP/J018143/1., IFAC-PapersOnLine, vol. 50, no. 1, pp. 7038–7043, Jul. 2017, doi: 10.1016/j.ifacol.2017.08.1349.

[259] Y. Wang, Z. Liu, Z. Zuo, Z. Li, L. Wang, and X. Luo, “Trajectory Planning and Safety Assessment of Autonomous Vehicles Based on Motion Prediction and Model Predictive Control,” in IEEE Transactions on Vehicular Technology, vol. 68, no. 9, pp. 8546–8556, Sep. 2019, doi: 10.1109/TVT.2019.2930684.

[260] S. Lefèvre, D. Vassez, and C. Laugier, “A survey on motion prediction and risk assessment for intelligent vehicles,” ROBOMECH Journal, vol. 1, no. 1, p. 1, Jul. 2014, doi: 10.1186/s41004-014-0001-z.

[261] J. Nilsson, M. Brännström, J. Fredriksson, and E. Coelingh, “Longitudinal and Lateral Control for Autonomous Yielding Maneuvers,” IEEE Transactions on Intelligent Transportation Systems, vol. 17, no. 5, pp. 1404–1414, May 2016, doi: 10.1109/TITS.2015.2504718.

[262] J. Nilsson, M. Brännström, E. Coelingh, and J. Fredriksson, “Lane Change Maneuvers for Automated Vehicles,” in IEEE Transactions on Intelligent Transportation Systems, vol. 18, no. 5, pp. 1087–1096, May 2017, doi: 10.1109/TITS.2016.2597966.

[263] R. R. Rajamani, Vehicle Dynamics and Control, Springer US, 2006, doi: 10.1007/978-0-387-28823-6.

[264] M. Brännström, E. Coelingh, and J. Sjöberg, “Model-based threat assessment for avoiding arbitrary vehicle collisions,” Trans. Intell. Transport. Sys., vol. 11, no. 3, pp. 658–669, Sep. 2010, doi: 10.1109/TITS.2010.2048314.

[265] N. Kaempchen, B. Schiele, and K. Dietmayer, “Situation assessment of an autonomous emergency braking for arbitrary vehicle-to-vehicle collision scenarios,” Trans. Intell. Transport. Sys., vol. 10, no. 4, pp. 678–687, Dec. 2009, doi: 10.1109/TITS.2009.2026452.

[266] A. Broadhurst, S. Baker, and T. Kanade, “Monte Carlo road safety reasoning,” in IEEE Proceedings of the Intelligent Vehicles Symposium, 2005, Jun. 2005, pp. 319–324. doi: 10.1109/IVS.2005.1505122.

[267] A. Eidehall and L. Petersson, “Statistical Threat Assessment for General Road Scenarios Using Monte Carlo Sampling,” Trans. Intell. Transport. Sys., vol. 9, no. 1, pp. 137–147, Mar. 2008, doi: 10.1109/TITS.2007.909241.

[268] K. Okamoto, K. Berntorp, and S. Di Cairano, “Driver-Intention Based Vehicle Threat Assessment Using Random Forests and Particle Filtering,” IFAC-PapersOnLine, vol. 50, no. 1, pp. 13860–13865, Jul. 2017, doi: 10.1016/j.ifacol.2017.08.2231.

[269] S. Ammour and F. Nashedi, “Real time trajectory prediction for collision risk estimation between vehicles,” in 2009 IEEE 5th International Conference on Intelligent Computer Communication and Processing, Cluj-Napoca, Romania, Aug. 2009, pp. 417–422, doi: 10.1109/ICCP.2009.5284727.

[270] D. Meissner, R. Sturutz, and K. Dietmayer, “Intersection-Based Road User Tracking Using a Classifying Multiple-Model PHD Filter,” IEEE Intelligent Transportation Systems Magazine, vol. 6, no. 2, pp. 21–33, Summer 2014, doi: 10.1109/MITS.2014.2304754.

[271] M. Althoff, O. Stursburg, and M. Buss, “Model-Based Probabilistic Collision Detection in Autonomous Driving,” IEEE Trans. In-
HESHAM ALGHODHAIFI (Student Member, IEEE) received the B.S. degree (Hons.) in physics from Al-Baath University, Homs, Syria, in 2011, and the M.S.E. degree (High Distinction) from the University of Michigan, USA, in 2017. He is currently pursuing the Ph.D. degree with the University of Michigan-Dearborn, USA. His current research focuses on image processing, machine learning, and autonomous vehicle modeling and evaluation.

SRIDHAR LAKSHMANAN Sridhar Lakshmanan received the B.S. degree in electronics and communications engineering from the Birla Institute of Technology, Mesra, India, in 1985, and the M.S. and Ph.D. degrees in electrical and computer engineering from the University of Massachusetts, Amherst, in 1987 and 1991, respectively. In 1991, he joined the faculty at the University of Michigan, Dearborn, as an Assistant Professor of Electrical and Computer Engineering. His current research interests are in the areas of game design, machine learning, optimization, intelligent systems, and wearable sensors. He is also a recipient of several research contracts that address problems in Intelligent Transportation Systems, and serves as a consultant to the industry and the government on these problems as well. In 1996, he founded M-Vision, Inc., a company specializing in the application of computer vision to automotive problems.

In May 1997, he was promoted to an Associate Professor, with tenure. His current research interests are in the areas of game design, machine learning, optimization, intelligent systems, and wearable sensors. He is also a recipient of several research contracts that address problems in Intelligent Transportation Systems, and serves as a consultant to the industry and the government on these problems as well. In 1996, he founded M-Vision, Inc., a company specializing in the application of computer vision to automotive problems.