RACL: Risk Aware Closed-Loop Agent Simulation with High Fidelity

Qiujing Lu Ruoxuan Bai Shu Li
Honglin He Shuo Feng[✉]
Department of Automation, Tsinghua University
{brx22, li-s23, hehl21}@mails.tsinghua.edu.cn
{qiujinglu, fshuo}@mail.tsinghua.edu.cn

Abstract

The progression in autonomous driving technologies has underscored the necessity for rigorous testing and validation methods to ensure safety and reliability. A primary challenge in establishing such a testing framework lies in the development of realistic and reactive simulated agents, for both routine driving scenarios and safety-critical situations. To address this, we introduce Risk Aware Closed-Loop agent simulation(RACL), a comprehensive framework that respects the safety critical states and captures the natural progression of risk. By distinctly modeling normal and risk states, we ensure efficient utilization of limited crash data and high fidelity for both modes. Additionally, our approach employs risk monitoring and mode transition strategies to enable a smooth shift from normal to risky scenarios. We have provided both quantitative and qualitative results to demonstrate the efficacy of our system.

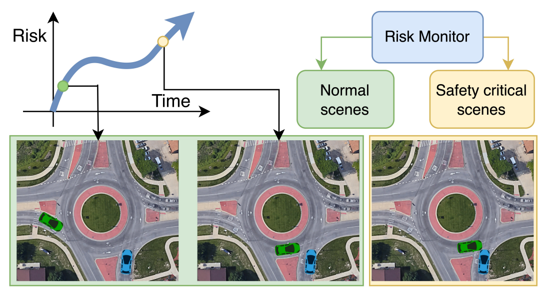


Figure 1. RACL is designed to enhance the generation of realistic driving scenarios by integrating risk assessment into the simulation process. It uses a learnt predictor to estimate the risk of the whole scene at each timestep. Once high risk is detected, it automatically switches from the normal driving policy to risky driving policy for generating realistic safety-critical scenes.

1. Introduction

The rise of autonomous vehicles (AVs) marks a major shift in transportation, aiming to improve mobility, traffic efficiency, and reduce accidents. Yet, realizing full autonomy poses challenges, especially in ensuring safety and reliability.Central to addressing these challenges is the development of simulation testing methodologies, considering the high costs (both financial and in terms of social trust) and inefficiencies associated with road testing. To make simulation a viable tool for the development and testing of AV, simulators need to produce agents that are both realistic and responsive. These agents should facilitate the creation of high-fidelity scenarios, encompassing both typical driving situations and safety-critical events, that mirror real-world occurrences. This capability is crucial for enabling the rigorous evaluation of AV performance across a broad spectrum of driving conditions.

However, human driving behavior is characterized by high uncertainty due to diverse intentions and driving styles, making the construction of such generative models a pressing scientific challenge that needs to be addressed. Adding to the complexity is the reality that human drivers exhibit an average accident rate of approximately 1.9×10^{-6} per mile [12]. The scarcity of safety incident data poses substantial hurdles in developing data-driven simulation models. Nevertheless, considering the societal risks involved, there is a critical need for these models to achieve extensive coverage and maintain a stringent low error tolerance.

To address these challenges, this paper introduces a novel approach to enable the realistic simulator for generating full spectrum realistic scenarios including normal driving and safety-critical states, leveraging both empirical data and advanced simulation techniques. Our contribution includes:

- Introduced a risk monitoring mechanism that tracks the evolution of risk throughout the simulation process.
- Proposed a framework for searching and generating risky states, utilizing accident data efficiently.
- Attained a high level of realism in simulating both normal driving conditions and safety-critical scenarios.

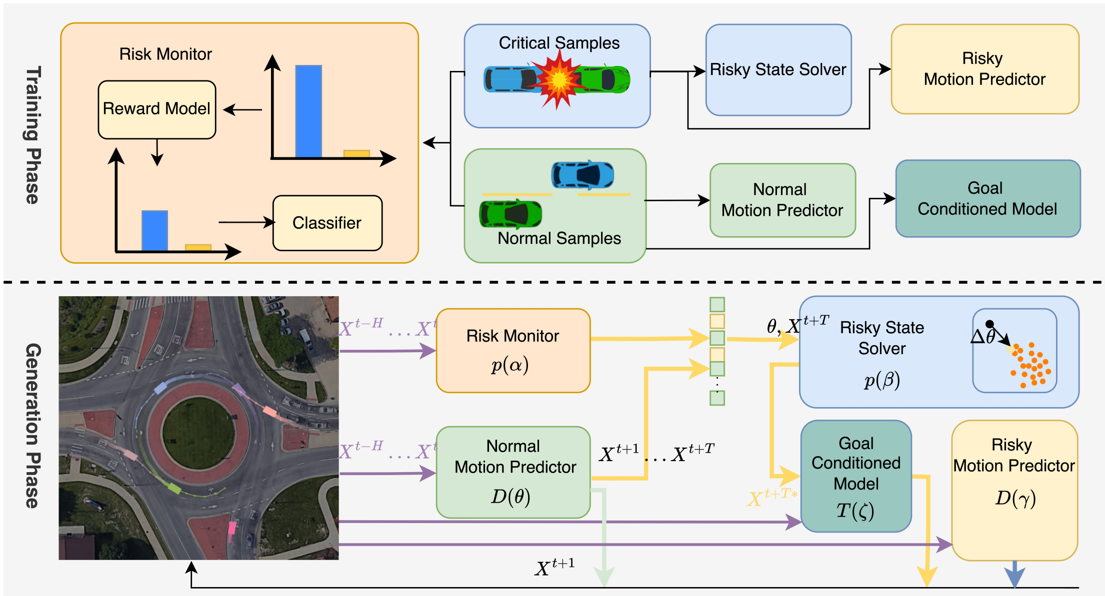


Figure 2. Overview of our proposed system, RACL. **Bottom:** During the generation phase, the risk monitor analyzes the current state and past information $X^t \dots X^{t-H}$. Depending on the assessed risk level, RACL chooses between two pathways:(1) For low risk, it employs a normal motion predictor $D(\theta)$ for next step's action.(2) For high risks, it utilizes a VAE encoder $p(\beta)$ and $D(\theta)$ to identify an optimal risky future state X^{t+T*}. Upon determining this state, the framework generates subsequent motions using a goal-conditioned model and a risky motion policy. **Top:** It illustrates our training pipeline and the flow of training data into each module.

2. Related work

Traffic simulation. Replaying logged behavior of all vehicles in the scene can ensure the realistic behavior of each agent while the whole environment can't respond to new behavior of AVs, leading to in simulation results. Creating reactive agents that can respond in a human-like way is crucial [15]. Traditional rule-based models, such as the Intelligent Driver Model (IDM) [20] and MOBIL [13] are overly accommodating and lack the diversity of human driving behaviors on a micro level, not strict and random enough to meet the current testing needs. Data-driven deep learning models are more promising than rule based ones considering their adaptability and capacity to learn complex patterns. TrafficSim[19] utilizes GRU [5] and CNN to mimic the behaviors of human drivers from real-world data. [23] applies Transformer[21] with receding horizon prediction mechanism for enhanced realism. TrafficGen[7] uses an autoregressive neural generative model and Symphony [11] combines learnt policies with a parallel beam search to improves realism. Most work are focusing on imitating human's driving behavior and evaluated on naturalistic scenarios collected from daily life with rare accidents.

Safety-critical scenario generation. Simultaneously, there have been several efforts aimed at creating safety-critical scenarios, addressing the need for efficient testing[6, 9]. Some work leverage prior knowledge: [16] uses equations to generate random risk scenes while [2] reduces the drivable area by manipulating the surrounding vehicles via constraint optimization. Additionally, Waymo adopts a data-driven approach to recreate fatal accidents by perturbing key parameters based on collected data, as detailed in [18]. Furthermore, to actively create risky scenarios, some methods challenge the AV system in an adversarial manner. AdvSim [22] crafts adversarial agent behaviors and updates AV's LiDAR sensor data accordingly to test system vulnerabilities. STRVE [17] learns a graph-based conditional VAE as traffic prior, optimizing each agent's behavior to provoke collisions with a rule-based AV planner while KING [10] utilizes kinematics gradients to find adversarial behaviors targeting the actual AV. [4, 25] apply diffusion models with a collision cost to promote adversarial agents during the inference stage. However, modifying initial conditions or employing adversarial tactics can diminish the realism of agent behaviors and, by extension, the entire simulation scenario.

While these strategies are effective for short-term testing, there is a growing need for a comprehensive framework capable of learning and adapting realistic agent reactions in both normal driving scene and evolving safety-critical driving scenarios for more extended simulation periods, as pointed out in [8, 24].

3. Method

We represent the states of N agents at time t as $X^t = [x_1^t, x_2^t, \dots, x_N^t]$, where each agent's state, x_i^t, encompasses its current 2D position and heading information. An accident that occurs after time step t is designated as event A_t. The risk level of the current situation is estimated by the probability of event A_t, denoted by $\mathcal{P}(A_t|X^t \dots X^{t-H})$, with the system having access to H steps of historical information. The set of risky states is defined as $S = \{X^t|\mathcal{P}(A_t|X^t \dots X^{t-H}) \geq \tau\}$ with τ as a specified threshold and the set of normal state as S^c correspondingly.

3.1. System Overview

As shown in Figure 2, there are three main components in our system: (1) A risk monitor that identifies the critical timestep t^c and risky agents within the scene. (2) two motion predictors which capture the intrinsic differences in driving strategies and uncertainty in human-driven vehicles, $D(\theta)$ for normal scenes and $D(\gamma)$ for safety critical scenes; (3) risky state model $p_\beta(X^t|X^t \in S)$ that learns the distribution of risky states, and a goal-conditioned model $T_\zeta(X^{t+1} \dots X^{t+T}|X^t \dots X^{t-H}, X_{goal})$ that imitates human driving behavior with the target state X_{goal} as constraint.

For a given moment t, the training goal is to maximize the likelihood of the subsequent T steps given the history of the past H steps. This can be formulated as Equation (1):

$$\arg\max_{\theta,\gamma,\zeta} \mathbb{E}_{X^{t+T}\dots X^{t-H} \sim p_{data}} \mathcal{L}(\theta, \gamma, \zeta|X^{t+T} \dots X^{t-H}) \tag{1}$$

where the likelihood function is detailed below:

$$\begin{aligned}
&\mathcal{L}(\theta, \gamma, \zeta|X^{t+T} \dots X^{t-H}) \\
=&p_{\theta,\gamma,\zeta}(X^{t+1} \dots X^{t+T}|X^t \dots X^{t-H}) \\
=&D_\gamma \mathbb{1}(X^t \in S) + D_\theta \mathbb{1}(X^{t+T} \in S^c) + T_\zeta \mathbb{1}(X^t \in S^c, X^{t+T} \in S)
\end{aligned}$$

We optimize θ, γ and ζ independently, each addressing a different aspect: normal driving behaviors, high-risk situations, and the transition phase, respectively. The architecture details are introduced in Sections 3.2 to 3.4

3.2. Risk Monitor

The risk estimator is trained independently from other componentsand operates through a dual-stage binary classification mechanism. In the initial phase, an unsupervised reward model is employed to distinguish accident data from typical driving data. Following this, potential accident data are further analyzed by another classifier. This classifier utilizes an enhanced Bilateral-Branch Network (BBN)[3, 26] with supervised training for a more nuanced classification, assessing the presence of risky situations. This multi-layered strategy extracts key features from imbalanced data, achieving high precision.

3.3. Motion predictor

For the normal motion predictor $D(\theta)$, Transformer is used as its backbone to learn from normal human driving data and more training details can be found in [24]. It learns to minic human driving behaviors in normal situations while adhering to vehicle dynamics constraints. i.e. $\arg\min_\theta \mathbb{E}_{X^{t+T}\dots X^{t-H} \sim \mathcal{P}(data|X^{t+T} \in S^c)} p_\theta(X^{t+1} \dots X^{t+T}|X^t \dots X^{t-H})$. Similarly, we trained D_γ with accident data.

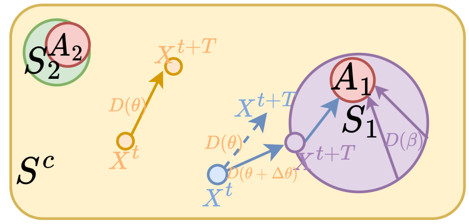


Figure 3. Demonstration for trajectories rolled out by normal policy (indicated in yellow) and risky policy activated upon risk detection (shown in blue). See more details in Section 3.4

3.4. Mode Transition

Risky State Solver We formulate the task of finding the most likely risky state at time step $t+T$ given $X^t \dots X^{t-H}$ as an optimization problem shown in Equation (2):

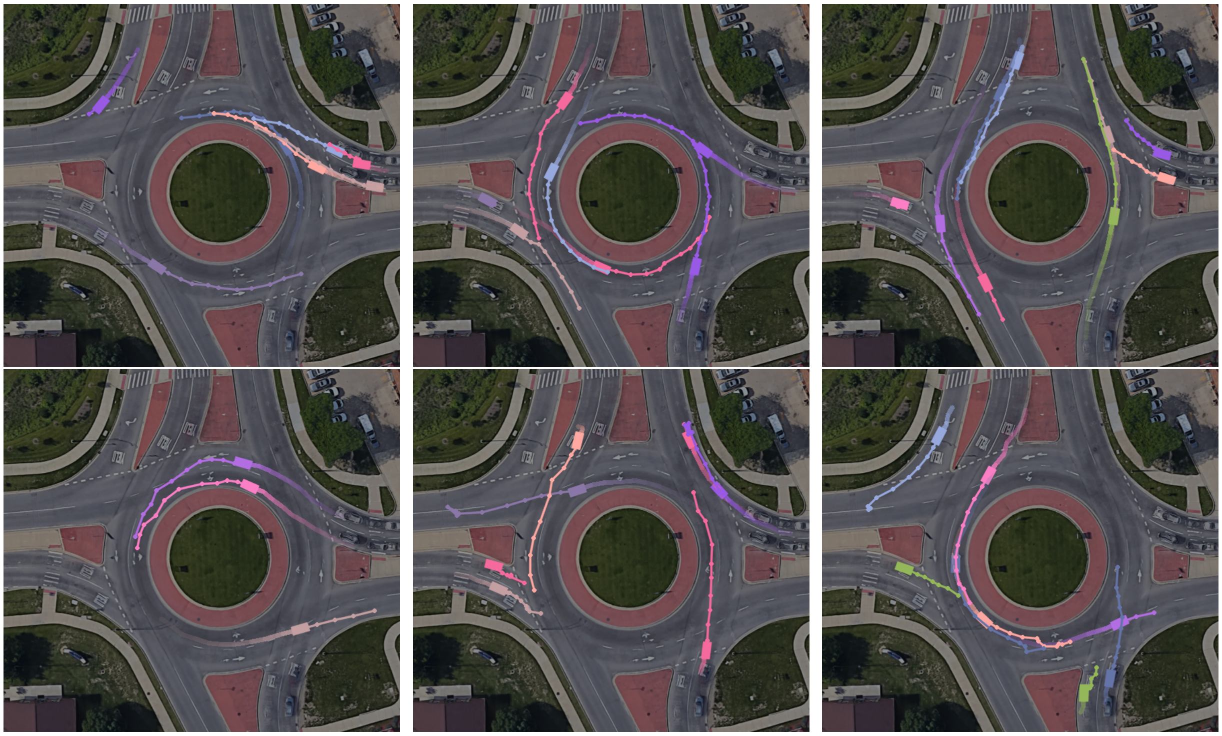
$$\begin{aligned}
\max_{\Delta\theta} \quad & p_\beta(X^{t+T}) \\
\text{s.t.} \quad & X^{t+1} \dots X^{t+T} \sim D_{\theta+\Delta\theta}(X^t \dots X^{t-H}) \\
& ||\Delta\theta|| \leq \epsilon
\end{aligned} \tag{2}$$

Previously trained motion predictor D_θ is used to construct the initial parameter space for locating the dynamically most likely states in the next T steps given the past H steps information. Simultaneously, a Variational Autoencoder (VAE) is learnt in an unsupervised manner to capture the distribution of risky states. The encoder of this VAE, acting as a proxy $p_\beta(X^{t+T})$ to guide the search for the most likely risky states near initial state X^{t+T}.
Goal Conditioned Model Once the optimal solution X^{t+T} is found for 2, a learnt goal conditioned model T_ζ transports the agent from current state X^t to target state X^{t+T} realistically, such that $X^{t+T} \in S$ and $X^{t+T} \approx X_{goal}$.

4. Experiments

4.1. Implementation Details

We chose roundabout as static environment for our experiment, recognizing it as an challenging urban driving environment for AVs, to validate our generation framework. This real-world dataset was collected from a two-lane roundabout located in Ann Arbor, Michigan, USA. All the training and evaluation were conducted on this dataset, adhering to the preprocessing steps outlined in NeuralNDE , which served as our baseline also.



(a) Sideswipe (b) Angle (c) Rear-end

Figure 4. Sampled simulation traces centered at 4 seconds before the noted accident with shadow lines to indicate past trajectories and dotted lines for future paths. The top section depicts Neural NDE, while the bottom section presents RACL.

Method	D_{KL}(Speed)↓	D_{KL}(Distance)↓	D_{KL}(Yielding$_v$) ↓	D_{KL}(Yielding$_d$) ↓	D_{KL}(PET)↓
SUMO*	0.126	0.089	0.087	0.107	-
NeuralNDE[24]	0.008	0.004	0.005	0.005	0.017
RACL	**0.005**	**0.003**	0.005	0.005	**0.016**

Table 1. Quantitative evaluation of simulation traces at the two-lane roundabout. Results of SUMO are queried from [24].

4.2. Results

Quantative results As shown in Table 1, the speed and distance distribution of the agents generated by our model are statistically closer to the ground truth than those produced by both SUMO [14] and NeuralNDE [24]. In yielding situations, our results significantly surpass those of SUMO, while showcasing distance and velocity metrics comparable to those generated by NeuralNDE, which aligns with expectations given NeuralNDE's proficiency in imitation the whole driving set. Our improvements are primarily evident in the transitions between different driving states. Utilizing Post Encroachment Time (PET)[1] to gauge scenario complexity and potential conflict, we outperformed NeuralNDE, showcasing our refined ability to handle and depict more intricate driving conditions.

Qualitative results As depicted in Figure 4, our approach successfully generates realistic interactions for various accident types. NeuralNDE often results in accidents within crowded scenarios, whereas our method enables a more diverse and complex range of interactions, effectively capturing the evolution of risk.

5. Conclusion and Future work

This paper presents a system capable of generating realistic scenarios featuring both safe and safety-critical situations. It opens up the opportunity to rapidly adapt to various maps and local accident data for specific autonomous driving tests. Future work will focus on introducing additional distributional metrics to quantify the behaviors of simulated agents on a finer grid. In the long term, we aim to use these learned agents to train more robust and safer autonomous driving systems, thereby closing the loop.

References

[1] Brian L Allen, B Tom Shin, and Peter J Cooper. Analysis of traffic conflicts and collisions. Technical report, 1978. 4

[2] Matthias Althoff and Sebastian Lutz. Automatic generation of safety-critical test scenarios for collision avoidance of road vehicles. In *2018 IEEE Intelligent Vehicles Symposium (IV)*, page 1326–1333, 2018. 2

[3] Ruoxuan Bai, Jingxuan Yang, Weiduo Gong, Yi Zhang, Qiujing Lu, and Shuo Feng. Accurately predicting probabilities of safety-critical rare events for intelligent systems. *arXiv preprint arXiv:2403.13869*, 2024. 3

[4] Wei-Jer Chang, Francesco Pittaluga, Masayoshi Tomizuka, Wei Zhan, and Manmohan Chandraker. Controllable safety-critical closed-loop traffic simulation via guided diffusion. *arXiv preprint arXiv:2401.00391*, 2023. 2

[5] Kyunghyun Cho, Bart Van Merriënboer, Caglar Gulcehre, Dzmitry Bahdanau, Fethi Bougares, Holger Schwenk, and Yoshua Bengio. Learning phrase representations using rnn encoder-decoder for statistical machine translation. *arXiv preprint arXiv:1406.1078*, 2014. 2

[6] Wenhao Ding, Chejian Xu, Mansur Arief, Haohong Lin, Bo Li, and Ding Zhao. A survey on safety-critical driving scenario generation—a methodological perspective. *IEEE Transactions on Intelligent Transportation Systems*, 24(7): 6971–6988, 2023. 2

[7] Lan Feng, Quanyi Li, Zhenghao Peng, Shuhan Tan, and Bolei Zhou. Trafficgen: Learning to generate diverse and realistic traffic scenarios. In *2023 IEEE International Conference on Robotics and Automation (ICRA)*, page 3567–3575, 2023. 2

[8] Shuo Feng, Haowei Sun, Xintao Yan, Haojie Zhu, Zhengxia Zou, Shengyin Shen, and Henry X Liu. Dense reinforcement learning for safety validation of autonomous vehicles. *Nature*, 615(7953):620–627, 2023. 2

[9] Daniel J Fremont, Edward Kim, Tommaso Dreossi, Shromona Ghosh, Xiangyu Yue, Alberto L Sangiovanni-Vincentelli, and Sanjit A Seshia. Scenic: A language for scenario specification and data generation. *Machine Learning*, 112(10):3805–3849, 2023. 2

[10] Niklas Hanselmann, Katrin Renz, Kashyap Chitta, Apratim Bhattacharyya, and Andreas Geiger. King: Generating safety-critical driving scenarios for robust imitation via kinematics gradients. In *Computer Vision – ECCV 2022*, page 335–352, Cham, 2022. Springer Nature Switzerland. 2

[11] Maximilian Igl, Daewoo Kim, Alex Kuefler, Paul Mougin, Punit Shah, Kyriacos Shiarlis, Dragomir Anguelov, Mark Palatucci, Brandyn White, and Shimon Whiteson. Symphony: Learning realistic and diverse agents for autonomous driving simulation. In *2022 International Conference on Robotics and Automation (ICRA)*, page 2445–2451, 2022. 2

[12] Nidhi Kalra and Susan M Paddock. Driving to safety: How many miles of driving would it take to demonstrate autonomous vehicle reliability? *Transportation Research Part A: Policy and Practice*, 94:182–193, 2016. 1

[13] Arne Kesting, Martin Treiber, and Dirk Helbing. General lane-changing model mobil for car-following models. *Transportation Research Record*, 1999(1):86–94, 2007. 2

[14] Pablo Alvarez Lopez, Michael Behrisch, Laura Bieker-Walz, Jakob Erdmann, Yun-Pang Flötteröd, Robert Hilbrich, Leonhard Lücken, Johannes Rummel, Peter Wagner, and Evamarie Wießner. Microscopic traffic simulation using sumo. In *2018 21st international conference on intelligent transportation systems (ITSC)*, pages 2575–2582. IEEE, 2018. 4

[15] Nico Montali, John Lambert, Paul Mougin, Alex Kuefler, Nicholas Rhinehart, Michelle Li, Cole Gulino, Tristan Emrich, Zoey Yang, Shimon Whiteson, et al. The waymo open sim agents challenge. *Advances in Neural Information Processing Systems*, 36, 2024. 2

[16] Ashish Rana and Avleen Malhi. Building safer autonomous agents by leveraging risky driving behavior knowledge. In *2021 International Conference on Communications, Computing, Cybersecurity, and Informatics (CCCI)*, page 1–6, 2021. 2

[17] Davis Rempe, Jonah Philion, Leonidas J. Guibas, Sanja Fidler, and Or Litany. Generating useful accident-prone driving scenarios via a learned traffic prior. In *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, page 17284–17294, New Orleans, LA, USA, 2022. IEEE. 2

[18] John M. Scanlon, Kristofer D. Kusano, Tom Daniel, Christopher Alderson, Alexander Ogle, and Trent Victor. Waymo simulated driving behavior in reconstructed fatal crashes within an autonomous vehicle operating domain. *Accident Analysis Prevention*, 163:106454, 2021. 2

[19] Simon Suo, Sebastian Regalado, Sergio Casas, and Raquel Urtasun. Trafficsim: Learning to simulate realistic multi-agent behaviors. page 10400–10409, 2021. 2

[20] Martin Treiber, Ansgar Hennecke, and Dirk Helbing. Congested traffic states in empirical observations and microscopic simulations. *Physical Review E*, 62(2):1805–1824, 2000. 2

[21] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. Attention is all you need. *Advances in neural information processing systems*, 30, 2017. 2

[22] Jingkang Wang, Ava Pun, James Tu, Sivabalan Manivasagam, Abbas Sadat, Sergio Casas, Mengye Ren, and Raquel Urtasun. Advsim: Generating safety-critical scenarios for self-driving vehicles. page 9909–9918, 2021. 2

[23] Yu Wang, Tiebiao Zhao, and Fan Yi. Multiverse transformer: 1st place solution for waymo open sim agents challenge 2023. (arXiv:2306.11868). 2

[24] Xintao Yan, Zhengxia Zou, Shuo Feng, Haojie Zhu, Haowei Sun, and Henry X Liu. Learning naturalistic driving environment with statistical realism. *Nature communications*, 14 (1):2037, 2023. 2, 3, 4

[25] Ziyuan Zhong, Davis Rempe, Danfei Xu, Yuxiao Chen, Sushant Veer, Tong Che, Baishakhi Ray, and Marco Pavone. Guided conditional diffusion for controllable traffic simulation. In *2023 IEEE International Conference on Robotics and Automation (ICRA)*, page 3560–3566, 2023. 2

[26] Boyan Zhou, Quan Cui, Xiu-Shen Wei, and Zhao-Min Chen. Bbn: Bilateral-branch network with cumulative learning for long-tailed visual recognition. In *Proceedings of*

the IEEE/CVF conference on computer vision and pattern recognition, pages 9719–9728, 2020. 3