

中国公路学报
China Journal of Highway and Transport
ISSN 1001-7372,CN 61-1313/U

《中国公路学报》网络首发论文

题目：用于自动驾驶风险场景演化的交通车辆对抗行为优化
作者：蔡翹，白雪松，任毅龙，于海洋，文啸，丁川，田长航，崔志勇
收稿日期：2026-02-25
网络首发日期：2026-08-20
引用格式：蔡翹，白雪松，任毅龙，于海洋，文啸，丁川，田长航，崔志勇. 用于自动驾驶风险场景演化的交通车辆对抗行为优化[J/OL]. 中国公路学报.
<https://link.cnki.net/urlid/61.1313.U.20260820.1253.002>



网络首发：在编辑部工作流程中，稿件从录用到出版要经历录用定稿、排版定稿、整期汇编定稿等阶段。录用定稿指内容已经确定，且通过同行评议、主编终审同意刊用的稿件。排版定稿指录用定稿按照期刊特定版式（包括网络呈现版式）排版后的稿件，可暂不确定出版年、卷、期和页码。整期汇编定稿指出版年、卷、期、页码均已确定的印刷或数字出版的整期汇编稿件。录用定稿网络首发稿件内容必须符合《出版管理条例》和《期刊出版管理规定》的有关规定；学术研究成果具有创新性、科学性和先进性，符合编辑部对刊文的录用要求，不存在学术不端行为及其他侵权行为；稿件内容应基本符合国家有关书刊编辑、出版的技术标准，正确使用和统一规范语言文字、符号、数字、外文字母、法定计量单位及地图标注等。为确保录用定稿网络首发的严肃性，录用定稿一经发布，不得修改论文题目、作者、机构名称和学术内容，只可基于编辑规范进行少量文字的修改。

出版确认：纸质期刊编辑部通过与《中国学术期刊（光盘版）》电子杂志社有限公司签约，在《中国学术期刊（网络版）》出版传播平台上创办与纸质期刊内容一致的网络版，以单篇或整期出版形式，在印刷出版之前刊发论文的录用定稿、排版定稿、整期汇编定稿。因为《中国学术期刊（网络版）》是国家新闻出版广电总局批准的网络连续型出版物（ISSN 2096-4188，CN 11-6037/Z），所以签约期刊的网络版上网络首发论文视为正式出版。

基金项目：国家重点研发计划、自主式交通系统技术集成应用（2024YFB4303103）；国家自然科学基金、面向自主式交通系统的知识增强异智主体多尺度交互协同（52441202）；北京航空航天大学敢为行动计划项目、生成式数据驱动的车路云一体化自动驾驶感知技术研究与应用（WZ2024-2-16）；国家自然科学基金卓越研究群体项目 A 类、自动驾驶交通智能控制（T2588101）；山东省重点研发计划（重大科技创新工程）、AI 驱动的智能网联汽车智能安全加速测试技术（2025CXGC010109）；中央高校基本科研业务费资助

作者简介：蔡翺（1998-），男，四川德阳，无，工学硕士学位（博士在读）

崔志勇（1989-），河北承德，教授（博导），博士学位

National Key R&D Program of China, [Grant No. 2024YFB4303103]; National Natural Science Foundation of China (project number: 52441202); Ganwei Program of Beihang University (project number: WZ2024-2-16); National Natural Science Foundation of China (Grant No. T2588101); Shandong Provincial Key R&D Program, [Grant No. 2025CXGC010109]; Fundamental Research Funds for the Central Universities

用于自动驾驶风险场景演化的交通车辆对抗行为优化

蔡翺¹, 白雪松¹, 任毅龙¹, 于海洋¹, 文啸², 丁川¹, 田长航¹,
崔志勇*¹

- (1. 北京航空航天大学交通科学与工程学院 车路一体智能交通全国重点实验室, 北京市, 100080
2. 香港科技大学土木与环境工程系, 中国香港, 999077)

摘要：风险场景是评估自动驾驶系统鲁棒性的关键数据。针对自然驾驶数据中长尾边缘场景稀缺，以及实车道路测试成本高、风险不可控等问题，利用仿真技术生成高价值风险场景已成为重要途径。然而，现有生成方法难以兼顾场景的挑战性、丰富性、真实性、可控性与可解性，导致生成数据在算法闭环验证中的可用率较低。为此，本研究提出一种面向异质驾驶风格的多任务协同风险场景演化方法。该方法受贝叶斯概率分解思想启发，将生成过程组织为挑战对手选择、风格化行为控制与风险场景合成三个子任务。首先，构建基于监督深度学习的对抗性预测网络，精准锚定潜在风险最大的背景车辆作为交互对手；其次，利用最大熵逆强化学习方法量化并建模多维度的驾驶风格，实现对手行为的风格化多样性控制；最后，基于运动学优化器对手车辆进行增量控制序列优化，生成兼具运动合理性与对抗性的仿真场景。多场景拓扑下的实验结果表明，与现有前沿方法相比，本研究在各项性能指标上表现出最优的综合平衡能力。进一步将生成数据混合用于端到端自动驾驶策略的模仿学习微调，结果显示自动驾驶系统的碰撞率显著下降且性能可控，验证了该方法在自动驾驶测试评估与模型训练中的应用价值。

关键词：汽车工程；自动驾驶测试；边缘场景生成；多任务协同；对抗行为优化

Adversarial Behavior Optimization of Traffic Vehicles for Risky Scenario Evolution in Autonomous Driving

Xuan Cai¹, Xuesong Bai¹, Yilong Ren¹, Haiyang Yu¹, Xiao Wen², Chuan Ding¹,
Changhang Tian¹, Zhiyong Cui^{*1}

- (1. State Key Laboratory of Intelligent Transportation Systems, School of Transportation Science and Technology, Beihang University, Beijing 100080, P.R.China.
(2. Department of Civil and Environmental Engineering, the Hong Kong University of Science and Technology, Hong Kong, 999077, China.)

Abstract: Risk scenarios constitute critical data for evaluating the robustness of autonomous driving systems. Addressing the scarcity of long-tail edge cases in naturalistic driving data, as well as the prohibitive costs and uncontrollable risks associated with on-road testing, generating high-value risk scenarios via simulation has emerged as a pivotal approach. However, existing generation methods struggle to balance multiple performance metrics—namely challenge level, diversity, realism, controllability, and solvability—resulting in low usability of the generated data for closed-loop algorithm validation. To address this challenge, this study proposes a multi-task collaborative risk scenario evolution method for heterogeneous driving styles. Inspired by Bayesian factorization, the generation process is organized into three coordinated subtasks: challenging adversary selection, stylized behavioral control, and risk scenario synthesis. Specifically, an adversarial prediction network based on supervised deep learning is first constructed to accurately target the background vehicle with the highest potential risk as the interactive adversary. Subsequently, Maximum Entropy Inverse Reinforcement Learning (IRL) is utilized to quantify and model multi-dimensional driving styles, achieving stylized diversity control over the opponent's behavior. Finally, a kinematic optimizer is employed to refine the incremental control sequences of the opponent vehicle, generating simulation scenarios that possess both kinematic realism and adversarial characteristics. Experimental results across multiple road topologies demonstrate that, compared with existing state-of-the-art methods, the proposed method exhibits the best comprehensive balance across various performance metrics. Furthermore, the generated data are incorporated to fine-tune an end-to-end autonomous driving policy through imitation learning. The results demonstrate a significant reduction in the collision rate with controllable performance, validating the proposed method's value for both autonomous driving evaluation and model training.

Keywords: automotive engineering; autonomous driving testing; corner-case generation; multi-task collaboration; adversarial behavior optimization

Received 25 February 2026

中图分类号: U46

文献标志码: A

0 引言

自动驾驶 (Autonomous Driving, AD) 正处于由辅助驾驶向高级别自动化发展的关键阶段, 其在城市、高速及混合交通流等复杂场景下的安全性与鲁棒性是技术落地的核心。然而, 传统道路测试受“长尾效应”制约, 难以高效覆盖低频高危边缘场景 (Corner Cases), 同时还面临周期长、成本高、复现难和风险不可控等问题。因而, 场景驱动的仿真测试已成为自

自动驾驶验证的重要手段，可通过可控复现复杂交互工况，提高 AD 系统安全验证与迭代效率。

在仿真验证体系中，场景生成是决定测试有效性的核心环节，其旨在高效设计、合成或演化出兼具真实性和安全边界特征的测试用例，以充分暴露 AD 系统在任何可能现实条件下的潜在缺陷。综观现有研究，其研究范式可以分为 2 条主流路径：

① 证伪式：该方法聚焦于“点”，即主要考虑场景生成的挑战性，而场景的丰富性、真实性、可解性存在不足。例如，KING^[1]方法利用运动学梯度代理模型优化常规数据的挑战性，而由于梯度下降算法的局部最优特征容易导致场景较为单一。为解决该问题，有学者采用动态时间规整^[2]来提升生成代理的探索能力，但是可控性和计算负担问题又凸显^[3]，且难以扩展到多车环境^[4]。虽然有研究进一步采用大语言模型^[5-6]来增强其可控性并优化生成效率^[7]，但因为大模型的多模态理解能力不足，导致依然难以闭环调控对抗行为风格。

② 覆盖式：该方法聚焦于“域”，即主要考虑覆盖足够多的交互场景。例如，一些研究^[8-9]尽管通过离散场景空间并利用蒙特卡洛采样来生成高覆盖的场景参数，但往往忽视动力学约束而损失了真实性。为此，部分研究^[10-11]利用生成式模型^[12]或强化学习（Reinforcement Learning, RL）^[13]建模真实交通数据集，但是难以对抗强度进行精确引导，导致挑战性与可解性失衡。考虑到真实交通场景下人类驾驶不同的行为风格，有学者^[14]提出构建丰富的交互模式（如竞争、互惠、合作）以增强风险场景的丰富度，但由于缺乏有效的可解释性机制与控制变量，场景生成的行为可控性依然面临挑战。

在现有研究中，场景生成面临挑战性（边界极限逼近）、丰富性（行为分布多样）、真实性（物理一致性）、可控性（参数显式引导）与可解性（测试反馈闭环）之间的多维度权衡（Trade-off）。具体而言，当前研究仍存在以下核心瓶颈：① 高价值边界场景的有效利用率低下。现有对抗生成方法虽能触达高风险工况，但常导致场景“不可解”（如必撞点）或行为逻辑畸变，造成测试数据在算法闭环验证中的实际效能受限。② 交互主体行为风格的多样性缺失。多数方法将背景车辆视为单一的“对抗代理”^[15]，缺乏对驾驶员社会价值取向（如保守、激进、合作等）的显式建模，难以覆盖真实交通中异质主体触发的复杂风险分布。③ 生成任务与下游应用端的逻辑脱节。目前场景生成与自动驾驶算法训练（如模仿学习（Imitation Learning, IL）/RL 策略）之间缺乏耦合^[16]。鲜有研究验证所生成的对抗数据能否真实转化为下游模型性能的提升，导致生成端与验证端难以形成有效闭环。

针对现有方法难以兼顾场景生成中挑战性、真实性与可控性等的难题，受贝叶斯概率分解思想启发^[17]，本研究提出一种基于多任务协同的风险场景演化方法（Bayesian decomposition-based scenario generation）。该方法将复杂的生成任务解耦为 3 个串行子环节：首先，利用监督深度学习（Supervised Deep Learning, SDL）模型从真实轨迹数据集中精准锚定具有潜在高风险的背景车辆作为对抗目标；其次，通过最大熵逆强化学习（Maximum Entropy Inverse Reinforcement Learning, MaxEnt IRL）量化驾驶风格，构建风格化编辑损失函数以增强交互丰富性；最后，基于运动学优化器（Kinematics Optimizer, KO）对对抗车辆进行增量控制序列优化，从而在保证物理可解性的前提下演化出高价值风险场景。本文的主要创新点总结如下：

① 提出一种面向异质驾驶风格的风险场景演化方法。利用最大熵逆强化学习从自然驾驶轨迹中提取多维行为偏好，并将风格奖励嵌入对抗行为优化，实现具有差异化交互模式的可解释、可调控风险场景生成；

② 构建风险目标锚定、风格行为引导与运动学约束优化相协同的场景生成流程。受贝叶斯概率分解思想启发，将复杂生成过程组织为 3 个相互衔接的功能子任务，兼顾场景挑战性、丰富性、真实性与可解性；

③ 验证生成场景在自动驾驶测试评估与模型训练中的双重应用价值。实验结果表明，所生成场景在综合性能上优于对比方法，并能够有效提升自动驾驶策略在复杂交互工况下的安全性与鲁棒性。

1 场景生成问题建模与任务分解

本研究关心在观测到的交通环境 X 下发生高风险事件 χ_{nm} 的后验概率 $P(\chi_{nm}|X)$ 。将可能导致风险的潜在变量边缘化，并使用贝叶斯方程分解为联合分布和条件概率乘积：

$$\begin{aligned} P(\chi_{nm}|X) &= \int_{U_{ego}} \int_{U_{adv}} P(\chi_{nm}, U_{ego}, U_{adv}|X) dU_{ego} dU_{adv} \\ &= \int_{U_{ego}} \int_{U_{adv}} P(U_{adv}|X) P(U_{ego}|U_{adv}, X) P(\chi_{nm}|U_{ego}, U_{adv}, X) dU_{ego} dU_{adv} \end{aligned} \quad (1)$$

式中， $U_{ego} = \{u_{ego}^t\}_{t=0}^{T-1}$ ， $U_{adv} = \{u_{adv}^t\}_{t=0}^{T-1}$ 分别是自车（被测自动驾驶车）与对手车的未来控制序列（舵角/油门-制动）。由于自车是黑盒的，所以 U_{ego} 是未知的，本研究采取迭代优化获取 U_{ego} 历史值。为了最大化高风险事件概率：

$$\begin{aligned} \arg \max_{U_{adv}} P(\chi_{nm}|X) &= \arg \max_{U_{adv}} P(U_{adv}|X) \int_{U_{ego}} P(U_{ego}|U_{adv}, X) P(\chi_{nm}|U_{ego}, U_{adv}, X) dU_{ego} \\ &= \arg \max_{U_{adv}} \mathbb{C}_{adv} \int_{U_{ego}} P(U_{ego}|U_{adv}, X) P(\chi_{nm}|U_{ego}, U_{adv}, X) dU_{ego} \end{aligned} \quad (2)$$

式中， $\mathbb{C}_{adv} := P(U_{adv}|X)$ 。由于自车策略的黑盒性，其行为没有先验性，因此 $P(U_{ego}|U_{adv}, X)$ 是每次尝试 U_{adv} 后确定的，随后才能确定 $P(\chi_{nm}|U_{ego}, U_{adv}, X)$ ，因此可以通过该迭代过程优化对抗行为： $\max_{U_{adv}} \mathbb{C}_{adv} \rightarrow U_{adv} \rightarrow P(U_{ego}|U_{adv}, X) \rightarrow P(\chi_{nm}|U_{ego}, U_{adv}, X) \rightarrow \max_{U_{adv}} P(\chi_{nm}|X)$ 。

因此，场景生成任务可以分解为最大风险的对手车选择、对手车行为控制、自车行为获得、高风险场景合成，从而找到最大的 $P(\chi_{nm}|X)$ 。为了找到足够多的风险场景，需要具体设计具有不同风险特征的损失函数^[18]。基于上述条件关系，本研究从工程求解角度将场景生成过程组织为挑战对手选择、风格化行为控制和风险场景合成 3 个子任务，流程示意如图 1 所示。后文介绍本研究的具体方法细节。

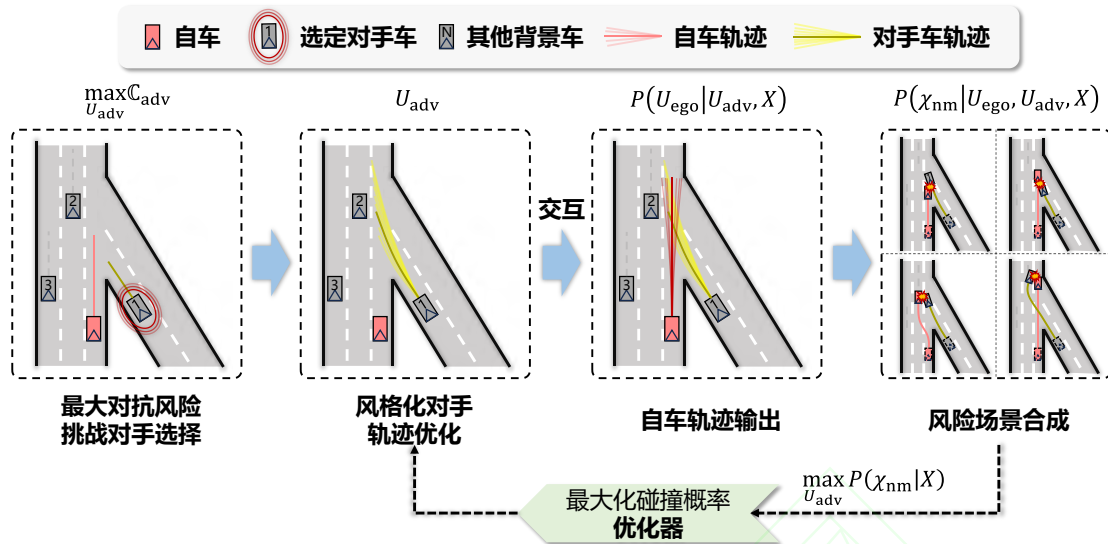


图 1 风险场景演变流程示意图

Fig 1 Schematic Diagram of the Risk Scenario Evolution Process

2 风险场景演变方法

2.1 模型框架概述

图 2 展示了本研究提出的风险场景演化框架。该框架以自然驾驶轨迹数据为输入，通过 3 个核心阶段实现从常规场景向高价值风险场景的演化：① 风险目标锚定：利用 SDL 模型对常规场景内的背景车辆进行危险性评估与分级，精准筛选出最具潜在威胁的个体作为对抗目标车；② 风格化行为演化：引入驾驶员社会行为风格表征，结合 IRL 构建风格编辑损失函数，诱导目标车产生具备特定行为倾向（如激进、保守）的对抗性交互行为；③ 多维场景合成：基于 KO 在物理约束下解算对抗轨迹，并整合自车响应轨迹与环境车辆约束，最终构建具备真实性与挑战性的风险场景库。

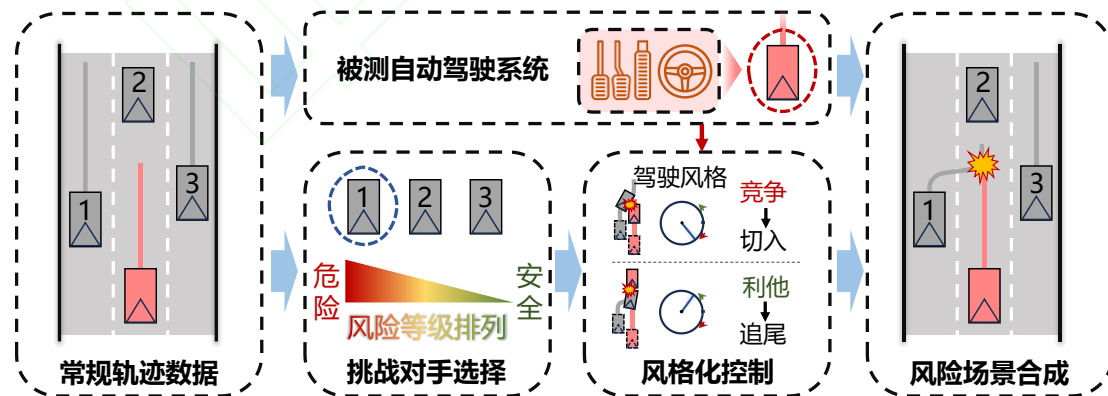


图 2 风险场景演变方法框架图

Fig 2 Framework of the Proposed Risk Scenario Evolution Method

2.2 对手车选择

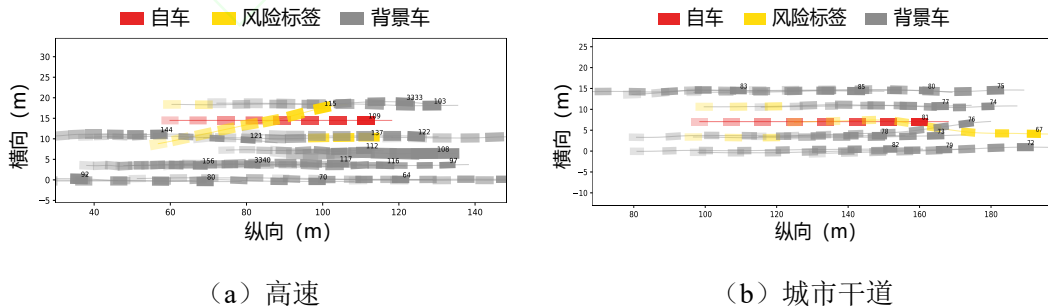
在复杂交通环境下,精准识别具备高风险制造潜力的背景车辆是构建挑战性场景的前提。传统的启发式方法通常仅依据空间邻近性(如距离自车最近)作为筛选准则,但这严重限制了风险场景的多样性。例如,在交叉口工况下,距离较远但处于冲突路径上的背景车往往具备更高的交互价值与威胁潜力。因此,为避免仅依据瞬时空间距离选择对手车,本研究采用融合历史轨迹与 BEV 上下文的学习式风险排序模型,对候选背景车辆的潜在交互风险进行连续量化。(更多实验细节详见附录 A¹)

① 启发式监督信号标注: 由于一定时间内车辆移动的空间局限性,空间关联依然是确定监督信号标注的主要影响因素。假设自车和第 k 个背景车在时间限制 $T = [t_0, t_1]$ 内的历史路点集合分别为 $\mathbf{s}_0 = \{(x_0, y_0, z_0, \varphi_0) \in \mathbb{R}^4 | t \in T\}_{t=t_0}^{t_1}$, $\mathbf{s}_k = \{(x_k, y_k, z_k, \varphi_k) \in \mathbb{R}^4 | t \in T\}_{t=t_0}^{t_1}$, 其中 x, y, z 是某一车辆的空间坐标(自车是第 0 个车辆), φ 是航向角, t 是时间步数。本研究考虑了空间相近和空间交叉 2 个因素, 则监督标签为:

$$Y = \begin{cases} 1, & (\min_{t \in \{1, \dots, T\}} (\|\mathbf{s}_0^t - \mathbf{s}_k^t\|) < L_0) \vee (\mathbf{s}_0 \cap \mathbf{s}_k \neq \emptyset) \\ 0, & \text{其他情况} \end{cases} \quad (3)$$

式中, L_0 为自车的车长。考虑条件之一是两车间距小于 L_0 , 最常出现于两车在相邻车道并排行驶时。另一个条件是 $\mathbf{s}_0$ 与 $\mathbf{s}_N$ 存在空间交叉, 可以通过纵向控制制造时间冲突。由于路点是离散的数据, 因而本研究通过叉积判断“跨立”关系, 再结合共线时的端点重叠情况来确定轨迹交叉情况。注意距离条件只适用于满足约束的帧, 而空间交叉条件适用于整条轨迹。

本研究对于存在潜在高风险的背景车生成了 1 的伪标签, 其他车辆置 0。在各种交通环境中构建充足的风险标签后, 对抗性预测网络可以学习车辆交互与风险的复杂隐式关联。图 3 所示为任意选取的 4 个场景案例, 红色方框为自车, 灰色方框为背景车, 亮黄色方框为置 1 的背景车风险标签对应位置。本研究对 4 种交通场景进行了标注^[19], 累计场景数量 2352 个, 标签数据 700 余万个, 随着时间推演方框颜色由浅及深, 采样步长为 1 秒。具有潜在风险的场景类型丰富, 例如背景车可能追尾自车、背景车突然切入自车车道等, 且对冲突车辆和时机的判定基本符合人类经验认知。相对于文献^[17], 本研究考虑了时序上每一离散位置的潜在交互可能性, 能够进一步评价背景车在整个可能的演进轨迹上与自车的交互程度, 而不是单纯的二元关系。



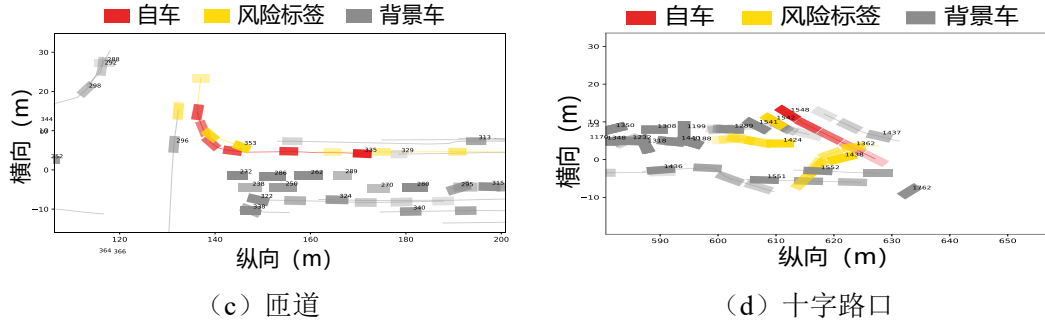


图 3 背景车风险标签启发式标注示例

Fig 3 Examples of Heuristic Annotation for Background-Vehicle Risk Labels

② 对抗性预测网络建模：本研究沿用了文献^[17]的方法。对于网络架构设计，本研究采用图神经网络 VectorNet^[20]来处理道路拓扑与交通参与者关联的矢量化特征；采用卷积神经网络 VGG16^[21]来处理栅格化的交通流鸟瞰图特征（224×224，3 通道），预测头包含一个多层感知机和全连接层。二者通过拼接（Concatenation）进行特征融合，具体技术细节见附录 B。

对于损失函数设计，我们使用了焦点损失来缓解正负样本之间的类别不平衡瓶颈：

$$L_{\text{focal}} = -\alpha_t(1 - p_t)^\beta \log(p_t) \quad (4)$$

式中， α_t 为平衡正负样本的权重，而 β 为调整模型对不同难度级别样本关注程度的权重， p_t 为真实类别的预测概率。

③ 挑战对手选择实验：为了检验对抗性预测网络的泛化性，本研究选取了部分不同于训练数据集的场景进行验证。训练场景采用开源自然数据集，而验证场景全部采用 CARLA 仿真随机生成的数据集，背景车数量、驾驶行为均不同于训练集。在推理过程中，网络给每辆背景车的每一帧赋予风险潜力权重，历史累计风险最高者被选择为挑战对手车。

如图 4 所示为仿真场景下的对抗性预测量化情况。背景车为亮黄色方框，自车为红色方框。亮黄色方框越深表示对抗性风险越高，总风险值被标记在方框旁。背景车的轨迹越深也表示在当前位置相对于自车的风险估计值（0~1）越高。例如，在（c）中，在自车左转过程中，可能会和预驶入车道的右侧车辆发生碰撞；在（d）中，自车在驶入环道内车道时，可能会和正在环道内行驶的车辆发生碰撞。可见，对抗性预测网络所选择的对手车辆和人类驾驶经验较为匹配，而且背景车的累计风险值符合直观判断。

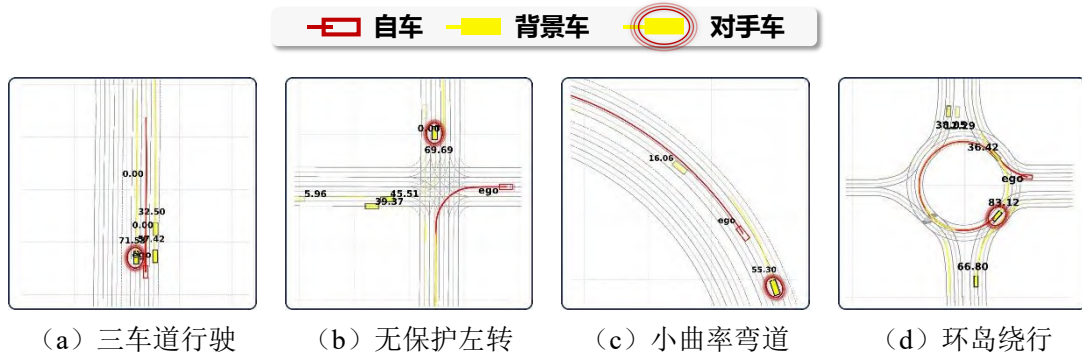


图 4 对抗性预测网络对手车选择推理示例

Fig 4 Examples of Adversary Selection by the Adversarial Prediction Network

2.3 风格化控制

① 风格化评价方法：启发自文献^[22]，本研究利用最大熵 IRL 来学习判断历史轨迹的风格。评估对手车轨迹的风格匹配度得分的流程如图 5 所示。本研究采用 NGSIM^[19]数据来构建数据驱动的风格评价模型。随后采用 K-Means 聚类将原驾驶风格分类为 4 种类型，可以理解为激进-竞争-利他-柔和等 4 种典型的驾驶风格。利用最大熵 IRL 分别针对 4 种数据进行奖励建模，获得 4 种个性化风格模型的权重 θ 。选择其中一种风格权重，基于当前的场景状态特征 f_ζ 便可以确定当前对象轨迹 ζ 的奖励得分 R_ζ 。



图 5 对手车轨迹风格匹配度得分估计 workflow

Fig 5 Workflow for Estimating Style-Matching Score of Adversary Trajectories

其中，最大熵 IRL 通过迭代优化 θ 来获得最能体现目标数据集中驾驶风格的权重，特征 f 继续选取为[速度，纵向加速度，横向加速度，纵向加加速度，前向风险，后向风险，碰撞，交互]，详细计算方式参考^[22]。图 6 所示为最大熵 IRL 的具体工作流程。首先，遵循最大熵 IRL 的一般过程，计算每条演示轨迹的特征参数 f_ζ ，根据当前的权重 θ 计算样本的特征期望值 $\tilde{f}$ ：

$$\tilde{f} = \sum_{i=1}^N \sum_j \frac{\exp(\theta^T f_{\zeta_i^j})}{\sum_j \exp(\theta^T f_{\zeta_i^j})} f_{\zeta_i^j} \quad (5)$$

式中， ζ_i^j 为第 i 条轨迹的第 j 个特征。然后，计算梯度值：

$$\nabla_{\theta} \mathcal{J}(\theta) = \bar{f} - \tilde{f} - 2\phi\theta \quad (6)$$

式中， $\bar{f}$ 为样本轨迹特征期望 $\bar{f} = \sum_{i=1}^N f_{\zeta_i}$ ， ϕ 为正则化参数。利用梯度上升方法，更新后的特征权重为：

$$\theta \leftarrow \theta + \varpi \cdot \nabla_{\theta} \mathcal{J}(\theta) \quad (7)$$

式中， ϖ 为更新步长。通过该迭代优化工作流可以使轨迹的特征期望与样本轨迹的特征期望相匹配，从而可以评价先验轨迹的风格匹配度。

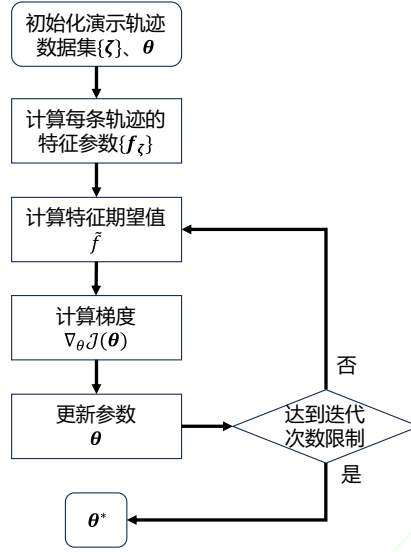


图 6 最大熵逆强化学习工作流程

Fig 6 Workflow of Maximum Entropy Inverse Reinforcement Learning

② 风格评价实验：本研究抽样了 500 个场景进行驾驶风格分类（聚类详细分析见附录 C）。如图 7 所示驾驶风格聚类的可视化二维散点图，绝大部分数据同属于“竞争”风格，对应关系是 0-柔和、1-激进、2-竞争、3-利他。如表 1 所示是通过最大熵 IRL 重建的个性化风格权重参数，横轴是特征 f ，纵轴是 4 种驾驶风格，权重绝对值越大表明越重视该特征。驾驶风格可以由各项参数进行解释，如激进驾驶风格下对速度的追求更加明显，青睐于与其他车辆进行交互，导致忽视风险。如图 8 所示，本研究选取了 1 种多车道行驶场景，利用 4 种不同风格化评价模型分别对相同的手车候选轨迹同时进行了打分，ADV4 为静止车辆，ADV2 被强制向左侧换道。在激进与竞争风格下，ADV2 倾向于在 EGO 到达 ADV4 位置前切入左侧车道；而柔和与利他风格下，ADV4 倾向于在本车道礼让 EGO 通过或者在 EGO 通过后再切换至左侧车道。可见，评价模型和人类直觉基本上是相同的，证明该方法可应用于轨迹的风格评价。

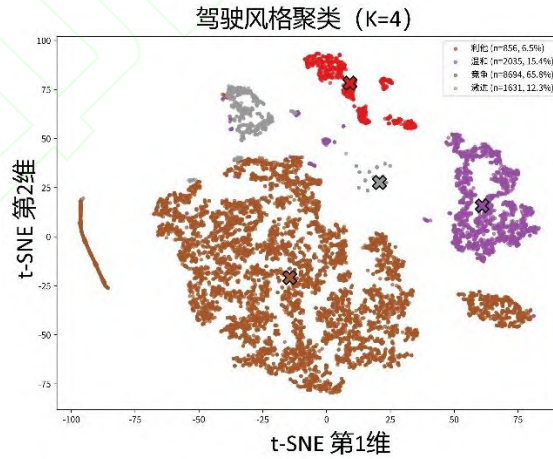


图 7 NGSIM 抽样数据集驾驶风格聚类二维散点（t-SNE 降维）

Fig 7 2D Scatter Plot of Driving-Style Clustering on the Sampled NGSIM Dataset Using t-SNE

表 1 个性化风格权重 θ
Table 1 Personalized Style Weights θ

θ	效率权重	舒适性权重			安全性权重		社会性权重
	速度	纵向加速度	横向加速度	纵向加速度	前向风险	后向风险	碰撞交互
激进 θ_A	9.842	-2.193	-1.085	-1.560	-1.570	-1.839	-9.989
竞争 θ_B	10.155	-3.967	-1.727	-0.233	-3.798	-4.373	-9.989
温和 θ_C	6.979	-3.120	-1.051	-1.385	-3.750	-3.452×10^{-1}	5.352×10^{-3}
利他 θ_D	3.865	-4.181	-1.254	-0.208	0.017	-0.536	-9.989

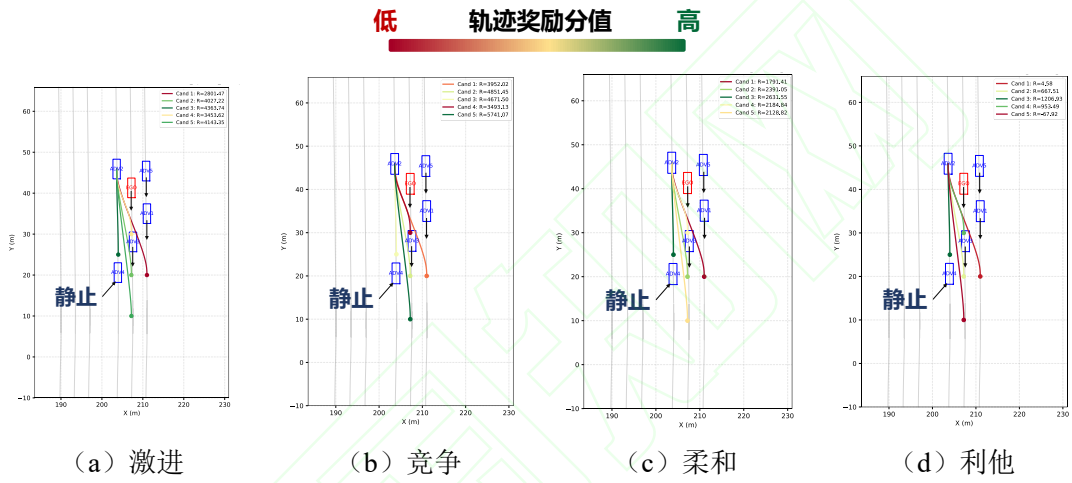


图 8 仿真场景中的风格化得分评价验证实验

Fig 8 Validation of Stylized Score Evaluation in Simulation Scenarios

2.4 风险场景合成

本研究在原始常规场景数据的基础上对挑战对手车进行增量控制，以维持自然真实性。在某个场景中，对手车在约定时间段内 $t \in T = [t_0, t_1]$ 的整个控制序列 $a_i \in \{a_k^t\}_{t=t_0}^{t_1-1}$ 都被扰动。假设车辆均服从运动学自行车模型：

$$\begin{cases} \dot{x} = v \cdot \cos(\varphi) \\ \dot{y} = v \cdot \sin(\varphi) \\ \dot{\varphi} = v \cdot \tan(\varphi)/l \\ \dot{v} = a \end{cases} \quad (8)$$

式中， v 为车辆的速度， l 为轴距。首先基于该模型，从 s_k 中恢复原有数据车辆的动作控制。对手车的控制动作包括纵向加速度和转向角，即 $a_k^t := [a_k^t, \delta_k^t]$ 。对抗轨迹演变过程为：① 应用一系列的边界扰动 $\{\Delta a_k^t\}_{t=t_0}^{t_1-1}$ 以获得新的增量控制动作；② 利用该运动学模型推演新的状态序列。则场景控制可以参数化为 $\vartheta = \{\Delta a_k^t\}_{t=t_0}^{t_1-1} \in \mathbb{R}^{T \times 2}$ ，同时限制扰动幅度在 Θ 之内。

自车对手车进行闭环迭代优化，优化的目标是：

$$\boldsymbol{\vartheta}^* = \arg \min_{\boldsymbol{\vartheta}} J_{\text{col}}^{\text{ego}}(\mathcal{S}) + \lambda \cdot J_{\text{col}}^{\text{adv}}(\mathcal{S}) + \gamma \cdot J_{\text{dev}}^{\text{ego}}(\mathcal{S}) + \xi \cdot J_{\text{sty}}^{\text{adv}}(\mathcal{S}) \quad (9)$$

式中， $J_{\text{col}}^{\text{ego}}(\mathcal{S})$ 、 $J_{\text{col}}^{\text{adv}}(\mathcal{S})$ 、 $J_{\text{dev}}^{\text{ego}}(\mathcal{S})$ 和 $J_{\text{sty}}^{\text{adv}}(\mathcal{S})$ 分别表示鼓励对手车和自车发生碰撞、排斥和其他背景车发生碰撞、排斥驶离道路、以及风格判断项。 λ 、 γ 和 ξ 分别是子目标函数的权重系数。 $\mathcal{S} = \{\mathbf{s}_0, \mathbf{s}_1, \dots, \mathbf{s}_N\}$ 。

具体来说，诱导对手车向自车靠近是生成风险场景的必要条件（从“碰撞概率最大化”到“距离势场代理代价最小化”的等价推导见附录 D），因此 $J_{\text{col}}^{\text{ego}}(\mathcal{S})$ 可以定义为：

$$J_{\text{col}}^{\text{ego}}(\mathcal{S}) = \min \sqrt{\|\mathbf{s}_0 - \mathbf{s}_k\|^2} \quad (10)$$

该式表示取两车最小欧式距离为目标函数值，以鼓励对手车靠近自车。

对于 $J_{\text{col}}^{\text{adv}}(\mathcal{S})$ ，背景车之间避免发生碰撞是维持真实性的关键条件：

$$J_{\text{col}}^{\text{adv}}(\mathcal{S}) = -\min(\min_{k_1, k_2 \in \{1, \dots, N\}} \sqrt{\|\mathbf{s}_{k_1} - \mathbf{s}_{k_2}\|^2}, \Gamma_{\text{col}}^{\text{adv}}) \quad (11)$$

式中，阈值 $\Gamma_{\text{col}}^{\text{adv}}$ 确保只有当一对背景车的距离比安全裕度更小时，软约束才是有效的。

为避免背景车偏离道路，对其进行约束：

$$J_{\text{dev}}^{\text{ego}}(\mathcal{S}) = \sum_{i=0}^N \sum_{t=t_0}^{t_1} g(\mathbf{s}_k, \mathcal{M}) \quad (12)$$

式中， $g(\mathbf{s}_k, \mathcal{M})$ 为对第 i 个车辆在地图 $\mathcal{M}$ 下的高斯势场力，即道路外存在排斥力。

为了将对手车的行为匹配预设的驾驶风格，本研究约定风格奖励相关的目标函数：

$$J_{\text{sty}}^{\text{adv}}(\mathcal{S}) = -\min(R_{\mathbf{s}_k}, \Gamma_{\text{col}}^{\text{adv}}) \quad (13)$$

式中， $R_{\mathbf{s}_k}$ 为对抗时序轨迹和期望的驾驶风格之间的匹配度奖励值，即 $R_{\mathbf{s}_k}$ 越大表明对抗轨迹越接近特定的驾驶风格。该值由本研究的 2.3 节中训练好的风格化匹配度评估模型计算得到。

图 9 所示为本研究风险场景演变方法的完整流程图，使用 Pytorch 的 Adam 优化器来迭代挑战对手车的控制序列。

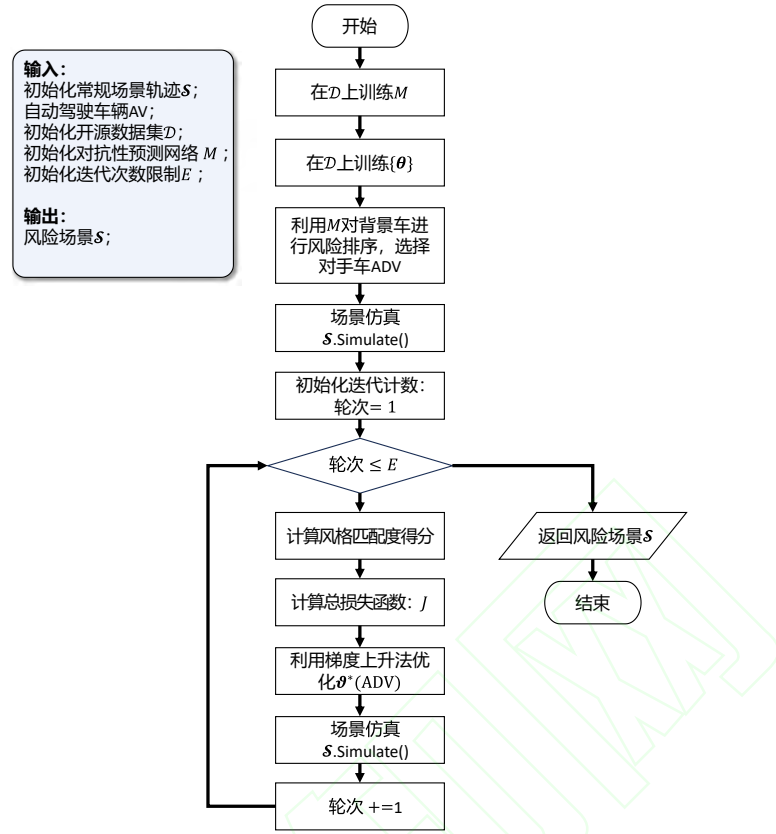


图 9 总流程图

Fig 9 Overall Flowchart

3 实验结果分析

3.1 实验配置

① 交通环境仿真器：本研究采用 CARLA 仿真器作为动力学基础，自动驾驶算法采用 AIM-BEV^[23-24]端到端框架（可迁移至其他 ADS 策略，详见附录 E），常规交通场景数据采用 Carla 自带的 Traffic Manager 随机生成，累计生成了 100 个场景，均保证自车能够完成仿真，仿真时间设置为 20-120 秒（依据场景的初始状态和预期路线完成时间设置）；

② 对抗场景生成基准方法：本研究选择 STRIVE^[25]、AST^[26]、KING^[1]、BBO、BGA 和 BRS 等 6 种方法作为基准。其中 STRIVE 使用条件变分自编码器建模场景；AST 是基于强化学习的压力测试场景生成技术（沿用双延迟深度确定性策略梯度算法）；KING 是基于运动学梯度代理的黑盒优化方法；后三者是 AdvSim^[27]的变体，分别使用了贝叶斯优化、遗传算法和随机搜索技术。

本研究提出的方法记为“BESTO”（Bayesian dEcomposition-based Scenario generaTiOn）表示，针对 4 种风格分别优化出不同的场景生成模型，同时融合 4 种模型生成的场景数据集进行平均场景的评价（标记为“BESTO-Ave”）。

③ 参数配置：如表 2 所示为本研究采取的参数配置，其中 ξ 在不同的场景下选择了不同的值。

表 2 参数配置

Table 2 Parameter Configurations

参数	含义	取值
α_t	对抗性网络损失函数超参数	0.25
β	对抗性网络损失函数超参数	2
λ	风险场景合成超参数	5
γ	风险场景合成超参数	20
ξ	风险场景合成超参数	$\xi \in \{3,5,10\}$
l_r	学习率	1×10^{-3}

3.2 实验指标

3.2.1 场景生成性能测试评价

本研究组织挑战性、丰富性、真实性、可控性和可解性等多方面的性能评价。

① 挑战性：本研究采用正确碰撞率（Right Collision Rate, RCR，即 adv-ego 碰撞率：对于优化型方法，采用优化稳定后的随机抽样迭代场景进行统计；对于学习型方法，随机采样训练完成后的推理生成场景进行统计），碰撞时间（Collision Time, CT，从仿真开始到碰撞的时间），碰撞速度（Collision Speed, CS，碰撞时刻速度）来评价所生成的场景挑战性；

② 丰富性：采用自车的动态时间规整（Dynamic Time Warping, DTW）、最终位移多样性（Final Displacement Diversity, FDD）来评价相同初始条件下生成不同场景的丰富性；

③ 真实性：采用道路偏离率（Off-Road Rate, ORR）、错误碰撞率（Wrong Collision Rate, WCR，包括背景车-ego，背景车-背景车，背景车-adv 3 类碰撞）来评价场景的合理性；

④ 可控性：本研究利用人类评估一致性（评价指标是一致性占比（Consistency Proportion, CP）和组内相关系数（Intraclass Correlation Coefficient, ICC））来检验生成场景是否符合对手车选择和风格化控制的目标；CP 比较生成风格与人类期望之间的一致性占比；ICC 用于评估通过重复测量（包括多名评估者对同一变量进行评分）所获得的定量数据的一致性或可靠性。由于其他基准方法没有风格控制能力，因此没有可控性，不予评价。本研究筛选了 3 名领域专家和工程师对场景风格进行评价。

⑤ 可解性：所有场景中基于规则专家的可求解率占比（Solvability Proportion, SP）。

3.2.2 自动驾驶鲁棒训练性能评价

本研究利用各种方法合成的对抗交通场景，与原始常规场景数据集混合形成再训练数据集，对 AIM-BEV 端到端自动驾驶方法进行鲁棒训练微调。随后，在常规-对抗混合场景中进行测试，检验 AIM-BEV 的路线完成度（Route Completion, RC）、合规率（Infraction Score, IS）和驾驶得分（Driving Score，路线完成度×合规率），以及碰撞率（Collision Rate, CR）和任务路线完成时间（Route Completion Time, RCT），以评价其不同场景数据集下的对抗训练效果。

3.3 实验结果及分析

3.3.1 实验结果分析

首先对生成的典型风险场景进行定性分析。如图 10 所示，本文方法在“前车切入”“匝道汇入”“路口穿行”和“调头”等 4 类强交互场景中生成了不同风格的风险交互轨迹。图中自车、对手车及背景车分别以红色、黄色和灰色表示，红色方框标记潜在冲突区域。对手车由 2.2 节的对抗性预测网络确定，并在原始常规场景轨迹基础上进行增量控制优化；其他背景车保持历史动作回放，以维持场景环境的真实性。

可以看出，不同风格参数下的对手车呈现出明显差异化的交互行为。以“前车切入”为

例，较激进风格（ θ_A ）下的对手车更倾向于提前抢行切入并增强与自车的交互，而较温和风格下的对手车切入更平缓、冲突强度相对较低。类似现象也出现在匝道汇入、路口穿行和调头场景中：风格越激进，对手车越倾向于主动压缩自车通行空间或增加交互频次。这表明本文风格化控制模型能够将不同驾驶风格映射为差异化的风险行为偏置，从而提升生成场景的丰富性和可控性。这表明 2.3 节风格化控制模型捕捉到了现实驾驶情况下不同的驾驶风格形式化描述方法并将其推广到了场景风格生成的偏置诱导上，这有助于更加丰富的场景数据生成，提升训练的鲁棒性，优化测试的全面性，在 3.5 节的实验也证明了这一点。

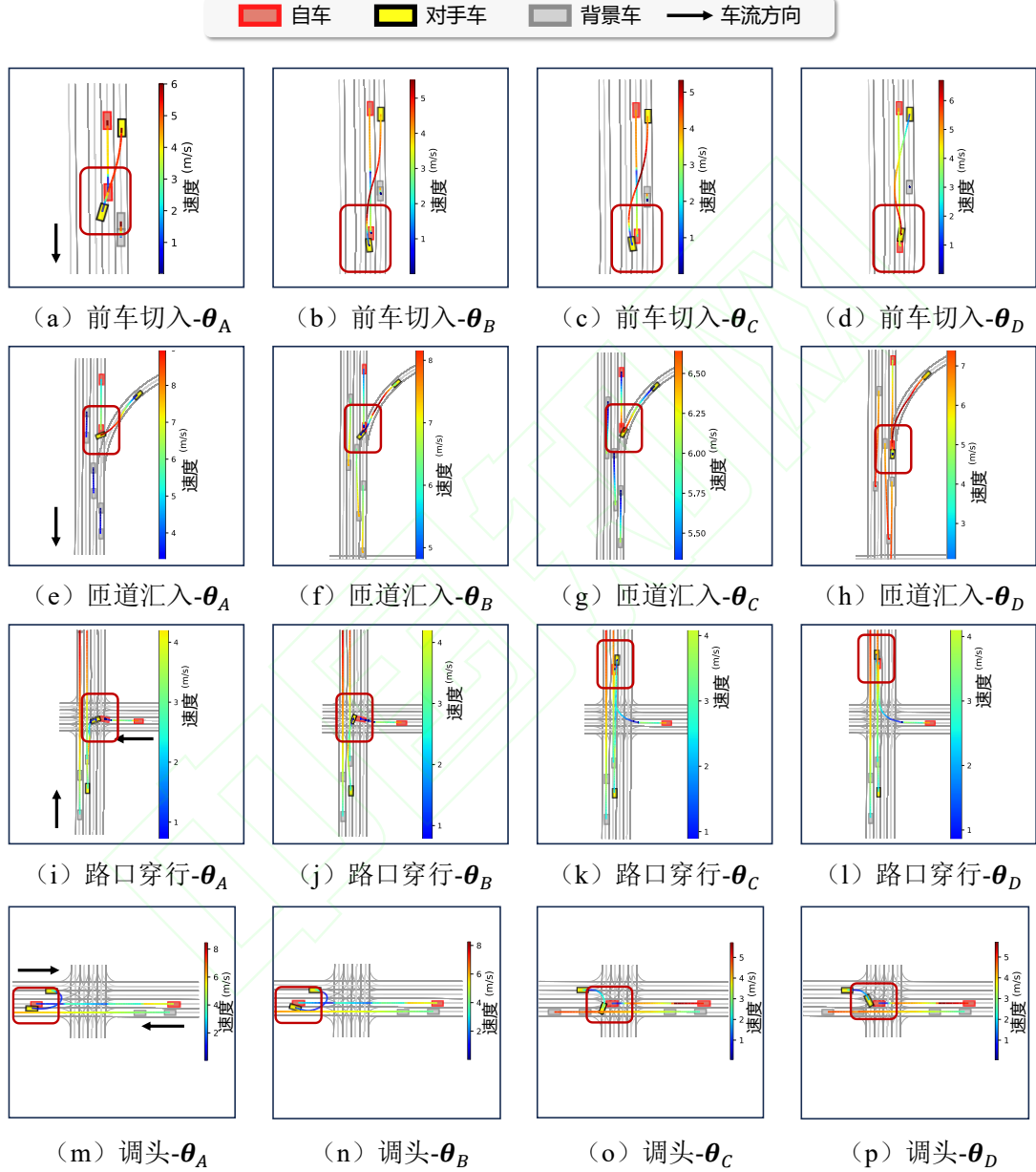


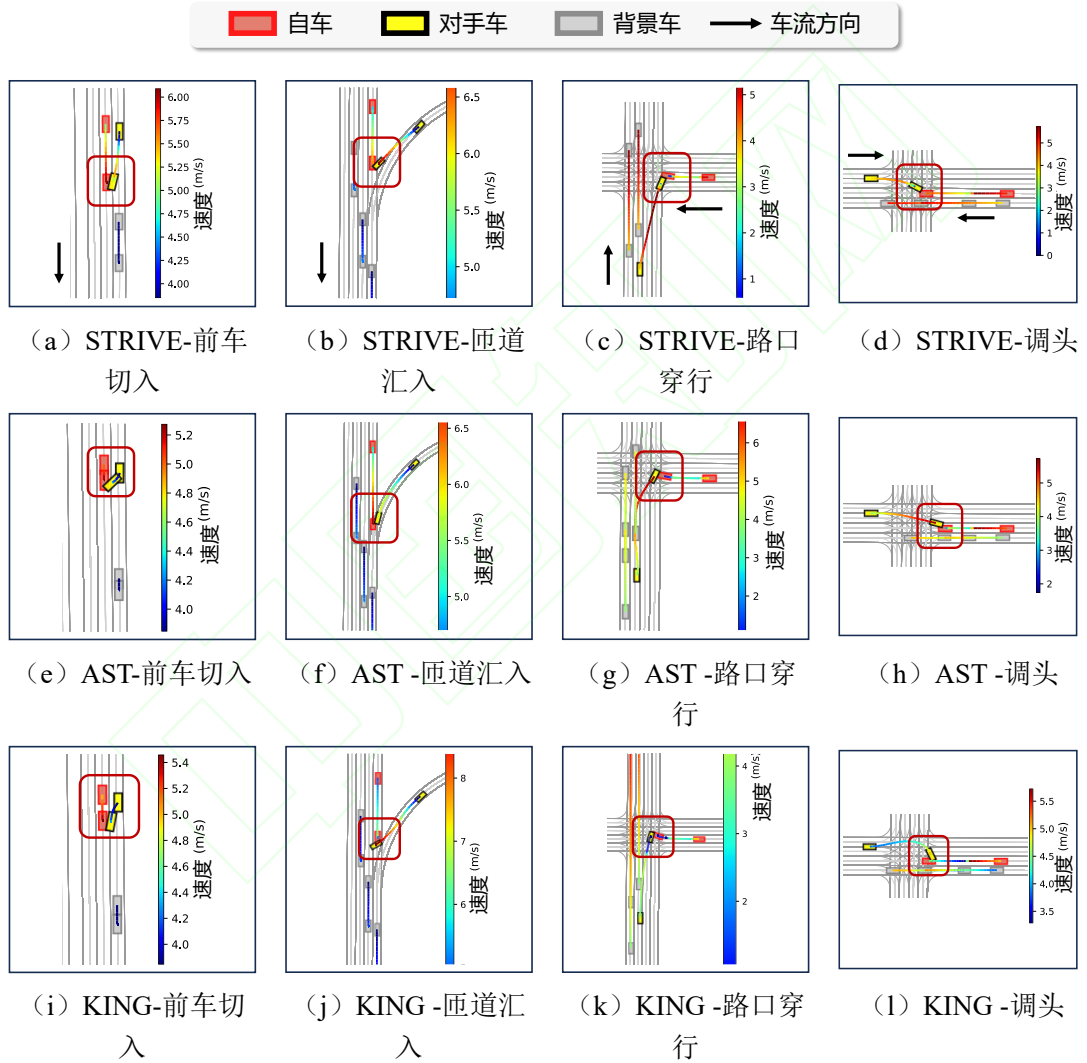
图 10 本研究方法生成的场景 BEV

Fig 10 BEVs of Scenarios Generated by the Proposed Method

为了验证本文方法的有效性，本节在完全一致的场景初始条件下，将其与 3.1 节所述的基准方法进行了对比实验。考虑到基准方法原有的对抗车辆筛选机制较为简单，在相同常规场景下可能产生目标选择的不确定性，为排除该因素干扰、确保对比的公平性，所有基准方

法均统一采用本文提出的对抗车辆筛选策略进行验证。关于该筛选策略的具体影响分析，详见 3.3.3 节的消融实验部分。

如图 11 所示，各基准方法生成的对抗车辆虽基本满足道路边界约束，但其行为模式普遍较为刻板，常表现为从初始位置直接逼近自车以制造碰撞。这主要是由于现有方法多侧重对抗性冲突诱导，缺乏驾驶风格与行为合理性约束，导致生成轨迹相对于原始场景偏离较大，且对抗行为单一、可控性不足。具体而言，STRIVE 虽引入对抗行为变化的软约束，但轨迹扰动仍较明显；AdvSim 系列方法受黑盒优化机制影响，生成结果易趋于同质化；KING 缺少风格化约束项 $J_{sty}^{adv}(\mathcal{S})$ ，难以体现拟人化交互差异；AST 虽具有一定样本多样性，但搜索过程缺乏理性行为约束，导致场景真实性和演化方向控制能力不足。相比之下，本文方法通过风格化约束与增量运动优化共同作用，能够生成更具策略性和行为差异性的风险交互场景。



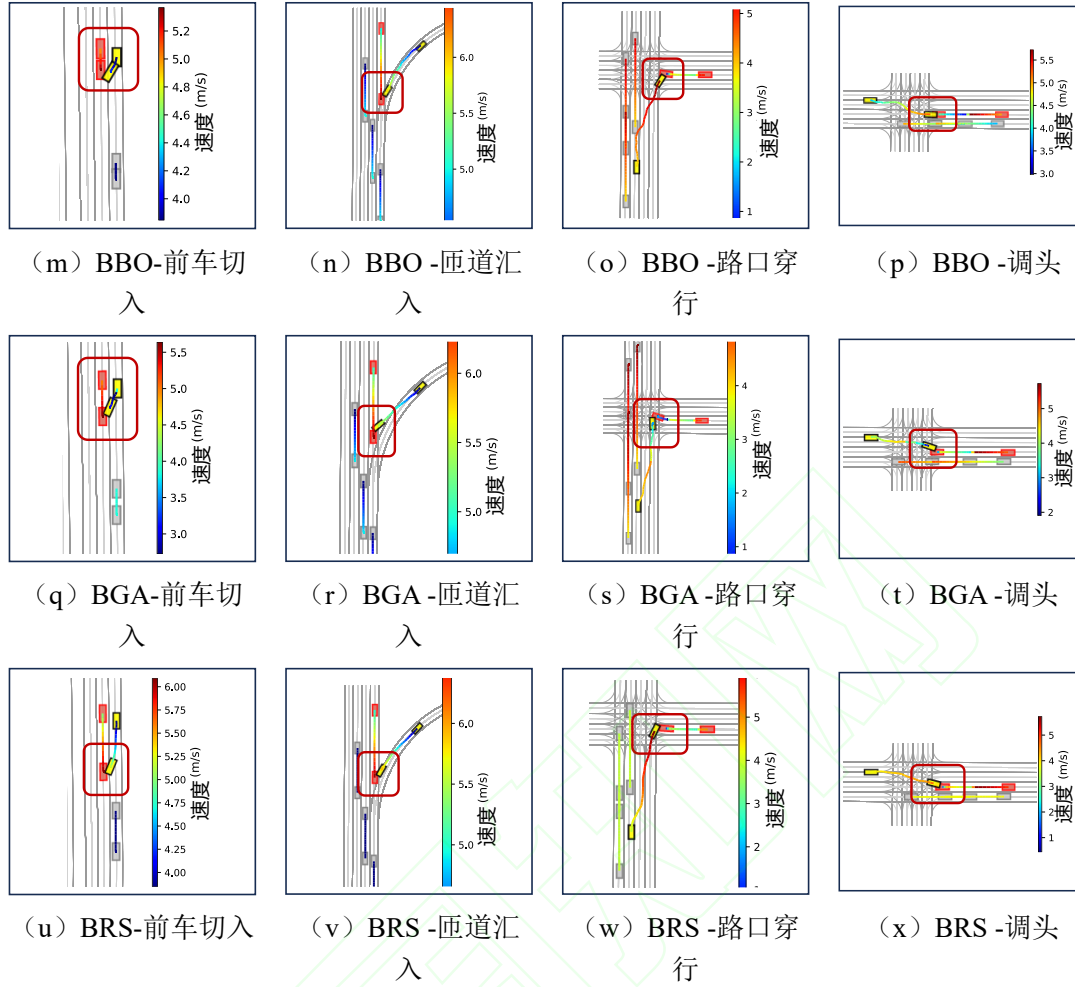


图 11 不同场景生成方法生成的场景 BEV 比较

Fig 11 Comparison of BEVs Generated by Different Scenario Generation Methods

3.3.2 数据比较分析

为了定量评估各方法在 3.2.1 节所述关键性能指标上的差异,本研究对每种方法在 20 个常规场景下进行了风险场景生成。待迭代或训练稳定后,选取后 20 组数据进行指标统计,结果如表 3 所示。表中列出了本研究所提出的不同风格模型 θ 及平均数据集(BESTO-Ave,由 4 种风格模型各取 25%构成)的分析数据。每种方法均采用 3 个随机初始化种子,置信度区间标注于数据下方。各列指标的最优值以粗体标识,次优值以灰色底框标识。

总体而言,本文方法在挑战性、丰富性、真实性、可控性和可解性之间取得了较好的综合平衡。挑战性方面, θ_C 和 θ_D 的碰撞成功率较高,与 KING、BGA 等优化方法接近;虽然本文方法的碰撞速度等指标略低于部分基准方法,但这是由于其在诱导风险的同时兼顾了场景真实性与可解性,避免生成过强但低价值的“必撞”场景。丰富性方面,受不同风格模型的偏置作用影响,BESTO-Ave 取得较高的 DTW 和 FDD,显著优于多数基准方法;真实性方面,本文方法在 ORR 和 WCR 上保持较低水平,说明风格化评价与运动约束有助于抑制不合理轨迹。可控性方面,本文方法的 CP 和 ICC 分别达到 0.87 和 0.80,表明生成风格与人类认知具有较高一致性;可解性方面, θ_C 和 θ_D 的 SP 均超过 80%,说明适度风格偏置能够生成更具训练价值的风险场景。此外,附录 F 中的真实数据驱动初始场景实验进一步表明,

在保持优化目标、优化预算和自车策略一致的条件下,本文方法仍能有效提升场景冲突风险,说明其并非仅依赖 CARLA 随机配置。综上,不同风格模型的偏置与混合机制使本文方法能够在综合性能稳定的基础上,实现对特定场景生成目标的可控调节。

表 3 不同方法关键性能的比较评价

方法	挑战性数值			丰富性数值		真实性数值		可控性数值		可解性数值
	RCR↑	CT	CS	DTW↑	FDD↑	ORR↓	WCR↓	ICC↑	CP↑	SP↑
	(%)	(s)	(m/s)	($\times 10^5$)	($\times 10^2$)	(%)	(%)	(-)	(-)	(%)
STRIVE	72.69	30.67	4.92	0.20	0.15	17.65	7.66	\	\	73.58
	± 1.56	± 1.86	± 0.25	± 0.00	± 0.00	± 0.51	± 0.17			± 2.07
AST	78.64	35.40	5.31	1.76	1.59	9.85	10.15	\	\	71.86
	± 0.89	± 1.07	± 0.29	± 0.01	± 0.00	± 0.02	± 0.59			± 2.08
KING	80.48	32.07	4.98	0.37	0.25	5.64	9.54	\	\	74.82
	± 2.33	± 2.00	± 0.25	± 0.00	± 0.00	± 0.26	± 0.71			± 3.10
BBO	79.72	32.68	4.80	0.16	0.11	5.54	7.47	\	\	74.55
	± 1.09	± 1.74	± 0.21	± 0.00	± 0.00	± 0.01	± 0.66			± 2.47
BGA	80.39	31.73	3.86	0.12	0.11	0.00	3.27	\	\	65.98
	± 0.97	± 1.99	± 0.16	± 0.00	± 0.00	± 0.00	± 0.13			± 2.09
BRS	4.00	34.55	5.04	0.46	0.37	93.47	2.53	\	\	77.07
	± 0.02	± 1.64	± 0.25	± 0.00	± 0.00	± 2.61	± 0.02			± 0.97
BESTO-Ave	80.24	43.95	4.69	1.73	1.38	4.32	10.15	0.87	0.80	80.89
	± 1.71	± 2.23	± 0.13	± 0.01	± 0.01	± 0.01	± 0.05	± 0.01	± 0.02	± 1.97
BESTO- θ_A	80.17	35.49	4.78	0.36	0.28	7.05	6.78	\	\	72.83
	± 2.91	± 1.69	± 0.17	± 0.00	± 0.00	± 0.02	± 0.00			± 1.61
BESTO- θ_B	80.33	37.80	4.14	0.35	0.25	3.20	14.47	\	\	78.64
	± 0.68	± 2.06	± 0.13	± 0.00	± 0.00	± 0.02	± 0.03			± 2.10
BESTO- θ_C	82.10	41.43	4.24	0.39	0.27	3.81	9.09	\	\	82.77
	± 3.02	± 2.39	± 0.19	± 0.00	± 0.00	± 0.01	± 0.02			± 2.69
BESTO- θ_D	80.52	49.52	5.16	0.38	0.22	3.18	10.30	\	\	82.64
	± 1.25	± 2.82	± 0.39	± 0.00	± 0.00	± 0.01	± 0.02			± 1.91

3.3.3 消融研究

为区分各模块的功能贡献,本文从风险目标锚定、风格行为引导和控制参数化 3 个维度开展消融实验。实验针对 2.2 节、2.3 节及控制模型,分别采用“随机对抗车辆选择”、“ $J_{\text{sty}}^{\text{adv}}(\mathcal{S}) = 0$ (即退化为增量控制式 KING 方法)”及“绝对控制模型”作为消融对照组。如表 4 所示,表中列出了消融后各指标相对于 BESTO-Ave 的百分比变化,其中降幅显著者加粗标识、增幅显著者以灰色底框标识。

实验分析表明: ① 采用“随机对抗车辆选择”时,碰撞时间(CT)明显缩短。这是由于随机策略忽略了对未来交互强度的预测,选定的车辆较本文预测网络更趋近于主车,同时也导致丰富性指标(DTW、FDD)分别下降了 14.5%和 8.7%。② 当设置 $J_{\text{sty}}^{\text{adv}}(\mathcal{S}) = 0$ 时,由

于缺失风格化偏置的引导,生成的冲突行为过于直接且缺乏策略性,导致丰富性骤降(DTW、FDD 降幅分别达 79.2%和 81.2%);由于消解了风格化项与真实性项之间的梯度冲突,ORR 呈现增长态势。③ 对于“绝对控制模式”,虽然丰富性与真实性指标因控制自由度无约束而大幅上升,但其代价是收敛时间显著增加(见图 11)。此外,过高的自由度易使生成场景严重偏离原始分布,导致行为特征与预期不符。消融结果表明,风格化控制是行为丰富性与可控性的主要贡献来源;对手车选择主要改善优化对象的交互价值;增量控制则用于约束场景偏离并提高求解效率。

表 4 消融研究数据结果
Table 4 Quantitative Results of the Ablation Study

方法	挑战性数值			丰富性数值		真实性数值		可控性数值		可解性数值
	RCR↑ (%)	CT (s)	CS (m/s)	DTW↑ ($\times 10^5$)	FDD↑ ($\times 10^2$)	ORR↓ (%)	WCR↓ (%)	ICC↑ (-)	CP↑ (-)	SP↑ (%)
BESTO- 随机对手 选择	79.64 -0.8%	37.05 -15.7%	4.77 +1.7%	1.48 -14.5%	1.26 -8.7%	4.10 -5.1%	16.26 +60.2%	0.88 +1.2%	0.80 +0	76.95 -4.9%
BESTO- $J_{\text{sty}}^{\text{adv}}(\mathcal{S}) = 0$ / KING	81.29 +1.3%	30.57 -30.4%	5.27 +12.4%	0.36 -79.2%	0.26 -81.2%	8.88 +105.6%	9.83 -3.2%	\	\	73.34 -9.3%
BESTO- 绝对控制 模式	78.96 -1.6%	43.59 -0.8%	4.58 -2.4%	1.93 +11.6%	1.51 +9.4%	5.46 +26.4%	15.58 +53.5%	0.87 +0	0.80 +0	79.75 -1.4%

图 12 展示了不同方法在风险场景生成过程中的收敛时间(基于损失函数或奖励函数的收敛状态判定)。其中,蓝色柱条代表基准方法,红色柱条代表本研究方法及其消融变体。总体而言,STRIVE、KING 及 BBO 等优化方法因其损失函数构造较为精简,计算开销最低;相比之下,本研究方法由于引入了多目标约束,且需在复杂博弈中平衡梯度冲突,故收敛时间相对较长。在基准方法中,BGA 由于遗传算法的种群演化特性,需执行更大规模的仿真采样,导致计算成本显著增加;而 AST 受强化学习“试错-交互”机制的影响,其收敛周期最长。此外,BRS 作为无目标的随机搜索方法,不具备常规意义上的收敛特征。在消融对比中,绝对控制模式(BESTO-Absolute Control)的计算耗时约为增量控制模式的 1.8 倍,这主要是由于绝对控制下的行动空间自由度更高,导致模型达到稳态所需的搜索时间显著延长。

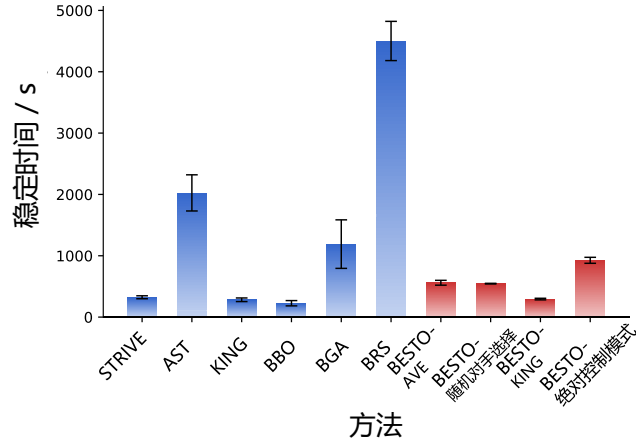
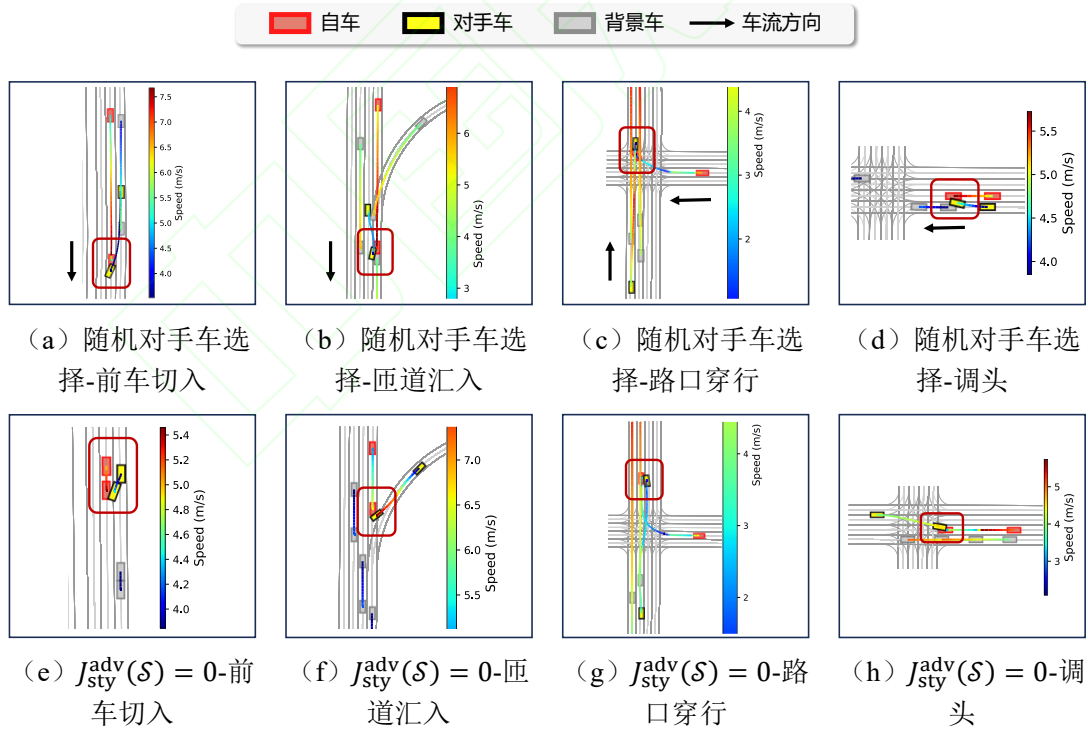


图 12 不同方法下的收敛时间

Fig 12 Convergence Time of Different Methods

图 13 进一步展示了消融 3 个核心模块后，本文方法生成的场景鸟瞰图（BEV）。在“随机对抗车辆选择”模式下，对抗行为呈现出显著的单调性与直接性，例如在“前车切入”、“匝道汇入”和“调头”等场景中，均表现为从相邻车道直接切入的同质化模式。当设置 $J_{\text{sty}}^{\text{adv}}(\mathcal{S}) = 0$ 时，本文方法发生退化，其生成的攻击模式与 KING 方法高度相似，缺乏驾驶风格的差异化特征。在“绝对控制”模式下，虽然对抗模式与全功能方法相比无明显变化，但考虑到其计算开销激增（见图 12），本文最终采用增量控制模式，这不仅提升了生成效率，更能确保生成的风险场景较原始常规场景在运动学轨迹上保持更高的保真度。



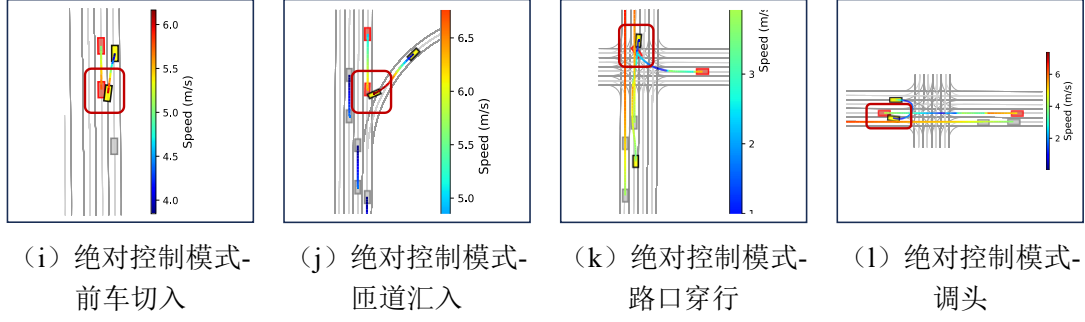


图 13 消融研究 BEV

Fig 13 BEV Visualization of the Ablation Study

3.4 可扩展性研究

本研究可以扩展到多车合作对抗，发现一些单车对抗模式下难以发现的冲突形式。一般情况下，产生对抗行为的车辆在自车周围。假设需要有 $m + 1$ 个背景车辆产生合作对抗行为，则损失项 $J_{col}^{ego}(\mathcal{S})$ 更改为：

$$J_{col}^{ego}(\mathcal{S}) = \sum_{n=k}^{n=k+m} \min(\sqrt{\|s_0 - s_n\|^2}) \quad (14)$$

如图 14 所示，除原始对抗车辆（黄色标识）外，本研究进一步选取距离主车最近的一辆背景车作为协同对象，即 $m = 1$ 。例如，在“前车切入”场景中，主车左侧的对抗车辆与协同背景车同时向主车车道实施变道，诱发了多车连续追尾碰撞。上述工况均体现了双背景车协同诱发的复杂对抗模式，此类高度耦合的博弈行为在单车对抗控制框架下难以表征（对比图 10）。

多车协同对抗场景具有重要的 ADS 测评价值，但其优化变量和交互约束随参与车辆数量显著增加，导致场景求解和下游策略微调均更加困难。因此，本文仅将双车协同对抗作为方法可扩展性的初步验证，而非主要性能贡献。实验结果显示，在双车协同对抗数据上微调后，ADS 碰撞率有所下降，但路径完成时间明显增加，说明高挑战场景可能诱导策略产生过度保守的避障行为，进而影响任务完成效率。该结果表明，多车协同对抗场景虽能增强测试强度，但其稳定优化和训练收益仍存在不确定性。本文已在附录 G 中补充复杂度、优化稳定性与失败案例分析，并将更大规模多车协同优化作为后续研究方向。（更多实验细节详见附录 G）

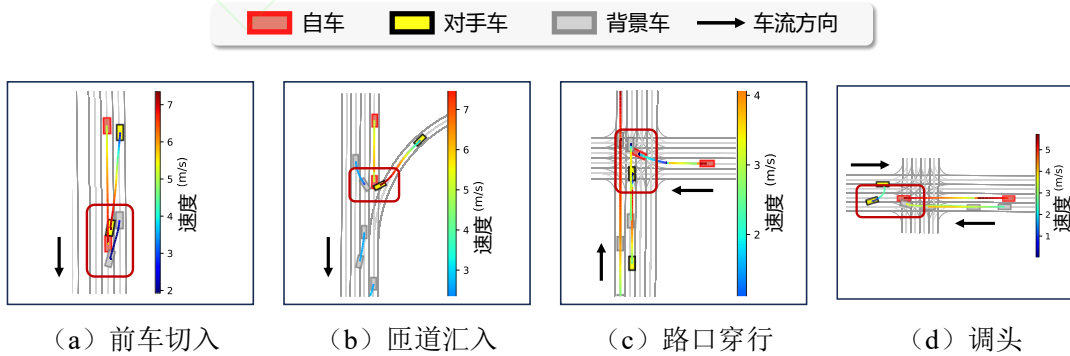


图 14 多对手车对抗场景优化 BEV

Fig 14 BEV of Multi-Adversary Scenario Optimization

3.5 自动驾驶重训练实验与分析

为进一步验证生成场景对自动驾驶模型训练的增益效果，本研究构建了如图 15 所示的自动驾驶增强训练技术框架。首先，提取本文方法及各基准方法生成的关键风险场景，并利用专家控制器（BEVDataAgent）在风险工况下进行操控，以获取保证安全任务目标的专家示教数据集（涵盖 BEV 图像序列、对应的未来轨迹路点及交通风险标识）。随后，将常规场景集与对抗场景集按预设比例加权混合，构建面向重训练的增强数据集。在模型训练阶段，设计基于 L1/L2 范数的路点监督学习损失函数，利用 Aim-BEV 模型执行端到端的监督学习，旨在输出具备鲁棒性的预测轨迹。最后，采用 PID 控制器实现对预测路点的精确跟踪控制。

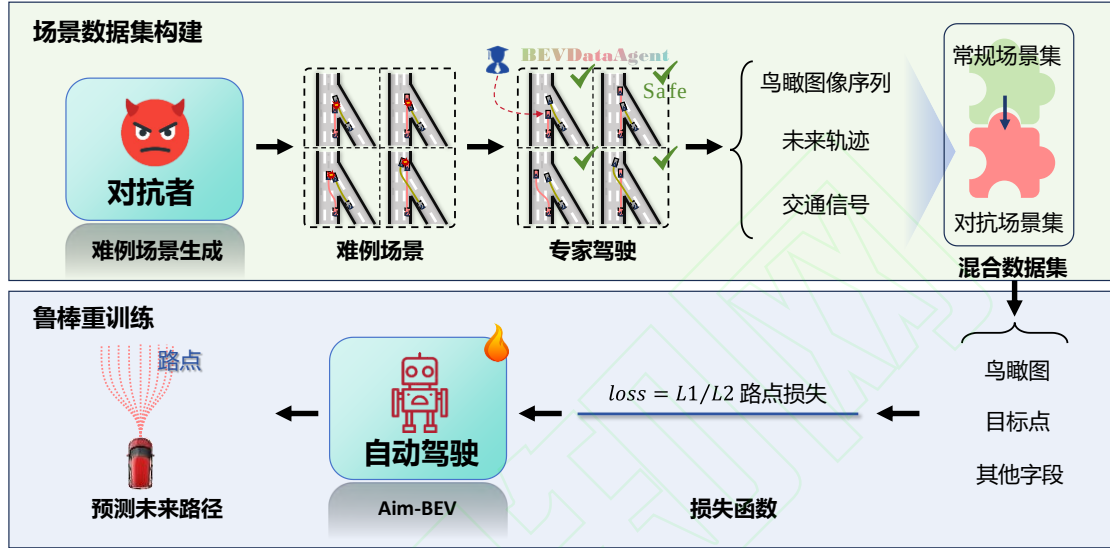


图 15 自动驾驶重训练的技术框架

Fig 15 Technical Framework for Autonomous Driving Retraining

重训练数据集由 50% 的 CARLA 常规场景与 50% 的 CARLA 对抗场景混合构成，训练阶段不包含 Argoverse2 场景。测试集由 CARLA 随机道路拓扑场景和基于 Argoverse2 真实交通数据重构的未见道路拓扑场景构成，二者各占 50%。训练集与测试集在道路拓扑、初始交通状态、场景实例及对抗轨迹层面均保持隔离。测试场景进一步覆盖 RAW、STRIVE、AST、BGA 及 BESTO-Ave 5 类生成方式，各占 20%，用于综合检验模型对未见道路拓扑和不同风险生成模式的适应能力。表 5 展示了 AIM-BEV 模型在不同重训练配置下的微调性能。本文重点分析了 4 类驾驶风格的不同混合比例对重训练效果的影响（见表 5 倒数第 2 至 6 行）。实验结果显示，基准方法的驾驶得分（DS）普遍偏低，尤其是 RAW 与 BRS 方法由于缺乏高质量对抗场景，导致碰撞率（CR）居高不下。值得注意的是，AST 取得了较高的 DS，其生成场景具有较高丰富性，这可能有助于提升 AIM-BEV 对不同交互模式的适应能力。（详细实验配置见附录 H）

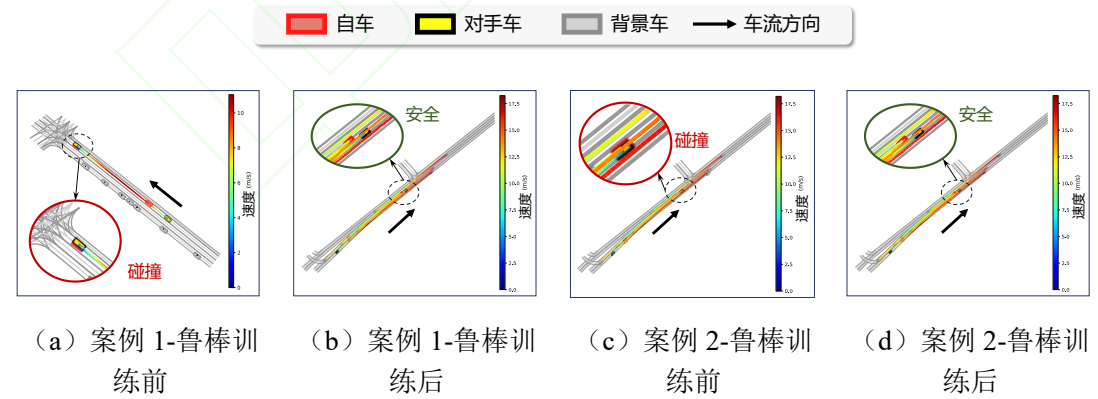
在本文方法组中，当风格参数 θ_c 权重为 0.4 时，DS 达到峰值 65.76，说明适当提高该类风格样本的比例有助于改善模型在混合分布测试集上的综合表现。此外，相较于 RAW 配置，该配置下的 CR 由 0.43 下降至 0.18，但 RCT 由 57.92 s 增加至 68.34 s，反映出模型在提高安全性的同时表现出一定的保守驾驶倾向。总体而言，本文方法在安全性与通行效率的综合表现上与 AST 相当，但本方法通过灵活调节风格比例，能够实现对训练倾向的精准引导，展现出更强的可控性。此外，在“BESTO-多车合作对抗”场景下，DS 并不理想且 RCT 显

著增加，原因在于多车协同极大提升了环境复杂度与求解难度，导致模型策略过于保守，难以在规定时间内完成驾驶任务。

表 5 AIM-BEV 自动驾驶在不同的重训练数据集下微调的测试结果
Table 5 Test Results of AIM-BEV Fine-Tuned on Different Retraining Datasets

数据集	RC↑ (-)	IS↑ (-)	DS↑ (0-100)	CR↓ (-)	RCT↓ (s)
RAW	0.25±0.01	0.32±0.00	8.72±0.12	0.43±0.06	57.92 ±3.60
STRIVE	0.52±0.02	0.63±0.00	33.76±0.14	0.19±0.04	66.87±3.33
AST	0.79±0.03	0.78±0.01	62.83±1.80	0.13 ±0.02	58.60±3.65
KING	0.65±0.01	0.75±0.00	47.53±0.93	0.22±0.03	66.24±2.69
BBO	0.60±0.01	0.76±0.02	46.60±1.36	0.25±0.00	66.21±3.97
BGA	0.64±0.00	0.72±0.03	46.48±1.05	0.21±0.00	64.77±3.20
BRS	0.33±0.00	0.37±0.00	12.11±0.89	0.41±0.03	57.93±3.68
BESTO-Ave 0.{25×4}	0.78±0.02	0.81±0.04	61.81±0.90	0.16±0.04	63.75±1.64
BESTO-0.{4,3,2,1}	0.73±0.02	0.83 ±0.01	59.95±2.28	0.14±0.02	63.27±0.49
BESTO-0.{3,4,2,1}	0.78±0.03	0.80±0.02	61.42±1.44	0.17±0.03	65.98±3.80
BESTO-0.{1,2,4,3}	0.83 ±0.02	0.79±0.01	65.76 ±1.31	0.18±0.02	68.34±2.76
BESTO-0.{1,2,3,4}	0.79±0.04	0.80±0.01	62.08±0.90	0.17±0.02	64.06±3.87
BESTO-多车合作对抗	0.59±0.03	0.80±0.02	48.83±0.96	0.19±0.01	72.95±1.15

图 16 对比了 AIM-BEV 模型在执行重训练前后的 Argoverse2 的部分 BEV 视图（展示场景与附录 F 相对应）。实验观察显示，重训练前的自动驾驶模型在面对对抗车辆攻击时，普遍表现为响应缺失或紧急避障动作滞后，从而引发碰撞。经过基于风险场景的增强训练后，模型具备了更敏锐的风险感知与决策能力，能够通过采取紧急制动或小幅度避让转向来有效缓解冲突。



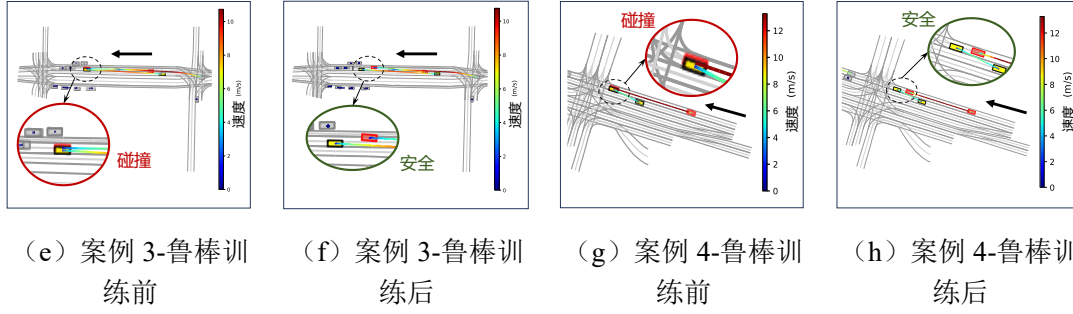


图 16 Argoverse2 场景下的 AIM-BEV 重训练微调结果 BEVs
Fig 16 BEVs of AIM-BEV Before and After Retraining in the Argoverse2 Dataset

4 结语

为实现风险场景挑战性、丰富性、真实性、可控性与可解性的综合平衡，本研究提出一种面向异质驾驶风格的多任务协同风险场景演化方法，并受贝叶斯概率分解思想启发，将生成过程组织为对抗车辆筛选、风格化行为建模与风险场景合成 3 个阶段。首先，利用有监督深度学习模型精准识别具备潜在碰撞风险的对抗车辆；其次，基于最大熵逆强化学习评估车辆轨迹的驾驶风格特征；最后，通过构建主车与对抗车辆的交互动力学模型，利用运动学优化器迭代生成兼顾行为挑战度与物理真实性的仿真场景。实验结果表明，本文方法在场景挑战性、丰富性、真实性、可控性与可解性之间取得了较好的综合平衡，并具有向多车协同对抗场景扩展的潜力。进一步利用生成数据微调自动驾驶策略后，本方法在包含 CARLA 与 Argoverse2 未见道路拓扑的混合测试集上将碰撞率由原数据集的 0.43 降低至 0.16，并将驾驶得分由 8.72 提高至 61.81，表明所生成场景能够改善自动驾驶策略对复杂交互风险的适应能力。

本研究为自动驾驶安全测评与强化训练提供了关键的高价值数据支撑，有效缓解了自然驾驶数据中边缘场景稀缺的问题。尽管本文在多项性能指标上取得了较好的平衡，但仍存在优化空间：一方面，受限于原始数据集的覆盖度，未来可结合生成式模型进一步扩充初始场景的多样性；另一方面，本研究目前侧重于机动车行为优化，未来将进一步引入行人、非机动车等异构交通参与者，探索复杂混合交通环境下的行为优化机制。

参考文献：

References:

- [1] HANSELMANN N, RENZ K, CHITTA K, et al. KING: Generating safety-critical driving scenarios for robust imitation via kinematics gradients[C]//AVIDAN S, BROSTOW G, Cissé M, et al., eds. Computer Vision—ECCV 2022: 17th European Conference, Tel Aviv, Israel, October 23–27, 2022, Proceedings, Part XXXIX. Cham: Springer Nature Switzerland, 2022: 335-352.
- [2] HU L, LU T, LI G, et al. Automatic generation of intelligent vehicle testing scenarios at intersections based on natural driving datasets[J]. IEEE Transactions on Intelligent Vehicles, 2024, 9(9): 5448-5460.
- [3] 邓伟文,李江坤,任秉韬,等. 面向自动驾驶的仿真场景自动生成方法综述[J]. 中国公路学报,2022,35(1):316-333.
DENG Wei-wen, LI Jiang-kun, REN Bing-tao, et al. A survey on automatic simulation scenario generation methods for autonomous driving[J]. China Journal of Highway and Transport, 2022, 35(1): 316-333.
- [4] 余荣杰, 田野, 孙剑. 高等级自动驾驶汽车虚拟测试: 研究进展与前沿[J]. 中国公路学报, 2020, 33(11): 125-138.
YU Rong-jie, TIAN Ye, SUN Jian. Highly automated vehicle virtual testing: A review of recent developments and research frontiers[J]. China Journal of Highway and Transport, 2020, 33(11): 125-138.
- [5] PENG M X, XIE Y T, GUO X S, et al. LD-scene:LLM-guided diffusion for controllable generation of adversarial safety-critical driving scenarios[J]. Transportation Research Part C: Emerging Technologies, 2026, 188: 105694.
- [6] WEI Y X, WANG Z, LU Y F, et al. Editable scene simulation for autonomous driving via collaborative LLM-agents[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ: IEEE, 2024: 15077-15087.
- [7] CAI X, BAI X S, CUI Z Y, et al. Text2Scenario:text-driven scenario generation for autonomous driving test [J]. Automotive Innovation, 2026, 9(1): 102-127.
- [8] 李鹏辉,董倩茹,袁赫男,等.面向自动驾驶仿真测试的高覆盖切入场景库生成方法[J]. 中国公路学报, 2024, 37(7): 237-249.
LI Peng-hui,DONG Qian-ru,YUAN He-nan,et al.High-coverage cut-in scenario library generation for automated driving simulation testing [J]. China Journal of Highway and Transport, 2024, 37(7): 237-249.

- [9] 朱宇, 赵祥模, 徐志刚, 等. 基于蒙特卡洛模拟的无人车高速公路变道虚拟测试场景自动生成算法[J]. 中国公路学报, 2022, 35(3): 89-100.
- ZHU Yu, ZHAO Xiang-mo, XU Zhi-gang, et al. Automatic generation algorithm of lane-change virtual test scenario on highways for automated vehicles using Monte Carlo simulation [J]. China Journal of Highway and Transport, 2022, 35(3): 89-100.
- [10] 谈东奎, 朱波, 胡旭东. 基于近邻目标区域表征与 CVAE 的智能汽车测试场景生成方法 [J]. 中国公路学报, 2024, 37(3): 170-180.
- TAN Dong-kui, ZHU Bo, HU Xu-dong. Test scenario generation method for intelligent vehicles based on neighbor object region representation and conditional variational autoencoder [J]. China Journal of Highway and Transport, 2024, 37(3): 170-180.
- [11] 马峻岩, 田叶凡, 赵祥模, 等. 基于自然驾驶数据挖掘的二阶车辆与行人交互测试场景 [J]. 中国公路学报, 2022, 35(3): 139-152.
- MA Jun-yan, TIAN Ye-fan, ZHAO Xiang-mo, et al. Two-stages test scenarios for interaction of vehicles and pedestrians based on naturalistic driving data mining [J]. China Journal of Highway and Transport, 2022, 35(3): 139-152.
- [12] XU C J, PETIUSHKO A, ZHAO D, et al. DiffScene: Diffusion-based safety-critical scenario generation for autonomous vehicles[C]//Proceedings of the AAAI Conference on Artificial Intelligence. [S.l.]: AAAI Press, 2025, 39(8): 8797-8805.
- [13] LI Q Y, PENG Z H, FENG L, et al. MetaDrive: Composing diverse driving scenarios for generalizable reinforcement learning[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2023, 45(3): 3461-3475.
- [14] MA Y N, JIANG W, ZHANG L T, et al. Evolving testing scenario generation and intelligence evaluation for automated vehicles[J]. Transportation Research Part C: Emerging Technologies, 2024, 163: 104620.
- [15] CHEN K Y, LEI Y H, CHENG H, et al. FREA: Feasibility-guided generation of safety-critical scenarios with reasonable adversariality[C]//AGRAWAL P, KROEMER O, BURGARD W, eds. Proceedings of the 8th Conference on Robot Learning. [S.l.]: PMLR, 2025: 566-586.
- [16] CHEN K Y, SUN W C, CHENG H, et al. RIFT: Closed-Loop RL Fine-Tuning for Realistic and Controllable Traffic Simulation[J]. ArXiv, 2025, DOI: 10.48550/arXiv.2505.03344.
- [17] LI Y Z, ZHANG L R, XIA B, et al. DeepMF: Deep motion factorization for closed-loop safety-critical driving scenario simulation[C]//Proceedings of the 2025 International Joint Conference on Neural Networks (IJCNN). Piscataway, NJ: IEEE, 2025: 1-8.

- [18] TUNCALI C E, FAINEKOS G, ITO H, et al. Simulation-based adversarial test generation for autonomous vehicles with machine learning components[C]//Proceedings of the 2018 IEEE Intelligent Vehicles Symposium (IV). Piscataway, NJ: IEEE, 2018: 1555-1562.
- [19] U.S. DEPARTMENT OF TRANSPORTATION FEDERAL HIGHWAY ADMINISTRATION. Next Generation Simulation (NGSIM) Vehicle Trajectories and Supporting Data[DB/OL]. Washington, D.C.: U.S. Department of Transportation Federal Highway Administration, 2016[2026-01-07].
- [20] GAO J Y, SUN C, ZHAO H, et al. VectorNet: Encoding HD maps and agent dynamics from vectorized representation[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ: IEEE, 2020: 11525-11533.
- [21] QASSIM H, VERMA A, FEINZIMER D. Compressed residual-VGG16 CNN model for big data places image recognition[C]//Proceedings of the 2018 IEEE 8th Annual Computing and Communication Workshop and Conference (CCWC). Piscataway, NJ: IEEE, 2018: 169-175.
- [22] HUANG Z Y, WU J D, LV C. Driving behavior modeling using naturalistic human driving data with inverse reinforcement learning [J]. IEEE Transactions on Intelligent Transportation Systems, 2022, 23(8): 10239-10251.
- [23] CHITTA K, PRAKASH A, GEIGER A. NEAT: Neural attention fields for end-to-end autonomous driving[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV). Piscataway, NJ: IEEE, 2021: 15793-15803.
- [24] BEHL A, CHITTA K, PRAKASH A, et al. Label efficient visual abstractions for autonomous driving[C]//Proceedings of the 2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). Piscataway, NJ: IEEE, 2020: 2338-2345.
- [25] REMPE D, PHILION J, GUIBAS L J, et al. Generating useful accident-prone driving scenarios via a learned traffic prior[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ: IEEE, 2022: 17305-17315.
- [26] CAI X, BAI X S, CUI Z Y, et al. Adversarial stress test for autonomous vehicle via series reinforcement learning tasks with reward shaping [J]. IEEE Transactions on Intelligent Vehicles, 2025, 10(2): 832-847.
- [27] WANG J K, PUN A, TU J, et al. AdvSim: Generating safety-critical scenarios for self-driving vehicles[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ: IEEE, 2021: 9909-9918.