

DK-Port: 基于大语言模型和强化学习的港口自动驾驶仿真环境构建与验证

娄云洁¹, 艾明飞¹, 庄术洁¹, 于海¹, 王鑫¹, 滕出¹, 王江成², 沈甜雨³, 郝坤坤², 崔文^{2,4,5}

1. 青岛前湾联合集装箱码头有限责任公司, 山东 青岛 266520;
2. 安徽深信科创信息技术有限公司, 安徽 合肥 230088;
3. 北京化工大学信息科学与技术学院, 北京 100029;
4. 交叉信息核心技术研究院(西安)有限公司, 陕西 西安 710077;
5. 西安电子科技大学前沿交叉研究院, 陕西 西安 710071)

摘要: 与常规驾驶环境相比, 港口环境具有车辆作业繁忙、道路定制化以及人机车辆混行等特点。为了解决港口自动驾驶数据缺乏和算法泛化性问题, 缩短算法开发周期和降低开发成本, 以具身智能设计理念为依托, 以提供逼真可控环境、车辆交互过程为目标, 提出了DK-Port港口自动驾驶仿真环境的构建与验证方法。首先, 基于调查问卷和专家经验, 采用零次提示和思路链等技术, 使通用大语言模型参与奖励函数的设计过程; 然后, 基于复杂道路上车辆的行驶数据, 构建丰富且逼真的人机混行仿真交互场景, 并利用PPO深度强化学习算法训练对抗驾驶员模型, 以揭示自动驾驶算法的安全隐患; 最后在直道和交叉路口等4种典型场景下进行了对比实验。结果表明DK-Port能够有效生成更符合港口实际特性的多类型对抗驾驶行为, 如危险超车、紧急切入、抢占路口交汇点等。在保证关键指标分布合理的前提下, 直道场景下的变道次数是基准方法的2.9倍, 交叉路口场景下的紧急制动率提升了46.7%。

关键词: 港口自动驾驶; 仿真环境构建; 大语言模型; 强化学习; 对抗场景生成

中图分类号: TP39

文献标志码: A

doi: 10.11959/j.issn.2096-6652.202509

DK-Port: construction and validation of port autonomous driving simulation environment based on large language models and reinforcement learning

LOU Yunjie¹, AI Mingfei¹, ZHUANG Shujie¹, YU Hai¹, WANG Xin¹, TENG Chu¹, WANG Jiangcheng², SHEN Tianyu³, HAO Kunkun², CUI Wen^{2,4,5}

1. Qingdao Qianwan United Container Terminal Co., Ltd., Qingdao 266520, China
2. Anhui Shenxin Kechuang Technology Co., Ltd., Hefei 230088, China
3. College of Information Science and Technology, Beijing University of Chemical Technology, Beijing 100029, China
4. Institute of Interdisciplinary Information Core Technology (Xi'an) Co., Ltd., Xi'an 710077, China
5. Academy of Advanced Interdisciplinary Research, Xidian University, Xi'an 710071, China

Abstract: Port environments differ significantly from conventional driving environments, characterized by high vehicle activity, customized roadways, and mixed traffic involving human-operated and autonomous vehicles. To address the lack

收稿日期: 2024-08-05; 修回日期: 2024-10-22

通信作者: 崔文, cuiw@iisict.com; 郝坤坤, haokunkun@synkrotron.ai

基金项目: 西安交通大学人机混合增强智能全国重点实验室开放课题 (No.HMHAI-202416)

Foundation Item: Open Project of the National Key Laboratory of Human-Machine Hybrid Augmented Intelligence, Xi'an Jiaotong University (No.HMHAI-202416)

of port-specific autonomous driving data and improve algorithm generalization while reducing development costs and time, this paper proposes a simulation environment for port autonomous driving called DK-Port. Leveraging embodied intelligence principles, DK-Port provides a realistic and controllable environment for vehicle interactions. This paper incorporates survey-based expert input, employing zero-shot prompting and chain-of-thought reasoning techniques to enable large language models to assist in designing reward functions efficiently. Then, human-machine mixed-traffic scenarios are constructed using complex road data, and adversarial driver models are trained with the PPO reinforcement learning algorithm to identify safety vulnerabilities in autonomous driving systems. Comparative experiments in four typical scenarios, including straight roads and intersections, demonstrate that DK-Port generates diverse adversarial driving behaviors aligned with real port characteristics, such as dangerous overtaking and abrupt lane cutting. Under the premise of ensuring a reasonable distribution of key metrics, the number of lane changes in the straight road scenario is 2.9 times that of the baseline method, and the emergency braking rate in the intersection scenario is increased by 46.7%.

Key words: ports autonomous driving, simulation environment construction, large language model, reinforcement learning, adversarial scenario generation

0 引言

港口负责全球超过80%的商品运输,其运营效率的波动直接影响到食品、能源等关键物资的转运,并对全球经济产生深远影响。然而,受到装卸作业和港口调度等因素的制约,船只到港后通常需等待4~8 h,严重阻碍了贸易的高效进行^[1-2]。为解决这一问题,港口自动化改造已成为业界的广泛共识^[3]。麦肯锡研究显示,自动化改造完成后,港口的运营成本有望降低55%,运营效率提升35%,显著减轻港口的吞吐压力^[4]。自1990年起,港口自动化首先在堆场决策管理领域得到应用,随后逐步扩展到其他运营领域^[5]。近年来,随着自动驾驶和人工智能技术的快速发展,自动驾驶运输车已成为港口地面运输的关键组成部分,吸引了全球许多先进港口的广泛关注^[6-9]。

相较于传统城市或高速道路的自动驾驶场景,港口的特殊环境带来了更多技术挑战^[10]。港口是封闭的商业环境,数据采集被严格限制,导致训练数据严重缺乏。并且港口的路网结构、驾驶行为和交通规则与常规交通场景存在差异,使现有的自动驾驶数据集难以直接应用于港口场景,相关的自动驾驶算法也难以达到传统场景下的性能水平^[11-12]。这些因素导致港口自动驾驶技术目前仍主要依赖于

基于巡线的自动导向车,在灵活性和应用性上大打折扣,影响了自动驾驶技术在港口的广泛应用^[13-14]。与处于封闭环境的仓库或矿山不同,港口空间受限且改动成本巨大,不能轻易划分出专门的自动驾驶行驶区域,致使自动驾驶车辆与人工驾驶车辆混行的现象普遍存在^[15-17]。在实际操作中,如图1所示,为了加快运输任务的完成,人类驾驶员常采取突然变道、紧急制动或路口交汇不让行等激进驾驶行为,甚至通过强行变道和超车等方式抢占自动驾驶车辆的行驶路权^[18-20],造成自动驾驶车辆运营低效。自动驾驶车辆在港口的实际表现远低于预期,其应用前景不明晰,使港口管理方难以对更广泛的自动驾驶技术升级进行投资决策,进一步延缓了港口自动驾驶技术的提升及其商业化进程^[21-22]。

为了全面验证港口环境下自动驾驶算法的安全性和可靠性,同时提升算法在复杂的人机混行驾驶环境中的表现,本文提出了D(data-driven)K(knowledge-driven)-Port港口自动驾驶仿真环境的构建和验证方法。该方法受具身智能设计理念启发,着重于强调构建仿真环境中可与被测车辆充分交互的驾驶员智能体,充分参与自动驾驶算法的测试与训练过程。DK-Port通过灵活、多样且逼真的驾驶员模型,全面模拟港口内人机车辆的交互行为,旨在加速港口自动化的进程,形成新时代平行



图1 港口环境激进驾驶行为示意

智能^[23-24]。区别于针对常规道路和驾驶情况的现有仿真方法^[25-28], DK-Port 结合专家经验和复杂道路行驶数据, 利用通用大语言模型和强化学习技术来应对港口数据集缺失和人机车辆混行验证不充分的挑战^[29-30]。首先, 向资深港口车队(运营200辆以上集装箱卡车)的管理人员发放调查问卷, 收集港口车辆行为特征数据, 并结合零次提示和思路链技术, 利用大语言模型的逻辑推理和语义扩展能力^[31-32], 设计奖励函数生成工作流 Port-Bot 归纳问卷内容, 从而增强通用大语言模型在港口领域的知识理解能力。其次, 在 Port-Bot 总结的内容的基础上, 进一步挖掘通用大语言模型的代码和功能函数合成能力^[33-35], 引导 Port-Bot 完成强化学习奖励函数的设计, 为模拟人类驾驶的对抗驾驶员模型提供训练基础。最后, 以青岛港前湾联合集装箱码头^[36]作为具体应用案例, 并结合NGSIM^[37]和INTERACTION^[38]两大公开数据集中的典型复杂场景, 基于所设计的奖励函数和强化学习算法^[39]训练对抗场景驾驶员模型, 完成在直道和交叉路口等4种港口典型场景中的激进驾驶行为仿真与算法性能验证。实验表明, DK-Port 能生成符合港口实际特性的多种类型的对抗驾驶行为, 如危险超车、紧急切入、强占路口交汇点等。在保证关键指标分布合理的情况下, 直道场景下的变道次数是基准方法的2.9倍, 交叉路口的紧急制动率提升了46.7%, 为港口自动驾驶算法的验证与提升提供了高效的解决方案。本文主要贡献如下。

(1) 构建了结合数据驱动与知识驱动的DK-Port港口自动驾驶仿真环境。该环境通过多样且逼真的驾驶员模型, 全面模拟港口内人机混行驾驶行为, 解决了港口自动驾驶领域的数据缺乏和算法适用性问题, 显著提升了仿真环境的真实性和适用性。

(2) 开发了奖励函数生成工作流 Port-Bot, 增强了驾驶员智能体的行为生成能力。在DK-Port仿真环境的具体实现上, 利用大语言模型和强化学习技术, Port-Bot 系统性整理并生成港口车辆的行为特征, 设计了适应港口环境的奖励函数, 提升了仿真环境中驾驶员智能体的行为多样性和逼真性。

(3) 实验验证了DK-Port在实际港口环境中的有效性。以青岛港前湾联合集装箱码头为案例, DK-Port 能够生成多种符合港口实际的激进行为, 显著提升了自动驾驶算法的验证效率, 为港口自动化进程提供了高效的解决方案。

文章的结构如下: 第1节介绍基于大语言模型的奖励函数生成工作流 Port-Bot; 第2节详述大语言模型对强化学习奖励函数的整合与实现细节; 第3节介绍基于复杂道路行驶数据的强化学习算法训练流程与仿真环境输入输出内容; 第4节进行实验介绍与结果分析; 第5节总结全文, 并展望未来工作。

1 基于通用大语言模型的强化学习奖励函数设计

模拟港口复杂混行驾驶环境的关键在于能否通过强化学习训练出逼真的人类驾驶员模型, 以实现对抗车辆与被测车辆之间复杂的交互过程。奖励函数在这一过程中起着决定性作用, 因为它可通过奖励或惩罚的方式对驾驶行为进行反馈, 从而影响人类驾驶员模型的训练效果, 实现模型训练的目标。传统奖励函数设计主要分为2种方式。一种是基于专家的设计, 这依赖于专家的前期经验积累^[40]。由于不同港口场景的复杂多变, 这种方法常常需要投入大量的人力和时间来不断调整奖励函数。另一种基于逆向强化学习, 它依赖海量数据来求解与数据集规律相符的奖励函数^[41]。虽然这种方法减少了对经验的依赖, 但其对数据的强依赖性限制了它在数据稀缺的港口环境中的直接应用。

不同于上述2种方法, 通用大语言模型已具备广泛的基础经验知识, 可减少对专家经验和数据集的依赖, 因此基于大语言模型的强化学习奖励函数设计方法近年来受到更多关注。然而, 这种方法也存在一些不足: 大语言模型虽然知识广泛但在专业领域相对较弱, 需要通过多轮专家反馈校正模型, 这将引入额外的模型设计成本; 在制定奖励函数的过程中与大语言模型的交互复杂度高, 容易引发模型产生幻觉^[42], 导致产生大量错误信息。此外, 沈甜雨等^[43]提出, 智能驾驶系统在复杂环境中的适应能力不仅依赖于模型的感知和控制, 还需要通过持续交互实现系统的长期进化。

在上述分析的基础上, 提出了如图2所示的奖励函数生成工作流 Port-Bot。Port-Bot 将调查问卷结果和专家的经验知识作为输入, 利用零次提示和思路链技术来提升大语言模型对港口领域知识的理解与推理能力, 以进一步扩充问卷信息。这种方法不仅提高了奖励函数设计的适应性, 还提升了其准确性, 从而支持面向复杂港口条件的自动驾驶仿真环境开发。

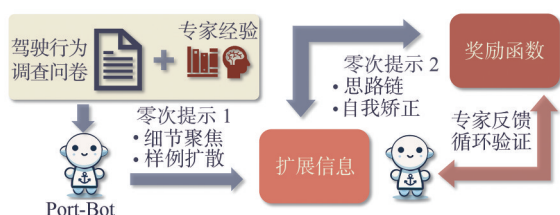


图2 奖励函数生成 workflow Port-Bot

1.1 面向驾驶员行为和心理特征的问卷调研

若要构建逼真的人类驾驶员模型，了解驾驶员的行为和心理特征至关重要。鉴于大语言模型在专业知识方面的局限性以及港口数据的匮乏，通过向资深港口车队的管理人员发放调查问卷，以收集港口车辆驾驶行为的特征数据。问卷设计采用横截面调查方法^[44]，该方法被广泛应用于医学和商业研究，能够在单一时间点通过较少的问题快速捕捉被调查对象的特征。问卷涉及以下内容。

- 涵盖话题：包括风险行为（如超速、使用手机）、与自动驾驶车辆的互动（如超车、抢道等激进行为）以及对交通效率的整体影响等问题。调查特别关注强化学习奖励函数中需要设定的激进驾驶行为。

- 回答格式：问题主要采用李克特量表（Likert scale）^[45]，例如“从来没有”“很少”“有时”“经常”，这种格式适合于精确测量感知和态度相关的问题。

1.2 问卷信息扩展与丰富

通过问卷，初步了解港口内司机的驾驶行为偏好，具体的分析结果将在实验部分详细介绍。然而，横截面调查方法通常要求问题保持精简，因此该方法获得的数据量相对较少。为了丰富和深入理解这些调查结果，传统方法是邀请领域专家进行信息的补充和完善，但其灵活性较低，不利于广泛应用。与此相比，本文设计的Port-Bot workflow将在问卷调研结束后以大语言模型为支撑完成信息的扩展和丰富，以生成更为详尽的内容。具体的实施方式如下。

- 使用零次提示进行信息扩展。在Port-Bot workflow中，采用零次提示技术来扩展调查结果。通过针对性设计零次提示词，可向通用大语言模型一次性提供完整详尽的信息，从而有效生成丰富且准确的内容。此方法将简化与大语言模型多轮逻辑交互的过程，可显著降低内容生成的复杂性，同时避免因多轮交互带来的误差累积，降低模型产生幻觉

的风险。

- 借鉴认知科学概念优化提示词。在设计提示词时，借鉴了认知科学的部分技巧。例如，在使用思路链（chain-of-thought, CoT）^[46]技术时，通过简单指令（如“let’s think step by step”）加强模型的逻辑推理能力。此外，还采用了“样例扩散”概念，引导大语言模型使用具体样例（如在提示词中加入“vivid examples or anecdotes”），促使模型在现有信息基础上进行创造性扩展和全面化分析，提高生成内容的实用性。同样，“细节聚焦”概念也被应用于指导模型更加关注描述的细节（如指令中加入“pay close attention to detail in your descriptions”），使模型捕捉到更多关于人类驾驶员行为模式和交互的细节，增强奖励函数设计的全面性，提高其精准度。

1.3 面向奖励函数实现的提示词设计

在上述信息基础上，大语言模型已具备生成奖励函数的前置条件，可以通过直接输入提示词引导模型生成具体的奖励函数或相应源码。然而，前期实验中发现此操作会导致以下问题。①奖励函数与输入信息不匹配导致错误输出。例如，本研究旨在生成能模拟人类驾驶员行为的模型，以完成与自动驾驶车辆的对抗驾驶。这与常规的自动驾驶模型训练所关注的小心驾驶行为规划有本质区别。直接输入会导致大语言模型产生信息混淆，错误地鼓励小心驾驶行为而惩罚对抗驾驶，从而违背了设计目标。②奖励函数直接生成的源码与代码环境不匹配^[47-48]。代码环境的更新速度通常快于大语言模型的训练更新速度，直接输入会导致合成代码滞后。此外，目前的代码合成主要针对常用代码库，对自动驾驶对抗训练等特定领域的适配能力较弱，特别是在调用平台中不常见但关键的函数时极易出错。

为了解决这些问题，本文的Port-Bot workflow采取以下关键步骤。

- 使用零次提示进行自我反思和CoT提示。在设计大语言模型输出答案之前，增加描述解决问题的步骤和CoT逻辑过程，以加强输出内容的丰富度。同时，明确增加内部校验过程，主动监测和评估答案的质量^[49]。这种方法可使模型在提供答案之前展示其思考和校验过程，该过程不仅会增强回答内容的透明度，还允许模型在逻辑推断中进行自我修正，减少错误的发生。

- 适时引入专家接入以闭环代码合成。大语

言模型在设计初步奖励函数方案后,专家可根据自己的专业知识和实际经验对模型输出进行评审和优化。专家可以指出回答的不足或错误,并提供具体反馈以指导模型进行更精确的调整。经过1~2轮反馈后,由专家完成代码合成的最后环节,确保代码的质量和 workflows 的实用性。

2 奖励函数整合与具体实现

第1节详细探讨了如何利用通用大型语言模型的综合知识和推理能力来减少错误信息的生成。在Port-Bot工作流中,接下来的关键任务是确定何时由专家介入以及如何完成奖励函数的最终代码实现。经研究发现,大型语言模型在以数学表达式和语言解释的形式呈现奖励函数时最为有效,这是因为直接的代码实现可能由于所依赖平台的差异而频繁出现编码错误。专家可以以大语言模型的输出为指导,在大幅降低奖励函数设计的难度和提高准确性的同时,也确保了代码编写过程的实用性。以下将详细介绍奖励函数的具体实现方式。此外,为了准确训练对抗驾驶员模型,本文采用了在直道和交叉路口分别进行训练的方式,这种方法能更真实地反映港口环境中不同交通场景下车辆的交互特点,从而有效地模拟现实驾驶环境中的各种挑战。

2.1 直道场景下的奖励函数设计与整合

2.1.1 大语言模型为主的直道奖励函数设计

针对港口直道场景,为了鼓励对抗车辆产生具有风险或不合规的行为(如超速),从而为自动驾驶车辆提供更严格的测试环境,Port-Bot从以下4个方面进行奖励函数的设计。

- 攻击性奖励 R_{attack} : 奖励对抗车辆进行攻击性操作,如突然切入或车道侵占,以测试自动驾驶车辆在处理交通环境突变时的能力。

- 干扰交通奖励 $R_{\text{disruption}}$: 奖励那些干扰正常交通流的行为,以测试自动驾驶车辆适应其他道路使用者不可预测行为和不合作行为的能力。

- 风险参与奖励 R_{risk} : 鼓励从事风险行为,如近距离超车或突然刹车,目的是测试自动驾驶系统对突发操作的反应能力。

- 速度激励奖励 R_{speed} : 奖励对抗车辆策略性地控制速度,以模拟某些人类驾驶员的竞争性驾驶行为。

最终的直道奖励函数 R_{straight} 为:

$$R_{\text{straight}} = \omega_1 R_{\text{attack}} + \omega_2 R_{\text{disruption}} + \omega_3 R_{\text{risk}} + \omega_4 R_{\text{speed}} \quad (1)$$

其中,权重 ω_1 、 ω_2 、 ω_3 和 ω_4 用来平衡不同类型的对抗行为,反映每种行为在模拟场景中的相对重要性,旨在更好地模拟真实世界中的复杂驾驶环境。

2.1.2 基于专家经验的直道场景奖励函数整合

(1) 攻击性奖励 R_{attack}

在直道场景中,“突然切入”行为是一类比较典型的攻击性场景,可对其进行攻击性奖励设计,具体实现如式(2)所示。

$$R_{\text{attack}} = \frac{1}{2} \left(e^{-\frac{|x_{\text{adv}} - x_{\text{ego}} - d_{\text{safe}}|}{d_{\text{safe}}}} + e^{-\frac{|y_{\text{adv}} - y_{\text{npc}}|}{w_{\text{lane}}}} \right) \quad (2)$$

其中, x_{adv} 、 x_{ego} 分别表示对抗车和被测车的横坐标, y_{adv} 表示对抗车的纵坐标, y_{npc} 表示协同对抗车的纵坐标, d_{safe} 和 w_{lane} 分别表示安全距离和车道宽度。

该函数的设计包含以下几个关键思路。一是场景中引入对抗车(adv)与协同对抗车(npc),模拟自然“突然切入”行为。当对抗车同车道前方的协同对抗车急刹车或临时停车时,对抗车被迫突然

切入被测车所在车道,通过 $e^{-\frac{|y_{\text{adv}} - y_{\text{npc}}|}{w_{\text{lane}}}}$ 确保 adv 与 npc 尽可能地处于同一车道,激发自然的“突然插入”行为。二是为了进一步确保“突然插入”行为

的自然性,通过 $e^{-\frac{|x_{\text{adv}} - x_{\text{ego}} - d_{\text{safe}}|}{d_{\text{safe}}}}$ 鼓励对抗车辆与被测车辆至少保持一个最小的安全距离,以防止不可避免的碰撞,保持测试的有效性;三是本文使用指数函数表示奖励函数,将变量范围限制在(0,1)区间内,侧重奖励符合条件的行为,同时对不符合条件的行为给予较小的奖励或无奖励。

(2) 风险参与奖励 R_{risk}

在直道场景中,“突然刹车”是一类比较典型的危险行为,可对其进行风险参与奖励设计,具体实现如式(3)所示。

$$R_{\text{risk}} = \begin{cases} R_{\text{same_lane}} + R_{\text{TTC}}, & \text{LaneID}_{\text{adv}} = \text{LaneID}_{\text{ego}} \\ R_{\text{same_lane}}, & \text{其他} \end{cases} \quad (3)$$

其中, $\text{LaneID}_{\text{adv}}$ 、 $\text{LaneID}_{\text{ego}}$ 分别代表对抗车和被测车所在的车道编号, $R_{\text{same_lane}}$ 是当对抗车与被测车处于同一车道时的基础奖励,设定为0.1。TTC表示车辆在保持当前的朝向与速度后产生碰撞所需的时间, R_{TTC} 为TTC的指数化度量,可表示为

$$R_{\text{TTC}} = e^{-\frac{\text{TTC}}{\text{TTC}_{\text{max}}}}, \text{ 其中, } \text{TTC}_{\text{max}} \text{ 根据场景的道路特}$$

性(如限速规则)设定。

该函数的设计思路如下: ①利用 R_{TTC} 对更小的 TTC 进行奖励, 以增强对抗车的风险驾驶特性, 即通过缩短被测车辆的响应时间来测试其反应能力; ②该函数将测试场景尽可能地限制在同一车道内, 突出“突然刹车”行为发生的自然性, 同时避免对抗车辆跨车道直接碰撞被测车, 规避无法避免的碰撞场景。

(3) 干扰交通奖励 $R_{disruption}$

根据干扰交通奖励的描述, 本文选择复用风险参与奖励的方式来实现干扰交通行为。此设计的考虑如下: ①缩短 TTC, 增加对抗车辆的突发性行为频率, 从而测试被测车辆的反应能力和适应交通干扰的能力; ②复用风险参与奖励机制, 不仅可以降低设计复杂度, 还可避免奖励欺骗^[50]现象, 加速对抗学习模型对目标行为的收敛。

(4) 速度激励奖励 R_{speed}

根据速度激励奖励的描述, 借鉴攻击性奖励 R_{risk} 的设计思路, 其具体奖励函数实现如式 (4) 所示。

$$R_{speed} = \frac{1}{2} \left(e^{-\frac{|x_{adv} - x_{npc} - v_{diff}|}{v_{diff}}} + e^{-\frac{|y_{adv} - y_{npc}|}{w_{lane}}} \right) \quad (4)$$

其中, v_{diff} 表示对抗车和协同对抗车的车速差。该设计确保两车在同一车道上维持一定的速度差, 从而持续对被测车辆施加压力, 考验其在变速驾驶环境下的性能和安全性。

2.2 交叉路口场景下的奖励函数设计与整合

2.2.1 大语言模型为主的路口奖励函数设计

交叉路口场景与直道场景的主要构成相似, 不同点在于路口场景不再强调速度的影响, 取而代之的是对无红绿灯路口路权的协商。交叉路口场景下奖励函数 $R_{intersection}$ 为:

$$R_{intersection} = \omega_1 R_{negotiation} + \omega_2 R_{attack} + \omega_3 R_{disruption} + \omega_4 R_{risk} \quad (5)$$

其中, $R_{negotiation}$ 表示路权抢占奖励, 奖励对抗车在路口通过加减速行为来抢占路口关键点位, 从而通过对路口的被测车辆构成安全威胁。

2.2.2 基于专家经验的路口场景奖励函数整合

(1) 路权抢占奖励 $R_{negotiation}$

在交叉路口场景中, “路口抢行”行为是一类比较典型的路权抢占场景, 其具体奖励函数实现如式 (6) 所示。

$$R_{negotiation} = \begin{cases} e^{-|t_{ego} - t_{adv}|}, & t_{ego} \geq t_{adv} \\ -1 + e^{-|t_{ego} - t_{adv}|}, & t_{ego} < t_{adv} \end{cases} \quad (6)$$

其中, t_{ego} 和 t_{adv} 为被测车辆和对抗车辆到达关键点的时间, 关键点可为路口处指定的必经点。该函数的设计思路为鼓励对抗车优先抢占关键点, 从而验证被测车辆能否灵活调整路线并规避风险。

(2) 攻击性奖励 R_{attack}

根据攻击性奖励的描述, 可延续关键点抢占这一具体实现目标, 其具体奖励函数实现如式 (7) 所示。

$$R_{attack} = \frac{1}{2} \left(\left(1 - e^{-\max(v_{ego} - v_{adv}, 0)} \right) + e^{d_{adv}} \right) \quad (7)$$

其中, v_{ego} 和 v_{adv} 分别表示被测车和对抗车的速度, d_{adv} 表示对抗车与关键点的距离。该函数的设计思路如下: ①与路权抢占从时间上约束相似, 此次从空间角度鼓励对抗车抢占关键点; ②鼓励对抗车与被测车保持较大速度差, 并且对抗车相对缓行, 从而构造关键点被占用的情形, 测试被测车能否应对这种情形。

(3) 干扰交通奖励 $R_{disruption}$

根据干扰交通奖励的描述, 可以复用攻击性奖励内容降低奖励函数复杂度, 其内容中已包含低速占用道路的含义。

(4) 风险参与奖励 R_{risk}

在交叉路口场景中, “碰撞”是一类最为严重的风险行为, 可把其作为风险参与奖励设计的依据, 具体实现如式 (8) 所示。

$$R_{risk} = \begin{cases} -1, & \text{碰撞} \\ 0, & \text{其他} \end{cases} \quad (8)$$

该函数的设计思路为: 考虑到路口场景下车辆的车距普遍较近, 不宜再用安全距离来避免不可避免的碰撞, 因此可直接对碰撞这一状态进行惩罚, 增强对抗场景的自然性。

3 基于复杂道路行驶数据的强化学习算法训练流程

复杂道路行驶数据是强化学习驾驶员模型训练的基础, 依靠具有多种属性的数据所构建的仿真环境, 动态模拟港口中车辆的交互过程。由于缺失港口环境数据, 考虑港口实际地图具有直道、交叉路口等不同的路网形态, 本文采用真实复杂交通数据

集NGSIM和INTERACTION来弥补港口数据集不足的问题,实现对港口路网环境的有效对齐。基于仿真环境、复杂交通数据和专家知识,强化学习算法训练流程,如图3所示。

3.1 强化学习中的策略网络与价值网络

在DK-Port的设计中,本文选定近端策略优化(proximal policy optimization, PPO)^[51]网络作为具体实现,PPO网络通过在目标函数中引入CLIP函数的方式来限定网络的更新幅度,可有效保证网络更新速度和幅度的均衡,适用于自动驾驶场景的对抗场景构建。

PPO网络主要包含策略网络 $\pi_\theta(a_t|s_t)$ 和价值网络 $V_\phi(s_t)$,策略网络根据仿真环境反馈的状态信息 s_t 生成新的动作 a_t ,价值网络负责评估策略网络生成动作的期望回报并反馈,提升策略网络生成能力。

PPO网络的整体目标函数如式(9)所示。

$$L^{\text{Total}}(\theta, \phi) = \mathbb{E}_t[L^{\text{CLIP}}(\theta) - c_1 L^{\text{VF}}(\phi) + c_2 S[\pi_\theta](s_t)] \quad (9)$$

其中, $S[\pi_\theta](s_t)$ 表示网络的熵,用来鼓励策略网络在行为域内进行充分的探索,防止策略陷入局部最优; c_1 和 c_2 为均衡参数; $L^{\text{CLIP}}(\theta)$ 和 $L^{\text{VF}}(\phi)$ 分别为策略网络和价值网络的目标函数。

$$L^{\text{CLIP}}(\theta) = \mathbb{E}_t[\min(r_t(\theta)A_t, \text{CLIP}(r_t(\theta), 1-\varepsilon, 1+\varepsilon)A_t)] \quad (10)$$

$$L^{\text{VF}}(\phi) = \mathbb{E}_t[(V_\phi(s_t) - V_t^{\text{targ}})^2] \quad (11)$$

其中, A_t 为执行行为 a_t 较其他行为的优势, $r_t(\theta)$ 为新策略下执行行为 a_t 的概率与旧策略的概率之比,如果 $r_t(\theta) > 1$,则意味着新策略比旧策略更有可能执行行为 a_t ;在策略网络的目标函数中,利用 $\text{CLIP}(r_t(\theta), 1-\varepsilon, 1+\varepsilon)$ 中的超参数 ε 来限制 $r_t(\theta)$ 的范围,并最终控制策略更新的幅度,如果 $r_t(\theta)$ 超出这个区间,它就会被裁剪到区间的边界上, ε 一般取0.1或0.2; $\mathbb{E}[\cdot]$ 为小批量数据的经验均值估计, $V_\phi(s_t)$ 为价值网络所评估状态 s_t 的价值, V_t^{targ} 为基

于当前实际奖励推算的可得价值,价值网络通过降低二者的差距来提升行为价值评估的准确度,从而更好地为策略网络提供反馈。

3.2 仿真环境的输入与输出

仿真环境是强化学习网络获取行为反馈的关键组成部分,通常根据仿真目标的不同,其侧重点也会有所变化。例如,注重感知功能的仿真环境侧重全面的渲染效果,而面向规控功能的仿真环境更注重动力学属性。定制仿真环境的输入与输出,可以更准确地评估模型性能,本文关注的输入和输出内容如下。

输入: 仿真环境的输入通常仅包括由强化学习策略生成的行为^[52],这些行为的执行效果由内置的奖励函数来判定。为了精确模拟港口环境的特性,DK-Port将复杂的交通数据和专家知识整合为输入,参与奖励函数和仿真参数的配置。第2节已介绍了奖励函数的设计细节,此处主要阐述如何将复杂交通数据应用于仿真配置中,以提高模型训练的适应性和精度。

在强化学习训练过程中,完整的仿真环境由受控车辆或对抗车辆以及其他交通参与者共同组成。为了精确模拟车辆间的交互,参与者的行为描述至关重要。现有方法通常采用IDM^[53]和MOBIL^[54]模型来控制车辆的直行和变道行为,实现在计算复杂度较低的情况下的车辆行驶模拟。通过对现有数据特征的分析 and 港口环境特征的匹配,本文对IDM和MOBIL模型的关键参数进行了调整,包括直行行驶的目标速度、安全车距、最大加速度、舒适减速速度,以及路口的最小加速度增益和最大制动等。详细的调整方法将在实验部分展示。

输出: 仿真环境的输出主要包括状态和奖励。状态输出取决于仿真目标,可选择不同侧重点的仿真平台以提供相应的状态信息;奖励输出则通过设计的奖励函数来实现,该函数能有效地反映港口特

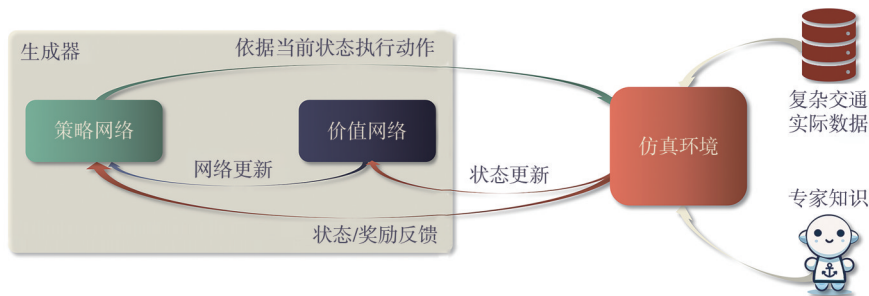


图3 强化学习算法训练流程

性,确保从具体行为到奖励值的有效映射,同时这一过程不受特定仿真平台的限制。

4 实验及结果分析

4.1 调查问卷结果分析与大语言模型提示词示例

4.1.1 问卷分析

本文通过对资深港口车队的管理者进行问卷调查,了解司机对特定驾驶行为的态度,以及在人工驾驶车辆与自动驾驶车辆互动中可能出现的复杂动态。理解司机的行为模式,有助于模拟真实的交互场景,提高仿真系统的适应性。

图4展示了对特定高风险驾驶行为的问卷调查结果,并汇总了车队管理者基于长期经验对自动驾驶技术引入的看法。结果显示,尽管港口内车辆倾向于安全驾驶,超速和抢道/别车的行为仍然存在,因此在模型设计中加入一定量的激进行为可以反映真实的驾驶状况。管理者普遍对自动驾驶的引入表示担忧,认为自动驾驶车辆与人类驾驶员将处于竞争关系,人类驾驶员可能考虑自身利益而不礼让自动驾驶车辆。此外,自动驾驶算法在港口环境下尚未成熟,这种不礼让行为可能会进一步影响自动驾驶的运行效率,从而降低整个港口的作业效率。

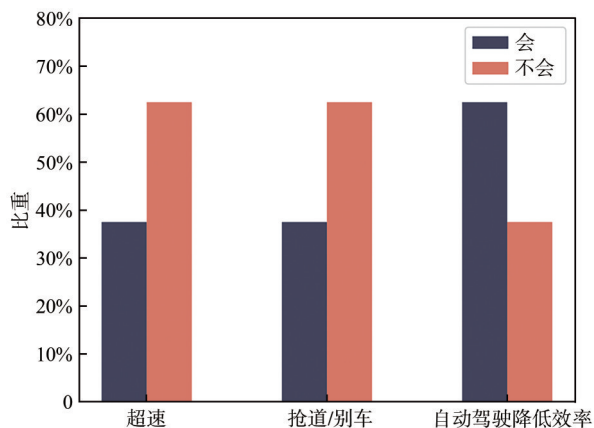


图4 高风险驾驶行为问卷调查结果

本文统计了视线偏离驾驶路线、操作其他车上设备、使用手机、急刹车/突然加速、无故缓行/停车/逆行和频繁超车等典型的违规或激进驾驶行为,如图5所示。尽管港口内的驾驶违规或激进行为总体较少,但在特定情境下这些行为仍然存在。例如:司机在长时间工作后可能会出现懈怠、操作其他设备或使用手机来消磨时间;在高峰期交通或紧急情况下,急刹车/突然加速、无故缓行/停车/逆行

等行为也会更加突出。本文在设计奖励函数时特别考虑了这些行为对交通参与者的实际影响。

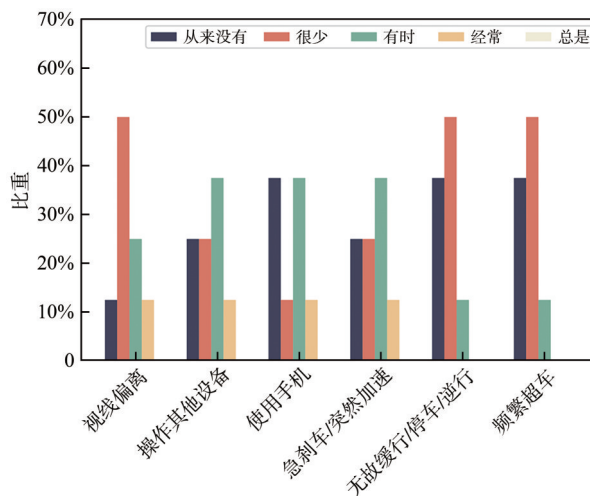


图5 典型违规/激进驾驶行为统计

4.1.2 大语言模型关键提示词

第1节描述了与大语言模型的交互过程,本节对重点提示词进行展示,其主要包含3个部分。

(1) 聚焦问卷细节,完成内容补充

提示词中主要包含以下内容:根据港口驾驶行为的问卷结果生成详细叙述,包含典型驾驶行为的具体例子,特别是自动驾驶卡车与人工驾驶卡车之间的交互;深入刻画交互细节,展现行为的细微差别;分析行为背后的心理动机,如竞争性、风险偏好及对技术的适应或抗拒等;通过连贯的叙述分析不同驾驶风格对港口交通动态的影响,特别是在安全性和效率方面的表现。

(2) 利用思路链与自我矫正完善输出内容

提示词中主要包含以下内容:首先总结问题或任务,确保对问题的理解清晰;然后明确解决问题的步骤,并解释每一步的必要性及其对解决问题的贡献;在执行每一步时,详细说明思考过程,突出假设、选择或替代方案;提供基于分析的全面解决方案,确保直接回应问题;最后对解决方案进行自我反思,审视可能的缺陷或改进点,并总结从中学到的经验以及其在类似问题中的应用潜力。

(3) 整合问答过程,形成奖励函数最终输出

提示词中主要包含以下内容:结合聊天中提供的信息和讨论,制定用于仿真环境中对抗驾驶员训练的奖励函数,这些奖励函数将用于训练自动驾驶卡车。奖励函数分为2类:一类用于交互场景,一

类用于直道场景。每个奖励函数以数学形式表示，并附上详细描述。建模的场景应为一个采取正确行动可避免的安全关键场景。

本文采用的通用大语言模型以 GPT-4 turbo 为主，鉴于该模型使用的基础训练集英文数据的丰富度较中文更高，其对英文的理解和生成能力相对更加出色。因此，为了进一步优化中文问答的准确性和自然度，后续可以考虑结合其他专注于中文的大型模型，如文心一言、星火大模型，或者尝试使用新型模型（如 o1-mini、o1-preview）进行联合验证，这样的模型整合可以提升中文处理的性能，使其更适应中文语境和用户需求。

4.2 复杂交通数据处理与分析

复杂交通数据集通常包括大量帧采集的车辆轨迹采样点，如 NGSIM、INTERACTION 等数据集，详细描述了车辆在交通环境中的动态行为。数据集中交通数据具有高维度、复杂多样以及噪声多和异常值频发的特点，需要在预处理中对数据进行有效分类，以便后续算法训练使用。相比传统基于距离或密度的聚类算法，HDBSCAN^[55]算法采用层次密度估计，能自适应地识别数据中的各种局部密度，无须预定义簇数量。因此，本文采用 HDBSCAN 算法进行场景数据的聚类，并生成相应的场景类别标签。

(1) NGSIM 数据集场景聚类分析

NGSIM 数据集提供了 US-101 高速公路上的车

辆行为记录，本文主要关注与变道相关的行为场景。首先，通过监测车辆的车道 ID 变化，初步识别出变道行为。为了提高数据筛选的准确性，将频繁变化车辆 ID 的数据剔除，这类异常数据多由车辆长时间在车道线上行驶导致。然后，在变道轨迹的关键特征提取方面，采用车辆的历史、当前及未来车头朝向角序列，以及变道前后车辆的车头朝向角序列，共 20 维特征来描述该场景。最终，利用 HDBSCAN 聚类算法对这些特征进行聚类，准确地识别出向左和向右的变道场景。图 6（a）描述了 NGSIM 数据集中两类变道场景的分布。

(2) INTERACTION 数据集场景聚类分析

INTERACTION 数据集主要关注交叉路口车辆行为的交互，涵盖了多样、复杂和关键的驾驶场景，可用于运动预测、模仿学习以及决策和规划的验证。首先，从数据集中提取所有在路口两两交互的轨迹，并通过计算这些轨迹的交点和潜在碰撞时间（potential encounter time, PET）来识别包含强交互行为的场景。然后，获取每条轨迹的起始与终止点的朝向角，根据起始朝向角与终止朝向角之间的夹角将轨迹区分为直行、左转和右转 3 类场景。图 6（b）展示了 INTERACTION 数据集中 3 类路口交互场景的分布。最后，基于上述 3 类场景，根据主车与背景车轨迹在交汇点的朝向角以及轨迹分段的朝向角特征，利用 HDBSCAN 聚类方法将交叉路口场景分类为“主车直行/背景车直行”“主车左

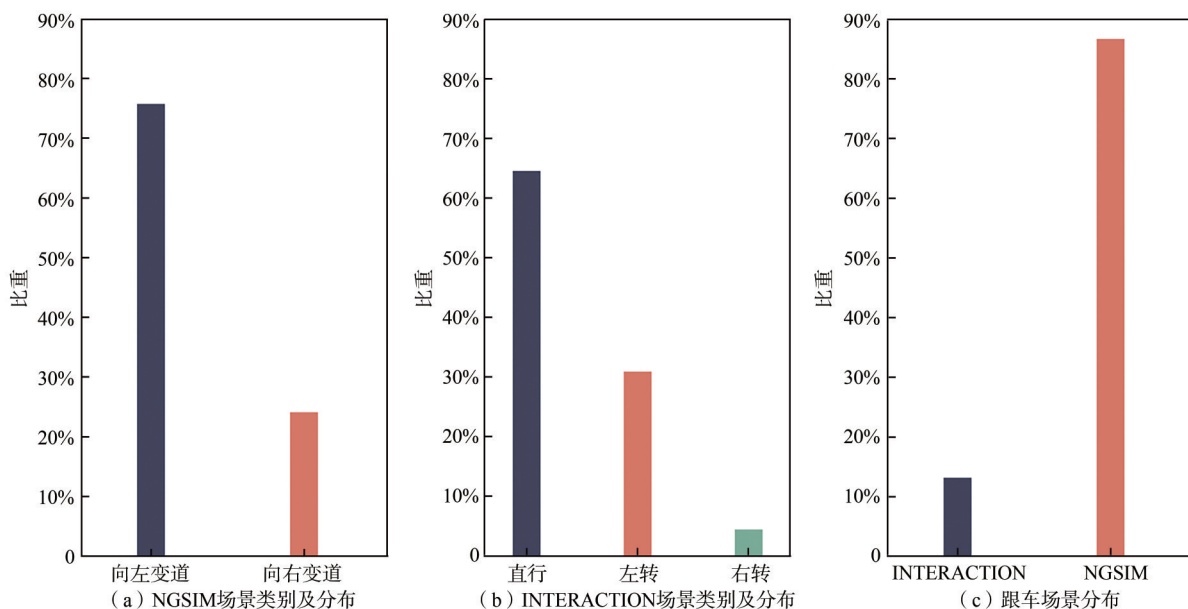


图6 不同场景类别及分布

转弯/背景车直行”和“主车右转弯/背景车直行”3类典型交互场景。此外,在处理数据时,发现数据集中存在一些轨迹的实际运动方向与数据中记录的朝向角不一致的现象,推测可能存在数据捕获或录入错误的问题,为确保分析的准确性,过滤这类轨迹。

(3) 跟车场景提取分析

为了丰富模型训练场景,本文基于车辆ID是否发生变化、前方50 m是否存在车辆以及轨迹持续时长等场景特征,从INTERACTION和NGSIM数据集中提取了770条和5 030条跟车轨迹,其比重如图6(c)所示。

本文共获取687个变道场景、5 689个路口交互场景以及5 800个跟车场景,可为模型提供更准确的参考数据集,促使模型学习更逼真的驾驶行为。

4.3 模型训练与验证

4.3.1 实验配置

在获取直道和路口典型场景之后,参考文献[26]对IDM与MOBIL模型进行参数校正,以在SUMO仿真器中建立逼真的交通流仿真环境,从而为对抗驾驶员模型训练及人机混行交互模拟提供基础。直道和路口场景IDM与MOBIL参数标定结果如表1所示。其中,基线模型以标定的IDM与MOBIL模型参数为基础,通过调整加速度、安全时距及最小加速度增益等参数,确定出一组无碰撞且变道率高的激进参数,用于与DK-Port方法进行对比。

表1 直道和路口场景IDM与MOBIL模型参数
标定结果

模型	参数	直道场景 (NGSIM)	路口场景 (INTERACTION)	基线 模型
IDM	目标速度/($\text{m}\cdot\text{s}^{-1}$)	15.0	9.0	15.0
	安全时距/s	1.0	1.0	0.5
	最大加速度/($\text{m}\cdot\text{s}^{-2}$)	3.0	3.0	5.0
	舒适减速度/($\text{m}\cdot\text{s}^{-2}$)	3.5	3.5	5.0
	安全车距/m	2.0	1.0	1.0
MOBIL	最小加速度增益/($\text{m}\cdot\text{s}^{-2}$)	0.2	0.2	0.1
	最大制动/($\text{m}\cdot\text{s}^{-2}$)	2.0	2.0	5.0
	礼貌系数	0.2	0.2	0.1

在模型训练与验证开始之前,选择两个最具代表性的路网结构来反映港口道路状况,如图7所示。图7(a)是青岛港前湾联合集装箱码头俯视图

图,前湾港是中国大陆最大的集装箱中转港,码头泊位效率连续数年排名全球第一,其吞吐量且运行效率高。图7(b)显示了对应的SUMO路网结构。图7(c)是CARLA开源自动驾驶仿真器的TOWN05地图。TOWN05地图包含多个无保护路口、直道以及复杂的交叉路段,这些特征与港口内集装箱码头的道路设计非常契合。在实验过程中,可以充分利用CARLA的动态交通生成、天气变化模拟以及传感器数据输出等特性,通过在TOWN05地图中选择多个具有代表性的区位,如主要进出口及关键转运节点等,全面测试和验证算法在不同复杂场景下的性能表现。

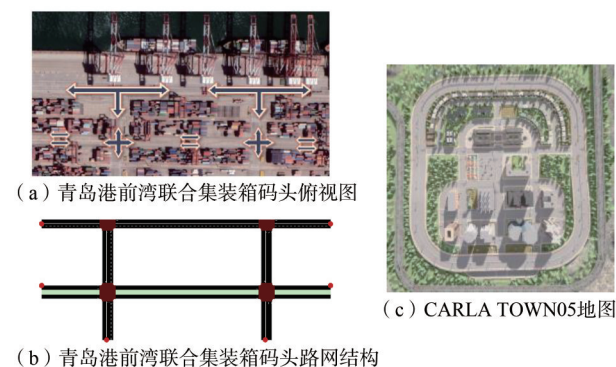


图7 模型训练与验证路网示意

对于实验运行环境,本文选用了SUMO 1.19.0作为主要仿真平台,配合TraCI 1.19.0接口,实现对复杂城市交通环境的精确模拟与动态控制。实验在配备Intel i7-13700KF CPU、Nvidia 4070 GPU、32 GB内存及12 GB显存的计算平台上进行执行。

4.3.2 模型训练

模型训练的主要参数设置如表2所示。在对抗模型训练过程中,针对直道场景与路口场景,其输入状态主要以对抗车为参考,包括对抗车和其他背景车的位置、速度、航向角差以及左右侧是否有车道、安全距离等信息;对抗模型输出行为由左变道、保持、右变道、加速、减速、左偏移、右偏移等动作组成。路口场景和直道场景模型训练过程中平均奖励随训练次数的变化如图8所示,2个场景均在训练大概1 500次后达到稳定状态,说明了训练参数设置的合理性以及模型参数已达到最优状态。

4.3.3 模型测试

为了验证DK-Port方法的有效性,针对直道和

表2 模型训练参数设置

参数	数值
训练次数	1 500
批次大小	1 024
学习率	0.001
探索率	0.8
折扣率	0.99
最大梯度	10

路口场景,利用基线模型和DK-Port对抗模型对每类场景分别进行100次仿真测试实验,评价指标为平均变道次数(即每类场景仿真测试的变道次数总和除以仿真次数)。在每次实验测试中,场景中初始车辆位置和速度进行随机初始化,最终以测试车的变道率、速度、加速度、TTC和PET等统计结果作为评价指标。

(1) 直道场景测试

针对直道场景,本文选择了2种典型场景作为测试场景,如图9所示。其中白色车辆为背景车,其通过标定的IDM和MOBIL模型控制,作为环境的相关干扰因素;蓝色车辆为测试车;绿色车为模仿人类驾驶的对抗车,其通过基线模型或DK-Port模型控制。为对抗车辆建模真实且激进的人类驾驶行为,可暴露测试车在决策控制方面的缺陷。

首先,利用场景平均变道次数来验证基线模型和DK-Port的有效性,自然的IDM和MOBIL模型(以下简称自然模型)、基线模型、DK-Port对抗模型在超车场景和切入场景下的平均变道次数如表3所示。在2种测试场景中,基线模型的平均变道次数相比自然模型均有所提升,表明了基线模型具有

一定的激进驾驶行为;DK-Port对抗模型的平均变道次数分别是基线模型的4倍和1.81倍,在两种直道场景中平均为基线模型的2.9倍,表明了所提方法能够充分模拟人类驾驶的激进行为,从而更有效地产生与测试车的交互行为。另外,3个模型在切入场景的平均变道次数均比超车场景的高,表明在超车场景中进行变道更为困难,与实际经验相符。

为了进一步验证DK-Port的有效性,本文统计了不同方法在超车场景和切入场景中测试车的TTC、速度和加速度,如图10和图11所示。通过计算可知,在2种场景下,相比基线模型,DK-Port对抗模型对测试车产生的TTC分布在 $[0, 5]$ 区间的占比分别为4.41%和4.15%,而基线模型仅为2.28%和1.56%;对于加速度制动占比,DK-Port分别达到27.24%和41.79%,显著高于模型算法的12.0%和9.56%。以上结果表明,DK-Port在生成危险场景方面具有显著优势。

(2) 路口场景测试

与直道场景类似,本文选择了直行交互和转弯交互这2种典型场景作为路口测试,如图12所示。图中绿色车辆仍为对抗车,不同之处在于黄色车辆表示对抗车的不同潜在初始位置,通过调整初始位置可体现复杂多样的路口交互场景。

在路口直行交互和路口转弯交互场景中,不同方法测试车PET、速度和加速度的测试结果如图13和图14所示。从速度及加速度分布可以看出,DK-Port能够促使测试车产生更多的减速制动行为,减少了加速行为。对图13和图14的分布结果进行计算可得,在2种路口场景下,DK-Port的加速度制

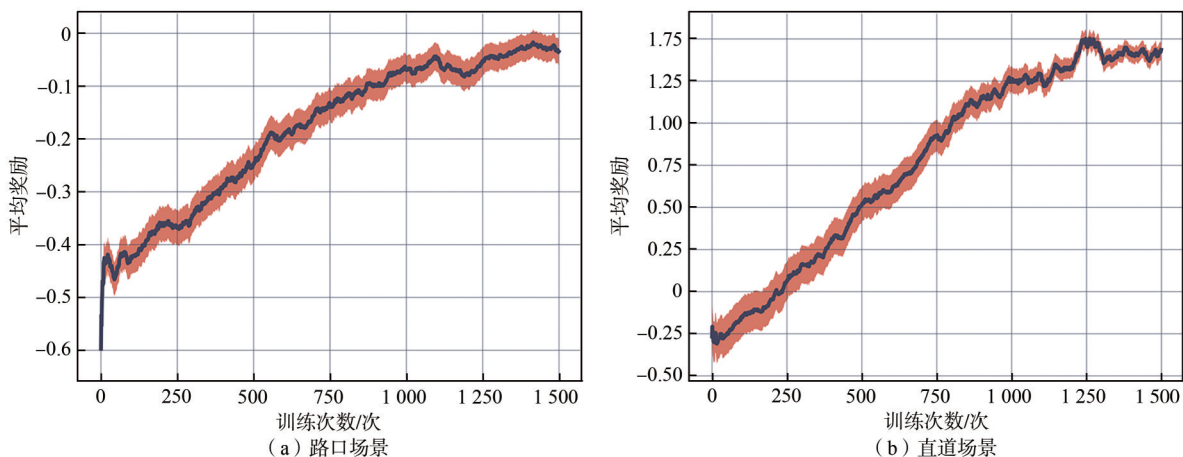


图8 模型训练过程中的平均奖励变化

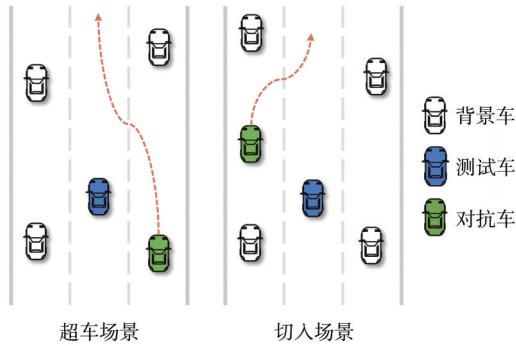


图9 直道测试场景

表3 不同方法在超车场景和切入场景下的平均变道次数

场景	平均变道次数		
	自然模型	基线模型	DK-port
超车场景	0.12	0.31	1.24
切入场景	0.72	0.73	1.32

动占比达到74.2%和85.1%，而基线模型的加速度制动占比仅为14.9%和51%，其制动占比相较于基线模型平均提升了46.7%，显示出更强的对抗效果。根据PET分布可知，DK-Port相比基线模型为测试车带来了更少的潜在碰撞时间，从而能够生成更多的安全关键场景。

(3) 关键交互场景示例

为了进一步展示DK-Port的实际效果，结合

CARLA 仿真平台，截取了TOWN05中部分路段的关键场景对抗过程，如图15所示，主要包括类人对抗模型的左/右侧直道超车场景以及路口抢占交汇点场景。对抗车（红色）能够根据测试车（蓝色）的当前状态持续调整驾驶行为，制造对测试车的潜在安全威胁，迫使其做出紧急响应动作，在仿真环境中完成人机混行驾驶行为的模拟，实现真实港口危险工况的仿真测试。

综上所述，在直道场景下，DK-Port模型的平均变道次数是基线模型的2.9倍，表明其能够模拟激进的驾驶行为，迫使测试车在复杂环境中频繁进行决策调整，暴露其控制缺陷。在路口场景中，DK-Port使测试车的加速度制动占比提高了46.7%，并减少了潜在碰撞时间，能够有效模拟高风险交互情境，增强测试车的应对能力。这些结果表明，DK-Port不仅能够提升仿真环境的真实性和测试的严苛性，还可以通过生成高风险场景增强自动驾驶系统的鲁棒性和安全性。

5 总结与展望

自动驾驶技术在港口的应用将为港口水平运输作业的效率提升和安全保障提供强有力的支撑。本文利用大语言模型，提出了将港口专家经验整合到强化学习模型设计中的方法，并根据港口特有的场

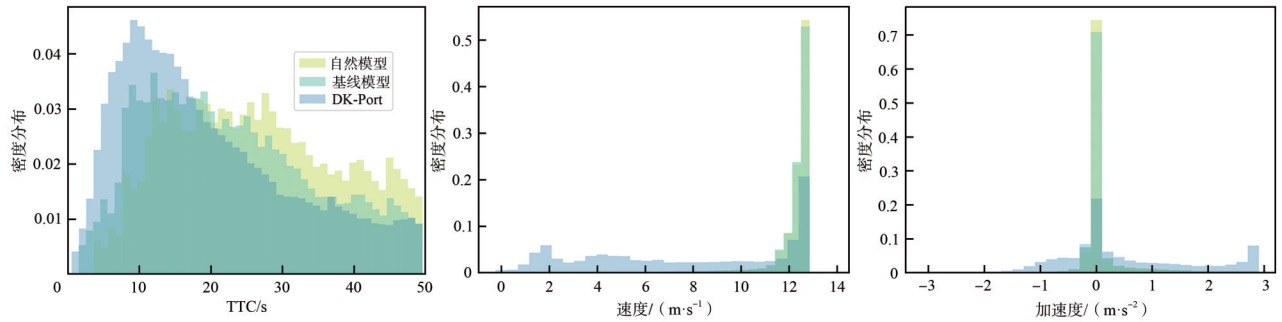


图10 超车场景测试车TTC、速度和加速度分布

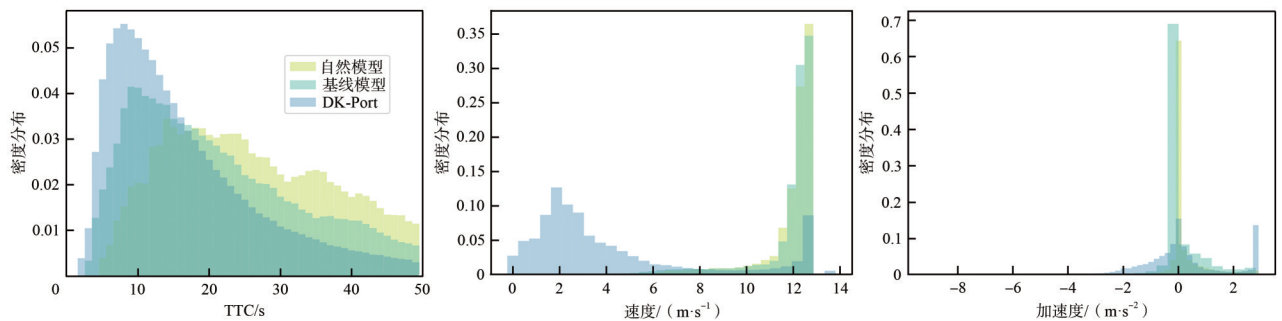


图11 切入场景测试车TTC、速度和加速度分布



图12 路口测试场景示意

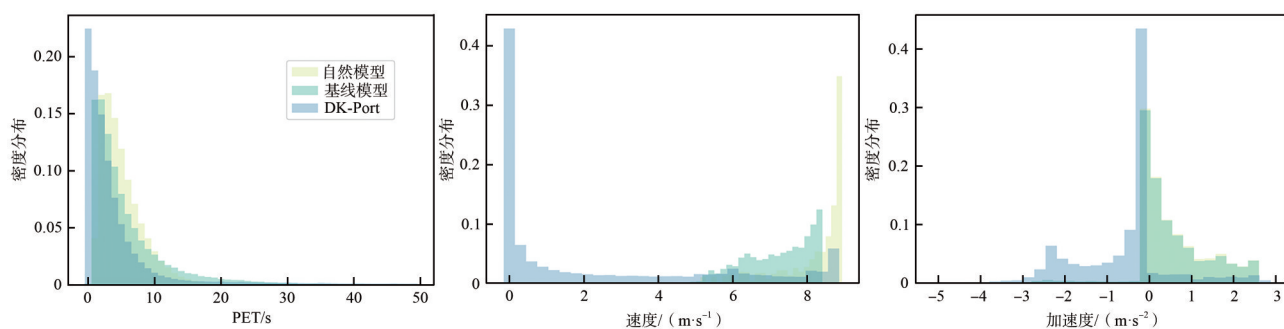


图13 路口直行交互场景测试车PET、速度和加速度分布

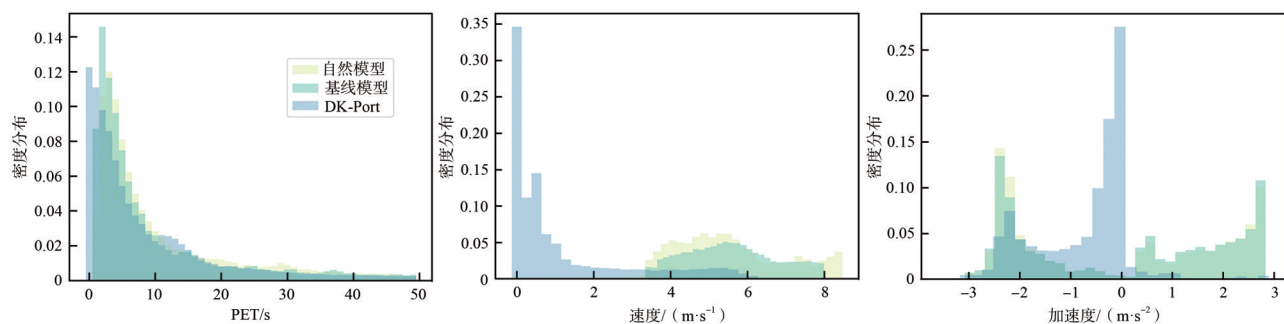


图14 路口转弯交互场景测试车PET、速度和加速度分布

景状态，优化了强化学习模型的关键过程和参数。实验表明，DK-Port模型的对抗车在TTC、PET、速度和加速度等主要指标上较传统方法展现出更强的对抗效果，能够充分暴露测试车的安全隐患，具有较强的实用价值。

在本研究的基础上，未来工作可以通过实地数据监测与分析来扩展数据集，并利用符合港口实际通行情况的真实数据集提升模型的表现力，从而使

训练出的对抗模型更好地适应不同港口的特性并应用于实际测试。此外，还可重点扩展仿真平台场景覆盖度，以更全面地模拟港口复杂多变的运行环境（如大雾、雨雪天气以及昼夜交替等场景），还可以补充多种类型的港口道路形式和交通规则，如不同类型的码头泊位、货车进出通道、集散中心及临时堆场等，以更真实地反映港口内部多样化的交通流和复杂的道路结构。

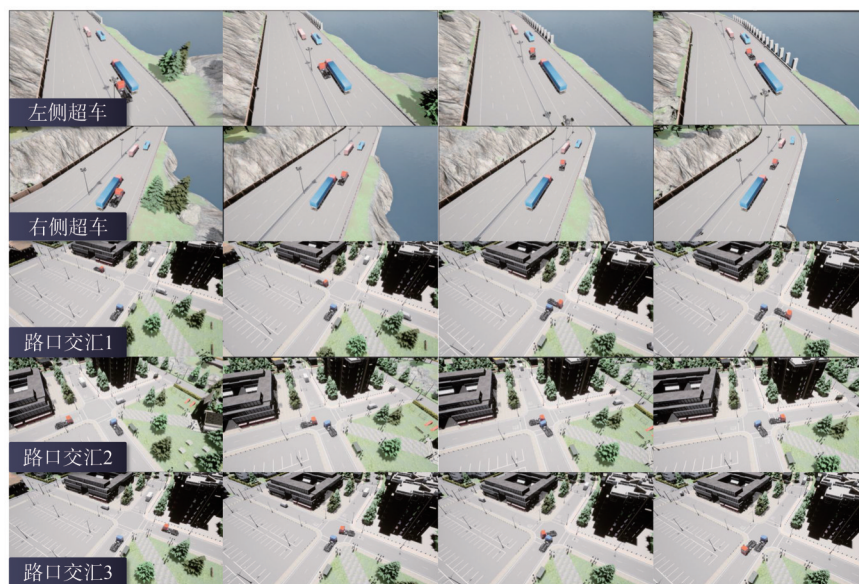


图15 CARLA TOWN05不同场景下DK-Port关键交互场景示例

参考文献:

- [1] The United Nations Conference on Trade and Development. Review of maritime transport 2023[R]. 2023.
- [2] 程国政. 我国港口集装箱运输发展现状与趋势[J]. 中国航务周刊, 2024(30): 45-47.
CHENG G Z. Development status and trends of container transportation in China's ports[J]. China Shipping Gazette, 2024(30): 45-47.
- [3] 柳长满, 张传捷, 陈微波, 等. 国内沿海自动化集装箱码头关键技术的发展趋势[J]. 中国港口, 2021(1): 17-23.
LIU C M, ZHANG C J, CHEN W B, et al. Key technology development trends of domestic coastal automated container terminal[J]. China Ports, 2021(1): 17-23.
- [4] CHU F, GAILUS S, LIU L, et al. The future of automated ports[R]. 2018.
- [5] RODRIGUE J P. The geography of transport systems[M]. New York: Routledge, 2020.
- [6] DE LA PEÑA ZARZUELO I, FREIRE SOEANE M J, LÓPEZ BERMÚDEZ B. Industry 4.0 in the port and maritime industry: a literature review[J]. Journal of Industrial Information Integration, 2020, 20: 100173.
- [7] CHEN X Q, LIU S H, ZHAO J S, et al. Autonomous port management based AGV path planning and optimization via an ensemble reinforcement learning framework[J]. Ocean & Coastal Management, 2024, 251: 107087.
- [8] LI X. Trends and prospects of port digital transformation[J]. Highlights in Business, Economics and Management, 2024, 24: 227-233.
- [9] 林维猛. 我国港口自动驾驶的趋势和挑战[J]. 中国港口, 2022(11): 1-5.
LIN W M. Trends and challenges of autonomous driving in China's ports [J]. China Ports, 2022(11): 1-5.
- [10] 王媛, 席芳, 汤伊琼. 新形势下我国智慧港口建设现状与发展趋势[J]. 水运工程, 2024(8): 224-228, 287.
WANG Y, XI F, TANG Y Q. Current situation and development trends of smart port construction in China[J]. Port & Waterway Engineering, 2024(8): 224-228, 287.
- [11] VAQUERO V, REPISO E, SANDELIU A, et al. Low cost, robust and real time system for detecting and tracking moving objects to automate cargo handling in port terminals[C]//Proceedings of Robot 2015: Second Iberian Robotics Conference: Advances in Robotics. New York: Springer International Publishing, 2016, 2: 491-502.
- [12] CHEN X Q, WANG Z C, HUA Q Z, et al. AI-empowered speed extraction via port-like videos for vehicular trajectory analysis[J]. IEEE Transactions on Intelligent Transportation Systems, 2023, 24(4): 4541-4552.
- [13] HU Y J, DONG L C, XU L. Multi-AGV dispatching and routing problem based on a three-stage decomposition method[J]. Mathematical Biosciences and Engineering, 2020, 17(5): 5150-5172.
- [14] HU H T, YANG X R, XIAO S C, et al. Anti-conflict AGV path planning in automated container terminals based on multi-agent reinforcement learning[J]. International Journal of Production Research, 2023: 61: 65-80.
- [15] QIN K, WANG B, ZHANG H, et al. Research on application and testing of autonomous driving in ports[J]. SAE Technical Paper, 2020.
- [16] KIM E, KIM Y, PARK J. The necessity of introducing autonomous trucks in logistics 4.0[J]. Sustainability, 2022, 14(7): 3978.
- [17] CHOTTANI A, HASTINGS G, MURNANE J, et al. Distraction or disruption? Autonomous trucks gain ground in US logistics[R]. 2018
- [18] MOORE D, REBECCA C, MICHAEL S, et al. Defense against the dark cars: design principles for grieving of autonomous vehicles[C]//Proceedings of the 2020 ACM/IEEE International Conference on Human-Robot Interaction. [S.l.:s.n.], 2020: 201-209.
- [19] LEE Y C, MOMEN A, LAFRENIERE J. Attributions of social interactions: driving among self-driving vs. conventional vehicles[J]. Technology in Society, 2021, 66: 101631.
- [20] YAM K C, TANG P M, JACKSON J C, et al. The rise of robots increases job insecurity and maladaptive workplace behaviors: multi-method evidence[J]. Journal of Applied Psychology, 2023, 108(5): 850-870.
- [21] International Transport Forum. Container port automation: impacts and implications[R]. 2021.

- [22] 和福建, 马文博, 田晓笛. 港口自动驾驶现状及典型场景测试技术研究[J]. 汽车电器, 2022(8): 1-2, 5.
HE F J, MA W B, TIAN X D. Research on the status of port autonomous driving and testing technology of typical scenes[J]. Auto Electric Parts, 2022(8): 1-2, 5.
- [23] 王飞跃, 王雨桐. 数字科学家与平行科学: AI4S 和 S4AI 的本源与目标[J]. 中国科学院院刊, 2024, 39(1): 27-33.
WANG F Y, WANG Y T. Digital scientists and parallel sciences: the origin and goal of AI for science and science for AI[J]. Bulletin of Chinese Academy of Sciences, 2024, 39(1): 27-33.
- [24] WANG X, HUANG J, TIAN Y L, et al. Parallel driving with big models and foundation intelligence in cyber-physical-social spaces[J]. Research. 2024, 7(3): 1-16.
- [25] YAN X T, ZOU Z X, FENG S, et al. Learning naturalistic driving environment with statistical realism[J]. Nature Communications, 2023, 14: 2037.
- [26] CHEN D, ZHU M X, YANG H, et al. Data-driven traffic simulation: a comprehensive review[J]. IEEE Transactions on Intelligent Vehicles, 2024, 9(4): 4730-4748.
- [27] HAO K K, CUI W, LUO Y G, et al. Adversarial safety-critical scenario generation using naturalistic human driving priors[J]. IEEE Transactions on Intelligent Vehicles, 2024: 1-16.
- [28] HAO K K, LIU L, CUI W, et al. Bridging data-driven and knowledge-driven approaches for safety-critical scenario generation in automated vehicle validation[EB]. arXiv preprint, 2023: 2311.109371.
- [29] 王祥, 谭国真. 基于知识与大语言模型的高速环境自动驾驶决策研究[J]. 系统仿真学报, 已录用, doi: 10.16182/j.issn1004731x.joss.24-0065.
WANG X, TAN G Z. Research on decision-making of autonomous driving in highway environment based on knowledge and large language model[J]. Journal of System Simulation, accepted, doi: 10.16182/j.issn1004731x.joss.24-0065.
- [30] 时高松, 赵清海, 董鑫, 等. 基于 PPO 算法的自动驾驶人机交互式强化学习方法[J]. 计算机应用研究, 2024, 41(9): 2732-2736.
SHI G S, ZHAO Q H, DONG X, et al. Human-machine interactive reinforcement learning method for autonomous driving based on PPO algorithm[J]. Application Research of Computers, 2024, 41(9): 2732-2736.
- [31] SHINN N, CASSANO F, GOPINATH A, et al. Reflexion: language agents with verbal reinforcement learning[EB]. arXiv preprint, 2023, arXiv: 2303.11366v4.
- [32] MADAAN A, TANDON N, GUPTA P, et al. Self-refine: iterative refinement with self-feedback[EB]. arXiv preprint, 2023, arXiv: 2303.17651v2.
- [33] XIE T B, ZHAO S H, WU C H, et al. Text2Reward: reward shaping with language models for reinforcement learning[C]//Proceedings of the Twelfth International Conference on Learning Representations. [S.l.:s.n.], 2024.
- [34] KWON M, XIE S M, BULLARD K, et al. Reward design with language models[EB]. arXiv preprint, 2023, arXiv: 2303.00001.
- [35] LI Y J, CHOI D, CHUNG J, et al. Competition-level code generation with AlphaCode[J]. Science, 2022, 378(6624): 1092-1097.
- [36] XUE B Z, PARK N K. A study on the efficiency of container terminals in Korea and China[J]. The Asian Journal of Shipping and Logistics, 2016, 32(4): 213-220.
- [37] COIFMAN B, Li L Z. A critical evaluation of the next generation simulation (NGSIM) vehicle trajectory dataset[J]. Transportation Research Part B: Methodological, 2017, 105: 362-377.
- [38] ZHAN W, SUN L T, WANG D, et al. INTERACTION dataset: an international, adversarial and cooperative motion dataset in interactive driving scenarios with semantic maps[EB]. arXiv preprint, 2019, arXiv: 1910.03088.
- [39] HO J, ERMON S. Generative adversarial imitation learning[EB]. arXiv preprint, 2016, arXiv: 1606.03476.
- [40] AMODEI D, OLAH C, STEINHARDT J, et al. Concrete problems in AI safety[EB]. arXiv preprint, 2016, arXiv:1606.06565.
- [41] ARORA S, DOSHI P. A survey of inverse reinforcement learning: challenges, methods and progress[J]. Artificial Intelligence, 2021, 297: 103500.
- [42] HUANG L, YU W J, MA W T, et al. A survey on hallucination in large language models: principles, taxonomy, challenges, and open questions[EB]. arXiv preprint, 2023, arXiv: 2311.05232v2.
- [43] 沈甜雨, 李志伟, 范丽丽, 等. 具身智能驾驶: 概念、方法、现状与展望[J]. 智能科学与技术学报, 2024, 6(1): 17-32.
SHEN T Y, LI Z W, FAN L L, et al. Embodied intelligent driving: concept, methods, the state of the art and beyond[J]. Chinese Journal of Intelligent Science and Technology, 2024, 6(1): 17-32.
- [44] SPECTOR P E. Do not cross me: optimizing the use of cross-sectional designs[J]. Journal of Business and Psychology, 2019, 34(2): 125-137.
- [45] JOSHI A, KALE S, CHANDEL S, et al. Likert scale: explored and explained[J]. British Journal of Applied Science & Technology, 2015, 7(4): 396-403.
- [46] KOJIMA T, GU S S, REID M, et al. Large language models are zero-shot reasoners[J]. Advances in Neural Information Processing Systems, 2022, 35: 22199-22213.
- [47] ZHOU X, KIM K, XU B W, et al. The devil is in the tails: how long-tailed code distributions impact large language models[C]//Proceedings of the 2023 38th IEEE/ACM International Conference on Automated Software Engineering (ASE). Piscataway: IEEE Press, 2023: 40-52.
- [48] JIANG Z, XU F F, GAO L, et al. Active retrieval augmented generation[EB]. arXiv preprint, 2023, arXiv:2305.06983.
- [49] XIE Z K. Order matters in hallucination: reasoning order as benchmark and reflexive prompting for large-language-models[EB]. arXiv preprint, 2024, arXiv: 2408.05093.
- [50] SKALSE J, HOWE N, KRASHENINNIKOV D, et al. Defining and characterizing reward gaming[J]. Advances in Neural Information Processing Systems, 2022, 35: 9460-9471.
- [51] SCHULMAN J, WOLSKI F, DHARIWAL P, et al. Proximal policy optimization algorithms[EB]. arXiv preprint, 2017, arXiv: 1707.06347.
- [52] LEURENT E. An environment for autonomous driving decision-making[R]. 2018.
- [53] KESTING A, TREIBER M, HELBING D. Enhanced intelligent driver model to access the impact of driving strategies on traffic capacity[J]. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 2010, 368(1928): 4585-4605.
- [54] KESTING A, TREIBER M, HELBING D. General lane-changing model MOBIL for car-following models[J]. Transportation Research Record: Journal of the Transportation Research Board, 2007, 1999(1): 86-94.
- [55] MCINNES L, HEALY J, ASTELS S. Hdbscan: Hierarchical density based clustering[J]. The Journal of Open Source Software, 2017, 2(11): 205.

[作者简介]



娄云洁（1971- ），男，青岛前湾联合集装箱码头有限责任公司总经理，主要研究方向为集装箱码头的运营和管理。



滕出（2001- ），男，青岛前湾联合集装箱码头有限责任公司自动化研创中心数字化工程师，主要研究方向为自动化轨道吊、无人集卡技术等。



艾明飞（1980- ），男，青岛前湾联合集装箱码头有限责任公司副总经理，主要研究方向为码头大型设备管理、绿色港口建设等。



王江成（1997- ），男，安徽深信科创信息技术有限公司算法工程师，主要研究方向为自动驾驶仿真测试、对抗场景生成、智能体训练等。



庄术洁（1992- ），男，青岛前湾联合集装箱码头有限责任公司自动化研创中心数字化部主任，主要研究方向为智慧港口建设、智能驾驶及自动化等。



沈甜雨（1996- ），女，博士，北京化工大学信息科学与技术学院副教授，主要研究方向为机器视觉、智能感知、智能机器人系统等。



于海（1972- ），男，青岛前湾联合集装箱码头有限责任公司自动化研创中心主任，主要研究方向为传统集装箱码头转型升级全场景自动化、无人化关键技术研究与应用等。



郝坤坤（1991- ），男，安徽深信科创信息技术有限公司高级算法工程师，主要研究方向为自动驾驶预期功能安全场景生成、对抗场景生成、仿真交通流构建等。



王鑫（1991- ），男，青岛前湾联合集装箱码头有限责任公司自动化研创中心智能设备部主任，主要研究方向为电控系统。



崔文（1989- ），男，博士，安徽深信科创信息技术有限公司博士后，主要研究方向为智能网联技术，如安全关键测试场景生成、车辆通信网络优化和车辆状态感知等方向。