

中文引用格式:周扬,陈运星,吴玲.用于自动驾驶测试的车辆危险切入策略研究[J].中国安全科学学报,2025,35(1):112-119.

英文引用格式:ZHOU Yang, CHEN Yunxing, WU Ling. Research on vehicle hazardous cut-in strategy used in autonomous driving test[J]. China Safety Science Journal, 2025, 35(1): 112-119.

用于自动驾驶测试的车辆危险切入策略研究*

周扬^{1,2}副教授,陈运星^{**2,3}副教授,吴玲¹副教授

(1 西安航空学院 车辆工程学院,陕西 西安 710077;2 湖北文理学院 纯电动汽车动力系统设计与测试湖北省重点实验室,湖北 襄阳 441053;3 湖北文理学院 汽车与交通工程学院,湖北 襄阳 441053)

中图分类号:X951;U467.13

文献标志码:A

DOI: 10.16265/j.cnki.issn1003-3033.2025.01.0540

基金项目:国家自然科学基金资助(51908054);陕西省科技厅自然科学基金基础研究计划项目(2024JC-YBMS-301);湖北省技术创新计划科技重大项目(2024BAA011);纯电动汽车动力系统设计与测试湖北省重点实验室开放基金资助(ZDSYS202310)。

【摘要】 为提高车辆切入测试场景中交通车的交互能力,提出一种基于深度强化学习方法的车辆危险切入策略设计方法。首先,基于可扩展多智能体强化学习培训学校(SMARTS)仿真平台构建仿真环境;然后,采用双延迟深度确定性策略梯度算法(TD3)训练智能体危险切入随机选定的目标车辆,将该算法与近端策略优化算法(PPO)和深度确定性策略梯度算法(DDPG)进行对比,在7种不同车辆密度的场景中测试训练后的模型;最后,构建多智能体测试环境,将所训练模型用于智能驾驶策略的验证。结果表明:模型在训练中的危险切入成功率达80.35%,优于2种对比方法;在模型测试中,除2700辆/h测试场景外,该模型在另外3个未在训练中使用的测试场景均达到80%以上的危险切入成功率,显示出良好的泛化能力。同时,切入时刻与目标车的碰撞时间值显示95%集中在0~6s,取值在(0,2]、(2,4]和(4,6]s的占比分别为60%、30%和5%,可覆盖具有不同碰撞风险的测试工况。在智能驾驶策略验证中,采用所训练模型控制的交通车能主动切入至待测车辆前方,使待测车辆面临追尾风险,有助于发现智能驾驶策略的安全隐患。

【关键词】 自动驾驶; 车辆危险切入; 虚拟测试; 危险场景; 强化学习

Research on vehicle hazardous cut-in strategy used in autonomous driving test

ZHOU Yang^{1,2}, CHEN Yunxing^{2,3}, WU Ling¹

(1 School of Vehicle Engineering, Xi'an Aeronautical Institute, Xi'an Shaanxi 710077, China;

2 Hubei Key Laboratory of Power System Design and Test for Electrical Vehicle, Hubei

University of Arts and Science, Xiangyang Hubei 441053, China; 3 School of Automotive

and Traffic Engineering, Hubei University of Arts and Science, Xiangyang Hubei 441053, China)

Abstract: To improve the interaction ability of traffic vehicles in the cut-in scenario, a method for constructing a vehicle hazardous cut-in strategy based on deep reinforcement learning was proposed. Firstly, a simulated environment was built based on scalable multi-agent reinforcement learning training

* 文章编号:1003-3033(2025)01-0112-08; 收稿日期:2024-08-11; 修稿日期:2024-10-20

** 通信作者:陈运星(1987—),男,湖北荆门人,博士,副教授,主要从事驾驶行为感知、智能驾驶技术等方面的研究。E-mail:chenyunxing@chd.edu.cn。

school (SMARTS) simulation platform. Then, twin delayed deep deterministic policy gradients (TD3) algorithm was adopted to train an agent to cut in a randomly chosen target vehicle hazardously. The algorithm was compared with proximal policy optimization (PPO) and deep deterministic policy gradient (DDPG) algorithms. The trained model was tested in seven different scenarios with varying traffic densities. Finally, a multi-agent testing environment was built, and the trained model was applied to validate intelligent driving strategies. The results show that the success rate of hazardous cut-ins reaches 80.35% in model training with TD3 algorithm, outperforming both comparative methods. In model testing, except for the 2 700 vehicle/h test scenario, the model achieves a hazardous cut-in success rate of over 80% in the other three test scenarios that were not used in training, demonstrating good generalization ability. Meanwhile, the time to collision values between the ego vehicle and the target vehicle at the moment of lane changes are concentrated within the range of 0 to 6 seconds, with 95% falling within this bracket. The proportions of time to collision values in the intervals of (0,2], (2,4], (4,6]s are 60%, 30%, and 5% respectively, covering test conditions with different collision risk. In the validation of intelligent driving strategies, the traffic vehicle controlled by the trained model can actively perform cut-ins in front of the test vehicles, exposing it to the risk of a rear-end collision and helping in identifying safety vulnerabilities in intelligent driving strategies.

Keywords: autonomous driving; vehicle hazardous cut-in; virtual tests; hazardous scenarios; reinforcement learning

0 引言

自动驾驶汽车 (Autonomous Vehicles, AVs) 为解决传统汽车存在的交通拥堵、事故频发和提升驾驶舒适度提供了新的解决方案,成为近年来的研究焦点^[1]。在实际部署 AVs 前,为确保其安全性而进行以安全性为主的性能测试至关重要。在开放道路环境下测试是 AVs 测试的常用方法,由于开放道路环境对 AVs 安全性产生挑战的危险场景极为罕见,因此,AVs 需行驶数十亿英里才能证明其安全性是否低于人类驾驶。从测试成本、效率以及安全性等多方面考虑,开放道路测试已无法满足 AVs 技术研发和迭代的要求^[2]。虚拟测试具有场景配置灵活、测试成本低及重复性强的优势,在虚拟条件下可构建开放道路中小概率发生的危险场景,已成为 AVs 测试的主要方式^[3]。在虚拟测试研究中,基于危险场景的 AVs 虚拟测试有利于缩短测试周期,实现对 AVs 的加速测试。因此,构建对 AVs 具有挑战的危险测试场景已成为重要的研究内容。

在相关研究中,郭柏仓等^[4]提取了 59 例实车变道切入数据片段,采用危险感知系数表征场景的危险程度,分析提取了与场景危险程度显著相关的场景要素,基于 K-means 算法聚类得到 4 类城市道路车辆变道切入场景。赵祥模等^[5]提出了一种紧急变道轨迹对抗生成方法,基于公开数据集中的紧急变道数据进行训练,构建了包含 5 万条变道轨迹的

车辆危险变道测试场景库。朱冰等^[6]提出了一种基于危险边界搜索的 AVs 加速测试方法,以车辆切入场景为例,采用所提出方法可得到所有碰撞场景的危险边界,使测试时长缩短 73%。为使所构建的危险测试场景中交通车具备交互能力,相关研究开始采用机器学习方法构建具有连续决策能力的交通车模型,生成危险测试场景测试 AVs 的连续决策能力。其中,FENG Shuo 等^[7]设计了用于自动驾驶虚拟测试的自然-对抗驾驶环境 (Naturalistic and Adversarial Driving Environment, NADE),在构建的自然驾驶环境训练场景中交通车在特定时刻执行特定动作,生成自然而具有挑战的测试环境,有效降低 AVs 的测试里程。FENG Shuo 等^[8]设计了自演绎的 AVs 测试环境,环境中的背景车辆采用提出的密集深度强化学习算法训练,该算法保留了数据中的安全关键事件信息,起到了 AVs 安全性加速测试的作用。SUN Haowei 等^[9]提出了一种用于生成自动驾驶测试边界用例的统一框架,采用深度强化学习方法构建场景中背景车辆的行驶策略,背景车辆采取攻击性驾驶风格与被测车辆进行互动,以产生更多的边界测试场景。李江坤等^[10]提出了一种 AVs 加速测试的危险场景强化生成方法,利用神经网络来建立场景的动力学系统模型,通过强化学习算法构建场景控制器,生成了具有交互博弈的危险测试场景。以上研究主要基于采样的方法生成车辆切入危险测试场景,场景中交通车采用预设的行驶轨迹,存

在交通车交互能力不足的缺陷。为使交通车具备交互能力,相关研究开始采用强化学习方法构建交通车控制策略,然而,目前对于在车辆切入场景中构建兼具危险切入及交互能力的交通车控制策略的研究仍较为缺乏。

鉴于此,笔者拟采用深度强化学习方法构建具备交互能力的车辆危险切入策略,通过可扩展多智能体强化学习培训学校平台建立车辆切入仿真环境,基于双延迟深度确定性策略梯度(Twin Delayed deep deterministic policy gradients, TD3)算法构建车辆危险切入模型,使用训练后的模型测试被测系统的安全性,以期为 AVs 危险测试场景中交通车控制模型的构建提供参考。

1 基于 TD3 的车辆危险切入策略

建立具有较高逼真度的仿真环境是模型训练的前提,提高仿真环境的多样性是确保模型能适应环境变化的基础。

1.1 车辆切入仿真环境搭建

选择基于 SMARTS 平台^[11]构建车辆切入仿真环境。SMARTS 通过 PyBullet 物理引擎提供真实的车辆运动仿真,基于开源交通仿真软件——城市交通能力仿真(Simulation of Urban Mobility, SUMO)来生成仿真道路和交通流,SMARTS 还提供基于网页的三维可视化,方便用户查看仿真结果。构建如图 1 所示的车辆切入环境,车道长度为 200 m,场景中背景车辆车速通过 SUMO 控制。具体而言,背景车辆会综合车道限速值、最高期望车速及 SUMO 内置的跟车模型所输出车速值,选择其中的最低值作为目标车速。自行车由强化学习智能体控制。

交通流是场景的核心要素之一^[12],为增加场景中交通流的多样性,采取如下措施:①参考胡祥旺等^[13]的研究,为更加接近真实交通环境中不断变化的交通流密度,生成 3 种不同流量水平的训练场景,分别对应低、中、高流量。设定这 3 种场景在单车道上的车流量分别为 1 200、1 800 和 2 400 辆/h,随机生成场景中交通车的初始间距、车速及位置,3 种场景中共包含 1 200 种不同初始化的交通流。②在智能体与环境交互的每一回合开始阶段,随机选择交通流中的任意一辆交通车作为目标车;自行车从起始位置出发,驶入目标车所在车道并位于目标车辆前方时,认定自行车完成切入。根据周文帅等^[14]的研究,在车辆切入场景中,若自行车在切入

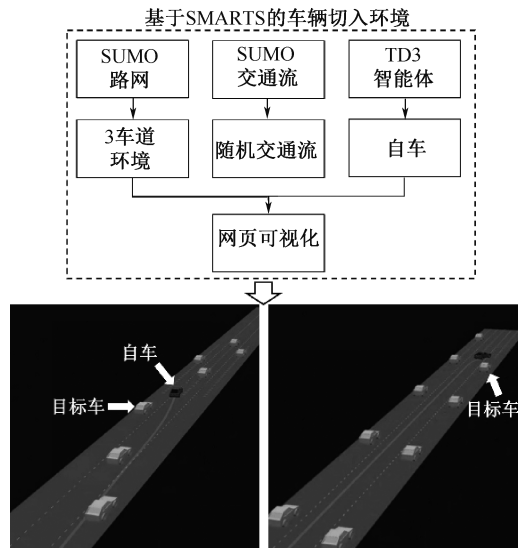


图 1 基于 SMARTS 的车辆切入环境

Fig. 1 Vehicle cut-in environment based on SMARTS

时刻与目标车的碰撞时间为 0~6 s,则定义为高风险切入场景。因此,当自行车成功切入目标车前方,且切入时刻的碰撞时间值为 0~6 s,则认为自行车危险切入成功。

1.2 车辆危险切入策略设计

强化学习算法通过智能体与环境互动以及接收环境反馈来进行试错学习,从而赋予智能体与环境的交互能力。通过设计特定的奖励函数,可引导智能体展现出期望的行为。因此,为构建同时具备危险切入及交互能力的交通车控制策略,选择深度强化学习算法中代表方法 TD3 来构建车辆危险切入策略。

1) TD3 算法。TD3 算法作为深度确定性策略梯度算法(Deep Deterministic Policy Gradient, DDPG)算法的改良版本,包括双 Q 网络、延迟策略更新和目标策略平滑的改进措施,有效改善了 DDPG 算法的性能和稳定性^[15]。为加强智能体对环境的探索,对 TD3 模型的主策略网络所输出动作 a_t 添加从正态分布中采样的噪声,见下式:

$$a_t = \text{clip}[\pi_\omega(s_t) + \mathcal{N}(\mu, \sigma), a_{\min}, a_{\max}] \quad (1)$$

式中: t 为时刻; π_ω 为 TD3 策略模型, ω 为参数; $\mathcal{N}(\mu, \sigma)$ 表示均值 μ 、方差 σ 的正态分布,二者取值分别为 0 和 0.1; s_t 为 t 时刻的状态; a_{\min} 和 a_{\max} 分别为动作取值的下限和上限; clip 函数截取了添加噪声后的动作,使其处于 $[a_{\min}, a_{\max}]$ 的区间。

2) 状态表示和动作空间。所采用的状态表示由 3 部分构成,见表 1。第 1 部分为自行车特征,共 4 项;第 2 部分是附近车辆特征,为提取与自行车决策

最相关的信息,过滤 SMARTS 所提供的周围车辆信息,仅提取与自车距离最近的 6 辆车(图 2 中除自车外的其他车辆),共计 18 项特征;第 3 部分是目标车与目标点特征,由于与自车距离最近的 6 辆附近车辆中可能并不包含目标车,因此,需额外提取目标车信息,目标点(用星形标志表示)是指目标车所在车道正前方 10 m 的一个预设点。

表 1 状态表示

Table 1 State representation

类型	特征描述
自车	纵向车速、横向车速、朝向、车道偏移距离
附近车辆 (i 为车辆数)	距离 d^i 、纵向距离 d_x^i 、横向距离 d_y^i
目标点与目标车 (T 为目标车, m 为目标点)	与目标点的距离为 d^m , 自车与目标车的纵向距离 d_x^T , 横向距离 d_y^T , 相对车速 Δv , 碰撞时间 $o = \frac{d_x^{T2} + d_y^{T2}}{\Delta v}$

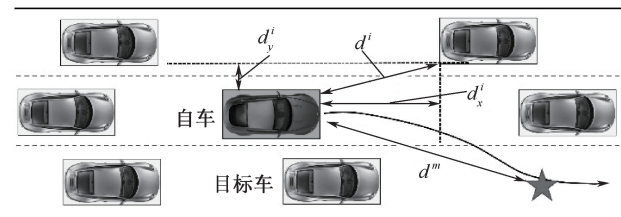


图 2 附近车辆的提取与目标点

Fig. 2 Extraction of adjacent vehicles and target point

由于 o 的计算结果取值范围较广,为减小模型训练中参数的搜索空间,将 o 计算值限制在 $(0, 20]$ s 范围。采用连续动作空间,TD3 模型输出每个时刻车辆的加速度和横摆角速度值。为将模型的输出动作值限制在合理范围,将加速度范围设为 $[-5, 5]$ m/s^2 ,将横摆角速度范围设为 $[-0.5, 0.5]$ rad/s ,当动作值超过设定区间的上限或下限值时,则取对应的上、下限值。

3) 奖励函数。奖励函数是强化学习的核心,设计合适的奖励可加快模型收敛并提升模型的性能表现。所设计奖励函数包括危险切入即时奖励、引导自车危险切入奖励和不当行为惩罚 3 部分。

危险切入即时奖励。根据 1.1 节中所定义的危险切入判定标准,记自车切入目标车所在车道且位于目标车前方时刻为 t_c ,若该时刻自车与目标车的碰撞时间 $o_{t_c} \in (0, 6]$ s,则给予智能体较大奖励;若

o_{t_c} 未落入上述区间,则依据自车切入时刻与目标车的相对车速 Δv_{t_c} 确定奖励值, Δv_{t_c} 越大,目标车追尾自车的风险则更高,奖励值相应增加,危险切入即时奖励值 r_{dc} 见下式:

$$r_{dc} = \begin{cases} 1000 & 0 < o_{t_c} \leq 6 \\ 10 + \Delta v_{t_c} & \text{其他} \end{cases} \quad (2)$$

引导自车危险切入奖励。为引导智能体控制自车行驶至目标车前方,在目标车所在车道正前方距离 10 m 处设置一个虚拟的目标点,根据自车与该目标点的距离 d^m 设计如下的距离奖励值 r_d 。随着 d^m 减小, r_d 增大,即自车与目标点距离越近,值越高。因此, r_d 鼓励智能体控制自车行驶至该目标点,到达目标车前方。

$$r_d = 1 - \left(\frac{d^m}{n_1} \right)^{0.4} \quad (3)$$

式中 n_1 为对 d^m 进行归一化所采用的系数,取 50。

由于强化学习训练是以最大化累积奖励为目标,仅设置距离奖励会导致智能体以尽可能快的车速行驶至目标车前方以收集更多奖励,导致自车在切入至目标车前方时容易出现车速大于目标车车速的情况,不能满足文中所定义危险切入的要求。为解决该问题,定义碰撞时间折扣系数 D ,当自车与目标点的距离小于 5 m 时,按照下式计算其值,可以看出, o 值越小则 D 越大。当自车与目标点较远时,将该值定为 1。

$$D = 1 - \left(\frac{o}{n_2} \right)^{0.4} \quad (4)$$

式中 n_2 为对 o 进行归一化的系数,取 20。若 o 为负值,则取 20 代入式(4)计算。

最后,通过下式计算得到引导自车危险切入奖励项 r_{yd} 。

$$r_{yd} = r_d \cdot D \quad (5)$$

不当行为惩罚。为防止自车在行驶过程中出现如驶出道路、倒车或反向行驶、停止不动及主动与交通车碰撞 4 种不当行为,将这些行为的奖励值设为 -10,作为不当行为惩罚项 r_p 。由于自车在危险切入时很容易造成目标车追尾,为不影响智能体对环境的探索,当自车被交通车追尾时不予惩罚。

所采用的奖励函数为上述 3 部分奖励的线性叠加,见下式:

$$R = \theta_1 \cdot r_{dc} + \theta_2 \cdot r_{yd} + \theta_3 \cdot r_p \quad (6)$$

式中 θ_1 、 θ_2 和 θ_3 为 r_{dc} 、 r_{yd} 和 r_p 等 3 种奖励对应的权重系数。

2 车辆危险切入策略模型

2.1 训练参数设置及对比方法选择

参照原算法,所构建 TD3 模型中的 6 个神经网络模型均采用 2 层全连接神经网络结构,神经元数量分别为 400 和 300,模型训练所采用的超参取值见表 2,其余参数设置均与原算法保持一致。

表 2 TD3 超参取值

Table 2 Hyperparameters of TD3

参数	取值	参数	取值	参数	取值
学习率	3×10^{-4}	经验回放池大小	1×10^6	软更新率	0.005
批大小	100	折扣因子	0.99	总步数	5×10^5

由于 TD3 算法是针对 DDPG 算法的改进,故将 DDPG 算法作为比较方法。DDPG 模型中,神经网络的结构及神经元数量、模型训练所采用超参取值均与 TD3 算法保持一致。

将基于策略梯度的深度强化学习代表算法近端策略优化算法 (Proximal Policy Optimization, PPO) 作为比较方法,PPO 算法通过限制策略更新的幅度来稳定训练过程。PPO 模型中的 Actor 和 Critic 网络结构与 TD3 保持一致,模型训练所采用的学习率和总训练步数与 TD3 的设置相同,其他超参的选择参照了 ZHOU Yang 等^[16]的研究。

2.2 模型训练及测试结果分析

为确定奖励权重系数 θ_1 、 θ_2 和 θ_3 的最佳取值,在 3 种奖励权重系数的不同取值下训练模型。在训练过程中,统计每个回合中自车完成危险切入的次数,并将成功危险切入的回合数除以记录间隔内的总回合数,以此得到模型在训练过程中的危险切入成功率,作为模型性能的评价指标。在选取 8 种不同奖励权重系数组合训练后发现,3 种奖励权重系数的取值均为 1/3 时,危险切入成功率达到了较高水平。由于奖励值的大小不会对模型性能产生影响,为便于计算,将 3 种奖励权重系数的取值均设为 1。

1) 模型训练结果。在训练过程中记录了智能体在每个回合累积奖励平均值的变化,结果如图 3 所示。可以发现,在训练过程中,TD3 智能体每回合累积的平均奖励值呈上升趋势。当累积步数达到 3.5×10^5 步时,累积奖励的平均值趋于平稳,表明 TD3 模型已经收敛。

图 4 为 TD3 模型在训练过程中危险切入成功率的变化,为说明本文设计的奖励函数对智能体完

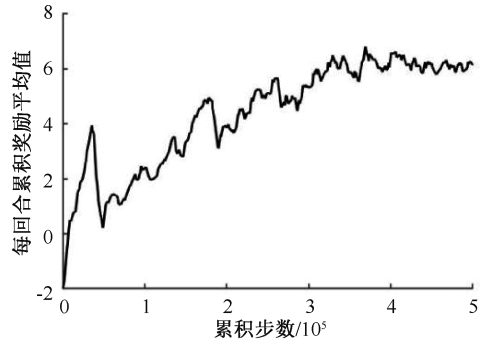


图 3 训练过程中累积奖励平均值变化

Fig. 3 Change in mean accumulated reward during training

成危险切入任务的影响,将 TD3 算法在其他 2 种奖励函数设置下的训练结果作为对比。图中 3 种奖励函数 R_1 、 R_2 和 R_3 的区别仅在于第 2 项,其中,奖励 R_1 采用式(6)的完整奖励函数,奖励 R_2 相比 R_1 在奖励组成的第 2 部分仅采用距离奖励值 r_d ,未采用式(5)中的 TTC 折扣系数缩减距离奖励,奖励 R_3 采用 SMARTS 内置的代表智能体每步行驶距离的 r_e 作为奖励的第 2 部分。

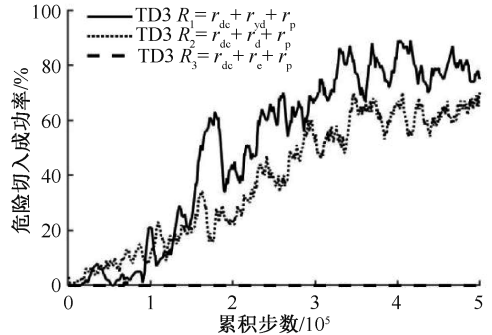


图 4 训练过程中自车危险切入成功率的变化

Fig. 4 Change in success rate of ego vehicle's hazardous cut-in during training

从训练结果看出,TD3 模型在训练过程中最高可达到 80% 左右的危险切入成功率。对比 TD3 算法在 3 种不同奖励函数设置下的训练结果,可以发现,相比采用本文所提出的奖励函数 R_1 ,采用奖励 R_2 时,虽然智能体仍能学习并不断提升,但最终仅能达到 65% 左右的危险切入成功率,而当采用奖励 R_3 时,智能体的危险切入成功率一直为 0,说明智能体不能学到危险切入行为。上述结果表明:文中所设计的奖励函数有利于智能体探索环境中的高奖励状态区域,使智能体能够尽快掌握本文所定义的危险切入行为,最终也取得了更好的性能表现。使用相同的奖励函数对 DDPG 和 PPO 模型进行了训练,

将不同模型在收敛状态下的危险切入成功率结果列于表 3 中。结果显示,采用 TD3 算法所训练模型取得了最高的 80.35% 的危险切入成功率,采用 PPO 算法所训练的模型的危险切入成功率略低于 TD3 模型。与二者相比,DDPG 模型仅能达到 29.57% 的危险切入成功率。由于 TD3 算法在 DDPG 算法的基础上进行了多方面的改进,在文中所定义的车辆危险切入任务中,TD3 算法的表现显著优于 DDPG 算法。

表 3 不同算法所训练模型的危险切入成功率对比

Table 3 Comparison of success rate of ego vehicle's hazardous cut-in among models trained with different algorithms

算法	危险切入成功率/%
TD3	80.35
DDPG	29.57
PPO	68.85

2) 模型测试结果。为了说明模型的泛化能力,将训练得到的最优模型在包括 900、1 200、1 500、1 800、2 100、2 400、2 700 辆/h 的 7 种不同车辆密度场景中进行测试,每种场景均测试 100 个回合。

图 5 为模型在 7 种车辆密度场景中危险切入成功率的对比,其中,在包括 900~2 400 辆/h 的 6 种场景中,模型均能达到 80% 以上的危险切入成功率,仅在车辆密度达到 2 700 辆/h 后,危险切入成功率下降至 63%。随着场景车辆密度提高,自车完成危险切入任务的难度也随之提高,危险切入成功率随之下降,尤其是在车辆密度达到 2 700 辆/h 后,经常出现目标车与相邻车道车辆并行行驶的情况,自车无法行驶至目标车前方。由于模型训练仅采用了 1 200、1 800、2 400 辆/h 3 种场景,对于其余 4 种训练中未使用的测试场景,除 2 700 辆/h 场景中的危险切入成功率较低外,在其余场景中的危险切入成功率均与训练结果相似,说明所训练模型具有较好的泛化能力。

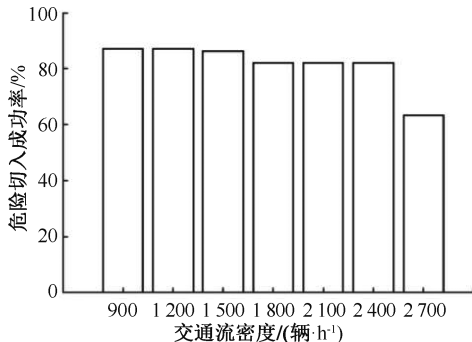


图 5 模型在不同交通流密度场景中的测试结果

Fig. 5 Test results of model in different traffic flow density scenarios

图 6 为模型在测试中的 o_{t_c} 分布,自车在危险切入时该值有约 95% 比例的取值在 $(0, 6]$ s 内,其中, $(0, 2]$ s 占 60%, $(2, 4]$ s 占 30%, $(4, 6]$ s 占 5%,结果表明:所建模型能在切入过程中对目标车产生不同程度的风险,从而覆盖具有不同风险的测试工况。

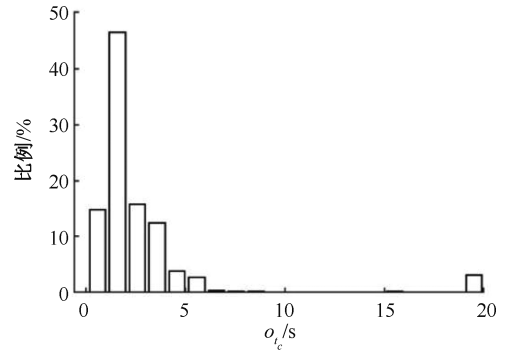


图 6 模型测试中自车在切入时刻与目标车的碰撞时间分布

Fig. 6 Distribution of time to collision between ego vehicle and target vehicle when ego vehicle cut in during model test

3) 智能驾驶策略测试。为将所构建的车辆危险切入策略模型用于 AVs 安全性能测试,在所建立 3 车道环境的基础上构建了包含 2 个智能体的多智能体测试环境,其中一个智能体采用待测智能驾驶策略控制待测车辆,另一智能体采用训练得到的车辆危险切入策略控制一辆测试车,以待测车辆作为目标车。环境中的其他交通车由 SMARTS 内置交通车模型来控制。

为验证该环境的测试效果,选用智能驾驶员模型 (Intelligent Driving Model, IDM) 作为被测智能驾驶策略。IDM 模型是典型的跟车模型,也被其他研究选作被测系统的控制策略^[6-7],其计算方法见下式。为防止 IDM 模型输出加速度值超过合理范围,将 IDM 模型输出的加速度值限制在 $[-7, 7]$ m/s² 范围。IDM 模型参数设置参考 CHEN Baiming 等^[17]的研究。

$$a(t) = \alpha_{\max} \left[1 - \left(\frac{v(t)}{\tilde{v}} \right)^{\beta} - \left(\frac{\tilde{h}(t)}{h(t)} \right)^2 \right] \quad (7)$$

$$\tilde{h}(t) = h_0 + v(t) \tilde{T} + \frac{v(t) \Delta v(t)}{2 \sqrt{a_{\max} a_{\text{com}}}} \quad (8)$$

式中: a 和 h 为自车的加速度和与前车间距,单位分别为 m/s² 和 m; v 和 Δv 为自车车速及自车与前车的相对车速, m/s; a_{\max} 为自车最大加速度或减速度,取 1 m/s²; a_{com} 指舒适减速度,取 1.67 m/s²; \tilde{v} 为期望速度,取 10 m/s; \tilde{T} 为期望车头时距,取 1.5 s;

h_0 为阻塞间距,取 2 m; β 为加速度指数,取 4。

在所构建的多智能体测试场景中测试被测智能驾驶策略,采用 SMARTS 支持的 Visdom 可视化工具对测试过程进行可视化,将摄像头固定于被测车辆(最左侧虚线框)。图 7 为测试过程中的一个片段,可以发现,在仿真初始时刻(图 7a),采用车辆危险切入策略所控制的测试车(中间车道实线框)位于被测车辆后方的中间车道,2 车的纵向初始间距为 6 m;测试车在超过被测车辆一定距离后(图 7b),开始主动降低车速,以确保在切入过程中对被测车辆产生碰撞风险;测试车于 $t=4.9$ s 开始换道(图 7c),最终于 $t=5.5$ s 左右切入被测车辆前方(图 7d),尽管被测车辆在 $t=5.5$ s 迅速采取制动措施,制动减速度达到所设置减速度的上限 -7 m/s^2 ,如图 8 所示。然而,被测车辆最终未能避免与测试车产生碰撞。测试结果表明:采用所提方法可生成对 AVs 安全性具有挑战的车辆切入危险测试场景,有助于发现 AVs 控制策略的安全隐患。

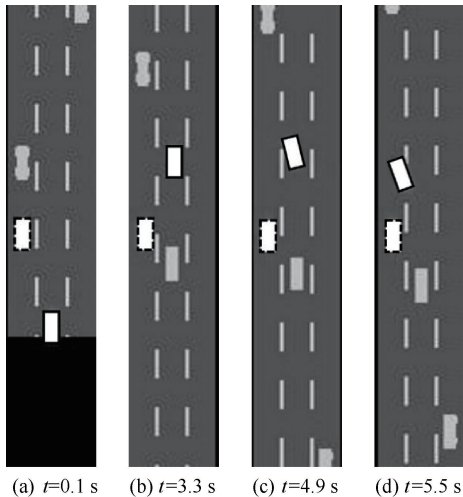


图 7 测试实例

Fig. 7 A test example

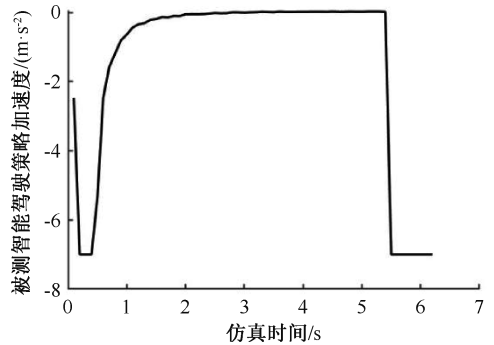


图 8 被测智能驾驶策略输出加速度变化

Fig. 8 Change of output acceleration of tested intelligent driving strategy

3 结论

1) 与 DDPG 和 PPO 这 2 种比较方法相比,TD3 算法在训练过程中能达到更高的危险切入成功率。将所训练模型在 7 种不同车辆密度场景中测试,除车流密度较高的 2 700 辆/h 场景外,在其余场景均取得较高的危险切入成功率,说明模型泛化性能良好。模型所控制自车在切入时刻与目标车的碰撞时间值主要集中在 $(0,6]$ s 范围内,能全面涵盖不同碰撞风险的测试工况。

2) 将所训练模型用于智能驾驶策略的验证,模型所控制交通车能主动危险切入至待测车前方,使待测车辆产生追尾风险,能够起到对待测智能驾驶策略的安全性进行验证的效果。

3) 在后续研究中,将考虑把训练得到的车辆危险切入策略模型集成到自动驾驶虚拟测试平台中,生成前车危险切入场景,以便测试自动驾驶策略。将考虑在包括弯道、十字路口等的复杂场景,以及基于自然驾驶数据的真实交通流中训练模型,提高模型对多样化环境的适应性。

参考文献

- [1] 朱冰,张培兴,赵健,等. 基于场景的自动驾驶汽车虚拟测试研究进展[J]. 中国公路学报, 2019, 32(6): 1-19. ZHU Bing, ZHANG Peixing, ZHAO Jian, et al. Review of scenario-based virtual validation methods for automated vehicle [J]. China Journal of Highway and Transport, 2019, 32(6): 1-19.
- [2] 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 Weiwen, LI Jiangkun, REN Bingtao, 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.
- [3] 徐向阳,胡文浩,董红磊,等. 自动驾驶汽车测试场景构建关键技术综述[J]. 汽车工程, 2021, 43(4): 610-619. XU Xiangyang, HU Wenhao, DONG Honglei, et al. Review of key technologies for autonomous vehicle test scenario

- construction [J]. *Automotive Engineering*, 2021, 43(4): 610-619.
- [4] 郭柏苍, 雒国凤, 金立生, 等. 面向自动驾驶虚拟测试的变道切入场景库构建方法 [J]. *吉林大学学报:工学版*, 2023, 53(11): 3 130-3 140.
GUO Baicang, LUO Guofeng, JIN Lisheng, et al. Construction method of cut-in scenario library for automatic driving virtual tests [J]. *Journal of Jilin University: Engineering and Technology Edition*, 2023, 53(11): 3 130-3 140.
- [5] 赵祥模, 赵玉钰, 景首才, 等. 面向自动驾驶测试的危险变道场景泛化生成[J]. *自动化学报*, 2023, 49 (10): 2 211-2 223.
ZHAO Xiangmo, ZHAO Yuyu, JING Shoucai, et al. Generalization generation of hazardous lane-changing scenarios for automated vehicle testing [J]. *Acta Automatica Sinica*, 2023, 49 (10): 2 211-2 223.
- [6] 朱冰, 范天昕, 赵健, 等. 基于危险边界搜索的自动驾驶系统加速测试方法 [J]. *吉林大学学报:工学版*, 2023, 53(3): 704-712.
ZHU Bing, FAN Tianxin, ZHAO Jian, et al. Accelerate test method of automated driving system based on hazardous boundary search [J]. *Journal of Jilin University: Engineering and Technology Edition*, 2023, 53(3): 704-712.
- [7] FENG Shuo, YAN Xintao, SUN Haowei, et al. Intelligent driving intelligence test for autonomous vehicles with naturalistic and adversarial environment [J]. *Nature Communications*, 2021, 12(1): DOI: 10.1038/s41467-021-21007-8.
- [8] FENG Shuo, SUN Haowei, YAN Xintao, et al. Dense reinforcement learning for safety validation of autonomous vehicles [J]. *Nature*, 2023, 615(7953): 620-627.
- [9] SUN Haowei, FENG Shuo, YAN Xintao, et al. Corner case generation and analysis for safety assessment of autonomous vehicles [J]. *Transportation Research Record*, 2021, 2675(11): 587-600.
- [10] 李江坤, 邓伟文, 任秉韬, 等. 基于场景动力学和强化学习的自动驾驶边缘测试场景生成方法[J]. *汽车工程*, 2022, 44 (7): 976-986.
LI Jiangkun, DENG Weiwen, REN Bingtao, et al. Automatic driving edge scene generation method based on scene dynamics and reinforcement learning [J]. *Automotive Engineering*, 2022, 44(7): 976-986.
- [11] ZHOU Ming, LUO Jun, VILLELLA J, et al. Smarts: an open-source scalable multi-agent reinforcement training school for autonomous driving [C]. *Conference on Robot Learning*, 2021: 264-285.
- [12] 罗崎瑞, 张道文, 周华, 等. 面向智能汽车预期功能安全的驾驶场景评价[J]. *中国安全科学学报*, 2022, 32(8): 140-145.
LUO Qirui, ZHANG Daowen, ZHOU Hua, et al. Evaluation on driving scenarios for safety of intended functionality of intelligent vehicles [J]. *China Safety Science Journal*, 2022, 32(8): 140-145.
- [13] 胡祥旺, 倪颖, 孙剑. 车联网环境下匝道汇入区瓶颈换道优化[J]. *同济大学学报:自然科学版*, 2023, 51(9): 1 424-1 432.
HU Xiangwang, NI Ying, SUN Jian. Freeway merging area lane changing advisory optimization under connected vehicles environment [J]. *Journal of Tongji University: Natural Science*, 2023, 51(9): 1 424-1 432.
- [14] 周文帅, 朱宇, 赵祥模, 等. 面向高速公路车辆切入场景的自动驾驶测试用例生成方法[J]. *汽车技术*, 2021(1): 11-18.
ZHOU Wenshuai, ZHU Yu, ZHAO Xiangmo, et al. Vehicle cut-in test case generation methods for testing of autonomous driving on highway [J]. *Automobile Technology*, 2021(1): 11-18.
- [15] FUJIMOTO S, HOOF H, MEGER D. Addressing function approximation error in actor-critic methods [C]. *International Conference on Machine Learning (ICML)*, 2018: 1 587-1 596.
- [16] ZHOU Yang, CHEN Yunxing. Learning to drive in the NGSIM simulator using proximal policy optimization [J]. *Journal of Advanced Transportation*, 2023, 2023: DOI: 10.1155/2023/4127486.
- [17] CHEN Baiming, CHEN Xiang, WU Qiong, et al. Adversarial evaluation of autonomous vehicles in lane-change scenarios [J]. *IEEE Transactions on Intelligent Transportation Systems*, 2021, 23(8): 10 333-10 342.

作者简介: 周扬 (1989—),男,陕西汉中,博士,副教授,主要从事人-车-路系统安全、自动驾驶测试等方面的研究。E-mail:297399014@qq.com。

