文章编号：1673-0291(2026)03-0138-14 DOI：10.11860/j.issn.1673-0291.20250179

融合SENet的深度强化学习快速路匝道信号控制模型

纪裕伟[a]，孙 健[a,b]，于珂伟[a]，李子豪[a]，赵昱霖[a]
（长安大学 a.未来交通学院，b.运输工程学院，西安 710064）

摘 要：针对城市快速路入口匝道合流区交通拥堵问题，以徐州市北三环快速路中山北路入口匝道为实证案例，提出集成挤压激励网络（Squeeze-and-Excitation Network，SENet）注意力机制的模块化深度强化学习（Deep Reinforcement Learning，DRL）框架，该框架可以优化匝道信号控制，将动态复杂控制问题建模为马尔可夫决策过程（Markov Decision Process，MDP）.设计全面状态空间，涵盖合流区网格交通流分布特征、信号相位状态、时间特征及网联自动驾驶车辆（Connected and Automated Vehicles，CAV）实时交通量；构建多目标奖励函数，平衡实时微观反馈与长期系统目标.为确保仿真环境贴近实际交通场景，基于微观交通仿真平台（Simulation of Urban MObility，SUMO）构建路网模型时，采用回归正交法校准交通流模型，使合流区上下游、中段仿真换道比例与现场监测数据误差控制在±10%以内.训练阶段采用双深度Q网络（Double Deep Q-Network，DDQN），结合优先经验回放、渐进式交通流学习策略提升学习效率与稳定性.研究结果表明：所提方法与传统固定配时控制相比，高峰时段平均延误率降低40.5%、排队长度减少29.5%，优于基准DDQN模型；随着CAV渗透率从20%提升至80%，平均延误率进一步降低44.7%，在混合交通环境中展现出较好的拥堵缓解效果与强适应性.

关键词：交通工程；深度强化学习；匝道信号控制；SENet注意力机制；多目标优化；渐进式交通流学习策略

中图分类号：U491 **文献标志码**：A

A deep reinforcement learning model incorporating SENet for freeway on-ramp signal

JI Yuwei[a]，*SUN Jian*[a,b]，*YU Kewei*[a]，*LI Zihao*[a]，*ZHAO Yulin*[a]
（a. School of Future Transportation，b. School of Transportation Engineering，Chang'an University，Xi'an 710064，China）

Abstract：To address traffic congestion in the merging area of urban expressway on-ramps，this paper takes the Zhongshan North Road on-ramp of the Xuzhou North Third Ring Expressway as an empirical case study. A modular Deep Reinforcement Learning (DRL) framework integrating a Squeeze-and-Excitation Network (SENet) attention mechanism is proposed to optimize ramp signal control，formulating this dynamic and complex control problem as a Markov Decision Process (MDP). A comprehen-

收稿日期：2025-11-07；**修回日期**：2026-04-09
基金项目：国家自然科学基金（52172319，71971138）
Foundation items：National Natural Science Foundation of China（52172319，71971138）
第一作者：纪裕伟（2000—），男，河南安阳人，硕士生.研究方向为交通管理与控制，交通仿真.**email**：2020900069@chd.edu.cn.
通信作者：孙健（1977—），男，安徽芜湖人，教授，博士.**email**：jiansun@chd.edu.cn.

引用格式：纪裕伟，孙健，于珂伟，等.融合SENet的深度强化学习快速路匝道信号控制模型[J].北京交通大学学报，2026，50(3)：138－151.
JI Yuwei，SUN Jian，YU Kewei，et al.A deep reinforcement learning model incorporating SENet for freeway on-ramp signal[J].Journal of Beijing Jiaotong University，2026，50(3)：138－151.(in Chinese)

sive state space is designed, encompassing the grid-based traffic flow distribution characteristics of the merging area, signal phase states, temporal features, and the real-time traffic volume of Connected and Automated Vehicles (CAVs). Furthermore, a multi-objective reward function is constructed to balance real-time microscopic feedback with long-term system objectives. To ensure the simulation environment closely reflects actual traffic scenarios, an orthogonal regression method is adopted to calibrate the traffic flow model during the construction of the road network in the Simulation of Urban MObility (SUMO) microscopic traffic simulation platform. Consequently, the discrepancies between the simulated lane-changing ratios at the upstream, mid-section, and downstream segments of the merging area are controlled within ±10%. compared with field monitoring data. During the training phase, a Double Deep Q-Network (DDQN) is employed alongside prioritized experience replay and a progressive traffic flow learning strategy to enhance learning efficiency and stability. Experimental results indicate that, compared to traditional fixed-time signal control, the proposed method reduces the average delay rate by 40.5% and shortens the queue length by 29.5% during peak hours, outperforming the baseline DDQN model. Furthermore, as the CAV penetration rate increases from 20% to 80%, the average delay rate decreases by an additional 44.7%, demonstrating the model's excellent congestion mitigation capabilities and strong adaptability in mixed traffic environment.

Keywords: traffic engineering; deep reinforcement learning; ramp signal control; SENet attention mechanism; multi-objective optimization; progressive traffic flow learning strategy

城市快速路匝道合流区作为连接地面道路与快速路的关键节点,其固有交通流线复杂性及交通需求时变特性,严重制约快速路系统整体运行效率,交通组织与控制面临巨大挑战.因此,针对城市快速路匝道及合流区的相关研究具有重要的现实意义和应用价值.

基于数据驱动的匝道入口主被动控制被广泛研究.文献[1]提出了ALINEA控制,通过反馈机制调整前一时刻匝道流入率,以期达到理想占有率.ALINEA控制在缓解入口匝道合流区拥堵及保证交通系统稳定方面效果显著,且容易实施,在许多城市快速路及高速公路系统得到应用[2].文献[3]对快速路合流区主线到达流率、汇入流量和合流区通行能力三者间的关系进行研究,借鉴RWS-C控制策略,提出基于影响通行能力的匝道控制方法.文献[4]分析了交通流崩溃现象与合流区交通拥堵现象的关系,提出交通流崩溃事件探测算法.文献[5]从高速公路交通流运行时空特性角度出发,构建以通行效率最大和平均延误最小为控制目标的高速公路主线与匝道协调最优控制模型.文献[6]构建了多车道分配与入口多匝道协同控制模型,采用入口多匝道控制协调匝道合流区各汇入匝道车辆,实现主线和匝道的通行效率最大化.

深度强化学习(Deep Reinforcement Learning, DRL)在传统入口匝道控制领域带来了革新性机遇.一方面,为提高基于DRL算法的匝道控制效果,通过改进算法,使模型在有限特征下尽可能挖掘潜在信息.文献[7]提出一种基于深度Q网络(Deep Q-Network, DQN)的入口匝道控制方法,实现对入口匝道车流的智能实时控制.文献[8]采用了深度确定性策略梯度(Deep Deterministic Policy Gradient, DDPG)算法,适用于连续动作空间问题,克服了仅以固定周期为前提的动作输入导致输出方案可能非最优解的问题.文献[9]提出了一种基于多智能体深度强化学习(Multi-Agent Deep Reinforcement Learning, MADRL)的自改进入口匝道算法,为入口匝道信号灯实时生成优化调节方案.文献[10]针对单入口匝道场景提出了DQNOP(Deep Q-Network Overflow Protection)算法,设计溢出保护模块自适应解决入口匝道溢出问题.此外,DRL模型的决策过程难以解释,这在需要高度安全性和可靠性的交通控制领域可能成为一个障碍.为解决这些问题,研究者尝试在DRL中融入物理信息[11-12],或将DRL与模型预测控制(Model Predictive Control, MPC)等传统模型控制方法相结合[13].文献[14]提出了一种基于DRL的鲁棒自适应匝道控制模型DRLARM(Deep Reinforcement Learning-Based Adaptive Ramp Metering),设计了多种交通状态进行研究.文献[15]采用基于图像卷积神经网络的匝道控制DRL算法,建立了基于图像卷积神经网络的连续时空交通状态解析.另一方面,为提高基于DRL算法的交通信号控制策略的学

习和控制效果，可通过结合注意力机制探索不同特征对环境信息的表达效果，从而用较少的特征全面描述环境.文献[16]将注意力机制融入基于DRL的协同入口匝道控制策略中，使其能够捕捉入口匝道和主线段交通状态之间的时空相关性.文献[17]融合异构图注意力机制与Actor-critic算法，通过对交通状态信息分类分配不同注意力嵌入向量.文献[18]运用基于注意力机制的DRL信号控制算法DQN_AM(Deep Q Network_Attention Module)，降低状态向量设计难度，但未探索其他DRL算法结合注意力机制能否达到该目的.文献[19]采用自注意力网络DRL算法，使用非均匀量化和独热编码方法从有限特征中挖掘潜在交通状态信息，从而降低交通状态设计难度.

智能网联车辆(Connected and Automated Vehicles, CAV)的普及将导致道路上长期存在CAV与传统人工驾驶车辆(Human-driven Vehicles, HV)混合行驶的交通场景.文献[20]解析了CAV渗透率混合交通流基本图，改进了标准入口匝道控制策略.文献[21]提出了一种动态协同汇入辅助策略以适应变化的交通需求.文献[22]分析了明尼阿波利斯一个入口匝道混合自主交通对入口匝道控制性能的影响，为实际应用提供了理论基础.文献[23]基于博弈论车辆协同控制模型，根据纳什均衡原理求解同一对局中车辆的最优控制策略.文献[24]设计了CAV集中控制场景中考虑特种车辆优先通行的协同控制方法.文献[25]针对反馈控制、最优控制和强化学习3种不同的控制范式，对CAV使用的融合策略进行了评估并给出了具体的改进措施.

对于传统模型，反馈控制存在明显时延，可能导致拥堵快速扩散；基于模型的预测控制依赖精准的交通预测模型，在复杂混合交通环境中难以达到控制精度要求.对于DRL，状态表征存在不完备性，研究场景单一，缺乏跨场景泛化设计，研究聚焦于组合有限特征全面表达交通状态或引入新技术提升DRL算法的特征提取能力，忽视了对状态表征的挖掘及其本身的可迁移性设计，缺少两者结合的研究，导致模型高度依赖训练场景的几何与流量特征；奖励函数虽会考虑长期回报，但多个随机变量导致合流区整体情况描述存在偏差，奖励设计本身缺乏对多样化交通场景的归一化设计与自适应考量，使得优化结果缺乏精细化与精准性，限制了策略在交通构成与需求模式变化时的泛化能力.

综上，本文将混合交通环境下匝道流量调节控制问题抽象为马尔可夫决策过程(Markov Decision Process, MDP).针对状态表征环节，引入通道注意力机制(Squeeze-and-Excitation Network, SENet)的合流区网格化交通流建模方法，通过SENet动态校准网格特征的重要性，智能体能够自适应聚焦合流区拥堵区域，提升状态信息决策效用和可迁移性.在此基础上，进一步将实时CAV渗透率信息纳入状态空间，设计多目标奖励函数以平衡系统长期目标与即时反馈，采用双深度Q网络(Double Deep Q-Network, DDQN)算法求解该问题.为验证方法的有效性与泛化能力，选取徐州市北三环快速路中山北路入口匝道合流段数据，基于城市交通仿真系统(Simulation of Urban MObility, SUMO)构建混合交通流下城市快速路合流区的微观交通仿真环境，训练并测试优化DDQN模型，探究SENet注意力机制模块对控制性能的提升能力，验证所提优化DDQN模型在不同CAV渗透率下的优越性，并证明融入基于网格的注意力机制可进一步提升控制性能，为模型跨场景迁移提供基础.

1 马尔可夫决策过程建模

1.1 状态空间与预处理

状态元素包括基于网格的交通流分布特征、信号状态特征、时间特征及CAV传输的匝道交通量信息，状态S为

$$S=[\boldsymbol{G},\boldsymbol{P},T,V_{\mathrm{CAV}}^{\mathrm{Up}},V_{\mathrm{CAV}}^{\mathrm{R}}] \tag{1}$$

式中：$\boldsymbol{G}$为合流区基于网格的交通流特征矩阵；$\boldsymbol{P}$为当前信号相位的独热编码向量；T为当前相位的归一化时长；$V_{\mathrm{CAV}}^{\mathrm{Up}}$为上游主线CAV交通量；$V_{\mathrm{CAV}}^{\mathrm{R}}$为匝道CAV交通量.交通流特征建模为三维张量$\boldsymbol{G}\in R^{L\times C\times 2}$，其中，$L$为受控车道数量，$C$为每条车道的网格单元数量，第3维度包含网格占用指示符与归一化速度两个通道.网格单元长度由$L_{\mathrm{cell}}=L_{\mathrm{lane}}/C$计算，$L_{\mathrm{cell}}$为单个网格长度，$L_{\mathrm{lane}}$为受控车道实际长度，网格数量根据最小检测距离确定.对于第i条车道上的第j个网格单元，定义二元网格占用变量$g_{i,j,0}$与归一化速度特征$g_{i,j,1}$为

$$g_{i,j,x}=\begin{cases}1 & x=0\text{且网格}i\text{被占用}\\ 0 & x=0\text{且网格}i\text{未占用}\\ \min\left(\dfrac{v_{i,j}}{v_{\max}},1.0\right) & x=1\end{cases} \tag{2}$$

式中：$v_{i,j}$为网格单元(i,j)内车辆的实际速度；$v_{\max}$为道路设计最高速度；x为特征通道索引变量.

信号相位采用独热编码表征当前相位状态.在

匝道控制场景下,信号控制器遵循四相位逻辑[26].独热编码p_k为

$$\sum_{k=0}^{U-1} p_k = 1 \qquad p_k \in \boldsymbol{P} \tag{3}$$

式中:U为相位总数;p_k为该向量的第k个元素,当前信号相位为相位k时取值为1,否则取值为0.对黄灯相位,其编码需使智能体识别此类过渡相位,并在此期间避免执行无效相位切换动作.

相位时长特征T作为表征当前相位已持续时长的补充时间信息定义为

$$T = \min\left(\frac{t_{\text{current}}}{t_{\max}}, 1.0\right) \tag{4}$$

式中:t_{current}为当前相位激活后已持续时长;$t_{\max}$为该类型相位最大允许时长,固定为40 s.该特征为相位演变提供时间参考,辅助智能体判断是否需要切换相位.为确保与其他特征的量纲一致性,将该时间特征归一化至[0,1]区间.

1.2　动作空间

相位切换遵循严格的逻辑顺序,其中黄灯相位为过渡相位,时长固定且不可调整.因此,需排除物理上不现实的相位跳变,动作空间A定义为

$$A = \begin{cases} 0 & \text{保持当前相位} \\ 1 & \text{切换下一相位} \end{cases} \tag{5}$$

双动作设计在保证交通信号安全运行的同时,仍赋予智能体足够的决策自由度,且简化后的动作空间有助于降低学习复杂度.若当前相位为绿灯且选择动作1,则发起相位切换请求;黄灯相位期间,强制忽略所有动作输入,相位保持固定时长;黄灯相位结束后,新相位计时开始.

1.3　多目标奖励函数

为实现即时微观交通状态改善与长期系统运行优化的动态平衡,兼顾路网通行效率、驾驶舒适性与信号控制安全性的多目标协同,奖励函数R_t被定义为

$$\begin{aligned} R_t = {} & \omega_1 \cdot \Delta Q_t + \omega_2 \cdot \Delta W_t + \omega_3 \cdot Q_t^{\text{norm}} + \\ & \omega_4 \cdot W_t^{\text{norm}} + \omega_5 \cdot T_t^{\text{norm}} + \omega_6 \cdot r_{\text{d}} \end{aligned} \tag{6}$$

式中:ω_1至ω_6分别为各分量的权重.

ΔQ_t为合流区平均排队长度趋势奖励,聚焦排队长度变化趋势,ΔQ_t被定义为

$$\Delta Q_t = \frac{\bar{Q}_{t-1} - \bar{Q}_t}{\max Q(t) - \min Q(t)} \tag{7}$$

式中:$\bar{Q}_t$为t时刻合流区平均排队长度.该奖励项引导智能体快速响应拥堵,当平均排队车辆减少时给予正向奖励,鼓励智能体主动采取减堵动作,侧重动态优化.

ΔW_t为合流区平均车辆延误趋势奖励,聚焦行车体验改善趋势,ΔW_t被定义为

$$\Delta W_t = \frac{\bar{W}_{t-1} - \bar{W}_t}{\max W(t) - \min W(t)} \tag{8}$$

式中:$\bar{W}_t$为t时刻合流区平均车辆延误.该奖励项关注总延误下降幅度,当车辆延误减少时,给予正向奖励,强化快速消除延误的动作,侧重动态优化.

Q_t^{norm}为合流区排队长度绝对奖励,聚焦当前时刻合流区排队状态,Q_t^{norm}被定义为

$$Q_t^{\text{norm}} = -\frac{\sum_{i \in L} Q_i(t)}{\max Q(t)} \tag{9}$$

式中:$Q_i(t)$为t时刻车道i的排队长度;L为合流区车道集合;$\max Q(t)$为滑动窗口内合流区最大排队长度.该奖励项约束智能体避免长期拥堵,强制智能体将排队长度控制在安全范围.

W_t^{norm}为合流区车辆延误绝对奖励,聚焦于当前时刻合流区的延误状态,W_t^{norm}被定义为

$$W_t^{\text{norm}} = -\frac{\sum_{i \in L} W_i(t)}{\max W(t)} \tag{10}$$

式中:$W_i(t)$为t时刻车道i的车辆延误;L为合流区车道集合;$\max W(t)$为在滑动窗口内合流区的最大车辆延误.该奖励项聚焦个体累积延误的绝对量,避免智能体为追求总等待时间下降而牺牲部分车辆的行程体验.

T_t^{norm}为合流区通行能力奖励,聚焦路网容量最大化,T_t^{norm}被定义为

$$T_t^{\text{norm}} = \frac{\sum_{i \in L} n_i(t)}{N} \tag{11}$$

式中:$n_i(t)$为t时刻车道i的车辆通过数;L为合流区车道集合;N为道路通行能力.通过量化当前时刻通过交通量与道路容量的比值,对合流区的通行效率进行宏观判断,引导智能体让更多车辆尽快驶出合流区域.

r_{d}为相位切换奖励,仅当智能体触发相位切换动作时,才会根据当前合流区排队长度计算切换惩罚,避免无意义的惩罚叠加,确保惩罚仅针对实际发生的切换行为,r_{d}被定义为

$$r_{\text{d}} = -(1 + Q_t^{\text{norm}}) I(a_t = 1) \tag{12}$$

式中:$I(a_t = 1)$为切换指示函数.各分量均通过滑动窗口进行归一化处理保证量纲无关性.

1.4　融入SENet注意力机制状态空间优化

图1为城市快速路合流区交通流快照.由图1可知,拥堵更易发生在靠近合流区下游端的位置,且

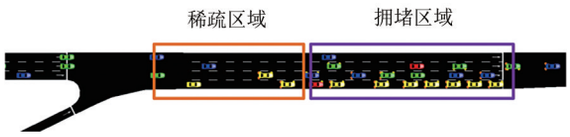


图 1 拥堵时刻交通流快照

Fig.1 Snapshot of traffic flow during congestion

该现象对智能体的策略学习影响更大.

基于此,SENet通过“压缩-激励”操作自适应增强关键特征通道的表征力,提升状态信息的决策效用.其结构简洁、计算高效,易于嵌入DRL框架而不影响实时性,能够抑制无关特征[27].相较于空间注意力(在网格化表征中冗余)、自注意力(计算开销大)或混合注意力(结构复杂),SENet在保障对特征通道进行精准重校准的前提下,更好地实现了模型性能与计算效率的平衡.其压缩操作与激励操作的原理为

$$S_c=\frac{1}{H\times W}\sum_{m=1}^{H}\sum_{n=1}^{W}F_c(m,n) \tag{13}$$

$$Z_c=\sigma(\boldsymbol{W}_2\cdot\mathrm{ReLU}(\boldsymbol{W}_1\cdot S_c+b_1)+b_2) \tag{14}$$

式中:$F_c=(m,n)$为第c个通道在空间位置(m,n)处的特征值;H和W分别为特征图的高度与宽度;S_c为第c个通道的全局特征标量(聚合该通道的空间信息);$\boldsymbol{W}_1\in R^{(C/r)\times C}$和$\boldsymbol{W}_2\in R^{(C/r)\times C}$分别为降维与升维权重矩阵(其中$r$为压缩比);$b_1$和$b_2$为偏置项;$\sigma$为Sigmoid激活函数;$Z_c$为第$c$个通道的注意力权重(取值范围为[0,1]),权重越大表明该通道的特征越重要.

SENet本身不显式建模空间关系,故设计网格化表征编码车辆的空间分布,强化空间感知.网络首先通过卷积层从交通流网格中提取空间特征,并基于网格设计所包含的两类信息执行分离操作.随后,网络进入注意力机制的核心模块,该模块包含3个步骤.

步骤1:压缩阶段采用全局平均池化,将每个特征通道的空间信息聚合为全局描述子,生成通道维度的统计特征.

步骤2:激励阶段由两个全连接层构成的瓶颈结构学习通道间的非线性依赖关系.通过应用Sigmoid激活函数,生成取值范围为0~1的通道注意力权重.

步骤3:重校准阶段将学习到的注意力权重与原始特征图在通道维度上以逐元素方式相乘,随后进行组合.这一过程实现了特征的自适应增强,使网络能够根据实时交通状况动态调整感知重点,并提升关键特征的表征能力.

SENet为后续的强化学习决策提供了更具区分度的状态表征,其结构如图2所示.

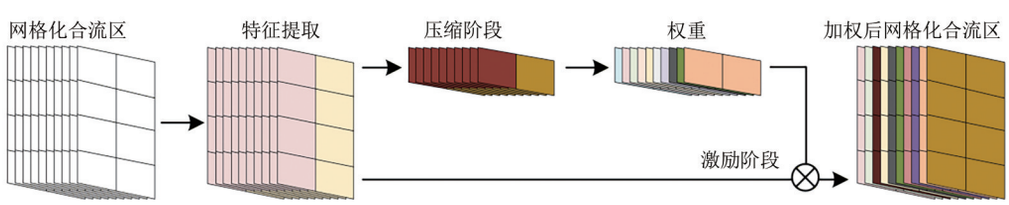


图 2 SENet注意力机制网络架构图

Fig.2 Network architecture of the SENet attention mechanism

2 深度强化学习方法

DRL旨在实现从交通状态精准感知到信号动作智能决策的闭环优化,该框架分为3个协同作用层.状态感知层:以SENet为核心,接收网格化交通流数据,输出强化状态特征向量.网络学习层:以DDQN为核心,通过神经网络拟合最优动作价值函数,并基于时序差分(Temporal Difference, TD)误差进行更新.训练优化层:包含优先经验回放机制、ε-贪心动作选择策略、渐进式交通流学习策略3个模块,旨在高效、稳定地训练上述状态表征与决策网络.模型架构如图3所示.

2.1 双深度Q学习更新算法

DDQN利用深度神经网络逼近状态-动作价值函数,由DQN算法改进而来.DQN算法的动作选择与评估仅依赖目标网络参数,易产生高估问题.DDQN算法能够处理由网格化合流区、信号相位及CAV数据构成的高维状态输入,通过双网络架构缓解高估问题,为需要高稳定性的交通控制场景提供保障.在匝道交通信号控制场景中,Q函数表示在状态s下执行动作a所能获得的期望累积折扣奖励,记为$Q(s,a)$.Q学习的核心更新逻辑基于TD误差,更新规则为

$$Q(s,a)\leftarrow Q(s,a)+\alpha[r+\gamma\max_{a'}Q(s',a')-Q(s,a)] \tag{15}$$

式中:$\alpha\in(0,1)$为学习率;$\gamma\in(0,1)$为折扣因子;r为即时奖励;s'为执行动作a后到达的新状态.在DQN中,该更新规则转化为损失函数的最小化问题,计算式为

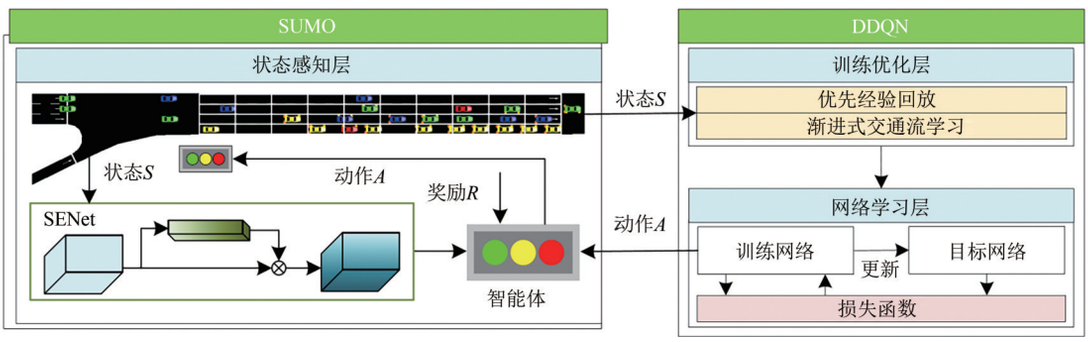


图 3　基于DDQN的匝道信号控制框架

Fig.3　DDQN-based ramp signal control framework

$$L(\theta)=E\left[(y_t-Q(s_t,a_t;\theta))^2\right] \tag{16}$$

式中:目标值为 $y_t=r_t+\gamma\max Q(s_{t+1},a';\theta^-)$;$\theta$ 为当前网络的参数,θ^- 为目标网络的参数.该目标网络与当前网络架构相同,但参数保持固定;经过固定次数的迭代后,再从当前网络复制参数更新目标网络.这一机制可有效降低连续样本间的相关性,提升训练稳定性.

2.2　模块化优化策略

DQN普遍存在训练过程不稳定、对探索策略设计敏感及性能高估等挑战,为此,模型设计了3个策略以应对这些局限.

策略1:优先经验回放机制.在城市快速路匝道信号控制场景中,一个仿真周期内连续时刻的状态信息存在强相关性,导致样本数据间相关性过高,且部分关键时间步可能对训练产生较大影响,不利于智能体决策.针对上述问题,引入优先经验回放机制[28].智能体在每个时间步获取经验四元组样本(s_t,a_t,r_t,s_{t+1}),并将其存储至经验池中.新经验的初始优先级设为当前经验池中所有经验的最高优先级,确保新经验至少被采样一次.采样过程中,通过TD误差衡量经验的重要性:TD误差越大,表明对应经验对模型修正的潜力越强,因此获得更高优先级.

策略2:ε-贪心动作选择策略.采用ε-贪心策略平衡“对当前估计最优动作的利用”与“对状态空间的持续探索”.训练过程中,ε-贪心策略以概率$1-\varepsilon$选择估计价值最高的动作,以概率ε随机选择动作.为促进模型收敛,引入衰减因子,每经过一个仿真步,随机选择动作的概率便会降低.该机制使智能体在训练初期倾向于执行探索性随机动作,在训练后期则倾向于选择估计价值或奖励最高的贪心策略动作.最终实现智能体在利用已知信息的同时,持续探索环境中未被访问的状态.

策略3:渐进式交通流学习策略.在实际交通场景中,高峰交通流量易达到3 000~4 000 veh/h,会导致仿真反复触发车辆瞬移与紧急制动,破坏奖励信号稳定性.基于此,采用渐进式交通流学习策略,引入调整因子以优化决策过程.初始仅输入20%低交通流量,第1~60个训练回合线性提升至50%,使系统先学习基础绿灯放行策略;第61~200个训练回合,将交通流量线性提升至100%,使智能体掌握重度拥堵下相位切换机制;最后100个训练回合保持100%交通流量输入,此时网络已能获得稳定参数.

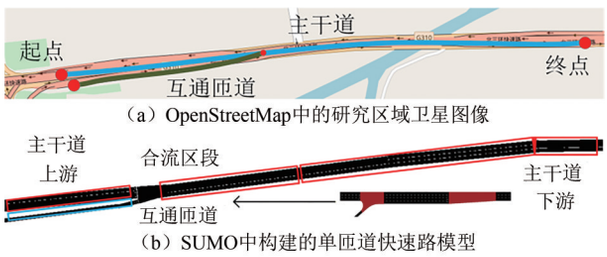
2.3　模型设计泛化性能分析

框架以模块化和参数化为设计原则,增强模型泛化性.在状态空间采用网格化表征,通过调整车道数与网格数适配不同合流区的几何布局.奖励函数所有指标均通过滑动窗口在线归一化处理,使智能体聚焦优化相对性能而非绝对数值,确保不同交通需求水平下目标一致性.同时,SENet动态聚焦关键区域,渐进式交通流学习策略通过从简单到复杂的训练过程,提升模型对不同交通需求的泛化能力.

3　仿真模型设置

3.1　路网建模

仿真以中国江苏省徐州市北三环快速路中山北路入口匝道合流段为研究对象.该快速路主线限速为80 km/h,匝道限速为40 km/h,所选取路段主线为3车道,全长1.36 km;匝道为单车道,长度为200 m;合流区为4车道,长度为100 m.研究区域范围如图4所示.实际交通需求来源于2023年3月28日至2023年4月2日该路段环形监测器记录数据.经交通需求分析,6:30至8:30和16:00至18:00为高负荷时间段,其他为低负荷时间段,即高峰时段主线小时交通量为3 785辆,匝道小时交通量为1 101辆;平峰时段主线小时交通量为1 250辆,匝道小时交通量为302辆.



（a）OpenStreetMap中的研究区域卫星图像

（b）SUMO中构建的单匝道快速路模型

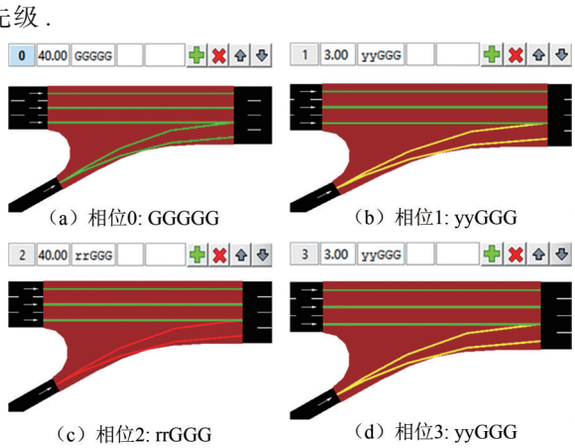
图 4 研究区域说明

Fig.4 Overview of the study area

针对该合流区的匝道信号控制系统，定义4个相位，对应的相位状态模型为

$$P_t \in \{0, 1, 2, 3\} \tag{17}$$

不同相位下合流区通行规则如图5所示．对于匝道受控车道，相位0为绿灯（40 s），相位2为红灯（40 s），相位1与相位3为黄灯（3 s）．为贴合实际运行规则，各相位下主线进口道均保持绿灯；且当主线与匝道同时具备通行权时，主线车辆享有更高优先级．



（a）相位0: GGGGG （b）相位1: yyGGG

（c）相位2: rrGGG （d）相位3: yyGGG

图 5 匝道信号控制系统相位配置

Fig.5 Phase configuration of the ramp signal control system

3.2 交通流与网络模型校准

通过回归正交法，结合正交试验设计与回归分析进行参数校准[29]．通过建立回归模型估计变量与参数之间的函数关系，确定各参数最优值[30]．在SUMO中，车辆的纵向运动由跟驰模型（Krauss模型）控制，横向合流行为由换道模型（LC2013）控制．仿真校准需使仿真能够准确反映合流区车辆交互的微观动态，模型选取的关键校准参数如表1所示．选取匝道车辆驶入主线后的换道位置比例作为核心校准指标，主要基于以下考虑：首先，换道行为是匝道合流过程最直接的微观表现，其空间分布影响合流区车流的稳定性，是连接车辆决策与宏观交通状态的关键变量．其次，该指标具有良好的可观测性，且不易受瞬时交通波动影响，具备作为稳定校准基准的条件．最后，SUMO中的换道模型参数对车辆换道决策位置敏感，校准该比例能有效反映真实的驾驶行为特征．

表 1 SUMO交通流模型参数校准范围

Tab.1 Calibration ranges for SUMO traffic flow model parameters

模型	属性	描述	下限	默认	上限
Krauss (car-following)	tua	安全车头时距	0.5	1.0	1.5
	sigma	驾驶误差	0.2	0.5	0.8
	accel	最大加速度	1.6	2.0	2.4
	decel	最大减速度	3.6	4.5	5.4
LC2013 (lane-changing)	lcStrategic	策略性换道倾向	0.4	1.0	1.6
	lcCooperative	合作性换道倾向	0.2	0.5	0.8
	lcSpeedGain	速度增益阈值	0.5	1.0	1.5

针对Krauss跟驰模型和LC2013换道模型的7个参数，采用$L_{18}(3^7)$正交表设计实验，将各参数设计为低、中、高3个水平，统计区域为合流区上游50 m至加速车道下游50 m的范围．对每个实验方案进行仿真，记录换道比例，以确定最优参数组合，使仿真输出的换道比例与现场观测值之间的误差在±10%以内．多元线性回归模型为

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_7 x_7 + \varepsilon \tag{18}$$

式中：y为加速车道中段匝道车辆的换道比例；x_1～x_7为表3中的7个参数；β_0为常数项；β_1～β_7为各参数的回归系数；ε为服从正态分布的随机误差项．最优参数组合如表2所示，仿真输出的换道比例与实测数据的对比如表3所示，所有区段的绝对误差均小于7%．

表 2 SUMO交通流模型参数的最优校准集合

Tab.2 Optimal calibration set for SUMO traffic flow model parameters

模型	属性	校准值
Krauss (car-following)	tua	1.1
	sigma	0.5
	accel	2.2
	decel	4.8
LC2013 (lane-changing)	lcStrategic	1.3
	lcCooperative	0.4
	lcSpeedGain	1.2

为进一步从宏观角度验证校准效果，引入总行程耗时（Total Time Spent，TTS）作为补充验证指标．TTS定义为在仿真时段内，所有车辆在路网中行驶时间的总和，在相同的早高峰交通需求输入下，分别使用真实监测数据和校准后的最优参数集进行对比，并计算主路和匝道车辆的TTS绝对误差．结果如表4所示，匝道车辆、主路车辆的仿真值与实际值绝对误差分别为7.72%、9.60%，基于换

表 3　最优校准参数组的SUMO仿真结果对比

Tab.3　Comparison of SUMO simulation results for the optimal calibrated parameter set

车辆类型	换道位置区段	实际比例/%	校准后比例/%	绝对误差/%
匝道车辆	上游区段	24.15	25.34	4.93
	中段	47.24	45.86	2.92
	下游区段	28.61	28.8	0.66
主路车辆	上游区段	58.32	56.28	3.5
	中段	35.21	37.56	6.67
	下游区段	6.47	6.16	4.79

道比例校准的模型不仅复现了微观行为特征,更在宏观系统性能上更贴近实际情况,为后续强化学习训练提供了可靠的仿真环境.

表 4　最优校准参数组TTS结果对比

Tab.4　Comparison of TTS results for the optimal calibrated parameter set

车辆类型	对比数据类型	上游区段总行程耗时/(veh·h)
匝道车辆	仿真值	141.43
	实际值	153.27
主路车辆	仿真值	115.13
	实际值	127.36

此外,DDQN算法性能受多个超参数影响,学习率设为0.001,折扣因子设为0.9,探索率设置为从1衰减至0.05,目标网络更新频率设为500步.经验池大小设为100 000.批处理大小设为32,优先调整因子设为0.6.

4　实验验证与结果分析

实验首先在20%CAV渗透率下,评估不同组件对匝道信号控制的影响,随后分析所提模型在不同CAV渗透率下的控制性能.

4.1　奖励函数敏感性分析及权重设置

奖励函数包括一系列权重,需要针对影响控制性能的奖励项进行灵敏度分析,根据控制效果确定奖励项的权重分配.为实现模型有效训练,首先通过小规模网格搜索[31]构建一组稳定收敛至较高奖励函数值的基准权重参数组,如表5所示.灵敏度分析仿真均基于经校准的SUMO模型开展,输入数据采用真实早高峰交通需求,仿真时长为3 600 s,并进行1 000 s交通流预热.

表 5　奖励函数基准权重参数组

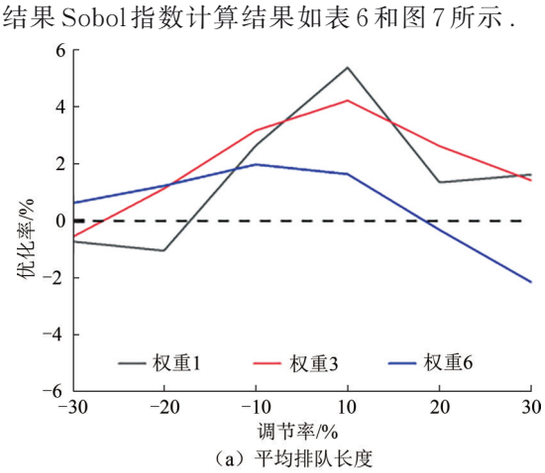
Tab.5　Baseline parameter set for the reward function weights

权重符号	权重值	权重符号	权重值
ω_1	0.282	ω_4	0.076
ω_2	0.221	ω_5	0.253
ω_3	0.101	ω_6	0.067

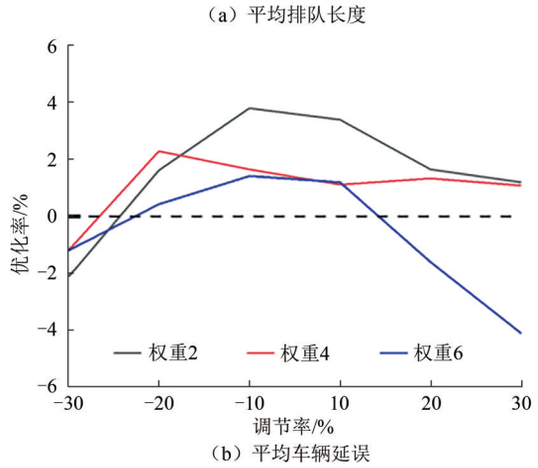
采用基于单因子扰动的局部敏感性分析与基于Sobol方差法的全局敏感性分析相结合的方法进行研究.在局部分析中[32],通过固定其他奖励权重,使每个奖励权重单独扰动±30%,以10%为步长训练智能体并记录对应性能指标变化.图6展示了不同权重变化对其关联合流区性能的局部敏感性分析结果.

由图6可知:过高或过低的奖励函数权重设置,都会降低模型优化效果;ω_1和ω_3在提升10%时优化效果最佳,ω_2和ω_4在降低10%时优化效果最佳;过高的调整权重ω_5,平均通行能力的优化率基本保持不变;ω_6在调节率为−10%时整体效果最佳.

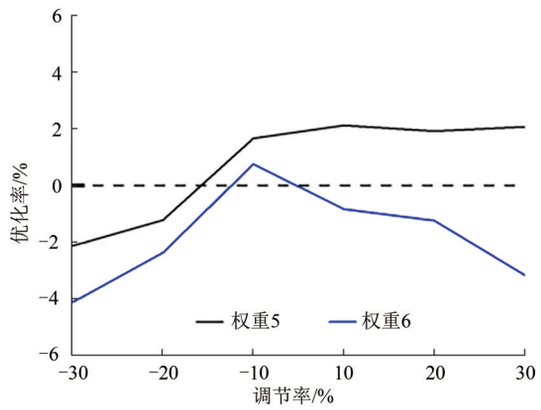
Sobol方差法是一种基于方差分解的全局敏感性分析方法,其核心思想是将模型输出的总方差分解为各输入参数单独作用及其相互作用所贡献的方差之和[33].该方法通过计算两类关键指数来量化参数重要性:一阶Sobol指数S_1,反映单个参数单独变化对输出方差的贡献程度,衡量独立影响;总效应Sobol指数S_T,涵盖该参数自身及其与所有其他参数交互作用共同导致的方差贡献,反映全局影响.根据文献[34]的研究,对于中等维数的问题,基础样本量$N=64$足以获得稳定的敏感性排序结果,对于$k=6$个权重,总仿真次数为$N\times(k+2)=512$次,仿真结果Sobol指数计算结果如表6和图7所示.



(a) 平均排队长度



(b) 平均车辆延误



（c）平均通行能力

图 6 奖励函数权重局部敏感性分析

Fig.6 Local sensitivity analysis of reward function weights

表 6 奖励函数权重 Sobol 指数分析

Tab.6 Sobol index analysis of the reward function weights

权重	平均车辆延误		平均排队长度		平均通行能力	
	S_1	S_T	S_1	S_T	S_1	S_T
ω_1	0.38	0.46	0.42	0.56	0.14	0.27
ω_2	0.22	0.35	0.21	0.37	0.22	0.32
ω_3	0.11	0.18	0.06	0.12	0.14	0.14
ω_4	0.02	0.09	0.04	0.08	0.03	0.06
ω_5	0.17	0.26	0.12	0.22	0.32	0.48
ω_6	0.06	0.14	0.11	0.24	0.06	0.13

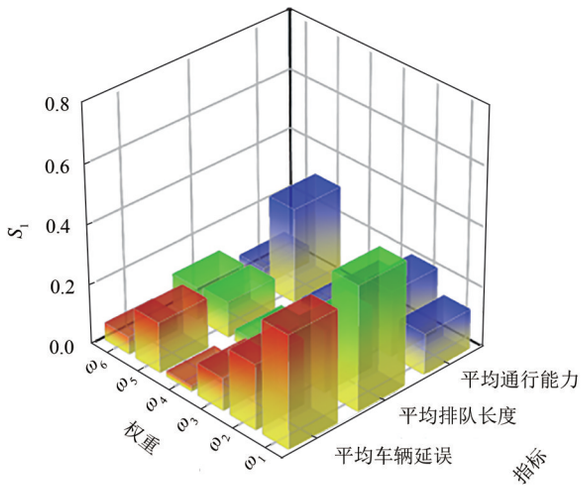
由表 6 和图 7 可知：ω_1 和 ω_5 为主导权重，构成了系统优化的两个主要维度，两者相对大小决定了控制策略的倾向性，且通过交互作用共同影响系统行为；ω_2 和 ω_4 起到辅助调节作用，两者通过交互作用（S_T 与 S_1 差值较高）调节 ω_1 和 ω_5 的影响，特别是权重 ω_2，较高 S_T 值体现了其在缓解主线延误与排队间的桥梁作用；权重 ω_3 和 ω_6 对稳定性提供保障，两者提供了持续状态反馈和基本约束机制，防止系统行为过度极端化．最终权重配置如表 7 所示．

4.2 消融实验

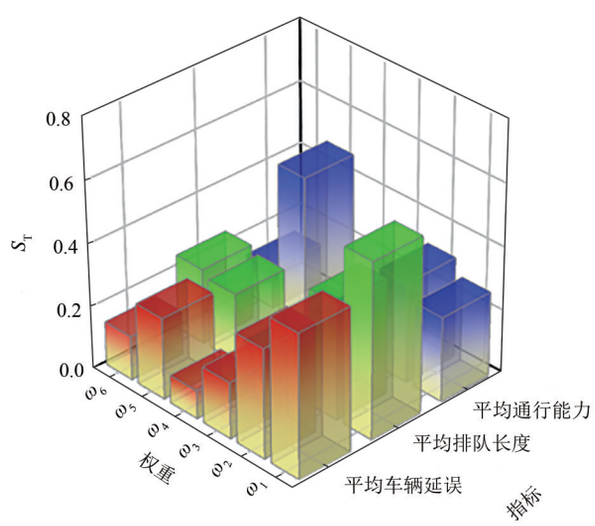
实验以经过奖励校准的 DDQN 模型为基准，仿真实验均基于经校准的 SUMO 模型开展，输入数据采用真实早高峰交通需求，单轮仿真时长为 3 600 s，并进行 1 000 s 交通流预热，训练 300 轮，评估基于 SENet 的状态空间设计、渐进式交通流学习策略以及融合这两个模块的集成方法．图 8 呈现了 3 种实验条件下奖励函数在训练回合中的演化曲线，表 8 展示了消融实验结果数据对比．

由图 8 和表 8 可知：DDQN-SENet＋Train 实验组取得了最佳的综合性能，表明渐进式交通流学习策略为模型提供了一个平滑的学习曲线，避免训练初期因直接面对高峰流量而产生的剧烈振荡和收敛困难；基于渐进式交通流训练策略，SENet 充分学习如何从交通流网格中提取与下游拥堵最相关的特征．优先经验回放机制确保了这些包含关键状态转换的经验能被高效学习，ε-贪心动作选择策略维持了必要的探索．

注意力热力图见图 9，注意力热力图直观呈现了该模块在平峰和高峰两种模拟交通场景下的关注



（a）一阶指数



（b）总效应指数

图 7 奖励函数权重 Sobol 分析

Fig.7 Sobol analysis of reward function weights

表 7 奖励函数权重的设计值

Tab.7 Design values for the reward function weights

权重符号	权重对应奖励项描述	取值
ω_1	合流区平均排队长度趋势奖励 ΔQ_t	0.30
ω_2	合流区平均车辆延误趋势奖励 ΔW_t	0.20
ω_3	合流区排队长度绝对奖励 Q_t^{norm}	0.10
ω_4	合流区车辆延误绝对奖励 W_t^{norm}	0.05
ω_5	合流区通行能力奖励 T_t^{norm}	0.30
ω_6	相位切换奖励 r_d	0.05

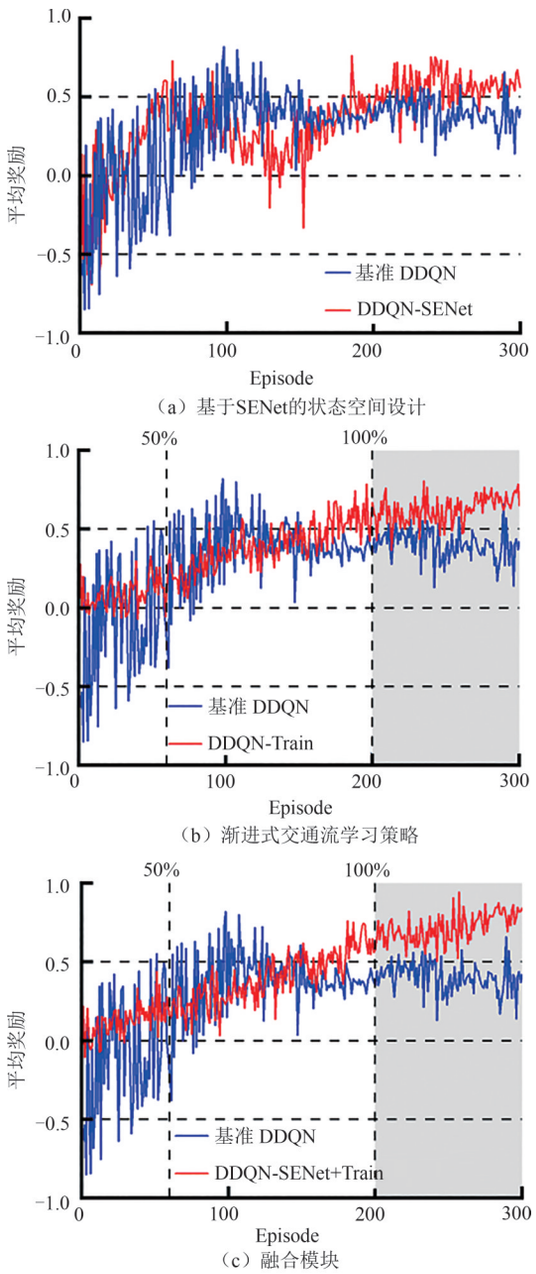


图 8　消融实验奖励函数对比

Fig.8　Comparison of reward functions in ablation experiments

区域变化.由图9可知:平峰时段,注意力分布较为广泛;当拥堵加剧至中度水平时,合流区下游区域精准形成并逐渐强化1个热点.高峰交通场景下,注意力集中于合流区尾部的严重瓶颈处.

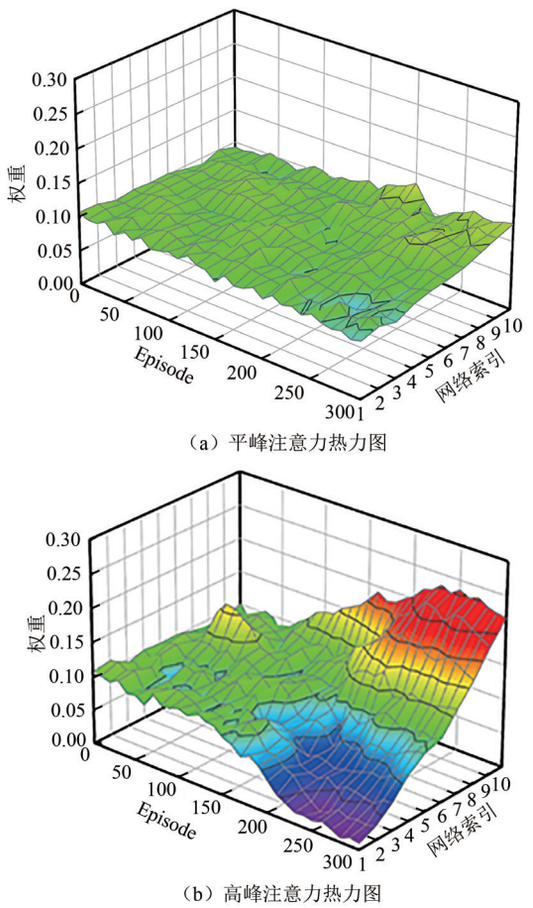


图 9　注意力热力图

Fig.9　Attention heatmap

4.3　不同模型对比

设计高峰与平峰交通流两组主要实验以验证所设计模型的泛化能力.每组实验包含3个实验方案.

方案1:基于ALINEA的原始固定配时控制(基准组).交通流输入采用渐进式交通流学习策略,传统信号控制系统则采用ALINEA进行固定参数匝道配时.

方案2:基准DDQN模型.仅保留状态与奖励函数设置,不包含模型优化策略.

方案3:优化DDQN模型.包含基于SENet的状态空间、多目标奖励架构与模型优化策略.

在实验分析中,延误率DR被定义为

表 8　消融实验结果对比

Tab.8　Comparison of ablation experiment results

实验方案	平均车辆延误率		平均通行能力		平均排队长度数值	
	数值	优化率/%	数值/(veh/步)	优化率/%	数值/m	优化率/%
基准DDQN	0.56	—	32.34	—	35.12	—
基于SENet的状态空间设计方案(DDQN-SENet)	0.49	12.5	35.67	10.3	32.42	7.7
渐进式交通流学习策略方案(DDQN-Train)	0.53	5.4	33.84	4.6	33.24	5.4
融合模块方案(DDQN-SENet+Train)	0.47	16.1	36.13	11.7	31.63	9.9

$$DR = 1 - \frac{v_{current}}{v_{max}} \qquad DR \in [0,1] \qquad (19)$$

式中：$v_{current}$ 为车辆当前实际行驶速度（由 TraCI 实时返回数据）；v_{max} 为道路最高限速；DR 为延误率，该数值越接近 1，表明合流区拥堵程度越高，越接近 0，表明合流区通行越顺畅.

图 10 为高峰和平峰状态下实验性能结果对比，最后 100 个回合对比数据如表 9 所示.

由图 10 和表 9 可知：①方案 3 在不同交通需求下展现出稳定的有效性和鲁棒性，优于方案 1.②方案 3 在所有指标上均实现较高提升，平均延误率降低 40.5%、通行能力提升 26.2%、排队长度减少 29.5%. 这些提升幅度超过方案 2 所实现的效果，证实了方案 3 在严重拥堵场景下的优越性.③在控制效率方面，方案 1 存在明显的滞后性，随着后期瓶颈效应加剧，这种滞后性会导致控制失效，使其无法满足高峰时段交通环境的优化需求.④对于方案 2，由

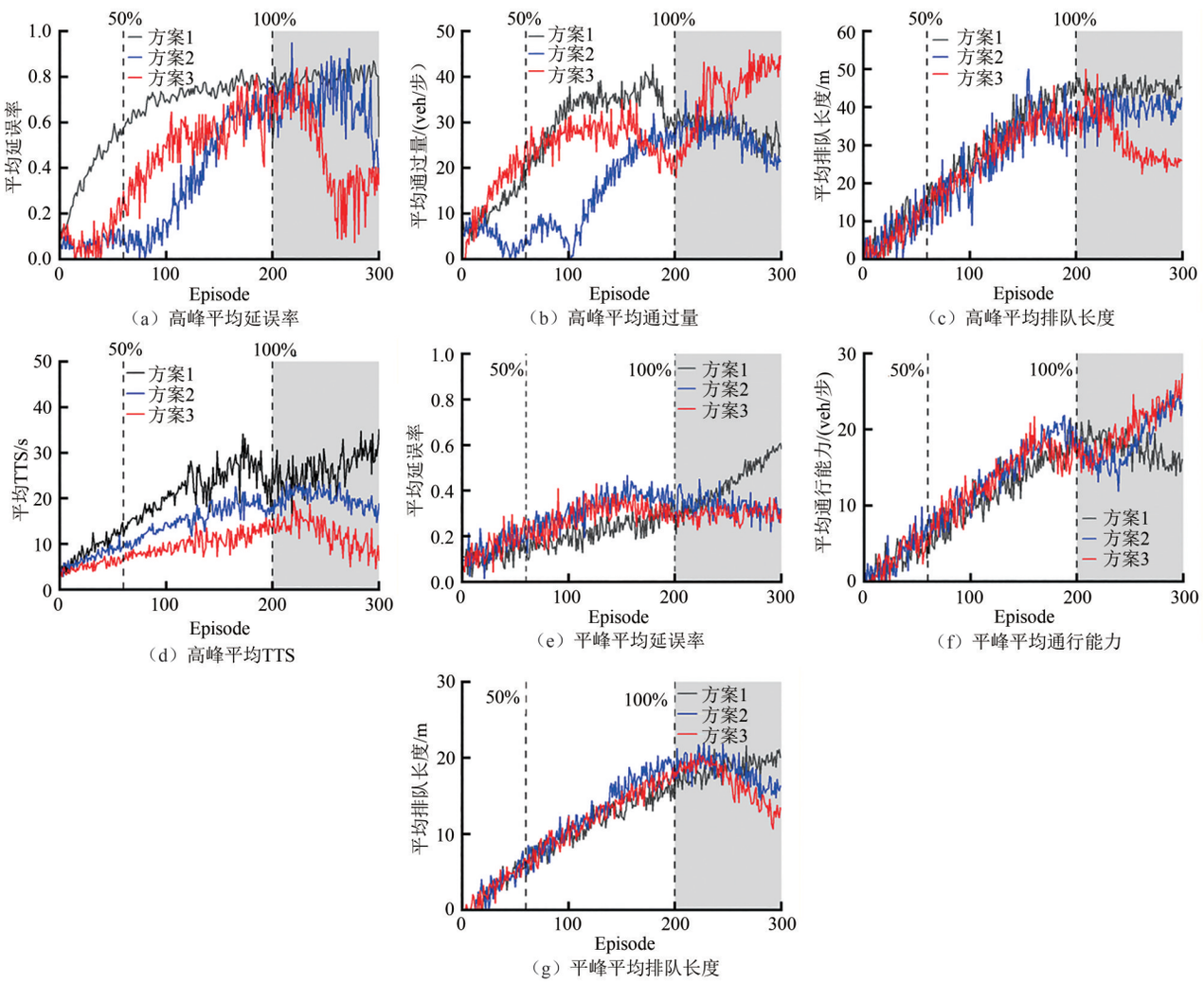


图 10　实验性能对比

Fig.10　Comparison of experimental performance

表 9　不同模型的综合数据

Tab.9　Comprehensive data of different models

流量情况	实验方案	平均延误率		平均通行能力		平均排队长度	
		数值	优化率/%	数值/(veh/步)	优化率/%	数值/m	优化率/%
高峰流量输入实验组	1	0.79	—	28.62	—	44.89	—
	2	0.56	29.1	32.34	13	35.12	21.8
	3	0.47	40.5	36.13	26.2	31.63	29.5
平峰流量输入实验组	1	0.43	—	17.5	—	18.51	—
	2	0.33	23.3	18.63	6.5	17.38	6.1
	3	0.3	30.2	20.13	15.3	16.81	9.2

于其在状态空间获取和学习策略方面存在局限性,在训练初期和中期的学习稳定性较差,相较于方案3,其对交通环境的稳定控制效果更差.⑤在平峰时段,方案3仍保持较高性能,证明其在高峰时段之外同样具备较强的适用性和有效性.⑥相比高峰不同方案TTS结果,DDQN模型缩短了合流区的出行时间,这间接表明合流区的交通效率得到提升,同时,方案3最终出行时间约为7 s,其速度接近道路限速的75%.

4.4　不同CAV渗透率对比

在状态空间设计中,主线上游与匝道的交通信息均通过CAV获取,因此控制性能会随CAV渗透率的变化而不同.考虑到CAV在低、中、高密度混合交通流中的影响,分别设置20%、40%、60%、80%共4种渗透率开展实验.不同CAV渗透率输入场景下优化DDQN模型控制效果如表10所示.

由表10可知:优化DDQN模型在不同CAV渗透率环境下均实现了理想的控制效果.在总输入流量相同的前提下,随着CAV渗透率的提高,主线上游与匝道的状态信息准确性逐步提升,进而使合流区的拥堵程度与平均延迟率相应降低,控制效果持续优化.

表 10　单一仿真回合内不同CAV渗透率下的匝道信号控制性能评估

Tab.10　Performance evaluation of ramp signal control under different CAV penetration rates in a single simulation episode

渗透率/%	平均延误率		平均通行能力		平均排队长度	
	数值	优化率/%	数值/(veh/步)	优化率/%	数值/m	优化率/%
20	0.47	—	36.13	—	31.63	—
40	0.43	8.5	39.81	10.2	28.64	9.5
60	0.37	21.3	43.21	19.6	24.87	21.47
80	0.26	44.7	48.54	34.4	18.63	41.1

5　结论

1) 优化DDQN模型构建了从交通状态感知到信号决策的闭环控制流程,实验证明其具备应对合流区动态拥堵的能力.

2) SENet注意力机制与优先经验回放机制、渐进式交通流学习策略可提升模型的感知能力、训练稳定性与收敛速度.随着拥堵加剧,SENet注意力机制能够动态聚焦于下游瓶颈,为控制决策提供了可解释的依据.

3) 实验验证该框架在高峰、平峰不同交通场景下均表现优异,控制效果优于传统固定配时控制与基准DDQN模型,在应对交通需求波动时表现出较高的适应性与控制性能.

4) 优化DDQN模型适配CAV渗透率变化,CAV提供的高精度交通数据可提升状态表征准确性,为其在智能交通场景中的应用提供扩展空间.

研究仅针对单匝道独立控制,未考虑多匝道交通流耦合效应;优化目标仅聚焦通行效率,未纳入燃油消耗、碳排放等环境指标;研究采用的轻量化SENet注意力机制有效提升了状态表征的效能,但其在显式建模空间依赖关系方面存在固有局限;CAV仅作为数据采集载体,未充分利用其可控性.尽管如此,所提出的方法为模型泛化提供了结构性基础,实验表明模型对交通需求及网联水平的变化具有良好的鲁棒性.

参考文献(References):

[1] PAPAGEORGIOU M, HADJ-SALEM H, MIDDELHAM F. ALINEA local ramp metering: Summary of field results[J]. Transportation Research Record: Journal of the Transportation Research Board, 1997, 1603(1): 90−98.

[2] SMARAGDIS E, PAPAGEORGIOU M. Series of new local ramp metering strategies: Emmanouil smaragdis and markos papageorgiou [J]. Transportation Research Record: Journal of the Transportation Research Board, 2003, 1856(1): 74−86.

[3] 刘丽琴, 邵长桥, 杨超. 基于匝道控制的快速路合流区通行能力分析[J]. 交通信息与安全, 2013, 31(3): 39−42.
LIU Liqin, SHAO Changqiao, YANG Chao. Capacity of expressway merging area based on ramp metering methodology [J]. Journal of Transport Information and Safety, 2013, 31(3): 39−42. (in Chinese)

[4] 李林恒, 李宗平, 李远辉. 合流区匝道交通量控制指标计算方法研究[J]. 交通运输工程与信息学报, 2015, 13(3): 64−69.
LI Linheng, LI Zongping, LI Yuanhui. Computational method of control indicators for ramp traffic volume in the merging area [J]. Journal of Transportation Engineering and Information, 2015, 13(3): 64−69. (in Chinese)

[5] 马明辉, 杨庆芳, 梁士栋, 等. 高速公路主线与匝道合流区协调控制方法[J]. 哈尔滨工程大学学报, 2015, 36(12): 1603−1608.
MA Minghui, YANG Qingfang, LIANG Shidong, et al. Coordination control of mainline control and ramp metering in freeway merging area [J]. Journal of Harbin Engineering University, 2015, 36(12): 1603−1608. (in Chinese)

[6] 张军，宋浪，俞山川，等. 高速公路车道分配与入口多匝道协同控制方法[J]. 公路交通技术，2023，39(3)：166－175.
ZHANG Jun，SONG Lang，YU Shanchuan，et al. A collaborative control method of lane assignment and entrance multi-ramp on expressway[J]. Technology of Highway and Transport，2023，39(3)：166－175.(in Chinese)

[7] XU Q，LIU Z G，XU Z G. A novel ramp metering algorithm based on deep reinforcement learning[C]//2022 2nd International Conference on Algorithms，High Performance Computing and Artificial Intelligence (AHPCAI). Guangzhou，2023：128－133.

[8] 丁鸿志，彭安琪，张忠宇，等. 基于深度强化学习的快速道路入口匝道信号控制方法研究[J]. 河南科学，2023，41(10)：1424－1432.
DING Hongzhi，PENG Anqi，ZHANG Zhongyu，et al. Ramp signal control method of expressway entrance based on deep reinforcement learning[J]. Henan Science，2023，41(10)：1424－1432.(in Chinese)

[9] DENG F W，JIN J D，SHEN Y，et al. Advanced self-improving ramp metering algorithm based on multi-agent deep reinforcement learning[C]//2019 IEEE Intelligent Transportation Systems Conference (ITSC). Auckland，2019：3804－3809.

[10] CHENG J W，YE C L，WANG N N，et al. A deep reinforcement learning based ramp metering control method considering ramp outflow[J]. IFAC-PapersOnLine，2024，58(10)：200－205.

[11] HAN Y，WANG M，LI L H，et al. A physics-informed reinforcement learning-based strategy for local and coordinated ramp metering[J]. Transportation Research Part C：Emerging Technologies，2022，137：103584.

[12] BINJAKU K，MEÇE E K，PASQUALE C，et al. AI-based predictive ramp-metering control for freeway traffic systems[C]//2024 IEEE 27th International Conference on Intelligent Transportation Systems (ITSC). Edmonton，2025：870－875.

[13] AIRALDI F，DE SCHUTTER B，DABIRI A. Reinforcement learning with model predictive control for highway ramp metering[J]. IEEE Transactions on Intelligent Transportation Systems，2025，26(5)：5988－6004.

[14] 章立辉，余宏鑫，熊满初，等. 平衡主线和匝道交通运行的强化学习型匝道控制研究[J]. 重庆交通大学学报(自然科学版)，2023，42(4)：87－97.
ZHANG Lihui，YU Hongxin，XIONG Manchu，et al. Reinforcement learning ramp metering to balance mainline and ramp traffic operations[J]. Journal of Chongqing Jiaotong University (Natural Sciences)，2023，42(4)：87－97.(in Chinese)

[15] 戴昇宏，李志斌. 基于图像卷积神经网络的匝道控制深度强化学习算法研究[J]. 交通工程，2019，19(4)：1－6.
DAI Shenghong，LI Zhibin. Research on ramp metering based on deep reinforcement learning with image convolutional neural network[J]. Journal of Transportation Engineering，2019，19(4)：1－6.(in Chinese)

[16] YU S X，HAN Y. Coordinated ramp metering strategy based on deep reinforcement learning incorporating attention mechanism[J]. IEEE Transactions on Intelligent Transportation Systems，2026，27(2)：2794－2806.

[17] YANG S T，YANG B. An inductive heterogeneous graph attention-based multi-agent deep graph infomax algorithm for adaptive traffic signal control[J]. Information Fusion，2022，88：249－262.

[18] 任安妮，周大可，冯锦浩，等. 基于注意力机制的深度强化学习交通信号控制[J]. 计算机应用研究，2023，40(2)：430－434.
REN Anni，ZHOU Dake，FENG Jinhao，et al. Attention mechanism based deep reinforcement learning for traffic signal control[J]. Application Research of Computers，2023，40(2)：430－434.(in Chinese)

[19] 张玺君，聂生元，李喆，等. 基于自注意力机制的深度强化学习交通信号控制[J]. 交通运输系统工程与信息，2024，24(2)：96－104.
ZHANG Xijun，NIE Shengyuan，LI Zhe，et al. Traffic signal control with deep reinforcement learning and self-attention mechanism[J]. Journal of Transportation Systems Engineering and Information Technology，2024，24(2)：96－104.(in Chinese)

[20] SHANG M F，WANG S A，STERN R E. Extending ramp metering control to mixed autonomy traffic flow with varying degrees of automation[J]. Transportation Research Part C：Emerging Technologies，2023，151：104119.

[21] LI L H，QIAN C，GAN J，et al. DCoMA：a dynamic coordinative merging assistant strategy for on-ramp vehicles with mixed traffic conditions[J]. Transportation Research Part C：Emerging Technologies，2024，165：104700.

[22] ZARE A，STERN R. Adapting ramp metering control for mixed autonomy traffic[C]//2025 IEEE Conference on Control Technology and Applications (CCTA). San Diego，2025：584－589.

[23] 蔡科，聂振宇，马庆禄，等. 智能网联环境下高速公路匝道合流区域协同控制[J]. 兰州交通大学学报，2022，41(1)：25－33.
CAI Ke，NIE Zhenyu，MA Qinglu，et al. Collaborative control of highway ramp confluence area under intelligent network environment[J]. Journal of Lanzhou Jiaotong University，2022，41(1)：25－33.(in Chinese)

[24] 王秋玲, 赵祥模, 徐志刚, 等. 特种CAV优先的集中式匝道合流协同控制方法[J]. 交通运输工程学报, 2022, 22(1): 263－272.
WANG Qiuling, ZHAO Xiangmo, XU Zhigang, et al. Centralized ramp confluence cooperative control method with special connected and automated vehicle priority[J]. Journal of Traffic and Transportation Engineering, 2022, 22(1): 263－272. (in Chinese)

[25] 李春, 吴志周, 曾广, 等. 合流区智能网联汽车协同控制方法综述[J]. 计算机工程与应用, 2024, 60(12): 1－17.
LI Chun, WU Zhizhou, ZENG Guang, et al. Review of connected autonomous vehicle cooperative control at on-ramp merging areas[J]. Computer Engineering and Applications, 2024, 60(12): 1－17. (in Chinese)

[26] WEI H, ZHENG G J, YAO H X, et al. IntelliLight: a reinforcement learning approach for intelligent traffic light control [C]//Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. London ,2018: 2496－2505.

[27] MNIH V, KAVUKCUOGLU K, SILVER D, et al. Human-level control through deep reinforcement learning[J]. Nature, 2015, 518(7540): 529－533.

[28] DU Y, SHANGGUAN W, RONG D C, et al. RA-TSC: learning adaptive traffic signal control strategy *via* deep reinforcement learning[C]//2019 IEEE Intelligent Transportation Systems Conference (ITSC). Auckland, 2019: 3275－3280.

[29] MORRIS M D. Factorial sampling plans for preliminary computational experiments[J]. Technometrics, 1991, 33(2): 161－174.

[30] MCKAY M D, BECKMAN R J, CONOVER W J. A comparison of three methods for selecting values of input variables in the analysis of output from a computer code[J]. Technometrics, 2000, 42(1): 55－61.

[31] 吴鉴庭, 佟宁, 董芃. 深度强化学习驱动的应急车辆优先信号控制策略[J]. 软件导刊, 2025, 24(12): 8－17.
WU Jianting, TONG Ning, DONG Peng. Deep reinforcement learning-driven priority signal control strategy for emergency vehicles [J]. Software Guide, 2025, 24(12): 8－17. (in Chinese)

[32] 张浩然, 王嘉文, 周丽萍. 基于深度强化学习PPO的匝道混合交通流合流控制方法[J]. 交通运输工程与信息学报, 2026, 24(1): 116－130.
ZHANG Haoran, WANG Jiawen, ZHOU Liping. Deep reinforcement learning PPO-based ramp merging control approach for mixed traffic flow[J]. Journal of Transportation Engineering and Information, 2026, 24(1): 116－130. (in Chinese)

[33] SOBOL L M. Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates[J]. Mathematics and Computers in Simulation, 2001, 55(1/2/3): 271－280.

[34] SALTELLI A, ANNONI P, AZZINI I, et al. Variance based sensitivity analysis of model output. Design and estimator for the total sensitivity index[J]. Computer Physics Communications, 2010, 181(2): 259－270.