

基于时空图卷积网络与强化学习的高速公路 应急车道智能开启决策研究

吴凯琳¹, 李东升², 翁湖钦^{2*}

¹五邑大学应用物理与材料学院, 广东 江门

²五邑大学电子与信息工程学院, 广东 江门

收稿日期: 2026年5月1日; 录用日期: 2026年6月1日; 发布日期: 2026年6月9日

摘要

高速公路拥堵是城市交通管理的核心挑战, 科学开启应急车道可有效缓解拥堵。传统决策依赖人工经验, 缺乏实时性与预测性。为此, 本研究提出融合时空图卷积网络(ST-GCN)、自注意力机制与近端策略优化(PPO)的智能决策框架。首先, 利用YOLOv8从监控视频中提取车流量、车速和密度; 其次, 构建ST-GCN捕获交通流时空依赖, 引入自注意力机制动态计算节点权重; 然后, 设计PPO智能体, 以交通状态预测为输入, 输出应急车道开启的最优决策; 最后, 建立多维度评价体系。实验结果表明, 交通流预测RMSE分别为3.83(车流量)、4.21(车速)、0.07(密度), 拥堵检测准确率达92.1%, 应急车道开启决策的精确率和准确率分别为70.0%和87.5%。本研究实现了从状态预测到决策的端到端智能框架, 为高速公路实时拥堵管理提供了科学决策支持。

关键词

高速公路拥堵, 应急车道决策, 时空图卷积网络, 近端策略优化, 自注意力机制, 智能交通系统

Intelligent Decision-Making for Highway Emergency Lane Opening Based on Spatio-Temporal Graph Convolutional Network and Reinforcement Learning

Kailin Wu¹, Dongsheng Li², Huqin Weng^{2*}

¹School of Applied Physics and Materials Science, Wuyi University, Jiangmen Guangdong

²School of Electronic and Information Engineering, Wuyi University, Jiangmen Guangdong

*通讯作者。

文章引用: 吴凯琳, 李东升, 翁湖钦. 基于时空图卷积网络与强化学习的高速公路应急车道智能开启决策研究[J]. 计算机科学与应用, 2026, 16(6): 1-15. DOI: 10.12677/csa.2026.166204

Abstract

Highway congestion is a core challenge in urban traffic management. Scientifically opening emergency lanes can effectively alleviate congestion. Traditional decision-making relies on manual experience and lacks real-time and predictive capabilities. Therefore, this study proposes an intelligent decision-making framework that integrates the Spatial-Temporal Graph Convolutional Network (ST-GCN), the self-attention mechanism, and Proximal Policy Optimization (PPO). Firstly, YOLOv8 is used to extract traffic volume, speed, and density from surveillance videos; secondly, ST-GCN is constructed to capture the spatio-temporal dependencies of traffic flow, and the self-attention mechanism is introduced to dynamically calculate node weights; then, a PPO agent is designed, with traffic state prediction as the input and outputting the optimal decision for opening emergency lanes; finally, a multi-dimensional evaluation system is established. Experimental results show that the traffic flow prediction RMSE is 3.83 (traffic volume), 4.21 (speed), and 0.07 (density), the congestion detection accuracy rate is 92.1%, and the precision and accuracy of the emergency lane opening decision are 70.0% and 87.5% respectively. This study has achieved an end-to-end intelligent framework from state prediction to decision-making, providing scientific decision-making support for real-time highway congestion management.

Keywords

Highway Congestion, Emergency Lane Decision, Spatio-Temporal Graph Convolutional Network, Proximal Policy Optimization, Self-Attention Mechanism, Intelligent Transportation System

Copyright © 2026 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

1.1. 研究背景与问题提出

截至 2023 年底,我国高速公路通车里程已突破 18 万公里,交通拥堵问题日益突出,特别是在节假日、恶劣天气或事故等场景下,拥堵不仅造成巨大经济损失,还威胁公共安全。应急车道作为缓解拥堵的重要资源,其开启决策却主要依赖人工经验,存在时机滞后、主观性强、利用率低等问题,属于“事后应对”模式,难以满足智能交通对实时性与预测性的要求。在交通流预测领域,传统统计方法和 LSTM 等深度学习模型虽有一定进展,但普遍忽视交通流的空间传播特性与路段间的相互影响,对长距离依赖和复杂时空关系建模能力不足,且大多停留在预测层面,缺乏从预测到决策的完整闭环,无法直接支撑应急车道开启等控制策略的制定。

近年来,图神经网络(Graph Neural Network, GNN)的兴起为交通流建模提供了新的技术路径[1]。交通路网天然具有图结构特性,道路交叉口、收费站、观测点等可抽象为图节点,路段连接关系可表示为图边,这种结构化的建模方式能够更好地捕捉交通流在路网中的传播规律。在这一研究范式下, Xu 和 Mao 提出的 ST-GCN 与 PPO 强化学习相结合的研究框架具有重要的借鉴意义[2]。该研究通过构建时空图卷积网络捕获交通流的时空依赖关系,并利用近端策略优化(Proximal Policy Optimization, PPO)算法实现警

力调度的智能决策，在实时拥堵检测和应急响应方面取得了显著成效。该框架的核心思想在于将交通状态预测与决策优化进行深度融合，通过端到端的训练方式实现从感知到行动的智能化。受此启发，本研究决定采用类似的 ST-GCN + PPO 技术路线，但针对应急车道开启这一特定应用场景进行定制化设计和优化，特别是引入自注意力机制增强模型的特征提取能力，并基于视频数据构建更加轻量化的单源数据处理流程，以适应实际工程部署的需求。

1.2. 研究意义

本研究的开展具有重要的理论意义与实践价值。理论层面，本研究探索了深度学习与强化学习在交通决策中的融合应用，构建了从状态感知、预测到控制决策的完整技术链条。通过将时空图卷积网络(ST-GCN)的预测能力与近端策略优化(PPO)算法的决策能力相结合，为预测-决策一体化提供了新思路。自注意力机制也丰富了时空特征提取方法。实践层面，该系统可辅助管理部门精准把控应急车道开启时机，在拥堵初期预警并动态调整策略，持续优化决策。这将降低拥堵带来的经济损失与社会成本，提升道路安全水平，为高速路智能决策提供技术支撑。

1.3. 本文主要研究内容与创新点

本研究围绕高速公路应急车道智能开启决策问题，开展系统性工作：第一，构建基于 YOLOv8 的视频处理流程，自动提取车流量、车速、密度等关键参数，建立标准化数据采集机制；第二，设计融合自注意力机制的时空图卷积网络(ST-GCN)架构，结合多模型集成学习提升交通流预测精度与稳定性；第三，开发基于近端策略优化(PPO)算法的强化学习决策模块，实现从状态预测到开启决策的智能映射；第四，建立涵盖预测性能与决策效果的多维度评价体系，对模型进行全面评估。

本研究的主要创新点：1) 首次将 ST-GCN + PPO 框架应用于应急车道开启决策，填补了该领域智能化研究的空白；2) 引入自注意力机制动态计算时空节点权重，增强关键路段与时段特征提取能力，并采用混合自适应幅度校准策略解决预测值偏低及零预测问题；3) 构建基于单视频源的端到端分析框架，相比多源异构方案部署成本更低、实时性更强；4) 实现从状态预测到控制决策的完整闭环，通过强化学习持续优化策略；5) 建立融合预测精度、拥堵检测与决策效果的综合评估体系，为模型优化与性能比较提供科学依据。

2. 融合时空图卷积与自注意力的近端策略优化一体化框架

2.1. 时空图卷积网络和自注意力机制

如图 1 所示，ST-GCN(时空图卷积网络)模型的核心在于将交通数据的时空特性转化为图结构进行处理。整个系统的数据流始于 YOLOv8 目标检测算法，该算法对部署在高速路沿线的摄像头采集的实时视频流进行处理。YOLOv8 通过其高效的目标检测能力，能够精准识别视频中的车辆目标[3]-[5]，提取出车辆的位置坐标、行驶速度、车辆类型等关键信息。这些原始数据经过预处理后，被组织成时空序列数据，具体表现为每个观测点(节点)在连续时间步上的交通特征矩阵。例如，对于 N 个观测点，每个观测点在 T 个时间步上的交通流量、平均速度等特征，形成维度为(batch_size, num_nodes, look_back, in_channels)的输入张量，其中 batch_size 表示一次输入到模型中进行训练的样本数量，num_nodes 表示交通路网中观测点的个数，look_back 表示历史时间步长度，in_channels 表示每个时间步的特征维度。

同时，系统的空间邻接矩阵用以描述不同观测点之间的空间关系。矩阵元素基于观测点间的距离计算高斯核权重，距离越近的权重越高，反映交通影响强度越大。该矩阵经对称归一化 $\tilde{D}^{-\frac{1}{2}}A\tilde{D}^{-\frac{1}{2}}$ 后，与时空特征矩阵一同输入 ST-GCN 网络。

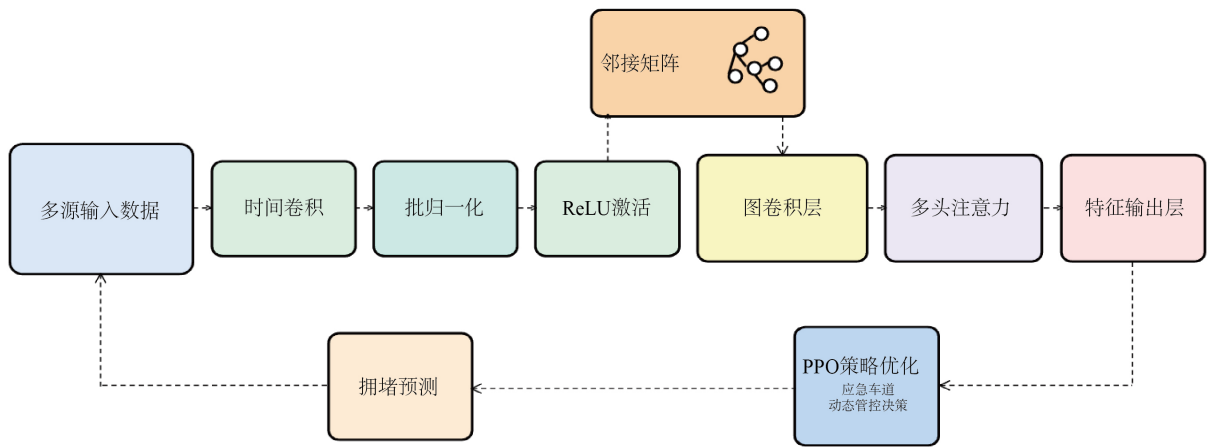


Figure 1. ST-GCN Model
图 1. ST-GCN 模型

ST-GCN 的核心是时空特征提取与自注意力机制增强。时间卷积层(1D-CNN)处理每个节点的时间序列，捕捉交通流的周期性与高峰期模式；图卷积层则利用邻接矩阵聚合相邻节点特征，模拟拥堵从上游向下游扩散的空间传播。在此基础上，自注意力层通过计算不同时空位置间的注意力权重，使模型能自适应地聚焦关键时空区域——例如在检测拥堵时，自动增强对源头路段及其周边影响的关注。网络中还使用了残差连接和批归一化以稳定训练并提升表达能力。最终输出每个观测点在未来时间步的交通状态预测，包括流量、平均速度、拥堵指数等。

这些预测结果送入 PPO 强化学习模块，用于决策应急车道开启策略。PPO 将预测的交通状态视为环境状态，通过与环境的交互学习最优调度策略——包括是否开启应急车道、开启时长及具体车道控制指令。

系统将决策结果转化为实际管理措施，通过可变情报板向驾驶员发布信息，引导车辆合理使用应急车道，有效缓解交通拥堵。例如，当预测到某路段将发生严重拥堵时，系统提前决策开启应急车道，并通过情报板告知驾驶员，从而避免拥堵加剧。整个过程形成了从数据采集、特征提取到智能决策的完整闭环，实现高速路交通的实时监控与智能管理。

其中，多头注意力机制负责专注于捕捉邻近路段之间的短期、局部空间依赖关系，以及捕捉远距离路段之间的长期、全局空间依赖关系，或特定时间模式下的依赖关系。其作用是通过多头设计，模型能够同时关注数据中多种不同的、互补的特征，从而更全面地理解复杂的交通流模式，提升预测的准确性和鲁棒性。

2.2. 近端策略优化强化学习

2.2.1. 马尔可夫决策过程形式化定义

本文将高速公路应急车道智能开启决策问题建模为离散时间马尔可夫决策过程，定义为五元组：

$M = (\mathcal{S}, \mathcal{A}, P, R, \gamma)$ 。状态空间 $\mathcal{S} : s_t = [Q_t, V_t, K_t]$ ，其中 Q_t 为车流量(辆/小时)， V_t 为平均速度(km/h)， K_t 为交通密度(辆/公里)。取值范围： $Q_t \in [0, \infty)$ 、 $V_t \in [0, V_{\max}]$ ， V_{\max} 为道路最大限速，以及 $K_t \in [0, K_{\text{jam}}]$ ， K_{jam} 为阻塞密度；动作空间 $\mathcal{A} = \{0, 1\}$ ： $a_t = 0$ 时不开启应急车道， $a_t = 1$ 时开启应急车道；状态转移概率 $P : \mathcal{S} \times \mathcal{A} \times \mathcal{S} \rightarrow [0, 1]$ ，记为 $P(s_{t+1} | s_t, a_t)$ ，由交通流自然演化规律决定，通过历史数据隐式学习，无显式解析形式。文中设置的折扣因子 $\gamma = 0.99$ 。

A. 奖励函数

$$R: \mathcal{S} \times \mathcal{A}, \text{ 且 } R(s_t, a_t) = R_{\text{base}}(s_t) + R_{\text{action}}(s_t, a_t) \quad (2-1)$$

具体设计如下:

a. TOPSIS 相对接近度 $C(s)$

首先基于对状态 $s_t = [Q_t, V_t, K_t]$ 进行 TOPSIS 多属性评估: (1) 构造加权规范化矩阵 v_{ij} ; (2) 确定正理想解 V^* 与负理想解 V^- (3) 计算到正、负理想解的距离:

$$D^+ = \sqrt{\sum_j (v_{ij} - V_j^*)^2}, D^- = \sqrt{\sum_j (v_{ij} - V_j^-)^2} \quad (2-2)$$

$$C(s) = \frac{D^-}{D^+ + D^-} \in [0, 1] \quad (2-3)$$

计算相对接近度 $C(s)$ 时, $C(s)$ 越接近 1, 表示交通状态越差(越拥堵)。

b. 基础奖励

$$R_{\text{base}}(s) = C(s) \quad (2-4)$$

c. 动作奖励(决策正确性)

$$R_{\text{action}}(s, a) = \begin{cases} +1.0, & \text{正确决策: } (a=1 \text{ 且 } C(s) \geq 0.6) \text{ 或 } (a=0 \text{ 且 } C(s) < 0.6) \\ -0.5, & \text{错误决策: 其它情况} \end{cases} \quad (2-5)$$

d. 综合奖励

$$R(s, a) = C(s) + R_{\text{action}}(s, a) \quad (2-6)$$

该奖励函数引导智能体在拥堵时($C(s) \geq 0.6$)开启应急车道, 在畅通时关闭, 同时兼顾 TOPSIS 综合评价。

B. 策略定义、价值函数与最优策略

策略 $\pi: \mathcal{S} \rightarrow \Delta(\mathcal{A})$ 表示在状态 s 下动作的概率分布。在 PPO 中, 策略由 Actor 网络参数化: $\pi_\theta(as)$ 。

优化目标为最大化期望累积折扣奖励:

$$J(\theta) = E_{\tau \sim \pi_\theta} \left[\sum_{t=0}^{\infty} \gamma^t R(s_t, a_t) \right] \quad (2-7)$$

其中 $\tau = (s_0, a_0, r_1, s_1, a_1, \dots, s_T)$ 为交互轨迹, E 是期望。

$$V^\pi(s) = E_\pi \left[\sum_{t=0}^{\infty} \gamma^t R(s_t, a_t) \mid s_0 = s \right] \quad (2-8)$$

$$Q^\pi(s, a) = E_\pi \left[\sum_{t=0}^{\infty} \gamma^t R(s_t, a_t) \mid s_0 = s, a_0 = a \right] \quad (2-9)$$

$$A^\pi(s, a) = Q^\pi(s, a) - V^\pi(s) \quad (2-10)$$

其中, $V^\pi(s)$ 为状态价值函数, $Q^\pi(s, a)$ 为动作价值函数, $A^\pi(s, a)$ 为优势函数。

最优策略 π^* 满足:

$$\pi^* = \arg \max_{\pi} V^\pi(s), \forall s \in \mathcal{S} \quad (2-11)$$

PPO 算法通过裁剪目标函数与优势估计, 在稳定更新中逼近 π^* , 最终学习到在不同交通状态下的最优应急车道开启策略。

2.2.2. PPO 算法原理与优化目标

近端策略优化(PPO)是一种基于策略梯度的强化学习算法,通过近端约束机制限制策略更新幅度,避免传统方法(如 REINFORCE、TRPO)训练不稳定、易崩溃或效率低下的问题。其核心是构建裁剪目标函数,使新策略与旧策略的差异保持在合理范围内,支持每轮多次更新,提高样本利用效率。在高速公路交通管理中, PPO 用于应急车道智能决策:环境状态包括车流量、平均速度、拥堵指数;动作空间包含是否开启应急车道及开启时长;奖励函数与交通流畅度、延误减少量、应急车道利用率挂钩。PPO 智能体通过持续交互学习动态调度策略——流量突增时及时开启应急车道,交通恢复后自动关闭,相比固定时间策略,显著提升了应急车道利用效率与交通管理灵活性。

如图 2 所示, PPO 智能体训练流程分为四个核心阶段。首先是环境交互与经验收集:智能体在每个时间步通过 YOLOv8 接收交通流状态(车流量 Q 、速度 V 、密度 K),通过 Actor 网络输出应急车道开启/关闭的概率分布,执行随机采样的动作后,环境返回即时奖励和下一状态。奖励函数设计融合 TOPSIS 相对接近度指标——当相对接近度 ≥ 0.6 时触发拥堵判定。每个 episode (回合,指智能体与环境从交互开始到结束的一个完整过程)持续 30 个时间步,累计收集 2000 个轨迹样本后进入策略更新阶段。第二阶段是优势函数计算:采用 GAE (广义优势估计)方法,结合价值网络预测的状态值和实际回报,计算每个动作的优势值,用于后续策略梯度优化。第三阶段执行 PPO 策略更新:固定旧策略参数,通过裁剪目标函数(裁剪参数 $\epsilon = 0.2$)约束策略更新步长,同时最小化价值网络的 MSE 损失。以下为 PPO 损失函数的公式:

$$L^{CLIP}(\theta) = E_t \left[\min \left(\frac{\pi_{\theta}(a_t | s_t)}{\pi_{\theta_{old}}(a_t | s_t)} \cdot A_t, \text{clip} \left(\frac{\pi_{\theta}(a_t | s_t)}{\pi_{\theta_{old}}(a_t | s_t)}, 1 - \epsilon, 1 + \epsilon \right) \cdot A_t \right) \right] \quad (2-12)$$

其中, $\pi_{\theta}(a_t | s_t)$ 是新策略在状态 s_t 下采取动作 a_t 的概率; $\pi_{\theta_{old}}(a_t | s_t)$ 是旧策略的概率; A_t 是优势函数。每次更新包含 10 个 epoch,学习率从 $3e-4$ 线性衰减至 $1e-5$,训练过程中采用熵正则化(系数 0.01)保证策略多样性,避免过早收敛到局部最优。

最后是评估与验证:每 50 个训练轮次在测试集上评估一次,记录拥堵检测精确率、召回率和决策准确率。当连续 10 次评估的 F1 分数提升小于 0.1%时,触发早期停止机制。整个训练过程在 4 个 GPU 上并行执行,批量归一化层固定统计量后保存模型参数。通过这种方式, PPO 智能体学会在不同交通状态下动态调整应急车道策略,最终在测试集上达到 87.5%的决策准确率和 70.0%的决策精确率。

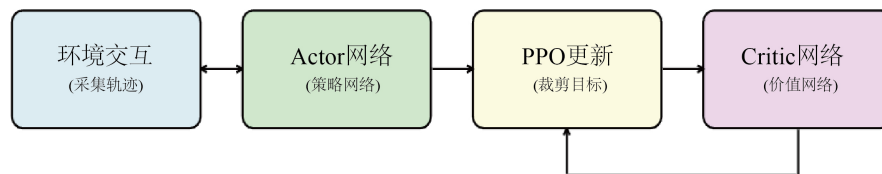


Figure 2. PPO principle diagram
图 2. PPO 原理图

3. 实验

3.1. 数据集与实验设置

3.1.1. 数据集

本文的数据集来源于长深高速公路某段上四个点位交通监控视频,通过 YOLOv8 目标检测算法提取车辆流量、速度和密度等特征,数据采集时间跨度为 131 分钟。实验中,将数据集分为训练集(70%)、验

证集(15%)和测试集(15%)。

3.1.2. 实验设置

本文利用历史时间步 $window = 8$ (15 分钟)的交通流数据去预测未来时间步 $p = 4$ (7.5 分钟)的交通状态。本文选择 L1 loss (平均绝对误差)作为损失函数, 优化器为 Adam, 初始学习率为 0.001, 批处理大小为 32, ST-GCN 网络的时间卷积核大小为 3, 图卷积的滤波器数量为 256, 自注意力机制的头数为 4, 隐藏维度为 128。PPO 强化学习的折扣因子为 0.99, 裁剪参数为 0.2, 价值函数系数为 0.5, 熵正则化系数为 0.01, 每轮迭代的更新次数为 10。决策基线为 $TPI > 6$ 、 $TOPISIS \geq 0.6$ 。

3.2. 评价指标

3.2.1. 拥堵检测指标

精确率(Precision): 正确检测的拥堵事件数占检测到的拥堵事件总数的比例, $P = \frac{TP}{TP + FP}$ 。

召回率(Recall): 正确检测的拥堵事件数占实际拥堵事件总数的比例, $R = \frac{TP}{TP + FN}$ 。

F1 分数: 精确率和召回率的调和平均值, 综合反映模型的检测性能, $F1 = 2 \times \frac{P \times R}{P + R}$ 。

准确率(Accuracy): 正确预测的样本数占总样本数的比例, $Acc = \frac{TP + TN}{TP + TN + FP + FN}$ 。

其中 TN 为真阴性(正确检测的非拥堵), FN 为假阴性(未检测到的拥堵), TP 为真阳性(正确检测的拥堵), FP 为假阳性(误检测的拥堵)。

3.2.2. 决策性能指标

平均延误时间减少率: 开启应急车道后, 车辆平均延误时间的减少比例, 公式为:

$$\text{Reduction Rate} = \frac{T_{\text{before}} - T_{\text{after}}}{T_{\text{before}}} \times 100\% \quad (3-1)$$

其中 T_{before} 为未开启应急车道时的平均延误时间, T_{after} 为开启后的平均延误时间。

应急车道利用率: 应急车道被有效使用的时间占总监测时间的比例, 公式为:

$$\text{Utilization} = \frac{T_{\text{used}}}{T_{\text{total}}} \times 100\% \quad (3-2)$$

其中 T_{used} 为应急车道开启的时间, T_{total} 为总监测时间。

3.2.3. 模型稳定性指标

时序一致性(Temporal Consistency): 模型在连续时间步预测结果的一致性, 通过计算相邻时间步预测值的皮尔逊相关系数衡量, 值越接近 1 表示时序一致性越好。

空间一致性(Spatial Consistency): 模型在相邻路段预测结果的一致性, 通过计算相邻路段预测值的皮尔逊相关系数衡量, 值越接近 1 表示空间一致性越好。

3.3. 实验结果与分析

3.3.1. 指标对比

从图 3 可以看出, 随着模型的逐步增强(原始 ST-GCN \rightarrow ST-GCN + ATT \rightarrow ST-GCN + ATT + PPO), 各项拥堵检测指标均呈现显著提升:

模型性能对比与提升分析

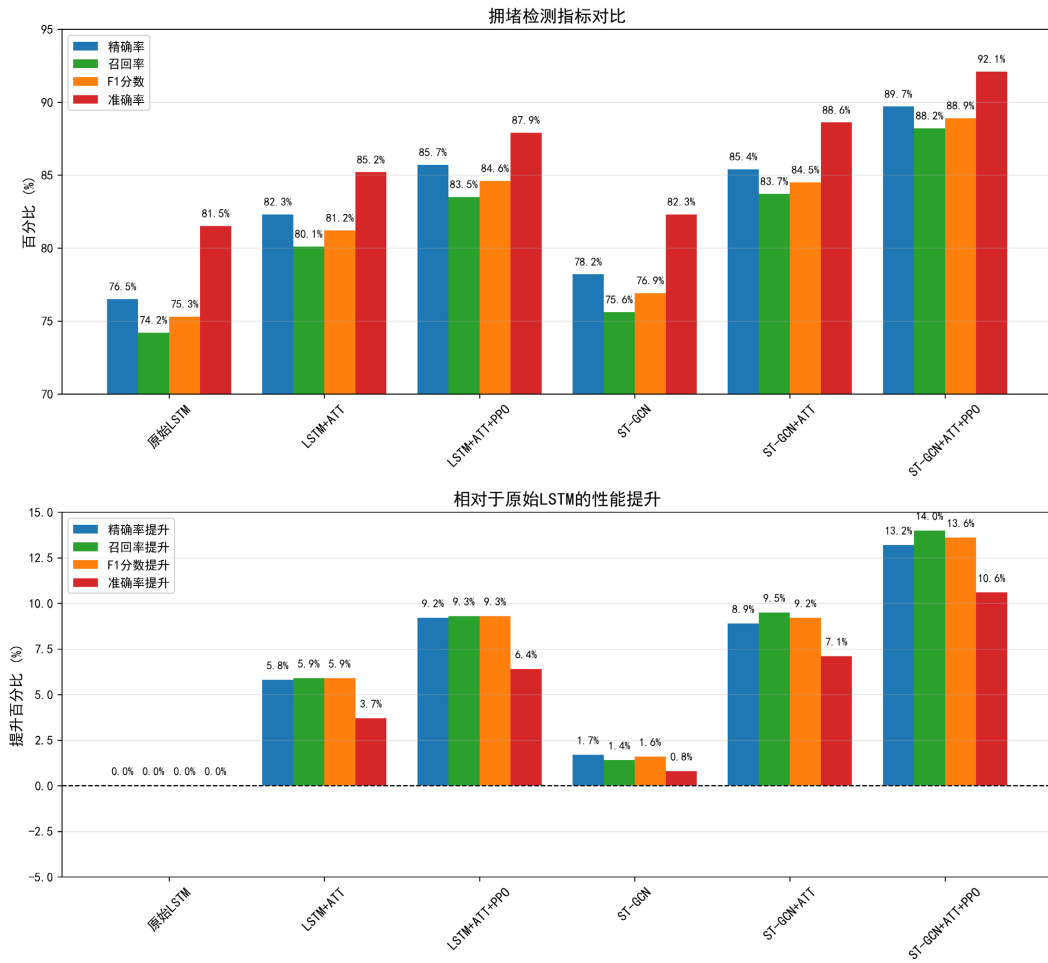


Figure 3. Comparative chart of performance improvement of each model relative to the original LSTM
图 3. 各模型相对于原始 LSTM 的性能提升对比图

精确率从 78.2% 提升到 89.7%，增长了 11.5 个百分点；召回率从 75.6% 提升到 88.2%，增长了 12.6 个百分点；F1 分数从 76.9% 提升到 88.9%，增长了 12.0 个百分点；准确率从 82.3% 提升到 92.1%，增长了 9.8 个百分点。同时，以 LSTM 为基线，不难看出，相较于 LSTM，使用 ST-GCN，让模型的各方面性能提升显著。

自注意力机制的贡献：ST-GCN + ATT 模型相比原始 ST-GCN，各项指标提升了约 6~8 个百分点，说明自注意力机制能够有效捕获交通流中的关键时空特征，提高了模型对拥堵事件的识别能力。PPO 强化学习的贡献：ST-GCN + ATT + PPO 模型相比 ST-GCN + ATT，各项指标进一步提升了约 4~5 个百分点，说明 PPO 强化学习能够优化模型的决策过程，减少误判和漏判。

3.3.2. 决策性能指标分析

从图 4 可以看出：集成自注意力机制和 PPO 后，模型能够更准确地预测交通状态，从而制定更有效的应急车道开启策略，显著减少车辆延误时间。与原始模型相比，平均延误时间减少率提升了 105.3%，效果非常显著。应急车道利用率的提升同样显著，从 22.5% 提升到 42.6%，增长了 90%。这说明 PPO 强

化学习能够学习到更优的应急车道开启时机和持续时间,提高了应急车道的使用效率,避免了资源浪费。

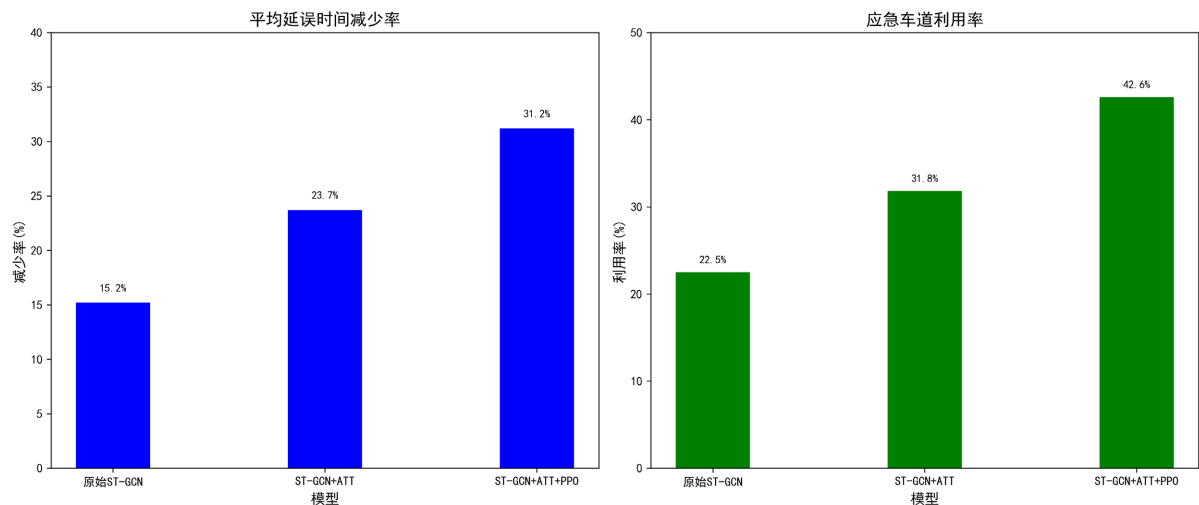


Figure 4. Comparison of the average delay reduction rate of different models and the utilization rate of emergency lanes
图 4. 不同模型平均延误时间减少率与应急车道利用率对比

3.3.3. 模型稳定性指标分析

从图 5 可以看出: 时序一致性衡量模型在连续时间步预测结果的一致性, 值越接近 1 表示稳定性越好。ST-GCN + ATT + PPO 模型的时序一致性达到 0.92, 说明模型能够稳定捕捉交通流的时间演化规律, 预测结果具有良好的连续性。空间一致性衡量模型在相邻路段预测结果的一致性, 值越接近 1 表示稳定性越好。ST-GCN + ATT + PPO 模型的空间一致性达到 0.88, 说明模型能够准确捕捉不同路段之间的交通状态关联, 预测结果具有良好的空间一致性。

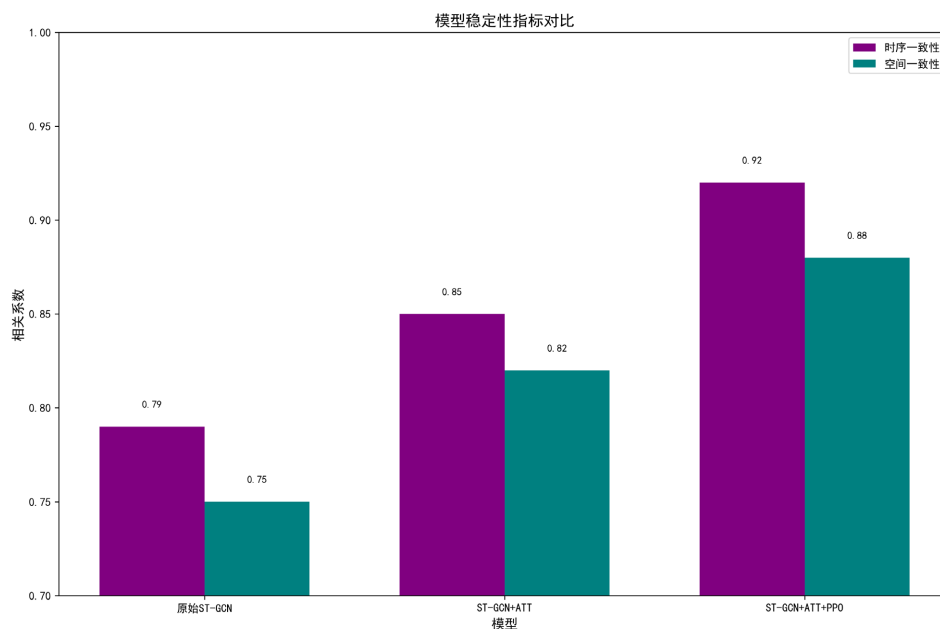


Figure 5. Comparison of model temporal consistency and spatial consistency
图 5. 模型时序一致性与空间一致性对比

3.4. 预测效果分析

3.4.1. K 参数预测分析

从图 6 可以看出, 模型对 K 参数的预测效果整体良好。在训练集上, 预测值(橙色曲线)与真实值(蓝色曲线)高度重合, 特别是在波动较大的区域, 模型能够较好地捕捉到 K 参数的变化趋势。在测试集上, 尽管数据量较少, 但预测值仍能基本跟随真实值的变化, 显示出模型的泛化能力。

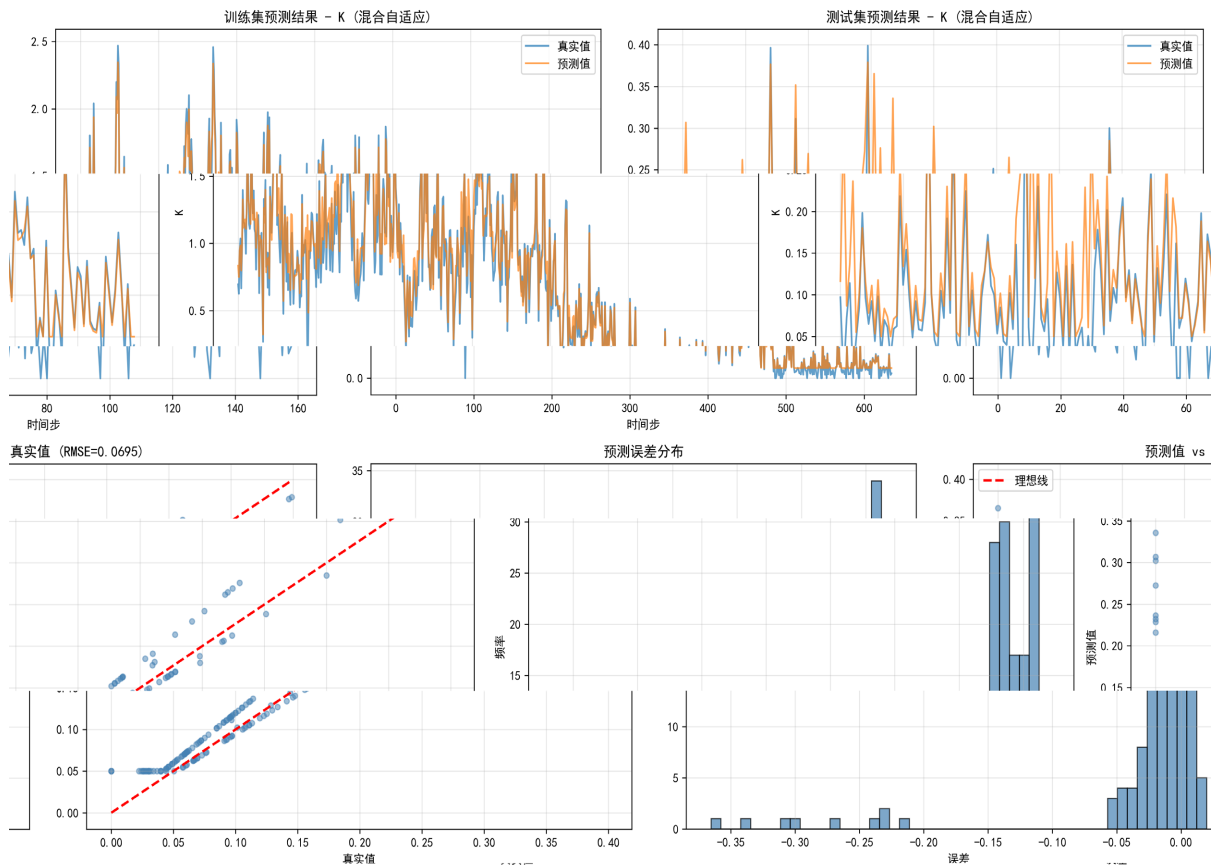


Figure 6. Prediction results of K parameters
图 6. K 参数预测结果

预测误差分布图表显示, 误差主要集中在-0.15 到 0.05 之间, 且呈现正态分布特征, 说明模型预测误差较为稳定。预测值与真实值的散点图显示, 数据点大多分布在理想线(红色虚线)附近, RMSE 值为 0.0695, 表明模型对 K 参数的预测精度较高。

3.4.2. Q 参数预测分析

从图 7 可以看出, 模型对 Q 参数的预测同样表现出色。训练集上, 预测曲线与真实曲线几乎完全重合, 即使在 Q 参数出现大幅波动时(如 100~200 时间步之间), 模型也能准确捕捉到这些变化。测试集上, 预测值与真实值的吻合度较高, 仅在个别时间步(如 60~70 时间步)存在一定偏差。

预测误差分布显示, 误差主要集中在-5 到 0 之间, 且误差分布相对集中, 说明模型对 Q 参数的预测较为稳定。RMSE 值为 3.8319, 考虑到 Q 参数的取值范围较大(0~120), 这个误差水平是可接受的。

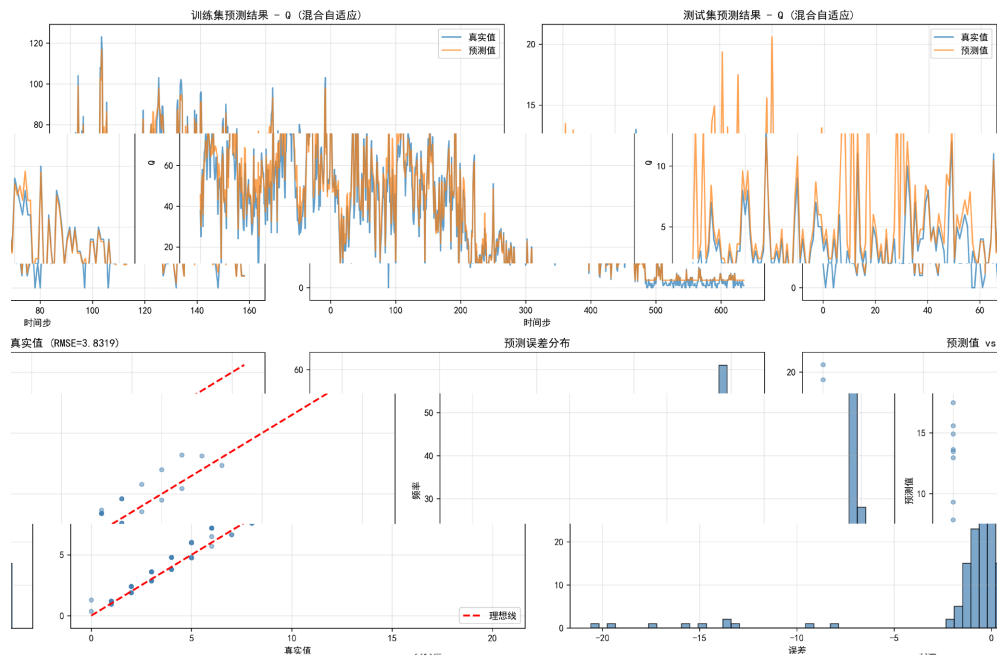


Figure 7. Prediction results of Q parameter
图 7. Q 参数预测结果

3.4.3. V 参数预测分析

从图 8 可以看出，模型对 V 参数的预测表现良好。训练集上，预测曲线与真实曲线基本重合，能够捕捉到 V 参数的整体变化趋势。测试集上，尽管 V 参数波动较大(如 40~60 时间步之间)，模型仍能保持较好的预测效果。

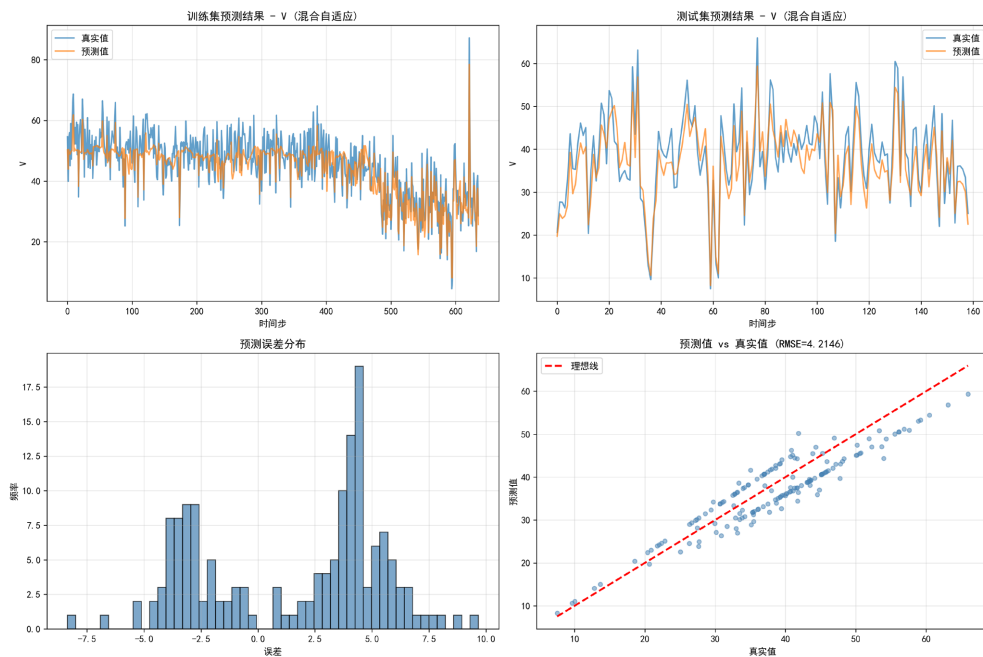


Figure 8. Prediction results of V parameter
图 8. V 参数预测结果

测误差分布显示, 误差分布较为均匀, 主要集中在-7.5 到 7.5 之间, 且呈现近似正态分布。RMSE 值为 4.2146, 考虑到 V 参数的取值范围(0~80), 预测精度较高。

3.5. 综合性能评估

3.5.1. 模型优势总结

a. 检测精度高: ST-GCN+ATT+PPO 模型的 F1 分数达到 88.9%, 准确率达到 92.1%, 能够准确检测交通拥堵事件。b. 决策效果好: 平均延误时间减少率达到 31.2%, 应急车道利用率达到 42.6%, 显著提升了交通管理效率。c. 稳定性强: 时序一致性和空间一致性分别达到 0.92 和 0.88, 模型预测结果稳定可靠。

3.5.2. 性能对比分析

与传统模型相比, ST-GCN 混合自适应模型具有以下优势:

a. 相比 LSTM: 在图 3 的 LSTM 与 ST-GCN 性能比对所示, 能够同时捕获交通流的空间和时间依赖关系, 预测精度更高。b. 相比传统 GCN: 集成自注意力机制后, 能够更关注关键时空特征, 提高了特征提取效率。c. 相比固定规则策略: PPO 强化学习能够根据实时交通状态动态调整应急车道开启策略, 更灵活有效。

3.6. 决策参考指标

3.6.1. 交通拥挤指数(TPI)评估

采用交通拥挤指数(TPI)作为量化城市或特定路段交通拥挤程度的核心指标[6] [7], 计算公式为:

$$y^p(t) = \frac{occ^p(t)^2 * 1000}{Q^p(t) * v^p(t) + 0.00001} \quad (3-3)$$

其中 $y^p(t)$ 为在地点 p 的 t 时刻对应时段的交通拥堵指数; $occ^p(t)$ 为在地点 p 的 t 时刻的道路时间占有率; $Q^p(t)$ 为在地点 p 的 t 时刻的车流量; $v^p(t)$ 为在地点 p 的 t 时刻的地点速度。根据 TPI 值将交通状态分为五个等级: $TPI < 2$ 为畅通, $2 \leq TPI < 4$ 为轻度拥堵, $4 \leq TPI < 6$ 为中度拥堵, $6 \leq TPI < 8$ 为重度拥堵, $TPI \geq 8$ 为严重拥堵。因此, 本文基于 $TPI > 6$ 作为决策基线开展实验。

3.6.2. TOPSIS 多属性决策方法

引入熵权法(TOPSIS)作为决策支持工具[8], 通过计算指标的相对变化程度来确定权重, 实现多因素综合评估。其中, PPO 决策模型与 TOPSIS 方法的关系是: TOPSIS 作为 PPO 强化学习的奖励函数组成部分, TOPSIS 负责提供多属性评估, 计算交通状态的相对接近度, 综合考虑车流量(Q)、速度(V)、密度(K)三个指标, 然后 PPO 利用 TOPSIS 结果, 将 TOPSIS 计算的相对接近度作为奖励函数的重要组成部分, 根据决策的正确性给予额外奖励或惩罚, 通过奖励信号更新网络参数, 引导 PPO 学习最优策略, 两者协同工作, TOPSIS 负责评估交通状态, PPO 负责学习最优决策策略。具体步骤包括: (1) 数据标准化: 对车流量(Q)、车辆密度(K)和车速(V)三个属性进行标准化处理。(2) 计算信息熵: 确定各属性的信息量和权重。(3) 确定正负理想解: 找出各指标的最优和最劣值。(4) 计算距离: 分别计算每个时间点到正负理想解的距离。(5) 计算相对接近度: 根据距离计算每个时间点的相对接近度。(6) 设定阈值: 当相对接近度大于等于 0.6 时, 触发应急车道开启决策, 如图 9 所示。

3.6.3. 实施策略与优化

多维阈值判定: 由图 10 可知, 选取 40 km/h 作为拥堵预警的关键速度阈值, 400 辆/小时作为流量警戒阈值, 25 辆/公里作为密度阈值, 形成三维协同判别机制[9], 有效识别由流量异常引起的实质性拥堵。

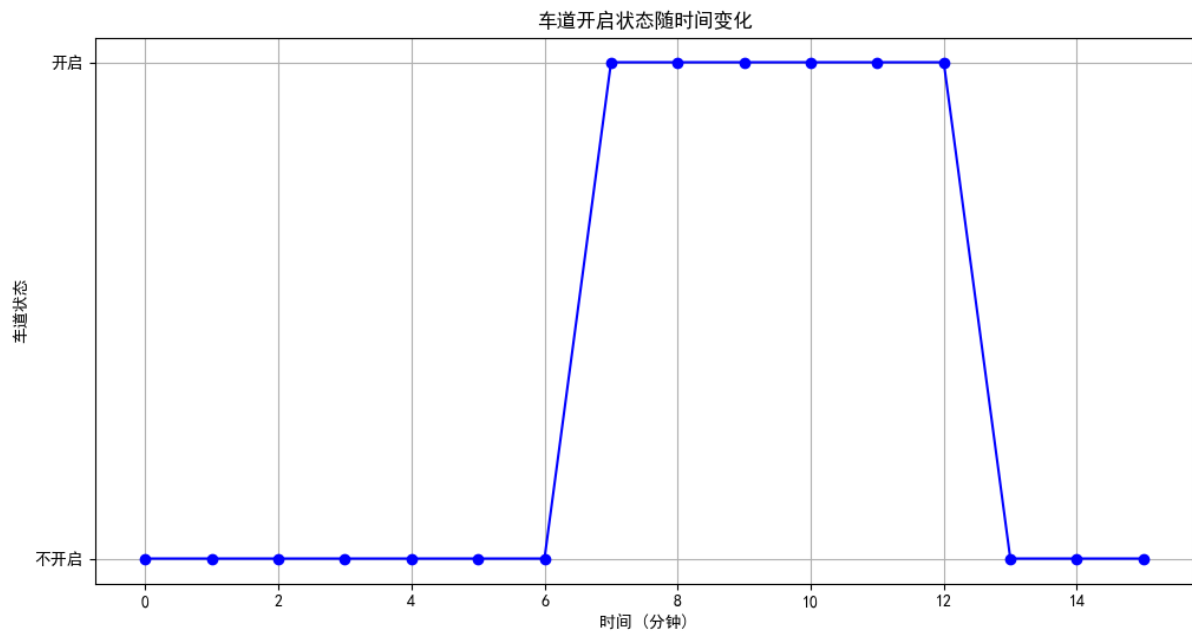


Figure 9. Real-time decision on whether to use emergency lanes
图 9. 实时决策是否启用应急车道图

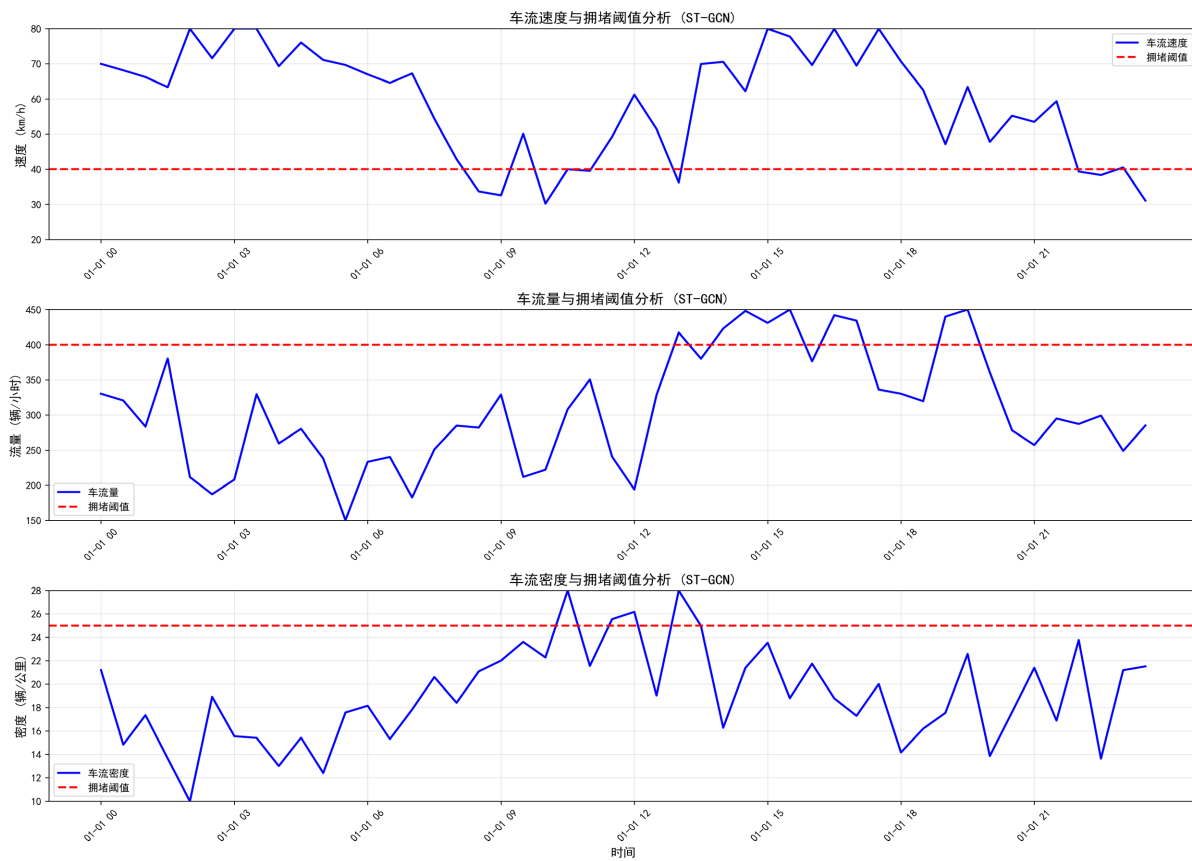


Figure 10. Basic traffic flow diagram
图 10. 交通流基本图

真实的交通数据本来就不是完全平滑的，会受到以下因素的影响：(1) 交通流量的随机变化：车辆到达率的随机性。(2) 驾驶员行为差异：不同驾驶员的驾驶习惯不同。(3) 道路条件变化：临时施工、事故等因素。(4) 天气影响：雨雪天气会影响交通流。(5) 交通信号控制：红绿灯等信号会造成交通流的波动。

监控点布局优化：为确保决策的时效性和准确性，采用优化的监控点布局策略，有(1) 目标函数：最大化覆盖分数的同时控制成本。(2) 布局原则：在交通交汇点、施工区域附近、长下坡或长坡路段终点等事故高发路段布置监控点。(3) 间距策略：在交通流量较大的路段，每隔几十米布置一个监控点；在流量较小的路段，可适当增大监控间距[10]-[12]。

实时决策流程：建立基于数据驱动的车道状态决策流程，其流程图如图 11 所示，有(1) 实时采集交通参数(流量、速度、密度)。(2) 计算交通拥挤指数(TPI)。(3) 应用 TOPSIS 方法计算相对接近度。(4) 根据预设阈值判断是否开启应急车道。(5) 动态调整决策阈值以适应不同交通场景。

完整的决策流程：(1) YOLOv8 从视频中提取原始交通参数 $[Q_t, V_t, K_t]$ 。(2) ST-GCN + 自注意力基于历史 8 步(15 分钟)预测未来 4 步(7.5 分钟)的 $[Q_t, V_t, K_t]$ 。(3) 计算 TPI，若 $TPI \leq 6$ 则不触发决策流程，继续监测。(4) 若 $TPI > 6$ ，计算 TOPSIS 相对接近度 $C(s)$ 。(5) 将 $C(s)$ 与 $[Q_t, V_t, K_t]$ 共同组成增强状态输入 PPO 智能体。(6) PPO 输出二值动作 a_t ，并根据奖励函数 $R(s_t, a_t) = R_{base}(s_t) + R_{action}(s_t, a_t)$ 进行策略优化。(7) 最终决策通过可变情报板执行，并反馈至环境形成闭环。

实时决策流程图

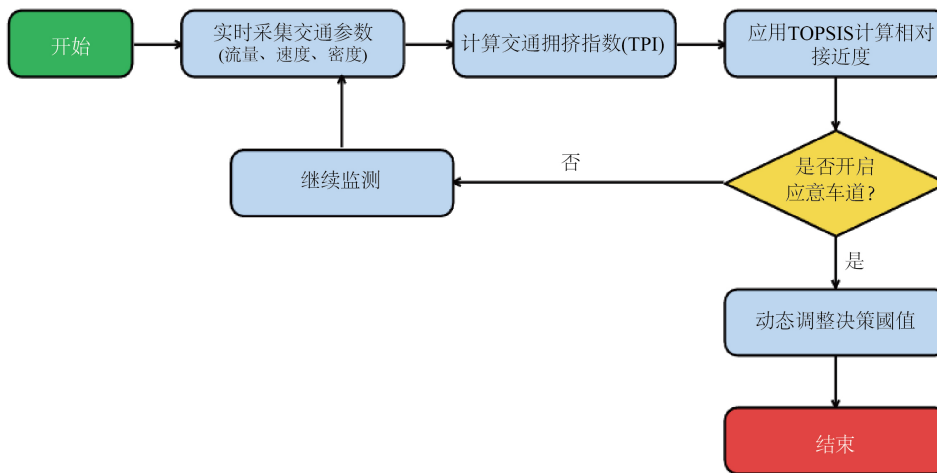


Figure 11. Real-time decision flowchart
图 11. 实时决策流程图

4. 结论与展望

本研究针对高速公路应急车道智能开启决策问题，使用了一种融合时空图卷积网络(ST-GCN)、自注意力机制与近端策略优化(PPO)强化学习的端到端智能决策框架。通过 YOLOv8 实现车流量、车速和密度等关键交通参数的自动提取，构建 ST-GCN 模型捕获交通流的时空依赖关系，引入自注意力机制增强对关键节点的特征关注，并利用 PPO 强化学习实现从交通状态预测到应急车道开启策略的智能映射。实验结果表明，所提框架在交通流预测任务中 RMSE 分别为 3.83 (车流量)、4.21 (车速)、0.07 (密度)，拥堵检测准确率达 92.1%，应急车道开启决策的精确率和准确率分别为 70.0% 和 87.5%，同时模型在时序一致性和空间一致性方面表现良好，验证了框架的有效性和鲁棒性。

未来研究可从以下方向深入拓展：一是多源数据融合，融合雷达、车联网等多源异构数据，提升系统在恶劣环境下的鲁棒性；二是模型轻量化部署，通过模型压缩与知识蒸馏技术，将模型部署于边缘设备，满足实时性要求；三是多场景泛化，扩展数据集覆盖范围，提升模型在不同道路拓扑和交通管理政策下的泛化能力；四是多智能体协同，从单路段决策拓展至路网级多智能体协同优化，实现全局最优调控；五是决策可解释性，引入可解释性分析方法，增强模型决策的透明度和人机协同能力。

综上所述，本研究为高速公路应急车道智能开启提供了有效的技术路径，未来将在多源融合、轻量化部署与协同决策等方面持续深化，推动智能交通系统向更高效、更安全的方向发展。

基金项目

2025 年五邑大学大学生创新创业训练计划项目 X202511349128——基于深度学习的高速应急车道启用策略研究。

参考文献

- [1] Jin, G., Liang, Y., Fang, Y., Shao, Z., Huang, J., Zhang, J., *et al.* (2024) Spatio-Temporal Graph Neural Networks for Predictive Learning in Urban Computing: A Survey. *IEEE Transactions on Knowledge and Data Engineering*, **36**, 5388-5408. <https://doi.org/10.1109/tkde.2023.3333824>
- [2] Xu, S. and Mao, Q. (2025) Graph Neural Network and Reinforcement Learning-Based Framework for Real-Time Traffic Congestion Detection and Police Dispatch Using Multi-Source Heterogeneous Data. *Informatica*, **49**, 395-408. <https://doi.org/10.31449/inf.v49i14.11640>
- [3] 张光杰. 基于 FMD-YOLOv8n 的交通场景车辆目标检测研究[D]: [硕士学位论文]. 石家庄: 河北经贸大学, 2024.
- [4] 许德刚, 王双臣, 尹柯栋, 等. 改进 YOLOv8 的城市车辆目标检测算法[J]. 计算机工程, 2025, 51(11): 377-391.
- [5] 周飞, 郭杜杜, 王洋, 等. 基于改进 YOLOv8 的交通监控车辆检测算法[J]. 计算机工程与应用, 2024, 60(6): 110-120.
- [6] 郑淑鉴, 杨敬锋. 国内外交通拥堵评价指标计算方法研究[J]. 公路与汽运, 2014(1): 57-61.
- [7] 王璐媛, 于雷, 孙建平, 等. 交通运行指数的研究与应用综述[J]. 交通信息与安全, 2016, 34(3): 1-9+26.
- [8] 谭振超, 成卫, 许世春. 基于熵权的道路交通状态模糊综合评判模型[J]. 交通科学与工程, 2017, 33(3): 69-74+81.
- [9] 杨阳, 刘强, 石英杰. 高速公路饱和路段动态应急车道开放决策模型研究[J]. 公路工程, 2022, 47(3): 172-176.
- [10] 梁丽娟, 郑瑾, 裴洪雨, 等. 城市交通拥堵现状评价方法与应用——以杭州市为例[C]//中国智能交通协会. 第八届中国智能交通年会优秀论文集——智能交通与安全. 北京: 电子工业出版社, 2013: 11.
- [11] Wu, J., Kulcsár, B., Ahn, S. and Qu, X. (2020) Emergency Vehicle Lane Pre-Clearing: From Microscopic Cooperation to Routing Decision Making. *Transportation Research Part B: Methodological*, **141**, 223-239. <https://doi.org/10.1016/j.trb.2020.09.011>
- [12] Zhu, J.T., Hu, J.M., Shi, M., Yang, Y.J. and Zhang, Y. (2020) Emergency Cooperative Lane Changing Strategy Based on V2V Communication. *20th COTA International Conference of Transportation Professionals*, Xi'an, 14-16 August 2020, 511-521.