

# 控制与决策

Control and Decision

## 融合多源环境感知的强化学习能耗预测模型

彭自然, 杨肖阳, 舒中宾

引用本文:

彭自然, 杨肖阳, 舒中宾. 融合多源环境感知的强化学习能耗预测模型[J]. *控制与决策*, 2026, 41(6): 1731–1742.

在线阅读 View online: <https://doi.org/10.13195/j.kzyjc.2025.1052>

---

## 您可能感兴趣的其他文章

### Articles you may be interested in

#### 一种基于池计算的宽度学习系统

A broad learning system based on reservoir computing

控制与决策. 2021, 36(9): 2203–2210 <https://doi.org/10.13195/j.kzyjc.2019.1729>

#### 基于近端强化学习的股价预测方法

Method of stock prices forecast based on proximal reinforcement learning

控制与决策. 2021, 36(4): 967–973 <https://doi.org/10.13195/j.kzyjc.2019.1245>

#### 基于MCPDDPG的智能车辆路径规划方法及应用

The method and application of intelligent vehicle path planning based on MCPDDPG

控制与决策. 2021, 36(4): 835–846 <https://doi.org/10.13195/j.kzyjc.2019.0460>

#### 一种基于深度学习的时间序列预测方法

A time series prediction method based on deep learning

控制与决策. 2021, 36(3): 645–652 <https://doi.org/10.13195/j.kzyjc.2019.0809>

#### 基于深度学习的行人轨迹预测方法综述

Survey of pedestrian trajectory prediction methods based on deep learning

控制与决策. 2021, 36(12): 2841–2850 <https://doi.org/10.13195/j.kzyjc.2020.1841>

# 融合多源环境感知的强化学习能耗预测模型

彭自然, 杨肖阳, 舒中宾<sup>†</sup>

(湖南工业大学 交通与电气工程学院, 湖南 株洲 412000)

**摘要:** 针对现有电动汽车实时能耗预测模型环境感知能力不足、动态校准机制缺失等问题, 提出一种融合环境感知与强化学习的能耗预测模型. 首先, 为增强模型对于复杂工况的感知和理解, 设计基于对比学习与耦合强化学习协同训练的路况感知算法, 并引入多尺度图像特征融合机制, 有效提取与车辆能效高度相关的环境特征, 从而提升对于非平稳工况的感知精度; 然后, 构建马尔可夫实时能效估计模型并将其映射至强化学习框架中, 引入基于折扣未来能耗的时序一致性正则项 (其中  $Q$  函数仅用于能耗响应评估), 实现模型的自校准式优化, 从而在动态场景下显著增强预测的鲁棒性和自适应性 (模型不产生控制输出); 同时, 结合场景感知的优先经验回放机制, 强化模型对坡度突变、急加减速等关键工况的识别和学习, 进一步提高复杂环境下的特征提取和模型泛化能力; 最后, 通过场景感知的优先采样策略优化训练样本分布, 提升强化学习的收敛速度和训练效率. 实验结果表明, 所提出方法在所测试的两款车型以及多种仿真工况下均表现出优越的鲁棒性和稳定性, 其 MAE 低于 0.2%, RMSE 低于 0.3%,  $R^2$  超过 99.5%. 与现有 Transformer、Informer、Mamba 以及 LSTM 模型相比, 平均误差分别降低约 40% ~ 70%, 收敛速度提升约 30%, 在复杂工况下能耗预测精度显著提高.

**关键词:** 电动汽车实时能耗预测; 深度强化学习; 实时路况感知; 主动进化学习

中图分类号: TM715

文献标志码: A

DOI: 10.13195/j.kzyjc.2025.1052

引用格式: 彭自然, 杨肖阳, 舒中宾. 融合多源环境感知的强化学习能耗预测模型 [J]. 控制与决策, 2026, 41(6): 1731-1742.

## Reinforcement learning-based energy consumption prediction model with multi-source environmental perception

PENG Zi-ran, YANG Xiao-yang, SHU Zhong-bin<sup>†</sup>

(School of Transportation and Electrical Engineering, Hunan University of Technology, Zhuzhou 412000, China)

**Abstract:** To address the limitations of existing real-time energy consumption prediction models for electric vehicles — Particularly their insufficient environmental perception and lack of dynamic calibration mechanisms — this study proposes an energy consumption prediction model integrating environmental perception with reinforcement learning. First, to enhance the model's capability of understanding complex driving conditions, a road-condition perception algorithm is designed based on contrastive learning and coupled reinforcement learning, together with a multi-scale image feature fusion mechanism. This design effectively extracts environment features highly correlated with vehicle energy efficiency, thereby improving perception accuracy under non-stationary operating conditions. Second, a Markov-based real-time energy efficiency estimation model is constructed and mapped into a reinforcement learning framework. A temporal consistency regularization term based on discounted future energy consumption (where the  $Q$ -function is used solely as an energy-response evaluator) is introduced to achieve self-calibrated optimization, significantly enhancing prediction robustness and adaptability in dynamic scenarios (without generating control outputs). Meanwhile, a scenario-aware prioritized experience replay mechanism is incorporated to strengthen the model's ability to recognize and learn from key driving conditions such as slope mutations and rapid acceleration/deceleration events, further improving feature extraction and generalization in complex environments. Finally, a scenario-aware prioritized sampling strategy is employed to optimize the distribution of training samples, improving the convergence rate and efficiency of the reinforcement learning process. Experimental results demonstrate that the proposed method exhibits excellent robustness and stability across two vehicle types and multiple simulated

收稿日期: 2025-10-11; 录用日期: 2026-02-13.

责任编辑: 马大中.

<sup>†</sup>通信作者. E-mail: 123998439@qq.com.

driving scenarios, achieving an MAE below 0.2%, an RMSE below 0.3%, and an  $R^2$  above 99.5%. Compared with existing Transformer, Informer, Mamba, and LSTM models, the proposed approach reduces average prediction error by approximately 40% ~ 70% and improves convergence speed by about 30%, yielding significantly higher prediction accuracy under complex driving conditions.

**Keywords:** real-time energy consumption prediction for electric vehicles; deep reinforcement learning; road condition awareness; proactive adaptive learning

## 0 引言

电动汽车 (EVs) 作为交通领域的重要发展方向, 是应对环境污染和能源危机的关键途径<sup>[1]</sup>. 随着普及率的提高, 精准能耗预测已成为提升用户体验和使用效率的核心问题<sup>[2]</sup>. 实时能耗预测可为车主提供充电和出行决策依据, 但是受驾驶习惯、天气、路况以及载荷等多因素影响, 预测仍具挑战<sup>[3]</sup>.

目前, 电动汽车能耗预测方法主要包括基于物理模型和基于数据驱动两类<sup>[4]</sup>: 前者依赖电池特性建模 (如电压、电流、温度等) 计算能耗<sup>[5]</sup>, 但是因假设简化, 难以应对实际运行中的复杂非线性因素<sup>[6]</sup>, 如驾驶行为和交通状况, 预测精度有限<sup>[7]</sup>; 相比之下, 基于数据驱动的方法利用历史数据训练模型, 可自动捕捉行驶过程中的多种影响因素, 从而显著提升预测准确性<sup>[8]</sup>. 早期研究多采用线性回归、支持向量机 (SVM) 等传统机器学习算法<sup>[9]</sup>. 文献 [10] 采用了梯度提升决策树算法, 进而考虑到大量传统回归方法无法涵盖的因素, 该方法的预测误差最大为 1.58 km, 最小为 -1.41 km, 平均预测误差约为 0.7 km, 有效提升了其预测性能. 但是, 其存在计算复杂度高、模型解释性差、对于异常值敏感以及对于数据质量依赖性强等问题<sup>[11]</sup>.

近年来, 随着深度学习技术的突破性进展, 基于深度神经网络的电动汽车能耗预测方法为续航里程估算开辟了新途径<sup>[12]</sup>. 循环神经网络 (RNN)<sup>[13]</sup>、长短期记忆网络 (LSTM)<sup>[14]</sup> 和门控循环单元 (GRU)<sup>[15]</sup> 等模型, 凭借自动特征提取和非线性建模能力, 显著提升了能耗预测的精度. 文献 [16] 提出了一种结合多尺度线性回归与神经网络的能耗预测方法, 利用实时路况、地理和天气数据动态预测能耗, 预测误差仅为 12% ~ 14%; 文献 [17] 针对电池荷电状态 (SOC) 对于续航预测的影响, 采用 LSTM 模型实现了 SOC 的高精度预测, 从而提升了剩余里程估计的准确性.

然而, 传统深度学习模型仍然受梯度消失、计算复杂度高等问题制约, 难以兼顾预测精度与实时性<sup>[18]</sup>. 文献 [19] 提出的 LSTM-DNN 混合模型虽然通过整合实时交通与车辆状态数据提高了预测精度, 但是在动态环境下的自适应性不足; 文献 [20] 利用 LSTM 与 Transformer 相结合的能耗预测框架, 增强

了对驾驶风格和环境变化的建模能力, 有效缓解了长依赖问题. 但是现有研究仍然存在以下两方面主要局限:

1) 环境感知和动态自适应能力不足: 现有能耗预测模型普遍依赖固定特征或静态数据建模, 难以充分感知环境突变 (如道路坡度变化、交通流密度波动以及气候条件变化等), 缺乏面向非平稳工况的主动进化和自校准机制, 导致模型在复杂动态场景下预测性能显著下降<sup>[21]</sup>.

2) 长时序依赖建模和计算效率受限: Transformer 及其变体虽然在序列建模中表现优异, 但是在处理车辆运行的长时序能耗数据时, 仍然存在计算开销大、位置编码精度不足以及远程依赖衰减等问题, 难以兼顾实时性与预测精度, 限制了其在复杂交通场景下的实时能耗预测能力<sup>[22]</sup>.

针对上述问题, 本文提出一种融合多尺度环境建模与强化学习的实时能耗预测模型 EARL-ECP, 以弥补现有模型在环境感知、自适应性以及时序一致性建模方面的不足. 本文主要内容如下:

1) 创新性地融合环境感知与强化学习机制: 提出环境感知驱动的强化学习能耗预测框架, 将对比学习与多模态感知相结合, 使得模型能够自适应复杂非平稳环境下的能耗特征变化.

2) 提出时序一致性正则化的自校准预测机制: 以折扣未来能耗为价值函数约束, 实现模型在多场景切换和动态路况下的自进化式校准, 提高预测的稳定性和鲁棒性.

3) 构建场景感知的多尺度特征融合和优先采样策略: 通过多尺度融合强化模型对关键工况 (如坡度突变、急加减速) 的感知能力, 显著提升能耗预测的精度和泛化性能.

## 1 环境感知强化学习的能耗估计框架

### 1.1 场景分析

电动汽车实时能耗受多种动态因素影响, 其中复杂多变的路况环境和驾驶控制策略对于电动汽车的实时能效表现有显著作用, 直接导致状态电量 SOC 动态评估中的累积误差. 影响电动汽车能耗的 3 种典型场景如下:

1) 在城市拥堵环境中, 车辆频繁启停, 能耗受到交通信号、前车行为和突发状况等高频动态因素影响. 电动汽车通过感知模块提取高维观测中的潜在动态状态, 可显著提升能耗估计的即时性和精度.

2) 在高速公路匀速行驶场景下, 尽管速度稳定, 但是坡度、风阻等因素仍然会导致能耗波动. 电动汽车通过感知前方隧道、限速标志和坡道信息, 结合强化学习不断优化路况状态-能耗估值函数, 从而能够实时修正异常路况下能耗波动造成的预测误差.

3) 在乡村以及非结构化道路环境中, 大量不可预知因素, 如急转弯、泥泞路面等不可预知因素对于能耗影响显著, 对此, 电动汽车通过端到端决策模型做出最佳能耗估计, 提升能耗预测的鲁棒性.

### 1.2 电动汽车能耗预测系统框架

综上所述, 尽管路况环境复杂多变, 电动汽车实时能耗波动显著, 但是通过智能驾驶辅助系统感知路况并结合决策推理模型推算环境状态对应的能耗

增量, 便可实现能耗的精准预测和评估.

鉴于此, 本文提出端到端能耗预测模型 EARL, 其包含两部分, 如图 1 所示: 1) 环境感知模块, 负责特征提取和状态表征; 2) 强化学习模块, 负责策略学习和优化. 感知模型首先对路况图像、行车状态等多模态高维观测信息  $o_t$  进行编码, 利用编码器  $\phi_{enc}$  提取与能耗行为高度相关的特征, 并映射至低维表征空间. 经编码后的表征数据作为状态信息输入强化学习网络进行实时能耗增量预测的评估. 强化学习模块采用时间差分 (TD) 学习策略, 当前  $Q$  网络与目标  $Q$  网络的参数在经验数据引导下交替优化, 有效增强模型对于动态环境的响应能力. 其次, 本文任务为点预测, 不涉及控制指令的生成和执行. 强化学习相关术语仅用于构建价值一致性正则: 以“折扣未来能耗的负值”作为训练信号, 学习一个能耗响应评估器, 其梯度作为时序正则回流至编码和预测头; 推理阶段无策略/控制头和闭环执行.

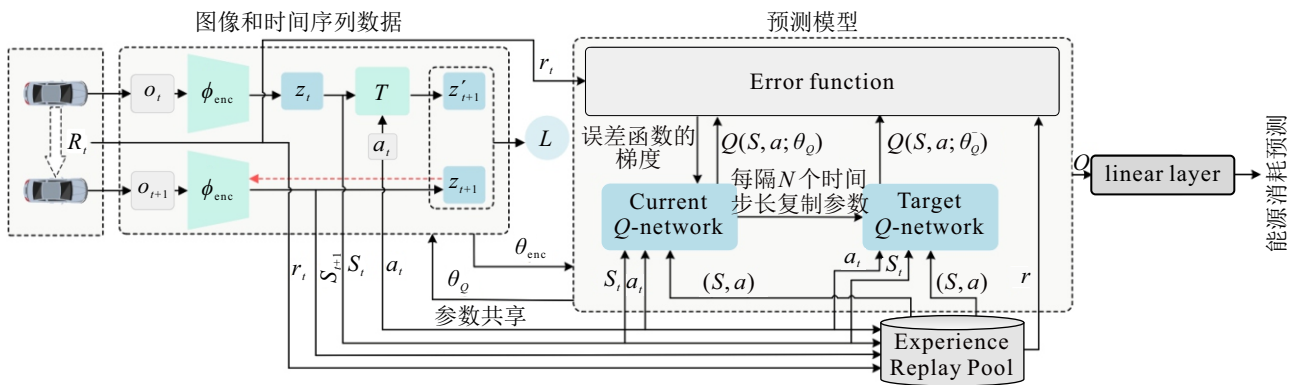


图1 网络整体框架

### 1.3 理论分析

为论证强化学习在电动汽车实时能耗预测中的有效性以及算法的收敛性, 本节从理论视角深入剖析马尔可夫决策 (MDP) 过程和强化学习建模机制. 在训练过程中, 当前状态  $S_t$  和动作  $a_t$  共同输入  $Q$  网络, 生成  $Q$  值估计  $Q(S_t, a_t)$ , 表示在该状态下采取相应动作后预计产生的累积能耗. 经固定时间间隔  $\Delta t$ , 系统根据函数  $S_{t+1} = \Gamma(S_t, a_t)$  转移至下一状态, 并返回即时奖励  $R_{t+1}$ . 然后, 模型采用均方误差构造损失函数  $L(\theta)$ , 以目标  $Q$  值与当前  $Q$  值间的偏差为优化目标, 梯度为

$$\nabla_{\theta} L(\theta) = (Q(S_t, a_t; \theta) - y_t) \nabla_{\theta} Q(S_t, a_t; \theta).$$

根据梯度下降法,  $\theta_{new} \leftarrow \theta - \alpha \nabla_{\theta} L(\theta)$ , 因此, 对于足够小的学习率  $\alpha$ , 参数更新后的损失函数近似满足  $L(\theta_{new}) \approx L(\theta) - \alpha \|\nabla_{\theta} L(\theta)\|^2$ . 由于  $\|\nabla_{\theta} L(\theta)\|^2 \geq 0$ , 可得到  $L(\theta_{new}) \leq L(\theta)$ , 从而保证损失函数单调下

降, 网络逐步收敛至最优解.

为实现端到端的估计目标, 模型输出为状态与动作的联合函数  $L(\theta)$ , 通过联合最小化  $L_{\theta}$  和最大化累积奖励  $G_t$ , 实现能效估计精度与策略性能的同步优化, 计算如下所示:

$$L(\theta) = \frac{1}{T} \sum_t (\hat{E}_t - E_t)^2 - \lambda \sum_t G_t. \quad (1)$$

其中:  $\hat{E}_t$  为时刻  $t$  的能耗预测值,  $E_t$  为实际能耗,  $G_t$  为对应的累积奖励,  $\lambda$  为精度和收益的权衡系数.

综上所述, 基于环境感知的强化学习能耗估计模型不仅能够适应动态环境变化, 还能在复杂工况下精确预测能耗.

## 2 融合深度强化学习的路况感知

电动汽车感知、理解复杂路况环境是提升预测精度的关键. 所提出模型通过与强化学习深度耦合的表征学习, 实现对工况变化的感知和理解.

## 2.1 感知编码器的对比学习与耦合强化训练

为从高维噪声观测中提取与电动汽车能耗预测高度相关的潜在特征, 本文通过对比学习与耦合强化学习对感知编码器进行训练, 原理结构如图1所示. 针对每个时刻 $t$ 的图像数据 $o_t^{\text{img}}$ 和时序数据 $o_t^{\text{ts}}$ , 首先, 通过预融合函数 $F(\cdot)$ 得到融合数据 $o_t = F(o_t^{\text{img}}, o_t^{\text{ts}})$ , 再经 $\phi_{\text{enc}}$ 映射为潜在表征 $z_t = \phi_{\text{enc}}(o_t; \theta_{\text{enc}})$ ,  $\theta_{\text{enc}}$ 为编码器的网络参数. 考虑系统动态演化, 引入潜在前向模型 $T$ 描述潜在表征的时间演化, 得到 $z'_{t+1} = T(z_t, a_t; \theta_T)$ . 其中:  $a_t$ 为当前汽车加速度,  $\theta_T$ 为前向模型的网络参数.

通过优化辅助损失函数 $L$ 最小化预测潜在状态与真实潜在状态间的差异, 使得该模块能够准确捕捉并表征系统动态规律. 具体计算如下所示:

$$L(\theta_{\text{enc}}, \theta_T) = \mathbb{E}_{o_t, a_t, o_{t+1}} \sim D[\|z_{t+1} - z'_{t+1}\|^2]. \quad (2)$$

同时, 为增强感知结果的决策适应性, 本文将编码器与强化学习(DQN)进行耦合训练, 如图2所示. 潜在表征 $z_t$ 作为状态输入, 通过与策略网络参数共享实现协同优化, 提升训练效率与感知一致性. DQN中估计动作价值函数 $Q(s_t, a_t)$ 实际上为

$$Q(z_t, a_t; \theta_Q) = Q(\phi_{\text{enc}}(o_t), a_t; \theta_Q). \quad (3)$$

其中:  $\phi_{\text{enc}}$ 在表征学习模块和DQN主网络、目标网络中共享,  $\theta_Q$ 为DQN中除编码器外的参数.

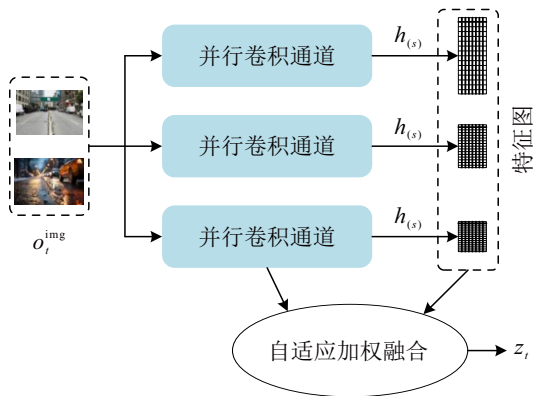


图2 多尺度图像特征融合机制

训练过程中, DQN 损失函数产生的梯度通过反向传播经 DQN 模块到达编码器 $\theta_{\text{enc}}$ , 从而实现编码器参数更新, 过程如下所示:

$$\begin{cases} \theta_Q \leftarrow \theta_Q - \alpha \nabla_{\theta_Q} (Q(z_t, a_t) - y_t)^2, \\ y_t = r_t + \gamma \max_{a'} Q(z_{t+1}, a'; \theta_Q^-). \end{cases} \quad (4)$$

其中 $\theta_Q^-$ 为DQN目标网络的参数.

通过将感知编码器与强化学习进行耦合训练学习, 编码器能够提取决策优化过程中最敏感的信息进行编码, 从而实现从高维复合信息到精准的能效估计决策.

## 2.2 多尺度特征融合提升图像时序编码性能

为增强感知编码器 $\phi_{\text{enc}}$ 在处理复杂图像观测时对不同尺度动态信息的建模能力, 本文引入多尺度特征融合机制(MSFF). 该机制通过在编码器中设置多个并行卷积通道, 分别提取局部细节和全局趋势信息, 并以权重融合的方式生成统一的潜在表征, 从而实现多粒度时空特征的协同建模, 具体原理如图2所示.

设原始观测图像帧为 $o_t^{\text{img}}$ , 则其在第 $s$ 个尺度通道上的卷积特征提取为

$$h_t^{(s)} = f_{\text{conv}}^{(s)}(o_t), \quad (5)$$

其中 $f_{\text{conv}}^{(s)}$ 为第 $s$ 个尺度下的卷积通道, 具备特定大小的感受野. 然后, 通过加权融合策略将多尺度特征集成为统一的潜在表示 $z_t$ , 计算如下所示:

$$z_t = \sum_{s=1}^S \alpha_s \cdot h_t^{(s)}, \quad \sum_{s=1}^S \alpha_s = 1, \quad (6)$$

这里权重系数 $\alpha$ 由轻量级门控网络根据输入图像的全局特征动态生成, 体现模型对于各尺度特征贡献度的自适应感知能力.

该机制可以显著增强模型在工况突变与模态干扰场景下的表征鲁棒性和时序建模一致性, 为后续决策网络提供更可靠的输入基础.

## 3 能耗预测强化学习

本文将能效估计建模为MDP, 以车辆实际加速度 $a_t$ 作为动作输入, 以行车真实能耗作为奖励信号 $R_t$ , 并将该过程映射至深度强化学习框架以实现能耗的实时精确预测. 需要强调的是, 本文并非将能耗预测视为一个控制或决策问题, 而是将其建模为一个带有时序一致性约束的预测问题. 引入强化学习相关符号和训练机制的目的, 并非学习最优动作策略, 而是利用TD学习在折扣未来累积量建模方面的优势, 为能耗预测网络提供一种结构化的时序正则化形式. 与传统的LSTM/Transformer仅依赖监督误差不同, TD正则能够显式约束“当前预测”与“未来预测+实际能耗观测”间的一致性关系, 从而在非平稳工况下提升模型的稳定性和自校准能力. 建模分析如下.

### 3.1 模型参数建模与分析

#### 1) 状态、动作建模.

状态信息涵盖天气、路况、总行驶距离 $d_t$ 、瞬时速度 $v_t$ 、电压、电流、温度、SOC等, 均来自电池管理系统BMS, 以及车辆自身特性(如踏板状态、车重、迎风面积等), 即 $S_t = \{\text{天气, 路况, } d_t, v_t, \text{电池状态, 车辆信息}\}$ . 其均为CARLA仿真平台BMS的标准

特征, 相关作用机制已有验证, 因此, 未再单独分析. 在本文中,  $Q(S_t, a_t)$  并不表示“采取动作的最优回报”, 而是表示在给定状态和观测加速度条件下, 未来折扣能耗的条件期望值, 用于预测网络的时序一致性评估函数.

2) 奖励建模.

在能耗预测任务中, 模型需要准确预测未来能耗曲线. 为确保预测结果真实反映电能消耗变化, 奖励函数  $R_{t+1}$  定义为  $S_t$  到  $S_{t+1}$  实际能耗的负值, 以最小化预测误差, 如下所示:

$$R_{t+1} = - \int_t^{t+\Delta t} P(\tau) d\tau, \quad (7)$$

其中  $P(\tau)$  为瞬时功率, 由 BMS 直接测量.

由于奖励被定义为每步的负电能消耗, 所学的价值函数  $Q$  为

$$Q(S_t, a_t; \theta) = \mathbb{E} \left[ \sum_{k=0}^{\infty} \gamma^k R_{t+k} | S_t, A_t \right]. \quad (8)$$

3.2 基于场景感知的优先级采样机制

本文针对传统优先级经验回放 (PER) 在极端路况样本稀缺和平稳数据价值挖掘不足的双重缺陷, 提出一种场景感知的优先采样机制 (SAPSM). 该方法保留了 PER 基于 TD 误差优先学习框架, 显著增强了对路况突变、关键转移的识别能力, 有效提升了能耗预测模型在复杂环境下的学习效率和泛化性能.

首先, 对每条经验转移  $(S_t, A_t, R_t, S_{t+1})$  计算其 TD 误差, 如下所示:

$$\delta_i = r_t + \gamma Q_{\theta'_i}(s_{t+1}, \pi_{\phi}(s_{t+1})) - Q_{\theta_i}(s_t, a_t), \quad (9)$$

并据此赋予静态优先级权重, 即

$$P_i^{\text{static}} = (|\delta_i| + \varepsilon)^\alpha. \quad (10)$$

其中:  $\alpha \in (0, 1]$ ;  $\varepsilon$  为一个微小的正数常量, 用于防止极小误差导致采样概率为 0.

然而, 传统 TD 误差优先级评估在处理数值平缓但是包含关键驾驶特征的转移 (如急加减速、坡度突变) 时易出现识别不足. 为提升对此类关键场景的关注, 本节引入场景指示变量  $\xi$  (剧烈行驶特征为 1), 构建融合 TD 误差和场景价值的综合优先级评分机制, 表述如下:

$$g_i = 1 + (\beta - 1)\xi_i, P_i = P_i^{\text{static}} \times g_i. \quad (11)$$

其中:  $\beta$  为场景增益因子, 用于控制关键场景样本的重要性放大程度.  $\xi_i$  为场景指示变量, 当样本处于急加减速、坡度突变等复杂工况时取 1; 否则, 取 0.

在此基础上, 采用分层采样策略, 先按照“平稳行驶”和“剧烈行驶”划分经验池, 再在子集中按照综合优先级进行抽样. 该策略既保证了关键转移的

学习频次, 又通过重要性权重校正, 兼顾采样多样性与训练效率. 表述如下:

$$W_j = \left( \frac{1}{N \cdot P(j)} \right)^\mu, P(i) = \frac{p_i}{\sum_j p_j}. \quad (12)$$

其中:  $W_j$  为重要性采样权重,  $\mu$  为重要性采样指数,  $p_j$  为样本抽样概率.

3.3 Off-Policy 能耗预测强化学习算法

首先, 经验池  $D$  中存储每一时刻交互生成的四元组  $(S_t, A_t, R_t, S_{t+1})$ , 在每次参数迭代更新时从中采样出一批大小为  $N$  的样本, 并网络预测得出在当前状态  $S_t$  采取动作  $a_t$  后, 预测得到  $Q(S_t, a_t; \theta) = f(S_t, a_t; \theta)$ . 为了获得更稳定的目标估计, 引入与主网络结构相同的目标网络, 目标网络输出的目标  $Q$  值用于网络的更新, 计算如下所示:

$$y_t = R_{t+1} + \gamma \max_{a'} Q(S_{t+1}, a_{t+1}; \theta^-), \quad (13)$$

其中  $\max_{a'} Q(S_{t+1}, a_{t+1}; \theta^-)$  为下一状态的最大  $Q$  值, 由目标网络 (参数为  $\theta^-$ ) 计算.

接下来, 计算每个样本的时序差分误差, 表示为目标  $Q$  值与当前  $Q$  值间的均方误差, 并以此作为均方误差损失函数, 计算如下所示:

$$L(\theta) = \mathbb{E}[(y_t - Q(S_t, a_t; \theta))^2]. \quad (14)$$

主网络的参数  $\theta$  通过最小化该损失函数更新, 计算如下所示:

$$\theta \leftarrow \theta - \alpha \nabla_{\theta} \mathbb{E}[(R_{t+1} + \gamma \max_{a'} Q(S_{t+1}, a'; \theta^-) - Q(S_t, a_t; \theta))^2]. \quad (15)$$

为确保训练的稳定性, 目标网络的参数  $\theta^-$  并非每次完全复制主网络, 而是定期采用软更新方式与主网络同步更新, 即  $\theta^- \leftarrow \tau \theta + (1 - \tau) \theta^-$ ,  $\tau$  为软更新系数 ( $0 < \tau \ll 1$ ), 控制更新速度. 配合  $\varepsilon$ -贪心策略兼顾探索与利用, 该框架可在以加速度为外生输入的条件下学习能耗响应评估器, 用于提升能耗预测的时序一致性和泛化能力, 进而实现对车辆能耗的精准预测和优化. 所提出强化学习算法以加速度作为外生输入, 用于学习状态-动作条件下的能耗价值函数, 以增强预测的时序一致性和自校准能力; 该

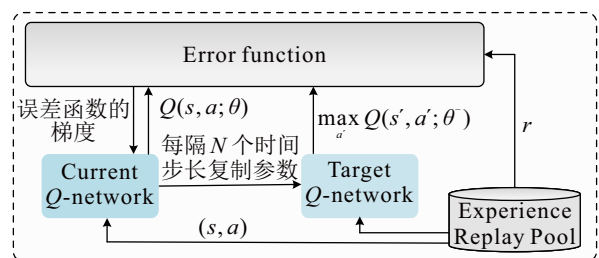


图3 整体算法原理框架

表1 EARL-ECP 算法伪代码

```

算法1 EARL-ECP.
Initialize parameters  $\theta, \theta^-, \theta_{enc}$ 
for each data-collection window do
#使用日志和仿真轨迹(非策略控制)填充经验池
    Using current policy, collect and store  $\langle S_t, a_t, R_t, S_{t+1} \rangle$  in  $D$ 
    for  $updatestep = 1, 2, \dots, K$  do
    Sample a mini-batch from  $D$  #点预测损失(能耗/里程)

    Calculate  $Q(S_t, a_t) = f(S_t, a_t); y_t = R_{t+1} + \gamma \max_{a'} Q(S_{t+1}, a_{t+1})$ 
    Calculation of batch loss:  $L(\theta); L$ 
    end for

 $\theta \leftarrow \theta - \alpha \nabla_{\theta} L(\theta); \theta_{enc} \leftarrow \theta - \alpha \nabla_{\theta_{enc}} L$ 

 $\theta^- \leftarrow \tau \theta + (1 - \tau) \theta^-$ 
end for

```

框架不涉及控制策略学习,也不对车辆未来行驶轨迹产生影响.整体框架和算法的伪代码分别如图3和表1所示.

## 4 实验分析

为全面验证模型的泛化和稳定性能,本节在虚拟仿真平台上进行多场景实验,涵盖不同车辆、地图、天气(晴、雨、雾、昼夜)以及多种交通密度设置,以反映模型在复杂动态环境下的预测适应能力.

### 4.1 实验设置与能耗模块设计

#### 4.1.1 车辆运行工况设置

为验证模型在典型驾驶场景下的泛化能力,本文在CARLA平台中构建4种代表性工况,包括城市启停、高速巡航、山路坡道和乡村变速,覆盖从低速频繁制动到高速匀速、从平路到坡道的典型能耗特征,工况设置和参数配置如表2所示.

表2 工况设置和参数配置

工况名称	平均速度 (km/h)	加速度范围 (m/s <sup>2</sup> )	路况类型	特征说明
城市启停	15	± 2.5	城市道路	拥堵频繁启停
高速巡航	90	± 0.5	高速公路	稳定匀速
山路坡道	45	± 1.8	山路/坡道	高低起伏明显
乡村变速	40	± 1.2	郊区/乡道	转弯与减速交替频繁

在每种工况下,分别选用两种车型(Tesla Model 3和Audi E-Tron)进行仿真,运行地图和轨迹分别如

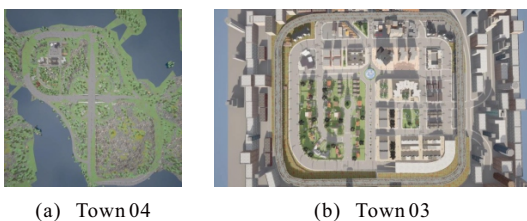


图4 仿真训练测试运行地图



(a) 山区运行场景



(b) 城市运行场景

图5 训练过程运行轨迹

图4和图5所示.

模型训练阶段采用城市启停和乡村变速工况数据,并按照7:2:1划分训练集、验证集和测试集;高速巡航和山路坡道工况数据则作为独立测试集,用于评估模型在非训练场景下的能耗估计准确性和泛化能力.选取平均绝对误差(MAE)和均方根误差(RMSE)作为模型的评价指标,具体计算如下所示:

$$\begin{cases} \text{MAE} = \frac{1}{n} \sum_{i=1}^n (|y_i - \hat{y}_i|), \\ \text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}. \end{cases} \quad (16)$$

其中:  $y_i$  为真实值,  $\hat{y}_i$  为预测值,  $n$  为数据点的数量.

#### 4.1.2 平台设置参数

实验平台基于CARLA 0.9.13版本,并通过Python API进行环境交互和数据采集.具体信息如表3所示.

表3 CARLA 平台参数配置

参数名称	值
地图环境	Town03, Town04
控制频率	20 Hz
仿真步长	0.05 s
采样字段	位置信息、速度、SOC、电量、油门等
车辆模型	Tesla Model 3, Audi E-Tron
摄像头配置	前视RGB相机(1280×720),安装于车头中央,视角90°;车辆自带BMS与环境传感器同步采集
数据规模	约38 h ≈ 2.7 × 10 <sup>6</sup> 帧,清洗后保留2.6 × 10 <sup>6</sup> 帧
采样和清洗	20 Hz采样; 剔除静止段以及传感器异常值,滑动均值平滑噪声
网络结构	感知编码器采用5层CNN(通道数64-128-256-256-512),后接双层Transformer编码器(隐藏维度512,注意力头数8),强化学习模块为双层MLP(隐藏层128、64)
超参数设置	学习率1 × 10 <sup>-4</sup> , batch size 128, 折扣因子 $\gamma = 0.98$ , soft update $\tau = 0.005$ , $\epsilon$ 初始0.9逐步衰减至0.05
训练时长	训练2 × 10 <sup>4</sup> 轮,约7 h
早停和调参范围	验证集R <sup>2</sup> 连续500轮无提升即早停; 学习率[1e-5, 5e-4], batch [64, 256]

#### 4.1.3 能耗模块设置

CARLA 仿真平台缺乏内置能耗模型,无法生成

实时能耗数据,因此,本文基于车辆动力学原理和车辆参数,自主构建实时能耗模型.根据车辆动力学方程,车辆所需牵引力包括空气阻力、滚动阻力、坡度阻力以及加速阻力,表述如下:

$$F_{\text{traction}} = \frac{1}{2}\rho C_d A v^2 + mg \cos(\theta) C_r + mg \sin(\theta) + ma. \quad (17)$$

其中:  $C_d$ 为空气阻力系数,  $A$ 为车辆迎风面积,  $v$ 为当前速度,  $m$ 为整车质量,  $\theta$ 为道路坡度角,  $C_r$ 为滚动阻力系数.

进而计算牵引功率  $P_{\text{traction}} = F_{\text{traction}} \cdot v$ , 考虑电机效率  $\eta$ , 驱动能耗为  $P_{\text{drive}} = P_{\text{traction}}/\eta$ . 在减速工况下, 引入再生制动机制, 回收功率为  $P_{\text{regen}} = \eta_{\text{regen}} \times mav$ . 此外, 辅助系统能耗  $P_{\text{aux}}$  作为固定值参与总功率消耗.

最终单位时间能耗计算如下所示:

$$E_{\text{total}} = \int \left( \frac{F_{\text{traction}} \cdot v}{\eta} + P_{\text{aux}} - P_{\text{regen}} \right) dt. \quad (18)$$

该模型通过 CARLA 的 Python API 实时获取车辆速度、加速度、道路坡度等信息, 实现不同驾驶行为和环境条件下的能耗动态估算, 提升仿真系统的物理一致性和工程参考价值.

#### 4.1.4 随机种子重复实验与统计分析

为进一步验证模型在不同随机初始化条件下的稳定性和统计可靠性, 本文针对 EARL-ECP 模型进行 5 次不同随机种子下的独立重复实验, 并计算主要评价指标的均值、标准差以及 95% 置信区间. 实验结果如表 4 所示.

表4 不同随机种子下 EARL-ECP 模型性能的统计结果

随机种子(序号)	MAE/%	RMSE/%	$R^2$ /%
0	0.172	0.243	99.66
1	0.181	0.256	99.75
2	0.188	0.246	99.73
3	0.171	0.238	99.68
4	0.179	0.257	99.77
均值 ± 标准差	0.178 ± 0.006	0.248 ± 0.007	99.72 ± 0.11
95%置信区间	± 0.005	± 0.006	± 0.10

由表 4 可见, EARL-ECP 模型在多次独立训练中均保持高度一致的性能, 其 MAE 为  $0.178 \pm 0.006\%$ , RMSE 为  $0.248 \pm 0.007\%$ , 决定系数  $R^2$  为  $99.72 \pm 0.11\%$ , 95% 置信区间分别控制在  $\pm 0.005\% \sim 0.01\%$  范围内. 实验结果表明, 该模型在随机初始化和经验采样扰动下几乎不受影响, 预测性能稳定, 具有优良的收敛性和鲁棒性.

为进一步验证模型收敛趋势的稳定性, 本文绘

制不同模型的学习曲线(训练轮次 vs  $R^2$ ), 结果如表 5 所示. 由表 5 可知, EARL-ECP 在约 5000 轮训练后即可达到  $R^2 \approx 95.7\%$ , 并在 10000 轮内迅速收敛至  $R^2 \approx 99.8\%$ , 相较于 Mamba、Informer、Transformer 等模型提前约 30% 达到收敛稳定阶段; LSTM 收敛最慢且波动较大. 该结果表明, EARL-ECP 不仅在精度上具有优势, 同时, 在训练效率和稳定性方面亦优于其他模型.

表5 不同模型在训练轮次下的  $R^2$  变化对比 (学习曲线数据)

Epoch	EARL-ECP	Mamba	Informer	Transformer	LSTM
0	65.2	60.1	59.8	57.4	50.7
2000	88.5	84.1	80.6	78.9	72.3
5000	95.7	91.3	88.2	86.9	82.5
10000	98.2	94.8	93.1	91.4	85.9
15000	99.3	96.5	95.4	93.6	87.2
20000	99.8	97.1	96.0	94.1	88.0

为进一步诊断过拟合, 本文对 EARL-ECP、DQN 与 TD3 在约 20000 轮训练过程中的平均奖励变化曲线进行对比分析, 结果如图 6 所示.

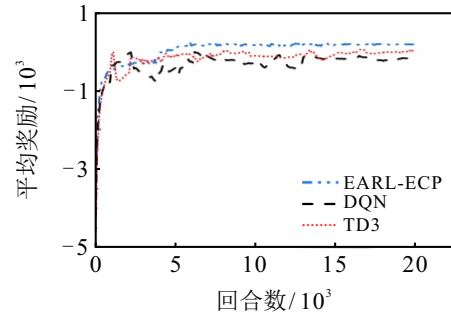


图6 模型训练收敛对比

由图 6 可见: EARL-ECP 在训练初期即实现显著奖励提升, 并于较短周期内平稳收敛, 曲线无明显震荡或过度上升, 表明模型未出现过拟合; 相比之下, DQN 与 TD3 虽然在早期波动较大, 但是最终趋于稳定. 总体而言, EARL-ECP 以更快的收敛速度和更高的稳定性表现出优异的学习效率和泛化能力, 验证了其在复杂环境和电池退化条件下的稳健性和应用潜力.

## 4.2 预测效果及其对比实验

### 4.2.1 车辆行驶能耗预测效果

为评估模型在电池退化条件下的鲁棒性, 本节引入不同 SOH 水平, 模拟轻中度衰减场景, 通过预测行驶过程中的功率消耗(包括驱动功率  $P_{\text{battery}}$  和辅助功率  $P_{\text{aux}}$ ), 计算车辆在特定时间段内的总能耗  $E$ , 计算如下所示:

$$E = \int (P_{\text{battery}} + P_{\text{aux}}) dt. \quad (19)$$

随后,依据能耗与电池容量的关系,反推出 SOC 的变化过程,具体计算如下所示:

$$\text{SOC}_{t+1} = \text{SOC}_t - \alpha \cdot \frac{E}{Q_{\text{nominal}}} \cdot 100\%, \quad (20)$$

其中 $\alpha$ 为可调节系数,用于校准能耗-电量关系.

基于上述 SOC 估算过程,为了更直观地展示模型在电池性能退化条件下的估计能力,本文进一步结合当前预测的 SOC 值和能耗水平,估算车辆的剩余可行驶里程 $D_{\text{remain}}$ ,计算如下所示:

$$D_{\text{remain}} = \frac{\text{SOC}_{t+1} \cdot Q_{\text{actual}}}{\bar{P}/v_{\text{avg}}}. \quad (21)$$

其中: $Q_{\text{actual}}$ 为退化后电池的实际可用容量, $\bar{P}$ 为平均功率消耗, $v_{\text{avg}}$ 为平均车速.

图7和图8均为电池健康状态变化时,模型在两种车型上的估计表现,验证其在退化工况下的适应性和应用潜力.

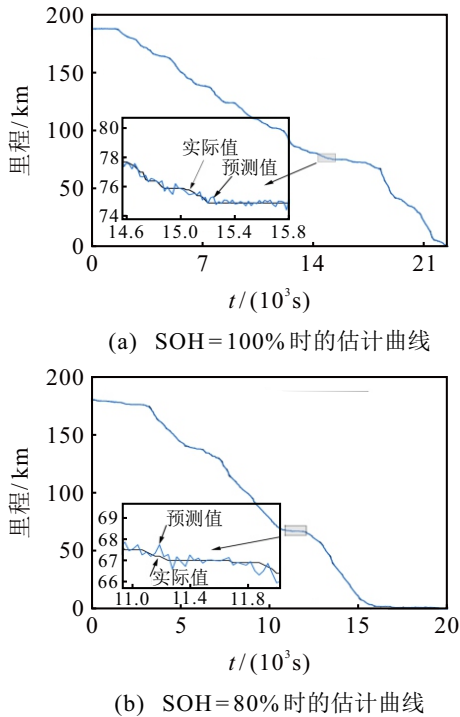
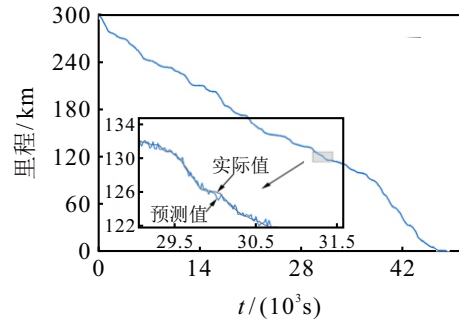


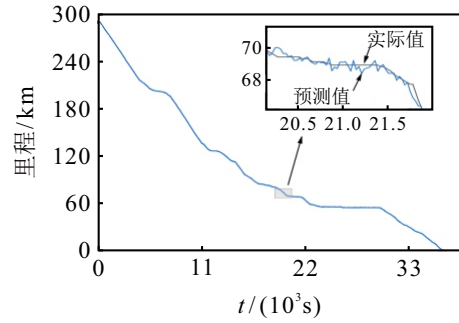
图7 车型 A 剩余里程估计曲线

结合图7和图8中两种车型在SOH = 100%和SOH = 80%条件下的预测结果可见,在本文所选的两款车型(轿车和SUV)中,即便在不同SOH状态下,模型的估计曲线仍然与仿真参考值保持了较高一致性,误差始终处于较低水平.这表明模型在所测试的多类工况和健康状态下具备良好的精度和稳定性.需要指出的是,其在其他类型车辆上的适用性仍然需要进一步验证.

为全面评估所建立网络的估计性能,本节对其



(a) SOH=100%时的估计曲线



(b) SOH=80%时的估计曲线

图8 车型 B 剩余里程估计曲线

在不同车型与不同SOH条件下的MAE、RMSE和决定系数 $R^2$ 进行统计分析,其结果如表6所示.整体而言,网络在各工况下均保持较低的估计误差和接近1的 $R^2$ ,体现出良好的精度、鲁棒性和泛化能力.进一步结合近似 $t$ 检验结果可见,模型在不同SOH条件下对两种车型均表现出较强的一致性:在SOH为100%和90%时,两种车型在MAE和RMSE上的差异均不显著( $p > 0.05$ ),表明所提出模型在高健康状态下对不同车型均能够保持稳定且一致的估计性能;仅在SOH为90%的 $R^2$ 上出现显著差异,反映了少量车型差异对拟合程度产生的轻微影响.

表6 不同SOH条件下模型估计性能及其显著性分析

SOH/%	指标/%	A车	B车	$t$ 值	$p$ 值	显著性
100	MAE	0.173	0.170	0.847	0.401	-
	RMSE	0.239	0.248	-1.789	0.079	-
	$R^2$	99.77	99.81	-1.936	0.058	-
90	MAE	0.171	0.178	-1.942	0.057	-
	RMSE	0.252	0.253	-0.192	0.849	-
	$R^2$	99.75	99.69	2.905	0.005	**
80	MAE	0.184	0.167	4.684	$1.7 \times 10^{-5}$	**
	RMSE	0.255	0.243	2.332	0.023	*
	$R^2$	99.63	99.83	-9.682	$1.0 \times 10^{-13}$	**

当SOH降至80%时,3项指标均呈现显著性差异( $p < 0.05$ 或 $p < 0.01$ ),表明电池退化后不同车型间的表现分化更加明显.但是即便在此状态下,整体误差仍然处于较低水平,表明所提出模型在实验两

种车型以及不同衰退程度下依然具备良好的适用性.

### 4.2.2 多源数据输入对预测性能与实时性的影响分析

为进一步验证多源输入对于能耗预测模型在性能和实时性方面的影响,本文设计对比实验,分别采用传统 BMS 时序数据与融合摄像头图像的多源数据作为输入,比较两种方式在不同 SOH 条件下的预测精度、参数量以及推理时间差异.实验结果如表 7 所示.从整体结果来看,多源输入(图像 + BMS)在各 SOH 状态下均取得更低的 MAE、RMSE 以及更高的  $R^2$ ,相比于传统输入表现更优,表明引入图像信息能够显著增强模型的特征感知能力和估计性能.基于  $t$  检验结果,多数指标均达到显著或高度显著差异:在 SOH 为 100% 和 80% 条件下,MAE、RMSE 和  $R^2$  全部呈现高度显著提升 ( $p < 0.01$ ),表明无论电池处于健康或退化状态,多源输入均能够明显提高预测精度;在 SOH = 90% 时,虽然 MAE 差异不显著,但是 RMSE 与  $R^2$  仍然表现出显著差异,表明图像信息的融合在中等健康状态下依然有助于提升模型稳定性和拟合能力.尤其是在  $R^2$  指标上,多源模型在各 SOH 区间均保持显著优势,其值在 SOH = 80% 时仍然可达 99.88%,显示模型对于退化特征的捕捉能力较强.

表7 多源与单源输入模型的性能及统计显著性对比

SOH/%	指标	传统输入	多源输入	$t$ 值	$p$ 值	显著性
100	MAE%	0.202	0.176	6.644	$1.2 \times 10^{-8}$	**
	RMSE%	0.284	0.239	8.300	$1.9 \times 10^{-11}$	**
	$R^2$ %	97.81	99.59	-86.174	$<1e-12$	**
	推理时间/ms	46	69	-	-	-
90	MAE%	0.188	0.185	0.779	0.439	-
	RMSE%	0.271	0.257	2.566	0.013	*
	$R^2$ %	97.58	99.54	-94.888	$<1e-12$	**
	推理时间/ms	41	66	-	-	-
80	MAE%	0.195	0.173	5.778	$3.2 \times 10^{-7}$	**
	RMSE%	0.268	0.238	5.731	$3.8 \times 10^{-7}$	**
	$R^2$ %	98.05	99.88	-88.594	$<1e-12$	**
	推理时间/ms	44	69	-	-	-
参数量/M		2.5	3.7	-	-	-

在性能提升的同时,多源模型的参数量虽然由 2.5 M 增至 3.7 M,但是推理时间保持在 70 ms 以内,且不同 SOH 条件下的实时响应均满足应用需求.其推理时间  $T_{infer}$  与模型规模  $P$  和输入维度  $D$  的关系近似为线性关系,即

$$T_{infer} \approx \alpha \cdot P + \beta \cdot D + \epsilon. \quad (22)$$

其中:  $\alpha$ 、 $\beta$  为与硬件相关的常数,  $\epsilon$  为常量偏移项.该

分析可进一步佐证所提出方法在确保预测精度的同时,仍然具备良好的计算效率和实时性保障.

### 4.2.3 不同算法估计效果对比

为全面评估所提出 EARL-ECP 模型的性能,本节选取工程领域常用的预测方法 —— LSTM、XGBoost、Informer 与 Mamba 进行对比.

上述方法广泛应用于电动汽车能耗和电池状态预测,能够从不同角度反映模型在时序建模、非线性表达以及复杂工况适应性方面的能力.其在不同 SOH 条件下的关键性能指标如表 8 所示.为便于直观地展示模型在实际工况下的估计能力,本节重点绘制 A 车型在 SOH = 90% 条件下,各算法的剩余里程预测曲线(如图 9 所示).该 SOH 水平对应电池中度衰减阶段,可较真实地反映模型在退化工况下的泛化性能和鲁棒性.由图 9 可见,各方法在 SOH = 90% 时的预测精度差异明显:LSTM 曲线波动较大,多处出现尖峰和滞后;XGBoost 能够跟随整体趋势,但是在局部区间存在持续偏差;Informer 和 Mamba 拟合较为平稳,其中 Mamba 在常用方法中表现最佳;相比之下,所提出 EARL-ECP 曲线始终与真实值高度重合,即使在放大区间亦保持最小误差,展现出更强的环境感知能力和时序稳定性.

表8 不同算法关键性能指标

算法	车型	A			B		
	SOH/%	MAE/%	RMSE/%	$R^2$ /%	MAE/%	RMSE/%	$R^2$ /%
LSTM	100	0.694	1.138	90.72	0.684	1.091	91.03
	90	0.739	1.204	85.02	0.744	1.121	84.51
	80	0.781	1.309	76.02	0.778	1.321	77.44
XGBoost	100	0.258	0.339	96.72	0.271	0.351	96.41
	90	0.332	0.401	90.84	0.318	0.389	91.02
	80	0.362	0.453	83.21	0.347	0.449	83.55
Informer	100	0.249	0.332	98.10	0.245	0.329	98.16
	90	0.279	0.366	92.21	0.273	0.359	92.74
	80	0.349	0.395	88.37	0.336	0.381	88.62
Mamba	100	0.238	0.311	98.72	0.239	0.314	98.65
	90	0.261	0.346	98.32	0.264	0.352	98.04
	80	0.272	0.352	97.63	0.283	0.382	96.98
EARL-ECP	100	0.173	0.239	99.77	0.170	0.248	99.81
	90	0.171	0.252	99.75	0.178	0.253	99.69
	80	0.184	0.255	99.63	0.167	0.243	99.83

结合表 7 的性能指标进一步验证了上述结论:LSTM 随 SOH 降低误差显著增大(MAE 从 0.694 升至 0.781,  $R^2$  下降了 14.7%),稳定性较差;XGBoost 在 3 种 SOH 下整体误差偏高,尤其是在 SOH = 80% 时 MAE 超过 0.36%,难以处理复杂非线性影响;Informer 与 Mamba 均保持较好的拟合能力,其中

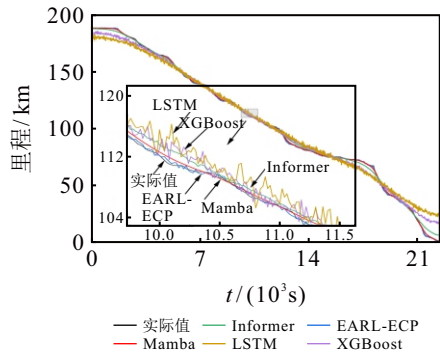


图9 不同算法剩余里程估计曲线

Mamba 在传统方法中最为稳定 (SOH = 90% 时  $R^2$  仍然超过 98%); 相比之下, EARL-ECP 在不同 SOH 条件下均取得了最优结果, A 车 SOH = 90% 时 MAE 仅为 0.171%,  $R^2$  达到 99.75%, 且随 SOH 衰减性能

变化极小, 展现出突出的泛化性和自校准能力.

综上, 常用方法在电池衰减和复杂工况下均存在不同程度的性能下降, 而所提出模型在工况波动、噪声干扰以及电池退化的综合影响下仍然保持高精度预测, 充分验证了所提出环境感知强化学习框架的有效性.

### 4.3 消融实验

为验证融合环境感知强化学习的电动汽车能耗预测 (EARL-ECP) 算法进行实时能耗估计时关键模块的有效性, 本文设计消融实验. 以传统的 Transformer 为基础, 通过替换或增加创新模块, 测试不同车型在不同工况下的性能指标. 表 9 消融实验结果显示, 各模块对于能耗预测的贡献度存在差异. 下面以车型 A 的实验数据为例进行具体分析.

表9 消融实验结果

网络	改进									
	DQN	ERL	SAPSM	车型 A			B			
				SOH/%	MAE/%	RMSE/%	$R^2$ /%	MAE/%	RMSE/%	$R^2$ /%
D	√			100	0.271	0.341	96.02	0.276	0.344	95.98
				90	0.285	0.343	95.49	0.262	0.326	96.61
				80	0.266	0.326	96.72	0.281	0.321	96.09
E-T		√		100	0.259	0.316	97.10	0.279	0.319	96.32
				90	0.295	0.344	93.21	0.294	0.346	93.02
				80	0.309	0.371	91.62	0.320	0.386	91.09
S-T			√	100	0.252	0.306	97.88	0.263	0.310	97.55
				90	0.288	0.334	93.93	0.283	0.333	93.63
				80	0.304	0.378	91.72	0.293	0.368	91.97
DE	√	√		100	0.259	0.317	97.68	0.220	0.254	98.16
				90	0.269	0.320	97.05	0.271	0.318	97.52
				80	0.239	0.300	97.70	0.210	0.247	98.24
DS	√		√	100	0.254	0.314	97.15	0.256	0.315	97.33
				90	0.259	0.300	97.74	0.231	0.267	97.95
				80	0.258	0.322	97.19	0.259	0.305	97.39
EARL-ECP(本文)	√	√	√	100	0.173	0.239	99.77	0.170	0.248	99.81
				90	0.171	0.252	99.75	0.178	0.253	99.69
				80	0.184	0.255	99.63	0.167	0.243	99.83

在进行改进模块有效性验证前, 首先对各对比网络的组成进行说明: D 表示仅使用 DQN; E-T 表示引入所提出路况感知模块但是未使用 DQN (以 Transformer 作为基础结构); S-T 表示使用优先采样机制但是未使用 DQN 和路况感知模块 (同样基于 Transformer); DE 和 DS 分别表示 DQN 与路况感知模块、DQN 与优先采样机制的组合. 该设置旨在清晰评估不同模块对于预测性能的独立和联合贡献.

为验证各改进模块的效果, 对比表 9 中 D、E-T 与 S-T 的结果可见: D 网络在不同 SOH 条件下变化最小, 如 A 车  $R^2$  从 96.02% 仅提升至 96.72% (增幅

约 0.7%), MAE 与 RMSE 的波动亦保持在 3% 以内; 而 E-T 和 S-T 随 SOH 降低均出现 6% ~ 10% 的误差增幅以及 4% ~ 6% 的  $R^2$  下降, 显示其对于衰减工况更为敏感. 表明 DQN 能够有效稳定基础网络的时序一致性和动态适应性.

进一步比较加入 ERL 的 DE 与加入 SAPSM 的 DS, 可以观察到两者均显著优于 D 网络. DE 在 SOH = 100% 时  $R^2$  相比于 D 提升了约 1.7%, 在 SOH = 80% 时 MAE 和 RMSE 分别降低了 8% ~ 12%; DS 在 SOH = 90% 时 A 车 MAE 和 RMSE 分别降低了约 9% 和 11%,  $R^2$  提升了约 2.3%. 这表明 ERL 可增强环境

感知和自校准能力,而 SAPSM 能够有效强化关键场景学习。

综合来看,EARL-ECP 在所有 SOH 条件下均取得了最优结果。如:在 SOH = 100% 时, A 车  $R^2$  达到了 99.77%, 相较于 DE 与 DS 分别提升了 2.1% 与 2.4%, MAE 和 RMSE 降幅均超过 25%; 在 SOH = 80% 时, EARL-ECP 仍然保持  $R^2 \approx 99.6%$ , 变化幅度不足 0.2%, 显著优于其余模型。结果表明, DQN + ERL + SAPSM 的联合机制在提升模型精度、稳定性和跨工况泛化能力方面具有明显协同增益。

## 5 结论与展望

本文提出了一种面向电动汽车的能耗预测模型, 通过将折扣未来能耗的估计引入为时序一致性正则项 (而非控制回路中的优化目标), 并结合多源环境感知与 TD 学习, 实现了对能耗/续航里程的高精度点预测和跨工况自校准能力。所提出融合环境感知与强化学习的预测框架, 在多尺度环境表征、动态决策建模和关键场景识别的基础上, 引入了参数共享机制以提升特征提取与策略学习的协同效率。所提出方法能够有效应对复杂工况下的能耗预测挑战, 并在所测试车型和工况范围内展现出良好的泛化性能 (更多车型仍然需要进一步验证)。实验结果表明, 所提出模型在多种典型工况下的预测精度均优于现有主流方法, 为电动汽车能耗智能管理提供了可靠支撑。通过多模态特征的深度融合, 模型能够动态适应不同道路和环境条件, 进一步提升了预测精度和稳健性。未来, 该技术可拓展至车-路-云协同能源管理场景, 支持智能充电调度、换电策略优化等应用, 为新能源汽车朝更高水平的智能化、自主化方向发展提供基础。

## 参考文献 (References)

- [1] 曹慕昆, 王逸诚, 刘新龙, 等. 轨迹数据驱动的新能源汽车充电桩选址模型研究[J]. 控制与决策, 2025, 40(9): 2739-2747.  
(Cao M K, Wang Y C, Liu X L, et al. Research on location model of new energy vehicle charging station driven by trajectory data[J]. Control and Decision, 2025, 40(9): 2739-2747.)
- [2] Peng Z R, He Z Y. Optimization of regenerative braking control strategy for dual-motor electric vehicles based on deep reinforcement learning[J]. IEEE Transactions on Intelligent Transportation Systems, 2025, 26(7): 10954-10967.
- [3] 彭自然, 杨肖阳, 肖伸平. 基于 EKF-HInformer 模型估计汽车动力电池的 SOC & SOH[J]. 电子测量与仪器学报, 2025, 39(3): 21-33.  
(Peng Z R, Yang X Y, Xiao S P. SOC and SOH of the battery are estimated based on the EKF-HInformer model[J]. Journal of Electronic Measurement and Instrumentation, 2025, 39(3): 21-33.)
- [4] Armenta-Déu C, Boucheix B. Evaluation of lithium-ion battery performance under variable climatic conditions: Influence on the driving range of electric vehicles[J]. Future Transportation, 2023, 3(2): 535-551.
- [5] 彭自然, 王顺豪, 肖伸平, 等. 基于 KA Informer 的电动汽车动力电池荷电状态和健康状态估算[J]. 电工技术学报, 2025, 40(19): 6378-6394.  
(Peng Z R, Wang S H, Xiao S P, et al. State of charge and state of health estimation of electric vehicle power battery based on KA informer model[J]. Transactions of China Electrotechnical Society, 2025, 40(19): 6378-6394.)
- [6] 邢玉伟, 王展华, 杨华龙. 需求可离散拆分电动汽车充电策略和路径优化问题[J]. 控制与决策, 2025, 40(3): 987-995.  
(Xing Y W, Wang Z H, Yang H L. Electric vehicle charging strategies and routing optimization under discrete split demands[J]. Control and Decision, 2025, 40(3): 987-995.)
- [7] Ullah I, Liu K, Yamamoto T, et al. A comparative performance of machine learning algorithm to predict electric vehicles energy consumption: A path towards sustainability[J]. Energy & Environment, 2022, 33(8): 1583-1612.
- [8] Peng Z R, Yang X Y. Short-and medium-term power load forecasting model based on a hybrid attention mechanism in the time and frequency domains[J]. Expert Systems with Applications, 2025, 278: 127329.
- [9] 田慧欣, 秦鹏亮, 李坤, 等. 基于 HI-DD-AdaBoost. RT 的锂离子动力电池 SOH 预测[J]. 控制与决策, 2021, 36(3): 686-692.  
(Tian H X, Qin P L, Li K, et al. Prediction of Li-ion battery SOH based on HI-DD-AdaBoost. RT[J]. Control and Decision, 2021, 36(3): 686-692.)
- [10] Bi J, Wang Y X, Sai Q Y, et al. Estimating remaining driving range of battery electric vehicles based on real-world data: A case study of Beijing, China[J]. Energy, 2019, 169: 833-843.
- [11] Sun S, Zhang J, Bi J, et al. A machine learning method for predicting driving range of battery electric vehicles[J]. Journal of Advanced Transportation, 2019: 4109148.
- [12] LeCun Y, Bengio Y, Hinton G. Deep learning[J]. Nature, 2015, 521(7553): 436-444.
- [13] Chen J X, Zhang Y, Wu J, et al. SOC estimation for lithium-ion battery using the LSTM-RNN with extended input and constrained output[J]. Energy, 2023, 262: 125375.
- [14] Hong J C, Liang F W, Yang H X, et al. Multi-forward-step state of charge prediction for real-world electric vehicles battery systems using a novel LSTM-GRU hybrid neural network[J]. eTransportation, 2024, 20: 100322.

- [15] 彭自然, 王顺豪, 肖伸平, 等. 一种精确估算电动汽车动力电池 SOC & SOH 的循环门控模型[J]. *电子测量与仪器学报*, 2024, 38(9): 11-23.  
(Peng Z R, Wang S H, Xiao S P, et al. Cycle gating model for accurate estimation of SOC & SOH of power battery in electric vehicles[J]. *Journal of Electronic Measurement and Instrumentation*, 2024, 38(9): 11-23.)
- [16] de Cauwer C, Verbeke W, Coosemans T, et al. A data-driven method for energy consumption prediction and energy-efficient routing of electric vehicles in real-world conditions[J]. *Energies*, 2017, 10(5): 608.
- [17] Mohanty P K, Jena P, Padhy N P. Electric vehicle state-of-charge prediction using deep LSTM network model[C]. *IEEE International Conference on Power Electronics, Drives and Energy Systems*. Jaipur, 2022: 1-6.
- [18] Ren W J, Xie X Y, Yi Y, et al. The LiFePO<sub>4</sub> battery lifespan SoC estimation using Ham-Informer and internal pressure[J]. *Journal of Energy Storage*, 2024, 88: 111474.
- [19] Kim D, Shim H G, Eo J S. A machine learning method for EV range prediction with updates on route information and traffic conditions[C]. *Proceedings of the AAAI Conference on Artificial Intelligence*. Online, 2022: 12545-12551.
- [20] Feng Z Y, Zhang J, Jiang H, et al. Energy consumption prediction strategy for electric vehicle based on LSTM-transformer framework[J]. *Energy*, 2024, 302: 131780.
- [21] Zhao T T, Li G X, Zhao T, et al. Learning explainable task-relevant state representation for model-free deep reinforcement learning[J]. *Neural Networks*, 2024, 180: 106741.
- [22] Bao Z Y, Nie J H, Lin H P, et al. TTSNet: State-of-charge estimation of Li-ion battery in electrical vehicles with temporal transformer-based sequence network[J]. *IEEE Transactions on Vehicular Technology*, 2024, 73(6): 7838-7851.

### 作者简介

彭自然 (1969-), 男, 副教授, 博士, 主要研究方向为人工智能、信号处理、智能检测仪表, E-mail: pengziran@hut.edu.cn;

杨肖阳 (2000-), 男, 硕士生, 主要研究方向为电动汽车动力电池荷电状态、健康状态以及能耗的预测, E-mail: 2873633605@qq.com;

舒中宾 (1970-), 男, 副教授, 硕士, 主要研究方向为电力电子与电力传动, E-mail: 123998439@qq.com.