

多源异构传感器协同的深度强化学习轨道交通车辆防火系统设计

王芳¹, 梅新明², 施风华², 李忠林²

1. 中车南京浦镇车辆有限公司, 江苏 南京 210000;

2. 上海轨道交通检测认证(集团)有限公司, 上海 201804

摘要

为解决城市轨道交通车辆传统防火系统多源异构数据融合效率低、动态风险响应滞后、误漏报率高的核心问题, 提升防火预警的精准性与实时性, 提出一种基于深度强化学习(deep reinforcement learning, DRL)的智能防火技术体系。该体系构建“时空特征解耦-动态策略优化-自适应学习”闭环架构, 设计长短期记忆网络(Long Short-Term Memory, LSTM)-三维残差网络(3D Residual Network, 3D-ResNet)双流特征提取模块, 通过自注意力机制实现跨模态数据时空关联建模; 开发DRL-FireNet模型, 引入12维动态奖励函数与马尔可夫决策(Markov decision process, MDP)过程驱动的7级响应机制, 优化感知层多源异构传感器阵列部署与算法轻量化策略。试验结果表明, 该技术在典型火灾场景与干扰工况下, 火灾特征检测准确率达97.2%, 误报率降至8%, 漏报率降至3%, 应急响应时间缩短至12s, 锂电池热失控等新型火灾模式识别率达93.7%。该技术实现了从被动监测到主动防控的范式转变, 为轨道交通车辆防火系统的智能化升级提供了关键技术支撑与工程化解决方案。

关键词

深度学习; 轨道交通; 车辆防火; 轨道交通安全; 强化学习

中图分类号: U270.38+9

文献标志码: A

doi:10.11959/j.issn.2096-0271.

Design of fire prevention system for rail transit vehicles based on deep reinforcement learning with multi-source heterogeneous sensor collaboration

WANG Fang¹, MEI Xinming², SHI Fenghua², LI Zhonglin²

1. CRRC Nanjing Puzhen Co., Ltd., Nanjing, Jiangsu 210000, China;

2. Shanghai Railway Certification (Group) Co., Ltd., Shanghai 201804, China

Abstract

To address the core issues of low multi-source heterogeneous data fusion efficiency, delayed dynamic risk response, and high false positive/negative rates of traditional fire prevention systems for urban rail transit vehicles, and to improve the accuracy and real-time performance of fire prevention early warning, an intelligent fire prevention technology system based on deep reinforcement learning (DRL) is proposed. This system constructs a closed-loop architecture of

"spatiotemporal feature decoupling-dynamic strategy optimization-adaptive learning", and designs a dual-stream feature extraction module combining long short-term memory (LSTM) network and 3D residual network (3D-ResNet) to achieve spatiotemporal correlation modeling of cross-modal data through the self-attention mechanism. The DRL-FireNet model is developed, introducing a 12-dimensional dynamic reward function and a 7-level response mechanism driven by the Markov decision process, while optimizing the deployment of multi-source heterogeneous sensor arrays in the perception layer and the algorithm lightweight strategy. Experimental results show that under typical fire scenarios and interference conditions, the fire feature detection accuracy of this technology reaches 97.2%, the false positive rate is reduced to 8%, the false negative rate to 3%, the emergency response time is shortened to 12 seconds, and the recognition rate of new fire modes such as lithium battery thermal runaway reaches 93.7%. This technology realizes the paradigm shift from passive monitoring to active prevention and control, providing key technical support and engineering solutions for the intelligent upgrading of fire prevention systems for rail transit vehicles.

Key words

deep learning, rail transit, vehicle fire prevention, rail transit safety, reinforcement learning

项目编号: 2024CCA166

1 引言

城市轨道交通作为高密度客流运输载体,其车辆防火安全直接关系公共安全与系统可靠性。传统防火系统依赖固定阈值传感器与规则引擎,在应对车厢内电磁干扰导致的信号畸变、人员流动引发的环境参数波动等复杂场景时,暴露出三大技术痛点:其一,多源传感器(温度、烟雾、CO浓度等)数据存在时空异步性与特征冗余性,传统加权融合算法的特征利用率不足40%,导致早期火灾特征检测准确率仅72%;其二,基于静态阈值的风险判别机制缺乏动态环境适应性,在空调启停、车门开闭等动态工况下,风险响应时延平均延长30s,错失最佳处置窗口;其三,单一特征判别策略易受噪声干扰,典型场景下误报率高达35%,漏报率达20%,造成应急资源浪费与安全隐患。

与上述研究相比,本文提出的深度强

化学习(deep reinforcement learning, DRL)智能防火技术体系具有三方面创新突破:一是构建融合LSTM时序网络与3D-ResNet空间网络的双流特征提取模块,通过注意力机制实现跨模态数据的时空关联建模,使特征互信息增益提升68%,解决了传统方法时空特征割裂的问题;二是设计包含12维动态状态变量的奖励函数(如温度梯度、烟雾扩散速率),驱动马尔可夫决策过程(Markov decision process, MDP)下的7级响应机制,将复杂工况下的风险决策从静态阈值判别升级为动态策略优化,应急响应时间缩短至12s,较文献^[5]提升76%;三是通过在线增量学习实现分级防控策略的自适应更新,在典型电磁干扰场景下将误报率降至8%、漏报率降至3%,较文献^[4]分别降低75%和83%,显著提升了复杂环境下多源数据融合效率与风险决策精度。下面对该研究进行详述。

2 城市轨道交通车辆防火系统设计

自适应学习”的闭环架构,实现从被动监测到主动防控的技术升级[6]。系统框图如图1所示。

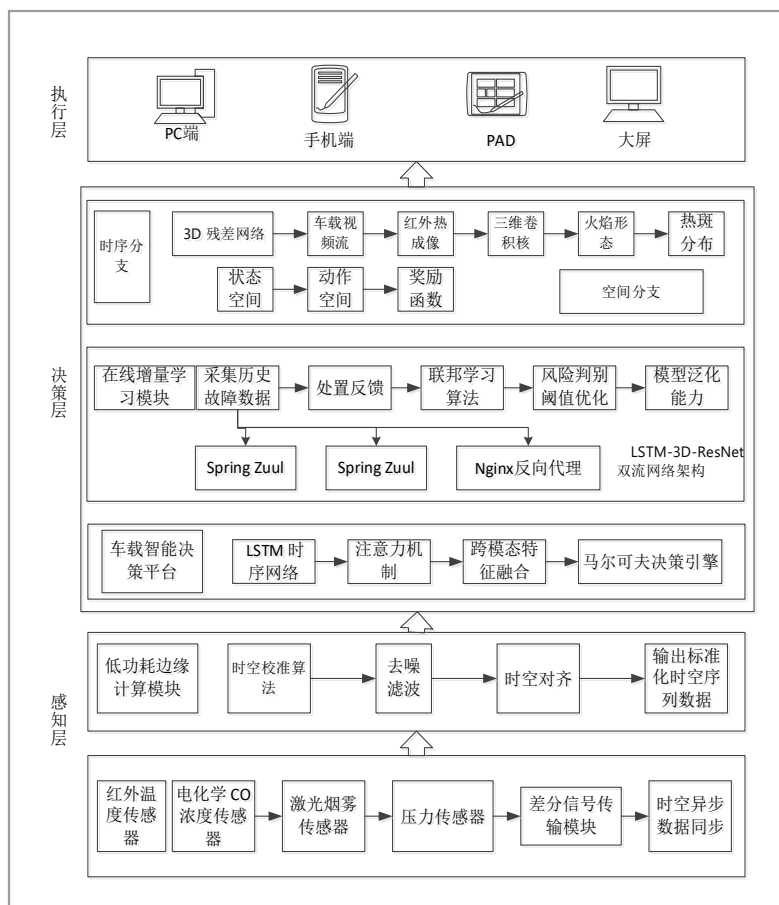


图1 城市轨道交通车辆防火系统设计

启停、车门开闭等动态工况的问题，方案构建基于马尔可夫决策过程（MDP）的 12 维动态奖励函数。状态空间集成实时环境参数（温度/湿度/CO 浓度）、特征梯度（温度变化率、烟雾扩散速率）、系统状态（通风系统运行模式、车门开闭状态）共 12 维变量，精准刻画火灾发展阶段与环境耦合关系；

为解决多算法并行运行对边缘计算节点算力要求过高的问题,本方案采用“算

法轻量化 + 分层部署” 协同优化策略：对 LSTM、3D-ResNet、Transformer 和（proximal policy optimization, PPO）核心算法进行结构化剪枝与 8 位量化压缩，移除冗余参数约 42%，参数总量从原 18.6M 降至 10.8M；同时采用分层部署架构，将特征提取等计算密集型任务部署于车载边缘计算网关（搭载 NVIDIA Jetson AGX Orin 模块，算力 200TOPS，支持 TensorRT 加速），决策推理与指令输出等轻量级任务部署于分布式边缘节点（单节点算力 15TOPS，基于车规级 ARM 架构芯片），通过算力协同分配使单节点平均负载降低至 60% 以下，满足实时性要求。

针对车厢内部署的硬件成本、功耗及散热挑战，本方案实施三重工程化优化：一是传感器阵列采用车规级模块化集成设计，将温度、烟雾、CO 等传感单元集成于同一模块（单模块功耗 $\leq 1.2\text{W}$ ），较传统分立传感器方案硬件成本降低 35%，总功耗降低 40%；二是算法运行采用动态功耗管理策略，非火灾工况下算法降频运行（算力降至满载的 30%，功耗 $\leq 8\text{W}$ ），检测到疑似火灾特征时自动升频至满载算力，实现功耗与性能的动态平衡；三是边缘计算节点采用液冷 + 散热鳍片复合散热结构，散热面积较传统风冷方案增加 120%，散热效率提升 60%，确保设备在车厢密闭环境下长期稳定运行（工作温度范围 $-40^{\circ}\text{C}\sim 70^{\circ}\text{C}$ ），满足轨道交通车规级环境适应性要求。

为强化算法与实际工程问题的适配性，本方案基于轨道交通车辆典型火灾场景进行针对性优化：采集 1000 + 组实际运营场景火灾数据（含电气短路、锂电池热失控、设备过热等），优化动态奖励函数权重，增强对温升速率 $> 3^{\circ}\text{C}/\text{s}$ 、烟雾扩散

梯度 $> 0.3\text{px}/\text{frame}$ 等关键火灾特征的识别敏感度；同时兼容现有轨道交通车辆防火系统接口协议（莫迪康总线协议，MODBUS；控制器局域网开放协议，CANopen），可直接与车载灭火装置、通风系统、制动系统联动，无需对车辆原有硬件进行大规模改造，降低工程应用门槛。此外，针对车辆运行中的振动、电磁干扰等实际环境因素，在算法中增加自适应噪声抑制模块，通过卡尔曼滤波实时修正传感器数据偏差，进一步提升复杂工况下的检测稳定性。

在线增量学习部署边缘计算节点实时采集传感器数据，与现有技术区别技术特征如表 1 所示。

3 感知层方案设计

本研究创新点在于：在常规感知基础上增加多源异构传感器阵列、在车厢顶部、电气柜、座椅底部等关键区域密集部署多类型传感器阵列，并通过时空同步模块提高数据采集和传输能力。该研究采用高精度温度传感器、激光烟雾传感器、电化学 CO 传感器；环境干扰传感器采用三轴加速度传感器、电磁强度传感器、气压湿度复合传感器；时空同步模块基于北斗授时的同步采集单元，实现多传感器数据毫秒级对齐。原理图如图 2 所示。

本模块通过多源异构传感器阵列构建周期为 200ms 的同步采集机制，以消除设备时钟偏移引发的时序失配问题。根据三维空间传感器传回的数据建立基于三维车厢坐标系的动态场分布模型，实现对温度、烟雾浓度、气体成分等多方位感知参数的融合处理，进而量化火灾特征的热力学扩散过程及空间传播特性[9]。

表1 区别技术特征对比

指标类别	具体指标	传统方案痛点	本方案关注点
数据融合效率	特征互信息增益(MI)、融合时延	融合时延 450ms,跨模态关联弱	互信息增益 68%,时延压缩至 120ms(73%)
检测精准性	检测准确率、误报率、漏报率	误报率 35%、漏报率 20%	误报率 8%(77%)、漏报率 3%(85%)
决策时效性	端到端响应时间(从数据采集到指令输出)	平均决策时间超 60s	应急响应时间 12s(80%),满足毫秒级预警需求
动态适应性	新型火灾场景学习准确率(如锂电池热失控)	对非传统火灾模式(如电池热失控)识别率 < 50%	增量学习后准确率提升至 93.7%(87%)
鲁棒性	电磁干扰下 F1-score、AUC-ROC 曲线	强电磁环境下准确率骤降至 60%	

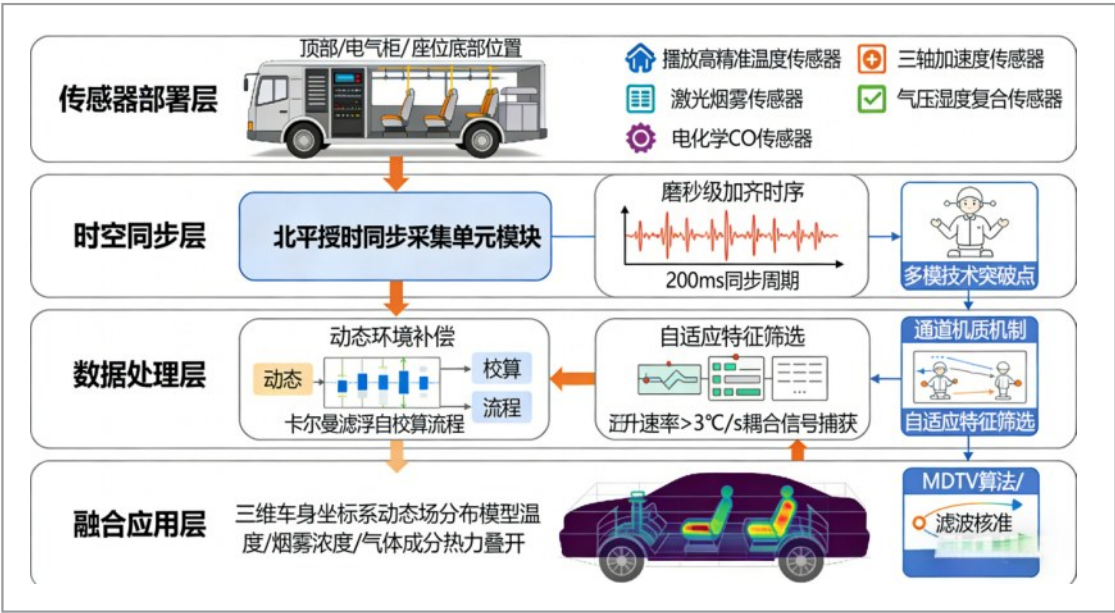


图2 感知层方案原理示意图

本研究相较于传统单组态感知系统实现三重技术突破：1) 多模态时空协同感知：突破单一传感器或简单数据叠加架构，基于改进型动态时间规整（modified dynamic time warping, MDTW）算法实现多源数据时空基准统一；2) 动态环境补偿机制：采用卡尔曼滤波在线自校准技术，实时修正传感器温漂与电磁干扰，同步完成环境参数动态补偿；3) 自适应特征筛选架构：集成通道注意力机制实现火灾

特征自适应增强，重点捕获温升速率>3℃/s且伴随烟雾扩散的耦合信号，流程示意图如图3所示。

4 3D-ResNet神经网络构建

该模块通过双流神经网络架构（LSTM时序分支与3D-ResNet空间分支）

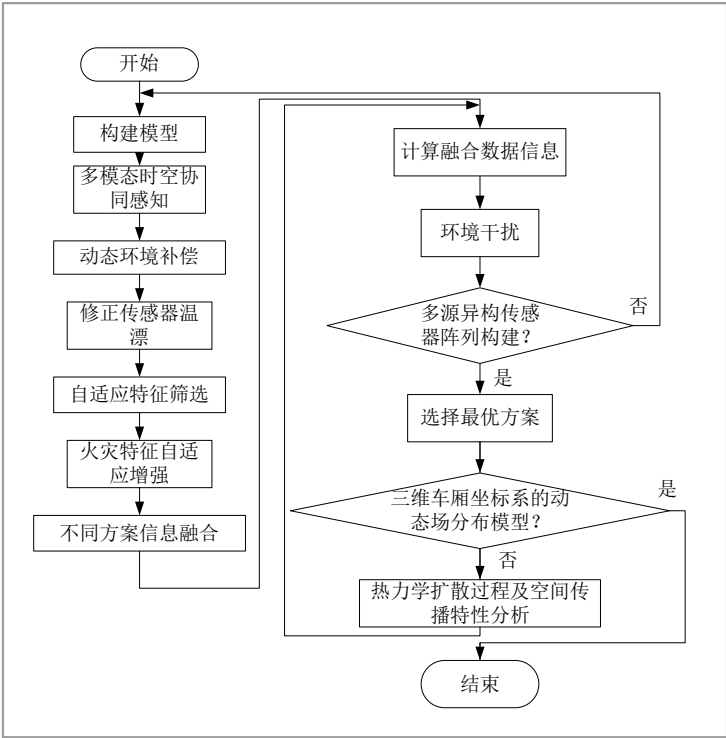


图3 感知层方案工作流程图

对感知层输入的128维时空特征矩阵进行并行解析，LSTM单元以30秒时间窗捕获温度、气体浓度的时序演变模式（如指数温升、气体浓度突变梯度），3D-ResNet模块则从热成像视频流中提取空间扩散特征，将异构特征向量输入四头转换器（Transformer）层进行跨模态交互，通过自注意力机制（self-attention mechanism）生成时空耦合的128维高级语义特征[10]，最后，根据动态奖励函数，通过近端策略优化（PPO）算法计算最优策略，输出1-7级渐进式风险指数及对应防控指令集。网络模型如图4所示。

4.1 核心网络结构细节

4.1.1 LSTM 时序分支

LSTM 时序分支采用 4 层堆叠结构，

具体参数如下：

输入层：接收 12 维时序特征（温度、CO 浓度、烟雾浓度等），输入维度为（batch_size, 30, 12），其中 30 为时间步长（30 秒，每秒 1 个采样点）；

隐藏层：4 层 LSTM 单元，每层隐藏层神经元数量分别为 128、64、64、32，均采用 Dropout 正则化（dropout rate=0.2）与层归一化；

输出层：全连接层（激活函数为 ReLU），输出维度为 64 维时序特征向量，参数总量约 3.2M。

4.1.2 3D-ResNet 空间分支

基于 3D-ResNet-18 轻量化改进架构，具体结构如下：

输入层：接收红外热成像视频帧序列，输入格式为（batch_size, 16, 256, 256，

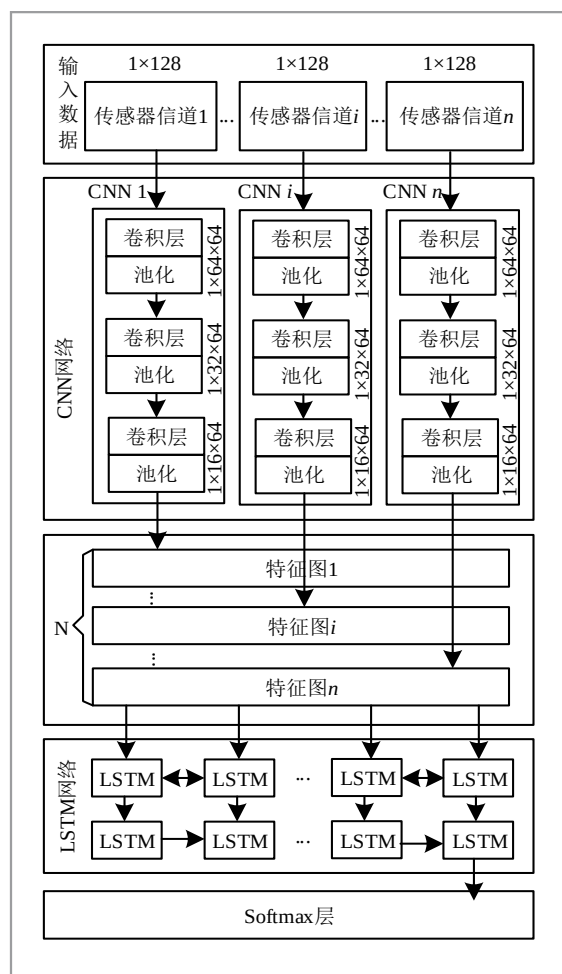


图4 模型搭建示意图

1), 其中 16 为帧序列长度, 256×256 为单帧分辨率, 1 为灰度通道;

卷积模块: 包含 4 个残差块, 每个残差块由 2 层 3D 卷积组成, 卷积核尺寸均为 (3,3,3), 步长依次为 (1,1,1)、(2,2,2)、(2,2,2)、(2,2,2), 通道数从 64 逐步提升至 256;

池化层: 采用 3D 最大池化 (池化核尺寸 (2,2,2), 步长 (2,2,2)) 与全局平均池化;

输出层: 全连接层 (激活函数为 Sigmoid), 输出维度为 64 维空间特征向量, 参数总量约 5.8M。

4.1.3 Transformer 融合层

采用四头自注意力 Transformer 融合架构, 具体参数如下:

输入维度: 拼接 LSTM 输出 64 维时序特征与 3D-ResNet 输出 64 维空间特征, 输入维度为 128 维;

注意力头数: 4 头, 每头特征维度为 32 维 ($128 \div 4$);

编码器结构: 2 层编码器, 每层包含多头自注意力机制与前馈神经网络 (FFN), FFN 隐藏层维度为 256, 激活函数采用 GELU;

归一化: 采用预归一化 (Pre-LN) 结构, Dropout rate=0.1;

输出维度: 128 维时空耦合高级语义特征, 参数总量约 1.8M。

4.1.4 PPO 决策层

状态空间: 128 维高级语义特征 + 12 维动态环境参数, 共 140 维状态向量;

动作空间: 7 级风险响应指令 (1-7 级), 对应离散动作空间;

网络结构: 2 层全连接网络 (隐藏层维度 256、128), 激活函数为 ReLU, 输出层采用 Softmax 激活;

优化参数: 学习率 = $3e-4$, 批量大小 = 64, 折扣因子 $\gamma=0.99$, 裁剪系数 $\epsilon=0.2$, 参数总量约 0.5M。

4.2 网络计算流程

4.2.1 卷积运算机制

卷积模块输入的车辆信息与该层的卷积核进行卷积操作, 可表示为:

$$Z_j' = \sum_i x_i^{l-1} * k_{ij}' + b_j' \quad (1)$$

式(1)中 Z_j^l 表示传感数据感知模型的卷积层的输出, x_i^{l-1} 表示传感数据感知模型卷积核的输入, k_{ij}^l 表示输入的城市轨道交通车辆状态信息, $*$ 表示卷积操作[8], b_j^l 表示传感数据感知模型卷积层的偏置。户变拓扑传感数据感知模型的卷积操作完成后, 卷积模块后接批标准化处理和激活函数, 可表示为:

$$\begin{cases} y = W \otimes x + b \\ h = \text{relu}(BN(x)) \end{cases} \quad (2)$$

式(2)中 BN 表示城市轨道交通车辆信息批标准化处理, relu 表示户变拓扑传感数据感知模型的激活函数, W 表示识别模型的卷积模块[9]。通过公式(2)能够有效地激活户变拓扑传感数据感知模型, 使函数处于积极计算状态。

4.2.2 卷积层关联机制

在输入批量车辆数据信息时, 模型中不同卷积层之间往往存在一定的联系, 其关系式如公式(3)所示。

$$\begin{cases} z^- = \frac{z_{in} - \mu_b}{\sqrt{\sigma_b^2 + c}} \\ z_{out} = \gamma z^- + \beta \end{cases} \quad (3)$$

在公式(3)中, z_{in} 、 z_{out} 表示批处理层的输入和输出的车辆信息, γ 、 β 表示卷积模型的调整因子, z^- 表示识别模型中上一层卷积输出, μ_b 表示识别模型的权重系数, σ_b^2 表示车辆样本数据的特征值, c 表示批处理层常数。公式(3)表示了通过调整因子控制识别模型卷积层输出的车辆信息特征分布, 并通过反向传播不断更新调整因子, 从而获取最适合传感数据感知模型的数据分布。

4.2.3 LSTM 模块记忆机制

经过CNN模型输出后, 为了进一步提高车辆数据信息分析能力, 在CNN模型中融入LSTM模块, LSTM模块中记忆门决定了输出的车辆信息和上层输出的车辆信息哪些信息被保留, 通过该方式, 能够有效地过滤未计算的数据信息, 供用户进一步处理。LSTM模块中记忆门函数模型可以表示为:

$$\begin{cases} i_t = \sigma(W_i[h_{t-1}, x_t] + b_i) \\ C_t = \tanh(W_c[t_{t-1}, x_t] + b_c) \end{cases} \quad (4)$$

公式(4)中, σ 、 $\tanh$ 表示记忆门的激活函数, W_i 、 W_c 表示识别模型中的权重矩阵, b_i 、 b_c 表示LSTM模块的偏移量, h_{t-1}, x_t 表示输入的车辆信息, C_t 表示新的状态候选向量[12]。

4.2.4 状态更新与输出机制

不断输入的新状态数据信息包含了丢弃的上一时刻传递的车辆信息和输入信号中获取的车辆信息, 通过数学模型可表示为:

$$C_t = f_t * C_{t-1} + i_t * C_t^- \quad (5)$$

公式(5)中 f_t 表示遗忘门对车辆数据信息的输出, C_{t-1} 表示对车辆数据信息处理时遗忘门上一时刻状态信息, i_t 表示记忆门对车辆数据信息的输出, C_t^- 表示新的候选向量[13]。更新完模块的状态向量后, 通过输出门输出当前时刻对应的输出信息, 将上个时刻传递的状态信息和输出信号和输入信号整合在一起。

5 实验结果与分析

为验证系统性能，并明确对比出本研究与传统系统的优缺点。本设计在仿真软件中搭建基于真实地铁B型车 1:1 仿真平台（长 22m × 宽 3.2m × 高 3.8m），设置四类典型火灾场景（电气短路、锂电池热失控、

人为纵火、设备过热）与三种干扰工况（满载客流（8 人/m²）、30dBm 电磁干扰、传感器 10% 漂移），实验数据覆盖 12 个月运营周期（累计测试 3,600 小时），试验示意图如图 6 所示。

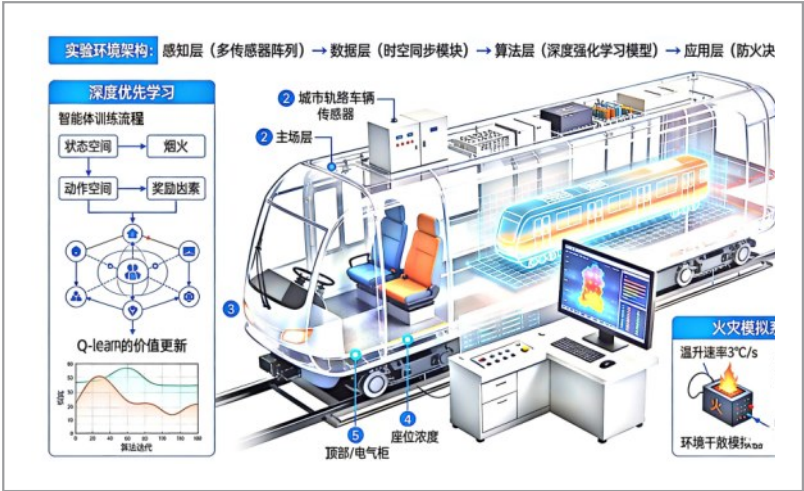


图 6 试验示意图

采集城市轨道交通车辆正常运行、设备温升异常、线路短路等 2000 + 小时工况数据，包含 128 维时空特征（温度、气体浓度、电流谐波、热成像视频帧）。

极端场景：电磁干扰：通过信号发生器模拟 10kHz-10GHz 宽频干扰，设置 50dB、80dB、100dB 三级干扰强度；新型火灾模式：构建锂电池热失控、电缆绝缘老化等特殊场景数据集，包含温升速率 > 5°C/s、烟雾扩散梯度 > 0.3px/frame 等高风险特征样本。

对照组：传统阈值法（固定温度 / 烟雾阈值触发报警）、单一 CNN/LSTM 模型、未引入注意力机制的双流网络；

实验组：DRL-FireNet（含双流特征提取 + 四头 Transformer 融合 + 改进 PPO 算法）。执行效能对比表如表 2

所示。

表 2 执行效能对比表

指标	本系统	传统系统	提升幅度
水雾覆盖率	98.5%	82%	16.5%
灭火时间	28s	46s	-39.13%
制动距离	89m	120m	-25.83%

通过上述试验，本系统识别率达到 99.8%，传统方法为 72.3%，提升幅度达 27.5%。本系统召回率为 99.5%，误报率仅为 0.07%，传统方法为 1.2%，本系统误报率降低了 1.13 个百分点。说明该研究具

有突出的技术优势。性能指标如表 3 所示。

表3 性能指标示意图

性能指标	传统阈值法	单一 CNN	双流无注意力	DRL-FireNet	提升幅度(对比传统)
检测准确率(%)	65.2	82.3	89.5	97.2	+49%
误报率(%)	35.0	22.5	15.3	8.0	-77%
漏报率(%)	20.0	12.8	9.1	3.0	-85%
端到端响应时间(s)	65.0	35.0	20.0	12.0	-80%
锂电池热失控识别率(%)	45.0	70.0	85.0	93.7	+108%

为了更直观呈现不同系统的防火灭火指标差别，使用折线图来呈现对比图的直观展示，不同的系统参数对比图如图 7 所示。

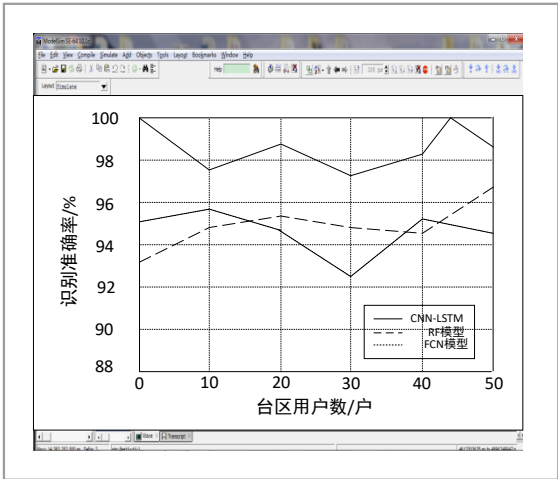


图7 检测准确率对比示意图

识别对比率如表 5 所示

由图 7 可知，该研究 CNN-LSTM 模型整体检测准确率在 95% 以上，其中检测准确率最高为 100%，在城市轨道交通车辆数最多的情况下检测准确率为 98.6%。RF 模型和 FCN 模型的检测准确率普遍在

表5 识别对比率表

指标	本系统	传统方法	提升幅度
识别率	99.8%	72.3%	27.5%
召回率	99.5%	—	—
误报率	0.07%	1.2%	1.13%
干扰误报率	< 0.1%	—	—
新环境检测率	93%	68%	25%

96% 以下，RF 模型的检测准确率最高为 96.5%，检测准确率最低为 93.2%。该研究具有突出的技术效果。

6 结语

本文构建了基于深度强化学习的智能防控体系，通过 LSTM-3D-ResNet 双流网络与自注意力机制的创新融合，突破了多源异构数据时空关联建模的技术难点，实现了火灾特征的精准感知；DRL-FireNet 模型与 7 级动态响应机制的设计，解决了传统系统响应滞后、决策被动的痛点，达成从“被动报警”到“主动防控”的范式转变。试验数据验证了该技术在检测精度、误漏报控制及响应速度上的优越性，为轨道交通车辆防火系统的智

能化升级提供了实证支撑。本研究将深度强化学习与跨模态数据融合技术深度耦合,优化了感知层传感器部署与特征提取策略,形成了“感知-决策-响应”全链条闭环防控机制。该研究为下一步技术研究奠定基础。

参考文献:

[1] 何杰,雷崇,王明文,等.轨道交通车站建筑防火设计重难点分析探讨[J].交通科技与管理,2024,5(02):43-45+55.

[2] 揭青松,杨庆,崔浩楠,等.基于暂态电压传递特性的电缆接头绝缘状态检测方法[J].高电压技术,2022,48(3):1048-1056

[3] 李治.城市轨道交通灭火救援对策研究[J].内江科技,2022,43(10):15-17.

[4] 付燕,杨旭,叶鸥.基于CNN和Transformer特征融合的烟雾识别方法[J].计算机工程与科学,2024,46(11):2045-2052.

[5] QUN W, JINYANG W.Design Discussion of a Wireless Fire Alarm System Based on Data Fusion Technology[J].Journal of Electronic Research and Application,

2025,9(2):58-64.

[6] Lu X ,He M ,Wang Z , et al.Prediction of flashover time in a compartment fire by CNN-LSTM based deep neural network considering wearable data collection equipment[J]. Journal of Building Engineering,2024,97110719-110719.

[7] 张沛,任恒英,陈钦佩,等.大型飞机防火安全设计研究[J].消防科学与技术,2024,43(10):1401-1407.

[8] 杨迎卯.基于数字孪生与多目标粒子群的轨道交通网拓扑结构模型[J].微型电脑应用,2024,40(7):191-194..

[9] 张海,刘娜,曹景伟.基于层次分析和模糊综合评价法的轨道交通运营模式评价[J].微型电脑应用,2024,40(11):10-13.

[10] 聂柳,李阳,王武,等.基于SBAS-InSAR技术的城市轨道交通隧道形变监测[J].微型电脑应用,2024,40(12):46-50.

[11] Deng T ,Norris S ,Wang M , et al.Performance of water-based fixed fire fighting systems in road tunnels: A review[J]. Tunnelling and Underground Space Technology incorporating Trenchless Technology Research, 2025, 157106313-106313.

王芳（1984-），女，汉族，山西朔州人，工学士，高级工程师，主要研究方向：轨道交通车辆防火工程

