

基于深度强化学习的新能源汽车动力电池组 降温自适应控制系统设计

张丽芝

(西安思源学院 工学院, 西安 710000)

摘要: 在新能源汽车高负荷运行场景下, 动力电池组面临复杂电热耦合环境, 其内部电化学反应剧烈、热积累效应显著, 易引发温场能效失衡问题, 难以实现热状态的精准感知与实时控制; 为了解决该问题, 设计了基于深度强化学习的动力电池组降温自适应控制系统; 以数据驱动决策、闭环保障精准为核心, 构建了3层协同的全流程闭环总体架构, 围绕深度强化学习算法决策中枢, 联动数据采集、信息交互、执行控制3大模块, 实现动力电池热特性与冷却控制的自适应匹配; 硬件上, 采用“测温+测流”双组件协同设计的热特性多维度采集模块, 确保数据精准量化; 通过CAN总线实现整车工况通信模块的实时数据交互; 自适应冷却指令执行模块则将控制指令转化为实际冷却动作, 形成闭环控制; 在软件上, 通过计算动力电池温度偏差变化率, 量化产热、温度偏差与趋势递进关系, 结合深度强化学习算法, 构建“状态输入—行为输出—奖励函数—策略更新”闭环控制框架, 实现温场能效平衡, 完成动力电池热特性的自适应降温; 在高温爬坡极端工况下进行系统测试, 结果表明, 该方法在温度分布均匀性(温差 $\leq 1^{\circ}\text{C}$)与响应延迟($\leq 5\text{ ms}$)方面均显著优于现有主流控制方法, 展现出良好的控制性能。

关键词: 深度强化学习; 新能源汽车; 动力电池组; 降温控制; 极端工况

Design of Adaptive Control System for New Energy Vehicle Power Battery Pack Cooling Based on Deep Reinforcement Learning

ZHANG Lizhi

(School of Engineering, Xi'an Siyuan University, Xi'an 710000, China)

Abstract: In high-load operating scenarios, the power battery pack of new energy vehicles faces a complex electric thermal coupling environment, where intense internal electrochemical reactions and significant thermal accumulation effects can easily lead to temperature field energy efficiency imbalance, making it difficult to achieve accurate perception and real-time control of thermal states. To address this issue, a deep reinforcement learning based adaptive control system for cooling power battery packs was designed. With data-driven decision-making and precise closed-loop as the core, a three-layer collaborative full process closed-loop overall architecture has been constructed. With the deep reinforcement learning algorithm as the decision-making center, the data acquisition, information exchange, and execution control are integrated to adaptively match the thermal characteristics of power batteries and cooling control. At the system hardware, through the collaboration of temperature measurement and flow measurement components, a multi-dimensional acquisition module for thermal characteristics is designed to accurately quantify the data; Through the CAN bus, the real time data exchange of the vehicle operating condition communication module is achieved; The adaptive cooling command execution module converts the control command into the actual cooling action, forming a closed-loop control. At the system software, by calculating the temperature deviation change rate of the power battery, quantifying the relationships between the heat generation, temperature deviation, and trend progression, and combining the deep reinforcement learning algorithm, a closed-loop control framework of “state input-behavior output-reward function-strategy update” is constructed to achieve an energy efficiency balance in the

收稿日期:2025-11-26; 修回日期:2026-01-07。

基金项目:西安思源学院校长基金科研项目(XASYB24ZD09)。

作者简介:张丽芝(1978-),女,大学本科,副教授。

引用格式:张丽芝. 基于深度强化学习的新能源汽车动力电池组降温自适应控制系统设计[J]. 计算机测量与控制, 2026, 34(8): 130-138.

temperature field, achieving adaptive cooling for thermal characteristics of power battery. System testing is conducted on the extreme conditions of high temperature climbing, and the results show that this method significantly outperforms existing mainstream control methods in terms of temperature distribution uniformity (temperature difference $\leq 1^{\circ}\text{C}$) and response delay ($\leq 5\text{ ms}$), demonstrating a good control performance.

Keywords: deep reinforcement learning; new energy vehicles; power battery pack; cooling control; extreme operating conditions

0 引言

在新能源汽车的运行过程中,高负荷、极端工况十分常见,例如持续高速爬坡、在高温环境下长时间行驶等。在这些工况下,动力电池组内部的化学反应极为剧烈,容易快速积聚大量热能。倘若这些热能无法及时散发出去,就会破坏电池的电化学环境,导致电池内阻增大、充放电效率降低,进而影响电池的使用性能,具体表现为车辆动力下降、续航里程缩短。此外,这还会加速电池的老化进程,缩短其使用寿命,增加用户的使用成本和维修频率^[1]。为解决动力电池组的散热难题,部分研究者提出了具有高应变响应特性的电池冷却控制方法。该方法以双流冷却管路布置为核心,深入探究各因素对散热特性的影响,构建精准的散热模型,从而实现根据实际情况自适应调整冷却参数、精准降低电池温度。而且,在车辆急加速或制动时,电池的产热状态会发生急剧变化,传统正交试验获取数据的能力有限,难以有效控制电池温度,这会进一步加剧电池性能的衰减^[2]。因此,开展新能源汽车动力电池组降温自适应控制研究十分必要。

在当前的新能源汽车动力电池组降温控制研究中,文献[3]提出了基于分类回放双延迟贝叶斯深度的控制方法,考虑电池组非线性、不确定性及环境条件干扰等特性,设计融入贝叶斯神经网络、分类经验回放等技术,能依据电池实时状态与环境变化,自适应调整降温策略,精准调控冷却系统,有效提升电池组降温控制性能,保障其安全稳定运行。该方法虽能自适应调控,但算法复杂度高,计算量大,对硬件要求苛刻,可能增加系统成本与能耗。且面对极端复杂工况,算法稳定性与可靠性有待进一步验证;文献[4]提出了基于多频电化学阻抗谱的控制方法,利用电池表面与环境温度初步估计内部温度并选定频率,基于电化学阻抗谱对内部温度二次估计,以此精准掌握电池热状态。通过研究内部温度、荷电状态(SOC, state of charge)对电池阻抗的影响,在忽略SOC影响时获取内部温度与阻抗实部关系,进而实现自适应降温控制,提升电池组散热效率与安全性。该方法虽精度高、范围广,但多频测量过程复杂,对测量设备精度要求严苛,成本较高。且忽略SOC影响,在复杂工况下,温度估计的准确性可能受

影响,进而影响降温控制效果;文献[5]提出了基于变论域的控制方法,借助Matlab/Simulink工具包搭建动力电池组降温变论域模糊控制器,运用面向生产过程的通信(OPC, OLE for process control)通信技术,将西门子S7-200PLC与Simulink相连,达成实时数据交互。通过建立半实物仿真系统,能依据电池实时温度、环境条件等动态变化,在线整定控制参数,使冷却系统自适应调整工作状态,精准控制电池组温度,保障其安全高效运行。该方法构建半实物仿真系统成本较高,且对通信稳定性要求严苛,一旦出现通信故障,可能导致数据传输中断,影响在线参数整定,进而使降温控制失效,危及电池组安全;文献[6]提出了基于APSO-BP-PID控制方法,采用基于自适应粒子群优化(APSO, adaptive particle swarm optimization)算法优化的反向传播(BP, back-propagation neural network)神经网络与比例-积分-微分(PID)控制相结合的策略,实时调控电池组工作温度。通过系统性研究,获得电池组在不同环境和降温状态下的温度响应和内部平均温度的动态变化,并筛选出对温度变化具有较高敏感性的关键温度响应特征参量。在此基础上,利用基于APSO算法的BP神经网络,对PID控制参数进行动态调整,从而达到对冷却系统参数进行精确、实时调整的目的。该方法严重依赖于APSO-BP-BP网络的学习精度,如果模型训练数据有偏差或者算法没有得到充分的优化,将会造成参数整定不准,从而降低控温的准确性。

为了解决复杂工况下易出现热量累积的问题,设计了一种基于深度强化学习的动力电池组降温自适应控制系统。该系统通过热特性多维度采集模块、整车工况通信模块与自适应冷却执行模块的协同工作,结合深度强化学习算法实现冷却策略的实时动态优化。

1 动力电池降温三元闭环硬件控制架构设计

架构以数据驱动决策、闭环保障精准为核心逻辑,构建3层协同的全流程闭环的总体架构,围绕深度强化学习算法决策中枢,联动数据采集、信息交互、执行控制3大模块,实现动力电池热特性与冷却控制的自适应匹配。

由图1可知,通过测温、测量双模块协作获取热特性数据,实时测量单体电池的温度和充放电电流,经干

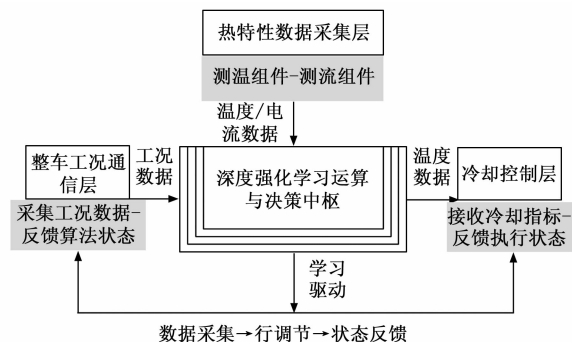


图 1 基于深度强化学习的降温自适应控制总体架构

扰处理与信号校正后，得到准确的热特性参数。借助统一总线协议与整车联网进行工况信息交互，采集车辆运行速度、电耗等运行状况信息，并将算法运行状态反馈给整车控制器，完成“热力特征—运行状态”双维输入，实现双向通信。在此基础上，分析电机速度和驱动保护情况，将所得散热命令转化为对冷却液流动的调控作用，并实时反馈其运行状态至决策中心，形成“命令—执行—反馈”的闭环。3 个部分的功能模块通过数据交互和状态反馈深度联动，构建自适应控制闭环，确保复杂工况下动力电池组的温度稳定。

在硬件控制系统架构设计中，本系统围绕感知精准、交互实时、执行可靠的核心原则，构建了 3 层闭环协同体系。其技术深度主要体现在以下方面：

在感知层，热特性多维度采集模块通过 ZAM6228 温度采集芯片与 ACS773 霍尔电流传感器协同工作，实现了高精度数据同步采集。ZAM6228 基于 PT100 传感器阵列和 24 位 ADC，可实现 $\pm 0.1\text{ }^{\circ}\text{C}$ 的测温精度；ACS773 则提供 $\pm 1\%$ 的全量程电流测量精度。二者通过 SPI 接口以 1 kHz 的同步采样率工作，确保了温度与电流信号的时空对齐。在交互层，整车工况通信模块基于 CAN FD 协议构建了高速通信通道。采用 TJA1043 收发器配合 SJA1000 协议控制器，实现 5 Mbps 的高速数据交换。该模块通过硬件滤波机制实时筛选关键工况参数，并将电池热状态以诊断帧形式反馈至整车控制器，形成双向验证闭环。在执行层，自适应冷却执行模块采用 DRV8313 驱动芯片配合 OWP-BL93-301 无刷直流水泵，通过空间矢量 PWM 算法实现 0.1 L/min 分辨率的流量精确控制。模块内置三重保护电路，可在 100 μs 内响应异常情况，确保系统安全运行。

该硬件架构通过专业级芯片选型与精心设计的同步机制，为深度强化学习算法提供了高质量的数据基础与可靠的控制执行通道。

1.1 强化学习驱动的热特性多维度采集模块

多维度热特性采集模块（ZAM6228+ACS773）是

基于深度强化学习的动力电池降温自适应控制系统的核心热特性数据源，负责实时采集单体电池温度与充放电工作电流，为单体生热功率计算、模组平均温度计算提供精准原始数据，也是实现动力电池热特性递进表征、支撑强化学习算法状态输入特征集构建的基础单元。

模块采用“测温+测流”双组件协同设计，二者通过串行外设接口（SPI，serial peripheral interface）接口与系统微处理器联动。其中，ZAM6228 以 8 通道 PT100 测温模块为核心，配套工业级 PT100 铂电阻传感器阵列、仪表放大器及高精度模数转换器（ADC，analog-to-digital converter），基于铂电阻正温度系数特性，通过恒流源激励将电阻变化转为电压信号，经仪表放大器调理、滤波后由 ADC 转换为数字信号，再经模块内部固件完成异常值剔除与校准，最终输出单体实时温度^[7]；同时采用“模拟地—数字地”分离布线、隔离电源配置等抗干扰设计，安装时严格避开工业动力线缆，且传感器线缆选用屏蔽双绞线，进一步降低电磁辐射干扰^[8]。ACS773 作为汽车级霍尔电流传感器，配套信号处理单元与数据转换模块，基于霍尔效应，当电流通过传感器的载流导体时产生磁场，霍尔元件将磁场转换为霍尔电压，经内部放大器放大、温度补偿后输出模拟信号，再由外部数据转换模块转换为数字信号传输至微处理器，为单体生热功率计算中的欧姆热量变化提供关键电流数据；其采用高防护等级封装，能适应车载潮湿、振动环境，同时内置过流保护电路，当电流超过安全范围时输出报警信号，保护后续电路安全。热特性多维度采集结构如图 2 所示。

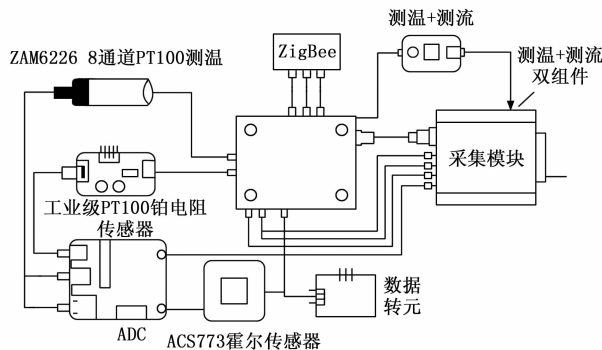


图 2 热特性多维度采集结构框图

两组件协同实现对“温度—电流”双维度热特性数据的精准量化，模块的抗干扰设计与信号处理机制确保热特性数据信噪比满足要求，避免车载电磁干扰对数据精度的影响；采集数据经 SPI 接口传输至强化学习运算模块后，与整车工况数据融合构建强化学习算法的状态输入特征集，为算法提供“热特性—工况”双维度输入。

1.2 CAN 闭环驱动的整车工况通信模块

整车工况通信模块（TJA1043）是以深度强化学习为核心的动力电池降温自适应控制系统中的工况信息交互核心单元。它负责对整车行驶速度、车载用电设备功率等工况数据进行实时采集，并将自身运行状态反馈给整车控制器，是连接整车系统与强化学习算法、支撑“热特性—工况”二维状态输入特征集的关键单元^[9]。该模块通过控制器局域网（CAN，controller area network）总线连接整车 CAN 网络，并与整车控制器（VCU，vehicle control unit）以及车载用电设备控制器（如空调、音响控制器等）进行通信连接。由 SJA1000 协议控制器依据 CANFD 数据帧的完整结构（包括仲裁区、控制区、数据区、循环冗余校验（CRC，cyclic redundancy check）区、确认（ACK，acknowledgement）区和帧终区）对车辆运行状态数据进行解析，精确提取出车辆行驶速度、车载用电设备功率等关键信息。在此基础上，将分析得到的工况数据转化为可识别的二进制数字格式（由 NVIDIA Jetson Orin NX 处理），并通过隔离电路将其传送到操作模块的 CAN 接口中，以保证数据的实时传输，进而实现“整车工况采集—算法状态反馈”双向闭环控制。其控制结构如图 3 所示。

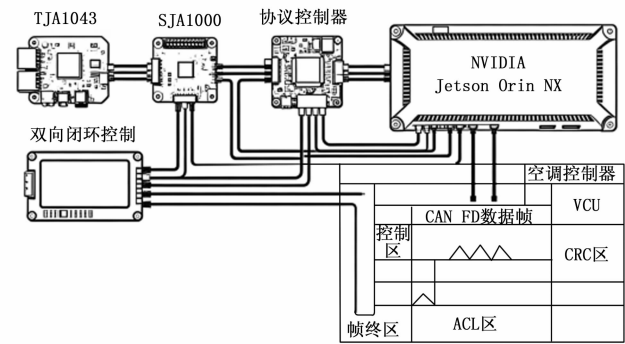


图 3 整车工况 CAN 闭环驱动通信控制框图

1.3 自适应冷却指令执行与闭环控制模块

自适应冷却执行模块（OWP-BL93-301）作为系统的控制执行终端，功能是将强化学习运算模块输出的冷却控制指令（如冷却液流速增量）转化为实际冷却动作。该模块以“脉冲宽度调制（PWM，pulse width modulation）控制无刷直流（BLDC，brushless direct current）电机与车规级驱动电路”为核心设计架构，核心硬件型号为车载 PWM 控制 BLDC 电子水泵 OWP-BL93-301 与驱动芯片 DRV8313，同时配套无刷直流电机本体、位置传感器（霍尔传感器 AH3144）、PWM 接收电路、过流/过热保护电路及电源接口、冷却液接口，如图 4 所示。

由图 4 可知，各组件协同配合，实现“指令接收—

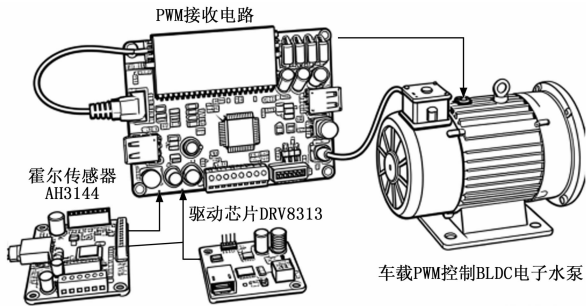


图 4 自适应冷却指令执行与闭环控制模块组成架构

电机驱动—流量调节—状态反馈”的全流程闭环控制。其中，OWP-BL93-301 的 BLDC 电机本体可兼容车载电源，能通过 PWM 占空比实现转速线性调节，进而改变冷却液流量，以此满足不同热特性场景下的冷却需求；DRV8313 作为车规级三相半桥驱动芯片，支持 PWM 驱动模式，且内置过流、过热、欠压保护功能，可保障电机安全稳定运行^[10]；AH3144 霍尔传感器安装在电机转子附近，能输出三路互差 120° 的位置信号，为电机换相提供精准依据^[11]；PWM 接收电路由运算放大器搭建的信号调理电路组成，主要负责接收来自强化学习运算模块的 PWM 占空比指令，并对指令进行滤波处理，确保指令精准传输至驱动芯片。

2 动力电池热特性递进的深度强化学习自适应降温控制

新能源汽车动力电池在高速加速、紧急制动、大功率电器运行等复杂工况下，电池内部电化学反应、欧姆损耗和极化效应会产生大量热量，若不能及时调控，将导致电池性能下降、寿命缩短，甚至出现热失控。传统的动力电池降温控制方法存在两个核心缺陷：1）依靠单一控制参数的温度偏差，无法预知温度的动态变化趋势，在工况突变时容易出现控制滞后现象；2）依赖于准确的热管理系统数学模型，但复杂工况下动力电池热特性参数时变易造成模型失配，导致控制精度下降。针对以上问题，开展基于深度强化学习的动力电池组自适应降温控制方法研究。

在本文设计的系统中，深度强化学习算法输出的动作与控制执行之间建立了精确的量化映射关系，构成自适应降温的直接驱动核心。具体而言，策略网络输出的标准化动作值并非抽象指令，而是对应冷却系统的具体物理调节量。该值通过线性映射关系转换为实际的冷却液流速增量，其中标准化动作值由 OWP-BL93-301 无刷直流水泵的性能曲线确定。流速增量通过预标定的 PWM 占空比—流量特性曲线，转化为对 DRV8313 驱动芯片的具体 PWM 控制信号，最终实现冷却液流量的精确调节。这种从“策略输出一物理量转换—执行器驱

动”的完整链条，确保了智能体决策能够实时、准确地转化为物理系统的调控行为。更重要的是，动作输出的连续性与平滑性设计，避免了传统开关控制或阶跃调节带来的机械冲击与温度波动，使冷却功率的变化与电池产热率的动态特性实现渐进式匹配，这正是系统能在高温爬坡工况下实现温差 $\leq 1^\circ\text{C}$ 精细控制的关键物理实现机制。深度强化学习架构如图 5 所示。

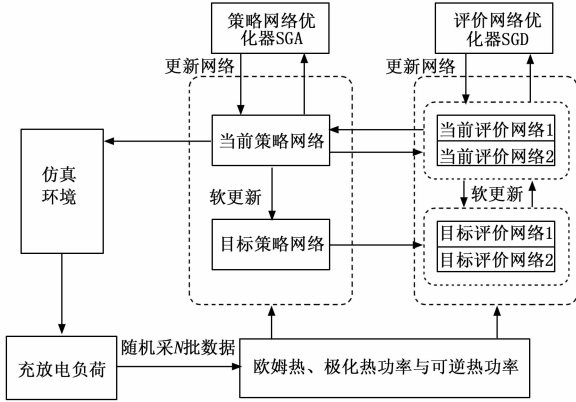


图 5 深度强化学习架构图

在本文研究中，DDPG 算法完美匹配了动力电池热管理的连续控制需求。与输出离散动作的算法不同，DDPG 能够直接输出冷却液流速调节量这样的连续值动作，这对应了实际冷却系统中无级调节的执行机构特性，避免了传统离散控制带来的阶跃响应和机械冲击，为实现温差 $\leq 1^\circ\text{C}$ 的精细控制提供了算法基础。且 DDPG 将深度学习强大的感知能力与强化学习的序列决策能力相结合，使其能够从高维的、非线性的传感器数据中自动提取特征，并学习复杂的冷却策略。规避了传统方法依赖精确物理建模的难题，实现了从状态感知到最优动作的端到端映射，显著提升了热管理系统的自适应能力。

2.1 动力电池温度偏差变化率计算

动力电池热特性表征的核心目标是量化产热、温度偏差与趋势递进关系，为深度强化学习提供物理意义明确的输入参量。单体电池生热功率需综合欧姆热、极化热功率与可逆热功率，计算公式如下：

$$P_a = I_a^2 \cdot R_a \cdot (1 + L_h) + P_o + P_v \quad (1)$$

式中， I_a 为电池工作电流，由整车控制器实时采集，反映电池的充放电负荷^[12]； R_a 为电池基准内阻，表征电池的固有电阻特性^[13]； L_h 为欧姆热损比； P_o 为极化热功率，由电极电化学极化与浓度差极化产生，其值与电流变化率正相关^[14]； P_v 为可逆热功率，由电化学反应熵变导致，放电时若反应放热则 $P_v > 0$ ，若吸热则 $P_v < 0$ 。通过 $R_a \cdot (1 + L_h)$ 建立热损量化与生热功率的关联，使 P_a 能实时响应工况突变引发的热损波动，为模组温度

计算提供精准的产热数据源。

电池模组由多个单体串联/并联组成，其平均温度需综合各单体生热功率与模组散热特性，计算公式如下：

$$T_v = \frac{\sum_{i=1}^N \left(\frac{P_{a,i} \cdot \epsilon}{c_a \cdot m_a} + T_{0,i} \cdot S_i \right)}{\sum_{i=1}^N S_i} \quad (2)$$

式中， N 为模组内单体电池数量； ϵ 为温度采样时间步长，确保与整车控制周期同步； c_a 为单体电池比热容，由电池材料特性决定^[15]； m_a 为单体电池质量； $T_{0,i}$ 为第 i 个单体的初始温度^[16]； S_i 为第 i 个单体的散热面积，由单体几何尺寸与模组装配方式决定^[17]。通过生热、散热平衡关系，将单体层面的 $P_{a,i}$ 转化为模组层面的平均温度 T_v ，解决了单一单体温度无法表征模组整体热状态的问题，为后续温度偏差计算奠定基础。

温度偏差变化率是预判温度动态趋势的关键参量，计算公式如下：

$$e_p = \frac{(T_v - T_u)_t - (T_v - T_u)_{t-1}}{\epsilon} \quad (3)$$

式中， T_u 为电池最优工作温度； $(T_v - T_u)_t$ 、 $(T_v - T_u)_{t-1}$ 分别为当前时刻的温度偏差、前一采样时刻的温度偏差^[18]。该公式的核心作用是解决仅靠 $T_v - T_u$ 无法应对的工况突变问题：当 $T_v - T_u > 0$ 且 $e_p > 0$ 时，温度偏差持续增大，需激进调节以抑制温度飙升；当 $T_v - T_u > 0$ 但 $e_p < 0$ 时，温度偏差开始减小，需温和调节以避免过度降温；当 $e_p \approx 0$ 时，温度趋于稳定，可维持当前冷却强度。该参量直接决定了降温控制的执行强度，是实现自适应的核心依据。

2.2 深度强化学习闭环驱动动力电池热特性自适应降温

深度强化学习是在不依赖精确系统模型的前提下，通过主体和环境的交互来学习最优控制策略。结合动力电池热特征参数，构建“状态输入—行为输出—奖励函数—策略更新”闭环控制框架，实现温场能效平衡，完成动力电池的低温自适应控制。结合动力电池温度偏差变化率，形成状态特征集，表达式如下：

$$Z = [L_h, P_a, e_p, v, P_{\max}] \quad (4)$$

式中， v 为车辆行驶速度^[19]； $P_{\max}$ 为车载用电设备的最大总功率^[20-21]。将分散的热特性参量与工况信息转化为智能体可识别的标准化特征向量，确保输入信息的完整性与一致性。

动作输出为冷却系统的调节量，需与热特性参量的动态变化适配。以冷却液流速增量为例，基于深度强化学习的策略网络构建动作输出映射，表达式如下：

$$\Delta v_c = \tanh[\omega_2 \cdot \sigma(\omega_1 \cdot Z + b_2) + b_1] \cdot \Delta v_{c,\max} \quad (5)$$

式中， ω_1 、 ω_2 分别为策略网络隐藏层与输出层的权重矩阵； b_1 、 b_2 分别为对应层的偏置向量； $\sigma(\cdot)$ 为 ReLU 激活函数； $\tanh[\cdot]$ 为输出激活函数，将结果映射至 $[-1, 1]$ 区间； $\Delta v_{c, \max}$ 为冷却液流速的最大调节增量，由冷却系统硬件性能决定。当 e_p 较大时， Δv_c 趋近于 $\Delta v_{c, \max}$ （增强降温）；当 e_p 较小时， Δv_c 趋近于 0（维持或减弱降温）。该映射实现了状态、动作的自适应匹配，无需人工预设控制规则。

奖励函数是引导智能体学习最优策略的核心，需平衡温度控制精度与能耗经济性。设计多目标奖励函数，表达式如下：

$$\mathcal{J} = -\gamma_1 \cdot e_p^2 - \gamma_2 \cdot \dot{e}_p^2 - \gamma_3 \cdot (\Delta v_c)^2 \tag{6}$$

式中， γ_1 、 γ_2 、 γ_3 分别为温度偏差、偏差变化率、冷却能耗的权重系数； $-\gamma_1 \cdot e_p^2$ 为温度偏差惩罚项， e_p 越大，惩罚越强，引导智能体减小温度偏离； $-\gamma_2 \cdot \dot{e}_p^2$ 为偏差变化率惩罚项， $\dot{e}_p$ 越大，惩罚越强，引导智能体抑制温度剧烈波动； $-\gamma_3 \cdot (\Delta v_c)^2$ 为冷却能耗惩罚项， Δv_c 越大，惩罚越强，避免过度降温导致的能耗浪费。智能体通过最大化累积奖励，自动学习温度稳定、能耗经济的平衡策略，既保证温度控制精度，又避免能耗消耗。

采用深度确定性策略梯度算法实现策略更新，该算法适用于连续动作空间，如冷却液流速增量。策略网络的更新通过最小化策略损失实现，损失函数表达式如下：

$$Loss = -E_{Z \sim U} \{ F[Z, \psi(Z)] \} \tag{7}$$

式中， E 表示状态—动作—奖励—下一状态交互样本的期望运算； U 为经验回放池； F 为评价网络，用于估计状态—动作对的价值； ψ 为策略网络。智能体通过与电池热环境交互，将样本存入 U ；从 U 中采样样本，更新 F 网络以准确估计价值；再基于 F 网络的价值反馈，更新 ψ 以输出更优动作。该迭代过程使策略 ψ 逐步收敛至最优策略，实现自适应控制。

3 高温爬坡极端工况下系统测试

3.1 测试环境

夏季 7 月午后 14:00，西南山区 G85 高速 6° 连续 3 km 爬坡路段，阳光直射下路面温度 42℃，环境温度 36℃，微弱南风（1.2 m/s）难以缓解热堆积。一辆续航 500 km 的纯电 SUV 以 90 km/h 行驶，驾驶员需保持 60% 油门开度维持车速，避免动力衰减。

车辆初始状态：动力电池电量 65%，温度 38℃（25~45℃ 最优区间上限），电池管理系统显示放电电流 80 A，冷却风扇中速运转（2 200 r/min），风道进风 3 m/s，产散热基本平衡。进入爬坡路段后，因对抗坡

度重力分力与风阻（风阻系数较平路提升 25%），电机输出功率从 45 kW 增至 90 kW，电池放电电流升至 120 A 并稳定。此时电池内阻 80 mΩ，欧姆热每小时产生 1 152 kJ，极化热每小时 480 kJ，总生热速率 456 W，远超冷却系统平路 320 W 的散热功率。

爬坡 5 min，电池温度升至 41℃，风扇提速至 3 500 r/min，但车头上扬使风道进风降至 1.5 m/s，实际散热功率仅 350 W；10 min 时温度达 47℃（超最优上限），电池管理系统记录异常，仪表盘显示淡黄色“电池温度监测中”；15 min 温度升至 50℃，3、4 模组局部达 53℃，散热鳍片吸入热空气致散热功率降至 330 W；20 min 平均温度达 52℃，仪表盘显示橙色“电池温度偏高”并蜂鸣，电池管理系统因未超 55℃ 保护阈值，未触动力限制，车辆仍正常爬坡，处于温度骤升状态，急需降温处理以维持车内温度稳定。

3.2 测试装置

为了确保测试的实用性和真实性，通过图 6 明确了电池在新能源车内的安装位置。

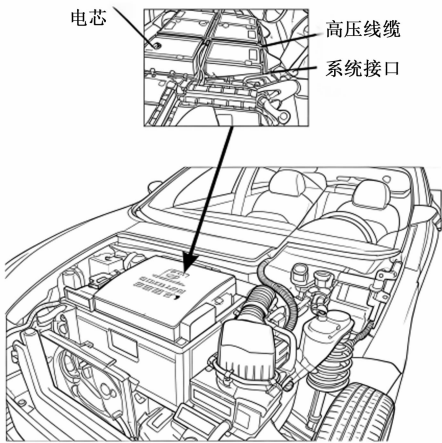


图 6 电池组现场检修架构图

通过对电池组的局部特写，可以清楚地看到电池模块、接线口等部件，硬件结构细节一目了然，方便维修时进行部件识别和接线。电池组的核心参数，如表 1 所示。

表 1 电池组的核心参数

名称	参数
单电芯容量	4.8 Ah
单电芯电压	3.7 V
4 模组总容量	57.6 Ah
排焊厚度	2 mm
散热片材质	铝合金 6061
绝缘等级	IP67

所列核心参数则为后续电池温度控制、热稳定性等

测试实验提供了明确的硬件基础与参数参考,有效确保了测试的可追溯性与准确性。动力电池热特性递进的深度强化学习自适应降温控制方法的具体实现步骤为:

- 1) 系统初始化与数据接口建立;
- 2) 热特性参量实时计算与状态构建;
- 3) DRL 算法框架部署与训练;
- 4) 自适应控制策略在线优化;
- 5) 闭环控制执行与安全监控;
- 6) 性能评估与模型迭代。

本实验深度强化学习策略基于深度确定性策略梯度算法构建,其策略网络与价值网络均采用两层 256 神经元的全连接结构,核心超参数设置为:经验回放池容量 10 万条、批量采样 128 条、策略与价值网络学习率分别为 1×10^{-4} 和 1×10^{-3} 、折扣因子 0.99 及目标网络软更新系数 0.005,并采用 Ornstein-Uhlenbeck 过程 ($\theta=0.15$, $\sigma=0.2$) 注入探索噪声,经 500 回合 WLTC 工况离线预训练后投入在线应用。为精准获取热管理关键数据,我们在 4 模组电池包内部署了 32 个 PT100 温度传感器阵列(覆盖电芯表面、流道壁面及环境),结合 ACS773 电流传感器与涡轮流量计,以 100 Hz 频率通过 ZAM6228 模块与高精度 ADC 同步采集温度、电流、流量及整车 CAN 信号,所有数据均打上微秒级时戳确保严格同步;基于此,系统实时通过焦耳定律与冷却液热焓差计算局部产热与散热功率,并采用克里金插值算法生成三维温度场云图,从中动态提取平均温度、空间最大温差(用于验证 $\leq 1^\circ\text{C}$ 的均匀性指标)及温度变化率等关键特征,构建出 5 维归一化状态向量,为深度强化学习智能体提供了高保真、高时空一致性的训练与决策数据基础。

3.3 测试指标

模组温度一致性负责验证系统冷却调节的空间均匀性,反映系统是否可避免局部过热、实现热量均匀分布,理想情况下的热量分布如图 7 所示。

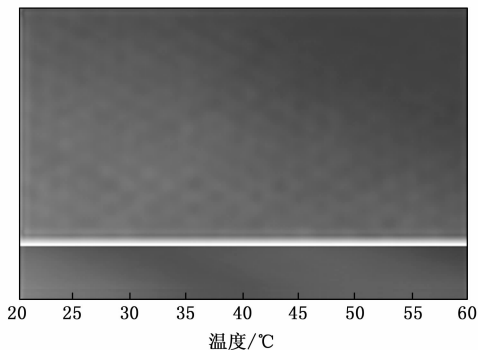


图 7 理想情况下的热量分布

由图 7 可知,温度空间分布通过颜色渐变(从深色

到浅色)呈现,结合整体视觉效果与刻度对应关系可见,模组内不同区域的温度呈极均匀的渐变趋势,这意味着任意两个区域之间的温度差 $\leq 1^\circ\text{C}$ 。这种均匀的温度分布,反映系统可有效避免局部过热,实现热量在空间上的均匀分配,保障电池模组在工作过程中各区域性能与寿命的一致性,同时降低局部高温引发的安全风险。

工况突变响应延迟负责验证系统的实时响应特性,高功率用电工况突变会使电池产热急剧变化,该指标体现决策速度与执行模块的响应效率,如图 8 所示。

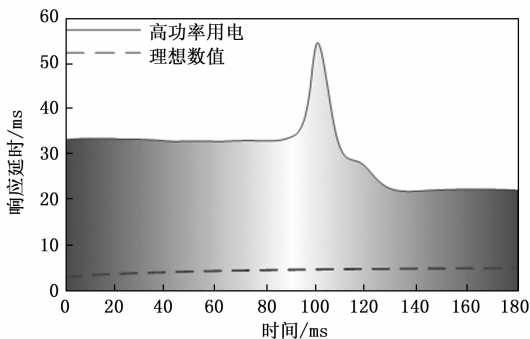


图 8 工况突变实际与理想情况下的响应延迟

由图 8 可知,实际响应延迟在时间 100 ms 时出现了峰值,高达 56 ms,其余时间段也超过了 20 ms,而理想数值控制在 5 ms 内,这说明系统在高功率用电工况突变时,决策速度与执行模块的响应效率极优,能在 ≤ 5 ms 内快速介入冷却调节,有效避免电池因产热急剧变化而出现温度失控。

3.4 测试结果与分析

对比分析基于分类回放双延迟贝叶斯深度的控制方法(文献[3]方法)、基于多频电化学阻抗谱的控制方法(文献[4]方法)、基于变论域的控制方法(文献[5]方法)、基于 APSO-BP-PID 控制方法(文献[6]方法)、基于深度强化学习(本文方法)的控制方法的热量分布情况,本文选取的 4 种对比方法均为电池热管理领域主流控制策略,其核心目标与本文一致:通过算法实时处理传感器数据并输出冷却控制指令,以维持电池温度稳定。这些方法在“输入状态—输出动作”的控制范式上具有内在一致性。图 9 对比的工况突变响应延迟指标,衡量的是从信号突变到控制指令输出的端到端系统总延迟,涵盖了所有方法共有的数据采集、算法计算与指令生成环节。在相同硬件平台与工况激励下的测试表明,各方法在此统一指标下具有直接可比性,而 DRL 方法展现出的超低延迟(≤ 5 ms),正凸显了其前向推理机制在实时决策上的独特优势。结果如图 9 所示。

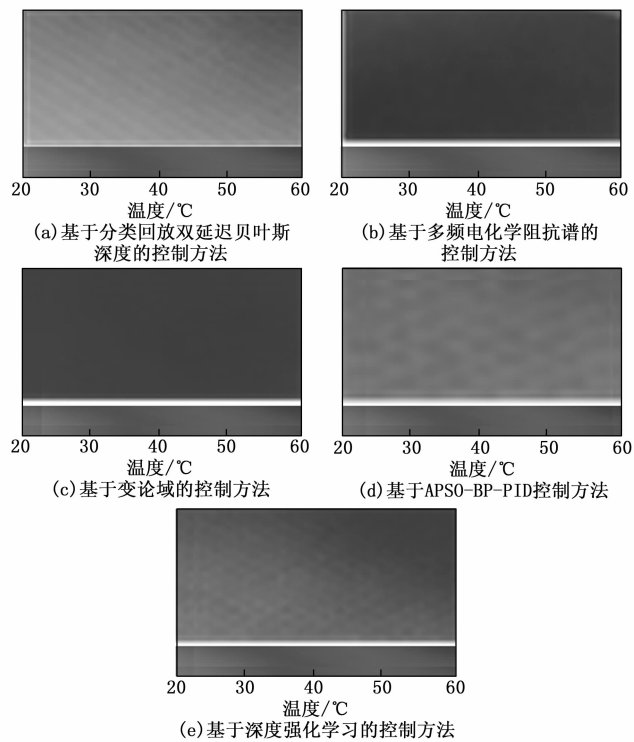


图 9 不同方法热量分布控制效果

由图 9 可知，使用基于分类回放双延迟贝叶斯深度的控制方法、基于多频电化学阻抗谱的控制方法、基于变论域的控制方法没有充分考虑到热量在空间上的传递和扩散特性，忽视了整体热量分布的均匀性要求，使得热量在空间上呈现出以深色为主的分布，即热量集中区域和低热量区域差异明显，与理想均匀分布不一致；使用基于 APSO-BP-PID 控制方法无法完全准确地描述热量传递的物理过程，方法中的一些假设和简化可能在实际应用中引入误差，导致对热量分布的预测和控制不够精确。即使颜色渐变呈现一定规律，但由于模型精度的限制，仍然无法实现与理想情况下完全一致的热量分布；使用基于深度强化学习的控制方法能够准确地捕捉热量传递过程中的复杂非线性关系，它可以对热量在空间上的分布进行精确建模，并根据模型预测结果制定相应的控制策略，确保热量在空间上实现均匀分配，实现任意两个区域之间的温度差 $\leq 1\text{ }^{\circ}\text{C}$ ，与理想情况一致。通过颜色渐变从深色到浅色的呈现方式，直观地展示了该方法在热量均匀分配方面的优异性能。

对比不同方法的工况突变响应延迟控制效果如图 10 所示。

由图 10 可知，基于分类回放双延迟贝叶斯深度的控制方法在 80~120 ms 区间响应延迟峰值超过 40 ms，远高于理想阈值。原因是参数更新和决策逻辑的计算耗时较长，导致工况突变时无法快速输出控制指令；基于多频电化学阻抗谱的控制方法响应延迟全程维持在

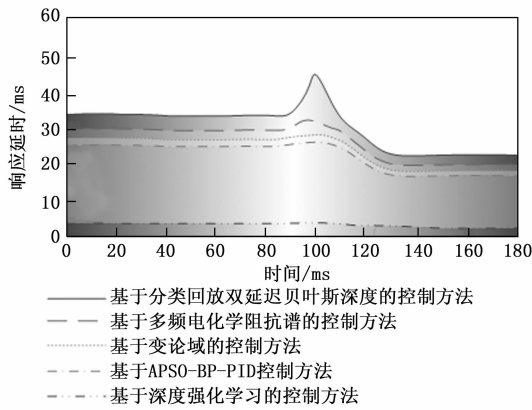


图 10 不同方法工况突变响应延迟控制效果

30 ms 左右，远超理想范围。原因是多频电化学阻抗谱需对多频段电信号进行采集、解析和特征提取，信号处理流程繁琐，难以在工况突变的瞬间完成控制决策；基于变论域的控制方法响应延迟稳定在 20~30 ms 区间，未达理想要求。原因是工况突变的动态场景下规则调整和论域变化的迭代过程耗时，无法实现毫秒级快速响应；基于 APSO-BP-PID 的控制方法响应延迟约 25 ms，仍超出理想范围。原因是工况突变时的优化/整定耗时，导致响应延迟偏高；基于深度强化学习的控制方法该方法的响应延迟全程 $\leq 5\text{ ms}$ ，完全满足“理想数值 $\leq 5\text{ ms}$ ”的性能要求。优势在于深度强化学习通过智能体与热管理系统的实时交互，以奖励机制直接学习最优控制策略，无需复杂的模型推理、信号解析或参数迭代过程，能在工况突变瞬间（亚毫秒级）快速输出控制指令，确保冷却调节的及时性，从根本上避免电池温度失控风险。

为验证所提出的 DRL 自适应控制系统在面对电池容量衰减（老化）、随机大功率负载冲击、以及传感器微小漂移等复合干扰时的控制性能与稳定性，在电池老化、负载突变和传感器噪声与漂移干扰下进行热安全等测试，结果如表 2 所示。

表 2 不同方法的热安全测试

评估指标	未采用深度强化学习方法	采用深度强化学习方法
温控均方根误差/ $^{\circ}\text{C}$	2.8	1.9
热安全/ $^{\circ}\text{C}$	48.2	46.5
累计时间/ms	312	185
负载突变后恢复时间/ms	45	32

基于上述实验表明，深度强化学习的控制方法在动力电池热管理中表现出色的核心，在于其数据驱动和端到端优化的本质。该方法通过深度神经网络，能够直接从运行数据中隐式学习电池包复杂的三维热动态与电热耦合关系，摆脱了对难以精确建模的物理模型的依赖。

这使得智能体能够基于全局温度场趋势进行决策,通过协调多个冷却执行单元,主动实现产热与散热的空间最优匹配,从而确保了优异的热量分布均匀性。

在响应速度上,该方法将决策过程简化为一次高效的前向网络计算,实现了感知即决策,避免了传统方法中多步骤串行处理带来的延迟。更重要的是,其策略是通过最大化长期奖励训练而来,内嵌了对未来状态变化的预测能力。因此,它能在电流骤增等工况突变征兆出现时,立即发出前瞻性的控制指令,而非被动响应,从而将响应延迟降至最低。

综上,深度强化学习的优势源于其将整体优化、实时推理与前瞻控制融为一体的自适应机制,完美契合了电池热管理高维、非线性、快时变的挑战。

4 结束语

针对新能源汽车高负荷运行下动力电池组面临的复杂电热耦合、温场失衡与精准控制难题,本文以数据驱动与闭环精准控制为核心,设计并实现了基于深度强化学习的动力电池组降温自适应控制系统。通过“热特性多维度采集—整车工况通信—自适应冷却执行”三层协同的闭环硬件架构,构建了数据基础与物理执行通路。软件层面,将动力电池温度偏差变化率作为关键状态特征,与工况信息融合,并利用深度确定性策略梯度算法构建了“状态—动作—奖励—更新”的闭环学习框架,实现了冷却策略的动态自寻优。实验结果表明,该系统模组内温度分布均匀(温差 $\leq 1^{\circ}\text{C}$),动态响应迅速(延迟 $\leq 5\text{ ms}$),显著优于传统控制方法,这得益于DRL策略对电池包整体热场空间关联性的隐式学习与协调控制能力,实现了产热与散热在空间上的最优匹配,有效避免了局部过热。综上所述,本文所提系统不仅完成了动力电池热特性与冷却控制自适应匹配的研究目标,更在实际极端场景验证中,在控制精度、响应速度与空间均匀性上实现了显著提升。

参考文献:

- [1] 张家金,林歆悠.融合电池温度和寿命的深度强化学习PHEV能量管理策略[J].福州大学学报:自然科学版,2023,51(1):89-96.
- [2] 陈鼎,马海玲,李京,等.真空蒸馏炉在新能源汽车锂离子电池负极材料中的应用[J].真空,2025,62(4):64-68.
- [3] 赵洪山,潘思潮,马利波,等.基于分类回放双延迟贝叶斯深度确定性策略梯度的燃料电池温度控制[J].电工技术学报,2024,39(13):4240-4256.
- [4] 沈珍华,戴锋,焦凤,等.基于多频电化学阻抗谱的锂离子电池内部温度估计新方法[J].中国电机工程学报,2025,45(5):1913-1921.
- [5] 张涛,范福磊,张博强,等.基于变论域的燃料电池温度模糊控制系统仿真研究[J].模糊系统与数学,2024,38(4):64-72.
- [6] 商蕾,杨萍,杨祥国,等.基于APSO-BP-PID控制的质子交换膜燃料电池热管理系统温度控制[J].吉林大学学报(工学版),2024,54(9):2401-2413.
- [7] 单春贤,杨鹏,唐爱坤,等.电动汽车电池组冷媒直冷系统工作特性的试验[J].江苏大学学报(自然科学版),2024,45(1):30-37.
- [8] 柳秀丽,李潇磊,陈学,等.GaSb电池热控温度与辐射源温度的匹配性分析[J].工程热物理学报,2024,45(8):2475-2482.
- [9] 申江卫,高承志,舒星,等.基于迁移模型的锂离子电池宽温度全寿命SOC与可用容量联合估计[J].电工技术学报,2023,38(11):3052-3063.
- [10] 陈轶嵩,兰子剑,蔡旭,等.车用燃料电池堆变负载电流下的出入口温度控制研究[J].汽车工程,2024,46(12):2303-2313.
- [11] 王世学,范玉聪,蹇季廷,等.固体氧化物燃料电池内温度分布测量技术进展[J].工程热物理学报,2024,45(10):2895-2920.
- [12] 朱泽华,吴笑宇,张恒运.锂离子电池模组热扩散结构的优化[J].机械设计与制造,2023(1):257-263.
- [13] 高帅,王俊博,朱佳慧,等.制冷剂相态变化对动力电池直冷系统温控性能的影响分析[J].制冷学报,2023,44(3):58-66.
- [14] 张小坤,吕大伟,尹中强,等.平板型固体氧化物燃料电池内温度分布规律[J].中国电力,2023,56(6):123-131.
- [15] 李军求,刘子鸣,刘正楠,等.基于EIS的锂离子电池内部温度估计及内短路诊断策略研究[J].汽车工程,2025,47(6):1112-1121.
- [16] 李先允,冯瀚飞.基于NSGA2算法的燃料电池系统净功率优化[J].电源技术,2023,47(1):83-87.
- [17] 李紫轩,荣浚合,王新改,等.基于NTC温度传感器的锂电池内部温度监测技术研究[J].电源技术,2024,48(2):276-283.
- [18] 张晖,李少鹏,丁乃侃.电动汽车电池温度一致性与驾驶行为关联特性研究[J].中国公路学报,2024,37(3):216-230.
- [19] 韩佳洁,苑清扬,李郁,等.基于少测点数据驱动的大规模锂电池模组温度实时预测[J].工程科学学报,2025,47(5):1103-1116.
- [20] 夏雪磊,杨博,赵志刚,等.基于MILS-MIUKF算法的锂离子电池宽温度全寿命荷电状态估计[J].中国公路学报,2024,37(8):289-301.
- [21] 刘志强,臧阿顺.基于BOA-CNN-BiLSTM的锂离子电池SOC估计[J].计算机仿真,2025,42(7):229-235.