

DOI: 10.20189/j.cnki.CN/61-1527/E.202603002

机器学习赋能有限样本数据下 锂离子电池健康状态估计方法综述

李爱华, 王鹏翔, 王涛*, 许晓东, 项纪钧

(火箭军工程大学, 西安 710025)

摘要: 锂离子电池健康状态 (state of health, SOH) 准确估计是保障电池系统安全运行与健康管理的核心技术。实际应用中, 电池全生命周期实验周期长、标注成本高, 且真实运行数据常表现为低标注样本、短时充电稀疏样本和工况覆盖不足, 有限样本数据已成为制约机器学习 SOH 估计精度与泛化能力的重要因素。基于此, 综述了有限样本数据下机器学习 SOH 估计方法的研究现状, 重点分析了回归方法、浅层学习、深度学习与迁移学习 4 类方法的性能差异, 并从估计精度、计算复杂度、部署条件和适用性等方面进行对比分析。结果表明, 回归方法具有概率推理优势但计算成本较高, 浅层学习结构简洁但外推能力有限, 深度学习特征提取能力强但高度依赖样本规模, 迁移学习在有限样本数据下表现更优, 但仍面临负迁移和训练稳定性不足等问题。未来应重点突破无标签数据挖掘、物理约束融合、多智能感知及实际工况部署等难题, 以实现稳健、可迁移且可解释的锂离子电池 SOH 估计。

关键词: 锂离子电池; 健康状态估计; 机器学习; 有限样本数据; 数据高效

中图分类号: TM912

文献标志码: A

开放科学标识码(OSID):

文章编号: 2097-2741(2026)03-0024-13



听语音
聊科研
与作者互动

随着全球清洁能源与低碳经济的快速推进^[1], 锂离子电池凭借高能量密度、长循环寿命和低自放电率等优势, 已成为电动汽车、储能系统和便携式电子设备的核心动力源^[2]。然而, 电池在服役过程中会不可避免地发生老化和性能衰退, 其性能衰退受内部结构演化、多老化机理耦合以及材料间物理化学反应共同作用。因此, 锂离子电池健康状态 (state of health, SOH) 的准确估计对保障电池系统运行安全、优化充放电策略、延长服役寿命及降低运维成本具有至关重要的意义。

目前, 锂离子电池 SOH 估计方法已形成多元技术体系, 其中机器学习方法因适应性强、建模灵活成为研究主流。但受限于模型表征能力及数据依赖特性, 此类方法在实际应用中仍存在诸多固有缺陷。例如, 高斯过程回归 (Gaussian pro-

cess regression, GPR)^[3]虽然能够通过核函数量化预测不确定性, 但高昂的矩阵求逆计算量以及对老化拐点追踪的迟滞, 限制了其在实时嵌入式平台上的部署。多层感知器 (multilayer perceptron, MLP)^[4]虽具备强非线性映射能力, 但在数据稀疏场景下极易受随机噪声影响, 从而陷入局部最优并引发过拟合。支持向量机 (support vector machine, SVM)^[5]通过结构风险最小化准则在小样本学习中表现较稳健, 但其性能高度依赖核函数与超参数选择, 难以在有限的搜索空间内准确映射多物理场耦合下的非平稳退化轨迹。深度学习范式如卷积神经网络 (convolutional neural network, CNN)^[6]、循环神经网络 (recurrent neural network, RNN)^[7]及其变体在处理高维时序特征方面优势显著, 但在小样本条件下因缺乏有效

收稿日期: 2026-04-03

修回日期: 2026-04-27

录用日期: 2026-05-08

基金项目: 陕西省自然科学基金研究计划 (2025JC-YBQN-686)

第一作者: 李爱华 (1966—), 男, 教授, 主要从事故障诊断、信号处理等研究。

***通信作者:** 王涛 (1977—), 男, 教授, 主要从事故障诊断、图像处理等研究。

引用格式: 李爱华, 王鹏翔, 王涛, 等. 机器学习赋能有限样本数据下锂离子电池健康状态估计方法综述[J]. 火箭军工程大学学报, 2026, 40 (3): 24-36. LI Aihua, WANG Pengxiang, WANG Tao, et al. Review of machine learning-based methods for lithium-ion battery state of health estimation with limited sample data[J]. Journal of Rocket Force University of Engineering, 2026, 40 (3): 24-36.

的归纳偏置, 易聚焦伪相关特征而忽略关键老化机理。迁移学习虽通过跨域知识共享缓解了标注数据不足的问题, 但源域与目标域不匹配时常引发负迁移风险。

上述各类机器学习方法在锂离子电池 SOH 估计中暴露出的过拟合、泛化能力不足、负迁移等一系列应用短板, 并非单纯的模型结构或算法参数问题, 其核心本质是实际工况下有限样本数据的客观条件所致, 这也是制约各类机器学习算法落地应用的共性问题。有限样本数据问题在实际工况下形成的原因主要有以下四方面: 一是电池全生命周期实验长达数月至数年, 测试设备与耗材成本高昂, 难以批量获取足量标注数据; 二是实际工况复杂多变, 难以覆盖电池老化完整状态空间; 三是退役及二次利用电池的历史运行数据缺失, 真实健康标签难以追溯; 四是规模化场景下实时标注成本过高, 无法支撑大规模高质量数据集构建。针对上述成因, 本文明确界定研究涉及的有限样本主要指两类典型场景: 一是标注比例通常低于 10% 的低标注样本; 二是单次充电时长短于 15 min 的稀疏样本。本文将重点围绕这两类场景下的研究现状展开综述, 以增强综述的针对性与参考价值。

然而, 现阶段针对上述特定有限样本场景的研究仍处于起步阶段, 多数文献仍偏重于通用估计方法的研究。例如, 金帅等^[8]聚焦于锂离子电池数据驱动的 SOH 估计研究; 陈宏宇等^[9]聚焦于退役动力电池 SOH 估计研究; Vignesh 等^[10]聚焦于二次寿命电池 SOH 估计研究。尽管上述研究从不同角度推进了 SOH 估计技术, 但在实际工况下, 有限样本数据的问题仍是各类方法共同面临的核心难题。机器学习方法虽在特征提取与非线性建模上具备天然优势, 但因样本依赖特性, 在有限样本数据下易出现泛化能力不足、过拟合风险加剧、物理一致性缺失等瓶颈, 制约了其在真实工况下的落地应用。鉴于此, 本文旨在系统性梳理有限样本下锂离子电池机器学习 SOH 估计方法的研究现状, 分析关键技术瓶颈, 并展望未来发展方向, 为数据高效的机器学习 SOH 估计技术研究提供理论参考与技术借鉴。

1 SOH 的定义与老化机理

1.1 SOH 的定义

在长期充放电循环与存储过程中, 电池容量

会发生不可逆衰减, 内部化学材料也会伴随性能退化。当容量与性能衰退至临界阈值时, 需及时更换电池以保障用电设备稳定运行。然而, 电池在材料体系、结构设计及工作环境上的显著异质性, 使得有限样本下快速、准确的 SOH 估计极具挑战性^[11]。在实际工况中, 通常将 SOH 的 80% 设为电池寿命结束 (end of life, EOL) 以及需要更换的界限, 以确保电池性能的有效发挥。而电池 SOH 可以从容量以及内阻^[12]两方面进行定义:

$$\text{SOH} = Q_n / Q_N \quad (1)$$

$$\text{SOH} = (R_{\text{EOL}} - R_n) / (R_{\text{EOL}} - R_N) \quad (2)$$

式中: Q_n 为当前循环时的最大放电容量; Q_N 为新电池的额定容量; R_{EOL} 、 R_n 和 R_N 分别为寿命结束时、当前时刻和新电池的内阻。

1.2 锂离子电池老化机理及影响因素

锂离子电池作为民用领域的主流电源, 受内部副反应与外部因素的共同作用持续老化, 主要表现为容量衰减与功率衰减^[13]。电池老化通常分为两类: 日历老化和循环老化^[14]。日历老化是指电池在静态或非工作状态下, 随时间推移而发生的不可逆的性能衰减; 循环老化是指电池在充放电循环工作过程中发生的性能衰减。为深入探究底层老化机理, 研究中通常将电池衰减分为三大核心模式^[15]: 可用锂离子损失 (loss of lithium inventory, LLI)、正极活性材料损失 (loss of electrode active material, LAM) 和阻抗增加 (resistance increase, RI)。

锂离子电池老化是一个复杂过程, 由作用于电池各关键组分的耦合机制共同驱动。深入分析组分是理解电池整体老化过程的核心前提, 下文将详细阐述正极、负极、电解质及其他关键组分的老化机理。

正极在老化过程中会发生多种化学和物理反应。其中, 化学老化反应主要包括过渡金属溶解^[16]、正极电解液界面 (catholyte electrolyte interface, CEI) 膜的形成与演化^[17]、气体生成以及正极活性材料溶解; 物理老化反应主要包括相变^[18]、颗粒开裂^[19]和晶体结构变化^[20], 这些均会破坏电极的结构完整性及离子或电子的导电性。

负极老化现象同样复杂多样。化学老化反应主要包括固态电解质界面 (solid electrolyte interphase, SEI) 膜的形成、增厚、破裂与重构^[21]、枝晶锂析出^[22]和气体生成; 在物理老化方面, 负极易受机械应力诱导的结构变化^[23]、相变和晶体结构改变影响, 会发生 LLI 和 RI。

电解质的老化过程涉及其成分与电池其他组分之间的众多化学反应。例如, 电解液中的主要成分的热分解与水解^[24]、有机溶剂的氧化还原反应以及电极与电解质界面处的不良反应^[25]。这些过程通常会导致气体产生、LLI以及RI。

其他关键组分, 如集流体、隔膜和黏结剂,

其劣化在电池整体老化过程中也起着重要的作用。物理老化反应主要包括集流体开裂^[26]、接触失效和隔膜孔堵塞^[27]。从化学角度来看, 这些组分会发生集流体腐蚀和黏结剂分解^[28], 从而进一步加剧电隔离和容量衰减。锂离子电池主要老化机理如图1所示。

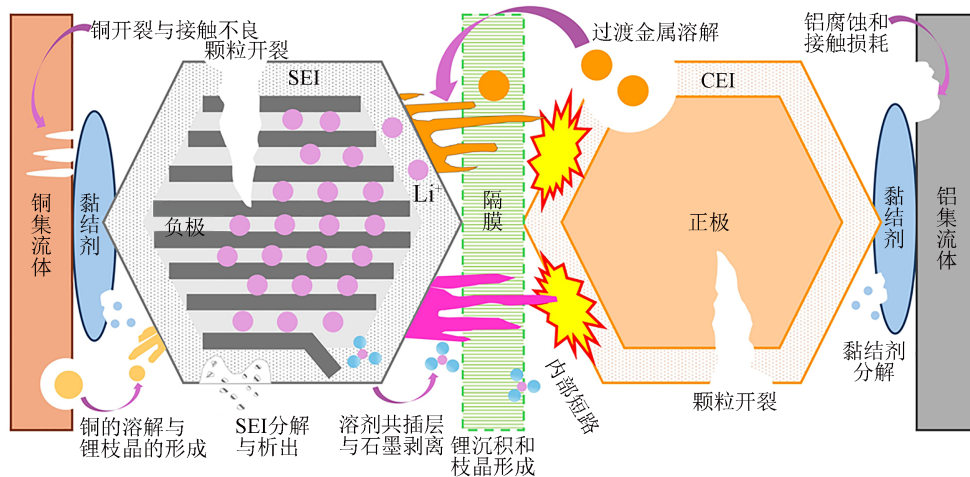


图1 锂离子电池的主要老化机理

Fig. 1 Main aging mechanisms for lithium-ion batteries (LIBs)

电池的老化并非由一种因素主导而成, 而是由内部组分在化学和物理等方面产生的老化反应以及外部因素共同造成。外部因素主要包含温度、放电深度、充放电速率、荷电状态 (state of charge, SOC) 等。其中, 温度对电池老化的影响最大^[29], 高温会加速电池容量的衰退, 发生的老化反应主要为SEI膜的增厚以及锂枝晶的形成, 这会给电池造成LLI以及RI, 甚至会发生内部短路问题; 而低温会影响锂离子的嵌入动力学、诱发析锂。高充放电速率会加速电池的老化, 并且随着充放电速率的增加, 电池的负极表面会发生较为显著的析锂以及结构的变化^[30], 造成显著的LAM以及LLI。此外, 放电深度对不同正极材料的影响差异较大^[31], 但高放电速度会加速电池结构改变, 如CEI膜的增厚等, 从而影响电池的老化。高SOC状态主要会加速SEI膜的生长以及增加过渡金属溶解的风险, 特别是当高SOC与高充放电速率相结合时, 析锂的风险会显著增加, 进而导致电池出现LAM与LLI^[32]。

2 基于机器学习的锂离子电池SOH估计方法研究现状

基于机器学习的各类方法在锂离子电池SOH估计中应用广泛, 本文选取回归方法、浅层学习、

深度学习、迁移学习4类主流且应用最广泛的核心方法, 梳理其在有限样本数据下的应用与瓶颈。

2.1 回归方法

锂离子电池SOH估计常使用到的回归方法包括线性回归及非线性回归等。其中, 线性回归在锂离子电池SOH估计中单独运用的情况较少, 一般都是和其他方法联合运用。例如, Ni等^[33]提出结合容量损失机理特征提取、变分模态分解去噪、线性回归与贝叶斯推理的混合模型, 实现老化拐点的准确识别。而非线性回归通过拟合容量、内阻等老化数据与循环次数的复杂非线性关系, 在线建立并更新SOH估计模型。如Lee等^[34]结合非线性回归与主成分分析, 全面提取并利用电池系统的各类衰减特征, 提升SOH预测性能。

在回归方法中, GPR在电池诊断领域应用广泛, 基于GPR的SOH估计框架如图2所示。该流程涵盖了从原始信号 (包括电压、电流、温度) 采集、容量增量分析曲线特征提取, 到利用训练特征矩阵与真实SOH标签来构建模型, 并最终基于新数据特征实现概率状态预测的全过程。GPR是基于贝叶斯理论的非参数概率推理模型, 通过核函数定义函数空间上的先验分布。与传统参数模型不同, GPR无需预定义解析函数形式, 而是通过随机过程理论捕捉输入空间的复杂非线性映射。GPR的核心优势在于能在点预测的同时量化

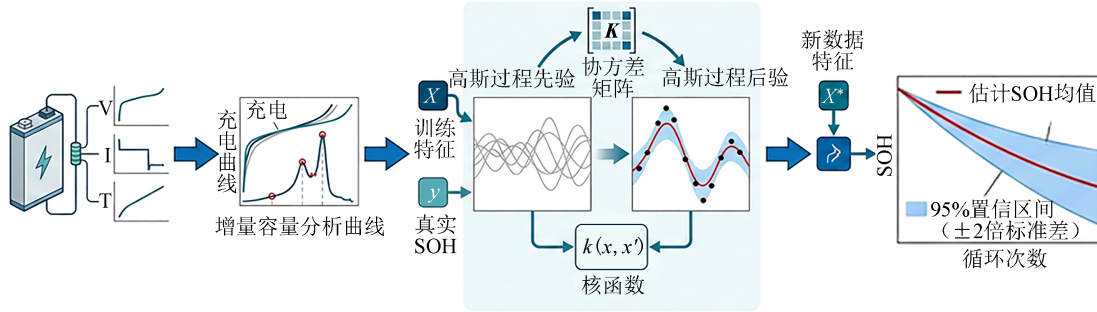


图2 基于GPR的锂离子电池SOH估计框架

Fig. 2 Framework of the SOH estimation method for LIBs based on GPR

估计的不确定性。例如, Pohlmann等^[35]提出基于自动相关性判定核函数GPR的多化学体系电池SOH估计框架, 其平均绝对误差低至1.33%。Yu^[36]开发3种集成多尺度逻辑回归与GPR的方法, 解决了SOH估计的复杂性难题。尽管GPR能从有限样本数据中提供概率性SOH估计, 但其在电池工程中的实际部署受高计算成本以及高敏感性的制约^[37]。首先, GPR需要对协方差矩阵求逆, 计算复杂度较高, 只能适应少量样本的数据集, 难以进行海量样本数据集的SOH估计。其次, GPR通常难以对非线性的老化过程进行SOH估计, 尤其是电池老化过程中的拐点。此外, GPR模型精度非常依赖于超参数辨识质量, 当训练数据较少时, 测量噪声对模型估计精度影响较大, 使其对噪声高度敏感。因此, 纯数据驱动的GPR模型在实际工况下的SOH估计精度不高, 若融入物理约束, 其精度以及鲁棒性能够得到较大提升。

2.2 浅层学习

锂离子电池的SOH估计是一个多因素耦合问题, 利用不同的模型对其进行建模估计可能会有不同的效果。在当前各类模型算法盛行的背景下, 浅层学习方法凭借其简洁的结构、较高的计算效率以及在小样本集上表现出的稳健性, 成为学术界与工业界关注的焦点。不同于依赖多隐藏层捕捉特征的深层架构, 浅层学习侧重于通过特征工程与统计学准则, 在有限的计算资源下实现对电池高度非线性演化轨迹的准确拟合。常见的浅层学习范式有MLP、SVM、随机森林 (random forest, RF)、极限学习机 (extreme learning machine, ELM)、相关向量机 (relevance vector machine, RVM), 其中MLP与SVM最具代表性^[8]。

2.2.1 多层感知器 (MLP)

MLP是一种典型的前馈人工神经网络模型^[38], 其核心价值在于通过模拟生物神经元的连接与激活方式, 实现从多维输入空间到输出空间的复杂

映射。图3(a)展示了用于锂离子电池SOH估计的MLP结构, 它由输入层、一个或多个隐藏层以及输出层构成, 其中 $x_1, x_2, \dots, x_n$ 代表输入的电压、电流、温度等表征电池退化的关键物理参数, w_1, w_2 为层间连接权重, $h_i (i = 1, 2, \dots, n)$ 为隐藏层神经元节点。输入参数经由隐藏层的非线性变换, 最终准确输出反映电池老化程度的SOH预测结果。

在早期研究中, Williard等^[39]利用MLP架构验证了处理电池状态预测任务的理论可行性。后续的研究则更多聚焦于将MLP与其他算法相结合, 以提升其在复杂环境下的参数辨识能力。例如, 针对传统电化学阻抗谱中特征冗余导致模型训练困难的问题, Wang等^[40]提出了一种融合弛豫时间分布物理特征与MLP的SOH估计框架, 通过特征选择算法提取阻抗谱中的关键峰值组合, 构建了高信息密度的输入向量, 充分利用了MLP在复杂非线性映射中的强拟合能力。

2.2.2 支持向量机 (SVM)

SVM由Cortes等^[41]提出, 并已成功应用于回归问题, 包括电网负载预测、故障诊断和图像处理等。SVM是一类按监督学习方式对数据进行二元分类的广义线性分类器, 其决策边界是对学习样本求解的最大边距超平面。SVM可以通过核方法进行非线性分类, 是常见的核学习方法之一。不同于其他机器学习方法, SVM遵循结构风险最小化准则, 通过协同优化拟合精度与模型复杂度, 有效解决了小样本场景下的过拟合难题, 这一特性使其在SOH估计中具备优异的泛化性。用于锂离子电池SOH估计的SVM模型结构如图3(b)所示, 其中 $K(x_i, x_j) (i = 1, 2, \dots, n)$ 代表核函数。

在锂离子电池SOH估计方面, SVM也有着广泛应用。例如, 针对SOH预测中常见的鲁棒性问题, Liu等^[42]提出了一种改进的非线性灰狼优化-内核支持向量机模型。实验结果表明, 这种改

进不仅适应变温工况,且在训练数据有限的情况下也能有效提升预测精度。针对退役电池在线 SOH 估计问题, Xiong 等^[43]针对老化非线性特征显著、样本稀缺的挑战,提出基于加权最小二乘 SVM 的解决方案。实验表明,该改进方法的 SOH 估计精度显著优于传统 GPR 模型,均方根误差稳定在 1.85% 以内。

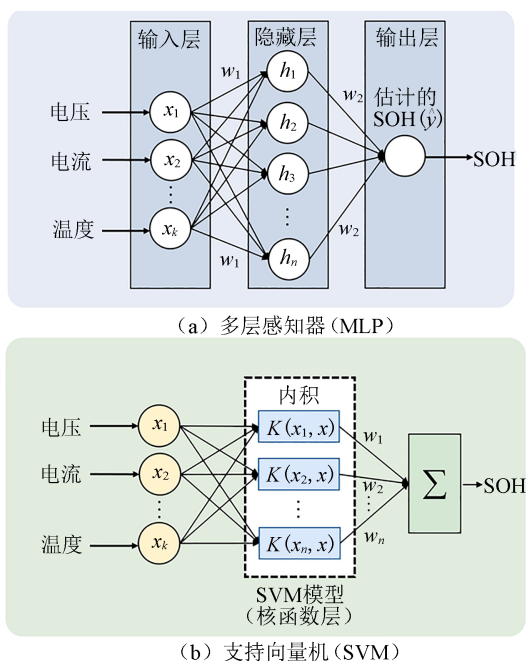


图 3 锂离子电池健康状态估计中典型的浅层学习模型

Fig. 3 Representative shallow learning models applied in LIB SOH estimation

在 SOH 估计的实际应用中,获取全生命周期老化数据通常需要数月乃至数年的实验周期,而有限样本数据是大多数锂离子电池 SOH 估计所面临的问题。尽管浅层学习中的 MLP 与 SVM 模型结构简洁,但在有限样本数据条件下,二者的建模局限性依然十分突出。MLP 容易在有限样本数据条件下发生过拟合现象,且基于梯度下降的优化算法在数据不足时极易陷入局部最优解。对于 SVM 而言,在有限样本数据条件下寻找能够准确反映非线性老化轨迹的核函数以及超参数的难度极大。同时,SOH 估计效果非常依赖于核函数的选择。其他的浅层学习算法也存在类似的局限性,因此,在有限样本数据下实现锂离子电池 SOH 的准确估计仍是浅层学习架构面临的核心挑战。

2.3 深度学习

深度学习技术凭借强大的特征自主提取与非线性建模能力,在众多工程领域得到广泛应用,在锂离子电池 SOH 估计领域也成为主流研究手段,

现阶段绝大多数相关研究均依托深度学习模型展开。目前,应用于锂离子电池 SOH 估计的深度学习模型种类丰富,主要包括 CNN、RNN、自编码器 (autoencoder, AE) 以及 Transformer 等模型。

2.3.1 卷积神经网络 (CNN)

CNN 是一类前馈神经网络,也是深度学习的核心模型之一,被广泛应用于图像识别与视觉任务。该模型通过卷积层与池化层的协同作用,实现电池老化特征的高效提取。卷积层借助局部感受野与参数共享机制,通过可学习卷积核对原始时序数据进行滑动扫描,不仅能显著降低模型计算复杂度,更能捕捉反映电池衰减的细微局部特征,包括电压平台演化、放电曲线斜率变化、极化效应引发的电压波动等。随后,池化层通过最大或平均下采样对特征图进行空间压缩,在实现数据降维的同时,能有效提取最具代表性的健康指示特征,抑制传感器采集过程中的随机噪声与信号冗余,显著提升模型在复杂动态工况下的鲁棒性与泛化能力。图 4 (a) 展示了基于二维卷积神经网络的锂离子电池 SOH 估计模型架构。

CNN 在锂离子电池 SOH 估计中应用广泛, Chemali 等^[44]构建了包含 1~6 层的 CNN 模型,通过输入充电过程多维监测数据实现端到端 SOH 估计,将误差控制在 1.6% 以内。针对实际应用中充电数据碎片化的挑战, Tian 等^[45]开发了深度学习方法,从不足 10 min 的稀疏输入中准确估计完整充电曲线,实现了高精度的衰减监测。

CNN 在有限样本数据条件下的局限主要表现为特征提取鲁棒性不足^[6]。CNN 依托卷积核提取局部感受野内的通用特征,该机制需要依托海量多样化样本完成滤波器训练,才能精准识别具备物理意义的电池老化特征。在训练数据稀缺时,卷积层易捕捉单个电池的制造差异或随机测量噪声,而非通用衰减特征,进而引发伪相关问题。同时,深度 CNN 架构理论上可提取高阶语义特征,但数据受限条件下过深的网络结构不仅会加大计算负担,还可能因特征空间过于稀疏导致特征退化,使提取的特征无法有效映射锂离子电池 SOH 真实演化轨迹。

2.3.2 循环神经网络 (RNN)

RNN 是一类专用于处理序列数据的深度学习模型,核心优势在于捕捉数据中的时序依赖关系,弥补传统深度神经网络与 CNN 无法利用历史信息的缺陷。RNN 由输入层、隐藏层与输出层构成,

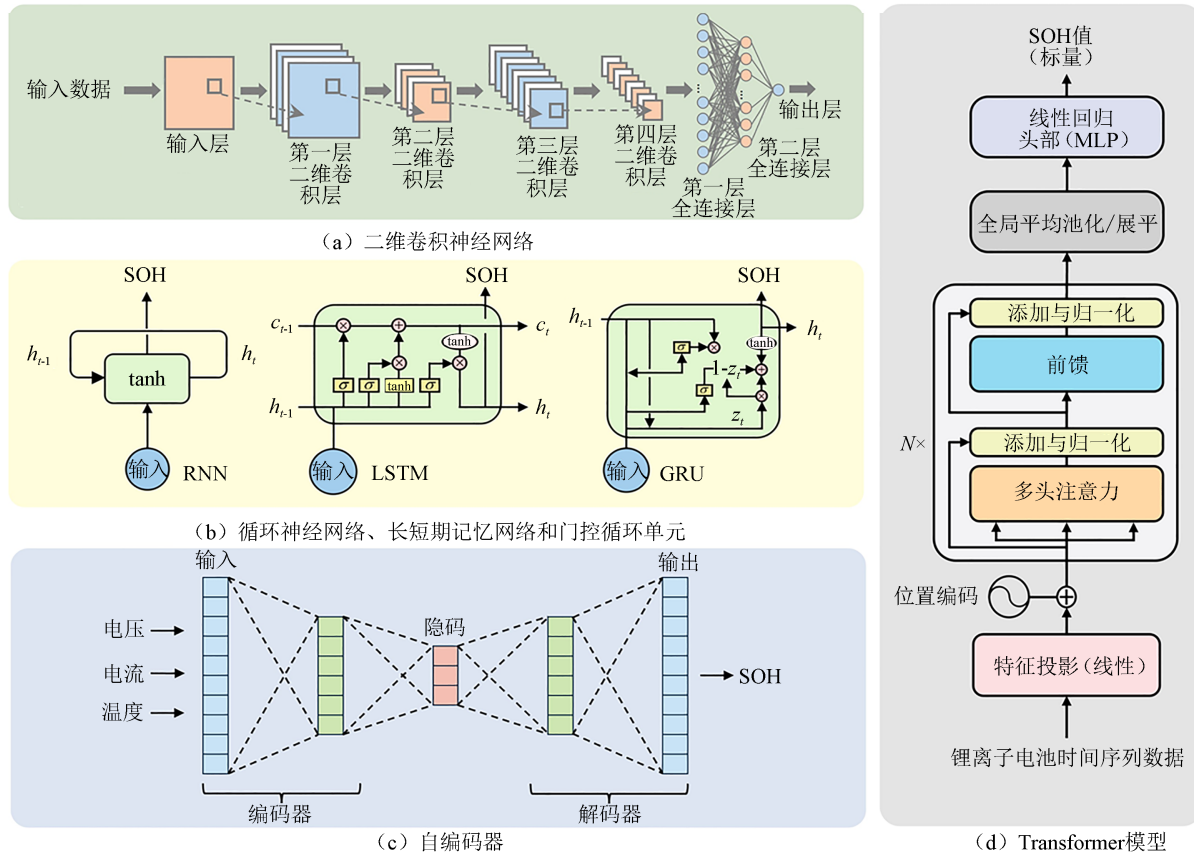


图4 用于锂离子电池健康状态估计的深度学习架构

Fig. 4 Deep learning architectures employed for LIB SOH estimation

凭借时序数据动态特征捕捉能力,被广泛应用于锂离子电池SOH序列预测任务。RNN通过时间步间权重共享实现历史信息记忆,但存在梯度消失与梯度爆炸的关键缺陷,限制了其对长序列数据的处理能力。为解决这一问题,长短期记忆网络(long short-term memory, LSTM)与门控循环单元(gated recurrent unit, GRU)等变体模型应运而生。LSTM引入遗忘门、输入门与输出门,能有效缓解梯度衰减问题,使梯度在时间步间稳定传播,准确捕捉长距离依赖关系;GRU则包含重置门与更新门,重置门用于捕捉短期依赖,更新门用于捕捉长期依赖。在有限样本数据条件下,LSTM与GRU均表现出优异性能^[46]。图4(b)展示了RNN、LSTM与GRU的基本结构。图中, h_{t-1} 与 h_t 分别代表上一时刻与当前时刻的隐含状态; σ 与 $\tanh$ 为激活函数; c_{t-1} 和 c_t 分别为上一时刻和当前时刻的细胞状态; z_t 为GRU的更新门状态。

RNN和LSTM、GRU的核心挑战在于模型复杂度与数据规模不匹配,容易出现过拟合问题^[47],门控结构虽解决了梯度传播问题,但遗忘门、更新门等结构会显著增加训练参数数量。小样本场景下,模型易记忆训练集中的噪声而非底

层老化趋势,导致未知工况下的预测精度大幅下降。此外,有限样本数据通常无法覆盖电池老化过程的长尾分布,使RNN模型在极早期或晚期非平稳衰减特征捕捉上稳定性不足,同时因缺乏充足验证集,超参数优化过程极度敏感,难以实现全局最优。

2.3.3 自编码器(AE)

AE作为包含编码器与解码器的人工神经网络结构,由Ackley等于1985年在关于玻尔兹曼机的研究中进行了探索^[48],并由Yann于1987年提出^[49]。AE的输入与学习目标一致,结构分为编码器与解码器两部分,图4(c)展示了用于锂离子电池SOH估计的AE基本结构。

在锂离子电池SOH估计中,AE主要用于数据降维与特征提取,凭借编解码架构保留SOH估计关键信息,实现高维数据的低维映射。根据学习范式不同,可将AE分为收缩自编码器(contractive autoencoder, CAE)、正则自编码器(regularized autoencoder, RAE)与变分自编码器(variational autoencoder, VAE)。前两者为判别模型,后者为生成模型,各类编码器在不同场景下的SOH估计具备差异化优势。例如,CAE可在数

据噪声显著时,通过压缩学习更稳健的SOH相关关键特征^[50];RAE在数据量有限时,能有效防止过拟合,提升SOH估计模型的泛化能力^[51];VAE则适用于数据不完整、需量化不确定性或生成衰减场景的SOH预测与故障预警^[52]。

以AE为代表的无监督或半监督模型,在有限样本数据条件下面临隐空间表征不连续的问题^[53]。AE通过重构误差学习数据的低维流形分布,若输入样本未覆盖电池老化完整状态空间,学习到的隐特征分布会出现碎片化或高度畸变,进而导致解码器在对未观测到的老化状态进行生成或推断时,产生显著的预测偏差。尤其是采用VAE时,有限样本数据无法支撑复杂后验分布估计,易引发后验坍塌现象,导致生成的健康因子失去物理可解释性,无法为后续SOH回归任务提供高质量输入特征。

2.3.4 Transformer 模型

Transformer模型架构由Vaswani等在2017年提出^[54],采用自注意力机制替代RNN与CNN,由编码器与解码器构成,编码器通过多层自注意力机制提取特征,解码器将输入与目标序列关联,并引入位置编码补充序列位置信息。图4(d)展示了基于Transformer模型的锂离子电池SOH估计架构, $N \times$ 表示由“多头注意力-添加与归一化-前馈-添加与归一化”组成的编码器模块重复堆叠次数和编码器层数。

近年来,基于Transformer的架构凭借注意力机制捕捉复杂依赖关系的能力备受关注,该架构性能优异,在自然语言处理与计算机视觉领域应用广泛,在锂离子电池SOH估计领域也展现出巨

大潜力。例如,Zhu等^[55]开发了基于Transformer的域自适应框架,解决了锂离子电池健康预测中的域偏移问题,通过融合CNN与Transformer架构实现高效跨域迁移,微调后实现SOH与剩余使用寿命高精度估计,在跨协议、跨化学体系场景下性能优于基准模型。Park等^[56]提出基于Transformer的通用健康评估框架,通过提取关键健康指标并利用Transformer编码器,实现了跨配置SOH估计且无需重新训练。

Transformer架构虽凭借自注意力机制在长序列建模中表现优异,但其对数据量的依赖程度远超其他模型^[57]。由于缺乏CNN的局部归纳偏置与RNN的时序归纳偏置,需要海量数据集从头学习位置关系与依赖模式。在小样本SOH估计任务中,Transformer难以构建稳定聚焦的特征图,注意力权重常因数据稀疏性而发散,模型无法关注关键电压或电流片段,这种数据密集特性使得Transformer在无大规模预训练或迁移学习辅助时,难以在有限样本数据上收敛至理想状态。

2.4 迁移学习

迁移学习是将源任务预训练模型应用于不同目标任务的核心机器学习范式,适用于多元场景,能有效缓解数据分布差异引发的问题,在数据可用性有限的条件下优势显著,通过将源域训练成熟的知识迁移至目标域,可大幅缩短训练时长、降低计算成本。针对有限样本数据条件下的锂离子电池SOH估计,迁移学习通过跨域知识共享,有效解决了目标域数据匮乏导致的模型泛化性差等痛点。典型的迁移学习SOH估计架构如图5(a)所示,主要分为两个阶段:首先,利用充足的

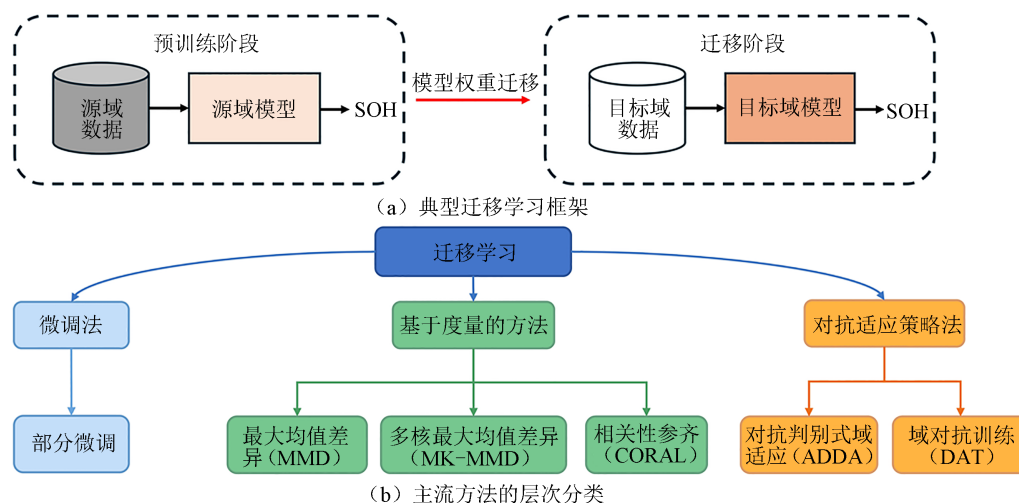


图 5 用于锂离子电池健康状态估计的迁移学习实现机制与系统分类方法

Fig. 5 Implementation mechanism and systematic taxonomy of transfer learning for LIB SOH estimation

源域数据训练基础模型;随后,通过权重迁移并结合少量目标域数据进行适配,从而实现对锂离子电池SOH的准确估计。基于迁移机制的不同,目前的主流方法可归纳为3类:微调法、基于度量的方法以及对抗适应策略法,如图5(b)所示。

微调法是在源域预训练模型后,在目标域重新开展训练,迁移过程中冻结预训练模型特定网络层的权重参数,仅利用目标域标注数据集优化未冻结层。针对传统机器学习方法的局限,Sun等^[58]利用6种正极材料的143节电池循环数据微调时序基础模型TimeGPT,在SOH估计任务中实现均方根误差与平均绝对误差分别低于1.06%与0.59%。

基于度量的方法通过直接最小化统计度量标准,缩小域间分布差异,主流技术包括最大均值差异(maximum mean discrepancy, MMD)^[59]、多核最大均值差异(multi-kernel maximum mean discrepancy, MK-MMD)^[60]与相关性对齐(correlation alignment, CORAL)^[61]。基于度量的迁移学习模型通常由特征提取器与预测器构成,特征提取器处理数据生成特征,特征输入预测器得到估计结果。该方法在迁移阶段需同时获取无标签源域数据与目标域数据,选择数学标准计算特征距离,量化域分布偏移并融入损失函数,通过随机梯度下降优化,通常仅目标特征提取器参数可训练,其余层冻结,度量标准项最小时即模型完成目标域适配。在SOH估计领域,Gui等^[62]提出多度量局部-全局卷积神经网络变换器模型,通过融合最大均值差异与相关对齐的多信息对齐策略,实现了跨域SOH高精度估计。

对抗适应策略法^[63]涵盖了对抗判别式域适应(adversarial discriminative domain adaptation, AD-DA)、域对抗训练(domain adversarial training, DAT)等方法。此类方法在标准特征提取器与预测器基础上引入域判别器,判别器用于区分源域与目标域数据,特征提取器则生成可欺骗判别器的特征,缩小域差异,这种极小极大博弈实现了与生成对抗网络核心概念类似的平衡状态^[64]。在SOH估计方面,Ye等^[65]开发了基于域对抗迁移学习的域对抗自适应网络模型,融合最大均值差异与相关对齐,将生成对抗学习与密集双向GRU特征生成器结合,有效提升了SOH估计精度与知识迁移性能。

尽管迁移学习为有限样本数据条件下的SOH估计提供了可行方案,但实际应用中仍面临源域

与目标域不匹配引发的负迁移风险^[66]。当正极材料或老化机理存在本质差异时,仅通过统计方法强制对齐特征分布无法弥合深层语义差距,甚至可能降低模型泛化性能。同时,有限样本数据会加剧参数优化与过拟合的矛盾,如微调法中的稀疏目标域数据可能无法支撑高维参数空间的细粒度调整,模型易收敛至局部最优解。最后,训练稳定性与计算成本会制约其工程部署,例如,对抗适应策略法难以维持生成器与判别器间的均衡以及基于度量的方法计算时需要源域数据的持续参与,这对电池管理系统存储资源与计算方面的能力提出了很高要求

2.5 机器学习SOH估计方法的对比分析

为了更直观地量化各类方法的性能差异,结合现有文献中基于公开数据集的典型验证结果,本文对上述方法在有限样本数据条件下的表现进行了对比分析。综合来看,各类算法在有限样本数据条件下的表现各有千秋。回归分析虽然能提供概率推理优化,但其计算复杂度较高,典型误差水平大约在1.3%~2.0%,高昂的矩阵运算成本限制了其在实时系统中的部署,通常需要依托高性能处理器。浅层学习结构简洁,计算复杂度低,可轻松部署于基础微控制器(microcontroller unit, MCU)上,但其在数据稀疏时易过拟合且外推能力受限,导致典型误差在1.5%~2.5%之间波动。深度学习因其极度的数据依赖性,在有限样本数据场景下的适用性相对较低,易出现特征退化,虽然其误差水平为1.0%~1.6%,但其落地部署高度依赖带有人工智能(artificial intelligence, AI)加速的专用芯片。相较之下,迁移学习通过领域自适应和微调,在有限样本数据下展现出较高的适用性,不仅能有效抑制分布偏移,将典型误差水平下探至0.8%~1.5%区间,且可通过云边协同架构进行部署,是目前缓解数据匮乏痛点的有效路径。各类方法的详细对比如表1所示。

3 挑战与展望

现阶段在锂离子电池SOH估计领域内,各类机器学习方法已形成多元技术体系,也涌现出诸多针对有限样本数据约束的优化思路与改进方案。但在工程应用中,有限样本数据带来的挑战依旧存在,现有方法在特征利用、机理融合、感知维度及工况适配等层面仍存在诸多亟待解决的难题,整体技术距离可以落地实施仍有较大差距。结合

本文研究梳理结果与锂离子电池健康管理实际需求,从无标签数据挖掘、物理约束融合、多智能感知、实际工况部署 4 个维度,剖析领域核心挑战,并展望未来可行的技术发展方向。

表 1 有限样本数据下锂离子电池各类机器学习 SOH 估计方法对比

Table 1 Comparison of machine learning methods for LIB SOH estimation with limited sample data

方法 范式	典型 算法	典型误 差水平	计算复 杂度	部署硬 件需求	适用 性	核心 瓶颈
回归 方法	GPR, 线性回归	约 1.3%~ 2.0%	较高	高性能 处理器	中	老化拐 点敏感
浅层 学习	MLP, SVM	约 1.5%~ 2.5%	低	基础 MCU	中	外推能 力受限
深度 学习	CNN, RNN	约 1.0%~ 1.6%	高	AI 加速 芯片	低	极端数 据依赖
迁移 学习	微调法, 对抗适应 策略法	约 0.8%~ 1.5%	较高	云边协 同架构	高	负迁移 风险

3.1 无标签数据挖掘

带标签的老化数据稀缺仍是影响 SOH 估计精度的核心,且锂电设备在实际运行中会产生海量无标签数据。而自监督学习范式^[67]为挖掘这些无标签数据提供了一条可行路径,其通过掩码序列重建或对比学习等前置任务从无标签数据中预训练提取通用退化特征,从而显著降低实验成本并提升泛化能力。此外,针对极端工况下样本数据缺失的挑战,扩散模型展现出巨大潜力,可以通过学习电池退化轨迹的复杂概率分布生成高保真合成数据,有效突破有限样本数据的约束,为模型训练提供高质量的数据增强支撑,从而实现从数据依赖向数据高效的转变。

3.2 物理约束融合

针对机器学习方法在有限样本下物理一致性缺失的局限,物理信息神经网络^[68]可通过融合电化学机理与深度学习范式提升物理可解释性以及泛化能力。然而,现有的实现路径多侧重于通过损失函数施加软约束,在数据稀疏场景下易引发梯度病态及收敛困难。未来的演进方向可聚焦于物理定律与网络架构的深度耦合。一方面,可利用电化学模型提取反映 LAM 或 LLI 的微观变量,将其作为物理特征输入网络,以增强对内部衰减状态的观测深度;另一方面,应探索将热-力耦合方程或半经验退化公式直接硬编码为计算图中的可微分模块。这种硬约束机制能通过物理先验严

格限制参数解空间,确保模型在超出训练域的外推预测中严格遵循质量与能量守恒,从根本上解决参数识别的不适定性问题,显著提升算法的工程落地价值。

3.3 多智能感知

若锂离子电池 SOH 估计仅依靠宏观电压、电流和温度等数据输入,会限制其对内部退化模式的可观测性。未来的传感框架应朝着异质信号的多模态^[69]融合方向发展,以弥补有限样本数据场景下的信息损失,包括将低频运行数据与高频响应特性进行协同,实现在线分离欧姆阻抗、电荷转移阻抗和扩散阻抗。此外,云边协同可以实现虚拟传感,即在云端运行的计算密集型电化学观测器会估算无法测量的内部状态,并将其传输到边缘端。将这些推断出的微观参数与可测量的宏观信号相融合,能够显著提高 SOH 估算对早期退化机制的敏感性,而这些早期退化机制在标准电压曲线中往往被掩盖。

3.4 实际工况部署

从实验室基准测试到实际部署的过渡中,要求算法具备极强的域适配性与泛化能力。由于电动汽车、储能电站等实际场景存在显著的域偏移,且电池很少经历完整充放电循环,后续研究应优先探索基于随机充电片段的特征重构技术。针对跨平台应用时面临的标签样本极度稀缺的问题,域适配技术^[70]与元学习、小样本学习范式^[71]将成为破局关键。通过在海量历史任务中提取具备跨体系泛化能力的元知识,模型有望在极少量目标样本下实现跨工况的快速精准适配。最后,算法还需构建能够量化不确定性的概率框架,以应对复杂工程环境下的非平稳特性。

4 结论

有限样本数据下的锂离子电池 SOH 准确估计,是直接关系到锂离子电池 SOH 估计研究真正落地的关键问题,更是解决当前锂离子电池健康管理技术推广难题的核心所在。通过对机器学习方法的分析发现,目前没有任何单一范式能为有限样本数据场景下的锂离子电池 SOH 估计难题提供有效的解决方案,作为主流的机器学习方法也存在泛化性不强以及物理约束性较弱等问题。最终,为了实现有限样本数据下准确的 SOH 估计需要实现从数据依赖到数据高效的转变。未来的范式突破应聚焦于协同自监督预训练技术释放无标

签运行数据的特征潜能、物理约束与机器学习方法的融合、利用多模态虚拟传感增强内部退化模式的观测性等,这些都是在目前实际工况下实现稳健、可迁移且可解释的生命周期管理的可行途径。

参考文献

- [1] Yan Jiajia, Jie Qiongnan. Mechanisms of policy intervention for China's transformation of the low-carbon economy[J]. *Journal of Cleaner Production*, 2025, 487: 144550.
- [2] Marri G K, Ee Y J, He Zhixiang, et al. Recent advancements in internal and external thermoregulation strategies for lithium-ion batteries[J]. *Renewable and Sustainable Energy Reviews*, 2026, 225: 116127.
- [3] 江银锋, 宋文祥, 施语. 基于随机充放电片段和稀疏高斯过程回归的电池健康状态估计[J]. *电工技术学报*, 2025, 40(19): 6359-6377.
Jiang Yinfeng, Song Wenxiang, Shi Yu. Battery state of health estimation using sparse Gaussian process regression based on random partial charge and discharge data[J]. *Transactions of China Electrotechnical Society*, 2025, 40(19): 6359-6377.
- [4] Zhao Xiaoyu, Wang Zuolu, Han Te, et al. A meta-learning method for few-shot multidomain state-of-health estimation of lithium-ion batteries[J]. *IEEE Transactions on Transportation Electrification*, 2025, 11(1): 4830-4840.
- [5] Wang Junping, Chen Quanshi, Cao Binggang. Support vector machine based battery model for electric vehicles[J]. *Energy Conversion and Management*, 2006, 47(7-8): 858-864.
- [6] Alzubaidi L, Zhang Jinglan, Humaidi A J, et al. Review of deep learning: concepts, CNN architectures, challenges, applications, future directions[J]. *Journal of Big Data*, 2021, 8(1): 53.
- [7] 吴起行. 基于改进型LSTM-RNN的锂离子电池SOH估计系统[J]. *电子制作*, 2021(9): 72-74.
- [8] 金帅, 董静. 基于数据驱动的锂离子电池健康状态估计研究进展综述[J]. *仪器仪表学报*, 2024, 45(3): 45-59.
Jin Shuai, Dong Jing. Review on progress of data-driven based health state estimation for lithium-ion batteries[J]. *Chinese Journal of Scientific Instrument*, 2024, 45(3): 45-59.
- [9] 陈宏宇, 陶志军, 朱永利, 等. 退役动力锂离子电池健康状态及剩余使用寿命预测技术研究[J]. *电力系统保护与控制*, 2025, 53(7): 174-187.
- Chen Hongyu, Tao Zhijun, Zhu Yongli, et al. Research on health state and remaining useful life prediction technology of retired power lithium-ion batteries[J]. *Power System Protection and Control*, 2025, 53(7): 174-187.
- [10] Vignesh S, Che Hangseng, Selvaraj J, et al. State of Health (SOH) estimation methods for second life lithium-ion battery: review and challenges[J]. *Applied Energy*, 2024, 369: 123542.
- [11] He Wenbin, Li Zongze, Liu Ting, et al. Research progress and application of deep learning in remaining useful life, state of health and battery thermal management of lithium batteries[J]. *Journal of Energy Storage*, 2023, 70: 107868.
- [12] Shu Xing, Shen Shiquan, Shen Jiangwei, et al. State of health prediction of lithium-ion batteries based on machine learning: advances and perspectives[J]. *iScience*, 2021, 24(11): 103265.
- [13] Vermeer W, Chandra Mouli G R, Bauer P. A comprehensive review on the characteristics and modeling of lithium-ion battery aging[J]. *IEEE Transactions on Transportation Electrification*, 2022, 8(2): 2205-2232.
- [14] Yarimca G, Cetkin E. Review of cell level battery (calendar and cycling) aging models: electric vehicles[J]. *Batteries*, 2024, 10(11): 374.
- [15] Chen B R, Walker C M, Kim S, et al. Battery aging mode identification across NMC compositions and designs using machine learning[J]. *Joule*, 2022, 6(12): 2776-2793.
- [16] Hawley W B, Parejiya A, Bai Yaocai, et al. Lithium and transition metal dissolution due to aqueous processing in lithium-ion battery cathode active materials[J]. *Journal of Power Sources*, 2020, 466: 228315.
- [17] Zhang Jienan, Li Qinghao, Wang Yi, et al. Dynamic evolution of cathode electrolyte interphase (CEI) on high voltage LiCoO₂ cathode and its interaction with Li anode[J]. *Energy Storage Materials*, 2018, 14: 1-7.
- [18] Tang Ming, Carter W C, Chiang Y M. Electrochemically driven phase transitions in insertion electrodes for lithium-ion batteries: examples in lithium metal phosphate olivines[J]. *Annual Review of Materials Research*, 2010, 40: 501-529.
- [19] Lin Feng, Zhao Kejie, Liu Yijin. Heterogeneous reaction activities and statistical characteristics of particle cracking in battery electrodes[J]. *ACS Energy Letters*, 2021, 6(11): 4065-4070.
- [20] Ren Dongsheng, Xie Leqiong, Wang Li, et al. A

- practical approach to predict volume deformation of lithium-ion batteries from crystal structure changes of electrode materials[J]. *International Journal of Energy Research*, 2021, 45 (5): 7732-7740.
- [21] Adenusi H, Chass G A, Passerini S, et al. Lithium batteries and the solid electrolyte interphase (SEI): progress and outlook[J]. *Advanced Energy Materials*, 2023, 13(10): 2203307.
- [22] Wang Shan, Ren Dongsheng, Xu Chengshan, et al. Lithium plating induced volume expansion overshoot of lithium-ion batteries: experimental analysis and modeling[J]. *Journal of Power Sources*, 2024, 593: 233946.
- [23] Han Xiang, Xu Min, Gu Lanhui, et al. Monolithic and conductive network and mechanical stress releasing layer on micron-silicon anode enabling high-energy solid-state battery[J]. *Rare Metals*, 2024, 43 (3): 1017-1029.
- [24] Campion C L, Li Wentao, Lucht B L. Thermal decomposition of LiPF₆-based electrolytes for lithium-ion batteries[J]. *Journal of the Electrochemical Society*, 2005, 152(12): A2327.
- [25] Dong Xiaoyu, Li Zhiwei, Ding Bing, et al. Electrolyte and electrode electrolyte interface for proton batteries: insights and challenges[J]. *ChemElectroChem*, 2024, 11 (2): e202300569.
- [26] Guo Liya, Thornton D B, Koronfel M A, et al. Degradation in lithium-ion battery current collectors[J]. *Journal of Physics: Energy*, 2021, 3(3): 032015.
- [27] Murali D R L, Banihashemi F, Lin J Y S. Zeolite membrane separators for fire-safe Li-ion batteries-effects of crystal shape and membrane pore structure[J]. *Journal of Membrane Science*, 2023, 680: 121743.
- [28] Chen Kai, Zhang Dian, Guo Jiaxin, et al. Hydrogen generated from binders: an overlooked thermal runaway source in lithium-ion batteries[J]. *Journal of Energy Chemistry*, 2025, 109: 602-608.
- [29] Xiong Rui, Wang Peng, Jia Yanbo, et al. Multi-factor aging in Lithium-ion phosphate batteries: mechanisms and insights[J]. *Applied Energy*, 2025, 382: 125250.
- [30] Saxena S, Xing Yinjiao, Kwon D, et al. Accelerated degradation model for C-rate loading of lithium-ion batteries[J]. *International Journal of Electrical Power & Energy Systems*, 2019, 107: 438-445.
- [31] Xiong Rui, Pan Yue, Shen Weixiang, et al. Lithium-ion battery aging mechanisms and diagnosis method for automotive applications: Recent advances and perspectives[J]. *Renewable and Sustainable Energy Reviews*, 2020, 131: 110048.
- [32] Wikner E, Thiringer T. Extending battery lifetime by avoiding high SOC[J]. *Applied Sciences*, 2018, 8(10): 1825.
- [33] Ni Yulong, Li Xiaoyu, Zhang He, et al. Online identification of knee point in conventional and accelerated aging lithium-ion batteries using linear regression and Bayesian inference methods[J]. *Applied Energy*, 2025, 388: 125646.
- [34] Lee P Y, Kwon S, Kang D, et al. Principle component analysis-based optimized feature extraction merged with nonlinear regression model for improved state-of-health prediction[J]. *Journal of Energy Storage*, 2022, 48: 104026.
- [35] Pohlmann S, Mashayekh A, Stroebl F, et al. State-of-Health prediction of lithium-ion batteries based on a low dimensional gaussian process regression [J]. *Journal of Energy Storage*, 2024, 88: 111649.
- [36] Yu Jianbo. State of health prediction of lithium-ion batteries: Multiscale logic regression and gaussian process regression ensemble[J]. *Reliability Engineering & System Safety*, 2018, 174: 82-95.
- [37] He Ye, Bai Wenyuan, Wang Lulu, et al. SOH estimation for lithium-ion batteries: an improved GPR optimization method based on the developed feature extraction[J]. *Journal of Energy Storage*, 2024, 83: 110678.
- [38] Pinkus A. Approximation theory of the MLP model in neural networks[J]. *Acta Numerica*, 1999, 8: 143-195.
- [39] Williard N, He Wei, Osterman M, et al. Comparative analysis of features for determining state of health in lithium-ion batteries[J]. *International Journal of Prognostics and Health Management*, 2013, 4(1): 1-7.
- [40] Wang Fang, Liu Shiqiang, Chen Shiqin, et al. SOH estimation for lithium-ion batteries using the distribution of relaxation time and feature optimized multi-layer perceptron[J]. *iScience*, 2025, 28(9): 113443.
- [41] Cortes C, Vapnik V. Support-vector networks[J]. *Machine Learning*, 1995, 20(3): 273-297.
- [42] Liu Shiyu, Fang Lide, Zhao Xiaoyu, et al. State-of-health estimation of lithium-ion batteries using a kernel support vector machine tuned by a new nonlinear gray wolf algorithm[J]. *Journal of Energy Storage*, 2024, 102: 114052.
- [43] Xiong Wei, Mo Yimin, Yan Cong. Online state-of-health estimation for second-use lithium-ion batteries based on weighted least squares support vec-

- tor machine[J]. IEEE Access, 2021, 9: 1870-1881.
- [44] Chemali E, Kollmeyer P J, Preindl M, et al. A convolutional neural network approach for estimation of Li-ion battery state of health from charge profiles[J]. Energies, 2022, 15(3): 1185.
- [45] Tian Jinpeng, Xiong Rui, Shen Weixiang, et al. Deep neural network battery charging curve prediction using 30 points collected in 10 min[J]. Joule, 2021, 5(6): 1521-1534.
- [46] Xie Yu, Luo Kai, Zheng Lihan, et al. A state of health estimation method for lithium-ion batteries based on initial charging segment and Gated Recurrent Unit neural network[J]. Journal of Power Sources, 2025, 638: 236607.
- [47] Dewi T, Mardiyati E N, Risma P, et al. Hybrid machine learning models for PV output prediction: harnessing random forest and LSTM-RNN for sustainable energy management in aquaponic system[J]. Energy Conversion and Management, 2025, 330: 119663.
- [48] Ackley D H, Hinton G E, Sejnowski T J. A learning algorithm for Boltzmann machines[J]. Cognitive Science, 1985, 9(1): 147-169.
- [49] Yann L C. Modeles connexionnistes de l'apprentissage[D]. Paris: Universite Pierre et Marie Curie, 1987: 114-143.
- [50] Rifai S, Vincent P, Muller X, et al. Contractive auto-encoders: explicit invariance during feature extraction[C]//Proceedings of the 28th International Conference on Machine Learning. Bellevue: Omni-press, 2011: 833-840.
- [51] Zheng Boyang, Ma Nanye, Tong Shengbang, et al. Diffusion transformers with representation autoencoders[EB/OL]. (2025-10-13) [2026-03-25]. <https://doi.org/10.48550/arXiv.2510.11690>.
- [52] He Xitian, Sun Bingxiang, Zhang Weige, et al. Inconsistency modeling of lithium-ion battery pack based on variational auto-encoder considering multi-parameter correlation[J]. Energy, 2023, 277: 127409.
- [53] Berahmand K, Daneshfar F, Salehi E S, et al. Autoencoders and their applications in machine learning: a survey[J]. Artificial Intelligence Review, 2024, 57(2): 28.
- [54] Vaswani A, Shazeer N, Parmar N, et al. Attention is all you need[C]//Advances in Neural Information Processing Systems. Long Beach, CA, USA: Curran Associates, 2017: 5998-6008.
- [55] Zhu Bo, Jia Li, Pan Quanke, et al. Cross-domain battery SOH and RUL estimation via domain-adaptive transformer[J]. Energy, 2025, 41(C): 139288.
- [56] Park S, Kim Y, Yoo R, et al. Transformer-based framework for configuration-agnostic Li-ion battery SOH estimation by exploring novel health indicators[J]. Chemical Engineering Journal, 2026, 528: 172197.
- [57] Khan S, Naseer M, Hayat M, et al. Transformers in vision: a survey[J]. ACM Computing Surveys, 2022, 54(10s): 1-41.
- [58] Sun Wenjie, Wu Chengke, Xie Chengde, et al. Fine-tuning enables state of health estimation for lithium-ion batteries via a time series foundation model[J]. Energy, 2025, 318: 134177.
- [59] Li Jiaqi, Fan Guodong, Zhang Xi. Hybrid end-to-end battery modeling and SOH estimation via physics-data fusion and maximum mean discrepancy minimization[J]. Energy, 2025, 340: 139253.
- [60] Zhu Yafei, Guo Tianyi, Li Xiang, et al. Domain generalization prognosis method for lithium-ion battery state of health with transformer and multi-kernel MMD[J]. Journal of Dynamics, Monitoring and Diagnostics, 2024, 3(4): 311-323.
- [61] Sun Baochen, Saenko K. Deep coral: correlation alignment for deep domain adaptation[C]//Computer Vision-ECCV 2016 Workshops. Cham: Springer, 2016: 443-450.
- [62] Gui Xuanang, Du Junrong, Wang Qianlong, et al. Multi-modal data information alignment based SOH estimation for lithium-ion batteries using a local global parallel CNN-Transformer Network[J]. Journal of Energy Storage, 2025, 129: 117178.
- [63] Tzeng E, Hoffman J, Saenko K, et al. Adversarial discriminative domain adaptation[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ, USA: IEEE, 2017: 7167-7176.
- [64] Goodfellow I, Pouget-Abadie J, Mirza M, et al. Generative adversarial networks[J]. Communications of the ACM, 2020, 63(11): 139-144.
- [65] Ye Zhuang, Yu Jianbo. State-of-health estimation for lithium-ion batteries using domain adversarial transfer learning[J]. IEEE Transactions on Power Electronics, 2022, 37(3): 3528-3543.
- [66] Hossen M I, Awrangzeb M, Pan Shirui, et al. Transfer learning in agriculture: a review[J]. Artificial Intelligence Review, 2025, 58(4): 97.
- [67] Che Yunhong, Zheng Yusheng, Sui Xin, et al. Supervised learning[J]. Journal of Energy Chemistry, 2023, 84: 335-346.
- [68] Wang Fujin, Zhai Zhi, Zhao Zhibin, et al. Physics-informed neural network for lithium-ion battery

- degradation stable modeling and prognosis[J]. Nature Communications, 2024, 15: 4332.
- [69] Huang Da, Qin Mengjiao, Liu Dan, et al. Multi-step ahead SOH prediction for vehicle batteries based on multimodal feature fusion and spatio-temporal attention neural network[J]. Journal of Energy Storage, 2025, 124: 116837.
- [70] Zhao Xiaoyu, Wang Zuolu, Miao Haiyan, et al. A label-free battery state of health estimation method based on adversarial multi-domain adaptation network and relaxation voltage[J]. Energy, 2024, 308: 132881.
- [71] Wang Runtong, Duan Bin, Zhao Guangcai, et al. Meta-learning driven small-sample lithium-ion battery state of health estimation based on Kolmogorov-Arnold network and long short-term memory network[J]. Journal of Energy Storage, 2026, 163: 122207.

Review of Machine Learning-Based Methods for Lithium-ion Battery State of Health Estimation with Limited Sample Data

LI Aihua, WANG Pengxiang, WANG Tao*,
XU Xiaodong, XIANG Jijun

(Rocket Force University of Engineering, Xi'an 710025, China)

ABSTRACT: Accurate state of health (SOH) estimation of lithium-ion batteries is a key technology for ensuring safe battery operation and health management. In practical applications, full-life-cycle battery experiments involve long durations and high labeling costs, whereas real-world operating data are often characterized by scarce labeled samples, sparse short-duration charging samples, and insufficient coverage of operating conditions. Therefore, limited sample data has become a major factor restricting the accuracy and generalization of machine learning-based SOH estimation. This study reviewed the current research status of machine learning-based SOH estimation methods with limited sample data, focusing on the performance differences of four categories of methods, including regression, shallow learning, deep learning, and transfer learning. The methods were compared in terms of their estimation accuracy, computational complexity, deployment requirements, and applicability. The results show that regression methods have advantages in probabilistic inference but suffer from high computational cost; shallow learning methods feature simple structures but are limited in their extrapolation capability; deep learning methods excel in feature extraction but rely heavily on the sample size; while transfer learning methods perform better under limited sample data, they still face problems such as negative transfer and training instability. Future research should focus on unlabeled data mining, physical constraint integration, multimodal sensing, and practical deployment to achieve robust, transferable, and interpretable SOH estimations for lithium-ion batteries.

KEYWORDS: lithium-ion battery; state of health estimation; machine learning; limited sample data; data efficiency

(编辑: 夏菁)