

面向锂电池少量循环的二维支持域直推式健康状态预测

王一航, 冯良骏, 赵春晖[†]

(浙江大学 控制科学与工程学院, 浙江 杭州 310027)

摘要: 锂离子电池的健康状态(SOH)是反映电池老化程度的关键指标,但由于电池老化的非线性和不确定性使得SOH难以精确估计,并且受到电池数据收集的高时间成本和容量再生现象的影响,传统的数据驱动方法在历史充放电循环数较少时效果较差. 针对上述问题,本文创新性地提出了一种二维支持域直推式学习(2D-RoSTL)建模思路,建立了数据由粗到细的精准划分方法,用于少量充放电循环下的SOH预测. 一方面,考虑同型号多块电池的批次特性,利用历史数据和批次数据构造二维支持域扩充模型信息来源,提供了粗范围的大量可供选择的样本;另一方面,首次尝试以直推式学习的方式解决SOH预测的任务,利用离线和在线样本特征空间的信息,对每个样本进行细致划分,提升少量充放电循环情况下模型的预测可靠性. 基于NASA的公开数据集,所提出的二维支持域直推式建模方法在4个电池上的预测误差均小于1.56%,实现了对锂电池充放电历史初期及再生点的精确预测.

关键词: 锂电池健康状态; 少量充放电循环; 二维支持域; 直推式学习

引用格式: 王一航, 冯良骏, 赵春晖. 面向锂电池少量循环的二维支持域直推式健康状态预测. 控制理论与应用, 2024, 41(3): 474 – 483

DOI: 10.7641/CTA.2023.20408

Lithium battery two-dimensional region of support transductive learning and state of health prediction oriented to few charge-discharge cycles

WANG Yi-hang, FENG Liang-jun, ZHAO Chun-hui[†]

(College of Control Science and Engineering, Zhejiang University, Hangzhou Zhejiang 310027, China)

Abstract: The state of health (SOH) of lithium-ion batteries is a key indicator reflecting the degree of battery aging. However, due to the nonlinearity and uncertainty of battery aging, it is challenging to estimate SOH accurately. Besides, affected by the high time cost of battery data collection and the phenomenon of capacity regeneration, the traditional data driving method is less effective when the number of historical charge-discharge cycles is small. To solve the above problems, this paper innovatively proposes a two-dimensional region of support transductive learning (2D-RoSTL) method, and establishes a precise data division method from coarse to fine, which is used for SOH prediction oriented to few charge-discharge cycles. On the one hand, considering the batch characteristics of the same model block battery, using historical data and batch data to construct a two-dimensional region of support to expand the information source of the model, providing an extensive range of samples for selection. On the other hand, we first attempt to solve the SOH prediction task by transductive learning method. Using offline and online information in the sample feature space, our model finely divides each sample to improve the online prediction reliability in the case of a few charge-discharge cycles. Based on the NASA's public data set, the proposed two-dimensional region of support transductive modeling method has a prediction error of less than 1.56% on the four batteries, and has realized the accurate prediction of the early history of lithium batteries and the point of regeneration.

Key words: lithium battery SOH; few charge-discharge cycles; 2-D region of support; transductive learning

Citation: WANG Yihang, FENG Liangjun, ZHAO Chunhui. Lithium battery two-dimensional region of support transductive learning and state of health prediction oriented to few charge-discharge cycles. *Control Theory & Applications*, 2024, 41(3): 474 – 483

收稿日期: 2022-05-18; 录用日期: 2023-03-31.

[†]通信作者. E-mail: chhzhao@zju.edu.cn; Tel.: +86 571-87951879.

本文责任编辑: 周平.

国家杰出青年科学基金项目(62125306), 国家自然科学基金重点项目(62133003), 工业控制技术国家重点实验室自主课题项目(ICT2021A15)资助. Supported by the National Science Fund for Distinguished Young Scholars (62125306), the National Natural Science Foundation of China (62133003) and the Research Project of the State Key Laboratory of Industrial Control Technology, Zhejiang University, China (ICT2021A15).

1 引言

近年来, 锂离子电池因其高能量密度、低自放电率、高电压、长寿命周期以及更高的可靠性被广泛使用^[1], 但同时, 其带来的安全问题不容小觑^[2]. 作为保障安全性的主要手段之一, 锂电池的健康状态(state of health, SOH)管理受到越来越多的关注. SOH作为一种评估电池老化程度的指标, 反映了电池实际容量与标称容量的衰减比率, 由式(1)表示, 其中 C_{act} 和 C_{norm} 分别代表电池的实际容量和标称容量^[3].

$$\text{SOH} = \frac{C_{\text{act}}}{C_{\text{norm}}}, \quad (1)$$

对于一个确定的电池, 其 C_{norm} 是确定的, 故对SOH的预测可以转化为对 C_{act} 的预测. 一个全新电池的SOH值为100%. 随着电池的不断使用, SOH会逐渐下降. 通常情况下, 当电池的容量下降到初始值的80%, 即SOH为80%时, 视为寿命终止(end of life, EOL), 应该更换电池^[4].

一般情况下, 准确的SOH只能在实验室条件下直接测量, 实际应用中只能通过其他的变量例如电压电流进行估算. 而由于电池老化的非线性和不确定性, SOH的预测变得十分困难^[5]. 为解决此问题, 近年来已经提出了很多预测SOH的方法, 主要可以分为两类, 机理分析法和数据驱动法^[6].

机理分析法使用微分方程等方式来模拟电池运行状态下的电池老化过程, 从而对SOH进行估计. Weng等^[7]借助电化学模型和恒定电流-电压充电方法, 使用变换函数和非线性最小二乘法对SOH进行预测. 该模型结合了化学/电动力学和线性传输关系, 以产生比经验模型更准确的预测. Li等^[8]使用单粒子模型(single particle model, SPM)对SOH进行预测, 该模型将阳极和阴极分别视为具有与电极相同的表面积的单颗粒. 卡尔曼滤波器(Kalman filter, KF)是一种自适应方法, 也经常被用于SOH估计. Wassiliadis等^[9]使用扩展卡尔曼滤波器进行SOH预测. 卡尔曼滤波器在电池管理系统(battery management system, BMS)中应用非常广泛, 其还被用于估计电荷状态(state of charge, SOC)、容量或者内部电阻^[10]. 然而, 基于微分方程的模型计算开销往往较大, 且对专业知识的要求较高, 很难精确地模拟整个老化进程.

数据驱动法通过利用历史充放电循环数据建立电池外部特征和SOH值的映射关系来预测SOH. 该类方法不要求使用者了解内部的复杂机理, 并且模型也可以在线应用^[11]. 因此, 近年来数据驱动的方法在学术界和工业界都越来越被广泛关注. 例如, You等^[12]使用神经网络方法预测SOH, 平均误差小于2.18%, Klass等^[13]提出了一种基于车载可用电池数据的SOH估计方法, 使用支持向量机方法进行预测. Li等^[14]通过

电池充电曲线提取了特征, 并且使用增量容量分析对特征进行选择, 最后使用随机森林方法进行SOH预测. 赵光财等^[15]使用贝叶斯正则化神经网络对锂电池SOH进行预测, 同时引入 K 临近-马尔科夫修正策略对预测结果进行修正. 组合预测模型对锂电池SOH的平均预测误差小于1%.

尽管很多的数据驱动方法已在SOH评估任务中得到了广泛关注, 但是该类方法需要依靠大量的历史数据进行训练. 而在实际场景中, 当历史充放电循环数据积累较少时, 往往已经迫切地需要给出后续的SOH预测, 但是锂电池的一次充放电时间往往需要花费数小时甚至数天, 如果希望收集足够的电池老化数据进行建模, 无疑需要巨大的时间成本^[16]. 此外, 锂电池在充放电循环中还会发生容量再生现象, 容量再生现象是锂离子电池退化过程中的一个典型特征, 如果电池在老化过程中的静置时间比正常的静置时间长, 则会开启电池的恢复过程, 从而可以增加下一个周期的可用容量, 导致模型的预测不可靠^[17]. 基于以上分析可以发现, 在真实场景中对锂电池进行SOH预测, 会受到典型的少数据问题的影响. 如何在少量充放电循环数据的情况下巧妙设计建模分析策略, 确保模型的预测精度, 以及如何充分利用训练样本的信息, 针对待预测样本有效进行建模, 是笔者迫切需要考虑的问题, 具有重要的理论研究价值和现实意义.

近年来, 随着一些迁移学习方法例如域自适应神经网络(domain adaptive neural network, DaNN)^[18]、细粒度域自适应对抗网络(fine-grained adversarial network-based domain adaptation, FANDA)^[19]、相关性对齐(correlation alignment, CORAL)^[20]和深度适配网络(deep adaptation network, DAN)^[21]的提出, 迁移学习算法已被广泛用于不同工业领域中. 例如, 在故障诊断方面, Chai等^[22]提出了一种故障原型适配网络用于工业智能故障诊断, 并利用多源域细化的对抗性适应策略解决域不一致和类别不一致下的故障诊断问题^[23]. 近期, 迁移学习算法也逐渐被用于锂电池SOH的少量数据预测问题. 例如, Shen等^[24]使用来自8个电池的源域数据训练了一个集成CNN模型, 再根据目标数据集对模型进行微调, Ye等^[25]使用对抗学习的方法使得源域和目标域通过特征提取器提取到的特征保持一致, 最后使用少量目标域的数据对模型进行微调, Han等^[26]提出了一种端到端的考虑域适应的增强长短期记忆(long short term memory, LSTM)迁移方法, 采用了数据增强的方法使得源域和目标域的数据量相当, 进而用最大平均差异(maximum mean discrepancy, MMD)约束实现域特征对齐. 上述方法的主体思路均是使用其他工况的源域数据进行预训练, 再利用目标域工况的数据进行迁移, 没有考虑利用同一工况下已有电池的信息, 在实际中不同工况的数据差异

极大, 电流、温度, 甚至充放电方式都可能有所改变, 导致模型的迁移困难. 在这种情况下, 如果能够利用已有的同一工况下其他电池的数据作为扩充, 根据待预测电池的特性对数据进行选择, 则会有充足的数据进行建模, 实现对SOH的精准预测.

针对上述问题, 本文提出了一种二维支持域直推式建模方法(two-dimensional region of support transductive learning, 2D-RoSTL), 通过对数据由粗到细的精确划分及模型搭建, 实现了少量充放电循环下SOH的准确预测. 一方面, 利用同一工况其他电池的相似特性作为数据扩充, 用批次数据和历史数据在两个维度上构造支持域, 对数据进行粗划分, 提供了充足可供选择的样本; 另一方面, 提出了一种直推式学习框架, 考虑离线样本和在线样本特征空间的相似性信息, 实现了少量循环数据下样本的精细化充分利用. 并且, 本文提出的方法不拘泥于特定数据, 是一种通用的结构范式, 其可以适用于任何具有批次数据的少量循环SOH预测场景中. 本文的贡献点总结如下:

1) 提出了一种针对少量充放电数据的二维支持域构造范式, 能够解决历史数据不充分的情况下模型过拟合问题以及容量再生现象导致的预测不可靠问题, 其思想方法为该领域典型少量数据问题的解决提供了新的研究思路;

2) 提出了一种关注样本特征空间内相似性信息度量的直推式框架, 扩充模型信息来源, 细化了样本的重要程度, 保证数据信息的充分利用, 从而实现了少量数据下SOH的精确预测.

本文结构如下: 第2节首先介绍了少量数据场景下的建模问题, 包括使用的数据以及待解决问题的阐述; 第3节介绍了本文提出的2D-RoSTL方法, 包括二维支持域的构造、特征提取以及直推式建模策略; 第4节通过实验证明了所提二维支持域直推建模方法的有效性; 第5节进行了总结与讨论.

2 少量数据下的建模问题

在本节中, 首先介绍了文中使用的4个NASA的电池数据, 进而从数据出发阐述了少量循环数据情况下的建模问题.

2.1 数据介绍

本文中使用的数据集来自NASA的爱达荷州国家实验室^[27], 使用的电池编号为No.5, No.6, No.7和No.18. 此组电池均在室温24°C条件下循环进行充电、放电和阻抗测试. 在充电过程中, 电池首先在恒流(constant current, CC)模式下以1.5 A电流进行充电, 直至电压上升至4.2 V, 然后进入恒压(constant voltage, CV)模式进行充电直至充电电流下降至20 mA. 放电过程包含CC模式, 以2 A恒定电流的方式进行放电, 直至电压下降至2.7 V, 2.5 V, 2.2 V和2.5 V.

图1展示了4个电池的老化曲线, 从图中可以看出, 电池的退化能力随着循环次数呈现不稳定的下降趋势, 并且在循环过程中还有一定的容量再生现象. 由于电池容量和电池循环次数之间的关系为非线性, 这就使得电池SOH的预测变得更加困难. 图2展示了No.6电池第10次充放电循环过程中的电压、电流和温度曲线, 对于每一次充放电循环, 采样点的个数都不尽相同, 如何选取合适的特征对SOH进行预测, 也是需要面临的一大挑战.

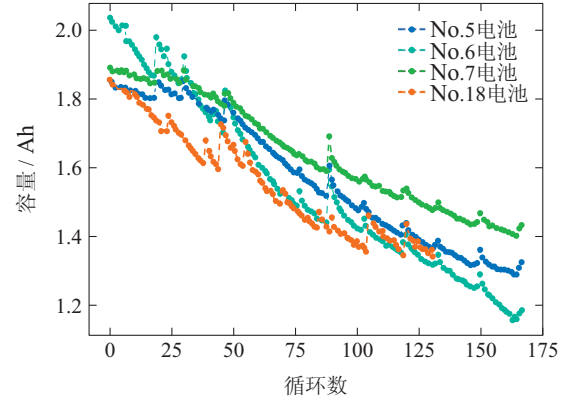


图1 4个锂电池的老化曲线

Fig. 1 Aging curves of four lithium batteries

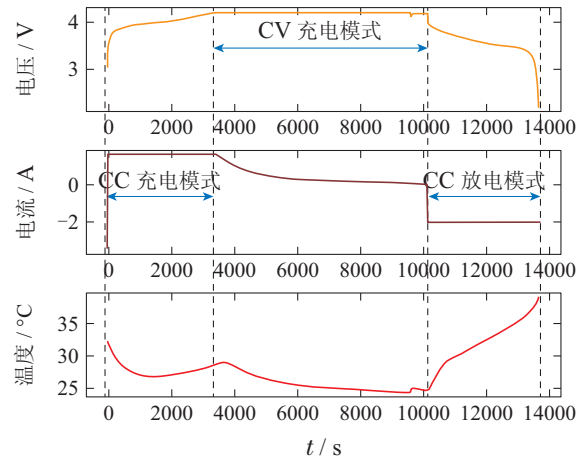


图2 No.6 电池第10次充放电循环电压、电流、温度曲线

Fig. 2 No.6 battery voltage, current, temperature curve during the tenth charge and discharge cycle

2.2 少量循环数据建模问题

对于样本的SOH值, 本文通过电压、电流、温度3个传感器的值对其进行估计. 在本文中用 y_i^j 表示第 j 个电池第 i 次循环的SOH值(后续均使用上标表示电池, 下标表示循环次数), 其中: $i \in \{1, 2, \dots, N\}$, N 为总循环数, $j \in \{1, 2, 3, 4\}$. V_i^j , I_i^j , T_i^j 分别表示第 i 次循环下电压、电流、温度曲线, $j \in \{1, 2, 3, 4\}$. V_i^j , I_i^j , $T_i^j \in \mathbb{R}^t$. t 代表该循环下采样点的个数, 三者均为时间的函数. 用 $F(\cdot)$ 表示整个预测框架, 则有

$$y_i^j = F(V_i^j, I_i^j, T_i^j). \quad (2)$$

注意到, V_i^j, I_i^j, T_i^j 均为 t 维, 如之前图2中所示, 一次充放电循环的采样点往往可以达到上万个, 直接将所有的维度作为输入会导致输入维度过高, 并且每一次循环 t 的长度也并不相同, 进而导致输入维度发生变化, 为模型的建立又增加了难度. 为了解决这一问题, 本文对曲线进行变换处理, 从变换后的曲线上提取一些和SOH值相关的特征点作为输入, 用 $G(\cdot)$ 表示变换以及提取的函数, x_i^j 表示输入特征点, $y_i^j \in \mathbb{R}^m$, m 表示特征点的维度, 则有

$$x_i^j = G(V_i^j(t), I_i^j(t), T_i^j(t)). \quad (3)$$

用 $\mathcal{D}$ 表示 x_i^j 到 y_i^j 的预测框架, 即

$$y_i^j = \mathcal{D}(x_i^j). \quad (4)$$

传统方法建立模型时使用的训练数据为之前积累的历史数据, 设之前共积累了 k 轮充放电循环的历史数据, 用 X_{history}^j 和 Y_{history}^j 分别表示历史数据的特征和标签, 有

$$X_{\text{history}}^j = \{x_1^j, x_2^j, \dots, x_k^j\}, \quad (5)$$

$$Y_{\text{history}}^j = \{y_1^j, y_2^j, \dots, y_k^j\}, \quad (6)$$

$$Y_{\text{history}}^j = \mathcal{D}(X_{\text{history}}^j). \quad (7)$$

在模型建立完成后, 如果希望预测第 n 轮充放电循环的SOH ($k < n \leq N$), 需要将特征点 x_n^j 输入模型, 则可以得到第 n 轮充放电循环后的SOH预测值, 其示意图如图3所示.

3 二维支持域直推式(2D-RoSTL)建模方法

在本节中, 首先介绍了本文所提出模型的整体结

构, 然后依次详细阐述了模型的各个部分, 包括二维支持域的构建方法、通过传感器信号提取输入特征的手段和直推式建模以及在线应用的过程.

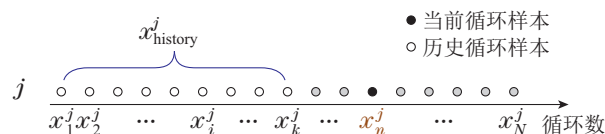


图3 第 n 轮充放电循环训练数据示意图

Fig. 3 Schematic diagram of training data for the n th round of charge-discharge cycles

3.1 整体结构

模型的整体结构如图4所示, 首先, 使用历史数据和批次数据在两个不同的维度上构造支持域, 对数据进行粗划分. 然后, 将所选数据进行特征提取, 通过滤波后的曲线提取了6个特征点. 最后, 直推式建模利用支持域中提供的数据, 使用离线和在线样本点特征空间的相似性信息, 对每个训练样本点再进行细划分, 赋予每个样本不同的权重. 模型建立完成后, 预测测试样本点的容量, 并计算SOH值. 二维支持域的构建为后续建模提供了充足的数据, 特征分析模块提取的特征点为模型提供了良好的输入支撑, 而直推式建模利用了新样本点特征空间的信息, 保证循环数据较少情况下信息的充分利用. 整体结构由粗到细, 模块之间相辅相成, 共同完成SOH预测模型的搭建, 以实现

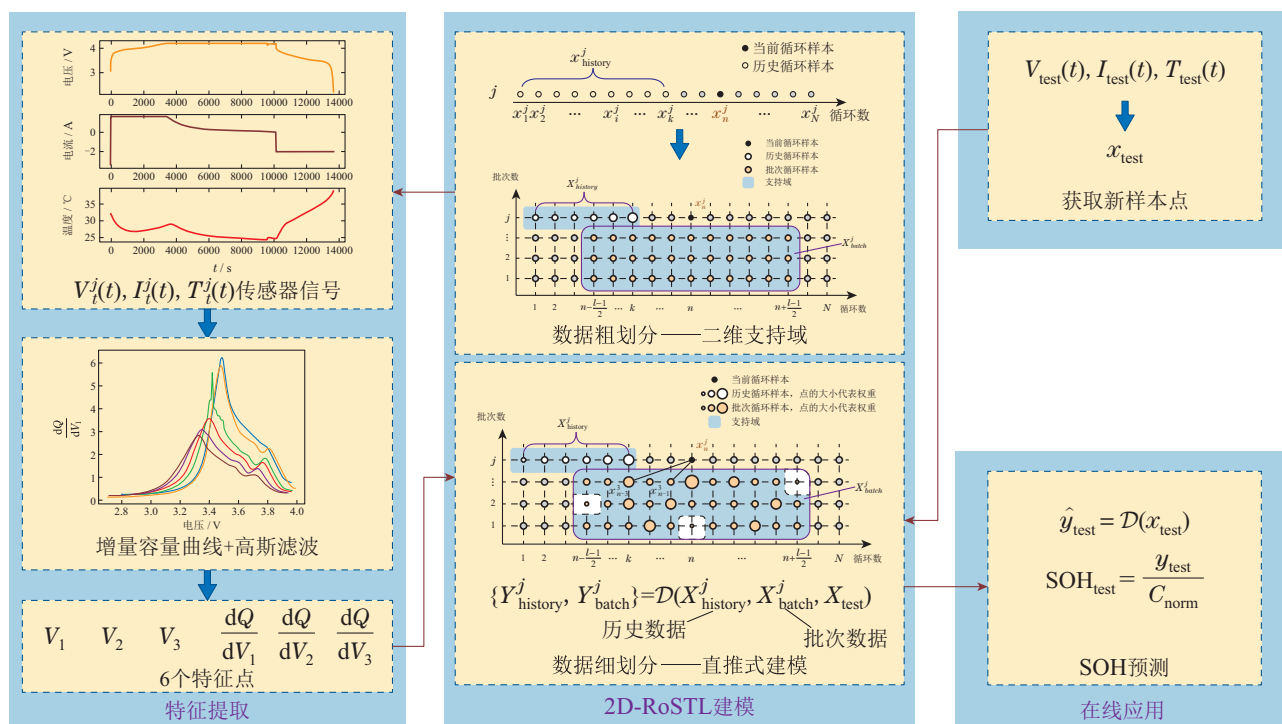


图4 建模应用流程图, 整体由特征提取、2D-RoSTL建模和在线应用3部分构成

Fig. 4 The modeling application flow chart is composed of three parts: feature extraction, 2D-RoSTL modeling and online application

3.2 二维支持域的构建

在少量充放电循环数据的条件下, 如何避免模型的过拟合及预测不可靠, 是笔者迫切要考虑的问题. 本文借鉴了工业过程中的批次思想^[28-31], 将其首次创新性引入到SOH研究中.

本文使用 x_n^j 表示当前第 j 个电池的的第 n 轮充放电循环数据的特征, 用 X_{history}^j 和 Y_{history}^j 分别表示历史数据的特征和标签. 此外, 本文还通过增加批次数据来扩充支持域, 使用 X_{batch}^j 和 Y_{batch}^j 来表示增加的批次循环数据的特征和标签. 为了更好地利用批次数据, 本文在当前循环周围通过滑窗进行选取, 数据窗口中不仅包括批次电池当前历史循环之前的数据, 也同时包括批次电池当前历史循环之后的数据, 通过这样选取为模型补充了更多有用的信息, 使其能够更好地应对未知的情况. 滑窗长度为 l , 由于滑窗是同时向两侧进行选取数据, 两侧各选取 $(l-1)/2$ 个样本, 故 l 应为

奇数. 有

$$X_{\text{batch}}^j = \{x_{n-\frac{l-1}{2}}^r, \dots, x_n^r, x_{n+1}^r, \dots, x_{n+\frac{l-1}{2}}^r\}, \quad (8)$$

$$Y_{\text{batch}}^j = \{y_{n-\frac{l-1}{2}}^r, \dots, y_n^r, y_{n+1}^r, \dots, y_{n+\frac{l-1}{2}}^r\}, \quad (9)$$

其中上标 r 表示除当前电池外其他的电池编号, $r \in \{1, 2, 3, 4\}$, 模型变为

$$\{Y_{\text{history}}^j, Y_{\text{batch}}^j\} = \mathcal{D}(X_{\text{history}}^j, X_{\text{batch}}^j). \quad (10)$$

图5展示了二维支持域的构建图, 其由两部分构成, 横向的历史循环数据和纵向的批次数据, 其中两部分张成的空间构成了数据矩阵. 相比传统建模方法仅使用历史数据 X_{history}^j 和 Y_{history}^j 建模, 引入批次思想构造建模数据不仅仅增加了数据量, 解决了模型训练数据不足的问题, 更是从另一个维度上为模型提供了其他电池的参考信息, 使得模型学习到更多不同情况下的知识, 增加模型的鲁棒性以及泛化性能.

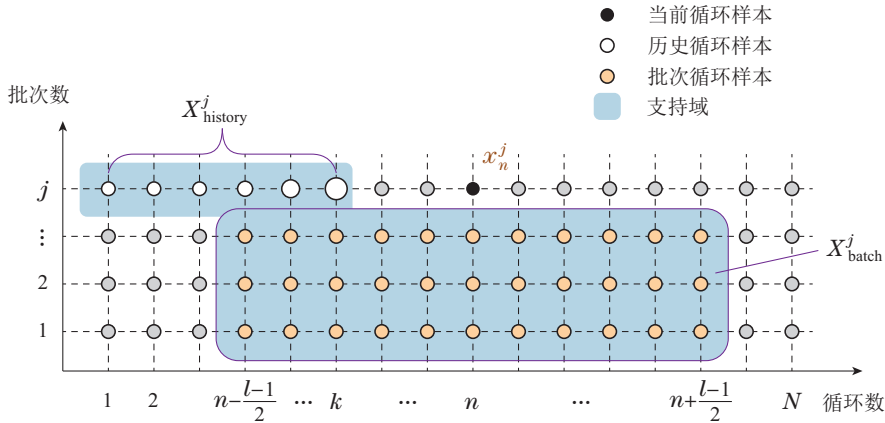


图5 二维支持域架构图

Fig. 5 2D region of support structure diagram

3.3 特征提取

通过构造二维支持域补充数据后, 还需要对数据进行处理. 每一次充放电循环都会产生3条 V, I, T 曲线, 考虑到电池老化的非线性, 需要从曲线中提取和电池容量密切相关的特征作为输入进行预测.

增量容量(incremental capacity, IC)曲线^[32]是近些年被用于提取SOH特征的重要手段, 其纵坐标是 $\frac{dQ}{dV}$, 横坐标是电压值, 可由下式计算:

$$\frac{dQ}{dV} = \frac{dI \times t}{dV} = \frac{I \times dt}{dV}. \quad (11)$$

由于IC曲线涉及到微分计算, 曲线中包含非常多的噪声信号, 噪声会影响本文对特征的提取, 所以, 如何去除噪声就变得尤为关键. 比较常用的方法有SG (Savitzky-Golay)滤波法^[33]和高斯滤波(Gaussian smoothing, GS)法^[34]. 图6比较了No.6电池第30个循环IC曲线由SG滤波器和GS滤波器滤波的效果. 可以看出, GS滤波器的滤波效果更加优秀, 其可以获得干净的峰

值信号. 故在本文中使用GS滤波器进行滤波, 其可以表示为

$$G(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(x-\mu)^2}{2\sigma^2}}, \quad (12)$$

其中: μ 是均值, σ 是标准差. 使用GS滤波器滤波之后, 从曲线上提取了6个特征点, 分别是两个波峰和一个波谷的电压值和 $\frac{dQ}{dV}$ 值, 使用这6个特征对SOH进行预测, 需要注意的是, 均值和标准差的选择必须谨慎, 如果过平滑可能使得曲线的波峰和波谷消失, 如果平滑不足可能使得波峰和波谷增多.

3.4 基于特征相似性的直推式建模

二维支持域的构建相当于是对整体数据做了一次粗划分, 为模型的建立提供了充足的样本, 但是由于模型中包含历史和批次两个维度的数据, 数据的地位并不等同, 并且在同一个维度内部, 例如建模的历史数据, 一个样本和一个样本之间的重要性也不是完全

相等的. 如果将全部的数据给予相同地位进行建模, 无疑会损失很多信息. 本文这里利用在线样本和离线样本特征空间的相似性, 扩充信息进行直推式建模, 为每一个样本赋予一个独立的权重, 细化每一个样本的重要程度, 整体由粗到细, 实现模型的精细化搭建和数据的充分利用. 用 x_{test} 表示在线样本, 模型变为

$$\{Y_{\text{history}}^j, Y_{\text{batch}}^j\} = \mathcal{D}(X_{\text{history}}^j, X_{\text{batch}}^j, x_{\text{test}}). \quad (13)$$

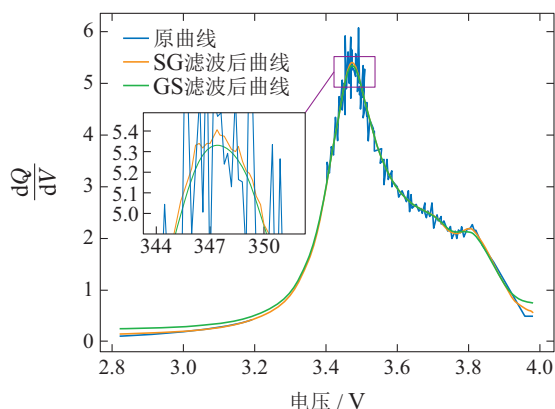


图6 No.6 电池第30个循环IC曲线SG滤波器和GS滤波器滤波效果

Fig. 6 No.6 battery 30th cycle IC curve SG filter and GS filter filtering effect

在SOH预测中, 希望更多的关注那些对预测更有帮助的信息, 更多的忽略那些可能携带噪声的信息. 所以, 这里采用特征空间中相似性度量的手段来判断训练样本的重要性, 和 x_{test} 更加相近的样本认为其更

加重要. 样本权重通过高斯核函数进行计算, 使用下标 q 代表训练集中的第 q 个样本, 用 w 表示样本的权重

$$w_q = e^{-\frac{(x_q - x_{\text{test}})^2}{2\tau^2}}, \quad (14)$$

其中: $x_q \in \{X_{\text{history}}^j, X_{\text{batch}}^j\}$, 代表训练集的第 q 个样本特征; w_q 代表该样本的权重; τ 为超参数.

图7展示了基于特征相似性的直推式建模方法, 其在构造二维支持域的基础上(图5所示)细化了支持域的范围, 并且增加了样本权重. 需要注意的是, 这里的远近并不是指的距离的远近, 而是通过式(14)使用特征之间的相似性进行度量, 循环相近的样本之间特征也有可能相差很远, 例如图中 x_{n-3}^3 和 x_{n-1}^3 两点权重不同, 循环更远的点 x_{n-3}^3 反而权重更大一些. 特别指出, 可通过事先定义一个衡量权重大小的阈值. 如果权重小到对建模没有影响, 即低于该阈值, 则这些数据点将会被忽略, 从而支持域的范围也会发生改变, 变为一个不规则的形状, 从图中可以看出, 支持域中已将权重小的点除去(如虚线部分所示), 这些点一般为批次数据中偏远的点.

由于提取到的特征和真实的SOH值相关性很高, 故本文建模方法采用的是最小二乘法, 损失函数取真实值和预测值的二范数平方, 有

$$\mathcal{L} = \frac{1}{2(k+3l)} \sum_{q=1}^{k+3l} w_q \times (y_q - \hat{y}_q)^2, \quad (15)$$

其中: $y_q \in \{Y_{\text{history}}^j, Y_{\text{batch}}^j\}$, 代表训练集中第 q 个样本的真实标签; $\hat{y}_q$ 代表该样本的预测值; $k+3l$ 是训练集全部的样本总数.

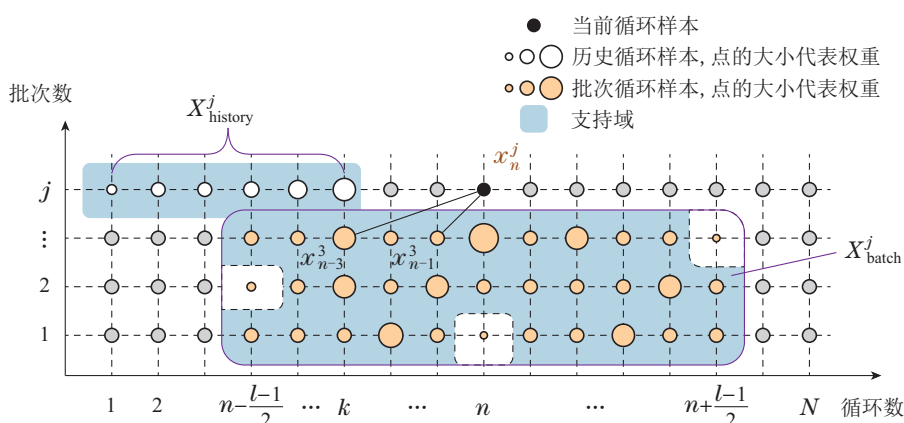


图7 基于特征相似性的直推式建模

Fig. 7 Transductive modeling based on feature similarity

3.5 在线应用

2D-RoSTL算法应用的整体过程如算法1所示, 当一次充放电循环完成后, 利用滑窗选取二维支持域, 得到批次数据, 加上原有的少量历史数据, 经过特征提取和特征相似度权重计算后建立最小二乘模型 $\mathcal{D}$, 最后将待预测样本进行特征提取后输入模型, 即可得

到预测的SOH值. 需要说明的是, 由于最小二乘模型建模及预测的时间为秒级, 故在线进行建模完全可以保证预测的及时性和准确性.

算法1 2D-RoSTL算法.

输入: 历史及其他电池充放电循环传感器信号 V, I, T , 每次循环所对应的SOH值 y , 待预测传感器信

号 $V_{\text{test}}, I_{\text{test}}, T_{\text{test}}$.

输出: 预测的SOH值 y_{test} .

步骤 1 通过二维支持域选取数据, 根据式(8)–(9)截取待预测循环前后其他电池的批次数据;

步骤 2 将选取的数据和该电池之前的历史数据使用增量容量曲线和高斯滤波进行特征提取, 得到 $X_{\text{history}}^j, Y_{\text{history}}^j, X_{\text{batch}}^j, Y_{\text{batch}}^j$ 用于训练;

步骤 3 将待测传感器信号进行特征提取, 得到 x_{test} ;

步骤 4 使用式(14)计算训练集中每个样本的权重;

步骤 5 建模, 最小化式(15)中的损失函数 $\mathcal{L}$, 模型收敛或迭代超过最大次数时停止迭代, 得到模型 $\mathcal{D}$;

步骤 6 通过模型 $\mathcal{D}$ 预测 x_{test} 对应的 SOH 值 y_{test} .

4 实验与分析

4.1 数据集及评价标准

本节利用NASA的开源数据^[27]对2D-RoSTL算法的有效性进行验证, 使用的电池编号为No.5, No.6, No.7和No.18, 具体充放电细节如第1.1节所述. 由于此组电池的实验条件均相同, 故将其视为同一工况下不同批次的数据. 每一次对比实验共进行4组, 每组选择一个电池作为测试电池(后续表中电池编号均指测试电池编号), 取其前20%的历史数据作为训练, 后80%的数据作为测试, 其余电池作为批次数据的补充, 共同进行2D-RoSTL建模预测, 其实验数据集划分如表1所示.

表 1 实验数据集划分

Table 1 Experimental dataset partition

实验	测试电池编号	其余批次电池编号
I	No.5	No.6, No.7, No.18
II	No.6	No.5, No.7, No.18
III	No.7	No.5, No.6, No.18
IV	No.18	No.5, No.6, No.7

实验评价标准选用均方根误差 (root mean square error, RMSE)和平均绝对误差(mean absolute error, MAE). RMSE可以衡量观测值同真值之间的偏差, 而MAE能更好地反映预测值误差的实际情况, 二者可以用下式表示:

$$\text{RMSE} = \sqrt{\frac{1}{M} \sum_{i=1}^M (y_i - \hat{y}_i)^2}, \quad (16)$$

$$\text{MAE} = \frac{1}{M} \sum_{i=1}^M |y_i - \hat{y}_i|, \quad (17)$$

其中: y_i 代表真实值, $\hat{y}_i$ 代表预测值, M 代表总预测样本数.

4.2 对比回归算法

为了验证所提2D-RoSTL算法的有效性, 首先, 将其与一些直接建模模型作对比实验, 选取了常用的线性回归(linear regression, LR)、支持向量回归(support vector regression, SVR)以及随机森林(random forest, RF)3种回归算法, 它们均已被广泛用于锂电池SOH预测^[16–17]. 实验中回归算法均使用测试电池前20%的历史数据和其余批次电池的全部数据集合作为训练, 而2D-RoSTL则是采用划分的支持域中的数据作为训练. 表2展示了几种算法对SOH的预测精度(均值最优用粗体表示), 2D-RoSTL在No.5和No.6电池上的预测有着最优的结果, 并且在4个电池平均的精度上也保持最优. 值得注意的是, 这里滑窗的长度仅为11, 相比较每个电池168次循环的数据长度, 建立支持域所使用的批次数据不足整体的10%, 但其效果却优于使用全部批次数据建模预测的其余回归算法, 充分体现了本文所提方法的优越性.

表 2 对比直接建模的回归算法预测性能

Table 2 Comparing the predictive performance with regression algorithms for direct modeling

性能指标	电池编号	LR	SVR	RF	2D-RoSTL
RMSE/%	No.5	1.37	1.69	1.25	0.92
	No.6	2.52	3.85	2.18	1.50
	No.7	1.27	1.02	1.58	1.56
	No.18	1.11	1.08	0.88	1.04
	均值	1.57	1.91	1.47	1.26
MAE/%	No.5	1.07	1.34	0.99	0.75
	No.6	2.26	3.06	1.55	1.17
	No.7	0.83	0.70	1.22	0.96
	No.18	0.80	0.94	0.71	0.78
	均值	1.24	1.51	1.12	0.92

图8展示了回归算法和2D-RoSTL在No.5电池上的SOH预测效果, 从图中可以看出, 2D-RoSTL的预测值(紫色)和真实的SOH值(蓝色)最相近, 其预测结果明显优于其余的回归算法.

4.3 对比迁移学习算法及支持域可视化

为了证明2D-RoSTL在少量数据方法中的优越性, 将所提算法与DaNN, CORAL和DAN等迁移学习方法进行对比实验, 迁移学习方法的源域为其他批次电池的全部数据, 目标域为测试电池的前20%历史数据.

表3展示了迁移学习方法和2D-RoSTL的预测性能, 在4组电池实验共10个RMSE和MAE指标中, 所提方法有7次均为最优, 其余性能也与最优结果相近, 且平均精度也为最优(RMSE均值为1.26%, MAE均值

为0.92%, 而其余迁移学习算法最优RMSE均值为1.74%, 最优MAE均值为1.47%。

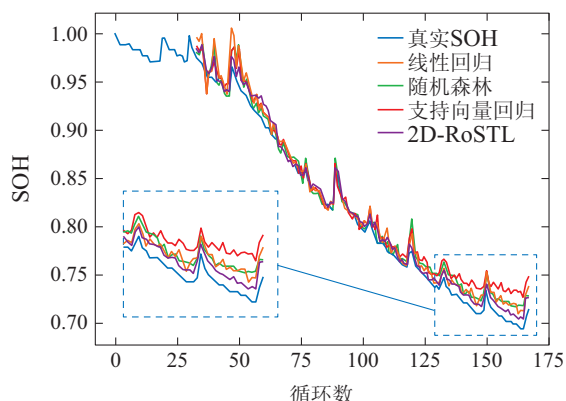


图8 回归算法和2D-RoSTL算法在No.5电池上的SOH预测效果

Fig. 8 The SOH prediction effect of the regression algorithm and the 2D-RoSTL algorithm on the No.5 battery

表3 对比迁移学习算法预测性能

Table 3 Comparing the predictive performance with transfer learning algorithms

性能指标	电池编号	DaNN	CORAL	DAN	2D-RoSTL
RMSE/%	No.5	2.99	3.44	1.63	0.92
	No.6	4.46	11.95	3.23	1.50
	No.7	1.50	7.40	1.27	1.56
	No.18	0.66	7.42	0.82	1.04
	均值	2.40	7.55	1.74	1.26
MAE/%	No.5	2.57	3.27	1.26	0.75
	No.6	4.24	11.74	3.09	1.17
	No.7	1.10	7.20	0.96	0.96
	No.18	0.50	7.19	0.55	0.78
	均值	2.10	7.35	1.47	0.92

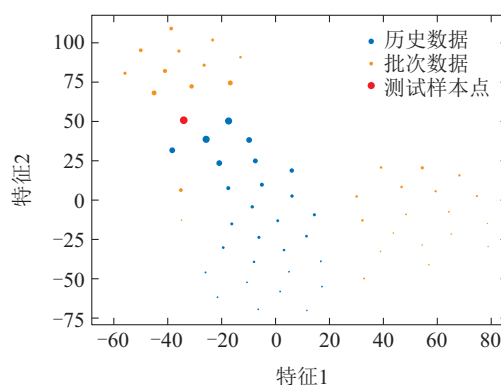
图9对No.6电池第35和第55次循环时的二维支持域分布以及训练数据样本权重进行可视化, 从图中可以看出, 对于不同的测试样本, 其二维支持域的选取不尽相同, 并且历史数据和批次数据的重要性也有所不同, 对于前期的待测样本点, 历史数据的重要性可能较大, 而对于中后期的样本点, 批次数据则是占主导地位, 二者共同构成的二维支持域将待测样本点囊括其中, 从而极大地提升了SOH预测的准确度。

4.4 消融实验

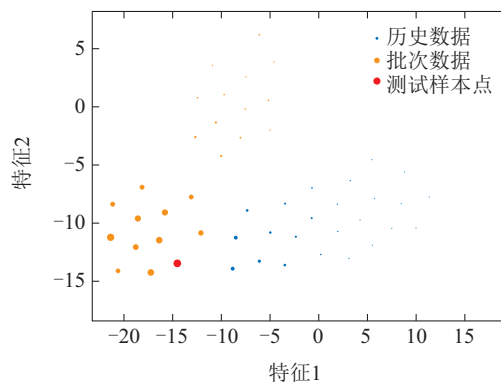
为了验证2D-RoSTL算法中各模块的有效性, 设计了消融实验进行对比, 其结果如表4所示。

表中二维支持域方法(2D-RoS)为本方法中去掉了直推式模块后的方法, 其直接使用构造的二维支持域进行建模, 不考虑样本权重。二维直推式方法(2D-TL)为本方法中去掉了二维支持域模块后的方法, 其

不进行二维支持域的构造, 而是直接使用全部的历史数据和批次数据进行建模。一维支持域直推式方法(1D-RoSTL)为本方法中将二维支持域替换为一维后的方法, 其不使用历史数据, 而是仅使用批次数据进行支持域的构造, 其中最优结果加粗表示, 可以看到, 2D-RoSTL在除No.7电池之外的其余所有实验组中均表现最优, 并且4组实验的误差均值也是最小, 对于No.7电池, 其与最优结果2D-TL误差接近, 但2D-TL算法建模是使用了全部的数据, 而2D-RoSTL在每次建模构造支持域时仅使用了很少的数据, 充分说明了2D-RoSTL各个模块的有效性。



(a) 第35次



(b) 第55次

图9 二维支持域可视化效果图

Fig. 9 The 2D region of support visualization

5 结论

本文针对锂电池SOH预测遇到的充放电循环历史数据不足与容量再生问题, 创新性地提出了一种2D-RoSTL方法, 通过对数据由粗到细的精准划分, 实现了基于少量充放电循环数据对SOH和再生现象的精确预测, 其中二维支持域的构造范式可以拓展到其他的少样本学习问题中。通过NASA的数据进行了实验验证, 2D-RoSTL方法对比其他的直接建模方法以及迁移学习方法都有显著的性能优越性, 其在4个电池中RMSE均小于1.56%, MAE均小于1.17%, 为少量数据的SOH预测提供了新思路。

表4 消融实验

Table 4 Ablation experiment

性能指标	电池编号	2D-RoS	2D-TL	1D-RoSTL	2D-RoSTL
RMSE/%	No.5	1.19	1.37	1.60	0.92
	No.6	3.58	2.52	3.04	1.50
	No.7	1.53	1.27	1.82	1.56
	No.18	1.06	1.11	1.94	1.04
	均值	1.84	1.57	2.10	1.26
MAE/%	No.5	0.92	1.07	1.39	0.75
	No.6	2.56	2.26	2.44	1.17
	No.7	0.94	0.83	1.55	0.96
	No.18	1.82	0.80	1.53	0.78
	均值	1.56	1.24	1.73	0.92

参考文献:

- [1] ZHENG Y, OUYANG M, HAN X, et al. Investigating the error sources of the online state of charge estimation methods for lithium-ion batteries in electric vehicles. *Journal of Power Sources*, 2018, 377: 161 – 188.
- [2] MANTHIRAM A. An outlook on lithium ion battery technology. *ACS Central Science*, 2017, 3(10): 1063 – 1069.
- [3] CHEN Y, MIAO Q, ZHENG B, et al. Quantitative analysis of lithium-ion battery capacity prediction via adaptive bathtub-shaped function. *Energies*, 2013, 6(6): 3082 – 3096.
- [4] FARMANN A, WAAG W, MARONGIU A, et al. Critical review of on-board capacity estimation techniques for lithium-ion batteries in electric and hybrid electric vehicles. *Journal of Power Sources*, 2015, 281: 114 – 130.
- [5] SARMAH S B, KALITA P, GARG A, et al. A review of state of health estimation of energy storage systems: Challenges and possible solutions for futuristic applications of Li-ion battery packs in electric vehicles. *Journal of Electrochemical Energy Conversion and Storage*, 2019, 16(4): 040801.
- [6] BEREICIBAR M, GANDIAGA I, VILLARREAL I, et al. Critical review of state of health estimation methods of Li-ion batteries for real applications. *Renewable and Sustainable Energy Reviews*, 2016, 56: 572 – 587.
- [7] WENG C, SUN J, PENG H. A unified open-circuit-voltage model of lithium-ion batteries for state-of-charge estimation and state-of-health monitoring. *Journal of Power Sources*, 2014, 258: 228 – 237.
- [8] LI J, ADEWUYI K, LOTFI N, et al. A single particle model with chemical/mechanical degradation physics for lithium ion battery state of health (SOH) estimation. *Applied Energy*, 2018, 212: 1178 – 1190.
- [9] WASSILIADIS N, ADERMANN J, FRERICKS A, et al. Revisiting the dual extended Kalman filter for battery state-of-charge and state-of-health estimation: A use-case life cycle analysis. *Journal of Energy Storage*, 2018, 19: 73 – 87.
- [10] ZHAO L, LIU Z, JI G. Lithium-ion battery state of charge estimation with model parameters adaptation using H_{∞} extended Kalman filter. *Control Engineering Practice*, 2018, 81: 114 – 128.
- [11] KHUMPROM P, YODO N. A data-driven predictive prognostic model for lithium-ion batteries based on a deep learning algorithm. *Energies*, 2019, 12(4): 660.
- [12] YOU G, PARK S, OH D. Real-time state-of-health estimation for electric vehicle batteries: A data-driven approach. *Applied Energy*, 2016, 176: 92 – 103.
- [13] KLASS V, BEHM M, LINDBERGH G. A support vector machine-based state-of-health estimation method for lithium-ion batteries under electric vehicle operation. *Journal of Power Sources*, 2014, 270: 262 – 272.
- [14] LI Y, ZOU C, BEREICIBAR M, et al. Random forest regression for online capacity estimation of lithium-ion batteries. *Applied Energy*, 2018, 232: 197 – 210.
- [15] ZHAO Guangcai, LIN Mingqiang, DAI Houde, et al. A modified strategy using the KNN-Markov chain for SOH estimation of lithium batteries. *Acta Automatica Sinica*, 2021, 47(2): 453 – 463. (赵光财, 林名强, 戴厚德, 等. 一种锂电池SOH估计的KNN-马尔科夫修正策略. *自动化学报*, 2021, 47(2): 453 – 463.)
- [16] MEI W, LI H, ZHAO C, et al. Numerical study on thermal characteristics comparison between charge and discharge process for lithium ion battery. *International Journal of Heat and Mass Transfer*, 2020, 162: 120319.
- [17] MESSING M, SHOA T, HABIBI S. Estimating battery state of health using electrochemical impedance spectroscopy and the relaxation effect. *Journal of Energy Storage*, 2021, 43: 103210.
- [18] GHIFARY M, KLEIJN W B, ZHANG M. Domain adaptive neural networks for object recognition. *Pacific Rim International Conference on Artificial Intelligence (PRICAI)*. Gold Coast, Australia: Springer, 2014, 8862: 898 – 904.
- [19] CHAI Z, ZHAO C. A fine-grained adversarial network method for cross-domain industrial fault diagnosis. *IEEE Transactions on Automation Science and Engineering*, 2020, 17(3): 1432 – 1442.
- [20] SUN B, SAENKO K. Deep CORAL: Correlation alignment for deep domain adaptation. *European Conference on Computer Vision (ECCV)*. Amsterdam, Netherlands: Springer, 2016, 9915: 443 – 450.
- [21] LONG M, CAO Y, WANG J, et al. Learning transferable features with deep adaptation networks. *International Conference on Machine Learning (ICML)*. Lille, France: PMLR, 2015, 37: 97 – 105.
- [22] CHAI Z, ZHAO C. Fault-prototypical adapted network for cross-domain industrial intelligent diagnosis. *IEEE Transactions on Automation Science and Engineering*, 2022, 19(4): 3649 – 3658.
- [23] CHAI Z, ZHAO C, HUANG B. Multisource-refined transfer network for industrial fault diagnosis under domain and category inconsistencies. *IEEE Transactions on Cybernetics*, 2022, 52(9): 9784 – 9796.
- [24] SHEN S, SADOUGHI M, LIM, et al. Deep convolutional neural networks with ensemble learning and transfer learning for capacity estimation of lithium-ion batteries. *Applied Energy*, 2020, 260: 114296.
- [25] YE Z, YU J. State-of-health estimation for lithium-ion batteries using domain adversarial transfer learning. *IEEE Transactions on Power Electronics*, 2021, 37(3): 3528 – 3543.
- [26] HAN T, WANG Z, MENG H. End-to-end capacity estimation of lithium-ion batteries with an enhanced long short-term memory network considering domain adaptation. *Journal of Power Sources*, 2022, 520: 230823.
- [27] SAHA B, GOEBEL K, POLL S, et al. Prognostics methods for battery health monitoring using a Bayesian framework. *IEEE Transactions on instrumentation and measurement*, 2008, 58(2): 291 – 296.
- [28] ZHAO Chunhui, WANG Fuli, YAO Yuan, et al. Phase-based statistical modeling, online monitoring and quality prediction for batch processes. *Acta Automatica Sinica*, 2010, 36(3): 366 – 374. (赵春晖, 王福利, 姚远, 等. 基于时段的间歇过程统计建模、在线监测及质量预报. *自动化学报*, 2010, 36(3): 366 – 374.)
- [29] ZHAO Chunhui, YU Wanke, GAO Furong. Data analytics and condition monitoring methods for nonstationary batch processes—current

- status and future. *Acta Automatica Sinica*, 2020, 46(10): 2072–2091. (赵春晖, 余万科, 高福荣. 非平稳间歇过程数据解析与状态监控—回顾与展望. *自动化学报*, 2020, 46(10): 2072–2091.)
- [30] LU Jingyi, CAO Zhixing, GAO Furong. Batch process control—overview and outlook. *Acta Automatica Sinica*, 2017, 43(6): 933–943. (卢静宜, 曹志兴, 高福荣. 批次过程控制—回顾与展望. *自动化学报*, 2017, 43(6): 933–943.)
- [31] LU N, YAO Y, GAO F, et al. Two-dimensional dynamic PCA for batch process monitoring. *AIChE Journal*, 2005, 51(12): 3300–3304.
- [32] LI X, JIANG J, CHEN D, et al. A capacity model based on charging process for state of health estimation of lithium ion batteries. *Applied Energy*, 2016, 177: 537–543.
- [33] SAVITZKY A, GOLAY M J. Smoothing and differentiation of data by simplified least squares procedures. *Analytical Chemistry*, 1964, 36(8): 1627–1639.
- [34] LI Y, ABDEL-MONEM M, GOPALAKRISHNAN R, et al. A quick on-line state of health estimation method for Li-ion battery with incremental capacity curves processed by Gaussian filter. *Journal of Power Sources*, 2018, 373: 40–53.

作者简介:

王一航 硕士研究生, 目前研究方向为锂离子电池的健康状态预测、迁移学习, E-mail: 22132123@zju.edu.cn;

冯良骏 博士研究生, 目前研究方向为零样本、少样本故障检测、软测量, E-mail: liangjunfeng@zju.edu.cn;

赵春晖 教授, 目前研究方向为机器学习、工业大数据解析与应用, E-mail: chhzhao@zju.edu.cn.