



## 基于 1DCNN-LSTM 的锂离子电池 SOH 预测

王英楷, 张 红, 王星辉

(福州大学物理与信息工程学院, 福建 福州 350116)

**摘 要:** 为了提高锂离子电池健康状态(SOH)的预测精准度和稳定性, 针对常规特征选取复杂且无法有效利用等问题, 提出了一种联合一维卷积(1DCNN)与长短记忆网络(LSTM)的电池 SOH 预测方法。首先采用多通道串联电压、电流、温度构建多维特征, 然后采用 1DCNN 从样本数据中提取高级数据特征输入 LSTM 中以有效利用历史信息, 最后通过全连接层输出电池 SOH 的预测结果。采用 NASA 锂离子电池容量衰减数据, 对所应用的联合算法进行验证, 结果表明, 相较于其他预测算法, 基于 1DCNN-LSTM 的算法具有更准确的 SOH 预测结果, 其平均绝对误差(MAE)为 0.01 左右, 且失效点误差周期(RUL)小于 2 个周期。

**关键词:** 1DCNN; LSTM; 锂电池; 多通道特征; 电池寿命

doi: 10.19799/j.cnki.2095-4239.2021.0250

中图分类号: TM 912

文献标志码: A

文章编号: 2095-4239 (2022) 01-240-06

## Hybrid 1DCNN-LSTM model for predicting lithium ion battery state of health

WANG Yingkai, ZHANG Hong, WANG Xinghui

(College of Physics and Information Engineering Fuzhou University, Fuzhou 350116, Fujian, China)

**Abstract:** To improve the prediction accuracy and stability of the lithium-ion battery state of health (SOH), a battery SOH prediction method combining one-dimensional convolution (1DCNN) and long and short-term memory network (LSTM) is proposed herein to solve the problems of the complex selection of conventional features and the inability to effectively use them. First, multichannel series voltage, current, and temperature are used to construct multi-dimensional features. Second, 1DCNN is used to extract advanced data features from the sample data and input them into the LSTM to effectively utilize historical information. Finally, the SOH prediction results of the battery are output through a full connection layer. The combined algorithm is verified by the capacity attenuation data of the NASA lithium battery. The results show that compared with other prediction algorithms, the algorithm based on the 1DCNN-LSTM has a more accurate prediction result of the SOH with a mean absolute error of 0.01 and a failure point error of less than two cycles.

**Key words:** 1DCNN; LSTM; lithium battery; battery life; multichannel feature

传统化石能源日益枯竭且带来了严峻的环境污染。中国碳中和的远景目标对可再生新能源的革命

带来了新的机遇与挑战。动力锂离子电池作为替代传统燃油汽车的主要方案, 是新能源革命的重中之

收稿日期: 2021-06-07; 修改稿日期: 2021-07-07。

基金项目: 国家自然科学基金 (83417013), 福建省科技厅 (2019J06008, 2018J01535)。

第一作者: 王英楷 (1993—), 男, 硕士研究生, 主要研究方向为新能

源与大数据, E-mail: 1258229203@qq.com; 通信作者: 张红, 副教授, 主要从事微纳材料与器件的研究, E-mail: zhanghong@fzu.edu.cn; 王星辉, 教授, 主要从事新型储能材料与器件领域的研究, E-mail: seaphy23@fzu.edu.cn。

重。为了有效地提升锂离子电池的利用率，在电池层面，可通过优化电池材料以提升能量密度和循环性能，在电池组层面，可通过有效监测锂电池健康状态(SOH)来优化电池管理系统<sup>[1]</sup>。

目前锂离子电池健康预测(SOH)中，主要是采用模型和数据驱动的方式<sup>[2-4]</sup>。焦自权等<sup>[5]</sup>对历史数据进行贝叶斯跟踪训练，优化辨识的电池老化模型参数，然后采用改进的 PF 算法对电池的健康状态进行预测，但电池健康衰减模型的选择影响着算法的精度。而数据驱动的方式，不需要通过深入电池内部的物理化学反应，仅通过外部的特征选取和特征表达，能很好地表征出电池 SOH 的变化情况，Choi<sup>[6]</sup>等联合多通道充电剖面与深度学习，提出的基于电压、电流和温度分布的多通道技术，结合长短期记忆网络(LSTM)处理时序趋势问题，提高了电池 SOH 预测精度。但构建只是充电的多通道，没有考虑放电过程中的特征表达，且直接输入长短期记忆网络中，无法充分利用数据与数据之间的高维特征。

针对上述分析，本文将结合一维卷积神经网络(1DCNN)与长短期记忆网络(LSTM)的优势，融合 CNN 处理高维度数据和 LSTM 处理时序问题的优势，预测锂离子电池 SOH<sup>[7]</sup>。传统预测性能的好坏很大程度取决于人工提取的数据特征的质量，为了获得相关度高的特征，往往在数据处理上需要专业的经验知识和人工的提取过程。但是面对较为复杂和庞大的数据库时，简单的模型无法深入提取高维的数据特征，且无法充分利用充放电过程中不同传感器采集的数据。为了更大程度地挖掘锂离子电池健康因子的数据信息，拟将每个充放电周期的电压、电流、温度在一定的采样周期内进行平均化处理，根据充放电的周期串联成高维度特征。结合 1DCNN 进行特征提取输入到长短期记忆网络(LSTM)中，该方法充分利用输入数据的关键特征信息，还能将历史信息和当前输入信息很好地映射到当前 SOH 的预测中。

## 1 电池 SOH 影响因素研究

锂离子电池虽然拥有高的能量密度、可重复利用的优点，但是在充放电能量交换过程中不可避免地产生电解液等材料的消耗，这个物理化学的反应过程无疑会引起锂离子电池的性能衰减，从而影响

锂离子电池的剩余寿命。锂离子电池的 SOH 定义式为

$$\text{SOH} = \frac{C_i}{C_e} (i = 1, 2, 3, \dots) \quad (1)$$

式中， $C_i$  为电池所能保持的实际电池容量； $C_e$  为电池的额定容量；SOH 会随着充放电的过程会发生明显的变化。

图 1 为锂离子电池在第 1 周期和第 160 周期的充电曲线，可见电压、电流、温度在第 1 个周期和第 160 个周期的充电曲线出现明显的偏移。此前王宇胜等<sup>[8]</sup>只对电压进行特征工程，仅仅采用单个周期的平均放电电压、电压变化率等作为主要特征向量，忽略电流、温度的影响。较为复杂的特征工程，往往需要进行人为的特征筛选和提取，不便于实际场所的使用。所以在电池健康的影响因素中引入多通道的特征，只需要将采样点平均化处理，串联成高维特征矩阵，避免对初始数据过多的修饰。

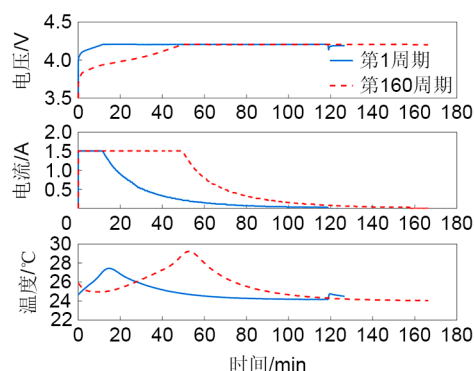


图 1 第 1 和 160 周期特征变化  
Fig. 1 Characteristic changes in cycles 1 and 160

## 2 算法介绍

### 2.1 卷积神经网络

卷积神经网络(CNN)<sup>[9]</sup>主要是用于模式的识别和特征提取，主要应用在图像识别、计算机视觉领域，CNN 可以通过卷积层的设置，提取到输入信息之间更高维度的信息特征。典型的 CNN 由输入层、卷积层、池化层、输出层构成，卷积层的滤波器通过合适的窗口大小及滑动步长进行特征提取，再经过池化层的平均化处理的高级特征输出到全连接层，实现回归预测或是分类。卷积的计算式如下

$$\mathbf{x}_i' = f(\mathbf{W}_i' * \mathbf{X}^{l-1} + \mathbf{b}_i') \quad (2)$$

式中， $\mathbf{x}_i'$  表示输出值的第  $l$  层的第  $i$  个特征向

量;  $W_i^l$  表示第  $l$  层的第  $i$  个滤波器的权重矩阵;  $*$  运算符表示卷积运算的过程;  $X^{l-1}$  表示  $l-1$  层的运算输出;  $b_i^l$  为运算式的偏置项;  $f$  表示激活函数, 这里选用 ReLu 作为激活函数, CNN 可以通过非线性的激活函数来近似拟合现实生活中的非线性问题。

池化层主要进行下采样, 通过对空间信息的合并, 在保证重要信息的基础上降低维度, 这里采用的是最大池化, 如式(3)所示,  $y_i^{l+1}$  指经过池化层后  $l+1$  层的第  $i$  个特征图的元素;  $D_j$  指池化的区域;  $x_j^l$  表示该池化区域内的元素。

$$y_i^{l+1}(j) = \max x_j^l(k), k \in D_j \quad (3)$$

## 2.2 长短记忆网络

LSTM<sup>[10]</sup> 是一种具有记忆功能的循环神经网络(RNN)算法, 它能有效避免 RNN 梯度消失或梯度爆炸的问题, LSTM 可以对时序数据进行精确的建模, 广泛用于自然语言处理领域。锂离子电池的 SOH 随着时间的推移会发生衰减, 故采用 LSTM 对锂电池的 SOH 进行预测。LSTM 单个单元如图 2 所示, 主要由 3 种门结构构成: 输入门、输出门和遗忘门。

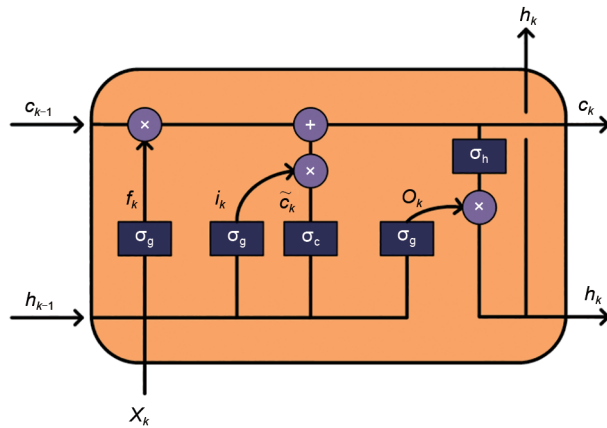


图2 LSTM 单个门  
Fig. 2 LSTM single gate

输入门主要是将当前信息储存在 LSTM 中

$$i_k = \sigma(W_i[h_{k-1}, x_k] + b_i) \quad (4)$$

式中,  $W_i$  输入节点的权重矩阵;  $h_{k-1}$  为  $k-1$  时刻的输出值;  $x_k$  为  $k$  时刻输入到 LSTM 中的输入值;  $b_i$  为偏置项;  $\sigma$  为激活函数, 这里选用 Sigmoid 函数。

输出门主要用于输出当前的 LSTM 单元的值

$$o_k = \sigma(W_o[h_{k-1}, x_k] + b_o) \quad (5)$$

式中,  $W_o$  指输出的权重矩阵;  $h_{k-1}$  为  $k-1$  时刻的输出值;  $x_k$  为  $k$  时刻输入到 LSTM 中的输入值;  $b_o$  为偏置项;  $\sigma$  为激活函数, 这里选用 Sigmoid 函数。

遗忘门主要是通过对历史信息的控制, 对当前时刻的 LSTM 单元产生影响

$$f_k = \sigma(W_f[h_{k-1}, x_k] + b_f) \quad (6)$$

式中,  $W_f$  为遗忘门的权重矩阵;  $h_{k-1}$  为  $k-1$  时刻的输出值;  $x_k$  为  $k$  时刻输入到 LSTM 中的输入值;  $b_f$  为遗忘门偏置项;  $\sigma$  为激活函数。

## 2.3 1DCNN-LSTM 算法

CNN 主要是采用二维卷积(2DCNN)处理二维图像问题, 而锂离子电池的衰老过程一般是依存于使用时间的连续演化过程, 因此锂离子电池健康因子一般是基于时间序列的一维数据。为了保留输入 LSTM 中的时序信息, 这里采用的是一维时间序列(1DCNN)对锂离子电池特征信息进行卷积处理, 避免了时间信息的割裂<sup>[11]</sup>。

CNN 一般在卷积层后连接 Flatten 层降维再输出到全连接层, 其中 Flatten 操作是将二维的数据平铺成一维的过程, 此时会损失部分特征。而本研究采用的 1DCNN-LSTM 算法中, 1DCNN 与 LSTM 之间仅用池化层连接, 实现最大程度高维特征的提取, 同时, 降低维度输入到 LSTM 当中。神经网络结构如图 3 所示,  $V$ 、 $I$ 、 $T$  表示当前充放周期的电压、电流和温度, 采样点平均化处理, 串联成 60 维度的输入矩阵。隐藏层由 1DCNN 和 LSTM 组成, 最后连接到全连接层, 输出为单个 SOH 的预测值, 中间引入随机丢弃机制层(dropout=0.3),

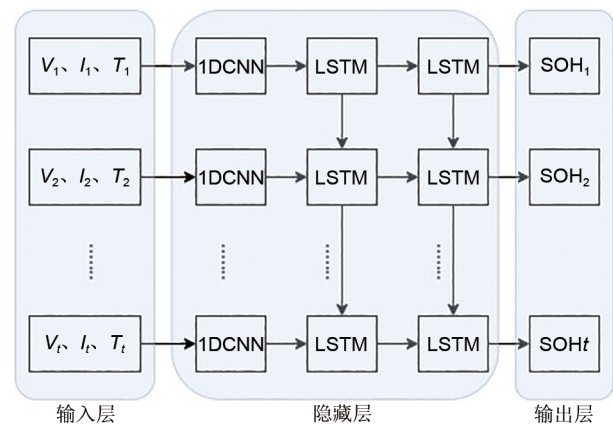


图3 1DCNN-LSTM 架构  
Fig. 3 1 DCNN LSTM architecture



该机制可以随机丢弃神经元之间的权重，降低对单一神经元的依赖，并降低输入的锂离子电池健康特征的噪声影响。

系统模型的流程如图 4 所示。首先，将 167 个周期的充放电数据中电压、电流、温度间隔平均化处理，再串联成 60 维度的特征，构成的多通道锂离子电池模型，充分考虑电压、电流、温度在不同周期的变化过程，且只需要进行平均化处理，无需人为的复杂特征工程，平均化处理的过程也能防止数据在短时间内的震荡。其次，根据滑动窗口 (TIMESTEP=3) 进行数据的构造，1DCNN 通过一维卷积层提取电池充放电的高级信息，并通过池化层 (Maxpooling) 降低特征维度，既避免了输入 LSTM 数据的复杂度，又加快 LSTM 网络的训练速度。再次，借由 LSTM 对于时序问题的优秀处理能力，LSTM 层设置两层，神经元个数为 256，并引入 dropout 层，设置为 0.3，优化器选用 Adam，最后得到最优模型。

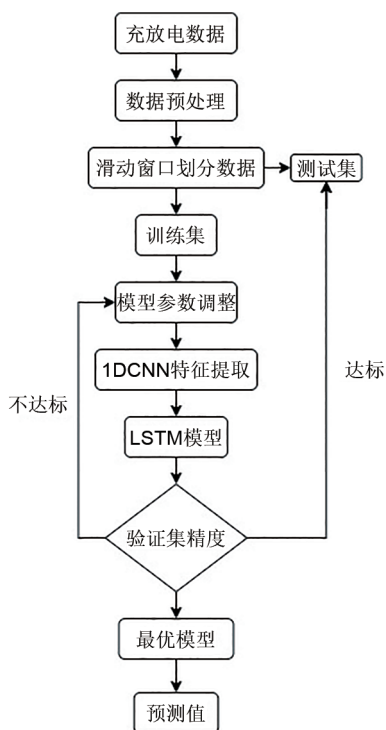


图 4 模型流程

Fig. 4 The process of the model

### 3 实验验证

#### 3.1 实验数据集以及预处理

本次研究采用的数据是来自美国航天局所提供

的 Battery Data Set 数据集<sup>[12]</sup>，电池的类型为 18650 钴酸锂电池，额定容量为 2 A·h。锂电池的老化实验主要经历两个过程：充电、放电。充电过程主要是以 1.5 A 的恒定电流模式充电至电压达到 4.2 V，后续用恒压模式继续充电直至电流下降到 20 mA；放电模式则是以恒定电流模式进行放电，直至电池电压达到设定值，根据设定值的不同，进行不同情况的老化，5 号电池截止电压为 2.7 V，6 号电池截止电压为 2.5 V，7 号电池截止电压为 2.2 V，如图 4 所示为不同电池的老化衰减情况，锂电池在 168 个充放电周期衰减到额定容量的 70%，通常 70% 的额定容量被认为是电池老化的阈值，锂电池在性能退化的过程中，会出现容量突变上升，是因为锂电池在休息期中物理化学反应的松弛，实现锂离子的再生，且同种电池会出现初始容量有差异的问题，这些不可避免的老化过程给电池健康的预测带来了困难。

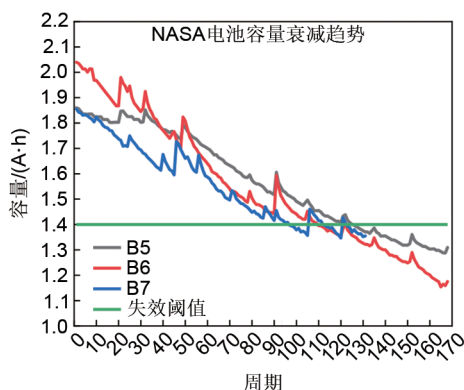


图 5 NASA 电池衰减趋势

Fig. 5 Decay trend for NASA battery

数据划分后，需要对特征进行归一化处理。电池健康 (SOH) 保持在 0~1，而电压、电流、温度的取值范围比较大，为了提高训练的效率，需要将训练集中的特征进行归一化处理，归一化公式如下

$$x' = (x - x_{\min}) / (x_{\max} - x_{\min}) \quad (7)$$

式中， $x'$  为归一化后的值； $x$  为样本值； $x_{\min}$  为样本的最小值； $x_{\max}$  为样本的最大值，归一化处理后可以保证样本在 0~1。由于可用的数据集是有限的，需要充分利用可用的数据集。因此，在 3 个数据集中，选择两个电池组以及剩余电池组前 80 个循环周期作为训练集，其余的一个电池组后 50 个周期作为测试集，并重复此过程得到 3 次测试结果。

### 3.2 实验结果和分析

为了更好地评估整个模型的性能指标,引入平均绝对误差(MAE)和平均百分比误差(MAPE)来评价锂离子电池健康状态预测模型的结果。评价公式如下

$$MAE(X, h) = \frac{1}{N} \sum_{i=1}^m |h(x^i) - y^i| \quad (8)$$

$$MAPE(X, h) = \frac{1}{N} \left| \frac{h(x^i) - y^i}{y^i} \right| \times 100\% \quad (9)$$

$$RUL = C_p - C_r \quad (10)$$

式中,  $h(x^i)$ 指模型对于SOH的预测值;  $y^i$ 为电池SOH实际值;  $N$ 为样本的总数;  $C_p$ 为模型对于失效点的预测周期;  $C_r$ 为电池真实失效点周期; RUL为失效点预测的误差。

表1为各个算法指标在不同电池之间的表现,

从中可见,在MAE和MAPE的指标上,1DCNN-LSTM算法上明显优于纯粹的LSTM算法。在B5电池上,MAE和MAPE上分别下降了0.0228%和1.68%,失效阈值的预测值误差,1DCNN-LSTM准确预测出失效点,而LSTM预测误差可达到8个周期;在B6电池上,MAE和MAPE上分别下降了0.013%和1.19%,失效阈值的预测值误差,1DCNN-LSTM预测误差基本在2个周期左右,而LSTM预测误差可达到7个周期;在预测周期为80个点时,相较于预测50个周期,少了30个循环周期作为训练样本,MAPE和MAE误差有所提高,对于LSTM算法预测拟合度和准确性有所下降,但是对于联合1DCNN-LSTM,仍能保持良好的预测性能。就总体而言,1DCNN-LSTM的联合算法明显更优于LSTM算法。

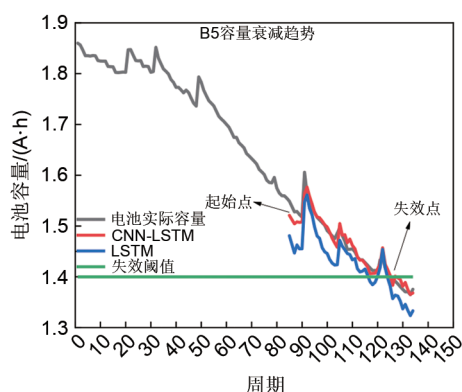
表1 算法指标

Table 1 Algorithm indicators

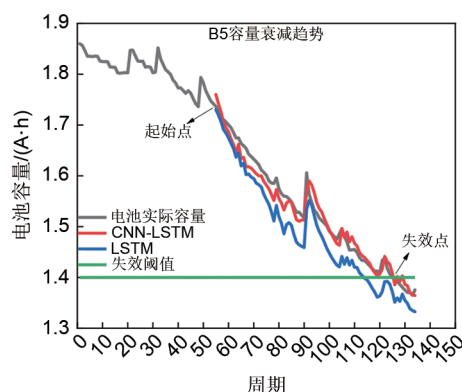
数据		50个预测点					80个预测点				
		MAE	MAPE/%	$C_r$	$C_p$	RUL	MAE	MAPE/%	$C_r$	$C_p$	RUL
B5	LSTM	0.032	2.3	125	117	-8	0.038	2.68	125	112	-13
	1DCNN-LSTM	0.0092	0.62	125	125	0	0.013	0.91	125	125	0
B6	LSTM	0.024	1.9	110	117	7	0.036	2.12	110	120	10
	1DCNN-LSTM	0.011	0.71	110	111	1	0.015	1.02	110	112	2
B7	LSTM	0.035	2.6	97	103	6	0.040	2.8	97	111	14
	1DCNN-LSTM	0.013	0.83	97	99	2	0.02	1.13	97	103	6

图6分别为B5、B6电池在50个预测点和80个预测点下的1DCNN-LSTM与LSTM拟合曲线,可见,在初始的容量衰减周期里,与电池的实际容量的曲线几乎重合,证明了多通道电池健康因子的可靠以及LSTM算法的短期记忆能力,但是随着预测

充放电周期的延长,纯粹的LSTM出现明显偏离实际的容量曲线,导致最终失效点的预测误差加大,而一维卷积(1DCNN)在处理一维的锂电池充放电数据上能保留更多的有效信息,使得预测结果能准确贴合电池容量衰减曲线,所以在锂电池健康状态估



(a) B5电池50个预测点



(b) B5电池80个预测点

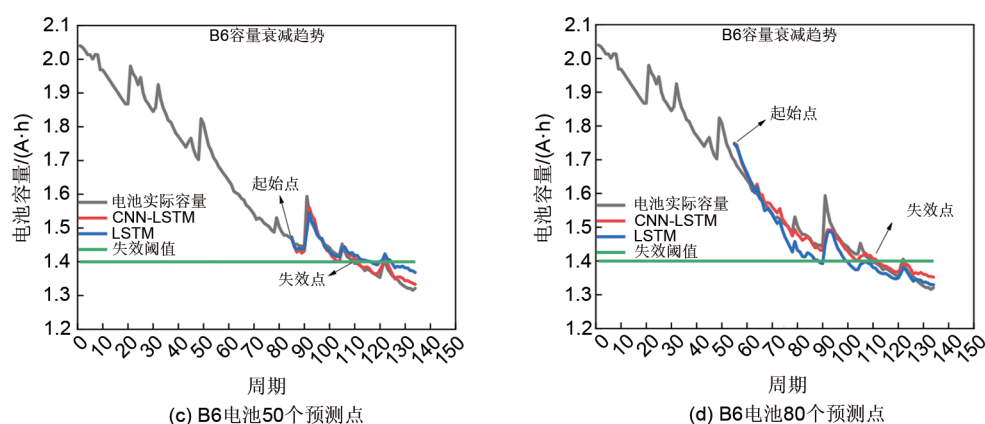


图6 电池SOH预测对比  
Fig. 6 Comparison of battery SOH prediction

计上有更明显的优势。

## 4 结 论

本文提出结合1DCNN和LSTM算法对锂离子电池的容量衰减趋势进行有效预测。主要结论如下。

(1) 文中提出利用充放电多通道特征信息, 构成了多维时间序列数据, 仅采用平均采样, 避免复杂的特征工程, 便于现实场景的应用;

(2) 采用了1DCNN进行时序特征的进一步提取, 结合LSTM保留历史信息实现对于电池健康状态的准确估计, 而且在训练集数据减少的情况下对失效点预测结果影响不大;

(3) 该方法相较于单一的LSTM在MAPE、MAE以及RUL指标上有更高的精度和更好的拟合度。在B5、B6数据集中, 1DCNN-LSTM在50个周期和80个周期的失效点预测误差在2个周期以内, MAPE和MAE分别低至1.12%和0.01。

## 参 考 文 献

- [1] 刘大同, 宋宇晨, 武巍, 等. 锂离子电池组健康状态估计综述[J]. 仪器仪表学报, 2020, 41(11): 1-18.  
LIU D T, SONG Y C, WU W, et al. Review of state of health estimation for lithium-ion battery pack[J]. Chinese Journal of Scientific Instrument, 2020, 41(11): 1-18.
- [2] 戴海峰, 张艳伟, 魏学哲, 等. 锂离子电池剩余寿命预测研究[J]. 电源技术, 2019, 43(12): 2029-2035.  
DAI H F, ZHANG Y W, WEI X Z, et al. Review of remaining useful life prediction for lithium ion battery[J]. Chinese Journal of Power Sources, 2019, 43(12): 2029-2035.
- [3] 苏伟, 钟国彬, 沈佳妮, 等. 锂离子电池故障诊断技术进展[J]. 储能科学与技术, 2019, 8(2): 225-236.  
SU W, ZHONG G B, SHEN J N, et al. The progress in fault

diagnosis techniques for lithium-ion batteries[J]. Energy Storage Science and Technology, 2019, 8(2): 225-236.

- [4] 田君, 高洪波, 张跃强, 等. 电动汽车动力锂离子电池寿命预测方法研究[J]. 电源技术, 2020, 44(5): 767-770.  
TIAN J, GAO H B, ZHANG Y Q, et al. Research of life prediction methods for power Li-ion battery in electric vehicles[J]. Chinese Journal of Power Sources, 2020, 44(5): 767-770.
- [5] 焦自权, 范兴明, 张鑫, 等. 基于改进粒子滤波算法的锂离子电池状态跟踪与剩余使用寿命预测方法[J]. 电工技术学报, 2020, 35(18): 3979-3993.  
JIAO Z Q, FAN X M, ZHANG X, et al. State tracking and remaining useful life predictive method of Li-ion battery based on improved particle filter algorithm[J]. Transactions of China Electrotechnical Society, 2020, 35(18): 3979-3993.
- [6] CHOI Y, RYU S, PARK K, et al. Machine learning-based lithium-ion battery capacity estimation exploiting multi-channel charging profiles[J]. IEEE Access, 2019, 7: 75143-75152.
- [7] 高昆伦, 杨帅, 刘思言, 等. 基于一维卷积神经网络的电力系统暂态稳定评估[J]. 电力系统自动化, 2019, 43(12): 18-26.  
GAO K L, YANG S, LIU S Y, et al. Transient stability assessment for power system based on one-dimensional convolutional neural network[J]. Automation of Electric Power Systems, 2019, 43(12): 18-26.
- [8] 王宇胜, 陈德旺, 蔡俊鹏, 等. 基于LSTM-SVR的锂电池健康状态预测研究[J]. 电源技术, 2020, 44(12): 1784-1787.  
WANG Y S, CHEN D W, CAI J P, et al. Research on lithium battery state of health prediction based on LSTM-SVR[J]. Chinese Journal of Power Sources, 2020, 44(12): 1784-1787.
- [9] 周飞燕, 金林鹏, 董军. 卷积神经网络研究综述[J]. 计算机学报, 2017, 40(6): 1229-1251.  
ZHOU F Y, JIN L P, DONG J. Review of convolutional neural network[J]. Chinese Journal of Computers, 2017, 40(6): 1229-1251.
- [10] 杨丽, 吴雨茜, 王俊丽, 等. 循环神经网络研究综述[J]. 计算机应用, 2018, 38(S2): 1-6, 26.  
YANG L, WU Y X, WANG J L, et al. Research on recurrent neural network[J]. Journal of Computer Applications, 2018, 38(S2): 1-6, 26.
- [11] SONG X B, YANG F F, WANG D, et al. Combined CNN-LSTM network for state-of-charge estimation of lithium-ion batteries[J]. IEEE Access, 2019, 7: 88894-88902.
- [12] SAHA B, Goebel "Battery data set" NASA AMES prognostics data repository[EB/OL]. 2007. <http://ti.arc.nasa.gov/tech/dash/groups/pcoe/prognostic-data-repository/#battery>.