Title

Deep transfer learning enabled online state-of-health estimation of lithium-ion batteries under small samples across different cathode materials, ambient temperature and charge-discharge protocols

Author name(s)

Xiaopeng Li[a], Minghang Zhao[a,c,*], Shisheng Zhong[a,b,c], Junfu Li[d], Zhiquan Cui[d], Song Fu[b], Zhiqi Yan[e]

Affiliation(s):

[a] Department of Mechanical Engineering, Harbin Institute of Technology, Weihai 264209, China

[b] School of Mechatronics Engineering, Harbin Institute of Technology, Harbin 150000, China

[c] Weihai Key Laboratory of Intelligent Operation and Maintenance, Harbin Institute of Technology, Weihai 264209, China

[d] School of Automotive Engineering, Harbin Institute of Technology, Weihai 264209, China

[e] College of Aeronautical Engineering, Civil Aviation University of China, Tianjin 300300, China

Abstract

State-of-health (SOH) estimation of lithium-ion batteries is crucial for ensuring their efficient and safe operation. However, the accurate SOH estimation under low temperatures and high discharge rates is still a challenging problem especially when facing insufficient early-stage data. To tackle this problem, this paper proposes a self-attention-based deep transfer learning (SDTL) approach that can be flexibly updated in respond to diverse cathode materials and varied operation conditions. First, an efficient self-attention-based feature learning model is constructed to capture diverse degradation patterns of batteries under different operating conditions. Second, deep transfer learning techniques are employed to achieve adaptable SOH estimation using previously learned degradation knowledge and the limited data of a new battery. To comprehensively validate the proposed approach, full life cycle tests on nickel cobalt manganese (NCM) batteries under 1C/2C conditions are conducted to supplement the nickel cobalt aluminum (NCA) public battery datasets. All the prepared battery datasets cover different cathode materials, charge-discharge rates, and ambient temperatures. Afterwards, eighteen health indicators are extracted and selected with Pearson correlation coefficient (PCC) to comprehensively characterize the statistical, electrochemical, and dynamic properties of batteries. Through comparisons with classical models that directly trained using state-of-the-art deep learning algorithms and other widely used deep transfer learning methods, the proposed SOH

estimation approach has shown wide generalizability as well as a positive accuracy improvement.

Keywords: State-of-health (SOH) estimation; Lithium-ion batteries; self-attention-based deep transfer learning (SDTL); insufficient early-stage data

Acknowledgment:

This work is supported by National Key R&D Program of China (2023YFB4302400).

1. Introduction

Lithium-ion batteries, as efficient and reliable energy storage devices, have been widely utilized in various fields, including electric vehicles, civil airlines and renewable energy storage systems [1–3]. With the growing energy demand and advancements in energy storage technology, accurately estimating the state-of-health (SOH) of batteries is crucial for ensuring device safety, extending battery lifespan, and reducing maintenance costs.

However, the complex electrochemical characteristics and nonlinear degradation trends of lithium-ion batteries make it challenging to accurately estimate their current SOH [4]. Key factors include solid electrolyte interphase (SEI) layer formation, lithium plating, and active lithium loss, compounded by side reactions like electrolyte decomposition and metal dissolution. These processes interact dynamically, leading to highly nonlinear degradation patterns [5]. To address this issue, numerous researchers have proposed various prediction methods in recent years, aiming to improve the reliability and accuracy of SOH estimation [6,7]. Currently, SOH estimation methods can be broadly categorized into two approaches: physical model-based methods and data-driven methods.

Physical model-based methods describe the state variations of batteries by constructing electrochemical models (EM) or equivalent circuit models (ECM). Sihvo et al. [8] employed a nonlinear ECM to describe the impedance characteristics of lithium-ion batteries, accounting for the constant phase element (CPE), charge-transfer processes, SEI effects, and ohmic and inductive resistances. Li et al. [9] proposed an EM based on the single particle (SP) model, which incorporates the growth of the SEI layer and the crack propagation caused by the volume expansion of anode and cathode materials during charging and discharging processes. Wang et al. [10] developed an electrochemical-thermal model based on pseudo-two-dimensional (P2D) [11] to improve the accuracy of estimation results with seven highly SOH-related electrochemical parameters. These methods offer physical interpretability but often rely on extensive experimental data and complex parameters like. Up to now, their applicability is still limited in real-world scenarios due to the influence of battery materials and operating conditions. In contrast, data-driven methods do not depend on specific physical models. Instead, they predict SOH by extracting features from historical data and constructing machine learning or deep learning models.

The predominant data-driven methods currently include convolutional neural networks (CNNs), recurrent

neural networks (RNNs), such as long short-term memory (LSTM) [12], gated recurrent units (GRU) [13], and Transformers [14]. Qian et al. [15] proposed a method based on a one-dimensional CNN to estimate lithium-ion battery SOH, where random segments of the charging curve are used as input. Xu et al. [16] improved the traditional convolutional neural network by integrating it with LSTM and introduced Skip Connections to enhance the accuracy of SOH estimation. Chen et al. [17] applied the Transformer network to SOH estimation for lithium-ion batteries and illustrated that Transformer models outperform RNN-based models, in terms of both prediction performance and computational efficiency. Jia et al. [18] proposed a network combining bidirectional GRU (Bi-GRU) and Transformers, showing that it offers higher accuracy, better robustness, and stronger generalization capability than traditional Transformer model.

However, whether data-driven or physical model-based methods, SOH estimation methods typically require training models again with new, extensive battery data for different tasks to achieve high accuracy and generalization. For instance, She et al. [19] proposed an offline-online blended machine learning approach for Li-ion battery SOH estimation, which combines incremental capacity analysis (ICA) with model migration and ensemble learning. While their method improves estimation accuracy, it still relies on multiple offline-trained models and requires periodic true SOH measurements for online adaptation, highlighting the persistent challenge of data dependency in conventional methods. Similarly, Tang et al. [20] developed a Bayesian-Monte Carlo framework for battery aging prediction that incorporates model migration and uncertainty quantification. Their approach showed improved generalization across different battery chemistries but still required substantial source domain data and suffered from computational complexity in real-time applications. Deep transfer learning (DTL) addresses this challenge by transferring knowledge learned by deep learning models from a large source domain to a target domain with limited data. This approach enhances the model's capability for small-sample learning while accelerating the training process, avoiding the need to train a deep learning model from scratch. This is particularly advantageous when target domain data is scarce or computational resources are limited. Consequently, DTL has garnered increasing attention in recent years and has been progressively adopted in battery SOH estimation research to tackle issues related to cross-domain data distribution discrepancies and high data requirements.

The DTL enabled SOH estimation methods can be broadly categorized into three techniques: domain adaptation, domain adversarial and fine-tuning [21–23].

Domain adaptation methods aim to reduce distribution discrepancies between source and target

domains, enabling models to transfer knowledge effectively. Lu et al. [24] proposed a deep transfer learning method based on domain adaptation to achieve accurate SOH estimation without requiring labeled data from the target battery domain. Ma et al. [25] introduced a domain adaptation technique combined with a CNN-based deep transfer learning model for SOH estimation. This method incorporates Maximum Mean Discrepancy (MMD) into the loss function during the forward propagation process, effectively reducing the discrepancy between the source and target domains and improving estimation accuracy. Yang et al. [26] proposed a joint adaptive deep transfer learning (JADTL) model, which utilizes an Ada-CNN-GRU adaptive recognition framework to identify data segments with the largest distribution differences. By minimizing MMD, the model reduces these differences, enabling more precise SOH estimation.

Domain adversarial methods leverage adversarial training to align feature distributions between domains and enhance model generalization. Su et al. [27] proposed a method based on generative adversarial networks (GANs) to construct a domain-adversarial neural network (DANN). This model integrates a deep feature extractor and a discriminator. By simultaneously minimizing the feature extraction and regression losses while maximizing the domain discrimination loss, the approach enables the model to effectively learn features from the source domain and adapt to the distribution of the target domain data.

Fine-tuning adapts pre-trained models to new tasks by updating parameters with limited target domain data, making it a straightforward and effective transfer learning technique. By freezing certain layers of the model and updating the parameters of the remaining layers using a small amount of target domain data, the model can be better adapted to the specific tasks of the target domain. Kim et al. [21] utilized an LSTM model and fine-tuned the parameters of the final fully connected layer to improve model`s performance. Similarly, Dong et al. [28] choose voltage sample entropy as a health indicator and fine-tuned the parameters of the last four layers of a Bi-LSTM model with an attention mechanism to adapt it to a new battery task. Yao et al. [29] proposed a deep transfer learning convolutional neural network (DTCNN) model based on incremental capacity curve analysis. They combined fine-tuning techniques to eliminate distribution differences between different battery types or charge-discharge protocols, thereby improving the accuracy of the estimation results.

Compared to domain adaptation and domain adversarial methods, fine-tuning avoids the assumption of significant similarity between source and target domain distributions, making it more suitable for scenarios where the source and target domains have minor differences. Additionally, fine-tuning is computationally

efficient and easier to implement, as it does not require complex loss functions or adversarial training, making it better aligned with the lightweight requirements of current battery tasks.

Therefore, this paper proposed an efficient self-attention-based deep transfer learning (SDTL) system for online SOH estimation, as illustrated in Fig. 1, to address the accuracy challenges faced by many existing approaches, including limited early-stage data and varying operating conditions. Initially, a full life aging test was conducted on batteries with different cathode materials and operation conditions. Secondly, eighteen HIs were extracted from charge-discharge current, voltage and incremental capacity curves. Subsequently, 3-4 HIs were selected by the Pearson correlation coefficient and processed into the form required by the attention mechanism through a sliding window segmentation approach. The treated data were categorized into four series according to operating conditions and materials. Then, the selected HIs were combined with the actual SOH to train the different models respectively. Finally, fine-tuning was applied to update the model parameters to better adapt to the degradation patterns under new battery operating conditions. In the fine-tuning process, only the final fully connected layer was unfrozen, which helps preserve the generic feature representations learned in the earlier layers, reduces the risk of overfitting, and minimizes computational complexity while enabling rapid adaptation to new data.

The originality of this work is summarized as follows:

(1) An efficient Self-attention-based Deep Transfer Learning (SDTL) approach is proposed, which utilizes a large amount of historical data from similar and previous tasks to pre-train the model. The pre-trained model will then be fine-tuned with a small amount of data to efficiently learn the battery degradation patterns of a new task, eliminating the need for retraining from scratch.

(2) External experiments on NCM batteries under 1C/2C conditions are conducted to supplement the public NCA battery datasets and all the datasets we prepared are collected across different cathode materials, charge-discharge rates, and ambient temperatures. To further enhance the performance of different models, 18 health indicators are extracted and selected with PCC to comprehensively characterize the statistical, electrochemical, and dynamic properties of batteries.

The subsequent sections are organized as follows: Section 2 describes the datasets, focusing on two types of lithium-ion batteries (NCM and NCA) and their specifications, along with experimental setups under various operating conditions. Section 3 detailed the proposed self-attention-based deep transfer learning method for SOH estimation. Section 4 outlines the experimental procedures and provides a comprehensive

analysis of results obtained from multiple trials. Finally, the conclusion and suggests directions for future research are summarized in Section 5.

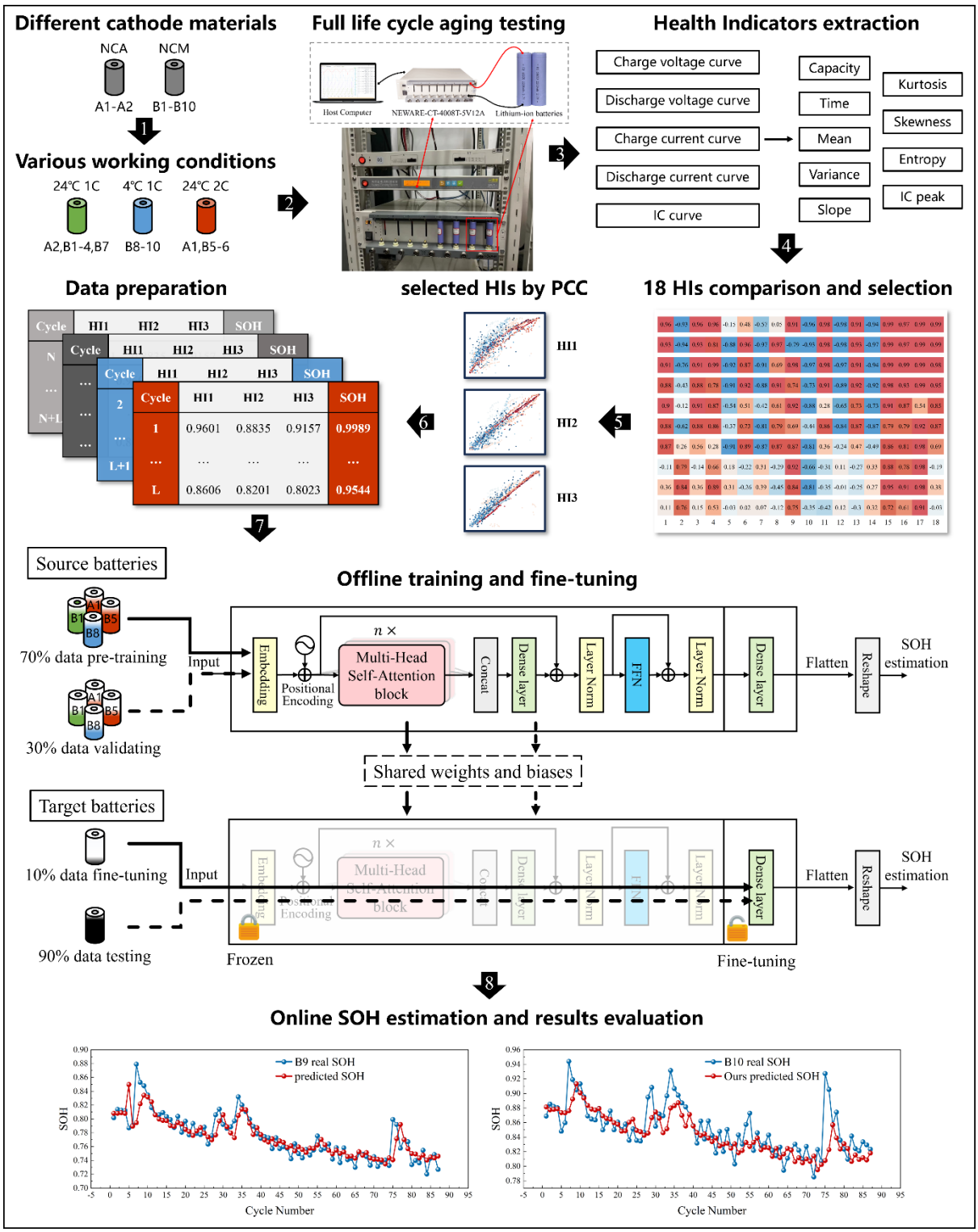

Fig. 1. Deep transfer learning enabled online SOH estimation framework.

2. Description of LIB datasets

This section introduces selected battery datasets under four kinds of operation conditions (series1-4), as shown in **Table 1.** These datasets encompass different battery materials, charge-discharge rates, and operating temperatures, all using 2Ah 18650 cylindrical batteries. Dataset A is collected from our laboratory, while dataset B is obtained from Prognostics Center of Excellence Data Set Repository of NASA.

Table. 1. Main characteristics of twelve battery cells.

Series	Batteries	Cathode material	Charge rate	Discharge rate	Temperature
1	A1	NCM	1C	**2C**	24℃
	A2	NCM	1C	1C	24℃
2	B1	NCA	0.75C	1C	24℃
	B2	NCA	0.75C	1C	24℃
	B3	NCA	0.75C	1C	24℃
	B4	NCA	0.75C	1C	24℃
3	B5	NCA	0.75C	**2C**	24℃
	B6	NCA	0.75C	**2C**	24℃
	B7	NCA	0.75C	1C	24℃
4	B8	NCA	0.75C	1C	**4℃**
	B9	NCA	0.75C	1C	**4℃**
	B10	NCA	0.75C	1C	**4℃**

2.1.1. NCM batteries under different discharge rate

The charge-discharge experiments on 18650 batteries with NCM material were conducted in our laboratory (as shown in **Fig. 2(a)**). The experimental setup consisted of a central control computer and a NEWARE-CT series testing system. Two NCM batteries with a nominal capacity of 2Ah (designated as A1 and A2) were utilized. During the charging process (as shown in **Fig. 2(b)**), the batteries were charged at a constant current of 1C until the voltage reached 4.2V, followed by constant voltage charging until the current dropped to 50mA. Subsequently, the batteries were rested for 15 minutes before undergoing constant current discharge until the voltage dropped below 2.5V, followed by another 15-minute rest. This process was repeated cyclically until the SOH reached the EOL threshold of 80% (as shown in **Fig. 2(c)**).

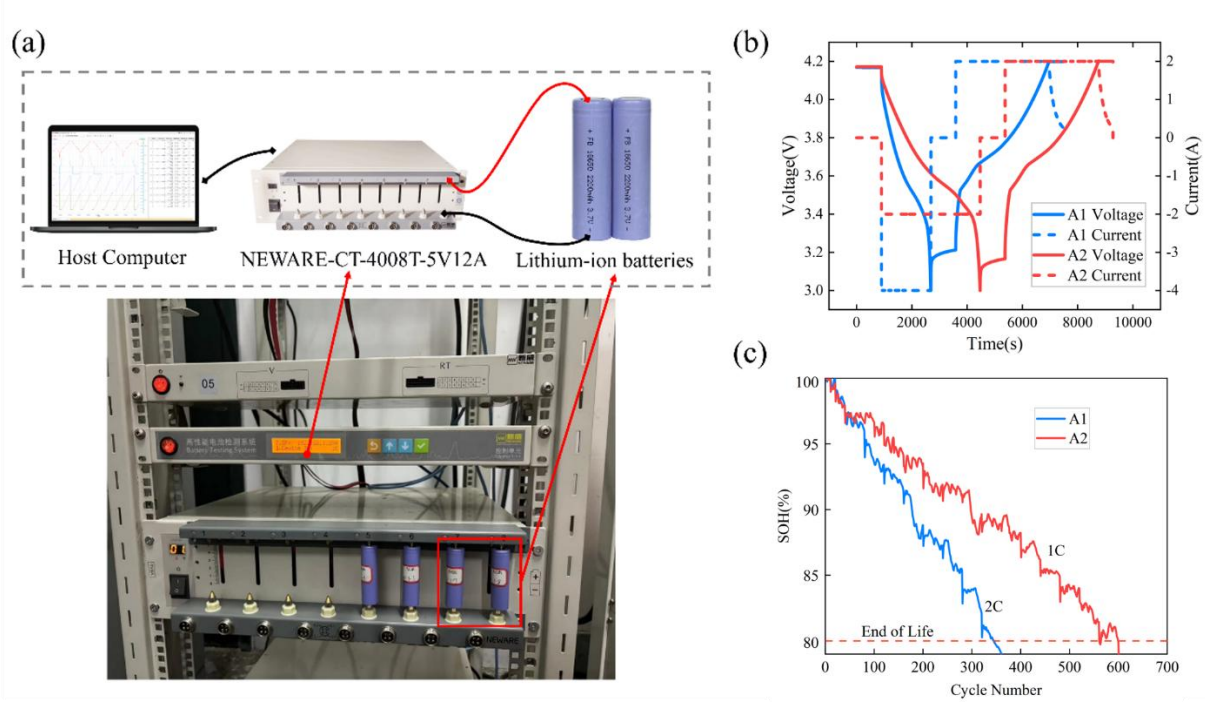

Fig 2. Full life cycle aging testing of A-series NCM batteries (a) Battery degradation platform in our laboratory for data collection of A1-2 batteries. (b) Charge-discharge protocols of two batteries. Solid lines represent the voltage curves of A1 and A2. Dashed lines represent the current curves of A1 and A2. (c) SOH degradation processes of A1 and A2 with different discharge rate.

2.1.2. NCA batteries under different discharge rate and temperature

The dataset is provided by the Prognostics Center of Excellence Data Set Repository at NASA [30,31], consisting of ten 18650 lithium-ion cylindrical batteries, sequentially named B1 through B10. The positive electrode material is lithium nickel cobalt aluminum oxide (NCA), differing from the A-series batteries. The SOH degradation curves are shown in **Fig. 3(a)**. The charging process for all batteries involves a constant current of 1.5A until the voltage reaches 4.2V, followed by constant voltage charging until the current decreases to 0.02A.

The primary differences among B1–B10 are as follows: (1) the operating temperature for B1–B7 is 24°C, while B8–B10 operate at 4°C; (2) the discharge rate for B5–B6 is 2C, whereas the remaining eight batteries discharge at 1C. These variations enable the evaluation of the model's performance under different discharge rates and operating temperatures.

3. Methodology

3.1. Health indicators (HIs) extraction and selection

The health indicators (HIs) are extracted from five typical curves: incremental capacity curve, charge-discharge voltage, and current curves. Each curve exhibits distinct degradation patterns. By extracting 18 HIs, as shown in **Table 2**, this paper aims to reveal the degradation patterns of different batteries under various operation conditions. Considering that batteries frequently operate in the mid-SOC range during daily use, we focus on the charge segment from 3.8V to 4.2V and the discharge segment from 3.8V to 3.4V for health indicator extraction. The time range for each cycle within the selected intervals is defined as $[t_{\text{start}},\ t_{\text{end}}]$, with the following calculation formulas for charging and discharge time, charge capacity, mean, variance, kurtosis, skewness, slope, and Shannon entropy:

$$\Delta t = t_{\text{end}} - t_{\text{start}} \tag{1}$$

$$\Delta Q = \int_{t_{\text{start}}}^{t_{\text{end}}} I \cdot dt \tag{2}$$

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x(i) \tag{3}$$

$$\sigma^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(x(i)-\bar{x}\right)^2 \tag{4}$$

$$\text{kurtosis} = \frac{\sum_{i=1}^{n}\left(x(i)-\bar{x}\right)^4}{(n-1)\sigma^4} \tag{5}$$

$$\text{skewness} = \frac{\sum_{i=1}^{n}\left(x(i)-\bar{x}\right)^3}{(n-1)\sigma^3} \tag{6}$$

$$\text{slope} = \frac{x_{\text{end}} - x_{\text{start}}}{t_{\text{end}} - t_{\text{start}}} \tag{7}$$

$$\text{entropy} = -\sum_{i=1}^{n} p_i \cdot \log_2(p_i) \tag{8}$$

where p_i represents the probability of each distinct voltage (or current) value in the normalized time series, and n represents the time-series size in current cycle.

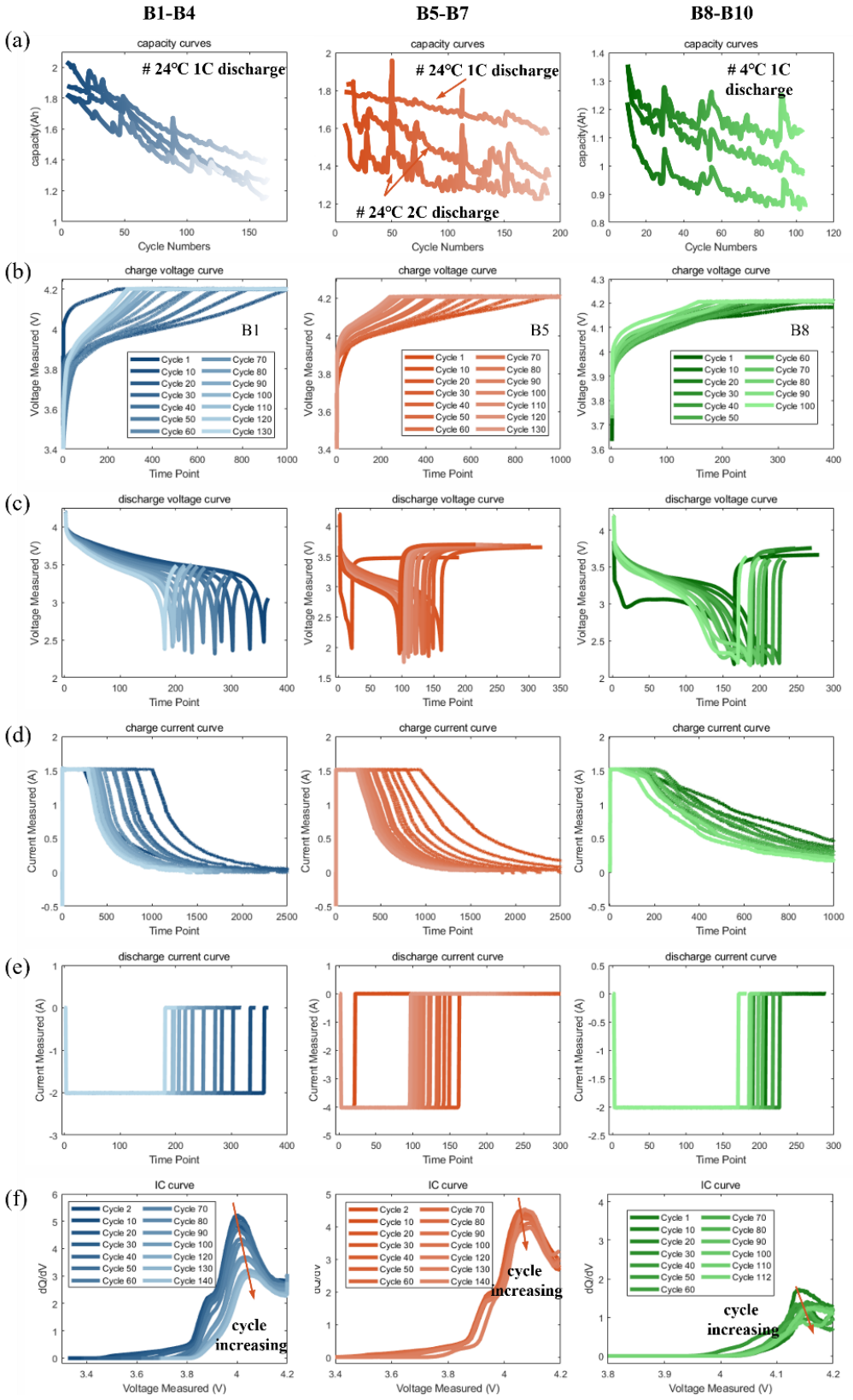

Fig. 3. Degradation curves of (a) capacity, (b) charge voltage, (c) discharge voltage, (d) charge current, (e) discharge current, and (f) incremental capacity (IC) for ten cells tested under different temperature (4 deg C or 24 deg C) or discharge rate (1C or 2C).

Table 2. Multi-source Health Indicators

No.	Description on Health Indicators
HI1	Constant current charge capacity within the 3.8-4.2V charge segment
HI2	Constant current charge time within the 3.8-4.2V charge segment
HI3	Constant voltage charge capacity within the 3.8-4.2V charge segment
HI4	Constant voltage charge time within the 3.8-4.2V charge segment
HI5	Constant current discharge time within the 3.8-3.4V discharge segment
HI6	Peak value of the IC curve within the 3.4-4.2 V charge segment
HI7	Voltage means within the 3.8–4.2 V charge segment
HI8	Voltage variance within the 3.8-4.2V charge segment
HI9	Voltage kurtosis within the 3.8-4.2V charge segment
HI10	Voltage skewness within the 3.8-4.2V charge segment
HI11	Voltage entropy within the 3.8-4.2V charge segment
HI12	Voltage slope within the 3.8-4.2V charge segment
HI13	Current means within the 3.8–4.2 V charge segment
HI14	Current variance within the 3.8-4.2V charge segment
HI15	Current kurtosis within the 3.8-4.2V charge segment
HI16	Current skewness within the 3.8-4.2V charge segment
HI17	Current entropy within the 3.8-4.2V charge segment
HI18	Current slope within the 3.8-4.2V charge segment

Fig. 3(a) illustrates the SOH degradation of the ten B-series batteries under different discharge rates and temperatures as the cycle number increases. In **Figs. 3(b)–(e)**, both the charging and discharging voltage and current curves shift to the left with an increasing number of cycles, indicating a gradual reduction in the batteries' operating time. **Fig. 3(f)** depicts the degradation of IC curves under varying discharge rates and temperatures. As the cycle number increases, the IC curves shift toward the lower-right corner, reflecting a gradual decrease in IC peak values. Notably, at low temperatures, the charging and discharging voltage and current curves become denser with narrower variations. This is attributed to increased polarization effects, higher internal resistance, and reduced efficiency of charge transfer and ion migration within the battery under low-temperature conditions.

Fig. 4(a)-(b) presents the Pearson correlation coefficients between HIs and true SOH, showing the influence of different discharge rates and temperatures on these HIs. Under the conditions of room temperature and a 1C discharge rate (Series 2: B1-B4), the lithium-ion migration efficiency and the reaction rate of active materials in the battery are relatively ideal. As the cycle number increases, a clear degradation trend becomes evident, resulting in strong correlations between most of the extracted indicators. Among these, HI5, HI6, HI17, and HI18 show significant potential as model inputs.

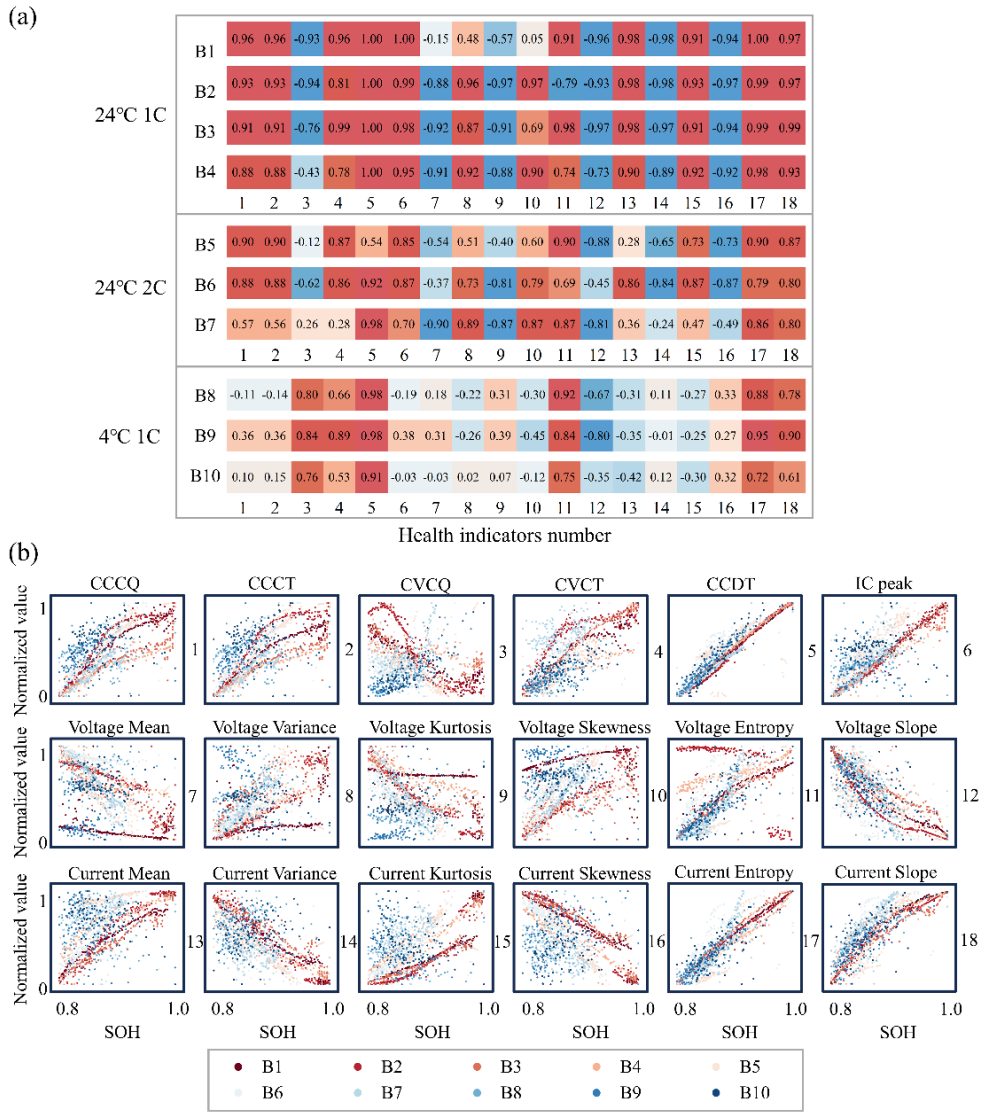

Fig. 4. Pearson correlation coefficients between the extracted HIs and SOH, shown through (a) a heatmap for an overview of the correlation strengths, and (b) eighteen scatter plots to depict the detailed relationships between HIs and SOH.

When the discharge rate increases to 2C (Series 3: B6-B7), the higher current density leads to increased internal heat accumulation, which in turn intensifies the non-linearity of the battery's reactions. This affects the correlations between certain features. In this case, HI6, HI17, and HI18 remain suitable as model inputs. Under low temperature conditions (4°C) in Series 4 (B8-B10), the number of indicators with a high correlation to SOH significantly decreases. This is due to the fact that low temperatures considerably slow down the electrochemical reaction rate within the battery, resulting in less pronounced voltage fluctuations and reducing the correlation between IC peak values and SOH. However, indicators such as HI3, HI5, and HI11 still perform well in terms of their correlation with SOH and can be used as model inputs.

3.2. Self-attention neural network construction

Self-attention mechanism calculates the similarity between each position in the input time-series sequence and all other positions, assigning different attention weights to emphasize the most relevant features. This mechanism, can be mathematically represented as follows:

$$\text{Attention}(Q,K,V)=\text{Softmax}\left(QK^T/\sqrt{d_k}\right)V \tag{9}$$

Here, Q, K, and V represent the query, key, and value matrices, respectively, obtained through linear transformations of the input sequence. The term d_k denotes the dimensionality of the key matrix and is used to scale the dot product to prevent excessively large values. The Softmax function converts the dot product results into a probability distribution, thereby determining the attention weights.

As shown in **Fig. 5**, self-attention neural network is designed to focus on features from different subspaces simultaneously by employing the multi-head self-attention mechanism. After dividing the input into multiple heads, each head performs independent attention computations, and the results are then concatenated. This approach allows the model to learn more diverse feature representations. The formula for the multi-head self-attention mechanism is as follows:

$$\text{MultiHead}(Q,K,V)=\text{Concat}(\text{head}_1,...,\text{head}_h)W^O \tag{10}$$

Here, the output of each head is $\text{head}_i=\text{Attention}(QW_i^Q,KW_i^K,VW_i^V)$, the results from each head are concatenated along the last dimension using the 'Concat' operation of Python. The concatenated output is then projected back to the original input dimension through a weight matrix W^O, typically with size $\mathbb{R}^{(h\cdot d_v)\times d_{\text{model}}}$, where h represents the number of attention heads, d_v is the output dimension of each head, and d_{model} is the original embedding dimension of the model.

Since the self-attention-based model does not involve iterative operations like RNNs, positional encoding is used to provide the model with positional information for each element in the time series. Positional encoding generates a unique encoding for each position in the sequence by applying linear transformations of sine and cosine functions. The formula is as follows:

$$PE(pos,2i)=\sin(pos/10000^{2i/d_{\text{model}}}) \tag{11}$$

$$PE(pos,2i+1)=\cos(pos/10000^{2i/d_{\text{model}}}) \tag{12}$$

Here, *pos* refers to the position of a specific element within a time series segment, i is the index of the dimension, d_{model} represents the original embedding dimension of the model's input.

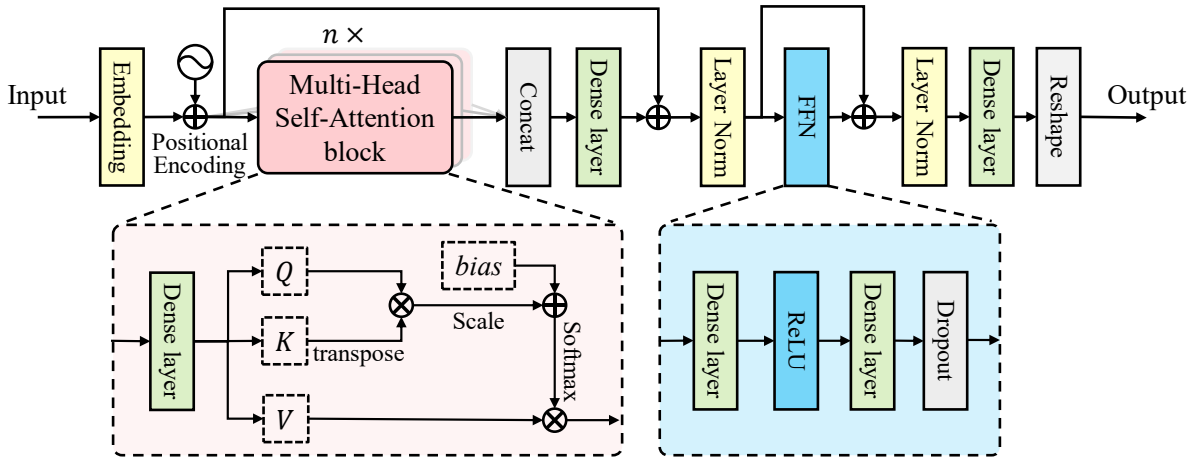

Fig. 5. Structure of self-attention-based model.

3.3. Deep transfer learning based on self-attention

This paper aims to achieve fast online SOH estimation of a new battery with few early data under non-consistent operating conditions and materials. To achieve this goal, this study adopts a deep transfer learning approach, which utilizes a model pre-trained on source domain data and fine-tunes the model parameters of upper layer using a small amount of target domain data.

The model involves optimization of the objective function in both the source domain pretraining phase and the transfer learning fine-tuning phase in the target domain.

During pretraining on the source domain $\mathcal{D}_S$, the model learns a set of parameters θ_S, including weights W and bias b, by minimizing the loss function $\mathcal{L}(f_S(X_S;\theta_S),Y_S)$ on the source domain data, This process enables the model to better fit the data features through multiple linear equations and nonlinear activation functions. The objective function optimization in the source domain can be simplified as follows:

$$\min_{\theta_S} \mathbb{E}_{(X_S,Y_S)\sim\mathcal{D}_S}\left[\mathcal{L}(f_S(X_S;\theta_S),Y_S)\right] \tag{13}$$

In the fine-tuning phase on the target domain $\mathcal{D}_T$, the pretrained model parameters θ_S are used as initial values. The model further optimizes the parameters θ_T by minimizing the loss function $\mathcal{L}(f_T(X_T;\theta_T),Y_T)$ on the target domain data. This process applies the knowledge learned from the source domain to the target task, adjusting the parameters to better fit the target task. The optimization of the objective function in the

target domain can be simplified as follows:

$$\min_{\theta_T} \mathbb{E}_{(X_T,Y_T)\sim\mathcal{D}_T}\left[\mathcal{L}(f_T(X_T;\theta_T),Y_T)\right] \tag{14}$$

here, θ_T represents the trainable parameters of the model for the target domain T, while $(X_T,Y_T)\sim\mathcal{D}_T$ indicates that the input data X_T and the corresponding ground-truth labels Y_T are sampled from the target domain distribution $\mathcal{D}_T$. The function $f_T(X_T;\theta_T)$ denotes the model's prediction given the input X_T and parameters θ_T. The loss function $\mathcal{L}$ quantifies the discrepancy between the predicted output and the true value Y_T, and the expectation operator $\mathbb{E}$ averages this loss over all samples from $\mathcal{D}_T$. The overall objective is to find the optimal parameters θ_T that minimize this expected loss.

The minimization in Eq. (14) is performed with the Adam optimizer of Python, which adaptively scales each parameter's update using estimates of the first and second moments of the gradients. In each iteration, a mini-batch $\left\{\left(X_T^{(b)},Y_T^{(b)}\right)\right\}_{b=1}^{B}$ is sampled from $\mathcal{D}_T$, and the batch gradient is computed as

$$g_t=\nabla_{\theta_T}\frac{1}{B}\sum_{b=1}^{B}\mathcal{L}\left(f_T\left(X_T^{(b)};\theta_T\right),Y_T^{(b)}\right) \tag{15}$$

Adam then updates its biased moment estimates:

$$m_t=\beta_1 m_{t-1}+(1-\beta)g_t \tag{16}$$

$$v_t=\beta_2 v_{t-1}+(1-\beta_2)g_t^2 \tag{17}$$

applies bias correlations

$$\hat{m}_t=\frac{m_t}{1-\beta_1^t} \tag{18}$$

$$\hat{v}_t=\frac{v_t}{1-\beta_2^t} \tag{19}$$

and finally adjusts the parameter via

$$\theta_T\leftarrow\theta_T-\alpha\frac{\hat{m}_t}{\sqrt{\hat{v}_t+\varepsilon}} \tag{20}$$

where α is the base learning rate, β_1 and β_2 control the decay of the moment estimates, and ε (typically 1e-8) prevents division by zero. This procedure converges rapidly and robustly to a minimum of the expected target-domain loss, even when data are noisy or limited.

4. **Experimental results and discussion**

In this section, the effectiveness and generalization capability of the proposed method are validated through experiments conducted on battery datasets A and B, which exhibit different materials, discharge rates, and operating temperatures.

4.1. Experiments setup

Due to the self-attention based model's requirement for smaller input segments, a sliding window approach is employed to segment the time-series data. Given the larger dataset size of Battery A compared to Battery B, a window size of 10 is used for Battery A series, while a window size of 5 is applied to Battery B series, with a step size of 1 in both cases. This method processes the input variables HI5, HI17, and HI18, while the output variable represents the current SOH of the battery. The SOH is calculated based on the rated capacity using the following formula:

$$\mathrm{SOH} = \frac{C_{\mathrm{current}}}{C_{\mathrm{rated}}} \times 100\%, \tag{21}$$

where C_{current} and C_{rated} are the current capacity and rated capacity, respectively.

The experimental process consists of three steps:

(1) Pretraining: For each condition, the first battery's data is divided into 70% for training and 30% for validation. The model is trained using a batch size of 16, and hyperparameters are tuned on the validation set. The final trained model serves as the pretrained model for transfer learning.

(2) Fine-tuning and Accuracy Testing: The pretrained model is fine-tuned using the first 10% of data from the second battery, while the remaining 90% is used to evaluate accuracy.

(3) Generalization Testing: The remaining batteries in the same condition are used as test sets to evaluate the model's generalization capability.

For example, the data from battery A1 is split into 70% for training and 30% for validation to pretrain the model. The pretrained model is fine-tuned using the first 10% of A2's data, and the remaining 90% is used for testing. In contrast, comparative models are trained without pretraining on A1 data. The learning rate and number of fine-tuning epochs are set to 1e-5 and 200, respectively.

The experiments are carried on an ASUS TUF A15 computer, which is equipped with an AMD Ryzen R7-4800H (2.90GHz), an NVIDA GeForce RTX 2060 and 24GB of RAM. The operating system is

Microsoft Windows 11 Professional. The prediction models are built by PyTorch 2.0.0, utilizing Python 3.9.16 as the programming language.

Two widely adopted evaluation metrics are selected to assess the models' performance: Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE) [32]. The mean values of all experimental results for each battery are labeled as the Average Mean Absolute Error (AMAE) and Average Root Mean Squared Error (ARMSE) to further assess the model's overall performance. The formulas for these metrics are as follows:

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n}\left|y_i - \hat{y}_i\right| \tag{22}$$

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \tag{23}$$

$$\mathrm{AMAE} = \frac{1}{m}\sum_{j=1}^{m}\frac{1}{n}\sum_{i=1}^{n}\left|y_i - \hat{y}_i\right| \tag{24}$$

$$\mathrm{ARMSE} = \frac{1}{m}\sum_{j=1}^{m}\sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(y_i - \hat{y}_i\right)^2} \tag{25}$$

where y, $\hat{y}$, n and m are the actual SOH, predicted SOH, the number of testing samples, and the number of experiments, respectively. The subscripts i denotes the i-th data point in the time-series of actual and predicted SOH, while the subscripts j indicates the j-th model testing experiment.

4.2. SOH prognostic for battery series A

This section presents the performance of the proposed method on Battery Series 1 (A1-A2), with the experimental results shown in **Fig. 6(a)** illustrates the line chart of estimation results using different training methods, including a SDTL model pre-trained on A1 and fine-tuned on 10% of A2 data (labeled as Proposed), as well as a Transformer model without fine-tuning (labeled as Transformer). **Fig. 6(b)** shows scatter plots comparing the predicted and actual SOH values, where points closer to the diagonal line indicate higher estimation accuracy. **Fig. 6(c)** depicts the estimation errors over the cycle numbers, highlighting deviations in predictions.

The MAE and RMSE of the Transformer model are 0.0044 and 0.0061, respectively. When using SDTL, the MAE and RMSE improved to 0.0018 and 0.0023, respectively. The experimental results illustrate that the fine-tuning technique in transfer learning achieves superior performance on Battery Series 1 with limited

data support (10%). More specifically, the proposed method can accurately estimate whether the battery's SOH has reached the EOL threshold (0.8), which is critical for mitigating risks and losses associated with battery performance degradation.

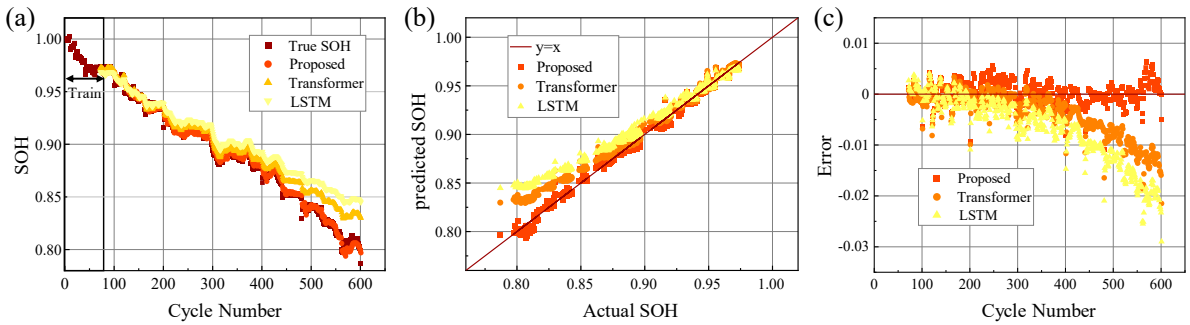

Fig. 6. Online SOH estimation results and errors for A2 (NCM) under 1C/24℃ conditions. (a) Line chart of estimation results using different training methods. (b) Scatter plots comparing predicted vs. actual SOH values. (c) Estimation errors as a function of cycle number.

To minimize the influence of random data partitioning variability, five experiments were conducted for each of the three models. The experimental results depicted in the bar charts shown in **Fig. 7**, show a comparative analysis of three models—SDTL, Transformer, and LSTM—across five trials, evaluated using Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE).

In the MAE analysis (left panel), the SDTL model outperformed the other two models across all trials, exhibiting the lowest error values. Specifically, the MAE for SDTL remained below 0.005 in all trials, while the Transformer model showed moderately higher errors (approximately 0.015), and the LSTM model showed the highest MAE values (around 0.025). This trend was similarly observed in the MSE analysis (right panel), where SDTL maintained the lowest error rates (below 0.0005), followed by the Transformer (around 0.0015) and LSTM (approximately 0.0025).

These results indicate that the SDTL model achieved superior predictive accuracy and robustness compared to the Transformer and LSTM models, as evidenced by lower MAE and MSE values across all experimental trials. The LSTM model, in contrast, exhibited the highest error rates, suggesting lower efficacy in this context. The performance of the SDTL model across multiple trials underscores its reliability and effectiveness in the evaluated tasks.

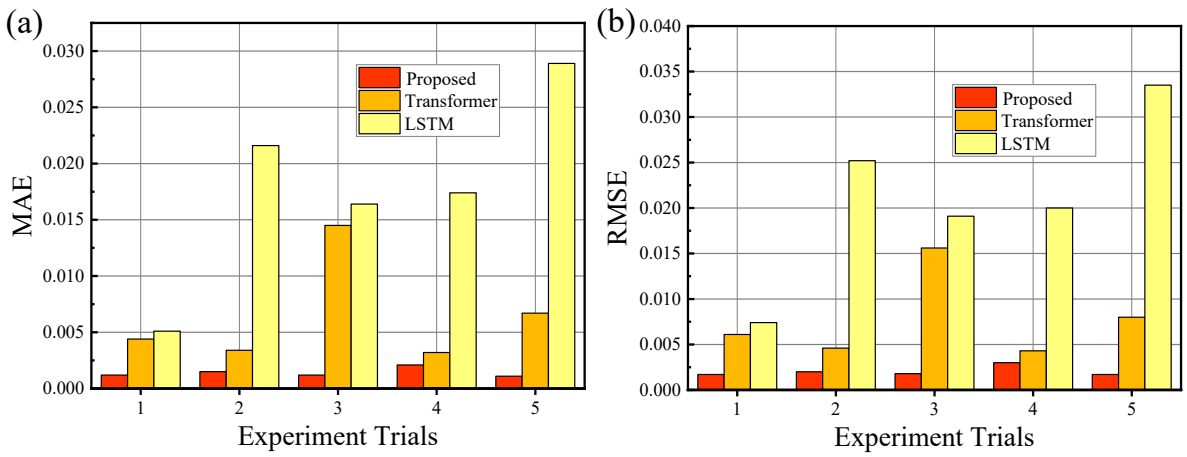

Fig. 7. Estimation errors of three methods on (a) MAE and (b) RMSE over five experiment trials

4.3. SOH prognostic for battery series B

To further investigate the effectiveness and generalization capability of proposed method under different battery anode materials, charge-discharge rates, and ambient temperatures, this section utilizes B-series battery datasets under three kinds of operation conditions, including 1C charge-discharge rate at 24°C, 2C charge-discharge rate at 24°C, and 1C charge-discharge rate at 4°C. Unlike the NCM material used in Battery A, these datasets feature NCA as the cathode material, enabling validation of the proposed method's applicability across different battery materials and operating conditions.

For clarity, the first battery in each B series (B1, B5, and B8) is divided into training and validation sets at a 7:3 ratio. Subsequently, the first 10% of data from the second battery in each series (B2, B6, B9) is used to fine-tune the model, which is then tested on the remaining 90% of the data. The model's performance is further evaluated through online SOH estimation and generalization validation on the remaining batteries (B3, B4, B7, and B10).

Table 3. SOH estimation errors on B-series batteries with three different approaches.

Series	Batteries	Proposed		Transformer		LSTM	
		AMAE	ARMSE	AMAE	ARMSE	AMAE	ARMSE
2	B2	**0.0214**	**0.0262**	0.0274	0.0295	0.0435	0.0439
	B3	**0.0108**	**0.0137**	0.0291	0.0331	0.0568	0.0614
	B4	**0.0143**	**0.0187**	0.0180	0.0235	0.0480	0.0520
3	B6	**0.0218**	**0.0319**	0.0272	0.0356	0.0506	0.0561
	B7	**0.0122**	**0.0176**	0.0149	0.0205	0.0279	0.0324
4	B9	**0.0092**	**0.0128**	0.0124	0.0148	0.0169	0.0209

B10	**0.0134**	**0.0187**	0.0138	0.0193	0.0214	0.0255

The results of three models on B series batteries are shown in **Table 3**. The metrics used for evaluation are the average Mean Absolute Error (AMAE) and the average Root Mean Square Error (ARMSE) from five experiments. The batteries are categorized into three series based on their operating conditions: Series 2 (B2-B4) operates at 1C and 24°C, Series 3 (B6-B7) at 2C and 24°C, and Series 4 (B9-B10) at 1C and 4°C. For Series 2 batteries, SDTL shows competitive performance, with AMAE values of 0.0214 for B2, 0.0108 for B3, and 0.0143 for B4, and ARMSE values of 0.0262, 0.0137, and 0.0187, respectively. These results are comparable to those of the Transformer and LSTM models, indicating that SDTL performs effectively under moderate operating conditions.

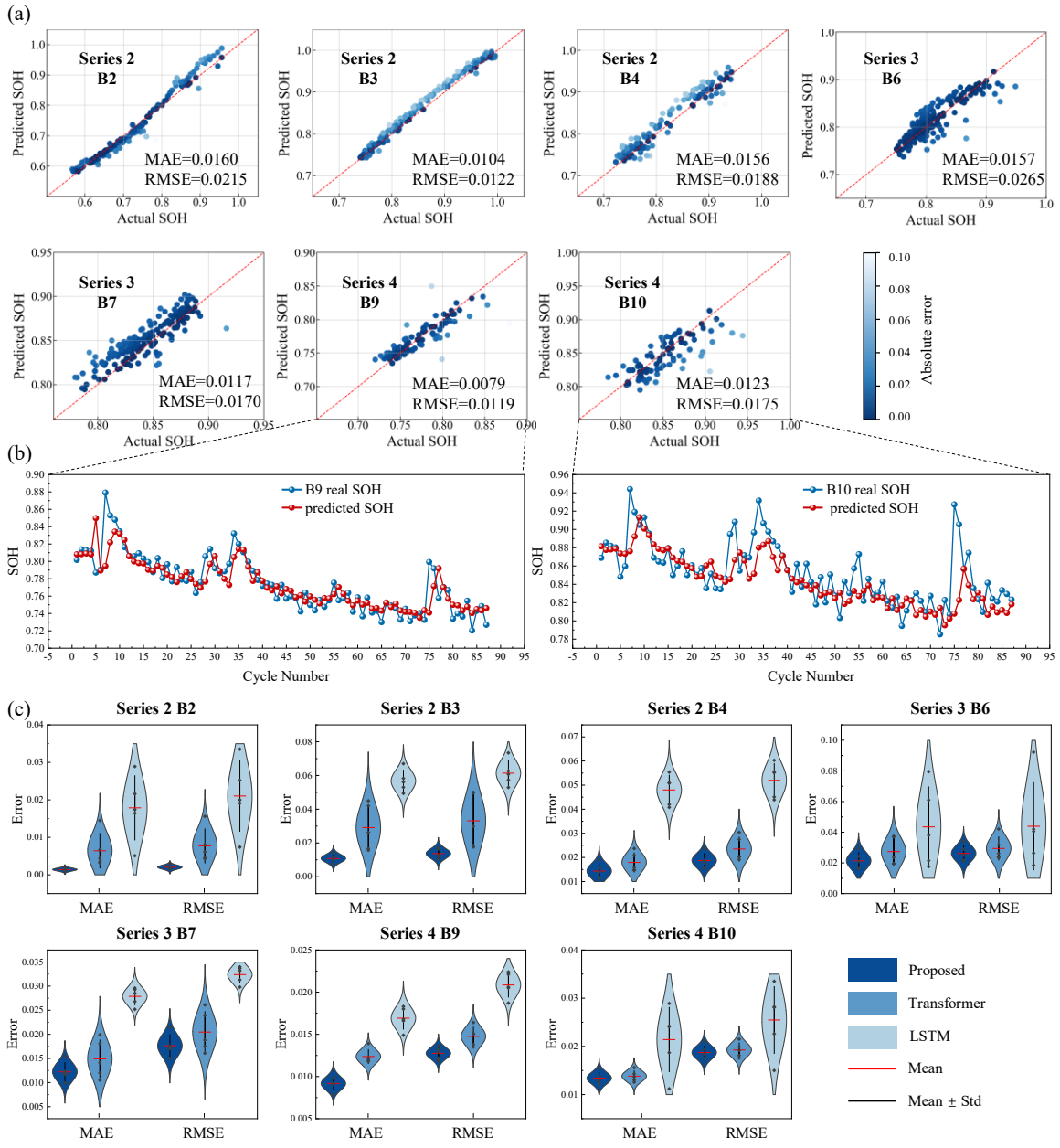

Fig. 8. SOH estimation results. (a) Estimation results of the proposed methods across three battery series under different work conditions. (b) Comparison between the estimated and true SOH for batteries B9 and B10 under 1C/4℃ conditions. (c) Distributions of mean absolute error (MAE) and root mean square error (RMSE) of 3 models (the proposed SDTL (Proposed), Transformer, LSTM)

In Series 3, which includes batteries operating at a higher rate (2C) and the same temperature (24°C), SDTL again shows robust performance. For B6, SDTL achieves an AMAE of 0.0218 and an ARMSE of 0.0319, while for B7, the AMAE is 0.0122 and the ARMSE is 0.0176. These results are superior to those of the Transformer and LSTM models, particularly for B7, where SDTL significantly outperforms the others.

Fig 8(a) further illustrates the distribution of predicted versus actual values, demonstrating that the predictions from the proposed method align closely with the diagonal, indicating superior estimation accuracy. Specifically, the proposed method achieves an average RMSE of 0.0319 on the high discharge rate condition (B6), with accuracy improvements of 10.39% and 43.14% over the two alternative models, respectively.

The B9 and B10 batteries, operating are utilized to validate the model performance under low-temperature conditions. The SOH estimation fitting performance of the proposed method is illustrated in Fig. 8(b). The average RMSE of the proposed SDTL is 0.0128 and 0.0187, respectively. This represents an accuracy improvement of 13.51% and 3.11% compared to Transformer, and 38.76% and 26.67% compared to LSTM. To further evaluate the stability of three models, the training and testing processes for each model are repeated ten times. The experimental results, as shown in Fig. 8(c), indicate that the prediction performance of Transformer and LSTM exhibits relatively poor stability. In contrast, the proposed SDTL model outperforms all others under various task conditions. This robustness is particularly critical for ensuring battery safety in practical applications.

4.4. SOH prognostic for different TL methods

Sections 4.2 and 4.3 have compared the performance of various deep learning models under different operating conditions. This subsection further extends the analysis by introducing a comprehensive comparison of current deep transfer learning techniques, aiming to validate the superiority of the proposed SDTL-based SOH estimation method.

The selected baseline methods include Domain-Adversarial (DAAD), Domain Adaptation (DAAP), fine-tuned Transformer (Transformer-TL), fine-tuned LSTM (LSTM-TL), and the proposed SDTL method. Specifically, DAAD adopts the commonly used Gradient Reversal Layer (GRL) to help the model learn domain-invariant features, while DAAP applies the widely adopted Maximum Mean Discrepancy (MMD) strategy to reduce the distribution discrepancy between source and target domains. For fine-tuning-based methods, all models share the same base architecture, with only the final layer parameters updated during transfer. For convenience, all methods are trained using the same health indicator HI5. In fine-tuning methods, the B1 dataset is used for pre-training, followed by fine-tuning with the first 10% of B2 data. The performance of all five methods is then evaluated on the B3 and B4 battery datasets. During training, a batch size of 32, a learning rate of 1e-3, and 200 epochs are used. For fine-tuning, the learning rate is set to 1e-5

with 100 training epochs.

In general, as shown in Fig 9, the proposed method outperforms the other four methods across all five trials on both B3 and B4 datasets. Compared to DAAD and DAAP, which exhibit larger errors and greater variability across different experimental trials, the proposed SDTL method achieves both higher accuracy and greater stability. Although Transformer-TL and LSTM-TL show more consistent performance than DAAD and DAAP, their error rates remain noticeably higher than those of the proposed method.

Specifically, as described in Table 3, the proposed method achieves the lowest average MAE and RMSE values, with 0.00536 and 0.00736 on B3, and 0.01238 and 0.01640 on B4, respectively, outperforming other four methods. Compared to Transformer-TL, which shows the next best performance, the proposed method reduces MAE and RMSE on B3 by 39.2% and 32.4%, respectively. On the more challenging B4 dataset, it still achieves a 12.0% reduction in MAE and a 5.9% reduction in RMSE. These results indicate that the proposed SDTL method not only provides high estimation accuracy but also shows good generalization and robustness when applied to different target domain.

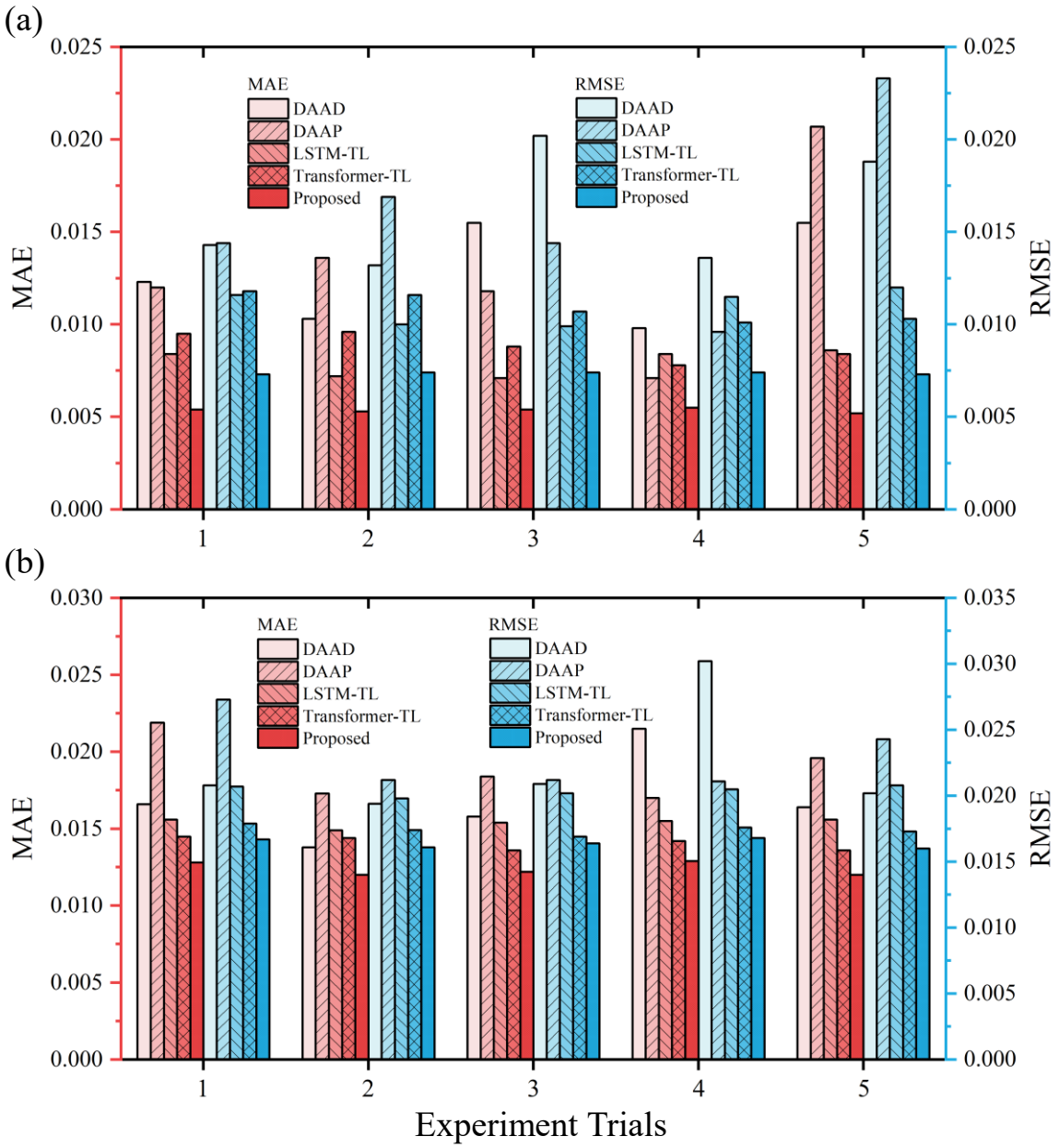

Fig. 9. Comparison of MAE and RMSE for five SOH estimation methods (DAAD, DAAP, Transformer-TL, LSTM-TL, SDTL) on (a) B3 and (b) B4, respectively.

Table 3. Average MAE and RMSE of five SOH estimation methods over five trials on B3 and B4 batteries.

Methods	B3		B4		Average	
	AMAE	ARMSE	AMAE	ARMSE	MAE	RMSE
DAAD	0.01268	0.01602	0.01682	0.02230	0.01475	0.01916
DAAP	0.01304	0.01572	0.01884	0.02302	0.01594	0.01937
LSTM-TL	0.00794	0.01100	0.01540	0.02040	0.01167	0.01570
Transformer-TL	0.00882	0.01090	0.01406	0.01742	0.01144	0.01416

Proposed	0.00536	0.00736	0.01238	0.01640	0.00887	0.01188

5. Conclusion

The present study introduces a self-attention-based Deep Transfer Learning (SDTL) approach for the online state-of-health (SOH) estimation of lithium-ion batteries under varying operating conditions, cathode materials, and ambient temperatures, particularly when faced with small sample sizes. The proposed approach leverages the strengths of deep transfer learning to address the challenges associated with limited early-stage data and differing operating environments, which are common in the practical application of battery health monitoring systems.

The experimental results, conducted on battery datasets A and B with different materials, discharge rates, and operating temperatures, illustrated the effectiveness and generalization capability of the proposed SDTL method. The method outperformed conventional Transformer and LSTM models, achieving superior accuracy with limited data support in these experiment trials. The robustness of SDTL against random data splits was also highlighted, which is crucial for ensuring battery safety in practical applications.

In comparison to other deep transfer learning techniques, such as Domain-Adversarial (DAAD), Domain Adaptation (DAAP), Transformer-TL and LSTM-TL, the proposed SDTL method showed superior performance in both accuracy and stability. While DAAD and DAAP exhibited higher variability and larger errors across trials, Transformer-TL and LSTM-TL, though more reliable, still lagged behind the proposed method in terms of error consistency. These findings highlight the SDTL method's superior ability to capture degradation trends and its stronger robustness and generalization across different conditions.

While the SDTL method has shown positive results, there are still several areas for future research to further enhance the approach: (1) Expanding the Dataset: Incorporating a broader range of batteries with different ages, chemistries, and from various manufacturers could further enhance the model's robustness and applicability across diverse scenarios. (2) Improving Domain Adaptation Strategies: Although the proposed method outperforms existing transfer learning techniques such as DAAD and DAAP, future work could explore more adaptive or task-specific strategies to further optimize performance under extreme data limitations or unseen domain shifts.

Reference

[1] B. Scrosati, J. Garche, Lithium batteries: Status, prospects and future, Journal of Power Sources 195 (2010) 2419–2430. https://doi.org/10.1016/j.jpowsour.2009.11.048.

[2] T. Kim, W. Song, D.-Y. Son, L.K. Ono, Y. Qi, Lithium-ion batteries: outlook on present, future, and hybridized technologies, J. Mater. Chem. A 7 (2019) 2942–2964. https://doi.org/10.1039/C8TA10513H.

[3] X. Li, M. Zhao, S. Zhong, J. Li, S. Fu, Z. Yan, BMSFormer: An efficient deep learning model for online state-of-health estimation of lithium-ion batteries under high-frequency early SOC data with strong correlated single health indicator, Energy 313 (2024) 134030. https://doi.org/10.1016/j.energy.2024.134030.

[4] F. Wang, Z. Zhai, Z. Zhao, Y. Di, X. Chen, Physics-informed neural network for lithium-ion battery degradation stable modeling and prognosis, Nat Commun 15 (2024) 4332. https://doi.org/10.1038/s41467-024-48779-z.

[5] Y. Wang, J. Tian, Z. Sun, L. Wang, R. Xu, M. Li, Z. Chen, A comprehensive review of battery modeling and state estimation approaches for advanced battery management systems, Renewable and Sustainable Energy Reviews 131 (2020) 110015. https://doi.org/10.1016/j.rser.2020.110015.

[6] X. Gu, K.W. See, P. Li, K. Shan, Y. Wang, L. Zhao, K.C. Lim, N. Zhang, A novel state-of-health estimation for the lithium-ion battery using a convolutional neural network and transformer model, Energy 262 (2023) 125501. https://doi.org/10.1016/j.energy.2022.125501.

[7] D. Roman, S. Saxena, V. Robu, M. Pecht, D. Flynn, Machine learning pipeline for battery state of health estimation, (2021). http://arxiv.org/abs/2102.00837 (accessed June 5, 2024).

[8] J. Sihvo, T. Roinila, D.-I. Stroe, SOH analysis of Li-ion battery based on ECM parameters and broadband impedance measurements, in: IECON 2020 The 46th Annual Conference of the IEEE Industrial Electronics Society, IEEE, Singapore, Singapore, 2020: pp. 1923–1928. https://doi.org/10.1109/IECON43393.2020.9254859.

[9] J. Li, K. Adewuyi, N. Lotfi, R.G. Landers, J. Park, A single particle model with chemical/mechanical degradation physics for lithium ion battery State of Health (SOH) estimation, Applied Energy 212 (2018) 1178–1190. https://doi.org/10.1016/j.apenergy.2018.01.011.

[10] D. Wang, Q. Zhang, H. Huang, B. Yang, H. Dong, J. Zhang, An electrochemical–thermal model of

lithium-ion battery and state of health estimation, Journal of Energy Storage 47 (2022) 103528. https://doi.org/10.1016/j.est.2021.103528.

[11] A. Jokar, B. Rajabloo, M. Désilets, M. Lacroix, Review of simplified Pseudo-two-Dimensional models of lithium-ion batteries, Journal of Power Sources 327 (2016) 44–55. https://doi.org/10.1016/j.jpowsour.2016.07.036.

[12] Y. Zhang, R. Xiong, H. He, M.G. Pecht, Long Short-Term Memory Recurrent Neural Network for Remaining Useful Life Prediction of Lithium-Ion Batteries, IEEE Trans. Veh. Technol. 67 (2018) 5695–5705. https://doi.org/10.1109/TVT.2018.2805189.

[13] J. Hong, W. Wang, Q. Chai, Q. Lin, F. Cai, State of health estimation of lithium-ion battery based on DAGRU, in: Fourth International Conference on Artificial Intelligence and Electromechanical Automation (AIEA 2023), SPIE, 2023: pp. 608–613. https://doi.org/10.1117/12.2684768.

[14] A. Vaswani, N.M. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A.N. Gomez, L. Kaiser, I. Polosukhin, Attention is All you Need, in: 2017. https://www.semanticscholar.org/paper/204e3073870fae3d05bcbc2f6a8e263d9b72e776 (accessed August 26, 2024).

[15] C. Qian, B. Xu, L. Chang, B. Sun, Q. Feng, D. Yang, Y. Ren, Z. Wang, Convolutional neural network based capacity estimation using random segments of the charging curves for lithium-ion batteries, Energy 227 (2021) 120333. https://doi.org/10.1016/j.energy.2021.120333.

[16] H. Xu, L. Wu, S. Xiong, W. Li, A. Garg, L. Gao, An improved CNN-LSTM model-based state-of-health estimation approach for lithium-ion batteries, Energy 276 (2023) 127585. https://doi.org/10.1016/j.energy.2023.127585.

[17] D. Chen, W. Hong, X. Zhou, Transformer Network for Remaining Useful Life Prediction of Lithium-Ion Batteries, IEEE Access 10 (2022) 19621–19628. https://doi.org/10.1109/ACCESS.2022.3151975.

[18] C. Jia, Y. Tian, Y. Shi, J. Jia, J. Wen, J. Zeng, State of health prediction of lithium-ion batteries based on bidirectional gated recurrent unit and transformer, Energy 285 (2023) 129401. https://doi.org/10.1016/j.energy.2023.129401.

[19] C. She, Y. Li, C. Zou, T. Wik, Z. Wang, F. Sun, Offline and Online Blended Machine Learning for Lithium-Ion Battery Health State Estimation, IEEE Trans. Transp. Electrific. 8 (2022) 1604–1618. https://doi.org/10.1109/TTE.2021.3129479.

[20] X. Tang, C. Zou, K. Yao, J. Lu, Y. Xia, F. Gao, Aging trajectory prediction for lithium-ion batteries via model migration and Bayesian Monte Carlo method, Applied Energy 254 (2019) 113591. https://doi.org/10.1016/j.apenergy.2019.113591.

[21] J. Kim, D. Han, P.-Y. Lee, J. Kim, Transfer learning applying electrochemical degradation indicator combined with long short-term memory network for flexible battery state-of-health estimation, eTransportation 18 (2023) 100293. https://doi.org/10.1016/j.etran.2023.100293.

[22] S.J. Pan, Q. Yang, A Survey on Transfer Learning, IEEE Trans. Knowl. Data Eng. 22 (2010) 1345–1359. https://doi.org/10.1109/TKDE.2009.191.

[23] C. Tan, F. Sun, T. Kong, W. Zhang, C. Yang, C. Liu, A Survey on Deep Transfer Learning, (2018). http://arxiv.org/abs/1808.01974 (accessed October 21, 2024).

[24] J. Lu, Deep learning to estimate lithium-ion battery state of health without additional degradation experiments, Nature Communications (2023).

[25] G. Ma, S. Xu, T. Yang, Z. Du, L. Zhu, H. Ding, Y. Yuan, A Transfer Learning-Based Method for Personalized State of Health Estimation of Lithium-Ion Batteries, IEEE Trans. Neural Netw. Learning Syst. 35 (2024) 759–769. https://doi.org/10.1109/TNNLS.2022.3176925.

[26] Y. Yang, Y. Xu, Y. Nie, J. Li, S. Liu, L. Zhao, Q. Yu, C. Zhang, Deep transfer learning enables battery state of charge and state of health estimation, Energy 294 (2024) 130779. https://doi.org/10.1016/j.energy.2024.130779.

[27] S. Su, W. Li, J. Mou, A. Garg, L. Gao, J. Liu, A Hybrid Battery Equivalent Circuit Model, Deep Learning, and Transfer Learning for Battery State Monitoring, IEEE Trans. Transp. Electrific. 9 (2023) 1113–1127. https://doi.org/10.1109/TTE.2022.3204843.

[28] G. Dong, N. Hua, H. Chen, Y. Lou, Deep Transfer Learning Enabled State of Health Estimation of Lithium-Ion Battery Using Voltage Sample Entropy Under Fast Charging Profiles, IEEE Trans. Transp. Electrific. (2024) 1–1. https://doi.org/10.1109/TTE.2024.3445344.

[29] J. Yao, T. Han, Data-driven lithium-ion batteries capacity estimation based on deep transfer learning using partial segment of charging/discharging data, Energy 271 (2023) 127033. https://doi.org/10.1016/j.energy.2023.127033.

[30] I. Obisakin, C.V. Ekeanyanwu, State of Health Estimation of Lithium-Ion Batteries Using Support Vector Regression and Long Short-Term Memory, Open Journal of Applied Sciences 12 (2022) 1366–

1382. https://doi.org/10.4236/ojapps.2022.128094.

[31] J. Liu, X. Liu, An improved method of state of health prediction for lithium batteries considering different temperature, Journal of Energy Storage 63 (2023) 107028. https://doi.org/10.1016/j.est.2023.107028.

[32] D. Chicco, M. Warrens, G. Jurman, The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation, PeerJ Computer Science 7 (2021) e623. https://doi.org/10.7717/peerj-cs.623.