



# Deep transfer learning enables battery state of charge and state of health estimation

Yongsong Yang<sup>a,b</sup>, Yuchen Xu<sup>c</sup>, Yuwei Nie<sup>a</sup>, Jianming Li<sup>a</sup>, Shizhuo Liu<sup>a,b</sup>, Lijun Zhao<sup>a</sup>,  
Quanqing Yu<sup>a,\*</sup>, Chengming Zhang<sup>d,\*\*</sup>

<sup>a</sup> School of Automotive Engineering, Harbin Institute of Technology, Weihai, Shandong, 264209, China

<sup>b</sup> BYD Company Limited, Shenzhen, Guangdong, 518118, China

<sup>c</sup> Key Laboratory of Advanced Manufacturing Technology for Automobile Parts, Ministry of Education, Chongqing University of Technology, Chongqing, 400054, China

<sup>d</sup> School of Electrical Engineering and Automation, Harbin Institute of Technology, Harbin, Heilongjiang, 150001, China

## ARTICLE INFO

Handling editor: X Ou

### Keywords:

Lithium-ion battery  
State of charge  
State of health  
Joint estimation  
Deep learning  
Transfer learning

## ABSTRACT

In the realm of lithium-ion battery state estimation, traditional data driven approaches face challenges in accurately estimating state of charge and state of health throughout the battery's life cycle under dynamic working condition, and there is still a lack of research on models that can fulfill these requirements simultaneously. To address these issues, this study proposes an adaptive convolutional gated recurrent unit with Kalman filter for state of charge estimation throughout battery's full life cycle, leveraging transfer learning and deep learning techniques. Additionally, an adaptive convolutional gated recurrent unit with average post-processor is developed to estimate the battery state of health under dynamic working conditions, using voltage, current, temperature, state of charge, and accumulated discharge capacity as input features. Furthermore, a joint adaptive deep transfer learning model is proposed for simultaneously state of charge and state of health estimation through battery's full life cycle under dynamic working conditions. Experimental results validate the feasibility, accuracy, and robustness of the proposed models.

## 1. Introduction

The growing demand for electric vehicles is driven by environmental concerns and the depletion of fossil fuels [1,2]. Lithium-ion batteries have emerged as the preferred power source for electric vehicles due to their outstanding performance [3]. However, challenges in accurately estimating battery states within the battery management system hinder widespread adoption of electric vehicles [4]. State of charge (SOC) and state of health (SOH) are crucial indicators within battery management system, but their estimation requires indirect features derived from external battery characteristic parameters such as voltage, current, and temperature [5,6].

SOC estimation methods mainly includes the coulomb counting method [7], open circuit voltage method [8], analytical model method [9] and data driven method [10]. The first three methods failed to meet the needs for high accuracy and robustness requirements in dynamic and intense application conditions. In contrast, data driven methods, with its

ability to handle complex nonlinear problems, have gained popularity in SOC estimation, especially with the advancements in computational power, such as graphics processing unit [11]. SOH estimation methods mainly includes coulomb counting method [12], impedance spectroscopy method [13], analytical model method and data driven method. Similarly, the first three methods face challenges in actual dynamic conditions for their operational procedures and limited performance, and the data driven methods, on the other hand, offering a promising direction for SOH estimation [14,15].

Up to now, researchers have applied various specific data driven methods like support vector machine [16], back propagation neural network and recurrent neural network (RNN) to estimate SOC and SOH. Huang et al. [17] proposed an SOC estimation method using the conventional neural network (CNN) and gated recurrent unit (GRU) model, achieving a mean absolute error (MAE) of 1.68% under Supplemental Federal Test Procedure-US06 (US06) driving cycle, which is not satisfying. Yang et al. [18] proposed an adaptive convolutional neural

\* Corresponding author. School of Automotive Engineering, Harbin Institute of Technology, Weihai, No. 2, West Wenhua Road, High-tech District, Weihai, 264209, Shandong, China.

\*\* Corresponding author.

E-mail addresses: [qqyu@hit.edu.cn](mailto:qqyu@hit.edu.cn) (Q. Yu), [cmzhang@hit.edu.cn](mailto:cmzhang@hit.edu.cn) (C. Zhang).

<https://doi.org/10.1016/j.energy.2024.130779>

Received 5 December 2023; Received in revised form 5 February 2024; Accepted 21 February 2024

Available online 27 February 2024

0360-5442/© 2024 Elsevier Ltd. All rights reserved.

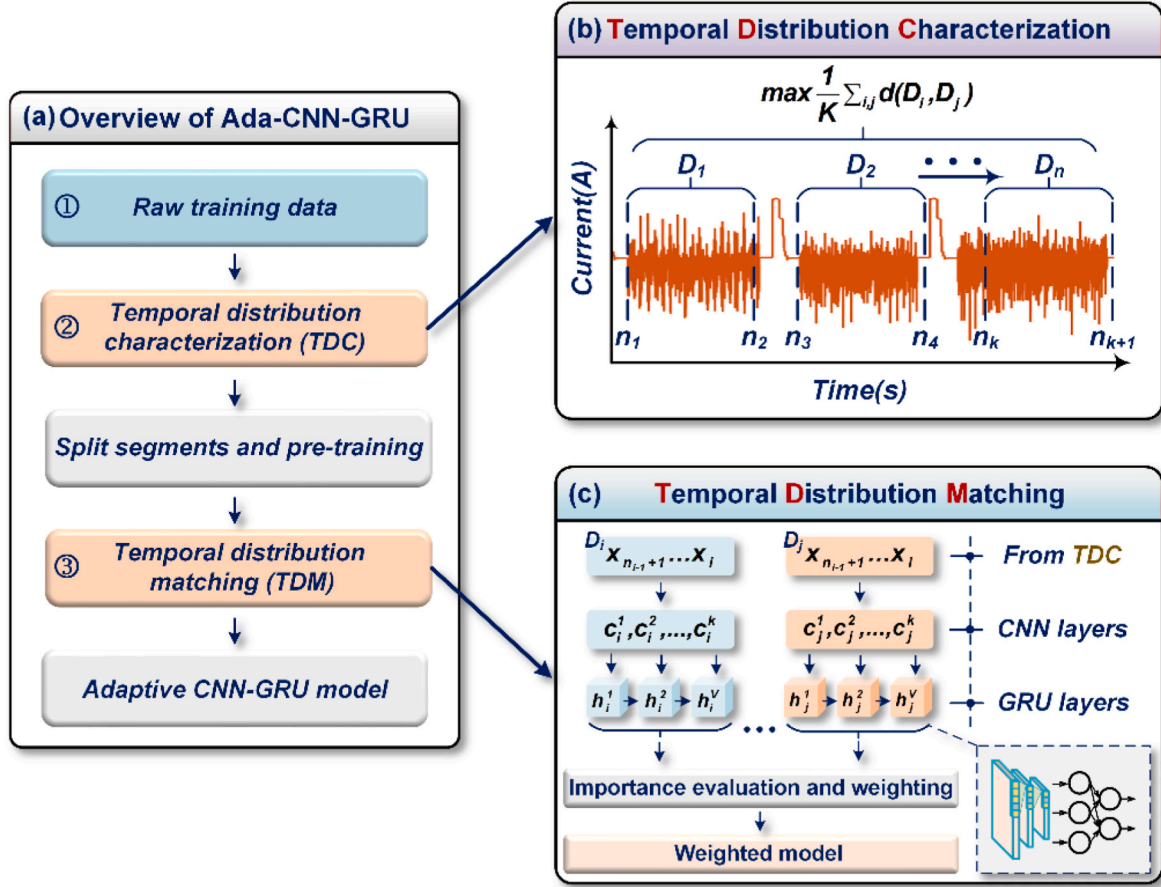


Fig. 1. Structure of Ada-CNN-GRU (a) Overview of Ada-CNN-GRU; (b) TDC; (c) TDM

network-gated recurrent unit with Kalman filter and feedback mechanism (Fb-Ada-CNN-GRU-KF) achieving an MAE around 0.8% under Urban Dynamometer Driving Schedule (UDDS) condition, but the model cannot be applied in the full life cycle of batteries for the model's complexity and lack of batteries' health state. For SOH estimation, Qiang et al. [19] achieved MAE of 0.3% utilizing features extracted from incremental capacity curve drawn from constant current constant voltage (CCCV) condition, which is unable to be applied in dynamic discharging process. Due to the strongly correlated relationship between SOC and SOH, various joint estimation methods have been designed and proposed [20–22]. Qiu et al. [23] estimated SOC and SOH jointly combining the backward smoothing square root cubature Kalman filter and the extended Kalman filter, but this method cannot be applied at intense working conditions. Li et al. [24] estimated SOC and SOH based on Gaussian process regression and CNN, which is easy to be applied in actual condition, but need additional expensive sensors to get more information.

In general, for SOC estimation, the conventional data driven methods are not applicable when considering battery degradation, and for SOH estimation, they are usually not suitable for dynamic discharging condition. In addition, the joint estimation methods mentioned above either cannot handle intense conditions or require additional sensors. To solve the problems mentioned above in data driven methods, a joint adaptive deep transfer learning (JADTL) model is proposed in this study to jointly estimate battery SOC and SOH. The contributions of this paper can be summarized as follows:

- (1) A novel SOC estimation model, Ada-CNN-GRU-KF, is proposed based on deep learning and transfer learning. This model achieves a balance between accuracy and computational burden by

considering both the distribution difference among segments and spatial features of the training datasets.

- (2) Similarly, a novel SOH estimation model, Ada-CNN-GRU-Ave, is proposed for dynamic working condition. This model takes into account the cross time scale characteristics of SOC and SOH, and takes dual time scale input and simplifies the traditional complex modeling process.
- (3) An SOC and SOH joint estimation model is proposed by combining the Ada-CNN-GRU-KF and Ada-CNN-GRU-Ave models. This model aims to achieve accurate and robust estimation of SOC and SOH throughout the full life cycle of batteries under dynamic working condition.

The following sections of this paper are arranged as follows: Section 2 introduces the methodology: the structure of the Ada-CNN-GRU, the Ada-CNN-GRU-KF model for SOC estimation, the Ada-CNN-GRU-Ave model for SOH estimation; and the JADTL model for joint estimation of SOC and SOH. Section 3 describes the battery datasets. Section 4 is the experimental evaluation of different models. Section 5 is the conclusion of this study.

## 2. Methodology

### 2.1. The architecture of the Ada-CNN-GRU

CNN is a deep learning neural network that offers the advantages of local connection, weight sharing, and down-sampling dimensional reduction for adaptive feature extraction [25]. GRU, as an enhanced RNN, tackles the issue of gradient disappearance [26] and enables units to adaptively capture sequence dependencies over varying time lengths

[27]. A CNN-GRU model comprises both CNN and GRU layers. At first, spatial information is extracted from different features by CNN, which is then fed into the GRU to capture the hidden temporal information within each feature's time series. This ensures comprehensive utilization of the raw experimental data.

Ada-RNN, proposed by Du et al. [28], is a deep transfer learning method. Initially, it identifies data segments with the greatest distribution differences in the training datasets. And by reducing the distribution discrepancies among these segments, higher levels of generalization and accuracy are achieved. Battery discharging process varies greatly under different working conditions, data from mild working condition can provide regular information and used to get stable model parameter, while data from intense working condition can provide various information and used to get a model suitable in different scene. Ada-RNN model can make a balance among different kind of data to get a model with high stability and generalization theoretically. Building upon Ada-RNN, Yang et al. [18] proposed Ada-CNN-GRU, which combines the advantages of CNN and GRU. The overview of Ada-CNN-GRU, illustrated in Fig. 1, consists primarily of two components: temporal distribution characterization (TDC) and temporal distribution matching (TDM).

The TDC algorithm aims to identify data segments with the most significant distribution differences and separate them based on the principle of maximum entropy. Given the clear distribution differences observed in battery data collected during laboratory experiments, which can be readily distinguished by working condition, manual splitting method will be adopted instead of a complex greedy algorithm. TDM, of which the structure is shown on the right of Fig. 1, is responsible for matching those split data segments and building a prediction model, and it is achieved through the minimization of loss function Eq. (1):

$$L(\theta, \alpha) = L_{pred}(\theta) + \lambda \frac{2}{K(K-1)} \sum_{i,j}^{i \neq j} L_{tdm}(D_i, D_j; \theta, \alpha) \quad (1)$$

where  $L_{pred}$  is loss of the prediction;  $\theta$  is learnable model parameters;  $L_{tdm}$  is loss of TDM;  $D_i$  and  $D_j$  represents different data segments;  $\alpha$  denotes the importance evaluation factor of each pair of data segments; and  $\lambda$  is a trade-off parameter.

And  $L_{pred}(\theta)$  can be computed by Eq. (2).

$$L_{pred}(\theta) = \frac{1}{K} \sum_{j=1}^K \frac{1}{|D_j|} \sum_{i=1}^{|D_j|} l(y_i^j, M(x_i^j; \theta)) \quad (2)$$

where  $l(\cdot, \cdot)$  is MSE loss function;  $M$  represents model prediction; and  $(x_i^j, y_i^j)$  denotes the  $i$ -th labeled segment from period  $D_j$ .

Given a period of  $(D_i, D_j)$ , loss of temporal distribution matching can be calculated through Eq. (3):

$$L_{tdm}(D_i, D_j; \theta, \alpha_{i,j}^t) = \sum_{t=1}^V \alpha_{i,j}^t d(h_i^t, h_j^t; \theta) \quad (3)$$

where  $(h_i^t, h_j^t)$  denotes the hidden state parameters of models trained from  $D_i$  and  $D_j$ , respectively;  $d(\cdot, \cdot)$  represents maximum mean discrepancy (MMD) in this paper, and it is calculated through Eq. (4).

$$d_{mmd}(x_s, x_t) = \frac{1}{n_s^2} \sum_{i,j=1}^{n_s} k(x_{s_i}, x_{s_j}) + \frac{1}{n_t^2} \sum_{i,j=1}^{n_t} k(x_{t_i}, x_{t_j}) - \frac{2}{n_s n_t} \sum_{i,j=1}^{n_s+n_t} k(x_{s_i}, x_{t_j}) \quad (4)$$

where  $x_s$  and  $x_t$  are the source data and target data, respectively;  $k(\cdot, \cdot)$  is the radial basis function kernel;  $n_s$  and  $n_t$  are the number of the two sets of data.

The evaluation factor for each pair of data segments,  $\alpha_{i,j}^t$ , with the objective of determining the relative significance of  $V$ -dimensional hidden states in the model, is computed through Eq. (5).

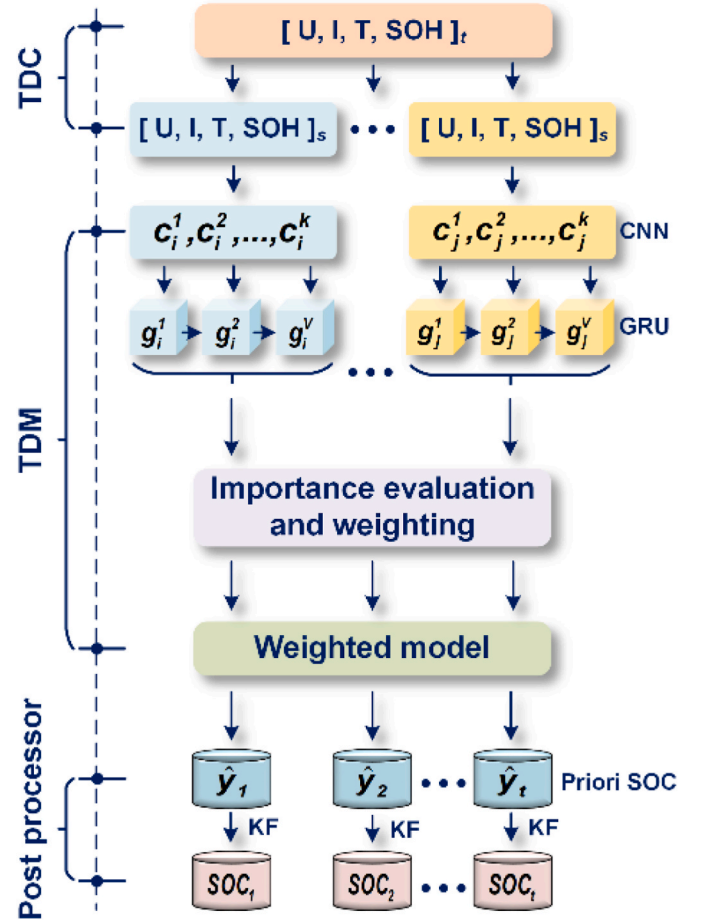


Fig. 2. Ada-CNN-GRU-KF model.

$$\alpha_{i,j}^{t,(n+1)} = \begin{cases} \alpha_{i,j}^{t,(n)} \times \left(1 + \sigma\left(d_{i,j}^{t,(n)} \times d_{i,j}^{t,(n-1)}\right)\right) & d_{i,j}^{t,(n)} \geq d_{i,j}^{t,(n-1)} \\ \alpha_{i,j}^{t,(n)} & \text{otherwise} \end{cases} \quad (5)$$

where  $d_{i,j}^{t,(n)}$  is the distribution distance at the time step  $t$  in epoch  $n$ .

## 2.2. Ada-CNN-GRU-KF model for SOC estimation

As the battery ages, its capacity gradually decreases, leading to a decline in the precision of SOC estimation. Most existing SOC estimation methods can only be operated at one single aging point, and cannot be applied throughout the entire lifespan of a battery. In view of this issue, an Ada-CNN-GRU-KF model is constructed in this section to estimate the SOC, while considering the battery's SOH, across the battery's full life cycle. By the way, for SOH changes slowly, it is a calculation waste to calculate SOH every seconds, thus aging point is leveraged to represent key time points among which there is a marked SOH difference. Fig. 2 illustrates the model structure, which primarily comprises the TDC module, the TDM module, and the KF post-processing module to provide stable prediction.

During the model training and validation process, the data is input using a sliding window approach, where each window consists of 120 time points and encompasses 4 features. The mathematical representation of the model's input and output is expressed as Eq. (6).

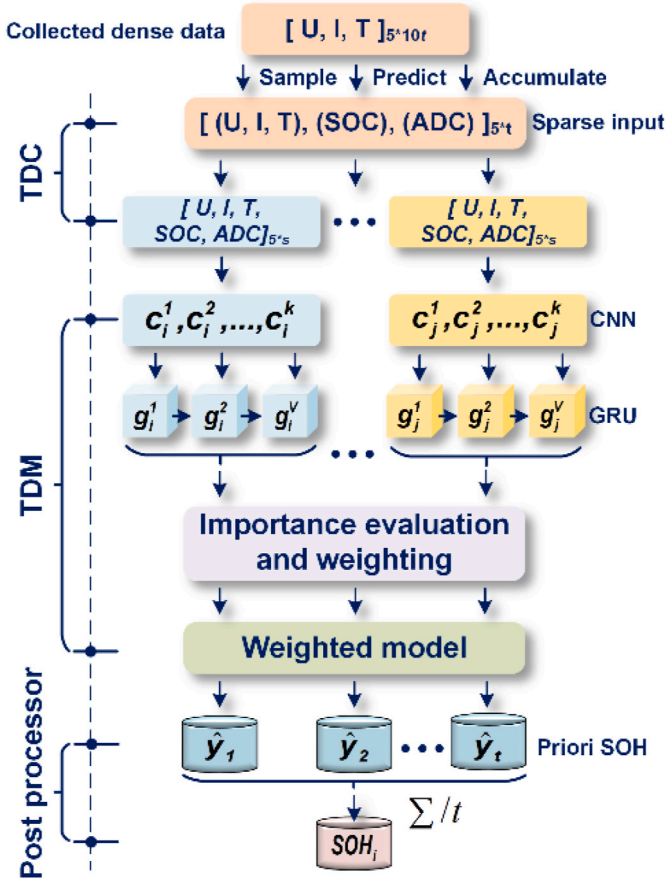


Fig. 3. Ada-CNN-GRU-Ave model.

$$\begin{cases} \text{input}_i = \begin{pmatrix} U_i & I_i & T_i & \text{SOH}_i \\ U_{i+1} & I_{i+1} & T_{i+1} & \text{SOH}_{i+1} \\ \vdots & \vdots & \vdots & \vdots \\ U_{i+119} & I_{i+119} & T_{i+119} & \text{SOH}_{i+119} \end{pmatrix}_{120 \times 4} \\ \text{output}_i = (\text{SOC}_i)_{1 \times 1} \end{cases} \quad (6)$$

The KF algorithm is composed of time update equation and state update equation, which will not be explained in detail here. The KF post-processing module takes the output of the Ada-CNN-GRU part as the observation and takes the Coulomb counting method as the update equation to obtain a smooth SOC output second by second online.

### 2.3. Ada-CNN-GRU-Ave model for SOH estimation

The structure of the Ada-CNN-GRU-Ave model is shown in Fig. 3, and it consists of the TDC module, the TDM module, and the mean post-processing module, taking voltage, current, temperature, SOC and accumulated discharged capacity (ADC) as input. Since SOH changes slowly throughout the battery's entire lifespan and the space usage of collected feature data is extremely large, there exists obvious redundancy in the feature data. Consequently, the data fed in Ada-CNN-GRU-Ave model is sampled at a frequency of 0.1 Hz, which implies that feature data collected at 1 Hz in BMS would be resampled every 10 s to facilitate the process of model training and application by reducing complexity.

The conventional deep learning models utilized for SOH estimation primarily rely on voltage, current, and temperature as input variables. However, these inputs may not be effective in accurately estimating SOH under intricate and dynamic operating conditions throughout the battery's entire lifespan. Because there is a strong correlation among SOH, SOC and ADC, so SOC and ADC are also taken as additional features alongside voltage, current, and temperature. Given the high precision of the laboratory testing equipment, to ensure a high-accuracy model to be trained, the SOC value is calculated using the Coulomb counting method in training process. ADC is obtained through Coulomb counting at a frequency of 1 Hz, while the sampling is performed at a frequency of 0.1 Hz. This approach ensures consistency in the dimensions of the input features while avoiding data redundancy. And the input of SOH estimation model can be represented by Eq. (7).

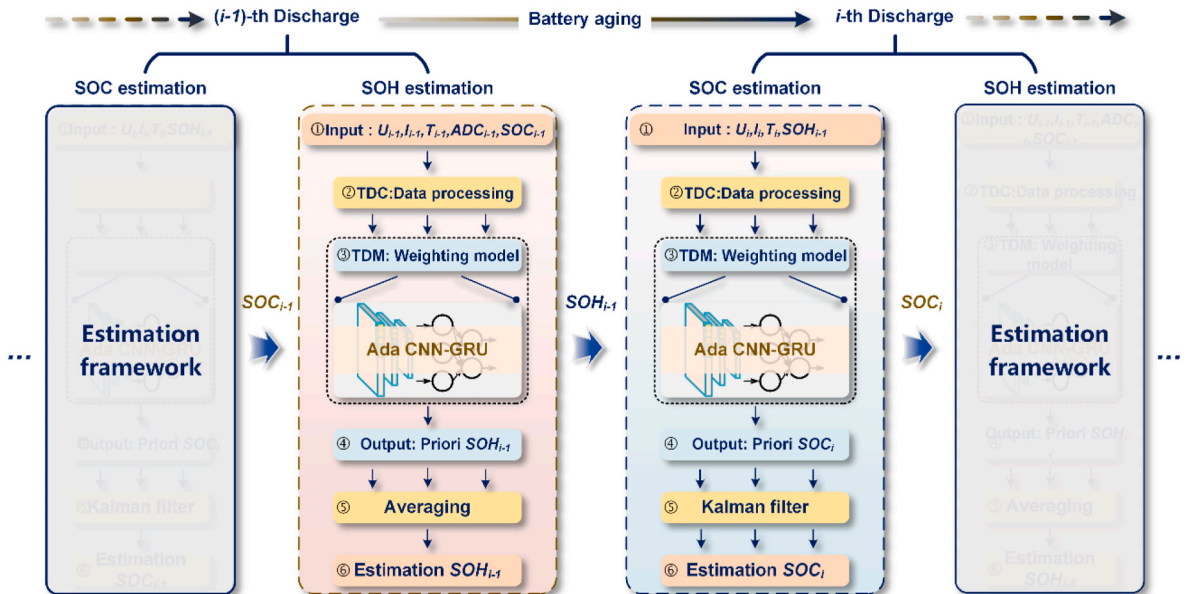


Fig. 4. JADTL model for joint estimation of SOC and SOH.



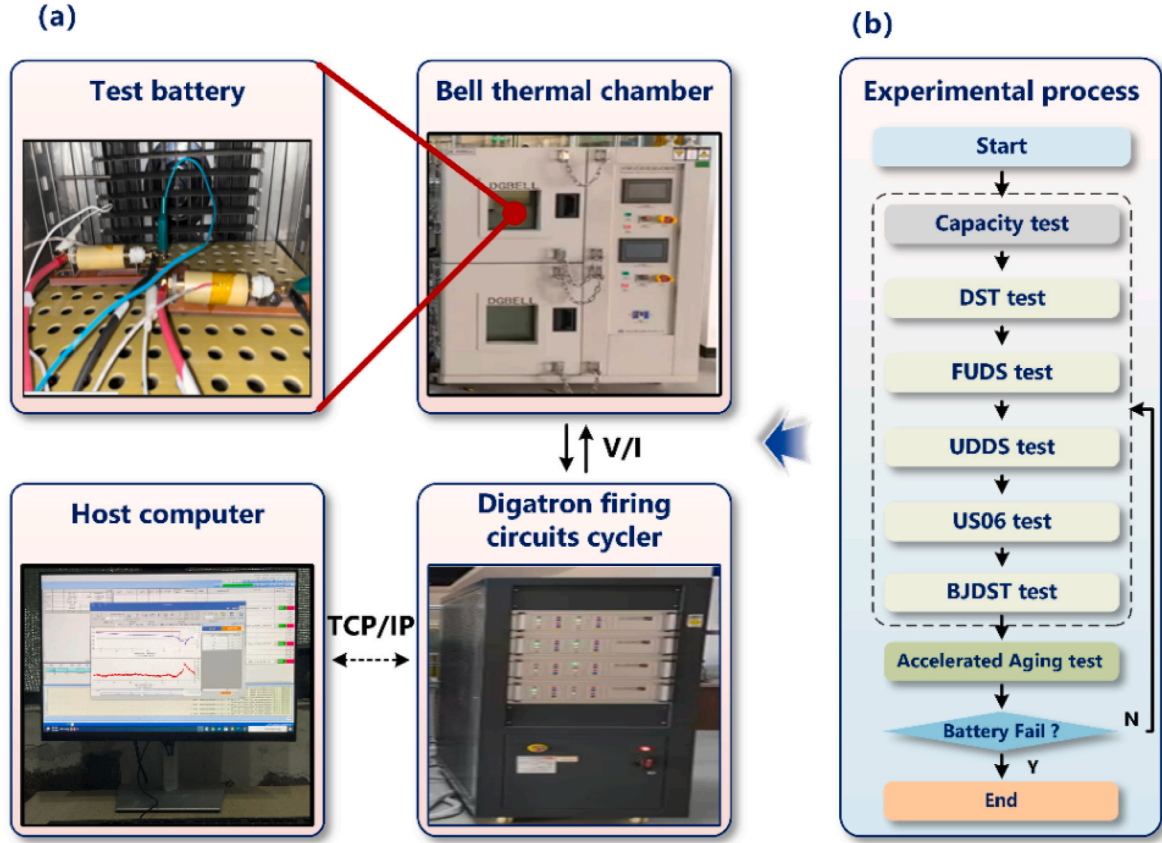


Fig. 5. (a) Battery testing platform (b) Experimental process.

$$\left\{ \begin{array}{l} input_i = \begin{pmatrix} U_i & I_i & T_i & SOC_i & ADC_i \\ U_{i+1} & I_{i+1} & T_{i+1} & SOC_{i+1} & ADC_{i+1} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ U_{i+29} & I_{i+29} & T_{i+29} & SOC_{i+29} & ADC_{i+29} \end{pmatrix}_{30 \times 5} \\ output_i = (SOH_i)_{1 \times 1} \end{array} \right. \quad (7)$$

where the time interval between  $i$  and  $i+1$  is 10 s. 10 s is selected as time interval according to hyper-parameter tuning experience, more appropriate interval need to be further studied in the future.

Furthermore, it is worth noting that the utilization of the Coulomb counting method during the generation of ADC values is less susceptible to inaccuracies and cumulative errors of the sensor in practical applications. This is because ADC can be reset to zero every time when the battery is fully charged.

The Ada-CNN-GRU component of the SOH estimation model operates in a many-to-one format, with a chosen input window size of 30. A complete discharge process consists of hundreds or thousands of data in sequence, which means that there would also be hundreds or thousands of SOH estimation values obtained at one single aging point. The length of the prior output of the model can be calculated by Eq. (8).

$$S_{pri} = S_t - S_{win} \quad (8)$$

where  $S_{pri}$  is the length of priori output, and  $S_t$  is the length of data from a complete discharge process, and  $S_{win}$  is the window size of input.

Due to the potential computational burden and the risk of generating unstable redundant output, this model adopts a strategy of averaging and updating the SOH value after completing a discharge cycle. The corresponding process can be mathematically expressed as Eq. (9).

$$SOH = \left( \sum_{i=1}^t \hat{y}_i \right) / t \quad (9)$$

#### 2.4. Joint model for SOC and SOH estimation

Despite completing the construction of the SOC estimation model and SOH estimation model for dynamic driving conditions throughout the battery's full life cycle, a comprehensive joint estimation model has not been fully established. Therefore, this section introduces the JADTL model, depicted in Fig. 4, for jointly estimating SOC and SOH. And adaptive means having ability to adapt to data with different distribution. During the full life cycle of the battery, the output of Ada-CNN-GRU-KF model and Ada-CNN-GRU-Ave model will serve as the input for the other. Specifically, for example, the SOH estimation results obtained by Ada-CNN-GRU-Ave in the  $(i-1)$ -th cycle will be utilized as input of Ada-CNN-GRU-KF model for the SOC estimation in the  $i$ -th cycle.

It is worth noting that under real-world conditions, the SOH value exhibits continuity between 2 adjacent cycles. Thus, it is feasible to take the average SOH value obtained from data of last cycle as SOH of current cycle.

### 3. Datasets description

In this study, battery experiments were conducted to collect discharge data from a Li-ion battery across different working conditions and aging points. The battery cell testing platform utilized in these experiments included a host, Digatron Firing Circuits cyler, Bell thermal chamber, and an A123-manufactured 18650 LiFePO<sub>4</sub> battery. In addition, it should be noted that, in theory, the proposed model can be applied to different types of batteries due to the strong non-linear fitting

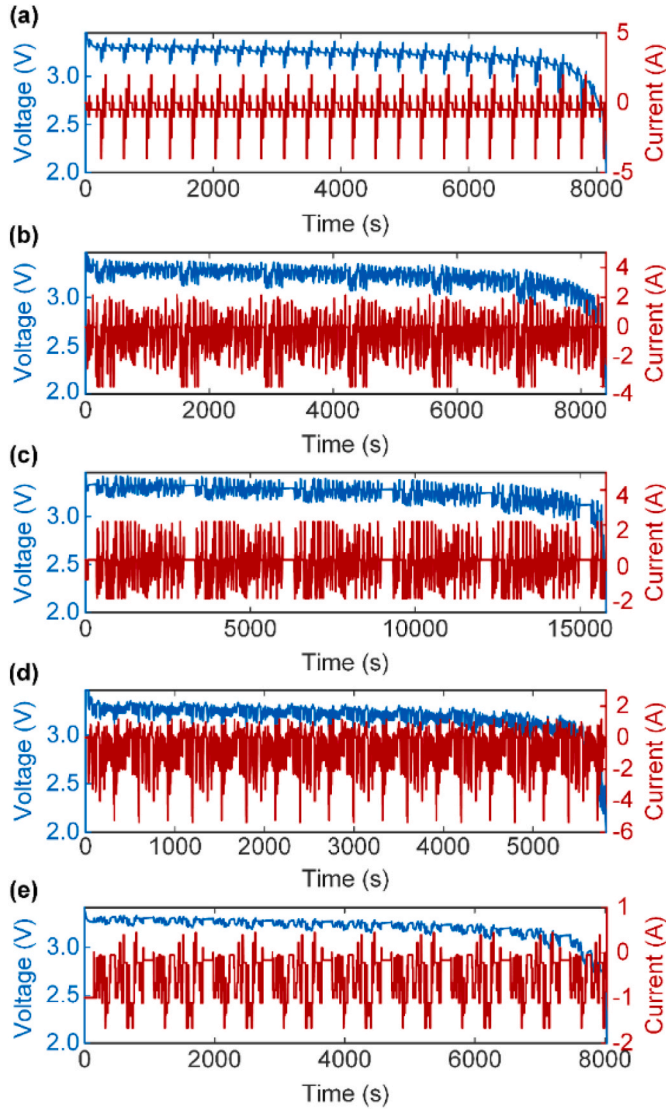


Fig. 6. Profiles under dynamic driving condition at initial aging point (a) DST (b) FUDS (c) UDSS (d) US06 (e) BJDST.

capabilities of its basic data-driven modules. In consideration of time cost, aging data are only collected from  $\text{LiFePO}_4$  battery at present in this paper, and data collection of other types of battery and corresponding validation would be conducted in the future. Fig. 5(a) illustrates the setup used for these experiments.

An experimental process plan, as depicted in Fig. 5(b), was devised. This plan includes a range of essential battery experiments, such as capacity testing, dynamic driving condition testing and battery accelerated aging cycling. Capacity testing procedure consists of the following steps: 1) Charging the battery to full using CCCV with a rate of 0.5C; 2) Resting the fully charged battery for an adequate duration; and 3) Discharging the battery at a constant current of 0.5C until it reaches the cut-off voltage.

Dynamic driving condition testing refers to tests simulating the battery working in real-world. In this study, five dynamic driving conditions were selected, including Dynamic Stress Test (DST), The Federal Urban Driving Schedule (FUDS), UDSS, US06, and Beijing Dynamic Stress Test (BJDST), which are the commonly used test procedures given in battery test procedure manuals.

The DST uses a 360 s sequence of power steps with seven discrete power levels. The DST is a typical driving cycle that is often used to evaluate various battery models and SOC estimation algorithms. The

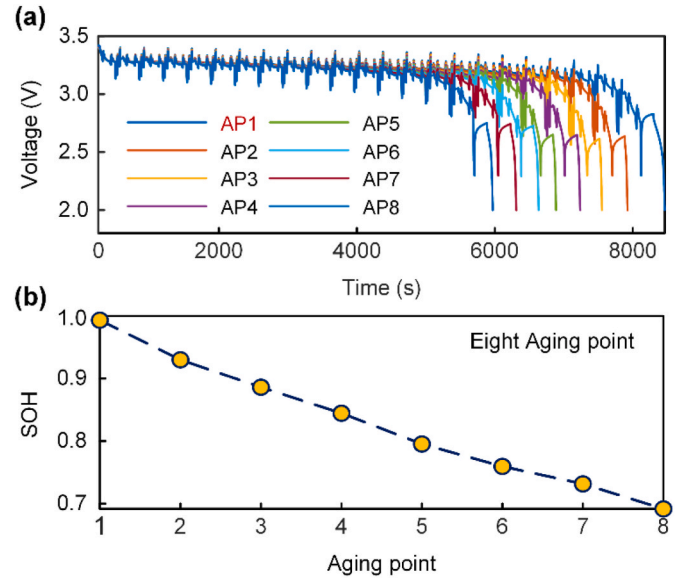


Fig. 7. (a) Voltage tendency at different aging point under DST (b) SOH tendency in full life cycle.

FUDS is a standard time-velocity profile for urban driving vehicles as well. The UDSS was used for the driving conditions of light vehicles. US06 was a driving condition of electric vehicles with high acceleration. The BJDST is obtained referencing DST and FUDS driving condition to reflect operating characteristics of electric bus in Beijing.

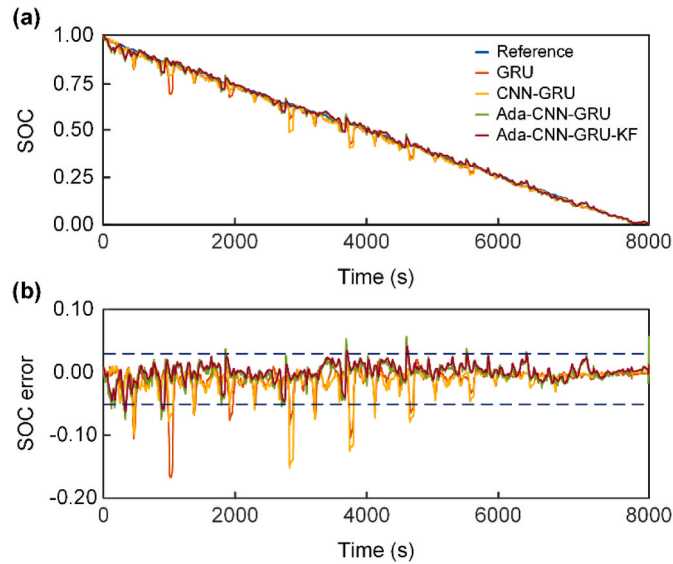
The voltage and current changes along time at the initial aging point is shown in Fig. 6. In the following sections, the term aging point denotes a specific time reference within the battery aging process, initially set at 1 for a new battery and increased by 1 after every 20 accelerated aging experiments.

Considering that the current recorded in the laboratory is generally regarded as accurate and reliable, the integral value of the current at the end of the full discharge process is designated as the aged capacity. The reference value for the SOH of the model can be obtained from Eq. (10).

$$SOH_{ref} = \frac{C_{aged}}{C_{fresh}} \times 100\% \quad (10)$$

In this study, accelerated aging tests were devised to efficiently acquire battery aging data throughout its entire lifespan through high-rate charging and discharging at low temperatures. The experimental procedure consisted of the following steps: 1) Charging the battery at a constant current of 2C until reaching cut-off voltage of 3.65 V, followed by a resting period of 0.5 h; 2) Discharging the battery at a constant current of 2C until reaching cut-off voltage of 2 V, followed by a resting period of 0.5 h; 3) Repeating steps 1 and 2 twenty times per aging point at a temperature of 10 °C until SOH of the battery fell below 70%.

As the natural charge and discharge aging process takes a long time, accelerated aging tests were designed in this study to quickly obtain battery aging data over its full life cycle through high-rate charging and discharging under low-temperature conditions. The specific experimental procedure is as follows: 1) Charge the battery at a constant current of 2C to the cut-off voltage of 3.65 V and let it rest for 0.5 h; 2) Discharge the battery at a constant current of 2C to the cut-off voltage of 2 V and let it rest for 0.5 h; 3) Repeat steps 1 and 2 20 times per aging point unit at 10 °C until the SOH of the battery drops below 70%. Taking the DST condition as an example, the voltage trajectory with aging is shown in Fig. 7(a). And the battery aging trend is shown in Fig. 7(b), in which the symbol AP means aging point.



**Fig. 8.** SOC estimation results at aging point 1: (a) estimated SOC (b) SOC estimation error.

## 4. Experimental evaluation

### 4.1. Verification of Ada-CNN-GRU-KF model

To assess the performance of the Ada-CNN-GRU-KF model in multi-aging point SOC estimation tasks, the datasets were initially partitioned into training, validation, and testing subsets. The training set comprised data from four different operating conditions: Cap, DST, UDDS, and FUDS. The validation set consisted of data from the US06 condition, while the testing set contained data from the BJDST condition. Additionally, during the TDC stage, time series segments were manually split based on their respective working conditions. It should be noted that the following verification of other models also obeys the same rule of datasets splitting, and data collected from BJDST is always selected as testing set.

The maximum absolute error (MAXE), MAE and the root mean square error (RMSE), as defined in Eq. (11), are chosen to assess the estimation performance of the proposed models in this paper.

$$\begin{cases} \text{MAXE} = \max(|y_i - \hat{y}_i|) \\ \text{MAE} = \frac{1}{m} \sum_{i=1}^m |y_i - \hat{y}_i| \\ \text{RMSE} = \sqrt{\frac{1}{m} \sum_{i=1}^m (y_i - \hat{y}_i)^2} \end{cases} \quad (11)$$

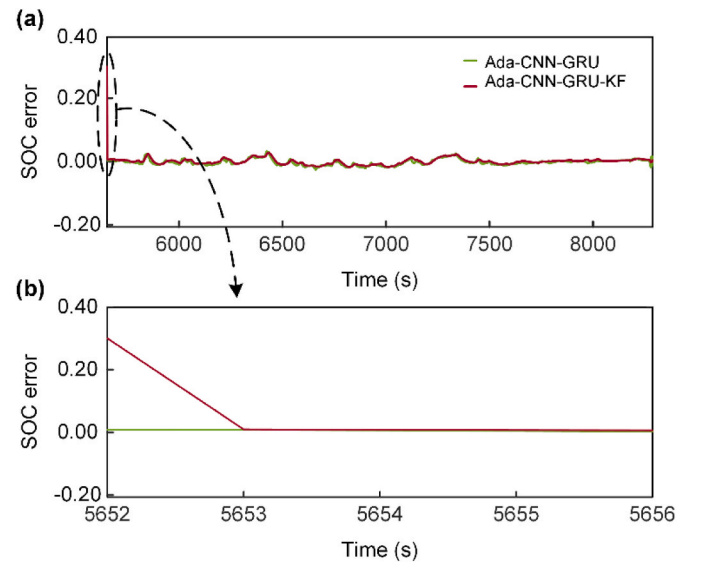
After conducting parameter tuning on the validation set, the measurement noise  $R$  for the KF was determined to be  $10^{-3}$ , and the process

noise  $Q$  was set to  $10^{-5}$ . The estimated SOC results, taking aging point 1 for example, are shown in Fig. 8, and the results throughout the full life cycle are presented in Table 1.

In the SOC estimation tasks, the Ada-CNN-GRU-KF model exhibits the capability to eliminate abnormal values, enhance result stability, and maintain a Maximum error within 10%. It achieves an average MAE value of 0.98% and an average RMSE value of 1.41% across eight aging points. In comparison to conventional deep learning models, the proposed model has been validated for its accuracy, generalization, and stability throughout the entire lifespan of the battery.

Furthermore, in order to evaluate the estimation performance under non-full charge conditions, an experiment was conducted with SOC at 0.3 as the starting point. And the green line in Fig. 9 presents the estimation results of the Ada-CNN-GRU model. Since the model parameters remain unchanged after training and the input data is confined to fixed-size windows not affected by data outside the window, the performance of the Ada-CNN-GRU model remains consistent across specific discharge intervals.

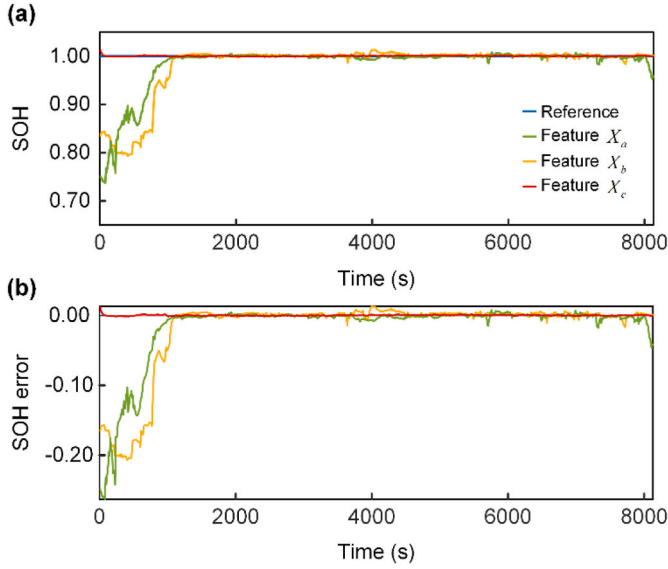
Simultaneously, to assess the robustness of the Ada-CNN-GRU-KF model when the KF initial observation value is incorrect, an SOC estimation experiment with extreme initial value bias was performed. For the MAXE of Ada-CNN-GRU model reaches 23.47%, the initial observation bias of KF is set to 30%. Specifically, the initial observation value was set to 0.6 while the reference value was 0.3. As depicted in Fig. 9, the estimation of the Ada-CNN-GRU-KF model converged to the normal range within 2 s, thereby validating the model's robustness.



**Fig. 9.** (a)SOC estimation error under conditions of unfully discharged and biased KF observation value at aging point 1 (b) detailed estimation error at the beginning.

**Table 1**  
SOC estimation results during full life cycle.

Aging point	Reference SOH	GRU			CNN-GRU			Ada-CNN-GRU			Ada-CNN-GRU-KF		
		MAE (%)	RMSE (%)	MAE (%)	RMSE (%)	MAE (%)	RMSE (%)	MAE (%)	RMSE (%)	MAXE (%)	MAE (%)	RMSE (%)	MAXE (%)
1	1.00	1.28	2.40	16.71	1.39	2.46	15.05	1.07	1.54	7.47	1.00	1.40	6.36
2	0.93	1.55	2.93	13.77	1.67	2.69	15.34	1.53	2.15	10.48	1.52	2.05	9.19
3	0.89	2.61	4.18	24.42	1.47	2.57	14.56	1.41	2.12	9.97	1.19	1.85	9.36
4	0.85	1.37	3.23	22.02	1.36	2.67	20.32	0.92	1.34	7.82	0.86	1.23	5.71
5	0.80	1.80	4.05	27.93	1.37	2.65	17.56	0.99	1.59	23.47	0.90	1.40	7.72
6	0.77	1.38	3.22	19.21	1.05	2.20	15.76	0.92	1.22	4.64	0.96	1.28	4.38
7	0.74	1.78	3.49	20.72	1.42	2.40	15.06	0.80	1.19	6.74	0.72	1.08	5.80
8	0.70	1.69	3.37	19.03	1.52	2.49	69.26	0.85	1.14	4.91	0.71	0.99	4.11



**Fig. 10.** SOH estimation results of CNN-GRU model at aging point 1 based on 3 different input features: (a) estimated SOH (b) SOH estimation error.

#### 4.2. Verification of Ada-CNN-GRU-Ave model

To achieve cost-effective estimation of SOH in dynamic driving conditions and establish a basis for the joint estimation of SOC and SOH, Ada-CNN-GRU-Ave model is constructed and assessed in this section.

Firstly, the collected data is processed by integrating and interval sampling to obtain feature data in required dimension at a frequency of 0.1 Hz. And then the data is divided into training, validation, and test sets using the same methodology as described in section 4.2.2, thus completing the initial data preparation.

In order to assess the impact of SOC and ADC as new features in the task of SOH estimation, three different types of input features are utilized for feature validation using the CNN-GRU model: a three-feature input  $X_a = [U, I, T]_{120 \times 3}$ , a four-feature input  $X_b = [U, I, T, SOC]_{120 \times 4}$ , and a five-feature input  $X_c = [U, I, T, SOC, ADC]_{120 \times 5}$ . The model structure closely resembles that of the CNN-GRU model outlined in section 4.2.1. The test results are presented in Fig. 10 and Table 2. Additionally, the Ave values in the table represent the average of the error indicators across eight aging points.

It can be observed from Table 2 that incorporating the SOC feature, in addition to voltage, current, and temperature, reduces the average MAE from 0.67% to 0.45% and the average RMSE from 1.4% to 1.03%. Furthermore, the inclusion of ADC significantly enhances the accuracy of the model estimation, resulting in an average MAE value of 0.14% and an average RMSE value of 0.37%. Moreover, Fig. 10 illustrates that the model with the novel input features exhibits a faster convergence speed, maintaining estimation errors within 0.5% for all eight aging points within 15 s.

**Table 2**

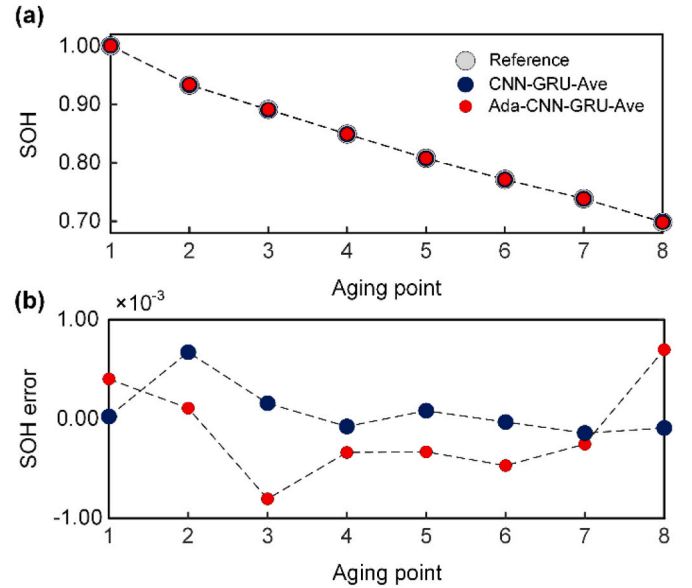
SOH estimation results of CNN-GRU model through full life cycle based on 3 different input features.

Aging point	Reference SOH	$X_a$			$X_b$			$X_c$		
		MAE (%)	RMSE (%)	MAXE (%)	MAE (%)	RMSE (%)	MAXE (%)	MAE (%)	RMSE (%)	MAXE (%)
1	1.00	2.15	5.69	20.67	1.70	5.03	26.31	0.06	0.09	1.27
2	0.93	0.37	0.67	6.72	0.25	0.68	7.18	0.19	0.63	6.68
3	0.89	0.59	0.91	9.91	0.30	0.49	3.24	0.18	0.41	4.31
4	0.85	0.51	0.73	5.02	0.27	0.42	3.72	0.15	0.38	4.13
5	0.80	0.51	0.87	7.18	0.24	0.37	2.67	0.17	0.39	4.04
6	0.77	0.41	0.86	6.19	0.15	0.20	0.74	0.15	0.33	3.39
7	0.74	0.47	0.67	3.61	0.42	0.61	6.29	0.14	0.29	3.04
8	0.70	0.42	0.80	5.07	0.29	0.50	3.75	0.15	0.40	3.76
Ave	—	0.67	1.40	8.04	0.45	1.03	6.73	0.14	0.37	3.83

It should be noted that although ADC value at the end of discharge is the same as the capacity value, this does not mean that the model is influenced by the value ADC at the very end of discharge process when estimating SOH. It can be clearly seen that the SOH value can converge to an accurate and credible range at the beginning of discharge, which means the ADC will not leak information, and the model can estimate SOH online reliably during testing. In addition, ADC will be recalculated every time when the battery is fully charged, so it is less affected by accumulated error caused by sensor accuracy.

To verify the superiority of the Ada-CNN-GRU model over the traditional CNN-GRU model and mitigate fluctuations and redundancy in the output SOH, both the Ada-CNN-GRU-Ave model and the CNN-GRU-Ave model are established. Apart from necessary dimension adjustments, the model structure remains consistent. The SOH estimation results under BJDST condition are displayed in Fig. 11, and the error indicators are summarized in Table 3.

According to Table 3 and it is evident that the Ada-CNN-GRU-Ave model outperforms the CNN-GRU-Ave model. Specifically, the MAE and RMSE values of the Ada-CNN-GRU-Ave model are 0.02% and 0.03% respectively, which are lower than the respective values of 0.04% and 0.05% obtained by the CNN-GRU-Ave model. Furthermore, the MAXE of the Ada-CNN-GRU-Ave model is also reduced compared to the CNN-GRU-Ave model. Thus, the precision and superiority of the Ada-CNN-GRU-Ave model has been verified, establishing a foundation for joint estimation of SOC and SOH.



**Fig. 11.** SOH estimation results through full life cycle: (a) estimated SOH (b) SOH estimation error.

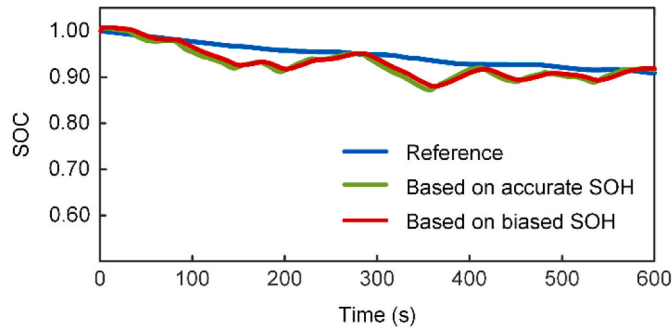


**Table 3**  
SOH estimation results through full life cycle.

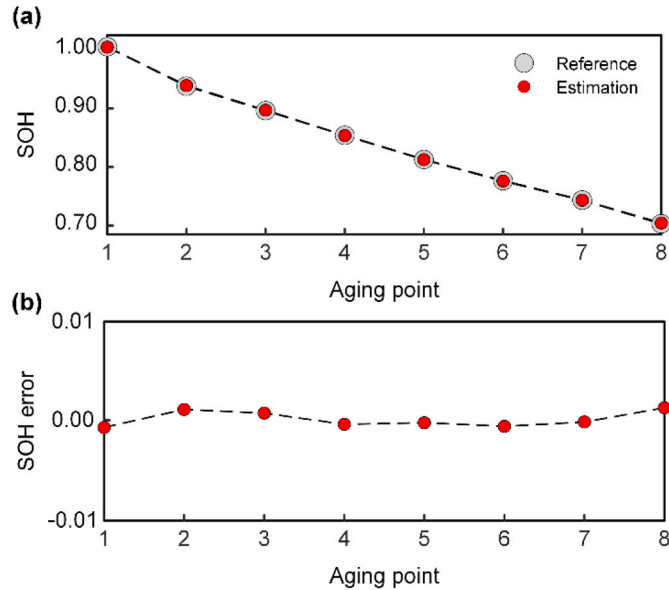
Model	MAE (%)	RMSE (%)	MAXE (%)
CNN-GRU-Ave	0.04	0.05	0.08
Ada-CNN-GRU-Ave	0.02	0.03	0.07

**Table 4**  
Settings of initial SOH value.

Initial SOH	Aging point							
	1	2	3	4	5	6	7	8
Reference	1.00	0.93	0.89	0.85	0.80	0.77	0.74	0.70
Biased	0.95	0.88	0.84	0.81	0.75	0.72	0.69	0.65



**Fig. 12.** Estimated SOC based on biased initial SOH in the first 600 s at aging point 1.

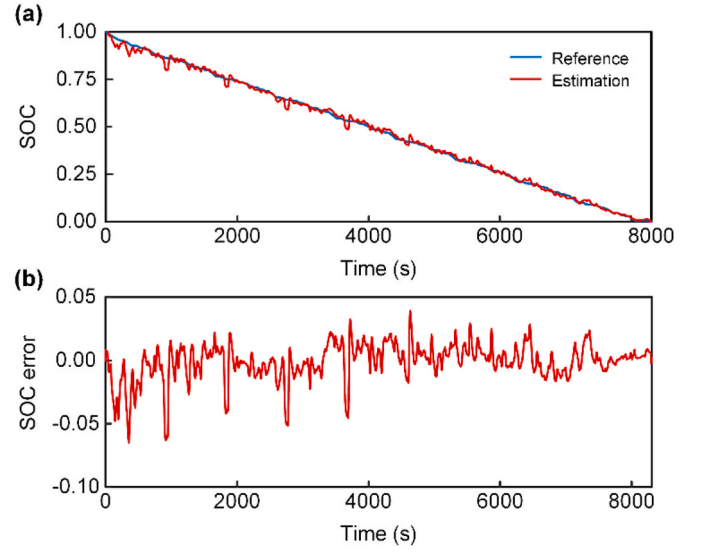


**Fig. 13.** SOH estimation results based on estimated SOC in simulated experiment: (a) estimated SOH (b) SOH estimation error.

#### 4.3. Verification of joint estimation model

Although the Ada-CNN-GRU-KF model incorporates SOH as an input feature, and the Ada-CNN-GRU-Ave model includes SOC in its input features, the input SOH in SOC estimation was ideal, so was SOC in SOH estimation.

In the battery aging experiment, there are 30 charge and discharge cycles between two adjacent aging points. As a result, the SOH values



**Fig. 14.** SOC estimation results in simulated experiment at aging point 1: (a) estimated SOC (b) SOC estimation error.

are discrete along aging points and the SOH value estimated based on previous cycle cannot be leveraged at current cycle. In order to assess the effectiveness of JADTL, inaccurate initial SOH values with a bias of 5% were set to obtain SOC estimation values with extreme errors during the first 600 s. Then the estimated SOC value was used to obtain the SOH value, which is closer to the value obtained in reality, and this was achieved using the Ada-CNN-GRU-Ave model. It should be noted that, for SOH estimation, the data collected during the first 600 s corresponds to 60 sampling points.

As for the reason the bias is set to 5%, given that the experimental battery undergoes 210 charge and discharge cycles throughout its full life cycle, and the MAE of Ada-CNN-GRU-Ave model is 0.02% under ideal conditions, so the cumulative error is  $210 \times 0.02\% = 4.2\%$ . To further increase the test difficulty and verify the convergence ability of Ada-CNN-GRU-Ave model given inaccurate SOC obtained from inaccurate SOH, the SOH initial bias is set to 5%. The specific settings are shown in Table 4.

During the initial 600 s, based on inaccurate SOH value, taking the 1st aging point for example, SOC is estimated and the results are shown in Fig. 12. It can be seen that despite the inaccurate initial SOH value, the estimated SOC can effectively depict the actual SOC trends. This suggests that, theoretically, it can be employed as one of the input features for SOH estimation.

Now that the SOC results obtained in extreme condition but closer to reality are obtained. And applying the SOC in joint estimation, the SOH results are obtained and shown in Fig. 13. It is evident that even under extreme conditions, the SOH values could converge to a relatively accurate range. The MAE and RMSE of SOH results obtained using an inaccurate SOC are both 0.07%, with MAXE of 0.13%.

Leveraging this SOH based on the JADTL to estimate SOC, taking aging point 1 for example, the SOC estimation results are shown in Fig. 14, while the comprehensive evaluation indicators throughout the battery's entire life are presented in Table 5.

In comparison to the SOC estimated under ideal conditions discussed in Section 4.2.1, the SOC estimated in simulated joint estimation experiment exhibits only marginal increases of 0.01%, 0.03%, and 0.46% in MAE, RMSE, and MAXE, respectively. This performance achievement is deemed satisfactory, verifying the feasibility of JADTL in joint estimation of SOC and SOH throughout the entire life of Li-ion batteries under dynamic conditions.

The code size of the proposed models is within 1 Mb. The trained FB-Ada-CNN-GRU-KF model occupies approximately 40 Mb of storage

**Table 5**

Comparison of SOC estimation results between ideal experiment based on accurate SOH and simulated experiment based on biased SOH.

Aging point	Reference SOH	Ideal experiment			Simulated experiment		
		MAE (%)	RMSE (%)	MAXE (%)	MAE (%)	RMSE (%)	MAXE (%)
1	1.00	1.00	1.40	6.36	1.02	1.45	6.50
2	0.93	1.52	2.05	9.19	1.46	2.03	9.78
3	0.89	1.19	1.85	9.36	1.28	1.98	9.91
4	0.85	0.86	1.23	5.71	0.85	1.25	6.34
5	0.80	0.90	1.40	7.72	0.88	1.45	8.54
6	0.77	0.96	1.28	4.38	0.85	1.13	4.28
7	0.74	0.72	1.08	5.80	0.71	1.10	6.41
8	0.70	0.71	0.99	4.11	0.83	1.12	4.62
Ave	–	0.98	1.41	6.58	0.99	1.44	7.04

space. The Ada-CNN-GRU-KF and Ada-CNN-GRU-Ave models both occupy around 13 Mb of storage space. The joint estimation model occupies approximately 26 Mb of space. The code and trained models could be loaded on the cloud server, and each estimation results would be calculated within 0.06 ms, which shows that the proposed model could be well applied on the task of joint estimation of SOC and SOH.

## 5. Conclusion

Firstly, an Ada-CNN-GRU-KF model that combines the advantages of transfer learning and deep learning is proposed with KF serving as a post-processor to provide stable and smooth SOC estimation results. Secondly, an Ada-CNN-GRU-Ave model is proposed, with dual-time scale input for estimating the SOH under dynamic driving conditions, and the model uses voltage, current, temperature, SOC, and ADC as input features. In the end, based on Ada-CNN-GRU-KF and Ada-CNN-GRU-Ave models, a JADTL model for joint estimation of SOC and SOH through battery's entire life is built. To verify the performance of joint estimation, a simulated joint estimation experiment under dynamic condition is conducted with initial SOH bias. And in the experiment, the MAE and RMSE of estimated SOH were both 0.07%, and the MAE and RMSE of estimated SOC were 0.99% and 1.44%, respectively. The generalization and accuracy of the joint estimation model were well verified.

According to authors knowledge, the proposed model mainly has 2 limitations. The first is that there are lot of hyper-parameters both in the neural network and the KF, it is a time consuming work to adjust these parameters and need a lot of experience. The second is that the proposed model will first train a set of hidden state parameters and weight them to a weighted model, so the model has more parameters than the model such as CNN-GRU, which would consume more computation resource and occupy more storage space. To address these problems, future work would (1) seek for intelligent algorithms for automatic parameter optimization, and (2) reduce the model size to enhance its suitability for vehicular applications.

## CRediT authorship contribution statement

**Yongsong Yang:** Writing – original draft, Methodology, Data curation. **Yuchen Xu:** Formal analysis, Software, Visualization. **Yuwei Nie:** Formal analysis, Data curation. **Jianming Li:** Visualization, Methodology. **Shizhuo Liu:** Validation. **Lijun Zhao:** Supervision. **Quanqing Yu:** Writing – review & editing, Supervision. **Chengming Zhang:** Supervision, Writing – review & editing.

## Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Data availability

Data will be made available on request.

## Acknowledgments

This work was supported by National Natural Science Foundation of China (Grant No. 52177210).

## References

- [1] Yu Q, Nie Y, Peng S, Miao Y, et al. Evaluation of the safety standards system of power batteries for electric vehicles in China. *Appl Energy* 2023;349:121674.
- [2] Huang R, Wei G, Zhou X, Zhu J, Pan X, Wang X, Jiang B, Wei X, Dai H. Targeting the low-temperature performance degradation of lithium-ion batteries: a non-destructive bidirectional pulse current heating framework. *Energy Storage Mater* 2024;65:103173.
- [3] Xiong R, Sun W, Yu Q, Sun F. Research progress, challenges and prospects of fault diagnosis on battery system of electric vehicles. *Appl Energy* 2020;279:115855.
- [4] Liu S, Nie Y, Tang A, Li J, Yu Q, Wang C. Online health prognosis for lithium-ion batteries under dynamic discharge conditions over wide temperature range. *eTransportation* 2023;18:100296.
- [5] Jiang B, Zhu Y, Zhu J, Wei X, Dai H. An adaptive capacity estimation approach for lithium-ion battery using 10-min relaxation voltage within high state of charge range. *Energy* 2023;263:125802.
- [6] Zhou Z, Liu Y, You M, Xiong R, Zhou X. Two-stage aging trajectory prediction of LFP lithium-ion battery based on transfer learning with the cycle life prediction. *Green Energy and Intelligent Transportation* 2022;1:100008.
- [7] Chen Ch, Xiong R, Yang R, Li H. A novel data-driven method for mining battery open-circuit voltage characterization. *Green Energy and Intelligent Transportation* 2022;1:100001.
- [8] Xing Y, He W, Pecht M, Tsui KL. State of charge estimation of lithium-ion batteries using the open-circuit voltage at various ambient temperatures. *Appl Energy* 2014; 113:106–15.
- [9] Pang H, Geng Y, Liu X, Wu L. A composite state of charge estimation for electric vehicle lithium-ion batteries using back-propagation neural network and extended Kalman particle filter. *J Electrochem Soc* 2022;169:110516.
- [10] Liu Y, He Y, Bian H, Guo W, Zhang X. A review of lithium-ion battery state of charge estimation based on deep learning: directions for improvement and future trends. *J Energy Storage* 2022;52:104664.
- [11] Xiong R, Li L, Tian J. Towards a smarter battery management system: a critical review on battery state of health monitoring methods. *J Power Sources* 2018;405: 18–29.
- [12] Yu Q, Nie Y, Liu S, Li J, Tang A. State of health estimation method for lithium-ion batteries based on multiple dynamic operating conditions. *J Power Sources* 2023; 582:233541.
- [13] Jiang B, Zhu J, Wang X, Wei X, Shang W, Dai H. A comparative study of different features extracted from electrochemical impedance spectroscopy in state of health estimation for lithium-ion batteries. *Appl Energy* 2022;322:119502.
- [14] Che Y, Deng Z, Li P, Tang X, Khosravinia K, Lin X, Hu X. State of health prognostics for series battery packs: a universal deep learning method. *Energy* 2022;238: 121857.
- [15] Chen Z, Zhao H, Zhang Y, Shen S, Shen J, Liu Y. State of health estimation for lithium-ion batteries based on temperature prediction and gated recurrent unit neural network. *J Power Sources* 2022;521:230892.
- [16] Hu X, Che Y, Lin X, Onori S. Battery health prediction using fusion-based feature selection and machine learning. *IEEE Trans. Transp. Electr.* 2021;7:382–98.
- [17] Huang Z, Yang F, Xu F, Song X, Tsui KL. Convolutional gated recurrent unit-recurrent neural network for state-of-charge estimation of lithium-ion batteries. *IEEE Access* 2019;7:93139–49.
- [18] Yang Y, Zhao L, Yu Q, Liu S, Zhou G, Shen W. State of charge estimation for lithium-ion batteries based on cross-domain transfer learning with feedback mechanism. *J Energy Storage* 2023;70:108037.

- [19] Liu Q, Kang Y, Qu S, Duan B, Wen F, Zhang C. An online SOH estimation method based on the fusion of improved ICA and LSTM. In: 2020 IEEE/IAS Ind Commer. Power Syst. Asia, I CPS Asia. 2020; 2020. p. 1163–7.
- [20] Wang Z, Feng G, Zhen D, Gu F, Ball A. A review on online state of charge and state of health estimation for lithium-ion batteries in electric vehicles. *Energy Rep* 2021; 7:5141–61.
- [21] Shrivastava P, Kok T, Yamani M, Bin I, Mekhilef S. Comprehensive co-estimation of lithium-ion battery state of charge, state of energy, state of power, maximum available capacity, and maximum available energy. *J Energy Storage* 2022;56: 106049.
- [22] Shen J, Ma W, Xiong J, Shu X, Zhang Y, Chen Z, Liu Y. Alternative combined co-estimation of state of charge and capacity for lithium-ion batteries in wide temperature scope. *Energy* 2022;244:123236.
- [23] Qiu X, Wu W, Wang S. Remaining useful life prediction of lithium-ion battery based on improved cuckoo search particle filter and a novel state of charge estimation method. *J Power Sources* 2020;450:227700.
- [24] Li Y, Li K, Liu X, Li X, Zhang L, Rente B, Sun T, Grattan KTV. A hybrid machine learning framework for joint SOC and SOH estimation of lithium-ion batteries assisted with fiber sensor measurements. *Appl Energy* 2022;325:119787.
- [25] Hossain Lipu MS, Hannan MA, Hussain A, Ayob A, Saad MHM, Karim TF, How DNT. Data-driven state of charge estimation of lithium-ion batteries: algorithms, implementation factors, limitations and future trends. *J Clean Prod* 2020;277:124110.
- [26] Song X, Yang F, Wang D, Tsui K-L. Combined CNN-LSTM network for state-of-charge estimation of lithium-ion batteries. *IEEE Access* 2019;7:88894–902.
- [27] Huang Z, Yang F, Xu F, Song X, Tsui K-L. Convolutional gated recurrent unit-recurrent neural network for state-of-charge estimation of lithium-ion batteries. *IEEE Access* 2019;7:93139–49.
- [28] Du Y, Wang J, Feng W, Pan S, Qin T, Xu R, Wang C. AdaRNN: adaptive learning and forecasting of time series. *Int Conf Inf Knowl Manag.* 2021:402–11.