Applied Energy 381 (2025) 125086

Contents lists available at ScienceDirect

Applied Energy

journal homepage: www.elsevier.com/locate/apenergy

A general framework for lithium-ion battery state of health estimation: From laboratory tests to machine learning with transferability across domains

Zhi Cao, Wei Gao, Yuhong Fu, Naser Vosoughi Kurdkandi, Chris Mi *

Department of Electrical and Computer Engineering, San Diego State University, San Diego, 92182, CA, USA

ARTICLE INFO

Keywords:
Lithium-on battery
State of Health (SOH)
Machine learning
Convolutional neural network (CNN)
Transfer learning

ABSTRACT

Accurate State of Health (SOH) estimation is crucial for safe, efficient, and optimal operation of lithium-ion batteries (LIBs), yet it remains challenging in real-world applications. In this regard, this paper presents a novel approach to estimating the SOH of lithium-ion batteries using a convolutional neural network (CNN) model enhanced with 3D histogram feature extraction and transfer learning. Unlike traditional models, our method is uniquely capable of handling varying lengths of input time series data with the varying sliding window, making it highly adaptable to real-world scenarios where data may be irregular or incomplete. The integration of transfer learning further enhances the model's adaptability, allowing it to efficiently generalize across different battery types and operational conditions with minimal retraining. Experimental results demonstrate the model's accuracy and robustness, with significant improvements over existing methods in terms of estimation accuracy, computational efficiency, and adaptability to new data. This research offers a practical and scalable solution for battery health monitoring, supporting the advancement of reliable and efficient battery management systems.

1. Introduction

Lithium-ion batteries (LIBs) are widely applied in diverse energy storage systems due to their high energy density, extended lifespan, low self-discharge rate, and lack of memory effect [1,2]. Nonetheless, capacity and power degradation occur throughout their lifetime, limiting the operational lifespan of LIBs in practical applications [3]. Consequently, accurately monitoring batteries' state-of-health (SOH) has become crucial for ensuring efficient, safe and optimal system operation, as well as proper planning of battery maintenance and retirement.

The SOH compares the current state of the battery to its initial state at the beginning of life (BOL) [4]. The capacity-based SOH of LIBs is defined as

$$\mathrm{SOH} = \frac{C_{now}}{C_0} \tag{1}$$

where C_{now} and C_0 are the current capacity and the initial capacity of the battery, respectively.

SOH can be directly measured through ampere-hour (Ah) integration by fully charging and discharging batteries. Although simple, this method is time-consuming and often unavailable in real-world applications [5]. As a result, researchers have turned to predictive models for the SOH of LIBs. Typically, these methods include model-based and data-driven approaches [6]. Model-based approaches estimate SOH by initially building models, such as empirical, electro-chemical (EM) and equivalent circuit models (ECM), to capture battery behaviors, and followed by the estimation of model parameters. However, the trade-off between model complexity and accuracy in actual applications remains challenging [7].

An emerging alternative to model-based methods is the data-driven method. It employs machine learning (ML) on vast battery aging datasets to map the nonlinear relationship between the specified inputs and SOH, which is model-free and requires no prior knowledge. With the availability of powerful computing resources and development of ML theories, recent studies have employed various ML models for battery SOH prediction. Those models range from relatively simple ML methods, such as linear regressing model (LR) [8], support vector machine (SVM) [9,10], Gaussian process regression (GPR) [11,12], and ensemble learning [13], to complex advanced deep learning (DL) techniques, like long short-term memory (LSTM) network [14,15], convolutional neural network (CNN) [16,17], generative adversarial network (GAN) [18], and their combinations [19,20], each contributing to more intelligent SOH predictions and yielding satisfactory results. Lee et al. conducted an extensive evaluation and comparison across seven ANN architectures [21]. They examined the impact of different cycling windows on predictive accuracy and found that a combination of window width 40 cycles and shift size 40 cycles had the best accuracy. Other advancements in battery health prediction models

* Corresponding author.
E-mail address: cmi@sdsu.edu (C. Mi).

https://doi.org/10.1016/j.apenergy.2024.125086
Received 9 July 2024; Received in revised form 26 November 2024; Accepted 2 December 2024

have focused on improving robustness and adaptability in real-world conditions. For instance, Wang et al. developed an anti-noise adaptive LSTM neural network to enhance the robustness of RUL predictions under noisy conditions [22]. Similarly, Wang et al. proposed a hybrid model combining singular filtering, GPR, and LSTM for estimating battery capacity throughout its lifecycle, demonstrating effectiveness in scenarios with fast aging and multi-current variations [23]. These studies highlight the importance of addressing noise and variability in battery data. However, all these models are typically designed for fixed-length input sequences with equal cycle intervals, limiting their flexibility in handling varying lengths of operational data.

Life cycle data of LIBs are critical to the efficacy of ML methods. In existing research, public datasets have been widely used for ML model training and validation. A review in [24] summarized public datasets associated with LIBs including those from NASA, CALCE, MIT, Oxford, and other databases. However, the content, integrity, and quality of these datasets vary significantly [25]. Some datasets only cover from BOL to 80% of nominal capacity [26]. Models trained on such datasets may exhibit decreased prediction performance when applied to scenarios where batteries are repurposed for less demanding applications after they are retired from their initial application. Some datasets lack operational data, including current, voltage, and temperature. Prediction models relying solely on capacity degradation information without input of any current, voltage, and temperature data could be less efficient in predicting SOH [27]. Additionally, most data come from batteries tested under constant cycling working conditions throughout their whole lifetime, but load profiles in real-world applications are non-constant time and space variants [28]. Models trained on these datasets for predicting SOH under varied working conditions have yet to be verified. Lastly, existing models are trained with data evenly sampled across the time horizon, thereby limiting their prediction flexibility regarding time horizon. In addition to public datasets, accelerated aging tests in laboratory environments can generate sufficient cycling data in short period, under various working stresses [29].

In response to these challenges, we propose a comprehensive framework for estimating the SOH of LIBs. First, we design extensive accelerated aging tests for LIBs under varying working conditions to gather comprehensive operational data and create a high-quality dataset for machine learning model. Subsequently, we use the varying sliding window technique to process the operational data, creating various data segments. The time series data in each window are transformed into 3D histogram features with consistent dimensions. These 3D histograms, along with time duration and initial SOH value in each window, are input into machine learning (ML) model. A deep CNN model is used to learn knowledge and map input features to predicted SOH values. This approach allows the model to handle any length of time-series data and estimate SOH from at any point in time. The use of sliding windows for histogram generation ensures adaptability to variable time durations, making the model highly flexible and suitable for real-world applications where operational conditions vary significantly. Furthermore, our framework advocates the integration of transfer learning as a standard practice in model development to generalize the model across different battery types and operational conditions with minimal retraining. This is particularly important in real-world applications where operational data can vary significantly.

This paper is organized as follows: Section 2 is devoted to battery aging test design, feature extraction, and the description of the proposed ML model. Section 3 presents the experiments, results, and discussion. Section 4 concludes the article.

2. Methodology

2.1. Overall framework

The proposed framework for SOH estimation is illustrated in Fig. 1. It consists of three parts: (1) data acquisition from accelerated battery aging tests under varying working conditions to cover a wide range of degradation scenarios; (2) Development and verification of a baseline ML model using laboratory datasets; (3) Application of transfer learning to adapt the model for unseen battery applications, utilizing pre-trained knowledge and targeted datasets.

Table 1
Nissan Leaf Gen3 battery cell specifications.

Parameter	Quantity
Nominal voltage	3.65 V
Maximum voltage	4.20 V
Minimum voltage	2.80 V
Nominal capacity	56.3 Ah
Length × Width × Thickness	261 mm × 216 mm × 7.91 mm
Mass	914 g
Energy density	224 Wh/kg

Table 2
Battery cell aging test conditions during initial 1080 cycles (T = 25 °C).

<table>
<tr><th>Group: Cells</th><th>1–240 cycles</th><th>241–480 cycles</th><th>481–1080 cycles</th></tr>
<tr><td>1: 1–6</td><td rowspan="6">DoD = 100%
1C/1C</td><td colspan="2">DoD = 80%
1C/1C</td></tr>
<tr><td>2: 7-9</td><td colspan="2">DoD = 80%
0.5C/0.5C</td></tr>
<tr><td>3: 10–12</td><td>DoD = 80%
0.5C/0.5C</td><td>DoD = 80%
0.5C/1C</td></tr>
<tr><td>4: 13–15</td><td>DoD = 100%
0.5C/0.5C</td><td>DoD = 100%
0.5C/1C</td></tr>
<tr><td>5: 16–18</td><td colspan="2">DoD = 100%
0.5C/0.5C</td></tr>
<tr><td>6: 19–24</td><td colspan="2">DoD = 100%
1C/1C</td></tr>
</table>

DoD is determined by the high and low cut-off voltages: a voltage range from 2.8 V to 4.2 V typically represents 100% DoD, while a range of 3.0 V to 4.05 V equates to 80% DoD.

2.2. Accelerated battery aging test

Extensive accelerated battery aging tests are designed in the laboratory environment to acquire adequate amount of data for training the ML model. These tests encompass multiple operation conditions in terms of charge and discharge currents, cut-off voltages and ambient temperatures.

Battery cell - The study employs twenty-four Nissan Leaf Gen3 battery cells, which were sourced from battery packs retired ahead of their anticipated service life due to warranty or shipping issues. However, these packs maintained over 90% SOH at the end of their service life. The main specifications of these cells are detailed in Table 1.

Equipment - The aging tests are performed in a Chroma 17 010 battery cycler, capable of delivering 100 A at 5 V. CSZ and THERMOTRON thermal chambers are used to maintain constant ambient temperatures at 35 °C and 10 °C, respectively.

Test procedure - To obtain comprehensive aging trajectories under varied conditions, twenty four cells are divided into six groups each with its distinct testing conditions in terms of C-rates, depth of discharge (DoD), and temperatures. The test protocol during the initial 1080 cycles are given in Table 2. After 1080 cycles, cells 03, 04, 07, and 11 are cycling at 35 °C high temperatures while cells 05, 06, 09, and 12 are tested at 10 °C low temperatures. Besides, cells upon reaching specific “aging knee points” are subjected to a protection cycling test mode. The conditions maintained during this protective mode include a voltage range from 3.0 V to 4.05 V and a constant current of 0.3C.

Characteristics test - Characterization cycles, such as capacity checks, are performed every N cycles, where N represents the interval between these characterization events. The value of N varies throughout the experiment. The characteristics test cycles conducted

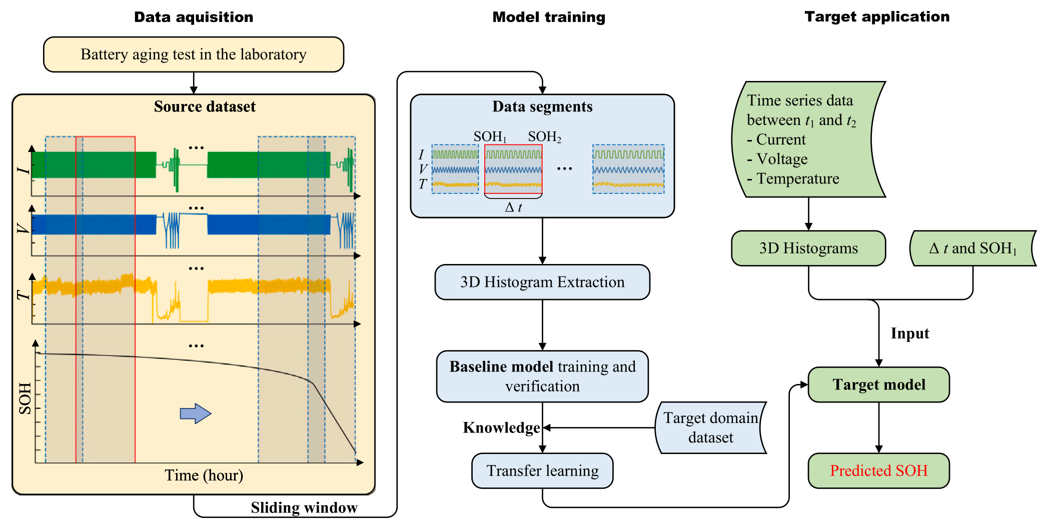

Fig. 1. Proposed framework for SOH estimation.

at 25 °C. The characteristic test cycles are designed to evaluate the degradation by capacity test. In the capacity test, three cycles are conducted, encompassing a constant current (CC) charge to the cell's maximum voltage and a CC discharge down to its minimum voltage.

2.3. Feature extraction

Feature extraction is a critical step in developing accurate machine learning models for the SOH estimation. By transforming raw operational data into meaningful features, we can capture the underlying patterns and stressors affecting battery aging, which are essential for precise SOH estimation.

Traditional feature extraction and ML models typically rely on evenly distributed, fixed-length input data. These methods may struggle to handle irregular or incomplete datasets commonly encountered in real-world applications. In contrast, our 3D histogram feature extraction method, combined with a varying sliding window, transforms irregular time-series data into a consistent and compact representation. This approach effectively captures the frequency and distribution of operational variables, providing both flexibility and robustness in handling varying operational conditions.

Operational data description - The aging rate of LIBs is influenced by operational conditions like current, temperature, DoD, and SoC as indicated in [7]. Therefore, in this study, we collect operational data including current, voltage, and temperature from lithium-ion batteries undergoing various aging tests. These data are easy to obtain and have direct impact on the battery's performance and degradation mechanisms. For instance, current and voltage profiles reflect the load conditions, while temperature and voltage provide insights into the thermal and charge–discharge cycles the battery undergoes.

Sliding window technique - To generate sequences of features over time, the sliding window technique is adopted. As shown in Fig. 2, a window slides over the time-series data with a shift distance. The window size and shift size can be varying, creating segments with various length and overlapping. The histogram features are extracted for each segment, and the duration Δt of each time window is another important input feature of the ML model.

By using this varying sliding window technique, the model can learn from both short-term and long-term dependencies in the data. Moreover, the model can process time-series data with different length as the input. As a result, it allows for SOH estimation from any point in time for various time intervals, thus providing flexibility in monitoring battery health over different periods. This capability ensures that the model can be applied dynamically in real-world scenarios where continuous monitoring is essential.

3D Histogram Feature Extraction - To effectively capture the stressors affecting battery aging in each window, 3D histograms are utilized to represent the operational data [30–32]. By converting the time series data of current, voltage, and temperature into 3D histograms, the distribution and frequency of these operational parameters over specific time windows are captured. This representation provide valuable insights into the conditions under which the battery operates. In histograms, higher frequency distributions in certain bins correspond to high temperature, high current and high depth of cycling conditions, which are known to accelerate battery degradation. This representation allows us to visualize and understand the impact of various operational stressors on battery health, thereby providing a comprehensive overview of the factors contributing to its degradation.

As shown in Fig. 3, the process involves dividing the time-series data of current, voltage, and temperature into discrete bins and counting the dwelling time within each bin. Mathematically, the 3D histogram H can be defined as:

$$H(i,j,k) = \sum_{t=t_1}^{t_2} \delta(C_t = i, V_t = j, T_t = k) \tag{2}$$

where C_t, V_t and T_t are the current, voltage, and temperature at time t, respectively, and δ is the indicator that is 1 if the condition is met and 0 otherwise. For each specified window, two 3D histogram tables are generated for charging and discharging phases, respectively. An example is given at the lower right corner in Fig. 3 to show the actual obtained 3D histogram tables.

This representation allows us to create a compact yet comprehensive feature set that captures the operational conditions over time, and it can transform different length of time series data into a universal format for the input of ML model.

2.4. Machine learning model

The architecture of the proposed ML model is depicted in Fig. 4. Inputs to the model include extracted 3D histogram tables, the initial value SOH_1 and duration Δt of each time window. The ML model learns the features that characterize the degradation under various conditions. Subsequently, it consolidates these insights in its final layer to forecast the future value SOH_2. In this study, SOH_1 refers to the measured SOH at the beginning of each sliding window. This value is used as the initial input for the model's SOH prediction. While our model is capable of using estimated SOH from previous sliding windows (SOH_2), for the results presented in this paper, we used measured SOH to ensure that the predictions were based on accurate initial conditions.

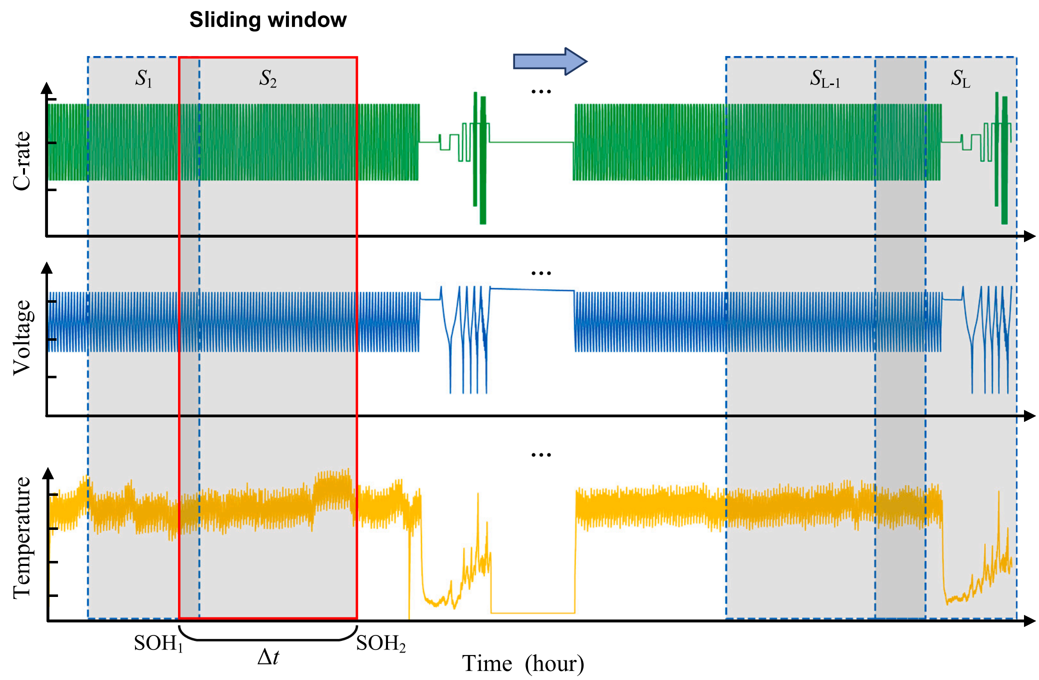

Fig. 2. Data-processing using the varying sliding window technique.

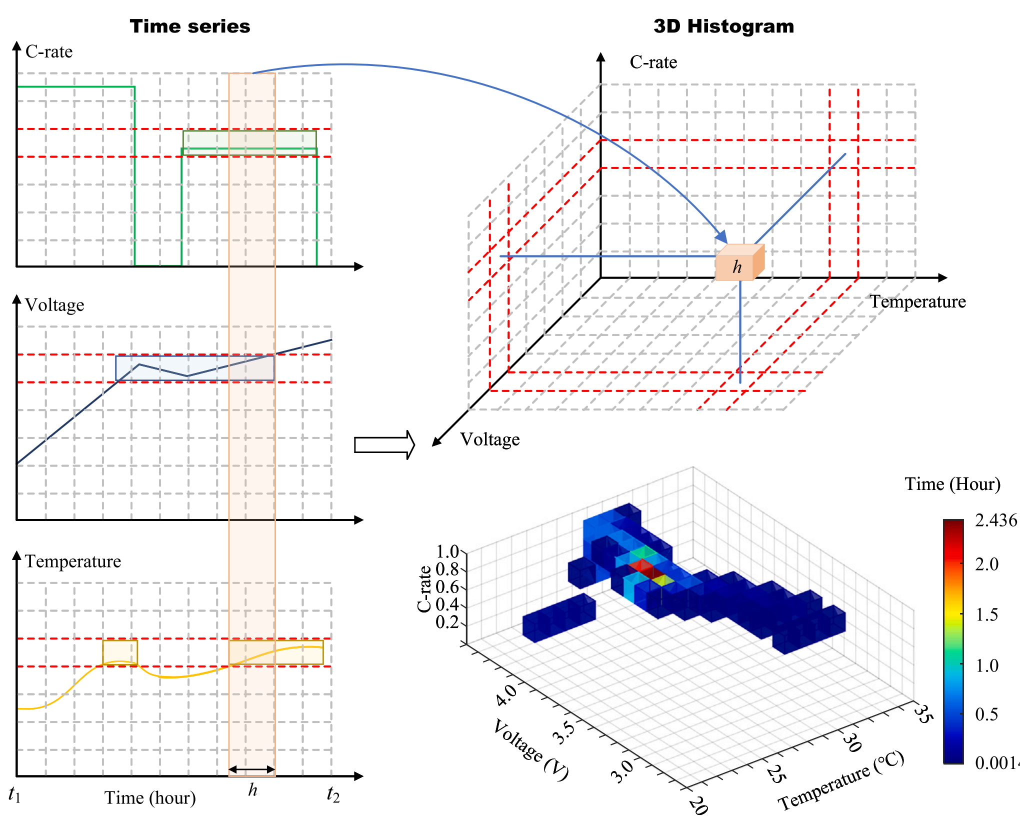

Fig. 3. Feature extraction of 3D histogram tables from time series of current, voltage and temperature.

CNNs are used in our framework thanks to their particular strength in extracting meaningful features from complex data [33]. The 3D histograms of current, voltage, and temperature provide a rich source of spatial information. CNNs are particularly adept at capturing local patterns and correlations within these histograms, which are critical for accurate SOH estimation.

Fig. 4 shows the process of the 3D CNNs. CNNs are composed of multiple convolutional and max-pooling layers. Convolutional layers use filters to highlight unique features in the tables. The Rectified Linear Unit (ReLU) activation function follows each convolutional layer, enhancing non-linearity, and a dropout layer is incorporated to mitigate overfitting by randomly nullifying input units. Max-pooling layers

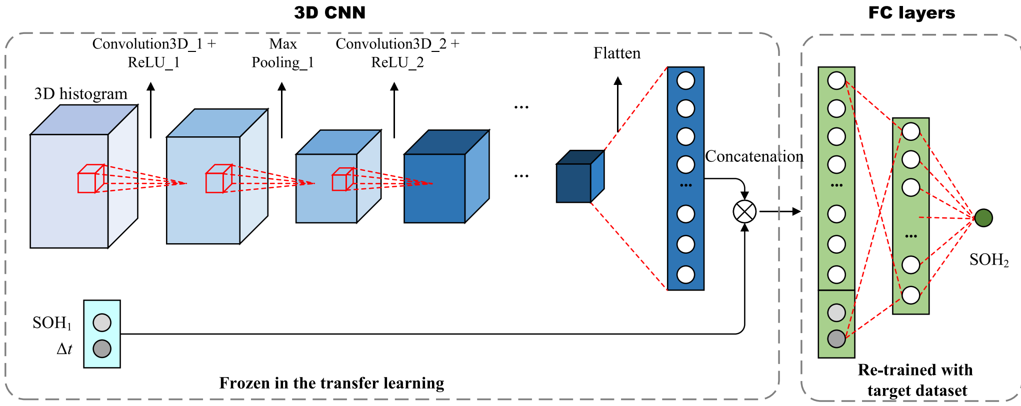

Fig. 4. Architecture of the deep CNN model.

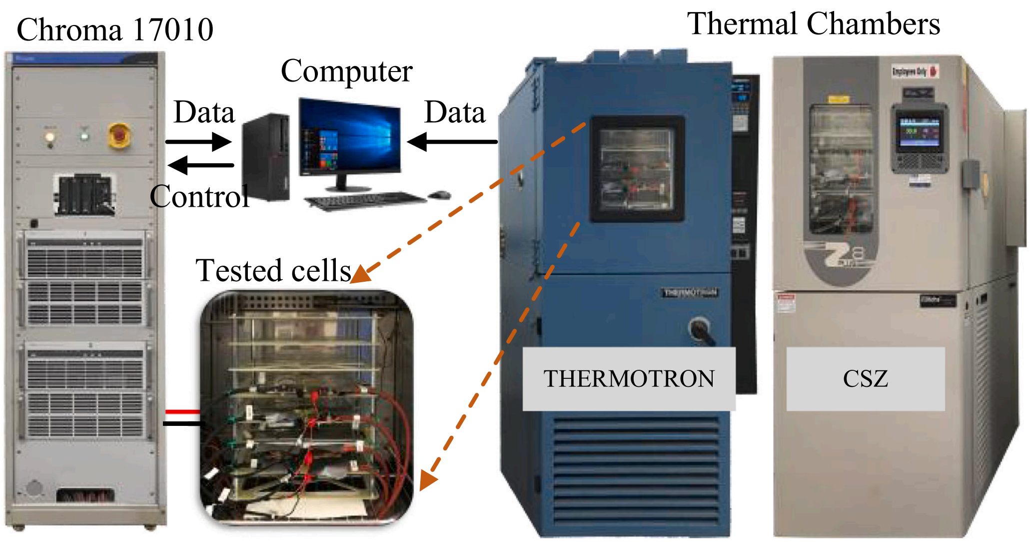

Fig. 5. Battery cell aging test bench.

reduce the spatial size of the feature tables and filter out irrelevant and redundant information.

After processing by CNN layers, the feature maps from the 3D histogram tables are flattened and concatenated with the initial SOH value SOH_1 and the time window duration Δt, forming a single vector containing all the information needed for predictions. Fully connected (FC) layers then use this vector to accurately map the features to desired output which is future SOH_2.

2.5. Transfer learning

The estimation performance of the model may deteriorate when applied to other batteries with different aging scenarios. To address this, transfer learning is adopted to refine the source model. This approach uses previously learned knowledge from one domain and applies it to different but the related domains [7]. Although transfer learning is widely used in SOH estimation models, its application in this work complements the unique capability of our model to handle varying lengths of time-series data. This flexibility allows for accurate SOH estimation in scenarios where data may be irregular or incomplete, which is often the case in real-world applications.

As shown in Fig. 4, the baseline model is established using datasets from laboratory tests. During the re-training process, the convolutional layers of the CNN, which have learned to identify relevant patterns in the operational data, are kept “frozen”. This means that the weights of these layers are not updated during the re-training phase. The rationale behind this is that the features learned by the CNN layers in the source domain are expected to be generalizable to the new target domain.

The fully connected (FC) layers, which map the extracted features to the predicted SOH, are then re-trained using the target dataset. The target dataset consists of a smaller set of data from different batteries or operational conditions that the model has not encountered before. The re-training process involves feeding the target dataset into the pre-trained model and updating the FC layers using a suitable optimization algorithm. The loss is computed based on the difference between the predicted SOH values and the actual SOH values from the target dataset. Through this re-training, the FC layers adapt to the new battery characteristics while the pre-trained CNN layers continue to provide relevant feature representations.

By freezing the CNN layers and only re-training the FC layers, the model efficiently adapts to new data with minimal computational cost and reduced risk of overfitting. This strategy uses the knowledge learned from the source domain while ensuring that the model can generalize effectively to the target domain.

3. Results and discussion

3.1. Battery aging test data

Battery aging tests were conducted on the test bench as shown in Fig. 5, and the operational data was collected and stored in the host computer at 5-s intervals. Fig. 6 illustrates the aging trajectories of 24

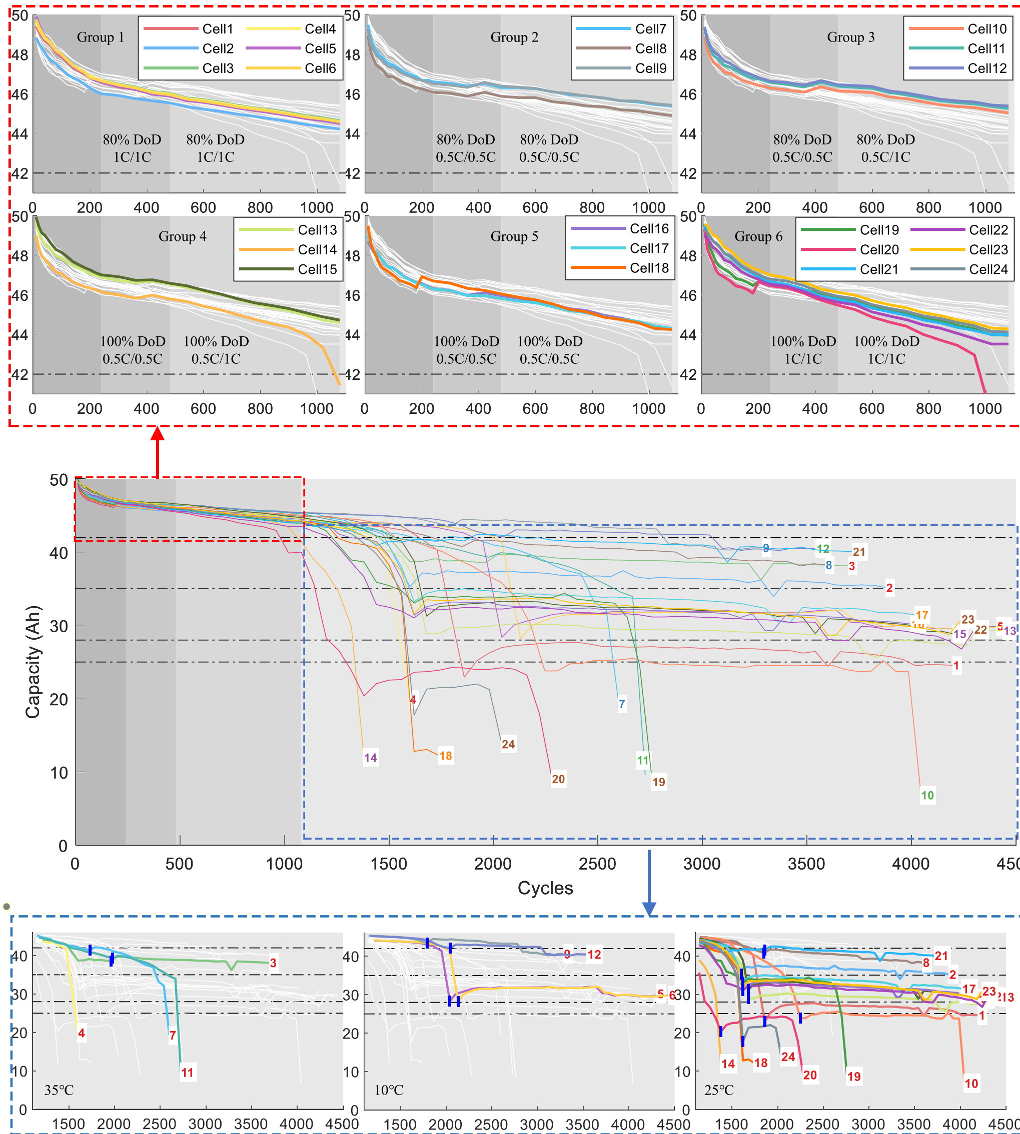

Fig. 6. Battery capacity trajectories obtained from laboratory tests, plotted against test cycles (The short blue line indicates enabling the protective cycling mode.).

battery cells derived from the test data. These curves show that battery degradation significantly depends on working conditions, evidenced by the abrupt changes in trajectory at various points. The non-monotonic features and strong nonlinearity of these trajectories indicate a complex relationship between the SOH and the test variables.

This dataset covers a wide range of battery life from BOL to 20% SOH and below, under varying working conditions. This comprehensive data can capture different degradation patterns and offer a rich source of aging insights for training the ML model.

3.2. Model configuration

ML model training and verification were carried out on a personal computer with an Intel Core i7-12700 (2.1 GHz) CPU, 16 GB of RAM, and Microsoft Windows 10 operating system. The models were developed using Python in Jupyter Notebook.

Selecting the appropriate bin size for histograms during feature extraction from time series data (current, voltage, and temperature) is crucial. Research [28] indicates that a finer signal interval can lead to overfitting, due to the curse of dimensionality, where increased number of features against a constant sample size impacts model performance. Besides, the number of samples depends on the window size and

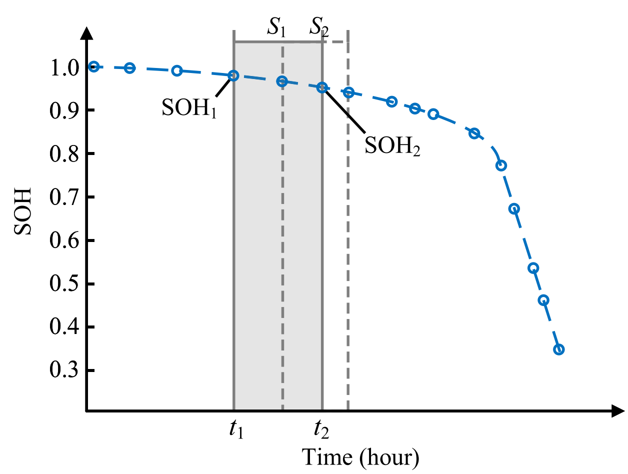

Fig. 7. Demonstration of the aging test data processing with the sliding window.

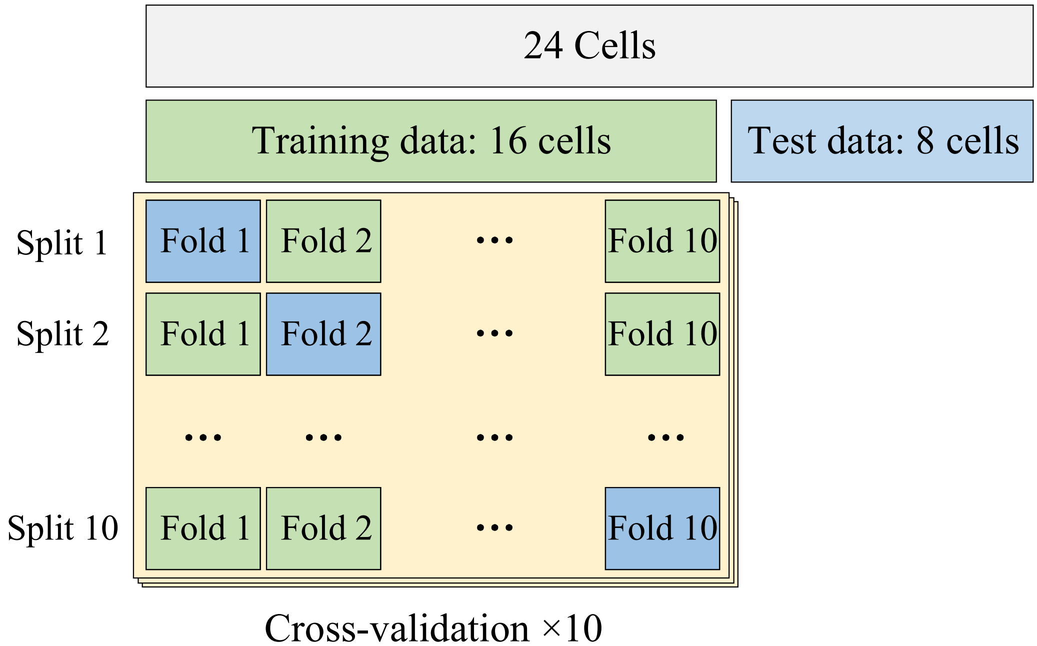

Fig. 8. Source dataset partition.

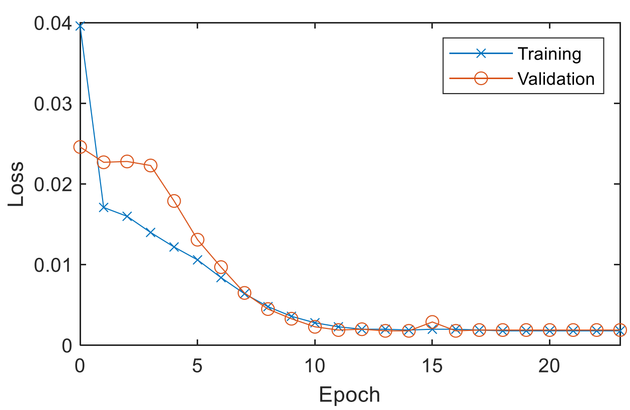

Fig. 9. Training loss and validation loss over the epochs of fold 1 in iteration 1.

Table 3
Parameter settings for feature extraction.

Parameter	Quantity
Bin edges of current (C-rate)	$[0,\ 0.2:0.2:0.8,\ 10]$
Bin edges of voltage (V)	$[0,\ 2.85:0.05:3.0,$ $3.1:0.1:4.0,$ $4.05:0.05:4.15,\ 5]$
Bin edges of temperature (°C)	$[0,\ 10:1:50,\ 100]$
Window size (Number of SOH measurement points)	$[1, 2, 4]$
Window shift (Number of SOH measurement points)	1

window shift. Thus, bin size has to been selected carefully together with window size and window shift. To enhance training generalization, different window sizes are combined to form samples across different time scales.

During aging tests, the battery capacity was measured at various intervals, which were not uniform across the testing period. Specifically, the number of cycles between each capacity test varied, and we chose to use “hours” as the time scale instead of cycles. This choice meant that the capacity data points were unevenly distributed along the time axis, as illustrated in Fig. 7. When preprocessing the experimental data, the window size and shift size were determined by the distribution of sample points in the SOH curve, which is plotted against time in hours. The sliding window approach involves selecting segments of this time-series data, and in our case, both the window size and shift size were varied to accommodate the non-uniform distribution of data points.

Table 3 gives the hyperparameters used for feature extraction in this study, with the resulting 3D histogram dimensions being $5 \times 42 \times 18$. The selected sliding window sizes (1, 2, and 4 measurements) are designed to address the variability in battery data collection intervals, which often differ across datasets. Smaller windows capture short-term trends and operational dynamics, while larger windows incorporate long-term degradation patterns. The hyperparameters for the ML model are listed in Table 4. The CNN hyperparameters, including filter sizes, strides, and pooling layers, were chosen based on established practices in similar applications [33]. These filter sizes are suitable for extracting spatial patterns from the 3D histograms while maintaining computational efficiency. The pooling layers reduce spatial dimensions, helping the model focus on high-level features without overfitting. While these choices were guided by literature and empirical observations, further optimization could enhance model performance and is planned for future work. Typically, the data is divided into training and test sets to evaluate the performance of a ML algorithm statistically. As shown in Fig. 8, 16 out of the 24 cells were allocated for model training and the rest (Cells 03, 08, 12, 14,16,18, 22, 24) were used for testing. 10 times 10-fold cross-validations are conducted on the training set. While 100 is set as the maximum number of training epochs, an early stopping technique is used to prevent overfitting and to ensure that training ceases once the model has converged to an optimal solution.

3.3. Baseline model performance

Evaluation criteria - The model performance in SOH estimation is assessed by root-mean-squared error (RMSE), mean absolute error (MAE), and R^2, which are defined as

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(\hat{y}_i - y_i\right)^2} \tag{3}$$

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n}\left|\hat{y}_i - y_i\right| \tag{4}$$

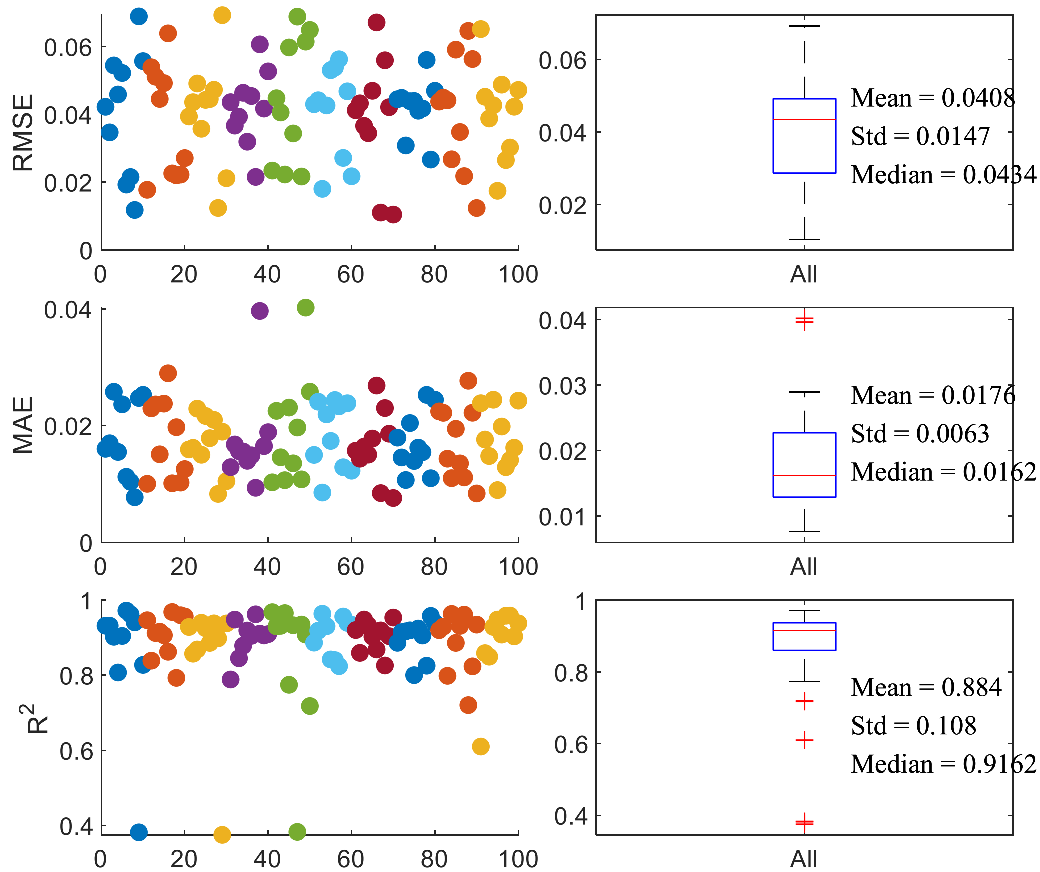

Fig. 10. Metrics distribution of 10 times 10-fold cross-validation.

Table 4
Hyper parameters settings for baseline ML model.

Parameter	Configuration
Convolution3D layer 1	Filter number: 8 Filter size: 2 × 3 × 3 Stride: [1 1 1]
Max Pooling 1	Pool size: 2 × 2 × 2 Stride: [2 2 2]
Convolution3D layer 2	Filter number: 16 Filter size: 2 × 3 × 3 Stride: [1 1 1]
Max Pooling 2	Pool size: 2 × 2 × 2 Stride: [2 2 2]
Convolution3D layer 3	Filter number: 32 Filter size: 2 × 3 × 3 Stride: [1 1 1]
Max Pooling 3	Pool size: 2 × 2 × 2 Stride: [2 2 2]
Hidden unit size of FC layers	150
Dropout rate	0.2
Learning rate	0.001
Batch size	64
Epochs	100
Optimizer	Adam

$$R^2 = 1 - \frac{\sum_{i=1}^{n}\left(\hat{y}_i - y_i\right)^2}{\sum_{i=1}^{n}\left(\bar{y} - y_i\right)^2} \tag{5}$$

where n is the number of test samples. $\hat{y}_i$ and y_i denote the estimated and real SOH for ith sample point, respectively. $\bar{y}$ is the mean value of the real SOH.

Cross-validation - To ensure the robustness and generalizability of the proposed model, we conducted 10 times 10-fold cross-validations on the training set. Fig. 9 shows the training and validation errors over the epochs of one fold in one iteration, demonstrating that the error stabilized and stopped improving before reaching the 100th epoch. This confirms that our approach effectively identified the optimal stopping point, ensuring that the model was neither undertrained nor overtrained.

The distributions of performance metrics for each fold in each iteration are presented in Fig. 10. Additionally, the average and standard deviation of these metrics are calculated to assess the model's stability and reliability across various initial conditions.

The results demonstrate that the model achieves consistent performance across different iterations and data splits, with low variability in the performance metrics. The average RMSE, MAE and R^2 values, along with their standard deviations, indicate the model's robustness and generalizability, reinforcing the reliability of our proposed framework for SOH estimation.

Test set validation - The comparison of predicted SOH values against actual SOH values for both training and test datasets is presented in Fig. 11, and the values of RMSE, MAE and R^2, as well as the distribution of absolute errors of predictions are also provided.

As Fig. 11 indicated, the model demonstrates commendable predictive performance, as evidenced by the metrics obtained on both training and test data. Specifically, the model achieves an RMSE of 0.0417, a MAE of 0.0170, and an R^2 value of 0.914 on the training set. The median absolute estimation error on the training set is 0.009261, increasing to 0.0471 when accounting for 95% of the results. These figures indicate a high degree of accuracy in the model's predictions, with a strong correlation between predicted and actual values as reflected by the R^2 value close to 1. When evaluated on the test set, the model exhibits a slight decrease in performance, yielding an RMSE of

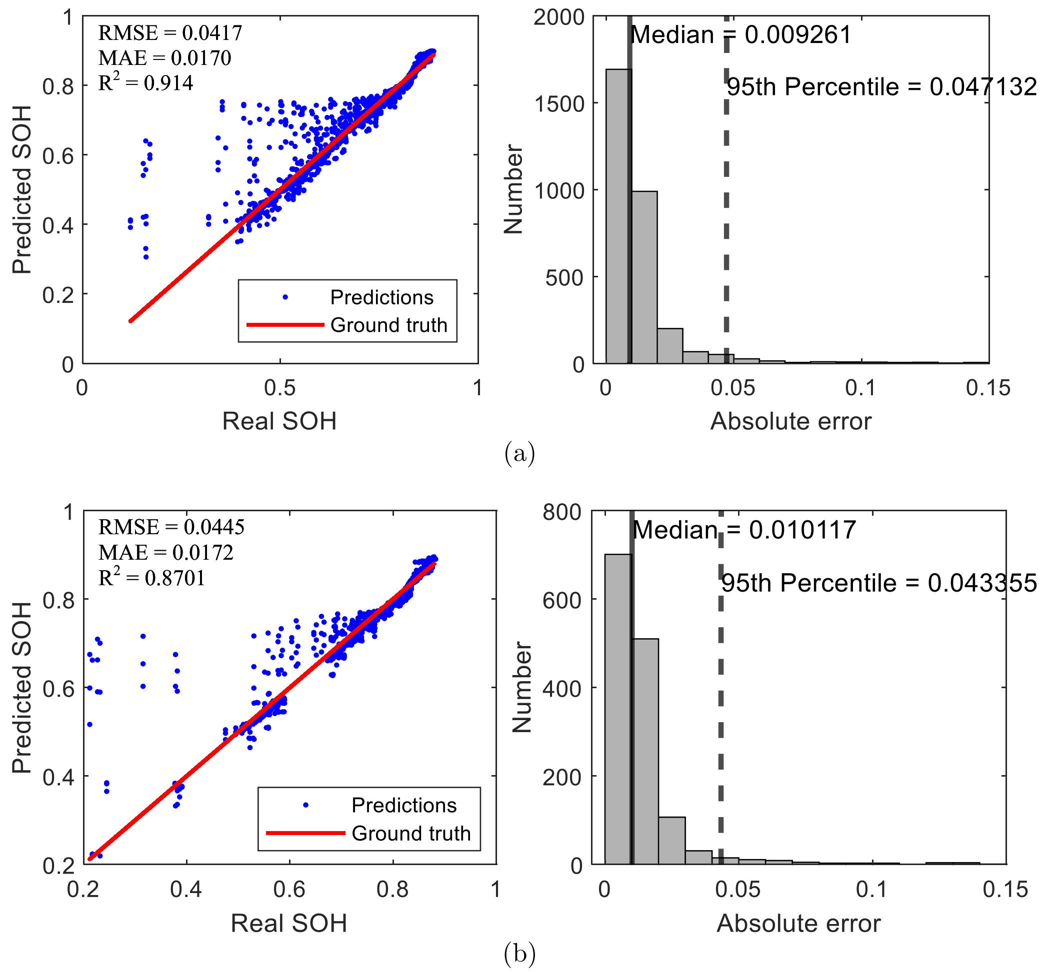

Fig. 11. Predicted values by ML model versus real SOH values: (a) training data; (b) test data.

0.0445, a MAE of 0.0172, and an R^2 of 0.87. Although these metrics are slightly lower compared to those of the training set, they still represent a robust predictive capability. The relatively high R^2 value on the test set suggests that the model retains a substantial proportion of its explanatory power when applied to new, unseen data, which is a crucial aspect of model generalization.

Fig. 12 further displays the predicted SOH curves for the test cells. Higher estimation errors occur in the low SOH segments due to the scarcity of low SOH data in the training dataset, resulting in insufficient generalization in that range. However, the overall predicted SOH curves by the proposed model are generally close to reference curves, which means the model can provide accurate SOH predictions.

3.4. Transfer learning performance

To assess the adaptability and transferability of the baseline model for unseen different batteries under different working scenarios, we used a dataset from [34] for our evaluation. This dataset contains 48 lithium-ion battery cells' time-series data (current, voltage and temperature) of cyclic aging tests. The cells used in this dataset are Sanyo/Pana-sonic UR18650E cylindrical cells. The dataset was partitioned into a training set, a validation set and a test set at a 20:20:60 ratio. The same preprocessing and feature extraction procedures were applied to this dataset.

To evaluate the effectiveness of the transfer learning approach, a comparative analysis was conducted with a purely supervised learning model. The supervised learning model was trained from scratch using only the target dataset and evaluated on the same dataset, while the transfer learning model was fine-tuned on the target training dataset using the pre-trained baseline model. The hyperparameters for transfer learning are given in Table 5.

Table 5
Hyper parameters settings for transfer learning.

Parameter	Configuration
Learning rate	0.001
Batch size	64
Epochs	10
Optimizer	Adam

Fig. 13 presents the overall estimation performance for both models. The results indicate that the transfer learning model outperforms the supervised learning model, achieving lower RMSE and MAE values, and higher R^2. This demonstrates that using pre-trained knowledge from the source dataset and fine-tuning it on the target dataset significantly enhances the model's estimation accuracy and generalization capability. Transfer learning helps the model better capture the characteristics even in the low SOH segments, thus reducing the errors. With the help of transfer learning, where only 20% of the new dataset is used to fine-tune the final layers of the model, a improvement is observed. It can be concluded that transfer learning can boost model adaptability and efficiency in new domains. This approach utilizes the existing knowledge and minimizes the need for extensive retraining.

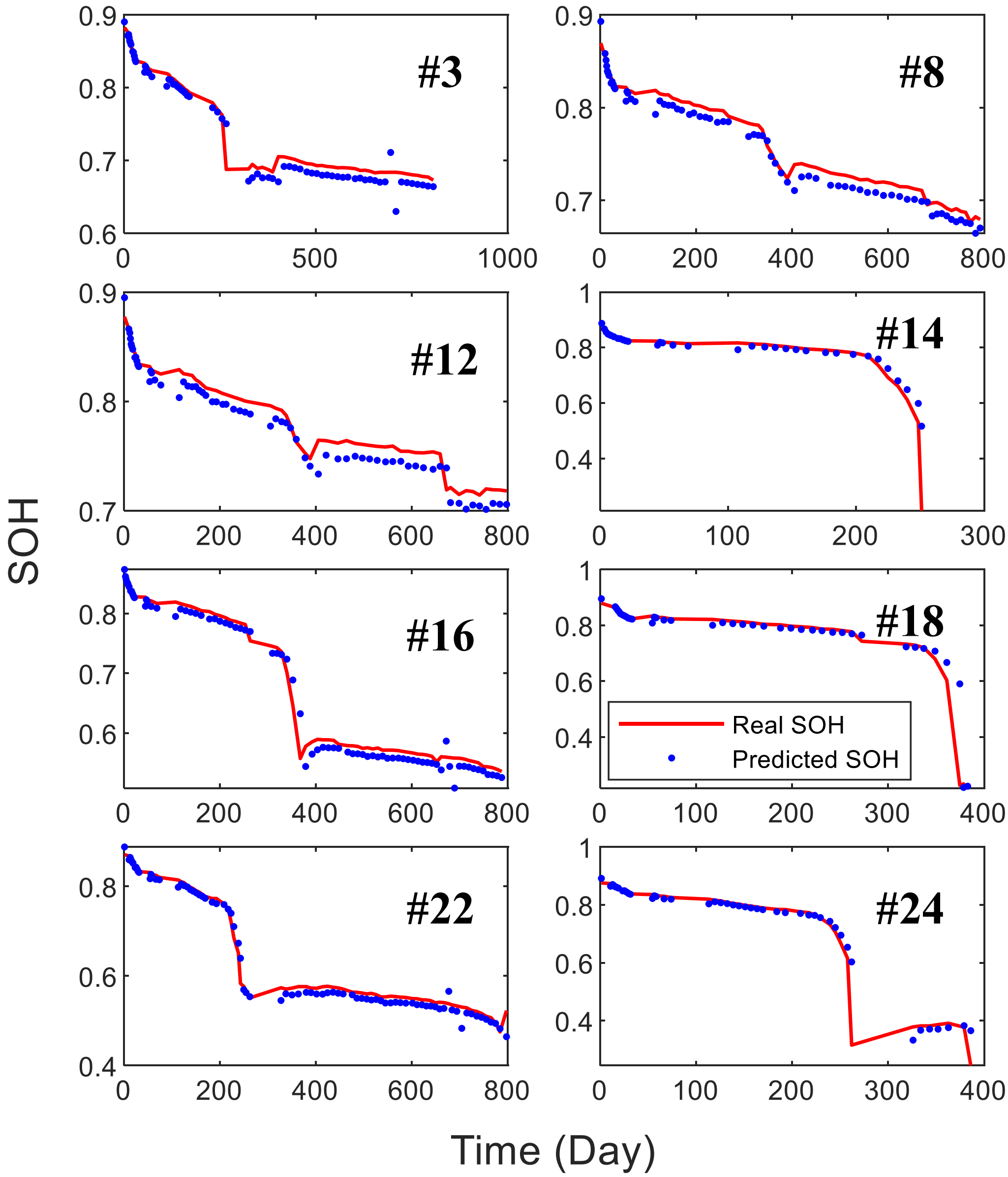

Fig. 12. Predicted versus real SOH curves of test cells for baseline model.

Table 6
Time consumption for model training.

Operation	Baseline model training	Transfer learning
Training time	10 min 10 s	3 s
Prediction time (ms/sample)	0.03	0.03
Training memory usage (MB)	5500	4028

3.5. Resource consumption

To evaluate the computational efficiency of the proposed model, we measured the training time, prediction time per sample, and memory usage during both training and prediction phases for the baseline model and transfer learning. The results are shown in Table 6. It is noteworthy that the integration of transfer learning requires significantly less time to train on new datasets compared to the baseline model. Additionally, the 3D histogram feature extraction technique provides a compact data representation, reducing memory usage during both training and prediction.

The reduced time associated with transfer learning suggest that this approach is more practical for real-time or large-scale battery monitoring applications. By minimizing the need for extensive retraining, transfer learning enables quicker adaptation to new data, making it a cost-effective solution for SOH prediction across different battery types and conditions.

3.6. Discussion of model application

In practical application, a battery manage system (BMS) can collect and aggregate voltage, current and cell temperature data to create real-time histogram information, with the window duration adjustable based on operational needs. This compact histogram data can be directly fed into the model for forecasting, offering a space-efficient alternative to raw time series data. When adapting the baseline model to new batteries, transfer learning requires reduced size of data and much less time for retraining. This is particularly advantageous in scenarios where data is scarce or model reusability is a priority.

Furthermore, with the adoption of cloud–edge technology, the model is trained in the cloud, updating with data from the edge, and

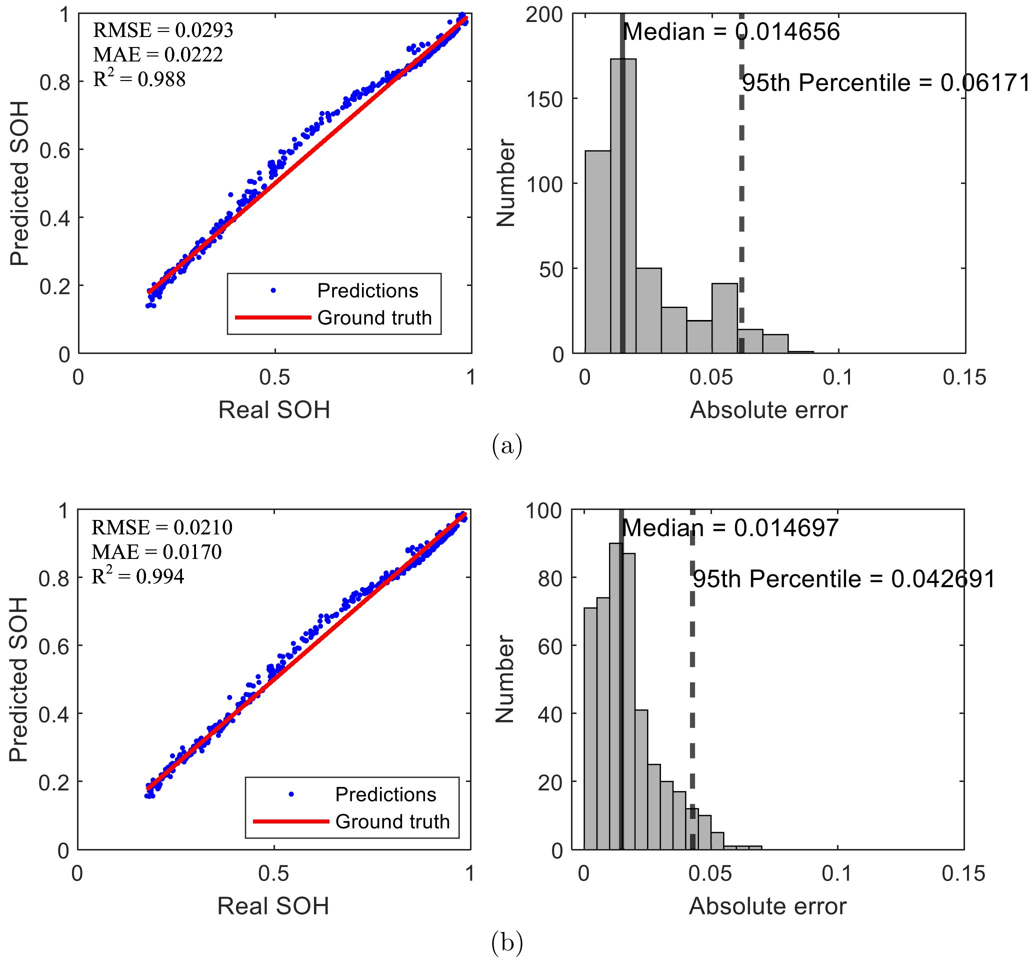

Fig. 13. Predicted values versus real SOH values for new test dataset: (a) supervised learning model; (b) transfer learning model.

sending predictions back for control [35]. The collaboration between cloud computing and onboard BMS enhances computational power, enabling rapid calculations, data storage, and ongoing model refinement for online deployment of advanced SOH prediction algorithms. Compact histogram information and reduced data for transfer learning decrease the need for data transmission, thereby enhancing the framework's adaptability even further.

4. Conclusion

In this study, we proposed a comprehensive framework for estimating the SOH of LIBs from laboratory aging tests to machine learning, complemented by transfer learning. Our methodology benefits from extensive laboratory aging tests, advanced data processing and feature extraction technique, a CNN-based model architecture, and strategic use of transfer learning to handle diverse operational conditions and battery types. The key contributions of our work include the introduction of a 3D histogram feature extraction technique with the varying sliding window, which captures the complex interactions between operational parameters, and the integration of transfer learning, which significantly improves the model's adaptability to different battery types and operational scenarios. Experimental results showcased a robust capability in accurately estimating the SOH of LIBs, whether from laboratory tests or unseen, indicating superior performance in terms of estimation accuracy, model generalization, and training efficiency.

The ability of our model to handle varying lengths of time-series data and adapt to different battery types underscores its practical relevance. In real-world applications, batteries are subjected to various operating conditions. The flexibility and generalizability of our model make it well-suited for these applications, offering a reliable tool for monitoring and predicting battery health in real-time. While our model has demonstrated strong performance, there is room for improvement. Future work will focus on enhancing scalability and practical applicability. Key directions include expanding the dataset to cover more battery chemistries, diverse operational conditions, and real-world dynamic scenarios. Additionally, we aim to systematically study how the amount of target data affects transfer learning performance. This will help establish guidelines for data requirements in applications where labeled data is scarce or expensive.

CRediT authorship contribution statement

Zhi Cao: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Methodology, Formal analysis, Conceptualization. **Wei Gao:** Writing – review & editing, Methodology, Data curation, Conceptualization. **Yuhong Fu:** Supervision, Methodology, Data curation. **Naser Vosoughi Kurdkandi:** Writing – review & editing, Investigation. **Chris Mi:** Writing – review & editing, Supervision, Resources, Project administration, Funding acquisition, Conceptualization.

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.

Acknowledgments

The authors would like to acknowledge the financial support of the California Energy Commission, USA under grant number EPC-19-053.

Data availability

Data will be made available on request.

References

[1] Masias A, Marcicki J, Paxton WA. Opportunities and challenges of lithium ion batteries in automotive applications. ACS Energy Lett 2021;6(2):621–30.
[2] Yang S, Zhang C, Jiang J, Zhang W, Zhang L, Wang Y. Review on state-of-health of lithium-ion batteries: Characterizations, estimations and applications. J Clean Prod 2021;314:128015.
[3] Xiong R, Pan Y, Shen W, Li H, Sun F. Lithium-ion battery aging mechanisms and diagnosis method for automotive applications: Recent advances and perspectives. Renew Sustain Energy Rev 2020;131:110048.
[4] von Bülow F, Meisen T. A review on methods for state of health forecasting of lithium-ion batteries applicable in real-world operational conditions. J Energy Storage 2023;57:105978.
[5] Hu X, Jiang J, Cao D, Egardt B. Battery health prognosis for electric vehicles using sample entropy and sparse bayesian predictive modeling. IEEE Trans Ind Electron 2015;63(4):2645–56.
[6] Yao L, Xu S, Tang A, Zhou F, Hou J, Xiao Y, Fu Z. A review of lithium-ion battery state of health estimation and prediction methods. World Electr Veh J 2021;12(3):113.
[7] Che Y, Hu X, Lin X, Guo J, Teodorescu R. Health prognostics for lithium-ion batteries: mechanisms, methods, and prospects. Energy Environ Sci 2023.
[8] Hosen MS, Youssef R, Kalogiannis T, Van Mierlo J, Berecibar M. Battery cycle life study through relaxation and forecasting the lifetime via machine learning. J Energy Storage 2021;40:102726.
[9] Klass V, Behm M, Lindbergh G. A support vector machine-based state-of-health estimation method for lithium-ion batteries under electric vehicle operation. J Power Sources 2014;270:262–72.
[10] Yang D, Wang Y, Pan R, Chen R, Chen Z. State-of-health estimation for the lithium-ion battery based on support vector regression. Appl Energy 2018;227:273–83.
[11] Yang D, Zhang X, Pan R, Wang Y, Chen Z. A novel gaussian process regression model for state-of-health estimation of lithium-ion battery using charging curve. J Power Sources 2018;384:387–95.
[12] Richardson RR, Birkl CR, Osborne MA, Howey DA. Gaussian process regression for in situ capacity estimation of lithium-ion batteries. IEEE Trans Ind Inf 2018;15(1):127–38.
[13] von Bülow F, Hahn Y, Meyes R, Meisen et al T. Transparent and interpretable state of health forecasting of lithium-ion batteries with deep learning and saliency maps. Int J Energy Res 2023. 2023.
[14] Che Y, Zheng Y, Forest FE, Sui X, Hu X, Teodorescu R. Predictive health assessment for lithium-ion batteries with probabilistic degradation prediction and accelerating aging detection. Reliab Eng Syst Saf 2024;241:109603.
[15] Tiane A, Okar C, Chaoui H. Adversarial defensive framework for state of health prediction of.
[16] Costa N, Sanchez L, Ansean D, Dubarry M. Li-ion battery degradation modes diagnosis via convolutional neural networks. J Energy Storage 2022;55:105558.
[17] Li Y, Li K, Liu X, Wang Y, Zhang L. Lithium-ion battery capacity estimation - a pruned convolutional neural network approach assisted with transfer learning. Appl Energy 2021;285:116410.
[18] Zhao G, Zhang C, Duan B, Shang Y, Kang Y, Zhu R. State-of-health estimation with anomalous aging indicator detection of lithium-ion batteries using regression generative adversarial network. IEEE Trans Ind Electron 2022;70(3):2685–95.
[19] Gao J, Yang D, Duan B, Wang S, Li Z, Wang L, Wang K. State of health estimation of lithium-ion batteries based on Mixers-bidirectional temporal convolutional neural network. J Energy Storage 2023;73:109248.
[20] Liu Y, Li Q, Wang K. Revealing the degradation patterns of lithium-ion batteries from impedance spectroscopy using variational auto-encoders. Energy Storage Mater 2024;69(3):103394.
[21] Lee J, Sun H, Liu Y, Li X. A machine learning framework for remaining useful lifetime prediction of li-ion batteries using diverse neural networks. Energy AI 2024;15:100319.
[22] Wang S, Fan Y, Jin S, Takyi-Aninakwa P, Fernandez C. Improved anti-noise adaptive long short-term memory neural network modeling for the robust remaining useful life prediction of lithium-ion batteries. Reliab Eng Syst Saf 2023;230:108920.
[23] Wang S, Wu F, Takyi-Aninakwa P, Fernandez C, Stroe DI, Huang Q. Improved singular filtering-Gaussian process regression-long short-term memory model for whole-life-cycle remaining capacity estimation of lithium-ion batteries adaptive to fast aging and multi-current variations. Energy 2023;284:128677.
[24] Dos Reis G, Strange C, Yadav M, Li S. Lithium-ion battery data and where to find it. Energy AI 2021;5:100081.
[25] Li C, Zhang H, Ding P, Yang S, Bai Y. Deep feature extraction in lifetime prognostics of lithium-ion batteries: Advances, challenges and perspectives. Renew Sustain Energy Rev 2023;184:113576.
[26] Severson KA, Attia PM, Jin N, Perkins N, Jiang B, Yang Z, Chen MH, Aykol M, Herring PK, Fraggedakis et al D. Data-driven prediction of battery cycle life before capacity degradation. Nat Energy 2019;4(5):383–91.
[27] Zhang Y, Zhao M. Cloud-based in-situ battery life prediction and classification using machine learning. Energy Storage Mater 2023.
[28] von Bülow F, Wassermann M, Meisen T. State of health forecasting of lithium-ion batteries operated in a battery electric vehicle fleet. J Energy Storage 2023;72:108271.
[29] Li R, Bao L, Chen L, Zha C, Dong J, Qi N, Tang R, Lu Y, Wang M, Huang R, et al. Accelerated aging of lithium-ion batteries: Bridging battery aging analysis and operational lifetime prediction. Sci Bull 2023.
[30] Nuhic A, Terzimehic T, Soczka-Guth T, Buchholz M, Dietmayer K. Health diagnosis and remaining useful life prognostics of lithium-ion batteries using data-driven methods. J Power Sources 2013;239:680–8.
[31] Greenbank S, Howey D. Automated feature extraction and selection for data-driven models of rapid battery capacity fade and end of life. IEEE Trans Ind Inf 2021;18(5):2965–73.
[32] von Bülow F, Mentz J, Meisen T. State of health forecasting of lithium-ion batteries applicable in real-world operational conditions. J Energy Storage 2021;44:103439.
[33] Goodfellow I, Bengio Y, Courville A. Deep learning. MIT Press; 2016, http://www.deeplearningbook.org.
[34] Sauer DU. Time-series cyclic aging data on 48 commercial NMC/graphite Sanyo/Panasonic UR18650E cylindrical cells. 2021, [Online]. Available: https://publications.rwth-aachen.de/record/818642.
[35] Shi D, Zhao J, Eze C, Wang Z, Wang J, Lian Y, Burke AF. Cloud-based artificial intelligence framework for battery management system. Energies 2023;16(11):4403.