ISSN: 2688-8645

# Lithium-ion Battery SOC Estimation based on CNN-Transformer Model

Yuwen Tao, Wenyuan Li, Zhendong Wang and Weijie Yang

School of Navigation, Shandong Jiaotong University, Weihai, Shandong, 264209, China

## Abstract

**The state of charge (SOC) of lithium-ion batteries is a critical state parameter in battery management systems (BMS), being closely associated with energy management, charge–discharge control and operational safety, yet it cannot be directly measured by conventional sensors. In recent years, deep learning methods have demonstrated considerable potential for SOC estimation owing to their strong capacity for nonlinear feature extraction. However, local dynamic characteristics and long-range temporal dependencies embedded in battery operating data are difficult to be modelled simultaneously by a single network architecture, thereby limiting SOC estimation accuracy under complex dynamic operating conditions. To address this limitation, a SOC estimation model integrating a multi-scale convolutional neural network (CNN) with a Transformer encoder is proposed. Local dynamic features from voltage and current sequences are extracted using multi-scale CNNs, while long-range temporal dependencies are further established by the Transformer encoder, thereby enhancing the representation of dynamic battery operating characteristics. During data preprocessing, SOC reference values are calculated using the ampere-hour integration method, and temporal input samples are constructed through normalization and a sliding-window strategy, providing a reliable basis for model training and prediction. Experiments are conducted using the publicly available CALCE INR 18650-20R lithium-ion battery dataset under four dynamic current profiles and three ambient temperatures. Model performance is evaluated using the root mean square error (RMSE), mean absolute error (MAE) and coefficient of determination ($R^2$). The proposed CNN-Transformer model achieves high SOC estimation accuracy, with an RMSE of 0.9371%, an MAE of 0.7913% and an $R^2$ of 0.9957 under the DST profile at 25 ° C. Compared with CNN, bidirectional long short-term memory (BiLSTM) and Transformer models, the RMSE is reduced by up to 52.57%. Further evaluation across different temperatures and dynamic operating conditions demonstrates the favourable environmental adaptability and dynamic-condition robustness of the proposed model. These results establish an effective data-driven framework for high-precision SOC estimation and state monitoring in lithium-ion battery management systems.**

## Keywords

**Lithium-ion Battery; State of Charge; SOC Estimation; Convolutional Neural Network; Transformer.**

## 1. Introduction

Lithium-ion batteries have become a key power source for electric vehicles and energy storage systems owing to their high energy density, long cycle life and low self-discharge rate. As one of the most important state variables in battery management systems (BMSs), the state of charge (SOC) is closely associated with charge–discharge control, remaining driving-range estimation and safety protection strategies. However, SOC cannot be directly measured by conventional sensors and must instead be estimated indirectly from measurable physical quantities, such as terminal voltage and current. Consequently, accurate SOC estimation has remained a longstanding challenge in battery management.

Conventional SOC estimation methods mainly include the ampere-hour integration method [1,2], the open-circuit voltage method [3] and Kalman-filter-based approaches and their variants [4]. Although the ampere-hour integration method is straightforward to implement, cumulative errors are inevitably introduced and an accurate initial SOC is required [5–7]. The open-circuit voltage method requires the battery to remain at rest for an extended period, making it unsuitable for real-time estimation [8,9]. Kalman-filter-based methods rely on accurate equivalent-circuit models, whereas the pronounced nonlinear and time-varying characteristics exhibited by batteries under practical operating conditions make their model parameters difficult to determine accurately [10–13]. These limitations have motivated the development of data-driven approaches for SOC estimation.

In recent years, deep learning techniques have demonstrated substantial potential for battery SOC estimation owing to their powerful nonlinear mapping capability and end-to-end learning characteristics [14–16]. Zhang Yuanjin et al. [17] developed a battery SOC estimation method based on an adaptive unscented Kalman filter (AUKF) combined with a back-propagation (BP) neural network. The estimation accuracy of the UKF was improved adaptively through a sampling strategy, and enhanced accuracy and robustness were achieved compared with conventional neural networks. Although such approaches can alleviate some limitations associated with recurrent neural networks (RNNs) and account for long-term dependencies within sequential data, insufficient attention has been paid to the spatially nonlinear local features contained in the input data. To address the susceptibility of BPNNs to local optima, Mao et al. [18] introduced a particle swarm optimization algorithm based on a Lévy flight strategy to optimize network weights and biases. Benefiting from the feature extraction capability of convolutional neural networks (CNNs), Fan et al. [19] developed a CNN-based model with a U-Net architecture, in which symmetric padding was employed to improve SOC estimation accuracy at sequence boundaries while variable-length input sequences were accommodated. Nevertheless, the U-Net architecture relies primarily on local receptive fields, and information loss may occur during the transmission of long-range dependencies. RNNs are also susceptible to vanishing and exploding gradients during backpropagation and exhibit limited capability in processing long sequences, which can compromise SOC estimation performance when large volumes of sequential data are involved. To address these limitations, Bian et al. [20] developed a dual-channel CNN–bidirectional RNN (MCNN-BRNN) model, in which local disturbance-resistant features were extracted by CNNs and time-varying information was captured by bidirectional RNNs. Chen et al. [21] proposed an SOC estimation algorithm based on an LSTM-RNN, in which estimation accuracy was subsequently improved through input expansion and a control-output strategy based on the ampere-hour integration method. Shen et al. [22] employed a Transformer architecture for battery SOC prediction and introduced an immersion-invariant adaptive observer to reduce estimation errors. It was demonstrated that, owing to the self-attention mechanism, the Transformer exhibited superior performance to RNNs in processing long sequential data.

On the basis of these considerations, a hybrid CNN–Transformer model is proposed for lithium-ion battery SOC estimation. Local feature extraction by the multi-scale CNN is integrated with global temporal modelling by the Transformer, enabling local dynamic characteristics and long-range temporal dependencies to be jointly represented. Deep feature fusion is achieved through residual connections, while an attention-pooling mechanism is introduced to adaptively aggregate the output representations of the Transformer. Voltage and current are used as the model inputs, thereby allowing SOC to be estimated directly from readily measurable electrical quantities.

ISSN: 2688-8645

# 2. Model Construction

## 2.1. CNN Model

Convolutional neural networks (CNNs) are a class of feed-forward neural networks characterized by convolutional operations and deep hierarchical architectures, and have been widely employed in image processing and temporal signal analysis. In particular, one-dimensional convolutional neural networks (1D-CNNs) are well suited to feature extraction from time-series data. Their fundamental operation is performed by sliding convolutional kernels along the temporal dimension, whereby signals within local receptive fields are subjected to weighted summation followed by nonlinear transformation. In this manner, local feature patterns with translational invariance can be extracted. Multiple convolutional kernels are typically employed within a convolutional layer, and convolution operations are performed on the input data to progressively derive higher-level and more expressive representations. A nonlinear transformation is subsequently applied to the convolutional outputs through an activation function, thereby enhancing the representational capacity of the network. This process can be expressed as Eq. (1), where $h_i$ denotes the features extracted by the $i$ th convolutional layer, $x_i$denotes the input to this layer, $\sigma$ denotes the activation function, * denotes the convolution operation, and $w_i$ and $b_i$ represent the convolutional kernel weight matrix and bias vector, respectively.

$$h_i = \sigma(x_i * x_i + b_i) \tag{1}$$

For battery SOC estimation, 1D-CNNs can effectively capture local variations in temporal signals such as voltage and current. For example, changes in the slope of voltage during plateau regions and response patterns before and after abrupt current variations can be represented through local convolutional operations. Multi-scale convolution represents an important extension of the conventional CNN architecture. By employing convolutional kernels with different receptive-field sizes in parallel, signal patterns across multiple temporal scales can be captured simultaneously: smaller kernels are primarily used to characterize fine-grained variations, whereas larger kernels are employed to capture broader temporal trends.

## 2.2. Transformer Model

The core of the Transformer architecture is the multi-head self-attention (MHSA) mechanism. Through self-attention, information from all other positions within a sequence can be dynamically attended to when each individual position is processed, while the relevance between different positions is quantified by learned attention weights. The multi-head mechanism further enables diverse dependencies to be captured from different representation subspaces through multiple attention heads operating in parallel. The corresponding computational procedure is given by Eqs. (2) and (3).

$$Q = ZW^Q, K = ZW^K, V = ZW^V \tag{2}$$

$$Q_{Attn} = (Q, K, V) = Soft\max(\frac{QK^T}{\sqrt{d_k}})V \tag{3}$$

$$Soft\max(x_i) = \frac{\exp(x_i)}{\sum_k \exp(x_k)} \tag{4}$$

First, the input incorporating positional encoding is linearly projected into the query matrix $Q$, key matrix $K$ and value matrix $V$. An attention score matrix is then obtained by computing the dot product between $Q$ and $K$, followed by scaling with the square root of the dimensionality of the key vectors to improve gradient stability. The scaled scores are subsequently normalized using the Softmax function to obtain the attention weights. Finally, the weighted values are multiplied by $V$ to produce the self-attention output. The Softmax function maps a $k$-

dimensional vector to a vector whose elements lie within (0,1) and sum to unity, as expressed in Eq. (4), where $x_i$ denotes the $i$ th component of the input vector. Through repeated linear transformations of $Q$, $K$ and $V$, multiple self-attention outputs are computed in parallel. These outputs are subsequently concatenated and subjected to a further linear transformation, yielding the final representation. In this manner, diverse temporal features can be captured from different representation subspaces.

Unlike LSTM and RNN architectures, which are characterized by inherently directional sequential propagation, the Transformer is not subject to an intrinsic constraint of stepwise information transmission. Instead, its attention mechanism enables contextual information both preceding and following the current time step to be considered simultaneously, thereby facilitating the modelling of global temporal dependencies.

For SOC estimation, the instantaneous SOC is associated not only with recent voltage and current observations but also with long-term dependencies arising from earlier charge–discharge histories. For example, voltage rebound following high-current discharge and differences in polarization effects induced by distinct charge–discharge histories require the capability to capture long-range temporal dependencies. The self-attention mechanism of the Transformer is therefore well suited to this class of problems, as information from the entire input sequence can be accessed directly at each time step, thereby mitigating the attenuation of long-range information commonly encountered in RNN-based architectures. In addition, the inherent parallelizability of the Transformer substantially improves computational efficiency during model training.

## 2.3. CNN–Transformer Hybrid Model

To fully exploit both the local variations embedded in lithium-ion battery operating data and the long-range temporal dependencies evolving over extended time scales, a CNN–Transformer hybrid model is developed for battery state-of-charge (SOC) estimation. In the proposed architecture, voltage and current measurements acquired during battery operation are used as the input features. Local temporal features are first extracted using a multi-scale convolutional neural network (CNN), after which global dependencies across different time steps are established using a Transformer encoder. SOC regression is subsequently performed through attention pooling followed by fully connected output layers. The overall architecture of the proposed model is illustrated in Fig. 1. The principal components comprise input feature construction, input feature mapping, multi-scale CNN feature extraction, residual feature fusion, positional encoding, Transformer encoding, attention pooling and SOC output.

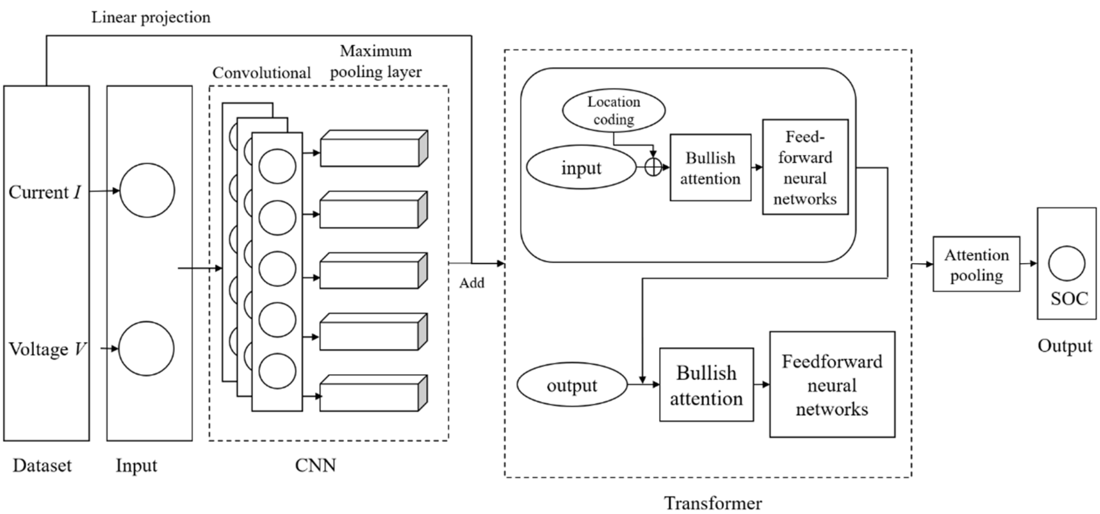



**Fig 1.** Overall architecture of the CNN-Transformer hybrid model

ISSN: 2688-8645

The proposed model follows a progressive processing pipeline comprising local feature extraction, global dependency modelling, information aggregation and regression output. First, the input sequences are processed using a multi-scale one-dimensional convolutional neural network, in which convolutional kernels with different receptive-field sizes are employed to capture local dynamic features of battery voltage and current across multiple temporal scales. The convolutionally extracted features are subsequently fused with the high-dimensional representations of the input features through residual connections, thereby enhancing feature representation while preserving information from the original inputs. The fused representations are then combined with positional encodings and fed into the Transformer encoder. Through the multi-head self-attention mechanism, relationships among different time steps are established, allowing long-range temporal dependencies associated with SOC evolution to be further captured.

Following Transformer encoding, attention pooling is employed to perform weighted aggregation of the features within each temporal window. In this manner, greater attention can be adaptively assigned to time steps that exert stronger influences on SOC estimation. Finally, a fully connected network is used to establish a nonlinear mapping from the high-dimensional representations to SOC, while a Sigmoid function is applied to constrain the output to the valid SOC range, thereby yielding the final SOC estimate. During model training, the mean squared error (MSE) is adopted as the loss function, and the network parameters are iteratively updated using the AdamW optimizer. The detailed model configuration is provided in Table 1.

**Table 1.** Parameter configuration of the CNN-Transformer hybrid model

| Type | Parameters | Value/Name |
|---|---|---|
| Data structure | Sliding window dimensions | 50 |
| | Data collection intervals | 1s |
| | Normalization method | Z-scoreStandardization |
| | Activation Function (Hidden Layer) | GELU |
| | Activation Function (Output Layer) | Sigmoid |
| Training process | Optimizer | AdamW |
| | Dropout | 0.2 |
| | Initial learning rate | 1e-4 |
| | Small batch sizes | 64 |
| | Training rounds | 150 |
| | Loss function | MSE |

## 3. Data Description

### 3.1. Dataset Description

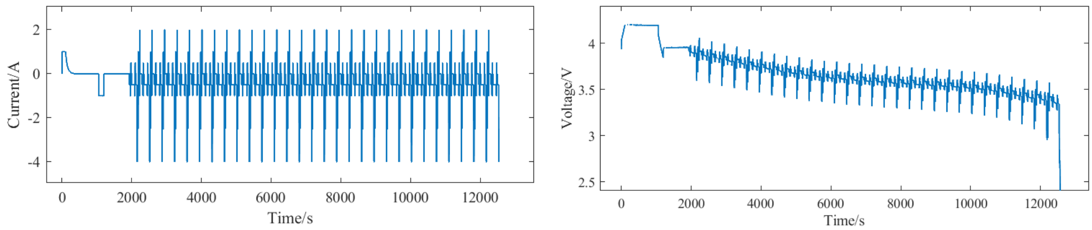
The publicly available lithium-ion battery dataset provided by the Center for Advanced Life Cycle Engineering (CALCE) at the University of Maryland is adopted in this study. The INR 18650-20R cell is selected as the battery under investigation, and its specifications are summarized in Table 2. This dataset has been widely used in lithium-ion battery modelling and state estimation studies and has served as an important benchmark dataset in the field of battery management.

The dataset comprises several dynamic current profiles, including the Dynamic Stress Test (DST), Federal Urban Driving Schedule (FUDS), US06 Highway Driving Schedule (US06) and Beijing Dynamic Stress Test (BJDST), with tests conducted at 0 °C, 25 °C and 45 °C under initial SOC conditions of 80% and 50%, respectively.

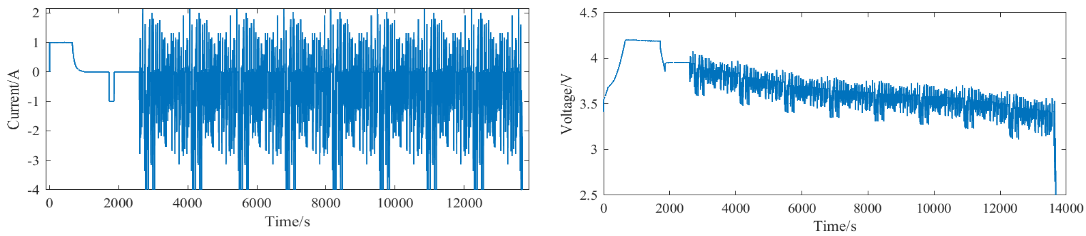
ISSN: 2688-8645

**Table 2.** INR 18650-20R Battery Specifications [23]

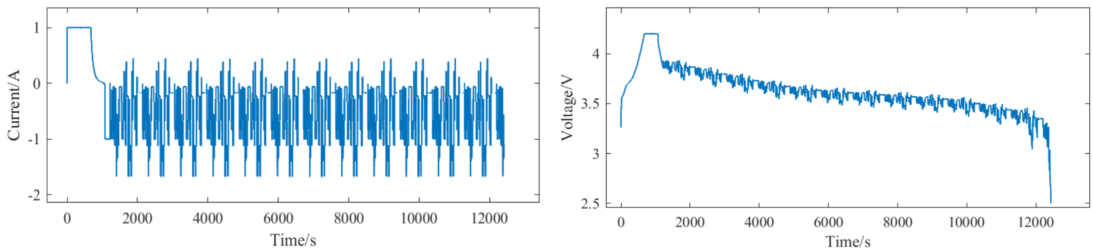
| Battery (Parameters) | Specifications (Value) |
|---|---|
| Capacity Rating | 2000 mAh |
| Cell Chemistry | LiNiMnCo/Graphite |
| Weight (w/o safety circuit) | 45 g |
| Diameter | 18.33 mm ± 0.07 mm |
| Length | 64.85 mm ± 0.15 mm |
| Special Notes | Tab length not included in the dimensions |



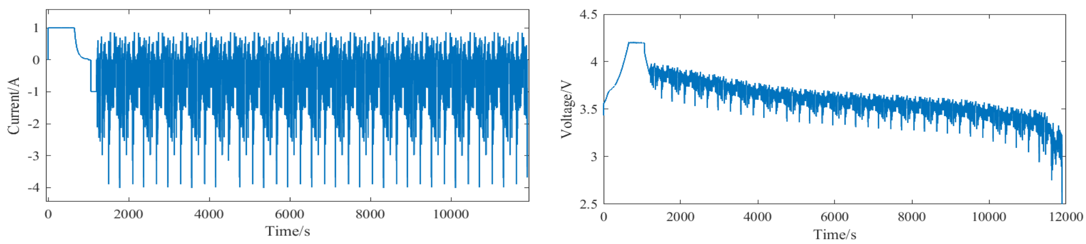

a) DST




b) FUDS




c) BJDST




d) US06

**Fig 2.** Current and voltage variation curves under different discharge conditions at 25°C

Among these profiles, the DST profile consists of multiple current steps with varying amplitudes and durations and is designed to evaluate the battery response under dynamic stress conditions [24]. Before the DST test, the battery is charged to a predetermined state and allowed to rest. The battery is subsequently adjusted to the specified initial SOC and subjected to dynamic discharge according to the prescribed DST current profile until the cutoff voltage is reached. During the test, battery voltage, current, temperature and SOC are recorded at fixed sampling intervals, providing time-series data for subsequent SOC estimation. In particular, the battery current, voltage and surface temperature are recorded at 1-s intervals. The voltage and current profiles of the four operating conditions at 25 °C are shown in Fig. 2.

## 3.2. Data Preprocessing

### 3.2.1. Ampere-hour Integration Method

As SOC cannot be directly measured by conventional sensors, the ampere-hour integration method is employed to calculate SOC during the charge–discharge process. In this manner, reliable SOC reference labels are provided for the dataset, thereby establishing the basis for model training and the evaluation of prediction performance.

During preprocessing of the raw experimental data, data corresponding to the battery resting periods are first removed, and only valid data collected during dynamic charge–discharge processes are retained, thereby reducing the interference of irrelevant samples with model training. Subsequently, the SOC is estimated using the ampere-hour integration method, with the corresponding calculation given in Eq. (5).

$$SOC_{\mathrm{t}} = SOC_{t_0} - \frac{\eta}{Q_n}\int_{t_0}^{t} I_t dt \tag{5}$$

Accordingly, $SOC(t)$ denotes the SOC at time $t$, $SOC(t_0)$ denotes the initial SOC, $Q_n$ denotes the rated battery capacity, $I(\tau)$ denotes the charge–discharge current at time $\tau$, with the discharge current defined as positive, and $\eta$ denotes the Coulombic efficiency, which characterizes the charge transfer efficiency during the battery charging and discharging processes.

To evaluate the estimation capability of the proposed model under different initial SOC conditions, data with an initial SOC of 80% are used as the training set, whereas data with an initial SOC of 50% are retained as an independent prediction set. This data partitioning strategy avoids direct overlap between the training and prediction sets while providing a more stringent assessment of the model's generalization capability across different initial SOC conditions.

### 3.2.2. Data Normalization

As voltage and current are characterized by different physical units and numerical ranges, direct input into the model may cause discrepancies in feature scales to affect the optimization of model parameters. Therefore, prior to model training, the input features are standardized as expressed in Eq. (6):

$$x^{'} = \frac{x-\mu}{\sigma} \tag{6}$$

Where $x$ denotes the original feature value, $\eta$ and $\sigma$ denote the mean and standard deviation, respectively, of the corresponding feature in the training set, and $x'$ denotes the standardized feature value. It should be noted that the prediction set is standardized using the normalization parameters derived from the training set, thereby preventing information leakage during data preprocessing.

Following this preprocessing procedure, the model input consists of two features, namely voltage and current, as expressed in Eq. (7):

$$X = [V, I] \tag{7}$$

### 3.2.3. Sliding-Window Strategy

As the SOC of lithium-ion batteries exhibits pronounced temporal correlations, the voltage and current measurements at a single time step alone are insufficient to fully characterize their dynamic evolution. Therefore, a sliding-window strategy is employed to reconstruct the continuous time-series data, as illustrated in Fig. 3.

ISSN: 2688-8645

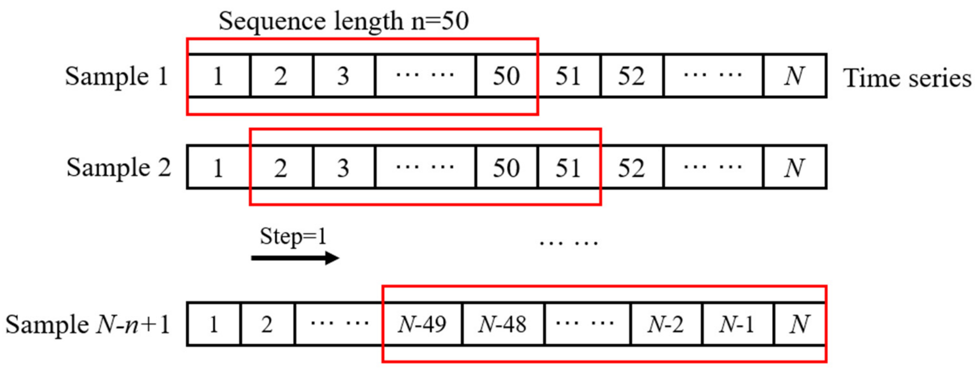



**Fig 3.** Sliding window sampling schematic

Assuming that the original data sequence has a length of $n$ and that the sliding-window length is $L$, $L$ consecutive time steps of voltage and current are selected at each iteration to form an input sequence, with the SOC at the corresponding time step used as the prediction target. The window is progressively shifted in chronological order, thereby transforming the original continuous measurements into sequential samples with explicit temporal dependencies. In this study, the window length is set to 50, such that each input sample contains voltage and current information from 50 consecutive sampling instants. This strategy provides an effective input representation for subsequent extraction of local temporal features by the CNN and modelling of long-range dependencies within the sequence by the Transformer.

### 3.3. Model Evaluation Metrics

Three commonly used regression metrics are employed to evaluate the SOC estimation performance of the proposed model:

(1) Mean Absolute Error (MAE)

$$MAE = \frac{1}{N}\sum_{i=1}^{N}\left|SOC_i - S\hat{O}C_i\right| \tag{8}$$

*MAE* represents the mean absolute difference between the estimated and reference SOC values. It is relatively insensitive to outliers and provides an intuitive measure of the overall magnitude of estimation errors.

(2) Root Mean Square Error (RMSE)

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(SOC_i - S\hat{O}C_i\right)^2} \tag{9}$$

*RMSE* imposes a greater penalty on large errors and is therefore more sensitive to substantial deviations in the estimation results.

(3) Coefficient of Determination ($R^2$)

$$R^2 = 1 - \frac{\sum_{i=1}^{N}\left(SOC_i - S\hat{O}C_i\right)^2}{\sum_{i=1}^{N}\left(SOC_i - S\bar{O}C_i\right)^2} \tag{10}$$

The coefficient of determination $R^2$ quantifies the extent to which the variance of the target variable is explained by the model. Its value ranges from (-∞, 1], with values closer to 1 indicating better agreement between the estimated and reference SOC values.

ISSN: 2688-8645

# 4. Experimental Results and Analysis

To evaluate the SOC estimation performance of the proposed CNN–Transformer model under different operating environments and dynamic operating conditions, and to further investigate the influence of model architecture on estimation accuracy, comparative experiments were conducted under different temperatures, dynamic operating conditions and model architectures. RMSE, MAE and $R^2$ were employed as the evaluation metrics, with RMSE and MAE used to characterize the magnitude of estimation errors and $R^2$ used to quantify the agreement between the estimated results and the reference SOC values. The experimental results are presented below.

## 4.1. Analysis of Results under Different Temperatures

To investigate the influence of ambient temperature variations on the SOC estimation performance of lithium-ion batteries, experiments were conducted using data from the DST dynamic operating condition at three temperatures, namely 0 °C, 25 °C and 45 °C. The corresponding estimation results under each temperature condition are shown in Fig. 4.

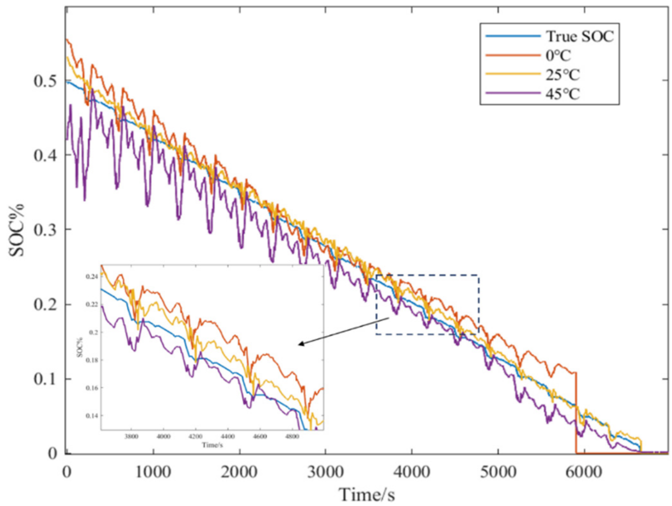



**Fig 4.** Comparison of SOC estimation at different temperatures

Temperature variations can alter the electrochemical reaction kinetics and polarization characteristics within lithium-ion batteries, thereby modifying the mapping relationship between voltage, current and SOC and increasing the difficulty of SOC feature extraction and state representation by data-driven models. As shown in Table 3, the optimal estimation performance is achieved at 25 °C, whereas estimation errors are increased under both low- and high-temperature conditions. Compared with 25 °C, the RMSE and MAE at 0 °C increase to 2.2921% and 1.9456%, respectively, while $R^2$ decreases to 0.9670. Under the high-temperature condition of 45 °C, the estimation errors are further increased, with RMSE and MAE reaching 3.0216% and 2.2610%, respectively, while $R^2$ decreases to 0.9586.

**Table 3.** Estimation errors for different temperatures

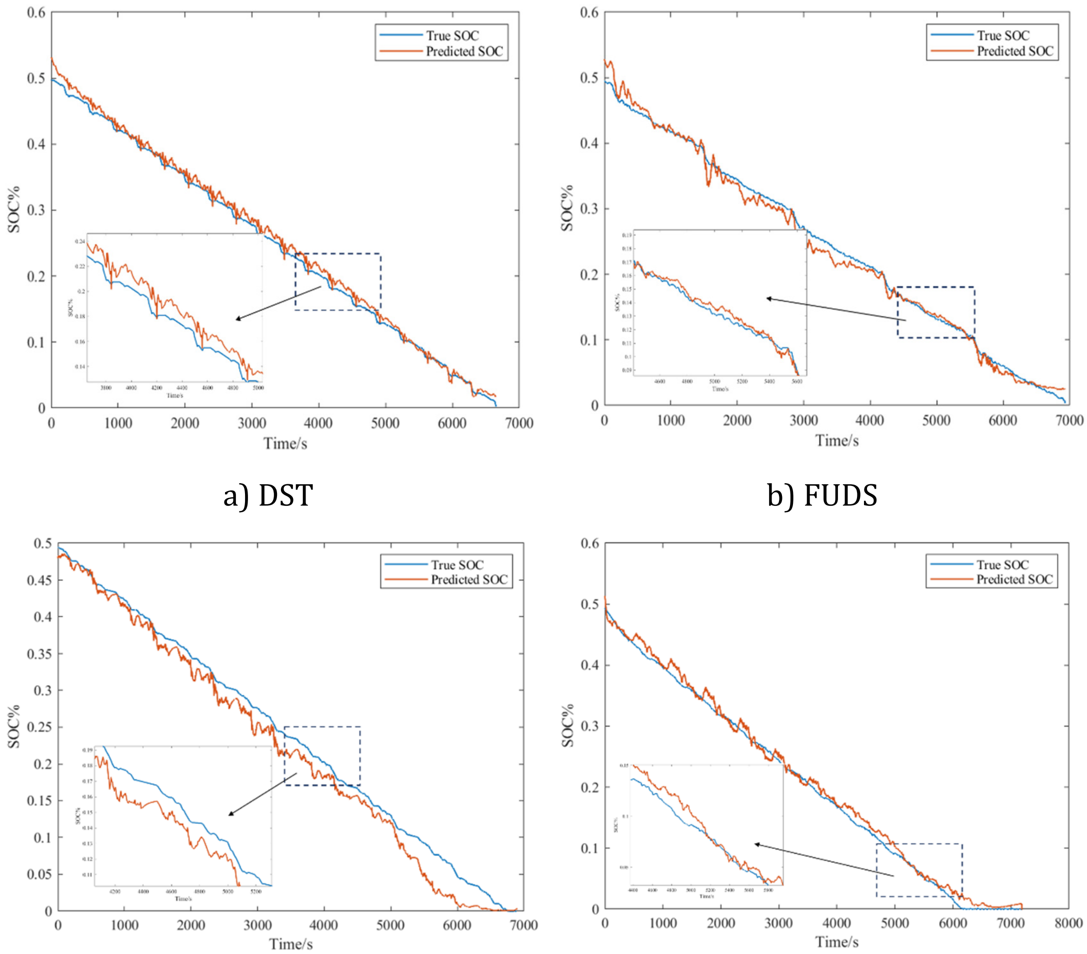
| Temperature | RMSE% | MAE% | $R^2$ |
|---|---|---|---|
| 0°C | 2.2921 | 1.9456 | 0.9670 |
| 25°C | 0.9371 | 0.7913 | 0.9957 |
| 45°C | 3.0216 | 2.2610 | 0.9586 |

ISSN: 2688-8645

Considering the experimental results obtained under different temperature conditions, the lowest RMSE and MAE and the highest $R^2$ are achieved at 25 °C. Therefore, all subsequent comparative experiments involving different dynamic operating conditions and model architectures are conducted at 25 °C, thereby minimizing the interference of temperature effects on model performance comparisons.

## 4.2. Analysis of Results under Different Dynamic Operating Conditions

To further evaluate the SOC estimation capability of the proposed model under different dynamic load conditions, experiments are conducted using four representative dynamic operating profiles, namely DST, FUDS, BJDST and US06, at an ambient temperature of 25 °C. The corresponding estimation results under the different operating conditions are shown in Fig. 5.




a) DST b) FUDS

c) BJDST d) US06

**Fig 5.** SOC estimation curves under different operating conditions

As different operating conditions are characterized by distinct current amplitudes, variation frequencies and dynamic load profiles, the nonlinear relationship between the terminal voltage response and SOC is correspondingly altered. Consequently, variations in estimation error are observed across different operating conditions. As shown in Table 4, relatively high $R^2$ values are achieved under all four dynamic operating conditions, with all values exceeding 0.97, indicating that the proposed model can effectively learn the temporal evolution of battery SOC under different dynamic load conditions. Among the four operating conditions, the best performance is achieved under the DST profile, whereas relatively larger estimation errors are observed under the BJDST profile.

**Table 4.** Estimation errors under different working conditions

| Operating conditions | RMSE% | MAE% | $R^2$ |
|---|---|---|---|
| DST | 0.9371 | 0.7913 | 0.9957 |
| FUDS | 1.2178 | 0.9280 | 0.9926 |
| BJDST | 2.1238 | 1.7682 | 0.9788 |
| US06 | 1.0559 | 0.9339 | 0.9945 |

The proposed model maintains a favourable SOC estimation capability across all four representative dynamic operating conditions, with particularly high fitting accuracy achieved under the DST profile. These results indicate that the model can effectively extract temporal features from voltage and current sequences and accurately characterize SOC variations under different dynamic load conditions.

## 4.3. Analysis of Results for Different Models

To further validate the effectiveness of the temporal feature extraction mechanisms incorporated into the CNN–Transformer model, comparative experiments were conducted using the DST profile at 25 °C. SOC estimation was performed using CNN, BiLSTM, Transformer and CNN–Transformer models, respectively. The experimental results obtained with the different models are shown in Fig. 6.

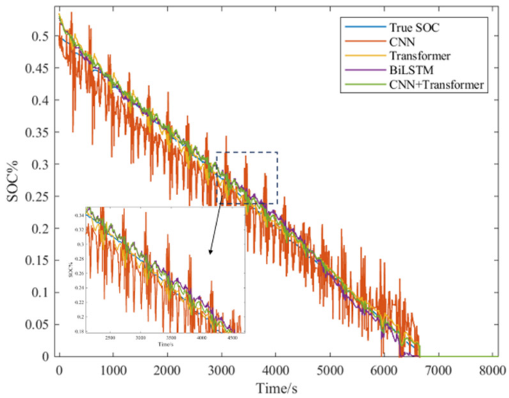



**Fig 6.** SOC estimation curves of different models

As shown in Table 5, all models achieve effective SOC estimation; however, the CNN–Transformer model achieves the best performance across all three evaluation metrics.

**Table 5.** Estimation errors of different models

| Model | RMSE% | MAE% | $R^2$ |
|---|---|---|---|
| CNN | 1.9758 | 1.6077 | 0.9924 |
| Transformer | 1.0189 | 0.8436 | 0.9948 |
| BiLSTM | 1.1583 | 0.9660 | 0.9934 |
| CNN-Transformer | 0.9371 | 0.7913 | 0.9957 |

Overall, the proposed CNN–Transformer model demonstrates an enhanced capacity to jointly exploit local feature extraction and long-range temporal dependency modelling, resulting in superior SOC estimation performance compared with the individual CNN, BiLSTM and Transformer models. The observed performance improvements are primarily attributed to the capability of CNNs to extract local variations in voltage and current, together with the ability of the Transformer to model long-range temporal dependencies. By integrating these complementary mechanisms, the temporal information embedded in the dynamic battery operating process can be more comprehensively exploited, thereby improving SOC estimation accuracy and the model's capability to characterize complex dynamic variations.

## 5. Conclusion

A SOC estimation model based on a multi-scale CNN integrated with a Transformer is proposed to address the challenges associated with the indirect measurement of lithium-ion battery SOC and the degradation of estimation accuracy under complex dynamic operating conditions. SOC reference values are obtained using the ampere-hour integration method, while normalization and a sliding-window strategy are employed to construct the time-series dataset, enabling SOC estimation from continuous voltage and current measurements. The experimental results demonstrate that the proposed model can effectively extract local dynamic features and long-range temporal dependencies from battery operating data, thereby achieving favourable SOC estimation performance.

Under different temperature conditions, the best performance is achieved at 25 °C, with RMSE, MAE and $R^2$ values of 0.9371%, 0.7913% and 0.9957%, respectively. At 0 °C and 45 °C, increased estimation errors are observed, indicating that temperature variations exert a certain influence on SOC estimation accuracy. Under different dynamic operating conditions, $R^2$ values of 0.9957, 0.9926, 0.9788 and 0.9945 are achieved under the DST, FUDS, BJDST and US06 profiles, respectively, demonstrating the favourable adaptability of the model to different dynamic loads. In addition, compared with the CNN, BiLSTM and Transformer models, the CNN–Transformer model achieves the best performance across all three evaluation metrics, with the RMSE being reduced by up to 52.57%. The effectiveness of integrating multi-scale CNNs with Transformers for SOC estimation is thereby demonstrated.

In summary, the proposed CNN–Transformer model enables local and global temporal information embedded in voltage and current sequences to be effectively exploited. High SOC estimation accuracy is maintained across different temperature and dynamic operating conditions, providing an effective technical reference for battery state monitoring and online SOC estimation in battery management systems.

## References

[1] Vidal, C., Malysz, P., & Naguib, M., et al. (2022). Estimating battery state of charge using recurrent and non-recurrent neural networks. Journal of Energy Storage, 47, 103660. https://doi.org/ 10. 1016/ j.est.2021.103660.

[2] Wang, Y., Zhang, X., Li, K., et al. (2023). Perspectives and challenges for future lithium-ion battery control and management. Elsevier. https://doi.org/10.1016/j.etran.2023.100260.

[3] He, H., Zhang, X., Xiong, R., et al. (2012). Online model-based estimation of state-of-charge and open-circuit voltage of lithium-ion batteries in electric vehicles. Energy, 39(1), 310-318. https:// doi. org/10.1016/j.energy.2012.01.009.

[4] Hossain, M., Haque, M. E., & Arif, M. T. (2022). Kalman filtering techniques for the online model parameters and state of charge estimation of the Li-ion batteries: A comparative analysis. Journal of Energy Storage, 51, 104174. https://doi.org/10.1016/j.est.2022.104174.

[5] Lan, Y., & Zhou, W. (2020). An improved method for estimating the charged state of power lithium batteries. International Core Journal of Engineering, 6(11), 311-320. https://doi.org/ 10.6919/ ICJE. 202011_6(11).0040.

[6] Streck, L., Roth, T., Bosch, H., et al. (2024). Self-discharge and calendar aging behavior of Li-ion and Na-ion cells. Journal of The Electrochemical Society, 171(8), 080531. https://doi. org/ 10. 1149/ 1945-7111/ad6cfd.

[7] Bao, H., & Yu, Y. (2013). Battery SOC estimation error correction based on ampere-hour integration method. Computer Simulation, 30(11), 148-151+159.

ISSN: 2688-8645

[8] Sundaresan, S., Devabattini, B. C., & Kumar, P., et al. (2022). Tabular open circuit voltage modelling of li-ion batteries for robust soc estimation. Energies, 15(23), 9142. https://doi.org/ 10. 3390/ en 15239142.

[9] Sundaresan, S., Devabattini, B. C., & Kumar, P., et al. (2022). Tabular open circuit voltage modelling of li-ion batteries for robust soc estimation. Energies, 15(23), 9142. https://doi.org/ 10. 3390/ en 15239142.

[10] Monirul, I. M., Qiu, L., & Ruby, R. (2024). Accurate SOC estimation of ternary lithium-ion batteries by HPPC test-based extended Kalman filter. Journal of Energy Storage, 92, 112304. https://doi. org/ 10. 1016/j.est.2024.112304.

[11] Naseri, F., Schaltz, E., Stroe, D. I., et al. (2022). An enhanced equivalent circuit model with real-time parameter identification for battery state-of-charge estimation. IEEE Transactions on Industrial Electronics, 69(4), 3743-3751. https://doi.org/10.1109/TIE.2021.3071679.

[12] Wang, Y., Cheng, Y., Xiong, Y., et al. (2022). Estimation of battery open-circuit voltage and state of charge based on dynamic matrix control - extended Kalman filter algorithm. Journal of Energy Storage, 52, 104860. https://doi.org/10.1016/j.est.2022.104860.

[13] Zhang, J., Xiao, B., Niu, G., et al. (2024). Joint estimation of state-of-charge and state-of-power for hybrid supercapacitors using fractional-order adaptive unscented Kalman filter. Energy, 294, 130942. https://doi.org/10.1016/j.energy.2024.130942.

[14] Shen, Y. Q. (2010). Adaptive online state-of-charge determination based on neuro-controller and neural network. Energy Conversion and Management, 51(5), 1093-1098. https://doi.org/ 10. 1016 / j.enconman.2009.12.015.

[15] Shi, Q. S., Zhang, C. H., & Cui, N. X. (2008). Estimation of battery state-of-charge using ν-support vector regression algorithm. International Journal of Automotive Technology, 9(6), 759-764. https: // doi.org/10.1007/s12239-008-0090-x.

[16] Sesidhar, D. V. S. R., Badachi, C., & Green, R. C. Ⅱ. (2023). A review on data-driven SOC estimation with Li-Ion batteries: Implementation methods & future aspirations. Journal of Energy Storage, 72, 108420. https://doi.org/10.1016/j.est.2023.108420.

[17] Zhang, Y. J., Wu, H. W., & Ye, C. J. (2021). Lithium-ion battery SOC estimation based on AUKF-BP neural network. Energy Storage Science and Technology, 10(1), 237-241. https://doi.org/ 10. 19799/ j.cnki.2095-4239.2020.0285.

[18] Mao, X., Song, S., & Ding, F. (2022). Optimal BP neural network algorithm for state of charge estimation of lithium-ion battery using PSO with Levy flight. Journal of Energy Storage, 49, 104139. https: // doi.org/10.1016/j.est.2022.104139.

[19] Fan, X., Zhang, W., Zhang, C., et al. (2022). SOC estimation of Li-ion battery using convolutional neural network with U-Net architecture. Energy, 256, 124612. https://doi.org/ 10.1016/ j. energy. 2022. 124612.

[20] Bian, C., Yang, S., Liu, J., et al. (2022). Robust state-of-charge estimation of Li-ion batteries based on multichannel convolutional and bidirectional recurrent neural networks. Applied Soft Computing, 116, 108401. https://doi.org/10.1016/j.asoc.2021.108401.

[21] Chen, J., Zhang, Y., Wu, J., et al. (2023). SOC estimation for lithium-ion battery using the LSTM-RNN with extended input and constrained output. Energy, 262, 125375. https://doi.org/ 10.1016/ j. energy. 2022.125375.

[22] Shen, H., Zhou, X., Wang, Z., et al. (2022). State of charge estimation for lithium-ion battery using Transformer with immersion and invariance adaptive observer. Journal of Energy Storage, 45, 103768. https://doi.org/10.1016/j.est.2021.103768.

[23] University of Maryland Center for Advanced Life Cycle Engineering. (n.d.). Battery data. https:// calce. umd.edu/battery-data.

[24] Zhu, B., Liu, X. Q., Wang, Y. H., et al. (Advance online publication). SOC prediction for lithium battery based on the model of EMD-LSTM-Attention. Journal of Power Supply. https://link.cnki. net/ urlid/ 12. 1420.TM.20240419.1720.008.