

## Article

# A Temporal Fusion Memory Network-Based Method for State-of-Health Estimation of Lithium-Ion Batteries

Kang Chen \* , Dandan Wang and Wenwen Guo

School of Information Engineering, Zhengzhou College of Finance and Economics, Zhengzhou 450054, China; wangdandan@zzife.edu.cn (D.W.); guowenwen@zzife.edu.cn (W.G.)

\* Correspondence: chenkang@zzife.edu.cn

**Abstract:** As energy storage technologies and electric vehicles evolve quickly, it becomes increasingly difficult to precisely gauge the condition (SOH) of lithium-ion batteries (LiBs) during rapid charging scenarios. This paper introduces a novel Time-Fused Memory Network (TFMN) for SOH estimation, integrating advanced feature extraction and learning techniques. Both directly measured and computationally derived features are extracted from the charge/discharge curves to simulate real-world fast-charging conditions. This comprehensive process captures the complex dynamics of battery behavior effectively. The TFMN method utilizes one-dimensional convolutional neural networks (1DCNNs) to capture local features, refined further by a channel self-attention module (CSAM) for robust SOH prediction. Long short-term memory (LSTM) modules process these features to capture long-term dependencies essential for understanding evolving battery health patterns. A multi-head attention module enhances the model by learning varied feature representations, significantly improving SOH estimation accuracy. Validated on a self-constructed dataset and the public Toyota dataset, the model demonstrates superior accuracy and robustness, improving performance by 30–50% compared to other models. This approach not only refines SOH estimation under fast-charging conditions but also offers new insights for effective battery management and maintenance, advancing battery health monitoring technologies.



**Citation:** Chen, K.; Wang, D.; Guo, W. A Temporal Fusion Memory Network-Based Method for State-of-Health Estimation of Lithium-Ion Batteries. *Batteries* **2024**, *10*, 286. <https://doi.org/10.3390/batteries10080286>

Academic Editor: Carlos Ziebert

Received: 28 June 2024

Revised: 4 August 2024

Accepted: 8 August 2024

Published: 10 August 2024



**Copyright:** © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

**Keywords:** lithium-ion battery; state of health; fast charge; channel self-attention module; long short-term memory

## 1. Introduction

To respond positively to the national vision of ‘peak carbon’ and ‘carbon neutrality’, the development of power energy storage systems is crucial. Power lithium-ion batteries (LiBs) have been the focus of close attention in electric power storage systems and electric vehicles (EVs) due to their high output power, high contained energy density, and long service life. However, due to the different fabrication processes of LiBs, noisy, inaccurate, or missing data are present in the data collection process, making it difficult to extract favorable features in each battery charge/discharge profile, and the steps such as feature selection and dimensionality reduction also require complex computational resources. Moreover, currently known methods are difficult to capture short-term patterns and long-term dependency issues in time-series data as well as different layers of parallel processing power and computational speed to such an extent that higher accuracy and robustness in estimating battery state of health (SOH) cannot be achieved [1,2]. Therefore, it is important to accurately and promptly detect and assess the SOH of LiBs. Traditional physical model-based methods are usually based on the equivalent circuit model of the battery, and state estimation is performed by measuring parameters such as voltage, current, and temperature of the battery, combined with electrochemical characteristics [3,4]. These methods have achieved some results in the field of LiB SOH estimation, but there are some problems and limitations. Firstly, complex battery systems and variable operating conditions lead

to challenging physical model building and parameter estimation. Secondly, physical models often require more a priori knowledge and experimental data support and are less adaptable to new battery materials and structures. In addition, physical models cannot fully exploit the potential patterns and features in battery data. Furthermore, understanding the impacts of abuse testing on battery performance, as highlighted by Gotz et al. [5], underscores the necessity of developing robust and reliable SOH estimation methods.

To assess the aging levels of LiBs, SOH is suggested as a critical metric for evaluating battery capacity [6,7]. In this study, SOH is interpreted through capacity metrics, where a fall in SOH below 80% indicates that the lithium battery is beyond its safe operational life and should be replaced [8,9]. Lately, the advancement of deep learning has opened up fresh avenues for SOH assessments of LiBs. Utilizing neural networks, deep learning is an approach in machine learning [10] capable of independently recognizing features and patterns by creating layered neural network models. This method exceeds conventional approaches by offering superior representation and broadening capabilities, enabling the identification of nonlinear relationships and intricate structures within data. Consequently, leveraging deep learning for SOH assessments in LiBs might enhance the precision and durability of these evaluations. SOH evaluation techniques have been outlined and reviewed by Xiong et al. [11] in 2023 and previously in 2015 by Lin et al. [12], who sorted these techniques into three types: (1) direct measurement methods; (2) model-based methods; (3) data-driven methods.

Methods for direct SOH measurement mainly involve measuring terminal voltages, currents, and impedances. For example, a common approach is to use the Coulomb counting method [13] or the electrochemical impedance spectroscopy (EIS) method [14]. Although the direct measurement method is simple in principle and can directly characterize the capacity and internal resistance of LiBs, this method requires expensive and complex equipment and experimental environments, which are less accessible for practical applications [15]. In contrast, model-based approaches focus on SOH estimation by establishing a mapping between measured variables and health states [16]. For example, equivalent circuit models [17] or electrochemical models [18] can be used. However, the accuracy of model-based approaches depends on the model chosen, and different models are selected for different battery types and must be iteratively corrected. Moreover, it is difficult to obtain accurate model parameters in reality [19].

With the advancement of the information era, data-driven methods are increasingly prevalent in both industry and academia. These data-driven methods generally depend on operational variables such as current, voltage, and temperature from LiBs, while often neglecting the aging mechanisms and chemical transformations within the batteries. Li et al. [20] implemented a two-layer integrated limit learning machine algorithm for the joint estimation of SOH and SOC in LiBs, which improved the estimation's accuracy and robustness. Wu et al. [21] refined a particle filtering algorithm for the joint estimation of SOH and SOC, achieving high accuracy through the use of internal resistance for feature extraction. Hsing et al. [22] introduced a Bayesian tracking method for SOH estimation using historical data, coupled with an enhanced optimized decision tree (DT), which ensures highly robust and accurate estimation results. Kong et al. [23] developed a hybrid model that merges long short-term memory with a deep convolutional neural network, integrating a unique health feature that connects the internal dynamics of LiBs with external aging features and has a strong correlation with capacity. The application of various models in machine learning and neural networks for the estimation of SOH in LiBs is on the rise [24–27]. Neural networks have become indispensable tools in machine learning, especially deep learning, where they solve complex nonlinear issues by interpreting patterns and features in data [27,28]. Prominent examples include recurrent neural networks (RNNs) [29], long short-term memory (LSTM) networks [30], and Transformers [31]. These models excel at utilizing historical data to uncover hidden patterns and significantly enhance the estimation accuracy of the SOH of LiBs.

Despite some advances in battery SOH estimation through 1D CNN, GRU, LSTM, and various models, opportunities still exist to enhance precision and robustness. To advance SOH estimation further, a novel strategy involving the introduction of a Temporal Fusion Memory Network (TFMN) is proposed. Initially, the framework incorporates a 1D CNN as an initial layer to derive deeper local insights from the input data. This 1D CNN is adept at identifying immediate patterns in time-series data, thereby enriching the data context for subsequent, more complex model layers. Next, data progress to a channel self-attention module (CSAM), which avoids the zero-denominator instability issue and enhances focus on pertinent attributes. Subsequently, the model maximizes LSTM capabilities for time-series management, simultaneously integrating the Transformer's self-attention features to elucidate series connections more effectively. This integration allows the LSTM to manage extended temporal dependencies, while the attention mechanism's parallel processing faculties expedite both the training and analysis phases. Ultimately, a comprehensive prediction is executed through two interconnected layers.

In summary, the contributions of this paper are fourfold:

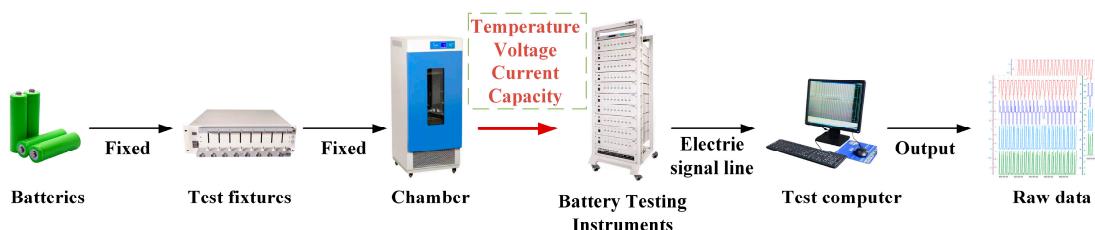
1. In this paper, a multi-feature hybrid feature extraction is proposed, i.e., time difference at equal voltage intervals, voltage difference at equal time intervals, cumulative integral of voltage change, and peak IC curve slope. These feature extraction methods effectively focus on factors affecting various aspects of battery aging. By simplifying the data processing process, the trend of battery health state over time can be effectively captured, providing a solid foundation for subsequent model construction.
2. In this paper, a new type of attention module, i.e., channel self-attention module (CSAM, channel self-attention module), is proposed. Tanh is chosen as the activation function. Compared with the self-attention module, this module does not have the instability problem of a zero denominator, allowing the model to pay more attention to the correct features while inhibiting the extraction of irrelevant features. Additionally, it does not have the problem of gradient disappearance or gradient explosion, thus improving the performance of the network.
3. In this paper, a novel framework for SOH estimation leveraging diverse features is introduced. Components such as 1D CNN, CSAM, LSTM, and multi-head self-attention are integrated to create a cohesive and efficient model architecture. The 1D CNN serves as the foundational layer, adeptly discerning the local characteristics of the battery data and converting these into more complex inputs, thereby establishing a robust groundwork for the model's efficacy. Additionally, the integration of CSAM with LSTM and the incorporation of the Transformer's self-attention feature into the framework facilitate a deeper understanding of the temporal connections within battery data, effectively addressing challenges of long-term dependencies. Furthermore, the LSTM's synergy with the attention mechanism permits an exhaustive analysis of varied information layers within sequential data, thereby enhancing the model's overall estimation capabilities. This multi-faceted and collaborative approach not only elevates the precision of estimations but also enhances the model's adaptability to diverse data variations.
4. The robustness of the proposed model was experimentally confirmed even with noise present. Noise was incorporated into the model inputs to mimic potential data fluctuations in an actual setting. Impressively, the model still excelled under noisy conditions, further substantiating its sturdiness and dependability. This indicates that the model not only elevates precision but also sustains top-notch performance in volatile settings, offering a steadfast approach for the battery SOH estimation challenge.

The remainder of this paper is organized as follows: In Section 2, the LiB experimental dataset and the four feature extraction methods employed are detailed. In Section 3, the proposed innovative model TFMN, which fuses LSTM and attention modules and combines the properties of 1D CNN to better solve the battery SOH estimation problem, is illustrated. Then, in Section 4, the experimental results are presented, along with an in-depth discussion

and analysis of the effectiveness of the proposed TFMN. Finally, conclusions that summarize the main contributions and findings of this paper are provided in Section 5.

## 2. Dataset and Feature Extraction

Datasets play a crucial role in machine learning, and having a sufficient dataset of LiBs helps to build robust general evaluation models. We established an experimental platform for battery cycle life testing with the Neware CT-4000 test equipment (Neware, Shenzhen, China), with a voltage and current error of 0.05%; two different charging and discharging conditions were designed using the platform, and the experiments were uniformly conducted in a temperature chamber (MGDW-408-20H, with a deviation of  $\leq \pm 1$  °C) at a constant temperature of 25 °C, as shown in Figure 1, and were performed on dataset A and dataset B. To verify the generality of the model, the open-source dataset C, the open-source LiB dataset of Severson et al. from the Toyota Research Institute (in collaboration with Stanford University and the University of Maryland), was used again [32,33]. Due to the limitations of capacity measurements in practical applications, in this section, four features are extracted for SOH estimation based on battery cycling curves reflecting the aging mechanism of the battery. The focus of this paper will concentrate on the dataset used and the extraction method of the health features. The specifications of the three datasets are shown in Table 1.



**Figure 1.** Schematic diagram of the experimental setup for battery cycle life testing.

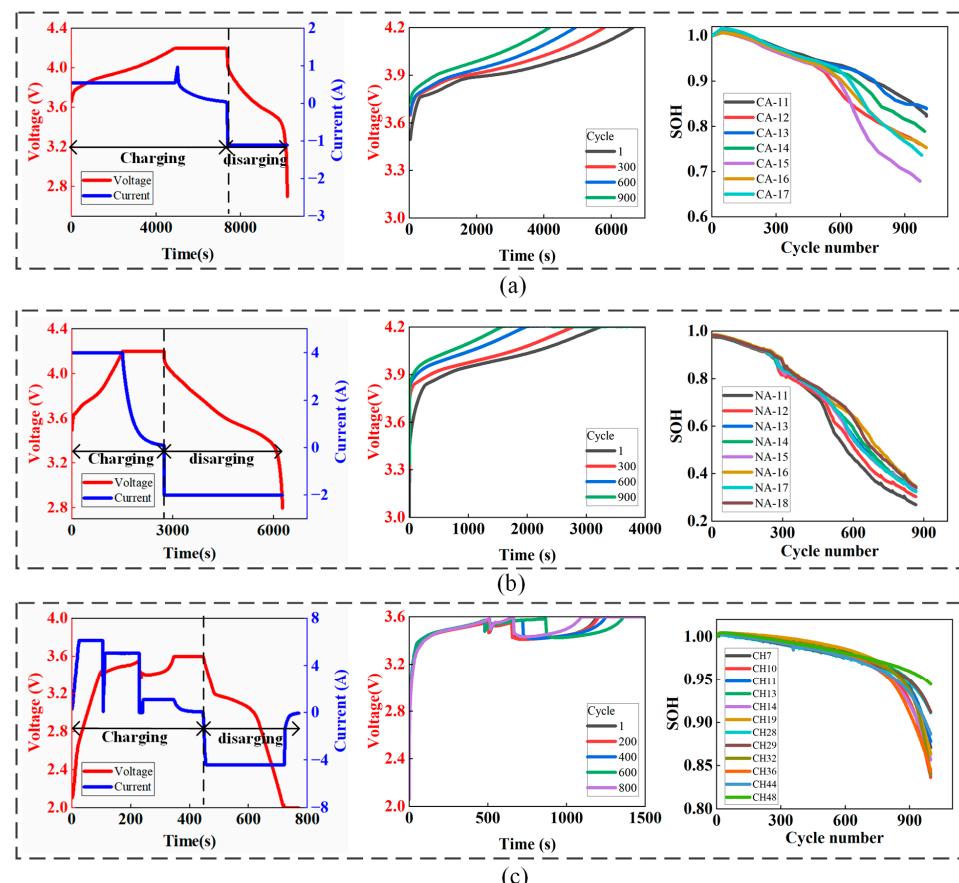
**Table 1.** Three dataset specification models.

Category	Dataset A	Dataset B	Dataset C
Battery type	CS2	ICR 18650 P	APR 18650 M1A
Cathode material	LiCoO <sub>2</sub>	LiNiCoMnO <sub>2</sub>	LiFePO <sub>4</sub>
Anode material	GraNAite	GraNAite	GraNAite
Quantities	7	8	12
Nominal capacity	1.1 Ah	2 Ah	1.1 Ah
Nominal voltage	3.6 V	3.6 V	3.3 V
Ambient temperature	25 °C	25 °C	30 °C

### 2.1. Dataset

1. **Dataset A:** The first dataset used in this study was measured on an experimental platform. Seven CS2 batteries from the same batch, labeled CA-11, CA-12, CA-13, CA-14, CA-15, CA-16, and CA-17, with a nominal capacity and nominal voltage of 2 Ah and 3.6 V, respectively, were selected. The experiments were carried out uniformly in a thermostat box (MGDW-408-20H, with a deviation of  $\leq \pm 1$  °C) at a constant temperature of 25 °C. The experimental setup included five battery pretreatments. The experimental setup consisted of 5 cycles of battery pretreatment, 50 cycles of battery aging, and 20 cycles of battery capacity calibration. As can be seen from each battery aging cycle curve, the battery was charged to 4.2 V at a constant current of 0.25 C and then continued to be charged at a constant voltage of 4.2 V until the current dropped to 0.05 A. After complete charging, the battery was depleted at a rate of 1 C until the terminal voltage dropped below the cut-off value of 2.7 V. The battery was fully charged and then depleted at a rate of 1 C until the terminal voltage dropped below the cut-off value of 2.7 V.

2. **Dataset B:** The experimental platform for the second dataset used in this study is the same as that of dataset A. The second dataset used in this study is the same as dataset A. Eight ICR18650P batteries from the same batch, labeled NA-11, NA-12, NA-13, NA-14, NA-15, NA-16, NA-17, and NA-18, with a nominal capacity and voltage of 2 Ah and 3.6 V, respectively, were selected and operated in the same steps. The batteries were first charged at a constant current of 2 C until the battery voltage reached 4.2 V. After a short rest, the batteries were fully charged using a constant voltage (CV) of 4.2 V with a cut-off current of 0.05 C and a cut-off voltage of 4.2 V. After resting for one hour, the batteries were discharged using 1 C until the voltage dropped to 2.7 V. The battery voltage was then reduced to 2.7 V.
3. **Dataset C:** The third dataset used in this paper uses the open-source lithium-ion battery dataset from the Toyota Research Institute (in collaboration with Stanford and Maryland) and Severson et al. [32]. The dataset obtained from the Toyota Research Institute contains information on various aspects of the aging process, such as voltage, current, temperature, and charging time, to give a side-by-side view of the battery aging process. Figure 2c shows the current and voltage variation curves, voltage curves, and capacity decay curves during a complete charge/discharge process. The work step setup firstly charges the battery with 5.5–6.1 C constant current charging until the voltage reaches 3.6 V. In the second step, the battery is charged with 2–4 C constant current until the voltage reaches 3.6 V. After that, the battery is charged with 1 C constant current and constant voltage. The upper and lower potentials are 3.6 volts and 2.0 volts, respectively, consistent with the manufacturer's specifications. After some cycles, the batteries are charged at a constant voltage. The upper limit potential may be reached during rapid charging. All batteries are discharged at 4 C.



**Figure 2.** Charge–discharge curves, charge–voltage curves, and SOH decay curves. (a) Dataset A; (b) dataset B; (c) Dataset C.

## 2.2. Feature Extraction

Battery cycle life can be expressed based on external characteristics such as voltage, internal resistance, and capacity, which can reflect the deterioration and health of a battery. To accurately estimate the SOH of a battery, it is crucial to extract characteristic variables that reliably reveal the aging trend of the battery.

Temporal features are the most direct indicators of changes in battery health status. The deep learning model learns complex nonlinear relationships from a large amount of data by learning time features associated with battery life trends. However, due to the short decay cycle of fast-charging batteries, the limited number of cycles and the amount of data limit the learning depth of the model and are among the main reasons for the gradient disappearance. In our battery dataset A, the SOH decayed to 80% in 6 months, and the fast-charging dataset B and dataset C took 1 month and 3 months, respectively. Temporal features are important in battery SOH estimation. They can provide information about battery degradation, historical behavior, dynamic changes, and failure warnings, helping the model to estimate battery health more accurately. Therefore, battery features such as the time difference of equal voltage intervals, the voltage difference of equal time intervals, the cumulative integral of the voltage change, and the peak slope of the IC curve were further extracted based on the time dimension through feature data analysis.

Fast-charging systems may use different charging strategies, including different combinations of voltage and current, depending on the chemistry of the battery, to optimize charging rate and battery life. However, fast charging may cause the battery to overheat, and to avoid safety issues and reduced life, the system monitors the battery temperature and adjusts the charging rate. This may lead to different stages in the charging process, and the fast-charging battery data are divided into multiple segments. Extracting features by time dimension leads to risks such as difficulty in model convergence, reduced generalization ability, and overfitting due to easy cross-confusion in extracting data from multiple segments together within the same cycle. Therefore, we utilized multi-segment data for feature extraction by automatically calculating the initial and end values of each segment of the cycle for battery data. Figure 3 shows the extraction process of four features, where Figure 3a,b are direct measurement features, which represent the time difference of equal voltage intervals and the voltage difference of equal time intervals, respectively. Figure 3c,d shows the secondary measurement features, which represent the cumulative integral and differential capacity curve slope peaks of the voltage change, respectively.

### 2.2.1. Time Difference between Equal Voltage Intervals

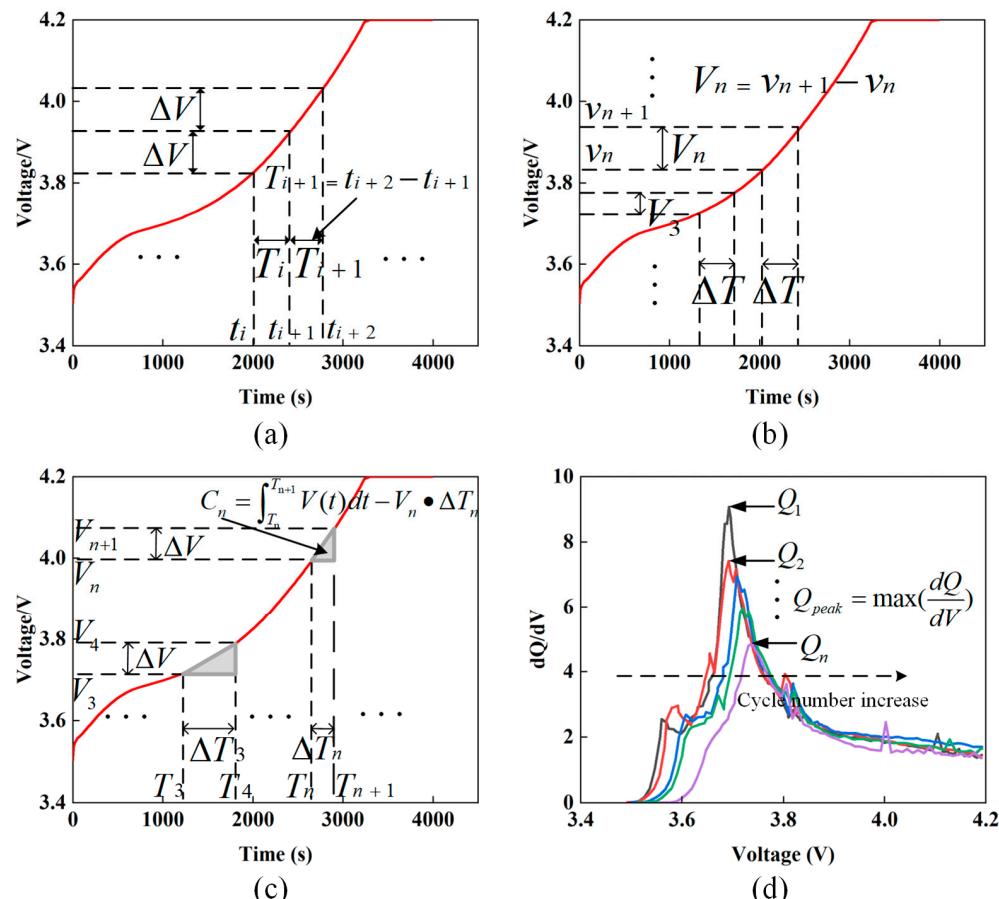
Each battery underwent charge/discharge cycles during an aging process experiment, and the related experimental data for these cycles were collected. A specific charging voltage profile from these data was selected, and the voltage spectrum during charging was segmented into equal parts. The duration allocated to each segment was then identified as a feature. As depicted in Figure 3a, a voltage spectrum from 3.6 to 4.2 V is established as the comprehensive interval of charging for both dataset A and dataset B, indicated by  $V_{all}$ .  $\Delta V$  represents the voltage span for each attribute, termed as the voltage sampling period or voltage resolution. The duration necessary for a consistent voltage sampling interval is selected as the attribute. Therefore, the total count of voltage sampling intervals, denoted as  $n$ , is referred to as the count of attributes, and its mathematical formulation is presented in Equation (1):

$$n = \left\lfloor \frac{V_{all}}{\Delta V} \right\rfloor, \quad (1)$$

The  $\lfloor \cdot \rfloor$  operation truncates the value to an integer. The entire voltage interval  $V_{all}$  is segmented into equal  $n$  subintervals, some of which may be unoccupied. Consequently, the characteristic  $T_i$  is illustrated in Equation (2):

$$T_i = \begin{cases} 0, & \text{if the subinterval } i \text{ is null;} \\ t_{i\_end} - t_{i\_start}, & \text{otherwise.} \end{cases}, \quad (2)$$

where  $t_{i\_start}$  and  $t_{i\_end}$  represent the beginning and conclusion of the  $i$  characterization subinterval, respectively. Consequently, each charging cycle of the battery is defined by  $n$  characteristic variables.



**Figure 3.** Feature extraction curves. (a) Time difference at equal voltage intervals; (b) voltage difference at equal time intervals; (c) cumulative integral of voltage change; (d) peak slope of differential capacity curve.

The same method is used for dataset C. The difference is that the voltage range of dataset C is 2.0–3.6 V.

## 2.2.2. Voltage Difference at Equal Time Intervals

Battery data, as time-series data, especially in the estimation of SOH decay, rely heavily on voltage features. The voltage difference in equal time intervals was calculated. From Figure 3a, with an increase in the number of cycles, the duration of charging for each cycle shortens, and the variation in voltage across different cycles within the identical charging time interval progressively grows. Therefore, it can be deduced that the voltage difference within equal time intervals contains features of battery degradation. As shown in Figure 3b, calculating the voltage difference allows these non-uniform changes to be captured more precisely.

The maximum and minimum values of time and the number of samples  $m$  are calculated for each cycle to determine the start and end voltages. The next feature extraction operation will be carried out within the voltage blocks divided based on the time intervals, and it should be noted that the calculation of the number of samples  $m$  has a direct impact

on the next feature extraction. The number of time sampling intervals  $m$  and the equation of the voltage change in equal time intervals are as follows:

$$m = \left\lfloor \frac{T_{all}}{\Delta T} \right\rfloor, \quad (3)$$

$$\Delta V_i = V_{i\_start} - V_{i\_end}, \quad (4)$$

where  $V_{i\_start}$  and  $V_{i\_end}$  are the start and end voltages of the  $i$  feature subinterval, respectively, and  $\Delta V_i$  is the voltage difference, a feature related to the voltage dynamics during charging.

### 2.2.3. Cumulative Integral of Voltage Change

It is easy to see in Figure 3a that the voltage curve is shifted to the upper left in the decay cycle of the fast-charging battery, and the slope of the voltage curve changes with it, with a slower voltage rise at the beginning of the charging period and a slower voltage rise when the charging is near completion. However, the simple instantaneous voltage value may not be able to fully reflect the state of the battery, while the cumulative integration can accumulate the voltage changes over a period, and the voltage changes are observed by means of cumulative integration, and the small voltage changes will be gradually amplified in the process of accumulation, thus more accurately characterizing the battery. In Figure 3c, it can be seen that the cumulative integral  $C_n$  of the voltage profile in the  $n$ th second time interval is to the upper right of the cumulative integral  $C_i$  at the initial stage, and there is a significant difference in size.

The  $C_n$  in Figure 3c represents the cumulative integral of the change in battery voltage during the charging cycle, i.e., the energy input of the battery during the charging process, calculated as follows:

$$C_n = \int_{T_n}^{T_{n+1}} V(t) dt - V_n \cdot \Delta T_n, \quad (5)$$

where  $\Delta T_n$  denotes the  $n$ th time block of division, where  $T_n$  and  $T_{n+1}$  are determined by the time block of division.

### 2.2.4. IC Curve Peaks

The IC curve, which is the curve of current versus voltage during battery charging and discharging, contains many intuitive aging features. These features are highly correlated with the SOH of the battery [34] and thus can be used to estimate the SOH of the battery. The area of the IC curve is used as a feature that is easy to obtain directly, but due to the small amount of data for a single cycle of a fast-charging battery, calculating the area of the integral may become difficult because the integral usually requires many data points to accurately approximate the curve of the function. The IC curve is relatively stable in the pre-cycling period. However, it appears to be extremely unstable in the later data, and it is difficult to obtain effective confirmation of both the area and the peak value. Moreover, the integration operation generates the problem of error accumulation and noise amplification, and this instability may lead to the problem of gradient explosion, which reduces the SOH prediction accuracy. Gradient explosion refers to the fact that during the training process of a deep neural network, the gradients of the network parameters become very large, leading to excessive weight updates, which results in a rapid expansion of the network's weight values and the loss of an effective representation of the data.

Thankfully, it was found that the position of the peaks during the charging process gradually shifted to the right compared to the first cycle, with the difference being obvious. Therefore, the calculation of area and peak height data, which are prone to anomalies, was avoided, and the most intuitive data, the peak horizontal position, was chosen. As

shown in Figure 3d, there is a clear horizontal displacement from  $Q_1$  to  $Q_5$ . The formula for calculating the horizontal value of the peak voltage slope is as follows:

$$Q_{peak} = \max\left(\frac{dQ}{dV}\right), \quad (6)$$

where  $Q_{peak}$  is used to measure the horizontal value of the maximum slope of the change in battery voltage, the horizontal value of the maximum value of the rate of voltage rise of the battery during charging.

### 2.3. Correlation Analysis

To establish the relationship between the extracted health features and the state of health (SOH) of lithium-ion batteries, we conducted a comprehensive correlation analysis using all battery data features described in Section 2.2. This analysis employed the Pearson correlation coefficient to quantify the linear relationships between these features and the SOH, aiming to enhance the accuracy and predictive power of our SOH estimation models.

Utilizing the Pearson correlation coefficient,  $r_{xy}$ , we calculated this statistic according to the following formula:

$$r_{xy} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}} \quad (7)$$

Here,  $x_i$  and  $y_i$  represent the feature values and SOH measurements, respectively, with  $\bar{x}$  and  $\bar{y}$  being their mean values over the dataset.

Given the consistent operational conditions for individual batteries within each dataset, a representative subset comprising two batteries from each dataset was selected for the correlation analysis. Table 2 displays the Pearson correlation coefficients for selected battery units CA-13, CA-16, NA-12, NA-18, CH-14, and CH-44. Evaluations were conducted on metrics including time differences, voltage differences, cumulative voltage, and IC curve peaks. The correlation coefficients, with absolute values ranging from 0.86 to 0.94, substantiate a significant linear relationship between the health features and the state of health (SOH), as presented in Table 2.

**Table 2.** Pearson correlation coefficient between health characteristics and SOH.

Features	CA-13	CA-16	NA-12	NA-18	CH-14	CH-44
Time difference between equal voltage intervals	0.87	0.92	0.89	0.94	0.88	0.91
Voltage difference at equal time intervals	-0.86	-0.90	-0.91	-0.88	-0.93	-0.90
Cumulative integral of voltage change	0.92	0.91	0.90	0.89	0.94	0.93
IC curve peaks	-0.88	-0.89	-0.92	-0.91	-0.89	-0.87

#### 1. Time Difference Between Equal Voltage Intervals

This feature exhibits high positive correlations across all datasets (ranging from 0.87 to 0.94), indicating that the time batteries spend at specific voltage intervals is closely linked to their health status. A greater time difference typically suggests a faster discharge rate at reaching the same voltage, which may be an early indicator of declining battery health.

#### 2. Voltage Difference at Equal Time Intervals

This feature shows high negative correlations (-0.86 to -0.93), meaning that the change in voltage over equal time intervals is inversely related to battery health deterioration. An increase in voltage difference may indicate an increase in internal impedance, reflecting accelerated aging of the battery.

### 3. Cumulative Integral of Voltage Change

This feature also demonstrates very high positive correlations (0.89 to 0.94) across all datasets, indicating a direct positive correlation with battery SOH. A higher cumulative integral of voltage change typically suggests that the battery releases energy more quickly during discharge, which may be related to a reduction in battery capacity.

### 4. IC Curve Peaks

This feature exhibits high negative correlations with battery SOH ( $-0.87$  to  $-0.92$ ), indicating that variations in IC curve peaks are inversely related to battery health. A decline in IC curve peaks usually reflects a decrease in chemical reactivity during the charging and discharging processes, which is a significant indicator of battery aging.

The analysis above reveals clear linear relationships between the extracted health features and the battery's state of health (SOH). The trends in these features not only provide a quantitative basis for assessing the health status of batteries but also help in predicting future battery performance. Therefore, integrating these features is crucial for enhancing the accuracy and predictive power of battery SOH estimation models.

#### 2.4. Sliding Window Module

In battery SOH (state of health) estimation, the sliding window method for time series is frequently employed to manage the sequential data and glean insights into battery behavior and performance [35]. The function of the sliding window is outlined as follows:

**Data sampling and processing:** The sliding window module is used to sample and process the historical data of the battery. It selects data over a period by sliding a fixed-size window, which will be used for subsequent feature extraction.

**Feature extraction:** The health state of the battery is influenced not only by present conditions but also by past actions. The sliding window technique in time series allows for the utilization of historical data as a framework to extract relevant features. These characteristics can offer valuable insights into the battery's condition and activities, facilitating a finer assessment of the SoH (state of health) of the battery. By modifying the dimensions of the sliding window, it becomes feasible to manage the equilibrium between leveraging historical data and sustaining immediate performance.

**Data Alignment:** The sliding window module can help align temporal data. In the battery SOH estimation task, batteries with different operating conditions have different charge/discharge cycles and sampling rates. Through the sliding window module, the data from different batteries can be aligned to the same time window to ensure the consistency of the subsequent SOH estimation.

**Sequence modeling:** A sliding window for time series can be employed to build sequence models, including recurrent neural networks (RNNs) and long short-term memory (LSTM) networks, among others. These models are capable of accounting for the temporal aspects of the data, thereby more effectively capturing the dynamics of battery actions. Utilizing data within a sliding window as input, the model is able to discern the evolving patterns of battery state changes over time, enabling it to forecast future battery health states.

Therefore, the use of time-series sliding windows in battery SoH prediction facilitates the extraction of useful features from time-series data and the building of appropriate sequential models to improve the accuracy and robustness of predictions.

In the experiments of this paper, our data consist of a  $c$  loop of feature vector  $x$  and SoH labels  $y$ , which can be represented by Equations (8) and (9):

$$x = \{x_1, x_2, \dots, x_j, \dots, x_c\}, j \in [1, c], \quad (8)$$

$$y = \{y_1, y_2, \dots, y_j, \dots, y_c\}, j \in [1, c], \quad (9)$$

$j$  denotes one of the cycles, and  $j \in [1, c]$ . The size of the time-series sliding window is then set to  $w$ , and the time step of the estimation is set to  $p$ . The sliding window was utilized to construct the inputs for model training as in Equations (10) and (11):

$$X_j = [x_{j-w}, \dots, x_j], j \in (w, c-p], \quad (10)$$

$$Y_j = [y_{j+p-w}, \dots, y_{j+p}], j \in (w, c-p], \quad (11)$$

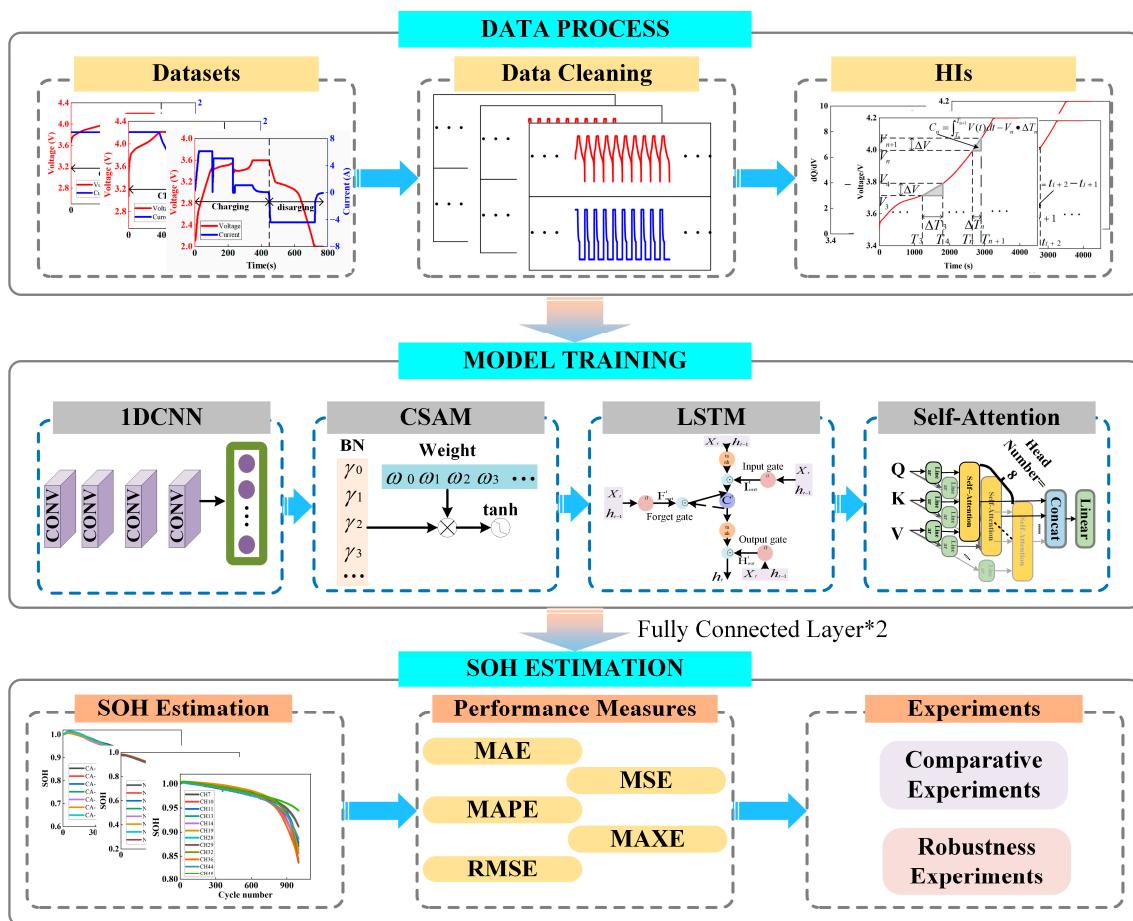
where  $X_j$  and  $Y_j$  denote the feature matrices and label vectors used to make predictions for cycle  $j$ , respectively, and  $p$  denotes the  $p$ -steps of advance prediction. Thereby, a set of sequence pairs comprising multivariate feature matrices and corresponding SOH label vectors is obtained. A set of training batches is created by combining the sequence pairs based on sequence lengths. In general, if the batch size is large, it may return a local optimum; if the batch size is too small, it may be difficult to achieve convergence. Therefore, it is important to set the appropriate batch size. In this paper, the batch size is set to 16.

### 3. Methodology

In this part, an in-depth explanation of the suggested model, Temporal Fusion Memory Network (TFMN), for estimating the SOH of LiBs is provided. Traditional models might struggle in certain scenarios, especially when managing long-term dependencies, and are often used to solve SOH estimation for datasets under standard charging conditions. There are very few models addressing SOH estimation for fast-charging datasets. To address this problem, an innovative SOH estimation model is proposed. TFMN effectively captures short-term patterns and long-term dependencies in battery data by fusing the powerful capabilities of a one-dimensional convolutional neural module, a channel self-attention mechanism, a long short-term memory module, and a multi-head self-attention module. With four feature extraction and sliding window techniques, a new efficient method for battery SOH estimation is provided.

The process of the TFMN in assessing the SOH of LiBs is depicted in Figure 4 and can be succinctly described as follows: Initially, data cleaning is applied to the raw data, followed by the retrieval of the four types of feature data. The one-dimensional convolutional neural module is utilized to thoroughly convolve the time-series data to extract the features, which are then handled using the time-series sliding window to derive the feature vectors. These vectors are subsequently inputted into the TFMN, where the model assimilates the correlation of the features with the battery SOH. Specifically, within the TFMN: (1) the feature vectors are standardized by splicing and acquiring temporal input data after filtering through the channel self-attention module (CSAM); (2) the temporal input data are refined through the long short-term memory module to generate the model data; (3) these model data are then advanced through the multi-head self-attention module to produce the resultant data; (4) ultimately, these resultant data are routed and transformed via the two-layer fully connected layer into the feature data, correlating it to the ultimate SOH prediction.

To address the problems in the prior art, a TFMN-based method for estimating the health state of LiBs is provided, which largely solves the problems of short-term patterns and long-term dependencies in the time-series data, improving the accuracy and robustness of SOH estimation.



**Figure 4.** Overall flowchart of the SOH estimation model.

### 3.1. 1DCNN Layer

In this study, the 1D convolutional neural network (1DCNN) not only proves essential for feature extraction but also significantly contributes to feature transformation. Our 1DCNN framework manipulates the input sequences through convolutional operations, converting the raw battery data into more complex feature representations. Specifically, we configure the output dimension of the 1DCNN to be four times the size of the input dimension. This amplification in dimension showcases the 1DCNN's exceptional capacity to identify intricate patterns and temporal variations. Via convolutional operations, the 1DCNN identifies and amplifies subtle features within the input sequence, enhancing its discriminatory and expressive power. Therefore, the augmentation of the output dimension of the 1DCNN transcends mere size enhancement; it embodies a comprehensive depiction of the data's inherent structure.

This enriched feature representation delivers extensive data in subsequent model layers, offering more detailed inputs to the neural network and boosting the model's expressive power. By enlarging the output dimension of the 1DCNN to four times the input, we capture more critical features within the battery data, thus furnishing more substantial inputs to the following Transformer encoder and LSTM networks, thereby enhancing the overall predictive accuracy.

### 3.2. Channel Self-Attention Module

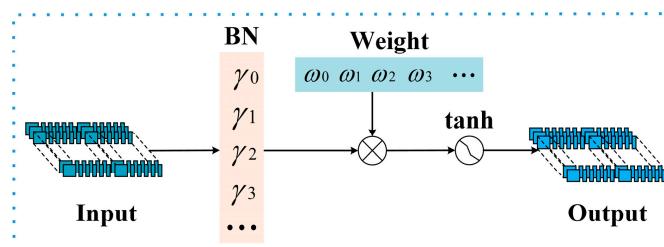
Attention mechanisms' different features or feature channels are weighted to make the model pay more attention to features that are more meaningful or relevant to the task at hand, and to dependencies between different locations or time steps in the data. It can also help the model to focus on important information, thus improving the performance

and accuracy of the model. It reduces the model's dependence on the overall input, focuses computational resources and attention on the most relevant parts, and reduces the impact on invalid or noisy information, which helps to reduce overfitting and improve model robustness.

A channel self-attention module, as shown in Figure 5, is proposed. Tanh is chosen as the activation function because it does not suffer from the instability problem of a zero denominator compared to the self-attention module. This allows the model to pay more attention to the correct features while suppressing the extraction of irrelevant features, without encountering the problem of vanishing or exploding gradients, thus improving network performance. The scale parameter in batch normalization signifies the significance of the weights. This factor represents the extent of variation for each channel and highlights the importance of the channel. The scaling factor, which is the variance in batch normalization, can react to the degree of change in the channel. If the variance is higher, the channel contains more information and is considered more significant. The result of the batch normalization operation  $C_{out}$  is as follows:

$$C_{out} = BN(C_{in}) = \gamma \frac{C_{in} - \mu_C}{\sqrt{\sigma_C^2 + \epsilon}} + \beta, \quad (12)$$

where  $\mu_C$  and  $\sigma_C$  are the mean and standard deviation of the small batch, respectively;  $\beta$  is the affine transform displacement parameter for training;  $BN$  is the batch normalization operation;  $C_{in}$  is the input to the channel self-attention module; and  $\epsilon$  is the hyperparameter that prevents the denominator of the formula from being zero.



**Figure 5.** Structure of channel self-attention module.

The weights and output operations of this module are shown in Equations (13) and (14):

$$\omega_i = \frac{\gamma_0}{\sum_{j=0} \gamma_j}, \quad (13)$$

$$M_c = \tanh(\omega_i(BN(C_{in}))), \quad (14)$$

where  $\omega_i$  is the weight,  $\gamma$  is the scaling factor of the channel, and  $\tanh$  is the activation function.

CSAM has the following advantages over self-attention modules:

1. CSAM can improve the stability and robustness of the model by normalizing the attention scores before calculating the attention weights; the self-attention module may face numerical instability when calculating the attention weights; for example, for the division problem when the denominator is close to zero, through the normalization operation, the attention scores can be kept in a more reasonable range and the numerical instability can be reduced.
2. CSAM can better control the flow of information in the sequence through the normalization operation, the tanh function in CSAM will asymptote to 1 and  $-1$  when the input tends to positive infinity or negative infinity, which ensures that the output value of the temporal transformation memory network is within a certain range, and will not suffer from the problem of vanishing or exploding gradient; when the input is in the range between  $[-1, 1]$ , the tanh function's function value changes are more sensitive than the Sigmoid function, making the model's performance more stable in

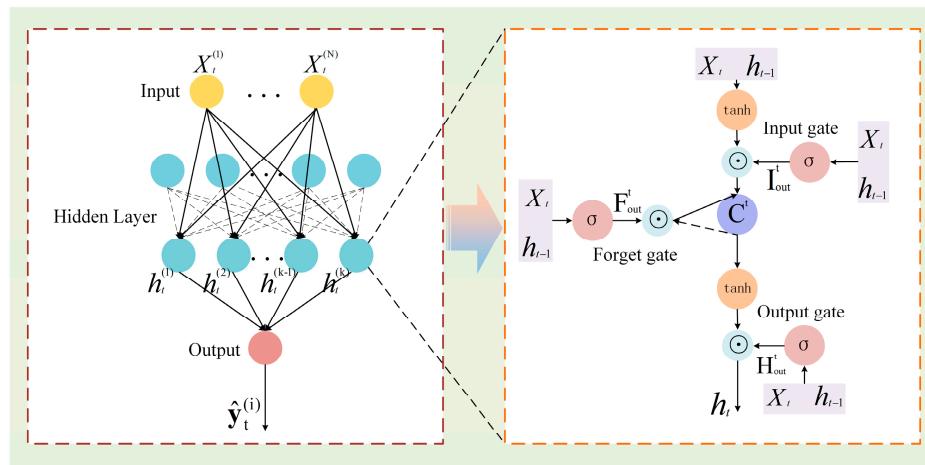
- this interval, enabling the information to be transmitted and concentrated effectively in the sequence.
3. CSAM has less influence on outliers; if there are outliers or noisy data, the self-attention module may pay too much attention to these outliers, resulting in a decrease in the model performance; the channel self-attention module can mitigate the influence of outliers by limiting the attention weights to a certain range, improving the robustness of the model.

### 3.3. Long Short-Term Memory

Long short-term memory (LSTM) [30] is a recurrent neural network framework specifically crafted for handling sequential data. It is depicted in Figure 6. Its fundamental architecture is composed of three elements: a forgetting gate, an input gate, and an output gate. A neural network layer possesses two properties, namely a weight vector  $W_i$  and bias vector  $b_i$ , and for each element  $A_i$  of the input vector  $A$ , the neural network layer conducts operations on it as detailed in Equation (15):

$$\text{Output}_i = f(W_i \cdot A_i + b_i), \quad (15)$$

where  $f(\cdot)$  represents the activation function; this study employs sigmoid and tanh functions as activation mechanisms.



**Figure 6.** LSTM neural network and LSTM units.

The key for LSTM to stand out from RNN lies in the hidden state (unit state) of the neuron in the blue circle in Figure 6; we can understand the hidden state of the neuron as the recurrent neural network's "memory" of the input data, and use  $C^t$  to indicate the neuron's "memory" after the  $t$  moment. This vector covers the recursive neural network's "summary" of all input information until moment  $t + 1$ .

These gates within the LSTM proficiently regulate the flow of feature information throughout the sequence, enabling the network to selectively retain and discard feature information across time steps, thus enhancing the capture of long-term dependencies. We detail the function of each gate individually below.

Within the LSTM, these gates efficiently manage the distribution of feature information throughout the sequence, thereby allowing the network to selectively preserve and omit feature information at different intervals, facilitating improved long-term dependency tracking. The functionality of each gate is explained separately below.

#### 1. Forget Gate:

The primary function of the forget gate is to determine whether the network should erase prior memories at the current time step. This gate computes its value by utilizing the current input  $X^t$  and the previous time step's hidden state  $h^{t-1}$ . The sigmoid function

yields an output for the forgetting gate ranging from 0 to 1, which governs the level of information retention within the memory cells, with 0 indicating total erasure and 1 ensuring full preservation. The output vector  $F_{out}^t$  is defined in Equation (16):

$$F_{out}^t = f_\sigma(W^f \cdot [h^{t-1}, X^t] + b^f), \quad (16)$$

where  $W^f$  and  $b^f$  are the weights and bias of the forgetting gate, respectively.  $f_\sigma(\cdot)$  denotes the sigmoid function.

## 2. Input Gate:

The input gate determines which new feature information is to be incorporated into the memory cell at the current time step. Analogous to the forgetting gate, the input gate calculates its values based on the current input  $X^t$  and the hidden state  $h^{t-1}$  from the previous time step. This gate generates a candidate memory cell ( $\tilde{C}_{out}^t$ ) designed to hold potential new feature information. The output from the input gate, which ranges from 0 to 1, specifies how much information from the candidate memory cell is integrated into the memory cell, with 0 representing total omission and 1 indicating complete retention. These dynamics are detailed in Equations (17) and (18):

$$I_{out}^t = f_\sigma(W^i \cdot [h^{t-1}, X^t] + b^i), \quad (17)$$

$$\tilde{C}_{out}^t = f_{tanh}(W^C \cdot [h^{t-1}, X^t] + b^C), \quad (18)$$

where  $W^i$  and  $b^i$  represent the weights and biases of the input gates, respectively.  $W^C$  and  $b^C$  constitute the weights and biases of the candidate memory cells, respectively. The function  $f_{tanh}(\cdot)$  denotes the tanh function.

## 3. Cell State Update:

Memory cells update their contents under the guidance of forgetting gates and input gates. The forgetting gate determines what information is to be discarded from the previous memory cell, and the input gate determines what new information is to be added to the memory cell. Memory cell  $C^t$  can be represented by Equation (19):

$$C_{out}^t = C^t = F_{out}^t \cdot C^{t-1} + I_{out}^t \cdot \tilde{C}_{out}^t, \quad (19)$$

where  $C^{t-1}$  denotes the candidate memory unit of the previous moment.

## 4. Output Gate:

The output gate determines how, at a given time step, information from the memory cell influences the next hidden state. This gate uses the current input  $X^t$  and the previous hidden state  $h^{t-1}$  to perform its calculation. It modulates the hidden state values using a tanh function. The result,  $H_{out}^t$ , is formulated as follows in Equation (20):

$$H_{out}^t = h^t = f_\sigma(W^o \cdot [h^{t-1}, X^t] + b^o) \cdot f_{tanh}(C^t), \quad (20)$$

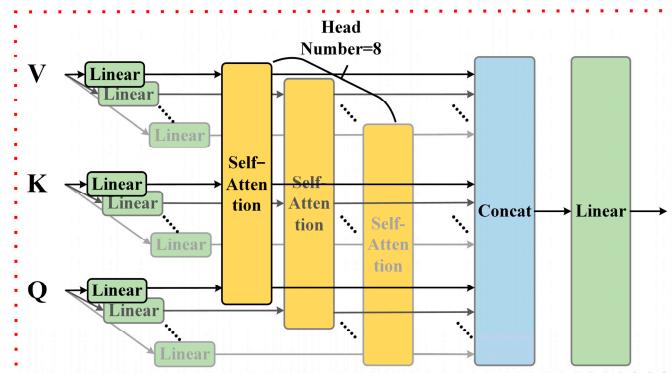
where  $h^t$  symbolizes the hidden state at this moment;  $W^o$  and  $b^o$  stand for the weight and bias of the output gate.

LSTM architecture includes forgetting gates, input gates, and output gates, which enable the network to selectively forget, retain, and communicate data. This structure is particularly effective for managing sequences that extend over long periods, making it adept at processing data with long-term implications, such as battery performance data.

### 3.4. Multi-Head Self-Attention

The model incorporates a multi-head self-attention (MHSA) mechanism [36], depicted in Figure 7. The fundamental concept of the multi-head self-attention mechanism involves

mapping the input features to various distinct attention subspaces, followed by the independent calculation of attention weights within each subspace. This approach enables the model to simultaneously focus on varied sequence positions and feature dimensions, thus enhancing its ability to fully grasp the sequence's structure and significance. In detail, for every head, the multi-head self-attention mechanism generates a collection of attention weights, which are amalgamated to produce the conclusive attention output.



**Figure 7.** Structure of multi-head self-attention.

The formula used for the attention mechanism of each head is presented in Equation (21), while the multi-head attention mechanism itself is detailed in Equation (22):

$$A\_Head(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{n'}}\right)V, \quad (21)$$

$$M\_Head(Q, K, V) = (A\_Head\_Q_1, K_1, V_1) \cup A\_Head\_Q_2, K_2, V_2 \cup \dots \cup W^M \quad (22)$$

where  $Q$ ,  $K$ , and  $V$  represent the linear transformations of the query, key, and value, respectively, and  $K^T$  represents the transposition of the key  $K$ .  $\sqrt{n'}$  represents the scaling factor, which controls the attention scores to remain at a stable gradient. The  $\text{softmax}(\cdot)$  function is used to compute the attention weights.  $W^M$  represents the attention weights in each of the subspaces.

In summary, the multi-head self-attention mechanism learns different feature representations from different subspaces by learning multiple attention heads in parallel, each of which can focus on different aspects or features in the sequence, thus providing richer expressive power; the multi-head self-attention mechanism simultaneously focuses on information at different locations in the sequence, and each of which learns dependencies at different granularities, which allows capturing the local and global dependencies in the input sequence, thus providing a more comprehensive understanding of the semantic structure of the sequence, speeding up model training and inference and improving model efficiency.

## 4. Result and Discussion

### 4.1. The Evaluation Criteria

In this study, several evaluation metrics are employed to provide a comprehensive assessment of the performance of the proposed TFMN estimation model. These metrics help to objectively measure the predictive ability of the model and reveal its accuracy and robustness from different perspectives. The following are the five evaluation metrics used:

1. Mean square error (MSE): The mean square error is a commonly used metric for assessing the difference between a model's predictions and the true values. It calculates the average of the squared differences between the predicted and true values to measure the average deviation of the predictions.

$$MSE = \frac{1}{n} \sum_{t=1}^n (f(x)_t - y_t)^2, \quad (23)$$

2. Root mean square error (*RMSE*): The *RMSE* is the square root of the mean square error, which represents the average difference between the predicted value and the true value. The *RMSE* is more sensitive to outliers and can be used to measure the accuracy of the model.

$$RMSE = \sqrt{\frac{\sum_{t=1}^n (f(x)_t - y_t)^2}{n}}, \quad (24)$$

3. Mean absolute error (*MAE*): The mean absolute error (*MAE*) is the average of the absolute differences between the forecast and the true value and measures the average error in the forecast. Unlike *MSE*, *MAE* does not amplify the effect of large errors and therefore better reflects the overall accuracy of the forecast.

$$MAE = \frac{1}{n} \sum_{t=1}^n |f(x)_t - y_t|, \quad (25)$$

4. Mean percentage absolute error (*MAPE*): The mean percentage absolute error (*MAPE*) is the average of the relative differences between the predicted and true values, expressed as a percentage. It measures the relative error of the model over different data ranges and can reflect the relative accuracy of the predictions.

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{f(x)_i - y_i}{y_i} \right|, \quad (26)$$

5. Maximum Absolute Error (*MAXE*): The maximum absolute error (*MAXE*) is the maximum value of the absolute difference between the predicted value and the true value, which identifies the model's worst-case prediction error. *MAXE* is particularly sensitive to outliers and is useful in understanding the model's maximum risk in making predictions.

$$MAXE = \max(|f(x)_t - y_t|), \quad (27)$$

where  $y_t$  is the actual value of time,  $f(x)_t$  is the  $t$  predicted value of time, and  $n$  is the number of predictions.

By using these evaluation metrics, we can comprehensively assess the performance of the TFMN estimation model to better understand its performance in battery health state assessment. A comprehensive analysis of these metrics will help us gain insight into the strengths and limitations of the model and provide valuable references for further research.

#### 4.2. Experimental Settings

In this paper, three datasets are used. In dataset A, CA-11, CA-12, CA-14, CA-16, and CA-17 are used as the training validation set, while CA-13 and CA-15 are used as the test set. In dataset B, NA-11, NA-12, NA-14, NA-16, NA-17, and NA-17 are used as the training validation set, while NA-03 and NA-15 are used as the test sets. For dataset C, CH7, CH10, CH11, CH14, CH19, CH28, CH29, CH32, CH36, and CH44 are used as the training validation set, while CH13 and CH48 are used as the test set. The training-validation set is randomly divided into a training set and a validation set in the ratio of 8:2.

For network training, the mean square error (*MSE*) was selected as the model's loss function. Subsequently, the gradient-driven AdamW optimization algorithm [37] was employed to adjust the weights and biases in the network model, aiming to reduce the loss function. An initial learning rate of 0.0035 was chosen. Finally, an early stopping mechanism was used to prevent model overfitting. Specifically, model training was termi-

nated if the validation loss did not decrease within the next 130 epochs. Notably, min–max normalization was performed on the data prior to feature extraction, as shown in the algorithm in Equation (28):

$$nor(X_i) = \frac{X_i - \min(X_i)}{\max(X_i) - \min(X_i)}; i \in \{1, \dots, n\}, \quad (28)$$

where  $X_i$  is the original data and  $n$  is the total amount of data.  $nor(\cdot)$  signifies that the data undergoes normalization prior to being input into the model, which facilitates quicker convergence of the model during training. Additionally, the equipment setup and model parameters utilized in the experiment are detailed in Table 3.

**Table 3.** Hardware setup and parameters for the model.

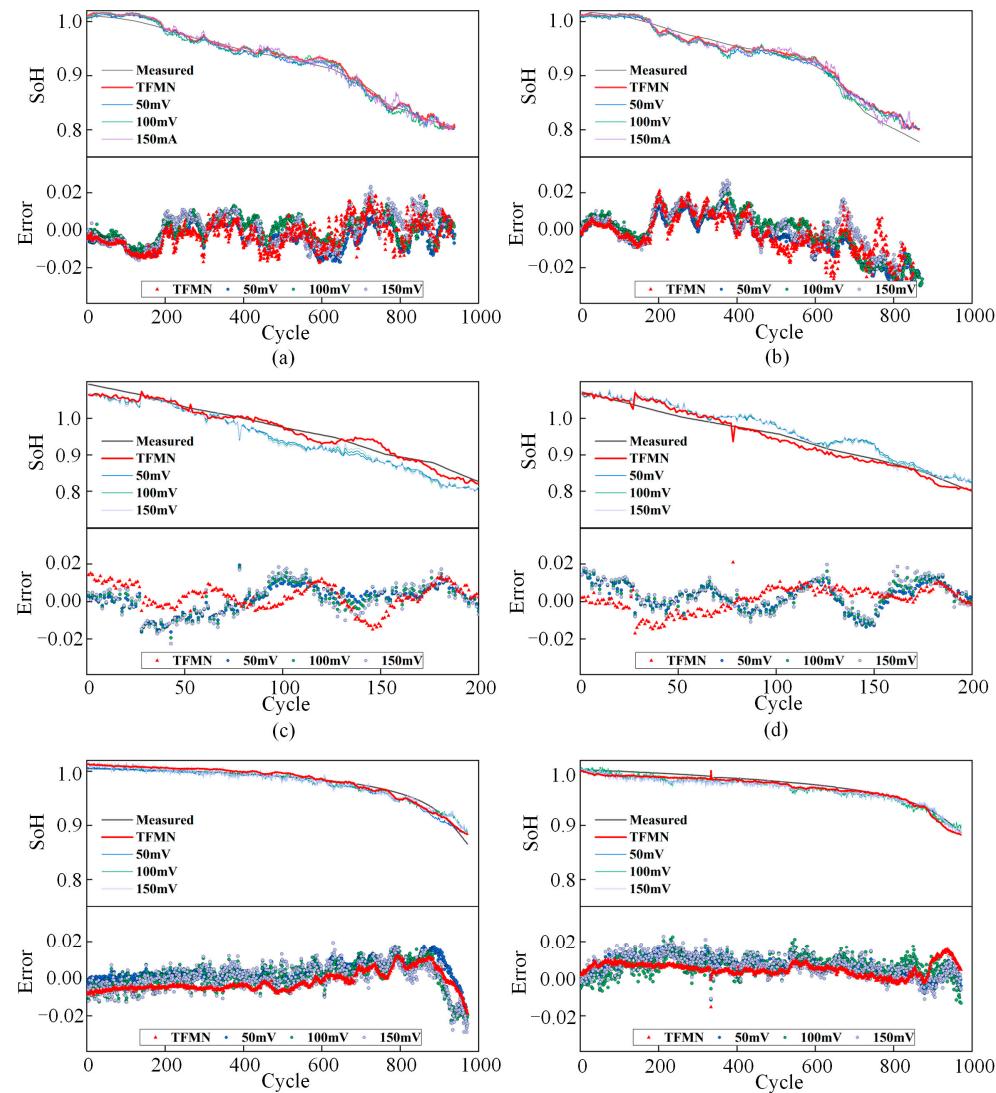
Class	Specifications and Parameters
CPU	Intel Xeon(R) Gold 6248R CPU @ 3.00 GHz × 96
GPU	Tesla V100s Pcie 32 GB
Memory	376.4 GiB
Environment	Pytorch_1.11.0 Torchvision_0.12.0
Model	Batch_sizes = 16 Epoch = 1000 Patience = 130 Learning rate = 0.00035 Optimizer = AdamW

#### 4.3. Robust Experiments

In robustness experiments, the effect of noise was assessed to determine the performance of the TFMN across varying noise intensities. Noise levels of 50 mV, 100 mV, and 150 mV were introduced to mimic the uncertainty typical of battery data. Analysis of the experimental outcomes provided insights into the stability and resilience of the TFMN when subjected to noise.

The results indicated that the efficacy of each model diminishes as the noise intensity escalates across all evaluation metrics. Despite this, the TFMN continues to exhibit high accuracy levels under various noise conditions. For instance, with the CA-15 dataset, as noise intensity increased from 50 mV to 150 mV, the MSE for TFMN rose only from 0.53 to 0.67, demonstrating TFMN's greater noise resistance and its ability to mitigate data uncertainty effects. To confirm the experiment's reliability, the parameter settings from Section 4.2 were applied, and the tests were performed under consistent conditions. The outcomes are documented in Table 4 and Figure 8. Additionally, TFMN demonstrated consistent performance on the CA-13, NA-13, and CH48 datasets. Although all models show decreased performance at higher noise levels, TFMN consistently recorded low scores on various evaluation metrics, affirming its ability to handle diverse data qualities effectively.

Taken together, the outcomes of these robustness trials underscore TFMN's superiority in managing noise interference. Compared with other models, TFMN shows stronger stability and robustness, which provides strong support for its reliability in practical applications, especially in battery data scenarios with uncertainty.



**Figure 8.** Robustness comparison for cases with noisy inputs of different intensities. (a,b) CA-13 and CA-15 for dataset A; (c,d) NA-13 and NA-15 for dataset B; (e,f) CH13 and CH48 for dataset C.

**Table 4.** Outcomes from incorporating noise.

LIB	Error Criteria	Results Analysis of Different Models			
		50 mV	100 mV	150 mV	TFMN
CA-13	MSE (%)	0.4545	0.5155	0.6153	0.4794
	RMSE (%)	0.6351	0.7286	0.7810	0.6996
	MAE (%)	0.5613	0.5561	0.6357	0.5628
	MAPE (%)	0.6651	0.6361	0.6589	0.6139
	R <sup>2</sup>	0.9627	0.9651	0.9153	0.9755
	MAXE (%)	1.5618	1.6878	1.9571	1.5073
CA-15	MSE (%)	0.5371	0.6156	0.6753	0.4974
	RMSE (%)	0.7651	0.8468	0.6678	0.7053
	MAE (%)	0.6561	0.6865	0.7858	0.5821
	MAPE (%)	0.6868	0.7668	0.8751	0.6458
	R <sup>2</sup>	0.9765	0.9661	0.9536	0.9786
	MAXE (%)	1.9354	2.1252	3.1245	1.8925

**Table 4.** Cont.

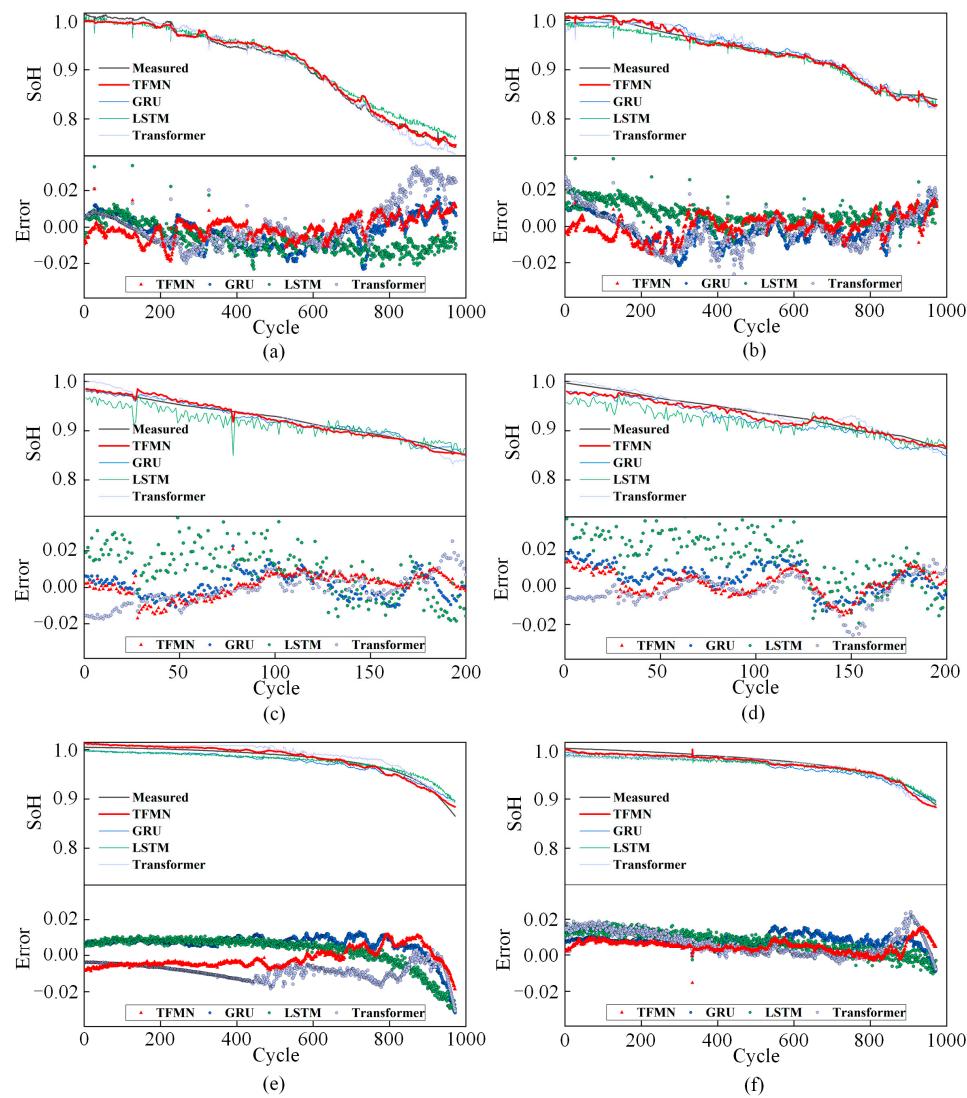
LIB	Error Criteria	Results Analysis of Different Models			
		50 mV	100 mV	150 mV	TFMN
NA-13	MSE (%)	0.4563	0.4964	0.7571	0.3059
	RMSE (%)	0.5935	0.5964	0.9686	0.5788
	MAE (%)	0.4856	0.5585	0.8864	0.5054
	MAPE (%)	0.5071	0.5217	0.6945	0.5861
	R <sup>2</sup>	0.8965	0.9875	0.9913	0.9756
	MAXE (%)	1.5512	2.0617	2.5523	1.4537
NA-15	MSE (%)	0.4680	0.4835	0.5422	0.3582
	RMSE (%)	0.7967	0.8991	0.7328	0.5879
	MAE (%)	0.8546	0.8823	0.6257	0.4636
	MAPE (%)	0.7964	1.0068	0.6633	0.5394
	R <sup>2</sup>	0.9765	0.9765	0.9765	0.9765
	MAXE (%)	1.8684	2.1786	2.2015	1.8966
CH13	MSE (%)	0.4448	0.5151	0.5340	0.3259
	RMSE (%)	0.6669	0.7177	0.7308	0.5709
	MAE (%)	0.4969	0.5501	0.5576	0.5030
	MAPE (%)	0.5172	0.5728	0.5793	0.5167
	R <sup>2</sup>	0.9551	0.9422	0.9389	0.9690
	MAXE (%)	1.9694	2.5451	2.8727	1.8671
CH48	MSE (%)	0.6735	0.6885	0.7575	0.3664
	RMSE (%)	0.8207	0.8297	0.8703	0.6053
	MAE (%)	0.7193	0.7176	0.7415	0.5196
	MAPE (%)	0.7321	0.7352	0.7559	0.5346
	R <sup>2</sup>	0.8810	0.8961	0.8680	0.9487
	MAXE (%)	1.7371	2.2433	2.2670	1.5635

#### 4.4. Comparative Experiments

In the comparative tests, the precision of the proposed TFMN was further confirmed through trials with several open-source models (GRU, LSTM, and Transformer), utilizing dataset A, dataset B, and a public dataset for individual assessments. Consistent experimental parameters and identical environmental conditions were upheld to ensure uniformity in experimental outcomes. The findings are displayed in Table 5 and Figure 9.

From the experimental results, it can be clearly observed that TFMN demonstrates significant advantages under all evaluation metrics. Taking the CA-15 dataset as an example, TFMN exhibits reductions of about 53.4%, 45.8%, 44.9%, 38.2%, and 52.7% in five evaluation metrics, namely MSE, RMSE, MAE, MAPE, and MAXE, respectively, with respect to GRU, LSTM, and Transformer models. Similarly, on the XQ-18 dataset, TFMN achieves about 47.8%, 36.8%, 34.5%, 32.9%, and 52.7% performance improvement with respect to GRU, LSTM, and Transformer models, respectively.

These findings reinforce the TFMN's advantage in terms of estimation precision. By delivering more precise estimation outcomes, TFMN offers a solid foundation for evaluating and managing battery health. An exhaustive review of the comparative testing results shows that TFMN yields more consistent and precise estimations across both our own dataset and the publicly accessible dataset. It exhibits a distinct edge over the open-source models GRU, LSTM, and Transformer across multiple evaluation metrics. This emphasizes TFMN's exceptional capability in assessing the health state of lithium-ion batteries, offering a robust resource and insights for enhancing battery management and maintenance tactics.



**Figure 9.** SOH prediction performance comparison of different models. (a,b) CA-13 and CA-15 for dataset A; (c,d) NA-13 and NA-15 for dataset B; (e,f) CH13 and CH48 for dataset C.

**Table 5.** Outcomes from comparative trials.

LIB	Error Criteria	Results Analysis of Different Models			
		LSTM	GRU	Transformer	TFMN
CA-13	MSE (%)	1.2834	1.5204	1.0425	0.4792
	RMSE (%)	1.1329	1.2330	1.0151	0.6922
	MAE (%)	0.8782	0.9489	0.8133	0.5686
	MAPE (%)	0.9590	1.0392	0.9713	0.6163
	R <sup>2</sup>	0.9156	0.9021	0.9388	0.9756
	MAXE (%)	4.7962	4.4144	2.6023	1.5071
CA-15	MSE (%)	0.6590	1.5595	0.6792	0.4978
	RMSE (%)	0.7252	2.0615	0.8235	0.7055
	MAE (%)	0.6248	1.5677	0.6535	0.5829
	MAPE (%)	0.7515	1.6519	0.7225	0.6452
	R <sup>2</sup>	0.9356	0.9369	0.9411	0.9784
	MAXE (%)	2.6867	2.5982	2.5186	1.8927

**Table 5.** Cont.

LIB	Error Criteria	Results Analysis of Different Models			
		LSTM	GRU	Transformer	TFMN
NA-13	MSE (%)	0.6348	0.7954	0.9938	0.3059
	RMSE (%)	0.4899	0.8938	0.9965	0.5788
	MAE (%)	0.6562	0.7215	0.8024	0.5054
	MAPE (%)	0.5682	0.7611	0.8515	0.5861
	R <sup>2</sup>	0.9510	0.9354	0.9111	0.9751
	MAXE (%)	2.0147	2.7117	3.5523	1.4537
NA-15	MSE (%)	0.5680	0.9835	0.5422	0.3582
	RMSE (%)	0.8967	0.9991	0.7328	0.5879
	MAE (%)	0.8546	0.8823	0.6257	0.4636
	MAPE (%)	0.6894	1.5414	0.6633	0.5394
	R <sup>2</sup>	0.9352	0.9125	0.9582	0.9761
	MAXE (%)	2.3684	3.3786	2.2015	1.8966
CH13	MSE (%)	0.7955	1.8423	0.8624	0.3259
	RMSE (%)	0.8138	0.9289	0.8574	0.5709
	MAE (%)	0.7398	0.7767	0.5584	0.5030
	MAPE (%)	0.7965	0.8086	0.9721	0.5167
	R <sup>2</sup>	0.9288	0.9252	0.9109	0.9690
	MAXE (%)	2.5563	2.8561	3.0534	1.8671
CH48	MSE (%)	0.7632	0.8565	0.8702	0.3664
	RMSE (%)	0.8824	0.8295	0.9157	0.6053
	MAE (%)	0.8335	0.7956	0.6176	0.5196
	MAPE (%)	0.7955	0.6841	0.6985	0.5346
	R <sup>2</sup>	0.9426	0.9407	0.9319	0.9487
	MAXE (%)	1.6268	1.8495	2.4002	1.5635

## 5. Conclusions

This study aims to overcome two major challenges in the health state assessment of lithium-ion batteries, namely insufficient accuracy and poor robustness under fast-charging conditions. With the introduction of an innovative TFMN estimation model, significant breakthroughs have been made in these aspects. This conclusion section will briefly summarize the research results and analyze the experimental findings.

The TFMN model effectively tackles the problems of limited precision and inadequate robustness in estimating battery health state by integrating various modules such as 1DCNN, CSAM, LSTM, and multi-head self-attention. Initially, the 1DCNN module adeptly captures both local and global features from the raw battery data, furnishing enriched inputs for the subsequent modeling stages. Moreover, the combination of LSTM and the attention mechanism allows the model to more accurately discern long-term dependencies within the sequential data, thereby enhancing the precision of battery health state estimations.

In robustness experiments, the performance of TFMN was evaluated in different noise environments. Although the predictive performance of the model decreases in high-noise situations, TFMN still maintains higher estimation accuracy relative to other models. This demonstrates the robustness of the model in coping with noise disturbances in real-world environments.

In comparison experiments, TFMN was evaluated against widely adopted models like GRU, LSTM, and Transformer. The test outcomes demonstrate that TFMN markedly surpasses other models across all evaluation metrics, achieving a reduction in MSE by approximately 31.1%, RMSE by 24.5%, MAE by 26.7%, MAPE by 18.9%, and MAXE by 34.5%. These findings robustly confirm the superior performance of the model in estimating battery health state.

In summary, TFMN has achieved remarkable results in Li-ion battery health state assessment. It effectively overcomes the challenges of insufficient accuracy and poor robustness of SOH estimation under fast-charging conditions, solves the problem of short-term

patterns and long-term dependencies in sequential data, accelerates the parallel processing of different levels of information, significantly improves the estimation performance and generalization ability of deep learning-based estimation methods, and overcomes the issues of insufficient accuracy, robustness, and real-time applicability faced by traditional methods. The model shows superior performance in different experiments. In the future, research will continue to deepen, further optimize the model, and apply it to practical scenarios of smart battery management and maintenance strategies to provide strong support and impetus for the development of the battery field.

## 6. Future Work

Our model has made great progress in SOH estimation under fast-charging conditions, and we decided to conduct the following studies in our future work:

(1) To satisfy SOH estimation under real working conditions, a battery test temperature sensor will be built to collect battery datasets under different working conditions at various temperatures. This will help understand the performance of the battery under different temperatures.

(2) Feature extraction is crucial for SOH estimation. Various features will be extracted, and an automated feature selection method will be used for correlation analysis to identify the features with the highest possible correlation.

(3) The processing performance of the model is a priority. Continuous optimization of the model will be conducted using a parameter optimization algorithm to find the optimal parameters for the model.

(4) Practical application: This result will be used as a standard to develop a complete practical application plan. The model will be integrated into the battery management system to verify its practicality in real-world battery management and maintenance, offering a more dependable tool for the battery industry.

**Author Contributions:** K.C.: conceptualization, methodology, validation, writing—review and editing, supervision, project administration, funding acquisition. D.W.: methodology, validation, writing—original draft, investigation, data curation. W.G.: investigation, visualization, writing—review and editing. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Data Availability Statement:** The relevant experiments of this study are still in progress. If you need our data for relevant research, you can contact the corresponding author of this article.

**Conflicts of Interest:** The authors declare that they have no known competing financial interests or personal relationships that may appear to influence the work reported in this paper.

## References

1. Oji, T.; Zhou, Y.; Ci, S.; Kang, F.; Chen, X.; Liu, X. Data-driven methods for battery SOH estimation: Survey and a critical analysis. *IEEE Access* **2021**, *9*, 126903–126916. [[CrossRef](#)]
2. Kumar, R.R.; Bharatiraja, C.; Udhayakumar, K.; Devakirubakaran, S.; Sekar, S.; Mihet-Popa, L. Advances in batteries, battery modeling, battery management system, battery thermal management, SOC, SOH, and charge/discharge characteristics in EV applications. *IEEE Access* **2023**, *11*, 105761–105809. [[CrossRef](#)]
3. Wu, M.; Wang, L.; Wu, J. State of health estimation of the LiFePO<sub>4</sub> power battery based on the forgetting factor recursive Total Least Squares and the temperature correction. *Energy* **2023**, *282*, 128437. [[CrossRef](#)]
4. Wu, L.; Lyu, Z.; Huang, Z.; Zhang, C.; Wei, C. Physics-based battery SOC estimation methods: Recent advances and future perspectives. *J. Energy Chem.* **2023**, *89*, 27–40. [[CrossRef](#)]
5. Gotz, K.; Hein, S.; Kunz, S.; Vetter, M.; Jossen, A.; Gasteiger, H.A. Studying Abuse Testing on Lithium-Ion Battery Packaging for Energy Storage Systems. *Sustainability* **2023**, *15*, 11545. [[CrossRef](#)]
6. Fan, Y.; Zhan, D.; Tan, X.; Lyu, P.; Rao, J. Optimization of cooling strategies for an electric vehicle in high-temperature environment. *Appl. Therm. Eng.* **2021**, *195*, 117088. [[CrossRef](#)]
7. Zhang, X.; Hou, J.; Wang, Z.; Jiang, Y. Joint SOH-SOC estimation model for lithium-ion batteries based on GWO-BP neural network. *Energies* **2022**, *16*, 132. [[CrossRef](#)]

8. Wang, Q.; Jiang, L.; Yu, Y.; Sun, J. Progress of enhancing the safety of lithium-ion battery from the electrolyte aspect. *Nano Energy* **2019**, *55*, 93–114. [[CrossRef](#)]
9. Kaur, K.; Garg, A.; Cui, X.; Singh, S.; Panigrahi, B.K. Deep learning networks for capacity estimation for monitoring SOH of Li-ion batteries for electric vehicles. *Int. J. Energy Res.* **2021**, *45*, 3113–3128. [[CrossRef](#)]
10. Liu, K.; Shang, Y.; Ouyang, Q.; Widanage, W.D. A data-driven approach with uncertainty quantification for predicting future capacities and remaining useful life of lithium-ion battery. *IEEE Trans. Ind. Electron.* **2020**, *68*, 3170–3180. [[CrossRef](#)]
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. [[CrossRef](#)]
12. Lin, C.; Tang, A.; Wang, W. A review of SOH estimation methods in Lithium-ion batteries for electric vehicle applications. *Energy Procedia* **2015**, *75*, 1920–1925. [[CrossRef](#)]
13. Ng, K.S.; Moo, C.S.; Chen, Y.P.; Hsieh, Y.C. Enhanced coulomb counting method for estimating state-of-charge and state-of-health of lithium-ion batteries. *Appl. Energy* **2009**, *86*, 1506–1511. [[CrossRef](#)]
14. Galeotti, M.; Cinà, L.; Giannamico, C.; Cordiner, S.; Di Carlo, A. Performance analysis and SOH evaluation of lithium polymer batteries through electrochemical impedance spectroscopy. *Energy* **2015**, *89*, 678–686. [[CrossRef](#)]
15. Vichard, L.; Ravey, A.; Venet, P.; Harel, F.; Pelissier, S.; Hissel, D. A method to estimate battery SOH indicators based on vehicle operating data only. *Energy* **2021**, *225*, 120235. [[CrossRef](#)]
16. Li, Z.; Shen, S.; Zhou, Z.; Cai, Z.; Gu, W.; Zhang, F. Novel method for modelling and adaptive estimation for SOC and SOH of lithium-ion batteries. *J. Energy Storage* **2023**, *62*, 106927. [[CrossRef](#)]
17. Huang, Z.; Best, M.; Knowles, J.; Fly, A. Adaptive piecewise equivalent circuit model with SOC/SOH estimation based on extended Kalman filter. *IEEE Trans. Energy Convers.* **2022**, *38*, 959–970. [[CrossRef](#)]
18. Laadjal, K.; Marques Cardoso, A.J. A review of supercapacitors modeling, SoH, and SoE estimation methods: Issues and challenges. *Int. J. Energy Res.* **2021**, *45*, 18424–18440. [[CrossRef](#)]
19. Guo, Y.; Yang, D.; Zhang, Y.; Wang, L.; Wang, K. Online estimation of SOH for lithium-ion battery based on SSA-Elman neural network. *Prot. Control. Mod. Power Syst.* **2022**, *7*, 40. [[CrossRef](#)]
20. Li, Y.; Li, K.; Liu, X.; Li, X.; Zhang, L.; Rente, B.; Grattan, K.T. 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. [[CrossRef](#)]
21. Wu, T.; Liu, S.; Wang, Z.; Huang, Y. SOC and SOH joint estimation of lithium-ion battery based on improved particle filter algorithm. *J. Electr. Eng. Technol.* **2022**, *17*, 307–317. [[CrossRef](#)]
22. Qiang, X.; Tang, Y.; Wu, L.; Lyu, Z. Li-Ion Battery State of Health Estimation Using Hybrid Decision Tree Model Optimized by Bayesian Optimization. *Energy Technol.* **2024**, *12*, 2301065. [[CrossRef](#)]
23. Kong, D.; Wang, S.; Ping, P. State-of-health estimation and remaining useful life for lithium-ion battery based on deep learning with Bayesian hyperparameter optimization. *Int. J. Energy Res.* **2022**, *46*, 6081–6098. [[CrossRef](#)]
24. Zhang, C.; Luo, L.; Yang, Z.; Zhao, S.; He, Y.; Wang, X.; Wang, H. Battery SOH estimation method based on gradual decreasing current, double correlation analysis and GRU. *Green Energy Intell. Transp.* **2023**, *2*, 100108. [[CrossRef](#)]
25. Sui, X.; He, S.; Vilnsen, S.B.; Meng, J.; Teodorescu, R.; Stroe, D.I. A review of non-probabilistic machine learning-based state of health estimation techniques for Lithium-ion battery. *Appl. Energy* **2021**, *300*, 117346. [[CrossRef](#)]
26. Akbar, K.; Zou, Y.; Awais, Q.; Baig, M.J.A.; Jamil, M. A machine learning-based robust state of health (SOH) prediction model for electric vehicle batteries. *Electronics* **2022**, *11*, 1216. [[CrossRef](#)]
27. Park, M.S.; Lee, J.; Kim, B.W. SOH estimation of Li-ion battery using discrete wavelet transform and long short-term memory neural network. *Appl. Sci.* **2022**, *12*, 3996. [[CrossRef](#)]
28. Lee, G.; Kwon, D.; Lee, C. A convolutional neural network model for SOH estimation of Li-ion batteries with Physical interpretability. *Mech. Syst. Signal Process.* **2023**, *188*, 110004. [[CrossRef](#)]
29. Van, C.N.; Quang, D.T. Estimation of SoH and internal resistances of Lithium-ion battery based on LSTM network. *Int. J. Electrochem. Sci.* **2023**, *18*, 100166.
30. Zhang, L.; Ji, T.; Yu, S.; Liu, G. Accurate prediction approach of SOH for lithium-ion batteries based on LSTM method. *Batteries* **2023**, *9*, 177. [[CrossRef](#)]
31. Gomez, W.; Wang, F.K.; Chou, J.H. Li-ion battery capacity prediction using improved temporal fusion transformer model. *Energy* **2024**, *296*, 131114. [[CrossRef](#)]
32. Severson, K.A.; Attia, P.M.; Jin, N.; Perkins, N.; Jiang, B.; Yang, Z.; Braatz, R.D. Data-driven prediction of battery cycle life before capacity degradation. *Nat. Energy* **2019**, *4*, 383–391. [[CrossRef](#)]
33. Attia, P.M.; Grover, A.; Jin, N.; Severson, K.A.; Markov, T.M.; Liao, Y.H.; Chueh, W.C. Closed-loop optimization of fast-charging protocols for batteries with machine learning. *Nature* **2020**, *578*, 397–402. [[CrossRef](#)] [[PubMed](#)]
34. Wen, J.; Chen, X.; Li, X.; Li, Y. SOH prediction of lithium battery based on IC curve feature and BP neural network. *Energy* **2022**, *261*, 125234. [[CrossRef](#)]
35. Shen, S.; Liu, B.; Zhang, K.; Ci, S. Toward fast and accurate SOH prediction for lithium-ion batteries. *IEEE Trans. Energy Convers.* **2021**, *36*, 2036–2046. [[CrossRef](#)]

36. Shi, D.; Zhao, J.; Wang, Z.; Zhao, H.; Wang, J.; Lian, Y.; Burke, A.F. Spatial-temporal self-attention transformer networks for battery state of charge estimation. *Electronics* **2023**, *12*, 2598. [[CrossRef](#)]
37. Jo, S.; Jung, S.; Roh, T. Battery state-of-health estimation using machine learning and preprocessing with relative state-of-charge. *Energies* **2021**, *14*, 7206. [[CrossRef](#)]

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.