International Journal of Innovative Computing, Information and Control
Volume **21**, Number **1**, February **2025**

ICIC International ©2025 ISSN 1349-4198
pp. **17–35**

STATE OF HEALTH PREDICTION OF LITHIUM-ION BATTERIES BASED ON TCN-TRANSFORMER

Jiayuan Li[1], Dezhi Xu[2,*], Tinglong Pan[1] and Dongnian Jiang[3]

[1]School of Internet of Things Engineering
Jiangnan University
No. 1800, Lihu Avenue, Wuxi 214122, P. R. China
6211920004@stu.jiangnan.edu.cn; tlpan@jiangnan.edu.cn

[2]School of Electrical Engineering
Southeast University
No. 2, Sipailou, Xuanwu District, Nanjing 210096, P. R. China
*Corresponding author: xudezhi@seu.edu.cn

[3]School of Electrical Engineering and Information Engineering
Lanzhou University of Technology
No. 36, Pengjiaping Road, Qilihe District, Lanzhou 730050, P. R. China
jiangdn@lut.edu.cn

Received February 2024; revised June 2024

Abstract. *Lithium-ion batteries are widely used in many fields because of their excellent performance, but in the long-term use process, lithium-ion batteries are facing problems such as capacity decline, which affects the performance of the equipment. Therefore, accurate prediction of battery health is an urgent problem to be solved. In this paper, a hybrid neural network model is constructed by combining temporal convolutional network (TCN) and Transformer. A lithium-ion battery state of health (SOH) prediction method combining XGBoost, T-SNE and TCN-Transformer was proposed. This paper collected nine lithium-ion battery health indicators (HI), and screened the importance of health indicators through XGBoost. T-SNE is used to reduce the dimension of the features, and the fused health indicators are input into the TCN-Transformer network to predict the SOH of lithium-ion batteries. The proposed method is validated using the NASA battery dataset, and compared with the LSTM model, GRU model, and RNN model. The experimental results show that the TCN-Transformer model can predict the SOH of lithium-ion batteries more accurately, and has good robustness and generalization.*
Keywords: Lithium-ion battery, State of health, XGBoost, Temporal convolutional network, Transformer

1. **Introduction.** With the ongoing advancement and creation of novel energy technologies, the manufacturing technology of lithium-ion batteries continues to mature. Lithium-ion batteries are extensively utilized in transportation, aviation, electronic products, power grid energy storage and other fields because of their high voltage, high energy density, excellent recyclability, and great safety. To ensure lithium-ion batteries function as intended, battery management technology is urged to monitor lithium-ion battery condition and enhance the safety and dependability of the battery system when it is operating [1, 2]. With the increase of the use time of lithium-ion batteries, the batteries' internal resistance will rise, its capacity will decrease, and a constant loss of lithium ions will result from the intricate electrochemical process within the battery. It affects the work of the equipment and even endangers the safety of human life [3]. As a result, understanding the SOH of

DOI: 10.24507/ijicic.21.01.17

lithium-ion batteries may assist lower the cost of equipment maintenance and increase system operating stability.

The two main methods presently utilized to predict the SOH of lithium-ion batteries are model-based and data-driven methods [4, 5, 6]. Model-based methods generally study the internal electrochemical principles of batteries, establish battery degradation models, and estimate SOH by parameter estimation of the models [7, 8]. In [9], electrochemical impedance spectroscopy (EIS) measurements are used to estimate SOH. In [10], the electrolyte polarization and aging mechanism (FOMeA) was taken into account using a fractional-order model. The link between the cycle number and the aging parameters is established by modeling the creation of the solid electrolyte interphase (SEI) layer and the side reactions of lithium plating, which cause battery aging. The model-based approach necessitates a deep understanding of physical chemistry, comprehension of the battery's internal reaction mechanism, precise description of the internal reaction's mathematical equations, and the creation of effective simulation models – all of which could be challenging in real-world scenarios [11]. Furthermore, the model-based method's capacity for generalization is weak.

In order to avoid complex battery modeling and make better use of the aging information of battery cells, data-driven approaches are increasingly being adopted. Examples include extreme learning machine (ELM) [12], support vector machine (SVR) [13], Gaussian process regression (GPR) [14, 15], and hybrid network model [16, 17]. The data-driven method maps the link between characteristics and predictions by treating the battery as a "black box", focusing solely on the data itself. In [18], it was suggested to use a battery SOH prediction model based on BP neural networks and incremental capacity analysis. By using the least squares approach, the mapping connection between temperature and the IC curve was established, and SOH prediction models at various temperatures were produced. By updating the feature data, the model is calibrated online in real time and the forecast accuracy is guaranteed. In [19], a fusion prediction model for SOH estimation combining CEMMDAN and gate recursive unit (GRU) is proposed. This method uses GRU network and ARIMA model respectively to predict decomposed local fluctuations and global degradation trends. In [20], an SOH prediction method based on CNN-BiLSTM and attention mechanism was proposed. The convolutional neural network (CNN) [21] was used to extract the features and reduce the dimensionality of the input factors of the model, which were input into the BiLSTM to learn the time correlation information of the input time series. The attention mechanism is introduced to strengthen the more important features in the input sequence. Graph neural networks (GNN) were introduced in [22] for battery SOH prediction. First, the time features were extracted by CNN and long short-term memory (LSTM) neural network. Spatial features are then obtained through graph sample aggregation (GraphSAGE) to propagate information on predefined graph structures, revealing deep information between HIs.

Compared to CNN, TCN has a stronger ability to capture long-term dependencies on data, and compared to traditional models such as LSTM and RNN, Transformer's self-attention mechanism allows the model to capture the dependencies between any two locations when processing sequence data. So, inspired by previous work, this study suggests a TCN-Transformer based SOH prediction model. By integrating the benefits of Transformer and TCN networks, the battery SOH will be predicted following the feature engineering treatment of health indicators. Compared with the traditional time series prediction model, the prediction performance of this strategy is improved. The following are this paper's primary contributions.

1) In order to reduce the redundancy of HIs, XGBoost was used to screen the health indicators, and T-SNE was used to reduce the dimension of the screened health indicators.

2) Use the advantages of TCN to learn the location and time information implied by HIs, and maintain the flow of location information between the post-sequence layers. The decoder module of Transformer model is used to learn the multi-dimensional feature information of data.

3) We can determine that the TCN-Transformer model's prediction effect is superior to the baseline model's by comparing its outcomes with each baseline model. This indicates the suggested method's rationality.

The structure of this article is as follows. In Section 2, the HI employed in this work is analyzed and the dataset used is described. Section 3 introduces the relevant algorithms used in this paper and explains the model's construction in this article. Section 4 compares the proposed model with the traditional model, and verifies the proposed method through experiments. Finally, we summarize the proposed methods in Section 5.

2. Extraction and Analysis of Battery Health Indicators.

With the continuous aging of the battery, its charging and discharging process will undergo some obvious changes. In this section, the dataset used in this paper is introduced, and the key features are extracted as health indicators reflecting the degree of battery aging [23].

2.1. NASA dataset introduction.

The publicly available lithium-ion battery cycle aging dataset provided by NASA was used in this study. The dataset was charged and discharged at 24°C on four type 18650 lithium phosphate batteries numbered B0005, B0006, B0007 and B0018 with nominal capacity of 2 Ah. There are 168 charge and discharge cycles for the batteries B0005, B0006, and B0007, and 132 cycles for the B0018 battery. The battery charge and discharge experiment follows these precise steps.

1) Charging stage: Charge the lithium-ion battery at a constant current (CC) of 1.5 A until the voltage reaches 4.2 V. Then, when the charging current drops to 20 mA, constant voltage (CV) mode charging is stopped.

2) Discharge stage: Discharge the battery with a constant current of 2 A. Due to the different cut-off voltage of each battery, the discharge will be stopped when the voltage of the B0005-B0018 battery drops to 2.7 V, 2.5 V, 2.2 V and 2.5 V, respectively. After charging and discharging, let the battery rest, and then proceed to the next cycle.

Figure 1 depicts the variations in voltage and current throughout the battery's charge and discharge cycle.

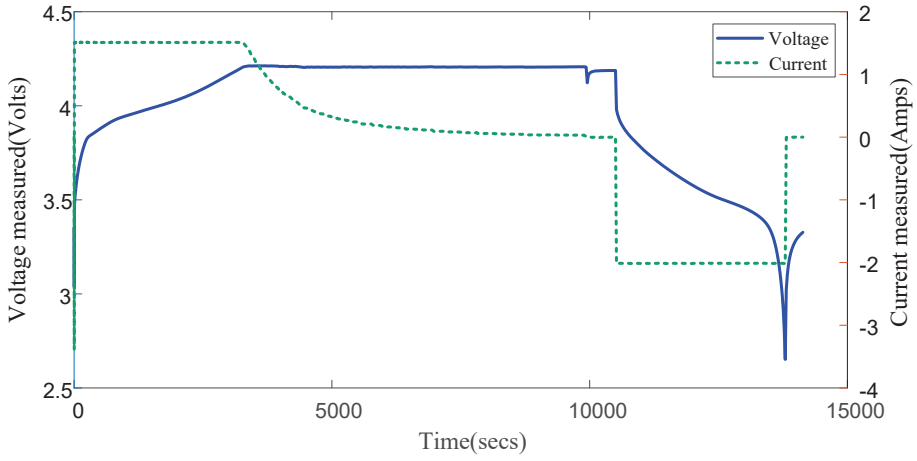


FIGURE 1. The charge and discharge cycle of the battery

While the battery is being used continuously, the four batteries will show different rates and different degrees of aging, and the battery capacity generally shows a downward trend. The capacity decay curve of B0005-B0018 is shown in Figure 2.

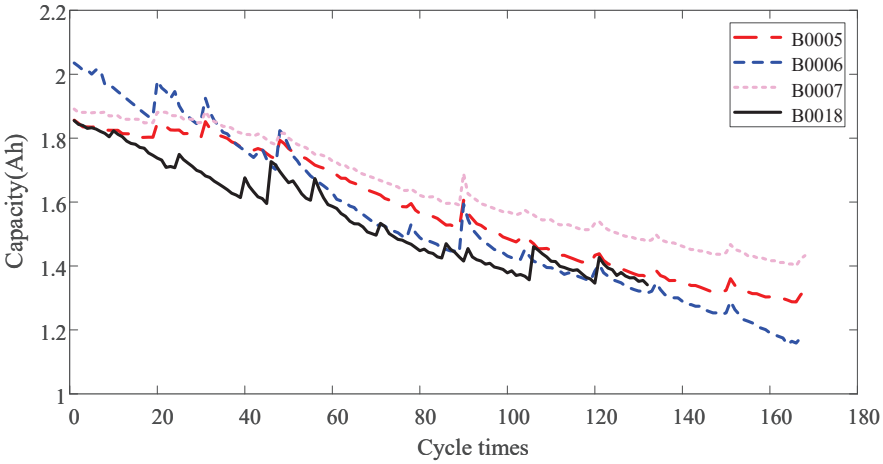


FIGURE 2. Capacity decline curve of B0005-B0018

Depending on the features of the battery, SOH is described variably; however, it is often characterized in terms of capacity and resistance. SOH is defined as follows in this article:

$$SOH(j) = \frac{C_j}{C_0} \tag{1}$$

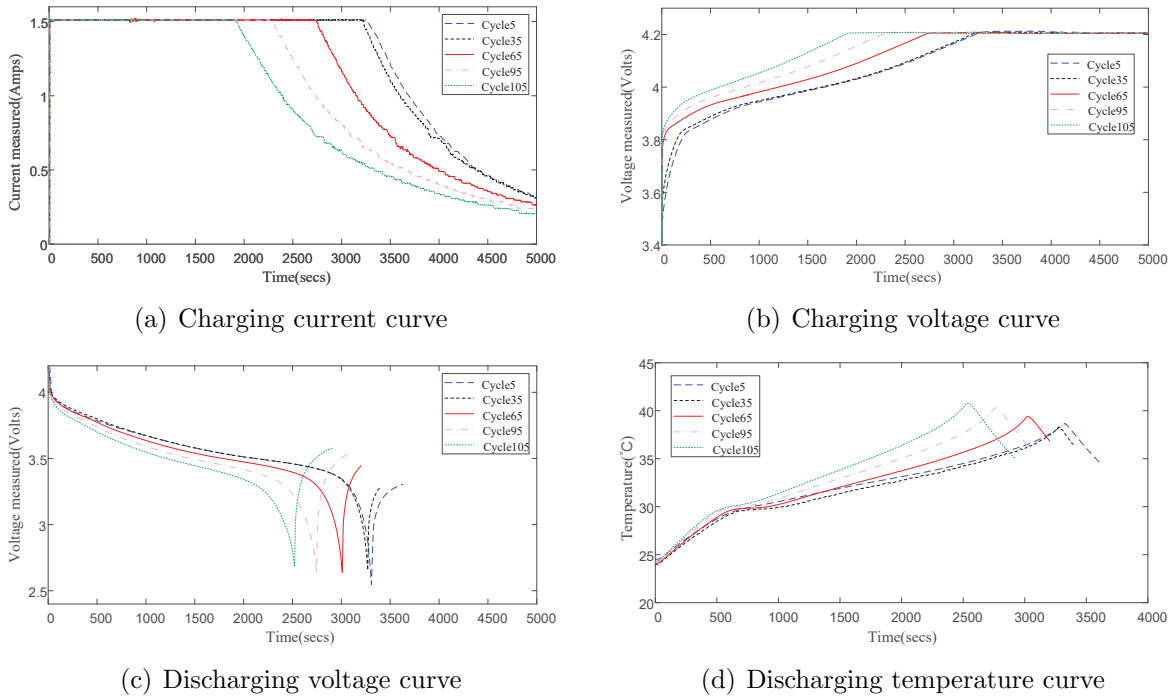
where C_j indicates the capacity of the battery at the jth cycle, and C_0 indicates the initial capacity of the battery.

2.2. **Feature extraction.** Nine characteristics that represent the battery's aging degree in terms of voltage, current, temperature, etc. were chosen to serve as the battery's health indicators. Each indicator is sensitive to battery decay and has obvious geometric characteristics.

Figure 3(a) describes the charging current changes of battery B0005 under different cycles. With the aging of the battery, the CC charging duration is gradually reduced, which can represent the change trend of battery SOH to a certain extent. Therefore, constant current charging time is chosen as HI1. The battery's CV charging time steadily increases as the CC segment's duration increasingly decreases. At CV stage, the charging current will gradually decrease, and with the deterioration of battery aging, the charging current starts to diminish more slowly. Therefore, the current drop time of battery constant voltage charging stage from 1.2 A to 0.5 A is chosen as HI2 in this paper. In addition, the integral of charging current curve over time was selected as the health index HI3 to measure the overall change of charging current during battery aging.

The charging voltage change of battery B0005 during various cycles is shown in Figure 3(b). With prolonged usage, the battery enters the constant voltage charging stage earlier and the voltage rise's slope steepens. In this paper, the isobaric rise time from 3.9 V to 4.1 V, which is the most gentle charging voltage, is selected as HF4.

Figure 3(c) describes the variation of discharge voltage of battery B0005 under different cycles. Under long-term use, the battery capacity becomes smaller, and it takes progressively less time for the discharge voltage to drop to its lowest point. The time when the battery discharge voltage falls to its lowest point during various cycles is chosen as HF5



(a) Charging current curve

(b) Charging voltage curve

(c) Discharging voltage curve

(d) Discharging temperature curve

FIGURE 3. Charge and discharge parameter curves of B0005 battery under different cycles

in this article. In addition, different cycles correspond to different voltage drop rates, so the voltage change rate of the battery during the discharge process can also represent the aging degree of the battery. In this paper, the time used for the discharge voltage to drop from 3.8 V to 3.5 V is taken as HF6.

Figure 3(d) describes the variation of discharge temperature of battery B0005 under different cycles. Under long-term use, the maximum battery discharge temperature generally shows an upward trend, and it takes less time to reach the maximum. In order to synthesize the two factors of the maximum temperature and the temperature rising speed, this paper selects the integral of the battery discharge temperature to the time under different cycles as HF7.

Incremental capacity (IC) analysis of batteries is widely used for SOH prediction of batteries [24]. It reflects the rate of capacity change under the change of unit voltage during the battery charging process. By differentiating the ratio of the charge capacity in the constant current charging phase to the terminal voltage, a smooth change in voltage can be converted into an easily identifiable peak on the IC curve. The capacity increment of the battery is calculated as shown in Equation (2).

$$\begin{cases} Q = It \\ IC_t = \dfrac{dQ_t}{dV_t} \approx \dfrac{\Delta Q_t}{\Delta V_t} = \dfrac{Q_t - Q_{t-1}}{V_t - V_{t-1}} \end{cases} \tag{2}$$

where t is the charging time, I is the constant charging current, IC_t indicates the capacity increment at time t, Q_t and Q_{t-1} indicate the battery capacity at time t and $t-1$ respectively, and V_t and V_{t-1} indicate the battery terminal voltage at time t and $t-1$ respectively.

Figure 4 shows the IC curves of the battery under different cycles. The IC peaks of different cycles have special positions, shapes and amplitudes. As the cycle increases, the IC curve shifts downward in general, which reflects the declining state of the battery.

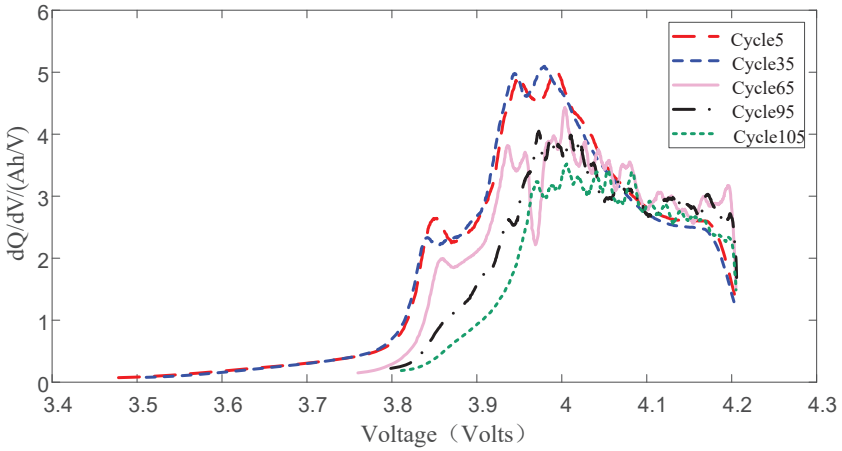


FIGURE 4. Discharging temperature curve of B0005 battery under different cycles

Therefore, the peak value is selected as HI8, and the voltage corresponding to the peak value is selected as HI9.

3. Theories of the Proposed SOH Prediction Method.

3.1. Importance evaluation of health indicators based on XGBoost. XGBoost is an extreme gradient boost decision tree based on Boosting algorithm, which is a strong learner composed of decision tree based learners. It improves the GBDT algorithm and carries out second-order Taylor expansion of the objective function to retain more information of the objective function. In this study, feature engineering in XGBoost method is used to sort the importance of features [25].

XGBoost can be expressed as

$$\hat{y}_i = \sum_{k=1}^{n} f_k(x_i), \quad f_k \in F \tag{3}$$

In the equation, $\hat{y}_i$ is the value after prediction; x_i is the ith sample; k is the number of decision trees; f_k is the structure and weight of decision tree. The decision tree's function space is denoted by F; n is the number of training samples.

The following equations illustrate the XGBoost objective function:

$$L(\varphi)^t = \sum_{i=1}^{n} l\left(y_i, \hat{y}_i\right) + \sum_{k=1}^{k} \Omega(f_k) \tag{4}$$

$$\Omega(f_k) = \gamma T + \frac{1}{2}\lambda\|\omega\|^2 \tag{5}$$

where $l\left(y_i, \hat{y}_i\right)$ is the loss function, namely the root-mean-square error; $\Omega(f_k)$ is the regularization term of each iteration; γ and λ are regular term coefficients, which are used to control the number and fraction of leaf nodes. T is the number of leaf nodes in the model; ω is the weight of leaf nodes.

The XGBoost algorithm considers the prediction residual in each fitting calculation to generate a new gradient lifting decision tree, and constantly changes the objective function according to the generated results. The objective function is expanded by Taylor series to get Formula (6).

$$L^{(t)} = \sum_{i=1}^{n} \left\{ l\left[y_i, \hat{y}_i^{(t-1)}\right] + g_i f_t(x_i) + \frac{1}{2} h_i f_t^2(x_i) \right\} + \Omega(f_t) \tag{6}$$

In the equation, $\hat{y}_i^{(t-1)}$ is the predicted value given by the model at step $t-1$; g_i and h_i are the first and second derivatives of the error function at $f_t = 0$, respectively.

The above formula is a convex function, and the derivation of it can obtain the extreme point of the objective function, which is the optimal solution:

$$\omega_i^* = -\frac{\sum_{i \in I_j} g_i}{\sum_{i \in I_j} h_i + \lambda} \tag{7}$$

where j is the leaf node; I_j is the instance set of the current node.

The two parameters of weight and gain used to build the tree reflect the correlation between a certain feature and the predicted target. The weight denotes how often a trait occurs. The gain is calculated by dividing the overall gain of the feature information by the total number of occurrences. The gain is calculated as follows:

$$\text{Gain} = \left[\frac{G_L^2}{H_L + \lambda} + \frac{G_R^2}{H_R + \lambda} - \frac{(G_L + G_R)^2}{H_L + H_R + \lambda} \right] - \gamma \tag{8}$$

The middle and lower corner marks L and R are left subtree and right subtree, respectively.

Figure 5 shows the feature importance distribution of nine health factors treated with XGBoost method, where the longer the bar chart, the higher the feature importance ranking in model training. As shown in the bar chart, HI1 and HI5 contribute the most to XGBoost model fitting, indicating that HI1 and HI5 are the key features that determine the prediction effect. Although HI3, HI4, HI6, HI7 and HI8 have not contributed much, they have all made some contributions in model training. HI2 and HI9 contribute the

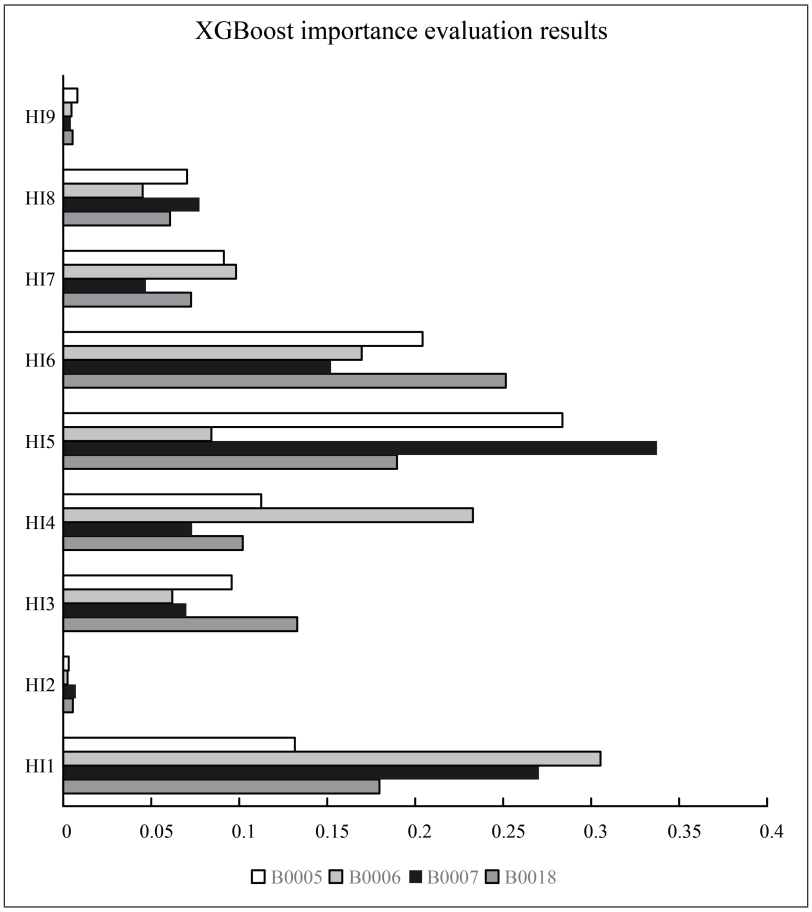


FIGURE 5. Importance evaluation results of HIS

least to XGBoost model fitting and are hardly used to build decision trees. Therefore, it is excluded in the subsequent research of this paper.

3.2. **Dimension reduction of health indicators based on T-SNE.** As an unsupervised nonlinear manifold learning algorithm, T-SNE can reduce the data dimension without changing the distribution of data points, and can fully extract low-dimensional sensitive feature information from high-dimensional data. The processes that are particular are as follows [26].

1) Let the high-dimensional space contain n data samples, expressed as $\{x_1, x_2, x_3, \ldots, x_n\}$. The conditional probability that data points x_i and x_j are adjacent points in the high-dimensional space can be expressed as

$$P_{j/i} = \frac{\exp\left(-\left\|x_i - x_j\right\|^2 / 2\sigma_i^2\right)}{\sum_{k \neq j} \exp\left(-\left\|x_i - x_k\right\|^2 / 2\sigma_i^2\right)} \tag{9}$$

Among them, σ_i is the standard deviation of Gaussian distribution centered on the data point x_i.

2) Calculate the joint distribution probability P_{ij} of x_i and x_j in space.

$$P_{ij} = \frac{P_{i/j} + P_{j/i}}{2n} \tag{10}$$

3) Set the data point in the low-dimensional space as $\{y_1, y_2, y_3, \ldots, y_n\}$, whose probability distribution matrix is Q. For the joint probability density q_{ij} of row i and column j in Q, the calculation formula is

$$q_{ij} = \frac{\left(1 + \left\|y_j - y_i\right\|^2\right)^{-1}}{\sum_{k \neq l} \left(1 + \left\|y_k - y_l\right\|^2\right)^{-1}} \tag{11}$$

4) The similarity between the probability distribution Q in low-dimensional space and the probability distribution P in high-dimensional space is represented by the Kullback-Leible (KL) divergence, which is used to calculate the similarity between the high- and low-dimensional spaces.

$$C = D_{\mathrm{KL}}(P \| Q) = \sum_i \sum_j P_{ij} \log_2 \frac{P_{ij}}{q_{ij}} \tag{12}$$

5) In order to obtain the best low-dimensional embedded data y_i, gradient descent is used to iterate the objective function repeatedly to obtain the minimum KL divergence expression.

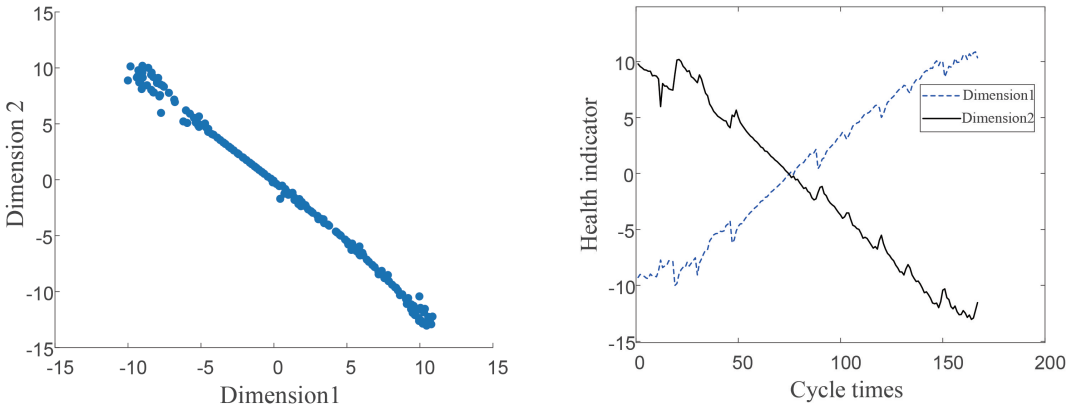
$$\frac{\delta C}{\delta y_i} = 4 \sum_j \left(P_{ij} - q_{ij}\right)\left(y_i - y_j\right)\left(1 + \left\|y_i - y_j\right\|^2\right)^{-1} \tag{13}$$

The iterative calculation formula is

$$y^t = y^{t-1} + \eta \frac{\delta C}{\delta y_i} + \alpha\left(y^{t-1} - y^{t-2}\right) \tag{14}$$

where η is the learning rate; α is the momentum factor. Iteratively, solve Equation (14) until $\frac{\delta C}{\delta y_i}$ is less than or equal to the set threshold.

Figure 6 shows the effect of T-SNE algorithm to reduce the original data dimension to two dimensions, where Figure 6(a) shows the joint distribution between the two dimensions, and Figure 6(b) shows the change trend of a single dimension. The data points in the joint distribution map are arranged along a clear descending trajectory, showing a continuous trend of change, and have strong one-dimensional characteristics. This continuous trend accurately reflects the continuity and process of actual battery capacity



(a) Joint distribution of dimensions 1 and 2 (b) Curves of dimensions 1 and 2

FIGURE 6. Dimension reduction effect of T-SNE

decline. Moreover, the fluctuation and trend of the curve in the health indicator graph are highly consistent with the original data and the capacity decline curve, which well completes the dimensional reduction while maintaining the spatial structure of the original data.

3.3. **TCN.** Temporal convolutional network (TCN) is proposed by [27], which is an improvement on one-dimensional convolutional neural networks. TCN uses full convolutional network to ensure that the input tensor and the output tensor have the same dimensions, and use causal convolution in order to establish a causal link between the network's layers. Furthermore, TCN broadens the network's receptive field to better capture the association between points in long-term time series when combined with dilated convolution. Figure 7 shows the structure of the expansion convolution in TCN. Its expression is as follows:

$$F(s) = (x *_d f)(s) = \sum_{i=0}^{k-1} f(i) \cdot x_{s-d \cdot i} \tag{15}$$

where $x \in R_n$ is a one-dimensional sequence input and d is the dilation rate. f is the convolution kernel. k is the number of convolutional kernels. s is the element of sequence.

In order to enable cross-layer information transmission and preserve network performance stability, the residual connection was created to address the issue of gradient disappearance or explosion that may arise from adding more network levels, as shown in

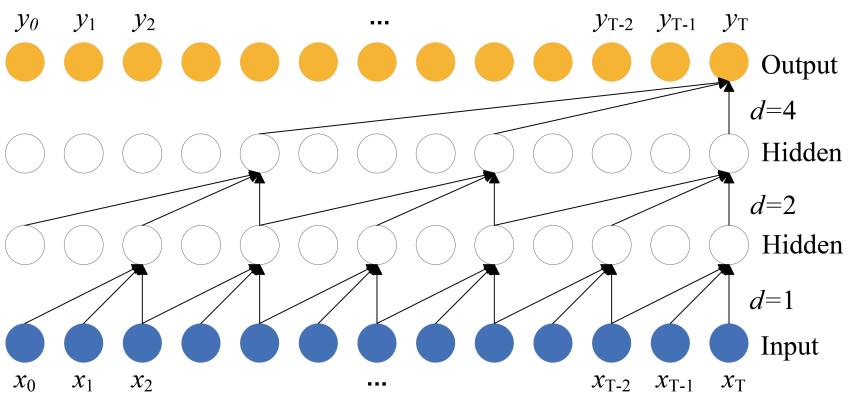


FIGURE 7. Dilated convolution structure diagram

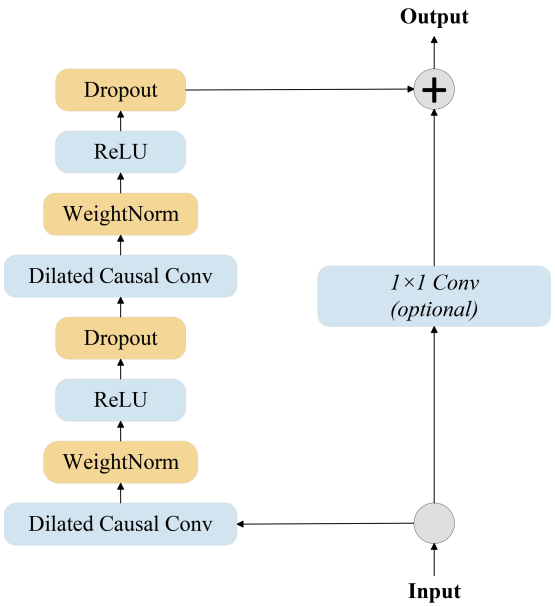


FIGURE 8. Residual connection structure diagram

Formula (16). Figure 8 shows the structure diagram of the residual connection. Residual connection can weigh and fuse the input x of the model into the output $F(x)$ of the model to obtain the final output.

$$O = \text{Activation}(x + F(x)) \tag{16}$$

where Activation$(\cdot)$ is the activation function.

3.4. **Transformer.** Transformer model has been widely used in text generation tasks due to its excellent predictive performance since it was proposed. Four components make up the general framework of the Transformer model: input, encoder, decoder, and output. The method in this paper adopts only the encoder part. Figure 9 displays the encoder's construction [28].

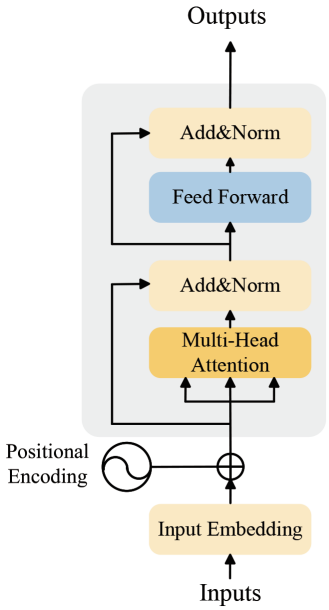


FIGURE 9. Transformer encoder structure diagram

As shown in the figure, Transformer's encoder module mainly consists of two parts: multi-head self-attention and feedforward neural network. The input vector is processed by the entity embedding layer into a new vector. It is then processed by positional encoding to ensure that the input data has relative position information in the sequence. The main reason why Transformer has better performance than other serial models is its multi-head attention module. The multi-head attention mechanism is composed of several self-attention mechanisms. The self-attention mechanism can obtain query matrix Q, key matrix K and value matrix V by linear transformation of input sequence. The attention weight of each input vector is obtained by the following equation.

$$\text{attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \tag{17}$$

where attention is the attention mechanism function; softmax is a normalized exponential function. d_k is the dimension of the key matrix.

The multi-head self-attention mechanism is defined as

$$\text{MultiHead}(Q, K, V) = \text{Concat}\left(\text{head}_1, \ldots, \text{head}_h\right)W_o \tag{18}$$

$$\text{head}_i = \text{attention}_i\left(QW_i^q, KW_i^k, VW_i^v\right) \tag{19}$$

where head_i is the ith head of attention; W_i^q, W_i^k, W_i^v are linear transformation weight matrices of query matrix, key matrix and value matrix obtained in the ith attention head; W_o is the linear transformation weight matrix of multi-head attention.

The residual connection in Transformer decoder module is used to solve the problem of gradient disappearance caused by the deepening of network hierarchy. Layer normalization speeds up convergence. The formula is as follows:

$$X_a = \text{LayerNorm}(\text{MultiHead}(X) + X) \tag{20}$$

$$X_{\text{en}} = \text{LayerNorm}\left(\text{GeLU}\left(0, X_aW_1 + b_1\right)W_2 + b_2 + X_a\right) \tag{21}$$

where X_a is the output of multi-head self-attention mechanism layer; X_{en} is the output of the feedforward network layer. LayerNorm() layer is normalization function.

3.5. **TCN-Transformer.** Figure 10 shows the proposed SOH prediction model for lithium-ion batteries. The model is divided into two components. The first section processes the battery aging data and extracts its features. Firstly, features that can characterize battery aging are extracted. XGBoost evaluates the importance of features, and to decrease the quantity of data and increase the network's training pace, the T-SNE approach is used to the dataset. The second part is the TCN-Transformer battery SOH prediction network, which uses TCN to capture the temporal and spatial characteristics of battery aging health indicators, and uses the multi-head attention mechanism in Transformer to strengthen and supplement the extracted features. The multi-scale time dependence relationship can be adaptive learned from the data of battery aging, and the aging information of battery can be extracted more accurately, so as to obtain better health state prediction results.

A residual module is used in the TCN model. In the residual module, the features and information contained in the battery aging data are extracted through expansion causal convolution, weight normalization is adopted to avoid the influence of excessive weight difference in training, and nonlinear factors are introduced into the data through ReLU. After adding the dropout layer, the overfitting problem can be alleviated to a certain extent. In Transformer, the output data of TCN is first encoded in position to enhance the timeliness of time series data, and then transferred to the attention mechanism in Transformer model. The results are then fed into the feedforward layer, these high-dimensional

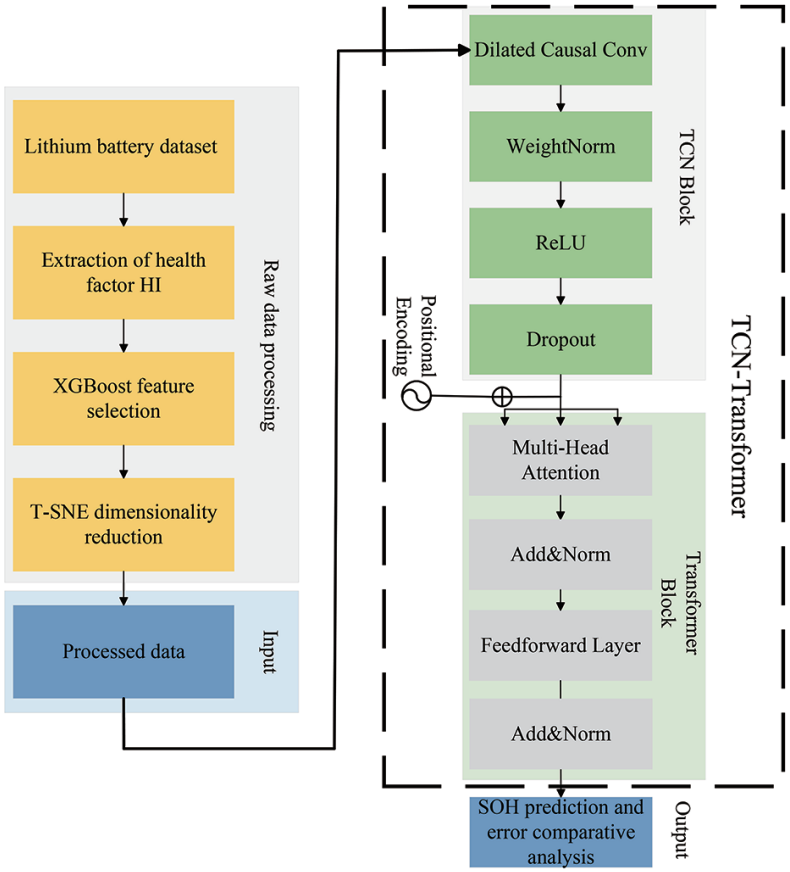


FIGURE 10. Overall framework for SOH prediction of lithium-ion batteries

features are compressed into the desired features, and the final prediction results are obtained.

4. Experiments and Result Analysis.

4.1. Evaluation criteria for battery SOH prediction accuracy.

The method's error was analyzed in this research using the following metrics: mean absolute percentage error (MAPE), root mean square error (RMSE), mean absolute error (MAE), and coefficient of determination (R^2). The following is the precise calculation formula:

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

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

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|(\hat{y}_i - y_i)| \tag{24}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y}_i)^2} \tag{25}$$

where n represents the number of samples; y_i indicates the real battery life, $\hat{y}_i$ represents the predicted battery life, and $\bar{y}_i$ represents the average of the raw data.

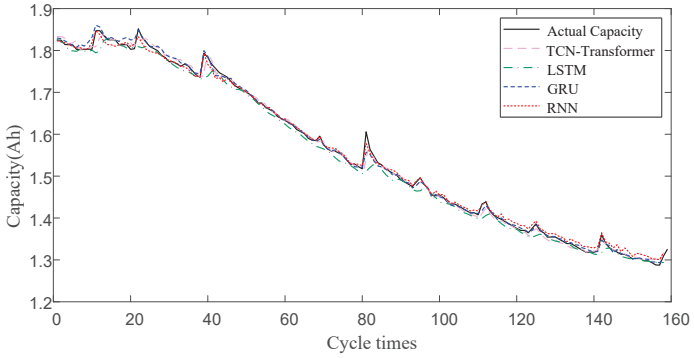
4.2. **Comparison of Li-ion battery SOH prediction results under various models.** Health indicators of batteries in the NASA dataset were collected for each charging and discharging cycle. After processed by XGBoost and T-SNE, the health indicators are used as input features, and the capacity of each cycle is used as a label to compose the lithium-ion battery life prediction dataset. To increase the content of data, the idea of cross-validation is adopted, that is, the data of B0006-B0018 are used as training sets to train the network, and the SOH of B0005 battery is predicted. Data from B0005, B0007, and B0018 are used as training sets to train the network, to predict the SOH of B0006, and so on.

$$X = \begin{bmatrix} x_1^1, x_2^1, \ldots, x_n^1 \\ x_1^2, x_2^2, \ldots, x_n^2 \\ \vdots \\ x_1^j, x_2^j, \ldots, x_n^j \end{bmatrix} \quad Y = \begin{bmatrix} y^1 \\ y^2 \\ \vdots \\ y^j \end{bmatrix} \tag{26}$$

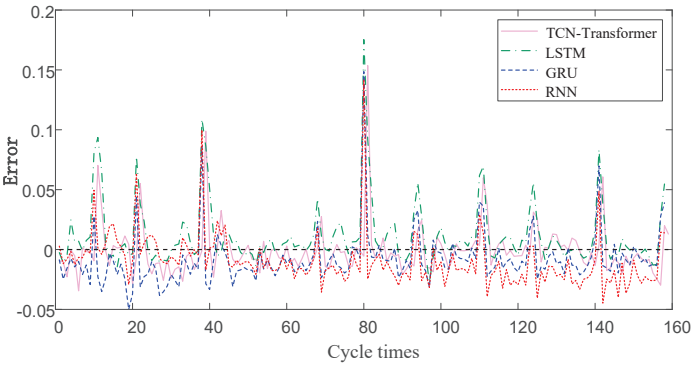
where x_n^j represents the nth health factor of the jth cycle, and y^j represents the capacity of battery of the jth cycle.

LSTM, GRU, and RNN networks with superior time series prediction ability are chosen as comparison networks of the proposed method in order to confirm the efficacy of the SOH prediction approach presented in this study, and all networks are trained and tested with the same dataset.

Figures 11-14 show the SOH prediction results and prediction errors of B0005-B0018 batteries. As shown in the figures, the LSTM method has a good forecasting performance in the stage of steady capacity decline, but it cannot effectively predict the capacity

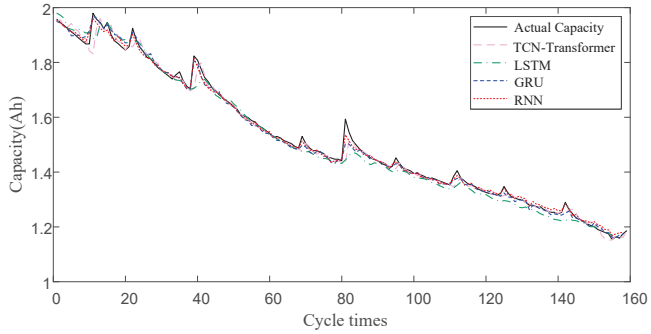


(a) SOH prediction results of B0005

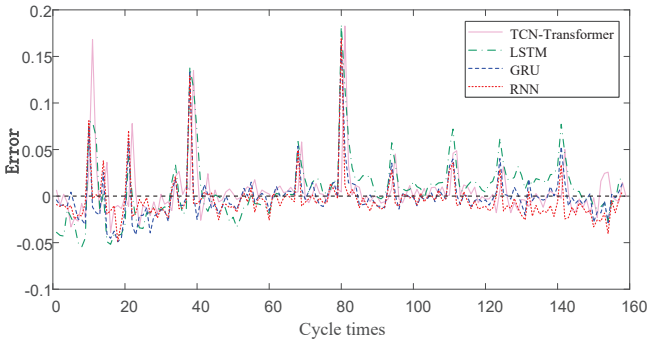


(b) Prediction error of B0005

FIGURE 11. Comparison of prediction results of B0005 under different models

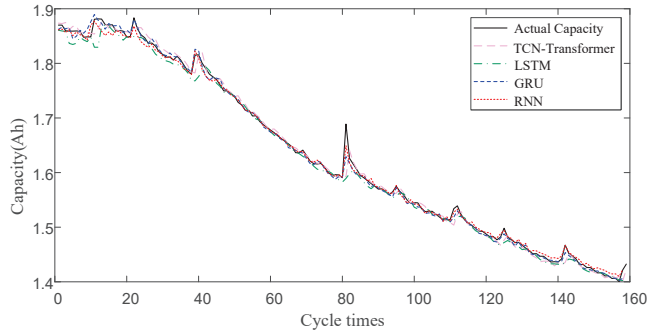


(a) SOH prediction results of B0006

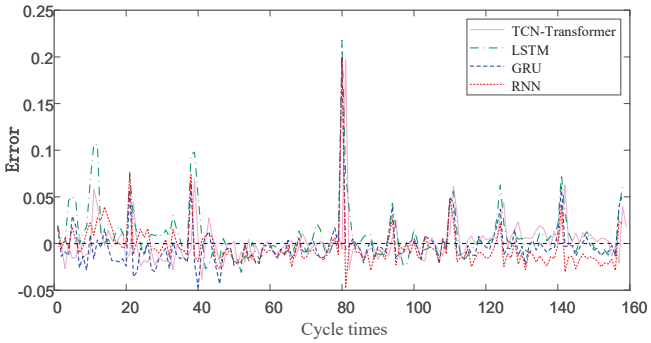


(b) Prediction error of B0006

FIGURE 12. Comparison of prediction results of B0006 under different models

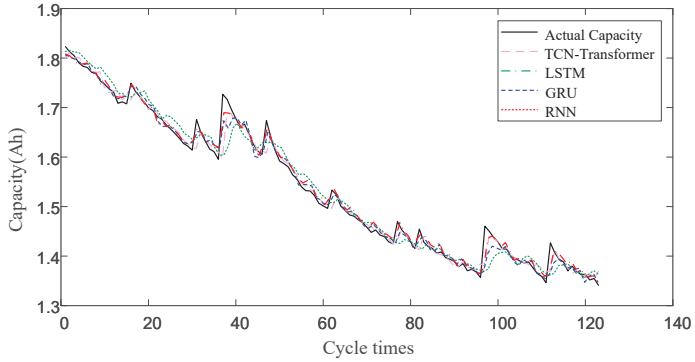


(a) SOH prediction results of B0007

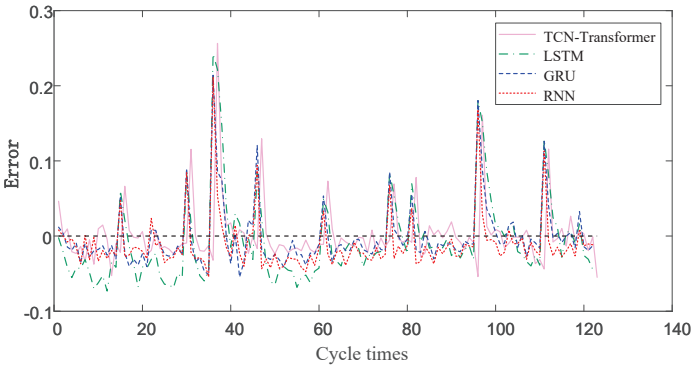


(b) Prediction error of B0007

FIGURE 13. Comparison of prediction results of B0007 under different models



(a) SOH prediction results of B0018



(b) Prediction error of B0018

FIGURE 14. Comparison of prediction results of B0018 under different models

rebound phenomenon of a large degree. GRU method has good prediction performance for the middle and late period of capacity decline, but it cannot predict accurately when the capacity fluctuates greatly in the early period. The RNN method has a good prediction effect when the capacity fluctuates greatly in the early and middle period, but the overall prediction result is too large in the late period of capacity attenuation. Although LSTM, GRU and RNN have their own advantages in forecasting, the TCN-Transformer model performs better in the first and intermediate phases of capacity decrease, and has the best overall forecasting performance among the four methods. We can also confirm the analysis just now from the prediction error graph. In general, LSTM method has the largest prediction error at the peak point, GRU method has a large error fluctuation at the early stage of capacity attenuation, and RNN method has a large error fluctuation at the late stage of capacity attenuation. On the other hand, the error of TCN-Transformer method is relatively stable in the early and late stages and always fluctuates around the 0 error line.

Table 1 shows the evaluation criteria under the four networks. According to the results in the table, except that the RMSE and R^2 results of B0006 battery are slightly better than those of TCN-Transformer when RNN method is used for prediction, TCN-Transformer of other batteries shows optimal indicators. The results demonstrate that, in comparison to the old model, the TCN-Transformer model's prediction accuracy has somewhat improved. Both the accuracy and prediction error have decreased.

5. **Conclusion.** A model based on TCN-Transformer is suggested to evaluate the SOH of Li-ion batteries, with the goal of improving the forecast accuracy. Firstly, nine health indicators HI1-HI9 were extracted from the voltage and current curves during charging and

TABLE 1. Evaluation results of SOH prediction

Battery	Model	MAE	RMSE	MAPE	R^2
B0005	TCN-Transformer	0.0075	0.0124	0.4792	0.9954
	LSTM	0.0089	0.0166	0.5645	0.9919
	GRU	0.0098	0.0133	0.6213	0.9947
	RNN	0.0104	0.0138	0.6915	0.9943
B0006	TCN-Transformer	0.0122	0.0220	0.7695	0.9909
	LSTM	0.0205	0.0307	1.3103	0.9823
	GRU	0.0127	0.0245	0.7991	0.9889
	RNN	0.0127	0.0213	0.8370	0.9915
B0007	TCN-Transformer	0.0068	0.0115	0.4134	0.9944
	LSTM	0.0086	0.0153	0.5174	0.9901
	GRU	0.0070	0.0117	0.4250	0.9942
	RNN	0.0076	0.0117	0.4695	0.9942
B0018	TCN-Transformer	0.0110	0.0196	0.7055	0.9800
	LSTM	0.0213	0.0288	1.3663	0.9568
	GRU	0.0136	0.0209	0.8748	0.9773
	RNN	0.0139	0.0203	0.9040	0.9791

discharging under different cycles, and then XGBoost was used to screen the importance of health indicators and eliminate the health indicators with low importance. The T-SNE algorithm was used to reduce dimensionality of multidimensional data to obtain indirect health indicators. Finally, the SOH of B0005-B0018 is predicted using the TCN-Transformer model. In the experiment, LSTM, GRU and RNN models are used to prove the effect of the proposed model with the same dataset. The findings demonstrate that the TCN-Transformer model, which has the best overall prediction accuracy and better evaluation indices, may continue to retain strong prediction performance in both the early and late phases of battery capacity attenuation.

The TCN-Transformer model is suitable for complex sequence processing tasks. However, when the sequence is long, the computational cost and memory consumption of the self-attention mechanism are high. Although TCN improves efficiency through convolution operations, the overall model still requires a lot of computational resources when processing sequences. Moreover, the model structure has high complexity, and more hyperparameters need to be adjusted, which increases the difficulty and time of model tuning. Optimizing computing resource consumption, simplifying model structure, improving generalization ability and tuning convenience are still important directions for future research and improvement.

Acknowledgment. This work is partially supported by the National Natural Science Foundation of China (Grant No. 62222307, No. 61973140) and the Natural Science Foundation of Jiangsu Province (Grant No. BK20211235). The authors also appreciatively acknowledge the helpful suggestions and comments of the reviewers, which have enhanced the presentation.

REFERENCES

[1] S. K. Pradhan and B. Chakraborty, Battery management strategies: An essential review for battery state of health monitoring techniques, *Journal of Energy Storage*, vol.51, 104427, 2022.

[2] X. Sui, S. He, S. B. Vilsen, J. Meng, R. Teodorescu and D.-I. Stroe, A review of non-probabilistic machine learning-based state of health estimation techniques for lithium-ion battery, *Applied Energy*, vol.300, 117346, 2021.

[3] S. A. Hasib, S. Islam, R. K. Chakrabortty, M. J. Ryan, D. K. Saha, M. H. Ahamed, S. Moyeen, I, S. K. Das, M. F. Ali, M. R. Islam, Z. Tasneem and F. R. Badal, A comprehensive review of available battery datasets, RUL prediction approaches, and advanced battery management, *IEEE Access*, vol.9, pp.86166-86193, 2021.

[4] M. Chen, G. Ma, W. Liu, N. Zeng and X. Luo, An overview of data-driven battery health estimation technology for battery management system, *Neurocomputing*, vol.532, pp.152-169, 2023.

[5] X. Wang, P. Ye, S. Liu, Y. Zhu, Y. Deng, Y. Yuan and H. Ni, Research progress of battery life prediction methods based on physical model, *Energies*, vol.16, no.9, 3858, 2023.

[6] J. Zhao, Y. Zhu, B. Zhang, M. Liu, J. Wang, C. Liu and X. Hao, Review of state estimation and remaining useful life prediction methods for lithium-ion batteries, *Sustainability*, vol.15, no.6, 5014, 2023.

[7] Y. Wang, G. Gao, X. Li and Z. Chen, A fractional-order model-based state estimation approach for lithium-ion battery and ultra-capacitor hybrid power source system considering load trajectory, *Journal of Power Sources*, vol.449, 227543, 2020.

[8] S. K. Rechkemmer, X. Zang, W. Zhang and O. Sawodny, Empirical Li-ion aging model derived from single particle model, *Journal of Energy Storage*, vol.21, pp.773-786, 2019.

[9] P. Iurilli, C. Brivio, R. E. Carrillo and V. Wood, Physics-based SOH estimation for Li-ion cells, *Batteries*, vol.8, no.11, 204, 2022.

[10] G. Zhu, C. Kong, J. Wang, V. J. Kang, Q. Wang and C. Qian, A fractional-order electrochemical lithium-ion batteries model considering electrolyte polarization and aging mechanism for state of health estimation, *Journal of Energy Storage*, vol.72, no.E, 108649, 2023.

[11] Z. Deng, L. Yang, H. Deng, Y. Cai and D. Li, Polynomial approximation pseudo-two-dimensional battery model for online application in embedded battery management system, *Energy*, vol.142, pp.838-850, 2018.

[12] X. Hou, X. Guo, Y. Yuan, K. Zhao, L. Tong, C. Yuan and L. Teng, The state of health prediction of Li-ion batteries based on an improved extreme learning machine, *Journal of Energy Storage*, vol.70, 108044, 2023.

[13] C. Wu, J. Fu, X. Huang, X. Xu and J. Meng, Lithium-ion battery health state prediction based on VMD and DBO-SVR, *Energies*, vol.16, no.10, 3993, 2023.

[14] J. Wang, R. Zhao, Q.-A. Huang, J. Wang, Y. Fu, W. Li, Y. Bai, Y. Zhao, X. Li and J. Zhang, High-efficient prediction of state of health for lithium-ion battery based on AC impedance feature tuned with Gaussian process regression, *Journal of Power Sources*, vol.561, 232737, 2023.

[15] R. Wang, Z. Hui and M. Yang, Gaussian process regression based on indirect health indicators for SOH estimation of lithium battery, *Energy Storage Science and Technology*, vol.12, no.2, pp.560-569, 2023.

[16] S. Buchanan and C. Crawford, Probabilistic lithium-ion battery state-of-health prediction using convolutional neural networks and Gaussian process regression, *Journal of Energy Storage*, vol.76, 109799, 2024.

[17] H. Xu, L. Wu, S. Xiong, W. Li, A. Garg and L. Gao, An improved CNN-LSTM model-based state-of-health estimation approach for lithium-ion batteries, *Energy*, vol.276, 127585, 2023.

[18] J. Wen, X. Chen, X. Li and Y. Li, Soh prediction of lithium battery based on IC curve feature and bp neural network, *Energy*, vol.261, no.A, 125234, 2022.

[19] S. Yu and X. Long, State-of-health prediction for lithium-ion batteries based on complete ensemble empirical mode decomposition with adaptive noise-gate recurrent unit fusion model, *Energy Technology*, vol.10, no.4, 2100767, 2022.

[20] S. Sun, J. Sun, Z. Wang, Z. Zhou and W. Cai, Prediction of battery SOH by CNN-BiLSTM network fused with attention mechanism, *Energies*, vol.15, no.12, 4428, 2022.

[21] K. Ummah, T. Karlita, R. Sigit, E. Yuniarno, I. Purnama and M. Purnomo, Effect of image preprocessing method on convolutional neural network classification of COVID-19 CT scan images, *nternational Journal of Innovative Computing, Information and Control*, vol.18, no.6, pp.1895-1912, 2022.

[22] X.-Y. Yao, G. Chen, M. Pecht and B. Chen, A novel graph-based framework for state of health prediction of lithium-ion battery, *Journal of Energy Storage*, vol.58, 106437, 2023.

[23] W. Zhou, Q. Lu and Y. Zheng, Review on the selection of health indicator for lithium ion batteries, *Machines*, vol.10, no.7, 512, 2022.

[24] A. Gismero, K. Norregaard, B. Johnsen, L. Stenhoj, D.-I. Stroe and E. Schaltz, Electric vehicle battery state of health estimation using incremental capacity analysis, *Journal of Energy Storage*, vol.64, 107110, 2023.
[25] R. Poornima, M. Elangovan and G. Nagarajan, Network attack classification using LSTM with XGBoost feature selection, *Journal of Intelligent & Fuzzy Systems*, vol.43, no.1, pp.971-984, 2022.
[26] Z. Zhang, B. Fan, Y. Liu, P. Zhang, J. Wang and W. Du, Rapid warning of wind turbine blade icing based on MIV-tSNE-RNN, *Journal of Mechanical Science and Technology*, vol.35, no.12, pp.5453-5459, 2021.
[27] S. Bai, J. Z. Kolter and V. Koltun, An empirical evaluation of generic convolutional and recurrent networks for sequence modeling, *arXiv Preprint*, arXiv: 1803.01271, 2018.
[28] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser and I. Polosukhin, Attention is all you need, *Advances in Neural Information Processing Systems*, vol.30, 2017.

Author Biography

Jiayuan Li received the B.Eng. degree in Electrical Engineering and Automation from Jilin University, Changchun, China in 2020. He is currently working toward the M.S. degree in Electrical Engineering with Jiangnan University, Wuxi, China.

His current research interests include the prediction of battery remaining using life and state of health.

Dezhi Xu received the Ph.D. degree in Control Theory and Control Engineering from Nanjing University of Aeronautics and Astronautics, China, in 2013.

He was a Visiting Fellow with the Department of Biomedical Engineering, City University of Hong Kong, China, from 2018 to 2019. He is currently a Professor and Doctoral Supervisor with the Southeast University. His research interests include data-driven control, fault diagnosis and fault-tolerant control, multi-agent systems and cyber-physical systems, technologies of renewable energy, motor control, and smart grid.

Prof. Xu was supported by the National Natural Science Fund for Excellent Young Scientists Fund Program in 2022. He was a recipient of the First Class Prize of Science and Technology Progression from the China General Chamber of Commerce in 2016, and the Best Young Scholar of Jiangnan University in 2022. He was a Guest Editor for the International Journal of Innovative Computing, Information and Control and the Electric Power. He currently serves as an Editorial Board Member for the International Journal of Innovative Computing, Information and Control, the Electric Power, the Electrotechnical Application and the Electrical Engineering. He is a Committee Member of the Association of Energy Internet, and Trusted Control in Chinese Association of Automation (CAA), and the Energy Storage in China Renewable Energy Society (CRES).

Tinglong Pan received his B.Eng. degree in Industrial Automation from China University of Mining and Technology, Xuzhou, China, in 1999, and the Ph.D. degree in Power Electronics and Power Drive from China University of Mining and Technology, Xuzhou, China, in 2004.

He is currently a Professor at Jiangnan University, where his research interests include microgrid control technology, power conversion technology, power drive system and its intelligent control technology.

Dongnian Jiang received his B.S. degree in Information and Computational Science from Xiamen University, Xiamen, China, in 2006, and the Ph.D. degree in Control Theory and Control Engineering from Lanzhou University of Technology, Lanzhou, China, in 2018.

He is currently an Associate Professor at Lanzhou University of Technology. His research interests include fault diagnosis and fault-tolerant control, advanced control of industrial processes.