batteries

MDPI

Article

A Hybrid Transformer Network Optimized by Kalman Filter for State of Health Prediction on Lithium-Ion Battery

Lei Xu [1,*], Peng Sun [2,*] and Nan Zhou [3]

[1] Shenyang Fire Science and Technology Research Institute of MEM, Shenyang 110034, China
[2] Key Laboratory of Electrochemical Energy Safety, Ministry of Emergency Management, XYZ Storage Technology Co., Ltd., Beijing 102400, China
[3] School of Mechanical Engineering and Automation, Northeastern University, Shenyang 110819, China
* Correspondence: syfri_xulei@126.com (L.X.); sunpeng@cpes.co (P.S.)

Abstract

State of Health (SOH) is a critical metric for evaluating the efficient and reliable operation of lithium-ion batteries (LIBs), although it cannot be directly measured. Accurate SOH prediction throughout the entire lifecycle of LIBs remains a significant challenge, primarily due to severe signal fluctuations and complex degradation mechanisms. In this paper, a novel hybrid Transformer-based architecture for SOH prediction is introduced, termed KF–SAMformer–GRU, which integrates a Kalman filter (KF) optimizer, a sharpness-aware minimization Transformer (SAMformer) model, and a multi-layer gated recurrent unit (GRU). To overcome the challenges of multivariate long-term forecasting, we innovatively integrate SAMformer to extract robust feature indicators, actively mitigating data distribution shifts via sharpness-aware minimization. Furthermore, reversible instance normalization (RevIN) is first introduced to tackle non-stationarity in multi-source datasets, effectively eliminating uncertainty and significantly boosting generalization capability. Complementing this, the KF mechanism is uniquely employed to fuse multi-dimensional features, reducing computational overhead while accelerating training. Finally, the multi-layer GRU precisely refines the mapping between SOH metrics and predicted values. Experimental validation on the NASA and CALCE datasets demonstrates the superiority of our approach, achieving a MAPE of 0.01, an RMSE of 0.91%, and an R^2 of 0.99. Notably, the method accurately captures phenomena such as battery capacity regeneration, exhibiting superior performance at peaks and valleys while maintaining high computational efficiency.

Keywords: lithium-ion battery; SOH prediction; Kalman filter; sharpness-aware minimization; transformer; GRU

Academic Editor: Federico Baronti

Received: 28 April 2026
Revised: 17 June 2026
Accepted: 30 June 2026
Published: 13 July 2026

1. Introduction

Lithium-ion batteries (LIBs) are widely used in various fields due to their high energy density, low self-discharge rate, and long cycle life [1,2]. During service, LIBs present non-uniform aging behaviors dominated by the physicochemical evolutions of the anode, cathode, and separator, which further determine their state-of-health (SOH) trajectories. Thus, the degradation of LIBs is generally characterized by a reduction in SOH [3]. The complete exhaustion of their remaining useful life may lead to severe consequences, ranging from thermal runaway to catastrophic safety incidents such as fires and explosions [4]. Therefore, the accurate prediction of SOH is a highly prominent topic that has driven

 https://doi.org/10.3390/batteries12070251

extensive research interest [5–7]. SOH is defined as the ratio of the current capacity to the initial capacity [8,9], described as:

$$SOH = C_{cur}/C_{init} \times 100\% \quad (1)$$

where C_{cur} is the current capacity, and C_{init} is the initial capacity of LIBs.

However, accurate prediction of SOH has been a complex task until now because it involves multiple factors such as sensing error, algorithm imperfection, and dataset incompleteness. First, the degradation of LIBs is a complex process, which is readily influenced by factors such as temperature, charge rate, and depth of discharge. Consequently, the SOH profile exhibits highly nonlinear behavior throughout its service life [10,11]. Second, capacity regeneration, a common phenomenon in the LIBs lifecycle, also affects the accuracy and stability of SOH prediction. It stems from internal electrochemical reactions during charge–discharge cycling, which partially restore battery capacity and cause a rapid capacity rise in the subsequent cycle [12]. Given that SOH cannot be measured directly, exhibits time-varying characteristics, and is susceptible to various disturbances [13,14]. High-precision SOH prediction has become a research hotspot in the field of lithium-ion batteries. To this end, a variety of SOH estimation methods have been developed, including model-based, data-driven, and hybrid fusion approaches [15,16].

According to the prediction time horizon and task objective, LIBs state forecasting and health estimation methods can be generally divided into four mainstream categories: nowcasting, very short-term prediction, short-term prediction, and long-term forecasting [17]. Nowcasting aims to estimate a battery's internal states at the current or immediate next sampling instance (milliseconds). To achieve this, model-based methods are widely employed, estimating degradation through physical and mathematical models derived from fundamental electrochemical principles. SOH is estimated by developing representative models, including empirical models [18], equivalent circuit models (ECM) [19–21] and electrochemical models (EM) [22], followed by the identification of their corresponding parameters. Very short-term prediction further enhances model-based methods for regenerative braking and the management of immediate power peaks by employing algorithms such as recursive least squares (RLS) [23] and the extended Kalman filter (EKF) [24,25]. Although SOH prediction has achieved remarkable progress, existing approaches are still hindered by excessive model complexity and poor parameter identifiability, particularly when applied to real-world scenarios.

Recently, data-driven approaches based on machine learning have been widely developed, establishing implicit statistical correlations between battery SOH and key operational variables to enable accurate short-term and long-term health forecasting. Early time series techniques have significantly advanced battery SOH estimation, with representative approaches including Gaussian process regression (GPR) [26], support vector machines (SVM) [27], and artificial neural networks (ANNs) [28]. Time-series analysis within deep learning frameworks is indispensable for accurately estimating the SOH of batteries, especially LIBs powering electric vehicles and energy storage systems. Recurrent neural networks (RNNs) and their improved variants, such as LSTM and GRU, are widely utilized in SOH estimation to extract sequential features while mitigating gradient vanishing or explosion issues [29]. To improve robustness against interference, Zhao et al. [30] utilized time-varying filter-based empirical mode decomposition to extract multi-frequency degradation features, leveraging a Bi–LSTM network for cross-dataset SOH estimation. Sahar et al. [31] analyzed the impact of operating conditions on battery degradation trajectories by combining GRU networks with PCA, K-means clustering, and Granger causality using a realistic forklift dataset. Tang et al. [32] proposed a novel paradigm for accurate and explainable PEMFC health prognostics by integrating electrochemical impedance spectroscopy

(EIS) with a gated recurrent unit (GRU) neural network, which leveraged high-fidelity dynamic data and distribution of relaxation times analysis to achieve highly precise SOH estimation and degradation quantification. While RNN variants demonstrate strong capabilities in SOH estimation, their inherent architecture is prone to losing early-stage sequence dependencies when handling long-term temporal data. In contrast, Transformer-based models [33–35] employ attention mechanisms to assign dynamic weights, which effectively resolve the challenge of modeling long-term temporal dependencies. To substantially improve feature extraction efficiency, Zhu et al. [36] introduced a cross-stage partial probabilistic sparse self-attention module utilizing staged sparsification and multi-scale temporal fusion. To simultaneously capture multi-dimensional dependencies, Zhang et al. [34] developed a Crossformer, which extended the standard Transformer via a cross-patch attention mechanism. Although Transformer variants excel at capturing global long-range dependencies in battery degradation sequences, their prediction performance is highly sensitive to input health feature quality and system complexity.

Physical methods typically utilize multiphysics coupling to simulate internal degradation and quantify battery health. As an early pioneer in physics-based modeling, Tang et al. [37] proposed an electrochemical thermal framework that integrates lithium inventory loss (LLI) and SEI layer growth to forecast LIB capacity degradation. Additionally, Li et al. [38] developed a chemically and mechanically coupled SPM to rapidly predict capacity fade and voltage evolution with an error of approximately 1%. The inherent multiphysics coupling significantly increases model complexity and computational costs, thereby compromising real-time capability and limiting practical online applications. To address the limitations of purely physics-based models, physics-informed neural networks (PINNs) integrating domain knowledge have become an emerging research focus in SOH prediction. Driven by their superior performance, PINNs have been widely applied to LIB SOH estimation [39,40]. To overcome data scarcity, researchers employ physics-informed data augmentation [41], which uses simulation data from multiphysics EMs to characterize aging mechanisms without direct measurements [39–42], thereby enhancing model accuracy and convergence. By fusing battery aging insights with deep learning, Tian et al. [42] constructed health monitoring models that precisely estimate both maximum and remaining capacities. To improve model generalization and interpretability, physics-informed loss functions in PINN-based SOH estimation have gained significant attention. Wang et al. [40] developed an explainability-driven framework that amplifies relevant features while suppressing irrelevant ones during training. Furthermore, Ye et al. [43] exemplified this paradigm by integrating statistical features with physics-constrained neural networks, thereby significantly boosting the interpretability and prediction accuracy of battery states. Lin et al. [44] proposed a lightweight two-stage physics-informed neural network (TSPINN) for SOH estimation of varying lithium-ion chemistries, which integrated physics-informed data augmentation derived from relaxation voltage and incremental capacity features with a mechanism-aware physics-informed loss function to accurately estimate SOH under diverse temperature conditions. Li et al. [45] proposed a unified SOH estimation framework that synergizes a computationally efficient reduced-order electrochemical model and subspace identification for degradation-sensitive feature extraction with Transformer-based learning to capture long-range aging dynamics.

Recent advancements in SOH estimation have focused on enhancing robustness and accuracy under complex operating conditions. To capture the cumulative nature of SOH degradation, hybrid models integrate extracted features to effectively characterize capacity fading trajectories. Wang et al. [46] coupled a GRU with a temporal convolutional attention (TCA) module to adaptively capture degradation dynamics and boost feature extraction. A hybrid framework [47] integrating bidirectional GRU (Bi-GRU) and bidirectional TCN

(Bi-TCN) was proposed to extract robust temporal features for precise SOH prediction. Zhao et al. [30] introduced a signal processing approach utilizing empirical mode decomposition with a time-varying filter coupled with Bi-LSTM to extract multi-frequency degradation features across diverse datasets. Zhou et al. [48] proposed a hybrid framework integrating MR–MLP with an LSTM–Transformer architecture to optimize multi-level feature extraction. Jin et al. [49] proposed a hybrid framework that integrates a conditional generative adversarial network (CGAN) with a Transformer model, thereby improving prediction accuracy and facilitating automatic hyperparameter optimization. Furthermore, Bai et al. [50] introduced the multi-view information perception framework (MVIP-Trans), a convolutional Transformer-based architecture designed to overcome the limitations of pure attention mechanisms in capturing local dynamic features. To address the challenges of real-world SOH estimation, Zhang et al. [51] proposed a novel Bi-GRU-Transformer model that leverages three characteristic parameters extracted from seven hierarchical discharge voltage intervals to reduce feature complexity and significantly improve estimation accuracy. Despite the robust performance of these hybrid models, their representational capacity remains limited in fully capturing the intricate non-linearity and time-varying dynamics inherent in battery degradation.

Based on the above analysis, this paper proposes a SOH prediction method using a fusion neural network, which embodies three major innovations. It integrates the sharpness-aware minimization (SAM) optimized Transformer framework, termed SAMformer [52], a multi-layer stacked GRU [53], and a Kalman filter (KF) for feature optimization of battery health indicators [24,25]. By employing sharpness-aware minimization (SAM), the model is optimized for SOH prediction using arbitrarily segmented charge/discharge data, a strategy that significantly alleviates the data dependency constraints typical of machine learning approaches. Specifically, the channel-attention mechanism embedded within the SAMformer effectively captures inter-channel dependencies while preserving the temporal dynamics of each channel. The SAMformer strategically incorporates reversible instance normalization (RevIN) to mitigate distributional discrepancies between the training and testing sets of battery sequence data. To address severe nonlinearity and high computational complexity, the KF is introduced to suppress noise and eliminate redundancy in the characterization of battery health indicators. The refined features derived from the KF process are subsequently fed into the GRU unit to learn the non-linear mapping relationship between the features and the SOH. The main contributions of this work are summarized as follows:

(1) A novel hybrid neural network, termed the KF–SAMformer–GRU model, is proposed to effectively integrate the coupled intra-variable and inter-variable information associated with degradation patterns, thereby significantly enhancing SOH prediction accuracy and robustness.

(2) Specifically, SAM is employed to optimize the base Transformer network, improving feature extraction capabilities and computational efficiency. To address the inherent non-stationarity of multi-source battery data, we strategically incorporate RevIN, which eliminates dataset uncertainty and substantially enhances model generalization. Additionally, the KF is introduced to filter noise and fuse multi-dimensional features through adaptive weighting. Furthermore, a three-layer stacked GRU network is utilized to output SOH estimates, which effectively capture the phenomenon of battery capacity regeneration.

(3) The proposed model was validated using the NASA and CALCE battery datasets, which encompass diverse battery types and aging trajectories. Experimental results, evaluated using MAPE, RMSE, MAE, and R^2, demonstrate that the proposed method achieves lower error rates and superior stability compared with existing approaches.

2. Related Work

Recent SOH prediction research focuses on accurate model construction, system state parameter prediction, and data architecture design, etc. [8,54,55]. SOH prediction methods can be categorized into model-based, data-driven, and fusion-based methods [6].

The model-based approaches are based on the degradation mechanisms of LIBs. Mathematical and physical models are developed to characterize the LIB degradation process and achieve SOH prediction. The estimation accuracy depends on the decay pattern of the model's key parameters representing the internal aging degree. These methods have several primary forms, such as the electrochemical impedance spectroscopy (EIS) model [56,57], the Thevenin model [58,59], and the multistage resistive circuit (RC) model [60,61]. The accuracy of these methods usually depends on model complexity and parameter identification. Model-based methods have insufficient immunity to interference in practical applications. To date, the principle of the model-based method is simple and easy to operate, but it is sensitive to the accuracy of model construction and parameter identification.

Data-driven methods [62,63] do not need to consider the complex electrochemical changes and fluctuations of correlated substances inside the LIBs. These methods have become a hot spot for research with data mining and extensive data analysis technologies [64,65]. By merging the extracted eigenvalues, parameters, and data mining algorithms, the relationship between the eigen parameters and the SOH is esablished. Battery aging data are essential for building the corresponding models, but they are expensive to collect. Datasets such as NASA and CALCE provide extensive data on the effects of different environmental stresses on battery life, which provide essential help for subsequent scientific development. In general, data-driven methods are divided into data analysis (DA)-based models [66–68] and machine learning (ML)-based models [69–73], which use data-driven SOH estimation methods in different aspects. DA-based models identify features from differential profiles by fitting an analytical function to the estimated data. For this purpose, the electrical, thermal, and mechanical behavior of battery correlations are introduced to draw the SOH. Due to their flexibility, ML-based models are widely used for data-driven SOH prediction of non-approximate matching capability. ML modeling is a specific data analysis methodology that learns from data and makes predictions with minimal human intervention. To date, ML-based models have become the most popular models for data-driven SOH prediction, such as RNN-based models [72–74], Multiple Regression (MR)-based models [75], and Transformer-based models [76]. Specifically, ML-based models aim to extract features affecting battery degradation from direct measurements such as terminal voltage, current, temperature, and recording time. Data-driven methods have strong robustness to overcome the influence of parameter fluctuations, but require massive datasets [5].

Recently, the fusion-based methods have generally caught many researchers' attention, and can be divided into the same-type methods, different-type methods [7,77] and combinations of methods [78]. Furthermore, there are hybrid approaches that integrate optimization algorithms with other techniques. These methods typically yield superior performance through the refinement of model parameters and thresholds [79]. The integration of fusion methods exhibits complementary advantages and is beneficial for addressing challenges such as the difficulty in measuring SOH decline and nonlinear changes. Compared to employing a single data-driven method, a scientifically designed fusion of multiple data-driven approaches can enhance data utilization and prediction accuracy [55]. Yassine [80] developed a combined CNN–LSTM model that achieves highly accurate SOH estimates. The Transformer architecture has also demonstrated effectiveness in tackling long-term SOH prediction challenges, exhibiting satisfactory performance -at capturing long-term dependencies in time-series data. The Bi-GRU–Transformer model boasts higher accuracy, better robustness, and superior generalization capabilities [53]. Lin et al. [44] proposed a lightweight two-stage physics-informed neural network (TSPINN) for SOH estimation of varying lithium-ion chemistries, integrating physics-informed data augmentation derived from relaxation voltage and incremental capacity features with a mechanism-aware loss function to achieve accurate estimation across diverse temperatures. Addressing the complexities of real-world SOH estimation, Zhang et al. [51] proposed a novel Bi-GRU–Transformer model. By leveraging three characteristic parameters extracted from seven hierarchical discharge voltage intervals, this approach effectively reduces feature complexity and significantly enhances estimation accuracy. However, existing SOH prediction methods remain constrained by the stringent data requirements of neural networks, limiting their applicability to batteries with ample data and rendering them less suitable for new batteries [81,82]. The strong nonlinear variations during battery degradation are not adequately captured by end-to-end neural network systems, leading to lower accuracy in predicting capacity regeneration phenomena.

In this paper, we propose a hybrid SOH prediction method based on fusion techniques. This method achieves excellent performance in SOH prediction with reduced prediction errors and effective capture of capacity regeneration. The proposed method demonstrates stable prediction performance across a wide range of battery materials and charging/discharging protocol datasets, effectively addressing the challenge of data distribution shifts encountered [83] in long-term SOH prediction.

3. Approach

3.1. Proposed Method

The overall structure of the KF–SAMformer–GRU network in this paper is shown in Figure 1, which contains three sections: data preprocessing, KF-SAMformer-based feature extraction, and GRU-based anticipation module.

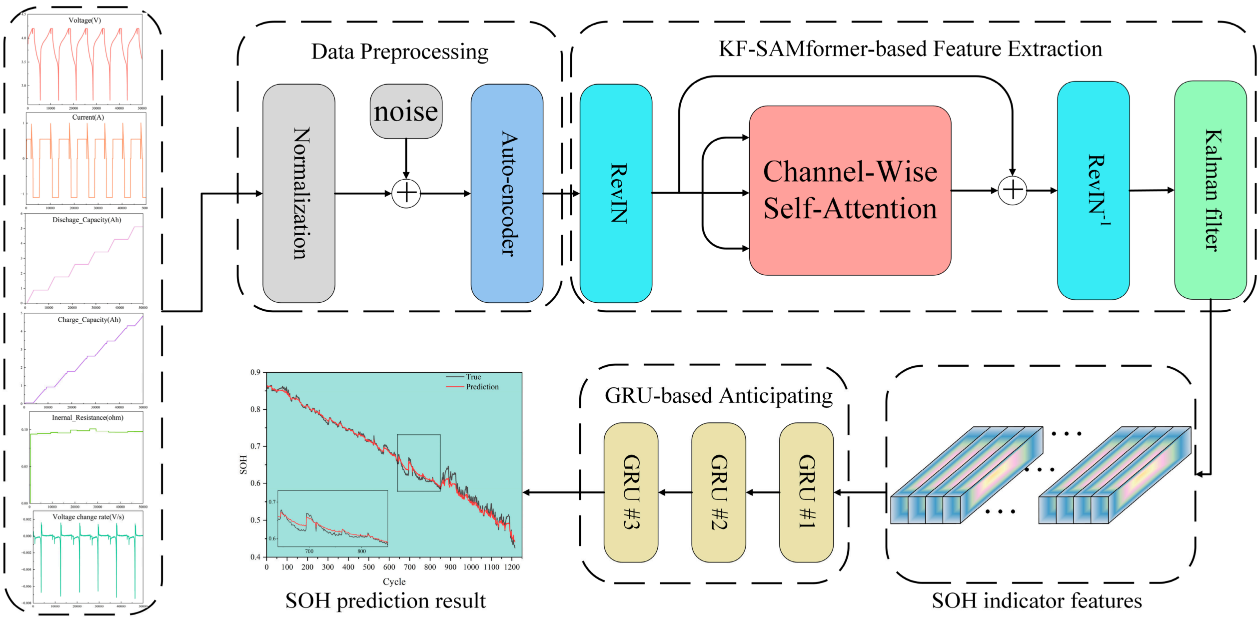

Figure 1. Overall flowchart of the proposed KF–SAMformer–GRU network for SOH estimation.

Data preprocessing: Raw multi-source data frequently suffer from common time-series issues, including missing values and duplicates, as well as inherent problems such as varying scales and inconsistent noise levels. Such challenges can hinder the optimal convergence of predictive models. Therefore, preprocessing the raw data prior to feature extraction is essential to mitigate these drawbacks. This step minimizes the extraction of irrelevant features and spurious dependencies, thereby enhancing both the accuracy and stability of the model while accelerating its convergence. First, to reduce the discrepancy in multi-dimensional data feature sizes and stabilize the training process, the battery sequence data are min–max normalized, followed by the addition of Gaussian white noise. A linear transformation is applied to normalize multi-dimensional data with varying value ranges into the [0, 1] interval.

$$x_i = (x - x_{\min})/(x_{\max} - x_{\min}) \tag{2}$$

where x_i is the normalized value, $x_{\max}$ and $x_{\min}$ correspond to the sample's maximum and minimum values, respectively, while x represents the initial value.

Next, Gaussian white noise is injected into the normalized data to train an autoencoder, which reconstructs the data to preserve as much useful information as possible.

$$x_i' = x_i + N\left(0, \sigma^2\right) \tag{3}$$

Here x_i is the normalized input, and N (0, σ^2) denotes zero-mean Gaussian noise with variance σ^2. $x_i{}'$ represents the noisy input fed into the autoencoder.

This preprocessing step mitigates training fluctuations caused by outliers in the raw data, retains information more relevant for prediction, and enhances the optimization efficiency of subsequent model training. The hidden layer output after processing by the autoencoder is formulated as:

$$x_o = f(kx_i + b) \tag{4}$$

where, x_o denotes the output, f $(\cdot)$ represents the activation function of the encoder, x_i is the input with added Gaussian noise, and b is the bias term.

KF–SAMformer–GRU is introduced to mitigate the instability inherent in battery sequence feature extraction, thereby alleviating the strict data requirements typically associated with data-driven methods. As a variant of the Transformer architecture, SAMformer incorporates state-of-the-art practices from the research community, including RevIN [84] and channel attention. The KF is integrated into SAMformer to mitigate noise and redundancy in both temporal and channel-wise data. This enhancement enables the network to process dynamic battery health information more effectively. A multi-layer GRU module is designed at the bottom of the KF-SAMformer to learn feature representations at varying depths, thereby enhancing the model's expressiveness. The multi-layer GRU structure then processes the temporal dependencies of the input features layer by layer. The KF–SAMformer–GRU process can be summarized as:

$$\widehat{SOH}(k) = \mathcal{G}multi(\mathcal{K}(\mathcal{S})(\mathcal{Z}(\mathcal{N}(\boldsymbol{X}_{\text{raw}})))) \tag{5}$$

Let $\boldsymbol{X}_{\text{raw}}$ denote the raw multi-source time-series input data, including battery voltage, current, internal resistance, and charge/discharge capacity. $\mathcal{N}(\cdot)$ represents the data normalization operation, and $\mathcal{Z}(\cdot)$ indicates the data reconstruction process via a noisy autoencoder. $\mathcal{S}(\cdot)$ corresponds to the feature encoding operation of SAMformer, while $\mathcal{K}(\cdot)$ denotes the feature refinement operation using the Kalman Filter. Furthermore, $\mathcal{G}multi(\cdot)$ represents the multi-layer GRU temporal prediction mapping operation, and $\widehat{\text{SOH}}(k)$ is the predicted value of the battery SOH at each time step.

3.2. Data Preprocessing

The dataset used in this paper contains current, voltage, and capacity information during the LIB charge/discharge cycle, serving as the raw input data for the network presented.

In time-series prediction, dirty data affects the final prediction results as time dependency plays a vital role in processing time-series data. To eliminate the impact of dirty data, preprocessing the raw data is necessary. The data are first normalized by linearly transforming the multi-dimensional data with various value scales to the interval [0, 1]. After adding Gaussian white noise to the normalized data, an autoencoder is generated to reconstruct the data to retain as much useful information as possible. Data preprocessing can reduce the training fluctuations caused by outliers in the original data, retain information that is more suitable for prediction, and improve the efficiency of training optimization for subsequent models. The output of the hidden layer after autoencoder processing is:

$$z = f(wx_t + b) \tag{6}$$

where z denotes the output, $f(\cdot)$ represents the activation function of the encoder, w is the encoder weight matrix, x_t is the input after adding Gaussian noise, and b is the bias.

The reconstructed input matrix is given by the following equation:

$$x'_t = f'(w'z + b') \tag{7}$$

here x'_t is the reconstructed output, $f'(\cdot)$ stands for the mapping function, w' represents the weight of the encoder output, and b' denotes the bias of the encoder output.

3.3. KF-SAMformer

The traditional Transformer structure performs independent encoding and decoding operations for each position in the input sequence and captures the long-term dependencies of the multivariate time-series with the help of a self-attentive mechanism. Despite the high expressive power of the Transformer, the open question of how to effectively train the Transformer architecture without a tendency to overfit remains. It has been shown that by removing the instability of training, Transformers perform well in long-term prediction [52]. Based on the existing research findings, we introduce SAMformer to remove instability during feature extraction of battery sequences, thus mitigating the stringent requirements of ML-based methods on input data. SAMformer, as a variant of the Transformer, incorporates best practices proposed by the research community, including RevIN [84] and channel-wise attention [85,86]. To enhance the generalization performance of the model for the challenges of complex battery materials and charging/discharging strategies, a SAM approach is used to optimize the model training process, minimizing training loss while correlating to the loss sharpness.

The overall architecture of SAMformer is shown in Figure 1, which aims to learn a stable mapping between sequence data and battery health features.

$$f(x'_t) = \left[x'_t + A(x'_t)x'_t w_v w_o\right] w_e \tag{8}$$

with w_v, w_o, w_e, $A(x'_t)$ being the attention matrices of an input sequence x'_t and defined as:

$$A(x'_t) = softmax\left(x'_t w_Q w_K^T {x'_t}^T / \sqrt{d_m}\right) \tag{9}$$

where the softmax is row-wise, w_Q and w_K are the attention matrices, and dm is the dimension of the model.

As a non-stationary dataset, the battery sequence dataset's training and test sets are divided over time, naturally introducing inconsistencies in their distributions. Additionally, there are inconsistencies in data distribution across different battery sequence datasets. Consequently, the battery SOH prediction model experiences inconsistent performance and limited generalization when transitioning from the training set to the test set. SAMformer

https://doi.org/10.3390/batteries12070251

is equipped with RevIN to handle shifts between training and test data in battery time series. Specifically, RevIN contains two parts: normalization and inverse normalization. Instance normalization is performed before inputting the battery data into the model, and the mean and variance are used as the guiding parameters, which are learned by the model and then inverse normalized to obtain the battery health indicator features. The input data is normalized as:

$$E\left[x_t'\right] = \frac{1}{T_x}\sum_{j=1}^{T_x} x_j' \tag{10}$$

$$Var\left[x_t'\right] = \frac{1}{T_x}\sum_{j=1}^{T_x}\left(x_t' - E\left[x_t'\right]\right)^2 \tag{11}$$

$$\hat{x}_t = \lambda\left(\frac{x_t' - E[x_t']}{Var[x_t'] + b''}\right) + \beta \tag{12}$$

where λ and β are learnable affine parameter vectors, $E[x_t']$ is the mean of the input x_t' and Var is the variance, $\hat{x}_t$ is the normalized data. The inverse normalization operation is implemented as follows:

$$\hat{y}_t = \sqrt{Var[x_t'] + b''} \cdot \left({}^{-\beta}\!\big/_{\lambda} \right) + E\left[x_t'\right] \tag{13}$$

where $\hat{y}_t$ is the final prediction of the instance normalization model.

In this paper, the KF is used to enhance SAMformer by reducing redundant noise in both time-dependent and channel-dependent data, allowing the network to better process dynamic battery health information. The KF assists the SOH prediction framework, which is highly nonlinear, in achieving more convergent parameter estimation, thereby improving prediction accuracy. Specifically, the output of SAMformer is fed into the KF, where the filtered battery health state features are weighted differently. This approach retains dependencies that are beneficial for enhancing the prediction accuracy of the proposed network.

3.4. GRU-Based Anticipation

GRU can handle the mapping of battery health state indicator features to SOH values while requiring only a small computational cost. In this paper, we design a multi-layer GRU module at the bottom of the KF-SAMformer to learn feature representations at different depths, thereby improving the model's expressive power. Specifically, the future SOH values are more highly correlated with the observed temporally nearest values; the battery health state indicator features output from the KF-SAMformer are used as the initial hidden states of the GRU, and the multi-layer GRU structure processes the time dependence in the input features layer by layer. The number of GRU iterations follows the discussion in the work [87] and is verified in the ablation experiments. As shown in Figure 1, the GRU is iterated three times, and the hidden state of the third iteration is the predicted target feature, which is input into the classifier to obtain the predicted value of SOH at the target moment. The whole process is as follows:

$$h_0 = X_{t_0};\ h_{i+1} = GRU(h_i),\ i = 1, 2, \cdots \tag{14}$$

where the SOH indicator feature X_{t_0} is used as the initial hidden state of the GRU, and h_{i+1} is the predicted feature of the i + 1-th iteration.

3.5. Loss Function and Model Training Strategy

Traditional optimization algorithms achieve convergence in models by seeking points with low loss within the training dataset. However, many modern neural networks can easily memorize the patterns of the training data, which frequently leads to overfitting when optimized using traditional methods. In this paper, we employ the sharpness-aware minimization (SAM) optimizer instead of conventional optimizers like SGD or Adam to train the neural network framework. By minimizing the sum of the training loss and the sharpness, SAM seeks flat minima in the loss landscape, thereby improving the model's generalization capability. Additionally, cosine annealing is introduced to adjust the learning rate during the training process. The principle of cosine annealing is expressed as:

$$\eta_t = \eta^i_{\min} + \left(\frac{1}{2}\eta^i_{\max} - \eta^i_{\min}\right)\left(1 + \cos\left(\frac{T_{cur}}{T_i}\pi\right)\right) \tag{15}$$

where i stands for the number of restarts, $\eta^i_{\max}$ and $\eta^i_{\min}$ correspond to the maximum and minimum learning rates, respectively. T_{cur} represents the current epoch, and T denotes the total number of epochs for the i-th restart.

The loss function employed in this paper is the mean squared error (MSE). The calculation formula for the loss function is expressed as follows:

$$L_{MSE} = \frac{1}{n}\sum\nolimits_{i=1}^{n} l\left(\widetilde{x}_i - x_{target}\right) \tag{16}$$

here, $l(\cdot)$ denotes the loss function, $\widetilde{x}_i$ and x_{target} represent the predicted and ground truth values, respectively.

Meanwhile, to achieve joint optimization of the encoder and the KF–SAMformer–GRU network during training, the network is trained using the objective function formulated as:

$$L = \phi\sum\nolimits_{i=1}^{n} I(\overline{x}_i - \hat{x}_i)^2 + \varepsilon R(\varphi) \tag{17}$$

where L is the network's objective function, φ serves as a parameter to control the denoising and prediction tasks, ε and R $(\cdot)$ correspond to the regularization parameter and the regularization term, respectively. φ represents the learnable parameters of the model.

4. Experimental Section

4.1. Datasets

The CALCE (CS2), CALCE (CX2), and NASA datasets have been used in the experiment; each dataset contains three types of data: discharge, charging, and impedance test. The technical specifications of the datasets are shown in Table 1.

Table 1. NASA and CALCE Datasets [88,89].

Dataset	NASA	CALCE (CS2)	CALCE (CX2)
Cell Chemistry	LiNiMnCo/Graphite	$LiCoO_2$	$LiCoO_2$
Nominal Voltage	3.7 V	3.7 V	3.7 V
Nominal Capacity	2 Ah	1.1 Ah	1.35 Ah
Operation Temperature	24 °C	1 °C	25, 35, 45, 55 °C
Upper Cut-off Voltage	4.2 V	4.2 V	4.2 V
Lower Cut-off Voltage	2.2 V, 2.5 V, 2.7 V	2.7 V	2.7 V
Charge/Discharge Rate	1.5 A (0.75 C)/2 A (1 C)	1.1 A (1 C)	0.675 A (0.5 C)

To demonstrate the robustness of the network proposed in this paper, three different battery datasets at temperatures of 24 °C, 1 °C, 25, 35, 45, and 55 °C are used. We divide the datasets into training and test sets in a 7:3 ratio. The prediction results show that our proposed method is superior across the three datasets, proving its accuracy and robustness.

4.2. Training Strategy

As is known, the training strategy is an important factor in deep learning, which consists of a loss function, learning rate, optimizer, and so on. The mean square error (MSE) is introduced to evaluate the loss, and the learning rate parameter is configured using the cosine annealing strategy. To cope with multi-dimensional battery sequence data, the SAM algorithm is introduced as the optimizer to replace the traditional Adam and SGD optimizers.

The SAM optimizer [90] is used to optimize the neural network framework while minimizing the total loss and sharp loss to obtain the optimal network training parameters. Cosine annealing is introduced to reduce the learning rate during the model training process. The cosine value first declines slowly, then accelerates the decline, and finally declines slowly again, and this descending can cooperate with the learning rate to produce good results in a very effective way. In gradient descent, the network may fall into a local minimum. Then a high learning rate can be used to jump out of the local minimum and find the path to the global minimum. Cosine annealing is shown in the following equation:

$$\eta_t = \eta_{\min}^i + \left(\frac{1}{2}\eta_{\max}^i - \eta_{\min}^i\right)\left(1 + \cos\left(\frac{T_{cur}}{T_i}\pi\right)\right) \tag{18}$$

where $\eta_{\max}^i$ and $\eta_{\min}^i$ denote the maximum and minimum values of the learning rate, respectively, T_{cur} is the number of epochs currently executed, and T_i can be expressed as the total number of Epochs in the i-time restart.

4.3. Predict Result

To assess the prediction performance, we use four evaluation metrics: the coefficient of determination (R^2), root mean square error (RMSE), mean absolute percentage error (MAPE) and mean absolute error (MAE). These metrics provide a more comprehensive assessment of the network's predictive performance:

$$R^2 = 1 - \left(\sum_{i=1}^{n}(\widetilde{y}_i - y_i)^2\right) / \left(\sum_{i=1}^{n}(y_i - \overline{y})^2\right) \tag{19}$$

$$RMSE = \sqrt{\frac{1}{n}\left(\sum_{i=1}^{n}(\widetilde{y}_i - y_i)^2\right)} \tag{20}$$

$$MAPE = \frac{100\%}{n}\sum_{i=1}^{n}\left|\frac{\widetilde{y}_i - y_i}{y_i}\right| \tag{21}$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|\widetilde{y}_i - y_i| \tag{22}$$

To demonstrate the effectiveness of our method, we compared its results with those of several methods. The networks were trained and tested using the NASA and CALCE datasets. The obtained R^2, RMSE, MAPE, and MAE values are presented in Tables 2 and 3. As shown in Table 2, our method has an average RMSE and MAPE of less than 0.007, an MAE of 0.0051, and an R^2 greater than 0.99, which is superior to traditional methods.

Table 2. Comparison of the method in this paper with other methods (NASA).

Method	LSTM	MLP	GRU	CNN-LSTM	Bi–GRU-Transformer	Ours
R^2	0.8785	0.8047	0.9172	0.9219	0.9629	**0.9921**
RMSE	0.0566	0.0959	0.0388	0.0472	0.0417	**0.0068**
MAPE	0.0628	0.0564	0.0375	0.0358	0.0191	**0.0065**
MAE	0.0490	0.0442	0.0296	0.0273	0.0133	**0.0051**

Table 3. Comparison of the method in this paper with other methods (CALCE).

Method	LSTM	MLP	GRU	CNN-LSTM	Bi–GRU-Transformer	Ours
R^2	0.6914	0.7092	0.8904	0.9719	0.9731	**0.9941**
RMSE	0.0711	0.0749	0.0355	0.0189	0.0297	**0.0091**
MAPE	0.0886	0.0926	0.0468	0.0210	0.0248	**0.0102**
MAE	0.0640	0.0674	0.0320	0.0170	0.0161	**0.0082**

The CNN–LSTM model outperforms the other three single ML-based models in terms of performance metrics. Furthermore, the Bi-GRU–Transformer surpasses the CNN–LSTM, demonstrating superior consistency in SOH prediction on the NASA dataset. Notably, our proposed method achieves significantly lower error values compared to all other models. The prediction curves of our method for the NASA dataset are shown in Figure 2.

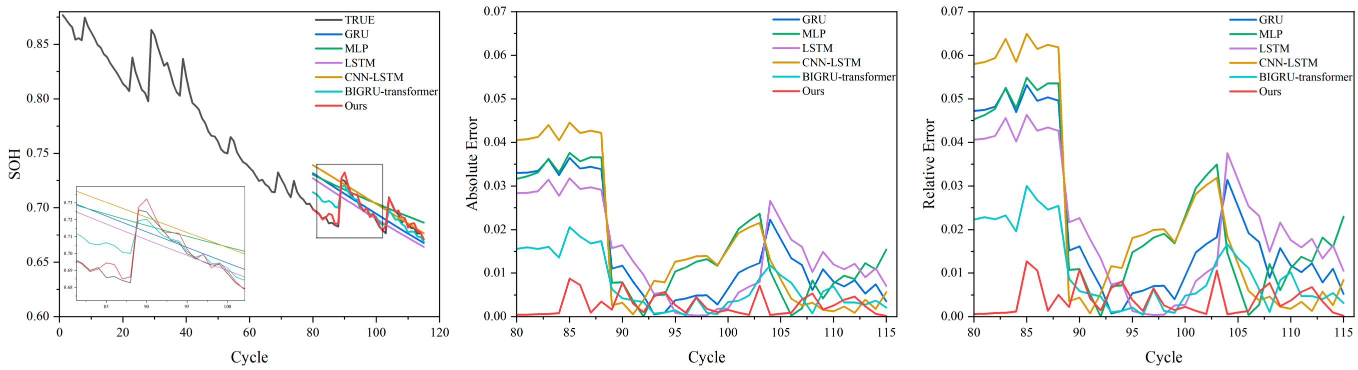

Figure 2. Prediction curves for the NASA 0018 Dataset.

As can be seen from the curves in the figure, our method accurately predicts the SOH of the NASA 0018 dataset during the charge/discharge cycle. Compared to other methods, our proposed method has the smallest absolute error. From the curves, the predicted values of the proposed method are highly consistent with the true values, and it has excellent regeneration capture capability, thereby obtaining accurate prediction results.

Table 3 demonstrates the comparative experimental prediction evaluation metrics for the CALCE dataset. From Table 3, the proposed method outperforms the rest of the comparative methods in the CALCE dataset, with RMSE and MAPE values less than 1.03% and an R^2 greater than 0.99. Test metrics similar to those of the NASA dataset are achieved for the CALCE dataset test, reflecting the stability of the proposed method under different materials, lithium-ion batteries, and charging/discharging protocols.

The prediction curves of our method for the CS2-37 and CX2-34 datasets are shown in Figures 3 and 4. Our prediction curves accurately reflect the trend of SOH curves in the CALCE dataset and capture the dramatic degradation fluctuations of battery health, proving that the prediction accuracy and robustness are satisfactory. As the number of battery charging and discharging cycles increases, the method proposed in this paper always maintains a high degree of consistency with the true values with respect to the trend of SOH changes. The predicted values of the comparison methods in the later stage increasingly deviate from the true values, while the absolute error between the proposed method and the true values is smaller than that of the comparison methods. When capacity regeneration occurs in the battery, this method can sensitively predict the trend of capacity regeneration. This proves the stability and accuracy of this method.

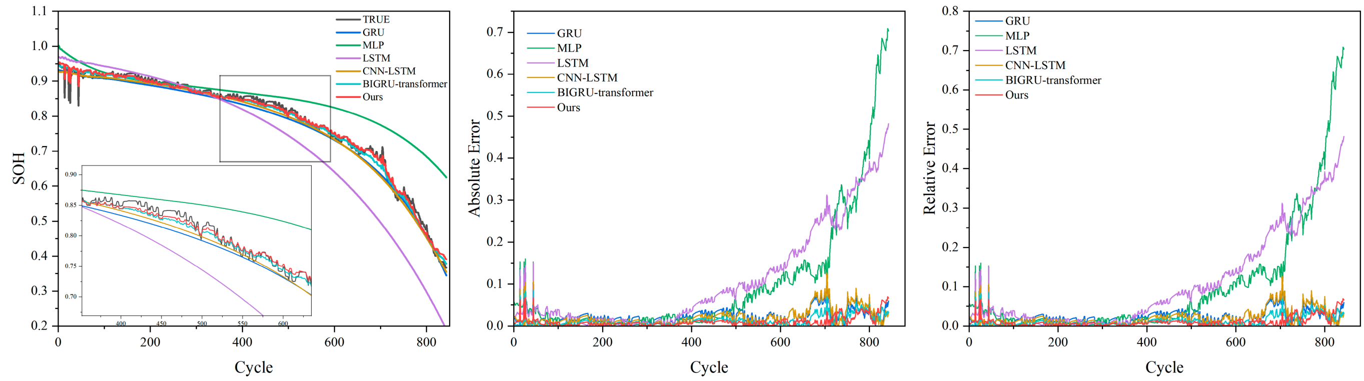

Figure 3. Prediction curves for CS2-37 Dataset.

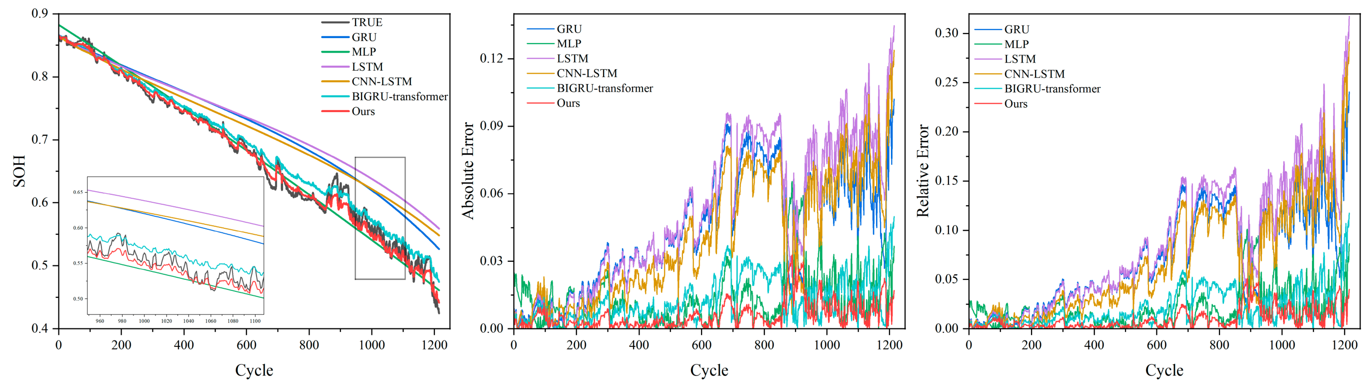

Figure 4. Prediction curves for CX2-34 Dataset.

From the experimental results, many mainstream data-driven methods currently fail to obtain stable results in both datasets. The proposed method in this paper combines SAMformer and the KF to realize a high-quality feature extraction process of battery health indicators, and then utilizes a GRU to solve the mapping of features to SOH values, and finally achieves incredible prediction results on both datasets. The methodology in this paper ensures prediction accuracy with brilliant stability, and still maintains a high level of prediction performance in test sets with significant differences in data features.

4.4. Ablation Experiment

To verify the contribution of the network in this paper and the rationality of the multilayer GRU design, two ablation experiments were executed. The results of the block contribution ablation experiment for the proposed network are shown in Table 4 and Figure 5, and the CALCE dataset was used as the validation.

Table 4. Results of block contribution ablation experiments (CALCE).

Method	Transformer	GRU	KF–Transformer	SAMformer–GRU	Ours
R^2	0.8561	0.8904	0.8952	0.9835	**0.9941**
RMSE	0.0934	0.0355	0.0493	0.0154	**0.0091**
MAPE	0.0732	0.0468	0.0488	0.0193	**0.0102**

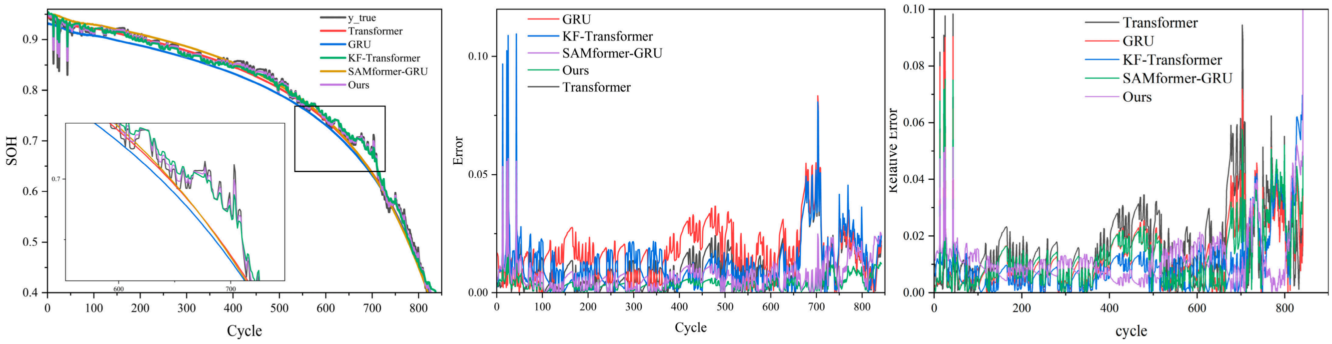

Figure 5. Ablation experiment prediction curves of block contribution.

We conducted an ablation study to evaluate the effectiveness of the KF. The corresponding results are presented in the first part of Section 4.4, as detailed in Table 4 and Figure 5. To examine the effectiveness of our structure, an ablation study was performed to validate the contributions of the SAMformer, KF, and GRU. For a fair comparison, the training settings and datasets were kept identical across all ablation experiments. The ablation experiments were performed on different key modules using the following control groups: (1) Transformer Model only, (2) GRU Model only, (3) KF–Transformer without SAMformer and GRU, and (4) SAMformer–GRU without KF, (5) KF–Transformer–GRU (ours). Transformer refers to the traditional Transformer time-series prediction model; KF–Transformer refers to the model that optimizes the traditional Transformer using the KF; and GRU refers to the gated recurrent unit prediction model. SAMformer–GRU refers to the model that combines the two data-driven methods of SAMformer and GRU, and ours is the method proposed in this paper. From the ablation test, it can be concluded that the method that concluded the fusion of SAMformer and GRU can achieve better performance than the traditional Transformer or GRU; the KF has a significant effect on the optimization of the network. The optimal performance was obtained by the hybrid network proposed in this paper using KF to optimize SAMformer and GRU, which proves the necessity of network structure in this paper.

For the multi-layer GRU design, the basis comes from the GRU layer ablation experiment, which was conducted on the CALCE dataset, and the results are shown in Table 5 and Figure 6. The worst prediction performance was achieved when the network used only one GRU layer, with both RMSE and MAPE at high levels. Given the complex pattern of LIB health-related sequence data and the highly nonlinear nature of the degradation curves, the prediction error was significantly reduced because the multi-layer GRU had a superior pattern-capturing capability compared to a single layer. The highest accuracy in SOH estimation was achieved by a three-layer GRU network. Performance slightly degraded when the number of layers exceeded three, indicating a potential model collapse caused by excessive feature abstraction.

Table 5. Results of GRU layer ablation experiments (CALCE).

Number of Layers	1	2	3	4	5
R^2	0.9565	0.9778	**0.9941**	0.9855	0.9799
RMSE	0.0334	0.0189	**0.0091**	0.0196	0.0206
MAPE	0.0412	0.0200	**0.0102**	0.0175	0.0282

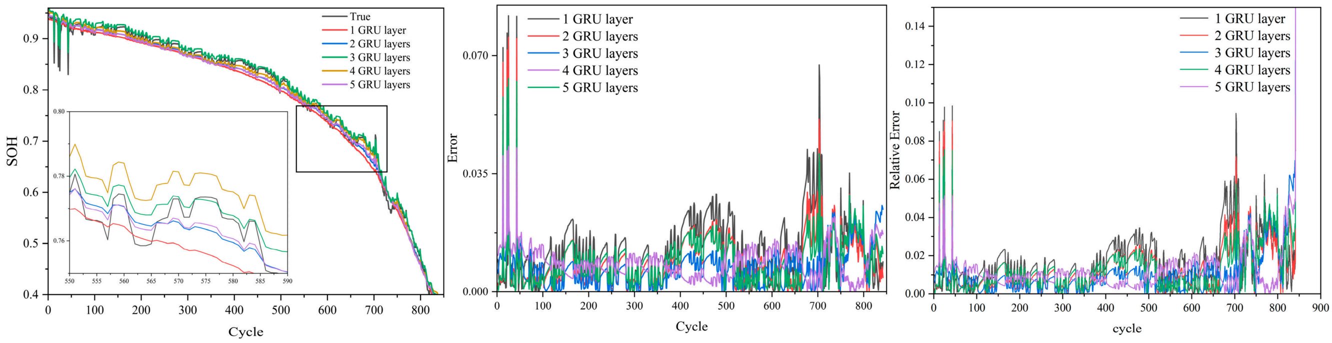

Figure 6. Ablation experiment Prediction curves of GRU layer.

5. Conclusions

Aiming at the SOH prediction issue, a hybrid neural network named KF–SAMformer–GRU is proposed, which captures the phenomenon of battery capacity regeneration and predicts battery capacity recovery. The Kalman filter is introduced to filter out noise and redundant features, and the sharpness-aware minimization-optimized Transformer improves the feature extraction effect and processing speed. Then the multi-layer GRU network is used to output SOH results with a superior pattern-capturing capability.

The proposed method is validated on various public battery datasets with different battery materials and degradation states. The results indicate that our model achieves a MAPE of 0.01, an MAE of 0.0051, an RMSE of 0.91%, and an R^2 of 0.99, outperforming other methods. This demonstrates that our method achieves higher precision in SOH prediction while maintaining excellent stability despite battery data distribution shift challenges. Ablation experiments confirm the rationality of the proposed framework and the number of layers of GRU. The hybrid architecture of KF–SAMformer–GRU significantly improves prediction accuracy compared to other mainstream methods, and the KF further optimizes the training process of the prediction network. Furthermore, the three-layer GRU framework has a superior pattern-capturing capability.

Despite the superior prediction accuracy achieved, our proposed model faces high computational complexity and memory consumption issues when dealing with long sequences of battery data. Despite the promising results, this study currently relies heavily on standardized public datasets (such as CALCE and NASA), which are collected under controlled laboratory conditions. Future research will focus on algorithmic lightweighting and SOH prediction under real charging and discharging conditions, advancing the application of this SOH prediction method in electric vehicles and electrochemical energy storage. Our ultimate goal is to enhance the model's generalization capability to ensure its seamless applicability to real-world battery management systems, facilitating reliable state estimation and health monitoring in practical automotive scenarios.

Author Contributions: L.X. has made contributions in conceptualization, methodology, experiments, initial draft writing, project management, funding acquisition, and revising. P.S. has made contributions to supervision, project management, funding acquisition, and conceptualization. N.Z. has made contributions to supervision, project management and conceptualization. All authors have read and agreed to the published version of the manuscript.

https://doi.org/10.3390/batteries12070251

Funding: This work was supported by the Key Laboratory of Electrochemical Energy Safety, Ministry of Emergency Management, No. EES2025KF09.

Data Availability Statement: The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Conflicts of Interest: Authors Peng Sun are employed by the Key Laboratory of Electrochemical Energy Safety, Ministry of Emergency Management, XYZ Storage Technology Co., Ltd. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

LIBs	lithium-ion batteries
BMS	Battery Management System
EIS	Electrochemical Impedance Spectroscopy
SOH	State of Health
RUL	Remaining Useful Life
KF	Kalman Filter
SAM	Sharpness-Aware Minimization
RevIN	Reversible Instance Normalization
GRU	Gated Recurrent Unit
LSTM	Long Short-Term Memory
MAPE	Mean Absolute Percentage Error
RMSE	Root Mean Square Error
MSE	Mean Square Error

The following Nomenclatures are used in this manuscript:

x_t	the output at moment t
$\hat{x}$	the normalized data of x_t
w	the weight of the auto-encoder output x_t
x'_t	the reconstructed output at moment t
w'	the weight of the auto-encoder output x'_t
w_e, w_v, w_o	attention matrix of an input sequence x'_t
d_m	the dimension of the model
β, λ	learnable affine parameter vectors
X_{t_0}	the initial hidden state of the GRU
h_{i+1}	the predicted feature of the i + 1-th iteration
$\eta^i_{\max}$	the maximum values of the learning rate
$\eta^i_{\min}$	the minimum values of the learning rate

References

1. Li, P.; Luo, S.; Zhang, L.; Liu, Q.; Wang, Y.; Lin, Y.; Xu, C.; Guo, J.; Cheali, P.; Xia, X. Progress, challenges, and prospects of spent lithium-ion batteries recycling: A review. *J. Energy Chem.* **2024**, *89*, 144–171. [CrossRef]
2. Wu, C.; Hu, W.; Meng, J.; Xu, X.; Huang, X.; Cai, L. State-of-charge estimation of lithium-ion batteries based on MCC-AEKF in non-Gaussian noise environment. *Energy* **2023**, *274*, 127316. [CrossRef]
3. Qian, C.; Xu, B.; Xia, Q.; Ren, Y.; Sun, B.; Wang, Z. SOH prediction for Lithium-Ion batteries by using historical state and future load information with an AM-seq2seq model. *Appl. Energy* **2023**, *336*, 120793. [CrossRef]
4. Li, X.; Zhang, L.; Wang, Z.; Dong, P. Remaining useful life prediction for lithium-ion batteries based on a hybrid model combining the long short-term memory and Elman neural networks. *J. Energy Storage* **2019**, *21*, 510–518. [CrossRef]
5. Chen, L.; Xie, S.; Lopes, A.M.; Li, H.; Bao, X.; Zhang, C.; Li, P. A new SOH estimation method for Lithium-ion batteries based on model-data-fusion. *Energy* **2024**, *286*, 129597. [CrossRef]
6. He, Y.; Bai, W.; Wang, L.; Wu, H.; Ding, M. SOH estimation for lithium-ion batteries: An improved GPR optimization method based on the developed feature extraction. *J. Energy Storage* **2024**, *83*, 110678. [CrossRef]

7. Yuan, Z.; Tian, T.; Hao, F.; Li, G.; Tang, R.; Liu, X. A hybrid neural network based on variational mode decomposition denoising for predicting state-of-health of lithium-ion batteries. *J. Power Sources* **2024**, *609*, 234697. [CrossRef]
8. Basia, A.; Simeu-Abazi, Z.; Gascard, E.; Zwolinski, P. Review on State of Health estimation methodologies for lithium-ion batteries in the context of circular economy. *CIRP J. Manuf. Sci. Technol.* **2021**, *32*, 517–528. [CrossRef]
9. Hossain Lipu, M.S.; Hannan, M.A.; Karim, T.F.; Hussain, A.; Saad, M.H.M.; Ayob, A.; Miah, M.S.; Indra Mahlia, T.M. Intelligent algorithms and control strategies for battery management system in electric vehicles: Progress, challenges and future outlook. *J. Clean. Prod.* **2021**, *292*, 126044. [CrossRef]
10. Dong, G.; Wei, J. A physics-based aging model for lithium-ion battery with coupled chemical/mechanical degradation mechanisms. *Electrochim. Acta* **2021**, *395*, 139133. [CrossRef]
11. Luo, F.; Huang, H.; Ni, L.; Li, T. Rapid prediction of the state of health of retired power batteries based on electrochemical impedance spectroscopy. *J. Energy Storage* **2021**, *41*, 102866. [CrossRef]
12. Huang, Y.; Zhang, P.; Lu, J.; Xiong, R.; Cai, Z. A transferable long-term lithium-ion battery aging trajectory prediction model considering internal resistance and capacity regeneration phenomenon. *Appl. Energy* **2024**, *360*, 122825. [CrossRef]
13. Liu, Z.; Zhao, J.; Wang, H.; Yang, C. A New Lithium-Ion Battery SOH Estimation Method Based on an Indirect Enhanced Health Indicator and Support Vector Regression in PHMs. *Energies* **2020**, *13*, 830. [CrossRef]
14. Hu, X.; Xu, L.; Lin, X.; Pecht, M. Battery Lifetime Prognostics. *Joule* **2020**, *4*, 310–346. [CrossRef]
15. Lu, J.; Xiong, R.; Tian, J.; Wang, C.; Sun, F. Deep learning to estimate lithium-ion battery state of health without additional degradation experiments. *Nat. Commun.* **2023**, *14*, 2760. [CrossRef] [PubMed]
16. Fu, J.; Wu, C.; Wang, J.; Haque, M.M.; Geng, L.; Meng, J. Lithium-ion battery SOH prediction based on VMD-PE and improved DBO optimized temporal convolutional network model. *J. Energy Storage* **2024**, *87*, 111392. [CrossRef]
17. von Bülow, F.; Meisen, T. A review on methods for state of health forecasting of lithium-ion batteries applicable in real-world operational conditions. *J. Energy Storage* **2023**, *57*, 105978. [CrossRef]
18. Li, R.; Bao, L.; Chen, L.; Zha, C.; Dong, J.; Qi, N.; Tang, R.; Lu, Y.; Wang, M.; Huang, R.; et al. Accelerated aging of lithium-ion batteries: Bridging battery aging analysis and operational lifetime prediction. *Sci. Bull.* **2023**, *68*, 3055–3079. [CrossRef] [PubMed]
19. Li, C.; Yang, L.; Li, Q.; Zhang, Q.; Zhou, Z.; Meng, Y.; Zhao, X.; Wang, L.; Zhang, S.; Li, Y.; et al. SOH estimation method for lithium-ion batteries based on an improved equivalent circuit model via electrochemical impedance spectroscopy. *J. Energy Storage* **2024**, *86*, 111167. [CrossRef]
20. Biju, N.; Fang, H. BattX: An equivalent circuit model for lithium-ion batteries over broad current ranges. *Appl. Energy* **2023**, *339*, 120905. [CrossRef]
21. Navas, S.J.; González, G.M.C.; Pino, F.J.; Guerra, J.J. Modelling Li-ion batteries using equivalent circuits for renewable energy applications. *Energy Rep.* **2023**, *9*, 4456–4465. [CrossRef]
22. Gao, Y.; Liu, K.; Zhu, C.; Zhang, X.; Zhang, D. Co-Estimation of State-of-Charge and State-of- Health for Lithium-Ion Batteries Using an Enhanced Electrochemical Model. *IEEE Trans. Ind. Electron.* **2022**, *69*, 2684–2696. [CrossRef]
23. Kang, E.; Lee, S.; Song, M.; Lee, J.; Song, J.; Kim, J. Advanced state-of-health estimation integrating partial capacity-based initialization and resistance-informed covariance correction in adaptive extended Kalman filter for lithium iron phosphate batteries. *Appl. Energy* **2026**, *419*, 128096. [CrossRef]
24. Li, M.; Li, C.; Zhang, Q.; Liao, W.; Rao, Z. State of charge estimation of Li-ion batteries based on deep learning methods and particle-swarm-optimized Kalman filter. *J. Energy Storage* **2023**, *64*, 107191. [CrossRef]
25. Fahmy, H.M.; Hasanien, H.M.; Alsaleh, I.; Ji, H.; Alassaf, A. State of health estimation of lithium-ion battery using dual adaptive unscented Kalman filter and Coulomb counting approach. *J. Energy Storage* **2024**, *88*, 111557. [CrossRef]
26. Goh, H.H.; Lan, Z.; Zhang, D.; Dai, W.; Kurniawan, T.A.; Goh, K.C. Estimation of the state of health (SOH) of batteries using discrete curvature feature extraction. *J. Energy Storage* **2022**, *50*, 104646. [CrossRef]
27. Chen, Z.; Sun, M.; Shu, X.; Xiao, R.; Shen, J. Online State of Health Estimation for Lithium-Ion Batteries Based on Support Vector Machine. *Appl. Sci.* **2018**, *8*, 925. [CrossRef]
28. Bhagavatula, S.; Yellamraju, R.; Eltem, K.; Bobba, P.; Marati, N. ANN based Battery Health Monitoring—A Comprehensive Review. *E3S Web Conf.* **2020**, *184*, 01068. [CrossRef]
29. Chen, H.; Chen, Y.; Sun, C.; Huo, L.; Zhang, W.; Shen, P.; Huang, L.; Luan, W.; Chen, H. Towards practical data-driven battery state of health estimation: Advancements and insights targeting real-world data. *J. Energy Chem.* **2025**, *110*, 657–680. [CrossRef]
30. Zhao, J.; Guo, W.; Pan, H.; Gao, Q.; Shi, P.; Min, Y. Lithium-ion battery state-of-health estimation based on TVFEMD and BiLSTM-Attention. *Energy* **2025**, *332*, 137172. [CrossRef]
31. Qaadan, S.; Alshare, A.; Popp, A.; Schmuelling, B. Prediction of Lithium-Ion Battery Health Using GRU-BPP. *Batteries* **2024**, *10*, 399. [CrossRef]
32. Tang, X.; Meng, X.; Liu, Z.; Miao, Y.; Hasanien, H.M.; Li, M.; Liu, J.; Sun, C. State of Health Estimation and Degradation Quantification of PEMFCs Using a Dynamic Coupling Method of Data-driven and Electrochemical Impedance. *IEEE Trans. Transp. Electrif.* **2026**. [CrossRef]

33. Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A.N.; Kaiser, L.U.; Polosukhin, I. Attention is All you Need. *Adv. Neural Inf. Process. Syst.* **2017**, *30*, 5998–6008.
34. Zhang, Y.; Yan, J. Crossformer: Transformer Utilizing Cross-Dimension Dependency for Multivariate Time Series Forecasting. In Proceedings of the International Conference on Learning Representations, Kigali, Rwanda, 1–5 May 2023.
35. Wu, H.; Xu, J.; Wang, J.; Long, M. Autoformer: Decomposition Transformers with Auto-Correlation for Long-Term Series Forecasting. *Adv. Neural Inf. Process. Syst.* **2021**, *34*, 22419–22430.
36. Zhu, X.; Xu, C.; Song, T.; Huang, Z.; Zhang, Y. Sparse Self-Attentive Transformer With Multiscale Feature Fusion on Long-Term SOH Forecasting. *IEEE Trans. Power Electron.* **2024**, *39*, 10399–10408. [CrossRef]
37. Tang, Y.; Jia, M.; Ai, L.; Yin, B.; Cheng, Y.; Liu, Y. Capacity Fade Analysis of the Lithium-Ion Power Battery Cycling Process Based on an Electrochemical-Thermal Coupling Model. *Energy Technol.* **2015**, *3*, 1250–1259. [CrossRef]
38. Li, J.; Adewuyi, K.; Lotfi, N.; Landers, R.G.; Park, J. A single particle model with chemical/mechanical degradation physics for lithium ion battery State of Health (SOH) estimation. *Appl. Energy* **2018**, *212*, 1178–1190. [CrossRef]
39. Wang, F.; Zhai, Z.; Zhao, Z.; Di, Y.; Chen, X. Physics-informed neural network for lithium-ion battery degradation stable modeling and prognosis. *Nat. Commun.* **2024**, *15*, 4332. [CrossRef] [PubMed]
40. Wang, F.; Zhao, Z.; Zhai, Z.; Shang, Z.; Yan, R.; Chen, X. Explainability-driven model improvement for SOH estimation of lithium-ion battery. *Reliab. Eng. Syst. Saf.* **2023**, *232*, 109046. [CrossRef]
41. Hervé de Beaulieu, M.; Jha, M.S.; Garnier, H.; Cerbah, F. Remaining Useful Life prediction based on physics-informed data augmentation. *Reliab. Eng. Syst. Saf.* **2024**, *252*, 110451. [CrossRef]
42. Tian, J.; Ma, L.; Zhang, T.; Han, T.; Mai, W.; Chung, C.Y. Exploiting domain knowledge to reduce data requirements for battery health monitoring. *Energy Storage Mater.* **2024**, *67*, 103270. [CrossRef]
43. Ye, J.; Xie, Q.; Lin, M.; Wu, J. A method for estimating the state of health of lithium-ion batteries based on physics-informed neural network. *Energy* **2024**, *294*, 130828. [CrossRef]
44. Lin, C.; Wu, L.; Tuo, X.; Liu, C.; Zhang, W.; Huang, Z.; Zhang, G. A lightweight two-stage physics-informed neural network for SOH estimation of lithium-ion batteries with different chemistries. *J. Energy Chem.* **2025**, *105*, 261–279. [CrossRef]
45. Li, Y.; Wang, H.; Wang, C.; Wang, L.; Liao, C.; Wang, L. Unified physics-informed subspace identification and transformer learning for lithium-ion battery state-of-health estimation. *J. Energy Chem.* **2026**, *112*, 350–369. [CrossRef]
46. Wang, X.; Dai, K.; Hu, M.; Ni, N. Lithium-ion battery health state and remaining useful life prediction based on hybrid model MFE-GRU-TCA. *J. Energy Storage* **2024**, *95*, 112442. [CrossRef]
47. Quan, R.; Zhou, Y.; Li, W.; Wan, H. A hybrid data-driven approach for SOH estimation of lithium-ion batteries based on health index extraction and QHDBO-BiTCN-BiGRU. *J. Energy Storage* **2025**, *125*, 116788. [CrossRef]
48. Zhou, J.; Rong, J.; Zhang, J.; Liu, C.; Yi, F.; Jiao, Z.; Zhang, C. Deep learning estimation of state of health for lithium-ion batteries using multi-level fusion features of discharge curves. *J. Power Sources* **2025**, *653*, 237781. [CrossRef]
49. Jin, H.; Ru, R.; Cai, L.; Meng, J.; Wang, B.; Peng, J.; Yang, S. A synthetic data generation method and evolutionary transformer model for degradation trajectory prediction in lithium-ion batteries. *Appl. Energy* **2025**, *377*, 124629. [CrossRef]
50. Bai, T.; Wang, H. Convolutional Transformer-Based Multiview Information Perception Framework for Lithium-Ion Battery State-of-Health Estimation. *IEEE Trans. Instrum. Meas.* **2023**, *72*, 2523312. [CrossRef]
51. Zhang, K.; Jiao, Z.; Ma, J.; Zhao, X.; Wang, S.; Yu, M.; Li, S. SOH estimation of battery based on the BIGRU-transformer algorithm considering the hierarchical voltage interval characterization of the discharge curve. *J. Energy Storage* **2025**, *132*, 117770. [CrossRef]
52. Ilbert, R.; Odonnat, A.; Feofanov, V.; Virmaux, A.; Paolo, G.; Palpanas, T.; Redko, I. SAMformer: Unlocking the Potential of Transformers in Time Series Forecasting with Sharpness-Aware Minimization and Channel-Wise Attention. 2024.
53. Jia, C.; Tian, Y.; Shi, Y.; Jia, J.; Wen, J.; Zeng, J. State of health prediction of lithium-ion batteries based on bidirectional gated recurrent unit and transformer. *Energy* **2023**, *285*, 129401. [CrossRef]
54. Mayemba, Q.; Mingant, R.; Li, A.; Ducret, G.; Venet, P. Aging datasets of commercial lithium-ion batteries: A review. *J. Energy Storage* **2024**, *83*, 110560. [CrossRef]
55. Tian, H.; Qin, P.; Li, K.; Zhao, Z. A review of the state of health for lithium-ion batteries: Research status and suggestions. *J. Clean. Prod.* **2020**, *261*, 120813. [CrossRef]
56. Huang, J.; Li, Z.; Zhang, J. Dynamic electrochemical impedance spectroscopy reconstructed from continuous impedance measurement of single frequency during charging/discharging. *J. Power Sources* **2015**, *273*, 1098–1102. [CrossRef]
57. Lvovich, V.; Wu, J.; Bennett, W.; Demattia, B.; Miller, T. Applications of AC Impedance Spectroscopy as Characterization and Diagnostic Tool in Li-Metal Battery Cells. *ECS Trans.* **2014**, *58*, 1. [CrossRef]
58. Tran, M.-K.; Mathew, M.; Janhunen, S.; Panchal, S.; Raahemifar, K.; Fraser, R.; Fowler, M. A comprehensive equivalent circuit model for lithium-ion batteries, incorporating the effects of state of health, state of charge, and temperature on model parameters. *J. Energy Storage* **2021**, *43*, 103252. [CrossRef]
59. Ding, X.; Zhang, D.; Cheng, J.; Wang, B.; Luk, P.C.K. An improved Thevenin model of lithium-ion battery with high accuracy for electric vehicles. *Appl. Energy* **2019**, *254*, 113615. [CrossRef]

60. Zhang, X.; Zhang, W.; Lei, G. A Review of Li-ion Battery Equivalent Circuit Models. *Trans. Electr. Electron. Mater.* **2016**, *17*, 311–316. [CrossRef]
61. Plett, G.L. High-Performance Battery-Pack Power Estimation Using a Dynamic Cell Model. *IEEE Trans. Veh. Technol.* **2004**, *53*, 1586–1593. [CrossRef]
62. Huang, Y.F.; Gong, X.J.; Lin, Z.Y.; Xu, L. A hybrid data-driven method for voltage state prediction and fault warning of Li-ion batteries. *Case Stud. Therm. Eng.* **2024**, *64*, 105420. [CrossRef]
63. Huang, Y.F.; Sun, J.; Xu, L. A multi-variable multi-step Seq2seq networks for the state of charge estimation of lithium-ion battery. *Case Stud. Therm. Eng.* **2023**, *51*, 103566. [CrossRef]
64. Chen, X.; Huang, Y.F. Memory-Oriented Unpaired Learning for Single Remote Sensing Image Dehazing. *IEEE Geosci. Remote Sens. Lett.* **2022**, *19*, 3511705. [CrossRef]
65. Huang, Y.F.; Lin, Z.Y.; Xiong, S.; Sun, T.T. Diffusion Models Based Null-Space Learning for Remote Sensing Image Dehazing. *IEEE Geosci. Remote Sens. Lett.* **2024**, *21*, 8001305. [CrossRef]
66. Li, Y.; Liu, K.; Foley, A.M.; Zülke, A.; Berecibar, M.; Nanini-Maury, E.; Van Mierlo, J.; Hoster, H.E. Data-driven health estimation and lifetime prediction of lithium-ion batteries: A review. *Renew. Sustain. Energy Rev.* **2019**, *113*, 109254. [CrossRef]
67. Vanem, E.; Salucci, C.B.; Bakdi, A.; Alnes, Ø.Å.s. Data-driven state of health modelling—A review of state of the art and reflections on applications for maritime battery systems. *J. Energy Storage* **2021**, *43*, 103158. [CrossRef]
68. Yang, D.; Lu, J.; Dong, H.; Hu, Z. Pipeline signal feature extraction method based on multi-feature entropy fusion and local linear embedding. *Syst. Sci. Control Eng.* **2022**, *10*, 407–416. [CrossRef]
69. Liang, C.-M.; Li, Y.-W.; Liu, Y.-H.; Wen, P.-F.; Yang, H. Segmentation and weight prediction of grape ear based on SFNet-ResNet18. *Syst. Sci. Control Eng.* **2022**, *10*, 722–732. [CrossRef]
70. Li, X.; Feng, S.; Hou, N.; Wang, R.; Li, H.; Gao, M.; Li, S. Surface microseismic data denoising based on sparse autoencoder and Kalman filter. *Syst. Sci. Control Eng.* **2022**, *10*, 616–628. [CrossRef]
71. Bamati, S.; Chaoui, H. Lithium-Ion Batteries Long Horizon Health Prognostic Using Machine Learning. *IEEE Trans. Energy Convers.* **2022**, *37*, 1176–1186. [CrossRef]
72. Sun, H.; Sun, J.; Zhao, K.; Wang, L.; Wang, K. Data-Driven ICA-Bi-LSTM-Combined Lithium Battery SOH Estimation. *Math. Probl. Eng.* **2022**, *2022*, 1–8. [CrossRef]
73. 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]
74. Li, J.; Jin, B.; Chang, J.; Tao, Z.; Liu, Z. Predicting the State of Health of Lithium Battery Based on Improved LSTM-RNN. In Proceedings of the 2023 International Symposium on Intelligent Robotics and Systems (ISoIRS), Changsha, China, 26–28 May 2023; pp. 52–55.
75. Wang, Z.; Ma, J.; Zhang, L. State-of-Health Estimation for Lithium-Ion Batteries Based on the Multi-Island Genetic Algorithm and the Gaussian Process Regression. *IEEE Access* **2017**, *5*, 21286–21295. [CrossRef]
76. Gu, X.; See, K.W.; Li, P.; Shan, K.; Wang, Y.; Zhao, L.; Lim, K.C.; Zhang, N. A novel state-of-health estimation for the lithium-ion battery using a convolutional neural network and transformer model. *Energy* **2023**, *262*, 125501. [CrossRef]
77. He, T.; Gong, Z. State of health estimation for lithium-ion batteries using a hybrid neural network model with Multi-scale Convolutional Attention Mechanism. *J. Power Sources* **2024**, *609*, 234680. [CrossRef]
78. Sun, H.; Yang, D.; Wang, L.; Wang, K. A method for estimating the aging state of lithium-ion batteries based on a multi-linear integrated model. *Int. J. Energy Res.* **2022**, *46*, 24091–24104. [CrossRef]
79. Ma, Y.; Li, J.; Gao, J.; Chen, H. State of health prediction of lithium-ion batteries under early partial data based on IWOA-BiLSTM with single feature. *Energy* **2024**, *295*, 131085. [CrossRef]
80. Toughzaoui, Y.; Toosi, S.B.; Chaoui, H.; Louahlia, H.; Petrone, R.; Le Masson, S.; Gualous, H. State of health estimation and remaining useful life assessment of lithium-ion batteries: A comparative study. *J. Energy Storage* **2022**, *51*, 104520. [CrossRef]
81. Huang, Y.F.; Xiong, S. Remote Sensing Image Dehazing Using Adaptive Region-Based Diffusion Models. *IEEE Geosci. Remote Sens. Lett.* **2023**, *20*, 8001805. [CrossRef]
82. Li, Z.Z.; Huang, Y.F.; Li, X.X.; Xu, L. Wildland Fire Burned Areas Prediction Using Long Short-Term Memory Neural Network with Attention Mechanism. *Fire Technol.* **2021**, *57*, 1–23. [CrossRef]
83. Kulinski, S.; Inouye, D.I. Towards Explaining Distribution Shifts. *arXiv* **2022**, arXiv:2210.10275.
84. Kim, T.; Kim, J.; Tae, Y.; Park, C.; Choi, J.; Choo, J. Reversible Instance Normalization for Accurate Time-Series Forecasting against Distribution Shift. In Proceedings of the International Conference on Learning Representations, Virtual Event, 25–29 April 2022.
85. Zamir, S.W.; Arora, A.; Khan, S.H.; Hayat, M.; Khan, F.S.; Yang, M.-H. Restormer: Efficient Transformer for High-Resolution Image Restoration. In Proceedings of the 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), New Orleans, LA, USA, 18–24 June 2022; pp. 5718–5729.

86. Zhang, H.; Wu, C.; Zhang, Z.; Zhu, Y.; Zhang, Z.-L.; Lin, H.; Sun, Y.; He, T.; Mueller, J.W.; Manmatha, R.; et al. ResNeSt: Split-Attention Networks. In Proceedings of the 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), New Orleans, LA, USA, 19–20 June 2020; pp. 2735–2745.
87. Karpathy, A.; Johnson, J.; Fei-Fei, L. Visualizing and Understanding Recurrent Networks. *arXiv* **2015**, arXiv:1506.02078.
88. Cheng, G.; Wang, X.; He, Y. Remaining useful life and state of health prediction for lithium batteries based on empirical mode decomposition and a long and short memory neural network. *Energy* **2021**, *232*, 121022. [CrossRef]
89. Jianing, X.; Yulong, N.; Chunbo, Z. Remaining useful life prediction for lithium-ion batteries based on improved support vector regression. *Trans. China Electrotech. Soc.* **2021**, *36*, 3693–3704.
90. Foret, P.; Kleiner, A.; Mobahi, H.; Neyshabur, B. Sharpness-Aware Minimization for Efficiently Improving Generalization. *arXiv* **2020**, arXiv:2010.01412.