Battery Prognostics and Health Management Using CNN-BiGRU with Temporal Attention on CS2 Cells

Amrit Dhakal
School of Civil Aviation
Northwestern Polytechnical University
Xi'an, China
amritdhakal@mail.nwpu.edu.cn

Yu Wu*
School of Civil Aviation
Northwestern Polytechnical University
Xi'an, China
yu.wu@nwpu.edu.cn

Umar Saleem
School of Civil Aviation
Northwestern Polytechnical University
Xi'an, China
engrumarsaleem15@nwpu.edu.cn

***Abstract*—Lithium-ion batteries (LiBs) are reliable and efficient energy sources; they are extensively employed worldwide in modern energy storage systems. The State of Health (SOH) and Remaining Useful Life (RUL) are considered some of the critical parameters for monitoring battery performance and safety. SOH and RUL reflect the battery's overall condition, functionality, and remaining lifespan. Their accurate estimation is essential for ensuring safe operation, effective energy management, and extension of battery's service life. However, to meet the demands of real-world applications, the accuracy of SOH and RUL estimation must be further improved. This research work proposes a Deep Learning (DL) based method that combines Convolutional Neural Network (CNN), Bidirectional Gated Recurrent Unit (BiGRU), and temporal attention for the SOH and RUL estimation of LiBs. The proposed architecture was tested through leave-one-out cross-validation (LOOCV). It achieved a Mean Absolute Error (MAE) of 0.0072, a Root Mean Squared Error (RMSE) of 0.0121, and a coefficient of determination (R^2) of 0.9965 on group I of the Center for Advanced Life Cycle Engineering (CALCE) CS2-type LiBs dataset. Compared prediction results with other DL based models, including CNN, Long Short-Term Memory (LSTM), CNN-Bidirectional LSTM (CNN-BiLSTM), and BiLSTM-Gated Recurrent Unit (BiLSTM-GRU), showing that our proposed method performed better in term of SOH and RUL estimation and also providing better solution for battery management system (BMS).**

Keywords-CNN; BiGRU; Temporal Attention; Lithium-ion battery; SOH; RUL

I. Introduction

In modern energy infrastructures, LiBs are indispensable due to their high energy density, long cycle life, and superior efficiency [1]. However, the concerns limiting the widespread use of LiBs are safety, health monitoring, and RUL prediction. BMS is dependent on accurate and real-time monitoring of the battery's internal condition. Among these, SOH and RUL are critical indicators for prognostics, such as accessing performance, planning maintenance , and ensuring operational reliability [2] [3]. SOH estimation reflects capacity fade and performance degradation, and RUL estimation supports long-term operational planning. Together, SOH and RUL form the backbone of intelligent BMS [4]. The methods for estimating SOH and RUL can be generally categorized into model-based and data-driven based [5]. Although model-based methods like resistor-capacitor circuit-based models, physics-based models, and direct measurement provide insightful information about what truly occurs inside the cell, these methods require researchers to have a deeper knowledge of the underlying physics, chemistry and numerous advanced experiments of battery in controlled lab conditions for identification of it's internal parameters. Moreover, these models are highly sensitive to environmental conditions such as temperature and humidity, which creates more difficulties in modeling, simulation as well as in real-world application [7] [8]. DL-based methods increase BMS prediction accuracy and effectiveness because they can handle large, complex data for accurate estimation of SOH and RUL [9] [10] [11]. Machine learning and DL offer greater adaptability by learning patterns from historical and real-time data without the need for detailed physical modeling [12]. Besides these benefits of using DL models, these methods are famous for their superior capability to learn non-linear patterns even in complex time series data [13] [14]. LSTM was applied in [15] for RUL prediction of LiBs. The authors used the resilient mean square back-propagation technique to train the LSTM neural networks, which lacked strong feature extraction. Similarly, in [16], CNN was used for RUL estimation on the National Aeronautics and Space Administration (NASA) dataset. Using a fixed-length sliding window, the authors also used auto-CNN to predict capacity deterioration. However, because the length of the window is not aligned with degradation characteristics during various battery life cycles, the actual capacity is either over- or under-predicted. A GRU-based SOH prediction was carried out in [17] but lacked the strong feature extraction technique. GRU combined with random forest was used to estimate SOH of the NASA battery dataset in [18], which also lacked the strong automatic feature extraction. A CNN-GRU framework was used in [19] to automatically extract features from voltage, current, and temperature data segments during charging. This eliminated the need for manual feature engineering, enabling a more adaptive and accurate SOH estimation. However, this framework lacked focus on more important time steps. Particle filter with LSTM was used in [20] for RUL prediction, which showed promising results, but it is computationally expensive. Similarly, mixers-bidirectional temporal CNN-based DL was used in [21], which lacked long-term feature extraction. A transformer-based multi-head attention was proposed in [22],

which has significant improvement in RUL prediction but increases the complexity of the model. To improve the SOH and RUL prediction accuracy and decrease computational complexity, we propose a novel method called CNN-BiGRU with a temporal attention. The main contributions of this work are as follows.

1) Application of deep neural network architecture through integration of CNN-BiGRU and temporal attention for estimating SOH and RUL.
2) The proposed method CNN–BiGRU-Temporal Attention was rigorously compared with other established DL architectures such as CNN, LSTM, CNN-BiLSTM, and BiLSTM-GRU.
3) The proposed model, along with other benchmark models, was evaluated using various metrics such as MAE, RMSE, and R^2, showing that our model has better results and contributes to developing reliable and safer BMS.

II. Dataset and Evaluation Criteria

A. CALCE dataset

In this research, we use the CALCE CS2 battery dataset. Four CS2 prismatic cells went through the same charging and discharging profile. Constant current, constant voltage (CCCV) was employed for charging and discharging at various temperatures. Additional technical information is provided in Table I.

TABLE I: CALCE Dataset Parameters for CS2 Battery Cells.

Parameters	Values
Dataset	CALCE CS2
Cells	CS2_35, CS2_36, CS2_37, CS2_38
Cell Chemistry	$LiCoO_2$cathode
Dimensions	$5.4 \times 33.6 \times 50.6$ mm
Nominal Voltage	4.2 V
Energy Capacity	1.1 Ah
Charging Current	0.55 A
Discharging Current	1.1 A
Voltage Range	4.2 V to 2.7 V (discharge)

B. Battery SOH and RUL

Over the lifetime of the battery, it undergoes charging and discharging cycles, which results in gradual degradation of its performance. This degradation in performance over time is called aging, which decreases the SOH of the battery. Capacity degradation of different cells can be seen in Fig. 1. We calculate SOH in its capacity form defined in (1). Similarly, true and predicted RUL are shown in (2) and (3), respectively.

$$\mathrm{SOH} = \frac{C_{\text{current}}}{C_{\text{rated}}} \times 100\%, \tag{1}$$

where C_{current} is the measured capacity of the battery and C_{rated} is the rated or initial capacity.

$$\mathrm{RUL}_{true} = N_{\text{EOL}}^{true} - N_{\text{current}}, \tag{2}$$

$$\mathrm{RUL}_{pred} = N_{\text{EOL}}^{pred} - N_{\text{current}}, \tag{3}$$

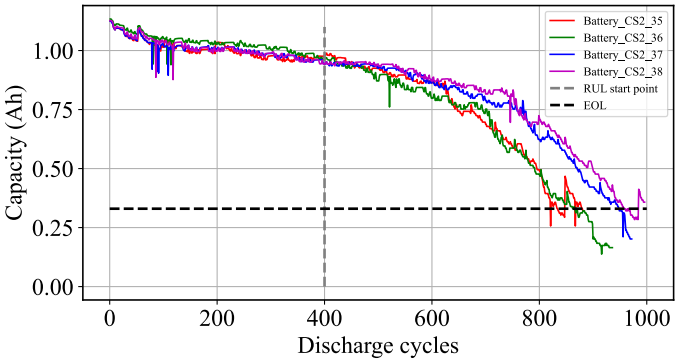

Figure 1: Capacity degradation of the CALCE dataset over discharge cycles. The 70% capacity threshold line indicates the commonly accepted EOL criterion for LiBs.

where N_{EOL}^{true} and N_{EOL}^{pred} are the true and predicted end-of-life cycles, respectively, and N_{current} is the RUL monitoring cycle.

C. Evaluation metrics

To measure the accuracy of the model in SOH and RUL estimation, RMSE, MAE, R^2, and RUL absolute error (RUL_{error}) are used. The corresponding formula is as follows:

$$\mathrm{RMSE}_{SOH} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(SOH_{true,i} - SOH_{pred,i})^2}, \tag{4}$$

$$\mathrm{MAE}_{SOH} = \frac{1}{n}\sum_{i=1}^{n}\left|SOH_{true,i} - SOH_{pred,i}\right|, \tag{5}$$

$$R^2_{SOH} = 1 - \frac{\sum_{i=1}^{n}(SOH_{true,i} - SOH_{pred,i})^2}{\sum_{i=1}^{n}(SOH_{true,i} - SOH_{true})^2}, \tag{6}$$

$$RUL_{Error} = |RUL_{pred} - RUL_{true}| \tag{7}$$

$SOH_{true,i}$ is the actual SOH value at index i, $SOH_{pred,i}$ is the predicted SOH value at index i, SOH_{true} is the mean of the actual SOH values, and n is the total number of data points. RUL_{Error} is the absolute error between RUL_{pred} and RUL_{true}.

III. Research Methodology

This section presents the modeling, development, and working of our proposed method. Below is the detailed explanation.

A. CNN for short-term feature extraction

CNN in the proposed method serves as a local feature extractor that identifies spatially correlated temporal patterns in the battery dataset. By applying one-dimensional (1D) convolutions, CNN captures short-term degradation signatures and learns hierarchical representations where deeper layers encode complex degradation behaviors. The convolutional operation is expressed as

$$h_i^{(k)} = f(W^{(k)} * x_{i:i+s-1} + b^{(k)}), \tag{8}$$

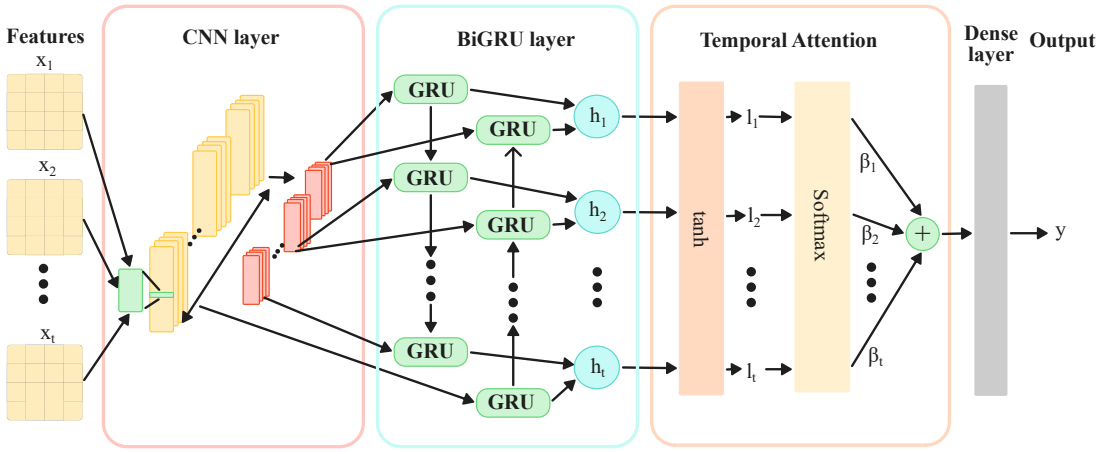

Figure 2: Proposed CNN-BiGRU architecture with temporal attention for accurate SOH and RUL prediction.

where $W^{(k)}$ and $b^{(k)}$ denote the kernel weights and bias of the k-th filter, s is the kernel size, and f is a nonlinear activation function. This process reduces noise, enhances generalization, and improves computational efficiency. The extracted feature maps are subsequently fed into the BiGRU.

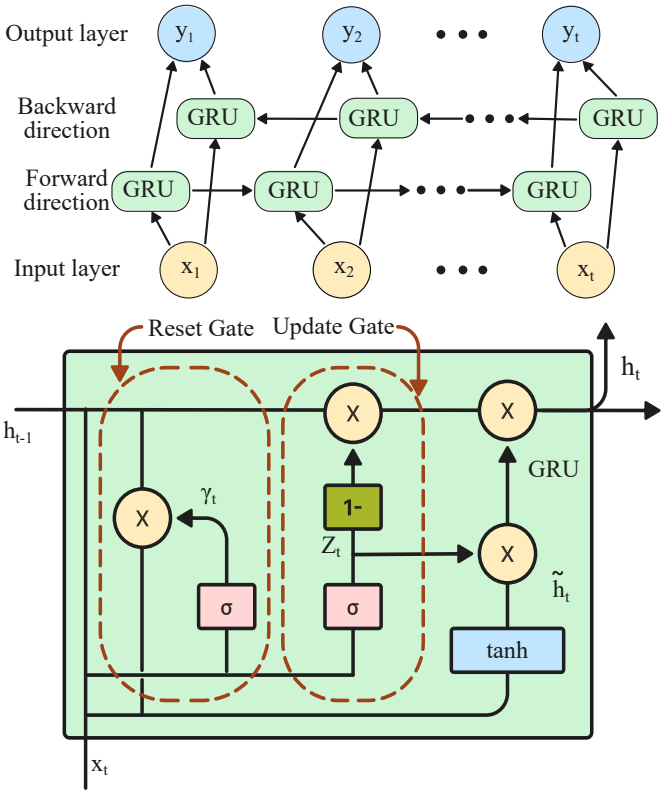

Figure 3: Bidirectional GRU (BiGRU) architecture.

B. BiGRU for long-term feature extraction

GRU networks are widely used for sequence modeling tasks and have shown strong performance in battery SOH estimation due to their ability to capture long-term dependencies from complex input patterns. In this work, a BiGRU architecture is adopted (see Fig. 3) to further improve the learning of temporal correlations by processing data in both forward and backward directions. This bidirectional structure enables the model to extract richer degradation features from cycling data, improving the robustness and accuracy of SOH prediction. GRU eliminates the need for a separate memory cell by updating a hidden state directly, making it particularly efficient for SOH estimation from time-series sensor data. The GRU's reset and update gates can prevent gradient dispersion, guarantee long-term memory capacity, and lessen the model's computational load [23].

The update gate Z_t regulates the amount of the prior hidden state that is kept, helping to maintain long-term dependencies that are crucial for monitoring battery deterioration. The amount of historical data that is disregarded is determined by the reset gate R_t. While the final hidden state h_t integrates past and current information for efficient temporal representation, the candidate hidden state $\hat{h}_t$ combines the reset-modulated memory and the current input. The sigmoid activation function is represented by σ, and the hyperbolic tangent activation function is represented by $tanh$, as shown in the equations below:

$$Z_t = \sigma(W_z \cdot [h_{t-1}, X_t]), \tag{9}$$

$$R_t = \sigma(W_r \cdot [h_{t-1}, X_t]), \tag{10}$$

$$\hat{h}_t = \tanh(W \cdot [R_t * h_{t-1}, X_t]), \tag{11}$$

$$h_t = (1 - Z_t) * h_{t-1} + Z_t * \hat{h}_t, \tag{12}$$

$$\sigma = \frac{1}{1 + e^{-x}}, \tag{13}$$

$$\tanh = \frac{e^{x} - e^{-x}}{e^{x} + e^{-x}}. \tag{14}$$

C. Temporal attention for highlighting important time steps

The attention module conveys the temporal relevance of attributes and boosts the BiGRU's potential to learn detailed input-output links. Relying on the most relevant parts of the input sequence is frequently vital in time-series modeling tasks, including battery SOH estimation. The temporal atten-

tion increases precision as well as understanding by enabling the model to select and attend to independent time steps through the distribution of important weights. The temporal attention suggested in [24], which has been tailored for time series data, strengthens neural networks' grasp in making use of the underlying trend patterns in these kinds of time series data. Because of this, time series data may be processed more precisely and individually. The basic temporal attention process for figuring out the SOH of LiBs comprises three stages: feature weighting, attention weight calculations, and SOH output generation [25].

Given an input sequence $\mathbf{H} = [h_1, h_2, \ldots, h_T] \in \mathbb{R}^{T \times d}$, where T is the sequence length and d is the hidden dimension, the temporal attention computes attention weights α_t for each time step t as in:

$$e_t = \tanh(W_a h_t + b_a), \tag{15}$$

$$\alpha_t = \frac{\exp(e_t^\top u_a)}{\sum_{k=1}^{T} \exp(e_k^\top u_a)}, \tag{16}$$

where W_a and b_a are the weight and the bias matrix, and u_a is a vector that acts like a "query" to measure importance. They are all trainable parameters. The alignment score e_t represents the relevance of the hidden state h_t with respect to the overall sequence. The attention weights α_t are then normalized via the softmax function to ensure they sum to one. The final context vector c is computed as a weighted sum of the hidden states as shown below:

$$c = \sum_{t=1}^{T} \alpha_t h_t. \tag{17}$$

This context vector captures the most relevant temporal features and is subsequently passed to the output layer for final SOH estimation. The output of the model, representing the predicted SOH Y_{SOH}, is obtained through a function $f(c)$ that maps the context vector via a fully connected layer as in:

$$Y_{\text{SOH}} = f(c). \tag{18}$$

D. Model Execution process

The DL architecture presented in this research is as shown in Fig. 2. The proposed method includes 1D convolutional layers to capture local features from the time-series battery cycling data. Two 1D CNN layers act as an automatic feature extractor for training data, eliminating the need for manual feature extraction. These extracted local features pass through the BiGRU layer, capturing long-term dependencies. For resource-heavy tasks like utilizing time-series data of battery cycling to learn long-term dependencies, BiGRU layers have a better computational edge over BiLSTM layers. The learned long-term dependencies from the BiGRU layer pass through the temporal attention layer. The temporal attention layer is used to capture the most informative time steps in battery cycling time-series data while keeping the complexity of the model to a minimum. Finally, the output from the temporal attention layer is passed through the dense layer. The dense layer is the final layer in the proposed method, which reduces the higher dimensionality from other layers to one estimating SOH as the final output. The network is trained across 1000 epochs with a batch size of 16 and verified with LOOCV. Table II shows how the CALCE data are divided into four different groups for model training and testing.

TABLE II: CALCE dataset's division into different group.

Group	Training Cells	Testing Cell
Group I	CS2_36, CS2_37, CS2_38	CS2_35
Group II	CS2_35, CS2_37, CS2_38	CS2_36
Group III	CS2_35, CS2_36, CS2_38	CS2_37
Group IV	CS2_35, CS2_36, CS2_37	CS2_38

The proposed method is trained using the mean squared error loss function. The AdamW optimizer is employed with a learning rate of 0.001. A dropout of 0.25 is used during training to mitigate overfitting by randomly deactivating a fraction of neurons in each iteration. Evaluation metrics such as MAE, RMSE, and R^2 are calculated on the test data for each group. General experimentation based on training loss and evaluation metrics is used to fine-tune hyperparameters.

IV. Results and Discussion

The performance of different approaches for the SOH and RUL prediction is examined in this section. Four data groups were analyzed for training and testing. The SOH and RUL prediction results for every group are covered using performance evaluation metrics. The battery capacity in ampere-hours (Ah) is presented on the y-axis, and the cycle number is presented on the x-axis. A threshold of 70% of the original capacity was used in this research to determine the EOL. In this research, RUL prediction started at the 400th cycle and ended at the EOL. The actual RUL values were then compared to the five DL methods: CNN, LSTM, BiLSTM-GRU, CNN-BiLSTM, and the suggested CNN-BiGRU with the temporal attention. The black line indicates the real SOH, while the red line denotes the proposed method. The comparison of several models to estimate SOH and RUL based on the group I dataset is displayed in Fig. 4.

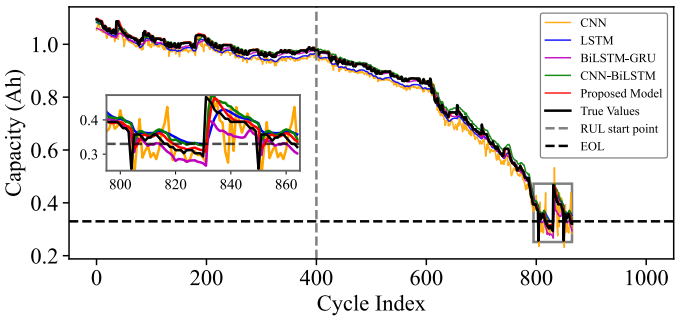

Figure 4: Prediction results of data group I (CS2_35).

TABLE III: Data group I (CS2_35) prediction results.

Model name	SOH Prediction			RUL Comparison		
	MAE	*RMSE*	R^2	*True*	*Pred.*	*Error*
CNN	0.0269	0.0313	0.9770	404	420	16
LSTM	0.0180	0.0216	0.9891	404	419	15
BiLSTM-GRU	0.0125	0.0172	0.9930	404	416	12
CNN-BiLSTM	0.0116	0.0171	0.9931	404	414	10
Proposed	**0.0072**	**0.0121**	**0.9965**	**404**	**406**	**2**

The comparison of prediction results utilizing multiple approaches is shown in Table III. The proposed model achieved an MAE of 0.0072, RMSE of 0.0121, an R^2 of 0.9965 and an RUL_{error} of 2 cycles on group I of the dataset where CS2_35 cell is used for testing and other cells are used for training.

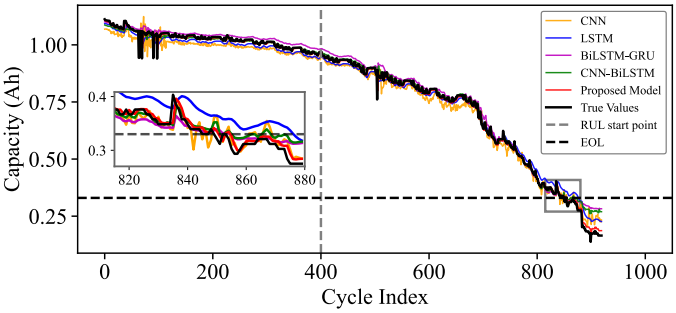

Figure 5: Prediction results of data group II (CS2_36).

TABLE IV: Data group II (CS2_36) prediction results.

Model name	SOH Prediction			RUL Comparison		
	MAE	*RMSE*	R^2	*True*	*Pred.*	*Error*
CNN	0.0264	0.0310	0.9855	450	463	13
LSTM	0.0208	0.0258	0.9900	450	457	7
BiLSTM-GRU	0.0131	0.0248	0.9910	450	455	5
CNN-BiLSTM	0.0199	0.0305	0.9859	450	442	8
Proposed	**0.0082**	**0.0145**	**0.9968**	450	**453**	**3**

Comparing the SOH and RUL predictions for group II, where CS2_36 is used as a testing cell and others are used for training, is shown in Fig. 5. The proposed method obtained the lowest SOH and RUL prediction error, with an SOH prediction MAE of 0.0082, an RMSE of 0.0145, and an RUL_{error} of 3 cycles, as shown in Table IV. Similarly, the proposed model's performance in group III of the dataset using CS_37 cell as the testing data is as shown in Fig. 6. The comparative error

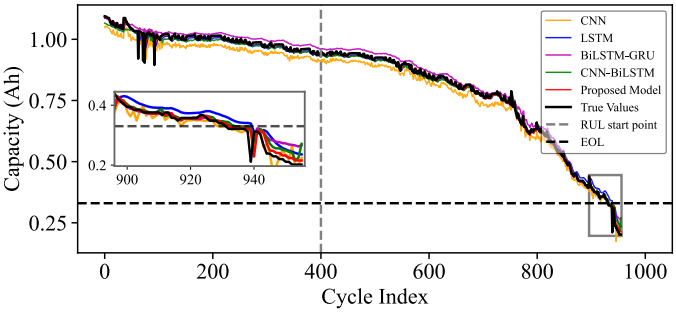

Figure 6: Prediction results of data group III (CS2_37).

analysis of multiple models using the CS2_37 cell's data for testing is as shown in Table V. The proposed method obtained good results, with an MAE of 0.0065, an RMSE of 0.0124, an R^2 of 0.9963, and an RUL_{error} of 1 cycle in group III, where the CS2_37 cell's data is used for testing and other cells' data are used for training. Similarly, the SOH and RUL

TABLE V: Data group III (CS2_37) prediction results.

Model name	SOH Prediction			RUL Comparison		
	MAE	*RMSE*	R^2	*True*	*Pred.*	*Error*
CNN	0.0297	0.0321	0.9759	532	551	19
LSTM	0.0112	0.0216	0.9941	532	537	5
BiLSTM-GRU	0.0110	0.0172	0.9944	532	536	4
CNN-BiLSTM	0.0173	0.0215	0.9891	532	541	9
Proposed	**0.0065**	**0.0124**	**0.9963**	532	**533**	**1**

predictions of various models were graphically compared for group IV of the dataset, where CS_38 cell is used as a testing and other cells are used for training as shown in Fig. 7, and evaluation metrics were compared in Table VI. Our model scored exceptionally well in SOH and RUL prediction with an MAE of 0.0075, an RMSE of 0.0131, an R^2 of 0.9958, and an RUL_{error} of 0 cycle.

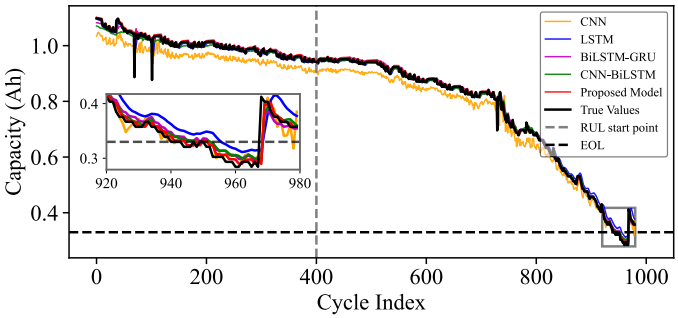

Figure 7: Prediction results of data group IV (CS2_38).

TABLE VI: Data group IV (CS2_38) prediction results.

Model name	SOH Prediction			RUL Comparison		
	MAE	*RMSE*	R^2	*True*	*Pred.*	*Error*
CNN	0.0264	0.0368	0.9671	540	553	13
LSTM	0.0095	0.0149	0.9946	540	547	7
BiLSTM-GRU	0.0083	0.0133	0.9956	540	550	10
CNN-BiLSTM	0.0084	0.0134	0.9856	540	548	8
Proposed	**0.0075**	**0.0131**	**0.9958**	540	**540**	**0**

V. Conclusion

An accurate estimation of SOH and RUL is critical for assessing battery condition, degradation, and its available lifespan. These key factors are essential to optimize the battery usage, enhance the safety, and reduce the operational costs. This work proposes a DL framework based on CNN-BiGRU with temporal attention, which effectively captures short- and long-term temporal dependencies. The proposed method achieved good prediction results compared with other methods such as CNN, LSTM, BiLSTM-GRU, and CNN-BiLSTM,

with the lowest MAE of 0.0072, an RMSE of 0.0121, and the highest R^2 of 0.9965 in group I, showing the proposed method improved the performance of SOH and RUL estimation. In future, this research can be integrated on diverse datasets that can ensure that the BMS adapts to various battery chemistries and usage patterns and facilitate more intelligent BMS in electric vehicles, aerospace systems, and renewable energy storage applications.

ACKNOWLEDGMENT

This research was partly supported by the Natural Science Basic Research Program of Shaanxi Province (2022JQ-595); Shanghai Sailing Program (No. 22YF1452300); Science and Technology Program of Suzhou (No.561366726361) ZXL2022457, and National Natural Science Foundation of China (NSFC, No.52302480).

REFERENCES

[1] Y. Yang, Y. Ye, Z. Cheng, G. Ruan, Q. Lu, X. Wang, and H. Zhong, "Life cycle economic viability analysis of battery storage in electricity market," *J. Energy Storage*, vol. 70, Oct. 2023

[2] K. S. Ng, C.-S. Moo, Y.-P. Chen, and Y.-C. Hsieh, "Enhanced coulomb counting method for estimating state-of-charge and state-of-health of lithium-ion batteries," *Appl. Energy*, vol. 86, no. 9, pp. 1506–1511, 2009.

[3] H. Dai, G. Zhao, M. Lin, J. Wu, and G. Zheng, "A novel estimation method for the state of health of lithium-ion battery using prior knowledge-based neural network and Markov chain," *IEEE Trans. Ind. Electron.*, vol. 66, no. 10, pp. 7706–7716, Oct. 2019.

[4] Y. Guo, D. Yang, Y. Zhang, L. Wang, and K. Wang, "Online estimation of soh for lithium-ion battery based on ssa-elman neural network," *Protection and Control of Modern Power Systems*, vol. 7, no. 3, pp. 1–17, 2022.

[5] Y. Chen, Y. Kang, Y. Zhao, L. Wang, J. Liu, Y. Li, Z. Liang, X. He, X. Li, N. Tavajohi, and B. Li, "A review of lithium-ion battery safety concerns: The issues, strategies, and testing standards," *Journal of Energy Chemistry*, vol. 59, pp. 83–99, 2021.

[6] W. Li, C. Lin, S. Hosseininasab, L. Bauer, and S. Pischinger, "Lithium-Ion Battery SOH Estimation Based on a Long Short-Term Memory Model Using Short History Data," *IEEE Trans. Power Electron.*, vol. 40, no. 5, pp. 7370–7384, May 2025.

[7] U. Saleem, W. Li, W. Liu, I. Ahmad, M. M. Aslam, and H. U. Lateef, "Investigation of Deep Learning Based Techniques for Prognostic and Health Management of Lithium-Ion Battery," in *Proc. 2023 15th Int. Conf. on Electronics, Computers and Artificial Intelligence (ECAI)*, Bucharest, Romania, pp. 01–06, 2023

[8] Y. Li et al., "SOH evaluation and RUL estimation of lithium-ion batteries based on MC-CNN-TimesNet model," *Reliability Engineering & System Safety*, vol. 261, p. 111125, Sep. 2025.

[9] [T. Oji, Y. Zhou, S. Ci, F. Kang, X. Chen, and X. Liu, "Data-Driven Methods for Battery SOH Estimation: Survey and a Critical Analysis," *IEEE Access*, vol. 9, pp. 126903–126916, 2021

[10] K. Luo, X. Chen, H. Zheng, and Z. Shi, "A review of deep learning approach to predicting the state of health and state of charge of lithium-ion batteries,"*Journal of Energy Chemistry*, vol. 74, pp. 159–173, 2022.

[11] S. Fu, S. Tao, H. Fan, et al.,"Data-driven capacity estimation for lithium-ion batteries with feature matching based transfer learning method," *Applied Energy,* vol. 353, Article 121991, 2024.

[12] S. Peng, J. Zhu, T. Wu, et al., "SOH early prediction of Lithium-ion batteries based on voltage interval selection and features fusion," *Energy,* vol. 308, Article 132993, 2024.

[13] U. Saleem, W. Liu, S. Riaz, M. M. Aslam, W. Li, and K. Wang, "EnerNet: Attention-based dilated CNN-BiLSTM for state of health prediction of CS2 prismatic cells in energy systems," *Electrochimica Acta*, vol. 512, p. 145454, Feb. 2025.

[14] Z. Zhu, Q. Yang, X. Liu, and D. Gao, "Attention-based CNN-BiLSTM for SOH and RUL estimation of lithium-ion batteries," *Journal of Algorithms & Computational Technology*, vol. 16, p. 17483026221130598, Jan. 2022.

[15] Y. Zhang, R. Xiong, H. He, and M. G. Pecht, "Long short-term memory recurrent neural network for remaining useful life prediction of lithium-ion batteries," *IEEE Trans. Veh. Technol.*, vol. 67, no. 7, pp. 5695–5705, Jul. 2018.

[16] L. Ren, J. Dong, X. Wang, Z. Meng, L. Zhao, and M. J. Deen, "A data-driven auto-CNN-LSTM prediction model for lithium-ion battery remaining useful life," *IEEE Trans. Ind. Inform.*, vol. 17, no. 5, pp. 3478–3487, May 2021.

[17] J. Lin, C. Wang, and G. Yan, "State-of-health prediction of lithium-ion battery based on improved gate recurrent unit," *J. Phys. Conf. Ser.*, vol. 2010, no. 1, Art. no. 012142, 2021.

[18] X. Wang, B. Hu, X. Su, et al.,"State of health estimation for lithium-ion batteries using random forest and gated recurrent unit,"*Journal of Energy Storage*, vol. 76, Article 109796, 2024.

[19] Y. Zheng, J. Hu, J. Chen, H. Deng, and W. Hu, "State of health estimation for lithium battery random charging process based on CNN-GRU method," *Energy Reports*, vol. 9, Supplement 3, pp. 1–10, 2023

[20] K. Xue, J. Yang, M. Yang and D. Wang, "An Improved Generic Hybrid Prognostic Method for RUL Prediction Based on PF-LSTM Learning," *IEEE Trans. on Instrumentation and Measurement*, vol. 72, pp. 1-21, 2023.

[21] J. Gao, D. Yang, S. Wang, Z. Li, L. Wang, and K. Wang, "State of health estimation of lithium-ion batteries based on mixers-bidirectional temporal convolutional neural network," *J. Energy Storage*, vol. 73, Article 109248, 2023.

[22] U. Saleem, W. Liu, S. Riaz, W. Li, G. A. Hussain, Z. Rashid, and Z. A. Arfeen, "TransRUL: A Transformer-Based Multihead Attention Model for Enhanced Prediction of Battery Remaining Useful Life," *Energies*, vol. 17, no. 16, Art. no. 3976, 2024.

[23] C. Zhang et al., "Battery SOH estimation method based on gradual decreasing current, double correlation analysis and GRU," *Green Energy and Intelligent Transportation*, vol. 2, no. 5, p. 100108, Oct. 2023

[24] S.-Y. Shih, F.-K. Sun, and H. Lee, "Temporal pattern attention for multivariate time series forecasting," *Machine Learning*, vol. 108, no. 8, pp. 1421–1441, Sep. 2019

[25] J. Huang et al., "A lithium-ion battery SOH estimation method based on temporal pattern attention mechanism and CNN-LSTM model," *Computers and Electrical Engineering*, vol. 122, p. 109930, Mar. 2025