

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/391781705>

# Towards Efficient SOH Estimation for Lithium-ion Batteries via Structural Re-parameterization

Conference Paper · November 2024

DOI: 10.1109/EI264398.2024.10990824

---

CITATIONS

0

---

READ

1

4 authors, including:



La Sandhu

Central South University

22 PUBLICATIONS 42 CITATIONS

SEE PROFILE

# Towards Efficient SOH Estimation for Lithium-ion Batteries via Structural Re-parameterization

Qian Zhou  
School of Automation  
Central South University  
Changsha, China  
224611074@csu.edu.cn

Yun Wang\*  
School of Automation  
Central South University  
Changsha, China  
wangyun19@csu.edu.cn

Guang Wu  
School of Automation  
Central South University  
Changsha, China  
wuguangcsuer@csu.edu.cn

Laceq Aslam  
School of Automation  
Central South University  
Changsha, China  
laceq.aslam.100@gmail.com

**Abstract**—Online state of health (SOH) estimation is critical for the safety and performance of lithium-ion batteries, especially in dynamic environments like electric vehicles and grid energy storage systems, where accurate monitoring prevents failures and enhances reliability. Convolutional Neural Networks have shown great potentials in SOH estimation. To improve the estimation accuracy, traditional methods focus on increasing the depth and width of the network. The resulting network structure is often complex and difficult to deploy on resource limited mobile devices. Based on the structure reparameterization technique, this paper simplifies the network structure from the training stage to the inference stage. Tuning the hyperparameters of a complex network is not an easy task. In this paper, an advanced hyperparameter optimization engine is used to automatically search for the optimal hyperparameters of the network structure. Finally, in the inference stage, the FLOPs of the model after structure reparameterization are reduced by 16.799% as least. The SOH estimation results on two battery aging datasets demonstrate the superiority and robustness of the proposed model, with a maximum average performance improvement of 21.92%.

**Keywords**—SOH estimation, Structural re-parameterization, Hyperparameter optimization, Lithium-ion batteries, Convolutional neural network.

## I. INTRODUCTION

The increased adoption of lithium-ion batteries in electric vehicles (EVs) and renewable energy systems has amplified the need for accurate and efficient state of health (SOH) estimation to ensure performance and safety [1]. Accurate SOH monitoring not only supports optimal energy usage but also minimizes risks associated with battery degradation, such as capacity reduction and increased resistance, which can lead to overheating and safety hazards [2]. Consequently, SOH estimation has become essential in extending the lifespan and reliability of EV batteries and grid storage systems, highlighting the importance of developing robust and efficient SOH estimation methods for these applications.

Traditionally, SOH estimation methods are divided into two primary categories: model-based and data-driven approaches [3]. Model-based methods involve constructing equivalent circuit models, electrochemical models, or empirical degradation

models to simulate battery behavior and monitor health indicators such as capacity retention and resistance change. Although these models can be highly accurate, they often require extensive domain knowledge and specialized parameterization, limiting their versatility across different battery chemistries and operating conditions. On the other hand, data-driven methods leverage historical battery data, bypassing the complexities of electrochemical modeling by directly mapping observable indicators to SOH predictions. This approach, supported by machine learning and deep learning techniques, has shown promise in capturing complex battery aging behaviors in real time, offering improved generalizability and adaptability across battery types and use cases.

Within data-driven SOH estimation, Convolutional Neural Networks (CNNs) have demonstrated significant potential, especially when combined with feature selection and model refinement techniques. Recent studies have proposed hybrid models, such as CNN-BiLSTM and transfer learning architectures [4], that integrate multiple data sources and leverage recurrent layers for temporal data processing, achieving state-of-the-art accuracy in SOH estimation with limited data. These models are particularly valuable for real-world applications, where acquiring extensive labeled datasets for different battery types is challenging. Moreover, incremental capacity and partial incremental capacity analysis, as explored in recent studies, have further refined feature selection by focusing on specific charge and discharge segments, enhancing model interpretability and accuracy [5]. However, these advanced models are often complex and computationally expensive, posing challenges for deployment on resource-limited devices commonly used in EVs and portable applications [6].

To address the challenges of model efficiency and computational complexity in SOH estimation, structural reparameterization techniques, exemplified by RepMLPNet [7], offer a promising solution. Structural reparameterization optimizes network architectures by converting complex structures used in the training phase into streamlined versions for inference, thereby reducing the computational load without sacrificing accuracy [8]. By adapting RepMLPNet with an improved channel attention mechanism, this study leverages a more efficient architecture suitable for real-time applications in EVs and IoT-based energy systems, enabling high-performance SOH estimation that remains lightweight and deployable on

This work was supported in part by the National Natural Science Foundation of China under Grant 62376289, and in part by the Natural Science Foundation of Hunan Province, China under Grant 2024JJ4069. Corresponding author: Yun Wang.

edge devices with limited resources. Furthermore, effective hyperparameter tuning is essential for maximizing model performance across diverse datasets and operating conditions. Utilizing the Optuna framework [9] within a carefully designed search space, this work automates the tuning process, achieving optimal parameter configurations that enhance both tuning efficiency and estimation accuracy. This approach enables the re-parameterized model to adapt robustly to different operational demands, further improving its generalizability and robustness.

In this paper, a structurally re-parameterized CNN model, termed r-RepMLPNet, is proposed for efficient SOH estimation of Lithium-ion batteries (LIBs). The specific contributions are as follows:

- The proposed r-RepMLPNet extends the original RepMLPNet model with an optimized channel attention mechanism, reinforcing feature extraction efficiency and simplifying the network structure, making it suitable for real-time SOH estimation tasks.
- Given a targeted search space, the Optuna framework efficiently tunes r-RepMLPNet, identifying optimal hyperparameters that enhance model accuracy and tuning efficiency, and making SOH estimation performance consistent across varying operational scenarios.
- By leveraging structural re-parameterization techniques during inference, r-RepMLPNet achieves significant FLOPs reduction without compromising accuracy, offering a high-performance, low-complexity model fit for deployment on resource-constrained devices.

## II. PRELIMINARY WORK

### A. SOH Definition

SOH is used to describe the health state of a battery which is defined as follows [10]:

$$SOH_i = \frac{Q_i}{Q_0}, \quad (1)$$

where  $Q_i$  is the maximum accessible capacity at the  $i^{th}$  cycle, and  $Q_0$  is the initial capacity of the battery.

### B. Structural Re-parameterization

Structural re-parameterization is a technique used to enhance the efficiency of deep neural networks by converting multi-branch structures into simpler, single-branch ones during inference. This process reduces computational overhead while retaining performance [11]. There are two representative cases of structural re-parameterization: first, fusing batch normalization (BN) parameters into conv (short for “convolutional”) layers; and second, merging bypass conv parameters into the main fully connected (FC) layers [7]. In this paper, a feature map is denoted as  $M \in \mathbb{R}^{n \times c \times h \times w}$ , where  $n$ ,  $c$ ,  $h$ ,  $w$  are the batch size, number of channels, height and width, respectively. Assuming that the size of the feature map remains unchanged after passing through the conv layer, this process can be expressed as:

$$M^{\text{out}} = \text{CONV}(M^{\text{in}}, F, p), \quad (2)$$

where  $M^{\text{out}} \in \mathbb{R}^{n \times o \times h \times w}$  is the output,  $o$  is the number of output channels,  $p$  is the number of pixels to pad, and  $F \in \mathbb{R}^{o \times c \times k \times k}$  is the kernel of conv.

#### 1) Fuse BN into conv

For a trained BN layer,  $\mu$ ,  $\sigma$ ,  $\gamma$ ,  $\beta \in \mathbb{R}^o$  are the accumulated mean, standard deviation, and learned scaling factor and bias, respectively. Then, to fuse the BN layer below a conv layer, the formula is given by

$$\begin{aligned} & \frac{\gamma_i}{\sigma_i} (\text{CONV}(M, F, p)_{:,i,:} - \mu_i) + \beta_i \\ & = \text{CONV}(M, F', p)_{:,i,:} + b'_i, \forall 1 \leq i \leq o, \end{aligned} \quad (3)$$

where  $F'$  and  $b'$  are the new kernel and bias, constructed by

$$F'_{i,:} = \frac{\gamma_i}{\sigma_i} F_{i,:}, \quad b'_i = -\frac{\mu_i \gamma_i}{\sigma_i} + \beta_i. \quad (4)$$

#### 2) Merge conv into FC

In this paper, a vector is denoted by as  $V \in \mathbb{R}^{n \times p}$ , the data flow through an FC layer (without bias) is formulated as

$$V^{\text{out}} = \text{MMUL}(V^{\text{in}}, W) = V^{\text{in}} \cdot W^{\text{T}}, \quad (5)$$

where  $V^{\text{out}} \in \mathbb{R}^{n \times q}$  is the output,  $q$  is the number of output neurons, and  $W \in \mathbb{R}^{q \times p}$  is the kernel of FC. When considering the input as a feature map, RS (short for “reshape”) is needed, i.e., the input is first flattened into  $n$  vectors of length  $chw$ , which is  $V^{\text{in}} = \text{RS}(M^{\text{in}}, (n, chw))$ , multiplied by the kernel  $W(ohw, chw)$ , then the output  $V^{\text{out}}(n, ohw)$  is reshaped back into  $M^{\text{out}}(n, o, h, w)$ . The whole process can be simplified as

$$M^{\text{out}} = \text{MMUL}(M^{\text{in}}, W). \quad (6)$$

In order to equivalently merge the conv layer into its parallel FC layer, it is necessary to construct such a  $W'$  which satisfies the following equation

$$\begin{aligned} & \text{MMUL}(M^{\text{in}}, W') \\ & = \text{MMUL}(M^{\text{in}}, W) + \text{CONV}(M^{\text{in}}, F, p). \end{aligned} \quad (7)$$

Considering the additivity of MMUL, the core of merging a conv layer into the FC layer lies in constructing a  $W^{F,p}$ , with the same shape of  $W$  to ensure that

$$\text{MMUL}(M^{\text{in}}, W^{(F,p)}) = \text{CONV}(M^{\text{in}}, F, p). \quad (8)$$

Equation (8) can be written in another way, that is

$$V^{\text{out}} = V^{\text{in}} \cdot W^{(F,p)\text{T}}. \quad (9)$$

For a given  $M^{\text{in}}$ ,  $F$ ,  $p$ , the corresponding  $W^{(F,p)}$  exists as a conv layer can be seen as a parameter-shared FC layer, structured liked a Toeplitz matrix. If inserting an identity matrix  $I(chw, chw)$  into (9), and then with explicit RS, the following formula can be derived:

$$V^{\text{out}} = V^{\text{in}} \cdot \text{RS}(I \cdot W^{(F,p)\text{T}}, (chw, ohw)). \quad (10)$$

Note that  $I \cdot W^{(F,p)\text{T}}$  is exactly a convolution with  $F$  on a feature map  $M^{\text{I}}$  which is reshaped from  $I$ , i.e.,

$$\begin{aligned} & I \cdot W^{(F,p)\text{T}} = \text{CONV}(M^{\text{I}}, F, p) \\ & = \text{CONV}(\text{RS}(I, (chw, c, h, w)), F, p). \end{aligned} \quad (11)$$

Actually, the formula to obtain  $W'$  given in [7], that is,

$$\begin{aligned} W' &= W^{(F,p)} + W \\ &= \text{RS}(\text{CONV}(M^I, F, p), (chw, ohw))^\top + W. \end{aligned} \quad (12)$$

### III. METHODOLOGY

#### A. *r-RepMLPNet: reinforced RepMLPNet*

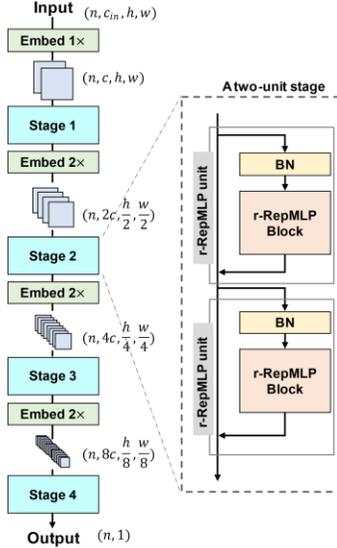


Fig. 1. Multi-stage *r-RepMLPNet*.

The architecture of the proposed *r-RepMLPNet* is shown in Fig. 1, which is largely consistent with the original *RepMLPNet*, but two changes are made to simplify the network structure and reduce the memory and computational overhead. First, the FFN-style blocks in the *r-RepMLP* unit are omitted. Second, the global perceptron in the *r-RepMLP* block, which can also be regarded as an Squeeze-Excitation (SE) Attention [12] module, is replaced with an Effective Channel Attention (ECA) [13] module. The details are elaborated in the following subsections.

#### 1) *ECA vs. SE Attention*

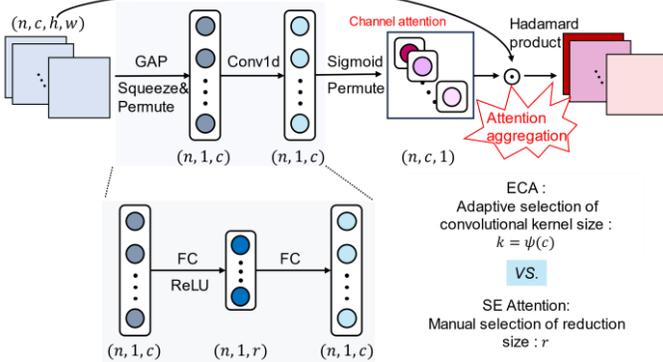


Fig. 2. The schemma of ECA and SE Attention.

SE Attention emphasizes relevant features by explicitly recalibrating channel weights through a squeeze-and-excitation operation, as shown in Fig. 2.

$$M^{\text{out}} = M^{\text{in}} \odot \sigma \left( f_{\{W_1, W_2\}} \left( g(M^{\text{in}}) \right) \right), \quad (13)$$

where  $g(M^{\text{in}}) = \frac{1}{w \times h} \sum_{i=1, j=1}^{w, h} M_{ij}^{\text{in}}$  is channel wise global average pooling,  $\sigma$  is a Sigmoid function,  $W_1 \in \mathbb{R}^{r \times c}$  and  $W_2 \in \mathbb{R}^{c \times r}$  are the kernel of two FC layers, respectively. Let  $y = g(M^{\text{in}})$ , then  $f_{\{W_1, W_2\}}(\cdot)$  is formulated as

$$f_{\{W_1, W_2\}}(y) = W_2 \text{ReLU}(W_1 y), \quad (14)$$

where ReLU indicates the Rectified Linear Unit.

ECA refines channel-wise dependencies by applying adaptive convolution, enhancing efficiency without dimensionality reduction. The mapping formula for  $k$  is written in (15)

$$k = \psi(c) = \left\lfloor \frac{|\log_2 c + b|}{\gamma} \right\rfloor_{\text{odd}}, \quad (15)$$

where  $\lfloor \cdot \rfloor_{\text{odd}}$  is the nearest odd number,  $\gamma$  and  $b$  are set to 2 and 1. Unlike SE, ECA omits dimensionality reduction, yielding a more lightweight and computationally efficient mechanism.

#### 2) *r-RepMLP Block*

As illustrated in Fig. 3, the *r-RepMLP* block integrates a multi-branch design in the training phase, then with structural re-parameterization, these branches are turned into a simpler, single-branch architecture for inference.

In its training-time configuration, the *r-RepMLP* Block is designed model the information on different levels by combining three key components: Global Perceptron, Channel Perceptron and Local Perceptron. During the inference phase, the parameters from the trained local perceptron are integrated into the channel perceptron through three main steps, this process can also be described as locality injection. First, the “conv-BN” structure in the local perceptron is transformed into a single “conv” structure, utilizing the method of “merge BN into conv”, as detailed in Section II. Second, three conv branches are combined into a single convolution layer by padding all kernels to the size of the largest kernel. Lastly, the “merge conv into FC” process mentioned in Section II is executed, which completes the locality injection, effectively streamlining the model for enhanced inference efficiency.

#### B. Hyperparameter Optimization

Optuna is an open-source hyperparameter optimization engine that has been successfully used to tune the hyperparameters of machine learning models. Among various samplers provided in Optuna, the Bayesian independent sampling algorithm, i.e., Tree-structured Parzen Estimator is used for hyperparameter searching. In Bayesian optimization, contrary to the grid or random search, the results of past evaluations are employed for building a probabilistic model that maps hyperparameters to the probability of a score in the objective function. In this way, the optimization process is more efficient as the next set of hyperparameters are selected in an informed manner. By setting an estimation metric on the validation set as the objective function, the tuning process is aiming to find the optimal combination of hyperparameters that minimizes this metric.

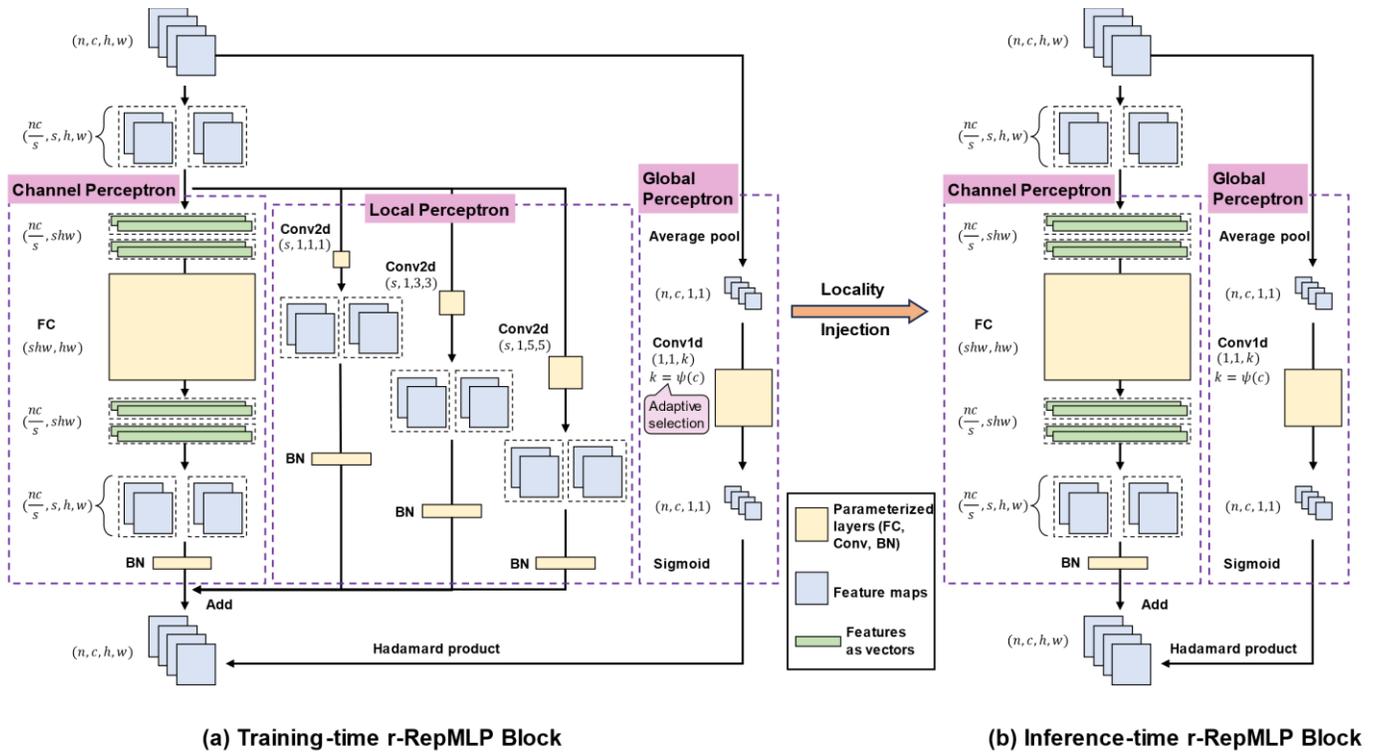


Fig. 3. The detailed structure of r-RepMLP Block.  $s$  is the number of share-sets [7].

### C. Implementation of r-RepMLPNet for SOH Estimation

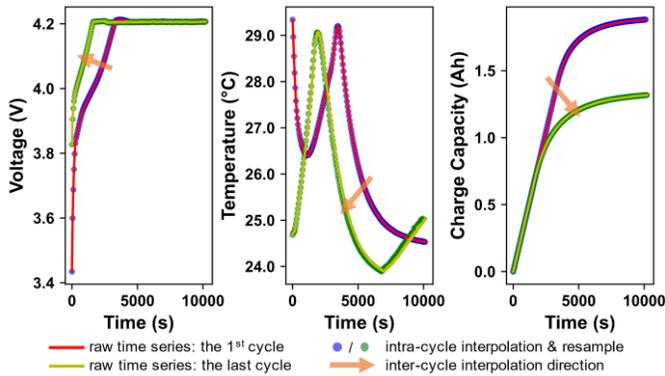


Fig. 4. The process of data augmentation.

The implementation of r-RepMLPNet for SOH estimation begins by acquiring battery aging data from two public datasets. Initially, the dataset is split into training and test sets. Given the limited sample size, data augmentation is applied to the training set as outlined in Fig. 4. The augmented training set is then further divided into a new training and validation set. Mean-variance normalization is performed on all sets for consistency. Next, time-series data are transformed into image formats for network training, using L1 loss as the loss function. Optuna is employed to optimize the network's hyperparameters. The best-performing model is selected based on validation results, after which locality injection is applied to yield a structurally re-parameterized network optimized for efficient model inference.

### IV. EXPERIMENTAL VALIDATION

#### A. Data Description

The NASA battery degradation dataset, from the NASA Ames Prognostics Center of Excellence, includes aging data for lithium cobalt oxide batteries with a rated capacity of 2 Ah [14]. This study used four LIBs (B5, B6, B7, and B18), charged and discharged at room temperature (24 °C). Charging followed a constant-current-constant-voltage (CC-CV) mode with a 1.5 A current until 4.2 V, then constant voltage until the current dropped below 20 mA. Discharging was at a constant current of 2 A until the voltage dropped to 2.7 V, 2.5 V, 2.2 V, and 2.5 V for B5, B6, B7, and B18, respectively. The experiment ended when the battery capacity decreased by 30% (from 2 Ah to 1.4 Ah), marking end of life.

Eight Kokam lithium-ion pouch cells (Cell1–Cell8), each with a 740 mAh capacity, were tested at 40 °C for the Oxford battery degradation dataset [15]. These cells have cathodes of lithium cobalt oxide and lithium nickel cobalt oxide, and anodes of graphite. They were charged using a CC-CV profile and discharged following a drive cycle from the urban Artemis profile. Characterization measurements were taken every 100 cycles to assess the SOH. All cells were charged at 1 C (0.74 A) until reaching 4.2 V, and discharged at 1 C until 2.7 V.

#### B. Evaluation Metrics

To evaluate the performance of the proposed model, mean absolute error (MAE), root mean squared error (RMSE), mean absolute percentage error (MAPE), and coefficient of determination ( $R^2$ ) are employed to measure the discrepancy

between the estimated SOH  $\hat{y}_i$  and the target SOH  $y_i$ . The formulas are given by:

$$MAE = \frac{1}{N} \sum_{i=1}^N |y_i - \hat{y}_i| \times 100\%, \quad (16)$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2} \times 100\%, \quad (17)$$

$$MAPE = \frac{1}{N} \sum_{i=1}^N \frac{|y_i - \hat{y}_i|}{|y_i|} \times 100\%, \quad (18)$$

$$R^2 = 1 - \frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{\sum_{i=1}^N \left( y_i - \frac{1}{N} \sum_{i=1}^N y_i \right)^2}, \quad (19)$$

where  $N$  is the number of test samples.

### C. Experimental Setting

Model parameters significantly influence the performance of data-driven methods. TABLE I. shows the hyperparameter configuration. Except the proposed r-RepMLPNet, a variant of ResNet [16] is used as a baseline model in this study. All deep learning models are uniformly configured with 100 training epochs and an early stopping patience of 20 epochs. The learning rate varies between  $1e-4$  and  $0.01$ , and batch size is set to 512. AdamW is used as the optimizer, with a weight decay of  $1e-4$  and momentum parameters  $\beta = (0.5, 0.999)$ .

TABLE I. HYPERPARAMETER SETTING

Source	Parameter	Search space
Model parameters: r-RepMLPNet	Number of stages	{1, 2}
	Number of r-RepMLP blocks	{1, 2}
	Number of share-sets	{1, 2, 4, 6, 8}
Model parameters: ResNet	Number of residual blocks	{1, 2, 3, 4}
Training parameter	Learning rate	{range ( $1e-4$ , 0.01)}

All cycles in both datasets were resampled to 196 points, expanding the training set to 20 times its size through data augmentation. For the Oxford dataset, Cell6 and Cell8 were target batteries; their full cycle data formed the test set, while data from other batteries were split 8:2 for training and validation. For the NASA dataset, B5 was the target battery, with the last 30% of its cycle data for testing, the prior 70% for validation, and data from all other batteries for training.

### D. Experimental Results

TABLE II. SOH ESTIMATION RESULTS OF THE PROPOSED MODEL ON TWO DATASETS

Model	NASA				Oxford							
	B5				Cell6				Cell8			
	MAE	RMSE	MAPE	R <sup>2</sup>	MAE	RMSE	MAPE	R <sup>2</sup>	MAE	RMSE	MAPE	R <sup>2</sup>
ResNet	0.501	0.726	0.683	0.8815	0.399	0.492	0.460	0.9923	0.532	0.626	0.618	0.9915
p-RepMLPNet	0.543	0.685	0.736	0.8945	0.359	0.484	0.415	0.9925	0.468	0.773	0.561	0.9870
RepMLPNet	0.464	0.688	0.640	0.8938	0.349	0.576	0.408	0.9894	0.432	0.478	0.504	0.9950
r-RepMLPNet	0.452	0.591	0.615	0.9216	0.283	0.460	0.328	0.9933	0.249	0.313	0.297	0.9979

Apart from evaluating the discrepancy between the proposed model's estimation and the true labels, a comparison is conducted with three other models: ResNet and two variations based on RepMLPNet. The primary difference is in the global perceptron; one model employs SE Attention as in the original RepMLPNet ("original"), while the other excludes the global perceptron entirely ("pruned"). The SOH estimation results for B5, Cell6, and Cell8 are illustrated in Fig. 5, Fig. 6, and Fig. 7.

TABLE II. presents the estimation accuracy metrics for four models across three LIBs. The r-RepMLPNet consistently outperforms the other three models, achieving higher estimation accuracy on each battery dataset. The average performance improvements of the proposed model on four error metrics are shown in TABLE III. It shows that the maximum and minimum improvements are 21.920% and 16.750%, respectively. This demonstrates the model's superior capability in accurately capturing the SOH, highlighting its effectiveness and robustness compared to ResNet and the other RepMLPNet-based variants.

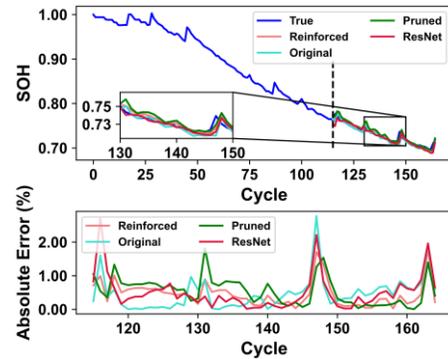


Fig. 5. Estimated SOH curves and the absolute errors obtained with various models for B5.

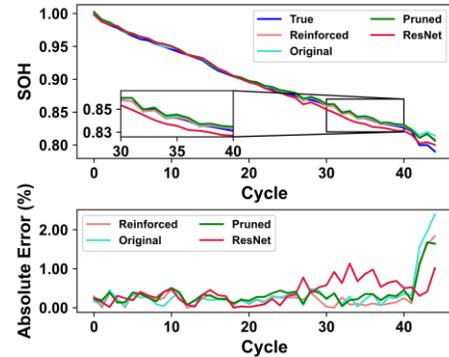


Fig. 6. Estimated SOH curves and the absolute errors obtained with various models for Cell6.

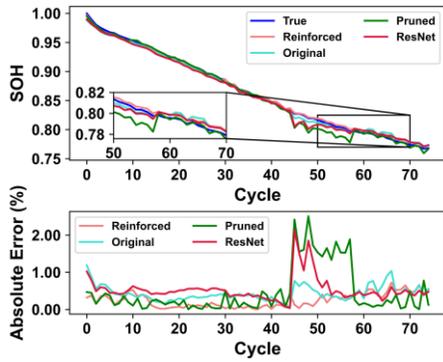


Fig. 7. Estimated SOH curves and the absolute errors obtained with various models for Cell8.

TABLE III. AVERAGE PERFORMANCE IMPROVEMENTS OF r-RepMLPNET

Model	Improvement of r-RepMLPNet (%)				
	MAE	RMSE	MAPE	R <sup>2</sup>	Average
ResNet	30.683	25.033	30.198	1.765	21.920
p-RepMLPNet	28.241	26.063	28.154	1.405	20.966
RepMLPNet	21.286	22.919	21.529	1.265	16.750

TABLE IV. summarizes the FLOPs for three RepMLPNet-based variants before and after locality injection, as well as the reduction rates achieved through structural re-parameterization. The results demonstrate a significant reduction in computational load, with all models showing a decrease of over 16% in FLOPs following locality injection.

TABLE IV. CHANGES IN FLOPs BEFORE AND AFTER LOCALITY INJECTION WITH B5 AS AN EXAMPLE

Model	p-RepMLPNet	RepMLPNet	r-RepMLPNet
Without locality injection	10.500	14.281	14.245
With locality injection	8.580	11.882	11.846
Reduction ratio	18.279%	16.799%	16.841%

## V. CONCLUSION

This study introduces the r-RepMLPNet, an advanced model for online SOH estimation of lithium-ion batteries, particularly designed for dynamic environments such as EVs and grid energy storage systems. By incorporating an optimized channel attention mechanism, the model enhances feature extraction efficiency while streamlining the network architecture for real-time applications. The use of the Optuna framework for hyperparameter optimization allows for effective exploration of the parameter space, resulting in improved model accuracy and consistency across various operational scenarios. Furthermore, structural re-parameterization techniques employed during the inference stage significantly reduce computational complexity without sacrificing accuracy. The proposed model demonstrates enhanced performance and robustness in SOH estimation tasks, providing a promising solution for high-performance, low-complexity applications on resource-constrained devices.

## REFERENCES

- [1] S. Karthikkumar, A. D. Mol, S. Arumugam, V. G. Shankar, S. Sharmila, and S. Subashini, "The Rise of AI-driven BMS: Revolutionizing Li-ion Battery Performance," in *2024 Third International Conference on Electrical, Electronics, Information and Communication Technologies (ICEEICT)*, Jul. 2024, pp. 1–7.
- [2] D. Ge, G. Jin, J. Wang, and Z. Zhang, "A novel data-driven IBA-ELM model for SOH/SOC estimation of lithium-ion batteries," *Energy*, vol. 305, p. 132395, Oct. 2024.
- [3] C. She, Y. Shen, G. Bin, J. Hong, and Y. Peng, "Accurate State of Health Estimation of Battery System based on Multi-stage Constant Current Charging and Behavior Analysis in Real-world Electric Vehicles," *IEEE Trans. Transport. Electric.*, pp. 1–1, 2024.
- [4] P. Li, Y. Cheng, K. Shan, and Y. Fang, "State-of-health estimation of lithium-ion battery based on feature transfer learning," in *2022 34th Chinese Control and Decision Conference (CCDC)*, Aug. 2022, pp. 889–894.
- [5] L. Cai, N. Cui, H. Jin, J. Meng, S. Yang, J. Peng, *et al.*, "A unified deep learning optimization paradigm for lithium-ion battery state-of-health estimation," *IEEE Trans. Energy Convers.*, vol. 39, no. 1, pp. 589–600, Mar. 2024.
- [6] H. Dai, J. Wang, Y. Huang, Y. Lai, and L. Zhu, "Lightweight state-of-health estimation of lithium-ion batteries based on statistical feature optimization," *Renewable Energy*, vol. 222, p. 119907, Feb. 2024.
- [7] X. Ding, H. Chen, X. Zhang, J. Han, and G. Ding, "RepMLPNet: Hierarchical Vision MLP with Re-parameterized Locality," Mar. 30, 2022, *arXiv: arXiv:2112.11081*.
- [8] P. K. A. Vasu, J. Gabriel, J. Zhu, O. Tuzel, and A. Ranjan, "FastViT: A Fast Hybrid Vision Transformer Using Structural Reparameterization," presented at the Proceedings of the IEEE/CVF International Conference on Computer Vision, 2023, pp. 5785–5795. Accessed: Oct. 26, 2024.
- [9] T. Akiba, S. Sano, T. Yanase, T. Ohta, and M. Koyama, "Optuna: A Next-generation Hyperparameter Optimization Framework," in *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, in KDD '19. New York, NY, USA: Association for Computing Machinery, Jul. 2019, pp. 2623–2631.
- [10] Y. Zhang and Y.-F. Li, "Prognostics and health management of Lithium-ion battery using deep learning methods: A review," *Renewable Sustainable Energy Rev.*, vol. 161, p. 112282, Jun. 2022.
- [11] E. Zamfir, M. V. Conde, and R. Timofte, "Towards Real-Time 4K Image Super-Resolution," in *2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, Vancouver, BC, Canada: IEEE, Jun. 2023, pp. 1522–1532.
- [12] J. Hu, L. Shen, and G. Sun, "Squeeze-and-Excitation Networks," presented at the Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018, pp. 7132–7141. Accessed: May 15, 2024.
- [13] Q. Wang, B. Wu, P. Zhu, P. Li, W. Zuo, and Q. Hu, "ECA-Net: Efficient Channel Attention for Deep Convolutional Neural Networks," Apr. 07, 2020, *arXiv: arXiv:1910.03151*. Accessed: Dec. 14, 2023.
- [14] G. Orzech, "Prognostics Center of Excellence Data Set Repository," NASA. Accessed: Mar. 23, 2023.
- [15] C. Birkl, "Diagnosis and prognosis of degradation in lithium-ion batteries," <http://purl.org/dc/dcmitype/Text>, University of Oxford, 2017. Accessed: May 07, 2024.
- [16] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, NV, USA: IEEE, Jun. 2016, pp. 770–778.