



OPEN ACCESS

EDITED BY

Rui Long,
Huazhong University of Science and
Technology, China

REVIEWED BY

Shunli Wang,
Southwest University of Science and
Technology, China
Zhongwei Deng,
University of Electronic Science and
Technology of China, China

*CORRESPONDENCE

Tongtong Gao,
✉ gao_tongtong@fj.sgcc.com.cn

RECEIVED 11 December 2023

ACCEPTED 22 January 2024

PUBLISHED 07 February 2024

CITATION

Li Z, Chen J, Gao T, Zhang W, Chen D and
Gu Y (2024), Cross-scenario capacity
estimation for lithium-ion batteries via
knowledge query domain mixing-up network.
Front. Energy Res. 12:1353651.
doi: 10.3389/fenrg.2024.1353651

COPYRIGHT

© 2024 Li, Chen, Gao, Zhang, Chen and Gu.
This is an open-access article distributed
under the terms of the [Creative Commons
Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use,
distribution or reproduction in other forums is
permitted, provided the original author(s) and
the copyright owner(s) are credited and that
the original publication in this journal is cited,
in accordance with accepted academic
practice. No use, distribution or reproduction
is permitted which does not comply with
these terms.

Cross-scenario capacity estimation for lithium-ion batteries via knowledge query domain mixing-up network

Zhicheng Li¹, Jinyu Chen¹, Tongtong Gao^{1*}, Weijun Zhang¹,
Dawei Chen¹ and Yi Gu²

¹State Grid Fujian Electric Power Research Institute, Fuzhou, Fujian, China, ²State Key Laboratory of Electrical Insulation and Power Equipment, Xi'an Jiaotong University, Xi'an, Shanxi, China

Introduction: Deep learning has demonstrated exceptional prowess in estimating battery capacity. However, its effectiveness is often compromised by performance degradation under a consequence of varying operational conditions and diverse charging/discharging protocols.

Methods: To tackle this issue, we introduce the Knowledge Query Domain Mixing-up Network (KQDMN), a domain adaptation-based solution adept at leveraging both domain-specific and invariant knowledge. This innovation enriches the informational content of domain features by segregating the functions of feature extraction and domain alignment, enhancing the efficacy of KQDMN in utilizing diverse knowledge types. Moreover, to identify time-deteriorating features in battery time series data, we employ convolutional operations. These operations are pivotal in extracting multi-scale features from the battery's characteristic curves. Inspired by the Transformer model, we have developed a set of knowledge queries that integrate these multi-scale features seamlessly, thereby enabling extensive global feature extraction. To ensure the retention of domain-specific information, we have instituted two independent feature extraction pathways. Pursuing domain-invariant knowledge, this study introduces cross-attention as a mechanism to connect two domain spaces, effectively diminishing the disparity between source and target distributions.

Results and Discussion: This approach is crucial for accurately estimating capacity in batteries with diverse performance characteristics. The practicality and robustness of the proposed method are validated using the MIT battery aging dataset, yielding highly satisfactory outcomes. The Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and Coefficient of Determination (R^2) for our capacity estimation process are 0.19%, 0.23%, and 0.997, respectively, highlighting the precision and reliability of our approach.

KEYWORDS

lithium-ion battery, capacity estimation, domain adaptation, cross attention, transfer learning

1 Introduction

In recent years, global attention has increasingly focused on environmental pollution and the looming energy crises. Traditional fossil fuels are confronting depletion and pollution issues, posing significant threats to both human health and the environment. In response to these pressing challenges, lithium batteries have emerged as a favored solution, owing to their inherent advantages such as high energy density, extended cycle life, and environmental friendliness. However, due to the influence of factors like charge/discharge cycles and usage environments, electrochemical reactions and physical alterations result in irreversible changes, gradually deteriorating the electrical performance of these batteries. This underscores the critical importance of capacity estimation in battery management. It serves as a pivotal aspect, providing vital insights into the battery's health and remaining useful life. Accurate capacity estimation, in turn, fosters enhanced battery utilization, prolongs battery lifespan, and elevates system reliability. As battery technology advances rapidly and its applications continue to expand, numerous researchers have been energetically investigating and researching methods for predicting battery capacity. These efforts have yielded significant progress and results of note. These methods can be categorized into two main groups: model-based and data-driven approaches.

Model-based methods are utilized to delve into the intricate connection between internal electrochemical mechanisms and external conditions, offering the advantages of real-time and closed-loop feedback. These methods involve constructing a discrete expression model based on material properties, electrochemical reactions, and impedance changes to depict the capacity aging process. Consequently, these models facilitate the recursive estimation of the battery's state. The commonly used models can be roughly classified into three types: electrochemical models (EM) (Corno et al., 2014; Lyu et al., 2017), equivalent circuit models (ECM) (Guha and Patra, 2018; Hu et al., 2012), and electrochemical impedance models (EIM) (Alavi et al., 2015; Bartlett et al., 2015) presented a reduced-order electrochemical model to predict lithium concentration and current split in the electrode, which is then used in dual-nonlinear observers to estimate cell SOC and capacity. Using the fundamental electrochemical model, Samadi et al. (2012) presented a particle filter algorithm for state estimation and condition monitoring of Li-ion batteries. The algorithm can effectively obtain the spatial distribution of field variables in a battery. Utilizing an open circuit voltage (OCV) correction strategy, Liu et al. (2023) presented a specialized method for calculating the capacity of electric vehicles (EVs). This method establishes two degradation models for battery capacity, incorporating mileage and service life as variables. Bi et al. (2016) proposed a new method for state-of-health (SOH) dynamic estimation of power battery packs in battery electric vehicles (BEVs) using a genetic resampling particle filter (GPF). A second-order equivalent circuit model of the Resistance-Capacitance (RC) circuit was developed to identify unknown parameters, and a state-space model of GPF was used to solve the non-Gaussian problem. With adaptively obtained initial conditions, the complex-nonlinear-least-squares algorithm was used to fit pseudo-random-sequence impedance measurements to the ECM, and the ECM parameters could reflect battery capacity reduction (Sihvo et al., 2020; Zhang et al., 2022) developed a series

electrochemical impedance model. Which has high computational efficiency with simplifications on the lithium-ion transport path and the electrolyte diffusion. While model-based methods have made substantial advancements in recent years, they hinge on a profound comprehension of the battery's electrochemical processes and aging mechanisms. This reliance amplifies the challenges and intricacies of modeling, primarily due to the complex interplay of physical and chemical processes, noise factors, and the diverse environmental conditions affecting batteries. Additionally, this reliance may constrain the applicability of these models within certain contexts.

Data-driven methods involve the analysis of historical data, including parameters like current, voltage, capacitance, and impedance, to identify features associated with battery degradation state. These methods subsequently construct predictive models for assessing battery health and degradation. Unlike model-based approaches, data-driven methods do not necessitate the creation of explicit mathematical models to describe the progression of battery degradation. Zhu et al. (2022) extracted statistical features from the voltage profile after full charging and used XGBoost, ElasticNet, and support vector regression models to estimate the battery capacity. The best-performing model achieved a RMSE of 0.011. Wang et al. (2022) analyzed the aging characteristics of batteries from several perspectives and extracted health indicators (HI) from battery charging and discharging curves. Based on a novel integrated Gaussian process regression (GPR) model, satisfactory estimation results were achieved, reaching 0.0241 RMSE on the NASA public data set. Because of their powerful ability to process historical operation data and time series information, recurrent neural networks are widely used for battery state prediction. Qu et al. (2019) combined long short-term memory (LSTM) networks and attention mechanisms and adopted fully integrated empirical mode decomposition with adaptive noise to denoise the raw data, which improved SOH estimation accuracy. Based on an improved dual closed-loop observation modeling strategy, Wang et al. (2023b) proposed an improved anti-noise adaptive long short-term memory (ANA-LSTM) neural network with high-robustness feature extraction and optimal parameter characterization, aiming to achieve accurate RUL prediction. Wang et al. (2023c) proposed an improved robust multi-time scale singular filtering-Gaussian process regression-long short-term memory (SF-GPR-LSTM) modeling method to estimate the remaining capacity of lithium-ion batteries. Particularly, this approach establishes a theoretical foundation for estimating the remaining capacity of batteries throughout their entire life cycle in extremely low temperatures. Ungurean et al. (2020) proposed a gated recursive unit (GRU)-based neural network structure for online prediction of battery SOH. With automatic feature extraction and a low risk of overfitting, convolutional neural networks (CNN) show great potential for battery capacity. To further improve the accuracy of SOH estimation, Yanwen Dai (2022) selected health feature parameters from lithium-ion battery charging curves using Spearman correlation coefficients and adopted CNN to extract local features of health features and LSTM to mine time series features, respectively. Chemali et al. (2022) developed a convolutional neural network (CNN)-based framework for SOH estimation, which takes time domain voltage, current, and temperature measurements as inputs. In order to simulate a real-world system, the data was

augmented with some noise and error. Thus, estimation accuracy was improved. Li et al. (2022) combined CNN and LSTM to design a hybrid neural network to capture the hierarchical features and the temporal dependencies. Due to the powerful ability to extract valuable and important information, attention mechanisms have been widely used in battery SOH estimation. Zhao et al. (2023) presented a hybrid attention neural network, extracted the health features from differential temperature curves, and achieved superior prediction performance. Ji et al. (2023) proposed knee-point probability to detect knee points and improve state-of-health (SOH) predictive accuracy, and an attention mechanism was utilized for extracting features. To reveal the correlation among features, Wei and Wu (2023) leveraged the graph convolutional networks (GNN) with different attention mechanisms to predict battery SOH. Yao et al. (2023) utilized the Pearson's correlation coefficients to extract highly correlated HIs, built the graph sample aggregate to propagate messages and uncover the deep information among HIs. Taking the advantages of both CNNs and Transformers, Gu et al. (2023) realized the battery SOH estimation with high accuracy and stability. Pham et al. (2022) took advantage of self-attention and fixed-point positional encoding to capture the sequence nature of battery data, and experimental results showed significant accuracy.

Data-driven methods have achieved remarkable success in battery capacity estimation. However, such achievement relies heavily on a huge amount of labeled training data, which is difficult to obtain in many real-world applications. In addition, these approaches are based on the assumption that the training and test datasets share the independently identically distribution, which ensures that the model developed based on the training sets can be directly applied to the test sets. However, in many practical scenarios, the two domains inevitably drift away from each other due to various working conditions and different charge/discharge protocols. Therefore, the domain shift issue results in the performance degradation of the model trained on the source domain.

In recent years, transfer learning and domain adaptation have been utilized to transfer the domain invariant knowledge learned from the source domain and minimize the domain gap. During the past decade, many researchers have extensively investigated domain adaptation methods. The typical methods can be divided into two categories: adversarial training (Ganin et al., 2016; Wei et al., 2021; Zhang et al., 2019) and distribution discrepancy metrics (Borgwardt et al., 2006; Sun and Saenko, 2016; Peng et al., 2019). Inspired by the generative adversarial network (GAN), the former category usually leverages a domain discriminator and enforces a feature extractor to learn domain invariant representation to fool the discriminator in an adversarial manner. The latter category explicitly reduces the domain discrepancy based on some distribution distance metrics. Deng et al. (2024) developed a deep convolutional neural network for estimating battery state of health. The method employs distribution distance metrics to mitigate the distribution discrepancy between the source domain and target domain. Distribution distance metrics encompass the Maximum Mean Discrepancy (MMD) along with other second or higher-order statistics such as correlation alignment (CORAL).

This paper introduces a domain adaptation-based method aimed at achieving accurate capacity estimation to address the domain shift challenge in battery capacity prediction. Importantly,

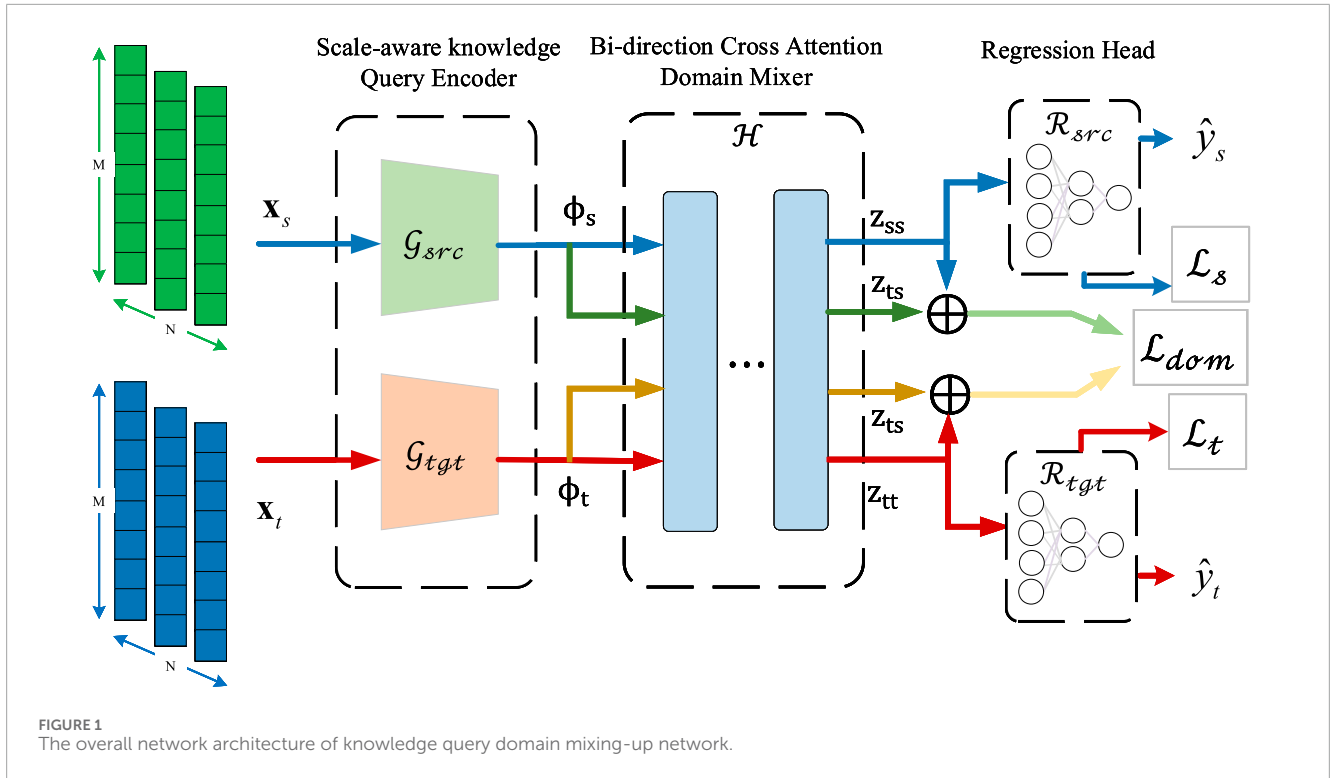
recognizing the limitations of true label supervision in the target domain, our approach promotes the sharing of domain knowledge while suppressing domain-specific features. However, this emphasis on alignment has the potential to constrain feature expressiveness within each domain, resulting in a bias towards the source domain and compromising the effectiveness of knowledge transfer. To mitigate this, we implement two separate pathways for individual feature extraction, allowing us to better preserve domain-specific characteristics. Furthermore, in order to enhance knowledge transferability and minimize domain discrepancies, we utilize the cross-attention mechanism to achieve bidirectional domain alignment. This mechanism acts as a bridge between different domains and implicitly facilitates domain mixing. In summary, the key contributions of this work are outlined as follows:

- We introduce a novel domain adaptation-based method for estimating battery capacity that effectively distinguishes between feature extraction and domain alignment. This approach allows for the strategic leveraging of both domain-specific and invariant knowledge, enhancing the precision of our estimations.
- To comprehensively preserve domain-specific information, we have developed two independent feature extraction pathways. Employing Convolutional Neural Networks (CNNs), we adeptly extract features indicative of temporal degradation. Moreover, through the use of self-attention mechanisms, we are able to distill global and diverse insights from these extracted features.
- In the quest to acquire domain-invariant knowledge, we implement cross-attention as an innovative bridging mechanism. This approach significantly reduces the disparity between source and target distributions, thus enabling us to predict battery capacity effectively in a wide range of scenarios.

The paper is organized as follows. In Section 2, the overall framework and the preliminaries of our methodology are presented. Section 3 will discuss the experimental details. The experimental results are given in Section 4. The discussions are given in Section 5. Finally, Section 6 gives the conclusion.

2 Methodology

For the capacity estimation, all the features are embedded in the sequence of charging or discharging curves. The relationship between battery V/I/T curves and battery health status is very difficult to establish in the full battery lifetime. The inter-cycle features are the sequences of voltage, current and temperature (V/I/T) curves in the battery charging/discharging processes measured by BMS within different cycles, which typically include $\mathbf{V} = \{V_1, V_2, \dots, V_N\}$, $\mathbf{I} = \{I_1, I_2, \dots, I_N\}$, $\mathbf{T} = \{T_1, T_2, \dots, T_N\}$, in the i th cycle. Although the BMS streaming are recorded at the same sampling rate, the total time duration could vary for different batteries and in different cycles. Given that the charged and discharged capacity increases monotonously, the battery V/I/T curves are normalized as the functions of capacity ratio. To account for differences in data volume and distribution arising from various operational conditions encountered in practical battery applications,



we categorize the battery data into the source data x_s and target data x_t . x_s consist of inter-cycle features from batteries operating under standard conditions, including current **I**, voltage **V** and temperature **T**. On the other hand, x_t consists of inter-cycle features from batteries operating under unique conditions, characterized by limited data availability and significant distribution variations in current **I**, voltage **V** and temperature **T**.

2.1 Overall architecture

The comprehensive network architecture is depicted in **Figure 1**. As observed, this architecture comprises three key components: the Scale-aware Knowledge Query Encoder, the Bi-directional Cross Attention Domain Mixer, and the Regression Head. The Scale-aware Knowledge Query Encoder receives raw data as input and is employed to generate domain-specific state-related embeddings. Since each domain provides battery timing data from different feature spaces and different aging patterns, this paper sets up private feature encoders for the source and target domains, which are denoted as G_{src} and G_{tgt} , respectively. The source and target data (x_s, x_t) are input into the encoders to yield the domain-specific features (ϕ_s, ϕ_t), respectively. The operation of domain-specific features can be expressed as Eq. 1.

$$\begin{aligned} \phi_s &= G_{src}(x_s) \\ \phi_t &= G_{tgt}(x_t) \end{aligned} \tag{1}$$

where G_{src} and G_{tgt} denote the feature encoders for the source and target domains, x_s and x_t denote the source and target data, ϕ_s and ϕ_t denote the domain-specific features.

Afterward, they are fed into the Bi-direction Cross Attention Domain Mixer to alleviate the domain gap and extract the domain

invariant feature. The operation of domain invariant feature can be expressed as Eq. 2.

$$(z_{ss}, z_{ts}, z_{ts}, z_{tt}) = \mathcal{H}(\phi_s, \phi_t) \tag{2}$$

where $\mathcal{H}(\bullet)$ represent the functions of the Bi-direction Cross Attention Domain Mixer, ($z_{ss}, z_{ts}, z_{ts}, z_{tt}$) denote the domain invariant feature.

In the end, the domain invariant features (z_{ss}, z_{tt}) are passed through the Regression Head to produce the estimation results (\hat{y}_s, \hat{y}_t), which is implemented as a 2-layer fully connected neural network. The operation of estimation results can be expressed as Eq. 3.

$$\begin{aligned} \hat{y}_s &= \mathcal{R}_{src}(z_{ss}) \\ \hat{y}_t &= \mathcal{R}_{tgt}(z_{tt}) \end{aligned} \tag{3}$$

where $\mathcal{R}(\bullet)$ represent the functions of the Regression Head, (\hat{y}_s, \hat{y}_t) denote the estimation results.

2.2 Domain-wise scale-aware knowledge query encoder

The domain-wise Scale-aware Knowledge Query Encoder is devoted to learning domain-specific feature space and modeling domain discriminative representation separately. Therefore, we use two private streams to learn domain-wise embedding functions G_{src} and G_{tgt} . Inspired by the powerful representation abilities of CNNs, such as parameter sharing, local information aggression, and spatial reduction, 1D CNN is adopted as the backbone to extract high-order features. As shown in **Figure 2**, the encoder consists of L

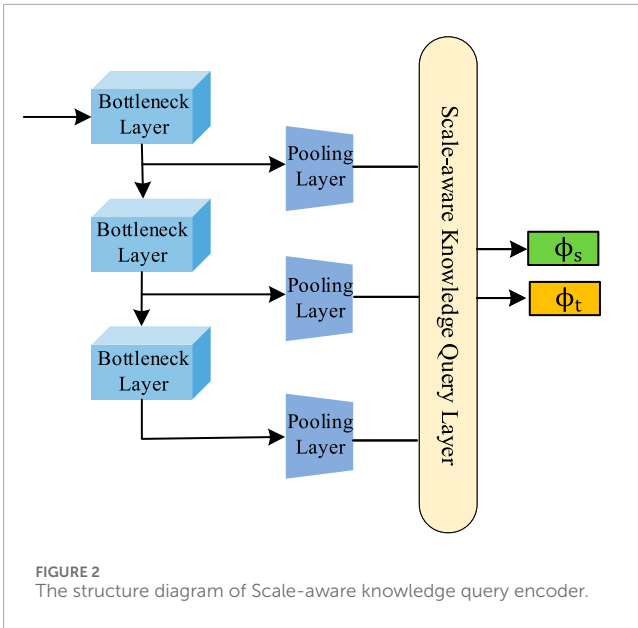


FIGURE 2 The structure diagram of Scale-aware knowledge query encoder.

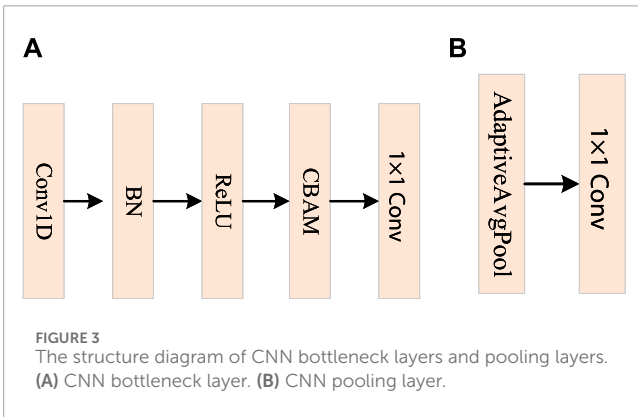


FIGURE 3 The structure diagram of CNN bottleneck layers and pooling layers. (A) CNN bottleneck layer. (B) CNN pooling layer.

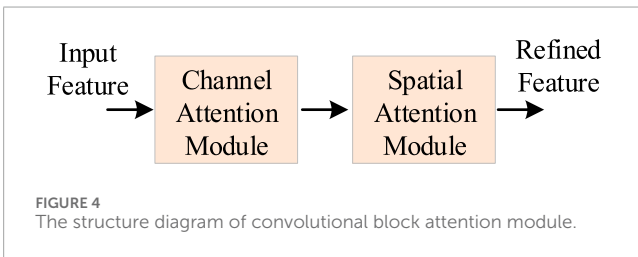


FIGURE 4 The structure diagram of convolutional block attention module.

stacked CNN bottleneck layers, L pooling layers, and one scale-aware knowledge query layer.

Given the source and target domain data $x_s, x_t \in \mathbb{R}^{M \times N}$, we adopt the L stacked CNN bottleneck layers to successively and alternatively extract features. As shown in Figure 3A, the CNN bottleneck layer comprises a one-dimensional convolution, Convolutional Block Attention Module (CBAM) block, and a point-wise convolution block. As shown in Figure 4, the CBAM comprises a channel attention module and a spatial attention module.

One-dimensional convolution is a method that conducts a sliding window convolution operation on signals, sequences, or time series data. This approach is highly effective in extracting features from the data, thereby enhancing the ability of machine learning

algorithms to comprehend and analyze the information. Specifically, suppose the size of the input feature map X is $C_{in} \times L$, the size of the convolution kernel W is $C_{out} \times C_{in} \times H$, and the size of the output feature map Y is $C_{out} \times L'$. The 1D convolution is calculated as Eq. 4.

$$Y_{k,i} = \sum_{c=1}^{C_m} \sum_{j=0}^{H-1} W_{k,c,j} \cdot X_{c,i+j} \quad (4)$$

where m denotes the length of the convolution kernel, k denotes the channel position in the output feature map Y , c denotes the channel position in the input feature map X , $X_{c,i}$ denotes the value of the c th channel in the input feature map X at position i and $Y_{k,i}$ denotes the pixel value of the k th channel in the output feature map Y at position i .

The Convolutional Block Attention Module (CBAM) is an attention mechanism utilized in computer vision. It is incorporated into convolutional neural networks to enhance their representation capabilities and bolster the robustness and generalization performance of features. The CBAM module comprises two key sub-modules: the Channel Attention Module and the Spatial Attention Module. The Channel Attention Module primarily serves to extract relationships between channels. It accomplishes this by globally pooling the feature maps from each channel to obtain a global feature vector. This vector is then fed into a fully connected network to learn the inter-channel relationships. Subsequently, the acquired weights are applied to the feature maps of each channel, effectively weighing the contributions of each channel on average. On the other hand, the Spatial Attention Module is focused on extracting spatial relationships. It employs various filters in the spatial dimension to learn features at different scales. Following this, a global feature vector is generated through the global pooling of these features, enabling the modeling of spatial relationships. Ultimately, the acquired weights are applied to features at individual spatial locations, effectively weighting the average of features at each location.

Point-wise convolution is also known as 1×1 convolution, whose convolution kernel has a size of 1. It is employed for feature fusion to aggregate different levels of feature maps and adjust the number of channels for the input feature map. The size of the convolution kernel W is $C_{out} \times C_{in}$. The operation of the point-wise convolution can be expressed as Eq. 5.

$$Y_{k,i} = \sum_{c=1}^{C_m} W_{k,c} \cdot X_{c,i} \quad (5)$$

To effectively address the issues of gradient vanishing and exploding, expedite convergence, and simultaneously enhance generalization capabilities, Batch Normalization (BatchNorm1d) is applied in 1D convolutional neural networks. The primary role of BatchNorm1d is to normalize each feature channel within small input data batches. This normalization process aims to bring the mean of the feature channel close to 0 and the variance close to 1, thereby accelerating the model's convergence. To further augment the network's nonlinear mapping capacity, the Rectified Linear Unit (ReLU) is employed as the activation function for all convolution layers. This choice of activation function is known to promote faster convergence compared to alternative functions.

To align the feature map size of the CNN encoder with that of the scale-aware knowledge query block, this paper uses the

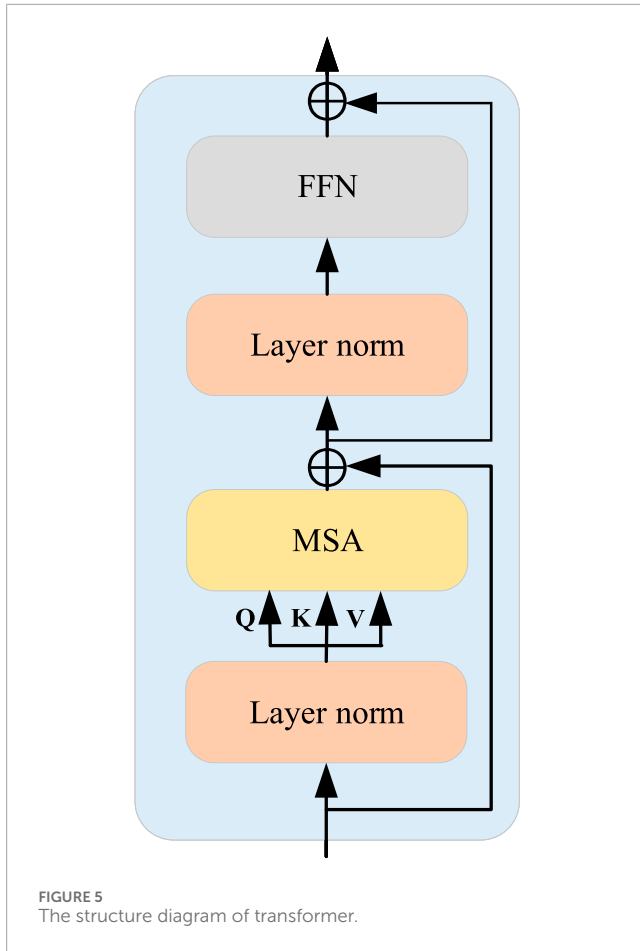


FIGURE 5 The structure diagram of transformer.

CNN pooling layer to transform the dimensionality of the output features from the CNN bottleneck layer. As shown in Figure 3B, AdaptiveAvgPool1d and point-wise convolution are combined to adapt to the length of the input tensor and reduce the number of channels. These two-dimensional vectors are then flattened into one-dimensional vectors along the channel. Following those steps, feature tokens $\{T_f^1, \dots, T_f^L\} \in \mathbb{R}^{L \times s}$ are extracted, where L is total number of time scales.

As one of the most important tools in the field of natural language processing, the transformer has excellent performance and generalization capabilities. The transformer is composed of Multi-Head Self Attention (MSA), Layer Normalization(LN), and Feed Forward Network (FFN). Suppose an input sequence X is given, which contains n vectors x_1, x_2, \dots, x_n , where the dimension of each vector is d . As shown in Figure 5, the output of the transformer is computed as Eq. 6.

$$Y' = MSA(LN(X)) + LN(X),$$

$$Y = FFN(LN(Y')) + LN(Y'). \tag{6}$$

where MSA , LN and FFN represent the functions of the Multi-Head Self Attention, the Layer Normalization and the Feed Forward Network, respectively.

Multi-head attention is able to generate a more accurate sequence representation by interacting within the sequence and capturing the relationships between different positions. For each vector x_i , we need to calculate its similarity to the other vectors with

scaled dot-product. The weighted sum of the value can be expressed as Eq. 7.

$$Q = XW_Q, \quad K = XW_K, \quad V = XW_V,$$

$$MSA(X) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \tag{7}$$

where Q, K , and V denote the query, key, and value vector matrices of the input sequence, $W_Q \in \mathbb{R}^{d_k \times d}$, $W_K \in \mathbb{R}^{d_k \times d}$, and $W_V \in \mathbb{R}^{d_v \times d}$ are learnable parameters, and typically d_v is equal to d_k .

Another noteworthy feature of the Transformer architecture is its ability to flexibly incorporate learnable classification tokens for various purposes. To encourage the features for alignment to gather global and diverse information in feature tokens $\{T_f^1, \dots, T_f^L\}$, we set a group of learnable tokens $\{T_k^1, \dots, T_k^N\} \in \mathbb{R}^{N \times s}$ as knowledge queries. These knowledge queries are concatenated with multi scale features as $\{T_k^1, \dots, T_k^N, T_f^1, \dots, T_f^L\} \in \mathbb{R}^{(N+L) \times s}$. These feature tokens will forward propagate into the transformer, and N features that correspond to the knowledge queries are utilized as domain-wise state knowledge, which is denoted as $\phi_{(\bullet)} = \{K_{(\bullet)}^1, \dots, K_{(\bullet)}^N\}$ where (\bullet) represents the source domain or the target domain.

2.3 Bi-direction cross attention domain mixer

While the previously mentioned self-attention module effectively harnesses domain-specific knowledge, it does not explore domain-invariant features. It is widely acknowledged that cross-attention enables information to flow between different sequences or tensors and can model relationships between diverse inputs. Consequently, it significantly enhances the accuracy of information transfer. As such, cross-attention is well-suited for mitigating domain discrepancies and acquiring domain-invariant representations.

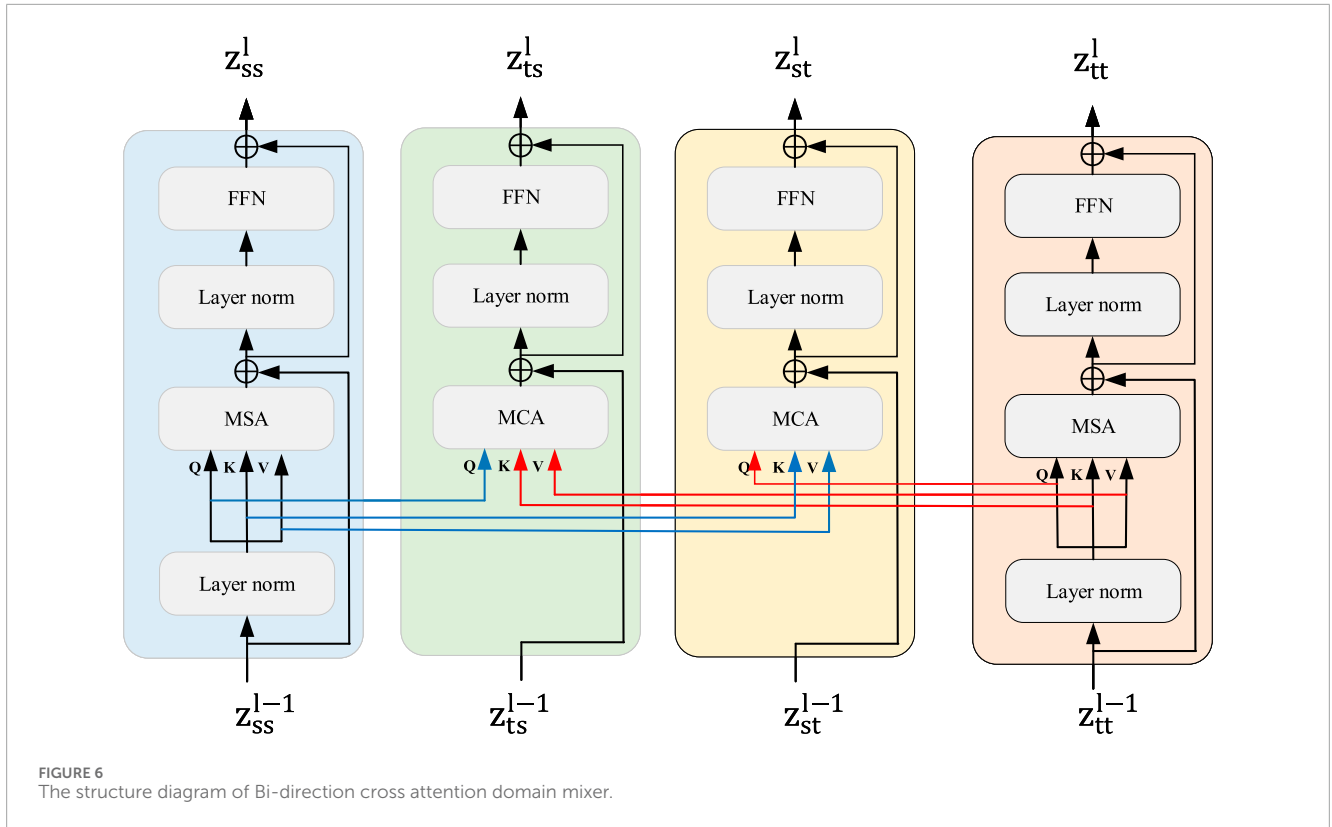
Cross-attention and self-attention are both attention mechanisms, but they are applied in slightly different scenarios and involve distinct computational methods. Self-attention primarily handles internal dependencies within a single input sequence or tensor. In contrast, cross-attention is predominantly employed to manage interactions between multiple input sequences or tensors. The computation of Multi-head Cross Attention (MCA) involves modeling the interaction between two input sequences or tensors as an attention distribution, which can be formulated as Eq. 8.

$$Q = SW_Q, \quad K = FW_K, \quad V = FW_V,$$

$$MCA(S, F) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \tag{8}$$

where S and F can be state knowledge from either the source domain or the target domain.

Here, this paper does not just keep ϕ_{src} unchanged and keep ϕ_{tgt} close to ϕ_{src} , but achieves domain alignment in a bidirectional way to transfer knowledge from the source domain. In this paper, multi-headed cross-attention is used as a bridge to connect different domains to achieve a mixture of different domain feature spaces. It can be seen from Figure 6 that four weight-sharing transformers are



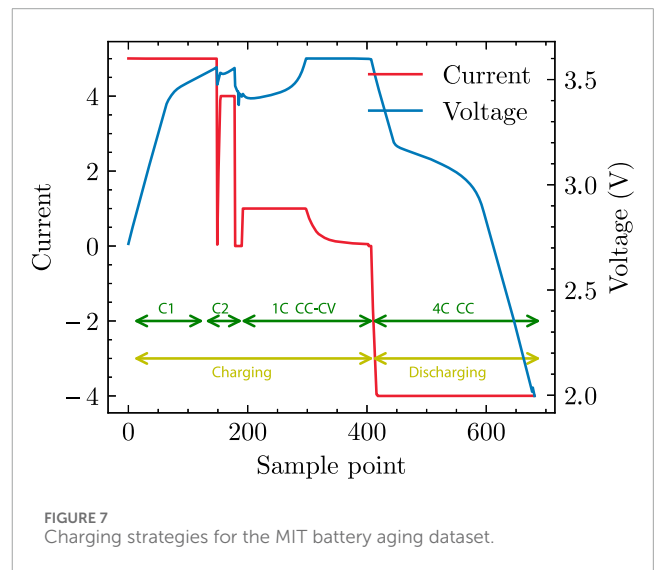
utilized to realize bi-direction domain alignment. State knowledge ϕ_{src} and ϕ_{trg} are fed into transformers with Multi-Head Self Attention to extract feature representation z_{ss} and z_{ts} , independently. In contrast, transformers with Multi-head Cross Attention are utilized to bring each domain feature representation closer to the other. Those can be formulated as Eq. 9.

$$\begin{aligned}
 \hat{z}_{ss}^{l-1} &= MSA(LN(z_{ss}^{l-1})) + LN(z_{ss}^{l-1}), \\
 z_{ss}^l &= FFN(LN(\hat{z}_{ss}^{l-1})) + LN(\hat{z}_{ss}^{l-1}), \\
 \hat{z}_{ts}^{l-1} &= MCA(z_{ss}^{l-1}, z_{tt}^{l-1}) + z_{ts}^{l-1}, \\
 z_{ts}^l &= FFN(LN(\hat{z}_{ts}^{l-1})) + LN(\hat{z}_{ts}^{l-1}), \\
 \hat{z}_{st}^{l-1} &= MCA(z_{tt}^{l-1}, z_{ss}^{l-1}) + z_{st}^{l-1}, \\
 z_{st}^l &= FFN(LN(\hat{z}_{st}^{l-1})) + LN(\hat{z}_{st}^{l-1}), \\
 \hat{z}_{tt}^{l-1} &= MSA(LN(z_{tt}^{l-1})) + LN(z_{tt}^{l-1}), \\
 z_{tt}^l &= FFN(LN(\hat{z}_{tt}^{l-1})) + LN(\hat{z}_{tt}^{l-1})
 \end{aligned} \tag{9}$$

where l represents the layer index of transformer block. And z_{ss}^0 and z_{ts}^0 are initialized to state knowledge ϕ_{src} , while z_{tt}^0 and z_{st}^0 are initialized to state knowledge ϕ_{trg} .

It is worth noting that only z_{ss} and z_{tt} are employed and fed into Regression Heads for final regression, respectively. To fully explore the supervised information in each domain, we employ a mean squared error (MSE) loss function to minimize the regression error of all labeled samples, which is calculated as Eq. 10.

$$\begin{aligned}
 \mathcal{L}_s &= \frac{1}{n_s} \sum_{i=1}^{n_s} \mathcal{E}(y_s, \hat{y}_s), \\
 \mathcal{L}_t &= \frac{1}{n_t} \sum_{i=1}^{n_t} \mathcal{E}(y_t, \hat{y}_t)
 \end{aligned} \tag{10}$$



where \mathcal{L}_s and \mathcal{L}_t denote the loss of the source and target Regression Head, y_s and y_t denote the true values, \hat{y}_s and \hat{y}_t denote the output of Regression Heads. $\mathcal{E}(\bullet, \bullet)$ denotes the predicted values, which is calculated as Eq. 11.

$$\mathcal{E}(y, \hat{y}) = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \tag{11}$$

where y denotes the true values, \hat{y} denotes the predicted values.

During the training stage, the combined features $\{z_{ss}, z_{ts}\}$ and $\{z_{tt}, z_{st}\}$ are utilized to reduce the domain discrepancy. Without

introducing the domain discriminator, distribution discrepancy metrics are employed to minimize the domain discrepancy between each pair of source and target domains. MMD is a widely used representative distribution discrepancy metric in domain adaptation, given a source features F_s , and a target features F_t , the MMD loss is defined as Eq. 12.

$$\mathcal{L}_{sup} = \mathcal{L}_s + \mathcal{L}_t \quad (12)$$

where \mathcal{L}_{sup} denotes the MMD loss.

But it only calculates first-order moments of inter-domain distance. Therefore, the correlation alignment (CORAL) loss is adopted, which measures the difference between the second-order statistics (covariance matrices) of the source and target domains. More formally, given a source features F_s , and a target features F_t , the CORAL loss is defined as Eq. 13.

$$\mathcal{L}_{dom} = \left| \frac{1}{n_s} \mathcal{X}_s \mathcal{X}_s^T - \frac{1}{n_t} \mathcal{X}_t \mathcal{X}_t^T \right|^2 \quad (13)$$

where \mathcal{L}_{dom} denotes the CORAL loss, \mathcal{X}_s and \mathcal{X}_t denote the features of the source and target domains respectively.

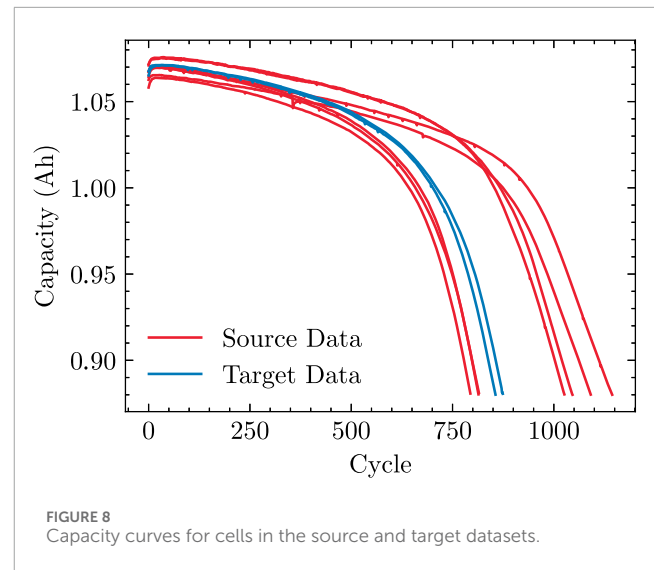
3 Experimental

3.1 Dataset description

The MIT Battery Aging Dataset is utilized to validate our proposed method, developed at MIT for battery aging prediction studies. The dataset consists of 124 commercial lithium-iron phosphate/graphite A123 APR18650M1A batteries with a nominal capacity of 1.1 amp-hours. As shown in Figure 7, all batteries were charged and discharged at constant current using a multi-step fast charging strategy "C1(S1)-C2" in a 30°C temperature chamber. The cells were first charged with current C1 until the state of charge (SOC) reached S1, then with current C2 until SOC reached 80%, and finally, the cells were charged from 80% to 100% SOC using a 1C rate (CC-CV) to a cut-off voltage of 3.6 V. These cells were discharged to 2.0 V at a constant current rate of 4C. Each battery data set captures full life cycle data from when the battery is new to when it fails, a time span ranging from a few days to several months. Each battery in the data set is measured at 10-s intervals, recording parameters such as battery voltage, current, and temperature, as well as information about the battery's capacity. Battery life is defined as the number of cycles in which the capacity of a battery is reduced to 80% of its nominal capacity. One hundred twenty-four battery samples have a life span ranging from 150 cycles to 2300 cycles.

In order to verify the effectiveness and superiority of the proposed domain adaptive method in predicting the capacity of Li-ion batteries under different data distribution and insufficient data scenarios, the seven batteries A1-A7 with the charging strategy "5.6C19%-4.6C" are used as the source datasets and two batteries B1-B2 with the charging strategy "5.9C15%-4.6C" are selected as the target datasets. Figure 8 shows the capacity decay curves for these batteries.

The voltage, current, and temperature profiles of a battery serve as indicators of its aging level. Generally, as a battery undergoes extended usage, its aging degree progressively increases, leading to



diminished performance, including voltage reduction and a decrease in current output capacity. Consequently, by scrutinizing the voltage, current, and temperature profiles, it becomes possible to assess the battery's aging level. Since the length of these distinct characteristic profiles varies with the aging degree, this paper employs a linear interpolation technique to resample and standardize these profiles to a fixed length. All the feature curves are stitched together to form an array $\mathbf{x} \in \mathbb{R}^{M \times N}$, where M is the number of feature curves, and N is the fixed length of each feature curve.

3.2 Data preprocessing

Various features typically exhibit different magnitudes, and issues stemming from these magnitude disparities often arise when using raw data for analysis and modeling. To enhance the stability of the model training process, this paper employs MinMax normalization. MinMax normalization serves to mitigate the magnitude of discrepancies among diverse features, thereby enhancing the efficiency, stability, and accuracy of the algorithm. The calculation formula for MinMax normalization is as Eq. 14.

$$\mathbf{x}' = \frac{\mathbf{x} - \min(\mathbf{x})}{\max(\mathbf{x}) - \min(\mathbf{x})} \quad (14)$$

where \mathbf{x} is the original data and \mathbf{x}' is the normalized data.

Substantial noise and irregularities within the data can result in overfitting, thereby impairing the model's generalization capabilities. In this study, outliers are substituted with the average of neighboring data points, and we employ Savitzky-Golay filtering for data smoothing. Savitzky-Golay filtering is a technique based on polynomial fitting aimed at noise reduction. It achieves this by fitting a sequence of data points and optimizing for various window sizes to minimize noise while retaining the inherent smoothing characteristics of the data. Since the length of different feature curves varies with age degree, these curves are resampled using linear interpolation and reshaped to the fixed length. All the feature curves are concatenated together to form an array $\mathbf{x} \in \mathbb{R}^{M \times N}$, where M is the number of feature curves, and N is the fixed length of each feature curve.

3.3 Model training

The proposed model is implemented using the PyTorch framework. We set the convolutional kernel sizes to 3 for both the CNN bottleneck layer and the CNN pooling layer, and we employ 128 convolutional kernels. The dimensionality of the model matches that of the scale-aware knowledge query block and the bidirectional cross-attention domain blender for multi-head self-attention and multi-head cross-attention. In this context, each feature token is transformed into a fixed dimensionality of 128. For multi-head self-attention and multi-head cross-attention, we use eight heads. The Regression Head comprises two fully connected layers, each with 128 and 64 neural nodes. During training, the model is trained for 200 epochs using the AdamW optimizer with a weight decay of $1e-3$ and a base learning rate of $8e-3$. We maintain a batch size of 8 for all experiments. To prevent overfitting, we apply a dropout operation with a ratio of 0.1.

Following the standard protocol for domain adaptation, we use all labeled source data and labeled target data of one battery from the target dataset for training, and the remaining batteries from the target dataset are used for performance validation. To mathematically validate the proposed method, *RMSE*, *MAE*, and R^2 are utilized to evaluate performance. The detailed formulations for these performance metrics are given as Eqs 15–17.

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (15)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (16)$$

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2} \quad (17)$$

where n is the number of samples, y_i and \hat{y}_i are the actual and predicted values, and \bar{y} denotes the mean value.

4 Results

In practical applications, there is often a discrepancy in the distributions of the training set and the test set due to differences in the internal chemistry of batteries and various operating conditions. We employ battery datasets with varied distributions to evaluate the performance of the domain adaptation-based approach proposed in this paper. The aim is to investigate whether the model can accurately estimate the capacity of target batteries even in scenarios with limited target data.

To tailor the models with various strategies, this paper introduces KQDMN-s, KQDMN-t, KQDMN-mmd, and KQDMN-coral. KQDMN-s and KQDMN-t employ single-domain supervised learning strategies using source and target datasets for training, respectively. Utilizing domain adaptive techniques, KQDMN-mmd and KQDMN-coral minimize feature differences between each source and target domain by applying the domain difference metrics MMD and CORAL, respectively. This aims to enhance predictive performance and generalization of the

TABLE 1 Statistical results of capacity estimation for target datasets B1 and B2 under different strategies.

Test cells	Strategies	MAE(%)	RMSE(%)	R^2
B1	KQDMN-s	1.56	1.87	0.825
	KQDMN-t	1.57	1.94	0.812
	KQDMN-mmd	0.64	0.76	0.971
	KQDMN-coral	0.54	0.39	0.985
B2	KQDMN-s	2.26	2.72	0.630
	KQDMN-t	1.80	1.95	0.801
	KQDMN-mmd	0.54	0.60	0.982
	KQDMN-coral	0.19	0.23	0.997

models on the target dataset. It is important to note that, for a fair comparison, this paper replaces the Multi-Channel Attention (MCA) components of the bidirectional cross-attention domain mixers in KQDMN-s and KQDMN-t with multi-headed self-attention.

Figures 9, 10 illustrate the prediction outcomes for lithium battery capacity using various strategies. These figures reveal a substantial difference in data distribution between the training and test sets, resulting in poor model performance on the test set. Consequently, KQDMN-s, which relies solely on a single-domain supervised learning strategy, exhibited the least favorable prediction performance. Similarly, KQDMN-t also struggled on the test set due to limited training data, leading to inadequate capturing of the true data distribution and resulting in under-fitting. In contrast, domain adaptive methods, such as KQDMN-mmd and KQDMN-coral, based on distribution distance metrics, displayed a notable improvement in performance on the test set. This improvement stems from their ability to leverage knowledge from the source domain data to adapt to the distribution of the target domain data. This adaptation mitigates the decline in generalization performance caused by disparities in data distribution. To be specific, in comparison to KQDMN-s and KQDMN-t, KQDMN-mmd achieved a reduction of 0.919 and 0.920 in the MAE for prediction results on test set B1 and a reduction of 1.721 and 2.261 in MAE for prediction results on test set B2, respectively. The prediction results of KQDMN-coral on test set B1 showed similar improvements. These findings indicate that the domain adaptive method proposed in this paper offers significant advantages for battery capacity prediction in scenarios with limited data and varying data distributions.

Table 1 shows the comparison of the prediction performance of the different training strategies on the test datasets B1 and B2. As shown in the table, both KQDMN-mmd and KQDMN-coral achieve large performance improvements relative to KQDMN-s and KQDMN-t.

Table 2 presents a comparison of prediction performance across various references. As indicated in the table, KQDMN-coral exhibits significant improvements in performance in

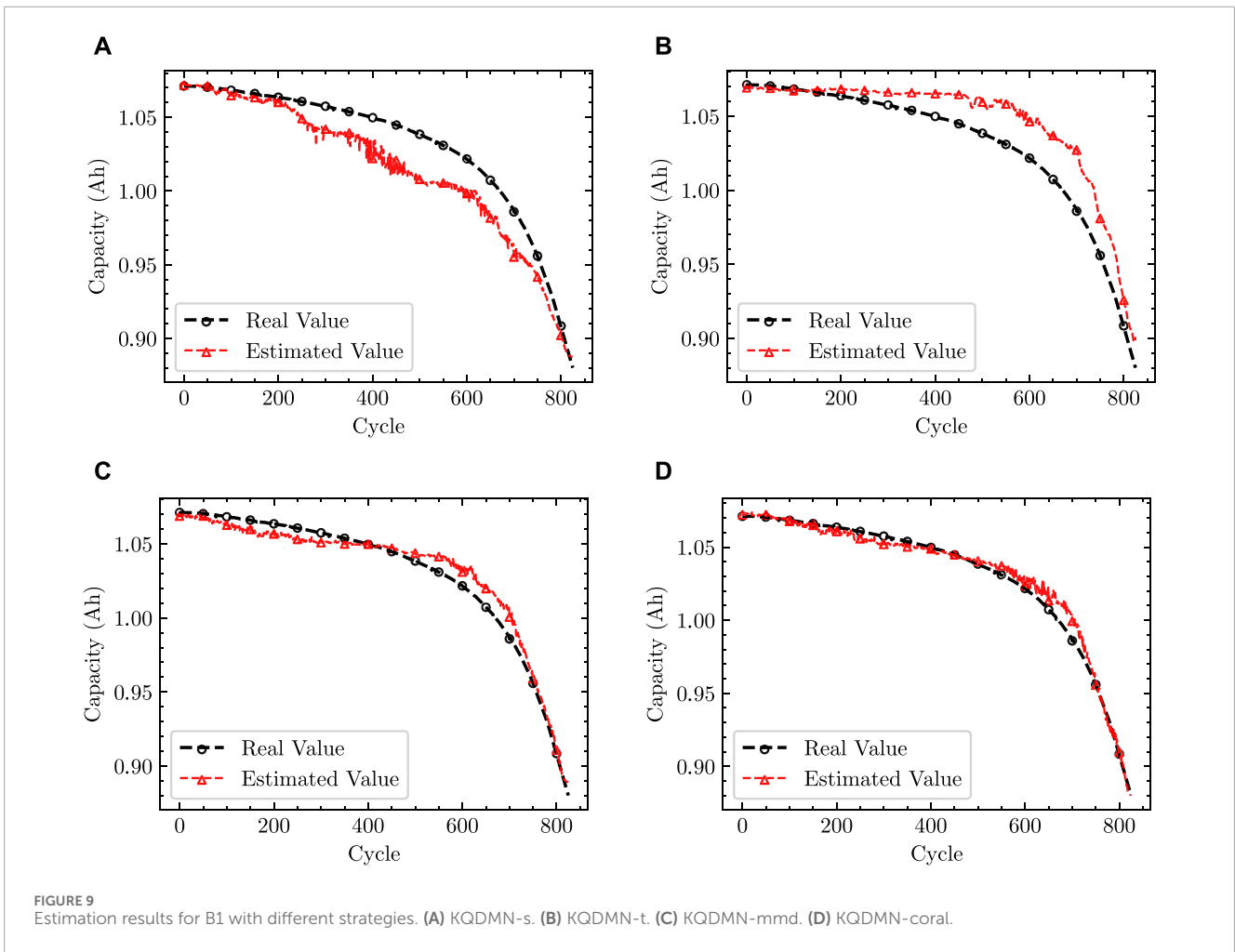


TABLE 2 Statistical results for capacity estimation across different references.

Strategies	MAE(%)	RMSE(%)
KQDMN-coral	0.19	0.23
Deng et al. (2024)	3.22	3.90
He and Wu (2023)	1.49	1.18
Han et al. (2022)	-	2.03
Wang et al. (2023a)	1.34	-

comparison to the four machine learning capacity estimation methods used in the references (Deng et al., 2024; He and Wu 2023; Han et al., 2022; Wang et al., 2023a). This demonstrates that the domain adaptive approach proposed in this paper can effectively improve the accuracy and generalization performance of battery capacity prediction. The excellent performance on the test set provides strong support for the method's feasibility in practical applications.

TABLE 3 Comparison results with private and shared encoders on test set B1.

Strategies	MAE(%)	RMSE(%)	R ²
shared encoder	1.45	1.15	0.895
private encoder	0.54	0.39	0.985

5 Features alignment and discussions

We investigate the ablation effects of the private domain-wise encoders on extracting domain-specific information, and compare two variants on test set B1: a) extracting information with a shared encoder, b) extracting information with two private encoders. The comparison results are given in Table 3. Compared with the variant a, the variant b brings 0.91 RMSE, 0.76 MAE and 0.09 R² improvement. This result demonstrates the essential of the private domain-wise encoders on extracting domain-specific information. Due to the larger domain gap, the private encoders can provide more capacity of information in each domain.

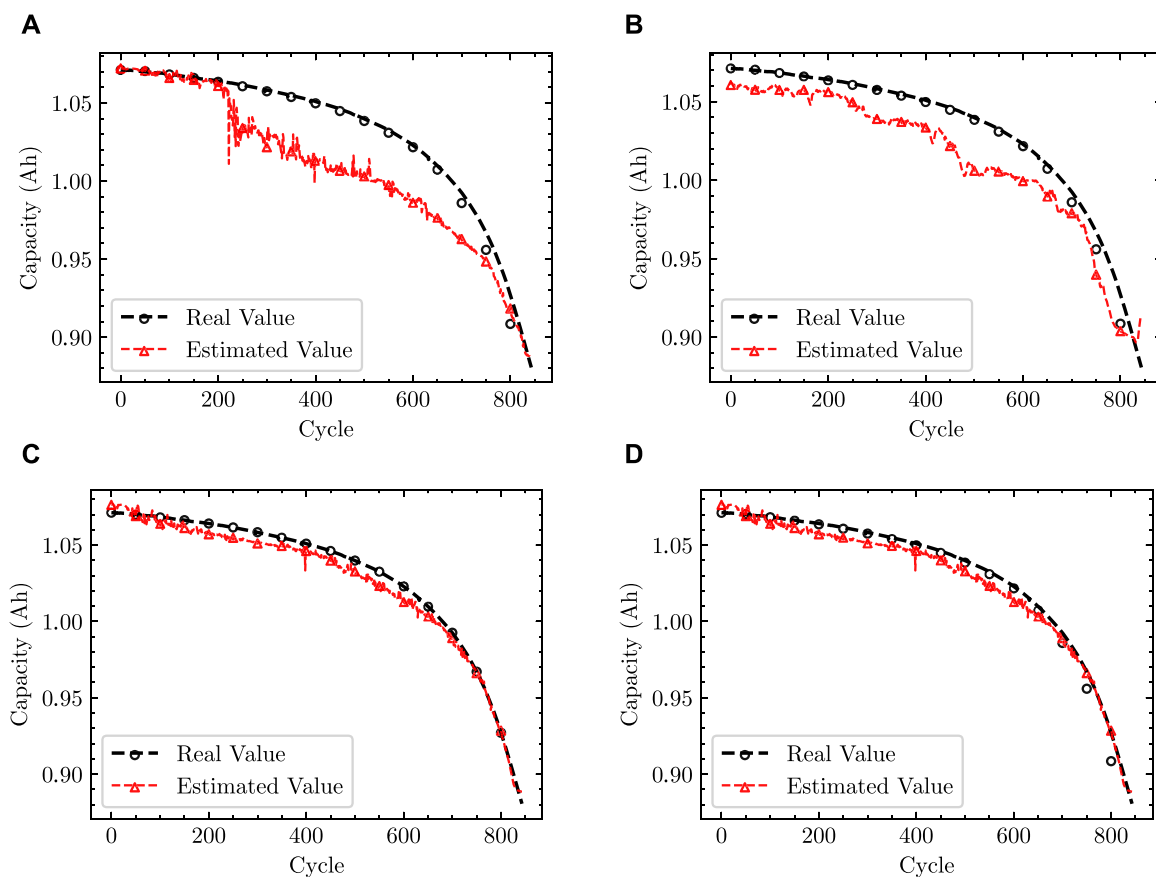


FIGURE 10 Estimation results for B2 with different strategies. (A) KQDMN-s. (B) KQDMN-t. (C) KQDMN-mmd. (D) KQDMN-coral.

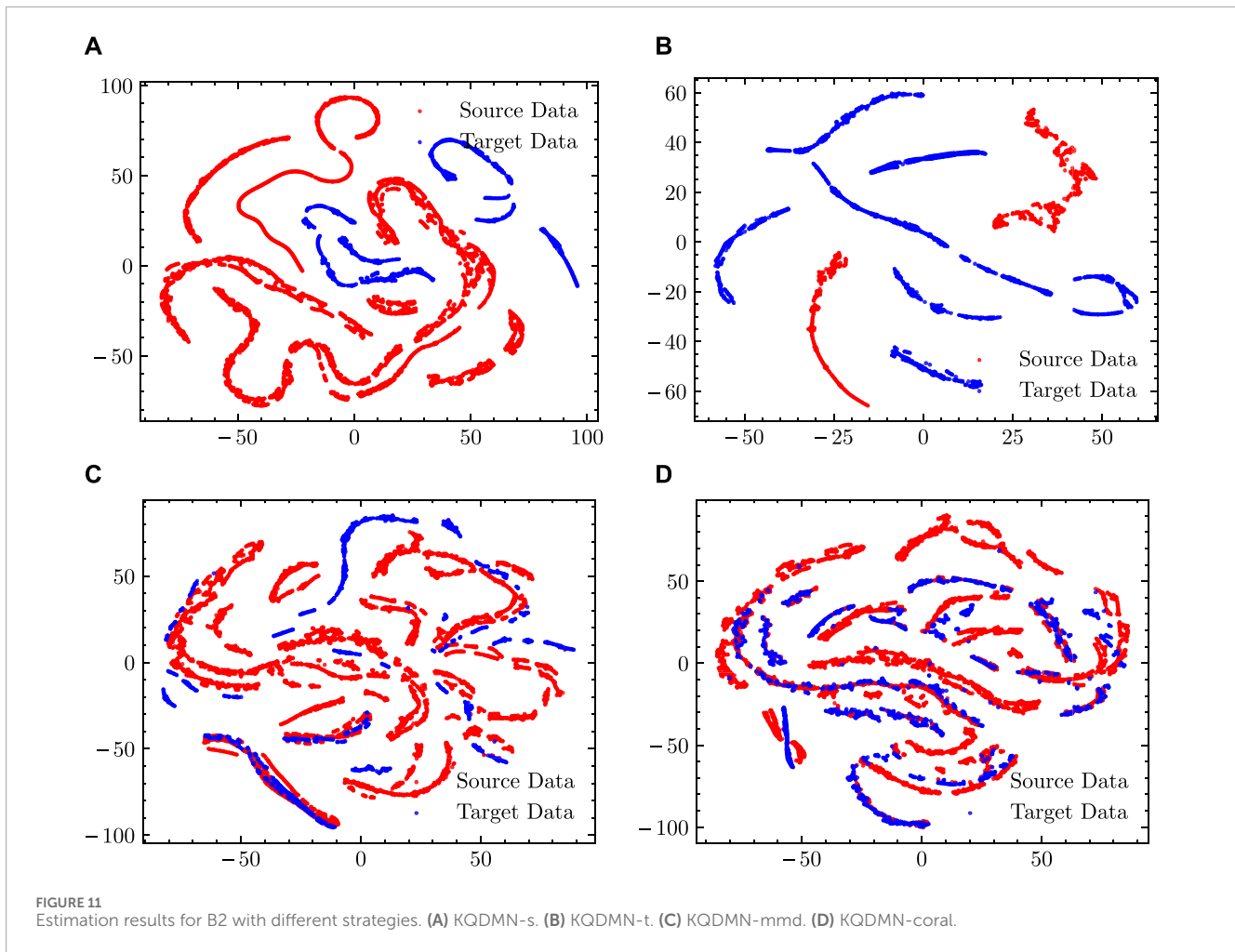
Features Alignment is important for cross-scenario capacity estimation via knowledge query domain mixing-up network. To investigate the impact of the proposed method on the source and target data features and the properties of aligned features, this paper compares the features extracted by a bidirectional cross-attentive Domain mixer under different training strategies using the t-Distributed Stochastic Neighbor Embedding (t-SNE) technique. t-SNE is a popular non-linear dimensionality reduction algorithm that can map high-dimensional data into a low-dimensional space for visualization and clustering (Maaten and Hinton, 2008). The basic idea of the algorithm is to construct a low-dimensional representation by measuring the similarity between high-dimensional data points and then attempting to minimize the KL scatter between that representation and the original data to preserve similarity. It is able to transform high-dimensional data points into points in two or three dimensions such that similar data points are close to each other in the low-dimensional space, while dissimilar data points are far from each other.

Figure 11 shows the visualization results of the features extracted by the bidirectional cross-attention domain blender in KQDMN-s, KQDMN-t, KQDMN-mmd, and KQDMN-coral, respectively, with the red and blue points representing data from the source and target domains, respectively. As can be seen from the figure, the source and target data feature spaces are more different due to

the different charging strategies used, and the source and target features learned by KQDMN-s and KQDMN-t are more dispersed and distributed in different regions. In contrast, due to the domain adaptation technique, the differences between the KQDMN-mmd and KQDMN-coral source data features and target data features are smaller, and the source data features and target data features are more closely distributed with features overlapping together. In addition, the distribution of features learned by KQDMN-coral was tighter compared to KQDMN-mmd. This is because MMD only reduces the first-order distance between domains, whereas CORAL loss can better align domains with second-order quantities. This is also reflected in Table 1, where the CORAL loss gives some performance gain due to the second-order distance measure.

6 Conclusion

In this paper, we have introduced the Knowledge Query Domain Mixing-up Network (KQDMN), a novel domain adaptation-based approach specifically designed for predicting lithium battery capacity. The approach addresses two key challenges in the field: the adaptability of predictive models to diverse domain features and the inherent issues of data scarcity and domain shifts that are commonly encountered in practical applications. The KQDMN is a



significant advancement in the realm of battery capacity estimation. By separating feature extraction from domain alignment, we have enhanced the model's ability to adapt to various domain features. This separation is crucial in dealing with the complexities and nuances of battery data, which often vary significantly across different application domains. One of the core components of the approach is the development of a set of knowledge queries inspired by the Transformer model. These queries are specifically designed to capture both temporal and global features of battery data. By integrating features obtained through convolutional operations, the KQDMN can effectively process and analyze data that encompasses a wide range of battery usage scenarios and conditions. To address the issue of domain bias, a common obstacle in battery capacity prediction, we have employed a cross-attention mechanism. This mechanism serves as a bridge between different domains, aiming to minimize the disparity between source and target domains. The bidirectional nature of this approach is crucial for facilitating the learning of domain-invariant feature representations, a key aspect in achieving accurate and reliable predictions.

The performance of the proposed model is demonstrated through its Mean Absolute Error (MAE) of 0.19%, Root Mean Square Error (RMSE) of 0.23%, and a Coefficient of Determination (R^2) of 0.997. These metrics highlight the precision and reliability of the KQDMN in estimating battery capacity. The experimental

results further reinforce the effectiveness of the method in a variety of practical scenarios, especially those characterized by limited data and significant variations in data distribution.

In conclusion, the Knowledge Query Domain Mixing-up Network represents a significant leap forward in the field of battery capacity prediction. Its innovative approach to feature extraction and domain adaptation, combined with the implementation of advanced machine learning techniques like convolutional operations and cross-attention mechanisms, makes it a powerful tool for accurately predicting lithium battery capacity. This method has shown high accuracy and robustness, particularly in challenging scenarios with limited data availability. We believe that the proposed approach can be a valuable asset in advancing the technology for lithium battery management and in contributing to more efficient and sustainable energy systems.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: <https://doi.org/10.1038/s41586-020-1994-5>.

Author contributions

ZL: Conceptualization, Investigation, Methodology, Project administration, Resources, Supervision, Visualization, Writing—original draft, Writing—review and editing. JC: Conceptualization, Data curation, Methodology, Writing—review and editing. TG: Formal Analysis, Validation, Writing—original draft. WZ: Resources, Validation, Writing—original draft. DC: Validation, Visualization, Writing—original draft. YG: Writing—review and editing, Formal Analysis.

Funding

The author(s) declare that no financial support was received for the research, authorship, and/or publication of this article.

References

- Alavi, S., Birkel, C., and Howey, D. (2015). Time-domain fitting of battery electrochemical impedance models. *J. Power Sources* 288, 345–352. doi:10.1016/j.jpowsour.2015.04.099
- Bartlett, A., Marcicki, J., Onori, S., Rizzoni, G., Yang, X. G., and Miller, T. (2015). Electrochemical model-based state of charge and capacity estimation for a composite electrode lithium-ion battery. *IEEE Trans. Control Syst. Technol.* 24, 384–399. doi:10.1109/TCST.2015.2446947
- Bi, J., Zhang, T., Yu, H., and Kang, Y. (2016). State-of-health estimation of lithium-ion battery packs in electric vehicles based on genetic resampling particle filter. *Appl. Energy* 182, 558–568. doi:10.1016/j.apenergy.2016.08.138
- Borgwardt, K. M., Gretton, A., Rasch, M. J., Kriegel, H.-P., Schölkopf, B., and Smola, A. J. (2006). Integrating structured biological data by kernel maximum mean discrepancy. *Bioinformatics* 22, e49–e57. doi:10.1093/bioinformatics/btl242
- Chemali, E., Kollmeyer, P. J., Preindl, M., Fahmy, Y., and Emadi, A. (2022). A convolutional neural network approach for estimation of li-ion battery state of health from charge profiles. *Energies* 15, 1185. doi:10.3390/en15031185
- Corno, M., Bhatt, N., Savaresi, S. M., and Verhaegen, M. (2014). Electrochemical model-based state of charge estimation for li-ion cells. *IEEE Trans. Control Syst. Technol.* 23, 117–127. doi:10.1109/tcst.2014.2314333
- Deng, Z., Xu, L., Liu, H., Hu, X., Wang, B., and Zhou, J. (2024). Rapid health estimation of in-service battery packs based on limited labels and domain adaptation. *J. Energy Chem.* 89, 345–354. doi:10.1016/j.jechem.2023.10.056
- Ganin, Y., Ustinova, E., Ajakan, H., Germain, P., Larochelle, H., Laviolette, F., et al. (2016). Domain-adversarial training of neural networks. *J. Mach. Learn. Res.* 17, 2096–2030.
- Gu, X., See, K., Li, P., Shan, K., Wang, Y., Zhao, L., et al. (2023). A novel state-of-health estimation for the lithium-ion battery using a convolutional neural network and transformer model. *Energy* 262, 125501. doi:10.1016/j.energy.2022.125501
- Guha, A., and Patra, A. (2018). Online estimation of the electrochemical impedance spectrum and remaining useful life of lithium-ion batteries. *IEEE Trans. Instrum. Meas.* 67, 1836–1849. doi:10.1109/tim.2018.2809138
- Han, T., Wang, Z., and Meng, H. (2022). End-to-end capacity estimation of lithium-ion batteries with an enhanced long short-term memory network considering domain adaptation. *J. Power Sources* 520, 230823. doi:10.1016/j.jpowsour.2021.230823
- He, J., and Wu, L. (2023). Cross-conditions capacity estimation of lithium-ion battery with constrained adversarial domain adaptation. *Energy* 277, 127559. doi:10.1016/j.energy.2023.127559
- Hu, X., Li, S., and Peng, H. (2012). A comparative study of equivalent circuit models for li-ion batteries. *J. Power Sources* 198, 359–367. doi:10.1016/j.jpowsour.2011.10.013
- Ji, S., Zhu, J., Lyu, Z., You, H., Zhou, Y., Gu, L., et al. (2023). Deep learning enhanced lithium-ion battery nonlinear fading prognosis. *J. Energy Chem.* 78, 565–573. doi:10.1016/j.jechem.2022.12.028
- Li, P., Zhang, Z., Grosu, R., Deng, Z., Hou, J., Rong, Y., et al. (2022). An end-to-end neural network framework for state-of-health estimation and remaining useful life prediction of electric vehicle lithium batteries. *Renew. Sustain. Energy Rev.* 156, 111843. doi:10.1016/j.rser.2021.111843
- Liu, H., Deng, Z., Yang, Y., Lu, C., Li, B., Liu, C., et al. (2023). Capacity evaluation and degradation analysis of lithium-ion battery packs for on-road electric vehicles. *J. Energy Storage* 65, 107270. doi:10.1016/j.est.2023.107270
- Lyu, C., Lai, Q., Ge, T., Yu, H., Wang, L., and Ma, N. (2017). A lead-acid battery's remaining useful life prediction by using electrochemical model in the particle filtering framework. *Energy* 120, 975–984. doi:10.1016/j.energy.2016.12.004
- Maaten, L. v. d., and Hinton, G. (2008). Visualizing data using t-sne. *J. Mach. Learn. Res.* 9, 2579–2605.
- Peng, X., Bai, Q., Xia, X., Huang, Z., Saenko, K., and Wang, B. (2019). “Moment matching for multi-source domain adaptation,” in Proceedings of the IEEE/CVF international conference on computer vision, 1406–1415.
- Pham, T., Truong, L., Bui, H., Tran, T., Garg, A., Gao, L., et al. (2022). Towards channel-wise bidirectional representation learning with fixed-point positional encoding for soh estimation of lithium-ion battery. *Electronics* 12, 98. doi:10.3390/electronics12010098
- Qu, J., Liu, F., Ma, Y., and Fan, J. (2019). A neural-network-based method for rul prediction and soh monitoring of lithium-ion battery. *IEEE Access* 7, 87178–87191. doi:10.1109/access.2019.2925468
- Samadi, M. F., Alavi, S. M., and Saif, M. (2012). “An electrochemical model-based particle filter approach for lithium-ion battery estimation,” in 2012 IEEE 51st IEEE conference on decision and control (CDC) (IEEE), 3074–3079.
- Silho, J., Roinila, T., and Stroe, D.-I. (2020). “Soh analysis of li-ion battery based on ecm parameters and broadband impedance measurements,” in IECON 2020 The 46th Annual Conference of the IEEE Industrial Electronics Society, 1923–1928. doi:10.1109/IECON43393.2020.9254859
- Sun, B., and Saenko, K. (2016). “Deep coral: correlation alignment for deep domain adaptation,” in Proceedings, Part III 14 Computer Vision—ECCV 2016 Workshops, Amsterdam, The Netherlands, October 8–10 and 15–16, 2016 (Springer), 443–450.
- Ungurean, L., Micea, M. V., and Cârstoiu, G. (2020). Online state of health prediction method for lithium-ion batteries, based on gated recurrent unit neural networks. *Int. J. Energy Res.* 44, 6767–6777. doi:10.1002/er.5413
- Wang, F., Zhao, Z., Zhai, Z., Guo, Y., Xi, H., Wang, S., et al. (2023a). Feature disentanglement and tendency retainment with domain adaptation for lithium-ion battery capacity estimation. *Reliab. Eng. Syst. Saf.* 230, 108897. doi:10.1016/j.res.2022.108897
- Wang, J., Deng, Z., Yu, T., Yoshida, A., Xu, L., Guan, G., et al. (2022). State of health estimation based on modified Gaussian process regression for lithium-ion batteries. *J. Energy Storage* 51, 104512. doi:10.1016/j.est.2022.104512
- Wang, S., Fan, Y., Jin, S., Takyi-Aninakwa, P., and Fernandez, C. (2023b). Improved anti-noise adaptive long short-term memory neural network modeling for the robust remaining useful life prediction of lithium-ion batteries. *Reliab. Eng. Syst. Saf.* 230, 108920. doi:10.1016/j.res.2022.108920
- Wang, S., Wu, F., Takyi-Aninakwa, P., Fernandez, C., Stroe, D.-I., and Huang, Q. (2023c). Improved singular filtering-Gaussian process regression-long short-term memory model for whole-life-cycle remaining capacity estimation of lithium-ion batteries adaptive to fast aging and multi-current variations. *Energy* 284, 128677. doi:10.1016/j.energy.2023.128677

Conflict of interest

The 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.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Wei, G., Lan, C., Zeng, W., Zhang, Z., and Chen, Z. (2021). Toalign: task-oriented alignment for unsupervised domain adaptation. *Adv. Neural Inf. Process. Syst.* 34, 13834–13846.

Wei, Y., and Wu, D. (2023). Prediction of state of health and remaining useful life of lithium-ion battery using graph convolutional network with dual attention mechanisms. *Reliab. Eng. Syst. Saf.* 230, 108947. doi:10.1016/j.res.2022.108947

Yanwen Dai, A. Y. (2022). Combined cnn-lstm and gru based health feature parameters for lithium-ion batteries soh estimation. *Energy Storage Sci. Technol.* 11, 1641. doi:10.19799/j.cnki.2095-4239.2021.0623

Yao, X.-Y., Chen, G., Pecht, M., and Chen, B. (2023). A novel graph-based framework for state of health prediction of lithium-ion battery. *J. Energy Storage* 58, 106437. doi:10.1016/j.est.2022.106437

Zhang, Q., Wang, D., Yang, B., Dong, H., Zhu, C., and Hao, Z. (2022). An electrochemical impedance model of lithium-ion battery for electric vehicle application. *J. Energy Storage* 50, 104182. doi:10.1016/j.est.2022.104182

Zhang, Y., Tang, H., Jia, K., and Tan, M. (2019). “Domain-symmetric networks for adversarial domain adaptation,” in Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, 5031–5040.

Zhao, H., Chen, Z., Shu, X., Shen, J., Lei, Z., and Zhang, Y. (2023). State of health estimation for lithium-ion batteries based on hybrid attention and deep learning. *Reliab. Eng. Syst. Saf.* 232, 109066. doi:10.1016/j.res.2022.109066

Zhu, J., Wang, Y., Huang, Y., Bhushan Gopaluni, R., Cao, Y., Heere, M., et al. (2022). Data-driven capacity estimation of commercial lithium-ion batteries from voltage relaxation. *Nat. Commun.* 13, 2261. doi:10.1038/s41467-022-29837-w