This article has been accepted for publication in IEEE Transactions on Transportation Electrification. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TTE.2024.3434553

Deep learning powered lifetime prediction for lithium-ion batteries based on small amounts of charging cycles

Yunpeng Liu, *Member, IEEE,* Moin Ahmed, Jiangtao Feng, Zhiyu Mao, *Member, IEEE,* and Zhongwei Chen

***Abstract*—The accurate lifetime prediction of lithium-ion batteries (LIBs) is essential to the normal and effective operation of electric devices. However, such estimation faces huge challenges due to the nonlinear capacity degradation process and uncertain LIBs' operating conditions. This paper proposes a novel end-to-end deep learning model, namely a dual-stream vision transformer with the efficient self-attention mechanism (DS-ViT-ESA), to predict the current cycle life (CCL) and remaining useful life (RUL) of the target battery. The local and global spatio-temporal features are effectively captured via the vision transformer with the efficient self-attention mechanism based on small amounts of charging cycles. Meanwhile, by serving the differences between each cycle as the supplementary model input, the inner-cycle and cycle-to-cycle aging information could be extracted and fused by a dual-stream structure to enhance prediction accuracy. Experiments exhibit that the proposed model only needs 15 charging cycles (about 1∼3% along the trajectory to failure) while ensuring the lifetime prediction accuracy (RUL error: 5.40%, CCL error: 4.64%, early lifetime prediction error: 2.16%). Meanwhile, the model also shows the effective zero-shot generalization capacity for the charging strategies not appearing in the training dataset.**

***Index Terms*—Battery lifetime prediction, Small amounts of charging cycles, Vision transformer, Dual-stream structure, Efficient self-attention mechanism.**

NOMENCLATURE

Parameters

$\mathbf{Q}, \mathbf{K}, \mathbf{V}$	Matrices that pack together sets of queries, keys, and values, respectively
C_{patch}	Aging cycles included in one patch
d_k, d_v	Dimensionalities of $\mathbf{K}$ and $\mathbf{V}$, respectively.
d_m	Dimension of embedding
E_{patch}	Point amounts of one curve included in one patch
h	Number of attention heads
n	Input size
n_c	Serial number of the current charging cycle
V, I, T, Q	Terminal voltage, current, temperature, and capacity
x_c	Picked cycle amount in the current stage
x_i	Picked cycle amount in the initial stage

Abbreviations

BMSs	Battery management systems
CCL	Current cycle life
CNN	Convolutional neural network
DS-ViT-ESA	Dual-stream ViT with the ESA mechanism
DS-ViT-SA	Dual-stream ViT with the SA mechanism
EOL	End of life
ESA	Efficient self-attention
LIBs	Lithium-ion batteries
LSTM	Long short-term memory
MAE	Mean absolute error
MHESA	Multi-head efficient self-attention
MHSA	Multi-head self-attention
RMSE	Root mean square error
RUL	Remaining useful life
SA	Self-attention
ViT	Vision Transformer
ViT-ESA	ViT with the ESA mechanism
ViT-SA	ViT with the SA mechanism

This work was supported in part by the Strategic Priority Research Program of the Chinese Academy of Sciences (Grant No. XDB0600400), the National Natural Science Foundation of China (Grant No. 52102112), and the Fundamental Research Funds for the Central Universities, China (Grant No. xzy012022072). (Corresponding authors: Jiangtao Feng, Zhiyu Mao, Zhongwei Chen)

Yunpeng Liu, Zhiyu Mao, and Zhongwei Chen are with the Power Battery & Systems Research Center, the State Key Laboratory of Catalysis, Dalian Institute of Chemical Physics, Chinese Academy of Sciences, Dalian 116023, China. (e-mail: liuyunpeng1994@dicp.ac.cn; zhymao@dicp.ac.cn; zwchen@dicp.ac.cn).

Moin Ahmed is with the Department of Chemical Engineering, University of Waterloo, 200 University Avenue West, Waterloo, ON N2L 3G1, Canada (e-mail: moin.ahmed@uwaterloo.ca).

Jiangtao Feng is with the Department of Environmental Science & Engineering, School of Energy and Power Engineering, Xi'an Jiaotong University, Xi'an, 710049, China (e-mail: fjtes@xjtu.edu.cn).

I. INTRODUCTION

LITHIUM-ION batteries (LIBs) are widely employed in numerous energy storage scenarios, such as aerospace, electronic product, and electric vehicles, owing to their strengths, including high energy density, long cycle life, and high output power [1, 2]. However, the concomitant internal side reactions of LIBs will inevitably lead to battery performance degradation, increased equipment maintenance costs, and destructive device malfunction. When the actual battery performance declines to a particular failure threshold of approximately 70%∼80% of the rated capacity, LIBs can be considered to reach their end of life (EOL) [3]. If the failed LIBs run continuously, uncontrollable safety problems would surge, leading to short circuits and even spontaneous combustion of electric vehicles. Therefore, battery management systems (BMSs) are widely employed in all kinds of electric devices to estimate accurately the LIBs' health status by some diagnosis indexes, such as state of health (SOH), current cycle life (CCL), and remaining useful life (RUL) [4]. Unfortunately,

Authorized licensed use limited to: Xian Jiaotong University. Downloaded on October 26,2024 at 02:33:52 UTC from IEEE Xplore. Restrictions apply.
© 2024 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Transactions on Transportation Electrification. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TTE.2024.3434553

accurate lifetime estimation faces huge challenges because LIBs' aging behavior is a non-linear electrochemical process with peculiarly complex intrinsic properties and varies considerably under uncertain operating conditions [5].

Generally, the available prediction methods for LIBs' lifetime can be divided into model-based methods and data-driven methods [6]. The model-based methods could deeply investigate the electrochemical mechanism and build the mathematical model of LIBs' aging process [7]. Such the modeling process is usually complex and requires prior knowledge about the battery aging behavior, which greatly restricts the practical applications [8]. Data-driven approaches have displayed enormous potential for estimating the battery's health status because these approaches could avoid complicated mathematical or physical modeling [9]. Generally, the data-driven approaches could be subdivided into health-indicator (HI)-based methods and end-to-end deep learning (DL) methods. For HI-based methods, the correlations between battery lifetime and extracted HI will directly affect the prediction performance [10]. Recently, some HIs have been widely employed to study the batteries' aging mechanisms, such as incremental capacity analysis (ICA), differential capacity analysis (DQA), differential voltage analysis (DVA), differential thermal voltammetry (DTV), etc. [11]. Wei et al. [12] employed the Gaussian process regression (GPR) method based on the ICA to accurately predict LIBs' RUL. Afshari et al. [13] used sparse Bayesian learning (SBL) to investigate a battery's early lifetime prediction based on the DQA and DVA curves. Li et al. [14] established the battery aging model with the multi-output GPR and particle filter (PF) algorithms, and the hidden features were extracted from the DTV curves. Although these approaches perform well, manual operation is unavoidable in the HI extraction process, which is time-consuming and needs vast domain knowledge.

End-to-end DL methods have drawn more and more attention because they can automatically grasp the hidden features and directly map the underlying relationship between the raw data and the battery lifetime. Various DL algorithms have recently been employed to study LIBs' RUL prediction performance. Ma et al. [15] proposed a DL framework in which the convolutional neural network (CNN) was combined with the long short-term memory (LSTM), and this framework was adapted for long-term battery capacity prediction. Wei et al. [16] proposed a model combining Monte Carlo dropout and gated recurrent unit (GRU) to describe the uncertainty of LIBs RUL estimation. Jiao et al. [17] proposed a PF framework based on a conditional variational autoencoder (CVAE) and a reweighting strategy to predict the batteries' RUL. Hong et al. [18] proposed an approach that used a dilated CNN to realize the swift RUL prediction with a few cycles of the target battery. Compared with the HI-based method, the above end-to-end deep models exhibit similar characteristics, namely, no feature engineering and less prior knowledge, enhancing onboard BMSs' practicability.

Three perspectives could be considered carefully to improve the prediction performance and universality of the end-to-end DL models: 1) A model input with more battery aging information. Past investigations usually served the collected inner-cycle signals, i.e., terminal voltage, current, temperature, or capacity ($V/I/T/Q$) of each cycle, as the model input [19, 20], and even only utilized the SOH sequences to predict subsequent degradation trajectory [21, 22]. However, the terminal parameter differences along cycles contain abundant information about the batteries' health status. The cycle-to-cycle aging information is accessible but less utilized for battery lifetime prediction; 2) The enhanced generalization ability for the battery aging condition does not appear in the training dataset. The models in the past literature are often applied to the target batteries under a similar charge-discharge strategy with the training dataset. Meanwhile, the fine-tuning process also has to be conducted to ensure RUL prediction accuracy, meaning that an amount of historic aging data of tested batteries is necessary [23–25]; 3) The innovation of the DL model structures. The DL methods for the battery lifetime prediction primarily focus on the CNN model, LSTM model, and their variants, but their inherent drawbacks are insurmountable. For example, the inductive bias of the CNN model (locality and spatial invariance) would negatively influence the extraction of global features [26], which might lower the information interaction between the battery aging cycles. The LSTM model could not capture long-range context dependencies, especially when the battery raw data is extremely long [27]. Besides, the low parallelism and the overfitting problem for traditional temporal networks also restrict the performance enhancement of RUL prediction. Hence, a novel DL model structure is urgently needed to be designed.

To handle the above issues, a novel DL network (noted as DS-ViT-ESA) for battery lifetime prediction is proposed in this paper. Firstly, to utilize inner-cycle and cycle-to-cycle aging information more efficiently, a dual-stream parallel framework is constructed to extract and fuse the multi-timescale hidden features. Secondly, to decrease the negative influences of the common models' inductive bias on the LIBs' lifetime prediction, a vision transformer (ViT) is introduced because its self-attention (SA) mechanism ensures flexibility and performance in feature acquisition. This structure has been utilized in other classical models (CLIP [28], BLIP [29], and DeiT [30]) as an image encoder. Thirdly, based on the fact that the matrix multiplication is associative, the efficient self-attention (ESA) mechanism is introduced into the ViT structure, which exhibits linear memory and computational complexity relative to the input size. Meanwhile, this mechanism has an equivalent representational capacity as the widely used dot-product attention mechanism. Besides, more attention modules are integrated into a neural network, which would bring superior estimation performance [31, 32].

Unlike the previous LIBs' lifetime prediction model, the battery data matrix is split into a series of patches, similar to the multi-channel images' splitting process. Then, a stack of transformer encoders could extract the high-level spatio-temporal features in inner-patch and patch-to-patch information more comprehensively [33]. Therefore, the prediction accuracy about battery lifetime could be enhanced significantly. Three main contributions of this work are listed as follows:

1) For many previous battery lifetime studies, a mass of

Authorized licensed use limited to: Xian Jiaotong University. Downloaded on October 26,2024 at 02:33:52 UTC from IEEE Xplore. Restrictions apply.
© 2024 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Transactions on Transportation Electrification. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TTE.2024.3434553

the target battery's aging cycles must be collected. By contrast, this work just needs fewer aging cycles, which consist of a few initial cycles and several current cycles during the charging process. This setting could guarantee the randomness of the working load and decrease the operation cost for the onboard aging test.

2) A DS-ViT-ESA network, which integrates a dual-stream framework, ViT, and ESA mechanism, is proposed to map the relationship between original charging data and battery lifetime under various charging strategies. The dual-stream framework could capture and fuse the multi-timescale hidden feature from inner-cycle and cycle-to-cycle aging information. ViT structure could extract the local and global spatio-temporal degradation features. Introducing the ESA mechanism in the ViT structure could optimize computation complexity and enhance estimation performance.
3) The proposed approach achieves an outstanding battery lifetime prediction, which has been testified in an open dataset with many batteries. Then, this approach also exhibits the effective zero-shot generalization ability when predicting the target battery with the charging strategy not appearing in the training dataset. Besides, early lifetime prediction of one fresh battery could also be achieved when the initial cycles serve as the model input, expressing the practical application value.

The rest of this article is organized as follows. Section II introduces the battery dataset. In Section III, the proposed DS-ViT-ESA model is outlined. Section IV details the experiment results and discussion. Conclusions are drawn in Section V.

II. Data Analysis

A. Battery data description and data cleaning

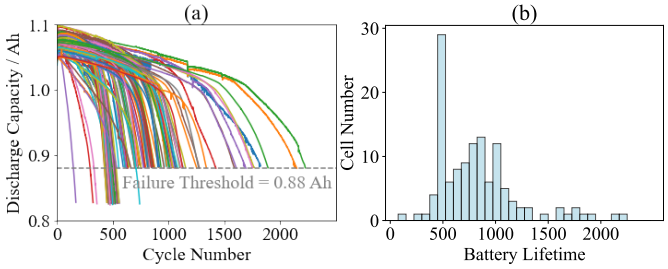

Fig. 1. Degradation trajectories (a) and lifetime distribution statistical histogram (b) of all batteries in the dataset.

Considering the sample amount of battery dataset and the various charging strategies in practice, the experimental LIBs datasets involving 72 different fast-charging strategies, coming from Massachusetts Institute of Technology and Stanford University were used in this study, which contains 124 high-power $LiFePO_4$ (LFP)/graphite batteries with a rated capacity of 1.1 Ah, and its upper-cutoff voltage (UCV) and lower-cutoff voltage (LCV) are 3.6 V and 2.0 V, respectively [34]. All the aging processes were conducted at 30 °C in a thermal chamber until the capacity reached 80% of the rated capacity. A series of fast charging strategies were operated, which is denoted as "C1(Q1%)-C2(80%)", where C1 and C2 represent the first and second fast charging rates, respectively, and Q1 represents the state of charge (SOC) at which C1 switches to C2. The multi-stage constant current continued until the SOC reached 80%, and then the batteries were charged to the UCV using a 1C constant current-constant voltage (CC-CV) strategy. Meanwhile, all the batteries were discharged to the LCV at the same discharge rate (4C). Fig. 1a shows the aging trajectories of 124 batteries, and all the capacity fades are negligible in the first 100 cycles and accelerate near the EOL. Fig. 1b displays the corresponding lifetime distribution statistical histogram. All the batteries' lifetimes range from 150 to 2200 cycles, and they are principally distributed between 500 and 1000 cycles.

The data cleaning of battery raw data was operated in advance to optimize some existing data fluctuations in this work, such as outliers and unexpected data slump or boom. Firstly, the outliers of the terminal parameters should be picked up and replaced with the average value of closely related points. Secondly, the terminal data curves should be smoothed to denoise via the savitzky-golay filter. Thirdly, to decrease the computational complexity and meet the requirement of the model's input, the terminal data curves should also be downsampled via the linear interpolation method. In this study, the point numbers of all terminal data curves are identical (160 points of one curve).

B. Battery charging data analysis

The battery's internal properties would determine its lifetime, which could be expressed by its electrical performance and thermal properties. Hence, there are strong correlations between battery lifetime and the monitored data. In this work, the terminal parameter data during the charging process is employed to describe the battery lifetime, which could be demonstrated as follows: First, it is relatively easy and precise to gain the terminal data during the practical charging process, which is peaceful and stable. In contrast, the terminal data during the discharging process could not be suitable and even unreliable for forecasting the battery lifetime due to the random workload. Therefore, the terminal charging data is more reasonable for predicting the battery lifetime in practice. Second, the batteries' external electrical and thermal properties with various lifetimes would be distinct, which could be reflected by the terminal curve during the charging process. Fig. 2 displays the $V/I/T$ curves changing with Q of the No.18 battery (the No.18 battery is an eighteenth battery in the "2017-05-12_batchdata_updataed_struct_errorcorrect.mat" data file provided by Ref. [34]) during the charging process. The battery lifetime of the No.18 battery is 782, under the "5.4C(70%)-3C(80%)" charging strategy. Both in V vs Q (V/Q) curves (Fig. 2a) and I vs Q (I/Q) curves (Fig. 2b), there are clear distinctions for different aging cycles, including the duration of the curve platforms and the position of the turning points, because the LIBs' internal properties would change while the aging process. Meanwhile, the value, rising, and falling rate of the T vs Q (T/Q) curves (Fig. 2c) also vary for different aging cycles. Although the chamber temperature

Authorized licensed use limited to: Xian Jiaotong University. Downloaded on October 26,2024 at 02:33:52 UTC from IEEE Xplore. Restrictions apply.
© 2024 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Transactions on Transportation Electrification. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TTE.2024.3434553

is fixed, the battery's surface temperature gradually rises along with the aging process due to the heat accumulation from the previous discharging stage. When the current maximum charging capacity is near 80% of the rated capacity, the surface temperature decreases clearly due to the sharply reduced discharging time. Therefore, the $V/I/T/Q$ data during the charging process serve as the model input in this study.

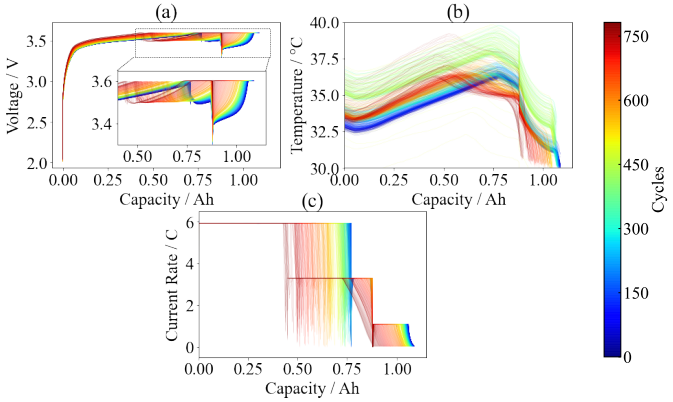

Fig. 2. Electrical performances and thermal properties of the No. 18 battery during the charging process: (a) V/Q curves; (b) I/Q curves; (c) T/Q curves.

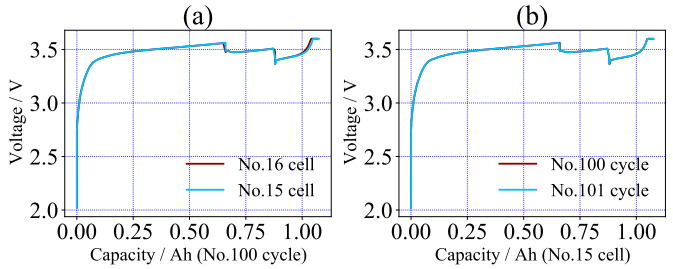

Fig. 3. (a) The V/Q curve comparisons of 100th cycle for No.15 battery and No.16 battery; (b) The V/Q curve comparisons of 100th cycle and 101th cycle for No.15 battery.

Interestingly, since certain tolerable deviations exist for industrial battery production, the batteries' lifetime might be distinct even though their charging-discharging strategies are identical. For example, the EOL values of the No.15 battery and No.16 battery (the No.15 battery and No.16 battery are the fifteenth battery and the sixteenth battery, respectively, in the "2017-05-12_batchdata_updataed_struct_errorcorrect.mat" data file provided by Ref. [34]), are 877 and 714 cycles, respectively. However, they are under the identical charging-discharging strategy (5.4C (60%)-3C (80%)). As shown in Fig. 3a, when the CCL value is 100 cycles, the two batteries' V/Q curves are almost identical, but the RUL values of the No.15 battery and No.16 battery are 777 cycles and 614 cycles, respectively. Besides, as displayed in Fig. 3b, the V/Q curves of the No.15 battery's neighboring aging cycles are also indistinguishable, which would cause unreliability and instability for the development of the DL models.

In this work, the cycle-to-cycle aging information is gained by calculating all the terminal parameter differences between each cycle and 1st aging cycle, and this supplement input has

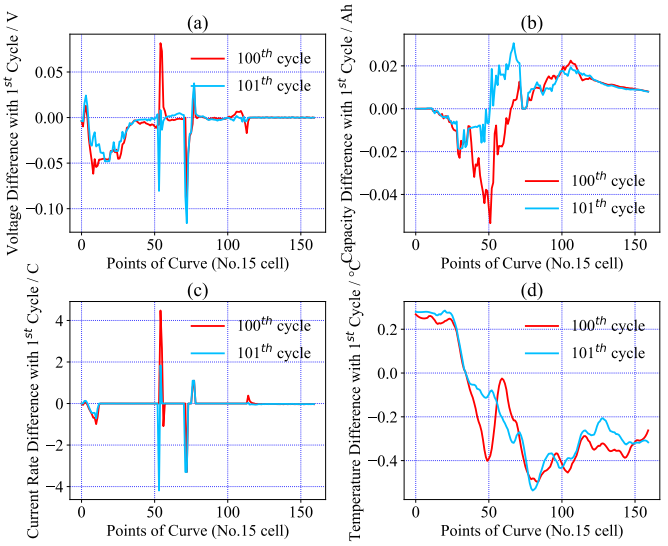

Fig. 4. All the parameter difference curves of the 100th cycle with the 1st cycle, and 101th cycle with the 1st cycle of the No.15 battery (a: V, b: Q, c: I, and d: T).

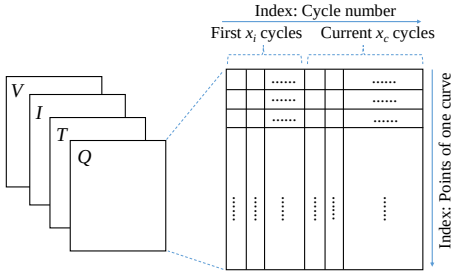

Fig. 5. Model input structure of V, I, T, and Q data.

been proven to work for LIBs' lifetime prediction by some literature [35, 36]. Fig. 4 displays the parameter difference curves (a: V, b: I, c: T, and d: Q) of the specified cycles with the 1st cycle for the No.15 battery. Compared with the indistinguishable V/Q curves (Fig. 3b), there are distinct differences, including the orientation, position, and value of the peaks for all the parameter differences curves. Therefore, all the terminal parameter differences would also serve as the supplement model input to achieve the end-to-end prediction and avoid the demand of the prior knowledge.

III. Proposed Approach

A. Input data generation

Based on the representations of Section II-B, the first x_i charging cycles reflecting the initial status and the current x_c charging cycles reflecting the current aging status are picked up for battery lifetime prediction. This data matrix consisting of initial and current cycles serves as the model input, in which the degradation characteristics contained in these aging cycles could be captured more preciously and robustly. The RUL and CCL of a running battery could be gained when the model input consists of the first x_i cycles and current x_c cycles, and the early lifetime prediction of one fresh battery also could be

Authorized licensed use limited to: Xian Jiaotong University. Downloaded on October 26,2024 at 02:33:52 UTC from IEEE Xplore. Restrictions apply.
© 2024 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Transactions on Transportation Electrification. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TTE.2024.3434553

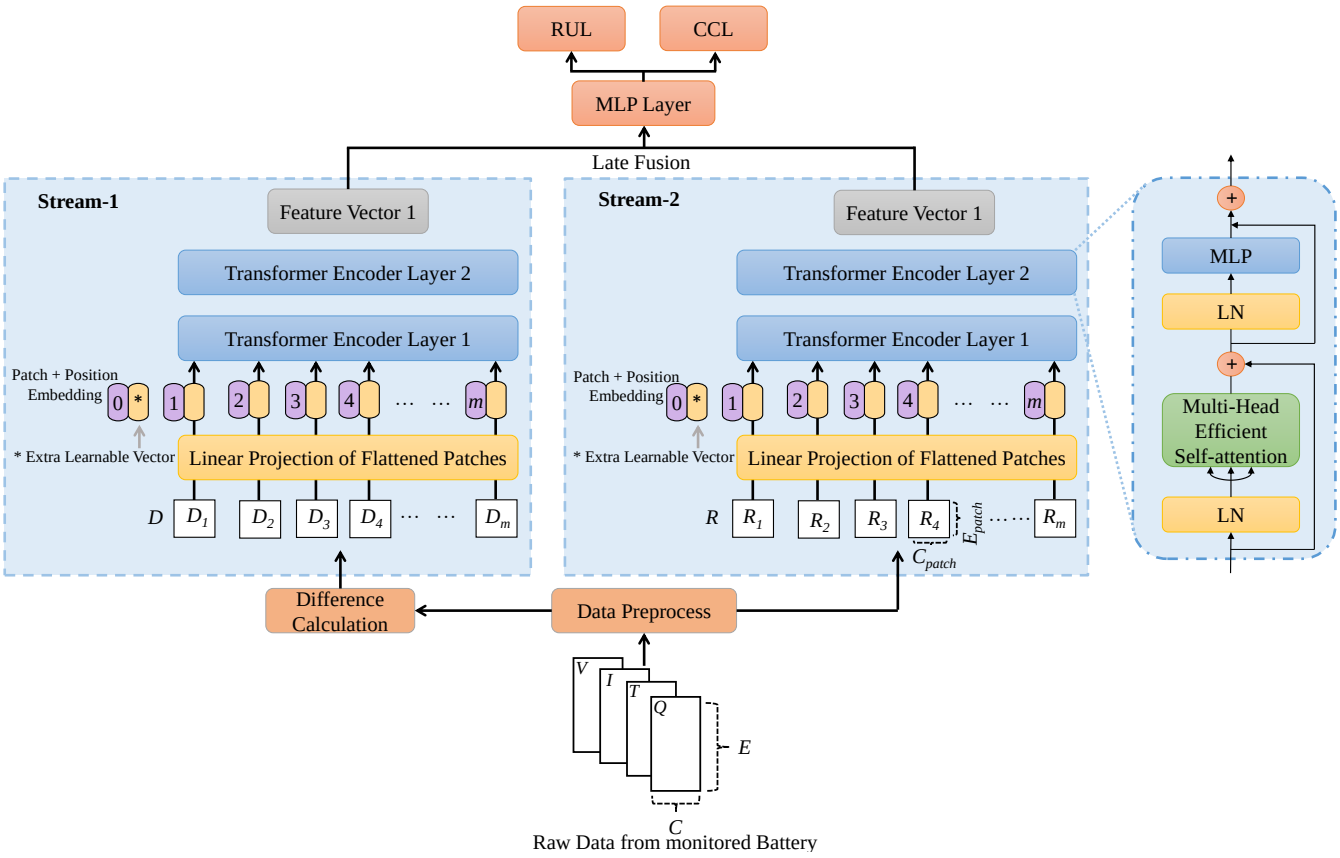

Fig. 6. Detailed illustration of DS-ViT-ESA algorithm.

achieved when the model input consists of the first $x_i + x_c$ cycles. Fig. 5 displays the specific structure of the input matrix, which is similar to the multi-channel image data, and this input matrix has four channels (c), i.e., V, I, T, and Q curve data during the charging process. Considering the computational complexity, this matrix's height (h) is set to 160, and the width (w) of this matrix is set to 15, in which x_i and x_c are set to 5 and 10, respectively. In detail, the initial aging data originate from the 1st to 5th charging cycles, and the current aging data originate from (n_c-9)th to n_cth cycle (n_c is the serial number of the current charging cycle; $15 \leq n_c \leq$ EOL of this battery). In other words, the measured CCL value is the n_c, and the measured RUL value is the difference between the EOL and n_c. Finally, the three-dimensional matrix ($c \times w \times h$: 4×15×160) is created as the model input. Besides, to speed up the model convergence rate, the input data, before being fed into the model, also needs to be normalized via z-score normalization.

B. DS-ViT-ESA structure

In 2020, the Google Brain team put forward the vision transformer, i.e., ViT, which employed a pure transformer directly to the sequence of image patches to classify the full image, and this has performed state-of-the-art accuracy on multiple image recognition benchmarks [37]. Because the battery data matrix is highly similar to the multi-channel image, the performance of the ViT variant applied for battery lifetime prediction is worth studying. In this work, the ViT with a dual-stream structure and the efficient self-attention mechanism (DS-ViT-ESA, as shown in Fig. 6) is proposed and applied to estimate the LIBs' RUL, CCL, and early lifetime effectively.

1) Embedding of battery data: After the data preprocessing, the raw data matrix (R) could be inputted into Stream-1 directly to extract the inner-cycle aging feature, and the parameter difference matrix (D) between each cycle and 1st aging cycle serves as the input of Stream-2, which would capture the cycle-to-cycle aging features. Before inputting in a stack of transformer encoder layers, the data matrix needs to be embedded into a sequence divided into patch embedding and position embedding. The two data matrices are split into a series of patches with the same size. Based on the description in Section III-A, the C and E are the w and h of one data matrix, respectively. The C_{patch} and E_{patch} are the w and h of one patch, respectively, representing the selected battery aging cycles and the portion of curves. So the data matrix is divided into $m = (C/C_{patch}) \times (E/E_{patch})$ patches and the m is the length of the input sequence of the subsequent transformer encoder.

Because the backbone of both the two streams is identical, the detailed structure of Stream-2 will be described. Every patch is flattened and then passes through a linear projection operation, changing its dimension to d_m. In a ViT structure, a classification (CLS, marked by "$*$") token plays a vital role because its introduction enables the model to perform

Authorized licensed use limited to: Xian Jiaotong University. Downloaded on October 26,2024 at 02:33:52 UTC from IEEE Xplore. Restrictions apply.
© 2024 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Transactions on Transportation Electrification. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TTE.2024.3434553

global semantic modeling of one image [37]. Therefore, an extra learnable vector with dimension d_m is joined with the patch embedding obtained before. This vector contains the learned aging features and would ultimately be fed into the output layer. Besides, the transformer-based model itself lacks an inherent mechanism to grasp the relative positional relationships between elements in the input sequence. However, the relative position of the terminal parameters is vital for precise battery lifetime prediction. Next, the learnable one-dimensional position embedding with dimension d_m is added with patch embedding to gain the final input sequence of a stack of transformer encoder layers.

2) Transformer encoder with the efficient self-attention mechanism: As displayed in the right of the Fig. 6, one transformer encoder with the ESA mechanism mainly consists of a multi-head efficient self-attention (MHESA) block, a two-layer multilayer perceptron (MLP), and a layer normalization (LN). The MHESA mechanism and structure will be discussed first.

It is well-known that the kernel of the classical transformer encoder is the multi-head self-attention (MHSA) mechanism, which would efficiently use parallel computing, thereby endowing the model with expressive power and learning capabilities [38, 39]. Through the collaborative manner of multiple attention heads, it could generate various attention representations, facilitating the simultaneous capture of features and relationships at different levels in the input sequence. The self-attention (SA) mechanism is essentially a scaled dot-product operation for each query with all the keys as follows:

$$D(\mathbf{Q},\mathbf{K},\mathbf{V}) = \rho\left(\frac{\mathbf{Q}\mathbf{K}^T}{\sqrt{d_k}}\right)\mathbf{V} \tag{1}$$

where $\mathbf{Q} \in \mathbb{R}^{n\times d_q}$, $\mathbf{K} \in \mathbb{R}^{n\times d_k}$, and $\mathbf{V} \in \mathbb{R}^{n\times d_v}$ represent matrices that pack together sets of queries, keys, and values, respectively. n represents the input size. d_q, d_k, and d_v are the dimensions of $\mathbf{Q}$, $\mathbf{K}$, and $\mathbf{V}$, respectively, in which $\sqrt{d_k}$ is employed to adjust the product to avoid extremely small gradients [40]. ρ donates the scaling normalization. For MHSA mechanism, $\rho(\mathbf{Y}) = \sigma_{row}(\mathbf{Y})$, in which σ_{row} donates applying the softmax function along each row of matrix $\mathbf{Y}$. The formula for the MHSA mechanism is built on the above SA weight matrix, as shown below:

$$Multihead(\mathbf{Q},\mathbf{K},\mathbf{V}) = Concat(head_1,...,head_h)\mathbf{W}^O \tag{2}$$

where $\mathbf{W}^O \in \mathbb{R}^{hd_v\times d_m}$ represents a projection matrix for the multi-head output, in which h represents the number of attention heads, and each $head_i$, $i = 1,..,h$, processes a single SA function characterized by its own learned projection matrices as:

$$head_i = D\left(\mathbf{Q}\mathbf{W}_i^Q, \mathbf{K}\mathbf{W}_i^K, \mathbf{V}\mathbf{W}_i^V\right) \tag{3}$$

where $\mathbf{W}_i^Q \in \mathbb{R}^{d_m\times d_q}$, $\mathbf{W}_i^K \in \mathbb{R}^{d_m\times d_k}$, and $\mathbf{W}_i^V \in \mathbb{R}^{d_m\times d_v}$ represent projection matrices which are utilized to create different subspace representations of the $\mathbf{Q}$, $\mathbf{K}$, and $\mathbf{V}$ in the i^{th} projection mode, respectively.

Fig. 7 displays the detailed architecture comparison between the MHSA and MHESA, in which the blue dotted boxes donate their respective attention structures. In the two blue dotted boxes, each small box represents the input, output, or intermediate matrix. Around the box is the name of the associated matrix, and inside it are both the variable name and the matrix's dimensions. $\mathbf{X}$, $\mathbf{D}$, and $\mathbf{E}$ denote the input and the output of the scaled dot-product attention and efficient attention, respectively. $\mathbf{S}$ and $\mathbf{G}$ donate the pairwise similarities and global context vectors, respectively. ρ, ρ_q, and ρ_k are the scaling normalization on $\mathbf{S}$, $\mathbf{Q}$, and $\mathbf{K}$, respectively. n, d, d_k, and d_v are input size and the dimensionalities of $\mathbf{X}$, $\mathbf{K}$, and $\mathbf{V}$, respectively. $\otimes$ and $\oslash$ denote the matrix multiplication and the scaled operation, respectively. As shown in Fig. 7a, for the classical SA mechanism, the pairwise similarities $\mathbf{S}$ between each pair of tokens should be calculated, leading to $O(n^2)$ memory complexity and $O(d_k n^2)$ computational complexity. Such networks' memory and computational cost grows quadratically with n [41]. It would cause an enormous resource burden for the embedded processors when the large-scale battery data is continuously inputted into BMSs [42]. To address this issue, attention mechanism variants, i.e., the ESA mechanism, which is inspired by the fact that matrix multiplication is associative, have been introduced in this work. Its detailed structure is shown in Fig. 7b. Compared with the dot-product attention, efficient attention avoids the attention map $\mathbf{S}$ about all among tokens, which is replaced by the generation of global context vector $\mathbf{G}$ with a global, semantic description of the input feature $\mathbf{X}$. The following equation defines the efficient attention mechanism:

$$E(\mathbf{Q},\mathbf{K},\mathbf{V}) = \rho_q(\mathbf{Q})\left(\rho_k(\mathbf{K})^T\mathbf{V}\right) \tag{4}$$

where ρ_q, and ρ_k are the scaling normalization on $\mathbf{Q}$, and $\mathbf{K}$, respectively. For this work, $\rho_q(\mathbf{Y}) = \sigma_{row}(\mathbf{Y})$, and $\rho_k(\mathbf{Y}) = \sigma_{col}(\mathbf{Y})$, in which σ_{row}, σ_{col} denote applying the softmax function along each row or column of matrix $\mathbf{Y}$, respectively. It can be seen that the main improvement of the ESA mechanism is the switching of the order from $(\mathbf{Q}\mathbf{K}^T)\mathbf{V}$ to $\mathbf{Q}(\mathbf{K}^T\mathbf{V})$. Hence, the memory complexity and computational complexity are decreased to $O(d_k d_v)$ and $O(d_k d_v n)$, respectively, and $d_k d_v$ is far less than n^2 in practical application, which enhances the calculation efficiency. Importantly, each row of $\sigma_{row}(\mathbf{Q}\mathbf{K}^T)$ sums up to 1, which represents a normalized attention distribution over all positions, and the matrix $\sigma_{row}(\mathbf{Q})\sigma_{col}(\mathbf{K})^T$ possesses a similar property [31], so it is almost equivalent between dot-product and efficiency attention. Besides, enhanced estimation performance is achieved by integrating additional attention modules into the neural network. The ESA mechanism could be introduced into the ViT with the dual-stream structure to decrease the computational complexity and enhance the battery lifetime prediction accuracy.

Finally, each transformer encoder layer contains a two-layer MLP with GELU non-linear activation function and a dropout layer, which could enhance the model's non-linearity ability and prevent overfitting. To address the issues of gradient vanishing while promoting information flow within the

Authorized licensed use limited to: Xian Jiaotong University. Downloaded on October 26,2024 at 02:33:52 UTC from IEEE Xplore. Restrictions apply.
© 2024 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Transactions on Transportation Electrification. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TTE.2024.3434553

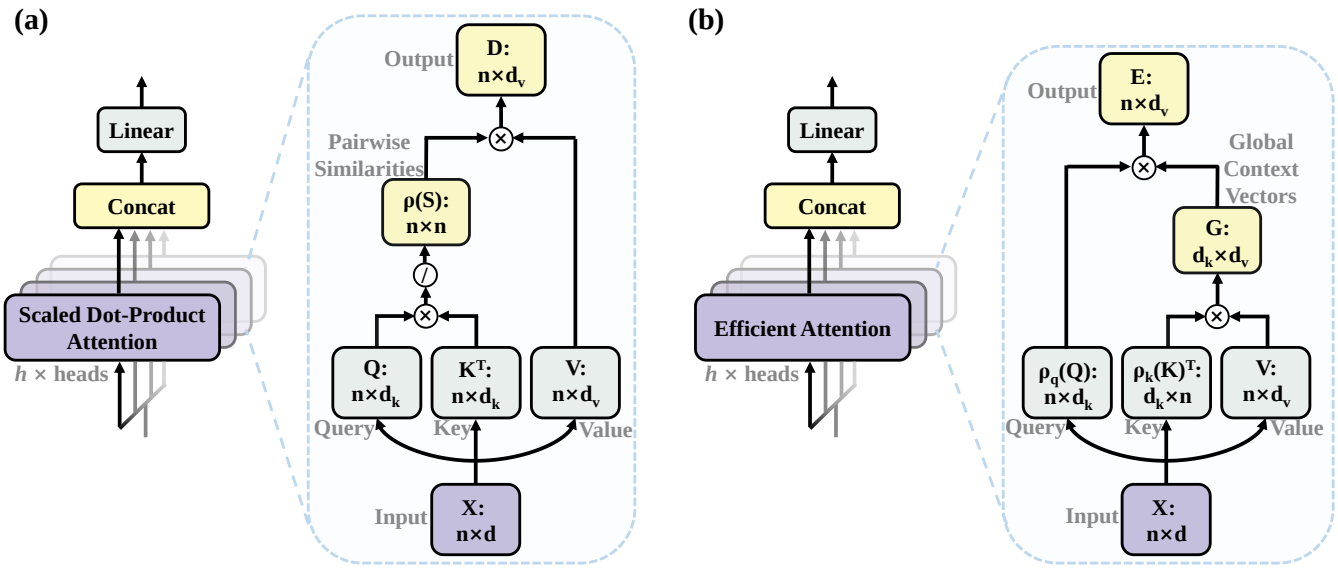

Fig. 7. Architecture of the MHSA and its scaled dot-product attention (blue dotted box) (a); Architecture of the MHESA and its efficient attention (blue dotted box) (b).

network, the residual connections are used before MHESA and MLP. Meanwhile, introducing the LN layer in the transformer encoder layer normalizes each feature dimension, enhancing the robustness and convergence of the model.

3) Late fusion of feature vectors from dual stream: As shown in the top of the Fig. 6, the two feature vectors could be gained after a stack of transformer encoders in the dual streams containing the inner-cycle and cycle-to-cycle battery aging information. The two-layer MLP with ReLU non-linear activation function is employed as late fusion in a dual-stream network and outputs the RUL and CCL to merge the two features. This study selects the joint mean square error (MSE) loss as the loss function because this model has two outputs. The model is trained with an Adam optimizer with an initial learning rate of 0.001, which would be gradually reduced based on the network's performance on the training dataset to avoid the overfitting problem.

IV. Results and Discussions

To ensure the prediction accuracy of the DS-ViT-ESA model, this work first uses grid search to gain the best hyperparameter combination of the DS-ViT-ESA model. Then, the DS-ViT-ESA model's sensitivity to the input data type is tested. Meanwhile, the ablation experiments of the DS-ViT-ESA model are also carried out. Finally, the prediction accuracy of the DS-ViT-ESA model is compared with other baseline methods. This study implements the models' simulations in Python 3.7 with the PyTorch packages on PyCharm. All run on a PC with an AMD EPYC 7402 48 cores (2.8 GHz) and an NVIDIA RTX 3080 GPU with a 24 GB GDDR6X.

The dataset described in Section II is randomly divided into two datasets: the training and test datasets. In the training dataset, 20% of the data are randomly selected as the validation dataset to search the model's hyperparameters and evaluate the model's estimation ability during the training process. Then, to display the prediction performance of the trained model, this work uses the 30 raw battery data as a test dataset, including 25 different charging strategies. Among these, 12 charging strategies do not appear in the training dataset.

A. Evaluation metrics

For algorithm evaluation, the three metrics are introduced as follows: Mean absolute error (MAE)

$$\mathrm{MAE} = \frac{1}{N}\sum_{i=1}^{N}|y_i - \hat{y}_i| \tag{5}$$

Mean absolute percentage error (MAPE)

$$\mathrm{MAPE} = \frac{1}{N}\sum_{i=1}^{N}\left|\frac{y_i - \hat{y}_i}{y_i}\right| \times 100\% \tag{6}$$

Root mean square error (RMSE)

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2} \tag{7}$$

where N is the number of data participating in the error calculation. y_i is the real value and $\hat{y}_i$ is the corresponding predicted value.

B. The configuration of DS-ViT-ESA

In this study, the crucial hyperparameters of the models are tuned using grid search, which includes Batch size $\in \{128, 256, \mathbf{512}, 1024\}$, $C_{patch} \in \{1, \mathbf{3}, 5, 15\}$, and $E_{patch} \in \{5, 10, \mathbf{16}, 32\}$, where the bold values indicate the selected optimal hyperparameters. Then, all the hyperparameters are displayed in Table I.

It is well-known that the model can learn high-level feature representations by increasing the depth of the network, so the influence of network depth on the prediction accuracy of DS-ViT-ESA is studied as shown in Table II. The enhanced

Authorized licensed use limited to: Xian Jiaotong University. Downloaded on October 26,2024 at 02:33:52 UTC from IEEE Xplore. Restrictions apply.
© 2024 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Transactions on Transportation Electrification. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TTE.2024.3434553

TABLE I
HYPERPARAMETERS SETTINGS.

Hyperparameters	Values
Batch size	512
Dropout	0.1
The aging cycles included in one patch C_{patch}	3
The point amounts of one curve E_{patch}	16
The dimension of embedding d_m	256
The number of attention heads h	8
Hidden neurons of the MLP layer in the encoder	256
Hidden neurons of the MLP layer in late fusion	512

expressive power enhances the RUL and CCL prediction accuracy when the network depth is increased from 1 to 2. When the network depth of the DS-ViT-ESA model is set to 2, the MAPE, RMSE, and MAE of RUL prediction are low as 5.40%, 15.45 cycles, and 10.28 cycles, respectively, and the MAPE, RMSE, and MAE of CCL prediction are low as 4.64%, 17.03 cycles, and 12.31 cycles, respectively. However, the prediction performance rapidly decreases while the network depth continues to increase, which might result from vanishing gradients or difficulties in information propagation in deep networks. Therefore, based on this result, the depth of the DS-ViT-ESA model is set to 2 unless otherwise stated in this work.

C. The battery RUL and CCL prediction performances of the DS-ViT-ESA model

Fig. 8a shows the RUL prediction result of 30 target batteries with 25 various charging strategies in the test dataset. Each battery is marked with a serial number (Battery ID), and the colours in the colour bar represent the Battery ID of 30 target batteries. The horizontal and vertical axes represent the measured and predicted RUL, respectively. The closer the point of one curve is to the black-dotted baseline, the lower the RUL prediction error, so the most predicted RUL values are extremely close to the corresponding real values, and the curves overlap each other. Fig. 8b illustrates the RUL prediction error result, and the proportion of RUL prediction error less than 40 cycles reaches up to 95% as shown in the frequency histogram.

Fig. 8c shows the CCL prediction results in the test dataset. Unlike the RUL prediction plot, the horizontal and vertical axes represent the measured CCL and the predicted CCL, respectively. A similar tendency also can be found, namely that the CCL prediction results nearly coincide with the black-dotted baseline. Fig. 8d illustrates the distribution of CCL prediction errors, and the proportion of CCL prediction errors less than 40 cycles reaches up to 95% as shown in the frequency histogram. Therefore, even though the lifetimes of target batteries vary from 400 to 1800 cycles, the model could mine the mapping relationship between the battery lifetime and the original terminal charging data.

Because some charging strategies in the test dataset don't appear in the training dataset, Table III displays the lifetime prediction result for target batteries with different charging strategies. Interestingly, these metrics for the batteries under the charging strategies that do not appear in the training dataset are just a little higher than those of the batteries with charging strategies that appear in the training dataset. Therefore, the DS-ViT-ESA model also displays effective zero-shot generalization capacity. Furthermore, a single battery from the tested batteries with charging strategies not appearing in the training dataset is picked up to observe the RUL (9a and b) and CCL (9c and d) prediction results. Most differences between predicted and true values are controlled within 20 cycles. Besides, the more accurate RUL and CCL predictions could be achieved when n_c is close to the EOL of this battery because the capacity degradation is distinct in the last stage, and there is more aging information.

D. The early battery lifetime prediction performance of the DS-ViT-ESA model

For one fresh battery, it is very prospective to accurately estimate its lifetime based on a few early aging cycles, so this proposed model's early lifetime prediction performance is studied in this work. To study the influence of the demand amount of tested aging cycles on the early prediction results, the different early aging cycles ($15 \leq n_c \leq 100$) are selected as the model input. Fig. 10a and b show the lifetime prediction and errors of 30 target batteries, in which the horizontal axis represents the true lifetime and vertical axis in Fig. 10a and b represent the predicted value, and vertical axis, respectively. Meanwhile, perpendicular to the horizontal axis, the dotted stub represents one battery. The dots' colour in one dotted stub is determined by the n_c displayed in the colour bar. Hence, the closer the point of one dotted stub is to the black-dotted baseline, the more accurate early lifetime prediction is. The detailed results of early lifetime prediction are displayed in Table IV, and the absolute error can be controlled within 5.0% for nearly all target batteries. Besides, as displayed in Fig. 10b, the predicted errors of the red points (n_c near 100) are closer to zero compared with those of the blue points (n_c near 15), indicating the bigger n_c is in favor of the improved lifetime prediction accuracy. This is consistent with the result of Section IV-C. As shown in Table IV, even when the fresh batteries just run for the first 15 aging cycles (n_c = 15), the accurate early lifetime prediction also could be realized with MAPE, RMSE, and MAE less than 2.16%, 25.88 cycles, and 20.52 cycles. Meanwhile, the most accurate early lifetime prediction of this model could be realized (MAPE: 1.82%; RMSE: 22.91 cycles; MAE: 18.05 cycles) when n_c increases to 80, but the test demand of initial aging cycles would boom from 15 to 80 and test duration would increase almost fivefold. Therefore, considering both the prediction accuracy and the operation cost, the n_c could be set to 15 for the early battery lifetime prediction.

E. The influence of different input format

Comparative experiments are carried out with the different inputs to study the influence of the model input format on the battery lifetime prediction. Namely, one parameter is individually removed from the $V/I/T/Q$ combination data. As shown in Table V, compared with the results of any three-parameter combinations, the best RUL and CCL prediction performance

Authorized licensed use limited to: Xian Jiaotong University. Downloaded on October 26,2024 at 02:33:52 UTC from IEEE Xplore. Restrictions apply.
© 2024 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Transactions on Transportation Electrification. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TTE.2024.3434553

TABLE II
THE INFLUENCE OF NETWORK DEPTH ON THE PREDICTION ACCURACY OF DS-VIT-ESA.

Depth	RUL prediction				CCL prediction			
	1	2	3	4	1	2	3	4
MAPE (%)	9.41	5.40	6.02	7.87	11.05	4.64	6.80	8.52
RMSE (Cycles)	28.33	15.45	16.14	20.07	34.20	17.03	21.07	25.36
MAE (Cycles)	20.40	10.28	11.70	14.42	24.85	12.31	16.14	20.75

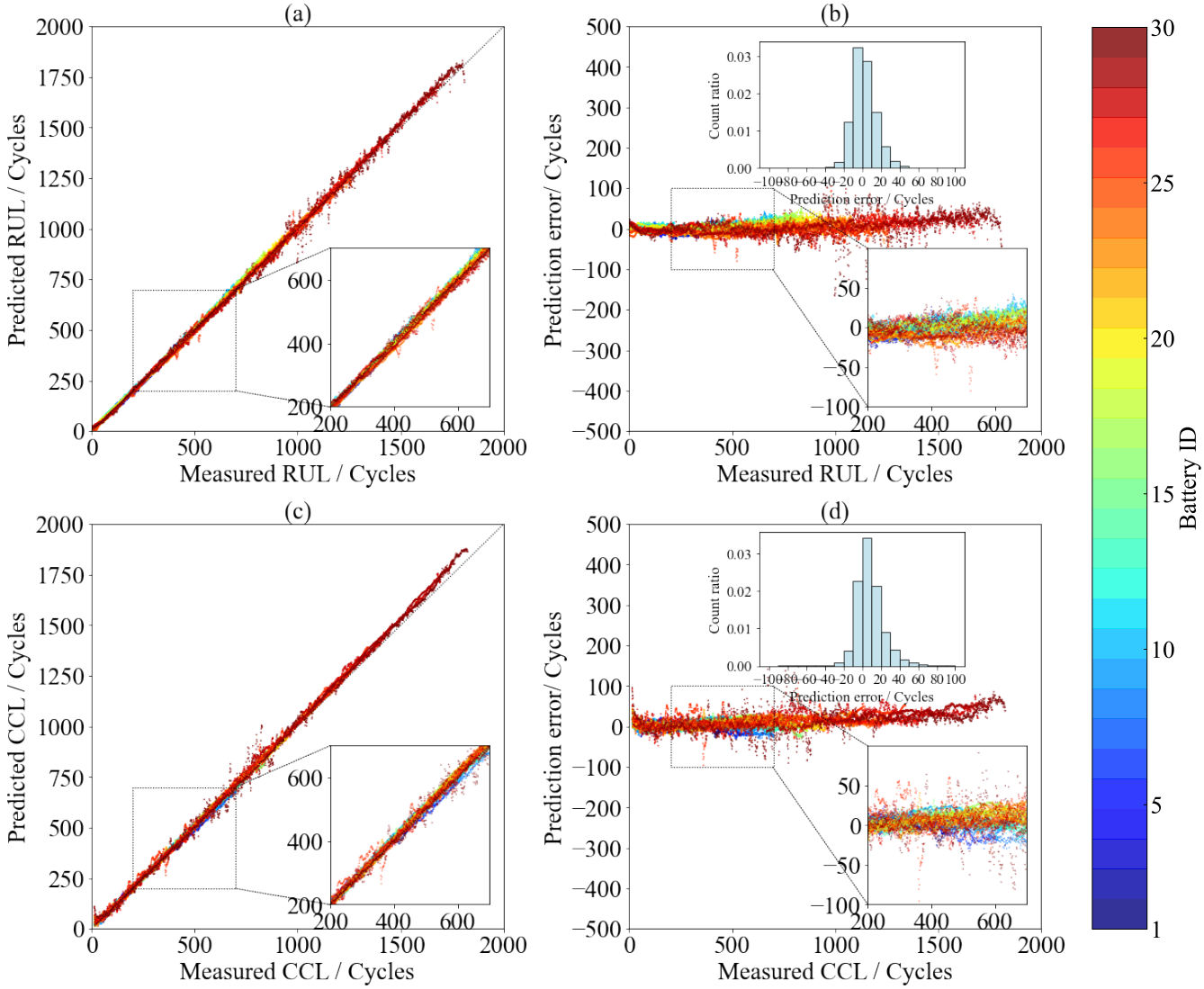

Fig. 8. Battery lifetime prediction of 30 target batteries: (a) RUL prediction results; (b) RUL prediction errors, and the inset is the frequency histogram of RUL prediction error; (c) CCL prediction results; (d) CCL prediction errors, and the inset is the frequency histogram of CCL prediction error.

could be realized when all terminal parameters are considered, so all terminal parameters should be collected and then serve as the model input. Besides, when the parameter T is not included in the model input, the evaluation metrics of RUL prediction exhibit the highest error (MAPE: 7.57%, RMSE: 20.63 cycles, MAE: 14.22 cycles). Meanwhile, when the parameter I is not included in the model input, the evaluation metrics of CCL prediction exhibit the highest error (MAPE: 9.81%, RMSE: 27.63 cycles, MAE: 21.37 cycles). Therefore, among these parameters, the parameter T and parameter I play vital roles in the accuracy of RUL and CCL prediction, respectively, and the relative terminal parameters should be paid more attention during the data collection.

F. Ablation experiment

To investigate the performance contribution of the dual-stream structure and the ESA mechanism, the ablation experiments are conducted via controlling the model structure: (a) dual-stream ViT with the self-attention mechanism (DS-ViT-SA); (b) ViT with the efficient self-attention mechanism (ViT-ESA); (c) ViT with the self-attention mechanism (ViT-SA). Meanwhile, the other baseline methods, (d) CNN model

Authorized licensed use limited to: Xian Jiaotong University. Downloaded on October 26,2024 at 02:33:52 UTC from IEEE Xplore. Restrictions apply.
© 2024 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Transactions on Transportation Electrification. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TTE.2024.3434553

TABLE III
THE LIFETIME PREDICTION RESULTS FOR TARGET BATTERIES WITH DIFFERENT CHARGING STRATEGIES.

Charging strategies	RUL prediction			CCL prediction		
	MAPE (%)	RMSE (Cycles)	MAE (Cycles)	MAPE (%)	RMSE (Cycles)	MAE (Cycles)
13 types appearing in the training dataset	4.17	14.25	9.39	4.14	14.71	10.48
12 types not appearing in the training dataset	6.89	16.36	11.54	5.36	19.94	14.84

TABLE IV
THE EARLY LIFETIME PREDICTION RESULTS OF THE DS-VIT-ESA MODEL WITH CERTAIN n_c.

	Early lifetime prediction for target batteries in the test dataset					
n_c	15	25	40	60	80	100
MAPE (%)	2.16	2.00	1.90	1.91	1.82	1.95
RMSE (Cycles)	25.88	25.66	24.16	23.96	22.91	23.14
MAE (Cycles)	20.52	19.59	18.75	18.38	18.05	18.73

TABLE V
THE INFLUENCE OF THE INPUT FORMAT ON THE BATTERY LIFETIME PREDICTION.

Different input	RUL prediction			CCL prediction		
	MAPE (%)	RMSE (Cycles)	MAE (Cycles)	MAPE (%)	RMSE (Cycles)	MAE (Cycles)
V/*I*/*T*/*Q*	5.40	15.45	10.28	4.64	17.03	12.31
I/*T*/*Q*	6.62	16.40	11.06	5.63	19.59	14.75
V/*T*/*Q*	7.42	19.20	13.19	9.81	27.63	21.37
V/*I*/*Q*	7.57	20.63	14.22	7.24	23.70	17.04
V/*I*/*T*	7.49	19.38	13.50	8.48	25.46	19.69

TABLE VI
THE INFLUENCE OF THE INPUT FORMAT ON THE BATTERY LIFETIME PREDICTION.

Models	RUL prediction			CCL prediction			Training time (h)
	MAPE (%)	RMSE (Cycles)	MAE (Cycles)	MAPE (%)	RMSE (Cycles)	MAE (Cycles)	
DS-ViT-ESA	5.40	15.45	10.28	4.64	17.03	12.31	17.31
DS-ViT-SA	5.88	16.17	10.60	5.22	18.80	14.59	19.57
ViT-ESA	8.94	24.85	18.51	7.02	23.39	16.15	13.26
ViT-SA	9.36	26.42	19.33	8.31	24.62	18.32	14.95
CNN	53.48	139.79	97.96	37.55	126.16	89.33	22.43
CNN-LSTM	27.27	71.61	55.23	72.68	98.51	72.39	24.67

and (e) CNN-LSTM model are also employed to verify the superior lifetime prediction performance of the DS-ViT-ESA model. Table VI exhibits these models' battery lifetime prediction results. Comparing the prediction results of DS-ViT-ESA and ViT-ESA, as well as the results of DS-ViT-SA and ViT-SA, the dual-stream structure could bring higher prediction accuracy than the single-stream structure because the multi-timescale aging information is captured and then fused. Besides, comparing the prediction results of DS-ViT-ESA and DS-ViT-SA, as well as the results of ViT-ESA and ViT-SA, the introduction of the ESA mechanism in the transformer encoder also could bring higher prediction performance, implying this variant promotes more attention modules into the neural network. Meanwhile, the models with the ESA mechanism exhibit relatively lower training time, demonstrating that the ESA mechanism could decrease the computational complexity of the model. Besides, the dual-stream structure would introduce more parameters, so their training time is higher than the single-stream structure. The CNN model is also usually utilized for battery lifetime prediction because it can effectively capture relationships between different parameters in the input sequences. Meanwhile, the LSTM units are used to replace the full connection layers of the CNN model because they effectively capture and remember long-term dependencies through a series of gate mechanisms. Compared with the two baselines, the DS-ViT-ESA model exhibits superior battery lifetime prediction accuracy because it can extract higher-level hidden features. Meanwhile, the DS-ViT-ESA model also displays less training cost due to its parallel structure. Hence, this model would possess an outstanding predictive behavior on batteries' big data compared to another traditional network.

G. Prediction results under other datasets

The experimental temperature and cathode material of the above dataset are fixed (30 °C and LFP, respectively), while the graphite anode dominates the batteries' degradation process, so it is promising to employ this approach to other graphite anode batteries. To testify to the generality of the proposed approach, the two other open datasets from the University of Maryland [43] and NASA [44] are employed to evaluate the developed

Authorized licensed use limited to: Xian Jiaotong University. Downloaded on October 26,2024 at 02:33:52 UTC from IEEE Xplore. Restrictions apply.
© 2024 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Transactions on Transportation Electrification. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TTE.2024.3434553

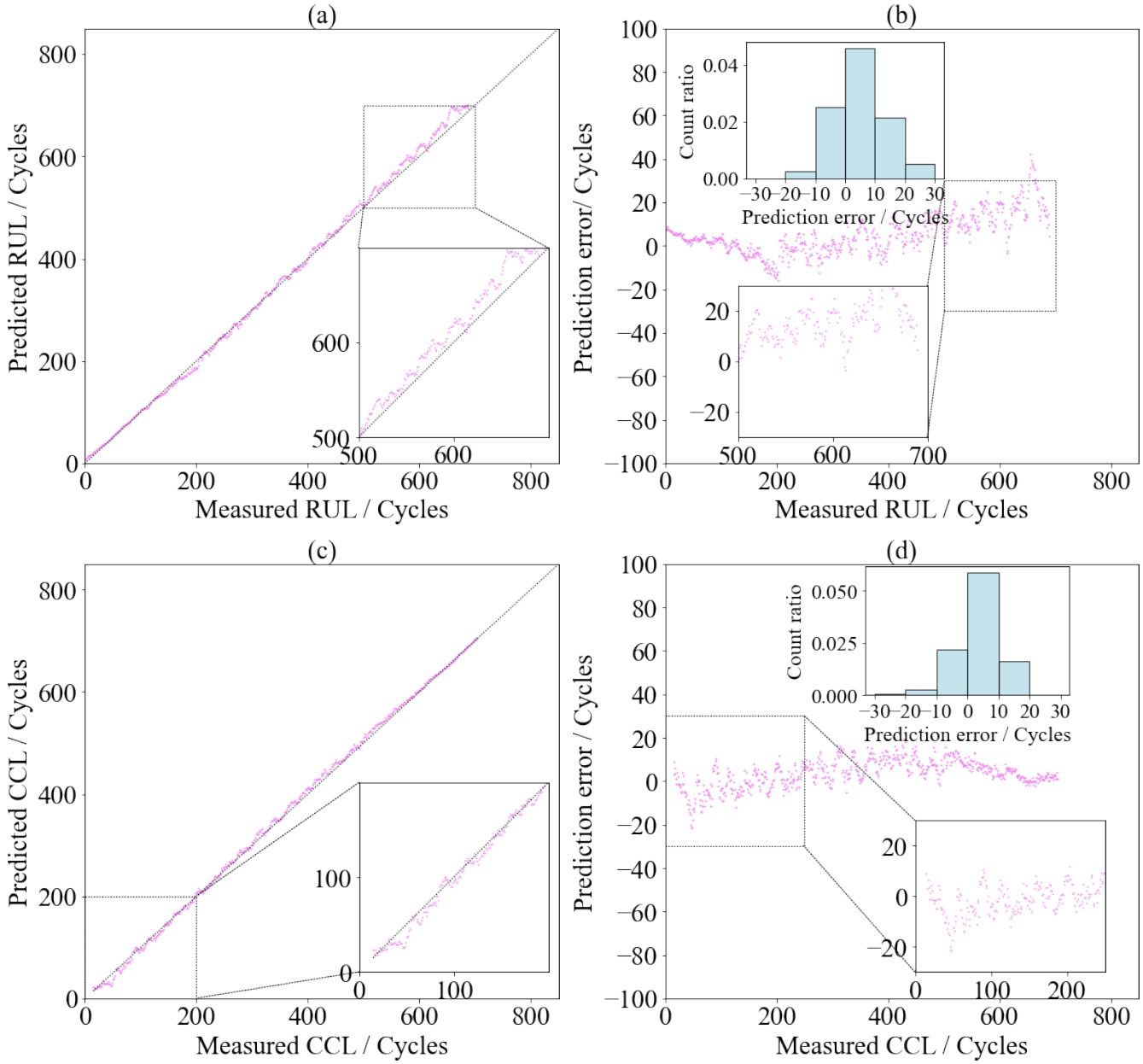

Fig. 9. An example of battery lifetime prediction of one battery: (a) RUL prediction results; (b) RUL prediction errors, and the inset is the frequency histogram of RUL prediction error; (c) CCL prediction results; (d) CCL prediction errors, and the inset is the frequency histogram of CCL prediction error.

approach. Table VII lists their experimental temperature and cathode material. The batteries with 1.10 Ah (No.35, No.36, No.37, and No.38) from the University of Maryland were tested at a standard CC-CV protocol with a current rate of 0.5C until the voltage reached 4.2 V. Then 4.2 V was sustained until the charging current dropped to below 0.05 A. This dataset did not collect the parameter T, so the other three parameters served as the model input. The batteries were discharged at the same rate (1C) to 2.7 V. The batteries (B0005, B0006, B0007, and B0018) from NASA have a rated capacity of 2 Ah, which were charged by the CC-CV protocol at the constant current of 1.5 A to the constant voltage of 4.2 V until the charge current dropped to below 20 mA, and then discharged at the constant current of 2 A to the different LCV (2.7 V for B0005, 2.5 V for B0006, 2.2 V for B0007, and 2.5 V for B0018). The rest of the battery data are utilized to train the proposed model from scratch, and the battery lifetime prediction results are also shown in Table VII. The prediction errors are still relatively low, even if the experimental condition changes drastically, indicating the proposed approach could apply to different environmental temperatures and cathode materials. Therefore, the results exhibit the outstanding robustness and generalization of the proposed model.

H. Compared with other published approaches.

To further study the prediction performances of the proposed model for battery lifetime, Table VIII lists some recently published models that exhibit good lifetime prediction accuracy.

Authorized licensed use limited to: Xian Jiaotong University. Downloaded on October 26,2024 at 02:33:52 UTC from IEEE Xplore. Restrictions apply.
© 2024 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Transactions on Transportation Electrification. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TTE.2024.3434553

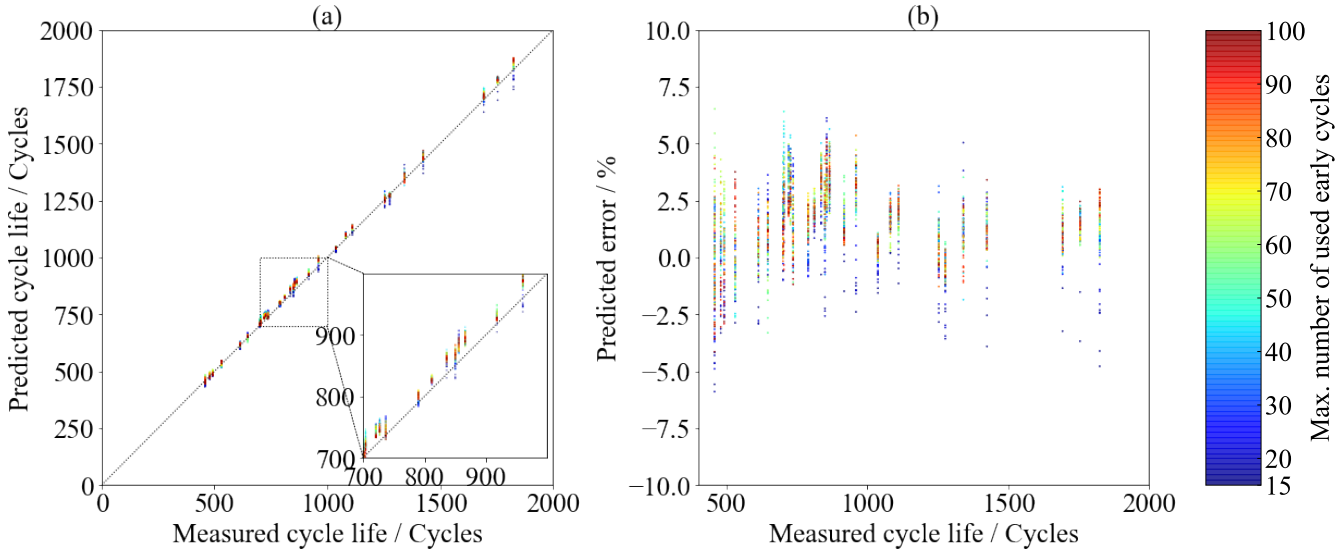

Fig. 10. Early lifetime prediction of 30 different batteries from their first 15∼100 charge cycles: (a) early lifetime prediction results; (b) prediction errors.

TABLE VII
THE BATTERY LIFETIME PREDICTION UNDER DIFFERENT EXPERIMENTAL CONDITIONS.

Battery	Cathode material	Environmental temperature (°C)	RUL prediction			CCL prediction		
			MAPE (%)	RMSE (Cycles)	MAE (Cycles)	MAPE (%)	RMSE (Cycles)	MAE (Cycles)
No.35	$LiCoO_2$	25∼30	4.69	5.82	4.74	4.59	5.97	5.09
B0018	$LiNiCoAlO_2$	24	3.54	4.65	3.44	4.05	4.95	4.31

Compared with HCNN and TOP-Net(100) models, the proposed model exhibits nearly state-of-the-art performance for early lifetime prediction. In this work, the minimum demand (15 cycles) of early aging cycles could be realized, which is more beneficial for practical battery design and optimization. Meanwhile, compared with the machine learning approach, the proposed method exhibits higher accuracy (implied by low MAPE) and stability (implied by low RMSE) even when the demand for early aging cycles is dropped from 100 to 15. Compared with the UKF-RVM method for RUL prediction, the proposed method exhibits a lower MAPE with less aging-cycle demand. The TOP-Net(100) method performs better RUL prediction than the proposed model, while its aging-cycle demand (100 cycles) is much higher than that of the proposed method. Besides, the 100-cycle full discharging data under constant current mode is necessary for the TOP-Net(100) method to realize good RUL accuracy. At the same time, it is impractical because the working load of the electric device is dynamic and random. Although the RUL prediction accuracy of the HCNN method is also better than that of the proposed method, the aging-cycle demand of the proposed model is lower than that of the HCNN method (20 cycles). Besides, this proposed approach's accurate CCL prediction ability could also help researchers further grasp the battery's health status.

V. CONCLUSION

This study proposes a novel DL method (donated as DS-ViT-ESA) to realize end-to-end LIBs' lifetime prediction, in which a small amount of charging data is collected. DS-ViT-ESA model integrates ViT structure, dual-stream framework, and ESA mechanism, which could more efficiently capture inner-cycle and cycle-to-cycle aging information. Experiments display that the proposed model produces an accurate battery lifetime prediction (RUL: 5.40%; CCL: 4.64%) using only 15 charging cycles, demonstrating the effectiveness and practicality of the model. The model also exhibits a low prediction error (RUL: 6.89%, CCL: 5.36%) for the charging strategies not appearing in the training dataset, implying the effective zero-shot generalization capacity. Moreover, the early lifetime prediction with an error as low as 2.16% is also achieved by utilizing only the first 15 charging cycles. Although the prediction performance of this proposed model is not state-of-the-art, the amount of demand for aging cycles is lowest while ensuring the prediction accuracy and training cost. More importantly, RUL and CCL could also be gained together, which is beneficial for researchers to design and optimize new batteries.

While this model has achieved relatively satisfied lifetime prediction performance at a lower aging cycle demand, there are still some limitations. Firstly, the approach works when all the charging data is certain throughout the battery aging process, especially for fixed application scenarios such as electric unmanned delivery cars and electric buses. However, the uncertain charging behavior which approximates real electric vehicles is not investigated. Then, the approach requires the initial aging cycle to form the difference matrix, but the first cycle is unknown sometimes, and it is hard to use this dual-stream structure to predict the lifetime accurately.

Authorized licensed use limited to: Xian Jiaotong University. Downloaded on October 26,2024 at 02:33:52 UTC from IEEE Xplore. Restrictions apply.
© 2024 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Transactions on Transportation Electrification. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TTE.2024.3434553

TABLE VIII
THE BATTERY LIFETIME PREDICTION UNDER DIFFERENT EXPERIMENTAL CONDITIONS.

Task	Evaluation metrics	Approaches				
		Proposed method	HCNN[b][35]	Machine-learning approach [34]	TOP-Net(100)[c][45]	UKF-RVM[d][46]
Early lifetime prediction	Early cycles used	15	60	100	100	-
	MAPE (%)	2.16	1.12	7.5, 10.7[f]	1.47	-
	RMSE (Cycles)	25.88	13	100, 214[f]	11	-
	MAE (Cycles)	20.52	11	-	9	-
RUL prediction	Requirement [a]/ %	1∼3	1∼4	-	4∼20	38∼96
	MAPE (%)	5.40	3.55	-	2.85	14.64
	RMSE (Cycles)	15.45	11	-	9	-
	MAE (Cycles)	10.28	9	-	7	-
CCL prediction	Requirement / %	1∼3	-	-	4∼20	-
	MAPE (%)	4.64	-[e]	-	1.25	-
	RMSE (Cycles)	17.03	-	-	8	-
	MAE (Cycles)	12.31	-	-	6	-

[a] the ratio of the required aging cycles for lifetime prediction into the EOL.
[b] a hybrid deep learning approach consisting of a CNN network and attention mechanism.
[c] a two-dimensional CNN and one-dimensional CNN parallel hybrid deep learning approach.
[d] an integrated Unscented Kalman Filter (UKF) and RVM method.
[e] the relative metric of the method was not conducted.
[f] the two values are gained from two tests, respectively.

Nevertheless, this model gains considerable lifetime prediction performance in a large-scale fast-charging battery dataset and has the potential to be extended to real applications.

REFERENCES

[1] Y. Liu, B. Hou, M. Ahmed, Z. Mao, J. Feng, and Z. Chen, "A hybrid deep learning approach for remaining useful life prediction of lithium-ion batteries based on discharging fragments," *Appl. Energy*, vol. 358, p. 122555, 2024.

[2] P. Li, Z. Zhang, R. Grosu, Z. Deng, J. Hou, Y. Rong, and R. Wu, "An end-to-end neural network framework for state-of-health estimation and remaining useful life prediction of electric vehicle lithium batteries," *Renewable Sustainable Energy Rev.*, vol. 156, p. 111843, 2022.

[3] W. Li, N. Sengupta, P. Dechent, D. Howey, and D. U. Sauer, "One-shot battery degradation trajectory prediction with deep learning," *J. Power Sources*, vol. 506, no. 1, p. 230024, 2021.

[4] G. Ma, Y. Zhang, C. Cheng, B. Zhou, and Y. Yuan, "Remaining useful life prediction of lithium-ion batteries based on false nearest neighbors and a hybrid neural network," *Appl. Energy*, vol. 253, p. 113626, 2019.

[5] X. Hu, Y. Che, X. Lin, and Z. Deng, "Health prognosis for electric vehicle battery packs: A data-driven approach," *IEEE-ASME Trans. Mechatron.*, vol. 25, no. 6, pp. 2622–2632, 2020.

[6] X. Zhang, Y. Qin, C. Yuen, L. Jayasinghe, and X. Liu, "Time-series regeneration with convolutional recurrent generative adversarial network for remaining useful life estimation," *IEEE Trans. Ind. Inform.*, vol. 17, no. 10, pp. 6820–6831, 2021.

[7] C. Hu, B. D. Youn, and J. Chung, "A multiscale framework with extended kalman filter for lithium-ion battery soc and capacity estimation," *Appl. Energy*, vol. 92, pp. 694–704, 2012.

[8] M. Catelani, L. Ciani, R. Fantacci, G. Patrizi, and B. Picano, "Remaining useful life estimation for prognostics of lithium-ion batteries based on recurrent neural network," *IEEE Trans. Instrum. Meas.*, vol. 70, pp. 1–11, 2021.

[9] C. Lin, J. Xu, J. Hou, Y. Liang, and X. Mei, "Ensemble method with heterogeneous models for battery state-of-health estimation," *IEEE Trans. Ind. Inform.*, vol. 19, no. 10, pp. 10 160–10 169, 2023.

[10] Z. Deng, X. Hu, X. Lin, L. Xu, Y. Che, and L. Hu, "General discharge voltage information enabled health evaluation for lithium-ion batteries," *IEEE-ASME Trans. Mechatron.*, vol. 26, no. 3, pp. 1295–1306, 2020.

[11] X. Hu, Y. Che, X. Lin, and S. Onori, "Battery health prediction using fusion-based feature selection and machine learning," *IEEE Trans. Transp. Electrif.*, vol. 7, no. 2, pp. 382–398, 2020.

[12] M. Wei, P. Balaya, M. Ye, and Z. Song, "Remaining useful life prediction for 18650 sodium-ion batteries based on incremental capacity analysis," *Energy*, vol. 261, p. 125151, 2022.

[13] S. S. Afshari, S. Cui, X. Xu, and X. Liang, "Remaining useful life early prediction of batteries based on the differential voltage and differential capacity curves," *IEEE Trans. Instrum. Meas.*, vol. 71, pp. 1–9, 2021.

[14] X. Li, C. Yuan, Z. Wang, J. He, and S. Yu, "Lithium battery state-of-health estimation and remaining useful lifetime prediction based on non-parametric aging model and particle filter algorithm," *eTransportation*, vol. 11, p. 100156, 2022.

[15] B. Zraibi, C. Okar, H. Chaoui, and M. Mansouri, "Remaining useful life assessment for lithium-ion batteries using cnn-lstm-dnn hybrid method," *IEEE Trans. Veh. Technol.*, vol. 70, no. 5, pp. 4252–4261, 2021.

[16] M. Wei, H. Gu, M. Ye, Q. Wang, X. Xu, and C. Wu, "Remaining useful life prediction of lithium-ion batteries based on monte carlo dropout and gated recurrent unit," *Energy Rep.*, vol. 7, pp. 2862–2871, 2021.

[17] R. Jiao, K. Peng, and J. Dong, "Remaining useful life prediction of lithium-ion batteries based on conditional variational autoencoders-particle filter," *IEEE Trans. Instrum. Meas.*, vol. 69, no. 11, pp. 8831–8843, 2020.

[18] J. Hong, D. Lee, E.-R. Jeong, and Y. Yi, "Towards the swift prediction of the remaining useful life of lithium-ion batteries with end-to-end deep learning," *Appl. Energy*, vol. 278, p. 115646, 2020.

[19] J. Wang and Y. Xiang, "Fast modeling of the capacity degradation of lithium-ion batteries via a conditional temporal convolutional encoder–decoder," *IEEE Trans. Transp. Electrif.*, vol. 8, no. 2, pp. 1695–1709, 2021.

[20] A. Tang, Y. Jiang, Y. Nie, Q. Yu, W. Shen, and M. G. Pecht, "Health and lifespan prediction considering degradation patterns of lithium-ion batteries based on transferable attention neural network," *Energy*, vol. 279, p. 128137, 2023.

[21] T. Tang and H. Yuan, "A hybrid approach based on decomposition algorithm and neural network for remaining useful life prediction of lithium-ion battery," *Reliab. Eng. Syst. Saf.*, vol. 217, p. 108082, 2022.

[22] Z. Chen, L. Chen, W. Shen, and K. Xu, "Remaining useful life prediction of lithium-ion battery via a sequence decomposition and deep learning integrated approach," *IEEE Trans. Veh. Technol.*, vol. 71, no. 2, pp. 1466–1479, 2021.

[23] D. Pan, H. Li, and S. Wang, "Transfer learning-based hybrid remaining useful life prediction for lithium-ion batteries under different stresses," *IEEE Trans. Instrum. Meas.*, vol. 71, pp. 1–10, 2022.

[24] Y. Che, Z. Deng, X. Lin, L. Hu, and X. Hu, "Predictive battery health management with transfer learning and online model correction," *IEEE Trans. Veh. Technol.*, vol. 70, no. 2, pp. 1269–1277, 2021.

[25] G. Ma, S. Xu, B. Jiang, C. Cheng, X. Yang, Y. Shen, T. Yang, Y. Huang, H. Ding, and Y. Yuan, "Real-time personalized health status prediction of lithium-ion batteries using deep transfer learning," *Energy Environ. Sci.*, vol. 15, no. 10, pp. 4083–4094, 2022.

[26] O. S. Kayhan and J. C. v. Gemert, "On translation invariance in cnns: Convolutional layers can exploit absolute spatial location," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2020, pp. 14 274–14 285.

[27] T. Shen, T. Zhou, G. Long, J. Jiang, and C. Zhang, "Bi-directional block self-attention for fast and memory-efficient sequence modeling," *arXiv preprint arXiv:1804.00857*, 2018.

[28] A. Radford, J. W. Kim, C. Hallacy, A. Ramesh, G. Goh, S. Agarwal, G. Sastry, A. Askell, P. Mishkin, J. Clark *et al.*, "Learning transferable visual models from natural language supervision," in *Proc. Int. Conf. Mach. Learn.* PMLR, 2021, pp. 8748–8763.

[29] J. Li, D. Li, C. Xiong, and S. Hoi, "Blip: Bootstrapping language-image pre-training for unified vision-language understanding and generation," in *Proc. Int. Conf. Mach. Learn*. PMLR, 2022, pp. 12 888–12 900.

[30] H. Touvron, M. Cord, M. Douze, F. Massa, A. Sablayrolles, and H. Jégou, "Training data-efficient image transformers & distillation through attention," in *Proc. Int. Conf. Mach. Learn*. PMLR, 2021, pp. 10 347–10 357.

[31] S. Zhuoran, Z. Mingyuan, Z. Haiyu, Y. Shuai, and L. Hongsheng, "Efficient attention: Attention with linear complexities," in *Proc. IEEE Winter Conf. Appl. Comput. Vis. (WACV)*, Jan. 2021, pp. 3530–3538.

[32] D. Han, X. Pan, Y. Han, S. Song, and G. Huang, "Flatten transformer: Vision transformer using focused linear attention," in *Proc. IEEE/CVF Int. Conf. Comput. Vis.(ICCV)*, October 2023, pp. 5938–5948.

[33] P. Fu, L. Chu, Z. Hou, J. Hu, Y. Huang, and Y. Zhang, "Transfer learning and vision transformer based state-of-health prediction of lithium-ion batteries," *arXiv preprint arXiv:2209.05253*, 2022.

[34] K. A. Severson, P. M. Attia, N. Jin, N. Perkins, B. Jiang, Z. Yang, M. H. Chen, M. Aykol, P. K. Herring, D. Fraggedakis *et al.*, "Data-driven prediction of battery

Authorized licensed use limited to: Xian Jiaotong University. Downloaded on October 26,2024 at 02:33:52 UTC from IEEE Xplore. Restrictions apply.
© 2024 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Transactions on Transportation Electrification. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TTE.2024.3434553

cycle life before capacity degradation," *Nat. Energy*, vol. 4, no. 5, pp. 383–391, 2019.

[35] Y. Yang, "A machine-learning prediction method of lithium-ion battery life based on charge process for different applications," *Appl. Energy*, vol. 292, p. 116897, 2021.

[36] Q. Zhang, L. Yang, W. Guo, J. Qiang, C. Peng, Q. Li, and Z. Deng, "A deep learning method for lithium-ion battery remaining useful life prediction based on sparse segment data via cloud computing system," *Energy*, vol. 241, p. 122716, 2022.

[37] A. Dosovitskiy, L. Beyer, A. Kolesnikov, D. Weissenborn, X. Zhai, T. Unterthiner, M. Dehghani, M. Minderer, G. Heigold, S. Gelly *et al.*, "An image is worth 16x16 words: Transformers for image recognition at scale," *arXiv preprint arXiv:2010.11929*, 2020.

[38] Z. Wan, Y. Kang, R. Ou, S. Xue, D. Xu, and X. Luo, "Multi-step time series forecasting on the temperature of lithium-ion batteries," *J. Energy Storage*, vol. 64, p. 107092, 2023.

[39] X. Gu, K. See, P. Li, K. Shan, Y. Wang, L. Zhao, K. C. Lim, and N. Zhang, "A novel state-of-health estimation for the lithium-ion battery using a convolutional neural network and transformer model," *Energy*, vol. 262, p. 125501, 2023.

[40] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, and I. Polosukhin, "Attention is all you need," *Adv. Neural Inf. Process. Syst.*, vol. 30, 2017.

[41] Y. Gao, M. Zhou, and D. N. Metaxas, "Utnet: a hybrid transformer architecture for medical image segmentation," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.* Cham, Switzerland: Springer, 2021, pp. 61–71.

[42] Z. Fei, Z. Zhang, and K.-L. Tsui, "Deep learning powered online battery health estimation considering multi-timescale ageing dynamics and partial charging information," *IEEE Trans. Transp. Electrif.*, Apr. 05, 2023.

[43] M. Pecht, "Battery data set." *CALCE, CALCE Battery Research Group, Maryland, MD*, 2017.

[44] B. Sahaand and K. Goebel, "Battery data set." *CALCE, CALCE Battery Research Group, Maryland, MD*, 2007.

[45] D. Chen, W. Zhang, C. Zhang, B. Sun, X. Cong, S. Wei, and J. Jiang, "A novel deep learning-based life prediction method for lithium-ion batteries with strong generalization capability under multiple cycle profiles," *Appl. Energy*, vol. 327, p. 120114, 2022.

[46] X. Zheng and H. Fang, "An integrated unscented kalman filter and relevance vector regression approach for lithium-ion battery remaining useful life and short-term capacity prediction," *Reliab. Eng. Syst. Saf.*, vol. 144, pp. 74–82, 2015.

Authorized licensed use limited to: Xian Jiaotong University. Downloaded on October 26,2024 at 02:33:52 UTC from IEEE Xplore. Restrictions apply.
© 2024 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.