Applied Energy 292 (2021) 116897

Contents lists available at ScienceDirect

Applied Energy

journal homepage: www.elsevier.com/locate/apenergy

A machine-learning prediction method of lithium-ion battery life based on charge process for different applications

Yixin Yang

School of Electronic and Information Engineering, Xi'an Jiao Tong University, Xi'an 710049, China

HIGHLIGHTS

- Established a hybrid CNN for both battery life early prediction and RUL prediction.
- Described battery life by several charge V, I and T curves and their difference.
- Applied a feature attention algorithm able to reduce errors by up to 2.7 times.
- Applied a cycle attention algorithm able to reduce errors by up to 3.3 times.
- Achieved 1.1% test error for battery life early prediction and 3.6% for RUL.

ARTICLE INFO

Keywords:
Cycle life early prediction
Remaining useful life prediction
Lithium-ion battery
Hybrid convolutional neural network
Feature and cycle attention

ABSTRACT

For accelerating the technology development and facilitating the reliable operation of lithium-ion batteries, accurate prediction for battery cycle life and remaining useful life (RUL) are both critical. However, diverse aging mechanisms, significant device variability and random working conditions have remained challenges. A reasonable description and an effective prediction algorithm are indispensable for achieving accurate prediction results. In this paper, battery terminal voltage, current and temperature curves from several charge cycles and especially their difference between these cycles are first utilized for description of battery cycle life and RUL. Moreover, a hybrid convolutional neural network (CNN), which is based on a fusion of three-dimensional CNN and two-dimensional CNN, is designed for their predictions. The battery charge voltage, current and temperature and their curves are first fused for considering the strong relationships between them. And the features hidden in the curves are extracted and modelled automatically. Furthermore, a feature attention algorithm and a multi-scale cycle attention algorithm are proposed to estimate the relationships between different features and cycles respectively for further heightening the prediction performance. Experiments and comparisons are conducted. The results show that the proposed method is an accurate method for different applications. It achieved 1.1% test error for battery cycle life early prediction of different batteries under different charge policies, and 3.6% for RUL prediction.

1. Introduction

Lithium-ion batteries are deployed in a wide range of applications due to their low pollution, high energy–density, high power-density and long lifetimes [1]. It is inevitable to evaluate the battery life completely and repeatedly during the development while the existing life test will take a long time [2]. As is the case with many chemical, mechanical and electronic systems, long battery lifetime entails delayed feedback of performance often many months to years [3]. And during the life span of a battery, its electrical performance would change with battery remaining useful life (RUL) [4]. When the performance degradation exceeds a certain threshold, lithium-ion batteries may suffer from severe consequences such as rapid performance degradation or catastrophic failure [5]. Therefore, an accurate prediction of battery life is very helpful for accelerating the battery development, and is essential for proposing a more scientific battery management system.

The early prediction, that is to predict battery lifetime using the early-cycle data at the early stage of battery, would unlock new opportunities in battery production, use and optimization. For example, if the life of a battery with final life of 2000 cycles can be predicted by using the early 100 cycles of the battery, 1900 cycles of testing can be avoided, thus saving 95% of the test time and cost. And online prediction of RUL, that is to predict the battery RUL during the lifetime of battery in

E-mail address: mjdfx090103@stu.xjtu.edu.cn.

https://doi.org/10.1016/j.apenergy.2021.116897
Received 26 October 2020; Received in revised form 19 March 2021; Accepted 30 March 2021
Available online 13 April 2021

Nomenclature

2DCNN	two-dimensional convolutional neural network
3DCNN	three-dimensional convolutional neural network
CC	constant current
CFA	charging feature attention
CNN	convolutional neural network
CV	constant voltage
DCN	dilated convolution network
DNN	deep neural network
HCNN	hybrid convolutional neural network
LSTM	long short-term memory
MAE	mean absolute error
MAPE	mean-absolute-percentage-error
MsCCA	multi-scale charging cycle attention
OCV	open circuit voltage
ReLU	rectified linear unit
RMSE	root-mean-squared-error
RUL	remaining useful life
SOC	state of charge
RNN	recurrent neural network
RVM	relevance vector machine
UKF	unscented Kalman filter
UPF	unscented particle filter

its operation in practical devices, can ensure the battery to run more safely and efficiently as well as help users establish better maintenance strategies. Therefore, the accurate prediction of lithium-ion battery life for these different applications is critically important, but challenging due to nonlinear degradation with cycling and wide variability, and such random operating conditions as the working loads and environment [6].

1.1. Literature review

Batteries with different lives or under different RULs would have different internal electrochemical properties. They would also have different electrical properties, such as battery charge/discharge performance, total available capacity, peak power and so on. In recent years, several descriptions have been utilized to characterize battery life. Sarasketa-Zabala et al. [7] applied exponential function to calculate the loss percentage of lithium-ion battery capacity, because capacity could gradually decrease with various aging process. He et al. [8] employed a sum of two exponential functions of discharge cycles to model the battery capacity fade as well. Furthermore, resistance can serve as a health indicator for describing the RUL of lithium-ion batteries. Eddahech et al. [9] used a single parameter identified from electrochemical impedance spectroscopy (EIS) tests to identify the battery RUL. Pattipati et al. [10] built a modified Randles circuit model, and used the high-frequency resistance and charge-transfer resistance to calculate the power fade and capacity fade for battery RUL description. However, the description of battery life using one or two battery parameter values may not be robust enough and may be lopsided. Zhou et al. [11] found a linear correlation between mean voltage falloff (MVF) and capacity, and then used the MVF as a health indicator to predict RUL. Sun et al. [12] considered 6 influential factors to identify RUL based on the evaluation of the battery performance characteristics, analyses on their disparities, and opinions of the experts. He et al. [13] described the battery aging state by several battery state of charges (SOCs) since battery with different aging states would have different SOC-OCV (open circuit voltage) curves. Severson et al. [3] proposed a feature-based approach and built a model with 9 features selected out of 20 for battery life early prediction. Different researchers proposed or selected different features for battery life description. What and how many features do we need to describe the battery life? There is still no definite answer due to the diverse aging mechanisms of lithium-ion battery. A few works, for example that by Ren et al. [14], tried to address this issue by using battery voltage curve under constant current to predict RUL. However, it is obvious that they cannot work well in actual operation where the current through battery usually varies. Therefore, a reasonable definition for battery life and RUL is essential for accurate prediction.

Based on a rational definition for battery life and RUL, a prediction algorithm is also quite necessary for battery life and RUL estimation. State estimation algorithms, including Kalman filter (KF) and particle filter (PF) have been widely used in battery RUL prediction. Dalal et al. [15] built a lumped parameter battery model and subsequently used in a PF framework to predict the battery RUL. A hybrid method developed by Chang et al. [16] in which an unscented KF (UKF) was combined with a double exponential model was adopted for RUL prediction. Zhang et al. [17] obtained the prediction results of battery RUL by using an improved PF algorithm-unscented particle filter (UPF) based on linear optimizing combination resampling, which can predict battery RUL with an error less than 5%. Recently, thanks to the advances in computational power and data generation, prediction methods based on data-driven and machine-learning techniques become attractive. The support vector machine (SVM) is a well-known machine learning method. Nuhic et al. [18] proposed an RUL monitoring method using SVM. Feng et al. [19] developed a predictive diagnosis method by comparing partial charging curves with the stored SVMs. Qin et al. [20] built an improved particle swarm optimization (PSO)–support vector regression (SVR) model to estimation RUL under different fault thresholds. A method which integrated relevance vector machine (RVM) and UKF to simulate battery degradation was proposed by Zheng et al. [21]. The Gaussian process regression (GPR) is a type of Bayesian non-parametric machine learning method, which can handle uncertainty in a complex model. Richardson et al. [22] proposed three multi-output GPR models for RUL prediction by incorporating data from multiple batteries. In their later work [23], a GPR transition model was proposed to predict the capacity degradation and battery RUL under dynamic conditions. In addition, neural networks appear to be promising for RUL predictions of lithium-ion batteries. The Recurrent Neural Network (RNN) is a commonly used method to predict unknown sequences. Liu et al. [24] confirmed that the adaptive RNN shows better a learning capability than classical training algorithms, including the RVM and PF methods. The long short-term memory (LSTM) was applied to predict lithium-ion battery RUL by Ma et al. [25] and Zhang et al. [26]. Liu et al. [27] applied LSTM model to predict the residual of GPR for learning the correlations of the capacity time series. A deep learning approach was proposed for RUL prediction of lithium-ion battery by integrating auto-encoder with deep neural network by Ren et al. [14]. Different than the above researches focused on battery RUL prediction, Severson et al. [3] built a linear model and used regularization techniques for model setting and model selection for battery life early prediction, which can predict battery life based on the first 100 discharge cycles and with 9.1% test error achieved.

However, r(1) these advanced nonlinear filters, data-driven and machine-learning techniques might be unable to achieve good predictions in practical applications. For example, in the battery RUL prediction: the UKF only works well when the variance of the observation noise is small, while the UPF is limited by the particle degeneracy issue [28]; the empirically determined sliding window size used by RVM may result in controversy; and in the LSTM, an under-fitting phenomenon commonly occurs when the battery data are noisy or the quantity of data is small. (2) Some parameters for prediction such as battery capacity [10], electrochemical impedance spectroscopy [12], and OCV [13] are hard to measure in practical operations, due to the requirements of either specific equipment or particular charge/discharge schedule. (3)

The definition of battery life using several battery parameter values, such as the test temperature and the SOC in Ref. [9], and the 8 features extracted from the discharge cycles in Ref. [29], may not be robust enough and may be lopsided, because all cycling data points contribute more or less to the battery life prediction accuracy. (4) The existing methods are commonly based on the working process (especially, the discharge process) of battery under constant current. They are not suitable for practice application very well, because such working process always under constant current rarely exists in practice, due to the random working loads of the devices (e.g. the electric vehicles) with the battery [30].

Thus the existing methods may be inadequate for the early prediction of battery life or the prediction of battery RUL in the practical operations of a device. Moreover, it is conceivable that for the early prediction of battery life and the prediction of battery RUL in different devices' applications, if the method and model can be unified, it will be of great significance. The prediction accuracy can be guaranteed and consistent in different applications. Moreover, the time and cost of developing the prediction systems for different applications can be saved. However, in the existing literatures, the researches on such unified method and model have not been found.

1.2. Contributions in the work

Focusing on those problems, a hybrid convolutional neural network (HCNN) method for both early prediction of the cycle life of lithium-ion batteries and prediction of their RULs is proposed in this paper. It is known that batteries with different lives or under different RULs would have different charge and discharge performance, which may lead to the variation of the shapes of charge and discharge curves. To avoid the influence of the random working loads of devices on the predictions, the proposed prediction is based on battery charge process, due to the fact that the battery charge process is always controlled in all devices. And the robustness and accuracy of the proposed method for different batteries with different lives under different charge polices are systematically verified by experiments. Three main contributions have been made in this paper.

(1) Unlike the existing method commonly using some selected features, the battery cycle life and RUL are all described by charge process and battery terminal voltage, current and temperature (V/I/T) curves from several charge cycles, and in particular by the difference between these cycles as well, to make the description robust and reasonable. In addition to avoiding the influence of the random working loads of devices on battery life prediction, it can also make the prediction under different charge policies and with high accuracy become possible. Experiments prove that, if without such difference as input, the battery life prediction errors will increase by more than 2 times.
(2) A HCNN is then proposed to simulate the relationship between battery charge V/I/T curves under different charge policies, battery cycle life and RUL. Different than the conventional CNN, the proposed HCNN is based on a fusion of three-dimensional CNN (3DCNN) and two-dimensional CNN (2DCNN) for the first time for battery life prediction. The introduced 3DCNN layer can fuse these V/I/T curves and their difference between cycles for considering the strong relationship between them. And the 2DCNN layers can extract and model the features hidden in these curves comprehensively and automatically. It is proved by experiments that: errors of battery life prediction might increase by up to 79% if without the help of the introduced 3DCNN layer.
(3) Moreover, because different charging features and cycles might exhibit varying degrees of salience to the prediction of battery cycle life and RUL, a charging feature attention (CFA) algorithm and a multi-scale charging cycle attention (MsCCA) algorithm are defined to further heighten the prediction performance. The experimental results indicate that: the battery life prediction errors will increase by up to 2.7 times if without the proposed CFA, and by up to 3.3 times if without the assistance of the proposed MsCCA.

Then, by a set of weight and bias values which make up the proposed model, not only can the battery life be predicted early before battery capacity degradation for accelerating battery development, but also the battery RUL can be predicted during battery whole useful lifetime range for different devices' practical applications. And due to the ability of the proposed method to accurately predict the battery life under different charge policies, it can also be used to find a suitable charge policy for battery life.

To the best of our knowledge, the proposed method is the first one that can be used for both battery life early prediction and RUL prediction, and with such outstanding performance as follows. According to the experiment results, for different batteries under different charge policies: in RUL prediction, the proposed method achieves low test error as 3.6%; in early prediction of battery cycle life, the test error is merely around 1% only using the charge data from the first 60 cycles (exhibiting a median increase of 0.1% from initial capacity); and when compared with other state-of-the-art methods, the root-mean-squared-error (RMSE) can be reduced by 8–17 times. The results indicate that the proposed method can offer prediction with high accuracy, strong stability and wide generality.

1.3. Organization of this paper

The rest of this paper is organized as follows. In Section 2, battery charge process in practical applications is analyzed, and then the capacity-indexed battery voltage, current and temperature curves from such processes are introduced for the description of battery life states. Afterward, the methodology for battery cycle life early prediction and RUL prediction is presented in Section 3. After introducing the input data generation, the HCNN is proposed, and the novel charging feature attention algorithm and multi-scale charging cycle attention algorithm are defined. To verify the proposed method, experiments and comparison with other methods are conducted and the results are demonstrated in Section 4. After the results are discussed in Section 5, conclusions are drawn in Section 6.

2. Description of battery life

Batteries with different lives or at different life states would have different performance. Based on this, a battery life description is proposed for battery life prediction in this paper.

2.1. Definition of battery life

Due to the irreversible electrochemical degradation, the discharge performance of battery will become poor, the battery goes towards to the end of its life. When the degradation reaches a predefined threshold (e.g., battery actual capacity drops below 80% of its initial value), the battery should be replaced. Accordingly, the life of lithium-ion battery is generally defined by the cycle number under experiment, in which the battery is charged with a certain charge policy and discharged with a constant current under a constant temperature. And the battery RUL in actual operations can be expressed as:

$$RUL = N_{EOL} - N_{ECL} \tag{1}$$

where N_{EOL} represents the battery life (also known as cycle life), achieved usually by battery life experiment. N_{ECL} is the equivalent cycle number of battery in actual applications. Obviously in actual operations, a battery might have polytrophic charge/discharge current rates and sometimes different environment temperatures. N_{ECL} is used to formu-

late the cycle number via actual operation into cycle number under the battery life experiment.

In order to predict the battery life early before the battery capacity degradation, the most important is to find out the early features hidden in the charge and/or discharge performance and their models to simulate N_{EOL}. And the battery RUL can be predicted similarly to avoid the influence of the random working loads and environments of battery in actual operations on the estimation of N_{ECL}.

2.2. Battery charge and discharge process

There are three status of battery in practice: charge, discharge and rest. Because the working loads of the devices such as the electric vehicles might change rapidly in their actual operations, the current passing through battery changes rapidly. It is hard to measure or calculate the internal and external parameters of battery accurately during such discharge process. In the status of rest, the battery parameters are generally unvarying or varying slowly, while parameters estimation cannot be calculated based on the indistinctive data. By contrast, the charge process is usually peaceful and always controlled either in the real applications on devices or in the development process of battery. As a result, the necessary external performance can be easily and accurately measured. Several battery parameters, such as the battery charged capacity, internal resistance and open circuit voltage, can be calculated or estimated during charging.

Charge process generally consists of several sub-processes (charge policy) in actual operation: constant current (CC) and constant voltage (CV). In practice, for fast charging as well as the battery life, multistage charge policy is usually used [31]. Fig. 1 shows an example. The charge current rate is firstly constant 3.6C (A rate of n C is the equal of a full charge or discharge in $1/n$ hours). After the charged capacity reaches QR1 (A ratio of the charged capacity of a battery to the nominal capacity given by the battery manufacturer), the charge current rate is switched to another constant 1C until battery terminal voltage reaches the cut-off voltage of 3.6 V. Afterward, battery is charged by a constant voltage of 3.6 V with a decreasing current and stops charging when the current is 0.02C. The battery terminal voltage clearly indicates the severe nonlinearity as illustrated in Fig. 1 (a). Different charge policies, which have different effects on battery life, might be employed for different batteries or the battery at different life states. Therefore, the battery life prediction needs to be suitable for different charge policies. While to the best of our knowledge it remains challenging, given their complex relationships with the typically nonlinear degradation process in a battery [32].

2.3. Battery life description based on charge process

It is known that in essence the battery cycle life or RUL is dominated by its internal properties, which would be reflected by its external electrical and thermal performance. Thus, the internal parameters of battery can be calculated based on the measurement of its external performance and parameters. By analogy, battery cycle life and RUL likewise have relationships with the measurable battery performance and parameters.

In this paper, the battery terminal voltage, current and temperature (V/I/T) curves during charge process are used to reflect the battery life. Some of the reasons for using this method can be expatiated as follows:

First, as stated in Section 2.2, the battery V, I and T can be measured easily and accurately during practical battery charge process because of the mild characteristic of this process. In particular, if the battery life is predicted based on the charge process, the influence of the random working loads of the devices on the prediction can be avoided. This therefore can make the prediction suitable for battery RUL prediction in practical devices' applications, also for the battery life early prediction.

Second, the battery would have diverse properties when its life differs or RUL varies. Accordingly, the internal property and external performance of batteries with different cycle lives or under different cycle numbers would be different, and then lead to different terminal voltages, currents, surface temperatures and curve shapes during charging. This is proved by Fig. 1 and Fig. 2.

Fig. 1 indicates the shapes of charge V/I/T curves of a battery under different cycle numbers. As illustrated in Fig. 1 (b) and (c), under the charge policy as shown in Fig. 1 (a): the voltage-curve characteristics, like value and duration of platform, voltage boost points (process that battery voltage reach to 3.6 V from the platform) and end of the curve are all diverse under different cycle numbers; and the temperature-curve characteristics, like rising rate, the apex, and decreasing rate are also various. They together reflect the battery cycle life, life degradation rate and RUL, so should be fused for battery life prediction.

Fig. 2 indicates the shapes of charge V/I/T curves of different batteries under different charge policies. We know that, under different charge polices, different batteries might have different lives, while different batteries might also have similar life. So here as an example, we choose two groups of LFP/graphite batteries to compare: each group has two batteries with similar life under different charge polices, but the battery lives of the two groups are different (one group with life of around 700 cycles and another with life of around 1050 cycles). The charge processes are indexed by charged capacity ratio, which is defined as the ratio of the capacity calculated by coulombic counting during charging to the nominal capacity. The different charge policies are illustrated in Fig. 2 (c). As shown in Fig. 2 (a), (c) and (d), the curves of battery V, I and T during charging and their relationship are complex and all diverse among batteries under different charge numbers. There is great difference between the curves of different batteries at the same cycle number, even if the battery life is similar. This further proves that the battery V, I and T during the whole charge process should be fused to reflect the difference for accurate and robust prediction of battery life.

Moreover, from Fig. 2 (b), which shows the difference of terminal voltages of four batteries between the 500th cycle and the 1st cycle, we notice that: because the battery cycle life and RUL vary with its internal properties, the difference between the curves of batteries with similar life reduces when compared with the curves in Fig. 2 (a), while the difference between batteries respectively with around 1050 and 700 cycle life becomes more evident. This capacity ratio-indexed description of such difference can mitigate the impact of different charging policies, so can help improve the suitability and accuracy of the battery life

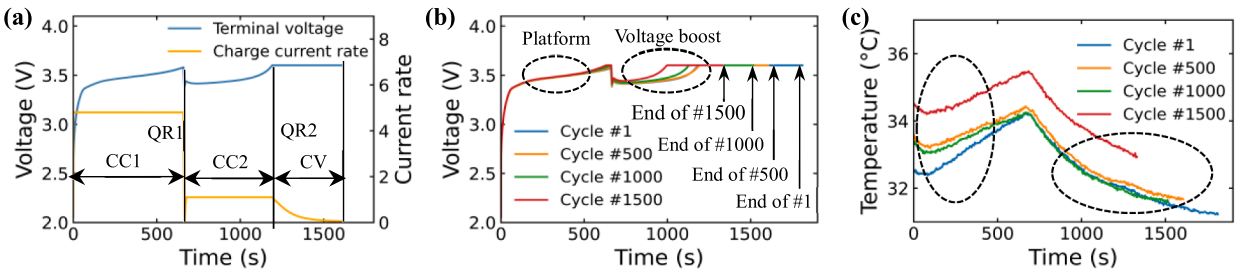

Fig. 1. Charging at different cycle numbers: (a) charge process; (b) terminal voltage curves; (c) battery chamber temperature curves.

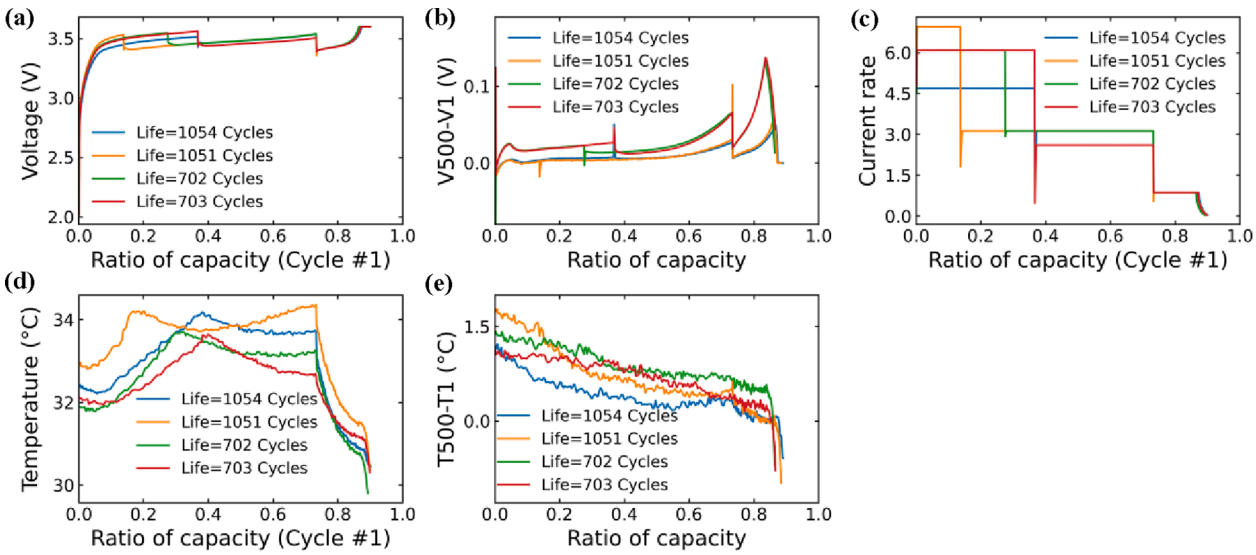

Fig. 2. Charging of different batteries with different charge policies at different cycle numbers: (a) terminal voltage curves of four batteries respectively with life = 1054, 1051, 702 and 703 at the 1st cycle; (b) curves of difference of terminal voltages of these four batteries between the 500th cycle and the 1st cycle, where the label of the ordinate, 'V500-V1', is the difference of the terminal voltage in the 500th cycle from the terminal voltage in the 1st cycle; (c) charge policies applied for these four batteries, respectively; (d) battery chamber temperature curves of these four batteries at the 1st cycle; (e) curves of difference of battery chamber temperatures of these four batteries between the 500th cycle and the 1st cycle., where the label of the ordinate, 'T500-T1′, is the difference of the temperature in the 500th cycle from the temperature in the 1st cycle. Here, the four batteries are the 10th, 42nd, 43rd and 36th cells in the "2017-05-12_batchdata_updataed_struct_errorcorrect.mat" data file provided by Ref. [3], respectively.

prediction for different charge policies.

In addition, due to the heat thermal effects from previous cycle, the surface temperatures of the battery at the beginning of charging might be different at different cycle numbers, although the environment is controlled as constant. So, the battery temperature at the same time under the larger cycle number may be higher or lower than that under the smaller cycle number as shown in Fig. 1 (c) and Fig. 2 (d). In fact, even the starting temperature of different cycle batteries is kept the same, the follow-up change curves of battery voltage and temperature will also be affected, due to the different internal parameters (such as the internal resistance) of battery at different aging states. Incorporating the charge temperature curve can help improve the battery life prediction performance.

Furthermore, the current curves of different charge policies are different. And even for the same charge policy, the current curve may also change during battery lifetime. For example, the charge process would automatically be switched to CV mode before the end of CC mode owing to degradation of battery capacity. So incorporating the charge current curve can make the prediction suitable for different charge policies and practical devices' actual applications become possible.

Thereupon in this paper, the battery cycle life and RUL are first described by the battery V/I/T curves during charge process from several cycles and the difference between these cycles in particular. As described above, the curves and relationship among different batteries under different cycle numbers and charge policies are complex. In order to predict the battery cycle life and RUL, a method for automatically learning the complex characteristics and modelling their relationship is required and will be introduced in Section 3.

3. Methodology

There is no explicit quantitative formula for relationship description between the lithium-ion battery life and curves of the battery V, I and T during the charge process. What features hidden in these curves which can accurately describe the lithium-ion battery life states is still not very clear yet. The relationship between battery V/I/T curves and battery cycle life (or RUL) is very difficult to perceive owing to the complex electrochemical reactions and mechanism inside the batteries during their lifetime. In this work, to predict the battery life and RUL which do not have a detailed expression, convolutional neural network (CNN) is applied to catch and model the features hidden in the curves comprehensively and automatically.

Because accurate prediction of battery cycle life and RUL needs to be based on proper input and algorithms, the input data generation is firstly presented in Section 3.1. Then, a HCNN structure and the algorithm are proposed in Section 3.2 to automatically extract and model the features hidden in the input for the predictions. A charging feature attention algorithm and multi-scale charging cycle attention algorithm are put forward in Section 3.3 and Section 3.4, respectively, for further heightening the prediction performance.

3.1. Input data generation

Based on the battery life description method presented in Section 2.3, here we take out the battery V/I/T curves during charge processes of the first m_i cycles which represent fresh state, and the other n_i cycles which reflect the cycled state, for battery life early prediction and RUL prediction. And as the charged capacity during charge process increases monotonously even under different charge policies, we consider the battery V, I and T as the functions of capacity ratio (so hereinafter, referred as the capacity-indexed battery terminal voltage, current and temperature), not as the functions of time usually used, to maintain a uniform description basis when applying different charge policies. Here, the nominal capacity (C_n) which the battery manufacturers provide is used considering there might be estimation error of real battery capacity in practice.

In this way, the evolution laws of battery life states along cycles and all the features hidden in these curves can be input into the battery life prediction for accuracy and robustness. And the battery life will be predicted through inputting the charge data of $m_i + n_i$ cycles. This means that: using the first $m_i + n_i$ cycles, the battery life can be early predicted before its capacity degradation; using the first m_i cycles and the latest n_i cycles, the battery RUL can be predicted in practice without being influenced by the random operation loads and environment of the device. The structure of the matrices of battery V, I and T is shown in Fig. 3 (a). The V, I and T curves during charge processes of the $m_i + n_i$ cycles of

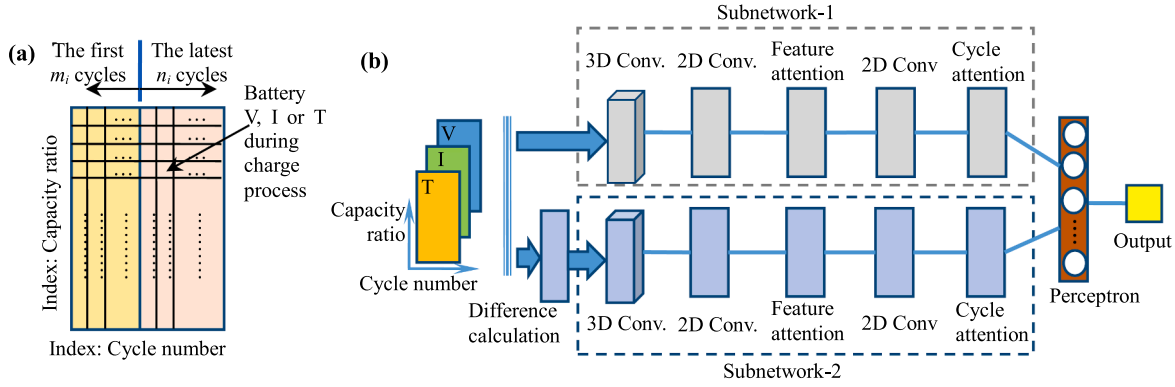

Fig. 3. (a) Structure of input matrices of battery V, I and T; (b) Structure of hybrid convolutional neural network.

a battery are treated as three matrices respectively, like those representing images. With the purpose of prediction and equilibrating computational complexity and accuracy, we select m_i as 5 and n_i as 15. Furthermore, a non-sampling interval is defined since batteries might be not fully discharged when charging started in the practical use. In this interval, the battery V/I/T during charging will not be sampled into these charge curves as the input. And on the basis of massive actual data analysis, the non-sampling interval is about from 0% to 20% battery capacity ratio in this paper. In addition, to make the prediction applicable to different capacity batteries, the battery I is expressed as the current rate (I/C_n).

Note that there are often some small errors in the original data, such as outliers, accidental reduction in the monotonic increase process. Data cleaning should be performed first. Before using the battery V, I and T data for the prediction, abnormal values are replaced by the average of highly correlated data. Then, Savitzky-Golay filter is used to reduce measurement noises of corresponding voltages, currents and temperatures [33]. In addition, it is known that battery V, I and T have different value ranges, which will have a negative impact on the training of the model. In order to eliminate this effect, after data cleaning, we proceed with data normalization to balance the data distribution. In this work, the Z-score normalization method [34] is used. The values of V, I and T in Fig. 3 (a) are the ones after Z-score normalization. When presenting the final predicted cycle life results, denormalization is performed.

3.2. Proposed hybrid convolutional neural network (HCNN)

CNN is widely used in image processing in machine learning because of its ability for automatic feature extraction and nonlinear simulation [35]. But unlike the R/G/B (Red, Green, and Blue) in a color image, the battery V, I and T have strong relationship. So, both they and their curves of different charge cycles need to be fused for battery life prediction. The three matrices generated from battery V/I /T curves in Section 3.1 should not be treated like an R/G/B color image as the input of a two-dimensional CNN (2DCNN). They need to be input respectively. But 2DCNN cannot capture the information on the input sequence very well. Three-dimensional CNN (3DCNN) can overcome this issue, but neural network with too many 3DCNN layers usually has too many parameters leading to the difficulty of training or applying the model. Therefore, in this paper, we propose a hybrid CNN (HCNN), which combines the advantages of 3DCNN and 2DCNN, to predict the battery N_{EOL} or RUL.

3.2.1. HCNN structure

Fig. 3 (b) indicates the framework of the proposed battery life prediction and the structure of the proposed HCNN. There are two subnetworks in parallel at first. Besides taking the three matrices generated from the battery V/I/T curves during charge processes of $m_i + n_i$ cycles as the input of subnetwork-1, to mitigate the impact of different charging policies, the difference between each curve and the first-cycle curve of the battery V/I/T is calculated respectively. They are input to subnetwork-2. Each subnetwork is constructed as follows. Because the evolution of battery life states along cycles has been expressed by input as stated in Section 3.1, only one 3DCNN layer is used as the input layer to fuse the curves considering the strong relationships between battery V, I and T. It is combined with two 2DCNN layers to extract and model the features hidden in the input curves automatically. The result of battery life prediction is finally output from the outputs of the two subnetworks through a perceptron. However, different features and different cycles would contribute to the battery cycle life or RUL prediction differently. But their relationships haven't been described yet. So, we propose charging feature attention and cycle attention algorithms and embed them among the layers as depicted in Fig. 3 (b) respectively, to perform the charging feature relevance estimation and cycle relevance evaluation for the prediction. It is expected that battery cycle life or RUL could be predicted when the parameters of the HCNN and attention models are determined from the training process.

To diminish the difficulty of neural network training and improve the accuracy of prediction result, the training data of HCNN should reflect the difference in the V/I/T curves among batteries with different charge policies under different cycle numbers. Moreover, the activation functions in the convolutional layers are all Leaky Rectified Linear Unit (Leaky ReLU) function with $\alpha = 0.05$, written as:

$$f(z) = \begin{cases} \alpha z, & z \leqslant 0 \\ z, & z > 0 \end{cases} \tag{2}$$

3.2.2. HCNN algorithm

After preparing the data as described in Section 3.1, the input matrices of battery voltages, currents and temperatures will have the same structure illustrated in Fig. 3 (a) and same size. Here, represent the size as $H^1 \times C^1$, where H^1 is the number of the indexed capacity ratios, C^1 is the number of the cycles as $m_i + n_i$.

As described above, for the two subnetworks in the proposed HCNN, although their inputs and so their functions are different, their structures are the same. Below in this section, we first present the algorithm of the proposed HCNN in detail with the example of subnetwork-1.

Firstly, as analyzed above, we need to fuse the battery V, I and T for considering their strong relationships. Thus, the matrices are input to 3DCNN layer as a 3-D matrix. Thanks to the battery V, I and T during charge processes under $m_i + n_i$ cycle numbers are arranged in three matrices respectively, the depth of the input data volume is only three. So, the computational burden of the 3DCNN layer can be significantly reduced. And the size of the input matrix will be $3 \times H^1 \times C^1$. For one measurement ($X^{(i)}$, $Y^{(i)}$), the 3-D matrix $X^{(i)}$ is the input, and $Y^{(i)}$ is the value of battery life (RUL or N_{EOL} for fresh battery). We define $X^{(i)}(m)$ is the m^{th} submatrix of $X^{(i)}$ with a size $3 \times h^1 \times c^1$, the 3D convolutional

result of $X^{(i)}(m)$ is:

$$u_m^{1,r_1} = W^{1,r_1} \odot X^{(i)}(m) + B^{1,r_1}, r_1 = 1, 2, \cdots, q^1 \quad (3)$$

where q^1 is the amount of convolutional kernels in the 3D convolutional layer, value B^{1,r_1} and matrix W^{1,r_1} are the bias and weights in the r_1^{th} 3-D kernel, respectively, and u_m^{1,r_1} is the convolutional result between $X^{(i)}(m)$ and the r_1^{th} kernel. Besides, "$\odot$" is dot product. Then the output will go through a Leaky ReLU function. After processing by this layer, the output feature volumes produced by this 3D convolutional layer can be obtained. The element of the r_1^{th} output feature volume (which we represent as p^{1,r_1}) can be described as:

$$p_m^{1,r_1} = f(u_m^{1,r_1}), r_1 = 1, 2, \cdots, q^1 \quad (4)$$

Note that, the kernel size is the same as the size of $X^{(i)}(m)$. And in order to avoid the loss of possible features hidden in the input battery V, I and T curves, the stride in convolutional operation is set as 1. So the size of p^{1,r_1} will be $1 \times v_H^1 \times v_C^1$, where $v_{\text{H}}^1 = H^1 - h^1 + 1$ and $v_{\text{C}}^1 = C^1 - c^1 + 1$. Therefore, by squeezing the first dimension of p^{1,r_1}, p^{1,r_1} can be easily transferred to 2-D feature map (here, represented as p_s^{1,r_1}) with a size $v_H^1 \times v_C^1$, which could be then fed into the first 2D convolutional layer.

Subsequently, to extract and model the charging features from the fused results through 3DCNN layer automatically for battery life prediction, the first 2D convolutional layer is used. Taking the $X^{(2)}(n, r_1)$ as the n^{th} submatrix of p_s^{1,r_1} with a size $h^2 \times c^2$, the 2D convolutional result of $X^{(2)}(n, r_1)$ is:

$$u_n^{2,r_2} = \sum\nolimits_{r_1=1}^{q^1} W^{2,r_2} \odot X^{(2)}(n, r_1) + B^{2,r_2}, r_2 = 1, 2, \cdots, q^2 \quad (5)$$

where q^2 is the amount of convolutional kernels in the first 2D convolutional layer, matrix W^{2,r_2} and value B^{2,r_2} are the weights and bias in the r_2^{th}2-D kernel, respectively, and u_n^{2,r_2} is the convolutional result between $X^{(2)}(n, r_1)$ ($r_1 = 1, 2, \cdots, q^1$) and the r_2^{th} kernel. Then the output will also go through a Leaky ReLU function as Eq. (2). After processing by this layer, we can get the output feature maps produced by the first 2D convolutional layer. And the element of the r_2^{th} output feature map (which we represent as p^{2,r_2}) can be described as:

$$p_n^{2,r_2} = f(u_n^{2,r_2}), r_2 = 1, 2, \cdots, q^2 \quad (6)$$

Note that, the kernel size is the same as the size of $X^{(2)}(n, r_1)$, and the stride in the 2D convolutional operation is also set as 1. So the size of p^{2,r_2} is $v_H^2 \times v_C^2$, where $v_{\text{H}}^2 = v_{\text{H}}^1 - h^2 + 1$ and $v_{\text{C}}^2 = v_{\text{C}}^1 - c^2 + 1$.

Afterwards, to estimate the relevance and salience of the features extracted by the first 2DCNN layer for battery life prediction, the proposed charging feature attention (CFA, as will be presented in Section 3.3) is used. And the second 2D convolutional layer is embedded after it as a transition layer to the followed multi-scale charging cycle attention (MsCCA, as will be explained in Section 3.4).

Similarly, the output feature maps produced by the second 2D convolutional layer can be calculated, which we represent as p^{3,r_3} ($r_3 = 1, 2, \cdots, q^3$). Here, q^3 is the number of convolutional kernels in this 2D convolutional layer. And the size of p^{3,r_3} is $v_H^3 \times v_C^3$, where $v_{\text{H}}^3 = v_{\text{H}}^{FA} - h^3 + 1$, $v_{\text{C}}^3 = v_{\text{C}}^{FA} - c^3 + 1$, $h^3 \times c^3$ is the size of the kernel of this layer, $v_{\text{H}}^{FA} \times v_{\text{C}}^{FA}$ is the size of the output map after feature attention (see Section 3.3). In the CNN process, the max pooling layer [36] is used because it can reduce the deviation of estimated mean caused by convolution layer parameter error, reducing the spatial size of the feature map for efficient calculation and providing robust learning results for input data.

Then, after the proposed MsCCA evaluates the relevance and salience of the multi-scale charging cycle features hidden in the output results of the second 2D convolutional layer, the output of the subnetwork-1 can be obtained (see Section 3.4).

The algorithms of the subnetwork-2 in the proposed HCNN are similar to those of the subnetwork-1 as presented above. So similarly, after calculating the difference between each curve and the first-cycle curve of the battery V/I/T, through subnetwork-2, its output of MsCCA can also be obtained. By concatenating the A^C outputs of these two subnetworks, the model records the prediction value A^{Cs} with a size of $1 \times l_{CAs}$.

Finally, by using a perceptron with two fully-connected layers, the final predicted battery life is determined:

$$N_{EOL} \text{ or } RUL = L\left(\sum_{j=1}^{d^o}\left(\beta_j \cdot f\left(\sum_{i=1}^{l_{CAs}} \alpha_{i,j} \cdot A_i^{Cs} + b_j\right)\right) + c\right) \quad (7)$$

where α and β are the weights, b and c are the bias values, d^o is the number of the transformation neurons, $L(\cdot)$ is a linear activation function.

3.3. Charging feature attention (CFA)

Different features during charging reflect different influence of battery life states. In order to dynamically estimate the salience and relevance of different features in charge process for the battery life prediction, we define and apply a charging feature attention algorithm to learn them. It means: let the model itself learn to pay attention to specific regions in the feature maps from an input, according to the correlation with the battery life prediction task, and incorporate attention for charging feature salience detection for different parts in an input may exhibit varying degrees of salience to the prediction. It is implemented as a feature map filter embedded after the first 2DCNN layer of both the two subnetworks in the proposed HCNN (see in Fig. 3 (b)) to preserve the information based on the calculated attention weight.

3.3.1. CFA algorithm

For the sake of generality and convenience, here represent the output volume of the previous convolutional layer as F, which has a size of $M \times N \times K$, with K feature maps of $M \times N$. For example, as presented in Section 3.2, the output of the first 2DCNN layer in the subnetwork-1 is p^{2,r_2} ($r_2 = 1, 2, \cdots, q^2$) with a size of $v_H^2 \times v_C^2$. Accordingly, for subnetwork-1, $K = q^2$, $M \times N = v_H^2 \times v_C^2$.

To fuse the different charging features in the different parts of the feature maps and score their relevance and salience for battery life prediction, the feature attention denoted by a matrix A is embedded after the convolutional layer. Here, the element A_{ij} of A is the attention weight for the feature vector F_{ij}, which is composed of the elements with indexes (i, j) in all the K feature maps. The matrix A has the same size of $M \times N$ as that of the feature map, and each feature vector corresponds to a certain part of the input curves (i.e., receptive field) of battery V, I and T. Therefore, when the attention model is inserted in latter layers, the receptive field of the attention can become larger and the features in the receptive field can be weighted for the prediction.

Considering the second 2-D convolutional layer is used as a transition layer as described in Section 3.2.2, we insert the feature attention after the first 2-D convolutional layer of each of the two subnetworks as shown in Fig. 3 (b). The feature vector F_{ij} is firstly expressed as:

$$F_{ij} = [F_{ij}^1, F_{ij}^2, \cdots\cdots, F_{ij}^K] \quad (8)$$

where F_{ij}^m is the element in the m^{th} feature map F^m in F, $m = 1, 2, \cdots, K$.

Then, to learn the attention weight matrix A, two feature attention models are employed: the first model is parameterized by individual weights for each entry of feature map, while the second model shares the transformation weights across the whole feature map. Here, we use two fully-connected layers to model the attention weight A. Thus, the mathematical equation of the attention weight A_{ij} for F_{ij} can be expressed as:

$$A_{ij} = f\left(\sum_{u=1}^{n_{FA}} \delta_{ij}^{u} g\left(\sum_{m=1}^{K} \omega_{ij}^{u,m} F_{ij}^{m} + b_{ij}^{u}\right) + c_{ij}\right) \quad (9)$$

where $\omega_{ij}^{u,m}$ is the transformation weight value connecting the m^{th} element F_{ij}^{m} in feature vector F_{ij} and the u^{th} transformation neuron, δ_{ij} is the weight value connecting the u^{th} transformation neuron and A_{ij} to fuse the information from different feature maps, b_{ij}^{u} is the threshold value of the u^{th} transformation neuron, and c_{ij} is the threshold value of neuron A_{ij}. Here, n_{FA} is the number of the transformation neurons, and hyperbolic tangent function $g(\cdot)$ and sigmoid function $f(\cdot)$ are used as the activation functions in the first and second feature attention models, as written in Eq. (10) and Eq. (11), respectively.

$$g(x) = \frac{e^{x} - e^{-x}}{e^{x} + e^{-x}} \quad (10)$$

$$f(x) = \frac{1}{1 + e^{-x}} \quad (11)$$

The sigmoid function constraints the attention weight to lie in the range [0,1]. And the obtained attention weight matrix A controls the information flowing into the subsequent layer by an element-wise multiplication to each feature map of F. The output of this feature attention, which will be fed into the subsequent layer, can be calculated as follows:

$$G^{m} = A. \times F^{m} \quad (12)$$

where G^{m} is the m^{th} of the output maps after feature attention, ". × " is element-wise multiplication.

3.3.2. CFA principle for battery life prediction improvement

The principle of the proposed feature attention for battery life prediction improvement can be explained as follows.

As mentioned earlier in this paper, battery cycle life or RUL is dominated by its internal properties in essence. Obviously, the internal electrochemical variation of the battery has effects on its external electrical and thermal performance. When battery cycle life differs or RUL varies, the battery would have diverse properties which can impact the battery charge/discharge performance in different degrees. These different internal properties and external performances of batteries with different lives or RULs would lead to overall and/or local different changes of their V/I/T curves during charging. They would contribute more or less to the battery life prediction accuracy. Therefore, capturing the local variations as well as the overall variations of the curves and evaluating them might be helpful for battery life prediction improvement.

In this attention algorithm, individual weights (ω_{ij}) are used for the first model and shared weights (δ_{ij}) for the second fusion model in this feature attention. The first model is expected to capture the local detailed variation while the second fusion model is expected to capture the global variation and smooth the attention distribution. And the spatial variation of the input, which are composed of the curves of battery voltage, current and temperature in different charge cycles, will not be blurred. Therefore, not only can the charging features be clearly distinguished from the input curves, but also the salient regions in the curves can be captured. And different charging features hidden in different parts of the input curves are automatically weighted by this feature attention, for helping improve the battery life prediction performance (accuracy and robustness).

3.4. Multi-scale charging cycle attention (MsCCA)

The difference between charge cycles reflects the evolution laws of battery life states. In order to dynamically estimate the salience and relevance of different cycles for battery life prediction, we define a charging cycle attention algorithm to learn them. It means: let the model itself learn to pay attention to different segments in a sequence of the charging cycles from an input according to the relevance in the battery life prediction task, and detect the charging cyclic salience in the sequential expression to determine the amount of information in each charging cycle to be incorporated into the prediction. This can be taken as a sequence-to-sequence learning problem. Usually, as the improved recurrent neural network (RNN), the LSTM neural network is used for such problem. But there is the issue of many parameters and long training time in LSTM. Therefore, other methods should be found to obtain the features between charge cycles. Inspired by the principle of dilated convolution network (DCN), we first introduce the DCN into our charging cycle attention.

3.4.1. MsCCA principle for battery life prediction improvement

Fig. 4 (a) indicates the basic principle of 1-D DCN as an example. From the process of the DCN, it can be seen that: the range of input neurons involved in DCN increases exponentially with the increase of network layers. In the HCNN, each column of the output feature map of the second 2D convolutional layer corresponds to the curves of V, I and T during charge process under different certain cycle numbers. Therefore, if introducing the 2-D DCN, the potential advantages can be that: (1) owing to the computational simplicity of convolution, the computational efficiency would be much higher than that by LSTM; (2) in particular, with the increase of network layers, through the dilation convolution at the width direction of the feature map, battery charge cycles in different ranges can be fused by different layers, and so the underlying information between different range of cycles can be captured; (3) furthermore, through the dilated convolution at the height direction of the feature map, the underlying information between different multiple periods in a cycle can be captured at the same time. In other words, not only the multi-scale charging features between different charge cycles (adjacent cycles or cycles at different intervals) of battery, but also the multi-scale charging features between different periods in a charge cycle of battery, can be obtained and evaluated automatically and efficiently by employing 2-D DCN for battery life prediction. Accordingly, we refer to the cycle attention defined in this section as MsCCA.

Fig. 4 (b) shows the principle of the defined MsCCA. The 2-D DCN is first used to fuse and extract the multi-scale charging features from different cycles of battery. Their relevance and salience information of each output from the DCN layers are then captured by a two-layer perceptron. An attention weight as the salience score for each output of the DCN layers is finally produced for filtering the output information flow from the DCN layers. To enlarge the receptive field of the cycle attention, the cycle attention is added after the second 2DCNN layer of both two subnetworks in the HCNN as shown in Fig. 3 (b).

As stated in Section 3.3.2, when battery cycle life differs or RUL varies, the effects of the diverse properties in battery could cause the changes of the external performance curves of the battery in different degrees. In addition, the changes would also change with the different life states of the battery or under different cycle numbers. In other words, the caused changes could be between different cycles as well as between different periods in a cycle. They would contribute more or less to the prediction accuracy of battery cycle life or RUL. Therefore, capturing and scoring these changes by realizing this multi-scale charging cycle attention may be useful for improving the battery life prediction.

3.4.2. MsCCA algorithm

As can be seen in Fig. 4 (b), in the 2-D DCN, the input of the first layer is the output feature maps of the previous layer, and the input of other layers is the output feature maps of its previous dilated convolutional layer. Also for the sake of generality and convenience, here suppose that there are L dilated convolutional layers in the 2-D DCN, and the n^{th}

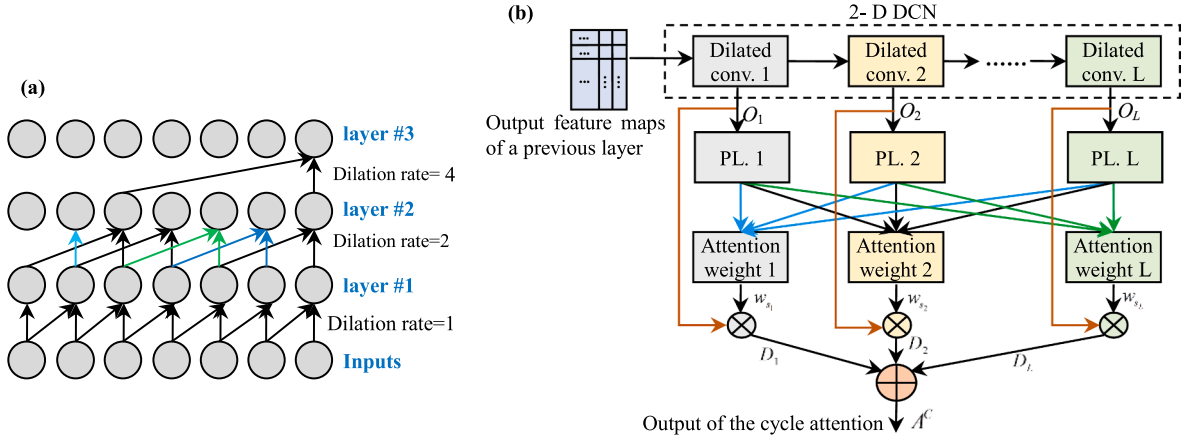

Fig. 4. Schematics of dilated convolution and proposed cycle attention: (a) 1-D dilated convolution; (b) multi-scale cycle attention, where "PL." is short for the perceptron layer.

dilated convolutional layer has the output $O_n(n = 1, 2\cdots, L)$, the input feature maps with a size of $M_n \times N_n$, the weights W_n and bias values B_n in the kernels with a size of $P_n \times Q_n$ and the dilation rate as $(d_{h,n}, d_{w,n})$. The output of this layer can be expressed as:

$$o_n^k(i,j) = f\Big(\sum_{r=1}^{R}\sum_{p=0}^{P_n-1}\sum_{q=0}^{Q_n-1}(x_{i+a,j+b}^r \cdot w_{p,q}^k) + b_k\Big), \quad \begin{cases} a = (p+1)d_{h,n} + 1 \\ b = (q+1)d_{w,n} + 1 \end{cases} \tag{13}$$

where $o_n^k(i,j)$ is the element at (i,j) in the k^{th} of output feature maps O_n of the n^{th} layer, $x_{i+a,j+b}^r$ is the element at $(i+a, j+b)$ in the r^{th} of R input feature maps from the $(n-1)^{\text{th}}$ layer, and $w_{p,q}^k$ is the element at (p,q) in the W_n^k. Therein, W_n^k and b_k are the weight matrix and bias value, respectively, in the k^{th} kernel of the n^{th} layer. And the size $(R_{h,n+1} \times R_{w,n+1})$ of the receptive field of the output of this layer can be calculated as:

$$R_{h,n+1} = R_{h,1} + \sum_{l=1}^{n} d_{h,l}(P_l - 1) \tag{14}$$

$$R_{w,n+1} = R_{w,1} + \sum_{l=1}^{n} d_{w,l}(Q_l - 1) \tag{15}$$

where $R_{h,1} \times R_{w,1} = \left(h^1 + h^2 + h^3 - 2\right) \times \left(c^1 + c^2 + c^3 - 2\right)$ is the size of the receptive field of the output of the second 2DCNN layer in the HCNN.

Based on the principles of 3DCNN, 2DCNN and the CFA proposed above, we know that $R_{h,1}$ and $R_{w,1}$ correspond, respectively, to a certain range of battery capacity ratios and to a certain range of battery cycle numbers in the input 3-D matrix as described in Section 3.1. Thus, the battery V, I and T curves in range of $R_{w,n+1}$ cycles in the input 3-D matrix can be fused by Eq. (13), and their features in the range of $R_{h,n+1}$ battery capacity ratios can be further extracted by Eq. (13) at the same time. For different layers $(n = 1, 2\cdots, L)$ in the 2-D DCN, the different ranges of battery cycles and capacity ratios are corresponded to, referring to Eq. (14) and Eq. (15). Thereby, the multi-scale features along the battery charging cycles and capacity ratios can be fused and further extracted through the 2-D DCN.

Then, for the output of each of the dilated convolutional layers, to estimate its relevance and salience for battery life prediction result, the cycle attention score s_n is modelled by a two-layer perceptron:

$$\begin{cases} s_n = f\left(\sum_{j=1}^{n_{CA}} v_n^j g\left(\sum_{i=1}^{I_n} z_n^{i,j} O_n^i + b_n^j\right) + c_n\right) \\ \text{where } n = 1, 2, \cdots, L \end{cases} \tag{16}$$

where I_n is the number of the elements after global average pooling of the feature maps in O_n, z_n and b_n are the weight and bias for the first perceptron layer, v_n and c_n are the weight and bias for the second perceptron layer. Here n_{CA} is a hyper-parameter that is the dimension of transformed mid-representation. Again, the hyperbolic tangent function $g(\cdot)$ and sigmoid function $f(\cdot)$ are used to model the attention score. So the score values are constrained between 0 and 1. The relevance of the output of each of the dilated convolutional layers to the battery life prediction is therefore measured by this perceptron.

Next, the attention score is normalized over the whole outputs of the dilated convolutional layers to get the final cycle attention weight w_{s_n} as salience score. And the obtained cycle attention weights are used to control how much information for each output of the dilated convolutional layers is taken into account as the attended output of this layer to perform the battery life estimation, as written in Eq. (17).

$$D_n = w_{s_n} \cdot O_n = \frac{s_n}{\sum_{l=1}^{L} s_l} \cdot O_n, \quad n = 1, 2, \cdots, L \tag{17}$$

Finally, by concatenating D_n $(n = 1,2\cdots,L)$, we get the final output of cycle attention as A^C.

It is worth noting that, in practice, we have selected the number of cycles in an input 3-D matrix as 20 (see Section 3.1). Through the previous layers in the HCNN, the corresponding dimension (width) of the input feature map of this cycle attention might reduce to about 10. Accordingly, a 2-D DCN with $L = 3$ dilated convolutional layers is employed. And the dilation rates at width dimension are selected as 1, 2 and 5 for the three layers, respectively. Thereby, three scales of charge cycles and the features (between cycles and between periods in a charge cycle) are fused and further extracted by the 2-D DCN simultaneously. And their relevance and salience for battery life prediction result can be then captured and weighted by Eq. (16) and Eq. (17) as A^C for improving the prediction performance.

Battery terminal voltages, currents and temperatures during charge process of each of charge/discharge cycles of batteries under different charge policies in the experimental situation are measured and treated for network training. The values of the weight and bias in the proposed HCNN with the two attentions are determined after the training process. Then, through the proposed model with these values, the charging features hidden in the input curves of battery V, I and T during charging will be automatically extracted and evaluated perfectly. So as will be illustrated later in this paper, the battery cycle life can be accurately predicted early before battery capacity degradation, and the RUL of battery during its whole cycle life can be also estimated accurately as well.

In this work, a mean-absolute-percentage-error (MAPE) is used for

optimization in network training. According to the framework above, we use the Adam optimizer for parameter training in Backpropagation to update weights. The test set will be substituted into the trained model to predict the cycle life of lithium-ion battery once the training work is done using training set. The entire training process of the proposed HCNN with charging feature attention and multi-scale charging cycle attention is summarized in Algorithm 1.

Algorithm 1. (*Outline of model training for early cycle life and RUL prediction of lithium-ion batteries*)

Input: The training set
Output: Trained HCNN weights
Repeat
Forward Propagation:
Do
Step1: Preform three-dimensional convolutional operation with the V/I/T data of different batteries at different cycles by Eqs. (3) and (4).
Step2: Conduct two-dimensional convolutional operation with the output of the 3DCNN layer by Eqs. (5) and (6).
Step3: Use the feature attention in Eqs. (8)–(12) to estimate the salience and relevance of different features in charge process for the battery life prediction.
Step4: Conduct two-dimensional convolutional operation with the output of the previous layer, similarly with Step 2.
Step5: Use the cycle attention in Eqs. (13), (16) and (17) to fuse the salience and relevance of different cycles for battery life prediction.
Step6: Use fully-connected layer to determine the output as Eq. (7).**Step7:** Calculate the MAPE between the predictions and target lives of batteries.
End
Backward Propagation:
Update network parameters.
until A predefined small loss.

4. Experimental verification and comparative analysis

4.1. Data set of lithium-ion batteries

In this paper, a dataset of LFP/graphite lithium batteries (A123 Systems, model APR18650M1A, 1.1 Ah nominal capacity and 3.3 V nominal voltage) is used to verify the effectiveness of the proposed battery life prediction method. The dataset is collected in a temperature chamber set to 30 °C, by cycling 124 commercial LFP/graphite batteries on a 48-channel Arbin LBT potentiostat. The batteries are cycled under different charging conditions but identical discharging conditions (4C to 2.0 V) to 80% of nominal capacity, obtaining their measured cycle lives.

In this dataset, there are 72 charge policies and approximately 96,700 cycles. Charge processes in various battery states, with average current rates ranging from 3.6C to 6C are contained, so that this dataset can be used to test whether the proposed method could accurately predict the cycle life and RUL of batteries under different charging policies for a wide sphere of application.

In addition, as shown in Fig. 5, the battery lives in this dataset have a wide range from approximately 150 cycles to 2300 cycles. Therefore, at the same time, this dataset can be used to measure whether the proposed method could perform well in different application scenarios with greatly various battery lives. The dataset is published by Severson et al. [3]. As the authors pointed out: this dataset is the largest publicly available for commercial lithium-ion batteries cycled under controlled conditions. Tests on different batteries will be conducted to systematically verify the prediction performance of battery cycle life and RUL of the proposed method.

4.2. Prediction verification

In this section, the proposed HCNN with the two attentions is exploited for life prediction of lithium-ion batteries. For both battery cycle life early prediction and RUL prediction, parameters including the

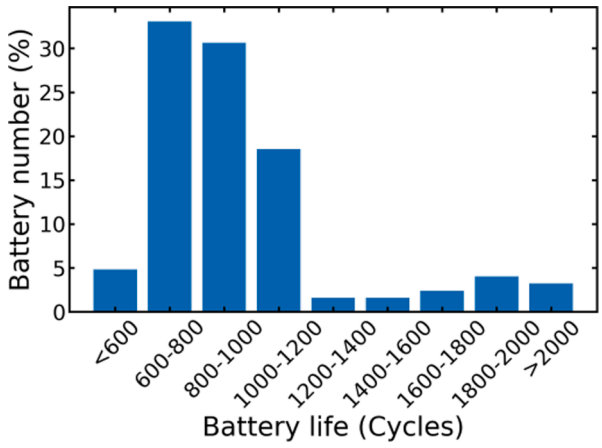

Fig. 5. Cycle lives' distribution of the lithium-ion batteries in the data set.

m_i and n_i mentioned in Section 3.1, kernel size, number of filters and hyper parameters of the attention algorithms, are all unified to guarantee the realizability and suitability, which means the trained model can be used for different prediction applications. And if enough offline data sets are provided, it will be an easy work to adjust the parameters to obtain possibly better prediction performance. The tuning parameters are listed in Table 1.

The dataset mentioned in Section 4.1 is randomly divided into two sets. According to the training process outlined in Algorithm 1, the proposed HCNN with the two attentions is firstly trained by the experiment charge data of 94 batteries in the training dataset, which includes 67 different charge policies. Then, the charge data of 30 other batteries in the test dataset, which includes 22 different charge policies, are applied to test the proposed method.

To evaluate the predictive performance, three metrics are used: RMSE with units of cycles as the steady metric; MAPE as the accuracy metric adopted as the relative error of the predicted life; and the mean absolute error (MAE) as error metric to describe the absolute error between the actual life and the predicted life.

4.2.1. Battery cycle life early prediction verification

To verify the performance of the proposed method for battery cycle life early prediction, after training the model as described above, tests on the different batteries with different lives and charge policies in the test dataset are conducted in this section.

According to the input data generation stated in Section 3.1, to verify the early prediction of battery cycle life, for a battery, the input consists of the V/I/T curves of the first 5 cycles and the other 15 early cycles from the $(n_e - 14)^{\text{th}}$ cycle to n_e^{th} cycle. Here, n_e is an integral number, being the maximum serial number of the used early cycles for battery life early prediction.

The number of the used early cycles might have influence on the prediction error. In order to identify the reasonable number of the used early cycles for the battery life early prediction, different n_e will be investigated. The range of n_e is set from 20 to 100 considering the time needed for conducting battery cycle tests. And each n_e would lead to a cycle life prediction result for a battery. Accordingly, comparison of the predictions with these different n_e is presented. The prediction results

Table 1
The training parameters in each layer of the proposed HCNN.

m_i	n_i	kernel size: height × width	filters in 3DCNN layer	filters in 2DCNN layer	hyper parameter in the feature attention, n_{FA}	hyper parameter in the cycle attention, n_{CA}
5	15	8 × 2	100	150	50	64

and errors of the 30 different batteries in the test dataset are shown in the Fig. 6 (a) and (b) respectively, where the abscissa values represent the measured final cycle lives of the 30 batteries. Note that each battery has a single final cycle life. And in all plots, the colors are determined based on the n_e, as shown by the color bar. Moreover, the numerical results are shown in Table 2.

As shown in Fig. 6, the cycle lives of the test batteries under different charge policies vary significantly from about 450–2000 cycles. While because the features which dominate the battery cycle life might have been successfully extracted from the input charge V/I/T data and perfectly modelled by the proposed method, the predicted results of all the batteries are almost identical with their real lives measured, respectively. So that in the figure, some predicted results are covered by others. The maximum absolute error of early prediction can be limited to a band of 6% for all the batteries. Moreover, as shown in Fig. 6 (b), on the whole, the blue dots which denote the errors using smaller n_e are further away from zero than the red dots which denote the errors using bigger n_e. This means that the prediction with more number of early cycles would be able to have better prediction performance on the whole. From Table 2, it can be seen that:

(1) only using the charge data of the first 40–100 cycles, the MAPE, RMSE and MAE of early prediction of battery cycle life, which are less than 1.3%, 19 cycles and 14 cycles respectively, can be achieved;
(2) in particular, even only using the charge data of the first cycles whose number is as little as 20, the proposed method can also make the early prediction of battery cycle life very accurate, with MAPE, RMSE and MAE less than 3.1%, 42 cycles and 33 cycles, respectively;
(3) while when n_e is 100, the best prediction performance is achieved. However, because the battery life characteristics and their evolution might have been captured and modelled perfectly enough when n_e increases to 60, the MAPE and MAE achieved when n_e is 60 are very close to those best ones while the test work can be reduced by 40%. An error of 1.1% can be achieved for all the test batteries under different charge policies. Therefore, to balance the demand of prediction accuracy and amount of test work, we would like to recommend using the first 60 cycles (i.e., n_e= 60, exhibiting a median increase of less than 0.1% from initial capacity) for the battery life early prediction.

These results indicate that the proposed method can achieve high accuracy which is reflected by the low MAPE and MAE, and strong stability which is illustrated through the RMSE, for battery cycle life early prediction before battery capacity degradation under different charge policies.

Table 2
The numerical results of the RUL prediction and the early life prediction with certain n_e.

	Early prediction of battery cycle life					RUL prediction	
						For all test batteries	For a single battery
n_e	20	40	60	80	100	/	/
MAPE (%)	3.08	1.30	1.12	1.21	1.12	3.55	1.35
RMSE (Cycles)	42	19	13	13	11	11	4
MAE (Cycles)	33	13	11	10	9	9	3

This can significantly promote the battery production and optimization. For example, to evaluate the cycle life of a battery which might be with a life span of 2000 cycles, it is well known that traditional method (which is through experimental cycling to the end of battery life [4]) often needs to test continuously for more than 6 months. While the proposed method only needs 1–3% of the test time for conducting the test of the first 20–60 cycles. In other words, we don't need to carry out hundreds or thousands of cycles of testing as usual. More than 97% test time and cost might be saved, manufacturers therefore can greatly accelerate the battery development cycle and perform rapid verification of new manufacturing.

4.2.2. Battery RUL prediction verification

After the battery cycle life early prediction verification presented above, this Section is aimed at verifying the performance of the proposed method for battery RUL prediction, using the same trained model as that in battery life early prediction by conducting the tests on the different batteries in the test dataset.

Similar to the battery life early prediction, the input consists of the V/I/T curves of the first 5 cycles and the other 15 early cycles from the $(n_e - 14)^{\text{th}}$ cycle to n_e^{th} cycle. But differently here, n_e ranges from 20 to the battery cycle life (in the unit of cycles), being the serial number of the current charge cycle.

Accordingly, the measured RUL of a battery is the difference of the battery cycle life and n_e. And each n_e would lead to a RUL prediction result for a battery. So, as the value of n_e changes, the RULs during the whole useful lifetime of a battery can be predicted. Fig. 7 (a) shows the RUL prediction results of the 30 test batteries in the test dataset, and Fig. 7 (b) illustrates the prediction errors. Here, the serial number of each battery is called the battery ID. And in all plots, the colors are determined based on the battery ID of the 30 batteries, as shown by the color bar. Moreover, the numerical results are also shown in Table 2.

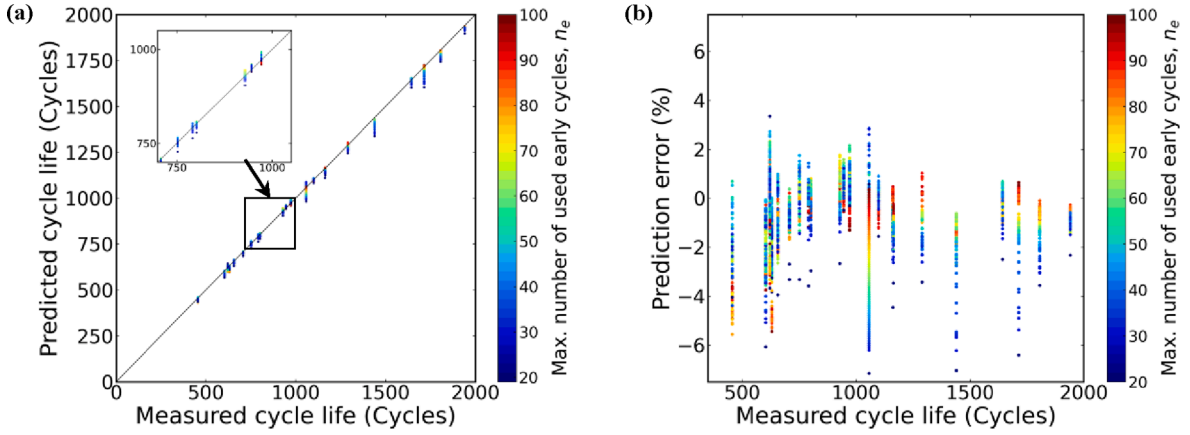

Fig. 6. Cycle life's early prediction of 30 different batteries from their first 20–100 charge cycles: (a) early prediction results of batteries' cycle lives; (b) prediction errors.

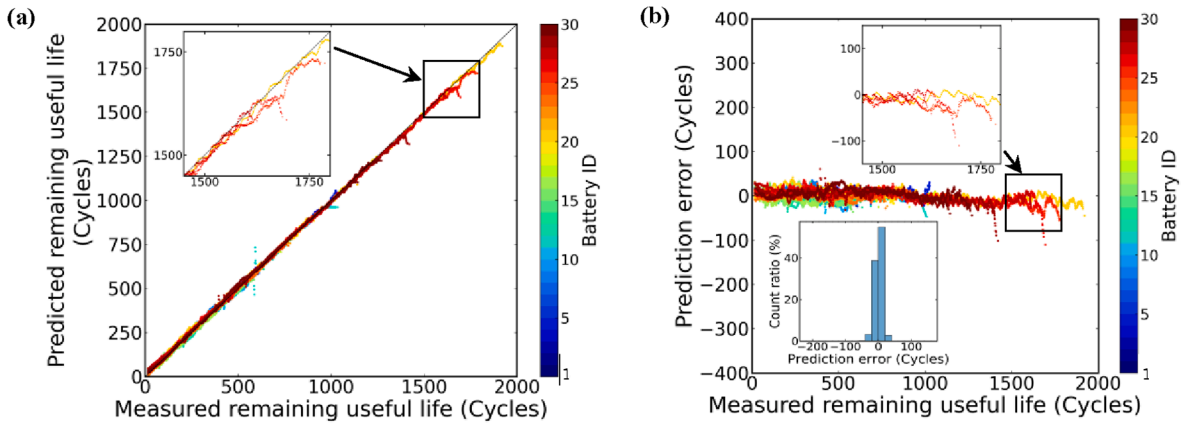

Fig. 7. Remaining useful cycle life prediction of 30 different batteries: (a) prediction results of batteries' remaining useful cycle lives; (b) prediction errors, where the inset shows the histogram of residuals (predicted – measured).

As shown in Fig. 7, for the 30 test batteries, the RUL of each battery varies greatly from 0 to X, where X represents the cycle life of battery which also varies greatly from about 450–2000 cycles under different charge policies. While because the features dominating battery RUL might have been successfully captured from the input charge V/I/T data and perfectly modelled by the proposed method, all the predicted RULs of all the batteries can be also very consistent with their measured real ones during their whole useful lifetime ranges, respectively. So that in the figure, some predicted results are covered by others. The count ratio of the prediction error less than 40 cycles can reach 99% as shown by the histogram analysis in Fig. 7 (b). Note that there might be some differences in the prediction error of RUL for different batteries. To illustrate the RUL prediction results for a single battery in detail, an example is shown in Fig. 8. By observing the curves, one can see that during the whole lifetime range of the battery, the error can be limited to a band of around 10 cycles, and the count ratio of the prediction error less than 10 cycles reaches 99% as shown by the histogram analysis in Fig. 8 (b). Encouraging prediction results of RUL are achieved both on the whole and for the single battery. As shown in Table 2, on the whole, the MAPE, RMSE and MAE of RUL prediction of all the different batteries are low as 3.6%, 11 cycles and 9 cycles, respectively. And as shown by the example for a single battery, the MAPE, RMSE and MAE might be as little as 1.4%, 4 cycles and 3 cycles, respectively.

Furthermore, to investigate the cases that the predicted battery failure life is earlier or later than the actual life, the errors of RUL prediction are divided into positive and negative errors here. For all tested batteries, calculated by the predicted RULs with positive errors, the MAPE, RMSE and MAE are 4.7%, 10 cycles and 8 cycles, respectively; while calculated by the predicted RULs with negative errors, the MAPE, RMSE and MAE are 1.8%, 14 cycles and 9 cycles, respectively. Compared with those shown in Table 2, one can see that they are almost at the same levels respectively. Therefore, in the following of this paper, the errors like those shown in Table 2 will be used to compare the proposed method with others.

These results fully proof that the proposed method can also enjoy high accuracy and strong stability for RUL prediction of different batteries under different charge polices during their whole useful lifetime ranges.

It is worth noting that, the working loads of different devices in practice might be different and random. Therefore, if battery RUL is predicted based on the discharge process like those existing methods, development of battery RUL prediction for different devices is usually required. In comparison, due to the generality for different charge policies and the avoidance of the influence of random working loads of devices on the RUL prediction, the proposed method can be applied for the practical applications of different devices, therefore, the huge cost and time of the development of battery RUL prediction for different devices can be avoided.

4.3. Comparison with the methods without MsCCA, CFA, 3DCNN layer or subnetwork-2

Different from the existing models, the MsCCA, CFA, 3DCNN layer and subnetwork-2 are proposed for battery life prediction in this work.

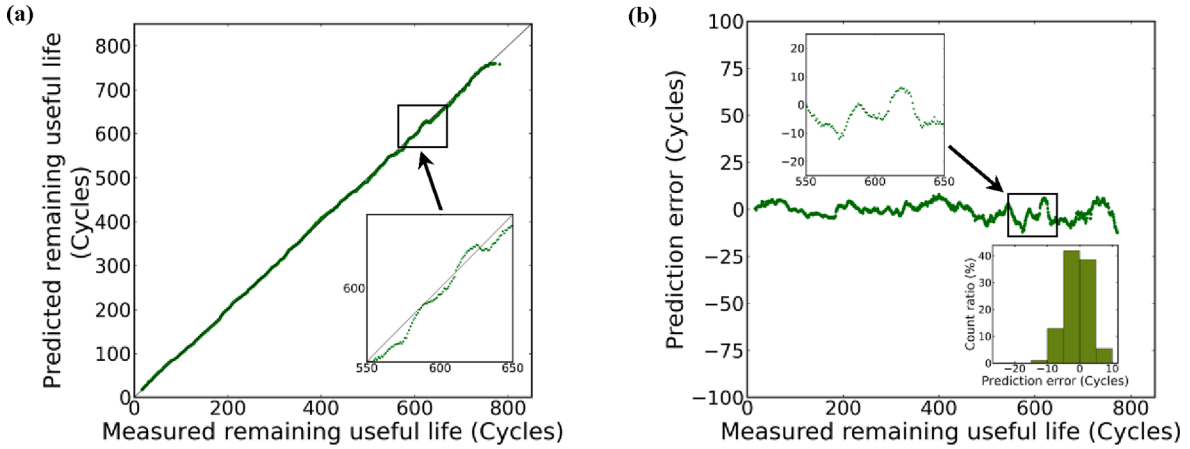

Fig. 8. An example of remaining useful cycle life prediction of a battery: (a) remaining useful cycle life prediction results; (b) prediction errors, where the inset shows the histogram of residuals (predicted – measured).

To analyze their effects on the prediction, we remove them respectively from our proposed full HCNN as shown in Fig. 3 (b), and compare their prediction results of battery cycle life and RUL in this section. For fair comparison, the parameters in these models are also tuned, and these models are also trained by the same training dataset, then tests of the battery cycle life early prediction and RUL prediction are conducted also based on the same test dataset as stated above. The numerical results are shown in Table 3. Obviously, when compared with the full HCNN, if the prediction error of a model is greater than that of the full HCNN, it means that the difference of the compared model from the full HCNN can reduce this error.

4.3.1. When without the proposed MsCCA or CFA

As one can see from Table 3, in the early prediction application: when without the MsCCA, since the relevance and salience of different charging cycles cannot be evaluated for battery life prediction, the three error indicators will increase by 50–66%; and by 6% on average when without the CFA, because the relevance and salience of different charging features cannot be estimated for battery life prediction. More importantly, in the RUL prediction application, due to much more predictions required for battery whole lifetime range, the lack of these two attentions would lead to a greater increase in prediction errors: the MAPE, RMSE and MAE may increase significantly by 1.4–2.7 times when without the MsCCA, and by 1.7–3.3 times when without the CFA. It is known that, the changes of battery characteristics in the early cycles are not as big as those in the subsequent cycles. So, the contribution of the CFA to the error reduction of the early prediction is not as huge as that of the RUL prediction. Results fully prove that these two proposed attentions all can remarkably heighten the prediction performance by evaluating the relationship between different features and cycles respectively.

4.3.2. When without the 3DCNN layer

It is known that, the earlier the battery life prediction during battery lifetime range, the more difficult it is to capture and evaluate the intrinsic characteristics of battery life states [3]. Therefore, the more important it is to fuse the battery V, I and T for battery life prediction when using the charge data from earlier cycles, due to the strong relationship between them. Consequently, as shown in Table 3, if without the help of the 3DCNN layer as the first layer in the model to fuse the battery charge V/I/T curves, the MAPE, RMSE and MAE may significantly increase by 55–79% in the battery life early prediction, and also by about 5% on average in the RUL prediction. Results indicate that the 3DCNN layer has strong ability to lower the errors when making early prediction of battery cycle life, and improves the prediction precision of RUL as well.

4.3.3. When without the subnetwork-2

As presented in Section 3.2, the subnetwork-2 takes difference between each curve and the first-cycle curve of the battery V/I/T as input. Therefore, model without subnetwork-2 is unable to learn from such difference for battery life prediction. Consequently, as indicated in Table 3, when being applied in early prediction, the model without subnetwork-2 has the MAPE, RMES and MAE about 2.5 times of those of the full model. And for RUL prediction, with assistance of subnetwork-2, the MAPE can decrease by about 62%, and the RMSE and MAE can be both cut down by more than 2 times. In addition, it can be easily noticed that the subnetwork-2 makes greater melioration in all the error indicators, except the MAPE in the RUL prediction (although it is also evidently reduced by subnetwork-2), than any other structures compared in the table. These results fully demonstrate the importance of using the differences of charge V/I/T curves between cycles to describe the battery life, and learning from these differences to obtain accurate prediction.

The above results indicate that, the MsCCA, CFA, 3DCNN layer and subnetwork-2 are of great significance for battery life early prediction and RUL prediction. The MsCCA and subnetwork-2 can significantly reduce the errors of both the battery cycle life early prediction and battery RUL prediction; the main contribution of the CFA is to remarkably reduce the error of battery RUL prediction, it also can slightly reduce the error of the battery cycle life early prediction; while the main contribution of the 3DCNN layer is to reduce the error of battery cycle life early prediction, it also has slight effect on reducing the error of battery RUL prediction. Therefore, one can see that, if without them, the very accurate prediction for both of such different prediction applications could not be achieved.

It is also worth pointing out that, generally, the calculation burden of 3DCNN is much greater than that of 2DCNN. However, in this paper, the 3DCNN layer is only used as the first layer of the model. And thanks to the input data generation as presented in Section 3.1 and the 3-D kernels used in the 3DCNN layer as stated in Section 3.2, like the 2DCNN layer there is no need to slide convolution kernel in depth direction. So the calculation time did not increase evidently. The algorithms were trained with a NVIDIA GPU of 1.60 GHz. For the HCNN without the 3DCNN layer, the calculation time of each iteration was 447 s, the number of iterations required to achieve convergence was 340, and the total model training time was 50 h; while for the full HCNN, they were 595 s, 352 and 58 h, respectively. Using the trained models, on a computer with a 2.2 GHz CPU processor (Intel(R) Core(TM) i5-5200U), the calculation time of one prediction of the full HCNN only increased to 0.091 s from 0.076 s of the HCNN without the 3DCNN layer. There is a slight increase of calculation time, while the prediction accuracy can be improved remarkably, especially for the battery cycle life early prediction.

4.4. Comparison with other published methods

To further verify the precision and stability of our proposed method, some other published methods are compared as shown in Table 4. Some works were not carried out among the published articles in Table 4, causing the unavailable comparison. Results show that the proposed method significantly outperforms the others. For example, the novel data-driven method of early prediction of lithium-ion battery cycle life [3] was recently published on the journal of Nature Energy. Based on the same dataset used above, the constant-current (CC) discharge data of the first 100 cycles are required for this method. It achieved a RMSE of 100 cycles in the primary test and 214 cycles in the secondary test. And the

Table 3
The numerical results of the RUL prediction and the early life prediction using different models.

Used Model		HCNN without MsCCA	HCNN without CFA	HCNN without 3DCNN layer	HCNN without subnetwork-2	Full HCNN
Early prediction of battery cycle life (n_e=60)	MAPE (%)	1.86	1.23	2.00	2.86	1.12
	RMSE (Cycles)	20	14	21	32	13
	MAE (Cycles)	17	11	17	26	11
RUL prediction for all the test batteries	MAPE (%)	9.66	11.79	3.76	9.25	3.55
	RMSE (Cycles)	17	21	12	25	11
	MAE (Cycles)	13	15	9	19	9

Table 4
Comparison with other published methods of battery life prediction.

Methods	Battery life early prediction			Battery RUL prediction	
	MAPE (%)	RMSE (Cycles)	Early cycles used	MAPE (%)	Ratio of required cycles for starting prediction [1] (%)
Proposed method	1.12 3.08	13 42	60 20	3.55	1–4
Data-driven prediction [3]	7.5, 10.7 [2]	100, 214 [2]	100	–	–
UKF- double exponential model [16]	– [3]	–	–	7.74	41–85
UKF-RVM [21]	–	–	–	14.64	38–96
UPF based on LOCR [4] [17]	–	–	–	4.27	30–35
CNN-LSTM [25]	–	–	–	6.43	23–63
LSTM-RNN [26]	–	–	–	6.99	21–24

[1] : its value "X" means that the required cycles for starting RUL prediction amounts to X% along the trajectory to battery failure. Here the ratio exists a range. This is because the prediction is conducted on different batteries, and they have different lives.
[2] : where the two values are obtained respectively in the primary test and in the secondary test in Ref. [3].
[3] : the symbol "-" denotes "unavailable".
[4] : "LOCR" is short for linear optimizing combination resampling.

test errors were 7.5% and 10.7% respectively in these two tests. To the best of our knowledge, these results are great in the published literatures. But by our proposed method, thanks to the different description of battery life and prediction model proposed in this paper, using the first 60 cycles, the test error is 1.1% and the RMSE is 13 cycles (see Table 2), reducing which the existing model achieves by 6.8–9.7 times, and 7.7–16.5 times, respectively. Especially, even only the first 20 cycles are used, a RMSE of around 45 cycles is achieved, and the mean error is also only 3.1%. That is to say that, compared with the existing method, just requiring a fifth of the cycle amount needed by the former, the proposed method can reduce the RMSE by 2.4–5.1 times, and the test error by 2.4–3.5 times.

In addition, the required data for RUL prediction in the current studies generally amounts to at least 20% along the trajectory to failure, whereas the proposed method can make the accurate prediction of battery life after only 20 cycles (about 1% along the trajectory to failure for a battery with life of about 2000 cycles). This indicates that the RUL prediction can be made much earlier than the existing methods. This might also be the reason why the existing RUL prediction methods were not used for battery cycle life early prediction, as shown in Table 4. Furthermore, by comparing the results in Table 4, one can see that the RUL prediction performance of the proposed method also improves remarkably, and MAPE reduces from 4.27–14.64% to 3.55%.

Moreover, the proposed method is based on the charge process which is controlled in all practical applications [32], and the battery charge V, I and T curves and their difference between several cycles are fused for the prediction, so its suitability for both battery cycle life early prediction and battery RUL prediction in practical devices can be guaranteed under different charge policies. While as mentioned earlier in this paper, due to the random working loads of devices in practical applications [30], it is almost impossible to keep constant current during discharge, so it is difficult to implement those existing RUL prediction methods [37] commonly based on CC discharge process in practical devices. This also is the reason why the existing battery life early prediction methods which are based on discharge process cannot be used for battery RUL prediction in real devices' applications, as illustrated in Table 4.

Note that, with the aim of the proposed method, a dataset in which batteries are charged with different charge policies is required. But I found, in the battery cycle life test dataset from NASA [38], all the batteries are cycled with the same charge policy (charged with constant current 1.5A to 4.2 V, then under constant voltage to 20 mA). And the same is for the dataset [39] from University of Maryland and others widely used, except for that provided in Ref. [3]. Therefore, in Table 4, the results of the proposed method using the dataset from Ref. [3] are compared. It is known that the indicator MAPE is comparable among different methods even if they use different datasets. Also, achieving accurate prediction under different charge policies is usually more difficult than that under the same charge or discharge conditions like those cited in Table 4, due to the different influences of different charge policies on battery life. The above comparison further highlighted the accurate prediction ability of proposed method.

5. Discussion and explanation

Above all, because the features hidden in the charge data of battery V, I and T and between cycles might have been perfectly fused, extracted and modelled, and the influence of the random working conditions of devices on the prediction can be avoided, the proposed method can achieve encouraging prediction results for both battery life early prediction before its capacity degradation, and RUL prediction during the whole useful lifetime of battery.

5.1. Discussion

We would like to point out that: since the graphite negative electrode dominates degradation in these batteries [31], the above results could be applicable to other types of lithium-ion batteries based on graphite. We note that the graphitic negative electrode is common to nearly all commercial lithium-ion batteries in use today [3].

We also would like to point out that: due to the lack of lithium-ion batteries data under different charge policies in different temperature environments, in Section 4, the proposed method is verified using the dataset collected in a controlled constant temperature chamber. However, as shown in Fig. 1 (c) and Fig. 2 (d), the temperatures of different batteries under different cycle numbers are different even in the constant temperature chamber. But from the verification presented above, the accuracy of the proposed method is very encouraging. Therefore, it can be concluded that when the temperature changes, the proposed method can still guarantee the effectiveness.

To proof this, and as a comparison, the accuracy of the proposed method in battery life estimation during discharge under different ambient temperatures is investigated here. The dataset from NASA [38] is selected, which is collected in temperature chamber by cycling commercial 18650 lithium-ion batteries. Specially, 15 batteries which are cycled to 70% of nominal capacity are selected here (in Section 4, this is 80%), for further testing the prediction accuracy of the proposed method when the criteria for end of life is different. The batteries are cycled with the same charge policy as stated above, while discharged in groups through different operational profiles. As illustrated in Table 5, there are 4 batteries (labeled as B0005, B0043, B0044 and B0055) with different discharge current profiles, respectively. During discharge, both B0043 and B0044 operate under multiple temperatures (22 and 4 °C), while B0005 and B0055 under 24 °C and 4 °C, respectively. And the cut-off voltages are also different. More details regarding the NASA test bench can be found in [38]. The proposed model is trained with the battery V, I and T data of the other selected batteries during discharge. The test results of the prediction of these four batteries are also summarized in Table 5. One can see that, as long as the discharge profiles of batteries can be controlled, the MAPE, RMSE and MAE can be less than 4%, 2 cycles and 2 cycles, respectively, by applying the proposed method. In other words, accurate prediction can also be achieved based on the controlled discharge process. Unfortunately, in the real operations, the discharge process of batteries cannot be controlled like that

Table 5
The investigation of the battery RUL prediction during discharge at different temperatures.

Battery ID	Ambient temperature (℃)	Discharge current (A)	lower cut-off voltage (V)	Capacity loss (end of life) (%)	MAPE (%)	RMSE (Cycles)	MAE (Cycles)
B0005	24	2	2.7	30	3.82	2	2
B0043	22 & 4 [1]	4 & 1 [2]	2.5	30	2.93	1	1
B0044	22 & 4 [1]	4 & 1 [2]	2.7	30	3.74	2	1
B0055	4	2	2.5	30	3.74	2	1

[1] : where the two values indicate the multiple ambient temperatures at which the battery was run.
[2] : where the values indicate the multiple load current levels which were used during discharging

during charging. This is why, as presented earlier, the prediction is proposed to be performed based on charge process in this paper. These results also proof that the proposed method can achieve accurate prediction under different temperatures.

Furthermore, because the battery V, I and T curves during charging are fused for the prediction, the changes of them can be automatically extracted and estimated for the prediction. So, the feature learning of the proposed method and its accuracy will not be affected. Specially, the prediction does not need the data during battery discharging, so the changes of battery discharge rate, temperature and level etc. will also not affect the feature learning of the proposed method and its accuracy.

It is noteworthy that, for both the battery cycle life early prediction and RUL prediction, the model and all the parameters are the same, as stated in Section 4.2. This means that the proposed method and the trained model can be used as a uniform method and model for different applications:

(1) not only can be used by battery manufactures, to early evaluate the battery cycle life before battery capacity degradation for accelerating their development of battery. As verified in Section 4.2.1, by applying the proposed method and trained model, very accurate prediction results of battery cycle life can be achieved only using the charge data from the first 60 cycles. This means that users only need to conduct such small number of test cycles (not all), then they can perform the prediction to evaluate the battery cycle life accurately. Therefore, a lot of time (usually from months to years) and cost of battery life test can be saved (as stated in Section 4.2.1), accordingly, unlocking new opportunities in battery production, use and optimization.
(2) but also can be written in those battery management systems with powerful controller, charger system or remote battery management platform deployed in the cloud, to accurately estimate battery RUL during the actual operation of different devices in different scenarios, such as the electric vehicles, hybrid electric vehicles, microgrids and so on. Different devices usually have different and random working loads, which have influences on RUL prediction. While by applying the proposed method, such influences can be avoided through making the prediction based on the charge process which is controlled in all devices. The RUL can be predicted based on the input generated from the data sampled in the battery charge processes of real devices. And as verified in Section 4.2.2, under different charge policies or current, accurate prediction of RUL can be achieved. Therefore, the proposed method and the trained model can be applied to different devices, achieving accurate prediction of RUL in actual operations of batteries in devices.
(3) moreover, can be used to find a suitable charging strategy for battery life. For example, as verified above, by applying the proposed method and trained model, the cycle life of the battery can be early predicted accurately only using the charge data from the first 60 cycles under different charge policies. Accordingly, by applying them, users can quickly evaluate the influences of different charge polices on battery life through a small number of early tests, then quickly find a suitable charge policy in battery development to provide guidance for battery charge strategy determination.

5.2. Explanation of the learned results

Earlier in this paper, we have explained why FCA, MsCCA, 3DCNN layer and subnetwork-2 are applied for battery life prediction, and quantitatively evaluated their contributions to the reduction of the prediction errors. In this section, trying to explain the learned results for the battery life prediction, the qualitative evaluation is performed on them by visualizing the learned weights.

Note that, as presented in Section 3, the battery V/I/T curves during charging from 20 cycles are inputted for each prediction, and the CFA, MsCCA as well as the 3DCNN layer are all included in the subnetwork-1 and subnetwork-2. So below in this section, the learned results of them in the two subnetworks are evaluated at the same time.

5.2.1. The learned results of CFA

To demonstrate the charging feature salience detection, in Fig. 9, three inputs for battery life prediction are presented to visualize the learned results of CFA. For each group corresponding to subnetwork-1 or subnetwork-2, we first visualize the attention values from our CFA directly, then we up-sample the attention weight distribution back to the size of input. This allows us to visualize the receptive field of attention.

Note that, as stated in Section 3.1, the row numbers of the input matrices correspond to the charge capacity ratios (which range is divided into 1024 points in practice, and the smaller ones denote smaller capacity ratios); and the column numbers correspond to the first 5 cycles and the last 15 cycles used for a prediction of battery life. These heat maps show that, both in the subnetwork-1 and 2, our CFA is able to not only distinguish the core regions from the input charge V/I/T curves effectively, but also capture the salient parts inside the regions. Specifically, the area with capacity ratio lower than 0.4 (corresponding to row number of about 250 in the up-sampled plots) is detected as salient part by the CFA in subnetwork-1; while the area with capacity ratio greater than 0.85 (corresponding to row number of about 840 in the up-sampled plots) is detected as salient part for battery life prediction by the CFA in subnetwork-2. It is reasonable since these areas tend to generate difference evidently when battery ages or life varies [31], which are discriminative charging features for the task of battery life prediction. In addition, some other local areas are also detected as salient parts by the two CFA modules, respectively. And different points, which represent the different features extracted from the input charge V/I/T curves, are weighted. It helps further improve the prediction accuracy, since each charging feature contributes less or more to the predictions of battery lives, and their contributions are all scored reasonably by the learned weights of CFA.

5.2.2. The learned results of MsCCA

To demonstrate the charging cyclic salience detection, the learned cyclic attention weights for several representative inputs for RUL prediction of a battery during its whole lifetime are shown in Fig. 10. Note that for each prediction, there are three MsCCA weight values (the w_{s1}, w_{s2} and w_{s3} in Eq. (17), which are represented as w_1, w_2 and w_3 in Fig. 10, respectively), because the three-layer DCN is applied in our

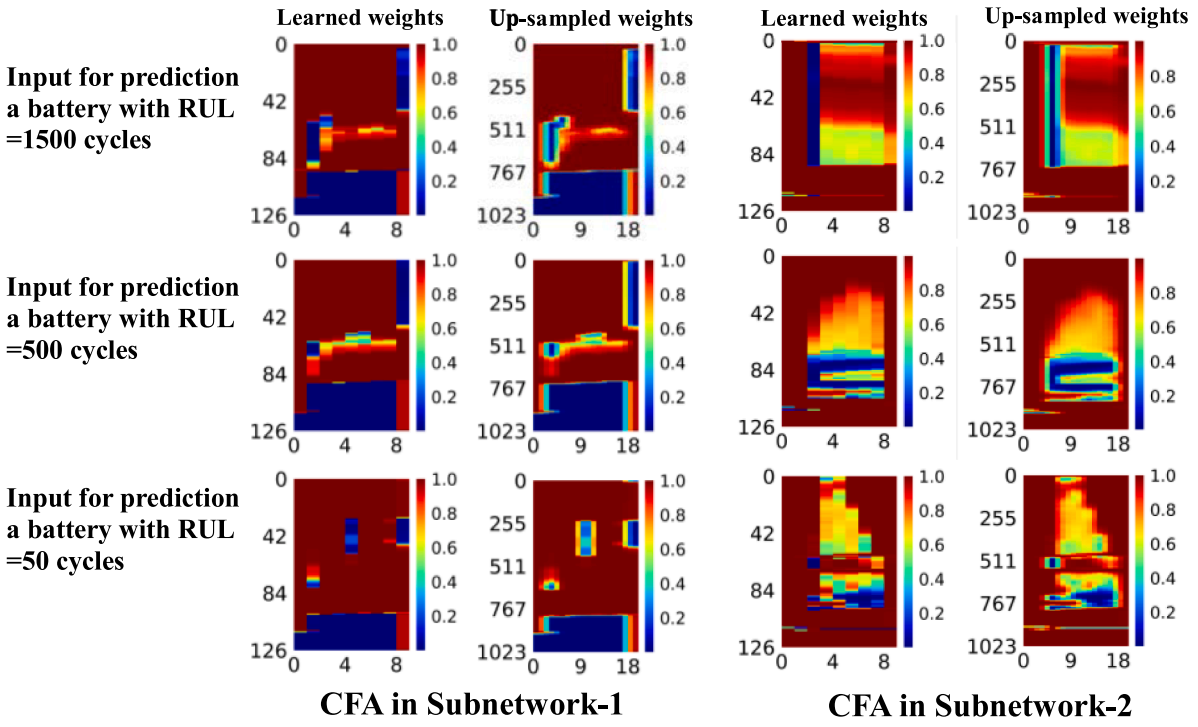

Fig. 9. Examples of heat map visualization of the learned charging feature attention weights by CFA in subnetwork-1 and subnetwork-2.

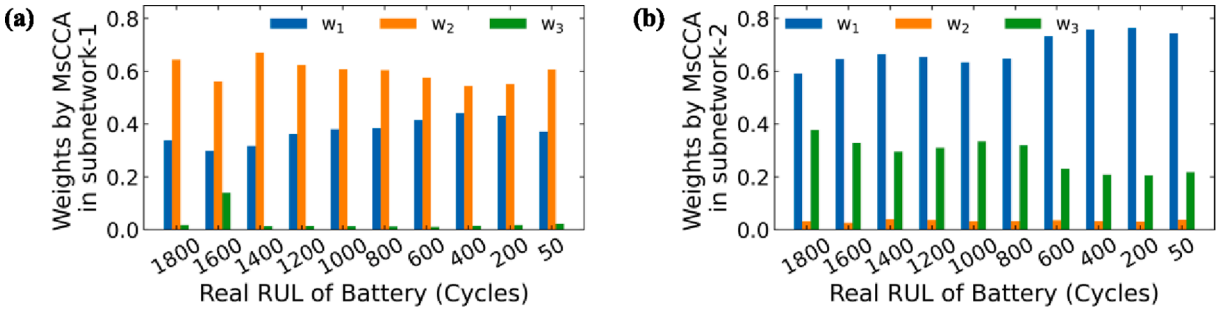

Fig. 10. Examples of visualization of learned charging cyclic attention weights by our MsCCA in different subnetworks: (a) in subnetwork-1; (b) in subnetwork-2. Note that the three attention weights for one RUL prediction are normalized (see those in Eq. (17)) to make them sum up to 1.

MsCCA in practice, as stated in Section 3.4. And for the three layers, the dilation rates at width dimension were selected as 1, 2 and 5. One can see that, in subnetwork-1: $w_2 > w_1 > w_3 > 0$; in subnetwork-2: $w_1 > w_3 > w_2 > 0$. The learned results indicate that our MsCCA is able to catch the multi-scale charging features between different cycles at different intervals, and detect their salience for the battery life prediction. In addition, in the RUL predictions of a battery during its whole lifetime, each of the three learned weights varies nonlinearly either in subnetwork-1 or in subnetwok-2. It is reasonable because the evolution of battery degradation is nonlinear, which also is the reason why the learned values of these three weights are different. w_3 learned in subnetwork-1 is small, but its value leaned in subnetwork-2 is big; w_2 learned in subnetwork-2 is small, but its value learned in subnetwork-1 is big; while the w_1 retains a relatively high value, both in subnetwork-1 and 2. The three weights learned by MsCCA modules in subnetwork-1 and 2, as a whole, play an important role which is helpful for accurate prediction of battery life.

5.2.3. The learned results of 3DCNN layer

The 3DCNN layer is used to fuse the battery charge V, I and T for battery life prediction. To demonstrate what it learns from the input of battery charge V, I and T curves and their difference between cycles, several 3-D kernels of the 3DCNN layer after training are randomly selected and visualized as examples, as shown in Fig. 11. Note that, the dimension in depth direction of the 3-D kernel is 3 as the same as that of the input. So, it can be considered to be composed of three matrices (here, denoted as $z = 1$, 2 and 3 respectively), which correspond to the three input matrices of battery voltages, currents and temperatures, respectively. This allows us to show the learned weights in the 3-D kernels clearly. As one can see that, in the 3-D kernels of both 3DCNN layer in subnetwork-1 and 2, the learned weights in their three matrices are different, and the weights in different 3-D kernels are also different. It means that, the battery V, I and T are fused with different weights, and the features hidden in the input curves are extracted and weighted differently for battery life prediction. It helps improve the battery life prediction accuracy because of the reasons as analyzed in Section 2.3 and explained in Section 3.2.

6. Conclusion

In this paper, a HCNN method is proposed to make battery cycle life early prediction and RUL prediction for lithium-ion batteries. Trying to make the prediction accurate and suitable for such different applications, the battery V, I and T curves of charge processes from several

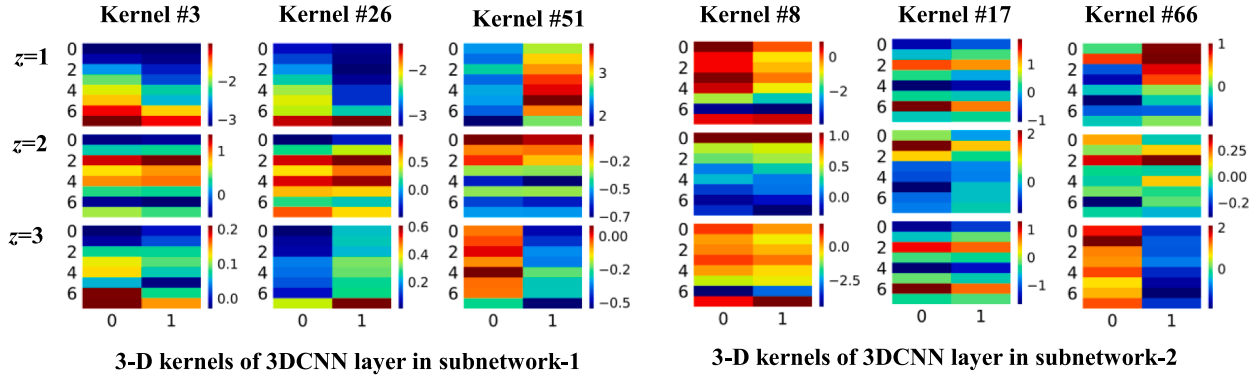

Fig. 11. Examples of visualization of learned weights by 3DCNN layer in subnetwork-1 and subnetwork-2.

cycles and especially the difference between these cycles are first utilized for the description of battery cycle life and RUL for reasonable and robust consideration. And a HCNN, which is based on a fusion of 3DCNN and 2DCNN with two subnetworks in particular, is designed. The battery V, I and T are first fused by a 3DCNN layer, and the features hidden in these curves and their difference between cycles are extracted and modelled automatically by two 2DCNN layers for the predictions. Furthermore, a charging feature attention algorithm and a multi-scale charging cycle attention algorithm are defined to estimate the relationship between different features and cycles respectively to further heighten the prediction performance.

To verify the proposed method, tests and numerical comparisons are carried out. The accuracy and robustness of the proposed method are systematically verified by experiments. Results show that, the proposed method has a reliable performance of battery cycle life early prediction and RUL prediction, and the trained model can be used as a unified model for such different applications. For different batteries under different charge policies: in the RUL prediction, the proposed method achieves 3.6% test error during the whole lifetime ranges of the batteries; and in the battery cycle life early prediction, it achieves 1.1% test error only using the charge data from the first 60 cycles. When compared with other published methods, the results demonstrate that the proposed method significantly outperforms the published methods in terms of precision, stability and suitability for different applications. For example, the proposed method can reduce the number of the required cycles for early prediction by 80%, and reduce RMSE by 8–17 times and mean error by 6.8–9.7 times. The proposed method can be applied for both battery cycle life early accurate prediction before battery capacity degradation and RUL accurate prediction for different devices' practical operations. Moreover, due to the ability to accurately predict the battery life under different charge policies, the proposed method can also be used to find the suitable charge policy for battery life.

The data and codes for the work are available at https://doi.org/10.17632/ms2f63ng3y.1.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors would like to thank K.A. Severson, P.M. Attia and N. Jin, et al. for providing the experimental datasets of lithium-ion batteries.

References

[1] Omariba Z, Zhang L, Sun D. Review on health management system for lithium-ion batteries of electric vehicles. Electronics 2018;7(5):72.
[2] Bloom I, Cole B, Sohn J, Jones S, Polzin E, Battaglia VS, et al. An accelerated calendar and cycle life study of Li-ion cells. J Power Sources 2001;101:238–47.
[3] Severson K, Attia P, Jin N, Perkins N, Jiang B, Yang Z, et al. Data-driven prediction of battery cycle life before capacity degradation. Nat Energy 2019;4:383–91.
[4] Wu L, Fu X, Guan Y. Review of the remaining useful life prognostics of vehicle lithium-ion batteries using data-driven methodologies. Appl Sci 2016;6:166.
[5] Wang Y, Zhang C, Chen Z. An adaptive remaining energy prediction approach for lithium-ion batteries in electric vehicles. J Power Sources 2016;305:80–8.
[6] Harris S, Harris D, Li C. Failure statistics for commercial lithium ion batteries: a study of 24 pouch cells. J Power Sources 2017;342:589–97.
[7] Sarasketa-Zabala E, Gandiaga I, Rodriguez-Martinez L, Villarreal I. Calendar ageing analysis of a LiFePO4/graphite cell with dynamic model validations: towards realistic lifetime predictions. J Power Sources 2014;272:45–57.
[8] Wang D, Miao Q, Pecht M. Prognostics of lithium-ion batteries based on relevance vectors and a conditional three-parameter capacity degradation model. J Power Sources 2013;239:253–64.
[9] Eddahech A, Briat O, Woirgard E, Vinassa J. Remaining useful life prediction of lithium batteries in calendar ageing for automotive applications. Microelectron Reliab 2012;52:2438–42.
[10] Pattipati B, Sankavaram C, Pattipati K. System identification and estimation framework for pivotal automotive battery management system characteristics. IEEE Trans Syst Man Cybern C: Appl Rev 2011;41:869–84.
[11] Zhou Y, Huang M, Chen Y, Tao Y. A novel health indicator for on-line lithium-ion batteries remaining useful life prediction. J Power Sources 2016;321:1–10.
[12] Sun B, Jiang J, Zheng F, Zhao W, Liaw B, Ruan H, et al. Practical state of health estimation of power batteries based on Delphi method and grey relational grade analysis. J Power Sources 2015;282:146–57.
[13] He Z, Gao M, Ma G, Liu Y, Chen S. Online state-of-health estimation of lithiumion batteries using Dynamic Bayesian Networks. J Power Sources 2014;267:576–83.
[14] Ren L, Zhao L, Hong S, Zhao S, Wang H, Zhang L. Remaining useful life prediction for lithium-ion battery: a deep learning approach. IEEE Access 2018;6:50587–98.
[15] Dalal M, Ma J, He D. Lithium-ion battery life prognostic health management system using particle filtering framework. Proc Inst Mech Eng O: J Risk Reliab 2011;225(1):81–90.
[16] Chang Y, Fang H, Zhang Y. A new hybrid method for the prediction of the remaining useful life of a lithium-ion battery. Appl Energy 2017;206:1564–78.
[17] Zhang H, Miao Q, Zhang X, Liu Z. An improved unscented particle filter approach for lithium-ion battery remaining useful life prediction. Microelectron Reliab 2018;81:288–98.
[18] Nuhic A, Terzimehic T, Soczkaguth T, Michael B, Klaus D. Health diagnosis and remaining useful life prognostics of lithium-ion batteries using data-driven methods. J Power Sources 2013:680–8.
[19] Feng X, Weng C, He X, Han X, Lu L, Ren D, et al. Online state-of-health estimation for Li-ion battery using partial charging segment based on support vector machine. IEEE Trans Veh Technol 2019;68(9):8583–92.
[20] Qin T, Zeng S, Guo J. Robust prognostics for state of health estimation of lithium-ion batteries based on an improved PSO–SVR model. Microelectron Reliab 2015;55:1280–4.
[21] Zheng X, Fang H. 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 Safe 2015;144:74–82.
[22] Richardson RR, Osborne MA, Howey DA, et al. Gaussian process regression for forecasting battery state of health. J Power Sources 2017:209–19.
[23] Richardson RR, Osborne MA, Howey DA. Battery health prediction under generalized conditions using a Gaussian process transition model. J Storage Mater 2019;23:320–8.
[24] Liu J, Saxena A, Goebel K, Saha B, Wang W. An adaptive recurrent neural network for remaining useful life prediction of lithium-ion batteries. In: Annual conference of the prognostics and health management society. 2010: 1–9.

[25] Ma G, Zhang Y, Cheng C, Zhou B, Hu P, Yuan Y. Remaining useful life prediction of lithium-ion batteries based on false nearest neighbors and a hybrid neural network. Appl Energy 2019;253:113626.
[26] Zhang Y, Xiong R, He H, Pecht M. Long short-term memory recurrent neural network for remaining useful life prediction of lithium-ion batteries. IEEE Trans Veh Technol 2018;67:5695–705.
[27] Liu K, Shang Y, Ouyang Q, Widanage WD. A data-driven approach with uncertainty quantification for predicting future capacities and remaining useful life of lithium-ion battery. IEEE Trans Ind Electron 2021;68(4):3170–80.
[28] Cheng Q, Bondon P. A new unscented particle filter. ICASSP 2008. In: IEEE international conference on acoustics, speech and signal processing, 2008. IEEE; 2008. p. 3417–20.
[29] Patil M, Tagade P, Hariharan K, Kolake S, Song T, Yeo T, et al. A novel multistage Support Vector Machine based approach for Li ion battery remaining useful life estimation. Appl Energy 2015;159:285–97.
[30] Yang L, Cai Y, Yang Y, Deng Z. Supervisory long-term prediction of state of available power for lithium-ion batteries in electric vehicles. Appl Energy 2020; 257:114006.
[31] Anseán D, Dubarry M, Devie A, Liaw B, Garcia V, Viera J, et al. Fast charging technique for high power LiFePO4 batteries: a mechanistic analysis of aging. J Power Sources 2016;321:201–9.
[32] Lee C, Wu Z, Hsu S, Jiang J. Cycle life study of li-ion batteries with an aging-level-based charging method. IEEE Trans Energy Convers 2020;35:1475–84.
[33] Richardson RR, Birkl CR, Osborne MA, Howey D. Gaussian process regression for in-situ capacity estimation of lithium-ion batteries. IEEE Trans Ind Inform 2019;15 (1):127–38.
[34] Li D, Yang L. Remaining useful life prediction of lithium battery using convolutional neural network with optimized parameters. In: IEEE 5th Asia Conference on Power and Electrical Engineering, 2020;840–44.
[35] Anagun Y, Isik S, Seke E. SRLibrary: Comparing different loss functions for super-resolution over various convolutional architectures. J Visual Commun Image Represent 2019;61:178–87.
[36] Szegedy C, Liu W, Jia Y, Sermanet P, Reed S, Anguelov D, et al. Going deeper with convolutions. In: Proceedings of the IEEE conference on computer vision and pattern recognition. 2015: 1–9.
[37] Li Y, Liu K, Foleyd AM, Zülkea A, Berecibarb M, Nanini-Mauryf E, et al. Data-driven health estimation and lifetime prediction of lithium-ion batteries: a review. Renew Sustain Energy Rev 2019;113:109254.
[38] Saha B, KG. Battery data set. NASA Ames Prognostics Data Repository. NASA Ames Research Center, Moffett Field, CA; 2007. <https://ti.arc.nasa.gov/tech/dash/groups/pcoe/prognostic-data-repository/>.
[39] He W, Williard N, Osterman M, Pecht M. Prognostics of lithium-ion batteries based on Dempster-Shafer theory and the Bayesian Monte Carlo method. J Power Sources 2011;196:10314–21.