A conditional random field based feature learning framework for battery capacity prediction

Hai-Kun Wang1,2*, Yang Zhang1 & Mohong Huang1

This paper proposes a network model framework based on long and short-term memory (LSTM) and conditional random field (CRF) to promote Li-ion battery capacity prediction results. The model uses LSTM to extract temporal features from the data and CRF to build a transfer matrix to enhance temporal feature learning for long serialization prediction of lithium battery feature sequence data. The NASA PCOE lithium battery dataset is selected for the experiments, and control tests on LSTM temporal feature extraction modules, including recurrent neural network (RNN), gated recurrent unit (GRU), bi-directional gated recurrent unit (BiGRU) and bi-directional long and short term memory (BiLSTM) networks, are designed to test the adaptability of the CRF method to different temporal feature extraction modules. Compared with previous Li-ion battery capacity prediction methods, the network model framework proposed in this paper achieves better prediction results in terms of root mean square error (RMSE) and mean absolute percentage error (MAPE) metrics.

Rechargeable lithium-ion batteries play a crucial role in many modern applications1,2, from portable electronics and medical devices to renewable energy integration in power grids and electric vehicles, due to their long cycle life, low self-discharge rate, wide range of applications, and low environmental pollution3. However, as a product of industrialization, the performance of lithium batteries4 decreases with increasing usage time. As the capacity of the battery decreases, the battery will eventually fail. Therefore, it is important to predict the future health5 of lithium batteries to improve the reliability of energy systems.

The future health of Li-ion batteries is usually measured by the indicator state of health (SOH), which reflects the battery's usage by calculating the decay process of the battery capacity and obtaining the difference between the battery's current usage state and its factory characteristics. When the current capacity decreases from 80 to 70% of the nominal capacity, the performance of the battery will decline rapidly. When the capacity of a lithium battery decreases to 70%, it is usually regarded as a battery failure. Therefore the accuracy of battery capacity prediction often affects the SOH calculation results, leading to bias in the estimation of battery usage and affecting the subsequent predictive maintenance strategy and energy system management. Li-ion battery capacity prediction6 is essentially a regression problem. Current methods for capacity decay estimation mainly include equivalent circuit models, electrochemical models and data-driven models7.

The equivalent circuit model uses circuit elements with empirical nonlinear parameters to build a mathematical model representing the battery system and combines measured data to estimate battery capacity and SOH8,9. Wei et al.10 used the Thevenin model by constructing the battery health index online, using the health index as the input parameter and the battery capacity decay as the output parameter in a gray neural network model to estimate the battery SOH. He et al.11 compared resistor-circuit equivalent circuit models of different orders, and adding RC series modules using the high nonlinearity of the battery operating characteristics can improve the accuracy and reliability of the model prediction, but reduces the applicability of real-time12. The capacity prediction method of the equivalent circuit model has the advantages of fewer parameters and high timeliness, but it is difficult to achieve higher prediction accuracy because it ignores the effects caused by environmental changes and data acquisition during the operation of lithium batteries13,14.

The electrochemical model establishes a physical model for SOH estimation and prediction by studying the chemical processes occurring inside the battery during operation15. Zhang et al.16 improved the prediction accuracy by analyzing the impedance characteristics based on the changes in the impedance properties of the battery. Ashwin et al.17 established an electrochemical battery aging model under cyclic loading conditions and constructed the capacity decay of the lithium-ion battery process. The electrochemical modeling method can accurately describe the movement pattern of positive and negative electrodes and the changing trend of active materials.

1School of Artificial Intelligence, Chongqing University of Technology, Chongqing 40400, China. 2Chongqing Industrial Big Data Innovation Center Co., Ltd., Chongqing 40400, China. *email: hkwang@cqut.edu.cn
Zhang et al.\(^{31}\) used LSTM and RNN networks to capture the long-standing relationship between lithium battery propagation neural networks (BP), gated recurrent units (GRU), and long short-term memory networks (LSTM). Work approaches mainly include artificial neural networks (ANN), convolutional neural networks (CNN), back propagation neural networks (BP), and neural network methods, and fusion methods. Statistical filtering is a method to extract and reproduce valid signals and waveforms from data containing a large number of noisy signals, and the best weighting factor with a strong target following ability is automatically calculated and determined by a recursive linear data processing algorithm\(^{20,21}\). He et al.\(^{12}\) used the extended Kalman filtering (EKF) algorithm to estimate the unknown parameters in the time degradation parameter model of lithium-ion battery capacity and obtained the future phase prediction results of the degradation trend of the remaining battery capacity. Support vector machine (SVM) as a nonlinear data analysis method, can not only provide relatively accurate estimation and prediction with a small amount of data but also improve the data quality to a certain extent and overcome the drawback that the model falls into the optimal local extremes. Fewer unknown parameters and high sparsity are the characteristics of this method\(^{20,21}\). Zhang et al.\(^{25}\) used to improve the prediction performance and operational efficiency of the battery by optimizing the relevance vector machine (RVM), by optimizing the RVM. Gao et al.\(^{26}\) proposed a single radial basis kernel function based on the novel multicore SVM based on polynomial kernel and radial basis kernel function for predicting the remaining useful life (RUL) of Li-ion batteries, which has better prediction accuracy and stronger generalization ability compared to SVM while reducing training time and computational complexity. A neural network is a nonlinear prediction method composed of many neurons according to certain rules. The neurons contained in the network model are trained to connect weights and thresholds to build an accurate estimation and prediction model\(^{20,21}\). Increasing the depth of the neural network can approach any nonlinear mapping with a simple structure and high learning ability\(^{20,21}\). Neural network approaches mainly include artificial neural networks (ANN), convolutional neural networks (CNN), back propagation neural networks (BP), and long short-term memory networks (LSTM). Zhang et al.\(^{11}\) used LSTM and RNN networks to capture the long-standing relationship between lithium battery capacity degradation for prediction. Fan et al.\(^{12}\) proposed a GRU-CNN network for learning shared information and time dependence of charging profiles, including characteristic variation profiles such as voltage, current, and temperature for estimating SOH. Zhou et al.\(^{33}\) improved the prediction accuracy of the model by capturing the local capacity regeneration phenomenon generated by the battery during charging and discharging through time convolutional networks (TCN). The fusion method is based on the characteristics of different algorithms, each taking their strengths for fusion improvement, which not only ensures the accuracy of the predicted data but also provides an accurate assessment of the prediction uncertainty. Liu et al.\(^{34}\) proposed a fusion algorithm based on least squares support vector regression (LSSVR) and hidden markov model (HMM) to predict the health status of rolling bearings, where LSSVR was used to predict the feature signal, and HMM was used to identify state features. Hong et al.\(^{35}\) proposed a fusion estimation method for SOH of lithium-ion batteries based on capacity incremental analysis and a weighted Kalman filter algorithm, which has higher prediction accuracy compared to the common Kalman filter method. The recent Li-ion battery capacity prediction models are detailed in Table 1.

### Table 1. Recent prediction models for lithium battery.

| Authors                     | Year   | Approach                                                                 |
|-----------------------------|--------|---------------------------------------------------------------------------|
| Zhang et al.\(^{31}\)       | 2016   | Relevance vector machine                                                  |
| Wang et al.\(^{37}\)        | 2017   | State space model                                                         |
| Gao et al.\(^{34}\)         | 2017   | Multi-kernel support vector machine with particle swarm optimization      |
| Zhang et al.\(^{38}\)       | 2018   | Particle filter and unscented Kalman filter                               |
| Zhang et al.\(^{35}\)       | 2018   | LSTM                                                                     |
| Ren et al.\(^{41}\)         | 2018   | Autoencoder with deep neural network                                      |
| Fang et al.\(^{56}\)        | 2019   | Double extended Kalman filter                                             |
| Deng et al.\(^{56}\)        | 2019   | Least squares support vector machine                                      |
| Fan et al.\(^{56}\)         | 2020   | GRU-CNN                                                                   |
| Zhou et al.\(^{33}\)        | 2020   | Temporal convolution network                                              |
| Song et al.\(^{36}\)        | 2020   | Principal component analysis and support vector machine                   |
| Ren et al.\(^{40}\)         | 2020   | CNN-LSTM                                                                  |
| Kodjo S.R.Mawonou et al.\(^{32}\) | 2021  | Random forest                                                            |
| Hong et al.\(^{54}\)        | 2021   | Locally linear embedding and isomap                                        |
| Jungsoo Kim et al.\(^{35}\) | 2022   | Genetic algorithm and pseudo-2-dimensional model                          |

Substances during the capacity decay of lithium ions by analyzing the detailed internal electrochemical reaction process and reaction intensity during the aging process of the battery. However, the electrochemical system is more complex and the characteristic parameters are coupled with each other, which makes its dynamic prediction accuracy poor and difficult to achieve wide applicability.

The data-driven model establishes the mapping relationship between the characteristic parameters and the health condition from the overall level of the data by extracting the characteristic values of the measured parameters\(^{18,19}\). Depending on the data mining methods, they are mainly divided into statistical filtering methods, support vector methods, neural network methods, and fusion methods. Statistical filtering is a method to extract and reproduce valid signals and waveforms from data containing a large number of noisy signals, and the best weighting factor with a strong target following ability is automatically calculated and determined by a recursive linear data processing algorithm\(^{20,21}\). He et al.\(^{12}\) used the extended Kalman filtering (EKF) algorithm to estimate the unknown parameters in the time degradation parameter model of lithium-ion battery capacity and obtained the future phase prediction results of the degradation trend of the remaining battery capacity. Support vector machine (SVM) as a nonlinear data analysis method, can not only provide relatively accurate estimation and prediction with a small amount of data but also improve the data quality to a certain extent and overcome the drawback that the model falls into the optimal local extremes. Fewer unknown parameters and high sparsity are the characteristics of this method\(^{20,21}\). Zhang et al.\(^{25}\) used to improve the prediction performance and operational efficiency of the battery by optimizing the relevance vector machine (RVM), by optimizing the RVM. Gao et al.\(^{26}\) proposed a single radial basis kernel function based on the novel multicore SVM based on polynomial kernel and radial basis kernel function for predicting the remaining useful life (RUL) of Li-ion batteries, which has better prediction accuracy and stronger generalization ability compared to SVM while reducing training time and computational complexity. A neural network is a nonlinear prediction method composed of many neurons according to certain rules. The neurons contained in the network model are trained to connect weights and thresholds to build an accurate estimation and prediction model\(^{20,21}\). Increasing the depth of the neural network can approach any nonlinear mapping with a simple structure and high learning ability\(^{20,21}\). Neural network approaches mainly include artificial neural networks (ANN), convolutional neural networks (CNN), back propagation neural networks (BP), and long short-term memory networks (LSTM). Zhang et al.\(^{11}\) used LSTM and RNN networks to capture the long-standing relationship between lithium battery capacity degradation for prediction. Fan et al.\(^{12}\) proposed a GRU-CNN network for learning shared information and time dependence of charging profiles, including characteristic variation profiles such as voltage, current, and temperature for estimating SOH. Zhou et al.\(^{33}\) improved the prediction accuracy of the model by capturing the local capacity regeneration phenomenon generated by the battery during charging and discharging through time convolutional networks (TCN). The fusion method is based on the characteristics of different algorithms, each taking their strengths for fusion improvement, which not only ensures the accuracy of the predicted data but also provides an accurate assessment of the prediction uncertainty. Liu et al.\(^{34}\) proposed a fusion algorithm based on least squares support vector regression (LSSVR) and hidden markov model (HMM) to predict the health status of rolling bearings, where LSSVR was used to predict the feature signal, and HMM was used to identify state features. Hong et al.\(^{35}\) proposed a fusion estimation method for SOH of lithium-ion batteries based on capacity incremental analysis and a weighted Kalman filter algorithm, which has higher prediction accuracy compared to the common Kalman filter method. The recent Li-ion battery capacity prediction models are detailed in Table 1.

To improve the accuracy of lithium battery capacity prediction, this paper proposes a Li-ion battery capacity prediction model with CRF as the core. CRF is a discriminative probabilistic model about the temporal sequence, which is widely used in natural language processing (NLP)\(^{46-47}\). CRF constructs the state transfer matrix by the trend of the changing relationship of the neighboring labels and obtains the probability distribution of the prediction sequence by reverse decoding, where the state sequence with the highest probability is the optimal prediction.
result. The model adds CNN networks to learn feature data at different scales and LSTM networks to collect time-series relationship information. The feasibility and effectiveness of the model were verified on the lithium battery dataset provided by NASA, and the prediction accuracy of the model was improved compared with other network models. The CNN-LSTM-CRF model provides a new idea for the lithium battery prediction problem.

The main contributions of this work are:

1. The CRF method is attempted to be introduced in the capacity prediction problem to calculate the observed state of the capacity prediction sequence by the offset matrix of the feature data, which more intuitively reflects the change of the capacity decline trend.

2. To improve the prediction accuracy of the CRF model, the study incorporates a CNN convolution module for collecting feature data at different time scales and an RNN time-linked module for capturing the changing trend of feature data on the before-and-after time difference and extracting its time-series relationship information. To verify the fit of the CRF prediction model to different time-linked modules, the study added GRU, LSTM, BiLSTM, and other control experiments, and the experimental results on the NASA lithium battery dataset showed that LSTM achieved better results.

**Methodology**

**Overall framework of model.** Lithium-ion battery residual life prediction is based on the analysis and processing of lithium battery use data to estimate the residual life of the battery. This paper studies how to make the prediction results more accurate and improve the robustness of the model. Since the test time points of each charge and discharge cycle are different, the test number of one cycle at the maximum collection point in the data set is taken as the standard, and the zero vector is used to supplement the insufficient ones. The collected data is first trained through the CNN model of the convolutional window, and then the extracted feature information vector containing the timing relationship is output to the LSTM network for training. After training, a complete implicit state sequence is obtained, namely the vector containing the timing sequence feature information of the charging-discharge cycle. Because the CRF has a good effect on time-series prediction, the vector with time series feature information trained by LSTM is input into CRF, and the final prediction result is obtained by CRF. The overall framework of the model is shown in Fig. 1.

**CNN network.** The CNN module mainly uses the convolutional layer in the convolutional neural network to capture the local features of the data, and uses a variety of different convolutional cores to carry out the convolution operation. Then, the Maxpooling operation is used to further extract the most effective features of the local features, while reducing overfitting. Then, the vector of local features of battery test data containing time-series relationship obtained after convolution and pooling is fused to obtain more effective feature information. The CNN model established in this paper is shown in Fig. 2.
(1) Input layer: This layer is mainly used to receive the initial battery characteristic data. The feature data matrix $R$ is obtained by two-dimensional reconstruction of multi-feature timing series test data. As shown in Eq. (1), $R$ is connected to the CNN model as the input layer matrix.

$$R = m \sum_{i=1}^{n} x_{n} \quad m \in \{t, f\}$$  

where $m$ represents the dimension selected for construction, $t$ represents the time dimension, $f$ represents the characteristic dimension, and $x_{n}$ represents the battery data measured in the NTH charge–discharge cycle.

(2) Convolutional layer: This layer can use different sizes of convolution windows to perform convolution operations. The parameters of the convolutional neural network are stored in the weight matrix and the bias matrix. The initial value is randomly generated and changed through training. Due to the difference in the size of the convolution kernel, through the convolution operation, various forms of local features can be extracted, as shown in Eq. (2).

$$G = f(a \ast c) + b$$  

Among them, $a$ is the weight, $c$ is the convolution vector matrix to be calculated, $b$ is the bias, and $f$ selects the ReLU activation function.

For all neurons in the next layer, they are calculated by the convolution kernel of the previous layer, so they represent the characteristics of the neurons in the previous layer detected from different positions. Since multiple convolution kernels are used in the CNN module to calculate the feature mapping matrix of the next layer, multiple feature mapping matrices $G_{w}$ of the next layer are obtained, where $w$ represents the type of convolution window size used, that is, the final CNN integrated Number.

(3) Pooling layer: This layer validates the information extracted from the convolutional layer matrix through maxpooling operation to obtain multiple feature mapping matrices $P_{w}$; then the pooled multiple feature matrices are compressed into a feature matrix $P$, this process is called It is CNN integration. The integration formula is shown in formula (3). The dimensions of the compressed matrix rows are the same as the initial input $X$ matrix, but the data in this matrix can express more characteristic information.

$$P = \frac{1}{m} \sum_{i=1}^{m} P_{n}$$  

In this way, the input word vector is subjected to multi-layer convolution and pooling operations, then an eigenvector matrix containing the timing relationship is obtained, finally this matrix is used as the input of the next layer of LSTM model.

**LSTM network.** The second layer of the model is the LSTM layer, which is used to deal with timing features. The core of LSTM has a four-layer structure, which mainly contains three gates (forgetting gate, input gate, output gate) and a memory unit. The LSTM network model is shown in Fig. 3.

LSTM uses the forget gate to determine what information can pass through the state unit. The forget gate determines how much information can pass through at the previous time based on the output $h_{t-1}$ and the current input $x_{t}$ at the previous time. The calculation of $f_{t}$ is shown in Eq. (4).

$$f_{t} = \sigma(W_{f}x_{t} + U_{f}h_{t-1} + b_{f})$$  

Through the input gate to generate new information that needs to be updated. This step consists of two parts: the first part is to determine the value for updating $i_{t}$ obtained from the input gate; the second part is to use the Tanh layer to generate a new candidate value $\tilde{C}_{t}$, which is added as the candidate value generated by the current layer to the state unit. Then combine the values generated by the two parts to update. The calculations of $i_{t}$ and $\tilde{C}_{t}$ are as follows:

$$i_{t} = \sigma(W_{i}x_{t} + U_{i}h_{t-1} + b_{i})$$  

$$\tilde{C}_{t} = \tanh(W_{C}x_{t} + U_{C}h_{t-1} + b_{C})$$
Combine the forget gate and input gate, that is, discard redundant information and add new information:

\[ C_t = f_t C_{t-1} + i_t \tilde{C}_t \]  
(7)

where \( \sigma \) is the sigmoid activation function; \( \tanh \) represents the hyperbolic tangent activation function; \( W_f, W_i, W_o, U_f, U_i, U_o \) represent the weight matrix of input gate, forget gate, and output gate respectively; \( b_f, b_i, b_o \) represents the bias vector of the input gate, forget gate, and output gate; \( h_t \) represents the output at time \( t \).

The CRF network. In the prediction task, LSTM is good at processing long-term series of test data, but it cannot coordinate the dependence between adjacent results of time series data, especially in the face of battery capacity regeneration. CRF can obtain an optimal prediction result through the relationship of neighboring data, and make up for the shortcomings of LSTM. For any sequence \( X = (x_1, x_2, \cdots, x_n) \), assume that \( p \) is the output matrix of the LSTM, and the size of \( p \) is \( n \times k \), where \( n \) is the time series prediction step size, \( k \) is the measurement feature information, and \( p_{ij} \) represents the \( j \)th measurement of the \( i \)-th time point word feature. For the prediction sequence \( Y = (y_1, y_2, \cdots, y_n) \), the score function to get it is:

\[ s(X, Y) = \sum_{i=0}^{n} A_{y_i,y_{i+1}} + \sum_{i=1}^{n} p_{i,j} \]  
(9)

\( A \) represents the transition score matrix, \( A \) represents the score which the predicted value \( i \) is transferred to \( j \), and the probability of the predicted sequence \( Y \) is:

\[ p(Y|X) = \frac{e^{s(X, Y)}}{\sum_{\hat{Y} \in Y_X} s(X, \hat{Y})} \]  
(10)

Take the logarithm at both ends to get the likelihood function of the predicted sequence:

\[ \ln(p(Y|X)) = s(X, Y) - \ln(\sum_{\hat{Y} \in Y_X} s(X, \hat{Y})) \]  
(11)

In the formula, \( \hat{Y} \) represents the real labeling sequence, and \( Y_X \) represents all possible labeling sequences. The output sequence with the largest score after decoding:

\[ Y^* = \arg \max_{\hat{Y} \in Y_X} s(X, \hat{Y}) \]  
(12)

The CRF model is shown in Fig. 4.

**Figure 4.** Structure of CRF model.

**Experiment**

**Description of lithium-ion battery datasets.** The data used in the experiment came from the NASA PCOE lithium-ion battery data set. A set of four Li-ion batteries (B05, B06, B07, and B18) were run through 3 different operational profiles (charge, discharge and impedance) at room temperature. Charging was carried out in a constant current mode at 1.5A until the battery voltage reached 4.2 V and then continued in a constant voltage mode until the charge current dropped to 20 mA. Discharge was carried out at a constant current level of
2A until the battery voltage fell to 2.7 V, 2.5 V, 2.2 V, and 2.5 V for batteries B05, B06, B07, and B18 respectively. Impedance measurement was carried out through an electrochemical impedance spectroscopy frequency sweep from 0.1 Hz to 5 kHz. Repeated charge and discharge cycles result in accelerated aging of the batteries while impedance measurements provide insight into the internal battery parameters that change as aging progresses. This dataset can be used for the prediction of both the remaining charge and remaining useful life. From the Fig. 5 that the capacity of the battery is gradually decreasing as the charging and discharging cycle continues. The sudden increase of points in the Fig. 5 is due to the capacity regeneration effect.

The data set contains six statistical test features, namely Voltage_Measured, Current_Measured, Temperature_Measured, Current_Load, Voltage_Load and Time.

Datasets segmentation. Normalization of data can be used in machine learning algorithms to eliminate the negative effects of different value ranges and to improve the convergence speed and accuracy of the model. The method used in this study is min–max normalization, which scales the data to the interval \([0, 1]\) through a linear transformation:

\[
\text{Norm}(X_f) = \frac{(X_f - X_{\min})}{X_{\max} - X_{\min}} + \varepsilon
\]

(13)

where \(X_f\) is the all readings of sensor \(f\) on all units, \(\varepsilon\) denotes a positive number that tends to 0 infinitely, preventing the case where the denominator is 0.

Datasets segmentation. In order to verify the generalizability of the prediction results of this framework, three sets of data are randomly selected from four battery datasets as the training set and another set as the validation and test set. Figure 6 details the overall process of datasets partitioning.

Experimental configuration. The experimental code in this paper runs in Python 3.7 environment; the deep learning frameworks are Tensorflow 1.15.2 and Keras 2.2.4; the experiments are implemented on a PC (Windows 10 OS, Intel (R) Core (TM) I9-10900 KF CPU 3.7 GHz, 24 Gbytes of RAM. NVIDIA GeForce RTX 3090 GPU).

Parameter configuration. The correct choice of network model parameters often affects the prediction results. The experiment sets the parameters of step size of the predicted time series, the number of neurons in the network layer, learning rate and batch_size as hyperparameters, and the detailed data are shown in Table 2. The ReLU activation function is selected in the convolutional layer, the linear activation function is selected in the fully connected layer, and the marginal learning mode is selected in the CRF.

To obtain the hyperparameters suitable for the network model faster, particle swarm optimization (PSO) was experimentally chosen as the parameter optimization algorithm. PSO is a swarm intelligence algorithm for finding optimal parameters, which is often used in the parameter finding the process of network models in battery prediction problems49,50. PSO completes the search process by the individual search for optimal values and population information sharing, and the Fig. 7 shows the parameter optimization process of the particle swarm algorithm in detail.

In the experiment is divided into the following specific steps:
Parameter initialization. Set the number of particles \( n = 10 \), the particle size \( D \) as the number of parameters to be optimized \( 10 \), the learning factor of particle update \( c_1 = 1 \) and \( c_2 = 0.5 \), the number of iterations \( M = 100 \), and the inertia weight parameter \( w = 0.8 \). Randomly generate the initial velocity information \( v_{ij} \) and position information \( x_{ij} \) of the particle.

**Figure 6.** Partition of battery datasets.

| Parameter        | Range                                      |
|------------------|--------------------------------------------|
| \( \text{cnn}_1 \) neurons | 2–16                                       |
| \( \text{cnn}_2 \) kernel | 2–18                                       |
| \( \text{cnn}_2 \) neurons | 2–36                                       |
| MaxPooling neurons | 2–\( \min(\text{cnn}_1, \text{cnn}_2, \text{cnn}_2) \) |
| RNN neurons      | 6–200                                      |
| CRF neurons      | 6–200                                      |
| Dense neurons    | 6–200                                      |
| Learning rate    | 0.01–0.0001                                |
| Batch size       | 1–200                                      |

**Table 2.** Hyperparametric range.

**Figure 7.** Process structure of PSO.
(2) The mean square error of the prediction result is used as the objective function of the particles, and the calculation formula is as follows.

$$MSE = \frac{1}{N} \sum_{t=1}^{N} (x'_t - x_t)^2$$  \hspace{1cm} (14)

(3) Calculate and update the current optimal solution $p_i$ and the global optimal solution $p'_i$ obtained from the particle calculation of the current iteration number.

(4) Update the velocity and position information of the particle, and update the formula as follows:

$$v_{ij}(t+1) = w \cdot v_{ij}(t) + c_1 r_1 [p_i(t) - x_{ij}(t)] + c_2 r_2 [p'_i(t) - x_{ij}(t)]$$  \hspace{1cm} (15)

$$x_{ij}(t+1) = x_{ij}(t) + v_{ij}(t+1)$$ \hspace{1cm} (16)

where $r_1$ and $r_2$ is taken as a uniform random number in the range of $[0–1]$, so that the particle swarm algorithm has the ability to search randomly to avoid falling into local optimum.

(5) Judge whether the loop reaches the termination condition maximum number of iterations $M$, reaches the termination condition then end the optimization process to get the optimization result $p'_i$.

The results of the PSO optimization parameters are shown in Table 3.

### Evaluation metrics.

To quantify the forecast results for comparison and analysis. RMSE and MAPE are used to evaluate the performance of the model in this paper\(^5\). The calculation methods of each evaluation index are as follows:

$$RMSE = \sqrt{\frac{1}{N} \sum_{t=1}^{N} (predicted_t - observed_t)^2}$$  \hspace{1cm} (17)

$$MAPE = \sum_{t=1}^{N} \left| \frac{predicted_t - observed_t}{observed_t} \right| \times \frac{100}{N}$$ \hspace{1cm} (18)

In the formula, $N$ is the total number of measurements predicted by the model. The results were averaged over several experiments.

### Experimental results and discussion

#### Results of time-linked module control experiment.

To test the effect of different temporal association modules on the prediction accuracy in CRF models, control experiments of LSTM, GRU, and BiLSTM were designed. The temporal association module compares the difference between using a single-layer CNN network to extract feature information as network input and using a two-layer CNN network. The output of the temporal association module is used as the input of the CRF model, and the experiments are done on B18, and the experimental results are shown in Table 4.

The experimental results surface that the prediction error of LSTM is smaller compared with other RNNs, which indicates that the temporal information extracted by LSTM is more adapted to the input with the CRF network. Compared with single-layer CNN networks, two-layer CNN networks can obtain better results. We found that this is because the two-layer CNN changes the length of the input to the network model using the pooling layer compared to the single-layer CNN, which enables the second layer CNN to extract a wider range of feature information.

| Parameter       | Range |
|-----------------|-------|
| cnn_1 neurons   | 9     |
| cnn_kernel      | 17    |
| cnn_2 neurons   | 6     |
| MaxPooling neurons | 5   |
| RNN neurons     | 192   |
| CRF neurons     | 76    |
| Dense neurons   | 199   |
| Learning rate   | 0.0005|
| Batch size      | 2     |

Table 3. Hyperparameter optimization results.
Results of CRF ablation experiments. To prove the effect of CRF on the model, the effects of adding CRF and not adding CRF on the CNN-LSTM-CRF predicted results were compared. The comparison results are shown in Table 5.

The experimental results found that CRF could improve the accuracy of network model prediction, and the RMSE and MAPE evaluation metrics on four datasets B05, B06, B07, and B18 improved by more than 20% on average compared with no CRF, with the MAPE metric of B06 dataset improving by 53% as the largest improvement of the experiment, which indicated the importance of CRF model, which was also proved on the subsequent experiments of the probability distribution of prediction results.

Results of capacity prediction. To intuitively reflect the prediction results of this method, Fig. 8 shows in detail the original measurement capacity and model prediction capacity of battery data sets B05, B06, B07, and B18(threshold value of precision region α is ± 2.5%).

The results from the figure show that most of the predicted results are within the error range of the true capacity. Of course, the predicted values at 78 cycles on the B05, B06, and B07 data sets show poorer prediction results, which is due to the capacity rebound caused by the capacity regeneration phenomenon during the charging and discharging process of Li-ion batteries, and the sudden change in capacity will bring a larger prediction error compared to the smooth state. This is because we use the complete battery dataset for model training to predict a different battery dataset, and the CRF model uses the feature offset matrix during the training process to count the overall trend of the training data and record the overall trend of the battery capacity so that when there is a large error, the error will be reduced in the subsequent prediction process according to the learning record. The error is reduced in the subsequent prediction process based on the learning records.

Comparison of the previous model. In order to verify the prediction superiority of the CNN-LSTM-CRF model proposed in this paper, comparative experiments were conducted with SVM, LSTM, and GRU models. The RMSE and MAPE results of the four models are compared in Table 6, which can more intuitively show the prediction accuracy of this algorithm.

It can be found from the table that for RMSE and MAPE metrics, the average metrics of the predicted value of the model in this paper are superior to the comparison model, illustrating the feasibility of the CNN-LSTM-CRF model proposed in this paper in the battery capacity prediction problem.

Conclusion

For the problem of lithium battery capacity prediction, this paper takes inspiration from the field of NLP and proposes a combined CNN-LSTM-CRF neural network prediction model, which is applied to the battery remaining life prediction for the first time. The model inputs continuous-time battery measurement data and predicts the output battery capacity situation at the current time point to obtain the remaining battery life at this time. Compared with the previous battery capacity prediction network model, the major difference in this model is the inclusion of CRF. The capacity prediction sequence is calculated by the offset matrix of the feature data, which more intuitively reflects the change of the decreasing trend of capacity. The CNN convolutional module is added to the model to collect the feature data, and the time-linked module captures the trend of feature data in the time dimension to extract the temporal information. Among them, LSTM achieves better results in the time-linked module control experiments. The ablation experiments demonstrate the effectiveness of the CRF network in the capacity prediction process. By comparing with previous models, our model achieves better prediction results.

Table 4. CNN and RNN compared the experimental prediction results.

| CNN   | RNN   | RMSE   | MAPE   |
|-------|-------|--------|--------|
| Single layer | LSTM  | 0.0367 | 0.0264 |
|       | GRU   | 0.0589 | 0.0352 |
|       | BiLSTM| 0.0496 | 0.0314 |
| Double layer | LSTM  | 0.0248 | 0.0141 |
|       | GRU   | 0.0316 | 0.0236 |
|       | BiLSTM| 0.0421 | 0.0272 |

Table 5. CRF comparison experiment predicted the results.

| Data | Cycle | CRF- | Without CRF |
|------|-------|------|-------------|
|      |       | RMSE | MAPE        | RMSE | MAPE |
| B05  | 168   | 0.0216 | 0.0161 | 0.0316 | 0.0253 |
| B06  | 168   | 0.0179 | 0.0081 | 0.0275 | 0.0152 |
| B07  | 168   | 0.0257 | 0.0193 | 0.0357 | 0.0284 |
| B18  | 132   | 0.0316 | 0.0230 | 0.0411 | 0.0325 |
Our model still has flaws. The large number of network structures combined makes the network depth and computation of the model huge, which will cost more computational resources and time. Future work can try to experiment with migration learning in the model learning process, and use the extracted trained network parameters to adjust the network model to make the real-time prediction of the model possible.
Data availability
NASA PCOE lithium-ion battery data used to support this study are available at https://ti.arc.nasa.gov/tech/ dash/groups/pcoe/prognostic-data-repository.

Received: 7 March 2022; Accepted: 26 July 2022
Published online: 02 August 2022

References
1. Zhang, Q. & White, R. E. Capacity fade analysis of a lithium ion cell. J. Power Sources 179(2), 793–798 (2008).
2. Wright, R. et al. Power fade and capacity fade resulting from cycle-life testing of advanced technology development program lithium-ion batteries. J. Power Sources 119, 865–869 (2003).
3. Ramadesigan, V. et al. Parameter estimation and capacity fade analysis of lithium-ion batteries using reformulated models. J. Electrochem. Soc. 158(9), A1048–A1054 (2011).
4. Cordoba-Arenas, A., Onori, S., Guenezenc, Y. & Rizzoni, G. Capacity and power fade cycle-life model for plug-in hybrid electric vehicle lithium-ion battery cells containing blended spinel and layered-oxide positive electrodes. J. Power Sources 278, 473–483 (2015).
5. Christensen, J. & Newman, J. A mathematical model for the lithium-ion negative electrode solid electrolyte interphase. J. Electrochem. Soc. 151(11), A1977–A1988 (2004).
6. Pinson, M. B. & Bazant, M. Z. Theory of SEI formation in rechargeable batteries: Capacity fade, accelerated aging and lifetime prediction. J. Electrochem. Soc. 160(2), A243–A250 (2013).
7. Wu, L., Fu, X. & Guan, Y. Review of the remaining useful life prognostics of vehicle lithium-ion batteries using data-driven methodologies. Appl. Sci. 6(6), 166 (2016).
8. Farman, A. & Sauer, D. U. Comparative study of reduced order equivalent circuit models for on-board state-of-available-power prediction of lithium-ion batteries in electric vehicles. Appl. Energy 225, 1102–1122 (2018).
9. Lai, X., Wang, S., Ma, S., Xie, J. & Zheng, X. Parameter sensitivity analysis and simplification of equivalent circuit model for the state of charge of lithium-ion batteries. ScienceDirect. Electrochem. Soc. 330, 135329 (2020).
10. Wei, H., Chen, X., Lü, Z., Wang, Z. & Chen, L. Online estimation of lithium-ion battery state of health using grey neural network. Sci. Rep. 12, 13221 (2022).
11. Lai, X., Wang, S., Ma, S., Xie, J. & Zheng, X. Parameter sensitivity analysis and simplification of equivalent circuit model for the state of charge of lithium-ion batteries. ScienceDirect. Electrochem. Soc. 330, 135329 (2020).
12. Barcellona, S. & Piegari, L. Lithium ion battery models and parameter identification techniques. Energies 2017, 10 (2007).
13. Figueroa-Santos, M. A., Siegel, J. B. & Stefanopoulou, A. G. Leveraging cell expansion sensing in state of charge estimation: Practi-
14. Bruch, M. et al. Novel method for the parameterization of a reliable equivalent circuit model for the precise simulation of a battery cell's electric behavior. J. Power Sources 490, 229513 (2021).
15. Wang, Y., Tian, J., Sun, Z., Wang, L. & Chen, Z. A comprehensive review of battery modeling and state estimation approaches for advanced battery management systems. Renew. Sustain. Energy Rev. 131, 110015 (2020).
16. Zhang, Q., Wang, D., Yang, B., Cui, X. & Li, X. Electrochemical model of lithium-ion battery for wide frequency range applications. Electrochimica Acta 343, 136994 (2020).
17. Ashwin, T. R., Chung, Y. M. & Wang, J. Capacity fade modelling of lithium-ion battery under cyclic loading conditions. J. Power Sources 328, 586–598 (2016).
18. Wang, S., Han, X. & Chen, L. Particle filtering for lithium-ion battery residual life prediction. Chin. J. Power Sources 44, 48–53 (2020).
19. Vidal, C., Malysz, P., Kollmeyer, P. & Emadi, A. Machine learning applied to electrified vehicle battery state of charge and state of health estimation: State-of-the-art. IEEE Access 8, 52796–52814 (2020).
20. Wei, Z., Zhao, J., Ji, D. & Tseng, K. J. A multi-timescale estimator for battery state of charge and capacity dual estimation based on an online identified model. Appl. Energy 204, 1264–1274 (2017).
21. Vichard, L., Ravey, A., Venet, P., Harel, F. & Hissel, D. A method to estimate battery soh indicators based on vehicle operating data only. Energy 225, 120235 (2021).
22. Wei, H. E., Williard, N., & Osterman, M. et al. Prognostics of lithium-ion batteries using extended Kalman filtering. in IMAPS Advanced Technology Workshop on High Reliability Microelectronics for Military Applications. 17–19. (Linthicum Heights: IMAPS, 2011).
23. Feng, X. et al. Online state-of-health estimation for li-ion battery using partial charging segment based on support vector machine. IEEE Trans. Veh. Technol. 68, 8583–8592 (2019).
24. Qin, T., Zeng, S., Guo, J. & Skaf, Z. A rest time based prognostic framework for state of health estimation of lithium-ion batteries with regeneration phenomena. Energies 9, 896 (2016).
25. Zhang, Z., Huang, M., Chen, Y. & Zhu, S. Prediction of lithium-ion battery’s remaining useful life based on relevance vector machine. SAE Int. J. Altern. Powertrains 5, 30–40 (2016).
26. Gao, D. & Huang, M. Prediction of remaining useful life of lithium-ion battery on multi-kernel support vector machine with particle swarm optimization. J. Power Electron. 17, 1288–1297 (2017).
27. Khunprom, P. & Yodo, N. A Data-Driven Predictive Prognostic Model for Lithium-Ion Batteries based on a Deep Learning Algorithm. Energies 12, 660 (2019).
28. Sterkens, W., Dillam, D. R., Goedemé, T., Dewulf, W. & Jef, R. P. Detection and recognition of batteries on X-ray images of waste vehicles. Energy Convers. Manage. 64, 113–112 (2012).
29. Zhang, Z., Huang, M., Chen, Y. & Zhu, S. Prediction of lithium-ion battery’s remaining useful life based on relevance vector machine. SAE Int. J. Altern. Powertrains 5, 30–40 (2016).
30. Fan, Y., Xiao, F., Li, C., Yang, G. & Tang, X. A novel deep learning framework for state of health estimation of lithium-ion battery. J. Energy Storage 32, 101741 (2020).
31. Zhou, D., Li, Z., Zhu, J., Zhang, H. & Hou, L. State of health monitoring and remaining useful life prediction of lithium-ion batteries based on temporal convolutional network. IEEE Access 8, 53307–53320 (2020).
32. Zhiquan, L. I. U. et al. A hybrid LSSVR/HMM-based prognostic approach. Sensors 13(5), 5542–5560 (2013).
33. Hong, S., Yue, T. & Liu, H. Vehicle energy system active defense: A health assessment of lithium-ion batteries. Int. J. Intel. Syst. 1, 1–19 (2020).
34. Fang, L., Li, J. & Peng, B. Online estimation and error analysis of both SOC and SOH of lithium-ion battery based on DEKF method. Energy Proc. 158, 3008–3013 (2019).
37. Wang, D. & Tsui, K. L. State Space Models Based Prognostic Methods for Remaining Useful Life Prediction of Rechargeable Batteries 307–334 (Springer, 2017).
38. Zhang, H., Miao, Q., Zhang, X. & Liu, Z. An improved unscented particle filter approach for lithium-ion battery remaining useful life prediction. Microelectron. Rel. 81, 288–298 (2018).
39. Song, Z., Gao, J., Pan, L. & Xi, J. Lithium-ion battery health status prediction based on principal component analysis and improved support vector machine. Automot. Technol. 11, 21–27 (2020).
40. Ren, L., Dong, J., Wang, X., Meng, Z. & Zhao, L. A data-driven auto-CNN-LSTM prediction model for lithium-ion battery remain useful life. IEEE Trans. Ind. Inform. 17, 3478–3487 (2020).
41. Ren, L. et al. Remaining useful life prediction for lithium-ion battery: A deep learning approach. IEEE Access 6, 50587–50598 (2018).
42. Kodjo, S. R. et al. State-of-health estimators coupled to a random forest approach for lithium-ion battery aging factor ranking. J. Power Sources 25, 229154 (2020).
43. Deng, Y. et al. Feature parameter extraction and intelligent estimation of the state-of-health of lithium-ion batteries. Energy 176, 91–102 (2019).
44. Hong, S. & Zeng, Y. A health assessment framework of lithium-ion batteries for cyber defense. Appl. Soft Comput. 101, 107067 (2021).
45. Kim, J. et al. Parameter identification of lithium-ion battery pseudo-2-dimensional models using genetic algorithm and neural network cooperative optimization. J. Energy Storage 45, 103571 (2022).
46. Li, M. et al. Automated data function extraction from textual requirements by leveraging semi-supervised CRF and language model. Inf. Softw. Technol. 143, 106770 (2022).
47. An, Y. et al. Chinese clinical named entity recognition via multi-head self-attention based BiLSTM-CRF. Artif. Intell. Med. 127, 102282 (2022).
48. Saha, B. & Göbel, K. Battery Data Set, NASA AMES Prognostics Data Repository. http://it.arc.nasa.gov/project/prognostic-data-repository. (NASA AMES, 2007).
49. Özşar, Ş et al. Adaptation of metaheuristic algorithms to improve training performance of anESZSL model. Turk. J. Electr. Eng. Comput. Sci. 29(3), 1781–1796 (2021).
50. Ren, X. et al. A method for state-of-charge estimation of lithium-ion batteries based on PSO-LSTM. Energy 234, 121236 (2021).
51. Zheng, S., Ristovski, K., Farahat, A. et al. Long short-term memory network for remaining useful life estimation. in IEEE International Conference on Prognostics and Health Management. 88–95. (2017).

Acknowledgements
This study was supported by the Scientific and Technological Research Program of Chongqing Municipal Education Commission (KJQN202001142), the Chongqing Research Program of Basic Research and Frontier Technology (Grant No. cstc2020jcyj-msxmX0352), the fellowship of China Postdoctoral Science Foundation (2021M700616), and the Chongqing University of Technology (2019ZD118).

Author contributions
Hai-Kun Wang: Conceptualization, Supervision. Yang Zhang: Methodology, Software, Data curation, Writing-Original draft preparation. Mohong Huang: Visualization and Reviewing.

Competing interests
The authors declare no competing interests.

Additional information
Correspondence and requests for materials should be addressed to H.-K.W.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2022