

Development of a Lithium-Ion Battery Lifetime Prediction Model Using Deep Learning for Short-Term Learning

Yu Fujitaki and Hiroyuki Kobayashi
Osaka Institute of Technology, Osaka, Japan

Keywords: Lithium-Ion Batteries, Deep Learning, Cnn, Prediction.

Abstract: We will use the open data utilized in Severson's research. This data consists of cycle data obtained from repeated charging and discharging of lithium-ion batteries, which will be analyzed. One issue is that the amount of cycle data is limited, which could lead to inadequate training. To address this problem, we have adopted a method that extracts multiple data points from a single battery dataset, thereby improving prediction accuracy. In this experiment, we compared data from 100 charge-discharge cycles with data from just 1 charge-discharge cycle.

1 INTRODUCTION

I am surrounded by a wide range of portable electronic devices, such as smartphones, all of which depend on lithium-ion batteries (LIBs). LIBs have been in use for about 25 years and are now a common component in our daily lives due to their high energy and power density, long lifespan, cost-effectiveness, and reliable safety compared to other commercially available batteries. Despite these advantages, the frequency of accidents involving LIBs has been increasing in recent years. To ensure their safe and reliable use in electric vehicles and other devices, it is essential to monitor various parameters including voltage, temperature, state of charge (SOC), state of health (SOH), remaining capacity, and cycle life. While parameters like voltage and temperature can be directly measured with sensors, others such as SOC and SOH require estimation through algorithms based on measurement characteristics. Predicting battery cycle life is crucial, but traditional methods for doing so are complex, relying on physics-based models and needing to account for a wide range of operating conditions and device variability, even among batteries from the same manufacturer. Recently, there has been increasing interest in using machine learning-based methods to predict battery behavior. These methods offer the potential for accurate early predictions of battery cycle life, which not only speeds up the validation of new manufacturing processes but also helps end-users detect performance degradation and replace failing batteries in time. (Schmush, 2018).

2 OBJECTIVE

In the prediction of the lifespan of lithium-ion batteries (LIBs), methods include not only simple empirical approaches but also physical models that hypothesize degradation phenomena and numerically solve electrochemical reaction equations. More recently, data-driven models that use machine learning to predict lifespan based on charge-discharge cycle data have emerged. The data-driven method that I am particularly focused on has been highly regarded for its ability to estimate LIB capacity, remaining life, and cycle life. However, this method is based on limited empirical or mechanical test data and does not take degradation modes into account.

In Severson's study, they predicted LIB cycle life using data from the first 100 cycles, during which degradation is minimal, achieving a very low test error of 9.1%. This research demonstrates the potential of data-driven approaches in predicting the behavior of complex nonlinear systems. However, a limitation of this method is that only one piece of data is generated per battery, leading to a shortage of data. Additionally, when predicting the performance of a new battery, it requires 100 charge-discharge cycles before a prediction can be made, which is another issue.

Therefore, my goal is to develop a model that can achieve more accurate predictions with fewer charge-discharge cycles by training on single-cycle data and extracting multiple data points from a single battery, rather than relying on multiple cycles. (Severson, 2019)

3 DATA SPLITTING METHOD

First, the data obtained from charging and discharging a single battery includes values such as voltage during charging and discharging, as well as discharge capacity. This data is recorded over the course of hundreds of charge-discharge cycles. Data collection continues until the battery's maximum capacity falls below 80% of its initial value, indicating significant degradation. The data obtained in this way is often used for predicting battery life and evaluating performance. By analyzing the changes in voltage and capacity for each cycle, degradation trends can be identified.

However, instead of using this dataset as-is, employing a data splitting method is crucial for efficient analysis. By appropriately splitting the large amount of data obtained from a single battery, we can increase the diversity of data used for training machine learning models, resulting in more accurate predictions.

Next, the data is divided by each cycle. For example, when using cycle data from a battery with a lifespan of 800 cycles, the data from the first cycle is used as the training data (Figure 1). At this time, the corresponding target data for the first cycle is "800." This means that the battery is still capable of 800 more charge-discharge cycles.

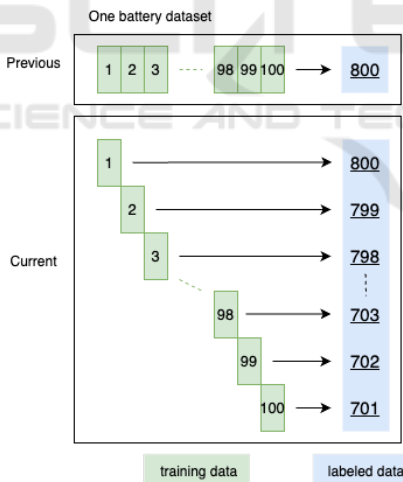


Figure 1: Data Splitting Method.

Similarly, when using the data from the second cycle as training data, the corresponding target data is "799." This indicates that at the second cycle, the battery has 799 charge-discharge cycles remaining. In this way, for each cycle's data, the remaining number of charge-discharge cycles (i.e., the remaining cycle life) is set as the target data, and the model is trained accordingly.

By using this method, the data obtained from a single battery can be split into multiple training

datasets, increasing the overall amount of data. As a result, the accuracy of the machine learning model is expected to improve.

4 LEARNING MODELS

In this study, we utilized a one-dimensional convolutional neural network (1D-CNN) to analyze battery charge-discharge cycle data. The 1D-CNN is particularly effective for processing time-series data due to its ability to perform convolutional operations along a single dimension. (Figure 2) This characteristic makes it especially suitable for handling sequential data, such as the voltage, current, temperature, and capacity measurements taken during battery cycles. Unlike traditional two-dimensional convolutional neural networks (2D-CNN), which are designed for spatial data, the 1D-CNN focuses on extracting features from a linear sequence of data points (Pelletier et al., 2019).

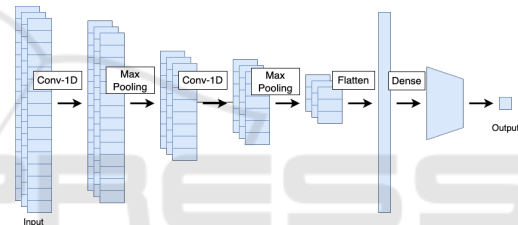


Figure 2: 1D-CNN.

The dataset for each cycle is represented as a two-dimensional array, where one dimension corresponds to the time steps within the cycle, and the other dimension represents various features related to the battery's performance. By applying convolutional operations across the feature dimension, the 1D-CNN effectively captures and learns significant patterns that are indicative of battery degradation and remaining lifespan. This approach allows the model to identify and understand the temporal dynamics and trends within the charge-discharge cycles. Table 1 lists the parameters of the 1D-CNN model used in this study.

Table 1: Model Architecture and Parameters.

Layer (type)	Param #
Conv1D (64 filters, kernel_size=50)	19,264
MaxPooling1D (pool_size=2)	0
Conv1D (64 filters, kernel_size=50)	204,864
MaxPooling1D (pool_size=2)	0
Flatten	0
Dense (512 units, ReLU)	918,016
Dense (1 unit, Linear)	513

CNNs, in general, are renowned for their ability to learn hierarchical features from data. They excel in recognizing patterns through convolutional layers, which apply filters to the input data to detect local features, followed by pooling layers that reduce the dimensionality while retaining critical information. Fully connected layers then integrate these features to perform the final predictive tasks. Although CNNs are widely used in image processing, their adaptation to one-dimensional data sequences proves valuable for predicting battery performance, as they can capture temporal relationships and nuances in the data (Simonyan and Zisserman, 2014).

In summary, by employing a 1D-CNN, our research aims to leverage the model’s capability to process and analyze time-series data from battery cycles, thereby enhancing the accuracy of predictions related to battery degradation and lifespan. This methodology offers a promising avenue for improving the reliability and efficiency of battery performance assessments.

5 EXPERIMENT

In this study, we used the dataset from the "Data-driven prediction of battery cycle life before capacity degradation" (Schmush, 2018). This dataset consists of data from 136 commercial lithium-ion batteries that were cycled between 150 and 2,300 times under 72 different fast-charging conditions, totaling approximately 96,700 cycles. The dataset includes 15 variables, such as voltage, charge capacity, discharge capacity, charge energy, discharge energy, and internal battery temperature (Table 2). The batteries are

Table 2: Data set variables.

Data Point	Test Time
Date Time	Step Time
Step Index	Cycle Index
Current	Voltage
Charge Capacity	Charge Energy
Discharge Capacity	Discharge Energy
dV/dt	Temperature
fff Internal Resistance	

lithium iron phosphate (LFP)/graphite cells manufactured by A123 Systems (APR18650M1A), cycled using a 48-channel Alvin LBT potentiometer in a forced convection temperature chamber set at 30°C. The purpose of this study was to optimize fast charging. All cells were charged using a one-step or two-step fast charging policy in the format of "C1(Q1)-C2." Charging was conducted at 1C CC-CV up to 80% SOC,

with upper and lower cutoff voltages set at 3.6V and 2.0V, respectively. All cells were discharged at 4C.

The dataset is divided into three batches, each based on the "batch date" and containing approximately 48 cells per batch. Temperature was measured using a T-type thermocouple; however, it is important to note that measurement accuracy varied, and contact may have been lost during cycling. Internal resistance was measured during charging at 80% SOC using ±3.6C pulses.

Additionally, the data used in this study was obtained from the publicly available Severson dataset, which contains data from 127 batteries with lifespans ranging from 450 to 1,325 cycles. The dataset includes six attributes: voltage, charge capacity, discharge capacity, charge energy, discharge energy, and internal battery temperature. For training purposes, data from the first 100 cycles of each battery was used, yielding a total of 12,700 data points (100 cycles × 127 batteries).

6 COMPARISON

To evaluate the proposed method, we compared it with the results of a previous study on battery life prediction. In that previous research, one data point was obtained from each battery using data from cycles 1 to 100. The learning model employed was a CNN, as in the current study.

In the proposed method, a 1D CNN was used, but when training on data from cycles 1 to 100, the convolutional operations on the features remained unchanged. However, since the current data consists of two-dimensional data with time steps and 100 cycles, a 2D CNN was utilized. In the next section, we will compare the prediction accuracy and data volume between the previous study and the proposed method.

7 RESULTS

For training data, 84 batteries were used, with 100 cycles of data collected from each battery, resulting in a total of 8,400 data points. The remaining 43 batteries were used for validation data, yielding a total of 4,300 data points from 100 cycles of data each. The current study (Figure 3) and the previous study (Figure 4) show graphs with the actual measured life cycles on the x-axis and the predicted life cycles on the y-axis. The closer the results are to the red line in the graph, the better. However, due to the difference in the amount of data between this study and the previous one, it may be difficult to discern from the graph.

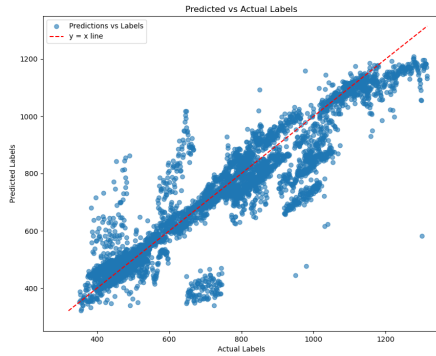


Figure 3: Current.

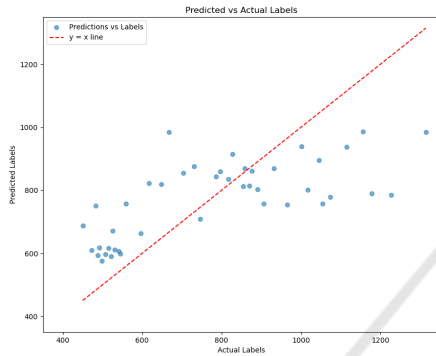


Figure 4: Previous.

To evaluate the battery life prediction results, we used the Mean Absolute Percentage Error (MAPE).

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| \times 100 \quad (1)$$

MAPE is an effective metric for assessing prediction accuracy, as it expresses the error between actual values and predicted values as a percentage, allowing us to understand the overall error trend. The results of the proposed method (table 3) are compared with the predictions from the previous research described earlier (table 4).

Table 3: Current.

Model	1D-CNN
Number of training samples	8400
Number of charge-discharge cycles	1
Number of features	6
MAPE	8.43%

By examining both the tables, it is evident that the proposed method in this study provides better accuracy compared to the previous approach.

Table 4: Previous.

Model	2D-CNN
Number of training samples	84
Number of charge-discharge cycles	100
Number of features	6
MAPE	17.40%

8 CONCLUSION

The results confirm that the proposed method of extracting multiple data points from a single battery cycle effectively increases the amount of training data. This approach also enables high-accuracy predictions for the lifespan of new batteries, even with a limited number of charge-discharge cycles. Furthermore, the improved prediction accuracy compared to conventional methods suggests that this technique could be highly useful in practical battery management and lifespan prediction. However, despite the improvement in accuracy, there is still a noticeable gap between the predicted and actual battery life, as shown in the graphs.

9 PERSPECTIVES

In future research, we will focus on developing a hybrid model that combines data-driven approaches with physics-based models that numerically solve electrochemical reaction equations, taking battery degradation mechanisms into account. In this study, we recognized the limitations of using only a data-driven approach, which led us to consider integrating a model-based approach. This hybrid model will incorporate predictions from the physics-based model as auxiliary data, feeding them back into the learning process of the machine learning model. Since many studies on model-based lifetime prediction estimate changes with each charge-discharge cycle, we are exploring methods that align with the single-cycle charge-discharge data proposed in this study. By doing so, we aim to leverage the strengths of both the physical and machine learning models to achieve greater accuracy and broaden the application of battery lifetime prediction.

REFERENCES

Pelletier, C., Webb, G. I., and Petitjean, F. (2019). Temporal convolutional neural network for the classifica-

tion of satellite image time series. *Remote Sensing*, 11(5):523.

Schmush, Richard, e. a. (2018). Performance and cost of materials for lithiumbased rechargeable automotive batteries. *Nature Energy* 3.4, page 267.

Severson, K. (2019). Data-driven prediction of battery cycle life before capacity degradation. *Nature Energy* 4.5, pages 383–391.

Simonyan, K. and Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*.

