Journal of Energy Storage 59 (2023) 106346

Contents lists available at ScienceDirect

Journal of Energy Storage

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

Research papers

Machine learning techniques suitability to estimate the retained capacity in lithium-ion batteries from partial charge/discharge curves

Hector Beltran [a,*], Emilio Sansano [a], Michael Pecht [b]

[a] Department of Industrial Engineering Systems and Design, Universitat Jaume I, Castelló de la Plana 12003, Spain
[b] Center for Advanced Life Cycle Engineering (CALCE), University of Maryland, College Park, MD 20742, USA

ARTICLE INFO

Keywords:
lithium-ion batteries
State-of-health
Deep learning
Convolutional neural networks

ABSTRACT

The accurate estimation of the retained capacity in a lithium-ion battery is an essential requirement for the electric vehicles. The aging of the batteries depends on parameters and factors that are not easily monitored by the battery management system. This paper analyzes the ability of various machine learning algorithms to deal with the data generated by the battery management system during the partial charging/discharging process to instantly diagnose and estimate the retained capacity of the battery. Experimental data from an online dataset containing thousands of battery cycles are used for training and validation of the different models. Results demonstrate that the developed convolutional neural network outperforms the rest of the machine learning algorithms implemented, regardless of the portion of the cycle registered by the battery management system. The estimates obtained outperform most previous references. However, the estimation error values registered when analyzing partial cycles with depths lower than 50 % (above 1.5 %) remain too high to validate any of the analyzed algorithms as a solution for commercial systems.

1. Introduction

The consolidation of lithium-ion batteries (LIBs) as a mature technology during the last decade has revolutionized our everyday lives, laying the foundation for a potentially fossil-fuel-free society based on renewable energies and electric mobility [1]. However, with the electric vehicle (EV) market accounting for almost 80 % of current LIB demand and a global total market of around 200 GWh in 2021 [2], LIBs have yet to reach their full potential as sources of global energy. Analysts project that between 2022 and 2030, the global demand for LIBs will increase tenfold, reaching over 3 terawatt-hours in 2030 [3]. Such a huge industry increase will stress the associated supply chain and the mining involved. In fact, serious concerns about the limited reserves available worldwide for some of the metals used in these batteries [4] suggest the sector's future hinges on developments in LIB recycling, battery life extension, and potential second life applications.

For both the extension of a battery's lifetime and its usefulness in second life applications, an estimation of its current state-of-health (SOH) with an accuracy below 1 % in terms of both mean absolute error (MAE) and root mean square error (RMSE) becomes very important at any time during service life [5]. This remains a persistent challenge—one to which the scientific community is devoting significant efforts [6]. LIBs degrade with time as a function of different stress factors, particularly temperature and usage pattern. Temperature and state-of-charge (SOC) affect LIBs when they are at rest (calendar aging) as well as when they are under operation (cycle aging) [7]. However, while cycle aging is associated with mechanical strain in the electrode active materials and lithium plating, the main causes of calendar aging are parasitic reactions between electrolyte and electrodes, such as solid electrolyte interface (SEI) creation and growth, electrolyte oxidation, and transition metals dissolution. All these mechanisms reduce the battery's capacity over time [8].

To determine the SOH of a LIB or to analyze its retained capacity, various methodologies have been proposed in the literature [9–12], including machine learning (ML) solutions [13–15]. These data-driven estimation options are advantageous because they do not require the deep understanding of the electrochemical reactions and their modeling, since they utilize only the degradation patterns registered in the data [16].

For instance, Roman et al. [17] designed and evaluated an ML pipeline to estimate the battery capacity fade on 179 cells cycled under various conditions based on segments of the charge voltage and current

* Corresponding author.
E-mail address: hbeltran@uji.es (H. Beltran).

https://doi.org/10.1016/j.est.2022.106346
Received 14 September 2022; Received in revised form 23 November 2022; Accepted 5 December 2022
Available online 23 December 2022

curves and using two parametric and two nonparametric algorithms: a Bayesian ridge regression (BRR), a Gaussian process regression (GPR), a random forest (RF), and a deep ensemble of neural networks (NNs). Their best model achieved a root mean squared percent error of 0.45 %.

Paulson et al. [18] analyzed 300 pouch cells to introduce an in-depth investigation of strategies for feature selection in battery lifetime prediction, ML models' generalization across multiple battery chemistries (including lasso regression, RF, GPR, and artificial NN regression), and predictions beyond the training set in the chemical space.

Yang et al. [19] introduced a three-layer back propagation neural network (BPNN) used to estimate the SOH based on the parameters of a first-order equivalent circuit model (ECM) of the cells under study. They obtained estimation errors around 5 %.

Fan et al. [20] proposed a hybrid NN called gate recurrent unit-convolutional neural network (GRU-CNN) to estimate the SOH from the observed voltage, current, and temperature curves during charging. The approach provided a maximum estimation error limited to within 4.3 %.

Another hybrid NN concatenating a one-dimensional convolutional NN (CNN) and active-state-tracking long–short-term memory (LSTM) NN is proposed in [21]. By using a Bayesian optimization algorithm to build the SOH model, authors achieve a global average RMSE of 0.0269 in the SOH estimation.

Kong et al. [22] developed a similar framework-combined deep CNN with a double-layer LSTM NN and obtained very accurate results, with RMSE as low as 0.0061. Still, Pradyumna et al. [23] present a method to estimate capacity using impedance curves obtained from an electrochemical impedance spectroscopy (EIS) test and a convolutional neural network (CNN). They claim RMS errors to be 0.233 % in capacity estimation.

Ungurean et al. [24] introduced deep learning techniques based on recurrent neural networks (RNNs) with memory, such as the LSTM and gated recurrent unit (GRU). The solutions implemented provided errors in the range between 5.5 % and 2 %. Also, an encoder-decoder model based on deep learning and using the battery charging curves were introduced in [25], where the proposed model adapted well to different types of batteries, was able to adapt to various sampling modes of charging curves and presented high estimation accuracies with MAE and RMSE values below 1 %.

Along the same lines to the previous references, this paper focuses on the estimation of the retained capacity of the LIBs by means of different ML algorithms and analyzes the suitability of the different algorithms developed to be accurate enough to be implemented in commercial solutions. The six ML algorithms developed and analyzed —a support vector regression (SVR), a K-nearest neighbors (KNN), an RF, a gradient boosting decision tree (XGB), a deep neural fully connected network (FCN), and a CNN— are trained and validated with data measured during the partial cycling (charging/discharging processes) of the cells, highlighting the use of portions of the incremental capacity (dQ/dV) curve obtained from the voltage profiles and for two different C-rates. This combination of factors represents a challenge that has not been addressed in the literature very often.

Several previous studies pursued a similar goal. Yang et al. [26] used a combination of a CNN and an RF model to extract changes in SOH between two successive charge/discharge partial cycles. They obtained accuracies in terms of MAE in the range between 0.62 % and 1.51 %, outperforming other models introduced in the same paper for comparison. Also, Wei et al. [27] proposes a multistage artificial NN method to estimate the SOH in practical scenarios including heavily partial charging of the batteries and with different initial charging voltages. The MAE results they achieve are confined in the range from 0.3 % to 2.54 % for NCA cells, according to their conclusions. Finally, a very similar proposal to the one presented in this paper is found in [28], where the authors estimated the SOH of the same LIB cells for segments of their charging curves. However, differences can be found with these works since the approach in this paper makes use of the dQ/dV curves, generated during both full and partial cycles, and analyzes the influence of the ΔV value selected to discretize these curves. Moreover, the data preparation performed in this work, together with the use of different sources of information supplied to the ML algorithms (such as the dQ/dV curve and the C-rate parameter during both the charge and discharge profiles), allowed us to significantly outperform the results in [28] and obtain them in close competence with those in [26] (achieved for a different dataset).

The rest of this paper is organized as follows. Section 2 is devoted to introducing the cells used for the analysis and the dataset preparation. Section 3 introduces the principles of each ML algorithm analyzed as well as the specific model implemented for SOH estimation. Section 4 presents the results and compares the performance of the various ML algorithms for both full cycles and partial cycles. Finally, Section 5 provides some concluding remarks.

2. Data preparation

2.1. Lithium-ion cells under analysis

The dataset used in this study is developed from the information available at the online database provided by the Center of Advanced Life Cycle Engineering (CALCE) of the University of Maryland [29]. Among the applicable cells, this study employs the experimental results uploaded for six different $LiCoO_2$ prismatic samples labeled as CS2 LIBs, according to the nomenclature in [29]. These are rated 1.1 Ah, and the specific samples used are CS2-33 to CS2-38. They were tested with an Arbin battery tester, cycled under a constant current-constant voltage (CC-CV) charging mode until the voltage reached 4.2 V and under a constant current discharge mode until the cutoff voltage of 2.7 V. However, while the first two cells (numbers 33 and 34) were cycled at 1C, the other four cells (numbers 35 and 38) were cycled at 0.5C. The total number of cycles and the total energy throughput experienced by each of these cells during the experiment, which lasted seven months, is summarized in Table 1.

2.2. Dataset preparation

For each of the six cells, we extracted the charging and discharging voltage profiles (cleaning some outliers that corresponded to cycles that have not been completed properly) and generated the corresponding dQ/dV curves. The evolving shape of these curves as the cell is cycled and degraded, anticipated in the literature [25], can be observed in Fig. 1 a) and Fig. 1 b) for the charging and discharging processes, respectively.

These dQ/dV curves were generated using MATLAB and computed by differentiating the battery capacity Q with respect to the battery voltage V (i.e., calculating how much charge was exchanged by the cell at a given voltage level). The derivation from the voltage profiles and the discretization of the curves were performed with four different ΔV: 5 mV, 10 mV, 20 mV, and 50 mV. Then, four different kinds of dQ/dV profiles were obtained, as depicted in Fig. 2. Each of the represented curves presents a different number of points along the x axes. The total charge exchanged by the cell along each of the four curves is the same (1.1 Ah).

In this way, four datasets (one per ΔV value) were assembled for the charging processes, and the other four were assembled for the discharging processes of the six cells. Before uploading them to the Python

Table 1
Information on the cycling test performed on the cells being used.

Cell label	CS-33	CS-34	CS-35	CS-36	CS-37	CS-38
Number of cycles	864	777	873	972	989	1028
Energy throughout (Ah)	725.58	728.69	756.34	787.99	807.09	870.72

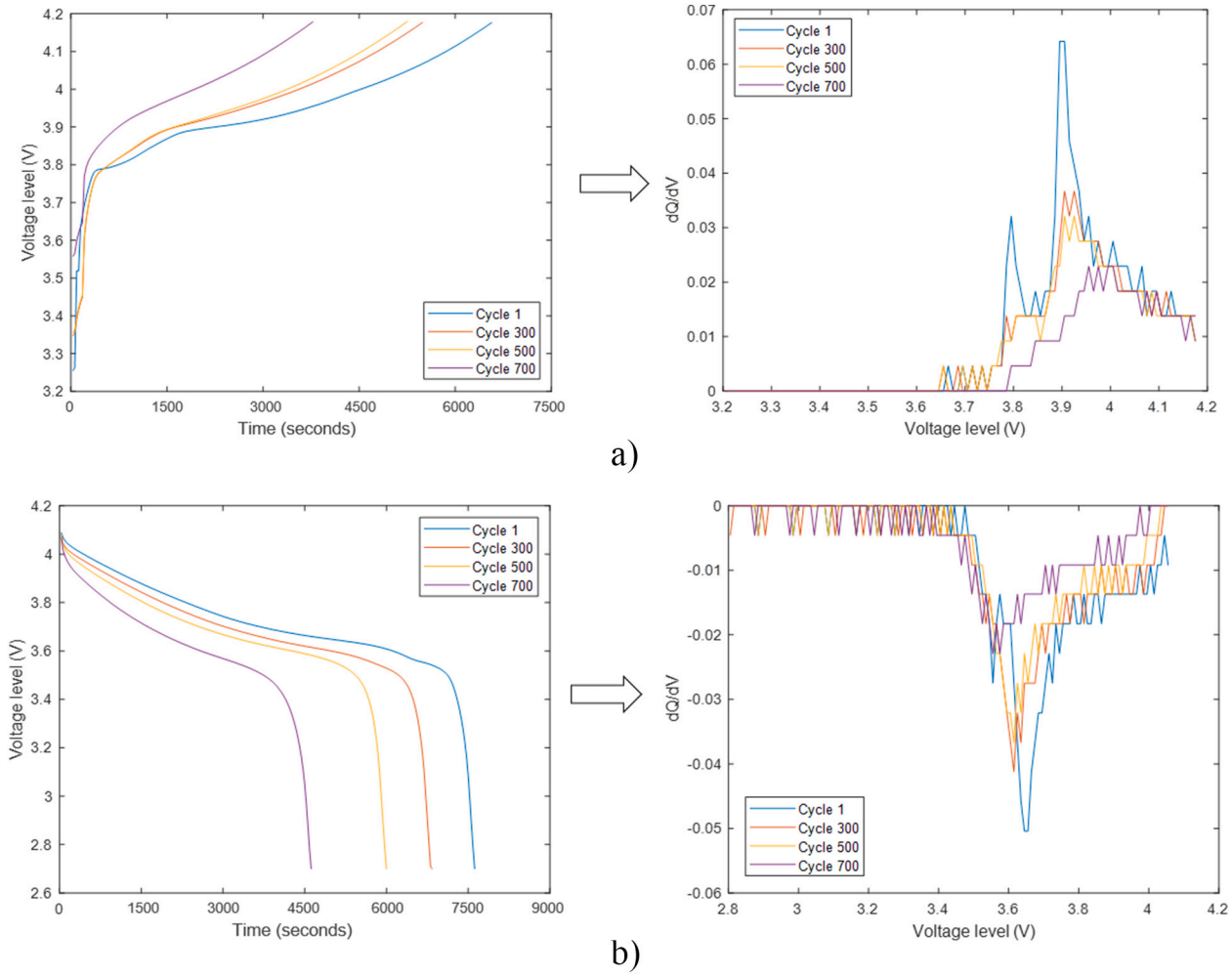

Fig. 1. Voltage profiles and dQ/dV curves derived for: a) charging and b) discharging semi-cycles during the experiment.

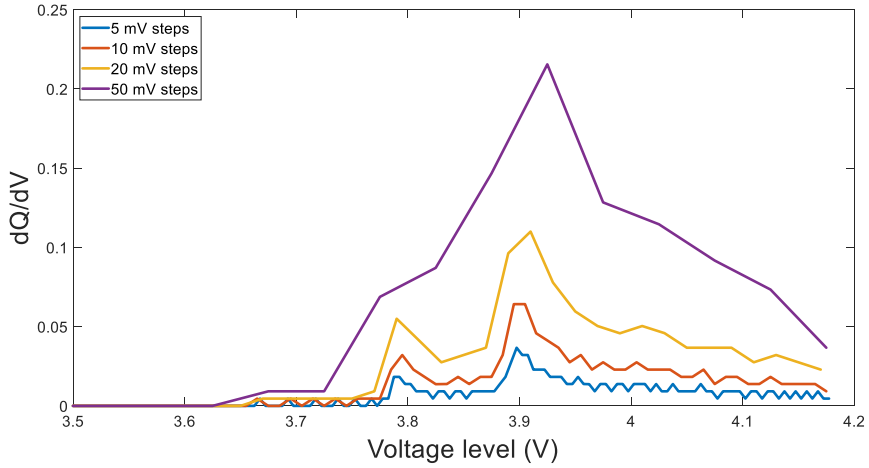

Fig. 2. Typical dQ/dV profiles obtained for the third cycle of a cell as a function of the ΔV value: 5 mV, 10 mV, 20 mV, and 50 mV.

environment, they were preprocessed. First, we truncated them below 3.48 V because lower voltages presented no dQ/dV value; hence, they did not contribute any information to the ML algorithms. Second, we took the absolute value of the profiles into account in order to work with positive values. And third, we normalized the data to make model training less sensitive to the scale of features. This allowed our models to converge to better weights and, in turn, achieve better accuracy. Finally, the dQ/dV profiles were complemented with the C-rate and the measured "retained capacity" for each cycle. This derived in matrixes with the structure like the one represented in Fig. 3 with sizes: 5486 × 253, 5486 × 128, 5486 × 65, and 5486 × 28, (as a function of the ΔV value).

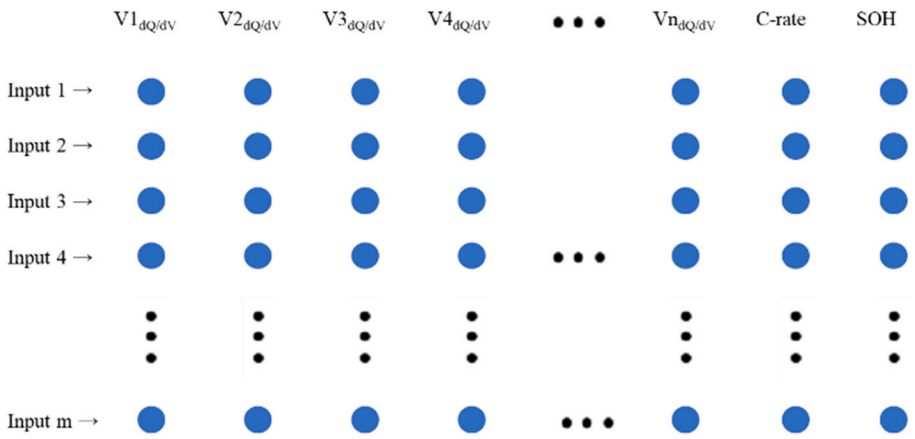

Fig. 3. Structure of the dataset used to train, validate, and test the ML algorithms. The number of columns "n" is dependent on the ΔV step value, while the number of rows "m" corresponds to the number of cycles registered for the six CS2 cells under consideration (5486 in all). The last two columns, beyond the "n" values of voltage, correspond to the C-rate and the measured retained capacity, respectively.

3. Machine learning algorithms

Six different ML algorithms have been considered in this work to estimate the SOH. The optimal architecture and/or hyperparameter for each algorithm has been optimized by means of Ray Tuner [30]. For the case of the NNs, the optimization has been done in terms of both MAE and RMSE. The algorithms implemented are the following.

3.1. Support vector regression

Support vector machine (SVM) is a well-known ML algorithm used in classification problems. It works by representing the data in a higher-dimensional space, by means of a kernel specified as a parameter, and by finding a soft margin hyperplane that successfully separates the observations into two classes. Support vector regression (SVR) models—as the one implemented in this work—use the same basic idea as SVM but apply it to predict values on a real scale, rather than for classification tasks. Here, support vectors are used to find the closest match between the data points and the actual function that is represented by them.

SVR models are usually defined by two parameters: the acceptable error margin (ϵ) and the tolerance of falling outside that acceptable error rate (C). Among the different combinations of these parameters analyzed, the one that performs best for our datasets is that with $\epsilon = 1$ and $C = 10$.

3.2. K-nearest neighbors

The K-nearest neighbors (KNN) algorithm is a supervised ML algorithm that works by finding the "k" most similar data points (nearest neighbors) to a given instance. KNN does not make any assumptions about the underlying data distribution. It is a nonparametric algorithm that relies on an item feature similarity metric. This metric, such as the Euclidean distance or the cosine similarity, is used to calculate the "distance" between the input and the data points in the training dataset. The target is then predicted by a majority vote, in the case of classification tasks, or by local interpolation of the targets associated with the nearest neighbors, in the case of the regression (as is used in this work).

KNN models are mainly defined by one operational parameter: the number of "k" most similar data points to be considered. Values from 1 to 9 have been explored, with results outstanding for $k = 5$.

3.3. Random forest

Random forest (RF) is another supervised ML algorithm based on independent decision trees (DTs). The "forest" in its name stands for an ensemble of such DTs usually trained with the bagging method, a combination of learning models that improves the overall result. Bagging generates models by randomly choosing a subset of features, which ensures low correlation among DTs. For classification tasks, the output of the RF is the class selected by most trees, while for regression tasks, the result is the average of the numerical outputs of the trees.

RF as a regressor presents many hyperparameters that can be tuned. We focused on just one: the number of estimators that, after some analysis, was fixed at 100.

3.4. Gradient boosting decision trees

Gradient boosting is also an ML technique, based on DTs, that produces a prediction model in the form of an ensemble of weak prediction models. Gradient boosting decision trees (known as GBDT or XGB) use each DT as the weak prediction model in gradient boosting. Its performance is based on the method of the additive training in which, at each iteration, a new tree learns the gradients of the residuals between the target values and the current predicted values from the previous tree, and then the algorithm conducts gradient descent based on the learned gradients. Therefore, it is a sequential algorithm that cannot be parallelized like the RF, and the algorithm execution reduces to parallel DT building. It is a slow-to-train model, but it is very fast to predict.

As with the RF algorithm, XGB presents many hyperparameters that can be tuned. We focused on just two of them: the number of estimators, which after some analysis was also fixed at 100, and the maximum tree depth for the base learners, which we set to 10.

3.5. Fully connected network

Deep neural network (DNN) models are composed of units (neurons) that combine multiple inputs and produce a single output. These units are arranged in layers. Hence, "deep" refers to models with high complexity in the number of layers and units per layer. The use of deep models allows for capturing higher levels of patterns in the data and thus obtaining more accurate results.

Feedforward networks are the most basic DNNs composed of a series of fully connected (FC) layers that form a network. Each successive layer is a set of nonlinear functions of a weighted sum of all the FC outputs of the previous layer. For regression tasks, these models have a single output whose numerical value represents the expected value of the approximate function for a given input.

After running around 1000 simulations with multiple combinations of the main hyperparameters of the fully connected network (FCN) (mainly the batch size, the number of FC layers, the number of neurons

in each layer, the number of outputs from each layer, and the learning rate), we selected the best-performing combination. Fig. 4 presents this architecture, with four dense layers (the first two included a 0.25 ratio dropout layer), a batch size of 32, and a learning rate of 0.001, trained during 100 epochs.

3.6. Convolutional neural network

A convolution is a mathematical operation often used for image processing and recognition since its effect is similar to filtering. The convolutional layer is the core building block of a convolutional neural network (CNN). This type of layer applies a set of learnable filters to the input, generally an image (2D), by executing convolution operations. These filters, or feature detectors, are applied to an area of the input as a sliding window, extracting high-level features that can be consequently passed on to the next layer. Typically, convolutional layers are followed by a pooling layer, which downsamples the outputs to provide local translation invariance by summarizing the presence of features in the previous convolutional output. By applying various convolutional filters, CNN models can capture a high-level representation of the input data, which is then sent to an FCN to reduce the dimension of features to the expected dimension of the output.

In this case, since the information is generated in a vector format, 1D convolutional layers were implemented. After the corresponding simulations run with Ray Tune (other 1000 trials) while varying the combinations of the main hyperparameters of the NN—mainly the number of layers (both 1D convolutional and FC), the number of neurons in each layer, the kernel composition, and the learning rate—we selected once again the best-performing combination. This corresponds to the architecture shown in Fig. 5, with four convolutional (with one pooling layer in the middle) and two dense layers, a batch size of 32, and a learning rate of 0.001, trained during 100 epochs. Also note how the C-rate is only provided to the NN after the convolutional operations and the dropout are executed.

4. Results

All the simulations and calculations were performed using the PyTorch framework and some specific Python libraries such as “pandas”, “numpy”, and “scikit-learn.” Datasets were divided into training (60 %), validating (20 %), and test (20 %) data. At this point, it is important to highlight that to avoid the variability in the results inevitably derived from the aleatory partitioning introduced to the datasets for training, a fivefold cross-validation strategy [31] was implemented with each of the models simulated.

Results were initially calculated for the datasets containing the complete dQ/dV curve. This allowed us to analyze how the various ΔV values selected influence the performance of the ML algorithms. Then, we evaluated the final performance of the algorithms with partial sections of the dQ/dV curves—just for the better-performing datasets—and compared with previous works. The results are as follows.

4.1. Full charge/discharge dQ/dV profiles

Fig. 6 and Fig. 7 display the resulting capacity fade estimation error in terms of MAE and RMSE, respectively, achieved by each of the five ML algorithms when using any of the four datasets under consideration (with ΔV being 5, 10, 20, and 50 mV).

It becomes clear that the NN models (represented as FCN and CNN in the figures) outperform the other algorithms when the whole dQ/dV profile is available. Among them, the SVR model works clearly worse than the rest, followed by the KNN (which gets much worse as the ΔV is reduced), the RF, and the XGB. If the two NN-based models are compared, results are similar in both MAE and RMSE. The best performance among them is the CNN, with an MAE value as low as 0.002 (i.e. 0.2 %, when considering the unit would be the whole initial capacity of the battery) and an RMSE as low as 0.003 (0.3 %). The good accuracy of the retained capacity estimation achieved with the CNN is evident in Fig. 8, in which the estimation values offered by the CNN with the fivefold cross-validation methodology are depicted over the actual retained capacity labels registered in the dataset for each of the test data instances. Note how the estimated capacity values, calculated as the average value of the estimations obtained for a given test instance at each of the cross-validation tests, almost overlap the actual values in all the cases with just a few exceptions in which the standard deviation of the estimation is observed. Such estimation error levels, obtained with

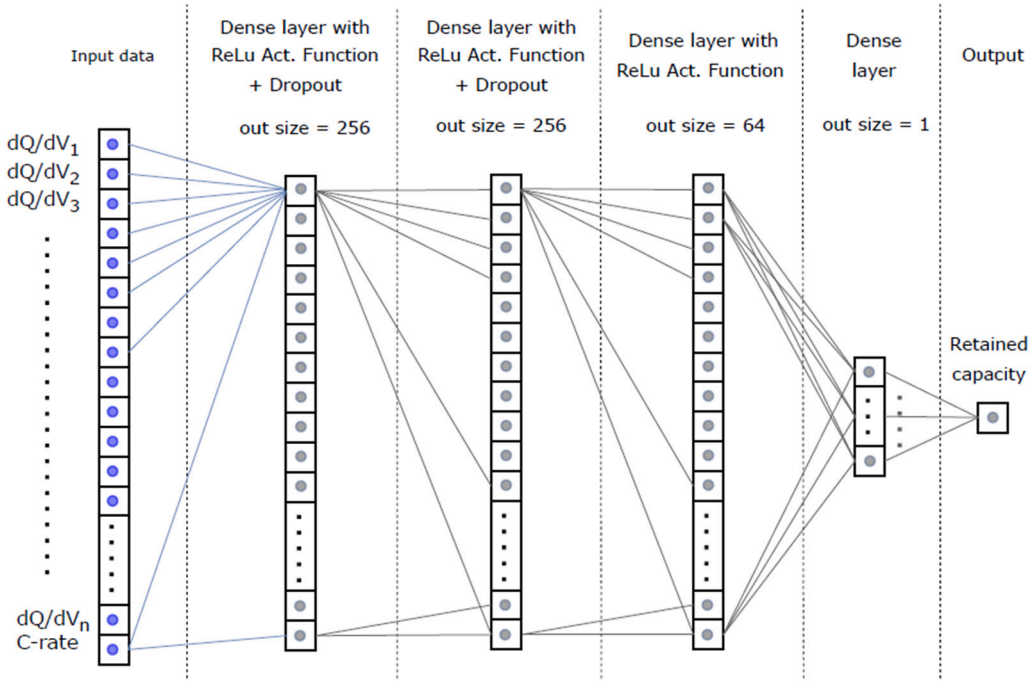

Fig. 4. Architecture of the proposed FCN with four FC layers and two dropout layers.

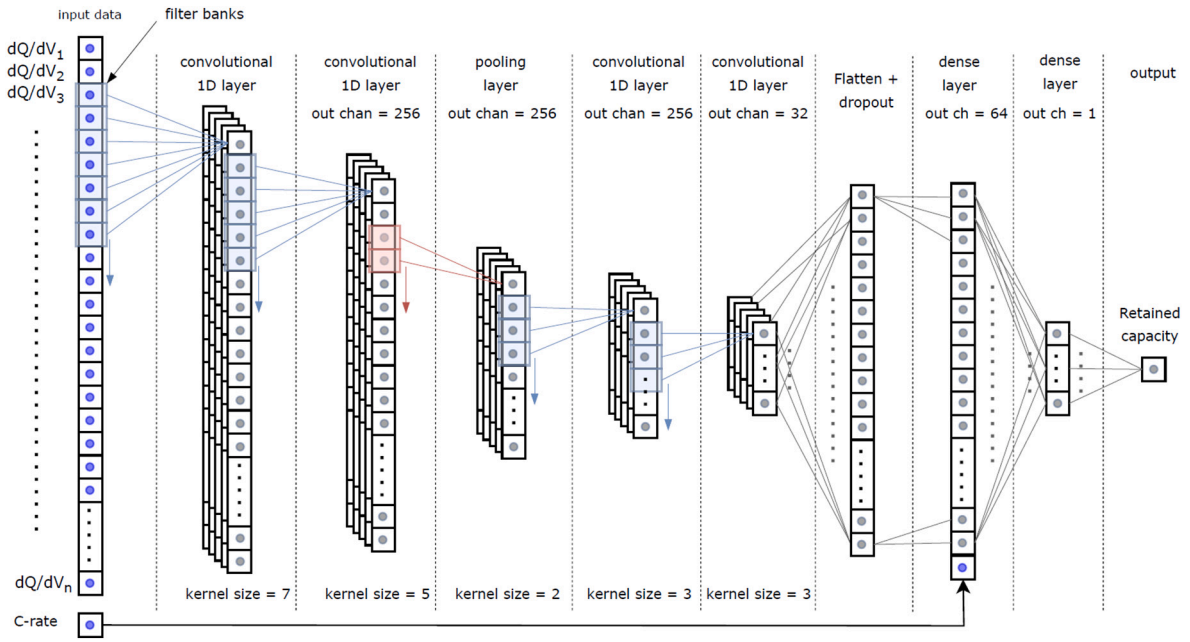

Fig. 5. Architecture of the CNN implemented with four convolutional 1D layers, one pooling, one flatten and one dropout layer, and two FC dense layers.

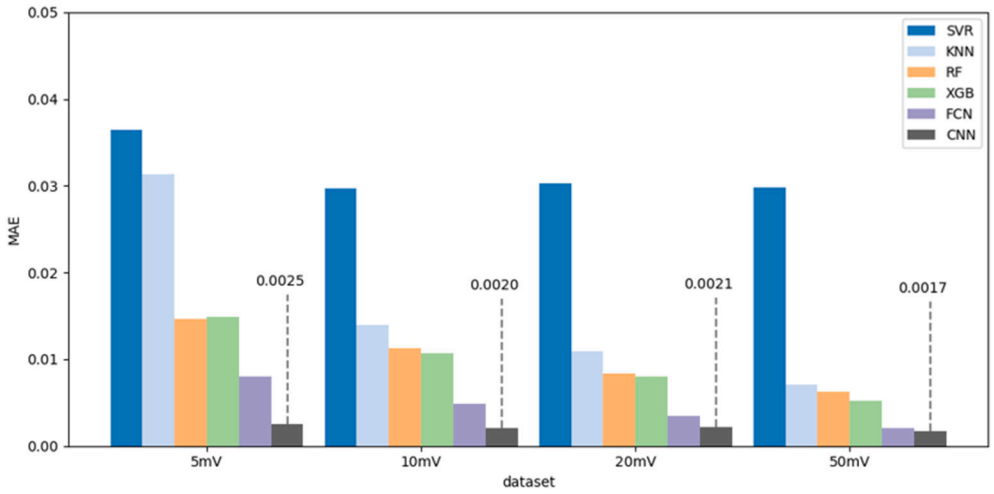

Fig. 6. Resulting MAEs of the retained capacity estimation results for the six ML algorithms with each of the four datasets under test.

the two NN models, clearly outperform those obtained in [28] and are in close competence with those in [26], which were achieved for a different type of Li-ion cells and are, therefore, more difficult to compare.

Apart from that, when comparing results among the four datasets, the larger the ΔV value the better the performance for all the ML algorithms because a larger ΔV implies shorter dQ/dV vectors and more pronounced/characteristic profiles, Fig. 2. Within this general trend, results for both the FCN and CNN models present similar responses to both the 20 mV and the 50 mV datasets in terms of both MAE and RMSE, producing the best resulting combinations of those analyzed in this work. Therefore, given the similarity in the best-case response for these two datasets, both are analyzed in the partial dQ/dV profiles evaluation.

4.2. *Partial charge/discharge dQ/dV profiles*

When the dQ/dV profiles are not complete and only a part is registered (obtained from a partial charge or discharge process), the results vary significantly. Five different scenarios have been analyzed. These correspond to the cases with portions of the dQ/dV, compiling the 20 %, 30 %, 40 %, 50 %, and 80 % portions of it. Fig. 9 and Fig. 10 summarize the respective MAE and the RMSE values obtained for each of the ML algorithms with the two datasets under consideration: the $\Delta V = 20$ mV and the $\Delta V = 50$ mV datasets.

When comparing the results among datasets, the 20 mV dataset allows the ML models to better estimate the retained capacity with a lower MAE and a similar RMSE. The difference gets more significant, especially for the MAE value, as the registered portions of the dQ/dV curve decreases below the 50 % of the whole cycle.

If we focus the analysis on comparing the performance of the different algorithms versus the same dataset, results indicate that, as for the case of the whole dQ/dV profiles, the SVR algorithm works significantly less effectively than the others (especially with the 50 mV dataset). Also, the CNN performs marginally better in this scenario than its FCN counterpart. Both the FCN and the CNN models are consistently superior to the rest of algorithms. Only when portions of the dQ/dV curve are below 40 % do the RF and the XGB models become competitors to the NNs, although even in the worst case (for 20 % portions) the CNN keeps being the best solution.

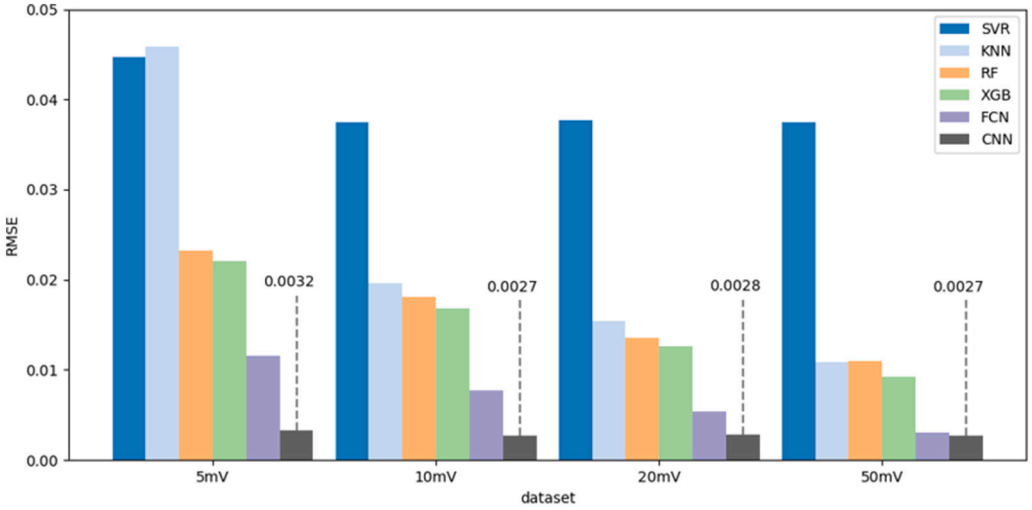

Fig. 7. Resulting RMSEs of the retained capacity estimation results for the six ML algorithms with each of the four datasets under test.

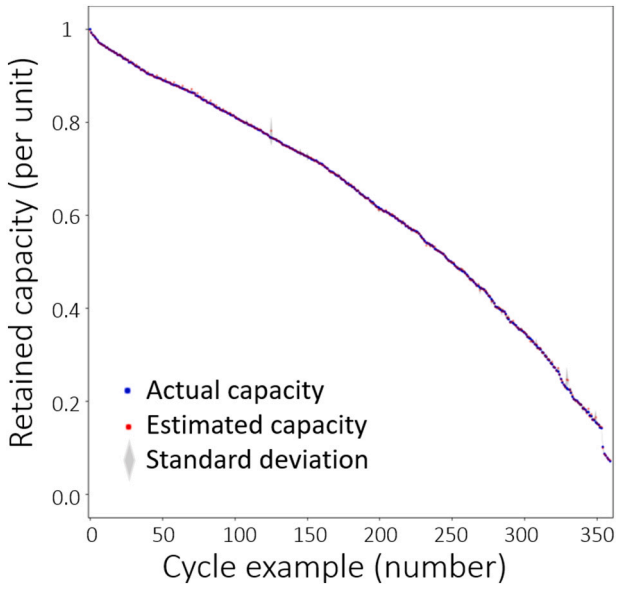

Fig. 8. Actual retained capacity labels (blue) vs. retained capacity estimations (red) and its corresponding standard deviation (grey band) achieved with the CNN model for the test data from the 50 mV dataset when the whole dQ/dV profile is available. MAE = 0.002, RMSE = 0.0032.

However, even though the CNN algorithm would be the selection of choice when just partial curves were available, the best CNN model explored would not appear to be a good solution for industrial applications due to the resulting MAE and RMSE values. It presents errors (in MAE) of 0.0106 (1.1 %), 0.0187 (1.9 %), 0.027 (2.7 %), and 0.035 (3.5 %) for the 50 %, 40 %, 30 %, and 20 % portions, respectively. Although these values surpass results offered in previous references, such as [28], they would still imply a poor estimate for a reliable commercial system such as the BMS of an EV.

The limited accuracy of the retained capacity estimation capability can be observed in Fig. 11. This has been obtained for the CNN with the 50 mV dataset at 20 % partial curves. With a MAE value of 0.038 (3.8 %) for the whole test dataset, estimated values (in red, calculated as the average of the capacity estimates obtained for each test instance during the fivefold cross-validation strategy) clearly differ along most of the capacity range from actual measured values (in blue). The standard deviation of the estimates (grey band around the red points) would be also too high to consider the estimation acceptable, or accurate, especially for low SOH values.

In fact, the performance of the algorithm gets clearly worse as the battery SOH diminishes, Fig. 11. While the battery operates within its healthy range (retained capacity above 80 %) the MAE is 0.028, which would still represent an acceptable estimation error value. However, when the retained capacity gets below 80 % the MAE drops to 0.061 and the estimation becomes highly inaccurate.

4.3. *Discussion about the results*

Results demonstrate that the ΔV value used to discretize dQ/dV curves influences the MAE and RMSE results provided by the different algorithms. Changing the ΔV value from 5 mV to 50 mV involves a reduction in the MAE error for full cycles from 0.25 % to 0.17 % when using the best-performing algorithm, the CNN. Differences between the 20 mV and the 50 mV datasets are not so significant and both have been considered for the partial cycles' analysis. Then, as the available portion of cycle decreases, the better the 20 mV dataset performs in relation to the 50 mV dataset. For the 20 % portions case, the 20 mV provides better results than those achieved for the 50 mV dataset.

Among the different ML algorithms under evaluation, those based on NN outperform the rest (SVR, KNN, RF, and XGB), by between 0.5 % and 4 % depending on the case, and they achieve so both when the whole dQ/dV curve is available and when just a portion of the curve (cycles ranging from 30 % to 80 % of full scale) is registered. Between them, the CNN model works better than the FCN counterpart for all the datasets and portions analyzed, with values as low as 0.002 (0.2 %) in MAE and 0.003 (0.3 %) in RMSE for the whole curve case, and 0.035 in MAE and 0.066 in RMSE for the cases with dQ/dV curve portions as low as 20 %. Therefore, when comparing the algorithms in terms of SOH estimation accuracy, the defined CNN model would be the best option to analyze retained battery capacities.

However, although retained capacity estimation outcomes with the whole dQ/dV curve result accurate (significantly below 1.5 %, as observed in Fig. 8) for most of the algorithms analyzed, the error level achieved with partial charge/discharge operations is still high for all of them, especially when SOH actual values get below 80 %. Even the CNN model provides estimations (0.014 (1.4 %), 0.020 (2.0 %), 0.027 (2.7 %), and 0.035 (3.5 %) for the 50 %, 40 %, 30 %, and 20 % portions, respectively) that would be, in our opinion, too inaccurate for a reliable diagnose (as observed in Fig. 11).

In terms of efficiency and time consumption, the different types of algorithms implemented would be valid for commercial products

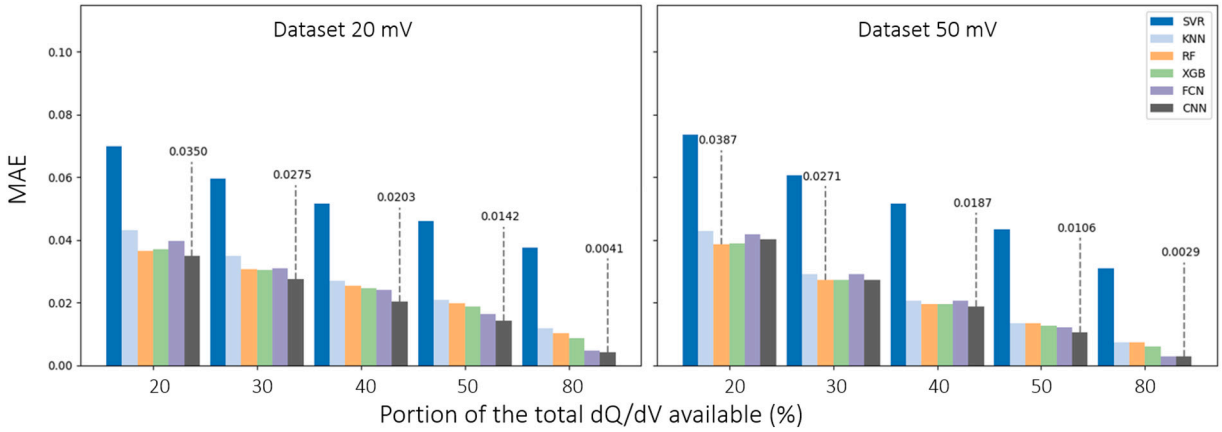

Fig. 9. Resulting MAEs of the retained capacity estimation results for the six ML algorithms with partial cycles.

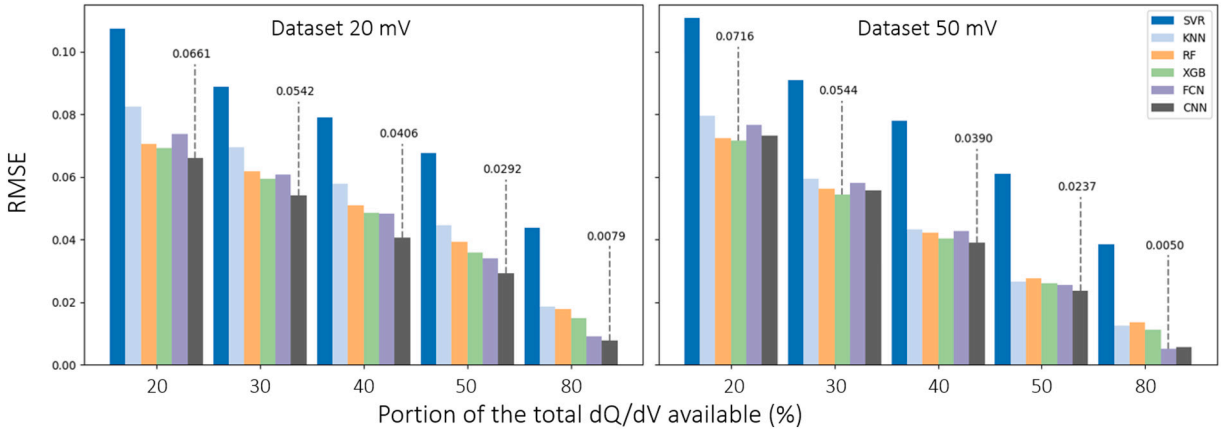

Fig. 10. Resulting RMSEs of the retained capacity estimation results for the six ML algorithms with partial cycles.

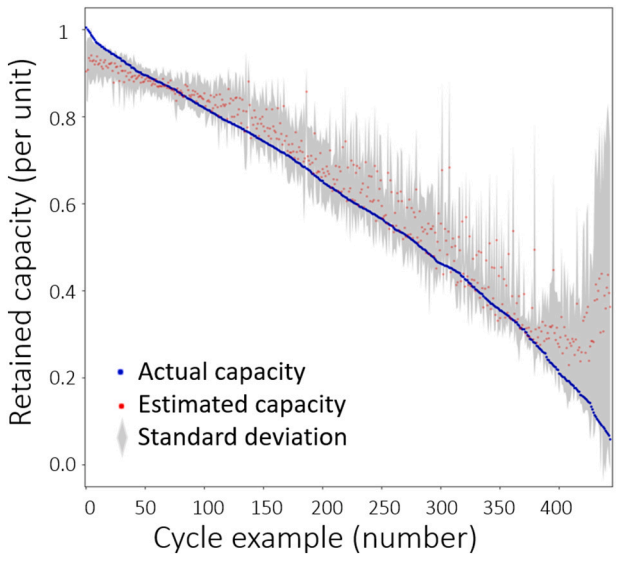

Fig. 11. Actual retained capacity labels (blue) vs. retained capacity estimations (red) and its corresponding standard deviation (grey band) achieved with the CNN model for the test data from the 50 mV dataset when just the 20 % of the dQ/dV profile is available. MAE = 0.042, RMSE = 0.072.

because, even though the training of the models can be time consuming (specially for the deep learning cases), all of them could provide retained capacity estimations in a timely manner without a great computing burden once they are trained. Therefore, although some of the algorithms would be more efficient in terms of time consumption and computing requirements, this would not be a critical parameter for the application under consideration.

Nonetheless, attending to the different simulations performed and to the results achieved for the various portions of curves and with the different algorithms, it can be concluded that none of the algorithms analyzed would conform by itself a reliable methodology to estimate from partial cycles smaller than 50 % the retained capacity of batteries. In fact, any SOH estimation method with MAE and RMSE values above 0.015 (1.5 %) would not seem an acceptable solution to implement in rigorous commercial systems or to analyze the potential use of the batteries, for instance, in second life applications. This implies the KNN, the RF and the XGB algorithms would only be acceptable when the whole charging cycles were available; the SVR model would not be suitable even in this case; and the FCN or CNN models could be valid potential solutions when the depth-of-discharge of the cycles was above 80 %.

In any case, ML SOH estimation models would require large cycles to provide accuracy and potentially complement the partial dQ/dV curve portion with further information when implemented in commercial solutions. Also, external influences on the dQ/dV curve profile such as the temperature would require some analysis. This is an open topic to be taken into consideration in future works.

5. Conclusions

This paper analyzed the suitability of six different ML algorithms (SVR, KNN, RF, XGB, FCN, and CNN) to diagnose and estimate the retained capacity of a Li-ion battery at a given moment during its operation lifetime using the data registered by the battery management system during the last partial or full charging/discharging process. Data available online from long-term degradation experiments on LCO prismatic cells have been used to generate the training datasets and test the models. These data allowed generating the dQ/dV curves (with four different formats) during full and partial charging/discharging processes which were combined with the C-rate values of the corresponding cycles. The main conclusions extracted are summarized as follows:

1) The ΔV step used to discretize the dQ/dV curve (5 mV, 10 mV, 20 mV, and 50 mV) influences the performance of the ML algorithms. The best results arise with the 20 mV and the 50 mV step-size to datasets for partial and full cycles, respectively.
2) Among the six different ML algorithms implemented, the defined CNN model would be the best option to analyze retained battery capacities with both full and partial cycles.
3) The proposed methods present strong robustness to the cell C-rate and state-average of-charge of the partial cycles since no significant differences are registered in the estimation results when varying these parameters.
4) Although the retained capacity estimation results become very accurate for most of the algorithms analyzed when the whole dQ/dV curve is registered, the error with partial charge/discharge operations is too high (above 1.5 %) even for the CNN model, especially when the battery SOH gets below 80 % and for partial cycles smaller than 80 %. Only within the framework with cycles larger than 80 % could the FCN or CNN models be valid potential solutions to be implemented in commercial applications. The KNN, RF and XGB algorithms would only be acceptable when the whole charging cycles were available and the SVR model would not be suitable even in that case.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Hector Beltran reports financial support and article publishing charges were provided by University Jaume I. Hector Beltran reports travel was provided by Generalitat Valenciana - Conselleria d'Innovació, Universitats, Ciència i Societat Digital.

Data availability

The datasets used in this study are available at: https://web.calce.umd.edu/batteries/data.htm

Acknowledgements

This work was supported in part by the University Jaume I under Grant UJI-B2021-35 and in part by the Generalitat Valenciana under grant CIBEST/2021/54.

CRediT authorship contribution statement

Hector Beltran: Investigation, Conceptualization, Methodology, Software, Data curation, Writing- Original draft preparation. **Emilio Sansano**: Software, Validation. **Michael Pecht**: Visualization, Writing Reviewing and Editing, Supervision.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.est.2022.106346.

References

[1] J.B. Goodenough, H. Gao, A perspective on the Li-ion battery, in: Science China Chemistry 62, Science in China Press, Dec. 01, 2019, pp. 1555–1556, https://doi.org/10.1007/s11426-019-9610-3, no. 12.
[2] Wood Mackenzie, Global lithium-ion battery capacity to rise five-fold by 2030, online, https://www.woodmac.com/press-releases/global-lithium-ion-battery-capacity-to-rise-five-fold-by-2030/, 2022. (Accessed 26 July 2022).
[3] Statista, Lithium-ion batteries - statistics & facts, online, https://www.statista.com/topics/2049/lithium-ion-battery-industry/, 2022. (Accessed 26 July 2022).
[4] Y. Miao, L. Liu, Y. Zhang, Q. Tan, J. Li, An overview of global power lithium-ion batteries and associated critical metal recycling, J. Hazard Mater. 425 (November 2021) (2022) 127900, https://doi.org/10.1016/j.jhazmat.2021.127900.
[5] B. Pan, et al., Aging mechanism diagnosis of lithium ion battery by open circuit voltage analysis, Electrochim. Acta 362 (Dec. 2020), 137101, https://doi.org/10.1016/J.ELECTACTA.2020.137101.
[6] Z. Wang, G. Feng, D. Zhen, F. Gu, A. Ball, A review on online state of charge and state of health estimation for lithium-ion batteries in electric vehicles, in: Energy Reports vol. 7, Elsevier Ltd, Nov. 01, 2021, pp. 5141–5161, https://doi.org/10.1016/j.egyr.2021.08.113.
[7] M. Ecker, et al., Calendar and cycle life study of Li(NiMnCo)O2 -based 18650 lithium- ion batteries, J. Power Sources 248 (2013) 839–851, https://doi.org/10.1016/j.jpowsour.2013.09.143.
[8] R. Xiong, Y. Pan, W. Shen, H. Li, F. Sun, Lithium-ion battery aging mechanisms and diagnosis method for automotive applications: recent advances and perspectives, in: Renewable and Sustainable Energy Reviews vol. 131, Elsevier Ltd, Oct. 01, 2020, p. 110048, https://doi.org/10.1016/j.rser.2020.110048.
[9] M.F. Ge, Y. Liu, X. Jiang, J. Liu, A review on state of health estimations and remaining useful life prognostics of lithium-ion batteries, Measurement 174 (Apr. 2021), https://doi.org/10.1016/j.measurement.2021.109057 (Lond).
[10] S.B. Sarmah, et al., A review of state of health estimation of energy storage systems: challenges and possible solutions for futuristic applications of Li-ion battery packs in electric vehicles, in: Journal of Electrochemical Energy Conversion and Storage vol. 16, American Society of Mechanical Engineers (ASME), Nov. 01, 2019, https://doi.org/10.1115/1.4042987 no. 4.
[11] H. Meng, Y.F. Li, A review on prognostics and health management (PHM) methods of lithium-ion batteries, Renew. Sust. Energ. Rev. 116 (January) (2019), 109405, https://doi.org/10.1016/j.rser.2019.109405.
[12] J. Peng, S. Jia, S. Yang, X. Kang, H. Yu, Y. Yang, State estimation of lithium-ion batteries based on strain parameter monitored by fiber Bragg grating sensors, J. Energy Storage 52 (Aug. 2022), https://doi.org/10.1016/j.est.2022.104950.
[13] X. Sui, S. He, S.B. Vilsen, J. Meng, R. Teodorescu, D.I. Stroe, A review of non-probabilistic machine learning-based state of health estimation techniques for lithium-ion battery, Appl. Energy 300 (January) (2021), https://doi.org/10.1016/j.apenergy.2021.117346.
[14] Y. Li, et al., Data-driven health estimation and lifetime prediction of lithium-ion batteries: a review, in: Renewable and Sustainable Energy Reviews vol. 113, Elsevier Ltd, Oct. 01, 2019, https://doi.org/10.1016/j.rser.2019.109254.
[15] X. Shu, et al., State of health prediction of lithium-ion batteries based on machine learning: advances and perspectives, iScience 24 (11) (Nov. 2021), 103265, https://doi.org/10.1016/J.ISCI.2021.103265.
[16] S. Amir, M. Gulzar, M.O. Tarar, I.H. Naqvi, N.A. Zaffar, M.G. Pecht, Dynamic equivalent circuit model to estimate state-of-health of lithium-ion batteries, IEEE Access 10 (2022) 18279–18288, https://doi.org/10.1109/ACCESS.2022.3148528.
[17] D. Roman, S. Saxena, V. Robu, M. Pecht, D. Flynn, Machine learning pipeline for battery state-of-health estimation, Nat. Mach. Intell. 3 (5) (2021) 447–456, https://doi.org/10.1038/s42256-021-00312-3.
[18] N.H. Paulson, J. Kubal, L. Ward, S. Saxena, W. Lu, S.J. Babinec, Feature engineering for machine learning enabled early prediction of battery lifetime, J. Power Sources 527 (February) (2022), 231127, https://doi.org/10.1016/j.jpowsour.2022.231127.
[19] D. Yang, Y. Wang, R. Pan, R. Chen, Z. Chen, A neural network based state-of-health estimation of lithium-ion battery in electric vehicles, Energy Procedia 105 (2017) 2059–2064, https://doi.org/10.1016/j.egypro.2017.03.583.
[20] Y. Fan, F. Xiao, C. Li, G. Yang, X. Tang, A novel deep learning framework for state of health estimation of lithium-ion battery, J. Energy Storage 32 (August) (2020), 101741, https://doi.org/10.1016/j.est.2020.101741.
[21] P. Li, et al., An end-to-end neural network framework for state-of-health estimation and remaining useful life prediction of electric vehicle lithium batteries, Renew. Sust. Energ. Rev. 156 (Mar. 2022), 111843, https://doi.org/10.1016/J.RSER.2021.111843.
[22] D. Kong, S. Wang, P. Ping, State-of-health estimation and remaining useful life for lithium-ion battery based on deep learning with Bayesian hyperparameter optimization, Int. J. Energy Res. 46 (5) (Apr. 2022) 6081–6098, https://doi.org/10.1002/er.7548.
[23] T.K. Pradyumna, K. Cho, M. Kim, W. Choi, Capacity estimation of lithium-ion batteries using convolutional neural network and impedance spectra, J.Power Electron. 22 (5) (May 2022) 850–858, https://doi.org/10.1007/s43236-022-00410-4.

[24] L. Ungurean, M.V. Micea, G. Casrtiou, Online state of health prediction method for lithium-ion batteries based on Intl J of Energy Research - 2020 - Ungurean.pdf, Int. J. Energy Res. (2020) 6767–6777, https://doi.org/10.1002/er.5413.
[25] Q. Gong, P. Wang, Z. Cheng, An encoder-decoder model based on deep learning for state of health estimation of lithium-ion battery, J. Energy Storage 46 (November 2021) (2022) 103804, https://doi.org/10.1016/j.est.2021.103804.
[26] N. Yang, Z. Song, H. Hofmann, J. Sun, Robust State of Health estimation of lithium-ion batteries using convolutional neural network and random forest, J. Energy Storage 48 (March 2021) (2022) 103857, https://doi.org/10.1016/j.est.2021.103857.
[27] Z. Wei, H. Ruan, Y. Li, J. Li, C. Zhang, H. He, Multistage state of health estimation of lithium-ion battery with high tolerance to heavily partial charging, IEEE Trans. Power Electron. 37 (6) (Jun. 2022) 7432–7442, https://doi.org/10.1109/TPEL.2022.3144504.
[28] C. Qian, et al., Convolutional neural network based capacity estimation using random segments of the charging curves for lithium-ion batteries, Energy 227 (2021), 120333, https://doi.org/10.1016/j.energy.2021.120333.
[29] CALCE Battery Research Group, CALCE Database, 2017.
[30] R. Liaw, E. Liang, R. Nishihara, P. Moritz, J.E. Gonzalez, I. Stoica, Tune: a research platform for distributed model selection and training, in: 2018 ICML AutoML Workshop, Jul. 2018, https://doi.org/10.48550/arXiv.1807.05118.
[31] M. Ojala, G.C. Garriga, Permutation tests for studying classifier performance, J. Mach. Learn. Res. 11 (2010) 1833–1863.