

Research papers

State of health estimation of lithium-ion batteries using EIS measurement and transfer learning

Yichun Li^{*}, Mina Maleki, Shadi Banitaan

Department of Electrical & Computer Engineering & Computer Science, University of Detroit Mercy, Detroit, MI, United States

ARTICLE INFO

Keywords:

Deep neural network
Electric vehicles
Electrochemical impedance spectroscopy
State of health
Transfer learning
Lithium-ion batteries

ABSTRACT

Accurately estimating the state of health (SOH) of lithium-ion batteries in real-world scenarios, especially for electric vehicles (EVs) is challenging due to dynamic operating conditions and limited battery usage data. To address these challenges, transfer learning, as an effective approach that uses the knowledge gained from a source task to improve efficiency and accuracy on a similar target task, has been combined with deep neural networks (DNN) to determine the SOH of Li-ion batteries at various temperatures with a limited amount of target data. In this study, the SOH of Li-ion batteries has been estimated using one of the largest Electrochemical Impedance Spectroscopy (EIS) datasets collected at three different temperatures, 25, 35, and 45 degrees Celsius. The source task is to estimate the SOH using DNN models at one or two operating temperatures (25, 35, and 45 degrees Celsius), while the target task is to estimate the SOH at the other temperature. The results demonstrate that the average mean squared error (MSE) and mean absolute percentage error (MAPE) for estimating the SOH using transfer learning-based DNN (DNN-TL) at 35 degrees Celsius have been reduced up to 72.63% and 48.43%, respectively, compared with the DNN model without transfer learning. Similarly, the average MSE and MAPE for estimating the SOH at 45 degrees Celsius using DNN-TL have been reduced up to 79.51% and 55.11%, respectively. Also, our findings demonstrate the robustness of DNN-TL models, requiring target data with only 6% of the size of source data for retraining, leveraging previously learned knowledge from the source data. Furthermore, the exclusion of fixed layers in the retraining process enhances the performance of the DNN-TL model.

1. Introduction

Due to their high power density, long life span, and low self-discharge rate, lithium-ion batteries are widely used in portable electronic devices and electric vehicles (EVs) as a promising energy storage solution. Battery capacity, as opposed to battery voltage, current, and temperature, cannot be directly measured with gauges in a battery management system (BMS) in real time. Battery maximum capacity decreases as a result of the aging of lithium-ion batteries over time, which is an inevitable process caused by the chemical changes inside the battery. In this regard, the primary objective and challenge of the BMS design is to develop advanced algorithms for accurate and efficient battery state estimations.

In order to determine the degradation level of the Li-ion batteries, the state of health (SOH), calculated as the ratio of the maximum battery charge Q_m to the rated capacity of the battery C_r , by the manufacturer as shown in Eq. (1), has to be estimated accurately to reflect the real-time status and performance of the batteries in BMS.

$$SOH = \frac{Q_m}{C_r} \quad (1)$$

Coulomb counting [1] has been widely used as a direct capacity measurement in a laboratory environment to determine the capacity of Li-ion batteries by integrating the battery current over time during the battery charging or discharging. Considering that a fully discharging and charging cycle can accelerate the aging of lithium-ion batteries, coulomb counting cannot be implemented in real-life applications in which real-time estimations are required. Therefore, designing robust and reliable battery management systems that accurately estimate battery capacity remains a challenging task. Besides the direct measurement like counting coulombs, there are mainly two methods that have been utilized to estimate the battery capacity recently, namely model-based and data-driven approaches.

In model-based approaches, the degradation trajectory of batteries is described mathematically through a mathematical model. Since the battery degradation process is non-linear and complex, it is often difficult to make a trade-off between model complexity and prediction accuracy [2]. As the battery operation conditions change in the real world, every change that occurs to the batteries will result in a loss of

^{*} Corresponding author.

E-mail addresses: liy18@udmercy.edu (Y. Li), malekimi@udmercy.edu (M. Maleki), banitash@udmercy.edu (S. Banitaan).

prediction accuracy, since those factors should be taken into account when deriving the mathematical equation that describes the relationship between various aging factors and battery capacity. Moreover, it is a time-consuming and computationally expensive process to derive a model-based method to predict the battery states since the electrochemical principles of battery and battery materials combinations are specifically required to build the model. This results in the issue that the model-based methods usually have poor generalization capacity when the battery types vary. Additionally, such a model cannot ultimately reflect the dynamics of a Li-ion battery in real life [3,4].

With the advancement of machine learning techniques and the availability of large amounts of battery data, data-driven approaches have recently gained considerable attention in order to overcome the limitations of direct measurement and model-based methods in capacity measurement. Also, data-driven approaches are much more feasible and practical as it is not required to have an understanding of the degradation process of the battery. As a result of using data-driven methods, it is possible to determine the relationship between capacity and other easily measurable features, such as voltage, current, temperature, impedance, and so on, without taking into account the mathematical model of the battery and the nonlinear and complex aging mechanisms of the battery. The difficulty of estimating the capacity of Li-ion batteries can be overcome by converting the impractical measurement of the capacity of Li-ion batteries into a machine learning regression problem, which is the main advantage of data-driven approaches. But as a consequence of using machine learning techniques, data-driven methods rely solely on experimental data collected from the battery, which results in its sensitivity to the quality of the dataset. In the case where the amount of battery usage data is limited or data are collected in a few operation conditions, the capacity of generalization may behave poorly since the model has not learned to predict the capacity at an operating condition that it has not learned before.

Generally, data-driven approaches are developed under the assumption that the data used for training and testing is drawn from the same data experiment and is expected to have the same distribution [5]. It follows that a statistical model trained on one dataset cannot be directly applied to another dataset with a different distribution, and the model must be reconstructed and retrained from scratch on the new dataset. Nevertheless, in real-world applications, obtaining sufficient data to retrain a new model is often time-consuming and costly, while the training process takes a considerable amount of time. To address such cases, transfer learning has been proposed, which reduces the need and effort for data recollection and model re-training by utilizing the knowledge gained from a source task to a different but related task, even though the data distributions of these two tasks may vary. As a result of the transfer learning, the latter task requires less data in order to retrain the model.

In order to implement the transfer learning technique to estimate battery degradation while leveraging the benefits of data-driven approaches, the SOH of Li-ion batteries has been estimated using one of the largest Electrochemical Impedance Spectroscopy (EIS) datasets collected at three different temperatures, 25, 35, and 45 degrees Celsius. Even though the Li-ion battery cells tested in this dataset are the same, testing at different temperature settings results in the diversity of the data distribution. As an example, the data-driven model trained with the EIS dataset at 25 and 35 degrees Celsius failed to predict battery capacity at 45 degrees Celsius, which results in the question of how the trained data-driven model can be applied to a wider range of operations scenarios as more and more battery usage data with various distributions become available over time. Results of the experiment indicated that the deep neural network (DNN) model does not perform well in predicting data with different distributions than it previously learned, which is consistent with the assumption of machine learning models. Hence, the paper investigates the effectiveness of transfer learning by retaining the model with previous knowledge to predict the data with different distributions for battery degradation estimation.

Also, the amount of new data needed to retrain the trained model will be investigated to see how transfer learning can improve efficiency and overcome critical battery operation scenarios.

The novelty and contributions of this paper are summarized as follows.

- In order to accommodate the varied operating temperature scenarios in real-world applications, the transfer learning-based DNN model (DNN-TL) model was proposed for the estimation of the SOH of Li-ion batteries by combining transfer learning with DNN models. It has been demonstrated that the accuracy of SOH estimation at 35 or 45 degrees Celsius has been improved over the stand-alone DNN model by utilizing the knowledge gained from DNN models pre-trained on EIS data at other temperatures using the DNN-TL model.
- The amount of data needed for retraining the DNN model has also been determined, with 80%, 50%, 35%, 20%, and 10% of target data being utilized for retraining the DNN model, respectively. The results demonstrated that the target dataset with only 6% of the size of source data is enough to retrain the DNN model using transfer learning. As well, the DNN model with transfer learning has shown less sensitivity to various amounts of target data than the DNN model without transfer learning.
- Additionally, the effectiveness of the transfer of features from the source task to the target task at each layer has been examined. When dealing with various operating temperature conditions, the results indicate that the transfer of four hidden layers from the pre-trained DNN to the DNN-TL model is essential in maintaining a stable and accurate SOH estimation.

The remainder of the paper is organized as follows: Section 2 summarizes the recent state-of-the-art research work regarding the SOH estimation of Li-ion batteries. Section 3 describes the utilized EIS dataset and the proposed transfer learning framework. Section 4 highlights the results and comparisons obtained from the employed models with the EIS feature set. Finally, the conclusion is drawn in Section 5.

2. Literature review

There have been a variety of approaches reported to provide accurate capacity estimates, and these approaches can be roughly divided into three categories, namely model-based, data-driven, and transfer learning-based approaches.

Model-based methods are highly dependent on the domain knowledge of the multi-physics phenomena of Li-ion batteries, including electrochemistry and aging characteristics. The use of model-based methods involves the development of complex mathematical models that are designed to account for the long-term dependencies of battery degradation and to describe the degradation process. Model-based methods are developed based on the battery modeling, which includes equivalent circuit models (ECMs) and electrochemical models (EMs) [6]. The Kalman filter-based models (KFs) have been proposed by many researchers, and most of them are based on the ECMs of lithium-ion batteries, where the battery has been characterized by resistance and RC circuits. One of the main limitations of using such a model will be the over-simplification of the battery since chemical interactions and battery aging will significantly alter the battery parameters. The identification of the battery parameters in ECMs is also dependent upon the battery experiments that have been implemented, which presents the challenge of considering all possible battery operation scenarios, especially for electric vehicles. As a result, model-based approaches will lose accuracy as the battery state changes. The ECMs take into account only a limited number of variables and parameters when formulating a mathematical description of battery aging, for instance, battery voltage, current and temperature, which allows them to provide sufficient precision while maintaining a relatively low computational complexity.

Thus, the ECMs [7] have been widely used to estimate Li-ion battery degradation. Using an ECM model, Daigle and Kulkarni [8] were able to describe how the key parameters of the battery change as battery degradation occurs, and predict the end of life of the battery through randomized discharged profiles. In [9], a dual ECM-based extended Kalman filter (EKF) model has been proposed for the estimation of the state of charge (SOC) and SOH of Li-ion batteries. By analyzing the drifts of open circuit voltage (OCV)-SOC curves at each state of battery aging, the effect of battery aging has been taken into account. Thus, the accuracy of SOC and SOH estimation has been improved by describing the correlation between SOC and SOH and identifying the updated battery parameters using recursive Least Squares (RLS) and parameter varying approach (PVA), respectively. The battery discharging curve was reproduced by using the parameters from the EIS measurement rather than using battery voltage measured from experiments as a basis for determining the internal impedance of the ECM model in [10]. There was a good correlation between the experimented discharging curve and the EIS-based ECM parameters that were used to reproduce the same discharging curve. As a result, the proposed ECM model was used to predict internal impedance and battery capacity when batteries were discharged with a maximum error of 3.73% for the SOH estimation. The extended Kalman filter (EKF) and an enhanced self-correcting equivalent circuit model were proposed by Plett in [11] to achieve an online capacity estimation of the Li-ion battery cell for both SOC and SOH estimations. In [12], An EKF was initially proposed to estimate the degradation of the Li-ion battery. In order to improve the performance of the EKF SOH estimator, the Sage-Husa adaptive algorithm and fuzzy controller were utilized to correct the state noise covariance and the observed noise covariance. According to experimental results, the adaptive dual extended Kalman filter-based fuzzy inference system (ADEKF-FIS) obtained a higher accuracy and faster convergence rate than the EKF algorithm.

Model-based estimators are highly dependent upon the accuracy of battery modeling. Consequently, EMs have been adopted as an alternative way to model the battery with more parameters added to reflect chemical interactions inside the battery than ECMs. The complexity of the EMs increases the accuracy of battery modeling. However, the difficulty of obtaining the parameters and computation burden of the partial differential equations involved in the EMs limits the applications of such model-based real-time SOH estimators in the BMS. In contrast to ECMs, where battery charging and discharging experiments are conducted to identify the battery parameters, EIS experiments were conducted extensively in EMs to identify additional parameters described by EMs [13], for instance, pseudo-two-dimensional model (P2D), and signal particle model (SPM) [14].

In order to study the capacity degradation under dynamic charge-discharge conditions, a P2D electrochemical lithium-ion battery model has been proposed in [15]. In addition to considering the heat generated within the battery, this study also considered the capacity fade and power fade caused by the solvent reduction reaction. In this electrochemical model, the electrochemistry, capacity fade, and their coupled effects with temperature were modeled to capture electrochemistry of the Li-ion battery. The P2D model was deployed in [16] to sufficiently describe the relevant temporal and spatial evolution of Li-ion concentration in each electrode. Two adaptation sliding mode observers (SMOs) were proposed for estimating Li-ion concentration and film resistance, as well as a parameter estimator for the solid-state diffusion coefficient of anodes. Results from simulations demonstrated the effectiveness of the proposed P2D for co-estimating the SOC and SOH.

P2D model is composed of partial differential equations (PDE) and nonlinear algebra with many parameters, hence the computation complexity limits its application for many real-world applications [17]. As a consequence, many scholars are actively seeking a reasonable and effective simplification method for P2D models, with the SPM being the most widely used [14]. It is assumed in the SPM that the liquid

phase concentration of lithium ions is constant throughout the battery, as well as that the solid phase's electric potential is constant throughout the electrode, thereby reducing the model complexity so that the SPM can be used in a wide range of applications. A trade-off between model complexity and model accuracy was achieved in [18] by simplifying the SPM from PDE to ordinary differential equations (ODE). The proposed SPM model was evaluated based on a comparison between ODE SPM and PDE SPM, as well as benchmark validation. The results demonstrated the effectiveness of the simplified SPM and the superiority of the proposed SOC/SOH estimation method. The enhanced single particle model (eSPM) has been proposed in [19] for estimating the remaining useful life (RUL) of Li-ion batteries based on vehicle charging data. In order to develop an eSPM that can predict the evolution of parameters related to battery aging, battery data collected from an aging study conducted on LMO-NMC-cathode graphite-anode batteries was used. Results of the experiment demonstrated that it was possible to infer battery SOH and RUL from charging data readily available in plug-in battery electric and hybrid vehicles using SPMs.

As mentioned earlier, model-based methods involve the development of complex mathematical models in order to account for the long-term dependencies of battery degradation and to describe the degradation process. The lack of domain knowledge regarding model construction, however, prevents the use of these methods in real-world applications due to the difficulty of identifying all the hidden complex and highly non-linear degradation features. As opposed to model-based approaches, data-driven models make use of machine learning techniques to provide an accurate estimation of the SOH of Li-ion batteries, thereby overcoming the limitations of model-based approaches for different types of batteries. Data-driven models can be applied in different scenarios without requiring an analysis of the battery mechanism system. Since there is no need to analyze the battery system mechanism, data-driven approaches only depend on empirical data. The effectiveness and richness of input data play an essential role in constructing a robust data-driven model for estimating the capacity of Li-ion batteries, which motivates a significant number of researchers to develop data-driven models for estimating the degradation of Li-ion batteries. Based on input features, such as battery terminal voltage, current, and ambient temperature from charging curves, feed-forward neural networks (FFNNs) have been employed in [20] to estimate the SOH of Li-ion batteries. As a result, the neural network can extract the dynamic characteristics of Li-ion batteries and correlate them with capacity by utilizing these input features. In [21], a gate recurrent unit-convolutional neural network (GRU-CNN) was proposed to estimate the state of health of Li-ion batteries. The network is capable of extracting shared information from the charging curve and reducing the maximum prediction error to 4.3%. Various CNN-based SOH estimation models have been developed and applied to the battery data derived from charging/discharging curves in [22], demonstrating the efficacy of the proposed models in predicting Li-ion battery SOH. In light of the fact that the terminal voltage of a battery suddenly drops at the end of a discharge cycle, it has been questioned whether charging/discharging curve-based features are reliable. In recent years, the EIS has garnered considerable attention due to its ability to accurately measure impedance across a wide range of frequencies and provide a comprehensive solution to characterize battery performance. A study conducted by [23] utilized EIS measurement data to estimate Li-ion battery state of charge, and the results showed that the EIS features led to improved accuracy and efficiency over current-voltage features. In [24], EIS measurement data have been adopted to train the machine learning models to estimate the state of charge of Li-ion batteries with an error rate of less than 5%. As a result of the failure to include a variety of temperatures and different SOC points in the original dataset, the model is not capable of dealing with dynamic scenarios. It has been proposed in [25] that a feature extraction method was proposed to decrease the computation burden by only considering highly correlated input features, as opposed to [24], in which the entire EIS impedance

spectrum was used to establish the mapping between impedance and SOC. This resulted in a relatively good accuracy achieved through the use of the Gaussian process regression model that had an error rate of less than 3.8%. Moreover, EIS measurement data have been combined with the cycle number of charging/discharging batteries to estimate the SOH of Li-ion batteries in [26], where simulation results indicated that regression models that incorporate cycle number as an input feature improve SOH prediction accuracy by up to 50% over those that rely exclusively on the EIS feature set. Using Nyquist plots of EIS measurements as an alternative method of representing EIS measurement data, the SOH of Li-ion batteries was estimated using a convolutional neural network in [27], where simulation results demonstrated the Nyquist plot of impedance provides more comprehensive information on battery aging than simple impedance values.

Even though the data-driven methods are capable of overcoming the drawbacks of model-based approaches, for example, the high dependency on battery domain knowledge, time-consuming battery experiments, and the inflexibility to apply when the battery types vary, they still have their shortcomings. The quality and availability of battery datasets make a substantial difference in model training and prediction since data-driven approaches learn patterns based on comprehensive battery testing data. In the real world, however, it can be challenging to obtain battery datasets of acceptable quality, and the EVs may be operated in various conditions, leading to the expectation that the data-driven system will be able to make adaptive adjustments to overcome critical conditions. A machine learning technique called transfer learning, which transfers previously learned experience to a similar task with only a small number of targets required, has garnered considerable attention in recent years. In [28], Firstly, a Convolutional Neural Network model was pre-trained on a large battery dataset and transferred to a small dataset of the targeted battery to improve its estimation accuracy. In the second step, a contribution-based neuron selection method was proposed to prune the transferred model using a fast recursive algorithm, which reduced its size and complexity while maintaining its performance. Using the proposed model, it is possible to estimate capacity in real-time, and its effectiveness has been tested on a data set collected from four lithium-ion phosphate batteries, and its performance has been compared to those of other convolutional neural network models. An innovative battery capacity prediction method based on TL and DNN has been proposed in [29]. It was conducted at three different ambient temperatures of 25 °C, 35 °C, and 45 °C in order to analyze the effects of EIS measurements directly, which are not dependent on the charging and discharging process and the IC curves. During the retraining of the DNN-TL, the number of fixed layers was also investigated. In [30], the authors proposed a novel LSTM-based transfer learning model for adaptive online prediction of battery capacity under fast charging with a novel feature constructed by integrating the 80% SOC charging voltage and the cycle number through SW technology. A transfer learning approach to battery capacity estimation has been proposed in [31], where a deep convolutional neural network model with transfer learning (DCNN-TL) was demonstrated to outperform a traditional machine learning method, Gaussian process regression, as well as the DCNN model without transfer learning. In [32], in order to pre-train a deep convolutional neural network (DCNN), ten-year cycling data from implantable Li-ion cells were used as the source dataset. This pre-trained DCNN was then transferred to a new DCNN model referred to as deep convolutional neural networks-transfer learning (DCNN-TL). DCNN-TL was then fine-tuned and retrained to produce accurate capacity estimation for a target dataset (NASA data). It has been demonstrated that in [33] using only small batch data, the proposed model was able to predict the SOH of a single battery cycle with a relative error of less than 2%. In addition, it can be used for estimating battery health in different loading modes by freezing the LSTM layer by using transfer learning. To evaluate the performance of lithium-ion batteries, a model-based transfer learning framework was proposed in [34]. Several comparison experiments

have been conducted using the NASA dataset in order to verify the effectiveness of the method. The experimental results indicated that the model-based transfer learning framework provided a more effective solution to the small sample problem arising in lithium-ion battery performance evaluation.

The novelty of applying transfer learning to the estimation of Li-ion battery states has resulted in a limited number of previous papers. To address the difficulty of collecting EIS measurement data in critical battery operating conditions, the transfer learning technique has been combined with DNN models in this work to estimate the degradation of Li-ion batteries using a reduced amount of EIS measurement data while maintaining a high estimation accuracy.

3. Material and methodology

In this section, the implementation of transfer learning with a deep neural network to estimate the SOH of Li-ion batteries at various temperatures will be discussed, and a general framework is presented in Fig. 1. There are shows two phases in the framework, the first one utilizes the source dataset to train the DNN model to estimate the SOH of Li-ion batteries at the temperature specified in the source dataset. However, when the temperature changes, the DNN model trained using the source dataset is not capable of estimating the SOH precisely, which results in the second phase, where the SOH of Li-ion batteries is expected to be estimated accurately at different temperatures. Transfer learning is implemented to fine-tune the trained DNN model from the source task, where the previously learned knowledge is expected to be transferred to estimate SOH at varied temperature settings. Additionally, a stand-alone DNN model, where all weights and biases are initially randomized, is trained using the target dataset in order to compare the transfer learning-based DNN model with the stand-alone model.

3.1. EIS dataset

As specified in [35], extensive EIS measurement experiments were conducted by continuously charging and discharging 12 Eunicell LR2032 lithium-ion coin cells, which are made of LiCoO₂/graphite. First, the battery cell is the minimal unit of the battery pack (e.g., from cells to modules and packs), where the cells are connected in series and parallel to increase the capacity of the battery to meet the requirements of various applications, for example, portable electronic devices, electric vehicles, and energy storage systems. As a result, investigating the degradation of battery cells offers detailed insight into the overall performance of a battery pack. As a real-world application, a BMS is used to record the battery data of each cell and balance the performance of each cell in order to maximize total battery performance. Low-capacity battery cells (e.g., coin cells (mAh)) are widely used in labs to simulate battery degradation as opposed to high-capacity cells (e.g., 52Ah LFP cells), which may have a cycle life of over a thousand cycles before reaching their end [10,34,35]. Second, EIS measurement with battery aging tests is difficult for high-capacity battery cells, especially when a high C-rate is required to accelerate the test. For example, a booster may be required for the EIS measurement devices to increase the maximum current they can provide, but in reality, it is extremely difficult to conduct the EIS measurement and cycle the high-capacity cells at high C-rates through an automated experiment setting in the laboratory environment. Due to the fact that coin cells require significantly fewer cycles to reach 80% of their rated capacity, the battery aging experiment has been accelerated significantly. A higher battery capacity results in a lower cell impedance when making a new battery since more anodes and cathodes are connected in parallel inside the battery cell, which explains why small cells have a higher impedance value than large ones.

Due to the difficulty of conducting battery aging experiments, researchers have evaluated their proposed prediction algorithms based on

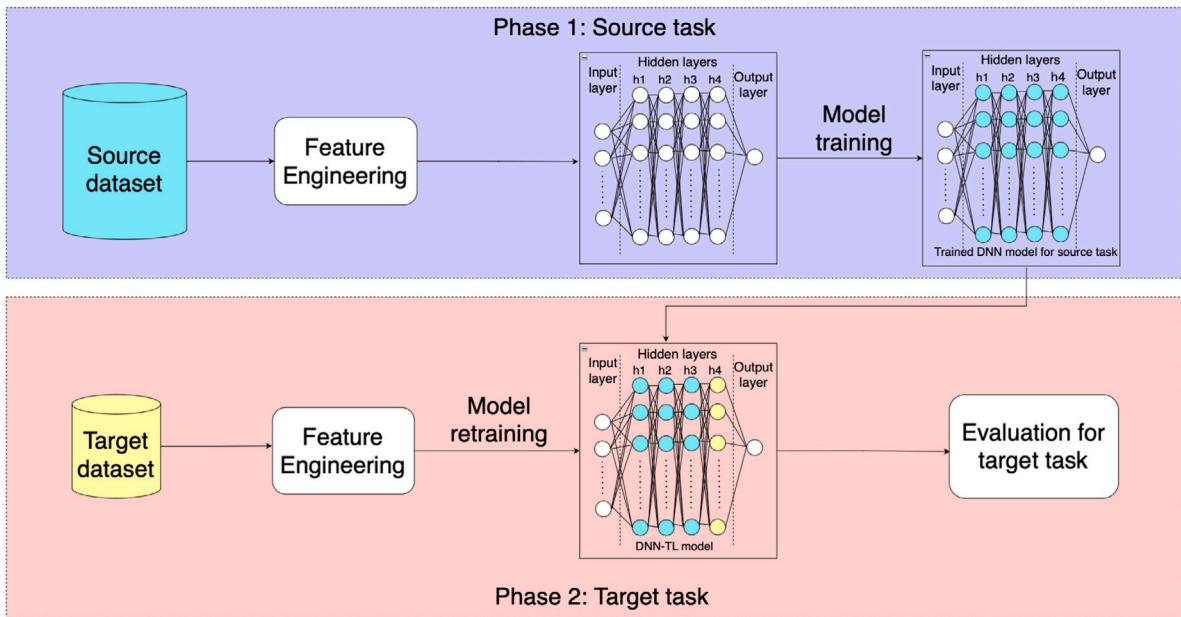


Fig. 1. The proposed transfer learning framework.

publicly available battery datasets. The degradation of Li-ion batteries is characterized by a variety of non-linear mechanisms and complex decline trajectories. It is essential to obtain reliable battery aging data in order to train a machine learning model that will accurately predict the SOH of Li-ion batteries. An important factor in determining battery performance is its internal impedance, which determines its operating voltage, rate capability, and efficiency, as well as its practical capacity. Generally, the measurement approach involves generating a sinusoidal current or voltage with a certain amplitude and frequency and measuring the output voltage or current's amplitude and phase shift. A characteristic impedance spectrum can be obtained by repeating this procedure at different frequencies, typically between KHz and mHz. There are more than 20,000 EIS spectra collected for 12 commercially 45 mA h Eunicell LR2032 lithium-ion coin cells, which are included in the EIS dataset [35]. The battery cells in Zhang's dataset [36] were cycled at different temperatures, specifically, eight cells were cycled at 25 °C, two cells at 35 °C, and the remaining cells were cycled at 45 °C. EIS measurement data are collected spanning a frequency range from 0.02 Hz to 20 KHz, where 60 sample points were conducted at each charging/discharging cycle.

3.1.1. Feature engineering

Identifying the change in the maximum available capacity of the battery when fully charged is critical to characterizing the battery degradation. To extract the EIS measurement data from the original dataset, state V, referred to 15 min of rest following a full charge of the battery cell, has been used. At each cycle (fully discharged and then charged to full), the collected EIS measurement data are composed of 60 complex values (real and imaginary parts), and each complex value is measured at a specific frequency. Therefore, the input feature at one specific cycle contains 120 different values, which are defined as EIS features. The output corresponding to the input features is the maximum capacity available at the time of the cycle. Table 1 summarizes the EIS measurement data collected in Zhang's dataset.

In addition, Table 2 presents the EIS measurement data summarized at different test temperatures.

In order to visualize EIS measurement data, Nyquist plots are commonly used to represent complex battery impedance values [37]. In Fig. 2, the real part of the impedance was plotted against the imaginary part of the impedance, each blue dot representing a complex value

Table 1

EIS measurement data.

Feature types	Impedance measured over the range from 0.02 Hz to 20 KHz (60 sample points)
Tested temperatures	25, 35, and 45 degree Celsius
Feature size	120 features (60 real parts and 60 imaginary parts)
Number of battery cells	12
Output	Battery capacity

Table 2

Data collected from EIS experiments at various temperatures.

Temperatures	Feature size	Number of samples	Number of cells tested
25 degree Celsius	120	1424	8
35 degree Celsius	120	598	2
45 degree Celsius	120	609	2

of the impedance, and 60 different blue dots were presented at each frequency point.

Several Nyquist plots with various SOH of Li-ion batteries were shown in Fig. 3, where the plot curves of impedance shift as a result of battery aging, which illustrated how the Nyquist plot reflects the degradation of Li-ion batteries.

3.1.2. Source and target datasets

EIS data collected at different temperatures show varied distributions owing to the fact that the battery cells in the dataset were tested at different temperature settings and the difference in each battery cell.

The machine learning approaches generally assume that the training and testing data share the same distributions [38]. However, in EV applications, the batteries are exposed to varying outside temperatures during operation, resulting in a variety of battery data distributions.

In order to simulate the distribution variations in battery data, the EIS dataset was divided into two different datasets, the source and target datasets. The EIS data and corresponding battery capacity were collected at 25 degrees Celsius in the source data, while the EIS dataset

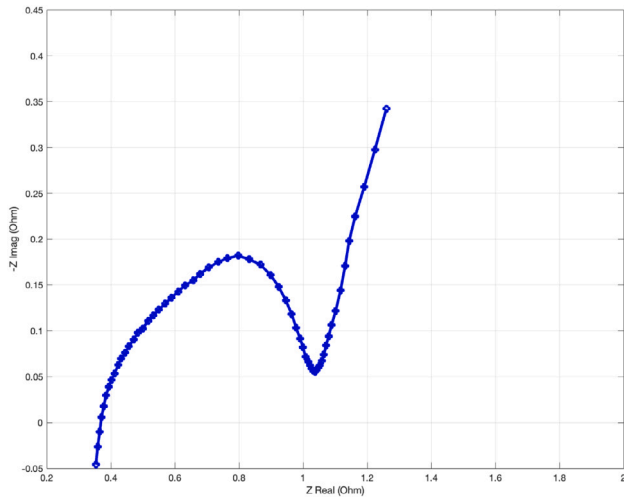


Fig. 2. Nyquist plot of impedance values.

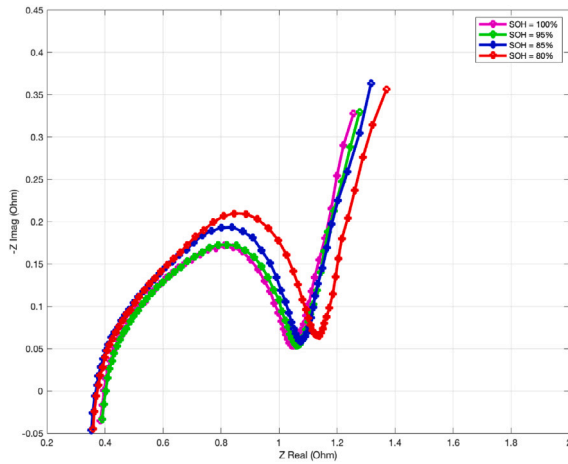


Fig. 3. Nyquist plots of impedance values at varied SOHs for a 45 mA h Eunicell LR2032 Li-ion coin cell.

and corresponding battery capacity were collected at 35 or 45 degrees Celsius in the target dataset. As a result of these differences in measured temperatures, there are different distributions of data. In Table 3, the source and target datasets are summarized, including the number of charging/discharging cycles tested in each battery cell, the battery test temperature, and the maximum capacity of each battery cell.

3.2. Source and target tasks

To investigate the effectiveness of transfer learning in estimating SOH at various temperature scenarios, the EIS measurement data collected at different temperatures will be grouped into source tasks and target tasks after the EIS features set has been selected and organized to reflect the degradation of Li-ion batteries. There are two types of tasks: source tasks and target tasks. Source tasks are tasks that the model was originally designed to address, whereas target tasks are tasks that were intended to be addressed by using the previously trained model from the source task. In order to simulate the ability of a trained SOH machine learning predictor at normal operating conditions (25 degrees Celsius) to predict the SOH under more critical operating conditions, such as 35 or 45 degrees Celsius, transfer learning is used to leverage

Table 3
Source and target datasets summary.

	Battery cell	Number of cycles	Collected temperature	Maximum capacity
Source dataset	1	200	25 °C	36 mA h
	2	250		
	3	229		
	4	81		
	5	275		
	6	212		
	7	140		
	8	37		
Target dataset 1	1	299	35 °C	40 mA h
	2	299		
Target dataset 2	1	299	45 °C	42 mA h
	2	310		

Table 4
Source and target tasks.

Source tasks	Target tasks
Pre-trained at 25 degree Celsius	Re-training at 35 degree Celsius
Pre-trained at 25 degree Celsius	Re-training at 45 degree Celsius
Pre-trained at 25 & 45 degree Celsius	Re-training at 35 degree Celsius
Pre-trained at 25 & 35 degree Celsius	Re-training at 45 degree Celsius

only a small amount of EIS measurement data at those temperatures in order to retrain the previously trained model.

SOH estimators trained on EIS measurement data at a particular temperature cannot be used to estimate the SOH of lithium-ion batteries at these very different temperatures. In reality, however, battery data in critical operation scenarios are not always available, resulting in the difficulty of training another SOH estimator from scratch.

In this work, the source tasks are to develop a trained SOH estimator that can estimate the SOH of Li-ion batteries under similar operating conditions. For example, a trained DNN model using EIS measurement data at 25 degrees Celsius will be capable of estimating SOH at 25 degrees Celsius in the future. However, the trained DNN model will lead to a significant error when it is used to estimate the SOH at different temperatures because the trained DNN model has never been trained on such data, which has been considered general limitations of machine learning techniques. Contrary to source tasks, the target tasks involve estimating the SOH of Li-ion batteries at various temperatures. In order to address that, there are two practical solutions. Considering the availability of battery data for target tasks, it is feasible to train a brand-new DNN model from scratch using the target data. Alternatively, the previously learned knowledge from a pre-trained DNN model based on source data can be transferred to the DNN-TL model retrained using target data. Table 4 shows a horizontal correlation between each source task and each target task. In source tasks, DNN models are pre-trained using source data at 25, 25 and 35, or 25 and 45 degrees Celsius, while in target tasks, the pre-trained DNN models obtained from source tasks are retrained using target data at 35 or 45 degrees Celsius to estimate the SOH at the respective temperatures.

3.3. Deep neural network

The DNN is one of the promising deep learning techniques, especially when a large amount of data is involved, which can be used to learn the dynamics and nonlinear degradation patterns of Li-ion batteries in order to perform regression-based SOH estimations of the batteries. In general, DNN models are simple feed-forward neural networks consisting of input, hidden, and output layers [39].

As a result of the selection of the number of layers and the number of neurons within each layer, there is a significant effect on the prediction performance of models, where it is important to carefully adjust the hyperparameters for a DNN model in order to achieve the

Table 5
DNN Hyperparameters.

Hyperparameter	Selection Options
Number of hidden neurons in first hidden layer	64
Number of hidden neurons in second hidden layer	32
Number of hidden neurons in third hidden layer	16
Number of hidden neurons in fourth hidden layer	8
Number of hidden layers	4
Optimization algorithm	Adam
Activation function for hidden & output layers	ReLU
Loss function	Mean squared error
Batch size	32
Validation split	0.1

best prediction performance. There are multiple neurons in the input layer of the DNN model, depending on the size of the input data, which are then sequentially fed into hidden layers to determine the correlation between the battery features and the estimated capacity of the battery. In order to characterize the nonlinearity of the battery degradation process, activation functions have been used both in the hidden and output layers. The weights and biases of the DNN model have also been adjusted using the Adam optimization algorithm that optimizes the loss value computed by the loss function in each iteration. Eventually, the DNN model that has been properly trained with battery degradation features and responses will be capable of predicting the state of Li-ion batteries. A deployed DNN model has been tuned based on the complexity of battery features and output. It has been identified that four fully connected hidden layers exhibit the best performance among a variety of hidden layer numbers. The detailed hyperparameters have been summarized in Table 5.

3.4. Transfer learning

The use of data mining and machine learning technologies has already yielded significant successes in a variety of areas of applied machine learning, including classification, regression, and clustering. In spite of this, many machine learning methods are only effective when both the training and test data originate from the same feature space and distribution. In order to stay adaptive to changes in training and testing data distributions, most statistical models must be reconstructed from scratch utilizing newly collected training data in order to accommodate changes in distributions [40]. This results in many real-world applications requiring recollecting target battery data and rebuilding estimation models, which can be very costly or even impractical. Transfer learning, which transfers knowledge from pre-trained machine learning models to address target tasks, has been utilized in deep neural networks to reduce the effort of recollecting data and improve the efficiency of retraining models.

3.4.1. Model pre-training for source tasks

In Fig. 4, the neurons in each hidden layer of the DNN model are represented in different colors, showing that the patterns of battery degradation learned by the DNN model can be transferred to a transfer learning-based DNN model using transfer learning. The blue neurons represent the knowledge acquired from the source task, which is represented as values of trained weights and biases and stored in hidden layers. The values stored in blue neurons in each hidden layer will remain unchanged when the neural networks are retrained. Each connection between neurons in each hidden layer has its own weights, which are continuously updated during the training phase using the Adam optimization algorithm until an optimal state is attained. In general, battery impedance increases as the battery ages. When pre-training the DNN model for SOH estimation, the DNN model learns the patterns from EIS measurement data and then maps the relationship between EIS and battery capacity. The information learned is then stored in the weights of the neurons in the hidden and output layers.

DNN models begin with randomly assigned weights for neurons, which are then tuned based on the training data. As a result of feeding the training data into the initial DNN model, the loss function is used to calculate the difference between the estimated battery capacity and the actual battery capacity. In addition, the loss value is optimized using an optimization algorithm for the backpropagation process. Once the DNN model has been trained for the source task, the weights stored in the neurons will reflect the best combination of parameters to describe the degradation of the battery. As a result, a trained model for EIS measurement at a specific temperature is obtained, which is defined as a pre-trained model for source task estimation.

3.4.2. Model re-training for target tasks

Using EIS measurement data at a different temperature than the EIS measurement data from the source task to estimate the SOH of Li-ion batteries is a great example of how the pre-trained DNN models can be re-trained using newly collected measurements at the target temperature for the target task when the target task differs from the source task. Fig. 4 shows all the possible re-training scenarios, where 0 fixed layer refers to the fact that all the weights of the neurons in the hidden layers can be updated from previously stored values during the model re-training phase. In addition, when the number of fixed layers equals four, all the weights in the neurons remain unchanged during the retraining process. Considering the similarity between the target and source tasks, only a few layers need to be fine-tuned to transfer previously acquired knowledge to the target task, as discussed in the result discussion. In order to achieve optimal performance, previously learned knowledge from the source task is preserved by fixing weights in hidden layers, thereby improving the efficiency of model training. By learning from newly collected data, the tunable weights in the hidden layers enable the accuracy to be improved for the target task. As a result, a trade-off between model accuracy and efficiency must be evaluated in terms of the number of fixed layers for the SOH estimation.

3.5. Evaluation metrics

To evaluate the performance of transfer learning when transferring previously learned knowledge to the target task, four different regression evaluation metrics have been used to evaluate the models' performance under different scenarios, including mean squared error (MSE), mean absolute error (MAE), R-squared, and mean absolute percentage error (MAPE).

3.5.1. Mean squared error

MSE is defined as the mean of the square of the difference between actual and estimated values in statistics. In our case, it is calculated by taking the mean of the square of the difference between the true capacity $y_{i_{true}}$ and the estimated one $y_{i_{estimated}}$ as shown in:

$$MSE = \sum_i^n \frac{(y_{i_{estimated}} - y_{i_{true}})^2}{n} \quad (2)$$

The smaller the MSE, the better the estimation accuracy.

3.5.2. Mean absolute error

MAE is the average of the summation of the absolute difference between actual and predicted values. In our case, it is calculated by taking the mean of the sum of the absolute value of the difference between estimated capacity $y_{i_{estimated}}$ and actual capacity $y_{i_{true}}$, which is shown in:

$$MAE = \sum_i^n \frac{|y_{i_{estimated}} - y_{i_{true}}|}{n} \quad (3)$$

The smaller MAE value indicates a better estimation result.

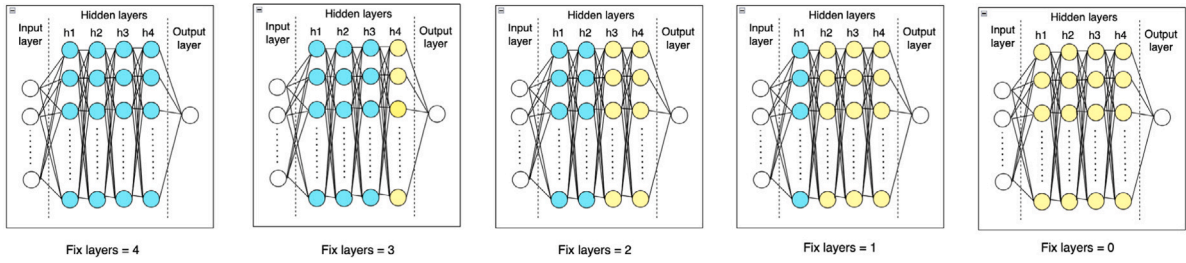


Fig. 4. DNN models with varied numbers of fixed layers.

3.5.3. R-squared

R-squared, or coefficient of determination, is another statistic used to measure regression model performance. It measures the amount of variance in the output that can be explained by the model's inputs. R-squared can be expressed mathematically as the ratio between the sum of squares regression (SSR) and the sum of squares total (SST). SSR refers to the total variation of all the predicted values from the mean value of all the values of response variables, while SST refers to the total variation of actual values from the mean value of all the values of response variables. The equation of R-squared has been shown in:

$$R^2 = \frac{SSR}{SST} = \frac{\sum_i^n (y_{i_{estimated}} - y_{i_{true}})^2}{\sum_i^n (y_{i_{true}} - y_{i_{average}})^2} \quad (4)$$

The greater R-squared value indicates a better estimation result.

3.5.4. Mean absolute percentage error

MAPE is another evaluation metric to evaluate the machine learning regression problem. It reflects the average percentage difference between predictions and their intended targets in the dataset. The formula for MAPE is defined as follows:

$$MAPE = \frac{1}{n} \sum_i^n \left| \frac{y_{i_{estimated}} - y_{i_{true}}}{y_{i_{true}}} \right| \quad (5)$$

4. Results and discussion

A data-driven approach to estimating the SOH of Li-ion batteries has the advantage of using neural networks to learn a correlation between EIS measurements and decreased battery capacity. Thus, it is an end-to-end estimation methodology that does not require the battery to be modeled in ECMs or EMs, which improves computation efficiency and prevents parameter identification errors in battery modeling. Data-driven approaches, especially for neural networks, have the general disadvantage of being difficult to interpret, particularly when a large number of hidden layers are involved. For instance, there are over 10,000 tunable parameters in the deployed DNN-TL models, it is not practical to assign meaning to each neuron in order to correlate the weights with the electrochemical composition of a battery. This section describes the results of the implementation of transfer learning with DNN models to estimate the SOH of Li-ion batteries at a variety of temperatures. As a first step, the effectiveness of transfer learning will be evaluated by comparing it with a stand-alone DNN model in which all weights and biases have been initially randomized. Moreover, the amount of target data that will be required to make an accurate estimation using transfer learning will also be analyzed. Following that, the number of fixed layers in DNN-TL will be examined with regard to transfer learning to determine how the knowledge learned from the source dataset influences the estimation of the target task.

4.1. Investigation on varied temperatures

In the source task, the DNN model was trained using source data where the EIS measurement data were measured and collected at 25 degrees Celsius, while in target tasks, the objective is to estimate the

SOH of Li-ion batteries at 35 or 45 degrees Celsius, respectively. This experiment was conducted to see how transfer learning can improve the SOH estimation at 35 or 45 degrees Celsius by fine-tuning the trained DNN model from the source task. With the same hyperparameters as the trained DNN model for the source task, but with all weights and biases initially randomized, a stand-alone DNN model was also trained on the same target dataset to draw the performance comparison. The results of DNN-TL and the stand-alone DNN model will be compared and discussed in detail with respect to two different target tasks, when there are no fixed layers in the DNN-TL and the target data are split into 80% for training and 20% for testing.

4.1.1. Estimate the SOH of Li-ion batteries at 35 degrees Celsius

The purpose of this experiment is to investigate how transfer learning improves SOH estimation accuracy when the estimation is performed at a different temperature. The DNN from the source task, where it was trained and tested to estimate the SOH of Li-ion batteries at 25 degrees Celsius, has been retained and used with the new data collected at 35 degrees Celsius to estimate the SOH at 35 degrees Celsius. Meanwhile, a stand-alone DNN model has been trained from scratch using the same target data collected at 35 degrees Celsius to show the performance comparison. Table 6 summarizes the results of SOH estimation using DNN-TL and stand-alone DNN models when there are no fixed layers in the DNN-TL and the target data is split into 80% for training and 20% for testing for both the DNN-TL and stand-alone DNN models. Given that the DNN-TL model has previously been trained on the source dataset at 25 degrees Celsius, transfer learning enables the pre-trained DNN-TL model to transfer the learned knowledge at 25 degrees Celsius to estimate the SOH at 35 degrees Celsius, which has been compared with the stand-alone DNN model that has only been trained on target data when the temperature is 35 degrees Celsius. The results demonstrated that the accuracy for SOH estimation at 35 degree Celsius has been improved up to 31.49% for MSE, 16.77% for MAE, 1.2% for R-squared, and 19.26% for MAPE by transferring the previously learned knowledge at 25 degree Celsius. The improvement rate in Tables 6, 7, 8 and 9 refers to the improvement in accuracy for DNN-TL models compared to DNN models (stand-alone DNN) trained using source EIS data from scratch. One of the major differences between the DNN-TL model and the stand-alone DNN model is that the weights in DNN-TL are pre-trained using EIS data from the source task, whereas the stand-alone DNN model has its weights randomized without any previous knowledge of the source task. Thus, the stand-alone DNN models have been used as a benchmark to demonstrate the performance improvement of DNN-TL models in estimating SOH from EIS measurement data when the temperature varies from source to target tasks.

In order to determine the generality of the DNN-TL model, the pre-trained DNN model using EIS features collected at 25 and 45 degrees Celsius was retrained using target data collected at 35 degrees Celsius and compared with the stand-alone DNN model. The results are summarized in Table 7 in terms of MSE, MAE, R-squared, and MAPE. As a result, the pre-trained DNN-TL model is capable of estimating the SOH of Li-ion batteries at 35 degrees Celsius when pre-trained on source data collected at 25 and 45 degrees Celsius. And the accuracy of the SOH

Table 6

Performance comparison between pre-trained DNN-TL at 25 degrees Celsius and stand-alone DNN when estimating SOH at 35 degrees Celsius.

	MSE	MAE	R ²	MAPE
DNN-TL	0.2796	0.3846	0.9768	0.0119
Stand-alone DNN	0.4082	0.4621	0.9652	0.0147
Improvement rate	31.49%	16.77%	1.2%	19.26%

Table 7

Performance comparison between pre-trained DNN-TL at 25 & 45 degrees Celsius and stand-alone DNN when estimating SOH at 35 degrees Celsius.

	MSE	MAE	R ²	MAPE
DNN-TL	0.1117	0.2423	0.9907	0.0076
Stand-alone DNN	0.4082	0.4621	0.9652	0.0147
Improvement rate	72.63%	47.56%	2.64%	48.43%

Table 8

Performance comparison between pre-trained DNN-TL at 25 degrees Celsius and stand-alone DNN when estimating SOH at 45 degrees Celsius.

	MSE	MAE	R ²	MAPE
DNN-TL	0.0903	0.2258	0.9885	0.0064
Stand-alone DNN	0.1299	0.2858	0.9828	0.0080
Improvement rate	30.47%	20.99%	0.57%	20.19%

estimation has been improved be up to 72.63% for MSE, 47.56% for MAE, 2.64% for R-squared, and 48.43% for MAPE, over the stand-alone DNN model that has been only trained on target data at 35 degrees Celsius. Further, the additional 45 degrees Celsius in the source dataset makes a difference in the estimation of SOH at 35 degrees Celsius, compared to the DNN-TL that is only pre-trained on the source dataset at 25 degrees Celsius.

4.1.2. Estimate the SOH of Li-ion batteries at 45 degrees Celsius

In the last section, the results have been summarized and discussed for SOH estimation at 35 degrees Celsius for the target task based on a comparison between the DNN-TL model and a stand-alone DNN model. In this section, similar implementations have been conducted for both the DNN-TL and stand-alone DNN models when the DNN-TL were pre-trained at different temperatures to estimate the SOH of Li-ion batteries at 45 degrees Celsius. This experiment aims to investigate the generality of the proposed DNN-TL model by varying the diversity of the source datasets (varied EIS measurement temperatures) utilized by the pre-trained models. The DNN-TL model was trained on source data at 25 degrees Celsius, then retrained using transfer learning on target data at 45 degrees Celsius to estimate the SOH of Li-ion batteries at 45 degrees Celsius. To evaluate DNN-TL performance on the target task, a stand-alone DNN model was trained from scratch using the same target data at 45 degrees Celsius. As shown in Table 8, MSE, MAE, R-squared, and MAPE were used to evaluate the performance of DNN-TL and standalone DNN models.

The results demonstrated that the DNN-TL leveraging the previous experience from source data at 25 degrees Celsius outperformed the stand-alone DNN model without previous knowledge. And the SOH estimation accuracy at 45 degrees Celsius has been improved up to 30.47% for MSE, 20.99% for MAE, 0.57% for R-squared, and 20.19% for MAPE.

Similarly, the DNN-TL model that has been previously trained on the EIS feature set at 25 and 35 degrees Celsius, has been retained using target data collected at 45 degrees Celsius to estimate the SOH of Li-ion batteries at 45 degrees Celsius. Furthermore, it has been compared with a stand-alone DNN model trained solely on target data at 45 degrees Celsius. Results were summarized in Table 9, which demonstrated that the DNN-TL improved SOH estimation accuracy by utilizing the previously acquired knowledge derived from the source data at 25 and 35 degrees Celsius.

Table 9

Performance comparison between pre-trained DNN-TL at 25 & 35 degrees Celsius and stand-alone DNN when estimating SOH at 45 degrees Celsius.

	MSE	MAE	R ²	MAPE
DNN-TL	0.0266	0.1288	0.9968	0.0036
Stand-alone DNN	0.1299	0.2858	0.9828	0.0080
Improvement rate	79.51%	54.93%	1.4%	55.11%

The results indicated that the MSE, MAE, R-squared, and MAPE for SOH estimation at 45 degrees Celsius have been improved up to 79.51%, 54.93%, 1.4%, and 55.11%, respectively. Moreover, compared with the DNN-TL having only the EIS features at 25 degree Celsius in the source dataset, the additional 35 degree Celsius source data provided the DNN-TL with a better insight into battery degradation and enhanced the prediction accuracy and generalizability of the DNN-TL model.

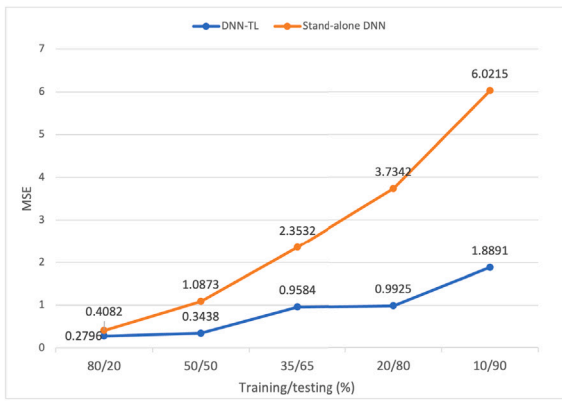
4.2. Investigation on the amount of target data needed

Previously, it was shown that transfer learning improved the accuracy of SOH estimation when Li-ion batteries were operated at temperatures different from those from the source dataset. In the real world, it can be difficult to collect battery data at critical temperatures; therefore, it is important to examine the efficiency and effectiveness of using transfer learning when the amount of target data is limited. In order to validate the effectiveness of transfer learning, the target data were split into training and testing at different splitting rates. In addition to the DNN-TL models, stand-alone DNN models have also been tested in order to compare their performance.

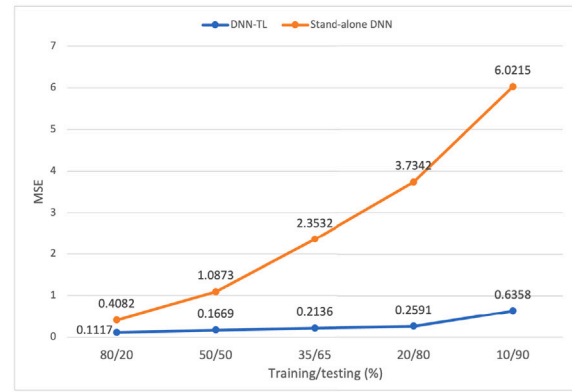
To evaluate the performance of DNN-TL and stand-alone DNN models under different target data splitting, Fig. 5(a) shows the DNN-TL model that has been previously trained on source data at 25 degrees Celsius and then retrained using target data at 35 degrees Celsius. A DNN-TL model which was previously trained using source data at 25 degrees and 45 degrees Celsius and then retrained using target data at 35 degrees Celsius was shown in Fig. 5(b). In addition, a stand-alone DNN model with all weights and biases initially randomized has been trained from scratch using target data at 35 degrees Celsius to estimate the SOH at 35 degrees Celsius. The performance of the stand-alone DNN model is also shown in Fig. 5 for comparison with the DNN-TL model.

In both scenarios, DNN-TL models which leverage the previously learned knowledge from the source task, are less sensitive to the amount of target data used to retrain the model, compared with the stand-alone DNN model. When both the EIS features at 25 and 45 degrees Celsius were involved in the source dataset, only 20% target data were enough to make the DNN-TL capable of estimating the SOH of Li-ion batteries at 35 degrees Celsius, which are 120 samples in the target dataset. In contrast, when only EIS features at 25 degrees Celsius are included in the source dataset, 50% of the target data will be sufficient for retraining the DNN-TL model to estimate the SOH of Li-ion batteries at 35 degrees Celsius, which is 299 samples in the target dataset.

Implementations have also been performed to evaluate the amount of target data required for DNN-TL and stand-alone DNN models to estimate the SOH at 45 degrees Celsius Using source data at 25 degrees Celsius or both 25 and 35 degrees Celsius, DNN-TL models were pre-trained and then retrained on target data at 45 degrees Celsius to estimate the SOH at 45 degrees Celsius, while the stand-alone DNN models were trained from scratch using target data at 45 degrees Celsius. Fig. 6 illustrates the results for those two DNN-TL models under different target data splittings. The performance of DNN-TL as shown Figs. 5(b) and 6(b) have been improved compared to the DNN-TL models in Figs. 5(a) and 6(a) due to the difference of amount and the diversity of source data used to pre-train the DNN-TL models in source tasks. For instance, in Figs. 5(a) and 6(a), only EIS measurement data

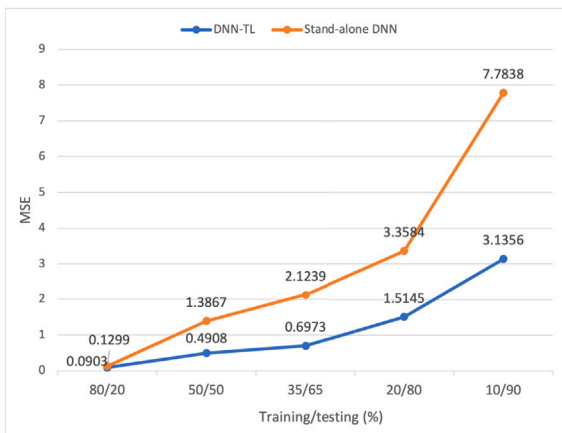


(a) Pre-trained DNN-TL at 25 degrees Celsius

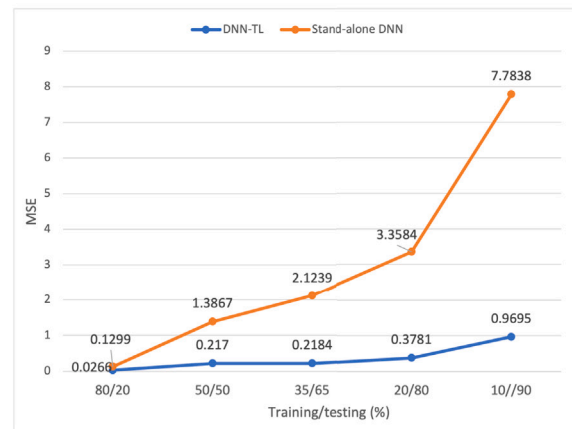


(b) Pre-trained DNN-TL at 25 & 45 degrees Celsius

Fig. 5. DNN-TL and stand-alone DNN models when estimating SOH at 35 degrees Celsius.



(a) Pre-trained DNN-TL at 25 degrees Celsius



(b) Pre-trained DNN-TL at 25 & 35 degrees Celsius

Fig. 6. DNN-TL and stand-alone DNN models when estimating SOH at 45 degrees Celsius.

at 25 degrees Celsius (1424 samples) were used to pre-train the DNN-TL models, while in Figs. 5(b) and 6(b), the DNN-TL models were pre-trained using EIS measurement data at 25 and 35 (1424+598 samples) or 25 and 45 degrees Celsius (1424+609 samples).

The results in Fig. 6 clearly demonstrate that the DNN with transfer learning is more robust than the stand-alone DNN model, especially when the amount of target data is limited. Moreover, when EIS features at 25 and 35 degrees Celsius are included in the source dataset, the DNN-TL exhibits greater robustness than a DNN-TL trained solely on data at 25 degrees Celsius. Additionally, only 20% of the target data (122 samples in the target dataset) are sufficient to retrain the DNN-TL model for estimating the SOH at 45 degrees Celsius, whereas the DNN-TL model previously trained on source data at 25 degrees requires 50% target data (305 samples in the target dataset) for maintaining accuracy in prediction.

4.3. Investigation on the number of fixed layers

The effectiveness of transfer learning in estimating the SOH at varied temperatures, as well as its robustness when the amount of target data is limited, have been demonstrated in previous sections. In the context of battery aging estimation with neural networks, the initial layers of a deep neural network, particularly when using transfer learning, are designed to capture foundational electrochemical features and principles learned from the source task. These initial layers are responsible for acquiring knowledge related to electrode materials behavior, charge/discharge profiles, and the impact of operating conditions on

battery performance. Based on the results of the implementations, DNN-TL has consistently demonstrated its best performance when no fixed layer is involved in the transfer learning, with all weights and biases in the DNN-TL being updated based on the prior knowledge learned from the source task. By contrast, the weights and biases of a stand-alone DNN model are always initially randomized.

Additionally, the results indicate that the DNN-TL is capable of maintaining an acceptable level of prediction accuracy when the number of fixed layers is less than 3, but when the number of fixed layers exceeds 2, the prediction error increases exponentially. It is this electrochemical knowledge from the EIS measurement dataset that provides the basis for accurate battery state estimation. As the network deepens, subsequent layers may introduce more abstract or specialized representations. However, they often build upon the fundamental electrochemical insights gained in the earlier layers. Therefore, the first two layers of the DNN-TL model may contain valuable information from the source task that provides the core electrochemical understanding essential for accurate battery analysis and prediction. This electrochemical foundation from the EIS measurement dataset informs the model's ability to make meaningful inferences about battery state and performance. To visualize the difference in performance between DNN-TL and DNN models with different numbers of fixed layers with respect to different source and target tasks, Figs. 7 and 8 have been presented.

4.3.1. Pre-trained DNN-TL on EIS features at 25 degrees Celsius

As shown in Fig. 7(a), the DNN-TL was previously trained on source data at 25 degrees Celsius and then retrained with target data at 35

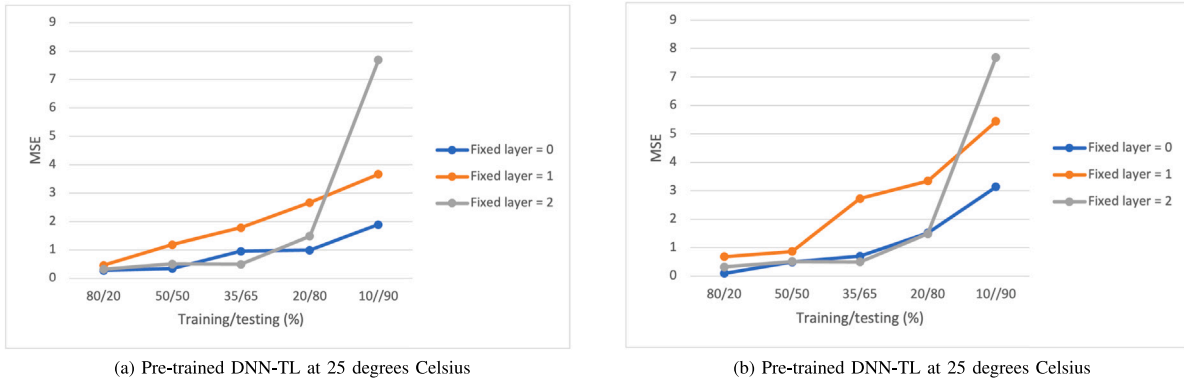


Fig. 7. DNN-TL models with different numbers of fixed layers.

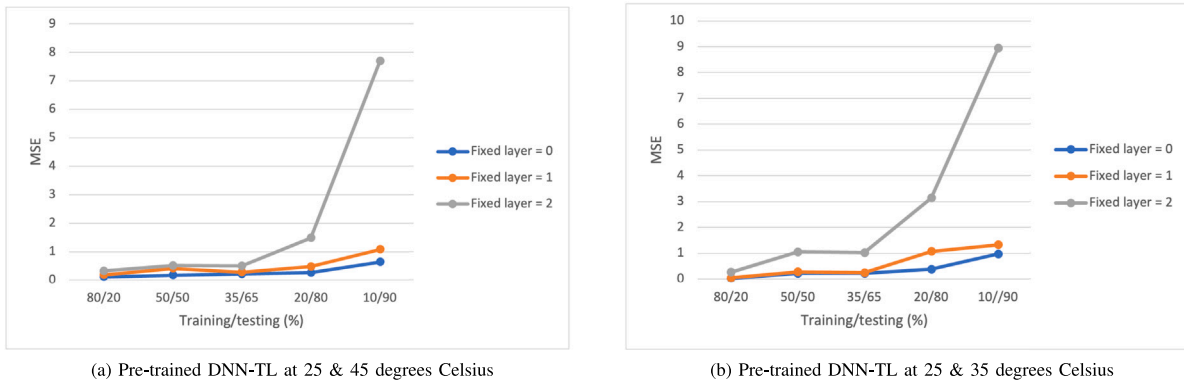


Fig. 8. DNN-TL models with different numbers of fixed layers.

degrees Celsius in order to estimate the SOH at 35 degrees Celsius, whereas Fig. 7(b) shows the DNN-TL that was previously trained on source data at 25 degrees Celsius and retrained using target data at 45 degrees Celsius to estimate the SOH at 45 degrees Celsius.

The figure illustrates that when the amount of target data used for training is more than 20%, 0 fixed layer and 2 fixed layers exhibit similar performance, otherwise the DNN-TL model with 0 fixed layer was more robust. Figs. 7(a) and 7(b) illustrate that the DNN-TL models are more robust to various data splitting and more accurate when the first two layers are fixed (gray line) during the retraining process as compared with only the first layer being fixed (orange line). DNN-TL model has the most accurate and robust performance when there is no fixed layer in the DNN-TL, which indicates all the information learned and stored from the source task can be reused for the target task. Consequently, it is evident that the first two layers of the DNN-TL model contain more useful information previously learned from the source task. Also, when only the first layer of DNN-TL is fixed, the DNN-TL cannot estimate the SOH for the target task accurately and robustly.

4.3.2. Pre-trained DNN-TL on EIS features at 25 and 35, or 25 and 45 degrees Celsius

In contrast, Figs. 8(a) and 8(b) show the performance difference when the DNN-TL models were pre-trained at 25 and 45 degrees Celsius or 25 and 35 degrees Celsius, respectively. When there is no fixed layer in the DNN-TL models, the re-trained DNN-TL models maintain a stable and accurate SOH estimation under various data splitting for model retraining. The results indicated that the information previously learned and stored in the first two layer of the DNN-TL model contains essential battery degradation information when the DNN-TL were pre-trained using EIS measurement data at 25 and 35 or 45 degrees Celsius.

As opposed to Fig. 7, where the pre-trained DNN-TL models were only able to learn from EIS measurements at 25 degrees Celsius, Fig. 8 indicates a better overall performance of DNN-TL models when dealing with a limited amount of target EIS measurement data, due to the pre-trained DNN-TL models that learned battery degradation patterns from dynamic EIS measurement temperatures. Additionally, the DNN-TL with 0 fixed layers slightly outperformed the DNN-TL with the first layer fixed when there are EIS features at two different temperatures in the source dataset.

5. Conclusion

Generally, machine learning approaches assume that training and testing data share the same distributions. However, Li-ion batteries in EV applications are exposed to varying temperatures during operation. It results in a variety of battery data distributions that need to be taken into consideration when estimating the SOH of Li-ion batteries. The first part of this study involves applying the transfer learning-based DNN model to estimate the SOH of Li-ion batteries using EIS measurement data at a different operation temperature than those previously learned by the DNN model. Results showed that DNN-TL models that utilized previously learned knowledge from source data outperformed the stand-alone DNN model trained from scratch. Moreover, in real-world applications, battery data in critical operating scenarios can be difficult to collect, which results in the investigation of the amount of target data needed to retrain the DNN-TL using transfer learning. According to the results, the target dataset, which was only 6.0% of the source dataset in size, was sufficient to retrain the DNN model using transfer learning. Additionally, the effectiveness of the transfer of features from the source task to the target task has been examined at each layer. The results indicated that the DNN-TL model is most effective when there is no fixed layer during retraining.

For future investigations, more critical scenarios for Li-ion batteries in EVs should be considered as a means of investigating the performance and robustness of transfer learning. Additionally, the feasibility of integrating automatic EIS measurements with dynamic charging and discharging profiles for high-capacity battery cells should be investigated to reduce the difficulty in collecting data on battery aging. Furthermore, different types of batteries have been used in EVs due to the development of battery technologies, which presents a challenge for data-driven state estimation.

CRedit authorship contribution statement

Yichun Li: Conceptualization, Methodology, Software, Writing – original draft. **Mina Maleki:** Conceptualization, Methodology, Writing – review & editing. **Shadi Banitaan:** Conceptualization, Methodology, Writing – review & editing.

Declaration of competing interest

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

Data availability

Data will be made available on request.

References

- [1] K.S. Ng, C.S. Moo, Y.P. Chen, Y.C. Hsieh, Enhanced coulomb counting method for estimating state-of-charge and state-of-health of lithium-ion batteries, *Appl. Energy* 86 (9) (2009) 1506–1511.
- [2] X. Hu, L. Xu, X. Lin, M. Pecht, Battery lifetime prognostics, *Joule* 4 (2) (2020) 310–346.
- [3] K.A. Smith, C.D. Rahn, C.Y. Wang, Control oriented 1D electrochemical model of lithium ion battery, *Energy Convers. Manag.* 48 (9) (2007) 2565–2578.
- [4] S.J. Moura, Estimation and control of battery electrochemistry models: A tutorial, in: 2015 54th IEEE Conference on Decision and Control, CDC, 2015, pp. 3906–3912.
- [5] S.J. Pan, Q. Yang, A survey on transfer learning, *IEEE Trans. Knowl. Data Eng.* 22 (10) (2009) 1345–1359.
- [6] I. Babaeiyazdi, A. Rezaei-Zare, S. Shokrzadeh, State of charge prediction of EV Li-ion batteries using EIS: A machine learning approach, *Energy* 223 (2021) 120116.
- [7] G. Saldaña, J.I. San Martín, I. Zamora, F.J. Asensio, O. Oñederra, Analysis of the current electric battery models for electric vehicle simulation, *Energies* 12 (14) (2019) 2750.
- [8] M. Daigle, C.S. Kulkarni, End-of-discharge and end-of-life prediction in lithium-ion batteries with electrochemistry-based aging models, in: AIAA Infotech@Aerospace, 2016, p. 2132.
- [9] S. Tong, M.P. Klein, J.W. Park, On-line optimization of battery open circuit voltage for improved state-of-charge and state-of-health estimation, *J. Power Sources* 293 (2015) 416–428.
- [10] M. Galeotti, L. Cina', C. Giammanco, S. Cordiner, A. Di Carlo, Performance analysis and SOH (state of health) evaluation of lithium polymer batteries through electrochemical impedance spectroscopy, *Energy* 89 (2015) 678–686.
- [11] G.L. Plett, Extended Kalman filtering for battery management systems of LiPB-based HEV battery packs: Part 3. State and parameter estimation, *J. Power Sources* 134 (2) (2004) 277–292.
- [12] K. Yang, Z. Chen, Z. He, Y. Wang, Z. Zhou, Online estimation of state of health for the airborne Li-ion battery using adaptive DEKF-based fuzzy inference system, *Soft Comput.* 24 (2020) 18661–18670.
- [13] A.M. Bizeray, J.H. Kim, S.R. Duncan, D.A. Howey, Identifiability and parameter estimation of the single particle lithium-ion battery model, *IEEE Trans. Control Syst. Technol.* 27 (5) (2018) 1862–1877.
- [14] J. Zhou, B. Xing, C. Wang, A review of lithium ion batteries electrochemical models for electric vehicles, in: E3S Web of Conferences, Vol. 185, EDP Sciences, 2020, p. 04001.
- [15] T.R. Ashwin, Y.M. Chung, J. Wang, Capacity fade modelling of lithium-ion battery under cyclic loading conditions, *J. Power Sources* 328 (2016) 586–598.
- [16] S. Dey, B. Ayalew, P. Pisu, Combined estimation of state-of-charge and state-of-health of Li-ion battery cells using SMO on electrochemical model, in: 2014 13th International Workshop on Variable Structure Systems, VSS, IEEE, 2014, pp. 1–6.
- [17] J. Tian, R. Xiong, W. Shen, A review on state of health estimation for lithium ion batteries in photovoltaic systems, *ETransportation* 2 (2019) 100028.
- [18] Z. Cen, P. Kubiak, Lithium-ion battery SOC/SOH adaptive estimation via simplified single particle model, *Int. J. Energy Res.* 44 (15) (2020) 12444–12459.
- [19] K.K. Sadabadi, X. Jin, G. Rizzoni, Prediction of remaining useful life for a composite electrode lithium ion battery cell using an electrochemical model to estimate the state of health, *J. Power Sources* 481 (2021) 228861.
- [20] H. Chaoui, C.C. Ibe-Ekeocha, H. Gualous, Aging prediction and state of charge estimation of a LiFePO4 battery using input time-delayed neural networks, *Electr. Power Syst. Res.* 146 (2017) 189–197.
- [21] 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 (2020) 101741.
- [22] 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 (2022) 103857.
- [23] Y. Li, M. Maleki, S. Banitaan, M. Chen, Data-driven state of charge estimation of Li-ion batteries using supervised machine learning methods, in: 2021 20th IEEE International Conference on Machine Learning and Applications, ICMLA, IEEE, 2021, pp. 873–878.
- [24] M. Messing, T. Shoa, R. Ahmed, S. Habibi, Battery SoC estimation from EIS using neural nets, in: 2020 IEEE Transportation Electrification Conference & Expo, ITEC, IEEE, 2020, pp. 588–593.
- [25] I. Babaeiyazdi, A. Rezaei-Zare, S. Shokrzadeh, State of charge prediction of EV li-ion batteries using EIS: A machine learning approach, *Energy* 223 (2021) 120116.
- [26] Y. Li, M. Maleki, S. Banitaan, M. Chen, State of health indicator modeling of lithium-ion batteries using machine learning techniques, in: 2022 IEEE International Conference on Electro Information Technology, EIT, IEEE, 2022, pp. 440–445.
- [27] Y. Li, M. Maleki, S. Banitaan, M. Chen, State of health estimation of lithium-ion batteries using convolutional neural network with impedance Nyquist plots, 2023.
- [28] Y. Li, K. Li, X. Liu, Y. Wang, L. Zhang, Lithium-ion battery capacity estimation—A pruned convolutional neural network approach assisted with transfer learning, *Appl. Energy* 285 (2021) 116410.
- [29] I. Babaeiyazdi, A. Rezaei-Zare, S. Shokrzadeh, Transfer learning with deep neural network for capacity prediction of Li-ion batteries using EIS measurement, *IEEE Trans. Transp. Electr.* (2022).
- [30] Z. Chen, W. Shen, L. Chen, S. Wang, Adaptive online capacity prediction based on transfer learning for fast charging lithium-ion batteries, *Energy* 248 (2022) 123537.
- [31] S. Shen, M. Sadoughi, C. Hu, Online estimation of lithium-ion battery capacity using transfer learning, in: 2019 IEEE Transportation Electrification Conference and Expo, ITEC, IEEE, 2019, pp. 1–4.
- [32] S. Shen, M. Sadoughi, M. Li, Z. Wang, C. Hu, Deep convolutional neural networks with ensemble learning and transfer learning for capacity estimation of lithium-ion batteries, *Appl. Energy* 260 (2020) 114296.
- [33] L. Yao, J. Wen, S. Xu, J. Zheng, J. Hou, Z. Fang, Y. Xiao, State of health estimation based on the long short-term memory network using incremental capacity and transfer learning, *Sensors* 22 (20) (2022) 7835.
- [34] G. Zou, Z. Yan, C. Zhang, L. Song, Transfer learning with CNN-LSTM model for capacity prediction of lithium-ion batteries under small sample, *J. Phys. Conf. Ser.* 2258 (1) (2022) 012042.
- [35] Y. Zhang, Q. Tang, Y. Zhang, J. Wang, U. Stimming, A.A. Lee, Identifying degradation patterns of lithium ion batteries from impedance spectroscopy using machine learning, *Nat. Commun.* 11 (1) (2020) 1706.
- [36] Yunwei Zhang, Qiaochu Tang, Yao Zhang, Jiabin Wang, Ulrich Stimming, Alpha A. Lee, Identifying Degradation Patterns of Lithium Ion Batteries from Impedance Spectroscopy using Machine Learning [Data Set], Zenodo, 2020, <http://dx.doi.org/10.5281/zenodo.3633835>.
- [37] E. Barsoukov, J.R. Macdonald, Impedance Spectroscopy Theory, Experiment, and Applications, second ed., John Wiley & Sons, Inc, Hoboken, NJ, 2005, p. 2005.
- [38] S.J. Pan, Q. Yang, A survey on transfer learning, *IEEE Trans. Knowl. Data Eng.* 22 (10) (2009) 1345.
- [39] G. Montavon, W. Samek, K.R. Müller, Methods for interpreting and understanding deep neural networks, *Digital Signal Process.* 73 (2018) 1–15.
- [40] T. Tommasi, F. Orabona, B. Caputo, Learning categories from few examples with multi model knowledge transfer, *IEEE Trans. Pattern Anal. Mach. Intell.* 36 (5) (2013) 928–941.