Comparative Analysis of Neural Networks Techniques for Lithium-ion Battery SOH Estimation

Comparative Analysis of Neural Networks Techniques for Lithium-ion Battery SOH Estimation

Alessandro Aliberti*, Filippo Boni*, Alessandro Perol‡, Marco Zampolli‡, Rémi Jacques Philibert Jaboeuf‡,
Paolo Tosco‡, Enrico Macii* and Edoardo Patti*,
*Politecnico di Torino, Turin, Italy. Email: name.surname@polito.it
‡Edison, Italy, Email: name.surname@edison.it

***Abstract*—Li-ion batteries have become the most important technology for electric mobility. One of the most pressing challenges is the development of reliable methods for battery state-of-health (SOH) diagnosis and estimation of remaining useful life. In electric mobility scenario, battery capacity degradation prediction is crucial to ensure service availability and life duration. This research work provides a comprehensive comparative analysis of neural networks for a data-driven approach suitable for SOH estimation on single cells, stressed under laboratory conditions. For this purpose, different neural networks (i.e., LSTM, GRU, 1D-CNN, CNN-LSTM) are trained and optimized on NASA Randomized Battery Usage dataset. Experimental results demonstrate that data-driven neural networks generally performed well SOH estimation on single cells. In detail, the 1D-CNN best predicts SOH and has the lowest variance in the output. The LSTM have the highest variance in estimating SOH, while GRU and CNN-LSTM tend to overestimate and underestimate the value of SOH, respectively.**

***Index Terms*—Li-ion battery, State-of-Health, Electric Vehicles, neural networks, Deep Learning**

I. Introduction

According to the European Union technical report [1], transport is one of the major responsible for greenhouse gas emissions in Europe. As a result, many countries are promoting the shift to electric vehicles (EVs) to address the emerging challenges of emissions reduction and the low-carbon economy. Indeed rechargeable battery and lithium-ion batteries have become the standard for energy storage thanks to their high energy and power density, low memory effects, and long service life. Along with the technological innovation in lithium-ion battery cell chemistry, great attention is now being paid to battery monitoring, estimation, and control technologies [2], to accurately assess battery capacity degradation and power degradation, the consequences of cell component degradation. Batteries manufacturers exploit a special electronic system called Battery Management System (BMS) that evaluates the health state of the battery over time. Indeed, BMS provides important information such as state-of-charge (SOC) and battery state-of-health (SOH). The SOC describes the difference between a fully charged battery and the same battery in use and it is defined as the ratio of the remaining charge in the battery, divided by the maximum charge that can be delivered by the battery [3], as in the following Eq. 1:

$$SOC = \frac{C_{actual}[Ah]}{C_{available}[Ah]} * 100\% \tag{1}$$

where C_{actual} is the residual capacity of an operating battery and $C_{available}$ is the maximal releasable capacity of a fully charged battery, when discharged at a constant reference current. In this work the $C_{available}$ is approximated with C_{rated} (defined in Eq. 2). While the research area of SOC estimation has been widely explored, the field of SOH is still very challenging since SOH does not correspond to any particular physical quality. The most commonly used SOH definition is expressed by a percentage of capacity loss with respect to the nominal capacity of the battery [4], described in Eq. 2:

$$SOH = \frac{C_{available}[Ah]}{C_{rated}[Ah]} * 100\% \tag{2}$$

where C_{rated} is the nominal capacity measured by the manufacturer on fresh cells discharged at constant current and will decline with the used time.

A robust SOH estimation is critical to ensure battery life duration and quantify the residual market value of batteries for safe and reliable second-life applications with lower power requirements. Moreover, it impacts also the estimation of service availability (i.e. km range for EVs), as SOH and SOC are strongly dependent one to another. Most of the existing approaches estimate SOH under controlled laboratory conditions and static charge/discharge conditions [5] that are significantly different from the working circumstances of real electric vehicles (i.e. in terms of temperature, rate of charge/discharge, depth of discharge, etc). An efficient solution suitable for real-world use cases needs to be i) *real-time*, since the battery SOH provides important information about the actual capacity of the battery and helps to estimate the available driving mileage of the EV; ii) *dynamic*, battery aging is affected by many components such as charge/discharge current, temperature, and voltage, which are very dynamic in a real-world application.

The objective of this research is to compare fully data-driven methodology based on artificial neural network (NN) meeting the real-time and dynamic requirements by exploiting the non-linear relationship between the battery parameters of voltage, current, temperature, and SOH. To find the best solution, we investigated and compared different neural architectures, respectively a Long Short-Term Memory (LSTM), a Gated Recurrent Unit (GRU), a One-dimensional Convolutional (1D-CNN), and a CNN-LSTM hybrid neural network. All the designed architectures are trained, tested and optimized exploiting the NASA Randomized Battery Usage dataset [6].

The final purpose is to verify which of the compared models best fits the complex physical behaviors that characterize a battery with respect to individual cells.

II. Related Works

Nowadays, battery SOH estimation represents a very challenging task. Literature presents a variety of studies and methodologies, which mainly fall into three macro-categories [7]: i) direct measurement-based methods, ii) model-based methods and iii) data-driven methods. The direct measurement-based methods rely on collecting data and measurements that can be used to understand and evaluate the battery capacity and internal resistance, and then obtain the SOH accordingly to Eq. 2. Two of the most used direct measurement methods are the *Coulomb Counting* [8] and *Impedance Spectroscopy* [5]. These methods are reliable and accurate, but they are highly dependent on the laboratory equipment and the stationary/controlled conditions of the test laboratory experiments, making them suitable only for off-line and off-board use.

In the model-based approach, the ageing process of batteries is evaluated by developing a physical battery model that can represent the battery characteristics. Two commonly used battery models are the equivalent circuit model (ECM) and the electrochemical model (EChM). The ECM model aims to determine parameters such as voltage, current and temperature using basic electrical elements (i.e. resistors, capacitors and voltage sources). The ECM model can be used for real-time on-board processing due to its simple implementation, but on the other hand, it is not able to express the complex internal dynamics of the battery, resulting in low accuracy [2], [4]. The EChM model is a model based on the complex physical laws that govern the internal chemical dynamics of the battery. Due to its high mathematical complexity and computational cost, this model is not suitable for a real-time application on the BMS logic [2], [4].

Data-driven methods for battery SOH estimation are gaining popularity thanks to their flexibility and their ability to establish nonlinear relationships among battery parameters by exploiting machine learning algorithms (ML). One of the main advantages of this approach is that the models are based only on collected data, without requiring expert knowledge of the degradation phenomena. Moreover, their design is suitable for on-board applications and can be easily executed on the BMS hardware. On the other hand, the drawbacks are not negligible: to achieve high accuracy, they require a large dataset of the same battery; the training is computationally intensive and they lack in ability to generalize to other types of batteries, different in chemistry and shape.

The main focus in recent years within the data-driven approach has been on deep learning (DL) methods, which have been shown to work well with time series [9]. In the battery SOH estimation domain, battery parameters of voltage (V), current (I), and temperature (T), which are measured by sensors and sampled at very high frequency during charge and discharge cycles, represent large time series data that are typically passed to NNs via sliding windows, i.e., the NN makes an estimate based on the features at time t plus their last N values. Vanilla Feedforward and Recurrent NNs, Convolutional Neural Networks (CNN) [10], Long Short-Term Memory (LSTM) [11], Gated Recurrent Units (GRU) [12], and more complex structures, such as Bidirectional Gated Recurrent Unit (Bi-GRU) [13] and Gate Recurrent Unit-Convolutional Neural Network (GRU-CNN) [14], have been explored in literature to find the best way to map the time series characteristics of charge voltage, current and temperature to the battery SOH, usually obtaining appreciable results. However, most of the reported works adopted collection of data derived from laboratory tests where battery cells were stressed under fixed charge/discharge profiles and fixed temperatures, which does not reflect the dynamic and random experience of real-world EV driving. This lack of generalisation from the controlled conditions of a lab test to the real world of EVs was highlighted in [4], where researchers used an advanced deep learning neural network, called Independent Recurrent Neural Network (IndRNN), capable of better capturing the nonlinear properties of lithium-ion batteries and a dataset collected in the lab by applying random charge/discharge profiles to the batteries to more closely approximate the dynamics of the real world.

Inspired by [4], the proposed work exploits the randomized battery usage dataset from the National Aeronautics and Space Administration's (NASA) Ames Prognostic Data Repository [6] to train and optimize deep models that reflect real-world electric vehicle driving patterns. Therefore, the most common NNs in time series analysis, i.e. LSTM, GRU, CNN, CNN-LSTM, are trained, tested, optimized and compared on the NASA dataset. The final objective is to provide a comprehensive comparative analysis of neural network for a data-driven approach suitable for on-board SOH estimation of EVs batteries.

III. Methodology

This Section describes the proposed methodology for SOH estimation. Figure 1 depicts the main stages: i) *Data Preprocessing*, ii) *Model Training and Optimization* and iii) *Performance Evaluation*. In *Data Preprocessing* phase, NASA Randomized Battery Usage Dataset was cleaned from outliers; battery features were extracted and data augmentation tecninque was applied to obtain a bigger dataset. Finally, data were split into training- and test-set. In *Model Training* phase, we trained and optimized the 1D-CNN, LSTM, GRU and CNN-LSTM. The hyper-parameters of the architectures were adjusted using the validation data. In the *Performance Evaluation* stage, the trained models were tested using the test set exploiting widely-used performance indexes (i.e. Root mean Squared Error (RMSE), Mean Absolute Error (MAE) and Coefficient of Determination (R2)).

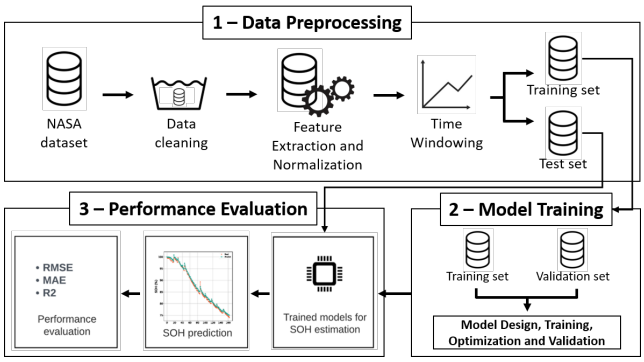

Fig. 1. Overview of the proposed process for SOH estimation

A. NASA Dataset and Data Preprocessing

In this research work, we exploit a dataset provided by NASA Ames Prognostics Center of Excellence Randomized Battery Usage Repository [6]. In [6], authors synthesized the whole dataset to better represent the practical operating conditions of a real EV battery cell by applying a dynamic load profile consisting of different battery cycles with randomly selected charge and discharge currents. This is ideal for training a data-driven model that is intended to be suitable for on-board use and for the dynamics of real electric cars. The tests were performed with 28 battery cells divided into 7 groups, and each group of cells was subjected to a different randomized cycling procedure, such as different types of randomized currents and temperatures. Cell specifications of NASA dataset are the following: i) the manufacturer is LG CHEM; ii) the form factor is 18650; iii) the chemistry is Lithium cobalt oxide vs. graphite; iv) the nominal capacity, the lower cut-of voltage and the upper threshold voltage are respectively 2.10 Ah, 3.2 V and 4.2 V.

For this work, we selected the cells that belong to *Group 3 (Cells 9-12)* since the charge/discharge profiles applied are the most representative of an EV usage in terms of randomness. The *Group 3* NASA dataset operating profiles operated using a sequence of charging/discharging currents between -4.5A and 4.5A. Each loading period lasted 5 minutes. Reference characterization carried out after 1500 periods (about 5 days). Group 3 contains the test results of four cells (identified as RW9, RW10, RW11, RW12) that were continuously operated using a sequence of charging and discharging currents, randomly selected among 12 values between -4.5A and 4.5A, at laboratory temperature conditions (negative values represent charge operations, positive ones represent discharge).

Then, in *Data Preprocessing* phase, the dataset is split into train- and test-set.

1) Data Cleaning: Focusing on RW9, RW10, RW11 and RW12 cells, we started removing noise and outliers. As reported by authors in [6], the cells profiles present two points where the capacity of the cell increases abnormally. The reason for this unexpected behaviour of the cell is not explained in the experimental guidelines, but since they occur for every cell in the same cycle, we decided not to discard them because they may reflect internal dynamics of the cell that are relevant to our task. However, RW11 has two noise spots at the end of the experiment (cycles 35 and 36), which we removed because they are likely due to measurement errors, since they only occur for this cell.

2) Feature Extraction and Normalization: NASA dataset presents measure of voltage (V), current (C), and temperature (T) sampled every second, resulting in random walking steps, rest steps, and reference charge and discharge profiles represented by vectors containing all measurements at 1-second intervals. The records of the resting steps are not taken into consideration in this work because the cell is not loaded during this 1-second period, so no information about its SOH can be derived. Since the battery capacity is used for the SOH estimation, the relationship between the input parameters and the battery capacity must be established for each RW step in order to build an accurate estimation model. For example, for each RW step in each cell profile, the measured parameters V_i, I_i, and T_i are mapped to a single multi-dimensional input vector x, that has the same dimension for each cell profile. In this experiment, we identified four most meaningful factors that contribute to SOH estimation. For each k-th step, the first three elements of the input vector are:

$$x_{k,1} = V_{avg}(k) = \frac{1}{n}\sum_{i=1}^{n} V(k)_i \tag{3}$$

$$x_{k,2} = I_{avg}(k) = \frac{1}{n}\sum_{i=1}^{n} I(k)_i \tag{4}$$

$$x_{k,3} = T_{avg}(k) = \frac{1}{n}\sum_{i=1}^{n} T(k)_i \tag{5}$$

where *n* represents the number of measurements in each step, which is typically 210, *k* represents the number of RW steps in each cell data set, $V(k)_i$, $I(k)_i$, $T(k)_i$ represent the voltage, current, and temperature, respectively, recorded each second during the step *k*.

The fourth element of the feature vector is the total time of the RW step (recall that the default duration of the step is 5 minutes, but can be interrupted if the upper/lower voltage thresholds are exceeded):

$$x_{k,4} = \Delta t(k) = t(k)_n - t(k)_1 \tag{6}$$

where $t(k)_n$ and $t(k)_1$ represent the time measured at the end and beginning of the RW step *k*.

Finally, the last element needed to build the estimation model, is the SOH associated to each RW step. The capacity of the cell at cycle *i* is computed and then divided by the nominal capacity of cell (i.e. 2.10 Ah), obtaining the SOH for the i-th cycle. However, reference discharges are run only at the end of every RW cycle, so that no SOH values will be available for the 1500 RW steps in between two reference discharges. To solve this problem and take advantage also of the steps measurements, we applied a linear interpolation of the SOH value between two reference discharges. The resulting 1500 SOH values per RW cycle can be directly associated with the 1500 RW step input vectors, obtaining a supervised learning

dataset. Linear interpolation is a form of data augmentation, techniques used to increase the amount of data. By using this method, we have gone from a very short dataset to a much larger one where all the steps can be used to create the model. The cell dataset are now structured as follows: i) RW9: 56340 supervised records, i) RW10: 54964 supervised records, i) RW11: 54242 supervised records, and i) RW12: 54514 supervised records. The adopted feature extraction approach is inspired by [4], where the authors use the RW cycles as "resolution" for feature extraction, identifying 18 input variables as representative for the entire cycle. Data normalization is a very important step in data preparation for DL models where raw data values are scaled to a particular range. Therefore, the normalization process of min-max is applied to scale the input sequence values in a range from 0 to 1, where the difference between the values is not distorted. The raw data are normalized using Eq. 7, where t refers to the t-th feature of the dataset, $\bar{x_t}$ is normalized data, x_t is unscaled data, while min_t and max_t are minimum and maximum values of the t-th feature in the input dataset:

$$\bar{x}_t = \frac{(x_t - min_t)}{max_t - min_t} \tag{7}$$

3) Time Windowing: Cell measurements are considered as multivariate time series that can describe the aging characteristics of batteries at different scales [15], the time step in our study being the RW step. The information provided by the cell dataset is reshaped into fixed time windows of fixed size that provides the model with the most complete information possible at a given time in the recent past to achieve accurate prediction.

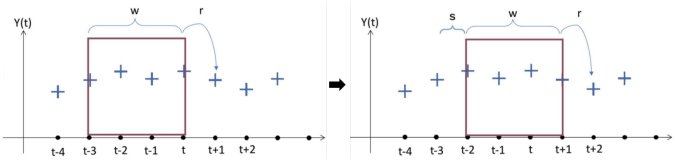

Fig. 2. Sliding windows mechanism on a univariate time series

Figure 2 explains the time window mechanism used for time series data: *w* represents the fixed time window step, *r* represents the response (the number of steps to predict), and *s* refers to the stride (how many steps go forward in generating the next time window). In this work, *r* and *s* are fixed values set to 1 and 20, respectively, while the dimension of the time window *w* is a hyperparameter that is optimized during the training of the Neural Networks.

B. Neural Networks

The proposed methodology exploits and compares different neural networks suitable for time-series applications, focusing on their usage in the Li-ion batteries world [16].

1) LSTM: Long Short-Term Memory belongs to recurrent neural networks (RNNs), neural architectures famous to allow information to persist. In RNNs, the output of each neuron depends not only on the current input, but also on the history of previous outputs of the hidden state. Depending on the number of time steps, RNNs can efficiently retain information about the past. This makes them suitable for time series prediction applications, such as SOH estimation, where the current output is predicted based on a history of the previous output and the current input over time [17]. Unfortunately, simple RNNs have been shown to lose their predictive capabilities when processing data with long-term dependencies, due to the vanishing gradient problem [18]. This problem occurs in neural networks trained with gradient-based methods and back-propagation, where the partial derivative can become smaller and smaller as it propagates through the layers, causing the neural network to train very slowly or never converge. Therefore, LSTMs were introduced to solve the problem of long-term dependencies [19]. They were specifically designed to overcome the vanishing gradient problem typical of RNNs. They avoid the decay of the gradient during back propagation by allowing the gradient to flow freely backward in time using certain types of gates that can regulate the flow of information. These gates contribute to the construction of the cell state, a "transport highway" that transmits relative information all the way down the sequence chain.

2) GRU: The gated recurrent unit is a special type of optimized recurrent neural network based on LSTM. The internal GRU unit is similar to that of the internal LSTM, except that the GRU combines the input gate and the forget gate in LSTM into a single update gate. They were first introduced by Junyoung Chung et al. in [20] to simplify the LSTM structure and its complicated and time-consuming internal state updates.

3) 1D-Convolutional Neural Network: 1D-convolutional neural networks are a class of Convolutional Neural Networks (CNN). CNN is a deep learning model for processing data composed by a grid pattern, such as images. It is inspired by the organization of animal visual cortex and is designed to automatically and adaptively learn spatial hierarchies of features, from low-level to high-level patterns [21]. A convolution operation is performed by sliding a window of weights over a matrix, where an output value generated at each position is a weighted sum of the input values covered by the window. The weights that parameterize the window remain the same throughout the scanning process. Therefore, convolutional layers can capture the displacement invariance of different patterns and learn robust features. CNN's convolution operations are usually exploited with spatial or 2-D data. However, for time series, instead of extracting spatial information, 1D-CNN extracts information along the time dimension by moving a fixed-size kernel.

4) CNN-LSTM: With respect the characteristics described in Section III-B1 and Section III-B3, this model was designed to combine the best aspects of the two models for multiple time series analysis. This hybrid model consists of a CNN model to extract important features from the observed time series and an LSTM to predict the future values, based on the described features [22]. The CNN-LSTM hybrid model consists of an input layer, two convolutional layers to capture features (i.e. in our case voltage, temperature, current, and relative time

input signals), a pooling layer to downscale the map of output features, and a flattening layer to convert these feature maps into one-dimensional vectors as input to the LSTM layer. The LSTM layer models the long-term dependencies in the feature maps, and the full connected layers make predictions according to these extracted features.

C. Models Training

In the model training phase, the training-set has been further split into two parts, 90% for training and 10% for validation, respectively. All models were trained using the Keras library with the TensorFlow backend on the Google Colab platform.

After a comprehensive campaign of trial-and-error, we found the best architecture for all models and we set some common parameters as follows: i) the batch size is 128, ii) the epochs are 50, iii) the optimizer is Adam, iv) the learning rate is 0.001, and v) time-lag sets to 250, 500, 750, 1000, 1500, respectively. We focus especially to find the best *time-lag* for each architecture. Indeed, many RW steps are required to observe a noticeable SOH degradation of the cells. However, increasing the input time-lag too much to get much more information from the past would lead to overfitting and high computational cost and training time, due to the high number of parameters.

Figure 3 depicts in detail the final optimized structure, for each neural model, obtained after the trial-and-error hyper-parameters and structure layer investigation. The symbol *?* refers to the time-lag dimension and *4* is the number of input features, namely voltage, current, temperature and time. The activation function for the Conv1D layers is *relu*. For the proposed LSTM and GRU, we selected *tanh*. Finally, for the 1D-CNN layers and LSTM layers we selected *relu* and *tanh*, respectively.

IV. Results

In this section, we discuss the experiments performed to assess the prediction models. To evaluate the prediction accuracy of the models, we exploited a set of metrics that are often used in SOH estimation: i) the Root Mean Square Error ($RMSE$) that measures the differences between values predicted and the values observed; ii) the Mean Average Error (MAE) that represents the average absolute difference between the actual and predicted values and is used to assess the accuracy of the SOH estimate; iii) the Coefficient of Determination R^2 used to verify the goodness of fit of the outcome w.r.t. the true regression line.

For any neural network, the performance of the training model is mainly determined by its hyperparameters, such as the number of hidden layers, the number of neurons in each hidden layer, the activation function, the optimizer, the learning rate, and the number of training epochs. These hyperparameters must be defined before the training process begins. However, for the models defined in this work, the focus is on applying and analyzing the performance of the different neural networks for SOH estimation under different time-lag sizes. The performance evaluation criteria are considered in determining the best time-lag size for each architecture. For each time-lag, SOH estimation on the test set was performed and compared to the true SOH values.

TABLE I
NETWORKS PREDICTION RESULTS ON THE *test_set*, I.E., RW9.

Model	Time-lag	Metrics		
		$RMSE$	MAE	R^2
1D-CNN	250	0.058	0.047	0.874
	500	0.066	0.056	0.833
	750	0.041	0.032	0.936
	1000	0.06	0.051	0.858
	1500	**0.035**	**0.028**	**0.951**
LSTM	250	0.068	0.053	0.811
	500	0.066	0.052	0.834
	750	0.065	0.05	0.822
	1000	**0.064**	**0.049**	**0.841**
	1500	0.067	0.051	0.818
GRU	**250**	**0.046**	**0.037**	**0.921**
	500	0.063	0.052	0.852
	750	0.084	0.073	0.732
	1000	0.096	0.086	0.639
	1500	0.065	0.055	0.829
CNN-LSTM	250	0.059	0.048	0.873
	500	0.051	0.041	0.904
	750	0.046	0.037	0.92
	1000	0.047	0.038	0.912
	1500	**0.04**	**0.033**	**0.933**

All networks are trained based on the *training-set* and the model weights and optimizer status of the best epoch are stored based on the *validation loss*. Once the training is complete, the best epoch model is loaded and tested using the *test-set*. Finally, the $RMSE$, MAE and R^2 are calculated. Table I shows all experimental results.

The best time-lag for the 1D-CNN is 1500 RW step, with the lowest values among all models, both in terms of metrics and test phase duration, indicating a robust and lightweight architecture. The $RMSE$ is 0.035 (3.5% of SOH), the MAE is 0.028 (2.8% of SOH), the R^2 is 0.95. The best time-lag of the LSTM is 1000 RW steps, with 0.064 $RMSE$ (6.4% of SOH), 0.049 MAE (4.9% of SOH). As it can be observed in Table I, all LSTM time-lag sizes are quite close to each other in metrics results. The main difference is in the complexity of the network, which increases significantly with the input delay leading to a duration increase. For the GRU, the best results are obtained with the 250 RW steps time-lag, with an $RMSE$ of 0.046 (4.6% of SOH), an MAE of 0.037 (3.7% of SOH), which increases with the time-lag and a R^2 of 0.92. Finally, the best results for the CNN-LSTM are observed with 1500 RW steps: 0.04 $RMSE$ (4% of SOH), 0.033 MAE (3.3% of SOH). In the CNN-LSTM, considerable improvements in metrics are observed with increasing time-lag, indicating a good ability to extract meaningful features (thanks to the CNN action) and retain past memory (thanks to the LSTM).

According to the results presented, all networks performed well in estimating SOH, with great help from the data augmentation applied to the datasets (Section III-A2). However, among all the tested models, the 1D-CNN seems to be the most promising, as it has achieved the best results in all metrics compared to all the other networks. The better predictive capability of the 1D-CNN can also be seen in the Figure 4,

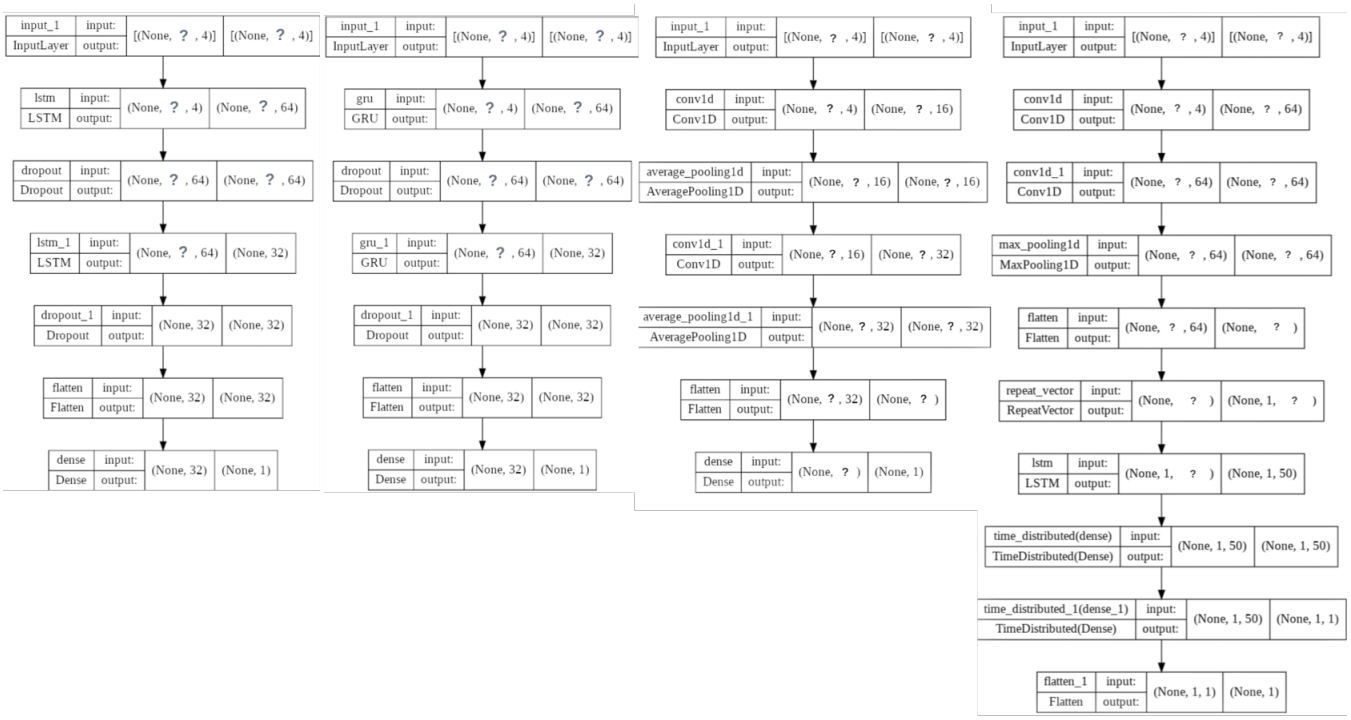

Fig. 3. Proposed models implementation details.

which compares the test results for each network with the benchmark SOH given by the reference discharges values (the predictions on the *test-set* was performed with their best time-lag configuration, that can be find bolded in Table I).

The Figure 4 shows that the 1D-CNN best approximates the real SOH with the highest R^2 and it has the lowest variance in the output. The LSTM shows the highest variance in estimating SOH, while GRU and CNN-LSTM tend to overestimate and underestimate the value of SOH, respectively.

V. Discussion

This comparative analysis research demonstrates that data-driven neural networks generally performed well in lithium-ion battery SOH estimation on single cells stressed under laboratory conditions. In detail, 1D-CNN outperforms all the other proposed and optimized neural networks, w.r.t. the exploited performances index. This work wants to represent a valid starting point to overcome the limitation in realizing an efficient data-driven model capable of estimating the state of an EV's battery. Indeed, the lack of lifetime data for an EV's battery that could fully represent the actual SOH regression line and the relationship between the variables of voltage, current, temperature and the degradation of the battery.

Unfortunately, scaling up from a single cell to a whole battery pack is not a linear task, as electrical and thermal dependencies among the cells differently connected (i.e. in series or in parallel) and packed in larger units (i.e. sub-modules, modules) heavily impact the pack behaviour, determining more complex responses on the overall SOH. However, we belive that much more information and signals are transmitted over the CAN bus of real world EV batteries. Further research can focus on determining which of these signal types contribute most to battery cells degradation, with the aim of implementing further level of complexity, for example by considering cell voltages, mileage and other temperature measures, to improve NNs response. Consequently, by including more features SOH battery estimation domain and possibly including real EVs data in the training, as a more accurate representation of the real-world could also lead to improvements. From this perspective, a larger amount of data will be required to assess network predictions reliability, as no coherent records were available for the same car model in literature to cover a significant period of battery degradation.

VI. Conclusion

In this research, we propose a data-driven approach to estimate the SOH of lithium-ion batteries used in electric vehicles by replicating the dynamic driving behavior of real EVs using NASA Randomized Battery Usage Dataset. For this purpose, we designed and optimized some state-of-the-art NNs in time-series scenario (i.e. 1D-CNN, GRU, LSTM and CNN-LSTM). At the beginning, to prove their SOH estimation capabilities, we trained and tested the models on the extended NASA dataset, with the main objective of finding the best input-lag dimension for them. All neural networks performed well when trained and tested on single cells stressed under laboratory conditions. However, among all the proposed models, the 1D-CNN achieved the best results in all metrics

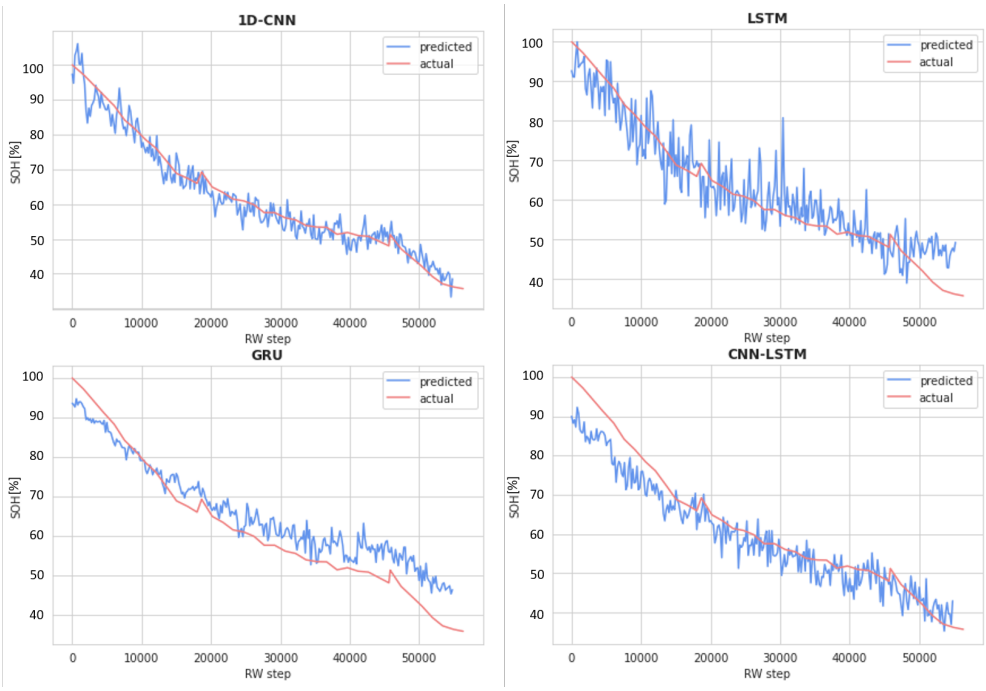

Fig. 4. SOH prediction of each model on the RW9 augmented *test-set*

compared to all the other networks, by fitting best the real SOH profile.

REFERENCES

[1] EU, "A european strategy for low-emission mobility," European Union, Tech. Rep. COM 501, Aug. 2016.

[2] J. Kim, J. Yu, M. Kim, K. Kim, and S. Han, "Estimation of li-ion battery state of health based on multilayer perceptron: as an ev application," *IFAC-PapersOnLine*, vol. 51, no. 28, pp. 392–397, 2018, 10th IFAC Symposium on Control of Power and Energy Systems CPES 2018. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2405896318334542

[3] M. Murnane and A. Ghazel, "A closer look at state of charge (soc) and state of health (soh) estimation techniques for batteries," 2017.

[4] P. Venugopal and V. T., "State-of-health estimation of li-ion batteries in electric vehicle using indrnn under variable load condition," *Energies*, vol. 12, no. 22, 2019. [Online]. Available: https://www.mdpi.com/1996-1073/12/22/4338

[5] L. Yao, S. Xu, A. Tang, F. Zhou, J. Hou, Y. Xiao, and Z. Fu, "A review of lithium-ion battery state of health estimation and prediction methods," *World Electric Vehicle Journal*, vol. 12, no. 3, 2021. [Online]. Available: https://www.mdpi.com/2032-6653/12/3/113

[6] B. Bole, C. Kulkarni, and M. Daigle, "Randomized Battery Usage Data Set," [Online; accessed April-2022]. [Online]. Available: http://ti.arc.nasa.gov/project/prognostic-data-repository

[7] S. Jiang and Z. Song, "A review on the state of health estimation methods of lead-acid batteries," *Journal of Power Sources*, vol. 517, p. 230710, 2022. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0378775321012052

[8] K. S. Ng, C.-S. Moo, Y.-P. Chen, and Y.-C. Hsieh, "Enhanced coulomb counting method for estimating state-of-charge and state-of-health of lithium-ion batteries," *Applied Energy*, vol. 86, no. 9, pp. 1506–1511, September 2009.

[9] H. Ismail Fawaz, G. Forestier, J. Weber, L. Idoumghar, and P.-A. Muller, "Deep learning for time series classification: a review," *Data mining and knowledge discovery*, vol. 33, no. 4, pp. 917–963, 2019.

[10] E. Chemali, P. J. Kollmeyer, M. Preindl, Y. Fahmy, and A. Emadi, "A convolutional neural network approach for estimation of li-ion battery state of health from charge profiles," *Energies*, vol. 15, no. 3, 2022. [Online]. Available: https://www.mdpi.com/1996-1073/15/3/1185

[11] P. Li, Z. Zhang, Q. Xiong, B. Ding, J. Hou, D. Luo, Y. Rong, and S. Li, "State-of-health estimation and remaining useful life prediction for the lithium-ion battery based on a variant long short term memory neural network," *Journal of Power Sources*, vol. 459, p. 228069, 2020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0378775320303724

[12] S. Cui and I. Joe, "A dynamic spatial-temporal attention-based gru model with healthy features for state-of-health estimation of lithium-ion batteries," *IEEE Access*, vol. 9, pp. 27 374–27 388, 2021.

[13] Z. Zhang, Z. Dong, H. Lin, Z. He, M. Wang, Y. He, X. Gao, and M. Gao, "An improved bidirectional gated recurrent unit method for accurate state-of-charge estimation," *IEEE Access*, vol. 9, pp. 11 252–11 263, 2021.

[14] S. Jo, S. Jung, and T. Roh, "Battery state-of-health estimation using machine learning and preprocessing with relative state-of-charge," *Energies*, vol. 14, no. 21, 2021. [Online]. Available: https://www.mdpi.com/1996-1073/14/21/7206

[15] X. Tan, X. Liu, H. Wang, Y. Fan, and G. Feng, "Intelligent online health estimation for lithium-ion batteries based on a parallel attention network combining multivariate time series," *Frontiers in Energy Research*, vol. 10, 2022. [Online]. Available: https://www.frontiersin.org/article/10.3389/fenrg.2022.844985

[16] H. Tian, P. Qin, K. Li, and Z. Zhao, "A review of the state of health for lithium-ion batteries: Research status and suggestions," *Journal of Cleaner Production*, vol. 261, p. 120813, 2020.

[17] A. Mittal, "Understanding rnn and lstm," 2019. [Online]. Available: https://aditi-mittal.medium.com/understanding-rnn-and-lstm-f7cdf6dfc14e

[18] R. Pascanu, T. Mikolov, and Y. Bengio, "On the difficulty of training recurrent neural networks," *30th International Conference on Machine Learning, ICML 2013*, 11 2012.

[19] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural computation*, vol. 9, pp. 1735–80, 12 1997.

[20] J. Chung, C. Gulcehre, K. Cho, and Y. Bengio, "Empirical evaluation of gated recurrent neural networks on sequence modeling," in *NIPS 2014 Workshop on Deep Learning, December 2014*, 2014.

[21] Y. R, N. M, D. RKG, and T. K., "Convolutional neural networks: an overview and application in radiology," 2018.

[22] H. Widiputra, A. Mailangkay, and E. Gautama, "Multivariate cnn-lstm model for multiple parallel financial time-series prediction," *Complexity*, vol. 2021, 10 2021.