

Transformer-based Capacity Prediction for Lithium-ion Batteries with Data Augmentation

Gift Modekwe* Saif Al-Wahaibi* Qiugang Lu*

* *Department of Chemical Engineering, Texas Tech University,
Lubbock, TX 79409, USA*

Abstract: Lithium-ion batteries are pivotal to technological advancements in transportation, electronics, and clean energy storage. The optimal operation and safety of these batteries require proper and reliable estimation of battery capacities to monitor the state of health. Current methods for estimating the capacities fail to adequately account for long-term temporal dependencies of key variables (e.g., voltage, current, and temperature) associated with battery aging and degradation. In this study, we explore the usage of transformer networks to enhance the estimation of battery capacity. We develop a transformer-based battery capacity prediction model that accounts for both long-term and short-term patterns in battery data. Further, to tackle the data scarcity issue, data augmentation is used to increase the data size, which helps to improve the performance of the model. Our proposed method is validated with benchmark datasets. Simulation results show the effectiveness of data augmentation and the transformer network in improving the accuracy and robustness of battery capacity prediction.

Keywords: Transformer, data augmentation, capacity prediction, lithium-ion battery

1. INTRODUCTION

Lithium-ion batteries have become increasingly popular across various sectors due to their numerous benefits, including a longer lifespan, higher specific weight and energy density, faster charging, and a lower self-discharge rate compared to other types of batteries (Horiba, 2014). Their widespread adoption is driven by the growing demand for portable electronics (Morita et al., 2021) and a shift towards more sustainable transportation (Miao et al., 2019). For the safety and optimal performance of these batteries, battery management systems are usually put in place to monitor the state of health (SOH). The SOH of a battery describes the current condition of the battery by comparing its present capacity to its nominal capacity (Yao et al., 2021). This is necessary because repeated charging, discharging, and internal chemical changes gradually diminish battery capacity, reducing the energy that it can deliver (i.e., lifetime) (Barcellona et al., 2022).

Generally, capacity estimation is carried out by two main approaches: model-based approaches and data-driven approaches (Lucaferri et al., 2023). Model-based approaches use physics-based models to describe the gradual loss of capacity observed in lithium-ion batteries. These models aim to provide fundamental insights into how the capacity fade occurs with an attempt to offer physical explanations about the observed changes. Models that have been explored include the electrochemical models, equivalent circuit models, and semi-empirical models (Guo et al., 2022). A significant drawback of model-based approaches is their high computational complexity due to the

large-scale coupled nonlinear partial differential equations (PDEs) involved, especially in the electrochemical models.

Alternatively, data-driven approaches directly utilize historical data for capacity prediction (Zhang et al., 2023). Examples include support vector machines, relevance vector machines, neural networks, and so on (Wu et al., 2016). Among others, neural networks (NNs) are usually preferred because they do not require manual feature engineering or additional processing that involves extensive expertise (Hannan et al., 2021). Various types of NNs have been adopted for capacity prediction, including artificial neural networks (Hussein, 2014), convolutional neural networks (CNNs) (Lee et al., 2023), recurrent neural networks (Catelani et al., 2021), and long short-term memory (LSTM) networks (Zhang et al., 2018). In particular, Park et al. (2020) examined the usage of LSTM by leveraging multi-channel battery profiles as a means of improving the remaining useful life prediction.

Despite the success achieved by these methods, capacity prediction of Lithium-ion batteries with recurrent networks often require extensive computation and may suffer from the inability to capture long-term dependencies in time-series data (Chen et al., 2022). In contrast, the transformer network possesses a significant advantage due to its self-attention and parallel processing ability, which makes it a good option for capturing long-range dependencies with computational ease (Vaswani et al., 2017). Consequently, the transformer architecture has been applied to solve time series problems, including battery management (Chen et al. (2022), Gu et al. (2023)).

Meanwhile, a significant limitation of NN-based approaches is the need for extensive data. Fan et al. (2022)

¹ Corresponding author: Qiugang Lu; Email: jay.lu@ttu.edu

suggested a data augmentation method that involves adding Gaussian noise to the battery dataset to improve the flexibility and robustness of the model. One motivation is to use these random noises to mimic capacity regeneration, fluctuations, and sensor inaccuracies occurring during the usage and monitoring of batteries, thus enhancing the adaptability and resilience of the developed SOH estimation method. In this paper, we will leverage and integrate the benefits of transformers and data augmentation to improve the capacity prediction of Lithium-ion batteries. Benchmark NASA Group 1 and University of Michigan (UofM) battery aging datasets will be used to validate the effectiveness of the proposed method.

2. PRELIMINARIES

2.1 Transformer Network

The transformer network has drawn widespread attention in the natural language processing community (Wolf et al., 2020). It adopts the well-known attention mechanism, which captures the relationship between input elements in a sequence (Vaswani et al., 2017). This mechanism enables the model to prioritize specific information, aiding in enhanced context comprehension and better handling of long-range relationships. The major parts of the transformer used in this study are discussed below.

Multi-head Attention (MHA): MHA uses the attention mechanism to learn from and simultaneously attend to different representational subspaces of the input at various positions (Vaswani et al., 2017). Due to its parallelizable nature, it captures more complex patterns and accelerates the training speed. Given an input sequence $X = \{x_1, x_2, \dots, x_n\}$, the attention mechanism projects the input elements into a subspace by using learned weight matrices W_i^Q, W_i^V, W_i^K , where i refers to the i -th head in the attention:

$$Q_i = XW_i^Q, \quad K_i = XW_i^K, \quad V_i = XW_i^V, \quad (1)$$

Q, K , and V stand for the query, key, and value matrices, respectively. A scaled-dot product is done to obtain the attention scores that stabilize the gradients during training. Softmax normalizes the values obtained by the scaled-dot product, and the attention matrix Z_i is thus obtained as (Vaswani et al., 2017):

$$Z_i = \text{softmax} \left(\frac{Q_i K_i^T}{\sqrt{d_k}} \right) V_i. \quad (2)$$

For m heads, the attention matrices are concatenated to form the multi-head attention matrix, which is then transformed back to the original input dimension by using a learnable weight matrix W^O :

$$Z = \text{concat}(Z_1, Z_2, \dots, Z_m)W^O. \quad (3)$$

Positional Encoding: The scaled-dot product of the transformer architecture does not take into account the order of the input sequence. To address this issue, positional encoding is used to include the position information into the input embedding. The sinusoidal function is commonly adopted to compute the positional encoding, as shown below (Vaswani et al., 2017):

$$PE(pos, 2j) = \sin \left(\frac{pos}{10000^{2j/d_{\text{model}}}} \right), \quad (4)$$

where $PE(pos, 2j)$ and $PE(pos, 2j + 1)$ represent the positional encoding for the even and odd index positions, respectively, pos is the position of the element in the sequence, j is the dimension ranging from 0 to $d_{\text{model}}/2$, and d_{model} is the dimension of the embedding.

Feed-Forward Network: The feed-forward neural network of the transformer is made up of an activated function, the Rectified Linear Unit (ReLU) and dense layers that separate two linear transformations. As a major part of the encoder, it is tasked with introducing non-linearity that can assist the transformer to learn intricate patterns for making better predictions.

Encoder Block: The transformer encoder consists of several identical blocks stacked on top of each other, with each block containing two layers (see Fig. 2). The first layer utilizes the multi-head attention for the positionally encoded input, while the second layer is a fully connected feed-forward neural network to process the extracted features from the attention mechanism (Vaswani et al., 2017). Skip connections are utilized within each layer to mitigate vanishing gradient issues (Liu et al., 2021), and layer normalization is also applied to enhance the model's stability and potentially speed up the training. Such a block is repeated several times to enhance the contextual understanding of the input sequence.

2.2 Data Augmentation

Data augmentation is commonly employed in machine learning (ML) to overcome the issue of data scarcity (Mumuni and Mumuni, 2022). It can be achieved through generating synthetic data and expanding existing dataset to increase the performance and robustness of ML models.

In this study, to address the data scarcity issue, data augmentation is done by adding Gaussian noise to the original dataset to obtain a variant of the input having a similar distribution to the original dataset. For simplicity, Gaussian Noise $G \sim \mathcal{N}(0, \sigma^2)$ is added to every individual sample to create a new dataset:

$$\bar{D} = D + E, \quad (5)$$

where $\bar{D}$ is the new dataset, D is the original dataset and E is the noise obtained from sampling a Gaussian distribution with mean 0 and standard deviation σ . The zero mean ensures that the central tendency of the data remains unchanged after augmentation. However, the introduction of noise with varying σ allows the model to accommodate potentially unseen variations in the data, which is particularly useful for handling real-world fluctuations in the measurements. Through this method, more variability is introduced and the training set is expanded.

3. BATTERY DATASET

To validate the proposed method, we use data from Group 1 of the NASA Ames Prognostic Center of Excellence Aging Dataset (Saha and Goebel, 2007) and 4 cells from the dataset generated by UofM (Mohtat et al., 2021). The NASA Group 1 dataset consists of 4 cells labeled as B0005, B0006, B0007, and B0018. These cells were charged using a constant-current constant-voltage (CC-CV) protocol where the current was kept constant until the

voltage reached 4.2V, and then the voltage was maintained while the current dropped to 20 mA. The experiment ended when the nominal voltage of the cells decreased to 2.7V, 2.5V, 2.2V, and 2.5V for the 4 cells, respectively. The experiment was carried out at an ambient temperature of 24°C. For the UofM dataset, the 4 cells considered were identified as M1, M4, M7, and M10. The cycling procedure includes charging the batteries with a CC-CV protocol similar to that of NASA, although in this case the current dropped to less than C/50 rate (i.e., fully charging the battery needs 50 hours). The experiment ends when the voltage decreased to 3.0V for the four cells and it was performed at room temperature (25°C). Note that each of the 4 cells has a different C-rate for their charge and discharge process. Fig. 1 below shows the capacity degradation over charging-discharging cycles for NASA Group 1 and UofM datasets, respectively.

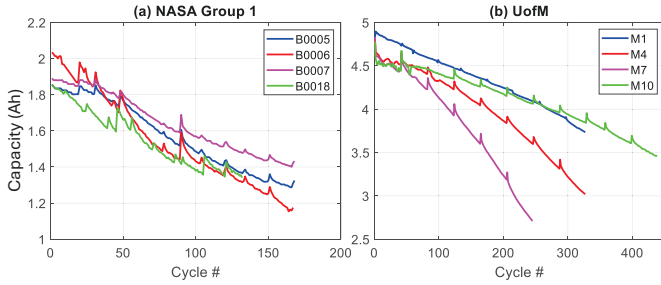


Fig. 1. Capacity degradation profile for the batteries used in this work.

In this study, we focus on the charging cycles as they offer a more stable environment for measuring the internal battery parameters, which are otherwise challenging to gauge during discharging because of the swift and unpredictable current variations. Additionally, the discharge rates are often dictated by the battery owner's usage patterns, which introduces a high degree of variability (Choi et al., 2019).

4. TRANSFORMER-BASED CAPACITY PREDICTION WITH MULTI-CHANNEL PROFILES

4.1 Multi-channel Profiles for Capacity Prediction

This study integrates multi-channel battery profiles, including current, voltage, and temperature, for a charging cycle to capture the predictive relationship between such profiles and battery capacity values. Our motivation for using multi-channel profiles arises from the fact that as battery degrades, the patterns of charging profiles will change, which can serve as an indicator of the degradation status. We adapt the transformer architecture to forecast future capacities by omitting the traditional decoder structure, as our output is a scalar value rather than a sequence.

As in Fig. 3, under the CC-CV protocol, the time it takes for the voltage to reach the upper threshold decreases as the charging cycles progress and the battery degrades. Also, the current will switch to the constant-voltage mode at an earlier time progressively, with the temperature reaching the maximum peak with less time. Note that for the CC-CV protocol, the duration of each charging cycle differs, resulting in different data length for each cycle.

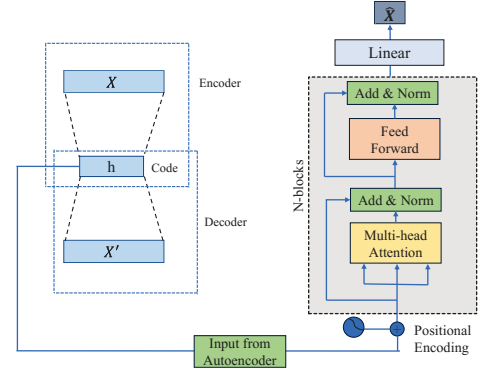


Fig. 2. The proposed auto-encoder and transformer architecture for battery capacity prediction.

To address this issue, we use the subsampling technique to downsample the profiles of each cycle to be the same length prior to being passed to the transformer network.

4.2 Architecture of the Proposed Transformer Network

The structure of the proposed transformer network is shown in Fig. 2. For a generic charging cycle t , we define the down-sampled current, voltage, and temperature profiles as $I_t \in \mathbb{R}^m$, $V_t \in \mathbb{R}^m$, and $T_t \in \mathbb{R}^m$, respectively, with a scalar capacity value $c_t \in \mathbb{R}$. Also, we define the stacked multiple profiles as $X_t = [I_t, V_t, T_t, c_t] \in \mathbb{R}^q$, $q = 3m + 1 \gg 1$, which will be used to predict future capacities c_k , $k > t$. To capture the dynamics in the multivariate time-series, we consider a moving window of past w cycles as the input sequence to the transformer: $\{X_{t-w+1}, \dots, X_t\}$. To tackle the potential high dimensions of X_i at each cycle $i \in \{t-w+1, \dots, t\}$, we first use an auto-encoder to reduce its dimension to a latent state $h_i \in \mathbb{R}^l$, $l \ll q$, with encoder and decoder respectively:

$$h_i = \mathcal{E}_{\theta_1}(X_i), \quad X'_i = \mathcal{D}_{\theta_1}(h_i), \quad i = t-w+1, \dots, t, \quad (6)$$

where X'_i is the reconstructed profile from the decoder $\mathcal{D}_{\theta_1}(\cdot)$, and θ_1 stands for the parameters of the auto-encoder network. We expect the extracted latent state h_i to be representative of the core features of multiple profiles at cycle i . Then, the latent state sequence $\{h_{t-w+1}, \dots, h_t\}$ is fed into the transformer model as the input. The positional encoding as in (4) is adopted to incorporate position information into the input. A number of transformer blocks, each block consisting of a multi-head attention mechanism and feed-forward network, is used to predict the next-step capacity value c_{t+1} . We define the series of transformer blocks as a mapping:

$$\hat{c}_{t+1} = \mathcal{T}_{\theta_2}(h_{t-w+1}, \dots, h_t), \quad (7)$$

where θ_2 is the collection of parameters of all blocks in the transformer architecture in Fig. 2, and $\hat{c}_{t+1}$ is the predicted next-cycle capacity value, which will be compared with the ground-truth capacity c_{t+1} . Ultimately, optimizing the parameters of the entire auto-encoder and transformer framework is formulated as:

$$\min_{\theta} \sum_{t=w}^{n-1} (c_{t+1} - \hat{c}_{t+1})^2 + \lambda \sum_{t=w}^{n-1} \sum_{i=t-w+1}^t (X'_i - X_i)^2, \quad (8)$$

where n is the available number of moving windows created from training data, $\theta = \{\theta_1, \theta_2\}$ stacks all network parameters, and $\lambda \geq 0$ is the weight controlling the

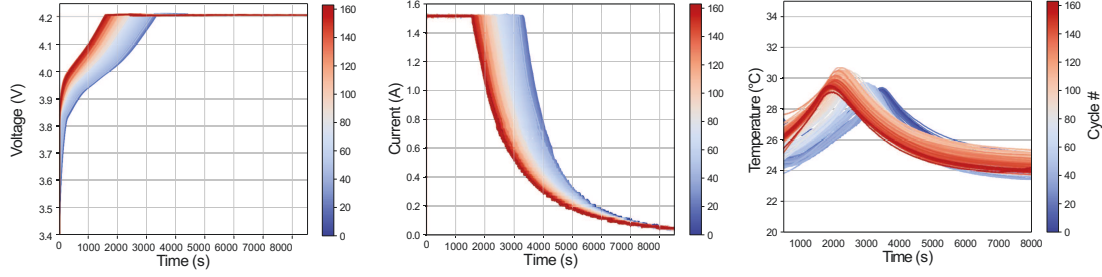


Fig. 3. Voltage, current and temperature profile for battery B0005.

relative importance between two objectives: transformer prediction error and auto-encoder reconstruction error.

5. CASE STUDIES AND DISCUSSIONS

In this section, we employ benchmark NASA Group 1 and UofM battery degradation datasets to validate the proposed methods in previous sections.

5.1 Setup of Case Studies

For these case studies, the hardware configurations include an Intel(R) Xeon(R) W-2265 CPU @ 3.50 GHz processor, 64.0 GB RAM, and an NVIDIA RTX A4000 graphics card. The software environment comprises the PyTorch operating on Python version 3.9. For each battery group, we predict the last 60 cycles of each cell while using the preceding cycles and data from the other 3 cells as the training data. We use the voltage, current, temperature, and the capacity values of each cycle. However, since each cycle has a different length of voltage, current, and temperature profiles, down-sampling is used to divide the profiles of each cycle to a desired dimension for the transformer. Specifically, we down-sample the profiles into $q = 48$ dimensions, consisting of 16 voltage, 16 temperature, 15 current points, and one capacity value for each charging cycle.

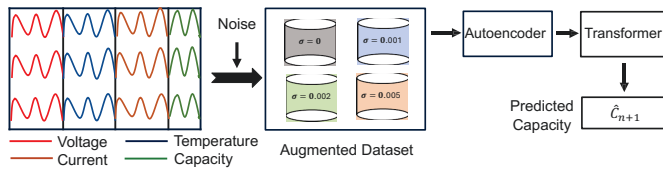


Fig. 4. Logic flow of the case studies.

To tackle the data scarcity, we use data augmentation by creating 3 additional artificial datasets, and each of them is obtained by adding white noise of a different level ($\sigma = 0.001, 0.002, 0.005$) to all the profiles (see Fig. 4). We observed that a large noise level resulted in decreased model performance. This decline is due to the fact that adding large white noise may destroy the inherent dynamics of the original data. For each dataset, a sliding window of $w = 16$ cycles is created on the temperature, voltage, and current profiles. For the autoencoder, the dimension of latent state is chosen as $l = 8$.

The prediction of future capacities is conducted in a rolling manner, where the predicted $\hat{c}_{n+1}$ (n is the current cycle index) will be incorporated to construct X_{n+1} , and I_{n+1} ,

V_{n+1} , T_{n+1} are assumed to be known, as in practice these profiles are easily acquired from the battery management system, in contrast to the capacity value.

5.2 Experimental Result and Analysis:

To assess the efficiency of the model, the mean absolute error (MAE) and root mean square error (RMSE) metrics below are used as performance indicators:

$$\text{MAE} = \frac{1}{p} \sum_{i=1}^p |(y_i - \hat{y}_i)|, \text{RMSE} = \sqrt{\frac{1}{p} \sum_{i=1}^p (y_i - \hat{y}_i)^2} \quad (9)$$

where p is the number of capacities to be predicted, y_i is the true capacity while $\hat{y}_i$ is the predicted capacity.

Figs. 5 and 7 show the predicted capacity values versus the ground truth values, while Figs. 6 and 8 show the absolute error plot for the NASA Group 1 and UofM datasets, respectively. Four methods are compared: the transformer with data augmentation (our proposed method), transformer without data augmentation, LSTM, and CNN. Notably, the tests for CNN and LSTM were performed using the augmented data. Tables 1-2 illustrate the specific metrics for the prediction errors from these methods. From these results, it is observed that the proposed transformer with data augmentation method consistently outperforms the other approaches. Moreover, overall, the transformer methods excel CNN and LSTM largely, which validates the benefits of using transformer to extract more useful (e.g., long-term) patterns for better capacity prediction.

Table 1. Capacity prediction for NASA data.

Battery	Criteria	CNN	LSTM	Transf.	Transf. w. aug.
B0005	MAE	0.0223	0.0156	0.0087	0.0074
	RMSE	0.2093	0.1466	0.0098	0.0078
B0006	MAE	0.0258	0.0223	0.0228	0.0223
	RMSE	0.2817	0.2534	0.0029	0.0026
B0007	MAE	0.0193	0.0111	0.0139	0.0131
	RMSE	0.1810	0.1105	0.0173	0.0157
B0018	MAE	0.0308	0.0220	0.0238	0.0235
	RMSE	0.2488	0.2185	0.0320	0.0275

6. CONCLUSION

In this work we proposed a framework integrating auto-encoder and transformer network for battery capacity prediction. To address the data scarcity issue, we use data augmentation to enrich the dataset by adding random noise to mimic the signal fluctuations and sensor inaccuracies occurring in real-life situations. We further use the well-known benchmark NASA Group 1 and UofM datasets

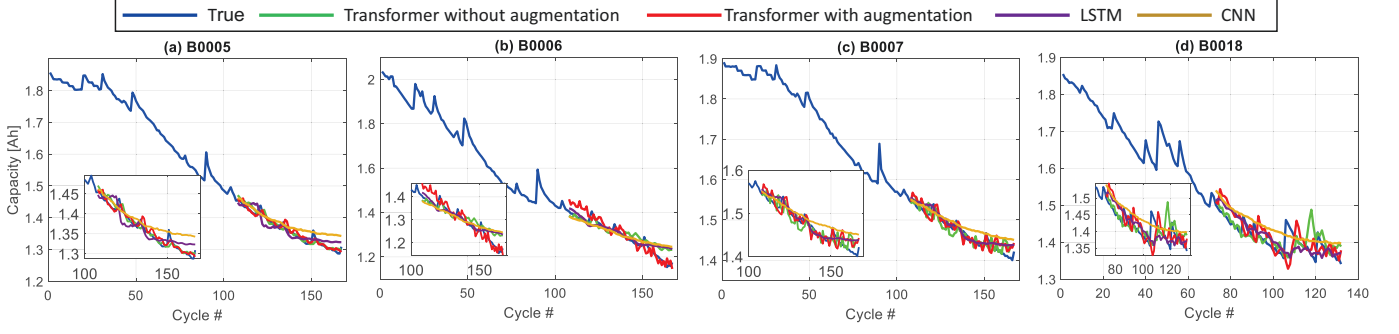


Fig. 5. Predicted capacity vs. ground-truth values for NASA Group 1 dataset.

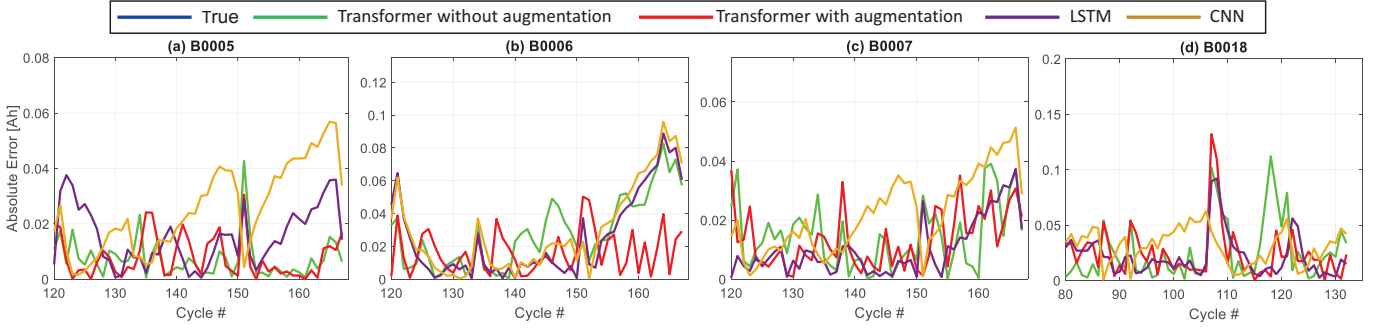


Fig. 6. Absolute prediction error of capacities for NASA Group 1 dataset.

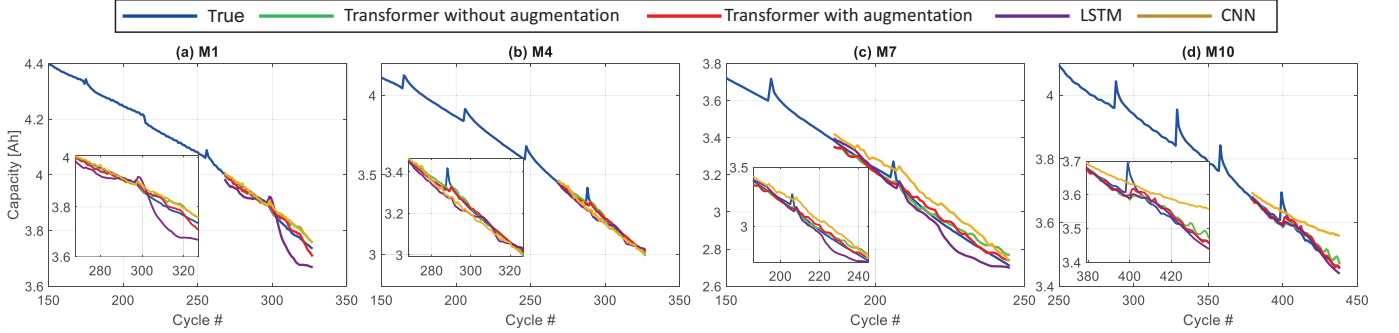


Fig. 7. Predicted capacity vs. ground-truth values for UofM dataset.

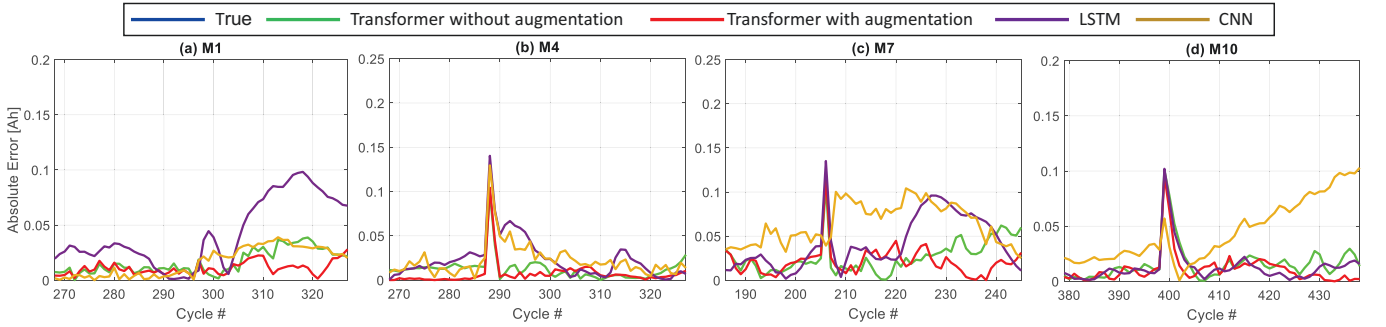


Fig. 8. Absolute prediction error of capacities for UofM dataset.

for validating the effectiveness of the proposed methods. Simulation results show that transformer networks consistently outperform the other deep learning methods such as LSTM and CNN in giving accurate battery capacity predictions. Moreover, with data augmentation, the performance of transformer networks can be further improved when trained with a larger data size. In this work, the variance of the random noise used to enrich the dataset

is kept small, so that the noise will not distort the inherent dynamic behaviors of battery signals. However, since random noise cannot represent the dynamics of the original signal, one future direction is to exploit other data augmentation techniques to increase the data size and richness to better address the data scarcity issue. Additionally, we hope to explore datasets collected under

Table 2. Capacity prediction for UofM data.

Battery	Criteria	CNN	LSTM	Transf.	Transf. w. aug.
M1	MAE	0.0163	0.0424	0.0165	0.0108
	RMSE	0.1202	0.1219	0.0197	0.0123
M4	MAE	0.0223	0.0236	0.0125	0.0072
	RMSE	0.2568	0.2967	0.0193	0.0159
M7	MAE	0.0642	0.0417	0.0262	0.0212
	RMSE	0.4429	0.4565	0.0322	0.0268
M10	MAE	0.0448	0.0129	0.0158	0.0110
	RMSE	0.2857	0.1624	0.0229	0.0183

different temperature conditions to further enhance the robustness and applicability of our model.

7. ACKNOWLEDGMENT

We acknowledge the support from NSF Grant 2340194.

REFERENCES

- Barcellona, S., Colnago, S., Dotelli, G., Latorrata, S., and Piegari, L. (2022). Aging effect on the variation of li-ion battery resistance as function of temperature and state of charge. *Journal of Energy Storage*, 50, 104658.
- Catelani, M., Ciani, L., Fantacci, R., Patrizi, G., and Picano, B. (2021). Remaining useful life estimation for prognostics of lithium-ion batteries based on recurrent neural network. *IEEE Transactions on Instrumentation and Measurement*, 70, 1–11.
- Chen, D., Hong, W., and Zhou, X. (2022). Transformer network for remaining useful life prediction of lithium-ion batteries. *IEEE Access*, 10, 19621–19628.
- Choi, Y., Ryu, S., Park, K., and Kim, H. (2019). Machine learning-based lithium-ion battery capacity estimation exploiting multi-channel charging profiles. *IEEE Access*, 7, 75143–75152.
- Fan, Y., Wu, H., Chen, W., Jiang, Z., Huang, X., and Chen, S.Z. (2022). A data augmentation method to optimize neural networks for predicting soh of lithium batteries. In *Journal of Physics: Conference Series*, volume 2203, 012034. IOP Publishing.
- Gu, X., See, K., Li, P., Shan, K., Wang, Y., Zhao, L., Lim, K.C., and Zhang, N. (2023). A novel state-of-health estimation for the lithium-ion battery using a convolutional neural network and transformer model. *Energy*, 262, 125501.
- Guo, W., Sun, Z., Vilsen, S.B., Meng, J., and Stroe, D.I. (2022). Review of “grey box” lifetime modeling for lithium-ion battery: Combining physics and data-driven methods. *Journal of Energy Storage*, 56, 105992.
- Hannan, M.A., How, D.N., Lipu, M.H., Mansor, M., Ker, P.J., Dong, Z., Sahari, K., Tiong, S.K., Muttaqi, K.M., Mahlia, T.I., et al. (2021). Deep learning approach towards accurate state of charge estimation for lithium-ion batteries using self-supervised transformer model. *Scientific Reports*, 11(1), 19541.
- Horiba, T. (2014). Lithium-ion battery systems. *Proceedings of the IEEE*, 102(6), 939–950.
- Hussein, A.A. (2014). Capacity fade estimation in electric vehicle li-ion batteries using artificial neural networks. *IEEE Transactions on Industry Applications*, 51(3), 2321–2330.
- Lee, G., Kwon, D., and Lee, C. (2023). A convolutional neural network model for SOH estimation of Li-ion batteries with physical interpretability. *Mechanical Systems and Signal Processing*, 188, 110004.
- Liu, F., Ren, X., Zhang, Z., Sun, X., and Zou, Y. (2021). Rethinking skip connection with layer normalization in transformers and resnets. *arXiv preprint arXiv:2105.07205*.
- Lucaferri, V., Quercio, M., Laudani, A., and Riganti Fulginei, F. (2023). A review on battery model-based and data-driven methods for battery management systems. *Energies*, 16(23), 7807.
- Miao, Y., Hynan, P., Von Jouanne, A., and Yokochi, A. (2019). Current li-ion battery technologies in electric vehicles and opportunities for advancements. *Energies*, 12(6), 1074.
- Mohtat, P., Lee, S., Siegel, J.B., and Stefanopoulou, A.G. (2021). Reversible and irreversible expansion of lithium-ion batteries under a wide range of stress factors. *Journal of The Electrochemical Society*, 168(10), 100520.
- Morita, Y., Saito, Y., Yoshioka, T., and Shiratori, T. (2021). Estimation of recoverable resources used in lithium-ion batteries from portable electronic devices in japan. *Resources, Conservation and Recycling*, 175, 105884.
- Mumuni, A. and Mumuni, F. (2022). Data augmentation: A comprehensive survey of modern approaches. *Array*, 16, 100258.
- Park, K., Choi, Y., Choi, W.J., Ryu, H.Y., and Kim, H. (2020). LSTM-based battery remaining useful life prediction with multi-channel charging profiles. *IEEE Access*, 8, 20786–20798.
- Saha, B. and Goebel, K. (2007). Battery data set. *NASA AMES Prognostics Data Repository*.
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, L., and Polosukhin, I. (2017). Attention is all you need. *Advances in Neural Information Processing Systems*, 30.
- Wolf, T., Debut, L., Sanh, V., Chaumond, J., Delangue, C., Moi, A., Cistac, P., Rault, T., Louf, R., Funtowicz, M., et al. (2020). Transformers: State-of-the-art natural language processing. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing: System Demonstrations*, 38–45.
- Wu, L., Fu, X., and Guan, Y. (2016). Review of the remaining useful life prognostics of vehicle lithium-ion batteries using data-driven methodologies. *Applied Sciences*, 6(6), 166.
- Yao, L., Xu, S., Tang, A., Zhou, F., Hou, J., Xiao, Y., and Fu, Z. (2021). A review of lithium-ion battery state of health estimation and prediction methods. *World Electric Vehicle Journal*, 12(3), 113.
- Zhang, M., Yang, D., Du, J., Sun, H., Li, L., Wang, L., and Wang, K. (2023). A review of SOH prediction of Li-ion batteries based on data-driven algorithms. *Energies*, 16(7), 3167.
- Zhang, Y., Xiong, R., He, H., and Pecht, M.G. (2018). Long short-term memory recurrent neural network for remaining useful life prediction of lithium-ion batteries. *IEEE Transactions on Vehicular Technology*, 67(7), 5695–5705.