Predicting Battery Cycle Life with Few-Shot Transfer Learning over Heterogeneous Datasets

Runyao Yu [1, 2, 3], Jiaqi Wang [1], Yongsheng Han [1], Chi Zhang [1], Teddy Szemberg O'Connor [3], Jochen L. Cremer [1, 2]
[1] Delft University of Technology, Delft, Netherlands
[2] Austrian Institute of Technology, Vienna, Austria
[3] Rimac Technology, Zagreb, Croatia
Email: runyao.yu@tudelft.nl

2024 IEEE PES Innovative Smart Grid Technologies Europe (ISGT EUROPE) | 979-8-3503-9042-1/24/$31.00 ©2024 IEEE | DOI: 10.1109/ISGTEUROPE62998.2024.10863102

***Abstract*—This paper presents an efficient approach to battery cycle life prediction through few-shot transfer learning, addressing the challenges of costly and limited battery aging data. Leveraging freely available datasets, a multi-layer perceptron (MLP) model was pretrained on diverse battery aging datasets to adapt to new prediction tasks with minimal training samples through few-shot fine-tuning techniques on the target data. The proposed fine-tuning strategy was validated using a heterogeneous aging dataset of 347 batteries, with cycle lives ranging from 144 to 4,052 cycles, incorporating batteries with lithium iron phosphate (LFP), lithium cobalt oxide (LCO), nickel cobalt aluminum oxide (NCA), and nickel manganese cobalt oxide (NMC) chemistries, which ensures robust validation of our methods. The results show that even with few samples of data from a target task, a comparable generalization performance to training from scratch with 100% data can be achieved, thus demonstrating its effectiveness in utilizing available resources for accurate cycle life prediction.**

***Index Terms*—Battery aging prediction, cycle life prediction, transfer learning, few-shot learning, fine-tuning, deep learning**

I. Introduction

Accurate prediction of battery cycle life is necessary for optimizing the performance of batteries across various applications such as energy storage systems and electric vehicles. These applications rely on precise cycle life estimation to ensure reliability, cost-effectiveness, and sustainability.

In recent years, data-driven aging models have garnered increasing attention in the field. Data-driven models have been extensively applied to enhance the performance of batteries [1], [2]. A particularly notable study used data from only the first 100 cycles of an aging campaign to predict battery cycle life with an impressively low error of 9.1% [3]. Recent study highlights possibilities and methods in predicting battery life from field data [4]. However, acquiring field data can be costly and time-consuming. Consequently, the freely available open-source aging data has become a useful source.

Transfer learning can be used to pretrain a model on one or several datasets and then transfer it to a target dataset. A study illustrated how a pretrained deep learning model, initially developed with 20 NCA cells, was successfully transferred to target NCM cells using only two samples of NCM data [5].

This paper proposes few-shot transfer learning on heterogeneous aging datasets to overcome the limitation of dataset sizes boosting predicting performance. Precisely, an MLP model is pretrained using partial battery aging datasets from multiple sources [3], [6], [7], [8], [9], [10], [11], [12], intentionally forming a mixed aging dataset with different chemistries including LFP, LCO, NCA, and NMC. A fine-tuning strategy is proposed to adapt the MLP using a few samples from the target dataset, and the MLP is optimized using Optuna [13]. The case studies assess the generalization performance as the fine-tuning samples increase. An overview of the experiment design can be found in Fig. 1. Compared to training a model with 100% data for a specific task, the proposed transfer learning strategy achieves comparable generalization performance and requires between only 8.6% to 48.8% of the data for different tasks, showing its effectiveness in utilizing the potential of open-sourced aging datasets for accurate cycle life prediction. In addition, we have prepared cleaned and merged feature sets to facilitate researchers in applying transfer learning to their datasets.

II. Dataset

To address the fragmented landscape of battery data formats and sources, the study BatteryML [14] introduced a unified data format. In this work, we form and utilize three major datasets following the BatteryML format. The distribution of cycle lives, representing the remaining useful life (RUL) of batteries, for three datasets, is shown in Fig. 2.

1) MATR: This contains 124 LFP cells [3]. The cells were cycled in a temperature-controlled environmental chamber (30 °C) under varied fast-charging conditions but identical discharging conditions until 80% capacity fade. The dataset was originally split into 33.1% (41 cells) for training, 34.7% (43 cells) for primary testing, and 32.3% (40 cells) for secondary testing. We chose the primary testing set to evaluate the performance of generalization, as the secondary testing set was created after the model was developed [3].

2) HUST: This contains 77 LFP cells [6], which underwent the same charging protocol but were subjected to various multi-stage discharge regimes, all conducted at a temperature of 30 °C. The dataset was split into 71.4% (55 cells) for training and 28.6% (22 cells) for testing.

3) MIX: This dataset combines several smaller battery collections from the CALCE, HNEI, RWTH, SNL, and UL

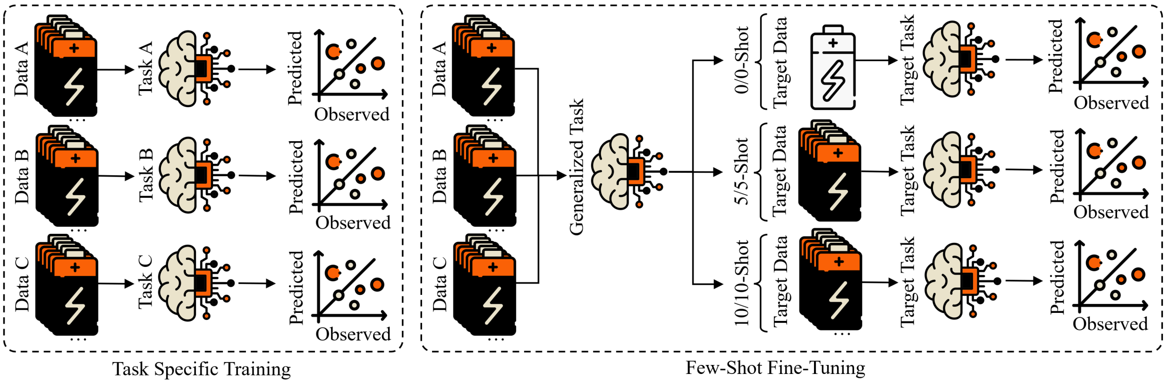

Fig. 1. Overview of the experiment design. Task specific training (left) involves training one model from scratch using 100% of the training data. Few-shot fine-tuning (right) entails pretraining one model with external datasets, and then fine-tuning with only a limited number of samples from the target data. After finishing training, different approaches will be evaluated and compared with the same testing data.

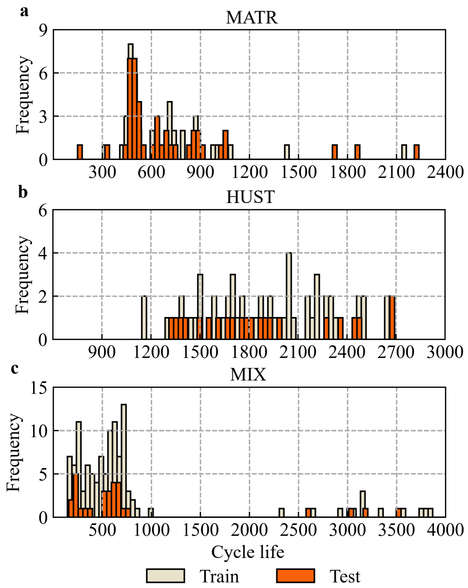

Fig. 2. Distribution of cycle lives for three datasets. The training set is depicted in orange and the testing set in tan. **a**, Distribution for MATR, cycle life varies from 148 to 2,238 cycles. **b**, Distribution for HUST, cycle life varies from 1144 to 2,691 cycles. **c**, Distribution for MIX, cycle life varies from 144 to 4,052 cycles.

PUR datasets [7], [8], [9], [10], [11], [12], encompassing 146 cells. The dataset features a variety of electrode chemistries, including LFP, LCO, NCA, and NMC. The data is divided into training and testing subsets, with 79.5% (116 cells) allocated for training and 20.5% (30 cells) designated for testing.

III. Methodology

A. Feature Engineering

The feature engineering is based on the quantified discharge linearity methodology, proposed in [3]. This involves a processing pipeline that interpolates the discharge voltage-capacity curve.

1) Interpolation: Each discharge voltage-capacity curve is subjected to a conditional interpolation. V_{C} and Q_{C} are defined as the sets of voltage and capacity values where the discharge current I is less than or equal to a small negative threshold $-\epsilon$, where $\epsilon \geq 0$ and is a very small value:

$$V_{\mathrm{C}} = V[I \leq -\epsilon], \quad Q_{\mathrm{C}} = Q[I \leq -\epsilon]. \tag{1}$$

The interpolation of discharge capacity Q_I is performed across a specified voltage interval $[V_{\min}, V_{\max}]$. This interval is divided into K even segments. The interpolation model is trained on data points $(V_{\mathrm{C}}, Q_{\mathrm{C}})$. It then predicts the discharge capacity at these K segmented voltage points to form K interpolated points of discharge capacity:

$$Q_I = \mathrm{Interpolate}(V_{\mathrm{C}}, Q_{\mathrm{C}}, [V_{\min}, V_{\max}], K) \tag{2}$$

2) Feature Compilation: The final features for each cell are constructed by calculating the difference between the values at cycle 100 and cycle 10 of the interpolated capacities. These differential features help emphasize the changes in battery behavior over cycles and are utilized as input in machine learning models. The features are expressed as follows:

$$\Delta(Q_I)_{100-10} = (Q_I)_{100} - (Q_I)_{10} \tag{3}$$

3) Label Extraction: In predictive modeling of battery health, RUL is a critical label that signifies the number of remaining cycles before a battery reaches its end-of-life (EoL). The EoL is defined as the point in its lifecycle when the battery retains only a certain percentage of its nominal capacity Q_{N}, here set at 80%.

B. Modeling

1) Base Model: The base model for the transfer learning approach is MLP. The MLP is designed with multiple layers, each consisting of neurons that apply a nonlinear activation function. Mathematically, MLP can be described as follows:

$$\begin{aligned} f(\mathbf{x};\Theta) = \sigma\Big(\mathbf{W}^{(L)}\sigma\Big(\mathbf{W}^{(L-1)}\cdots \\ \sigma(\mathbf{W}^{(1)}\mathbf{x}+\mathbf{b}^{(1)})\cdots\Big)+\mathbf{b}^{(L-1)}\Big)+\mathbf{b}^{(L)} \end{aligned} \tag{4}$$

TABLE I
TESTING RMSE AND MAE FOR PREDICTED CYCLE LIVES

Task Specific Training	Inference	
Data	RMSE	MAE
MATR → MATR	113.50	87.28
HUST → HUST	443.98	373.54
MIX → MIX	327.59	209.52

Few-Shot Fine-Tuning	0/0-Shot		5/5-Shot		10/10-Shot	
Data	RMSE	MAE	RMSE	MAE	RMSE	MAE
HUST + MIX → MATR	242.31	202.20	121.88	95.95	110.39	90.97
MIX + MATR → HUST	909.96	696.48	496.34	402.01	444.80	354.00
MATR + HUST → MIX	874.13	607.36	344.38	248.71	253.58	177.71

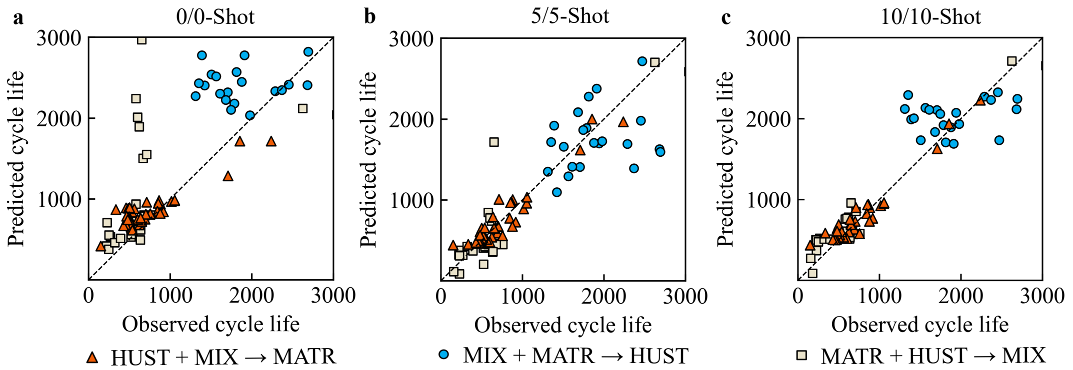

Fig. 3. Observed and predicted cycle lives using few-shot transfer learning. The fine-tuning strategy was evaluated on testing data from MATR (orange triangle), HUST (blue circle), and MIX (tan square). **a**, Usage of zero samples for both fine-tuning and hyperparameter optimization. **b**, Usage of five samples for each. **c**, Usage of ten samples for each.

where $\mathbf{x}$ is the input vector, $\mathbf{W}^{(l)}$ and $\mathbf{b}^{(l)}$ are the weights and biases at layer l, σ represents the activation function, and Θ denotes the collection of all weights and biases. The objective is to find:

$$\Theta^* = \arg\min_{\Theta} \left(\frac{1}{n} \sum_{i=1}^{n} (y_i - f(x_i; \Theta))^2 \right) \tag{5}$$

where y_i is the actual cycle life and n is the number of samples in the training set.

2) Transfer Learning: Transfer learning in the context of this study involves adapting a pretrained MLP to a target cycle life prediction task with minimal target battery data. The fine-tuning of the model weights can be expressed as:

$$\Theta' = \Theta^* + \Delta\Theta \tag{6}$$

where Θ' represents the updated parameter set, Θ^* the initial parameters from the pre-trained MLP, and $\Delta\Theta$ the adjustments computed through several iterations of gradient descent on the new dataset:

$$\Delta\Theta = -\eta \nabla_{\Theta} \mathcal{L}(f(\mathbf{x}; \Theta), \mathbf{y}) \tag{7}$$

where η is the learning rate, and $\mathcal{L}$ is the loss function. $\mathbf{x}$ and $\mathbf{y}$ represent input features and target outputs, respectively.

3) Fine-Tuning Strategy: Initially, the MLP is pretrained using all available data from two of the three datasets. Subsequently, the model is fine-tuned on M samples from the training set of the third dataset to adapt the pretrained model. To optimize the MLP's hyperparameters and prevent overfitting, other N samples from the training set of the third dataset are used for validation through Optuna [13]. We introduce this approach as the M/N-shot fine-tuning strategy.

The core idea behind employing this fine-tuning strategy is to emulate a realistic scenario in which the target dataset—identified as the third dataset in this study—represents project-specific in-house aging data, typically limited in availability. By pretraining on the more abundant and diverse data from the other two datasets, which we consider as freely available open-source data, we prepare the MLP to adapt to the scarce target data effectively.

C. Evaluation

To assess the performance of our models, we employ two standard metrics: mean absolute error (MAE) and root mean squared error (RMSE).

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^{n} |y_i - \hat{y}_i| \tag{8}$$

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2} \tag{9}$$

where y_i represents the observed cycle lives, $\hat{y}_i$ represents the predicted cycle lives, and n is the number of observations.

IV. EXPERIMENT

The objective of this experiment is to demonstrate that the few-shot transfer learning can achieve comparable predictive accuracy to a training method that utilizes the entire dataset.

A. HUST + MIX $\rightarrow$ MATR

The MLP is pretrained using data from 77 cells from HUST and 146 cells from MIX. Fine-tuning is performed with M samples and optimized with N samples from the original training dataset (41 cells) of MATR. Testing is conducted on the original testing dataset (43 cells) from MATR.

Our experiments demonstrate a clear trend: as we progress from 0/0-shot to 5/5-shot, and further to 10/10-shot, there is a consistent decrease in the loss metrics (RMSE and MAE), as can be seen in Table I, corroborating the efficacy of the M/N fine-tuning strategy. A comparison with task specific training reveals that the 10/10-shot fine-tuning, which takes only 48.8% of the data, yields similar performance to training models from scratch using 100% of the data. This indicates that the few-shot transfer learning approach effectively matches the predictive accuracy of the task specific training.

B. MATR + MIX $\rightarrow$ HUST

The MLP is pretrained using data from 124 cells from MATR and 146 cells from MIX. Fine-tuning is performed with M samples and optimized with N samples from the original training dataset (77 cells) of HUST. Testing is conducted on the original testing dataset (22 cells) from HUST.

Despite achieving similar predictive performance by applying a 10/10-shot fine-tuning strategy compared to task specific training using 100% training data, the HUST dataset presents a notable challenge, as indicated in Table I, which is reflected in higher loss metrics.

C. HUST + MATR $\rightarrow$ MIX

The MLP is pretrained using data from 77 cells from HUST and 124 cells from MATR. Fine-tuning is performed with M samples and optimized with N samples from the original training dataset (116 cells) of MIX. Testing is conducted on the original testing dataset (30 cells) from MIX.

The MIX dataset achieves comparable performance with 5/5-shot fine-tuning, utilizing only 8.6% of the training dataset, as opposed to training from scratch. However, as shown in Fig. 3, the model tends to overestimate the short cycle lives when the number of samples used for fine-tuning is small. Increasing the fine-tuning samples from MIX itself results in rapid calibration of the model, effectively mitigating overestimation for short cycle lives while also improving underestimation for longer cycle lives.

V. Conclusion

This study demonstrates the potential of leveraging large amounts of open-source aging data to transfer learning effectively to target tasks. The results indicate that the proposed M/N-shot fine-tuning strategy proves to be effective, enabling the achievement of performance comparable to models trained from scratch. This is particularly significant for applications where data acquisition is expensive and time-consuming. Furthermore, our cleaned and merged feature sets can be utilized to transfer learning to researchers' datasets.

Looking ahead, there is a clear avenue for advancing the accuracy of cycle life predictions through feature engineering. This could involve the integration of deeper domain-specific knowledge into the learning process to capture the complex behaviors of battery degradation more precisely. Additionally, exploring model-agnostic meta-learning could be worthwhile, as it may enable models to adapt rapidly with even fewer data points. Embracing such strategies could pave the way for more nuanced and robust predictive models in the field of battery health management.

References

[1] M. Zhang, D. Yang, J. Du, H. Sun, L. Li, L. Wang, and K. Wang, "A review of soh prediction of li-ion batteries based on data-driven algorithms," *Energies*, vol. 16, no. 7, 2023. [Online]. Available: https://www.mdpi.com/1996-1073/16/7/3167

[2] A. M. Billert, R. Yu, S. Erschen, M. Frey, and F. Gauterin, "Improved quantile convolutional and recurrent neural networks for electric vehicle battery temperature prediction," *Big Data Mining and Analytics*, vol. 7, no. 2, pp. 512–530, 2024.

[3] K. A. Severson, P. M. Attia, N. Jin, N. Perkins, B. Jiang, Z. Yang, M. H. Chen, M. Aykol, P. K. Herring, D. Fraggedakis *et al.*, "Data-driven prediction of battery cycle life before capacity degradation," *Nature Energy*, vol. 4, no. 5, pp. 383–391, 2019.

[4] V. Sulzer, P. Mohtat, A. Aitio, S. Lee, Y. T. Yeh, F. Steinbacher, M. U. Khan, J. W. Lee, J. B. Siegel, A. G. Stefanopoulou, and D. A. Howey, "The challenge and opportunity of battery lifetime prediction from field data," *Joule*, vol. 5, no. 8, pp. 1934–1955, 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2542435121002932

[5] Y. Wang, J. Zhu, L. Cao, B. Gopaluni, and Y. Cao, "Long short-term memory network with transfer learning for lithium-ion battery capacity fade and cycle life prediction," *Applied Energy*, vol. 350, p. 121660, 2023. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0306261923010243

[6] G. Ma, S. Xu, B. Jiang, C. Cheng, X. Yang, Y. Shen, T. Yang, Y. Huang, H. Ding, and Y. Yuan, "Real-time personalized health status prediction of lithium-ion batteries using deep transfer learning," *Energy & Environmental Science*, vol. 15, no. 10, pp. 4083–4094, Oct. 2022.

[7] Y. Xing, E. W. Ma, K.-L. Tsui, and M. Pecht, "An ensemble model for predicting the remaining useful performance of lithium-ion batteries," *Microelectronics reliability*, vol. 53, no. 6, pp. 811–820, 2013.

[8] A. Devie, G. Baure, and M. Dubarry, "Intrinsic variability in the degradation of a batch of commercial 18650 lithium-ion cells," *Energies*, vol. 11, no. 5, p. 1031, 2018.

[9] W. Li, N. Sengupta, P. A. Dechent, D. Howey, A. Annaswamy, and D. U. Sauer, "One-shot battery degradation trajectory prediction with deep learning," *Journal of power sources*, 2021.

[10] Y. Preger, H. M. Barkholtz, A. Fresquez, D. L. Campbell, B. W. Juba, J. Romàn-Kustas, S. R. Ferreira, and B. Chalamala, "Degradation of commercial lithium-ion cells as a function of chemistry and cycling conditions," *Journal of The Electrochemical Society*, vol. 167, no. 12, p. 120532, 2020.

[11] D. Juarez-Robles, J. A. Jeevarajan, and P. P. Mukherjee, "Degradation-safety analytics in lithium-ion cells: Part i. aging under charge/discharge cycling," *Journal of The Electrochemical Society*, vol. 167, no. 16, p. 160510, 2020.

[12] D. Juarez-Robles, S. Azam, J. A. Jeevarajan, and P. P. Mukherjee, "Degradation-safety analytics in lithium-ion cells and modules: Part iii. aging and safety of pouch format cells," *Journal of The Electrochemical Society*, vol. 168, no. 11, p. 110501, 2021.

[13] T. Akiba, S. Sano, T. Yanase, T. Ohta, and M. Koyama, "Optuna: A next-generation hyperparameter optimization framework," in *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, ser. KDD '19. New York, NY, USA: Association for Computing Machinery, 2019, p. 2623–2631. [Online]. Available: https://doi.org/10.1145/3292500.3330701

[14] H. Zhang, X. Gui, S. Zheng, Z. Lu, Y. Li, and J. Bian, "BatteryML:An Open-source platform for Machine Learning on Battery Degradation," Dec. 2023.