POLITECNICO MILANO 1863

SCUOLA DI INGEGNERIA INDUSTRIALE E DELL'INFORMAZIONE

Executive Summary of the Thesis

Domain Adaptation for Remaining Useful Life Estimation of Lithium-Ion Batteries

Laurea Magistrale in Computer Science and Engineering - Ingegneria Informatica

Authors: Giuseppe Meli, Luca Pasquarelli

Advisor: Prof. Francesco Amigoni

Co-advisors: Prof. Loredana Cristaldi, Dott. Davide Azzalini, Dott. Luca Martiri

Academic year: 2022-2023

1. Introduction

Nowadays there are far more electronic devices than human beings on Earth. These devices are usually powered using lithium-ion batteries. Lithium-ion batteries are widely employed due to their numerous advantages, such as high output power, extended cycle life, high energy density, and lower environmental impact. These batteries find applications in diverse industries, including electronics, aerospace, and military. However, each of these batteries can assume completely different behaviors from their peers based on usage, charging, and many other factors, leading to potential harm, unreliableness, and other major potential issues depending on the importance and purpose of the given device being powered. Ensuring batteries operate safely and effectively relies on the implementation of a Battery Management System (BMS). This system is pivotal in precisely assessing battery conditions, maximizing their efficiency, and adeptly identifying potential issues, thereby underpinning the core functions of battery management. Building on the essential role of a BMS, its ability to extend battery life and ensure safety is further enhanced by accurate predictions of the Remaining Useful Life (RUL). Nowadays, these predictions are given by ever-advancing data-driven models [5] that are surpassing the previously used physical-based models. It is not just about being able to predict the RUL for a specific set of batteries; equally important is developing a model that maintains reliable predictive performance even when applied to different sets of batteries than those it was initially trained on. This would broaden the scope of the BMS, enabling quicker deployment across various applications and contributing to cost efficiencies.

2. Aim of the Thesis

Our research aims to contribute to expedited deployment and cost reduction in the production and maintainment phase of electronic devices by building one model for RUL prediction trained on one specific dataset of batteries, that is capable of delivering good performances on a different set of batteries while accounting for an unavoidable loss due to unpredictability and diversity in battery behavior. It is also within our scope to portray a comparison of different methodologies that could be adopted to tackle domain adaptation in this specific context.

We employ a sophisticated neural network architecture, specifically utilizing a Convolutional Long Short-Term Memory Neural Network (ConvLSTM) with an attention layer, supported by a Domain Adversarial Neural Network (DANN) [1] in the training phase. This thesis introduces the DANN to align feature representations across different domains, effectively harmonizing feature distributions. To evaluate the effectiveness and versatility of our proposed methodology, we leverage the MIT-Toyota 2019 collaboration dataset, composed of a diverse set of batches and recognized as the most extensive lithium-ion batteries dataset publicly available at the time of our study [5]. In our work, we were able to assess strong performances using both transfer learning and DANN, allowing us to present an interesting comparison of the results from each method. Our research holds significant practical value, offering substantial advancements in RUL estimation methodologies for lithium-ion batteries, crucial for their safer and more efficient use. Notably, our emphasis on addressing challenges associated with domain adaptation introduces a pioneering approach in the field, as domain adversarial learning has not been explored yet in the context of RUL estimation for lithium-ion batteries [3]. The use of DANN offers a distinctive advantage, as its unsupervised learning approach eliminates the dependency on explicit labels from source or target domains. This is particularly valuable in scenarios where labeled data is limited or costly, facilitating practical implementation in real-world applications.

3. Problem Description

The State of Health (SoH) of a battery quantifies its current health and performance in comparison to its pristine state when new. It is expressed as $\text{SoH}_k = \frac{C_k}{C} \times 100$ for any cycle k, where C_k is the battery's capacity at cycle k and C is the original capacity. The SoH value ranges from 0 to 100, reflecting the percentage of the battery's residual health.

The RUL at a given cycle k is defined as the number of cycles remaining until the battery reaches its End of Life (EoL). This is calculated as $\text{RUL}_k = \text{cycle}_{\text{EOL}} - k$. For lithium-ion batteries, the EoL is typically identified when the SoH falls to 80% or lower.

Our estimation involves predicting the RUL based on data [5] such as current capacity and cell temperature, which are omnipresent features in battery datasets, and BMSs.

In **Remaining useful life** estimation acquiring sufficient battery aging data is a significant hurdle. This difficulty arises from the complex nature of battery operations and the extensive, labor-intensive nature of battery aging experiments. Consequently, available datasets often include only a limited number of batteries, necessitating the development of models based on sparse data. This scenario leads to the inherent challenges associated with training and testing models on datasets that exhibit varying data distributions.

Domain adaptation and transfer learning emerge as crucial strategies in RUL estimation, addressing the gap between limited data availability and the need for accurate predictive models. The ultimate goal is: to train a model on a comprehensive dataset and ensure its effective performance on a smaller target dataset. The choice between domain adaptation and transfer learning hinges on the specific attributes of the target dataset, especially the presence or absence of labeled data.

Domain adaptation involves training a model on a source domain, characterized by a substantial dataset, while integrating elements of the target domain data into the training. This approach enables the model to adapt to the unique characteristics and patterns of the target domain. It operates under the assumption that the source and target domains differ primarily in their data distributions and that a single hypothesis with minimal error is applicable across both domains.

Transfer learning, on the other hand, emphasizes the use of pre-trained models or their components. It involves transferring knowledge from a source task to a target task, with fine-tuning as a crucial element. Fine-tuning allows the model to adjust to the specifics of the target domain without necessitating complete retraining from the ground up.

Domain adversarial neural networks have emerged as an innovative approach in this field. DANNs aim to design a network's hidden layers to learn representations that are predictive of source labels while remaining neutral regarding the data's domain origin. It involves train-

ing a neural network for dual objectives: accurate label prediction in the source domain and, through adversarial learning, ensuring that the feature representations do not favor any specific domain.
Despite these advancements, the literature reveals limited exploration of domain adversarial learning applied to RUL estimation for batteries. Our research seeks to bridge this gap. We aim to develop a model, informed by the DANN methodology [1], that maintains high accuracy across various deployment scenarios, thus addressing the challenges of limited data diversity in battery RUL estimation.

4. Features

In our study, we propose a set of features derived from battery aging tests. These features are universally applicable and not tied to a specific dataset, enabling the application of our model across various datasets and under different conditions. The set includes the following key features:

- **Discharge Capacity:** This is the maximum discharge capacity of the battery at the cycle in question. It serves as an indicator of the battery's capacity degradation over its lifespan, reflecting how the battery's ability to hold charge diminishes with use and it is measured in milliampere-hour.
- **Cycle Number:** This represents the number of charging and discharging cycles the battery has undergone. It provides insight into the operational age of the battery, indicating how extensively it has been used.
- **Temperature:** The average temperature during the cycle, measured in degrees Celsius. Lithium-ion batteries are particularly sensitive to temperature fluctuations, making this a critical factor in assessing their health and performance.

These features collectively offer a comprehensive view of the battery's condition and are instrumental in our model's ability to evaluate battery health and predict its RUL under varying operational circumstances.

5. Model

In this research, we introduce a model characterized by a ConvLSTM model, enhanced with an attention mechanism for effectively predicting the RUL of batteries. The input to the model is a sliding window of adjustable size, encompassing features from the most recent cycles. The model's architecture is structured in three main components:

- **Convolutional layers:** convolutional layers are crucial for processing high-dimensional data. They efficiently extract and learn complex features from inputs like cycle numbers, temperature, and discharge capacity, making them vital in our deep-learning architecture.
- **LSTM layers with attention:** LSTM layers are adept at recognizing long-term patterns in sequential data. The integration of attention mechanisms allows the model to focus on specific segments of the sequence. This feature is particularly beneficial for understanding sequence-based insights in our model.
- **Fully connected layers for regression:** these layers process the features extracted by previous layers to produce the final output. In our model, they perform regression tasks, predicting continuous values based on the processed input, thus determining the RUL of the battery.

These modules form the backbone of our model, enabling it to reliably predict the RUL of batteries in varying operational scenarios.

5.1. Hyperparameter Tuning

The selection of appropriate hyperparameters is fundamental to the performance of deep learning models. Manual tuning can be cumbersome, and time-consuming, and may not always lead to good results. Automated tuning tools not only save time but also employ sophisticated algorithms to explore a broader range of hyperparameters more efficiently than manual tuning. Among these tools, **KerasTuner** stands out as an advanced method specifically tailored for the hyperparameter tuning of deep learning models implemented in Keras. With its array of features and its alignment with the Keras ecosystem, it emerged as the logical choice for ensuring our model reached its optimal performance. We employed Keras Tuner for systematic optimization, utilizing random search, Bayesian optimization, and hyperband strategies, with hyperband that, despite its longer search times, proved more effective due to its ability to rapidly identify promising hyperparameter configurations.

6. Domain Adversarial Neural Network

The core idea of DANNs is to train a network in such a way that it cannot distinguish between source and target domain data, thereby ensuring that the features it learns are equally applicable to both. A DANN consists of three primary components: a feature extractor, a regressor, and a domain classifier. Its strength lies in ensuring the feature extractor produces features beneficial for label prediction, while also being domain-agnostic. We can see the structure in Figure 1.

- **Working mechanism:** DANNs operate by integrating a domain classifier with a feature extractor and label predictor. The domain classifier, connected to the feature extractor via a gradient reversal layer, actively works against the feature extractor to ensure that the extracted features do not favor any particular domain. This is done by reversing the direction of the gradient during backpropagation from the domain classifier.
- **Application in RUL estimation:** the utility of DANN in RUL estimation lies in its ability to handle variations between different data sources, a common challenge in battery life prediction.
- **Implementation in thesis:** in our thesis, the domain adversarial neural network is directly connected to our previous model described in Section 5. Empirical observations indicate that connecting it to the output of the encoder in the feature extractor layer improves overall performance. The domain classification was effectively achieved using the sigmoid activation function.
- **The training strategy:** it involves an adversarial approach, balancing the minimization of classification loss while maximizing domain distinction loss. We fine-tuned the model's domain adaptation capabilities, through a λ parameter that affects the weight of the domain classifiers inverted gradient, for the efficient learning of the network.

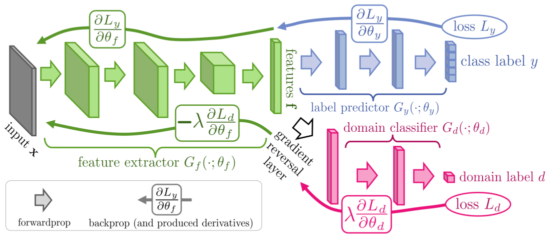

Figure 1: DANN basic implementation.

7. Experiments

7.1. Datasets

The *MIT-Toyota* dataset was collected for a joint study between the Massachusetts Institute of Technology and Toyota Motor Corporation. The dataset is composed of 124 commercial lithium-ion batteries that were cycled to failure through a fast charge policy to obtain the relevant data.
Since the objective was to have enough diverse data for the study, different charging policies have been applied. The conjunction of this factor with the strong nonlinearity of battery degradation led to a very heterogeneous dataset, where the life duration of batteries ranges from 150 to 2200 cycles. The dataset is divided into three batches, representing approximately 48 cells each. Each batch was defined by a batch date, the date the tests were started. The dataset has been sliced in three because the same testing equipment was used for all.
We simply kept the division in batches and used categorical numbering to identify them.
Each batch ends up having a relevant difference from the others in the feature's distribution, making the unique dataset more similar to 3 diverse datasets combined. This poses the challenge and the opportunity of training a model on one or two batches and applying the domain adaptation techniques. The diversity previously mentioned is reflected in the different distributions of the adopted features: *cycles to failure, average temperature, discharge capacity.* We can see an example of such diversity in the distribution of life cycles in Figure 2
In the preprocessing phase, the MIT Toyota Research Group initially removed cells not reaching 80% capacity. We further cleaned the data by correcting outliers and applied a smoothing algorithm to the charge-discharge curves. Our analysis focused on the last 800 cycles, prioritizing

Table 1: Performance on target domain (Batch 1) with a model trained on Batch 2-3.

Metrics	MAE	RMSE	MAE 50
test b2-b3	29.6	52.6	4.5
test b1	175.2	183.7	14.3

Table 2: Performance on target domain (Batch 1) after training on batches 2-3 and fine-tuning for 10 epochs on batch 1.

Metrics	MAE	RMSE	MAE 50
test b1	108.1	104	5.4

data toward the end of the batteries' operational life for accurate RUL estimation. Our research also underscored the efficacy of traditional data augmentation techniques, like jittering, in enhancing model performance [2].

7.2. Transfer Learning and Fine-Tuning

Our study employed transfer learning to adapt our model to the domain-specific characteristics of lithium-ion battery datasets, training initially on two batches. The fine-tuning process is conducted in two phases, aimed at refining the model's alignment with the third batch's data. The initial fine-tuning phase, involving 10 epochs, led to notable performance improvements. However, an extended phase of 20 epochs did not yield further significant enhancements, indicating the model's adaptation capabilities right from the initial phase. The changes in metrics across different stages of the study are illustrated in Table 1 to Table 3.

Transfer learning proved effective in our study

Table 3: Fine-tuning extended results, target domain (Batch 1) after +20 Epochs of fine-tuning.

Metrics	MAE	RMSE	MAE 50
(10 epochs)	108.1	104	5.4
(30 epochs)	108.9	104	4.6

due to its ability to leverage pre-trained models on the source domain, exploiting shared underlying relationships with the target domain. However, fine-tuning, especially with limited labeled data, excelled in adapting to the unique characteristics of the target domain, resulting in enhanced predictive accuracy.

7.3. DANN Training

The primary goal in this phase was to minimize prediction loss while simultaneously maximizing domain distinction loss, a process managed by the gradient reversal layer and its λ parameter. This approach led the feature extractor to produce domain-invariant features. A similar approach was pursued in [4] but with a slightly different model.

While trying to optimize DANN's performance, we tested various approaches for tuning the λ parameter. Starting with a fixed λ proved ineffective, as did dynamically modifying it during training. However, gradually decreasing λ from a high to a low value over time successfully led to model convergence.

A critical design decision was the attachment point of the domain classifier. After extensive testing, it was connected to the output of the encoder, excluding the attention and decoder layers. This choice was informed by the literature

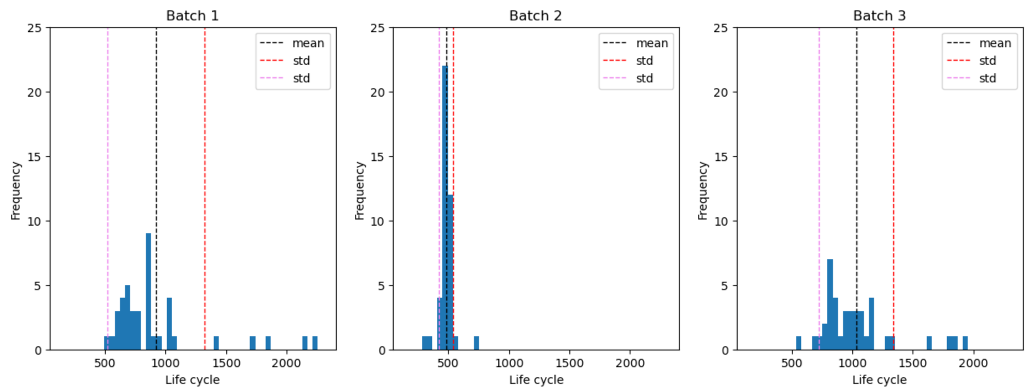

Figure 2: Distribution of the number of cycles to failure for each batch.

and empirical-based considerations, as the encoder's output was most suitable for our time series prediction task.

Table 4: DANN trained on batches 2-3.

Batch	RMSE	MAE	MAE 50
source b2-b3	90.6	64.8	14.6
target b1	166.1	124.4	10.0

7.4. Comparison

Upon training, the DANN's performance was juxtaposed against two other strategies, namely transfer learning and fine-tuning. The results presented in the above tables reveal that the performance is quite similar across the different methodologies adopted, with the fine-tuning model emerging as the leader in performance, closely followed by DANN and then basic transfer learning.

8. Conclusions

In this thesis, we critically examined domain adaptation in RUL estimation of lithium-ion batteries, with a focus on employing a DANN. Our study diverged from traditional transfer learning by leveraging DANN's unsupervised learning capabilities to align feature distributions across different domains. This method showed promise in domain adaptation, particularly in its ability to generalize without needing explicit labels.

However, we observed that in scenarios with limited data, such as our case study involving only few battery cells for the target domain, fine-tuning outperformed DANN. This was due to fine-tuning's capacity to adapt more closely to specific characteristics of the target domain, whereas DANN's emphasis on domain invariance could potentially overlook important domain-specific nuances.

Furthermore, we developed a systematic approach for hyperparameter optimization in our ConvLSTM model, which significantly improved prediction accuracy and efficiency.

The findings from our study underline the critical role of deep domain adaptation in enhancing the accuracy of RUL estimation, while also cautioning against the indiscriminate application of transfer learning due to potential performance losses.

Future directions for this research could include exploring new data augmentation methods, possibly using Generative Adversarial Networks (GANs) for synthetic data creation [2], or investigating the applicability of transformer architectures in battery health management, an area where they have not been extensively used but hold significant potential [4].

References

[1] Yaroslav Ganin, Evgeniya Ustinova, Hana Ajakan, Pascal Germain, Hugo Larochelle, François Laviolette, Mario Marchand, and Victor Lempitsky. Domain-adversarial training of neural networks. *The journal of machine learning research*, 17(1):2096–2030, 2016.

[2] Guillermo Iglesias, Edgar Talavera, Ángel González-Prieto, Alberto Mozo, and Sandra Gómez-Canaval. Data augmentation techniques in time series domain: a survey and taxonomy. *Neural Computing and Applications*, 35(14):10123–10145, May 2023.

[3] MS Hossain Lipu, MA Hannan, Aini Hussain, MM Hoque, Pin J Ker, Mohamad Hanif Md Saad, and Afida Ayob. A review of state of health and remaining useful life estimation methods for lithium-ion battery in electric vehicles: Challenges and recommendations. *Journal of cleaner production*, 205:115–133, 2018.

[4] Xin Liu, Changbo Yang, Yanmei Meng, Jihong Zhu, and Yijian Duan. Capacity estimation of li-ion battery based on transformer-adversarial discriminative domain adaptation. *AIP Advances*, 13(7), 2023.

[5] Kristen A Severson, Peter M Attia, Norman Jin, Nicholas Perkins, Benben Jiang, Zi Yang, Michael H Chen, Muratahan Aykol, Patrick K Herring, Dimitrios Fraggedakis, et al. Data-driven prediction of battery cycle life before capacity degradation. *Nature Energy*, 4(5):383–391, 2019.