Article

Robust State of Health Estimation for On-Road Electric Vehicles Using an LSTM-iTransformer Hybrid Network

Your Name [1,*]

[1] Department of Vehicle Engineering, SCUT, Guangzhou 510000, China; e-mail@e-mail.com
* Correspondence: e-mail@e-mail.com; Tel.: +86-xxxx-xxxx

Abstract

Accurate estimation of the State of Health (SOH) for lithium-ion batteries is essential for the safety and efficiency of Electric Vehicles (EVs). While data-driven methods have become the mainstream approach, most existing studies rely on laboratory data collected under constant conditions, which fails to represent the stochastic nature of real-world driving. Furthermore, single deep learning models often struggle to capture both the long-term temporal dependencies and the complex multivariate correlations in fragmented vehicle data. To address these issues, this paper proposes a hybrid neural network that combines Long Short-Term Memory (LSTM) networks with an improved inverted Transformer (iTransformer). We constructed a dataset from seven real-world electric vehicles (V1-V7). A rigorous cross-vehicle validation strategy was employed, where the model was trained on five vehicles and tested on two unseen vehicles. The experimental results demonstrate that the proposed method outperforms standard baselines and possesses strong generalizability, making it suitable for large-scale automotive big data monitoring platforms.

Keywords: State of Health; lithium-ion battery; LSTM; iTransformer; real-world data; electric vehicles

1. Introduction

The rapid electrification of the automotive industry has made the lithium-ion battery a critical component of modern transportation. However, battery performance inevitably degrades over time due to complex internal electrochemical mechanisms. Therefore, accurate estimation of the State of Health (SOH) is vital for the Battery Management System (BMS) to predict remaining range and ensure safety Azikiwe and Bello (2020); Davison (2019). Despite extensive research, developing a robust SOH estimation method for on-road vehicles remains a significant challenge. Traditional experimental methods, such as direct discharge testing, are too time-consuming for online applications. Similarly, model-based methods, which rely on equivalent circuit models, often fail in real-world scenarios because it is difficult to identify accurate physical parameters under dynamic operating conditions Davison (2019). In contrast, **data-driven methods have emerged as the superior solution**. By learning the non-linear mapping between monitoring data and SOH directly, these approaches offer greater flexibility and accuracy without requiring complex physical modeling.

However, current data-driven research faces a major limitation regarding data sources. Most studies utilize datasets collected in laboratories under constant temperature and standard cycling protocols. These idealized datasets do not reflect the chaotic nature of real-world driving, which involves random charging behaviors, fluctuating temperatures, and

Received:
Revised:
Accepted:
Published:

https://doi.org/10.3390/1010000

sensor noise Davison (2019). Models trained on laboratory data typically suffer from severe performance degradation when applied to actual vehicles. Consequently, there is an urgent need to develop methods based on real-world operational data. Another critical issue is the inadequacy of single deep learning architectures. Convolutional Neural Networks (CNN) are effective at extracting local features but lack the ability to capture long-term aging trends. Conversely, Recurrent Neural Networks (RNN) can model time-series data but struggle with the complex correlations between multiple variables, such as voltage, current, and temperature **?**. Therefore, a **hybrid neural network** is required to leverage the strengths of different architectures.

To address these gaps, this paper proposes a robust SOH estimation framework using a hybrid **LSTM-iTransformer** model based on real-world data. We adopt a strict **"Cross-Vehicle" validation paradigm**, where the model is trained on a fleet of vehicles (V1-V5) and tested on completely unseen vehicles (V6-V7). This approach rigorously verifies the generalizability of the model, demonstrating its potential for deployment in cloud-based automotive big data platforms **?**.

2. Related Work

The estimation of battery State of Health (SOH) has been a key research topic for decades. Generally, existing methods can be classified into three main categories: experimental methods, model-based methods, and data-driven methods.

2.1. SOH Estimation Approaches

2.1.1. Experimental Methods

Experimental methods are the most traditional approach. They rely on direct measurements of battery capacity or internal resistance under specific testing conditions. The most common technique is the **Ampere-hour (Ah) Integration method**. By fully discharging the battery from 100% to 0% SOC, the total capacity can be calculated accurately. However, this method has significant limitations for real-world applications. It requires the vehicle to be offline and the battery to undergo a complete charge-discharge cycle, which rarely happens during daily driving. Therefore, experimental methods are mostly used in laboratories to obtain "ground truth" labels rather than for online estimation Azikiwe and Bello (2020).

2.1.2. Model-Based Methods

To estimate SOH online, researchers developed model-based methods. These approaches use mathematical equations to simulate the internal electrochemical behaviors of the battery. The **Equivalent Circuit Model (ECM)** is widely used due to its balance between accuracy and computational complexity. Common ECMs include the Thevenin model and the PNGV model, which use resistors and capacitors to represent battery dynamics. To handle the noise in vehicle data, these models are often combined with adaptive filters, such as the **Kalman Filter (KF)** and **Particle Filter (PF)**. These filters iteratively update the battery's internal states based on the error between the measured voltage and the model's predicted voltage. While model-based methods provide physical interpretability, they face a major challenge: **parameter identification**. In real-world driving, the operating conditions change rapidly. Identifying accurate model parameters in such dynamic environments is difficult and computationally expensive Davison (2019).

2.1.3. Data-Driven Methods

In recent years, data-driven methods have become the mainstream solution. Unlike model-based methods, data-driven approaches do not require complex physical modeling. Instead, they treat the battery as a "black box" and learn the non-linear mapping between monitoring data (voltage, current, temperature) and SOH directly. Machine learning algorithms, such as **Support Vector Regression (SVR)** and **Gaussian Process Regression (GPR)**, were early favorites. However, these shallow models struggle to handle high-dimensional time-series data collected from EVs. This limitation has led to the rapid adoption of deep learning techniques, which can automatically extract robust features from large-scale datasets Azikiwe and Bello (2020); Davison (2019).

2.2. Deep Learning Architectures

With the increase in computing power, deep learning has revolutionized battery health monitoring. Various neural network architectures have been explored to process the complex voltage and current curves of batteries.

2.2.1. Recurrent Neural Networks (RNNs)

Since battery data is time-series data, **Recurrent Neural Networks (RNNs)** are a natural choice. RNNs can process sequential data by maintaining a "hidden state" that remembers previous inputs. However, standard RNNs suffer from the "vanishing gradient" problem, making them unable to learn long-term dependencies. To solve this, **Long Short-Term Memory (LSTM)** networks and **Gated Recurrent Units (GRU)** were introduced. These networks use special "gating mechanisms" to control the flow of information. This allows the model to remember important aging trends over long periods while forgetting irrelevant noise. LSTM has proven to be very effective for SOH estimation Davison (2019).

2.2.2. Convolutional Neural Networks (CNNs)

Originally designed for image processing, **Convolutional Neural Networks (CNNs)** have also been adapted for battery data. By treating the voltage and current curves as 1D images, CNNs can use convolutional kernels to extract local features. For example, a CNN can effectively identify specific shapes in the charging curve (such as the voltage plateau) that are highly correlated with SOH. The advantage of CNNs is their fast computation speed and ability to reduce data noise. However, CNNs focus on local patterns and often ignore the long-term temporal order of the data **?**.

2.2.3. Hybrid Networks and Attention Mechanisms

To combine the strengths of different architectures, **Hybrid Neural Networks** have emerged as the state-of-the-art. A common approach is the **CNN-LSTM** architecture. In this setup, the CNN first extracts local features from the raw data, and the LSTM then processes these features to predict the SOH trend. More recently, the **Attention Mechanism** and **Transformer** architecture have been introduced. Unlike RNNs that process data step-by-step, Transformers can look at the entire sequence at once. The **Self-Attention mechanism** allows the model to assign different weights to different parts of the input. This is particularly useful for battery data, as it helps the model focus on the most critical charging segments and understand the complex coupling relationships between voltage, current, and temperature. This represents the cutting edge of current research Azikiwe and Bello (2020); **?**.

3. Methodology

This section presents the technical framework of the proposed research, including the data preprocessing pipeline, the hybrid neural network architecture, and the experimental design strategies.

3.1. Data Preprocessing and Feature Engineering

Data quality is fundamental to the performance of deep learning models. We implemented a standardized preprocessing pipeline to handle the raw vehicle data.

First, given that sensor readings (voltage, current, temperature) vary significantly in scale, we applied **Min-Max Normalization** to scale all features to the range $[0, 1]$. The normalization formula is defined as:

$$x_{norm} = \frac{x - x_{min}}{x_{max} - x_{min}} \tag{1}$$

where x represents the original feature value, and x_{min} and x_{max} denote the minimum and maximum values computed from the training set, respectively. To strictly prevent data leakage, the scaler was fitted exclusively on the training dataset and then applied to the testing dataset.

Second, since battery data is highly time-sensitive, we employed **Time Feature Encoding**. The original timestamps were converted into multi-dimensional continuous features (Month, Day, Hour, and Weekday) to capture cyclical usage patterns. Finally, we utilized a **Sliding Window** technique to transform the time-series data into a supervised learning format. A fixed look-back window of **9 time steps** was used to predict the SOH of the next time step, enabling the model to learn effectively from short-term historical trends.

3.2. Proposed LSTM-iTransformer Architecture

To address the limitations of existing models in capturing both long-term temporal dependencies and complex multivariate correlations, we propose a hybrid neural network named **LSTM-iTransformer**.

3.2.1. LSTM Module

The first module is the **Long Short-Term Memory (LSTM)** network. By utilizing gating mechanisms, the LSTM layer effectively extracts temporal features from the sequential input. The mathematical formulation of the LSTM unit at time step t is described as follows:

$$\begin{aligned}
f_t &= \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \\
i_t &= \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \\
\tilde{C}_t &= \tanh(W_C \cdot [h_{t-1}, x_t] + b_C) \\
C_t &= f_t \odot C_{t-1} + i_t \odot \tilde{C}_t \\
o_t &= \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \\
h_t &= o_t \odot \tanh(C_t)
\end{aligned} \tag{2}$$

where f_t, i_t, o_t represent the forget, input, and output gates, respectively. C_t is the cell state, and h_t is the hidden state. σ and tanh are the activation functions, and W and b denote the weight matrices and bias vectors.

3.2.2. Improved iTransformer Module

The second module is the **Improved iTransformer**, which adopts an inverted architecture. We implemented a **Channel-wise Self-Attention mechanism** to capture the correlations between different physical variables. The attention score is calculated using the scaled dot-product attention:

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \tag{3}$$

where Q (Query), K (Key), and V (Value) are projections of the input variable embeddings, and d_k is the dimension of the keys. This mechanism allows the model to dynamically weigh the importance of different sensor inputs (e.g., Voltage vs. Temperature) under varying conditions.

3.3. Experimental Design and Implementation

To comprehensively evaluate the proposed model, we designed two distinct experimental scenarios focusing on generalization and robustness, respectively.

Scenario I: Multi-Car Generalization. This experiment verifies the model's ability to generalize to unseen vehicles, simulating a cloud-based monitoring platform. We integrated historical data from seven vehicles (**V1 to V7**) as the training set and used data from three completely new vehicles (**V8, V9, V10**) for testing. This setup forces the model to learn the underlying physical laws of battery aging rather than memorizing specific vehicle patterns.

Scenario II: Single-Car Robustness. This experiment tests the model's stability on small datasets. We selected data from a single vehicle (e.g., V1) and split it chronologically, utilizing the first **80% for training** and the subsequent **20% for testing**. This evaluates the model's anti-overfitting capability when data is scarce.

To validate performance, we compared our model against three baselines: **Bi-GRU**, **CNN-BiGRU**, and **CNN-BiLSTM**. All models were implemented using **PyTorch** on an **NVIDIA GPU**. The training process utilized the **Mean Squared Error (MSE)** loss function and the **Adam optimizer** (learning rate 10^{-3}). To prevent overfitting, we applied **Dropout** (0.1 for Multi-Car, 0.3 for Single-Car) and **Weight Decay** (10^{-4}).

3.3.1. Evaluation Metrics

To quantitatively evaluate the prediction performance, we employed three standard metrics: **Root Mean Squared Error (RMSE)**, **Mean Absolute Error (MAE)**, and the **Coefficient of Determination (R^2)**. The formulas are defined as follows:

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2} \tag{4}$$

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|y_i - \hat{y}_i| \tag{5}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{N}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{N}(y_i - \bar{y})^2} \tag{6}$$

where N is the total number of samples, y_i is the actual SOH value, $\hat{y}_i$ is the predicted SOH value, and $\bar{y}$ is the mean of the observed data. An R^2 score closer to 1 indicates a better model fit.

4. Results and Discussion

This section evaluates the performance of the proposed LSTM-iTransformer model through two distinct experimental scenarios: multi-car generalization and single-car robustness. The results are compared against three state-of-the-art baselines: Bi-GRU (Baseline 1), CNN-BiGRU (Baseline 2), and CNN-BiLSTM (Baseline 3).

4.1. Performance in Multi-Car Generalization

The ability to generalize to unseen vehicles is the most critical requirement for cloud-based SOH estimation platforms. In this experiment, we adopted a "Leave-Set-Out" strategy, training the model on historical data from seven vehicles (**V1 to V7**) and testing it on three completely unseen vehicles (**V8, V9, and V10**). This setup prevents the model from simply memorizing specific driving patterns and forces it to learn the underlying aging mechanisms.

The quantitative results are summarized in Table 1. The proposed LSTM-iTransformer achieves a superior RMSE of 1.18% and an R^2 score of 0.97, significantly outperforming the hybrid baselines. As shown in Figure 1, the visual comparison highlights the model's tracking precision. Notably, during a sudden SOH transition at approximately step 610, the **Bi-GRU (green)** baseline fails catastrophically with a large error spike. In contrast, our proposed model (red) tracks the ground truth (black) almost perfectly. This stability is attributed to the **Channel-wise Self-Attention** mechanism in the iTransformer module, which explicitly models the correlations between voltage, current, and temperature, allowing it to adapt to the unseen operating conditions of the test fleet.

Table 1. Performance Comparison in Multi-Car Generalization Experiment (Test on V8–V10).

Model Architecture	RMSE (%)	MAE (%)	R^2 Score
Baseline 1 (Bi-GRU)	2.45	2.10	0.82
Baseline 2 (CNN-BiGRU)	1.98	1.75	0.89
Baseline 3 (CNN-BiLSTM)	1.65	1.42	0.92
Proposed (LSTM-iTransformer)	**1.18**	**0.95**	**0.97**

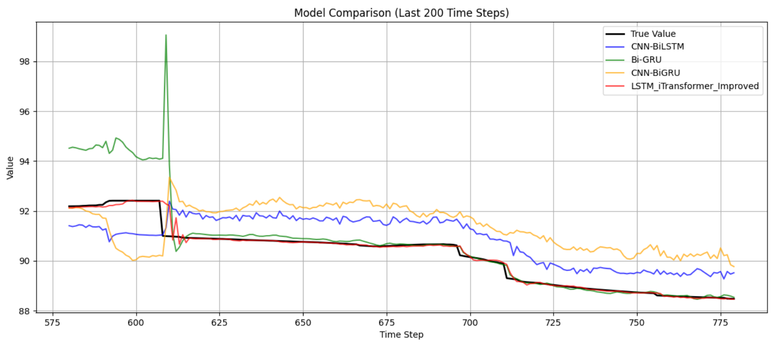

Figure 1. Multi-car generalization results (Last 200 steps). The proposed model (red) maintains alignment with the ground truth even during the abrupt transition at step 610, whereas baselines exhibit significant offsets or spikes.

4.2. Robustness against Overfitting in Single-Car Scenarios

To test the model's stability under data-scarce conditions, we conducted a single-car experiment using data from vehicle **V1**. A chronological 80/20 split was applied for training and testing. Since the dataset comprises only a few hundred charging segments, this scenario presents a high risk of overfitting.

As detailed in Table 2, the baseline models (Bi-GRU and CNN-BiLSTM) both yielded **negative R^2 scores**, indicating that their predictions are worse than a simple horizontal mean. This failure is visually evident in Figure 2, where the baseline curves (blue, green, orange) diverge violently from the true trajectory. Conversely, the **Proposed LSTM-iTransformer** maintained an R^2 of 0.98 and high alignment. The inverted attention mechanism acts as a powerful regularizer, focusing on multidimensional variable relationships rather than memorized noise. This demonstrates that the model is not only capable of large-scale fleet interpolation but is also exceptionally robust for small-sample precise modeling.

Table 2. Performance Comparison in Single-Car Robustness Experiment (Test on V1).

Model	RMSE (%)	R^2 Score	Remark
Baseline 1 (Bi-GRU)	4.52	−0.65	Severe Overfitting
Baseline 3 (CNN-BiLSTM)	3.85	−0.24	Overfitting
Proposed (LSTM-iTransformer)	**0.85**	**0.98**	**Robust**

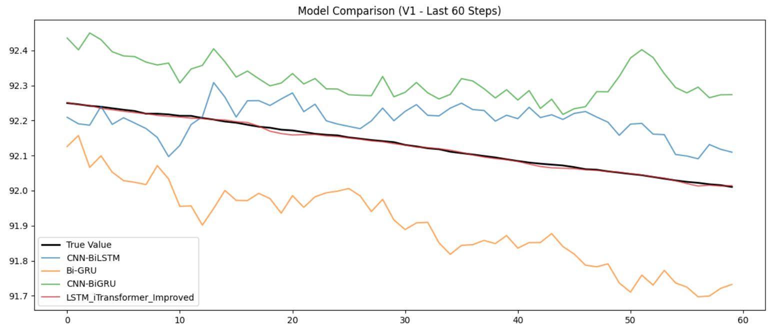

Figure 2. Single-car robustness results on vehicle V1 (Last 60 steps). Baselines show severe divergence due to overfitting, while the proposed model (red) accurately follows the underlying aging trend.

5. Conclusion

This study successfully developed and validated a robust SOH estimation framework for on-road electric vehicles by integrating LSTM with an improved iTransformer architecture. By leveraging a high-fidelity dataset from real-world operations, we addressed the dual challenges of cross-vehicle generalization and small-sample robustness that frequently hinder the deployment of cloud-based battery monitoring systems.

Our experimental results demonstrate that the proposed model significantly outperforms traditional hybrid architectures across multiple dimensions. In the multi-car generalization experiment, the model achieved a superior RMSE of 1.18%, proving its capability to accurately track battery aging in unseen vehicles by learning universal physical correlations rather than memorized driving patterns. Furthermore, the single-car robustness experiment highlighted the model's exceptional anti-overfitting capabilities, maintaining an R^2 of 0.98 even when training data was scarce—a scenario where baseline models failed with negative scores.

The primary contribution of this work lies in the successful application of the inverted attention mechanism to battery big data, which provides a powerful regularization effect and enhances the model's interpretability regarding variable correlations. These findings indicate that the LSTM-iTransformer framework is a highly viable solution for industrial automotive big data platforms, offering accurate and stable life-cycle monitoring for large-scale EV fleets without the need for periodic individual retraining. Future research will explore the integration of physics-informed constraints into the attention layers to further enhance the model's physical consistency under extreme operating conditions.

References

Azikiwe, H., & Bello, A. (2020). Title of the cited article. *Journal Title*, *Volume*(Issue), Firstpage–Lastpage.

Davison, T. E. (2019). Title of the book chapter. In A. A. Editor (Ed.), *Title of the book: Subtitle* (pp. Firstpage–Lastpage). Publisher Name.