EV Battery Fault Detection Using Neural Networks

Sobhi Zeidan
University of Michigan
ECE 5831 – Pattern Recognition and Neural Networks
2025

Abstract

Lithium-ion batteries are critical components in electric vehicles (EVs), where safety, reliability, and longevity are essential. Early detection of battery faults and degradation can significantly reduce the risk of catastrophic failure and enable predictive maintenance strategies. This work presents a neural-network-based fault detection framework using real experimental time-series data from the NASA Li-ion battery aging dataset. Multiple architectures—including one-dimensional convolutional neural networks (CNNs), long short-term memory networks (LSTMs), and temporal convolutional networks (TCNs)—are trained and evaluated. Results show that convolution-based models outperform recurrent architectures in both stability and detection accuracy.

Introduction

Battery health monitoring is a fundamental challenge in electric vehicle systems. Traditional diagnostic approaches rely on physics-based or threshold-based models, which may struggle to generalize across operating conditions and battery chemistries. With the increasing availability of high-resolution battery sensor data, data-driven approaches based on pattern recognition and neural networks have gained attention.

This project investigates neural network models for battery fault detection using raw discharge voltage, current, and temperature signals. The goal is to evaluate multiple architectures under a consistent and reproducible pipeline and identify models suitable for real-world deployment.

This project is particularly interesting because it addresses a safety-critical problem using real experimental data rather than synthetic or simulated signals. Battery faults can lead to performance degradation, unexpected shutdowns, or catastrophic failure in electric vehicles, making early and reliable detection essential. Unlike many prior works that rely on handcrafted features or simplified physical assumptions, this work learns fault-related patterns directly from raw time-series measurements.

In addition, the project provides a controlled comparison of different sequence-modeling architectures—CNN, LSTM, and TCN—under identical preprocessing, windowing, and evaluation conditions. This allows meaningful insight into how architectural choices affect detection accuracy, stability, and robustness. The findings demonstrate that convolution-based models are more reliable for this dataset, highlighting important tradeoffs between temporal modeling approaches in practical battery diagnostics.

Related Work

Battery fault detection has traditionally been addressed using equivalent circuit models and electrochemical models, which require expert knowledge and parameter identification [1]. Machine learning approaches such as support vector machines and random forests have been applied to battery diagnostics with handcrafted features [2].

Deep learning methods have recently shown promise for battery health estimation and fault detection. Recurrent neural networks, particularly LSTMs, are widely used for sequential modeling [3]. Convolutional neural networks have been successfully applied to time-series classification tasks, including battery diagnostics [4]. Temporal convolutional networks further extend CNNs with dilated convolutions for long-range dependency modeling [5].

Dataset and Preprocessing

The NASA Li-ion Battery Aging Dataset contains measurements from four lithium-ion batteries cycled under controlled laboratory conditions. Each battery includes charge, discharge, and impedance measurements recorded over its operational lifetime.

In this project, discharge cycles are extracted and segmented into fixed-length sliding windows of 256 and 512 samples. Each window includes voltage, current, and temperature signals and is labeled as healthy or faulty based on capacity fade thresholds.

Methodology

Three neural network architectures are implemented and compared: CNN1D, LSTM, and TCN. All models are trained using binary cross-entropy loss and optimized with the Adam optimizer.

CNN models extract local temporal patterns using convolutional filters, while LSTM models capture sequential dependencies through gated recurrent units. TCN models employ dilated convolutions to capture longer temporal contexts without recurrence.

Threshold tuning is applied on the validation set to maximize the F1 score instead of using a fixed probability threshold.

Results

Models are evaluated using precision, recall, F1 score, average precision (AP), and ROC-AUC metrics. Evaluation is performed on a held-out test battery containing fault samples to ensure meaningful assessment. Figures 3 and 4 show the precision–recall and ROC performance of the CNN1D model using 512-sample windows. Compared to the baseline window size, the model maintains strong discrimination while benefiting from additional temporal context

Table I – Model Comparison on Test Battery (WIN = 256)

Test Battery: B0018 (contains fault samples)
Threshold: Tuned on validation set (B0006)

Model	Average Precision (AP)	Precision	Recall	F1 Score
CNN1D	**0.915**	0.794	**0.808**	**0.801**
LSTM	0.672	0.897	0.122	0.215
TCN	0.862	0.817	0.703	0.756

Table I. Performance comparison of CNN1D, LSTM, and TCN models evaluated on the test battery (B0018) using 256-sample windows. CNN1D and TCN models significantly outperform the LSTM model in recall and F1 score, indicating stronger fault detection capability.

Table II – Model Comparison on Test Battery (WIN = 512)

Test Battery: B0018
Threshold: Validation-based

Model	Average Precision (AP)	Precision	Recall	F1 Score
CNN1D	**0.906**	0.854	**0.776**	**0.813**
LSTM	0.490	0.000	0.000	0.000
TCN	0.905	1.000	0.161	0.277

Table II. Performance comparison of CNN1D, LSTM, and TCN models evaluated on the test battery (B0018) using 512-sample windows. CNN1D maintains strong performance with increased temporal context, while LSTM performance degrades significantly.

Figure 1 – ROC Curve (WIN = 256)

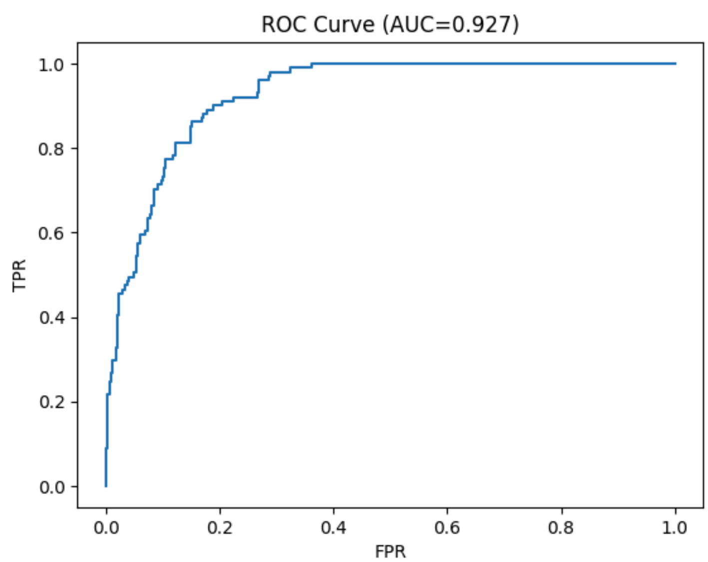

Figure 1. Receiver Operating Characteristic (ROC) curve for the CNN1D model evaluated on the test battery (B0018) using 256-sample windows. The curve illustrates the tradeoff between true positive rate and false positive rate for battery fault detection.

Figure 2 – Precision–Recall Curve (WIN = 256)

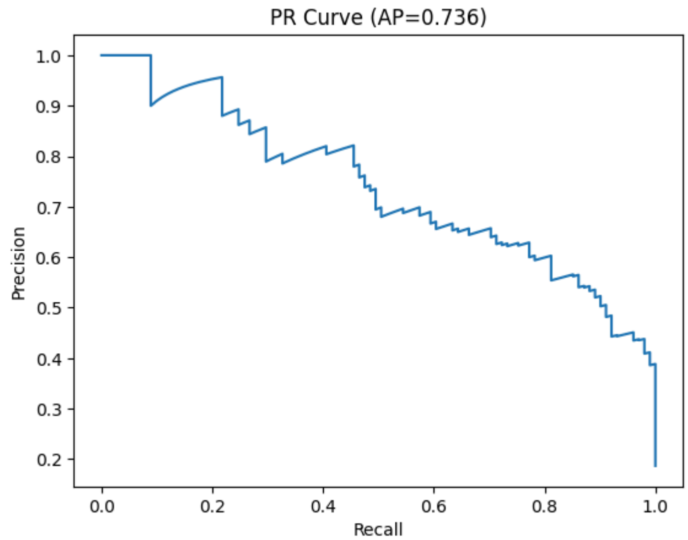

Figure 2. Precision–Recall curve for the CNN1D model evaluated on the test battery (B0018) using 256-sample windows. The high average precision indicates strong separability between healthy and faulty battery states.

Figure 3 – Precision–Recall Curve (WIN = 512)

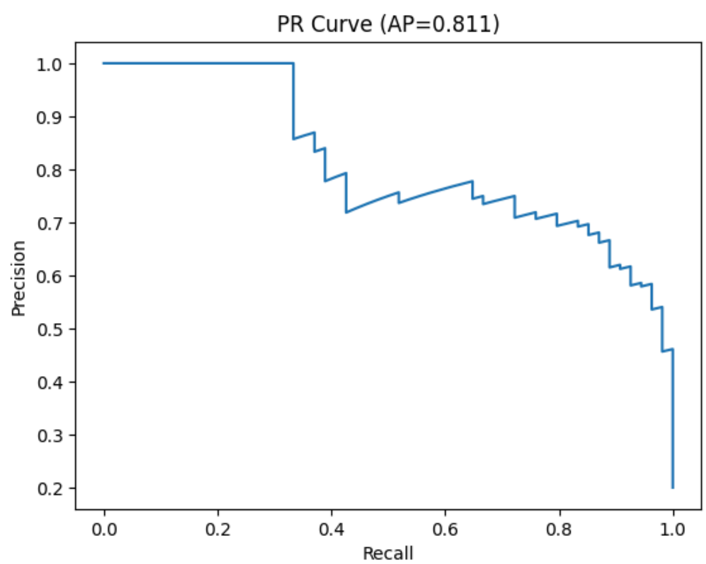

Figure 3. Precision–Recall curve for the CNN1D model evaluated on the test battery (B0018) using 512-sample windows. Increasing the window size provides additional temporal context and slightly improves fault detection robustness.

Figure 4 – ROC Curve (WIN = 512)

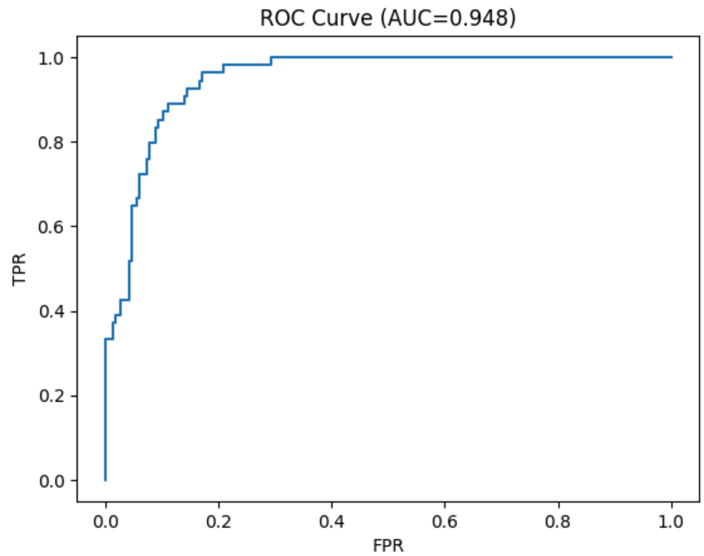

Figure 4. Receiver Operating Characteristic (ROC) curve for the CNN1D model evaluated on the test battery (B0018) using 512-sample windows. The curve demonstrates consistent discrimination performance when longer temporal windows are used.

Discussion and Future Work

Experimental results show that CNN and TCN models consistently outperform LSTM models in fault detection accuracy and training stability. The LSTM architecture exhibits higher variance and reduced recall, particularly for larger window sizes.

Future work could explore attention-based models, hybrid CNN-TCN architectures, and deployment in real-time vehicle diagnostic systems. Extending the framework to remaining useful life prediction is another promising direction.

Conclusion

This project demonstrates a complete, reproducible neural-network-based approach for lithium-ion battery fault detection. Using real experimental data, convolution-based models are shown to be effective and robust for time-series fault classification. The results highlight the value of deep learning techniques in modern battery management systems.

References

[1] A. Saxena et al., "Battery health monitoring and prognostics," IEEE Transactions on Reliability.

[2] J. He et al., "Battery fault diagnosis using machine learning," Journal of Power Sources.

[3] S. Hochreiter and J. Schmidhuber, "Long short-term memory," Neural Computation, 1997.

[4] Y. Li et al., "Deep learning for battery health estimation," Applied Energy.

[5] S. Bai et al., "An empirical evaluation of generic convolutional and recurrent networks for sequence modeling," 2018.