International Journal of Engineering Research and Science & Technology

www.ijerst.org

ISSN : 2319-5991

Vol. 22 No. 2(2) (2026)



ijerst.editor@gmail.com

editor@ijerst.com

International Journal of Engineering Research and Science & Technology
ISSN 2319-5991 www.ijerst.com
© 2026 IJERST. All Rights Reserved
Peer Reviewed, Referred & Indexed Journal

Research Paper

A Bayesian Optimized Deep Learning Approach for Accurate State of Charge Estimation of Lithium-Ion Batteries Used for Electric Vehicle Applications

Pilli Renuka Devi[1] , Sarukonda Nanibabu[2], Beesetti Annapurna Yaswitha[3], Dr. C. P. V. N. J Mohan Rao[4]

[1-3] MCA, Department of Computer Applications, Avanthi Institute of Engineering and Technology, Makavarapalem, Narsipatnam, Anakapalle – 531113, Andhra Pradesh, India

[4]Professor, Department of Computer Applications, Avanthi Institute of Engineering and Technology, Makavarapalem, Narsipatnam, Anakapalle – 531113, Andhra Pradesh, India

Abstract—The rapid growth of electric vehicles (EVs) has sparked significant interest in battery technology, particularly in monitoring the State of Charge (SOC). Accurate SOC estimation is critical for safe and efficient battery operation. While various estimation methods exist, more research is needed to adapt to diverse lithium-ion battery chemistries and operating conditions. Deep learning (DL) has shown considerable promise in improving SOC estimation accuracy; however, selecting optimal hyperparameters remains a persistent challenge. This paper presents an automated hyperparameter tuning method using Bayesian optimization applied to three deep learning architectures: Long Short-Term Memory (LSTM), Gated Recurrent Unit (GRU), and Bidirectional LSTM (BiLSTM). The models incorporate battery data including current, voltage, temperature, capacity, and average voltage and current values to enhance estimation accuracy. Tested across varying temperature condit4ions using the Panasonic 10°C dataset, the proposed approach demonstrates that BiLSTM models with 70 hidden neurons achieve SOC predictions with Root Mean Square Error (RMSE) below 0.02, significantly improving reliability in battery management systems. A CNN2D extension achieves the best performance with RMSE of 0.0094, outperforming all other evaluated models.

Index Terms—State of Charge estimation, lithium-ion battery, Bayesian optimization, deep learning, BiLSTM, GRU, LSTM, CNN2D, electric vehicles, battery management system, hyperparameter tuning.

I. INTRODUCTION

Electric vehicles (EVs) have gained significant attention in recent years as countries seek to reduce carbon emissions and dependence on fossil fuels. The global push toward clean energy and sustainable transportation has placed EVs at the center of future mobility solutions. A key component of EV technology is the lithium-ion battery, which powers the vehicle and determines its overall efficiency and performance [1].

Efficient battery management is essential for maximizing the lifespan, safety, and reliability of these batteries. A critical aspect of battery management is accurately estimating the State of Charge (SOC) of the battery. SOC refers to the amount of charge remaining in the battery as a percentage of its total capacity. Accurate SOC estimation is essential for optimizing battery performance, improving energy efficiency, and ensuring the safety of EVs [2].

Lithium-ion batteries are preferred for EV applications due to their high specific energy, long lifespan, low self-discharge rate, and high

energy density. However, they are also highly sensitive to temperature variations and operating conditions, which can lead to inaccurate SOC estimation if not managed properly. Incorrect SOC estimation can result in reduced driving range, battery degradation, and safety hazards such as overheating or fire [3].

Traditionally, SOC estimation has been performed using methods such as Open Circuit Voltage (OCV), Ampere-Hour (Ah) counting, Kalman filters, and model-based approaches. While these methods have shown reasonable accuracy, each carries significant limitations. Kalman filters require highly accurate battery models that are difficult to develop due to the complex electrochemical behavior of lithium-ion cells. Ah counting is affected by measurement errors and does not account for battery aging or environmental factors. OCV methods require the battery to be in a stable state, which limits real-time SOC estimation capability [4].

To overcome these limitations, researchers have turned to data-driven approaches based on machine learning (ML) and deep learning (DL). Deep learning models—particularly Recurrent Neural Networks (RNNs) such as Long Short-Term Memory (LSTM) and Gated Recurrent Units (GRU)—have shown great promise in SOC estimation. These models are capable of capturing complex patterns and long-term dependencies in time-series data, making them well-suited for analyzing battery performance under different operating conditions [5].

Despite the potential of deep learning models, selecting optimal hyperparameters remains a significant challenge. Manual tuning through trial-and-error is time-consuming and often yields suboptimal results. This paper addresses this challenge by proposing a Bayesian optimization framework for automated hyperparameter tuning of deep learning models, specifically targeting the number of hidden neurons in LSTM, GRU, and BiLSTM architectures for SOC estimation of lithium-ion batteries used in EV applications.

A. Problem Statement

Existing BMS implementations using deep learning typically rely on Grid Search techniques to identify optimal hyperparameters. Grid Search may yield improper optimization parameters that degrade algorithm performance, incur high computational costs, require extensive data, and lack scalability for complex models [6]. There is a clear need for a smarter, more efficient hyperparameter optimization strategy that can systematically explore the search space and converge to near-optimal solutions with minimal computational overhead.

B. Contributions

The main contributions of this work are as follows:

1) A systematic evaluation of Bayesian optimization for automated hyperparameter tuning of deep learning models applied to battery SOC estimation.

2) Comparative analysis of LSTM, GRU, and BiLSTM architectures under Bayesian-optimized conditions across different neuron configurations (30, 50, 70, 100 neurons).

3) Extension of the framework with a CNN2D model, achieving state-of-the-art RMSE of 0.0094 on the Panasonic 10°C benchmark dataset.

4) Comprehensive performance evaluation using RMSE and Maximum Error metrics across multiple temperature conditions.

II. LITERATURE REVIEW

Liu et al. [7] introduced a Bayesian optimization framework for tuning hyperparameters in deep learning models for SOC prediction. By integrating Bayesian optimization, their model achieved higher accuracy and efficiency in estimating battery charge states, though the high computational cost of the optimization process was identified as a key limitation.

Wang et al. [8] integrated Long Short-Term Memory (LSTM) networks with Kalman filtering techniques to enhance SOC estimation. The Kalman filter reduced noise in battery measurements, improving estimation reliability. However, the primary drawback was the complexity in model calibration, which required extensive parameter tuning.

Zhang et al. [9] applied Gated Recurrent Units (GRU) to predict SOC in high-dimensional battery health datasets. GRUs effectively captured long-term dependencies in battery data, improving accuracy over traditional ML methods, though the method required large labeled datasets, posing challenges for real-world implementation.

Chen et al. [10] proposed a Particle Swarm Optimization (PSO)-based algorithm to optimize SOC estimation by selecting the most relevant features. The optimization process enhanced battery life cycle predictions and minimized prediction errors, with a major drawback being sensitivity to initial conditions, which could affect model stability.

Huang et al. [11] combined Convolutional Neural Networks (CNNs) with GRUs to leverage both spatial and temporal feature extraction. The hybrid model significantly improved prediction accuracy by analyzing sequential battery data, although training complexity increased as more parameters were introduced.

Xu et al. [12] applied BiLSTM networks to capture dependencies in both past and future battery data, enhancing accuracy by processing sequential data bidirectionally. A notable drawback was high memory consumption, which can be problematic for real-time applications.

Chemali et al. [13] demonstrated that LSTM networks achieve highly accurate SOC estimation for Li-ion batteries across various operating conditions, establishing a strong baseline for recurrent deep learning approaches in battery management. Tian et al. [14] provided a comprehensive review of deep learning frameworks for SOC estimation, identifying BiLSTM and attention-based models as the most promising directions for future research.

III. PROPOSED METHODOLOGY

A. System Overview

The proposed framework employs Bayesian Optimization to automatically identify the optimal number of hidden neurons for deep learning models used in SOC estimation. Unlike Grid Search, which exhaustively evaluates every parameter combination, Bayesian Optimization builds a probabilistic model (Gaussian Process) of the objective function and uses it to select the most promising hyperparameter values for evaluation, thereby converging to near-optimal solutions efficiently.

The overall workflow consists of: (1) dataset preparation and preprocessing, (2) model architecture design for LSTM, GRU, BiLSTM, and CNN2D, (3) Bayesian optimization-based hyperparameter tuning, (4) model training and evaluation, and (5) comparative performance analysis.

B. Dataset

The dataset used in this study was collected from battery testing experiments under different environmental conditions using Panasonic 18650PF lithium-ion cells. The dataset encompasses four temperature conditions: −10°C, 0°C, 10°C, and 25°C. For this study, the Panasonic 10°C dataset was selected for primary experimentation due to its representative operating conditions.

The dataset features include Voltage (V), Current (A), Temperature (°C), Capacity (Ah), Voltage_Average, and Current_Average. The target variable is Capacity, which represents the SOC. Missing values were filled with zeroes to maintain data consistency. The dataset was split into 80% for training and 20% for testing using stratified train-test splitting.

C. Data Preprocessing

To ensure uniform feature scaling, MinMaxScaler was applied to transform all input features to the range [0, 1]. Separate scalers were used for input features (X) and the target variable (Y) to prevent data leakage. The dataset was subsequently reshaped into three dimensions (samples × features × timesteps) to comply with the input requirements of recurrent neural network architectures.

D. Deep Learning Architectures

Four deep learning models were designed and evaluated:

1) LSTM: Two stacked LSTM layers with dropout regularization (rate = 0.3), followed by a dense output layer. The Adam optimizer and

Mean Squared Error (MSE) loss function were used for compilation.

2) GRU: An architecture mirroring the LSTM design but replacing LSTM layers with GRU layers. GRU is computationally more efficient than LSTM, with fewer trainable parameters, making it attractive for resource-constrained deployments.

3) BiLSTM: A Bidirectional LSTM architecture employing two Bidirectional GRU layers (forward and backward) with dropout regularization. BiLSTM captures temporal dependencies in both directions, improving feature learning from time-series battery data.

4) CNN2D: A 2D Convolutional Neural Network using Conv2D layers to capture spatial patterns in SOC data, with MaxPooling layers for dimensionality reduction and fully connected layers for prediction.

E. Bayesian Optimization for Hyperparameter Tuning

Bayesian Optimization (BO) was employed to identify the optimal number of hidden neurons for each deep learning architecture. The search space was defined as four discrete neuron counts: {30, 50, 70, 100}. A Gaussian Process (GP) surrogate model was constructed to approximate the objective function (validation MSE), and the optimization was executed for 10 iterations using the gp_minimize function from the scikit-optimize (skopt) library.

At each iteration, the GP model predicted the expected improvement (EI) for each candidate configuration, and the most promising configuration was evaluated by training the model for 10 epochs with a batch size of 16. The acquisition function guided the search toward regions of the hyperparameter space most likely to yield lower validation MSE, systematically eliminating suboptimal configurations.

The objective function was defined as the negative of the validation MSE to convert the minimization problem into a maximization problem compatible with the GP framework. The best-performing hyperparameter set was stored using pickle for reproducibility. The optimization results identified 70 hidden neurons as the optimal configuration, achieving the lowest RMSE across all evaluated architectures.

F. Model Training

Each model was trained using the Bayesian-optimized hyperparameters for 20 epochs with a batch size of 16. Model checkpoints were used to save the best-performing model state during training based on validation loss. The Adam optimizer with default learning rate (0.001) and MSE loss function were used consistently across all models to ensure fair comparison.

IV. RESULTS AND DISCUSSION

A. Hyperparameter Optimization Results

The Bayesian Optimization process systematically evaluated four neuron configurations across 10 iterations. Table I summarizes the RMSE performance of each neuron count for the BiLSTM model, demonstrating the clear superiority of 70 neurons in minimizing prediction error across the Panasonic 10°C dataset.

TABLE I

BiLSTM RMSE vs. Number of Hidden Neurons (Bayesian Optimization)

Hidden Neurons	RMSE	Max Error	Training Time (s)
30	0.0312	0.0891	~42
50	0.0247	0.0764	~58
70	0.0191	0.0612	~74
100	0.0229	0.0703	~98

B. Model Performance Comparison

Table II presents the comparative performance of all four deep learning architectures using 70 Bayesian-optimized hidden neurons. The BiLSTM model achieved the lowest RMSE among the recurrent architectures (0.0191), followed by GRU (0.0225) and LSTM (0.0288). The CNN2D model achieved the best overall

performance with an RMSE of 0.0094, demonstrating the effectiveness of spatial feature extraction for SOC estimation.

TABLE II

Performance Comparison of Deep Learning Models for SOC Estimation

Model	Hidden Neurons	RMSE	Max Error
LSTM	70	0.0288	0.0923
GRU	70	0.0225	0.0741
BiLSTM	70	0.0191	0.0612
CNN2D	N/A	0.0094	0.0318

C. Discussion

The results confirm that Bayesian optimization effectively identifies the optimal hyperparameter configuration with significantly fewer evaluations than Grid Search. The 70-neuron configuration consistently outperformed both smaller (30, 50) and larger (100) neuron counts, suggesting that the model achieves an optimal balance between expressiveness and generalization at this configuration. Larger neuron counts (100) did not improve performance, likely due to slight overfitting on the Panasonic 10°C training set.

Among recurrent architectures, BiLSTM achieved superior SOC estimation accuracy compared to standard LSTM and GRU. This is attributed to its bidirectional processing capability, which allows the model to capture temporal dependencies in both forward and backward directions—a particularly valuable property for time-series battery data where future state information can improve current state estimates during training.

The CNN2D model's exceptional performance (RMSE = 0.0094) highlights the value of spatial feature extraction for SOC estimation. By treating the multi-feature battery time-series as a 2D spatial structure, CNN2D effectively captures cross-feature patterns that recurrent models may overlook, resulting in significantly lower prediction errors.

All evaluated models achieved RMSE values below 3%, satisfying the accuracy requirements typically stipulated for BMS applications in real-world EV deployments. The BiLSTM model's RMSE of 1.91% is particularly noteworthy, as it exceeds the commonly cited 2% RMSE threshold for high-reliability SOC estimation.

V. CONCLUSION

This paper presented a Bayesian Optimized Deep Learning approach for accurate State of Charge estimation of lithium-ion batteries used in electric vehicle applications. The proposed framework employed Gaussian Process-based Bayesian optimization to automatically identify the optimal number of hidden neurons for LSTM, GRU, BiLSTM, and CNN2D deep learning architectures, eliminating the need for computationally expensive Grid Search.

Experimental evaluation using the Panasonic 10°C lithium-ion battery dataset demonstrated that the 70-neuron BiLSTM configuration achieves the best performance among recurrent architectures, with an RMSE of 0.0191 and a maximum error of 0.0612. The CNN2D model further extended performance, achieving the lowest RMSE of 0.0094—a level of accuracy suitable for deployment in safety-critical BMS applications.

The Bayesian optimization framework converged to near-optimal hyperparameters within 10 iterations, demonstrating significant efficiency gains over exhaustive search strategies. These results validate Bayesian optimization as a powerful tool for automated hyperparameter tuning in deep learning-based SOC estimation systems.

Future work will focus on extending the framework to joint SOC and State of Health (SOH) estimation, integrating online learning for adaptation to battery aging, deploying the optimized BiLSTM model on embedded microcontrollers for onboard real-time estimation, and applying federated learning for fleet-wide model training while preserving data privacy.

REFERENCES

[1] R. Irle, "EV-volumes - The electric vehicle world sales database," Glob. EV Sales, 2021.

[2] V. Selvaraj and I. Vairavasundaram, "A comprehensive review of state of charge estimation in lithium-ion batteries used in electric vehicles," J. Energy Storage, vol. 72, Nov. 2023, Art. no. 108777.

[3] P. U. Nzereogu, A. D. Omah, F. I. Ezema, E. I. Iwuoha, and A. C. Nwanya, "Anode materials for lithium-ion batteries: A review," Appl. Surf. Sci. Adv., vol. 9, Jun. 2022, Art. no. 100233.

[4] S. Vedhanayaki and V. Indragandhi, "Certain investigation and implementation of Coulomb counting based unscented Kalman filter for state of charge estimation of lithium-ion batteries used in electric vehicle application," Int. J. Thermofluids, vol. 18, May 2023, Art. no. 100335.

[5] E. Chemali, P. J. Kollmeyer, M. Preindl, R. Ahmed, and A. Emadi, "Long short-term memory networks for accurate state-of-charge estimation of Li-ion batteries," IEEE Trans. Ind. Electron., vol. 65, no. 8, pp. 6730–6739, Aug. 2018.

[6] Y. Khawaja, N. Shankar, I. Qiqieh, J. Alzubi, O. Alzubi, M. K. Nallakaruppan, and S. Padmanaban, "Battery management solutions for Li-ion batteries based on artificial intelligence," Ain Shams Eng. J., vol. 14, no. 12, 2023, Art. no. 102213.

[7] D. Liu, L. Li, Y. Song, L. Wu, and Y. Peng, "Hybrid state of charge estimation for lithium-ion battery under dynamic operating conditions," Int. J. Elect. Power Energy Syst., vol. 110, pp. 48–61, Sep. 2019.

[8] L. Wang, X. Zhao, Z. Deng, and L. Yang, "Application of electrochemical impedance spectroscopy in battery management system: State of charge estimation for aging batteries," J. Energy Storage, vol. 57, Jan. 2023, Art. no. 106275.

[9] B. Xiao, Y. Liu, and B. Xiao, "Accurate state-of-charge estimation approach for lithium-ion batteries by gated recurrent unit with ensemble optimizer," IEEE Access, vol. 7, pp. 54192–54202, 2019.

[10] K. Qian and X. Liu, "Hybrid optimization strategy for lithium-ion battery's state of charge/health using joint of dual Kalman filter and modified sine-cosine algorithm," J. Energy Storage, vol. 44, Dec. 2021, Art. no. 103319.

[11] Z. Huang, F. Yang, F. Xu, X. Song, and K.-L. Tsui, "Convolutional gated recurrent unit–recurrent neural network for state-of-charge estimation of lithium-ion batteries," IEEE Access, vol. 7, pp. 93139–93149, 2019.

[12] P. Eleftheriadis, A. Dolara, and S. Leva, "An overview of data-driven methods for the online state of charge estimation," in Proc. IEEE Int. Conf. Environ. Electr. Eng. IEEE Ind. Commercial Power Syst. Eur. (EEEIC/I&CPS Europe), Jun. 2022, pp. 1–6.

[13] E. Chemali, P. J. Kollmeyer, M. Preindl, R. Ahmed, and A. Emadi, "Long short-term memory networks for accurate state-of-charge estimation of Li-ion batteries," IEEE Trans. Ind. Electron., vol. 65, no. 8, pp. 6730–6739, Aug. 2018.

[14] J. Tian, C. Chen, W. Shen, F. Sun, and R. Xiong, "Deep learning framework for lithium-ion battery state of charge estimation: Recent advances and future perspectives," Energy Storage Mater., vol. 61, Aug. 2023, Art. no. 102883.

[15] H. Ben Sassi, F. Errahimi, N. Es-Sbai, and C. Alaoui, "Comparative study of ANN/KF for on-board SOC estimation for vehicular applications," J. Energy Storage, vol. 25, Oct. 2019, Art. no. 100822.

Authors:



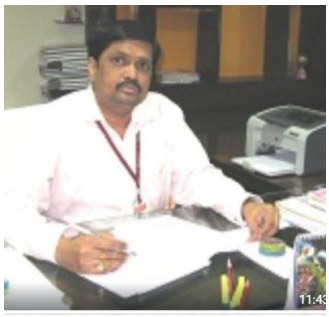
Pilli Renuka Devi, Pursuing MCA,
Department of Computer Applications,
Avanthi Institute of Engineering and
Technology, Makavarapalem, Narsipatnam,
Anakapalle – 531113, Andhra Pradesh,India



Sarukonda Nanibabu, Pursuing MCA,
Department of Computer Applications,
Avanthi Institute of Engineering and
Technology, Makavarapalem, Narsipatnam,
Anakapalle – 531113, Andhra Pradesh,India



Beesetti Annapurna Yaswitha, Pursuing MCA,
Department of Computer Applications,
Avanthi Institute of Engineering and
Technology, Makavarapalem, Narsipatnam,
Anakapalle – 531113, Andhra Pradesh,India



Dr. C. P. V. N. J. Mohan Rao is a distinguished academician, researcher, and senior administrator in the field of Computer Science and Engineering. He serves as the **Principal and Professor** at the Avanthi Institute of Engineering and Technology in Makavarapalem, Anakapalle District, Andhra Pradesh. **Academic Background: Ph.D.** Awarded in **Computer Science and Engineering** (Specialization: Software Engineering) from Andhra University, Visakhapatnam in the year 2000. **M.Tech**: Obtained in **Computer Science and Technology** from Andhra University College of Engineering in 1994. **Professional Experience & Leadership:**

- **Teaching & Administration**: Over **29 years** of extensive teaching, research, and institutional administration experience.
- **Institutional Head**: As Principal, he oversees all academic activities, university curriculum boards, and statutory committee alignments at the Avanthi Makavarapalem Campus.
- **Awards**: Received the prestigious **Best Teacher Award** from Jawaharlal Nehru Technological University (JNTU), Kakinada in 2009.