Hybrid 1D-CNN and CKF Model for Battery State of Charge Estimation in Electric Vehicles

1st Nacira BOUDRAA
Department of Electrical Engineering
University of Kasdi Merbah Ouargla
Ouargla, Algeria
boudraa.nacira@univ-ouargla.dz

2nd Hajer SALEM
AUDENSIEL
Boulogne-Billancourt, France
h.salem@audensiel.com

3rd Houari TOUBAKH
Department of Electrical Engineering
University of Kasdi Merbah Ouargla
Ouargla, Algeria
toubakh.houari@univ-ouargla.dz

4th Elhadi GASMI
Department of Electrical Engineering
University of Kasdi Merbah Ouargla
Ouargla, Algeria
Gasmi.elhadi@univ-ouargla.dz

5th Noureddine ZERHOUNI
FEMTO-ST Institute
UFC/Supmictrotech/UTBM, France
Besançon, France
zerhouni@ens2m.fr

6th Mohamed Redouane KAFI
Department of Electrical Engineering
University of Kasdi Merbah Ouargla
Ouargla, Algeria
kafi.redouane@univ-ouargla.dz

***Abstract*—Accurate State of Charge (SOC) estimation is essential for the safe and efficient operation of Electric Vehicles (EVs). State-of-the-art data-driven models, particularly Convolutional Neural Network (CNN)-based approaches, are capable of capturing the nonlinear dynamics of lithium-ion batteries. However, their performance may degrade under measurement noise and operating-condition-induced distribution shifts, which limits their robustness in real-world applications. This paper proposes a hybrid SOC estimation framework that combines a one-dimensional CNN with a Cubature Kalman Filter (CKF). The proposed approach first employs the One-Dimensional Convolutional Neural Network (1D-CNN) to extract temporal features from current, voltage, and temperature measurements and generate an initial SOC estimate. Subsequently, the CKF is integrated to refine the CNN predictions by filtering measurement noise and improving estimation stability and tracking capability. Experimental validation under multiple EV driving profiles demonstrates that the proposed hybrid framework achieves higher estimation accuracy, enhanced robustness, and improved reliability compared with the standalone CNN-based estimator. The results highlight the effectiveness of integrating deep learning and Bayesian filtering techniques for advanced battery management applications.**

***Index Terms*—Lithium-ion battery, SOC, 1D-CNN, CKF, EVs, Battery Management System (BMS).**

I. Introduction

Lithium-ion batteries are widely used in EVs due to their high energy and power densities, low self-discharge rate, and long service life. However, accurate SOC estimation remains challenging because of complex electrochemical behavior, strong nonlinearities, and varying operating conditions. As a key parameter in a BMS, precise SOC estimation is essential for safe operation, effective fault diagnosis, and the prevention of overcharging and deep discharging [1].

Existing SOC estimation methods can be broadly classified into conventional methods, model-based, data-driven, and hybrid approaches [2]. Conventional methods, such as coulomb counting and lookup tables, are computationally efficient but suffer from cumulative errors, sensitivity to initial conditions, and limited adaptability to battery aging and dynamic conditions [3]. Model-based methods combine mathematical battery models with observers or nonlinear filters, such as the Extended Kalman Filter (EKF) and Unscented Kalman Filter (UKF) [4], achieving good accuracy when model fidelity is high, but their performance degrades in the presence of modeling uncertainties and parameter variations. Data-driven approaches, including machine learning and deep learning techniques such as Support Vector Machine (SVM), CNN, and Long Short-Term Memory (LSTM) [5], learn nonlinear mappings directly from measurable battery signals to estimate SOC. These methods provide strong modeling flexibility and improved robustness to noise compared to conventional methods and model-based approaches. However, their performance is highly dependent on the availability of large, high-quality datasets and may suffer from limited generalization capability under unseen operating conditions or distribution shifts.

Hybrid SOC estimation frameworks have increasingly been explored to mitigate the limitations of individual methods. In such approaches, data-driven models are combined with model-based filtering techniques to exploit complementary strengths, thereby improving estimation accuracy, robustness, and adaptability under complex operating conditions. For example, in [6], hybrid Neural Network–Kalman Filter (NN–KF) methods for lithium-ion battery SOC estimation are employed, where initial SOC estimates are provided by Kalman filters using equivalent circuit models derived from voltage and current measurements, and neural networks subsequently refine these estimates by learning nonlinear error corrections. In [7], an EKF-CNN-Bidirectional Long Short-Term Memory (BiLSTM) framework is proposed, in which an EKF generates preliminary SOC estimates based on a Second-order RC model (2RC) model, followed by convolutional and bidirectional recurrent networks with an attention mechanism to capture spatial and temporal dependencies in the data. Similarly, in [8], an Adaptive Cubature Kalman Filter (ACKF) based on an improved generalized minimum error entropy criterion is introduced to estimate SOC, with adaptive noise covariance tuning to enhance robustness under non-Gaussian noise conditions.

Despite their demonstrated effectiveness, these hybrid SOC estimation methods remain constrained by the inherent challenges of battery modeling, including reliance on simplified equivalent circuit representations, sensitivity to model parameters, and limited capability to fully capture complex nonlinear electrochemical dynamics under diverse operating conditions. In this paper, a hybrid SOC estimation framework combining a 1D-CNN and a CKF is proposed. The 1D-CNN learns the nonlinear battery behavior directly from time-series measurements of current, voltage, and temperature, eliminating the need for equivalent circuit models. To improve robustness and reduce estimation noise, the CKF is integrated to filter and correct the 1D-CNN-generated SOC predictions. As a result, the proposed framework mitigates modeling uncertainties and parameter sensitivity while maintaining high estimation accuracy.

The remainder of this paper is organized as follows. Section II presents the SOC estimation problem formulation together with an overview of the adopted 1D-CNN architecture and CKF algorithm. Section III describes the datasets, experimental setup, and performance evaluation results. Finally, Section IV concludes the paper.

A. Main notation

The state of charge is denoted by SOC. The battery measurements include current I_k, voltage V_k, and temperature T_k. The system state is defined as $x_k = SOC_k$. The nonlinear dynamics are described by the state transition function $f_k(\cdot)$ and the measurement function $g(\cdot)$. Process and measurement noises are w_k and v_k, with covariance matrices Q_k and R_k, respectively. The maximum capacity is Q_{max}. The initial SOC estimate from the 1D-CNN is y_k, while the CKF-refined estimate is $\hat{x}_k$, with estimation error covariance P_k.

II. Problem Statement

Accurate SOC estimation in EVs batteries is critical for safety, longevity, and efficient energy management, yet remains challenging due to strong electrochemical nonlinearities and highly dynamic operating conditions. To address these limitations, hybrid approaches that combine deep learning with filtering techniques have emerged as a promising solution for robust and accurate SOC estimation.

A. 1D-CNN for SOC Estimation

CNN are fundamental to deep learning, and 1D-CNN layers are well suited to modelling temporal dependencies and local sequential structure in data such as time-series signals [9]. In this context, 1D-CNN can automatically extract hierarchical and local features from measurements like current, voltage, and temperature, effectively modelling temporal dependencies, short-term fluctuations, and long-term trends without manual feature engineering. This capability makes them especially effective for SOC estimation, where the relationship between measured signals and SOC is highly nonlinear and governed by complex electrochemical dynamics. The proposed 1D-CNN architecture consists of two successive one-dimensional convolutional layers with ReLU activations to extract temporal features from input sequences of length k. An adaptive average pooling layer then compresses the resulting feature maps, which are subsequently flattened and passed through fully connected layers with ReLU and sigmoid activations to produce the final SOC estimate. This architecture effectively captures hierarchical patterns and nonlinear relationships in sequential input data of length k for accurate SOC prediction. In the next step, the SOC predicted by the 1D-CNN is used as input to a CKF for smoothing and filtering , which further smooths the SOC estimates and improves the robustness of the SOC monitoring under measurement noise and transient fluctuations.

B. SOC Filtering Using a CKF

Filtering techniques have long been applied to state estimation from noisy data. Among these, the CKF is a nonlinear Bayesian filtering method that effectively smooths noisy predictions by optimally combining the model output with the system's underlying dynamics [10]. This approach mitigates the impact of measurement noise and transient fluctuations in battery signals, yielding a more stable and reliable SOC trajectory. To improve the stability of SOC predictions, a CKF is applied to the 1D-CNN model outputs. The CKF is a nonlinear Bayesian filter that reduces the effect of measurement noise and transient fluctuations [11]. Let y_k denote the raw SOC prediction from the trained 1D-CNN model at time step k, and let x_k denote the true state SOC. We define the SOC estimation problem as a state-space model as follows (Eq. (1)).

$$\begin{aligned} x_{k+1} &= f_k(x_k) + w_k \\ y_k &= g(x_k) + v_k \end{aligned} \tag{1}$$

where the $f(\dot{)}$ is the nonlinear state function given by Eq. (2)

$$f_k(x_k) = x_k - \frac{I_k \Delta t}{Q_{\text{max}}} \tag{2}$$

where I_k is the battery current, Q_{max} is the maximum battery capacity. The measurement function $g(x_k) = x_k$. The terms w_k and v_k represent the process noise and measurement noise, with covariance matrices Q_k and R_k, respectively.

To apply the CKF filtering, we assume the system has a Gaussian distribution; this assumption is commonly adopted in the literature to simplify Bayesian inference and filtering derivations, particularly in Kalman-based and Gaussian approximation methods [7], [8]. According to this assumption, the state vector x_k obeys a Gaussian distribution $\mathcal{N}(\hat{x}_k, P_k)$. Based on the third-degree spherical cubature integration rule, $2n$ sigma points are selected to approximate the probability distribution as follows (Eq. (3)):

$$X_{k-1}^{(i)} = \hat{x}_{k-1} + \sqrt{P_{k-1}}\,\psi^i, \quad i = 1, \dots, 2n. \tag{3}$$

Here, $\psi^{(i)}$ denote the unit sigma points defined as

$$\psi^i = \begin{cases} \sqrt{n}\, e_i, & i = 1, \dots, n, \\ -\sqrt{n}\, e_{i-n}, & i = n+1, \dots, 2n, \end{cases} \tag{4}$$

where e_i is the i-th unit vector, and n is number of state.

Using these sigma points, the one-step predicted state estimate $\hat{x}_{k|k-1}$ and the predicted error covariance matrix $P_{k|k-1}$ are computed as (Eq. (5))

$$\hat{x}_{k|k-1} = \frac{1}{2n}\sum_{i=1}^{2n} \hat{X}_k^{(i)},$$
$$P_{k|k-1} = \frac{1}{2n}\sum_{i=1}^{2n} \left(\hat{X}_k^{(i)} - \hat{x}_{k|k-1}\right)\left(\hat{X}_k^{(i)} - \hat{x}_{k|k-1}\right)^T + Q_{k-1}. \tag{5}$$

Then, the set of new sigma points obtained using the predicted mean and covariance in Eq. (5) and the measurement model $\hat{y}_k^{(i)}$. Thus, the predicted measurement mean and covariance and cross covariance are given as (Eq. (6))

$$\hat{y}_{k|k-1} = \frac{1}{2n}\sum_{i=1}^{2n} \hat{y}_k^{(i)},$$
$$S_{k|k-1} = \frac{1}{2n}\sum_{i=1}^{2n} \left(\hat{y}_k^{(i)} - \hat{y}_{k|k-1}\right)\left(\hat{y}_k^{(i)} - \hat{y}_{k|k-1}\right)^T + R_k,$$
$$P_{k|k-1}^{xy} = \frac{1}{2n}\sum_{i=1}^{2n} \left(\hat{X}_k^{(i)} - \hat{x}_{k|k-1}\right)\left(\hat{y}_k^{(i)} - \hat{y}_{k|k-1}\right)^T. \tag{6}$$

by computing the predicted mean and covariance Eq. (5) based on the spherical cubature integration rule in Eq. (4), the update step of the SOC estimation is similar to the standard Kalman filter (Eq.(7))

$$\begin{aligned} K_k &= P_{x_k y_k} S_{k|k-1}^{-1}, \\ \hat{x}_k &= \hat{x}_{k|k-1} + K_k\left(y_k - \hat{y}_{k|k-1}\right), \\ P_k &= P_{k|k-1} - K_k S_{k|k-1} K_k^T \end{aligned} \tag{7}$$

III. Experimental Evaluation and Results

This section presents the experimental protocol and the obtained results used to assess the proposed SOC estimation framework. The evaluation is structured into three parts: (i) the dataset and driving profiles employed for training and testing, (ii) the preprocessing pipeline and evaluation metrics, and (iii) a quantitative and qualitative analysis of SOC estimation performance under different temperatures and operating conditions.

A. Dataset Description

Accurate SOC estimation requires datasets that capture the diverse and dynamic operating conditions of EVs. To ensure robustness and generalization, this study employs controlled laboratory data obtained from $LiFePO_4$ batteries (1.1 Ah) for training and realistic driving data for testing, from the dataset reported in [12]. For training, Dynamic Stress Test (DST) profiles were recorded in a temperature chamber, where SOC, voltage, current, and temperature were measured at discrete temperatures of $-10°C$, $0°C$, $10°C$, $25°C$, $30°C$, and $50°C$. These DST profiles include various current steps and regenerative charging events, providing comprehensive coverage of SOC, voltage, and current dynamics over a wide thermal range. For testing, the trained model was evaluated under realistic driving conditions using the Supplemental Federal Test Procedure Driving Schedule (US06) and the Federal Urban Driving Schedule (FUDS), conducted at the same temperature conditions. These profiles exhibit rapid and complex variations in current and SOC, enabling an assessment of the model performance under near real-world EVs operating conditions.

B. Preprocessing and Metrics

Before training the 1D-CNN for SOC estimation, the battery data were preprocessed by dividing the time series into overlapping sequences of length $K = 20$. For each sequence starting at index i, the input is a matrix containing K consecutive samples of current (I), voltage (V), and temperature (T) (Eq.(8)):

$$X_i = \begin{bmatrix} I_i & V_i & T_i \\ I_{i+1} & V_{i+1} & T_{i+1} \\ \vdots & \vdots & \vdots \\ I_{i+K-1} & V_{i+K-1} & T_{i+K-1} \end{bmatrix}, \quad y_i = \mathrm{SOC}_{i+K}. \tag{8}$$

where the target y_i is the SOC at the last step of the sequence.

Each input feature $x \in \{I, V, T\}$ was normalized to the range $[-1, 1]$ using Min-Max scaling. The scaling parameters $(x_{\min}, x_{\max})$ were computed on the training set only and then applied unchanged to the test set, as defined in Eq. (9).

$$x_{\text{norm}} = 2 \cdot \frac{x - x_{\min}}{x_{\max} - x_{\min}} - 1 \tag{9}$$

where $x_{\min}$ and $x_{\max}$ denote the minimum and maximum values of the input in the dataset.

The performance of SOC estimation was evaluated using the Root Mean Square Error (RMSE) (Eq. (10)) and Mean Absolute Error (MAE) (Eq.(11)) as follows:

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\mathrm{SOC}_i - \widehat{\mathrm{SOC}}_i\right)^2}, \tag{10}$$

$$\mathrm{MAE} = \frac{1}{N}\sum_{i=1}^{N}\left|\mathrm{SOC}_i - \widehat{\mathrm{SOC}}_i\right|. \tag{11}$$

where SOC_i is the true SOC at time step i, $\widehat{\mathrm{SOC}}_i$ is the predicted SOC, and N is the total number of predictions. To evaluate the benefit of the hybrid approach, the metrics were calculated first using only the 1D-CNN predictions and then using the 1D-CNN predictions refined by the CKF (1D-CNN+CKF), allowing a quantitative comparison of SOC estimation accuracy with and without filtering enhancement.

C. Experiments Evaluation and Obtained Results

a) Objective of the experiments: The objective of this study is to evaluate the accuracy and robustness of the proposed hybrid 1D-CNN+CKF method for SOC estimation under realistic operating conditions. The experiments aim to assess the estimation performance under different driving profiles and temperatures, and to quantify the benefit of combining data-driven prediction with probabilistic filtering.

b) Training protocol and model parameters: The proposed 1D-CNN model was trained for 100 epochs with a batch size of 64 using the Adam optimizer and the mean squared error (MSE) as the loss function. The MSE is defined as:

$$\text{MSE} = \frac{1}{n}\sum_{i=1}^{n}\left(SOC_i^{\text{true}} - SOC_i^{\text{pred}}\right)^2, \tag{12}$$

where SOC_i^{true} and SOC_i^{pred} denote the true and predicted SOC at time step i, and n is the number of samples.

To improve estimation stability and reduce prediction noise, the 1D-CNN outputs were refined using a CKF, as described in Section II-B. The model was trained until convergence on the training dataset and evaluated on unseen data corresponding to the US06 and FUDS driving profiles. The reference SOC used for evaluation is obtained using the coulomb counting method, which estimates the battery charge level by tracking the amount of electric charge removed over time. In this approach, the battery is assumed to be fully charged at the initial time instant, i.e., $SOC(t_0) = 100\%$, and its charge decreases as current is drawn during discharge. The SOC at any time t is then computed as Eq.(13):

$$SOC(t) = SOC(t_0) - \frac{1}{Q_{\text{nom}}}\int_{t_0}^{t} I(\tau)\, d\tau, \tag{13}$$

where Q_{nom} denotes the nominal battery capacity and $I(\tau)$ is the measured current. The integral term represents the total charge extracted from the battery, while Q_{nom} normalizes it into a dimensionless SOC value. This method provides a reference SOC trajectory, which is used as ground truth to evaluate the performance of the proposed estimation method. Finally, estimation accuracy is assessed using the RMSE and MAE, which quantify the deviation between the estimated and reference SOC.

c) Results and discussion: The proposed hybrid approach achieved an RMSE of less than 4% and an MAE of less than 3% for both the US06 and FUDS datasets, indicating accurate SOC estimation under dynamic operating conditions.

Tables I and II show the influence of temperature and driving profiles. In all tested conditions, the hybrid 1D-CNN+CKF consistently outperforms the standalone 1D-CNN. For the US06 cycle, RMSE decreases from 5.50–6.09% to 2.46–3.21%, while MAE decreases from 4.08–4.69% to 1.84–2.37%. For the FUDS cycle, RMSE decreases from 4.90–5.77% to 1.94–2.72%, and MAE decreases from 3.54–4.14% to 1.42–2.04%. On average, the CKF refinement reduces the estimation error by approximately 50%.

Figures 1–4 corroborate the quantitative improvements reported in Tables I and II by illustrating the effect of CKF refinement on SOC trajectories. The plot shows the battery SOC (%) over discrete time steps. The red solid line represents the true SOC from Coulomb counting, smoothly decreasing during discharge. The black dashed line shows the standalone 1D-CNN estimate, exhibiting high-frequency fluctuations and transient deviations, particularly during rapid current variations such as acceleration and regenerative braking. The blue solid line depicts the hybrid 1D-CNN+CKF estimate, which closely follows the true SOC, attenuating noisy spikes and maintaining a smoother, physically consistent trajectory. This improvement arises from the CKF's Bayesian fusion mechanism, which treats the 1D-CNN output as a noisy measurement, reducing measurement noise and transient errors. The hybrid approach enhances stability and tracking accuracy under highly dynamic conditions and across temperatures, although relative gains decrease under thermal extremes (e.g., 50°C on US06) due to stronger parameter drift and nonlinearities not fully captured by a fixed filter. These findings motivate further work on temperature-aware modeling (e.g., $Q_{\text{nom}}(T)$ and polarization effects) and adaptive CKF tuning with temperature and SOC dependent noise covariances to preserve accuracy and reliability under extreme conditions.

TABLE I: SOC estimation performance of 1D-CNN and 1D-CNN+CKF models for the US06 driving profile at different temperatures.

Temp (°C)	**1D-CNN**		**1D-CNN+CKF**	
	RMSE (%)	**MAE (%)**	**RMSE (%)**	**MAE (%)**
25	5.51	4.26	2.46	1.84
30	6.09	4.69	2.63	2.02
50	5.50	4.08	3.21	2.37

TABLE II: SOC estimation performance of 1D-CNN and 1D-CNN+CKF models for the FUDS driving profile at different temperatures.

Temp (°C)	**1D-CNN**		**1D-CNN + CKF**	
	RMSE (%)	**MAE (%)**	**RMSE (%)**	**MAE (%)**
25	4.90	3.54	1.94	1.42
30	5.77	4.14	2.68	2.00
50	5.23	3.68	2.72	2.04

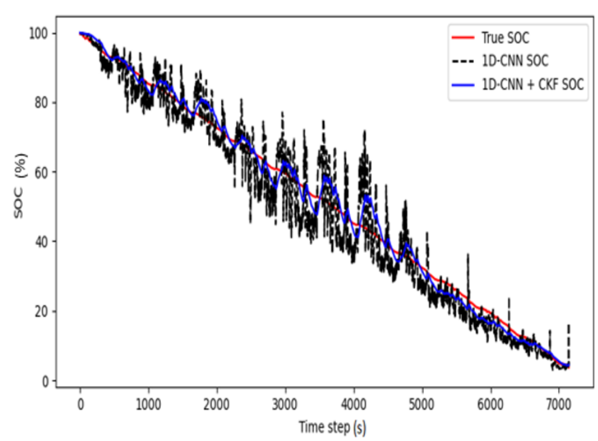

Fig. 1: SOC estimation at 30°C for the US06 profile.

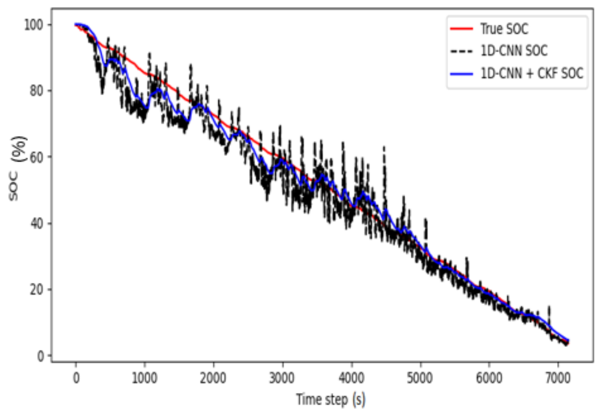

Fig. 2: SOC estimation at 50°C for the US06 profile.

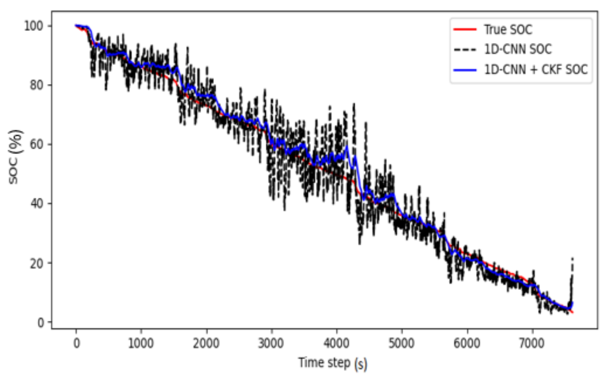

Fig. 3: SOC estimation at 30°C for the FUDS profile.

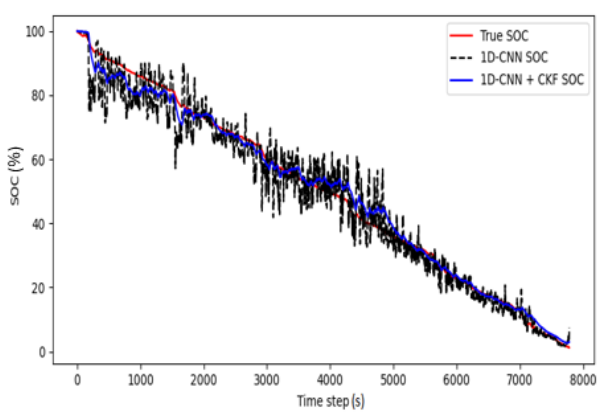

Fig. 4: SOC estimation at 50°C for the FUDS profile.

IV. Conclusion

This paper proposed a hybrid 1D-CNN+CKF framework for SOC estimation by combining data-driven learning with Bayesian filtering. The 1D-CNN captures nonlinear battery dynamics from time-series data, while the CKF improves stability by reducing noise and transient errors. Experimental results on US06 and FUDS driving cycles under various temperatures show that the proposed method consistently outperforms the standalone 1D-CNN, with around 50% reduction in RMSE and MAE, confirming its effectiveness and robustness in dynamic conditions. Despite these advantages, the approach is limited by its dependence on training data quality, fixed CKF covariance settings that do not fully capture temperature and aging effects, and the lack of real-time implementation validation. Future work will address these issues through adaptive covariance tuning, temperature-aware modeling, and embedded system deployment.

References

[1] Bhattacharjee, A., Verma, A., Mishra, S., Saha, T. K. (2021). Estimating state of charge for xEV batteries using 1D convolutional neural networks and transfer learning. IEEE Transactions on Vehicular Technology, 70(4), 3123-3135.

[2] Rouabah, B., Toubakh, H., Djemai, M., Ben-Brahim, L., Ghandour, R. (2023). Fault diagnosis based machine learning and fault tolerant control of multicellular converter used in photovoltaic water pumping system. IEEE Access, 11, 39013-39023.

[3] Ezukwoke, K., Toubakh, H., Hoayek, A., Batton-Hubert, M., Boucher, X., Gounet, P. (2021, August). Intelligent fault analysis decision flow in semiconductor industry 4.0 using natural language processing with deep clustering. In 2021 IEEE 17th international conference on automation science and engineering (CASE) (pp. 429-436). IEEE.

[4] He, H., Xiong, R., Zhang, X., Sun, F., Fan, J. (2011). State-of-charge estimation of the lithium-ion battery using an adaptive extended Kalman filter based on an improved Thevenin model. IEEE Transactions on vehicular technology, 60(4), 1461-1469.

[5] Chai, X., Li, S., Liang, F. (2024). A novel battery SOC estimation method based on random search optimized LSTM neural network. Energy, 306, 132583.

[6] Cui, Z., Dai, J., Sun, J., Li, D., Wang, L., Wang, K. (2022). Hybrid methods using neural network and Kalman filter for the state of charge estimation of lithium-ion battery. Mathematical Problems in Engineering, 2022(1), 9616124.

[7] Zhu, M., Wang, Y., Wang, D. (2025). SOC Prediction of Li-Ion Battery Based on EKF and CNN-BiLSTM-Attention. ACS omega, 10(47), 57157-57168.

[8] Z. Shi, J. Xu, M. Wu, L. Zeng, H. Zhang, Y. He, C. Liu. (2023). An improved adaptive square root cubature Kalman filter method for estimating state-of-charge of lithium-ion batteries. Journal of Energy Storage, 72, 108245.

[9] Mahboub, M. A., Rouabah, B., Kafi, M. R., Toubakh, H. (2022). Health management using fault detection and fault tolerant control of multicellular converter applied in more electric aircraft system. Diagnostyka, 23(2).

[10] Arasaratnam I, Haykin S. Cubature Kalman filters. IEEE Trans Autom Control 2009; 54 (6):1254–69.

[11] Boudraa, N., Salem, H., Toubakh, H., Henna, H., Kafi, M. R. (2024, May). Optimized Fault Diagnosis-Driven RUL Prediction for Lithium-Ion Batteries using Multivariate LSTM. In 2024 International Conference on Control, Automation and Diagnosis (ICCAD) (pp. 1-6). IEEE.

[12] He, W., Williard, N., Chen, C., Pecht, M. (2014). State of charge estimation for Li-ion batteries using neural network modeling and unscented Kalman filter-based error cancellation. International Journal of Electrical Power Energy Systems, 62, 783-791.