



HAL
open science

Deep Learning-Based Fault Diagnosis in Photovoltaic System

Mouaad Bougoffa, Samir Benmoussa, Mohand Djeziri, Thierry Contaret

► **To cite this version:**

Mouaad Bougoffa, Samir Benmoussa, Mohand Djeziri, Thierry Contaret. Deep Learning-Based Fault Diagnosis in Photovoltaic System. 2024 International Conference on Decision Aid Sciences and Applications (DASA), Dec 2024, Manama, France. pp.1-5, 10.1109/DASA63652.2024.10836268 . hal-04901924

HAL Id: hal-04901924

<https://amu.hal.science/hal-04901924v1>

Submitted on 20 Jan 2025

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Deep Learning-Based Fault Diagnosis in Photovoltaic System

Mouaad BOUGOFFA
LASA Laboratory, Badji
Mokhtar-Annaba University
P.O.BOX 12, Annaba,23000,
Algeria
mouaad.bougoffa@univ-
annaba.dz

Samir BENMOUSSA
LASA Laboratory, Badji
Mokhtar-Annaba University
P.O.BOX 12, Annaba,23000,
Algeria
Samir.benmoussa@univ-
annaba.dz

Mohand DJEZIRI
Aix Marseille Univ, Université de
Toulon, CNRS, IM2NP,
Marseille, France
mohand.djeziri@im2np.fr

Thierry CONTARET
Aix Marseille Univ, Université de
Toulon, CNRS, IM2NP
Marseille, France
Thierry.contaret@im2np.fr

Abstract— This study presents a methodology for diagnosing faults in photovoltaic (PV) systems using Convolutional Neural Networks (CNNs). By leveraging CNNs' ability to extract intricate features from PV data, including voltage, current, irradiation, and temperature, the proposed approach enables the accurate identification of various fault types such as short circuits, open circuits, wear, and shading. The contribution of this paper enhances PV system performance and reliability through early fault identification and a comprehensive framework for fault analysis. By effectively distinguishing between normal operating conditions and different fault occurrences, this methodology supports a more sustainable and efficient solar energy future. Through data preprocessing, optimization of the CNN architecture, and incorporation of a multi-class classification layer, the model achieved a high accuracy in fault detection, with an average detection and classification accuracy of 99.7%. The rigorous validation and testing employed in this study affirm the effectiveness and reliability of the CNN-based approach. Overall, this methodology offers a promising strategy for improving fault diagnosis in PV systems, advancing the reliability and efficiency of solar energy applications.

Keywords— Photovoltaic (PV), Diagnosis, Fault Detection, AI techniques, CNN.

I. INTRODUCTION

In recent years, the deployment of photovoltaic (PV) systems has increased globally as a clean and sustainable energy solution. However, ensuring the reliable operation of these systems is crucial for optimizing energy production and reducing maintenance expenses. Faults and failures in PV systems can significantly impact their performance, leading to reduced efficiency, increased downtime, and potential safety risks. Therefore, effective fault detection and classification methods are essential for maintaining the optimal functioning of PV installations [1].

Traditional fault detection methods often rely on manual inspection or rule-based algorithms, which can be time-consuming, labor-intensive, and susceptible to inaccuracies. With the advancements in deep learning, particularly Convolutional Neural Networks (CNNs), there is an opportunity

to revolutionize fault classification in PV systems. CNNs have demonstrated remarkable capabilities in recognition tasks, making them well-suited for analyzing array data generated by PV modules [2].

In recent years, the detection and categorizing faults in PV systems have been explored using various machine learning algorithms other than CNNs. Studies have investigated the efficacy of Support Vector Machines (SVM), Artificial Neural Networks (ANN), Multi-Layer Perceptrons (MLP), and Long Short-Term Memory (LSTM) networks in addressing the challenges of fault identification and classification in PV systems [3,4].

Support Vector Machines have been widely applied in fault detection tasks within PV systems. For example, in [5] the authors present SVM-driven fault diagnosis algorithm that utilizes SVM to identify short circuit, open circuit, and absence lack of irradiation faults, by analyzing the I–V characteristic curves of PV arrays.

Multi-Layer Perceptrons and Artificial Neural Networks have also been employed in fault classification tasks in PV systems [6]. In [7], the authors utilized MLP for fault detection by extracting features from sensor data, demonstrating its ability to accurately classify faults in PV systems. Similarly, in [8], Khelil et al. applied radial basis function (RBF) neural networks to identify and localize the most prevalent faults in photovoltaic (PV) generators, including short-circuit and open string faults. In the experimental trials, all fault categories were accurately classified.

Long Short-Term Memory networks, known for their ability to capture temporal dependencies in sequential data, have shown promise in fault detection and classification tasks in PV systems. For instance, in [9] Appiah et al. employed LSTM networks to analyze time-series data from PV modules for fault detection, demonstrating its effectiveness in identifying anomalies and predicting impending faults.

Study [10] introduced a model that utilized stacked autoencoders to detect faults in photovoltaic modules. This approach was validated using simulated data from PV systems.

Additionally, another study in [11] integrated nonlinear autoregressive exogenous (NARX) network and linguistic fuzzy rule-based systems to identify and categorize faults in PV systems. This method utilized Sugeno fuzzy inference and incorporated real-time sensor data to predict and observe the power generated by the PV array and the ambient environmental conditions.

In this paper, we propose a novel deep learning approach for fault classification in photovoltaic (PV) arrays, leveraging the power of Convolutional Neural Networks (CNNs) to accurately detect and categorize a variety of fault types. Using a robust dataset derived from real PV array operations, our method effectively captures and analyzes complex interactions between system parameters and fault conditions.

The proposed methodology comprises several critical phases: data preprocessing, CNN architecture design, model training, validation, and testing. Our carefully curated dataset, rich with operational parameters and corresponding fault labels, is used to train the CNN model to identify intricate patterns and relationships within the data. By learning these relationships, the CNN can precisely classify faults, even in previously unseen scenarios [12]. During training, the model iteratively optimizes its feature extraction and classification capabilities, progressively improving its accuracy in fault detection across diverse test samples.

The contributions of this work extend beyond merely developing a fault classification model for PV arrays. By showcasing the efficacy of CNN-based deep learning techniques in PV diagnostics, this research highlights the potential for integrating sophisticated ML tools into renewable energy systems. Automated fault detection systems powered by CNNs can substantially enhance the reliability, operational efficiency, and maintenance strategies of PV installations, ultimately supporting the broader adoption of solar energy as a sustainable and resilient power source.

The organization of this paper is outlined as follows: the subsequent section provides detailed insights into the methodology employed to develop this approach, Section 3 outlines the findings and initiates discussion, and lastly, the concluding section offers a summary of the paper.

II. METHODOLOGY

This section outlines our methodology for fault detection and classification in PV systems, utilizing Convolutional Neural Networks (CNN). The proposed approach begins with data preprocessing to enhance dataset quality. Then, CNNs are used, they allow feature extraction, and optimizing model parameters and architecture, ensuring thus accurate fault classification. Fig. 1 depicts the flowchart of the proposed methodology.

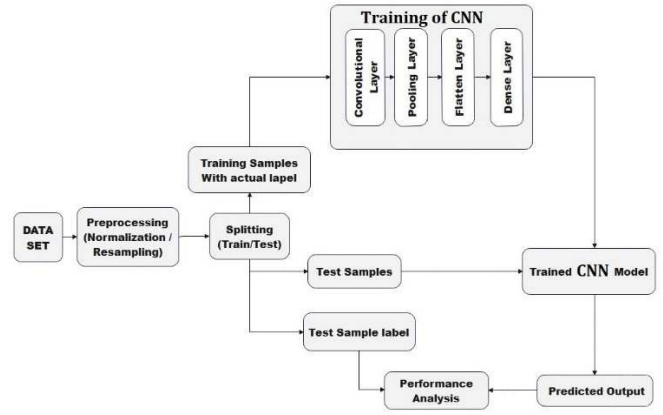


Fig.1. Proposed architecture for fault diagnosis within the system.

A. Photovoltaic system

The data used in this research are obtained from [13], originating from a real PV system that intentionally includes four different types of faults. Importantly, this dataset comes from an actual setup capable of accurately simulating faults, as illustrated in Fig. 2. The PV setup comprises two strings, each consisting of eight C6SU-330P PV modules labeled as PV₁ through PV₁₆. The output of each string is connected to a 5 kW grid-connected inverter (NHS Solar 5K-GDM1). Additionally, an electrical panel encompasses circuit protection components such as main circuit breakers (DM₁), circuit protection block, circuit breakers (DS_{1,2}), and fuses.

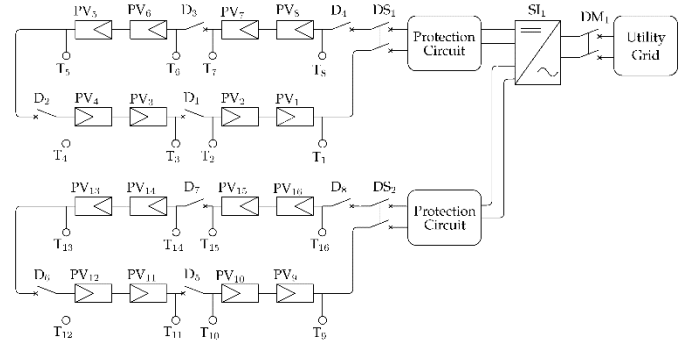


Fig. 2. Schematic representation of the solar PV system.

B. Dataset Description

The dataset was created by Lazzaretti et al. in 2020 and encompasses various parameters including voltage, current, temperature and the level of irradiation. It comprises data for five different operational conditions: open circuit (OC), short circuit (SC), partial shading (PS), degradation, and normal conditions. For our research, we specifically selected five photovoltaic features denoted as $x = [x_1, x_2, x_3, x_4, x_5]$ representing $[V_{dc1}, V_{dc2}, I_{dc1}, I_{dc2}, f_{nv}]$ respectively to be used as input for training our model Table I.

TABLE I. ELECTRICAL DATASET

Variables	Description
V_{dc1}	Voltage of String1
V_{dc2}	Voltage of String2
I_{dc1}	Current of String1
I_{dc2}	Current of String2
F_{nv}	Label Fault

C. Convolutional Neural Network algorithm

This section provides an extensive explanation of the proposed method for detecting and classifying faults. Researchers have introduced various methods for normalizing data, as well as for training, validating, and testing neural networks [8,14]. The flowchart given in Fig. 3 showcases a comprehensive illustration of the development convolutional neural networks algorithm.

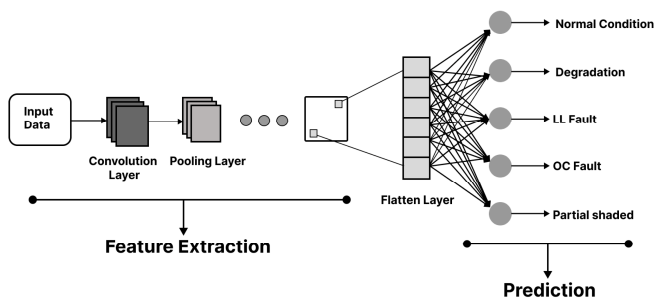


Fig. 3. The flowchart of the approach.

The CNN architecture designed for fault classification in PV panels undergoes preprocessing steps involving normalization and resampling before progressing through a convolution filter [15]. This filter efficiently extracts relevant features, generating a feature map. The convolution and pooling layers play an essential part in reducing the spatial dimension of the representation post convolution, effectively minimizing computational load and parameter count. Subsequently, the output layer, consisting of five nodes, is engaged for classification. These nodes correspond to distinct fault occurrences: normal condition (NC), degradation, short circuit faults, open circuit faults, and partial shading (PS), offering a comprehensive classification framework for fault analysis.

III. RESULTS AND DISCUSSION

This section presents a comparative analysis is conducted based on standard Machine Learning (ML) evaluation metrics, including precision, recall (sensitivity), specificity, and accuracy [16]. These metrics provide a comprehensive understanding of the algorithm performance in classifying faults across multiple classes, allowing for a quantitative analysis of its effectiveness.

The dataset was randomly divided into two portions: a training set, which consisted 70% of the data, and a testing set, which contained 30%. The segmentation facilitates a comprehensive evaluation of the algorithm's performance on novel data, typically represented by the testing dataset.

The performance of the algorithm was evaluated based on its ability to classify faults in solar photovoltaic systems. To assess this, the fault classifications generated by the algorithm were analyzed using a confusion matrix, which is depicted in Fig. 4.

Prior to testing, the algorithm was trained using labeled data to identify patterns in both electrical and environmental data collected from the PV systems. These training values enabled the algorithm to learn and assign fault labels to the measured values. The dataset, derived from the mention PV system, provides data points that distinguish between healthy and faulty operation.

Once trained, the algorithm was subjected to new test data to predict and categorize the corresponding fault types in the PV system. As our focus lies on classification algorithm, the evaluation metrics are associated with the accuracy of classification, distinguishing between true positives (TP) when the detection process correctly identifies a fault, true negatives (TN) when the detection process correctly identifies no fault, false positives (FP) when the detection process identifies a fault, and the system show no faults, and false negatives (FN) when the detection process show no fault, and the system experience a fault [17].

In the context of binary classification (where there are two classes: positive and negative), the recall (sensitivity) refers to the ability of a model to correctly identify all positive instances, on the other hand, specificity, is a measure of a model's ability to correctly identify all negative instances [18].

We then define accuracy as the measure of the overall correctness of the model's predictions. Unlike recall and specificity, which focus on specific aspects of the model's performance related to TP and TN, accuracy considers the overall correctness of the model across all classes [19].

Table II demonstrates the testing results based on these evaluation metrics.

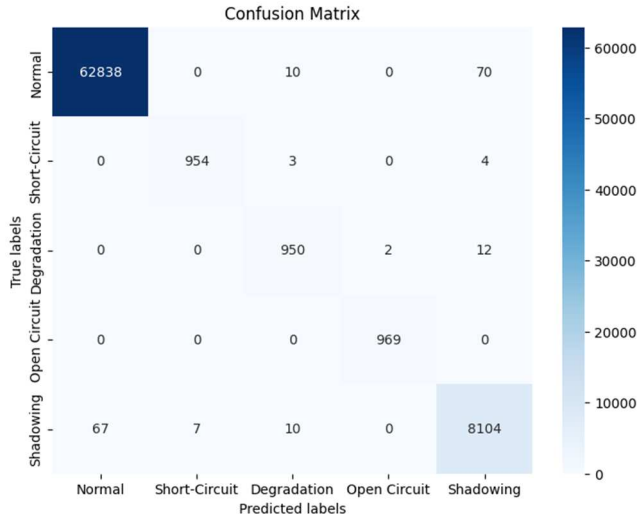


Fig. 4. The confusion matrix.

The confusion matrix summarizes the classification performance of a model across different classes. In this case, the model accurately classified all instances of the "normal" and "open circuit" classes. However, there were misclassifications observed for the "short circuit" and "degradation" classes, with some instances being incorrectly classified as "shadowing." These results highlight the model's strengths in certain classes but also indicate areas for improvement, particularly in distinguishing between similar fault types.

TABLE II. TESTING RESULTS

	Test 1	Test 2	Test 3	Average
Precision	99.6%	99.8%	99.6%	99.6%
Sensitivity	99.2%	99.4%	99.4%	99.3%
Specificity	100%	100%	100%	100%
Accuracy	99.6%	99.7%	99.8%	99.7%

Table II provides the testing results of the model across three different tests, along with their average values. The model demonstrates consistently high performance across all evaluation metrics, including precision, sensitivity, specificity, and accuracy. Precision ranges from 99.2% to 99.8%, sensitivity from 99.2% to 99.4%, specificity achieved 100%, and accuracy from 99.6% to 99.8%. These results indicate the model's robustness and reliability in fault detection within photovoltaic systems.

D. Comparative study

This section conducts a comparative study between our approach and recent models in the field of fault detection within photovoltaic systems using machine learning techniques such as SAE (Stacked Autoencoder), ANN, and MLP, as detailed in Table III. To evaluate how effective our suggested research is compared to previous studies, we compiled the table, outlining the identified fault types, the employed algorithms, and the accuracy of each method. Our method exhibits significantly superior testing accuracy when compared to these algorithms.

TABLE III. RESULTS OF THE COMPARATIVE STUDY

Classifier	Faults	Accuracy
SAE & clustering [20]	SC, PS, degradation	97.12 %
ANN [13]	OC, SC, PS, degradation	92.64 %
MLP [14]	SC, OC	97 %
Our method	OC, SC, PS, degradation	99.7 %

Table III compares the performance of different classifiers in fault detection within PV systems. SAE & clustering achieved 97.12% accuracy in identifying short circuits (SC), partial shading (PS), and degradation faults, while ANN achieved 92.64% accuracy in detecting open circuits (OC), SC, PS, and degradation faults. The accuracy of MLP is 97%. Notably, our method demonstrated the highest accuracy among all classifiers, at 99.7%, in identifying OC, SC, PS, and degradation faults. These results highlight the potential of our approach to significantly enhance fault detection accuracy in PV systems, thereby improving their reliability and efficiency.

IV. CONCLUSION

In conclusion, our methodology for detection faults and classification in PV systems, utilizing Convolutional Neural Networks (CNN), presents a systematic and effective approach. Beginning with meticulous data preprocessing to ensure dataset quality, we proceeded to utilize CNNs for feature extraction, optimizing model parameters and architecture to ensure precise fault classification. Rigorous validation and testing confirmed the reliability and effectiveness of our CNN-based approach. The utilization of real-world data from an actual PV system, intentionally incorporating various fault types, ensures the practicality and relevance of our research. Our accomplishments include enhanced performance and reliability, thus aiding in the progression of fault detection methods within the renewable energy domain. Our method presents a robust framework for improving both the performance and reliability of PV systems. To further enhance

the proposed methodology for diagnosis faults within PV systems, we will focus on the integration of additional data sources, such as weather conditions or panel orientation, to provide more comprehensive insights into fault occurrences and improve the model's predictive capabilities. Additionally, conducting real-world field tests and validations to assess the scalability and applicability of the proposed methodology in diverse operational environments.

ACKNOWLEDGMENT

This work was partially supported by the “PHC TASSILI +” programme (project number: 24MDU117), funded by the French Ministry for Europe and Foreign Affairs, the French Ministry for Higher Education and Research and the Algerian Ministry for Higher Education and Scientific Research (MESRS).

We thank the laboratory of automatic and signals at Annaba (LASA) for its interest and support to carry out this work.

REFERENCES

- [1] M. Mansouri, M. Trabelsi, H. Nounou, and M. Nounou, “Deep Learning-Based Fault Diagnosis of Photovoltaic Systems: A Comprehensive Review and Enhancement Prospects,” *IEEE Access*, vol. 9. Institute of Electrical and Electronics Engineers (IEEE), pp. 126286–126306, 2021. doi: 10.1109/access.2021.3110947.
- [2] D. Ruan, J. Wang, J. Yan, and C. Gühmann, “CNN parameter design based on fault signal analysis and its application in bearing fault diagnosis,” *Advanced Engineering Informatics*, vol. 55. Elsevier BV, p. 101877, Jan. 2023. doi: 10.1016/j.aei.2023.101877.
- [3] A. Abubakar, C. F. M. Almeida, and M. Gemignani, “Review of Artificial Intelligence-Based Failure Detection and Diagnosis Methods for Solar Photovoltaic Systems,” *Machines*, vol. 9, no. 12. MDPI AG, p. 328, Dec. 01, 2021. doi: 10.3390/machines9120328.
- [4] Z. Yuan, G. Xiong, and X. Fu, “Artificial Neural Network for Fault Diagnosis of Solar Photovoltaic Systems: A Survey,” *Energies*, vol. 15, no. 22. MDPI AG, p. 8693, Nov. 19, 2022. doi: 10.3390/en15228693.
- [5] J. Wang, D. Gao, S. Zhu, S. Wang, and H. Liu, “Fault diagnosis method of photovoltaic array based on support vector machine,” *Energy Sources, Part A: Recovery, Utilization, and Environmental Effects*, vol. 45, no. 2. Informa UK Limited, pp. 5380–5395, Sep. 29, 2019. doi: 10.1080/15567036.2019.1671557.
- [6] A. G. Olabi et al., “Artificial neural networks applications in partially shaded PV systems,” *Thermal Science and Engineering Progress*, vol. 37. Elsevier BV, p. 101612, Jan. 2023. doi: 10.1016/j.tsep.2022.101612.
- [7] B. Mouaad, B. Samir, and D. Mohand, “Dynamic Modeling for Fault Diagnosis in PV Systems Utilizing AI Techniques Based on Multilayer Perceptron (MLP),” 2024 ASU International Conference in Emerging Technologies for Sustainability and Intelligent Systems (ICETSIS). IEEE, Jan. 28, 2024. doi: 10.1109/icetsis61505.2024.10459490.
- [8] C. Kara Mostefa Khelil, B. Amrouche, K. Kara, and A. Chouder, “The impact of the ANN’s choice on PV systems diagnosis quality,” *Energy Conversion and Management*, vol. 240. Elsevier BV, p. 114278, Jul. 2021. doi: 10.1016/j.enconman.2021.114278.
- [9] A. Y. Appiah, X. Zhang, B. B. K. Ayawli, and F. Kyeremeh, “Long Short-Term Memory Networks Based Automatic Feature Extraction for Photovoltaic Array Fault Diagnosis,” *IEEE Access*, vol. 7. Institute of Electrical and Electronics Engineers (IEEE), pp. 30089–30101, 2019. doi: 10.1109/access.2019.2902949.
- [10] M. Bougoffa, S. Benmoussa, M. Djeziri, and T. Contaret, “Stacked AutoEncoder based diagnosis applied on a Solar Photovoltaic System,” *IFAC-PapersOnLine*, vol. 58, no. 4. Elsevier BV, pp. 384–389, 2024. doi: 10.1016/j.ifacol.2024.07.248.
- [11] S. Samara and E. Natsheh, “Intelligent PV Panels Fault Diagnosis Method Based on NARX Network and Linguistic Fuzzy Rule-Based Systems,” *Sustainability*, vol. 12, no. 5. MDPI AG, p. 2011, Mar. 05, 2020. doi: 10.3390/su12052011.
- [12] F. Aziz, A. Ul Haq, S. Ahmad, Y. Mahmoud, M. Jalal, and U. Ali, “A Novel Convolutional Neural Network-Based Approach for Fault Classification in Photovoltaic Arrays,” *IEEE Access*, vol. 8. Institute of Electrical and Electronics Engineers (IEEE), pp. 41889–41904, 2020. doi: 10.1109/access.2020.2977116.
- [13] A. E. Lazzaretti et al., “A Monitoring System for Online Fault Detection and Classification in Photovoltaic Plants,” *Sensors*, vol. 20, no. 17. MDPI AG, p. 4688, Aug. 20, 2020. doi: 10.3390/s20174688.
- [14] M. Hussain, M. Dhimish, S. Titarenko, and P. Mather, “Artificial neural network based photovoltaic fault detection algorithm integrating two bi-directional input parameters,” *Renewable Energy*, vol. 155. Elsevier BV, pp. 1272–1292, Aug. 2020. doi: 10.1016/j.renene.2020.04.023.
- [15] S. A. Memon, Q. Javed, W.-G. Kim, Z. Mahmood, U. Khan, and M. Shahzad, “A Machine-Learning-Based Robust Classification Method for PV Panel Faults,” *Sensors*, vol. 22, no. 21. MDPI AG, p. 8515, Nov. 04, 2022. doi: 10.3390/s22218515.
- [16] D. Adhya, S. Chatterjee, and A. K. Chakraborty, “Performance assessment of selective machine learning techniques for improved PV array fault diagnosis,” *Sustainable Energy, Grids and Networks*, vol. 29. Elsevier BV, p. 100582, Mar. 2022. doi: 10.1016/j.segan.2021.100582.
- [17] A. F. Amiri, S. Kichou, H. Oudira, A. Chouder, and S. Silvestre, “Fault Detection and Diagnosis of a Photovoltaic System Based on Deep Learning Using the Combination of a Convolutional Neural Network (CNN) and Bidirectional Gated Recurrent Unit (Bi-GRU),” *Sustainability*, vol. 16, no. 3. MDPI AG, p. 1012, Jan. 24, 2024. doi: 10.3390/su16031012.
- [18] G. Varoquaux and O. Colliot, “Evaluating Machine Learning Models and Their Diagnostic Value,” *Machine Learning for Brain Disorders*. Springer US, pp. 601–630, 2023. doi: 10.1007/978-1-0716-3195-9_20.
- [19] Z. B. Duranay, “Fault Detection in Solar Energy Systems: A Deep Learning Approach,” *Electronics*, vol. 12, no. 21. MDPI AG, p. 4397, Oct. 24, 2023. doi: 10.3390/electronics12214397.
- [20] Y. Liu et al., “Fault diagnosis approach for photovoltaic array based on the stacked auto-encoder and clustering with I-V curves,” *Energy Conversion and Management*, vol. 245. Elsevier BV, p. 114603, Oct. 2021. doi: 10.1016/j.enconman.2021.114603.