Al-Rafidain Journal of Engineering Sciences

Journal homepage https://rjes.iq/index.php/rjes

ISSN 3005-3153 (Online)

Artificial Intelligence and Machine Learning Techniques for Fault Detection and Predictive Maintenance in Photovoltaic Systems: A Comprehensive Review

Diana Hassan Katea[1], Iman Jabbar Awda[2]

[1]Energy Engineering department, University of Baghdad, Baghdad, Iraq
[2]Aeronautical Engineering Department, University of Baghdad, Baghdad, Iraq

ARTICLE INFO

Article history:

Received	27 December 2025
Revised	27 December 2025
Accepted	3 February 2026,
Available online	7 February 2026

Keywords:
Artificial intelligence
Machine learning
Photovoltaic systems
Fault detection
Predictive maintenance

ABSTRACT

This paper gives a comprehensive analysis of the Artificial Intelligence (AI) and Machine Learning (ML) strategies used for fault detection, diagnosis, and prediction in photovoltaic (PV) systems. This research compares different machine learning algorithms, such as Support Vector Machines (SVM), Artificial Neural Networks (ANN), and Decision Trees (DT), with more advanced Deep Learning (DL) techniques, such as the use of Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM), for fault classification in PV systems. Findings gathered through literature studies show there is substantial evidence that machine learning strategies can be used for fault classification in photovoltaic (PV) systems, allowing for a computationally effective way to accomplish this task. Moreover, more advanced AI-based predictive maintenance techniques incorporating Prognostics and Health Management (PHM) with strategies for calculating the life of PV systems have immense potential for improving system dependability. Despite the progress made in AI-based PV prediction techniques, key issues of concerns in research continue to be the problem of limited datasets, imbalanced datasets, and explainability.

1. Introduction

The swift global implementation of photovoltaic (PV) energy systems has been recognized as a notable marker of the transition to sustainable and low-carbon sources of energy production. With their modularity, scalability, and ever-decreasing cost of installation, PV energy systems are expected to find widespread implementation for residential, commercial, and large-scale applications. As such, the need for these PV systems to work flawlessly and last for a long time without affecting energy production was imperative [1].

PV systems are prone to failure for electrical, thermal, environmental, or aging-related reasons. Some potential problems, like the degradation of the PV system, partial shading, inverter failure, wiring, or hotspot issues, can result in a decline in the power output of the system. If problems are not identified, the power loss may exceed 20% within the grid-connected PV system, thereby emphasizing the need for timely detection of problems within the PV system [2].

Conventional PV fault detection and maintenance practices usually involve manual inspections, alarm threshold levels based on limits, and rule-based reasoning. Though appropriate for smaller to medium-sized

Corresponding author E-mail address: d.hassan1809@coeng.uobaghdad.edu.iq
https://doi.org/10.61268/6dqdn415

schemes, this process is less efficient in larger PV farms due to an inability to adequately model the nonlinear dynamics exhibited by PV systems when subjected to changes in levels of intensity and temperature [2].

Recently, the use of Artificial Intelligence (AI) methodologies in combination with Machine Learning (ML) methodologies has started to emerge as an effective way to identify faults in PV (photovoltaic) technologies. By having access to both historical and live data, ML algorithms can identify a variety of complex system behaviours and highlight potential fault events occurring within a PV system. Research has shown that the implementation of ML methodologies in PV Technology provides higher accuracy rates when classifying the different types of faults than traditional fault classification methods used in the past. [3]

With the ever-increasing availability of operational data and processing capacity, deep learning methodologies have become a key domain for upgrading the analysis of faults in solar PV systems. The use of Convolutional Neural Networks has been widely leveraged to find defects in Solar Panels with thermal and electroluminescence images. In addition, Recurrent Neural Networks (RNNs), including Long Short Term Memory (LSTM) networks, also offer great potential for analyzing the historical operational data of solar PV systems. CNN models have been successfully employed to accurately identify faulty solar cells, facilitating mass production processes [4].

In addition to fault detection, the use of predictive maintenance through AI technology is becoming increasingly important as a method to improve the availability of PV power plants and reduce operation expenditure. Predictive maintenance is an analytics-based approach for predicting the health status and life of PV power plant components to allow maintenance to be performed before a failure occurs in large-scale PV power plants [5]. Although the number of works pertaining to AI and ML in the PV community is steadily increasing, the majority of the existing research works are currently being performed for particular algorithms, faults, or small datasets. A methodical combination of traditional ML techniques with DL and PM strategies has not yet been accomplished. It is therefore essential to carry out a critical analysis of the existing research works for the purpose of identifying research gaps in order to chart the future course for developing reliable PV monitoring systems using intelligent techniques.

2. Overview of Photovoltaic System Faults

Photovoltaic (PV) systems may be deployed in diverse environmental and electrical conditions, exposing their components to varying fault mechanisms for the entirety of their operational lifetimes. Faults may develop as a result of electrical faults, thermal faults, environmental faults, or aging effects. A clear classification of PV system faults is necessary for the development of accurate AI and machine learning (ML)–based fault detection and predictive maintenance models [1,2].

2.1 Electrical Faults in PV Systems

Electrical faults are regarded as one of the most potentially dangerous types of failure in PV implementations. These occur generally because of insulation falling below specified levels, problems during installation, cuts to cables, or faults in the inverter. Open- and short-circuit faults directly disrupt the flow of current (or leave it at abnormally low/high levels), with the potential to fatally compromise the energy output or safety of the system. It is shown that if electrical faults occur in grid-connected electrical systems and the faults are not located early, these faults can lead to large amounts of lost energy. The development of automated diagnosis systems would be an absolute necessity due to this fact [6]. Examples of line-to-ground faults and line-to-line faults are also presented, which could ultimately lead to overheating of the inverter and/or overheating of the inverters located in the controller, creating the potential for fires due to the failure of the inverter or the failure of the controller and/or component overheating and the subsequent fire. Traditional electrical protection devices (overcurrent relays, circuit breakers, etc.) do not typically detect electrical faults caused by line-to-ground or line-to-line failure in a timely manner, particularly during periods of fluctuating irradiance.. Demonstrated that under these conditions, data-driven fault recognition approaches outperform conventional protection mechanisms [7].

To address these challenges, machine learning algorithms such as Support Vector Machines (SVM) and Artificial Neural Networks (ANNs) are increasingly deployed to classify patterns representing electrical faults by processing current, power, and voltage signals under both normal and faulty scenarios [3, 8].

2.2 Thermal Faults

Thermal anomalies in photovoltaic (PV) systems are generally related to irregularities in the distribution of temperatures, including cell mismatches, shading, defects, or aging. These anomalies often lead to hotspots, whatever their cause, that can reduce the life of the systems. Thermal anomalies in photovoltaic modules can be located by infrared thermography, even before the performance of the defective cell deteriorates [9].

Modern research shows that deep learning, specifically convolutional neural networks, has the ability to automatically extract thermal characteristics from IR images and then correctly categorize these characteristics as either normal or faulty. CNN algorithms for automatic defect detection for PV modules have been proved to be much more efficient than traditional image processing methods [10].

More recent developments have upheld the use of deep learning-optimized thermographic analysis for the diagnosis of faults. For instance, CNN models, in addition to Vision Transformers (ViT), performed better in detecting hotspots/microcracks in large-scale photovoltaic systems, thus facilitating scalable analysis [4].

Likewise, it was found that hybrid deep learning models comprising both CNN and LSTM improve the interpretability and robustness of the fault detection method using thermal images [11]. As a result, deep learning enables the accurate detection of the hotspot in PV systems.

2.3 Environmental and Operational Faults

Environmental and operational aspects are also a source of variability of performance and fault generation in PV systems. Partial shading of PV arrays due to trees or buildings, soiling, etc., causes mismatch losses and losses in performance across the PV strings due to the aforementioned imperfect synchronization of the PV cell performance.

Many environmental effects, such as sunlight variations or shadows, leave the PV system exhibiting highly nonlinear characteristics, and they require the use of complex methods to recognize faults instead of simple thresholding methods. [8]

Soiling or accumulation of dirt is known to produce heavy performance losses in PV systems, especially in arid and semi-arid zones where losses may be significant. Field measurements show that soiling can cause energy losses of several percent up to nearly 40% annually depending on local dust deposition and rainfall patterns: for example, data from the Atacama Desert record annual energy losses that peak near 39% due to soiling in dryer zones. Nature [12]

Additionally, reviews of PV soiling effects confirm that deposition of dust and atmospheric particulates significantly reduces module transmission of sunlight, which directly lowers output performance and increases mismatch losses across PV strings. [13]

Machine learning techniques–based models trained on such environmental-specific outputs may yield better insight into whether a performance drop is temporary (e.g., dust accumulation) or actually indicative of a fault. According to studies regarding the utilization of data to analyze and classify soiling patterns, machine learning (ML) technologies could be employed to accurately forecast and compute the soiling ratio while distinguishing the effect of soiling on solar array performance from other causes of performance variations, thus yielding improved results regarding both defective component/fault detection and maintenance scheduling compared to simple thresholds [14]. Soiling significantly reduces PV output; ML aids fault distinction.

2.4 Aging and Degradation-Related Faults

The long-term, continuous operation of a PV system results in the aging and degradation of components as power output and reliability gradually decrease. In distinction to random electrical faults, those arising from degradation evolve slowly over time and present a greater challenge for detection using standard monitoring procedures. They are heavily influenced by exposure, thermal cycling, humidity, and mechanical stress. Common degradation phenomena in PV modules include encapsulant sun-tanning and delamination, corrosion of metallic contacts and micro-cracking of cells, indicating that field-aged PV modules exhibit annual degradation rates of the order of 0.5% to 1% a year, depending on climate and module quality [15].

Voids and failure of soldered joints arising out of mechanical and thermal stress can gradually propagate and result in mismatch losses and hot spots. Studies such as this confirm that encapsulant delamination, discoloration, and solder fatigue significantly affect module reliability and cannot be reliably detected through visual inspection alone [16].

AI-based prognostics involving machine and deep learning are appropriate for analyzing degradation-related faults, as they can learn long-term deterioration trends from operational data. These models can shift maintenance strategies from reactive to predictive. demonstrated that data-driven AI models can effectively predict degradation patterns and identify early failure indicators in PV components [17].

2.5 Fault Classification Based on System Location

Viewed at a system level and schematized as such, faults in PV systems can be categorized according to their location and functional domain:

- DC side: module defects, string or cable disconnections, and cable insulation failures.
- AC side: inverters, or synchronization with the grid; transformers.
- System: sensors, communication, and data acquisition.

Recent research shows that fault classification accuracy improves when system-level information (combining DC, AC, and system sensor data) is incorporated into machine learning models. For example, they proposed a multi-output deep learning approach that detects, classifies, and locates different faults such as string-to-string (SS), string-to-ground (SG), and open-circuit (OC) conditions using CNN, LSTM, and Bi-LSTM, achieving high classification and localization accuracy in PV systems [18].

Another study using machine learning for PV fault classification evaluated multiple ensemble methods (such as random forest, K-nearest neighbors, and extreme gradient boosting) to identify and categorize several fault types with high accuracy, emphasizing the value of system-level operational data in training the diagnostic models [19]. System-level ML models significantly enhance PV fault classification accuracy.

3. Data Acquisition and Monitoring in Photovoltaic Systems

For successful implementation of AI and ML fault detection and predictive maintenance approaches reliable data acquisition is a fundamental requirement. The performance of the diagnostic and prognostic models are highly dependant on the nature, type and resolution of the data collected. The majority of the new PV monitoring systems will be composed of multiple sources of data such as (i) electrical measurements, (ii) environmental measurements, (iii) imaging techniques and (iv) supervisory control.

3.1 Electrical Measurements and SCADA Systems

Voltage, current, power output, and yield of energy are the main parameters for characterizing the electrical behavior of a PV system. Measurements are usually taken using Supervisory Control and Data Acquisition (SCADA) systems, gathering up-to-date and historical information on the operation of both local strings and at the inverter level.

proved that using SCADA for monitoring of the irradiance-induced power output produces early indication of abnormal equipment operation by comparing actual output with expected modeled output, where the produced difference between actual versus modeled power value indicates the presence of a fault [6].

SCADA data is often sufficient in order to achieve a machine learning system for fault detection of the PV cells, mostly due to its availability and the relatively low cost required to implement it, as demonstrated in studies that use SCADA signals as inputs for automated classification of faults with statistical and ML-based methods [20].

Some faults may, however, not be accurately localized using electrical data alone—for instance, thermal and degradation faults—and researchers have proposed combining SCADA electrical measurements with other diagnostic inputs (such as model-based prediction residuals or image-based sensors) to improve detection performance. This hybrid approach has been highlighted in recent fault detection literature as a way to compensate for the limitations of using electrical variables by themselves [21].

3.2 I–V Curve Measurement and Analysis

It has been shown that the analysis of the I-V curve can provide important information about the system operation, including the detection of particular defect states, such as partial shading, the effect of series resistance, as well as the occurrence of an open-circuit failure, based solely on the current-voltage (I-V) characteristics of the photovoltaic modules or strings [6].Pattern recognition tools combined with the current voltage measurement (IV) properties of PV modules will produce a very effective new method for identifying defects within PV modules, such as shading, degradation, and mismatch. Recent studies have shown that machine-learning techniques, for example, ANN-SVM, can be utilized to accurately classify defects in a PV module into short-circuit, shading, and increased resistance defects by applying these techniques to the extracted features from the current-voltage measurements. As a result, the classification accuracy of the machine learning techniques has been found to be very high [20].

3.3 Environmental and Meteorological Data

As extensively documented, environmental data supports the interpretation of photovoltaic (PV) systems' performance and helps prevent false identification of faults. Solar irradiance, ambient temperature, wind speed, and humidity all significantly impact PV generation output and the accuracy of fault diagnosis. By incorporating the above environmental measurements into a data-driven model, developers can significantly enhance their reliability for identifying faults by differentiating between weather-related changes in performance and true system faults. Furthermore, several Machine Learning (ML) models have outperformed more traditional ML models due to their reliance on environmental data as either a feature or normalization input when predicting PV output and reporting on PV faults.

3.4 Thermal Imaging and Infrared Thermography

IR thermography has suddenly found itself a popular non-destructive testing method for scanning hotspots in photovoltaic (PV) cells. IR thermography makes it possible to visualize the local differences in temperature, thereby detecting cell defects, partial shading, and degradation. The successful application of IR thermography has shown that this technology can detect defects caused by cell degradation before electrical measurements alone, in the life cycle of a PV cell [25]. The integration of the latest deep learning techniques in thermal imaging has further enhanced the technology’s potency to automatically classify defects in photovoltaic cells. Convolutional Neural Networks (CNN) have the potential to process the spatial information present in the IR images, making it possible to accurately detect the defects in the photovoltaic cell without manual dimensioning of the features [26]. Deep learning algorithms, like ensemble CNN and edge models, have shown the accuracy level of more than 95% in identifying various types of defects in photovoltaic cells, using IR images, thereby proving the potency of AI-based thermal image analysis techniques [27].

3.5 Advanced Sensing and Imaging Techniques

Apart from the infrared thermography technique, a number of modern image technologies have been designed in order to provide a more extensive evaluation of the quality of the photovoltaic (PV) module. For example, the integration of the electroluminescence image technology and the deep learning algorithm provides a remarkably accurate evaluation of microcracks and cell defects of crystalline silicon PV modules [28]. Similarly, the image technology of photoluminescence has been shown to be a reliable method for the evaluation of the long-term degradation of the PV module, in addition to the microstructural damage of the PV module, even in field inspections using a drone system [29]. Whereas these image technologies provide a profoundly precise evaluation, they still have a limited scope of application in the practical inspection of PV modules owing to the high cost of the image technology and the complexity of the system.

3.6 Data Requirements for AI-Based Fault Detection

In terms of AI and machine learning-based fault detection models for PV systems, for these models to be effective, there must be a sufficiently large operational database with high-quality labels and high sampling resolution. Because operational faults in end-user PV systems occur far less frequently than normal operational conditions, this results in a class imbalance that limits the reliability of supervised learning models unless proper precautions are taken. In support of this issue, studies have shown that PV condition monitoring can achieve high accuracy even when only a small portion of the dataset is annotated, highlighting both the difficulty of obtaining labeled data and the effectiveness of semi-supervised or weakly supervised learning methods for real-world PV monitoring applications [30].

Furthermore, large-scale PV plants primarily rely on SCADA and inverter telemetry data for continuous monitoring and predictive maintenance. AI-based predictive models depend strongly on the availability and continuity of these historical operational datasets to detect deviations from expected performance and predict emerging faults effectively [31]. Therefore, maintaining a robust and comprehensive data collection infrastructure is essential for the successful implementation of reliable AI-based fault detection in PV systems.

4. Machine Learning Techniques for Fault Detection in Photovoltaic Systems

Because of the ability of machine learning (ML) techniques to model the nonlinear relationship between system variables and operating conditions, they have been used extensively for fault detection and diagnosis in photovoltaic (PV) systems. Unlike rule-based systems, where rules are manually created or programmed by experts, ML algorithms learn fault patterns directly from data, making them considerably better adapted to changing irradiance and temperature conditions. Figure 1 illustrates a general classification of artificial intelligence algorithms commonly used for fault detection in photovoltaic systems, including classical machine learning methods, deep learning approaches, and emerging AI-based techniques.

4.1 Artificial Neural Networks (ANN)

Artificial neural networks (ANNs), one of the oldest types of artificial intelligence (AI), have been used to diagnose faults in photovoltaic (PV) systems for many years. The common application of ANNs to faults in PV systems is due to their ability to learn from historical information and model the interaction between temperature, irradiance, I-V parameters, and string output. Research studies have confirmed that ANNs can differentiate between normal operation and faulty operation of PV systems when the training of the ANN included wellrepresented data sets covering a variety of fault types. A previous study employed an ANN with multiple attributes to create a diagnostic algorithm capable of differentiating multiple types of PV faults at different irradiance and temperature levels. The findings of this study suggested that data representativeness and selecting appropriate features are critical factors to achieving reliable classifications for PV fault diagnosis. In addition, a recent systematic literature review found that the use of ANNs—especially hybrid and optimized architectures—is among the most reliable and flexible methods for diagnosing PV faults because of the ability of ANNs to analyze nonlinear multi-dimensional data types effectively.

4.2 Support Vector Machines (SVM)

Support vector machines, or SVMs, are an example of a supervised learning algorithm, which is a type of machine learning that uses labeled training data to solve problems related to classification and regression. In particular, SVMs have recently gained popularity in the PV (photovoltaic) community due to their effectiveness in detecting faults within a solar panel from low levels of training data and the high dimensionality of the feature space. Research has demonstrated that SVM-based classifier systems exhibit excellent capability in classifying faults in PV electrical data, particularly distinguishing between faults caused by shading and those caused by electrical disconnects [34,35]. Although SVM model performance is not as impacted by encountering local minima as it is with ANN models, the quality of the results produced by an SVM classifier is heavily dependent on both the choice of kernel and the careful selection of suitable parameter settings.

4.3 Decision Trees and Rule-Based Learners

The Decision Tree (DT) algorithm uses a series of decision rules to classify faults in an easy-to-understand hierarchical format. Because of their clear structure, DTs are ideal for use in PV Monitoring Systems that require an explanation of how classification decisions are made. Research indicates that SCADA information can successfully identify problems with PV plants when using decision tree-based algorithms, achieving high levels of accuracy and explainability of results that allow operators to be aware of the source of the problems due to the transparent nature of decision-making rules used in DT-based programming [36]. Also consistent with this research are results obtained from multiple studies showing that Decision Tree and ensemble methods (Extra Trees and Random Forest) 'classify PV faults such as partial shading, open-circuit and disconnection with exceptional levels of diagnostic accuracy while being user-friendly for real-time monitoring' [37]. A limitation of decision trees is that when using single-tree models, performance may degrade because of overfitting when working with noisy or unevenly distributed data. Overfitting can be minimized using pruning methods or building ensembles of trees.

4.4 Ensemble Learning Techniques

Ensemble learning techniques such as Random Forest (RF) and Gradient Boost both employ a combination of several base classifiers to improve the robustness and accuracy of classifying data. Researchers and users have recently taken an increased interest in using ensemble learning methods for photovoltaic (PV) fault detection. Compared to using just one single classifier, studies have demonstrated that using ensemble-based models like RF and gradient boosting greatly improves the accuracy of fault prediction for PV systems by reducing the variance of error through the increased level of generalization across a wider range of operating conditions [38, 39]. The primary advantages of using the ensemble classification techniques are the ability to use multiple base classifiers to combine strengths and reduce weaknesses of any one classifier in the process of analyzing heterogeneous data or recognizing complex patterns of failure within PV systems.

4.5 Comparative Performance of Classical ML Techniques

Multiple research papers have focused on the performance of various machine learning algorithms in terms of detecting faults in photovoltaic systems and how they compare to one another. All the papers came to the same conclusion: artificial neural networks and random forests provide the greatest accuracy for detecting faults in photovoltaic systems, whereas support vector machines work well when used with small sample sizes. Similarly, while the decision tree classification model can

provide greater interpretability than other models, it may sacrifice some degree of accuracy.

From the results of the comparative studies, it is clear that there is no single machine-learning technique that will always be preferred for every situation. Rather, the choice of algorithm used should depend on what type of data you have, how much time and money you are willing to invest, and what kind of fault you want to identify [40,37].

Table 1: Comparison of AI-Based Techniques for PV Fault Detection and Predictive Maintenance

Technique	Data Type	Accuracy	Data Requirement	Computational Cost	Interpretability	Typical Application
ANN	Electrical, environmental	High	Medium	Medium	Medium	Fault classification
SVM	Electrical	Medium–High	Low–Medium	Low	Low	Small datasets
Decision Trees	Electrical, SCADA	Medium	Low	Low	High	Explainable diagnosis
Random Forest	Multi-source	High	Medium	Medium	Medium	Robust fault detection
CNN	Thermal / image	Very High	High	High	Low	Hotspot & defect detection
LSTM	Time-series	High	High	High	Low	Degradation & RUL prediction
Autoencoders	Electrical, SCADA	Medium–High	Low	Medium	Low	Anomaly detection

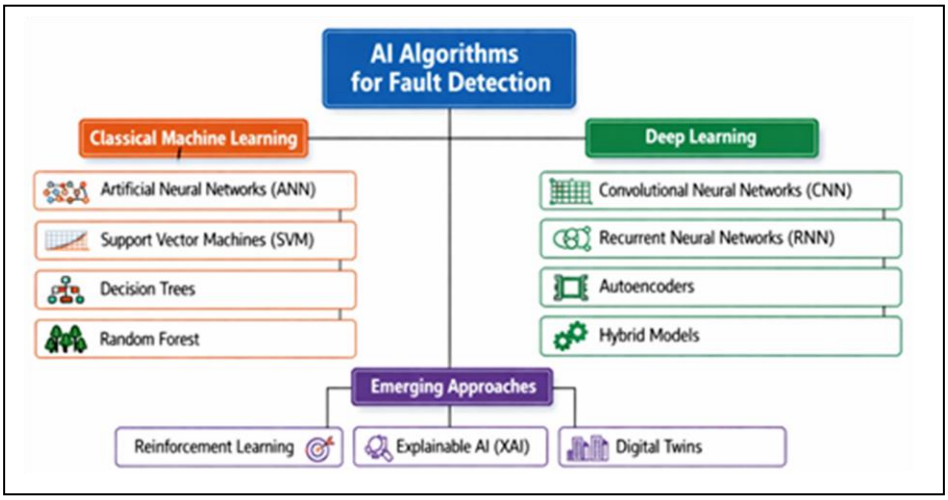

Figure 1. Classification of artificial intelligence algorithms used for fault detection

5. Deep Learning and Advanced AI Approaches for PV Fault Diagnosis

As a result of the rapid expansion of data available to researchers and developers, as well as advancements in computing power, many researchers and developers are now leveraging deep learning (DL) methods for fault detection and diagnosis in PV systems. In contrast to traditional machine learning algorithms, DL-based models can automatically identify higher-order, abstract features from unprocessed data, so they are particularly suited to working with extremely complex, multidimensional data (e.g., images, thermal images, and extended time series).

Several recent review papers have documented the increasing trend toward using DL-based techniques for fault diagnosis of PVs due to the superior capabilities of these techniques for processing highly complex, multidimensional data. conducted a comprehensive review of methodologies for fault diagnosis in PV systems based on deep learning, concluding that convolutional neural networks (CNNs) and recurrent neural networks (RNNs) consistently outperform classical machine learning algorithms when diagnosing faults from images or evaluating degradation over time [41].

Likewise, demonstrated the effectiveness of hybrid DL architectures combining CNN and Bi-GRU layers, achieving highly accurate classification of PV faults such as open circuits, short circuits, and partial shading [42]. Their findings reinforce the expectation that DL-based fault diagnosis models will continue to grow in PV monitoring systems as both data availability and computational resources improve.

5.1 Convolutional Neural Networks (CNN)

Among deep learning architectures, Convolutional Neural Networks (CNNs) have become the most commonly used models for detecting faults in photovoltaic (PV) systems using image-based data. CNNs can efficiently analyze electroluminescence (EL) and thermal infrared (IR) images, automatically extracting spatial features that correspond to cracks, hotspots, delamination, and micro-defects in PV modules. developed a lightweight CNN architecture for automatic defect detection in PV cell electroluminescence (EL) images, achieving 91.7% accuracy with reduced computational demand. Their work demonstrated that CNNs can accurately identify microcracks and localized cell defects without manual feature extraction, making them suitable for real-time industrial inspection [43]. Complementarily, applied CNNs to infrared drone imagery for large-scale PV plants. Their deep edge-based CNN framework achieved fast and accurate classification of defective modules and hotspot-related faults, confirming CNNs' superiority in processing complex IR data for automated PV field inspection [44].

5.2 Recurrent Neural Networks (RNN) and LSTM Models

Recurrent Neural Networks (RNNs) and their sophisticated version, Long Short-Term Memory (LSTM) networks, have proven to be useful tools in analyzing the data generated by photovoltaic systems in terms of their power outputs, irradiation, as well as SCADA telemetry data. These types of networks have the ability to analyze data to detect any unusual pattern as well as predict potential faults.

It was recently shown that an accuracy rate of over 91% can be achieved by an LSTM classifier in recognizing high-impedance fault occurrences within grid-connected PV systems, compared to conventional models like SVM and decision trees[45].

5.3 Autoencoders and Unsupervised Deep Learning

AutoEncoders refer to unsupervised deep learning models used for compressing data. By using data related to the normal functioning of PV systems, anomalies can be detected based on how they depart from the norm. Frameworks based on Stacked AutoEncoders have been designed for fault diagnosis and have achieved more than 99% accuracy in different fault models in PV systems, implying the aptness of unsupervised approaches on models for which data is limited in number [46].

5.4 Hybrid and Advanced AI Models

Hybrid models of artificial intelligence that combine classical machine learning algorithms with deep learning models have proved to be effective tools for diagnosing photovoltaic (PV) faults. This is achieved using a combination of feature extraction algorithms like CNNs with classifiers like SVM or Random Forest algorithms. For example, a hybrid deep learning algorithm that combines a long short-term memory (LSTM) network with a stacked autoencoder (SAE) outperforms both models when diagnosing symmetrical and asymmetrical faults in photovoltaics [47].

5.5 Advantages and Limitations of Deep Learning Approaches

The techniques of deep learning are highly accurate, able to automatically detect features, and scalable for handling large amounts of photovoltaic data. However, their performance is challenged by significant computational complexity, dependence on large amounts of labeled data, and lack of robustness in interpretation. This fact has been stressed, underlining that the lack of real-time interpretation and interpretation/interpretability are significant barriers in adopting models of deep learning for implementation in PV systems [41].

6. Predictive Maintenance in Photovoltaic Systems

Predictive maintenance has recently received significant attention for improving the reliability, availability, and economic efficiency of photovoltaic systems. Unlike traditional repair or scheduled maintenance, predictive maintenance relies on the constant monitoring of the condition of the components and the predictions of models that allow the failure of the components of a photovoltaic system to be forecasted in advance.

6.1 Maintenance Strategies in PV Systems

Typically, maintenance of photovoltaic (PV) systems is classified as either corrective, preventive, or predictive maintenance. The most advanced and cost-effective method for maintaining PV systems is through predictive maintenance, which utilizes continuous condition monitoring and analytical data to enable early detection of potential failures, intermediate fault detection, and optimal scheduling of preventive maintenance. Emphasizes the fact that by facilitating integration of predictive maintenance techniques and data analytics with risk-based planning methodologies, predictive maintenance systems provide greater reliability and availability to PV Systems [48]. Additionally, it states that it can significantly reduce the cost of operation and downtime for PV systems through the accurate, data-driven prediction of faults that would otherwise be undetectable [5].

6.2 Prognostics and Health Management (PHM)

PHM (Prognostics and Health Management) is an evolving area in photovoltaic (PV) systems that includes detecting faults, diagnosing problems, and estimating the useful life of a system in order to determine its health and provide predictive maintenance. To predict failure and degradation, PHM utilizes data gathered from sensors and uses various learning models to learn these patterns; this capability has been demonstrated, through recent research, to demonstrate PHM as a tool in enabling reliability-centered maintenance (RCM) of renewable energy systems, using earlier work that supports this vision, as well as AI-based prognostic tools in solar energy applications [49].

6.3 Remaining Useful Life (RUL) Estimation

An estimation of Remaining Useful Life (RUL) allows a manufacturing company to plan for the life of its products, allowing for scheduling preventive maintenance as well as reducing cost through the prediction of when a component will fail. The semiparametric framework presented here shows an improvement to the accuracy of predicting RUL of PV modules when accounting for their operating environment, which is known to cause significant variations in physical configurations and performance metrics [50]. This study utilizes weather data and degradation history to provide manufacturers with an accurate estimate of the RUL of a PV component throughout its operational life.

6.4 Data-Driven Predictive Maintenance Models

Data-driven predictive maintenance relies on real-time and historical data from SCADA, inverter telemetry, and environmental sensors to predict system failures before they occur. A scalable data-driven fault prediction model for utility-scale PV plants, achieving up to 95% fault prediction sensitivity using pattern recognition and unsupervised clustering based on SCADA data [51].

6.5 Benefits and Challenges of Predictive Maintenance in PV Systems

Predictive maintenance can help increase uptime, improve system reliability, and reduce costs; however, many challenges remain regarding data quality, model uncertainty, and making predictions based on these models. Accurate forecasts of environmental factors, such as irradiance and temperature, are critical to the success of predictive maintenance systems, along with strong data integration frameworks that minimize the uncertainty associated with model predictions [52].

7. Comparative Analysis and Discussion

This section will introduce a comparative assessment of Artificial Intelligence (AI), Machine Learning (ML), and Deep Learning (DL) methods in the application of fault detection and predictive maintenance within Photovoltaic (PV) systems. The focus will be placed upon assessing methods in terms of accuracy, data needs, complexity, explainability, and real-world feasibility. This comparative assessment is critical in establishing which technology would be best suited for PV systems depending upon their size.

7.1 Comparison between Classical ML and Deep Learning Techniques

Classical machine learning techniques, such as Artificial Neural Networks (ANN), Support Vector Machines (SVM), and Decision Trees (DT), can be successfully employed for photovoltaic (PV) fault detection through the appropriate use of feature engineering. They generally require less computational cost and data, which helps in their use for real-time analysis for small PV plants. However, the use of more recent approaches like Convolutional Neural Networks (CNN) and Long Short-Term Memory (LSTM) networks, which can handle large-scale data like thermal or EL images, has been observed to offer improved performance compared to the classical approaches. In this context, it has been found that the use of CNN-based models for fault analysis provides a higher accuracy level compared to the models based on SVM and Random Forest, which require more computational cost [53].

7.2 Diagnostic Accuracy and Data Requirements

Labeled datasets are a major determinant of the accuracy of ML machine learning models. Classical ML, such as Random Forest and ANN (Artificial Neural Networks), works well on structured data (electrical and environmental data), while deep learning outperforms these models on unstructured datasets (images, sequential data). A new deep stack-based ensemble learning (DSEL) model was developed by combining Deep Neural Networks (DNN) with Long Short-Term Memory Networks (LSTMs) and Bidirectional Long Short-Term Memory Networks (Bi-LSTMs). The DSEL model achieved an accuracy of 98.6% for PV fault detection in noisy datasets and was more accurate than classical ML methods. [54]

7.3 Computational Complexity and Interpretability

The low computational costs and high interpretability of traditional ML tools make them best suited for use in embedded or edge environments. Decision trees and random forest models are backbones of transparency and are therefore useful for embedded and edge systems.

By contrast, deep learning models often operate in a "black box" fashion, which limits their ability to be understood or interpreted by people. As a result, many users lack trust in ML-based models and do not want to use them in their industrial applications [55]. To address this issue, we have applied Explainable Artificial Intelligence (XAI) techniques (SHAP and DiCE) to the models developed for the fault detection of Photovoltaic (PV) installations, which has been shown to significantly increase end-user confidence in the model without compromising model diagnostic accuracy [57].

7.4 Applicability to Predictive Maintenance

LSTM and hybrid deep learning-based predictive maintenance solutions have outperformed traditional Machine Learning (ML) techniques for prediction of the Remaining Useful Life (RUL). A novel hybrid framework for predictive maintenance for the RUL of systems was developed, which employs a stack of sparse autoencoders (SSAE), and it was established that using optimal multi-layer feed-forward neural networks (OMLP) results in an exceptional 99.8% classification accuracy for fault diagnosis, outperforming commonly used techniques for predictive analytics. [56].

7.5 Discussion

At this point, the research does not indicate that any one type of AI approach will be the best way of detecting faults in PV systems or for predictive maintenance in the future.

Classical ML has advantages due to its relative simplicity, ability to understand results easily, and speed at which it can make predictions. However, DL methods provide superior accuracy and ability to detect complex patterns in large data sets and also work well for high dimensionality (e.g., billions) of data versus ML models that work best on low-dimensional data. Recent research shows that there is value in combining ML and DL methods in the construction of hybrid architectures to achieve a balance of accuracy, ability to explain results and computational efficiency. in their paper "Towards a Hybrid AI Method for Predictive Maintenance in the Solar Photovoltaic Sector" report that combining both types of machine learning techniques may produce better results than using ML or DL exclusively [57].

8. Challenges, Limitations, and Future Trends

Although significant improvements in such photovoltaic (PV) fault detection and maintenance tasks through the use of artificial intelligence (AI), machine learning (ML), and deep learning (DL) methods have been accomplished, several technological limitations continue to hinder their broader industrial applications. The limitations will be discussed, with emphasis on trends that could increase their level of readiness for industrial applications.

8.1 Key Challenges and Limitations

8.1.1 Data scarcity and class imbalance

The lack of a sufficient amount of labelled PV fault data causes a significant imbalance in the classes, therefore causing supervised models to be biased toward the label of normal operation. However, the utilization of semi-supervised and weakly supervised learning methods can help to alleviate this problem through the efficient use of the unlabelled data [30].

8.1.2 Data quality, missing values, and sensor uncertainty

Actual real-time PV data from SCADA and sensor networks commonly experience missing values, noise and drift in calibration that deteriorate the performance of the model. Effective methods for interpreting this type of data must also include techniques for handling uncertainty and long-term analysis of failed components through effective fault detection and diagnostic analysis [31].

8.1.3 Generalization across sites and climates

Because of the variation between modules, inverters, and environmental conditions, PV fault detection models developed for one site do not always perform well when applied to other sites. Therefore, using domain adaptation and transfer learning techniques can improve the ability of these models to adapt to new sites [44].

8.1.4 Lack of explainability and operator trust

The “black box” nature of deep learning models limits interpretability and operator trust in PV maintenance. Explainable AI (XAI) methods can visualize which parameters or image regions contribute most to fault diagnosis, improving user confidence [57].

8.1.5 Real-time deployment constraints

In many cases, photovoltaic systems require decisions in a real-time or near-real-time manner. Computations involved in deep learning inference can be resource-heavy, especially for image-related diagnostics. The resource limitations in edge devices, or edge sites, have led the way to lean model architectures.

8.1.6 Standardization and benchmark datasets

One area that has always presented a problem is the unavailability of common datasets and assessment procedures. The reported findings in the literature are often difficult to compare because of variations in the definition of faults, the sampling rate, and the evaluation criteria [41].

8.2 Future Trends and Research Directions

8.2.1 Hybrid AI frameworks

Hybrid approaches combining physical understanding, particularly photovoltaic (PV) models of solar performance with data-intensive methods of artifical intelligence, can be expected to strengthen robustness as well as reduce data requirements. Techniques encode computational constraints on the learning process with physical priors while retaining flexibility to address nonlinearities [7].

8.2.2 Transfer learning and domain adaptation

Transfer learning enables the use of pretrained CNNs or LSTMs to adapt fault detection models to new climates and configurations, reducing the need for large datasets. Domain adaptation methods will play a crucial role in adapting to models in diverse climatic conditions, module types, and sensing systems [58].

8.2.3 Digital twins for PV condition monitoring

The technology of digital twin is increasingly recognized as an effective means for intelligent monitoring and predictive maintenance of photovoltaic energy systems. In their work, they analyzed the application of digital twin technology in smart energy management and stressed its ability for real-time monitoring, fault simulation, and optimization of maintenance activities. The simulation of fault propagation and analysis of maintenance options based on AI models and digital twin technology can be achieved for PV energy systems [59].

8.2.4 Edge AI and IoT integration

Combining IoT sensing with edge AI allows for low-latency fault detection and scalability in large PV plants. Lightweight models and on-device inference reduce communication [44].

The state-of-the-art photovoltaic fault detection and predictive maintenance solutions using artificial intelligence face issues related to the availability, quality, and generalization capabilities of the data, as well as explainability and the ability for real-time execution.

9. Conclusion

The research considered in this document describes a detailed view of how Artificial Intelligence (AI) and Machine Learning (ML) methods are being incorporated into the monitoring of photovoltaic (PV) systems, the identification of faults in them, and the prediction of required maintenance on them. In addition, it illustrates that using new and previously established methods to process data—the traditional ML models previously mentioned, as well as the Deep Learning (DL) models previously mentioned—results in significant gains in diagnostic ability, dependability, and productivity in PV systems.

Classical Machine Learning (ML) techniques are still extremely useful for applications requiring little computational power while providing interpretable results in real-time settings; however, in dealing with the complexities of high-dimensional data (e.g., thermographic/electroluminescence images), Deep Learning (DL) will far exceed the performance of Classical (ML). The recent development of new "hybrid" models that combine aspects of both are likely to provide a good balance between cost, scalability, and accuracy of intelligent PV maintenance systems; therefore, these hybrid systems will be the next major development in this area.

The results of this investigation emphasize the potential for transformation through the use of AI technology for predictive maintenance through the application of PHM and RUL. There are, however, still a few persistent issues that need to be addressed with continued research and development, such as data imbalance, model explanations, and the ability to generalize across multi-site applications.

Future research can focus on developing open-source standardized datasets and hybrid AI models that are physics-informed and creating explainable AI models to build transparency and increase operator confidence in using these models. There are also many opportunities available for furthering the use of Edge AI and Digital Twin technologies in conjunction with Federated Learning to provide scalable, autonomous, and privacy-preserved solutions for photovoltaic (PV) maintenance.

Overall, the findings of this study further indicate that AI-enhanced predictive maintenance is pivotal to enabling the transition toward sustainable, intelligent, and self-optimizing solar energy systems globally, which corresponds to the goals of the renewable energy sector to align itself with the ongoing transformation resulting from the occurrence and proliferation of industry.

References

[1] E. H. Sepúlveda-Oviedo, L. Travé-Massuyès, A. Subias, M. Pavlov, and C. Alonso, “Fault diagnosis of photovoltaic systems using artificial intelligence: A bibliometric approach,” Heliyon, vol. 9, no. 11, p. e21491, Oct. 2023, doi: 10.1016/j.heliyon.2023.e21491. Available: https://doi.org/10.1016/j.heliyon.2023.e21491

[2] G. M. El-Banby, N. M. Moawad, B. A. Abouzalm, W. F. Abouzaid, and E. A. Ramadan, “Photovoltaic system fault detection techniques: a review,” Neural Computing and Applications, vol. 35, no. 35, pp. 24829–24842, Oct. 2023, doi: 10.1007/s00521-023-09041-7. Available: https://doi.org/10.1007/s00521-023-09041-7

[3] V. Khandeparkar, N. Shreshtha, and S. K. Ramu, “Effectiveness of supervised machine learning models for electrical fault detection in solar PV systems,” Scientific Reports, vol. 15, no. 1, p. 34919, Oct. 2025, doi: 10.1038/s41598-025-18802-4. Available: https://doi.org/10.1038/s41598-025-18802-4

[4] M. Jalal, I. U. Khalil, and A. U. Haq, “Deep learning approaches for visual faults diagnosis of photovoltaic systems: State-of-the-Art review,” Results in Engineering, vol. 23, p. 102622, Jul. 2024, doi: 10.1016/j.rineng.2024.102622. Available: https://doi.org/10.1016/j.rineng.2024.102622

[5] R. V. Vichare and S. R. Gaikwad, “AI-based predictive maintenance of solar photovoltaics systems: a comprehensive review,” Energy Informatics, vol. 8, no. 1, Oct. 2025, doi: 10.1186/s42162-025-00594-6. Available: https://doi.org/10.1186/s42162-025-00594-6

[6] A. Chouder and S. Silvestre, “Automatic supervision and fault detection of PV systems based on power losses analysis,” Energy Conversion and Management, vol. 51, no. 10, pp. 1929–1937, Mar. 2010, doi: 10.1016/j.enconman.2010.02.025. Available: https://doi.org/10.1016/j.enconman.2010.02.025

[7] A. Triki-Lahiani, A. B.-B. Abdelghani, and I. Slama-Belkhodja, “Fault detection and monitoring systems for photovoltaic installations: A review,” Renewable and Sustainable Energy Reviews, vol. 82, pp. 2680–2692, Oct. 2017, doi: 10.1016/j.rser.2017.09.101. Available: https://doi.org/10.1016/j.rser.2017.09.101

[8] A. Mellit and S. A. Kalogirou, “Artificial intelligence techniques for photovoltaic applications: A review,” Progress in Energy and Combustion Science, vol. 34, no. 5, pp. 574–632, Mar. 2008, doi: 10.1016/j.pecs.2008.01.001. Available: https://doi.org/10.1016/j.pecs.2008.01.001

[9] E. Kaplani, “Detection of Degradation Effects in Field-Aged c-Si Solar Cells through IR Thermography and Digital Image Processing,” International Journal of Photoenergy, vol. 2012, pp. 1–11, Jan. 2012, doi: 10.1155/2012/396792. Available: https://doi.org/10.1155/2012/396792

[10] M. Vlaminck, R. Heidbuchel, W. Philips, and H. Luong, “Region-Based CNN for anomaly detection in PV power plants using aerial imagery,” Sensors, vol. 22, no. 3, p. 1244, Feb. 2022, doi: 10.3390/s22031244. Available: https://doi.org/10.3390/s22031244

[11] M. Benghanem, A. Mellit, and C. Moussaoui, “Embedded hybrid model (CNN–ML) for fault diagnosis of photovoltaic modules using thermographic images,” Sustainability, vol. 15, no. 10, p. 7811, May 2023, doi: 10.3390/su15107811. Available: https://doi.org/10.3390/su15107811

[12] R. R. Cordero et al., “Effects of soiling on photovoltaic (PV) modules in the Atacama Desert,”

Scientific Reports, vol. 8, no. 1, p. 13943, Sep. 2018, doi: 10.1038/s41598-018-32291-8. Available: https://doi.org/10.1038/s41598-018-32291-8

[13] P. Borah, L. Micheli, and N. Sarmah, “Analysis of soiling loss in photovoltaic modules: A review of the impact of atmospheric parameters, soil properties, and mitigation approaches,” Sustainability, vol. 15, no. 24, p. 16669, Dec. 2023, doi: 10.3390/su152416669. Available: https://doi.org/10.3390/su152416669

[14] A. Kalimeris, I. Psarros, G. Giannopoulos, M. Terrovitis, G. Papastefanatos, and G. Kotsis, “Data-driven soiling detection in PV modules,” arXiv (Cornell University), Jan. 2023, doi: 10.48550/arxiv.2301.12939. Available: http://arxiv.org/abs/2301.12939

[15] D. C. Jordan and S. R. Kurtz, “Photovoltaic Degradation Rates—An analytical review,” Progress in Photovoltaics Research and Applications, vol. 21, no. 1, pp. 12–29, Oct. 2011, doi: 10.1002/pip.1182. Available: https://doi.org/10.1002/pip.1182

[16] L. Koester, S. Lindig, A. Louwen, A. Astigarraga, G. Manzolini, and D. Moser, “Review of photovoltaic module degradation, field inspection techniques and techno-economic assessment,” Renewable and Sustainable Energy Reviews, vol. 165, p. 112616, May 2022, doi: 10.1016/j.rser.2022.112616. Available: https://doi.org/10.1016/j.rser.2022.112616

[17] T. Rahman et al., “Investigation of Degradation of Solar Photovoltaics: A Review of Aging Factors, Impacts, and Future Directions toward Sustainable Energy Management,” Energies, vol. 16, no. 9, p. 3706, Apr. 2023, doi: 10.3390/en16093706. Available: https://doi.org/10.3390/en16093706

[18] Z. Mustafa, A. S. A. Awad, M. Azzouz, and A. Azab, “Fault identification for photovoltaic systems using a multi-output deep learning approach,” Expert Systems With Applications, vol. 211, p. 118551, Aug. 2022, doi: 10.1016/j.eswa.2022.118551. Available: https://doi.org/10.1016/j.eswa.2022.118551

[19] G. Nassreddine, A. E. Arid, M. Nassereddine, and O. A. Khatib, “Fault detection and classification for photovoltaic panel system using machine learning techniques,” Applied AI Letters, vol. 6, no. 2, Mar. 2025, doi: 10.1002/ail2.115. Available: https://doi.org/10.1002/ail2.115

[20] E. Garoudja, F. Harrou, Y. Sun, K. Kara, A. Chouder, and S. Silvestre, “Statistical fault detection in photovoltaic systems,” Solar Energy, vol. 150, pp. 485–499, May 2017, doi: 10.1016/j.solener.2017.04.043. Available: https://doi.org/10.1016/j.solener.2017.04.043

[21] Y. Gaaloul et al., “Faults detection and diagnosis of a Large-Scale PV system by analyzing power losses and electric indicators computed using random forest and KNN-Based prediction models,” Energies, vol. 18, no. 10, p. 2482, May 2025, doi: 10.3390/en18102482. Available: https://doi.org/10.3390/en18102482

[22] A. E. Lazzaretti et al., “A monitoring system for online fault detection and classification in photovoltaic plants,” Sensors, vol. 20, no. 17, p. 4688, Aug. 2020, doi: 10.3390/s20174688. Available: https://doi.org/10.3390/s20174688

[23] S. Yao, Q. Kang, M. Zhou, A. Abusorrah, and Y. Al-Turki, “Intelligent and Data-Driven fault detection of photovoltaic plants,” Processes, vol. 9, no. 10, p. 1711, Sep. 2021, doi: 10.3390/pr9101711. Available: https://doi.org/10.3390/pr9101711

[24] Monteiro, P., Lino, J., Araújo, R., & Costa, L. (2024). Comparison between LightGBM and other ML algorithms in PV fault classification. EAI Endorsed Trans. Energy Web, 11. https://doi.org/10.4108/ew.4865.

[25] S. Boubaker, S. Kamel, N. Ghazouani, and A. Mellit, “Assessment of machine and deep learning approaches for fault diagnosis in photovoltaic systems using infrared thermography,” Remote Sensing, vol. 15, no. 6, p. 1686, Mar. 2023, doi: 10.3390/rs15061686. Available: https://doi.org/10.3390/rs15061686

[26] C. Bu, T. Liu, T. Wang, H. Zhang, and S. Sfarra, “A CNN-Architecture-Based Photovoltaic cell fault classification method using thermographic images,” Energies, vol. 16, no. 9, p. 3749, Apr. 2023, doi: 10.3390/en16093749. Available: https://doi.org/10.3390/en16093749

[27] N. Drir, A. Mellit, and M. Bettayeb, “A novel ensemble CNN framework with weighted feature fusion for fault diagnosis of photovoltaic modules using thermography images,” IEEE Journal of Photovoltaics, vol. 15, no. 1, pp. 146–154, Nov. 2024, doi: 10.1109/jphotov.2024.3492283. Available: https://doi.org/10.1109/jphotov.2024.3492283

[28] F. M. A. Mazen, R. A. A. Seoud, and Y. O. Shaker, “Deep learning for automatic defect detection in PV modules using electroluminescence images,” IEEE Access, vol. 11, pp. 57783–57795, Jan. 2023, doi: 10.1109/access.2023.3284043. Available: https://doi.org/10.1109/access.2023.3284043

[29] B. Doll et al., “Aerial photoluminescence imaging of photovoltaic modules,” Physica Status Solidi (RRL) - Rapid Research Letters, vol. 17, no. 12, Apr. 2023, doi: 10.1002/pssr.202300059. Available: https://doi.org/10.1002/pssr.202300059

[30] K. Jaskie, J. Martin, and A. Spanias, “PV fault detection using positive unlabeled learning,” Applied Sciences, vol. 11, no. 12, p. 5599, Jun. 2021, doi: 10.3390/app11125599. Available: https://doi.org/10.3390/app11125599

[31] D. J. Benavides, P. Arévalo-Cordero, L. G. González, L. Hernández-Callejo, F. Jurado, and J. A. Aguado, “Method of monitoring and detection of failures in PV system based on machine learning,” Revista Facultad De Ingeniería Universidad De Antioquia, no. 102, pp. 26–43, Jul. 2020, doi: 10.17533/udea.redin.20200694. Available: https://doi.org/10.17533/udea.redin.20200694

[32] W. Chine, A. Mellit, V. Lughi, A. Malek, G. Sulligoi, and A. M. Pavan, “A novel fault diagnosis technique for photovoltaic systems based on artificial neural networks,” Renewable Energy, vol. 90, pp. 501–512, Jan. 2016, doi: 10.1016/j.renene.2016.01.036. Available: https://doi.org/10.1016/j.renene.2016.01.036

[33] B. Li, C. Delpha, D. Diallo, and A. Migan-Dubois, “Application of Artificial Neural Networks to photovoltaic fault detection and diagnosis: A review,” Renewable and Sustainable Energy Reviews, vol. 138, p. 110512, Nov. 2020, doi: 10.1016/j.rser.2020.110512. Available: https://doi.org/10.1016/j.rser.2020.110512

[34] L. C. Chen et al., “Fault diagnosis and classification for photovoltaic arrays based on principal component analysis and support vector machine,” IOP Conference Series Earth and Environmental Science, vol. 188, p. 012089, Oct. 2018, doi: 10.1088/1755-1315/188/1/012089. Available: https://doi.org/10.1088/1755-1315/188/1/012089

[35] Y. Zhong, B. Zhang, X. Ji, and J. Wu, “Fault diagnosis of PV array based on time series and support vector machine,” International Journal of Photoenergy, vol. 2024, pp. 1–10, Jan. 2024, doi: 10.1155/2024/2885545. Available: https://doi.org/10.1155/2024/2885545

[36] N. M. Castellà, N. K. Kampouropoulos, N. E. M. Urbano, and N. L. Romeral, “Supervision and fault detection system for photovoltaic installations based on classification algorithms,” Renewable Energy and Power Quality Journal, vol. 18, no. 4, Sep. 2020, doi: 10.24084/repqj18.337. Available: http://doi.org/10.24084/repqj18.337

[37] G. M. T. Tchio, J. Kenfack, J. Voufo, Y. A. Mindzie, B. F. Njoya, and S. S. Ouro-Djobo, “Diagnosing faults in a photovoltaic system using the Extra Trees ensemble algorithm,” AIMS Energy, vol. 12, no. 4, pp. 727–750, Jan. 2024, doi: 10.3934/energy.2024034. Available: https://doi.org/10.3934/energy.2024034

[38] K. Baradieh et al., “RSO based Optimization of Random Forest Classifier for Fault Detection and Classification in Photovoltaic Arrays,” The International Arab Journal of Information Technology, vol. 21, no. 4, Jan. 2024, doi: 10.34028/iajit/21/4/8. Available: https://doi.org/10.34028/iajit/21/4/8

[39] D. Adhya, S. Chatterjee, and A. K. Chakraborty, “Stacking ensemble based fault diagnosis approach for improved operation of photovoltaic arrays,” Energy Sources Part a Recovery Utilization and Environmental Effects, vol. 44, no. 2, pp. 5421–5439, Jun. 2022, doi: 10.1080/15567036.2022.2087804. Available: https://doi.org/10.1080/15567036.2022.2087804

[40] P. Monteiro, J. Lino, R. E. Araújo, and L. Costa, “Comparison between LightGBM and other ML algorithms in PV fault classification,” EAI Endorsed Transactions on Energy Web, vol. 11, Jan. 2024, doi: 10.4108/ew.4865. Available: https://doi.org/10.4108/ew.4865

[41] 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, pp. 126286–126306, Jan. 2021, doi: 10.1109/access.2021.3110947. Available: https://doi.org/10.1109/access.2021.3110947

[42] 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, p. 1012, Jan. 2024, doi: 10.3390/su16031012. Available: https://doi.org/10.3390/su16031012

[43] J. Zhang, X. Chen, H. Wei, and K. Zhang, “A lightweight network for photovoltaic cell defect detection in electroluminescence images based on neural architecture search and knowledge distillation,” arXiv (Cornell University), Feb. 2023, doi: 10.48550/arxiv.2302.07455. Available: http://arxiv.org/abs/2302.07455

[44] H. Ling, M. Liu, and Y. Fang, “Deep Edge-Based fault detection for solar panels,” Sensors, vol. 24, no. 16, p. 5348, Aug. 2024, doi: 10.3390/s24165348. Available: https://doi.org/10.3390/s24165348

[45] V. Veerasamy et al., “LSTM Recurrent Neural Network classifier for high impedance fault detection in solar PV integrated power system,” IEEE Access, vol. 9, pp. 32672–32687, Jan. 2021, doi: 10.1109/access.2021.3060800. Available: https://doi.org/10.1109/access.2021.3060800

[46] W. Gao, R.-J. Wai, and S.-Q. Chen, “Novel PV fault diagnoses via SAE and improved Multi-Grained Cascade Forest with string voltage and currents measures,” IEEE Access, vol. 8, pp. 133144–133160, Jan. 2020, doi: 10.1109/access.2020.3010233. Available: https://doi.org/10.1109/access.2020.3010233

[47] M. Alrifaey et al., “Hybrid Deep learning model for fault detection and classification of Grid-Connected photovoltaic System,” IEEE Access, vol. 10, pp. 13852–13869, Jan. 2022, doi: 10.1109/access.2022.3140287. Available: https://doi.org/10.1109/access.2022.3140287

[48] H. Abdulla, A. Sleptchenko, and A. Nayfeh, “Photovoltaic systems operation and maintenance: A review and future directions,” Renewable and Sustainable Energy Reviews, vol. 195, p. 114342, Feb. 2024, doi: 10.1016/j.rser.2024.114342. Available: https://doi.org/10.1016/j.rser.2024.114342

[49] A. M. Ahmed, L. Li, and K. Khalilpour, “Predictive Maintenance of Solar Photovoltaic Systems: A Comprehensive review,” IET Renewable Power Generation, vol. 19, no. 1, Jan. 2025, doi:

10.1049/rpg2.70152. Available: https://doi.org/10.1049/rpg2.70152

[50] Q. Liu, Q. Hu, J. Zhou, D. Yu, and H. Mo, “Remaining useful life prediction of PV systems under dynamic environmental conditions,” IEEE Journal of Photovoltaics, vol. 13, no. 4, pp. 590–602, May 2023, doi: 10.1109/jphotov.2023.3272071. Available: https://doi.org/10.1109/jphotov.2023.3272071

[51] A. Betti, M. L. Lo Trovato, F. S. Leonardi, G. Leotta, F. Ruffini, and C. Lanzetta, “Predictive Maintenance in Photovoltaic Plants with a Big Data Approach,” arXiv (Cornell University), Jan. 2019, doi: 10.48550/arxiv.1901.10855. Available: http://arxiv.org/abs/1901.10855

[52] J. Ramirez-Vergara, L. B. Bosman, E. Wollega, and W. D. Leon-Salas, “Review of forecasting methods to support photovoltaic predictive maintenance,” Cleaner Engineering and Technology, vol. 8, p. 100460, Mar. 2022, doi: 10.1016/j.clet.2022.100460. Available: https://doi.org/10.1016/j.clet.2022.100460

[53] A. Teta et al., “Enhanced fault diagnosis in Grid-Connected photovoltaic systems: leveraging transfer learning and ensemble methods for superior accuracy,” IEEE Access, vol. 12, pp. 194786–194803, Jan. 2024, doi: 10.1109/access.2024.3520490. Available: https://doi.org/10.1109/access.2024.3520490

[54] E. Lodhi et al., “A novel deep Stack-Based ensemble learning approach for fault detection and classification in photovoltaic arrays,” Remote Sensing, vol. 15, no. 5, p. 1277, Feb. 2023, doi: 10.3390/rs15051277. Available: https://doi.org/10.3390/rs15051277

[55] C. Utama, C. Meske, J. Schneider, R. Schlatmann, and C. Ulbrich, “Explainable artificial intelligence for photovoltaic fault detection: A comparison of instruments,” Solar Energy, 2023. doi: 10.1016/j.solener.2022.11.018.

[56] M. Bougoffa, S. Benmoussa, M. Djeziri, and O. Palais, “Hybrid deep learning for fault diagnosis in photovoltaic systems,” Machines, vol. 13, no. 5, p. 378, Apr. 2025, doi: 10.3390/machines13050378. Available: https://doi.org/10.3390/machines13050378

[57] K. Kull, M. A. Khan, B. Asad, M. U. Naseer, A. Kallaste, and T. Vaimann, “Adaptive deep belief networks and LightGBM-Based Hybrid fault diagnostics for SCADA-Managed PV Systems: a Real-World Case Study,” Electronics, vol. 14, no. 18, p. 3649, Sep. 2025, doi: 10.3390/electronics14183649. Available: https://doi.org/10.3390/electronics14183649

[58] G. Almonacid-Olleros, G. Almonacid, D. Gil, and J. Medina-Quero, “Evaluation of transfer learning and Fine-Tuning to nowcast energy generation of photovoltaic systems in different climates,” Sustainability, vol. 14, no. 5, p. 3092, Mar. 2022, doi: 10.3390/su14053092. Available: https://doi.org/10.3390/su14053092

[59] M. Yilmaz and A. A. Martinez-Morales, “Fault detection in photovoltaic panels using digital twin Technology: A comprehensive study,” The Journal of Cognitive Systems, vol. 8, no. 2, pp. 28–32, Dec. 2023, doi: 10.52876/jcs.1407133. Available: https://doi.org/10.52876/jcs.1407133