Review

Artificial Intelligence in Photovoltaic Fault Identification and Diagnosis: A Systematic Review

Mahmudul Islam 1, Masud Rana Rashel 2,*, Md Tofael Ahmed 2, A. K. M. Kamrul Islam 3 and Mouhaydine Tlemçani 2

1 Department of Computer Science and Engineering, Independent University, Dhaka 1229, Bangladesh; mahmud@iub.edu.bd
2 Instrumentation and Control Laboratory, Department of Mechatronics Engineering, University of Évora, 7000-671 Évora, Portugal; ahmed@uevora.pt (M.T.A.); tlem@uevora.pt (M.T.)
3 College of Engineering, North Carolina A&T State University, Greensboro, NC 27411, USA; akislam@ncat.edu
* Correspondence: mrashel@uevora.pt; Tel.: +351-920413407

Abstract: Photovoltaic (PV) fault detection is crucial because undetected PV faults can lead to significant energy losses, with some cases experiencing losses of up to 10%. The efficiency of PV systems depends upon the reliable detection and diagnosis of faults. The integration of Artificial Intelligence (AI) techniques has been a growing trend in addressing these issues. The goal of this systematic review is to offer a comprehensive overview of the recent advancements in AI-based methodologies for PV fault detection, consolidating the key findings from 31 research papers. An initial pool of 142 papers were identified, from which 31 were selected for in-depth review following the PRISMA guidelines. The title, objective, methods, and findings of each paper were analyzed, with a focus on machine learning (ML) and deep learning (DL) approaches. ML and DL are particularly suitable for PV fault detection because of their capacity to process and analyze large amounts of data to identify complex patterns and anomalies. This study identified several AI techniques used for fault detection in PV systems, ranging from classical ML methods like k-nearest neighbor (KNN) and random forest to more advanced deep learning models such as Convolutional Neural Networks (CNNs). Quantum circuits and infrared imagery were also explored as potential solutions. The analysis found that DL models, in general, outperformed traditional ML models in accuracy and efficiency. This study shows that AI methodologies have evolved and been increasingly applied in PV fault detection. The integration of AI in PV fault detection offers high accuracy and effectiveness. After reviewing these studies, we proposed an Artificial Neural Network (ANN)-based method for PV fault detection and classification.

Keywords: photovoltaic fault; Artificial Intelligence; machine learning; deep learning; Artificial Neural Network; Convolutional Neural Network; Recurrent Neural Network; computer vision; unmanned aerial vehicles; systematic review

1. Introduction

The global transition to sustainable energy has positioned photovoltaic (PV) systems at the top of renewable energy solutions. Robust fault detection and diagnosis systems are crucial to ensure the efficiency and longevity of PV systems. Historically, fault detection in PV systems was dependent on manual inspections and traditional electrical measurements [1]. However, with the vast arrays of panels installed, especially in large solar farms, this method proved to be inefficient, labor-intensive, and occasionally inaccurate. With the rapid progression in ML and DL techniques, researchers are exploring various computational methodologies to effectively identify and classify PV faults [2]. In recent times, many AI techniques have been developed for fault detection and diagnosis in PV Systems, but it is still unclear which one is the best. A significant challenge is the lack of
public datasets for PV fault detection [3]. This limits the potential of AI techniques in this field. There are some methods that are precise, but they are either too slow or too complex for real-world use [4]. Furthermore, there seems to be an imbalance in research attention, such as specific PV faults being extensively studied while others are overlooked [5].

To overcome these gaps and provide a holistic perspective, we were motivated to conduct this study. This study aims to provide a comprehensive overview of recent advancements in PV fault detection and diagnosis using Artificial Intelligence (AI). Initially, 142 papers were selected for review. However, after a rigorous filtering process adhering to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses statement (PRISMA) guidelines [6], 31 papers were retained for detailed examination. We found that many studies used I-V curves alongside ML for fault identification in PV arrays [7]. Techniques like Principal Component Analysis (PCA) [8] were commonly used for feature extraction, while complex PV systems adopted advanced methods, such as Recurrent Neural Networks (RNNs) [9] with satellite data and Convolutional Neural Networks (CNNs) [10], for analyzing voltage and current. Hybrid models, integrating conventional algorithms with deep learning techniques, gained better results [11]. Ensemble learning [12] and stacking classifiers offered refined diagnosis capabilities, while semi-supervised learning [13] methods showed promising results with limited data. Meanwhile, some studies introduced quantum circuits as a way to use Neural Networks, while others found that using infrared images, especially with methods like DeepLabV3+ and U-Net, was very helpful for identifying PV faults.

This study delivers a comprehensive analysis of PV fault detection and diagnosis using AI, aggregating insights from 31 research studies. This study also serves as a benchmark by providing a comparative evaluation of diverse AI techniques, from traditional methods to cutting-edge approaches like quantum circuits. We examined many studies and chose relevant ones from the 2021–2023 timeframe. In this study, we examined the following research questions.

RQ1: Which fault detection method is the most accurate and quickest for PV systems?
RQ2: How do different deep learning models perform in detecting PV faults?
RQ3: How effective is image-based fault detection compared to traditional methods?
RQ4: How does the quantity and quality of labeled data impact the accuracy of PV fault detection?
RQ5: How do weather conditions affect PV faults?
RQ6: How can we better differentiate and classify specific types of PV faults using machine learning?

A comparative analysis of this study with the existing review studies is presented in Table 1.

<table>
<thead>
<tr>
<th>Study</th>
<th>Year</th>
<th>PV Fault Detection</th>
<th>Image-Based Approach</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Traditional ML Approach</td>
<td>DL Approach</td>
</tr>
<tr>
<td>[14]</td>
<td>2021</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>[16]</td>
<td>2021</td>
<td>—</td>
<td>✓</td>
</tr>
<tr>
<td>[17]</td>
<td>2022</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>[18]</td>
<td>2023</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>[19]</td>
<td>2023</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>This study</td>
<td>2023</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>
The following are the contributions of this research study:
1. The identification of the latest trends in PV fault detection using AI techniques.
2. The identification of state-of-the-art AI techniques for identifying PV faults.
3. An AI-based technique is proposed to detect and classify PV faults.

In the following sections, the content of this article is organized as follows. Section 2 provides a comprehensive background on AI, and Section 3 describes the details of PV faults. In Section 4, the methodologies employed in this systematic review are explained, and Section 5 presents the findings of this study. Section 6 presents the proposed methodology, and finally, in Section 7, our conclusions are presented.

2. Artificial Intelligence

Using AI in the field of green energy, especially PV systems, has opened new opportunities for identifying and resolving issues. Data-driven decision-making systems can assist in efficient energy management [20]. This section highlights the different types of AI and their roles in PV fault detection.

2.1. Types of Artificial Intelligence

Machine Learning (ML): ML allows computers to learn from past data [21]. In PV systems, ML can analyze past performance data to predict and detect PV faults and abnormalities.

Deep Learning: A subset of ML, deep learning uses complex structures called Neural Networks [22]. It is especially useful in analyzing images of PV panels using Convolutional Neural Networks (CNNs) to find defects or performance issues.

Computer Vision: Computer Vision (CV) is about making computers interpret and act on visual data [23]. With PV systems, it can process images from drones or satellites to identify PV faults.

Machine Vision: Machine Vision is like computer vision, but it is mainly used in manufacturing [24]. For PV panels, it can ensure quality control during production.

Natural Language Processing (NLP): NLP is not related to PV fault detection, but it has the capacity to analyze textual data [25]. Therefore, NLP can be used to maintain reports, logs, and other textual information related to PV panels.

2.2. The Role of Artificial Intelligence in PV Systems

Image Analysis: Drones take high-quality pictures of PV farms. With deep learning, these pictures show small cracks or mismatched panels [26].

Predictive Maintenance: ML predicts possible faults before they happen by understanding past performance and current conditions. This allows for timely intervention.

Anomaly Detection: Modern PV systems are always monitored. AI-based technology helps these systems quickly find and point out any unusual changes, which ensures that no fault goes unnoticed [27].

Optimization: By analyzing environmental data, AI-based technology can suggest optimal operating conditions as optimal panel angles for PV panels. AI-based technology can also suggest cleaning schedules and energy storage strategies based on real-time data.

As AI technologies become more sophisticated, their capability to detect and even prevent faults in PV systems will grow. We can expect more automated solutions and a higher degree of reliability in PV installations.

3. Photovoltaic Faults

PV panels are an essential source of green energy, but they can face various types of issues that can degrade their performance. To obtain good performance from PV systems, it is essential to understand potential faults and strategies to mitigate them. This section discusses the types of PV faults, their causes, and potential mitigation techniques.
3.1. Types of Photovoltaic Faults

PV systems can encounter different types of faults that can negatively impact their performance. Some of the most common PV faults are described below.

Module Mismatching: This is usually caused by uneven aging or differences in manufacturing between modules. This can lead to reduced efficiency and hot spots [28].

Micro-Cracks: These often result from mechanical stress and can reduce the module’s performance over time [29].

Hot Spots: These are areas of localized heating in the PV module due to high resistance. They may arise from shading, dirt, or uneven aging [30].

Shadowing: Objects like trees or nearby buildings can block sunlight and shade panels. This causes a decrease in their efficiency [31].

Degradation: Over time, PV modules can degrade, which leads to a gradual drop in power output.

3.2. Causes of PV Faults

Module aging: Over time, modules experience natural degradation, which can affect their efficiency and output [32].

Manufacturing inconsistencies: Differences in the production process can lead to slight differences in module quality, which causes performance variations [33].

Temperature fluctuations: Rapid changes in temperature can make materials expand or shrink, which might damage the modules [34].

UV exposure: Continuous exposure to UV rays can deteriorate the protective layers of PV modules, which reduces their efficiency and lifespan [35].

Weather-related factors: Factors like hailstorms, snow, or persistent rain can directly or indirectly lead to faults such as micro-cracks or other structural faults in the PV system.

During Installation: Rough handling or physical pressure during the installation can cause defects in the PV panels.

3.3. Mitigating PV Faults

Efficient fault detection is the first step toward mitigation. Some strategies are discussed in this section to mitigate PV faults.

Image-Based Approaches: Aerial images, especially from Unmanned Aerial Vehicles (UAVs), combined with deep learning techniques like CNNs, have proven effective in detecting faults. Infrared and thermographic images can precisely locate hot spots and micro-cracks [36].

Machine Learning and Deep Learning: Methods like CNNs, RNNs, and quantum circuits have been used to analyze data and spot inconsistencies in PV outputs. These methods can be used effectively to indicate potential faults.

Routine Maintenance: Regularly cleaning PV panels to remove obstructions like dust and debris ensures consistent sunlight absorption. Periodic inspections help to identify defects or connection issues early and preserve the system’s efficiency and lifespan.

Optimal Panel Positioning: To maximize energy production, PV panels should be placed in locations with minimal shadowing from obstructions like trees or buildings. They should be angled and oriented to capture consistent sunlight throughout the day. Adjustments based on geographical and weather factors can enhance their efficiency [37].

To maintain the efficiency and longevity of PV systems, understanding potential faults and proactively addressing them is essential. With technologies like machine learning and regular maintenance, PV system issues can be managed effectively.

4. Systematic Review

4.1. Eligibility Criteria

To conduct a comprehensive and pertinent analysis of the existing literature on PV fault detection, we established a set of eligibility criteria. Papers had to focus on fault detection and diagnosis in PV systems, specifically discussing methodologies, algorithms,
and techniques for identifying, classifying, and diagnosing faults. In terms of methodology, we selected studies related to ML, DL, image-based approaches, and other AI techniques. We also considered research showing advancements in fault classification, especially those emphasizing closely related fault types of differentiation, like module mismatching and micro-cracks. To ensure the relevance of this review study, we restricted our focus to papers published from 2021 to 2023. The papers were extracted from different databases, such as ScienceDirect, MDPI, IEEE Xplore, Springer, and Wiley Online Library. Only peer-reviewed journal articles and conference papers were considered, and non-English research articles were excluded. Table 2 depicts the criteria for inclusion and exclusion that were used in this study.

Table 2. Criteria for inclusion and exclusion.

<table>
<thead>
<tr>
<th>Area</th>
<th>Inclusion Criteria</th>
<th>Exclusion Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>Search keyword</td>
<td>Photovoltaic, fault detection, machine learning, deep learning, machine vision</td>
<td>Hydroelectric, performance ratio analysis, electroluminescence imaging,</td>
</tr>
<tr>
<td>Publication type</td>
<td>Research articles</td>
<td>Editorial articles, dissertation articles, books</td>
</tr>
<tr>
<td>Area of interest</td>
<td>Photovoltaic fault detection, photovoltaic fault diagnosis, applied machine learning</td>
<td>Areas not in the inclusion criteria</td>
</tr>
<tr>
<td>Selected duration</td>
<td>2021–2023</td>
<td>Before 2021</td>
</tr>
<tr>
<td>Language</td>
<td>English</td>
<td>Non-English languages</td>
</tr>
</tbody>
</table>

4.2. Search Strategy

To thoroughly study fault detection in PV panels, we carefully created a detailed search string. This search string was crucial for ensuring thorough coverage of the targeted research domain. During the creation of the search string, we considered three aspects, the technology (‘Photovoltaic’), the goal (words such as ‘fault detection’), and common methods (like ‘machine learning’ or ‘deep learning’). From these areas, we formed the following search string:

\[
[\text{"Photovoltaic" OR "PV"} \text{ AND ("fault detection" OR "fault diagnosis" OR "fault classification" OR "fault identification" OR "fault analysis") AND ("machine learning" OR "deep learning" OR "neural networks" OR "machine vision" OR "computer vision" OR "image processing") NOT ("performance ratio analysis" OR "electroluminescence imaging")}]
\]

The above formulated search string helped us to find a comprehensive set of articles aligning with our review objectives. Each search string was customized to fit the specific syntax and capabilities of different research databases, such as Google Scholar, MDPI, Science Direct, and IEEE Xplore.

4.3. Data Extraction and Analysis

We followed the PRISMA guidelines for paper selection in this review study because it provides a standardized framework for reporting and assessing the quality of selected studies for review. From 142 initially identified articles, we narrowed down our selection to 71 through screening. Eventually, 31 research articles were selected for the review. The PRISMA flow diagram for this study is presented in Figure 1.

The research articles we reviewed were from 2021 to 2023, with 12 articles from 2021, 14 articles from 2022, and 6 articles from 2023. Figure 2 presents the number of selected articles by publication year. These research articles came from various publishers. In total, 8 articles came from Science Direct, 10 articles each from MDPI and IEEE, 1 article from Springer, and 3 articles from Wiley.
4.3. Data Extraction and Analysis

We followed the PRISMA guidelines for paper selection in this review study because it provides a standardized framework for reporting and assessing the quality of selected studies for review. From 142 initially identified articles, we narrowed down our selection to 71 through screening. Eventually, 31 research articles were selected for the review. The PRISMA flow diagram for this study is presented in Figure 1.

**Figure 1.** PRISMA diagram illustrating the process of the literature search. * denotes different databases and ** denotes exclusion of articles. 30 articles were excluded manually by human.

Figure 2 presents the number of selected articles by publication year. These research articles came from various publishers. In total, 8 articles came from Science Direct, 10 articles each from MDPI and IEEE, 1 article from Springer, and 3 articles from Wiley.

<table>
<thead>
<tr>
<th>Publisher</th>
<th># of Articles Selected at Initial Stage</th>
<th># of Articles Selected at the Screening Stage</th>
<th># of Articles Selected at the Inclusion Stage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Science Direct</td>
<td>44</td>
<td>21</td>
<td>8</td>
</tr>
<tr>
<td>MDPI</td>
<td>23</td>
<td>15</td>
<td>10</td>
</tr>
<tr>
<td>IEEE</td>
<td>57</td>
<td>26</td>
<td>10</td>
</tr>
<tr>
<td>Springer</td>
<td>5</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>Wiley</td>
<td>13</td>
<td>6</td>
<td>3</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>142</strong></td>
<td><strong>71</strong></td>
<td><strong>31</strong></td>
</tr>
</tbody>
</table>

We identified common trends, such as the frequent use of deep learning in fault detection. We also generated a geographic map showing where most of the research came from. Figure 3 depicts geographic mapping around the world for the 31 selected articles reviewed in our study. The map uses shades of red to represent research volume, whereby darker red means more research. China and the USA stood out as major contributors. By inspecting the details of each paper, we ensured our findings were consistent and accurate.

Table 3 presents articles selected from various digital libraries at different phases. For each research article, we noted essential details like the title, publication year, objectives, methods, and main findings, which are described in a later section. Our analysis combined both numbers and themes.
Table 3. Studies chosen from various digital libraries at different phases.

<table>
<thead>
<tr>
<th>Publisher</th>
<th># of Articles Selected at Initial Stage</th>
<th># of Articles Selected at the Screening Stage</th>
<th># of Articles Selected at the Inclusion Stage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Science Direct</td>
<td>44</td>
<td>21</td>
<td>8</td>
</tr>
<tr>
<td>MDPI</td>
<td>23</td>
<td>15</td>
<td>10</td>
</tr>
<tr>
<td>IEEE</td>
<td>57</td>
<td>26</td>
<td>10</td>
</tr>
<tr>
<td>Springer</td>
<td>5</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>Wiley</td>
<td>13</td>
<td>6</td>
<td>3</td>
</tr>
<tr>
<td>Total</td>
<td>142</td>
<td>71</td>
<td>31</td>
</tr>
</tbody>
</table>

We identified common trends, such as the frequent use of deep learning in fault detection. We also generated a geographic map showing where most of the research came from. Figure 3 depicts geographic mapping around the world for the 31 selected articles reviewed in our study. The map uses shades of red to represent research volume, whereby darker red means more research. China and the USA stood out as major contributors. By inspecting the details of each paper, we ensured our findings were consistent and accurate.

Figure 3. Geographic mapping of 31 selected articles for our review study.

5. Results

5.1. Comparison of Techniques

RQ1: Which Fault Detection Method Is the Most Accurate and Quickest for PV Systems?

The accuracy of fault detection methods varies. In terms of accuracy, the AdaBoost Ensemble model used in one study achieved a detection accuracy of 97.84% [38]. Convolutional Neural Networks (CNNs) are frequently adopted due to their high accuracy (100%), with several studies applying them for detecting a range of PV panel faults [39–41]. Additionally, hybrid models such as ensemble learning and stacking classifiers also reported significant accuracy enhancements [42,43]. Regarding speed and efficiency, the CNN approach in one study was particularly emphasized for its rapid fault identification [41]. Additionally, the automated inspection of PV panels using Unmanned Aerial Vehicles (UAVs) indicates a faster approach in certain studies [44,45]. Another paper highlighted...
the efficiency of CNNs for automating feature extraction, potentially accelerating the fault detection process [46].

5.2. Efficiency of Deep Learning Models
RQ2: How Do Different Deep Learning Models Perform in Detecting PV Faults?

Deep learning techniques are widely used for PV fault detection. CNNs are prominent in this domain. CNNs have been used to identify PV panel failure types using parameters like voltage and current [43], as well as for classifying multiple PV panel faults through infrared images [47]. An RNN-based method has also been adopted, capitalizing on satellite weather data for precise fault diagnosis [48]. Advanced models like DeepLabV3+, U-Net, and the Feature Pyramid Network (FPN) are increasingly combined for enhanced fault detection results [49,50]. Specific techniques, such as Ghost convolution integrated with You Only Look Once (YOLOv5), have also been introduced for PV fault detection [51]. Moreover, the Feature-Induced Augmentation (FIA) method showed improved results in identifying micro-cracks on PV surfaces [52]. Overall, while CNNs dominate PV fault detection [40,46,47], the selection of a model largely depends on the specific fault type and data available.

5.3. Image-Based Analysis Efficiency
RQ3: How Effective Is Image-Based Fault Detection Compared to Traditional Methods?

Image-based fault detection techniques in PV systems have seen a significant rise, especially with the advancements in computer vision and deep learning. One method that particularly stands out is the use of Convolutional Neural Networks (CNNs) to detect and classify PV panel faults via infrared images [47]. Further exploring the image-based techniques, the utilization of thermographic images taken by Unmanned Aerial Vehicles (UAVs) has proven beneficial in inspecting and classifying PV faults [44]. Moreover, deep learning-based photovoltaic fault detection techniques that particularly employ images are covered in [49]. There is also a focus on capturing images from UAVs to detect PV faults and classify them [45]. The automation of identifying visual faults in PV modules has been effectively addressed [53]. Traditional methods for fault detection often depend on electrical parameters. These include methodologies like analyzing I-V curves and extracting multivariate features using PCA, as mentioned in [54,55]. Image-based techniques are proving effective in identifying a wide array of faults when integrated with deep learning. Image-based techniques offer detailed spatial insights and outperform traditional methods in efficiency and depth of information. The choice between traditional and image-based techniques depends on specific contextual requirements, such as the nature of the fault and available resources.

5.4. Role of Data
RQ4: How Does the Quantity and Quality of Labeled Data Impact the Accuracy of PV Fault Detection?

The accuracy of PV fault detection is influenced by the quantity and quality of labeled data. In [56], a semi-supervised learning method was used that effectively detects solar faults, even when lesser labeled data were available, as compared to conventional techniques. This emphasizes the importance of labeled data volumes in achieving reliable fault detection results. Conversely, the accuracy and reliability of fault detection can be compromised if the data quality is poor, such as if data are improperly labeled or contain many outliers. It is a principle in machine learning that well-curated and accurately labeled data result in improved outcomes. The quantity of labeled data is crucial for developing robust models, while the quality of data is equally important to ensure accurate PV fault detection.
5.5. Impact of Environmental Factors
RQ5: How Do Weather Conditions Affect PV Faults?

Environmental factors play an important role in the performance and potential faults of PV panels. Satellite weather data, as utilized in [48], highlight the importance of environmental conditions in diagnosing PV system faults. Hot spots on PV panels, mentioned in [57], emphasize the sensitivity of PV systems to factors like shadowing, debris, and equipment issues, which can intensify under certain weather conditions. To account for these environmental variables, ML models can integrate parameters like temperature, sunlight intensity, and humidity. Data augmentation and hybrid modeling are also beneficial. These allow models to better predict how changing environmental conditions can impact PV system performance. Constantly updating these models to reflect real-world feedback is essential for their accuracy. While environmental factors significantly affect PV panels, machine learning models, when optimized using the right strategies, can effectively detect and predict associated faults.

5.6. Improving Fault Classification
RQ6: How Can We Better Differentiate and Classify Specific Types of PV Faults Using Machine Learning?

Fault classification in PV systems has evolved with the adoption of ML and DL models. Specific techniques, like the CNNs used in [40,47,49,58,59], have been particularly effective in leveraging visual data to identify and categorize diverse fault types. The differentiation between closely related faults, such as the module mismatching and partial shadowing mentioned in [60], remains challenging. However, combined methodologies like the hybrid system mentioned in [61] show potential in refining this differentiation. The success of fault detection is also related to the quality and quantity of data, as indicated by the semi-supervised approach in [56]. Combining various methodologies and leveraging quality data are crucial for enhancing accuracy in PV fault classification. Table 4 presents a summary of the selected studies.

Table 4. Summary of the selected papers.

<table>
<thead>
<tr>
<th>Paper Id</th>
<th>Year</th>
<th>Objective</th>
<th>Method</th>
<th>Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>P1 [54]</td>
<td>2021</td>
<td>To diagnose faults in PV panels.</td>
<td>The presented method utilizes I-V curves with ML techniques for fault detection in a PV array across eight scenarios.</td>
<td>The classifiers, calibrated using simulated samples, were verified using field-measured I-V curves. ANN achieved 100% accuracy in both simulated and real-world tests.</td>
</tr>
<tr>
<td>P2 [55]</td>
<td>2021</td>
<td>To develop an improved method for fault detection.</td>
<td>A fault detection and diagnosis (FDD) framework was proposed where PCA was used to extract the most pertinent multivariate features.</td>
<td>The proposed FDD framework was evaluated using various metrics, and random forest (RF) achieved 99.64% accuracy.</td>
</tr>
<tr>
<td>P3 [62]</td>
<td>2022</td>
<td>To develop an affordable approach for identifying different types of PV faults.</td>
<td>An incremental approach was followed to determine the best ML model that can detect faults using statistical testing.</td>
<td>In the original dataset, most models had 100% accuracy. But when data points were removed, their accuracy dropped, showing their vulnerability. However, Naïve Bayes (NB) maintained an accuracy of 94% to 100% on different datasets, which shows the reliability of NB.</td>
</tr>
<tr>
<td>P4 [39]</td>
<td>2022</td>
<td>To diagnose faults in PV systems.</td>
<td>An RNN-based approach using satellite weather data and inverter measurements.</td>
<td>The RNN-based approach identifies faults, leading to a minimum 5% power drop. Beyond classifying faults, the model estimates their severity and helps in the maintenance and detection of unknown faults.</td>
</tr>
<tr>
<td>Paper Id</td>
<td>Year</td>
<td>Objective</td>
<td>Method</td>
<td>Findings</td>
</tr>
<tr>
<td>----------</td>
<td>------</td>
<td>-----------</td>
<td>--------</td>
<td>----------</td>
</tr>
<tr>
<td>P5 [40]</td>
<td>2022</td>
<td>To identify faults in PV panels.</td>
<td>A CNN-based approach that considers different parameters such as voltage, current, etc., to identify failure types in PV panels.</td>
<td>The proposed CNN model outperformed previous models, achieving 95.20% accuracy.</td>
</tr>
<tr>
<td>P6 [63]</td>
<td>2022</td>
<td>To evaluate the resilience and efficiency of two control strategies against flying capacitor faults.</td>
<td>A k-nearest neighbor (KNN)-based approach was used.</td>
<td>Sliding mode control outperforms exact linearization control in response time, accuracy, and flying capacitor voltage fluctuations.</td>
</tr>
<tr>
<td>P7 [42]</td>
<td>2023</td>
<td>To diagnose faults in PV strings.</td>
<td>A Multi-Layer Stacking Classifier (MLSC) was proposed that merges features from different ML algorithms.</td>
<td>Light intensity affects PV short-circuit current, temperature affects open-circuit voltage. GridSearchCV set parameters in MLSC that optimize fault diagnosis and save time.</td>
</tr>
<tr>
<td>P8 [61]</td>
<td>2022</td>
<td>To detect faults in grid-linked PV systems.</td>
<td>A unique fault identification method was introduced using statistical signatures of PV operational states, as each fault distinctly affects the electrical system.</td>
<td>The random forest (RF) classifier, using the given signatures, identified all fault types.</td>
</tr>
<tr>
<td>P9 [64]</td>
<td>2021</td>
<td>To identify faulty PV modules.</td>
<td>A hybrid system was presented using three learning methods to accurately detect faulty PV modules.</td>
<td>The authors introduced three methods for solar module detection, where the first one combines enhanced gamma correction with a CNN; the second one uses a CNN with threshold preprocessing on IR temperatures; and the third one employs XGBoost with temperature statistics. All are efficient, with the hybrid approach being the most accurate.</td>
</tr>
<tr>
<td>P10 [56]</td>
<td>2021</td>
<td>To identify faults in PV arrays.</td>
<td>The authors presented an approach using a semi-supervised ML method. Positive unlabeled learning can efficiently detect solar faults with much fewer labeled data than conventional methods.</td>
<td>The authors designed a solar fault model with positive unlabeled learning that outperforms supervised classifiers using just 5% labeled data.</td>
</tr>
<tr>
<td>P11 [48]</td>
<td>2021</td>
<td>To identify faults in PV arrays.</td>
<td>The authors introduced a two-qubit quantum circuit as a Neural Network solution for PV fault detection.</td>
<td>The initial results from two-qubit implementation showed moderate fault detection compared to classical computation. Adding more qubits did not enhance accuracy, likely due to increased quantum noise in the simulation.</td>
</tr>
<tr>
<td>P12 [43]</td>
<td>2021</td>
<td>To diagnose faults in PV strings.</td>
<td>This study uses ensemble learning (EL) for PV system fault diagnosis, selecting features via a grid-search and optimizing the combined model for better accuracy.</td>
<td>The proposed EL method outperforms ML algorithms in PV system fault diagnosis, offering better classification and generalization. This system efficiently detects faults in small-scale PV systems with high accuracy and low cost.</td>
</tr>
<tr>
<td>P13 [58]</td>
<td>2021</td>
<td>To identify and categorize faults in PV arrays.</td>
<td>A Neural Network (NN)-based method was proposed for fault detection, using an auto-encoder and refining with concrete dropout.</td>
<td>Concrete dropout surpasses other methods with 89.87% accuracy. A 50% pruned network reduces accuracy by 3%, streamlining PV array fault detection tools.</td>
</tr>
</tbody>
</table>
Table 4. Cont.

<table>
<thead>
<tr>
<th>Paper Id</th>
<th>Year</th>
<th>Objective</th>
<th>Method</th>
<th>Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>P14 [47]</td>
<td>2021</td>
<td>To classify faults in PV modules.</td>
<td>A CNN-based approach was proposed to detect and classify PV faults using infrared images.</td>
<td>The model detected 92% of healthy and 93% of damaged modules using CNNs. Using oversampling with augmentation, the proposed CNN’s accuracy improved by 6% compared to under-sampling, enhancing its classification of PV degradation in IR images.</td>
</tr>
<tr>
<td>P15 [46]</td>
<td>2021</td>
<td>To classify faults in PV systems.</td>
<td>A meta-heuristic algorithm optimizes five PV model parameters. A new CNN method is introduced for fault classification, automating feature extraction and improving efficiency.</td>
<td>The CNN model achieved around 98.3% and 98.9% accuracy in simulations, and 96.76% and 97.41% in experiments. It also efficiently handles quick changes in PV system.</td>
</tr>
<tr>
<td>P16 [60]</td>
<td>2022</td>
<td>To identify partial shadowing and mismatch issues in PV arrays.</td>
<td>This article reveals that select points near peak power can identify module mismatching, eliminating the need for a GMPPT algorithm. Curvature changes are detected using techniques like decision trees and support vector machines.</td>
<td>SVM and one-layer multilayer perceptron perform better than other methods. However, the perceptron predicts faster than the support vector machine.</td>
</tr>
<tr>
<td>P17 [57]</td>
<td>2021</td>
<td>To identify faults and hot spots in PV panels.</td>
<td>Different deep learning (DL) approaches are used for PV fault and hot spot detection.</td>
<td>DL is successfully used in identifying faults across various electrical applications, with a notable emphasis on detecting issues within photovoltaic (PV) systems, such as the identification of hot spots.</td>
</tr>
<tr>
<td>P18 [38]</td>
<td>2022</td>
<td>To detect and classify PV faults.</td>
<td>An AdaBoost Ensemble model (AEM) was proposed.</td>
<td>AEM includes different weak base learners, stacked sequences that assist in learning from failures of previous weak learners, and develop a new improved model to identify and classify faults. The proposed AEM achieved 97.84% accuracy.</td>
</tr>
<tr>
<td>P19 [65]</td>
<td>2022</td>
<td>To identify faults in PV systems with better accuracy and less computational time.</td>
<td>In this study, RF and modified independent component analysis (MICA) are used.</td>
<td>This study deals with intermediate and maximum power point tracking. The proposed RF-MICA technique identified faults with an accuracy of 99.88% and 99.43% for two different scenarios, respectively.</td>
</tr>
<tr>
<td>P20 [44]</td>
<td>2021</td>
<td>To inspect PV panels automatically through Unmanned Aerial Vehicles.</td>
<td>This article inspects PV panels using thermographic images with the aid of UAVs. This study classifies ten common faults.</td>
<td>A computer vision tool was developed for semi-automatic processing that is based on thermographic videos taken from UAVs. It can detect ten common anomalies related to common PV faults.</td>
</tr>
<tr>
<td>P21 [66]</td>
<td>2022</td>
<td>To identify faults in PV systems.</td>
<td>Three different models are used in this study. They are DeepLabV3+, U-Net, and the Feature Pyramid Network (FPN). All three have different encoder architectures.</td>
<td>The three different proposed models identified defective panels from a large-scale solar plant using semantic segmentation. The U-Net stood out as the best model with an accuracy of 94% among them.</td>
</tr>
</tbody>
</table>
### Table 4. Cont.

<table>
<thead>
<tr>
<th>Paper Id</th>
<th>Year</th>
<th>Objective</th>
<th>Method</th>
<th>Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>P22 [49]</td>
<td>2021</td>
<td>To identify and classify faults in real time.</td>
<td>A hybrid DL model was proposed in this article.</td>
<td>This study uses wavelet packet transform for processing the data. The proposed DL architecture consists of an equilibrium optimizer algorithm and long short-term memory. Automatic feature extraction is improved using the proposed hybrid model.</td>
</tr>
<tr>
<td>P23 [41]</td>
<td>2021</td>
<td>To automatically identify faults in PV systems using thermographic images.</td>
<td>This article uses a CNN-based approach. The approach recognizes defects in PV modules with high accuracy.</td>
<td>From a dataset of 1000 images, the CNN achieved 99% accuracy. When tested on a smaller dataset of 200 images, the accuracy was 90%.</td>
</tr>
<tr>
<td>P24 [59]</td>
<td>2022</td>
<td>To identify and classify PV faults.</td>
<td>A multi-scale CNN was used in this study.</td>
<td>The proposed multi-scale CNN classified 11 different types of anomaly and the average accuracy was 97.32%.</td>
</tr>
<tr>
<td>P25 [67]</td>
<td>2022</td>
<td>To identify defects in large PV plants using visible and infrared images.</td>
<td>This study proposed a framework that incorporates image acquisition, segmentation, and fault orientation and provides warnings for PV defects.</td>
<td>The fifth version of You, YOLOv5, and ResNet are used in this study. The proposed framework has a strong capability to work under different brightness conditions and achieves 95% accuracy using infrared images.</td>
</tr>
<tr>
<td>P26 [45]</td>
<td>2023</td>
<td>To automatically detect PV faults using images that are taken from UAVs.</td>
<td>This article uses a technique named PV-YOLO. The CBAM attention technique is used for enhancing the effective features. It also classifies them into six groups of faults.</td>
<td>PV-YOLO and CBAM improved the performance, and the detecting accuracy was 92.5 percent. The proposed system can identify small objects.</td>
</tr>
<tr>
<td>P27 [53]</td>
<td>2023</td>
<td>To automatically identify and classify faults in PV modules.</td>
<td>DenseNet-201 was used for feature extraction. The most significant feature was selected using a decision tree algorithm (J48).</td>
<td>The combination of WiSARD and DenseNet 201 helped to achieve 100% accuracy in PV fault classification.</td>
</tr>
<tr>
<td>P28 [68]</td>
<td>2023</td>
<td>To identify faulty surfaces of PV panels.</td>
<td>Ghost convolution, BottleneckCSP, and YOLOv5 were used for fault detection.</td>
<td>The proposed method increased the accuracy of fault detection up to 27.8% compared to the existing methods. The highest attained mAp was 97.8%.</td>
</tr>
<tr>
<td>P29 [50]</td>
<td>2022</td>
<td>To identify micro-cracks in PV modules during the manufacturing period.</td>
<td>Feature-Induced Augmentation (FIA) shows improved results in identifying micro-cracks over a PV surface.</td>
<td>The used PV-CrackNet had 7.01 million learnable parameters and it achieved 97% accuracy during the test.</td>
</tr>
<tr>
<td>P30 [51]</td>
<td>2023</td>
<td>To identify faults in PV systems.</td>
<td>Backpropagation Neural Network (BPNN-PSO) and heuristic particle swarm optimization techniques were used.</td>
<td>The proposed method identifies faults in a PV system with an accuracy of 95%. BPNN-PSO conversion occurs after 250 steps.</td>
</tr>
<tr>
<td>P31 [52]</td>
<td>2023</td>
<td>To identify faults in PV systems using image processing techniques.</td>
<td>Deep learning techniques were used in this study.</td>
<td>The DeepCNN technique obtained the best accuracy (98.7%) in PV fault identification.</td>
</tr>
</tbody>
</table>

For PV fault detection, image-based methods using Unmanned Aerial Vehicles (UAVs) and deep learning models like U-Net, DeepLabV3+, and FPN have emerged as a significant trend. One of the main findings is that DL models, especially CNNs, have shown promising results and outperformed traditional techniques in accurately detecting PV faults. The role of data, both in terms of its quality and quantity, is crucial, though a semi-supervised ML approach has tackled the challenge of limited data. Environmental variables significantly influence PV performance. Table 5 presents the best accuracy obtained by algorithms or
models in different studies. It shows that DL models or algorithms achieved better accuracy compared to traditional ML algorithms or models.

<table>
<thead>
<tr>
<th>Algorithms/Models</th>
<th>Studies</th>
<th># of Studies</th>
<th>Obtained Best Performance in Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>NB [62]</td>
<td>1</td>
<td>94–100%</td>
<td></td>
</tr>
<tr>
<td>AEM [38]</td>
<td>1</td>
<td>97.84%</td>
<td></td>
</tr>
<tr>
<td>RF-MICA [65]</td>
<td>1</td>
<td>99.88%</td>
<td></td>
</tr>
<tr>
<td>RF [55,61]</td>
<td>2</td>
<td>99.64%</td>
<td></td>
</tr>
<tr>
<td>ANN [51,54,58]</td>
<td>3</td>
<td>100%</td>
<td></td>
</tr>
<tr>
<td>CNN (U-Net, ResNet, DenseNet, PV-CrackNet) [40,41,45,47,50,52,53,59,66,67]</td>
<td>9</td>
<td>100%</td>
<td></td>
</tr>
</tbody>
</table>

6. Proposed Method

It is crucial to identify PV faults in a faster way and to find them in real time, which can help to decrease downtime. Another important task is to classify PV faults. Using ML classification techniques, it is possible to predict expected solutions for a specific failure of a PV system. This section is going to describe an ML-based method for identifying failures in a PV system. We propose a method that consists of smart converters and inverters with a central AI system. It considers historical data to predict the actual power generation from PV and measure the actual power generation. We developed an IoT-based device named the Neural Device (ND). Our proposed system will analyze PV power generation using the ND. This system not only detects PV faults, but also classifies PV faults and suggests prospective solutions. The intelligent algorithm adapts to the embedded system to detect PV faults and classify them. The proposed system is named the Smart Neural Solar System (SNSS)

**Smart Neural Solar System**

Photovoltaic module power generation is not linear. It depends on the surrounding environmental factors, like metrological parameters (mainly irradiation), temperature, wind speed, and humidity. All of these metrological parameters’ values and the value from PV comprise the photovoltaic and metrological historical database (PMHD). This PMHD is the key to making a strong ANN model that evaluates the present power generation of PV. Based on the PMHD, the ANN model is being updated and is using more data to make it more precise to achieve better results. In this ANN model, it is possible to include more parameters that will enable the model to more accurately predict PV power generation. Continuous data will be provided as the input to the proposed ANN model. Historical data are important, as combining them with continuous data will help the ANN model to be improved. During the training period, the ANN model registers a fault situation when a specific type of fault occurs. In this way, it includes different types of faults in the model that enable the model to identify the faults and classify them.

Different types of technologies exist for PV systems in the market, but it is important to have specific PMHD information, as this information determines the appropriate technology for PV systems. Specific technology-based models can assist in predicting power from a specific PV system, but a PV system can have its own ANN model, depending on its PMHD dataset. The photovoltaic panel has a five-parameter model that simulates a solar cell. In a situation where it has only metrological data, it is possible to generate the power output using that five-parameter PV model.

Figure 4 depicts a single-diode-based five-parameter PV cell model. This cell model represents a solar cell. In Figure 4, the $I_L$ photocurrent is the current source of the model, $I_0$
is the current over the single diode $D$, $R_s$ is the series resistance, and $R_{sh}$ represents the shunt resistance that is in parallel.

$$I = I_L - I_0 \exp[(e(V + IR_s)/nkT)] - 1 - (V + IR_s)/R_{sh} \quad (1)$$

**Figure 4.** Single-diode five-parameter PV cell model [69].

This represents the output from the solar cell, which depends on the metrological data. The first step of the proposed method is to run the ANN model based on metrological data to predict the PV power generation in a particular time. Then, we make a comparison with the present real-time data generated from the PV panels. This comparison allows us to benchmark the currently obtained PV power in real time with the PV power obtained previously based on historical data. That gives us an idea of the different types of faults in the PV system.

Figure 5 illustrates the mesh network between PV modules integrated with NDs, each of them connected with one PV module. A single neural device can analyze data and send information to a central system through a gateway. The ND stores information generated from the paired PV module. Our proposed ANN model allows for the comparison of generated power and the identification of different types of PV faults based on PMHD. Mesh networks can communicate with each other in real time, allowing them to compare power generation among themselves while also sending data to a central system. The central system performs the final analysis using PMHD, providing results on module faults, their classification, and mitigation techniques.

**Figure 5.** Mesh network between PV modules integrated with NDs [70].
7. Conclusions

The need for robust solar panel systems in renewable energy is more important than ever. Traditional methods of finding faults have limitations, especially when large numbers of panels are involved. This systematic review followed PRISMA guidelines to narrow our literature search down from 142 studies to 31 papers published between 2021 and 2023. In this study, we found a growing use of AI methods to detect faults in solar panels. These range from simpler techniques like Principal Component Analysis to more advanced ones like Neural Networks. Combining traditional and AI methods also appears to give better results. Moreover, ensemble learning and stacking classifiers have been particularly effective in diagnosing faults. However, there are challenges to overcome. There is a shortage of publicly available data for testing these methods. Some complex algorithms may also be hard to apply in real-world settings. In this study, we found that some types of faults receive more attention in research than others. This study finds that many types of AI are being used to spot problems in solar panels, and deep learning methods perform better than traditional ones.

Through this study, gaps in the literature have been identified, such as the need for standardized evaluation metrics and more robust real-world testing. This study addresses shortcomings and offers specific recommendations, such as making datasets publicly accessible, simplifying complex algorithms for real-world applications, and using hybrid models that suit the particular nature of the problem. Finally, we introduce an ANN-based approach for the detection and classification of PV faults. Our method, named the Smart Neural Solar System (SNSS), employs a continuous input of PMHD to the ANN model. This enhances its accuracy with the extended utilization of PMHD data. SNSS is integrated into the PV module, which enables us to autonomously identify and categorize PV faults. We are currently experimenting with the efficiency of our proposed model, which is integrated into solar devices. In future, further extensive studies related to PV fault detection using AI techniques are necessary to uncover additional insights. Also, the proposed model should undergo additional experimentation and improvements to better integrate with PV, which will enhance the efficiency of solar devices.

**Funding:** This research received no external funding.

**Data Availability Statement:** Data sharing is not applicable to this article.

**Conflicts of Interest:** The authors declare no conflict of interest.

**Abbreviations**

- **PV** Photovoltaic
- **AI** Artificial Intelligence
- **ANN** Artificial Neural Network
- **PRISMA** Preferred Reporting Items for Systematic Reviews and Meta-Analyses Statement
- **KNN** k-Nearest Neighbor
- **CNN** Convolutional Neural Network
- **RNN** Recurrent Neural Network
- **ML** Machine Learning
- **PCA** Principal Component Analysis
- **CV** Computer Vision
- **NLP** Natural Language Processing
- **UAVs** Unmanned Aerial Vehicles
- **FPN** Feature Pyramid Network
- **YOLOv5** You Only Look Once
- **FIA** Feature-Induced Augmentation
- **FDD** Fault Detection and Diagnosis
- **NB** Naïve Bayes
- **MLSC** Multi-Layer Stacking Classifier
EL  Ensemble Learning
AEM  AdaBoost Ensemble Model
MICA  Modified Independent Component Analysis
BPNN  Backpropagation Neural Network
SNSS  Smart Neural Solar System
PMHD  Photovoltaic and Metrological Historical Data
ND  Neural Device

References
15. Kurukuru, V.S.B.; Haque, A.; Khan, M.A.; Sahoo, S.; Malik, A.; Blaabjerg, F. A review on artificial intelligence applications for grid-connected solar photovoltaic systems. Energies 2021, 14, 4690. [CrossRef]


64. Hwang, H.P.-C.; Ku, C.C.-Y.; Chan, J.C.-C. Detection of malfunctioning photovoltaic modules based on machine learning algorithms. *IEEE Access* 2021, 9, 37210–37219. [CrossRef]

65. Yang, N.-C.; Ismail, H. Robust intelligent learning algorithm using random forest and modified-independent component analysis for PV fault detection: In case of imbalanced data. *IEEE Access* 2022, 10, 41119–41130. [CrossRef]


68. Li, L.; Wang, Z.; Zhang, T. Gbh-yolov5: Ghost convolution with bottleneckcsp and tiny target prediction head incorporating yolov5 for pv panel defect detection. *Electronics* 2023, 12, 561. [CrossRef]


**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.