

AI-Driven Image-Based Fault Detection and Diagnosis in Photovoltaic Systems: A Bibliometric Analysis and Comprehensive Review

H. Ait Lhaj^{1*}, K. Amechnoue¹, R. Herbazi²

¹ Mathematics, Informatics and Applications Team, National School of Applied Sciences of Tangier, Abdelmalek Essaâdi University, Tetouan, Morocco.

² Intelligent Systems and Applications Laboratory, Moroccan School of Engineering Sciences, Tangier, Morocco.

Received Date 12 July 2026; Revised Date 09 August 2026; Accepted Date 17 October 2026

*Corresponding author: hichamaitlhaj52@gmail.com (H. Ait Lhaj)

Abstract

The growing emphasis on renewable energy has led to a notable increase in photovoltaic (PV) installations, highlighting the need for advanced diagnostic methods to enable efficient maintenance and performance optimization. This paper provides a comprehensive overview of recent developments in the use of artificial intelligence (AI) for image-based fault detection and diagnosis in PV systems, with particular focus on the influence of modality and task design on performance and real-world deployability. First, a bibliometric analysis was conducted to trace the field's evolution and structural trends, followed by an in-depth technical synthesis organized by fault type, learning paradigm, and application goal. The bibliometric data were sourced from Scopus and Web of Science (WoS). Using Bibliometrix/Biblioshiny, we observe rapid growth in publication activity, increasing collaboration, and a thematic evolution from early classification-oriented studies toward more deployment-relevant pipelines. Consequently, this review highlights persistent technical challenges as well as emerging research directions, providing a structured reference for both researchers and practitioners, and outlines actionable priorities for advancing reliable, field-deployable inspection systems.

Keywords: *PV fault diagnosis, Image-based PV inspection, AI-based fault detection, Machine learning, Photovoltaic.*

1. Introduction

The adoption of photovoltaic (PV) installations has positioned solar energy as the foundation of renewable energy. Ensuring the reliability and high performance of PV systems, especially in the long term, has become a critical priority [1]. Faults that occur at the module or cell level, such as hotspots, delamination, degradation, and soiling, can significantly reduce and disturb energy output. Moreover, these faults accelerate aging, thereby increasing maintenance costs. Traditional fault detection methods, including I-V curve analysis, power ratio monitoring, and string-level electrical supervision, provide valuable global indicators [2], [3], [4]. However, these approaches often lack the spatial granularity needed to detect and localize subtle faults, as they rely either on manual inspection or embedded sensors, which can be labor-intensive and cost-prohibitive for large-scale installations [5]. In this context, image-based inspection has gained significant traction due to its ability to capture patterns that directly reflect

physical degradation mechanisms using visual, thermal, and electroluminescence (EL) imagery [6]. Integrating these imaging techniques with computer vision and deep learning overcomes the limitations of manual inspection by enabling automated, scalable, and accurate diagnostic systems, particularly well-suited for large-scale PV plants. These technologies can be deployed on unmanned aerial vehicles (UAVs) and infrared cameras, making them standard tools for routine monitoring. Despite research efforts, the literature on image-based PV monitoring varies in terms of objectives, methodologies, and evaluation metrics.

Several reviews have examined photovoltaic fault detection and diagnosis; however, the literature remains fragmented across different perspectives. Some reviews address AI-based PV fault diagnosis using electrical measurements, including I-V curves, thermal imaging, and machine learning models, highlighting the efficiency of AI approaches and the limitations of manual inspection, especially for large-scale plants [7], [8]. Although they state that it remains unclear

which AI approach is most suitable—partly due to the lack of public datasets, the complexity of some models, and the imbalance of research attention across fault types—progress has been made. In parallel, reviews focused on vision-based techniques have shown that computer vision, UAV inspection, and deep learning models can improve the detection and localization of PV faults, but unresolved challenges related to real-time processing, scalability, multimodal fusion, and interpretability remain [8], [9]. Moreover, bibliometric analyses of AI-based PV fault diagnosis have demonstrated that conventional review approaches mainly compare algorithms by accuracy, despite differences in application settings and datasets, including weather conditions and PV module technologies [10]. These observations motivate the need for a unified synthesis specifically focused on AI-driven image-based PV fault detection and diagnosis.

Therefore, this review aims to provide a comprehensive and structured synthesis of image-based fault detection and diagnosis methods for PV systems. This work adopts a broad perspective, organizing the literature according to fault types, imaging modalities, learning approaches, algorithms, and application objectives, rather than focusing on a single methodological dichotomy. The study is limited to image-based approaches leveraging RGB, thermal infrared (IR), and EL imaging, with an emphasis on AI-driven techniques. Purely electrical diagnostics fall outside the scope of this work. Particular attention is given to how different studies formulate the detection task, whether as anomaly detection or multi-class fault identification, and how these choices influence data requirements, model complexity, and practical deployment. To contextualize the subsequent technical synthesis, this work first presents a bibliometric analysis of the image-based PV fault diagnosis literature. This quantitative mapping offers a macro-level view of publication growth, influential sources and contributions, collaboration patterns, and the thematic evolution of research topics. The bibliometric findings are then used to motivate and structure the remainder of the review, which proceeds with a fault-oriented and modality-aware synthesis. Accordingly, the objectives of this review are to: (i) develop a bibliometric profile of the field, highlighting its growth trajectory, leading contributors, and evolving themes; (ii) categorize PV fault types detectable through imaging; (iii) analyze image acquisition techniques and preprocessing strategies used in the literature; (iv) review and classify AI methods, particularly deep

learning approaches; and (vi) identify key research gaps and future directions.

2. Bibliometric analysis

Bibliometric analysis is a valuable tool for systematically mapping and evaluating scientific landscapes within a specific domain, enabling the identification of research trends, influential works, and emerging areas. This method facilitates understanding the evolution of research topics, collaboration patterns, and the impact of various publications within the field. By analysing publication outputs, citation networks, and keyword occurrences, this approach provides a quantitative overview of the intellectual structure and dynamics of research, complementing qualitative literature reviews. The term bibliometrics was introduced by Alan Pritchard in 1969 to denote the application of mathematical and statistical methods to books and other media [11]. To conduct a bibliometric analysis, a rigorous methodology is employed that involves selecting relevant databases such as Web of Science, Google Scholar, and Scopus for literature searches, then processing and visualising the extracted data after removing irrelevant and duplicate entries. The cleaned publications are then analysed using specialised software like VOSviewer or Bibliometrix to reveal publication trends, co-authorship networks, and keyword co-occurrence patterns [12], [13]. Bibliometrix is an R package that offers comprehensive tools for scientometrics and bibliometrics, including statistical and scientific mapping algorithms enhanced by the user-friendly web interface Biblioshiny for non-coders [14]. Finally, the results are interpreted to identify research gaps and future directions. This quantitative approach mitigates biases in manual reviews, providing a more comprehensive understanding of the field's conceptual and intellectual structure.

2.1. Data collection

To conduct the analysis, we chose Scopus and WoS databases as the principal data sources because of their extensive multidisciplinary coverage of peer-reviewed literature, encompassing a broad range of journals, conference proceedings, and scholarly materials that are essential for a thorough bibliometric investigation. This coverage is particularly strong for engineering and computer-science-adjacent topics, an important consideration for image-based PV inspection and AI methods, which often appear in conference proceedings as well as journals. Scopus offers large-scale citation

indexing (more than 100 million records and 2.4 billion cited references, as reported in its fact sheet), while WoS covers 254 subject areas including more than 99 million records for journal, conference, and book content. Combining both databases provides robust citation and network-based mapping. Metadata completeness and compatibility with bibliometric analysis software like Biblioshiny further enhance their suitability for this study. The corpus was retrieved through a carefully constructed search query designed to capture all relevant publications related to image-based photovoltaic fault detection and diagnosis, applying the query to the article title, abstract, and keywords fields to maximize recall while remaining topic-specific, as illustrated in Table I. This query structure was translated into the following database-specific syntax: ("PV" OR

"photovoltaic" OR "solar panel") AND ("AI" OR "artificial intelligence" OR "deep learning" OR "machine learning") AND ("fault detection" OR "fault classification" OR "diagnosis") AND ("image-based" OR "thermal image" OR "electroluminescence" OR "EL image" OR "UV image" OR "images"). The imaging terms clause pairs modality-specific labels (e.g., "thermal image," "EL image," "UV image") with the generic term "images" to maximize recall; RGB imaging, for instance, carries no dedicated label in the query and is captured only through this generic term. Modality scope is therefore not determined by whether a named term matches a record but is enforced downstream during eligibility screening against the RGB/IR/EL boundary stated in the introduction.

Table 1. Search query structure

Component	Terms
PV / Domain terms	"PV" OR "photovoltaic" OR "solar panel"
AI methods	"AI" OR "artificial intelligence" OR "deep learning" OR "Machine Learning"
Fault-related terms	"fault detection" OR "fault classification" OR "diagnosis"
Imaging terms	"image-based" OR "thermal image" OR "electroluminescence" OR "EL image" OR "UV image" OR "images"
Final query	(PV / Domain terms) AND (AI methods) AND (Fault-related terms) AND (Imaging terms)

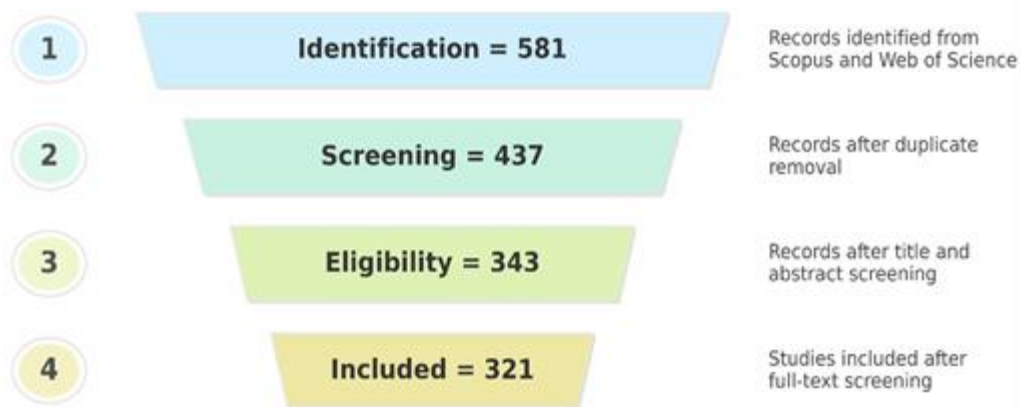


Figure 1. Data connection workflow

Abstract, and keywords fields to maximize recall while remaining topic-specific as illustrated in the Table . This query structure was translated into the following database-specific syntax: ("PV" OR "photovoltaic" OR "solar panel") AND ("AI" OR "artificial intelligence" OR "deep learning" OR "machine learning") AND ("fault detection" OR "fault classification" OR "diagnosis") AND ("image-based" OR "thermal image" OR "electroluminescence" OR "EL image" OR "UV image" OR "images") The imaging terms clause pairs modality-specific labels (e.g. "thermal image", "EL image", "UV image" with the generic term "images" to maximize recall; RGB imaging, for

instance carries no dedicated label in the query and is captured only through this generic term. Modality scope is therefore not determined by with named term a record matched, but enforced downstream during eligibility screening against the RGB/IR/EL boundary stated in the introduction.

Records retrieved were subsequently refined by excluding irrelevant document types, such as editorials, letters, and reviews to keep the technical synthesis grounded in original empirical contributions; model architectures, datasets, and reported results rather than secondary summaries of them. This also avoids double-counting citation and co-occurrence signals in the bibliometric analysis, since a review article's citation pattern

reflects its synthesis of the field rather than an independent methodological contribution, which would otherwise bias network centrality and co-citation measures toward the review itself. The exported information included core bibliographic data such as titles, abstracts, publication year, DOI, and source title. Authorship information, including author names, affiliations, and countries, was also extracted to facilitate co-authorship network analysis and geographical distribution mapping. Citation-related data such as cited references and Analysis and geographical distribution mapping were also performed. Citation-related data, such as cited references and citing articles, were collected to enable co-citation and bibliographic coupling analyses. Exports from both scholarly indexes were then merged into a unified dataset subjected to a structured cleaning workflow using the Zotero tool, where duplicates were identified and removed, and variations in author names and affiliations were standardized to ensure data consistency for subsequent analyses. Duplicate removal was performed using a hierarchical matching approach, including DOI, title, and Title + first author + year matching, resulting in the removal of 144 duplicates. A relevance eligibility check was applied to align the corpus with the scope of AI-driven image-based PV fault detection, classification, or diagnosis. Screening was conducted at two levels: first, removing records that clearly fall outside the scope (e.g., purely electrical signal-based PV diagnosis without imaging, non-PV solar applications, imaging modalities outside the RGB/IR/EL boundary defined in Section 1, or AI studies not addressing faults), which reduced the corpus to 343 records

(94 records excluded). Then, full-text verification was applied to borderline cases where the abstract did not clearly indicate whether imaging data and diagnostic objectives were central to the study, yielding a final corpus of 321 documents (22 records excluded at this stage). The entire process is summarized in Figure 1.

To enhance the accuracy of the co-occurrence analysis, keywords were normalized by consolidating synonyms and closely related terms to prevent fragmentation and improve the thematic representation of the dataset (e.g., photovoltaic cells and solar cells; EL and electroluminescence; IR thermal and thermal imaging).

2.2. Publication growth and source landscape

The cleaned dataset was analyzed using Bibliometrix/Biblioshiny. The dataset spans the period from 2017 to 2025 after eligibility screening, with records from 2015–2016 present in the initial 581-record extraction removed because these papers did not meet the imaging-modality and AI-methodology scope criteria described in Section 2.1. The final dataset comprises 321 documents distributed across 204 sources, indicating a broad and dispersed publication landscape. Additionally, the field is notably recent, as reflected by a low average document age (2.49 years) and a high annual growth rate (80.57%), which highlight the accelerating research attention and rapid development within the domain. In parallel, the corpus contains 1,687 keywords plus and 815 author keywords, emphasizing strong terminological diversity and the coexistence of multiple research paradigms (PV engineering, inspection/remote sensing, and AI/computer vision). These key metrics are presented in

Table , providing a comprehensive overview of the dataset's quantitative characteristics.

Table 2. Main information and descriptive statistics of the bibliometric dataset

Description	Results
MAIN INFORMATION ABOUT DATA	
Timespan	2017–2025
Sources (Journals, Books, etc)	204
Documents	321
Annual Growth Rate (%)	80.57
Document Average Age	2.49
Average citations per doc	18.15
References	3534
DOCUMENT CONTENTS	
Keywords Plus (ID)	1687
Author's Keywords (DE)	815
AUTHORS	
Authors	1102
Authors of single-authored docs	9
AUTHORS COLLABORATION	
Single-authored docs	11
Co-Authors per Doc	4.25

International co-authorships (%)	28.66
DOCUMENT TYPES	
Article	196
Conference paper	125

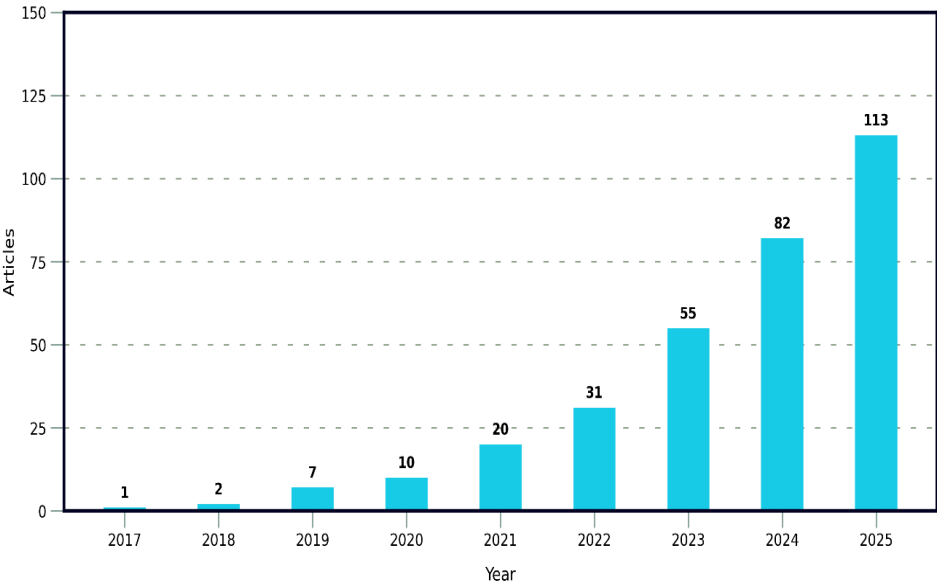


Figure 2. Annual Scientific Production

Annual scientific production shows a clear growth trajectory (Figure 2), with publications increasing from 1 in 2017 to 113 in 2025, with a marked inflection point after 2020. This growth pattern aligns with the broader diffusion of deep learning and the maturation of AI techniques for practical applications in renewable energy. The source landscape further confirms the interdisciplinary nature of the field (Figure 3), where the most productive sources include Energies, Solar Energy, IEEE Access, Scientific Reports, and Sensors, alongside energy systems venues (e.g., Energy

Conversion and Management) and proceedings series. Overall, this distribution across PV/energy journals, sensing/remote sensing, and IEEE sources indicates a robust integration of advanced computational methods with solar energy research, demonstrating the multidisciplinary engagement required for innovation in photovoltaic fault detection. Finally, document types reveal a strong conference presence, with 196 articles compared to 125 conference papers.

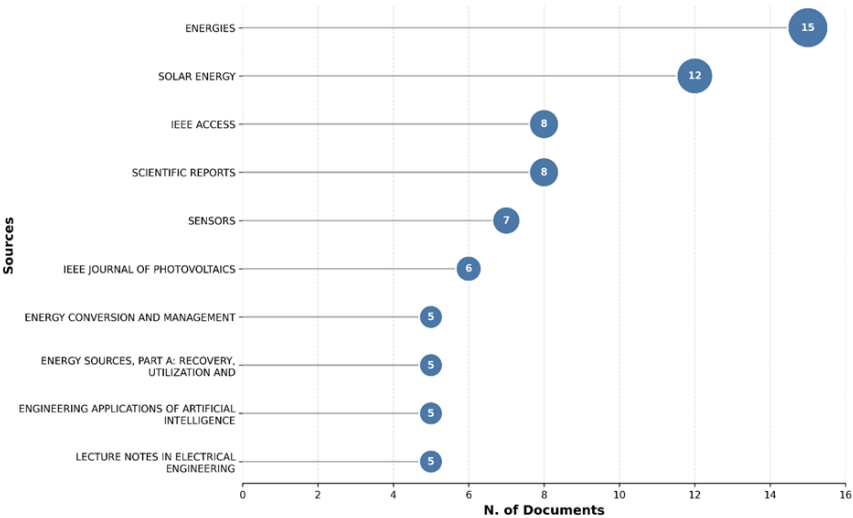


Figure 3. Most relevant sources

2.3. Influential documents and leading contributors

The results show a moderate-to-strong overall impact, with an average of 18.15 citations per document and 3,534 references cited across the dataset. The average citation rate indicates significant scholarly engagement within the domain, while the extensive reference list underscores the field’s reliance on a broad foundational knowledge base. The most globally cited documents often introduce novel architectural paradigms or foundational methodologies. Leading papers, such as those with 259 and 241 citations respectively [15], [16], have significantly influenced the development of deep learning applications such as papers with 259 and 241 citations respectively [15], [16]. The normalized citation indicator further highlights papers with

unusually rapid growth relative to their publication age. Notably, several items from 2024–2025 appear among the highest normalized total citations (TC), indicating their immediate and substantial impact on the current research trajectory and consolidation around new approaches and benchmarks, particularly those shaped by fast-evolving deep learning methods and inspection platforms. Regarding contributors, the dataset includes 917 authors (only 8 associated with single-authored documents), indicating a broad and collaborative research community. The most productive authors by publication count include “Mellit A” with 14 papers, and “S Nv” and “Sugumran S,” each contributing 13 papers, followed by several authors with 10, 8, and 7 publications (Figure 4). Importantly, fractionalized productivity suggests that leading contributors often operate within multi-institutional teams.

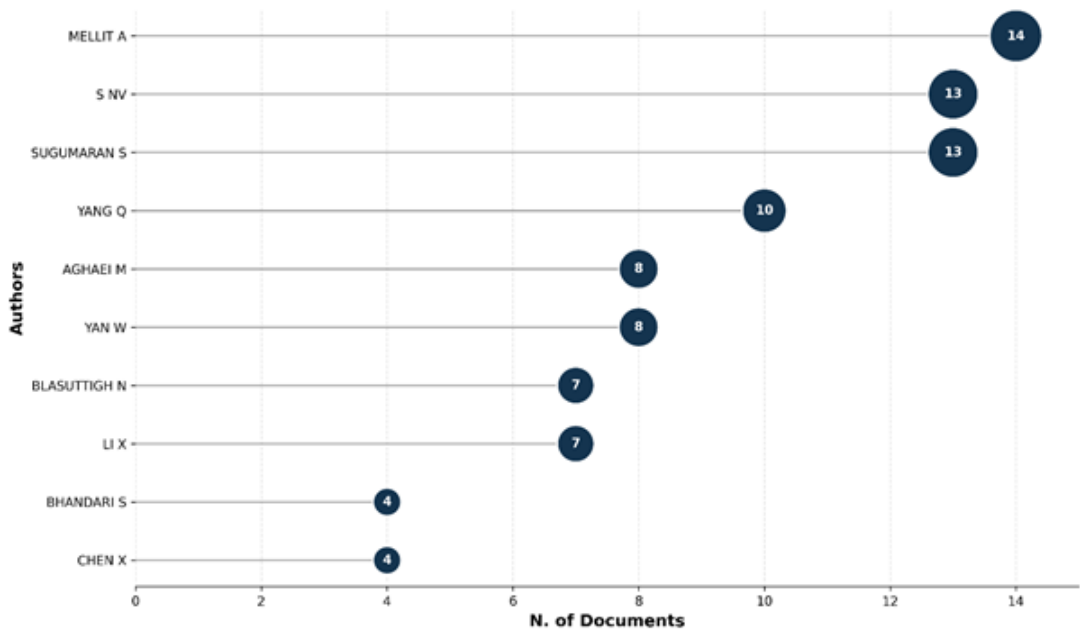


Figure 4. Most productive authors

2.4. Collaboration patterns across countries and institutions

Authorship statistics indicate a structurally collaborative research ecosystem: the corpus contains only 11 single-authored documents, with an average of 4.25 co-authors per paper. Moreover, international co-authorship reaches 28.66%, suggesting that cross-border partnerships are routine rather than exceptional modes of production. The collaboration map, as illustrated in Figure 5, reflects a connected network with frequent participation of countries such as China, India, the United States, and Algeria, alongside a broader set of European and Asian partners, as

shown in Figure 6: Co-occurrence network. This global interconnectedness underscores the reality that PV deployment and inspection research are geographically distributed across different climates and plant scales, which motivates collaborations for data access and validation. From an interpretive standpoint, the collaboration structure supports two important inferences for the remainder of the review: (i) multi-site and multi-condition evaluations are likely to become more common, and (ii) the persistence of international teams may accelerate the spread of shared practices (e.g., preferred imaging modalities, labeling protocols, and benchmark model families), even if open datasets remain unevenly available.

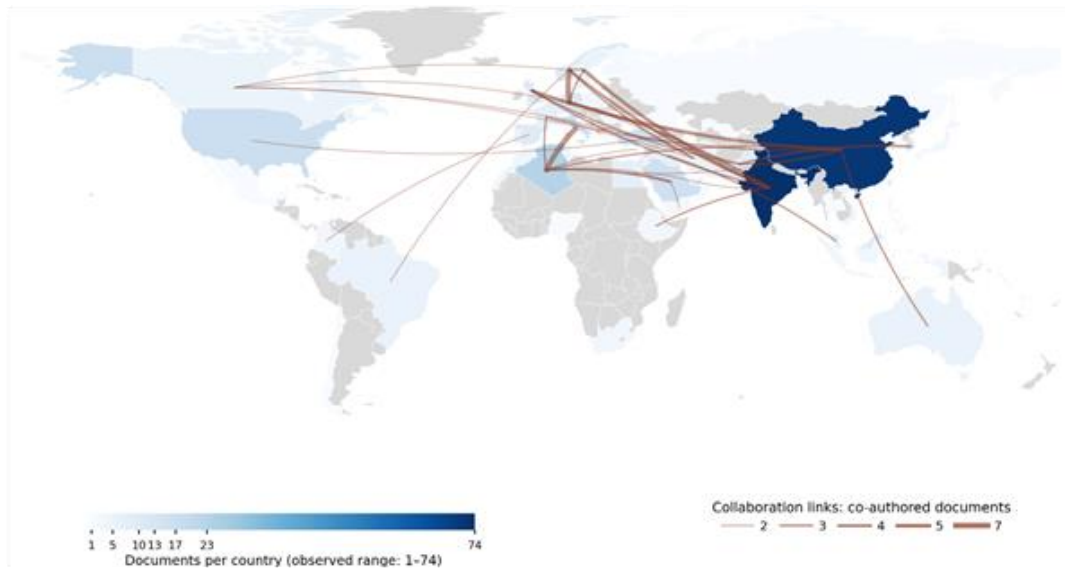


Figure 5. Country collaboration map

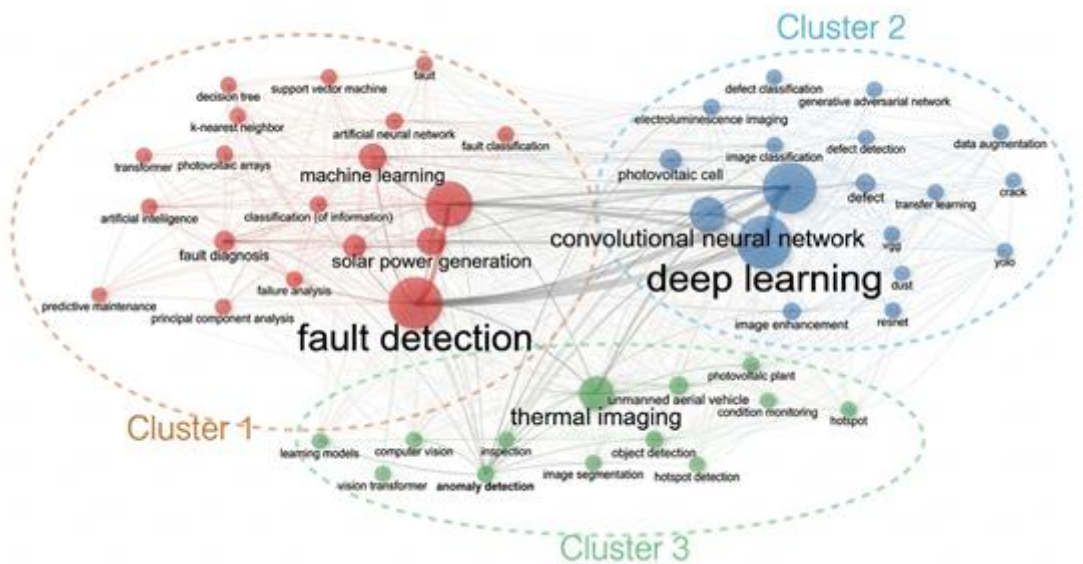


Figure 6. Co-occurrence network

2.5. Conceptual structure and thematic trends

The co-occurrence map reveals a field organized around three connected conceptual domains with 50 linked nodes, each representing a different way of framing image-based fault diagnosis in photovoltaic systems (Figure 6). The network was generated using Louvain community detection with inclusion-index normalization. The resulting clusters vary in size (Cluster 1: 20 terms; Cluster 2: 20 terms; Cluster 3: 10 terms) and in centrality: Cluster 2, anchored by "deep learning," "photovoltaic module," and "convolutional neural network," has the highest aggregate betweenness centrality (462.2 versus 321.2 for Cluster 1 and 29.5 for Cluster 3), indicating its terms serve disproportionately as bridges across the network.

"Fault detection" and "deep learning" stand out as the most central nodes in the network, which

indicates that the literature has consolidated around deep learning as the main paradigm used in the field.

The first cluster is anchored around terms such as "solar power generation," "predictive maintenance," and "machine learning (ML)." This configuration suggests a broad research stream that aligns with classical diagnostic thinking, where the emphasis is on embedding fault diagnosis within operational PV systems as energy-producing assets rather than solely recognizing faults from images. The presence of traditional machine learning algorithms like "decision tree (DT)" and "support vector machine (SVM)" indicates that part of the literature still relies on structured-feature approaches in studies with limited datasets or where interpretability remains important. Practically, Cluster 1 captures the segment of the field where AI is treated as a diagnostic tool within

a broader photovoltaic engineering framework.

Cluster 2 presents a different logic. It is centered on “deep learning” and “convolutional neural network (CNN),” surrounded by terms such as “photovoltaic cell,” “EL imaging,” and “defect detection.” The structure of this cluster clearly points to an image-centric research direction dominated by modern computer vision architectures. The keyword arrangement suggests that this cluster is especially associated with fine-grained defect recognition, often at the cell level, where EL imagery has become a major testing ground for deep architectures because it exposes cell-level defects requiring high visual sensitivity. The keyword arrangement suggests that this cluster is especially associated with fine-grained defect recognition often at the cell level where EL imagery has become a major testing

ground for deep architectures because it exposes some cell-level defects that demand high visual sensitivity.

Cluster 3, in parallel, reflects operational inspection, with thermal imaging occupying the central position because it is one of the most practical modalities for large-scale PV monitoring. It reveals electrically induced abnormalities, especially hotspots and mismatch-related faults. The surrounding terms like “UAVs” and “photovoltaic plant” indicate that this modality is used for inspection at the plant scale rather than solely at the cell scale. “Object detection” and “image segmentation” point to a shift from image classification toward localization tasks, aiming to identify where faults are and how they are spatially distributed.

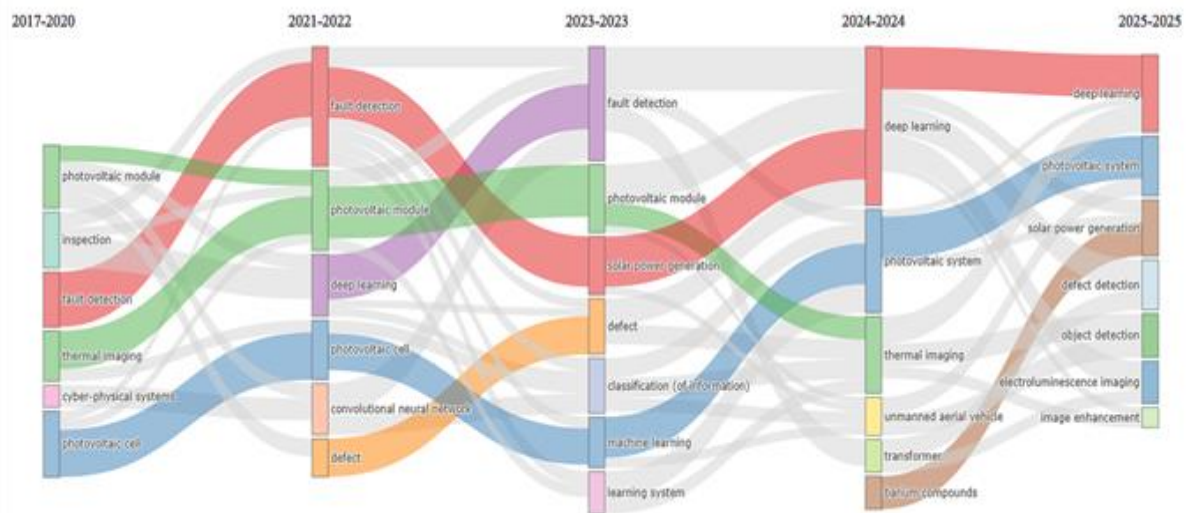


Figure 7. Thematic evolution map

Examining the thematic evolution in Figure 7, the earliest research is characterized by a relatively sparse and exploratory thematic profile. The 2017–2020 window is dominated by terms including *photovoltaic model*, *photovoltaic cell*, *fault detection*, and *inspection*, indicating that the field is still defining its vocabulary, with researchers establishing what to inspect, under which modality, and at which scale. The appearance of *cyber-physical systems* as a minor node reflects early ambitions toward integrated monitoring architectures.

In the 2021–2022 window, *fault detection* becomes the dominant thematic anchor, drawing the largest share of keyword flow from the previous period, while *deep learning* and *convolutional neural network* emerge as distinct, high-weight

nodes. This co-emergence indicates that the field has reached a point where CNN-based architectures are the accepted computational foundation for image-based PV diagnosis. *Photovoltaic cell* persists alongside *photovoltaic module*, suggesting that research is simultaneously operating at two spatial scales: the cell level, where EL-based defect recognition dominates, and the module level, where thermal inspection pipelines are maturing.

The 2023 snapshot, in contrast, is less concentrated, with *fault detection* decreasing in relative weight, while *classification of information*, *machine learning*, *solar power generation*, and *learning systems* appear as co-prominent nodes. This fragmentation reflects a transitional moment, suggesting that an increasing proportion of

publications embed fault diagnosis explicitly within the operational context of energy production rather than treating it as a standalone image recognition task.

By 2024, the thematic landscape reconsolidates around deep learning and photovoltaic systems. The shift from photovoltaic module reflects a scale transition from component-level recognition to plant-scale analysis. *Thermal imaging* remains a persistent modality-specific theme, confirming its role as the most field-deployable imaging technique. This period also marks the appearance of *transformers* as an independent node, indicating the diffusion of vision transformer (ViT) architectures into PV inspection tasks. "Transformer" and "vision transformer" emerge as two distinct, low-centrality nodes (zero betweenness, near-minimum closeness and PageRank among all retained terms) during this period, consistent with an emerging theme;

Table includes three transformer-based studies, providing a lower bound on this trend's representation in the technical literature.

The 2025 configuration reveals what this entire trajectory was building toward. Deep learning remains the primary organizing concept of the field, but the surrounding thematic structure is qualitatively different from 2021–2022, where *defect detection* and *object detection* appeared as distinct, high-weight nodes, indicating a consolidation around localization-oriented task formulations. EL imaging enters the map as an independent theme for the first time, reflecting the growing research commitment to cell-level, high-sensitivity diagnostics. Additionally, *image enhancement* appears, indicating that preprocessing is being treated as a first-class research problem.

The uneven period lengths in the thematic evolution map—four years for 2017–2020, two for 2021–2022, and one each thereafter—reflect the corpus's markedly uneven temporal density: given an annual growth rate of 80.57% (

Table), most of the included documents fall within the final three years of the search window. This period was specifically chosen to preserve visual resolution in the thematic map and avoid collapsing this high-volume period into a single overcrowded band. The apparent fragmentation of the 2023 window is consistent with genuine thematic diversification during a period of accelerating output.

3. Fault types

Photovoltaic modules constitute multi-layered assemblies comprising interconnected solar cells, metallic contact structures, polymeric encapsulants, a front-surface glass substrate, and a rear-side protective backsheet [17]. Degradation and component failures occurring at the cell, module, or system level may compromise overall performance metrics, which require systematic fault

identification as a fundamental component of photovoltaic operation and maintenance protocols [18]. Faults in photovoltaic systems can be classified through multiple schemes, including physical origin (cell-level, module-level, or system-level) [19], etiological factors (manufacturing anomalies, environmental causes, mechanical loading, or electrical transients) [20], or phenomenological manifestation (electrical, thermal, or optical signatures) [21]. From a diagnostic perspective, not all photovoltaic faults produce observable optical signatures. Many electrical anomalies remain undetectable via imaging methodologies and require direct electrical characterization. The choice of imaging technique is crucial in determining which faults can be observed. RGB imaging is particularly suited for surface-related and structural defects, while thermal IR imaging excels at revealing temperature anomalies associated with electrical mismatch or overheating, and EL imaging provides high sensitivity to cell-level defects such as micro-cracks and inactive regions. These characteristics influence the development of AI models, including network architecture, feature representation, and labeling strategies. This section provides an overview of the principal fault types commonly addressed in the literature, focusing on those detectable through visual, thermal, or EL imaging. This organization offers a structured foundation for linking physical degradation mechanisms to image-based analysis and AI-driven diagnostic approaches. Some faults manifest across multiple modalities simultaneously.

Table summarizes these cross-modal relationships. When two modalities are listed, the fault is not simply "more visible" in one; each modality captures a different stage or aspect of the same physical process. A crack is a structural discontinuity (EL: interrupted current collection) that may or may not be thermally active (Thermal: only once it disrupts current flow enough to create local resistive heating) or visually apparent (RGB: only once it propagates to the surface or triggers snail-trail discoloration).

3.1. Faults visible in RGB imaging

RGB imaging is the most accessible modality and is widely used for PV plant inspection because it relies on standard cameras mounted on either ground-based systems or UAVs [22], [23]. RGB images primarily capture surface appearance and structural integrity, making them suitable for identifying manifest faults. The common faults detectable through RGB imaging include soiling caused by dust, sand, bird droppings, or organic matter, which appear as non-uniform surface patterns and lead to irradiance losses received by the module. Nearby obstacles, vegetation, and structural elements cause temporary or permanent shading, which can be identified as darkened regions. Defects such as glass breakage, frame deformation, and delamination of encapsulation layers are often observable as cracks, bubbles, or cloudy regions on the module surface [24], [25].

Additionally, aesthetic defects like snail trails are visible as dark, meandering patterns. From an AI perspective, RGB-based fault analysis often relies on convolutional neural networks for image classification, object detection, or segmentation [26], [27], [28], [29], [30], [31]. However, RGB imaging is sensitive to lighting conditions, viewing angles, and environmental variability, which can complicate model generalization and necessitate robust preprocessing and data augmentation strategies [21], [22].

3.2. Faults visible in thermal imaging

Thermal infrared imaging captures temperature distributions across PV modules and is particularly effective for identifying faults associated with electrical mismatch and power dissipation. Thermal anomalies often serve as early indicators of malfunction that are not visually apparent in RGB images [32]. The most prominent thermal fault is the hotspot, characterized by temperature elevations resulting from cracked cells, shading, bypass diode failures, or severe mismatch conditions [33]. Bypass diode malfunctions can also be detected as distinct high-temperature regions affecting entire cell strings [6]. Using

thermal imaging can additionally reveal cell mismatch, interconnection defects, and string-level imbalances, which manifest as non-uniform heating patterns across modules or arrays. In image-based AI systems, thermal data are frequently employed for anomaly detection or binary classification, as abnormal temperature patterns tend to be more salient than subtle visual cues. Nevertheless, environmental conditions such as irradiance and wind speed can influence thermal signatures, which must be considered during data acquisition and model training.

3.3. Faults visible in EL imaging

Electroluminescence imaging is a high-resolution diagnostic technique that uncovers the internal electrical activity of solar cells by capturing emitted radiation when the module is electrically biased. EL imaging is particularly sensitive to cell-level defects and is widely regarded as the reference modality for identifying microstructural degradation [34].

Faults detectable via EL imaging include microcracks, inactive or electrically isolated cell areas, finger and busbar interruptions, and potential-induced damage.

Table 3. cross-modal fault matrix

Fault	Primary modality	Secondary manifestation	Mechanism (secondary)
Micro-crack	EL	Thermal (if crack disconnects circuit)	Localised loss of current path creates a small resistive/hotspot signature detectable thermally once the crack is electrically active, not merely structural.
Delamination	RGB	Thermal (subtle)	Encapsulant/backsheet separation reduces the conductive heat path, producing a small local temperature elevation, generally weaker than a hotspot from cell mismatch.
Glass breakage	RGB	EL, Thermal (if cell circuit affected)	Fracture propagating to interconnects or cell material can isolate a region electrically (EL: dark/inactive) or create a mismatch (Thermal: localised heating).
Snail trail	RGB	EL	Snail trails are a documented visual symptom of an underlying microcrack; the discoloration and the crack that produced it are frequently co-located.
Hotspot	Thermal	EL (if crack-driven), RGB (if chronic)	A crack-driven hotspot is visible in EL as the same disconnected region; long-standing hotspots can visibly discolour or blister the surface.
Shading	RGB	Thermal	A shaded cell in series can be forced into reverse bias by neighbouring cells, dissipating power as heat — shading and hotspot are frequently the same physical event observed at two points in its causal chain.
PID	EL	Thermal (variable)	Shunting associated with PID can alter local thermal behaviour, though the direction and magnitude vary by mechanism and are less consistently reported.

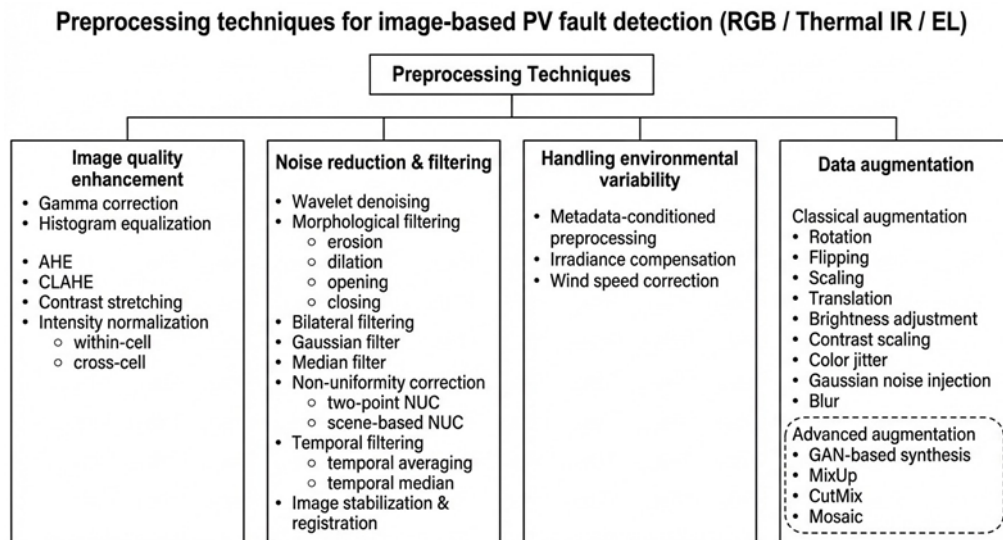


Figure 8. Preprocessing techniques summary

degradation (PID) [35], [36]. These defects appear as dark or low-intensity regions corresponding to reduced carrier recombination. Unlike RGB or thermal imaging, EL provides direct insight into electrical continuity at the cell level, enabling precise identification of defects. AI-based analysis of EL images often involves deep learning models for fine-grained classification or segmentation [37], [38], [37]. However, EL imaging typically requires controlled conditions, specialised equipment, and offline inspection, which limits its applicability for continuous PV systems diagnostic [35].

4. Preprocessing techniques

The efficacy of image-based fault detection systems for PV installations fundamentally depends on the quality and consistency of input data. The preprocessing stage is a critical part of the pipeline that transforms raw imagery into standardised and enhanced representations suitable for robust feature extraction and classification [26], [38], [39]. The different preprocessing techniques used in the literature are summarised in Figure .

4.1. Colour quality enhancement

Image acquisition suffers from temporal and spatial illumination variations due to solar position, cloud cover, and atmospheric scattering. Lighting normalisation techniques aim to standardise image appearance across diverse capture conditions. Gamma correction represents a fundamental approach, [20] demonstrates the effectiveness of applying a power-law transformation to normalise contrast across panels. For IR thermography, studies have reported that improved gamma correction preprocessing for infrared

thermography normalises contrast and temperature shading prior to CNN training contributing to high detection accuracy. Unlike RGB, thermal gamma must preserve temperature ordering while enhancing discrimination between operational and anomalous temperature ranges. Histogram equalisation and its adaptive variants contrast-limited adaptive histogram equalisation CLAHE offer alternative approaches for enhancing local contrast while mitigating global illumination effects and constrain the maximum slope of the cumulative histogram for EL images, while global histogram equalisation redistributes pixel intensities to achieve a uniform histogram, hence maximising the utilisation of available dynamic range. In addition, adaptive histogram equalisation (AHE) enables localised contrast enhancement. This transformation enhances overall contrast but may over-amplify noise in homogeneous regions and compress subtle variations in high-gradient areas. Intensity normalisation addresses cell-to-cell luminescence variations due to material properties, electrical characteristics, and imaging geometry. Within-cell normalisation computes statistics (mean, median) for each cell and applies scaling to achieve consistent intensity distributions. Cross-cell normalisation standardises intensity across all cells in a module, enabling fair comparison and preventing bias toward brighter or dimmer cells. Contrast enhancement amplifies temperature differences to improve visual discrimination and feature extraction. Linear contrast stretching maps the observed temperature range to the full dynamic range. Chen et al. applied grayscale conversion and data enhancement to IR datasets before classifier training, achieving 84.06% accuracy in fault detection [40].

4.2. Noise reduction

Aerial imagery frequently exhibits degradation from atmospheric turbulence, sensor noise, and motion blur. Moreover, thermal cameras exhibit characteristic noise patterns, including fixed-pattern noise, temporal noise, and non-uniformity artefacts. Preprocessing must suppress these artefacts while preserving edge-critical information and genuine temperature variations indicative of faults. For visible-light imagery, wavelet-based denoising has emerged as a particularly effective technique for PV inspection applications. Yang et al. demonstrated that Daubechies 4 (DB4) wavelet decomposition followed by soft-thresholding of detail coefficients effectively suppresses high-frequency noise [17]. Morphological filtering operations, including erosion, dilation, opening, and closing, provide complementary enhancement capabilities. These filters can emphasise defect geometries (e.g., cracks, delamination boundaries) while suppressing texture variations irrelevant to fault diagnosis. The selection of the structuring element shape and size is a critical design parameter. For linear crack detection, elongated structuring elements aligned with dominant crack orientations prove most effective [41]. Additionally, bilateral filtering reduces noise while preserving sharp boundaries between regions [42]. For PV applications, bilateral filtering has shown particular utility in preprocessing images before edge detection or segmentation operations. For thermal PV imagery, spatial filtering applies kernels to reduce high-frequency noise. Gaussian filters provide isotropic smoothing with a controllable bandwidth, while median filters suppress salt-and-pepper noise [43]. Non-uniformity correction (NUC) addresses systemic pixel-to-pixel sensitivity variations in thermal detector arrays. Two-point NUC employs gain and offset corrections derived from calibration against uniform-temperature blackbody sources, while scene-based NUC algorithms estimate and correct non-uniformity from operational imagery without external calibration; however, these methods assume temporal stationarity and may fail for dynamic scenes. Temporal filtering exploits redundancy across video sequences or multi-frame acquisitions. Temporal averaging reduces random noise proportionally to the square root of frame count, while temporal median filtering suppresses transient artefacts. For UAV-based thermal surveys, stabilisation and registration enable temporal filtering across sequential frames, substantially improving the signal-to-noise ratio

[44], [45].

4.3. Handling environmental variability

Environmental factors, including irradiance level, ambient temperature, and wind speed, substantially influence the RGB appearance and thermal signatures of PV modules. Effective preprocessing strategies must account for these systematic variations to ensure model generalization across deployment scenarios. Metadata-conditioned preprocessing leverages auxiliary sensor data (irradiance meters, temperature sensors, weather stations) to normalize images based on capture conditions [46]. For instance, image intensity can be adjusted according to measured solar irradiance to mitigate overall brightness fluctuations, while temperature-aware color corrections compensate for spectral variations in panel appearance. Such approaches require synchronized sensor integration but yield substantial improvements in cross-condition performance. Irradiance compensation accounts for the fundamental relationship between incident solar radiation and module temperature. Under steady-state conditions, panel temperature follows a linear relationship with irradiance: $T_{\text{module}} = T_{\text{ambient}} + kG$, where G represents irradiance and k is a module-specific thermal coefficient. Preprocessing can normalize thermal images by irradiance level, either through division by measured irradiance $T_{\text{norm}} = (T_{\text{image}} - T_{\text{ambient}})/G$ or $T_{\text{norm}} = T_{\text{image}}/G$, or by restricting analysis to time windows with stable, high irradiance (typically $> 700 \text{ W/m}^2$) [47]. Additionally, wind speed correction addresses convective cooling effects that reduce surface temperatures proportionally to wind velocity. Empirical models relate temperature reduction to wind speed. Some studies collect thermal-visible image pairs under varied wind conditions to train models robust to these effects, implicitly learning wind-invariant features rather than explicit compensation [48].

These compensation strategies assume conditions have stabilized long enough for the module to reach thermal equilibrium with its current irradiance, an assumption made explicit by the underlying steady-state relationship. This assumption is frequently violated during UAV-based thermal inspections, where a flight covering a large array can take several minutes, and cloud cover can change significantly within that window. A controlled drone IRT study spanning four irradiance levels and altitudes reported an unexplained approximately 2°C temperature deviation attributable to transient irradiance changes [49]. PV modules possess non-trivial

thermal mass, and their temperature response to a step change in irradiance lags rather than tracks it instantaneously; a module recently shaded by a passing cloud retains elevated temperature from its prior, higher-irradiance state, while one recently exposed retains residual cooling—plausible physical origins for exactly this kind of unexplained transient deviation. A transient, lag-driven temperature deviation and a real hotspot can be visually similar in a static thermal frame, differing mainly in properties a single frame cannot capture: persistence across successive frames at a fixed module location versus tracking the moving geometry of a cloud shadow edge. None of the studies evaluate model robustness under actively changing cloud cover; the prevailing strategy, reflected in the irradiance-restriction approach, is to exclude unstable acquisition windows rather than detect and correct for the resulting artifacts.

4.4. Data augmentation

Limited training data and class imbalance are pervasive challenges in supervised learning for PV fault detection. Data augmentation artificially expands training sets through label-preserving transformations, enhancing model generalisation and robustness to input variations. Zhang et al. highlight that dataset balancing via augmentation improved minority-class recognition and overall generalisation in PV fault classifiers [50]. Classical geometric augmentations include rotation (typically 180°), horizontal and vertical flipping, scaling (0.8–1.2 nominal resolution), and translation (10–20% of image dimensions). These transformations simulate variations in panel orientation, camera position, and viewing distance encountered during operational deployment. For PV applications, rotation augmentation is particularly valuable given the arbitrary orientations of panels within aerial frames. Furthermore, photometric augmentations manipulate pixel intensities to simulate illumination and atmospheric variations. Brightness adjustment (20–30% intensity shift), contrast scaling (0.7–1.3 multiplicative factor), and colour jittering (independent perturbation of RGB channels) enhance model robustness to lighting conditions. Gaussian noise injection and blur kernels simulate sensor noise and atmospheric degradation, preventing overfitting to pristine training images. Modern augmentation paradigms extend beyond classical geometric and photometric transforms, operating in feature or label space to improve generalisation. Generative Adversarial Networks (GANs) have been widely employed to synthesise realistic thermal, RGB, and

electroluminescence (EL) images of PV modules exhibiting diverse fault types [51], [52]. By learning the underlying distribution of real defect images, GANs can generate high-fidelity synthetic samples that enrich minority classes and mitigate data scarcity. In PV fault detection, GAN-based augmentation has been shown to significantly improve classification accuracy and recall for rare defects such as microcracks, hotspots, and partial shading. Romero et al. explicitly generated large synthetic EL datasets using a deep convolutional generative adversarial network (DCGAN). They produced 10,000 new EL cell images to extend a small real set and validated the images (IS=2.3, FID=15.8) to ensure realism, then used them to train machine learning models. This synthetic EL data can boost CNN training when measured data are limited [53]. Lekavičius et al. used a pix2pix GAN to synthesise overhead PV images for segmentation. By injecting 60% GAN-generated satellite/UAV images into training, their DeepLabV3 model saw IoU improve by +2.0% and F1 by +1.46% compared to classical augmentation [54]. GANs learn to reproduce the statistical distribution of their training images, not the physical mechanisms that produce a fault. Inception Score and Fréchet Inception Distance—metrics used to validate the synthetic data—quantify perceptual similarity to real images using features from a network trained on natural photographs; however, they do not verify whether a synthesised crack follows the fracture geometry of silicon under thermomechanical stress or whether a synthesised hotspot reproduces the thermal gradient physics of a shunted cell. A generator can therefore produce images that are statistically realistic according to these metrics while containing artefacts or feature combinations with no physical counterpart, and a downstream classifier trained on that data has no mechanism to distinguish a genuine fault signature from a generator-specific artefact correlated with the fault label.

This becomes measurable where accuracy is evaluated on held-out real data rather than performance validated on GAN-generated samples themselves. A Wasserstein GAN combined with a CNN classifier reached 86.98% test accuracy on an 11-class thermal fault taxonomy, compared to 89.83% for the same CNN paired with a denoising diffusion probabilistic model, indicating that synthetic samples from a GAN are not interchangeable with those from other generative approaches even when both serve the same augmentation purpose [51]. For minority class augmentation, a GAN trained to limited

convergence or on a small number of real class examples tends to reproduce a narrow subset of the true fault distribution rather than its full variability; thus, a dataset that appears larger after augmentation may not be meaningfully more diverse than before. None of the GAN-augmentation studies discussed above report accuracy on a real-only, GAN-excluded test set separately from performance on a test set drawn from the augmented distribution, which would isolate whether reported gains reflect genuine feature learning or partial overfitting to generator-specific characteristics.

Mixing-based augmentation techniques, including Mixup, CutMix, and Mosaic, create new training samples by combining multiple images or image regions. Mixup generates linear interpolations of image pairs and their corresponding labels, encouraging

Mixing-based augmentation techniques, including Mixup, CutMix, and Mosaic, create new training samples by combining multiple images or image regions. Mixup generates linear interpolations of image pairs and their corresponding labels, encouraging smoother decision

boundaries and reducing overfitting. CutMix and Mosaic, in contrast, replace or assemble image patches from different samples, forcing the model to focus on localised fault features rather than global appearance. These techniques have proven effective in PV fault detection tasks, particularly for object detection and segmentation in EL and thermal imagery, where defects often occupy small spatial regions. Jia et al. incorporated Mosaic and MixUp augmentation into a YOLOv5 detector for EL images. Mosaic concatenates tiles from multiple images, while MixUp linearly blends image pairs [55]. These augmentations increased visual variability and improved detection accuracy on microcrack and finger-line faults. In semi-supervised EL segmentation, Jha used CutMix, where random patches of an image are replaced with patches from another image (β -distributed). When combined with colour jitter, blur, and other augmentations, CutMix greatly diversified training views of cracks and interconnect failures. Such “mixing” augmentations effectively enforce the model to focus on different fault patterns and have been shown to improve robustness under limited labels [56].

Table summarises the applicability of each preprocessing technique across the three imaging modalities covered in this review, along with a qualitative rating of its inference-time computational cost.

Table 4. Preprocessing Techniques by Modality and Inference-Time Cost

Technique	RGB	Thermal	EL	Inference-time cost
CLAHE	✓	✓	Limited (low dynamic range)	Low -linear, tile-bounded
Wavelet denoising (DB4)	✓	Limited	✓	Moderate - multi-level decomposition overhead
Morphological filtering	✓	✓	✓	Low - scales with kernel size, not image content
Bilateral filtering	✓	✓	Rarely used	Higher - naive cost grows with window ²
Irradiance/temperature normalisation	-	✓	-	Low, but requires synchronised sensor metadata
Wind speed correction	-	✓	-	Low, but requires synchronised metadata

4.5. Computational cost of preprocessing

The preprocessing techniques discussed do not share a latency profile, and the distinction matters for real-time deployment. Data augmentation is a training-time cost only; once the model is trained, none of it affects the inference pipeline. Colour enhancement, denoising, and environmental normalization differ because a deployed model expects input distributed like its training data. Every field image must pass through the same preprocessing chain before inference, and this cost recurs on every frame.

Within that inference-time group, cost varies substantially. CLAHE operates on local tiles with bounded per-tile histogram computation, keeping the cost close to linear in pixel count and typically negligible relative to a CNN forward pass. Morphological operations scale with the size of the structuring element rather than image content and remain inexpensive for the small kernels used for

crack or delamination boundaries. Naive bilateral filtering does not share this property, as its cost grows with the square of the filter window and can become a bottleneck at UAV thermal imagery resolutions unless a separable or grid-based approximation is used. Wavelet-based denoising sits between the two: each decomposition level runs in linear time, but multi-level decomposition plus inverse reconstruction adds a fixed multiplicative overhead per image that a CLAHE-only pipeline does not incur.

Few reviewed studies isolate this cost. When full-pipeline latency is reported, it typically appears as a single end-to-end figure rather than broken down by stage: a hybrid framework combining four-level discrete wavelet decomposition with handcrafted and deep feature fusion reports 19–21 ms total inference time per image at a deployment size of roughly 87–107 MB [57].

Hand-engineered preprocessing versus

allowing the network to learn from minimally processed input remains under question. One direct comparison in this corpus pits a feature-engineered SVM pipeline against a CNN operating on lightly processed images, finding that the SVM pathway has the lower per-inference computational cost of the two, while the CNN achieves higher accuracy and is recommended for real-time, large-scale monitoring despite that cost gap [58]. This suggests that accuracy, not preprocessing cost alone, has driven architecture choice in the reviewed literature. However, because most studies report only accuracy with no timing breakdown, it is not possible to determine how large that gap typically is or at what point it would become prohibitive on UAV-class hardware.

5. Machine learning algorithms for fault detection

The transition from traditional rule-based diagnostic approaches to data-driven machine learning (ML) and deep learning (DL) methodologies has fundamentally transformed photovoltaic fault detection capabilities. Machine learning algorithms enable automated feature extraction, pattern recognition, and classification across diverse imaging modalities, offering scalability and accuracy that manual inspection cannot achieve. This section provides a comprehensive overview of ML approaches employed in PV fault detection, organised from classical methods to state-of-the-art deep architectures, with particular emphasis on their application to thermal, visible, and electroluminescence imagery. The technical details of the machine learning approaches presented in this section are listed in

Table .

Table 5. Summary of Photovoltaic Fault Detection Methods

Ref.	Fault category	Algorithm	Task formulation	Score	Metric
Thermal infrared (IR)					
[59]	Hotspot	HotSD-Net	Classification	97.9%	Utility score
[60]	Hotspot, Multi-cell, Microcrack, Diode fault, Shading, Shadowing, Soiling, Vegetation-induced defects, Offline no-power module	CNN - Adaptive Attention Pyramid (AAPN-Tiny)	Classification	96% / 92%	Acc. (binary / 12-class)
[61]	Hot spot, Open circuit, Substring in open circuit, Short circuit, Heated junction box	Mask R-CNN + U-Net	Segmentation	96%	mAP
[62]	Hotspot, shadowing, bird dropping, structural damage (IR + RGB)	YOLOv5 + ResNet	Detection	95%	Acc.
[63]	Sand accumulation, Dirt, Degradation, Junction box overheating	TinyCNN	Classification	94.8%	Acc.
[64]	Microcrack, hotspot, shadowing, soiling, vegetation	TinyTripleNet (lightweight CNN)	Classification	93.07%	Acc.
[65]	Ribbon burns, hotspots, bypass diode faults	DCNN + SVM / XGBoost	Classification	89%	Acc.
[66]	Hotspot	YOLOv8s	Detection	89.1%	F1
[33]	Hotspot	Modified MATL-CNN	Detection	--	Not reported
Electroluminescence (EL)					

Ref.	Fault category	Algorithm	Task formulation	Score	Metric
[67]	Microcrack, broken components	SqueezeNet	Classification	96.29%	Acc.
[68]	Finger crack, Black-core, Thick-line, Horizontal-dislocation, Short-circuit, Vertical-dislocation, Star-crack, Printing error	CNN-EML (SVM, RF, KNN ensemble)	Classification	94%	Acc.
[34]	Black core, Microcrack, Edge fault	VGG16	Classification	93%	Acc.
[52]	Microcrack, finger interruption	Transfer learning + GAN augmentation	Classification	90.13%	Acc.
[31]	Defective vs functional cells - binary (ELPV dataset)	MobileFFT (MobileViT + FFT-PC module)	Classification	87.4%	Acc. (F1: 0.858)
[69]	Microcrack, corrosion	MobileNetV1 / MobileNetV2	Classification	88.44% / 76.38%	Acc.
[56]	Contact near solderpad, Contact front grid interruption, Microcrack, Contact corrosion, Crack Resistive, Interconnect highly resistive	PV-SSS - Mean Teacher paradigm	Segmentation	88.9%	Precision
[55]	Microcrack, Think line, Black core, Finger scratch, Star crack, Dislocation, Short circuit	Improved VarifocalNet	Detetcion	85.7%	mAP
[70]	Microcrack, Finger, Black Core, Thick line, Star Crack, Corner, Fragment, Scratch, Dislocation, Printing Error, Short circuit	YOLO-CNN combined transformer (CCT)	Detection	77.9%	mAP50.
RGB / Visible					
[71]	Glass breakage, burn marks, Snail trail, Discolouration, Delamination	Ensemble DNN + Random Forest	Classification	99.68%	Acc.
[72]	Electrical damage, Physical damage, Dust, Bird-Drop, Snow-Covered	SolarFaultAttentionNet (multi-path CNN + InceptionV3)	Classification	99.14%	Acc.
[73]	Shading, Microcrack, delamination, discoloration	Vision Transformer (ViT)	Classification	93.73%	Acc.

Ref.	Fault category	Algorithm	Task formulation	Score	Metric
[31]	Dust, Bird-dropping, Physical damage, Electrical damage	MobileSFOT (MobileViT + slant transform module)	Classification	82.8%	Acc. (F1: 0.840)
[74]	Dust, bird-dropping (soiling)	SDS-YOLO	Detection	78.1%	F1

Acc. = accuracy; mAP = mean average precision; F1 = F1-score. Entries marked "—" did not report a numerical performance value.

The scores in the table span a 21.8%-point range, from 77.9% to 99.68%, and this spread should not be read as a ranking of algorithmic capability. For example, rows [70] and [71] both report accuracy, yet the former reaches 99.68% on a five-class RGB taxonomy (glass breakage, burn marks, snail trail, discolouration, delamination), while the latter reaches 77.9% on an eleven-class EL taxonomy that includes visually adjacent subtypes such as thick line, star crack, and finger crack. Task granularity and class separability differ as much between these two studies as the underlying models do. More broadly, each entry is drawn from a distinct, typically private dataset with its own acquisition protocol, class balance, and train/test split, and none of the reported figures were obtained under a shared benchmark or cross-validated against an external test set. Under these conditions, the table should be read as a compilation of self-reported, study-specific results rather than a comparative leaderboard.

Task formulation across the 22 entries in

Table is not distributed by fault type or modality alone, but by what a maintenance decision downstream of detection actually requires. Segmentation appears only twice, and both cases share a structural feature: the fault must be localised to a specific sub-component within a complex assembly like a solder joint, a junction-box element, or a particular interconnect where knowing which component failed matters as much as knowing that one did. Detection accounts for five or six entries, concentrated wherever a fault manifests as multiple discrete, countable instances within a single frame: several distinct crack subtypes as separate bounded objects, or discrete soiling spots where a single whole-image label would discard information a technician needs to act on any one instance. Classification dominates the remaining fourteen entries, covering both binary framings (functional versus defective cells) and coarse multi-class panel-level conditions (discolouration, delamination, snail trail) where the operative question is which condition a panel exhibits, not where within the image evidence for it appears.

The clearest illustration that this is a design

choice rather than a fault-intrinsic property is that EL crack detection appears under all three formulations across different studies in this corpus as multi-class classification, as bounding-box detection, and as the sub-component localisation problem underlying segmentation using substantially overlapping fault vocabularies. The determining factor is not the fault itself but what the study defines as an actionable output, reinforcing the localisation-first diagnostics priority raised elsewhere in this review: task formulation should be selected according to what a maintenance action requires a technician to know, not defaulted to classification because it is the simplest framing to report accuracy for.

5.1. Traditional machine learning methods

Before the widespread adoption of deep learning, traditional supervised learning algorithms such as Support Vector Machines, Random Forests, Decision Trees, k-Nearest Neighbors (kNN), and ensemble methods like XGBoost formed the foundation of automated PV fault detection systems. These classifiers operate on handcrafted features such as colour histograms, SIFT/SURF descriptors, colour moments, and visual vocabularies derived from K-Means clustering, extracted from thermal, visible, or electroluminescence images via conventional computer vision techniques or shallow neural networks [24], [74], [75].

Comprehensive reviews confirm their effectiveness, reporting median classification accuracies of 92.1% (interquartile range: 89.3–94.2%; F1 score: 0.90) on PV imagery including electroluminescence, infrared, and RGB images. They are particularly competitive on clean signals but are sensitive to imbalance and drift [76]. Traditional classifiers offer several practical advantages in PV applications. They require relatively small training datasets compared to deep networks, making them suitable for

scenarios with limited labelled fault examples [77]. Tree-based ensemble methods (Random Forest, XGBoost) provide interpretable feature importance rankings, enabling engineers to

understand which image characteristics most strongly indicate specific fault types [78]. This interpretability facilitates root-cause analysis and supports maintenance decision-making processes.

However, traditional methods exhibit fundamental limitations. Their performance depends significantly on the quality of feature engineering, with handcrafted features often failing to capture the complex, hierarchical patterns present in high-dimensional image data. Consequently, traditional classifiers generally underperform compared to end-to-end deep learning approaches on image-rich modalities such as thermal infrared and electroluminescence imaging [79].

5.2. Deep learning for image classification

Convolutional Neural Networks (CNNs) have emerged as the dominant paradigm for PV fault classification tasks, leveraging hierarchical feature learning to automatically extract discriminative patterns from raw image data. Multiple CNN architectures, including VGG, ResNet, Inception, EfficientNet, DenseNet, and MobileNet families, have been successfully adapted and fine-tuned for diverse PV imaging modalities (Kardas, & Atmaca, 2025; Sridharan et al., 2022; Nguyen-Vinh et al., 2023) [24], [76]. VGG networks, characterized by their deep, uniform architecture with small 3×3 convolutional filters, have been extensively applied to PV fault detection. VGG-16 and VGG-19 variants demonstrated high performance in detecting thermal hotspots, with VGG-16 showing particularly strong reliability on preprocessed thermal images compared to more recent architectures such as EfficientNet [80]. The architecture's depth enables learning of complex thermal patterns associated with electrical mismatch and cell degradation. VGG16 pre-trained on ImageNet has also been employed for aerial UAV imagery classification, successfully identifying multiple defect classes including burn marks, delamination, and snail trails [22]. Similarly, ResNet architectures, known for their residual connections that mitigate vanishing gradients in deep networks, have shown excellent performance in classifying various PV defects from electroluminescence images, achieving high accuracy in distinguishing between cracks, finger interruptions, and inactive regions (Jha et al., 2025) [56]. A combination of ResNet152 and Xception achieved 96.17% accuracy in binary classification and 92.13% in multiclass tasks for PV cell defect detection [52]. EfficientNet architectures employ compound scaling to balance network depth, width, and resolution, theoretically optimizing

computational efficiency. However, empirical results in PV applications show mixed performance. For instance, EfficientNetV2 achieved only 72.75% accuracy on visible-image datasets, where other architectures exceeded 90%. These results suggest that the compound scaling approach may not consistently translate to optimal performance in specialized PV fault detection tasks, possibly due to domain-specific characteristics that differ from natural image datasets used in architecture development. MobileNet architectures prioritize computational efficiency through depthwise separable convolutions, making them suitable for deployment on resource-constrained platforms such as UAV-mounted systems. MobileNetV3 achieved 92.04% accuracy in classifying PV panel defects in aerial imagery, demonstrating its viability for real-time inference in remote inspection scenarios. Vyas et al. demonstrate that lightweight architectures can maintain competitive performance while enabling real-time inference [73].

5.3. Vision transformers

While CNNs have long been the backbone of image processing, Vision Transformers (ViT) have recently emerged as a powerful alternative for PV defect detection, offering enhanced capabilities in capturing long-range dependencies and global contextual information across images. In a controlled comparison on the same visible-spectrum dataset, a ViT model achieved 93.73% accuracy (F1: 93.72%), outperforming EfficientNetV2 (72.75%), ResNet50v2 (91.55%), and MobileNetV3 (92.04%) evaluated under identical conditions [73].

The accuracy gap between ViT and MobileNetV3 in this study, for example, is 1.69 percentage points; the parameter gap between the architecture families is not small by comparison. A base-sized ViT typically contains on the order of 86 million parameters, roughly sixteen times MobileNetV3-Large's 5.4 million, and the quadratic cost of self-attention in the number of image patches causes both compute and memory requirements to grow faster with input resolution than a CNN's local convolutions. On a UAV-mounted platform, this difference in scale often determines feasibility. A lightweight hybrid approach exemplifies the other end of this trade-off: a wavelet/BoVW/CNN-feature fusion model reports 19–21 ms inference per image and an 87–107 MB deployment footprint at 94.35% accuracy, suggesting that classical feature fusion or distilled architectures remain competitive on constrained

hardware even where a full ViT reports marginally higher headline accuracy elsewhere [57].

5.4. Object detection methods

While classification networks assign labels to entire images, object detection frameworks simultaneously localise and classify multiple fault instances within a single image, a critical capability for large-scale PV inspections where multiple modules or cells may exhibit different fault types. Object detectors output bounding boxes with associated class probabilities, enabling spatial mapping of defects across PV arrays. YOLO architectures have demonstrated remarkable efficacy in real-time PV fault detection applications due to their single-stage detection approach that frames object detection as a regression problem, enabling fast inference speeds suitable for UAV-based inspections. Multiple YOLO variants, including YOLOv5, YOLOv7, YOLOv8, and YOLOv9, have been adapted for detecting various PV defects such as hotspots, cracks, and snail trails in thermal and visible spectrum imagery. Performance metrics for YOLO-based systems vary significantly depending on the specific dataset. A recent study employing YOLOv8 achieved an accuracy of 95% in classifying faults from thermal images [62]. This performance demonstrates the potential of YOLO architectures when provided with high-quality, diverse training data. While YOLO architectures dominate the literature, other object detection frameworks such as the R-CNN family (Faster R-CNN, Mask R-CNN), Single Shot Multibox Detector (SSD), and RetinaNet have also been adapted for PV defect detection, offering diverse approaches to balancing detection accuracy with computational efficiency [81]. For instance, Faster R-CNN, while generally slower, can achieve higher localisation precision for complex defect geometries due to its two-stage proposal generation and classification mechanism [82].

5.5 Edge deployment and hardware constraints

UAV-based inspection imposes constraints that accuracy metrics alone do not capture; onboard compute is limited to a few watts of power, flight time caps total inspection duration, and bandwidth to a ground station is often insufficient for raw image transmission, making onboard inference a practical requirement. Several PV-specific studies in this review report hardware-validated deployment figures rather than theoretical estimates. AAPN-Tiny, with 0.8 million parameters, achieves 40.2 ms inference time per image on a Google Coral Edge TPU [60].

TinyTripleNet, with 0.9 million parameters, reduces inference time by 75% and memory usage by over 80% relative to 29 baseline models including VGG16, ResNet50, and MobileNet, reaching 4.8 ms per image and a 1.7 MB runtime memory footprint on the same TPU class [64]. YOLOv11n-GhostLite, designed for EL-image defect localisation, has 2.34 million parameters and 6.2 GFLOPs, operating at approximately 2 W on a Coral Edge TPU while maintaining 97.2% mAP@50 [83]. A separate YOLOv8/MobileNetV2 hybrid for EL-image classification was deployed on an NVIDIA Jetson TX2 rather than a TPU, achieving 80.40% accuracy [84].

Quantisation is not metric-neutral. One Raspberry Pi 4 deployment explicitly compares three TensorFlow Lite optimisation strategies (simple conversion, dynamic-range quantisation, and float16 quantisation) and reports the resulting trade-off between model size and classification accuracy rather than treating quantisation as a black-box deployment step [85]. Most studies report a single post-hoc accuracy figure with no indication of whether it was measured on the full-precision training checkpoint or the quantised, edge-compiled model actually intended for field use. The two are not interchangeable, and the gap between them is frequently larger than the accuracy differences separating competing architectures. This represents a second, hardware-specific version of the comparability problem raised in the

Table discussion above.

5.6. Ensemble and hybrid approaches

Addressing the limitations of individual models, researchers are increasingly exploring ensemble and hybrid approaches to enhance the robustness and accuracy of PV fault detection systems. By combining the strengths of multiple models, these methods can take several forms: (i) combining multiple CNN architectures with different inductive biases, (ii) integrating deep features with traditional classifiers, and (iii) developing custom detector ensembles that augment baseline architectures [72].

Hybrid architectures that combine deep convolutional neural networks (DCNNs) for feature extraction with traditional classifiers for final decision-making represent an intermediate approach. Mellit et al. conducted a comparative study on infrared thermography datasets, demonstrating that DCNN features classified by SVM or XGBoost achieved approximately 87% accuracy, which is comparable to end-to-end DCNN performance [86]. In the same study, DCNN-RF, DCNN-k-NN, and DCNN-DT combinations yielded lower accuracies, highlighting the variability in the effectiveness of hybrid methods. These hybrid pipelines can be advantageous when training data is limited or when

interpretability requirements favor traditional classifiers, although they introduce additional complexity into the processing pipeline. Custom detector ensembles that modify baseline architectures represent another approach. The GELAN architecture, for example, incorporates improved gradient flow mechanisms that outperformed standard YOLO baselines, achieving a Mean Average Precision (mAP) of 70.4% [87]. Such architectural innovations demonstrate that domain-specific modifications to established frameworks can yield performance improvements tailored to PV fault detection characteristics.

5.7 foundation models, self-supervision, and diffusion-based methods

A cluster of methods outside the CNN/ViT/object-detection mainstream is beginning to appear in the literature, unevenly and mostly very recently. General-purpose multimodal foundation models, large vision-language systems trained without PV-specific data, have been tested zero-shot against RGB, EL, and IR fault imagery simultaneously, with Gemini 1.5 Pro reaching an F1-score of 0.97 across bird-dropping, crack, and hotspot detection without any task-specific fine-tuning [88]. This differs architecturally from segmentation-specific foundation models such as SAM or contrastive vision-language models such as CLIP, neither of which appears in the reviewed corpus, but demonstrates the same underlying shift toward general-purpose pretrained systems displacing task-specific training.

Self-supervised representation learning addresses labeling costs. A SimSiam-based approach to thermal hotspot detection and isolation achieves accuracy up to 97% and a Dice coefficient of 0.736 using self-supervised pretraining followed by fine-tuning, which is competitive with supervised baselines while requiring substantially fewer labeled examples [75]. This result is situated within a synthesis organized around supervised methods, since its central contribution is precisely reducing dependence on the labeled data that classification and detection pipelines elsewhere in this review assume are available.

Diffusion models appear in two distinct roles. As an augmentation source, a denoising diffusion probabilistic model combined with a CNN classifier outperforms both geometric transformation and GAN-based augmentation for an 11-class thermal fault taxonomy, a result extended in a UAV-oriented multi-stage pipeline combining DDPM augmentation with YOLOv8-based defect localization [51] [89]. More architecturally distinct is the use of diffusion

models to reconstruct a healthy counterpart of a faulty EL image, deriving a spatial attention map from the fault/reconstruction residual. This approach reports single-domain accuracy of 95.9% (detection) and 97.0% (classification), compared to 75.5% and 53.1% respectively under cross-domain evaluation—a 43.9%-point drop in classification accuracy, which is among the starkest domain-shift results in this review’s corpus, and a concrete illustration of the generalization problem raised in the discussion below [90].

No study in the corpus integrates physical constraints like thermal diffusion, electrical continuity, or fault-propagation physics directly into a neural architecture for image-based PV diagnostics. Physics-informed neural networks have seen substantial adoption in adjacent engineering domains, typically by embedding governing equations as additional loss terms rather than relying on data alone to learn physical consistency. Applied here, this might mean penalizing thermal predictions inconsistent with the heat equation or fault localizations inconsistent with expected current-path continuity across a cell string.

6. Discussion

The bibliometric mapping confirms that image-based PV fault diagnosis has shifted from a niche topic to a rapidly expanding research area, with clear consolidation around deep-learning and computer-vision terminology and a progressive move toward deployment-oriented objectives. This macro-level view aligns with the technical literature reviewed in the subsequent sections: early work frequently framed PV inspection as a classification problem under constrained datasets, whereas more recent studies increasingly adopt detection, segmentation, and anomaly-based formulations that better reflect operational PV inspection needs. Importantly, the bibliometric evolution does not merely indicate increased publication volume; it also highlights a thematic convergence toward “fault detection” accompanied by greater attention to diagnostic specificity (e.g., crack-related keywords and modality-specific themes such as EL imaging). When interpreted alongside the reviewed methods, this evolution can be understood as a transition from proof-of-concept recognition toward pipelines designed for spatial localisation, fine-grained defect taxonomy, and maintenance relevance.

A recurring finding across the reviewed studies is that the formulation of the diagnostic task is often as decisive as the chosen algorithm itself, particularly in determining the practical utility and

deployment feasibility of a given approach. Binary anomaly detection can be attractive for field inspection because it reduces labelling effort and can flag previously unseen faults, yet it is susceptible to domain shift and may offer limited interpretability for maintenance decisions if it cannot distinguish fault types. Multi-class detection, by contrast, aligns with fault taxonomies but typically demands balanced class coverage and consistent labelling standards—conditions that are difficult to meet in real PV plants where fault frequencies are uneven and confounded by environmental variability. Detection and segmentation methods occupy a practical middle ground because they provide localisation (critical for actionable maintenance) and can distinguish multiple faults within the same frame; however, they increase annotation cost and sensitivity to resolution, camera geometry, and defect size.

Consequently, the selection of an appropriate defect detection methodology must weigh the trade-offs between annotation complexity, the granularity of diagnostic output, and the robustness required for deployment across diverse environmental and operational contexts.

Imaging modality acts as a second organizing constraint that directly influences what models can learn and what a deployment can realistically achieve. RGB inspection is operationally attractive due to low cost and ease of acquisition, but its diagnostic ceiling is limited by lighting variability and the fact that many electrically driven defects lack a stable visual signature. Thermal IR improves fault saliency for mismatch- and dissipation-driven anomalies (e.g., hotspots), supporting anomaly-centric pipelines; nonetheless, thermal signatures are strongly modulated by irradiance, wind, and thermal transient effects, which makes robust acquisition protocols and metadata-aware normalisation important for generalisation. EL imaging provides the highest sensitivity to cell-level discontinuities and micro-defects and naturally supports fine-grained segmentation/classification; yet its controlled acquisition requirements limit routine field adoption. A key implication is that “best-performing” models in the literature may reflect the ease of signature capture as much as algorithmic superiority, which complicates cross-paper comparisons unless acquisition conditions and evaluation protocols are explicitly documented.

6.1. Data limitations and evaluation practices

Despite rapid model innovation, the dominant bottleneck remains data, including the scarcity of labelled faults, class imbalance, and cross-site

diversity. Many reported performances are obtained from narrowly scoped datasets (single plant, single camera setup, or restricted weather conditions), which inflate apparent generalisation and can mask deployment failure modes. This challenge is amplified in PV inspection because domain shift is structural: changes in module technology, mounting geometry, camera altitude/angle, irradiance regimes, and background clutter all alter the data distribution. As a result, the field’s movement toward detection/segmentation and toward “maintenance-relevant” outputs will be constrained until datasets and protocols support robust validation across conditions.

Evaluation practices also require stronger alignment with operational objectives. Accuracy remains common in the literature, yet it can be misleading under imbalance and does not express localisation quality for detection/segmentation tasks, as the accuracy range compiled in

Table illustrates. For deployable inspection, metrics should be chosen to reflect real costs (false alarms versus missed faults), spatial precision (IoU/mAP where appropriate), and performance stability under domain shift (cross-plant or cross-condition testing). Equally important is reporting latency and computational requirements for UAV/edge scenarios, since high-performing models may be impractical if inference cannot be achieved within operational inspection constraints. This aspiration is rarely matched by reported detail. When false-positive and false-negative behaviour is actually analysed rather than folded into a single accuracy or mAP figure, two failure patterns recur across otherwise unrelated studies. A domain-adaptation study using unsupervised anomaly detection across nine source-target PV plant combinations reports 79.4% and 77.1% correctly classified normal and anomalous images, respectively, at a fixed decision threshold, with AUROC ranging from 73.3% to 96.6% depending on the plant pair, and explicitly notes that most misclassified anomalies are low-severity cases such as hot diodes and small hotspots [91]. A separate soiling-detection study reports the same pattern for a different fault type and modality: small, low-contrast bird droppings are harder to detect than larger dust accumulations, requiring a dedicated attention mechanism specifically to close this gap [92]. False positives, by contrast, more often trace to acquisition geometry than to genuine visual ambiguity between fault classes. An online anomaly-detection framework for raw aerial thermography reports that its false alarms were attributable almost entirely to irregular image patterns at the edges of a PV array or to extreme

non-orthogonal viewing perspectives, rather than to any confusable fault signature, characterising this as a manageable error source that an appraiser can screen offline [93].

Together, these findings suggest that models tend to miss small or subtle faults before they miss large ones, and tend to generate false alarms on acquisition artefacts before confusing one fault type for another. Neither pattern is visible in a single aggregate accuracy figure. The domain-adaptation result further shows that this profile is not stable across deployment sites, where an AUROC range exceeding 23 percentage points across plant pairs—using the same architecture—means that a reported false-positive or false-negative rate from one site offers limited guarantee at another. Beyond the minimum reporting standard proposed in the Conclusion, task formulation should determine which metric is primary rather than treating accuracy as a default. Classification tasks under class imbalance are better served by F1-score or balanced accuracy than raw accuracy; detection tasks require mAP at a specified IoU threshold; segmentation tasks require IoU or Dice coefficient, the latter already used for hotspot isolation in this corpus; anomaly detection tasks, particularly under domain shift, are better characterised by AUROC across multiple threshold settings than by accuracy at a single fixed threshold.

On standardised benchmarks, PVEL-AD appears as an evaluation dataset in ten of the reviewed studies, making it the standard benchmark for EL-based defect detection; ELPV and PV Multi-Defect each appear in two studies; Kaggle-sourced RGB fault datasets, less standardised and more heterogeneous in composition, appear in four. This pattern suggests that the field does not lack a candidate standard so much as it lacks consistent adoption of the one already gaining traction. Future work reporting results on PVEL-AD or explicitly justifying a departure from it would materially improve cross-study comparability without requiring a new benchmark to be built.

On statistical comparison, k-fold cross-validation is used by a meaningful share of reviewed studies, providing a basis for variance estimation that a single train/test split does not. However, no study in this corpus reports a formal significance test when comparing competing architectures; no paired t-test, Wilcoxon signed-rank test, or McNemar's test appears anywhere in the corpus. Given that per-fold accuracy distributions are typically based on small samples and are not guaranteed to be normal, a paired

Wilcoxon signed-rank test across cross-validation folds is a more defensible default than an unpaired comparison of point estimates. The absence of such tests throughout the reviewed literature means that claims of one architecture "outperforming" another are rarely supported by a test of whether the difference exceeds what fold-to-fold variance alone would produce.

6.2. Toward operational PV diagnostics

A notable trend emerging from both bibliometric themes and technical methods is the gradual expansion from "model accuracy" toward system-level integration, including UAV-based acquisition, pipeline automation, and decision-support framing. However, many studies still treat the vision model as an isolated component, with limited attention to upstream and downstream factors that dominate real-world reliability: calibration, image quality assurance, georeferencing, module-level aggregation, and the mapping from detected anomalies to maintenance actions. Bridging this gap requires tighter coupling between CV outputs and PV engineering constraints: linking detections to expected power-loss impacts, prioritising interventions, and validating whether the identified fault type is actionable under field conditions. In this context, uncertainty-aware predictions and explainability are not optional "extras"; they can reduce unnecessary maintenance by communicating confidence and highlighting failure modes (e.g., ambiguous hotspots caused by transient shading or wind-driven cooling).

Two distinct capabilities are often conflated under this heading: post-hoc explanation of a given prediction and calibrated uncertainty over that prediction. This review's corpus has precedent for the former but not the latter. One study directly compares GPT-4o against ResNet for aerial infrared fault classification, finding GPT-4o offers natural-language explanations of its classifications, describing why an anomaly was flagged, at the cost of substantially higher computational demand than ResNet's opaque but efficient prediction [94].

The studies that do close this loop converge on a common architecture rather than a single implementation. An edge-computing system mounted on UAVs performs onboard inference and uses unsupervised clustering to flag defect types not yet present in the training database, addressing the constraint that new fault signatures continually appear in field data [95]. At the plant-diagnostics level, a modular software prototype combining aerial infrared thermography with deep learning explicitly targets this same integration gap, citing

data fusion with SCADA outputs and coupling with maintenance and end-of-life management tools as its next development stage rather than a solved problem [96]. Digital-twin platforms extend the same logic into simulation; one built for autonomous aerial PV monitoring generates synthetic imagery to train segmentation models under controlled, repeatable conditions before field deployment [97]. No reviewed study reports calibrated uncertainty like Monte Carlo dropout, deep ensembles, or conformal prediction for image-based PV fault classification specifically. This would mean a model outputting not "hotspot, 94% confidence" alone but a calibrated interval reflecting how often its 94%-confidence predictions are actually correct, which raw softmax outputs do not guarantee; an inspector could then triage low-confidence flags for manual review.

The image-to-SCADA link is therefore treated as a stated future direction rather than a modelled interface. Digital-twin frameworks are considerably more mature on the electrical side, where array-level twins compare real-time power signals against a simulated baseline to localise and classify faults; image-based diagnostics rarely feed that same comparison loop. Cyber-physical framings of PV monitoring date back at least to early combined sensing-and-actuation architectures [98], yet the vision component in these systems remains loosely coupled to the control layer rather than embedded within a shared physical-virtual state.

Several research directions emerge as particularly salient when combining the bibliometric evidence with the technical synthesis. A first priority is the development of standardised datasets and cross-site benchmarks. Progress is currently constrained by limited data diversity and inconsistent labelling practices; multi-plant datasets enriched with operational metadata (e.g., irradiance, wind conditions, and camera geometry) and shared annotation guidelines would enable fairer comparisons and more realistic tests of generalisation.

Furthermore, there is growing motivation for multimodal and multiscale fusion. When available, combining RGB, thermal IR, and EL can provide complementary fault signatures, while integrating imagery with operational data (e.g., SCADA) may improve diagnostic specificity. The key challenge is to design fusion strategies that are interpretable and that preserve the central role of image evidence rather than obscuring it within black-box aggregation.

A third gap relates to localisation-first

diagnostics. Many real maintenance decisions require not just identifying a fault category but also knowing where it occurs. This supports wider adoption of detection and segmentation pipelines that deliver actionable localisation, together with annotation-efficient strategies such as weak supervision, active learning, and physics-informed or synthetic data generation to reduce labelling costs.

Finally, the field is moving from detection toward decision support. Beyond reporting performance metrics, future studies should connect model outputs to fault severity, expected energy loss, and recommended interventions. This shift naturally aligns with predictive maintenance and digital-twin frameworks, where image-based diagnostics can serve as an evidence layer for actionable asset management decisions, but realising it requires defining what an operator-facing output actually contains, beyond a bounding box and a class label. Four elements recur in what plant operators need from an inspection system, and reviewed studies typically supply at most one or two: (i) fault classification mapped to an existing standard taxonomy rather than a model-specific label set, so that outputs are comparable across vendors and inspection cycles; IEC TS 62446-3 already defines such a taxonomy for thermographic inspection, making this a mapping problem, (ii) geolocation tied to a plant asset register (module, string, and inverter identifiers) rather than pixel coordinates within a single image, since a detection is only actionable once a technician can find the physical module it refers to, (iii) a severity or priority signal grounded in estimated power-loss impact feasible in principle, as EL-based PID assessment already yields power-loss estimates in the 25-75% range from EL current analysis, though this kind of fault-to-energy-loss mapping remains fault-specific rather than general, and (iv) a recommended action or triage category (monitor, schedule inspection, immediate intervention) consistent with existing maintenance workflows. Most studies stop at detection or segmentation output with a class label and confidence score, leaving the translation into an operator-actionable record as unstated future work rather than a design requirement.

The regulatory and certification path for automated interpretation is currently undefined. IEC TS 62446-3, the primary global standard for outdoor PV thermographic inspection, specifies measurement equipment, environmental conditions, and personnel qualification requirements for those interpreting the inspection, and IEC-compliant reports are what insurers, EPC

contractors, and investors currently accept as reliable documentation for warranty and fire-safety purposes. No provision in the standard addresses automated interpretation, and in the United States specifically, no equivalent body—such as NEC, ASNT, UL, or NFPA—has issued drone- or AI-specific thermography standards. An AI-based classification therefore has no defined path into the compliance chain governed by the standard: it is not currently a substitute for, nor a clearly valid input to, a certified inspector's sign-off, regardless of its measured accuracy.

7. Conclusion

This review examined AI-enabled, image-based fault detection and diagnosis for PV systems through a combined bibliometric mapping and technical synthesis. The bibliometric analysis shows a rapidly expanding research landscape that increasingly converges toward operationally actionable fault detection, while also branching into modality- and task-specific themes. Building on this context, the technical review highlighted how diagnostic performance and deployability are jointly shaped by fault type, imaging modality (RGB, IR, and EL), data preparation practices, and task formulation (classification, detection/segmentation, and anomaly detection).

Across the literature, the strongest constraints are not model availability but data quality and diversity, inconsistent labeling practices, and evaluation strategies that do not always reflect real maintenance objectives. Three of these constraints admit more concrete targets than "more data" and "better metrics." For instance, reported dataset sizes across the reviewed corpus range from 280 to 30,000 images, with widely used benchmarks such as ELPV (2,624 EL images) sitting near the middle of that range. Studies with fewer than approximately 1,000 images per class rely almost universally on transfer learning and aggressive augmentation, suggesting 1,000 images per class as a practical minimum for fine-tuning approaches and 5,000–10,000 per class as a target for training without heavy reliance on synthetic augmentation. Furthermore, multimodal fusion has been tested with a quantified ablation; parallel modality-specific backbones (e.g., one per RGB and thermal input) merged through a spatial pyramid pooling fusion layer produced a 10% improvement in detection rate over single-branch equivalents. This is a concrete, validated template; extending it to a third EL branch, where EL data are available, is architecturally straightforward but untested in the reviewed literature and remains an open empirical question. Also, a minimum reporting standard,

synthesized from the comparability failures identified throughout this review, would require: (i) task-appropriate metrics, not accuracy alone—mAP/IoU for detection and segmentation, F1 for imbalanced classification; (ii) disclosure of dataset size, class distribution, and whether the train/test split is single-site or cross-site; (iii) latency and model size whenever an edge-deployment claim is made, distinguishing full-precision from quantized measurements; and (iv) a false-positive/false-negative breakdown, not a single aggregate figure. None of these require new infrastructure to adopt; the barrier is reporting convention, not technical capability. Future progress depends on adopting benchmarks and reporting practices along these concrete lines, alongside the broader shift toward decision-support outputs that link visual evidence to fault severity, expected energy loss, and maintenance prioritization discussed in Section 6.2.

Priority directions for future work, in order of tractability: (1) adoption of PVEL-AD or an equivalent shared benchmark as a default reporting target, requiring no new infrastructure; (2) minimum evaluation reporting as specified above, likewise requiring convention; (3) validated dual-branch fusion architectures extended to a third EL branch where available; (4) operator-facing output standardization against IEC TS 62446-3's existing fault taxonomy; (5) integration of physics-informed constraints into image-based architectures, the least mature of the five and the one requiring genuinely new methodological work. Progress on the first two is achievable through reporting convention alone; the latter three require sustained collaboration between PV engineering groups, which hold the domain knowledge of fault mechanisms and access to field data, and machine learning groups, which hold the modeling expertise. This division of labor, as suggested by this review's corpus, is not yet occurring consistently, given how few reviewed studies combine deep physical grounding with state-of-the-art architectures.

References

- [1] R. Herbazi et al., « Performance evaluation and experimental validation of different empirical models for predicting photovoltaic output power », *International Journal of Ambient Energy*, vol. 43, n° 1, p. 7437– 7453, déc. 2022, doi: 10.1080/01430750.2022.2068069.
- [2] J.-M. Huang, R.-J. Wai, et W. Gao, « Newly-Designed Fault Diagnostic Method for Solar Photovoltaic Generation System Based on IV-Curve Measurement », *IEEE Access*, vol. 7, p. 70919– 70932, 2019, doi: 10.1109/ACCESS.2019.2919337.

- [3] M. Dhimish, V. Holmes, B. Mehrdadi, M. Dales, et P. Mather, « Photovoltaic fault detection algorithm based on theoretical curves modelling and fuzzy classification system », *Energy*, vol. 140, n° 1, p. 276- 290, août 2017.
- [4] E. Garoudja, F. Harrou, Y. Sun, K. Kara, A. Chouder, et S. Silvestre, « Statistical fault detection in photovoltaic systems », *Solar Energy*, vol. 150, p. 485- 499, juill. 2017, doi: 10.1016/j.solener.2017.04.043.
- [5] Q. Navid, A. Hassan, A. A. Fardoun, R. Ramzan, et A. Alraeesi, « Fault Diagnostic Methodologies for Utility-Scale Photovoltaic Power Plants: A State of the Art Review », *Sustainability*, vol. 13, n° 4, p. 1629, janv. 2021, doi: 10.3390/su13041629.
- [6] S. Boubaker, S. Kamel, N. Ghazouani, et A. Mellit, « Assessment of Machine and Deep Learning Approaches for Fault Diagnosis in Photovoltaic Systems Using Infrared Thermography », *Remote Sens.*, vol. 15, n° 6, 2023, doi: 10.3390/rs15061686.
- [7] M. Islam, M. R. Rashel, M. T. Ahmed, A. K. M. K. Islam, et M. Tlemçani, « Artificial Intelligence in Photovoltaic Fault Identification and Diagnosis: A Systematic Review », *Energies*, vol. 16, n° 21, nov. 2023, doi: 10.3390/en16217417.
- [8] A. Thakfan et Y. B. Salamah, « Artificial-Intelligence-Based Detection of Defects and Faults in Photovoltaic Systems: A Survey », *Energies*, vol. 17, n° 19, sept. 2024, doi: 10.3390/en17194807.
- [9] S. R. Mohanty, M. U. Maruf, V. Singh, et Z. Ahmad, « Machine learning approaches for automatic defect detection in photovoltaic systems », *Solar Energy*, vol. 298, p. 113672, sept. 2025, doi: 10.1016/j.solener.2025.113672.
- [10] E. H. Sepúlveda-Oviedo, L. Travé-Massuyès, A. Subias, M. Pavlov, et C. Alonso, « Fault diagnosis of photovoltaic systems using artificial intelligence: A bibliometric approach », *Heliyon*, vol. 9, n° 11, p. e21491, nov. 2023, doi: 10.1016/j.heliyon.2023.e21491.
- [11] I. Passas, « Bibliometric Analysis: The Main Steps », *Encyclopedia*, vol. 4, n° 2, p. 1014- 1025, juin 2024, doi: 10.3390/encyclopedia4020065.
- [12] N. van Eck et L. Waltman, « Software survey: VOSviewer, a computer program for bibliometric mapping », *Scientometrics*, vol. 84, n° 2, p. 523- 538, déc. 2009, doi: 10.1007/s11192-009-0146-3.
- [13] M. Aria et C. Cuccurullo, « bibliometrix: An R-tool for comprehensive science mapping analysis », *Journal of Informetrics*, vol. 11, n° 4, p. 959- 975, nov. 2017, doi: 10.1016/j.joi.2017.08.007.
- [14] B. Thomas, J. Joseph, et J. Jose, « Explorative Bibliometric Study of Medical Image Analysis: Unveiling Trends and Advancements », *Scientific Visualization*, vol. 15, n° 5, 2023, doi: 10.26583/sv.15.5.04.
- [15] M. Akram et al., « CNN based automatic detection of photovoltaic cell defects in electroluminescence images », *ENERGY*, vol. 189, déc. 2019, doi: 10.1016/j.energy.2019.116319.
- [16] W. Tang, Q. Yang, K. Xiong, et W. Yan, « Deep learning based automatic defect identification of photovoltaic module using electroluminescence images », *SOLAR ENERGY*, vol. 201, p. 453- 460, mai 2020, doi: 10.1016/j.solener.2020.03.049.
- [17] X. Yang, Q. Zhang, S. Wang, et Y. Zhao, « Detection of solar panel defects based on separable convolution and convolutional block attention module », *Energy Sources Recovery Util. Environ. Eff.*, vol. 45, n° 3, p. 7136- 7149, 2023, doi: 10.1080/15567036.2023.2218301.
- [18] C. Haitao, L. Caijia, W. Xin, et S. Qiang, « An automatic defects detection system for PV plants with data augmentation », in *2021 International Conference on Power System Technology (POWERCON)*, déc. 2021, p. 980- 984. doi: 10.1109/POWERCON53785.2021.9697898.
- [19] Y. Zhang, X. Zhang, et D. Tu, « Solar photovoltaic module defect detection based on deep learning », *Meas. Sci. Technol.*, vol. 35, n° 12, p. 125404, sept. 2024, doi: 10.1088/1361-6501/ad7d28.
- [20] H. P.-C. Hwang, C. C.-Y. Ku, et J. C.-C. Chan, « Detection of malfunctioning photovoltaic modules based on machine learning algorithms », *IEEE Access*, vol. 9, p. 37210- 37219, 2021, doi: 10.1109/ACCESS.2021.3063461.
- [21] Z. Barraz, I. Sebari, K. Ait El Kadi, et I. Ait Abdelmoula, « Towards a Holistic Approach for UAV-Based Large-Scale Photovoltaic Inspection: A Review on Deep Learning and Image Processing Techniques », *Technologies*, vol. 13, n° 3, p. 117, mars 2025, doi: 10.3390/technologies13030117.
- [22] J. Jerome Vasanth, N. V. S, S. Sugumaran, et V. S. Mahamuni, « Enhancing Photovoltaic Module Fault Diagnosis with Unmanned Aerial Vehicles and Deep Learning-Based Image Analysis », *Int. J. Photoenergy*, vol. 2023, 2023, doi: 10.1155/2023/8665729.
- [23] S. Prabhakaran, R. Uthra, et J. Preetha Roselyn, « Feature Extraction and Classification of Photovoltaic Panels Based on Convolutional Neural Network », *Comput. Mater. Continua*, vol. 74, n° 1, p. 1437- 1455, 2023, doi: 10.32604/cmc.2023.032300.
- [24] S. Naveen Venkatesh, B. Rebecca Jeyavadhanam, A. M. Moradi Sizkouhi, S. M. Esmailifar, M. Aghaei, et V. Sugumaran, « Automatic detection of visual faults on photovoltaic modules using deep ensemble learning network », *Energy Reports*, vol. 8, p. 14382- 14395, nov. 2022, doi: 10.1016/j.egyr.2022.10.427.
- [25] R. Pierdicca, M. Paolanti, A. Felicetti, F. Piccinini, et P. Zingaretti, « Automatic faults detection

of photovoltaic farms: Solair, a deep learning-based system for thermal images », *Energies*, vol. 13, n° 24, 2020, doi: 10.3390/en13246496.

[26] C.-F. Jeffrey Kuo, S.-H. Chen, et C.-Y. Huang, « Automatic detection, classification and localization of defects in large photovoltaic plants using unmanned aerial vehicles (UAV) based infrared (IR) and RGB imaging », *Energy Conversion and Management*, vol. 276, p. 116495, janv. 2023, doi: 10.1016/j.enconman.2022.116495.

[27] A. Rico Espinosa, M. Bressan, et L. F. Giraldo, « Failure signature classification in solar photovoltaic plants using RGB images and convolutional neural networks », *Renewable Energy*, vol. 162, p. 249- 256, déc. 2020, doi: 10.1016/j.renene.2020.07.154.

[28] D. Matusz-Kalász, I. Bodnár, et M. Jobbágy, « An Overview of CNN-Based Image Analysis in Solar Cells, Photovoltaic Modules, and Power Plants », *Applied Sciences*, vol. 15, n° 10, p. 5511, janv. 2025, doi: 10.3390/app15105511.

[29] S. Apak et M. Farsadi, « Multi-branch spatial pyramid dynamic graph convolutional neural networks for solar defect detection », *Comput Electr Eng*, vol. 121, 2025, doi: 10.1016/j.compeleceng.2024.109872.

[30] M. A. Gupta, M. B. Shah, et M. M. Z. Katlariwala, « Enhanced Detection of Solar Panel Defects Using VGG16-Based Convolutional Neural Networks », *International Journal of Scientific Research in Science and Technology*, vol. 12, n° 3, p. 1255- 1262, juin 2025, doi: 10.32628/IJSRST25123136.

[31] H. Gasparyan, S. Agaian, et S. Wu, « Efficient Lightweight Networks for Solar Panel Fault Classification Using EL and RGB Imagery », *IEEE Trans. Instrum. Meas.*, vol. 74, 2025, doi: 10.1109/TIM.2025.3548249.

[32] A. A. Raorane, D. B. Magare, et Y. D. Mistry, « Performance of fault classification on Photovoltaic modules using Thermographic images », *ITM Web of Conferences*, 2022, doi: 10.1051/itmconf/20224403065.

[33] A. Raorane, D. Magare, et Y. Mistry, « Thermal image-driven CNN for predicting solar photovoltaic module lifespan from hotspots », *Solar Energy*, vol. 302, p. 113965, déc. 2025, doi: 10.1016/j.solener.2025.113965.

[34] E. Ozturk, E. Ogliari, M. Sakwa, A. Dolara, N. Blasutigh, et A. M. Massi Pavan, « Photovoltaic modules fault detection, power output, and parameter estimation: A deep learning approach based on electroluminescence images », *Energy Convers. Manage.*, vol. 319, 2024, doi: 10.1016/j.enconman.2024.118866.

[35] V. E. Puranik et R. Gupta, « Analysis and insight of electroluminescence imaging in the assessment of potential-induced degradation in crystalline silicon photovoltaic module », *Engineering Failure Analysis*, vol. 134, p. 106027, avr. 2022, doi:

10.1016/j.engfailanal.2022.106027.

[36] B. Wang et al., « PVF-10: A high-resolution unmanned aerial vehicle thermal infrared image dataset for fine-grained photovoltaic fault classification », *Appl. Energy*, vol. 376, 2024, doi: 10.1016/j.apenergy.2024.124187.

[37] A. S. Al- Waisy et al., « Identifying defective solar cells in electroluminescence images using deep feature representations », *PeerJ Comput. Sci.*, vol. 8, p. e992, mai 2022, doi: 10.7717/peerj-cs.992.

[38] D. Manno, G. Cipriani, G. Ciulla, V. Di Dio, S. Guarino, et V. Lo Brano, « Deep learning strategies for automatic fault diagnosis in photovoltaic systems by thermographic images », *Energy Convers. Manage.*, vol. 241, 2021, doi: 10.1016/j.enconman.2021.114315.

[39] W. M. Shaban, « Detection and classification of photovoltaic module defects based on artificial intelligence », *Neural Comput. Appl.*, vol. 36, n° 27, p. 16769- 16796, 2024, doi: 10.1007/s00521-024-10000-z.

[40] H. Chen, A. Zhang, C. Gong, W. Liang, et Z. Wang, « Fault Diagnosis Method for Photovoltaic Panels Based on Improved ShuffleNet V2 and Infrared Images », in *2022 7th International Conference on Power and Renewable Energy (ICPRE)*, sept. 2022, p. 447- 451. doi: 10.1109/ICPRE55555.2022.9960396.

[41] N. L. Bin Zulkifli et al., « Optimizing Thermal Image Segmentation Using K-Means Clustering and Morphological Operations for Photovoltaic », in *IEEE Int. Conf. Control Syst., Comput. Eng., ICCSCE - Conf. Proc.*, Institute of Electrical and Electronics Engineers Inc., 2025, p. 169- 174. doi: 10.1109/ICCSCE65566.2025.11182688.

[42] X. Xu, M. Liu, Y. Nie, K. Wang, et W. Xu, « Photovoltaic Module Fault Detection Technology Based on Remote Sensing Technology and Deeplabv3+ Model », *IEEE Access*, vol. 12, p. 195472- 195482, 2024, doi: 10.1109/ACCESS.2024.3515165.

[43] S. P. Pathak, S. A. Patil, et D. Mishra, « Enhancing Solar Panel Fault Detection: An Efficient Multidomain Feature Analysis Model with Entropy-Guided Saliency Map Segmentation », *Int. J. Intelligent Eng. Syst.*, vol. 17, n° 4, p. 328- 339, 2024, doi: 10.22266/IJIES2024.0831.25.

[44] V. Carletti, A. Greco, A. Saggese, et M. Vento, « An intelligent flying system for automatic detection of faults in photovoltaic plants », *J Ambient Intell Human Comput*, vol. 11, n° 5, p. 2027- 2040, mai 2020, doi: 10.1007/s12652-019-01212-6.

[45] L. Bommès, T. Pickel, C. Buerhop, J. Hauch, C. Brabec, et I. Marius Peters, « Computer vision tool for detection, mapping, and fault classification of photovoltaics modules in aerial IR videos », *Prog Photovoltaics Res Appl*, vol. 29, n° 12, p. 1236- 1251, 2021, doi: 10.1002/pip.3448.

[46] W.-H. Chiang, H.-S. Wu, J.-S. Wu, et S.-J. Lin,

- « A Method for Estimating On-Field Photovoltaics System Efficiency Using Thermal Imaging and Weather Instrument Data and an Unmanned Aerial Vehicle », *Energies*, vol. 15, n° 16, p. 5835, janv. 2022, doi: 10.3390/en15165835.
- [47] Ö. Baltacı, Z. Kiral, K. Dalkılıç, et O. Karaman, « Thermal Image and Inverter Data Analysis for Fault Detection and Diagnosis of PV Systems », *Appl. Sci.*, vol. 14, n° 9, 2024, doi: 10.3390/app14093671.
- [48] A. Sohail, M. S. Rusdi, M. Z. Abdullah, et S. M. Sultan, « Numerical investigation of the thermal performance of PV system under variable wind speeds and ambient temperatures in the tropical climate of Malaysia », *Green Technologies and Sustainability*, vol. 3, n° 3, p. 100210, juill. 2025, doi: 10.1016/j.grets.2025.100210.
- [49] « (PDF) Influence of irradiance and drone altitude in infrared thermography inspections of photovoltaic plants », *ResearchGate*, avr. 2026, doi: 10.1051/epjpv/2025020.
- [50] W. Zhang, Y. Zhou, J. Wang, et Z. Li, « PHOTOVOLTAIC MODULES FAULT DETECTION MODEL BASED ON YOLOv7 », *Taiyangneng Xuebao*, vol. 46, n° 8, p. 333- 340, 2025, doi: 10.19912/j.0254-0096.tynxb.2024-0540.
- [51] C. R. da Silveira, C. E. R. Sousa, et R. H. Fonseca Alves, « Automatic Fault Classification in Photovoltaic Modules Using Denoising Diffusion Probabilistic Model, Generative Adversarial Networks, and Convolutional Neural Networks », *Energies*, vol. 18, n° 4, 2025, doi: 10.3390/en18040776.
- [52] M. A. Ebied, A. Munshi, S. A. Alhuzali, M. M. El-Sotouhy, A. I. Shehta, et M. S. Elborlsy, « Advanced deep learning modeling to enhance detection of defective photovoltaic cells in electroluminescence images », *Sci. Rep.*, vol. 15, n° 1, 2025, doi: 10.1038/s41598-025-14478-y.
- [53] H. F. M. Romero et al., « Synthetic Dataset of Electroluminescence Images of Photovoltaic Cells by Deep Convolutional Generative Adversarial Networks », *Sustainability*, vol. 15, n° 9, p. 7175, janv. 2023, doi: 10.3390/su15097175.
- [54] J. Lekavičius et V. Gružas, « Data Augmentation with Generative Adversarial Network for Solar Panel Segmentation from Remote Sensing Images », *Energies*, vol. 17, n° 13, p. 3204, janv. 2024, doi: 10.3390/en17133204.
- [55] Y. Jia, G. Chen, et L. Zhao, « Defect detection of photovoltaic modules based on improved VarifocalNet », *Sci Rep*, vol. 14, n° 1, p. 15170, juill. 2024, doi: 10.1038/s41598-024-66234-3.
- [56] A. Jha, Y. Rawat, et S. Vyas, « Advancing automatic photovoltaic defect detection using semi-supervised semantic segmentation of electroluminescence images », *Eng Appl Artif Intell*, vol. 160, 2025, doi: 10.1016/j.engappai.2025.111790.
- [57] A.-K. Hamid, M. Alrahhal, K. Obaideen, T. Bonny, Y. C. Tan, et M. I. Hussein, « Artificial intelligence for smart solar energy monitoring: Genetic attention-based hybrid deep-handcrafted fusion for faulty solar panel image classification », *Result. Eng.*, vol. 28, 2025, doi: 10.1016/j.rineng.2025.107900.
- [58] T. S. Babu, G. Sruthi, A. A. Goud, V. S. Manikanta, M. Paramasivan, et B. K. Chaitanya, « Fault Identification and Classification of Solar PV System », in *Int. Conf. Circuit, Power Comput. Technol., ICCPCT, Institute of Electrical and Electronics Engineers Inc.*, 2025, p. 1376- 1382. doi: 10.1109/ICCPCT65132.2025.11176670.
- [59] L. E. Montanez, J. Reyes-Luévano, L. M. Valentín, D. Moctezuma-Ochoa, F. Lezama, et Z. Vale, « HotSD-Net: Convolutional Neural Network for Estimation of Hot Spots in Photovoltaic Modules via Infrared Imaging », in *IEEE Symp. Comput. Intell. Energy, Transp. Environ. Sustain., CIETES, Institute of Electrical and Electronics Engineers Inc.*, 2025. doi: 10.1109/CIETES63869.2025.10995018.
- [60] R. D. Amar Raj, R. M. R. Yanamala, A. Pallakonda, et A. N. Kanasottu, « AAPN-Tiny: A Compact Edge-Deployable Adaptive Attention Pyramid Architecture for Multi-Class Fault Diagnosis in Solar Photovoltaic Modules », *IEEE. Lat. Am. Trans.*, vol. 23, n° 12, p. 1284- 1296, 2025, doi: 10.1109/TLA.2025.11231229.
- [61] D. Rocha et al., « Multidefect Detection Tool for Large-Scale PV Plants: Segmentation and Classification », *IEEE Journal of Photovoltaics*, vol. 13, n° 2, p. 291- 295, mars 2023, doi: 10.1109/JPHOTOV.2023.3236188.
- [62] F. Hong, J. Song, H. Meng, W. Rui, F. Fang, et Z. Zhang, « A novel framework on intelligent detection for module defects of PV plant combining the visible and infrared images », *Sol. Energy*, vol. 236, p. 406- 416, 2022, doi: 10.1016/j.solener.2022.03.018.
- [63] Z. Ksira, N. Blasutigh, A. Mellit, et A. M. Massi Pavan, « TinyML Model for Fault Classification of Photovoltaic Modules Based on Visible Images », in *Lect. Notes Networks Syst.*, M. Hatti, Éd., Springer Science and Business Media Deutschland GmbH, 2024, p. 373- 380. doi: 10.1007/978-3-031-60629-8_37.
- [64] K. Himavarshini et al., « TinyTripleNet: A lightweight architecture for solar photo voltaic fault detection optimized for edge tensor processing unit hardware deployment », *Electr Power Syst Res*, vol. 252, 2026, doi: 10.1016/j.eprsr.2025.112451.
- [65] A. Mellit, N. Blasutigh, et A. M. Massi Pavan, « TinyML for fault diagnosis of Photovoltaic Modules using Edge Impulse Platform », in *Int. Conf. Smart Grid, icSmartGrid, Institute of Electrical and Electronics Engineers Inc.*, 2023. doi: 10.1109/icSmartGrid58556.2023.10171088.

- [66] M. F. Bekciogullari, H. Acikgoz, et S. Ozbay, « Detection of Hot Spots in PV Modules with the CBAM-Enhanced Yolo Model », in ISAS - Int. Symp. Innov. Approaches Smart Technol., Proc., Institute of Electrical and Electronics Engineers Inc., 2025. doi: 10.1109/ISAS66241.2025.11101735.
- [67] A. Karakan, « Detection of Defective Solar Panel Cells in Electroluminescence Images with Deep Learning », *Sustainability*, vol. 17, n° 3, 2025, doi: 10.3390/su17031141.
- [68] N. Drir, F. Chekired, A. Mellit, et N. Blasutigh, « Hybrid CNN-EML model for fault diagnosis in electroluminescence images of photovoltaic cells », *Renew. Energy*, vol. 250, 2025, doi: 10.1016/j.renene.2025.123343.
- [69] S. Boubaker, C. Moussaoui, A. Mellit, et M. Benganem, « TinyML Model for Solar Cell Defect Classification Based on Electroluminescence Images », in *Lect. Notes Electr. Eng.*, B. Hajji, A. Gagliano, A. Mellit, A. Rabhi, et M. Calì, Éd., Springer Science and Business Media Deutschland GmbH, 2025, p. 511 - 519. doi: 10.1007/978-981-96-0644-3_46.
- [70] D. Lang et Z. Lv, « A PV cell defect detector combined with transformer and attention mechanism », *Sci. Rep.*, vol. 14, n° 1, 2024, doi: 10.1038/s41598-024-72019-5.
- [71] N. V. S, J. V. Joseph, S. Sugumaran, et M. Aghaei, « Weightless Neural Network-Based Detection and Diagnosis of Visual Faults in Photovoltaic Modules », *Energies*, vol. 16, n° 15, 2023, doi: 10.3390/en16155824.
- [72] M. Alanazi et Y. A. Alamri, « SolarFaultAttentionNet: Dual-Attention Framework for Enhanced Photovoltaic Fault Classification », *Inventions*, vol. 10, n° 5, 2025, doi: 10.3390/inventions10050091.
- [73] P. A. Vyas, S. Bhandari, et P. Sonowane, « Vision Transformer-Based Fault Diagnosis for Solar Panels Using Visible Image Dataset », in *Int. Conf. Energy, Power Environ., ICEPE*, Institute of Electrical and Electronics Engineers Inc., 2025. doi: 10.1109/ICEPE65965.2025.11139664.
- [74] K. Chahine, « Tree-Based Algorithms and Incremental Feature Optimization for Fault Detection and Diagnosis in Photovoltaic Systems », *Eng*, vol. 6, n° 1, p. 20, janv. 2025, doi: 10.3390/eng6010020.
- [75] S. Goyal et J. C. Rajapakse, « Self-supervised learning for hotspot detection and isolation from thermal images », *Expert Systems with Applications*, vol. 237, p. 121566, mars 2024, doi: 10.1016/j.eswa.2023.121566.
- [76] F. Z. Kardaş et A. Atmaca, « A REVIEW OF ARTIFICIAL INTELLIGENCE BASED STUDIES ABOUT FAULT-DETECTION AND PERFORMANCE ASSESSMENT IN SOLAR POWER PLANTS », *IJEES*, n° Advanced Online Publication, p. 16- 39.
- [77] A. Oulefki et al., « Unveiling the Invisible: Enhanced Detection and Analysis of Deteriorated Areas in Solar PV Modules Using Unsupervised Sensing Algorithms and 3D Augmented Reality », 12 juillet 2023, arXiv: arXiv:2307.05136. doi: 10.48550/arXiv.2307.05136.
- [78] A. Esteban, A. Cano, S. Ventura, et A. Zafra, « Simultaneous fault prediction in evolving industrial environments with ensembles of Hoeffding adaptive trees », *Appl Intell*, vol. 55, n° 13, p. 930, août 2025, doi: 10.1007/s10489-025-06786-7.
- [79] N. Sergin, J. Huang, T.-S. Chang, et H. Yan, « Image-based Novel Fault Detection with Deep Learning Classifiers using Hierarchical Labels », 26 mars 2024, arXiv: arXiv:2403.17891. doi: 10.48550/arXiv.2403.17891.
- [80] E. K. Ukiwe, S. A. Adeshina, T. Jacob, et B. B. Adetokun, « Deep learning model for detection of hotspots using infrared thermographic images of electrical installations », *Journal of Electrical Systems and Inf Technol*, vol. 11, n° 1, p. 24, juin 2024, doi: 10.1186/s43067-024-00148-y.
- [81] D. Rocha, M. Lopes, J. P. Teixeira, P. A. Fernandes, M. Morais, et P. M. P. Salomé, « A Deep Learning Approach for PV Failure Mode Detection in Infrared Images: First Insights », in *Conf Rec IEEE Photovoltaic Spec Conf*, Institute of Electrical and Electronics Engineers Inc., 2022, p. 630- 632. doi: 10.1109/PVSC48317.2022.9938524.
- [82] Y. Tao, Z. Xu, Q. Liu, L. Li, et Y. Zhang, « Improved faster R-CNN algorithm for defect detection of electromagnetic luminescence », in *Proc SPIE Int Soc Opt Eng*, H. Xia, L. X. Yang, et L. Yu, Éd., SPIE, 2021. doi: 10.1117/12.2617320.
- [83] N. S. Maddileti, R. Namburi, R. D. A. David Amar Raj, R. M. R. Yanamala, et A. Pallakonda, « DCGAN-Driven Minority Class Augmentation for Lightweight YOLO-Based Photovoltaic Defect Localization Suitable for Edge Deployment », *IEEE Trans. Device Mater. Reliab.*, vol. 25, n° 3, p. 742- 751, 2025, doi: 10.1109/TDMR.2025.3592416.
- [84] P. R. Sonawane et S. Bhandari, « Deep Learning-Based Fault Diagnosis System for Solar Photovoltaic Panels Using EL Images », in *Lect. Notes Electr. Eng.*, H. Mohamad, M. H. Hasan, S. J. Abdulkadir, et N. Shafiq, Éd., Springer Science and Business Media Deutschland GmbH, 2025, p. 279- 294. doi: 10.1007/978-981-96-5848-0_24.
- [85] A. Mellit, « An embedded solution for fault detection and diagnosis of photovoltaic modules using thermographic images and deep convolutional neural networks », *Eng Appl Artif Intell*, vol. 116, 2022, doi: 10.1016/j.engappai.2022.105459.
- [86] A. Mellit, C. Moussaoui, S. Pastore, et A. M. Pavan, « Fault Diagnosis of Photovoltaic Modules

Using Infrared Images and Hybrid DCNNs: A Comparative Study », in 2024 IEEE International Conference on Environment and Electrical Engineering and 2024 IEEE Industrial and Commercial Power Systems Europe (EEEIC / I&CPS Europe), juin 2024, p. 1- 6. doi: 10.1109/EEEIC/ICPSEurope61470.2024.10751391.

[87] P. Malik, V. Saxena, S. Raj, S. Singh, S. Kumar, et P. S. Ganaraj, « Fault Detection of the Solar Photovoltaic Modules Using YOLO Models », in IEEE Reg. 10 Symp., TENSYP, Institute of Electrical and Electronics Engineers Inc., 2024. doi: 10.1109/TENSYP61132.2024.10752194.

[88] B. Li, X. Chen, et A. Jain, « Exploring Capability of Multimodal Foundation Model for Image-based Fault Detection of Photovoltaic Modules », in Conf Rec IEEE Photovoltaic Spec Conf, Institute of Electrical and Electronics Engineers Inc., 2025, p. 988- 991. doi: 10.1109/PVSC59419.2025.11133222.

[89] Q. Hong, T. Huang, W. Yan, C. Ji, X. Li, et Z. Du, « Deep Learning Based Multi-Stage Defect Diagnosis Method of Photovoltaic Modules Using Aerial Images from UAV », in Chinese Control Conf., CCC, J. Sun et H. Yin, Éd., IEEE Computer Society, 2025, p. 7520- 7525. doi: 10.23919/CCC64809.2025.11179561.

[90] W.-Q. Lu, C. Delpha, D. Diallo, et A. Migan-Dubois, « Spatial Attention Mechanism Derived from Diffusion Model for Photovoltaic Cell Fault Diagnosis », in Proc. - EUROCON: Int. Conf. Smart Technol., I. Czarnowski et M. Jasinski, Éd., Institute of Electrical and Electronics Engineers Inc., 2025. doi: 10.1109/EUROCON64445.2025.11073476.

[91] L. Bommès et al., « Anomaly detection in IR images of PV modules using supervised contrastive learning », Prog Photovoltaics Res Appl, vol. 30, n° 6, p. 597- 614, 2022, doi: 10.1002/pip.3518.

[92] U. Naeem, K. Chadda, S. Vahaji, J. Ahmad, X. Li, et E. Asadi, « Aerial Imaging-Based Soiling Detection System for Solar Photovoltaic Panel Cleanliness Inspection », Sensors, vol. 25, n° 3, 2025, doi: 10.3390/s25030738.

[93] Q. Wang, K. Paynabar, et M. Pacella, « Online automatic anomaly detection for photovoltaic systems using thermography imaging and low rank matrix decomposition », J Qual Technol, vol. 54, n° 5, p. 503- 516, 2022, doi: 10.1080/00224065.2021.1948372.

[94] S. Gallmetzer, M. Sondoqah, E. Turri, L. Koester, A. Louwen, et D. Moser, « Redefining failure detection in PV Systems: a comparative study of GPT-4o and ResNet's computer vision in aerial infrared imagery analysis », EPJ Photovolt., vol. 16, 2025, doi: 10.1051/epjpv/2025010.

[95] X. Li, W. Li, Q. Yang, W. Yan, et A. Y. Zomaya, « Edge-Computing-Enabled Unmanned Module Defect Detection and Diagnosis System for Large-Scale Photovoltaic Plants », IEEE Internet Things J., vol. 7, n° 10, p. 9651- 9663, 2020, doi: 10.1109/JIOT.2020.2983723.

[96] J. I. A. Tsanakas et P. Marechal, « Decoding pixels: A modular software prototype for cognitive image-based diagnostics of PV plants », EPJ Photovolt., vol. 16, 2025, doi: 10.1051/epjpv/2025013.

[97] M. Kolahi, S. M. Esmailifar, A. M. Moradi Sizkouhi, et M. Aghaei, « Digital-PV: A digital twin-based platform for autonomous aerial monitoring of large-scale photovoltaic power plants », Energy Convers. Manage., vol. 321, 2024, doi: 10.1016/j.enconman.2024.118963.

[98] S. Rao et al., « A cyber-physical system approach for photovoltaic array monitoring and control », in Int. Conf. Inf., Intell., Syst. Appl., IISA, Institute of Electrical and Electronics Engineers Inc., 2017, p. 1- 6. doi: 10.1109/IISA.2017.8316458.