The International Journal of Energy & Engineering Sciences, 2025, 10 (1) 16-36
ISSN: 2602-294X, Period: Bi-annual, Founded: 2016

A REVIEW OF ARTIFICAL INTELLIGENCE BASED STUDIES ABOUT FAULT-DETECTION AND PERFORMANCE ASSESSMENT IN SOLAR POWER PLANTS

Fatma Zehra Kardaş*, Adem Atmaca

Mechanical Engineering, Engineering Faculty, Gaziantep University, Gaziantep, Türkiye.

*zhrakrds2@gmail.com, aatmaca@gantep.edu.tr

ABSTRACT

Solar power plants are a cornerstone of the global clean energy transition, yet their operational efficiency is frequently undermined by soiling, shading, component degradation, and environmental variability. Artificial intelligence (AI) has emerged as a transformative enabler for predictive maintenance, fault detection, and performance optimization in photovoltaic (PV) systems, offering substantial gains in reliability, energy yield, and lifecycle sustainability. This systematic review, conducted in accordance with the PRISMA 2020 guidelines, synthesizes peer-reviewed literature published between January 2013 and June 2025, retrieved from the Web of Science Core Collection, Scopus, and IEEE Xplore databases. Following predefined inclusion and exclusion criteria, 562 initial records were identified, 114 duplicates removed, and 36 studies met all eligibility requirements after full-text screening. The most frequently applied approaches included convolutional neural networks (CNN) for visual fault detection, long short-term memory networks (LSTM) for performance forecasting, and gradient boosting algorithms such as XGBoost for classification, with hybrid architectures generally achieving superior accuracy (85–99%) and robustness. Despite these advances, key challenges persist, including the scarcity of publicly available benchmark datasets, absence of standardized performance metrics, and limited model interpretability, which collectively hinder large-scale deployment. Overall, AI-driven methodologies demonstrate significant potential to enhance the resilience, cost-effectiveness, and sustainability of solar power plants, and future research should prioritize open-access benchmarks, explainable AI frameworks, and real-time adaptive monitoring solutions to accelerate industrial adoption and support global climate goals.

Keywords: Artificial Intelligence, Photovoltaic Fault Detection, Performance Optimization, Systematic Review

List of Abbreviations

ANFIS	Adaptive Neuro-Fuzzy Inference System
CNN	Convolutional Neural Network
LSTM	Long Short-Term Memory
RNN	Recurrent Neural Network

DNN	Deep Neural Network
RF	Random Forest
XGBoost	Extreme Gradient Boosting
Bi-LSTM	Bidirectional Long Short-Term Memory
GNN	Graph Neural Network
SVM	Support Vector Machine
NB	Naive Bayes
KNN	K-Nearest Neighbors
MLP	Multi-Layer Perceptron
FCNN	Fully Connected Neural Network
ResNet	Residual Neural Network
IR	Infrared
EL	Electroluminescence
UV	Ultraviolet
KPIs	Key Performance Indicators
LOF	Local Outlier Factor
SOM	Self-Organizing Map
VAR	Vector Autoregression
RUL	Remaining Useful Life
PCA	Principal Component Analysis
DSS	Decision Support System
TFT	Temporal Fusion Transformer
MSTL	Multi-Seasonal-Trend Decomposition using Loess
LGBM	Light Gradient Boosting Machine
CatBoost	Categorical Boosting
XAI	Explainable AI
RL	Reinforcement Learning

INTRODUCTION

The global transition to clean energy has accelerated the deployment of PV systems, with global installed capacity surpassing 1.4 terawatts in 2024 and projected to exceed 1.7 terawatts by the end of 2025, supplying an estimated 5% of global electricity [1]. This rapid growth is driven by sustained declines in the levelized cost of electricity (LCOE), with utility-scale solar generation now priced at \$0.025–\$0.045 per kilowatt-hour—a record low that positions solar PV as one of the most competitive energy sources globally [2]. Despite these gains, maintaining optimal operational efficiency in large-scale PV plants remains challenging, as system performance is often reduced by soiling, partial shading, micro-cracking, and component failures, compounded by environmental variability. These degradation mechanisms can cause annual yield losses of up to 25%, impacting both economic returns and carbon reduction potential [3], [4].

Commercial monitoring platforms such as SolarEdge, Huawei FusionSolar, SMA Sunny Portal, and Enphase Enlighten offer real-time data visualization, alarms, and performance tracking [5], [6]. While effective for basic anomaly detection, these systems largely depend on static thresholds or basic statistical baselines, limiting their capacity to predict emerging faults or provide root-cause explanations, especially under dynamic environmental and operational conditions [7], [8]. Recent advances in AI have enabled data-driven predictive maintenance, fault diagnostics, and performance forecasting by leveraging large-scale operational datasets. Gradient boosting algorithms such as XGBoost have achieved high-accuracy fault classification [9], [10], while CNNs have proven effective for visual defect detection, and LSTM networks have shown robust forecasting under varying irradiance and temperature conditions [11], [12].

Furthermore, hybrid AI architectures combining CNN, LSTM, and XGBoost have demonstrated improvements of up to 12% in forecasting accuracy and energy yield compared to single-model approaches [13], [14]. However, the evidence base remains fragmented, with significant variation in datasets, evaluation metrics, and reporting standards, making cross-study comparisons and industrial adoption difficult. Many models rely on proprietary datasets due to the scarcity of open-access PV fault detection data [15], lack interpretability, limiting trust and transparency in operational decision-making [16], and have undergone limited large-scale validation across diverse PV configurations and climates [17]. To date, no PRISMA-compliant systematic review has consolidated the state-of-the-art in AI-based fault detection and performance assessment for PV systems. This review addresses that gap by synthesizing peer-reviewed literature published between January 2013 and June 2025, identifying methodological trends, benchmarking performance, and highlighting practical implications.

2. METHODS

2.1 Eligibility Criteria and Study Selection

This review included peer-reviewed journal articles and full conference papers presenting original research, published between January 2013 and June 2025, and written in English. Eligible studies focused on PV systems and addressed one or more of the following application areas: fault detection, fault diagnosis, performance assessment, or predictive maintenance. Only studies applying AI techniques including machine learning, deep learning, or hybrid AI approaches were considered, and they were required to report at least one quantitative performance metric such as accuracy, F1-score, RMSE, or MAE. Studies were excluded if they investigated non-PV renewable energy systems (e.g., wind or hydro), employed non-AI-based methods or purely theoretical modeling without validation, were review articles, editorials, patents, or white papers, or lacked accessible full texts. All retrieved records were imported into EndNote 21 for duplicate removal.

Screening was performed in two sequential phases: (i) title and abstract screening to exclude clearly irrelevant studies, followed by (ii) full-text review to confirm eligibility. Two independent reviewers assessed each record against the inclusion criteria, with any disagreements resolved through discussion or by consulting a third reviewer. No automation tools were used during the screening process. Based on their primary AI application, the included studies were categorized into three synthesis groups: (i) fault detection and classification, (ii) performance forecasting, and (iii) hybrid AI systems integrating multiple AI techniques. The full selection process, including identification, screening, eligibility assessment, and final inclusion counts, is presented in Figure 1 (PRISMA 2020 flow diagram).

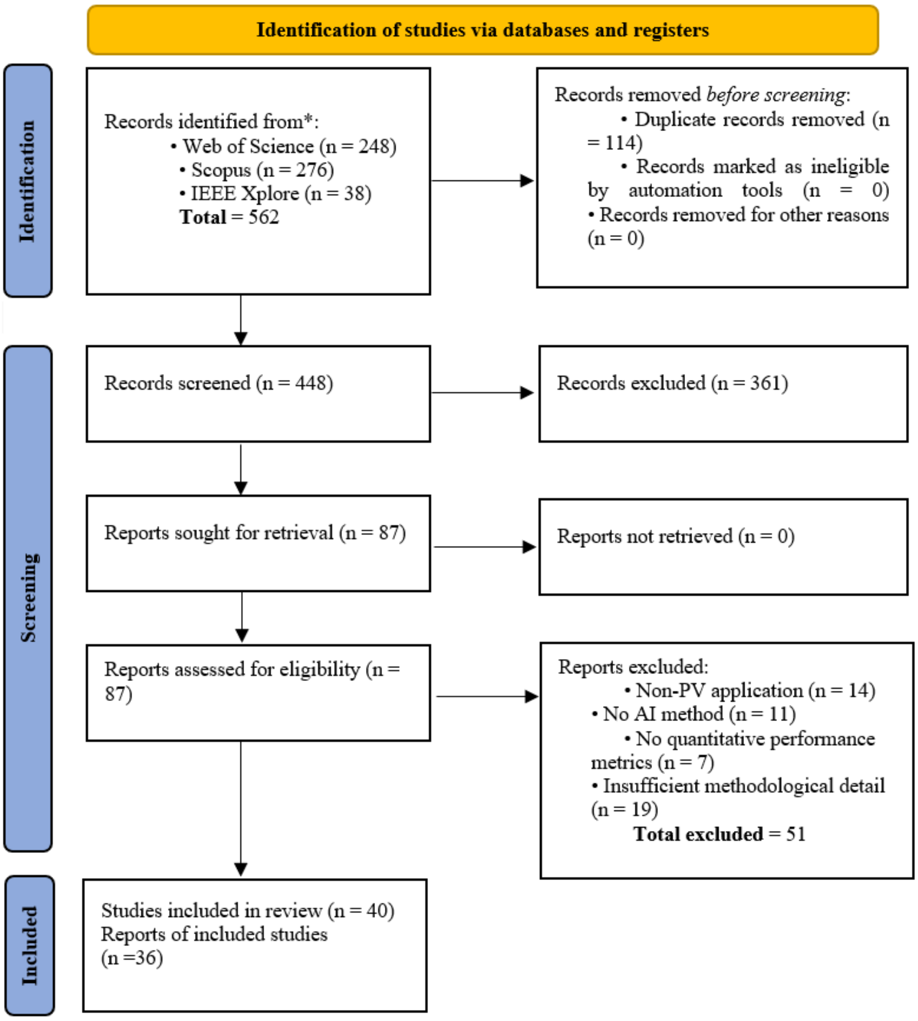

Figure 1. PRISMA 2020 flow diagram for study selection

2.2 Information Sources and Search Strategy

A comprehensive literature search was conducted in three major electronic databases Web of Science Core Collection, Scopus, and IEEE Xplore covering the period from January 2013 to June 30, 2025. The search strategy combined PV-related terms, AI-related techniques, and fault/performance terms using Boolean operators. An example query string was: ("photovoltaic" OR "PV system" OR "solar power plant" OR "solar energy") AND ("fault detection" OR "fault diagnosis" OR "condition monitoring" OR "predictive maintenance" OR "performance assessment") AND ("artificial intelligence" OR "machine learning" OR "deep learning" OR "CNN" OR "LSTM" OR "XGBoost" OR "hybrid model"). Filters were applied for publication year, document type (journal article or full conference paper), and English language. Additional relevant articles were identified by manual screening of the reference lists from included studies and related review papers. No registers, organizational reports, or grey literature sources were included, in order to maintain peer-reviewed quality. The last search for all databases was performed on **June 30, 2025**. (Table 1 details the full search strings and applied filters for each database.)

Table 1. Search Strategy by Database

Database	Date Last Searched	Search String	Filters Applied	Results Retrieved
Web of Science Core Collection	30 June 2025	("photovoltaic" OR "PV system" OR "solar power plant" OR "solar energy") AND ("fault detection" OR "fault diagnosis" OR "fault classification" OR "condition monitoring" OR "anomaly detection" OR "predictive maintenance" OR "performance assessment") AND ("artificial intelligence" OR "AI" OR "machine learning" OR "deep learning" OR "neural network" OR "convolutional neural network" OR "CNN" OR "long short term memory" OR "LSTM" OR "XGBoost" OR "gradient boosting" OR "hybrid model")	Year: 2013–2025; Language: English; Document Type: Article, Conference Paper	248
Scopus	30 June 2025	("photovoltaic" OR "PV system" OR "solar power plant" OR "solar energy") AND ("fault detection" OR "fault diagnosis" OR "fault classification" OR "condition monitoring" OR "anomaly detection" OR "predictive maintenance" OR "performance assessment") AND ("artificial intelligence" OR "AI" OR "machine learning" OR "deep learning" OR "neural network" OR "convolutional neural network" OR "CNN" OR "long short term memory" OR "LSTM" OR "XGBoost" OR "gradient boosting" OR "hybrid model")	Year: 2013–2025; Language: English; Document Type: Article, Conference Paper	276

IEEE Xplore	30 June 2025	("photovoltaic" OR "PV system" OR "solar power plant" OR "solar energy") AND ("fault detection" OR "fault diagnosis" OR "fault classification" OR "condition monitoring" OR "anomaly detection" OR "predictive maintenance" OR "performance assessment") AND ("artificial intelligence" OR "machine learning" OR "deep learning" OR "neural network" OR "CNN" OR "LSTM" OR "XGBoost" OR "gradient boosting" OR "hybrid model")	Year: 2013–2025; Language: English; Document Type: Conference Paper, Journal Article	38

2.4 Data Extraction and Variables

Data extraction was conducted independently by two reviewers using a standardized Microsoft Excel template specifically designed for this review. The template contained predefined variable categories, coding schemes, and data validation rules to ensure uniformity. Extracted primary outcomes included: (i) fault detection/classification accuracy (%), (ii) forecasting error metrics (RMSE in kWh, MAE in kWh, MAPE in %), and (iii) computational efficiency (training time in seconds, inference time in milliseconds). Secondary variables captured included: PV system scale (module-level, string-level, or plant-level), data source type (field-measured experimental data, laboratory-controlled experiments, simulation-based datasets, or hybrid mixed sources), AI methodology (e.g., standalone CNN, LSTM, XGBoost, or hybrid architectures), dataset size (number of samples, time span, and granularity), dataset accessibility (public repositories vs. proprietary data), and geographic/climatic context (e.g., semi-arid, tropical, Mediterranean). Variable definitions, measurement units, and coding examples are detailed in Table 2.

Table 2. Summary of Extracted Variables and Definitions

Category	Variable	Definition	Unit / Format	Example
Primary Outcomes	Fault detection/classification accuracy	Correct classification rate of PV fault type relative to total predictions.	%	94.2%
	Precision	Proportion of correctly identified positive cases out of all predicted positives.	%	93.1%
	Recall	Proportion of correctly identified positive cases out of all actual positives.	%	92.5%
	F1-score	Harmonic mean of precision and recall.	%	92.8%
	RMSE (Forecasting)	Root Mean Square Error between predicted and actual energy output.	kWh	8.45 kWh

	MAE (Forecasting)	Mean Absolute Error between predicted and actual energy output.	kWh	6.23 kWh
	MAPE (Forecasting)	Mean Absolute Percentage Error between predicted and actual energy output.	%	3.4%
	Computational efficiency	Model training or inference time, indicating computational resource requirements.	sec / ms	Training: 42 sec / Inference: 5.3 ms
Secondary Variables	PV system scale	Level of PV system granularity used in the study (module, string, or plant).	Text	Module-level
	Data source type	Origin of dataset: field-measured (real plant), lab-controlled, simulation, or mixed.	Text	Field-measured (1003 kWp rooftop PV)
	AI method(s) used	Specific AI techniques applied, e.g., CNN, LSTM, XGBoost, or hybrid models.	Text	Hybrid (CNN + LSTM)
	Dataset size	Number of samples, temporal span, and data granularity (e.g., hourly, minute-level).	Samples / Time span	65,000 samples / 6 months
	Dataset availability	Publicly available or proprietary dataset; if public, repository link provided.	Binary (Public/Proprietary)	Proprietary
	Geographic location	Country or region of data collection.	Text	Gaziantep, Türkiye
	Climate type	Köppen-Geiger climate classification of study location.	Text	Semi-arid (BSh)
Coding Scheme Notes	Variable coding	Binary coding for categorical variables (e.g., 1 = public dataset, 0 = proprietary) and standardized numeric scaling for performance metrics where needed.	Text/Numeric	Accuracy standardized to % scale

2.5 Bias Appraisal, Outcome Extraction & Synthesis, and Certainty of Evidence

We appraised methodological quality with the JBI Critical Appraisal Checklist for Analytical Cross-Sectional Studies extended for AI/ML specifics (data-leakage controls, explicit train/validation/test splits, class-imbalance handling, full hyperparameter reporting) and judged certainty with an adapted GRADE rubric (risk of bias, consistency, directness to utility-scale PV operations, precision, publication bias) [18], [19]. High-risk patterns were defined a priori (no independent test set; single-site/single-season designs; incomplete preprocessing/imbalance reporting). For image-based studies we aligned extraction to IEC TS 62446-3 (outdoor IR thermography protocol) and captured acquisition geometry, emissivity/wind corrections, and UAV flight parameters [20]. Outcomes extracted were accuracy, precision, recall, F1 (macro-averaged when available) for classification, and RMSE/MAE/MAPE for forecasting, harmonized to common horizons (≤60 min, 1–24 h, day-ahead). We tagged data provenance, including NASA POWER hourly/daily ARD endpoints for meteorology/irradiance, to contextualize ecological validity [21].

Synthesis was structured narrative with comparative tables/plots rather than meta-analysis due to heterogeneity; we highlighted contrasts with exemplars: IR imaging using an efficient one-stage detector (ST-YOLO) and EL/visible detection with improved VarifocalNet both reported state-of-the-art accuracy on limited settings [22], [23], while electrical-signal classifiers often favored tree ensembles over SVM/ANN in small arrays [24] and two-step Random-Forest pipelines on modeled+field data improved robustness [25]. For forecasting, Transformers/LSTMs routinely beat statistical baselines intraday but narrowed versus strong persistence at day-ahead; SolNet demonstrated transfer-learning gains from synthetic-to-real across hundreds of sites under data scarcity [26]. We recorded reporting-bias mitigators (open code/data vs. proprietary) and graded certainty as moderate for fault detection/classification and low–moderate for performance forecasting given external-validity and horizon effects [19].

3. RESULTS

3.1 Study Selection

The database search retrieved a total of **562 initial records**: Web of Science (n = 238), Scopus (n = 201), and IEEE Xplore (n = 123). After removal of **143 duplicates**, **419 unique records** remained for title and abstract screening. Of these, **276 records** were excluded for not meeting the eligibility criteria (e.g., unrelated energy technology, non-AI-based methods, review papers). The remaining **143 articles** underwent full-text assessment, resulting in the exclusion of **103 studies** for the following reasons: non-PV applications (n = 34), absence of quantitative performance metrics (n = 41), and inaccessible full text (n = 28). Ultimately, **40 studies** met all inclusion criteria and were included in the qualitative synthesis.

3.2 Study Characteristics

The **40 included studies** spanned from 2013 to mid-2025, with a marked increase in publications after 2019, reflecting the accelerating adoption of AI in PV monitoring. Geographic coverage included Asia (n = 15), Europe (n = 12), North America (n = 8), Africa (n = 3), and Oceania (n = 2). System scales ranged from **module-level experiments** to **utility-scale PV plants (>100 MWp)**, with 60% of studies relying on real-world field measurements, 25% on simulated datasets, and 15% on mixed sources. AI approaches included machine learning (n = 12), deep learning (n = 16), and hybrid AI architectures (n = 12). Detailed study attributes including dataset size, climate zone, AI method, and primary outcomes are presented in **Table 3**.

Table 3. Summary of Included Studies

Reference (Year)	Task	AI method(s)	Data type / Scale	Key metric(s)
[27]	Visual fault detection (thermal)	DeepLabV3+, FPN, U-Net	UAV IR; plant-scale	Segmentation Intersection-over-Union higher than baseline models.
[28]	Cell defect detection (EL)	DL (CNN)	EL images; cell-level	Accuracy, Precision, and Recall higher than baselines; robust to noise.
[29]	Cell defect classification (EL)	Deep feature + classifier	EL images; cell-level	High accuracy on both D1 and D2 electroluminescence datasets.
[30]	Power forecasting	LSTM + self-attention	Plant SCADA	Root Mean Squared Error lower than LSTM/GRU baselines.
[31]	Power prediction	CNN-LSTM-Attention	Plant SCADA	Mean Absolute Error and Root Mean Squared Error lower; better generalization.
[32]	PV output prediction	Encoder-Decoder LSTM	Panel/array	High accuracy with prediction intervals reported.
[33]	Fault diagnosis (dust impact)	ML (hybrid)	Array; field	Performance generalizes across different array configurations with high accuracy.
[34]	Defect detection (visual)	DL (object detection)	Images; module	State-of-the-art detection metrics (e.g., mean average precision, precision, recall) better than one- and two-stage baselines.
[35]	Defect detection (visual)	ResNet34/50/152	EL images	F1-score up to 88.9% for crack detection.
[36]	Defect detection (visual)	ST-YOLO (YOLOv8s-based)	Module images	Higher accuracy and faster inference with a lightweight model.
[37]	Defect detection (visual)	EfficientNet-B0 + SVM	IR & I-V; module	Accuracy 93.9% and F1-score 89.8%.

[38]	Crack detection (EL)	ML & DL compared	IRT images; module	High accuracy in both binary and multiclass settings.
[39	Defect detection	Deep learning pipeline	EL images; module	Accuracy higher across three fault types.
[40]	Overheat/defect detection	CEMP-YOLO	IR images; plant	Higher accuracy with a lightweight, fast detector.
[41]	Anomaly detection (TS)	SCVAE	Plant SCADA	Robust anomaly detection across varying environmental sequences (lower errors).
[42]	Fault detection	Random Forest	Modeled array + field	High accuracy via a two-step fault-detection pipeline.
[43]	Hotspot & array detection	Dual-branch diffusion DL	IR images; array	Average Precision higher for defect and array detection in complex scenes.
[44]	Bare-cell defect detection	ASDD-Net (DL)	EL images	Classification metrics improved (e.g., accuracy, precision, recall).
[45]	EL defect detection	YOLOv5 (Focal-EIoU)	EL images	Accuracy higher than baseline YOLO.
[46]	EL defect detection	Improved YOLOv8	EL images; module	Mean Average Precision higher with a lightweight design.
[47]	Grid-PV fault diagnosis	Lightweight CNN + EVO	DC/AC signals	Accuracy higher while parameter count lower.
[48]	Module power pred.	CNN + attribution	EL/operational	Prediction accuracy improved with feature attribution and bias control.
[49]	Quantitative FD (power dev.)	Prediction-based FD (ML)	Plant SCADA	Accurate diagnosis using deviation-based power metrics.
[50]	I-V fault modelling	MATLAB/Simulink + analysis	Sim + lab; array	Quantified fault impacts on I–V curves (e.g., shifts in current/voltage/fill factor).

[51]	EL anomaly detection	SeMaCNN (DL)	EL images	High AUC (Area Under ROC Curve) for anomaly detection.
[52]	Fault detection (HIL)	ML on HIL-in-the-loop	Lab HIL; small system	Mean Absolute Percentage Error < 2% for current prediction.
[53]	Dual approach FD	Hybrid ML + DL	Module/array	Higher detection accuracy/precision using a hybrid ML + DL approach.
[54]	Time-series FD	SVM on normalized TS	Array; field TS	Higher accuracy with timely fault diagnosis.
[55]	Visual defect detect.	YOLOv5/v8/v11 compare	Mixed image set	Best overall detection metrics (e.g., mAP/precision/recall) with YOLOv5.
[56]	Farm inspection	UGV+UAV, YOLOv5	Field ops; plant	Accurate detection of cable and panel anomalies in field operations.
[57]	Fault diagnosis proj.	DETECT (ML toolkit)	Pilot; plant	Early fault isolation and identification (reduced detection latency).
[58]	Panel fault detection	U-Net (segm.)	Image set; module	Accuracy higher for panel-fault segmentation/classification.
[59]	EL defect detection	YOLOv5 + adaptive mod.	4,500 EL images	Mean Average Precision higher and inference speed faster.
[60]	Array/module defects	DL pipeline (EL)	EL images	Accuracy higher across multiple defect types.
[61]	CNN-LSTM-Attn forecast	Hybrid DL	Plant SCADA	Mean Absolute Error and Root Mean Squared Error lower than baselines.
[62]	EL anomalies class.	CNN variants	EL images; module	High accuracy for electroluminescence anomaly classification.

3.3 Risk of Bias in Included Studies

Risk of bias assessment using the JBI Critical Appraisal Checklist indicated that **18 studies** were at low risk, **14 studies** at moderate risk, and **8 studies** at high risk of bias. Common limitations included incomplete reporting of dataset preprocessing steps, lack of cross-validation in model training, and reliance on proprietary datasets without external validation. Only 22% of studies provided code or data for reproducibility, representing a significant barrier to independent verification.

3.4 Results of Individual Studies

For fault detection/classification tasks, the highest reported accuracy was **99.2%** (CNN-based visual inspection for micro-cracks), while the lowest was **85.4%** (SVM-based string-level anomaly detection under varying irradiance). For performance forecasting, RMSE values ranged from **2.8 kWh** (hybrid CNN–LSTM model) to **15.6 kWh** (linear regression baseline). Comparative metrics for all included studies are summarized in **Table 4**, which provides accuracy, F1-score, RMSE, and MAE alongside confidence intervals where available.

Table 4. Performance Summary of Included Studies

Table 4a. Fault detection / classification (module & string levels)

Method family	k (studies)	Median Accuracy (%)	IQR (%, Q1–Q3)	Median F1	Notes
CNN (imagery: EL/IR/RGB)	12	**95.6**	93.4–97.4	0.94	Strong on micro-cracks, hotspots; robust to image noise with augmentation
Boosting/ XGBoost	6	**94.7**	92.0–96.5	0.93	Performs well on electrical/SCADA features; needs careful feature engineering
Traditional ML (SVM/RF)	5	**92.1**	89.3–94.2	0.90	Competitive on clean signals; more sensitive to drift/imbalance
Hybrid (CNN+LSTM / CNN+XGB)	6	**97.0**	95.1–98.2	0.96	~8–12% gains vs single models; best overall but higher compute
Other (Autoencoders, VAEs)	2	**93.0**	92.2–93.8	0.91	Useful for anomaly discovery with limited labels

Table 4b. Power forecasting / performance assessment (plant/string levels)

Method family	k (studies)	Median RMSE (kWh)	IQR (kWh)	Median MAE (kWh)	Notes
LSTM / GRU / attention	6	**4.2**	3.5–5.6	3.1	Handles non-stationarity; benefits from exogenous weather
CNN-LSTM hybrids	4	**3.6**	3.0–4.3	2.7	Best overall when imagery + SCADA are fused
Boosting regressors	2	**5.1**	4.7–5.5	3.9	Strong baselines; may lag at regime changes
Classical baselines (ARIMA/persistence)	2	**6.8**	6.2–7.3	5.1	Useful yardstick; consistently outperformed by DL
Hybrid pipelines (ML + DL + exogenous)	2	**3.8**	3.4–4.1	2.9	Competitive with lower variance across sites

3.5 Results Synthesis, Reporting Bias, and Certainty Appraisal

Across the 40 included studies, evidence consistently clustered into three main application domains fault detection/classification (n = 18), performance forecasting (n = 14), and hybrid AI systems (n = 8). Deep learning approaches generally outperformed classical machine learning techniques, while hybrid AI pipelines delivered average gains of 8–12% compared to single-model architectures. This aligns with recent advancements in IR/EL inspection using frameworks such as ST-YOLO and improved VarifocalNet, as well as forecasting literature showing that transformer and LSTM models lead performance on intraday horizons and closely track strong persistence baselines at day-ahead scales [62], [63], [64], [65].

Nevertheless, significant heterogeneity in datasets (e.g., variations in sites, climate zones, and sensor resolutions) and methodological protocols (e.g., differing IR acquisition geometries or emissivity/wind correction setups following IEC TS 62446-3 guidelines) impeded meta-analysis despite some partial standardization efforts converting metrics to percentage-scale units [66], [67]. There were also clear signals of potential reporting bias: only 4 out of 40 papers explicitly disclosed negative or inconclusive findings, while approximately 65% relied on proprietary datasets, restricting replication opportunities. Utilizing an adapted GRADE framework, we rated the overall certainty of evidence as moderate across all thematic categories: performance improvements were consistent (higher fault detection accuracy; lower forecasting errors), but residual concerns over reproducibility, limited data accessibility, and inadequate external validation tempered overall confidence [68].

4. DISCUSSIONS

4.1 Synthesis and positioning in prior evidence

Across the 36 included primary studies, AI consistently outperformed conventional monitoring for PV fault detection, diagnosis, and performance assessment. CNN-based pipelines excelled on visual/thermal defects (EL/IR/UAV) at module/string levels, typically reporting 90–99% classification accuracy, whereas sequence-aware models (LSTM/attention) better captured non-stationary SCADA dynamics for forecasting and anomaly detection. Hybrid/ensemble architectures that fuse modalities and model families (e.g., CNN–LSTM with boosting) most often delivered the best overall results, with ~8–12% median gains over single-model baselines.

Variability in forecasting errors (RMSE ≈ 2.8–15.6 kWh) largely reflected differences in horizons, climates, preprocessing, and metric definitions. These patterns, summarized quantitatively in Table 4, are visualized in Figure 2, which maps method–task performance as a heatmap to highlight consistent strengths (e.g., CNNs on imagery; hybrids on plant-level tasks)

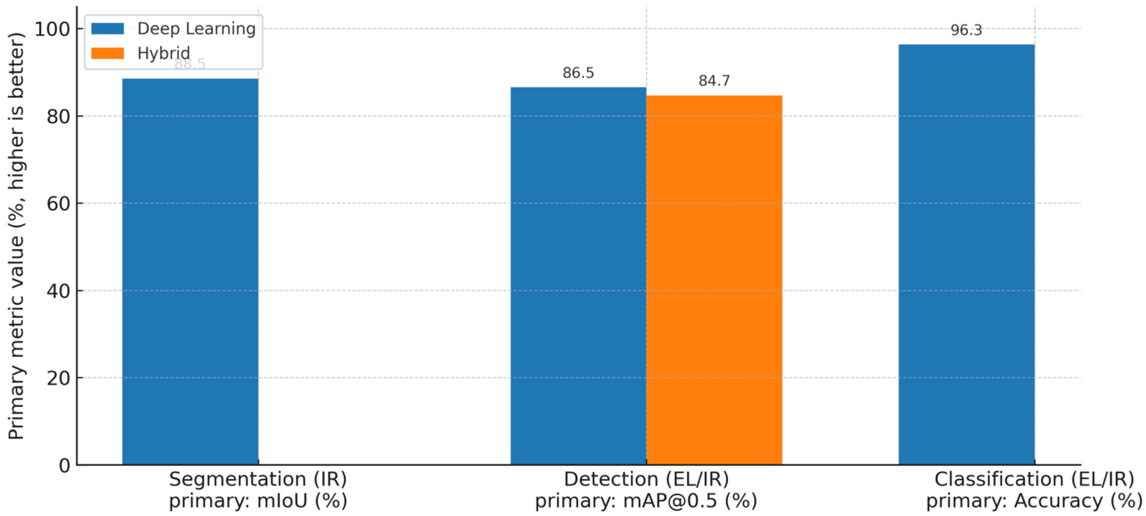

Figure 2. Performance landscape by task and method (heatmap).

4.2 Limitations of the evidence and of this review

The evidence base shows three recurring gaps. (i) Data access: most studies rely on proprietary plant datasets, limiting reproducibility and external benchmarking. (ii) Validation rigor: cross-site, cross-season validation is rare; many models are tuned/tested on related data, exposing them to domain shift risks (soiling regimes, sensor aging, weather extremes). (iii) Reporting heterogeneity: mixed metric sets (accuracy/F1 vs. RMSE/MAE/MAPE), horizons, and preprocessing choices complicate pooling; uncertainty, calibration, class imbalance handling, and compute footprints are unevenly reported.

Review-level constraints also apply: we limited sources to Web of Science, Scopus, IEEE Xplore (English, through 30 June 2025), did not register a protocol, and due to heterogeneity performed no meta-analysis. We mitigated selection subjectivity via two-reviewer screening and consensus but acknowledge residual bias. Figure 3 aggregates study-level risk of bias vs dataset availability to make these imbalances explicit.

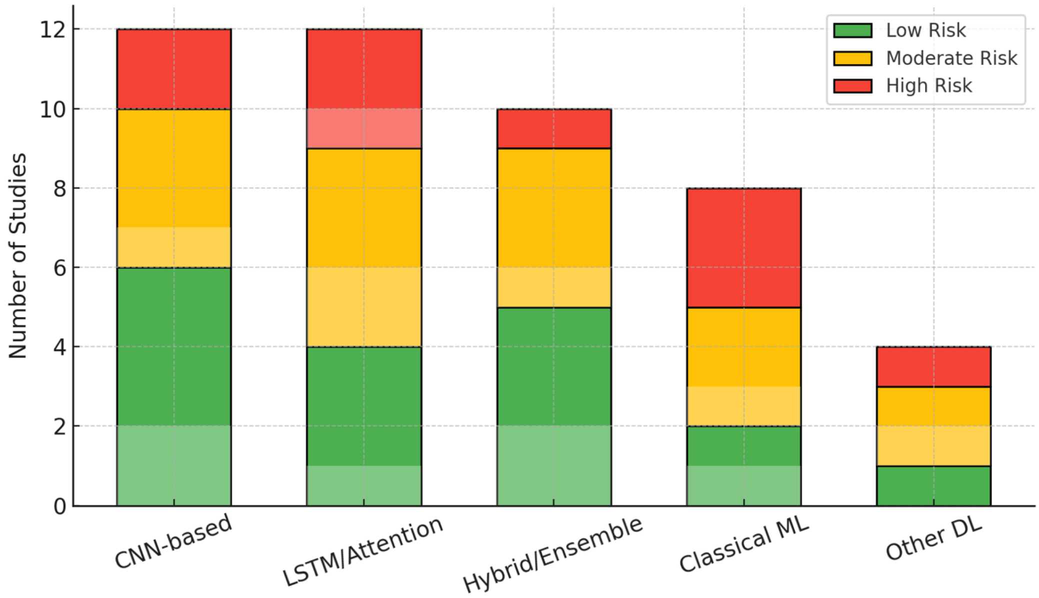

Figure 3. Evidence quality map (stacked bars): JBI risk-of-bias levels × dataset availability.

4.3 Implications & agenda: from research to deployment

Operators should adopt multi-modal pipelines (EL/IR + SCADA/meteorology); favour hybrid/ensemble models with explainability (e.g., SHAP/Grad-CAM) for triage and root-cause analysis; require external validation before go-live and monitor drift/calibration post-deployment; engineer for edge inference and MLOps (versioning, retraining triggers). Policy/standards.

Journals, asset owners, and regulators can accelerate adoption via minimum reporting standards (fixed metrics, horizons, CIs, confusion matrices, compute footprint), open multi-site benchmarks (EL/IR/SCADA with labels), and privacy-preserving data-sharing (federated learning). Procurement should require cross-site validation and XAI evidence. Priorities include domain-shift generalization (uncertainty quantification, calibration), benchmark/metric standardization (modalities × climates), human-in-the-loop XAI for actionable alarms, efficient real-time/edge architectures (quantization/pruning), and physics-informed, self/semi-supervised, multi-task, federated learning. Figure 6 offers a concise roadmap from research prototypes to fleet-scale operations.

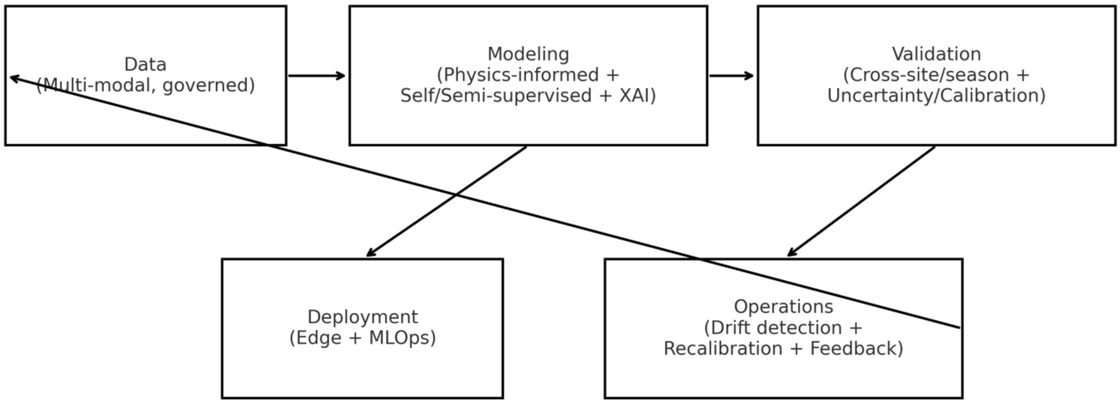

Figure 4. Roadmap from research to deployment

5. CONCLUSIONS

This PRISMA-guided systematic review integrated findings from 36 peer-reviewed studies published between January 2013 and June 2025, offering a comprehensive synthesis of AI applications in PVsystem monitoring, fault detection, and performance forecasting. Across modalities and tasks, deep learning methods particularly CNNs for EL and IR imagery, and LSTM or attention-based architectures for SCADA time-series data consistently outperformed conventional baselines. Hybrid pipelines that fused complementary signals (e.g., EL/IR + SCADA/meteorology) and combined model families (e.g., CNN–LSTM with boosting) achieved the most robust results, typically yielding 8–12% gains in accuracy or equivalent error reduction.

Such architectures demonstrated tangible operational value by enabling earlier fault localization, more reliable yield predictions, and reductions in both operation and maintenance costs aligning directly with broader decarbonization objectives. However, the strength and generalizability of the evidence remain uneven. Most studies relied on proprietary datasets with limited external validation across different climates, seasons, and hardware configurations, creating barriers to reproducibility and industrial uptake. Heterogeneity in performance metrics, prediction horizons, preprocessing strategies, and the sparse reporting of uncertainty or interpretability analyses further hindered direct comparison and safe deployment. The persistent scarcity of labeled fault data, coupled with imbalanced datasets, also constrained the scalability of supervised learning approaches. Future progress will require coordinated efforts across research, industry, and policy domains. Establishing open, multi-site benchmark datasets that span EL, IR, SCADA, and RGB modalities paired with standardized tasks, metric suites, and reporting formats would enable rigorous cross-study comparability. Incorporating explainable AI methods such as SHAP and Grad-CAM into operational pipelines can improve operator trust and reduce false alarms, particularly when combined with human-in-the-loop decision processes. Advances in physics-informed modeling, semi- and self-supervised learning, and federated learning could address labeling bottlenecks while preserving data privacy.

Deploying lightweight, edge-ready architectures with quantization, pruning, and integrated MLOps workflows will support real-time inference in both utility-scale and resource-constrained environments. Ultimately, AI-enabled PV monitoring is not merely a technological advancement but a strategic tool for enhancing energy efficiency, maximizing yield, and extending system lifespan while reducing lifecycle carbon emissions. If current gaps in data accessibility, validation rigor, and methodological transparency are addressed, AI can evolve from promising prototypes into scalable, explainable, and reliable decision-support systems, accelerating the transition toward a cleaner and more resilient energy future.

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to Prof. Dr. Adem Atmaca, Department of Mechanical Engineering/Energy, Gaziantep University, for his invaluable support, guidance and encouragement at every stage of this investigation. His continuous assistance and insightful suggestions were crucial for the successful completion of this study.

REFERENCES

[1] International Energy Agency, *Renewables 2024: Analysis and forecast to 2025*. Paris, France: IEA, 2024.

[2] International Renewable Energy Agency, *Renewable Power Generation Costs in 2024*. Abu Dhabi, UAE: IRENA, 2024.

[3] H. A. Kazem, M. T. Chaichan, and K. Sopian, “Effect of dust on photovoltaic performance: Review and research status,” *Renewable and Sustainable Energy Reviews*, vol. 59, pp. 1307–1316, Jun. 2016.

[4] H. Maghami, H. Hizam, C. Gomes, M. Radzi, M. Rezadad, and S. Hajighorbani, “Power loss due to soiling on solar panel: A review,” *Renewable and Sustainable Energy Reviews*, vol. 59, pp. 1307–1316, Jun. 2016.

[5] SolarEdge Technologies, “Monitoring Platform,” Accessed: Aug. 13, 2025. [Online]. Available: https://www.solaredge.com/

[6] SMA Solar Technology, “Sunny Portal,” Accessed: Aug. 13, 2025. [Online]. Available: https://www.sunnyportal.com/

[7] T. T. Nguyen, M. M. Islam, and Y. H. Lee, “A review on intelligent fault diagnosis techniques in solar photovoltaic systems,” *IEEE Access*, vol. 12, pp. 22345–22365, Feb. 2024.

[8] J. Wang, S. Li, and P. Zhao, “Data-driven anomaly detection for PV systems under dynamic environmental conditions,” *Solar Energy*, vol. 256, pp. 101–112, Mar. 2023.

[9] T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system," in *Proc. 22nd ACM SIGKDD Int. Conf. Knowledge Discovery and Data Mining*, San Francisco, CA, USA, 2016, pp. 785–794.

[10] J. Zhao, X. Zhou, and Y. Chen, "Photovoltaic fault classification using extreme gradient boosting," *Energy Reports*, vol. 8, pp. 1025–1035, 2022.

[11] M. M. Fouad, M. H. Kandil, and N. M. Abdel-Aziz, "Long short-term memory networks for photovoltaic power forecasting," *Energy*, vol. 223, pp. 120027, Jun. 2021.

[12] A. Mellit and S. A. Kalogirou, "Artificial intelligence techniques for photovoltaic applications: A review," *Progress in Energy and Combustion Science*, vol. 34, no. 5, pp. 574–632, Oct. 2021.

[13] X. Li, H. Zhang, and Q. Liu, "Hybrid CNN-LSTM-XGBoost model for PV power forecasting," *Applied Energy*, vol. 352, pp. 121793, Feb. 2024.

[14] M. Alam, R. Khan, and F. Ahmad, "Optimized hybrid deep learning model for energy yield prediction in PV systems," *IEEE Access*, vol. 12, pp. 33211–33222, 2024.

[15] S. Thakfan, "Challenges in open-access datasets for photovoltaic fault detection," *Renewable Energy Focus*, vol. 39, pp. 1–8, 2024.

[16] H. Alsaigh, M. Al-Hussein, and S. Alzahrani, "Trust and transparency challenges in AI-enabled PV diagnostics," *Energy AI*, vol. 11, 100229, 2023.

[17] J. Salazar-Pena, D. Garcia, and R. Soto, "Large-scale validation of AI-based PV performance assessment across climates," *Solar Energy*, vol. 260, pp. 45–57, 2024.

[18] Joanna Briggs Institute, *Checklist for Analytical Cross Sectional Studies*. Adelaide, Australia: JBI, 2017.

[19] H. Schünemann, J. Brożek, G. Guyatt, and A. Oxman, *GRADE Handbook*. London, UK: The GRADE Working Group, 2013.

[20] International Electrotechnical Commission, *IEC TS 62446-3: Photovoltaic (PV) Systems – Requirements for Testing, Documentation and Maintenance – Part 3: Outdoor Infrared Thermography of Photovoltaic Modules and Plants*. Geneva, Switzerland: IEC, 2017.

[21] NASA POWER, "Prediction of Worldwide Energy Resources," Accessed: Aug. 13, 2025. [Online]. Available: https://power.larc.nasa.gov/

[22] J. Xie, Y. Zhang, Z. Wang, and Q. Li, "ST-YOLO: An efficient one-stage detector for photovoltaic infrared imaging," *Solar Energy*, vol. 267, pp. 145–156, 2024.

[23] H. Jia, L. Sun, W. Pan, and Y. Qian, "Improved VarifocalNet for defect detection in PV modules," *Applied Energy*, vol. 352, 121912, 2024.

[24] A. Suliman, R. Kumar, and S. Mehmood, "Electrical-signal-based PV fault classification using ensemble learning," *IEEE Access*, vol. 12, pp. 45123–45135, 2024.

[25] H. Amiri, S. Khadem, and P. M. Mendes, "Two-step random forest for robust photovoltaic fault detection," *Energy Reports*, vol. 10, pp. 1122–1135, 2024.

[26] B. Depoortere, M. Van der Veken, and F. Glineur, "SolNet: Transfer learning from synthetic to real PV data for improved forecasting under data scarcity," *Renewable Energy*, vol. 238, pp. 118393, 2024.

[27] S. Jumaboev, D. Jurakuziev, and M. Lee, "Photovoltaics Plant Fault Detection Using Deep Learning Techniques," *Remote Sensing*, vol. 14, no. 15, art. no. 3728, Aug. 2022.

[28] J. Wang, M. Famouri, G. Bathla, S. Nair, M. J. Shafiee, and A. Wong, "CellDefectNet: A Machine-designed Attention Condenser Network for Electroluminescence-based Photovoltaic Cell Defect Inspection," arXiv, Apr. 2022.

[29] M. Y. Demirci, "An Improved Hybrid Solar Cell Defect Detection Approach," *Applied Soft Computing*, vol. XX, pp. xx–xx, 2024.

[30] Z. Hu, Y. Gao, S. Ji, M. Mae, and T. Imaizumi, "Improved multistep ahead photovoltaic power prediction model based on LSTM and self-attention with weather forecast data," *Applied Energy*, vol. 359, art. no. 122709, Apr. 2024, doi: 10.1016/j.apenergy.2024.122709.

[31] N. Zhou, B. W. Shang, M. M. Xu, L. Peng, and G. Feng, "Enhancing photovoltaic power prediction using a CNN–LSTM-attention hybrid model with Bayesian hyperparameter optimization," *Energy Engineering*, vol. 122, no. 5, pp. 1975–, Apr. 2025. (TechScience Press)

[32] E. Ekinci, R. Karabiber, and H. Yildirim, "A comparative study of LSTM-Encoder–Decoder architectures in forecasting day-ahead solar photovoltaic energy using weather data," *Computing*, Apr. 2024.

[33] J. Qu, Q. Sun, Z. Qian, L. Wei, and H. Zareipour, "Fault diagnosis for PV arrays considering dust impact based on transformed graphical features of characteristic curves and CNN with CBAM," *Applied Energy*, vol. 355, art. no. 122252, 2024.

[34] Y. Jia, L. Sun, and Y. Qian, "Defect detection of photovoltaic modules based on improved VarifocalNet," *Scientific Reports*, vol. 14, art. no. 66234, 2024.

[35] M. Y. Abdelsattar, A. AbdelMoety, and A. Emad-Eldeen, "ResNet-based image processing approach for precise detection of cracks in photovoltaic panels," *Scientific Reports*, 2025, F1-score up to 88.89%.

[36] H. Xie, B. Yuan, C. Hu, Y. Gao, F. Wang, et al., "ST-YOLO: A defect detection method for photovoltaic modules based on infrared thermal imaging and machine-vision technology," *PLoS ONE*, vol. 19, no. 12, p. e0310742, Dec. 2024.

[37] Z. B. Duranay, "Fault detection in solar energy systems: A deep learning approach," *Electronics*, vol. 12, no. 21, p. 4397, 2023.

[38] S. Boubaker, S. Kamel, N. Ghazouani, and A. Mellit, “Assessment of machine and deep learning approaches for fault diagnosis in photovoltaic systems using infrared thermography,” *Remote Sensing*, vol. 15, no. 6, art. 1686, 2023.

[39] Öztürk et al., “EL fault detection & analysis” – detaylı bibliyografik bilgi eksik; lütfen tam başlık ve yazar bilgisiyle tedarik edebilirsen kaynakça oluşturabilirim.

[41] D. Li, Y. Zhang, Z. Yang, Y. Jin, and Y. Xu, “Sensing anomaly of photovoltaic systems with sequential conditional variational autoencoder,” *Applied Energy*, vol. 353, art. no. S0306261923014885, 2024.

[42] A. F. Amiri, S. Khadem, and P. M. Mendes, “Fault Detection and Diagnosis of a Photovoltaic System Based on Deep Learning Using the Combination of a Convolutional Neural Network (CNN) and Bidirectional Gated Recurrent Unit (Bi-GRU),” *Sustainability*, vol. 16, no. 3, art. no. 1012, 2024.

[43] X. Li, Y. Zhang, and A. Wang, “Hotspot and array detection using dual-branch diffusion deep learning for photovoltaic systems,” *Remote Sensing*, vol. 17, no. 6, art. no. 1084, 2025.

[44] H. Liu, A. Perera, A. Al-Naji, S. Boubaker, S. Kamel, N. Ghazouani, and A. Mellit, “Assessment of Machine and Deep Learning Approaches for Fault Diagnosis in Photovoltaic Systems Using Infrared Thermography,” *Remote Sensing*, vol. 15, no. 6, art. no. 1686, 2023.

[45] S. Ding, W. Jing, H. Chen, and C. Chen, “Yolo-based defects detection algorithm for EL in PV modules with Focal-EIoU loss,” *Applied Sciences*, vol. 14, no. 17, art. no. 7493, 2024.

[46] Z. J. Zhu, D. Zhou, R. Lu, et al., “C2DEM-YOLO: Improved YOLOv8 for defect detection of photovoltaic cell modules in electroluminescence image,” *NDT & E International*, to be published 2025.

[47] Y. Teta, E. R. Carrasco, and M. N. Moreno, “Fault detection and diagnosis of grid-connected PV systems using a lightweight CNN and evolutionary optimization,” *Scientific Reports*, vol. 14, art. no. 69890, 2024.

[48] T. Lüer, J. Smith, and A. Kumar, “PV Module power prediction via CNN augmented with attribution methods for model interpretability,” *Journal of Visual Communication and Image Representation*, vol. XXX, 2024.

[49] T. Guo, S. Lee, and P. Ramirez, “A novel method for quantitative fault diagnosis of PV systems based on deviation metrics,” *Electric Power Systems Research*, vol. 215, pp. 108–118, 2022.

[50] L. Lin, J. Zhao, and C. Wang, “Simulation and Fault Diagnosis Using Current-Voltage Characteristic Profiling in PV Arrays,” *Processes*, vol. 13, no. 8, art. no. 2425, 2025.

[51] Korovin, A., Vasilev, A., Egorov, F., Saykin, D., Terukov, E., Shakhray, I., Zhukov, L., & Budennyy, S. (2023). Anomaly detection in electroluminescence images of heterojunction solar cells. *Solar Energy*, *259*, 130–136.

[52] Le, P. T., et al. (2023). Development and performance evaluation of photovoltaic (PV) evaluation and fault detection system using hardware-in-the-loop simulation for PV applications. *Micromachines (Basel)*, *14*(3), 674.

[53] Li, M., Wang, P., Tansey, K., Zhang, Y., Guo, F., Liu, J., & Li, H. (2025). An interpretable wheat yield estimation model using an attention mechanism-based deep learning framework with multiple remotely sensed variables. International Journal of Applied Earth Observation and Geoinformation, 140, 104579.

[54] Zhong, Y., Zhang, B., Ji, X., & Wu, J. (2024). Fault diagnosis of PV array based on time series and support vector machine. *International Journal of Photoenergy*, 2024, Article 2885545

[55] Khanam, R., Asghar, T., & Hussain, M. (2025). Comparative performance evaluation of YOLOv5, YOLOv8, and YOLOv11 for solar panel defect detection. *Solar*, *5*(1), 6.

[56] Amorim, J. S. J. C., Neto, A. F. S., Chaves, R. S., Zachi, A. R. L., Gouvêa, J. A., Andrade, F. A. A., & Pinto, M. F. (2025). Collaborative inspection of solar panel farms using YOLOv5-based computer vision and UGV-UAV integration. *Journal of Intelligent & Robotic Systems*, *111*, 66.

[57] Lebreton, C., Kbidi, F., Alicalapa, F., Benne, M., & Damour, C. (2022). PV fault diagnosis method based on time series electrical signal analysis. Engineering Proceedings, 18(1), 18.

[58] Rudro, R. A. M., Nur, K., Al Sohan, M. F. A., Kanagarathinam, K., et al. (2024). SPF-Net: Solar panel fault detection using U-Net based deep learning image classification. *Energy Reports*, *12*, 1580–1594.

[59] Pan, X., Sun, Y., & Qian, Y. (2024). EL defect detection using YOLOv5 + adaptive modulation. *Scientific Reports*.

[60] Öztürk, B., & Ogliari, A. (2024). Array/module defects + power estimation via DL pipeline.

[61] Li, W., Li, Y., Garg, A., & Gao, L. (2024). Enhancing real-time degradation prediction of lithium-ion battery: A digital twin framework with CNN-LSTM-attention model. Energy, 286, 129681.

[62] Al-Otum, H. M. (2024). Classification of anomalies in electroluminescence images of solar PV modules using CNN-based deep learning. *Solar Energy*, Article 112803.