IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 28 April 2026, accepted 18 May 2026, date of publication 26 May 2026, date of current version 2 June 2026.

Digital Object Identifier 10.1109/ACCESS.2026.3697265

RESEARCH ARTICLE

A Compact 1D-CNN for Photovoltaic Fault Diagnosis With Leakage-Aware Validation: A Comparative Study

SYED ZULQADAR HASSAN[1], (Member, IEEE), BABAR SANA[2], TARIQ KAMAL[3,4], (Member, IEEE), ARSALAN MASOOD[2], MUHAMMAD AHMAD[5], (Member, IEEE), AND MD. SHAFIULLAH[5,6,7], (Senior Member, IEEE)

[1]Department of Computer Science, Faculty of Information Technology and Computer Science, University of Central Punjab, Lahore 54000, Pakistan
[2]Department of Electrical Engineering, University of Sialkot, Sialkot 51310, Pakistan
[3]Department of Electrical Engineering, School of Technology and Innovations, University of Vaasa, 65200 Vaasa, Finland
[4]Arcada University of Applied Sciences, 00560 Helsinki, Finland
[5]Interdisciplinary Research Center for Sustainable Energy Systems, King Fahd University of Petroleum and Minerals, Dhahran 31261, Saudi Arabia
[6]Department of Control and Instrumentation Engineering, King Fahd University of Petroleum and Minerals, Dhahran 31261, Saudi Arabia
[7]Interdisciplinary Research Center for Smart Mobility and Logistics, King Fahd University of Petroleum and Minerals, Dhahran 31261, Saudi Arabia

Corresponding author: Tariq Kamal (tariq.kamal@uwasa.fi)

The research activities of Tariq Kamal were supported by the University of Vaasa, Finland, through Business Finland under the project Grid Code Certification by Simulation, Grant No. 2452/31/2024. The open access publication fee was covered through the FinELib consortium's agreement with IEEE.

ABSTRACT Reliable photovoltaic fault diagnosis depends on classifiers that maintain high accuracy under varying environmental conditions while remaining suitable for lightweight implementation. This study presents a one-dimensional convolutional neural network (1D-CNN) for six-class photovoltaic fault detection using normalized electrical features and compares its performance with alternative methods including a linear multiclass support vector machine, inverse-distance k-nearest neighbors, a compact feedforward neural network and a wavelet-assisted decision tree. A photovoltaic array test system was simulated using 39 months of irradiance and temperature data comprising 105,213 records. Six operating states, namely normal operation, small line-to-line fault, large line-to-line fault, open-circuit fault, partial shading and bypass-diode anomaly, were generated using 46,334 operating points obtained after daylight filtering with $G \geq 80$ W/m^2. To prevent data leakage, samples derived from the same irradiance–temperature condition were assigned exclusively to a single subset among training, validation, or testing. The proposed 1D-CNN achieved an accuracy of 98.95% and a macro-F1 score of 98.95% on the held-out test set, with a serialized model size of 0.162 MB. These results indicate that compact one-dimensional convolution can effectively support the proposed weather-driven photovoltaic fault test system. However, validation using measured field-fault data remains necessary before establishing long-term deployment viability.

INDEX TERMS Photovoltaic fault diagnosis, 1D-CNN, multiclass classification, wavelet features, deep learning, renewable energy monitoring.

I. INTRODUCTION

The raising global demand for clean and sustainable energy has made renewable energy a key part of the energy transition. Among renewable energy technologies, solar photovoltaic (PV) generation has become one of the fastest-growing contributors to new renewable power capacity worldwide [1], [2]. The growing penetration of PV systems has also increased research on grid integration, maximum power point tracking (MPPT), intelligent control and optimization of renewable energy systems [3], [4], [5]. According to the International Energy Agency (IEA), annual solar PV and wind capacity

The associate editor coordinating the review of this manuscript and approving it for publication was Riccardo Carotenuto.

additions are projected to reach almost 710 GW by 2028, with solar PV remaining the largest contributor to renewable capacity growth [6]. Similarly, in the European Union, solar accounted for about one-fifth of renewable electricity generation in 2022 and remained the fastest-growing renewable electricity source [7]. Due to the continued expansion of PV systems, long-term system reliability, safety and efficiency have become significant priorities.

PV installations are exposed to many electrical, environmental and degradation-related faults, including open-circuit (OC) faults, line-to-line (LL) faults, partial shading, hot spots, bypass-diode anomalies and mismatch conditions. These faults can reduce energy production, accelerate module aging, increase maintenance costs and add safety risks in severe cases [8]. Recent PV fault-diagnosis studies also classify faults by their evolution, such as intermittent, abrupt, or incipient faults and show that these faults can happen across PV modules, connection lines, converters, inverters and grid-side components [9]. Electrical faults, particularly LL faults, are important because they can cause voltage collapse, current redistribution and abnormal mismatch among strings. Besides, MPPT operation may reduce fault-current levels, allowing some faults to remain hidden for longer periods and making detection more difficult under practical operating conditions [10]. Chronic operations such as partial shading and bypass-diode anomalies are also important because they may create mismatch losses and distinct electrical or thermal behavior; for instance, bypass-diode failure can make reverse-current and local-heating risks that differ from ordinary shading behavior [11].

Conventional PV fault-detection methods, including visual inspection, thermal imaging, electrical characterization and threshold-based comparisons, offer useful diagnostic capability. However, they are often limited in continuous, scalable and low-cost monitoring applications. Visual inspection is labor-intensive and difficult to apply continuously, while thermal imaging involves additional instrumentation and can be affected by environmental and operating conditions [12], [13]. Electrical characterization and threshold-based methods are attractive because they depend on readily available measurements. However, their performance can weaken when irradiance, temperature, shading effects and fault signatures overlap [14]. These limitations become more significant in large PV plants, where automated fault detection and classification must remain reliable under changing weather conditions, sensor noise and variable operating points.

Recent advances in machine learning (ML) and deep learning (DL) have improved the potential for automated PV fault detection and diagnosis. Classical ML models, including support vector machines (SVM), decision trees, random forests, k-nearest neighbors (KNN) and ensemble learning, have been widely applied to classify normal and faulty PV operating states [15]. Hybrid signal-processing and ML approaches have also been studied. For example, the Stockwell transform combined with SVM, decision trees, random forests and KNN has been used to identify partial shading, open-circuit conditions and degradation faults by extracting informative time–frequency features [16]. These studies show that feature quality is central to reliable PV fault diagnosis. In addition, soft-computing and rule-discovery methods have also been applied to obtain interpretable classification rules and support data-driven decision-making in engineering applications [17].

More recently, DL-based methods have been introduced to capture nonlinear and coupled fault signatures in PV measurements. A hybrid CNN–BiGRU architecture was proposed in [18] to capture both spatial and temporal dependencies in PV signals for the classification of open-circuit, short-circuit and shading faults. Similarly, stacked sparse autoencoders combined with an optimized multilayer perceptron were explored in [19], achieving high diagnostic accuracy across several PV fault categories. These studies show the diagnostic potential of intelligent learning models for PV systems. However, review studies also emphasize that the performance of intelligent FDD models depends strongly on the quality of extracted features, the availability of representative data and the ability of the model to handle noisy, uncertain and non-stationary operating conditions [9]. Feature-engineering and one-dimensional convolutional neural network (1D-CNN) approaches are particularly relevant in this context because normalized electrical features can improve model adaptability and enable compact classification architectures for PV fault diagnosis [20].

Despite these advances, several challenges remain. First, measured and well-labeled PV fault datasets are difficult to obtain because many fault events are rare, unsafe to reproduce, or poorly documented in field maintenance records. Second, models trained under limited laboratory or synthetic conditions may not generalize well to long-term environmental variability, including changes in irradiance, temperature, soiling and partial shading. Third, although many DL-based methods report high classification accuracy, they often provide limited information on deployment-related requirements such as model size, parameter count, inference speed and computational cost.

A further methodological concern is the validation protocol itself. In ML-based science, data leakage is a recognized source of overoptimistic results when information from the test distribution is inadvertently introduced during data collection, preprocessing or model evaluation [21]. This issue is particularly relevant when multiple PV operating classes are generated from the same irradiance-temperature operating point, because random sample-level splitting can place closely related variants in both training and testing subsets. Therefore, leakage-aware partitioning is necessary to obtain a more credible estimate of held-out model performance.

To address these issues, this paper presents a compact 1D-CNN-based comparative study for PV fault diagnosis using a leakage-aware validation strategy. The study uses a

39-month Lahore, Pakistan irradiance-temperature profile to represent outdoor environmental variability. These measured environmental data drive a 4 × 4 PV array, from which six operating states are generated: normal operation, small LL fault, large LL fault, OC fault, partial shading and bypass-diode anomaly. It is important to clarify that the environmental profile is measured, whereas the fault labels are generated through a physics-guided model. Therefore, this study focuses on test-system-based fault separability under realistic weather variation, rather than direct validation on measured field-fault data.

The proposed 1D-CNN is used as a compact feature-wise classifier. Unlike temporal CNNs applied to long time-series windows, the model in this work learns local interactions among normalized electrical descriptors, including irradiance, temperature, voltage, current, power and healthy-reference ratio features. The proposed model is compared with four baseline methods: linear multiclass SVM, inverse-distance KNN, a compact feedforward neural network and a wavelet-assisted decision tree. To reduce data leakage, grouped train/validation/test partitioning is applied so that all class variants generated from the same irradiance-temperature operating point remain within the same subset. In addition to accuracy and macro-F1, training time, inference throughput, serialized model size and parameter count are reported to support deployment-oriented model selection.

The remainder of this paper is arranged as follows. Section II explains the PV fault classes considered in this study. Section III presents the fault-detection techniques and candidate classifiers. Section IV provides the research methodology, including data preparation, PV array modelling, feature generation and grouped validation. Section V discusses the classification results and deployment trade-offs. Section VI concludes the paper and outlines future research directions.

II. FAULTS IN PV ARRAY SYSTEMS

PV arrays experience both acute electrical faults and chronic performance faults. Acute faults occur suddenly and often produce pronounced electrical deviations. On the other hand, chronic faults may develop gradually and remain partly masked via variations in irradiance and temperature. Recent review studies commonly group PV abnormalities into external, internal, and system-level categories. In these studies, fault detection, classification and localization are treated as related but distinct diagnostic tasks [8], [9], [14], [22].

In this work, both abrupt and chronic operating states are considered. Among the six operating classes, the normal class represents the healthy reference condition of the 4 × 4 PV array under the corresponding measured irradiance-temperature operating point. The acute electrical fault classes include small LL faults, large LL faults and OC faults. LL faults occur when conductors or nodes at different potentials become unintentionally connected, producing voltage collapse, reverse-current redistribution and mismatch between adjacent strings [8], [23]. Their detection can be difficult in practical PV arrays because MPPT operation may reduce fault-current magnitudes and reshape the post-fault operating point [10]. OC faults interrupt the current path of a module or string, which reduces the delivered current while largely maintaining the OC voltage envelope [22]. In this study, LL behavior is divided into small-LL and large-LL classes to represent fault severity rather than fault location.

The chronic classes maintained in this study are partial shading and bypass-diode anomaly. Partial shading introduces nonuniform irradiance across cells or strings, creates multi-peak power characteristics and usually affects current ratios more strongly than voltage ratios. Bypass-diode-related anomalies distort the voltage transfer characteristics of the affected string and are practically important because they do not always produce the same abrupt current signatures as hard wiring faults [14]. Experiments also shows that partial shading and bypass-diode failure can produce different electrical and thermal signatures; in particular, bypass-diode failure may produce reverse current and severe local heating under certain operating conditions [11]. In this study, the bypass-diode class represents a controlled electrical anomaly associated with altered string-voltage behavior. It is not intended to differentiate between specific diode failure mechanisms, such as OC, short-circuit, or thermally degraded bypass diodes.

Although other chronic conditions such as degradation, hot spots, soiling, snow cover, and corrosion are also important in the broader PV literature [12], [13], [24], they are not included as separate classes in this study. These conditions often require thermal imagery, visual inspection, insulation measurements, long-term degradation records or field maintenance labels for reliable diagnosis, which are outside the available low-dimensional electrical-feature scope. The selected six-class test system therefore provides a practical balance between diagnostic relevance, reproducibility and compatibility with the available low-dimensional electrical descriptors.

Similar classification difficulty appears in other electrical assets, where chronic insulation defects can produce weak and highly similar signal signatures. For example, STFT scalograms combined with CNN classification have been used to distinguish cable partial-discharge defects such as crack, contamination, air void, microcrack, and composite defects [25]. This supports the future extension of the proposed PV framework toward additional chronic fault classes when temporal electrical, thermal or field-maintenance labels become available.

A more generalized classification of the main PV faults is illustrated in figure 1. In this work, a 39-month irradiance–temperature dataset from Lahore, Pakistan, define the environmental operating profile, whereas the six fault labels are generated using the physics-guided PV model described in the methodology section. This feature avoids

overstating field-fault validation while still exposing the classifiers to realistic daily and seasonal environmental variability.

III. FAULT DETECTION TECHNIQUES

PV faults on the DC side remain harder to diagnose than AC-side disturbances because the measured voltage, current and power of a PV array change continuously with irradiance and temperature. Recent PV diagnosis reviews therefore organize the field into measurement-comparison methods, energy-loss methods, thermal and imaging methods, signal-injection methods and data-driven methods used for detection, classification and localization [14], [22], [26], [27], [28]. Direct difference measurement compares measured and expected electrical quantities, adjacent-string comparison uses neighboring strings as references, PV energy-loss analysis tracks abnormal loss patterns, temperature-based methods exploit thermal changes and external-injection methods study the response of the array to an auxiliary signal [14], [22], [26]. In this study, the final decision stage is data-driven and compares five classifiers: a linear multiclass SVM, inverse-distance KNN, ANN-light, a wavelet-assisted decision tree and a 1D-CNN.

At each operating point, the common electrical descriptor is written as

$$\mathbf{x} = [G, T, V, I, P, V_r, I_r, P_r]^\top$$
$$V_r = \frac{V_f}{V_n}, \quad I_r = \frac{I_f}{I_n}, \quad P_r = \frac{P_f}{P_n} \tag{1}$$

where G and T are irradiance and temperature, (V, I, P) are the measured electrical quantities and (V_n, I_n, P_n) denote the corresponding healthy operating values. All feature-space models are trained with z-score standardization,

$$\tilde{x}_{i,j} = \frac{x_{i,j} - \mu_j}{\sigma_j} \tag{2}$$

where μ_j and σ_j are computed from the training split only.

Linear multiclass SVM: Support vector machines remain attractive for PV fault classification when a compact and interpretable decision boundary is needed [23], [29]. The classifier used here follows a one-vs-all linear structure,

$$f_c(\tilde{\mathbf{x}}) = \mathbf{w}_c^\top \tilde{\mathbf{x}} + b_c, \qquad \hat{y}(\tilde{\mathbf{x}}) = \arg\max_c f_c(\tilde{\mathbf{x}}) \tag{3}$$

and each binary learner minimizes the soft-margin objective

$$\min_{\mathbf{w},b,\boldsymbol{\xi}\geq 0} \frac{1}{2}\|\mathbf{w}\|_2^2 + C\sum_{i=1}^{n}\xi_i \quad \text{s.t.} \quad y_i\left(\mathbf{w}^\top \tilde{\mathbf{x}}_i + b\right) \geq 1 - \xi_i \tag{4}$$

Because the feature vector is already low-dimensional and standardized, the linear form offers a useful baseline for how well the six operating classes can be separated without nonlinear expansion.

Inverse-distance KNN: KNN labels a query sample from the class composition of its nearest neighbors in normalized feature space and is often used as a simple local baseline in PV diagnosis studies [22], [30]. With Euclidean distance and $k = 5$,

$$d(\tilde{\mathbf{x}}, \tilde{\mathbf{x}}_i) = \|\tilde{\mathbf{x}} - \tilde{\mathbf{x}}_i\|_2, \quad \text{score}_c(\tilde{\mathbf{x}}) = \sum_{i\in\mathcal{N}_5(\tilde{\mathbf{x}})} \frac{\mathbb{1}\{y_i = c\}}{d(\tilde{\mathbf{x}}, \tilde{\mathbf{x}}_i) + \varepsilon} \tag{5}$$

and the final decision is

$$\hat{y}(\tilde{\mathbf{x}}) = \arg\max_c \ \text{score}_c(\tilde{\mathbf{x}}) \tag{6}$$

where $\varepsilon > 0$ prevents division by zero. This model is easy to interpret but becomes sensitive when chronic and acute classes overlap in the same local neighborhood.

ANN: Compact feedforward ANN models continue to be relevant in nonlinear PV applications because they offer strong approximation ability without the overhead of heavier deep networks [18], [19], [31]. The ANN used in this work receives the standardized eight-dimensional vector and applies two hidden layers with 24 and 12 neurons:

$$\begin{aligned} \mathbf{h}_1 &= \sigma(\mathbf{W}_1\tilde{\mathbf{x}} + \mathbf{b}_1) \\ \mathbf{h}_2 &= \sigma(\mathbf{W}_2\mathbf{h}_1 + \mathbf{b}_2) \\ \mathbf{p} &= \text{softmax}(\mathbf{W}_3\mathbf{h}_2 + \mathbf{b}_3) \end{aligned} \tag{7}$$

with predicted class $\hat{y} = \arg\max_c p_c$. Training minimizes cross-entropy,

$$\mathcal{L}_{\text{CE}} = -\sum_{c=1}^{C} t_c \log(p_c) \tag{8}$$

using scaled conjugate gradient updates with validation-based early stopping. This choice keeps the model compact while preserving nonlinear decision capability.

Wavelet-assisted decision tree: Recent PV studies have shown that wavelet descriptors can improve sensitivity to environmental fluctuations and fault-related transients, especially when paired with lightweight classifiers [16], [32]. In this study, irradiance and temperature sequences are decomposed by a level-3 db4 maximal-overlap wavelet transform and summarized through four descriptors:

$$\begin{aligned} E_G &= \sum_{\ell=1}^{3}\left|W_G^{(\ell)}\right|, \quad E_T = \sum_{\ell=1}^{3}\left|W_T^{(\ell)}\right|, \\ D_G &= \left|W_G^{(1)}\right|, \quad D_T = \left|W_T^{(1)}\right| \end{aligned} \tag{9}$$

which extend the base feature vector as

$$\mathbf{x}_w = [\mathbf{x}, E_G, E_T, D_G, D_T]^\top, \quad \text{Gini}(S) = 1 - \sum_{c=1}^{C} p_c^2 \tag{10}$$

A decision tree is then trained on $\mathbf{x}_w$ using the Gini impurity criterion. This branch remains interpretable and computationally light while bringing limited temporal context into the classifier.

1D-CNN: Modern PV fault studies increasingly favor 1D-CNN models because they learn local cross-feature

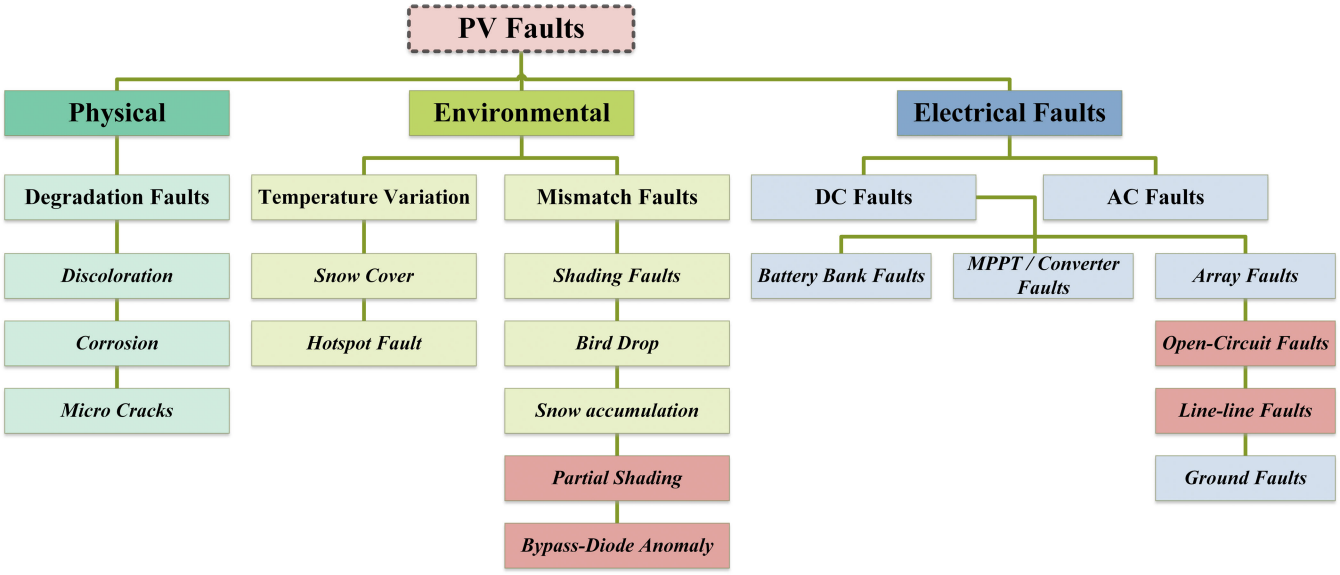

FIGURE 1. Generalized PV fault taxonomy [8].

patterns directly from low-dimensional signals and can be deployed on constrained hardware [18], [20], [33], [34]. Here, the standardized feature vector is treated as a short one-dimensional sequence and the forward path is

$$\begin{aligned}\mathbf{u}_1 &= \mathrm{ReLU}\big(\mathrm{Conv}_1(\tilde{\mathbf{x}})\big)\\ \mathbf{u}_2 &= \mathrm{ReLU}(\mathrm{Conv}_2(\mathbf{u}_1))\\ \mathbf{g} &= \mathrm{ReLU}\big(\mathbf{W}_{fc}\,\mathrm{Dropout}(\mathbf{u}_2; 0.10) + \mathbf{b}_{fc}\big)\\ \mathbf{p} &= \mathrm{softmax}(\mathbf{W}_o\mathbf{g} + \mathbf{b}_o)\end{aligned} \tag{11}$$

with $\hat{y} = \arg\max_c p_c$. The network uses two convolutional layers, a compact fully connected layer and Adam optimization. This design is intended to keep memory usage low while improving class separation beyond what dense or distance-based models can provide.

IV. RESEARCH METHODOLOGY

This section describes the data preparation, PV array modeling, feature generation and training procedure used in the study. The environmental operating points drive a 4×4 series-parallel PV array model and are expanded into six operating classes: normal, small line-to-line, large line-to-line, open circuit, partial shading and bypass diode. A grouped split is then applied so that all class variants created from the same operating point remain in one partition only, which prevents leakage across environmental states.

A. PV ARRAY MODELING AND ITS INTERFACE TO THE CLASSIFIERS

The 4×4 series-parallel PV array used in this work is retained because it gives a simple but representative platform for generating consistent fault signatures. The healthy operating point is computed from the module maximum-power quantities rather than from a long one-diode derivation.

$$\begin{aligned}I_{\mathrm{mp,mod}} &= I_{\mathrm{mp,stc}}\frac{G}{1000}\left[1 + \alpha_I(T - 25)\right]\\ V_{\mathrm{mp,mod}} &= V_{\mathrm{mp,stc}}\left[1 + \beta_V(T - 25)\right]\end{aligned} \tag{12}$$

where $I_{\mathrm{mp,stc}} = 7.41$ A, $V_{\mathrm{mp,stc}} = 29.0$ V, $\alpha_I = 5 \times 10^{-4}$ and $\beta_V = -4 \times 10^{-3}$. For $N_s = 4$ and $N_p = 4$.

$$I_n = N_p I_{\mathrm{mp,mod}}, \quad V_n = N_s V_{\mathrm{mp,mod}}, \quad P_n = V_n I_n \tag{13}$$

These healthy values act as the reference for the six operating classes. The acute electrical states are assigned as

$$\begin{aligned}\text{Normal: } V_f &= V_n,\ I_f = I_n\\ \text{Small LL: } V_f &= 0.75V_n,\ I_f = 1.05I_n\\ \text{Large LL: } V_f &= 0.50V_n,\ I_f = 1.10I_n\\ \text{Open Circuit: } V_f &= 0.98V_n,\ I_f = 0.75I_n\end{aligned} \tag{14}$$

while the chronic classes are generated through bounded severity factors

$$\begin{aligned}s_{\mathrm{ps}} &= \mathrm{clip}\big(0.25 + 0.35r_1 + 0.15\bar{G},\ 0.20,\ 0.75\big)\\ s_{\mathrm{bd}} &= \mathrm{clip}\big(0.35 + 0.25r_2 + 0.20\bar{T},\ 0.25,\ 0.85\big)\end{aligned} \tag{15}$$

where r_1 and r_2 are uniform random variables, $\bar{G} = \mathrm{clip}((G - 200)/800, 0, 1)$ and $\bar{T} = \mathrm{clip}((T - 10)/45, 0, 1)$. The chronic operating states are then written as

$$\begin{aligned}\text{Partial Shading: } V_f &= V_n(0.88 - 0.04s_{\mathrm{ps}}),\\ I_f &= I_n(0.55 - 0.18s_{\mathrm{ps}})\\ \text{Bypass Diode: } V_f &= V_n(0.62 - 0.10s_{\mathrm{bd}}),\\ I_f &= I_n(0.90 - 0.05s_{\mathrm{bd}})\\ P_f &= V_f I_f.\end{aligned} \tag{16}$$

Bounded measurement noise is finally injected into V_f, I_f and P_f to emulate nonideal sensing. The resulting feature

TABLE 1. Grouped split counts for the six operating classes.

Class	Train	Validation	Test
Normal	27569	4865	13900
Small LL	27569	4865	13900
Large LL	27569	4865	13900
Open Circuit	27569	4865	13900
Partial Shading	27569	4865	13900
Bypass Diode	27569	4865	13900
Total samples	165,414	29,190	83,400

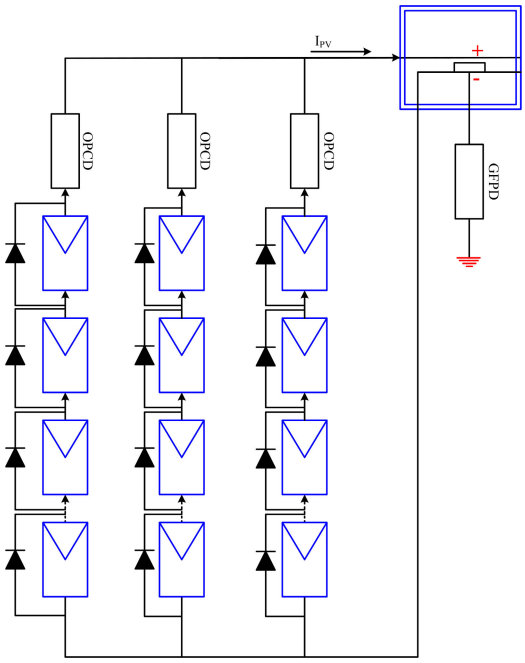

FIGURE 2. 4 × 4 PV array in normal operation.

set is $[G, T, V, I, P, V_r, I_r, P_r]$, where the ratio terms are referenced to (V_n, I_n, P_n).

The grouped partitioning follows the operating-condition index rather than the sample index. A 30% grouped holdout is reserved for testing and the remaining 70% is further divided with a 15% validation fraction. This yields 27,569 training operating points, 4,865 validation operating points and 13,900 test operating points. Since each operating point is replicated across six classes, the final sample counts are summarized in Table 1. This split keeps all class variants generated from the same environmental point in the same partition and therefore avoids leakage across train, validation and test sets.

The grouped split keeps all class variants generated from the same environmental point in the same partition and therefore avoids leakage across train, validation and test sets. The healthy 4 × 4 PV array reference used for normal operation is shown in Figure 2, while Figure 3 shows the corresponding representative faulted operating condition used to generate the six-class fault dataset.

B. MULTI CLASS SVM TRAINING

The SVM branch uses the shared eight-feature vector $\mathbf{x} = [G, T, V, I, P, V_r, I_r, P_r]^\top$ after z-score normalization with

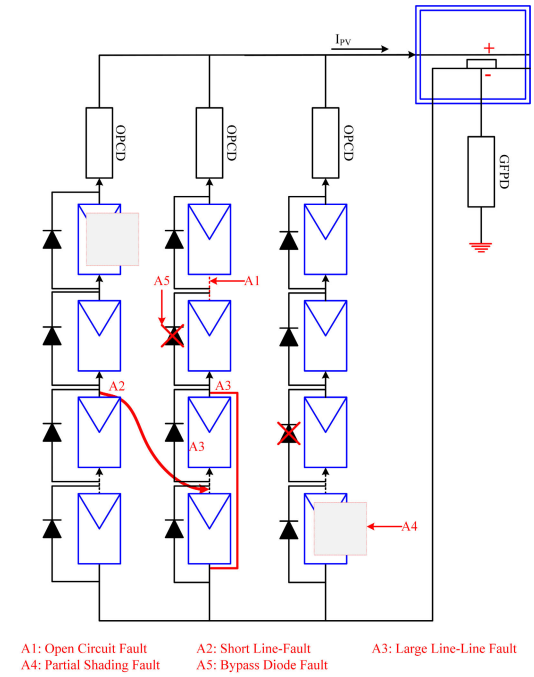

FIGURE 3. 4 × 4 PV array under representative faulted operation.

training-set statistics only. A linear error-correcting output code (ECOC) structure with one-vs-all coding is adopted because it is consistent with the low-dimensional feature space and provides a strong, reproducible baseline. The main steps are:

1) Load the grouped train, validation and test partitions generated from the operating-condition index.
2) Standardize the eight predictors using the training mean and standard deviation and apply the same transform to validation and test samples.
3) Train a linear multiclass SVM with box constraint $C = 1$ using one-vs-all coding.
4) Select the final model on the grouped validation split without mixing samples derived from the same environmental operating point.
5) Evaluate the grouped test split using confusion matrices, class-wise predictive values, overall accuracy, macro-F1, train time and inference throughput.

Figure 4 summarizes the SVM training flow from grouped data loading through standardized test-set evaluation.

C. K NEAREST NEIGHBORS

The KNN branch uses the same eight standardized predictors and preserves the same grouped split. The final configuration uses Euclidean distance, inverse-distance weighting and $k = 5$. The main steps are:

1) Import the grouped train, validation and test sets and standardize the predictors with training statistics.
2) Compute Euclidean distances in the normalized feature space and identify the $k = 5$ nearest neighbors for each query sample.

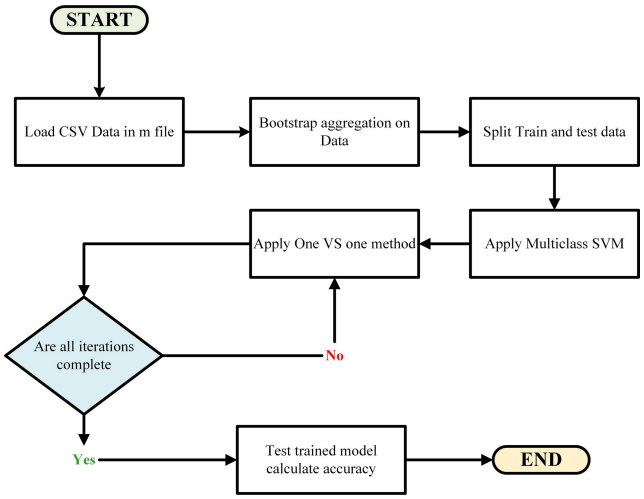

FIGURE 4. SVM flow chart.

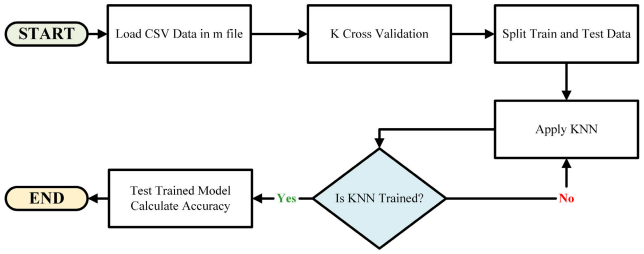

FIGURE 5. KNN flow chart.

3) Weight neighbor votes by inverse distance so that nearer operating points contribute more strongly than farther ones.
4) Tune and confirm the final k and weighting choice on the grouped validation data.
5) Evaluate the grouped test set using the same metrics reported for SVM.

Figure 5 summarizes the KNN workflow from grouped partition loading to inverse-distance voting and final evaluation.

D. ARTIFICIAL NEURAL NETWORK

The ANN branch is kept as a compact nonlinear baseline. The network receives the same eight standardized predictors, uses two hidden layers with 24 and 12 neurons and is trained by scaled conjugate gradient backpropagation with cross-entropy loss. Validation-based early stopping with a patience of 12 epochs is used to limit overfitting. Figure 6 summarizes the ANN training workflow. The procedure is:

1) Prepare the grouped train, validation and test partitions using the common eight-feature vector.
2) Standardize the input features with the training-set statistics only.
3) Train a two-hidden-layer ANN with architecture [24, 12], scaled conjugate gradient updates and cross-entropy loss.

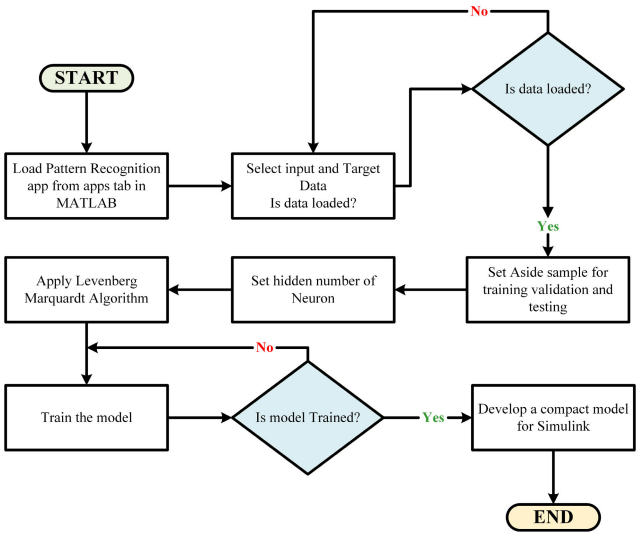

FIGURE 6. ANN working flow chart.

4) Monitor the validation subset and stop when the validation loss fails to improve for 12 consecutive checks.
5) Store the final trained network and report accuracy, macro-F1, serialized model size, parameter count and inference throughput in the results section.

E. WAVELET-ASSISTED DECISION TREE

The wavelet-assisted branch is added to introduce low-cost temporal context from the irradiance and temperature sequences. Four descriptors, namely WG_{Energy}, WT_{Energy}, WG_{d1} and WT_{d1}, are extracted from a level-3 db4 maximal-overlap wavelet transform and appended to the base electrical feature vector. This produces a 12-dimensional input for the tree classifier. The training sequence is:

1) Apply a level-3 db4 wavelet decomposition to the irradiance and temperature series.
2) Compute four summary descriptors from the detail and accumulated wavelet magnitudes.
3) Append these descriptors to the common eight-feature vector to form the 12-feature input.
4) Train a decision tree using the Gini split criterion with minimum leaf size 12.
5) Evaluate the grouped test split with the same confusion-matrix and deployment metrics used for the other models.

Figure 7 summarizes the wavelet-assisted decision-tree workflow, including wavelet descriptor extraction, feature concatenation and tree-based classification.

F. PROPOSED 1D-CNN

The proposed deep model receives the same standardized eight-feature vector but interprets it as a short one-dimensional sequence. The final architecture uses two convolutional layers with 16 and 32 filters, kernel size 3,

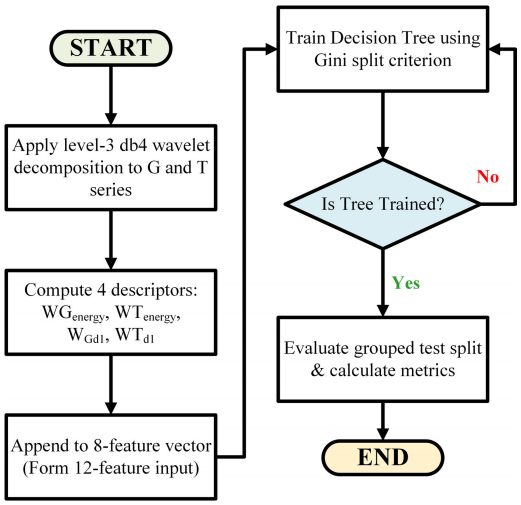

FIGURE 7. Wavelet-assisted decision-tree workflow.

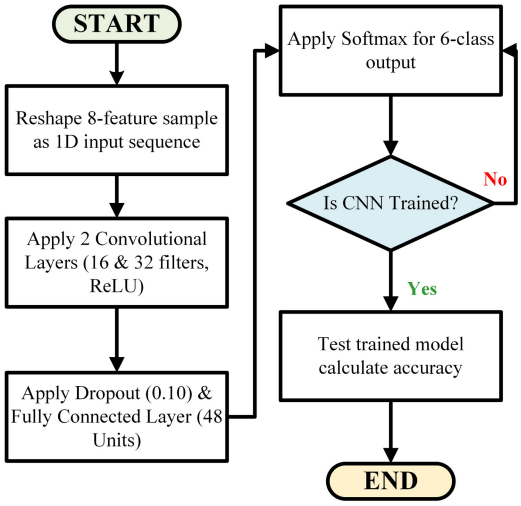

FIGURE 8. 1D-CNN architecture used for PV fault classification.

dropout 0.10, a fully connected layer of width 48 and a softmax output layer. Training is performed with Adam for 25 epochs using a mini-batch size of 64. The main steps are:

1) Reshape each standardized eight-feature sample as a short one-dimensional input sequence.
2) Apply two stacked convolutional layers with ReLU activations to learn local cross-feature patterns.
3) Use dropout and a compact fully connected layer to control model size and reduce overfitting.
4) Optimize the network with Adam on the grouped training split and monitor the grouped validation split during training.
5) Evaluate the grouped test split and report classification performance together with train time, inference throughput, serialized size and parameter count.

Figure 8 summarizes the 1D-CNN architecture used to capture local dependencies among normalized electrical features.

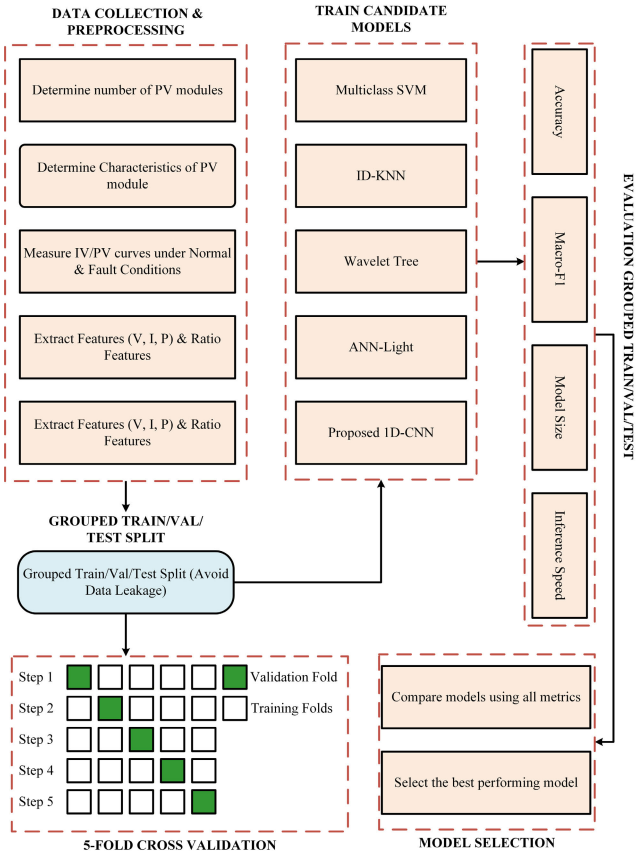

FIGURE 9. Flow chart for the five-technique PV fault classification pipeline.

G. MODEL WORKING FLOW CHART

The complete pipeline is summarized in figure 9. The flow starts from the 39-month irradiance-temperature file, generates the six operating classes, extracts the common electrical and ratio features, adds wavelet descriptors for the tree branch and then trains the five candidate models on grouped train, validation and test splits. The full execution order is:

1) Load the 39-month irradiance and temperature file and remove nighttime operating points using $G \geq 80$ W/m^2.
2) Compute the healthy V, I and P operating values of the 4×4 series-parallel PV array and generate six operating classes from the same environmental point.
3) Form the shared feature vector [G, T, V, I, P, V_r, I_r, P_r] and append four wavelet descriptors for the wavelet-assisted tree branch.
4) Create grouped train, validation and test partitions so that no operating point appears in more than one split.
5) Train the five candidate models: linear multiclass SVM, inverse-distance KNN, ANN-light, wavelet-assisted decision tree and 1D-CNN.
6) Evaluate all models using confusion matrices, predictive values versus false discovery rates, accuracy, macro-F1, train time, inference throughput, model size and parameter count.

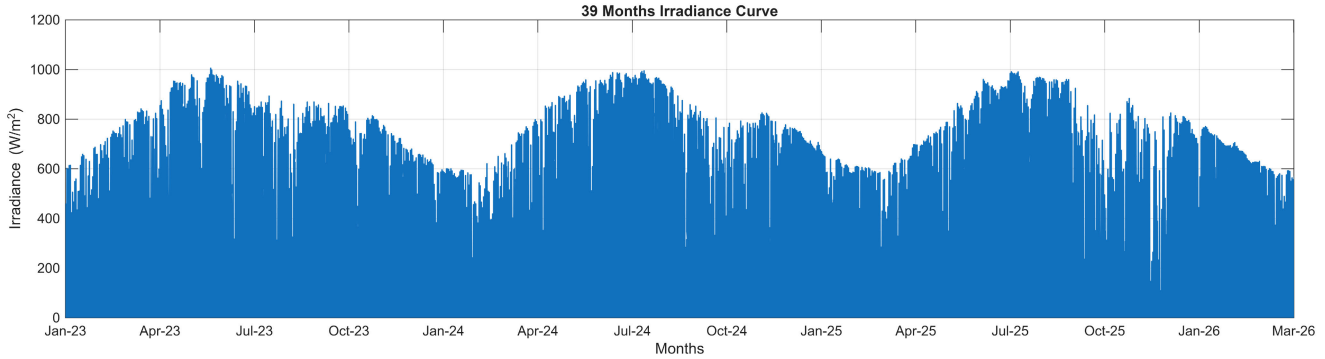

FIGURE 10. 39-month irradiance profile used to drive the PV fault classification study.

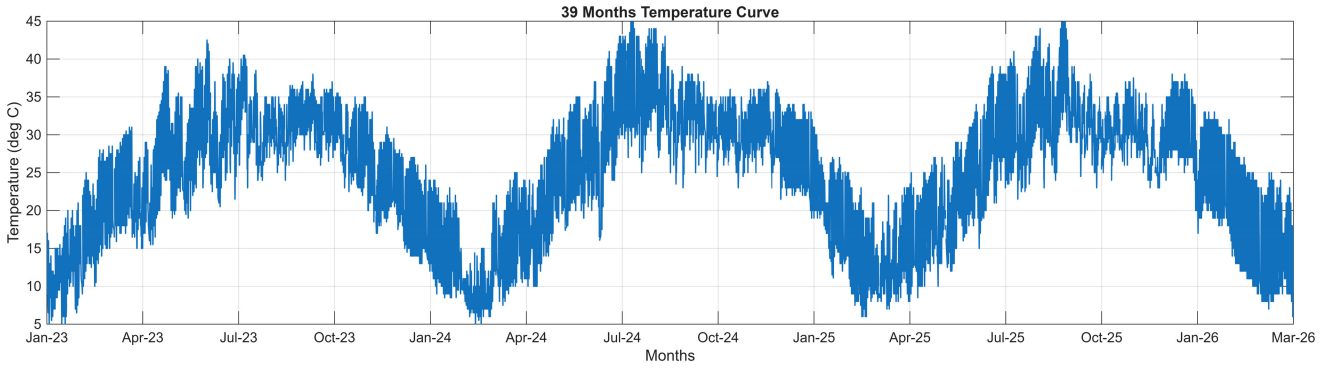

FIGURE 11. 39-month temperature profile used to drive the PV fault classification study.

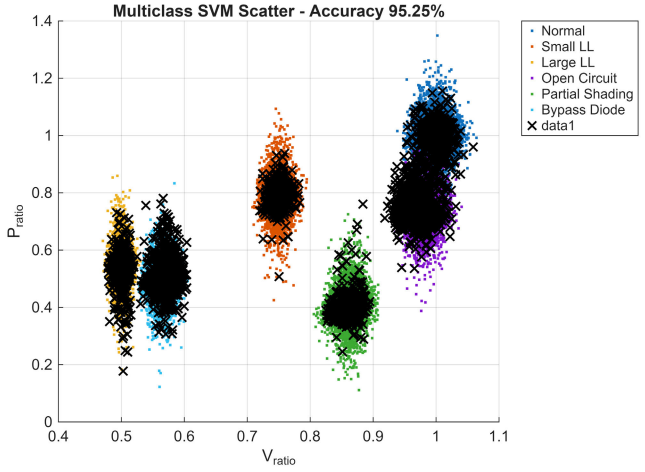

FIGURE 12. SVM scatter plot for the six-class study. Crosses indicate misclassified test samples.

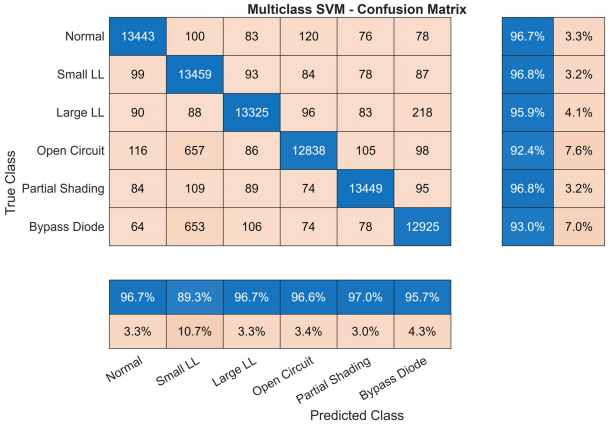

FIGURE 13. SVM confusion matrix for the six-class study.

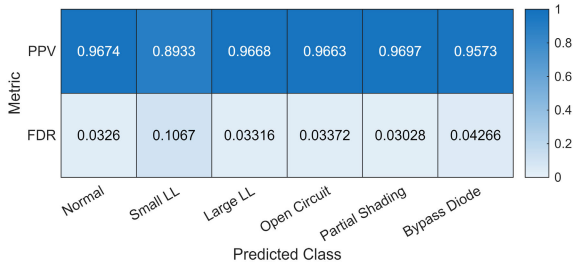

FIGURE 14. SVM predictive values versus false discovery rates for the six-class study.

The performance indices used throughout the paper are

$$\text{Accuracy} = \frac{\sum_{c=1}^{C} C_{cc}}{\sum_{i=1}^{C}\sum_{j=1}^{C} C_{ij}}$$

$$\text{Macro-F1} = \frac{1}{C}\sum_{c=1}^{C}\frac{2\,\text{Precision}_c\,\text{Recall}_c}{\text{Precision}_c + \text{Recall}_c} \tag{17}$$

where C_{ij} is the confusion-matrix entry for true class i predicted as class j. If N_{test} is the number of test samples and t_{inf} is the measured inference time, the throughput is

$$\text{Throughput} = \frac{N_{\text{test}}}{t_{\text{inf}}} \tag{18}$$

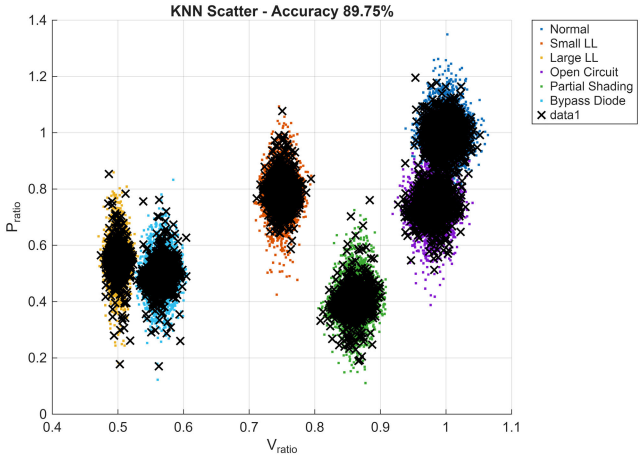

FIGURE 15. KNN scatter plot for the six-class study.

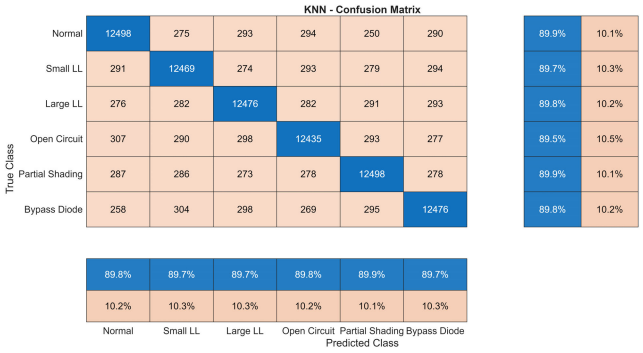

FIGURE 16. KNN confusion matrix for the six-class study.

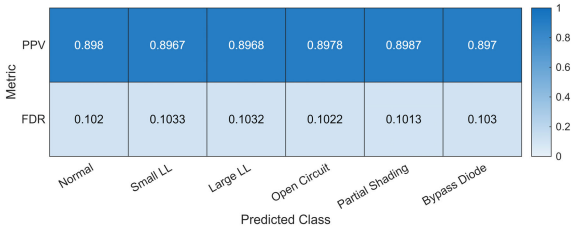

FIGURE 17. KNN predictive values versus false discovery rates for the six-class study.

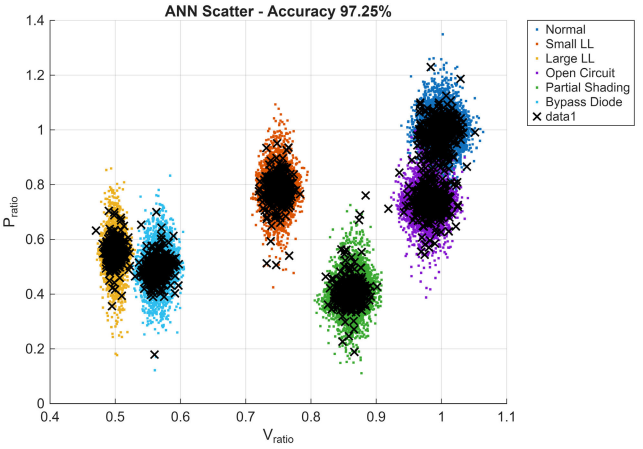

FIGURE 18. ANN scatter plot for the six-class study.

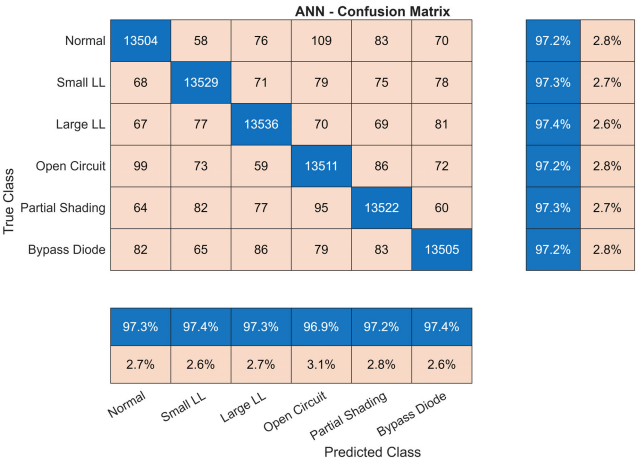

FIGURE 19. ANN confusion matrix for the six-class study.

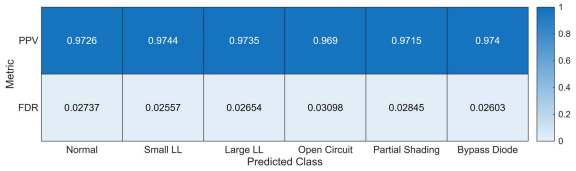

FIGURE 20. ANN predictive values versus false discovery rates for the six-class study.

These metrics are reported together so that classification quality and deployment cost can be compared on the same footing.

V. RESULTS AND DISCUSSION

Temperature data at latitude 31.5497° and longitude 74.3436° (i.e Lahore, Pakistan) were extracted from Weatherbit at sub-hourly resolution using its API [35]. Solar irradiance data were derived from open-source cloud-cover data obtained from the Copernicus Atmosphere Monitoring Service (CAMS) Atmosphere Data Store for the same site [36]. This study uses 39-month irradiance-temperature measurements containing 105,213 records at 15-minute resolution. After daylight filtering ($G \geq 80$ W/m^2), 46,334 operating points are transformed into a six-class criterion comprising normal, small line-to-line, large line-to-line, open-circuit, partial-shading and bypass-diode conditions. Each operating point is replicated across the six classes and kept in one partition only, which gives 165,414 training samples, 29,190 validation samples and 83,400 test samples, or 13,900 test samples per class. This grouped split ensures that samples generated from the same environmental state do not appear across multiple partitions. Figures 10 and 11 show the 39-month irradiance and temperature profiles used in the study.

1) Multiclass SVM Results: The linear multiclass SVM achieved 95.25% accuracy with 95.26% macro-F1 on the grouped test partition. Figure 12 shows that most

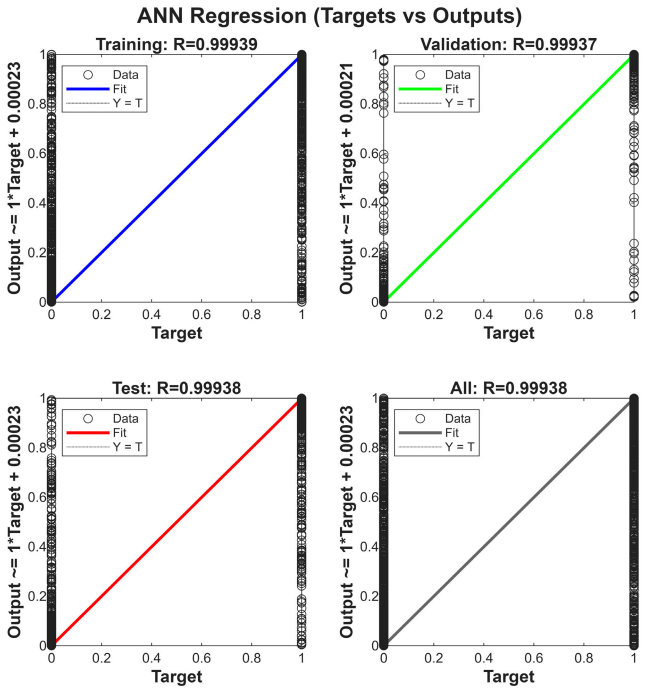

FIGURE 21. ANN regression plot.

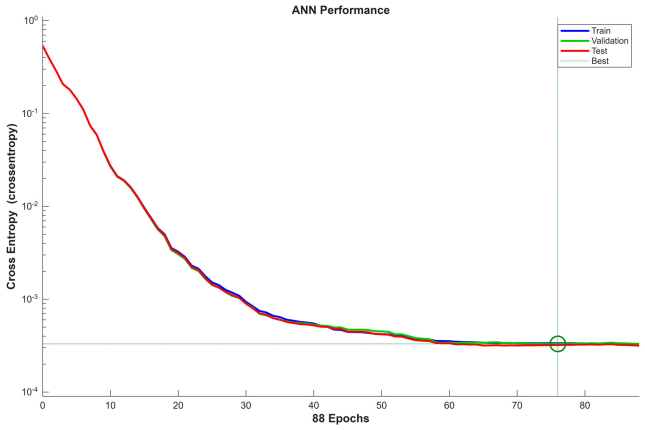

FIGURE 22. ANN performance curve.

samples remain well separated in the reduced ratio-feature space, although the crossed markers become denser around the reduced-voltage fault regions. The confusion matrix in Figure 13 remains strongly diagonal but the main off-diagonal leakage occurs among the small line-to-line, open-circuit and bypass-diode classes. The predictive-value versus false-discovery-rate map in Figure 14 confirms that the SVM is a strong and interpretable baseline but it is less selective than the neural models when class boundaries begin to overlap.

2) KNN Results: The inverse-distance KNN model achieved 89.75% accuracy with 89.75% macro-F1. The scatter plot in Figure 15 contains visibly more crossed samples than the SVM case, which shows that local voting is more sensitive to overlap between chronic and acute fault signatures. Figure 16 still shows a preserved diagonal but the

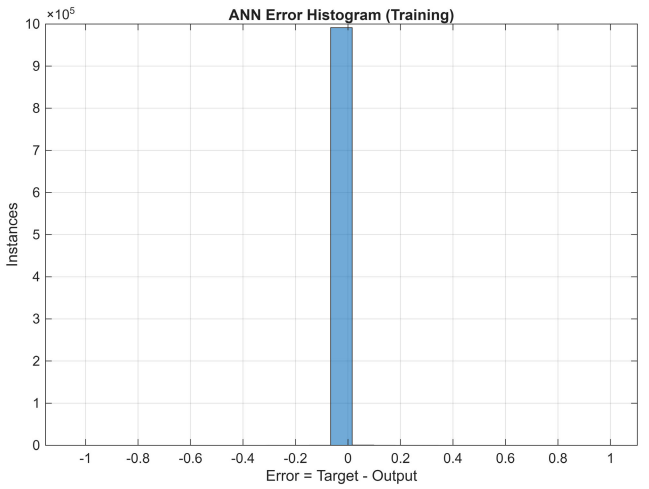

FIGURE 23. ANN error histogram.

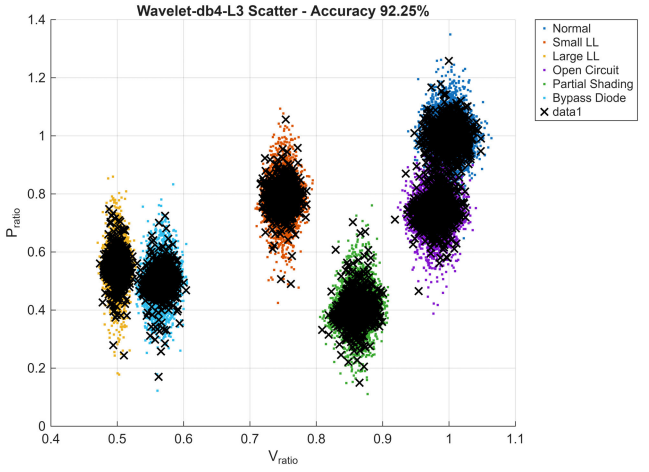

FIGURE 24. Wavelet-assisted tree scatter plot for the six-class study.

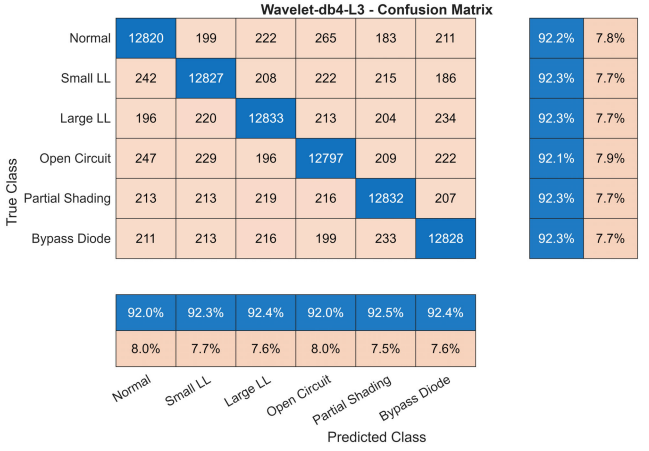

FIGURE 25. Wavelet-assisted tree confusion matrix for the six-class study.

off-diagonal leakage is broader across all six classes. The predictive-value plot in Figure 17 confirms that KNN is the least selective of the five techniques in this study.

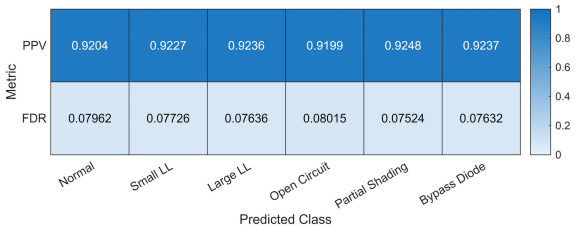

FIGURE 26. Wavelet-assisted tree predictive values versus false discovery rates for the six-class study.

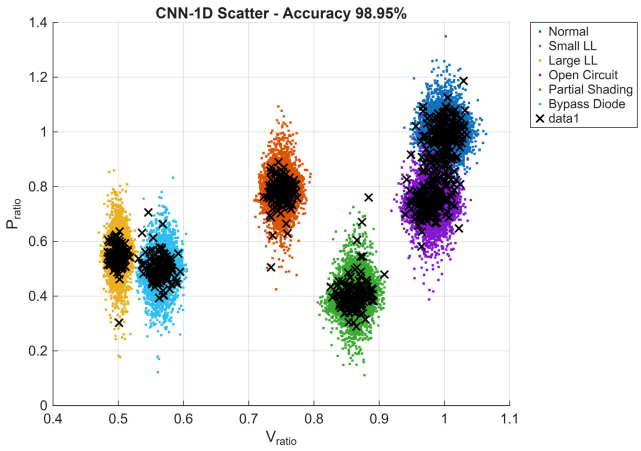

FIGURE 27. 1D-CNN scatter plot for the six-class study.

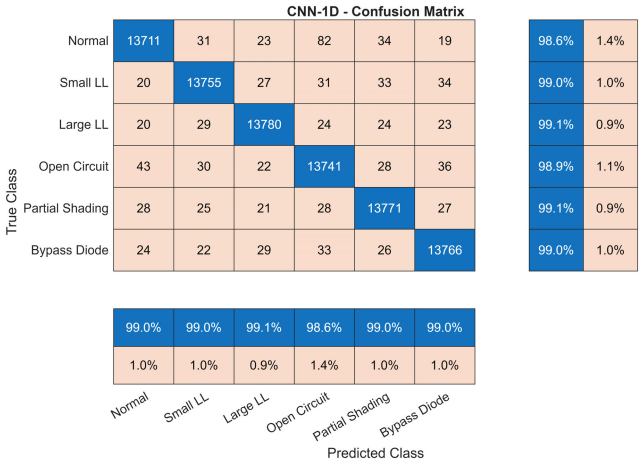

FIGURE 28. 1D-CNN confusion matrix for the six-class study.

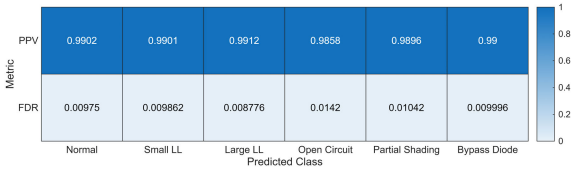

FIGURE 29. 1D-CNN predictive values versus false discovery rates for the six-class study.

3) Artificial Neural Network Results: The ANN-light model improved the classification performance to 97.25% accuracy with 97.25% macro-F1. Figure 18 shows a much tighter concentration of correctly classified samples than

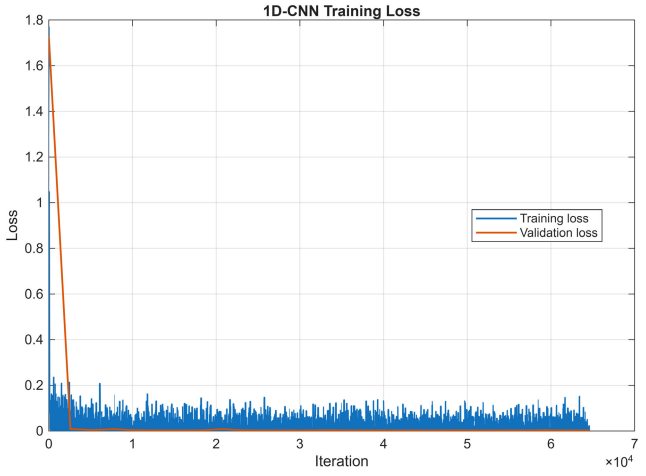

FIGURE 30. Training-loss curve of the proposed 1D-CNN.

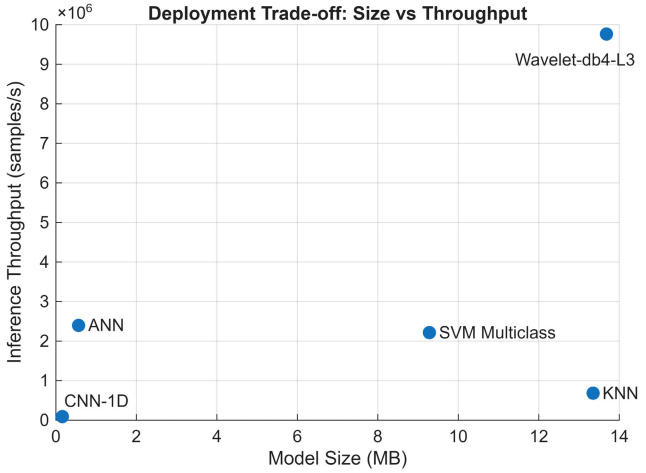

FIGURE 31. Deployment trade-off between serialized model size and measured inference throughput.

the classical baselines. The confusion matrix in Figure 19 remains balanced across all six classes with only small residual confusion counts. The predictive-value plot in Figure 20 shows consistently strong class-wise selectivity. In deployment terms, the trained ANN occupies 0.563 MB, contains 594 trainable parameters and reaches approximately 2.40×10^6 samples/s on the held-out test set. The regression plot in Figure 21, the performance curve in Figure 22 and the error histogram in Figure 23 provide additional convergence diagnostics and support stable optimization with low residual error.

4) Wavelet-Assisted Tree Results: The wavelet-assisted decision tree achieved 92.25% accuracy with 92.25% macro-F1. This is better than KNN but lower than the ANN. Figure 24 shows that the wavelet descriptors help preserve the coarse clustering structure, yet some boundary overlap still remains. The confusion matrix in Figure 25 indicates moderate confusion spread across all six classes, which suggests that the appended wavelet descriptors improve environmental awareness but do not fully resolve the nonlinear

TABLE 2. Comparative results for the five-technique PV fault classification study.

Model	Features	Accuracy (%)	Macro-F1 (%)	Train (s)	Throughput (samples/s)	Size (MB)	Params
Linear multiclass SVM	8	95.25	95.26	125.68	2,212,417	9.283	–
Inverse-distance KNN	8	89.75	89.75	0.20	681,982	13.349	–
ANN-light	8	97.25	97.25	19.14	2,396,385	0.563	594
Wavelet-assisted tree	12	92.25	92.25	0.18	9,766,121	13.682	–
Proposed 1D-CNN	8	98.95	98.95	412.95	93,229	0.162	14,262

class interactions. The predictive-value map in Figure 26 reflects the same middle-ground behavior. An important practical advantage of this model is speed, since it provides the highest measured inference throughput at approximately 9.77×10^6 samples/s.

5) Proposed 1D-CNN Results: The proposed 1D-CNN achieved the best overall result, namely 98.95% accuracy with 98.95% macro-F1. The scatter plot in Figure 27 shows that only a small number of samples fall outside their dominant class regions. The confusion matrix in Figure 28 remains sharply diagonal for all six classes and the predictive-value plot in Figure 29 confirms that the model is both accurate and selective. The training-loss curve in Figure 30 also shows stable optimization behavior. In deployment terms, the trained 1D-CNN is the smallest model in memory at only 0.162 MB, although its convolutional operations reduce the measured throughput to approximately 9.32×10^4 samples/s. Even so, the model offers the best balance when classification quality and memory footprint are both important.

The comparative performance of the five techniques is clear: the proposed 1D-CNN achieves the highest accuracy and macro-F1 score, followed by ANN-light, the linear SVM, the wavelet-assisted tree, and finally KNN. This ordering is consistent with the per-model figures. KNN suffers the broadest boundary leakage, the wavelet-assisted tree improves the classical local-voting baseline but does not close the gap to the neural models, the ANN provides a strong compromise between accuracy and throughput and the 1D-CNN yields the cleanest class separation.

From a deployment perspective, the choice depends on the target constraint. If the highest classification accuracy is required, the proposed 1D-CNN is the best option. If fast inference is the main requirement, the wavelet-assisted tree is attractive. If a neural model with strong accuracy and moderate resource demand is preferred, ANN-light offers the most practical balance. Figure 31 illustrates this trade-off between serialized model size and measured inference throughput.

The outdoor data provide a realistic operating envelope but the class labels are generated through structured fault perturbations rather than measured field-fault campaigns. For that reason, the reported accuracies demonstrate discriminative power on an outdoor-driven study, not a guaranteed fixed accuracy over the full 20–25 year lifetime of a PV asset. In practical long-term deployment, periodic recalibration, retraining or domain adaptation would still be required as modules age, sensors drift and additional chronic fault modes emerge.

Although the proposed 1D-CNN has lower throughput than the tree and ANN baselines, its measured throughput remains suitable for PV monitoring intervals that typically operate at sub-second, second-level, or minute-level sampling rates. Reporting inference speed and memory footprint follows recent real-time fault-classification practice, where deployment feasibility is assessed using latency together with diagnostic accuracy [37].

Table 2 consolidates the final comparison and explains the deployment meaning of the ranking. Accuracy and macro-F1 identify the 1D-CNN as the best classifier, but throughput and model size show different operating trade-offs. The wavelet-assisted tree is the fastest option at 9,766,121 samples/s, ANN-light gives a high-accuracy neural baseline with 0.563 MB memory use and the 1D-CNN gives the highest accuracy with the smallest serialized size of 0.162 MB. Therefore, the final model choice should depend on whether the monitoring unit prioritizes classification margin, inference speed or memory footprint.

VI. CONCLUSION

This work presented a weather-driven photovoltaic fault diagnosis study using a 39-month Lahore irradiance-temperature profile and a physics-guided 4×4 PV array PV system. Five models were compared: linear multiclass SVM, inverse-distance KNN, ANN-light, wavelet-assisted decision tree, and the proposed compact 1D-CNN. The study considered six operating states, including normal operation, small LL fault, large LL fault, OC fault, partial shading, and bypass-diode anomaly. To reduce data leakage, a grouped train/validation/test split was applied so that all class variants generated from the same environmental operating point remained within the same subset.

Among all tested models, the proposed compact 1D-CNN achieved the strongest held-out performance, with 98.95% accuracy and 98.95% macro-F1, while maintaining a serialized model size of 0.162 MB. ANN-light delivered the second-best classification result and provided a practical neural trade-off between speed and memory footprint. The linear SVM remained competitive as a simple and interpretable baseline, while the wavelet-assisted decision tree achieved the highest inference throughput. These results indicate that compact 1D convolution over normalized electrical descriptors is a promising approach for weather-driven PV fault classification when both accuracy and deployability are considered.

This study should be interpreted as a weather-driven test-system evaluation rather than a validation using measured

field-fault data. Although the irradiance–temperature profile is based on measured data, the fault labels are generated using a physics-guided PV model. Future work should extend the proposed framework through long-term field campaigns with measured fault labels, additional chronic degradation classes, external weather profiles, sensor-drift analysis, and domain-adaptive retraining strategies. Such extensions are needed before fixed long-term deployment accuracy can be claimed for practical PV assets.

CONFLICT OF INTEREST

No conflict of interest was declared by the authors.

DATA AVAILABILITY

The weather data used in this research is openly accessible in [35] and [36].

REFERENCES

[1] F. Mehmood, N. Ashraf, L. Alvarez, T. N. Malik, H. K. Qureshi, and T. Kamal, "Grid integrated photovoltaic system with fuzzy based maximum power point tracking control along with harmonic elimination," *Trans. Emerg. Telecommun. Technol.*, vol. 33, no. 2, Feb. 2022, Art. no. e3856.

[2] T. Kamal, M. Karabacak, F. Blaabjerg, S. Zulqadar Hassan, and L. M. Fernández-Ramírez, "A novel Lyapunov stable higher order B-spline online adaptive control paradigm of photovoltaic systems," *Sol. Energy*, vol. 194, pp. 530–540, Dec. 2019.

[3] T. Kamal, M. Karabacak, S. Z. Hassan, H. Li, and L. M. Fernández-Ramírez, "A robust online adaptive B-spline MPPT control of three-phase grid-coupled photovoltaic systems under real partial shading condition," *IEEE Trans. Energy Convers.*, vol. 34, no. 1, pp. 202–210, Mar. 2019.

[4] B. Butt, T. Kamal, S. Z. Hassan, and A. M. A. Masood, "Robust MPPT for photovoltaics: Interval type-2 fuzzy logic with experimental validation," in *Proc. Horizons Inf. Technol. Eng. (HITE)*, 2025, pp. 1–6.

[5] T. Kamal and S. Z. Hassan, "Special issue 'applications of advanced control and optimization paradigms in renewable energy systems,'" *Energies*, vol. 16, no. 22, 2023, Art. no. 7551.

[6] International Energy Agency (IEA). (2016). *Renewable Electricity Capacity Additions By Technology and Segment, 2016–2028*. [Online]. Available: https://www.iea.org/energy-system/renewables/solar-pv

[7] Eurostat. (2024). *Wind and Water Provide Most Renewable Electricity; Solar is the Fastest-Growing Energy Source, 2024*. [Online]. Available: https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Renewableenergystatistics#Windandwaterprovidemostrenewableelectricity.3Bsolaristhefastest-growingenergysource

[8] D. S. Pillai and N. Rajasekar, "A comprehensive review on protection challenges and fault diagnosis in PV systems," *Renew. Sustain. Energy Rev.*, vol. 91, pp. 18–40, Aug. 2018.

[9] M. Mansouri, M. Trabelsi, H. Nounou, and M. Nounou, "Deep learning-based fault diagnosis of photovoltaic systems: A comprehensive review and enhancement prospects," *IEEE Access*, vol. 9, pp. 126286–126306, 2021.

[10] D. S. Pillai and N. Rajasekar, "An MPPT-based sensorless line–line and line–ground fault detection technique for PV systems," *IEEE Trans. Power Electron.*, vol. 34, no. 9, pp. 8646–8659, Sep. 2019.

[11] C. G. Lee, W. G. Shin, J. R. Lim, G. H. Kang, Y. C. Ju, H. M. Hwang, H. S. Chang, and S. W. Ko, "Analysis of electrical and thermal characteristics of PV array under mismatching conditions caused by partial shading and short circuit failure of bypass diodes," *Energy*, vol. 218, Mar. 2021, Art. no. 119480.

[12] M. Cubukcu and A. Akanalci, "Real-time inspection and determination methods of faults on photovoltaic power systems by thermal imaging in Turkey," *Renew. Energy*, vol. 147, pp. 1231–1238, Mar. 2020.

[13] E. Kaplani, "Detection of degradation effects in field-aged c-si solar cells through IR thermography and digital image processing," *Int. J. Photoenergy*, vol. 2012, pp. 1–11, May 2012.

[14] G. M. El-Banby, N. M. Moawad, B. A. Abouzalm, W. F. Abouzaid, and E. A. Ramadan, "Photovoltaic system fault detection techniques: A review," *Neural Comput. Appl.*, vol. 35, no. 35, pp. 24829–24842, Dec. 2023.

[15] A. A. Al-Katheri, E. A. Al-Ammar, M. A. Alotaibi, W. Ko, S. Park, and H.-J. Choi, "Application of artificial intelligence in PV fault detection," *Sustainability*, vol. 14, no. 21, p. 13815, Oct. 2022.

[16] C. Saiprakash, S. R. Kumar Joga, A. Mohapatra, and B. Nayak, "Improved fault detection and classification in PV arrays using stockwell transform and data mining techniques," *Results Eng.*, vol. 23, Sep. 2024, Art. no. 102808.

[17] I. A. Fadel, H. Alsanabani, C. Öz, T. Kamal, M. İskefiyeli, and F. Abdien, "Hybrid fuzzy-genetic algorithm to automated discovery of prediction rules," *J. Intell. Fuzzy Syst.*, vol. 40, no. 1, pp. 43–52, Jan. 2021.

[18] A. F. Amiri, S. Kichou, H. Oudira, A. Chouder, and S. Silvestre, "Fault detection and diagnosis of a photovoltaic system based on deep learning using the combination of a convolutional neural network (CNN) and bidirectional gated recurrent unit (Bi-GRU)," *Sustainability*, vol. 16, no. 3, p. 1012, Jan. 2024.

[19] M. Bougoffa, S. Benmoussa, M. Djeziri, and O. Palais, "Hybrid deep learning for fault diagnosis in photovoltaic systems," *Machines*, vol. 13, no. 5, p. 378, Apr. 2025.

[20] Y. M. Ali and L. Ding, "Fault diagnosis in photovoltaic arrays: A robust and efficient approach using feature engineering and 1D-CNN," *Int. J. Electr. Power Energy Syst.*, vol. 170, Sep. 2025, Art. no. 110863.

[21] S. Kapoor and A. Narayanan, "Leakage and the reproducibility crisis in machine-learning-based science," *Patterns*, vol. 4, no. 9, Sep. 2023, Art. no. 100804.

[22] Y.-Y. Hong and R. A. Pula, "Methods of photovoltaic fault detection and classification: A review," *Energy Rep.*, vol. 8, pp. 5898–5929, Nov. 2022.

[23] A. Eskandari, J. Milimonfared, and M. Aghaei, "Line-line fault detection and classification for photovoltaic systems using ensemble learning model based on I-V characteristics," *Sol. Energy*, vol. 211, pp. 354–365, Nov. 2020.

[24] K. AbdulMawjood, S. S. Refaat, and W. G. Morsi, "Detection and prediction of faults in photovoltaic arrays: A review," in *Proc. IEEE 12th Int. Conf. Compat., Power Electron. Power Eng. (CPE-POWERENG)*, Apr. 2018, pp. 1–8.

[25] M. H. Saad, S. Hashima, A. I. Omar, M. M. Fouda, and A. Said, "Deep learning approach for cable partial discharge pattern identification," *Electr. Eng.*, vol. 107, no. 2, pp. 1525–1540, Feb. 2025.

[26] Z. Hassan, A. Amir, J. Selvaraj, and N. A. Rahim, "A review on current injection techniques for low-voltage ride-through and grid fault conditions in grid-connected photovoltaic system," *Sol. Energy*, vol. 207, pp. 851–873, Sep. 2020.

[27] A. Thakfan and Y. Bin Salamah, "Artificial-intelligence-based detection of defects and faults in photovoltaic systems: A survey," *Energies*, vol. 17, no. 19, p. 4807, Sep. 2024.

[28] M. M. Rahman and A. K. Srivastava, "Detection, classification, and localization of faults and failures in photovoltaic arrays: A review, comparative analysis, and path forward," *Discover Energy*, vol. 6, no. 1, p. 26, Feb. 2026.

[29] J. Cervantes, F. Garcia-Lamont, L. Rodríguez-Mazahua, and A. Lopez, "A comprehensive survey on support vector machine classification: Applications, challenges and trends," *Neurocomputing*, vol. 408, pp. 189–215, Sep. 2020.

[30] R. K. Halder, M. N. Uddin, M. A. Uddin, S. Aryal, and A. Khraisat, "Enhancing K-nearest neighbor algorithm: A comprehensive review and performance analysis of modifications," *J. Big Data*, vol. 11, no. 1, pp. 1–55, Aug. 2024.

[31] N. F. Ibrahim, M. M. Mahmoud, A. M. H. A. Thaiban, A. B. Barnawi, Z. M. S. Elbarbary, A. I. Omar, and H. Abdelfattah, "Operation of grid-connected PV system with ANN-based MPPT and an optimized LCL filter using GRG algorithm for enhanced power quality," *IEEE Access*, vol. 11, pp. 106859–106876, 2023.

[32] A. Teta, B. Korich, D. Bakria, N. Hadroug, A. Rabehi, M. Alsharef, M. Bajaj, I. Zaitsev, and S. S. M. Ghoneim, "Fault detection and diagnosis of grid-connected photovoltaic systems using energy valley optimizer based lightweight CNN and wavelet transform," *Sci. Rep.*, vol. 14, no. 1, p. 18907, Aug. 2024.

[33] B. Aljafari, P. R. Satpathy, S. B. Thanikanti, and N. Nwulu, "Supervised classification and fault detection in grid-connected PV systems using 1D-CNN: Simulation and real-time validation," *Energy Rep.*, vol. 12, pp. 2156–2178, Dec. 2024.

[34] Y. M. Ali, L. Ding, and S. Qin, "An efficient approach for diagnosing faults in photovoltaic array using 1D-CNN and feature selection techniques," *Int. J. Electr. Power Energy Syst.*, vol. 166, May 2025, Art. no. 110526.

[35] *Weatherbit Historical Subhourly Weather Data Api*, Weatherbit.io, Ukraine, 2026.

[36] Copernicus Atmosphere Monitoring Service. (2026). *Cams Solar Radiation Time-Series*. [Online]. Available: https://ads.atmosphere.copernicus.eu/cdsapp#!/dataset/cams-solar-radiation-timeseries
[37] E.-S. Ibrahim Abo Elkheir, M. A. Abd-Allah, S. Aly Hussien, and A. I. Omar, "Real-time microgrid fault management: Leveraging hybrid deep learning for superior detection and classification performance," *IEEE Access*, vol. 14, pp. 9606–9623, 2026.

SYED ZULQADAR HASSAN (Member, IEEE) received the Ph.D. degree in electrical engineering from Chongqing University, China, in 2018.

He is currently an Associate Professor with the University of Central Punjab, Pakistan. His research interests include modeling, control, design, techno-economic analysis, and the energy management of alternative energy systems and microgrids, with a particular focus on demand response, load management, smart grids, and microgrid applications. He has published more than 80 research outputs, including journal articles, book chapters, conference papers, and workshop contributions. His work has appeared in reputable international venues, including IEEE and Springer Nature publications, and covers fuzzy logic, artificial intelligence, neural networks, and adaptive control for renewable energy and energy storage systems, including solar PV, wind, fuel cells, hydropower, electric vehicle charging, hydrogen systems, supercapacitors, and batteries.

Dr. Hassan also serves as a Reviewer for several leading international journals, including IEEE Transactions on Energy Conversion and Applied Energy.

BABAR SANA received the B.S. degree in telecommunications engineering from FAST University Islamabad, in 2011, and the M.S. degree in electrical engineering from the University of Sialkot, in 2024. He is currently a Lecturer in electronics technology with the Government College of Technology Gujranwala. His major research interests include renewable energy resources and artificial intelligence.

TARIQ KAMAL (Member, IEEE) received the dual Ph.D. degree in energy and sustainability engineering from the University of Cádiz, Spain, Technical University of Munich, Germany, and degree in electrical engineering from Sakarya University, Türkiye.

He is currently an Assistant Professor in electrical engineering and the Head of the Power Electronics and Control Research Group, School of Technology and Innovations, University of Vaasa, Finland. He is also an Experienced Researcher, a Technical Coordinator, and the Project Leader in power electronics, intelligent control, smart grids, renewable energy integration, hybrid energy storage systems, electric vehicles, maritime electrification, and AI-enabled energy systems. He has contributed to and led several major European and international research projects, including Horizon Europe initiatives, with cumulative project funding of approximately €20 million from European Commission, DAAD Germany, TÜBİTAK Türkiye, and Spanish Ministry of Science and Innovation. His work focuses on real-time control, advanced power conversion, flexibility markets, distributed energy resources, peer-to-peer energy trading, the digitalization of energy systems, and AI-based optimization for sustainable energy transition. He has authored or co-authored more than 100 peer-reviewed publications in international journals, conferences, books, and book chapters, and regularly reviews for leading journals, including IEEE Transactions on Industrial Electronics, IEEE Transactions on Smart Grid, and *Applied Energy*. His international collaborations span Europe, North America, Asia, and Australia, involving universities, research organizations, and industry partners. His research interests include intelligent electrification, clean energy transition, and the development of reliable, efficient, and future-ready power and energy systems.

ARSALAN MASOOD received the M.S. degree in electrical engineering from COMSATS University Islamabad and the Ph.D. degree in electrical engineering from North China Electric Power University.

He is currently affiliated with the University of Sialkot. He has authored and co-authored several peer-reviewed publications in international journals and conferences in the areas of smart grids, energy optimization, and low-carbon power systems. His research interests include distributed energy resource integration, electric vehicles in power systems, demand-side flexibility, transactive energy systems, electricity market modeling, and the analytical and optimization-based coordination of flexible energy resources in modern power systems.

MUHAMMAD AHMAD (Member, IEEE) was born in Pakistan, in 1991. He received the B.Sc. degree in electrical engineering with a specialization in power system engineering from the Swedish College of Engineering and Technology (SCET), Rahim Yar Khan, Pakistan, and the M.Sc. and Ph.D. degrees in electrical engineering, focusing on electrical drive systems and fault-current limiters (FCL) for VSC-HVDC transmission systems from Shanghai Jiao Tong University (SJTU), China.

He received scholarships from SJTU for his M.Sc. and Ph.D. degrees. He is currently a Postdoctoral Research Fellow with the Interdisciplinary Research Center for Sustainable Energy Systems, King Fahd University of Petroleum and Minerals, Saudi Arabia. He has published several high-impact papers in prestigious journals and conferences. His research interests include, but not limited to, power electronics, motor drives, FCLs, hybrid circuit breakers for VSC-HVDC, renewable energy assessment, policies, and power markets.

MD. SHAFIULLAH (Senior Member, IEEE) received the B.Sc. and M.Sc. degrees in electrical and electronic engineering from Bangladesh University of Engineering and Technology (BUET), Bangladesh, in 2009 and 2013, respectively, and the Ph.D. degree in electrical power and energy systems from the King Fahd University of Petroleum and Minerals (KFUPM), Saudi Arabia, in 2018.

He was a Faculty Member of the Department of Electrical and Electronic Engineering, International Islamic University of Chittagong (IIUC), Bangladesh, from 2009 to 2013. He is currently an Assistant Professor with the Control and Instrumentation Department and a Research Affiliate with the Interdisciplinary Research Center for Sustainable Energy Systems (IRC-SES), KFUPM. His research interests include grid fault diagnosis, the grid integration of renewable energy resources, power quality analysis, power system control and stability, evolutionary algorithms, and machine learning techniques.