Advances in Sustainable Energies & Environment

Journal homepage: https://A-SEE.MazUST.ac.ir/

Research Article

Integrated phaselet-neural network-based approach to fault detection & localization in photovoltaic arrays

Erfan Dolati, Navid Ghaffarzadeh *

Department of electrical engineering, Imam Khomeini international university, Qazvin, 3414896818, Iran

ARTICLE INFO

Article history:
Received: 2026-06-07
Revised: 2026-06-29
Accepted: 2026-06-29
Published: 2026-07-01

Keywords:
Fault detection;
Localization;
PV systems;
Phaselet;
Neural network.
.

ABSTRACT

To ensure reliable operation and prevent energy loss, equipment damage, or fire hazards in PV systems, timely and accurate fault detection and localization are essential. Conventional methods relying solely on total current or power often fail to distinguish faults that produce similar power drops — for instance, a line-to-ground fault on the first panel of different substrings may result in identical total power reduction. To overcome this limitation, this paper proposes a high-accuracy fault diagnosis method based on the Phaselet Transform of the instantaneous DC power signal and individual substring currents. A feedforward neural network is trained on several distinct scenarios (including healthy operation and various line-to-line and line-to-ground faults) and achieves an outstanding accuracy of 99.85% on the test dataset. For rare ambiguous cases where the network output does not uniquely identify a single fault location, a thermal imaging camera is suggested as a cost-effective secondary verification tool.

1. Introduction

Photovoltaic (PV) systems offer numerous advantages over conventional power generation technologies. They produce clean energy with zero greenhouse gas emissions, have no fuel costs, require minimal maintenance due to the absence of moving parts, and can be installed at various scales from small rooftop systems to large utility-scale power plants. Moreover, PV systems reduce geopolitical dependency on fossil fuel-rich regions, thereby enhancing national energy security.

In contrast, fossil fuel power plants suffer from severe drawbacks. They are major sources of carbon dioxide and other greenhouse gases, contributing significantly to global warming and climate change. They also release harmful pollutants such as sulfur dioxide, nitrogen oxides, and particulate matter, which cause acid rain, smog, and serious respiratory diseases. Furthermore, fossil fuel power plants consume enormous amounts of water for cooling, pollute water resources through leaks and ash disposal, and their operation depends on finite and depleting resources with volatile prices.

Due to these environmental, economic, and geopolitical factors, PV systems are rapidly replacing fossil fuel power plants worldwide. The installed capacity of PV systems has grown exponentially over the past decade, and this trend is expected to continue in the coming years.

However, PV panels are exposed to harsh environmental conditions such as dust, temperature variations, humidity, hail, and mechanical stresses. These factors can lead to various faults, including line-to-line and line-to-ground faults. Such faults cause a reduction in output power and energy efficiency. If these faults are not detected, localized, and repaired in time, they lead to decreased system reliability, significant power loss, and accelerated aging and degradation of PV panels, ultimately reducing their useful lifetime.

To address these challenges, various fault detection and localization methods have been proposed in the literature. Several studies have explored these techniques using imaging and signal-based approaches. One approach combines thermal and visual data imagery using a Canny edge detector, Gaussian filter, histogram equalization, and seed pixels, which enables

*** Corresponding author:**
E-mail address: *ghaffarzadeh@eng.ikiu.ac.ir*

Cite this article as:
Dolati, E., & Ghaffarzadeh N. (2026). Integrated phaselet-neural network-based approach to fault detection & localization in photovoltaic arrays. *Advances in Sustainable Energies & Environment 1 (3) 41 – 51.* https://doi.org/10.22034/a-see.2026.2090932.1026

real-time monitoring capabilities and can detect various types of faults; however, it does not specifically address dust-related issues [1]. Another study proposed surface defect detection using Alex Net with a convolutional neural network (CNN) based on online visual images, introducing an innovative concept, but it relies on manual feature extraction during the detection stage, and relying on online data collection may limit the capacity of the classification model [2]. Dust and soil detection using a gray-level co-occurrence matrix (GLCM) with an RGB camera is relatively straightforward to implement, yet it does not account for other classes of anomalies such as shadow areas, broken panels, or wet panels [3].

Further research has applied segmentation techniques for PV solar panels, including filtering by area and the active contours level-set method (ACM LS), refined using morphological operations and the Hough transform, achieving a Dice score of 0.97 and an IoU of 0.94, though small statistical differences exist between the two segmentation methods and the analysis is limited to photovoltaic system analysis by cell [4]. An autonomous drone-based solution employing RGB and thermal cameras with an automatic flight path planning algorithm and image processing algorithm enables automatic detection and localization of faulty PV modules and precise location estimation, but it requires testing and validation for different power plant locations and environmental conditions [5]. Thermographic non-destructive tests (TNDTs) combined with a convolutional neural network (CNN) for image classification have achieved 98% accuracy in tests for automatic classification of thermographic images; nevertheless, this method is limited to classification of dust and hotspots and requires further evaluation for other faults [6]. An image processing technique using maximally stable extremal regions (MSER) and a homography translation technique applied to thermal and visible images of PV modules provides automatic identification of hotspots with 97% consistency with visual evaluation, although it is limited to the identification of hotspots and related abnormalities [7]. Deep learning approaches such as Mask R-CNN, UNet, FPNet, and LinkNet applied to UAV-based thermal imaging have achieved values of 0.741 and 0.841 on the Jaccard and Dice indices, but face challenges in binary segmentation, multi-class segmentation, dataset improvement, and real-time measurements [8]. A deep learning unified model using various modalities (thermal to visual images) for various energy installations has demonstrated an accuracy of 0.79 in surface damage detection, yet it remains limited to surface damage detection and specific energy installations [9]. The use of drones, thermal images, and MATLAB image analysis including image acquisition, grayscale conversion, filtering, and 3D image construction for solar modules installed on buildings has improved inspection efficiency and enhanced defect diagnosis capability, but it is limited to thermal image-based defects and has potential dependency on image quality [10]. A method incorporating Ghost Convolution, BottleneckCSP, a tiny target prediction head, YOLOv5, Feature Pyramid Network (FPN), and Path Aggregation Network (PAN) applied to PV panel surface images has improved accuracy in tiny defect detection and enhanced model inference speed, though it is limited to PV panel surface defect detection with potential scalability limitations [11].

Other techniques include electroluminescence imaging using a high-resolution infrared camera, which can locate faults but struggles with identification and requires expensive equipment [12,13]. Similarly, thermal imaging with a thermal camera also can locate faults but struggles with identification and requires expensive equipment [14–16]. Ultrasonic methods using ultrasonic equipment likewise can locate faults but struggle with identification and require expensive equipment [17]. An energy loss detection method employing ambient temperature and irradiance sensors can carry out continuous fault diagnosis and identification while the PV array is working; however, diagnosis accuracy depends on the accuracy of the simulation model and the method cannot locate the fault [18–21]. I-V curve based diagnosis using I-V curve measurement equipment accurately diagnoses faults by measuring output characteristics of PV arrays but lacks localization capability [22–24]. A sequential voltage and current detection method using voltage and current sensors for online measurement can diagnose and identify faults while the PV array is working without needing to measure irradiance and temperature, yet it cannot locate the fault [25–28]. A power switch and diode-based method with voltmeters offers simplicity, accuracy, and module-level fault localization, but it is unfeasible for large-scale systems due to the high cost of switches, wiring, maintenance, operation, and voltage drop issues [29,30]. Virtual imaging using infrared thermography provides high accuracy but requires precise and costly cameras, and continuous monitoring is essential [31–33]. Model-based methods are interpretable and effective under high-irradiance conditions, but their accuracy is sensitive to environmental variations, depends on precise system modeling, and is challenging in practice [34,35]. Data-driven methods using machine learning or artificial neural networks offer higher diagnostic accuracy and adaptability to varying operating conditions, yet they require significant computational resources and long training times [36–39].

Additional protection and inspection methods present further limitations. Undervoltage or overcurrent protection lacks the required sensitivity for detecting high-resistance faults or is unreliable due to communication delay and failure [40,41]. The rate of change in current or voltage protection similarly lacks sensitivity for detecting high-resistance faults or is unreliable to communication delay and failure [42]. Directional current-based protection is unreliable when communication fails, and operation of high-bandwidth fiber optic communication is expensive and not justifiable for low-voltage DC grids [43]. Manual field inspection is costly, time-consuming, subjective, and limited by the expertise of the inspector [44]. Finally, thermal imaging and electromagnetic imaging (electroluminescence or photoluminescence) are costly or require unusual equipment [45,46].

Conventional methods mainly rely on monitoring the total output current or voltage. While simple and low-cost, they often fail to distinguish faults that produce similar drops in total power. For example, a line-to-ground fault occurring on the first panel of different substrings may result in nearly identical total power reduction. This ambiguity leads to delayed or incorrect fault localization, directly reducing system reliability and accelerating panel degradation.

Compared with image-based inspection methods, the proposed signal-processing approach does not require expensive cameras, continuous visual monitoring, or external environmental visibility. In addition, unlike conventional power- or current-only methods, the integration of Phaselet-based features from both total power and substring currents significantly reduces ambiguity by providing complementary information from different electrical measurements. Compared with deep learning-based approaches, the proposed method also offers a simpler and more interpretable solution, as it combines Phaselet-based feature engineering with a lightweight feedforward neural network. This reduces the dependence on large training datasets and high computational resources, while still providing strong fault discrimination and localization capability.

To overcome the limitations of existing methods, this paper proposes a novel fault detection and localization technique based on the Phaselet Transform of both the instantaneous DC power signal and individual substring currents. The proposed method extracts Phaselet coefficients from the total power signal (eight coefficients: four sine and four cosine) and from each of the three substring currents (eight coefficients per substring, totaling 24 coefficients).

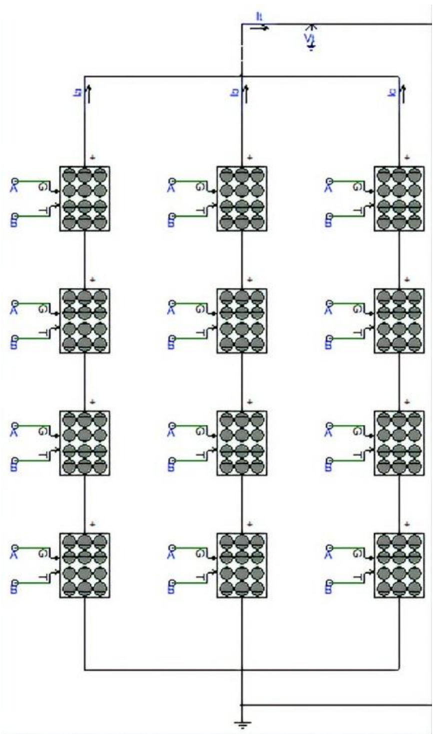

Fig. 1.PV panel configuration

This results in a rich 32-dimensional feature vector per time sample. Adding substring current features significantly reduces ambiguities, enabling faster and more accurate fault detection and localization. Consequently, system reliability is improved, and the useful lifetime of PV panels is extended. A feedforward neural network is trained on several distinct scenarios (including healthy operation and various line-to-line and line-to-ground faults) and achieves 99.84% accuracy on the test dataset. For rare ambiguous cases where the network output does not uniquely identify a single fault location, a thermal imaging camera is suggested as a cost-effective secondary verification tool. The remainder of this paper is organized as follows. Section 2 describes the proposed method, including PV system simulation in PSCAD, Phaselet feature extraction in MATLAB, and the neural network architecture. Section 3 presents simulation results and discussion. Finally, Section 4 concludes the paper.

2. Materials and methods

2.1. System Simulation in PSCAD

The proposed PV system consists of 12 panels arranged in three substrings, with each substring containing 4 panels connected in series. The total simulation time is 1 second, the fault occurs at t = 0.5 s, and the system provides 1112 samples per second for each measured parameter. All three substrings are connected in parallel to form the DC bus. For a general PV array with N_s panels in series and N_p strings in parallel, the total voltage and current are given by:

$$V_{total} = N_s \times V_{panel} \quad (1)$$

$$I_{total} = N_p \times I_{string} \quad (2)$$

$$P_{total} = V_{total} \times I_{string} = (N_s \times V_{panel}) \times (N_p \times I_{string}) \quad (3)$$

where V_{panel} and I_{string} are the output voltage of a single panel and the output current of a single string, respectively. In this work, $N_s = 4$ (four panels per substring) and $N_p = 3$ (three substrings in parallel). The parameters in Table 1 are defined in accordance with the panel configuration previously illustrated in Fig. 1.

The system is simulated in the PSCAD/EMTDC environment. Various fault scenarios, including line-to-line faults and line-to-ground faults, are implemented at different locations. For each scenario, the instantaneous DC values of total power, total voltage, total current, and the current of each individual substring are recorded for further analysis.

Table 1. Parameter specifications

Parameter	Description	Value in this work
N_s	Number of panels connected in series in each substring	4
N_p	Number of parallel substrings (strings)	3
V_{panel}	Output voltage of a single panel (V)	Depends on operating point
I_{string}	Output current of one substring (A)	Depends on operating point
V_{total}	Total DC voltage of the entire PV array	$V_{total} = N_s \times V_{panel}$
I_{total}	Total DC current of the entire PV array	$I_{total} = N_p \times I_{string}$
P_{total}	Total DC power of the entire PV array	$P_{total} = V_{total} \times I_{string}$

The equivalent circuit of a single PV cell is described by the well-known single-diode model. The current–voltage characteristic of a PV cell is given by:

$$I = I_{ph} - I_0(\exp(\frac{V + IR_S}{nV_t}) - 1) - \frac{V + IR_s}{R_{sh}} \tag{4}$$

The parameters of the single-diode PV cell model equation are described in Table 2. Here, each substring consists of 4 series-connected panels, and the three substrings are connected in parallel.

Therefore, $N_s = 4$ and $N_p = 3$. The values of V_{panel} and I_{string} vary depending on the type and location of the fault.

For a PV panel consisting of N_s cells in series, the output voltage is multiplied accordingly. In this work, each panel has $N_s = 4$ cells (for simplicity, without loss of generality). The total DC power of the entire array is calculated as:

$$P_{total}(t) = V_{total}(t) \times I_{total}(t) \tag{5}$$
$$V_a = V_{panel1} + V_{panel2} + V_{panel3} + V_{panel4} \tag{6}$$
$$V_b = V_{panel5} + V_{panel6} + V_{panel7} + V_{panel8} \tag{7}$$
$$V_c = V_{panel9} + V_{panel10} + V_{panel11} + V_{panel12} \tag{8}$$
$$V_{total} = V_a = V_b = V_c \tag{9}$$
$$I_a = I_{panel1} = I_{panel2} = I_{panel3} = I_{panel4} \tag{10}$$
$$I_b = I_{panel5} = I_{panel6} = I_{panel7} = I_{panel8} \tag{11}$$
$$I_c = I_{panel9} = I_{panel10} = I_{panel11} = I_{panel12} \tag{12}$$
$$I_{total} = I_a + I_b + I_c \tag{13}$$

In a series-parallel PV array, the total voltage V_{total} is forced to be equal across all parallel substrings. Consequently, when a fault occurs, the faulty substring's internal impedance changes, but the healthy substrings try to maintain the common bus voltage.

As a result, currents I_a, I_b, and I_c become different for each substring, while V_{total} alone does not reveal the fault location. Thus, substring currents provide the discriminative information needed to detect and localize faults, even when the total voltage remains common.

where $V_{total}(t)$ and $I_{total}(t)$ are the instantaneous total DC voltage and current of the whole PV array, respectively. In addition, the currents of the three individual substrings are denoted as $I_{sub1}(t)$, $I_{sub2}(t)$ and $I_{sub3}(t)$. All signals are sampled and stored for subsequent feature extraction.

2.2. Phaselet Feature Extraction

Phaselet transform was selected for feature extraction due to the non-stationary nature of the monitored signal and the transient behavior of fault-induced disturbances. Its inherent time-frequency localization enables the capture of abrupt local variations immediately after fault inception, without requiring prior knowledge of the exact fault time. Consequently, the extracted features exhibit high sensitivity to fault occurrence, making the proposed approach well suited for online fault detection and localization.

Table 2. Parameters of the single-diode PV cell model

Parameter	Description
I	Output current of the cell (A)
V	Output voltage of the cell (V)
I_{ph}	Photo-generated current (proportional to irradiance)
I_0	Reverse saturation current of the diode
R_S	Series resistance (ohm)
R_{sh}	Shunt resistance (ohm)
n	Diode ideality factor
V_t	Thermal voltage $V_t = kT/q$ (V)
k	Boltmann constant (1.38×10^{-23} J/K)
T	Cell temperature (K)
q	Electron charge (1.6×10^{-19}C)

To capture time-frequency characteristics of the fault transients, a Phaselet Transform is applied to the recorded signals. For a discrete-time signal $x_c[i]$ (where i=0, 1, …, N−1), the sine and cosine coefficients at harmonic index m and reference time k are computed as:

$$S_{c,m}[k] = \sum_{i=0}^{N-1} x_c[i] \times \sin(\frac{2\pi m(k-i)}{N}) \tag{14}$$

$$C_{c,m}[k] = \sum_{i=0}^{N-1} x_c[i] \times \cos(\frac{2\pi m(k-i)}{N}) \tag{15}$$

Table 3 describes the parameters of the Phaselet transform equations.

In this work, four different signals are processed:

- Total DC power of the entire PV array, denoted as:

$$x_1[i] = P_{total}[i] \tag{16}$$

- Current of substring 1, denoted as:

$$x_2[i] = I_{sub1}[i] \tag{17}$$

- Current of substring 2, denoted as:

$$x_3[i] = I_{sub2}[i] \tag{18}$$

- Current of substring 3, denoted as:

$$x_4[i] = I_{sub3}[i] \tag{19}$$

For each of these four signals, four sine coefficients and four cosine coefficients (corresponding to harmonics m=1,2, 3,4) are computed for every time index k. Therefore, the total number of features per time sample is:

$$4\,signals \times (4\,\text{sine} + 4\,cosine) = 32\,feature \tag{20}$$

Table 3. Parameters of the discrete Phaselet transform used in this work

Parameter	Description
$x_c[i]$	Input signal (total power or substring current) at time instant i
N	Window length = 1112
m	Harmonic index (m=1, 2, 3, 4 for the first four harmonic)
k	Reference time index (k=0,1,…,N-1)
$S_{c,m}[k]$	Sine coefficient for channel c, harmonic m, at time k
$C_{c,m}[k]$	Cosine coefficient for channel c, harmonic m, at time k

The resulting 32-dimensional features vector for a single time sample is:

$$
\begin{aligned}
f[k] \\
= [&S_{1,1}[k], S_{1,2}[k], S_{1,3}[k], S_{1,4}[k], C_{1,1}[k], C_{1,2}[k], C_{1,3}[k], C_{1,4}[k], \\
&S_{2,1}[k], S_{2,2}[k], S_{2,3}[k], S_{2,4}[k], C_{2,1}[k], C_{2,2}[k], C_{2,3}[k], C_{2,4}[k], \\
&S_{3,1}[k], S_{3,2}[k], S_{3,3}[k], S_{3,4}[k], C_{3,1}[k], C_{3,2}[k], C_{3,3}[k], C_{3,4}[k], \\
&S_{4,1}[k], S_{4,2}[k], S_{4,3}[k], S_{4,4}[k], C_{4,1}[k], C_{4,2}[k], C_{4,3}[k], C_{4,4}[k]\ ]
\end{aligned} \tag{21}
$$

where the first subscript (1 to 4) indicates the signal channel (1 = total power, 2–4 = substring currents). All feature vectors from all scenarios and all time instants are collected to form the final dataset.

2.3. Selection of Distinct Scenarios

Initially, a total of 49 scenarios were simulated: one healthy case and 48 fault cases (line-to-line and line-to-ground faults at different locations). However, after extracting the 32-dimensional Phaselet features, it was observed that some faults produced identical or nearly identical feature vectors. This indicates that the extracted Phaselet features are highly sensitive to fault location, but in a few cases, the overlap between certain scenarios makes unique separation impractical. These redundant scenarios were removed from the training set to avoid unnecessary complexity.

Therefore, such scenarios were merged into a single class, resulting in 31 distinct classes in the final dataset.

The final training set consists of 31 distinct scenarios: one healthy class and 30 unique fault classes (out of the original 48).

This selection significantly reduces the number of training samples while preserving all distinguishable fault patterns. As a result, the neural network is trained only on distinct feature vectors, which improves convergence and classification accuracy.

2.4. Neural Network Architecture

In this study, a Feedforward Neural Network (FNN) is employed for multi-class classification. The network architecture consists of an input layer with 32 neurons, two hidden layers with 350 and 180 neurons respectively (both using tansig activation), and an output layer with 31 neurons (softmax activation). The network is trained using the Scaled Conjugate Gradient (trainscg) algorithm with a cross-entropy loss function. The dataset is divided into 70% training, 15% validation, and 15% testing. To ensure a balanced evaluation, the samples were split within each class into training, validation, and testing subsets. Redundant scenarios with nearly identical Phaselet feature vectors were merged before the split, reducing similarity between the subsets and improving the reliability of the reported accuracy.

After Z-score normalisation, the network achieved a near-perfect test accuracy of 99.85%. The detailed parameters are summarized in Table 1.

Table 4. Neural Network Specifications

Parameter	Value
Network Type	Feedforward Neural Network (FNN)
Number of Layers	4 (Input + 2 Hidden + Output)
Number of Hidden Layers	2
Input Layer Neurons	32
Hidden Layer 1 Neurons	350
Hidden Layer 2 Neurons	180
Output Layer Neurons	31
Hidden Layers Activation Function	tansig
Output Layer Activation Function	softmax
Optimisation Algorithm	Scaled Conjugate Gradient (trainscg)
Loss Function	Cross-entropy
Number of Epochs	2400
Early Stopping (Max Fail)	480
Learning Rate	0.008
Regularisation	0
Normalisation Method	Z-score
Training	70%
Validation	15%
Test Split	15%
Total Samples	34,472
Features	32
Classes	31 of 49
Training Time	18.3 minutes

Activation for hidden layer 1:

$$tanh(z) = \frac{e^{z_1} - e^{-z_1}}{e^{z_1} + e^{-z_1}} \tag{22}$$

Activation for hidden layer 2:

$$tanh(z) = \frac{e^{z_2} - e^{-z_2}}{e^{z_2} + e^{-z_2}} \tag{23}$$

Activation for output layer :

$$softmax(z_m) = \frac{e^{z_m}}{\sum_k e^{z_k}} \tag{24}$$

Formula for hidden layer 1:

$$h^{(1)} = \tanh(W^{(1)}x + b^{(1)}) \tag{25}$$

Formula for hidden layer 2:

$$h^{(2)} = \tanh(W^{(2)}h^{(1)} + b^{(2)}) \tag{26}$$

Formula for output layer :

$$y = softmax(W^{(out)}h^{(2)} + b^{(out)}) \tag{27}$$

$$h^{(1)} = \frac{e^{W^{(1)}x+b^{(1)}} - e^{-(W^{(1)}x+b^{(1)})}}{e^{W^{(1)}x+b^{(1)}} + e^{-(W^{(1)}x+b^{(1)})}} \tag{28}$$

$$h^{(2)} = \frac{e^{W^{(2)}h^{(1)}+b^{(2)}} - e^{-(W^{(2)}h^{(1)}+b^{(2)})}}{e^{W^{(2)}h^{(1)}+b^{(2)}} + e^{-(W^{(2)}h^{(1)}+b^{(2)})}} \tag{29}$$

$$y = \frac{e^{W^{(out)}h^{(2)}+b^{(out)}}}{\sum_{k=1}^{31} e^{z_k}} \tag{30}$$

Table 5 presents the proposed neural network equations in a step-by-step manner. Table 6 describes the parameters of the proposed neural network equations.

Table 5. Neural Network Mathematical Formulations

Layer	Activation Function	Formula	Output Range
Hidden Layer 1 (350 neurons)	$\tanh(z) = \frac{e^{z_1}-e^{-z_1}}{e^{z_1}+e^{-z_1}}$	$h^{(1)} = \tanh(W^{(1)}x + b^{(1)})$	[-1,1]
Hidden Layer 2 (180 neurons)	$\tanh(z) = \frac{e^{z_2}-e^{-z_2}}{e^{z_2}+e^{-z_2}}$	$h^{(2)} = \tanh(W^{(2)}h^{(1)} + b^{(2)})$	[-1,1]
Output Layer (31 neurons)	$softmax(z_m) = \frac{e^{z_m}}{\sum_k e^{z_k}}$	$y = softmax(W^{(out)}h^{(2)} + b^{(out)})$	[0,1]

Table 6. Parameter specifications of the neural network mathematical formulations

Symbol	Description
x	Input feature vector
$\mathbf{h}^{(1)}$	Output vector of first hidden layer
$\mathbf{h}^{(2)}$	Output vector of second hidden layer
y	Output probability vector (softmax output)
$\mathbf{W}^{(1)}$	Weight matrix (input → hidden layer 1)
$\mathbf{W}^{(2)}$	Weight matrix (hidden layer 1 → hidden layer 2)
$\mathbf{W}^{(out)}$	Weight matrix (hidden layer 2 → output)
$\mathbf{b}^{(1)}$	Bias vector of first hidden layer
$\mathbf{b}^{(2)}$	Bias vector of second hidden layer
$\mathbf{b}^{(out)}$	Bias vector of output layer
e	Euler's number ≈ 2.71828

2.5. Ambiguity Resolution with Thermal Imaging

During the testing phase, if the network outputs a class label that does not uniquely identify a single fault location (i.e., the same feature vector may correspond to two different physical fault positions), a thermal imaging camera is used as a secondary verification tool. The operator first inspects the most probable fault location; if no fault is found, the other possible location is checked. This hybrid approach (neural network + thermal imaging) enables faster fault detection and localization while maintaining high reliability.

The key advantage of the proposed method is that it significantly reduces the time required to identify the type and location of a fault compared to conventional methods that rely solely on post-event inspection or thermal scanning of the entire array.

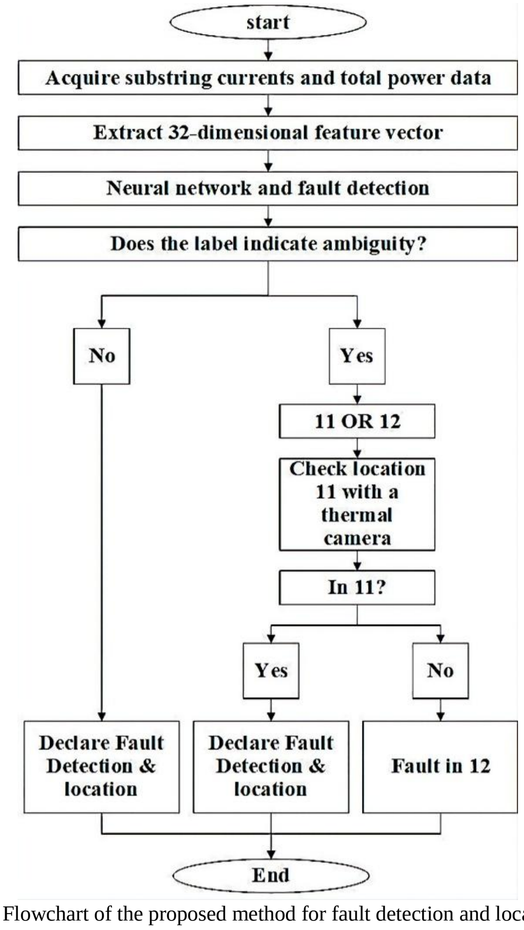

Fig. 2. Flowchart of the proposed method for fault detection and localization

3. Results

In this section, the performance of the proposed fault detection and localization method is evaluated. A total of 49 scenarios were initially simulated: one healthy case and 48 fault cases (line-to-line and line-to-ground faults at different locations). After extracting the 32-dimensional Phaselet features, some faults produced identical feature vectors and were therefore merged. As a result, the final dataset consists of 31 distinct classes: one healthy class and 30 unique fault classes (out of the original 48). The neural network was trained and tested using this 31-classes dataset, as described in Section 2. The results are presented in terms of the confusion matrix, performance metrics, and comparison with an approach that uses only total power (without substring currents).

3.1. Classification Performance

After training, the network was tested on the 15% unseen test dataset. The overall confusion matrix is shown in Fig. 3.

Fig. 3 shows the confusion matrix of the proposed neural network for the 31 classes (one healthy + 30 unique faults). Diagonal entries represent correctly classified samples, and only eight off-diagonal misclassifications are observed (each from a different class). The network correctly classifies 5,169 out of 5,177 test samples, achieving a near-perfect accuracy of 99.85%. Table 7 presents the final neural network results for the proposed method in full.

In the unlikely event that the network output corresponds to a class label that does not uniquely identify a single fault location (i.e., the same feature vector may arise from two different physical fault positions), a thermal imaging camera is used as a secondary verification tool. The operator first inspects the most probable location; if no fault is found, the other possible location is examined. This hybrid approach resolves the ambiguity in minimal additional time, preserving the overall fast fault diagnosis capability of the proposed method.

3.2. Comparison with Total-Power-Only Approach

If only the total power signal were used (8 Phaselet features), the number of distinguishable fault classes would decrease dramatically because many different fault locations produce nearly identical total power drops (e.g., a single-phase-to-ground fault on the first panel of different substrings). Consequently, fault detection and localization would be delayed due to increased ambiguity, as the neural network would not be able to uniquely identify the fault location. Resolving such ambiguities would require time-consuming post-inspection (e.g., thermal imaging of multiple possible locations), defeating the purpose of fast automated fault diagnosis.

Table 7. Neural Network Results

Metric	Value (%)
Overall Accuracy	99.8455
Macro-Average Precision	~99.85
Macro-Average Recall	~99.85
Macro-Average F1-Score	~99.85
Classes with 100% Accuracy	23 out of 31
Total test samples	5,177
Total misclassified samples	8 of 5,177

To quantify this, we trained the same neural network using only the 8 total-power Phaselet features. The resulting accuracy was only 14.58%, and the network could not reliably distinguish between fault classes that had identical power drops. In contrast, the proposed method (32 features) achieves 99.85% accuracy and uniquely identifies all 31 classes, as shown in Fig. 4 (The class that is correctly predicted is the healthy state). Table 8 presents the comprehensive final results of the neural network when only the total output power of the PV panels is considered.

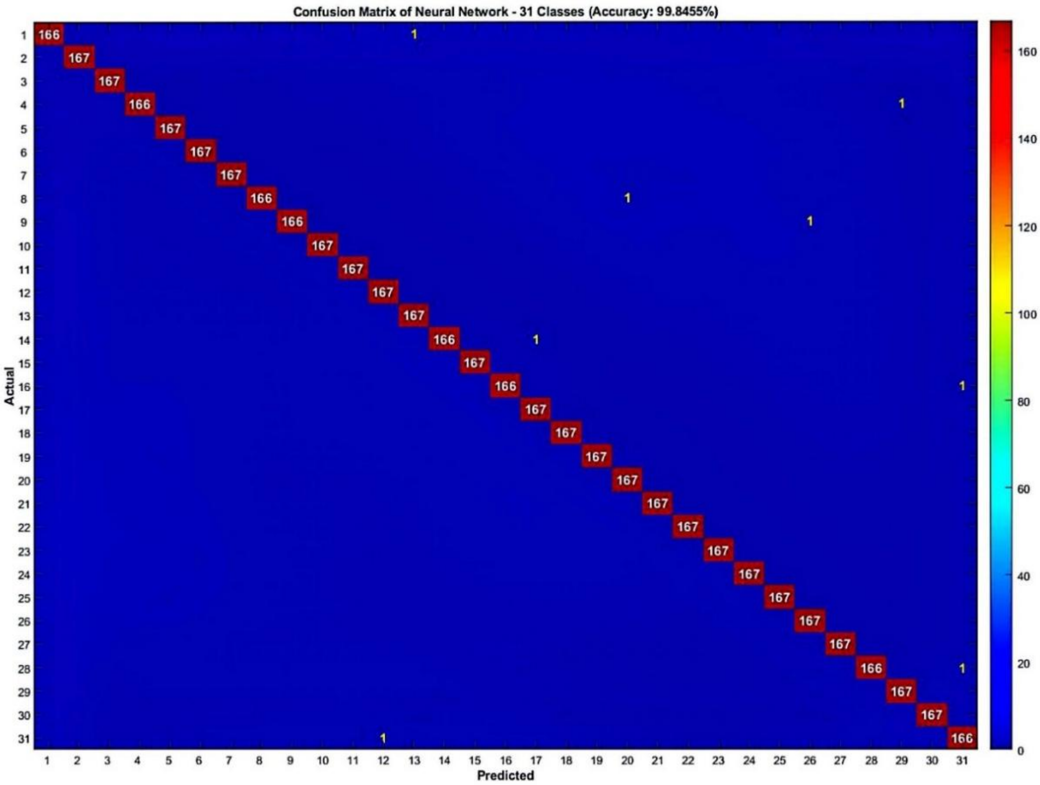

Fig. 3. Confusion matrix of the proposed neural network output for classification of the 31 classes based on Phaselet features extracted from total power and substring currents.

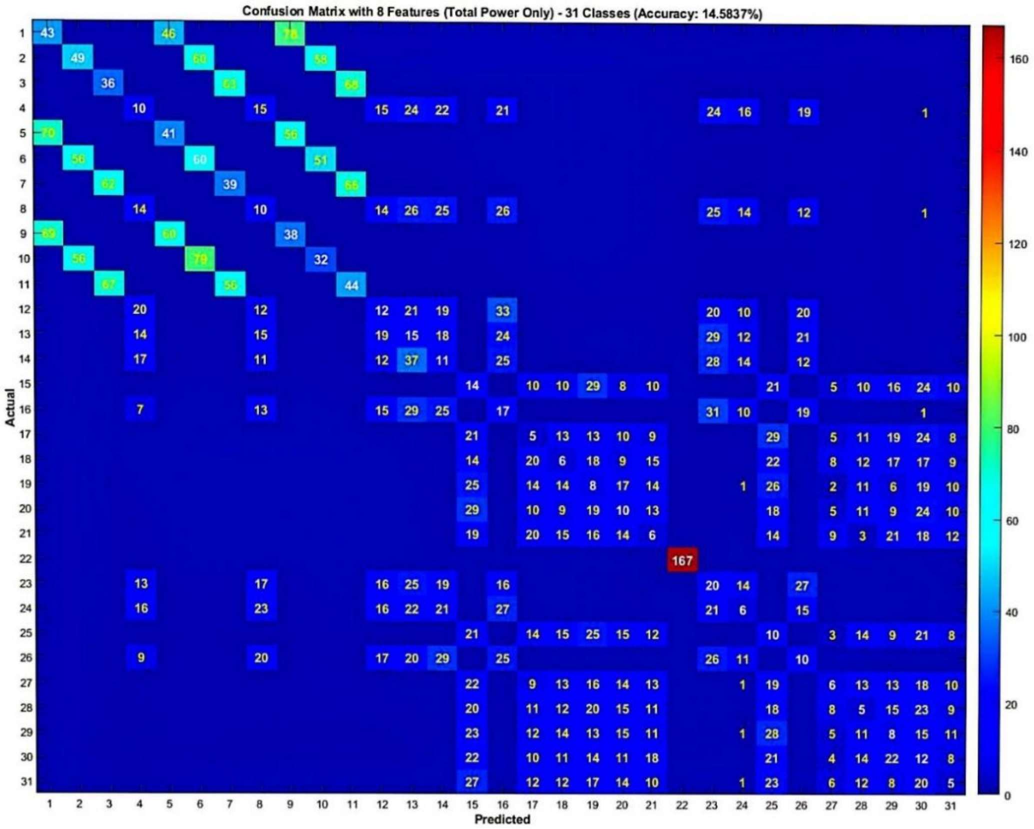

Fig. 4. Confusion matrix of the total-power-only network output for the 31-class test set based on Phaselet features extracted from the total power signal.

Table 8. Neural Network Results with just total power

Metric	Value (%)
Overall Accuracy	14.5837
Macro-Average Precision	~14.62
Macro-Average Recall	~14.58
Macro-Average F1-Score	~14.48
Classes with 100% Accuracy	1 out of 31
Total test samples	5,177
Total misclassified samples	4,422 of 5,177

Furthermore, adding the substring current features significantly increases the dimensionality of the feature space (from 8 to 32 features), which provides the neural network with richer discriminative information. This resolves the ambiguity problem and improves overall classification accuracy.

3.3. Why Power Instead of Voltage or Current?

Faults in PV systems cause a decrease in both output voltage and current, leading to a drop in overall system efficiency. While both voltage and current are sensitive to fault occurrence and are readily available in most PV installations, neither alone provides a complete picture. As illustrated in Fig. 5 to 7, during a fault event, the current drops by approximately 21%, the voltage drops by approximately 18%, while the power drops by approximately 36% — nearly the sum of the individual drops. The power signal therefore amplifies the fault signature by capturing the simultaneous effect on both voltage and current, making it more robust against measurement noise and minor environmental variations than either quantity alone. For this reason, the instantaneous DC power is chosen as the primary signal for feature extraction in this work.

As shown Fig. 5 to 7, the proposed method is based on the signal with the highest sensitivity: during a fault, power drops by 34% (from 938 W to 620 W), while current drops by 23% (from 7.3 A to 5.6 A) and voltage drops by 14% (from 128 V to 110 V). Therefore, power is selected as the primary signal for feature extraction.

3.4. Why Substring Currents Instead of Substring Powers and Voltages?

The power of each individual substring could theoretically be used instead of its current. However, computing substring power requires measuring both voltage and current for each substring, doubling the number of sensors (one voltmeter and one ammeter per substring). In contrast, using only the substring currents requires only one ammeter per substring, while the voltage information is already captured by the total power signal (which needs only one voltmeter for the entire array). Moreover, installing voltmeters on each substring is not cost-effective, as voltmeters are generally more expensive than ammeters. Therefore, using substring currents instead of substring powers is the most economical choice. It provides sufficient discriminative information with minimal additional hardware, making the proposed method practical for real-world PV installations.

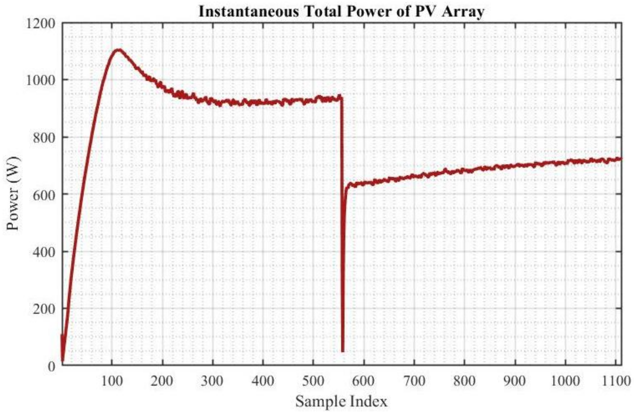

Fig. 5. Total power of the PV array

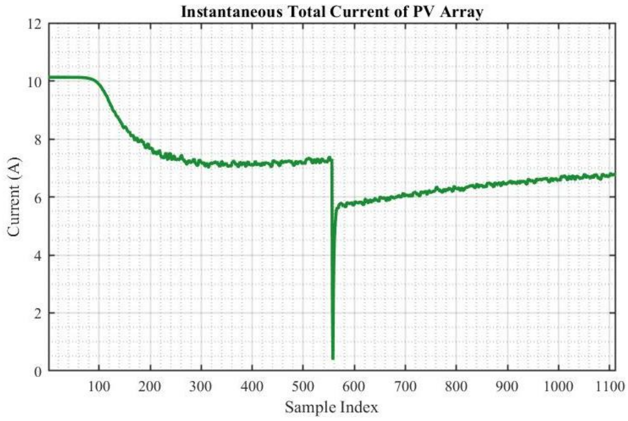

Fig. 6. Total current of the PV array

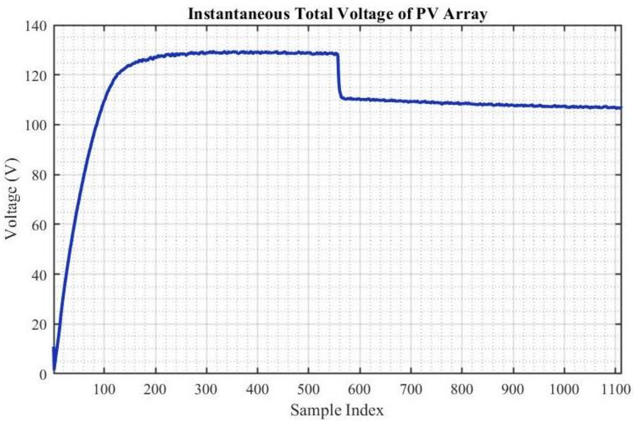

Fig. 7. Total voltage of the PV array

3.5. Discussion

The results presented above confirm the effectiveness of the proposed fault detection and localization method. Several key observations can be highlighted:

First, the choice of the total power signal as the primary feature source is justified by its superior sensitivity (34% drop during a fault) compared to current (23%) or voltage (14%). This makes the method robust against measurement noise and minor environmental variations.

Second, while total power alone fails to uniquely distinguish faults with identical power drops, its classification accuracy on the same 31-class dataset is only 14.58% (as reported in Section 3.2). In contrast, the proposed method — which adds Phaselet features of the three substring currents

— achieves 99.85% accuracy on the same 31 classes. This demonstrates that high sensitivity and high selectivity are not mutually exclusive; they are achieved simultaneously by the proposed 32-feature set.

Third, the use of substring currents instead of substring powers is a deliberate cost-effective choice. It avoids the need for additional voltmeters (which are more expensive than ammeters) while still providing sufficient discriminative information.

Fourth, the overall accuracy of 99.85% on 31 distinct fault classes, with only eight misclassifications out of 5,177 test samples (0.15% error rate), demonstrates the high reliability of the proposed method. The rare ambiguous cases can be resolved quickly using a thermal imaging camera without compromising the fast diagnosis capability. This high classification accuracy demonstrates that the proposed Phaselet-based framework is well suited to the present fault diagnosis problem, achieving 99.85% accuracy across 31 classes with a relatively simple and interpretable feedforward architecture. Compared with thermal- or image-based inspection methods, the proposed approach does not require expensive cameras or continuous visual monitoring. In contrast to model-based methods, it does not depend on precise system modeling or environmental assumptions. Compared with machine-learning-based approaches that require large datasets and higher computational cost, the proposed framework uses a compact Phaselet feature set and a lightweight feedforward network. Moreover, unlike I-V curve-based diagnosis, the proposed method achieves both fault detection and localization by combining total power and substring current information. The results indicate that the proposed method provides an effective balance between diagnostic performance, computational complexity, and model interpretability, which is particularly attractive for practical online fault detection and localization.

Fifth, the proposed Phaselet representation was more effective than conventional time-domain or frequency-domain features for capturing transient fault signatures. Unlike purely time-domain features, it provides localized transient information, and unlike frequency-only features, it preserves the timing of abrupt fault-induced changes. Compared with commonly used wavelet-based features, Phaselet offered a more compact and discriminative representation of the fault transients in this study. This led to a more compact and discriminative feature set and improved classifier responsiveness.

Sixth, although some fault scenarios yielded identical or nearly identical feature vectors, these cases are limited in number and correspond to physically ambiguous operating conditions. In practice, a thermal imaging camera can be used as a secondary verification tool to inspect only the few candidate locations instead of the entire array, thereby preserving the speed of fault diagnosis while resolving residual ambiguity.

Finally, the proposed method is generic and scalable. It does not depend on the specific PV array size (number of substrings) or the exact fault locations used in this study. The same feature extraction and neural network framework can be applied to larger systems by adding more substring current channels. The scalability of the proposed method is supported by the fact that it relies on measurable power and substring current signals, which are available in on-grid, off-grid, and hybrid PV systems and can be extended to larger arrays by incorporating additional current channels without changing the core feature-extraction or classification framework.

Although the present study is based on simulation results, the proposed phaselet-based features exhibit strong separability. Therefore, moderate measurement noise and environmental variations may slightly affect the extracted feature values, but they are not expected to significantly degrade the overall classification performance. The proposed method relies on relative feature discrimination rather than exact waveform matching, which supports its robustness under practical operating conditions.

4. Conclusion

This paper proposed a novel fault detection and localization method for photovoltaic (PV) panels based on Phaselet analysis of the total DC power and individual substring currents. The main novelty of this work lies in the simultaneous use of total power Phaselet coefficients (high sensitivity) and substring current Phaselet coefficients (high selectivity). Unlike conventional approaches that rely solely on total current or total power — which cannot distinguish faults with identical power drops — the proposed method extracts a rich 32-dimensional feature set per time sample:

Eight Phaselet coefficients (four sine and four cosine) from the total power signal, and eight coefficients from each of the three substring currents. This combination is the key innovation: total power provides strong fault signatures (34% drop), while substring currents resolve location ambiguities that total power alone cannot.

A total of 49 scenarios (one healthy and 48 faults) were initially simulated in PSCAD/EMTDC. After removing redundant cases with identical feature vectors, the final dataset consisted of 31 distinct classes: one healthy class and 30 unique fault classes. A feedforward neural network with a 32-350-180-31 architecture (two hidden layers) was trained on 34, 472 samples (1,112 per class) and achieved an outstanding 99.85% accuracy on the test dataset. Eight misclassifications were observed out of 5,177 test samples, demonstrateing the excellent discriminative power of the proposed feature set.

The superiority of the power signal over current or voltage alone was demonstrated: power drops by approximately 34% during a fault, compared to 23% for current and 14% for voltage. However, total power alone fails to distinguish faults that produce identical power drops (e.g., a line-to-ground fault on the first panel of different substrings). The key contribution of this work is the addition of Phaselet features of the three substring currents, which resolves this ambiguity completely, enabling unique and highly accurate localization of all 31 classes (99.85% accuracy).

Using substring currents instead of substring powers is a deliberate cost-effective choice. It avoids the need for additional voltmeters (which are more expensive than ammeters) while providing sufficient discriminative information. For the rare ambiguous cases where the network output does not uniquely identify a single fault location, a thermal imaging camera is suggested as a secondary verification tool, resolving the ambiguity in minimal additional time.

The proposed method is generic and scalable. It does not depend on the specific PV array size or the exact fault locations used in this study. The same framework can be applied to larger systems by adding more substring current channels. Future work includes testing the method on a real PV installation, incorporating other fault types (e.g., high-impedance faults, arc faults), and exploring deep learning architectures (e.g., CNNs or LSTMs) for automatic feature extraction from raw signals.

In summary, the key novelty and contributions of this paper are:

1. Phaselet-based feature extraction from both total power and individual substring currents.
2. Resolution of the "identical power drop" ambiguity by adding substring current features.
3. High 99.85% classification accuracy on 31 distinct classes.
4. A cost-effective sensor configuration (only ammeters on substrings, one voltmeter for total voltage).

5. References

[1] M. Alsafasfeh, I. Abdel-Qader, B. Bazuin, Q. Alsafasfeh, W. Su, Unsupervised fault detection and analysis for large photovoltaic systems using drones and machine vision, Energies 11 (2018) 2252, https://doi.org/10.3390/en11092252.

[2] I. Zyout, A. Oatawneh, Detection of PV solar panel surface defects using transfer learning of the deep convolutional neural networks, in: 2020 Advances in Science and Engineering Technology International Conferences (ASET), IEEE, 2020, pp. 1–4, https://doi.org/10.1109/ASET48392.2020.9118249.

[3] K.A. Abuqaaud, A. Ferrah, A novel technique for detecting and monitoring dust and soil on solar photovoltaic panel, in: 2020 Advances in Science and Engineering Technology International Conferences (ASET), IEEE, 2020, pp. 1–6, https://doi.org/10.1109/ASET48392.2020.9118306.

[4] E. Alfaro-Mejía, H. Loaiza-Correa, E. Franco-Mejía, L. Hernández-Callejo, Segmentation of thermography image of solar cells and panels, in: Ibero-American Congress of Smart Cities, Springer, 2019, pp. 1–8, https://doi.org/10.1007/978-3-030-34576-3.

[5] C. Henry, S. Poudel, S.-W. Lee, H. Jeong, Automatic detection system of deteriorated PV modules using drone with thermal camera, Appl. Sci. 10 (2020) 3802, https://doi.org/10.3390/app10113802.

[6] G. Cipriani, A. D'Amico, S. Guarino, D. Manno, M. Traverso, V. Di Dio, Convolutional neural network for dust and hotspot classification in PV modules, Energies 13 (2020) 6357, https://doi.org/10.3390/en13236357.

[7] H. Jeong, G.-R. Kwon, S.-W. Lee, Deterioration diagnosis of solar module using thermal and visible image processing, Energies 13 (2020) 2856, https://doi.org/10.3390/en13112856.

[8] R. Pierdicca, M. Paolanti, A. Felicetti, F. Piccinini, P. Zingaretti, Automatic faults detection of photovoltaic farms: Solair, a deep learning-based system for thermal images, Energies 13 (2020) 6496, https://doi.org/10.3390/en13246496.

[9] A. Shihavuddin, M.R.A. Rashid, M.H. Maruf, M.A. Hasan, M.A. ul Haq, R.H. Ashique, A. Al Mansur, Image based surface damage detection of renewable energy installations using a unified deep learning approach, Energy Rep. 7 (2021) 4566–4576, https://doi.org/10.1016/j.egyr.2021.07.045.

[10] K.-C. Liao, H.-Y. Wu, H.-T. Wen, Using drones for thermal imaging photography and building 3D images to analyze the defects of solar modules, Inventions 7 (2022) 67, https://doi.org/10.3390/inventions7030067.

[11] L. Li, Z. Wang, T. Zhang, Photovoltaic panel defect detection based on ghost convolution with BottleneckCSP and tiny target prediction head incorporating YOLOv5, arXiv preprint, arXiv:2303.00886, 2023.

[12] G. Alves dos Reis Benatto, C. Mantel, S. Spataru, A.A. Santamaria Lancia, N. Riedel, S. Thorsteinsson, P.B. Poulsen, H. Parikh, S. Forchhammer, D. Sera, Drone-based daylight electroluminescence imaging of PV modules, IEEE J. Photovolt. 10 (2020) 872–877, https://doi.org/10.1109/JPHOTOV.2020.2980721.

[13] E. Kaplani, Detection of degradation effects in field-aged c-Si solar cells through IR thermography and digital image processing, Int. J. Photoenergy 2012 (2012) 396792, https://doi.org/10.1155/2012/396792.

[14] O. Breitenstein, J. Bauer, T. Trupke, R.A. Bardos, On the detection of shunts in silicon solar cells by photo- and electroluminescence imaging, Prog. Photovolt. Res. Appl. 16 (2008) 325–330, https://doi.org/10.1002/pip.803.

[15] M. Cubukcu, A. Akanalci, Real-time inspection and determination methods of faults on photovoltaic power systems by thermal imaging in Turkey, Renew. Energy 147 (2020) 1231–1238, https://doi.org/10.1016/j.renene.2019.09.080.

[16] X. Guo, J. Cai, Optical stepped thermography of defects in photovoltaic panels, IEEE Sensors J. 21 (2021) 490–497, https://doi.org/10.1109/JSEN.2020.3014117.

[17] S.K. Chakrapani, M.J. Padiyar, K. Balasubramaniam, Crack detection in full size CZ-silicon wafers using Lamb wave air coupled ultrasonic testing (LAC-UT), J. Nondestruct. Eval. 31 (2011) 46–55, https://doi.org/10.1007/s10921-011-0124-8.

[18] M. Dhimish, V. Holmes, B. Mehrdadi, M. Dales, Simultaneous fault detection algorithm for grid-connected photovoltaic plants, IET Renew. Power Gener. 11 (2017) 1565–1575, https://doi.org/10.1049/iet-rpg.2017.0366.

[19] M. Dhimish, V. Holmes, B. Mehrdadi, M. Dales, P. Mather, Photovoltaic fault detection algorithm based on theoretical curves modelling and fuzzy classification system, Energy 140 (2017) 276–290, https://doi.org/10.1016/j.energy.2017.08.070.

[20] R. Hariharan, M. Chakkarapani, G. Saravana Ilango, C. Nagamani, A method to detect photovoltaic array faults and partial shading in PV systems, IEEE J. Photovolt. 6 (2016) 1278–1285, https://doi.org/10.1109/JPHOTOV.2016.2582918.

[21] F. Harrou, B. Taghezouit, Y. Sun, Improved KNN-based monitoring schemes for detecting faults in PV systems, IEEE J. Photovolt. 9 (2019) 811–821, https://doi.org/10.1109/JPHOTOV.2019.2895901.

[22] Z. Chen, Y. Chen, L. Wu, S. Cheng, P. Lin, Deep residual network based fault detection and diagnosis of photovoltaic arrays using current-voltage curves and ambient conditions, Energy Convers. Manag. 198 (2019) 111793, https://doi.org/10.1016/j.enconman.2019.111793.

[23] J.-M. Huang, R.-J. Wai, W. Gao, Newly-designed fault diagnostic method for solar photovoltaic generation system based on IV-curve measurement, IEEE Access 7 (2019) 70919–70932, https://doi.org/10.1109/ACCESS.2019.2919645.

[24] S. Spataru, D. Sera, T. Kerekes, R. Teodorescu, Diagnostic method for photovoltaic systems based on light I–V measurements, Sol. Energy 119 (2015) 29–44, https://doi.org/10.1016/j.solener.2015.06.022.

[25] B.P. Kumar, G.S. Ilango, M.J.B. Reddy, N. Chilakapati, Online fault detection and diagnosis in photovoltaic systems using wavelet packets, IEEE J. Photovolt. 8 (2018) 257–265, https://doi.org/10.1109/JPHOTOV.2017.2773165.

[26] Z. Yi, A.H. Etemadi, Line-to-line fault detection for photovoltaic arrays based on multiresolution signal decomposition and two-stage support vector machine, IEEE Trans. Ind. Electron. 64 (2017) 8546–8556, https://doi.org/10.1109/TIE.2017.2698362.

[27] Z. Yi, A.H. Etemadi, Fault detection for photovoltaic systems based on multi-resolution signal decomposition and fuzzy inference systems, IEEE Trans. Smart Grid 8 (2017) 1274–1283, https://doi.org/10.1109/TSG.2016.2598215.

[28] S. Liu, L. Dong, X. Liao, Y. Hao, X. Cao, X. Wang, A dilation and erosion-based clustering approach for fault diagnosis of photovoltaic

arrays, IEEE Sensors J. 19 (2019) 4123–4137, https://doi.org/10.1109/JSEN.2019.2896412.

[29] M. Aghaei, A. Dolara, S. Leva, F. Grimaccia, Image resolution and defects detection in PV inspection by unmanned technologies, in: Proc. 2016 IEEE Power Energy Soc. Gen. Meet. (PESGM), IEEE, Boston, MA, USA, 2016, pp. 1–5, https://doi.org/10.1109/PESGM.2016.7741811.

[30] F. Grimaccia, S. Leva, A. Niccolai, PV plant digital mapping for modules' defects detection by unmanned aerial vehicles, IET Renew. Power Gener. 11 (2017) 1221–1228, https://doi.org/10.1049/iet-rpg.2016.0944.

[31] P.B. Quater, F. Grimaccia, S. Leva, M. Mussetta, M. Aghaei, Light unmanned aerial vehicles (UAVs) for cooperative inspection of PV plants, IEEE J. Photovolt. 4 (2014) 1107–1113, https://doi.org/10.1109/JPHOTOV.2014.2323715.

[32] A. Sinha, O.S. Sastry, R. Gupta, Nondestructive characterization of encapsulant discoloration effects in crystalline-silicon PV modules, Sol. Energy Mater. Sol. Cells 155 (2016) 234–242, https://doi.org/10.1016/j.solmat.2016.06.005.

[33] International Electrotechnical Commission (IEC), IEC 61646 Ed2.0 - Thin-film terrestrial photovoltaic (PV) modules - Design qualification and type approval, IEC, Geneva, Switzerland, 2008.

[34] B. Taghezouit, F. Harrou, Y. Sun, W. Merrouche, Model-based fault detection in photovoltaic systems: A comprehensive review and avenues for enhancement, Results Eng. 21 (2024) 101835, https://doi.org/10.1016/j.rineng.2024.101835.

[35] M.H. Ali, A. Rabhi, A.E. El Hajjaji, G.M. Tina, Real-time fault detection in photovoltaic systems, Energy Procedia 111 (2017) 914–923, https://doi.org/10.1016/j.egypro.2017.03.256.

[36] M. Mansouri, M. Trabelsi, H. Nounou, M. Nounou, Deep learning-based fault diagnosis of photovoltaic systems: A comprehensive review and enhancement prospects, IEEE Access 9 (2021) 126286–126306, https://doi.org/10.1109/ACCESS.2021.3112065.

[37] F. Suliman, F. Anayi, M. Packianather, Electrical faults analysis and detection in photovoltaic arrays based on machine learning classifiers, Sustainability 16 (2024) 1102, https://doi.org/10.3390/su16031102.

[38] G.M. Toche Tchio, J. Kenfack, D. Kassegne, F.-D. Menga, S.S. Ouro-Djobo, A comprehensive review of supervised learning algorithms for the diagnosis of photovoltaic systems, proposing a new approach using an ensemble learning algorithm, Appl. Sci. 14 (2024) 2072, https://doi.org/10.3390/app14052072.

[39] R. Moyo, M. Dewa, H.M. Romero, V.A. Gómez, J. Aragonés, L. Hernández-Callejo, An adaptive neuro-fuzzy inference scheme for defect detection and classification of solar PV cells, Renew. Energy Sustain. Dev. 10 (2024) 218–232, https://doi.org/10.21622/resd.2024.10.1.218.

[40] R. Mohanty, U.S.M. Balaji, A.K. Pradhan, An accurate noniterative fault-location technique for low-voltage DC microgrid, IEEE Trans. Power Deliv. 31 (2016) 475–481, https://doi.org/10.1109/TPWRD.2015.2456934.

[41] N. Chaudhuri, B. Chaudhuri, R. Majumder, A. Yazdani, Multi-terminal direct-current grids: Modeling, analysis, and control, Wiley-IEEE Press, 2014, https://doi.org/10.1002/9781118960530.

[42] S.D.A. Fletcher, P.J. Norman, S.J. Galloway, P. Crolla, G.M. Burt, Optimizing the roles of unit and non-unit protection methods within DC microgrids, IEEE Trans. Smart Grid 3 (2012) 2079–2087, https://doi.org/10.1109/TSG.2012.2198499.

[43] A.A.S. Emhemed, K. Fong, S.D.A. Fletcher, G.M. Burt, Validation of fast and selective protection scheme for an LVDC distribution network, IEEE Trans. Power Deliv. 32 (2017) 1432–1440, https://doi.org/10.1109/TPWRD.2016.2609224.

[44] P. Nooralishahi, C. Ibarra-Castanedo, S. Deane, F. López, S. Pant, M. Genest, N.P. Avdelidis, X.P.V. Maldague, Drone-based non-destructive inspection of industrial sites: A review and case studies, Drones 5 (2021) 106, https://doi.org/10.3390/drones5040106.

[45] K. Masita, A. Hasan, T. Shongwe, H. Abu Hilal, Deep learning in defects detection of PV modules: A review, Sol. Energy Adv. 5 (2025) 100090, https://doi.org/10.1016/j.seja.2025.100090.

[46] I. Polymeropoulos, S. Bezyrgiannidis, E. Vrochidou, G.A. Papakostas, Enhancing solar plant efficiency: A review of vision-based monitoring and fault detection techniques, Technologies 12 (2024) 175, https://doi.org/10.3390/technologies12100175.