An adaptive four-layer digital twin with segmented diode model for real-time fault diagnosis and output characterization of PV modules

Yihan Chen [a], Mingyao Ma [a,*], Wenting Ma [a], Xilian Zhou [a], Rui Zhang [b], Zhenyu Fang [b]

[a] National and Local Joint Engineering Laboratory for Renewable Energy Access to Grid Technology, Hefei University of Technology, 230041, Hefei, China
[b] Sungrow Smart Maintenance Technology Co., Ltd, 231200, Hefei, China

A B S T R A C T

With the rapid growth of photovoltaic (PV) power generation, accurate fault diagnosis and output characterization of PV modules have become crucial challenges. Existing PV models often lack real-time adaptability, limiting their effectiveness under complex and dynamic operational conditions. To address these limitations, this paper proposes a novel four-layer digital twin (DT) model architecture specifically designed for PV modules. The framework uniquely integrates a physics-informed segmented diode model with a hybrid Fungal Growth-Differential Evolution (FG-DE) optimization algorithm, enabling high-precision modeling and robust multi-source data fusion. Unlike traditional models, the proposed segmented diode model is dynamically adjusted to varying fault conditions – both mismatch and non-mismatch – providing detailed insights into fault mechanisms and output behavior. The virtual layer continuously synchronizes with physical module data, while the decision layer employs a hierarchical machine learning strategy for accurate, real-time fault classification and performance prediction. Experimental validation across four large-scale PV power stations demonstrates that the model achieves an average fault diagnosis accuracy of 98.49%, reflecting a 2.37% improvement over traditional methods. In terms of output characteristic representation, the proposed method attains an average root mean square error (RMSE) of 0.0555 A and a coefficient of determination R^2 of 0.9715, indicating a 70.10% reduction in error and a 10.88% improvement in goodness of fit, respectively. These results underscore the model's significant potential for deployment in intelligent PV systems, offering a scalable, data-driven solution for real-time monitoring, diagnosis, and control.

1. Introduction

The global energy landscape is shifting toward renewable sources, with photovoltaic (PV) technology playing a central role. According to the International Energy Agency, renewable energy's proportion in global power generation is expected to expand from 30% in 2023 to 37% by 2026, driven primarily by advances in solar PV technology [1].

PV modules are critical components in PV power generation process, and their accurate modeling is essential for reliable control, fault diagnosis, operation, and maintenance [2–5]. Existing approaches primarily rely on a traditional single-diode model (SDM) [6] and double-diode model (DDM) [7,8], wherein artificial intelligence algorithms are employed to optimize model parameters and enhance parameter extraction accuracy [7,9,10]. In [11], the SDM parameters were extracted using a I–V translation method combined with a modified Grey Wolf Optimization algorithm, while two Random Forest Classifiers were trained for fault detection and classification. While effective under standard conditions, these methods are inadequate for

current mismatch faults, which introduce inflection points in the I–V curve that static models fail to capture or adapt to. With the advancement of deep learning technologies, data-driven models are capable of accurately capturing and representing complex nonlinear input–output relationships. In [12], a I–V curve-based model using convolutional neural network (CNN) with Convolutional Block Attention Modules was proposed to Analyze the fault characteristics. In [13], an Unsupervised clustering and probabilistic neural networks were employed to establish fault diagnosis model under different operating conditions. In [14], a fault diagnosis model based on stacked auto-encoders and clustering algorithms was designed to automatically extract effective features from I–V curves. In [15,16], data-driven PV output characteristic models were developed using different artificial neural networks (ANN). Despite their high accuracy, these methods often demand extensive datasets and face significant challenges in transferability and generalization, limiting their practicality for real-world deployment. In [17],

Table 1
Benefits and drawbacks of DT.

Benefits	Limitations
Real-time monitoring and control	High implementation costs
Predictive maintenance and operational optimization	Complexity in system integration
Enhanced collaboration and communication	High communication costs
Simulations for operational planning	Cybersecurity risks
Real-time data-driven decision-making	Ongoing maintenance and upgrades

a mathematical model for PV panels was developed under dust accumulation conditions, considering non-mismatch faults, but this model is limited to dust accumulation faults. In [18], a reverse-bias model for PV modules was established using an equivalent circuit model under partial shading conditions, addressing current mismatch faults caused by irradiance non-uniformity. However, it is not applicable to other types of current mismatch faults. Traditional static models are inadequate for real-time adjustments based on the physical fault conditions of PV modules, which are necessary for accurately characterizing output characteristics and diagnosing faults under varying operational conditions. These limitations reveal a critical gap in existing PV modeling approaches: the absence of dynamic, real-time adaptability to both environmental variability and evolving fault mechanisms. This underscores the need for a next-generation modeling framework that can bridge physical and virtual systems seamlessly.

The concept of Digital Twin (DT), originating from the Apollo program and formally introduced in 2002, has bridged physical systems and virtual models through real-time synchronization and multi-source data integration [19,20]. With core capabilities such as intelligent decision support, dynamic monitoring and adjustment, DT presents a compelling solution for overcoming the static limitations of traditional models. In [21,22], key benefits and limitations of DT implementation have been identified, as presented in Table 1.

Existing research has made significant strides in DT applications for PV systems. A DT model of the photovoltaic energy conversion unit was established in [23], and the fault diagnosis was conducted by evaluating discrepancies between the predicted output of DT model and the measured output of physical entities. In [24], a DT model was developed for the PV system under sunny weather, which predicts power and performs fault diagnosis by subtracting the estimated from the measured outputs. In [25], a DT model was established for building-attached photovoltaic systems under partial shading conditions and used ANN for fault diagnosis. In [26], the DT model of power plants was divided into three parts: PV panels, DC–DC converters, and grids based on different time resolutions, where each model represented a specific component of solar PV power generation. This method can detect the performance and health of each component and improve the speed of the system, but the modeling part in the literature is not comprehensive enough. Additionally, in [27] a hybrid neural network was employed to simulate P-V characteristics of PV panels, creating a DT model capable of accommodating diverse environmental variations. However, this approach does not account for fault conditions in PV panels. While several studies have explored DT implementations at the system level, few have addressed the core PV module in sufficient depth, especially under complex fault scenarios.

Most existing DT models target entire PV systems, often overlooking the granularity and complexity of individual PV modules. Moreover, current approaches lack modularity and hierarchical structure, limiting diagnostic precision and real-time adaptability. In addition, many recent studies rely primarily on scalar electrical features, failing to exploit the rich diagnostic information embedded in the complete

I–V curve. Moreover, most existing models treat normal and faulty conditions separately, lacking a unified physical framework capable of consistently representing both states. Motivated by the above findings, this paper proposes a four-layer DT model for PV modules. Through the integration of DT concepts, Fungal Growth-Differential Evolution (FG-DE) optimization algorithm, and multi-source data fusion, the model identifies mismatch and non-mismatch faults, accurately representing PV module output characteristics under varied conditions. Unlike prior DT-based approaches that focus on system-level modeling or rely solely on deep learning, our method enables module-level analysis by leveraging the entire I–V curve and introducing a unified segmented diode model applicable to both normal and fault conditions. This approach lays the groundwork for modular DT-based PV power station design, advancing fault diagnosis accuracy and operational efficiency across diverse scenarios. The major contributions of this study are summarized as follows:

1. **Modular Four-Layer DT Architecture**: A four-layer DT model for PV modules is designed to facilitate dynamic real-time adjustments in response to environmental factors and operational conditions. Its modular architecture ensures high adaptability and scalability, enabling seamless integration across diverse PV systems;
2. **Segmented Diode Model with Physics-Based Fault Inversion**: The physics-based segmented diode model for PV modules facilitates the inversion of fault mechanisms and ensures precise fault diagnosis and accurate representation of output characteristics across varied scenarios. The model parameters are derived from multi-source data and identified using the FG-DE optimization algorithm;
3. **Hierarchical Fault Diagnosis Using HSVM-DT**: A hierarchical Support Vector Machine with Decision Tree (HSVM-DT) structure is proposed for fault detection, utilizing selected parameters from the diode model to classify both mismatch and non-mismatch faults with high accuracy;
4. **High-Fidelity Output Characterization**: The characterization of output characteristics across various conditions facilitates an accurate representation of the real-world performance of PV modules under fault scenarios, thereby enhancing control and scheduling capabilities.

The remainder of this paper is structured as follows: Section 2 outlines the DT framework for PV modules, covering the physical, connection, virtual, and decision layers. Section 3 examines fault mechanisms of PV modules in the physical layer, particularly current mismatch faults, and analyzes the output characteristics using equivalent circuit models. In Section 4, the connection layer bridges the physical and virtual layer. In Section 5, a physics-based segmented diode model is introduced in the virtual layer, along with the FG-DE optimization algorithm for parameter identification, achieving an RMSE of 0.0585 and a correlation coefficient R^2 of 0.9922. Section 6 describes the decision layer, which utilizes the identified parameters from the virtual layer for fault diagnosis and PV module output characterization, attaining an average fault diagnosis accuracy of 98.49% and an average RMSE of 0.0555 A in output characterization. Finally, Section 7 outlines the principal findings and contributions of this study.

2. DT architecture for PV modules

The reliability of PV systems critically depend on the accurate modeling and monitoring of PV modules, which are highly sensitive to environmental variations and complex fault mechanisms. To address these challenges, a novel four-layer DT architecture is proposed, which is tailored specifically for PV modules. This modular structure enables real-time fault diagnosis, dynamic output characterization, and seamless integration within PV systems, as presented in Fig. 1.
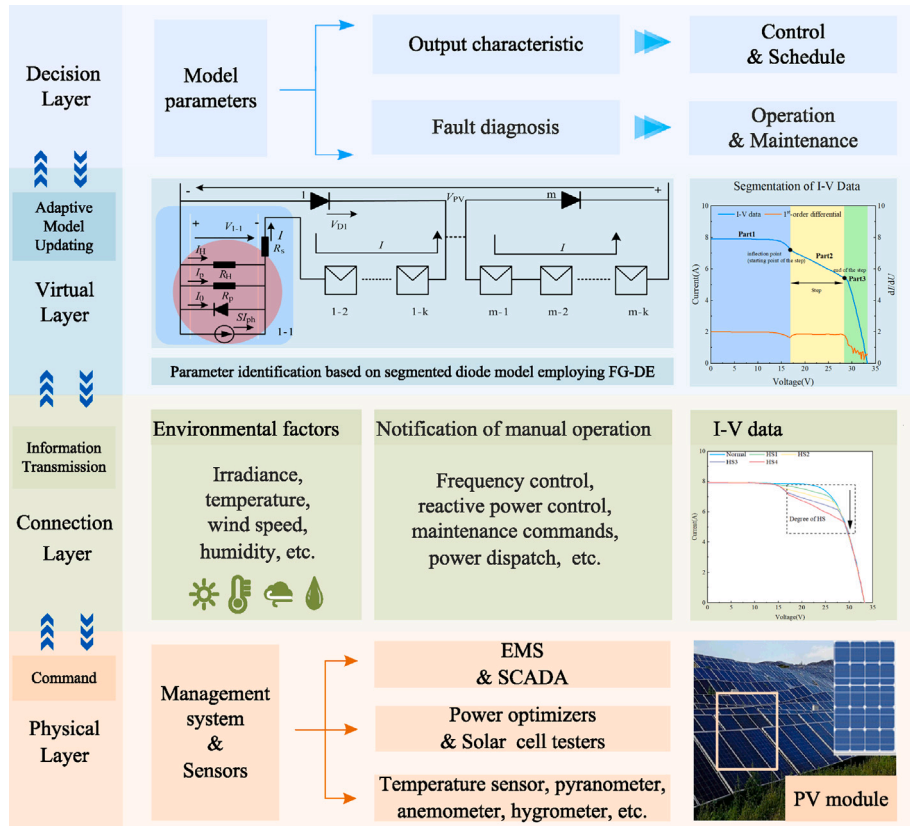
Fig. 1. DT architecture for PV modules.

Physical Layer: At the foundation, this layer encompasses the physical entity of PV modules with a comprehensive sensor network managed by an Energy Management System (EMS) and Supervisory Control and Data Acquisition (SCADA) system. Designed to capture the real-time operational state of PV modules, the sensor network reflects real-world performance under various environmental conditions by gathering critical operational data. In contrast to conventional models, a physics-informed equivalent circuit model is proposed to characterize and diagnose PV modules mismatch faults, offering deeper insights into fault propagation mechanisms;

Connection Layer: Serving as the system's communication interface, this layer relays data streams such as I–V data, environmental factors, and operational commands from the physical layer to the virtual layer. Compared to existing DT frameworks that often lack efficient data handling for PV modules, this layer is optimized for the high-dimensional and heterogeneous data encountered in PV operations. This robust connectivity ensures that real-time physical data is promptly integrated into the model in the virtual layer;

Virtual Layer: A high-fidelity digital replica of the physical entity is developed, dynamically adapting to real-time data received from the connection layer. Compared to conventional models, a novel segmented diode model employing FG-DE optimization algorithm is employed to simulate the output characteristics of PV modules under both normal and fault conditions. This innovation enables real-time fault detection, dynamic adjustment, and enhanced the efficiency of PV systems;

Decision Layer: This layer is instrumental in fault diagnosis and operational optimization by designating selected model parameters as diagnostic fault vectors, significantly enhancing diagnostic accuracy and aiding in proactive maintenance strategies. By leveraging predictive insights from the virtual layer, this layer facilitates intelligent control and adaptive scheduling, ensuring real-time system adjustments under both normal and fault scenarios. In contrast to conventional approaches, this novel architecture integrates data-driven diagnostics

with predictive control mechanisms, thereby enhances fault tolerance and optimizes overall PV system performance.

The architecture implements a sophisticated feedback loop mechanism wherein continuous data generation from the physical layer flows through real-time transmission protocols to comprehensive processing within the virtual layer. This processed information enables advanced decision-making in the decision layer, ultimately resulting in optimized command execution through the system hierarchy. This continuous feedback mechanism ensures system adaptability and operational optimization under varying conditions.

3. Physical layer: Operational states and fault mechanisms of PV modules

This section describes the operational states and fault mechanisms in PV modules, beginning with a mathematical model for the I–V characteristics of PV cells based on their physical properties, which is then extended to modules. Faults are classified into two primary categories: current mismatch faults and non-current mismatch faults. A detailed analysis is provided for three prevalent mismatch faults. Finally, the equivalent circuit model is utilized for assessing the output characteristics of PV modules under each fault type.

3.1. Normal PV module model

In PV systems, individual cells, the smallest power-generating units, are commonly represented by SDM and DDM as presented in Fig. 2. Nevertheless, both models remain insufficient for dynamically modeling fault-induced nonlinearity, motivating the development of the segmented diode model proposed in this paper. According to Kirchhoff's current law, the output equations for SDM and DDM are derived as follows:

$$I = I_{\mathrm{ph}} - I_0 \left[\exp\left(\frac{V + IR_{\mathrm{s}}}{nV_{\mathrm{t}}} \right) - 1 \right] - \frac{V + IR_{\mathrm{s}}}{R_{\mathrm{p}}} \tag{1}$$

(a) Equivalent circuit based on SDM

(b) Equivalent circuit based on DDM

Fig. 2. The equivalent circuit model of PV cell.

(a) PV station in Shaoguan, Guangdong Province

(b) PV station in Huai'an, Jiangsu Province

(c) PV station in Hefei, Anhui Province

(d) PV station in Suzhou, Anhui Province

Fig. 3. Actual scene of the PV power station in China.

$$I = I_{ph} - I_{01}\left[\exp\left(\frac{V + I R_s}{n_1 V_t}\right) - 1\right] - I_{02}\left[\exp\left(\frac{V + I R_s}{n_2 V_t}\right) - 1\right] - \frac{V + I R_s}{R_p} \quad (2)$$

where I is the PV cell's output current, I_{ph} is the photocurrent, I_0 is the diode's reverse saturation current, R_s is the series resistance, R_p is the parallel resistance, V_t is the thermal voltage and n is the diode ideality factor. For PV modules and arrays, the equations scale by the number of cells in series (N_s) and parallel (N_p) connections. The physical layer was experimentally investigated at four actual PV power stations in different provinces of China, as shown in Fig. 3.

3.2. Non-current mismatch faults

Several types of non-mismatch faults commonly affect PV modules, altering their electrical characteristics without inducing distinctive inflection points in the I–V curve.

Abnormal aging, caused by prolonged field exposure or accelerated testing (e.g., thermal cycling and damp-heat stress), leads to chemical degradation of contact materials and encapsulant browning. This increases the series resistance R_s, thereby degrading current output. Fig. 4(a) shows darkened areas in Electroluminescence (EL) images, indicating material-level degradation. Dust fault, whether natural or artificially applied, attenuates incident irradiance, resulting in reduced photocurrent I_{ph}. Fig. 4(b) illustrates the uniform surface coverage used in experimental simulations. Short-circuit faults, induced via direct wiring between cells or bypassed traces, significantly lower the

Table 2
Effect of non-mismatch fault.

Fault type	Fault mechanism	Parameter variation
Abnormal aging	Chemical degradation of contact materials and encapsulant browning	R_s: increase
Dust	Dust accumulation attenuates incident irradiance	I_{ph}: reduce
Short-circuit	Low-resistance connections between cells	V_{OC}: reduce
PID	Shunting effects within the p–n junction caused by sustained voltage differences	V_{OC}: reduce

open-circuit voltage V_{OC}. This is demonstrated in Fig. 4(c). Potential-Induced Degradation (PID) results from sustained high-voltage stress across the p–n junction, forming unwanted leakage pathways along cell edges. Fig. 4(d) shows corresponding edge darkening in EL images. This reduces the parallel resistance R_p. Table 2 summarizes the effect of each non-mismatch fault.

For each of these fault types, I–V data were collected and compared to that of healthy modules. These faults primarily manifest as gradual parameter drifts within the SDM or DDM framework, rather than abrupt
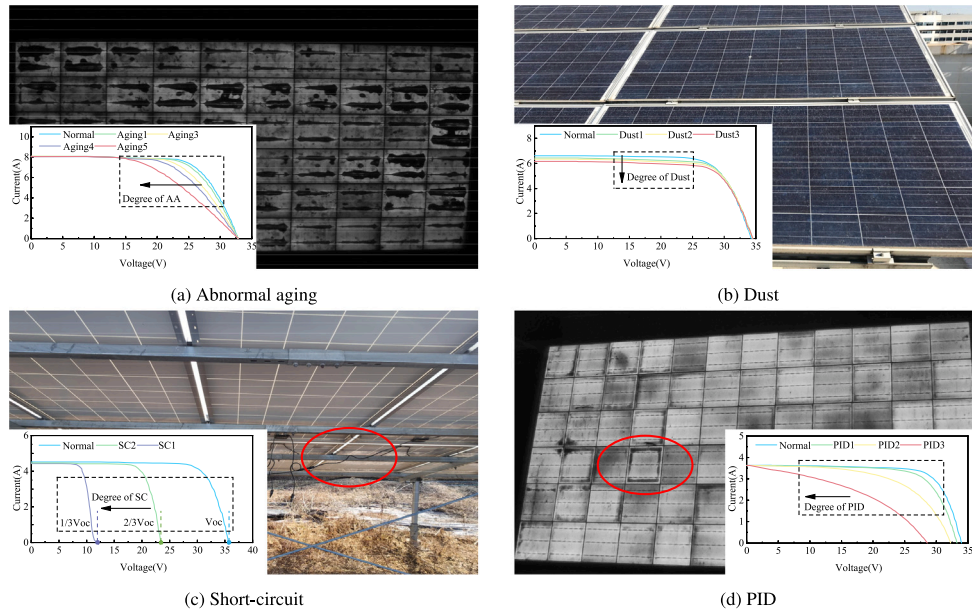
(a) Abnormal aging

(b) Dust

(c) Short-circuit

(d) PID

Fig. 4. PV modules with non-mismatch faults.

structural changes in the curve shape. Therefore, while their effects are non-negligible, they do not necessitate a segmented modeling approach and are not further elaborated in this study.

3.3. Current mismatch faults

Unlike non-mismatch faults, current mismatch faults introduce clear inflection points in the I–V curve. These characteristics make them detectable through slope variation and curve segmentation, forming the basis for the proposed segmented diode model. EL images and infrared (IR) thermography are employed to capture the characteristic fault images in Fig. 5. Specifically, Fig. 5(a) demonstrates partial shading conditions, where irradiance is obstructed, leading to a non-uniform power generation profile. Fig. 5(b) highlights the presence of hot spot through elevated temperatures detected via IR thermography. Fig. 5(c) presents EL images where cracked cells are clearly visible. Additionally, the I–V data corresponding to each fault type are collected and compared against normal PV modules. Mismatch faults lead to unique inflection points in the I–V curve, which deviate from standard SDM and DDM behaviors but apply to individual cells. Each fault's mechanism and characteristic step in the I–V curve are analyzed below.

3.3.1. Partial shading

Partial shading reduces the photocurrent of shaded cells, while a simplified example is shown in Fig. 6(a) where one cell in the first substring is shaded. Depending on the bypass diode state, the I–V characteristics vary significantly, which can be classified into the following scenarios:

(1) $0 < I \leq S I_{SC}$ (where S is the shading factor, I_{SC} is the short-circuit current)

Full Generation Mode: All cells generate power without diode conduction.

$$V_1 = V_{1\text{-}1} + \sum_{i=2}^{k} V_{1\text{-}i} = f(I)\Big|_{I_{ph}=I_{ph}\cdot S} + (k-1)f(I)\Big|_{I_{ph}=I_{ph}} \quad (3)$$

(2) $S I_{SC} < I \leq I_{SC}$ and $V_1 > V_{bd}$ (where V_{bd} is the conduction voltage of the bypass diode)

Reverse Biased Shaded Cell: The shaded cell becomes reverse biased, with the diode remaining off.
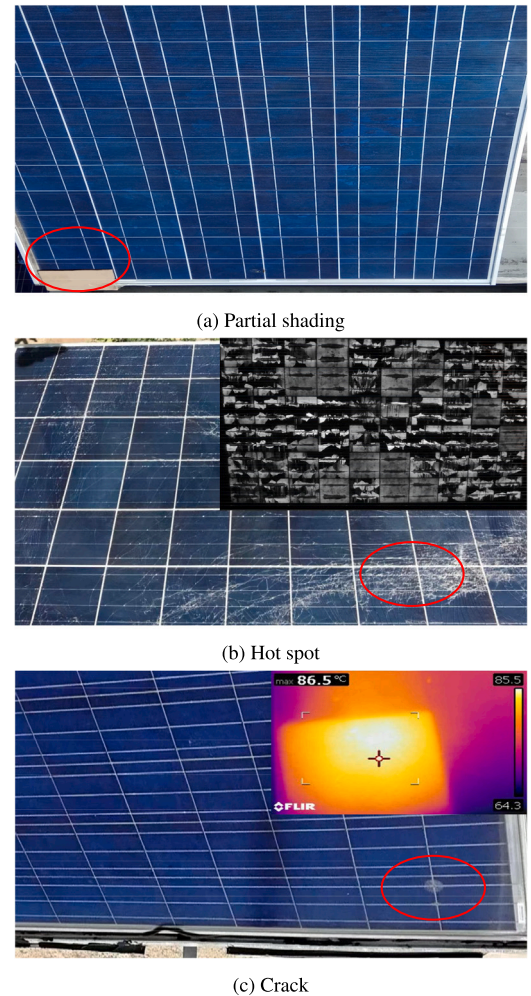

(a) Partial shading

(b) Hot spot

(c) Crack

Fig. 5. PV Modules with Mismatch Faults.
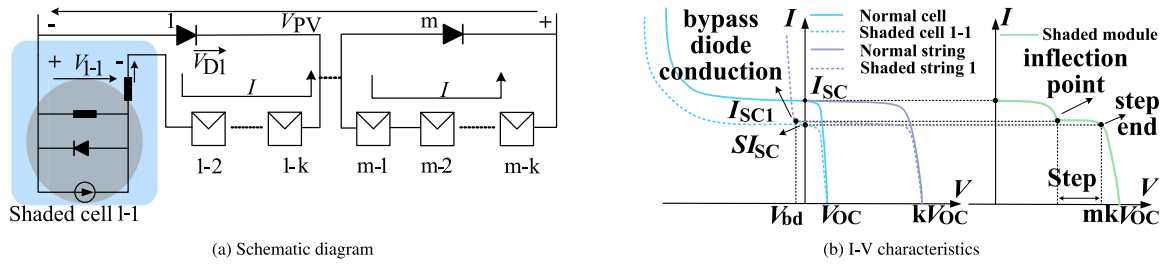
(a) Schematic diagram

(b) I-V characteristics

Fig. 6. Schematic diagram and I–V characteristics of PV module under partial shading faults.

(a) Schematic diagram

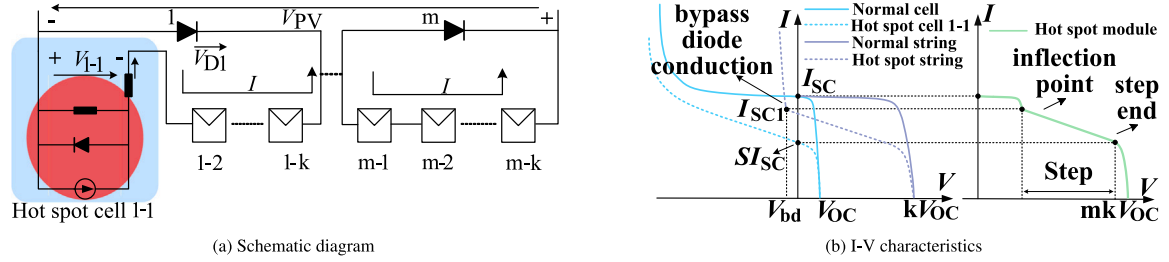(b) I-V characteristics

Fig. 7. Schematic diagram and I–V characteristics of PV module under hot spot faults.

$$V_1 = V_{1\text{-}1} + \sum_{i=2}^{k} V_{1\text{-}i} = f(I)\Big|_{I_{ph}=I_{ph}\cdot S} + (k-1)f(I)\Big|_{I_{ph}=I_{ph}} \quad (4)$$

Despite the voltage of the shaded cell is negative, the actual relationship still follows the diode model.

(3) $SI_{SC} < I \le I_{SC}$ and $V_1 = V_{bd}$

The shaded cell experiences reverse voltage, and the bypass diode is on the verge of conducting.

(4) $SI_{SC} < I \le I_{SC}$ and $V_1 < V_{bd}$

The shaded cell remains in reverse bias, with reverse voltage applied. The bypass diode conducts, and the I–V characteristic of PV string 1 is given by:

$$V_1 = V_{D1} \quad (5)$$

The voltage V_1 is clamped to the bypass diode voltage V_{D1}, typically around -0.7 V.

The I–V characteristics of the shaded module and its internal shaded and normal cells are shown in Fig. 6(b). For cases where multiple PV strings are shaded, the theoretical analysis is similar and will not be reiterated here.

3.3.2. Hot spot

The hot spot studied here refers primarily arises from low-resistance defects which can be modeled by a parallel resistor R_H in the equivalent circuit of the hot spot cell. This represents the multilevel trap-assisted tunneling effect, which generates leakage current. As presented in Fig. 7(a), the schematic diagram for hot spot module resembles that of the shaded module, with shaded cells replaced by hot spot cells. This fault mechanism introduces an inflection point in the module's I–V curve, followed by a steep drop in current, as shown in Fig. 7(b), which distinguishes hot spot faults from partial shading fault. Such complex, segmented behaviors in the I–V curve cannot be adequately captured by conventional single-parameter diode models that assume global continuity and smoothness.

3.3.3. Crack

The crack studied here primarily refers to cases where the PV cells are damaged but their electrical connections remain intact. The schematic of PV module under crack faults is presented in Fig. 8(a). The I–V curves of a cracked module and its internal normal and cracked cells are shown in Fig. 8(b). The cracked string exhibits distinct

reverse bias characteristics in different stages, with the I–V curve slope decreasing progressively.

Assume that the equivalent circuit for the cracked cells includes additional parallel resistances(R_{C1}, R_{C2}, R_{C3}), with varying degrees of photocurrent and damage represented by factors (S_1, S_2, S_3), where $S_1 > S_2 > S_3$.

(1) $0 < I \le S_3 I_{SC}$

All cells operate in forward mode, and the bypass diode is not conducting. The I–V output characteristic of the cracked string is given by:

$$V_1 = V_{1\text{-}1} + \sum_{i=2}^{k} V_{1\text{-}i} = f(I)\Big|_{I_{ph}=I_{ph}\cdot S_q} + (k-1)f(I)\Big|_{I_{ph}=I_{ph}} \quad (6)$$

where p and q represent the indices of normal and cracked cells, respectively.

(2) $S_3 I_{SC} < I \le S_2 I_{SC}$

Cracked cells 1–3 are in reverse bias and dominate power dissipation, while the remaining cells generate power. The I–V characteristic of the cracked string follows Eq. (7), but $V_{1\text{-}3}$ is negative.

(3) $S_2 I_{SC} < I \le S_1 I_{SC}$

Cracked cells 1–2 and 1–3 are in reverse bias and dominate power dissipation, while other cells generate power.

(4) $S_1 I_{SC} < I \le SI_{SC}$ and $V_1 > V_{bd}$

All cracked cells are in reverse bias and dominate power dissipation, and the bypass diode is not conducting.

(5) $S_1 I_{SC} < I \le SI_{SC}$ and $V_1 = V_{bd}$

The bypass diode is on the verge of conducting, operating in maximum reverse bias.

(6) $S_1 I_{SC} < I \le SI_{SC}$ and $V_1 < V_{bd}$

The bypass diode conducts, and the voltage of the cracked PV string is clamped to V_{D1}.

In summary, this analysis of fault mechanism provides a robust theoretical foundation for the unified modeling of both normal and fault PV modules. Mismatch and non-mismatch faults induce distinct alterations in PV module electrical characteristics. Specifically, mismatch faults create observable inflection points and segmented behavior in the I–V curve. Despite these discontinuities, e segment between these
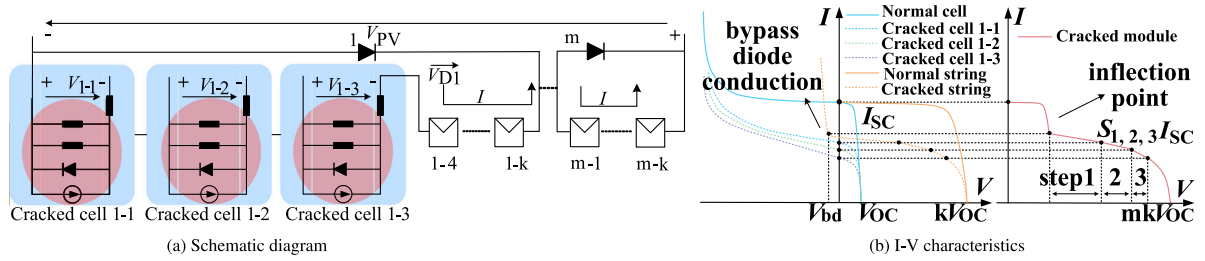
Fig. 8. Schematic diagram and I–V characteristics of PV module under crack faults.

inflection points can still be accurately described by a standard diode-based model, indicating that the fundamental physical behavior of each sub-region remains consistent with the underlying PV principles. By modeling each segment independently with dedicated parameter sets, the approach enables precise identification of fault-induced variations and accurate characterization of the module's electrical output. Importantly, the core model structure remains unified for both normal and faulty modules, the distinction lies in the number of segments and the corresponding parameter values in each segment.

4. Connection layer: Data acquisition and transmission

The connection layer bridges the physical and virtual layers by enabling efficient data collection and transmission, which is essential for the operation of the proposed DT model. As the proposed DT framework heavily relies on accurate I–V curve data, the feasibility and practicality of obtaining such data are of paramount importance. Traditionally, I–V curves are acquired through the use of power optimizers and portable solar cell testers. Meanwhile, the widespread deployment of smart inverters has significantly enhanced the feasibility and convenience of acquiring I–V curves at both the module and string levels. Without the need to disconnect strings or interrupt grid-tied operations, these inverters enable full-range I–V curve sampling with second-level resolution, thereby minimizing power generation losses [28,29]. Most modules employ SP-375 W power optimizers or SG series smart inverters by Sungrow Power Supply Co., Ltd., facilitating rapid I–V data acquisition. For modules without optimizers and smart inverters, I–V data are recorded with the AV6591 portable solar cell tester. The data are subsequently transmitted to the cloud platform and high-performance embedded processing unit based on the Raspberry Pi, enabling further processing at the virtual layer.

5. Virtual layer: Segmented diode model

This section presents the core contribution of this study: a segmented diode model that dynamically adapts to both normal and faulty conditions of PV modules. Traditional diode models rely on a single set of parameters, limiting their ability to represent complex behaviors. The proposed model divides the I–V curve into physically interpretable segments, each governed by locally optimized parameters. This allows precise modeling of inflection-point-induced nonlinearity, particularly prevalent under mismatch faults, which is essential for accurate real-time fault diagnosis within the DT framework.

The construction of the segmented diode model commences with the acquisition and pre-processing of raw data from real-world PV power plants. Subsequently, a slope-based method is applied to the I–V curve to accurately locate the inflection point (step onset) and step end associated with mismatch faults. Each segment is further analyzed using a FG-DE optimization algorithm to perform parameter identification, enabling the extraction of distinct physical parameters for each stage and facilitating real-time representation of the full-range output characteristics, as presented in Fig. 9.
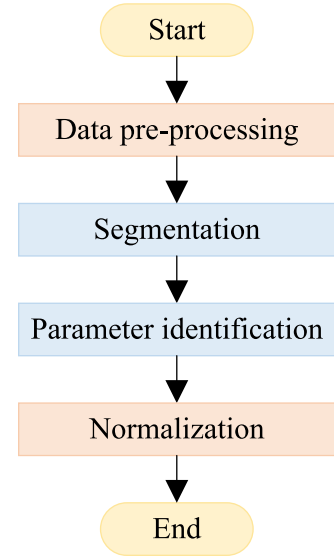

Fig. 9. Flowchart of the segmented diode model.

5.1. Data pre-processing

Accurate I–V data collection requires rigorous pre-processing to address outliers, noise, and redundancy arising from environmental fluctuations and sensor noise. However, due to the extremely fast acquisition speed of I–V curves sampling with second-level resolution, the impact of typical weather variations during data capture can be largely neglected. For the small subset of special weather conditions and the unavoidable sensor noise that do affect the I–V curves, the Locally Estimated Scatterplot Smoothing (LOESS) method [30] is employed in this study for I–V data pre-processing. LOESS, a weighted local polynomial fitting technique, is particularly suitable for handling nonlinear variations and offers strong robustness against outliers. This robustness makes LOESS an ideal choice for I–V data preprocessing in PV parameter identification studies. Sharp variations observed in the first derivative of the raw data indicate abrupt changes that may introduce errors in identifying reflection points critical for accurate analysis. For example, in the analysis of hot spot faults, a comparative illustration of the data before and after LOESS preprocessing is presented in Fig. 10, demonstrating the smoothing effect and improved reliability of the processed data.

5.2. Segmentation of I-V data

Mismatch faults exhibit segmented characteristics in the I–V curve due to variations in diode conduction states, motivating the development of a segmented diode model that divides the curve into multiple
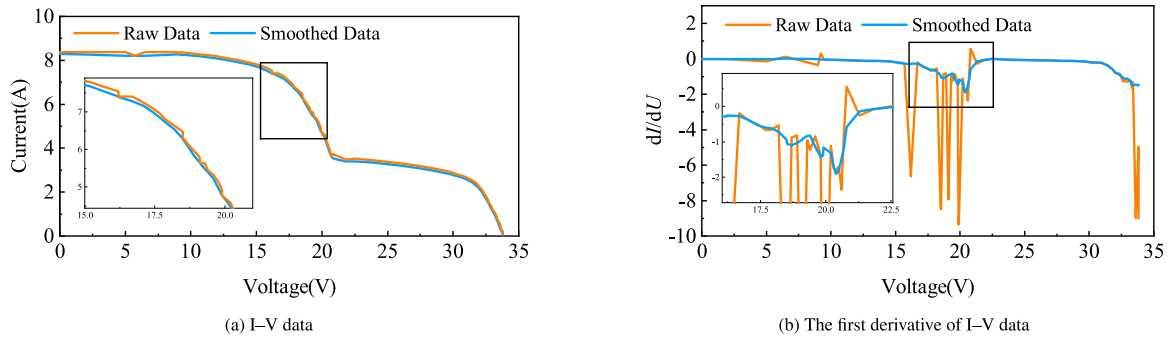
(a) I–V data

(b) The first derivative of I–V data
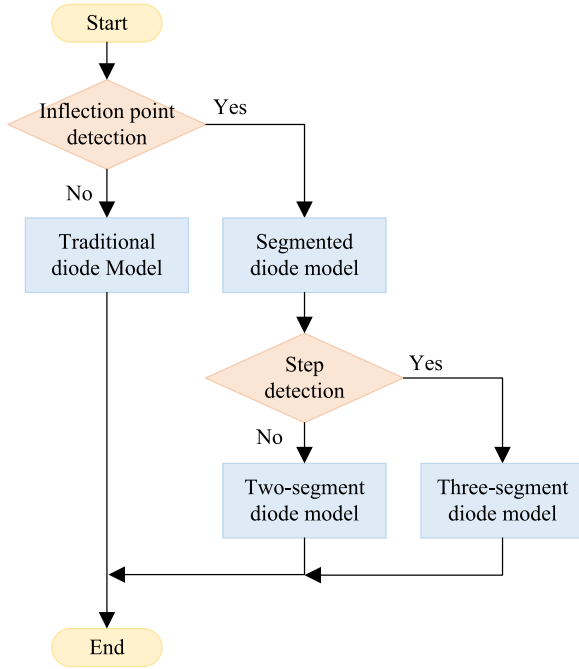
Fig. 10. Original data and smoothed data.

Fig. 11. Flowchart of the segmentation process.

regions. The I–V curve can be segmented into three segments based on cell state and diode conduction, with segmentation boundaries defined at the inflection point (step onset) and the step end, as shown in Fig. 11. Precise segmentation hinges on accurately identifying these critical points.

5.2.1. Inflection point

For mismatch faults, the reverse-bias stress on the faulty cell becomes progressively more severe, eventually driving the voltage across the bypass diode beyond its forward conduction threshold. This transition induces a discontinuity in the I–V curve, which manifests as an inflection point. The inflection point marks a shift in the curve's concavity and is typically associated with a localized extremum in the first derivative. To accurately detect this transition, a slope-based method is used to identify the inflection point, which alters the convexity of the I–V curve, while the slope k_i between two adjacent data points (U_{i-1}, I_{i-1}) and (U_i, I_i) is:

$$k_i = \frac{I_i - I_{i-1}}{U_i - U_{i-1}} \tag{7}$$

A sharp change in slope identifies the inflection point, with the first negative peak in the slope curve marking this position. The first

derivative curve is shifted upwards by two units for clarity, as shown in Fig. 12(a).

5.2.2. Step end

In crack faults, segmentation can be performed solely based on the inflection point, as identifying the step end is unnecessary. However, for hot spot and partial shading faults, defining the step end is essential. In these cases, when the faulty cell enters a reverse-biased state while the associated bypass diode remains non-conductive, the diode branch effectively behaves as an open circuit. This configuration constrains current flow through the affected cell, resulting in a high-impedance path. Consequently, the corresponding segment of the I–V curve exhibits a near-linear behavior, characterized by a relatively constant slope and minimal curvature. The first derivative within this step phase remains nearly constant, with sharp changes marking the step end. This often manifested as a discontinuity or abrupt deviation in the first derivative of I–V curve, marking the boundary between regions governed by different equivalent circuit dynamics, as illustrated in Figs. 12(b) and 12(c).

For both partial shading and hot spot mismatch faults, the I–V curves can be divided into three segments, leading to a three-segment diode model. In contrast, for crack faults, the slope of each segment in the I–V curve gradually decreases, eventually aligning with the slope beyond the step end. The progressively shorter segment lengths cause these sections to be treated as a single segment, forming a two-segment diode model. Regardless of whether the three-segment or two-segment diode model is used, the second segment following the inflection point contains significant fault characteristic information. Therefore, the parameters of this segment are selected for fault diagnosis, while the remaining segments are utilized for output characteristic representation. This inflection-point-guided segmentation offers an effective mechanism to localize fault-induced nonlinearities, a feature not available in conventional diode models or purely data-driven approaches.

5.3. Segmented parameter identification

In PV modeling, parameter identification methods based on the diode model typically involve analytical method and numerical optimization methods. The analytical method constructs the parameters based on open-circuit voltage point, short-circuit current point, and maximum power point. Numerical optimization methods estimate parameters by minimizing the error between actual measurements and model calculation, while incorporating artificial intelligence techniques to enhance the accuracy. Particle Swarm Optimization (PSO) is a population-based optimization algorithm, but often suffers from limited solution accuracy. The Differential Evolution (DE) optimizer algorithm achieves global optimization through mutation, crossover, and selection, but its performance depends on the selection of appropriate initial values. The algorithm employs population size N_{DE} to enhance the diversity of candidate solutions within the search space and evolves
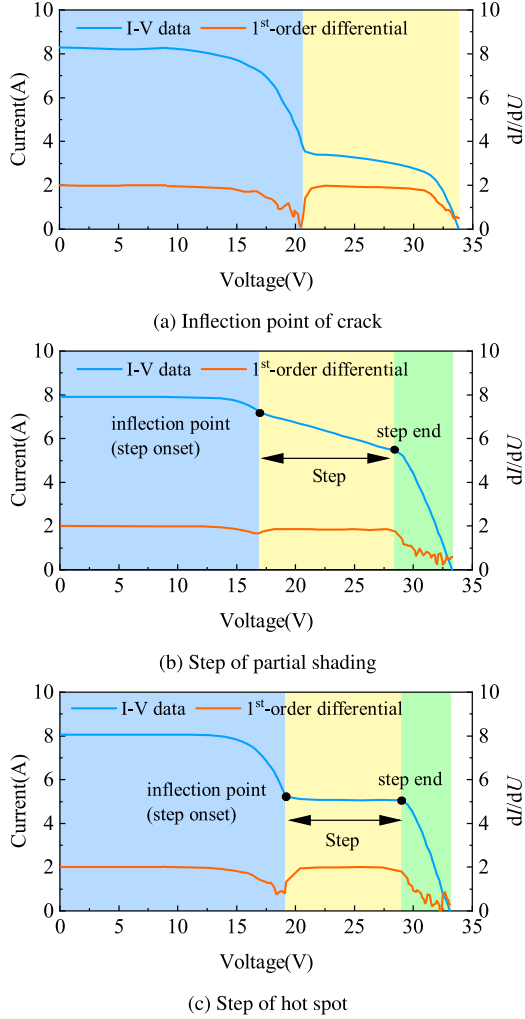
(a) Inflection point of crack

(b) Step of partial shading

(c) Step of hot spot

Fig. 12. Identification of inflection point and step end.

the population over maximum number of generations G_{DE} to facilitate progressive local refinement. The core operators are governed by the mutation factor F and crossover rate C_r, which jointly regulate the algorithm's convergence speed and population diversity. The Fungal Growth (FG) optimizer algorithm simulates fungal growth behaviors, which has fast computational speed but relatively poor performance in handling objective functions represented by transcendental equations. Similarly, the algorithm employs population size N_{FG} and maximum number of generations G_{FG} to maintain sufficient initial diversity and reduce computational overhead. The growth dynamics are governed by the growth modulation factor M and exploration probability E_p, which balance exploitation and exploration. Considering the strengths and limitations of the above algorithms, this study proposes an improved FG-DE optimizer algorithm. By exploiting the high computational efficiency of the FG optimizer algorithm, parameter identification is conducted to generate high-quality initial solutions for the DE optimizer algorithm, thereby enhancing the FG-DE optimizer algorithm's optimization performance. The hyperparameter configuration of the FG-DE optimizer in is shown in Table 3. The dataset used for evaluation is derived from I–V curve measurements under fault and non-fault conditions, with interpolation applied to standardize each curve to 50 sampling points. The full dataset is stratified and split into training and testing sets in an 70:30 ratio, ensuring balanced class representation.

Table 3
Hyperparameter configuration of the FG-DE optimizer.

Parameter	Value	Description
N_{FG}	10	Population size for FG optimizer
T_{FG}	500	Maximum number of iterations for FG optimizer
M	0.4	Growth modulation factor controlling exploration capability
E_p	0.3	Elite retention rate in the FG optimizer
N_{DE}	80	Population size for DE optimizer
G_{DE}	1200	Number of generations for DE local refinement
F	0.6	Mutation factor determining search step size in DE
C_r	0.9	Crossover rate controlling offspring generation in DE

To quantify accuracy, the RMSE between actual current and diode model current at corresponding voltages serves as the objective function, which is defined as:

$$\min \ RMSE = \sqrt{\frac{1}{k}\sum_{i=1}^{k}\left(I_{DM} - I_{real}\right)^2} \qquad (8)$$

where k is the number of I–V data points, I_{DM} is the diode model current (Eqs. (1) and (2)), and I_{real} is the actual current. By minimizing the RMSE through iterative adjustments, the calculated current values are closer to the measurements, yielding the optimal model parameters. Parameter identification is conducted using the aforementioned methods, with ten representative datasets selected for each fault type. The average RMSE, computational time t and coefficient of determination R^2 are computed and presented in Table 4.

According to the results in Table 4, the analytical method exhibits poor accuracy and computational efficiency due to the transcendental nature of the objective function. Furthermore, in the step segment of mismatch faults, the method fails due to its inability to identify the key points. While the PSO algorithm offers advancement in both precision and computational time in contrast to the analytical method, its performance remains suboptimal. In contrast, the proposed FG-DE algorithm utilizes the high computational efficiency of the FG algorithm to generate high-quality initial solutions, which are subsequently refined by the DE algorithm to improve accuracy. As a result, the proposed method achieves a lower average RMSE of 0.0585 A compared to 0.0594 A with the traditional DE algorithm, while computation time decreases from 0.3044 s to 0.2956 s. Furthermore, the R^2 value improves from 0.9917 to 0.9922, demonstrating a stronger correlation between predicted and actual values, thereby ensuring a more accurate and reliable modeling performance. The process of parameter identification is shown in Fig. 13.

The parameters of PV modules under various fault conditions are normalized under standard test conditions based on irradiance and temperature measurements at the physical layer using the following correction formulas:

$$
\begin{aligned}
I_{ph\_STC} &= I_{ph}(S,T)\frac{S_{STC}}{S} - \alpha_I\left(T - T_{STC}\right)\\
I_{0\_STC} &= \frac{I_0(S,T)}{e^{\left(\frac{E_{g\_STC}}{knT_{STC}} - \frac{E_g}{knT}\right)}}\left(\frac{T_{STC}}{T}\right)^3\\
R_{s\_STC} &= R_s(S,T)\\
R_{p\_STC} &= R_p(S,T)\frac{S}{S_{STC}}\\
n_{STC} &= n(S,T)\\
E_g(S,T) &= E_g(0) - \frac{\alpha T^2}{T + \beta}
\end{aligned}
\qquad (9)
$$

Table 4
Comparison of different methods for parameter identification.

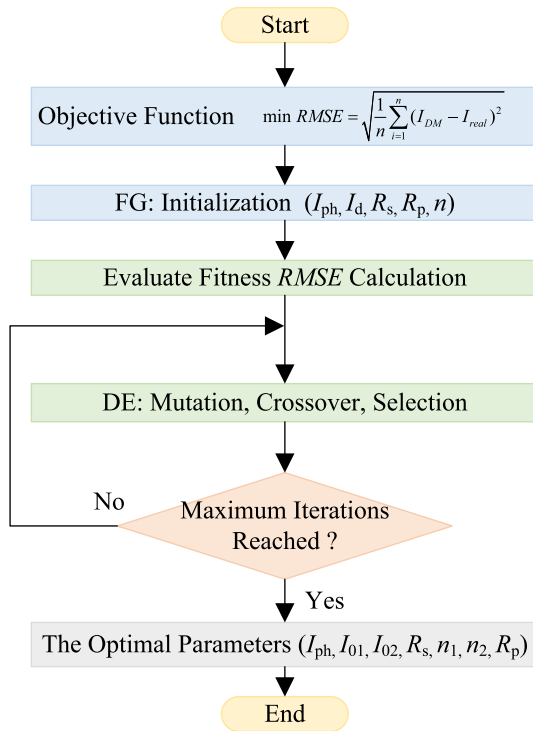| Method	Evaluation metric	Normal	Dust	PID	Short circuit	Abnormal aging	Partial shading	Hot spot	Crack	Average
Analytical method	RMSE (A)	0.6022	0.4968	0.9574	0.9794	0.7777	5.7868	6.0517	1.2111	2.1079
	t (s)	1.3260	1.3620	1.4801	1.3882	1.5091	1.1367	1.2909	1.2348	1.3410
	R^2	0.8845	0.8525	0.6994	0.6546	0.8864	−2.58e3	−401.86	−0.5448	−372.76
PSO	RMSE (A)	0.1528	0.5079	0.0909	0.2352	0.1113	0.1129	0.1111	0.1080	0.1788
	t (s)	0.1186	0.4389	0.1378	0.1016	0.1250	0.0447	0.0489	0.0823	0.1372
	R^2	0.9953	0.6725	0.9906	0.9581	0.9979	0.2304	0.7872	0.9831	0.8269
FG	RMSE (A)	0.1476	0.1564	0.0379	0.2410	0.1002	0.0333	0.0480	0.2326	0.1246
	t (s)	0.0190	0.0590	0.0237	0.0166	0.0195	0.0106	0.0124	0.0174	0.0223
	R^2	−3.68e4	0.9931	−136.21	0.9899	0.9945	0.9268	0.9810	0.9769	−4.62e3
DE	RMSE (A)	0.0808	0.0563	0.0202	0.1210	0.0577	0.0250	0.0419	0.0722	0.0594
	t (s)	0.2584	0.9413	0.2987	0.2277	0.2705	0.1224	0.1254	0.1915	0.3045
	R^2	0.9986	0.9984	0.9999	0.9957	0.9994	0.9594	0.9859	0.9967	0.9917
FG-DE	RMSE (A)	0.0789	0.0550	0.0201	0.1201	0.0575	0.0241	0.0411	0.0712	0.0585
	t (s)	0.2526	0.9213	0.2951	0.2178	0.2664	0.1091	0.1182	0.1840	0.2956
	R^2	0.9987	0.9986	0.9999	0.9957	0.9994	0.9624	0.9858	0.9968	0.9922

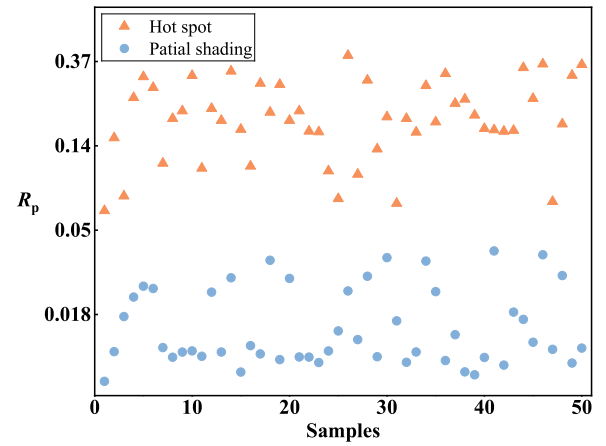Fig. 13. The process of parameter identification based on FG-DE algorithm.

Fig. 14. R_p values of partial shading and hot spot fault.

where α_I is the temperature coefficient, $q = 1.6 \times 10^{-19}$ C, $k = 1.38 \times 10^{-23}$ J/K, $E_g(0) = 1.12$ eV, $\alpha = 7.021 \times 10^{-4}$ eV/K, and $\beta = 1108$ K for silicon.

To ensure consistency across different PV module types, the fault model parameters are further normalized by dividing each parameter by its corresponding value from a normal module. This normalization facilitates generalization and enables the deployment of the proposed model across heterogeneous PV systems, as presented in Table 5.

Since mismatch faults and non-mismatch faults can be clearly distinguished based on inflection point characteristics, this study analyzes them separately.

For non-mismatch faults, extensive research has been conducted, and this study provides a brief analysis based on experimental data. Both potential induced degradation and short-circuit faults lead to a reduction in V_{OC}, albeit to different extents. The maximum and minimum

V_{OC} values for PID faults are selected as representative data, ranging from 0.9495 to 0.8403. This corresponds to a voltage reduction of approximately 5% to 20%, as shown in Table 5. In contrast, the typical V_{OC} values for short-circuit faults are 0.6526 and 0.3193, exhibiting voltage reduction characteristics that follow standard integer multiples. Dust faults cause a decrease in I_{ph} due to irradiation obstruction, while abnormal aging faults result in an increase in R_s. The parameters of the components with the lowest levels of dust and abnormal aging are selected as representative data. The maximum I_{ph} value for the dust component is 0.9675, with a degradation exceeding 4%. The minimum R_s value for the abnormal aging component is 1.1530, reflecting an increase of more than 15%. Both parameters exhibit clear distinctions from the other three types of non-mismatch faults.

For mismatch faults, crack faults can be directly distinguished from the other two types owing to the absence of the step end. While partial shading and hot spot faults exhibit similar linear step characteristics, but their slope differences are primarily reflected in the variation of the R_p parameter. To further quantify this difference, this study selects the mildest and most severe cases of partial shading and hot spot faults as representative samples, characterizing the maximum and minimum values of the R_p parameter, as shown in Table 5. Specifically, the minimum R_p value for partial shading faults is 0.0627, which is approximately 1.61 times the maximum R_p value for hot spot faults. Furthermore, 50 sets of R_p values for both partial shading and hot spot faults are collected, as illustrated in Fig. 14. To enhance the

Table 5
Parameter identification results after normalization.

Fault type	Case	I_{ph}	$I_{01} + I_{02}$	$n_1 + n_2$	R_s	R_p	V_{OC}
PID	Case 1	1.0065	4.11e5	2.5075	3.5e−15	0.3390	0.9495
	Case 2	1.0585	3.86e6	3.6393	4.4e−15	0.0608	0.8403
Dust	Case 1	0.9675	0.3543	0.9502	1.0631	0.9823	1.0029
	Case 2	0.9575	0.3016	0.9424	1.0311	0.4746	1.0000
Short circuit	Case 1	0.9919	4087.85	0.5903	2.49e−14	2.0169	0.3193
	Case 2	1.0075	7.2506	0.7368	0.4818	1.9802	0.6526
Abnormal aging	Case 1	1.0019	0.7838	0.9849	1.2644	0.7061	1.0000
	Case 2	1.0038	3.3333	1.0896	1.1530	0.4412	0.9969
Partial shading	Case 1	0.8949	4.64e−4	0.6529	1.43e−3	0.0627	–
	Case 2	0.8823	8.18e−6	0.5733	0.0403	0.3949	–
Hot spot	Case 1	0.9064	3.53e−11	3.0973	3.01e−15	8.29e−3	–
	Case 2	0.6196	1.28e−12	2.5343	6.19e−15	0.0389	–
Crack	Case 1	0.9612	0.2298	0.9309	0.0148	0.0254	–
	Case 2	0.7536	0.3729	0.9864	4.2e−5	0.0176	–

clarity of comparison, the scale of the figure is transformed using the natural logarithm. The results indicate a significant difference in R_p values between the two types, confirming R_p as an effective fault characteristic parameter.

6. Decision layer: Fault diagnosis and output characteristic representation

6.1. Method for fault classification

Existing fault diagnosis methods typically rely on manually defined thresholds or heuristic rules, leading to subjective uncertainty. To overcome these limitations, we propose a novel Hierarchical Support Vector Machine combined with Decision Tree (HSVM-DT) structure, which is optimized through information entropy principles, as depicted in Fig. 15.

This architecture is specifically designed to handle the multi-class nature of PV fault diagnosis, integrating domain knowledge (such as RMSE patterns and segmented diode model parameters) directly into the classification hierarchy. By incorporating physical insights into the data-driven process, the HSVM-DT achieves improved accuracy, robustness, and interpretability. The diagnostic process proceeds as presented in Fig. 16.

The fault dataset used for testing contains a total of 377 I–V samples, divided into seven classes: partial shading (61), hotspot (84), crack (35), dust (123), short circuit (2), PID (35) and abnormal aging (37). Each sample consists of a 50-point voltage–current curve. The extracted model parameters from all previous steps are fed into an HSVM-DT model, optimized using information entropy principles. Final classification results are visualized through a confusion matrix, as shown in Fig. 17. The confusion matrix reveals an overall fault diagnosis accuracy of 97.35%. While this metric provides a general performance overview, it may obscure the accuracy for specific classes, especially when samples are imbalanced. Thus, a micro-analysis is conducted to evaluate performance metrics for each class individually using the proposed segmented diode model. This analysis is compared with traditional methods [31] and deep learning-based method: the five-point detection linear method (Method 1), the I–V curve segmentation and feature point method (Method 2), the I–V data concavity detection method (Method 3) and the CNN-based fault diagnosis method (Method 4), as summarized in Table 6.

With the advancement of deep learning techniques, a wide range of deep learning-based approaches have been applied to the fault diagnosis of PV modules. Most of these approaches aim to extract diagnostic features directly from I–V curves for classification. Commonly used architectures include CNNs, long short-term memory networks (LSTMs), and Transformer-based models [32–34]. However, I–V curves

Table 6
Comparison of proposed and traditional methods in terms of accuracy.

Accuracy (%)	Proposed method	Method 1	Method 2	Method 3	Method 4
Abnormal aging	100	92.00	88.00	–	97.04
PID	100	92.50	87.50	–	94.74
Dust	95.12	–	–	–	–
Short-circuit	100	100	100	–	–
Partial shading	96.72	100	100	100	97.31
Hot spot	97.62	95.00	93.80	95.83	94.48
Crack	100	91.40	93.30	92.00	97.52
Average	98.49	95.15	93.77	95.94	95.78

do not exhibit strong temporal dependencies, limiting the effectiveness of LSTM-based models. Additionally, Transformer models typically require large-scale labeled datasets for training, but obtaining sufficient fault-related I–V data in real-world PV systems remains challenging. Therefore, CNNs are generally preferred for PV fault diagnosis due to their ability to extract spatially localized features from I–V curves with relatively small datasets. By incorporating encoder–decoder structures and automatic feature extraction, CNN-based methods have demonstrated strong potential to improve diagnostic accuracy. The inability of traditional methods and CNN-based method to detect dust faults is evident from the absence of accuracy values for this category. Furthermore, traditional approaches exhibit lower accuracy in identifying non-mismatch faults compared to the proposed method. In the case of abnormal aging, PID and short-circuit faults, the proposed method yields a perfect accuracy of 100%, while the highest accuracy among traditional methods CNN-based method is 97.04%, 94.74% and 100%. Similarly, for hot spot and crack faults, the proposed method attains accuracy of 97.62% and 100%, outperforming the best-performing traditional methods and CNN-based method, which achieve 95.83% and 97.52%, respectively. Despite its advantages, the proposed method exhibits minor misclassifications, particularly in distinguishing crack faults from partial shading faults, leading to slight limitations. However, the proposed method achieves an average fault diagnosis accuracy of 98.49%, surpassing the four comparative methods, which achieve 95.15%, 93.77%, 95.94% and 95.78%, respectively. This represents an improvement of at least 2.55% in average fault diagnosis accuracy.

Although accuracy is a fundamental metric for assessing computational efficiency, it does not fully capture performance in imbalanced datasets. Precision, which indicates the effectiveness in minimizing false positives, and recall, which quantifies the capacity to reduce false negatives, are crucial for a comprehensive evaluation. To address the trade-off between these metrics, the F1-score, defined as the harmonic mean of precision and recall, provides a balanced and robust assessment. To ensure a more rigorous comparison between the
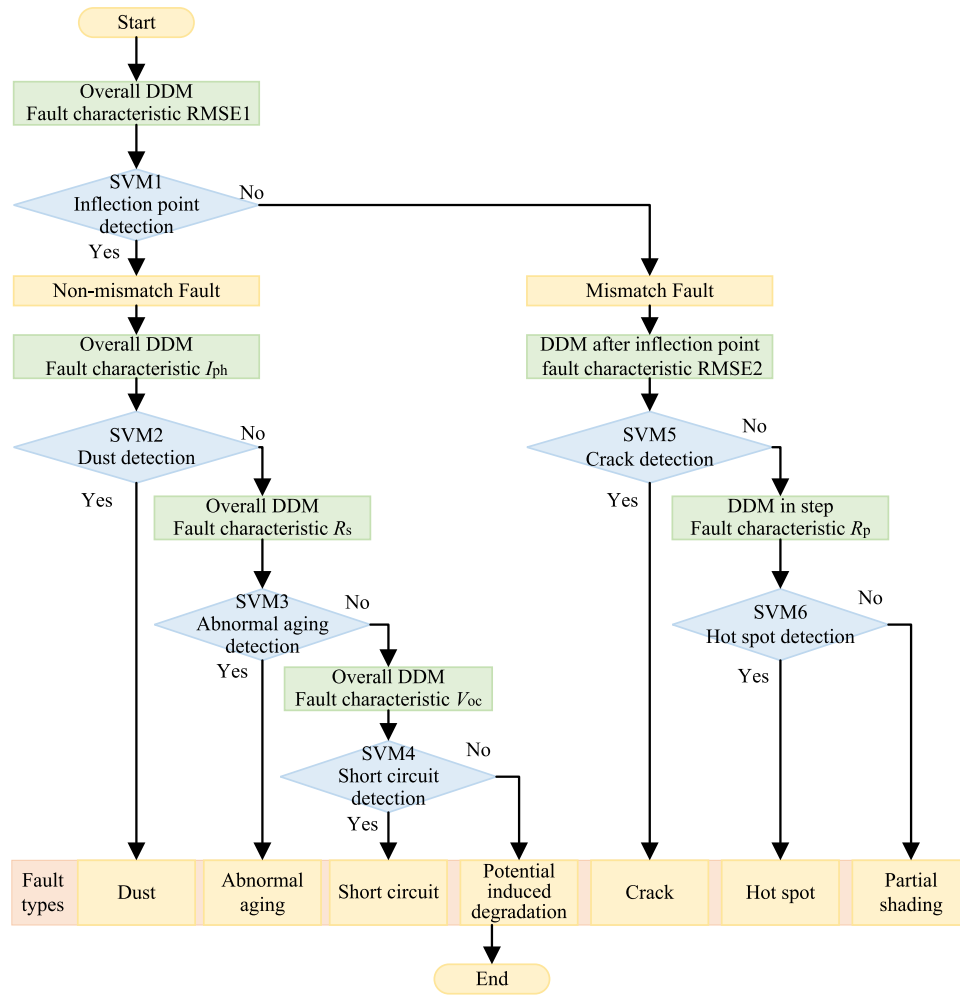
Fig. 15. Structural diagram of HSVM-DT.

Table 7
Comparison of precision, recall, and F1-score for different fault types.

Fault type	Precision (%)		Recall (%)		F1-score (%)	
	Proposed method	Method 4	Proposed method	Method 4	Proposed method	Method 4
Abnormal aging	100	97.04	86.05	97.06	93.02	97.05
PID	100	94.74	100	97.06	100	95.90
Dust	95.12	–	100	–	97.56	–
Short-circuit	100	–	100	–	100	–
Partial shading	96.72	97.31	96.72	97.02	96.72	97.16
Hot spot	97.62	94.48	100	95.00	98.80	94.74
Crack	100	97.52	94.29	93.65	97.22	95.58
Average	98.49	96.21	96.72	95.95	97.62	96.08

proposed and comparative methods, these four methods are applied to typical mismatch faults and compared with the best-performing CNN-based method (Method 4), thereby mitigating the limitations of single-metric evaluation. Although Method 3 achieves a higher overall accuracy than Method 4 by 0.16%, Method 4 is capable of diagnosing a broader range of fault types, making it a more suitable benchmark for comprehensive comparison. The proposed method outperforms Method 4 across all three key metrics according to experimental results in Table 7. Specifically, it achieves an average precision of 98.49%, reflecting a 2.37% improvement over Method 4 (96.21%). Similarly, the proposed method attains an average recall of 96.72%, surpassing Method 4 (95.95%) by 0.80%, indicating enhanced fault detection capability. Furthermore, the average F1-score of the proposed method reaches 97.62%, exceeding Method 4 (96.08%) by 1.60%, highlighting its superior balance between precision and recall. Above results underscore the

robustness and strength of the proposed diagnostic method in precisely identifying mismatch and non-mismatch faults. Although the proposed method achieves a relatively modest improvement of 2.37% in diagnostic accuracy compared to existing approaches, this enhancement becomes highly significant when considered in the context of large-scale PV systems. For example, a 2.5 MW PV plant typically comprises over 22,230 modules [35]. Simultaneously, field studies have reported annual energy losses due to faults as high as 18.9% [36]. Therefore, even a slight increase in diagnostic accuracy can significantly reduce the number of undetected or misclassified faults, thereby minimizing cumulative energy losses across thousands of modules and improving operational efficiency at scale. Furthermore, the proposed framework requires virtually no additional sensors, thereby avoiding increased system complexity or associated costs. It fully utilizes the existing digital infrastructure, including power optimizers and smart inverters that

---

Algorithm 1: Fault Diagnosis Based on HSVM-DT Structure

    Input: I-V curve data of PV module

    Output: Fault classification result

1 **Step 1: Global Fitting and RMSE Calculation**
2 Fit the entire I-V curve using DDM optimized by FG-DE algorithm;
3 Compute global RMSE1;
4 **Step 2: Initial Fault Type Classification (SVM1)**
5 **if** *Inflection point detected in I–V curve* **then**
6   Fault type ← Mismatch;
7 **else**
8   Fault type ← Non-mismatch;

9 **Step 3: Non-Mismatch Fault Classification (if applicable)**
10 **if** *Fault type == Non-mismatch* **then**
11   **if** *SVM2 detects decrease in I_{ph}* **then**
12   Fault ← Dust;
13   **else if** *SVM3 detects increase in R_s* **then**
14   Fault ← Abnormal aging;
15   **else if** *SVM4 detects decrease in V_{oc}* **then**
16   **if** *Magnitude of drop is large* **then**
17   Fault ← Short-circuit;
18   **else**
19   Fault ← PID;

20 **Step 4: Mismatch Fault Classification (if applicable)**
21 **if** *Fault type == Mismatch* **then**
22   Perform segmented fitting after inflection point to obtain RMSE2;
23   **if** *SVM5 detects slope change* **then**
24   Fault ← Crack;
25   **else if** *SVM6 evaluates R_p variation* **then**
26   **if** *R_p is high* **then**
27   Fault ← Partial shading;
28   **else**
29   Fault ← Hot spot;

30 **Step 5: Integrated Decision**
31 Combine all SVM results using HSVM-DT for final classification;

Fig. 16. Pseudocode of HSVM-DT.

Fig. 17. Confusion matrix of sample test results.

Table 8
Comparison of proposed and traditional methods in terms of RMSE and R^2.

Fault type	RMSE (A)		R^2	
	Proposed method	Traditional method	Proposed method	Traditional method
Crack	0.0863	0.3223	0.9905	0.9288
Partial shading and hot spot	0.0247	0.0488	0.9525	0.8236
Average	0.0555	0.1856	0.9715	0.8762

To assess the edge computing capabilities of the proposed method, both the proposed diagnostic framework and a conventional CNN-based fault diagnosis approach (Method 4) were deployed on a Raspberry Pi 4 embedded platform. This platform is equipped with a Broadcom BCM2711 quad-core Cortex-A72 (ARM v8) 64-bit SoC running at 1.5 GHz and 8 GB RAM. A series of experiments were conducted to systematically compare their computational performance across several key metrics, including training time, inference latency, model size, Resident Set Size (RSS), CPU utilization, and device temperature increase. Notably, RSS serves as a critical indicator of the physical memory footprint, reflecting the portion of memory actually residing in RAM and thus offering a direct measure of real-time memory usage. The experimental results demonstrate that Method 4 required a training time of 431.86 s and an average inference time of 1.02 s.
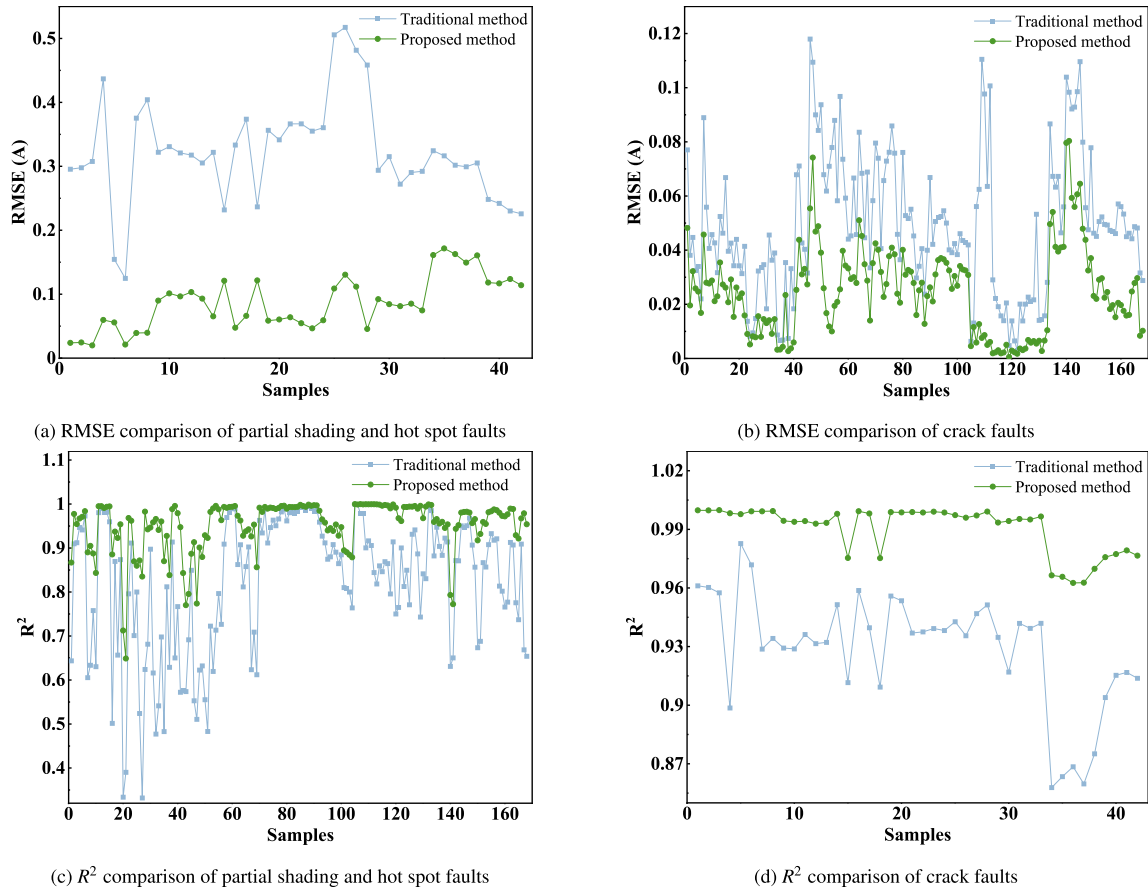
feature integrated I–V curve scanning functionality, which are increasingly adopted as standard components in modern PV installations. This inherent compatibility ensures minimal integration overhead while preserving high diagnostic effectiveness. Even for systems employing solar cell testers, the process typically completes within seconds, resulting in negligible power loss during detection.

(a) RMSE comparison of partial shading and hot spot faults

(b) RMSE comparison of crack faults

(c) R^2 comparison of partial shading and hot spot faults

(d) R^2 comparison of crack faults

Fig. 18. Output characterization comparison of proposed method and traditional method.

Its model size was 4266.02 KB, with an RSS of 263.11 MB. Furthermore, it led to a device temperature increase of 24.90 °C and an average CPU utilization of 97.02%. In contrast, the proposed method significantly outperformed Method 4 across all evaluated dimensions. It achieved a reduced training time of only 23.19 s, corresponding to a speedup of 18.62, and a near-instantaneous average inference time of 1.98e−3 s, representing a 515.15 times acceleration. Moreover, the proposed model's size was merely 0.04 KB, indicating a dramatic reduction by a factor of 106,650. Its RSS was also substantially lower at 126.76 MB, corresponding to a 51.83% reduction compared to Method 4. Additionally, the proposed framework resulted in a minimal device temperature increase of only 3.40 °C, representing an 86.3% decrease, and a significantly lower average CPU utilization of 44.12%, indicating a 52.90% reduction in computational load. These findings collectively highlight the superior computational efficiency, memory economy, and thermal stability of the proposed framework, underscoring its suitability for real-time fault diagnosis in resource-constrained edge computing environments.

However, certain limitations should be acknowledged. Although the proposed approach significantly improves diagnostic accuracy, it remains dependent on the availability and quality of I–V curve data, the acquisition of which can still be time-consuming or resource-intensive in certain PV system configurations. Furthermore, while the model exhibits strong performance in identifying existing faults, its ability to provide predictive warnings prior to fault manifestation remains limited. In addition, the current implementation primarily addresses diagnostic inference and lacks integration with automated control or scheduling systems. While the decision layer has been conceptually

formulated, it does not yet form a closed-loop with physical-layer actuators or supervisory systems. This limits the system's ability to autonomously trigger maintenance actions or reconfigure PV array operations based on diagnosed faults. Future work will focus on developing feedback-enabled architectures that integrate additional time-series features and environmental parameters to enhance the predictive diagnostic capabilities of the proposed DT framework.

6.2. Output characterization of PV modules

Characterizing the output characteristics of PV modules is critical for performance assessment and lifespan prediction. Traditional methods use linear and polynomial functions to represent mismatch faults, capturing only a limited phase of the I–V relationship. This approach fails to describe the full range of mismatch fault behaviors and lacks physical interpretability and theoretical support due to the purely mathematical model. In this study, the improved segmented diode model, based on the FG-DE algorithm, is deployed to fit the output characterization under mismatch faults. A comparison is conducted with traditional linear and polynomial-based characterization methods. RMSE and R^2 are chosen as the feature vector to assess the accuracy of these output characterizations. Over 200 sets of typical I–V data are verified, with experimental results illustrated in Fig. 18. As shown in Table 8, for the three-segment diode model including partial shading and hot spot faults, traditional linear method obtains an average RMSE of 0.0488 A, while the improved model achieves 0.0247 A, representing a 49.39% error reduction. For the two-segment diode model including crack faults, traditional polynomial fitting method produces an RMSE

of 0.3223 A, whereas the improved model achieves 0.0863 A, marking a 73.22% reduction. In summary, the proposed method achieves an average RMSE of 0.0555 A, representing a 70.10% reduction compared to 0.1856 A obtained by the traditional method. Although RMSE effectively quantifies absolute prediction errors, it does not fully reflect the fitting accuracy. To address this, the coefficient of determination R^2 is further analyzed. Similarly, the proposed segmented diode model achieves an average R^2 of 0.9715 for mismatch faults, representing a 10.88% reduction compared to 0.8762 obtained by the traditional method.

These results demonstrate that the improved segmented diode model offers superior accuracy and enhanced physical relevance in output characterization compared to traditional approaches.

7. Conclusion

This paper has presented a novel four-layer DT framework for PV modules, integrating a segmented diode model, a hybrid FG-DE optimization algorithm, and a HSVM-DT classifier. This approach enables dynamic adaptation to real-time operating conditions, precise fault diagnosis, and accurate output characterization. A key advancement is the segmented diode model, which for the first time captures inflection-point-induced non-linearity caused by mismatch faults in PV modules, enabling segment-wise parameter identification and high-fidelity I–V curve reconstruction under fault conditions. Combined with the FG-DE optimization and multi-source data fusion, the proposed DT framework has achieved a fault diagnosis average accuracy of 98.49% and a 70.10% reduction in modeling error, outperforming existing methods in both accuracy and robustness. In future studies, the DT framework will be extended to address complex multi-fault scenarios, incorporating fault propagation modeling, and expanding the architecture to include other critical PV system components such as DC–DC converters and inverters. Efforts will also focus on enhancing predictive capabilities through the integration of time-series and environmental data, and on establishing closed-loop architectures with real-time control systems to enable autonomous fault response and system reconfiguration under operational conditions.

CRediT authorship contribution statement

Yihan Chen: Writing – original draft, Methodology. **Mingyao Ma:** Funding acquisition, Conceptualization. **Wenting Ma:** Software, Formal analysis. **Xilian Zhou:** Writing – review & editing, Validation. **Rui Zhang:** Formal analysis, Data curation. **Zhenyu Fang:** Investigation, Data curation.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

References

[1] International Energy Agency, Renewables 2024, 2024, (accessed 15 April 2025) URL https://www.iea.org/reports/renewables-2024.

[2] P. Zhang, W. Li, S. Li, Y. Wang, W. Xiao, Reliability assessment of photovoltaic power systems: Review of current status and future perspectives, Appl. Energy 104 (2013) 822–833, http://dx.doi.org/10.1016/j.apenergy.2012.12.010.

[3] H. Yang, K. Ding, X. Chen, M. Jiang, Z. Yang, J. Zhang, R. Gao, Fast simulation modeling and multiple-ps fault diagnosis of the pv array based on i–v curve conversion, Energy Convers. Manage. 300 (2024) 117965, http://dx.doi.org/10.1016/j.enconman.2023.117965.

[4] L. Wan, L. Zhao, W. Xu, F. Guo, X. Jiang, Dust deposition on the photovoltaic panel: A comprehensive survey on mechanisms, effects, mathematical modeling, cleaning methods, and monitoring systems, Sol. Energy 268 (2024) 112300, http://dx.doi.org/10.1016/j.solener.2023.112300.

[5] W. Ma, M. Ma, Z. Zhang, J. Ma, R. Zhang, J. Wang, Anomaly detection of mountain photovoltaic power plant based on spectral clustering, IEEE J. Photovoltaics 13 (4) (2023) 621–631, http://dx.doi.org/10.1109/JPHOTOV.2023.3267222.

[6] S. Lidaighbi, M. Elyaqouti, D. Ben Hmamou, D. Saadaoui, K. Assalaou, E. Arjdal, A new hybrid method to estimate the single-diode model parameters of solar photovoltaic panel, Energy Convers. Manag.: X 15 (2022) 100234, http://dx.doi.org/10.1016/j.ecmx.2022.100234.

[7] K. El Ainaoui, M. Zaimi, E. Assaid, Innovative approaches to extract double-diode model physical parameters of a pv module serving outdoors under real-world conditions, Energy Convers. Manage. 292 (2023) 117365, http://dx.doi.org/10.1016/j.enconman.2023.117365.

[8] K. Hong Min, H. eun Song, M. Gu Kang, S. Hee Lee, S. Park, Double-diode model carrier lifetime-based internal recombination parameter analysis and efficiency prediction of crystalline si solar cells, Sol. Energy 277 (2024) 112697, http://dx.doi.org/10.1016/j.solener.2024.112697.

[9] R. Mohamed, M. Abdel-Basset, K.M. Sallam, I.M. Hezam, A.M. Alshamrani, I.A. Hameed, Novel hybrid kepler optimization algorithm for parameter estimation of photovoltaic modules, Sci. Rep. 14 (1) (2024) 3453, http://dx.doi.org/10.1038/s41598-024-52416-6.

[10] Q. Liu, C. Zhang, Z. Li, T. Peng, Z. Zhang, D. Du, M.S. Nazir, Multi-strategy adaptive guidance differential evolution algorithm using fitness-distance balance and opposition-based learning for constrained global optimization of photovoltaic cells and modules, Appl. Energy 353 (2024) 122032, http://dx.doi.org/10.1016/j.apenergy.2023.122032.

[11] A.F. Amiri, H. Oudira, A. Chouder, S. Kichou, Faults detection and diagnosis of pv systems based on machine learning approach using random forest classifier, Energy Convers. Manage. 301 (2024) 118076, http://dx.doi.org/10.1016/j.enconman.2024.118076.

[12] J. Qu, Q. Sun, Z. Qian, L. Wei, H. Zareipour, Fault diagnosis for pv arrays considering dust impact based on transformed graphical features of characteristic curves and convolutional neural network with cbam modules, Appl. Energy 355 (2024) 122252, http://dx.doi.org/10.1016/j.apenergy.2023.122252.

[13] H. Zhu, L. Lu, J. Yao, S. Dai, Y. Hu, Fault diagnosis approach for photovoltaic arrays based on unsupervised sample clustering and probabilistic neural network model, Sol. Energy 176 (2018) 395–405, http://dx.doi.org/10.1016/j.solener.2018.10.054.

[14] Y. Liu, K. Ding, J. Zhang, Y. Li, Z. Yang, W. Zheng, X. Chen, Fault diagnosis approach for photovoltaic array based on the stacked auto-encoder and clustering with i-v curves, Energy Convers. Manage. 245 (2021) 114603, http://dx.doi.org/10.1016/j.enconman.2021.114603.

[15] T. Khatib, A. Ghareeb, M. Tamimi, M. Jaber, S. Jaradat, A new offline method for extracting I-V characteristic curve for photovoltaic modules using artificial neural networks, Sol. Energy 173 (2018) 462–469, http://dx.doi.org/10.1016/j.solener.2018.07.092.

[16] R. Castro, Data-driven PV modules modelling: Comparison between equivalent electric circuit and artificial intelligence based models, Sustain. Energy Technol. Assessments 30 (2018) 230–238, http://dx.doi.org/10.1016/j.seta.2018.10.011.

[17] H.M. Khalid, Z. Rafique, S.M. Muyeen, A. Raqeeb, Z. Said, R. Saidur, K. Sopian, Dust accumulation and aggregation on PV panels: An integrated survey on impacts, mathematical models, cleaning mechanisms, and possible sustainable solution, Sol. Energy 251 (2023) 261–285, http://dx.doi.org/10.1016/j.solener.2023.01.010.

[18] S.W. Ko, Y.C. Ju, H.M. Hwang, J.H. So, Y.-S. Jung, H.-J. Song, H. e. Song, S.-H. Kim, G.H. Kang, Electric and thermal characteristics of photovoltaic modules under partial shading and with a damaged bypass diode, Energy 128 (2017) 232–243, http://dx.doi.org/10.1016/j.energy.2017.04.030.

[19] M. Liu, S. Fang, H. Dong, C. Xu, Review of digital twin about concepts, technologies, and industrial applications, J. Manuf. Syst. 58 (2021) 346–361, http://dx.doi.org/10.1016/j.jmsy.2020.06.017.

[20] M. Kolahi, S. Esmailifar, A. Moradi Sizkouhi, M. Aghaei, Digital-pv: A digital twin-based platform for autonomous aerial monitoring of large-scale photovoltaic power plants, Energy Convers. Manage. 321 (2024) 118963, http://dx.doi.org/10.1016/j.enconman.2024.118963.

[21] C. Ghenai, L.A. Husein, M. Al Nahlawi, A.K. Hamid, M. Bettayeb, Recent trends of digital twin technologies in the energy sector: A comprehensive review, Sustain. Energy Technol. Assessments 54 (2022) 102837, http://dx.doi.org/10.1016/j.seta.2022.102837.

[22] A.I. Khalyasmaa, A.I. Stepanova, S.A. Eroshenko, P.V. Matrenin, Review of the Digital Twin Technology Applications for Electrical Equipment Life-cycle Management, Math. 11 (6) (2023) 1315, http://dx.doi.org/10.3390/math11061315.

[23] P. Jain, J. Poon, J.P. Singh, C. Spanos, S.R. Sanders, S.K. Panda, A digital twin approach for fault diagnosis in distributed photovoltaic systems, IEEE Trans. Power Electron. 35 (1) (2020) 940–956, http://dx.doi.org/10.1109/TPEL.2019.2911594.

[24] D.E. Guzman Razo, B. Müller, H. Madsen, C. Wittwer, A genetic algorithm approach as a self-learning and optimization tool for PV Power Simulation and digital twinning, Energies 13 (24) (2020) 6712, http://dx.doi.org/10.3390/en13246712.

[25] Y.-Y. Hong, R.A. Pula, Diagnosis of photovoltaic faults using digital twin and PSO-optimized shifted window transformer, Appl. Soft Comput. 150 (2024) 111092, http://dx.doi.org/10.1016/j.asoc.2023.111092.

[26] T. Yalçin, P. Paradell Solà, P. Stefanidou-Voziki, J.L. Domínguez-García, T. Demirdelen, Exploiting digitalization of solar PV plants using machine learning: Digital twin concept for operation, Energies 16 (13) (2023) 5044, http://dx.doi.org/10.3390/en16135044.

[27] G. Zhang, X. Wang, Digital twin modeling for photovoltaic panels based on hybrid neural network, in: 2021 IEEE 1st International Conference on Digital Twins and Parallel Intelligence, DTPI, 2021, pp. 90–93, http://dx.doi.org/10.1109/DTPI52967.2021.9540210.

[28] Z. He, H.C. Chen, S. Shan, Y. Hu, K. Zhang, H. Wei, Shading type and severity diagnosis in photovoltaic systems via i-v curve imaging and two-stream deep neural network, Energy Convers. Manage. 324 (2025) 119311, http://dx.doi.org/10.1016/j.enconman.2024.119311.

[29] Z. He, P. Chu, C. Li, K. Zhang, H. Wei, Y. Hu, Compound fault diagnosis for photovoltaic arrays based on multi-label learning considering multiple faults coupling, Energy Convers. Manage. 279 (2023) 116742, http://dx.doi.org/10.1016/j.enconman.2023.116742.

[30] P. Li, Y. Luo, X. Xia, X. Gao, R. Chang, Z. Li, J. Zheng, W. Shi, Z. Liao, Factors and quantitative impact on electrical yield in fishery complementary photovoltaic power plant under different cloud cover conditions, Energy 309 (2024) 133079, http://dx.doi.org/10.1016/j.energy.2024.133079.

[31] Z. Zhang, M. Ma, H. Wang, H. Wang, W. Ma, X. Zhang, A fault diagnosis method for photovoltaic module current mismatch based on numerical analysis and statistics, Sol. Energy 225 (2021) 221–236, http://dx.doi.org/10.1016/j.solener.2021.07.037.

[32] Y.-Y. Hong, R.A. Pula, Detection and classification of faults in photovoltaic arrays using a 3d convolutional neural network, Energy 246 (2022) 123391, http://dx.doi.org/10.1016/j.energy.2022.123391.

[33] F. Cui, Y. Tu, W. Gao, A photovoltaic system fault identification method based on improved deep residual shrinkage networks, Energies 15 (11) (2022) 00, http://dx.doi.org/10.3390/en15113961, URL https://www.mdpi.com/1996-1073/15/11/3961.

[34] Q. Chang, Y. Luo, Q. Wang, B. Ren, N. Zhou, Fault diagnosis method for photovoltaic string based on global i-v curve feature extraction using swin transformer, Trans. China Electrotech. Soc. 250431, http://dx.doi.org/10.19595/j.cnki.1000-6753.tces.242073.

[35] M. Chandel, G. Agrawal, S. Mathur, A. Mathur, Techno-economic analysis of solar photovoltaic power plant for garment zone of jaipur city, Case Stud. Therm. Eng. 2 (2014) 1–7, http://dx.doi.org/10.1016/j.csite.2013.10.002.

[36] M. Aghaei, A. Fairbrother, A. Gok, S. Ahmad, S. Kazim, K. Lobato, G. Oreski, A. Reinders, J. Schmitz, M. Theelen, P. Yilmaz, J. Kettle, Review of degradation and failure phenomena in photovoltaic modules, Renew. Sustain. Energy Rev. 159 (2022) 112160, http://dx.doi.org/10.1016/j.rser.2022.112160.