

Article

Fault Detection and Classification of CIGS Thin-Film PV Modules Using an Adaptive Neuro-Fuzzy Inference Scheme

Reham A. Eltuhamy ^{1,2}, Mohamed Rady ^{3,*} , Eydhah Almatrafi ³ , Haitham A. Mahmoud ⁴ 
and Khaled H. Ibrahim ⁵

¹ Mechanical Engineering Department, Faculty of Engineering, Helwan University, Cairo 11795, Egypt

² Mechanical Engineering Department, Ahran Canadian University, Cairo 12451, Egypt

³ Mechanical Engineering Department, Faculty of Engineering at Rabigh, King Abdulaziz University, Jeddah 21589, Saudi Arabia

⁴ Industrial Engineering Department, College of Engineering, King Saud University, Riyadh 11421, Saudi Arabia

⁵ Electrical Power Department, Faculty of Engineering, Fayoum University, El-Fayoum 63514, Egypt

* Correspondence: maradhi@kau.edu.sa

Abstract: The use of artificial intelligence to automate PV module fault detection, diagnosis, and classification processes has gained interest for PV solar plants maintenance planning and reduction in expensive inspection and shutdown periods. The present article reports on the development of an adaptive neuro-fuzzy inference system (ANFIS) for PV fault classification based on statistical and mathematical features extracted from outdoor infrared thermography (IRT) and I-V measurements of thin-film PV modules. The selection of the membership function is shown to be essential to obtain a high classifier performance. Principal components analysis (PCA) is used to reduce the dimensions to speed up the classification process. For each type of fault, effective features that are highly correlated to the PV module's operating power ratio are identified. Evaluation of the proposed methodology, based on datasets gathered from a typical PV plant, reveals that features extraction methods based on mathematical parameters and I-V measurements provide a 100% classification accuracy. On the other hand, features extraction based on statistical factors provides 83.33% accuracy. A novel technique is proposed for developing a correlation matrix between the PV operating power ratio and the effective features extracted online from infrared thermal images. This eliminates the need for offline I-V measurements to estimate the operating power ratio of PV modules.

Keywords: CIGS thin film; PV modules; adaptive neuro-fuzzy inference system; operating power ratio



Citation: Eltuhamy, R.A.; Rady, M.; Almatrafi, E.; Mahmoud, H.A.; Ibrahim, K.H. Fault Detection and Classification of CIGS Thin-Film PV Modules Using an Adaptive Neuro-Fuzzy Inference Scheme. *Sensors* **2023**, *23*, 1280. <https://doi.org/10.3390/s23031280>

Academic Editors: Yongliang Xie, Liang Gong and Chang'e Cai

Received: 25 December 2022

Revised: 18 January 2023

Accepted: 18 January 2023

Published: 22 January 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

A crucial technology for a sustainable energy supply is the adoption of PV modules. According to recent statistics, the reliance on PV modules' capacity has increased globally from 17 GW in 2010 to 139 GW in 2020 and has reached 760 GW at the end of 2020 [1]. Several techniques have been proposed for fault detection and diagnosis in PV modules; examples of these techniques include visual examination, infrared thermography (IRT), electroluminescence (EL), photoluminescence (PL), and measurements of static characteristics [2]. Based on the previously mentioned PV fault detection techniques, different types of PV module defects have been addressed, such as cracks, delamination, burn marks, potential-induced degradation (PID), soiling, snail trails, hotspots, faulty interconnections, back sheet defects, corrosion, and shunts [3–5]. The continuous, safe, reliable, and effective functioning of PV systems depends on the development of automatic diagnosis and classification techniques for PV monitoring systems. The huge number of solar modules used in large-scale PV facilities has made it difficult to detect and classify faults [6–9]. The automation of detection, diagnoses, and classification of PV system faults, using artificial

intelligence (AI) approaches, has attracted a great deal of research interest. Many AI techniques have been suggested in the literature, such as particle swarm optimization [10], ant colony optimization [11], support vector machine [12], k-nearest neighbors [13], linear regression [14], decision tree [15], naive Bayes (NB) [16], neural network systems [17], and fuzzy logic [18]. Additionally, hybrid systems, such as ANFIS [19] and artificial neural network with genetic algorithm (ANN-GA) [20], have also been examined.

Previous research demonstrates a considerable interest in the development of intelligent systems for automatic classification of faults in PV modules. Table 1 provides a general review of the application of several AI approaches in fault detection and diagnosis in PV plants. The previous work is categorized, based on the input datasets, into thermography techniques (Table 1a), PV modules electrical I-V characteristics (Table 1b), PV modules electrical I-V characteristics and environmental conditions (Table 1c), and using thermography and I-V measurements (Table 1d). The input parameters, fault types, classification accuracy, and limitations are summarized for each technique. The previously described methods, however, are restricted to specific fault types and/or a single fault per module. In some cases, the proposed techniques are capable of defining only a faulty module without being capable of detecting the fault type. Additionally, I-V measurements are often carried out when the PV module is offline; this introduces a significant difficulty for automation of fault detection and classification processes. The use of IRT offers the advantage of possible online operation of the diagnosis process. However, to the authors' knowledge, there is no research work on correlating the infrared (IR) images of faulty modules with their I-V measurements. Such a correlation is very useful in developing an online-detection classification system that is useful for maintenance planning. The importance of accurately classifying faults lies in determining the required maintenance action and calculating the remaining useful life (RUL) of the PV module. Moreover, previous studies in classifying faults have focused on polycrystalline and monocrystalline PV technologies, and barely any of this research has dealt with PV modules of thin-film technology.

Table 1. An overall review of using AI techniques for fault detection and diagnosis in PV plants.

Methodology	Reference	Fault Classification	Accuracy of Classification	Remarks
(a) using thermography techniques.				
Texture feature extraction (TFE) and support vector machine (SVM)	[12]	Cracks, hot spots due to shading and soiling. Categorize solar modules into defective and non-defective.	97%	
K-nearest neighbor (KNN)	[21]		80.3%	- Limited to a single type of fault per PV module. - Only categorize modules into defective and non-defective
Support vector machine (SVM)			56.8%	
Neural network			92.8	
Support vector machine (SVM)	[22]	Categorize solar modules into defective and non-defective.	91.2%	- No correlation with module output power
Deep-learning convolutional neural network (CNN)	[22]		89.5%	
n Bayes: a binary class density-based classifier	[23]		98.4%	
The automated edge detection technique	[24,25]	Defective solder junctions, short circuits, and bypassed substrings.	Not reported	

Table 1. Cont.

Methodology	Reference	Fault Classification	Accuracy of Classification	Remarks
(a) using thermography techniques.				
Deep learning neural network	[26]	Cracks, shadowing, diode, soiling, hotspots, and offline module.	Classify 12 anomaly types with an average of 86%	
(b) with input datasets from PV modules electrical I-V characteristics.				
Multi-class adaptive neuro-fuzzy classifier	[19]	Partial shading, increased series resistance, bypass diode short-circuited, bypass diode impedance, PV module short-circuited.	65–100% depending on fault type	- I-V datasets obtained using a real time PV emulator. Datasets from real systems are not considered.
Principal component analysis (PCA)	[27]	Shading faults.	97%	- Real systems configurations are not considered.
AI nonlinear autoregressive exogenous neural network (NARX)	[28]	Open and short-circuit degradation, faulty MPPT, partial shading (PS).	98.2%	- Applicable for small systems.
Multilayer neural network with a scaled conjugate gradient algorithm (SCG)	[29]	Short circuits, aging, shading faults, and bypass diode faults.	99.6%	- Consider polycrystalline and thin-film PV technologies.
Convolutional neural networks (CNN)	[30]	Partial Shading (PS), high impedance, low location mismatch, maximum power point tracking (MPPT).	73.53%	- 2D scalograms generated from I-V characteristics using continuous wavelet transform. (CWT) - Low classification accuracy.
Multiclass adaptive boosting (AdaBoost) algorithm, using multiclass exponential (SAMME) loss function based on the classification and regression tree (CART)	[31]	Short-circuit faults (SCF), partial shading with the bypass-diode on (PSBO), partial shading with the bypass-diode reversed (PSBR), and abnormal aging faults (AAF).	99.4%	- The results consider only faulty modules. - Time-consuming, suitable for a small number of modules.
Radial basis function (RBF) kernel extreme learning machine (ELM) optimized by simulated annealing algorithm,	[32]	Short circuits, shading faults, and aging.	Shadows 91.55% Short circuits 93.64% Aging 90.91%	Need real outdoor experiments.

Table 1. Cont.

Methodology	Reference	Fault Classification	Accuracy of Classification	Remarks
(b) with input datasets from PV modules electrical I-V characteristics.				
Artificial neural network	[33]	Partial shading	Not reported	A single type of fault.
Multiclass adaptive neuro-fuzzy classifier (MC-NFC) and ANN	[19]	Partial shading, high series resistance, bypass diode impedance and short circuits.	Not reported	The MC-NFC outperforms the ANN-classifier.
(c) with input datasets from PV modules electrical I-V characteristics and environmental conditions.				
Backward propagation NN optimized by genetic algorithm	[20]	Short circuits, local material aging, shading.	78% for short circuits, 97% for aging, 100% shadows	<ul style="list-style-type: none"> - Expensive real time PV modules output current, voltage, irradiation, ambient temperature monitoring system for each module. - Limited number of faults.
Neuro-fuzzy and simulation	[34]	Upper and lower earth faults, diode short-circuit faults, partial shading.	Not reported	Limited number of PV module circuit faults.
Cursive linear model and an ANN	[35]	Short circuits, open circuits, partial shading, and degradation.	92.64%	Limited number of PV module circuit faults.
ANNs	[36]	Disconnected modules.	97%	<ul style="list-style-type: none"> - Using solar radiation and PV measured output power. - Limited to one type of fault.
(d) with input datasets from thermography analysis and PV modules electrical I-V characteristics.				
Statistical features extraction and electrical measurements characteristics	[3]	Cracks, delamination, burn marks, PID, soiling, and open strings.	Not reported	Applied for CIGS PV modules.
Fuzzy inference system (FIS) using Mamdani-type fuzzy controller	[37]	Identify the six main types of hotspots that influence PV modules.	96.7%	Inability to detect hot spots when there is a lot of partial shading.
Novel feature extraction based on mathematical parameters	[38]	Cracks, delamination, burn marks, PID, soiling, and open strings.	Not reported	Detect all types of CIGS thin-film PV modules, detect modules with multi-faults.

To this end, this paper proposes using an ANFIS for automation of fault classification of copper indium gallium selenide (CIGS) thin-film PV modules. The proposed ANFIS approach is based on statistical and mathematical features of outdoor IRT and I-V measurements of PV modules. A correlation matrix is proposed to be developed between the operating power ratio (P_r) and the effective features extracted from thermal images. The main scientific contributions made in this work are summarized as follows: (1) an automatic ANFIS fault classification technique is developed for the detection and classification of PV module faults; (2) new effective features for each type of fault are identified that are highly correlated with the P_r of a PV module; (3) a correlation matrix is developed to address

correlations between the P_r and the proposed effective features extracted from thermal images. Thus, the need for offline I-V measurements is eliminated, and (4) the effectiveness of the proposed scheme is evaluated using collected datasets from a typical plant with PV modules with multiple faults.

2. Methodology

The procedure implemented in developing the present ANFIS scheme for fault classification using IRT and I-V measurements is shown in Figure 1. The datasets are prepared by acquisition of input data of PV module thermal images and the corresponding I-V curves. Thermal images are processed using image filtering, segmentation, and reshaping. Image filtering and panels reshaping are performed using pixel-shifting techniques. In our methodology, a moving average filter and multidimensional filter are used. The geometric transformations are used to correct distortions caused by viewing geometry using pixel-shifting techniques, and reshape that panel into a rectangular shape. The destination image is filled by regular scan lines, taking the values from the source image by bi-cubic interpolation. This is achieved by applying geometric transformations to images. After getting the rectangular shape, the cropping process is easily completed to remove unwanted segments. The next step uses feature extraction methods. Statistical and mathematical feature extraction techniques of IR images and I-V curves are employed. Once feature extraction techniques have been deployed, ANFIS is utilized to classify CIGS module failures. ANFIS is learnt using the training dataset, just like artificial neural networks. Faults are classified according to their types (A, B, C, D, E, F, and G). The most effective features, correlated with the power ratio, are identified to develop a correlation matrix between these effective features and PV module operating power ratio. A regression model is applied, utilizing the correlation matrix, to predict the PV module's operating power ratio for each type of fault. The value of the operating power ratio of each module is then used for evaluating the overall plant power and for management of the power plant's operation and maintenance.

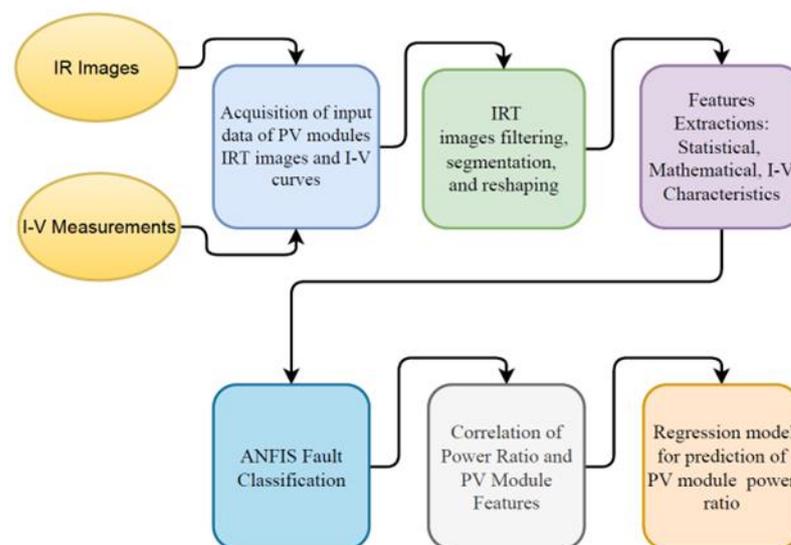


Figure 1. Procedure of development of ANFIS scheme for automation of fault detection and classification.

3. Collection of Input Datasets

The datasets for this investigation were gathered from a PV power plant located in Bani Mazar, Egypt. The plant consists of 84 TianWei solar films comprised of TW-SF-W100 CIGS thin-film modules. I-V curves were obtained using a variable resistor and a potentiometer. A total of 84 IRT images were captured with a Fluke Ti3 IR camera. The IEC62446 [39] requirements for PV thermography testing were followed. Before the tests, the junction boxes and all electrical connections were inspected. The solar irradiance (E)

was around 900 W/m^2 during the testing, and the ambient temperature was 19°C , with a low wind speed of about 18 km/hr . Cloud shading was avoided during filming to avoid the appearance of warmer reflected areas and the resulting effect of fault misunderstanding [39]. The emissivity of PV panels was also considered, and all shots were taken from a 50 to 60° angle. The plant's defective modules were previously identified; most of the modules that were tested contained multiple faults.

The current study focuses on the most frequent defects that can occur in PV modules. Table 2 shows how the faults are grouped. For each type of fault, an IRT image and I-V measurements were acquired for each of the 84 modules. IRT imaging and I-V measurements were conducted three times for each module to reduce measurement uncertainty. In [3,38], the present authors have published a detailed set of IR pictures and I-V curves. Figure 2 shows an example of IRT image and I-V curve for a PV module with Type-B fault.

Table 2. Categories of CIGS PV module faults.

Category/Type	Description
A	Soiling
B	Cracking and soiling
C	Cracks, burn marks, and soiling
D	Potential-induced degradation (PID)
E	PID and cracks
F	PID, cracks, and delamination
G	Open strings (HM)
H	Dead modules

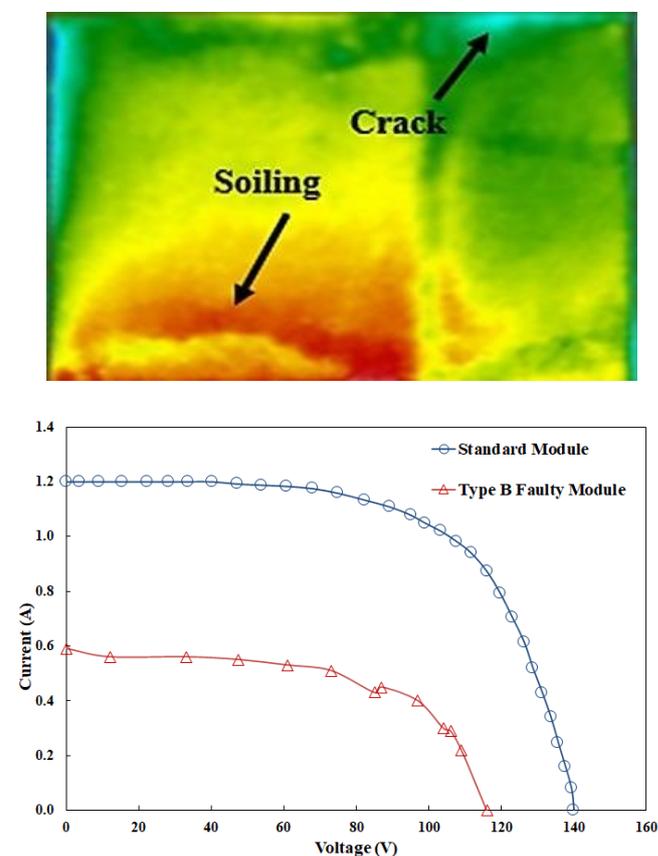


Figure 2. IRT image and I-V curve for a PV module with Type-B fault.

4. Features Extraction Techniques

The feature extraction techniques applied in this study include statistical feature extraction, mathematical parameters-based feature extraction, and electrical measurements-based feature extraction. The parameters of IR image features are summarized in the following subsections (Sections 4.1–4.3). The process of features extraction was carried out using the MATLAB platform. Detailed description of these features is reported in [3,38]. The ratio of operating voltage, current, power, fill factor, and efficiency are all features of the I-V characteristic curve. They are derived by comparing the PV module operational value to its healthy equivalent. The investigated module is denoted by the subscript o, while the healthy module is denoted by the subscript h. Because the healthy values of these parameters alter with solar irradiation and electrical demand, they might be considered the maximum operational voltage, current, and power of all PV modules in use. The mean, standard deviation, skewness, and kurtosis are all statistical properties of an IRT image. The capability of employing I-V measurements and statistical aspects of IR images for CIGS module fault classification is discussed in [3]. Statistical features and I-V measurements are used to generate a generic classification matrix for fault identification and diagnosis. Statistical features, on the other hand, have demonstrated a limited ability to distinguish and detect problems in defective modules with many types of defects occurring at the same time. Faulty modules of the types (A, B), (B, E), (D, E), (E, F), (E, G), (E, H), (E, I), and (G, H) are examples.

4.1. Statistical Parameters of IR images

The following equations summarize the statistical parameters of IR images features.

$$\bar{\chi} = \frac{1}{N} \cdot \sum_{i=1}^N \chi_i \quad (1)$$

$$\sigma = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (\chi_i - \bar{\chi})^2} \quad (2)$$

$$\gamma_1 = \frac{1}{N} \sum_{i=1}^N \left[\frac{\chi_i - \bar{\chi}}{\sigma} \right]^3 \quad (3)$$

$$\gamma_2 = \frac{1}{N} \sum_{i=1}^N \left[\frac{\chi_i - \bar{\chi}}{\sigma} \right]^4 - 3 \quad (4)$$

where Equations (1)–(4) represent the mean, the standard deviation, the skewness, and the kurtosis of the images' feature parameters. N is the total number of pixels, χ_i is the pixel value, and σ is standard deviation of the image.

4.2. Mathematical Parameters of IR Images

The mathematical parameters of IR include the peak-to-peak value (PP), flatness density measure (FDM), flatness continuity measure (FCM), global form factor (GFF), maximum form factor in vertical and horizontal (FF_{maxv} , FF_{maxh}), mean form factor (FF_{mnv} , FF_{mnh}), first-order zero temperature change rate (ω), second-order zero temperature change rate ($\dot{\omega}$), and cold area percentages measure (CPM). They are calculated using equations 5 to 17, where N_w is the number of rows, N_l is the number of columns, T_{ij} is the pixel temperature at (i, j), and (dTH, d2TH, dTV, d2TV) are the first- and second-order derivatives in the horizontal and vertical axes using $\varepsilon \approx 0.01$.

$$PP = T_{max} - T_{min} \quad (5)$$

$$FDM = ((\sum \text{peak pixels}) / (\text{total number of pixels})) \quad (6)$$

$$FCM = \frac{\sum \text{peak pixels}}{\text{No of flat Portions}} \quad (7)$$

$$GFF = \frac{N \times \text{peakvalue}}{\sum_{j=1}^{N_l} \sum_{i=1}^{N_w} T_{i,j}} \quad (8)$$

$$FF_{maxv} = \frac{N_l \times \max\{\max_j T_{i,j}\}}{\sum_{i=1}^{N_l} \{\max_j T_{i,j}\}} \quad (9)$$

$$FF_{maxh} = \frac{N_w \times \max\{\max_i \{T_{i,j}\}\}}{\sum_{j=1}^{N_w} \{\max_i T_{i,j}\}} \quad (10)$$

$$FF_{mnv} = \frac{N_l \times \max\{\sum_{i=1}^{N_w} T_{i,j}\}}{N_w \times \sum_{j=1}^{N_l} \sum_{i=1}^{N_w} T_{i,j}} \quad (11)$$

$$F_{mnh} = \frac{N_w \times \max\{\sum_{j=1}^{N_l} T_{i,j}\}}{N_l \times \sum_{i=1}^{N_w} \sum_{j=1}^{N_l} T_{i,j}} \quad (12)$$

$$\omega_v = \frac{\text{No.of}\{d2TV < \varepsilon\}}{N}, \omega_h = \frac{\text{No.of}\{d2TH < \varepsilon\}}{N} \quad (13)$$

$$\bar{\omega}_h = \frac{\text{No.of}\{d2TH < \varepsilon\}}{N}, \bar{\omega}_v = \frac{\text{No.of}\{d2TV < \varepsilon\}}{N} \quad (14)$$

$$CPM = ((\sum \text{coldpixels}) / (\text{total number of pixels})) \quad (15)$$

4.3. Electrical Parameters of I-V Measurements

The electrical parameters are calculated using Equations (16)–(20). They include the operating current ratio (I_r), operating voltage ratio (V_r), operating power ratio (P_r), fill factor (FF), and PV efficiency (η). P_{in} is the incident solar radiation and A is the module area. Subscripts o and h refer to the operating value and the value of the healthy PV module.

$$I_r = I_o / I_h \quad (16)$$

$$V_r = V_o / V_h \quad (17)$$

$$P_r = P_o / P_h \quad (18)$$

$$FF = (I_o \times V_o) / (I_{SC} \times V_{OC}) \quad (19)$$

$$H = (I_{SC} \times V_{OC} \times FF) / (P_{in} \times A) \quad (20)$$

The use of mathematical parameters for IR image features extraction was studied in [38]. These features are calculated using the pixel values of an IR image, as detailed in Section 4.1. The physical significance of each of these features is explained in [38]. The GFF, which analyses the temperature variation of pixels depending on the type of fault, is found to be one of the most useful parameters of the diagnosis technique. The temperature of modules varies depending on the fault type: low for dirty modules, moderate for PID fault modules, and high for hot spot modules. For dirty modules, the average value of GFF rises when multiple faults are observed. Assessing the maximum and mean temperature variation in both the horizontal and vertical axes improves the accuracy of fault diagnosis.

The proposed features are demonstrated to be independent of temperature fluctuations that may arise owing to the use of different infrared camera sensors because they do not directly deal with pixel temperatures or values. They can also describe the shape of fault patterns in cold and hot areas. A broad classification matrix summarizes the relationships between fault type and associated variables. This matrix can be used to schedule maintenance and automate fault classification.

5. ANFIS Fault Classification Technique

After deploying feature extraction methods, ANFIS is used to categorize CIGS module faults. Fuzzy logic and ANNs are combined in ANFIS. Fuzzy if-then rules with appropriate membership functions are derived using the ANN's learning capabilities [40]. The ANFIS structure and its implementation of automatic fault detection are described in the following subsections.

5.1. ANFIS Structure

A typical ANFIS structure with two input variables and five levels is shown in Figure 3. From a set of inputs, the ANFIS algorithm produces only one output. The output is determined through consecutive stages including establishing fuzzy rules, fuzzifying the inputs with membership functions, defining rule strength and assessing its implications with a specific input dataset. The curve's parameters of the membership functions are determined using product values among the preset learning rules and the relevant weighted values. The parameters of membership functions are used to calculate the ratio between the individual and total weighted values. Finally, ANFIS foresees the target by calculating an overall gain value, which serves as the output value.

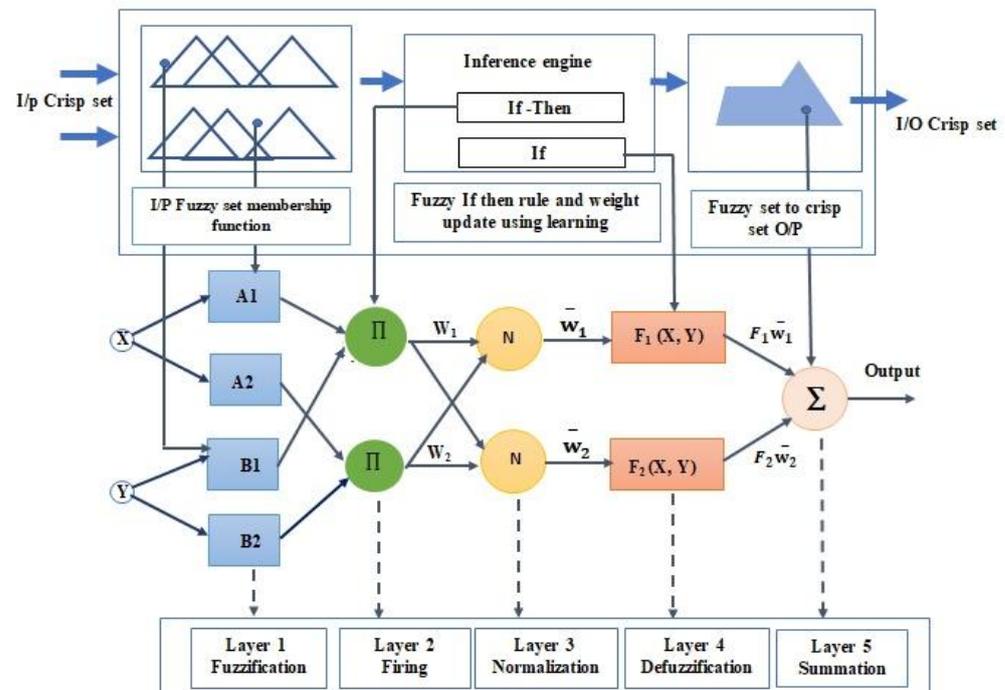


Figure 3. Proposed ANFIS structure.

To attain the desired values, the input and output membership function parameters are adjusted during the learning process. The ANFIS method is a hybrid algorithm that employs both back-propagation and least-square estimation techniques. These techniques are combined in an artificial neural network to offer a considerably faster and more accurate output to the ANFIS target. The linguistic values of the relationship between the input data rows (variable data) are determined by the rule “IF-THEN”.

The ANFIS design architecture comprises two inputs (x, y), two rules ($R1, R2$), and five layered feed-forward networks, as shown in Figure 3. The design also indicates adaptive (square) and non-adaptive (circle) nodes with one output. The learning algorithm is used to update the adaptive node parameters by minimizing tracking errors between target data and ANFIS output. The two fuzzy if-then rules are determined as function of the inputs (x, y) [40]. $A1, B1, A2,$ and $B2$ are the fuzzy sets and F is the number of membership functions. A typical ANFIS structure includes five layers. Layer 1 is the fuzzification layer,

and every node I is an adaptive node in this layer. This layer's outputs are the inputs' fuzzy membership grades. Layer 2 is called the rule layer (the firing). It is a fixed node layer, labelled II . To obtain the output of layer 2, which is the product of all the incoming signals, the firing strength is computed. Layer 3 is the normalization layer. It has a fixed node, labeled N . Normalized firing strengths w_i are the outputs of this layer. It determines the ratio of the i^{th} rule's firing strength to the total firing strength. Layer 4 is the defuzzification layer; it is a part that takes the fuzzy set inputs generated by the inference engine and transfers the fuzzy values into crisp values. Layer 5 is a summation neuron, a stationary node that adds all input signals together to compute the final output.

5.2. Application of ANFIS for Classification of Faults

The Jang model [40] is adopted in the present study; it is one of the most common fuzzy system architectures with less time-consumption and works well when large data are available for training. The input data are built from the feature extraction methods described in Section 4. The available input data are randomly divided into 50% as training data, 25% as check data, and 25% as test data. At the beginning, input and output datasets are used to plan the ANFIS architecture. Using feature extraction methods, the principal component analysis (PCA) is estimated for each input dataset. PCA is a technique for reducing the dimensionality of a dataset containing many interconnected elements. This was achieved by using a covariance matrix of extracted features to extract the more effective features that are uncorrelated with one another. Therefore, features reduction using principal component analysis decreases the computation volume and enhances the classification accuracy. According to the present results, the form factors (GFF , $FF_{\max h}$, $FF_{\max v}$, $FF_{\min h}$, $FF_{\min v}$, flatness continuity measure (FCM), statistical features extraction, operating power ratio (Pr) and efficiency (η) are found to be the most effective features.

The data extracted using principal components analysis are divided into three categories for each feature extraction method: training, checking, and testing. The goal of utilizing a checking dataset is to avoid overfitting the training dataset. In theory, as training advances, the model error tends to decrease, leading to overfitting, and an abrupt increase in model error occurs. Overfitting is accounted for by testing the FIS trained on the training data against the checking data. Furthermore, if these errors indicate model overfitting, the membership function parameters associated with the smallest checking error are chosen. The testing dataset is used to evaluate the FIS's generalization capability. Eight membership functions are selected for inputs and the same type of membership functions for the output is applied during the generation of a fuzzy inference system. The objective of increasing the number of fuzzy memberships per variable is to improve the matching accuracy of the proposed ANFIS. As the number of fuzzy memberships per variable increases, the number of robust rules increases, resulting in higher matching accuracy. The total number of samples (PV modules) is slightly low, 84 samples, compared to other ANFIS applications; thus, the number of fuzzy memberships is increased to overcome this problem. The system was prepared with training data and checked with a test dataset. The ANFIS model is implemented using MATLAB[®] software. The number of nodes and fuzzy rules is determined, and a neural network training algorithm is utilized to refine the rules and decrease errors. Finally, the output of the classifications is obtained.

5.3. Analysis of ANFIS Results

In this section, the results of the performance of ANFIS models, to classify the faults of CIGS PV module, are illustrated. For each type of feature extraction, a series of trials was applied to the dataset to select an adequate membership function type for the best classification. The data were divided into three sets, each of which was mutually exclusive. There were 512 rules, three inputs, one output, and eight membership functions per input in the final network. The ANFIS system parameters are shown in Table 3. The results of each feature extraction method are addressed in the following section. After many trials, the best

membership function type that obtained a high classification percentage for each type of feature extraction is shown in Table 4. Also, all features were trained and tested together.

Table 3. ANFIS system parameters.

Item	Number of
Nodes	1078
Linear parameters	2048
Nonlinear parameters	48
Training data pairs	36
Checking data pairs	27
Fuzzy rules	512

Table 4. Type of membership and accuracy for type of feature extraction method.

Type of Feature Extraction (FE) Methods	Type of Membership Function	Accuracy
Statistical (FE)	Triangle	83.33%
I-V measurement (FE)	Gaussian	100%
Mathematical parameter (FE)	All type	100%

A general classification matrix (GCM) was proposed in [3,38] for classification of faults. It has been shown that mathematical parameter-based feature extraction exhibited a higher capability in classification than statistical feature extraction and electrical measurement-based feature extraction. However, the GCM takes more time in classification, since it compares each type of fault with other types using feature extraction methods. On the other hand, the ANFIS approach proposed in the present study is very fast in the classification process and detects the type of fault for the classified module. For each type of fault, the proper required maintenance action can be scheduled/implemented once the data are trained. This represents an advantage as compared to previous studies on fault classification that define only the module to be either defective or non-defective and fails to determine the fault type and corresponding proper maintenance action.

6. Estimation of Operating Power Ratio

The operating power ratio (P_r) of PV modules represents the ratio of PV module power output and the healthy module power output (refer to Equation (18)). The previously reported research efforts were concerned with the utilization of IRT images for the detection and classification of PV faults. No research efforts have considered the possibility of correlating the features of IRT images to the operating power ratio. Such a correlation is very useful in predictive maintenance applications; it can be used as a measure for deciding a maintenance action or even module replacement. Assessment of the operating power ratio of a given PV module in the plant is normally performed using I-V measurements in offline conditions. Correlation of IRT image features to the P_r of PV modules can be performed online without any disturbance in plant operation. Studying the effect of faults on power loss is very useful to correlate fault types with module reliability, degradation, and remaining useful life, and obtaining a future picture of the performance of the PV power plant. Figure 4 shows a box plot of power loss due to faults, with the Y-axis representing the operating power ratio of CIGS PV modules and the X-axis representing the box plot of each type of fault.

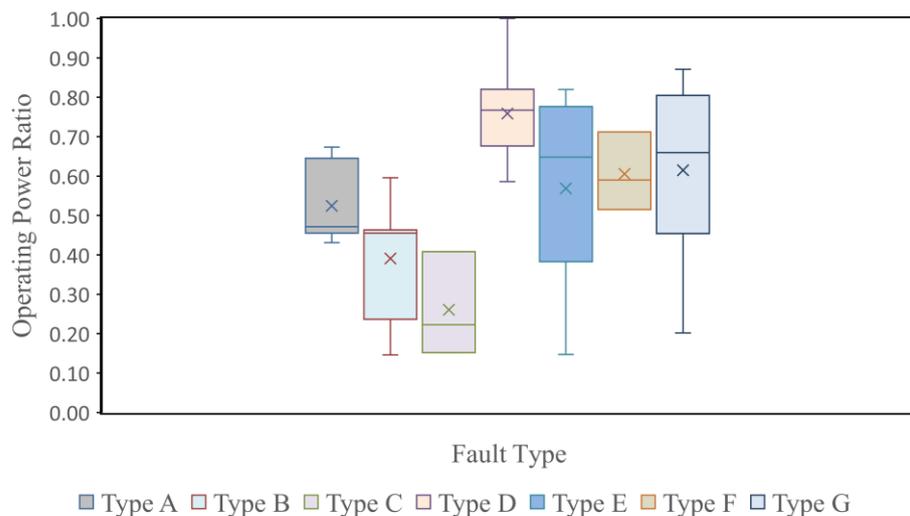


Figure 4. Box plot of operating power ratio for different fault types.

In the present article, the correlation between the P_r and the mathematical and statistical features is obtained for each type of fault. The features with high correlation with the power ratio are determined. Then, regression analysis is used to obtain the correlation equation between the high correlation features and the operating power ratio using Minitab 19 software. Figure 5 shows the correlation matrix between features with the P_r for each type of fault. The high correlation values are marked with red circles. For each type of fault, there is a feature that gives a high correlation with the P_r . Since the maintenance action of faulty modules with fault type H is to remove the module, this type of fault was not considered in the correlation with the P_r . It can be observed that the features that have a high correlation with the P_r include skewness (γ_1), FDM, ω , and PP.

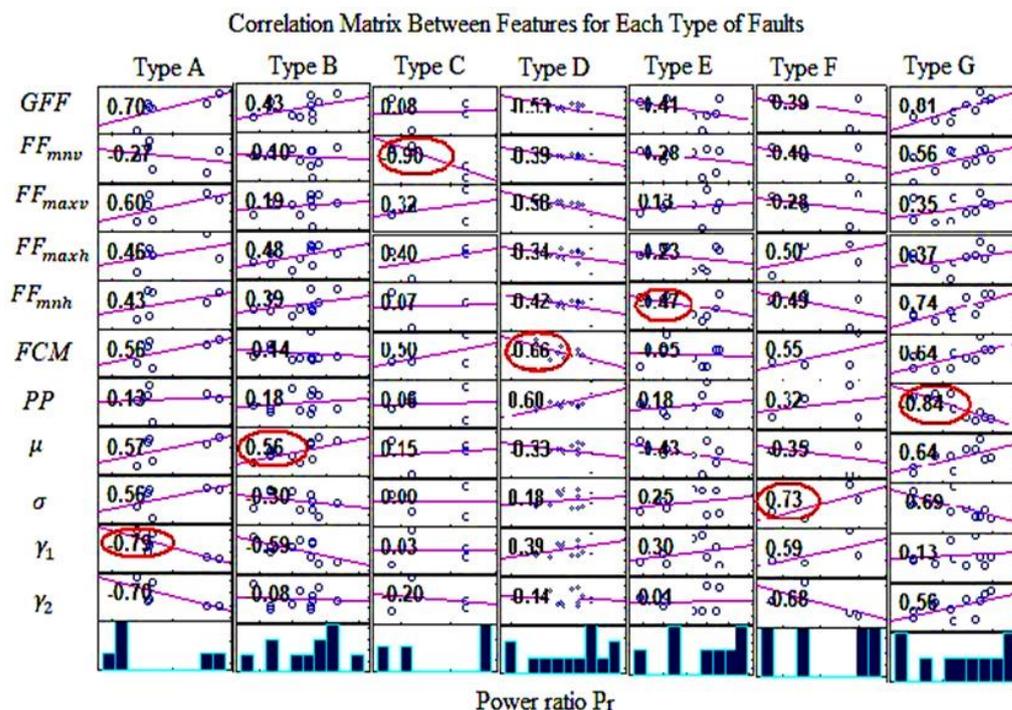


Figure 5. Correlation matrix between features and operating power ratio for each type of fault.

The results of regression analysis using Minitab software are shown in Figure 6. The regression model equation and the values of the R-squared and p -value for each fault type

are shown in Table 5. The value of R-squared (0–100%) represents the scatter around the regression line. A p -value between 0 and 1 is commonly used to represent the level of statistical significance. The lower the p -value, the more evidence there is for rejecting the null hypothesis. A statistically significant p -value is usually less than 0.05. the combination of low p -value and high R-squared indicates that changes in the predictors are related to changes in the response variable and that the regression model explains a lot of the response variability. The combination of low p -value and low R-squared indicates that the regression model has significant variables but explains little of the variability. The results in Table 5 reveal that features with faults of type A, C, D, F, and G have high R-squared values. Type B and F faults, on the other hand, have low R-squared values. All types of faults have a low p -value, indicating strong evidence against the null hypothesis. The current study advises that more research should be carried out on the analysis of the operational power ratio for type B and F faults. Figure 6 shows that the values of P_r decrease with the increase in γ_1 for A and B type faults. The variation of P_r with PP is dependent upon the fault type. The value of P_r decreases with the increase in PP for fault type D and decreases with the decrease in PP for fault type C.

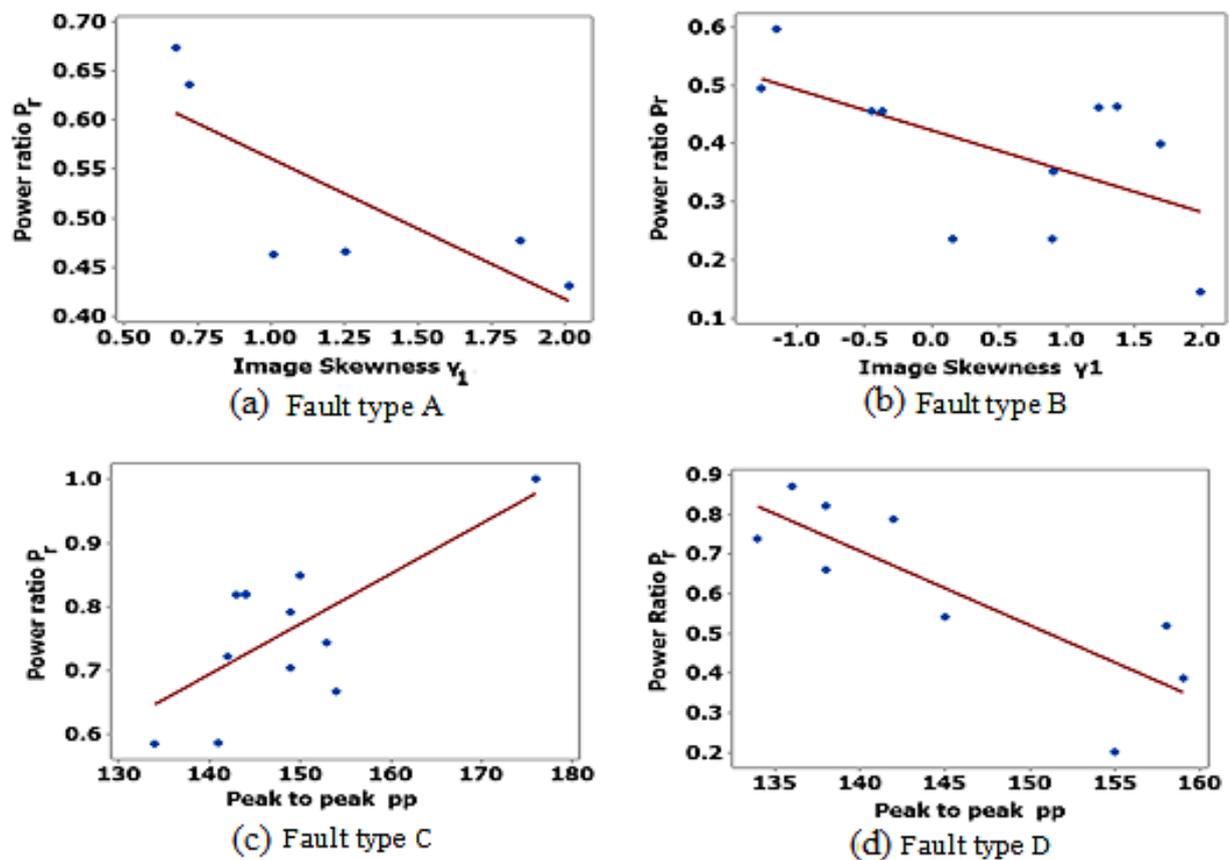


Figure 6. Regression analysis for correlation of operating power ratio for different types of faults.

Table 5. Regression model parameters for different types of faults.

Fault Type	Regression Model	R-sq %	p -Value
A	$P_r = 0.7035 - 0.1428 \gamma_1$	62.34	0.062
B	$P_r = 0.4227 - 0.07026 \gamma_1$	34.83	0.056
C	$P_r = 0.5929 - 19.28 \text{FDM}$	99.99	0.06
D	$P_r = -0.4062 + 0.007860 \text{pp}$	48.67	0.012
E	$P_r = -5.292 + 6.747 \omega$	57.41	0.029
F	$P_r = 0.5444 + 0.6615 \text{FDM}$	77.43	0.315
G	$P_r = 3.332 - 0.0187 \text{pp}$	69.24	0.005

7. Conclusions

An ANFIS is developed for automation of the fault classification process of PV modules. Statistical and mathematical features of outdoor IRT and I-V measurements of thin-film PV modules of a typical CIGS PV plant are used to provide the dataset for the implementation and analysis of the proposed scheme. The proposed features are proven to be unaffected by temperature variations that may occur due to the usage of different infrared camera sensors.

Many tests have been carried out to determine the best type of membership function for high classification accuracy. PCA is used to reduce dimensionality and speed up the classification process. This was achieved by using a covariance matrix of extracted features to obtain the more effective features that are uncorrelated with one another. According to the results, the most effective classification features are found to be global form factor (GFF), maximum form factor for horizontal and vertical (FF_{maxh} , FF_{maxv}), mean form factor for horizontal and vertical (FF_{mnh} , FF_{mnv}), flatness continuity measure (FCM), statistical features, operating power ratio (P_r) and efficiency (η). The evaluation of the proposed approach shows that the accuracy of classification reaches 100% using feature extraction methods that are based on mathematical parameters and I-V measurements, and 83.33% using features based on statistical parameters.

A novel approach is proposed for developing a correlation matrix between the P_r and the effective features extracted online from IRT images. Thus, the need for offline I-V measurements to estimate the P_r of PV modules is eliminated. Effective features that are highly correlated with the operating power ratio of PV modules are determined for each type of fault. The features of IR images that have a high correlation with the P_r include skewness (γ_1), mean (μ), mean form factor in the vertical direction (FF_{mnv}), FCM, mean form factor in the horizontal direction (FF_{mnh}), σ , and PP. The results of a regression analysis of the P_r demonstrate that features with faults of type A, C, D, F, and G have high R-squared values. Type B and F faults, on the other hand, have low R-squared values. All types of faults have a low p -value, indicating strong evidence against the null hypothesis. The present analysis suggests the need for further research work on the analysis of P_r for faults of types B and F.

Author Contributions: Conceptualization, R.A.E., M.R., H.A.M. and K.H.I.; Methodology, R.A.E., M.R., H.A.M. and K.H.I.; Software, R.A.E., H.A.M. and K.H.I.; Validation, R.A.E. and K.H.I.; Formal analysis, R.A.E. and H.A.M.; Writing—Original draft preparation, R.A.E. and M.R.; Supervision, M.R. and H.A.M.; Project administration, Funding acquisition, M.R. and E.A.; Writing—Review and Editing, E.A., H.A.M. and K.H.I.; Resources, E.A.; Investigation, E.A. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Institutional Fund Projects, King Abdulaziz University, under grant no. [IFPIP-1059-829-1442]. And the APC was funded by [IFPIP-1059-829-1442].

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data that support the findings of this study are available from the corresponding author, upon reasonable request.

Acknowledgments: This research work was funded by Institutional Fund Projects under grant no. (IFPIP-1059-829-1442). Therefore, the authors gratefully acknowledge the technical and financial support from the Ministry of Education and King Abdulaziz University, Jeddah, Saudi Arabia.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Feldman, R.; Wu, K.; Margolis, R. Solar Industry Update. NREL/PR-7A40-80427; 2021. Available online: <https://www.nrel.gov/docs/fy21osti/80427.pdf> (accessed on 12 April 2022).
2. Haque, A.; Bharath, K.V.S.; Khan, M.A.; Jaffery, Z.A. Fault diagnosis of photovoltaic modules. *Energy Sci. Eng.* **2019**, *7*, 622–644. [[CrossRef](#)]
3. Ibrahim, K.; Eltuhamy, R.; Rady, M.; Mahmoud, H. Failure Mode and Effects Analysis of CIGS Thin Film PV Modules Using Thermography Analysis and IV Measurements. *Int. J. Energy Convers. (IRECON)* **2021**, *9*, 17–28. [[CrossRef](#)]
4. Köntges, M.; Kurtz, S.; Packard, C.; Jahn, U.; Berger, K.A.; Kato, K.; Friesen, T.; Liu, H.; Iseghem, M.V. *Review of Failures of Photovoltaic Modules*; Report IEA-PVPS T13-01:2014; International Energy Agency: Paris, France, 2014.
5. Buerhop, C.; Schlegel, D.; Niess, M.; Vodermayr, C.; Weißmann, R.; Brabec, C. Reliability of IR-imaging of PV-plants under operating conditions. *Sol. Energy Mater. Sol. Cells* **2012**, *107*, 154–164. [[CrossRef](#)]
6. Wang, P.; Zheng, S. Fault Analysis of Solar PV Array Based on Infrared Image. *Acta Energ. Sol. Sin.* **2010**, *31*, 197–202.
7. Patel, H.; Agarwal, V. MATLAB-Based Modeling to Study the Effects of Partial Shading on PV Array Characteristics. *IEEE Trans. Energy Convers.* **2008**, *23*, 302–310. [[CrossRef](#)]
8. Kumar, A.; Pachauri, R.K.; Chauhan, Y.K. Experimental analysis of SP/TCT PV array configurations under partial shading conditions. In Proceedings of the 2016 IEEE 1st International Conference on Power Electronics, Intelligent Control and Energy Systems (ICPEICES), Delhi, India, 4–6 July 2016; pp. 1–6. [[CrossRef](#)]
9. Bechouat, M.; Younsi, A.; Sedraoui, M.; Soufi, Y.; Yousfi, L.; Tabet, I.; Touafek, K. Parameters identification of a photovoltaic module in a thermal system using meta-heuristic optimization methods. *Int. J. Energy Environ. Eng.* **2017**, *8*, 331–341. [[CrossRef](#)]
10. Lydia, M.; Sindhu, K.; Gagan, K. Analysis on Solar Panel Crack Detection Using Optimization Techniques. *J. Nano Electron. Phys.* **2017**, *9*, 02004. [[CrossRef](#)]
11. Titri, S.; Larbes, C.; Toumi, K.Y.; Benatchba, K. A new MPPT controller based on the Ant colony optimization algorithm for Photovoltaic systems under partial shading conditions. *Appl. Soft Comput.* **2017**, *58*, 465–479. [[CrossRef](#)]
12. Natarajan, K.; Pala, P.K.; Sampath, V. Fault Detection of Solar PV system using SVM and Thermal Image Processing. *Int. J. Renew. Energy Res.* **2020**, *10*, 967–977.
13. Madeti, S.R.; Singh, S.N. Modeling of PV system based on experimental data for fault detection using kNN method. *Sol. Energy* **2018**, *173*, 139–151. [[CrossRef](#)]
14. Shin, J.H.; Kim, J.O. On-Line Diagnosis and Fault State Classification Method of Photovoltaic Plant. *Energies* **2020**, *13*, 4584. [[CrossRef](#)]
15. Zhao, Y.; Yang, L.; Lehman, B.; de Palma, J.-F.; Mosesian, J.; Lyons, R. Decision tree-based fault detection and classification in solar photovoltaic arrays. In Proceedings of the Twenty-Seventh Annual IEEE Applied Power Electronics Conference and Exposition (APEC), Orlando, FL, USA, 5–9 February 2012; pp. 93–99. [[CrossRef](#)]
16. Sun, J.; Sun, F.; Fan, J.; Liang, Y. Fault Diagnosis Model of Photovoltaic Array Based on Least Squares Support Vector Machine in Bayesian Framework. *Appl. Sci.* **2017**, *7*, 1199. [[CrossRef](#)]
17. Pierdicca, R.; Malinverni, E.S.; Piccinini, F.; Paolanti, M.; Felicetti, A.; Zingaretti, P. Deep convolutional neural network for automatic detection of damaged photovoltaic cells. *ISPRS Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2018**, *XLII-2*, 893–900. [[CrossRef](#)]
18. Jaffery, Z.A.; Dubey, A.K.; Irshad; Haque, A. Scheme for predictive fault diagnosis in photo-voltaic modules using thermal imaging. *Infrared Phys. Technol.* **2017**, *83*, 182–187. [[CrossRef](#)]
19. Belaout, A.; Krim, F.; Mellit, A.; Talbi, B.; Arabi, A. Multiclass adaptive neuro-fuzzy classifier and feature selection techniques for photovoltaic array fault detection and classification. *Renew. Energy* **2018**, *127*, 548–558. [[CrossRef](#)]
20. Lin, H.; Chen, Z.; Wu, L.; Lin, P.; Cheng, S. Online Monitoring and Fault Diagnosis of PV Array Based on BP Neural Network Optimized by Genetic Algorithm. *Multi-Discip. Trends Artif. Intell.* **2015**, *9426*, 102–112. [[CrossRef](#)]
21. Kurukuru, V.S.B.; Haque, A.; Khan, M.A.; Tripathy, A.K. Fault classification for Photovoltaic Modules Using Thermography and Machine Learning Techniques. In Proceedings of the International Conference on Computer and Information Sciences (ICCIS), Sakaka, Saudi Arabia, 3–4 April 2019; pp. 1–6.
22. Dunderdale, C.; Brettigny, W.; Clohessy, C.; Dyk, E.E. Photovoltaic defect classification through thermal infrared imaging using a machine learning approach. *Prog. Photovolt. Res. Appl.* **2020**, *28*, 177–188. [[CrossRef](#)]
23. Niazi, K.; Akhtar, W.; Khan, H.A.; Sohaib, S.; Nasir, A.K. Binary Classification of Defective Solar PV Modules Using Thermography. In Proceedings of the 2018 IEEE 7th World Conference on Photovoltaic Energy Conversion (WCPEC) (A Joint Conference of 45th IEEE PVSC, 28th PVSEC & 34th EU PVSEC), Waikoloa, HI, USA, 10–15 June 2018; pp. 753–757. [[CrossRef](#)]
24. Leotta, G.; Pugliatti, P.; Di Stefano, A.; Aleo, F.; Bizzarri, F. Post processing technique for thermographic images provided by drone inspections. In Proceedings of the 31st European Photovoltaic Solar Energy Conference and Exhibition (EU PVSEC), Hamburg, Germany, 14–18 September 2015. [[CrossRef](#)]
25. Rasch, R.; Behrens, G.; Hamelmann, F.; Hantelmann, S.; Dreimann, R.; Weicht, J. Automated Thermal Imaging for Fault Detection on PV-Systems. In Proceedings of the 31st European Photovoltaic Solar Energy Conference and Exhibition (EU PVSEC), Hamburg, Germany, 14–18 September 2015. [[CrossRef](#)]
26. Le, M.; Luong, V.S.; Nguyen, D.K.; Dao, V.-D.; Vu, N.H.; Vu, H.H.T. Remote anomaly detection and classification of solar photovoltaic modules based on deep neural network. *Sustain. Energy Technol. Assess.* **2021**, *48*, 101545. [[CrossRef](#)]

27. Fadhel, S.; Delpha, C.; Diallo, D.; Bahri, I.; Migan, A.; Trabelsi, M.; Mimouni, M. PV shading fault detection and classification based on I-V curve using principal component analysis: Application to isolated PV system. *Sol. Energy* **2018**, *179*, 1–10. [[CrossRef](#)]
28. Samara, S.; Natsheh, E. Intelligent PV Panels Fault Diagnosis Method Based on NARX Network and Linguistic Fuzzy Rule-Based Systems. *Sustainability* **2020**, *12*, 2011. [[CrossRef](#)]
29. Ul Haq, A.; Sindi, H.; Gul, S.; Jalal, M. Modeling and Fault Categorization in Thin-Film and Crystalline PV Arrays through Multilayer Neural Network Algorithm. *IEEE Access* **2020**, *8*, 102235–102255. [[CrossRef](#)]
30. Aziz, F.; Haq, A.U.; Ahmad, S.; Mahmoud, Y.; Jalal, M.; Ali, U. A Novel Convolutional Neural Network Based Approach for Fault Classification in Photovoltaic Arrays. *IEEE Access* **2020**, *8*, 41889–41904. [[CrossRef](#)]
31. Huang, J.M.; Wai, R.J.; Gao, W. Newly Designed Fault Diagnostic Method for Solar Photovoltaic Generation System Based on IV-Curve Measurement. *IEEE Access* **2019**, *7*, 70919–70932. [[CrossRef](#)]
32. Wu, Y.; Chen, Z.; Wu, L.; Lin, P.; Cheng, S.; Lu, P. An Intelligent Fault Diagnosis Approach for PV Array Based on SA-RBF Kernel Extreme Learning Machine. *Energy Procedia* **2017**, *105*, 1070–1076. [[CrossRef](#)]
33. Mekki, H.; Mellit, A.; Salhi, H. Artificial neural network-based modelling and fault detection of partial shaded photovoltaic modules. *Simul. Model. Pract. Theory* **2016**, *67*, 1–13. [[CrossRef](#)]
34. Bonsignore, L.; Davarifar, M.; Rabhi, A.; Tina, G.M.; Elhajjaji, A. Neuro-Fuzzy Fault Detection Method for Photovoltaic Systems. *Energy Procedia* **2014**, *62*, 431–441. [[CrossRef](#)]
35. Lazzaretti, A.E.; Da Costa, C.H.; Rodrigues, M.P.; Yamada, G.D.; Lexinoski, G.; Moritz, G.L.; Oroski, E.; De Goes, R.E.; Linhares, R.R.; Stadzisz, P.C.; et al. A Monitoring System for Online Fault Detection and Classification in Photovoltaic Plants. *Sensors* **2020**, *20*, 4688. [[CrossRef](#)]
36. Hussain, M.; Dhimish, M.; Titarenko, S.; Mather, P. Artificial neural network based photovoltaic fault detection algorithm integrating two bi-directional input parameters. *Renew. Energy* **2020**, *155*, 1272–1292. [[CrossRef](#)]
37. Dhimish, M.; Badran, G. Photovoltaic Hot-Spots Fault Detection Algorithm Using Fuzzy Systems. *IEEE Trans. Device Mater. Reliab.* **2019**, *19*, 671–679. [[CrossRef](#)]
38. Eltuhamy, R.A.; Rady, M.; Ibrahim, K.H.; Mahmoud, H.A. Novel features extraction for fault detection using thermography characteristics and IV measurements of CIGS thin-film module. *Instrum. Mes. Métrologie* **2020**, *19*, 311–325. [[CrossRef](#)]
39. International Electrotechnical Commission. *Photovoltaic (PV) Systems—Requirements for Testing, Documentation and Maintenance—Part 3: Photovoltaic Modules and Plants—Outdoor Infrared Thermography*; IEC TS 62446-3: 2007; International Electrotechnical Commission (IEC): Geneva, Switzerland, 2017.
40. Jang, J.S.R. ANFIS: Adaptive-network-based fuzzy inference system. *IEEE Trans. Syst. Man Cybern.* **1993**, *23*, 665–685. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.