

Review

A review of automated solar photovoltaic defect detection systems: Approaches, challenges, and future orientations



Ula Hijjawi ^{a,b,*}, Subhash Lakshminarayana ^a, Tianhua Xu ^a, Gian Piero Malfense Fierro ^b, Mostafizur Rahman ^b

^a School of Engineering, University of Warwick, Coventry CV4 7AL, United Kingdom

^b The Manufacturing Technology Centre (MTC), Ansty Park, Coventry CV7 9JU, United Kingdom

ARTICLE INFO

ABSTRACT

Keywords:
 Photovoltaic systems
 Solar module
 Defect detection
 Imaging-based techniques
 Electrical testing techniques
 Artificial Intelligence
 Fault diagnosis
 Machine Learning

The development of Photovoltaic (PV) technology has paved the path to the exponential growth of solar cell deployment worldwide. Nevertheless, the energy efficiency of solar cells is often limited by resulting defects that can reduce their performance and lifespan. Therefore, it is crucial to identify a set of defect detection approaches for predictive maintenance and condition monitoring of PV modules. This paper presents a comprehensive review of different data analysis methods for defect detection of PV systems with a high categorisation granularity in terms of types and approaches for each technique. Such approaches, introduced in the literature, were categorised into **Imaging-Based Techniques (IBTs)** and **Electrical Testing Techniques (ETTs)**. Although several review papers have investigated recent solar cell defect detection techniques, they do not provide a comprehensive investigation including IBTs and ETts with a greater granularity of the different types of each for PV defect detection systems. Types of IBTs were categorised into Infrared Thermography (IRT), Electroluminescence (EL) imaging, and Light Beam Induced Current (LBIC). On the other hand, ETts were categorised into Current-Voltage (I-V) characteristics analysis, Earth Capacitance Measurements (ECM), Time Domain Reflectometry (TDR), Power Losses Analysis (PLA), and Voltage and Current Measurements (VCM). Approaches based on digital/signal processing and Machine Learning (ML) models for each method are included where relevant. Moreover, the paper critically analyses the advantages and disadvantages of each of the adopted techniques, which can be referred to by future studies to identify the most suitable method considering the use-case's requirements and setting. The adoption of each of the reviewed techniques depends on several factors, including the deployment scale, the targeted defects for detection, and the required location of defect analysis in the PV system, which are expanded further in the presented analysis. From a high-level perspective, while IBTs provide a high-resolution visual representation of the module surface, allowing for the detection and diagnosis of small structural defects that may be missed by other techniques, ETts can detect electrical faults beyond the PV module's surface. On the IBT level, the most notable adopted techniques in the literature are IRT- and EL-based. While IRT techniques are more practical for large-scale applications than EL imaging, the latter is considered a non-intrusive technique that is highly efficient in localising defects of solar cells. The paper also discusses challenges observed in the state-of-the-art related to data availability, real-time monitoring, accurate measurements, computational efficiency, and dataset distribution, and reviews data pre-processing and augmentation approaches that can address

Abbreviations: 1D, 1-Dimensional; IBT, Imaging-Based Technique; 2D, 2-Dimensional; ICA, Independent Component Analysis; ADC, Analog-to-Digital Converter; IR, Infrared; AI, Artificial Intelligence; IRT, Infrared Thermography; aIRT, Aerial Infrared Thermography; I-V, Current-Voltage; ANN, Artificial Neural Network; KNN, K-Nearest Neighbour; BOS, Balance-of-System; LBIC, Light Beam Induced Current; CCD, Charge Coupled Device; ML, Machine Learning; CE, Conversion Efficiency; MLP, Multilayer Perceptron; CIGS, Copper-indium-Gallium-disulfide; MPPT, Maximum Power Point Tracker; CNN, Convolutional Neural Network; MSE, Mean Square Error; CS-LBIC, Compressive Sensing-based LBIC; NDIT, Non-Destructive Testing Technique; CV, Computer Vision; NLP, Natural Language Processing; DFT, Discrete Fourier Transform; PLA, Power Losses Analysis; DIP, Digital Image Processing; PSC, Perovskite Solar Cell; DL, Deep Learning; PV, Photovoltaic; EBIC, Electron Beam Induced Current; R-CNN, Region-based Convolutional Neural Network; ECM, Earth Capacitance Measurements; RF, Random Forest; EL, Electroluminescence; SSTDR, Spread Spectrum Time Domain Reflectometry; ETT, Electrical Testing Technique; SVM, Support Vector Machine; FNR, False Negative Rate; TDR, Time Domain Reflectometry; GAN, Generative Adversarial Networks; TIP, Thermal Image Processing; GPU, Graphics Processing Unit; VAE, Variational Autoencoder; HOG, Histogram of Oriented Gradient; VCM, Voltage and Current Measurements.

* Corresponding author.

E-mail address: ula.hijjawi@warwick.ac.uk (U. Hijjawi).

<https://doi.org/10.1016/j.solener.2023.112186>

Received 17 August 2023; Received in revised form 10 October 2023; Accepted 9 November 2023

Available online 22 November 2023

0038-092X/© 2023 The Authors. Published by Elsevier Ltd on behalf of International Solar Energy Society. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

some of these challenges. Furthermore, potential future orientations are identified, addressing the limitations of PV defect detection systems.

1. Introduction

Different statistical outcomes have affirmed the significance of Photovoltaic (PV) systems and grid-connected PV plants worldwide. Surprisingly, the global cumulative installed capacity of solar PV systems has massively increased since 2000 to 1,177 GW by the end of 2022 [1]. Moreover, installing PV plants has led to the exponential growth of solar cell deployment worldwide. For example, the cumulative number of solar PV installations in the UK has boomed from 29,320 in 2010 to 1,249,761 by the end of 2022 [2]. Nevertheless, the energy efficiency of solar cells is often limited by resulting defects that can reduce their performance and lifespan. Defects can disseminate power by creating new recombination pathways (losses), allowing the light to generate heat rather than electricity, or even consuming power stored in the battery bank, degrading the PV module's efficiency [3,4]. Moreover, the new generations of solar cells, such as Copper-indium-Gallium-disulfide (CIGS) and Perovskite solar cells (PSCs), come with emerging challenges related to increasing their power-conversion efficiency, reducing the fabrication cost and reducing the environmental impact when using toxic materials [5]. For example, recent research and the manufacturing sector have shown a growing interest in the development of PSC technology due to the ease of its fabrication process and higher conversion efficiency. In fact, reports of KRICT and MIT have recently verified an efficiency of 25.2 % for PSCs [6,7]. However, one of the main challenges encountered in PSCs is their low stability, which can be influenced by intrinsic or extrinsic factors, including soaking, thermal stress, and/or humid conditions leading to chemical and structural changes [5,8]. This can eventually lead to higher costs and a limited lifetime. Therefore, it is crucial to identify a set of defect detection approaches for predictive maintenance and condition monitoring of PV modules.

Several studies have classified defects into different groups based on their causes, effects, and locations. Rana and Arora [9] categorised defects into manufacturing and environmental defects. Inspection of manufacturing defects is critical in the early stages to prevent delivering defective modules to the following process. They are caused by mechanical pressure or improper manipulation during manufacturing, yielding scratches in the cells of PV modules. Cracks are also often induced during manufacturing for various reasons, including stress produced during soldering, needle pressing on a silicon wafer (known as a cross crack line), or applying heavy loads during transportation, handling, and assembly [10]. Environmental defects, on the other hand, are generated due to post-installation conditions. For example, soil deposition on the module's glass surface due to dust accumulation or humidity leads to partial shadowing. Heating of a module's parts results in discolouration or burn marks. Exposure to high temperature, UV, and water degrades the adhesive material between the glass and the cell, causing cell yellowing. The degradation of sealing between the module's parts leads to corrosion and faulty interconnections between the modules or strings of a module [9]. A similar approach to defect classification was introduced by Kurukuru et al., [11].

While the defects above alter the appearance of the PV module's surface, common failures of PV systems that may be invisible were classified by Mansouri et al., [12] into three main areas depending on the affected component during the operation: 1) PV module failures (e.g., bypass diode, mismatch, partial shading, and line-line faults), 2) power electronics interface failures at the DC side of the PV system (e.g., wear-out, open circuit, and short circuit faults due to power semiconductor failures), and 3) grid-side failures at the AC side of the PV system (e.g., islanding and wiring degradation). Triki-Lahiani et al., [13] classified PV system failures as Module Failures (occurring at the generator level), Inverter Failures, and others, including mismatch,

ground, and line-line faults.

Faults of PV systems were also classified into permanent and temporal by Mellit et al., [14]. Examples of permanent defects are delamination, bubbles, yellowing, scratches and burnt marks, and can be eliminated by replacing the affected components. However, temporal defects, such as partial shading and dust accumulation, can be removed by operators without replacing the faulty module.

Although the terms 'defects' and 'faults' were interchangeably used in the literature, it was observed that the reference to 'defects' was typically related to the physical components or materials used in the PV system, such as physical anomalies in PV modules (e.g., cracks, hotspots, delamination, disconnections, etc.). On the other hand, 'faults' can be typically related to malfunctions affecting the PV system's electrical performance, such as power electronics or Balance-of-System (BOS) failures (e.g., open circuits, short circuits, ground faults, inverter failures, etc.). The categorisations of failures in PV systems can be summarised as illustrated in Fig. 1.

Many approaches were introduced in the literature to investigate defect detection in PV systems, which are categorised in this article into Imaging-Based Techniques (IBTs) and Electrical Testing Techniques (ETTs). In IBTs, high-resolution images of the PV modules are captured, whereas time-series data of electrical signals are used in ETTs to analyse defects.

Recent state-of-the-art research has focused on Artificial intelligence (AI) and Machine Learning (ML) techniques for condition monitoring of PV modules to detect defects accurately. Such automatic defect detection systems would save time-consuming manual inspection efforts requiring intensive analysis of images captured by remote cameras [4]. They can also improve the PV panels' reliability and durability, help manage their deterioration, and enhance their long-term performance [5].

Several review papers have investigated recent techniques for solar cell defect detection. Mansouri et al., [12] have only reviewed fault diagnosis and detection techniques based on Deep Learning (DL) for PV systems from the perspective of methodology and five basic architectures: stacked autoencoder network, deep belief network, Convolutional Neural Network (CNN), recurrent neural network, and deep transfer learning. ML-based techniques for surface defect detection of solar cells were reviewed by Rana and Arora [9], of which were only imaging-based techniques. Similarly, Al-Mashhadani et al., [10] have reviewed DL-based studies that adopted only imaging-based techniques. Oliveira et al., [15] have only reviewed aerial Infrared Thermography (IRT) for PV plant inspection. Similarly, Herraiz et al., [16] have reviewed studies based on solar thermography for condition monitoring of PV plants. Mellit et al., [14] have only reviewed electrical testing techniques for fault diagnosis, focusing on faults occurring in PV arrays. Similarly, Triki-Lahiani et al., [13] reviewed fault detection and monitoring systems focusing on electrical signal approaches and electrical circuit simulation methods.

Nevertheless, review papers proposed in the literature need to provide a comprehensive review or investigation of all the existing data analysis methods for PV system defect detection, including imaging-based and electrical testing techniques with greater granularity of each category's different types of techniques. Moreover, a critical analysis of the advantages and disadvantages of each of the adopted techniques has yet to be addressed.

In this paper, different data analysis methods for the defect detection of PV systems are reviewed and discussed. The main contributions of this review paper are:

- A comprehensive investigation of data analysis methods for PV systems defect detection, including imaging-based and electrical testing techniques with a greater categorisation granularity in terms of types and approaches for each technique.
- Critical analysis of each of the different data analysis methods, including the advantages and disadvantages, can be referred to by future studies to identify the most suitable technique considering the use-case's requirements and setting.
- Review of data pre-processing and augmentation approaches in PV systems defect detection.
- Discussion of challenges related to the state-of-the-art followed by an outline of future orientations addressing the limitations of PV defect detection systems.

The paper is organised into seven sections: [Section 2](#) provides an overview of the categorised data analysis methods for PV system defect detection including Imaging-Based and Electrical Testing Techniques along with their corresponding types and approaches. [Section 3](#) reviews IBTs including Infrared Thermography (IRT), Electroluminescence (EL) Imaging, and Light Beam Induced Current (LBIC). Approaches based on digital image processing as well as ML are included. It also reviews automated image pre-processing pipelines required for data analysis. [Section 4](#) reviews ETTs including Current-Voltage (I-V) characteristics analysis, Earth Capacitance Measurements (ECM), Time Domain Reflectometry (TDR), Power Losses Analysis (PLA), Voltage and Current Measurements (VCM), and AI-based approaches. A comparative analysis of the reviewed studies on PV system defect detection and diagnosis is discussed in [Section 5](#) in addition to a critical analysis of the advantages and disadvantages of each method. Current challenges and future orientations are discussed in [Section 6](#). Data augmentation approaches are also reviewed in [Section 6](#) addressing some of the discussed challenges. [Section 7](#) concludes the paper.

2. Overview of data analysis methods

In this paper, data analysis methods for solar cell defect detection are categorised into two forms: 1) IBTs, which depend on analysing the deviations of optical properties, thermal patterns, or other visual features in images, and 2) ETTs, which depend on comparing the deviations of the module's measured electrical parameters from the expected electrical behaviour for detecting faults. Furthermore, both IBT and ETT adopted in the literature are investigated and categorised into a greater granularity of types of approaches as illustrated in [Fig. 2](#). IBTs are categorised into three different types: Infrared Thermography (IRT), Electroluminescence (EL) imaging, and Light Beam Induced Current (LBIC). On the other hand, ETTs are categorised into Current-Voltage (I-V) characteristics analysis, Earth Capacitance Measurements (ECM), Time Domain Reflectometry (TDR), Power Losses Analysis (PLA), Voltage and Current Measurements (VCM), and AI-based approaches.

It is also worth mentioning that the selection of these data analysis

methods has been guided by a comprehensive review of the existing literature, in which an extensive analysis of the current state of research has been concluded. Based on this analysis, the presented data analysis methods were further categorised into different types and corresponding approaches.

3. Imaging-based techniques (IBTs)

3.1. Infrared thermography (IRT)

IRT is considered one of the widely used, non-invasive techniques, in which the radiation emitted by the surface of any body is processed in the infrared wavelength spectrum between 1.4 and 15 μm [15]. In this method, thermal cameras are used, which are usually capable of measuring wavelengths in the mid-infrared wavelength range of 7–14 μm [15,17]. To carry out data collection, only Infrared (IR) cameras are required without the involvement of other equipment or interrupting the operation of the PV system [17].

In the case of a healthy PV module, the resulting temperature distribution in the thermal image is observed to be homogeneous on the module's surface. On the other hand, defects in the PV module having a significant influence on its thermal behaviour can be detected by observing variations in the temperature yielding inhomogeneities thermal distributions [18]. Thermal variations are caused due to heat produced by irradiance, or photons, that are not converted into electricity. Consequently, the accumulation of heat increases the temperature leading to defects appearing in the acquired IRT images, which can be characterised by increased series resistance, disconnected or shunted cells, or others affecting the thermal distribution [11,15]. For instance, defects appearing in the thermal image can be disconnected modules or strings recognised by hotter areas, shadowing recognised by hotspots, and cell cracks recognised by elongated cell heating [11]. Therefore, temperature differences form 'temperature signatures' shown by different colours or varying brightness levels, and provide information on the exact physical location of an occurring defective region [17,19].

According to Herraiz et al., [16,20], IRT for PV modules can be carried out using two main approaches:

- 1) Active IRT: in which additional energy is supplied to the targeted object by employing an external source, leading to an increase in temperature. Examples of such techniques include:
 - a) Pulsed Thermography: considered the most popular active IRT technique, in which a heat pulse is applied and data is collected after the temperature decreases [21].
 - b) Lock-In Thermography: a periodic external excitation, i.e., an oscillating temperature field, is applied to the PV module causing it to heat up. The resulting thermal patterns on the PV module's surface can then be captured using an IR camera. This technique is more popular for detecting shunt faults or failures in the modules

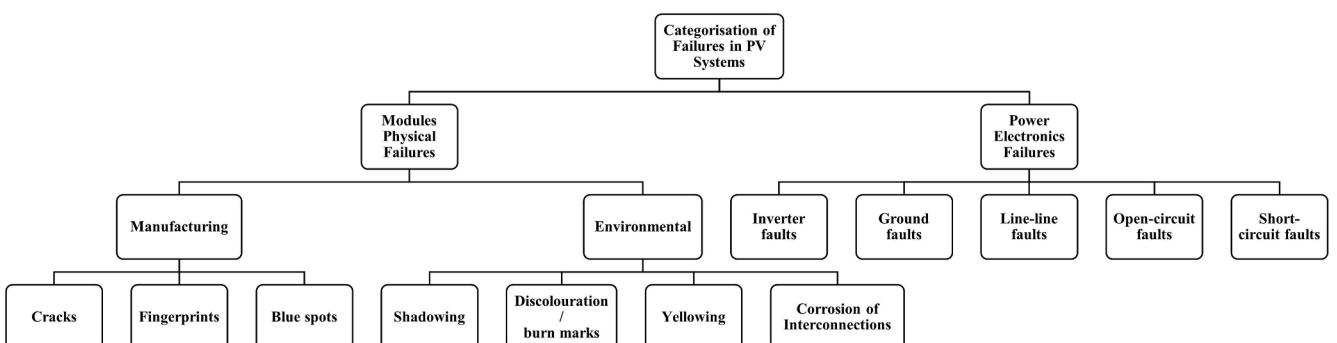


Fig. 1. Categorisation of failures in PV systems.

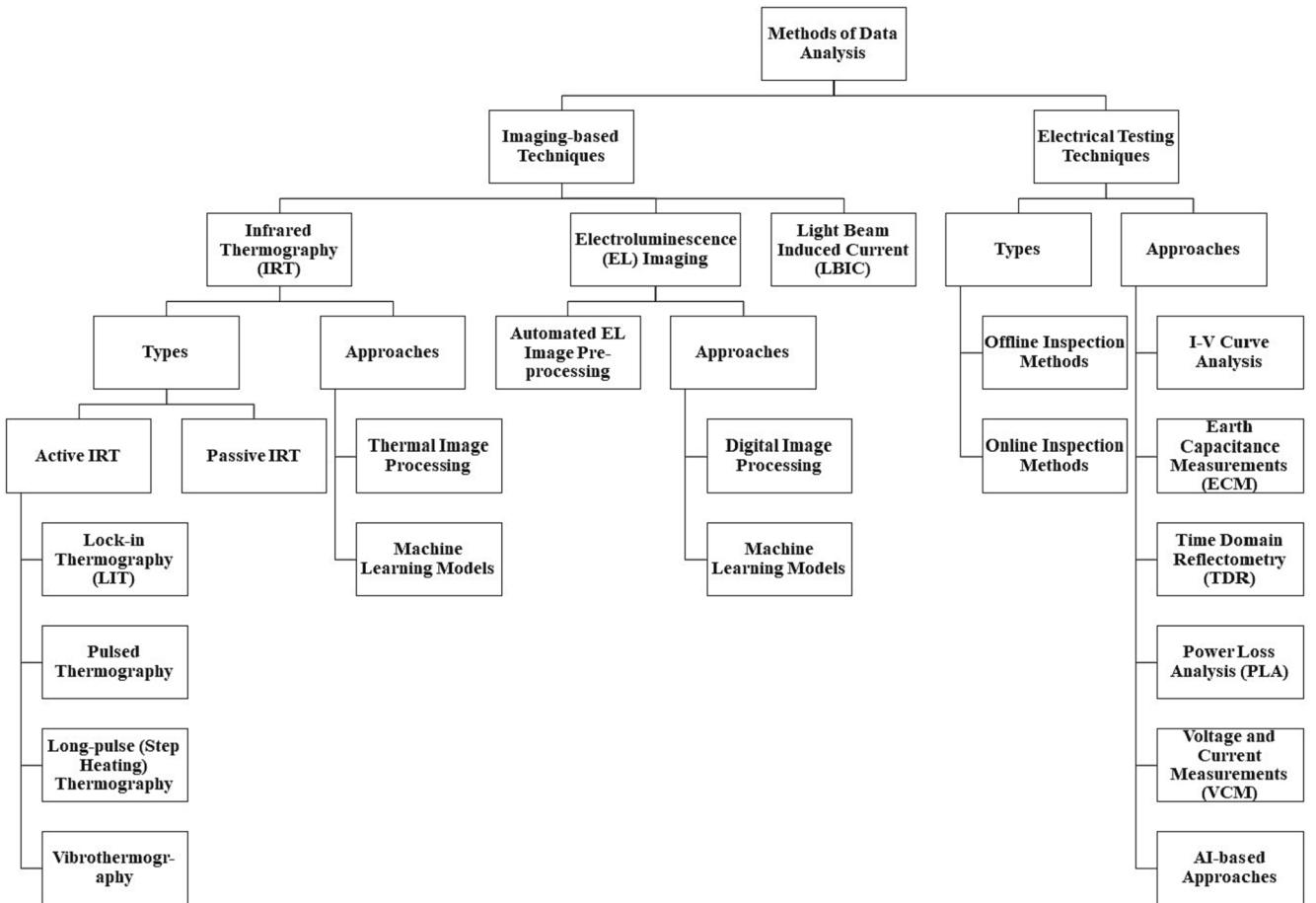


Fig. 2. A high-level overview of data analysis types and approaches for solar PV defect detection systems.

generated during cutting or laser scribing in the industrial fabrication process of solar cells.

- c) Long-pulse, or Step Heating, Thermography: a continuous low-power heat source is applied. This technique assesses the cooling process, whereas the pulsed technique is focused on the heating process. Nevertheless, employing this technique is difficult for analysing large PV modules [16,22].
- d) Vibrothermography: a mechanical excitation, i.e., vibrations, is applied to the module's surface to induce localised heating near defects, as the mechanical energy is converted into thermal energy inside of the material [16]. The resulting thermal patterns on the module's surface during the mechanical excitation can then be captured using an IR camera. By analysing the obtained temperature variations, defects can be identified.
- 2. Passive IRT: the natural IR radiation emitted by the PV module is directly used in this technique to capture thermal patterns on the module's surface, and therefore, it does not require employing an external source of heat. Passive IRT is considered the most widely

adopted technique for inspecting large PV plants as an IR camera is the only equipment required to carry out the inspection making the implementation of this technique simpler and more cost-effective than Active IRT techniques. Examples of IR images of PV modules are shown in Fig. 3.

3.1.1. Approaches based on thermal image processing

Alajmi et al., [23] proposed a thermal energy analysis technique based on Thermal Image Processing (TIP) integrated with voltage sensor-based frameworks for hotspot fault detection and localisation. The study also aimed at identifying whether the hotspot fault was permanent or transient due to environmental factors. Pre-processing steps were first applied including noise filtering using edge-preserving and a smoothing filter, RGB-HSV conversion, and generating the threshold masks for the HSV channels before the fault classification. The proposed algorithm claimed to achieve an accuracy of 100 %.

Similarly, another TIP-based algorithm was proposed by Kurukuru

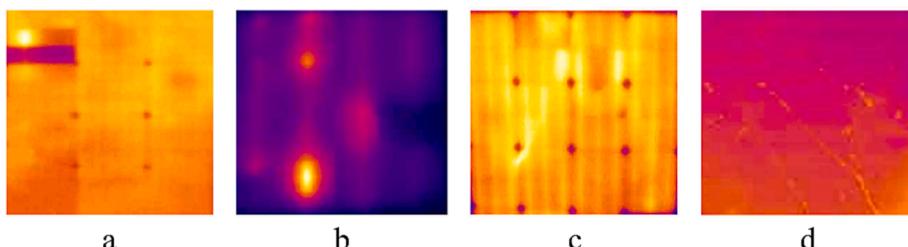


Fig. 3. Defects shown in IR images of PV modules: (a) failed cell interconnection (b) cell cracking (c) localised shunting in cell (d) breakage of module glass [19].

et al., [11] based on Canny edge detection and Hough transform. Feature extraction was first applied to the image, which was then passed through a classification algorithm for localising and identifying the fault type. To produce an Edge Map, noise filtering was first applied using a Gaussian filter followed by an intensity gradient computation. Pixels that did not contribute to the edge were removed using a non-maximum suppression followed by hysteresis thresholding to identify the pixel gradient. To separate the panels, lines in the image were detected by deploying a Hough Transform. A testing accuracy of 93.1 % was obtained.

On the other hand, Jiang et al., [24] have proposed a curve fitting and peak detection technique to reduce noise in thermal images and identify hotspots. The technique is performed on the histograms of the thermal image, followed by applying a threshold to extract hotspots from the background.

Oliveira et al., [15] presented a review of aerial IRT (aIRT) for automatic PV plant inspection summarising Digital Image Processing (DIP) and DL-based algorithms. The review investigated three main aspects of the procedure of automated aIRT: optimisation of the flight path, the *ortho*-mosaicking of the PV plant, and detection of soiling rather than faults, which were found addressed by only a few studies in the literature.

3.1.2. Approaches based on machine learning models

A wide range of traditional and more advanced ML algorithms have been deployed for IRT-based PV module defect detection. Ali et al., [25] have targeted hotspot detection and classification by employing the Support Vector Machine (SVM) model. A hybrid-feature vector was proposed to train the SVM model composed of RGB, texture, the Histogram of Oriented Gradient (HOG), and Local Binary Pattern (LBP). The pre-processing steps consisted of grey-scale conversion, histogram equalisation, and noise filtering (using a Gabor filter). The hybrid feature extraction was applied before SVM training. The SVM model achieved a testing accuracy result of 92 % and outperformed other conventional ML algorithms when compared against Quadratic Discriminant Analysis (QDA), naïve-Bayes (n-Bayes), K-Nearest Neighbour (KNN), and Bagging Ensemble (BE) at higher computational and storage efficiency.

Haidari et al., [26] have modified the VGG16 architecture for PV powerhouse applications, mainly addressing three classes: healthy, hotspots, and substrings. The dataset was collected from an Unmanned Aerial Vehicle (UAV) and hand-portable IR cameras. A total of 1116 thermal images have been manually labelled, of which the majority is of the ‘healthy’ class. In this study, hyperparameters of the original VGG16 network, which consisted of 13 CNNs in 5 blocks [27], have been modified, achieving an overall accuracy of 98 %.

Another CNN-based approach employing passive IRT was proposed by Huerta-Herranz et al., [20]. The study integrated two novel Region-based CNNs (R-CNNs) to detect and localise hotspots. A system composed of an IR camera and UAV (IR-UAV) was deployed to collect not only thermal images indicating hot regions, but also telemetry data to locate hotspots. Examples of such data are Global Positioning System (GPS), coordinates, altitude, orientation, etc. Another relevant study [28] concluded that the GPS approach is more effective than UAV systems based on aerial triangulation [20].

On the other hand, Waqar Akram et al., [19] addressed more advanced IRT-based techniques for PV module defects using transfer learning to improve performance. The CNN model is first pre-trained on EL images, of which the learned features are then repurposed, or transferred, to another target model for a new dataset or task. An IR dataset was collected with normal operation and artificially induced defects of PV modules. The fine-tuning of IR images achieved an accuracy of 99.23 % at a real-time prediction speed.

Nevertheless, several surrounding factors can influence the temperature leading to incorrect variations in thermographic measurements including emissivity variations (i.e., the capacity of bodies to irradiate infrared energy), reflections, vision angle (as an IR camera can provide

different temperatures according to the vision angle), camera malfunctions (caused by the machine operator or due to the IR camera limitations), and environmental interferences (e.g., wind, solar light, and humidity) [16]. Moreover, false fault identification can also occur due to interference from other heat-emitting bodies [11].

3.2. Electroluminescence (EL) imaging

In PV modules, the incident solar energy is converted to electricity at a certain range of Conversion Efficiency (CE). The energy CE of a solar cell is defined as the ratio between the maximum electrical power that can be delivered to the load and the power of the incident radiation over the device [29]. For instance, a commercial cell of a CE of 15 % means that, for a cell surface of 1 m², only 15 W would be delivered to the rest of the circuit for every 100 W/m² [29]. EL imaging is a technique in which the module’s surface defects are inspected by examining the PV module’s areas of varying CE, in which an area of low CE can indicate the presence of defects. In an EL imaging system, an external DC power supply is connected to a PV module in a darkened room, as shown in Fig. 4, to pass through an electric current and exhibit the photoemission of the excited PV module using an infrared-sensitive camera. In the sensed EL image, brighter luminescence is exhibited in areas of high CE, whereas inactive or areas of low light emission indicate defects appearing as dark regions [4,30]. In fact, recent studies have quantitatively estimated electrical performance parameters of solar cells in PV modules using EL imaging. Rajput et al., [31] have computed the I-V characteristics of PV modules, and hence, the efficiency for individual cells based on EL imaging. This was achieved by first extracting the series resistance and dark saturation current density from solar cells.

Sample examples of defects detected in EL images are visualised in Fig. 5. The examples are of different anomalies including cracks, such as the line crack (Fig. 5 (a)) and star crack (Fig. 5 (b)); finger interruptions (Fig. 5 (c)); thick lines (Fig. 5 (d)); scratches (Fig. 5 (e)), and fragments (Fig. 5 (f)). These samples were presented as part of the PV EL Anomaly Detection (PVEL-AD) dataset [32], which contains more than 37,000 near-infrared images with various internal defects and heterogeneous backgrounds.

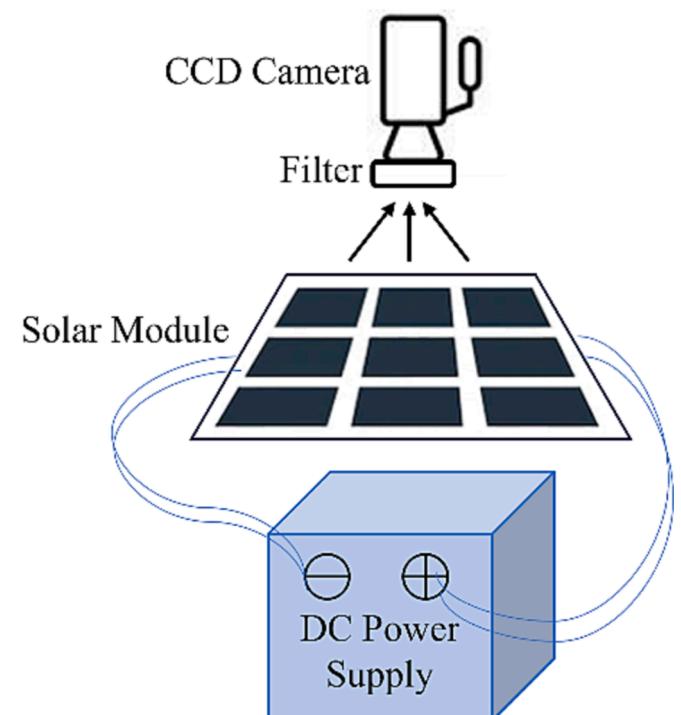


Fig. 4. Basic setup of the EL imaging technique.

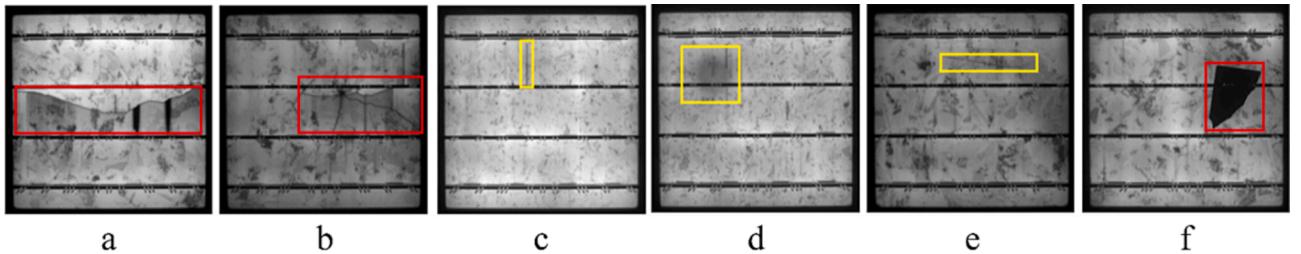


Fig. 5. Defects shown in EL images of PV modules: (a) line crack (b) star crack (c) finger interruption (d) thick line (e) scratch (f) fragment [32].

It is evident that the EL imaging technique is more effective when compared to images captured by typical CCD cameras and lighting systems. In CCD images, automated defect detection cannot be effectively identified due to the random crystal grain surface of a solar cell or interior defects that do not visually appear on the cell surface. An example of CCD and EL images captured from a defective PV module is illustrated in Fig. 6, in which inner micro-cracks and other various defects cannot be detected in the CCD image (Fig. 6 (a)), but can be identified in the EL image (Fig. 6 (b)) [4].

3.2.1. Automated image pre-processing pipeline

To address the importance of high-quality data for higher performance of the detection algorithm, several pre-processing pipelines have been proposed in the literature to facilitate the automated analysis of high-volume throughput of data. Bartler et al., [33] have addressed the automated pre-processing pipeline for EL imaging with a particular focus on highly imbalanced datasets. The first step of the pipeline is correcting the intrinsic distortion caused by the camera lens. This is followed by image segmentation to extract the raw contour of the module, and then, perspective correction given that the module's dimensions and edge coordinates are known. Finally, the cell extraction is performed from the corrected image.

Another automated EL image pre-processing pipeline has been proposed by Fada et al., [34] and Karimi et al., [35] in stages as depicted in Fig. 7. An EL image is first converted to grayscale (Fig. 7 (a)). This is followed by correcting the radial distortion (or the barrel distortion) caused by the inherent property of the camera lens deforming the image. The correction is done employing Computer Vision (CV) via OpenCV and mathematical equations representing the radial distortion division model [35]. Following the distortion correction, perspective transformation is performed. For this step, any noise spikes across the image are reduced using a convolving Median filter of size three on the image.

To identify the module region, the background in the image is removed. A histogram is first used by mapping the spectral colour of the pixel intensity values to the binned colour ranges. This yields a background of colour purple (Fig. 7 (b)). To identify the module region (active region) in the image, the background colour values are then set to 0 (i.e., black) keeping the module region of non-zero values (Fig. 7

(c)) for further processing purposes. The four corners of the module region are then identified using linear regression fits (Fig. 7 (d)). Using the identified four corners of the active region, a perspective-transformed image is produced (Fig. 7 (e)). Finally, solar cell images are extracted by slicing the module image (Fig. 7 (f)), given that the module's dimensions are known. A popular approach for deploying the described image pre-processing pipeline is using the open-source Python coding language with additional packages of NumPy, SciPy, scikit-image, and OpenCV.

3.2.2. Approaches based on digital image processing

Methods based on CV are considered one of the main directions in DIP of EL imaging for automated solar cell defect detection [36]. They have been mainly categorised into three groups: 1) local scheme-based methods, in which the local image features (e.g., gradient, texture, local intensity contrast, etc.) are extracted, and defective regions are then segmented using a threshold. 2) Global scheme-based methods, in which the image is integrally transformed using a Fourier transform, wavelet transform, or matrix factorisation, and the defect detection results are obtained by inverse transformation. Lastly, 3) local-global scheme-based methods, where both local and global scheme-based methods are integrated to achieve a more balanced performance for a variety of defect types [36].

Based on a local scheme-based method, gradient features were adopted by Anwar et al., [37], in which each region is either sharpened or smoothed depending on its gradient value to segment the defective regions from their background. The detection results are then obtained by applying morphological operations.

Based on a global scheme-based method, Fourier image reconstruction has been proposed by Tsai et al., [30], given that defects can be identified as they appear darker than their surroundings in EL images. The Fourier image reconstruction process was applied by first assigning all frequency components associated with defects to zero values. A spatial image representing the 'defect-free' image was then generated by back-transforming the spectral image using the inverse discrete Fourier transform. The 'defect-free' image was then subtracted from the input image to detect defective regions including those with inhomogeneous backgrounds. Another work based on binary and discrete Fourier Transform (DFT) has been introduced by Dhimish and Holmes [38] targeting cracks in EL images.

Based on a local-global scheme-based method, Independent Component Analysis (ICA) was proposed as a signal processing technique. Tsai et al., [4] have adopted ICA to address defects that may not necessarily be detected by extracting distinct local, geometric, and intensity features, as well as the computational cost required for large, high-resolution EL images. In this study, the ICA technique was composed of two stages: learning and inspection. In the learning stage, a set of 2-dimensional (2D) defect-free cell images was reshaped into 1-dimensional (1D) signals to find a set of statistically independent basis images. In the inspection stage, the basis images from the learning stage were used to reconstruct a test solar cell image as a linear combination. To detect defects, the deviation between the test image and the reconstructed one derived from the ICA basis images is then evaluated by computing the reconstruction error. Limitations of the proposed method

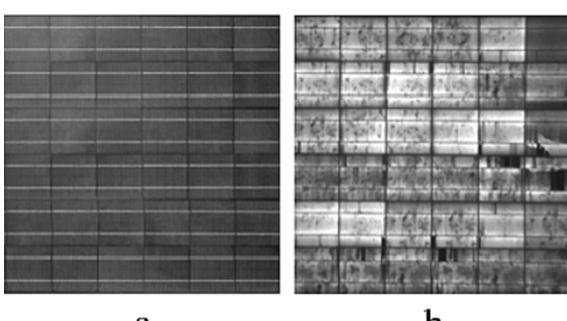


Fig. 6. A defective 36-cell PV module as a (a) CCD-captured image and an (b) EL-captured image [4].

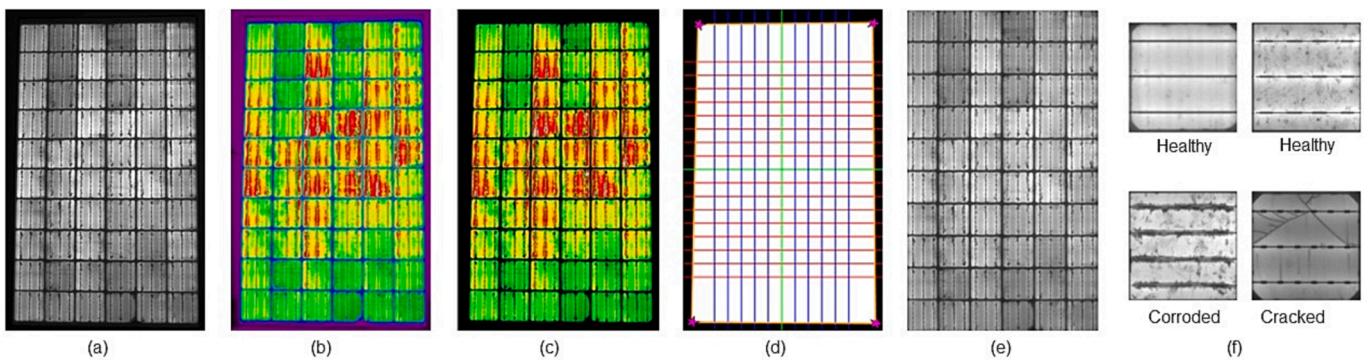


Fig. 7. An automated EL image pre-processing pipeline for solar cell defect detection [18].

include a lack of ability to identify the shape and location of defects.

Moreover, a common characteristic found in EL images is the heterogeneously textured background in the cell images making automatic crack detection a challenging task due to the low contrast between the crack and the surrounding background. Heterogeneity in the background texture is mainly formed due to the randomly distributed nature of crystal grains in multi-crystalline solar cells. To address this challenge, Chen et al., [39] have first developed a steerable evidence filter to improve the contrast between the crack and its surroundings. Second, the complete crack was extracted based on segmentation using a local threshold. Subsequently, the isolated non-crack pixels were removed using morphological operations.

3.2.3. Approaches based on machine learning models

Fada et al., [34] have addressed the problem using statistical learning methodologies and compared the training performance between SVM, Random Forest (RF), and Artificial Neural Networks (ANNs). A Multilayer Perceptron (MLP) was adopted as the ANN method (MLP-ANN). The targeted defect types were cracks and busbar corrosion. The results showed that SVM achieved a slightly higher prediction accuracy than the MLP-ANN model, whereas the RF performance ranked the lowest. The same work was improved afterwards by Karimi et al., [35] in another study using a CNN rather than an MLP-ANN, and it was shown that CNN outperformed the performance of SVM and RF.

Deitsch et al., [40] have also compared the performance of SVM and CNNs in supervised classification frameworks for solar cell defect detection. Both approaches differed in the hardware efficiency needed, as SVM yielded a performance accuracy of around 82 % with low computational requirements, whereas CNN achieved a more accurate performance of around 88 % but required a Graphics Processing Unit (GPU). The study also contributed with a dataset of around 2600 annotated EL solar cell images. Nevertheless, as available data was limited, transfer learning was adopted for the CNN training by adapting the VGG19 architecture [41] originally trained on the ImageNet dataset to the EL dataset.

Bartler et al., [33] have addressed the application of CNNs for solar cell defect detection using EL imaging for the first time with special care to imbalanced datasets. The study tackled a binary classification task adapting the VGG16 architecture by reducing the number of filters and fully connected layers, hence, the total number of parameters. The performance yielded a False Negative Rate (FNR) of around 13 %.

The authors in [42] portray a DL-based PV detection system using Generative Adversarial Networks (GANs). The system first generates a dataset of high-resolution EL images using a low number of existing images. Following this, a CNN classifier is trained on the generated dataset for feature extraction. It is claimed that this approach can boost classification robustness with 83 % classification accuracy, where the system is tested against common models including MobileNet, Inception V3, VGG16, and ResNet50.

Moreover, Akram et al., [43] developed a lightweight CNN

classification system for defect detection. Focusing on EL images, data augmentation is performed to compensate for sparsity in available data. The light CNN model is trained on the publicly available dataset and then evaluated and tested, achieving up to 93.02 % accuracy. In addition, the model can classify one EL image in about 8 ms. However, the authors mentioned the issue of overfitting impacting the overall robustness of the model.

In the same light, Chen et al., [44] developed a fast PV defect diagnosis system to detect multiple defect types. The system uses a multi-algorithm EL image classifier for training. The PVLP dataset, of around 2.4 M cell images, has been trained over the Random Forest, ResNet, and YOLO models. Current results show a classification accuracy of 84 % and an inference time as low as 0.5 s.

The work of Fiorese et al., [45] provided an example of the combined use of image processing and ML for PV defect diagnosis. A model was proposed that can differentiate between four defect types in multi-crystalline and monocrystalline cells. Using a combination of DeepLabv3 and ResNet-50, the model is trained on more than 17,000 images. Test results achieved an accuracy of approximately 95 %. It is worth noting that some images in the training are generated to mitigate class imbalance.

A classification system based on a Complementary Attention Network (CAN) was proposed by Su et al., [46]. This method was used to mitigate noise features and emphasise defective features accordingly. After training the model on a 3600 EL-image dataset, a classification accuracy of up to 97 % was achieved.

3.3. Light beam induced current (LBIC)

The Electron Beam Induced Current (EBIC) technique is deployed by measuring the electric current variations in the solar cell induced by an electron beam that is emitted by a Scanning Electron Microscope (SEM) [47]. The technique can accomplish a high detection spatial resolution, yet it is limited by the SEM's lens properties and impractical vacuum atmosphere requirements amongst others.

Therefore, overcoming the limitations of the EBIC, the LBIC technique has been proposed in the literature as non-intrusive by nature. LBIC can potentially yield comprehensive diagnoses for structural and process-based solar cell defects. Unlike EBIC, this method flows photo-generated current in solar cells by scanning the solar module's surface using a focused light beam from an isolated source while simultaneously measuring the generated photocurrent. A basic LBIC setup is illustrated in Fig. 8.

LBIC is also configurable by wavelength, light intensity, and voltage bias, which facilitates a more comprehensive analysis of defects in various solar cells. Nevertheless, limitations lie in the need to mechanically stabilise the laser beam to ensure accuracy in addition to the relative slowness of the technique to detect defects [47].

Considering the constraints of conventional LBIC, Gan et al., [47] proposed an improved method for anti-noise defect detection by

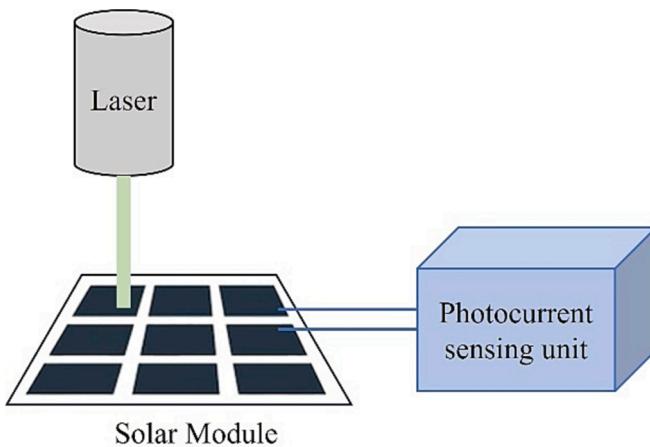


Fig. 8. Basic setup of the LBIC imaging technique.

employing orthogonal light points in a pattern-controlled structure resulting in a more robust and accurate performance. The deployment also involved software-based signal processing to remove erroneous parts of the captured signals to produce a more accurate current map of the PV unit. An experimental setup was developed in a dark room involving an industrial projector and a polycrystalline solar cell. Defects were simulated on the cell by covering certain areas of the cell. An Analog-to-Digital Converter (ADC) was used to measure the resultant current and feed it into a MATLAB program for further processing. The method signified that the improved LBIC technique detects faults more accurately and consistently compared with conventional LBIC.

With the use of image processing techniques, Quan et al., proposed a Compressive Sensing-based LBIC (CS-LBIC) technique to detect solar cell defects [48]. The method aimed to resolve the complexity problem caused by the PV modules' busbars and fingers, which increase the complexity of signals captured of normal CS-LBIC and may affect the detection accuracy. Hence, an optical camera was introduced to capture an image of PV units and feed it to a computer program that extracts busbar and finger structures, resulting in a simpler defect detection problem. The experimental validation focused on the detection speed, which is considered about five times faster than the reported literature, achieving a detection speed on the order of milliseconds.

A critical analysis comparing different IBTs with respect to their advantages and disadvantages is discussed in Section 5.

4. Electrical testing techniques (ETTs)

ETTs assess the electrical characteristics and performance of the modules by comparing anomalies with normal conditions and identifying their deviations to indicate potential faults or defects. ETTs are primarily of two types: offline and online fault inspection methods [49]. Common ETTs utilised in the literature for fault detection in PV systems can be categorised into: Current-Voltage (I-V) Curve Analysis, Earth Capacitance Measurements (ECM), Time Domain Reflectometry (TDR), Power Losses Analysis (PLA), Voltage and Current Measurements (VCM), and AI-based approaches.

To prevent PV modules from accidents, electronic protection devices, such as fuses and circuit breakers, are traditionally used on the DC side to protect against over-current faults, grounding faults, and arcing faults [49,50]. Moreover, Maximum Power Point Trackers (MPPTs) are applied in PV systems to optimise the power generation whenever there is a drop in power such that maximum power can be delivered [51]. However, MPPTs may impede correct fault detection with the electronic protection devices when the output current and voltage of the PV system deviate from those of the normal condition. Therefore, offline inspection methods have been adopted in several studies to overcome limitations imposed by MPPTs. However, these methods require interrupting the

normal operation of the PV system and are usually carried out overnight. On the other hand, online fault detection is proposed in the literature addressing operational PV systems under MPPT conditions and involving continuous real-time monitoring of PV modules' health. Moreover, AI-based approaches are proposed in the literature considering steady-state and time-domain analysis methods [49].

4.1. Current-voltage (I-V) curve analysis

Current mismatches lead to a decrease in the generated power, causing distortions in the I-V characteristic curves of the PV module. A particular work by Zhang et al., [52] explored current mismatch faults, particularly associated with partial shading, hotspots, and cracks of PV modules, and evaluated the modules' I-V characteristics. A numerical analysis of the collected test data is then concluded to perform statistical diagnostics of low- and high-voltage areas of the curve. The performance indicators 'fault detection rate' for correctly identified faults and 'false detection rate' for falsely identified faults were used for evaluation.

Wang et al., [53] utilised the dynamic I-V characteristics to determine the PV panel's intrinsic parameters while tracking the Maximum Power Point (MPP) simultaneously using a nonintrusive fault diagnosis technique. The estimated parameters were then communicated to the central control unit via a Power Line Communication (PLC) module. The study utilises four 80-W PV panels, of which two are healthy, and the other two have different levels of crack damage. After testing the proposed approach, results showed a significant drift in the parameters of the cracked panels from their original values indicating the presence of a panel failure.

Similarly, another contribution that utilised MPP tracking was depicted by Solórzano and Egido [54]. A fault diagnosis system was proposed at the module level, namely DC-to-DC converters, and inverters. A prototype built in MATLAB and tested on a building's rooftop has demonstrated the successful detection of fixed object shading, generalised and localised dirt, potential hotspots, PV degradation, and significant DC cable loss.

Analogously, it was argued that automated monitoring systems are significant for PV yield evaluation, and considerable losses can be avoided if fault detection models were put in place in industrial production plants. In this context, an I-V curve-based fault detection model was devised with the ability to detect short circuits, partial shading, and dust faults, as well as ageing [55]. The solution was evaluated in MATLAB and experimentally. As the methodology is model-based, the model for the PV module under test is developed and calibrated with its current electrical parameters using MATLAB-Simulink. Following this, the behaviour of a reference I-V curve was simulated. The simulated curves were then compared against the experimental I-V curve measurements in each faulty case. Consequently, the faults were diagnosed by the algorithm by observing the different fault signatures on the I-V curve [55].

4.2. Earth capacitance measurements (ECM)

Earth Capacitance Measurements (ECM) are usually applied to detect the disconnection position on power lines. For PV modules, the concept of this method relies on the capacitance measurement between the module's terminal and the grounding terminal to detect capacitance formed due to the presence of defects [56]. Therefore, this technique is independent of irradiance changes. For example, Takashima et al., [56] have developed a fault diagnosis system targeting the PV array to detect disconnections. Takashima et al., [57] have also addressed the detection and location of disconnections in strings between the PV modules by measuring the earth capacitance. The value was then compared against a threshold value of undefective strings.

4.3. Time domain reflectometry (TDR)

On the contrary of measuring capacitance, the Time Domain Reflectometry (TDR) method relies on the measurement of impedance change between an input signal to the PV module's strings and the reflected signal that is caused by the presence of a defect or an abnormality [58]. The reflected signal is then compared against the reference signal in terms of different electrical characteristics, including 1) the signal shift from the input signal indicating the fault position in the string, 2) the waveform change indicating the mismatch nature (e.g., open circuit, short circuit, resistance increase, etc.), and 3) the impedance change indicating the severity of the fault [57]. To detect impedance mismatches in the string, Takashima et al., [57] have proposed a TDR-based diagnosis method for periodic inspections using a step voltage input signal. When the study compared TDR against ECM, TDR was found more appealing as not only the disconnection was detected, but also the impedance changes with degradation. However, efficient fault detection based on reflectometry can be limited by the frequent variations of impedance throughout the PV array. This is due to the occurrence of multiple reflections at different mismatches leading to difficult interpretation of the time-domain reflection to detect the presence of ground faults in PV arrays. To address this challenge, Roy et al., [59] have proposed the Spread Spectrum TDR (SSTDR) diagnosis method that is independent of fault-current magnitudes and the presence of solar irradiation. In principle, the SSTDR method uses a pseudorandom binary signal referred to as the pseudonoise code (PN code) as an incident signal consisting of randomly generated ones and zeroes (each of which is referred to as a chip). The incident signal is generated by modulating the PN code with a carrier sine wave, of which its frequency is known as the centre frequency of SSTDR. For easy system calibration, the frequency of the carrier sine wave and the chip rate of the PN code are maintained the same. By using a variable phase-delay generator, the reflected signal is cross-correlated with the delayed copies of the incident signal. The characteristic impedance at the load terminal is then observed. In case a mismatch is present, a lobe at a time delay that corresponds to the distance from the source terminal is produced in the autocorrelation plot [59].

4.4. Power losses analysis (PLA)

In this technique, the actual power output of the PV module is assessed by comparing it with the simulated power output of a healthy PV module. The resulting error deviations are then analysed to identify attributed defects [13]. Stauffer et al., [60] have proposed a self-adaptive PV production model to continuously monitor the produced PV power against the expected power. In case of detected discrepancies, a warning flag is raised. However, the fault type and location were not determined. Statistical methods were applied by Dhimish and Holmes [61] on the measured PV power data for an automatic fault detection system. These data were then compared with the simulated theoretical performance using a statistical *t*-test. Moreover, the ratio between the measured and theoretical DC power and voltage was monitored to identify the fault location.

On the other hand, an integrated solution based on PLA and I-V curve analysis was proposed by Chen et al., [62] for an online and real-time fault diagnosis system. The simulated PV module's power during the operational status was compared with the measured value taking into consideration the variations of temperature and irradiation. As a result, abnormalities in the power loss values were identified. The change of output voltage was then assessed to determine whether a short circuit fault had occurred and the number of short-circuited cells.

4.5. Voltage and current measurements (VCM)

Under low irradiance conditions and with a running MPPT, DC-side short circuit faults can be nearly undetectable. Addressing this

challenge, a signal processing-based approach by means of utilising a signal decomposition pattern recognition system has been proposed by Yi and Etemadi [63]. A data collection setup for measuring the output voltage and current signals was implemented for a grid-connected PV system to extract their features. This is followed by a fuzzy inference system to inform of a fault occurrence, particularly line-line and line-ground faults.

4.6. Artificial intelligence (AI)-based approaches

Recent research has demonstrated the capabilities of AI-based techniques for PV module fault diagnosis systems using ETTs. ANNs were used in [64] with genetic algorithms as an optimisation methodology for dynamic diagnosis and repair. The study claimed that the proposed algorithm overcame cost and management limitations in complex PV systems and outperformed other diagnostic PV systems when compared with fuzzy-based and traditional neural networks.

An MLP neural network structure was proposed by Mekki et al., [65] to detect several shading patterns. In this analysis, electrical and non-electrical parameters including solar irradiance, current, voltage, and temperature data of a given PV cell, were used. To diagnose faults under MPPT conditions, a CNN-based fault detection system with sequential current and voltage data of transient in the time domain was developed [49]. The network was able to extract useful features from the collected electrical time-series data and classify the features according to the selected fault labels: open circuit on two strings, open circuit on one string, 16.67 % mismatch short circuit, 33.33 % mismatch short circuit, and normal. After validating the system on installed PV panels, an average of 99 % accuracy was achieved.

5. Discussions

This paper reviews all analysis methods of imaging-based and electrical testing techniques for solar cell defect detection in PV systems. This section introduces a comparative analysis of the surveyed studies in the literature. Moreover, a critical analysis of the presented techniques is discussed in terms of their advantages and disadvantages.

5.1. Comparative analyses of surveyed studies in the literature

In general, the most prominent types of IBTs are EL imaging and IRT, whereas common types of ETTs are based on I-V Curve Analysis and PLA. Recent studies have shown the high capabilities of advanced ML models, such as CNNs, R-CNNs, ResNet, YOLO, and VGG16, to identify defect types when compared against traditional and digital signal processing techniques. Nevertheless, few studies have addressed defect localisation within the PV module in large-scale applications.

A comparative analysis of surveyed studies is summarised in Table 1 with respect to the adopted data analysis method, technique type, type of detected defects, and performance results. The disparity in the volume of the reviewed literature related to each of IBTs and ETAs can be rationalised by the sheer volume of relevant studies available on some techniques. For example, literature relevant to some of the ETT-based techniques, such as TDR and ECM, was found to be relatively very limited compared to other techniques. On the other hand, a greater depth of knowledge and research was undertaken in the literature relevant to other ETT and IBT-based techniques. On the IBT level, it can be observed from Table 1 that the most notable techniques adopted in the literature are IRT- and EL-based, whereas the least are LBIC-based.

5.2. Advantages and disadvantages of imaging-based techniques

Different IBTs were adopted in the literature for defect detection in PV modules. The advantages and disadvantages of each technique differ in terms of the setup, scale of deployment, and data processing, as summarised in Table 2.

Table 1

Reviewed papers on defect detection and diagnosis methods for PV systems.

Data Analysis Method	Method Type	Reference	Year	Technique	Fault/Defect Type	Best Results
Imaging-Based Techniques (IBTs)	IRT	[11]	2019	TIP	Hotspot.	Acc (testing): 93.1 %
		[19]	2020	CNN	Binary classification (defective, non-defective).	Acc: 99.23 %
	EL	[20]	2020	R-CNNs	Hotspot.	Acc: 99 %
		[23]	2019	TIP	Permanent hotspot, Transient hotspot.	Acc: 100 %
		[25]	2020	SVM	Non-faulty hotspot, Faulty PV panels.	Acc (testing): 92 %
		[26]	2022	CNN (VGG16)	Hotspot, Hot substring.	Acc: 98 %
		[33]	2018	CNN (VGG16)	Binary classification (defective, non-defective).	FNR: 12.96 %
	DIP	[34]	2017	SVM, RF, MLP-ANN	Cracks, Busbar corrosion.	RF Acc: 96.90 % MLP-ANN Acc: 98.13 %SVM Acc: 98.77 %
		[35]	2019	SVM, RF, CNN	Cracks, Busbar corrosion.	RF Acc: 97.46 % SVM Acc: 99.43 %CNN Acc: 99.71 % Acc: ~99 %
		[36]	2017	DIP	Binary classification (defective, non-defective).	
		[37]	2014	DIP	Micro-cracks.	Acc: 88 % Sensitivity: 97 %
		[30]	2012	DIP	Binary classification (defective, non-defective).	
		[4]	2013	DIP	Binary detection (defective, non-defective).	Mean recognition rate: 93.4 %
		[39]	2019	DIP	Binary detection (cracks, non-defective).	Average detection rate: 94.4 %
		[40]	2019	SVM, CNN (VGG19)	Binary classification (defective, non-defective).	SVM Acc: 82.44 %CNN Acc: 88.42 %
		[42]	2020	CNN	Micro-crack, Finger-interruption, Break.	CNN Acc: 83 %
		[43]	2019	CNN	Binary classification (defective, non-defective).	CNN Acc: 93 %
		[44]	2022	RF, ResNet-18, YOLO	Cracks, Intra-cell defects, Oxygen-induced defects, Solder disconnections.	ResNet-18 Acc: 84 %YOLO Acc: 78 %
		[45]	2022	Deeplabv3, ResNet-50	Cracks, Contact interruptions, Cell interconnect failures, Contact corrosion.	Acc: 95 %
		[46]	2021	Complementary Attention Network (CAN).	Line-like cracks, Star-like cracks, Line-like finger interruption, Blob-like black core defects.	Acc: 97 %
Electrical Testing Techniques (ETTs)	LBIC	[47]	2022	–	Binary classification (defective, non-defective).	
		[48]	2019	–	Binary classification (defective, non-defective).	Detection speed ~ 5x faster than literature
I-V Curve Analysis	Online	[52]	2021	Online	Non-mismatch, Partial shading, Hotspots, Cracks.	Fault detection rates: 91.08–100 % False detection rates: 0–2.45 %
		[53]	2016	Online	Cracks (two different damage levels).	–
		[54]	2013	Online	Fixed object shading, Localised dirt, Possible hotspots, Module degradation, Excessive losses in DC cables.	Comparison of measured parameters
		[55]	2018	Online	Uniform dust faults, Partial shading faults, Short circuit faults, Ageing.	Comparison of measured vs. reference parameters
		[57]	2009	(n/a)	Disconnection between strings, Impedance mismatch (resistance increase)	Comparison method
TDR PLA	Online	[59]	2018	–	Ground fault.	
		[60]	2015	Online	Binary detection (defective, non-defective).	Comparison method
		[61]	2016	Online	Faulty PV module, Faulty PV string, Faulty MPPT unit.	Detection error: ~ 0

(continued on next page)

Table 1 (continued)

Data Analysis Method	Method Type	Reference	Year	Technique	Fault/Defect Type	Best Results
		[62]	2016	Online	Partial shadow, Short circuits, Ageing.	Comparison method
VCM		[63]	2017	Online	Line-line fault, Line-ground fault.	Acc: 84.1–96.1 %
AI-based		[49]	2019	CNN (Online)	Line-line fault, Open circuit fault.	Acc: 99 %
		[64]	2015	ANNs with GAs	Degradation state, Short circuit, Open circuit, Shading, Charger failure, Battery failure.	Acc: 98.9 %
		[65]	2016	MLP (Online)	Partial shading.	Comparison method (estimated parameters: correlation coefficient and mean absolute error).

Table 2

Advantages and disadvantages of IRT and EL-based imaging for solar PV defect detection systems.

Imaging-based Technique (IBT)	Advantages	Disadvantages
Infrared Thermography (IRT)	Large-scale applications. Non-destructive technique. Real-time imaging. Efficient detection of hotspots.	Expensive. Experienced labour is needed. External factors (e.g., emissivity variations, reflections) may affect accuracy. Reliability affected by sunlight.
Electroluminescence (EL) Imaging	Defect localisation in the solar cells. Non-invasive. Techno-economically effective. Real-time imaging. Defect detection of internal solar cells.	Challenging for large-scale applications. Limited by optical degradation (e.g., delamination or glass breakage).
Light Beam Induced Current (LBIC)	Non-intrusive technique. Defect localisation in the solar cells.	Long processing time. Accuracy affected by unstable laser beam. Impractical for industrial applications due to long processing time.

According to several studies, IRT techniques are notably efficient in detecting hotspots amongst other solar cell defects [58]. They are more suitable for large-scale applications than EL imaging, as IRT is a Non-Destructive Testing Technique (NDTT), and can be deployed with minimal instruments and at a higher level of safety [19,66]. IRT can also support real-time monitoring during data acquisition [16,66]. Nevertheless, the IRT setup can be expensive due to the costs of professional and reliable IR cameras needed for accurate measurements. Moreover, the configuration of parameter settings in IR cameras requires expert or trained operators. Moreover, there are other external factors that may affect the measurement accuracy, such as emissivity variations and reflections. The reliability of IRT techniques can also be affected by sunlight [58].

On the other hand, EL imaging is a non-intrusive technique that is highly efficient in localising defects, such as micro-cracks, cracks, and shunts [19]. Moreover, defects in internal solar cells can be easily detected from EL images [58]. It is also considered techno-economically effective and appealing for small-scale deployments [18]. However, limitations in EL imaging include detecting optical degradation, such as delamination or glass breakage [66]. In addition, EL imaging can be practically challenging for large-scale outdoor deployments due to the involvement of a large power supply making the setup complex and

costly in terms of energy consumption [19]. It is also commonly deployed outdoors overnight, or indoors, as the crystalline silicon luminescence signal is several orders of magnitude lower than sunlight [67].

Several studies have targeted these challenges in EL imaging by developing accurate defect detection systems under daylight and outdoor conditions. For instance, Johnston and Silverman [68] used InGaAs uncooled and InSb-cooled detector cameras. Other studies [67,69] proposed drone-based daylight EL imaging systems using fast pulsed EL imaging with InGaAs detector cameras. Guada et al., [70] have proposed the bias switching method using an InGaAs camera.

As LBIC is based on the measurement of the induced current under excitation, the main advantages include the ability to provide a comprehensive evaluation of a solar cell's performance by varying parameters, such as wavelength, voltage bias, and light intensity [71]. Nevertheless, LBIC is impractical for industrial applications, as it is limited by the relatively long processing time [71]. It also exhibits slowness in detecting defects as the laser beam needs to be mechanically stabilised to ensure accuracy [47].

5.3. Advantages and disadvantages of imaging-based and electrical testing techniques

Indeed, both categories of data analysis methods provide advantages depending on the requirements of the application, as summarised in Table 3. IBTs depend on analysing the deviations of optical properties, thermal patterns, or other visual features. Unlike ETTs, IBTs are considered more non-intrusive, as they do not require any physical contact with the PV module nor need to cut the electrical connection of the PV system to the grid for measurement acquisition. Moreover, IBTs provide high-resolution data allowing for the detection of small defects

Table 3

Advantages and disadvantages of Imaging-based and Electrical Testing Techniques for PV defect detection systems.

Data Analysis Method	Advantages	Disadvantages
Imaging-based Techniques (IBTs)	• Non-contact and non-intrusive. High-resolution of small defects. Visual representation of structural defects.	• Electrical faults may not be detected. Equipment cost. Knowledge-based.
Electrical Testing Techniques (ETTs)	• Defect detection of internal PV system or electrical faults.	• Intrusive. Lower resolution. Fault detection challenges under MPPT conditions. Equipment cost.

that may be missed by other techniques. In addition, they provide a visual representation of the module surface making it easier for operators to interpret and diagnose defects including structural defects (e.g., cracks, delamination, hotspots, etc.).

However, IBTs have a limited ability to detect defects in the internal PV system or electrical faults, as they can only detect defects that are located on or near the surface of the module. Moreover, IBTs require specialised equipment, such as cameras and software, making it expensive. The accuracy of IBTs can also be affected by the skill and experience of the operator.

On the other hand, ETTs depend on comparing the deviations of the measured operating states or electrical parameters of the modules from the expected (i.e., simulated) electrical behaviour. Unlike IBTs, they can detect electrical faults beyond the PV module's surface (e.g., bypass diode failures, faulty interconnections, open circuits, short circuits, etc.) focusing more on the electrical integrity of the module rather than external defects.

However, ETTs are considered intrusive requiring the module to be interrupted and physically disconnected for inspection. They also have a limited ability to detect small or subtle defects. Depending on the plant size, ETTs may require the deployment of a large number of sensors including electrical and meteorological sensors increasing the cost of the system [52]. Other challenges associated with electrical signal measurements include real-time monitoring of PV modules under MPPT conditions, which may impede fault detection with the presence of electronic protection devices when the PV system's output current and voltage deviate from the expected condition.

Overall, IBTs can provide a high-resolution investigation of external defects on or near the module's surface level, whereas ETTs utilise electrical parameters to assess internal electrical faults. The adoption of each of these techniques depends on the deployment scale (e.g., plant size), the targeted defects for detection, and the required location of defect analysis in the PV system.

6. Current challenges and future orientations

6.1. Challenges in State-Of-The-Art

Based on the conducted review of prior work related to defect detection in PV systems, the main challenges can be summarised as illustrated in Fig. 9. One of the key challenges is data availability and obtaining high-quality data for training ML algorithms. The performance of any defect diagnosis method heavily relies on the quality of the available processing data. However, high levels of noise, autocorrelation, and errors in PV measurements can mask important features in data [12]. Moreover, the effectiveness of a defect diagnosis system can be limited by inverters, as they do not normally provide precise

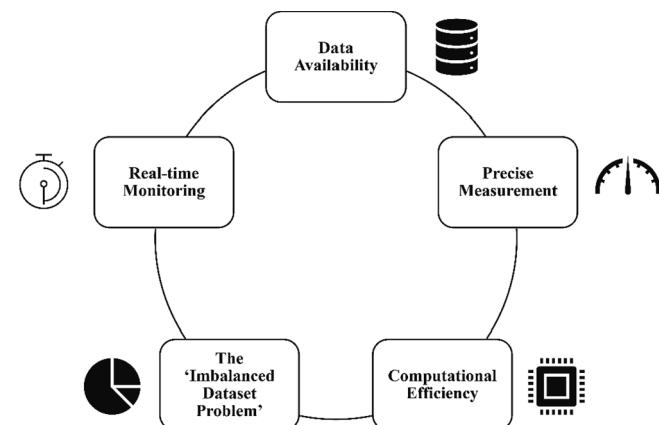


Fig. 9. Main challenges of defect detection in PV systems.

measurements, which necessitates the deployment of additional sensors to collect more accurate data [14]. These limitations hinder efficient monitoring techniques for accurate defect detection.

Although data availability improves the performance of defect diagnosis systems, big data or large training datasets can degrade computational efficiency, and therefore, the effectiveness of these systems. This limits the deployment of DL-based techniques in practical applications with big data. To address this challenge, different frameworks are considered in the literature for reducing data size in future studies, such as K-means metric, Hierarchical K-means clustering, and Euclidean distance [12].

Real-time monitoring is significant for deploying practical applications of PV system defect detection techniques. Nevertheless, it requires integrating these techniques, especially the ones based on AI, into cost-effective circuits, which can be challenging with complex and computationally demanding algorithms hindering high-performance accuracy.

On the other hand, the 'imbalanced dataset' problem remains a standing issue, in which the data distribution of different class categories is highly unequal leading to data scarcity of certain classes [72]. As defects are relatively rare and random amongst the distributed modules, it is commonly observed that the 'defect' class is the minority, whereas the 'healthy' class is the majority. This affects the robustness of the ML model in classifying the minority class and yields overfitting or bias towards the majority classes. One of the most notable solutions considered in the literature is adopting data augmentation approaches in addition to cost-sensitive learning, imbalanced learning, and other models applied as inputs into existing DL-based models.

6.2. Data augmentation techniques

To address several challenges that were presented in Section 6.1, data augmentation approaches have been investigated in the literature using different techniques. These approaches can be utilised to combat the 'imbalanced dataset problem' (i.e., unequal data distribution), availability of high-quality data, and availability of precise measurements by enhancing both the size and the quality of the training dataset. In general, data augmentation approaches can be grouped into two main categories: classical image manipulation techniques, and ML-based techniques, as illustrated in Fig. 10.

6.2.1. Classical image manipulation techniques for data augmentation

One of the main traditional techniques for data augmentation is increasing the size of the dataset by applying simple manipulations, such as geometrical transformations, colour-space transformations, and kernel filtering, on existing images. Examples of geometric transformations include flipping, rotation, translation, noise injection, and random cropping [72]. Colour-space transformations can be used to address lighting challenges found in images, such as altering the values of pixels by a constant value, segregation of the RGB matrices, or aggregation of the RGB matrices to produce a colour histogram and apply filters. Finally, kernel filters can be applied to blur images to increase resistance to motion blur or sharpen images using contrast edge filters to manifest details of interest.

In recent studies, Bartler et al., [33] first randomly oversampled the minority class images by generating random copies to balance the class distribution. Nevertheless, overfitting was easily produced due to the presence of duplicate images. To overcome overfitting, data augmentation was further performed by applying geometric transforms on the oversampled images. Such transforms included random horizontal and vertical image flipping, rotation, translation, and shearing performed within limited, pre-defined ranges. The study [33] has also compared the effect of oversampling only and when integrated with augmentation transforms. Evidently, integrating oversampling and data augmentation does not only increase the minority class size, but also improve the generalisation capability or diversity and robustness of the algorithm for unseen data. Another example of applying geometric augmentation

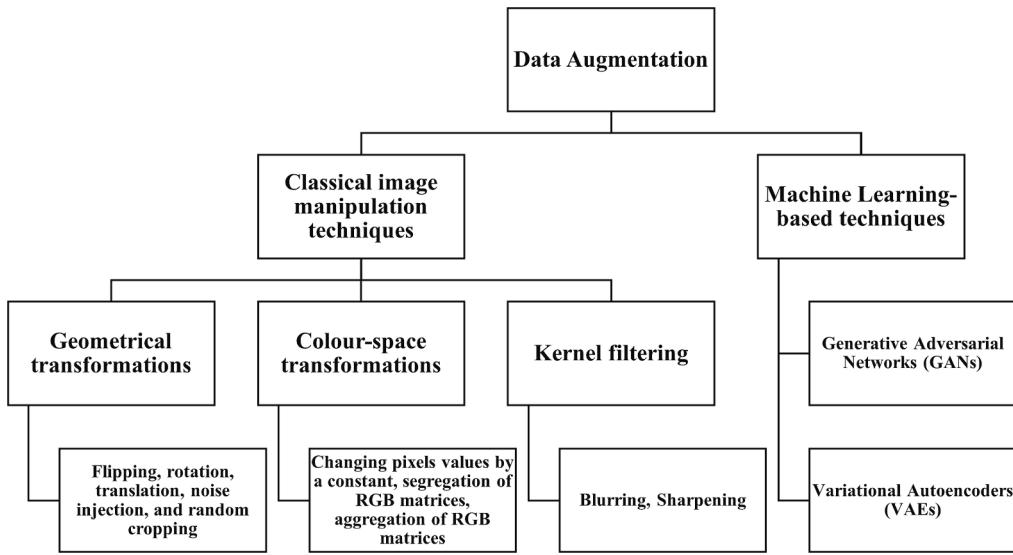


Fig. 10. Image augmentation approaches.

transforms on EL imaging was introduced in [40], where rotation, random flipping, and translation were performed within limited ranges. Similarly, Fada et al., [34] have augmented EL cell images to generate four times the dataset size by means of flipping about the x and y axes and rotation at 180 degrees.

Although image manipulation techniques can reduce positional biases and be easily implemented, they require memory resources, high computational costs, and increasing training time. Moreover, they can be non-label preserving techniques as they may change the meaning of a sample if not properly verified. For example, rotating an image with the class of number ‘6’ to ‘9’ can lead to another meaning interpretation. In addition, a colour-space transformation may neglect important colour information [33,72].

6.2.2. Machine learning approaches for data augmentation

On the other hand, more complex DL-based techniques have emerged in recent studies, such as Generative Adversarial Networks (GANs), as state-of-the-art algorithms for data generation [73]. GANs have demonstrated high capabilities in different applications including synthetic data generation, classification, and regression for both semi-supervised and unsupervised learning. A GAN mainly consists of two neural networks: a generator and a discriminator. Both components run in competition with one another through deriving backpropagation signals to produce synthetic data that are indistinguishable from real data in the training set [74]. Eventually, GANs can estimate the density or distribution of real data, and new data can be generated accordingly

[75]. As illustrated in Fig. 11, the generator acts as the ‘forger,’ as it is trained to produce forgeries by mapping a noise sample to a synthetic data sample that can ‘fool’ the discriminator. The discriminator acts as the ‘judge,’ as it is trained to receive both real samples and forgeries and try to distinguish them apart. This is an iterative backpropagation process, in which the generator aims to produce realistic data [73,74].

To target data generation of high-resolution images in the dataset, Luo et al., [76] have adopted a GAN variant, called the Progressive Growing of GANs (PGGANs) [77], to improve the CNN classification performance for a defective solar cell in EL images. In this study, the GAN model was first trained with low-resolution samples followed by increasing the size progressively to yield more stable training results. Romero et al., [73] proposed Deep Convolutional GANs (DCGANs) to generate a synthetic dataset starting from a small EL image set of PV cells. On the other hand, to compare images, the Mean Square Error (MSE) method was used by several studies as the loss function, which can impose inaccuracy in describing real data and cause image distortion. To address this problem, Shou et al., [78] have proposed a GAN-based detection system that calculates the Structural Similarity Index (SSIM) to discriminate between real and synthetic EL images and reduce image distortion.

Other different applications and methods for training GANs have been reviewed by Creswell et al., [74], in which open challenges related to training and evaluating GANs are also presented.

On the other hand, Variational Autoencoders (VAEs) can also be used as an unsupervised learning method for data augmentation. A

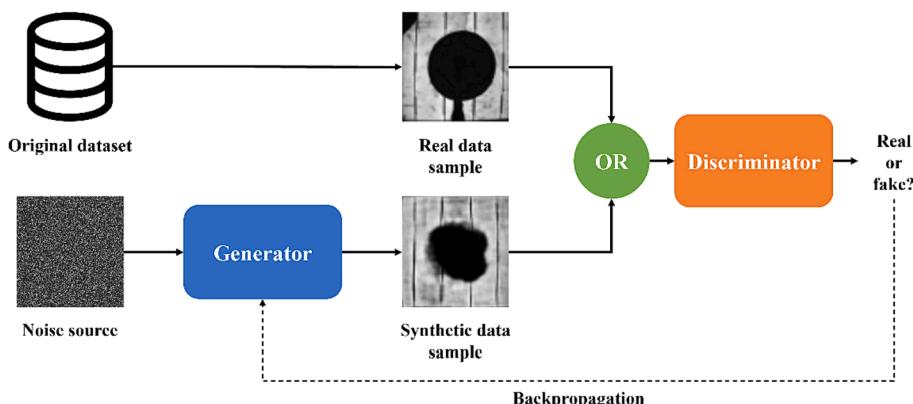


Fig. 11. The training process for a GAN [73,74].

traditional Autoencoder (AE) is a generative neural network composed of two parts: an encoder and a decoder. The encoder learns via deterministic mapping from encoding the input data, such as images, into a latent representation. The decoder maps the latent representation back to a space representation producing a new reconstruction that is not identical to the original. The two mappings are learnt such that the reconstructed image is as similar as possible to the input [74]. VAEs expand on this concept by following a probability distribution in the latent representation (e.g., a Gaussian distribution) rather than an arbitrary vector space [79].

For instance, VAEs were introduced by Westraadt et al., [79] to expand the training dataset and improve the classification performance for detecting faults in thermal solar cell images. In this study, an additional 900 thermal images were generated for each fault class. Three CNN models, namely InceptionV3, ResNet50, and Xception, trained on the VAE-augmented dataset showed higher accuracies in PV fault detection and classification tasks when compared to the performance using the original datasets.

Moreover, Conditional VAEs (CVAEs) were used for augmenting time-series data in a study proposed by Gong et al., [80], in which power curves are reconstructed for electricity theft detection. In this case, the CVAE structure needed to be redesigned to accommodate the processing of power curves, whereas existing structures were able to support the processing of 2D data [80]. The results showed that the accuracy values with CVAEs-augmented data outperformed other data augmentation methods, such as Conditional GANs (CGANs).

6.3. Future orientations

Potential future directions are identified to address the limitations of PV defect detection systems as illustrated in Fig. 12. As defect detection algorithms can be computationally demanding, especially with large datasets, model acceleration is considered a key area for enabling efficient real-time monitoring. This involves reducing the detection time and adopting lightweight algorithms to improve the efficiency. In addition to model acceleration, solutions based on edge computing, involving processing data at the device level, can be developed to reduce the latency and enable real-time monitoring. Moreover, the optimisation of selected hardware contributes to faster and more efficient defect detection.

Another key area is enhancing data pre-processing. One of the approaches considered for combating the 'imbalanced dataset' problem is data augmentation. Many samples can be obtained through classical

manipulation techniques; however, more complex DL-based techniques have been investigated in the literature for generating high-quality synthetic data with defects resembling the original data distribution. Future work is expected to target improving the quality of GAN samples and testing their effectiveness on a wide range of datasets. Moreover, the combinatorics of GAN samples with other augmentation techniques need to be explored. Other difficulties concerning GANs include achieving high-resolution output generation [72].

Apart from data augmentation, transfer learning can be developed for PV module defect detection. This technique leverages knowledge from one task to improve learning in another. It can be applied when available data is insufficient and extensive labelled datasets are needed to compare the results with existing approaches.

On the other hand, transformers have been originally developed for Natural Language Processing (NLP) and have shown capability in image generation tasks. However, the literature has scarcely covered the utilisation of transformers for data augmentation, and more exploration is needed to identify their effectiveness for PV defect diagnosis systems.

On the AL algorithm level, defect generalisation is considered crucial in practical, uncontrolled environments. This involves the ability to identify multiple class defects and detect emerging ones. In addition, enhancing the robustness and sensitivity of the algorithm is required to adapt to the variations in data under changing conditions. Another critical aspect of DL-based approaches is the selection of optimised parameters for extracting features and reconstructing inputs. This requires the ability to reduce tremendous work related to the manual tuning of parameters by using optimisation tools to achieve an improved diagnosis performance.

Other future work is expected to expand on defect localisation and size identification, integration with the industrial manufacturing processes of PV panels, and rising privacy and data security while still providing accurate and effective results.

7. Conclusions

In this review, a comprehensive review of the different data analysis methods of PV defect detection systems has been presented. The review included all approaches related to the main two categories of both imaging-based and electrical testing techniques with a greater categorisation granularity in terms of types and methods for each technique. Moreover, a critical analysis of each of the different data analysis methods including the advantages and disadvantages is performed, which can be referred to by future studies to identify the most suitable

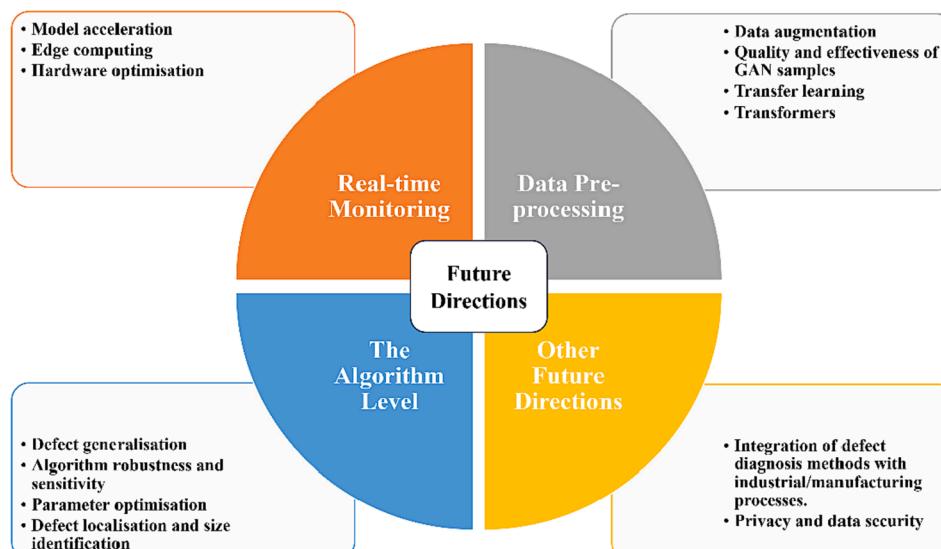


Fig. 12. Future directions for defect detection in PV systems.

method considering the use-case's requirements and setting.

The adoption of each of the reviewed techniques depends on several factors including the deployment scale, the targeted defects for detection, and the required location of defect analysis in the PV system. From a higher perspective, IBTs can be considered for analysing optical properties, thermal patterns, or other visual features, whereas ETTs detect anomalies by comparing the deviations of the module's measured operating states or electrical parameters from the expected electrical behaviour. While IBTs provide a high-resolution visual representation of the module surface allowing for the detection and diagnosis of small, structural defects that may be missed by other techniques, they have a limited ability to detect defects of the internal PV system or electrical faults. On the other hand, although ETTs can detect electrical faults beyond the PV module's surface, they are considered intrusive requiring the module to be interrupted and physically disconnected for inspection.

On the IBT level, it can be observed from Table 1 that the most notable adopted techniques in the literature are IRT- and EL-based, whereas the least are LBIC-based. IRT techniques are more practical for large-scale applications than EL imaging, as the latter would involve a large power supply making the setup more complex and costly. However, the IRT setup can also be expensive due to the costs of professional and reliable IR cameras for accurate measurements, which may require experienced operators for configuration. Moreover, the reliability of IRT techniques can be affected by external factors that may impact the measurement accuracy, such as emissivity variations and reflections. EL imaging is considered a non-intrusive technique that is highly efficient in localising defects and detecting defects of internal solar cells. They can be techno-economically effective for small-scale deployments. Recent studies [67,70] have shown that challenges in EL imaging, such as the requirement of a darkened space for image acquisition, can be overcome and EL images can be captured under daylight conditions by using InGaAs uncooled and InSb cooled detector cameras.

Challenges observed in the state-of-the-art literature have also been identified to be related to data availability, real-time monitoring, accurate measurements, computational efficiency, and dataset distribution. To address some of these challenges, classical and ML-based data augmentation approaches have been reviewed. In addition, potential future directions are identified addressing the limitations of PV defect detection systems. Future advancements are expected to be in the areas of real-time monitoring (e.g., model acceleration, edge computing, and hardware optimisation), data pre-processing (e.g., data augmentation, improving the quality of GAN samples, transfer learning, and use of transformers), the algorithm development (e.g., defect generalisation, improvement of robustness and sensitivity, and defect localisation), system integration with manufacturing processes, and privacy and data security.

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.

Acknowledgements

This research has been co-funded by The Manufacturing Technology Centre (MTC) and the University of Warwick.

References

- [1] Global Market Outlook For Solar Power 2023-2027 n.d. <https://www.solarpowereurope.org/insights/outlooks/global-market-outlook-for-solar-power-2023-2027> (accessed October 8, 2023).
- [2] Solar photovoltaics deployment. GOVUK 2023. <https://www.gov.uk/government/statistics/solar-photovoltaics-deployment> (accessed October 8, 2023).
- [3] H. Ullah, B. Marí, L.M.S. Ruiz, Effect of defects on the performance of some photovoltaic solar cells: an introduction to research methods to engineering students. 44th SEFI Conference, 2016.
- [4] D.-M. Tsai, S.-C. Wu, W.-Y. Chiu, Defect detection in solar modules using ICA basis images, IEEE Trans. Ind. Inf. 9 (2013) 122–131, <https://doi.org/10.1109/TII.2012.2209663>.
- [5] A. Bouich, Study and characterization of hybrid perovskites and copper-indium-gallium selenide thin films for tandem solar cells, Tesis Doctoral. Universitat Politècnica De València (2021), <https://doi.org/10.4995/Thesis/10251/160621>.
- [6] A. Bouich, J. Marí-Guaita, F. Baig, Y. Hameed Khattak, B.M. Soucase, P. Palacios, Investigation of the surface coating, humidity degradation, and recovery of perovskite film phase for solar-cell applications, Nanomaterials 12 (2022) 3027, <https://doi.org/10.3390/nano12173027>.
- [7] M.M. Byranvand, M. Saliba, Charge carrier management for developing high-efficiency perovskite solar cells, Matter 4 (2021) 1758–1759, <https://doi.org/10.1016/j.matt.2021.04.020>.
- [8] A. Bouich, J.C. Torres, H. Chfii, J. Marí-Guaita, Y.H. Khattak, F. Baig, et al., Delafossite as hole transport layer a new pathway for efficient perovskite-based solar cells: insight from experimental, DFT and Numerical Analysis. Solar Energy 250 (2023) 18–32, <https://doi.org/10.1016/j.solener.2022.12.022>.
- [9] N. Rana, S. Arora, A Review on Surface Defect Detection of Solar Cells Using Machine Learning, in: N. Marriwala, C.C. Tripathi, S. Jain, S. Mathapathi (Eds.), Soft Computing for Intelligent Systems, Springer, Singapore, 2021, pp. 385–395, https://doi.org/10.1007/978-981-16-1048-6_29.
- [10] R. Al-Mashhadani, G. Alkawsi, Y. Baashar, A.A. Alkahtani, F.H. Nordin, W. Hashim, Deep learning methods for solar fault detection and classification: a review, Information Sciences Letters 10 (2021).
- [11] K.V.S. Bharath, A. Haque, M.A. Khan, Fault classification for photovoltaic modules using thermography and image processing, 2019 IEEE Industry Applications Society Annual Meeting (2019) 1–6, <https://doi.org/10.1109/IAS.2019.8912356>.
- [12] M. Mansouri, M. Trabelsi, H. Nounou, M. Nounou, Deep learning-based fault diagnosis of photovoltaic systems: a comprehensive review and enhancement prospects, IEEE Access 9 (2021) 126286–126306, <https://doi.org/10.1109/ACCESS.2021.3110947>.
- [13] A. Triki-Lahiani, A. Bennani-Ben Abdelghani, I. Slama-Belkhodja, Fault detection and monitoring systems for photovoltaic installations: a review, Renew. Sustain. Energy Rev. 82 (2018) 2680–2692, <https://doi.org/10.1016/j.rser.2017.09.101>.
- [14] A. Mellit, G.M. Tina, S.A. Kalogirou, Fault detection and diagnosis methods for photovoltaic systems: a review, Renew. Sustain. Energy Rev. 91 (2018) 1–17, <https://doi.org/10.1016/j.rser.2018.03.062>.
- [15] A.K.V. de Oliveira, M. Aghaei, R. Rüther, Automatic inspection of photovoltaic power plants using aerial infrared thermography: a review, Energies 15 (2022) 2055, <https://doi.org/10.3390/en15062055>.
- [16] Á.H. Herráiz, A.P. Marugán, M.F. Chapter, 7 - A review on condition monitoring system for solar plants based on thermography, in: M. Papaelias, F.P.G. Márquez, A. Karyotakis (Eds.), Non-Destructive Testing and Condition Monitoring Techniques for Renewable Energy Industrial Assets, Butterworth-Heinemann, Boston, 2020, pp. 103–118, <https://doi.org/10.1016/B978-0-08-101094-5.00007-1>.
- [17] I. Tsanakas (John), G. Vannier, A. Plissonnier, Fault diagnosis and classification of large-scale photovoltaic plants through aerial orthophoto Thermal Mapping, (2015).
- [18] J.A. Tsanakas, L. Ha, C. Buerhop, Faults and infrared thermographic diagnosis in operating c-Si photovoltaic modules: a review of research and future challenges, Renew. Sustain. Energy Rev. 62 (2016) 695–709, <https://doi.org/10.1016/j.rser.2016.04.079>.
- [19] M.W. Akram, G. Li, Y. Jin, X. Chen, C. Zhu, A. Ahmad, Automatic detection of photovoltaic module defects in infrared images with isolated and develop-model transfer deep learning, Sol. Energy 198 (2020) 175–186, <https://doi.org/10.1016/j.solener.2020.01.055>.
- [20] Á. Huerta Herráiz, A. Pliego Marugán, F.P. García Márquez, Photovoltaic plant condition monitoring using thermal images analysis by convolutional neural network-based structure, Renew. Energy 153 (2020) 334–348, <https://doi.org/10.1016/j.renene.2020.01.148>.
- [21] F. Lopez, C. Ibarra-Castanedo, N.V. de Paulo, X. Maldague, Optimization of pulsed thermography inspection by partial least-squares regression, NDT and E Int. 66 (2014) 128–138, <https://doi.org/10.1016/j.ndteint.2014.06.003>.
- [22] A. Sinha, O.S. Sastry, R. Gupta, Detection and characterisation of delamination in PV modules by active infrared thermography, Nondestructive Testing and Evaluation 31 (2016) 1–16, <https://doi.org/10.1080/10589759.2015.1034717>.
- [23] M. Alajmi, K. Awedat, M.S. Aldeen, S. Alwagdani, IR thermal image analysis: an efficient algorithm for accurate hot-spot fault detection and localization in solar photovoltaic systems, IEEE Int. Conf. Electro Inform. Tech. (EIT) 2019 (2019) 162–168, <https://doi.org/10.1109/EIT.2019.8833855>.
- [24] L. Jiang, J. Su, X. Li, Hot spots detection of operating PV arrays through IR thermal image using method based on curve fitting of gray histogram, MATEC Web of Conferences 61 (2016) 06017, <https://doi.org/10.1051/matecconf/20166106017>.
- [25] M.U. Ali, H.F. Khan, M. Masud, K.D. Kallu, A. Zafar, A machine learning framework to identify the hotspot in photovoltaic module using infrared thermography, Sol. Energy 208 (2020) 643–651, <https://doi.org/10.1016/j.solener.2020.08.027>.
- [26] P. Haidari, A. Hajiahmad, A. Jafari, A. Nasiri, Deep learning-based model for fault classification in solar modules using infrared images, Sustain. Energy Technol. Assess. 52 (2022), 102110, <https://doi.org/10.1016/j.seta.2022.102110>.
- [27] K. Gopalakrishnan, S.K. Khaitan, A. Choudhary, A. Agrawal, Deep Convolutional Neural Networks with transfer learning for computer vision-based data-driven

- pavement distress detection, *Constr. Build. Mater.* 157 (2017) 322–330, <https://doi.org/10.1016/j.conbuildmat.2017.09.110>.
- [28] J.A. Tsanakas, L.D. Ha, S.F. Al, Advanced inspection of photovoltaic installations by aerial triangulation and terrestrial georeferencing of thermal/visual imagery, *Renew. Energy* 102 (2017) 224–233, <https://doi.org/10.1016/j.renene.2016.10.046>.
- [29] B.-Rújula AA. Chapter, 8 - Solar photovoltaics (PV), in: F. Calise, M.D. D'Accadìa, M. Santarelli, A. Lanzini, D. Ferrero (Eds.), *Solar Hydrogen Production*, Academic Press, 2019, pp. 237–295, <https://doi.org/10.1016/B978-0-12-814853-2.00008-4>.
- [30] D.-M. Tsai, S.-C. Wu, W.-C. Li, Defect detection of solar cells in electroluminescence images using Fourier image reconstruction, *Sol. Energy Mater. Sol. Cells* 99 (2012) 250–262, <https://doi.org/10.1016/j.solmat.2011.12.007>.
- [31] A.S. Rajput, J.W. Ho, Y. Zhang, S. Nalluri, A.G. Aberle, Quantitative estimation of electrical performance parameters of individual solar cells in silicon photovoltaic modules using electroluminescence imaging, *Sol. Energy* 173 (2018) 201–208, <https://doi.org/10.1016/j.solener.2018.07.046>.
- [32] B. Su, Z. Zhou, H. Chen, PVEL-AD: a large-scale open-world dataset for photovoltaic cell anomaly detection, *IEEE Trans Ind Inf* 19 (2023) 404–413, <https://doi.org/10.1109/TII.2022.3162846>.
- [33] A. Bartler, L. Mauch, B. Yang, M. Reuter, L. Stoicescu, Automated detection of solar cell defects with deep learning, 2018 26th european signal processing conference (EUSIPCO), 2018, p. 2035–9, 10.23919/EUSIPCO.2018.8553025.
- [34] J.S. Fada, M.A. Hossain, J.L. Braid, S. Yang, T.J. Peshek, R.H. French, Electroluminescence image processing and cell degradation type classification via computer vision and statistical learning methodologies, 2017 IEEE 44th Photovoltaic Specialist Conference (PVSC), 2017, p. 3456–61, 10.1109/PVSC.2017.8366291.
- [35] A.M. Karimi, J.S. Fada, M.A. Hossain, S. Yang, T.J. Peshek, J.L. Braid, et al., Automated pipeline for photovoltaic module electroluminescence image processing and degradation feature classification, *IEEE J. Photovoltaics* 9 (2019) 1324–1335, <https://doi.org/10.1109/JPHOTOV.2019.2920732>.
- [36] X. Qian, H. Zhang, H. Zhang, Y. Wu, Z. Diau, Q.-E. Wu, et al., Solar cell surface defects detection based on computer vision, *Int. J. Performability Eng.* 13 (1048) (2017), <https://doi.org/10.23940/ijpe.17.07.p6.10481056>.
- [37] S.A. Anwar, M.Z. Abdullah, Micro-crack detection of multicrystalline solar cells featuring an improved anisotropic diffusion filter and image segmentation technique, *J Image Video Proc* 2014 (2014) 15, <https://doi.org/10.1186/1687-5281-2014-15>.
- [38] M. Dhimish, V. Holmes, Solar cells micro crack detection technique using state-of-the-art electroluminescence imaging, *J. Sci.: Adv. Mater. Devices* 4 (2019) 499–508, <https://doi.org/10.1016/j.jamtd.2019.10.004>.
- [39] H. Chen, H. Zhao, D. Han, K. Liu, Accurate and robust crack detection using steerable evidence filtering in electroluminescence images of solar cells, *Opt. Lasers Eng.* 118 (2019) 22–33, <https://doi.org/10.1016/j.optlaseng.2019.01.016>.
- [40] S. Deitsch, V. Christlein, S. Berger, C. Buerhop-Lutz, A. Maier, F. Gallwitz, et al., Automatic classification of defective photovoltaic module cells in electroluminescence images, *Sol. Energy* 185 (2019) 455–468, <https://doi.org/10.1016/j.solener.2019.02.067>.
- [41] Simonyan K, Zisserman A. Very Deep Convolutional Networks for Large-Scale Image Recognition 2015.
- [42] W. Tang, Q. Yang, K. Xiong, W. Yan, Deep learning based automatic defect identification of photovoltaic module using electroluminescence images, *Sol. Energy* 201 (2020) 453–460, <https://doi.org/10.1016/j.solener.2020.03.049>.
- [43] M.W. Akram, G. Li, Y. Jin, X. Chen, C. Zhu, X. Zhao, et al., CNN based automatic detection of photovoltaic cell defects in electroluminescence images, *Energy* 189 (2019), 116319, <https://doi.org/10.1016/j.energy.2019.116319>.
- [44] X. Chen, T. Karin, A. Jain, Automated defect identification in electroluminescence images of solar modules, *Sol. Energy* 242 (2022) 20–29, <https://doi.org/10.1016/j.solener.2022.06.031>.
- [45] J. Fiorese, D.J. Colvin, R. Frota, R. Gupta, M. Li, H.P. Seigneur, et al., Automated defect detection and localization in photovoltaic cells using semantic segmentation of electroluminescence images, *IEEE J. Photovoltaics* 12 (2022) 53–61, <https://doi.org/10.1109/JPHOTOV2021.3131059>.
- [46] B. Su, H. Chen, P. Chen, G. Bian, K. Liu, W. Liu, Deep learning-based solar-cell manufacturing defect detection with complementary attention network, *IEEE Trans. Ind. Inf.* 17 (2021) 4084–4095, <https://doi.org/10.1109/TII.2020.3008021>.
- [47] D. Gan, L. Quan, F. Zhu, K. Xie, J. Bai, Orthogonal modulation based light beam induced current method for anti-noise defect detection in photovoltaic cells, *Sol. Energy* 245 (2022) 158–164, <https://doi.org/10.1016/j.solener.2022.08.058>.
- [48] L. Quan, K. Xie, Y. Liu, H. Zhang, Camera enhanced compressive light beam induced current sensing for efficient defect detection in photovoltaic cells, *Sol. Energy* 183 (2019) 212–217, <https://doi.org/10.1016/j.solener.2019.02.055>.
- [49] X. Lu, P. Lin, S. Cheng, Y. Lin, Z. Chen, L. Wu, et al., Fault diagnosis for photovoltaic array based on convolutional neural network and electrical time series graph, *Energ. Convers. Manage.* 196 (2019) 950–965, <https://doi.org/10.1016/j.enconman.2019.06.062>.
- [50] A. Bakdi, W. Bounoua, A. Guichi, S. Mekhilef, Real-time fault detection in PV systems under MPPT using PMU and high-frequency multi-sensor data through online PCA-KDE-based multivariate KL divergence, *Int. J. Electr. Power Energy Syst.* 125 (2021), 106457, <https://doi.org/10.1016/j.ijepes.2020.106457>.
- [51] D.S. Pillai, N. Rajasekar, A comprehensive review on protection challenges and fault diagnosis in PV systems, *Renew. Sustain. Energy Rev.* 91 (2018) 18–40, <https://doi.org/10.1016/j.rser.2018.03.082>.
- [52] 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, <https://doi.org/10.1016/j.solener.2021.07.037>.
- [53] W. Wang, A.-C.-F. Liu, H.-S.-H. Chung, R.-W.-H. Lau, J. Zhang, A.-W.-L. Lo, Fault diagnosis of photovoltaic panels using dynamic current-voltage characteristics, *IEEE Trans. Power Electron.* 31 (2016) 1588–1599, <https://doi.org/10.1109/TPEL.2015.2424079>.
- [54] J. Solórzano, M.A. Egido, Automatic fault diagnosis in PV systems with distributed MPPT, *Energ. Convers. Manage.* 76 (2013) 925–934, <https://doi.org/10.1016/j.enconman.2013.08.055>.
- [55] S. Sarikh, M. Raoufi, A. Bennouna, A. Benlarabi, B. Ikken, Fault diagnosis in a photovoltaic system through I-V characteristics analysis, 2018 9th International Renewable Energy Congress (IREC), 2018, p. 1–6, 10.1109/IREC.2018.8362572.
- [56] T. Takashima, J. Yamaguchi, M. Ishida, Disconnection detection using earth capacitance measurement in photovoltaic module string, *Prog. Photovolt. Res. Appl.* 16 (2008) 669–677, <https://doi.org/10.1002/pip.860>.
- [57] T. Takashima, J. Yamaguchi, K. Otani, T. Oozeki, K. Kato, M. Ishida, Experimental studies of fault location in PV module strings, *Sol. Energy Mater. Sol. Cells* 93 (2009) 1079–1082, <https://doi.org/10.1016/j.solmat.2008.11.060>.
- [58] Z. Chen, Y. Chen, L. Wu, S. Cheng, P. Lin, Deep residual network based fault detection and diagnosis of photovoltaic arrays using current-voltage curves and ambient conditions, *Energ. Convers. Manage.* 198 (2019), 111793, <https://doi.org/10.1016/j.enconman.2019.111793>.
- [59] S. Roy, M.K. Alam, F. Khan, J. Johnson, J. Flicker, An irradiance-independent, robust ground-fault detection scheme for PV arrays based on spread spectrum time-domain reflectometry (SSTDR), *IEEE Trans. Power Electron.* 33 (2018) 7046–7057, <https://doi.org/10.1109/TPEL.2017.2755592>.
- [60] Y. Stauffer, D. Ferrario, E. Onillon, A. Hutter, Power monitoring based photovoltaic installation fault detection, *Int. Conf. Renew. Energ. Res. Appl. (ICRERA)* 2015 (2015) 199–202, <https://doi.org/10.1109/ICRERA.2015.7418695>.
- [61] M. Dhimish, V. Holmes, Fault detection algorithm for grid-connected photovoltaic plants, *Sol. Energy* 137 (2016) 236–245, <https://doi.org/10.1016/j.solener.2016.08.021>.
- [62] Y.H. Chen, R. Liang, Y. Tian, F. Wang, A novel fault diagnosis method of PV based-on power loss and I-V characteristics, *IOP Conf Ser: Earth Environ Sci* 40 (2016), 012022, <https://doi.org/10.1088/1755-1315/40/1/012022>.
- [63] Z. Yi, A.H. Etemadi, Fault detection for photovoltaic systems based on multi-resolution signal decomposition and fuzzy inference systems, *IEEE Trans. Smart Grid* 8 (2017) 1274–1283, <https://doi.org/10.1109/TSG.2016.2587244>.
- [64] H. Mohamed, A.M. Nassar, New algorithm for fault diagnosis of photovoltaic energy systems, *IJCA* 114 (2015) 26–31, <https://doi.org/10.5120/20008-1959>.
- [65] H. Mekki, A. Mellit, H. Salhi, Artificial neural network-based modelling and fault detection of partial shaded photovoltaic modules, *Simul. Model. Pract. Theory* 67 (2016) 1–13, <https://doi.org/10.1016/j.simpat.2016.05.005>.
- [66] A.W. Kandeel, M.R. Elkadeem, A. Kumar Thakur, G.B. Abdelaziz, R. Sathyamurthy, A.E. Kabeel, et al., Infrared thermography-based condition monitoring of solar photovoltaic systems: a mini review of recent advances, *Sol. Energy* 223 (2021) 33–43, <https://doi.org/10.1016/j.solener.2021.05.032>.
- [67] A. Dos Reis, G. Benatto, C. Mantel, S. Spataru, A.A. Santamaría Lancia, N. Riedel, S. Thorsteinsson, et al., Drone-based daylight electroluminescence imaging of PV modules, *IEEE J Photovoltaics* 10 (2020) 872–877, <https://doi.org/10.1109/JPHOTOV.2020.2978068>.
- [68] S. Johnston, T. Silverman. Photoluminescence and Electroluminescence Outdoor Module Imaging n.d.
- [69] G.A. Dos Reis Benatto, J. Vedde, M. Petersen, H. Voss, M. Messerschmidt, H. Parikh, et al., Development of outdoor luminescence imaging for drone-based PV array inspection, 2017 IEEE 44th Photovoltaic Specialist Conference (PVSC), Washington, DC: IEEE; 2017, p. 2682–7, 10.1109/PVSC.2017.8366002.
- [70] M. Guada, Á. Moretón, S. Rodríguez-Conde, L.A. Sánchez, M. Martínez, M.A. González, et al., Daylight luminescence system for silicon solar panels based on a bias switching method, *Energy Sci. Eng.* 8 (2020) 3839–3853, <https://doi.org/10.1002/ese3.781>.
- [71] L. Quan, K. Xie, R. Xi, Y. Liu, Compressive light beam induced current sensing for fast defect detection in photovoltaic cells, *Sol. Energy* 150 (2017) 345–352, <https://doi.org/10.1016/j.solener.2017.04.059>.
- [72] C. Shorten, T.M. Khoshgoftaar, A survey on image data augmentation for deep learning, *J Big Data* 6 (2019) 60, <https://doi.org/10.1186/s40537-019-0197-0>.
- [73] H.F.M. Romero, L. Hernández-Callejo, M.Á.G. Rebollo, V. Cardeñoso-Payo, V. A. Gómez, H.J. Bello, et al., Synthetic dataset of electroluminescence images of photovoltaic cells by deep convolutional generative adversarial networks, *Sustainability* 15 (2023) 7175, <https://doi.org/10.3390/su15097175>.
- [74] A. Creswell, T. White, V. Dumoulin, K. Arulkumaran, B. Sengupta, A.A. Bharath, Generative adversarial networks: an overview, *IEEE Signal Process Mag* 35 (2018) 53–65, <https://doi.org/10.1109/MSP.2017.2765202>.
- [75] Y. Chen, Y. Ding, F. Zhao, E. Zhang, Z. Wu, L. Shao, Surface defect detection methods for industrial products: a review, *Appl. Sci.* 11 (2021) 7657, <https://doi.org/10.3390/app11167657>.
- [76] Z. Luo, S.Y. Cheng, Q.Y. Zheng, GAN-based augmentation for improving CNN performance of classification of defective photovoltaic module cells in electroluminescence images, *IOP Conf Ser: Earth Environ Sci* 354 (2019), 012106, <https://doi.org/10.1088/1755-1315/354/1/012106>.
- [77] T. Karras, T. Aila, S. Laine, J. Lehtinen, Progressive growing of GANs for improved quality, stability, and variation 2018, 10.48550/arXiv.1710.10196.
- [78] C. Shou, L. Hong, W. Ding, Q. Shen, W. Zhou, Y. Jiang, et al. Defect detection with generative adversarial networks for electroluminescence images of solar cells. 2020

- 35th Youth Academic Annual Conference of Chinese Association of Automation (YAC), 2020, p. 312–7. 10.1109/YAC51587.2020.9337676.
- [79] E.J. Westraadt, W.J. Brettenny, C.M. Clohessy, Deep learning for photovoltaic defect detection using variational autoencoders, *S. Afr. J. Sci.* 119 (1–8) (2023), <https://doi.org/10.17159/sajs.2023/13117>.
- [80] X. Gong, B. Tang, R. Zhu, W. Liao, L. Song, Data augmentation for electricity theft detection using conditional variational auto-encoder, *Energies* 13 (2020) 4291, <https://doi.org/10.3390/en13174291>.