Energy Conversion and Management: X 27 (2025) 101079

Contents lists available at ScienceDirect

Energy Conversion and Management: X

journal homepage: www.sciencedirect.com/journal/energy-conversion-and-management-x

Innovative dust detection and efficient cleaning of PV Panels: A CNN-RF approach using I–V curve data transformed into RGB mosaics

Safia Babikir Bashir [a], Mena Maurice Farag [a,b], Abdul-Kadir Hamid [a,b], Ali A. Adam [a], Ramesh C. Bansal [a,c], Nsilulu T Mbungu [a], A. Elnady [a], Ahmed G. Abo-Khalil [a], Mousa Hussein [d,*]

[a] Smart Grid Research Group, Sustainable Energy and Power Systems Research Center, Research Institute of Sciences and Engineering (RISE), Department of Electrical Engineering, University of Sharjah, Sharjah 27272, United Arab Emirates
[b] Smart Automation and Communication Technologies Research Center (SACT), University of Sharjah 27272 Sharjah, United Arab Emirates
[c] Department of Electrical, Electronic and Computer Engineering, University of Pretoria, Pretoria, South Africa
[d] Department of Electrical and Communication Engineering, United Arab Emirates University, Al Ain, United Arab Emirates

ARTICLE INFO

Keywords:
Automated Cleaning
Convolutional Neural Networks
Dust Detection
Machine Learning
Photovoltaics

ABSTRACT

Photovoltaic (PV) panels are vital for renewable energy generation, yet their efficiency is critically hindered by environmental challenges such as dust accumulation, especially in arid regions like the UAE. Dust buildup can reduce efficiency by up to 30% within a month, threatening the sustainability of solar power, which is projected to supply 10% of global energy by 2030. Existing cleaning methods are unsustainable, consuming an estimated 10 billion gallons of water annually, enough to meet the drinking needs of 2 million people, necessitating the development of a cost-effective, resource-efficient alternative. This research presents a novel machine learning-based system to automate dust detection and optimize cleaning, significantly reducing water consumption while improving power generation efficiency. The methodology transforms I-V curve electrical parameters into RGB mosaic images, enabling precise classification of operational states such as normal operation, dust accumulation, shading, and faults without relying on external imaging devices. The system is built on a hybrid model combining Convolutional Neural Networks (CNN) and Random Forest (RF) classifiers (CNN-RF), where the CNN extracts high-level features from RGB mosaic images, and the RF classifier accurately categorizes operational states. Upon detecting dust accumulation, a secondary CNN-RF model classifies the severity into low, moderate, or heavy, guiding an optimized cleaning process that minimizes water usage while maintaining cleaning effectiveness. The primary CNN-RF model achieved 100% accuracy in classifying operational states using RGB mosaic images, surpassing the 97% accuracy achieved by I–V curve-based methods. Furthermore, the secondary CNN-RF model for dust severity classification attained an accuracy of 98% using RGB mosaic images, compared to only 68% when using traditional I–V curves, highlighting the superior performance of RGB mosaic images in detecting fine-grained dust levels. This optimized classification approach guides an automated cleaning system that minimizes water usage while maintaining PV panel efficiency.

1. Introduction

The rapid consumption of fossil fuels has escalated environmental concerns, which has motivated the global initiative for renewable energy sources [1–3] Renewable energy sources have been introduced in various forms such as solar, wind, geothermal, hydro, biomass, and hybrid renewable energy systems [4,5]. Solar energy has been predominantly utilized amongst other clean energy resources due to its scalability in regions with abundant solar irradiance [6,7], through the development of photovoltaic (PV) power plants [8]. Reports indicate that solar energy could meet the demand for global power generation by 10 % by 2030 [9], with major power plants installed in isolated and

* Corresponding author.
E-mail addresses: safia.mohamed@sharjah.ac.ae (S.B. Bashir), u20105427@sharjah.ac.ae (M.M. Farag), akhamid@sharjah.ac.ae (A.-K. Hamid), aismail@sharjah.ac.ae (A.A. Adam), rbansal@sharjah.ac.ae (R.C. Bansal), anady@sharjah.ac.ae (A. Elnady), aabokhalil@sharjah.ac.ae (A.G. Abo-Khalil), mihussein@uaeu.ac.ae (M. Hussein).

https://doi.org/10.1016/j.ecmx.2025.101079
Received 20 March 2025; Received in revised form 9 May 2025; Accepted 21 May 2025
Available online 25 May 2025

deserted regions with abundant exposure to sunlight [10,11]. However, PV module's performance depends on several environmental circumstances, such as shading [12], dust accumulation [13], overheating [14], etc., leading to degradation in their performance and stability. Amongst these circumstances, dust accumulation poses a substantial challenge in maintaining consistent PV power generation [15,16].

Dust accumulation on the surface of PV modules forms a fine layer that obstructs solar irradiance from reaching the PV cells, leading to a notable reduction in DC current generation, DC power generation, and energy conversion efficiency [17]. Studies reported that periodic dust accumulation can severely reduce the output of PV modules by up to 30 % during a single month [18]. Therefore, the dust accumulation not only undermines the energy yield but also exposes a risk of having hotspots and overheating, which potentially impact the longevity and lifetime of PV systems [19].

Several studies have examined the effect of dust in harsh and arid environmental conditions. The properties of dust in Iran's desert region and their impact on PV module energy efficiency were investigated by Ali Sadat et al. [20]. Furthermore, several models have been developed to examine the process of dust collection and the resulting reduction in PV module efficiency. To identify the best strategy depending on the type of dust, Kazem et al. [21] analyzed the effects of various dust compositions on the performance of monocrystalline and polycrystalline panels in addition to evaluating several hand-cleaning techniques. Research results show that the efficiency loss associated with dust that accumulates progressively grows exponentially with its bulk [22]. Al-Ghussain [23] has evaluated the impact of operational parameters on the yearly energy output of photovoltaic power plants in arid areas. The ideal yearly tilt and azimuth angles for maximum power output, as well as the panel's cleaning frequency, ambient temperature, wind speed, and dust collection rate, were all measured during this study. These studies demonstrated the impact of dust accumulation on the overall yield generated by PV systems.

In principle, the typical response to dust accumulation involves periodic cleaning, which is cost-intensive and environmentally challenging [24]. Common methods of cleaning such as water cleaning, have been reported to consume 10 billion gallons of water annually, which is approximately sufficient to supply drinking water for 2 million people [25]. Alternative methods for cleaning have been further studied, however, they require a large workforce and may cause damage to the PV module's surface such as dry cleaning methods, which negatively affect the PV's efficiency [26]. Therefore, early dust detection must be addressed by proposing efficient methods for managing and detecting dust in order to maximize resource consumption and preserve PV systems' ideal performance [27–29].

Regular visual inspections have historically been the most efficient way to detect dirt and dust on photovoltaic (PV) modules; however, new technologies such as intelligent image-processing algorithms and sensor-based remote monitoring are being developed more and more to increase the efficiency and accuracy of detection [30].

Photodiodes, phototransistors, or optical sensors are used in sensor-based dust detection approaches to track the amount of light that reaches the PV panels. Dust buildup on the panels prevents light from getting to the sensors, which causes the output of the sensors to noticeably decline. Comparably, Malik et al. [31] presented a technique to automatically identify the organization of dust on photovoltaic panels by integrating a brightness sensor with a decision algorithm that takes into account various daytime and weather circumstances. The ultimate product was an Arduino-based embedded system prototype that could notify the maintenance operator in real time and transfer data to the cloud. A sensor-based monitoring system using an Arduino microcontroller was created by Mohammed et al. [32] in a related work to track the power generation of PV modules and to start cleaning when needed. Hussain et al. [33] used sensors to monitor important electrical performance indicators like voltage, current, and power to investigate the effect of ambient dust on the energy loss of PV modules. According to their research, dust deposition in desert areas might cause an electrical efficiency loss of up to 60 %. Furthermore, Pérez-Anaya et al. [34] presented a machine learning-based methodology that uses a variety of sensors, including voltage, current, temperature, and solar radiation, to evaluate different degrees of dust collection on PV panels. The accuracy of the suggested method exceeded 94 %, demonstrating accurate dust collection during the entire period.

Moreover, cameras or other imaging equipment are used in image-based detection systems to take pictures of PV panels, which are then analyzed using algorithms to identify and categorize dust particles. These techniques distinguish dust from other flaws or artifacts by examining characteristics such as particle size, shape, color, and texture. For example, to detect dust and grime on PV panels, Abuqaaud and Ferrah [35] used a computer vision-based approach with the Gray Level Cooccurrence Matrix (GLCM) technology, attaining high identification rates and simple implementation. In a similar vein, Sriram and Sudhakar [36] used thermal imaging to build a solar panel soiling detection system that eliminates the need for in-person inspections of big solar systems. Additionally, other research used image processing in conjunction with current and voltage sensors to evaluate how dirt and defects affect the energy efficiency of the panels [37].

Dust particles on photovoltaic panels may be reliably detected and classified using deep learning techniques. Rico Espinosa et al. [38], for example, presented an automated fault classification technique for PV plants that uses CNN for semantic classification and PV picture segmentation. The technique achieves up to 75 % prediction accuracy for two-class outputs. Unluturk et al. [39] proposed two methods, Gray Level Cooccurrence Matrix (GLCM) and CNN, developed to measure dust levels on solar panels. To test the efficacy of the strategy, Fan et al. [40] conducted a study in which they recommended the use of a deep residual neural network (DRNN) to recognize uneven dust states, such as concentration and distribution on PV panels. Similar to this, Tan et al. [41] developed a denoising CNN to evaluate the state of dust collection on photovoltaic panels; however, more work needs to be done to increase accuracy and efficiency. As a component of an integrated dust recognition system, Cao et al. [42] developed an end-to-end convolutional neural network (CNN) that showed efficiency and effectiveness. Furthermore, a study that coupled CNN with Support Vector Machine (SVM) was able to diagnose flaws in electroluminescence images of PV panels [43]. Seven distinct CNN-based power generation efficiency prediction models were examined by Zhang et al. [44], where the V2 model outperformed the others. Lastly, Onim et al. [45] achieved a prediction accuracy of 98.2 % when they constructed a specific CNN architecture called SolNet for detecting dust formation in solar panels.

Using drones or robots, autonomous solutions for PV panel cleaning and monitoring have become the subject of more and more research in recent years. To efficiently navigate and clean panels without causing any damage, these systems frequently combine sensors with image-based detection techniques [46,47]. Computer vision and machine learning techniques are also being used by several research institutes and businesses to advance solar panel cleaning technologies [48].

Despite these significant advancements, most current methods fall short in offering a comprehensive solution. Many focus solely on detecting dust or categorizing the operational state of the PV panel, but few provide a full system that can both classify various operational states, such as normal operation, shading, faults, and dust accumulation, and classify the level of dust accumulation. Additionally, most of these methods rely on actual images from cameras or sensor-based systems, which can be influenced by environmental conditions and may require complex setups. Furthermore, current approaches do not integrate optimization strategies to manage the water usage required for cleaning. A summary of the discussed literature is demonstrated in Table 1.

To address these challenges, this research introduces an integrated machine learning-based approach that not only detects and classifies various operational states of PV panels but also distinguishes between different dust accumulation levels (low, moderate, heavy). Unlike

Table 1
Summary of reported literature review on PV dust detection techniques.

Category	Reference	Focus	Methodology	Findings	Research Gap
Sensors	[31]	Dust detection on PV panels.	Brightness sensors integrated with decision algorithm, Arduino-based prototype.	Real-time dust detection with cloud data transfer; notifies maintenance staff.	Limited scalability; no integration with cleaning automation.
	[32]	Automated cleaning based on dust detection.	Sensor-based monitoring using an Arduino microcontroller.	Cleaning initiation is automated based on sensor data.	No classification of dust levels or resource optimization.
	[33]	Effect of dust on PV performance in desert areas.	Sensors track voltage, current, and power.	Dust deposition caused up to 60 % electrical efficiency loss.	Lacks predictive analytics and advanced detection systems.
	[34]	Evaluating degrees of dust collection.	Multi-sensor machine learning (voltage, current, temperature, solar radiation).	Achieved > 94 % accuracy in evaluating dust collection.	No integration with cleaning strategies or resource optimization.
Images	[35]	Dust detection using computer vision.	GLCM technology for analyzing particle characteristics.	High identification rates; simple implementation.	Focuses solely on detection; no operational state classification or cleaning strategies.
	[36]	Soiling detection in large PV systems.	Thermal imaging-based detection.	Eliminated the need for in-person inspections for large systems.	Requires specialized imaging equipment; lacks adaptability to environmental conditions.
	[38]	Fault classification for PV plants.	CNN-based semantic classification and image segmentation.	75 % accuracy for two-class classification.	Does not explore multi-class operational state classification or cleaning automation.
	[41]	Dust collection evaluation on PV panels.	Denoising CNN for image processing.	Improved dust detection accuracy but requires further enhancement in efficiency.	Focuses on detection only; lacks scalability and integration with cleaning optimization.
	[42]	End-to-end dust recognition system for PV panels.	Convolutional Neural Network (CNN).	Efficient dust recognition using end-to-end CNN.	Limited integration with optimization strategies or cleaning systems.
	[45]	Dust formation detection on PV panels.	SolNet (CNN-based architecture).	Achieved 98.2 % accuracy in dust detection.	Does not address cleaning or resource management for maintenance.
Machine Learning	[34]	Dust collection using sensors and ML.	Machine learning to evaluate sensor data.	Achieved > 94 % accuracy in dust evaluation.	No direct integration with cleaning strategies or water usage optimization.
	[40]	Uneven dust distribution and states on PV panels.	Deep residual neural network (DRNN).	Recognized dust concentration and distribution patterns.	Limited exploration of real-time application and cleaning integration.
	[44]	Efficiency prediction of PV panels.	CNN-based power generation efficiency prediction models.	Identified the V2 model as the best-performing for efficiency prediction.	Limited exploration of dust-specific classification or cleaning system optimization.
Drones/ Robots	[46,47]	Autonomous cleaning and monitoring.	Drones and robots are integrating sensors and image-based techniques.	Efficient navigation and cleaning without damage.	High implementation cost; no integration with water-efficient strategies.
	[48]	Advancing solar panel cleaning technologies.	Combining computer vision and machine learning for improved cleaning systems.	Promising results in cleaning technology enhancements.	Limited scalability and reliance on specific environmental conditions.

traditional methods that rely on camera or sensor images, the proposed method uses IV curve parameters (voltage, current, and power) converted into RGB mosaic images, which eliminates the need for external imaging devices and reduces dependency on environmental factors. In recent years, hardware solutions integrated at the inverter level have become commercially available, allowing I–V curves to be measured periodically at the power plant level, making this approach highly feasible for large-scale implementation.

The proposed method leverages a hybrid model combining CNN for feature extraction with Random Forest (RF) for classification. This system not only classifies the operational state of the PV panels but also categorizes the severity of dust accumulation. Furthermore, an optimization model is incorporated to automate the cleaning process and minimize water usage. By adjusting water consumption based on the detected dust level, the method ensures resource efficiency, reducing both operational costs and environmental impact. This comprehensive solution offers a cost-effective and scalable alternative to existing methods.

The main contributions of this article are summarized as follows:

- Introduces a novel hybrid CNN-RF model for precise classification of photovoltaic (PV) panel operational states, utilizing IV curve parameters converted to RGB mosaic images, thereby enhancing diagnostic accuracy and efficiency.
- Develops an advanced automated dust detection system that categorizes dust accumulation levels, enabling timely and targeted cleaning to optimize panel performance.
- Implements a water-efficient cleaning strategy that automates cleaning processes based on dust detection, minimizing water consumption and operational costs while maintaining optimal power output.

The paper is divided as follows. Section 2 discussed the impact of several operating scenarios and their effect on the I-V curve characteristics of PV systems. Section 3 details the proposed methodology, which involves transforming I-V curves into RGB mosaic images for enhanced classification using a hybrid CNN-RF model. This section also introduces the automated cleaning system enabled by dust level classification. Moreover, the experimental setup demonstrated in this study is depicted in Section 4. The results and key findings are presented in Section 5. Finally, Section 6 concludes the paper and discusses potential directions for future work.

2. Impact of different operation conditions on PV electrical characteristics

The I-V curve is a vital diagnostic tool for understanding the electrical performance of photovoltaic (PV) panels. It represents the relationship between current and voltage under varying operating conditions, providing insights into the panel's operational state. The I-V curve is obtained by varying the load across the PV panel while measuring the resulting current and voltage at each step. This is typically done using a high-precision I-V curve tracer, which sweeps the voltage from zero to the open-circuit voltage (V_{oc}) while simultaneously

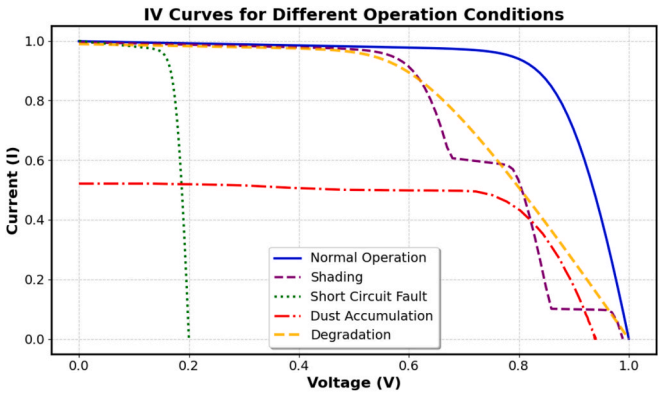

Fig. 1. PV I–V curves under different operational conditions.

recording the corresponding current, starting from the short-circuit current (I_{sc}). In ideal circumstances, as shown by the blue curve in Fig. 1 (Normal Operation), the I-V curve progresses smoothly from a high short-circuit current at zero voltage to the open-circuit voltage with zero current. Between these two points lies the maximum power point (MPP), where the product of current and voltage is maximized, indicating optimal panel performance.

When environmental and system-level factors come into play, the I-V curve exhibits distinct signatures that correspond to specific issues. Shading introduces uneven illumination across the PV cells, often resulting in multiple steps or kinks along the curve as bypass diodes activate to prevent reverse current flow. These characteristics, shown in Fig. 1, illustrate the substantial reduction in power output caused by partial or complete shading.

Dust accumulation, a common issue in arid environments, forms a thin layer on the PV panel surface, reducing sunlight penetration. This primarily affects the current output, as seen in the vertical section of Fig. 1, while the open-circuit voltage remains largely stable. The diminished current generation leads to significant losses in power output and efficiency.

Short-circuit faults, arising from unintended electrical connections, produce abrupt changes in the I-V curve, with sharp drops in both current and voltage. These discontinuities, evident in Fig. 1, indicate abnormal operation that demands immediate corrective measures to avoid further system damage.

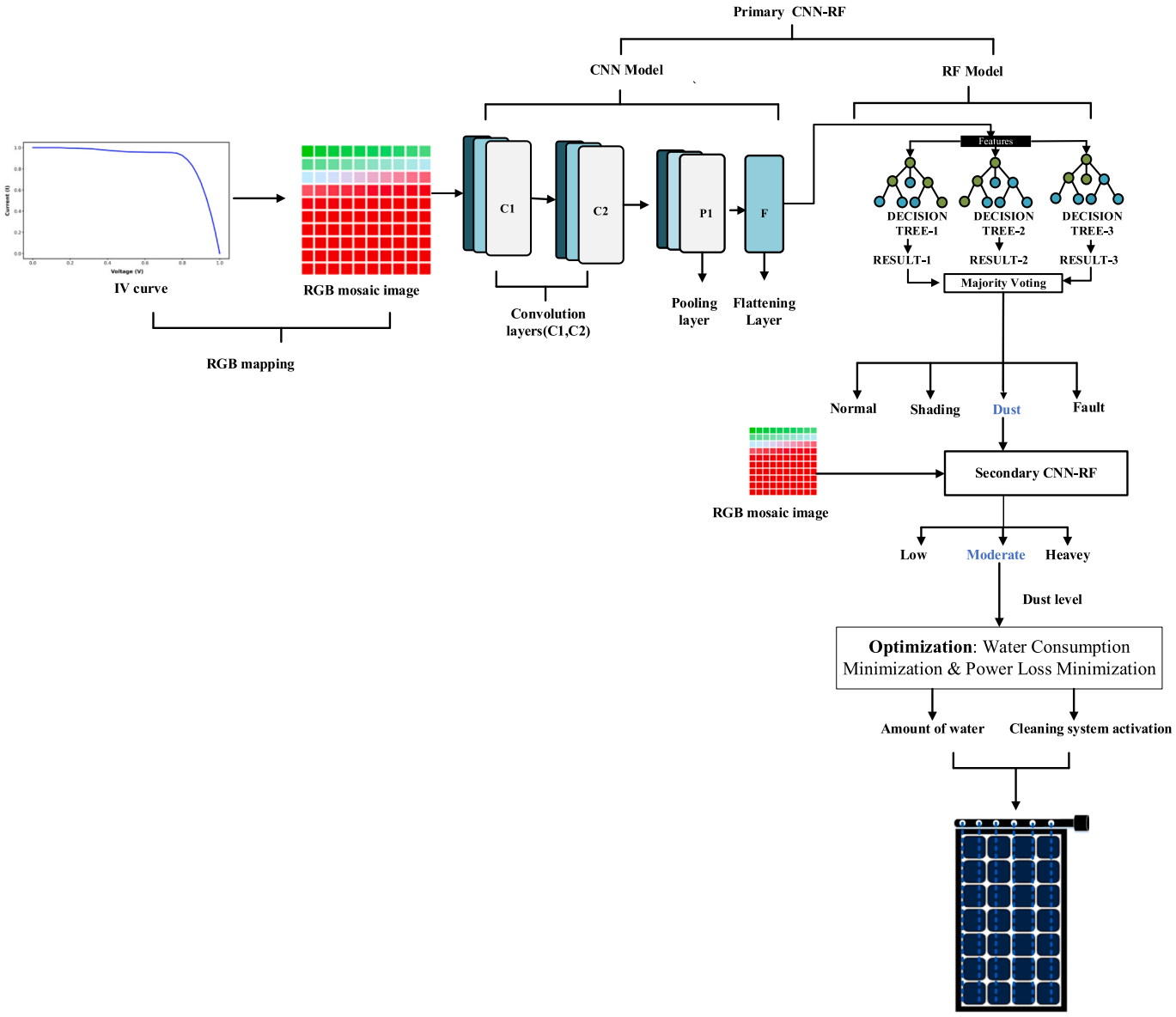

Fig. 2. Workflow of the proposed system for automated PV panel cleaning.

Degradation is another factor that gradually shifts the I-V curve downward due to material wear and prolonged environmental exposure. Unlike abrupt changes caused by shading or faults, degradation results in a smoother and more uniform decline in performance, as depicted in Fig. 1.

3. Methodology

This section presents the proposed methodology for enhancing photovoltaic (PV) panel efficiency through automated dust detection and optimized cleaning. The approach integrates three main stages: transforming I–V curve data into RGB mosaic images, classifying panel states using machine-learning models, and minimizing water consumption during cleaning. The complete workflow is illustrated in Fig. 2.

The process begins with the acquisition of I-V curve data, which includes voltage V_i, current I_i, and calculated power P_i under various operational conditions. Each (V, I, P) triplet is normalized to the range 0–255 and mapped to the red, green, and blue channels, respectively, to create RGB color triplets. These triplets are arranged into a $G \times G$ grid, where $G = \sqrt{N}$and N is the number of data points. Each grid cell is then upsampled into a larger tile, producing RGB mosaic image. This image preserves a one-to-one correspondence between raw data and tiles, providing sufficient resolution for both human inspection and convolutional neural network (CNN) feature extraction.

The resulting mosaic images are fed into a hybrid CNN–Random Forest (CNN-RF) model for primary classification of PV panel states, which include normal operation, shading, dust accumulation, short-circuit faults, and degradation. If dust accumulation is detected, a secondary CNN-RF model classifies dust severity as low, moderate, or heavy. Finally, an optimization procedure minimizes water consumption and power loss during cleaning, ensuring efficient panel operation while conserving resources.

The following subsections provide detailed explanations of each step. Section 3.2 describes the mapping of I-V curve data into RGB-encoded mosaic images. Section 3.3 explains the classification process using the hybrid CNN-RF model, and Section 3.4 discusses the optimization of the cleaning process based on classified dust levels.

3.1. Mapping I–V curve data to upsampled RGB encoded mosaic images

The transformation of PV I-V curve data into an RGB mosaic image enables visualization of PV panel operational states. This process converts electrical parameters (voltage, current, and power) into color-coded, tile-based images, as shown in Fig. 3. The following steps outline the procedure, with an example provided for clarity.

Step 1: Extracting Electrical Parameters from the IV Curve.

Each data point on the I-V curve is characterized by three key electrical parameters: voltage V_i, current I_i, and power P_i, where power is calculated as $P_i = V_i \times I_i$. These parameters form a three-dimensional data vector: $E_i = [V_i, I_i, P_i]$. This vector captures the essential electrical properties of the PV panel at each operating point.

Step 2: Normalization and RGB Mapping.

Since the RGB color space requires values in the range of 0–255, the raw electrical parameters are normalized as illustrated in (1). The normalization process ensures that all values fit within the RGB range while maintaining the proportional relationships between voltage, current, and power.

$$V_{norm} = 255^{*}\frac{V_{i-}V_{min}}{V_{max} - V_{min}}$$

$$I_{norm} = 255^{*}\frac{I_{i-}I_{min}}{I_{max} - I_{min}} \quad (1)$$

$$P_{norm} = 255^{*}\frac{P_{i-}P_{min}}{P_{max} - P_{min}}$$

Where $I_{max}, I_{min}, V_{max}, V_{min}, P_{max}, P_{min}$ are the minimum and maximum values of current, voltage, and power, respectively.

The normalized values are then assigned to the RGB channels, where each parameter corresponds to one color channel:

- **Red (R):** Normalized voltage value(V_{norm}).
- **Green (G):** Normalized current value (I_{norm}).
- **Blue (B):** Normalized power value (P_{norm}).

The combination of the three normalized channels forms an RGB triplet that defines the pixel's color, directly representing the relative magnitudes of the normalized voltage (V_{norm}), current (I_{norm}) and power (P_{norm}). For instance, a high (V_{norm}) compared to (I_{norm}) and (P_{norm}) produces a predominantly red pixel. Similarly, elevated I_{norm} results in a green hue, and high (P_{norm}) yields a blue hue. When two or more parameters share comparable intensities, mixed hues emerge, reflecting a more balanced contribution among voltage, current, and power.

Step 3: Grid Formation and Pixel Placement.

After converting each data point into an RGB triplet, the triplets are arranged into a two-dimensional grid to form the RGB mosaic image. The size of the grid (G) is determined by the equation:

$$G = \left\lceil \sqrt{N} \right\rceil \quad (2)$$

Where N is the total number of I–V curve data points. The square root provides the side length of the smallest square grid that can contain all samples. The ceiling function $[\cdot]$ ensures the grid size is large enough to accommodate all (N) data points, even if (N) is not a perfect square. This results in a $G \times G$ grid with sufficient cells, leaving some empty if (N) is less than G^2.

Each data point (with index $k = 0, 1, \cdots\cdots N-1$) is mapped to a grid position with row i and column j using:

$$i = \left[\frac{k}{G}\right], j = Kmode\left(G\right) \quad (3)$$

Instead of representing each cell as a single pixel, the RGB triplet assigned to cell (i, j) is upsampled into a uniform tile (e.g., 50×50,). This tile–based rendering preserves the one–to–one mapping between data and image while producing a resolution high enough for both human inspection and CNN feature extraction.

To illustrate the complete transformation process, consider a

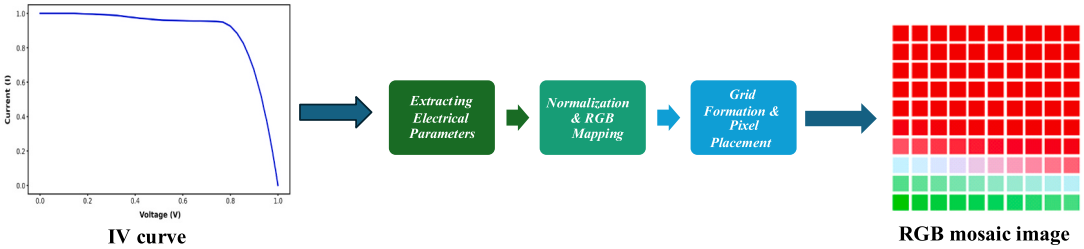

Fig. 3. Converting I–V Curve Data into an RGB mosaic image.

Table 2
I–V data points with corresponding RGB triplets after normalization.

K	V_i	I_i	P_i	(V_{norm},I_{norm},P_{norm}) RGB Triplet
0	20	4	80	(0,255,0)
1	40	3	120	(85,170,255)
2	60	2	120	(170,85,255)
3	80	1	80	(255,0,0)

Table 3
Grid positions for each data point index (k) calculated using Eq. (3).

K	$i = \left[\frac{k}{G}\right]$	$j = Kmode(G)$	Grid Position
0	$\left[\frac{0}{2}\right] = 0$	$0mode(2) = 0$	Row 0, Column 0
1	$\left[\frac{1}{2}\right] = 0$	$1mode(2) = 1$	Row 0, Column 1
2	$\left[\frac{2}{2}\right] = 1$	$2mode(2) = 0$	Row 1, Column 0
4	$\left[\frac{3}{2}\right] = 1$	$3mode(2) = 1$	Row 1, Column 1

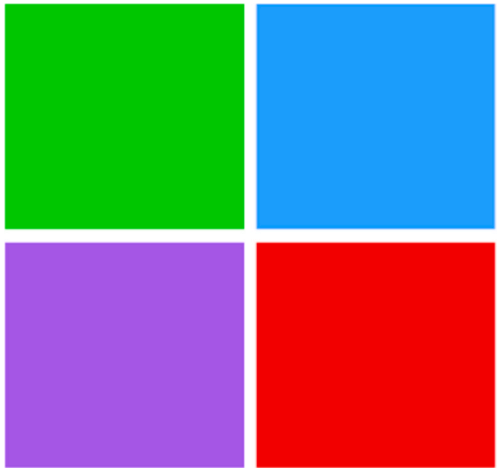

Fig. 4. RGB mosaic image resulting from the 4-point example.

simplified example with four I–V data points obtained from a PV panel. Table 2 presents the measured voltage V_i, current I_i, and power P_i , along with their normalized RGB values. These values are normalized to the range 0–255 using Eq. (1), assuming the following bounds:

$V_{min} = 20v$, $V_{max} = 80v$

$I_{min,} = 1A$, $I_{max} = 4A$

$P_{min} = 80\text{W}$, $P_{max} = 120W$

The grid size is then calculated as:

$$G = \left[\sqrt{4}\right] = 2$$

This results in a 2 × 2 grid with four cells. Each RGB triplet obtained from the normalized voltage, current, and power values is placed into the grid based on its index K, using the coordinate mapping defined in Eq. (3), arranged in row-major order as shown in Table 3.

Each grid cell is then upsampled into a 50 × 50 pixel block, resulting in a 2 × 2 RGB-encoded mosaic image as illustrated in Fig. 4.

Applying this RGB mosaic visualization method generally produces distinct patterns that effectively differentiate various PV panel states, as demonstrated in Fig. 5. Under normal operational conditions, the mosaic presents a continuous vertical gradient. Specifically, the upper rows predominantly exhibit green hues, representing high current at lower voltages, transitioning toward reds and blues in the lower rows, indicative of increased voltage and power. Conversely, short-circuit faults initially produce green-dominant tiles similar to the normal condition; however, a noticeable decrease in red intensity appears in the lower mosaic rows, clearly reflecting a sudden voltage collapse. This similarity in visual patterning between normal and fault conditions can initially appear ambiguous, yet Fig. 6 clarifies this by showing the numeric RGB channel intensities for each tile, confirming a significant suppression in voltage and power during short-circuit conditions.

In cases of shading, the resulting mosaic images display irregular, patchy patterns dominated by purples and pinks. These visual discontinuities correspond directly to uneven illumination conditions and frequent activation of bypass diodes, disrupting the smooth gradient typically observed during normal operation. When dust accumulation occurs, the mosaic images become predominantly red. This effect arises from a reduced current, represented by a diminished green channel, while the voltage, encoded as red, remains relatively stable. This disparity becomes particularly pronounced toward the lower portion of the mosaic image. Finally, long-term degradation manifests as uniformly muted colors of purples and blues, reflecting a gradual and simultaneous reduction in both voltage and current levels.

By translating each I–V data point into a distinct RGB color tile, subtle variations in electrical parameters are transformed into immediate visual distinctions. Consequently, this RGB mosaic visualization enables rapid differentiation among shading conditions, electrical faults, dust deposition, and long-term degradation features often challenging to discern in traditional I–V curve plots. Furthermore, the two-dimensional RGB representation maintains essential spatial gradients, textures, and anomaly signatures, substantially enhancing the performance of machine-learning models and image-analysis algorithms. As a result, this approach provides improved accuracy and robustness in automated classification and anomaly detection of PV panel operational state.

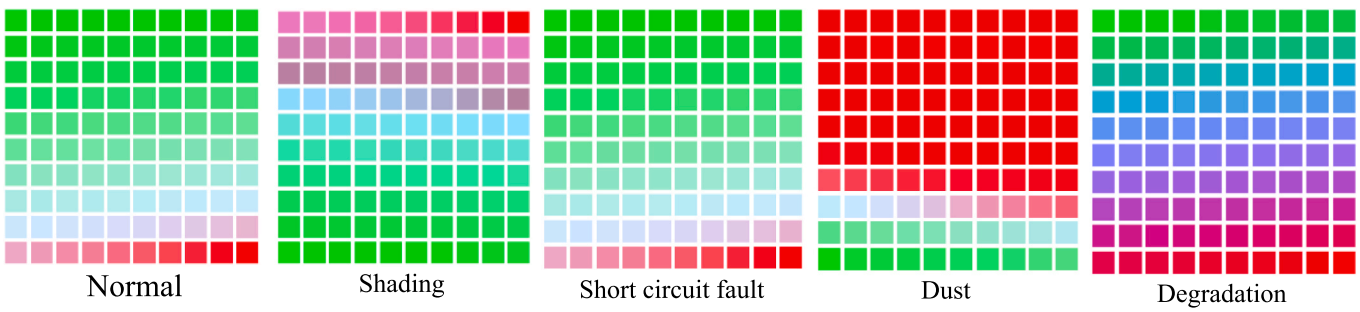

Fig. 5. RGB mosaic images of different PV operation conditions.

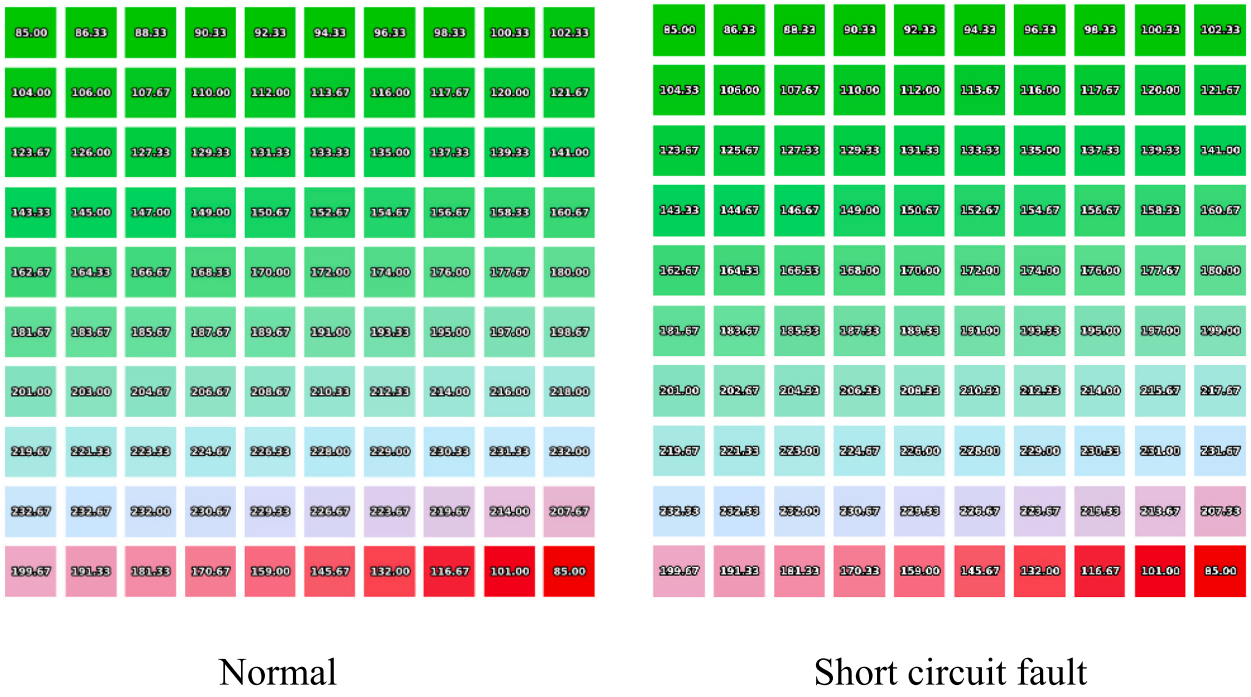

Fig. 6. RGB mosaic images with numeric intensities for normal and short-circuit conditions.

3.2. Classification using hybrid CNN-RF model

The RGB-encoded mosaic images generated from the PV I-V curves are utilized to classify the operational states of PV panels. This is achieved through a hybrid classification model that combines a two-dimensional CNN (2D-CNN) with an RF classifier. The CNN acts as a feature extractor, leveraging its convolutional and pooling layers to process the RGB mosaic images and capture meaningful spatial and contextual patterns as shown in Fig. 2. These extracted features are then fed into the RF classifier, which categorizes the operational state into predefined classes, including normal operation, shading, degradation, short-circuit faults, and dust accumulation. If the primary classification detects dust accumulation, a second CNN-RF model is employed for secondary classification. This secondary model further refines the analysis, categorizing the severity of dust accumulation into low, moderate, or heavy levels. Each CNN-RF model is independently trained and optimized for its specific classification task, ensuring precise identification of both the operational state and the severity of dust accumulation.

In the proposed hybrid CNN-RF model, the 2D-CNN takes the RGB mosaic images denoted as

$I(x,y,c)$ where x and y represent the spatial coordinates of the image grid and c denotes the color channels (red, green, blue) as input. The CNN processes $I(x,y,c)$ through a series of convolutional, pooling, and fully connected layers to extract a feature vector that captures the spatial and contextual patterns indicative of PV panel performance. The convolutional layer, the core of the CNN, applies filters to detect patterns within the RGB mosaic images. For the $k-th$ filter in the $m-th$ convolutional layer, the convolution operation is expressed as:

$$F_m(x,y,k) = \varnothing\left(\sum_{c=1}^{3} I(x,y,c) * W_m(c,k) + b_m(K)\right) \tag{4}$$

Where:

- $F_m(x,y,k)$: The feature map value at position at position (x,y) in the k-th layer filter in the m-th layer.
- $I(x,y,c)$: The input RGB-encoded mosaic images value at spatial position (x,y) in color channel c
- $W_m(c,k)$: Convolutional kernel for the c-th color channel in the k-th filter of the m-th layer
- $b_m(k)$: Bias term for the k-th filter in the m-th convolutional layer.
- $\varnothing$: Activation function (ReLU).

In this study, the Rectified Linear Unit (ReLU) activation function is used, which introduces non-linearity into the network to model complex relationships in the data. The ReLU activation function is defined as:

$$\varnothing(z) = \max(0,z) \tag{5}$$

Where z is the input to the activation function. The ReLU operation ensures that only positive feature values are passed forward, mitigating the vanishing gradient problem and improving computational efficiency.

Following the convolutional layer, a pooling layer is applied to reduce the spatial dimensions of the feature maps while preserving the most salient features. Maximum pooling is employed in this study, and the operation is expressed as:

$$P_m(x,y,k) = \max(F_m(x,y,k)) \tag{6}$$

Where $P_m(x,y,k)$ is the pooled feature map value at position (x,y) for the k-th filter in m-th layer and $F_m(x,y,k)$ is the convolutional layer output. The pooling operation reduces the computational complexity and improves the model's ability to focus on the most critical features. After multiple convolutional and pooling layers, the resulting feature maps are flattened into a one-dimensional feature vector (V)

$$V = Flatten(P_m) \tag{7}$$

This feature vector V is then passed to the RF classifier, which serves as the final decision-making component in the hybrid model. The RF classifier, an ensemble of decision trees, predicts the operational state by aggregating the outputs of individual trees trained on random subsets of V. The final class prediction is determined by majority voting:

$$C = mode\{C_t(V)|t = 1, 2, \cdots, T\}) \tag{8}$$

Where:

- $C_t(V)$: The class prediction by the t-th decision tree based on feature V.
- C: The final predicted class
- T: The number of decision trees
- *mode*: The most frequent class across all trees

The proposed hybrid CNN-RF model implementation involves two sequential phases. First, the CNN is independently trained on RGB-encoded mosaic images, enabling it to learn and extract highly discriminative spatial and contextual features. Following this training, the fully connected classification layers of the CNN are removed, leaving only the convolutional, pooling, and flattening layers dedicated to feature extraction. In the second phase, the RF classifier independently utilizes these extracted features to perform classification tasks, benefiting from its robust ensemble structure to improve generalization and interpretability. The parameters utilized in the CNN-RF model are summarized in Table 4. As depicted in Fig. 2.

If the primary CNN-RF model identifies the operational state dust it triggers the secondary CNN-RF model. The secondary model follows the same architecture but is trained specifically to classify dust severity levels (low, moderate, heavy) using the same RGB mosaic images input $(I(x,y,c))$. The extracted features from the secondary CNN are similarly fed into its RF classifier to determine the dust level, enabling the optimization of the cleaning process outlined in Section 3.3.

Table 4
CNN-RF model parameters.

Model	Layer	Parameter	Value
CNN	Conv 2D (1)	Activation	ReLU
		Filters	32
		Kernel Size	(3, 3)
	MaxPooling 2D (1)	Pool Size	(2, 2)
	Conv 2D (2)	Activation	ReLU
		Filters	64
		Kernel Size	(3, 3)
	MaxPooling 2D (2)	Pool Size	(2, 2)
	Conv 2D (3)	Activation	ReLU
		Filters	128
		Kernel Size	(3, 3)
	MaxPooling 2D (3)	Pool Size	(2, 2)
	Fully Connected (1)	Number of Units	128
		Activation	ReLU
		Dropout Rate	0.5
	Fully Connected (2)	Number of Units	Number of Classes
		Activation	Softmax
RF	–	Number of Trees	100
		Random State	42

3.3. Optimization of PV panel cleaning

The cleaning process for PV panels is optimized based on dust accumulation severity, classified using a secondary CNN-RF model. This methodology integrates dust severity classification with associated power loss data to minimize water usage and maintain optimal panel performance. An optimization framework evaluates the cleaning cost against the cost of power loss due to dust accumulation, formulated through the following objective function:

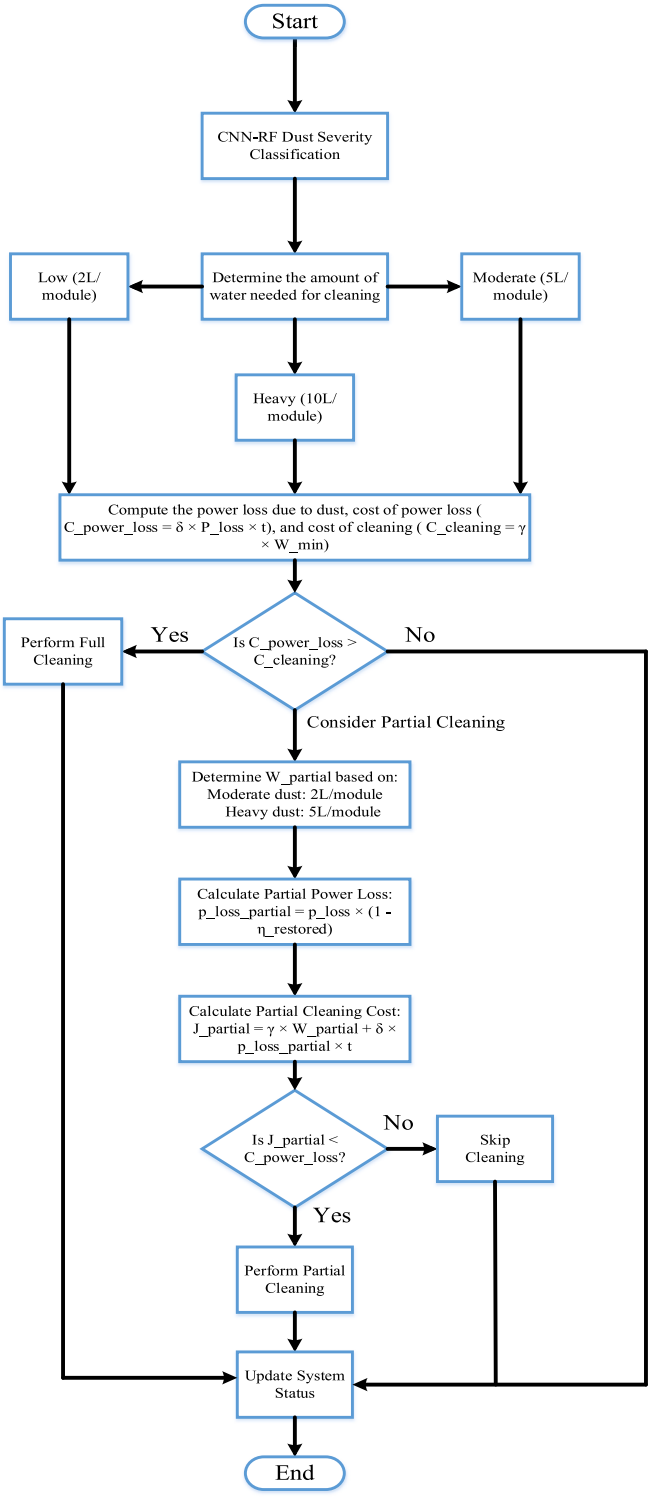

Fig. 7. Flow chart of the automated cleaning process.

Table 5
Description of PV module specifications.

Model No.	EGE-55P-36
Maximum Power at Standard Test Conditions (STC)	55 W
Maximum Operating Voltage (Vmpp)	19.11 V
Maximum Operating Current (Impp)	2.88 A
Open Circuit Voltage (Voc)	23.41 V
Short Circuit Current (Isc)	3.08 A
Temperature Coefficient of Power (%/°C)	−0.41 %/°C
Temperature Coefficient of Voltage (%/°C)	−0.31 %/°C
Temperature Coefficient of Current (%/°C)	+0.06 %/°C

$$\min(J(W) = \gamma \bullet W + \delta \bullet P_{loss} \bullet t) \tag{9}$$

Where:

- W: is the amount of water used for cleaning (in liters).
- P_{loss} : denotes the power loss caused by dust accumulation (in watts).
- γ: is the water cost factor (AED/Liter)
- δ: is the electricity cost factor (AED/kWh)
- t: is the time period over which the cost of power loss is calculated.

Cleaning is recommended when the cost of power loss $C_{PoweLoss}$ exceeds the cost of cleaning $C_{cleaning}$, as expressed in the following conditions:

$$C_{PoweLoss} = \delta \bullet P_{loss} \bullet t \tag{10}$$

$$C_{cleaning} = \gamma \bullet W \tag{11}$$

Hence, Cleaning is performed if:

$$C_{PoweLoss} > C_{cleaning} \tag{12}$$

It is worth noting that in the cost model, only water consumption is considered as a variable cleaning cost. This is because the proposed system operates autonomously without human intervention, rendering labor costs inapplicable. Moreover, the system uses only water for cleaning, without the need for chemical agents or consumables. In addition, equipment maintenance costs are excluded from the decision-making process, as they are periodic and not directly influenced by the frequency of cleaning operations. Therefore, the decision logic focuses on dynamically balancing water cost against energy loss to ensure practical, efficient, and real-time cleaning optimization.

In this study, water usage is specifically defined for the PV modules under analysis based on the dust severity level:

$$W_{min} = \begin{cases} 2liters/module\ (low\ dust), \\ 5liters/module\ (moderate\ dust), \\ 10liters/module\ (heavy\ dust), \end{cases} \tag{13}$$

For cases where full cleaning is not immediately cost-effective ($C_{PoweLoss} < C_{cleaning}$), partial cleaning is considered. In this study, partial cleaning uses:

$$W_{partial} = \begin{cases} 2liters/module\ (moderate\ dust) \\ 5liters/module\ (heavy\ dust) \end{cases} \tag{14}$$

And restores 50 % of the efficiency loss, represented as:

$$p_{loss.partial} = p_{loss} \bullet (1 - \eta_{restored}) \tag{15}$$

Where $\eta_{restored} = 50\%$ is specifically chosen for this case based on the observed system performance. The total cost after partial cleaning is calculated as:

$$J_{partial} = \gamma \bullet W_{partial} + \delta \bullet p_{loss.partial} \bullet t \tag{16}$$

The values used in this study, including water usage and partial efficiency restoration, are specific to the PV modules and operational scenarios analyzed, ensuring tailored optimization for the system's characteristics. This approach balances operational costs and resource efficiency, guiding decisions between full and partial cleaning based on the recalculated costs and observed severity levels. The automated cleaning process is further illustrated in Fig. 7.

4. Experimental setup

To further validate the feasibility and performance of the proposed method in practical applications, photovoltaic (PV) modules were installed on the roof of the W-12 building at the main campus of the University of Sharjah, located at a latitude of 25.2879° N and a longitude of 55.4784° E in Sharjah, UAE. The PV modules, each with a capacity of 55 W, were tested under various operation conditions, including normal operation, dust deposition, shading, degradation, and short-circuit faults. The detailed specifications of the PV modules used in this study are provided in Table 5. Moreover, the weather conditions under which the experiment was conducted are depicted in Fig. 8.

The measurement of I-V characteristics under these experimental conditions was conducted using a high-precision Profitest device, which is essential for accurately obtaining I-V curves. The Profitest device also functioned as an electrical load, ensuring stable and consistent measurements across varying environmental conditions. With a measuring accuracy of ± 5 %, the Profitest provided reliable data collection over extended periods, making it suitable for long-term monitoring and analysis.

In this experimental setup, a dedicated Profitest PV analyzer was used for each PV module to obtain high-precision I–V curve measurements under varying weather conditions. This approach was appropriate for research purposes, where detailed validation was required. However, in practical large-scale deployments, a single I–V curve tracer can be used to monitor entire arrays or individual strings, significantly reducing equipment costs. Moreover, many modern PV inverters now

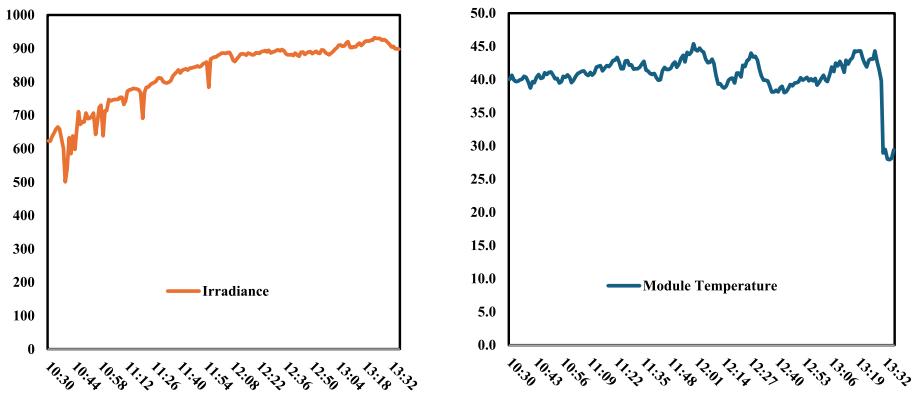

Fig. 8. Weather conditions during the experimental period.

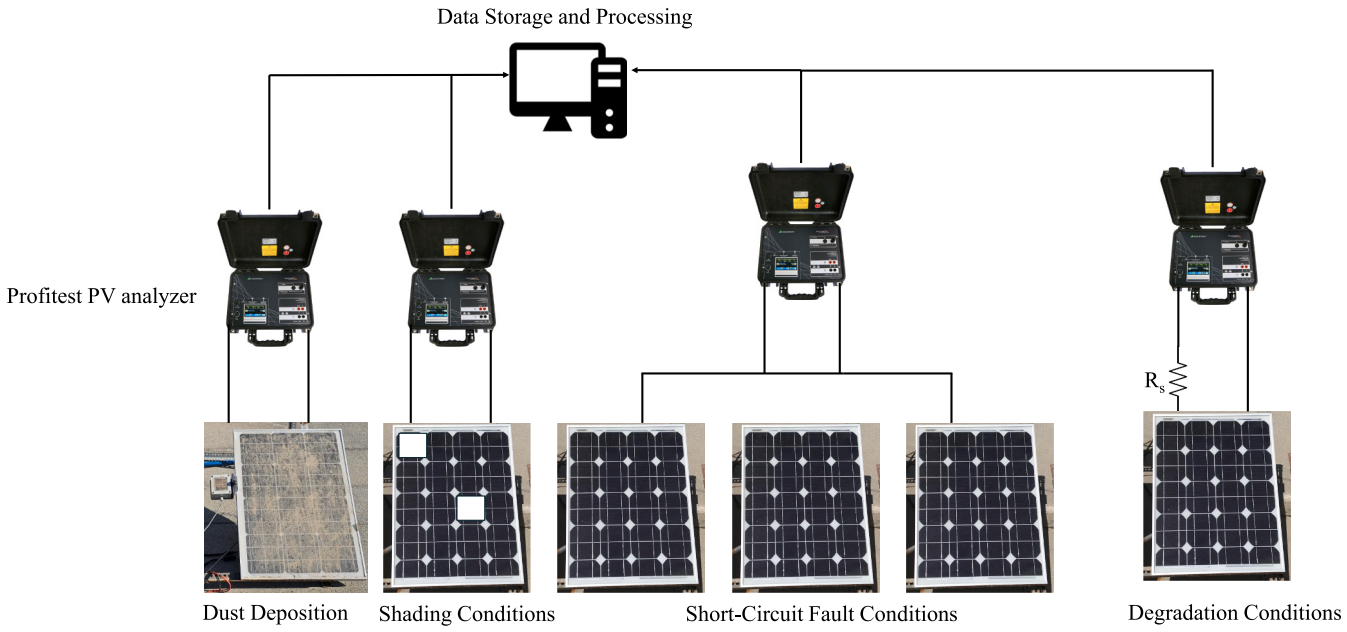

Fig. 9. Experimental Setup demonstrating various operating conditions.

Table 6
Sample Count of data utilized for classification.

Operating Condition		Sample Count
Standard Conditions		534
Shading Condition		534
Short Circuit Fault		534
Degradation		534
Dusty Condition	Low	178
	Medium	178
	Heavy	178
	Total	534

include built-in I–V curve tracing functionality, enabling seamless and low-cost acquisition of I–V data. This ensures that the proposed method remains scalable and cost-effective for real-world applications.

Several scenarios were applied to the PV modules to assess their performance under different conditions. Dust deposition was monitored over a period of more than 60 days, during which three distinct levels of dust accumulated: low, medium, and heavy. Each level of dust introduced varying impacts on the I-V characteristics, with noticeable reductions in current generation, thereby affecting power output.

Partial shading was simulated using cardboard sheets, applied selectively across single, double, or multiple cells to create multiple peaks in the I-V curves. These variations reflect the behavior of the panels under different shading patterns, helping to assess the system's response to real-world conditions. Additionally, short-circuit faults were introduced using MC4 connector wires to short between PV modules, resulting in voltage drops and corresponding reductions in power output. Additionally, degradation was simulated by connecting a cement resistor in series with the PV module, introducing resistance that mimics the effects of long-term wear and tear on the module's electrical performance.

The illustration of the experimental setup is provided in Fig. 9, which highlights the arrangement of PV modules, environmental simulation components, and the Profitest device used for continuous measurements.

The collected I-V curves were transformed into RGB mosaic images, which were used as input to the hybrid CNN-RF model for classification. These RGB mosaic images enabled the identification of different operational states and faults. Table 6 summarizes the number of RGB mosaic images generated for each operational scenario.

5. Results

The proposed method was implemented using Python, with Google Colab serving as the development platform to leverage its computational resources for deep learning. TensorFlow and Keras were utilized as the primary frameworks for constructing and training the CNN models, while Scikit-learn was employed for specific tasks such as training the RF classifier, generating classification reports, calculating accuracy scores, and performing train-test splits. Additionally, core Python libraries were used for data manipulation and processing.

The dataset, comprising RGB mosaic images generated from the I-V curves of PV modules under various operational states, was organized into training, validation, and testing directories. Each image was resized to ensure uniform input dimensions for the CNN models.

Table 7
Classification Report of the PV Operational States.

CNN-RF					
Operational States	**Precision**	**Recall**	**F1-Score**	**Support**	**Overall Accuracy**
Normal Operation	1.0	1.0	1.0	107	1.0
Short Circuit Fault	1.0	1.0	1.0	107	
Shading	1.0	1.0	1.0	107	
Dust Accumulation	1.0	1.0	1.0	107	
Degradation	1.0	1.0	1.0	107	
CNN					
Operational States	**Precision**	**Recall**	**F1-Score**	**Support**	**Overall Accuracy**
Normal Operation	0.5	1.0	0. 67	107	0.80
Short Circuit Fault	0	0	0	107	
Shading	1.0	1.0	1.0	107	
Dust Accumulation	1.0	1.0	1.0	107	
Degradation	1.0	1.0	1.0	107	

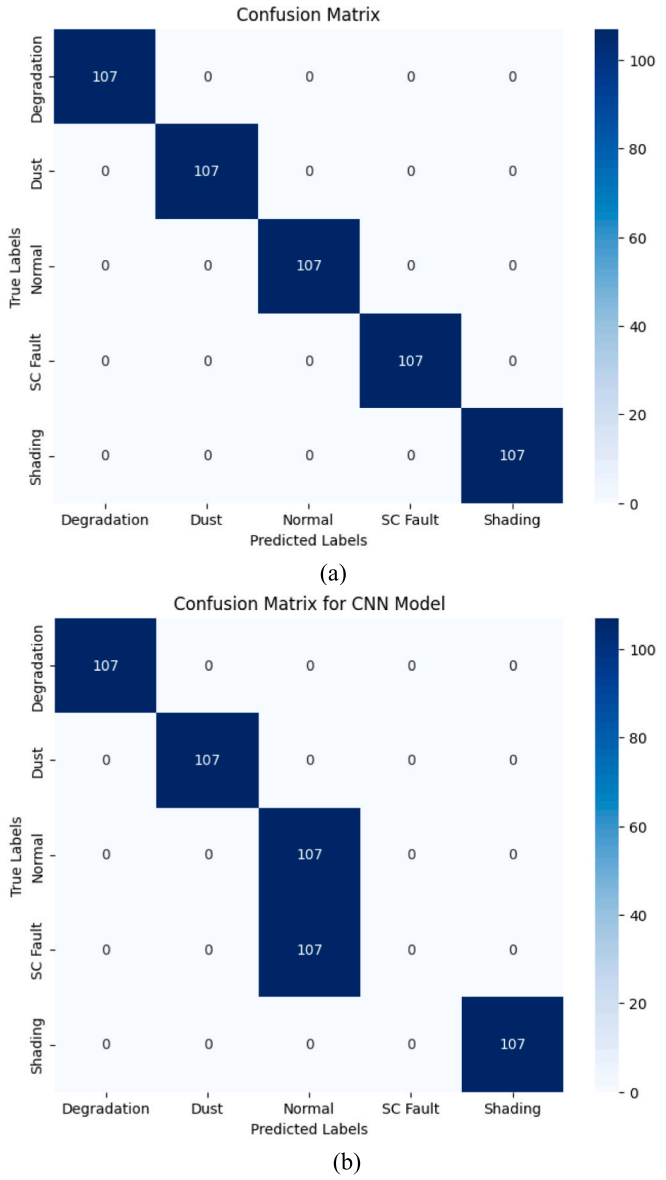

Fig. 10. Operational States confusion matrix: (a) CNN-Rf (b) CNN.

The primary CNN-RF model demonstrated exceptional performance in classifying PV panel operational states, including normal operation, shading, degradation, short-circuit faults, and dust accumulation. As shown in Table 7, the model achieved 100 % accuracy across all operational states, with precision, recall, and F1-scores of 1.0 for each category. The confusion matrix in Fig. 10 (a) confirms zero misclassifications, underscoring the effectiveness of combining CNN's feature extraction capabilities with RF's robust decision-making. To validate that the model's 100 % classification accuracy was not a result of overfitting, we conducted a performance consistency check across the training, validation, and test datasets. As shown in Fig. 11, the training and validation loss curves converge steadily, exhibiting similar trends with no significant divergence. This indicates that the CNN component generalizes well during feature extraction. Furthermore, the Random Forest classifier, trained on the features extracted by the CNN, achieved identical accuracy scores across all datasets consisting of training, validation, and testing.

For comparison, a standalone CNN model was trained and tested on the same dataset. The standalone CNN achieved an overall accuracy of 80 %, significantly lower than the CNN-RF model. The confusion matrix in Fig. 10 (b) highlights the CNN's struggle to correctly classify short-

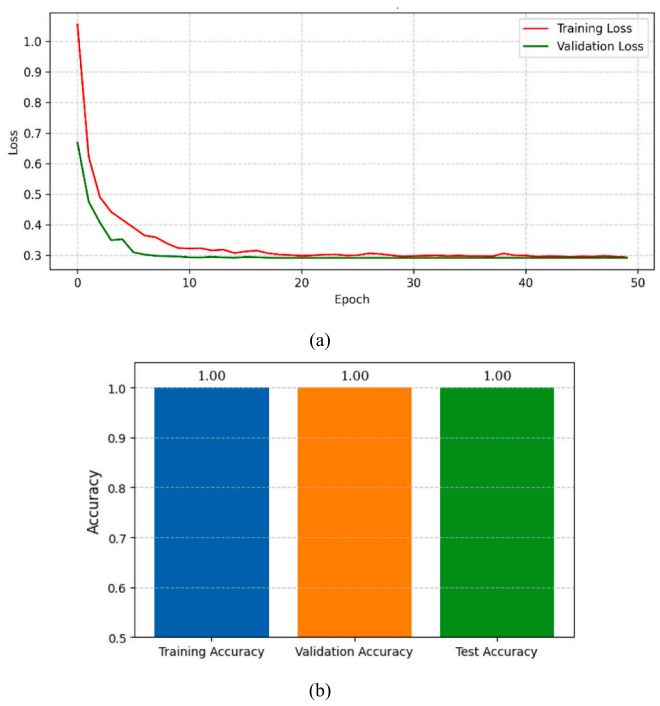

Fig. 11. Model evaluation for CNN-RF using RGB mosaic images: (a) CNN loss curves (b) RF accuracy across datasets.

circuit faults. This comparison highlights the superior classification performance of the CNN-RF model, which effectively reduces misclassifications by leveraging the complementary strengths of both models.

After the primary CNN-RF model detected dust accumulation, a secondary CNN-RF model was employed to classify the dust levels into low, moderate, and heavy categories. This secondary model also demonstrated strong performance, achieving an overall accuracy of 98 %, as shown in the confusion matrix in Fig. 12 (a) and Table 8. Despite the high accuracy, the model misclassified 2 images: 1 moderate dust image was classified as low, and 1 low dust image was classified as moderate. These misclassifications were minimal, demonstrating the model's reliability in categorizing dust levels, thereby supporting optimized maintenance strategies that reflect the actual severity of dust accumulation.

The standalone CNN model was also evaluated for dust level classification and achieved an overall accuracy of 72 % as presented in Table 8. The confusion matrix in Fig. 12(b) reveals that the CNN struggled significantly with distinguishing between low and moderate dust levels. The model misclassified 20 Moderate dust images as Low and 8 Low dust images as Moderate, along with 2 Moderate dust images misclassified as Heavy. This results in a total of 30 misclassifications, highlighting a notable gap in performance compared to the CNN-RF model. These misclassifications underscore the limitations of the standalone CNN in accurately differentiating dust levels, particularly in scenarios involving moderate dust accumulation.

To justify the use of RGB mosaic images rather than I–V curve images, the I–V curves were likewise employed to classify five operational states in Stage 1, then extended to dust severity classification in Stage 2. In Stage 1, the CNN-RF model trained on I–V curve images achieved a respectable 97 % overall accuracy, with most classes accurately identified. However, 16 short circuit fault samples were misclassified as normal, as illustrated in Fig. 13(a), causing a slight decrease in recall (0.85) and F1 score (0.92) for the short circuit fault class, as shown in Table 9. These findings confirm the superiority of the RGB mosaic images, which attained a 100 % accuracy rate, and suggest that the richer visual cues in RGB mosaic images provide clearer insight into certain PV

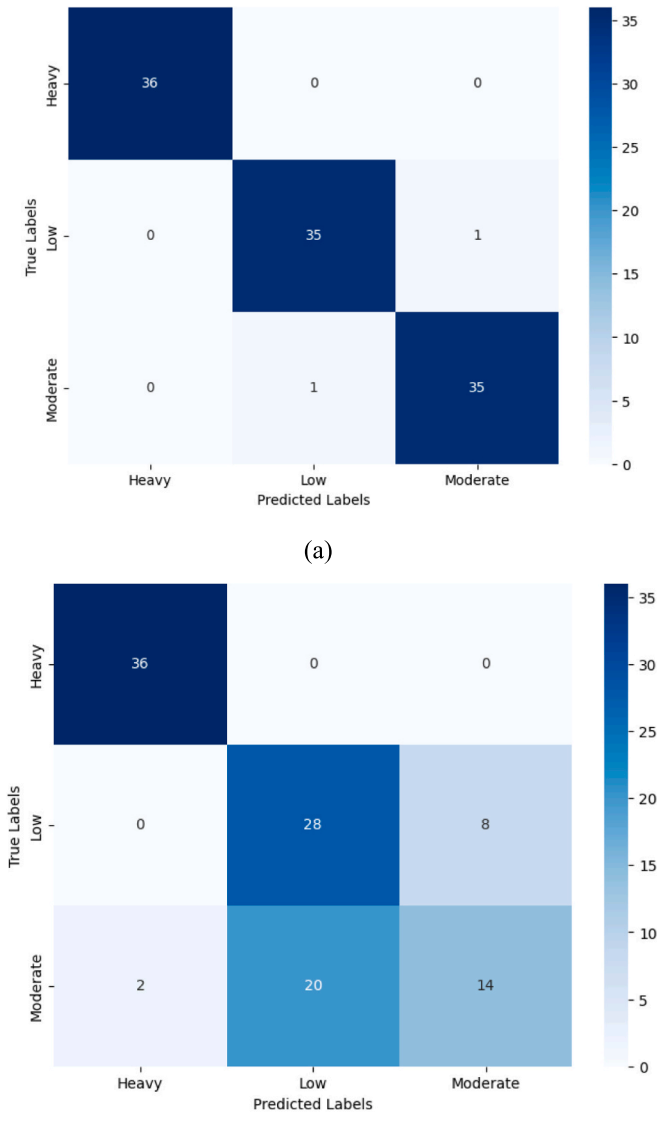

Fig. 12. Dust level confusion matrix (a) CNN-RF (b) CNN.

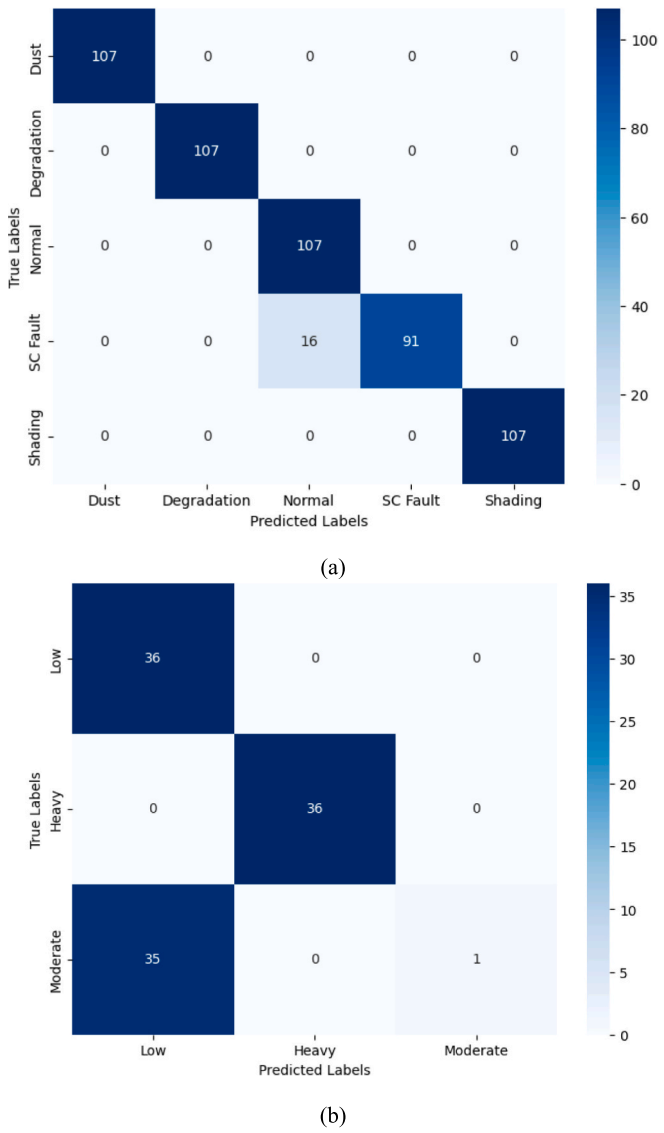

Fig. 13. IV curves confusion matrix: (a) PV Operational States (b) Dust Level.

Table 8
Dust Level Classification Report.

CNN-RF					
Dust Level	**Precision**	**Recall**	**F1-Score**	**Support**	**Overall Accuracy**
Heavy	1.0	1.0	1.0	36	0.98
Moderate	0.97	0.97	0.97	36	
Low	0.97	0.97	0.97	36	
CNN					
Operational States	**Precision**	**Recall**	**F1-Score**	**Support**	**Overall Accuracy**
Heavy	0.95	1.0	0.97	36	0.72
Moderate	0.64	0.39	0.48	36	
Low	0.58	0.78	0.67	36	

Table 9
I-V Curves Classification Report.

PV Operational States					
Operational States	**Precision**	**Recall**	**F1-Score**	**Support**	**Overall Accuracy**
Normal Operation	0.87	1.0	0.93	107	0.97
Short Circuit Fault	1.0	0.85	0.92	107	
Shading	1.0	1.0	1.0	107	
Dust Accumulation	1.0	1.0	1.0	107	
Degradation	1.0	1.0	1.0	107	
Dust Level					
Operational States	**Precision**	**Recall**	**F1-Score**	**Support**	**Overall Accuracy**
Heavy	1.0	1.0	1.0	36	0.68
Moderate	1.0	0.03	0.05	36	
Low	0.51	1.0	0.67	36	

Table 10
Performance comparison of the proposed CNN-RF (RGB) framework against existing dust detection and cleaning methods.

Technique	Accuracy	Computational Resources	Water Usage Efficiency	Scalability	Cost-effectiveness
Proposed Method (RGB mosaic images)	High (100 %)	Moderate (CNN-RF model)	High (optimized cleaning)	High	High
Thermal Imaging	Moderate	High (image processing)	Low	Moderate	low
Drone-based Imaging	High	High (drone operation and image processing)	Low	Low	Low
Pure I-V Curve Analysis	Moderate (97 %)	Low	N/A	High	High
Sensor-based Methods	Moderate	Low	N/A	Moderate	Moderate

operational characteristics compared to images derived from the electrical curve.

In Stage 2, the I–V based CNN-RF model attained 67.59 % accuracy, with frequent misclassification of Moderate dust as Low as shown in Fig. 13 (b). This substantially lowered the Moderate class's F1-score (0.05) as depicted in Table. 9, underscoring the difficulty of using IV curve images for fine-grained dust differentiation.

To demonstrate the adaptability of the proposed automated cleaning method on large-scale PV systems, the experimental system demonstrated in Fig. 9 was scaled up to 5 kW, to investigate practical and economic viability. The assumed large-scale PV system consisted of 91 PV modules, each with a power output of 55 W. According to Sharjah Electricity and Water Authority (SEWA), the tariffs considered were AED 0.028 per liter for water and AED 0.845 per kWh for electricity.

The power loss for each module was initially calculated as the difference between the reference power output of a clean panel and the output of a panel classified as dusty, based on the level determined by the secondary CNN-RF model (low, moderate, or heavy dust). The total power loss for the system was then obtained by multiplying the per-module power loss by 91 to reflect the cumulative impact across all modules. These scenarios evaluate both conditions where cleaning is required and where it is not recommended, depending on the balance between the cost of cleaning and the cost of power loss. The dust scenarios for the PV modules are evaluated daily every 24 h at noon time, to compute the power losses, determining the required cleaning action. Therefore, this method provides a cost-effective solution by optimizing the cleaning operation. This fixed daily evaluation interval was selected to align with peak solar activity, as measurements taken at noon typically reflect the highest irradiance levels and thus the maximum potential power loss due to dust. By keeping the analysis to a consistent reference time, the proposed method ensures a reliable and reproducible assessment of energy loss and cleaning decisions. Although a 24-hour interval is used in this study for demonstration, the method can be extended to shorter or adaptive intervals based on real-time monitoring or environmental variability if higher temporal resolution is desired.

Scenario 1: Low dust level

- Cleaning Required: The system exhibited a total power loss of 500 W. With a water tariff of AED 0.028 per liter and an electricity tariff of AED 0.845 per kWh, the cleaning cost using 2 L of water per module (182 L total) was AED 5.10. The power loss cost over 24 h was AED 10.14. Since the cost of power loss exceeded the cleaning cost, cleaning was recommended, requiring 182 L of water.
- Cleaning Not Required: For a lower power loss of 250 W, the power loss cost over 24 h was AED 5.07, less than the cleaning cost of AED 5.10. In this case, cleaning was not recommended.

Scenario 2: Moderate dust level

- Cleaning Required: With a power loss of 1.25 kW, the cleaning cost using 5 L of water per module (455 L total) was AED 12.74, while the power loss cost over 24 h was AED 25.38. Since the cost of power loss exceeded the cleaning cost, cleaning was recommended.
- Cleaning Not Required: For a power loss of 600 W, the power loss cost over 24 h was AED 12.17, less than the cleaning cost of AED 12.74. Cleaning was therefore not performed.

Scenario 3: Heavy dust level

- Cleaning Required: A power loss of 2.5 kW resulted in a cleaning cost of AED 25.48 using 10 L of water per module (910 L total). The power loss cost over 24 h was AED 50.76, justifying full cleaning.
- Cleaning Not Required: For a power loss of 1.5 kW, the power loss cost over 24 h was AED 30.46, slightly exceeding the cleaning cost of AED 25.48. In this case, cleaning was deferred, and partial cleaning using 5 L of water per module was considered, reducing costs and restoring efficiency.

The cost-effectiveness of the assumed large-scale system is enhanced by its compatibility with commercially available inverters that include I-V curve measurement capabilities, eliminating the need for multiple dedicated analyzers as used in the experimental setup in Fig. 9. Moreover, the automated cleaning system will further reduce the operational costs by optimizing water consumption based on dust severity levels. For the assumed 5 kW system, scenarios demonstrated that cleaning is only initiated when economically justified, ensuring a balance between maintenance costs and power output recovery. This approach makes the system viable for practical deployment in arid regions with high dust accumulation.

To further contextualize the performance of the proposed method, Table 10 compares it with other techniques reported in the literature, including thermal imaging, drone-based imaging, pure I-V curve analysis, and sensor-based methods, across key metrics such as accuracy, computational resources, water usage efficiency, scalability, and cost-effectiveness.

6. Conclusion

In conclusion, this study introduced an innovative hybrid machine learning approach utilizing a combination of CNN and RF classifiers to effectively address the critical challenge of dust accumulation on PV panels, particularly in arid climates. By transforming electrical I–V curve parameters (voltage, current, power) into RGB mosaic images, the proposed method significantly enhanced the precision of detecting various PV panel operational states, including normal operation, dust accumulation, shading, short-circuit faults, and degradation.

The primary CNN-RF model demonstrated superior performance by achieving a perfect accuracy rate of 100 % in classifying these operational states using RGB mosaic images, notably surpassing the 97 % accuracy achieved when traditional I–V curve images were utilized. Furthermore, the secondary CNN-RF model effectively classified dust severity into low, moderate, and heavy levels with an impressive accuracy of 98 %, a substantial improvement compared to the 68 % accuracy obtained using conventional I–V curve-based methods. This explicit comparison highlights the significant advantage and added value of utilizing RGB mosaic images representations over traditional I–V curves for fine-grained operational state and dust-level classifications.

The study also integrated a cost-optimization framework to minimize water usage during cleaning processes while maintaining PV performance. Scenarios were analyzed to demonstrate the adaptability of the proposed method under varying dust levels and power loss conditions. Results highlighted that cleaning is recommended when the cost of power loss exceeds the cleaning cost, with tailored water usage based on

dust severity levels, offering a practical, resource-efficient solution.

This hybrid approach not only outperforms traditional methods in accuracy and resource management but also establishes a scalable and sustainable framework for PV system maintenance. Future work will focus on extending the methodology to larger datasets, diverse climatic conditions, and real-time monitoring for predictive maintenance. Additionally, integrating dynamic cleaning schedules and exploring the method's applicability in hybrid renewable energy systems will further enhance its contribution to sustainable energy generation.

CRediT authorship contribution statement

Safia Babikir Bashir: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Resources, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Mena Maurice Farag:** Writing – review & editing, Writing – original draft, Visualization, Validation, Methodology, Formal analysis, Data curation. **Abdul-Kadir Hamid:** Supervision, Project administration, Funding acquisition, Formal analysis, Conceptualization. **Ali A. Adam:** Writing – review & editing, Writing – original draft, Supervision, Formal analysis. **Ramesh C. Bansal:** Writing – review & editing, Supervision, Project administration, Funding acquisition. **Nsilulu T Mbungu:** Writing – review & editing, Software, Formal analysis. **A. Elnady:** Writing – review & editing, Project administration, Funding acquisition. **Ahmed G. Abo-Khalil:** Writing – review & editing, Project administration. **Mousa Hussein:** Writing – review & editing, Supervision, Funding acquisition.

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

The authors would like to acknowledge the technical and financial support provided by the United Arab Emirates University and the - University of Sharjah.

Data availability

Data will be made available on request.

References

[1] Hassan Q, Viktor P, J. Al-Musawi T, Mahmood Ali B, Algburi S, Alzoubi HM,, et al. The renewable energy role in the global energy Transformations. Renew Energy. Focus 2024;48:100545. https://doi.org/10.1016/j.ref.2024.100545.
[2] Karlilar Pata S, Balcilar M. Decarbonizing energy: Evaluating fossil fuel displacement by renewables in OECD countries. Environ Sci Pollut Res 2024;31:31304–13. https://doi.org/10.1007/s11356-024-33324-8.
[3] Hamid A-K, Farag MM, Hussein M. Enhancing photovoltaic system efficiency through a digital twin framework: A comprehensive modeling approach. Int J Thermofluids 2025;26:101078. https://doi.org/10.1016/j.ijft.2025.101078.
[4] Khurshid H, Mohammed BS, Al-Yacouby AM, Liew MS, Zawawi NAWA. Analysis of hybrid offshore renewable energy sources for power generation: A literature review of hybrid solar, wind, and waves energy systems. Dev Built Environ 2024;19:100497. https://doi.org/10.1016/j.dibe.2024.100497.
[5] Wang G, Zhang Z, Lin J. Multi-energy complementary power systems based on solar energy: A review. Renew Sustain Energy Rev 2024;199:114464. https://doi.org/10.1016/j.rser.2024.114464.
[6] Bashir SB, Farag MM, Hamid AK, Adam AA, Abo-Khalil AG, Bansal R. A novel hybrid CNN-XGBoost model for photovoltaic system power forecasting. In: 2024 6th Int. Youth Conf. Radio Electron. Electr. Power Eng., IEEE; 2024, p. 1–6. https://doi.org/10.1109/REEPE60449.2024.10479878.
[7] Kakran S, Rathore JS, Sidhu A, Kumar A. Solar energy advances and CO2 emissions: A comparative review of leading nations' path to sustainable future. J Clean Prod 2024;475:143598. https://doi.org/10.1016/j.jclepro.2024.143598.
[8] Bouraima MB, Ayyıldız E, Badi I, Özçelik G, Yeni FB, Pamucar D. An integrated intelligent decision support framework for the development of photovoltaic solar power. Eng Appl Artif Intell 2024;127:107253. https://doi.org/10.1016/j.engappai.2023.107253.
[9] Molla S, Farrok O, Alam MJ. Electrical energy and the environment: Prospects and upcoming challenges of the World's top leading countries. Renew Sustain Energy Rev 2024;191:114177. https://doi.org/10.1016/j.rser.2023.114177.
[10] Farag MM, Bansal RC. Solar energy development in the GCC region – a review on recent progress and opportunities. Int J Model Simul 2023;43:579–99. https://doi.org/10.1080/02286203.2022.2105785.
[11] Salameh T, Farag MM, Hamid A-K, Hussein M. Adaptive neuro-fuzzy inference system for accurate power forecasting for on-grid photovoltaic systems: A case study in Sharjah. UAE Energy Convers Manag X 2025;26:100958. https://doi.org/10.1016/j.ecmx.2025.100958.
[12] Farag MM, Patel N, Hamid A-K, Adam AA, Bansal RC, Bettayeb M, et al. An Optimized Fractional Nonlinear Synergic Controller for Maximum Power Point Tracking of Photovoltaic Array Under Abrupt Irradiance Change. IEEE J Photovoltaics 2023;13:305–14. https://doi.org/10.1109/JPHOTOV.2023.3236808.
[13] Elamim A, Sarikh S, Hartiti B, Benazzouz A, Elhamaoui S, Ghennioui A. Experimental studies of dust accumulation and its effects on the performance of solar PV systems in Mediterranean climate. Energy Rep 2024;11:2346–59. https://doi.org/10.1016/j.egyr.2024.01.078.
[14] Farag MM, Hamid A-K, AlMallahi MN, Elgendi M. Towards highly efficient solar photovoltaic thermal cooling by waste heat utilization: A review. Energy Convers Manag X 2024;23:100671. https://doi.org/10.1016/j.ecmx.2024.100671.
[15] Hassan G, Sami Yilbas B, Al-Sharafi A, Al-Sulaiman F, Abdulhamid AA. Dust mitigation strategies concerning solar energy applications: A comprehensive review. Sol Energy 2024;277:112728. https://doi.org/10.1016/j.solener.2024.112728.
[16] Mbungu NT, Bashir SB, Michael NE, Farag MM, Hamid A-K, Ismail AAA, et al. Predictive control technique for solar photovoltaic power forecasting. Energy Convers Manag X 2024;24:100768. https://doi.org/10.1016/j.ecmx.2024.100768.
[17] Wan L, Zhao L, Xu W, Guo F, Jiang X. Dust deposition on the photovoltaic panel: A comprehensive survey on mechanisms, effects, mathematical modeling, cleaning methods, and monitoring systems. Sol Energy 2024;268:112300. https://doi.org/10.1016/j.solener.2023.112300.
[18] Kayri İ, Bayar MT. A new approach to determine the long-term effect of efficiency losses due to different dust types accumulation on PV modules with artificial neural networks. J Clean Prod 2024;434:140282. https://doi.org/10.1016/j.jclepro.2023.140282.
[19] Al SA, Issa UH, Sayed ET, Rezk H, Abdelkareem MA, Miky Y, et al. Modeling and analysis of risk factors affecting operation of photovoltaic power plants. Ain Shams Eng J 2024;15:102812. https://doi.org/10.1016/j.asej.2024.102812.
[20] Ali Sadat S, Faraji J, Nazififard M, Ketabi A. The experimental analysis of dust deposition effect on solar photovoltaic panels in Iran's desert environment. Sustain Energy Technol Assessments 2021;47:101542. https://doi.org/10.1016/j.seta.2021.101542.
[21] Kazem HA, Chaichan MT, Al-Waeli AHA, Sopian K. Effect of dust and cleaning methods on mono and polycrystalline solar photovoltaic performance: An indoor experimental study. Sol Energy 2022;236:626–43. https://doi.org/10.1016/j.solener.2022.03.009.
[22] Jaszczur M, Teneta J, Styszko K, Hassan Q, Burzyńska P, Marcinek E, et al. The field experiments and model of the natural dust deposition effects on photovoltaic module efficiency. Environ Sci Pollut Res 2019;26:8402–17. https://doi.org/10.1007/s11356-018-1970-x.
[23] Al-Ghussain L, Taylan O, Abujubbeh M, Hassan MA. Optimizing the orientation of solar photovoltaic systems considering the effects of irradiation and cell temperature models with dust accumulation. Sol Energy 2023;249:67–80. https://doi.org/10.1016/j.solener.2022.11.029.
[24] Liu Y, Li H, Li L, Yin X, Wu X, Su Z, et al. Solar photovoltaic panel soiling accumulation and removal methods: A review. IET Renew Power Gener 2024. https://doi.org/10.1049/rpg2.12940.
[25] Hosseini F, Sheikholeslami M. Effects of self-cleaning technique and nanofluid cooling on performance of photovoltaic solar unit employing sinusoidal surfaces. Appl Therm Eng 2024;240:122223. https://doi.org/10.1016/j.applthermaleng.2023.122223.
[26] Ndeto MP, Njoka F, Wekesa DW, Kinyua R. Mechanisms and economics of a self-powered, automated, scalable solar PV surface cleaning system. Renew Energy 2024;226:120477. https://doi.org/10.1016/j.renene.2024.120477.
[27] Shao Y, Zhang C, Xing L, Sun H, Zhao Q, Zhang L. A new dust detection method for photovoltaic panel surface based on Pytorch and its economic benefit analysis. Energy AI 2024;16:100349. https://doi.org/10.1016/j.egyai.2024.100349.
[28] Cruz-Rojas T, Franco JA, Hernandez-Escobedo Q, Ruiz-Robles D, Juarez-Lopez JM. A novel comparison of image semantic segmentation techniques for detecting dust in photovoltaic panels using machine learning and deep learning. Renew Energy 2023;217:119126. https://doi.org/10.1016/j.renene.2023.119126.
[29] Halwani S, Farag MM, Hamid A-K, Hussein M. An automated and cost-efficient method for photovoltaic dust cleaning based on biaxially oriented polyamide coating material. 7th Int Conf Renew ENERGY Gener Appl, 2024:316–23. https://doi.org/10.21741/9781644903216-41.
[30] He B, Lu H, Zheng C, Wang Y. Characteristics and cleaning methods of dust deposition on solar photovoltaic modules-A review. Energy 2023;263:126083. https://doi.org/10.1016/j.energy.2022.126083.
[31] Malik H, Alsabban M, Qaisar SM. Arduino Based Automatic Solar Panel Dust Disposition Estimation and Cloud Based Reporting. Procedia Comput Sci 2021;194:102–13. https://doi.org/10.1016/j.procs.2021.10.063.
[32] Mohammed HA, Al- Hilli BAM, Al-Mejibli IS. Smart system for dust detecting and removing from solar cells. J Phys Conf Ser 2018;1032:012055. https://doi.org/10.1088/1742-6596/1032/1/012055.

[33] Hussain A, Batra A, Pachauri R. An experimental study on effect of dust on power loss in solar photovoltaic module. Renewables Wind Water, Sol 2017;4:9. https://doi.org/10.1186/s40807-017-0043-y.
[34] Perez-Anaya E, Elvira-Ortiz DA, Osornio-Rios RA, Antonino-Daviu JA. Methodology for the Identification of Dust Accumulation Levels in Photovoltaic Panels Based in Heuristic-Statistical Techniques. Electronics 2022;11:3503. https://doi.org/10.3390/electronics11213503.
[35] Abuqaaud KA, Ferrah A. A Novel Technique for Detecting and Monitoring Dust and Soil on Solar Photovoltaic Panel. Adv. Sci. Eng. Technol. Int. Conf. IEEE 2020; 2020:1–6. https://doi.org/10.1109/ASET48392.2020.9118377.
[36] Sriram A, D. Sudhakar T.. Photovoltaic Cell Panels Soiling Inspection Using Principal Component Thermal Image Processing. Comput Syst Sci Eng 2023;45: 2761–72. https://doi.org/10.32604/csse.2023.028559.
[37] Parenti M, Fossa M, Delucchi L. A model for energy predictions and diagnostics of large-scale photovoltaic systems based on electric data and thermal imaging of the PV fields. Renew Sustain Energy Rev 2024;206:114858. https://doi.org/10.1016/j.rser.2024.114858.
[38] Rico Espinosa A, Bressan M, Giraldo LF. Failure signature classification in solar photovoltaic plants using RGB images and convolutional neural networks. Renew Energy 2020;162:249–56. https://doi.org/10.1016/j.renene.2020.07.154.
[39] Unluturk M, Kulaksiz AA, Unluturk A. Image Processing-based Assessment of Dust Accumulation on Photovoltaic Modules. 1st Glob. Power, Energy Commun. Conf. IEEE 2019;2019:308–11. https://doi.org/10.1109/GPECOM.2019.8778578.
[40] Fan S, Wang Y, Cao S, Zhao B, Sun T, Liu P. A deep residual neural network identification method for uneven dust accumulation on photovoltaic (PV) panels. Energy 2022;239:122302. https://doi.org/10.1016/j.energy.2021.122302.
[41] Tan Y, Liao K, Bai X, Deng C, Zhao Z, Zhao B. Denoising Convolutional Neural Networks Based Dust Accumulation Status Evaluation of Photovoltaic Panel. IEEE Int. Conf. Energy Internet. IEEE 2019;2019:560–6. https://doi.org/10.1109/ICEI.2019.00105.
[42] Cao S, Ma X, Fan S, Wang T. IDS-Net: Integrated Network for Identifying Dust State of Photovoltaic Panels. Int. Conf. Electron. Inf. Eng. Comput. Sci. IEEE 2021;2021: 89–92. https://doi.org/10.1109/EIECS53707.2021.9588136.
[43] Et-taleby A, Chaibi Y, Allouhi A, Boussetta M, Benslimane M. A combined convolutional neural network model and support vector machine technique for fault detection and classification based on electroluminescence images of photovoltaic modules. Sustain Energy, Grids Networks 2022;32:100946. https://doi.org/10.1016/j.segan.2022.100946.
[44] Zhang T, Chen L, Ma H, Xue X. Power Generation Efficiency Prediction Methods of Photovoltaic Panel under Dust Based on Convolutional Neural Network. IEEE 5th Conf. Energy Internet Energy Syst. Integr. IEEE 2021;2021:3329–34. https://doi.org/10.1109/EI252483.2021.9713321.
[45] Onim MSH, Sakif ZMM, Ahnaf A, Kabir A, Azad AK, Oo AMT, et al. SolNet: A Convolutional Neural Network for Detecting Dust on Solar Panels. Energies 2022; 16:155. https://doi.org/10.3390/en16010155.
[46] Jeffrey Kuo C-F, Chen S-H, Huang C-Y. Automatic detection, classification and localization of defects in large photovoltaic plants using unmanned aerial vehicles (UAV) based infrared (IR) and RGB imaging. Energy Convers Manag 2023;276: 116495. https://doi.org/10.1016/j.enconman.2022.116495.
[47] Olorunfemi BO, Nwulu NI, Ogbolumani OA. Solar panel surface dirt detection and removal based on arduino color recognition. MethodsX 2023;10:101967. https://doi.org/10.1016/j.mex.2022.101967.
[48] Dwivedi D, Babu KVSM, Yemula PK, Chakraborty P, Pal M. Identification of surface defects on solar PV panels and wind turbine blades using attention based deep learning model. Eng Appl Artif Intell 2024;131:107836. https://doi.org/10.1016/j.engappai.2023.107836.