Water Utility Journal (2026) 35:5

Fault detection and predictive maintenance of solar panels using improved open-sourced image processing algorithms

Lorena Aguilar Abril, Raufdeen Rameezdeen, Jing Gao and Christopher W.K. Chow*

Sustainable Infrastructure and Resource Management (SIRM), UniSA STEM, Adelaide University, Mawson Lakes, SA 5095, Australia

* *e-mail: chris.chow@adelaide.edu.au*

Abstract: Solar PVs' (panels') performance and efficiency can be impacted by surface conditions. This study explores aerial imaging as a technique to foster predictive maintenance. An extensive review was conducted on open-sourced algorithms for processing thermal images using Matlab, comparing similarities and differences. The information was organized into stages: pre-processing, processing, segmentation, and validation. This research formed the baseline for testing four types of algorithms on 277 images, evaluating their repeatability and reproducibility using the Coefficient of Variation (CV). Algorithm 2 achieved a CV of 4.3%, utilizing four filters during processing and applying equalization at each stage. For segmentation, a thresholding technique was used, and the red and green channels were employed to binarize the image for area calculation. The area calculation, along with technical parameters such as voltage and current from the solar panels, were used to determine efficiency. This combined with the percentage of degradation, allowed for calculating the impact of hotspots on efficiency. The efficiency of the analysed solar panel ranged from 15% to 22%, with hotspots potentially reducing efficiency by 1.3% to 1.5% annually, thereby affecting the energy supply.

Key words: Asset condition assessment; solar panel fault detection; image analysis.

1. INTRODUCTION

Assets are the lifeblood of most organisations and are critical to their daily operations. Effective condition monitoring techniques should enable the early detection of asset faults and provide accurate diagnoses of potential failure risks. Hence, there is growing interest in developing advanced technologies to enhance condition monitoring and diagnostic capabilities. Solar panels, or solar photovoltaic (PV) systems are an important new asset class that water utilities must manage due to their increasing reliance on off-grid renewable energy sources. Asset managers in water utilities need to be equipped with modern tools and techniques to effectively monitor, maintain, and integrate these systems into their asset management portfolios. Developing in-house capabilities for fault detection and rectification can reduce operational expenditure (OpEx) while improving maintenance efficiency. The timely detection of solar module faults can extend the service life of PV systems and enhance their operational performance (Liao & Lu, 2020). The motivation for this study came from a project the authors undertook in collaboration with the local water utility that developed a solar farm on a reservoir that could be only accessed using Unmanned Aerial Vehicle (UAV). The utility was keen to explore how image processing could be used for fault detection and predictive maintenance of these solar panels.

Solar panel efficiency impacts energy production: a 21% efficient panel generates 50% more kWh than 14% efficiency under identical conditions (Dhanraj et al., 2021). Faults in panels, influenced mainly by temperature and humidity (Kirubakaran et al., 2022), lead to reduced efficiency, energy loss and increased maintenance cost (Balakrishnan et al., 2023). Currently, traditional solar panel monitoring is performed manually, which is labour-intensive, time-consuming, and prone to systematic errors (Liao & Lu, 2020). New techniques and AI capabilities are constantly sought, such as utilising UAV combined with image processing technique for condition monitoring. One of the techniques widely used without shutdown (Pathak & Patil, 2023) is thermal imaging, a non-invasive and non-destructive technique. In terms of failures to detect with this technique, findings have been reported for single and multiple hotspots (Constantin et al., 2023). This technique converts object appearances into thermal images, revealing temperature patterns. The bright colour shows hot portions, while dark, cold portions (Kirubakaran et al., 2022). UAVs equipped with these cameras navigate solar arrays and capture those images; UAVs have the benefit of performing defect detection in a short time by its large area coverage, lightweight, high flexibility and speed of operation (Liao & Lu, 2020). Organisations are embracing artificial intelligence (AI) as part of the condition monitoring process. AI algorithms such as artificial neural network (ANN), fuzzy logic system (FLS), genetic algorithms (GA) and support vector machine (SVM) etc., are well-developed image processing techniques. Matlab software is used to analyse thermal images obtained by UAVs. This is a powerful tool specialised for image processing due to its extensive library of functions and algorithms (Raikwar et al., 2023).

While past research has reported UAV-based image processing techniques and algorithms for solar panel fault detection, there is a lack of empirical evidence regarding their practical implementation. This research aims to address that gap. It seeks to analyse solar panel hotspots using aerial thermal images, with the goal of establishing a baseline for the water utility to enhance the operational efficiency of solar panels within their asset management system. Finally, the research evaluates the efficiency loss of solar panels affected by hotspots. It focuses on analysing hotspots faults in solar panels using aerial thermal images utilising open-sourced codes and Matlab. The outcomes allow for the establishment of a consolidated baseline for the water utility to analyse their own thermal images.

2. METHODOLOGY

The first part of the methodology section focuses on obtaining current knowledge of how solar panel monitoring and maintenance in different countries were conducted, including analysing thermographic images as part of asset management. It outlines the approach and methods used to gather information and prepare this overview. The second part of the methodology section focuses on the development of flowcharts used to develop algorithms aimed at detecting dark areas, which indicate high temperatures on the surface of solar panels. Additionally, it highlights the key aspects considered during the trial-and-error process of algorithm creation to achieve the desired outcomes.

2.1 A global overview of solar panel maintenance: A benchmarking exercise

A comprehensive questionnaire survey was administered among maintenance experts to evaluate how solar panel maintenance is conducted in various countries. The outcome from the survey was recorded in a form created using Microsoft Forms, focusing on key aspects such as demographic information, technical maintenance details, challenges and solutions, economic aspects and management, future improvements in solar panel maintenance, and specific questions involving smart image processing techniques with drones. The questionnaire was divided into six sections with 19 Likert-scale questions, designed to be completed within a short period, ensuring it would attract participants by being short, easy to understand, and respectful of their time.

After the questionnaire was finalized, a targeted outreach was conducted using email addresses gathered from scientific articles, LinkedIn, and various databases, focusing on individuals from academia, industry, and consulting. The survey was also posted in various solar groups on LinkedIn, and to the membership in the International Solar Energy Society (ISES). Partially completed responses, as well as those submitted by participants who did not meet the minimum experience requirements, were excluded from the analysis. Following the data screening process, a total of 33 valid and fully completed responses were retained for analysis and are reported in this study. The geographical distribution of respondents was diverse, with the largest proportion originating from Asia (48.5%), followed by Oceania (21.2%), Europe (18.2%), and Latin America

(12.1%). This distribution reflects a broad international representation of professionals responded to the questionnaire survey. Respondents' professional backgrounds had a range of sectors: approximately a quarter each represented the energy industry, the automation, and consulting and academia, indicating a balanced distribution across these key and relevant professional domains. The questionnaire comprised demographic questions and a series of Likert-scale items designed to assess current solar PV maintenance practices and perceptions. Responses to the Likert-scale items were analysed using descriptive statistics to benchmark prevailing solar PV maintenance methods across different countries. Demographic variables were examined to characterise the sample and to explore similarities and differences across geographical regions and professional sectors.

2.2 Analysis of thermographic images captured using UAVs

A dataset published by Alfaro-Mejía et al. (2019) presenting 277 thermographic aerial images was selected for this study. The thermal images were with adequate quality and resolution, taken under specific environmental conditions and following UAV recommendations for angle positioning when capturing images. It was fully characterised to identify snail trails and hotspot failures using various filtering and segmentation techniques. The completed dataset included environmental measurements such as temperature, wind speed, and irradiance at the time the images were taken. Table 1 provides the characterization data for these images. All 277 images comprising 174 left and 103 right sides were included in the analysis.

Table 1. Characterization data of thermographic images

Date	# images		Irradiance (W/m^2)	Temperature (°C)	Wind Speed (km/h)
	Left side	Right Side			
27/04/2018	13	n/a	412	n/a	n/a
28/04/2018	13	n/a	136	n/a	n/a
4/5/2018	29	n/a	150	n/a	n/a
20/12/2018	24	36	196	34	34
21/12/2018	26	22	122	23	23
16/01/2019	30	18	248	34	34
19/01/2019	39	27	227	40	40

Prior to building an improved algorithm using Matlab (MathWorks Image Processing Toolbox) for detecting solar panel hotspots, an extensive literature review was carried out to find similarities, functions, and capabilities, and to assess all factors necessary to create a generalized algorithm to foster better accuracy in the outcomes. Matlab is selected as the processing software because previous literature has demonstrated good results for detecting hotspots in solar panels (Alajmi et al., 2019; Kirubakaran et al., 2022; Liao & Lu, 2020; Pathak & Patil, 2023; Raikwar et al., 2023).

The approach included a qualitative analysis in the literature review, utilizing comparative tables organized by stages to highlight similarities and differences in the algorithm flowcharts during execution. Additionally, scientific data fundamentals and programming code were used to develop and execute the algorithm in Matlab. A flowchart of the methodology is shown in Figure 1.

The first step was to identify all articles related to algorithm using the flowcharts and extract important information by stages, focusing on understanding how they worked. This process led to the creation of four algorithm flowcharts that maintain the core stages and principles of the main functions but vary in the sequence of certain functions and the set values, such as threshold values, required for image analysis. Due to the necessity of testing the algorithm as it was being developed, one image was taken and tested at each step, including visualization, to observe the results and the effects of applied filters, manual thresholds, and thresholding per colour channel. Once all the functions of the algorithm had been tested, the next step was to integrate them into a loop. The user would specify the file containing all the images, and the algorithm would automatically run on each image, calculating and displaying the identified hotspot areas in a matrix.

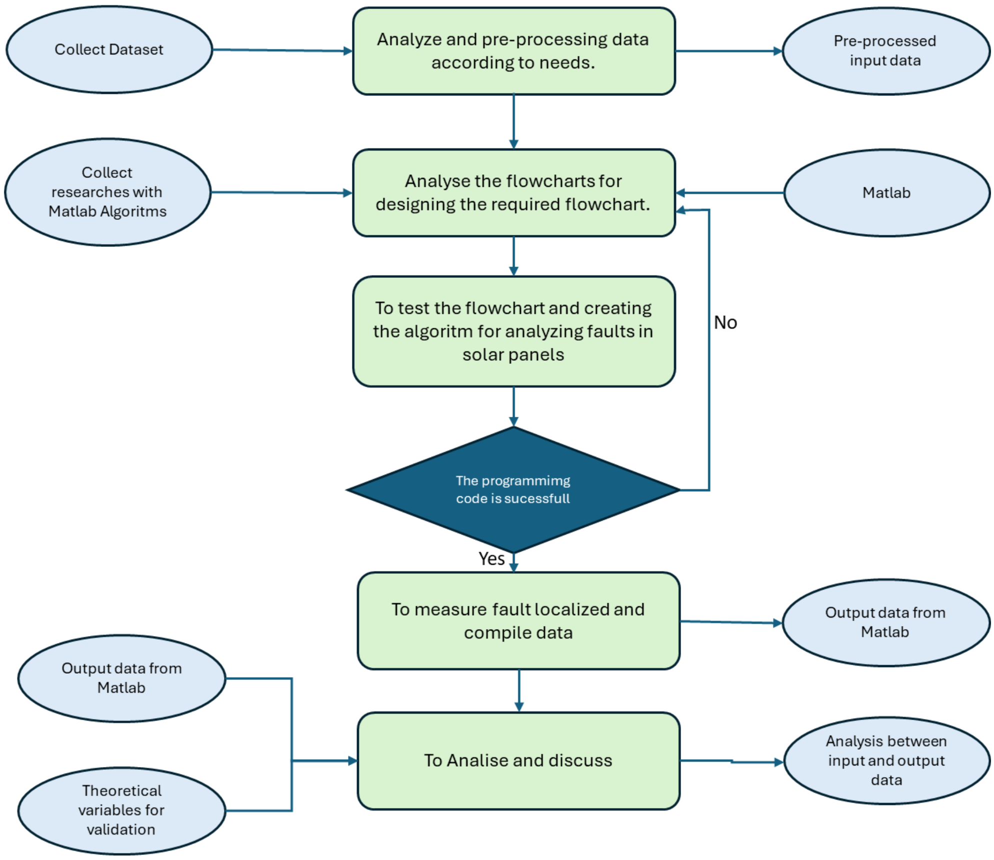


Figure 1. Proposed methodology for the development of the algorithm to detect hotspots in solar panels using Matlab

Algorithm 1 was developed following the flowchart shown in Figure 2. Subsequently, three additional algorithms were developed using a similar methodology, with their details provided in the Results and Discussion section. Overall, this study evaluated four algorithms.

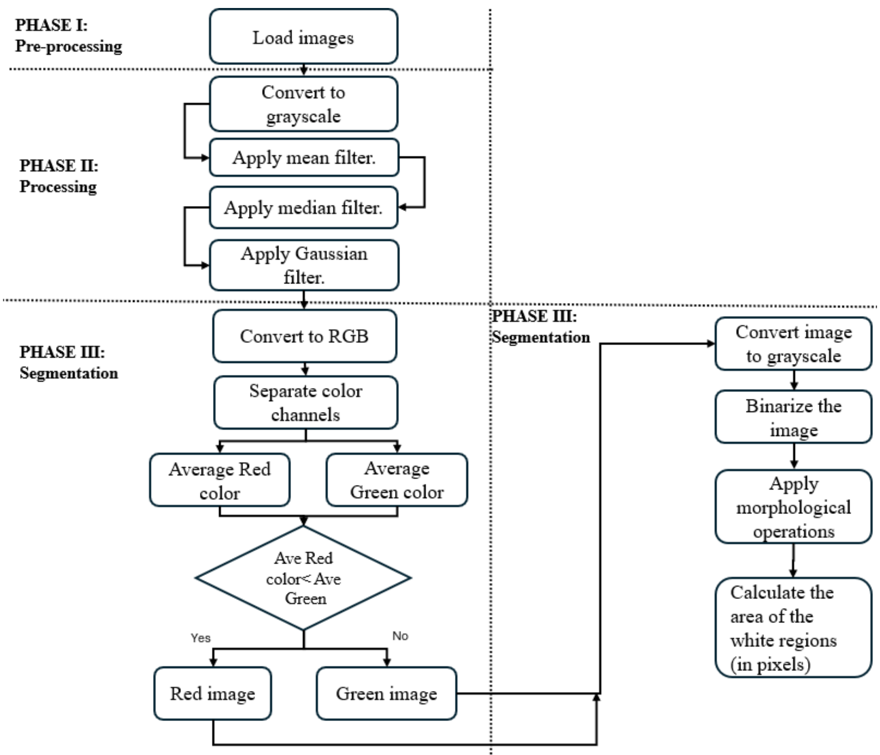


Figure 2. Flowchart for algorithm 1 in the calculation of hot spots using Matlab

The detection predictability of the four algorithms was compared and statistically evaluated using the coefficient of variation (CV), which served as an indirect indicator of algorithm reliability.

3. RESULTS AND DISCUSSION

3.1 A global overview of solar panel maintenance by experts

The questionnaire responses indicate that preventive maintenance is the predominant solar PV maintenance strategy, with 56% of respondents reporting that it is the most frequently adopted approach. In comparison, 25% indicated that they primarily rely on corrective maintenance, while only 9% reported using predictive maintenance. These findings suggest that current maintenance practices remain largely dependent on preventive and corrective approaches, with comparatively limited adoption of predictive maintenance strategies. However, predictive maintenance is becoming a more important key component for enhancing both short- and long-term power supply. This approach would need to be implemented promptly to reduce maintenance costs. Although the frequency of solar PV maintenance is ultimately determined by organisational policies and operational requirements, the survey responses indicate a preference for relatively frequent maintenance intervals. Specifically, 37% of respondents recommended conducting maintenance every three months, while 30% favoured monthly inspections. In comparison, 15% recommended maintenance every six months, and 18% suggested annual maintenance to be sufficient. Overall, the survey indicates that the majority of respondents support maintenance schedules of three months or less, reflecting a preference for more proactive maintenance practices. When asked about what diagnostic technique is used in their organisation to identify faults in solar panels, measurement of current and voltage (36%) followed by visual inspections (25%). Infrared thermography is not popular with only 11% usage reported.

The most significant challenge for organisations in maintaining solar panels is ensuring that maintenance is appropriately prioritised within their asset management plans. When respondents were asked about the challenges their organisations face in maintaining solar panel systems, two issues emerged as equally important: the inability to prioritise maintenance activities and the lack of staff education and training, with each accounting for 25% of responses. Resource allocation dependencies shared spare parts, and budget constraints in large organizations make this task difficult. Additionally, the lack of integrated systems results in the loss of valuable and key data for decision-making added difficulties. For organization preparing themselves to apply predictive maintenance, they need to adopt an integrative knowledge plan, with training and mentoring as a top priority (28% of responses), that includes data analytics and analysis.

Respondents perceived the payback from regular solar panel maintenance to be substantial, with approximately 72% rating it as either favourable or highly favourable. According to the respondents, the most representative indicators for assessing solar panel maintenance performance are the efficiency of the solar panel itself (44%), downtime and inactivity (28%), rather than maintenance costs (25%). More than 50% of the respondents indicated that their organisations struggle planning solar panel maintenance due to lack of integration in their information management system. The respondents believe that the future of solar panel maintenance would require the prompt integration of technology, such as artificial intelligence (50% of responses), to detect and diagnose faults in PV systems. If a company were to incorporate thermographic imaging techniques for fault detection in solar panels, the main challenge would be understanding how this technique can help measure lifespan and improve performance assessment. 85% of respondents have not implemented drone inspections using thermographic image analysis for solar panels. Companies that have adopted thermographic imaging report that the technique is fast, reliable, cost-effective, and allows for quicker identification of panel locations and faults.

After an extensive review of fault classification in solar panels, a comparative analysis was conducted, resulting in a generalized classification outlined in Faults Classification in PV Systems (Figure 3).

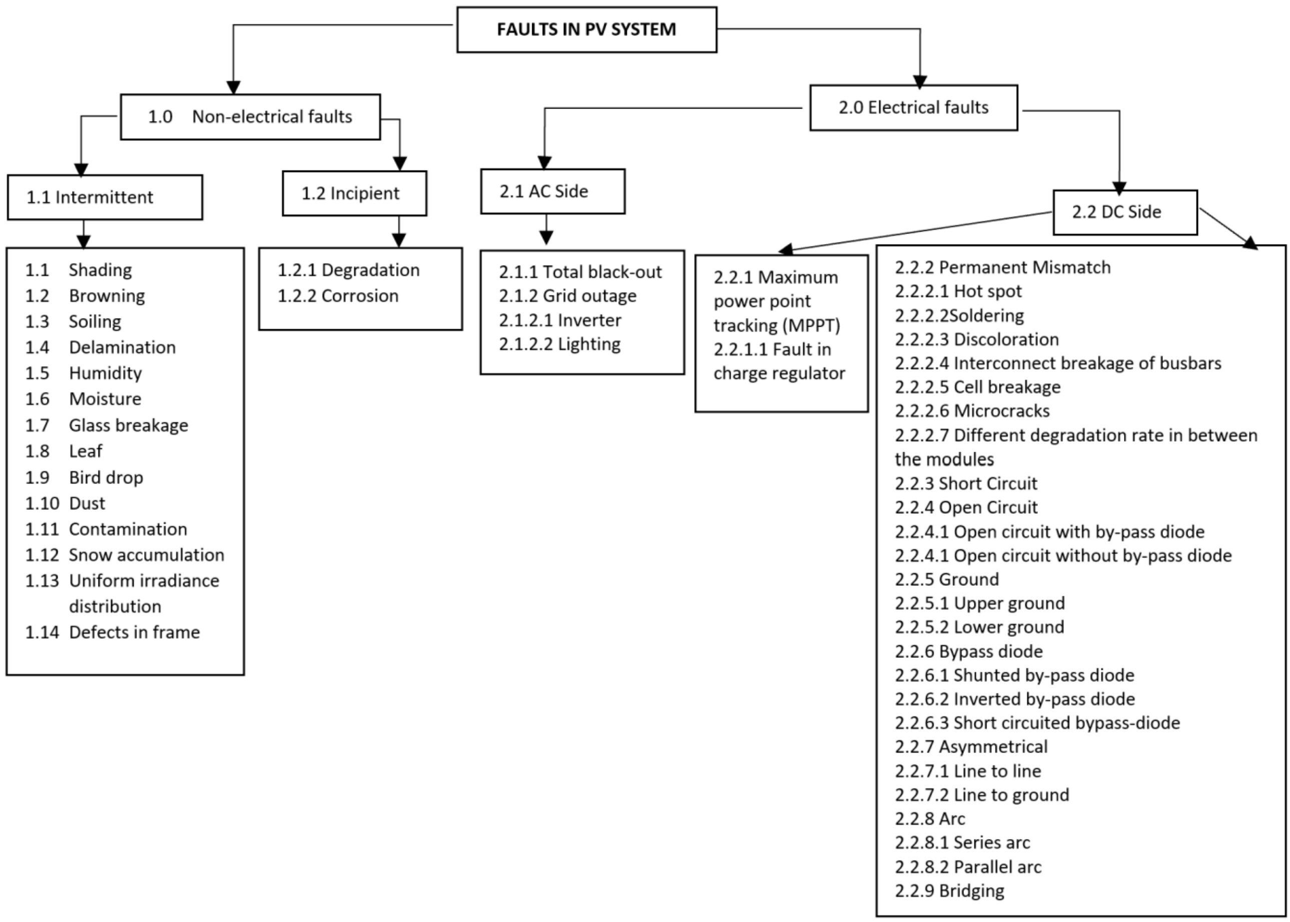


Figure 3. Faults classification in PV systems

A total of 41 potential faults were identified. These were divided into two main categories: non-electrical and electrical faults. Regarding non-electrical faults, initial visual inspections are crucial for detecting visible issues such as yellowing, delamination, and cracks (Karim, 2015). For electrical faults, microscopic techniques like scanning electron microscopy (SEM), attenuated total reflectance infrared (ATIR), and X-ray microtomography are used to analyse DC-side problems. (Dhanraj et al., 2021; Madeti & Singh, 2017).

In the survey, the most applied technique for fault analysis was found to be current and voltage data measurement, with 31% of respondents indicating this as their preferred method. This was followed by visual inspections at 22% and performance sensors at 19%. Less frequently used techniques included drones with cameras (6%) and infrared thermography (9%). These results align with a focus on preventive maintenance for assessing current performance. However, the preference for current and voltage measurement is a positive indicator for the future shift toward predictive maintenance. Predictive maintenance involves monitoring systems processed through software employing machine learning algorithms, with fuzzy logic, neural networks, and deep learning being the most recommended techniques (Rahmoune et al., 2023; Dhanraj et al., 2021; Han et al., 2021). In addition to this, the predictive maintenance can reduce the maintenance cost and reducing downtime, but requires a significant investment in data collection and analysis tools (Orosz et al., 2024).

Additionally, the use of drones with cameras depends on the location of the solar panels. It is reasonable to assume that among the respondents, their solar panels are in areas where drone usage is not necessary. Consequently, the application of thermography, which is cost-effective and fast for inspection (Hernández-Callejo et al., 2019).

3.2 Analysis of thermographic images captured with UAVs

Numan et al. (2021) reported that hotspots are location overheating and represents one of the main components for the reliability and safety of C-Si cells. It relies on a risk for the photovoltaic module lifetime and a decrease in its operational efficiency. If the shading is not removed, the temperature of the PV cells reach critical level causing the PV cells break down in seconds (Alajmi et al., 2019). Hotspots cannot be easily identified by naked-eye, however, processing of thermal images using image processing tools with Matlab (Kirubakaran et al., 2022) is highly recommended.

Thermography through thermal images develop a thermo-gram which is a variation of colour patterns in terms of the temperature, where the dark colour represent a colder area while bright colour a hot area (Chaudhary & Chaturvedi, 2018). Liao & Lu (2020) identified gray images in a thermo-gram in an interval range between 0 (means black) and 255 (means white). The image in Figure 4 shows the thermos-gram of a sample of left solar panel, where this variation is seen through the colour bar in the right hand of the image.

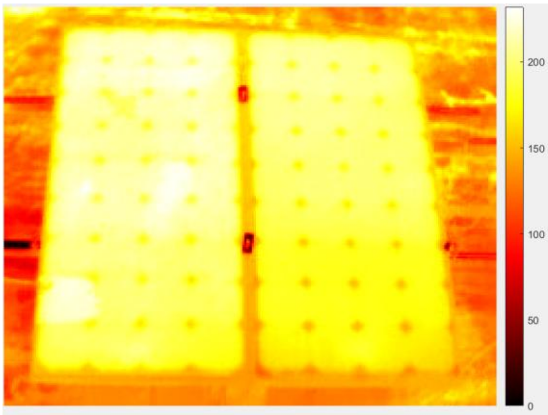


Figure 4. Thermo-gram of left solar panel in 2019-Jan-19 DJI_0373

As seen in Figure 4, the image shows colour patterns where the lighter areas are more prone to hotspots. However, this observation is purely qualitative, with shading spots above the 200-colour level. To accurately identify these shaded areas as affected, it is necessary to quantify the area using different algorithms designed for this purpose.

The process of creating an algorithm, a comparison table of six past studies that used Matlab to identify hotspots was developed. It was found that the comparison would be more useful if the algorithms had common elements. Therefore, five key stages of the algorithms were identified: pre-processing, processing, segmentation, and validation. Table 2 presents a comparison of these stages in detecting hotspots in solar panels.

Once these stages were identified, the functions and important parameters used in each algorithm were extracted. This allowed for the development of new algorithms by testing combinations of these functions and parameters, as well as incorporating newly created parameters. In the pre-processing stage, the aim is to improve image data by enhancing image features and eliminating unwanted distortions (Pathak & Patil, 2023). Image enhancement focuses on making images more readable and easier to interpret (Ujjwal et al., 2022). Across all stages, a common approach was to convert the images to grayscale before proceeding with any other steps, as enhancing RGB images is not very effective (Alajmi et al., 2019). Grayscale enhance the colour difference so that defect location can be observed (Liao & Lu, 2020).

In the processing stage, the goal is to remove noise by applying filters (Alajmi et al., 2019). The most common filters are mean, median, Gaussian, and bilateral. According to Pathak & Patil (2023), the mean filter is used for smoothing. It reduces noise in images by decreasing the intensity variation between a pixel and its neighbours, with larger kernels resulting in more aggressive smoothing. A 7x7 kernel provides more accurate fault detection. The median filter is superior to the mean filter for various segmentation techniques, but it is computationally expensive and slow to process. The Gaussian filter is used to reduce image noise by blurring the image. The bilateral filter, a nonlinear filter, is effective for reducing additive noise while preserving edges in images.

In the segmentation stage, the two most common methods are thresholding and K-means clustering (Ujjwal et al., 2022). For thresholding, the objective is to binarize the images. This can be done manually or by using histograms to determine the appropriate pixel value for the cutoff. On the other hand, K-means clustering uses machine learning to efficiently compute image segments. After this method is applied, morphological dilation of image, whose purpose is to remove unnecessary parts, was carried out (Kirubakaran et al., 2022).

In the verification stage, there are no details, but it was found that the area calculated indicating the hot spots to correlate it with the impact on solar panel efficiency (Raikwar et al., 2023). According to Pathak & Patil (2023) the Intersect of Union (IoU) is a variable to overlap two images: one labelled by experts and the other one with the results obtained in Matlab to the respective deviations.

As Matlab considers the entire image's pixels, it is important that areas outside the object of interest be removed, otherwise those areas could affect the Matlab results. Therefore, it was necessary to manually adjust the image to focus solely on the solar panel, removing the extended area outside of it. This operation had to be performed on all 277 images, the study objects used. This step is necessary because the area out of the object's interest is noise, as highlighted by Ujjwal (2022), and the optimisation of the image must be done for analysis.

Based on the theoretical information and testing with a set of images, Algorithm 1 was developed (Figure 2). In general, the algorithm consists of the four previously mentioned stages, with the processing phase involving the application of three types of filters: mean, median, and Gaussian. During the segmentation stage, the image was converted to red, green, and blue (RGB), and the colours were separated. The average pixel values of these colours were compared to identify either a red or green image. The identified image was then binarized, morphological operations were applied, and the respective area was calculated. In this algorithm the threshold was set manually as given in Table 3.

Table 2. Comparison table between different stages for detecting hot spots in solar panels

Stage	**Using Matlab real-time image analysis for solar panel fault detection with UAV (Liao & Lu, 2020)**	**Solar hotspot analysis using image processing in Matlab (Raikwar et al., 2023)**	**Processing of thermal satellite Images using Matlab (Ujjwal et al., 2022)**	**IR Thermal Image Analysis (Alajmi et al., 2019)**	**Infrared thermal images of solar PV panels for fault identification using image processing technique (Kirubakaran et al., 2022)**	**Evaluation of effect of pre-processing techniques in solar panel fault detection (Pathak & Patil, 2023)**
Pre-processing	To convert the image from RGB to grayscale.	The image must first be changed from RGB to grayscale	Image rectification and restoration: -Radiometric pre-processing. -Geometric pre-processing. -Flat field correction.	Thermal image to convert it into gray scale. It crops the image to one selected panel	Thermal image to convert it into gray scale.	Thermal image to convert it into gray scale.
Processing	To convert the image to Binary image The mean filtering. The median filtering. The Gaussian filtering *Note: The filtering are*	Gaussian filters. Binary image Image enhancement: Removing edges detection	*Denoising:* (Gaussian smoothing function). -Median filtering. -Wiener filtering *Normalisation Equalisation:* Histogram equalisation contrast limited adaptive	Edge-preserving Smoothing filter RGB colour space	Intensity adjusted image Morphological dilated image. Cell structure remove image. Mask for optical disk Edges of image Circular border Optical removed filter image	Filtering: Various filters such as Mean: kernel of size 7×7 is used which resulted in more accurate fault detection. Median: kernel of size 7×7 is used. Bilateral: kernel of

	applied to the original binary image not subsequent.		histogram equalisation (CLAHE). *Image enhancement:* Edge Detection: Sobel filter, Prewitt filter, Canny filter, Laplacian method *False colouring Thermal colour palettes*			size 3×3, 5×5, window size of [5 5] and σ =2.5 and is used Gaussian: Typical value of σ =2.5 and kernel of size 7×7 Image Enhancement: histogram equalization *Note: It seems histogram equalization is applied to each filtering separately.*
Segmentation	Convert 3D image Convert RGB Green: Cold Area (Normal) Red: Hot Area (Maintenance)	Used Pixel region tool on Normal image using "Used Pixel region tool on Normal image". Fill holes and remove unwanted regions	Thresholding (Otsu method). K-means (Machine learning) Multiclass classification (Machine learning)	- Convert the colour space to (HSV) space. -Threshold mask for three channels (HSV): $0.15 \leq H \leq 0.591$ $0.013 \leq S \leq 0.62$ $0.911 \leq V \leq 1.00$ -Slide the mask on the HSV image (Binary image). Value of 1: Hotspot fault. Value of 0: Not a hotspot.	Border remove image Fault cell detection. Performance analysis	- Histogram-based colour thresholding. After each filter separately, images are converted from RGB plane to HSV plane. $0.095 \leq H \leq 0.422$ $0.000 \leq S \leq 0.455$ $0.867 \leq V \leq 1.000$ Apply morphological opening on threshold image: specific structuring element (SE) of size 9× 9 - Separate colour channel-based thresholding Th: (Tot # Pixels/ #Pixels in new image) x 2.354
Validation	No information	Area calculation	No information	No information	Correlation quality (CQ) Correlation coefficient (CC) Normalised cross-correlation (NCC). IF PSNR SNR MSE NMSE	Intersect of Union (IoU) calculation Fault detection

Table 3. Manual umbral for left panel side applying algorithm 1

Sample date	Manual umbral
2018 April 27	0.71
2018 April 28	0.75
2018 May 4	0.76
2018 December 20	0.90
2018 December 21	0.93
2019 January 16	0.90
2019 January 19	0.80

For the second algorithm, after converting the image to grayscale, four filters are applied: mean, median, Gaussian, and bilateral. Then, equalization is applied to each filter separately. The results are converted into three channels, transformed to RGB, and then to HSV. The colour channels are separated, and a threshold is defined for each channel to segment the original image. By calculating the average of the red and green images, the appropriate one is selected to be converted to grayscale, binarized, and processed with morphological operations. Finally, the area of the white regions is calculated. The flowchart is shown in the following figure (Figure 5).

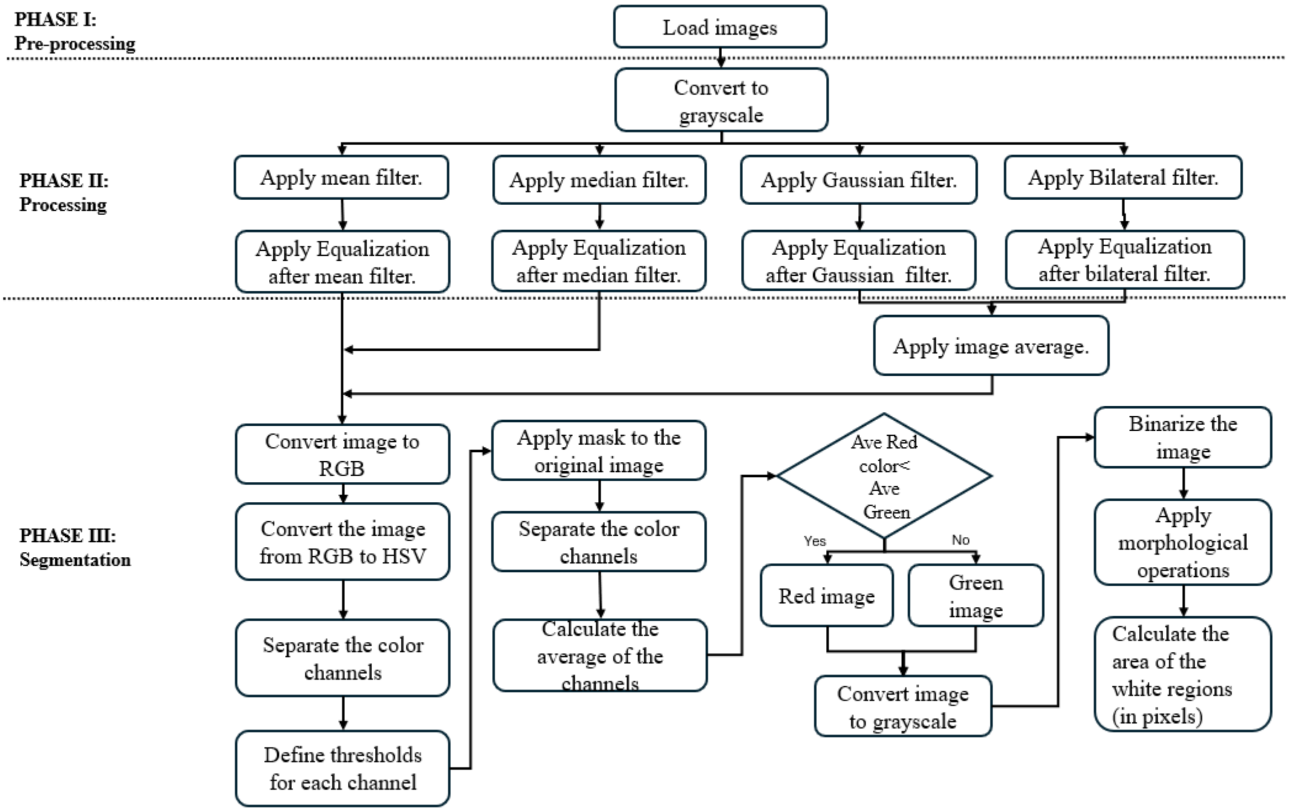


Figure 5. Flowchart for algorithm 2 in the calculation of hot spots using Matlab

For the third and fourth algorithms, the flowchart was almost identical (Figure 6), with the main difference being that the filters were applied sequentially before equalization was performed on the final filter. The key distinction between the two algorithms was the application of different low- and high-level thresholds in the HSV channel (hue, saturation, and value).

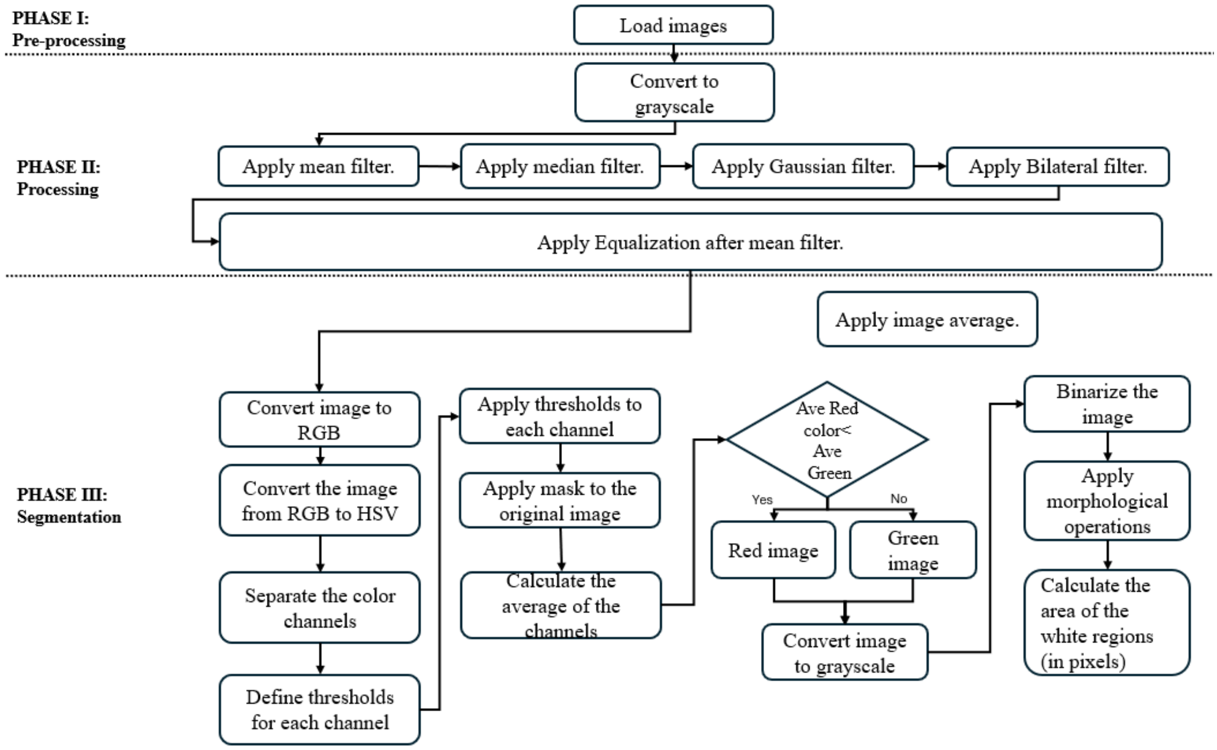
Each of the four algorithms was applied to the entire dataset, consisting of 277 images. However, for practicality, only the first image of each set is shown, with the results presented in the following figures and tables. Once the algorithms are executed, Matlab gives the results in terms of pixels; in addition to this, the sizing is different per image due to those images were adjusted manually as was previously mentioned; therefore, the following equations were established for the area calculation of each image.

Alfaro-Mejía et al. (2019) indicates that per 36 cells the sizing is 1186mm x 551mm x 35mm, so, due to there is in total 72 cells, the sizing is multiplied by 2.

$$Area\ solar\ panel\ (field) = \left(1186mm\ x\ \frac{0.1cm}{1mm}\ x\ 551mm\ x \frac{0.1cm}{1mm}\right)\ x2 = 13069.72cm^2 \tag{1}$$

Matlab is going to provide pixels of the white area and the pixels of the whole image, so:

$$Area\ calculated\ (cm^2) per\ image = \frac{Pixels\ of\ white\ area}{Pixels\ of\ the\ total\ image} \times Area\ solar\ panel\ (field) \tag{2}$$



Note: The difference between algorithms 3 and 4 will depend on the threshold value for each channel

Figure 6. Flowchart for algorithm 3 and 4 in the calculation of hot spots using Matlab

In hotspot detection, the region of interest is identified and assigned unique labels through region-labelling techniques (Raikwar et al., 2023). Evaluation of hotspot detection algorithms is normally performed by comparing the detected regions with expert-annotated ground-truth data using metrics such as Intersection over Union (IoU). However, the dataset used in this study did not include expert-labelled hotspot annotations, making the application of IoU or other ground-truth-based accuracy metrics infeasible. Therefore, this study employed the CV as a comparative measure to evaluate the consistency and robustness of algorithm performance across the dataset. A lower CV indicates that the algorithm produces more stable and less variable results under varying imaging conditions, thereby providing an indirect assessment of its reliability. While CV does not quantify detection accuracy or localisation performance, it enables a relative comparison of algorithm stability when objective ground-truth annotations are unavailable. This is a limitation of this study and should be considered when interpreting the results.

$$\mu = \frac{\sum_{i=1}^{N} x_i}{N} \quad (3)$$

$$\sigma^2 = \frac{\sum_{i=1}^{N} (x_i - \mu)^2}{N} \quad (4)$$

$$\sigma = \sqrt{\frac{\sum_{i=1}^{N} (x_i - \mu)^2}{N}} \quad (5)$$

$$CV = \frac{\sigma}{\mu} \quad (6)$$

After analysing the CV, the dataset with the least variability was observed on 20 December 2018, while the greatest variability occurred on 19 January 2019 (Table 4). In terms of algorithm performance, there was only a small difference between Algorithm 2 and Algorithm 4, with Algorithm 2 showing the lowest average CV (4.3%). This suggests that, among the algorithms evaluated, Algorithm 2 produced the most stable and consistent results for these image samples. Accordingly, Algorithm 2 was selected for the subsequent analyses. It should be noted that this selection is based on the observed stability of the results rather than implying that Algorithm 2 is inherently the most accurate or universally suitable algorithm. Furthermore, as each image was considered unique and analysed individually, the same algorithm was applied consistently across all date groups to ensure a uniform basis for comparison, following the recommendation of Kirubakaran et al. (2022).

Table 4. Coefficient of variation (CV) per different types of algorithms for left panel

Date	CV Algorithm 1	CV Algorithm 2	CV Algorithm 3	CV Algorithm 4	Average between CV and dates
2018 Apr 27	12.9%	7.2%	4.5%	*5.4%*	7.5%
2018 Apr 28	30.9%	*3.4%*	9.6%	4.5%	12.1%
2018 May 4	22.0%	*3.4%*	7.1%	5.0%	9.4%
2018 Dec 20	9.1%	1.9%	6.5%	*1.6%*	4.8%
2019 Dec 21	15.8%	*2.1%*	6.1%	3.5%	6.9%
2019 Jan 16	35.4%	6.9%	5.8%	8.7%	14.2%
2019 Jan 19	47.3%	5.4%	8.6%	5.4%	16.7%
Average between CV per algorithms.	24.8%	*4.3%*	6.9%	4.9%	

Firstly, it is notable that Algorithm 1 presents a high CV of around 25%. The manual threshold varied significantly across datasets, ranging from 0.7 to 0.9. The following images in Figure 7 show how the area of interest changes with different thresholds, clearly illustrating the impact of threshold selection. Therefore, having a reference image labelled by experts to identify hotspots is beneficial when results can be highly affected by manual categorization. This algorithm is less suitable for this type of detection under the conditions examined in this study.

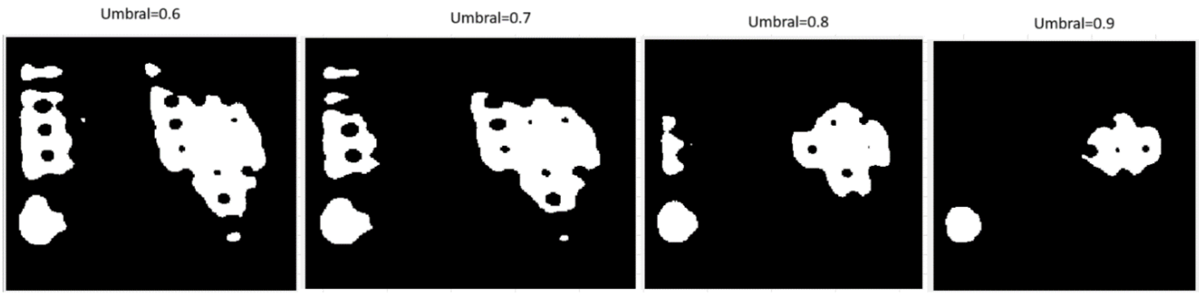


Figure 7. Interest area varying umbral (setpoint) manually

As seen from Figures S1 to S4 / Table S1 to S4 (Appendix), the white areas represent the hotspots. Despite all images showing the same panel on different dates, the white areas varied in both size and location. As explained earlier, Matlab can process images by analysing contrast, brightness, and pixels, but the results depend heavily on the input data. As demonstrated in these figures, the input data lacks consistency, even though the same solar panel was imaged on different dates. Some images were taken closer, others farther away, and the capture angles and brightness levels differed significantly. While these parameters are critical to this research, they were constrained by the characteristics of the available dataset. As the images had already been acquired, standardising the image capture conditions across all dates was beyond the scope of this study.

The key distinction of Algorithm 2 was the application of an additional filter, specifically a bilateral filter combined with histogram equalization (HE), to each of the existing filters. This supports the findings of Pathak & Patil (2023), who confirmed that using a bilateral filter alongside HE improves accuracy in fault detection during the pre-processing stage, resulting in better overall outcomes. The authors also explained that HE enhances the quality of low-contrast images by increasing the intensity difference between objects and the background, leading to a more uniform distribution of pixel intensity values and achieving the highest possible contrast. For the algorithm 2, 3, 4, a conversion from RGB to HSV was made, it was meanly because Alajmi et al. (2019) confirmed that is difficult to detect hotspot in RGB colour because the difficulty of separate colours and enable the threshold mask to detect the hot spot. For this research, the threshold mask values for HSV channels were stablished by Alajmi et al. (2019) and Pathak & Patil (2023).

The research found that in the segmentation process, most thermal images used for fault detection show relevant information in the green and blue channels, while the red channel does not provide useful data. In this case, the results indicate that the red and green channels are used for thresholding, while the blue channel does not offer valuable information. Pathak & Patil (2023). When the HSV image is converted to binary image, pixel value equals 1 is a hot-spot fault whereas value equals to 0 is not a hot-spot (Alajmi et al., 2019).

3.3 Impact of solar panel efficiency by hotspots

In evaluating the deposition effect of tree shading, the current and voltage output were measured under different environmental conditions. The current output as well as the output power of solar PV reduces when solar panels come under shading, dust, depositions according to the atmospheric conditions, those depositions increase the temperature of solar panel surface representing hotspots. (Chaudhary & Chaturvedi, 2018). Due to the unavailability of field data for those values, the output generated by Matlab, specifically in terms of hotspot areas, was to analyse the impact on solar panel efficiency.

Solar panels efficiency is an important metric to measure the performance and feasibility of solar panels. It is used for assessing a solar panel´s proficiency in converting sunlight into electrical energy. The equation to calculate this is (Solar Control, n.d.):

$$Efficiency(\%) = \frac{Output\ power\ (Watts)}{Input\ Power\ (Sunlight, in\ Watts\ per\ square\ \ meter)} \times Area(m^2) \times 100 \tag{7}$$

Due to the lack of voltage and current measurements needed in the field to assess the impact of hotspots on solar panels, the study hypothetically calculated the impact on efficiency using theoretical data from the dataset and the affected area calculated in Matlab.

The dataset of this research, the values of voltage (21.78 V) and current (5.13 A), is extracted from Alfaro-Mejía et al. (2019), in terms of calculating the range of efficiency of this solar panel:

$$Output\ power\ (Watts) = Voltage\ (V) \times Current(A) \tag{8}$$

$$Output\ power\ (Watts) = 21.78 \times 5.13 = 111.73W \tag{9}$$

$$Input\ Power = \frac{500W}{m^2} / \frac{1000W}{m^2} \tag{10}$$

$$Area = 1186mm \times \frac{1cm}{10mm} \times 551mm \times \frac{1cm}{10mm} \times 2 \times \frac{1m^2}{100cm^2} = 1.307m^2 \tag{11}$$

$$Efficiency(\%) = \frac{111.73W}{500\frac{W}{m^2}} \times 1.307(m^2) \times 100 = 17.10\% \tag{12}$$

$$Efficiency(\%) = \frac{111.73W}{1000\frac{W}{m^2}} \times 1.307(m^2) \times 100 = 8.55\% \tag{13}$$

According to those results the range of efficiency is between 9% and 17%; however, a review of solar panels on the market revealed that the efficiency ranged between 15% and 22%. Therefore, having this data as market reference will be taken for next calculations the irradiance of 500 W/m^2.

Although the solar panels of this research are in Colombia (Table 5), it was found solar panels in Australia had efficiencies between 12% and 17%.

Table 5. Solar system´s efficiency reported in Imteaz & Ahsan (2018), table has been modified with just the key information related to this paper

Manufacturer	Efficiency (%)
1	17.2
2	16.1
3	15.9
4	15.8
5	15.7
6	15.7
7	15.7
8	14.7
9	14.2
10	14.2
11	13.7
12	12.3

The percentage of degradation of solar cell due to hotspot is calculated using the following formula (Chaudhary & Chaturvedi, 2023):

$$Percentage\ average\ degradation = \frac{Area_{White\ area}}{Area_{Total}} \quad (14)$$

In that sense the percentage of degradation for left panel with 4 algorithms is calculated as follows:

Table 6. Percentage of degradation for left panel

Date	Percentage of degradation Algorithm 1	Percentage of degradation Algorithm 2	Percentage of degradation Algorithm 3	Percentage of degradation Algorithm 4
2018 Apr 27	2.4%	14.8%	9.0%	*14.4%*
2018 Apr 28	5.2%	*15.7%*	9.7%	15.0%
2018 May 4	9.1%	*15.5%*	9.6%	14.7%
2018 Dec 20	14.2%	15.0%	9.6%	*14.1%*
2019 Dec 21	4.7%	*15.5%*	9.4%	14.1%
2019 Jan 16	2.6%	16.9%	9.3%	15.7%
2019 Jan 19	*5.0%*	*17.2%*	*9.5%*	*16.4%*
STD	4.2%	0.9%	0.2%	0.9%

From the obtained results in Table 6, it is evident that Algorithms 1 and 3 have the lowest percentage of degradation. However, as previously explained in Table 4, the average CV per algorithm is more favourable for Algorithms 2 and 4. Therefore, it is not sufficient to check the percentage of degradation on a single date; it is important to analyse the trend over a period.

In that sense, the impact of hotspots will be analysed only for the Algorithm 2, which was the proper algorithm selected for this analysis.

The impact on efficiency by hotspots is calculated with the following formula:

$$\Delta e = e_0 \times \%\ degradation\ \times C \quad (15)$$

where,

$$\Delta e = loss\ of\ efficiency$$

$C = Severity\ coefficient$

The severity coefficient ranges from 0 to 1. As no established reference values were identified in the literature for this coefficient, a midpoint value of 0.5 was assumed for the purposes of this study. This assumption represents a hypothetical scenario and was adopted to provide a representative estimate of efficiency loss. The resulting efficiency loss estimates should therefore be interpreted within the context of this assumption.

Table 7. Loss of efficiency by Hotspots, Efficiency: 17.10% and Severity Coefficient: 0.5

Date	Delta	%	Real efficiency
2018 Apr 27	0.01265	1.3%	15.8%
2018 Apr 28	0.01340	1.3%	15.8%
2018 May 4	0.01325	1.3%	15.8%
2018 Dec 20	0.01284	1.3%	15.8%
2019 Dec 21	0.01323	1.3%	15.8%
2019 Jan 16	0.01444	1.4%	15.7%
2019 Jan 19	0.01466	1.5%	15.6%

Based on the assumptions adopted in this study, the identified hotspots are estimated to have reduced the initial efficiency of the solar panels by approximately 1.3% to 1.5% over one-year period (Table 7). These values represent an estimated impact derived from the proposed methodology rather than direct field measurements of current, voltage, or power output.. These values are consistent with the findings of Numan et al. (2021), where a single effect in solar panels reduces the power output by 1.5% per affected substring. Another research reported substantially greater power losses under more severe operating conditions when the hottest cell is shaded by 25%, 50%, 75% and 100% of the shaded area, the power losses were 37.2%, 50.1%, 48.6% and 52.9%, respectively (Skomedal et al., 2020).

The results of this study demonstrate that hotspot detection is not perfect, as the performance of the algorithms varied depending on image characteristics and processing parameters. From a practical perspective, the proposed image-processing approach could serve as a decision-support tool within a broader asset condition monitoring and maintenance framework of water utilities. For example, utilities could use the algorithm to screen large numbers of solar panels and identify those that require further inspection, enabling maintenance resources to be prioritised more efficiently. However, before such a method can be relied upon for operational maintenance decisions, further validation against expert-labelled reference data and direct electrical performance measurements (e.g., current, voltage, and power output) is required to establish its accuracy and reliability under a range of operating conditions.

Several limitations of this study should also be acknowledged. First, the analysis was conducted using a single dataset, which limits the generalisability of the findings. Second, no expert-labelled ground truth was available to validate the detected hotspots. Third, manual threshold adjustment was required for some algorithms, introducing a degree of subjectivity into the analysis. Fourth, variations in image acquisition conditions, including camera distance, viewing angle, and lighting, affected the consistency of the input data. Finally, the efficiency loss estimates were based on assumed parameter values rather than direct field measurements and should therefore be interpreted as indicative estimates rather than measured performance losses.

4. CONCLUSIONS

In South Australia, one state water utility alone integrates 360,000 solar panels generating 242 GWh annually and meeting 70% of its energy needs, primarily in wastewater pumping and treatment operations. To gain an overview of solar panel maintenance, assess efficiency, identify failures that lead to losses, an extensive literature review has been conducted to evaluate the current state and future perspectives of solar panel maintenance. Followed by the evaluation and information collection, aerial thermal imaging as a technique has been explored to foster predictive maintenance, a better approach rather than preventive or corrective maintenance. This paper investigates the whole adoption process to identify the critical successful factors on how AI technologies can be applied in the next generation condition monitoring process. In addition, the search for open-source codes for image analysis is one of the key goals to assess the applicability and user experience.

Water utility organisations are asset rich. This provides opportunities for the AI technology to be adopted widely to enhance the current asset condition monitoring process, leading to a preventative maintenance strategy. However, the design of an AI adoption process is a complex subject, requiring thorough understanding of data management, sensor and data acquisition capabilities and advanced analytics techniques. Therefore, this project aims at identifying a framework to design an AI adoption process to enhance the current condition monitoring process, ultimately forming a preventative asset maintenance strategy. This research reported a general overview of solar panel maintenance by experts, including current practices, future expectations, and the adoption of thermal imaging for detecting failures in solar systems, as well as reviewing the literature on algorithms for identifying hotspots using Matlab, and studying four algorithmic options for analysis in addition to the impact of hotspots on the efficiency loss of solar panels. Among the algorithms evaluated, Algorithm 2 showed the lowest coefficient of variation (4.3%), indicating the most stable and consistent performance for the image dataset considered in this study. This algorithm included the application of equalization after each filter (mean, median, Gaussian, and bilateral), followed by converting RGB channels to HSV, using a threshold technique in the segmentation stage. Additionally, the red and green channels were adjusted before binarizing the image for morphological operations, which calculated the area of white regions. Using the proposed methodology and the assumptions adopted in this study, hotspot-related efficiency losses were estimated to be between 1.3% and 1.5% over the one-year period. These estimates should be interpreted as indicative because they are based on image analysis and assumed model parameters rather than direct electrical performance measurements.

The findings demonstrate that UAV-based thermal imaging combined with AI-assisted image processing provides a promising basis for supporting solar panel condition monitoring and maintenance planning. However, further research is required to validate the proposed workflow using larger and more diverse datasets, expert-labelled reference images, and direct field measurements of electrical performance.

DISCLOSURE STATEMENT

No potential conflict of interest was reported by the authors and contributors.

APPENDIX: SUPPLEMENTARY INFORMATION

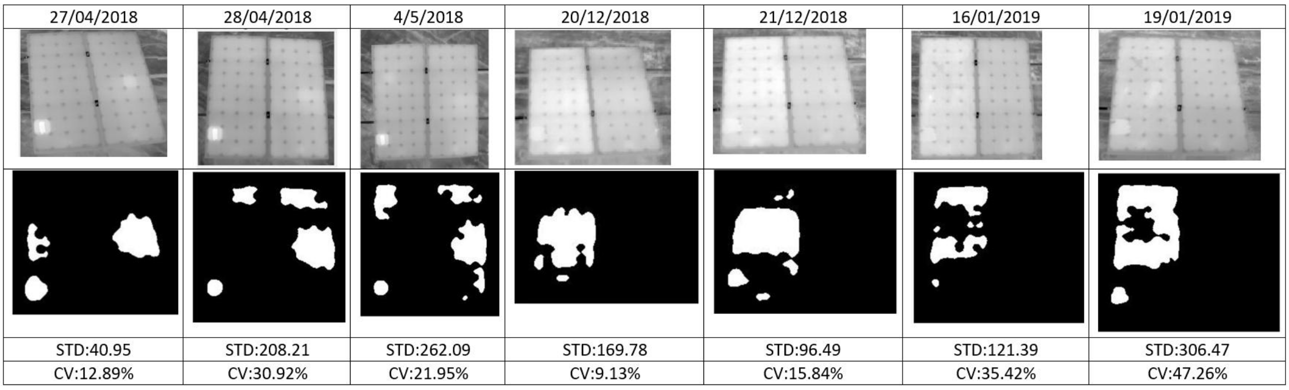


Figure S1. Graphical representation for left panel side applying Algorithm 1

Table S1. Variation coefficient for Algorithm 1

Date	Area (cm^2)	STD	CV
2018 Apr 27	317.54	40.95	12.89%
2018 Apr 28	673.36	208.21	30.92%
2018 May 4	1194.08	262.09	21.95%
2018 Dec 20	1860.37	169.78	9.13%
2019 Dec 21	609.03	96.49	15.84%
2019 Jan 16	342.72	121.39	35.42%
2019 Jan 19	648.44	306.47	47.26%

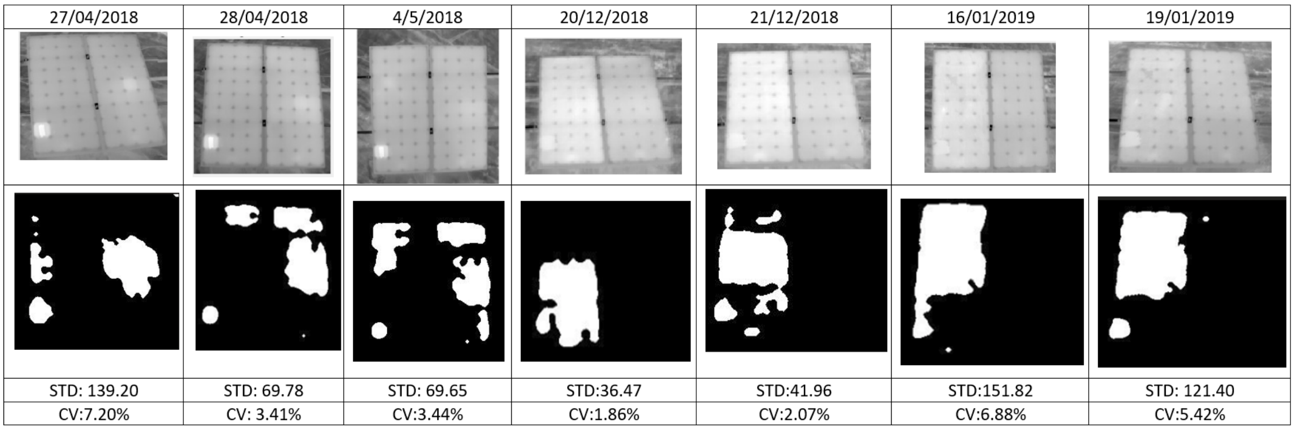


Figure S2. Graphical representation for left panel side applying Algorithm 2

Table S2. Variation coefficient for Algorithm 2

Date	Area (cm^2)	STD	CV
2018 Apr 27	1933.97	139.20	7.20%
2018 Apr 28	2047.85	69.78	3.41%
2018 May 4	2025.03	69.65	3.44%
2018 Dec 20	1962.00	36.47	1.86%
2019 Dec 21	2023.10	41.96	2.07%
2019 Jan 16	2207.66	151.82	6.88%
2019 Jan 19	2241.01	121.40	5.42%

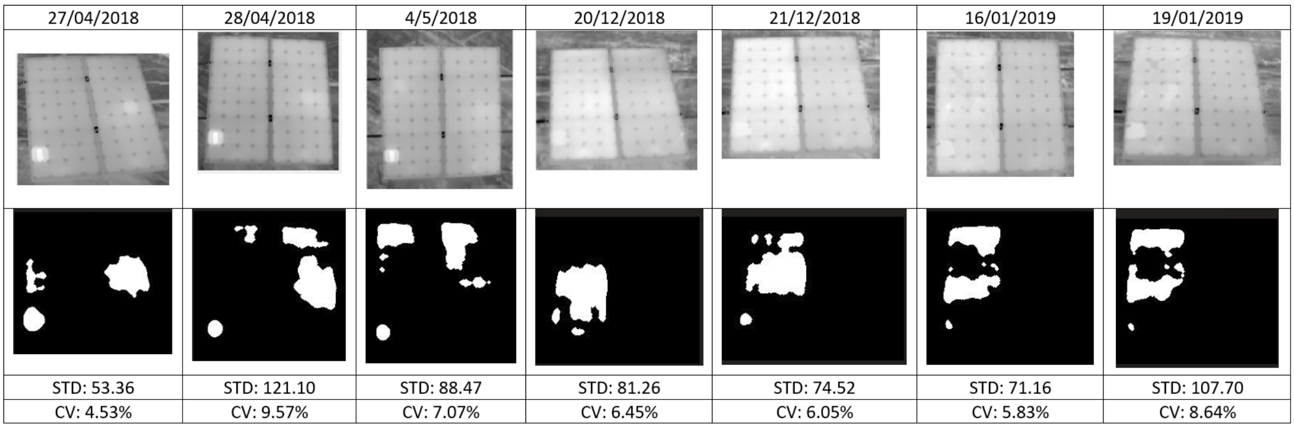


Figure S3. Graphical representation for left panel side applying Algorithm 3

Table S3. Variation coefficient for Algorithm 3

Date	**Area (cm^2)**	**STD**	**CV**
2018 Apr 27	1177.45	53.36	4.53%
2018 Apr 28	1264.93	121.10	9.57%
2018 May 4	1251.02	88.47	7.07%
2018 Dec 20	1259.09	81.26	6.45%
2019 Dec 21	1230.81	74.52	6.05%
2019 Jan 16	1220.90	71.16	5.83%
2019 Jan 19	1246.07	107.70	8.64%

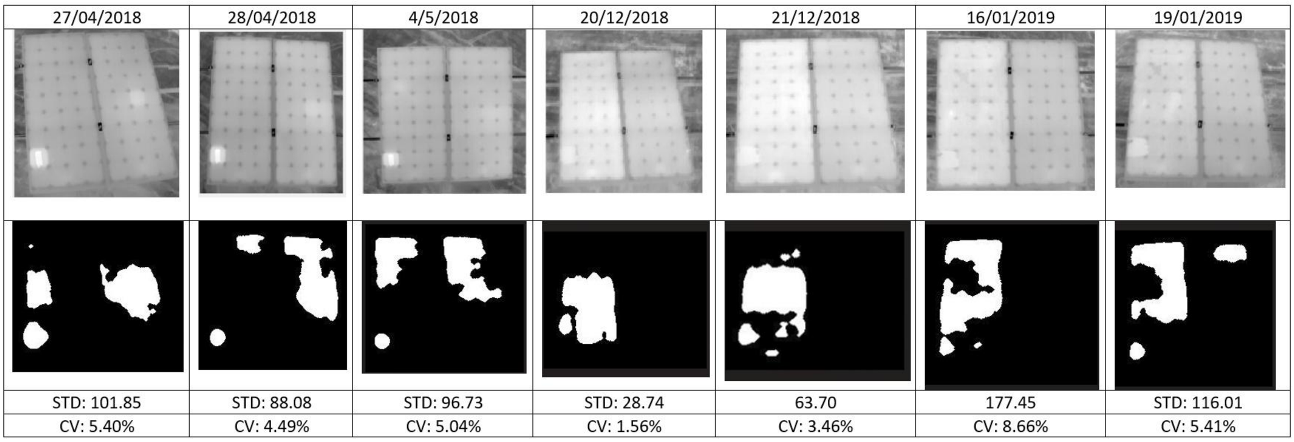


Figure S4. Graphical representation for left panel side applying Algorithm 4

Table S4. Variation coefficient for Algorithm 4

Date	**Area (cm^2)**	**STD**	**CV**
2018 Apr 27	1885.14	101.85	5.40%
2018Apr 28	1962.33	88.08	4.49%
2018 May 4	1917.87	96.73	5.04%
2018 Dec 20	1841.05	28.74	1.56%
2019 Dec 21	1840.82	63.70	3.46%
2019 Jan 16	2047.96	177.45	8.66%
2019 Jan 19	2142.68	116.01	5.41%

REFERENCES

Alajmi, M., Awedat, K., Aldeen, M.S., & Alwagdani, S. (2019). IR Thermal Image Analysis: An Efficient Algorithm for Accurate Hot-Spot Fault Detection and Localization in Solar Photovoltaic Systems. 2019 IEEE International Conference on Electro Information Technology (EIT), 162–168. https://doi.org/10.1109/EIT.2019.8833855.

Alfaro-Mejía, E., Loaiza-Correa, H., Franco-Mejía, E., Restrepo-Girón, A.D., & Nope-Rodríguez, S.E. (2019). Dataset for recognition of snail trails and hot spot failures in monocrystalline Si solar panels. Data in Brief, 26, 104441. https://doi.org/10.1016/j.dib.2019.104441.

Balakrishnan, D., Raja, J., Rajagopal, M., Sudhakar, K., & Janani, K. (2023). An IoT-Based System for Fault Detection and Diagnosis in Solar PV Panels. E3S Web of Conferences, 387, 05009. https://doi.org/10.1051/e3sconf/202338705009.

Chaudhary, A.S. & Chaturvedi, D.K. (2018). Analyzing Defects of Solar Panels under Natural Atmospheric Conditions with Thermal Image Processing. International Journal of Image, Graphics and Signal Processing, 10(6), 10–21. https://doi.org/10.5815/ijigsp.2018.06.02.

Chaudhary, A.S., & Chaturvedi, D.K. (2023). Live Hotspots Visualization and Degradation Analysis of Solar Photovoltaic Panels Using QR Code and Thermal Image Processing. In S. Tiwari, M. C. Trivedi, M. L. Kolhe, & B. K. Singh (Eds.), Advances in Data and Information Sciences (Vol. 522, pp. 23–33). Springer Nature Singapore. https://doi.org/10.1007/978-981-19-5292-0_3.

Constantin, A.-I., Iosif, G., Chihaia, R.-A., Marin, D., Abu Shehadeh, G.U., Karahan, M., Gerikoglu, B., Stavrev, S. (2023). Preventive Maintenance in Solar Energy Systems and Fault Detection for Solar Panels based on Thermal Images. Electrotehnica, Electronica, Automatica, 71(1), 1–12. https://doi.org/10.46904/eea.23.71.1.1108001.

Dhanraj, J.A., Mostafaeipour, A., Velmurugan, K., Techato, K., Chaurasiya, P. K., Solomon, J. M., Gopalan, A., & Phoungthong, K. (2021). An Effective Evaluation on Fault Detection in Solar Panels. Energies, 14(22), 7770. https://doi.org/10.3390/en14227770.

Han, S.H., Rahim, T., & Shin, S.Y. (2021). Detection of Faults in Solar Panels Using Deep Learning. 2021 International Conference on Electronics, Information, and Communication (ICEIC), 1–4. https://doi.org/10.1109/ICEIC51217.2021.9369744.

Hernández-Callejo, L., Gallardo-Saavedra, S., & Alonso-Gómez, V. (2019). A review of photovoltaic systems: Design, operation and maintenance. Solar Energy, 188, 426–440. https://doi.org/10.1016/j.solener.2019.06.017.

Karim, I.A. (2015). Fault analysis and detection techniques of solar cells and PV modules. 2015 International Conference on Electrical Engineering and Information Communication Technology (ICEEICT), 1–4. https://doi.org/10.1109 /ICEEICT.2015.7307349.

Kirubakaran, V., Preethi, D.M.D., Arunachalam, U., Rao, Y.K.S.S., Gatasheh, M.K., Hoda, N., & Anbese, E.M. (2022). Infrared Thermal Images of Solar PV Panels for Fault Identification Using Image Processing Technique. International Journal of Photoenergy, 2022, 1–10. https://doi.org/10.1155/2022/6427076.

Liao, K.C., & Lu, J.H. (2020). Using Matlab real-time image analysis for solar panel fault detection with UAV. Journal of Physics: Conference Series, 1509(1), 012010. https://doi.org/10.1088/1742-6596/1509/1/012010.

Madeti, S.R., & Singh, S.N. (2017). A comprehensive study on different types of faults and detection techniques for solar photovoltaic system. Solar Energy, 158, 161–185. https://doi.org/10.1016/j.solener.2017.08.069.

Numan, A.H., Hussein, H.A., & Dawood, Z.S. (2021). Hot Spot Analysis of Photovoltaic Module under Partial Shading Conditions by Using IR-Imaging Technology. Engineering and Technology Journal, 39 (09), 1338-1344. http://doi.org/10.30684 /etj.v39i9.841.

Orosz, T., Rassõlkin, A., Arsénio, P., Poór, P., Valme, D., & Sleisz, Á. (2024). Current Challenges in Operation, Performance, and Maintenance of Photovoltaic Panels. Energies, 17(6), 1306. https://doi.org/10.3390/en17061306.

Pathak, S.P., & Patil, S.A. (2023). Evaluation of Effect of Pre-Processing Techniques in Solar Panel Fault Detection. IEEE Access, 11, 72848–72860. https://doi.org/10.1109/ACCESS.2023.3293756.

Rahmoune, B.M., Iratni, A., Amari, A.S., Hafaifa, A., & Colak, I. (2023). Fault detection and diagnosis of photovoltaic system based on neural networks approach. Diagnostyka, 24(3), 1–10. https://doi.org/10.29354/diag/166428.

Raikwar, R., Kamble, D., Kamble, M., Kamat, N., Kamble, A., & Kamble, P. (2023). Solar Hotspot Analysis using Image Processing in Matlab. International Journal for Research in Applied Science and Engineering Technology, 11(10), 1073–1076. https://doi.org/10.22214/ijraset.2023.56080.

Skomedal, Å.F., Aarseth, B.L., Haug, H., Selj, J., & Marstein, E.S. (2020). How much power is lost in a hot-spot? A case study quantifying the effect of thermal anomalies in two utility scale PV power plants. Solar Energy, 211, 1255–1262. https://doi.org/10.1016/j.solener.2020.10.065.

Solar Control. (n.d.). How to calculate solar panel efficiency. Solar Control. https://www.solarctrl.com/blog/how-to-calculate-solar-panel-efficiency/

Ujjwal, K., Sheshadri, S.M., Sabhari, D., Dhanuka, A., & Singh, A. (2022). Processing of Thermal Satellite Images Using Matlab. 2022 IEEE Aerospace Conference (AERO), 1–10. https://doi.org/10.1109/AERO53065.2022.9843260.