Taylor & Francis
Taylor & Francis Group

JOURNAL OF QUALITY TECHNOLOGY

Journal of Quality Technology

A Quarterly Journal of Methods, Applications and Related Topics

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/ujqt20

# Online automatic anomaly detection for photovoltaic systems using thermography imaging and low rank matrix decomposition

Qian Wang, Kamran Paynabar & Massimo Pacella

**To cite this article:** Qian Wang, Kamran Paynabar & Massimo Pacella (2021): Online automatic anomaly detection for photovoltaic systems using thermography imaging and low rank matrix decomposition, Journal of Quality Technology, DOI: 10.1080/00224065.2021.1948372

**To link to this article:** https://doi.org/10.1080/00224065.2021.1948372

JOURNAL OF QUALITY TECHNOLOGY
https://doi.org/10.1080/00224065.2021.1948372

CASE REPORT

# Online automatic anomaly detection for photovoltaic systems using thermography imaging and low rank matrix decomposition

Qian Wang[a], Kamran Paynabar[a], and Massimo Pacella[b]

[a]Georgia Institute of Technology, ISyE, Atlanta, Georgia; [b]Dipartimento di Ingegneria dell'Innovazione, Universita del Salento, Lecce, Italy

**ABSTRACT**

Faults occurred during the operational lifetime of photovoltaic (PV) systems can cause energy loss, system shutdown, as well as possible fire risks. Therefore, it is crucial to detect anomalies and faults to control the system's performance and ensure its reliability. Comparing to traditional monitoring techniques based on an on-site visual inspection and/ or electrical measuring devices, the combination of drones and infrared thermography imaging evidently provides the means for faster and less expensive PV monitoring. However, the literature in this area lacks automatic and implementable algorithms for PV fault detection, particularly, using raw aerial thermography, with precise performance evaluation. The objective of this paper is, thus, to build a fully automatic online monitoring framework. We propose an analytical framework for online analysis of the raw video streams of aerial thermography. This framework integrates image processing and statistical machine learning techniques. We validate the effectiveness of the proposed framework and provide sufficient details to facilitate its implementation by practitioners. Two challenges hinder direct fault detection on raw PV images. One is that raw PV images often have non-smooth backgrounds that can impact the detection performance. This background needs to be removed before fault detection. However, this is a daunting task given the perspective of images. To deal with this challenge, we utilize the Transform Invariant Low-rank Textures (TILT) method to orthogonalize the perspective before applying edge detection to crop out the background and aligning the cropped images. The other issue is that the regular hot spots at the bottom edges of the solar panels are normal and should not be detected as anomalies. This makes the intensity-based detection method in the literature fail. These hot spots are part of the low-rank pattern of the image sequence. On the other hand, the hot spots caused by anomalies deviate from the normal low-rank pattern of the PV cells. Therefore, we propose a methodology that relies on Robust Principal Component Analysis (RPCA), which can separate sparse corrupted anomalous components from a low-rank background. The RPCA is applied to the PV images for simultaneous detection and isolation of anomalies. In addition to RPCA, we suggest a set of post-processing procedures for image denoising, and segmentation. The proposed algorithm is developed using 20 normal (with no anomalies) training samples and 100 test samples. The results showed that the algorithm successfully detects the anomalies with a recall of 0.80 and detects the significant anomalies with the maximum recall of 1. Our method outperforms two benchmark methods in terms of F1 score by 44.5% and 114.3%. The small number of false alarms is mostly due to irregular image patterns at the end of a PV array or an extreme non-orthogonal perspective. Since the number of false alarms is not large, it does not disrupt the inspection process, and they can easily be identified by an appraiser offline. The average computation time is 6.32 sec/image, which enables online automatic inspection of PV panels.

**KEYWORDS**

Aerial thermography; automatic anomaly detection; robust PCA; PV system

## 1. Problem description

Nowadays, there is a growing interest in renewable energy and photovoltaic (PV) systems. PV systems are pollution-free, noiseless, modular, and easily installable in several locations. In a PV system, numerous solar cells convert solar energy into current through the photoelectric effect of doped semiconductor material. Simplifying, a solar cell is a diode designed to absorbs photons with energy. This creates a difference of potential in the internal junction zone. The

CONTACT Kamran Paynabar kamran.paynabar@isye.gatech.edu Georgia Institute of Technology, ISyE, Atlanta, GA 30332.
Supplemental data for this article is available online at https://doi.org/10.1080/00224065.2021.1948372.

polarization of the junction zone prevents electrons from crossing it, hence producing a voltage between the diode terminals.

The output voltage of a single solar cell is about half a volt. Since this voltage is too low to use in most applications, solar cells are electrically connected in series. A set of 15 to 25 solar cells is encapsulated into a single and long-lasting unit, known as a PV module (the most basic and building block in a PV system). The encapsulation structure enables PV modules to withstand adverse environmental conditions. When a single solar cell fails, the entire PV module housing the failed cell has to be replaced. Faulty modules unnoticed for long periods of time may conduct to a destructive failure of PV panels. This may cause a large amount of energy loss, system shut down, and possible fire risks. Therefore, effective fault diagnosis is required to ensure reliable and safe operations in PV power plants.

Lorenzo et al. (2020) reviewed several causes of faults that may occur in PV modules, based on data acquired from a database of more than 80 PV power plants. Hot spots represent a commonly reported fault, which occurs when there is at least one solar cell with a short-circuit current much smaller than the others. In such cases, the defective cell is forced to pass a current higher than its generation capabilities, thus dissipating power in the form of heat instead of producing electrical power. This power dissipation occurring in a small area results in a local overheating, namely the hot spot. Hot spots correspond to those areas in PV modules characterized by the higher temperature. This may cause destructive effects such as cell or glass cracking, sintering, excessive degradation of the PV module, and eventually fire risks.

Various methodologies have been reported in the literature for automatically classify the PV condition as defective or healthy. These methodologies can be broadly distinguished as (i) signal processing-based (SP-based), mostly electrical signals, and (ii) image processing-based (IP-based), mostly infrared thermography.

The signal-based methods use fault signatures from sensor data to detect faults (Madeti and Singh 2017). Commonly used SPl-based methods include statistical signal processing, I-V (current-voltage) characteristics analysis, and electrical power loss analysis. More recently, SP-based methods, which use available diagnosis data and machine learning techniques, have shown improvements in detection accuracy due to their highest model representation capabilities. The data are first used to identify the model in the training phase, which is then used to diagnose faults in the testing phase. These methods leverage labeled data and powerful models from machine learning techniques to perform multi-class regression or classification, which are quite important for detecting faults. In the literature, methods using Gaussian process regression (Fazai et al. 2019) and random forest (Dhibi et al. 2020) have been proposed. Machine learning methods, which use the PCA tool for feature extraction and selection, have also been discussed for PV fault detection (Hajji et al. 2021).

When compared with standard model-based approaches, these methods provide good modeling and prediction performances. However, detecting signal measurements can only be applied to string terminals of multiple PV modules (AbdulMawjood, Refaat, and Morsi 2018). For example, string terminal voltage and current can be measured and tested to identify possible defects like disconnection or degradation because such defects disturb the PV string I-V curve. Using an SP-based method, one cannot identify the particular faulty PV modules within the string.

Recently, IP-based methods, such as infrared thermography, have been successfully applied for PV inspection. Through these methods, it is possible to identify local anomalies in a PV module and to hypothesize possible types of fault. Additional advantages are contactless detection, a fast inspection of large plants, and continuity of the plant operation during diagnosis (Teubner et al. 2017), which make IP-based approaches more time efficient and less expensive compared with IP-based ones.

The most comprehensive work in comparing SPl-based and IP-based methods for PV faults detection is given by Tsanakas, Ha, and Buerhop (2016), where the employment of such methods is widely described and discussed, starting from a classification of possible types of fault. In this review, an extensive classification study between thermal image pattern, I-V curve shape, and possible fault type, is also reported. This fault classification can be of help for the fast identification of fault typologies.

### 1.1. Literature analysis

Conventionally, faults in PV installations have been detected by applying electrical tests. Hence, various studies were conducted on developing statistical analysis for monitoring the quality of PV power plants by measuring electrical outputs on the terminals of the modules. Madeti and Singh (2017) provided a comprehensive state-of-the-art review on PV monitoring systems based on electrical testing. Besides data analysis methods, they discussed various practical aspects

crucial for the development of monitoring systems. These include working principles and operating characteristics of various PV systems, selection and operation of monitoring instruments, communication and storage in data acquisition systems, and challenges and opportunities. Ventura and Tina (2015) applied this approach for monitoring the electrical performance ratio of a PV site. Zhao et al. (2014) proposed a fault detection method based on three different outlier detection rules. Dhimish and Holmes (2016) presented a PV fault diagnostic algorithm that can diagnose faults based on the comparison between the theoretical and measured electrical data using statistical techniques. Garoudja et al. (2017) used a simulation model to predict the maximum current, voltage, and power generated from the PV system. The differences in values between the measured and predicted electrical performances were considered to signal a fault. Under normal operating conditions, these differences are expected to vary with a mean value of zero and constant variance. An EWMA control chart was employed to detect faults in the monitored PV system. Harrou et al. (2018) discussed a similar statistical control chart methodology for monitoring PV systems based on an electrical simulation model. In particular, they used a multivariate EWMA (MEWMA) control chart to detect potential faults and a univariate EWMA to identify the type of detected faults. Mansouri et al. (2018) proposed a wavelet approach combined with an EWMA technique for monitoring several variables such as currents, voltages, and power.

Faults in a PV installation can also be located through visual assessment of modules by an operator. Imaging techniques may be used as a supporting tool for analysis during a visual inspection by the operator. Such imaging techniques include infrared thermography, ultraviolet fluorescence, and electroluminescence. Malfunctioning PV modules typically have a higher temperature than adjacent normal ones regardless of the fault reason. This makes infrared thermography, which measures higher or lower heat emissivity of areas or points, the most commonly used imaging technique. The manual inspection of the sequence of thermal images, in this case, is a time-consuming operation. The main issue of manual inspection, performed by an operator equipped with a thermal camera, is the significant inspection time. This time can be even longer for rooftop systems. This is a critical aspect for inspection of PV plants as to obtain reliable results from the thermographic images, high solar irradiation is required. This issue has recently been addressed by using aerial thermography based on a drone equipped with a thermal camera. Aerial thermography by drones (UAVs - Unmanned Aerial Vehicles or RPAs - Remotely Piloted Aircrafts) has become popular for acquiring images and videos related to a PV power plant. The development of electronic sensors and instrumentation involved in aerial thermography has increased the potential of this technique. Rahaman, Urmee, and Parlevliet (2020) conducted a comprehensive review on different aspects of aerial thermography for inspection of PV power plants. Two main advantages of aerial thermography are i) identification and localization of failed modules without prior planning, and ii) inspection without interrupting normal operation of the PV power plant, thus reduction of maintenance and operational costs and increase of the plant capacity factor.

Aerial thermography is also reported in the literature for monitoring PV power plants (Rahaman, Urmee, and Parlevliet 2020). The idea of integrating infrared thermography into a novel inspection method that uses a drone started to appear in 2012. In particular, Denio III (2012) proposed aerial thermography to easily reach PV modules and/on rooftops or large areas. Subsequently, Grimaccia et al. (2015, 2017) and Buerhop et al. (2018) conducted several research studies to prove the feasibility of aerial thermography inspection of PV sites. Some advantages of this method are large area coverage, fast detection, cost-effectiveness, time-saving, and multiple sensors data collection. Buerhop et al. (2012, 2018) used aerial thermography and employed a remote-controlled helicopter to investigate several PV installations with different characteristics. The authors showed that mounting the thermal camera onto a remote-controlled drone facilitates the inspection and makes it possible to detect many types of faults and failures. To automate processes in large PV sites, Tsanakas, Ha, and Shakarchi (2017) proposed two different techniques for the mapping of aerial thermography inspection: the aerial triangulation that uses image data obtained with a drone to generate orthophoto mosaics; and terrestrial georeferencing that associates terrestrial images taken in different positions of the PV site with geographic data.

As evident, traditional monitoring techniques based on onsite visual inspection or electrical tests are slow and expensive. Comparing to traditional techniques, the combination of drones and thermal cameras provides the means for faster and less expensive PV monitoring. This motivates the automatic aerial inspection of PV plants using image processing algorithms. Colaprico et al. (2018) developed an image-processing based commercial software module called DUBIO,

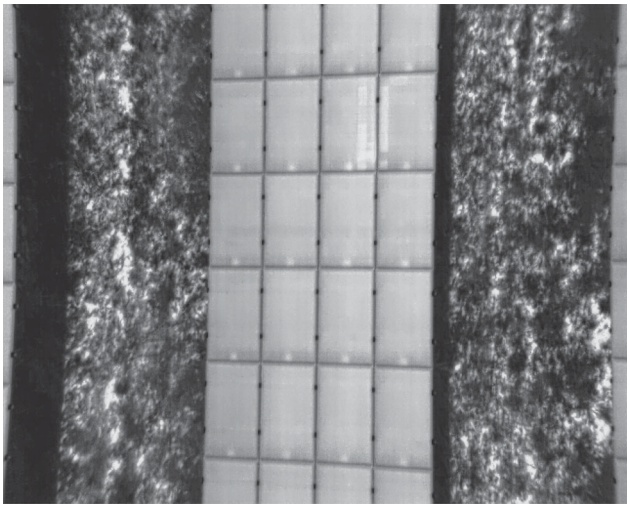

**Figure 1.** Raw image directly from drone-based infrared thermography.

which can automatically elaborate infrared thermography images acquired by drones over large PV sites. However, the authors did not provide details or systematic evaluation of the implemented algorithms. Alsafasfeh et al. (2018) proposed a real-time inspection of the solar gardens based on the use of a drone equipped with two cameras, a thermal and a charge-coupled device mounted on a drone. In the paper, a simple algorithm for image processing is described, in which a segmentation based on hot pixels is introduced. However, the algorithm details required to implement their method are not provided in the paper. Hence, their analysis is not repeatable, and practitioners may not be able to utilize this method. Libra et al. (2019) discussed different drone-based systems equipped with a thermal camera for continuous monitoring of PV plants. These methods are effective for damage identification that may be difficult, e.g., a cell cracking that is not visible through visual inspection and can only be detected through electroluminescence or lock-in thermography. However, they discussed the aerial inspection method from an engineering point-of-view, without providing an automatic detection algorithm. Li et al. (2020), proposed an algorithmic solution for the rapid detection of multiple visible defects in PV modules by using the true color images taken from a drone-based inspection system. The paper mainly focuses on defect classification and is based on true-color images instead of infrared images. Aghaei et al. (2015) proposed an automatic detection method utilizing the pixel intensity distribution of each PV image to determine hot areas, and consequently, the degradation percentage. The anomaly is detected if the degradation percentage is larger than a pre-specified threshold. The detection method by Aghaei et al. (2015) is the most repeatable in the literature as it includes sufficient details about the algorithm. However, the paper does not report sufficient numerical results on detection performance. Moreover, their detection algorithm requires clean and cropped images as input. In reality, as shown in Figure 1, the raw thermal images taken by drones include the irrelevant soil background and have perspective distortion.

To the best of our knowledge, there is a lack of online and automatic algorithms in the literature, for anomaly detection in PV panels using aerial infrared images.

### 1.2. Main purpose

Given the growth in the size of modern PV power plants, the main inspection challenge is to automate and expedite the online detection of anomalies using the video stream of thermal cameras. A research effort to automate data collection and preparation, data analysis, and post-processing steps in aerial thermography is necessary to move toward a fully automatic monitoring process.

The reviewed methods in this area either fail to provide algorithmic details for implementation or do not accurately evaluate the performance of detection algorithms using real data. Moreover, some cannot deal with the raw image feed from the camera. This motivates us to build a fully automatic online image-based monitoring system for anomaly detection in PV panels by integrating image processing and statistical machine learning techniques. We validate the proposed framework by evaluating its detection performance using a sufficiently large sample of real data. Additionally, we provide the required details and codes for implementing this framework in practice.

## 2. Data collection and preparation

The image dataset was generated using drone-based infrared thermography, as shown in Figure 1. The basis of the measurement system was a remote-controlled quadcopter, specifically the DJI MATRICE 210 RTK drone depicted in Figure 2, which is generally used for inspection of power lines, wind turbines, bridges, and solar panels. The model is capable of a 24-minute flight (full payload) on a single battery. It performs autonomous flight along the pre-programmed route. The flight controller consists of two barometer sensors and a global positioning system (GPS) module. The barometer sensors help the quadcopter flying at a constant height and hovering, which is an important prerequisite

**Figure 2.** The DJI MATRICE 210 RTK drone with infrared thermal camera and remote controller.

in aerial imaging. The GPS module also allows the quadcopter to fly the programmed route.

A lightweight thermal camera is installed beneath the quadcopter. The camera is characterized by an uncooled vanadium oxide microbolometer with the possibility of carrying out the measurement temperature in the range of from $-25^{\circ}$ C to $135^{\circ}$ C. The main features of this camera are: resolution $640 \times 512$ pixels, lens 19 mm $32^{\circ}$ (H) $\times$ $26^{\circ}$ (V) and thermal sensibility $< 0.05^{\circ}$ C. Image stabilization during the flight is provided by a gimbal device, which integrates images from an infrared thermal camera and a visual camera.

Test objects include the PV panels operational in a PV power plant located in northern Italy, with an overall peak power of approximately 6 MWp. In this plant, the PV panels are installed on flat terrain. Each panel consists of 60 PV modules using as embedding material ethylene vinyl acetate. The dimension of each panel is 1650 mm long and 992 mm wide. The PV panels are installed at an azimuth angle of $180^{\circ}$ (South) and a tilt angle of $20^{\circ}$.

The flight path of the quadcopter had an altitude of approximately 20 m above the PV panels. The observation angle of the infrared camera was as orthogonal as possible to the ground of the plant. The observation angle of the infrared camera toward the glass surface of the PV panels was less than $30^{\circ}$. In our research, we extracted infrared images for each panel at the original resolution of $640 \times 512$ pixels from recorded video sequences (see Figure 1 for an example).

In the present study, we consider 120 solar panel images, 14 of which are defective, and the remaining 106 are normal. We randomly choose 20 normal images for training. So, the size of the test dataset is 100 (86 normal and 14 defectives).

## 3. Analysis and interpretation

### 3.1. Overview of proposed algorithm

As discussed earlier, there are two analytical challenges in developing an automatic online monitoring system: One is that raw PV images often have soil backgrounds that can impact the detection performance. This background needs to be removed before fault detection, which is challenging due to the image perspective. To deal with this issue, we utilize the Transform Invariant Low-rank Textures (TILT) algorithm (Zhang et al. 2012) to orthogonalize the perspective before applying edge detection to crop out the background and aligning the cropped images. The other issue is that the regular hot spots at the bottom edges of the solar panels are normal and should not be detected as anomalies. This makes the intensity-based detection methods in the literature fail. These hot spots are part of the low-rank pattern of the image sequence. On the other hand, the hot spots caused by anomalies deviate from the normal low-rank pattern of the PV cells. Therefore, we propose a methodology that relies on Robust Principal Component Analysis (RPCA) (Candès et al. 2011), which can separate sparse corrupted anomalous components from a low-rank background. The RPCA is applied to the PV images for simultaneous detection and isolation of anomalies.

Overall, our proposed framework consists of three main stages. 1) *Preprocessing* of images for perspective transformation using the TILT, cropping, and alignment. 2) *Anomaly separation* from the background using RPCA. 3) *Postprocessing* for filtering residual noises and identifying the location of defects. The proposed framework integrates techniques from image processing, optimization, and statistical machine learning. Figure 3, shows a flowchart that details the proposed framework.

The preprocessing stage aims to deal with the first challenge. It begins with image transformation from a perspective view to an orthographic view. This allows removing the image perspective caused by the angle deviation of the camera from the ground line. After this transformation, the boundaries of PV panels in the images are parallel and/or orthogonal to each other. This makes the background cropping and image alignment possible. Figure 4 shows a sample of raw and transformed images. The next step of preprocessing is to remove the image background and crop images along the PV panel borders. To identify the PV panel borders, we apply image segmentation and binarization based on the Otsu's method (Otsu 1979). Since the PV panels have a higher temperature, their

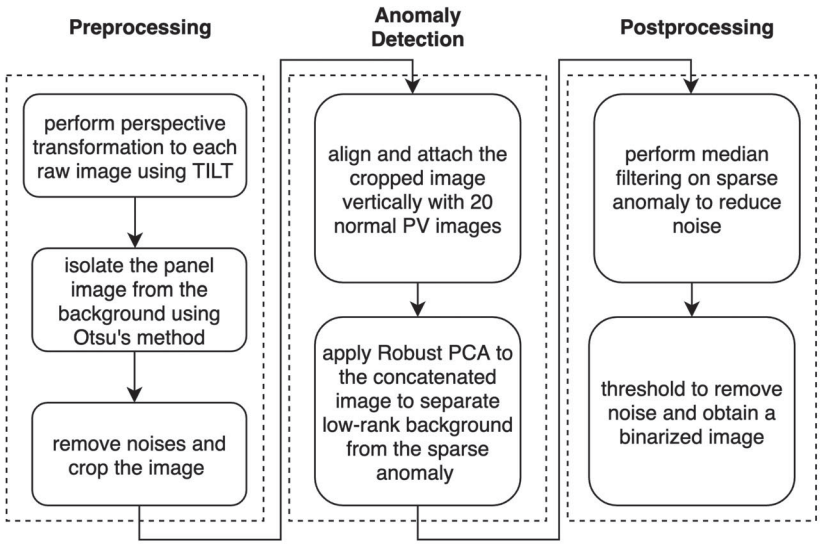


**Figure 3.** Flowchart of the proposed algorithm.

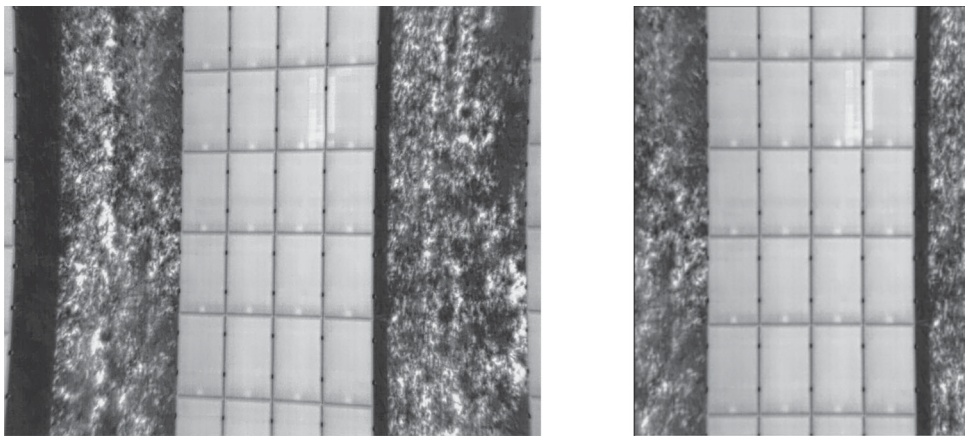

**Figure 4.** Left: The original infrared PV panel image; Right: After the perspective transformation.

corresponding pixels are much brighter than the dark background. This light contrast helps to identify the boundaries. Segmentation of panels and background is followed by thresholding and filtering. This to clean up the hot spots in the background. Figure 5 shows examples of this preprocessing step. Next, using the resulting binary image, the bounding box of the PV region is defined, as shown in Figure 6. Then, the orthographic images obtained from the previous step are cropped along the edges of the box.

After three pre-processing steps, the original images are transformed into the format compatible with RPCA requirements. Anomaly separation using RPCA is the core stage of our anomaly detection framework. Since both the regular and irregular (anomalous) hot spots have similarly high pixel intensities, the intensity-based method tends to detect both as anomalies or normal. PV images are low-rank because of the organization of the PV modules cells in the panels. The regular hot spots close to the bottom edges are a part of this low-rank pattern. This is the key to separating the regular hot spots from anomalies. We intend to separate anomalous hot spots from the underlying low-rank pattern of the PV images, from which we can decide whether the anomalies are significant enough to trigger an alarm. Classic PCA is one way to mine an underlying low-rank pattern. However, its performance is influenced by outliers. Since anomalies in PV images have a large magnitude (high temperature), the classic PCA is not suitable to capture the low-rank pattern. In this case, we choose RPCA, which is robust to outliers and anomalies. After mining the pattern, the anomalies are filtered out by subtracting this pattern from the original image.

To apply RPCA, each pre-processed test image is concatenated with 20 pre-processed normal training images along the $y$ axis. Note that the width of the PV region after the cropping step may vary. Thus, further

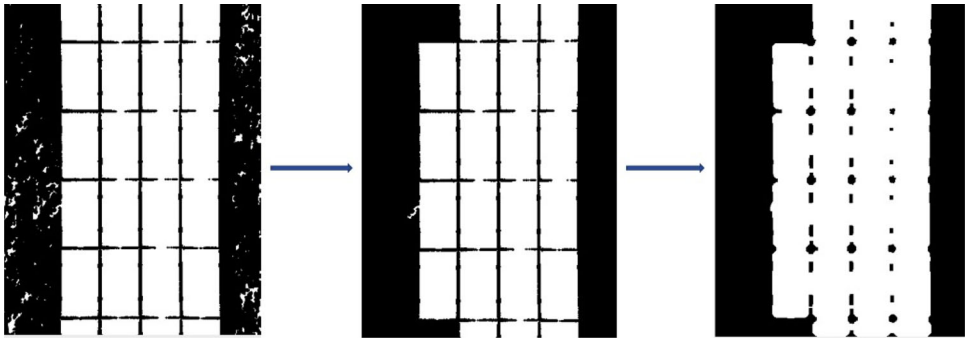

Figure 5. Left: After binarization; Middle: After removing small connected components; Right: After median filtering.

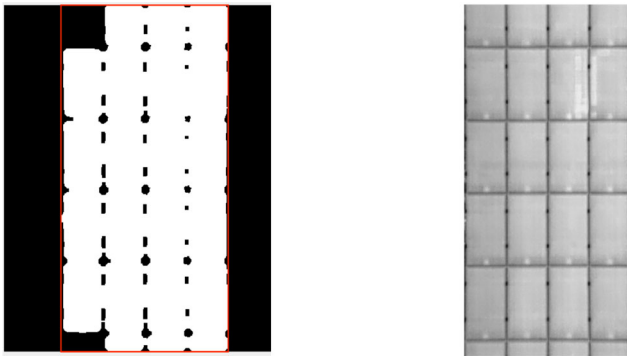

Figure 6. Left: The minimal bounding box; Right: Crop using the bounding box.

alignment of the image width is necessary. The alignment is done by firstly performing image resizing using bicubic interpolation, which changes the image width to a fixed number of pixels. Then, all the resized images are vertically attached together into a long matrix. Subsequently, the concatenated images are fed into the anomaly separation stage using RPCA, leading to the sparse anomalous regions separated from the low-rank structure of the image up to some noise.

In the post-processing stage, the residual noises from the separation stage are removed by performing median filtering. Thresholding is finally applied on the filtered sparse regions detected by RPCA to generate the final detection region.

In the subsequent sections, each step of the proposed framework is elaborated.

### 3.2. Preprocessing

#### 3.2.1. Transformation to orthographic view using TILT algorithm

The images of PV panels taken by drones are in a perspective view, as shown in Figure 4. Therefore, to facilitate the image cropping and recover the low-rank structure for RPCA, we consider a transformation from a perspective to the orthographic view. For this purpose, we utilize the Transform Invariant Low-rank Textures (TILT) algorithm proposed by Zhang et al. (2012).

TILT can recover the plan view of the PV panels as the plan view is intrinsically low-rank. Define $I^0$ as the intrinsic low-rank image with the orthographic view, and $I$ as the perspective image. We assume the perspective image $I$ is obtained by adding a set of sparse noises $E$ to the original low-rank ortho-image $I^0$ before applying inverse transformation $\tau^{-1}$. That is, $I^{\circ}\tau = I^0 + E$, where $^{\circ}\tau$ is the transformation operator. To recover $I^0$ from perspective image $I$, Zhang et al. (2012) proposed to solve the following optimization problem:

$$\min_{I^0, E, \tau} \|I^0\|_* + \lambda \|E\|_1 \quad \text{s.t.} \quad I^{\circ}\tau = I^0 + E. \tag{1}$$

The objective function is a linear combination of the nuclear norm (i.e. $||.||_*$), a convex proxy of the non-convex rank function that encourages the low-rank structure, and the $l_1$ norm (i.e. $||.||_1$) that encourages the noise sparsity and its small magnitudes. The transformation operator, $\tau : \mathbb{R}^2 \to \mathbb{R}^2$ belongs to the homography group $\in GL(3)$ with 8 transformation parameters $a, b, c, d, e, f, g, h$ that is represented by

$$\tau(x, y, 1) = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \tag{2}$$

$$= (ax + by + c, dx + ey + f, gx + hy + 1) \tag{3}$$

$$= \left( \frac{ax + by + c}{gx + hy + 1}, \frac{dx + ey + f}{gx + hy + 1}, 1 \right), \tag{4}$$

under the homogeneous coordinate system.

The intuition of the algorithm is to find an orthographic transformation under a noisy setting such that the transformed image $I^0$ has the lowest rank. The resulting $I^0$ is likely the plan view of the original image $I$ due to the low-rank structure of the PV panels.

To solve the foregoing optimization problem Zhang et al. (2012) proposed an optimization

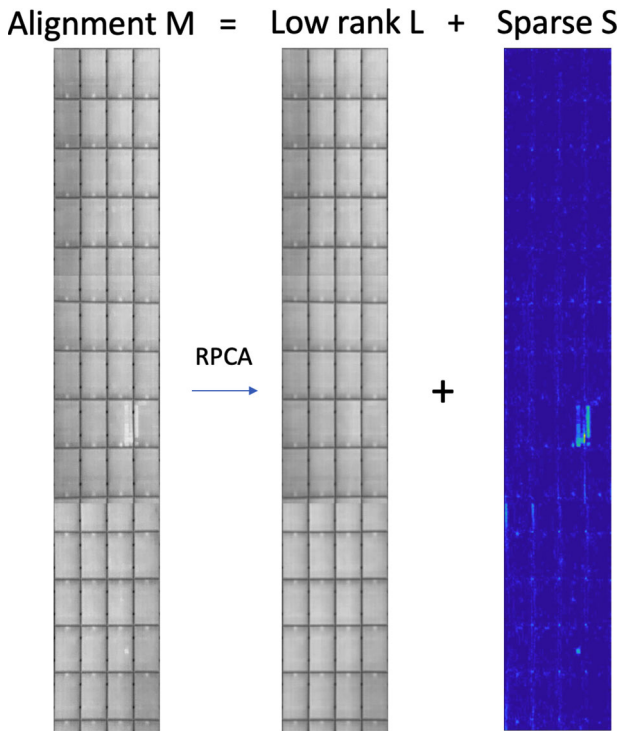


**Figure 7.** Robust PCA decomposition.

algorithm that involves the local linearization of the constraint as well as the use of the alternating direction method of multipliers (ADMM) approach discussed by Boyd et al. (2010). A summary of the optimization algorithm is given in the online appendix, while the detailed information can be found in Zhang et al. (2012). After solving the optimization problem, the obtained transformed image $I^0$ is used in the next step. A sample recovered image with a plan view is given in Figure 4.

#### 3.2.2. Automatic cropping and image alignment

The next step is to crop the rectangular panel region and separate it from the ground area. To detect the PV panel borders, we first apply the well-known Otsu's method (Otsu 1979) that clusters an image into binary pixels. Darker regions associated with the ground and the panel edges are assigned to the cluster with the pixel value of 0, while the brighter regions associated with the panels are assigned to another cluster with the pixel value of 1. This is done by finding a threshold on the pixel intensity that maximizes the interclass variance of two clusters given by

$$\omega_0(t)\omega_1(t)(\mu_0(t) - \mu_1(t))^2, \tag{5}$$

where weight $\omega_i(t)$ is the probability of the pixel intensity belonging to class $i$, $t$ is the intensity threshold, and $\mu_i(t)$ is the mean intensity of class $i$.

Since the images are grayscale, the pixel intensity ranges between 0 and 255. If a pixel intensity is less than the obtained threshold, the pixel is assigned to cluster 0. Otherwise, it belongs to cluster 1. An example of a binarized image is given in the left panel of Figure 5.

After obtaining the binary images, we perform the following tasks to clean up each image:

1. Connected pixels with an area smaller than half of the PV cell are removed. This can help to remove most of the noisy white patches in the ground region. An example is given in the middle panel of Figure 5.
2. The median filtering is applied to remove possible small appendices attached to a unit. An example of this is given in the right panel of Figure 5.

After cleaning up each image, the minimal bounding box of the white region is found. It corresponds to the PV panel and is hence used to crop the image. The bounding box is found by finding the minimum and maximum of $x$ and $y$ coordinates of white pixels, respectively. See Figure 6 for a graphical illustration.

Following the cropping step, each test image is vertically attached to a sequence of 20 normal training images as shown in the left panel of Figure 7. These normal images, corresponding to a "non-anomalous" panel recognized by human experts offline, are used to enhance the low-rank pattern of normal PV panels, boosting the performance of RPCA.

### 3.3. Anomaly detection using robust PCA

After aligning each cropped images with 20 normal images, a vertically long matrix, denoted by $M$ is obtained. Note that in the absence of anomalies, the matrix $M$ is low rank because of the linear dependency among its rows and its columns resulting from a similar pattern that normal images share. This inspires us to apply the Robust PCA method that decomposes the matrix $M$ into a low-rank matrix $L$ and a sparse anomaly matrix $S$ using the following optimization problem (Candès et al. 2011):

$$\min_{L,S} ||L||_* + \lambda||S||_1 \quad \text{s.t.} \quad L + S = M, \tag{6}$$

where $M$ is the input image, $||.||_*$ is the nuclear norm, encouraging the recovery of the low-rank matrix $L$, $||.||_1$ is the $L_1$ norm, helping with the sparse estimation of the anomalous regions, and $\lambda$ is a tuning parameter adjusting the level of the sparsity. The formulation is under the weak assumptions that the

low-rank component $L$ is not sparse, and the sparse anomaly matrix $S$ is not low-rank (i.e. corruption pattern is selected uniformly at random). Both assumptions are valid in our application, making RPCA an appropriate tool for anomaly detection.

To understand the intuition behind RPCA, let us plug the constraint into the objective function by replacing $L$ with $M - S$. The resulting problem is given by $\min_S ||M - S||_* + \lambda||S||_1$. It is clear from this optimization problem that RPCA tries to estimate a sparse anomaly matrix such that the residual image after removing the sparse anomaly is low-rank, i.e., it is like a normal image.

The RPCA optimization problem can be solved using the ADMM. The detailed algorithm is given in Algorithm 1. This is a two-step iterative algorithm that involves the soft thresholding function given by $\mathcal{S}_\gamma(x) = \text{sign}(x)\max(|x| - \gamma, 0)$ with $\gamma$ as the threshold, and the singular value thresholding operator $\mathcal{D}_\gamma(X) = U\mathcal{S}_\gamma(\Sigma)V^T$ where $X = U\Sigma V^T$, is the singular value decomposition. The algorithm solves the augmented Lagrangian objective function in alternating directions, i. e.,

$$l(L, S, Y) = ||L||_* + \lambda||S||_1 + \langle Y, M - L - S\rangle + \frac{\mu}{2}||M - L - S||_F^2 \quad (7)$$

where $\mu > 0$ is the coefficient associated with the augmentation term to ensure the strong convexity property without changing the minimizer of the objective function, which can speed up the convergence.

In practice, $\mu$ is selected as $\mu = n_1 n_2/4||M||_1$ where $n_1$ and $n_2$ are the length and width of the image. The stopping criterion is chosen as $||M - L - S||_F \leq \delta||M||_F$ with $\delta = 10^{-7}$. Also, $\lambda$ is selected universally to be $1/\sqrt{\max(n_1, n_2)}$, which is derived from mathematical analysis to ensure good theoretical properties (Candès et al. 2011).

**Algorithm 1** Solve Robust PCA via ADMM

1: **procedure** ROBUST PCA
2: $S_0 = Y_0 = 0, \lambda > 0, \mu > 0$
3: **while** not converged **do**
4: $L_{k+1} = \mathcal{D}_{\mu^{-1}}(M - S_k + \mu^{-1}Y_k)$
5: $S_{k+1} = \mathcal{S}_{\lambda\mu^{-1}}(M - L_{k+1} + \mu^{-1}Y_k)$
6: $Y_{k+1} = Y_k + \mu(M - L_{k+1} - S_{k+1})$
7: **return** $L$, $S$

An example of a concatenated image $M$ decomposed into the low-rank matrix $L$, and the sparse anomaly $S$ is shown in Figure 7. Note that the edges between PV panels, the brighter spots in the bottom middle of each cell, and the shadows due to illumination may cause some noises in the decomposed $S$. To remove such noises, we perform a postprocessing stage described in the subsequent section.

### 3.4. Postprocessing

To remove the noise in the estimated $S$, we perform median filtering followed by thresholding to binarize the image and highlight the anomalous region. The median filtering operator replaces each pixel with the median of neighbor pixels. While median filtering can remove the noise to a large extent, it does not remove the anomalies because the anomalous pixels are often clustered, thus, robust to median filtering. A graphical illustration of postprocessing is shown in Figure 8.

The proposed anomaly detection framework is summarized in Algorithm 2. The MATLAB packages are used to implement the proposed framework. It is worth noting that the whole algorithm can automatically be streamlined without human interaction.

**Algorithm 2** Anomaly Detection for PV data

1: **procedure** ANOMALYDETECTION
2: Perform perspective transformation to each raw image using TILT
3: Use Otsu's method to binarize the images
4: Remove small connected components using median filtering and thresholding
5: Find the smallest bounding box of the solar panel region to crop the images
6: Concatenate and align each image vertically with a set of 20 normal images
7: Apply Robust PCA to the resulting image matrix $M$ and get the low-rank texture $L$ and sparse anomaly $S$
8: Perform median filtering followed by the thresholding to obtain a binary image with white area denoting the anomalous regions.

### 3.5. Performance evaluation of the proposed algorithm

To evaluate the performance of the proposed algorithm, a set of 120 infrared images are used. Out of 120 images, 14 contain anomalies, and the rest are normal. We randomly choose 20 out of 106 normal images as the training set to attach to the end of each test image and the remaining 100 images, including 14 anomalous ones, as the test set. For each image in the test set, we follow the algorithm mentioned above with the 20 training images attached to detect anomalies in the test image and calculate relevant evaluation metrics. We repeat the evaluation process 50 times

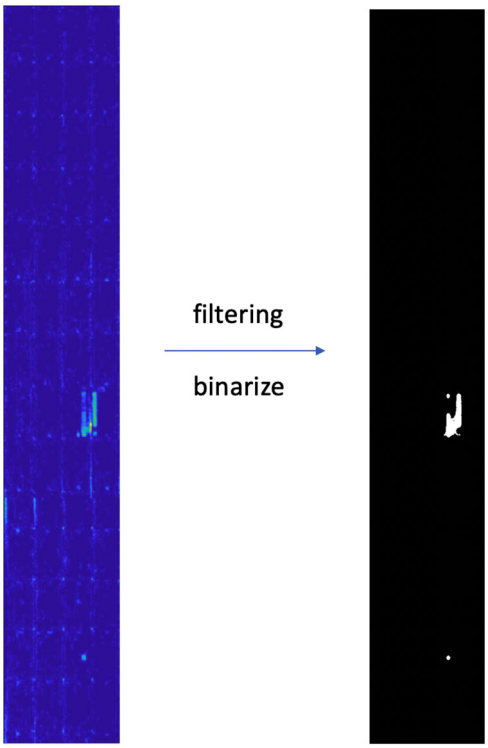


**Figure 8.** Postprocessing.

and report the average of the evaluation metrics and run time. The program is run on a 2.6 GHz 6-Core CPU with and memory size of 16 GB.

Moreover, we adopt a benchmark from Aghaei et al. (2015) and compare its performance with the RPCA component of the proposed framework. The benchmark method requires clean and cropped thermal images and does not perform well on raw images. To emphasize the importance of our pre-processing method, we consider two scenarios; one with pre-processed and cleaned images using our pre-processing steps, designated by B1, and another with raw images, designated by B2. We compare our method's performance with that of the benchmark under both scenarios. Aghaei et al. (2015) define hot areas in each image by following the rule" $IF\text{Lum}(x,y) > \text{Lum}_{AVG} + 5\sigma$ THEN HotArea $(x,y) = 1$ ELSE HotArea $(x,y) = 0$", where $\text{Lum}(x,y)$ represents the pixel intensity at $(x, y)$ and $\text{Lum}_{AVG}$, $\sigma$ represent the mean and standard deviation of pixel intensities of the PV image, respectively. Then, they calculate the degradation percentage by

$$\%Degradation = \frac{\text{Hot Area}}{\text{Total Area}} 100, \tag{8}$$

and classify a PV module as anomalous if the degradation percentage exceeds 5%. Since the image intensity in our dataset is different from that discussed in their paper, to achieve the best performance, we slightly modify their method by defining a hot area if" $\text{Lum}(x,y) > \text{Lum}_{AVG} + 2\sigma$". For each replication of the experiment, we randomly choose 20 in-control training images to calculate their degradation percentages before finding the $95^{th}$ percentile as the threshold. Then, for the remaining 100 test images, we compare the degradation percentage with the threshold and raise alarms if a degradation percentage exceeds the threshold. The experiment is also replicated 50 times, and the average performance metrics are reported.

**Table 1.** Confusion tables for the evaluation result.

| | | Predicted | | |
|---|---|---|---|---|
| | | Anomalous | Normal | Total |
| **Proposed Method** | | | | |
| True | Anomalous | 11.2 | 2.8 | 14 |
| | Normal | 3.52 | 82.48 | 86 |
| | Total | 14.72 | 85.28 | 100 |
| **Benchmark with Pre-processing (B1)** | | | | |
| True | Anomalous | 7.06 | 6.94 | 14 |
| | Normal | 5.62 | 80.38 | 86 |
| | Total | 12.68 | 87.32 | 100 |
| **Benchmark without Pre-processing (B2)** | | | | |
| True | Anomalous | 4.26 | 9.74 | 14 |
| | Normal | 4.86 | 81.14 | 86 |
| | Total | 9.12 | 90.88 | 100 |

**Table 2.** Relevant metrics for the evaluation result.

| Accuracy | Precision | Recall | F1 Score | FPR | FNR |
|---|---|---|---|---|---|
| **Proposed Method** | | | | | |
| 0.9368 | 0.7722 | 0.8000 | 0.7823 | 0.0409 | 0.2000 |
| **Benchmark with Pre-processing (B1)** | | | | | |
| 0.8744 | 0.6156 | 0.5043 | 0.5413 | 0.0653 | 0.4957 |
| **Benchmark without Pre-processing (B2)** | | | | | |
| 0.8540 | 0.6121 | 0.3043 | 0.3651 | 0.0565 | 0.6957 |

The confusion tables describing average false positive (FP) and average false negative (FN) of the proposed framework and benchmarks across 50 iterations are shown in Table 1. Some evaluation metrics including the false positive rate (FPR) defined as $\frac{\text{FP}}{\text{FP+TN}}$, false-negative rate (FNR) defined as $\frac{\text{FN}}{\text{FN+TP}}$, precision defined as $\frac{\text{TP}}{\text{TP+FP}}$, recall defined as $\frac{\text{TP}}{\text{TP+FN}}$, and F1 score defined as the harmonic mean of precision and recall i.e., $2 \cdot \frac{\text{precision} \cdot \text{recall}}{\text{precision} + \text{recall}}$ are shown in Table 2. These values are the average values across 50 iterations. As can be seen from the tables, the FPR and FNR of our proposed method are 4% and 20% respectively, resulting in an accuracy of 93.7%Ṫ ur method's accuracy is 7.2% higher than that of the benchmark with pre-processing (B1) and 9.7% higher than that of the benchmark without pre-processing (B2). The recall of 80% indicates that most of the anomalous images can be successfully detected, while the precision of 77% implies that the false positives are not problematic, both beating their counterparts' metrics for the benchmarks. The low FPR (around

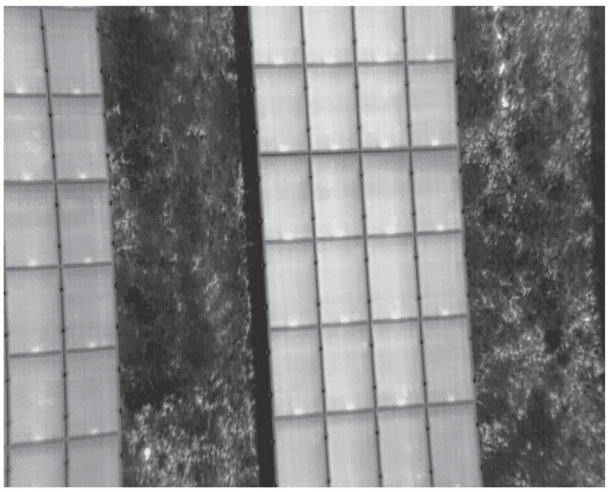

Figure 9. Non-orthogonal perspective.

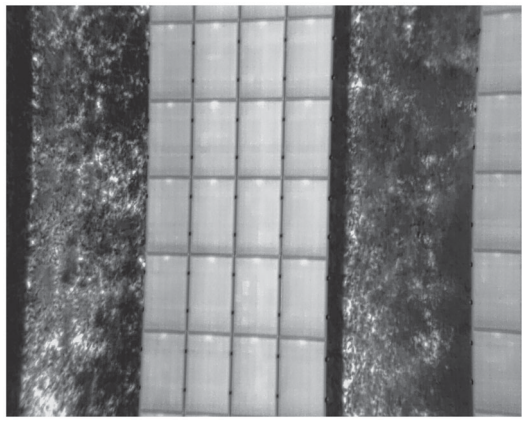

Figure 10. Faint anomaly.

4%) also indicates the algorithm is reliable. F1 score is usually used as an evaluation of the overall performance of classification algorithms. In terms of the F1 score, our method outperforms B1 by 44.5% and B2 by 114.3%.

The run time of our framework per evaluation iteration is about 13.4 minutes on average. This includes the preparation of the training images (orthogonalization and cropping) and the test phase. The average run time of the pre-processing phase is 1.94 s/image, the average run time of the RPCA is 4.35 s/image, and the average run time of the post-processing phase is 0.0346s/image. The average overall run time is about 6.32 s/image, which makes the online automatic detection of the anomaly possible. The running time for B1 is shorter since the Robust PCA step is replaced by much simpler thresholding, and there is no post-processing involved. The thresholding time is negligible, $2.2047 \times 10^{-4}$ s/image, which results in an average run time of 1.97 s/image, faster than our method. The speed of B2 is even faster without pre-processing, resulting in $2.20 \times 10^{-4}$ s/image. However, considering the performance improvement obtained by the pre-processing and RPCA, our method is preferred over both of the benchmarks.

There are three points to be noticed from investigating the FP and FN values, which are also helpful for practitioners in using the method.

- The false positive are likely due to the highly non-orthogonal perspective of the image like Figure 9. The non-orthogonality assumes robust PCA that the PV image follows a low-rank pattern fails. Although the TILT algorithm can fix a slight non-orthogonal perspective, it is helpless in extreme cases. An accurate selection of the transformed area of TILT can improve the result. However, since we aim for an automatic algorithm for anomaly detection, the selection area should be predefined. To solve this problem, a more stable image stream from the drone is expected.
- The case of false-negative is likely due to the faint and unclustered anomaly of the image like Figure 10. However, they can be detected by the RPCA and are removed during the post-processing, which primarily aims to reduce the noise. Note that reducing the strength of the noise reduction step will cause a high false-positive rate.
- The main purpose of the algorithm is to reduce the burden of human inspection. Indeed, the algorithm achieves automated detection, high recall of all anomalies, and full recall of severe anomalies. Also, since the anomaly detection for PV images does not need an instant reaction and may be done only once a day, it provides enough time for a human to identify the small number of false positives.

## 4. Conclusion

In this paper, we successfully overcome two difficulties arising from our real data and built a fully automatic online monitoring system to automate and expedite the process of detecting anomalies. This is done frame by frame from the raw video stream of the thermal camera using a hybrid approach involving image processing and statistical machine learning techniques. The results showed that the algorithm could successfully detect the anomalies with a recall of 0.80 (can detect significant anomalies with a high recall of 1) and the high speed of 6.32 seconds/image. Our method outperformed benchmark 1 by 44.5%

and benchmark 2 by 114.3% in terms of the F1 score. Although there are false alarms, they happen mainly due to the irregular pattern of images at the two ends of a PV module or an extreme non-orthogonal perspective. Since the number of false alarms is not high, they can be identified by human experts offline.

In short, the high accuracy and reasonable computation time of our proposed framework makes it feasible to implement online anomaly detection in practice. Additionally, as our algorithms do not require ortho-images, it would not be necessary to continuously adjust the angle of drones to force them to take ortho-images. This would potentially reduce imaging and inspection costs. Integration of this innovative statistical framework with aerial thermography techniques facilitates the automatic online detection of anomalies in PV systems with minimal human inspection efforts.

To help practitioners implement the proposed framework, the algorithms and codes are provided in the online appendix, that could be applied to PV cell images that present some low-rank structure, as in the present paper. Nevertheless, some parts of the codes, which depend on the specific data in use, should be adapted accordingly. The following guidelines facilitate the implementation of the algorithms.

- To implement perspective transformation, the TILT algorithm requires the practitioners to specify the rectangular region of interest (ROI), where most of the PV cells lie. In practice, it is not desirable to select the whole image as the ROI since the soil background does not belong to the category of low-rank structure, and the performance is thus not the best. Our implementation finds a common ROI where, for most of the images, the PV cells lie (this largely depends on the specific data in use.) In our code function named *MYTRANSFORM*, the third argument of the TILT function specifies the location of the ROI. The first column represents the *xy* coordinates of the top-left corner of ROI, while the second column represents those of the bottom-right corner. The practitioners should adjust these parameters according to the data in use and find the best ROI works for all images. The MATLAB function *IMPIXELINFO* could be used to find the common ROI. This function helps to determine which *xy* coordinates the cursor lies in the image.
- To eliminate background noise during automatic cropping, small connected components were removed before the median filtering was applied. The threshold to remove small connected components depends on the panel size. Current practice consists of removing those with an area smaller than about half of a unit panel, depending on the specific data in use. To achieve the desired performance, also the size of the median filter should be carefully tuned, depending on the specific data set.
- When implementing post-processing, after tuning the size of median filtering, a threshold needs to be determined to remove the remaining noise. This threshold level could be determined by finding some percentile, for example, 95% percentile of pixel intensities of the image, or be carefully tuned to achieve the best performance.

## About the authors

**Qian Wang** is currently a Ph.D. student in Machine Learning at H. Milton Stewart School of Industrial and Systems Engineering (ISyE) at Georgia Tech. His research focuses on the analysis of high-dimensional functional data including profiles, images and point cloud data using statistical machine learning tools for anomaly detection, quality control or causal analysis purposes in manufacturing.

**Kamran Paynabar** is the Fouts Family Early Career Professor and Associate Professor in the H. Milton Stewart School of Industrial and Systems Engineering at Georgia Tech. He received his Ph.D. in IOE and M.A. in Statistics from The University of Michigan. His research interests comprise both applied and methodological aspects of machine-learning and statistical modeling integrated with engineering principles. He is a recipient of the Data Mining Best Student Paper Award, the Best Application Paper Award from IIE Transactions, the Best QSR refereed paper from INFORMS, and the Best Paper Award from POMS. He has been recognized with the GT campus level 2014 Junior Faculty Teaching Excellence Award and Provost Teaching and Learning Fellowship. He served as the chair of QSR of INFORMS, and the president of QCRE of IISE. He is an Associate Editor for *Technometrics* and *IEEE-TASE*, a Department Editor for *IISE-Transactions* and a member of the editorial board for the *Journal of Quality Technology*.

**Massimo Pacella** received the M.Sc. degree in computer engineering from the University of Salento (Italy) in 1998, and the Ph.D. degree in manufacturing and production systems from the Polytechnic University of Milan (Italy) in 2003. He received the Fulbright Fellowship in 2009. Currently, he is an Associate Professor with the Department of Engineering for Innovation, University of Salento. His main research interests are functional data processing, profile monitoring, design of experiments, manufacturing process control, and coordinate metrology, including machine learning techniques and methods of applied statistics. He is a member of the Italian Association for Manufacturing Technology (AITeM).

## Acknowledgement

The authors would like to thank the editor, the editor of the Case Study section, and two anonymous referees for their constructive comments and suggestions that have considerably improved the article. The work of Paynabar was partially supported by National Science Foundation Grants CMMI-1839591. The authors are thankful to PVK S.r.l. industry (Italy) for providing the data set of the case study.

## References

AbdulMawjood, K., S. S. Refaat, and W. G. Morsi. 2018. Detection and prediction of faults in photovoltaic arrays: A review. In 2018 IEEE 12th International Conference on Compatibility, Power Electronics and Power Engineering (CPE-POWERENG 2018), 1–8, doi: 10.1109/CPE.2018.8372609.

Aghaei, M., F. Grimaccia, C. A. Gonano, and S. Leva. 2015. Innovative automated control system for PV fields inspection and remote control. *IEEE Transactions on Industrial Electronics* 62 (11):7287–96. doi: 10.1109/TIE.2015.2475235.

Alsafasfeh, M., I. Abdel-Qader, B. Bazuin, Q. Alsafasfeh, and W. Su. 2018. Unsupervised fault detection and analysis for large photovoltaic systems using drones and machine vision. *Energies* 11 (9):2252. doi: 10.3390/en1109.

Boyd, S., N. Parikh, E. Chu, B. Peleato, and J. Eckstein. 2010. Distributed optimization and statistical learning via the alternating direction method of multipliers. *Foundations and Trends® in Machine Learning* 3 (1): 1–122. doi: 10.1561/2200000016.

Buerhop, C., R. Weißmann, H. Scheuerpflug, R. Auer, and C. Brabec. 2012. Quality control of PV-modules in the field using a remote-controlled drone with an infrared camera. In European Photovoltaic SolarEnergy Conference and Exhibition (EUPVSEC), 3370-3. Frankfurt, Germany: EU PVSEC. doi: 10.4229/27thEUPVSEC2012-4BV.2.43.

Buerhop, C., F. Fecher, T. Pickel, A. Häring, T. Adamski, C. Camus, J. Hauch, and C. Brabec. 2018. Verifying defective PV-modules by IR-imaging and controlling with module optimizers. *Progress in Photovoltaics: Research and Applications* 26 (8):622–30. doi: 10.1002/pip.2985.

Candès, E. J., X. Li, Y. Ma, and J. Wright. 2011. Robust principal component analysis? *Journal of the ACM* 58 (3): 1–37. doi: 10.1145/1970392.1970395.

Colaprico, M., M. De Ruvo, G. Leotta, F. Bizzarri, S. Vergura, and F. Marino. 2018. Dubio: A fully automatic drones cloud based infrared monitoring system for large-scale PV plants. In Proceedings - 2018 IEEE International Conference on Environment and Electrical Engineering *and 2018 IEEE Industrial and Commercial Power Systems Europe, EEEIC/I and CPS Europe 2018*, doi: 10.1109/EEEIC.2018.8493736.

Denio, H. III. 2012. Aerial solar thermography and condition monitoring of photovoltaic systems. In *38th IEEE Photovoltaic Specialists Conference*, 613–618. doi: 10.1109/PVSC.2012. 6317686.

Dhibi, K., R. Fezai, M. Mansouri, M. Trabelsi, A. Kouadri, K. Bouzara, H. Nounou, and M. Nounou. 2020. Reduced kernel random forest technique for fault detection and classification in grid-tied pv systems. *IEEE Journal of Photovoltaics* 10 (6):1864–71. 3011068. doi: 10.1109/JPHOTOV.2020.

Dhimish, M., and V. Holmes. 2016. Fault detection algorithm for grid-connected photovoltaic plants. *Solar Energy* 137:236–45. doi: 10.1016/j.solener.2016.08.021.

Fazai, R., K. Abodayeh, M. Mansouri, M. Trabelsi, H. Nounou, M. Nounou, and G. Georghiou. 2019. Machine learning-based statistical testing hypothesis for fault detection in photovoltaic systems. *Solar Energy* 190: 405–13. URL http://www.sciencedirect.com/science/article/pii/S0038092X19308126. doi: 10.1016/j.solener.2019.08.032.

Garoudja, E., F. Harrou, Y. Sun, K. Kara, A. Chouder, and S. Silvestre. 2017. Statistical fault detection in photovoltaic systems. *Solar Energy* 150:485–99. doi: 10.1016/j.solener.2017.04.043.

Grimaccia, F., M. Aghaei, M. Mussetta, S. Leva, and P. Quater. 2015. Planning for PV plant performance monitoring by means of unmanned aerial systems (UAS). *International Journal of Energy and Environmental Engineering* 6 (1):47–54. doi: 10.1007/s40095-014-0149-6.

Grimaccia, F., S. Leva, A. Dolara, and M. Aghaei. 2017. Survey on PV modules' common faults after an O&M flight extensive campaign over different plants in Italy. *IEEE Journal of Photovoltaics* 7 (3):810–6. doi: 10.1109/JPHOTOV.2017.2674977.

Hajji, M., M.-F. Harkat, A. Kouadri, K. Abodayeh, M. Mansouri, H. Nounou, and M. Nounou. 2021. Multivariate feature extraction based supervised machine learning for fault detection and diagnosis in photovoltaic systems. *European Journal of Control* 59:313–3580. http://www.sciencedirect.com/science/article/pii/S0947358019304054. doi: 10.1016/j.ejcon.2020.03.004.

Harrou, F., Y. Sun, B. Taghezouit, A. Saidi, and M.-E. Hamlati. 2018. Reliable fault detection and diagnosis of photovoltaic systems based on statistical monitoring approaches. *Renewable Energy*. 116:22–37. doi: 10.1016/j.renene.2017.09.048.

Li, X., W. Li, Q. Yang, W. Yan, and A. Zomaya. 2020. An unmanned inspection system for multiple defects detection in photovoltaic plants. *IEEE Journal of Photovoltaics* 10 (2):568–76. doi: 10.1109/JPHOTOV.2019.2955183.

Libra, M., M. Daneček, J. Lešetický, V. Poulek, J. Sedláček, and V. Beránek. 2019. Monitoring of defects of a photovoltaic power plant using a drone. *Energies* 12 (5):795. doi: 10.3390/en12050.

Lorenzo, G., R. Araneo, M. Mitolo, A. Niccolai, and F. Grimaccia. 2020. Review of O&M practices in PV plants: Failures, solutions, remote control, and monitoring tools. *IEEE Journal of Photovoltaics* 10 (4):914–26. doi: 10.1109/JPHOTOV.2020.2994531.

Madeti, S. R., and S. Singh. 2017. Monitoring system for photovoltaic plants: A review. *Renewable and Sustainable Energy Reviews* 67:1180–207. http://www.sciencedirect.com/science/article/pii/S1364032116305792. doi: 10.1016/j.rser.2016.09.088.

Mansouri, M.,. A. Al-Khazraji, M. Hajji, M. Harkat, H. Nounou, and M. Nounou. 2018. Wavelet optimized

EWMA for fault detection and application to photovoltaic systems. *Solar Energy* 167:125–36. doi: 10.1016/j.solener.2018.03.073.

Otsu, N. 1979. A threshold selection method from gray-level histograms. *IEEE Transactions on Systems, Man, and Cybernetics* 9 (1):62–6. doi: 10.1109/TSMC.1979.4310076.

Rahaman, S., T. Urmee, and D. Parlevliet. 2020. PV system defects identification using remotely piloted aircraft (RPA) based infrared (IR) imaging: A review. *Solar Energy* 206:579–95. doi: 10.1016/j.solener.2020.06.014.

Teubner, J., I. Kruse, H. Scheuerpflug, C. Buerhop-Lutz, J. Hauch, C. Camus, and C. J. Brabec. 2017. Comparison of drone-based IR-imaging with module resolved monitoring power data. *Energy Procedia* 124:560–6. http://www.sciencedirect.com/science/article/pii/S1876610217339383. doi: 10.1016/j.egypro.2017.09.094.

Tsanakas, J. A., L. Ha, and C. Buerhop. 2016. Faults and infrared thermographic diagnosis in operating c-Si photovoltaic modules: A review of research and future challenges. *Renewable and Sustainable Energy Reviews* 62: 695–709. http://www.sciencedirect.com/science/article/pii/S1364032116301629. doi: 10.1016/j.rser.2016.04.079.

Tsanakas, J., L. Ha, and F. A. Shakarchi. 2017. Advanced inspection of photovoltaic installations by aerial triangulation and terrestrial georeferencing of thermal/visual imagery. *Renewable Energy*. 102:224–33. doi: 10.1016/j.renene.2016.10.046.

Ventura, C., and G. Tina. 2015. Development of models for on-line diagnostic and energy assessment analysis of PV power plants: The study case of 1 MW sicilian PV plant. *Energy Procedia* 83:248–57. doi: 10.1016/j.egypro.2015.12.179.

Zhang, Z., A. Ganesh, X. Liang, and Y. Ma. 2012. TILT: Transform invariant low-rank textures. *International Journal of Computer Vision* 99 (1):1–24. doi: 10.1007/s11263-012-0515-x.

Zhao, Y., F. Balboni, T. Arnaud, J. Mosesian, and B. M. Ball, and R. Lehman. 2014. Fault experiments in a commercial-scale PV laboratory and fault detection using local outlier factor. In *IEEE 40th Photovoltaic Specialist Conference, PVSC 2014*, 3398–3403. doi: 10.1109/PVSC.2014.6925661.