

**SUPSI**

# End-to-End Pipeline for Thermal Anomaly Detection in PV Plants via Unmanned Aerial Vehicle (UAV) Thermography

---

Student/s

**Rava Miro**

Supervisor

**Giusti Alessandro**

---

Co-supervisor

**Abbate Gabriele**

---

Host

-

---

Course

**Data Science and Artificial Intelligence**

Module

**Diploma Thesis**

---

Year

**2025**

---

Date

**August 22, 2025**



*This thesis is dedicated to my parents, my girlfriend Arianna, and all of my friends. Thank you for your continuous support and for your patience during the countless times I had to decline an evening out for a cold beer to work on this project.*

*I would also like to extend my gratitude to my classmates and professors. These past three years of my bachelor's degree would not have been nearly as enjoyable or enriching without your camaraderie and guidance*



# Contents

<b>1</b>	<b>Introduction</b>	<b>3</b>
<b>2</b>	<b>The Problem Domain: PV Inspection, Defectology and Thermography</b>	<b>5</b>
2.1	Importance of Reliability in Photovoltaic Systems . . . . .	5
2.2	Impact of Performance Loss in PV Plants . . . . .	5
2.2.1	Economic Implications of Underperformance and Degradation . . . . .	6
2.2.2	Environmental and Life Cycle Considerations . . . . .	6
2.3	Taxonomy of Photovoltaic Defects and Failure Modes . . . . .	7
2.3.1	Cell-Level Defects and Anomalies . . . . .	7
2.3.1.1	Hotspots . . . . .	7
2.3.1.2	Micro-cracks . . . . .	7
2.3.1.3	Snail Trails . . . . .	8
2.3.1.4	Potential-Induced Degradation (PID) . . . . .	8
2.3.1.5	Interconnect Failures . . . . .	9
2.3.2	Module-Level Degradation and Failures . . . . .	9
2.3.2.1	Encapsulant Delamination and Bubbles . . . . .	9
2.3.2.2	Encapsulant Discoloration (Browning/Yellowing) . . . . .	9
2.3.2.3	Backsheet Degradation . . . . .	10
2.3.2.4	Glass Breakage . . . . .	10
2.3.3	Balance-of-System (BOS) and Component Faults . . . . .	10
2.3.3.1	Bypass Diode Failure . . . . .	10
2.3.3.2	Junction Box and Connector Faults . . . . .	11
2.3.3.3	String and Inverter-Level Anomalies . . . . .	11
2.4	Thermography Standards and Data Availability . . . . .	11
2.4.1	The IEC 62446-3 Standard for Outdoor Infrared Thermography . . . . .	11
2.4.2	Data Acquisition: Manual vs Autonomous Aerial Systems . . . . .	12
2.4.3	Analysis of Public Datasets for Model Development . . . . .	12
2.5	Research Opportunities . . . . .	14
<b>3</b>	<b>State of the Art in Automated PV Plant Analysis</b>	<b>15</b>
3.1	Introduction . . . . .	15

3.2	Computer Vision Approaches for Panel Detection and Localization . . . . .	15
3.2.1	Traditional Image Processing Techniques . . . . .	16
3.2.2	Deep Learning-Based Object Detection . . . . .	17
3.3	Machine Learning for Thermal Anomaly Classification . . . . .	17
3.3.1	Supervised Learning Techniques . . . . .	17
3.3.2	Unsupervised Learning Techniques . . . . .	18
3.4	Review of Existing Systems and Solutions . . . . .	18
3.4.1	Academic and Open-Source Implementations . . . . .	18
3.4.2	Commercial Cloud-Based Platforms . . . . .	19
3.4.2.1	Criticalities of Commercial Solutions . . . . .	20
3.5	Gap Analysis and SoTA Limitations . . . . .	21
<b>4</b>	<b>Software Interface and Methods</b>	<b>23</b>
4.1	Orthophoto-Centric Workflow Rationale . . . . .	23
4.2	System Architecture and Data Flow . . . . .	24
4.3	The Graphical User Interface (GUI) . . . . .	25
4.4	Module 1: Input and Alignment . . . . .	26
4.4.1	Data Ingestion . . . . .	26
4.4.2	Manual Image Alignment . . . . .	26
4.5	Module 2: The Panel Detection Pipeline . . . . .	27
4.5.1	Core Detection via Normalized Cross-Correlation . . . . .	28
4.5.2	Advanced Detection Refinement . . . . .	29
4.6	Module 3: Segmentation and Feature Extraction . . . . .	30
4.7	Module 4: Machine Learning Model Training . . . . .	33
4.8	Module 5: Classification and Analysis . . . . .	33
4.9	Module 6: Outputs and Reporting . . . . .	34
<b>5</b>	<b>Experiment and Results</b>	<b>37</b>
5.1	Experimental Setup . . . . .	37
5.1.1	Dataset and Hardware . . . . .	37
5.1.2	System Performance and Scalability . . . . .	38
5.2	Evaluation Metrics . . . . .	38
5.3	Results of Panel Detection . . . . .	39
5.4	Results of Anomaly Classification . . . . .	40
5.4.1	Methodology and Ground Truth Creation . . . . .	40
5.4.2	Model Performance . . . . .	41
5.5	Synthesis of Experimental Findings . . . . .	42

---

<b>6 Discussion</b>	<b>45</b>
6.1 Interpretation of System Performance . . . . .	45
6.2 Innovations and Contribution . . . . .	45
6.3 Limitations and Future Directions . . . . .	46
6.4 Comparison with the State of the Art . . . . .	46
<b>7 Conclusions</b>	<b>49</b>



# List of Figures

2.1	Example of a hotspots defects from soiling in the bottom corners of solar panels	7
2.2	Potential-Induced Degradation (PID). (Top) Disposition of the panels in the string. (Bottom) Infrared thermogram showing the characteristic "checker-board" pattern of moderately heated cells near the negative side of the string.	8
2.3	Encapsulant discoloration. (Left) Visual image showing browning of the encapsulant over a cell. (Right) Thermogram confirming that the discolored cell is operating as a hotspot. . . . .	9
2.4	Bypass diode failure. The thermogram shows one-third of the module uniformly heated, a classic signature of a short-circuited or active bypass diode. .	10
2.5	Comparison of thermographic inspection practices. (Right) An image acquired under standardized conditions (high irradiance, no clouds, near-perpendicular view). (Left) A poor quality image compromised by bad camera angle and sun glares, rendering it unusable for diagnostics. . . . .	12
2.6	Contrast between data types. (Left) A true radiometric thermal image where pixel values correspond to temperature, enabling quantitative analysis. (Right) A typical 8-bit JPEG from a public dataset, where colors are arbitrary and prevent robust, physics-based analysis. . . . .	13
3.1	How template matching works, a template image is taken and in our case is moved across the target image. The points where the match was higher are identified as having a match. . . . .	16
3.2	A screenshot of the GUI of the open-source PV Hawk viewer . . . . .	19
3.3	A sample of the available commercial cloud-based SaaS solutions: Drone Deploy (Left), SiteMark (Center) and RaptorMaps (Right) . . . . .	20
4.1	The critical impact of reconstruction quality on data fidelity. A poor reconstruction (left) exhibits significant sun glare and artifacts, while a high-quality reconstruction (right) provides clean data suitable for analysis. . . . .	24
4.2	A high-level diagram illustrating the system's architecture and the logical flow of data between its core processing modules. . . . .	25

4.3	The main welcome screen of the Solar PV Anomaly Detection Suite, outlining the six primary workflow sections accessible via the left-hand navigation panel.	26
4.4	A projective transformation: placing four pair of points is able to unequivocally define a transformation for the entire image. . . . .	27
4.5	The user interface for the Alignment module. The operator places corresponding control points (blue markers) on the RGB (left) and thermal (right) orthophotos to enable a precise affine transformation. . . . .	27
4.6	A special implementation of the NMS algorithm was used since the images are very large, for a small plant more than 2 million panels where detected and doing an IoU check for all of them is exponentially slower. . . . .	28
4.7	The UI for the core detection step. The operator defines a grid to create templates and adjusts parameters for template matching and Non-Maximum Suppression (NMS). The initial raw detections are shown as small green boxes.	29
4.8	Example of thermal (left) and RGB (right) panel templates automatically extracted by the system for use in the NCC algorithm. . . . .	29
4.9	The various algorithms removing unwanted detections (Top). The UI for the advanced detection refinement pipeline, showing the controls for DBSCAN/Grid Fitting and Border Outlier Removal. The map displays the refined panel detections. (Bottom) . . . . .	30
4.10	The interactive analysis dashboard within the Segmentation module. The interface displays thematically colored panels, a plant-wide thermal distribution histogram with user-controlled thresholds, and a detailed information card for the selected panel. . . . .	31
4.11	A detailed view of a single panel pair, showing the radiometric thermal image with its temperature scale and the corresponding high-resolution RGB image. In this case the Hot-Spot is caused by a physical obstruction like soiling or bird deposits. . . . .	32
4.12	The user interface for the Machine Learning Models module. This view shows the results of training and testing a Binary Random Forest classifier, including performance metrics and a confusion matrix. . . . .	33
4.13	The Classification module interface. The map visualizes the model's predictions, with a detailed pop-up showing the predicted class ('Hot Spot') for the selected panel, alongside its thermal data and detailed classification results on the right. . . . .	34
4.14	The Outputs section, presenting a final, exportable report of all detected anomalies with their type, confidence level, and location. . . . .	35

## List of Tables

5.1	Summary of the aggregated dataset used for system evaluation. . . . .	38
5.2	Representative performance breakdown for a 6,857-panel PV plant. . . . .	38
5.3	Panel detection performance across all 13 PV plants, highlighting variation between installation types. . . . .	40
5.4	Model performance on the <b>training set</b> . . . . .	41
5.5	Model performance on the unseen <b>test set</b> . . . . .	42
5.6	Representative confusion matrix for the XGBoost multiclass classifier on a balanced test set of 3,930 panels. . . . .	42



### Abstract

Thermography for photovoltaic (PV) plants is hindered by a critical trade-off: operators must choose between opaque, proprietary cloud platforms that compromise data sovereignty and fragmented, impractical open-source tools. This thesis resolves this dilemma by introducing the *PV Anomaly Detection Suite*, a novel, end-to-end software solution architected for local, privacy-preserving operation.

The system leverages a robust orthomosaic-centric workflow, combining computationally efficient computer vision for panel detection with machine learning for radiometric anomaly classification on standard hardware. Validated across 13 diverse PV plants, the suite achieved a panel detection rate exceeding 99.42% and demonstrated robust defect classification. The primary contribution is a practical, user-friendly tool that provides operators with immediate, on-site analytics, bridging the gap between commercial services and academic projects while empowering future research through a powerful dataset creation framework.

### Abstract

La termografia per gli impianti fotovoltaici (FV) è ostacolata da un compromesso critico: gli operatori devono scegliere tra piattaforme cloud proprietarie e opache, che compromettono la sovranità dei dati, e strumenti open-source frammentati e poco pratici. Questa tesi risolve tale dilemma introducendo la *PV Anomaly Detection Suite*: una nuova soluzione software end-to-end, progettata per un'operatività locale e a tutela della privacy.

Il sistema si basa su un robusto flusso di lavoro incentrato su ortomosaici, combinando algoritmi di computer vision computazionalmente efficienti per il rilevamento dei pannelli con il machine learning per la classificazione di anomalie radiometriche su hardware standard. Validata su 13 diversi impianti FV, la suite ha raggiunto un tasso di rilevamento dei pannelli superiore al 99.42% e ha dimostrato una classificazione dei difetti robusta. Il contributo principale è uno strumento pratico e di facile utilizzo che fornisce agli operatori analisi immediate e in loco, colmando il divario tra i servizi commerciali e i progetti accademici e, al contempo, potenziando la ricerca futura tramite un potente framework per la creazione di dataset.



# Chapter 1

## Introduction

The proliferation of utility-scale photovoltaic (PV) plants, has introduced a critical challenge: ensuring decades-long operational reliability. The economic viability and environmental sustainability of these assets are directly contingent on maximizing their energy yield and operational lifespan, which necessitates advanced methods for Operations and Maintenance (O&M).

Unmanned Aerial Vehicle (UAV) thermography has become the industry standard for rapid, large-scale PV inspection, providing the raw data needed to identify performance-inhibiting thermal anomalies. This technological advance, however, has shifted the bottleneck from data acquisition to data analysis. The sheer volume of imagery generated makes manual inspection untenable, creating an urgent need for automated, algorithm-driven diagnostic systems.

The current ecosystem of automated solutions presents operators with no choice. Commercial, cloud-based platforms offer polished user experiences but operate as "black boxes," requiring the surrender of sensitive operational data to third-party servers and introducing concerns of privacy and vendor lock-in. Conversely, the existing open-source alternatives are often fragmented, lack robust end-to-end functionality, and remain at the level of academic proof-of-concept. This creates a significant research and utility gap for a tool that is at once powerful, private, transparent, and practical for field deployment.

This thesis bridges this gap by presenting the design, implementation, and validation of a novel, end-to-end *PV Anomaly Detection Suite*. It delivers a complete, locally-executable software solution that provides operators with powerful analytical capabilities without compromising data sovereignty. By integrating efficient computer vision and machine learning within a user-friendly, orthomosaic-centric workflow, this work provides a practical alternative to existing solutions. This dissertation will detail the problem domain, review the state of the art, present the system's methodology and experimental results, and conclude by discussing its possible future contributions to the field of automated PV diagnostics.



## Chapter 2

# The Problem Domain: PV Inspection, Defectology and Thermography

### 2.1 Importance of Reliability in Photovoltaic Systems

The installed PV capacity has grown exponentially in the last few years, surpassing 2 terawatts (TWp)<sup>1</sup>. This expansion highlights the critical importance of ensuring the long-term reliability, performance, and safety of these assets<sup>2</sup>. The economic viability and environmental promise of solar energy are intrinsically linked to the operational integrity of PV systems. Financial models depend on predictable and uninterrupted energy generation, while the technology's green credentials hinge on a long operational life to compensate the initial environmental investment<sup>3:4</sup>. Defects and premature failures pose a systemic threat to this paradigm<sup>5</sup>.

This chapter presents a comprehensive review of these threats, focusing on the thermal anomalies that serve as their primary indicators<sup>6</sup>. It aims to establish the theoretical foundation for advanced diagnostic methodologies by first quantifying the economic and environmental impacts of performance loss<sup>7</sup>. Subsequently, a detailed taxonomy of PV defects is provided, outlining their physical origins and characteristic thermal signatures<sup>8</sup>. The principles of thermographic inspection, as governed by international standards, highlighting the crucial limitations of existing public datasets that hinder the development of robust artificial intelligence models<sup>9:10</sup>.

### 2.2 Impact of Performance Loss in PV Plants

The significant consequences of PV system under-performance, which affect financial stability, environmental lifecycles, and operational integrity, highlights the need for robust defect detection<sup>5</sup>.

### 2.2.1 Economic Implications of Underperformance and Degradation

The PV global industry lost from underperforming assets an estimated \$4.6 billion in 2023 and this number is projected to approach \$15 billion in 2025<sup>11</sup>. The average annual revenue loss per megawatt (MW) has surged from \$977/MW in 2020 to \$5,720/MW in 2024, indicating a worsening financial impact per unit of capacity<sup>12</sup>. This underperformance directly erodes project profitability; a 100 MW site with a 5.77% underperformance rate could see a 249 basis point reduction in its internal rate of return (IRR)<sup>12</sup>.

The Levelized Cost of Energy (LCOE) is highly sensitive to degradation rates<sup>13</sup>. While financial models often assume a 0.5% annual degradation, field data suggests a median system-level rate of 1.3%/year<sup>3</sup>. An increase in the degradation rate from 0.5% to 2.5% can elevate the LCOE by over 30% in favorable economic conditions and by 103.2% in less favorable scenarios<sup>13;14</sup>.

Losses are distributed throughout the system, with inverters (2.13%), string-level issues (1.26%), and combiner boxes (1.04%) being the largest contributors, challenging the perception that modules are the sole concern<sup>12</sup>. In the U.S., inverters are responsible for 39% of total power loss<sup>12</sup>. This problem is exacerbated by a widening "O&M Gap": for example from 2018 to 2023, U.S. solar capacity grew by 182%, while O&M employment grew by only 91%<sup>12</sup>. This structural imbalance highlights the unscalability of manual maintenance and creates a powerful economic need for automated diagnostic technologies<sup>15</sup>.

### 2.2.2 Environmental and Life Cycle Considerations

The environmental benefits of PV technology depend on a long and productive operational life<sup>4</sup>. The manufacturing process is energy- and resource-intensive, requiring materials like high-purity silicon and involving hazardous chemicals such as lead<sup>4;16</sup>. Consequently, a solar panel has an energy payback time (EPBT) of 2-3 years, the period it must operate to offset the energy consumed during its production<sup>4</sup>. Defects that cause premature failure may prevent a panel from repaying this environmental "debt"<sup>17</sup>.

This issue is more evident at the end-of-life (EOL) stage, with the PV waste stream projected to reach 78 million tons by 2050<sup>4</sup>. Current recycling infrastructure is inadequate, and most EOL panels end up in landfills, where hazardous materials can leach into the environment<sup>18</sup>. From a Life Cycle Assessment (LCA) perspective, module reliability is an imperative, as the "use" phase is where the initial environmental impacts are reduced the most<sup>19;20</sup>. Defects shorten this phase, increasing the relative environmental burden of manufacturing and EOL management<sup>21</sup>. Therefore, asset longevity is a fundamental prerequisite for solar energy to be a truly sustainable solution<sup>4</sup>.

## 2.3 Taxonomy of Photovoltaic Defects and Failure Modes

A good understanding of failure modes is essential for developing effective diagnostic tools<sup>6</sup>. This section provides a detailed taxonomy of PV defects, from the cell to the system level, outlining their physical origins, impacts, and characteristic visual and thermal signatures<sup>8</sup>.

### 2.3.1 Cell-Level Defects and Anomalies

#### 2.3.1.1 Hotspots

A hotspot is a localized area of a module operating at a significantly elevated temperature<sup>22</sup>. It occurs when current from healthy cells is forced through a damaged or shaded cell, causing it to become reverse-biased and dissipate power as heat<sup>23</sup>. Common causes include micro-cracks, partial shading, soiling and manufacturing defects<sup>8</sup>. Hotspots are a severe failure mode, capable of causing permanent burn marks, glass breakage, and fire risk<sup>8</sup>. Thermally, a hotspot appears as a well-defined square area of high temperature; a  $\Delta T$  greater than 15 °C above adjacent cells is considered a serious defect<sup>24;25</sup>.

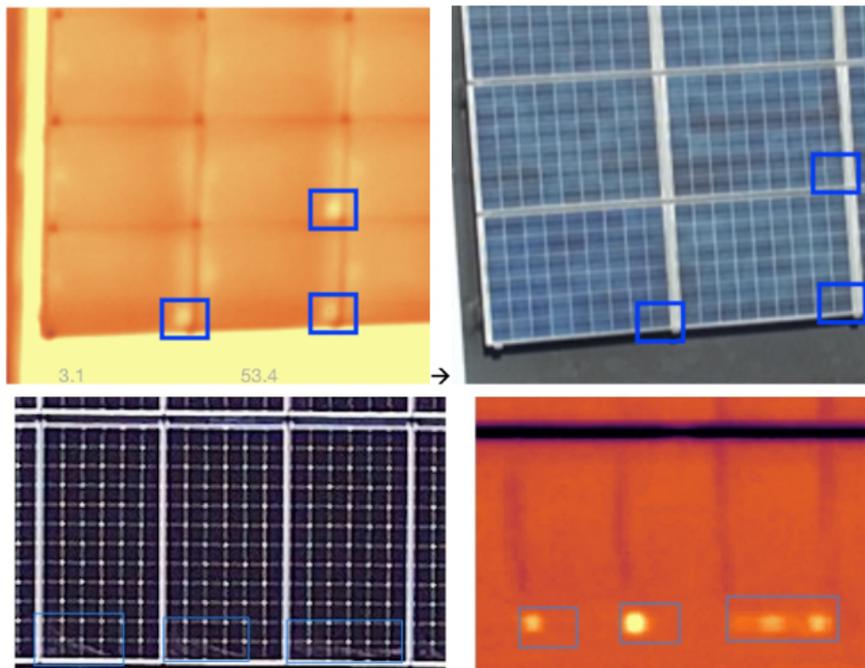


Figure 2.1: Example of a hotspots defects from soiling in the bottom corners of solar panels

#### 2.3.1.2 Micro-cracks

Micro-cracks are microscopic fractures in the silicon wafer caused by mechanical or thermomechanical stress during manufacturing, transport, or operation (e.g., from heavy snow

or wind)<sup>8;26</sup>. While some cracks have a negligible effect, those that electrically isolate a portion of the cell cause direct power loss and can lead to severe hotspot formation<sup>8;27</sup>. They are typically invisible to the naked eye but can be detected using electroluminescence (EL) imaging, where they appear as dark, non-emissive lines<sup>28;29</sup>. Their thermal signature is often indirect, manifesting as the hotspot they induce<sup>24;30</sup>.

### 2.3.1.3 Snail Trails

Snail trails are a form of discoloration on the cell surface resulting from a chemical reaction between moisture, the silver paste of the grid fingers, and the encapsulant<sup>31;32</sup>. Moisture ingress often occurs through pre-existing micro-cracks<sup>33</sup>. While the direct power loss from snail trails is minor (<2%), their presence is a strong indicator of underlying micro-cracks, which can cause significant power loss and hotspots<sup>34;35</sup>. Visually, they appear as thin, dark lines following the cell's grid pattern<sup>28</sup>. Thermally, the signature is that of the underlying defect, often a hotspot<sup>36;37</sup>.

### 2.3.1.4 Potential-Induced Degradation (PID)

PID is caused by leakage currents flowing from the cells to the grounded module frame, driven by a high voltage potential<sup>8</sup>. It is exacerbated by high temperature and humidity and can lead to severe power loss, in some cases exceeding 50%<sup>8;38</sup>. While invisible in its early stages, PID exhibits a highly characteristic thermal pattern: a "checkerboard" or patchy arrangement of heated cells, typically starting at the modules located at the electrically negative end of a string of panels in series<sup>24;28</sup>.

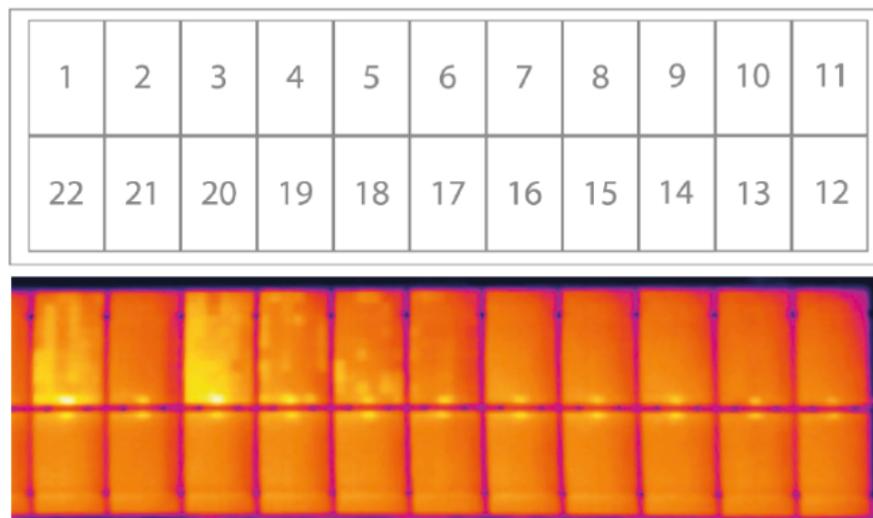


Figure 2.2: Potential-Induced Degradation (PID). (Top) Disposition of the panels in the string. (Bottom) Infrared thermogram showing the characteristic "checkerboard" pattern of moderately heated cells near the negative side of the string.

### 2.3.1.5 Interconnect Failures

Failures in the solder joints connecting interconnect ribbons to cell busbars can result from manufacturing defects or thermomechanical fatigue<sup>8</sup>. A disconnected ribbon creates an open circuit, causing a significant drop in power output ( 35% for one failed ribbon)<sup>8</sup>. An incomplete break can create a point of high resistance, leading to arcing, intense heat, and a high fire risk<sup>38</sup>. The thermal signature is typically a very localized and intense hotspot at the point of the failed connection<sup>28</sup>.

## 2.3.2 Module-Level Degradation and Failures

### 2.3.2.1 Encapsulant Delamination and Bubbles

Delamination is the loss of adhesion between module layers (e.g., glass-EVA or EVA-cell), often initiated by manufacturing defects and exacerbated by moisture and heat<sup>23;38</sup>. It reduces light transmission and creates pathways for moisture ingress, leading to corrosion<sup>8;39</sup>. Visually, it appears as "milky" or cloudy patches<sup>23</sup>. Thermally, delaminated areas appear as irregularly shaped warmer regions due to the insulating effect of trapped air impeding heat dissipation<sup>23</sup>.

### 2.3.2.2 Encapsulant Discoloration (Browning/Yellowing)

Long-term exposure to UV radiation and high temperatures can chemically degrade the EVA encapsulant, causing it to turn yellow or brown<sup>8;40</sup>. This discoloration reduces transparency, leading to a gradual reduction in power output<sup>41</sup>. Localized discoloration is a strong indicator of an underlying hotspot, as the higher temperature accelerates the chemical degradation of the encapsulant above it<sup>42</sup>.

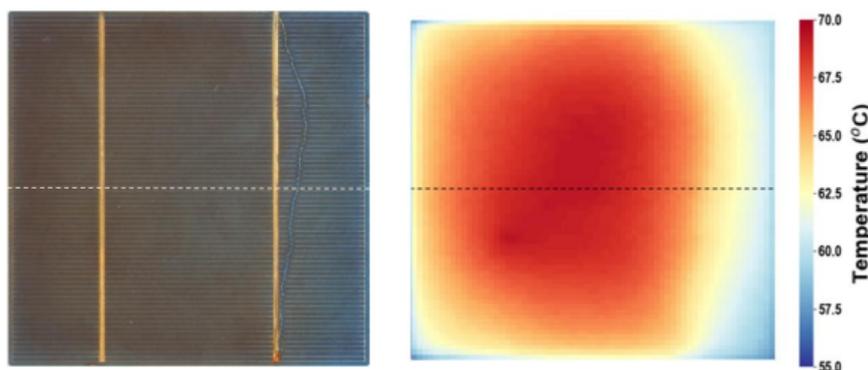


Figure 2.3: Encapsulant discoloration. (Left) Visual image showing browning of the encapsulant over a cell. (Right) Thermogram confirming that the discolored cell is operating as a hotspot.

### 2.3.2.3 Backsheet Degradation

The backsheet provides mechanical protection and electrical insulation<sup>8</sup>. Degradation from UV, temperature, and humidity can cause cracking, delamination, or chalking<sup>43</sup>. This is a critical failure, as it compromises the module's primary insulation, creating a severe electric shock hazard and allowing moisture ingress that can lead to ground faults and string outages<sup>44;45</sup>. While it has no direct thermal signature detectable from the front, its secondary effects (hotspots, hot strings) are visible<sup>28</sup>.

### 2.3.2.4 Glass Breakage

Module glass can break from external impacts (e.g., hail) or manufacturing flaws<sup>46;47</sup>. This results in a complete loss of environmental protection, leading to rapid corrosion, short circuits, and a significant safety hazard<sup>8</sup>.

## 2.3.3 Balance-of-System (BOS) and Component Faults

### 2.3.3.1 Bypass Diode Failure

Bypass diodes protect cell substrings from hotspot formation when shaded<sup>24</sup>. A short-circuited diode causes a fixed power loss (typically 33% or 66%) and produces a clear thermal signature: the entire substring appears as a uniformly heated rectangle<sup>24</sup>.

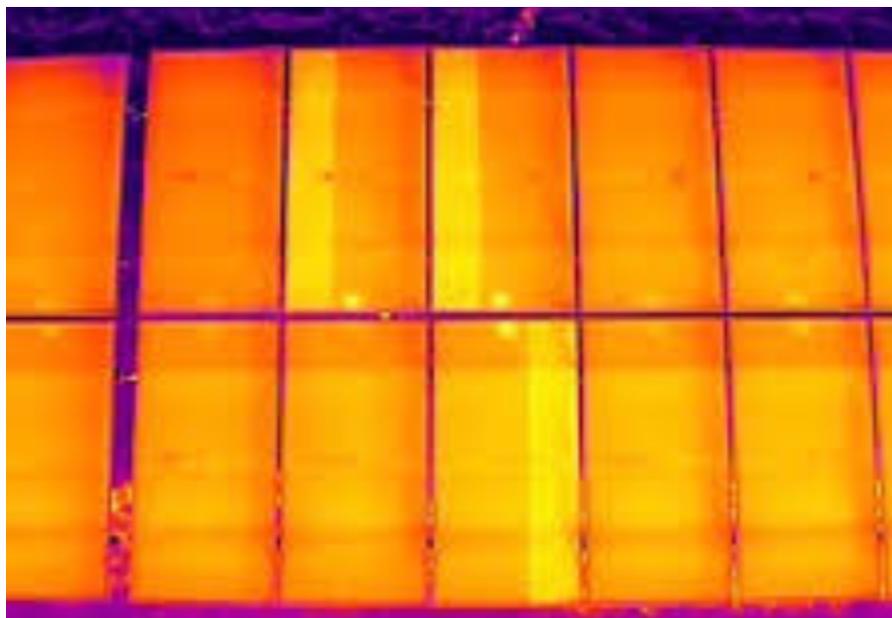


Figure 2.4: Bypass diode failure. The thermogram shows one-third of the module uniformly heated, a classic signature of a short-circuited or active bypass diode.

### 2.3.3.2 Junction Box and Connector Faults

High-resistance connections in junction boxes or module connectors, caused by poor installation or corrosion, dissipate energy as heat<sup>8</sup>. This can lead to significant power loss and represents a major fire hazard<sup>38</sup>. The thermal signature is a distinct, intense, and highly localized hotspot centered on the faulty component<sup>25</sup>.

### 2.3.3.3 String and Inverter-Level Anomalies

An entire string of modules can go offline due to a blown fuse, disconnected cable, or inverter malfunction<sup>5</sup>. This results in a total loss of power from that string<sup>5</sup>. The thermal signature is unambiguous: the entire string of modules appears uniformly hot, as the sun's energy is converted in heat instead of electricity, creating a stark contrast with adjacent, operational strings<sup>24</sup>.

## 2.4 Thermography Standards and Data Availability

Infrared thermography is the preeminent non-destructive technique for inspecting PV plants, as nearly all failure modes alter a module's thermal behavior<sup>48;49</sup>. However, acquiring reliable data requires adherence to rigorous standards<sup>9</sup>.

### 2.4.1 The IEC 62446-3 Standard for Outdoor Infrared Thermography

To ensure data is valid and comparable, inspections must follow the IEC 62446-3 standard<sup>50</sup>. Key requirements include:

- **Solar Irradiance:** A minimum of 600 W/m<sup>2</sup> is required to ensure sufficient current flow to reveal thermal anomalies<sup>51</sup>.
- **Environmental Conditions:** Inspections need clear sky conditions, as passing clouds create thermal transients. High wind speeds (> 10 m/s) should be avoided as they can cool the module surface and mask defects<sup>51</sup>.
- **Viewing Angle and Reflections:** The camera should be positioned as close to perpendicular to the module surface as possible to avoid emissivity errors and, crucially, to prevent solar reflections from obscuring the thermal data<sup>25</sup>.

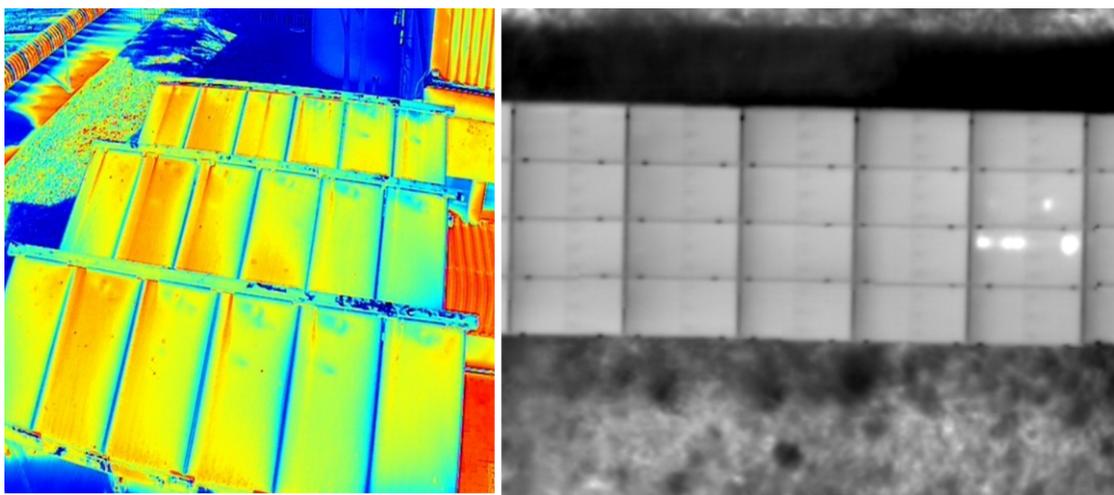


Figure 2.5: Comparison of thermographic inspection practices. (Right) An image acquired under standardized conditions (high irradiance, no clouds, near-perpendicular view). (Left) A poor quality image compromised by bad camera angle and sun glares, rendering it unusable for diagnostics.

#### 2.4.2 Data Acquisition: Manual vs Autonomous Aerial Systems

Manual, handheld inspections are time-consuming, costly, and pose safety risks, often leading to incomplete sampling of large plants<sup>15;52</sup>. Unmanned Aerial Vehicles (UAVs) equipped with radiometric thermal cameras have revolutionized PV inspection<sup>53</sup>. Drones offer high speed (10 minutes/MW vs. 2-5 hours/MW for manual), 100% module coverage, highly repeatable data quality ideal for repeated analysis in time, and significantly enhanced safety<sup>53-56</sup>. This automated, standards-compliant data acquisition is an essential prerequisite for robust AI-based analysis<sup>57;58</sup>.

While drone-based thermography solves data acquisition, it creates a data analysis bottleneck, creating the need for research into automated pipelines using machine learning<sup>59</sup>. However, the progress of these researches is critically hampered by the quality of public datasets<sup>10;60</sup>.

#### 2.4.3 Analysis of Public Datasets for Model Development

Robust and generalizable AI models depend on large, diverse, high-quality public datasets, which represents a major bottleneck in the PV diagnostics field<sup>61-63</sup>. A review of existing public datasets reveals several pervasive limitations:

- **Lack of Radiometric Data:** Most public datasets consist of 8-bit JPEG or PNG images where pixel values represent a color map, not temperature<sup>64-68</sup>. A model trained on these learns color patterns, not physical temperature patterns, making it extremely

brittle and unable to perform quantitative analysis<sup>69</sup>. The scarcity of public datasets containing true radiometric temperature data is a major impediment to progress<sup>70</sup>.

- **Lack of Standardization:** Many datasets are collected without adherence to the IEC 62446-3 standard, introducing significant noise from variable environmental conditions<sup>71;72</sup>. Models trained on such inconsistent data learn spurious correlations and fail to generalize to real-world field conditions<sup>73;74</sup>.
- **Poor Annotation and Imbalance:** Datasets often suffer from inaccurate labeling, a limited number of defect classes, and severe class imbalance, making it difficult to train models to identify rare but critical failure modes<sup>60;62</sup>.
- **Lack of Diversity:** Existing datasets are often collected from a small number of sites with similar technologies and climates, limiting the generalizability of models trained on them<sup>75;76</sup>.

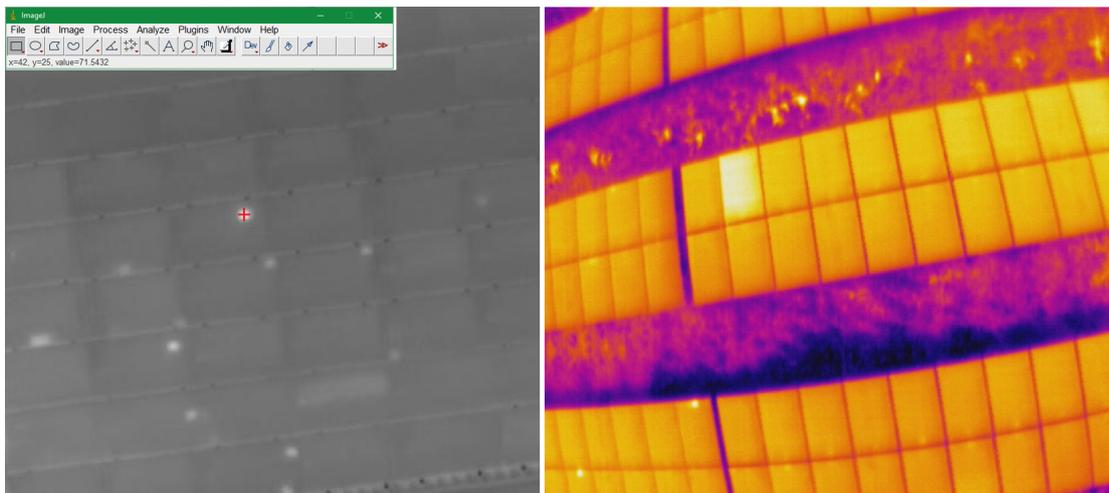


Figure 2.6: Contrast between data types. (Left) A true radiometric thermal image where pixel values correspond to temperature, enabling quantitative analysis. (Right) A typical 8-bit JPEG from a public dataset, where colors are arbitrary and prevent robust, physics-based analysis.

The confluence of these issues suggests a potential generalization problem, where high accuracy figures reported in studies may be artifacts of testing on small, proprietary datasets<sup>10</sup>. Without a high-quality, radiometric, and standardized benchmark dataset, it is impossible to meaningfully compare methodologies and transition research into industry-ready tools<sup>69</sup>.

## 2.5 Research Opportunities

This review has established the profound economic and environmental stakes of PV system underperformance, with losses mounting into the billions of dollars and a widening "O&M Gap" that necessitates automated solutions<sup>12</sup>. The complex, interconnected nature of defect cascades underscores the need for early and accurate diagnostics<sup>38</sup>. While drone-based thermography guided by the IEC 62446-3 standard is the definitive method for scalable data collection, its potential is unrealized due to an analysis bottleneck<sup>9,53</sup>.

With this review in problem domain and defectology a part of the research gap has been identified: is the critical lack of large-scale, public, and standardized datasets containing essential **radiometric** thermal data<sup>63,69</sup>. The current landscape of 8-bit, non-standardized datasets risks a reproducibility crisis, producing models that are not robust or generalizable to real-world conditions<sup>10</sup>. This dissertation is positioned to address this gap by developing and validating novel detection algorithms designed specifically to operate on radiometrically accurate thermal data collected in strict accordance with international standards.

## Chapter 3

# State of the Art in Automated PV Plant Analysis

### 3.1 Introduction

This chapter provides a review of the technical state of the art in the automated analysis of UAS imagery for photovoltaic (PV) plant inspection. Building upon the economic and environmental importance and public data need established in Chapter 2, this section evaluates the algorithmic foundations and existing system-level solutions for panel detection and anomaly classification.

A central theme of this review is the dichotomy between computationally intensive, proprietary, cloud-based commercial systems and the emerging need for transparent, privacy-preserving, and locally executable open-source alternatives. Commercial platforms offer polished services at the cost of data sovereignty, algorithmic transparency, and significant processing delays<sup>77</sup>. Conversely, open-source tools, while philosophically aligned with local control, are often fragmented or exist as academic proofs-of-concept rather than robust, deployable systems<sup>78</sup>. This analysis will systematically identify the technical and operational gaps, thereby positioning the unique contribution of this thesis: to bridge this divide by developing a practical, open-source, and locally-runnable inspection framework.

### 3.2 Computer Vision Approaches for Panel Detection and Localization

The primary challenge in automated inspection is the accurate identification and delineation of individual PV panels from aerial imagery, a critical prerequisite for any subsequent anomaly analysis. The field is dominated by two main families of techniques: traditional image processing methods, which are computationally lean and well-suited for local deployment, and deep learning methods, which represent the performance-driven state of the art but carry a significant computational burden.

### 3.2.1 Traditional Image Processing Techniques

These methods are foundational in computer vision and are particularly relevant to this thesis due to their low computational overhead and their potential for implementation on standard local hardware without reliance on cloud infrastructure. Their principles are well-understood, making them transparent and auditable.

- **Template Matching:** The principle of template matching is to locate instances of a small template image within a larger source image by sliding the template across the source and calculating a similarity metric at each position<sup>79</sup>. The choice of metric is crucial; Normalized Cross-Correlation (NCC) is particularly important for aerial PV inspection, as it provides a degree of invariance to the global changes in brightness and contrast ubiquitous in outdoor imagery<sup>79</sup>. While basic template matching is sensitive to scale and rotation, it can be quite powerful in cases where the pattern to match is repeated and always very constant in shape and dimensions (like a PV module).

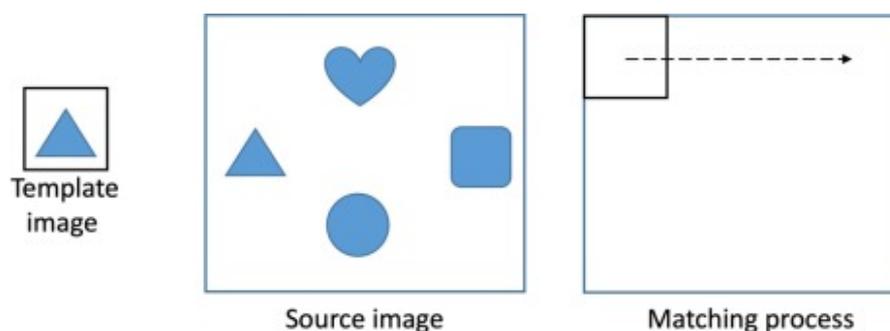


Figure 3.1: How template matching works, a template image is taken and in our case is moved across the target image. The points where the match was higher are identified as having a match.

- **Hough Transform:** The Hough Transform is a powerful feature extraction technique used to find imperfect instances of shapes, such as lines, through a voting procedure in a parameter space<sup>80</sup>. Its primary application in PV inspection is the detection of the straight lines that form the rectangular boundaries of PV modules, typically after an edge detection step using the Canny algorithm<sup>81</sup>. A critical application of this technique is the correction of geometric distortions inherent in aerial imagery. In the orthomosaic reconstruction phase, by detecting the dominant sets of horizontal and vertical lines, the Hough Transform allows for the computation of the necessary homography to apply a corrective warp, simplifying subsequent segmentation and analysis<sup>81</sup>.

### 3.2.2 Deep Learning-Based Object Detection

Methods based on Convolutional Neural Networks (CNNs) represent the current state of the art in object detection accuracy but require significant computational power, typically necessitating GPUs and favoring cloud-based processing. This paradigm contrasts with the local-first objectives of this thesis.

The "You Only Look Once" (YOLO) family of algorithms has revolutionized real-time object detection, with successive versions continuously improving the speed-accuracy trade-off<sup>82;83</sup>. In PV inspection, these models have demonstrated exceptional performance, with mean Average Precision (mAP) for panel detection often exceeding 95%<sup>83</sup>. For instance, the ST-YOLO model reported an mAP@0.5 of 96.6% for defect detection<sup>84</sup>. This high accuracy, however, is computationally demanding; a model like YOLOv8s contains over 11 million parameters and requires 28.4 GFLOPs for inference, making real-time processing on standard hardware challenging<sup>84</sup>.

The choice of a detection algorithm, therefore, represents a fundamental engineering trade-off. Deep learning models are architected for a "GPU-first" or "cloud-native" deployment, while traditional methods like template matching are "local-first" and well-suited for applications where accessibility and data privacy are paramount. The selection of template matching in this thesis is a deliberate decision that prioritizes these practical deployment constraints over achieving the absolute peak accuracy reported in deep learning literature.

## 3.3 Machine Learning for Thermal Anomaly Classification

Once panels are localized, the subsequent task is to classify them based on their thermal signatures to identify potential defects. This section reviews the two dominant paradigms for this task: supervised learning, which requires pre-labeled training data, and unsupervised learning, which identifies anomalies by detecting deviations from a learned model of normal behavior.

### 3.3.1 Supervised Learning Techniques

Supervised methods learn a direct mapping from input data to a set of predefined anomaly classes. Their efficacy is fundamentally tied to the quality and diversity of the labeled dataset used for training. For the direct analysis of thermal imagery, **Convolutional Neural Networks (CNNs)** are the preeminent approach because they can automatically learn the complex spatial patterns and thermal signatures that define defects<sup>85</sup>. Architectures ranging from custom networks to established models like VGG16 and ResNet have demonstrated consistently high performance, achieving fault classification accuracies of over 97% for various defect types<sup>85-87</sup>.

Beyond direct image analysis, other techniques excel at classifying pre-extracted thermal

features. For instance, both **Random Forest (RF)** and **Support Vector Machines (SVMs)** have been effectively used to classify specific faults, such as inverter issues or partial shading, based on statistical data derived from the thermal images<sup>88;89</sup>. A common hybrid strategy also involves using a pre-trained CNN solely as a powerful feature extractor, with its output then fed into a less computationally intensive classifier like an SVM for the final decision<sup>90</sup>.

### 3.3.2 Unsupervised Learning Techniques

Unsupervised methods are of practical importance where labeled data is scarce. They operate by learning the statistical characteristics of "normal" PV panel behavior and flagging significant deviations as anomalies. Several techniques are employed to build this model of normalcy. For example, a **One-Class SVM (1SVM)** can be trained to define a boundary that encloses the feature space of healthy operations, flagging any panel whose thermal signature falls outside this boundary as anomalous<sup>91</sup>.

Similarly, generative models like **Variational Autoencoders (VAEs)** are trained to accurately reconstruct thermal images of healthy panels; a high reconstruction error on a new panel therefore signals a potential defect<sup>92</sup>. The **Isolation Forest** algorithm operates on a related principle, identifying anomalies not by reconstruction but by their ease of isolation. Since defects are "few and different," their corresponding data points are more easily separated from the dense cluster of normal panel data in a decision tree structure<sup>92</sup>.

The scarcity of comprehensive, labeled public datasets for PV defects poses a significant challenge for supervised learning. Unsupervised methods circumvent this bottleneck, offering a highly practical path for deployment by requiring only data from normal operations, which is abundant.

## 3.4 Review of Existing Systems and Solutions

This section transitions from individual algorithms to an analysis of integrated systems, evaluating the current landscape of academic, open-source, and commercial solutions.

### 3.4.1 Academic and Open-Source Implementations

The open-source domain provides invaluable building blocks but generally lacks the end-to-end integration and robustness of commercial products, the only complete software pipeline is PV Hawk:

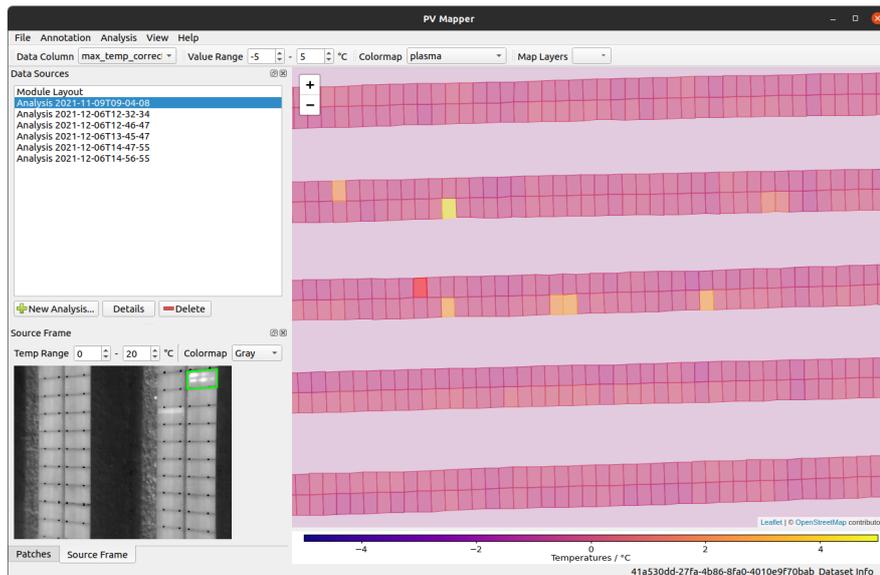


Figure 3.2: A screenshot of the GUI of the open-source PV Hawk viewer

PV-Hawk stands out as a comprehensive open-source system for automated PV inspection. It is a command-line tool that takes drone video and GPS data, performs a 3D reconstruction with OpenSfM, segments modules, and extracts their geocoordinates for mapping, all designed for local processing<sup>78;93</sup>. A companion desktop application allows for visualization and analysis<sup>94;95</sup>. While demonstrating the feasibility of a local pipeline, its author notes its status as a "proof-of-concept" with potential instability, highlighting a clear gap for a more robust open-source alternative<sup>78</sup>.

The main difference between the approach taken by PV Hawk and this Thesis Project is in the starting data: PV Hawk directly uses the Images or video frame from the drone, this approach needs a Structure-From-Motion reconstruction to geolocate the modules correctly but it can be hindered by various factors. Many researches have highlighted that the use of orthomosaics is much more accurate and much less error prone<sup>96;97</sup>, so in the following Chapter 4 precisely this method was used.

Beyond PV-Hawk, the open-source landscape is fragmented. Numerous repositories exist but typically serve narrow purposes, such as providing datasets<sup>98</sup> or specific algorithmic implementations presented as feasibility studies<sup>99</sup>. This fragmentation underscores the lack of a unified, production-ready open-source system for PV plant operators.

### 3.4.2 Commercial Cloud-Based Platforms

The dominant commercial model is the polished, cloud-based Software-as-a-Service (SaaS) platform, which abstracts away technical complexities but introduces trade-offs in cost, transparency, turnaround time, and data privacy.

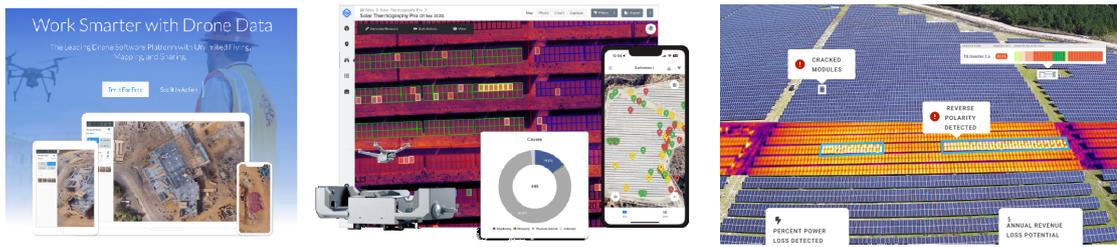


Figure 3.3: A sample of the available commercial cloud-based SaaS solutions: Drone Deploy (Left), SiteMark (Center) and RaptorMaps (Right)

The most used leading platforms include DroneDeploy, a versatile platform serving multiple industries<sup>100</sup>; Raptor Maps, a highly specialized platform focused exclusively on the solar industry<sup>101</sup>; and Sitemark, another solar-focused platform offering unique rapid-analysis solutions<sup>102</sup>.

#### 3.4.2.1 Criticalities of Commercial Solutions

A critical examination of these platforms reveals a consistent operational model with several key characteristics:

- **Processing Model:** All major commercial solutions are fundamentally cloud-based; users must upload their imagery to the company's servers for processing<sup>100</sup>.
- **Turnaround Time:** While premium offerings promise results in hours<sup>103;104</sup>, standard service tiers can be substantially slower, with some platforms specifying turnaround times of 7 to 20 days<sup>105</sup>. This long wait stands in stark contrast to the near-instant feedback possible with local processing.
- **Algorithmic Transparency:** These platforms market their "proprietary AI" as "black boxes," providing no details about the models used, the data they were trained on, or their performance metrics<sup>101</sup>. An open-source solution offers complete auditability.
- **Data Privacy and Sovereignty:** This is a critical distinction. The privacy policies of major platforms explicitly state that customer data may be used to "monitor, analyze and improve the performance of the Services"<sup>106;107</sup>. This confirms that the client's operational data is a raw material used to enhance the vendor's core AI asset. For plant owners who consider their data proprietary, this is a significant concern<sup>108</sup>. A local system provides absolute data sovereignty.

This analysis reveals that the business model of commercial platforms is centered on their proprietary AI, which is continuously improved using client data. The proposed system in this thesis offers a fundamentally different value proposition: it is a tool, not a service, where the user retains complete ownership and control of their data.

### 3.5 Gap Analysis and SoTA Limitations

This final section synthesizes the findings from the preceding review to construct a compelling argument for the novelty and necessity of the research presented in this thesis.

The review of the state of the art reveals a significant disparity in the landscape of automated PV inspection tools. On one side, commercial solutions offer polished, end-to-end systems but are built on a proprietary, cloud-based, "black-box" model, introducing critical issues regarding data privacy, transparency, and cost<sup>100</sup>. On the other side, academic and open-source solutions, while philosophically aligned with local control, are fragmented and often immature, lacking a unified, robust, and user-friendly system for day-to-day operations<sup>78</sup>.

The unmet need is for a solution that is simultaneously open-source, end-to-end, locally-executable on standard hardware, compliant with IEC standards, that uses Orthomosaic images, and privacy-preserving, ensuring sensitive data remains entirely within the owner's control.



## Chapter 4

# Software Interface and Methods

This chapter provides a comprehensive exposition of the system architecture and the multi-stage methodology developed for the automated detection of thermal anomalies in photovoltaic (PV) plants. The system is implemented as a cohesive software solution, the "Solar PV Anomaly Detection Suite," featuring a Graphical User Interface (GUI) designed to guide an operator through a structured workflow. The methodology emphasizes a modular, local-first processing pipeline, user-in-the-loop control, and the synergistic integration of advanced image processing with machine learning techniques.

### 4.1 Orthophoto-Centric Workflow Rationale

As discussed in the preceding State of the Art review, existing open-source frameworks such as PV-Hawk often operate directly on raw aerial videos and their associated metadata. This project deliberately diverges from that approach by adopting an orthophoto-centric workflow. The system is architected to ingest pre-processed orthomosaics, which are generated beforehand using specialized photogrammetry software. While the commercial package Agisoft Metashape was utilized for this research, the workflow is compatible with any tool capable of producing high-quality GeoTIFF orthomosaics, including open-source solutions like WebODM.

This strategic decision to decouple the analysis pipeline from the initial Structure from Motion (SfM) and orthomosaic reconstruction phase offers several advantages. It allows operators to use best-in-class, established tools for photogrammetry and focuses the scope of this project entirely on the novel aspects of panel detection, analysis, and classification.

However, the importance of the reconstruction quality cannot be overstated. The fidelity of the input orthomosaics is imperative for the success of the entire downstream pipeline. A high-quality survey and reconstruction, conducted in strict adherence to the IEC 62446-3 standard, is essential for minimizing artifacts such as sun glare. Specular reflection from panel surfaces can saturate both RGB and thermal sensors, corrupting the underlying data and rendering accurate detection and quantitative thermal analysis impossible.

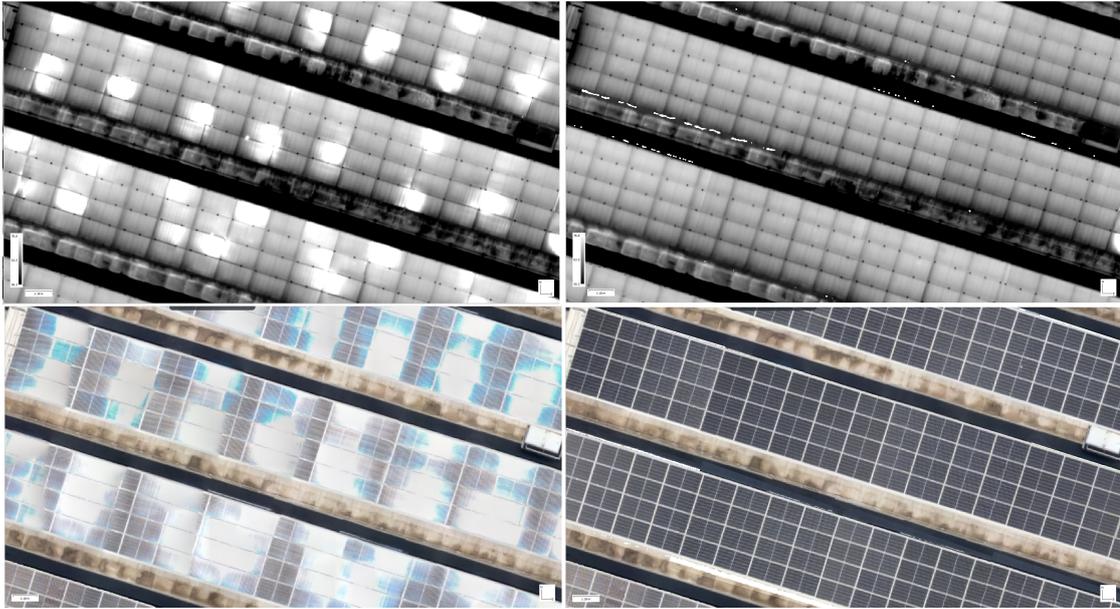


Figure 4.1: The critical impact of reconstruction quality on data fidelity. A poor reconstruction (left) exhibits significant sun glare and artifacts, while a high-quality reconstruction (right) provides clean data suitable for analysis.

## 4.2 System Architecture and Data Flow

The "Solar PV Anomaly Detection Suite" is designed as a modular desktop application that guides the user through a sequential analysis pipeline. The architecture, as depicted in Figure 4.2, is structured around a series of distinct processing modules, each corresponding to a major stage of the workflow. This modularity ensures a logical progression of tasks, from data ingestion to the final generation of reports. The primary data objects—RGB and thermal image layers, panel detections, and thermal statistics—flow between these modules, with each stage enriching the data for the next. The system is designed for local execution, ensuring that all data and processing remain on the operator's machine, thus guaranteeing data sovereignty and privacy.

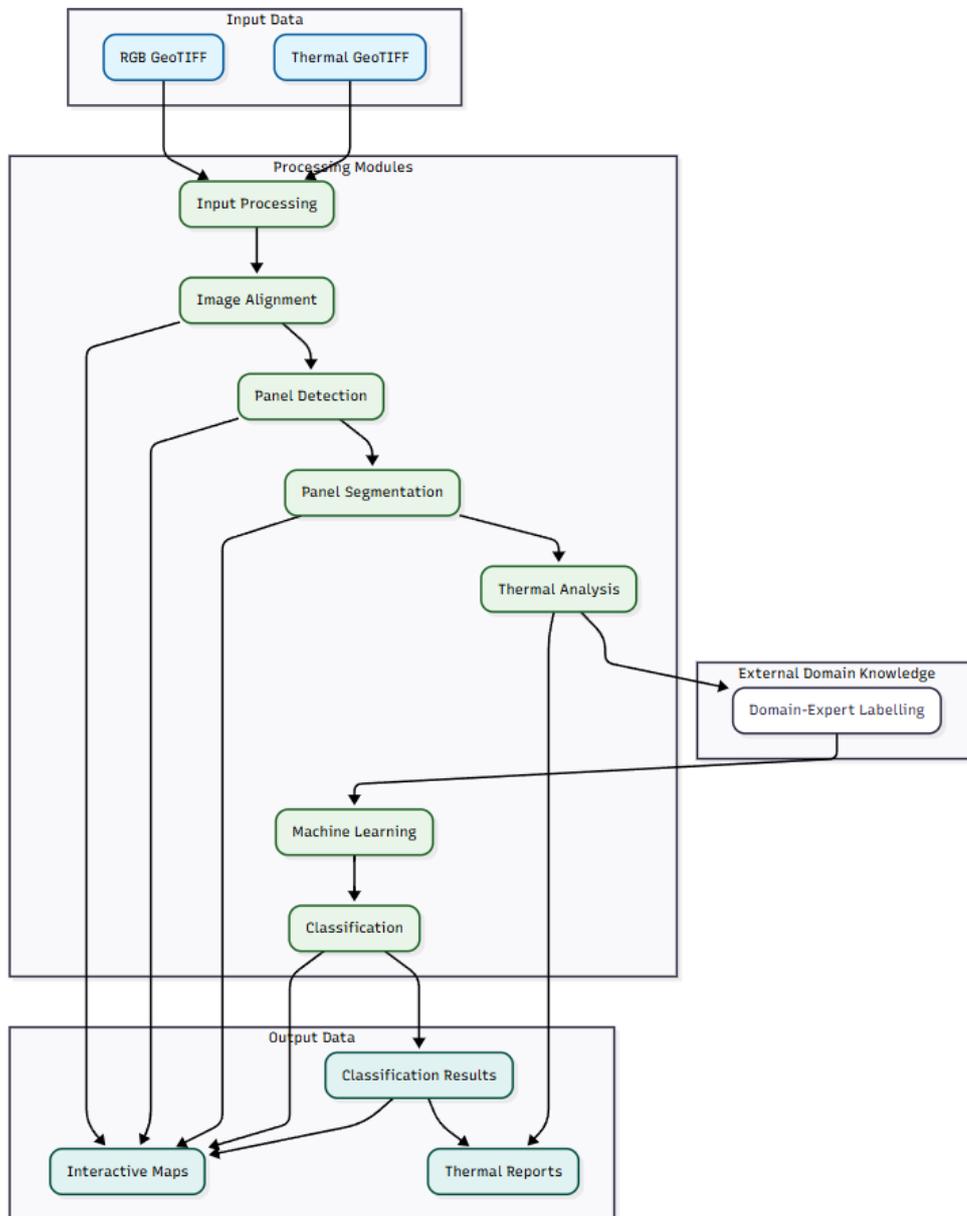


Figure 4.2: A high-level diagram illustrating the system's architecture and the logical flow of data between its core processing modules.

### 4.3 The Graphical User Interface (GUI)

The entire system is encapsulated within a user-friendly GUI, which serves as the primary command and control center for all operations. The interface is organized into three main areas: a persistent "Workflow" panel on the left, a main interactive view in the center, and a context-sensitive control panel that appears on the right. The Workflow panel provides access to the distinct modules of the application: Inputs, Alignment, Detection, Segmen-

tation, Models, Classification, and Outputs. This structure provides a clear and intuitive step-by-step guide for the user, from getting started to exporting the final results.

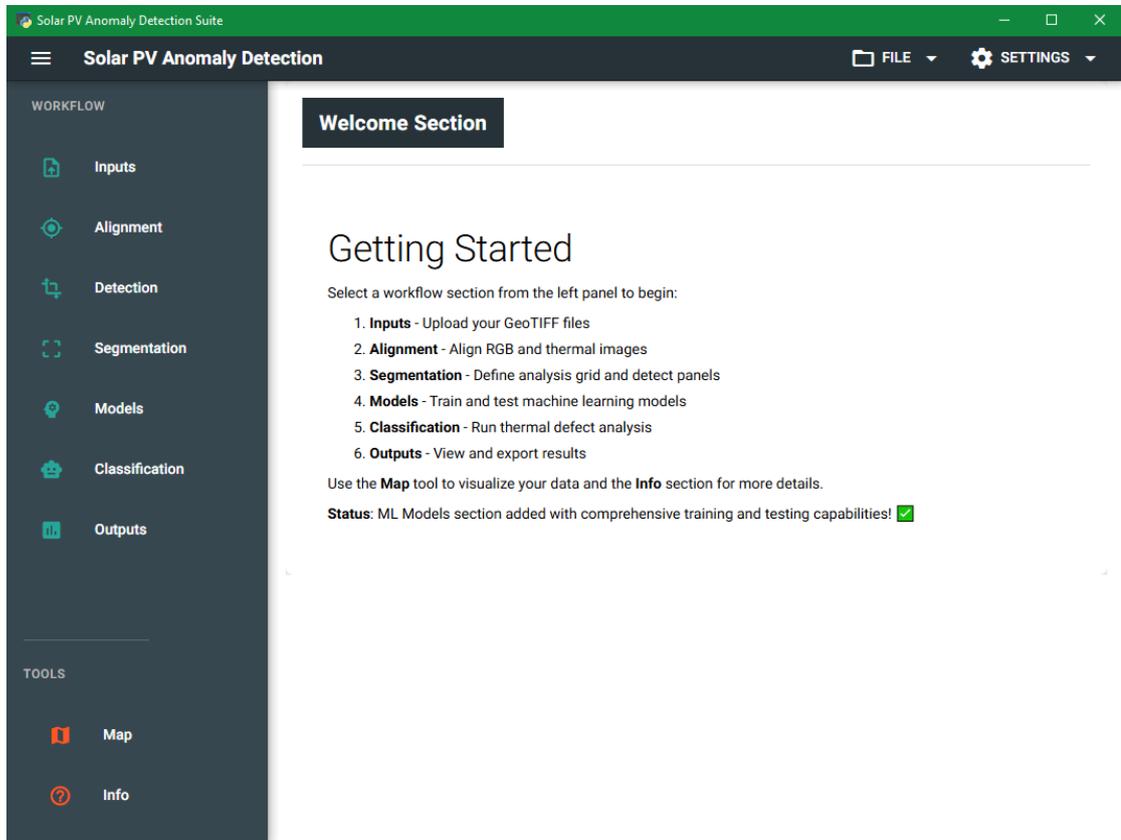


Figure 4.3: The main welcome screen of the Solar PV Anomaly Detection Suite, outlining the six primary workflow sections accessible via the left-hand navigation panel.

## 4.4 Module 1: Input and Alignment

### 4.4.1 Data Ingestion

The workflow begins in the "Inputs" section, where the user uploads the two foundational GeoTIFF files: the three-channel RGB orthophoto and the single-channel radiometric thermal orthophoto. The system validates these inputs to ensure they are in the correct format before proceeding.

### 4.4.2 Manual Image Alignment

Upon successful data ingestion, the operator moves to the "Alignment" module. Here, the GUI presents the two orthophotos side-by-side on an interactive map interface. To rectify minor misalignments between the layers, the system implements a user-driven alignment procedure based on a 2D projective transformation.

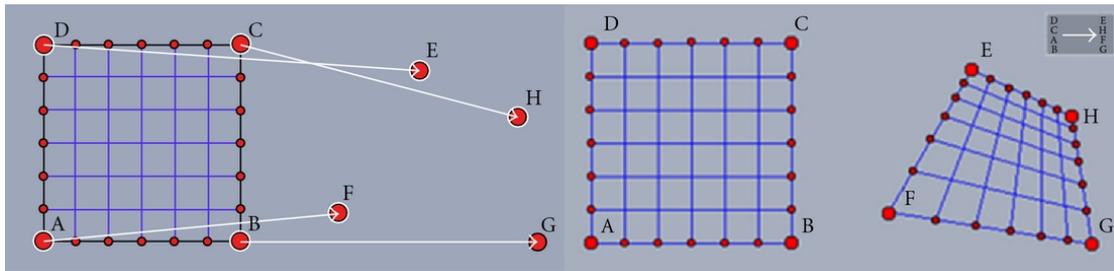


Figure 4.4: A projective transformation: placing four pair of points is able to unequivocally define a transformation for the entire image.

The user is prompted to place at least four homologous control points on each image. The more precise the placement of these points, the better the resulting alignment. Once the points are set, the system computes the transformation matrix that warps the thermal layer coordinates system to precisely match the one of the RGB layer. This accurate, user-verified alignment is fundamental for all subsequent stages.

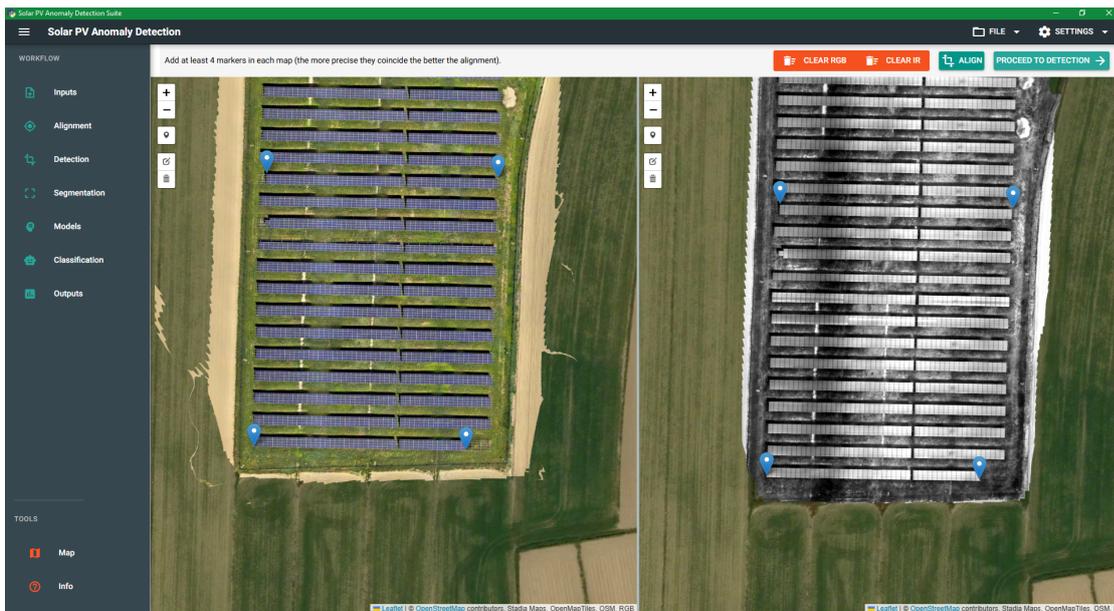


Figure 4.5: The user interface for the Alignment module. The operator places corresponding control points (blue markers) on the RGB (left) and thermal (right) orthophotos to enable a precise affine transformation.

## 4.5 Module 2: The Panel Detection Pipeline

The core of the system’s analytical capability resides in the "Detection" module. A key architectural decision was to execute the detection process independently on the RGB and thermal layers to ensure that the bounding box for each detected panel is perfectly tailored

to its respective image data.

#### 4.5.1 Core Detection via Normalized Cross-Correlation

The primary detection algorithm is a robust implementation of template matching based on Normalized Cross-Correlation (NCC). The process shown in 4.7 is initiated by the user defining a panel group and its configuration (e.g., rows and columns) on the map. This generates a grid from which multiple panel templates are automatically extracted. Using multiple templates builds a more resilient matching model that can account for variations in panel appearance.

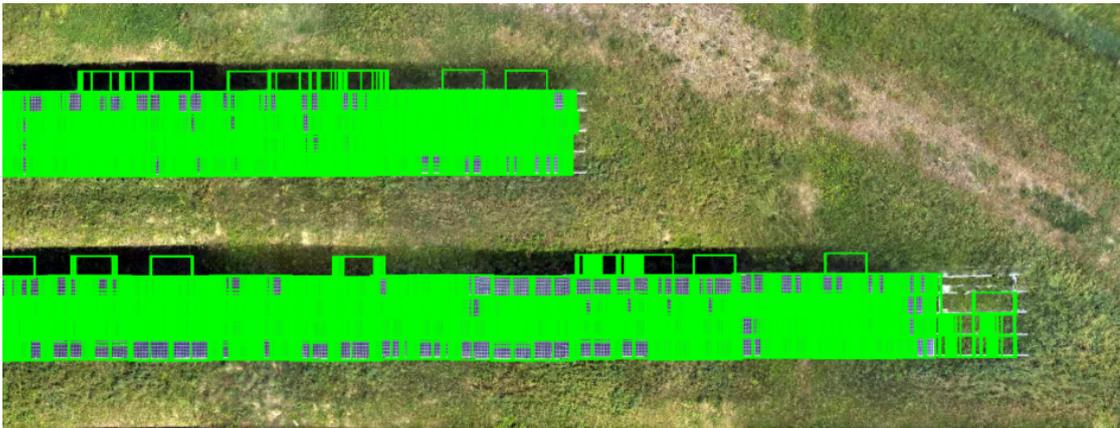


Figure 4.6: A special implementation of the NMS algorithm was used since the images are very large, for a small plant more than 2 million panels where detected and doing an IoU check for all of them is exponentially slower.

The operator can then fine-tune the detection using sliders for ‘Template Matching Threshold’ and ‘NMS IoU Threshold’ before running the main matching process. As seen in 4.6, During the NMS (Non-Maximum Suppression) phase, each detection is spatially binned together with others nearby, this allows for an exponentially faster IoU check between the detections instead of doing the check on all detections in a single bin.

Furthermore a novel feature was introduced for the thermal detection phase to enhance robustness against hotspots, allowing an operator to set a temperature cap for the NCC calculation, preventing high-temperature anomalies from causing a template mismatch.

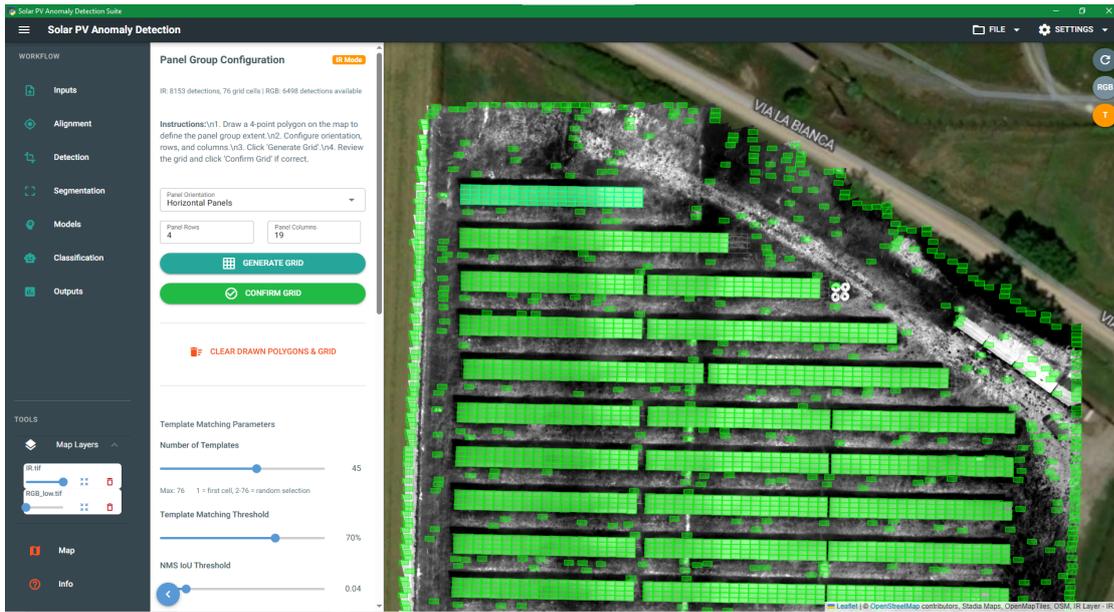


Figure 4.7: The UI for the core detection step. The operator defines a grid to create templates and adjusts parameters for template matching and Non-Maximum Suppression (NMS). The initial raw detections are shown as small green boxes.



Figure 4.8: Example of thermal (left) and RGB (right) panel templates automatically extracted by the system for use in the NCC algorithm.

### 4.5.2 Advanced Detection Refinement

To maximize detection accuracy, the system provides a suite of advanced algorithms to refine the initial results. These tools leverage the inherent spatial structure of PV plants. As shown in Figure 4.9, the refinement steps include:

- **Spatial Clustering and Grid Fitting:** This algorithm first employs the DBSCAN algorithm to group detected panels into spatially coherent blocks. Subsequently, a custom algorithm fits a regular grid to each cluster, allowing the system to programmatically identify and remove false positives and infer the locations of missed panels. The user can control parameters such as 'Clustering Distance' and 'Delta Jitter'.
- **Border Outlier Removal:** A subsequent step allows for the removal of isolated outlier panels based on a confidence threshold.

- **Cellular Automaton for Gap Filling:** A novel algorithm inspired by Conway's Game of Life is implemented to intelligently fill small, isolated gaps within a regular panel array based on the number of neighboring panels.



Figure 4.9: The various algorithms removing unwanted detections (Top). The UI for the advanced detection refinement pipeline, showing the controls for DBSCAN/Grid Fitting and Border Outlier Removal. The map displays the refined panel detections. (Bottom)

## 4.6 Module 3: Segmentation and Feature Extraction

The "Segmentation" module is where the detections from the two layers are unified and analyzed. With complete and verified detection sets, the first step is to establish a one-to-

one correspondence between them by calculating the Intersection over Union (IoU) of the pairs of bounding boxes (Thermal and RGB). A pair is successfully established if their IoU exceeds a predefined threshold.

For each successfully matched pair, the system extracts the cropped RGB and thermal images and computes a vector of key thermal statistics (mean, median, min/max temperature, standard deviation, and variance). All associated data is then stored in a structured database.

This module also functions as an interactive decision-support and labeling tool. The GUI presents the plant on the map, with each panel color-coded based on a selected thermal indicator. A plant-wide histogram with interactive sliders allows the operator to dynamically define thresholds for "nominal" and "anomalous" ranges, providing immediate visual feedback. By clicking on a panel, the operator can view its detailed information, including its pairing quality, thermal data, and the cropped thermal and RGB images. The user can then save labels based on these visual analytics, creating a dataset for domain-expert labeling and subsequent model training.

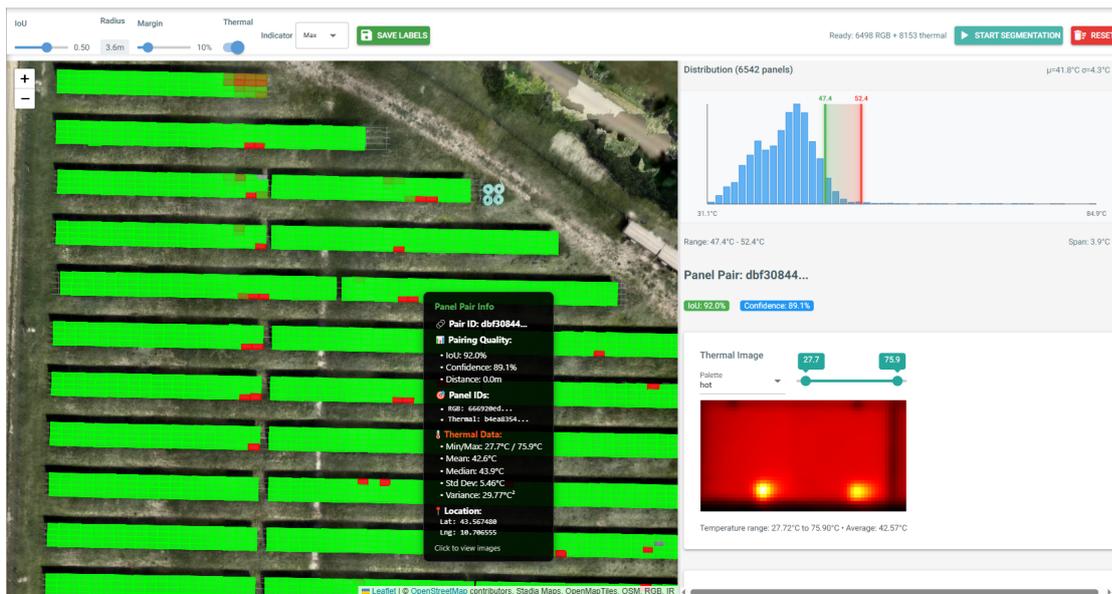


Figure 4.10: The interactive analysis dashboard within the Segmentation module. The interface displays thematically colored panels, a plant-wide thermal distribution histogram with user-controlled thresholds, and a detailed information card for the selected panel.

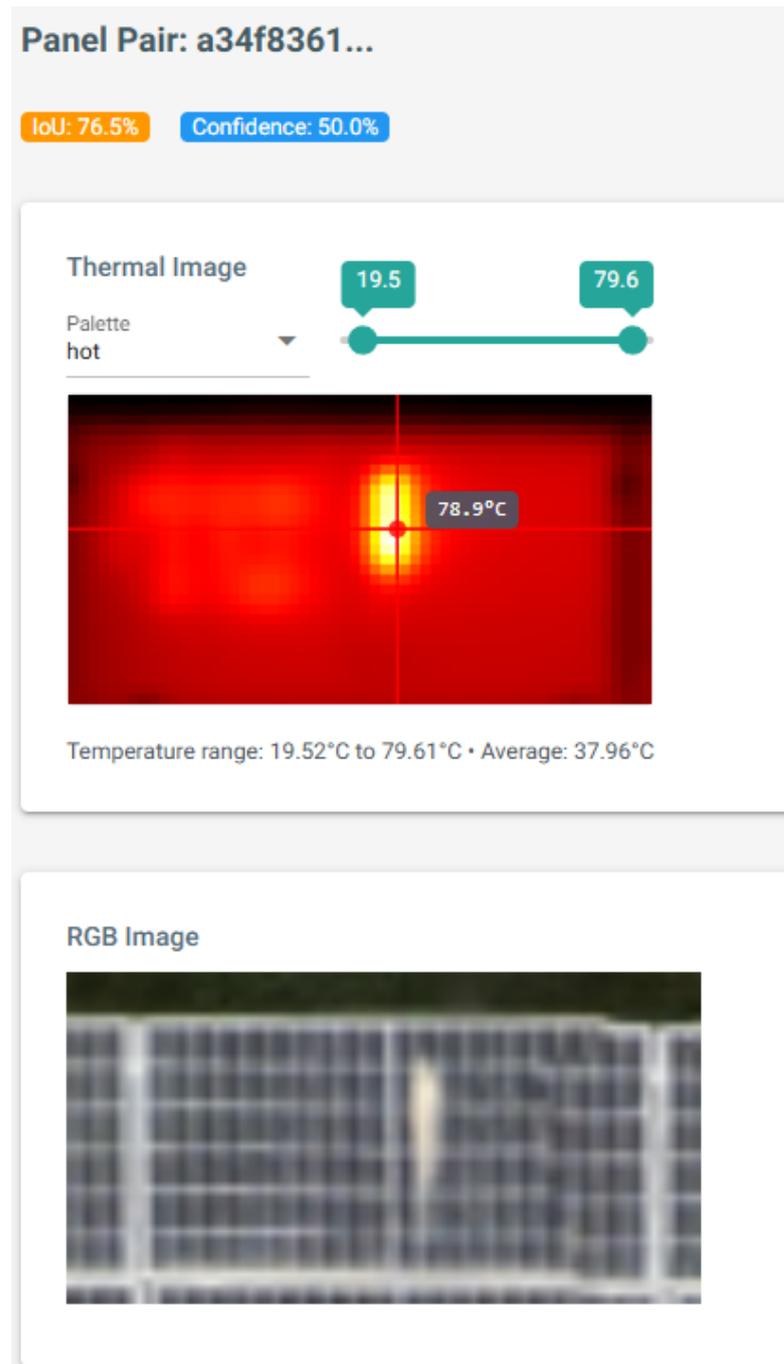


Figure 4.11: A detailed view of a single panel pair, showing the radiometric thermal image with its temperature scale and the corresponding high-resolution RGB image. In this case the Hot-Spot is caused by a physical obstruction like soiling or bird deposits.

## 4.7 Module 4: Machine Learning Model Training

The "Models" section of the suite is more intended for the users that want to use the tool as a dataset generation tool for research purposes, it provides a comprehensive environment for training, testing, and evaluating various machine learning models. The user begins by loading a labeled CSV file (generated in the previous step or prepared externally). The interface allows for the training of multiple model types—including binary, multi-class, and regression models—on the thermal feature data. For each model, such as the Binary Random Forest shown in Figure 4.12, the UI displays training progress, training time, and a full suite of performance metrics for both training and testing sets, including accuracy, precision, recall, and F1 score. A confusion matrix is also generated to provide a clear visual assessment of the model's classification performance.

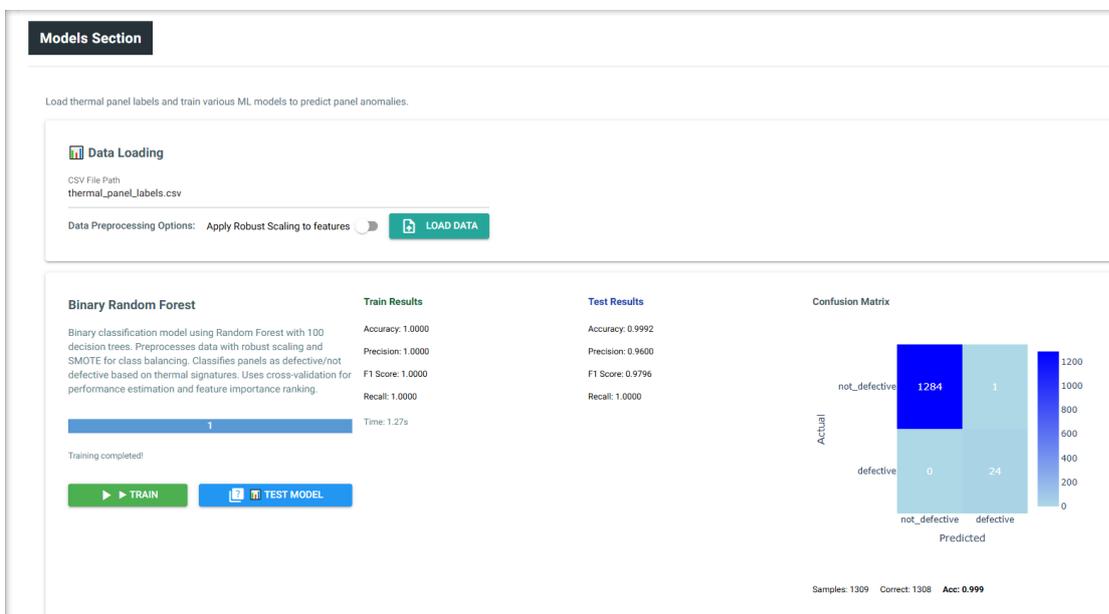


Figure 4.12: The user interface for the Machine Learning Models module. This view shows the results of training and testing a Binary Random Forest classifier, including performance metrics and a confusion matrix.

## 4.8 Module 5: Classification and Analysis

Once a model has been trained and saved, the "Classification" module is used to apply it to the entire dataset of detected panels. The user can load a pre-trained model from the previous step or a custom one and run the classification process. The results are displayed interactively on the map, with each panel color-coded according to its predicted class (e.g., green for 'not defective', red for 'hot spot'). Clicking on a panel reveals its full thermal data alongside the predictions from multiple models (e.g., 'multiclass\_nn', 'binary\_rf',

'regression\_nn'). This provides a powerful tool for rapidly assessing the health of the entire plant based on the learned knowledge encapsulated in the trained models (even if they were trained from a different plant).

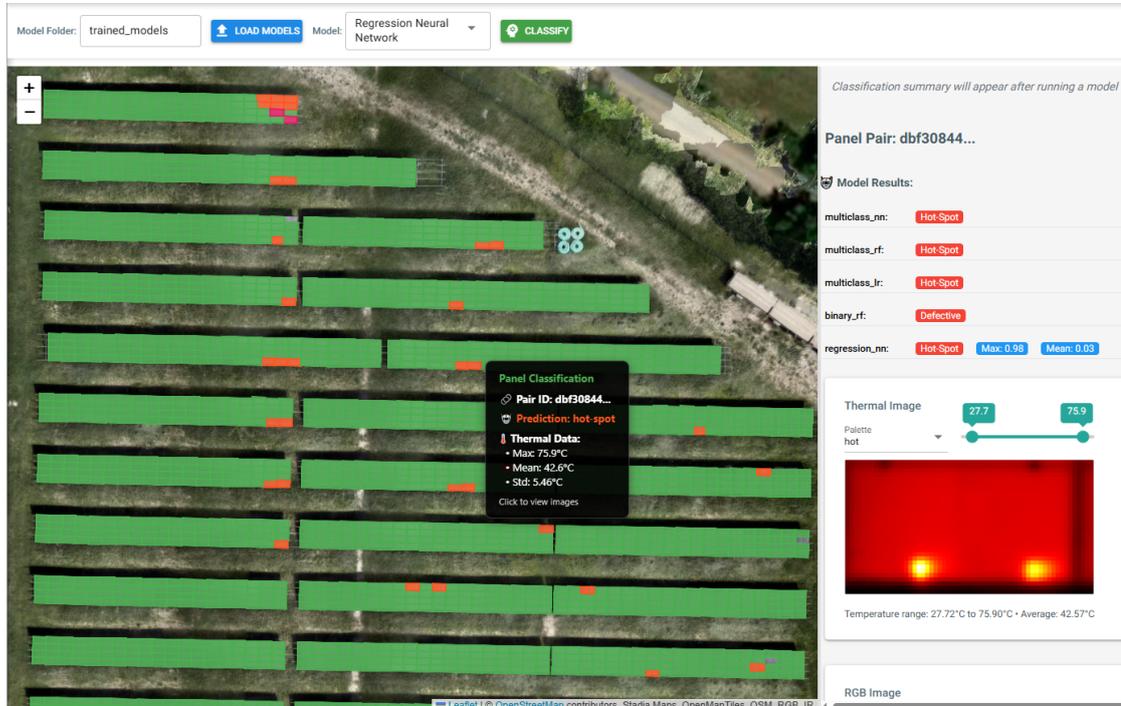


Figure 4.13: The Classification module interface. The map visualizes the model's predictions, with a detailed pop-up showing the predicted class ('Hot Spot') for the selected panel, alongside its thermal data and detailed classification results on the right.

## 4.9 Module 6: Outputs and Reporting

The final "Outputs" module provides a clean, tabular summary of all detected anomalies from the classification step. The report lists each anomalous panel's ID, its predicted anomaly type, the model's confidence in that prediction, and its precise geographic location. This actionable report can be exported and used by maintenance teams to locate and address specific defects in the field, thus closing the loop from automated data analysis to physical remediation.

## Outputs Section

View, analyze, and export the results of AI detections.

This includes reports, annotated imagery, and statistical summaries.

Detected Anomalies			
ID	Anomaly Type	Confidence	Location
P001-A01	Hotspot	95%	43.568201746846, 10.707515646128
P003-A01	Module Failure	88%	43.568201746846, 10.70751564635
P007-A02	Soiling	75%	43.568201746846, 10.70754587925
P009-A03	Diode Failure	65%	43.568201746846, 10.707576112172
P011-A04	Hotspot	97%	43.568201746846, 10.707606345096
P013-A05	Module Failure	86%	43.568201746846, 10.70763657802
P015-A06	Soiling	75%	43.568201746846, 10.707666810944
P017-A07	Diode Failure	64%	43.568201746846, 10.707697043868
P019-A08	Hotspot	92%	43.568201746846, 10.707727276792
P021-A09	Module Failure	89%	43.568201746846, 10.707757509716

Figure 4.14: The Outputs section, presenting a final, exportable report of all detected anomalies with their type, confidence level, and location.



## Chapter 5

# Experiment and Results

This chapter presents a comprehensive evaluation of the developed *Solar PV Anomaly Detection Suite*. The primary objective of the experiments is to validate the efficacy and efficiency of the entire end-to-end pipeline, from data ingestion to final classification. The evaluation is structured to assess each major component of the system, focusing first on the experimental conditions, followed by the performance of the panel detection and anomaly classification modules. The results underscore the system's capacity not only as a functional diagnostic tool but also as a powerful framework for creating high-quality, labeled datasets for future research.

### 5.1 Experimental Setup

#### 5.1.1 Dataset and Hardware

The experiments were conducted on a diverse dataset aggregated from 13 different photovoltaic plants located in Italy. These sites, summarized in Table 5.1, comprise 10 large-scale ground-mounted fields and 3 industrial rooftop installations, ensuring the system was tested against a variety of layouts, panel types, and environmental conditions.

For each site, paired RGB and thermal aerial surveys were conducted. The raw imagery, totaling approximately 73,000 images (36,500 per sensor type), was processed separately using Agisoft Metashape to generate two coregistered orthomosaics per plant. The resulting orthophotos exhibit a high spatial resolution, with an average Ground Sampling Distance (GSD) of approximately 0.5 cm/pixel for the RGB layer and 1.5 cm/pixel for the thermal layer.

All processing tasks, from orthomosaic image alignment to model inference, were executed on a mid-range laptop equipped with an 8-core AMD CPU (2020) and 16 GB of RAM. No dedicated GPU was required for the core pipeline, underscoring the system's accessibility and computational efficiency.

Table 5.1: Summary of the aggregated dataset used for system evaluation.

Parameter	Value
Number of PV Plants	13
- Ground-Mounted	10
- Industrial Rooftop	3
Total Raw Images	~73,000
Output Orthomosaics	26 (13 RGB, 13 Thermal)
Average RGB GSD	~0.5 cm/pixel
Average Thermal GSD	~1.5 cm/pixel
Processing Hardware	Mid-range Laptop (No dGPU)

### 5.1.2 System Performance and Scalability

The entire software pipeline was tested on all 13 orthomosaic pairs. The system's architecture, while currently executed as a single-threaded desktop application, was designed with modularity to facilitate future optimization for vertical or horizontal scaling. Table 5.2 details a representative performance breakdown for the processing of a large-scale plant comprising 6,857 panels. The processing times demonstrate the system's efficiency, with the entire workflow from orthophoto alignment to classification completing in under four minutes. The most time-consuming step remains the initial template matching, whereas subsequent refinement and analysis stages are comparatively rapid.

Table 5.2: Representative performance breakdown for a 6,857-panel PV plant.

Processing Module	Execution Time (s)
Image Alignment	1
Panel Detection	
- Template Matching	142
- DBSCAN Clustering	56
- Refinement	1
Panel Segmentation	43
Classification (Inference)	2-10
<b>Total Approximate Time</b>	<b>~250</b>

## 5.2 Evaluation Metrics

To quantitatively assess the system's performance, a set of standard evaluation metrics was employed for the distinct tasks of detection and classification.

For **panel detection**, the primary metric is the True Positive Rate (TPR), also known as Recall or Sensitivity. It is defined as:

$$TPR = \frac{TP}{TP + FN}$$

where  $TP$  (True Positives) is the number of correctly identified panels and  $FN$  (False Negatives) is the number of missed panels.

For **anomaly classification**, models were evaluated using two primary metrics:

- **Accuracy:** The ratio of correctly classified instances to the total number of instances.
- **F1-Score (Weighted Average):** The weighted average of the F1-scores for each class, which is the harmonic mean of precision and recall. This metric is particularly useful for imbalanced datasets.

For the **regression models**, which were trained to predict a continuous value (expert-assigned confidence), two additional metrics were used:

- **R-squared ( $R^2$ ):** The coefficient of determination, representing the proportion of the variance in the dependent variable that is predictable from the independent variables.
- **Mean Absolute Error (MAE):** The average of the absolute differences between the predicted values and the actual values.

For direct comparison with classifiers, the continuous output of the regression models was subsequently label-encoded into discrete classes to calculate accuracy and F1-scores.

### 5.3 Results of Panel Detection

The panel detection module, combining template matching with the advanced refinement algorithms, demonstrated exceptionally high performance across all tested sites. Table 5.3 highlights the results on all the plants and the average performance across the entire dataset.

Table 5.3: Panel detection performance across all 13 PV plants, highlighting variation between installation types.

Plant Type	Total Panels	Detected (TP)	Missed (FN)	TPR
Ground-Mounted (Site A)	6,857	6,848	9	99.87%
Ground-Mounted (Site B)	11,255	11,236	19	99.83%
Ground-Mounted (Site C)	3,657	3,657	0	100.00%
Ground-Mounted (Site D)	8,421	8,415	6	99.93%
Ground-Mounted (Site G)	12,877	12,871	6	99.95%
Ground-Mounted (Site I)	11,298	11,284	14	99.88%
Ground-Mounted (Site J)	8,935	8,928	7	99.92%
Ground-Mounted (Site K)	10,880	10,874	6	99.94%
Ground-Mounted (Site L)	19,650	19,638	12	99.94%
Ground-Mounted (Site M)	7,542	7,539	3	99.96%
Roof-Mounted (Site E)	4,776	4,516	260	94.56%
Roof-Mounted (Site F)	1,244	1,209	35	97.19%
Roof-Mounted (Site H)	16,521	16,179	342	97.93%
<b>Overall (All 13 Sites)</b>	<b>123,913</b>	<b>123,194</b>	<b>719</b>	<b>99.42%</b>

The results strongly suggest that for structured environments like PV plants, the implemented classical computer vision approach is highly effective. It achieves near-perfect detection rates without the significant computational overhead associated with deep learning models such as YOLO or Mask R-CNN. These alternative models would require substantial GPU memory and processing power for inference on orthomosaics that can be as large as  $40,000 \times 26,000$  pixels.

The slightly lower average TPR of 99.42% is primarily attributed to the challenges posed by rooftop installations, where metallic roofing elements, HVAC units, and complex shadowing occasionally led to false negative or false positive detections. Nonetheless, the overall performance robustly validates the chosen detection strategy.

## 5.4 Results of Anomaly Classification

### 5.4.1 Methodology and Ground Truth Creation

A key objective of this work was to validate a pipeline that could facilitate the rapid creation of labeled datasets. The anomaly classification experiment was designed to reflect this. Instead of relying on a purely Computer Vision (CV) approach, which would necessitate extensive manual labeling of images, this Thesis focused on training simple, fast machine learning models using statistical thermal features. The features used for training were:  $T_{max}$ ,

$T_{mean}$ ,  $T_{median}$ ,  $T_{std}$ ,  $T_{var}$ , and  $T_{min}$ .

These variables were calculated during the panel pairing phase for each panel found with the template matching.

The ground truth dataset was created using a semi-automated, expert-in-the-loop workflow enabled by the software's UI:

1. **Candidate Pre-selection:** Using the interactive histogram tool, loose thermal thresholds were applied to an unlabeled dataset of over 10,000 panels. This step filtered the dataset down to approximately 800 panels exhibiting potentially anomalous thermal signatures.
2. **Expert Labeling:** This reduced set of "suspicious" panels was exported and provided to a certified expert in PV thermography, who assigned a ground truth label to each one. The multiclass labels included: *Normal*, *Hot-spot*, *Module-Defect*, *Diode*, and *PID*.
3. **Dataset Balancing:** The expert-labeled dataset was then re-balanced with a random sample of panels previously classified as "normal" to create a final, representative training dataset.

This methodology drastically reduces the manual labeling burden and demonstrates a practical path to creating large-scale, high-quality ground truth datasets. The final dataset was appropriately scaled and normalized with a robust scaler before being used for model training and testing.

## 5.4.2 Model Performance

The prepared dataset was used to train and test a suite of models directly within the application's UI. The models were chosen for their simplicity and inference speed, aligning with the goal of demonstrating pipeline viability. The performance of these models on the training and test sets is summarized in Table 5.4 and Table 5.5, respectively.

Table 5.4: Model performance on the **training set**.

Model	Accuracy (Weighted)	F1-Score (Weighted)	R <sup>2</sup>	MAE
RF Binary Classifier	0.996	0.996	N/A	N/A
RF Multiclass	0.998	0.998	N/A	N/A
NN Multiclass	0.995	0.995	N/A	N/A
LR Multiclass	0.979	0.979	N/A	N/A
NN Regression Multiclass	0.989	0.912	0.96	0.04
RF Regression Multiclass	0.994	0.956	0.98	0.02
XGBoost Regression Multiclass	0.997	0.963	0.98	0.01

Table 5.5: Model performance on the unseen **test set**.

Model	Accuracy (Weighted)	F1-Score (Weighted)	R <sup>2</sup>	MAE
RF Binary Classifier	0.921	0.720	N/A	N/A
RF Multiclass	0.925	0.794	N/A	N/A
NN Multiclass	0.938	0.828	N/A	N/A
LR Multiclass	0.865	0.712	N/A	N/A
NN Regression Multiclass	0.889	0.817	0.79	0.28
RF Regression Multiclass	0.917	0.856	0.85	0.19
XGBoost Regression Multiclass	0.918	0.872	0.88	0.13

Table 5.6: Representative confusion matrix for the XGBoost multiclass classifier on a balanced test set of 3,930 panels.

True Class	Predicted Class					Total
	Normal	Hot-Spot	PID/Multi-HS	Diode	Offline-Module	
<b>Normal</b>	<b>1953</b>	38	0	0	0	1991
<b>Hot-Spot</b>	95	<b>1110</b>	42	11	0	1258
<b>PID/Multi-HS</b>	8	75	<b>319</b>	5	0	407
<b>Diode</b>	0	28	12	<b>135</b>	4	179
<b>Offline-Module</b>	0	1	3	0	<b>91</b>	95
<b>Total</b>	2056	1252	376	151	95	<b>3930</b>

## 5.5 Synthesis of Experimental Findings

The experimental results presented in this chapter successfully validate the core hypotheses of this thesis. The panel detection pipeline, which combines classical computer vision techniques with novel refinement algorithms, achieved an outstanding average True Positive Rate of 99.42% across a diverse dataset of 13 PV plants. This performance confirms that for highly structured environments like PV installations, a computationally efficient, non-deep-learning approach can yield results comparable to the state-of-the-art while running on standard local hardware.

Furthermore, the anomaly classification experiments demonstrated the viability of using statistical thermal features for rapid and accurate defect identification. The good performance of models like the XGBoost Regressor (F1-Score of 0.872 on the test set) underscores that even the simple feature based approach is quite powerful. Critically, the entire experimental process showcased the system's dual capability: not only as a robust analysis tool but also as a powerful, semi-automated framework for the rapid creation of high-quality, expert-

labeled datasets from radiometric data. The implications of these findings, the innovations they represent, and the limitations of the current work will be critically examined in the following chapter.



## Chapter 6

# Discussion

This chapter provides a critical interpretation of the experimental results, contextualizing them within the research objectives and the state-of-the-art. The analysis focuses on the meaning behind the performance metrics, the core innovations of the developed system, its inherent limitations, and its specific positioning against existing solutions.

### 6.1 Interpretation of System Performance

The system's panel detection performance, with a 99.42% average True Positive Rate, strongly validates the decision to use a computationally efficient, classical computer vision pipeline. For a geometrically regular problem like locating PV modules, the implemented hybrid approach of template matching and spatial refinement proves to be a highly effective alternative to resource-intensive deep learning models. This result demonstrates that tailored, "local-first" algorithms can achieve state-of-the-art accuracy on standard hardware, a key objective of this thesis.

The anomaly classification results further support this philosophy. The success of models using simple statistical thermal features (e.g., XGBoost Regressor with a 0.872 F1-score on the test set) confirms that a full image-based CNN is not always necessary for accurate defect identification. The performance drop from the training to the test set is an expected outcome in machine learning, highlighting not a system failure, but the crucial need for diverse training data to improve model generalization. The system's strength, therefore, lies in its demonstrated ability to serve as a platform for creating and iteratively improving these models as more data becomes available.

### 6.2 Innovations and Contribution

The primary contribution of this research is the development of a solution that occupies a unique and under-served gap in the PV inspection landscape. Its key innovations are:

- **A Privacy-Preserving, Local-First Architecture:** In direct contrast to the dominant cloud-based commercial model, this work delivers a fully local application. This design choice grants the operator absolute data sovereignty, resolving the critical privacy and data ownership concerns associated with uploading sensitive operational data to third-party servers.
- **A Robust Orthomosaic-Centric Workflow:** By standardizing on high-quality orthomosaics as input, the system decouples analysis from the complex photogrammetry process. This modularity allows operators to leverage best-in-class tools for reconstruction and ensures a more reliable and accurate data foundation than systems that process raw video feeds directly.
- **A Dual-Purpose Analysis and Dataset Creation Framework:** The system functions not only as a diagnostic tool but also as a powerful framework for addressing the well-documented lack of high-quality, radiometric public datasets. Its semi-automated, expert-in-the-loop labeling capabilities provide a practical pathway for the rapid generation of large, verified datasets, a significant contribution to the broader research community.

### 6.3 Limitations and Future Directions

A critical evaluation of this work must acknowledge its limitations, which naturally point toward avenues for future research. The system's performance is fundamentally dependent on the quality of the input orthomosaics; artifacts from poor surveys or reconstructions can degrade accuracy. While the current single-threaded implementation is highly performant for most plants, its scalability for mega-scale installations could be enhanced through parallelization. Finally, the generalizability of any machine learning model is bound by its training data; future work should focus on using the tool to aggregate data from a wider variety of sites and technologies to build more universally robust models, and even better the implementation of Computer Vision models that use the segmented panels for defect classification, offering a much more robust approach that will probably be able to outperform the simple models trained on tabular data.

### 6.4 Comparison with the State of the Art

When positioned against existing solutions, the system developed in this thesis fills a clear gap. Compared to commercial platforms like Raptor Maps, it offers a fundamentally different value proposition: it is a *tool*, not a service, prioritizing data control, transparency, and immediate on-site results over a polished but opaque cloud service. In contrast to open-source proofs-of-concept like PV-Hawk, this thesis delivers a complete, user-friendly and

robust end-to-end system with a GUI, representing a significant step forward in usability and practical deployment. It successfully combines the functionality of a commercial product with the transparent, accessible, and privacy-centric philosophy of open-source software, without compromising on data acquisition standards and radiometric processing.



## Chapter 7

# Conclusions

The rapid global expansion of photovoltaic installations has created an urgent need for scalable, efficient, and reliable methods for plant inspection and maintenance. Performance losses due to undetected defects lead to billions of dollars in economic damages annually and challenge the long-term environmental sustainability of solar assets. The widening "O&M Gap," where capacity growth outpaces the growth of skilled maintenance personnel, underscores the unscalability of manual inspection and necessitates the development of automated solutions.

This challenge was addressed through the design, implementation, and validation of a *PV Anomaly Detection Suite*, a novel, end-to-end software solution for automated thermal anomaly detection. The system was architected with a "local-first" philosophy, ensuring all processing and data remain on the operator's machine, thereby guaranteeing absolute data sovereignty and privacy. By adopting an orthomosaic-centric workflow and integrating a hybrid of classical computer vision and modern machine learning techniques, the suite provides a robust, computationally efficient pipeline accessible to users without specialized GPU hardware.

The experimental validation, conducted on a diverse dataset of 13 PV plants, confirmed the system's efficacy. The panel detection module achieved an outstanding average True Positive Rate, demonstrating that tailored classical algorithms can rival the accuracy of deep learning models for structured environments. Furthermore, classification models trained on simple, extracted thermal features achieved good performance validating a rapid and efficient approach to anomaly identification.

Furthermore, this Thesis moves beyond a purely analytical function, providing an effective framework for the rapid, semi-automated creation of high-quality, radiometric, and expert-labeled datasets, a valuable resource that can fuel future research and development in the field.

Future work should focus on several key enhancements. System scalability for mega-scale plants can be improved through parallel processing, and workflow efficiency can be increased by automating some of the manual steps. Critically, the system's dataset creation

capabilities should be leveraged to train advanced computer vision models, such as YOLO like models, U-Nets or other Convolutional Neural Networks on the segmented panel images. This would enable classification based on learned spatial patterns, like the distinct signatures of PID or diode failures, moving beyond statistical features to achieve more nuanced and accurate defect identification. Concurrently, continued aggregation of data from diverse sites will be essential for building more universally robust models.

In conclusion, this thesis has successfully developed and validated a comprehensive solution that lowers the barriers to adopting advanced automated diagnostics. By providing an open, private, and effective tool, this work contributes to ensuring the long-term reliability, profitability, and sustainability of the global solar energy infrastructure.

# Bibliography

- [1] REN21. *RENEWABLES 2025 GLOBAL STATUS REPORT*. REN21, 2025. URL [https://www.ren21.net/gsr-2025/downloads/pdf/go/GSR\\_2025\\_G0\\_2025\\_Full\\_Report.pdf](https://www.ren21.net/gsr-2025/downloads/pdf/go/GSR_2025_G0_2025_Full_Report.pdf).
- [2] International Energy Agency. *Global energy review 2025*, 2025. URL <https://www.iea.org/reports/global-energy-review-2025>.
- [3] D. C. Jordan et al. *System-level performance and degradation of 21 gwdc of utility-scale pv plants in the united states*. Technical report, National Renewable Energy Laboratory (NREL), 2020. URL <https://docs.nrel.gov/docs/fy20osti/77257.pdf>.
- [4] Environmental Protection. *Examining the environmental impact of solar panels*, April 2025. URL <https://eponline.com/articles/2025/04/04/examining-the-environmental-impact-of-solar-panels.aspx>.
- [5] TST Engineering. *The hidden costs of solar farm underperformance: How equipment issues impact revenue*, March 2025. URL <https://www.tstpros.com/2025/03/05/the-hidden-costs-of-solar-farm-underperformance-how-equipment-issues-impact-revenue/>.
- [6] Solar Energy Engineering. *Failures & defects in pv systems: Typical methods for detecting defects and failures*, 2025. URL <https://www.study-solar.com/blog-article/failures-defects-in-pv-systems-typical-methods-for-detecting-defects-and-failures>.
- [7] Phoenix Strategy Group. *Ultimate guide to solar asset valuation with degradation*, 2025. URL <https://www.phoenixstrategy.group/blog/ultimate-guide-to-solar-asset-valuation-with-degradation>.
- [8] Encyclopedia MDPI. *Failures of the photovoltaic module components*. Online, 2023. URL <https://encyclopedia.pub/entry/56103>.
- [9] Thermography Services. *What is iec 62446-3:2017, and why is it important?*, 2025. URL <https://thermographyservices.co.uk/faq-items/what-is-iec-62446-3-2017-and-why-is-it-important/>.

- [10] Ali Thakfan and Yasser Bin Salamah. Artificial-intelligence-based detection of defects and faults in photovoltaic systems: A survey. *Energies*, 17(19), 2024. ISSN 1996-1073. doi: 10.3390/en17194807. URL <https://www.mdpi.com/1996-1073/17/19/4807>.
- [11] Mercom India. Solar industry loses \$4.6 billion in 2023 due to equipment underperformance, 2024. URL <https://www.mercomindia.com/solar-industry-lose-equipment-underperformance>.
- [12] pv magazine USA. U.s. solar facilities lost \$5,720 per mw to equipment underperformance in 2024, March 2025. URL <https://pv-magazine-usa.com/2025/03/05/u-s-solar-facilities-lost-5720-per-mw-to-equipment-underperformance-in-2024/>.
- [13] Fahad Almubaddel et al. The impact of pv panel degradation rate, initial system efficiency, and interest rate on the levelized cost of energy for pv projects: Saudi arabia as a benchmark. *ResearchGate*, 2024. URL [https://www.researchgate.net/publication/385934261\\_The\\_Impact\\_of\\_PV\\_Panel\\_Degradation\\_Rate\\_Initial\\_System\\_Efficiency\\_and\\_Interest\\_Rate\\_on\\_the\\_Levelized\\_Cost\\_of\\_Energy\\_for\\_PV\\_Projects\\_Saudi\\_Arabia\\_as\\_a\\_Benchmark](https://www.researchgate.net/publication/385934261_The_Impact_of_PV_Panel_Degradation_Rate_Initial_System_Efficiency_and_Interest_Rate_on_the_Levelized_Cost_of_Energy_for_PV_Projects_Saudi_Arabia_as_a_Benchmark).
- [14] Fahad Almubaddel et al. The impact of pv panel degradation rate, initial system efficiency, and interest rate on the levelized cost of energy for pv projects. *Sustainability*, 16(22):10012, 2024. URL <https://www.mdpi.com/2071-1050/16/22/10012>.
- [15] Raptor Maps. 5 benefits of using drones for pv system inspections, 2021. URL <https://raptormaps.com/blog-posts/5-benefits-of-using-drones-for-pv-system-inspections>.
- [16] U.S. Energy Information Administration (EIA). Solar energy and the environment, 2025. URL <https://www.eia.gov/energyexplained/solar/solar-energy-and-the-environment.php>.
- [17] Student. Life-cycle assessment of photovoltaic systems. Technical report, Lund University Publications, 2014. URL <https://lup.lub.lu.se/student-papers/record/7471285/file/7471400.pdf>.
- [18] U.S. Environmental Protection Agency (EPA). End-of-life solar panels: Regulations and management, 2025. URL <https://www.epa.gov/hw/end-life-solar-panels-regulations-and-management>.
- [19] Various Authors. Review on life cycle assessment of solar photovoltaic panels. *ResearchGate*, 2020. URL [https://www.researchgate.net/publication/338384189\\_Review\\_on\\_Life\\_Cycle\\_Assessment\\_of\\_Solar\\_Photovoltaic\\_Panels](https://www.researchgate.net/publication/338384189_Review_on_Life_Cycle_Assessment_of_Solar_Photovoltaic_Panels).

- [20] Various Authors. Life cycle assessment for photovoltaic structures—comparative study of rooftop and free-field pv applications. *Sustainability*, 15(18):13692, 2023. URL <https://www.mdpi.com/2071-1050/15/18/13692>.
- [21] Various Authors. Impact assessment of photovoltaic panels with life cycle analysis techniques, 2024. URL <https://www.preprints.org/manuscript/202402.1160/v1/download>.
- [22] S. D’Addona et al. Photovoltaic hot spot analysis for cells with various reverse-bias characteristics through electrical and thermal simulation. *Illinois Experts*, 2018. URL <https://experts.illinois.edu/en/publications/photovoltaic-hot-spot-analysis-for-cells-with-various-reverse-bia>.
- [23] Various Authors. Investigations of delamination fault in photovoltaic panel using... *International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering*, 2017. URL [https://www.ijareeie.com/upload/2017/april/66\\_Investigations.pdf](https://www.ijareeie.com/upload/2017/april/66_Investigations.pdf).
- [24] Sitemark Help Center. Thermal anomaly types, 2025. URL <https://support.sitemark.com/en/articles/4643878-thermal-anomaly-types>.
- [25] HeatSpring Magazine. Identifying issues on installed pv systems: A thermal imaging guide for solar professionals, 2025. URL <https://blog.heatspring.com/identifying-issues-on-installed-pv-systems-a-thermal-imaging-guide-for-solar-professionals/>.
- [26] AJG United States. Micro-fractures in solar modules: Causes, detection and prevention, 2025. URL <https://www.ajg.com/news-and-insights/micro-fractures-in-solar-modules-causes-detection-and-prevention/>.
- [27] BrightSpot Automation. Detection and impact of cracks hidden near interconnect wires in silicon solar cells. Technical report, 2023. URL <https://brightspotautomation.com/wp-content/uploads/2023/07/Hidden-Cracks-PVSC2023-BrightSpot.pdf>.
- [28] Various Authors. A review on detection of solar pv panels failures using image processing techniques. *ResearchGate*, 2024. URL [https://www.researchgate.net/publication/379064824\\_A\\_Review\\_on\\_Detection\\_of\\_Solar\\_PV\\_Panels\\_Failures\\_Using\\_Image\\_Processing\\_Techniques](https://www.researchgate.net/publication/379064824_A_Review_on_Detection_of_Solar_PV_Panels_Failures_Using_Image_Processing_Techniques).
- [29] A1 SolarStore. Most common solar panel defects and how to deal with them, 2025. URL <https://a1solarstore.com/blog/most-common-solar-panel-defects-and-how-to-deal-with-them.html>.

- [30] M. B. Diaz et al. Micro-cracks detection in photo-voltaic cells by infrared thermography. *ResearchGate*, 2016. URL [https://www.researchgate.net/publication/305183609\\_Micro-cracks\\_detection\\_in\\_photo-voltaic\\_cells\\_by\\_infrared\\_thermography](https://www.researchgate.net/publication/305183609_Micro-cracks_detection_in_photo-voltaic_cells_by_infrared_thermography).
- [31] Maysun Solar. Snail trails on the solar panels, 2025. URL <https://www.maysunsolar.com/blog-snail-trails-on-the-solar-panels/>.
- [32] PV-Manufacturing.org. Snail trails, 2025. URL <https://pv-manufacturing.org/snail-trails/>.
- [33] S. Kajari-Schroder et al. A defect formation as snail trails in photovoltaic modules. *ResearchGate*, 2015. URL [https://www.researchgate.net/publication/282409505\\_A\\_Defect\\_Formation\\_as\\_Snail\\_Trails\\_in\\_Photovoltaic\\_Modules](https://www.researchgate.net/publication/282409505_A_Defect_Formation_as_Snail_Trails_in_Photovoltaic_Modules).
- [34] M. Koehl et al. Understanding snail trails on pv modules. Technical report, PV-Tech, 2014. URL <https://www.pv-tech.org/wp-content/uploads/legacy-publication-pdfs/84b14d2965-understanding-snail-trails-on-pv-modules.pdf>.
- [35] ECOPROGETTI. Snail track effect on photovoltaic modules, 2025. URL <https://ecoprogetti.com/snail-track-effect-photovoltaic-modules/>.
- [36] Various Authors. Photovoltaic failure diagnosis using imaging techniques and electrical characterization. *EPJ Photovoltaics*, 2024. URL [https://www.epj-pv.org/articles/epjpv/full\\_html/2024/01/pv20240020/pv20240020.html](https://www.epj-pv.org/articles/epjpv/full_html/2024/01/pv20240020/pv20240020.html).
- [37] Various Authors. Analysis and characterization of pv module defects by thermographic inspection. *ResearchGate*, 2019. URL [https://www.researchgate.net/publication/333505248\\_Analysis\\_and\\_characterization\\_of\\_PV\\_module\\_defects\\_by\\_thermographic\\_inspection](https://www.researchgate.net/publication/333505248_Analysis_and_characterization_of_PV_module_defects_by_thermographic_inspection).
- [38] IEA-PVPS. Review of failures of photovoltaic modules. Technical report, International Energy Agency Photovoltaic Power Systems Programme, 2014. URL [https://iea-pvps.org/wp-content/uploads/2020/01/IEA-PVPS\\_T13-01\\_2014\\_Review\\_of\\_Failures\\_of\\_Photovoltaic\\_Modules\\_Final.pdf](https://iea-pvps.org/wp-content/uploads/2020/01/IEA-PVPS_T13-01_2014_Review_of_Failures_of_Photovoltaic_Modules_Final.pdf).
- [39] Various Authors. Delamination between the front cover glass and encapsulant. *ResearchGate*, 2021. URL [https://www.researchgate.net/figure/Delamination-between-the-front-cover-glass-and-encapsulant-a-Early-state-b-Massive\\_fig1\\_348854638](https://www.researchgate.net/figure/Delamination-between-the-front-cover-glass-and-encapsulant-a-Early-state-b-Massive_fig1_348854638).
- [40] Sinovoltaics. Eva browning, 2025. URL <https://sinovoltaics.com/learning-center/materials/eva-browning/>.

- [41] Shashwata. Encapsulant discoloration in photovoltaic modules. Technical report, EE IIT Bombay, 2015. URL [https://www.ee.iitb.ac.in/~anilk/Shashwata\\_thesis.pdf](https://www.ee.iitb.ac.in/~anilk/Shashwata_thesis.pdf).
- [42] Various Authors. Browning area percentage (%) versus browning aspect ratio (ar) plot for both overall. *ResearchGate*, 2021. URL [https://www.researchgate.net/figure/Browning-area-percentage-versus-browning-aspect-ratio-AR-plot-for-both-overall\\_fig4\\_352090143](https://www.researchgate.net/figure/Browning-area-percentage-versus-browning-aspect-ratio-AR-plot-for-both-overall_fig4_352090143).
- [43] Envista Forensics. Pv module reliability issues: Understanding backsheets degradation and claim disputes, 2025. URL <https://www.envistaforensics.com/knowledge-center/insights/articles/pv-module-reliability-issues-understanding-backsheet-degradation-and-claim-disputes/>.
- [44] J. A. del Cueto et al. Differential degradation patterns of photovoltaic backsheets at the array level. Technical report, OSTI.GOV, 2018. URL <https://www.osti.gov/servlets/purl/1424906>.
- [45] Various Authors. Review of unmanned ground vehicles for pv plant inspection. *Murdoch University Research Portal*, 2022. URL [https://researchportal.murdoch.edu.au/view/pdfCoverPage?instCode=61MUN\\_INST&filePid=13174582830007891&download=true](https://researchportal.murdoch.edu.au/view/pdfCoverPage?instCode=61MUN_INST&filePid=13174582830007891&download=true).
- [46] Various Authors. Tough break: Many factors make glass breakage more likely. Technical report, National Renewable Energy Laboratory (NREL), 2025. URL <https://docs.nrel.gov/docs/fy25osti/91695.pdf>.
- [47] Saur Energy International. Top 5: Factors responsible for glass breakage in solar modules, 2023. URL <https://www.saurenergy.com/solar-energy-blog/top-5-factors-responsible-for-glass-breakage-in-solar-modules>.
- [48] Various Authors. Analysis and characterization of pv module defects by thermographic inspection. *Scielo*, 2019. URL [http://www.scielo.org.co/scielo.php?script=sci\\_arttext&pid=S0120-62302019000400092](http://www.scielo.org.co/scielo.php?script=sci_arttext&pid=S0120-62302019000400092).
- [49] Optris. Steady state thermography for non-intrusive solar module defect detection, 2025. URL <https://optris.com/us/application/solar/steady-state-thermography-for-non-intrusive-defect-detection-for-solar-modules/>.
- [50] The Drone Life NJ. What is an iec compliant drone solar inspection?, 2025. URL <https://thedronelifenj.com/what-is-an-iec-compliant-drone-solar-inspection/>.
- [51] Solar Best Practices. Detail - common tests and inspections, 2025. URL <https://solarbestpractices.com/guidelines/detail/common-tests-and-inspections>.

- [52] vHive. Understanding thermal imaging inspections: The ultimate beginner's guide, 2025. URL <https://www.vhive.ai/thermal-imaging-inspections/>.
- [53] Advexure. Faster fault detection: Thermal drones for solar farm inspections, 2025. URL <https://advexure.com/blogs/news/faster-fault-detection-thermal-drones-for-solar-farm-inspections>.
- [54] DJI. Shining a light on drone technology in the solar energy revolution, 2025. URL <https://enterprise-insights.dji.com/user-stories/drone-solar-inspection-saves-time>.
- [55] FEDS Group. Leveraging drone technology for more efficient solar panel inspections, 2025. URL <https://feds.group/leveraging-drone-technology-for-more-efficient-solar-panel-inspections>.
- [56] Air Provence. Photovoltaic panels : the advantages of drone inspection, 2025. URL <https://www.air-provence.com/en/photovoltaic-panels-advantages-drone-inspection>.
- [57] DJI. Solar asset inspections with aeroprotechnik, 2025. URL <https://enterprise-insights.dji.com/user-stories/drone-solar-inspections-aeroprotechnik>.
- [58] Various Authors. A review of ir thermography applied to pv systems. *ResearchGate*, 2013. URL [https://www.researchgate.net/publication/254039824\\_A\\_review\\_of\\_IR\\_thermography\\_applied\\_to\\_PV\\_systems](https://www.researchgate.net/publication/254039824_A_review_of_IR_thermography_applied_to_PV_systems).
- [59] Various Authors. Automated detection-classification of defects on photo-voltaic modules assisted by thermal drone inspection. *MATEC Web of Conferences*, 2021. URL [https://www.matec-conferences.org/articles/mateconf/pdf/2021/18/mateconf\\_iceaf2021\\_03015.pdf](https://www.matec-conferences.org/articles/mateconf/pdf/2021/18/mateconf_iceaf2021_03015.pdf).
- [60] Various Authors. Detection of visual faults in photovoltaic modules using a stacking ensemble approach. *Scientific Reports*, 14(1), 2024. URL <https://pmc.ncbi.nlm.nih.gov/articles/PMC10944267/>.
- [61] Various Authors. Automatic inspection of photovoltaic power plants using aerial... *Energies*, 15(6):2055, 2022. URL <https://www.mdpi.com/1996-1073/15/6/2055>.
- [62] Various Authors. Pv fault detection using positive unlabeled learning. *Applied Sciences*, 11(12):5599, 2021. URL <https://www.mdpi.com/2076-3417/11/12/5599>.
- [63] NREL. Open data sets for assessing photovoltaic system reliability. Technical report, National Renewable Energy Laboratory, 2025. URL <https://docs.nrel.gov/docs/fy25osti/95519.pdf>.

- [64] PK Darabi. Thermal solar pv anomaly detection dataset, 2025. URL <https://www.kaggle.com/datasets/pkdarabi/solarpanel>.
- [65] Marcos Gabriel. Photovoltaic system thermography, 2025. URL <https://www.kaggle.com/datasets/marcosgabriel/photovoltaic-system-thermography>.
- [66] Various Authors. Thermal imaging dataset for hotspot detection on solar panels: Impact of bird droppings on efficiency, 2023. URL <https://data.mendeley.com/datasets/8sxfmrpfpv>.
- [67] Various Authors. Photovoltaic module dataset for automated fault detection and analysis in large photovoltaic systems using photovoltaic module fault detection, 2024. URL <https://data.mendeley.com/datasets/5ssmfprpc>.
- [68] Alicja. Pv panel defect dataset, 2025. URL <https://www.kaggle.com/datasets/alicjalena/pv-panel-defect-dataset>.
- [69] Various Authors. Radiometric infrared thermography of solar photovoltaic systems: An explainable predictive maintenance approach for remote aerial diagnostic monitoring. *Drones*, 7(3):53, 2024. URL <https://www.mdpi.com/2624-6511/7/3/53>.
- [70] Data.gov. radiometrics - dataset, 2025. URL <https://catalog.data.gov/dataset/?tags=radiometrics>.
- [71] Various Authors. Detection and analysis of deteriorated areas in solar pv modules using unsupervised sensing algorithms and 3d augmented reality. *Scientific Reports*, 14(1), 2024. URL <https://pmc.ncbi.nlm.nih.gov/articles/PMC10963330/>.
- [72] Josep Blazquez Folch. Detection of artifacts in infrared images and video of pv solar panels. Technical report, UPCommons, 2024. URL <https://upcommons.upc.edu/bitstream/handle/2117/420925/Detection%20of%20artifacts%20in%20infrared%20images%20and%20video%20of%20PV%20solar%20panels%20-%20Josep%20Blazquez%20Folch.pdf?sequence=3>.
- [73] Various Authors. Multi-resolution dataset for photovoltaic panel segmentation from satellite and aerial imagery. *Earth System Science Data*, 13:5389–5403, 2021. URL <https://essd.copernicus.org/articles/13/5389/2021/>.
- [74] Various Authors. Photovoltaic module dataset for automated fault detection and analysis in large photovoltaic systems using photovoltaic module fault detection. *Scientific Data*, 11(1), 2024. URL <https://pmc.ncbi.nlm.nih.gov/articles/PMC11683254/>.
- [75] binyisu. Pvel-ad: Photovoltaic cell defect detection, 2025. URL <https://github.com/binyisu/PVEL-AD>.

- [76] Various Authors. Drone-based thermal imaging inspection of solar energy equipment. In *World Conference of the Society for Industrial and Systems Engineering, 2022*. URL <https://www.ieworldconference.org/content/WP2022/Papers/49-GDRKM-CC-22.pdf>.
- [77] Reddit User. Alternatives to dronedeploy, 2025. URL [https://www.reddit.com/r/UAVmapping/comments/og8qz/alternatives\\_to\\_dronedeploy/](https://www.reddit.com/r/UAVmapping/comments/og8qz/alternatives_to_dronedeploy/).
- [78] Lukas Bommers. Pv-hawk: Tool for the extraction and mapping of photovoltaics modules from ir drone videos of utility-scale pv plants, 2025. URL <https://github.com/LukasBommers/PV-Hawk>.
- [79] Roboflow Blog. What is template matching? an introduction., 2025. URL <https://blog.roboflow.com/template-matching/>.
- [80] Wikipedia. Hough transform, 2025. URL [https://en.wikipedia.org/wiki/Hough\\_transform](https://en.wikipedia.org/wiki/Hough_transform).
- [81] J. T. St-Onge et al. Automatic processing and solar cell detection in ... *arXiv*, 2018. URL <https://arxiv.org/pdf/1807.10820>.
- [82] Various Authors. Ai-based pv panels inspection using an advanced yolo algorithm. *ResearchGate*, 2025. URL [https://www.researchgate.net/publication/382266791\\_AI-Based\\_PV\\_Panels\\_Inspection\\_using\\_an\\_Advanced\\_YOLO\\_Algorithm](https://www.researchgate.net/publication/382266791_AI-Based_PV_Panels_Inspection_using_an_Advanced_YOLO_Algorithm).
- [83] Various Authors. Detecting defects in solar panels using the yolo v10 and v11 algorithms. *Electronics*, 14(2):344, 2025. URL <https://www.mdpi.com/2079-9292/14/2/344>.
- [84] Various Authors. St-yolo: A defect detection method for photovoltaic modules ... *PubMed Central*, 2025. URL <https://pmc.ncbi.nlm.nih.gov/articles/PMC11637253/>.
- [85] Various Authors. An overview of cnn-based image analysis in solar cells ... *Applied Sciences*, 15(10):5511, 2025. URL <https://www.mdpi.com/2076-3417/15/10/5511>.
- [86] Various Authors. A cnn-architecture-based photovoltaic cell fault classification method using thermographic images. *Energies*, 16(9):3749, 2025. URL <https://www.mdpi.com/1996-1073/16/9/3749>.
- [87] Various Authors. Deep learning-based recognition and classification of soiled photovoltaic modules using halcon software for solar cleaning robots. *PubMed Central*, 2025. URL <https://pmc.ncbi.nlm.nih.gov/articles/PMC11902315/>.

- [88] Various Authors. An svm-based anomaly detection method for power system ... *Processes*, 13(2):549, 2025. URL <https://www.mdpi.com/2227-9717/13/2/549>.
- [89] Various Authors. Faults detection and diagnosis of a large-scale pv system by ... *Energies*, 18(10):2482, 2025. URL <https://www.mdpi.com/1996-1073/18/10/2482>.
- [90] Various Authors. Machine learning-based detection of solar panel surface defects using deep features from inceptionv3. *ResearchGate*, 2025. URL [https://www.researchgate.net/publication/393711156\\_Machine\\_Learning-Based\\_Detection\\_of\\_Solar\\_Panel\\_Surface\\_Defects\\_Using\\_Deep\\_Features\\_from\\_InceptionV3](https://www.researchgate.net/publication/393711156_Machine_Learning-Based_Detection_of_Solar_Panel_Surface_Defects_Using_Deep_Features_from_InceptionV3).
- [91] Various Authors. An unsupervised monitoring procedure for detecting anomalies in photovoltaic systems using a one-class support vector machine. *ResearchGate*, 2025. URL [https://www.researchgate.net/publication/329788025\\_An\\_unsupervised\\_monitoring\\_procedure\\_for\\_detecting\\_anomalies\\_in\\_photovoltaic\\_systems\\_using\\_a\\_one-class\\_Support\\_Vector\\_Machine](https://www.researchgate.net/publication/329788025_An_unsupervised_monitoring_procedure_for_detecting_anomalies_in_photovoltaic_systems_using_a_one-class_Support_Vector_Machine).
- [92] Various Authors. Unsupervised machine learning for anomaly detection in solar power generation: Comparative insight. *ResearchGate*, 2025. URL [https://www.researchgate.net/publication/384052746\\_Unsupervised\\_Machine\\_Learning\\_for\\_Anomaly\\_Detection\\_in\\_Solar\\_Power\\_Generation\\_Comparative\\_Insight](https://www.researchgate.net/publication/384052746_Unsupervised_Machine_Learning_for_Anomaly_Detection_in_Solar_Power_Generation_Comparative_Insight).
- [93] Lukas Bommers. Tutorial — pv hawk documentation, 2025. URL <https://lukasbommers.github.io/PV-Hawk/tutorial.html>.
- [94] Lukas Bommers. Pv hawk — pv hawk documentation, 2025. URL <https://lukasbommers.github.io/PV-Hawk/>.
- [95] Lukas Bommers. Pv-hawk-viewer, 2025. URL <https://github.com/LukasBommers/PV-Hawk-Viewer>.
- [96] Michiel Vlaminck, Rugen Heidbuchel, Wilfried Philips, and Hiep Luong. Region-based cnn for anomaly detection in pv power plants using aerial imagery. *Sensors*, 22(3), 2022. ISSN 1424-8220. doi: 10.3390/s22031244. URL <https://www.mdpi.com/1424-8220/22/3/1244>.
- [97] Gonzalo Luzardo, Michiel Vlaminck, Dionysios Lefkaditis, Wilfried Philips, and Hiep Luong. Gps-assisted feature matching in aerial images with highly repetitive patterns. In *2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS*, pages 8348–8351. IEEE, 2021.
- [98] binyisu. Pvel-ad: Photovoltaic cell defect detection, 2025. URL <https://github.com/binyisu/PVEL-AD>.

- [99] cdeldon. thermography: Automatic detection of defected solar panel modules, 2025. URL <https://github.com/cdeldon/thermography>.
- [100] DroneDeploy. Dronedeploy robotic industrial inspection: Automated asset monitoring & inspection software, 2025. URL <https://www.dronedeploy.com/product/robotic-industrial-inspection>.
- [101] V7 Labs. How raptor maps uses v7 darwin to reduce time-to-model by 83%, 2025. URL <https://www.v7labs.com/customer-stories/raptor-maps>.
- [102] DJI Enterprise. Sitemark: Ai & robotics software for solar -cosystem solution catalogue, 2025. URL <https://enterprise.dji.com/ecosystem/sitemark>.
- [103] Scopito. The solar inspection software of your dreams, 2025. URL <https://scopito.com/solar-pv-inspection-software/>.
- [104] SolaR. Corporate overview. Technical report, 2025. URL <https://f.hubspotusercontent30.net/hubfs/6021737/Marketing%20Materials/Corporate%20Brochure/Corporate%20Overview.pdf>.
- [105] Sitemark Help Center. How long does it take to process my data?, 2025. URL <https://support.sitemark.com/en/articles/5633717-how-long-does-it-take-to-process-my-data>.
- [106] DroneDeploy. Privacy policy, 2025. URL <https://www.dronedeploy.com/legal/privacy>.
- [107] Raptor Technologies. Privacy policy, 2025. URL <https://apps.raptortech.com/About/Privacy>.
- [108] Drone Universities. The ethics of drone use: Privacy and legal considerations, 2025. URL <https://www.droneuniversities.com/the-ethics-of-drone-use-privacy-and-legal-considerations/>.

