SCIENTIFIC CULTURE
www.sci-cult.com

SCIENTIFIC CULTURE, Vol. 12, No. 2.1, (2026), pp. 10576-10589
Open Access. Online & Print

DOI: 10.5281/zenodo.20067683

HIGH-FIDELITY SOLAR PANELS DEFECT DIAGNOSIS: A YOLOV9 FRAMEWORK LEVERAGING A LARGE CURATED THERMAL DATASET

Zeeshan Haider[1*], Ahmad Taher Azar[2] and Yasmin Adel Hagag[3]

[1]Automated Systems Computing Lab (ASCL), Prince Sultan University, Riyadh 11586, Saudi Arabia. Email: zhaider@psu.edu.sa
[2]College of Computer and Information Sciences, Prince Sultan University, Riyadh 11586, Saudi Arabia, Email: aazar@psu.edu.sa
[3]Faculty of Engineering, Electrical Power and Machine Department, Cairo University, Giza, 12613, Egypt. Email: yasmin.adel.hagag@gmail.com

Received: 06/04/2026
Accepted: 29/04/2026

Corresponding Author: Zeeshan Haider (zhaider@psu.edu.sa)

ABSTRACT

Automated defect detection in photovoltaic (PV) modules is a critical component of renewable energy infrastructure maintenance. However, the efficacy of such systems is frequently limited by the quality of training data rather than architectural constraints. A recurring issue in public repositories is data fragmentation and inconsistent labelling, which impedes the development of generalized models. This study introduces a data-centric methodology focused on the curation and unification of a comprehensive PV thermal dataset from multiple sources. To resolve taxonomic discrepancies such as conflicting definitions of" hotspots" and" cracks", a binary classification protocol (Defect vs. Background) is adopted. This approach eliminates label noise, establishing a consistent foundation for pre-training. The YOLOv9 architecture and its variants (Gelan-c, Gelan-e, Yolov9-c, and Yolov9-e) are employed to validate the dataset's integrity. Empirical analysis confirms that this unified strategy effectively captures fundamental defect characteristics, providing a robust baseline for future PV monitoring applications.

KEYWORDS: Photovoltaic Defects; Defects Detection; Renewable Energy; Thermography; Yolov9; Solar Panel Defects.

1. INTRODUCTION

The escalation of global energy requirements has positioned photovoltaic (PV) systems as a primary component of sustainable power generation infrastructures (El-Rashidy, 2022; GUO et al., 2023; Hagag, ElSobky, Ibrahim, Azar, & Haider, 2025; Kesavan, Subramaniam, Almakhles, & Selvam, 2023; Martínez-Duart & Hernandez Moro, 2013; Spajic´, Talajic´, & Mrsˇic´, 2024). With the expansion of these installations, the maintenance of operational efficiency presents a substantial challenge. Structural impairments and electrical anomalies are known to compromise both performance and longevity (Smida, Azar, Sakly, & Hameed, 2024; Spajic´ et al., 2024), necessitating the deployment of reliable diagnostic mechanisms. Conventional manual inspections, characterized by labor intensity and susceptibility to error, are increasingly being supplemented by non-invasive techniques such as infrared thermography (IRT), which facilitates the detection of thermal irregularities indicative of underlying faults (de Oliveira, Aghaei, & Ruther, 2022; Zhu, Zhang, & Deng, 2013).

Thermographic analysis elucidates the operational status of PV modules through surface temperature mapping, where localized heating typically signals malfunction (Akram, Li, Jin, & Chen, 2022). Such thermal signatures are associated with a spectrum of defects, including micro-cracks, grid fractures, potential induced degradation (PID), and bypass diode failures. For instance, current obstruction caused by fissures may induce localized temperature elevation, whereas partial shading or soiling can precipitate hotspots, accelerating module degradation. While expert analysis remains valid, its scalability is limited in the context of extensive solar farms. Consequently, automation via computer vision and machine learning has emerged as a viable alternative (Høiaas et al., 2022).

Convolutional Neural Networks (CNNs) have demonstrated particular efficacy in the automated analysis of thermal imagery for fault identification (Mazen, Seoud, & Shaker, 2023; Rogotis, Ioannidis, Tsolakis, Tzovaras, & Likothanassis, 2014; Spajic´ et al., 2024). The interpretation of thermographic data is complicated by environmental variables such as reflection, shadowing, and meteorological fluctuations (Akram et al., 2022). Distinguishing genuine defects from environmental noise requires robust analytical frameworks. Deep learning methodologies, specifically CNNs, offer the capacity to extract complex features from raw data, thereby decoding the intricate thermal patterns associated with PV defects (Ab Hamid et al., 2024; Ren, Yu, Li, & Zhang, 2020; Spajic´ et al., 2024; Yang et al., 2019). Numerous studies have corroborated the high accuracy of CNNs in PV fault detection, often surpassing traditional image processing techniques (Ameerdin, Jamaluddin, Shukor, Kamaruzaman, & Mohamad, 2024; Han Wang, Chen, Li, & Piao, 2023; Hong et al., 2022; Tajwar et al., 2021; J. Wang, Zhang, & Zheng, 2023; Wu & Hao, 2023).

However, the predominant research focus has been model-centric, prioritizing architectural modifications over the optimization of the training data itself. A significant impediment to the deployment of CNNs in thermal PV inspection is the scarcity of large, diverse, and accurately annotated datasets (Gevorgian, Pernigotto, & Gasparella, 2024). Variations in sunlight intensity and ambient temperature further exacerbate the variability of fault presentation (Bosman, Leon-Salas, Hutzel, & Soto, 2020). Certain defects, such as micro-cracks, exhibit faint thermal signatures that challenge both human and automated detection capabilities (Thakfan & Bin Salamah, 2024).

In this study, a Data-Centric AI perspective is adopted. Rather than proposing a novel architecture, the primary contribution lies in the curation, validation, and unification of a comprehensive pre-training dataset derived from disparate sources. A critical issue identified in existing public repositories is label inconsistency, where varying nomenclatures such as hotspot, crack, and overheating are applied to similar phenomena. This is addressed through the implementation of a strict binary classification protocol: Defect vs. Background. This standardization reduces label noise, establishing a robust foundation for subsequent fine-tuning. The YOLOv9 family (Gelan-c, Gelan-e, Yolov9-c, Yolov9-e) is utilized as a validation instrument to assess the quality and learnability of this unified dataset.

The remainder of this paper is structured as follows: Section 2 provides a review of PV defect detection and deep learning. Section 3 details the data unification process and validation models. Section 4 outlines the experimental configuration. Section 5 presents the performance analysis. Finally, Section 6 discuss broader implications and summarizes the findings.

2. PV DEFECT DETECTION - RELATED WORK

The rapid expansion of the photovoltaic (PV) industry to address global renewable energy requirements (Byrne & Kurdgelashvili, 2011; Hoffmann, 2006) has rendered scalable maintenance strategies indispensable. Infrared thermography (IRT) has been established as the standard for rapid, non-destructive diagnostics (Glavasˇ, Vuko-bratovic´, Primorac, & Mustran, 2017). However, the efficacy of automated detection systems is frequently compromised by inconsistencies in the underlying training data. Consequently, contemporary research is transitioning from purely architectural modifications toward data-centric methodologies that prioritize the construction of robust training datasets.

Historically, the analysis of thermal imagery relied upon conventional image processing and elementary machine learning techniques (Kendig, Alers, & Shakouri, 2010). These approaches utilized handcrafted features such as texture statistics and temperature gradients to identify anomalies (Tsai, Chang, & Chao, 2010). A significant limitation of these shallow models was their susceptibility to failure under variable environmental conditions, including fluctuating irradiance and wind speed. To address these challenges, spatio-temporal techniques were introduced. For instance, the Panel Energy Image (PEI) method (Rogotis et al., 2014) was developed to differentiate true thermal defects from transient environmental noise.

The advent of Deep Learning, particularly Convolutional Neural Networks (CNNs), marked a significant shift by eliminating the requirement for manual feature engineering. These networks are capable of learning hierarchical patterns directly from raw thermal data (Han et al., 2023; J. Wang et al., 2023; Wu & Hao, 2023). Various architectures have demonstrated efficacy; for example, ResNet-50 has been utilized to classify anomalies in aerial videography with high accuracy (Bommes et al., 2021). Such automation provides tangible economic advantages, including reduced inspection durations and minimized energy yield losses (Spajic´ et al., 2024). Furthermore, VGG-16 models have been shown to robustly detect faults under severe climatic conditions (Kellil, Aissat, & Mellit, 2023). Real-time systems have also been developed to address prevalent issues such as partial shading and diode failures (Mellit, 2022).

Despite these advancements, a substantial obstacle persists: the fragmentation and inconsistency of public datasets. Studies attempting granular classification (e.g., distinguishing eleven defect types) frequently encounter conflicting labeling schemas (Alves, de Deus Junior, Marra, & Lemos, 2021). Terminological discrepancies, where a defect is labeled hotspot in one dataset and overheating or crack in another, are common. In this study, a data-centric solution is proposed to mitigate this issue. Rather than contending with noisy, granular labels, diverse datasets are aggregated into a single, comprehensive pre-training corpus utilizing a simplified binary protocol: Defect vs. Background. This strategy reduces label noise, establishing a consistent foundation that facilitates subsequent fine-tuning for specific tasks.

To evaluate the quality of this unified dataset, the YOLO series serves as the benchmark. YOLO models have undergone significant evolution, progressing from the original single-stage YOLOv1 (Redmon, Divvala, Girshick, & Farhadi, 2016) to the multi-scale enhancements of YOLOv2 (Redmon & Farhadi, 2017) and YOLOv3 (Redmon & Farhadi, 2018). Subsequent iterations, such as YOLOv4 (Bochkovskiy, Wang, & Liao, 2020) and YOLOv5 (Jocher et al., 2020), introduced improved feature aggregation, while recent versions, YOLO-R (C.- Y. Wang, Yeh, & Liao, 2021), YOLOvX (Ge, Liu, Wang, Li, & Sun, 2021), YOLOv6 (Li et al., 2022), and YOLOv7 (C. Wang, Bochkovskiy, & Liao, 2023) have focused on scaling and computational efficiency. YOLOv8 (Ter-ven, Co´rdova-Esparza, & Romero-Gonza´lez, 2023) and YOLO-NAS (Deci-Team, 2023) further refined architectural parameters and inference speed. In this work, the state-of-the-art YOLOv9 (C.-Y. Wang, Yeh, & Liao, 2024), which incorporates Programmable Gradient Information (PGI) to preserve input data integrity, is employed. The utilization of YOLOv9 is not intended to claim architectural novelty, but rather to serve as a reliable instrument for validating the quality of the unified dataset.

Table 1: Five Curated Datasets for PV Defect Detection Using Thermal Images.

Name	Train	Validation	Test	Source
Guangfu Pid Dataset	2647	234	–	(shuju, 2024)
Tunghai University Dataset	11394	121	80	(University, 2023)
Panel Dataset	1940	184	95	(S, 2023)
Thermal Defects Dataset	4788	262	263	(Solveview, 2023)
Beatrix Vision Dataset	1425	121	54	(Redefiningeducation, 2023)
Total	22194	922	492	

3. METHODOLOGY

This section delineates the data-centric methodology employed to construct a robust pre-training dataset for PV thermal defect detection. The procedural framework encompasses three primary components: the aggregation and unification of heterogeneous public datasets, the utilization of the YOLOv9 architecture for validation, and the evaluation across multiple model variants.

3.1. Curation And Unification of the Pre-Training Dataset

The central contribution of this research is the generation of a comprehensive, unified dataset, synthesized from five distinct public repositories. As detailed in Table 1, this consolidated corpus comprises 22,194 training images, a volume significantly exceeding typical datasets in this domain. The primary objective is to rectify the fragmentation and labeling inconsistencies that characterize individual sources.

The integration of these datasets presented challenges due to divergent labeling taxonomies. For instance, the Guangfu dataset (shuju, 2024) categorizes defects into PID, hotspots, and bypass-diode failures (Figure 1), whereas the Panel dataset (S, 2023) utilizes labels such as module-missing and vegetation occlusion (Figure 2). Similarly, the Thermal Defects (Solveview, 2023) and Beatrix Vision (Redefiningeducation, 2023) datasets predominantly focus on hotspots and PID (Figures 3 and 4), and the Tunghai University dataset (University, 2023) employs numerical codes rather than descriptive nomenclature (Figure 5)..

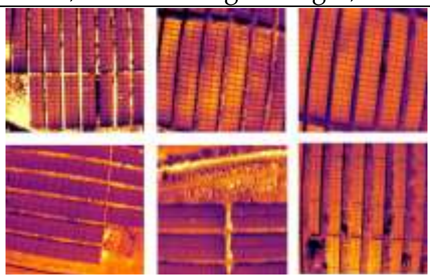

Figure 1: Batch of Six Images for the Guangfu Dataset.

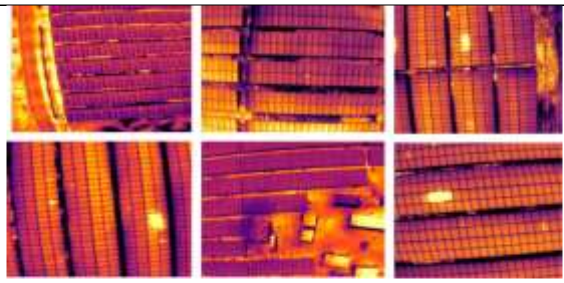

Figure 2: Batch of Six Images for the Panel Dataset.

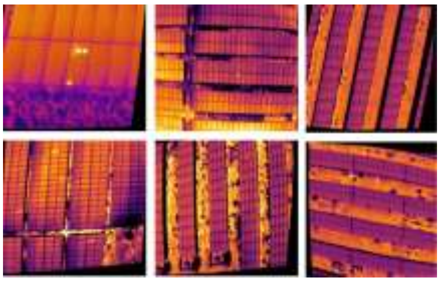

Figure 3. Batch Of Six Images for the Thermal Dataset.

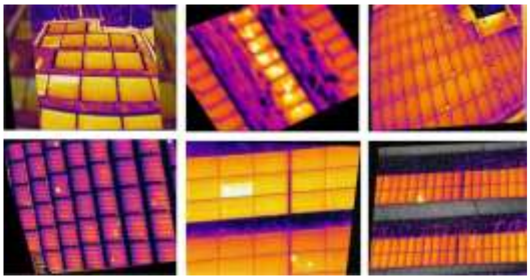

Figure 4: Batch of Six Images for the Beatrix.

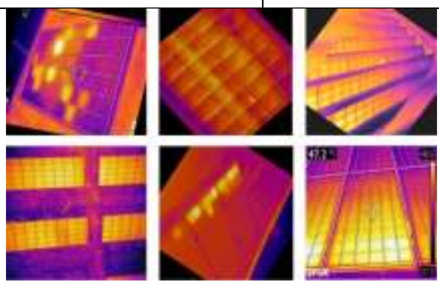

Figure 5: Batch of Six Images Tunghai University Dataset.

To mix these granular classes, which would otherwise result in a noisy and incoherent dataset, a binary classification protocol was implemented: Defect vs. Background. All specific defect types ranging from hotspots and cracks to overheating were mapped to a single Defect category. This strategy directs the model to learn fundamental defect characteristics rather than overfitting to specific sub-types. The outcome is a generalized pre-trained model capable of effective fine-tuning for specialized downstream tasks.

However, it is important to acknowledge the inherent limitations of reducing multi-class defects into a single binary category. While this homogenization is instrumental for robust pre-training and primary screening, it restricts the model's diagnostic granularity. By collapsing distinct anomalies into a uniform "Defect" class, the model loses the capacity to differentiate between specific physical or electrical faults, such as distinguishing a hotspot induced by partial shading from a micro-crack or a bypass diode failure. Consequently, while the proposed framework excels as an initial anomaly detection tool, translating these detections into targeted maintenance interventions requires a secondary phase of fine-tuning utilizing granular, multi-class annotations.

3.2. Validation Framework: Yolov9 Architecture

The YOLOv9 architecture (C.-Y. Wang et al., 2024) is employed to rigorously evaluate the quality of the unified dataset. This architecture was selected for its advanced capability to preserve information flow, a critical factor in assessing the learnability of the data.

Deep neural networks frequently experience information degradation as data propagates through successive layers (Tishby & Zaslavsky, 2015). Although mechanisms such as reversible architectures (Cai et al., 2023) or deep supervision (Shen et al., 2020) exist, they often incur high computational costs. YOLOv9 addresses this via Programmable Gradient Information (PGI). PGI utilizes an auxiliary branch to generate reliable gradients, ensuring that deep features retain essential information without compromising inference speed. This facilitates an assessment of the dataset's intrinsic potential by minimizing information loss during the training phase.

Furthermore, YOLOv9 incorporates the Generalized Efficient Layer Aggregation Network (GELAN), an evolution of ELAN (C.-Y. Wang, Liao, & Yeh, 2022). GELAN optimizes the trade-off between parameters, inference speed, and accuracy. The utilization of this efficient architecture ensures that the evaluation reflects the quality of the dataset rather than merely the capacity of the model. Figure 6 shows the YOLOv9 structure, including the backbone, auxiliary, neck, and head blocks.

3.2.1. Programmable Gradient Information (PGI)

As illustrated in Figure 7(d), PGI integrates an auxiliary reversible branch. Critically, this branch is active solely during the training phase to guide gradient updates and is excised during inference, thereby incurring no additional computational cost in production.

This mechanism mitigates the information bottleneck by providing a direct path for gradient propagation. It ensures that the model effectively learns defect characteristics from the unified dataset. Additionally, multilevel auxiliary information prevents error accumulation in deep networks, facilitating the training of both shallow and deep architectures on the large-scale dataset.

3.2.2. Generalized ELAN (GELAN)

GELAN synthesizes the advantages of CSPNet (C. Wang et al., 2020) and ELAN (C.-Y. Wang et al., 2022) to establish a flexible and efficient design. As depicted in Figure 8, it accommodates various computational blocks while maintaining a lightweight and accurate model structure. This efficiency is instrumental in evaluating the unified dataset across diverse model configurations.

3.3. Experimental Variants

To conduct a comprehensive evaluation, four variants of YOLOv9 were trained: Gelan-c, Gelan-e, Yolov9-c, and Yolov9-e. These variants span a spectrum from compact, high-speed models (Gelan-c, Yolov9-c) to larger, more complex architectures (Gelan-e, Yolov9-e). Table 2 provides the architectural specifications. This diver- sity enables an assessment of the unified pre-training dataset's performance across different computational constraints.

Table 2: Yolov9 Architecture Specification.

Variant	Layers	Parameters	GFLOPs
Gelan-c	621	25441698	103.2
Gelan-e	1228	58057378	192.2
Yolov9-c	962	51011140	238.9

Yolov9-e	1475	69415556	244.9

3.4. Deployment Considerations

The ultimate objective of this dataset is to facilitate scalable, real-time inspection systems. In a practical deployment scenario, thermal imagery from PV arrays would be processed by models pre-trained on the unified Defect class. This initialization provides a robust general understanding of thermal anomalies, thereby simplifying fine-tuning for specific requirements and supporting a flexible, data-driven maintenance strategy for the solar industry.

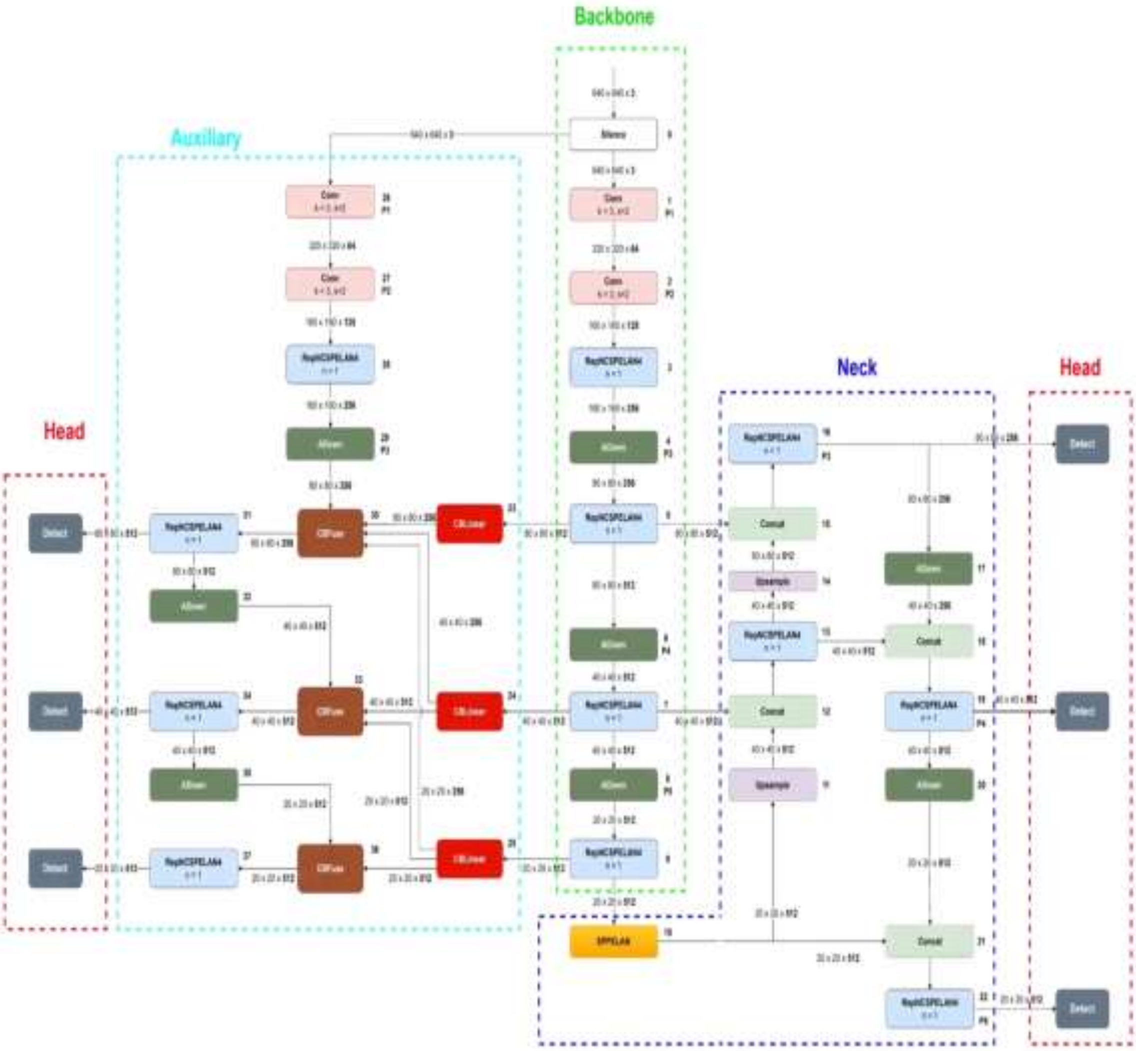

Figure 6: Block Diagram of Yolov9 Architecture Comprising Backbone, Auxiliary, Neck, And Head Blocks (Hidayatullah, 2024).

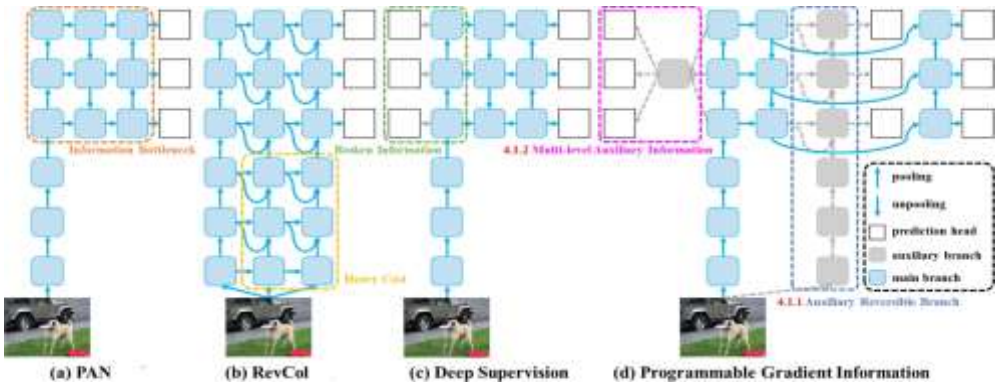

Figure 7 Comparison of Yolov9 PGI With Other Approaches (C.-Y. Wang Et Al., 2024).

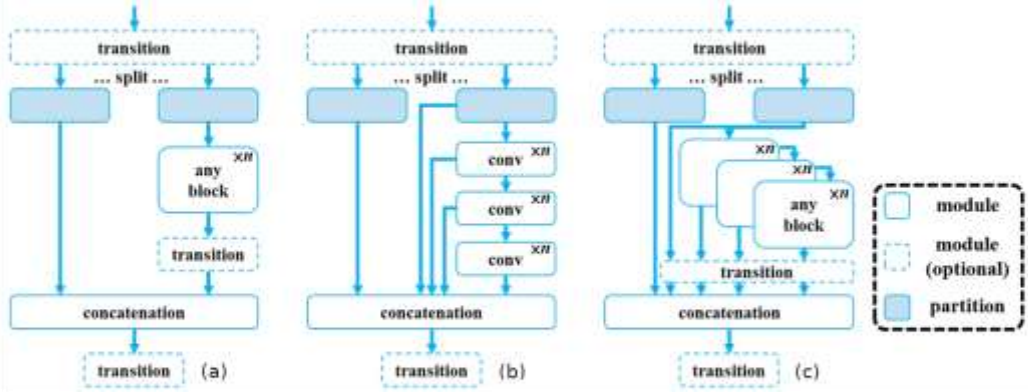

Figure 8: GELAN And Related Works A) Cspnet B) ELAN C) GELAN.

4. EXPERIMENTAL SETUP

A rigorous experimental framework was established to evaluate the efficacy of the unified pre-training dataset. The objective was to assess the extent to which this dataset facilitates the learning of generalized features across diverse architectures. Four YOLOv9 variants—Gelan-c, Gelan-e, YOLOv9-c, and YOLOv9-e, served as the validation instruments. These models were utilized in their standard configurations, functioning as benchmarks to quantify the quality and utility of the consolidated dataset.

To ensure reproducibility and experimental consistency, a standardized training protocol was applied across all trials. A linear warm-up phase was implemented during the initial three epochs to stabilize the training process. Subsequently, learning rate decay was adjusted according to model complexity. To refine detection accuracy in the final stages, Mosaic data augmentation was disabled for the last 15 epochs, thereby preventing the introduction of distortion artifacts during convergence.

Each experimental run was configured for a maximum of 60 epochs, incorporating an early stopping mechanism to terminate training upon the cessation of validation loss improvement, thus mitigating overfitting. Hyperparameters were maintained at constant values, with an initial learning rate of 0.01 and a batch size of 4. All computational tasks were executed on an NVIDIA RTX 3060 Ti GPU, ensuring sufficient processing capability for the large-scale dataset.

Table 3: Model Performance on Guangfu Pid Dataset.

Model	Dataset	Precision	Recall	mAP@0.5
gelan-c	Guangfu Pid Dataset	0.461	0.666	0.512
gelan-e	Guangfu Pid Dataset	0.422	0.675	0.474
yolov9-c	Guangfu Pid Dataset	0.455	0.683	0.505
yolov9-e	Guangfu Pid Dataset	0.450	0.676	0.509

5. RESULTS AND ANALYSIS

This section presents the evaluation of the unified pre-training dataset through the deployment of four YOLOv9 variants. The analysis focuses on the dataset's capacity to facilitate generalized feature learning across diverse data subsets, as quantified by standard object detection metrics.

Performance is assessed using Precision (P) eq. (1), Recall (R) eq. (2), Average Precision (AP), Mean Average Precision (mAP) eq. (3). Precision quantifies the accuracy of positive predictions, while Recall measures the proportion of actual defects identified. The F1-score provides a harmonic mean of these two metrics (Fan et al., 2020). Crucially, the Mean Average Precision at an Intersection over Union (IoU) of 0.5 (mAP@0.5) is utilized as the primary indicator of localization accuracy. Given the operational significance of even minor defects in PV systems, mAP@0.5 serves as the most reliable metric for assessing the dataset's fidelity in capturing defect severity and spatial location.

$$P = \frac{TP}{TP + FP} \quad (1)$$

$$R = \frac{TP}{TP + FN} \quad (2)$$

$$\text{mAP} = \frac{1}{N}\sum_{n=1}^{N} AP_n \quad (3)$$

$$F1 = 2\frac{P \times R}{P + R} \quad (4)$$

Here $\text{AP} = \int_0^1 \text{P}(R)\, dR$, TP = True Positive, FP = False Positive and FN = False Negative.

Validation of the dataset's integrity was conducted through the evaluation of YOLOv9 variants on both individual constituent data sources and the aggregated repository. This stratified analysis confirms the efficacy of the Single Class unification strategy across heterogeneous defect typologies, ranging from hotspots to cracks.

Quantitative results for the Guangfu PID dataset are presented in Table 3. The Gelan-c model attained the highest mAP@0.5 of 0.512 on the 234 test images, indicative of effective PID feature acquisition notwithstanding the simplified binary labeling schema. Regarding the Tunghai University dataset (Table 4), the YOLOv9-c variant exhibited superior performance, yielding a mAP@0.5 of 0.719. This suggests the dataset's efficacy persists even in the presence of ambiguous source annotations.

Performance metrics derived from the Panel dataset (Table 5) indicate that the lightweight Gelan-c model surpassed more complex architectures, achieving a mAP@0.5 of 0.406. This observation implies that the unified dataset facilitates efficient feature extraction, thereby obviating the necessity for extensive computational resources. A comparable trend was noted in the Thermal Defects dataset (Table 6), wherein Gelan-c again secured the premier position with a mAP@0.5 of 0.634. These findings were corroborated by results from the Beatrix Vision dataset (Table 7), with Gelan-c achieving a mAP@0.5 of 0.666.

Finally, an evaluation of the combined repository, comprising 471 images, was conducted (Table 8) to assess holistic performance. Gelan-c emerged as the most effective variant, attaining a mAP@0.5 of 0.591. These empirical results underscore the central thesis that the aggregation of diverse, inconsistent datasets into a single binary class establishes a robust foundation for learning. Through training on this comprehensive unified collection, models acquire fundamental defect characteristics with broad applicability, thereby offering advantages extending beyond architectural optimization.

Figure 10 illustrates the training trajectory of the Gelan-c model, evidencing steady improvements in precision, recall, and mAP. The precision curve (Figure 12) demonstrates the model's confidence in detections, while the recall curve (Figure 13) confirms the identification of 87% of actual defects even at lower thresholds. Qualitative validation in Figure 9 displays the model's capability to correctly identify both macroscopic and microscopic defects across disparate sources. This visual evidence confirms that the binary classification strategy preserves the spatial fidelity required for effective PV monitoring.

Table 4: Model Performance on Tunghai University Dataset.

Model	Dataset	Precision	Recall	mAP@0.5
gelan-c	Tunghai University Dataset	0.718	0.728	0.695
gelan-e	Tunghai University Dataset	0.663	0.648	0.630
yolov9-c	Tunghai University Dataset	0.761	0.738	0.719
yolov9-e	Tunghai University Dataset	0.737	0.734	0.696

Table 5: Model Performance on Panel Dataset.

Model	Dataset	Precision	Recall	mAP@0.5
gelan-c	Panel Dataset	0.449	0.577	0.406
gelan-e	Panel Dataset	0.433	0.593	0.399
yolov9-c	Panel Dataset	0.370	0.638	0.314
yolov9-e	Panel Dataset	0.394	0.624	0.329

Table 6: Model Performance on Thermal Dataset.

Model	Dataset	Precision	Recall	mAP@0.5
gelan-c	Thermal Dataset	0.560	0.667	0.634
gelan-e	Thermal Dataset	0.520	0.691	0.624
yolov9-c	Thermal Dataset	0.472	0.714	0.606
yolov9-e	Thermal Dataset	0.498	0.711	0.622

Table 7: Model Performance on Beatrix Vision Dataset.

Model	Dataset	Precision	Recall	mAP@0.5
gelan-c	Beatrix Vision Dataset	0.628	0.683	0.666
gelan-e	Beatrix Vision Dataset	0.617	0.690	0.628
yolov9-c	Beatrix Vision Dataset	0.624	0.696	0.633
yolov9-e	Beatrix Vision Dataset	0.595	0.636	0.604

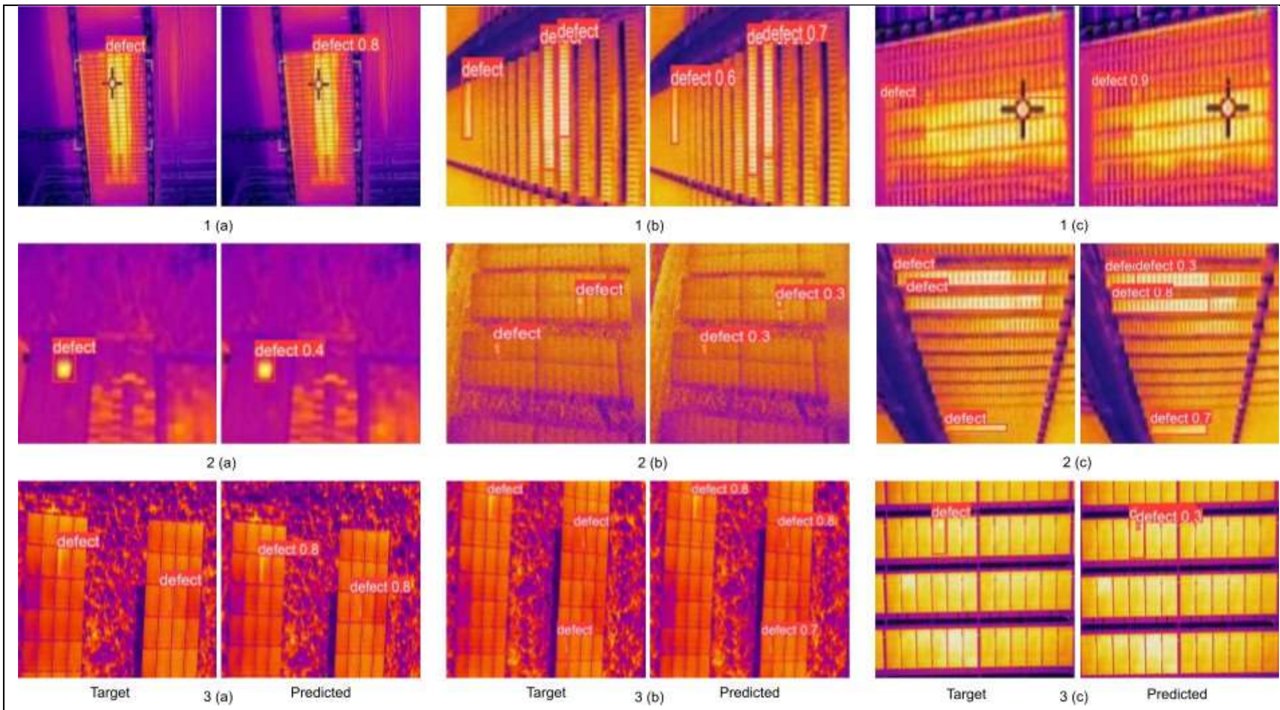

Figure 9: Qualitative Results for the Curated Dataset. Three Rows, Each Row Has Three Actual Defects Juxtaposed with Predicted Defects from Gelan-C Model.

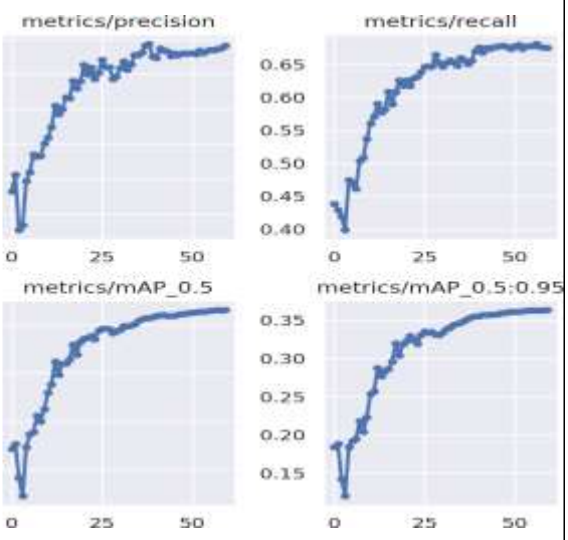

Figure 10: Gelan-C Model Metric Curves for Precision, Recall, Map@0.5 And Map0.5-0.95 On Evaluation Dataset.

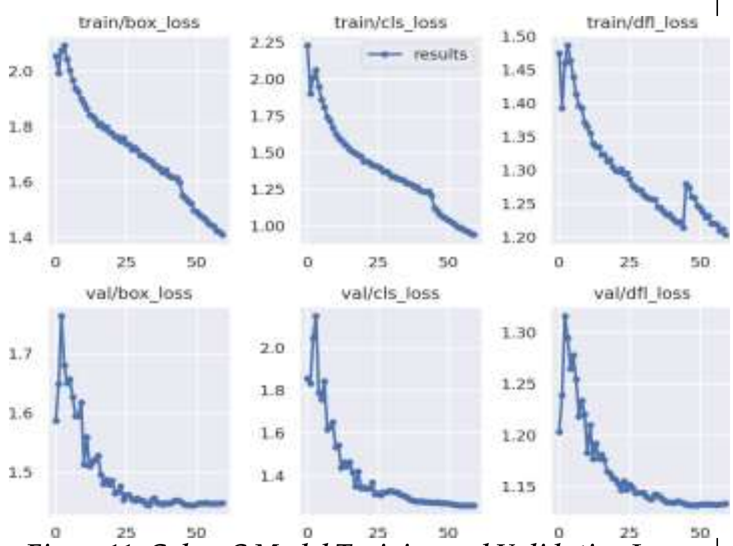

Figure 11: Gelan-C Model Training and Validation Loss Curves for Evaluation Dataset.

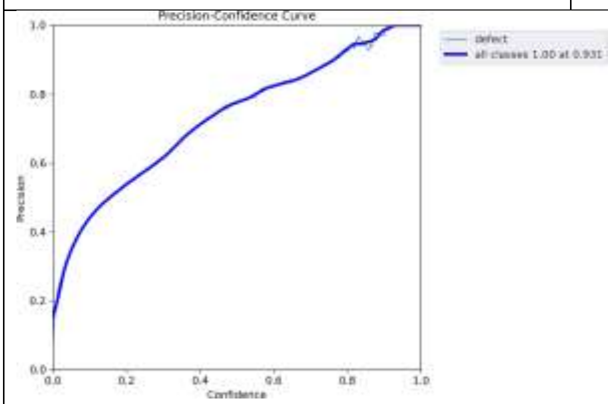

Figure 12: Gelan-C Model Precision Curve Trained on Curated Dataset.

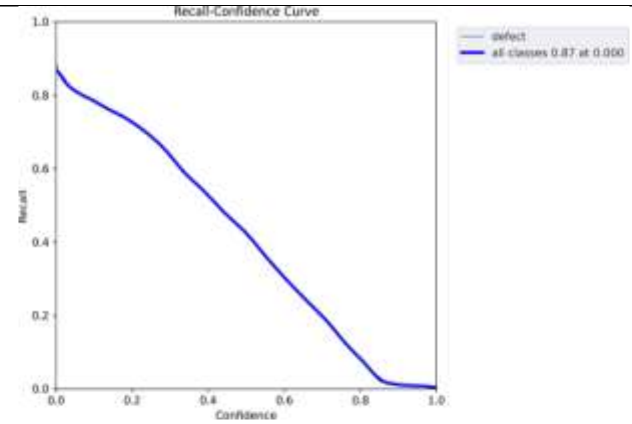

Figure 13: Gelan-C Model Recall Curve Trained on Curated Dataset.

Table 8: Model Performance on Curated Dataset.

Model	Dataset	Precision	Recall	mAP@0.5
gelan-c	All Datasets	0.558	0.666	0.591
gelan-e	All Datasets	0.537	0.657	0.573
yolov9-c	All Datasets	0.495	0.687	0.562
yolov9-e	All Datasets	0.505	0.690	0.570

6. CONCLUSION AND FUTURE WORK

This study addressed the challenges associated with detecting photovoltaic (PV) defects across varying environmental conditions through the implementation of a data-centric methodology. The primary contribution of this work involves the development and validation of a comprehensiv Pre-training PV Thermal Dataset, synthesized from five distinct open-source repositories. To mitigate the inconsistencies inherent in heterogeneous labeling taxonomies wherein classifications vary from specific hotspots to generalized faults, a binary classification strategy (Defect vs. Background) was adopted. This single-class formulation prioritizes the acquisition of fundamental defect characteristics, thereby circumventing the ambiguities associated with inconsistent nomenclature and establishing a robust foundation for subsequent model development. The efficacy of this unified dataset was substantiated through evaluation utilizing four YOLOv9 model variants (Gelan-c, Gelan-e, YOLOv9-c, and YOLOv9-e). These architectures were employed not to demonstrate architectural superiority, but to validate the hypothesis that data unification enhances feature learning. Empirical results demonstrate that the large-scale, curated dataset facilitates superior generalization capabilities. Notably, the Gelan-c variant emerged as a balanced and effective validation instrument. These findings indicate that a comprehensive, unified dataset effectively ameliorates the limitations associated with small-scale, isolated datasets, providing a reliable baseline for computer vision applications within the energy sector.

Future research will focus on leveraging this pre-trained foundation for transfer learning applications. It is anticipated that fine-tuning models initialized with this unified dataset will enhance the accuracy of specific defect type identification while reducing the dependency on extensive labeled examples. Additionally, investigations will be conducted into the environmental implications of training on large-scale datasets versus model optimization, alongside explorations into the deployment of this system on real-time edge computing devices to minimize operational expenditures.

Acknowledgments: This study was supported by a research grant provided by the Research, Development, and Innovation Authority (RDIA), Kingdom of Saudi Arabia, with Grant No. 13382-psu-2023- PSNU-R-3-1-EI-. The authors would like to acknowledge the support of Prince Sultan University, Riyadh, Saudi Arabia, in paying the article processing charges of this publication. In addition, this research is supported by the Automated Systems and Computing Lab (ASCL), Prince Sultan University, Riyadh, Saudi Arabia.

Fundings: This work is funded by Prince Sultan University, Riyadh, Saudi Arabia.

Conflict Of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Author Contributions:
Conceptualization: Zeeshan Haider, Ahmad Taher Azar
Investigation: Zeeshan Haider, Yasmin Adel Hagag
Validation: Ahmad Taher Azar, Yasmin Adel Hagag
Methodology: Zeeshan Haider, Ahmad Taher Azar
Formal analysis: Zeeshan Haider, Ahmad Taher Azar, Yasmin Adel Hagag
Writing–original draft: Zeeshan Haider
Writing–review & **editing**: Zeeshan Haider, Ahmad Taher Azar, Yasmin Adel Hagag

Consent for publication: “Not applicable”.

Availability Of Data: The datasets used and analyzed during the current study will be available on request.

Further Disclosure: This manuscript is being submitted exclusively to this journal.
(i) The findings presented in this manuscript have not been presented at any conference, academic meeting, congress, or similar public venue.
(ii) This paper has **not** been uploaded to or deposited in any preprint server.

REFERENCES

Ab Hamid, M. Z., Daud, K., Soh, Z. H. C., Osman, M. K., Isa, I. S., & Ishak, N. H. (2024). Review of deep learning-based hotspot detection in solar photovoltaic arrays. In 2024 ieee 4th international conference in power engineering applications (icpea) (pp. 332–337).

Akram, M. W., Li, G., Jin, Y., & Chen, X. (2022). Failures of photovoltaic modules and their detection: A review. Applied Energy, 313, 118822.

Alves, R. H. F., de Deus Junior, G. A., Marra, E. G., & Lemos, R. P. (2021). Automatic fault classification in photovoltaic modules using convolutional neural networks. Renewable Energy, 179, 502–516.

Ameerdin, M. I., Jamaluddin, M. H., Shukor, A. Z., Kamaruzaman, L. A. H., & Mohamad, S. (2024). Towards efficient solar panel inspection: A yolo-based method for hotspot detection. In 2024 ieee 14th symposium on computer applications & industrial electronics (iscaie) (pp. 367–372).

Bochkovskiy, A., Wang, C., & Liao, H. M. (2020). Yolov4: Optimal speed and accuracy of object detection. CoRR, abs/2004.10934. Retrieved from https://arxiv.org/abs/2004.10934

Bommes, L., Pickel, T., Buerhop-Lutz, C., Hauch, J., Brabec, C., & Peters, I. M. (2021). Computer vision tool for detection, mapping, and fault classification of photovoltaics modules in aerial ir videos. Progress in Photovoltaics: Research and Applications, 29(12), 1236–1251.

Bosman, L. B., Leon-Salas, W. D., Hutzel, W., & Soto, E. A. (2020). Pv system predictive maintenance: Challenges, current approaches, and opportunities. Energies, 13(6), 1398.

Byrne, J., & Kurdgelashvili, L. (2011). The role of policy in pv industry growth: past, present and future. Handbook of photovoltaic science and engineering, 39–81.

Cai, Y., Zhou, Y., Han, Q., Sun, J., Kong, X., Li, J., & Zhang, X. (2023). Reversible column networks. In The eleventh international conference on learning representations, ICLR 2023, kigali, rwanda, may 1-5, 2023. OpenReview.net. Retrieved from https://openreview.net/pdf?id=Oc2vlWU0jFY

Chen, Y., Liu, Y., Jiang, D., Zhang, X., Dai, W., Xiong, H., & Tian, Q. (2022). Sdae: Self-distillated masked autoencoder. In S. Avidan,

de Oliveira, A. K. V., Aghaei, M., & Ruther, R. (2022). Automatic inspection of photovoltaic power plants using aerial infrared thermography: a review. Energies, 15(6), 2055.

Deci-Team. (2023). Yolo-nas by deci achieves sota performance on object detection using neural architecture search. https://deci.ai/blog/yolo-nas-object-detection-foundation-model/. (Accessed: 2024-05-21)

El-Rashidy, M. A. (2022). An efficient and portable solar cell defect detection system. Neural Computing and Applications, 34(21), 18497–18509.

Fan, S., Li, J., Zhang, Y., Tian, X., Wang, Q., He, X., . . . Huang, W. (2020). On line detection of defective apples using computer vision system combined with deep learning methods. Journal of Food Engineering, 286, 110102. Retrieved from https://www.sciencedirect.com/science/article/pii/S0260877420302004 doi: https://doi.org/10.1016/j.jfoodeng.2020.110102

G. J. Brostow, M. Cisse, G. M. Farinella, & T. Hassner (Eds.), ´ Computer vision - ECCV 2022 - 17th european conference, tel aviv, israel, october 23-27, 2022, proceedings, part XXX (Vol. 13690, pp. 108–124). Springer. Retrieved from https://doi.org/10.1007/978-3-031-20056-4 7 doi: 10.1007/978-3-031-20056-4\ 7

Ge, Z., Liu, S., Wang, F., Li, Z., & Sun, J. (2021). Yolox: Exceeding yolo series in 2021.

Gevorgian, A., Pernigotto, G., & Gasparella, A. (2024). Addressing data scarcity in solar energy prediction with machine learning and augmentation techniques. Energies, 17(14), 3365.

Glavas, H., Vukobratovi ˇ c, M., Primorac, M., & Mu ´ stran, D. (2017). Infrared thermography in inspection of photovoltaic panels. In ˇ 2017 international conference on smart systems and technologies (sst) (pp. 63–68).

GUO, X., LI, X., Mohammed, F. A. H., Alenizi, F. A., Alasedi, K. K., Mohsen, K. S., . . . Bostani, A. (2023). Green certificate-driven photovoltaic promotion in distribution networks hosting hydrogen fueling stations for future sustainable transportation: A risk adjusted dominance analysis. Sustainable Cities and Society. doi: https://doi.org/10.1016/j.scs.2023.104911

Hagag, Y. A., ElSobky, M., Ibrahim, A. M., Azar, A. T., & Haider, Z. (2025). Automated defect detection in solar module electroluminescence images using yolov9 variants. International Journal of Automation and Control. doi: https://doi.org/10.1504/IJAAC.2025.147209

Han, X., Wang, X., Chen, C., Li, G., & Piao, C. (2023). Hot spot detection of thermal infrared image of photovoltaic power station based on multi-task fusion. Journal of Information Processing Systems, 19(6).

Hidayatullah, P. (2024). Yolov9 architecture explained. https://youtu.be/oZ6I1VHpil0?si=gIPwGfS BTY3VOaa. (Accessed: 2024-07-21)

Hoffmann, W. (2006). Pv solar electricity industry: Market growth and perspective. Solar energy materials and solar cells, 90(18-19), 3285–3311.

Høiaas, I., Grujic, K., Imenes, A. G., Burud, I., Olsen, E., & Belbachir, N. (2022). Inspection and condition monitoring of large-scale photovoltaic power plants: A review of imaging technologies. Renewable and Sustainable Energy Reviews, 161, 112353. Retrieved from https://www.sciencedirect.com/science/article/pii/S1364032122002647 doi: https://doi.org/10.1016/j.rser.2022.112353

Hong, F., Song, J., Meng, H., Wang, R., Fang, F., & Zhang, G. (2022). A novel framework on intelligent detection for module defects of pv plant combining the visible and infrared images. Solar Energy, 236, 406–416.

Jocher, G., Stoken, A., Borovec, J., NanoCode012, ChristopherSTAN, Changyu, L., . . . Rai, P. (2020, October). ultralytics/yolov5: v3.1 Bug Fixes and Performance Improvements. Zenodo. Retrieved from https://doi.org/10.5281/zenodo.4154370 doi: 10.5281/zenodo.4154370

Kellil, N., Aissat, A., & Mellit, A. (2023). Fault diagnosis of photovoltaic modules using deep neural networks and infrared images under Algerian climatic conditions. Energy, 263, 125902.

Kendig, D., Alers, G. B., & Shakouri, A. (2010). Characterization of defects in photovoltaics using thermoreflectance and electroluminescence imaging. In 2010 35th ieee photovoltaic specialists conference (pp. 001733–001736).

Kesavan, P. K., Subramaniam, U., Almakhles, D. J., & Selvam, S. (2023). Modelling and coordinated control of grid connected photovoltaic, wind turbine driven pmsg, and energy storage device for a hybrid dc/ac microgrid. Protection and Control of Modern Power Systems. doi: https://doi.org/10.23919/PCMP.2023.000272

Li, C., Li, L., Jiang, H., Weng, K., Geng, Y., Li, L., . . . Wei, X. (2022). Yolov6: A single-stage object detection framework for industrial applications. CoRR, abs/2209.02976. Retrieved from https://doi.org/10.48550/arXiv.2209.02976 doi: 10.48550/ARXIV.2209.02976

Mart́ınez-Duart, J. M., & Hernandez-Moro, J. (2013). Commentary: Photovoltaics firmly moving to the terawatt scale. ́ Journal of Nanophotonics, 7(1), 078599–078599.

Mazen, F. M. A., Seoud, R. A. A., & Shaker, Y. O. (2023). Deep learning for automatic defect detection in pv modules using electroluminescence images. IEEE Access, 11, 57783–57795.

Mellit, A. (2022). An embedded solution for fault detection and diagnosis of photovoltaic modules using thermographic images and deep convolutional neural networks. Engineering Applications of Artificial Intelligence, 116, 105459.

Redefining education. (2023, nov). Beatrix vision dataset [Open Source Dataset]. https://universe.roboflow.com/redefiningeducation/beatrix-vision. Roboflow. Retrieved from https://universe .roboflow .com/redefiningeducation/beatrix -vision (visited on 2024-12-22)

Redmon, J., & Farhadi, A. (2017). YOLO9000: better, faster, stronger. In 2017 IEEE conference on computer vision and pattern recognition, CVPR 2017, honolulu, hi, usa, july 21-26, 2017 (pp. 6517–6525). IEEE Computer Society. Retrieved from https://doi.org/10.1109/ CVPR.2017.690 doi: 10.1109/CVPR.2017.690

Redmon, J., & Farhadi, A. (2018). Yolov3: An incremental improvement. CoRR, abs/1804.02767. Retrieved from http://dblp.uni-trier.de/db/journals/corr/corr1804.html#abs-1804-02767

Redmon, J., Divvala, S. K., Girshick, R. B., & Farhadi, A. (2016). You only look once: Unified, real-time object detection. In 2016 IEEE conference on computer vision and pattern recognition, CVPR 2016, las vegas, nv, usa, june 27-30, 2016 (pp. 779–788). IEEE Computer Society. Retrieved from https://doi.org/10.1109/CVPR.2016.91 doi: 10.1109/CVPR.2016.91

Ren, Y., Yu, Y., Li, J., & Zhang, W. (2020). Design of photovoltaic hot spot detection system based on deep learning. In Journal of physics: Conference series (Vol. 1693, p. 012075).

Rogotis, S., Ioannidis, D., Tsolakis, A., Tzovaras, D., & Likothanassis, S. (2014). Early defect diagnosis in installed pv modules exploiting spatio-temporal information from thermal images. In Proceedings of the 12th quantitative infrared thermography conference (qirt) (pp. 7–11).

S, N. (2023, jun). panel dataset [Open-Source Dataset]. https://universe.roboflow.com/natarajan-s-krt0e/panel-u4vhy. Roboflow. Retrieved from https://universe.roboflow.com/natarajan-s-krt0e/panel-u4vhy (visited on 2024-12-22)

Shen, Z., Liu, Z., Li, J., Jiang, Y., Chen, Y., & Xue, X. (2020). Object detection from scratch with deep supervision. IEEE Trans. Pattern Anal. Mach. Intell., 42(2), 398–412. Retrieved from https://doi.org/10.1109/TPAMI.2019.2922181 doi: 10.1109/TPAMI.2019.2922181

shuju. (2024, may). guangfupid dataset [Open-Source Dataset]. https://universe.roboflow.com/shuju/guangfu-pid. Roboflow. Retrieved from https://universe.roboflow.com/shuju/guangfu-pid (visited on 2024-12-22)

Smida, M. B., Azar, A. T., Sakly, A., & Hameed, I. A. (2024). Analyzing grid connected shaded photovoltaic systems with steady state stability and crow search mppt control. Frontiers in Energy Research. doi: https://doi.org/10.3389/fenrg.2024.1381376

Solveview. (2023, aug). Thermal defects dataset [Open-Source Dataset]. https://universe.roboflow.com/solveview/thermal-defects. Roboflow. Retrieved from https://universe.roboflow.com/solveview/thermal-defects (visited on 2024-12-22)

Spajic, M., Talajić, M., & Mršić, L. (2024). Using cnns for photovoltaic panel defect detection via infrared thermography to support industry 4.0. Business Systems Research Journal, 15(1), 45–66. Retrieved from https://doi.org/10.2478/bsrj-2024-0003 doi: doi:10.2478/bsrj-2024-0003

Tajwar, T., Mobin, O. H., Khan, F. R., Hossain, S. F., Islam, M., & Rahman, M. M. (2021). Infrared thermography-based hotspot detection of photovoltaic module using yolo. In 2021 ieee 12th energy conversion congress & exposition-asia (ecce-asia) (pp. 1542–1547).

Terven, J., Cordova-Esparza, D.-M., & Romero-González, J.-A. (2023, November). A comprehensive review of yolo architectures in computer vision: From yolov1 to yolov8 and yolo-nas. Machine Learning and Knowledge Extraction, 5(4), 1680–1716. Retrieved from http://dx.doi.org/10.3390/make5040083 doi: 10.3390/make5040083

Thakfan, A., & Bin Salamah, Y. (2024). Artificial-intelligence-based detection of defects and faults in photovoltaic systems: A survey. Energies, 17(19), 4807.