International Journal for Multidisciplinary Research (IJFMR)
E-ISSN: 2582-2160 ● Website: www.ijfmr.com ● Email: editor@ijfmr.com

HelioIQ: AI-Powered Solar Panel Fault Detection and Predictive Maintenance

Dr. G. Fathima[1], Sonha P[2], Tejashwini S B[3], Prathibha G[4]

[1]Professor, Department of Computer Science and Engineering, Adhiyamaan College of Engineering, Hosur, India
[2,3,4]UG Scholars, Department of Computer Science and Engineering, Adhiyamaan College of Engineering, Hosur, India

Abstract
HelioIQ is an intelligent, AI-powered platform designed to detect solar panel faults and support maintenance by combining multiple data sources into a unified system. Solar panels can experience conditions such as dust accumulation, bird droppings, snow cover, shading, overheating, loose connections, cracks, and electrical damage, which may reduce power generation and efficiency. Existing monitoring methods often rely on manual inspection or a single data source, limiting their ability to identify different fault conditions and link diagnostic results to maintenance actions. HelioIQ addresses these limitations through a multimodal approach that uses solar panel images and sensor readings, including voltage, current, power, panel temperature, and light intensity. A Convolutional Neural Network (CNN) developed using TensorFlow and Keras analyses panel images to classify visible conditions, while an image validation component rejects inputs that do not contain solar panels. An Extreme Gradient Boosting (XGBoost) model analyses sensor data to identify abnormal operating conditions. A decision-level fusion mechanism combines the outputs to determine the overall panel status, fault category, severity level, confidence score, and maintenance recommendation. Cases involving conflicting model outputs are flagged for manual review. The system is implemented as a full-stack web application using FastAPI, JWT-based authentication, a MySQL database, and a React-based dashboard. Its modules include Image Diagnosis, Sensor Diagnosis, Combined Analysis, Solar Array Grid, Prediction History, Work Orders, and Reports. By maintaining diagnostic records and linking detected conditions to structured maintenance tasks, HelioIQ provides an integrated approach to solar panel monitoring and maintenance support.

Keywords: Solar Panel Fault Detection, Predictive Maintenance, Multimodal Fusion, Convolutional Neural Network, XGBoost, Condition Monitoring, Renewable Energy.

1. Introduction

Solar photovoltaic (PV) systems are affected by dust accumulation, shading, physical damage and electrical faults, which can reduce their performance and increase maintenance requirements. Image-based inspection helps identify visible panel defects, while sensor data provides information about electrical and operating conditions. Combining these approaches can support more comprehensive fault diagnosis and maintenance planning [1]-[6].

HelioIQ is a web-based system developed to integrate image-based fault classification and sensor-based diagnosis within a single platform. The image module uses a convolutional neural network (CNN) to

classify six panel conditions, while the sensor module uses XGBoost to classify seven operating conditions. A rule-based fusion module combines the outputs, assigns severity levels and supports maintenance activities through prediction history, work orders and reports.

The project was developed as an undergraduate mini project using image and sensor datasets. The two models were evaluated separately, and their results are reported independently. The available evaluation does not establish improved accuracy through fusion or validate performance under real-world operating conditions.

The main contributions of this work are:

1. A CNN-based image classification module for six solar panel conditions.
2. An XGBoost-based sensor diagnosis module for seven operating conditions.
3. A rule-based fusion mechanism for combining diagnostic outputs and assigning fault severity.
4. A web application integrating diagnosis, prediction history, maintenance work orders and reporting.

The remainder of this paper is organised as follows. Section 2 reviews related work. Section 3 describes the system architecture, including the image diagnosis, sensor diagnosis, decision-level fusion and maintenance modules. Section 4 presents the datasets, preprocessing and data splits. Section 5 describes the model training and evaluation of the image and sensor classifiers. Section 6 explains the system implementation, including the technology stack and modules. Section 7 presents the results and discussion. Section 8 concludes the paper and outlines future work.

2. Related Work

Previous studies have explored image-based fault detection, sensor-based diagnosis and machine-learning approaches for photovoltaic (PV) systems. The following review summarises the studies relevant to HelioIQ and identifies the gap addressed by the proposed system.

2.1 Image-Based Fault Detection

Boubaker et al. [3] investigated machine-learning and deep-learning methods for PV fault diagnosis using infrared thermography. Pathak et al. [4] explored deep learning for solar-panel hotspot localisation and fault classification. Deitsch et al. [5] studied automated identification of defective solar cells using electroluminescence images, and Buerhop-Lutz et al. [8] presented a benchmark for the visual identification of defective solar cells in electroluminescence imagery. These studies demonstrate the potential of image analysis for identifying PV defects, although their imaging methods and target conditions differ.

2.2 Sensor-Based Fault Diagnosis

Ledmaoui et al. [1] reviewed predictive maintenance approaches for solar plants. Amiri et al. [6] investigated a deep-learning approach combining CNN and Bi-GRU for PV fault diagnosis. These studies provide context for using operational data and machine-learning techniques to support fault identification and maintenance. Chen and Guestrin [9] introduced XGBoost, a gradient-boosting method suitable for structured data classification.

2.3 Multimodal Diagnosis and Research Gap

Afrasiabi et al. [7] proposed a multimodal deep-learning approach for PV fault diagnosis, highlighting research into combining different information sources. Abdessadak et al. [2] reviewed the effects of dust on PV performance, further motivating the need to detect panel conditions that affect operation. XGBoost [9] and MobileNetV2 [10] provide methodological foundations for structured-data classification and efficient image feature learning, respectively.

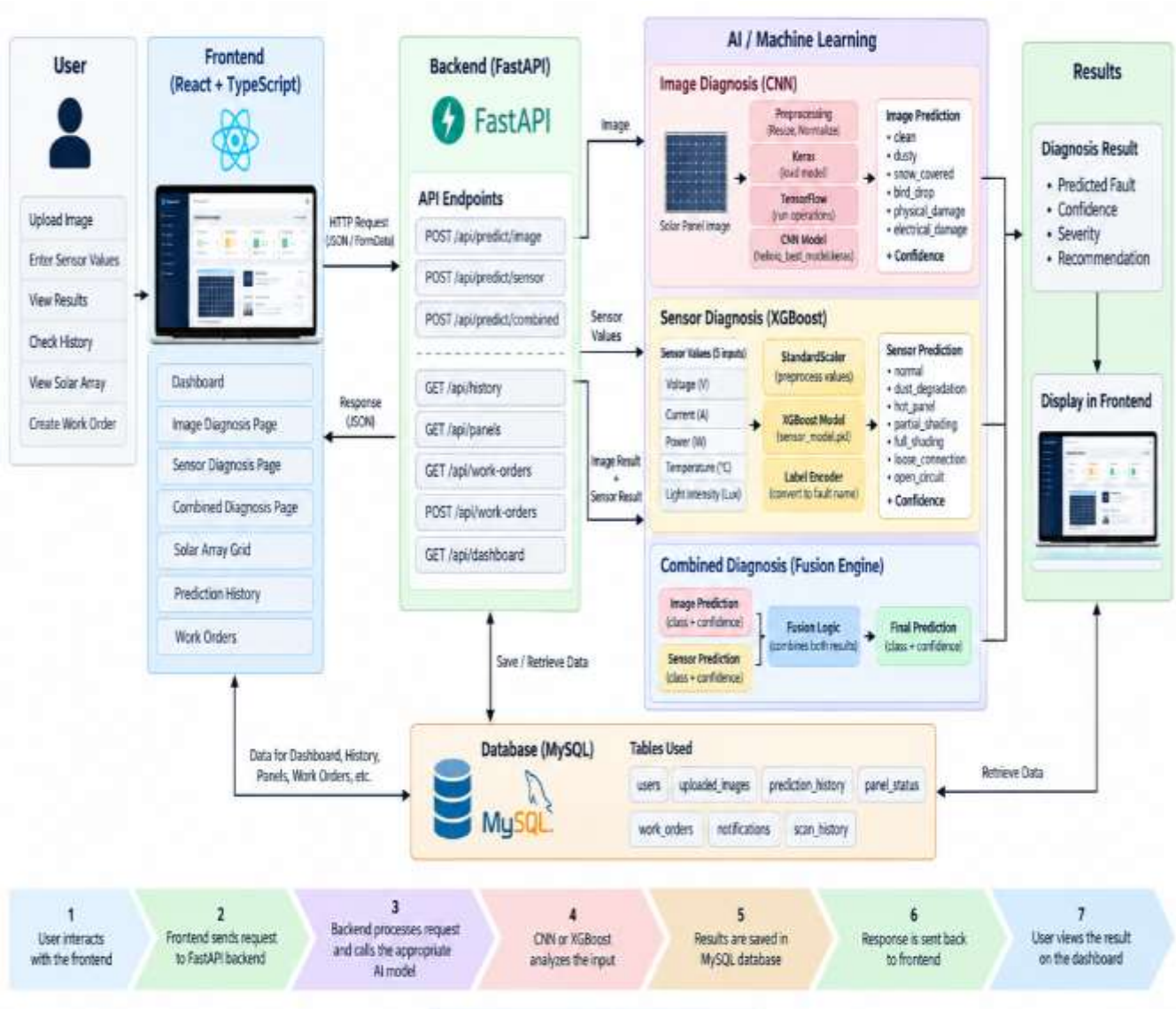
Unlike studies focused mainly on fault-detection models, HelioIQ integrates image classification, sensor diagnosis and rule-based decision support in a web application. Its contribution is this integrated workflow, including severity assignment and maintenance management. The proposed system integrates image-based and sensor-based predictions through a rule-based fusion mechanism. The combined analysis provides a unified fault assessment and maintenance recommendations.

3. System Architecture

3.1 Architectural Overview

HelioIQ follows a layered architecture, shown in Figure 1. The presentation layer comprises a React.js web dashboard developed with TypeScript, Vite, and Tailwind CSS. The service layer is implemented with FastAPI served through Uvicorn, with Pydantic schemas for request validation and JSON Web Token (JWT) based role-based access control (RBAC). The intelligence layer contains the MobileNetV2-based image validator, the CNN image classifier, the XGBoost sensor classifier, and the decision-level fusion engine. The data layer is a MySQL database that stores the prediction_history, work_orders, panel_status, and uploaded_images tables.

Figure 1: Architectural Workflow of the HelioIQ System



3.2 Image Diagnosis

The image diagnosis module processes an uploaded solar panel image by converting it to RGB format, resizing it to 256 x 256 pixels and normalizing pixel values to the range 0–1. The preprocessed image is passed to the trained CNN model, which predicts one of six conditions: clean, dusty, bird droppings, electrical damage, physical damage or snow-covered. Before classification, each uploaded image is passed to a MobileNetV2-based validator [10], which checks whether the image contains a solar panel. Images that are not recognised as solar panels are rejected.

3.3 Sensor Diagnosis

The sensor diagnosis module receives five inputs: voltage, current, power, panel temperature and light intensity. Preprocessing and feature engineering prepare the input features for the trained XGBoost classifier, which predicts one of seven conditions: normal, dust degradation, hot panel, partial shading, full shading, loose connection or open circuit. The output includes the predicted fault class and confidence score. The implemented feature engineering includes a power-to-illuminance proxy, rolling voltage and current standard deviations, power deviation and temperature change.

3.4 Decision-Level Fusion

As shown in Figure 1, the fusion engine combines the predictions from the image and sensor diagnosis modules to generate an overall panel diagnosis. The system considers the predicted fault categories and their confidence levels to provide a combined result, including panel status, fault severity, and a maintenance recommendation. When the image and sensor predictions indicate different fault conditions, the system can flag the result for manual review based on the implemented decision rules. This approach provides a unified view of panel condition by considering both visual and sensor-based information.

3.5 Monitor and Maintenance

The maintenance recommendation workflow uses the diagnostic status and severity assigned by the fusion engine to support maintenance planning. As shown in Figure 1, the results are displayed on the web dashboard and can be associated with prediction history and work orders. The application maps severity levels to work-order priorities: Critical to Urgent, High to High, Medium to Medium and Low to Low. Configured response periods are one day, three days, seven days and fourteen days, respectively. These periods represent application settings rather than universally established maintenance standards.

4. Experimental Dataset And Preprocessing Methodology

4.1 Datasets

The image classifier was developed using the publicly available Solar Panel Images dataset hosted on the Kaggle repository. The image dataset contains 6,815 solar panel images belonging to six classes: clean, dusty, bird droppings, electrical damage, physical damage and snow-covered. These images are used to train and evaluate the CNN-based image classification model.

The sensor classifier was developed using a publicly available solar panel sensor dataset hosted on Kaggle. The dataset initially contained 152,973 records representing seven operating conditions: normal, dust degradation, hot panel, partial shading, full shading, loose connection and open circuit. After removing 13 records labelled as sensor spikes, 152,960 records remained for model development.

4.2 Data Preprocessing

For image processing, each image is converted to RGB format, resized to 256 x 256 pixels and normalized to the range 0–1 using: Normalized pixel value = Original pixel value ÷255

For sensor data, preprocessing includes handling missing values, scaling the input features and engineering five additional features from the electrical and environmental measurements. These features represent power relative to light intensity, voltage and current variability, power deviation, and temperature change.

4.3 Training and Testing Data Split

The image dataset was divided into training, validation, and testing subsets in the proportions 70%, 15.00%, and 15.00%, containing 4,770, 1,022, and 1,023 images respectively. The class-wise composition of the subsets is given in Table 1.

Table 1: Training, Validation, and Testing Split of the Image Dataset

Class Label	Training	Validation	Testing	Total
bird_drop	1,082	232	232	1,546
clean	1,034	222	222	1,478
dusty	617	132	132	881
electrical_damage	409	87	88	584
physical_damage	835	179	179	1,193
snow_covered	793	170	170	1,133
Total	**4,770**	**1,022**	**1,023**	**6,815**

The sensor dataset was divided into training, validation and testing subsets in the proportions of approximately 64.00%, 16.00% and 20.00%, containing 97,894, 24,474 and 30,592 records, respectively. The class-wise distribution of the subsets is presented in Table 2.

Table 2: Class Distribution and Split of the Sensor Dataset

Fault Class	Training	Validation	Testing	Total
normal	88,676	22,170	27,711	138,557
dust_degradation	4,066	1,016	1,271	6,353
hot_panel	1,934	484	604	3,022
partial_shading	1,518	379	474	2,371
full_shading	1,026	256	321	1,603
loose_connection	354	89	111	554
open_circuit	320	80	100	500
Total	**97,894**	**24,474**	**30,592**	**152,960**

5. Model Training And Evaluation

5.1 Image Classification Model

The CNN model classifies solar panel images into six categories: bird droppings, clean, dusty, electrical damage, physical damage and snow-covered. Before training, images are resized to 256 × 256 pixels and normalized to the range of 0 to 1. The model was trained for 25 epochs using the Adam optimizer with a

learning rate of 1×10^{-5}. The final training accuracy was 70.06%, while the validation accuracy was 73.09%, as summarized in Table 3. The training and validation accuracy and loss curves are presented in Figure 2.

The model achieved an overall test accuracy of 74.58%. Figure 3 presents the Receiver Operating Characteristic (ROC) curves for the six image classes. The Area Under the Curve (AUC) values indicate varying classification performance across the categories, with the highest AUC of 0.93 for electrical damage and the lowest AUC of 0.50 for dusty images.

Table 3: Training configuration and final recorded metrics.

Training measure	Recorded value
Epochs	25
Learning rate	1×10^{-5}
Training accuracy	70.06%
Validation accuracy	73.09%
Training loss	0.8041
Validation loss	0.7018

Figure 2 : Training and validation accuracy and loss.

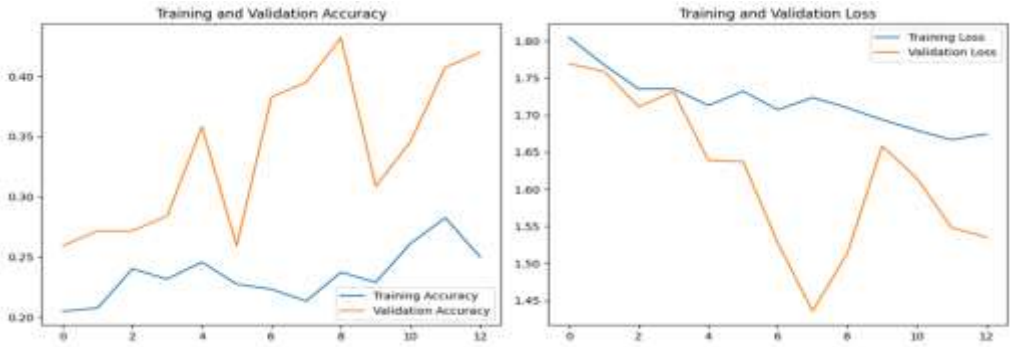
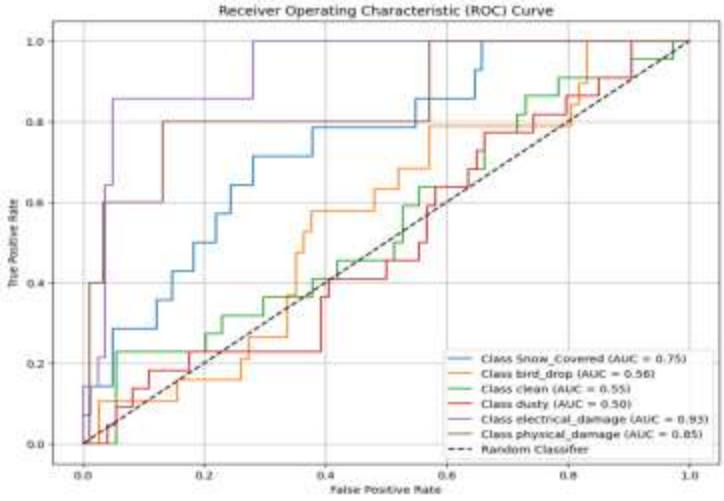


Figure 3: ROC curves and AUC for six classes.



5.2 Sensor Classification Model

The sensor classification model was developed using the XGBoost algorithm to identify solar panel conditions and faults based on sensor readings such as voltage, current, power, temperature, and light intensity. It classifies seven categories: normal, dust degradation, hot panel, partial shading, full shading, loose connection, and open circuit. The model achieved an overall test accuracy of 99.71%.

Figure 4 shows the training and validation loss of the XGBoost model, while Figure 5 presents the confusion matrix, illustrating its classification performance across the sensor fault categories.

Figure 4: Training and validation loss of the XGBoost model.

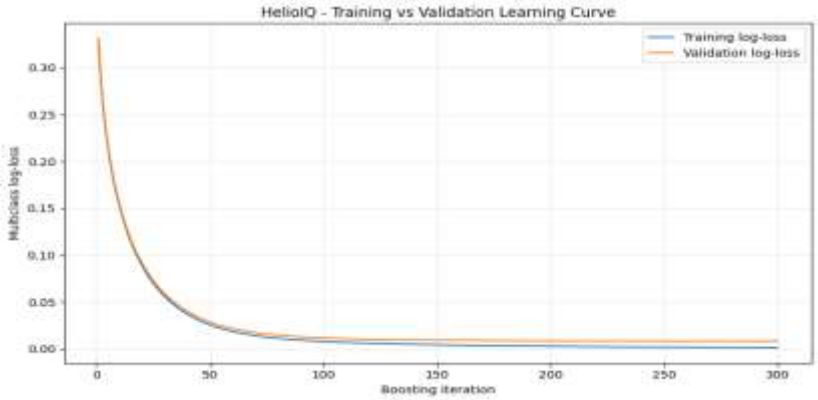


Figure 5: Confusion matrix for sensor fault classification.



6. System Implementation

6.1 Overall Technology Stack

HelioIQ was developed as a web-based application using frontend, backend, database, and machine learning technologies. The frontend uses React, TypeScript, Vite, and Tailwind CSS to provide the user interface. The backend is developed using Python FastAPI, which handles application requests and connects the diagnostic models. MySQL is used for data storage, while TensorFlow/Keras CNN and XGBoost are used for image-based and sensor-based fault classification, respectively. JWT authentication and role-based access control are used to manage user access.

6.2 System Modules

HelioIQ is developed as an integrated web-based platform that combines artificial intelligence techniques with maintenance management features for solar panel fault detection and predictive maintenance. The system is organized into seven major modules, each performing a specific function, including image-based diagnosis, sensor-based fault identification, combined analysis, prediction history, solar array monitoring, work order management, and report generation. These modules work together to help users analyse solar panel conditions, identify possible faults, review diagnostic results, and manage maintenance activities through a centralized web interface. The integration of CNN and XGBoost models enables the system to analyse both visual images and sensor measurements, while the decision-level fusion module combines their outputs to provide a comprehensive diagnosis and maintenance recommendation.

1. **Image Diagnosis Module:** Allows users to upload solar panel images. The CNN model analyses the images and identifies visual conditions such as bird droppings, dust, electrical damage, physical damage, and snow coverage.
2. **Sensor Diagnosis Module:** Processes sensor readings such as voltage, current, power, panel temperature, and light intensity. The XGBoost model identifies possible operating conditions and sensor-based faults.
3. **Combined Analysis Module**: Integrates the outputs of the image and sensor diagnosis models using decision-level rules to generate an overall panel status, fault severity, confidence level, and maintenance recommendation.
4. **Prediction History Module:** Maintains previous diagnosis records so users can review earlier image-based and sensor-based predictions.
5. **Solar Array Grid Module**: Provides an interface to view and monitor solar array information and related panel records.
6. **Work Order Management Module:** Supports the management of maintenance tasks based on identified panel conditions and recommended actions.
7. **Report Generation Module:** Generates downloadable reports containing relevant diagnosis and maintenance information.

The frontend communicates with the FastAPI backend to process image uploads and sensor readings. The backend invokes the corresponding CNN or XGBoost model and receives their predictions. The decision-level fusion module combines the available outputs according to the implemented rules. The resulting diagnosis and maintenance recommendations are then displayed on the web dashboard, with relevant records stored in the MySQL database.

7. Results And Discussion

7.1 Image Model Results

The CNN-based image classification model predicts the most likely solar panel condition from the six available categories. As shown in Figure 6, the system displays the predicted fault category along with its confidence score. In the example shown, the model identifies the input image as electrical damage. This output helps users understand the detected panel condition and supports maintenance personnel in deciding whether further inspection or corrective action is required.

Figure 6: Example image diagnosis output for electrical damage.



7.2 Sensor Model Results

The XGBoost-based sensor classification model identifies seven solar panel operating conditions using sensor measurements and engineered features. The model achieved an overall test accuracy of 99.71%. The model predicts normal operation with a displayed confidence of 99.87%. The corresponding maintenance recommendation is to continue regular monitoring while the panel operates normally. This demonstrates how the sensor diagnosis module provides fault predictions along with maintenance recommendations.

Table 4: Sensor Prediction and Maintenance Recommendation

Diagnosis	Displayed confidence	Maintenance recommendation
normal	99.87%	Panel operating normally. Continue regular monitoring.

7.3 Fusion and System-Level Results

The HelioIQ dashboard presents the results of image-based diagnosis, sensor-based diagnosis, and combined analysis. Figure 7 shows the image-based diagnosis result, where the uploaded solar panel image is analysed by the CNN model to identify the predicted condition and confidence score. Figure 8 presents the sensor-based diagnosis result generated by the XGBoost model using the provided sensor readings.

Figure 9 shows the combined analysis results, where the outputs of both diagnostic modules are processed through the decision-level fusion module. The system provides an overall panel status, fault severity, confidence level, and maintenance recommendation based on the implemented decision rules. These dashboard results demonstrate how HelioIQ brings image and sensor diagnosis together in a single web-based platform.

Figure 7: Image-Based Solar Panel Fault Diagnosis.

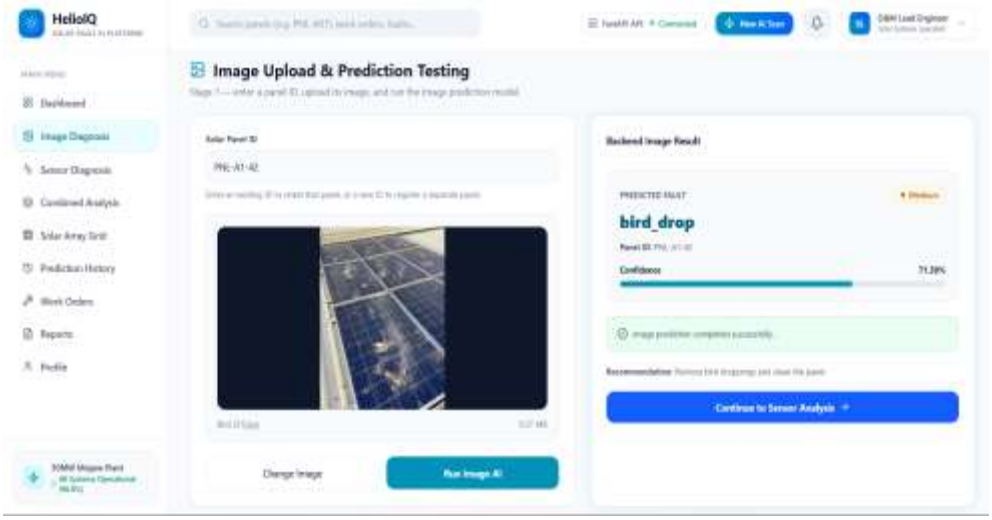


Figure 8: Sensor-Based Solar Panel Fault Diagnosis.

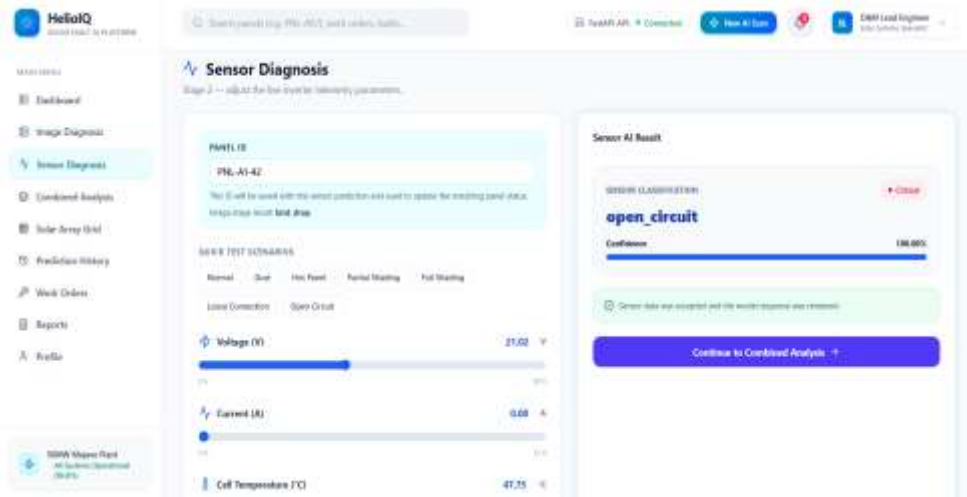
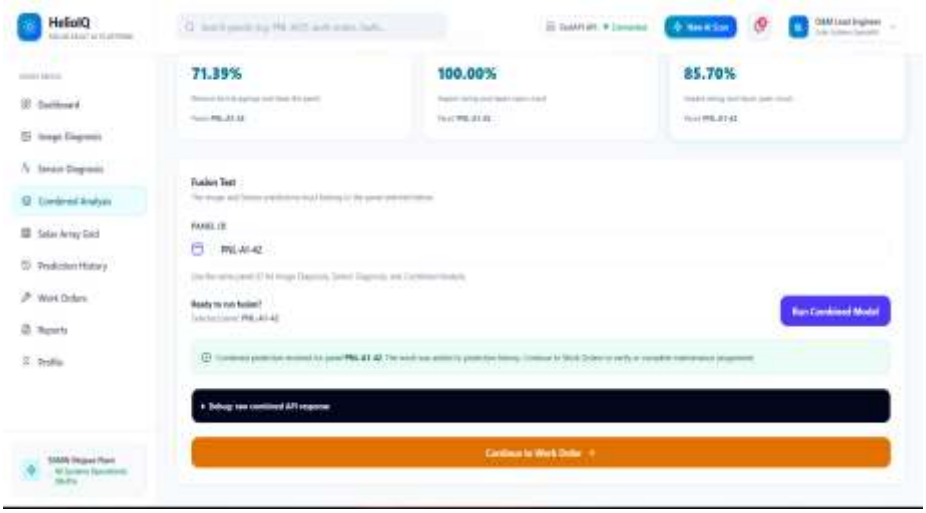


Figure 9: Combined Analysis Diagnostic Results.



8. Conclusion and Future Scope

HelioIQ presents an AI-powered approach for solar panel fault detection and predictive maintenance by combining image-based and sensor-based diagnosis. The CNN model achieved an overall classification accuracy of 74.58%, while the XGBoost model achieved 99.71% accuracy on their respective test datasets. The system integrates both diagnostic modules through decision-level fusion to provide panel status, fault severity, confidence levels, and maintenance recommendations. The web-based application also provides features such as prediction history, solar array grid, work order management, and report generation. The results demonstrate the potential of combining visual and sensor information to support solar panel fault identification and maintenance activities.

Future work will focus on improving image classification performance, evaluating the models using larger and more diverse real-world datasets, and testing the system with actual solar panel installations. Further improvements may include collecting paired image and sensor data for a more comprehensive evaluation of the fusion module, enhancing fault prediction capabilities, and integrating real-time sensor monitoring.

Acknowledgement

The authors thank Dr. G. Fathima, Professor and Head of the Department of Computer Science and Engineering, for her guidance as supervisor, Dr. R. Radhakrishnan, Principal, and the faculty of Adhiyamaan College of Engineering for the facilities and support provided during this work.

References

1. Y. Ledmaoui, A. El Maghraoui, M. El Aroussi, R. Saadane, "Review of Recent Advances in Predictive Maintenance and Cybersecurity for Solar Plants", Sensors, 2025, 25 (1), Article 206. https://doi.org/10.3390/s25010206
2. A. Abdessadak et al., "Assessing the Effects of Dust on Solar Panel Performance: A Comprehensive Review and Future Directions", Engineering Proceedings, 2025, 112 (1), Article 9. https://doi.org/10.3390/engproc2025112009
3. S. Boubaker, S. Kamel, N. Ghazouani, A. Mellit, "Assessment of Machine and Deep Learning Approaches for Fault Diagnosis in Photovoltaic Systems Using Infrared Thermography", Remote Sensing, 2023, 15 (6), Article 1686. https://doi.org/10.3390/rs15061686
4. S.P. Pathak, S. Patil, S. Patel, "Solar Panel Hotspot Localization and Fault Classification Using Deep Learning Approach", Procedia Computer Science, 2022, 204, 698–705. https://doi.org/10.1016/j.procs.2022.08.084
5. S. Deitsch, V. Christlein, S. Berger, C. Buerhop-Lutz, A. Maier, F. Gallwitz, C. Riess, "Automatic Classification of Defective Photovoltaic Module Cells in Electroluminescence Images", Solar Energy, 2019, 185, 455–468. https://doi.org/10.1016/j.solener.2019.02.067
6. A.F. Amiri, S. Kichou, H. Oudira, A. Chouder, S. Silvestre, "Fault Detection and Diagnosis of a Photovoltaic System Based on Deep Learning Using the Combination of a Convolutional Neural Network (CNN) and Bidirectional Gated Recurrent Unit (Bi-GRU)", Sustainability, 2024, 16 (3), Article 1012. https://doi.org/10.3390/su16031012
7. S. Afrasiabi, S. Allahmoradi, M. Afrasiabi, X. Liang, C.Y. Chung, J. Aghaei, "A Robust Multi-Modal Deep Learning-Based Fault Diagnosis Method for PV Systems", IEEE Open Access Journal of Power and Energy, 2024, 11, 583–594. https://doi.org/10.1109/OAJPE.2024.3497880

8. C. Buerhop-Lutz, S. Deitsch, A. Maier, F. Gallwitz, S. Berger, B. Doll, J. Hauch, C. Camus, C.J. Brabec, "A Benchmark for Visual Identification of Defective Solar Cells in Electroluminescence Imagery", Proceedings of the 35th European Photovoltaic Solar Energy Conference and Exhibition, Brussels, Belgium, 2018, 1287–1289. https://doi.org/10.4229/35thEUPVSEC20182018-5CV.3.15
9. T. Chen, C. Guestrin, "XGBoost: A Scalable Tree Boosting System", Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, August 2016, 785–794. https://doi.org/10.1145/2939672.2939785
10. M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, L.-C. Chen, "MobileNetV2: Inverted Residuals and Linear Bottlenecks", Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, June 2018, 4510–4520. https://doi.org/10.1109/CVPR.2018.00474