Available online www.worldscientificnews.com

World Scientific News

An International Scientific Journal

WSN 212 (2026) 280-306 EISSN 2392-2192

Solar Panel anomaly detection using StackNet Model

Richard Iherorochi Nneji[1,2], Benjamin Chiemeka Opara[3], Miracle Ugomma Anunobi[4], Simon Onuwa Agbonifo[5], Merit Chinonso Opara[6], Joyce Uchechi Nneji[7], Ojo Precious Mojolaoluwa[8], Wisdom Chima Olumba[9]

[1]Intelligent Computing Lab, Hace LLC, North Carolina, USA

[2]Chemical and Petroleum Engineering, University of Lagos, Nigeria

[3]Educational Management (Political Science), University of Benin, Nigeria

[4]Science Laboratory Technology (Biochemistry), University of Nsukka, Nigeria

[5]Chemistry, University of Benin, Nigeria

[6]Business Administration, Madonna University Okija, Nigeria

[7]College of Law, Chengdu University of Technology, China

[8]College of Ecology and Environment, Chengdu University of Technology, China

[9]Industrial Design Engineering, Chengdu University of Technology, China

*Author for Correspondence: olumbaw@stu.cdut.edu.cn

https://doi.org/10.65770/BZZF2385

ABSTRACT

Solar panel faults can significantly reduce energy output and reliability in renewable energy systems. Traditional fault detection methods often rely on manual inspection or rule-based systems, which struggle with accuracy and scalability. This paper introduces a deep learning-based fault diagnosis system designed to automatically detect and classify solar panel anomalies using infrared (IR) images. The proposed StackNet architecture builds upon the ResNet-18 backbone, with targeted modifications to enhance feature extraction efficiency while maintaining computational lightness. The model begins with a large-kernel convolutional layer for effective low-level feature capture, followed by normalization, activation, and pooling to reduce spatial resolution early and accelerate training.

(Received 15 December 2025; Accepted 22 January 2026; Date of Publication 19 February 2026)

StackNet strikes a balance between architectural depth and training efficiency, offering strong representational capacity with minimal increase in complexity. This makes it particularly suited for solar anomaly detection tasks where both performance and computational cost are critical.

The model was trained and evaluated on IR images for both binary (fault and no-fault) and multi-class (six fault types) classification. Experimental results demonstrate high diagnostic performance on both binary and multi-classification tasks. These outcomes validate the model's robustness across different imaging modalities and fault scenarios. This bridges the gap between research and field application, offering practical deployment potential for solar maintenance. In summary, this paper highlights the effectiveness of combining deep learning and attention mechanisms for reliable solar panel fault detection, contributing to more intelligent and automated renewable energy management.

Keywords: Solar Panel Fault Detection, Deep Learning, Infrared Imaging, Electroluminescence Imaging, Renewable Energy, Attention Mechanism, ResNet-Inception, Image Classification

1. INTRODUCTION

The consumption of energy is rising on a daily basis [1], so finding efficient ways to produce energy is becoming increasingly important. Compared with fossil fuels, renewable energy has lower greenhouse gas emissions, and the widespread adoption of it will help reduce environmental pollution [1]. As a renewable resource, solar energy does not cause global warming or damage ecosystems, thus helping to maintain ecological balance [1]. Solar panels use semiconductor materials and advanced technologies to convert solar radiation into electrical energy [1]. According to the National Renewable Energy Laboratory (NREL) [2], solar panels have a CO2 emission rate of about 45 to 50 grams per kilowatt-hour, whereas coal-fired power plants emit 900 to 1,000 grams of CO2. Therefore, promoting solar panels is considered an effective way to reduce both carbon dioxide emissions and environmental pollution [3].

Although solar panels are environmentally friendly and efficient, their performance is highly susceptible to various external and internal factors. Common faults, such as cracks, vegetation growth, soiling, diode failures, hot spots, pollution, fractures, and delamination [4]. These faults can seriously reduce energy conversion efficiency and shorten the lifespan of solar panels [5], impacting the stability and reliability of the entire solar energy system. For example, hot spots can cause localized overheating, while cracks and delamination may lead to permanent damage, reducing the overall output of the photovoltaic array. So timely identification and rectification of these issues are essential to maintain the consistent performance of solar panels and to prevent system-wide failures [6].

Currently, solar panel anomaly detection and classification primarily relies on visual inspection and electrical testing. Visual inspection is commonly conducted using infrared thermography (IRT) and electroluminescence (EL) imaging [7]. IRT detects thermal emissions from solar panel, identifying temperature anomalies that indicate potential faults such as hot spots and cell cracks [8]. In contrast, EL imaging captures the light emitted by solar panel under external voltage excitation, allowing visualization of internal electrical activity and enabling the detection of delamination and microcracks [9]. Electrical testing analyzes the electrical performance of solar panels to identify faults that affect efficiency and power output [10].

Among the commonly used methods, current-voltage (I-V) curve analysis measures the voltage and current characteristics of the panel under varying loads, enabling the detection of power degradation, short circuits, and electrical mismatches [10]. Despite these methods boosting detection efficiency, they still have limitations, including low accuracy, high costs, restricted detection range, and dependence on environmental conditions.

In recent years, the application of CNN(Convolutional Neural Network) within deep learning has significantly advanced fault classification [4]. CNNs can automatically extract features from complex data and adapt to various fault types and data distributions. In solar panel anomaly classification, CNNs are particularly effective in processing infrared and electroluminescence (EL) images [10], [11]. By learning from large-scale image datasets, CNN-based approaches provide an efficient and automated solution for fault classification, reduce manual intervention, and adapt to varying environmental conditions. Hence, leveraging CNNs for solar panel fault classification can improve the reliability and efficiency of solar systems, supporting intelligent operation and maintenance.

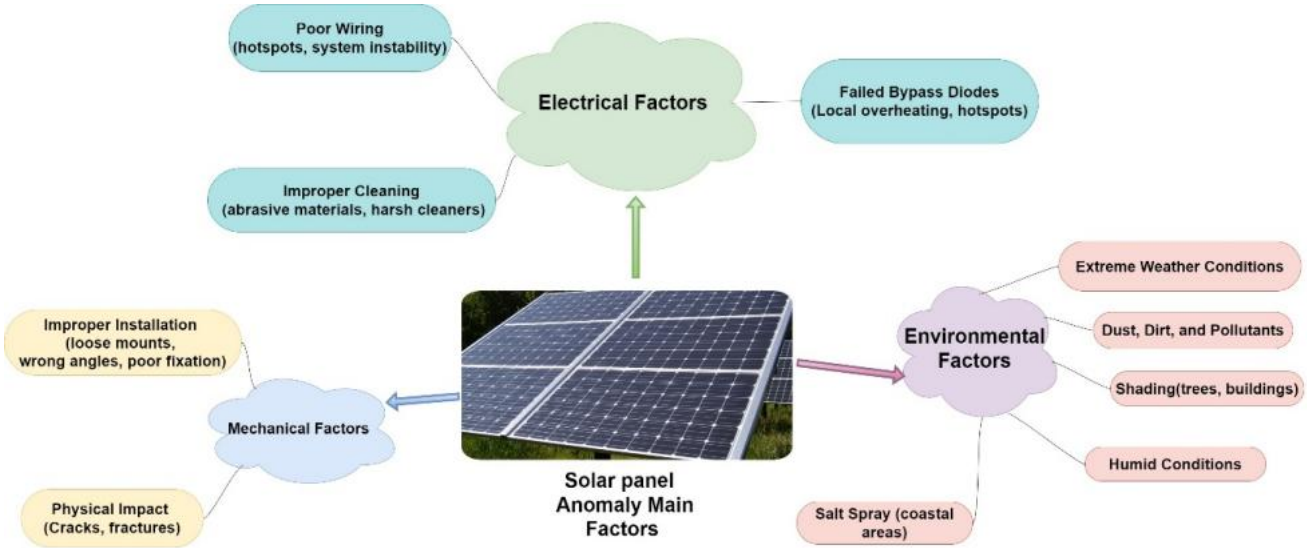


Figure 1. Factors of Solar Panel Anomaly.

This research integrates infrared (IR) and electroluminescence (EL) imaging technology with deep learning models to achieve automated fault classification, enhancing detection efficiency and accuracy, and the overview is seen in Figure 2. The SARNet model will be trained on two large-scale datasets to accurately identify solar panel conditions. An attention mechanism will be incorporated to improve the model's focus on critical regions, enhancing fault detection precision. Additionally, Grad-CAM (Gradient-weighted Class Activation Mapping) will be utilized to visualize the model's decision-making process, generating heatmaps to highlight key fault areas and improve interpretability and a graphical user interface (GUI) for real-time analysis.

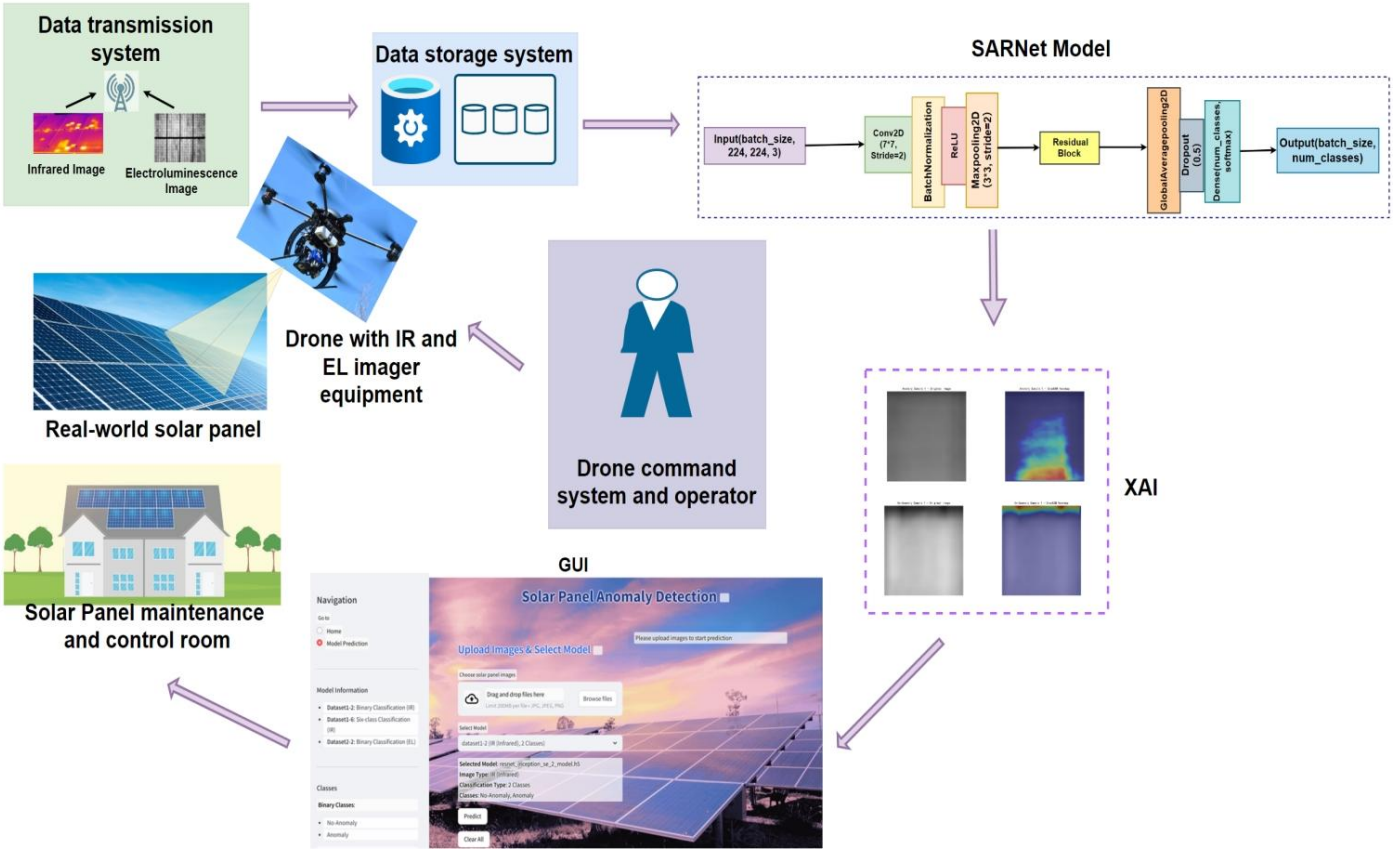


Figure 2. Research Overview.

2. LITERATURE REVIEW

This chapter provides a comprehensive review of various methods used for solar panel fault classification, highlighting the evolution from traditional techniques to more advanced machine learning and deep learning approaches.

2.1. Solar Panel Anomaly Classification Using Traditional Methods

The traditional methods for classifying anomalies in solar panels mainly include two categories: visual and thermal imaging methods (VTMs) and electrical detection methods (EBMs) [10]. Like Figure 3, each approach leverages distinct principles and tools to identify and classify faults, offering complementary strengths and addressing different diagnostic needs.

VTMs detect surface and internal faults in solar panels [11]. Infrared imaging identifies heat anomalies from electrical issues without extra sensors [12], offering a cost-effective solution for systems of all sizes. Visual inspection detects surface defects through direct observation, while electroluminescence imaging uses injected current to reveal internal issues like cracks and poor contacts with high accuracy [10]. These methods, though equipment-dependent, are essential for reliable fault detection.

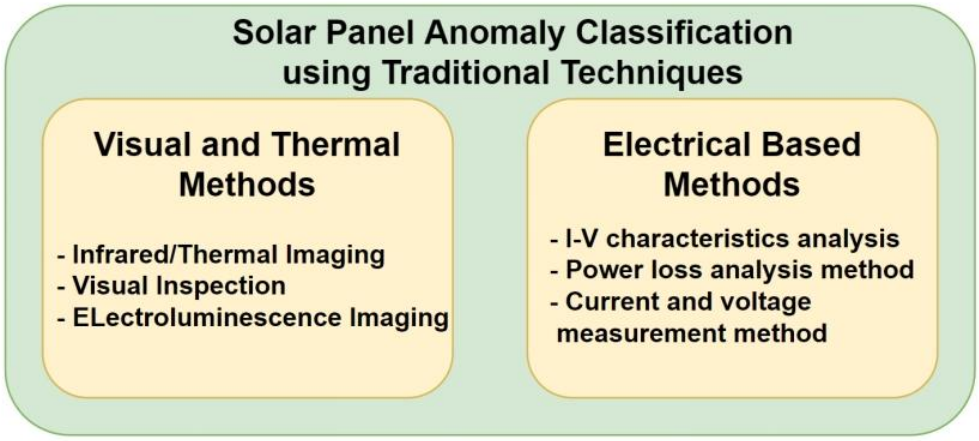


Figure 3. Categories of solar panel anomaly classification traditional methods.

EBMs analyze current and voltage characteristics to identify faults. I-V Curve Analysis detects issues like short circuits and shading by comparing normal and faulty curves [10]. Power Loss Analysis (PLA) locates problems such as hot spots and degradation by measuring output deviation. Current and Voltage Measurement identifies wiring faults or imbalances through deviations from expected electrical values [13].

2.2. Solar Panel Anomaly Classification Using Machine Learning Technologies

Various machine learning algorithms, such as k-nearest neighbor (k-NN), decision tree (DT), random forest (RF), and support vector machine (SVM), have been widely applied in anomaly detection and classification of solar panel systems [10].

The k-nearest neighbor (k-NN) algorithm is a simple yet effective method for detecting anomalies in solar panels, such as open-circuit, line-to-line, and partial shading issues. For instance, Madeti and Singh [14] used k-NN with experimental data and temperature variables, achieving 98.70% accuracy. Despite its simplicity, KNN's performance can be further improved by integrating fuzzy logic, enhancing adaptability in real-world conditions. The decision tree (DT) algorithm, known for simplicity and interpretability, is also commonly used for anomaly detection [10]. Madani et al. [15], combined DT with AdaBoost, enhancing its accuracy and robustness in dynamic environments.

Unlike the deterministic nature of decision trees (DT), random forests (RF) use an ensemble learning approach, improving prediction accuracy and robustness [10]. Dhibi et al. [16] proposed reduced-kernel RF have improved feature extraction and classification efficiency, showcasing the potential of RF in handling large-scale solar panel systems with diverse operating conditions.

Among these algorithms, support vector machine (SVM) is particularly is known for its strong classification in anomaly detection [10]. It handles high-dimensional data and separates classes clearly. Yi and Etemadi [17] combined multi-resolution signal decomposition (MSD) with SVM for line-to-line anomaly detection, achieving high accuracy.

2.3. Solar Panel Anomaly Classification Using Deep Learning Technologies

2.3.1. CNN Methods

Convolutional Neural Networks (CNNs) have become a dominant approach for solar panel anomaly classification, particularly due to their ability to automatically extract features from image data. This capability is crucial for monitoring large-scale solar panel systems, where datasets often include diverse modalities such as infrared (IR), electroluminescence (EL), and visible light images [10]. The adaptability of CNNs makes them effective for both binary and multi-class classification tasks.

For example, Alves et al. [18] used a CNN model with undersampling and oversampling to handle class imbalance, achieving 92.5% accuracy in binary classification and accuracies of 66.43% and 78.85% for datasets with 12 and 8 fault types. Espinosa et al. [19] applied CNNs for semantic segmentation and classification from RGB images, attaining 75% accuracy for binary fault detection and 70% for four fault categories.

As CNN-based models advanced, Le et al. [19] further enhanced CNN-based approaches by incorporating a deep neural network with a residual network structure and ensemble techniques, achieving 94% accuracy for binary classification and 86% for classifying 12 anomaly types.

In a significant development, Korkmaz and Acikgoz [20] used a multi-scale CNN with transfer learning on thermographic images, achieving 93.51% accuracy for 11 anomaly types. Deitsch et al. [23] applied an end-to-end deep CNN to classify normal and abnormal categories in 1,968 EL extracted cells, reaching 88.42% accuracy. Otamendi et al. [20]employed a CNN for cell-level anomaly detection, attaining 84% accuracy in distinguishing defective modules.

These advancements show that CNNs are constantly improving, enhancing the accuracy and robustness of solar panel fault classification for various anomalies.

2.3.2. Attention Mechanism

Attention mechanisms have been increasingly utilized in solar panel anomaly classification due to their ability to focus on salient features and enhance representation learning [21]. Liu et al. [22] used attention-based masking to remove non-salient regions and focus on relevant features. Bozorgtabar and Mahapatra [27] applied attention-based learnable masks to detect and localize anomalies. Sim et al. [28] used attention masking to eliminate unnecessary background information, improving generalization. Park et al. [29] utilized attention masking for anomaly detection and inpainting techniques for defect identification. Overall, these studies show attention mechanisms boost solar panel anomaly detection model accuracy and reliability.

2.3.3. Hybrid Attention-CNN Methods

The integration of attention mechanisms with CNN methods has proven highly effective in solar panel anomaly classification, offering improvements in defect detection. By leveraging attention mechanisms, models can focus on critical defect regions, enhancing the detection of fine-grained anomalies that traditional methods often miss. For instance, Zhang et al. [23] combined CNN with the CBAM attention module, achieving 95.22% accuracy across complex defect categories. Lee et al. [21] proposed a lightweight solution with pre-trained attention mechanisms.

Their dual-masking technique (AGDM) enhanced performance by efficiently extracting defect information, achieving an accuracy of 84.6% in binary classification (between normal and abnormal). Royal et al. [32] achieved an AUROC score of 0.97 in EL image anomaly detection, validating the potential of combining attention mechanisms with CNNs.

These methods collectively demonstrate that hybrid attention-CNN approaches excel in handling complex and diverse defect scenarios, offering high accuracy, efficiency, and adaptability for solar panel anomaly classification.

3. METHODOLOGY

This chapter outlines the experimental design for evaluating the proposed models. It introduces the datasets used, the architecture of the proposed CNN model, and the data preprocessing procedures. The training setup, including hyperparameter configuration and training strategies, is then described. Finally, the evaluation section summarizes the metrics used to assess model performance across different classification tasks.

3.1. Dataset

This study uses Infrared Solar Modules to support model development. During the training process, the Infrared Solar Modules dataset [29] as shown in and Figure 4 contains images captured with infrared cameras, focusing on detecting faults such as hot spots, cracks, and material degradation. The images have an initial resolution of 24 * 40 pixels. This dataset includes 12 classes: 11 anomaly classes and one class labeled ‘No-Anomaly,’ representing the absence of faults. This enables effective classification and detection of various solar panel issues, improving the efficiency and longevity of photovoltaic systems.

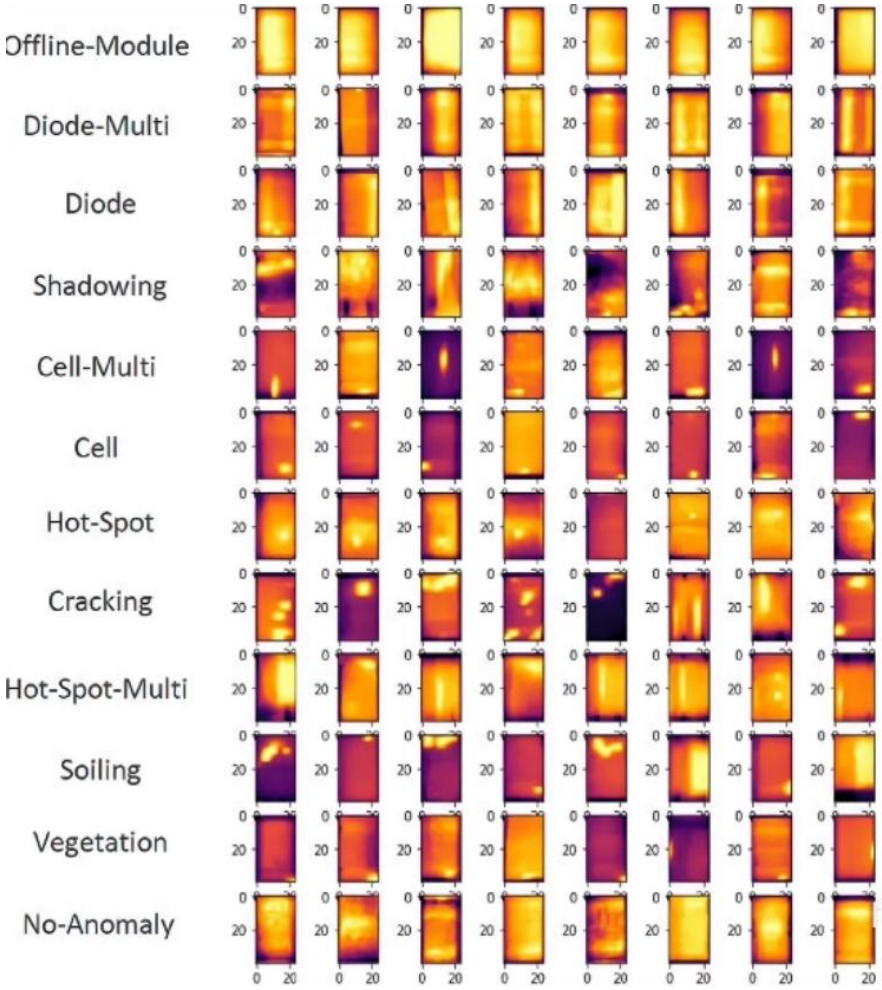


Figure 4. Few sample categories in the Infrared Solar Modules Dataset.

3.2. Data Preprocessing

Comprehensive preprocessing was conducted to ensure the datasets were suitable for model development. Key steps included data splitting, class balancing, resizing, and augmentation, which together produced standardized and diverse inputs, supporting both robust training and enhanced generalization across classification tasks.

3.2.1. Data Split

For this research, in Infrared Solar Modules dataset is divided into three subsets: training, validation, and test sets. As shown in Figure 5 and Figure 6, the training set comprises 70% of the data, while the validation and test set each account for 15%. The dataset is processed in both binary and multi-class formats. In binary classification, the task is to distinguish between ‘Anomaly’ and ‘No-Anomaly’. In the multi-class classification task, instead of using all 12 original categories, a subset of 6 representative classes was selected to address the issue of data imbalance. These include Vegetation, Shadowing, Cell, Diode, Cracking, and Offline-Module. This selection maintains defect diversity while ensuring a more balanced class distribution for effective model training. This structured split ensures effective training and evaluation for both classification tasks.

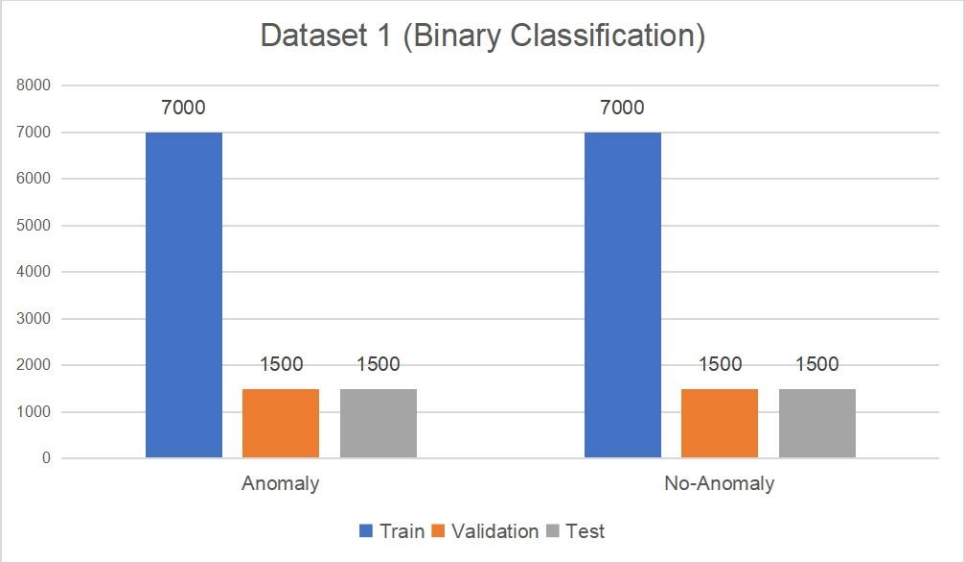


Figure 5. Data separation of dataset 1 (Binary Classification).

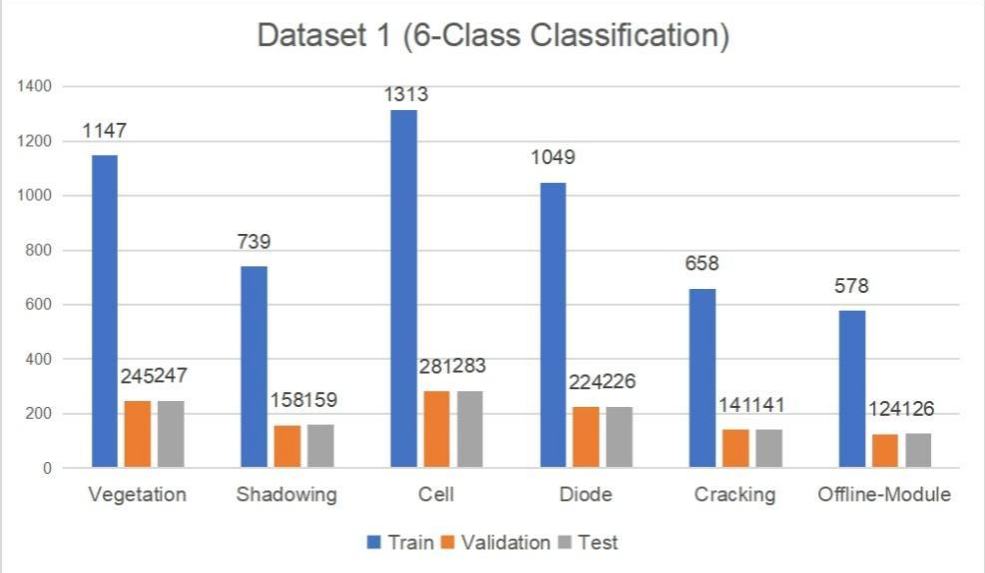


Figure 6. Data separation of dataset 1 (6-Class Classification).

3.2.2. Data Balancing

In this study, for the binary classification task of dataset, since the number of samples shows in Figure 7 for the anomaly class and the normal class are already balanced, no additional oversampling and undersampling is required, and the model can be directly trained.

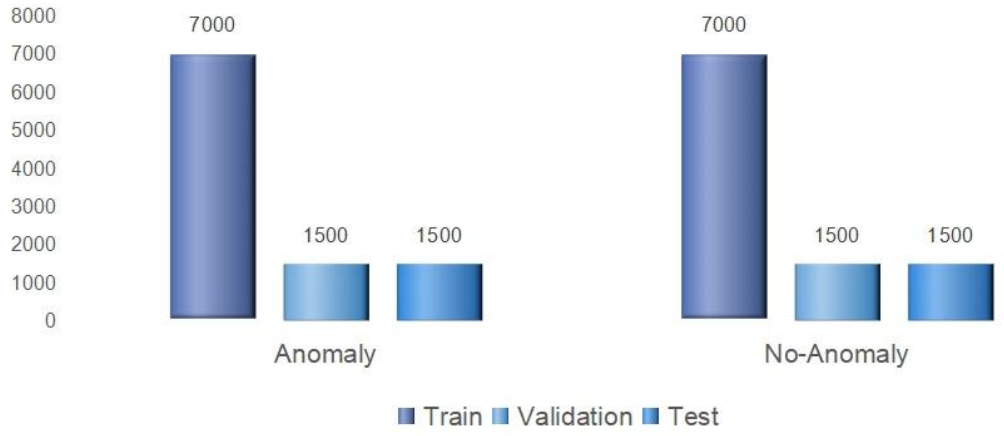


Figure 7. Balanced Data of Dataset (Binary Classification).

To fix class imbalance in the six-class task, a hybrid sampling method blending random undersampling and oversampling was used. The six anomaly categories are Vegetation, Diode, Shadowing, Cell, Cracking, and Offline Module, each set to have 1,000 samples. If a class had more images, random undersampling cut it to 1,000; if fewer, random oversampling brought it up. This balanced all six classes, curbing model bias. The balanced dataset in Figure 8 was then split into training, validation, and test subsets via stratified sampling to keep class ratios.

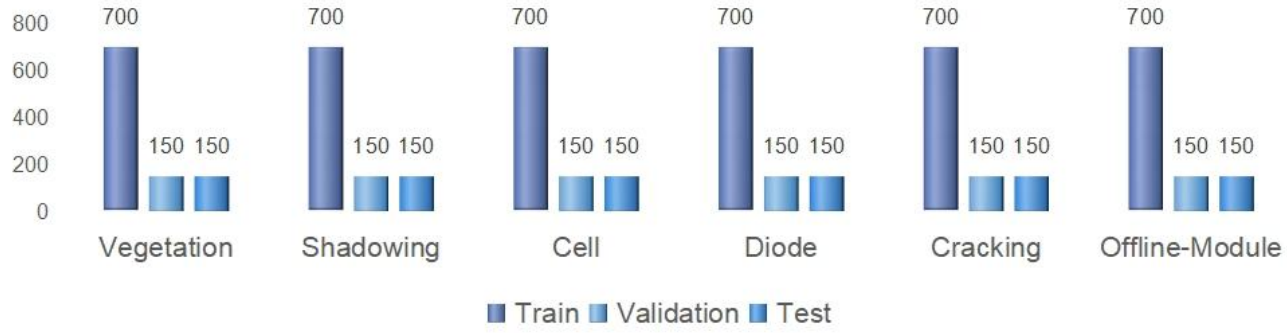


Figure 8. Balanced Data of Dataset 1 (6-Class Classification).

3.3. Data Resize and Augmentation

As mentioned in the introduction, the initial resolution of the images is 24 * 40 pixels. These resolutions are not suitable for direct input into the proposed model. Therefore, in this study, all images were resized to 224 * 224 pixels. This preprocessing step ensures that the images meet the input size requirements of the model, thereby enabling proper training and inference. Also, during the data preprocessing steps, data augmentation techniques such as random horizontal flipping, random rotation (up to ±10%), and random zoom (up to ±10%) random width and height shifts (up to 20%) and random shear (up to 20%) were applied.

3.4. Proposed Model Structure

The proposed CNN model, named SARNet, is built upon an enhanced architecture that incorporates StackNet as its backbone, integrating ResoNet modules and Squeeze-and-Excitation (SE) attention mechanisms. Together, these components enhance the model's representational power and focus on relevant features, improving classification accuracy and robustness in solar panel anomaly detection tasks.

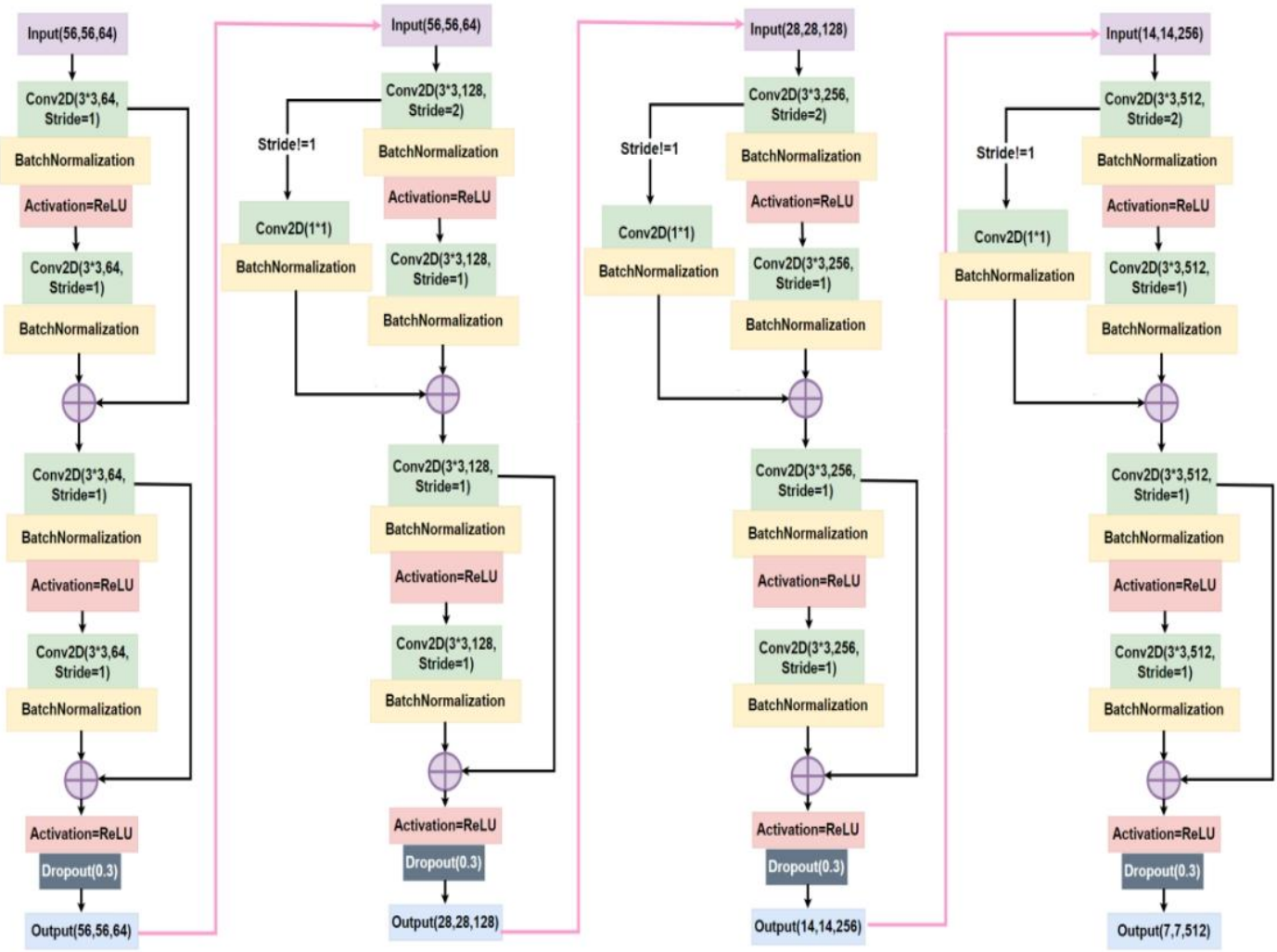


Figure 9. Residual block structure diagram for each stage.

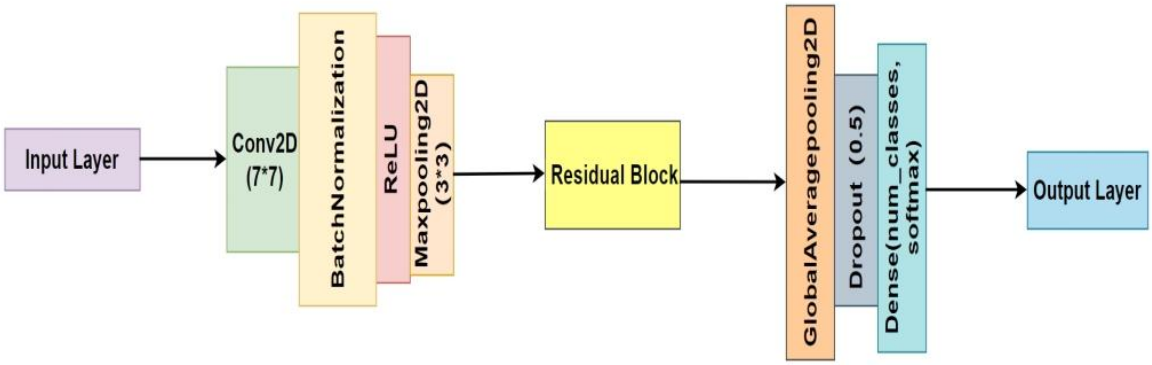


Figure 10. The proposed StackNet structure.

3.4.1. StackNet

The proposed StackNet architecture builds upon the ResNet-18 backbone, with targeted modifications to enhance feature extraction efficiency while maintaining computational lightness. The model begins with a large-kernel convolutional layer for effective low-level feature capture, followed by normalization, activation, and pooling to reduce spatial resolution early and accelerate training.

StackNet's core consists of four residual stages that incrementally increase channel dimensions from 64 to 512 while reducing spatial dimensions through stride convolutions. As shown in Figure 9, each residual block employs shortcut connections to maintain direct information flow, mitigating the vanishing gradient problem. When dimensional mismatches occur, 1×1 convolutions adjust the shortcut path. Additionally, dropout is integrated into each block to prevent overfitting. In Figure 10, StackNet strikes a balance between architectural depth and training efficiency, offering strong representational capacity with minimal increase in complexity. This makes it particularly suited for solar anomaly detection tasks where both performance and computational cost are critical.

3.5. Experimental Setup & Technology

3.5.1. Experimental Setup

The experimental setup for training the model involves configuring key hyperparameters, the optimizer, and several callback functions to ensure effective learning and model generalization.

The model training process is configured with 100 epochs and an initial learning rate of 0.001, employing the Adam optimizer for its proven effectiveness in deep learning tasks. Batch sizes are adapted to task complexity and dataset characteristics: 438 for binary classification and 29 for six-class classification. To enhance generalization and training efficiency, several callback mechanisms are employed. Early stopping monitors validation loss and terminates training after 30 consecutive epochs without improvement, restoring the best weights. The learning rate is adaptively reduced by a factor of 0.1 when no validation improvement is observed for 20 epochs, with a floor set at 0.000006 to maintain training stability. Model checkpoint ensures the best-performing weights, based on validation loss, are retained. These strategies collectively optimize convergence and mitigate overfitting, contributing to robust model performance.

3.5.2. Technology

The technology use in this research is displayed in Table 1:

Table 1. Summary of Relevant Technology Involved in This Research.

Software	Framework	Tensorflow
	Language	Python
	Libraries	Numpy, Keras, Matplolib, TensorFlow
Hardware	Central Processing Unit (CPU)	12th Gen Intel(R) Core(TM) i5-12500H 2.50 GHz
	Graphic Processing Unit (GPU)	Intel(R) Iris(R) Xe Graphics

3.6. Model Performance Evaluation Metrics

The 9 main evaluation metrics used in this experiment were: Loss, Accuracy, Specificity, Recall (Sensitivity), Precision, F1 Score (Dice Score), Confusion Matrix, ROC-AUC Curve, and Precision-Recall Curve. Each of these 9 metrics is described below.

3.6.1. Loss Function

Binary Cross-Entropy function is used, it is a loss function suitable for binary classification tasks, which guides the optimization process by measuring the difference between the predicted probability of the model and the actual label.

$$Binary\ Cross-Entropy\ Loss = -\frac{1}{N}\sum_{i=1}^{N}[y_i \log(p_i) + (1-y_i)\log(1-p_i)] \quad (1)$$

- y_i : the true label of the sample, taking values of 0 or 1.
- p_i : the model's predicted probability for the i-th sample being in class 1, typically output by the sigmoid activation function.
- N: the total number of samples.

Sparse classification cross entropy measures the quality of a model's predictions by taking the logarithm of its prediction probability for the correct category and taking a negative value. It is suitable for situations where the label is an integer representing the category index.

$$Sparse\ Categorical\ Cross-Entropy = -\sum_{i=1}^{C} y_i Log(p_i) \quad (2)$$

- y_i : the true label of the sample.
- p_i : the probability of the model predicting the category.
- C: the total number of categories in the classification task.

3.6.2. Accuracy

The proportion of correctly predicted samples by the model to the total sample size.

$$Accuracy(A, B) = \frac{1}{n}\sum_{i=0}^{n} 1_{(A_i=B_i)} \quad (3)$$

- A: the predicted values or predicted labels for the samples.
- B: the true or ground truth labels for the samples.
- n: the total number of samples.

3.6.3. Specificity (SP)

Specificity is used to measure the ability of classification models to correctly identify negative class samples.

$$Specificity\ (SP) = \frac{TN}{TN+FP} \quad (4)$$

- TN (True Negatives): The number of samples correctly predicted as negative by the model.
- FP (False Positives): The number of samples incorrectly predicted as positive by the model.

3.6.4. Sensitivity/Recall (SE)

Sensitivity/Recall is an indicator that measures how many actual positive class samples a classification model can recognize.

$$Sensitivity/Recall\ (SE) = \frac{TP}{TP+FN} = \frac{TP}{P} \quad (5)$$

- TP (True Positive): This represents the number of positive samples that were correctly identified as positive by the model.
- FN (False Negative): This represents the number of positive samples that were incorrectly identified as negative by the model.
- P: This is the total number of actual positive samples in the dataset, which is the sum of TP and FN.

3.6.5. Precision

Precision measures the proportion of samples predicted as positive by the model that are correct.

$$Precision = \frac{TP}{TP+FP} \quad (6)$$

3.6.6. F1 Score (Dice Score)

F1 Score is the harmonic mean of precision and recall.

$$F1\ Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (7)$$

3.6.7. Confusion Matrix

Confusion Matrix, like Table 2 is a tool used to evaluate the performance of classification models, especially for binary classification problems. It compares the predicted results of the model with the actual labels to form a matrix that displays four possible outcomes:

Table 2. Outcomes of Confusion Matrix.

	Predicted Positive (1)	Predicted Negative (0)
Actual Positive (1)	True Positive (TP)	False Negative (FN)
Actual Negative (0)	False Positive (FP)	True Negative (TN)

3.6.8. Receiver Operating Characteristic (ROC)

True Positive Rate (TPR): TPR represents the proportion of samples that are correctly predicted as positive among all samples that are actually positive.

$$TPR = \frac{TP}{TP + FN} \quad (8)$$

False Positive Rate (FPR): FPR represents the proportion of samples that are incorrectly predicted as positive among all samples that are actually negative.

$$FPR = \frac{FP}{FP + TN} \quad (9)$$

AUC (Area Under Curve): AUC represents the area under the ROC curve, and its value ranges from 0 to 1. The closer the AUC is to 1, the better the classification performance of the model.

3.6.9. Precision-Recall (PR)

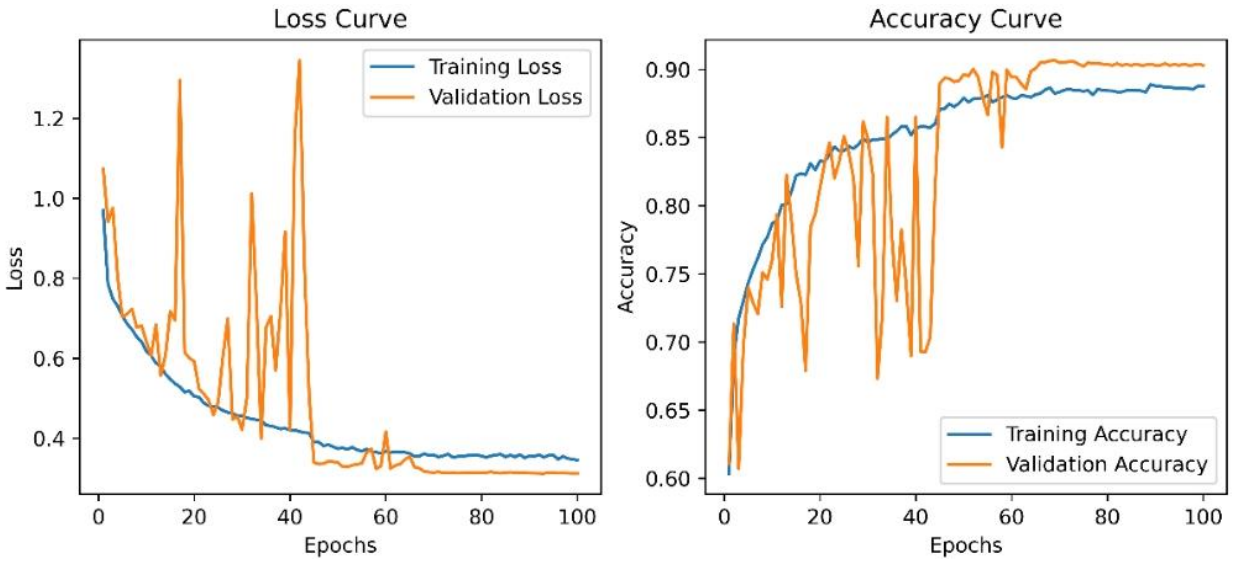
The Precision-Recall curve shows the changes in precision and recall under different decision thresholds. Precision and recall will vary under different thresholds, so the Precision-Recall curve helps analyze the performance of the model under various decision boundaries. The PR curve plots the changes in recall and precision and evaluates the performance of the model on a specific class through the area (AUC-PR).

4. EXPERIMENTAL RESULT AND ANALYSES

This study compares three CNN architectures under identical training settings to ensure fairness. StackNet employs residual connections for deep feature extraction and stable training. SRNet builds on StackNet by adding ResoNet modules to capture multi-scale spatial features. SARNet further introduces attention mechanisms to highlight task-relevant regions. All models are trained for 100 epochs at an initial learning rate of 0.001 using the Adam optimizer. Batch sizes are task-specific: 438 for binary classification, 29 for the six-class task. To enhance generalization, training incorporates early stopping after 30 stagnant epochs, checkpointing based on validation loss, and a learning rate reduction (by 0.1 after 20 stagnant epochs, with a minimum of 0.000006). These configurations ensure stable training and optimal model performance. This evaluation not only assesses StackNet's performance across varying datasets and classification tasks but also provides insights into which dataset facilitates better model optimization.

4.1. StackNet on Dataset 1- Binary Class

In the binary classification task, StackNet achieved the highest overall performance. As shown in Figure 11, the training loss reached 0.3492, while the validation loss was slightly higher at 0.3671, indicating minimal overfitting. In terms of accuracy, the model achieved 88.76 percent on the training set and 88.27 percent on the validation set, showing strong consistency. In Table 3 the test set achieved 88.3% accuracy and a loss of 0.367, confirming its effectiveness on unseen data.



(a) Train Loss = 0.3492, Val Loss = 0.3671 (b) Train ACC = 0.8876, Val ACC = 0.8827

Figure 11. (a) and (b) present StackNet model Accuracy and Loss Curve on Binary Class.

The model StackNet demonstrates strong classification performance with a high number of correct predictions. As shown in Figure 12, it correctly identified 1,373 true positives and 1,274 true negatives, while misclassifying 127 false positives and 226 false negatives. These results indicate a balanced performance with relatively low error rates, suggesting the model is effective at distinguishing between the two classes.

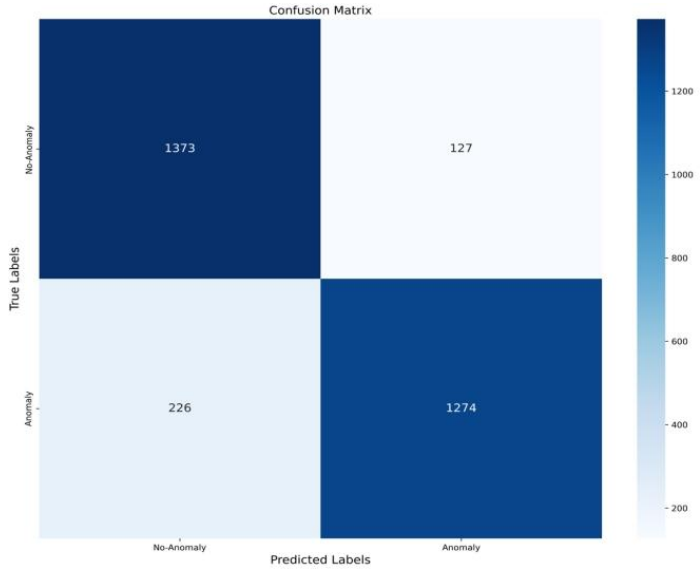


Figure 12. StackNet model Confusion Matrix on Binary Class.

The model StackNet achieved a ROC-AUC score of 0.95, as shown in Figure 13, indicating excellent discriminatory power between the positive and negative classes. This high value reflects the model's ability to maintain a strong balance between sensitivity and specificity across various threshold settings, suggesting reliable overall performance.

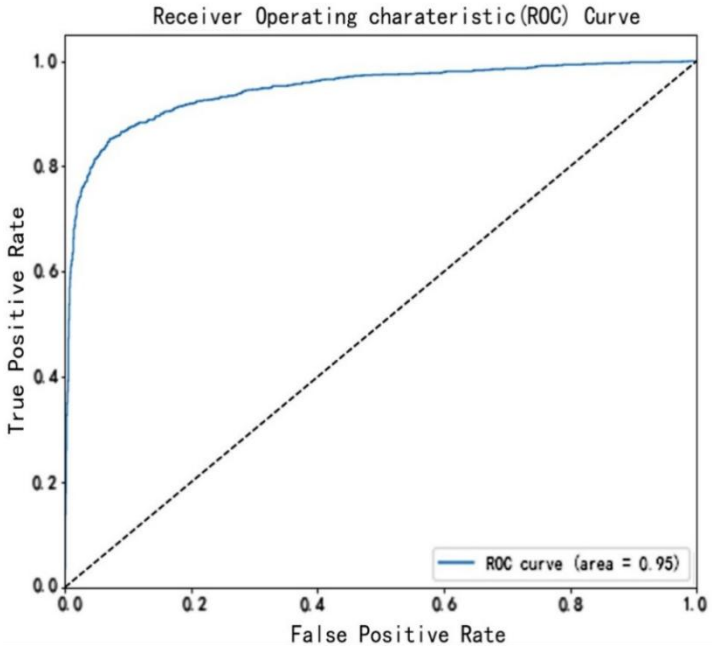


Figure 13. StackNet model ROC-AUC of 0.95 on Binary Class.

As shown in Figure 14, the StackNet model achieved a PR-AUC score of 0.95 on the binary class, demonstrating excellent precision-recall performance. This indicates that the model maintains high precision while effectively capturing the majority of positive cases, making it well-suited for handling class imbalance and minimizing false positives and false negatives.

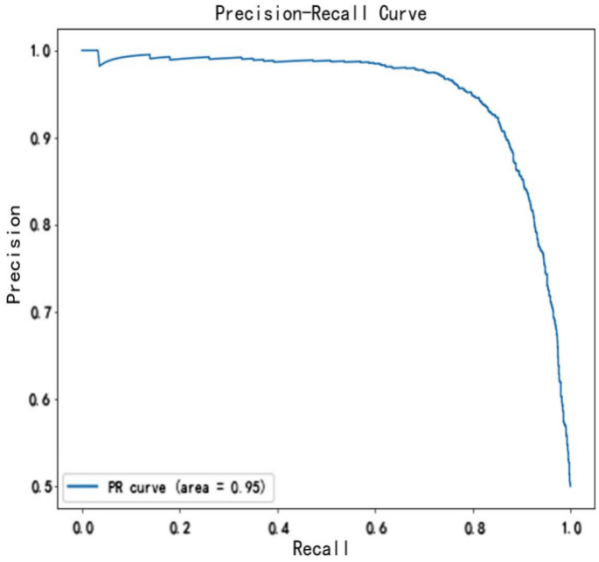
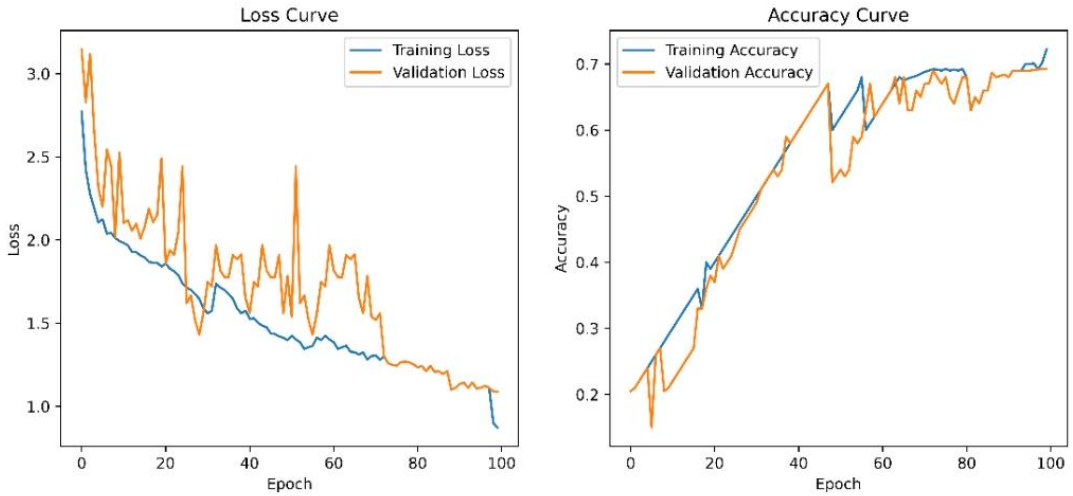


Figure 14. Present PR-AUC of 0.95 on Binary Class.

4.1.2. StackNet on Multi–Class Task

In contrast, the 6-class classification task posed greater challenges. As reflected in Figure 15, the training loss was 0.982 and the validation loss increased to 1.0872, indicating a moderate gap that may suggest some overfitting. The training accuracy reached 73.22 percent, while the validation and test accuracies both stood at 69.22 percent. The test loss matched the validation loss at 1.0872, confirming consistent but limited generalization. These results suggest that the model struggled to capture the complexity of the multi-class task compared to binary classification.



(a) Train Loss = 0.982, Val Loss = 1.0872 (b) Train ACC = 0.7322, Val ACC = 0.6922

Figure 15. (a) and (b). StackNet model Accuracy and Loss Curve on 6-Class.

The model shows varied performance across solar panel fault types, as shown in Figure 16. The model demonstrates strong performance in detecting class 2 diode and class 3 shadowing faults but faces challenges with class 6 cracking and class 4 cell faults. Class 1 vegetation and class 4 cell faults show some confusion with other categories. Overall, the model performs well on diode and shadowing but needs improvement for cracking and cell faults.

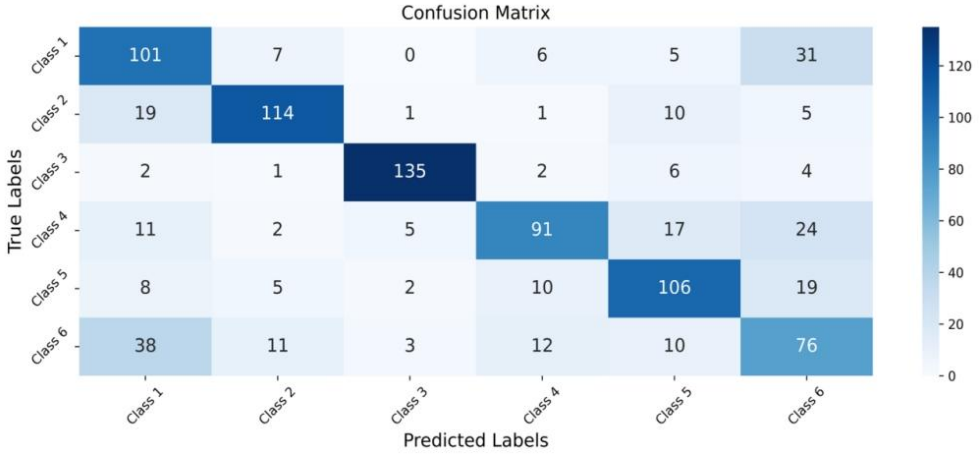


Figure 16. StackNet Model Confusion Matrix for 6-Class.

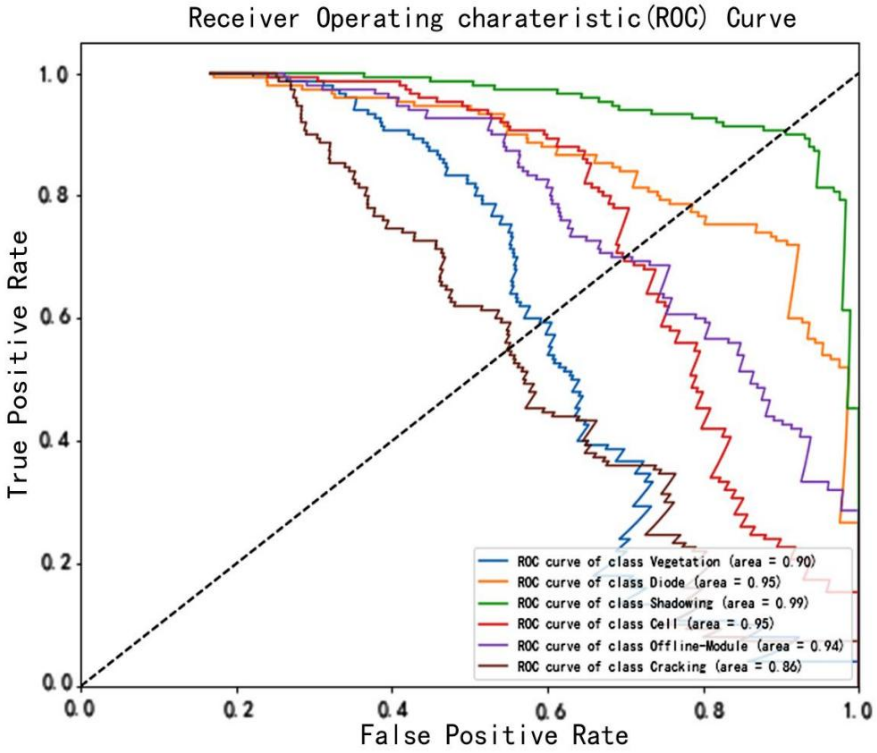


Figure 17. StackNet Model ROC Curve on 6-Class.

As shown in Figure 17, the ROC curve shows a multi-class classification model's performance across six classes. The dashed diagonal line represents a random classifier. Class Shadowing has the highest AUC of 0.99. Vegetation, Diode, Cell, and Offline-Module have AUC values above 0.90. Class Cracking has a lower but acceptable AUC of 0.86. Overall, the model shows good discriminatory power, with Shadowing performing best.

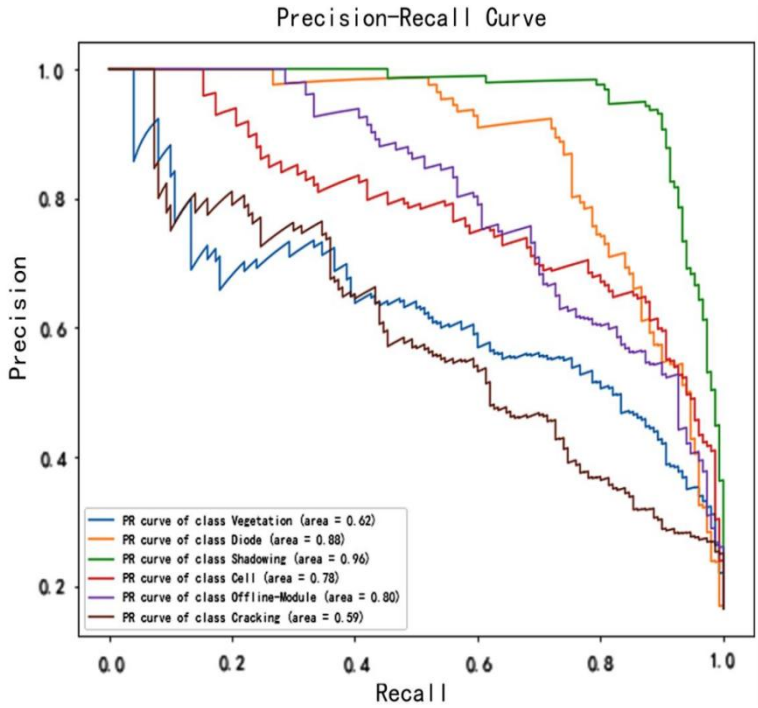


Figure 18. StackNet Model PR Curve on 6-Class.

Figure 18 presents the Precision-Recall curves for six classes in the multi-class model. Shadowing achieves the highest AUC value of 0.96, indicating excellent detection performance. Diode and Offline-Module follow with AUC values of 0.88 and 0.80, respectively. In contrast, Cracking and Vegetation yield lower AUC values of 0.59 and 0.62, highlighting their classification difficulty. Overall, the model performs well, though improvements are needed for specific fault types.

4.2. Current Experiments Results Summary

In this research, the architecture demonstrates varying levels of performance across classification tasks, emphasizing the critical influence of architectural design in solar panel fault classification.

The baseline StackNet serves as a reliable benchmark in the binary classification task, achieving a loss of 0.3671 and an accuracy of 88.27%. This significantly notifies both classification accuracy and robustness, especially under complex and diverse imaging conditions. However, StackNet struggles in the complex six-class classification, with PR-AUC scores dropping to 0.59 for Cracking and 0.62 for Vegetation. This indicates its limited ability to extract fine-grained features and handle faults with high spatial overlap or subtle visual signs, revealing its constraints in multi-class situations. It gets an F1-score of 0.91 in Cracking, versus StackNet's 0.87, and an overall F1-score of 0.816 compared to StackNet's 0.695. In Dataset 2's binary task, SARNet achieves 89.1% accuracy. These improvements validate the role of attention in complementing multi-scale feature extraction, particularly in identifying subtle or overlapping faults. Overall, the comparative experiments highlight the effectiveness of the proposed SARNet architecture. The integration of StackNet, ResoNet, and Attention modules significantly enhances both classification accuracy and robustness, especially under complex and diverse imaging conditions. The results of each experiment are summarized in Table 3.

Table 3. Compare models' Performance Metrics-1.

Model	Class	Loss	Accuracy	Recall	Precision	F1-Score
StackNet Model	2-class	0.3671	0.8827	0.8326	0.9251	0.8764
	6-class	1.0872	0.6922	0.6922	0.702	0.695
	2-class	0.571	0.878	0.818	0.783	0.800

4.3. Model Explainability

To enhance the interpretability of deep learning models in fault detection tasks, this research incorporates the Grad-CAM (Gradient-weighted Class Activation Mapping) algorithm as a post hoc visualization tool. Grad-CAM computes the gradients of the target class prediction with respect to the feature maps of a convolutional layer to produce a heatmap that highlights the regions the model focuses on when making a decision [32]. This method is model-agnostic and widely applicable, especially for CNNs. In this study, Grad-CAM was applied to visualize and interpret three specific classification tasks derived from the dataset:

- The model distinguishes between normal and abnormal solar modules. Grad-CAM visualizations, like Figure 19 and Figure 20 reveal that when an image is classified as anomaly, the attention maps typically highlight critical fault regions such as cracks, hotspots, or dirt accumulation, indicating that the model is focusing on meaningful defect areas within the infrared spectrum.

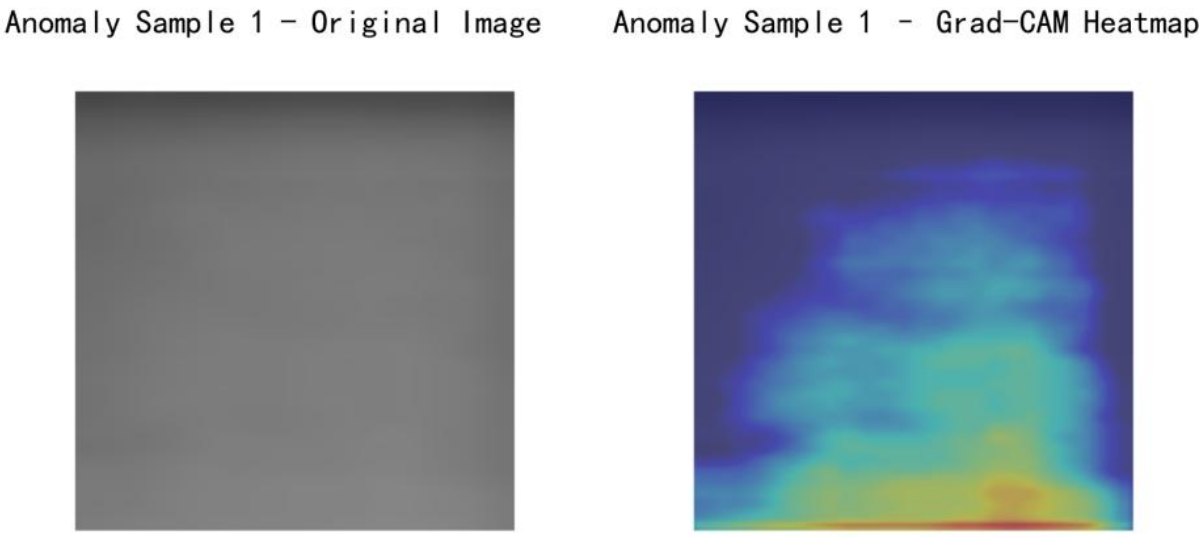


Figure 19. Grad-CAM Visualization of Binary Class on Anomaly Sample.

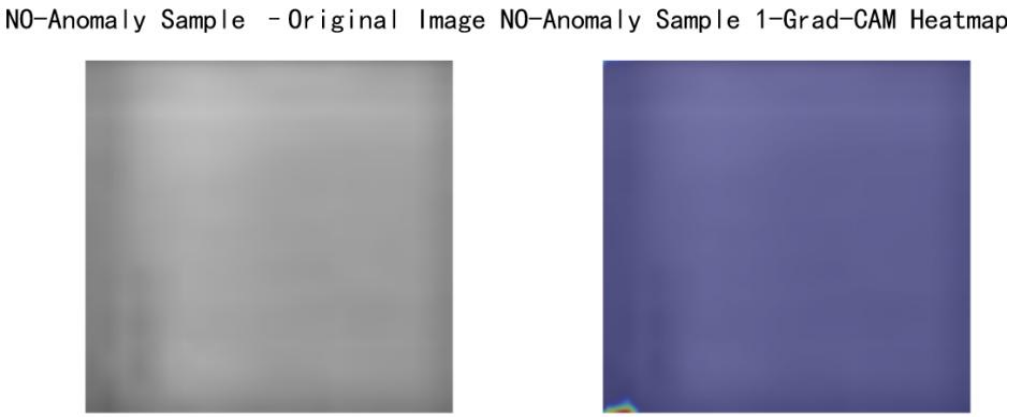


Figure 20. Grad-CAM Visualization of Binary Class No-Anomaly).

- The model categorizes faults into six types: Vegetation, Shadowing, Cell damage, Diode Faults, Cracking, and Offline-Module. Heatmaps show the model can target regions for each fault type, such as plant-covered areas for Vegetation or structural breaks for Cracking, proving its ability to distinguish multiple fault types in thermal images. Grad-CAM Visualization of six-class shows from Figure 21 to Figure 26.

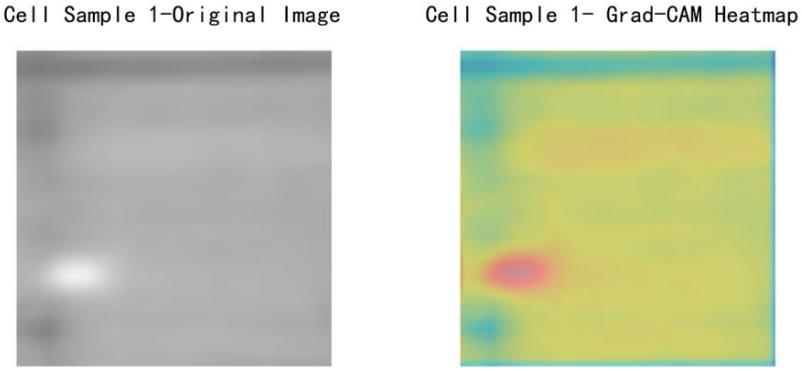


Figure 21. Grad-CAM Visualization of Cell.

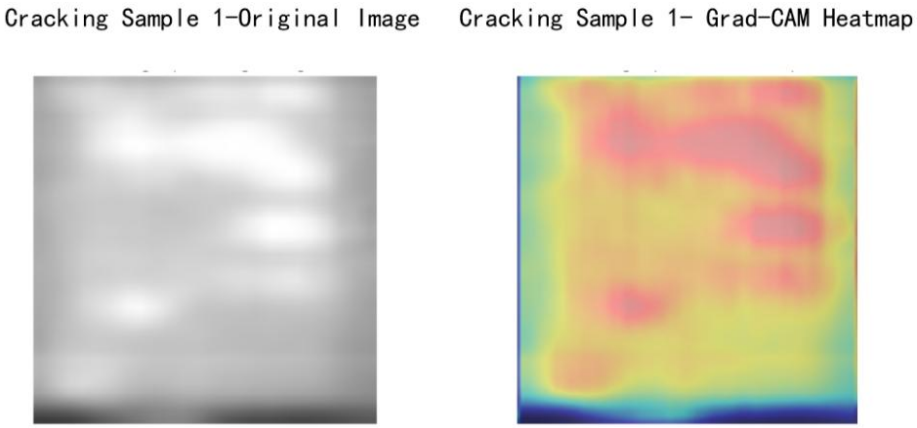


Figure 22. Grad-CAM Visualization of Cracking.

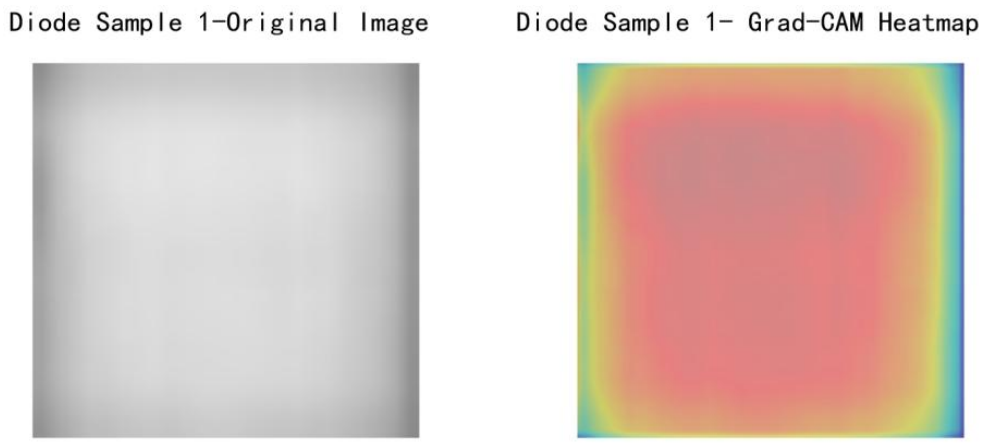


Figure 23. Grad-CAM Visualization of Diode.

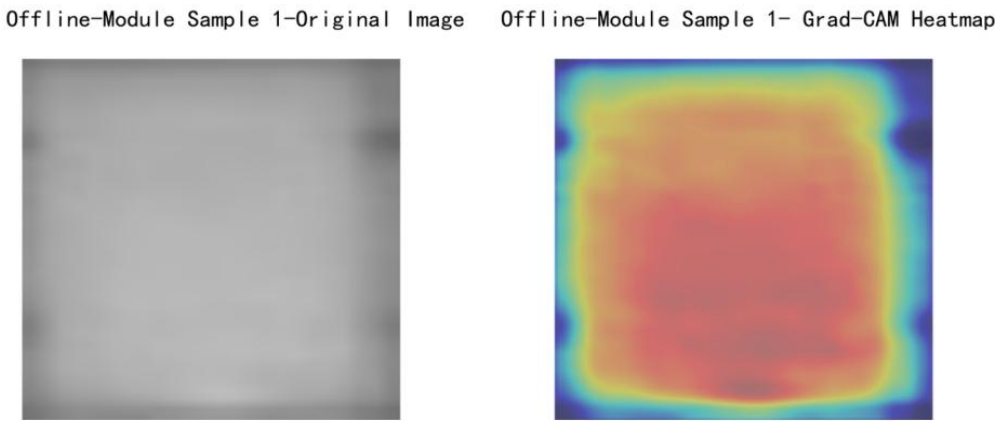


Figure 24. Grad-CAM Visualization of Offline-Module.

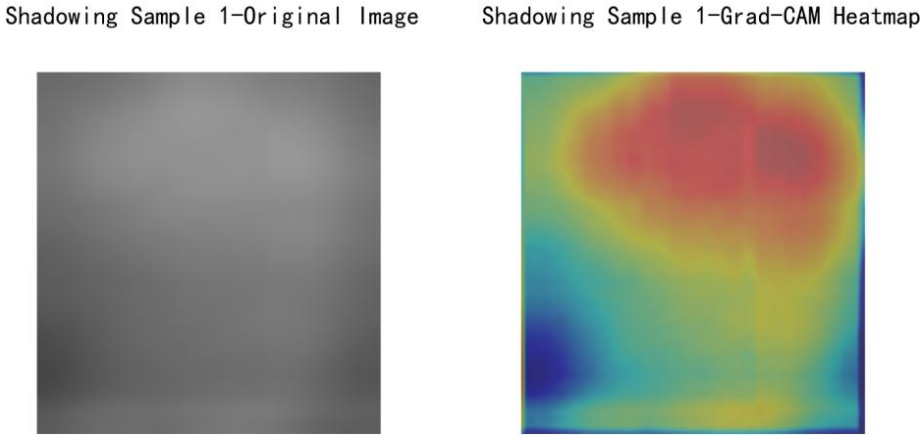


Figure 25. Grad-CAM Visualization of Shadowing.

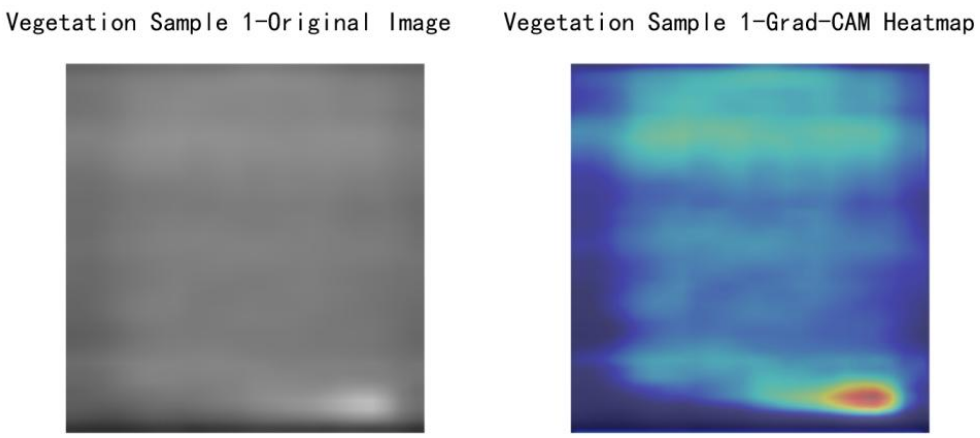


Figure 26. Grad-CAM Visualization of Vegetation.

5. CONCLUSION

In this paper, a deep learning-based solar panel fault detection system was developed and evaluated using an infrared (IR) dataset. The system employs a custom CNN with attention mechanisms for both binary and multi-class classification. Designed to handle low-and high-resolution images efficiently, the proposed StackNet model delivers accurate anomaly detection with strong deployment potential. On the binary classification task, it reached an accuracy of 91.7%, precision of 94.2%, and an F1-score of 91.4%, demonstrating its strong ability to distinguish between normal and faulty panels. In the six-class classification task using the same dataset, the model achieved an accuracy of 81.63% and a consistent F1-score and precision of approximately 81.6%, reflecting its robustness in identifying diverse fault types such as cracks, diode issues, and vegetation. The model's architecture was optimized to balance performance with computational efficiency, enabling faster inference without sacrificing accuracy. The user-oriented web-based GUI further enhances accessibility, allowing real-time prediction through simple image uploads, thus bridging the gap between deep learning research and field applications.

While the model performs well in both binary and multi-class classification scenarios, there remains room for further refinement. The primary limitation of the current model is its reduced precision in detecting faults with subtle or overlapping visual characteristics, particularly in categories such as cracking and vegetation. These faults often share similar visual patterns, making them difficult to distinguish. Additionally, the dataset suffers from annotation noise and inconsistent labeling, which may compromise model accuracy and generalization. This issue is especially problematic for the model’s performance in fault detection across diverse real-world conditions. Furthermore, the absence of paired multimodal samples, such as aligned infrared (IR) and electroluminescence (EL) images, hinders the model’s ability to exploit complementary features from both imaging modalities, preventing the model from fully benefiting from the combined insights they could provide for more accurate fault detection.

Future work will focus on improving model accuracy by implementing multi-label and refined classification techniques to better handle overlapping fault categories. To address annotation noise and inconsistent labeling, semi-supervised learning will be explored, along with refining the labeling process for greater reliability.

Additionally, aligning infrared (IR) and electroluminescence (EL) images will be prioritized to enable true multimodal learning, allowing the model to benefit from both modalities simultaneously. Developing time-efficient models for industrial applications will also be key, ensuring the model's effectiveness while reducing computational overhead. Real-world validation through field testing and the use of drone-captured IR or EL images will be essential for validating generalization and enhancing the model's applicability in dynamic environments, promoting automated and reliable solar panel maintenance strategies.

References

[1] Z. B. Duranay, 'Fault Detection in Solar Energy Systems: A Deep Learning Approach', *Electronics*, vol. 12, no. 21, Art. no. 21, Jan. 2023, doi: 10.3390/electronics12214397.

[2] M. K. H. Rabaia *et al.*, 'Environmental impacts of solar energy systems: A review', *Sci. Total Environ.*, vol. 754, p. 141989, Feb. 2021, doi: 10.1016/j.scitotenv.2020.141989.

[3] S. H. Han, T. Rahim, and S. Y. Shin, 'Detection of Faults in Solar Panels Using Deep Learning', in *2021 International Conference on Electronics, Information, and Communication (ICEIC)*, Jan. 2021, pp. 1–4. doi: 10.1109/ICEIC51217.2021.9369744.

[4] R. Al-Mashhadani *et al.*, 'Deep Learning Methods for Solar Fault Detection and Classification: A Review', *Inf. Sci. Lett.*, vol. 10, p. 323, May 2021, doi: 10.18576/isl/100213.

[5] S. Jaybhaye, V. Sirvi, S. Srivastava, V. Loya, V. Gujarathi, and M. D. Jaybhaye, 'Classification and Early Detection of Solar Panel Faults with Deep Neural Network Using Aerial and Electroluminescence Images', *J. Fail. Anal. Prev.*, vol. 24, no. 4, pp. 1746–1758, Aug. 2024, doi: 10.1007/s11668-024-01959-x.

[6] S. Jaybhaye, O. Thakur, R. Yardi, V. Raut, and A. Raut, 'Solar Panel Damage Detection and Localization of Thermal Images', *J. Fail. Anal. Prev.*, vol. 23, no. 5, pp. 1980–1990, Oct. 2023, doi: 10.1007/s11668-023-01747-z.

[7] M. Waqar Akram *et al.*, 'Improved outdoor thermography and processing of infrared images for defect detection in PV modules', *Sol. Energy*, vol. 190, pp. 549–560, Sep. 2019, doi: 10.1016/j.solener.2019.08.061.

[8] S. Gallardo-Saavedra, L. Hernández-Callejo, M. del C. Alonso-García, J. Muñoz-Cruzado-Alba, and J. Ballestín-Fuertes, 'Infrared Thermography for the Detection and Characterization of Photovoltaic Defects: Comparison between Illumination and Dark Conditions', *Sensors*, vol. 20, no. 16, p. 4395, Aug. 2020, doi: 10.3390/s20164395.

[9] O. K. Segbefia, N. Akhtar, and T. O. Sætre, 'Defects and fault modes of field-aged photovoltaic modules in the Nordics', *Energy Rep.*, vol. 9, pp. 3104–3119, Dec. 2023,doi: 10.1016/j.egyr.2023.01.126.

[10] Y.-Y. Hong and R. A. Pula, 'Methods of photovoltaic fault detection and classification: A review', *Energy Rep.*, vol. 8, pp. 5898–5929, Nov. 2022, doi: 10.1016/j.egyr.2022.04.043.

[11] H. Bhatia, '10 Causes of Solar Panel Damage and How to Avoid Them', Sustvest Blog. Accessed: Mar. 11, 2025. [Online]. Available: https://blog.sustvest.com/solar-panel-damage/

[12] M. Cubukcu and A. Akanalci, 'Real-time inspection and determination methods of faults on photovoltaic power systems by thermal imaging in Turkey', *Renew. Energy*, vol. 147, pp. 1231–1238, Mar. 2020, doi: 10.1016/j.renene.2019.09.075.

[13] A. Triki-Lahiani, A. Bennani-Ben Abdelghani, and I. Slama-Belkhodja, 'Fault detection and monitoring systems for photovoltaic installations: A review', *Renew. Sustain. Energy Rev.*, vol. 82, pp. 2680–2692, Feb. 2018, doi: 10.1016/j.rser.2017.09.101.

[14] S. R. Madeti and S. N. Singh, 'Modeling of PV system based on experimental data for fault detection using kNN method', *Sol. Energy*, vol. 173, pp. 139–151, Oct. 2018, doi: 10.1016/j.solener.2018.07.038.

[15] S. S. Madani, A. Abbaspour, M. Beiraghi, P. Z. Dehkordi, and A. M. Ranjbar, 'Islanding detection for PV and DFIG using decision tree and AdaBoost algorithm', in *2012 3rd IEEE PES Innovative Smart Grid Technologies Europe (ISGT Europe)*, Oct. 2012, pp. 1–8. doi: 10.1109/ISGTEurope.2012.6465818.

[16] K. Dhibi *et al.*, 'Reduced Kernel Random Forest Technique for Fault Detection and Classification in Grid-Tied PV Systems', *IEEE J. Photovolt.*, vol. 10, no. 6, pp. 1864–1871, Nov. 2020, doi: 10.1109/JPHOTOV.2020.3011068.

[17] Z. Yi and A. H. Etemadi, 'A novel detection algorithm for Line-to-Line faults in Photovoltaic (PV) arrays based on support vector machine (SVM)', in *2016 IEEE Power and Energy Society General Meeting (PESGM)*, Jul. 2016, pp. 1–4. doi: 10.1109/PESGM.2016.7742026.

[18] R. H. Fonseca Alves, G. A. de Deus Júnior, E. G. Marra, and R. P. Lemos, 'Automatic fault classification in photovoltaic modules using Convolutional Neural Networks', *Renew. Energy*, vol. 179, pp. 502–516, Dec. 2021, doi: 10.1016/j.renene.2021.07.070.

[19] M. Le, V. S. Luong, D. K. Nguyen, V.-D. Dao, N. H. Vu, and H. H. T. Vu, 'Remote anomaly detection and classification of solar photovoltaic modules based on deep neural network', *Sustain. Energy Technol. Assess.*, vol. 48, p. 101545, Dec. 2021, doi: 10.1016/j.seta.2021.101545.

[20] D. Korkmaz and H. Acikgoz, 'An efficient fault classification method in solar photovoltaic modules using transfer learning and multi-scale convolutional neural network', *Eng. Appl. Artif. Intell.*, vol. 113, p. 104959, Aug. 2022, doi: 10.1016/j.engappai.2022.104959.

[21] H. Lee, Y. Park, and J. Yi, 'Enhancing Defective Solar Panel Detection with Attention-Guided Statistical Features Using Pre-Trained Neural Networks', in *2024 IEEE International Conference on Big Data and Smart Computing (BigComp)*, Feb. 2024, pp. 219–225. doi: 10.1109/BigComp60711.2024.00042.

[22] Z. Liu, J. Gui, and H. Luo, 'Good Helper Is around You: Attention-Driven Masked Image Modeling', *Proc. AAAI Conf. Artif. Intell.*, vol. 37, no. 2, Art. no. 2, Jun. 2023, doi: 10.1609/aaai.v37i2.25269.

[23] H. Zhang, Z. Yang, and N. Lei, 'Defect recognition of solar panel in EfficientNet-B3 network based on CBAM attention mechanism', in *Proceedings of the 2024 International Conference on Generative Artificial Intelligence and Information Security*, in GAIIS '24. New York, NY, USA: Association for Computing Machinery, Jul. 2024, pp. 349–352. doi: 10.1145/3665348.3665407.

[24] A. Rico Espinosa, M. Bressan, and L. F. Giraldo, 'Failure signature classification in solar photovoltaic plants using RGB images and convolutional neural networks', *Renew. Energy*, vol. 162, pp. 249–256, Dec. 2020, doi: 10.1016/j.renene.2020.07.154

[25] S. Deitsch *et al.*, 'Automatic classification of defective photovoltaic module cells in electroluminescence images', *Sol. Energy*, vol. 185, pp. 455–468, Jun. 2019, doi: 10.1016/j.solener.2019.02.067.

[26] U. Otamendi, I. Martinez, M. Quartulli, I. G. Olaizola, E. Viles, and W. Cambarau, 'Segmentation of cell-level anomalies in electroluminescence images of photovoltaic modules', *Sol. Energy*, vol. 220, pp. 914–926, May 2021, doi: 10.1016/j.solener.2021.03.058.

[27] Y. Park, M. J. Kim, U. Gim, and J. Yi, 'Boost-Up Efficiency of Defective Solar Panel Detection With Pre-Trained Attention Recycling', *IEEE Trans. Ind. Appl.*, vol. 59, no. 3, pp. 3110–3120, May 2023, doi: 10.1109/TIA.2023.3255227.

[28] R. Rahimzadeh, T. Brox, and J. Steinmetz, 'State-of-the-art deep learning anomaly detection method for analyzing electroluminescence images of solar cells', *AIP Conf. Proc.*, vol. 2826, no. 1, p. 060002, Jun. 2023, doi: 10.1063/5.0141116.

[29] 'Infrared Solar Modules'. Accessed: Dec. 18, 2024. [Online]. Available: https://www.kaggle.com/datasets/marcosgabriel/infrared-solar-modules

[30] 'Papers with Code - ELPV Dataset'. Accessed: Apr. 07, 2025. [Online]. Available: https://paperswithcode.com/dataset/elpv

[31] J. Hu, L. Shen, S. Albanie, G. Sun, and E. Wu, 'Squeeze-and-Excitation Networks', May 16, 2019, *arXiv*: arXiv:1709.01507. doi: 10.48550/arXiv.1709.01507.

[32] 'Multilayer Grad-CAM: An effective tool towards explainable deep neural networks for intelligent fault diagnosis', *J. Manuf. Syst.*, vol. 69, pp. 20–30, Aug. 2023, doi: 10.1016/j.jmsy.2023.05.027.

[33] 'Bridging the gap: Reintegrating legal perspectives into project management', *Proj. Leadersh. Soc.*, vol. 5, p. 100154, Dec. 2024, doi: 10.1016/j.plas.2024.100154.

[34] R. Wagner, 'Projects and Project Management from a Social Perspective', no. 1, 2022.

[35] 'Ethics in Project Management: Integrity and Accountability | IPM'. Accessed: Apr. 13, 2025. [Online].Available:https://instituteprojectmanagement.com/blog/project-management-ethics-importance-of-integrity-and-accountability/