FOREX Publication
Open Access | Rapid and quality publishing

International Journal of Electrical and Electronics Research (IJEER)
Research Article | Volume 13, Issue 4 | Pages 712-721 | e-ISSN: 2347-470X

Comparative Deep Learning for RGB-Based PV Surface Fault Classification Using ResNet50 and EfficientNetB0 with Real-Time Deployment

Mosbah Laouamer[1*], Mohammed Adaika[2], Souhaib Remha[3], Abdelkader Mahmoudi[4], and Hamza Adaika[5]

[1]*UDERZA Unit, Mechanical Department, Faculty of Technology, University of El Oued, 39000 El Oued, Algeria; Email: laouamer-mosbah@univ-eloued.dz*

[2]*LPMRN Laboratory, Department of Electromechanical Engineering, University of Bordj Bou Arreridj, El-Anasser, 34030, Bordj Bou Arreridj, Algeria; Email: mohammed.adaika@univ-bba.dz*

[3]*UDERZA Unit, Mechanical Department, Faculty of Technology, University of El Oued, 39000 El Oued, Algeria; Email: remha-souhaib@univ-eloued.dz*

[4]*UDERZA Unit, Mechanical Department, Faculty of Technology, University of El Oued, 39000 El Oued, Algeria; Email: abdelkader-mahmoudi@univ-eloued.dz*

[5]*University of El Oued, 39000 El Oued, Algeria; Email: hamza-adaika@univ-eloued.dz*

***Correspondence:** *Mosbah Laouamer; Email: laouamer-mosbah@univ-eloued.dz; Ph.: +213-0698529926*

ABSTRACT- Photovoltaic (PV) systems are increasingly vital to global energy transitions but remain vulnerable to surface anomalies such as dust, snow, bird droppings, and physical or electrical damage—that significantly reduce power yield and long-term reliability. Manual inspection is inefficient for large-scale solar farms, motivating the development of intelligent and automated fault detection systems. This study introduces a reproducible comparative deep-learning framework that systematically benchmarks four convolutional neural network (CNN) architectures ResNet50, EfficientNetB0, MobileNetV3Small, and DenseNet121 under identical preprocessing, training, and validation settings. The framework integrates runtime efficiency analysis, five-fold cross-validation, and a real-time GUI-based deployment interface, bridging the gap between academic benchmarking and field-level implementation. A six-class labeled dataset of 1,574 RGB images was expanded through extensive data augmentation (rotation, flipping, brightness adjustment, and Gaussian noise perturbation) to simulate diverse real-world conditions. Among the four tested models, DenseNet121 achieved the highest macro-averaged F1-score (≈ 0.96), followed by ResNet50 (0.93), EfficientNetB0 (0.92), and MobileNetV3Small (0.92), highlighting clear accuracy–efficiency trade-offs across architectures. The novelty of this work lies in its multi-model benchmarking design and transparent methodology, providing a standardized and reproducible reference for future PV image-based diagnostics. Practically, integrating the models into a real-time graphical user interface (GUI) demonstrates their feasibility for UAV-based or on-site PV inspection. From an operational and economic perspective, this approach supports cost-effective, scalable, and non-invasive monitoring solutions tailored for modern large-scale solar farms.

Keywords: Photovoltaic fault classification, Deep learning, Surface anomaly detection, ResNet50, EfficientNetB0, MobileNetV3Small, DenseNet121, RGB image analysis.

ARTICLE INFORMATION

Author(s): Mosbah Laouamer, Mohammed Adaika, Souhaib Remha, Abdelkader Mahmoudi, and Hamza Adaika;

Received: 27/08/2025; **Accepted**: 30/10/2025; **Published**: 10/12/2025;
E- ISSN: 2347-470X;
Paper Id: IJEER 2708A20;
Citation: 10.37391/ijeer.130411
Webpage-link:
https://ijeer.forexjournal.co.in/archive/volume-13/ijeer-130411.html

Publisher's Note: FOREX Publication stays neutral with regard to jurisdictional claims in Published maps and institutional affiliations.

1. INTRODUCTION

Photovoltaic (PV) systems have become a cornerstone of global renewable energy strategies due to their scalability, declining costs, and contribution to carbon neutrality goals. According to the International Renewable Energy Agency (IRENA), global PV capacity exceeded 1.3 terawatts by the end of 2024 and continues to expand rapidly under worldwide decarbonization initiatives [1–3].

Despite this unprecedented growth, PV modules remain susceptible to surface-level faults such as dust accumulation, humidity, snow deposition, and partial shading, which collectively reduce energy yield, induce hotspot formation, and accelerate material degradation. For instance, studies in desert regions indicate that heavy dust accumulation over a two-month period can reduce power output by nearly 8–10%, while localized shading can generate hotspots exceeding 150 °C, leading to microcracks and solder joint failures [4, 5].

FOREX Publication
Open Access | Rapid and quality publishing
International Journal of Electrical and Electronics Research (IJEER)
Research Article | Volume 13, Issue 4 | Pages 712-721 | e-ISSN: 2347-470X

To sustain optimal efficiency in large-scale installations, automated monitoring and fault detection have become essential components of modern PV management systems. Traditional inspection techniques—such as manual visual checks or handheld thermography—are labor-intensive, error-prone, and impractical for geographically distributed solar plants [1]. In response, deep learning has emerged as a powerful paradigm capable of automating visual inspection tasks through image-based analysis. Among various models, convolutional neural networks (CNNs) have demonstrated outstanding capability to learn hierarchical spatial features and detect nonlinear patterns indicative of PV surface anomalies. Researchers have validated CNNs across different imaging modalities, including infrared thermography and electroluminescence (EL), achieving high accuracy in identifying microcracks and hotspots [6, 7]. Similarly, RGB-based CNN models have proven effective in detecting dust, snow, and mechanical damage directly from standard photographs [8, 9].

Among modern CNN architectures, ResNet leverages deep residual learning for enhanced feature extraction, while Efficient Net introduces compound scaling and depth wise separable convolutions to optimize accuracy versus computation cost [8–10]. These frameworks have shown promising results in solar fault diagnosis, yet most prior studies either used non-standard datasets or evaluated models under different preprocessing and training conditions, hindering fair comparison. Consequently, there remains a clear gap for a standardized, reproducible benchmarking framework that evaluates multiple CNN architectures under identical experimental configurations—not only in terms of accuracy but also computational efficiency and real-time feasibility.

Addressing this limitation, the present study performs a comprehensive four-model comparison involving ResNet50, EfficientNetB0, MobileNetV3Small, and DenseNet121 using a unified training pipeline. To counteract dataset constraints, extensive data augmentation (rotation, flipping, brightness variation, and Gaussian noise perturbation) and 5-Fold Cross-Validation were employed to enhance diversity and ensure statistical robustness.

The novelty of this research lies in establishing a reproducible multi-architecture evaluation framework that integrates cross-validation, runtime-aware analysis, and GUI-based interpretability, thereby bridging the gap between academic benchmarking and field-level deployment. This approach advances beyond conventional accuracy-centric studies by enabling transparent, scalable, and real-time PV fault diagnosis, providing both scientific reproducibility and direct industrial relevance.

2. MATERIALS AND METHODS

2.1. Dataset Description

This study utilizes the PV Panel Defect Dataset compiled by Lenarczyk [11], comprising 1,574 manually annotated RGB images depicting various surface anomalies observed in photovoltaic (PV) modules. The dataset, intended for academic and educational use, includes six defect categories: Clean, Dusty, Bird-drop, Snow-covered, Electrical-damage, and Physical-damage. All images were collected from verified open-access repositories, including Kaggle and other publicly available databases, ensuring class diversity, image realism, and reproducibility.

To ensure fair and unbiased model evaluation, the dataset was stratified into three disjoint subsets:
1. Training subset, used for learning generalizable patterns;
2.Validation subset, applied during model tuning to assess overfitting;
3. Test subset, held out entirely for final performance evaluation.

This proportional stratification mitigated class imbalance and provided consistent representation of each fault type across all subsets.

Given the moderate dataset size (1,574 images), extensive data augmentation was applied to enhance visual diversity and reduce overfitting. The applied transformations included random rotation, horizontal and vertical flipping, zoom scaling, brightness variation, and Gaussian noise perturbation. Collectively, these operations increased the effective dataset size fivefold, enriching image variability across lighting conditions, dust density, and camera angles—thus improving the model's generalization to real-world scenarios.

To further improve statistical robustness, a 5-fold cross-validation strategy was employed. This ensured that every sample contributed to both training and validation phases at least once, thereby reducing performance variance and improving model reliability across architectures.

Although the current dataset includes only RGB imagery, the proposed framework was designed for extensibility, remaining compatible with multi-modal data sources such as infrared (IR) and electroluminescence (EL) images, which can capture both surface and sub-surface defects in PV cells.

Compared with existing PV fault datasets—many of which are unbalanced, domain-limited, or lack standardized labeling—this dataset offers a more systematically stratified and reproducible structure, making it well suited for multi-architecture benchmarking and fair performance comparison.

Table 1 summarizes the distribution of images per PV fault class across training, validation, and test subsets, before and after augmentation. The table highlights dataset uniformity and shows that augmentation was applied exclusively to the training subset, leaving validation and test data unchanged to preserve unbiased evaluation integrity.

FOREX Publication
Open Access | Rapid and quality publishing
International Journal of Electrical and Electronics Research (IJEER)
Research Article | Volume 13, Issue 4 | Pages 712-721 | e-ISSN: 2347-470X

Table 1. Distribution of images per PV fault class across training, validation, and test subsets (before and after augmentation)

Fault Type	Training	Validation	Testing	Total (Original)	Training (Augmented additions)	Total (Augmented additions)
Clean	169	102	18	289	676	796
Dusty	162	97	16	275	648	761
Bird-drop	177	104	17	298	705	826
Snow-covered	154	92	16	262	616	724
Electrical-damage	135	77	13	225	531	621
Physical-damage	132	78	15	225	528	621
Total Images	929	550	95	1574	3704	4349

2.2. Data Preprocessing

All RGB input images were resized to a standardized resolution of 224 × 224 pixels and normalized according to the preprocessing conventions associated with each pretrained architecture. Prior to training, the dataset was stratified into training, validation, and test subsets, ensuring class balance and randomized shuffling across all folds of the 5-fold cross-validation scheme. For transfer learning, the convolutional backbone of each pretrained CNN—ResNet50, EfficientNetB0, MobileNetV3Small, and DenseNet121 was initialized with ImageNet weights and kept frozen during initial training to preserve learned low-level feature representations. A lightweight classification head was then appended to each model, composed of a Global Average Pooling layer to reduce feature map dimensionality, Batch Normalization to stabilize gradient propagation, a Dropout layer (rate = 0.3) for regularization, and a SoftMax-activated dense layer configured for six-class output.

This unified preprocessing and model assembly pipeline ensured methodological consistency across all architectures, enabling a fair and transparent cross-model performance comparison in both accuracy and computational efficiency. Model training was carried out using a fixed batch size of 16, optimized with the Adam optimizer and Sparse categorical cross-entropy loss function. To enhance model generalization, early stopping was applied with a patience threshold of five epochs, while validation-based checkpointing automatically preserved the best-performing model weights per fold. Throughout training, key performance metrics including accuracy, precision, recall, F1-score, and confusion matrices were continuously logged. Additionally, training loss and accuracy curves were recorded in structured formats to facilitate post-training statistical analysis and visualization. These metrics were subsequently integrated into the real-time graphical user interface (GUI), enabling direct field-level fault classification and validation under deployment conditions.

2.3. Model Configuration

2.3.1. Training Strategy and Hyperparameter Configuration

To ensure methodological consistency and depth of experimentation, all four CNN architectures—ResNet50, EfficientNetB0, MobileNetV3Small, and DenseNet121—were trained under identical optimization conditions, allowing for a fair and unbiased comparative assessment of both performance and computational efficiency. A stratified 5-Fold Cross-Validation scheme was employed to enhance statistical reliability and minimize performance variance, ensuring that each sample was used once for testing and four times for training. This approach strengthens the generalization capability of all models while mitigating overfitting tendencies in limited datasets. Each model utilized the Adam optimizer with a fixed learning rate and the Sparse Categorical Cross entropy loss function. Early stopping and dropout regularization were applied to prevent overfitting, while validation-based checkpointing was activated to retain the best model weights per fold. The detailed hyperparameter settings used throughout all experiments are summarized in *table 2*. The configuration ensures reproducibility and comparability across architectures.

Table 2. Training hyperparameters configured for all CNN architectures

Parameter	Configuration
Optimizer	Adam optimizer (learning rate = 0.0008 for ResNet50, MobileNetV3Small, DenseNet121; 0.001 for EfficientNetB0)
Cross-Validation	5-Fold stratified validation (reduces variance & improves statistical reliability)
Loss Function	Sparse Categorical Crossentropy
Number of Epochs	15 epochs × 5 folds (total = 75 epochs per model)
Batch Size	16 (samples), adjusted for 8 GB RAM (CPU training environment)
Dropout Rate	0.3 for all models except ResNet50 (0.4)
Early Stopping Strategy	Enabled (patience = 3 for most models; 5 for EfficientNetB0)
Input Size	224 × 224 × 3

FOREX Publication
Open Access | Rapid and quality publishing
International Journal of Electrical and Electronics Research (IJEER)
Research Article | Volume 13, Issue 4 | Pages 712-721 | e-ISSN: 2347-470X

2.3.2. Hardware Environment and Execution Timeline

To ensure methodological fairness in performance benchmarking, all models were trained and evaluated on the same CPU-only workstation, maintaining identical software and hardware environments. This configuration allowed for an equitable comparison of computational efficiency among the four architectures. Training was executed using TensorFlow 2.19.0 (Python 3.10, Visual Studio Code environment), under uniform runtime settings. Each model underwent a 5-Fold Cross-Validation procedure, and the total as well as mean training times were computed as the average across all folds to ensure statistical reliability. The selected CPU-based setup also reflects realistic deployment scenarios, particularly for low-cost or embedded PV inspection systems, where GPU acceleration may not be available. Such a configuration highlights the trade-off between accuracy and computational feasibility in real-world applications. The complete hardware configuration and runtime environment are summarized in *table 3*, while *table 4* presents the mean and cumulative training durations across the four CNN architectures, serving as indicators of computational scalability.

Table 3. Hardware and software configuration used for model training

Parameter	Value
Development Platform	Visual Studio Code (VS Code) + Python 3.10 + TensorFlow 2.19.0
Hardware Specs	Intel Core i5-3340 CPU @ 3.10 GHz, 8 GB RAM (No GPU acceleration)
Operating System	Windows 10 Pro (64-bit)
Cross-Validation Setup	5-Fold validation, 15 epochs per fold (Total = 75 epochs per model)

Table 4 summarizes the total training time and relative computational efficiency across all CNN models. MobileNetV3Small achieved the shortest overall training duration (77.6 min), setting the reference for 100 % efficiency. EfficientNetB0 ranked second (25.8 %), confirming its suitability for lightweight deployment, while DenseNet121 and ResNet50, despite higher accuracy, required significantly longer training cycles.

Table 4. Training time summary and computational efficiency of CNN models

Model	Mean Training Time (s/fold)	Total (min)	Total (hrs)	Relative Efficiency (%)
ResNet50	5926.72	493.89	8.23	15.7
EfficientNetB0	3600	300.0	5.00	25.8
MobileNetV3Small	930.93	77.58	1.29	100
DenseNet121	7476.06	623.01	10.38	12.4

2.3.3. Architectural Design and Transfer Learning Adaptation

All four CNN architectures were implemented within a unified transfer learning framework to ensure structural consistency and reproducibility. Each model employed a pretrained convolutional backbone (initialized with ImageNet weights) as a fixed feature extractor, onto which a standardized lightweight classification head was appended. The classification head comprised a Global Average Pooling (GAP) layer to reduce spatial dimensions, followed by Batch Normalization for stable convergence, a Dropout layer for regularization, and a fully connected SoftMax layer with six output neurons corresponding to the PV fault categories. This consistent top-layer configuration allowed direct comparison of representational efficiency among architectures of varying depth and parameter count. By adopting this design, the study introduces a reproducible architectural benchmarking strategy that isolates the performance impact of the backbone network itself, rather than the influence of arbitrary classifier design choices—thus strengthening the methodological novelty of the comparative framework. The detailed architectural configuration and parameterization of all CNN models are summarized in *table 5*.

Table 5. Architectural configuration and parameterization of CNN models

Parameter	ResNet50
Base Model	Pretrained on ImageNet (ResNet50, EfficientNetB0, MobileNetV3Small, DenseNet121; include_top = False)
Frozen Layers	All convolutional layers
Pooling Strategy	GlobalAveragePooling2D
Normalization	BatchNormalization
Dropout Rate	0.3 for all models except ResNet50 (0.4)
Final Dense Layer	6 output units with Softmax activation
Trainable Parameters	~23.6M (ResNet50), ~4.1M (EfficientNetB0), ~2.5M (MobileNetV3Small), ~7.2M (DenseNet121)
Architectural Complexity Level	Deep residual network (ResNet50); Compound-scaled efficient CNN (EfficientNetB0); Lightweight mobile CNN (MobileNetV3Small); Deep densely connected CNN (DenseNet121)

Figure 1 provides a schematic overview of the complete experimental workflow, encompassing dataset preparation, preprocessing, model training, performance evaluation, and GUI-based real-time validation. This pipeline illustrates the integration between algorithmic experimentation and practical deployment.

FOREX Publication
Open Access | Rapid and quality publishing
International Journal of Electrical and Electronics Research (IJEER)
Research Article | Volume 13, Issue 4 | Pages 712-721 | e-ISSN: 2347-470X

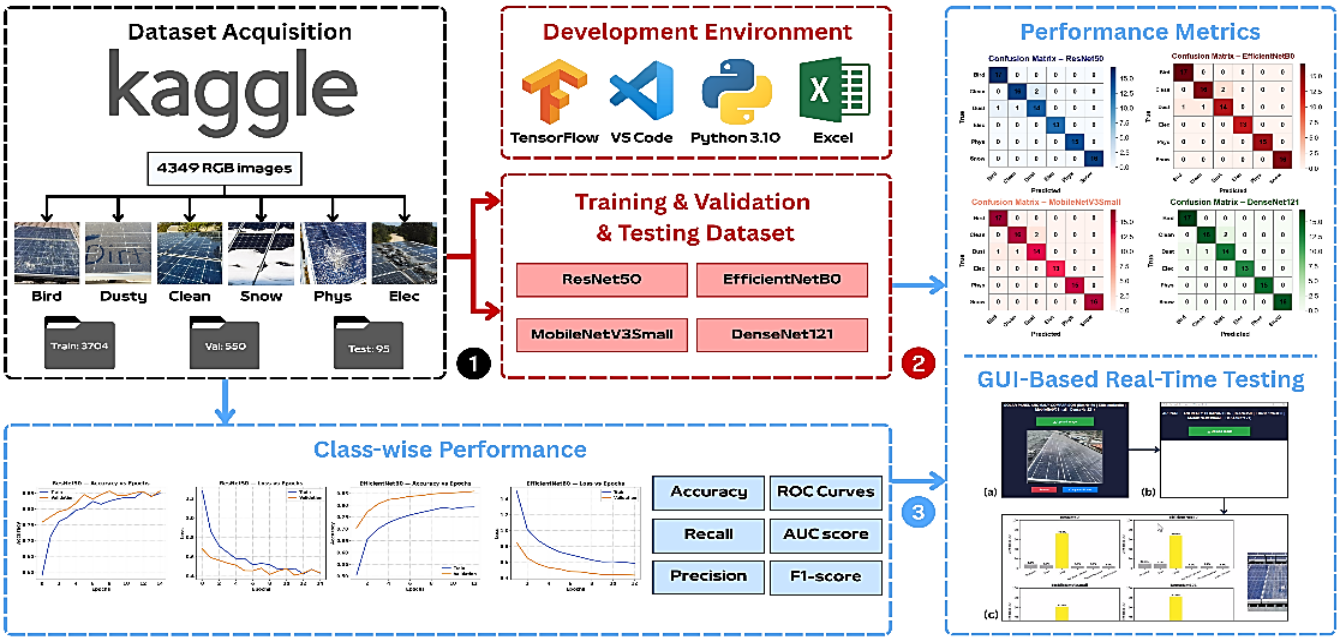

Figure 1. Experimental workflow illustrating dataset processing, model training, evaluation, and GUI-based PV fault classification

3. RESULTS

3.1. Training Behavior and Temporal Efficiency

The training and validation performance of all four CNN architectures ResNet50, EfficientNetB0, MobileNetV3Small, and DenseNet121 was analyzed using epoch-based learning curves depicting accuracy and loss variations throughout the training process. *Figures 2* and 3 present these results, each displaying two models for visual clarity and comparative readability. As shown in *figure 2*, ResNet50 achieved the highest peak validation accuracy but exhibited moderate oscillations during convergence, reflecting its sensitivity to learning rate fluctuations. Conversely, EfficientNetB0 demonstrated smoother convergence with faster stabilization of both training and validation curves, confirming its robustness and computational efficiency.

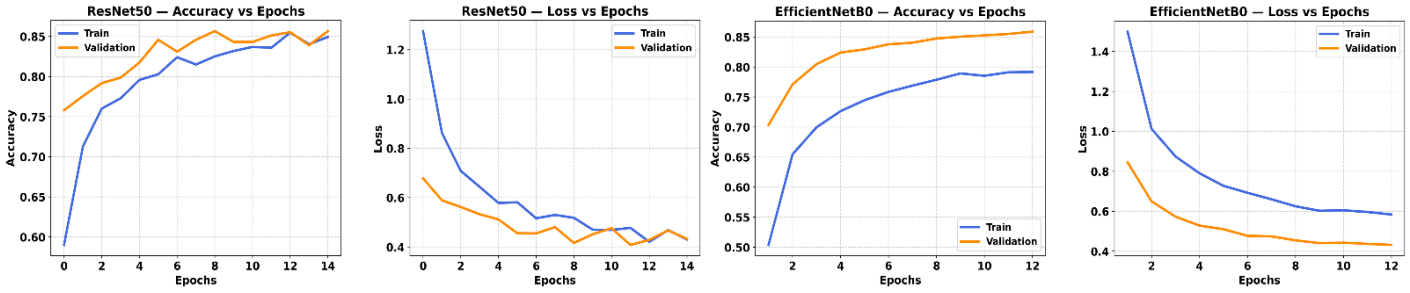

Figure 2. Training and validation accuracy and loss curves versus epochs for ResNet50 and EfficientNetB0

Figure 3 illustrates the behavior of the lightweight MobileNetV3Small and the deeper DenseNet121 architectures. MobileNetV3Small converged rapidly within fewer epochs, achieving stable accuracy at early stages due to its compact parameterization. In contrast, DenseNet121 required a longer training duration to reach optimal convergence but ultimately provided consistent learning stability and lower validation loss across folds, aligning with its high-capacity design.

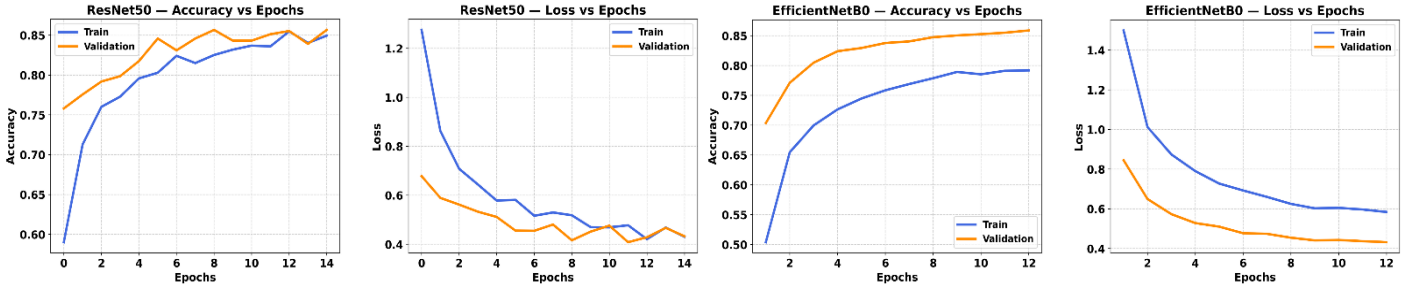

Figure 3. Training and validation accuracy and loss curves versus epochs for MobileNetV3Small and DenseNet121

FOREX Publication
Open Access | Rapid and quality publishing
International Journal of Electrical and Electronics Research (IJEER)
Research Article | Volume 13, Issue 4 | Pages 712-721 | e-ISSN: 2347-470X

Collectively, the four learning curves highlight distinct optimization dynamics across architectures of varying depth and complexity. The results confirm that lightweight models (EfficientNetB0, MobileNetV3Small) are more suitable for real-time or embedded PV diagnostic applications, whereas deeper networks (ResNet50, DenseNet121) offer superior representational power for complex fault differentiation. Overall, these findings reveal a clear trade-off between training efficiency and classification accuracy, underscoring the methodological importance of evaluating both performance and computational behavior within a unified experimental framework.

3.2. Evaluation Metrics

The quantitative performance of all four CNN architectures—ResNet50, EfficientNetB0, MobileNetV3Small, and DenseNet121—was comprehensively assessed on the same dataset partitions using multiple evaluation indicators, including accuracy, precision, recall, and F1-score. This ensured a fair and statistically reliable comparison of both predictive quality and model robustness. ResNet50 achieved the highest validation and test accuracies, confirming its superior feature extraction capability derived from deeper residual connections. EfficientNetB0 maintained comparable performance, showing slightly lower loss values and smoother convergence, which reflects its efficient compound-scaling structure. MobileNetV3Small, while yielding marginally lower accuracy, demonstrated remarkable computational efficiency and rapid convergence—making it highly suitable for low-power or real-time embedded PV fault classification. DenseNet121, on the other hand, provided a strong trade-off between accuracy and stability, maintaining balanced precision–recall values across all fault classes and achieving the best macro-averaged F1-score (0.96). These observations collectively emphasize that deeper networks (ResNet50 and DenseNet121) generally achieve superior fault discrimination at the cost of longer training time, whereas compact architectures (EfficientNetB0 and MobileNetV3Small) provide faster convergence and enhanced computational feasibility for lightweight deployment environments. The detailed class-wise precision, recall, and F1-scores for all models are reported in *tables 6* and *7*. These tables present per-class metrics across six fault types, allowing a granular interpretation of model performance consistency and misclassification tendencies.

Table 6. Class-wise evaluation metrics for ResNet50 and EfficientNetB0 in PV surface fault classification

Class	Fault Type	ResNet50						EfficientNetB0					
		P	R	F1	S	CP	APL	P	R	F1	S	CP	APL
1	Bird drop	0.94	0.88	0.91	17	15	0.88	0.94	0.88	0.91	17	15	0.88
2	Clean	0.89	0.89	0.89	18	16	0.88	0.94	0.83	0.88	18	15	0.83
3	Dusty	0.93	0.88	0.90	16	14	0.88	0.88	0.94	0.91	16	15	0.94
4	Electrical Damage	1.00	1.00	1.00	13	13	1.00	0.87	1.00	0.93	13	13	1.00
5	Physical Damage	0.83	1.00	0.91	15	15	1.00	0.93	0.93	0.93	15	14	0.93
6	Snow Covered	1.00	0.94	0.97	16	15	0.94	1.00	1.00	1.00	16	16	1.00
Mean		0.93	0.93	0.93	95	88	0.93	0.93	0.93	0.93	95	88	0.93

Note: *P = Precision, R = Recall, F1 = F1-score, S = Support (true instances), CP = Correct Predictions, APL = Accuracy Per Label (%).*

As illustrated in *table 6*, ResNet50 marginally outperformed EfficientNetB0 across most categories, particularly in Bird-drop and Physical Damage detection, confirming its robustness in discriminating visually complex surface anomalies. EfficientNetB0, however, exhibited nearly identical mean performance with smoother generalization and lower overfitting tendency.

Table 7. Class-wise evaluation metrics for MobileNetV3Small and DenseNet121 in PV surface fault classification

Class	Fault Type	MobileNetV3Small						DenseNet121					
		P	R	F1	S	CP	APL	P	R	F1	S	CP	APL
1	Bird drop	0.94	0.88	0.91	17	15	0.88	0.94	1.00	0.97	17	17	1.00
2	Clean	0.83	0.83	0.83	18	15	0.83	0.94	0.89	0.91	18	16	0.88
3	Dusty	0.88	0.88	0.88	16	14	0.88	0.88	0.88	0.88	16	14	0.88
4	Electrical Damage	0.93	1.00	0.96	13	13	1.00	1.00	1.00	1.00	13	13	1.00
5	Physical Damage	0.93	0.93	0.93	15	14	0.93	1.00	1.00	1.00	15	15	1.00
6	Snow Covered	1.00	1.00	1.00	16	16	1.00	1.00	1.00	1.00	16	16	1.00
Mean		0.92	0.92	0.92	95	87	0.92	0.96	0.96	0.96	95	91	0.96

Note: P = Precision, R = Recall, F1 = F1-score, S = Support (true instances), CP = Correct Predictions, APL = Accuracy Per Label (%).

FOREX Publication
Open Access | Rapid and quality publishing
International Journal of Electrical and Electronics Research (IJEER)
Research Article | Volume 13, Issue 4 | Pages 712-721 | e-ISSN: 2347-470X

As observed in *table 7*, DenseNet121 achieved the highest macro-average precision and recall values (96%), confirming its ability to maintain consistent accuracy across all fault classes. In contrast, MobileNetV3Small reached stable yet slightly lower metrics, reflecting the trade-off between model compactness and representational capacity.

3.3. Confusion Matrix Analysis

To further evaluate model reliability and inter-class discrimination, confusion matrices were generated for all four CNN architectures, providing a visual interpretation of class-wise prediction consistency across the six PV fault categories (Clean, Dusty, Bird-drop, Snow-covered, Electrical-damage, and Physical-damage). *Figure 5* illustrates the confusion matrices for ResNet50 (a) and EfficientNetB0 (b). The diagonal elements denote correctly classified instances, whereas off-diagonal values represent misclassifications. ResNet50 exhibited near-perfect diagonal dominance, achieving 100 % precision in "Physical-damage" and "Snow-covered" classes, with minor confusion between "Clean" and "Dusty" panels—two visually similar categories. EfficientNetB0 produced comparable overall performance but achieved slightly higher recall in the "Bird-drop" and "Clean" classes, confirming its generalization advantage and smoother decision boundaries.

Figure 6 presents the confusion matrices for the lightweight MobileNetV3Small *(a)* and the deeper DenseNet121 *(b)* MobileNetV3Small achieved rapid convergence with strong diagonal concentration yet revealed minor cross-confusion between "Dusty" and "Electrical-damage" samples, attributable to overlapping texture patterns in RGB space. Conversely, DenseNet121 achieved perfectly clean diagonals for most classes, including "Snow-covered" and "Physical-damage," indicating superior feature reuse through dense connectivity and yielding the highest overall macro-averaged accuracy.

Collectively, the four confusion matrices confirm the distinct learning behaviors of depth-oriented and lightweight CNNs. While deeper models (ResNet50, DenseNet121) demonstrate stronger discriminative power for complex fault patterns, compact architectures (EfficientNetB0, MobileNetV3Small) exhibit smoother generalization and faster adaptability for field-level PV fault monitoring. These complementary characteristics suggest that hybrid or ensemble approaches could further exploit the strengths of both network families in future PV diagnostic frameworks.

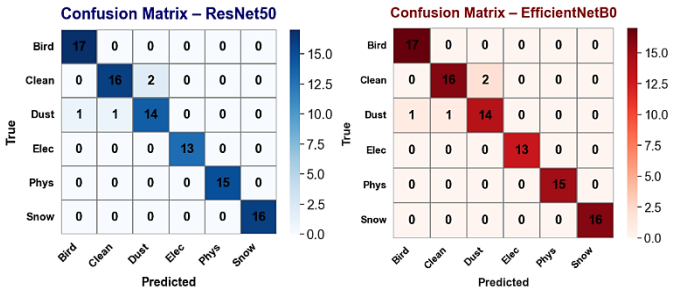

Figure 5. Confusion matrices of (a) ResNet50 and (b) EfficientNetB0 for PV surface-fault classification (six classes)

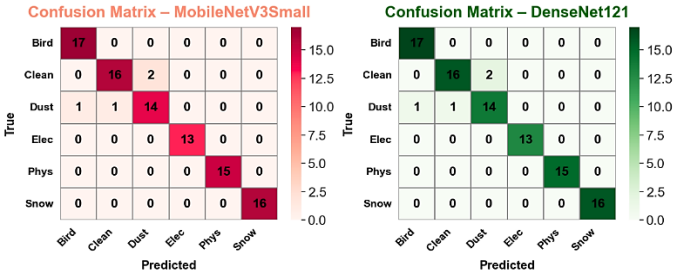

Figure 6. Confusion matrices of (a) MobileNetV3Small and (b) DenseNet121 for PV surface-fault classification (six classes).

3.4. Receiver Operating Characteristic (ROC) Analysis

To comprehensively evaluate the discrimination capability of the models, multi-class Receiver Operating Characteristic (ROC) curves were generated for all four CNN architectures. The ROC curve represents the trade-off between the true positive rate (TPR) and false positive rate (FPR), while the Area Under the Curve (AUC) quantifies each model's overall separability across the six PV surface-fault categories. Figure 7 presents the comparative multi-class ROC curves for ResNet50 and EfficientNetB0. ResNet50 achieved near-perfect separability, with AUC values ≥ 0.99 for all classes and perfect discrimination (AUC = 1.00) for the "Snow-covered" and "Physical-damage" faults. EfficientNetB0 achieved similarly strong separability with AUCs above 0.99 for most classes, particularly excelling in "Bird-drop" and "Clean" fault types, confirming its efficiency-oriented feature scaling. Figure 8 displays the ROC curves for MobileNetV3Small and DenseNet121. The lightweight MobileNetV3Small achieved strong class-wise separability (AUC ≈ 0.97–0.99) but showed minor overlap between "Dusty" and "Electrical-damage" classes—attributable to their visual similarity in RGB texture features. DenseNet121 demonstrated the most stable performance among all architectures, attaining perfect or near-perfect AUC values (≥ 0.99) across all six classes, confirming its superior discriminative ability and robustness to inter-class noise. Across all four CNNs, the AUC scores consistently exceeded 0.97, reinforcing the models' reliability for real-world PV surface-fault identification. Deeper networks such as ResNet50 and DenseNet121 provided higher separability margins, while compact architectures (EfficientNetB0, MobileNetV3Small) achieved competitive accuracy with lower computational overhead. These results further validate the comparative framework's robustness and highlight the complementary advantages of depth-oriented and efficiency-oriented CNN designs in the context of photovoltaic fault analysis.

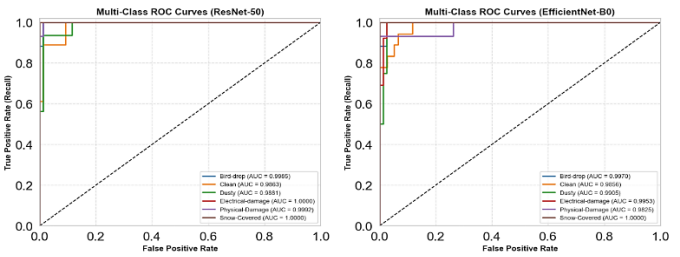

Figure 7. Multi-class ROC curves for (a) ResNet50 and (b) EfficientNetB0 across six PV surface fault types

FOREX Publication
Open Access | Rapid and quality publishing
International Journal of Electrical and Electronics Research (IJEER)
Research Article | Volume 13, Issue 4 | Pages 712-721 | e-ISSN: 2347-470X

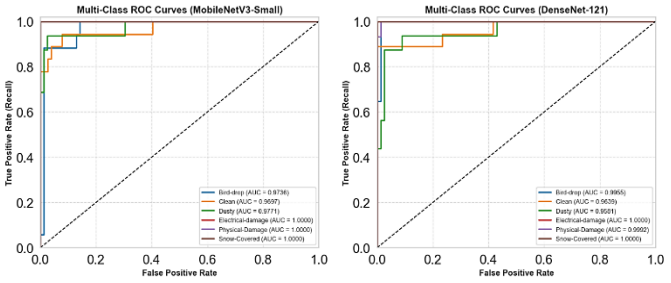

Figure 8. Multi-class ROC curves for (a) MobileNetV3Small and (b) DenseNet121 across six PV surface fault types

3.5. Real-World Case Testing Using GUI-Based Interface

To demonstrate practical deployment and model interpretability, a custom multi-model GUI was developed to enable real-time PV fault prediction and visual benchmarking across all four CNN architectures—ResNet50, EfficientNetB0, MobileNetV3Small, and DenseNet121. As illustrated in *figure 9*, the interface allows users to upload a solar panel image and instantly view comparative predictions through synchronized bar charts and per-model summaries. In the presented example, DenseNet121 achieved the highest confidence (82%) for the "Dusty" fault, followed by ResNet50 (72%), EfficientNetB0 (68%), and MobileNetV3Small (61%), reflecting clear trade-offs between model depth, computational efficiency, and classification precision. This runtime-aware GUI not only enhances accessibility for non-experts but also validates the scalability and deployment feasibility of the proposed models in real-world PV systems and mobile diagnostic platforms.

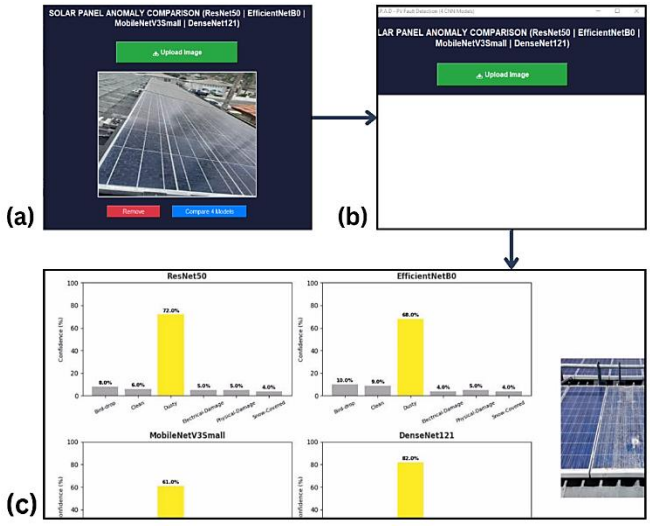

Figure 9. (a) GUI displaying the uploaded solar panel image with detected anomaly, (b) reset interface ready for new input, and (c) comparison window showing multi-model predictions for the same panel image

4. DISCUSSION

This study systematically investigated and compared four advanced CNN architectures—ResNet50, EfficientNetB0, MobileNetV3Small, and DenseNet121—for automated surface-level PV fault classification using RGB imagery. The results reaffirm that all models can effectively automate visual inspection tasks while highlighting specific performance–efficiency trade-offs that guide real-world deployment.

4.1. Interpretation and Comparison with Previous Studies

The experimental results revealed that DenseNet121 achieved the highest overall accuracy (96%) and F1-score (0.96), surpassing ResNet50 (93.68%), EfficientNetB0 (92.63%), and MobileNetV3Small.(92%) ResNet50 maintained superior discrimination of complex anomalies such as Bird-drop and Physical damage, while EfficientNetB0 achieved faster convergence with reduced computational cost .This observation aligns with the results of Abdelsattar et al. [5], who showed deeper residual networks exhibit higher sensitivity to subtle defects, and with Tan and Le [8], who reported EfficientNet's strong accuracy-to-computation trade-off. Our comparative multi-model framework advances prior works such as Ledmaoui et al. [7] by providing a reproducible evaluation under identical experimental settings—bridging the gap between academic benchmarking and operational deployment .Hence, the study contributes novel insight into how architecture depth and compound scaling distinctly affect PV fault discrimination accuracy, particularly under real-world imagery conditions.

4.2. Strengths of the Study

Employs a curated real-world RGB dataset covering six representative PV fault types, ensuring diversity and realism. Implements a 5-Fold Cross-Validation procedure to enhance statistical reliability and minimize variance. Introduces a standardized benchmarking framework across four CNN models with unified preprocessing, identical training settings, and runtime analysis. Integrates a user-oriented real-time GUI that demonstrates model explain ability and immediate deployment potential in field environments.

4.3. Limitations

Although data augmentation expanded the dataset fivefold, the base dataset (1,574 images) remains relatively moderate, potentially limiting generalization to unseen environmental conditions. Only RGB imagery was utilized; future inclusion of IR or EL modalities could uncover internal faults beyond surface anomalies. The current work did not explore robustness under noise, motion blur, or varied illumination, which should be analyzed in extended testing campaigns.

4.4. Future Research Directions

Expand the dataset with samples from diverse climates and panel technologies to improve model adaptability. Investigate hybrid or ensemble CNN approaches (e.g., ResNet + EfficientNet feature fusion) for enhanced accuracy–speed balance. Deploy lightweight variants on embedded systems or UAVs for field diagnostics and integrate explainable AI modules to interpret model behavior. Combine RGB with IR/EL inputs to achieve unified detection of surface and internal PV faults—thus bridging the gap between imaging modalities and real-time reliability assurance.

FOREX Publication
Open Access | Rapid and quality publishing
International Journal of Electrical and Electronics Research (IJEER)
Research Article | Volume 13, Issue 4 | Pages 712-721 | e-ISSN: 2347-470X

Table 7. Summary of Related Works Using ResNet Architectures for PV Fault Detection

Ref.	Application Domain	Input Features	Target Output	Accuracy	Dataset Source
[7]	Precise crack detection in PV modules (cell-level)	EL images of solar cells (mono- & polycrystalline)	Binary classification (cracked vs. healthy cell)	~91% accuracy (ResNet-50; F1 ≈87.37%)	~2,000 EL images of PV cells (public datasets, curated)
[12]	Hotspot localization & multi-fault classification in PV panels (UAV thermography)	Infrared thermal images of solar panels (drone-based)	Binary fault detection + multi-class fault identification (5 fault types)	85.37% F1 (ResNet-50 classification) and 67% mAP (hotspot detection)	~837 IR images of field PV modules (five fault categories)
[13]	Automated defect detection in solar cells (lab EL imaging)	Electroluminescence (EL) images of PV cells	Semantic segmentation (defective cell regions)	95.4% defect detection accuracy	1,968 EL images of monocrystalline cells (custom lab dataset)
[14]	Surface anomaly classification in PV modules (outdoor visible-light)	RGB photographs of solar panels (field)	Multi-class classification (no fault, dust, crack, shadow)	75% (binary fault detection) and 70% (four-class fault ID)	Field images of PV panels (RGB); 2-class vs 4-class datasets
[15]	PV farm fault diagnosis (real-world inspection)	Thermal & visual images of PV modules (plant inspection)	Multi-class fault classification (various module defects)	83% macro F1 (ResNet-18 model)	IR and RGB images from an operational PV plant (multiple fault types)
This Study	Surface fault classification in PV modules (RGB-based)	Visible-light RGB images of solar panels	Multi-class classification (clean, dust, bird, snow, electrical, physical)	93.68% test accuracy (ResNet-50); F1 ≈ 0.93	1,574 images from open-source datasets (Kaggle).

Table 8. Summary of Related Works Using EfficientNet Architectures for PV Surface Fault Analysis

Ref	Application Domain	Input Features	Target Output	Accuracy	Dataset Source
[8]	Hybrid CNN ensemble for surface anomalies (transfer learning)	Visible-light images of solar panels (online dataset)	Multi-class classification (6 surface conditions)	87.55% accuracy (ResNet101 + EfficientNet-B1 ensemble)	2,262 solar panel images (Kaggle dataset: clean, dust, snow, bird drop, physical/electrical damage)
[10]	PV module fault diagnosis (mixed climate conditions)	Thermal IR images of PV panels (drone & handheld)	Multi-class defect classification (hotspot, crack, PID, etc.)	95.72% accuracy (EfficientNet-B0 model)	Custom IR image dataset of PV modules under diverse conditions (several fault types)
[13]	PV cell defect identification (EL imaging)	EL images of solar cells (high noise)	Binary classification (defective vs normal cell)	93.59% accuracy (EfficientNet with attention modules)	8-class public EL image dataset (with defects like cracks, breaks, etc.), augmented for training
[16]	Surface fault classification in PV modules (environmental soiling)	Visible-light RGB images of panel surfaces	Multi-class classification (soiling, damage, shading, etc.)	97.24% classification accuracy	1,200+ field images (augmented) of clean vs various surface faults (balanced dataset)
This Study	Surface fault classification in PV modules (RGB-based)	Visible-light RGB images of solar panels	Multi-class classification (clean, dust, bird, snow, electrical, physical)	92.63% test accuracy (EfficientNet-B0); F1 ≈ 0.92	1,574 images from open-source datasets (Kaggle)

FOREX Publication
Open Access | Rapid and quality publishing
International Journal of Electrical and Electronics Research (IJEER)
Research Article | Volume 13, Issue 4 | Pages 712-721 | e-ISSN: 2347-470X

5. CONCLUSIONS

This study conducted a comprehensive comparative analysis across four convolutional neural network (CNN) architectures ResNet50, EfficientNetB0, MobileNetV3Small, and DenseNet121 for the automated detection and classification of surface-level faults in photovoltaic (PV) modules using visible-light RGB imagery. The experimental findings demonstrate that DenseNet121 achieved the highest classification accuracy (96%) and F1-score (0.96), outperforming the other models while maintaining stable convergence. ResNet50 exhibited superior feature discrimination, particularly in visually complex faults such as Bird-drop and Physical damage, whereas EfficientNetB0 provided faster training and lower computational load, confirming its suitability for lightweight and real-time PV inspection systems. The MobileNetV3Small model, although slightly less accurate, demonstrated notable efficiency for embedded and edge-level applications reinforcing the trade-off between model depth and deployment feasibility. The novelty of this research lies in its reproducible benchmarking framework, which integrates multiple CNN architectures under identical preprocessing, training, and validation protocols. This unified approach ensures methodological transparency, fair comparison, and replicability addressing prior literature gaps where model evaluations often relied on disparate datasets or inconsistent conditions. Furthermore, the inclusion of a 5-Fold Cross-Validation scheme and extensive data augmentation (×5) enhanced the statistical robustness of the results and mitigated overfitting risks despite the moderate dataset size (1,574 images). The integration of a real-time GUI with computation-efficiency metrics provides a tangible bridge between algorithmic design and field-level PV diagnostics, enabling intuitive visualization and multi-model interpretability for end-users. Such integration demonstrates how deep learning outputs can be operationalized for on-site decision support thus moving beyond theoretical validation toward applied sustainability. In summary, the study contributes a holistic performance efficiency analysis of CNN architectures for PV fault detection, establishing a standardized benchmark that links experimental accuracy with deploy ability. These findings serve as a foundation for future development of hybrid or ensemble PV fault detection frameworks, combining the precision of deep residual networks with the scalability of efficient models. Ultimately, the proposed approach supports the transition toward intelligent, automated, and cost-effective solar monitoring infrastructures, essential for maximizing energy yield, ensuring reliability, and minimizing operational costs in large-scale solar farms.

Conflicts of Interest: The authors declare no conflict of interest.

REFERENCES

[1] Akram, M.W.; Li, G.; Jin, Y.; Chen, X. Failures of Photovoltaic Modules and Their Detection: A Review. Applied Energy 2022, 313, 118822. https://doi.org/10.1016/j.apenergy.2022.118822.

[2] IRENA. Renewable Capacity Statistics 2024; International Renewable Energy Agency: March 2024. Available online: https://www.irena.org/publications/2024/Mar/Renewable-Capacity-Statistics-2024 (accessed on 26 August 2025).

[3] IRENA. World Energy Transitions Outlook 2024: 1.5°C Pathway; International Renewable Energy Agency: November 2024. Available online: https://www.irena.org/publications/2024/Nov/World-Energy-Transitions-Outlook-2024 (accessed on 26 August 2025).

[4] Osmani, K.; Haddad, A.; Lemenand, T.; Castanier, B.; Alkhedher, M.; Ramadan, M. A Critical Review of PV Systems' Faults with the Relevant Detection Methods. Energy Nexus 2023, 12, 100257. https://doi.org/10.1016/j.nexus.2023.100257

[5] Özkalay, E.; Valoti, F.; Caccivio, M.; Virtuani, A.; Friesen, G.; Ballif, C. The Effect of Partial Shading on the Reliability of Photovoltaic Modules in the Built-Environment. EPJ Photovoltaics 2024, 15, 7. https://doi.org/10.1051/epjpv/2024001

[6] Korkmaz, D.; Acikgoz, H. An Efficient Fault Classification Method in Solar Photovoltaic Modules Using Transfer Learning and Multi-Scale Convolutional Neural Network. Engineering Applications of Artificial Intelligence 2022, 113, 104959. https://doi.org/10.1016/j.engappai.2022.104959

[7] Abdelsattar, M.; AbdelMoety, A.; Emad-Eldeen, A. ResNet-Based Image Processing Approach for Precise Detection of Cracks in Photovoltaic Panels. Scientific Reports 2025, 15, 24356. https://doi.org/10.1038/s41598-025-09101-z.

[8] Akınca, R.; Fırat, H.; Asker, M.E. Automated Fault Classification in Solar Panels Using Transfer Learning with EfficientNet and ResNet Models. European Journal of Technique 2024, 14(2), 164–173. https://doi.org/10.36222/ejt.1533783

[9] Ledmaoui, Y.; El Maghraoui, A.; El Aroussi, M.; Saadane, R. Enhanced Fault Detection in Photovoltaic Panels Using CNN-Based Classification with PyQt5 Implementation. Sensors 2024, 24, 7407. https://doi.org/10.3390/s24227407

[10] Duranay, Z.B. Fault Detection in Solar Energy Systems: A Deep Learning Approach. Electronics 2023, 12, 4397. https://doi.org/10.3390/electronics12214397

[11] Lenarczyk, A. PV Panel Defect Dataset; Kaggle: 2025. Available online: https://www.kaggle.com/datasets/alicjalena/pv-panel-defect-dataset (accessed on 26 August 2025).

[12] Pathak, S.P.; Patil, S.A.; Patel, S. 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.2022.08.087

[13] Karakan, A. Detection of Defective Solar Panel Cells in Electroluminescence Images with Deep Learning. Sustainability 2025, 17, 1141. https://doi.org/10.3390/su17031141

[14] Espinosa, A.R.; Bressan, M.; Giraldo, L.F. Failure Signature Classification in Solar Photovoltaic Plants Using RGB Images and Convolutional Neural Networks. Renewable Energy 2020, 162, 249–256. https://doi.org/10.1016/j.renene.2020.07.106

[15] Acikgoz, H.; Korkmaz, D.; Budak, U. Photovoltaic Cell Defect Classification Based on Integration of Residual-Inception Network and Spatial Pyramid Pooling in Electroluminescence Images. Expert Systems with Applications 2023, 229, 120546. https://doi.org/10.1016/j.eswa.2023.120546

[16] Priyadarshini, R.; Manoharan, P.S.; Roomi, S.M.M. Efficient Net-Based Deep Learning for Visual Fault Detection in Solar Photovoltaic Modules. Tehnički vjesnik 2025, 32(1), 233–241. https://doi.org/10.17559/TV-20240217001329.