Black Sea Journal of Engineering and Science
doi: 10.34248/bsengineering.1927407

Open Access Journal
e-ISSN: 2619 – 8991

Research Article
Volume 9 - Issue 3: 1444-1453 / May 2026

A LIGHTWEIGHT CONVNEXT-BASED ARCHITECTURE WITH HYBRID FEATURE AGGREGATION FOR MULTICLASS PHOTOVOLTAIC PANEL DEFECT CLASSIFICATION

Berkay EREN[1]*

[1]*İskenderun Technical University, Department of Mechatronics Engineering, 31200, İskenderun, Hatay, Türkiye*

Abstract: Photovoltaic panel defect classification is essential for reliable solar energy systems, particularly in industrial inspection. This study proposes a lightweight ConvNeXt-based architecture with a hybrid feature aggregation mechanism that combines Global Average Pooling and Global Max Pooling, followed by a compact classification head for multiclass defect classification. Unlike standard pooling strategies, the proposed aggregation enhances feature representation by capturing both global distribution and salient responses. The proposed model is evaluated through ablation studies and compared with baseline configurations under a unified training pipeline. Experimental results show that the proposed architecture achieves a weighted F1-score of approximately 0.96, outperforming baseline ConvNeXt variants and demonstrating improved feature representation capability. These findings indicate that effective yet simple feature aggregation strategies can significantly enhance classification performance while retaining reasonable inference speed for industrial applications, making the model suitable for real-world PV inspection applications.

Keywords: Photovoltaic panel defect classification, ConvNeXt, Transfer learning, Feature aggregation, Deep learning

***Corresponding author:** İskenderun Technical University, Department of Mechatronics Engineering, 31200, İskenderun, Hatay, Türkiye
E mail: berkay.eren@iste.edu.tr (B. EREN)
Berkay EREN https://orcid.org/0000-0002-7019-124X

Received: April 10, 2026
Accepted: May 13, 2026
Published: May 15, 2026

Cite as: Eren, B. (2026). A lightweight convnext-based architecture with hybrid feature aggregation for multiclass photovoltaic panel defect classification. *Black Sea Journal of Engineering and Science*, 9(3), 1444-1453.

1. Introduction

Photovoltaic (PV) systems have become a fundamental component of sustainable energy production due to their scalability and environmental benefits. However, the efficiency and reliability of PV systems are highly dependent on the early detection of surface defects such as cracks, hotspots, and contamination. Undetected defects may lead to significant performance degradation and long-term system failure (Pathak and Patil, 2023; Masita et al., 2025). Therefore, accurate and automated defect classification methods are essential, particularly for large-scale industrial inspection scenarios (Hijjawi et al., 2023).

In traditional PV inspection processes, manual visual assessment and rule-based image processing techniques have been widely used. However, these approaches are often time-consuming and sensitive to variations in illumination, noise, and environmental conditions (Khanam et al., 2024). In recent years, deep learning-based methods, particularly convolutional neural networks (CNNs), have demonstrated strong performance in image classification tasks (Cao et al., 2023; Eren, 2026). Transfer learning models such as DenseNet, ResNet, Xception, and EfficientNet have been extensively applied in PV defect detection due to their ability to leverage pre-trained feature representations (Ahmed et al., 2021; Thakfan and Bin Salamah, 2024).

Despite their success, most existing approaches rely on standard feature extraction and pooling strategies, which may limit the model's ability to capture both global contextual information and localized discriminative features . In particular, Global Average Pooling (GAP) is commonly used to reduce feature dimensions, but it may suppress high-response features. Conversely, Global Max Pooling (GMP) emphasizes salient features while ignoring global distribution characteristics (Sun et al., 2025; Zhang et al., 2025). This limitation becomes more critical in multiclass classification problems where subtle differences between defect categories must be distinguished.

Recent studies have explored various improvements in PV defect classification, including lightweight architectures, attention mechanisms, and optimization-based approaches (Jiang and Zhao, 2022; Xiao et al., 2023; Huang et al., 2025). However, many of these methods either increase model complexity or do not explicitly address the limitations of feature aggregation. Moreover, comparative studies often lack a unified training framework, making fair evaluation between models difficult (Ejiyi et al., 2026). Ibn Malek and Imtiaz, (2025) investigated transfer learning combined with electroluminescence (EL) image processing,

demonstrating that deep neural networks significantly improve defect identification accuracy in PV modules under varying operating conditions. Similarly, Munawer Al-Otum, (2024) developed a CNN-based framework for classifying anomalies in EL images, showing that deep architectures can effectively distinguish multiple defect types when supported by appropriate dataset preprocessing and augmentation strategies.

In addition, Shaban, (2024) proposed an artificial intelligence-based method for PV defect detection, integrating CNN-based feature extraction method to enhance classification performance. Tella et al., (2025) explored ensemble deep learning techniques for PV cell defect classification and demonstrated that combining multiple models improves robustness and generalization across different defect categories. Likewise, Akram and Bai, (2025) introduced a deep learning framework incorporating image-to-image generation and transfer learning to improve defect detection performance, particularly in cases of limited and imbalanced datasets.

Feature representation and aggregation remain critical challenges in CNN-based PV inspection systems. Demir, (2025) emphasized that transfer learning-based CNN models can effectively capture defect features in EL images; however, their performance strongly depends on how spatial features are aggregated. Similarly, Ledmaoui et al., (2024) employed a VGG16-based CNN model with data augmentation and achieved high classification accuracy for multiple PV fault types. Despite these advancements, existing methods often rely on standard feature aggregation techniques and may not fully exploit the complementary characteristics of global and local feature representations. Therefore, there remains a need for lightweight and efficient architectures that can effectively integrate diverse feature aggregation strategies to improve classification performance in multiclass PV defect detection tasks.

In this study, a lightweight ConvNeXt-based architecture is proposed for multiclass PV panel defect classification. The proposed model introduces a hybrid feature aggregation strategy by combining GAP and GMP, enabling the extraction of complementary feature information. In addition, a compact bottleneck classification head is designed to improve discriminative capability while maintaining computational efficiency. The proposed approach is evaluated through ablation studies and compared with widely used transfer learning models under a unified experimental setup.

The performance of the proposed model is assessed using multiple evaluation metrics, including accuracy, precision, recall, and F1-score. Experimental results demonstrate that the proposed architecture achieves superior performance compared to baseline configurations, highlighting the effectiveness of hybrid feature aggregation for improving classification accuracy and robustness in PV inspection applications.

Based on the above discussion, there is a clear need for lightweight yet effective architectures that can improve feature representation without increasing model complexity.

The main contributions of this study can be summarized as follows:

- Instead of introducing a more complex architecture, this study focuses on improving feature aggregation within a ConvNeXt-based framework, showing that performance gains can be achieved without introducing substantially more complex architectural components.
- A hybrid feature aggregation strategy is proposed by combining GAP and GMP, enabling the model to simultaneously capture global feature distribution and highly activated discriminative responses.
- It is experimentally demonstrated that naive combination of GAP and GMP degrades performance, while the proposed structured integration with a lightweight bottleneck head significantly improves classification accuracy.
- A lightweight refinement stage (Batch Normalization + GELU + Dropout) is introduced after feature aggregation, showing that controlled feature transformation is critical for effective fusion of dual pooling representations.
- The proposed design achieves superior performance compared to both conventional CNNs and the baseline ConvNeXt model, indicating that feature aggregation plays a more decisive role than backbone selection in this task.

2. Materials and Methods

2.1. Proposed Framework Overview

In this study, a deep learning-based classification framework is proposed for PV panel defect detection. The model is designed as a lightweight yet effective architecture that combines a convolutional backbone with a compact classification head. The overall pipeline consists of four main stages: data augmentation, feature extraction, feature aggregation, and classification.

The overall architecture of the proposed model is illustrated in Figure 1. Initially, input images are processed through a lightweight augmentation module to improve generalization capability. These augmented images are then passed to a pretrained ConvNeXt-Tiny backbone, which serves as the feature extraction module and generates deep hierarchical feature representations.

To obtain a more informative global representation, a dual pooling strategy is employed. Specifically, GAP is used to capture overall feature distribution, while GMP highlights the most discriminative activations. The outputs of these two operations are concatenated to form a fused feature vector that encodes both global context and salient local information.

The fused features are subsequently processed by a lightweight classification head consisting of batch normalization, a fully connected layer with GELU activation, and dropout regularization. This design

enables effective feature refinement while maintaining low computational complexity. Finally, class probabilities are generated using a softmax layer for multi-class PV defect classification. Unlike conventional classification approaches that rely on a single global descriptor, the proposed framework integrates complementary feature representations through dual pooling. In addition, a unified architecture design is adopted to ensure fair and consistent comparison across different backbone models.

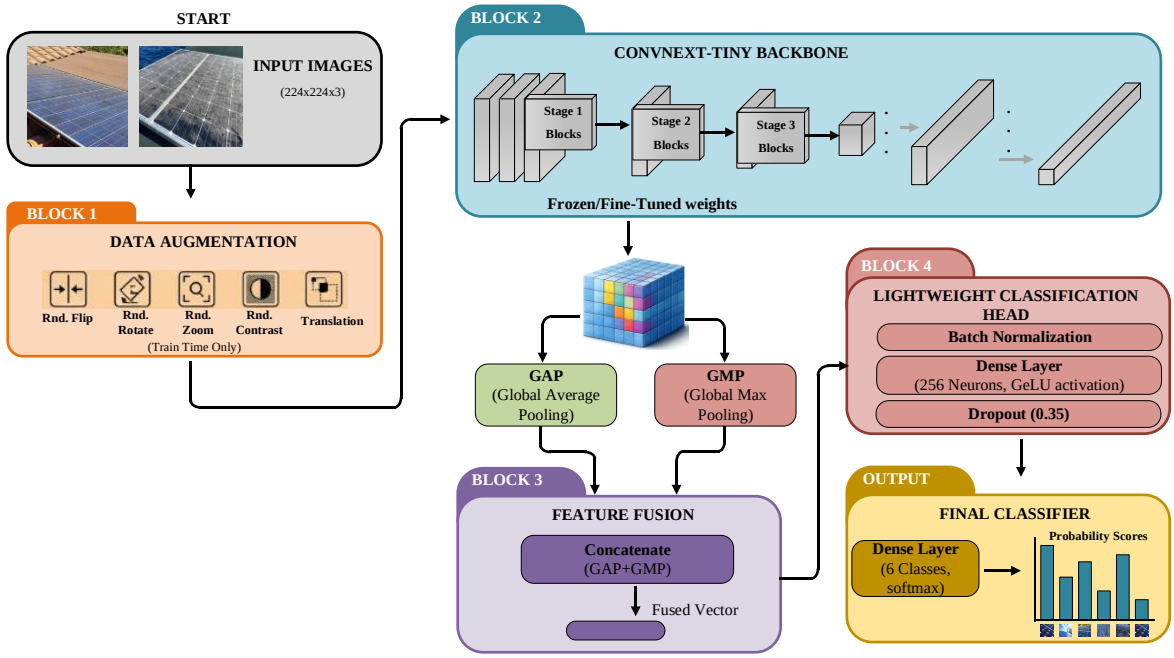

Figure 1. Overall architecture of the proposed PV defect classification model based on ConvNeXt-Tiny backbone and dual pooling feature aggregation.

2.2. Feature Extraction Backbone

In the proposed model, a convolutional neural network is used as a feature extraction backbone to obtain deep representations of the input image.

Specifically, the ConvNeXt-Tiny architecture is employed due to its strong representation capability and efficient design.

Given an input image x_i, the backbone extracts hierarchical features (equation 1):

$$z_i = \Phi(x_i) \tag{1}$$

where

- $\Phi(\cdot)$denotes the backbone network
- $z_i \in \mathbb{R}^{h \times w \times d}$ represents the extracted feature map

The backbone is initialized with pretrained ImageNet weights to leverage prior visual knowledge and improve generalization performance.

During training, the backbone is initially frozen to allow the classification head to adapt to the target task. In later stages, partial fine-tuning is applied to improve feature specialization.

2.3. Dual Feature Aggregation

To obtain a robust global representation, the proposed model employs a dual pooling mechanism combining GAP and GMP. The GAP operation is defined as (equation 2):

$$g_{avg} = \frac{1}{h \cdot w} \sum_{p=1}^{h} \sum_{q=1}^{w} z_i(p,q) \tag{2}$$

which captures the overall distribution of feature responses.

The GMP operation is defined as (equation 3):

$$g_{max} = \max_{p,q} z_i(p,q) \tag{3}$$

which emphasizes the most salient activations.

These two feature descriptors are concatenated (equation 4):

$$g_i = [g_{avg} || g_{max}] \tag{4}$$

where $g_i \in \mathbb{R}^{2d}$.

This fusion allows the model to simultaneously capture both global contextual information and prominent discriminative features, improving classification performance.

2.4. Lightweight Bottleneck Classification Head

Following feature aggregation, a lightweight bottleneck structure is applied to refine the feature representation.

First, batch normalization is applied (equation 5):

$$\tilde{g}_i = BN(g_i) \tag{5}$$

Then, a nonlinear transformation is performed using a fully connected layer (equation 6):

$$h_i = \sigma(W_1 \tilde{g}_i) \tag{6}$$

where $\sigma(\cdot)$ denotes the GELU activation function.

To reduce overfitting, dropout is applied (equation 7):

$$h_i' = Dropout(h_i) \quad (7)$$

Finally, the output logits are computed (equation 8):

$$o_i = W_2 h_i' + b \quad (8)$$

and class probabilities are obtained via (equation 9):

$$\hat{y}_i = softmax(o_i) \quad (9)$$

This lightweight design significantly reduces the number of parameters while maintaining strong discriminative capability.

2.5. Training Strategy

The training process of the proposed model is carried out in two stages to ensure both stable convergence and effective feature adaptation.

In the first stage, the backbone network is kept frozen, and only the classification head is trained. This allows the model to learn task-specific decision boundaries without altering the pretrained feature representations. Since the backbone is initialized with ImageNet weights, this stage helps preserve the general visual features while adapting the classifier to PV defect categories.

In the second stage, a fine-tuning process is applied by unfreezing the upper layers of the backbone network. Specifically, the last 20% of the backbone layers are set to be trainable, while the remaining layers remain frozen. This strategy enables the model to refine high-level feature representations without causing overfitting or instability during training. Also, a fixed random seed (42) was used throughout all experiments to ensure reproducibility.

The optimization is performed using the Adam optimizer. A relatively higher learning rate is used during the initial training stage to accelerate convergence of the classification head, while a lower learning rate is adopted during fine-tuning to ensure stable updates of the backbone parameters.

To further improve training stability and prevent overfitting, early stopping and learning rate scheduling techniques are employed based on validation performance.

2.6. Dataset and Experimental Setup

The dataset used in this study consists of 1,569 PV panel images belonging to six defect categories: Bird-drop, Clean, Dusty, Electrical-damage, Physical-Damage, and Snow-Covered. The dataset used in this study is publicly available and has been widely used in photovoltaic defect classification researches (Diaconu, 2026; Laouamer et al., 2025). The dataset is divided into training, validation, and test subsets, containing 929, 308, and 332 images, respectively. The class-wise distribution is as follows: Bird-drop (296), Clean (286), Dusty (275), Electrical-damage (225), Physical-Damage (225), and Snow-Covered (262). This distribution indicates a relatively balanced multi-class classification setting, although slight variations between class frequencies are present.

All images are resized to 224 × 224 pixels before being fed into the network. The dataset is organized using a directory-based class structure, and class labels are automatically assigned during loading through the TensorFlow data pipeline.

To improve model generalization, an online data augmentation strategy is applied only to the training images. This augmentation pipeline includes horizontal flipping, slight rotation, random zooming, contrast adjustment, and small spatial translations. These transformations are intentionally kept limited in order to preserve the intrinsic visual characteristics of PV defects while still increasing training diversity. It should be noted that augmentation is performed dynamically during training and does not alter the original number of images in the dataset.

The proposed framework is implemented using TensorFlow/Keras. A ConvNeXt-Tiny backbone pretrained on ImageNet is employed as the feature extractor. The training process is performed in two stages. In the first stage, the backbone is frozen and only the classification head is optimized for 15 epochs. In the second stage, the upper 20% of the backbone layers are unfrozen for fine-tuning, and training is continued for an additional 10 epochs.

The model is trained using the Adam optimizer with an initial learning rate of 1×10^{-3}. The batch size is set to 8, and the loss function is defined as sparse categorical cross-entropy. To improve convergence behavior and reduce overfitting risk, several callback strategies are employed, including early stopping, model checkpointing, and ReduceLROnPlateau based on validation performance. All experiments were conducted on a system equipped with an Intel Core i5-9300H CPU and an NVIDIA GeForce GTX 1650 GPU.

The final performance is evaluated on the independent test set using accuracy, precision, recall, and F1-score. In addition, confusion matrix analysis is conducted to examine class-wise prediction behavior in more detail.

3. Results

The performance of the proposed ConvNeXt-based architecture is evaluated on the test dataset using multiple classification metrics. To provide a structured and comprehensive analysis, the results are presented under three main components: overall performance evaluation, comparison with state-of-the-art models, and ablation study.

3.1. Overall Performance Evaluation

The performance of the proposed model is analyzed in detail through class-wise evaluation metrics, including precision, recall, F1-score, and support values, as presented in Table 1. Here, the support value represents the number of true samples belonging to each class in the test dataset.

Table 1. Class-wise performance of the proposed model

Class	Precision	Recall	F1-score	Support
Bird-drop	0.9615	0.9259	0.9434	54
Clean	0.963	0.9123	0.9369	57
Dusty	0.8644	0.9623	0.9107	53
Electrical-damage	0.9831	0.9831	0.9831	59
Physical-Damage	0.9808	0.9444	0.9623	54
Snow-Covered	0.9821	1	0.991	55
Overall (Weighted Avg.)	0.9566	0.9548	0.9551	332

The proposed model demonstrates consistently high performance across all PV defect categories, achieving an overall weighted F1-score of 0.9551, which indicates strong and balanced classification capability.

Among all classes, the Electrical-damage category achieves near-perfect performance, with precision, recall, and F1-score values of 0.9831, indicating that the model effectively captures distinctive structural defect patterns. Similarly, the Snow-Covered class achieves the highest performance with an F1-score of 0.9910 and a perfect recall of 1.000, suggesting that this class is highly distinguishable due to its prominent visual characteristics.

The Physical-Damage class also exhibits strong performance, with a precision of 0.9808 and an F1-score of 0.9623, demonstrating the model's capability to detect structural anomalies with high reliability. For contamination-related classes, Bird-drop and Clean achieve F1-scores of 0.9434 and 0.9369, respectively. These results indicate that the model successfully differentiates between clean surfaces and contamination patterns under varying conditions.

The lowest performance is observed in the Dusty class, where the model achieves an F1-score of 0.9107. Although the recall is relatively high (0.9623), the lower precision (0.8644) indicates that the model occasionally confuses dusty surfaces with other contamination-related categories. This behavior can be attributed to the visual similarity between certain defect types, particularly under challenging lighting and noise conditions. Overall, the class-wise evaluation confirms that the proposed model maintains a well-balanced performance across all categories without significant bias toward any specific class. The combination of high precision and recall values across diverse defect types demonstrates the effectiveness of the proposed hybrid feature aggregation mechanism in capturing both global contextual information and fine-grained visual features.

The training behavior of the proposed model is illustrated in Figure 2. As shown in the figure, both training and validation accuracy curves exhibit a stable increasing trend, while the loss values decrease consistently over epochs. The absence of significant divergence between training and validation curves indicates that the model does not suffer from severe overfitting and generalizes well to unseen data. In addition, although a temporary fluctuation is observed during the fine-tuning stage, this behavior is mainly associated with the partial unfreezing of the backbone layers and the subsequent adaptation of pretrained features to the target dataset. Nevertheless, the overall convergence behavior confirms the effectiveness of the two-stage training strategy. A transition point between the initial training phase and the fine-tuning stage can also be observed in the training curves, where a slight change in the learning dynamics occurs. This transition reflects the moment when the backbone layers are partially unfrozen, allowing the model to refine higher-level feature representations and further improve classification performance.

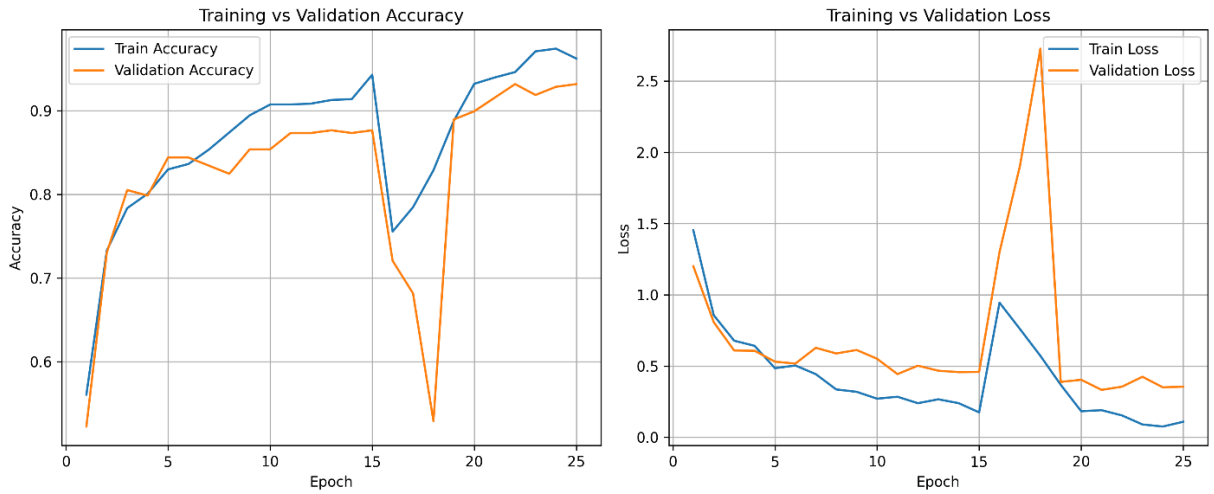

Figure 2. Training and validation accuracy and loss curves of the proposed ConvNeXt-based model.

3.2. Comparison with State-of-the-Art Models

To evaluate the effectiveness of the proposed architecture, a comparative analysis is conducted against widely used deep learning models, including DenseNet121, ResNet50, Xception, ConvNeXt-Tiny and EfficientNetB0. The results are summarized in Table 2.

As shown in Table 2, the proposed model achieves the best overall performance with an accuracy of 0.9548 and a weighted F1-score of 0.9551. In contrast, DenseNet121 and Xception exhibit significantly lower performance, with accuracies of 0.8795 and 0.8916, respectively. These results indicate that conventional architectures may struggle to capture fine-grained defect characteristics in PV images.

EfficientNetB0 demonstrates improved performance compared to DenseNet121 and Xception, achieving an accuracy of 0.9006. However, it still falls behind the proposed model, suggesting that standard feature extraction alone is not sufficient for optimal defect classification. ResNet50 provides competitive results with an accuracy of 0.9518 and a weighted F1-score of 0.9525. Nevertheless, the proposed model slightly outperforms ResNet50 across all evaluation metrics.

Table 2. Performance comparison with state-of-the-art models

Model	Test Loss	Acc.	Macro F1	W. F1	FPS
DenseNet121	0.3996	0.8795	0.8799	0.8795	36.95
ResNet50	0.2125	0.9518	0.9519	0.9525	60.71
Xception	0.3305	0.8916	0.8909	0.8921	57.94
EfficientNetB0	0.2833	0.9006	0.9010	0.9015	45.95
ConvNeXt-Tiny	0.1665	0.9457	0.9453	0.9457	42.47
Proposed Model	0.2146	0.9548	0.9546	0.9551	38.10

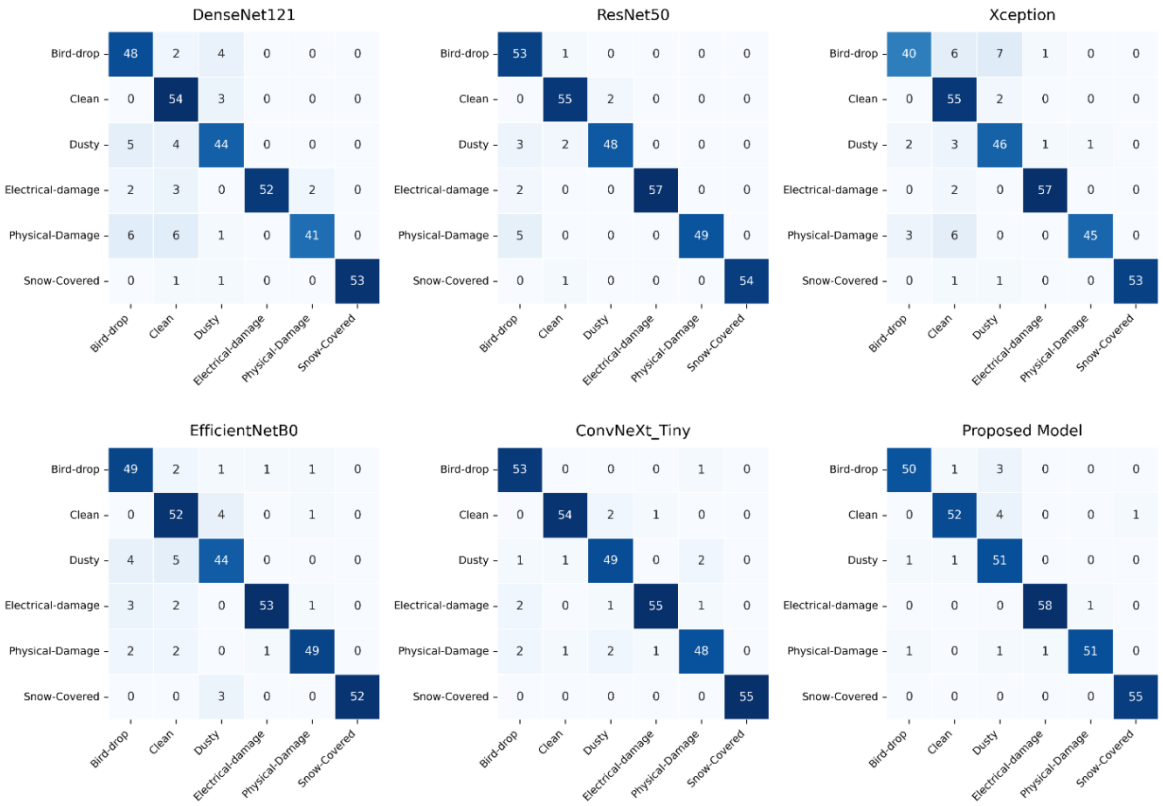

Figure 3. Confusion matrices of six deep learning models evaluated on the PV defect classification task.

ConvNeXt-Tiny, as a more recent and advanced architecture, achieves strong performance with an accuracy of 0.9457 and a weighted F1-score of 0.9457. While it surpasses several conventional models, it still remains slightly below the proposed model. This result highlights that modern backbone architectures alone are not sufficient, and that the proposed feature aggregation strategy plays a critical role in improving classification performance. This improvement, although moderate, is significant given that the proposed approach focuses on enhancing feature aggregation rather than increasing model complexity.

Inference speed analysis shows clear differences across models. ResNet50 achieves the highest speed (60.71 FPS), followed by Xception (57.94 FPS). In contrast, ConvNeXt-Tiny and the proposed model operate at 42.47

FPS and 38.10 FPS, respectively. Although the proposed model is slightly slower than some baseline architectures, it achieves the best accuracy and weighted F1-score, indicating a favorable trade-off between classification performance and inference speed for practical applications.

Figure 3 presents the confusion matrices of the evaluated models, providing a detailed view of class-wise prediction behavior. The proposed model demonstrates a strong diagonal dominance, indicating that most samples are correctly classified across all defect categories.

In particular, the Electrical-damage and Snow-Covered classes show near-perfect classification performance, with very limited misclassification. This suggests that these defect types exhibit distinctive visual characteristics that are effectively captured by the model. Similarly, the Physical-Damage class maintains high prediction accuracy, with only minor confusion observed with visually similar categories. However, a small degree of misclassification is observed in the Dusty class, where some samples are incorrectly predicted as Bird-drop or Clean. This behavior can be attributed to the visual similarity between contamination-related defects, especially under varying illumination conditions.

Compared to baseline models, the proposed architecture exhibits a more concentrated distribution along the diagonal and fewer off-diagonal errors, confirming its superior discriminative capability. These results further validate that the proposed hybrid feature aggregation strategy enhances class separability and reduces inter-class confusion.

To further illustrate the model behavior, representative examples of correct and incorrect predictions are presented in Figure 4. As shown in the figure, the model successfully identifies most defect categories with high confidence. Misclassifications mainly occur between visually similar classes, such as Dusty and Bird-drop, which share similar texture patterns under certain illumination conditions. This observation is consistent with the confusion matrix analysis and highlights the inherent difficulty of distinguishing subtle defect variations. This further indicates that misclassification mainly occurs in borderline cases rather than systematic errors.

Overall, these findings confirm that the proposed hybrid feature aggregation mechanism effectively improves feature representation by combining global contextual information and salient local features, leading to superior classification performance.

Figure 4. Representative examples of correctly classified (top row) and misclassified (bottom row) samples from the test dataset.

3.3. Ablation Study

To analyze the contribution of different architectural components, an ablation study is conducted under a unified experimental setup. The results of different feature aggregation and refinement configurations are summarized in Table 3.

To analyze the contribution of the proposed architectural components, an ablation study is conducted under a unified experimental setup. The quantitative results are presented in Table 3, where different configurations of the model are systematically evaluated.

The baseline ConvNeXt model using only GAP achieves an accuracy of 0.9247 and a weighted F1-score of 0.9243, indicating a strong baseline performance. However, when GMP is directly combined with GAP without any additional processing, the performance drops to an accuracy of 0.8976. This result demonstrates that naive feature fusion is not sufficient and may negatively affect the representation quality.

To address this limitation, a lightweight bottleneck classification head is introduced after feature aggregation. With this design, the proposed model achieves the highest performance, reaching an accuracy of 0.9548 and a weighted F1-score of 0.9551. This

improvement confirms that controlled feature refinement after dual pooling is essential for effective feature integration.

In addition to the proposed configuration, alternative feature refinement modules are also investigated. One of these approaches employs a Dynamic Progressive Gated Refinement (DPGR) mechanism, which aims to progressively refine feature representations by applying gated transformations at different stages. This approach achieves an accuracy of 0.9307, improving upon the baseline but remaining below the proposed model.

Another variant utilizes a Lightweight Global Fusion (LGF) module, designed to enhance feature interaction by combining global contextual information with channel-wise representations in a computationally efficient manner. This configuration achieves an accuracy of 0.9398, indicating that global feature interaction contributes positively to performance.

Despite these improvements, both DPGR and LGF-based configurations fail to surpass the proposed model. This suggests that introducing additional complexity through alternative modules does not necessarily lead to better performance unless feature aggregation is properly structured.

Overall, the ablation results presented in Table 3 clearly demonstrate that the proposed hybrid feature aggregation strategy, combined with a lightweight refinement mechanism, provides the most effective and balanced solution for PV defect classification. The findings highlight that the synergy between dual pooling and controlled feature refinement is more critical than the use of more complex architectural components.

Table 3. Ablation study results of the proposed model

Model	Accuracy	Macro F1	Weighted F1
ConvNeXt (GAP)	0.9247	0.9234	0.9243
GAP + GMP (no refinement)	0.8976	0.8979	0.8985
Dual pooling + bottleneck (proposed)	0.9548	0.9546	0.9551
DPGR-based refinement	0.9307	0.9305	0.9308
LGF-based refinement	0.9398	0.9399	0.9399

4. Discussion

The experimental results demonstrate that the proposed ConvNeXt-based architecture achieves superior performance compared to widely used deep learning models in PV defect classification. This improvement is mainly attributed to the proposed hybrid feature aggregation strategy, which combines GAP and GMP.

Unlike conventional approaches relying on a single pooling mechanism, the proposed method benefits from the complementary nature of GAP and GMP. The ablation results indicate that neither individual pooling nor naive fusion is sufficient, while the proposed structured aggregation with a lightweight bottleneck significantly improves performance.

The model also shows balanced performance across all defect categories. While classes such as Electrical-damage and Snow-Covered achieve near-perfect results due to their distinctive patterns, relatively lower performance in the Dusty class is observed. This is mainly due to visual similarity between contamination-related categories, which is a known challenge in PV defect classification.

Furthermore, alternative refinement modules such as DPGR and LGF provide improvements over the baseline but do not surpass the proposed configuration. This suggests that increasing architectural complexity alone is not sufficient, and that effective feature aggregation plays a more critical role.

Despite the strong performance, the dataset size is relatively limited, which may affect generalization in more diverse real-world scenarios. Future work may focus on larger datasets and integrating more advanced feature modeling techniques.

Overall, the results confirm that the proposed approach provides an effective and efficient solution, highlighting the importance of feature aggregation in PV defect classification tasks.

5. Conclusion

In this study, a lightweight ConvNeXt-based architecture is proposed for multiclass PV defect classification. The model introduces a hybrid feature aggregation strategy that combines GAP and GMP, along with a lightweight bottleneck classification head for effective feature refinement.

Experimental results demonstrate that the proposed model achieves superior performance compared to widely used deep learning architectures, reaching an accuracy of 0.9548 and a weighted F1-score of 0.9551. The ablation study further confirms that the proposed feature aggregation mechanism plays a key role in improving classification performance.

The findings indicate that effective feature aggregation is more critical than increasing architectural complexity for PV defect classification. The proposed approach provides a favorable balance between classification accuracy and computational cost, making it suitable for real-world industrial applications.

Despite the strong performance, the proposed method has some limitations. First, the dataset used in this study is relatively limited in size, which may affect the generalization capability of the model in more diverse

real-world scenarios. Second, certain defect classes such as Dusty, Bird-drop, and Clean exhibit visual similarities, leading to occasional misclassifications. These challenges indicate that the model may require more diverse data or additional feature modeling strategies to further improve robustness in complex industrial environments.

Future work may focus on extending the proposed framework to larger datasets and integrating advanced feature modeling techniques to further improve robustness and generalization.

Author Contributions

The percentages of the author' contributions are presented below. The author reviewed and approved the final version of the manuscript.

	B.E.
C	100
D	100
S	100
DCP	100
DAI	100
L	100
W	100
CR	100
SR	100
PM	100
FA	100

C= concept, D= design, S= supervision, DCP= data collection and/or processing, DAI= data analysis and/or interpretation, L= literature search, W= writing, CR= critical review, SR= submission and revision, PM= project management, FA= funding acquisition.

Conflict of Interest

The author declared that there is no conflict of interest.

Ethical Consideration

Ethics committee approval was not required for this study because of there was no study on animals or humans.

References

Ahmed, W., Hanif, A., Kallu, K. D., Kouzani, A. Z., Ali, M. U., & Zafar, A. (2021). Photovoltaic panels classification using isolated and transfer learned deep neural models using infrared thermographic images. *Sensors*, *21*(16), Article 5668. https://doi.org/10.3390/s21165668

Akram, M. W., & Bai, J. (2025). Defect detection in photovoltaic modules based on image-to-image generation and deep learning. *Sustainable Energy Technologies and Assessments*, *82*, Article 104441. https://doi.org/10.1016/j.seta.2025.104441

Cao, Y., Pang, D., Yan, Y., Jiang, Y., & Tian, C. (2023). A photovoltaic surface defect detection method for building based on deep learning. *Journal of Building Engineering*, *70*, Article 106375. https://doi.org/10.1016/j.jobe.2023.106375

Demir, F. (2025). Enhancing defect classification in solar panels with electroluminescence imaging and advanced machine learning strategies. *IEEE Access*, *13*, 58481–58495. https://doi.org/10.1109/ACCESS.2025.3551749

Diaconu, B. M. (2026). Diagnosing shortcut learning in CNN-based photovoltaic fault recognition from RGB images: A multi-method explainability audit. *AI*, *7*(3), Article 94. https://doi.org/10.3390/ai7030094

Ejiyi, C. J., Cai, D., Johnson, N., Osei-Mensah, E., Eze, F., Asare, S. K., Staffell, I., & Bamisile, O. O. (2026). SolarSynthNet (SSN): A deep learning framework for binary and multiclass classification of damaged or obstructed solar panels using images. *Renewable Energy*, *256*, Article 124224. https://doi.org/10.1016/j.renene.2025.124224

Eren, B. (2026). Deep learning approaches for weld defect detection: A comprehensive review of models, applications, and future directions. *Computers & Industrial Engineering*, *212*, Article 111725. https://doi.org/10.1016/j.cie.2025.111725

Hijjawi, U., Lakshminarayana, S., Xu, T., Piero Malfense Fierro, G., & Rahman, M. (2023). A review of automated solar photovoltaic defect detection systems: Approaches, challenges, and future orientations. *Solar Energy*, *266*, Article 112186. https://doi.org/10.1016/j.solener.2023.112186

Huang, J., Ariffin, S. A., Chen, Y., Lin, J., & Xu, W. (2025). A novel MoCo-based self-supervised learning framework for solar panel defect detection. *IEEE Access*, *13*, 22977–22988. https://doi.org/10.1109/ACCESS.2025.3529701

Ibn Malek, I., & Imtiaz, H. (2025). Advanced fault detection in PV panels using deep neural networks: Leveraging transfer learning and electroluminescence image processing. *Energy Advances*, 180–193. https://doi.org/10.1039/d5ya00239g

Jiang, Y., & Zhao, C. (2022). Attention classification-and-segmentation network for micro-crack anomaly detection of photovoltaic module cells. *Solar Energy*, *238*, 291–304. https://doi.org/10.1016/j.solener.2022.04.012

Khanam, R., Hussain, M., Hill, R., & Allen, P. (2024). A comprehensive review of convolutional neural networks for defect detection in industrial applications. *IEEE Access*, *12*, 94250–94295. https://doi.org/10.1109/ACCESS.2024.3425166

Laouamer, M., Adaika, M., Remha, S., Mahmoudi, A., & Adaika, H. (2025). Comparative deep learning for RGB-based PV surface fault classification using ResNet50 and EfficientNetB0 with real-time deployment. *International Journal of Electrical and Electronics Research*, *13*(4), Article 712. https://doi.org/10.37391/ijeer.130411

Ledmaoui, Y., El Maghraoui, Adila., El Aroussi, M., & Saadane, R. (2024). Enhanced fault detection in photovoltaic panels using CNN-based classification with PyQt5 implementation. *Sensors*, *24*(22), Article 7407. https://doi.org/10.3390/s24227407

Masita, K., Hasan, A., Shongwe, T., & Hilal, H. A. (2025). Deep learning in defects detection of PV modules: A review. *Solar Energy Advances*, *5*, Article 100090. https://doi.org/10.1016/j.seja.2025.100090

Munawer Al-Otum, H. (2024). Classification of anomalies in electroluminescence images of solar PV modules using CNN-based deep learning. *Solar Energy*, *278*, Article 112803. https://doi.org/10.1016/j.solener.2024.112803

Pathak, S. P., & Patil, S. A. (2023). Evaluation of effect of pre-processing techniques in solar panel fault detection. *IEEE Access*, *11*, 72848–72860. https://doi.org/10.1109/ACCESS.2023.3293756

Shaban, W. M. (2024). Detection and classification of photovoltaic module defects based on artificial intelligence.

Neural Computing and Applications, *36*(27), 16769–16796. https://doi.org/10.1007/s00521-024-10000-z

Sun, Y., Huang, G., Xu, C., Guo, H., & Feng, Y. (2025). Photovoltaic cell surface defect detection via subtle defect enhancement and background suppression. *Micromachines*, *16*(9), Article 1003. https://doi.org/10.3390/mi16091003

Tella, H., Hussein, A., Rehman, S., Liu, B., Balghonaim, A., & Mohandes, M. (2025). Solar photovoltaic panel cells defects classification using deep learning ensemble methods. *Case Studies in Thermal Engineering*, *66*, Article 105749. https://doi.org/10.1016/j.csite.2025.105749

Thakfan, A., & Bin Salamah, Y. (2024). Artificial-intelligence-based detection of defects and faults in photovoltaic systems: A survey. *Energies*, *17*(19), Article 4807. https://doi.org/10.3390/en17194807

Xiao, M., Yang, B., Wang, S., Zhang, Z., & He, Y. (2023). Fine coordinate attention for surface defect detection. *Engineering Applications of Artificial Intelligence*, *123*, Article 106368. https://doi.org/10.1016/j.engappai.2023.106368

Zhang, Z., Yang, Y., & Jian, X. (2025). Defect analysis using electroluminescence imaging in photovoltaic modules. *Measurement Science and Technology*, *36*(8), Article 085001. https://doi.org/10.1088/1361-6501/adf657