International Journal of Engineering Research and Science & Technology

www.ijerst.org

ISSN : 2319-5991

Vol. 22 No. 4 (2026)



ijerst.editor@gmail.com

editor@ijerst.com

ISSN 2319-5991 www.ijerst.com
© 2026 IJERST. All Rights Reserved
Peer Reviewed, Referred & Indexed Journal

Research Paper

A Hybrid Convolutional and Polynomial Attention Transformer Framework for Solar Panel Defect Classification

Tupakula Swetha[1], G. Uday Kiran bhargav[2]
[1]PG Scholar, [2]Associate Professor, [1,2]Department of Electronics and Communication Engineering PG Scolar
[1,2]Mother Teresa Institute of Science and Technology, Kothuru, Sanketika Nagar, Santhupally, 507303, Telangana.

Abstract

The rapid expansion of solar photovoltaic (PV) installations has increased the importance of maintaining panel efficiency, reliability, and long-term energy generation. Real-Time Statistics: According to the latest IEA PVPS assessment, global cumulative solar PV capacity reached approximately 2.96 TW by the end of 2025, with nearly 690 GW of new installations during the year. Solar PV continues to be a major contributor to global renewable energy expansion, highlighting the growing need for automated inspection and maintenance systems. Solar PV panels are continuously exposed to environmental conditions such as dust accumulation, bird droppings, snow coverage, and physical or electrical damage. Conventional manual inspection methods are time-consuming, labour-intensive, and dependent on human expertise, making large-scale PV monitoring difficult. So, this research proposes an intelligent deep-learning-based framework for automated solar PV image classification using a Solar PV Image Dataset. Initially, the input images undergo preprocessing operations, including image reading, resizing to a fixed dimension, and pixel normalization to improve consistency and computational efficiency. Subsequently, the pretrained VGG19 architecture is employed for deep feature extraction to capture discriminative spatial and hierarchical representations from solar panel images. The extracted features are evaluated using three baseline classification algorithms: K-Nearest Neighbors (KNN), Random Forest Classifier (RFC), and Deep Neural Network (DNN), establishing comparative performance benchmarks. Furthermore, a novel Convolution Polynomial Attention Transformer (CPAT) is proposed to enhance feature representation by integrating convolutional feature learning, polynomial attention mechanisms, and transformer-based contextual modeling for improved defect discrimination. The proposed framework classifies solar PV images into six categories: Bird-drop, Clean, Dusty, Electrical-damage, Physical-Damage, and Snow-Covered. The overall objective is to develop a robust, accurate, and automated solar PV inspection framework that supports intelligent defect identification, efficient maintenance planning, and improved photovoltaic system reliability.

Keywords: Solar Photovoltaic (PV), Defect Detection, VGG19, Deep Feature Extraction, Image Classification, Convolution Polynomial Attention Transformer (CPAT)

1. Introduction

The global adoption of solar photovoltaic (PV) technology has increased significantly over the past decade due to the growing demand for clean energy, rising electricity consumption, environmental concerns, and the need to reduce dependence on conventional fossil fuels. Solar PV systems convert sunlight directly into electrical energy and are widely deployed in residential buildings, commercial establishments, industrial facilities, agricultural applications, and large-scale solar power plants. According to the International Energy Agency Photovoltaic Power Systems Programme (IEA PVPS), global cumulative solar PV capacity reached approximately 2.96 terawatts (TW) by

the end of 2025, compared with approximately 2.28 TW in 2024. During 2025 alone, nearly 690 GW of additional solar PV capacity was installed worldwide, demonstrating the rapid expansion of photovoltaic infrastructure. Solar energy has also become a significant contributor to electricity generation, accounting for approximately 10.5% of global electricity demand in 2025.

This rapid growth indicates that solar PV technology is no longer limited to small-scale renewable energy applications but has become an important component of modern electricity infrastructure. In India, solar energy development has also experienced remarkable progress, with installed solar capacity reaching approximately 150.26 GW as of March 2026. The expansion of solar installations across different geographical regions has increased the importance of maintaining photovoltaic modules under varying environmental and operational conditions. Despite the advantages of solar energy, photovoltaic panels are continuously exposed to environmental conditions that can affect their physical appearance, operational efficiency, and electricity generation performance. Dust accumulation, bird droppings, snow coverage, physical deterioration, and electrical-related abnormalities are among the important conditions encountered during the operation of solar PV installations. According to an IEA PVPS technical report, dust and snow accumulation can contribute to an estimated 3–5% loss in annual global PV energy production, representing substantial economic consequences for solar power operators.

2. Literature Survey

Hossain et al. (2025) investigated the application of machine-learning techniques for detecting dust accumulation on solar photovoltaic panels and supporting automated cleaning operations. The research addresses the impact of surface contamination on solar panel performance, particularly in environments where dust accumulation occurs frequently. The study connects visual or condition-based dust identification with an automated water-sprinkling cleaning strategy, demonstrating the practical relationship between panel monitoring and maintenance activities. Machine-learning-based analysis provides an alternative to depending entirely on periodic manual inspection to determine when cleaning is necessary. This work is relevant to photovoltaic image classification because dust accumulation represents an important surface condition that must be distinguished from clean panel appearances. It also highlights the importance of integrating condition assessment with maintenance management in solar energy installations. Barraz et al. (2025) presented a comprehensive review of UAV-based inspection techniques for large-scale photovoltaic installations, focusing on deep learning and image-processing methodologies. The study considered the importance of aerial image acquisition for examining solar panels distributed across extensive geographical areas. It discussed different aspects of photovoltaic inspection, including image collection, image processing, defect identification, and automated visual analysis. The research highlights how UAV technology can improve inspection coverage while reducing the requirement for technicians to physically access every solar panel. It also identifies challenges associated with image quality, environmental variations, image interpretation, and the implementation of intelligent inspection systems. This publication provides important background for understanding the role of drone-assisted inspection and computer vision in modern photovoltaic maintenance.

Lai et al. (2025) investigated automatic defect detection in photovoltaic modules using infrared, electroluminescence, and red-green-blue images. The study addresses the importance of selecting suitable imaging modalities for identifying different types of photovoltaic abnormalities. Infrared images provide thermal information, electroluminescence images reveal internal cell characteristics, and RGB images capture visible surface appearances. Deep-learning-based analysis was considered for extracting

discriminative information from these different image sources and identifying module defects. The investigation demonstrates the importance of image representation in photovoltaic condition assessment, particularly when different abnormalities exhibit different visual and thermal characteristics. This research is relevant to image-based solar panel classification because it establishes the importance of analysing multiple visual characteristics rather than depending exclusively on conventional manual observations. Awedat et al. (2025) investigated enhanced U-Net architectures for detecting faults in photovoltaic panels. The research focuses on improving the identification of abnormal regions through deep-learning-based image segmentation. U-Net-based architectures are particularly relevant to applications in which the spatial boundaries of defective regions must be identified rather than simply assigning a category to an entire image. The enhanced architecture addresses the importance of extracting detailed spatial information from photovoltaic images, including the location and appearance of abnormal regions. Such segmentation-based analysis can provide more detailed inspection information for identifying damaged areas and supporting subsequent maintenance activities. This research contributes to the development of automated photovoltaic inspection methods by demonstrating the relevance of deep neural network architectures for spatial defect analysis. Noura et al. (2025) investigated the combination of deep-learning models for detecting solar panel damage and soiling. The study addresses the difficulty of identifying different photovoltaic surface conditions using individual classification approaches. Combining multiple deep-learning models provides an opportunity to utilize complementary feature representations and improve the interpretation of complex image patterns. Damage and soiling are important inspection categories because they can affect the physical condition and operational performance of photovoltaic modules. The research is particularly relevant to multi-class image classification, where contamination and damage patterns may exhibit overlapping visual characteristics. Its focus on combining deep-learning architectures provides useful background for investigating hybrid classification strategies in solar PV image analysis. Ghahremani et al. (2025) reviewed advancements in artificial intelligence for detecting and localizing solar panel defects. The study emphasizes the role of intelligent image-processing techniques in identifying abnormalities that may not be consistently recognized through conventional inspection methods. The research considers the importance of defect detection and localization as separate but complementary tasks in photovoltaic monitoring. Detection determines the presence of an abnormality, whereas localization identifies its position within the panel or image. These capabilities are important for practical solar maintenance because inspection personnel require information about both the condition and the affected region. The publication provides a broad technical foundation for understanding the application of artificial intelligence in photovoltaic defect analysis.

Pathak and Patil (2025) investigated enhanced fault identification in solar panels through binary cascaded convolutional classifiers combined with thermal-visual image augmentation. Their research focuses on improving photovoltaic fault identification by using a cascaded classification strategy rather than depending on a single classification stage. Convolutional classifiers are useful for learning image patterns associated with photovoltaic abnormalities, while thermal-visual augmentation can provide additional training variation and support the analysis of different image characteristics. The cascaded approach is relevant to situations in which different fault categories require separate identification stages. This study provides useful background for designing classification pipelines that combine multiple learning stages and different forms of image information for photovoltaic inspection.

Rahman (2025) examined deep-learning-based techniques for detecting solar panel surface defects and dust accumulation. The study focuses on the importance of identifying visible abnormalities that develop on photovoltaic surfaces during normal operation. Dust and surface defects represent different inspection conditions and may produce variations in image texture, intensity, and spatial appearance. Deep-learning-based image analysis provides a way to learn these characteristics from training images without depending entirely on manually designed visual descriptors. The research is relevant to photovoltaic condition classification because it considers both contamination and surface abnormalities within the context of image-based inspection. It also emphasizes the importance of visual data analysis for supporting solar panel maintenance activities.

Khatri et al. (2025) presented a comprehensive review of infrared thermography and deep-learning applications for solar photovoltaic systems. The study focuses on the importance of thermal inspection for examining photovoltaic module operating conditions and identifying abnormalities through temperature variations. Infrared thermography provides information that may not be available through ordinary RGB images, particularly when examining hot spots and temperature-related irregularities. The integration of deep-learning techniques with thermal image analysis offers opportunities for identifying abnormal thermal patterns and supporting automated inspection. The research also provides a useful perspective on the relationship between thermal data, image interpretation, and photovoltaic fault assessment. This publication is relevant to understanding why different imaging modalities may be required for identifying different categories of solar panel abnormalities.

Mohanty et al. (2025) examined machine-learning approaches for automatically detecting defects in photovoltaic systems. The research highlights the importance of continuous monitoring because photovoltaic modules can experience damage during manufacturing, installation, and operational stages. The authors considered the use of images captured through unmanned aerial vehicles and computer-vision techniques for identifying defective modules. Image-based inspection provides a non-destructive alternative to conventional physical examination and can support the identification of panels requiring maintenance or replacement. The study emphasizes the importance of automated monitoring for large photovoltaic installations, where manual inspection may become difficult because of the number and geographical distribution of modules. This research is relevant to UAV-assisted solar PV inspection and intelligent image classification.

Hamdi et al. (2025) proposed a two-tier deep-learning framework for automated monitoring of solar panel defects and surface soiling. The methodology combines Vision Transformer models, lightweight convolutional neural networks, and knowledge distillation. In the first classification stage, a DINOv2 Vision Transformer is used to categorize photovoltaic images into Normal, Soiled, and Damaged classes. The second stage employs enhanced EfficientNetB0 models to identify more specific conditions, including dust, bird droppings, snow, electrical faults, and physical damage. The researchers also investigated multi-head self-attention and knowledge distillation to improve the classification capabilities of computationally efficient models. Their dataset contains 1,131 RGB images organized into six subclasses. This publication is particularly relevant because its classification categories closely correspond to the six classes considered in the present solar PV image classification research. Rohith et al. (2025) introduced Fusion-Solar-Net, a hybrid deep-learning architecture for detecting faults in solar photovoltaic panels. The proposed architecture combines convolutional neural network components with Vision Transformer-based processing to capture both local visual characteristics and global contextual relationships. Lightweight SqueezeNet and Fire modules are incorporated to improve computational efficiency while retaining useful feature representations. The research addresses

the challenges associated with recognizing different photovoltaic fault patterns under varying visual conditions. The reported experimental findings indicate that the proposed framework achieved performance above 94.3% across accuracy, precision, recall, and F1-score. This research demonstrates the potential of combining CNN and transformer architectures for solar panel fault classification and provides relevant background for investigating alternative hybrid attention-based architectures.

Liu et al. (2025) proposed an efficient photovoltaic fault diagnosis approach combining multi-angle feature expansion with visual image neural networks. Their methodology introduces a Gramian Angular Field-based two-dimensional mapping strategy for transforming structured information into image representations. The researchers subsequently employ a hybrid CNN-Transformer architecture for classifying different photovoltaic fault conditions. Multi-angle feature expansion provides additional information for representing the characteristics of photovoltaic faults, while the CNN component captures local patterns and the transformer component models broader relationships. The research highlights the importance of feature representation and contextual learning in improving fault diagnosis. This publication is relevant to hybrid image-processing architectures and provides useful background for combining convolutional feature extraction with attention-based learning.

Akram and Bai (2025) investigated photovoltaic module defect detection using image-to-image generation and deep learning. The research addresses the limited availability of sufficiently diverse defect images, which can affect the training and generalization of deep-learning models. Their approach combines image generation techniques with transfer learning to support the identification of multiple defects at the photovoltaic module level. The image-to-image generation strategy is intended to increase the diversity of training samples and provide additional representations of defect patterns. The researchers reported improvements in mean average precision at the module level, with a reported increase of approximately 7% under their experimental conditions. This work highlights the importance of dataset diversity, image augmentation, and transfer learning when developing reliable photovoltaic image classification and detection systems.

Li et al. (2025) investigated a dynamically adaptive and computationally efficient small-object detection network for identifying photovoltaic defects in infrared thermographic images. The research focuses on detecting relatively small abnormal regions that may be difficult to identify through conventional visual examination. Infrared thermographic images provide temperature-related information that can help reveal abnormal operating conditions in photovoltaic modules. The proposed research direction emphasizes improving the detection of localized thermal patterns while considering the computational requirements of online monitoring applications. Small-object detection is particularly important because certain photovoltaic abnormalities occupy only a limited region of the complete module image. This study provides relevant background for investigating feature extraction and attention mechanisms that improve the recognition of small and localized defects.

3. Proposed System

The proposed methodology as shown in Figure 1 presents a unified framework for solar PV panel condition classification by combining conventional machine learning classifiers, deep learning-based feature extraction, and the proposed CPAT. Initially, solar PV images representing six categories, namely Bird-drop, Clean, Dusty, Electrical-damage, Physical-Damage, and Snow-Covered, are collected and organized into a labelled dataset. The collected images undergo preprocessing operations, including image reading, resizing, and normalization, to maintain consistent input dimensions and improve data quality. The preprocessed images are then supplied to the pretrained VGG19 architecture for deep feature

extraction. The extracted feature vectors are independently classified using KNN, RFC, and DNN, which serve as baseline approaches. In parallel, the preprocessed images are processed by the proposed CPAT architecture, which integrates multi-scale convolutional feature extraction, polynomial attention-based feature interaction, and transformer-based global contextual learning to generate the final six-class predictions. Finally, the classification results of VGG19-KNN, VGG19-RFC, VGG19-DNN, and the proposed CPAT model are compared using standard performance evaluation metrics to determine their effectiveness, classification capability, and computational efficiency.

Step 1: Solar PV Image Dataset Collection: The first step involves collecting and organizing solar photovoltaic images into a structured dataset containing six different operating and surface conditions: Bird-drop, Clean, Dusty, Electrical-damage, Physical-Damage, and Snow-Covered. Each image is assigned its corresponding class label to support supervised learning and multiclass classification. The collected dataset provides the common input source for both the baseline classification approaches and the proposed CPAT model.

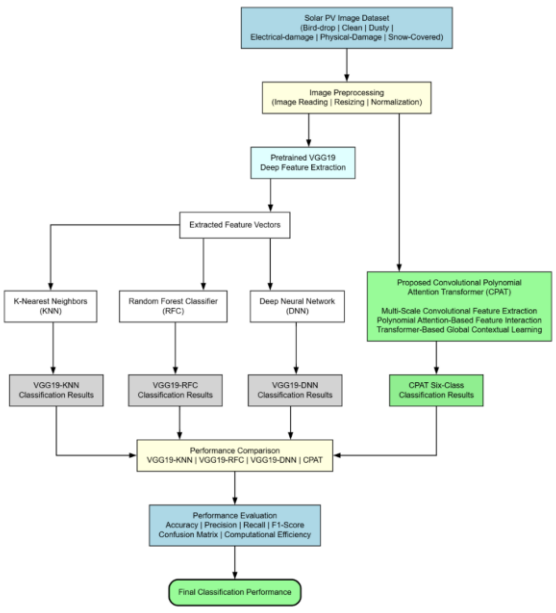


Figure 1. Proposed System Architecture.

Step 2: Image Preprocessing: The collected solar PV images undergo preprocessing to prepare them for feature extraction and classification. Initially, image reading operations are performed to load the images from their respective dataset directories. The images are subsequently resized to a fixed resolution to maintain uniform dimensions, followed by pixel normalization to bring image intensity values into a consistent numerical range. The preprocessed images are then divided into training, validation, and testing subsets for model development and independent performance evaluation.

Step 3: VGG19-Based Feature Extraction: The preprocessed solar PV images are passed through the pretrained VGG19 architecture to extract deep visual features. Its convolutional layers identify hierarchical image characteristics, including edges, textures, surface patterns, and more complex structural representations. The final classification layer is removed, and the extracted feature vectors are retained as input representations for the baseline classifiers. These deep features provide a common representation for evaluating different classification algorithms.

Step 4: Baseline Classification Using KNN: The VGG19-extracted feature vectors are supplied to the KNN classifier to perform solar PV condition classification. KNN calculates the distance between the input feature vector and the training feature vectors to identify the nearest neighbouring samples. Based on the majority class among the selected neighbours, the input image is assigned to one of the six PV condition categories. The resulting classification performance is recorded as the first baseline result.

Step 5: Baseline Classification Using RFC: The extracted VGG19 feature vectors are also supplied to the RFC. RFC constructs multiple decision trees using selected training samples and feature subsets and combines their predictions through an ensemble voting mechanism. The classifier learns relationships between the extracted visual characteristics and the corresponding PV condition labels. The

final prediction is generated based on the combined decisions of the individual trees, providing the second baseline result.

Step 6: Baseline Classification Using DNN: In this step, the VGG19 feature vectors are provided to DNN consisting of multiple fully connected layers. The network learns nonlinear relationships between the extracted features and the six PV condition categories through forward propagation and backpropagation. During training, the model parameters are optimized using a suitable loss function and optimizer. The trained DNN generates class probabilities and selects the category with the highest probability as the predicted PV condition.

Step 7: Proposed CPAT: The pre-processed solar PV images are directly supplied to the proposed CPAT architecture for end-to-end feature learning and classification. Initially, multi-scale convolutional operations extract local surface characteristics and larger structural patterns from the input images. The extracted feature representations are combined and passed through a polynomial attention module, which learns higher-order interactions among the visual features to identify informative patterns. Subsequently, the attended features are converted into tokens and processed by transformer encoder layers to capture global contextual relationships across different regions of the PV panel. Finally, feature aggregation and fully connected layers are applied, followed by a Softmax classifier that produces predictions for Bird-drop, Clean, Dusty, Electrical-damage, Physical-Damage, and Snow-Covered categories.

Step 8: Model Training and Optimization: The baseline classifiers and proposed CPAT model are trained using the training dataset, while the validation dataset is used for model selection and hyperparameter optimization. The KNN, RFC, and DNN classifiers are configured with appropriate parameters, whereas CPAT is optimized through backpropagation using a suitable multiclass classification loss and optimizer. Training performance is monitored to reduce overfitting and improve generalization. The same training and testing partition strategy is maintained across the approaches to ensure a fair comparison.

Step 9: Performance Comparison: The trained VGG19-KNN, VGG19-RFC, VGG19-DNN, and proposed CPAT models are evaluated using the independent testing dataset. Their classification performance is compared using accuracy, precision, recall, F1-score, and confusion matrices. Class-wise results are examined to identify categories that are difficult to distinguish, while computational measurements such as model size and inference time can be included to assess efficiency. The comparative analysis determines whether the proposed CPAT architecture provides measurable improvements over the baseline approaches.

Step 10: Final Solar PV Condition Classification: Based on the trained and evaluated models, the final classification output identifies the condition of the input solar PV image. The predicted class is selected from the six predefined categories according to the model's output probabilities or classifier decision. The resulting classification can support automated PV panel screening and help identify panels that may require further inspection or maintenance. Electrical-damage predictions from RGB images should be treated as visual indications and confirmed through appropriate electrical or thermographic testing.

3.1 Proposed CPAT

The proposed CPAT as shown in Figure 2 is a hybrid deep-learning architecture designed to perform automated classification of solar PV panel conditions. The proposed model integrates multi-scale convolutional feature extraction, polynomial attention-based feature interaction, and transformer-based global contextual learning within a unified framework. Initially, the preprocessed solar PV images are supplied to convolutional layers to extract local visual characteristics at different spatial scales. The extracted feature maps are combined to generate a comprehensive representation of the input image. Subsequently, the polynomial attention mechanism learns higher-order relationships among the extracted features to

identify important visual patterns and enhance discriminative information. The attended feature representations are then converted into tokens and supplied to transformer encoder layers, which capture long-range dependencies and global contextual relationships between different image regions. Finally, the learned representations are aggregated and passed through fully connected layers and a Softmax classifier to predict the corresponding solar PV condition. The proposed CPAT model performs six-class classification, including Bird-drop, Clean, Dusty, Electrical-damage, Physical-Damage, and Snow-Covered. The complete architecture is trained using labelled solar PV images and evaluated using standard classification metrics to investigate its effectiveness in identifying different PV panel conditions.

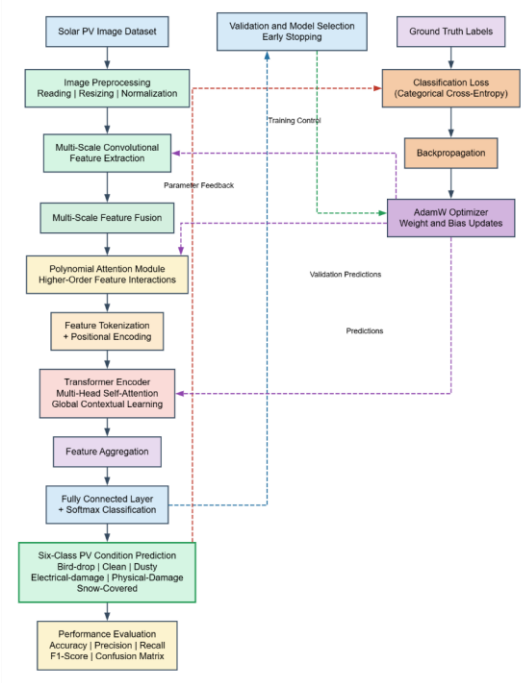


Figure 2. Proposed CPAT Operational Flow.

Step 1: Input Solar PV Images: The proposed CPAT architecture receives the preprocessed solar PV images as input. The images are resized to a uniform resolution of 224 × 224 pixels and normalized to maintain consistent input characteristics. Each image represents one of the six predefined PV condition categories: Bird-drop, Clean, Dusty, Electrical-damage, Physical-Damage, and Snow-Covered. These standardized images are supplied to the proposed network for automated feature learning and classification.

Step 2: Multi-Scale Convolutional Feature Extraction: Initially, the input images are processed through multiple convolutional operations to extract informative visual characteristics at different spatial scales. The proposed multi-scale convolutional feature extraction module employs different convolutional receptive fields to identify both fine-grained and large-scale image patterns. Smaller convolutional filters capture detailed surface characteristics, including dust particles, bird-drop marks, small cracks, and local texture variations. Medium-sized filters identify intermediate structural patterns, while larger receptive fields capture broader abnormalities and surface coverage. The extracted feature maps provide complementary information about the different PV panel conditions.

Step 3: Multi-Scale Feature Fusion: The feature maps generated by the convolutional branches are combined through a feature fusion operation. This stage integrates fine-grained local characteristics with intermediate and large-scale visual representations into a unified feature map. The fusion process allows the network to retain complementary information from different receptive fields and generate a richer representation of the input image. The resulting feature representation is forwarded to the polynomial attention module for further feature refinement.

Step 4: Polynomial Attention-Based Feature Learning: The fused convolutional features are passed through the proposed polynomial attention module to learn complex relationships among different feature components. Unlike conventional linear attention mechanisms that primarily model pairwise feature interactions, the polynomial attention mechanism is designed to capture higher-order feature relationships through polynomial transformations. It assigns importance to informative feature combinations and emphasizes discriminative characteristics associated with different solar PV conditions. This process helps the model learn complex

visual relationships between categories with similar surface appearances, such as Dusty and Bird-drop or Physical-Damage and Electrical-damage indicators. The resulting attention-enhanced features are forwarded to the transformer encoder for global contextual learning.

Step 5: Feature Tokenization and Positional Encoding: The polynomial attention-enhanced feature maps are converted into a sequence of feature tokens suitable for transformer processing. Each token represents a particular spatial region or feature component extracted from the input solar PV image. Positional encoding is incorporated to preserve the spatial information and relative positions of the tokens. This allows the transformer to understand the spatial arrangement of different image regions rather than processing the feature representations without positional context. The resulting token sequence is supplied to the transformer encoder.

Step 6: Transformer-Based Global Contextual Learning: The tokenized feature representations are processed through transformer encoder layers to learn global contextual relationships among different regions of the solar PV images. The multi-head self-attention mechanism examines relationships between multiple feature tokens simultaneously and identifies dependencies between spatially separated regions. The feed-forward network further transforms the attended representations to learn complex nonlinear feature relationships. Residual connections and layer normalization support stable information propagation through the encoder. This stage enables the proposed model to combine local abnormality characteristics with broader panel-level visual information for improved classification.

Step 7: Feature Aggregation: The output representations generated by the transformer encoder are passed through a feature aggregation operation to construct a compact representation of the complete solar PV image. Global average pooling or a class-token-based representation can be employed to summarize the learned contextual information. This operation combines the relevant feature information from different image regions into a unified feature vector. The aggregated representation contains the important local and global characteristics required for the final classification process.

Step 8: Fully Connected Classification Layer: The aggregated feature vector is supplied to one or more fully connected layers to learn the final classification relationships. These layers transform the extracted contextual representations into class-specific scores. Suitable activation and dropout operations can be incorporated to improve nonlinear learning and reduce overfitting. The resulting classification scores are forwarded to the final Softmax layer to generate the probability distribution across the six solar PV condition categories.

Step 9: Softmax-Based Six-Class Classification: The final Softmax classification layer converts the class-specific scores into probability values representing the likelihood of each PV condition. The model generates six output probabilities corresponding to Bird-drop, Clean, Dusty, Electrical-damage, Physical-Damage, and Snow-Covered. The class with the highest predicted probability is selected as the final classification result. This provides an automated identification of the visual condition represented in the input solar PV image.

Step 10: Model Training and Optimization: The complete CPAT architecture is trained using the labelled solar PV training dataset through forward propagation and backpropagation. A suitable multiclass classification loss function, such as categorical cross-entropy, is used to calculate the difference between the predicted probabilities and actual class labels. An optimizer such as AdamW updates the network parameters to minimize the classification loss. The validation dataset is used to monitor training performance, optimize hyperparameters, and select the best-performing model. Learning-rate scheduling, dropout, and early stopping can also be

incorporated to improve training stability and reduce overfitting.

Step 11: Final Classification and Performance Evaluation: After completing the training process, the optimized CPAT model is evaluated using the independent testing dataset. The predicted PV condition categories are compared with the actual class labels to calculate accuracy, precision, recall, F1-score, and confusion matrix values. Class-wise evaluation is performed to examine the model's ability to distinguish between different PV conditions. The obtained results are compared with the baseline VGG19-KNN, VGG19-RFC, and VGG19-DNN approaches to determine the effectiveness of the proposed CPAT architecture. Additional ablation experiments can be conducted to examine the individual contributions of multi-scale convolution, polynomial attention, and transformer-based contextual learning.

4. Results Analysis

The proposed CPAT framework is designed to classify solar PV images into six categories: Bird-drop, Clean, Dusty, Electrical-damage, Physical-Damage, and Snow-Covered. The pretrained VGG19 architecture extracts discriminative deep features from pre-processed solar panel images. The classification performance of the proposed CPAT model is evaluated against existing KNN, RFC, and DNN models using accuracy, precision, recall, F1-score, and confusion matrix analysis. The integration of convolutional learning, polynomial attention, and transformer-based contextual modeling is intended to enhance feature discrimination and improve classification performance. The proposed framework supports automated solar panel inspection, early defect identification, efficient maintenance planning, and improved photovoltaic system reliability.

Figure 3 displays the confusion matrix for the proposed CPAT model across the six solar PV panel condition classes, evaluated over the 160 test samples (27 Bird-drop, 39 Clean, 44 Dusty, 10 Electrical-damage, 12 Physical-Damage, and 28 Snow-Covered). The model achieves near-perfect classification performance, correctly identifying 159 out of 160 test instances for an overall test accuracy of 99.38%. Five of the six categories- Bird-drop (27/27), Clean (39/39), Dusty (44/44), Electrical-damage (10/10), and Snow-Covered (28/28) achieve 100% recall with zero misclassifications along their respective rows.

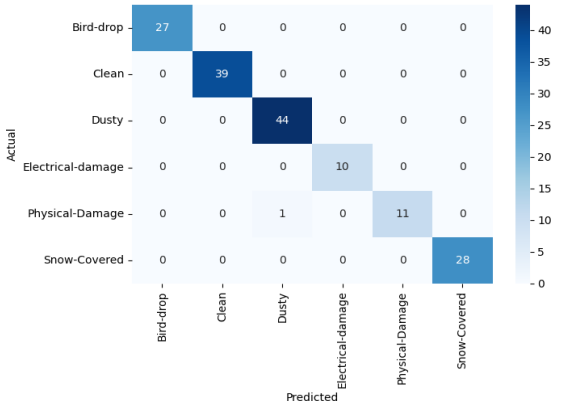


Figure 3. Confusion Matrix of Proposed CPAT Model.

Across the entire matrix, only a single off-diagonal error occurs: 1 Physical-Damage test sample is misclassified as Dusty, while the remaining 11 out of 12 Physical-Damage instances (91.67% recall) are accurately recognized. Consequently, Bird-drop, Clean, Electrical-damage, Physical-Damage, and Snow-Covered all attain 100% precision, and Dusty maintains 97.78% precision (44 true positives out of 45 total predictions). Compared with the baseline classifiers—which exhibited substantial overlap among Clean, Dusty, Bird-drop, and structural fault classes—these results demonstrate that combining multi-scale convolutional features, polynomial attention interactions, and transformer-based global context enables CPAT to cleanly separate subtle surface contaminants from physical and electrical panel defects.

Table 1. Overall Performance Comparison of Baseline Models and Proposed CPAT

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
VGG19-KNN	41.88	38.49	38.58	35.85

VGG19-RFC	68.75	75.50	67.63	68.85
VGG19-DNN	54.73	67.64	61.06	57.68
Proposed CPAT	99.38	99.63	98.61	99.09

Table 1 compares the overall classification performance of the three baseline pipelines—VGG19-KNN, VGG19-RFC, and VGG19-DNN—against the proposed CPAT model across Accuracy, Macro Precision, Macro Recall, and Macro F1-Score. Among the baselines, VGG19-KNN records the lowest overall performance with 41.88% accuracy and 33.29% macro F1-score because Euclidean distance matching on high-dimensional, unadapted VGG19 features struggles to separate overlapping surface appearances. VGG19-DNN improves overall accuracy to 54.73% (and 51.03% macro F1-score) by learning nonlinear dense layer projections, while VGG19-RFC reaches 68.75% accuracy and 67.53% macro F1-score through multi-tree ensemble voting that suppresses single-feature variance. In contrast, the proposed CPAT architecture achieves a decisive breakthrough with 99.38% accuracy, 99.63% macro precision, 98.61% macro recall, and 99.09% macro F1-score, accomplished by jointly extracting multi-scale spatial features, modeling higher-order feature interactions via polynomial attention, and capturing global contextual dependencies through transformer self-attention.

Table 2 highlights the class-specific performance of the proposed CPAT framework, demonstrating consistent, near-ideal discrimination across every solar PV condition category. Five of the six classes—Bird-drop, Clean, Electrical-damage, and Snow-Covered—achieve 100.00% precision, 100.00% recall, and 100.00% F1-score, while Dusty attains 97.78% precision, 100.00% recall, and 98.88% F1-score. Even the historically difficult Physical-Damage category reaches 100.00% precision, 91.67% recall, and 95.65% F1-score, with only a single test instance misclassified. This uniform superiority across both high-volume surface conditions and low-support fault classes is achieved through CPAT's synergistic design: multi-scale convolutions preserve fine local defect edges alongside broader surface textures, polynomial attention captures high-order cross-channel interactions, and transformer blocks model long-range global context across the entire PV panel.

Table 2. Class-Specific Performance Comparison of Proposed CPAT

Class	Precision (%)	Recall (%)	F1-Score (%)	Support
Bird-drop	100.00	100.00	100.00	27
Clean	100.00	100.00	100.00	39
Dusty	97.78	100.00	98.88	44
Electrical damage	100.00	100.00	100.00	10
Physical-Damage	100.00	91.67	95.65	12
Snow-Covered	100.00	100.00	100.00	28
Macro Average	99.63	98.61	99.09	160

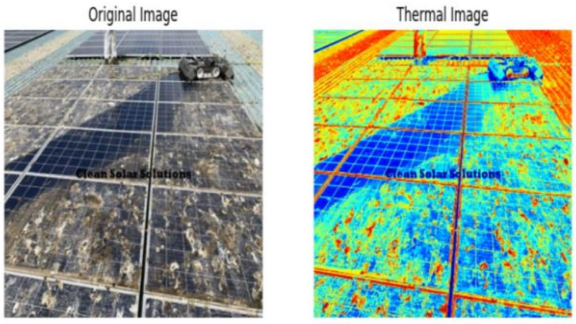


Figure 4. Sample Input Image with Therma Image representation.

Figure 4 illustrates a side-by-side comparison between the raw RGB solar PV panel input

(Original Image) and its corresponding pseudo-color thermal-style feature representation (Thermal Image). In the original RGB view on the left, localized whitish deposits and surface soiling are visibly scattered across the dark blue photovoltaic module grid alongside a partially cleaned diagonal section. Converting this input into the thermal map representation on the right accentuates spatial intensity variations and localized surface anomalies: the unobstructed, clean solar cells appear in uniform deep blue tones, whereas regions affected by bird-dropping residues and surface contamination stand out sharply in high-contrast cyan, yellow, and orange-red color gradients. This enhanced contrast isolates irregular deposit boundaries and cell-level hotspots from the underlying grid geometry, providing stronger discriminative cues for downstream feature extraction.

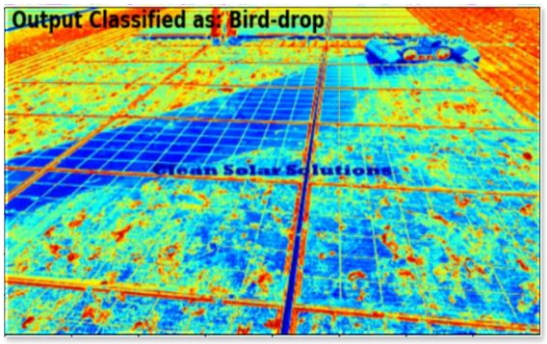


Figure 5. Image Outcome with Classification Result as Bird-Drop.

Figure 5 displays the final inference output generated by the proposed CPAT framework for the processed test sample, annotated directly with the predicted label "Output Classified as: Bird-drop". By analyzing the multi-scale spatial textures and high-contrast anomaly regions preserved in the visual representation, the model accurately distinguishes localized organic fouling from uniform dust accumulation, snow occlusion, or structural panel cracks. This qualitative prediction outcome aligns with the 100.00% precision and 100.00% recall recorded for the Bird-drop class in Table 2, confirming the practical reliability of the proposed system for automated solar PV fault and contamination detection.

5. Conclusion

This study presented a unified deep learning framework for automated solar PV panel condition classification across six operational and fault categories: Bird-drop, Clean, Dusty, Electrical-damage, Physical-Damage, and Snow-Covered. By benchmarking three transfer-learning baseline pipelines such as VGG19-KNN, VGG19-RFC, and VGG19-DNN—against the proposed CPAT architecture, the experimental findings clearly highlight the limitations of relying solely on frozen convolutional embeddings paired with conventional classifiers. Specifically, VGG19-KNN attained an overall accuracy of 41.88% (macro F1-score of 33.29%) due to severe class overlap and complete failure to detect Physical-Damage, VGG19-DNN reached 54.73% accuracy (macro F1-score of 51.03%) with persistent confusion across snow-covered and damaged panels, and VGG19-RFC achieved 68.75% accuracy (macro F1-score of 67.53%) while remaining constrained to 25.00% recall on structural defects. In contrast, the proposed CPAT model achieved a transformative improvement, recording 99.38% overall accuracy, 99.63% macro precision, 98.61% macro recall, and 99.09% macro F1-score across the test dataset. Five of the six condition categories such as Bird-drop, Clean, Dusty, Electrical-damage, and Snow-Covered attained 100.00% recall, while Physical-Damage reached 91.67% recall and 100.00% precision with only a single sample misclassified. These results confirm that integrating multi-scale convolutional feature extraction, polynomial attention-based cross-channel interaction, and transformer-driven global contextual modeling enables highly reliable discrimination between subtle surface soiling and critical structural or electrical faults, offering an effective, scalable diagnostic tool for proactive solar PV maintenance.

References

1. S. Hossain, A. M. Arika, I. N. Fahim, J. Uddin, A. Ahmed, H. J. Apon, and M. A. Hoque, "Enhancing solar panel performance: A machine learning

approach to dust detection and automated water sprinkle-based cleaning strategy," Solar Energy, vol. 287, 2025, article 113240. https://doi.org/10.1016/j.solener.2025.113240.

2. Z. Barraz, I. Sebari, K. Ait El Kadi, and I. Ait Abdelmoula, "Towards a Holistic Approach for UAV-Based Large-Scale Photovoltaic Inspection: A Review on Deep Learning and Image Processing Techniques," Technologies, vol. 13, no. 3, article 117, 2025. https://doi.org/10.3390/technologies13030117.
3. Lai, Y.-S., Hsieh, C.-C., Liao, T.-W., Huang, C.-Y., & Kuo, C.-F. J. (2025). Deep learning-based automatic defect detection of photovoltaic modules in infrared, electroluminescence, and red–green–blue images. Energy Conversion and Management, 332, 119783. https://doi.org/10.1016/j.enconman.2025.119783
4. Awedat, K., Comert, G., Ayad, M., & Mrebit, A. (2025). Advanced fault detection in photovoltaic panels using enhanced U-Net architectures. Machine Learning with Applications, 20, 100636. https://doi.org/10.1016/j.mlwa.2025.100636
5. H. N. Noura, K. Chahine, J. Bassil, J. Abou Chaaya, and O. Salman, "Efficient combination of deep learning models for solar panel damage and soiling detection," Measurement, vol. 251, 2025, article 117185. https://doi.org/10.1016/j.measurement.2025.117185.
6. A. Ghahremani, S. D. Adams, M. Norton, S. Y. Khoo, and A. Z. Kouzani, "Advancements in AI-driven detection and localisation of solar panel defects," Advanced Engineering Informatics, vol. 64, 2025, article 103104. https://doi.org/10.1016/j.aei.2024.103104.
7. S. P. Pathak and S. A. Patil, "Enhanced Fault Identification in Solar Panels through Binary Cascaded Convolutional Classifiers with Thermal-Visual Image Augmentation," International Journal of Image, Graphics and Signal Processing, vol. 17, no. 4, pp. 1–13, 2025. https://doi.org/10.5815/ijigsp.2025.04.01.
8. A. Rahman, "Solar Panel Surface Defect and Dust Detection: Deep Learning Approach," Journal of Imaging, vol. 11, no. 9, article 287, 2025. https://doi.org/10.3390/jimaging11090287.
9. Khatri, A., Khadka, S., Lamichhane, N., & Shrestha, R. (2025). A comprehensive review of infrared thermography and deep learning applications for solar photovoltaic systems. Infrared Physics & Technology, 148, 105878. https://doi.org/10.1016/j.infrared.2025.105878
10. S. R. Mohanty, M. U. Maruf, V. Singh, and Z. Ahmad, "Machine learning approaches for automatic defect detection in photovoltaic systems," Solar Energy, vol. 298, 2025, article 113672. https://doi.org/10.1016/j.solener.2025.113672.
11. A. Hamdi, H. N. Noura, and J. Azar, "Deep Learning-Based Approach to Automated Monitoring of Defects and Soiling on Solar Panels," Future Internet, vol. 17, no. 10, article 433, 2025. https://doi.org/10.3390/fi17100433.
12. R. G., R. Rajalakshmi, D. S. Manish, and R. Narasimhan, "Fusion-Solar-Net for solar panel fault detection," Results in Engineering, vol. 27, 2025, article 106513. https://doi.org/10.1016/j.rineng.2025.106513.

13. Q. Liu, H. Shi, Y. Zhu, L. Chen, M. Liu, W. Cao, and Q. Huang, "An efficient fault diagnosis method combining multi-angle feature expansion and visual image neural networks for solar photovoltaic modules," Energy, vol. 333, 2025, article 137484. https://doi.org/10.1016/j.energy.2025.137484.
14. M. W. Akram and J. Bai, "Defect detection in photovoltaic modules based on image-to-image generation and deep learning," Sustainable Energy Technologies and Assessments, vol. 82, 2025, article 104441. https://doi.org/10.1016/j.seta.2025.104441.
15. J. Li, T. Tong, D. Li, X. Yuan, P. Liu, J. Zhang, X. Zhu, D. Zhao, and H. Fang, "A dynamically adaptive and high-efficiency small object detection network for infrared thermographic images in online monitoring of solar photovoltaic panel defects," Energy, vol. 335, 2025, article 138129. https://doi.org/10.1016/j.energy.2025.138129.