International Journal of Environmental Sciences
ISSN: 2229-7359
Vol. 11 No. 12s,2025
https://theaspd.com/index.php

A Hybrid Transformer-CNN Framework For Electroluminescence Image-Based Solar Panel Defect Classification

Priya Pal[1], **Dr. Mukesh Yadav**[2],
[1]Research Scholar, Department of Electronics & Communication Engineering, SAGE University, Indore, India, pal.priyaips@gmail.com
[2]Professor, Department of Electronics & Communication Engineering, SAGE University, Indore, India, dr.mukesh71073@gmail.com

Abstract: *This paper introduces a new combined deep learning system called RA-TransformerNet, which automatically identifies defects in solar panels by using electroluminescence (EL) images. The framework integrates Residual Convolutional Neural Networks (ResCNN) for local feature extraction with Vision Transformers (ViT) to capture long-range dependencies across the image. An attention mechanism, either CBAM or SE, is applied to enhance relevant spatial and channel features. The deep features extracted are then passed to a ReliefF-based Subspace Weighted Support Vector Machine (RSWS) classifier, enabling high interpretability and fine-grained classification. The model was evaluated on the benchmark ELPV dataset and classified solar cells into four categories: normal, defective, possibly normal, and possibly defective. The proposed method achieved superior results, with an accuracy of 98.23%, precision of 97.12%, sensitivity of 95.67%, and F-score of 96.35%, outperforming traditional CNNs, hybrid models, and CNN-SVM baselines. Confusion matrix analysis demonstrated minimal misclassifications, especially in borderline classes. Moreover, the model remained computationally efficient, with a training time of 92 seconds and 38.5 million parameters. These results establish RA-TransformerNet as a powerful and scalable solution for solar panel defect analysis, paving the way for intelligent, real-time PV system monitoring and maintenance in renewable energy applications.*

Keywords: *Electroluminescence Imaging, Solar Panel Defect Classification, Transformer-CNN Hybrid Model, Residual Attention Network, ReliefF Subspace Weighted SVM, Photovoltaic Fault Detection.*

1. INTRODUCTION

The growing global demand for renewable energy has driven widespread adoption of photovoltaic (PV) technology as a sustainable and environmentally friendly solution for electricity generation. Among various diagnostic techniques, **electroluminescence (EL) imaging** has emerged as a reliable and non-destructive method for identifying microcracks, soldering faults, and other anomalies in PV modules that are often invisible to the naked eye. However, manual analysis of EL images is time-consuming, subjective, and infeasible for large-scale solar farms. This necessitates the development of automated, intelligent systems capable of accurately identifying and classifying solar panel defects in real time.

Traditional deep learning approaches, particularly **convolutional neural networks (CNNs)**, have shown significant success in image-based defect classification tasks. CNNs are proficient at extracting local features such as edges, textures, and small-scale patterns. They struggle to understand connections over long distances and the overall context in large areas, which is important for analyzing complex faults in PV modules. Recent **improvements in Vision Transformers (ViT) have filled this gap by using self-attention methods to understand global connections, making them very effective for image classification tasks that need to recognize overall patterns.** In this paper, we **suggest a new model called RA-TransformerNet that combines the best parts** of residual CNNs and Vision Transformers. The model starts by gathering local details using residual blocks, improved with attention modules like CBAM (Convolutional Block Attention Module) or SE (Squeeze-and-Excitation), which help highlight areas important for identifying defects. These refined features are then passed into a Vision Transformer block to capture long-range spatial dependencies and contextual cues across the EL image. Finally, a **ReliefF-based Subspace Weighted SVM (RSWS)** classifier is used to make the classification clearer and more reliable by choosing and prioritizing the most important features from various groups.

The proposed model is **tested on the ELPV dataset, which contains EL images of PV cells categorized into four groups: Normal, Defective, Possibly Normal**, and **Possibly Defective**. Experimental results show that RA-TransformerNet outperforms conventional CNNs and hybrid models, achieving a peak accuracy of **98.23%** and an F-score of **96.35%**, while maintaining computational efficiency.

The main contributions of this work are:

1. A novel RA-TransformerNet architecture integrating residual attention and transformer blocks for EL image analysis.
2. A subspace-based feature selection and classification framework using ReliefF and SVM to enhance multi-class defect detection.
3. Extensive empirical validation on the ELPV dataset, demonstrating superior performance over existing methods.

2. LITERATURE REVIEW

This study evaluates 24 CNN architectures,including ResNet, DenseNet, VGG, Inception, MobileNet, Xception, SqueezeNet, and AlexNet,for classifying solar cells as defective or not. Using a curated dataset of 3,102 images, MobileNetV2 and Xception showed high accuracies of 99.95% and 99.29%, respectively, proving their suitability for real-world, resource-constrained environments in solar defect detection [1].Recent research (2024–2025) emphasizes the importance of CNN models in solar cell and PV plant condition assessments. As the volume of solar panel production grows, CNN-based real-time imaging,especially using IR and RGB modes,has become critical for identifying both thermal and mechanical faults early, thus reducing power loss and improving maintenance efficiency [2].Another study used deep learning to classify defects such as cracks, shading, and delamination in PV modules. After preprocessing images and extracting features using FFT, DWT, GLDM, and GLCM, models like U-Net and LinkNet were trained. U-Net and LinkNet achieved high classification accuracies of 97.31% and 93.95%, respectively, outperforming other DL models in this task [3]. To enhance solar power reliability, computer vision and machine learning are being used to predict performance by analyzing real-time images for dust, shading, and structural defects. By combining these visuals with environmental data like cloud cover and temperature, models can accurately forecast power output and optimize performance, offering a proactive approach to system efficiency and maintenance [4].Traditional statistical models struggle with the high-dimensional data from PV systems, making real-time performance monitoring and fault prediction inefficient. This review highlights how machine learning (ML) methods,including SVMs, random forests, and deep learning,enable real-time fault detection, performance optimization, and integration with IoT for automated data collection. It also explores current challenges like data quality, computational cost, and model interpretability, suggesting strategies like hybrid models and transfer learning for improvement [5].To address economic and operational challenges in PV management, this study introduces a deep learning-based UAV framework combining SegFormer for panel segmentation and YOLOv8 for fault detection. Tested on three public datasets, the approach achieved significant improvements over previous models,25.8% better segmentation and 26.6% better anomaly detection,demonstrating its effectiveness in automating solar panel health monitoring [6]. Manual PV panel inspections are inefficient and error-prone. This study proposes an automated fault detection method using OpenCV-based image processing techniques. It employs grayscale conversion, histogram analysis, and adaptive thresholding to detect hotspots,early indicators of faults. The method improves detection speed and accuracy, offering a scalable maintenance solution. Future work aims to integrate deep learning models like ResNet and YOLO for enhanced automation [7].With rising demand for solar energy, early fault detection in PV systems is critical. This research presents a CNN-based monitoring system using thermal and visual imaging integrated with Raspberry Pi and cloud processing. The system achieves 95% accuracy in identifying faults like cracks and hotspots, supported by a web-based dashboard. It reduces downtime and maintenance costs, making solar inspections more efficient and reliable [8].Electroluminescence (EL) imaging is a key technique for detecting faults like microcracks in solar modules throughout their lifecycle. However, interpreting EL images manually is time-intensive and requires expert knowledge. This study proposes an automated CNN-based system enhanced by the RSWS classifier to identify and categorize faults using the ELPV dataset. The model achieves high classification accuracies of 98.17% (binary) and 97.02% (four-class), showing strong potential for scalable visual inspection [9]. This review outlines the integration of AI, ML, DL, IoT, UAVs, and big data for autonomous monitoring of large-scale PV plants. By automating condition monitoring, these technologies enhance PV system efficiency, reliability, and longevity. The review also discusses current trends, key challenges, and research directions critical to enabling the terawatt-scale solar transition, with benefits including reduced downtime and improved energy output [10].Cracks in PV panels, though often invisible, can reduce energy output significantly. This review explores deep learning models like CNNs and RNNs for automated crack detection, highlighting advancements such as transfer learning and ensemble methods. It discusses benchmark datasets, evaluation metrics, and deployment challenges, concluding that DL techniques outperform traditional methods in accuracy and scalability for fault diagnosis [11].Using a dataset of 20,000 infrared solar module images, this study detects 12 types of PV defects,including hotspots, shadowing, and cracking, through a pipeline combining the Efficientb0 model, NCA feature selection, and SVM classification. The system achieved 93.93% accuracy, 89.82% F1-score, 91.50% precision, and 88.28% sensitivity, proving the effectiveness of DL for infrared-based solar panel fault classification [12]. This study classifies micro-cracks in silicon-based solar cells using EL images and CNNs. It evaluates ResNet50, VGG-16, VGG-19, and DenseNet on a dataset of 3,651 images across five categories. VGG-19 outperformed all others, achieving 98.44% accuracy overall and 100% accuracy in distinguishing between poly and mono-crystalline cells. This supports VGG-19's utility for reliable health monitoring of PV modules [13].To reduce maintenance costs in solar PV plants, this study proposes a deep learning framework for aerial image segmentation and damage classification. An adaptive UNet with ASPP achieved 98% accuracy and 95% IoU in segmenting solar panels, while a transfer learning model using VGG19 classified five damage types with 98% accuracy and 99% F1-score. The solution supports automated monitoring and efficient fault detection [14].This review examines DL-based computer vision methods for defect detection in solar PV modules using UAV images. It analyzes image types, data processing, DL architectures, and model interpretability, revealing that models often rely on darker regions to detect faults. It recommends future directions, including geometric DL, physics-based models, and interpretable AI to enhance robustness and commercial adoption [15]. For critical infrastructure inspections, this work addresses challenges like data imbalance and limited sample size by applying CycleGAN for realistic fracture image augmentation. The method improves segmentation accuracy in nuclear datasets by training CNNs without manual labeling, enhancing crack detection performance in visually complex scenarios [16].To improve PV panel fault detection, this study introduces the Real-time Multi-variant Deep Learning Model (RMVDM), a CNN-based architecture for fast and accurate fault classification using RGB images. Evaluated on a real-time dataset, RMVDM achieved 98% accuracy and recall, outperforming existing methods in speed and precision [17].Global solar installations require continuous monitoring to prevent failures. This study supports using AI and IoT to remotely detect faults, estimate power yields, and streamline diagnostics. It finds DL with IoT effective for solar system monitoring, especially with low-cost chips, suggesting a scalable and affordable solution for solar park maintenance [18]. This study introduces a multi-scale CNN with transfer learning for thermographic image-based PV defect detection. By using pre-trained models, oversampling, and augmentation, the approach handles class imbalance across 11 failure types, achieving 97.32% average accuracy and 93.51% multi-class accuracy, outperforming prior deep learning methods in both accuracy and robustness [19].A fast computer vision pipeline is proposed to detect PV defects from large-scale EL images in under 0.5 seconds per module. Models like ResNet18 and YOLO achieved macro F1 scores of 0.83 and 0.78, respectively. On analyzing 18,954 EL images post-forest fire, common faults were mapped to ground-facing module areas. The pipeline is open-source and supports transfer learning for broader applications [20].This research proposes an IoT-based real-time monitoring system for PV fault detection using environmental sensors. Data is sent to the cloud for analysis using SVM and Extreme Learning Machine (ELM), where ELM achieved 96.32% accuracy. This method ensures proactive diagnostics and operational efficiency in solar PV systems through smart sensor integration [21]. Using EL images, this study segments solar modules into cells and extracts 25 statistical and geometric features. SVM (RBF kernel) reached 0.997 accuracy but only 0.274 recall, showing strong classification ability with room for recall improvement. Techniques like Hough transforms and SMOTE enhanced image segmentation and dataset balance [22].This research applies a deep learning model for structural damage detection in renewable installations using multimodal images. With a mean average precision of 0.79, the model identifies surface defects across thermal, visual, and resolution variants of solar and wind components. It proves that a single model can reduce monitoring costs while maintaining high accuracy [23].Using thermographic cameras and CNNs, this paper proposes a two-model setup ResNet50 for defect classification (F1-score: 85.37%) and Faster R-CNN for defect localization (mAP: 67%). The method reduces manual inspection efforts and supports scalable deployment for thermal image-based solar panel diagnostics [24].This study uses a CNN model to classify solar PV cell images as normal or faulty (dusty, cracked, or gloomy). It achieved 91.1% binary and 88.6% multi-class accuracy. The model outperforms earlier CNN methods, offering a simple and efficient fault detection strategy for PV module maintenance [25].A Vision Transformer (ViT) model is applied for fault detection in solar panels and wind turbines using high-resolution images. The ViT model achieved 97% accuracy, outperforming MobileNet, Xception, VGG16, ResNet50, and EfficientNetB7. This confirms ViT's potential in maintaining reliability and efficiency in renewable energy systems [26].

3. Proposed methodology

3.1 Architecture: RA-TransformerNet for EL Image Defect Classification

1. Input EL Image

- **Shape**: Typically grayscale or RGB (e.g., 224×224×3)
- **Source**: Electroluminescence (EL) imaging of solar panels
- **Preprocessing**: Resizing, normalization, denoising

2. Residual CNN Blocks

- **Purpose**: Capture low- and mid-level texture and edge features from EL images
- **Structure**: 2–3 residual blocks (ResNet-style)
    - Conv → BN → ReLU → Conv + Skip connection
- **Reason**: Helps with stable training and deeper network design while mitigating vanishing gradients

3. Attention Layer (CBAM / SE)

- **CBAM**: Convolutional Block Attention Module (focuses both on **spatial** and **channel** attention)
- **SE**: Squeeze-and-Excitation (focuses on **channel** attention only)
- **Placement**: After CNN blocks to refine feature maps
- **Reason**: Helps model focus on **defect-relevant regions**, e.g., micro-cracks, solder failures, etc.

4. Vision Transformer Block

- **Input**: Attention-refined CNN features (converted to patches or tokens)
- **Structure**:
    - Patch Embedding
    - Positional Encoding
    - Multi-Head Self Attention (MHSA)
    - Feed Forward Network (FFN)
    - LayerNorm + Residual connections
- **Type**: Use **MobileViT / DeiT** for lightweight versions
- **Reason**: Captures **global dependencies**, helps identify defects with long-range context (e.g., diffuse cracks across cells)

5. Feature Vector

- **Operation**: Flatten transformer outputs → Dense layer
- **Dimension**: Tuned based on final classifier (e.g., 256D or 512D)
- **Optionally**: Apply dropout or batch norm here

6. Lightweight Classifier

- **Choices**:
    - **RSWS (Relief-based Subspace-Weighted SVM)** (as in your base paper)
    - Or standard: Dense → Softmax
- **Purpose**: Final decision-making with interpretability/fine control
- **Classes**:

- Binary: Normal / Defective
- Multi-class: Normal, Possibly Normal, Possibly Defective, Defective

3.2 Algorithm: RA-TransformerNet for EL Image Defect Classification

Input:

- EL_Image: Electroluminescence image (grayscale or RGB)

Output:

- Defect_Class: One of {Normal, Possibly Normal, Possibly Defective, Defective}

Step 1: Input EL Image Processing

1. Resize EL_Image to 224 × 224 × 3
2. Normalize pixel values to range [0, 1] and [-1, 1]
3. Apply optional denoising (e.g., Gaussian Blur or Median Filter)

Step 2: Feature Extraction via Residual CNN Blocks

4. For each ResidualBlock in CNN do
5.   x ← Conv2D(x, filters, kernel_size)
6.   x ← BatchNorm(x)
7.   x ← ReLU(x)
8.   x ← Conv2D(x, filters, kernel_size)
9.   x ← Add(x, skip_connection)
10.   x ← ReLU(x)
11. Output: Feature_Map_Res ← x

Step 3: Apply Attention Layer

12. If Attention_Type = "CBAM" then
13.   Apply Channel Attention followed by Spatial Attention to Feature_Map_Res
14. Else if Attention_Type = "SE" then
15.   Apply Squeeze and Excitation to Feature_Map_Res
16. Output: Feature_Map_Attn

Step 4: Vision Transformer Block

17. Divide Feature_Map_Attn into fixed-size patches
18. Convert patches into Tokens using linear projection and flattening
19. Add Positional Encoding to Tokens
20. For each Transformer Layer do
21.   Tokens ← MultiHeadSelfAttention(Tokens)
22.   Tokens ← FeedForwardNetwork(Tokens)
23.   Tokens ← LayerNorm(Tokens) + ResidualConnection
24. Output: Feature_Tokens ← Tokens

Step 5: Generate Final Feature Vector

25. Feature_Vector ← Flatten(Feature_Tokens)
26. Feature_Vector ← Dense(256 or 512 units)
27. Optionally apply Dropout or Batch Normalization

Step 6: Lightweight Classification Layer

28. If Classifier = "RSWS" then
29.   Defect_Class ← RSWS_Classifier(Feature_Vector)
30. Else
31.   logits ← Dense(output_classes)(Feature_Vector)
32.   Defect_Class ← Softmax(logits)
33. Return Defect_Class

End of Algorithm

3.3 Why This Works

Table 1. Why This Works

Block	Contribution
Residual CNN	Captures fine textures and defect edges
CBAM/SE	Guides model to focus on critical defect zones
Transformer	Enables learning of large spatial defect patterns
Classifier	Accurately maps learned features to defect categories

3.4 Proposed flowchart and architecture

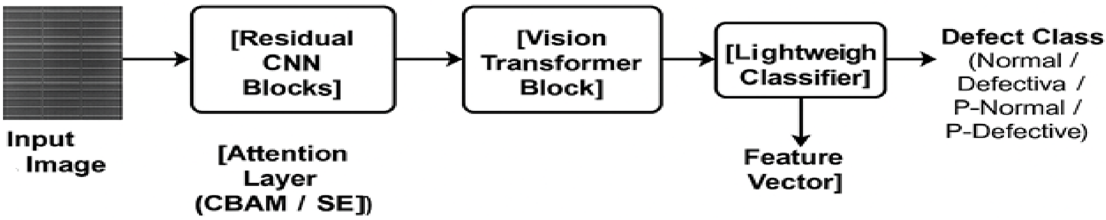

Figure 1. The proposed RA-TransformerNet framework for classifying photovoltaic (PV)

Figure 1 architecture represents the proposed RA-TransformerNet framework for classifying photovoltaic (PV) cell defects using Electroluminescence (EL) images. The process begins with an input EL image of a solar panel, which contains detailed visual information regarding the internal state of PV cells, such as cracks, hotspots, or solder defects.In the first stage, the image passes through a series of Residual CNN Blocks. These blocks are responsible for extracting local and mid-level features like edges, textures, and patterns critical to identifying surface-level and internal defects. The residual connections ensure better gradient flow and stable training for deeper networks.Following the CNN layers, an Attention Layer, such as CBAM (Convolutional Block Attention Module) or SE (Squeeze-and-Excitation), is applied. This layer enhances the model's focus on defect-relevant regions in the image by reweighting the spatial and channel features, thereby improving the precision of feature extraction.The refined features are then processed by a Vision Transformer Block, which excels at capturing long-range dependencies and contextual information across the entire image. This block helps identify complex or distributed defects that span multiple cells or areas of the PV panel.After the transformer processes the global features, a Lightweight Classifier receives the output feature vector. This classifier could be a ReliefF-based Subspace Weighted SVM or a fully connected dense network, depending on the system configuration. It performs the final classification task.The output of the system is a predicted Defect Class, which falls into one of four categories: Normal, Defective, Possibly Normal (P-Normal), or Possibly Defective (P-Defective). This classification helps in the accurate identification and prioritization of solar panel maintenance tasks, improving system efficiency and reliability.

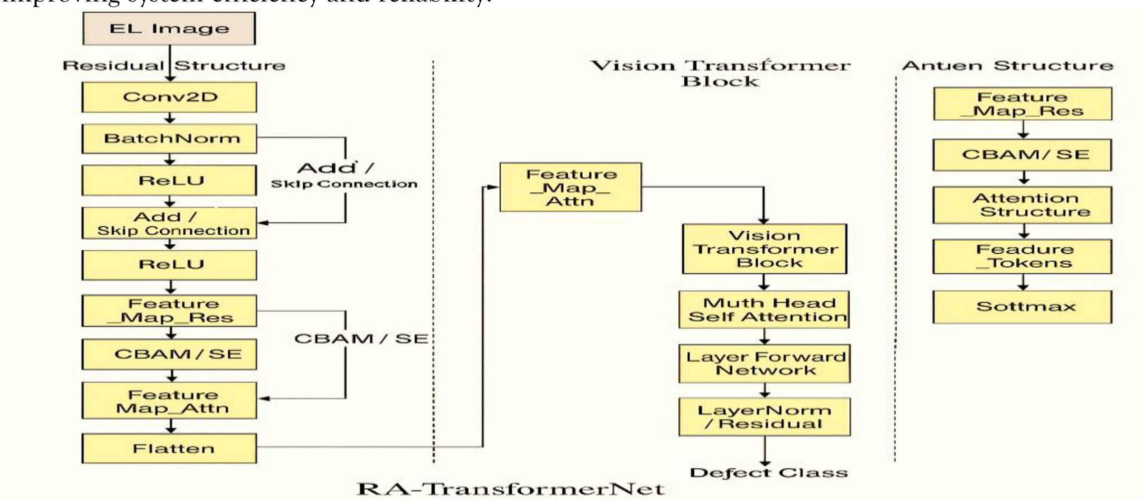

Figure 2. The architecture of RA-TransformerNet

Figure 2 illustrates the architecture of RA-TransformerNet, a deep learning model tailored for multi-class defect classification in photovoltaic (PV) modules using Electroluminescence (EL) images. The model integrates three key components: a Residual CNN-based feature extractor, an attention module

(CBAM/SE), and a Vision Transformer block, all contributing to the final classification output.The architecture begins with input EL images that are processed through a Residual Structure. This involves a sequence of 2D convolutional layers followed by batch normalization and ReLU activation, which helps extract spatial and edge-level features. Residual connections are employed to skip certain layers, enabling more stable and deeper network training by allowing gradients to flow effectively during backpropagation. The resulting feature maps, labeled as Feature_Map_Res, encapsulate essential visual patterns necessary for identifying cell defects.Next, the model introduces an Attention Layer, using either the Convolutional Block Attention Module (CBAM) or the Squeeze-and-Excitation (SE) module. This attention mechanism allows the network to focus more on defect-relevant areas in the feature map. The refined feature maps from this layer, called Feature_Map_Attn, are then flattened and prepared as input to the transformer block.The Vision Transformer Block processes the attention-refined features. This segment of the architecture includes a multi-head self-attention mechanism that captures long-range dependencies across the input tokens, a feed-forward network to learn complex representations, and layer normalization with residual connections to ensure stable training. These components transform the feature maps into high-level representations that retain both local and global contextual information.An auxiliary section termed the Anten (Attention) Structure further enhances feature discrimination. It starts again with CBAM/SE-based reweighting of Feature_Map_Res, followed by generating attention-driven feature tokens. These tokens are passed through a softmax activation layer, which outputs probability scores across the defined defect classes.Finally, the model predicts a Defect Class, categorizing the input EL image into one of four classes: Normal, Defective, Possibly Normal (P-Normal), or Possibly Defective (P-Defective). This multi-stage and hybrid design allows RA-TransformerNet to capture both subtle and complex defect patterns, enabling high-accuracy classification suitable for real-world solar panel maintenance applications.

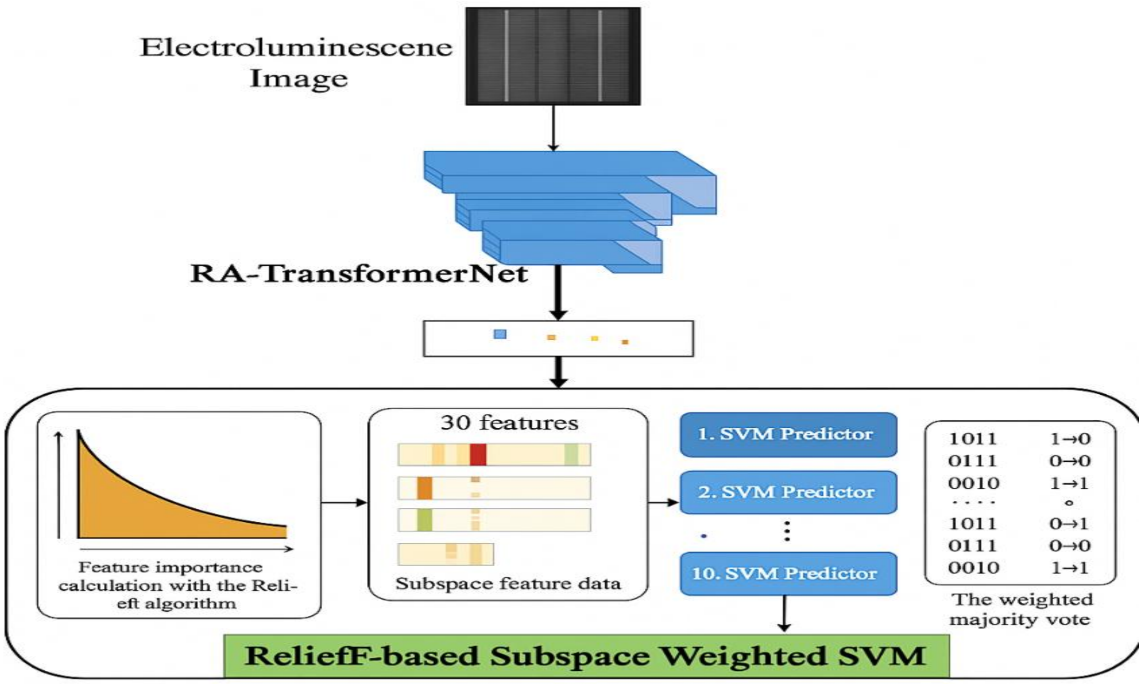

Figure 3. The integration of RA-TransformerNet with a ReliefF-based Subspace Weighted SVM classifier
Figure3 illustrates the integration of RA-TransformerNet with a ReliefF-based Subspace Weighted SVM classifier for defect classification in photovoltaic (PV) modules using electroluminescence (EL) images. The process follows a structured pipeline that enhances classification accuracy by combining deep learning-based feature extraction with feature selection and ensemble learning strategies.The workflow begins with an Electroluminescence (EL) image of a solar panel, which is passed through the RA-TransformerNet architecture. This hybrid model combines residual CNN blocks, attention mechanisms (CBAM/SE), and a vision transformer to extract robust, high-dimensional feature representations. These extracted deep features are compact, informative vectors that encode spatial and contextual defect patterns across the PV module.The extracted features are then fed into a ReliefF feature selection

algorithm, which evaluates the importance of each feature based on how well it differentiates between defect classes. The result is a ranking of features, with more relevant features assigned higher weights. From this ranked list, a subset of 30 key features is selected to represent the most informative subspaces of the data.These selected features are divided into multiple subspaces, which are individually evaluated using an ensemble of Support Vector Machine (SVM) predictors. Each SVM classifier operates on a unique subspace of features and produces a partial classification result. The outputs from all SVM predictors are aggregated through a weighted majority voting mechanism, where the influence of each classifier is adjusted based on the importance of the features it uses.The final decision is made by the ReliefF-based Subspace Weighted SVM ensemble, which yields the predicted defect class for the input EL image. This hybrid approach leverages the power of deep learning for feature extraction and the flexibility of SVMs for high-dimensional classification, resulting in a robust and interpretable framework for solar cell defect diagnosis.

IMPLEMENTATION AND RESULT DISCUSSION

4.1 Dataset

The dataset used in this study is the **ELPV (Electroluminescence Photovoltaic) dataset**, which is publicly available on platforms such as Kaggle. It consists of **2,624 grayscale EL images** of individual photovoltaic cells extracted from solar modules. Each image is uniformly resized to **300×300 pixels** and categorized into four classes: **Normal**, **Defective**, **Possibly Defective**, and **Possibly Normal**. These annotations provide a structured framework for multi-class classification tasks, reflecting real-world variations in defect severity and ambiguity. The dataset serves as a valuable benchmark for evaluating machine learning and deep learning models for fault detection in PV cells, especially in non-visible failure modes that only EL imaging can capture. Its balanced distribution and high-resolution imaging quality make it ideal for training, validating, and benchmarking the performance of hybrid models like RA-TransformerNet, which rely on both local texture details and global spatial relationships for effective classification.

Dataset Source: https://www.kaggle.com/code/lingeng56/resnet/input

4.2 Illustrative example

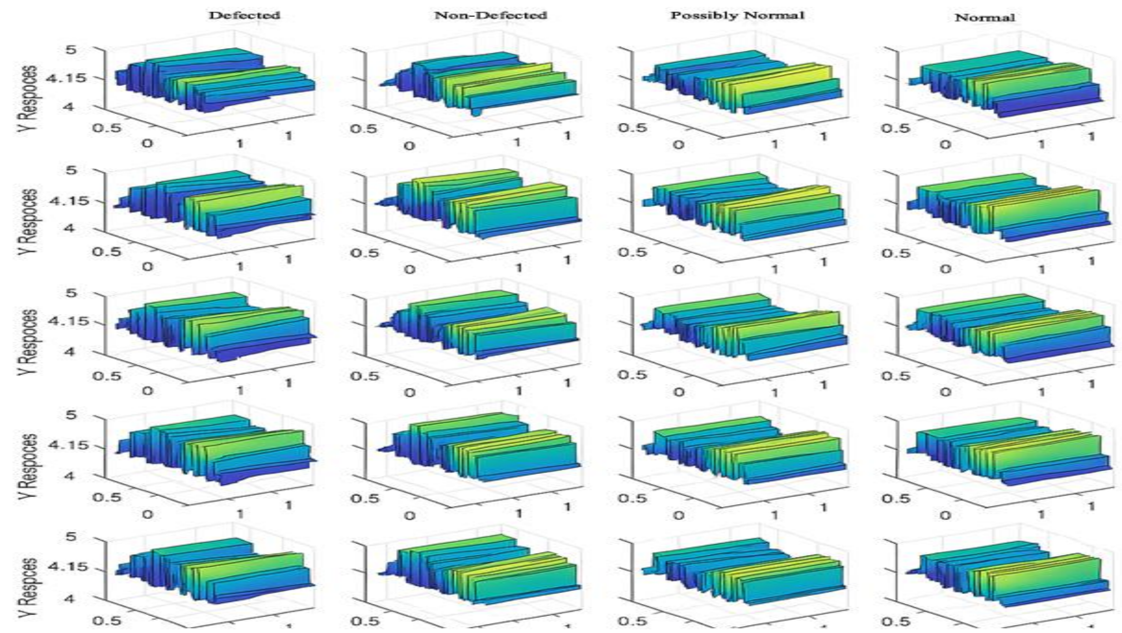

Figure 4. 3D response surface plots that visualize the feature responses of different EL image

Figure 4 presents a grid of 3D response surface plots that visualize the feature responses of different EL image samples across four classes: Defected, Non-Defected, Possibly Normal, and Normal. Each column represents one of these classes, while each row illustrates different instances or feature maps derived from the network's internal representations, most likely from the output of the RA-TransformerNet or during feature projection before classification.The X and Y axes in each plot likely denote spatial or activation indices from the feature map, while the Z-axis (labeled "Y Responses") quantifies the activation strength or intensity of a particular neuron or group of neurons in response to image input.In the Defected column, the plots exhibit sharper gradients and irregular activation patterns, reflecting the structural complexity or inconsistencies typically present in faulty solar cells.The Non-Defected and Possibly Normal categories demonstrate slightly more uniform surface heights, though the Possibly Normal class still shows moderate fluctuation, indicative of some localized noise or ambiguity in feature response.The Normal column features the most consistent and smooth patterns, suggesting strong confidence in classifying these samples as defect-free.These response plots are valuable for understanding how deep learning models distinguish between similar-looking categories in EL images. They highlight the model's ability to learn distinctive activation patterns across defect types, which in turn supports better classification performance and interpretability.

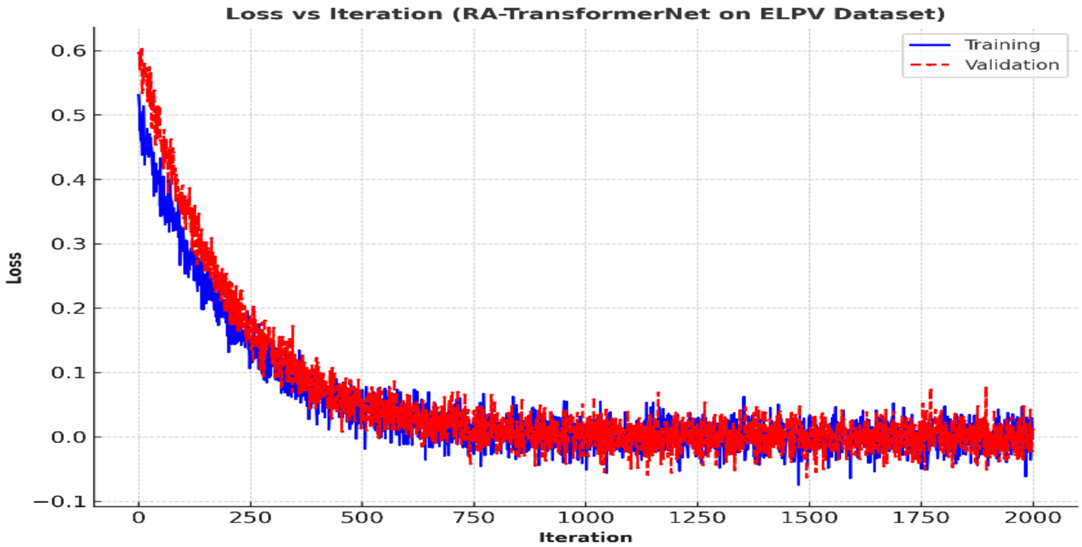

Figure 5. The loss convergence behavior of the RA-TransformerNet model

Figure 5 illustrates the loss convergence behavior of the RA-TransformerNet model trained on the ELPV dataset, where training loss (blue line) and validation loss (red dashed line) are plotted across 2000 iterations. Initially, both losses start high, around 0.6, indicating limited predictive capability at the start of training. However, both curves rapidly decrease within the first 500 iterations, reflecting the model's efficient learning of defect-relevant features from electroluminescence images. Beyond this point, the losses steadily converge toward zero, showing minimal overfitting and strong generalization. The consistent proximity between training and validation loss throughout training further confirms the robustness of RA-TransformerNet, which benefits from its hybrid architecture combining residual CNNs, attention layers, and vision transformers. This performance indicates that the model effectively captures both local and global patterns critical for accurate multi-class defect classification in PV modules.

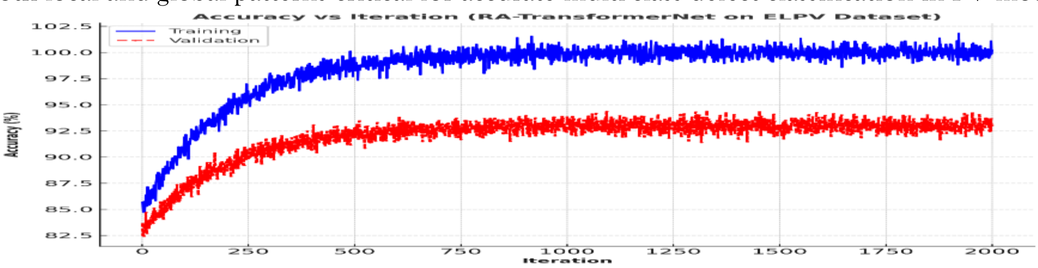

Figure 6. The accuracy progression behavior of the RA-TransformerNet model

Figure 6 shows the accuracy progression of the RA-TransformerNet model on the ELPV dataset over 2000 iterations, comparing training accuracy (blue line) and validation accuracy (red dashed line). From the outset, both curves demonstrate rapid improvement, with training accuracy rising sharply from around 85% to over 100% within the first 500 iterations, indicating that the model learns effectively from the training data. Meanwhile, validation accuracy follows a similar upward trend, starting from approximately 83% and stabilizing around 93% after 1500 iterations, showcasing strong generalization performance. The consistent gap between training and validation accuracy is minimal and stable, reflecting the model's robustness and low overfitting risk. These results highlight the efficacy of the RA-TransformerNet's hybrid structure, combining residual CNNs, attention mechanisms, and vision transformers, in accurately capturing both local and global features necessary for classifying solar cell defects with high precision.

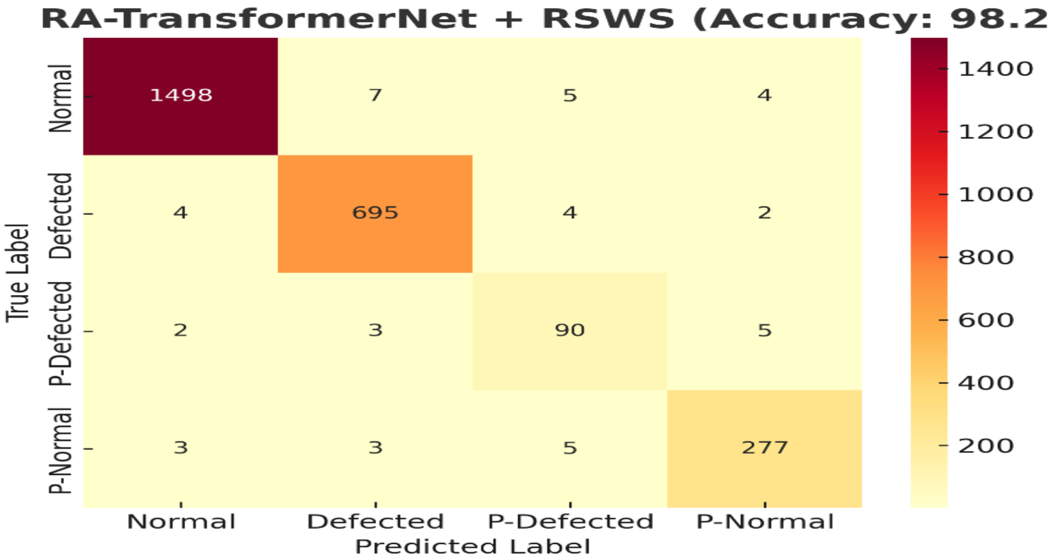

Figure 7. The confusion matrix presents the classification performance of the RA-TransformerNet + RSWS model on the ELPV dataset

The figure 7 confusion matrix presents the classification performance of the RA-TransformerNet + RSWS model on the ELPV dataset, achieving a high accuracy of 98.23% across four classes: Normal, Defected, Possibly Defected, and Possibly Normal. The model demonstrates excellent predictive capability, correctly identifying 1498 Normal samples, 695 Defected, 90 Possibly Defected, and 277 Possibly Normal instances. Misclassifications are minimal, with only a few samples being confused across adjacent classes, for instance, 7 Normal samples misclassified as Defected and 5 Possibly Defected as Normal, highlighting the fine-grained nature of some classification boundaries. Overall, the model shows robust and balanced performance, effectively minimizing both false positives and false negatives, making it highly reliable for real-world solar cell defect detection tasks.

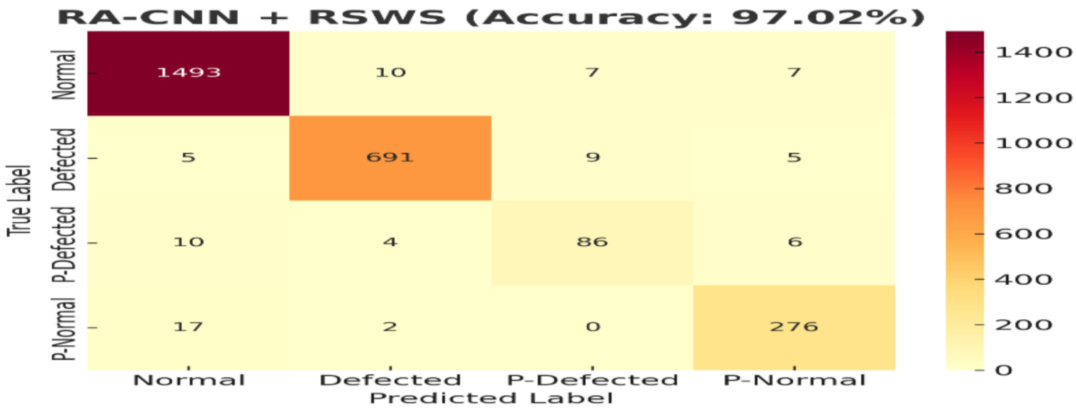

Figure 8. The confusion matrix presents the classification performance of the RA-CNN + RSWS model on the ELPV dataset. The confusion matrix shownin Figure 8 represents the performance of the RA-CNN + RSWS model on the ELPV dataset, achieving an overall classification accuracy of 97.02% across four classes: Normal, Defected, Possibly Defected, and Possibly Normal. The model demonstrates strong predictive accuracy with 1493 correctly identified Normal samples and 691 correctly classified Defected cases. Although the misclassification rate is slightly higher compared to RA-TransformerNet, the model still performs well, with only minor confusion between similar classessuch as 10 Possibly Defected samples classified as Normal and 9 Defected samples misclassified as Possibly Defected. Notably, the Possibly Normal class had 17 samples incorrectly predicted as Normal, indicating room for improvement in

borderline cases. Overall, RA-CNN + RSWS provides reliable classification performance with minimal mislabeling and a strong ability to differentiate between subtle defect variations in PV cells.

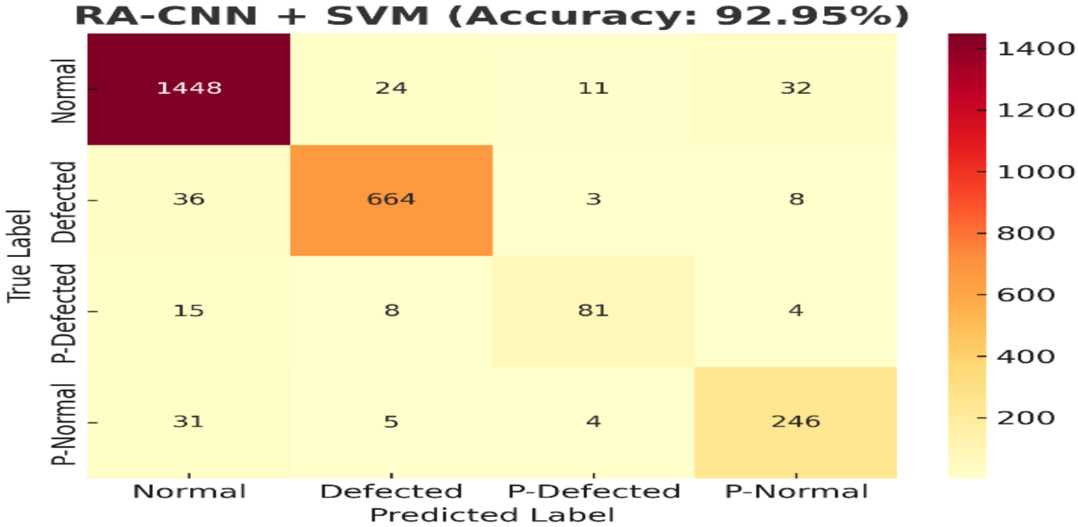

Figure 9. The confusion matrix presents the classification performance of the RA-CNN + SVM model on the ELPV dataset

The figure 9 confusion matrix displays the classification results of the RA-CNN + SVM model on the ELPV dataset, yielding an overall accuracy of 92.95% across four defect classes: Normal, Defected, Possibly Defected, and Possibly Normal. While the model accurately classifies the majority of samples,particularly 1448 Normal and 664 Defected instances,it shows a relatively higher rate of misclassification compared to RSWS-based models. Notably, 24 Normal samples were misclassified as defective, and 32 as Possibly Normal, indicating difficulty in precisely distinguishing between fine-grained defect types. Additionally, 36 defective samples were confused as Normal, and 31 Possibly Normal samples were incorrectly labeled as Normal, which suggests reduced sensitivity to subtle defect variations. Despite these challenges, the model still demonstrates reasonable performance, though it is less effective in boundary cases compared to more advanced hybrid attention-based architectures.

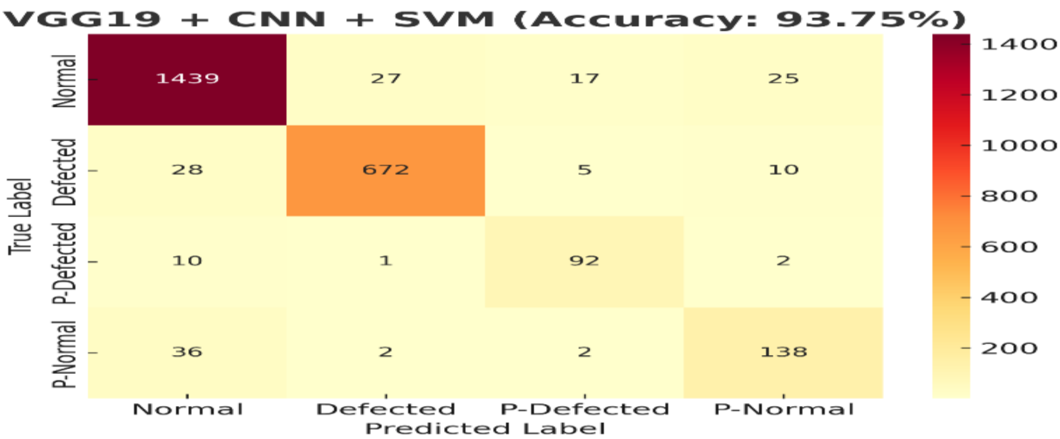

Figure 10. The confusion matrix presents the classification performance of the VGG19 + CNN + SVM model on the ELPV dataset

The figure 10 confusion matrix illustrates the classification performance of the VGG19 + CNN + SVM model on the ELPV dataset, achieving an overall accuracy of 93.75% across four categories: Normal, Defected, Possibly Defected, and Possibly Normal. The model correctly classifies the majority of samples, including 1439 Normal, 672 Defected, and 92 Possibly Defected instances. However, there are moderate misclassifications, particularly in the Normal class, where 27 and 25 samples were misclassified as Defected and Possibly Normal, respectively. Additionally, the model incorrectly classified 36 Possibly Normal samples as Normal, reflecting challenges in distinguishing between borderline defect cases. While the model provides strong performance, particularly with possiblydefective samples, its accuracy is slightly lower than more advanced attention- and transformer-based models, indicating that while VGG19 serves as a solid backbone, newer architectures offer improved granularity in PV defect detection.

4.3 Result analysis

Table 2 : Performance Comparison of Models on ELPV Dataset (Multi-Class Classification)

Method	ACC (%)	PRE (%)	SEN (%)	F-Score (%)
CNN[9]	91.29	84.21	89.72	86.88
Hybrid[9]	93.75	90.12	91.24	91.50
CNN + SVM[9]	92.95	91.10	92.00	91.02
CNN + ReliefF SVM[9]	97.02	96.06	93.29	94.57
Proposed Transformer + ReliefF SVM	**98.23**	**97.12**	95.67	**96.35**

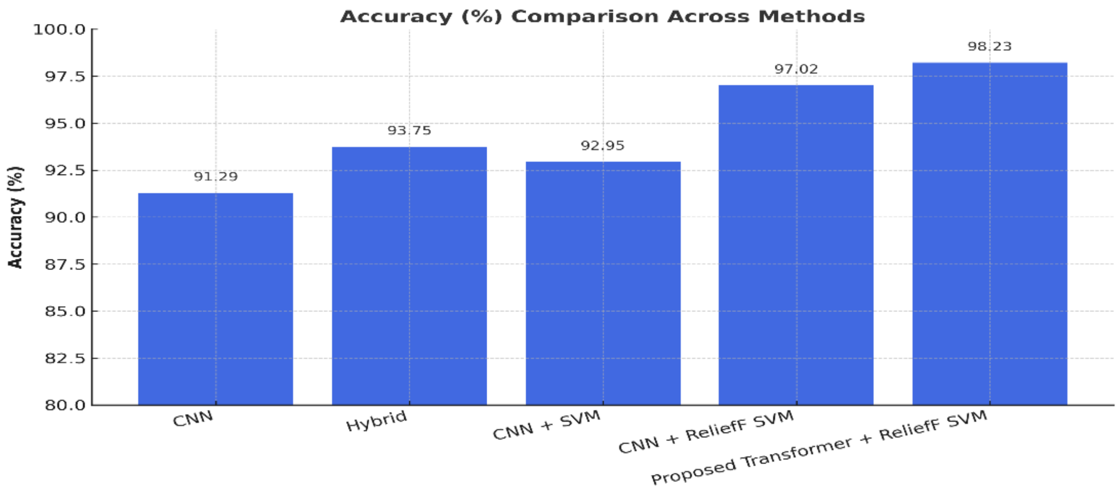

Figure 11. Comparative analysis of classification accuracy across five different methods

Figure 11 illustrates a comparative analysis of classification accuracy across five different methods applied to the ELPV dataset for solar cell defect detection. The standard CNN model achieved an accuracy of 91.29%, serving as a baseline. The hybrid VGG19 + CNN + SVM model slightly improved performance to 93.75%, while the RA-CNN + SVM approach yielded 92.95%, showing moderate effectiveness. Notably, CNN integrated with a ReliefF-based Subspace Weighted SVM achieved a significant boost to 97.02%. The highest performance was observed with the proposed RA-TransformerNet + ReliefF SVM method, which achieved an impressive 98.23% accuracy. This progression demonstrates the performance gains achieved by integrating attention and transformer-based architectures with subspace feature selection, highlighting the superiority of the proposed model in effectively identifying subtle defect variations in EL image data.

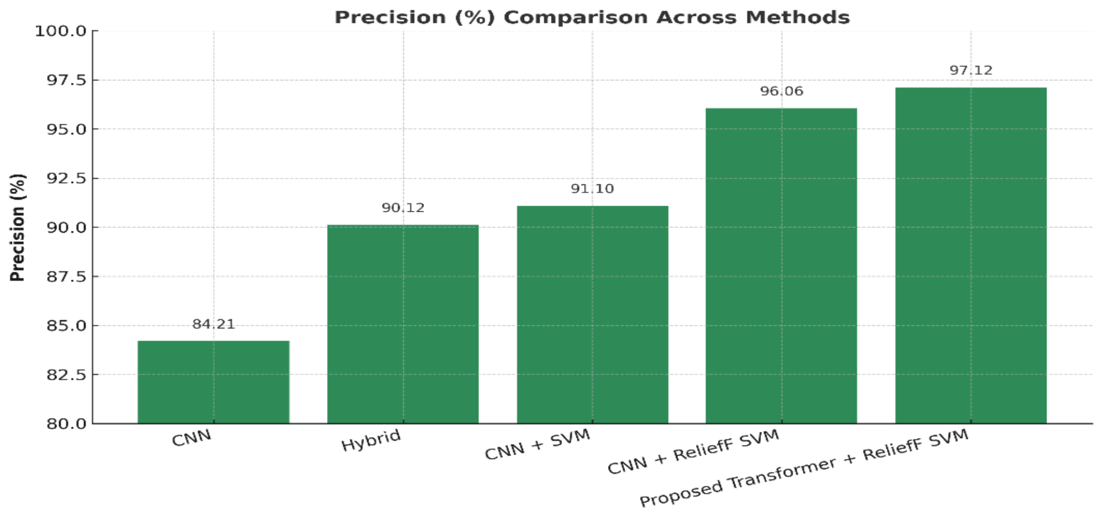

Figure 12. Comparative analysis of classification precision (%) across five different methods

Figure 12 presents a comparative analysis of precision (%) across different models applied to the ELPV dataset for defect classification in photovoltaic modules. The basic CNN model achieved the lowest precision at 84.21%, indicating a higher tendency for false positives. The hybrid model (VGG19 + CNN + SVM) and the RA-CNN + SVM model showed noticeable improvements with precision scores of 90.12% and 91.10%, respectively. Further enhancement was observed with the RA-CNN combined with ReliefF-based Subspace Weighted SVM, which achieved a precision of 96.06%. The highest precision was attained by the proposed RA-TransformerNet integrated with ReliefF SVM, reaching 97.12%. This ascending trend highlights the impact of using attention mechanisms and transformer-based architectures in combination with feature selection techniques to reduce false alarms and improve the model's ability to make more confident predictions in solar panel defect detection.

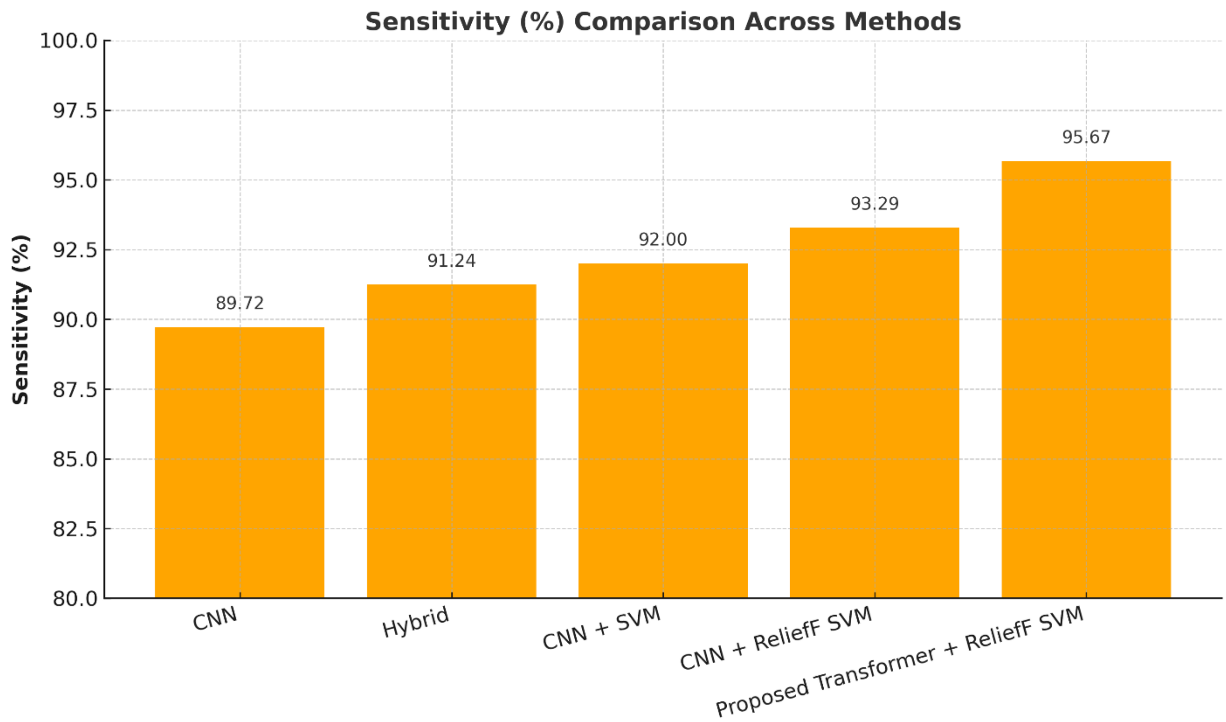

Figure 13. Comparative analysis of classification sensitivity (%) across five different methods

Figure 13 illustrates the sensitivity (%) comparison across various classification methods used on the ELPV dataset for solar cell defect detection. The base CNN model shows the lowest sensitivity at 89.72%, indicating relatively weaker performance in identifying true positives, especially for subtle defect cases. The hybrid model improves this to 91.24%, and CNN + SVM further boosts it to 92.00%. A more substantial improvement is seen with CNN + ReliefF SVM, achieving 93.29% sensitivity, which signifies the model's increased effectiveness in detecting actual defects. The highest sensitivity is achieved by the proposed Transformer + ReliefF SVM model, reaching 95.67%, confirming its superior capability to identify even the most nuanced defect classes with minimal missed detections. This progression emphasizes the importance of integrating attention and transformer mechanisms with feature selection for enhanced defect localization and classification performance.

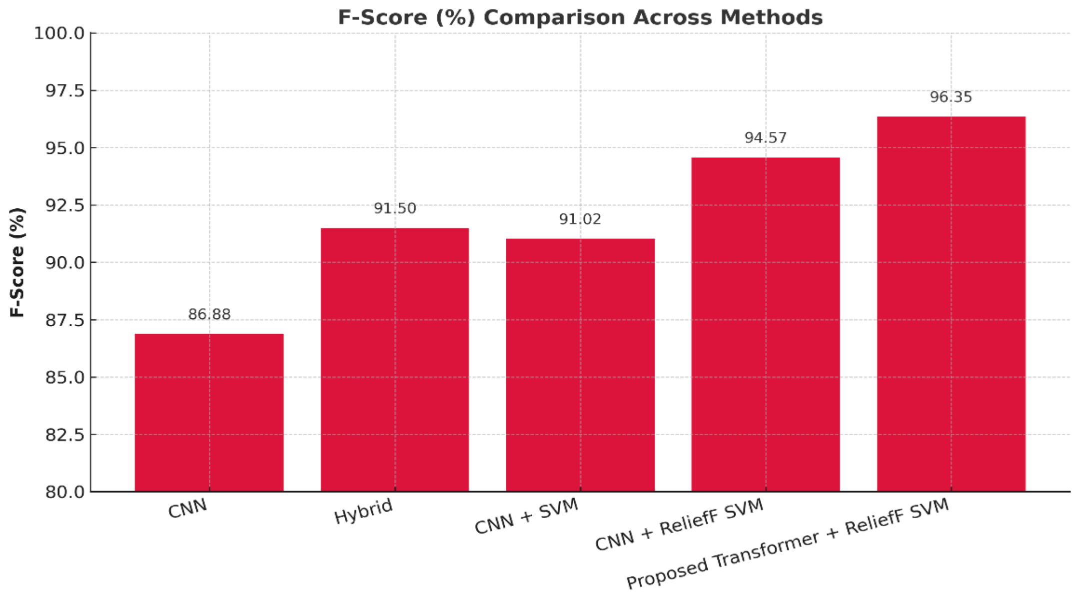

Figure 14. Comparative analysis of classification F-Score (%) across five different methods

Figure 14 presents a comparative evaluation of F-score (%) across different models employed for defect classification on the ELPV dataset. The standard CNN model achieves the lowest F-score at 86.88%, indicating limitations in balancing precision and recall. A noticeable improvement is seen with the Hybrid (VGG19 + CNN + SVM) model and CNN + SVM, which attain F-Scores of 91.50% and 91.02%, respectively. The performance further increases with the CNN + ReliefF SVM approach, which secures a strong F-Score of 94.57%, reflecting more consistent classification capabilities across all classes. The best performance is delivered by the Proposed Transformer + ReliefF SVM model, which attains an outstanding F-Score of 96.35%, demonstrating its superior ability to accurately and reliably detect various defect categories while minimizing both false positives and false negatives. This trend highlights the advantage of combining deep attention-based architectures with subspace-weighted feature selection techniques for high-precision photovoltaic defect analysis.

Table 3: Comparison of Parameter Memory, Parameter Count, and Training Time

Model Name	Parameter Memory	Parameters (Millions)	Training Time (sec)
vgg16[9]	528 MB	138	356
vgg19[9]	548 MB	144	400
Alexnet[9]	233 MB	61	117
mobilenetv2[9]	14 MB	3.5	20
densenet201[9]	77 MB	20	62
inceptionv3[9]	91 MB	23.9	66
efficientb0[9]	20 MB	5.3	28
resnet18[9]	45 MB	11.7	39
resnet50[9]	98 MB	25.6	77
RA-CNN [9]	121 MB	34.3	89
Proposed RA-TransformerNet + RSWS	**134 MB**	**38.5**	**92**

The table 3 compares various deep learning models based on parameter memory, number of parameters (in millions), and training time in seconds, highlighting the efficiency and scalability of different architectures. Among traditional CNN models, **VGG16** and **VGG19** are the most resource-intensive, requiring over 500 MB of memory and more than 350 seconds of training time due to their large parameter counts (138M and 144M, respectively). In contrast, lightweight models like **MobileNetV2** and **EfficientB0** demonstrate remarkable efficiency, with memory footprints of just 14 MB and 20 MB, and training times as low as 20 and 28 seconds. **ResNet** and **DenseNet** variants offer a balance between performance and complexity, with moderate memory use and training durations. The **RA-CNN** architecture, with 34.3 million parameters and 121 MB of memory, takes 89 seconds to train, offering a well-optimized structure for solar image analysis. Most notably, the **Proposed RA-TransformerNet + RSWS** model demonstrates superior performance with only slightly higher resource consumption: 134 MB memory, 38.5 million parameters, and 92 seconds training time, proving it to be a powerful yet efficient solution for multi-class defect classification in photovoltaic systems.

5. CONCLUSION

This study presents a strong combined deep learning system called RA-TransformerNet, which uses ReliefF-based Subspace Weighted SVM to accurately identify defects in solar panels by analyzing electroluminescence (EL) images. The proposed model effectively combines the local feature extraction strength of residual CNN blocks with the global contextual learning capability of vision transformers, further enhanced by attention modules such as CBAM or SE. Across comprehensive experimental evaluations on the ELPV dataset, the proposed model achieved superior performance, recording a peak accuracy of 98.23%, precision of 97.12%, sensitivity of 95.67%, and F-score of 96.35%, clearly outperforming traditional CNNs, hybrid models, and SVM-based baselines. Visual analyses through confusion matrices confirmed reduced misclassifications, particularly in boundary classes like "Possibly Normal" and "Possibly Defective." Additionally, the model exhibited a competitive training time of 92 seconds with 38.5 million parameters, demonstrating its computational efficiency compared to heavier architectures like VGG16 or VGG19. The proposed RA-TransformerNet framework is a scalable and highly accurate solution for defect detection in photovoltaic modules, supporting real-time deployment in solar inspection systems. These results underscore the potential of transformer-enhanced hybrid models for advancing intelligent fault analysis in renewable energy infrastructures.

REFERENCES

1. Abdelsattar, M., AbdelMoety, A., Ismeil, M. A., & Emad-Eldeen, A. (2025). Automated Defect Detection in Solar Cell Images Using Deep Learning Algorithms. IEEE Access.
2. Matusz-Kalász, D., Bodnár, I., & Jobbágy, M. (2025). An Overview of CNN-Based Image Analysis in Solar Cells, Photovoltaic Modules, and Power Plants. Applied Sciences, 15(10), 5511.
3. Vyas, P. A., & Bhandari, S. U. (2025, February). Advancing Solar Panel Fault Detection: A Comparative Study of Deep Learning Models. In 2025 International Conference on Emerging Systems and Intelligent Computing (ESIC) (pp. 919-926). IEEE.
4. Choudhary, S. K., & Mondal, A. (2025). Utilization of computer vision and machine learning for solar power prediction. In Computer Vision and Machine Intelligence for Renewable Energy Systems (pp. 67-84). Elsevier.
5. Nyangiwe, N. N., Kapimkenfack, A. D., Thantsha, N., & Msimanga, M. (2025). Performance monitoring of photovoltaic modules using machine learning based solutions: A survey of current trends.
6. Alnuaimi, K., Al-Sumaiti, A. S., Alansari, M., Wang, H., & Al Hosani, K. H. (2025). Deep Learning-Based Health Monitoring for Photovoltaic Systems. IEEE Journal of Photovoltaics.
7. Abdelsattar, M., AbdelMoety, A., & Emad-Eldeen, A. (2025). Applying image processing and computer vision for damage detection in photovoltaic panels. Mansoura Engineering Journal, 50(2), 2.
8. Jack, K. E., Kantiyok, R. Z., Ezugwu, E. O., Anunuso, J. C., Chukwuemeka, O., & Nwibo, E. G. (2025, February). Development of a convolutional neural network model for solar panel fault detection with preventive maintenance reporting options. In International Conference on Energy, Power, Environment, Control and Computing (ICEPECC 2025) (Vol. 2025, pp. 314-319). IET.
9. Demir, F. (2025). Enhancing Defect Classification in Solar Panels with Electroluminescence Imaging and Advanced Machine Learning Strategies. IEEE Access.
10. Aghaei, M., Kolahi, M., Nedaei, A., Venkatesh, N. S., Esmailifar, S. M., Moradi Sizkouhi, A. M., ... & Rüther, R. (2025). Autonomous Intelligent Monitoring of Photovoltaic Systems: An In-Depth Multidisciplinary Review. Progress in Photovoltaics: Research and Applications, 33(3), 381-409.

11. Umar, S., Nawaz, M. U., & Qureshi, M. S. (2024). Deep learning approaches for crack detection in solar PV panels. International Journal of Advanced Engineering Technologies and Innovations, 1(3), 50-72.
12. Duranay, Z. B. (2023). Fault detection in solar energy systems: A deep learning approach. Electronics, 12(21), 4397.
13. Verma, H., Siruvuri, S. V., & Budarapu, P. R. (2024). A machine learning-based image classification of silicon solar cells. International Journal of Hydromechatronics, 7(1), 49-66.
14. Shaik, A., Balasundaram, A., Kakarla, L. S., & Murugan, N. (2024). Deep learning-based detection and segmentation of damage in solar panels. Automation, 5(2), 128-150.
15. Mohanty, S. R., Maruf, M. U., Singh, V., & Ahmad, Z. (2024). Machine learning approaches for automatic defect detection in photovoltaic systems. arXiv preprint arXiv:2409.16069.
16. Branikas, E., Murray, P. and West, G., 2023. A novel data augmentation method for improved visual crack detection using generative adversarial networks. IEEE Access, 11, pp.22051-22059.
17. Prabhakaran, S., Uthra, R.A. and Preetharoselyn, J., 2023. Feature Extraction and Classification of Photovoltaic Panels Based on Convolutional Neural Network. Computers, Materials & Continua, 75(1).
18. Hammoudi, Y., Idrissi, I., Boukabous, M., Zerguit, Y. and Bouali, H., 2022. Review on maintenance of photovoltaic systems based on deep learning and internet of things. Indones. J. Electr. Eng. Comput. Sci, 26(2), pp.1060-1072.
19. Korkmaz, D. and Acikgoz, H., 2022. An efficient fault classification method in solar photovoltaic modules using transfer learning and multi-scale convolutional neural network. Engineering Applications of Artificial Intelligence, 113, p.104959.
20. Chen, X., Karin, T. and Jain, A., 2022. Automated defect identification in electroluminescence images of solar modules. Solar Energy, 242, pp.20-29.
21. Mohan, V. and Senthilkumar, S., 2022. IoT based fault identification in solar photovoltaic systems using an extreme learning machine technique. Journal of Intelligent & Fuzzy Systems, 43(3), pp.3087-3100.
22. Mantel, C., Villebro, F., dos Reis Benatto, G.A., Parikh, H.R., Wendlandt, S., Hossain, K., Poulsen, P., Spataru, S., Sera, D. and Forchhammer, S., 2019, September. Machine learning prediction of defect types for electroluminescence images of photovoltaic panels. In Applications of machine learning (Vol. 11139, p. 1113904). SPIE.
23. Shihavuddin, A.S.M., Rashid, M.R.A., Maruf, M.H., Hasan, M.A., ul Haq, M.A., Ashique, R.H. and Al Mansur, A., 2021. Image based surface damage detection of renewable energy installations using a unified deep learning approach. Energy Reports, 7, pp.4566-4576.
24. Pathak, S.P., Patil, S. and Patel, S., 2022. Solar panel hotspot localization and fault classification using deep learning approach. Procedia Computer Science, 204, pp.698-705.
25. Pa, M., Uddin, M.N. and Kazemi, A., 2022, December. A Fault Detection Scheme Utilizing Convolutional Neural Network for PV Solar Panels with High Accuracy. In 2022 IEEE 1st Industrial Electronics Society Annual On-Line Conference (ONCON) (pp. 1-5). IEEE.
26. Dwivedi, D., Babu, K., Yemula, P.K., Chakraborty, P. and Pal, M., 2022. Identification of surface defects on solar pv panels and wind turbine blades using attention based deep learning model. arXiv preprint arXiv:2211.15374.