International Journal of Basic and Applied Sciences, 14 (2) (2025) 535-541

International Journal of Basic and Applied Sciences

Website: www.sciencepubco.com/index.php/IJBAS
https://doi.org/10.14419/1h6tfk39
Research paper

Solar ViT: Vision Transformer for Fault Detection in Solar PV Systems

Pushpa Makwane [1], Yogesh Kumar [2], Alok Srivastava [2], Madhuri Gokhale [3] *, Saurabh Singh [3], Varsha Sisodiya [4]

[1] *Government Holkar Science College, Indore, M.P., India*
[2] *GD Goenka University, Gurugram, Haryana, India*
[3] *Jabalpur Engineering College, Jabalpur, M.P., India*
[4] *State Street Corporation Pvt. Ltd., Bengaluru, India*
**Corresponding author E-mail: mgokhale@jecjabalpur.ac.in*

Received: May 18, 2025, Accepted: June 11, 2025, Published: June 30, 2025

Abstract

Clean, renewable, and sustainable energy is vital for advancing social, economic, and environmental well-being, ultimately fostering productivity and long-term development. As solar photovoltaic (PV) systems become a cornerstone in global efforts to combat climate change and promote environmental sustainability, ensuring their reliable operation is increasingly important. Artificial Intelligence (AI) and Deep Learning (DL) have emerged as powerful enablers, offering advanced capabilities in monitoring, fault detection, and predictive maintenance of solar energy systems. These intelligent technologies enhance operational efficiency, minimize downtime, and support sustainable energy management. Blending AI and renewable energy supports key Sustainable Development Goals like clean energy, innovation, and climate action, driving both technological progress and environmental protection. This paper introduces SolarViT, a Vision Transformer-based deep learning model designed for precise fault detection in solar PV panels. By detecting issues like micro cracks and hotspots early, SolarViT helps prevent efficiency loss and reduce maintenance costs. Leveraging transfer learning, data augmentation, and ensemble methods, it delivers robust diagnostics while addressing interpretability, efficiency, and real-time deployment. This supports smarter solar infrastructure and promotes broader adoption of clean energy for sustainable development.

Keywords: *Artificial Intelligence; Convolutional Neural Networks; Crack Detection; Vision Transformer; Photovoltaic*

1. Introduction

Solar power systems have become an essential component of today's electrical grids as the trend toward renewable energy grows. Solar photovoltaic (PV) technology is preferred for sustainable energy generation due to its versatility and environmental friendliness. Figure 1 depicts how PV modules are formed by connecting several solar cells in series and parallel. Despite their benefits, these systems' performance and dependability can be hampered by issues such as panel degradation, hotspots, shading effects, and inverter failures [1]. Faults that go undetected or unresolved lower energy output and pose threats to system stability and lifespan, making it difficult to integrate solar energy into the power system [2][3].

Fig. 1: The Architecture of Solar Power Plant.

In Solar photovoltaic (PV) systems, electrical and structural failures are vulnerabilities that can reduce their performance and lifespan. Structural issues such as delamination, bubbles, yellowing, burns, deterioration, hotspots, scratches, and cracks can cause significant long-term damage to solar cells. Meanwhile, electrical problems, including open or closed circuits, potential-induced degradation (PID), junction box faults, diode malfunctions, and inverter failures, can disrupt the overall functionality of the system. These persistent faults often necessitate the replacement of affected modules. Furthermore, PV modules are susceptible to defects like microcracks, material delamination, corrosion, and oxidation, which can undermine their durability [4 – 6]. Micro cracks can be difficult to detect but may spread over time, worsening the module's condition. Environmental factors such as shading, soiling from dirt or snow, and hotspots can also reduce energy output, but these temporary defects can typically be resolved without removing the modules. Together, these structural and electrical challenges highlight the need for effective monitoring and maintenance to maintain the efficiency and reliability of PV systems [7]. Deep learning models have emerged as effective methods for detecting cracks in PV modules, enhancing their performance and lifespan [9] [10]. These structures enhance PV module inspection accuracy and efficiency, allowing for prompt maintenance and repairs while optimizing solar energy output capacity. Convolutional Neural Networks (CNNs) use convolutional, pooling, and fully connected layers with ReLU activation to detect defects in solar photovoltaic (PV) modules. Prominent CNN architectures include LeNet for digit recognition [11], AlexNet for pioneering deep learning in ImageNet [12], VGG for its deep, uniform filter design [13], GoogLeNet with inception modules [14], ResNet for addressing the vanishing gradient issue [15], and SqueezeNet for high accuracy with fewer parameters [16].

Transformer networks have recently emerged as effective models for various AI applications. In areas like computer vision, the Vision Transformer (ViT) has earned significant attention for its improved performance, inspired by the self-attention mechanism of transformer networks [17]. The Vision Transformer (ViT) leverages self-attention mechanisms to effectively model long-range dependencies and discern intricate spatial relationships within input data. Complementing this, the Multi-Layer Perceptron (MLP) component enables the learning of broad and adaptable patterns directly from raw inputs. Due to its architectural advantages, ViT has emerged as a prominent approach in a wide range of computer vision domains, such as image classification, object detection, semantic segmentation [18], image colorization [19], low-level vision tasks [20], and video analysis [21].

Furthermore, studies reveal that ViT models exhibit prediction errors more akin to human errors compared to CNNs [22 – 25].

The following are the main elements of this work.

- This is one of the methods to employ a vision transformer in PV fault detection.
- Comprehensive Fault Detection: The proposed method identifies Environmental faults (like Dust, snow, and bird droppings), physical damage, and internal electrical faults on solar panels.
- Optimized Precision: Utilizes advanced classifier optimization and feature engineering to improve impurity detection accuracy.
- Efficiency Improvement: Enhances tracking systems to determine cleaning or repair needs, boosting energy production efficiency.
- Sustainable and Cost-Effective Maintenance: Reduces maintenance costs and conserves resources by minimizing unnecessary cleaning and repairs.

The structure of the paper is as follows: Section 2 evaluates the relevant literature. A comprehensive explanation of the procedure, including the proposed structure, is provided in Section 3. A brief description of the dataset, experimental outcomes, and assessment measures, along with the results, has been provided in Section 4. Section 5 discusses the conclusion and future directions of the paper.

2. Related Work

With solar photovoltaic systems developing at a rapid pace and with a share in sustainable energy increasing rapidly, the need for efficient fault detection mechanisms and predictive energy management solutions has come into focus. The synergy of advanced machine learning, deep learning, and artificial intelligence with IoT-based systems has greatly contributed to augmenting this domain. The extent of efficiency and cost of maintenance of the Solar panels depend on the ability to detect accumulation of dirt, debris, cracks, and damage on the solar panel surface. These factors can cause a variety of issues, like an increase in the cost of maintenance, risk factors, a decrease in electricity production, and inaccurate estimations of production capabilities.

There were various approaches employed in the past for fault detection and maintaining the PV panels, like the use of Deep Learning Models with the use of CNNs, which were used for extracting features [47,48] from thermal and visual images taken from the UAVs with the help of camera, sensors and IR technologies. Each previous study conducted has also been summarised in Table 1. Models used for the study were ResNet, DenseNet, EfficientNetB0, MobileNetV3, ShuffleNet, and InceptionV3, whose primary applications were to identify hot spots, cracking, bypass diode faults, and cell anomalies [8]. Recent advancements in fault detection for photovoltaic (PV) systems have yielded various hybrid and deep learning models. Rudro et al. [34] introduced SPF-Net, a hybrid architecture combining U-Net and InceptionV3 with enhancements such as Squeeze-and-Excitation blocks, residual connections, and global average pooling. This model achieved a validation accuracy of 98.34% and a test accuracy of 94.35%, along with an F1-score of 94%. Duranay [8] proposed an efficient model utilizing depth-wise separable convolutions, which attained an accuracy of 93.93\%, offering faster feature extraction and reduced computational costs. Adhya et al. [26] employed LGBM, XGBoost, and CatBoost with real-time environmental data to achieve 99.996% precision, though they encountered challenges like data imbalance and implementation complexity.

Other models include Rao et al.'s [29] K-means clustering approach (99.72% precision), El-Banby's [36] rule-based and statistical hybrid model, and Memon et al.'s [37] CNN trained on historical PV data. Abubakar et al [38] combined ENN with BTA and statistical learning, while Eltuhamy et al. [39] adopted a neuro-fuzzy inference system for CIGS thin-film PVs. Amaral et al. [40] integrated image processing and PCA for PV tracker diagnostics, and Chen et al. [41] used autoregressive models with GLLR for rapid fault detection.

Badr et al. [28] explored multiple ML classifiers (DT, KNN, SVM) optimized via Bayesian techniques. Pamungkas et al. [27] developed UndenseNet, a lightweight model achieving high accuracy across multiple output classes, further enhanced by GAN-based augmentation. Since convolutional operations are essentially local and concentrate on neighboring pixel regions, they are unable to adequately capture long-range dependencies throughout the image [19]. Transformer models, on the other hand, are made to discover global contextual relationships [20]. The transformer architecture was initially created for machine translation and other natural language processing applications, but it has proven to be adaptable in a variety of fields. Kellil et al. [30] made a fine-tuned VGG-16 CNN on thermal images, reaching 99.91% precision under class-balanced conditions. Similarly, Shihavuddin et al. [35] applied VGG-16 on 3,336 thermal images for accurate fault detection.Venkatesh [31] applied a pre-trained VGG16 CNN to classify six PV panel conditions using SoftMax activation. Fazai et al. [33] combined Gaussian Process Regression and GLRT for fault detection, while Sridharan and Sugumaran [32] used AlexNet for feature extraction, J48 for feature selection, and random forest for classification, achieving 98.25% accuracy in under a second. The paper [43] proposes DPiT, a novel transformer-based model enhanced with convolution and cross-window self-attention for accurate solar cell defect detection. It achieves 91.7% top 1 accuracy on the ELPV dataset, outperforming existing methods with minimal computational overhead. The study [44] presents a Vision Transformer-based deep learning model for precise fault detection in PV modules using infrared images, achieving up to 98.23% accuracy. It outperforms existing methods in classifying both binary and multi-class PV anomalies, demonstrating strong potential for early fault detection. This paper [45] introduces a Transformer-based deep learning model for proactive fault prediction in PV systems, using the rate of change of solar parameters instead of historical trends. The model achieves a low MAE of 0.09377 and outperforms traditional methods like KNN, SVM, and NN in forecasting and fault severity classification. This study proposes a SegFormer-based framework for automated defect detection and pseudocolorization in PV modules [46], classifying defects into five types. It outperforms state-of-the-art models like UNET and DeepLab v3+, achieving 96.24% pixel-wise accuracy and strong F1-scores, demonstrating its effectiveness for precise visual inspection.4

Table 1: Key aspects of models proposed by various authors for fault detection in solar PV systems.

Author(s)	Techniques	Objectives	Method	Accuracy	Additional Notes
Adhya et al. [26] (2022)	XGBoost, LGBM, CatBoost	Fault detection using real-time data	Temperature & irradiance analysis	Precision: 99.996%	Addressed data imbalance issues
Pamungkas et al. [27] (2023)	CNN, SVM, Decision Trees	Fault diagnosis in PV systems	Historical database classification	High precision & recall	Focused on system-level faults
Badr et al. [28] (2021)	ImageNet, KNN, SVM	Automatic PV fault diagnosis	Model-based image classification	-	Multi-fault classification
Rao et al. [29] (2019)	K-Means Clustering	PV Array power production monitoring	Data noise handling and clustering	Accuracy 99.72%	Handled noisy data well
Kellil et al. [30] (2023)	CNN, VGG-16	Thermal image-based fault detection	Pre-trained CNN models	Precision: 99.91%	Five fault classifications
Naveen & Sugumaran [31] (2021)	CNN, UAV	Image Analysis	Fault classification using aerial images & UAV based data collection & analysis	Improved fault classification	Handled multiple fault types
Salaki et al. [9] (2024)	CNN, ResNet-50	Fault detection in solar panels	Image classification, fault detection using thermal imagery	ResNet-50: 85% accuracy, 42% loss Compared to VGG-16 with 80% accuracy	-
Sridharan et al. [32] (2021)	Multi-variant DL, RMVDM	Defect localization and detection	Histogram based pre-processing	Improved detection precision	Applied signal-based filtering
Fazai et al. [33] (2019)	GPR, GLRT	Fault detection using statistical models	Power-voltage current analysis	High detection rate	Focused on power system metrics

3. Proposed Method

The detailed methodology of the proposed framework is outlined in this section. An overview of the framework design is presented in Figure 2.

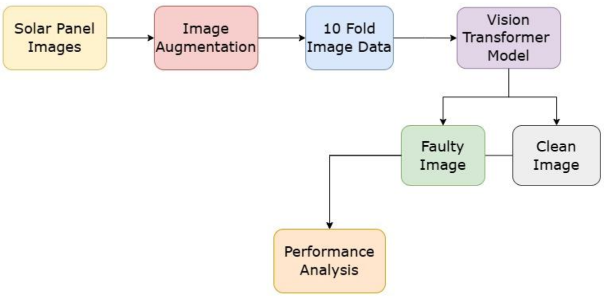

Fig. 2: Schematic Block Diagram of the Proposed Method

3.1 Processioning

The first step of image pre-processing is cropping the image to 256 × 256 pixels. To scale the range between 0 and 1, pixel values are normalized by dividing them by 255. Data augmentation is then used to improve the datasets for machine learning training. The proposed approach combines established approaches with random geometric transformations, including 20° to +20° rotations, 0.2 factor zooming, and horizontal/vertical flips.

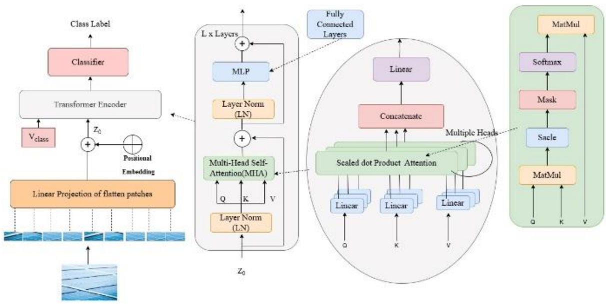

Fig. 3: Architecture of the Proposed SolarViT

3.2 SolarViT: solar vision transformer for fault detection

A transformer-based design called the Vision Transformer (ViT) [42] is highly effective in classifying images using a self-attention mechanism and retrieving features from them. Figure 3 illustrates the architecture of the SolarViT model. A key element of SolarViT is the combination of Multi-Head Self-Attention (MHA) blocks and a Multilayer Perceptron (MLP), which together enable the network to capture important global features effectively. The input image is initially split into fixed-size, non-overlapping patches, which are then flattened. These flattened patches are linearly projected and combined with the positional embeddings to preserve their spatial layout within the original image. Afterwards, the patches pass through a series of transformer blocks. Each block primarily includes MHA and MLP modules, both of which are preceded by normalization layers and linked together with residual connections.
In the Multi-Head Attention (MHA) mechanism, each input vector is projected into three distinct vectors: query (Q), key (K), and value (V). These vectors are grouped into corresponding matrices using the following linear transformations:

$$Q = XW_Q, \quad K = XW_K, \quad V = XW_V$$

Where the trainable weight matrices are WQ, WK, and WV. The similarity between embedded patches is then captured by computing the attention scores by taking the dot product of the Q and K matrices. This score matrix is passed through a SoftMax function to normalize the values, yielding an attention weight.
The self-attention output is calculated using the following formula:

$$\text{Attention}(Q,K,V)=\text{SoftMax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \tag{1}$$

Here, dk is the dimensionality of the key vectors and is used to scale the dot product, stabilizing the training process.
These attention outputs are then combined and passed through a linear projection layer and subsequently into the regression or classification head. The self-attention mechanism plays a crucial role in identifying important semantic features across different regions of the input image. A transformer encoder can contain multiple self-attention heads.
The output of a transformer block that includes normalization, multi-head attention, and residual connection is computed as:

$$MHA_{out}=MHA\big(NORM(x_{in})\big)+x_{in} \tag{2}$$

Where NORM represents layer normalization, xin is the input vector, and MHAout is the result of the attention block.
Following the attention layer is a Multi-Layer Perceptron (MLP), which utilizes the GeLU (Gaussian Error Linear Unit) activation function. The activation is computed by applying a Gaussian-based probabilistic weighting to the input values. Each transformer block includes skip connections, which help maintain gradient flow and improve model convergence. The final output of a transformer block is defined as:

$$T_{out}=MLP\big(NORM(MHA_{out})\big)+MHA_{out} \tag{3}$$

In this work, the SolarViT model is designed based on the ViT architecture with key adaptations. An improvement on the original ViT's 16 × 16 patch size, the source image is divided into non-overlapping 8 × 8 patches. Each patch is flattened and linearly projected, resulting in a 224 × 224 × 3 feature vector. This vector is passed through four transformer blocks, each with 8 self-attention heads, as determined by an ablation study.
For training, the model uses input images of dimension 224 × 224 × 3, which undergo pre-processing before being fed into the network. After passing through the transformer layers, the output is condensed into a one-dimensional vector, passed through SoftMax activation, and directed to a fully connected layer. The number of neurons in the output layer matches the number of target classes in the dataset. The model is trained using the categorical cross-entropy loss function and the AdamW optimizer. The model is trained over a maximum of 100 epochs using a learning rate of 0.0001, weight decay set to 0.001, and a batch size of 30. After training, its average classification performance is assessed using standard accuracy evaluation metrics. To reduce the risk of overfitting, observed from the performance gap between training and test accuracy, additional regularization techniques were incorporated into the SolarViT architecture. Dropout layers with a rate of 0.2 were applied within the transformer encoder blocks, and L2 regularization was used to penalize overly complex weight distributions. These enhancements, combined with an expanded data augmentation pipeline—including random brightness adjustment and noise injection- contributed to improved model generalization on unseen data.

4. Experimental Results and Discussion

Deep learning models such as VGG16, MobileNetV3, DenseNet, CNN, VGG19, InceptionV3, ResNet50, and the proposed vision transformer (ViT) are employed for detecting faults in solar panels. These models demonstrate strong capabilities in autonomously extracting meaningful features from images, which contributes to more accurate and efficient fault detection. By enabling automated and scalable surveillance of extensive solar installations, they help maintain stable operation even under fluctuating environmental conditions. Leveraging diverse architecture contributes to a more robust and comprehensive approach for identifying issues in solar panels.

4.1 Experimental datasets

To assess the effectiveness of the proposed approach, a publicly available dataset was utilized. An overview of the dataset, including the number of fault-specific images, is presented in Table 2. The dataset of PV panels of 885 images was partitioned into training, validation, and testing sets in a 70%-10%-20% ratio, yielding 620 images for training, 88 for validation, and 177 for testing. In terms of dataset characteristics, the SolarViT model was trained on a dataset comprising 885 labeled images representing a variety of fault types such as microcracks, hotspots, dust accumulation, snow coverage, and partial shading. These images were collected using aerial drones and infrared imaging devices across solar farms located in India and Germany. The dataset includes a diverse range of environmental conditions (e.g., varying light, dust, and weather scenarios), which enhances its representativeness. A detailed fault-wise distribution is provided in Table 2, and image examples are shown in Figure 4.

Table 2: Details of images in solar panel fault detection dataset.

Font Size	Total Images	Training	Validation	Testing
Fault detection	885	620	88	177

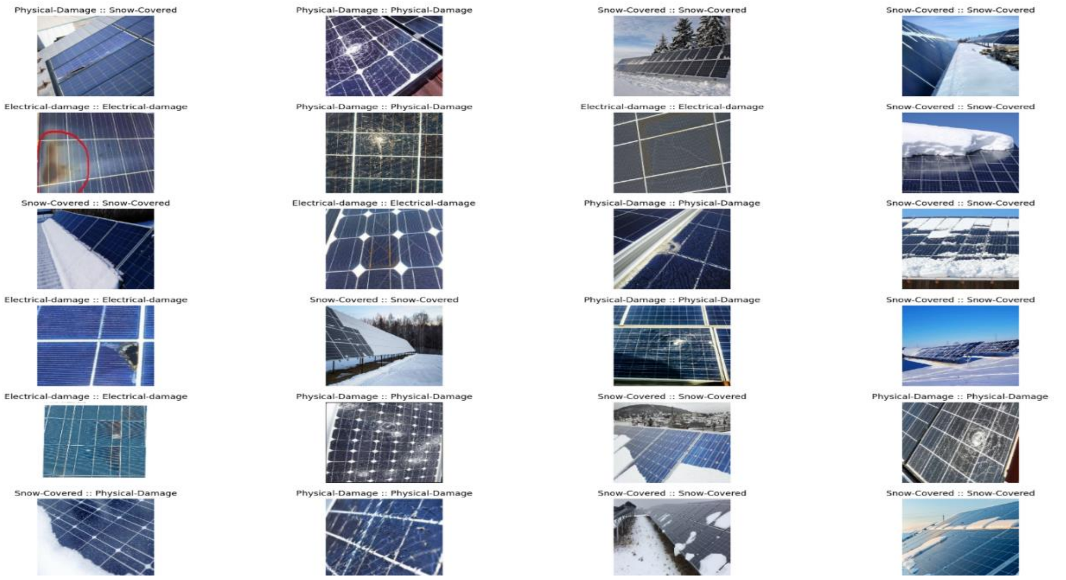

Fig. 4: Various Defects that occur in Solar Panel

4.2 Implementation and evaluation

The proposed SolarViT model, along with other comparative models, was implemented using the Keras framework with TensorFlow as the backend. All models were developed in Python, utilizing the Scikit-learn library for classification tasks. Experiments were executed on a system equipped with an Intel Core i7 processor running at 5 GHz and 8 GB of RAM.

To assess model performance, standard classification metrics such as Accuracy, Precision, Recall, F1-score, and Confusion Matrix were used. The results were averaged over a 10-fold cross-validation process to ensure robustness and reliability.

The evaluation metrics are defined as follows:

$$\begin{aligned} \text{Accuracy} &= \frac{\text{TP}+\text{TN}}{\text{TP}+\text{TN}+\text{FP}+\text{FN}} \\ \text{Precision} &= \frac{\text{TP}}{\text{TP}+\text{FP}} \\ \text{Recall} &= \frac{\text{TP}}{\text{TP}+\text{FN}} \\ \text{F1} &= 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \end{aligned} \tag{4}$$

Were,

- TP (True Positives): Correctly predicted positive samples
- TN (True Negatives): Correctly predicted negative samples
- FP (False Positives): Incorrectly predicted positive samples
- FN (False Negatives): Incorrectly predicted negative samples.

4.3 Performance analysis and comparison

This section evaluates deep learning models' performance during testing and validation. To assess the models' accuracy and dependability, critical evaluation metrics such as F1 Score, Recall, and Precision were systematically calculated.

An analysis of different deep learning models' training accuracies highlights the change in the effectiveness of learning from training datasets, as summarized in Table 3. DenseNet shows notable underfitting, with training and testing accuracies of only 21.28% and 21.00%, respectively, revealing it is unable to effectively identify deeper patterns in the solar dataset. In comparison, MobileNetV3 and VGG19 exhibit a reasonable level of learning efficiency, attaining testing and training accuracy within the range of 70% to 77%, reflecting balanced performance without signs of overfitting. CNN and VGG16 achieve accuracies close to 80%, showing robust learning capabilities, though they approach a threshold where overfitting could become a concern. ResNet50 exhibits a similar pattern. InceptionV3 and the proposed SolarViT model, however, achieve outstanding training accuracy exceeding 90%. Notably, SolarViT achieves a remarkable 97.01% training accuracy and 94.35% testing accuracy, raising concerns about potential overfitting due to its very high performance on the training dataset. It is important to achieve a balance between generalization and learning efficiency for models that have exceptionally high training model accuracy.

When the training and validation datasets are compared, it can be shown that although all models perform well on the training data, accuracy noticeably reduces during validation. Since the models have trouble generalizing new data, this drop suggests possible overfitting. Performance during the validation phase varies across models. Notably, the proposed SolarViT model achieves superior results, with a validation accuracy of approximately 98.34% and training Precision scores ranging from 20% to 99%, where SolarViT records the highest precision (99%), while DenseNet exhibits the lowest (20%). Recall values show a similar trend, ranging from 20% to 98%, with DenseNet again at the lower end and SolarViT achieving a recall of 99%. With an impressive score of 98%, SolarViT outperforms other models in the F1 score, which measures a balance between precision and recall.

This analysis highlights the importance of achieving a balance between precision and recall, optimizing model performance for real-world applications. SolarViT's exceptional performance underscores its potential, but attention must be paid to ensure it maintains robust generalization capabilities.

Table 3: Comparative performance metrics of various deep learning models on solar panel fault dataset.

Models	Training				Validation	Testing			
	Accuracy	Precision	Recall	F1 score	Accuracy	Accuracy	Precision	Recall	F1 score
Dense-Net	21.28	21.00	20.00	20.00	20.90	21.00	19.00	21.00	19.00
MobileNetV3	70.71	70.00	71.00	70.00	69.99	70.04	66.00	66.00	66.00
VGG19	76.89	75.00	76.00	75.00	76.01	77.00	79.00	76.00	78.00
CNN	76.24	76.00	75.00	75.00	74.86	79.40	75.00	77.00	75.00
VGG16	80.16	81.00	81.00	80.00	79.09	80.00	79.00	87.00	80.00
Resnet50	80.19	80.00	80.00	80.00	78.29	80.79	81.00	81.00	81.00
InceptionV3	91.74	92.00	91.00	91.00	89.87	90.19	91.00	92.00	90.00
SolarViT (Proposed)	99.01	99.00	99.00	98.00	93.34	94.35	94.00	94.00	94.00

Although the current implementation of SolarViT demonstrates high accuracy and robustness in fault classification, its real-time deployment and system-wide scalability are yet to be realized. These aspects are critical for translating research into practical solar PV monitoring solutions. In this context, future efforts will focus on evaluating SolarViT's performance on edge computing devices, such as NVIDIA Jetson Nano, to examine latency, computational load, and power efficiency under field conditions. To facilitate deployment in low-resource environments, techniques like model pruning and quantization will be investigated. Moreover, the integration of SolarViT within IoT-enabled infrastructures will be explored to enable automated, real-time fault detection across large-scale solar farms. Such integration has the potential to support remote diagnostics and predictive maintenance via cloud-based platforms. These planned directions are essential for ensuring the scalability and operational relevance of the proposed model in diverse, real-world solar energy applications.

5. Conclusion

In conclusion, this study presents SolarViT, a novel Vision Transformer-based model developed to improve the accuracy and robustness of solar panel fault detection. The approach begins with comprehensive data preprocessing, including augmentation to expand the training dataset and normalization to standardize inputs. Augmentation techniques such as rotation, shifting, shearing, zooming, and brightness modulation are used to improve model performance. Next proposed SolarViT model divides input images into 8 × 8 patches and processes them through two multilayer perceptrons with 2096 and 1048 neurons, respectively. Its core architecture incorporates four transformer blocks combined with an 8-head multi-head self-attention mechanism, enabling effective feature extraction. The Adam optimizer is used to minimize categorical cross-entropy loss while optimizing with a learning rate of 0.0001. The model achieves superior results, with a validation accuracy of 98.34%, test accuracy of 94.35%, and precision, recall, and F1 scores all at 0.94, clearly outperforming existing methods for solar panel fault detection. Future studies could extend SolarViT's use to other energy-efficient technologies, like hydroelectric power plants and wind turbines, to evaluate its wider utility. Additionally, adding real-time problem detection capabilities could boost the model's usefulness in preserving sustainable energy infrastructure.

References

[1] Joshua, S. R., Park, S., & Kwon, K. (2024). H2 URESONIC: Design of a Solar-Hydrogen University Renewable Energy System for a New and Innovative Campus. Applied Sciences, 14(4), 1554. https://doi.org/10.3390/app14041554

[2] Kabir, E., Kumar, P., Kumar, S., Adelodun, A. A., & Kim, K. H. (2018). Solar energy: Potential and future prospects. Renewable and Sustainable Energy Reviews, 82, 894-900. https://doi.org/10.1016/j.rser.2017.09.094

[3] Sovacool, B. K., Mullard, S., & Ceballos, J. C. (2024). "Made for corruption?" Private sector actors, renewable energy, and corruption risks for wind power in Mexico and solar electricity in Kenya. The Electricity Journal, 37(7-10), 107448. https://doi.org/10.1016/j.tej.2024.107448

[4] Izam, N. S. M. N., Itam, Z., Sing, W. L., & Syamsir, A. (2022). Sustainable development perspectives of solar energy technologies with focus on solar Photovoltaic—A review. Energies, 15(8), 2790. https://doi.org/10.3390/en15082790

[5] Hasan, M. M., Hossain, S., Mofijur, M., Kabir, Z., Badruddin, I. A., Yunus Khan, T. M., & Jassim, E. (2023). Harnessing solar power: a review of photovoltaic innovations, solar thermal systems, and the dawn of energy storage solutions. Energies, 16(18), 6456. https://doi.org/10.3390/en16186456

[6] Saffari, M., Khodayar, M., & Khodayar, M. E. (2022). Deep recurrent extreme learning machine for behind-the-meter photovoltaic disaggregation. The Electricity Journal, 35(5), 107137. https://doi.org/10.1016/j.tej.2022.107137

[7] Sharif, A., Meo, M. S., Chowdhury, M. A. F., & Sohag, K. (2021). Role of solar energy in reducing ecological footprints: An empirical analysis. Journal of Cleaner Production, 292, 126028. https://doi.org/10.1016/j.jclepro.2021.126028

[8] Duranay, Z. B. (2023). Fault detection in solar energy systems: A deep learning approach. Electronics, 12(21), 4397. https://doi.org/10.3390/electronics12214397

[9] Joshua, S. R., Yeon, A. N., Park, S., & Kwon, K. (2024). A Hybrid Machine Learning Approach: Analyzing Energy Potential and Designing Solar Fault Detection for an A IoT-Based Solar–Hydrogen System in a University Setting. Applied Sciences, 14(18), 8573. https://doi.org/10.3390/app14188573

[10] Joshua, S. R., Junghyun, Y., Park, S., & Kwon, K. (2024). Integrating Deep Learning and Energy Management Standards for Enhanced Solar–Hydrogen Systems: A Study Using MobileNetV2, InceptionV3, and ISO 50001: 2018. Hydrogen, 5(4), 819-850. https://doi.org/10.3390/hydrogen5040043

[11] El-Sawy, A., El-Bakry, H., & Loey, M. (2017). CNN for handwritten arabic digits recognition based on LeNet-5. In Proceedings of the International Conference on Advanced Intelligent Systems and Informatics 2016 2 (pp. 566-575). Springer International Publishing. https://doi.org/10.1007/978-3-319-48308-5_54

[12] Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. Advances in neural information processing systems, 25. https://proceedings.neurips.cc/paper_files/paper/2012/file/c399862d3b9d6b76c8436e924a68c45b-Paper.pdf

[13] Pasban, S., Mohamadzadeh, S., Zeraatkar-Moghaddam, J., & Shafiei, A. K. (2020). Infant brain segmentation based on a combination of VGG-16 and U-Net deep neural networks. IET Image Processing, 14(17), 4756-4765. https://doi.org/10.1049/iet-ipr.2020.0469

[14] Sam, S. M., Kamardin, K., Sjarif, N. N. A., & Mohamed, N. (2019). Offline signature verification using deep learning convolutional neural network (CNN) architectures GoogLeNet inception-v1 and inception-v3. Procedia Computer Science, 161, 475-483. https://doi.org/10.1016/j.procs.2019.11.147

[15] Borawar, L., & Kaur, R. (2023, March). ResNet: Solving vanishing gradient in deep networks. In Proceedings of International Conference on Recent Trends in Computing: ICRTC 2022 (pp. 235-247). Singapore: Springer Nature Singapore. https://doi.org/10.1007/978-981-19-8825-7_21
[16] Tsivgoulis, M., Papastergiou, T., & Megalooikonomou, V. (2022). An improved SqueezeNet model for the diagnosis of lung cancer in CT scans. Machine Learning with Applications, 10, 100399. https://doi.org/10.1016/j.mlwa.2022.100399
[17] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... & Polosukhin, I. (2017). Attention is all you need. Advances in neural information processing systems, 30. https://proceedings.neurips.cc/paper_files/paper/2017/file/3f5ee243547dee91fbd053c1c4a845aa-Paper.pdf
[18] Zheng, S., Lu, J., Zhao, H., Zhu, X., Luo, Z., Wang, Y., ... & Zhang, L. (2021). Rethinking semantic segmentation from a sequence-to-sequence perspective with transformers. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (pp. 6881-6890). https://doi.org/10.48550/arXiv.2012.15840
[19] Kumar, M., Weissenborn, D., & Kalchbrenner, N. (2021). Colorization transformer. arXiv preprint arXiv:2102.04432. https://doi.org/10.48550/arXiv.2102.04432
[20] Chen, H., Wang, Y., Guo, T., Xu, C., Deng, Y., Liu, Z., ... & Gao, W. (2021). Pre-trained image processing transformer. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (pp. 12299-12310). https://doi.org/10.48550/arXiv.2012.00364
[21] Arnab, A., Dehghani, M., Heigold, G., Sun, C., Lučić, M., & Schmid, C. (2021). Vivit: A video vision transformer. In Proceedings of the IEEE/CVF international conference on computer vision (pp. 6836-6846). https://doi.org/10.48550/arXiv.2103.15691
[22] Ramadan, E. A., Moawad, N. M., Abouzalm, B. A., Sakr, A. A., Abouzaid, W. F., & El-Banby, G. M. (2024). An innovative transformer neural network for fault detection and classification for photovoltaic modules. Energy Conversion and Management, 314, 118718. https://doi.org/10.1016/j.enconman.2024.118718
[23] De Luis, A., Tran, M., Hanyu, T., Tran, A., Haitao, L., McCann, R., ... & Le, N. (2024, May). Solarformer: Multi-scale transformer for solar pv profiling. In 2024 International Conference on Smart Grid Synchronized Measurements and Analytics (SGSMA) (pp. 1-8). IEEE. 10.1109/SGSMA58694.2024.10571461
[24] Dwivedi, D., Babu, K. V. S. M., Yemula, P. K., Chakraborty, P., & Pal, M. (2024). Identification of surface defects on solar pv panels and wind turbine blades using attention based deep learning model. Engineering Applications of Artificial Intelligence, 131, 107836. https://doi.org/10.1016/j.engappai.2023.107836Get rights and content
[25] Xie, X., Liu, H., Na, Z., Luo, X., Wang, D., & Leng, B. (2021). DPiT: Detecting defects of photovoltaic solar cells with image transformers. IEEE Access, 9, 154292-154303. 10.1109/ACCESS.2021.3119631
[26] Adhya, D., Chatterjee, S., & Chakraborty, A. K. (2022). Performance assessment of selective machine learning techniques for improved PV array fault diagnosis. Sustainable Energy, Grids and Networks, 29, 100582. https://doi.org/10.1016/j.segan.2021.100582
[27] Pamungkas, R. F., Utama, I. B. K. Y., & Jang, Y. M. (2023). A novel approach for efficient solar panel fault classification using coupled udensenet. Sensors, 23(10), 4918. https://doi.org/10.3390/s23104918
[28] Badr, M. M., Hamad, M. S., Abdel-Khalik, A. S., Hamdy, R. A., Ahmed, S., & Hamdan, E. (2021). Fault identification of photovoltaic array based on machine learning classifiers. IEEE Access, 9, 159113-159132. 10.1109/ACCESS.2021.3130889
[29] Rao, S., Spanias, A., & Tepedelenlioglu, C. (2019, May). Solar array fault detection using neural networks. In 2019 IEEE international conference on industrial cyber physical systems (ICPS) (pp. 196-200). IEEE. 10.1109/ICPHYS.2019.8780208
[30] Kellil, N., Aissat, A., & Mellit, A. (2023). Fault diagnosis of photovoltaic modules using deep neural networks and infrared images under Algerian climatic conditions. Energy, 263, 125902. https://doi.org/10.1016/j.energy.2022.125902
[31] Venkatesh, S. N., & Sugumaran, V. (2021). Fault Detection in aerial images of photovoltaic modules based on Deep learning. In IOP Conference Series: Materials Science and Engineering (Vol. 1012, No. 1, p. 012030). IOP Publishing. 10.1088/1757-899X/1012/1/012030
[32] Sridharan, N. V., & Sugumaran, V. (2021). Visual fault detection in photovoltaic modules using decision tree algorithms with deep learning features. Energy Sources, Part A: Recovery, Utilization, and Environmental Effects, 1-17. https://doi.org/10.1080/15567036.2021.2020379
[33] Fazai, R., Abodayeh, K., Mansouri, M., Trabelsi, M., Nounou, H., Nounou, M., & Georghiou, G. E. (2019). Machine learning-based statistical testing hypothesis for fault detection in photovoltaic systems. Solar Energy, 190, 405-413. https://doi.org/10.1016/j.solener.2019.08.032
[34] Rudro, R. A. M., Nur, K., Al Sohan, M. F. A., Mridha, M. F., Alfarhood, S., Safran, M., & Kanagarathinam, K. (2024). SPF-Net: Solar panel fault detection using U-Net based deep learning image classification. Energy Reports, 12, 1580-1594. https://doi.org/10.1016/j.egyr.2024.07.044
[35] Shihavuddin, A. S. M., Rashid, M. R. A., Maruf, M. H., Hasan, M. A., ul Haq, M. A., Ashique, R. H., & Al Mansur, A. (2021). Image based surface damage detection of renewable energy installations using a unified deep learning approach. Energy Reports, 7, 4566-4576. https://doi.org/10.1016/j.egyr.2021.07.045
[36] El-Banby, G. M., Moawad, N. M., Abouzalm, B. A., Abouzaid, W. F., & Ramadan, E. A. (2023). Photovoltaic system fault detection techniques: a review. Neural Computing and Applications, 35(35), 24829-24842. https://doi.org/10.1007/s00521-023-09041-7
[37] Memon, S. A., Javed, Q., Kim, W. G., Mahmood, Z., Khan, U., & Shahzad, M. (2022). A machine-learning-based robust classification method for PV panel faults. Sensors, 22(21), 8515. https://doi.org/10.3390/s22218515
[38] Abubakar, A., Jibril, M. M., Almeida, C. F., Gemignani, M., Yahya, M. N., & Abba, S. I. (2023). A novel hybrid optimization approach for fault detection in photovoltaic arrays and inverters using AI and statistical learning techniques: a focus on sustainable environment. Processes, 11(9), 2549. https://doi.org/10.3390/pr11092549
[39] Eltuhamy, R. A., Rady, M., Almatrafi, E., Mahmoud, H. A., & Ibrahim, K. H. (2023). Fault Detection and Classification of CIGS Thin-Film PV Modules Using an Adaptive Neuro-Fuzzy Inference Scheme. Sensors, 23(3), 1280. https://doi.org/10.3390/s23031280
[40] Amaral, T. G., Pires, V. F., & Pires, A. J. (2021). Fault detection in PV tracking systems using an image processing algorithm based on PCA. Energies, 14(21), 7278. https://doi.org/10.3390/en14217278
[41] Chen, L., Li, S., & Wang, X. (2016). Quickest fault detection in photovoltaic systems. IEEE Transactions on Smart Grid, 9(3), 1835-1847. 10.1109/TSG.2016.2601082
[42] Sharir, G., Noy, A., & Zelnik-Manor, L. (2021). An image is worth 16x16 words, what is a video worth?. arXiv preprint arXiv:2103.13915. https://doi.org/10.48550/arXiv.2103.13915
[43] Xie, X., Liu, H., Na, Z., Luo, X., Wang, D., & Leng, B. (2021). DPiT: Detecting defects of photovoltaic solar cells with image transformers. IEEE Access, 9, 154292-154303. https://doi.org/10.1109/ACCESS.2021.3119631
[44] Ramadan, E. A., Moawad, N. M., Abouzalm, B. A., Sakr, A. A., Abouzaid, W. F., & El-Banby, G. M. (2024). An innovative transformer neural network for fault detection and classification for photovoltaic modules. Energy Conversion and Management, 314, 118718. https://doi.org/10.1016/j.enconman.2024.118718
[45] Khalil, I. U., Haq, A. U., & ul Islam, N. (2024). A deep learning-based transformer model for photovoltaic fault forecasting and classification. Electric Power Systems Research, 228, 110063. https://doi.org/10.1016/j.epsr.2023.110063
[46] Mahboob, Z., Khan, M. A., Lodhi, E., Nawaz, T., & Khan, U. S. (2024). Using segFormer for effective semantic cell segmentation for fault detection in photovoltaic arrays. IEEE Journal of Photovoltaics. https://doi.org/10.1109/JPHOTOV.2024.3450009.
[47] Y. Kumar, S. K. Verma, and S. Sharma, “Multi-pose facial expression recognition using hybrid deep learning model with improved variant of gravitational search algorithm.,” Int.Arab J. Inf. Technol., vol. 19, no. 2, pp. 281–287, 2022.
[48] Kumar, Y., Verma, S.K. and Sharma, S., 2023. Facial expression recognition of multiple stylised characters using deep convolutional neural network. International Journal of Advanced Intelligence Paradigms, 26(3-4), pp.362-391.