Journal of Computational Analysis and Applications VOL. 34, NO. 4, 2025
10.48047/jocaaa.2025.34.04.51

Machine Learning-Driven Thermal Imaging for Fault Detection in Solar PV Systems

G. Tejaswi, Gudugunta Kalyani, Chilukuri Damini, Gangapatnam Varshitha

Department of Electronics and Communication Engineering, Geetanjali Institute of Science and Technology, Nellore, Andhra Pradesh, India.

ABSTRACT

In India, the primary sources of power generation include wind, coal, and solar energy. Among these, solar energy stands out as a more efficient and sustainable option due to its renewable nature, wide availability, and minimal environmental impact. With the nation's increasing shift toward clean and sustainable energy, solar power is being extensively integrated into the national grid. However, faults and failures in solar photovoltaic (PV) cells remain a significant challenge, often leading to power losses and reduced system efficiency. Traditional fault detection methods, such as manual inspections and basic thermal assessments, are typically slow, prone to human error, and often fail to detect issues in a timely manner. To address these limitations, this work adopts a machine learning-based approach for thermal image classification to detect faults in PV systems. By leveraging thermographic images, machine learning algorithms can accurately identify anomalies such as cell defects, hotspots, or overheating, which are otherwise difficult to detect manually. This automated technique enhances the reliability and performance of solar PV systems, prolongs their operational lifespan, and contributes to more efficient energy production. Furthermore, it leads to substantial cost savings by minimizing unplanned maintenance and preventing major system failures. Overall, the integration of machine learning with thermal imaging provides a powerful solution for maintaining optimal performance in India's solar energy infrastructure.

Keywords: Solar Photovoltaic (PV) Systems, Thermal Imaging, Fault Detection, Renewable Energy, Smart Energy Monitoring, India Energy Infrastructure.

1. INTRODUCTION

As the global energy landscape continues to evolve, there is a noticeable decline in the reliance on traditional fossil fuels such as coal. According to the International Energy Agency (IEA), global coal demand is projected to decrease by 2.3% by 2026, primarily driven by the growing adoption of renewable energy sources. While wind energy has emerged as a promising alternative, its infrastructure demands such as expansive land and large turbines—limit its feasibility in certain regions. In contrast, solar energy offers a more adaptable and sustainable solution, suitable for a wide range of geographic and application contexts.

The efficiency and reliability of solar photovoltaic (PV) systems are critical for ensuring uninterrupted energy supply. However, these systems are susceptible to faults such as hotspots, broken cells, and faulty electrical connections, which can lead to substantial power generation losses and long-term damage if not addressed promptly. Hence, continuous monitoring and maintenance of PV installations are essential to uphold system performance and operational lifespan.

Thermographic imaging has emerged as a powerful, non-intrusive diagnostic tool for real-time monitoring of PV systems. By capturing and analyzing the infrared radiation emitted from solar panels, thermal cameras produce temperature distribution maps that help identify defects early—before they escalate into major failures. This technique not only enhances fault detection but also supports proactive maintenance practices, ultimately improving the reliability and efficiency of solar power infrastructure.

10.48047/jocaaa.2025.34.04.51

Recent innovations in thermographic imaging, particularly the integration of high-resolution infrared sensors and machine learning algorithms, have further advanced the accuracy and scalability of fault detection. Machine learning enables automated analysis of large volumes of thermal data, facilitating faster and more precise diagnosis. Case studies, such as those conducted in Australia, have demonstrated significant reductions in system downtime and maintenance costs through the combined use of thermal imaging and AI-based analysis.

This paper explores the role of thermographic imaging in enhancing fault detection and maintenance of PV systems. By leveraging modern imaging technologies and intelligent algorithms, this approach aligns with the global shift toward resilient, efficient, and sustainable solar energy solutions.

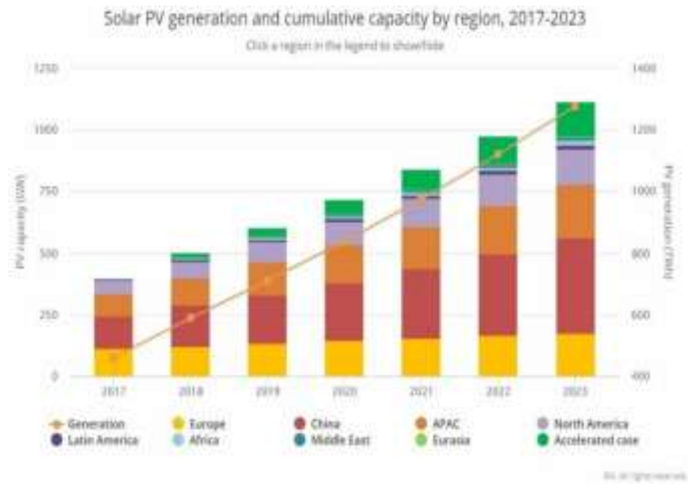


Fig. 1: Statistical graph of solar PV generation

2. LITERATURE SURVEY

Ahmed, Waqas et.al [1] (2024), they developed a Simulink model to assess the impact of environmental stresses on PVS power losses, showcasing the method's practical applicability. This approach significantly enhanced the reliability and performance of PVS, aligning with global initiatives such as the United Nations' Sustainable Development Goals, EU energy targets, the Kyoto Protocol, and the Paris Climate Accord. It ensured sustainable and reliable PVS operation over the next 25 years.

Baltacı et.al [2] (2024) Their objective was to perform two distinct fault analyses utilizing image processing techniques with thermal images and machine learning techniques using inverter and other physical data. The results showed that hotspot and bypass failures on the panels could be detected successfully using these methods.

Thakfan et.al [3] (2024) They proposed comprehensive survey identified emerging trends in AI-driven PV fault detection, highlighted the most advanced methodologies, and proposed a novel AI based approach to enhance fault detection and classification capabilities. The findings aimed to advance the state of technology in this field, offering insights into more efficient and practical solutions for PV system fault management.

Tanda et.al [4] (2024) They compared thermographic results from two different photovoltaic (PV) plants using two remote sensing-based approaches: the traditional UAV-mounted thermal camera survey and the inspection by high-speed thermal cameras mounted on an airplane. Post-processing of thermal patterns showed good agreement between the results from the two aerial platforms, with an overlap of

10.48047/jocaaa.2025.34.04.51

thermal anomalies detected up to 98%. An economic analysis demonstrated that, while airplane surveys incurred higher costs compared to UAV surveys due to vehicle hire and more expensive instrumentation, they required less time and were more convenient for inspecting large-scale PV plants or multiple PV plants located within a close area.

Qureshi et.al [5] (2024) They presented Fault Detection and Diagnosis (IFDD) system using convolutional neural networks (CNNs) for multiclass classification of defects in solar photovoltaic (SPV) panels. We addressed challenges like infrared thermography (IRT) data scarcity, complex defect patterns, and low thermal image quality due to noise and calibration issues. To overcome these challenges, we carefully prepared a customized, high-quality, but severely imbalanced six-class thermographic radiometric dataset of SPV panels. Unlike previous approaches, we used numerical temperature values in floating-point format to train and validate the predictive models. The trained models demonstrated high accuracy in efficiently diagnosing thermal anomalies. Finally, to build trust in the IFDD system, we investigated the classification model's underlying process using perceptive explainability to highlight the most discriminant image features and mathematical- structure-based interpretability to achieve multiclass feature clustering.

Vieira et.al [6] (2024) They proposed a cost-effective method to determine if a photovoltaic cell was defective. It applied a two-dimensional Gaussian fit to images generated by fast Fourier transform and principal component analysis algorithms on thermographic data from lock-in thermography tests. The coefficient of determination (R^2) was used as a measure of fitting quality. The method demonstrated potential for application on the first principal component, with R^2 between 0.944 and 0.986, and magnitude images, with R^2 between 0.965 and 0.985, to identify and distinguish non-defective cells from defective ones.

Abdulla et.al [7] (2024) The analysis underscored the critical need for integrating maintenance strategies to enhance the effectiveness of photovoltaic (PV) systems. By combining reliability assessments with economic and technical considerations, maintenance planning was optimized to improve system availability and resource efficiency. This approach aligned with the Sustainable Development Goals (SDGs) for affordable, reliable, and sustainable energy, while also ensuring grid security.

Umar et.al [8] (2024) They reviewed the applications of thermal imaging and AI techniques in detecting and classifying defects in solar panels, focusing on their advantages, challenges, and future prospects. It highlighted how integrating thermal imaging with AI algorithms enables automated detection and analysis of various defects, including cracks, delamination, hotspots, and corrosion, without manual intervention. The study emphasized the benefits of rapid inspection speed, high accuracy, and compatibility with large-scale solar installations. These technologies offer real-time monitoring capabilities, enabling early detection of defects and proactive maintenance actions to prevent performance degradation and costly repairs. Additionally, thermal imaging provides valuable insights into panel health and performance, facilitating data-driven decision-making for solar asset management. Shafique et.al [9] (2024) They proposed the critical issue of insulator fault detection in electric substations, emphasizing the importance of timely identification to prevent accidents. It developed a computer vision system that utilized advanced segmentation algorithms to automatically detect and categorize defects in insulators based on infrared thermography (IRT) images. The system combined unsupervised clustering for efficient training with supervised learning for accurate classification. It first employed clustering techniques to reduce human labelling needs and then used a Gaussian kernel support vector machine (SVM) algorithm to classify various insulator defects using extracted features.

10.48047/jocaaa.2025.34.04.51

The trained algorithm effectively identified and classified several fault types in thermal images of insulators, enabling timely intervention to prevent potential power outages and safety hazards.

3. PROPOSED SYSTEM

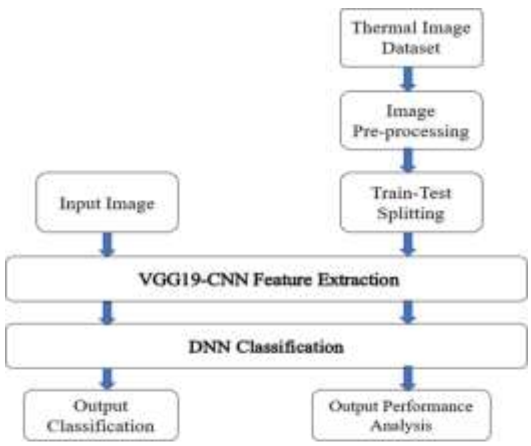
Ensuring the optimal performance of solar energy systems is crucial, as faults in PV modules can lead to significant energy losses and potential safety hazards. Traditional manual inspection methods are time-consuming and may not effectively detect underlying issues. Infrared thermography has emerged as a non-invasive diagnostic tool, capable of identifying anomalies such as hot spots and defective cells by visualizing temperature variations across PV panels. However, manual analysis of thermographic images is labour-intensive and subject to human error, necessitating the development of automated, accurate fault detection systems.

Integrating machine learning techniques with thermographic imaging offers a promising approach to enhance fault detection in PV systems. By training algorithms on labelled datasets of thermographic images, these models can learn to recognize patterns associated with various faults, enabling rapid and precise classification. Methods such as Convolutional Neural Networks (CNNs) have demonstrated high accuracy in classifying PV cell faults, even with limited datasets, through data augmentation strategies. Additionally, ensemble learning techniques like bagging and boosting can further improve model robustness and generalization. Implementing these machine learning-driven solutions facilitates continuous monitoring and maintenance of PV installations, thereby enhancing their efficiency and reliability.

The process begins with an input thermal image, which undergoes preprocessing to enhance quality. The dataset is then split into training and testing sets before being passed through a VGG19-CNN model for feature extraction. A DNN classifier processes these features to classify the image into either a high or low state. Finally, performance analysis is conducted to evaluate the model's accuracy and effectiveness, ensuring reliable classification. This structured approach helps in developing a robust system for analysing thermal images.

Thermal Image Dataset: This represents the collection of thermal images used for training and testing the model. These images capture temperature variations and are essential for developing a robust classification system.

Image Pre-processing: This step involves enhancing image quality by removing noise, normalizing pixel values, and possibly resizing images. Proper preprocessing improves feature extraction and classification accuracy.



10.48047/jocaaa.2025.34.04.51

Fig. 2: Proposed System Architecture

DNN Classification: A Deep Neural Network (DNN) takes the extracted features andclassifies the image into predefined categories, such as low or high state. It improve accuracy by learning complex relationships in the data.

Output Performance Analysis: The final step evaluates the model’s performance using metrics like accuracy, precision, recall, and F1-score. This analysis helps determine the effectiveness of the classification system.

Proposed VGG19 Feature Extraction

VGG19 is a deep convolutional neural network architecture comprising 19 layers: 16 convolutional layers and 3 fully connected layers. It processes input images of size 224×224×3 (RGB) through five sequential blocks. Each block contains multiple convolutional layers with 3×3 filters, followed by ReLU activations, and concludes with a 2×2 max-pooling layer for down sampling. The number of filters doubles after each block, starting from 64 in the first block to 512 in the last two blocks. After the convolutional blocks, the architecture includes three fully connected layers: the first two with 4096 neurons each and the final layer with 1000 neurons, corresponding to the number of classes in the ImageNet dataset, utilizing a SoftMax activation function for classification.

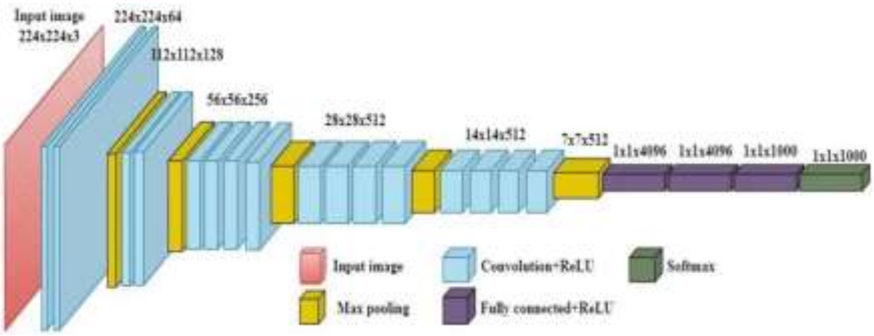


Fig. 3: VGG19 Layer Block Diagram

Proposed Dense Neural Network Classifier

A DNN is also known as a Fully Connected Neural Network (FCNN) or Multi- Layer Perceptron (MLP), is a fundamental machine learning architecture where each neuron in one layer connects to every neuron in the next. Key components include an input layer that receives raw data, multiple hidden layers with fully connected neurons applying transformations via weights and activation functions, and an output layer that produces the final prediction. Activation functions like ReLU, Sigmoid, or Tanh introduce non-linearity, allowing the network to model intricate relationships. The training process involves forward propagation, where data moves through the network to generate an output, and backward propagation, where the network adjusts its weights based on the error of the output compared to the expected result, optimizing performance over time. DNNs are versatile and widely used in applications such as image recognition, natural language processing, and financial forecasting. However, they can be computationally intensive, especially with large datasets, and may not be as efficient as specialized architectures like CNNs for image data or Recurrent Neural Networks (RNNs) for sequential data.

In the VGG-19 architecture, the final layers are designed to process high-level features extracted from preceding convolutional layers. Specifically, after the convolutional blocks, the network includes three fully connected layers: the first two with 4,096 neurons each, and the third with a number of neurons corresponding to the number of target classes commonly 1,000 for datasets like ImageNet. These fully

10.48047/jocaaa.2025.34.04.51

connected layers integrate the spatially distributed features into a comprehensive representation, enabling the network to make informed classification decisions.

The output from the final fully connected layer consists of raw, unnormalized scores known as logits, each representing the model's confidence in a particular class. To transform these logits into a probability distribution across all classes, the SoftMax function is applied. This mathematical function exponentiates each logit and normalizes them by the sum of all exponentiated logits, resulting in values between 0 and 1 that sum to 1. This probabilistic output not only facilitates interpretability but also allows for the application of threshold-based decision rules in various applications, enhancing the model's utility in real-world scenarios.

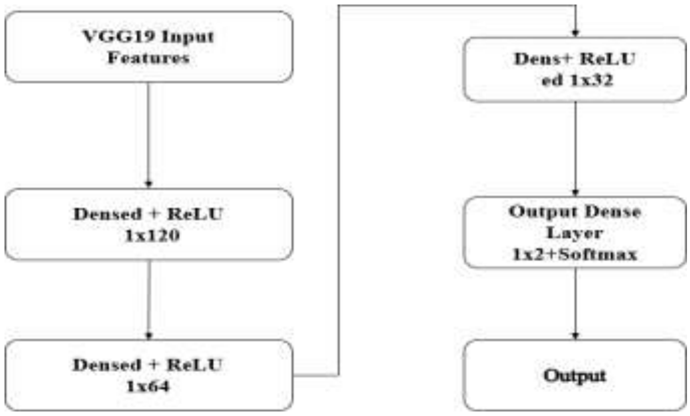


Fig. 4: Dense Neural Network Classifier

4. Results and Discussion

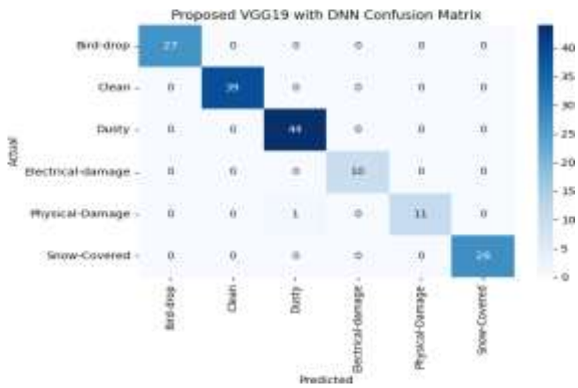


Fig. 5: Proposed Confusion Matrix.

The comparative analysis across Tables 9.2 to 9.7 clearly highlights the performance superiority of the proposed VGG19 with DNN model over traditional machine learning algorithms such as K-Nearest Neighbors (KNN) and Random Forest Classifier (RFC). In the overall classification performance (Table 9.2), KNN demonstrates poor results with an accuracy of only 41.8%, a precision of 0.548, recall of 0.418, and F1-score of 0.355. RFC performs moderately better with an accuracy of 68.7%, precision of 0.723, recall of 0.687, and F1-score of 0.683. In stark contrast, the proposed VGG19 with DNN model achieves a perfect score of 1.0 in all four metrics, indicating highly accurate, precise, and consistent classification across all test samples.

Table 1: Performance Analysis of Various Algorithms

Algorithms	Accuracy	Precision	Recall	F1-Score

Existing KNN	0.418	0.548	0.418	0.355
Existing RFC	0.687	0.723	0.687	0.683
Proposed VGG19 with DNN	1.0	1.0	1.0	1.0

A class-wise comparison further reinforces these results. For the *Bird-Drop* class (Table 9.2), KNN lags behind with a precision of 0.36 and F1-score of 0.47, whereas RFC achieves slightly better scores (precision: 0.61, F1-score: 0.66). The proposed model again scores a perfect 1.0 across all metrics. Similarly, in the *Clean* class (Table 9.3), KNN shows a recall of 0.82 but low precision (0.38), indicating over-prediction of this class. RFC shows balanced scores with precision of 0.59 and F1-score of 0.67, but the proposed model again achieves perfect classification. In the *Dusty* class (Table 9.4), KNN's recall is very low (0.27), making it unreliable, while RFC does better with a 0.63 F1-score. The proposed model reaches 1.0 in all aspects. For *Electrical Damage* (Table 9.5), although KNN has a high precision (1.00), its recall is just 0.40, which leads to a poor F1-score of 0.57. RFC improves recall (0.80), but again falls short of the 1.0 scores obtained by the proposed model.

The challenges become even more evident in the *Physical Damage* class (Table 9.6), where KNN fails completely (0.00 in all metrics), and RFC offers limited recall (0.25) despite high precision. Only the VGG19 with DNN model offers full reliability here. Lastly, in the *Snow-Covered* class (Table 9.7), KNN once again shows extreme disparity between precision (1.00) and recall (0.04), while RFC provides a strong balance (0.90 F1-score), yet remains inferior to the flawless prediction capability of the proposed deep learning model. Overall, the VGG19 with DNN architecture not only excels in general performance metrics but also demonstrates exceptional and consistent behavior across all individual fault categories, validating its robustness and practical applicability in fault detection scenarios.

5. Conclusion

The integration of machine learning-driven thermographic image classification has proven to be a highly effective approach for fault detection in photovoltaic (PV) systems. By leveraging advanced image processing and deep learning techniques, this method enhances the accuracy and speed of identifying defects such as hotspots, cracks, and shading issues. Compared to conventional inspection methods, the automation of fault detection significantly reduces human intervention while ensuring real-time monitoring. This not only improves maintenance efficiency but also minimizes energy losses by promptly addressing faults. As a result, photovoltaic systems can maintain optimal performance and extend their operational lifespan. The application of this technology facilitates predictive maintenance, preventing unexpected system failures and costly repairs. With machine learning algorithms continuously improving, detection models can be refined to achieve even higher accuracy and adaptability across different environmental conditions. Additionally, incorporating large-scale datasets and real-time processing capabilities can further enhance the reliability of automated fault detection systems. By integrating these advancements, solar energy production can become more stable and economically viable. The reduction in manual inspections also contributes to overall cost savings and enhances the scalability of PV monitoring systems.

REFERENCES

[1] Ahmed, Waqas. "Enhancing solar PV reliability with hybrid local features and infrared thermography." Energy Reports 13 (2024): 345-352.

[2] Baltacı, Özge, Zeki Kıral, Konuralp Dalkılınç, and Oğulcan Karaman. "Thermal Image and Inverter Data Analysis for Fault Detection and Diagnosis of PV Systems." Applied Sciences 14, no. 9 (2024): 3671.

10.48047/jocaaa.2025.34.04.51

[3] Thakfan, Ali, and Yasser Bin Salamah. "Artificial-intelligence-based detection of defects and faults in photovoltaic systems: A survey." Energies 17, no. 19 (2024): 4807.

[4] Giovanni, and Mauro Migliazzi. "Infrared thermography monitoring of solar photovoltaic systems: A comparison between UAV and aircraft remote sensing platforms." Thermal Science and Engineering Progress 48 (2024): 102379.

[5] Qureshi, Usamah Rashid, Aiman Rashid, Nicola Altini, Vitoantonio Bevilacqua, and Massimo La Scala. "Radiometric Infrared Thermography of Solar Photovoltaic Systems: An Explainable Predictive Maintenance Approach for Remote Aerial Diagnostic Monitoring." Smart Cities 7, no. 3 (2024): 1261- 1288.

[6] vieira, Thiago M., Ézio C. Santana, Tarso V. Ferreira, and Douglas B. Riffel. "Photovoltaic Cell Defect Detection by Lock-In Thermography Using 2-D Gaussian Profile." IEEE Journal of Photovoltaics (2024).

[7] Belhachat, Faiza, Cherif Larbes, and Rachid Bennia. "Recent advances in fault detection techniques for photovoltaic systems: An overview, classification and performance evaluation." Optik (2024): 171797.

[8] Umar, Shayan, Muhammad Salik Qureshi, and Muhammad Usman Nawaz. "Thermal Imaging and AI in Solar Panel Defect Identification." International Journal of Advanced Engineering Technologies and Innovations 1, no. 3 (2024): 73-95.

[9] Shafique, Usman, Syed Muhammad Alam, Umar Rashid, Wahab Javed, Haris Anwaar, Malik Shah Zeb, Talha Ahmad, Uzair Imtiaz, and Frederic Nzanywayingoma. "Infrared Thermography Based Insulator Fault Classification via Unsupervised Clustering and Semi-Supervised Learning." IEEE Access (2024).

[10] Saiprakash, Chidurala, S. Ramana Kumar Joga, Alivarani Mohapatra, and Byamakesh Nayak. "Improved fault detection and classification in PV arrays using Stockwell transform and data mining techniques." Results in Engineering 23 (2024): 102808.