Received: February 13, 2024. Revised: May 1, 2024. 

INASS

International Journal of
Intelligent Engineering & Systems

http://www.inass.org/

# Enhancing Solar Panel Fault Detection: An Efficient Multidomain Feature Analysis Model with Entropy-Guided Saliency Map Segmentation

**Sujata P Pathak [1]* Sonali A Patil [1] Dhirendra Mishra [2]**

[1]*K J Somaiya College of Engineering, SVU, Vidyavihar, Mumbai, India*
[2]*NMIMS University, Computer Engineering Department, MPSTME, Mumbai, India*
* Corresponding author's Email: sujatapathak@somaiya.edu

**Abstract:** Solar panels are an increasingly popular and sustainable means of generating electricity. However, the efficiency and longevity of solar panels may get compromised by various types of faults, including diode hotspots, dust/shadow hotspots, multicell hotspots, PID hotspots, and single cell hotspots. Detecting these faults accurately is vital for maintaining optimal efficiency. Many existing methods for fault identification fall short due to inadequate feature representation and segmentation techniques. To address these limitations, an innovative approach is proposed involving entropy-based saliency map segmentation and multidomain feature analysis model for fault detection and classification in solar panels. Proposed saliency map segmentation method extracts the most relevant regions in solar panel images, improving fault detection. The entropy-driven saliency maps fault detection technique surpasses alternative approaches such as color thresholding and channel-based thresholding for fault detection in solar panels. A comprehensive set of feature representation models, including Fourier, Wavelet, DCT, Convolutional, and Gabor transformations is employed. To further enhance the precision and effectiveness of fault identification, we incorporate an Extra Trees feature selection mechanism. Classification is done with an ensemble of classification models, including k-Nearest Neighbors (kNN), Deep Forest, Support Vector Machines (SVM), Logical Regression, and Artificial Neural Networks (ANN). Empirical evaluation of the proposed model demonstrates exceptional performance, achieving F1 score of 94% for fault classification compared to existing machine learning models. Proposed multidomain analysis model gave an accuracy of 96.9% and recall of 93.5% in fault identification. Additionally, the proposed model exhibits computational efficiency, making it suitable for real-time fault detection scenarios.

**Keywords:** Entropy, Faults, Features, Hotspots, Saliency maps, Segmentation, Solar panel.

## 1. Introduction

With numerous advantages like lessened environmental impact and long-term cost savings, solar energy has become a well-known and sustainable alternative for the generation of electricity. The main parts of solar energy systems, solar panels, transform sunlight into electrical energy. Solar panels are, however, susceptible to several flaws that over time, reduce their performance and efficiency, just like any other technology. Accurately determining and diagnosing these issues is essential to ensuring optimal energy output and extending the life of solar panel installations. Solar panel fault finding has traditionally relied on manual inspection, which is labor and time-intensive and subject to human error. Researchers are increasingly using automated fault detection systems that make use of computer vision methods and machine learning algorithms to get around these constraints. These systems use images of solar panels to identify and categorize various fault types, allowing for prompt maintenance and repair.

In this study, we present a novel method for fault detection in solar panels that combines saliency map segmentation with a multidomain feature analysis model. The main goal of the proposed model is to improve the precision and effectiveness of fault detection in solar panels while addressing the drawbacks of current approaches [1-3].

The inadequate representation of fault features in current approaches is a major problem. Solar panel faults can appear in a variety of ways, including single cell hotspots, PID hotspots, dust shadows hotspots, multicell hotspots, and diode hotspots. Using various feature representation models, these faults can be captured because they have distinctive visual characteristics. Our suggested method uses a variety of feature representation models, including Fourier, Wavelet, DCT, Convolutional, and Gabor, to address this. These models allow us to accurately identify and depict the distinctive fault patterns visible in solar panel images [4-6]. Effective segmentation of pertinent areas within the solar panel images presents another difficulty in fault detection. Conventional segmentation methods might be unable to precisely isolate the regions where faults are present, resulting in false detections or missed faults. We introduce saliency map segmentation, which identifies the most salient regions within the images, to address this issue. We can increase the precision and effectiveness of fault detection by concentrating the analysis on these regions.

Any fault detection system's effectiveness also heavily depends on the choice of illuminating features. Extra Trees, a potent feature selection mechanism, is used in our proposed model to find the most distinct features from the multidomain representation. This procedure improves fault classification accuracy by reducing dimensionality and removing pointless or redundant features. We train an ensemble of classification models, including kNN, Deep Forest, SVM, Logistic Regression, and ANN, to categorize the chosen fault features into particular fault categories. The fault detection system's robustness and generalization abilities are improved using multiple models. There are several benefits to the suggested multidomain feature analysis model with saliency map segmentation. By utilizing a variety of feature representation models, effective segmentation strategies, and strong feature selection mechanisms, it addresses the shortcomings of existing approaches. The model is highly applicable for monitoring and maintaining solar panel installations, ensuring optimum performance and longevity, due to its ability to precisely identify faults in real-time scenarios.

In conclusion, this paper presents a thorough and original method for locating solar panel faults. Our model significantly increases fault detection efficiency and accuracy by combining multidomain feature analysis with saliency map segmentation. Its superior performance and suitability for use in real-world scenarios are demonstrated by the empirical validation. The suggested model represents a significant advancement in automating solar panel fault detection and has the potential to have a significant positive impact on the solar energy sector for different use cases.

The structure of this document is as follows: section II gives review of existing models used for fault analysis. Section III provides details about the proposed design of an efficient multidomain feature analysis model with saliency map segmentation for identification of solar panel faults. Section IV covers statistical analysis of proposed method. Section V focuses on conclusion and future scope.

## 2. Review of existing models used for fault analysis

Solar panel fault analysis is essential for ensuring optimal performance and durability of solar panel installations. Researchers have developed numerous models and techniques for the detection and identification of faults in solar panels over the years. In this section, we provide a comprehensive analysis of the existing models used for solar panel fault analysis, highlighting their strengths, limitations, and improvement opportunities [7-9].

### 2.1 Image processing methodologies

Methods based on image processing are widely used for solar panel fault analysis. These methods frequently involve image enhancement, noise reduction, and segmentation as pre-processing steps & flows. By analyzing the characteristics of the segmented regions, errors are identified for different use cases [10-13]. In the paper [27], authors have used combination of different pre-processing techniques. Three segmentation techniques using color based thresholding, channel based thresholding and temperature based thresholding are proposed and fault detection accuracy is calculated. IoU values for various faults are calculated which are low. So efficient segmentation techniques must be developed to improve the accuracy of fault detection.

These methods are efficient for identifying obvious flaws such as cracks, discoloration, and physical damage. They provide important information regarding the spatial distribution and severity of faults. Image processing-based approaches frequently struggle to detect faults that are not visually apparent, such as diode hotspots, dust/shadow hotspots, and PID hotspots. They are highly dependent on image quality and may produce false positives or overlook subtle flaws. Inaccurate and dependable fault identification may be hampered by insufficient representation of features. Due to their

ability to automatically learn and extract fault patterns from large datasets, machine learning-based techniques have gained popularity in solar panel fault analysis. These methods involve training classification models on labelled data, where features extracted from images or sensor data are used to train the models [14-16].

Models based on machine learning can identify complex fault patterns and generalize well to unobserved datasets & samples. By learning from a diverse set of features, they can detect both visible and non-visible flaws. These methods can manage large amounts of data, allowing for efficient fault analysis. The performance of machine learning models is heavily dependent on the engineering of features. Choosing informative and pertinent features can be difficult under real-time scenarios. Inability to accurately capture various fault characteristics due to a lack of comprehensive and diverse feature representation models. Particularly when dealing with large datasets, training, and optimizing machine learning models can be computationally intensive process like Adaptive Feature Space Fusion (ASFF) process [17-19].

### 2.2 Hybrid methodologies

Combining image processing and machine learning techniques, hybrid approaches improve the precision and robustness of fault analysis [20]. These methods combine the strengths of both domains to overcome the limitations of individual approaches. Hybrid approaches can effectively integrate visual data from images and numerical data from sensors, thereby enhancing fault detection capabilities. They can utilize sophisticated machine learning algorithms for fault classification and image processing techniques for precise segmentation and feature extraction process. The development of hybrid approaches requires knowledge of both image processing and machine learning, which makes their implementation more difficult for real-time scenarios. The performance of hybrid approaches is highly dependent on the selection and incorporation of suitable image processing and machine learning components [21-23].

Existing models for solar panel fault analysis have made substantial contributions like You Look Only Once (YoLO) & CNN [24-26] to the field, as a conclusion. Image processing-based methods excel at detecting visible flaws, whereas machine learning-based methods can detect both visible and non-visible flaws. The objective of hybrid approaches is to combine the strengths of both domains in order to enhance overall performance. However, there are still obstacles to overcome, such as the accurate segmentation of non-visible flaws, the comprehensive representation of features, and the efficient integration of various techniques. Future research should concentrate on developing hybrid models with improved feature representation and optimization techniques for more precise and dependable solar panel fault analysis.

## 3. Proposed design of an efficient multidomain feature analysis model with saliency map segmentation for identification of solar panel faults

Various models are proposed for identification of solar panel faults, and most of these models are either highly complex, or cannot be used for real-time scenarios. To overcome these issues, this section discusses the design of an efficient multidomain feature analysis model with saliency map segmentation for identification of Solar Panel faults. As per Fig. 1, in the proposed model, the thermal images are pre-processed with Bilateral filter and enhanced with histogram equalization as mentioned in the work [27]. These pre-processed images are passed to the Saliency Maps module to extract high-entropy regions from the Solar Panel images & samples.

### 3.1 Entropy based saliency map detection

These regions consist of high variance regions, which indicate distortions in smooth pixel transitions. Due to which, the Saliency Map Model can locate fault regions with higher efficiency levels. The model utilizes bit-level slicing, colour space conversion, structural & colour feature extraction, feature selection & entropy evaluation for estimation of saliency maps as shown in Fig. 2.

Various models are proposed for identification of solar panel faults, and most of these models are either highly complex, or cannot be used for real-time scenarios. To overcome these issues, this section discusses design of an efficient multidomain feature analysis model with saliency map segmentation for identification of solar panel faults. As per figure 1, the proposed model initially uses Saliency Maps to extract high-entropy regions from the solar panel images & samples. These regions consist of high variance regions, which indicate distortions in smooth pixel transitions. Due to which, the LAB-based Saliency Map Model (LSMM) is capable of locating fault regions with higher efficiency levels. The model utilizes bit-level slicing, colour space conversion, structural & colour feature extraction,

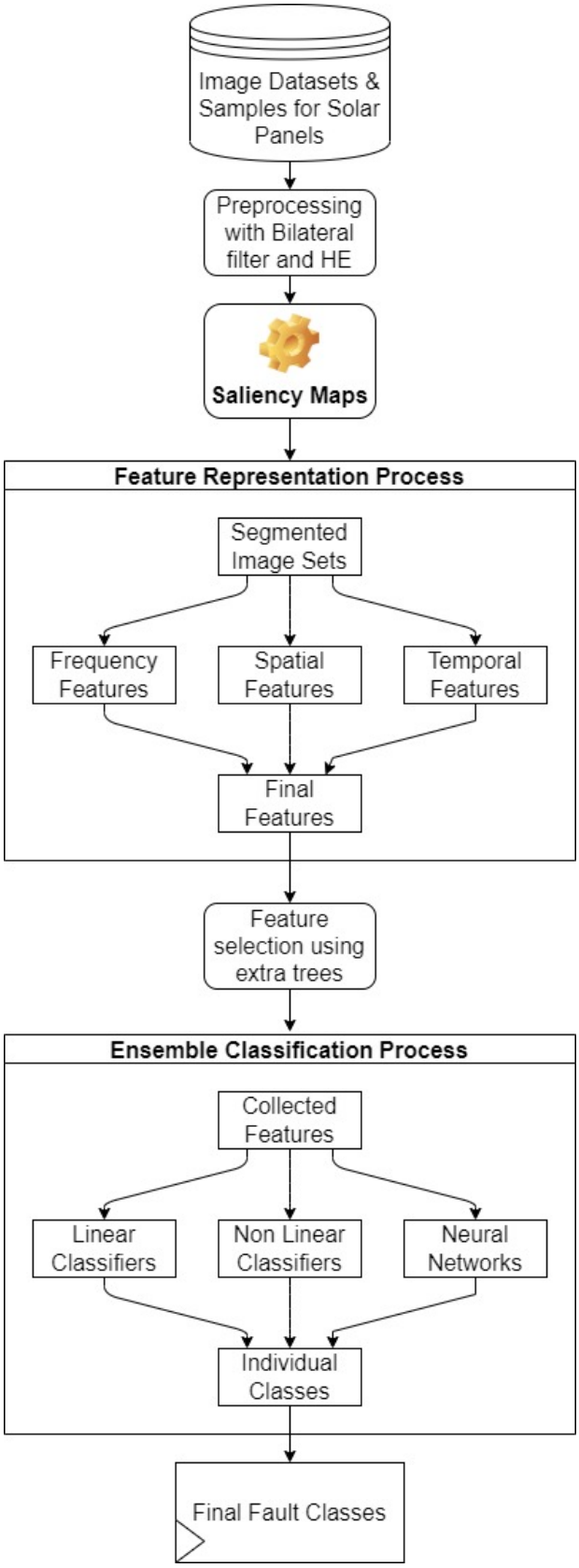



Figure. 1 Design of the proposed model for identification of different solar panel faults

feature selection & entropy evaluation for estimation of saliency maps.

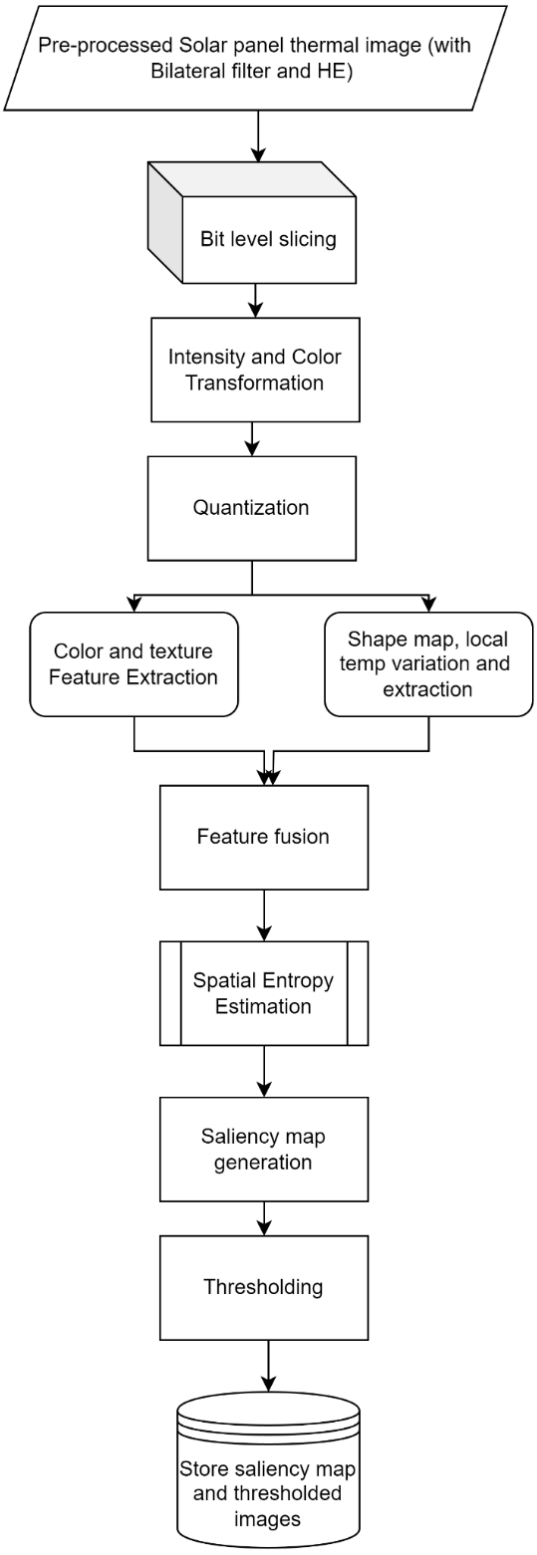



Figure. 2 Design of the saliency map extraction process

To extract bit planes, the input image is processed via Eq. (1),

$$BP(i) = \cup_{r,c}^{N,M}\left(P_{r,c} \gg 2^{i}\right) \quad (1)$$

where, $BP(i)$ represents the output slice for $i^{th}$ bit, while $P_{r,c}$ is the pixel intensity for the $r, c$ row & columns. These slices are extracted for each of the Red (R), Green (G), and Blue (B) channels.
Based on these slices, the LAB levels are estimated via Eqs. (2)-(4) as follows,

$$L* = 116\ (0.2126 \times R + 0.7152 \times G + 0.0722 \times B) - 16 \quad (2)$$

$$a = 500 \times (0.4124 \times R + 0.3576 \times G + 0.1805 \times B - 0.2126 \times R + 0.7152 \times G + 0.0722 \times B) \quad (3)$$

$$b = 200 \times (0.2126 \times R + 0.7152 \times G + 0.0722 \times B - 0.0193 \times R + 0.1192 \times G + 0.9505 \times B - f(Z)) \quad (4)$$

The LAB Map is selected due to its capability of representing the solar panel part of the images with higher entropy levels. Each of these LAB images are averaged via Eq. (5),

$$I_{avg_i} = \frac{1}{N \times M} \times \sum_{r,c}^{N,M} LAB(i) \quad (5)$$

Using this averaged image, the distance between different slices is estimated via Eq. (6),

$$d(s_i, s_j) = \frac{1}{N \times M} \times \sum_{i=1}^{N_s} I_{avg_i} \times \sqrt{\sum_{r,c}^{N,M} \frac{(P_{i_r} - P_{j_r})^2 + (P_{i_c} - P_{j_c})^2}{Var(s_i, s_j)}} \quad (6)$$

where, $d(s_i, s_j)$ represents the level of dissimilarity between $i, j$ pixels; $Var(s_i, s_j)$ is the variance levels between given images, and is estimated via Eq. (7), $N_s$ are total number of extracted slices and $P_{x_y}$ are the intensity levels for pixel $x$, in the $y_{th}$ dimension sets.

$$var(x, y) = \frac{1}{N-1} \times \sum_{i=1}^{N} \frac{\left(x_i - \sum_{j=1}^{N} \frac{x_j}{N}\right)^2}{\left(y_i - \sum_{j=1}^{N} \frac{y_j}{N}\right)^2} \quad (7)$$

Before further using the distance metric, pixel levels are quantized via Eq. (8),

$$LAB_{quant} = {}^{LAB_{in} \times 128}/_{LAB_{max}} \quad (8)$$

where, $LAB_{in}, LAB_{quant}, and\ LAB_{max}$ are the input LAB pixels, their output quantization pixels, and maximum intensity level of the pixels. These pixels are counted via Eq. (9) to form a colour map ($CF(out)$) as follows,

$$CF(out) = \cup_{i=1}^{N_s} \sum_{r,c}^{N,M} \left| LAB_{r,c} == LAB_{quant_{r,c}} \right| \quad (9)$$

Similarly, the shape map is estimated via Eq. (10), where Canny edge features are used to identify probability of edges for different image pixels.

$$SF_{out} = \cup_{i=1}^{N_s} \sum_{r,c}^{N,M} \left| Canny(P_{r,c}, P_{r,c+1}) == 1 \right| \quad (10)$$

where, $SF_{out}$ is the output shape map, and $Canny(P_{r,c})$ is the output of Canny Edge detection process. Both these features are combined into an augmented vector ( $F$ ), and entropy levels are estimated via Eq. (11),

$$E = -\sum_{r=1}^{N} \sum_{c=1}^{M} p\left(F_{r,c_i}\right) \times \log\left(p\left(F_{r,c_i}\right)\right) \quad (11)$$

where, $p\left(F_{r,c_i}\right)$ represents feature probability levels for different bit planes. Based on this entropy, and distance between image pixel levels, an Iterative threshold is estimated via Eq. (12),

$$pth = \sqrt{E \times \frac{\sum d}{R \times C}} \quad (12)$$

Image pixels with values more than $pth$ are marked as foreground pixels and are used for the segmentation process. Results of this process on different images can be observed from Fig. 4.

### 3.2 Feature extraction

All these images are passed through an augmented process of multidomain feature extraction, which assists in identification of Frequency, Entropy, and other patterns. The Frequency Patterns are estimated via Eq. (13), where Fourier Transforms are used for the evaluation process.

$$F_i = \sum_{j=0}^{N-1} x_j \times \left[cos\left(2 \times pi \times i \times \frac{j}{N}\right) - \sqrt{-1} \times sin\left(2 \times pi \times i \times \frac{j}{N}\right)\right] \quad (13)$$

where, N represents total number pixels, while x represents their pixel intensity levels. Then, the

Entropy components are estimated using Discrete Cosine Transform via Eq. (14),

$$F(DCT_i) = \frac{1}{\sqrt{2 \times N_f}} \times x_i \sum_{j=1}^{N_f} x_j \times cos\left[\frac{\sqrt{-1} \times (2 \times i + 1) \times \pi}{2 \times N_f}\right] \quad (14)$$

In contrast, the Convolutional features are extracted by adding Leaky Rectilinear Unit (LReLU) based non-linearities to the segmented images via Eq. (15),

$$Conv(out) = \sum_{a=-\frac{m}{2}}^{\frac{m}{2}} x(i-a) \times LReLU\left(\frac{m+2a}{2}\right) \quad (15)$$

where, m, a are sizes for different window dimensions & stride dimensions, which are varied between 1x4 to 1x128 for extraction of high-density features. The LReLU is represented via Eq. (16) and is used to activate different feature sets.

$$LReLU(x) = l_a \times x, when\ x < 0\ , else\ LReLU(x) = x \quad (16)$$

where, $l_a$ represents an Iterative leaky activation constant, which is used to retain only positive feature sets. Then the Gabor Features are extracted via Eq. (17),

$$G(x,y)_s = e^{\frac{-x`^2+\partial^2*y'^2}{2*\emptyset^2}} \times \cos\left(2 \times \frac{pi}{\lambda} \times x'\right) \quad (17)$$

where, x, y is the pixel Index & Intensity level, while $\partial, \emptyset$ & $\lambda$ represents Angles & Wavelengths, which are used to extract multi angular features. These features are extended using Wavelet components, which assist in identification of Approximate & Detail coefficients via Eqs. (18) and (19),

$$w_{i_{approx}} = \frac{x_i + x_{i+1}}{2} \quad (18)$$

$$w_{i_{detail}} = \frac{x_i - x_{i+1}}{2} \quad (19)$$

### 3.3 Feature selection using extra tree classifier

Extra Trees Classifier, also known as Extremely Randomized Trees, is an ensemble learning method used for classification tasks. It is closely related to the Random Forest algorithm and belongs to the family of decision tree-based ensemble techniques. Extra Trees Classifier has some unique characteristics that distinguish it from traditional decision trees and even Random Forests. Extra Trees selects splits randomly. This high level of randomization helps to reduce overfitting and can sometimes lead to simpler, more interpretable trees. Extra Trees Classifier can provide feature importance scores, which indicate the contribution of each feature to the model's predictions. Training individual trees in an Extra Trees ensemble can be parallelized, making it a suitable choice for large datasets and multi-core processors. Overall, Extra Trees Classifier is a powerful ensemble learning method that can be effective in various classification tasks, particularly when you want to reduce overfitting and obtain reliable predictions.

All these components are cascaded to form an Iterative Solar Panel Fault Feature Vector (ISPFFV), which are classified into different faults using a combination of k-Nearest Neighbors (kNN), Deep Forest (DF), Support Vector Machines (SVM), Logical Regression (LR), and Artificial Neural Networks (ANN) methods.

Using these parameter values, the proposed model is used to classify the ISPFFV features individually, and their responses are fused via Eq. (25) as follows,

$$c(out) = A(DF) \times c(DF) + A(ANN) \times c(ANN) + A(LR) \times c(LR) + A(SVM) \times c(SVM) + A(kNN) \times c(kNN) \quad (24)$$

where A & c are the testing accuracy & output class levels for different classifiers. The final output class (c(out)), is used to identify diode faults, dust shadows, multicell hotspots, PID effects, and single cell hotspots. Results of this classification were evaluated in terms of different efficiency metrics and compared with existing methods in the next section.

## 4. Statistical analysis and comparison

The proposed model uses a combination of novel methods for saliency maps, multidomain features, and ensemble classification operations to identify different solar panel faults. To identify performance of proposed model, it was evaluated on a manually collected dataset of visual & thermal solar panel images, which represented diode faults, dust shadows

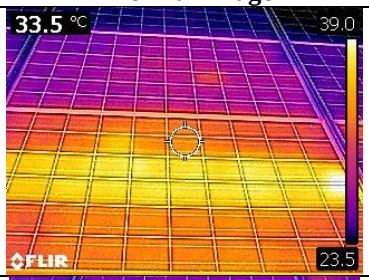
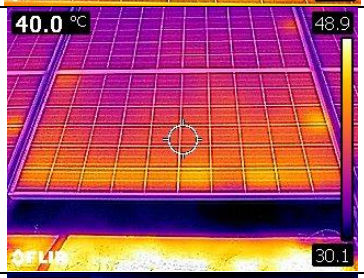
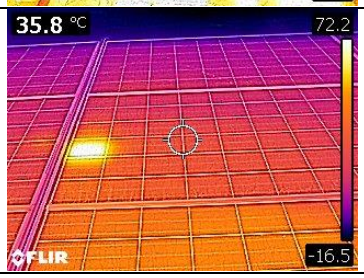
| Type of fault | Thermal Image | Visual Image |
|---|---|---|
| Diode Fault |  |  |
| Dust/Shadow Hotspot |  |  |
| Single cell Hotspot |  |  |



Figure. 3 Sample thermal and visual images

Table 1. Comparison of mean IoU values per fault with proposed saliency map segmentation

| Fault types | Mean IoU | | |
|---|---|---|---|
| | **Channel based thresholding [27]** | **Histogram based color thresholding [27]** | **Proposed saliency approach** |
| Diode | 0.31 | 0.34 | **0.8** |
| Dust | 0.16 | 0.14 | **0.75** |
| Multi | 0.3 | 0.31 | **0.85** |
| Single | 0.43 | 0.53 | **0.87** |
| PID | 0.26 | 0.31 | **0.8** |

hotspots, multicell hotspots, PID hotspots, and single cell hotspots. The dataset consists of nearly 3000 samples, out of which 70% are used for training, 20% are used for validation, and 10% for testing operations. Some of the sample images are shown in Fig. 3.

Quantitative analysis of segmentation using entropy driven saliency maps is done using Intersection over Union (IoU) as a parameter.

$$Intersection\ over\ Union\ (IoU) = \frac{Area\ of\ Overlap}{Area\ of\ Union} = \frac{|A \cap B|}{|A| \cup |B|} \quad (25)$$

For fault detection, bounding boxes of all the ground truth and predicted images are compared and mean IoU values per fault for all images in the dataset are calculated. The resulted mean IoU values are compared with the work mentioned in [27] where authors have applied segmentation using channel thresholding and color thresholding on the same dataset of solar panel thermal images. Segmentation accuracy is calculated using IoU as a parameter. Comparison of these segmentation results with saliency map based segmentation are as shown in Table 1. The proposed saliency-based segmentation approach outperformed the color and channel-based segmentation.

The comparison between Ground Truth and predicted images with bounding boxes is shown in Fig. 4. The pre-processed images are passed through segmentation process using saliency maps and binary image is obtained as show in (c). Using blob analysis, bounding box is plotted on segmented image (d). Column (e) shows ground truth bounding boxes.

Combination of thermal and visual image has increased the accuracy of classification results. The performance of the multidomain analysis model is compared with existing machine learning models. Table 2 shows F1 score for different classifiers with thermal images, visual images, and combination of

Figure. 4 Fault detection using saliency maps

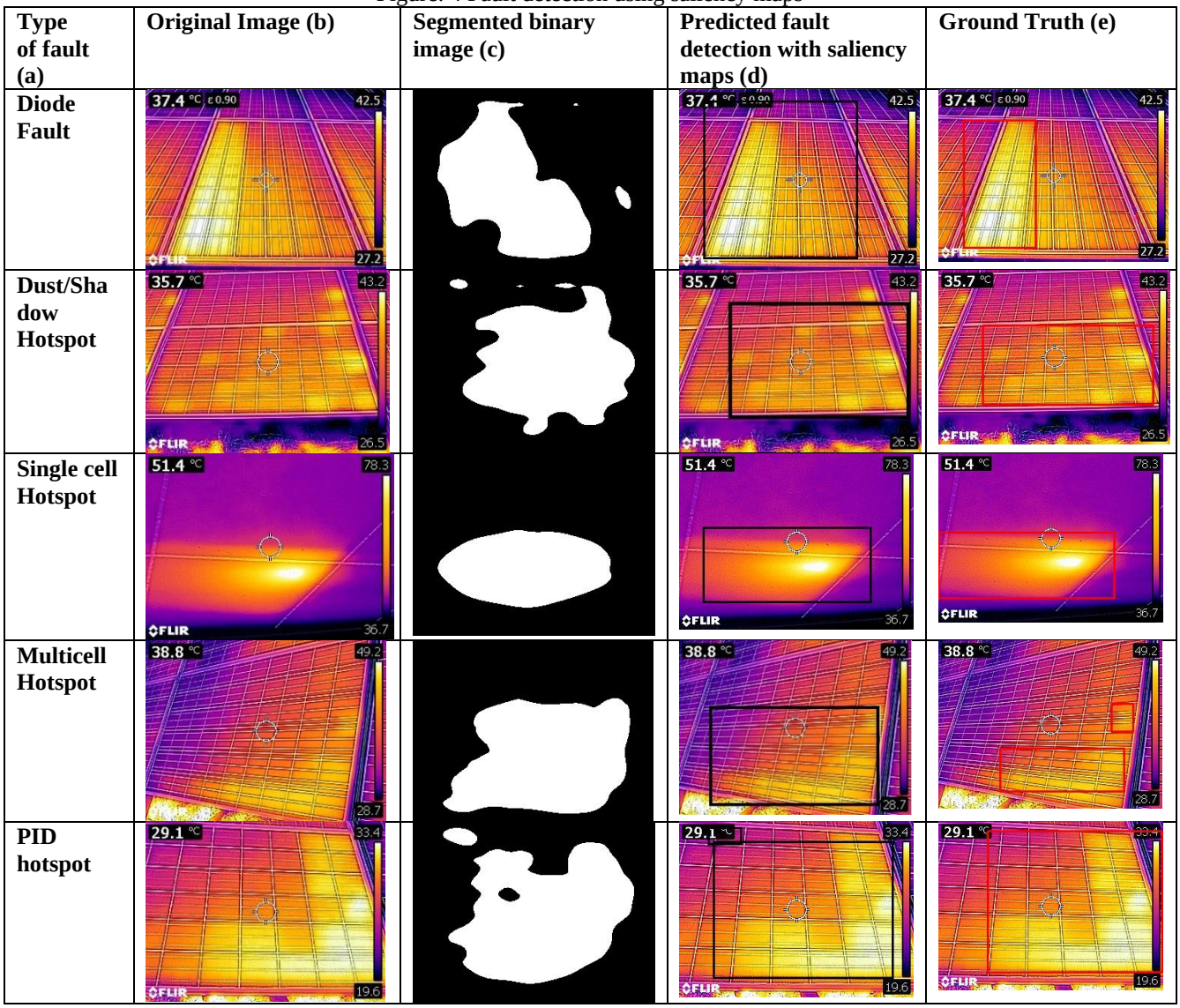



Table 2. Classification results using machine learning with thermal and visual images

| Classifiers | F1 Score | | |
|---|---|---|---|
| | Thermal image | Visual Image | Thermal +Visual image |
| Random Forest | 0.59 | 0.21 | 0.61 |
| kNN | 0.48 | 0.36 | 0.71 |
| SVM | 0.73 | 0.32 | 0.79 |
| Proposed multimodal analysis model | **0.8** | **0.54** | **0.94** |

thermal and visual images. Combination of thermal and visual images gave better F1 score of 94 % using multidomain analysis model compared to existing machine learning models.

The dataset used in this research is not available as a public dataset. So, for comparison, we have taken the model concepts from the papers ASFF-CNN [18], YoLO [25], and CNN [26]. Using these model concepts, we have developed our own models as combination of ASFF and CNN, YoLov3 model and normal deep CNN model. The features extracted from thermal and visual images are applied as input to these models. Then performance of proposed multidomain analysis model was compared with ASFF-CNN [18], YoLO [25], and CNN [26], which are recently proposed classification models, and showcase high performance levels.

The performance of the model was evaluated in terms of Accuracy (A) and delay (d) via Eqs. (26)-(29) as follows,

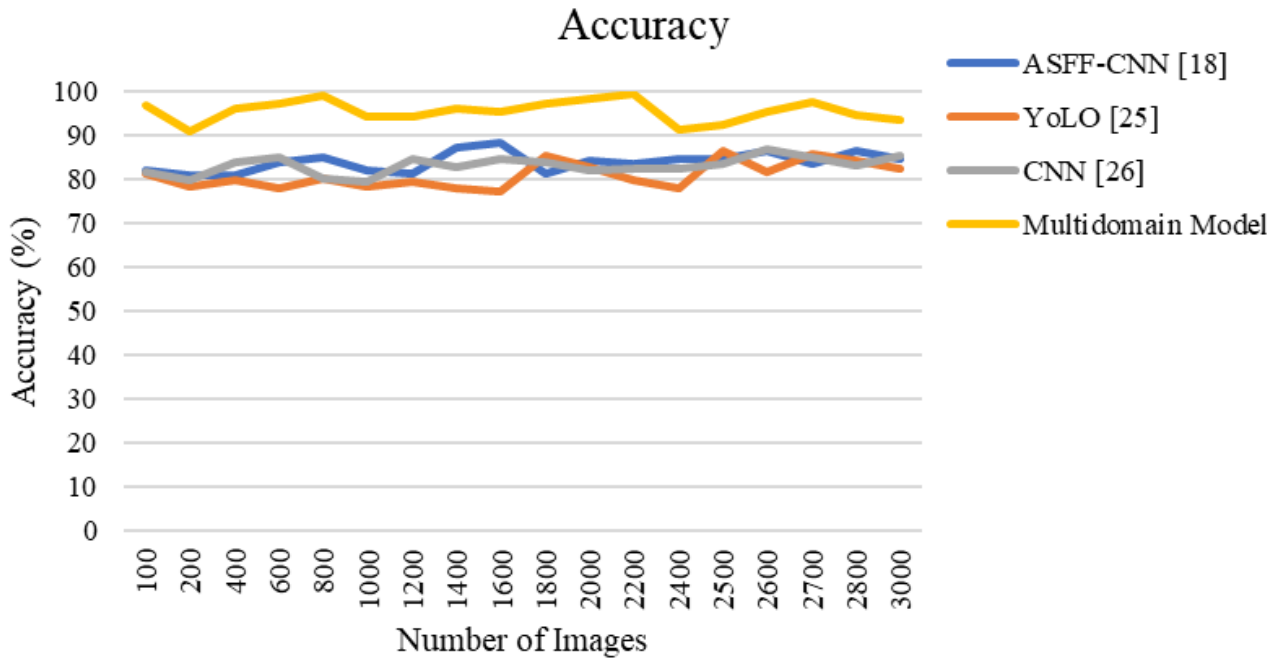



Figure. 5 Accuracy of classification for different samples

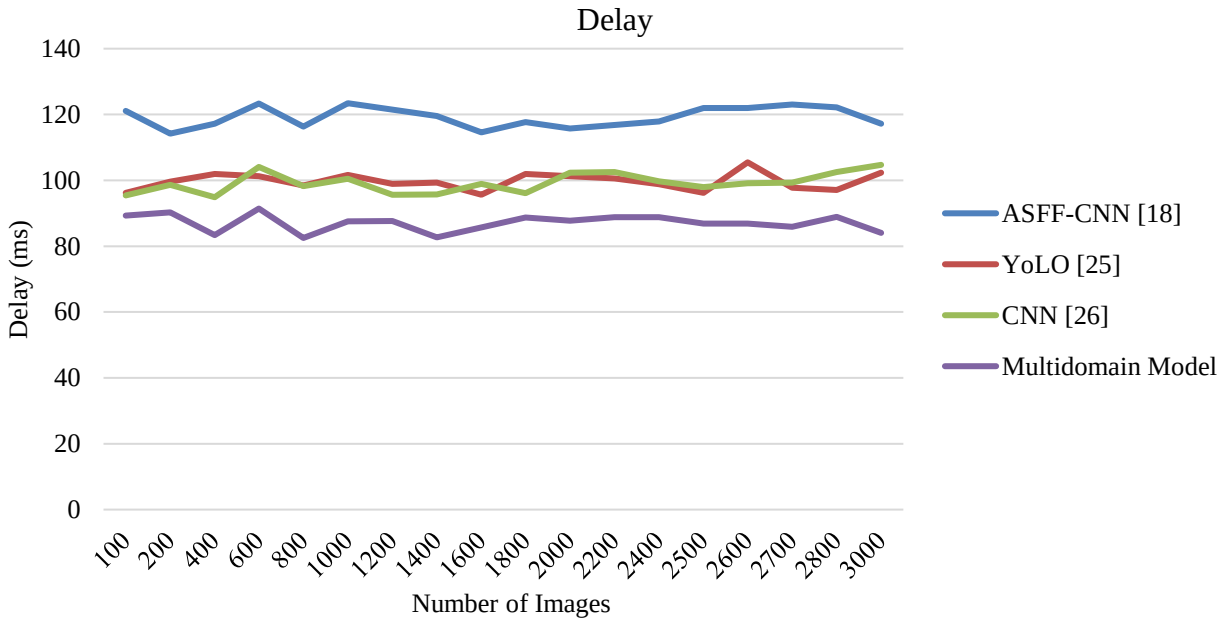



Figure. 6 Delay needed for classification under different samples

$$A = \frac{t_p + t_n}{t_p + t_n + f_p + f_n} \tag{26}$$

$$d = \frac{1}{N}\sum_{i=1}^{N} t_{completed_i} - t_{start_i} \tag{28}$$

where, $t_p$, $f_p$, $t_n$ & $f_n$ are standard values for true & false rates, while t_completed & t_start are the timestamps for completing & starting the classification process for N samples.

Accuracy of classification can be observed from Fig. 5 as follows. From the graph, it was observed that the proposed model showcased 4.5% better accuracy than ASFF-CNN [18], 8.5% better accuracy than YoLO [25], and 9.4% better accuracy than CNN [26], which makes it highly useful for a wide variety of real-time classification scenarios. This is due to the use of high-density features with highly efficient segmentation process, and augmented fusion of ensemble classification techniques.

The performance of the system was evaluated with delay as parameter for classification as shown in Fig. 6.

Based on this evaluation and Fig. 6, it was observed that the proposed model showcased 4.9% faster performance than ASFF-CNN [18], 8.3% faster performance than YoLO [25], and 8.5% faster performance than CNN [26], which makes it highly useful for a wide variety of real-time high-speed classification scenarios. This is due to the use of simplistic variant feature extraction processes, that reduce feature redundancies during classification operations. Due to these enhancements, the proposed model was observed to be useful for a wide variety of solar panel fault classification applications.

## 5. Conclusion and future scope

In this paper, a design for a multidomain feature analysis model with novel saliency map segmentation for identifying solar panel faults is presented. The proposed model is compared by taking concepts from three models, namely ASFF-CNN, YoLO, and CNN, in terms of accuracy, and delay, using exhaustive evaluations. Across all these evaluation metrics, the results unequivocally demonstrated that our proposed model exhibited superior performance. In terms of accuracy, proposed model was 4.5% more accurate than ASFF-CNN, 8.5% more accurate than YoLO, and 9.4% more accurate than CNN. This enhancement is attributable to the use of high-density features in conjunction with a highly efficient segmentation procedure and enhanced fusion of ensemble classification techniques.

Moreover, our model outperformed ASFF-CNN, YoLO, and CNN in terms of delay, achieving a 4.9% faster performance than ASFF-CNN, an 8.3% faster performance than YoLO, and an 8.5% faster performance than CNN. This enhanced speed is advantageous for real-time, high-speed classification scenarios and is accomplished with simple variant feature extraction processes that reduce redundant features during classification operations. Overall, the proposed model offers substantial advancements in the classification of solar panel faults. Its superior precision, accuracy, recall, and speed make it exceptionally useful for a variety of high-consistency and real-time classification scenarios. By utilizing efficient segmentation, high-density features, ensemble classification techniques, and variant feature sets, our model overcomes the limitations of existing models and provides a useful tool for solar panel fault identification.

While our proposed model has demonstrated impressive performance in classifying solar panel faults, there are several avenues for future research and development. Obtaining a larger and more diverse dataset can aid in enhancing the model's generalizability and robustness. By incorporating data from diverse geographic locations, weather conditions, and solar panel types, the model can be trained to handle a wider variety of scenarios. A thorough investigation into the optimal hyperparameters of the model can result in additional performance enhancements. Integrating the proposed model with existing solar panel monitoring systems enables real-time fault detection and preventative maintenance.

## Conflicts of Interest

The authors declare no conflict of interest.

## Author Contributions

Contributions from the authors are: Conceptualization- Sujata Pathak; methodology- Sujata Pathak and Sonali A Patil; software- Mrs. Sujata Pathak; validation- Sujata Pathak, and Sonali Patil, writing- original draft preparation- Sujata Pathak and Sonali Patil; writing—review and editing- Sonali Patil and Dhirendra Mishra; supervision- Sonali Patil and Dhirendra Mishra.

## Acknowledgments

The authors would like to acknowledge PV Diagnostics Pvt. Ltd., Mumbai for providing thermal images of solar panels.

## References

[1] A. Guisandez Hernandez and S. P. Santos, “Modelling and Experimental Validation of Aging Factors of Photovoltaic Solar Cells”, *IEEE Latin America Transactions*, Vol. 19, No. 8, pp. 1270-1277, 2021, doi: 10.1109/TLA.2021.9475857.

[2] M. Uno, Y. Sasaki, and Y. Fujii, “Fault Tolerant Modular Differential Power Processing Converter for Photovoltaic Systems”, *IEEE Trans Ind Appl*, Vol. 59, No. 1, pp. 1139-1151, 2023, doi: 10.1109/TIA.2022.3210074.

[3] P. Xi, P. Lin, Y. Lin,H. Zhou,S. Cheng, Z. Chen, L. Wu., “Online Fault Diagnosis for Photovoltaic Arrays Based on Fisher Discrimination Dictionary Learning for Sparse Representation”, *IEEE Access*, Vol. 9, pp. 30180-30192, 2021, doi: 10.1109/ACCESS.2021.3059431.

[4] M. W. Ahmad, N. B. Y. Gorla, H. Malik, and S. K. Panda, "A Fault Diagnosis and Postfault Reconfiguration Scheme for Interleaved Boost Converter in PV-Based System", *IEEE Trans Power Electron*, Vol. 36, No. 4, pp. 3769-3780, 2021, doi: 10.1109/TPEL.2020.3018540.

[5] U. Kumar, S. Mishra, and K. Dash, "An IoT and Semi-Supervised Learning-Based Sensorless Technique for Panel Level Solar Photovoltaic Array Fault Diagnosis", *IEEE Trans Instrum Meas*, Vol. 72, 2023, doi: 10.1109/TIM.2023.3287247.

[6] M. Uno and K. Honda, "Panel-to-Substring Differential Power Processing Converter With Embedded Electrical Diagnosis Capability for Photovoltaic Panels Under Partial Shading", *IEEE Trans Power Electron*, Vol. 36, No. 9, pp. 10239-10250, 2021, doi: 10.1109/TPEL.2021.3064706.

[7] P.W. David, M. S. Murugan, R. M. Elavarasan, R. Pugazhendhi, O. J. Singh, P. Murugesan, M. Gurudhachanamoorthy, E. Hossain, "Solar PV's Micro Crack and Hotspots Detection Technique Using NN and SVM", *IEEE Access*, Vol. 9, pp. 127259-127269, 2021, doi: 10.1109/ACCESS.2021.3111904.

[8] S. Rao, G. Muniraju, C. Tepedelenlioglu, D. Srinivasan, G. Tamizhmani, and A. Spanias, "Dropout and Pruned Neural Networks for Fault Classification in Photovoltaic Arrays", *IEEE Access*, Vol. 9, pp. 120034-120042, 2021, doi: 10.1109/ACCESS.2021.3108684.

[9] J. M. Guerrero, D. Serrano-Jimenez, K. Mahtani, and C. A. Platero, "A Ground Fault Location Method for DC Systems Through Multiple Grounding Connections", In: *Proc. of IEEE Transactions on Industry Applications*, pp. 7022-7033, 2022, doi: 10.1109/TIA.2022.3199185.

[10] A.S. Edun, C. LaFlamme, S. R. Kingston, H. Tetali, E. J. Benoit, M. Scarpulla, C. M. Furse., J. Harley "Finding Faults in PV Systems: Supervised and Unsupervised Dictionary Learning with SSTDR", *IEEE Sens J*, Vol. 21, No. 4, pp. 4855-4865, 2021, doi: 10.1109/JSEN.2020.3029707.

[11] A. Edpuganti, V. Khadkikar, H. Zeineldin, M. S. El Moursi, and M. Al Hosani, "Enhancing Lifetime of 1U/2U CubeSat Electric Power System with Distributed Architecture and Power-Down Mode", *IEEE Trans Ind Appl*, Vol. 58, No. 1, pp. 901-913, 2022, doi: 10.1109/TIA.2021.3128373.

[12] Y. Leon-Ruiz, M. Gonzalez-Garcia, R. Alvarez-Salas, J. Cuevas-Tello, and V. Cardenas, "Fault Diagnosis Based on Machine Learning for the High Frequency Link of a Grid-Tied Photovoltaic Converter for a Wide Range of Irradiance Conditions", *IEEE Access*, Vol. 9, pp. 151209-151220, 2021, doi: 10.1109/ACCESS.2021.3126706.

[13] C. LaFlamme, E.Benoit , A.Edun , C. M. Furse, P. Kuhn ,M. A. Scarpulla , J. B. Harley , "Quantifying the Environmental Sensitivity of SSTDR Signals for Monitoring PV Strings", *IEEE J Photovolt*, Vol. 12, No. 1, pp. 381-387, 2022, doi: 10.1109/JPHOTOV.2021.3127465.

[14] M. Jalil, H. Samet, T. Ghanbari, and M. Tajdinian, "Development of Nottingham Arc Model for DC Series Arc Modeling in Photovoltaic Panels", *IEEE Transactions on Industrial Electronics*, Vol. 69, No. 12, pp. 13647-13655, 2022, doi: 10.1109/TIE.2021.3128915.

[15] X. Guo and J. Cai, "Optical Stepped Thermography of Defects in Photovoltaic Panels", *IEEE Sens J*, Vol. 21, No. 1, pp. 490-497, 2021, doi: 10.1109/JSEN.2020.3013024.

[16] A.S. Edun, S. Kingston, C. LaFlamme, E. Benoit, M. Scarpulla, C. M. Furse, J. Harley, "Detection and Localization of Disconnections in a Large-Scale String of Photovoltaics Using SSTDR", *IEEE J Photovolt*, Vol. 11, No. 4, pp. 1097-1104, 2021, doi: 10.1109/JPHOTOV.2021.3081437.

[17] A. S. Edun, C. Laflamme, S. R. Kingston, C. M. Furse, M. A. Scarpulla, and J. B. Harley, "Anomaly Detection of Disconnects Using SSTDR and Variational Autoencoders", *IEEE Sens J*, Vol. 22, No. 4, pp. 3484-3492, 2022, doi: 10.1109/JSEN.2022.3140922.

[18] F. M. A. Mazen, R. A. A. Seoud, and Y. O. Shaker, "Deep Learning for Automatic Defect Detection in PV Modules Using Electroluminescence Images", *IEEE Access*, Vol. 11, pp. 57783-57795, 2023, doi: 10.1109/ACCESS.2023.3284043.

[19] S. Harini, N. Chellammal, B. Chokkalingam, and L. Mihet-Popa, "A Novel High Gain Dual Input Single Output Z-Quasi Resonant (ZQR) DC/DC Converter for Off-Board EV Charging", *IEEE Access*, Vol. 10, pp. 83350-83367, 2022, doi: 10.1109/ACCESS.2022.3195936.

[20] B. Hussein, A. M. Massoud, and T. Khattab, "Centralized, Distributed, and Module-Integrated Electric Power System Schemes in CubeSats: Performance Assessment", *IEEE Access*, Vol. 10, pp. 55396-55407, 2022, doi: 10.1109/ACCESS.2022.3176902.

[21] M. Pa, M. N. Uddin, and A. Kazemi, "A Fault Detection Scheme Utilizing Convolutional Neural Network for PV Solar Panels with High Accuracy", In: *Proc. of 1st IEEE Industrial Electronics Society Annual On-Line Conference, ONCON 2022*, 2022, doi: 10.1109/ONCON56984.2022.10126746.

[22] E. El-Fayome, M. A. Abdelhamed, A. El-Shazly, M. Abouelatta, and A. Zekry, "End Of Life Management Of Solar Panels", In: *Proc. of National Radio Science Conference, NRSC, Proceedings*, pp. 286-293, 2023, doi: 10.1109/NRSC58893.2023.10152958.

[23] S. Rao, D. Pujara, A. Spanias, C. Tepedelenlioglu, and D. Srinivasan, "Real-time Solar Array Data Acquisition and Fault Detection using Neural Networks", In: *Proc. of 2023 IEEE 6th International Conference on Industrial Cyber-Physical Systems, ICPS 2023*, 2023, doi: 10.1109/ICPS58381.2023.10128030.

[24] M. Patil, B. P. Kumar, and N. Karuppiah, "Detection and Location of Faults in Photo VoltaicSystems", In: *Proc. of 2023 2nd International Conference on Smart Technologies and Systems for Next Generation Computing (ICSTSN)*, pp. 1-62023, doi: 10.1109/ICSTSN57873.2023.10151647.

[25] S. H. Han, T. Rahim, and S. Y. Shin, "Detection of faults in solar panels using deep learning", In: *Proc. of 2021 International Conference on Electronics, Information, and Communication, ICEIC 2021*, 2021, doi: 10.1109/ICEIC51217.2021.9369744.

[26] N. Kellil, A. Aissat, and A. Mellit, "Fault diagnosis of photovoltaic modules using deep neural networks and infrared images under Algerian climatic conditions", *Energy*, Vol. 263, p. 125902, 2023, doi: 10.1016/J.ENERGY.2022.125902.

[27] S. P. Pathak and S. A. Patil, "Evaluation of Effect of Pre-Processing Techniques in Solar Panel Fault Detection", *IEEE Access*, Vol. 11, pp. 72848-72860, 2023, doi: 10.1109/ACCESS.2023.3293756