

# Advancements in Photovoltaic Panel Fault **Detection Techniques**

# **Junyao Zheng**

Duke University, Durham, North Carolina, USA Email: jzheng6657@gmail.com

How to cite this paper: Zheng, J.Y. (2024) Advancements in Photovoltaic Panel Fault Detection Techniques. Journal of Materials Science and Chemical Engineering, 12, 1-11. https://doi.org/10.4236/msce.2024.126001

Received: April 21, 2024 Accepted: June 11, 2024 Published: June 14, 2024

Copyright © 2024 by author(s) and Scientific Research Publishing Inc. This work is licensed under the Creative Commons Attribution International License (CC BY 4.0). http://creativecommons.org/licenses/by/4.0/ •

**Open Access** 

# Abstract

This paper examines the progression and advancements in fault detection techniques for photovoltaic (PV) panels, a target for optimizing the efficiency and longevity of solar energy systems. As the adoption of PV technology grows, the need for effective fault detection strategies becomes increasingly paramount to maximize energy output and minimize operational downtimes of solar power systems. These approaches include the use of machine learning and deep learning methodologies to be able to detect the identified faults in PV technology. Here, we delve into how machine learning models, specifically kernel-based extreme learning machines and support vector machines, trained on current-voltage characteristic (I-V curve) data, provide information on fault identification. We explore deep learning approaches by taking models like EfficientNet-B0, which looks at infrared images of solar panels to detect subtle defects not visible to the human eye. We highlight the utilization of advanced image processing techniques and algorithms to exploit aerial imagery data, from Unmanned Aerial Vehicles (UAVs), for inspecting large solar installations. Some other techniques like DeepLabV3+, Feature Pyramid Networks (FPN), and U-Net will be detailed as such tools enable effective segmentation and anomaly detection in aerial panel images. Finally, we discuss implications of these technologies on labor costs, fault detection precision, and sustainability of PV installations.

## **Keywords**

Photovoltaic Panels, Fault Detection, Deep Learning, Image Processing

# **1. Introduction**

Photovoltaic (PV) panels have been growing in popularity as an energy source over the past two decades. Their prevalence, particularly for large scale energy generation, has been increasing as pressures to develop green energy sources increase and the impacts of climate change on human health and society become more obvious. As energy producers build larger solar panel complexes, improving the efficiency of these fields is critical. Such PV panel complexes typically require employees to survey the fields on a regular basis to determine if any of the panels are experiencing faults. Fault detection is a process in which the surveyor must evaluate the power producing character of the solar panel and assess its efficiency. If a panel is deemed to be faulty or underperforming, maintenance is them conducted to improve the panel's function.

As the size of solar fields increase, the need for more high throughput methods for fault detection evaluation becomes even more evident. In recent years, computer scientists have begun applying image processing and deep learning to establish methodologies for fault detection in solar fields [1]. Here, we present the various strategies employed by the field and describe the mathematical underpinnings of these strategies. Though these strategies are primarily proof of concept in nature, energy producers are beginning to adopt these techniques into their practices. We envision that this review will serve as an introduction to the scientific community and to the PV industry to the variety of computational technologies available for fault detection.

## 2. Machine Learning Approaches in Fault Detection

## 2.1. Selecting a Template

Within the solar power industry, the most common techniques for fault detection are fairly archaic in the age of big data. Typically, I-V curves and other electrical parameters of PV panels are manually analyzed and fault detection is determined based on subjective parameters, introducing room for error and utilizing significant amounts of time. The application of machine learning, especially kernel-based extreme learning machines and support vector machines, to this problem of accurate fault detection is, therefore, sufficiently motivated [1] [2]. These models learn from historical data captured on PV systems. These models can then be used to interpret real-time data from PV panels and identify anomalous patterns, thereby reducing the rate of false negatives that may be captured during manual I-V curve analysis. Machine-learning algorithms can thereby provide insights on PV systems that require maintenance and minimize the downtime of that system. In most cases, the best models are trained on data that incorporate a variety of PV fault types that capture the spectrum of possible system errors [3].

The efficiency of PV systems is influenced by the electronic health of individual cells or modules. Any one of these cells may experience a number of flaws, such as shading, soiling, and physical damage; all of which can result in substandard performance and significant energy loss over long periods of time. To overcome these challenges in PV fault detection, machine learning approaches have started gaining ground by aiding in accurate fault detection and fault type classification using electrical characteristic analysis of PV system components, in particular, the current-voltage (I-V) characteristics [4].

## 2.2. I-V Curve Analysis for Fault Detection

The electrical performance of a PV cell can be captured in its I-V curve. Notably, these characteristic behaviors can be influenced by environmental conditions and internal cell properties. The typical I-V curve of a healthy PV cell can be modeled by the single-diode equation:

$$I = I_{ph} - I_0 \left[ \left( \exp \frac{V + I \cdot R_s}{n \cdot V_t} \right) - 1 \right] - \frac{V + I \cdot R_s}{R_{sh}}$$
(1)

Here, *I* is the current,  $I_{ph}$  is the photo-generated current,  $I_0$  is the diode reverse saturation current, *V* is the output voltage,  $R_s$  is the series resistance, *n* is the ideality factor,  $V_t$  is the thermal voltage, and  $R_{sh}$  is the shunt resistance [4]. Faults in the PV system can alter the shape and characteristics of the I-V curve significantly. Simple algorithmic models can be designed to calculate deviations in observed I-V curves relative to the theoretical model. However, simply calculating deviations from the theoretical model alone may not capture experimental conditions. Therefore, machine learning models which can take processed data that has been accurately labeled as operating properly or experiencing a fault to derive features that are predictors of fault behavior.

Typical data processing approaches that optimize for the utilization of machine learning approaches involve extracting features from the I-V curves that are potentially indicative of different types of faults [5]. Such features might include parameters like the short-circuit current, open-circuit voltage, maximum power point, fill factor, or I-V curve shape/concavity. These features can then be used to train a machine learning classifier, such as a Support Vector Machine (SVM) or a Neural Network (NN), to distinguish between healthy and faulty conditions in new data samples from PV systems of unknown operational status. Focusing in on the SVM, these types of models aim to find the hyperplane that best separates the feature space into classes representing different health statuses of the PV system. The decision function for a binary classification SVM can be expressed as:

$$f(x) = \operatorname{sign}\left(\sum_{i=1}^{N} \alpha_{i} y_{i} K(x_{i}, x) + b\right)$$
(2)

Here, K would be an input function, a and y would be weights, and b would be a bias term. By optimizing the feature weights, the SVM model can classify the I-V characteristics into healthy or faulty operational status. Further models can be developed to subdivide faulty operation into shaded, damaged, soiled, etc. with more granular labeling of the input data.

#### 2.3. Training and Validation

The effectiveness of the machine learning model is highly dependent on the quality and quantity of the training data. Typically, a dataset comprising several thousand I-V curves, each labeled with the correct fault category, is required for

training a robust model. The data is usually split into a training set (e.g., 80% of the data) and a test set (e.g., 20% of the data) to assess the model's performance. Key performance metrics include accuracy, precision, recall, and the F1 score.

## 3. Deep Learning Enhancements in PV Fault Analysis

Deep learning—a subset of machine learning—has been particularly impactful for advancing PV panel fault detection. Making use of convolutional neural networks (CNNs) like "Efficientb0," deep learning algorithms have proved to work well for the analysis of infrared images of solar modules to detect defects [6] [7]. Convolutional neural networks enable the handling of complex data structures and the derivation of predictive models from that data. These trained network models can identify complex patterns, codified in a series of data processing layers and weights, to identify abnormalities within panel images that correlate with operational status. Often times, as with other use cases of deep learning models, these insights or classifications would escape the human eye or remain undiscoverable with standard machine learning techniques. The typical architecture of a CNN used in image analysis, such as EfficientNet-B0, comprises several layers.

#### 3.1. Input Layer

The input layer takes an image data in multidimensional arrays (encoding pixel position and pixel values). Here, we utilize a model involving an input array with 3 dimensions (an x and y directional value of the pixel of interest and a pixel depth/value). Here, H, W, and D represent the length of the image height, image width, and pixel values, respectively.

$$X \in \mathbb{R}^{H \times W \times D} \tag{3}$$

#### **3.2. Convolution Layers**

These layers apply several parallelized filters F to the input layer to extract features from the data. Similarly, X is the input function, Y is the output,  $\sigma$  is a scaling factor, and  $b_k$  are a set of bias terms.

$$Y_{k} = \sigma \left( \sum_{i=1}^{f} \sum_{j=1}^{f} F_{ijk} \cdot X_{(s \cdot i, s \cdot j)} + b_{k} \right), k = 1, 2, \cdots, K$$
(4)

The main mathematical operation of a convolutional layer is discrete convolution-a method in which the filter is convolved over the width and height of the input array. At each position, a dot product is computed between the filter coefficients and the part of the image in the overlap. This operation is repeated on a sliding scale over the entire image, systematically grouping local patches of pixels to generate a feature map. This transform is designed to derive image features using small subsets of pixels from the input data while preserving the spatial relationship between pixels. In many cases, these convolutional layers incorporate a bias term that is added to the result to provide the model with an extra degree of freedom to shift the resulting activation function. The mathematical analog of this approach is a linear regression in which we fit individual slopes to a given independent-dependent variable pairing but add a "y-intercept" or constant term to account for any shifts of variables relative to each other. The activation function, typically a Rectified Linear Unit (ReLU), can convert these linear outputs of the convolution layers to non-linear outputs (for example, sigmoidal shapes) to capture more complex relationships between features and labels.

## 3.3. Pooling Layer

Pooling reduces the dimensionality of each feature map but retains the most important information. Max pooling is a common choice in the field:

$$P_k = \max\left(Y_k\right) \tag{5}$$

Here,  $P_k$  is the output and  $Y_k$  are the input layers. Pooling layers down sample the feature maps produced by convolutional layers to reduce their dimensionality and to create an aggregated summary of the features detected. The most common form is max pooling, where only the maximum value in a certain area (pooling window) is retained. This process is equivalent to blurring algorithms that retain a single value for a segment of a given image and report that averaged value as the output. This process reduces the sensitivity of the output to the exact location of features in the input, enforcing translation invariance and lowering the computational burden for subsequent layers.

#### 3.4. Fully Connected Layer

The final layers of the network, where all neurons are connected to all activations in the previous layer, are typically used for classification (e.g., operational status). Once again, W is the scaling factor, P is the output of pooling layer, and c is the bias term.

$$Z = \operatorname{softmax} \left( W \cdot \operatorname{flatten} \left( P \right) + c \right)$$
(6)

Fully connected layers take the high-level features extracted by convolutional and pooling layers and translate them into the final output, such as classification scores. These layers flatten the two-dimensional feature maps into a one-dimensional vector that acts as input. A weight matrix then transforms this vector into another vector of raw scores, one for each class in the classification task. A bias vector is typically added, like the bias correction seen in the convolutional layers, to these scores. Then, a *softmax* function converts them into normalized probabilities, with the highest probability indicating the model's prediction for that classification. The *softmax* function exponentiates each raw score and then normalizes by dividing by the sum of all exponentiated scores. This operation ensures that the output of the fully connected layer is a valid probability distribution, which quantifies the network's confidence in each potential class for the input image.

# 4. Image Processing Techniques for Fault Identification

The shift towards image processing techniques has provided a new dimension to

PV fault detection. High-resolution aerial images captured by unmanned aerial vehicles (UAVs) equipped with infrared sensors offer a comprehensive view of large solar installations. This is particularly relevant to large scale energy production farms in which energy loss in any one solar panel can impart significant financial losses and output losses to the operating firm. Advanced image processing algorithms, including DeepLabV3+, Feature Pyramid Networks (FPN), and U-Net, have been adapted to process these aerial images, and identify panel defects in new aerial images [8]-[15]. This approach not only enhances the accuracy of fault detection but also significantly reduces the time and labor costs associated with manual inspections performed by purveyors or other employees.

#### 4.1. Segmentation with U-Net and the Dice Coefficient

U-Net is structured for effective segmentation of images, especially useful in identifying small anomalies in large images such as PV panels. However, it is often recommended to use a statistic like the Dice coefficient to assess the effectiveness of the U-Net algorithm in properly segmenting the image at hand [16].

The Dice coefficient is used to evaluate segmentation performance, particularly handling unbalanced data where the area of interest (faults) is much smaller compared to the background. In these situations, the signal to noise ratio is low and therefore segmentation checks need to be performed to ensure that information is not lost. The Dice coefficient is calculated as follows:

$$\text{Dice} = \frac{2 \times |X \cap Y|}{|X| + |Y|} \tag{7}$$

Here, X and Y are two samples that we wish to compare. The Dice coefficient is a similarity measure between two samples that are regularly used in statistics, image processing, and other related computational analysis. In the context of image processing, these samples are usually binary images or segmentation maps. This measures the similarity among segmentations. In practice, Dice coefficients are considered an important measure of performance in applications where accurate image segmentation is especially critical to model performance, such as in medical imaging. High values of the Dice coefficient indicate good performance of the model, *i.e.*, the algorithm used for segmentation accurately captures the desired features. Lower scores would suggest that a review in the model or methodology is necessary to identify the source of discrepancies between the prediction and ground truth.

### 4.2. Feature Pyramid Networks (FPN) for Object Detection

FPNs utilize an architecture with lateral connections to generate an accurate model for object detection or classification over multiple scales. The network boosts feature extraction at every scale which is critical in the context of PV characterization where image arrays may span small or large plots of land [12]. FPN architecture works both bottom-up and top-down pathways in parallel to

effectively make use of semantic information being extracted at every scale. This is a remarkable advantage for the application of fault detection in solar panels because shapes and sizes can vary tremendously across various arrays. The bottom-up processing pathway enables processing of input images at increasing levels of abstraction, while the top-down processing pathway refines information by using higher-resolution features from the lower levels. With this, FPNs can form lateral connections between corresponding levels of these two pathways, enriching the semantic value of the high-resolution features.

Researchers can train the FPN on a dataset of aerial images of solar panels where faults have been labeled at varying scales. Adaptive wavelet analysis makes it possible for the network to work at different resolutions. The result is that not only small cracks but also large, damaged areas, which may emerge on different feature sets on their own, can be detected by the same model. Multi-scale feature extraction of FPN gives improved performances over single-scale models, which have the disadvantage of generating false positives or false negatives if the model is trained on a skewed dataset.

#### 4.3. DeepLabV3+ for Semantic Segmentation

DeepLabV3+ applies atrous, or dilated, convolution to capture the multi-scale context of input data and classifications by adopting multiple atrous rates, which aids in capturing spatial hierarchies in input images [6]. Atrous convolution involves inserting unitary holes into the convolutional filters to increase the field of view. By modulating these unitary hole sizes, the model can control the resolution at which feature responses are computed. Therefore, it can generate feature weights at various scale for classification.

$$Output[i] = \sum_{k} Input[i + r \cdot k] \cdot Filter[k]$$
(8)

The equation above defines the output of the operation of atrous convolution. The output, defined at position i of the input feature map, is calculated by summing over the element-wise product of the input feature map and the filter weights. The dilation rate r scales the index k in determining the step by which to sample the input feature map. This allows the filter to increase its field of view without increasing the number of parameters or the computational cost.

DeepLabV3+ is an advanced network neural network architecture that has been proposed for semantic segmentation with the goal of labeling every pixel of the image with a class. Semantic segmentation models are very relevant for applications such as autonomous driving, medical image analysis, or landscape categorization. In each of these applications, multi-scale context emerges as an important criterion to capture the semantics of the segmentation as they involve features of interest in images only notable at different scales or at several resolutions. Atrous convolution allows DeepLabV3+ to perform multi-scale context aggregation by modifying the dilation rate, which specifies how many pixels in a kernel should be skipped while performing convolution. This makes the model adaptive in such a way that it can control the level of the computed feature responses and results, thereby effectively capturing spatial hierarchies. This allows the exponential growth of the receptive field without loss of either resolution or coverage, one of the most important advantages of atrous convolution. This demonstrates the advantageous nature of such convolutional networks when segmenting large objects in such a way that the details of the smaller objects are preserved. It is a powerful tool that guarantees feature maps of high resolution are maintained throughout the network. In practice, the atrous convolution is applied as a module within DeepLabV3+ architecture and is known as the Atrous Spatial Pyramid Pooling (ASPP). ASPP probes an incoming feature set using filters at multiple sampling rates and effective fields of views to capture objects and image context in multiple scales. In practice, this really needs to be carefully weighed against the trade-off of receptive field sizes and feature map resolutions. It is, therefore, upon the research team to make model decisions that ensure that the selected rates of atrous convolution are in alignments with the scales of the objects of interest in the dataset, and thus, optimizing segmentation performance.

# 5. Data Mining and Artificial Intelligence in PV Monitoring

The integration of artificial intelligence (AI) techniques and data mining in photovoltaic (PV) monitoring systems marks significant progress in the development of renewable energy management systems [2] [3] [4] [7] [15] [16]. Application of such advanced computational technologies enables fault detection pinpointing in PV systems and prediction of probable failures prior to such occurrences possible. With such tools, proactive maintenance of PV systems has become possible. In the PV monitoring context, data mining offers an important way to filter and process enormous operational data recorded in data loggers and collected from solar panels. This data includes electrical parameters, such as voltage, current, and temperature, in addition to the environmental parameters, such as sunlight intensity and meteorological conditions. Advanced data mining algorithms are then applied for the purpose of identifying useful correlations and patterns. For instance, the use of clustering algorithms may be helpful in grouping PV cells or modules based on performance or degradation similarities. This helps in the isolation of particular cells or units with the greatest likelihood of failure.

Faults prediction, specifically through the use of machine learning models, is fundamental to shifting the PV monitoring paradigm from a reactive approach to proactive maintenance strategies. Models such as neural networks, decision trees, and ensemble methods like random forests have been used to forecast failures with high accuracy levels. These models are trained on historical performance data and are capable of processing non-linear relationships and complex interactions between various electrical and environmental parameters. For example, a neural network could be developed to forecast the chances of inverter failure by learning from patterns of fluctuating voltage levels and high surrounding temperatures. These proactive maintenance deployments avoid infrastructure downtimes and additional maintenance costs.

#### **Real-Time Monitoring and Anomaly Detection**

Real-time monitoring is an area of particular interest in which machine learning models can prove their benefit [2] [7] [16]. Anomaly detection algorithms are designed to incorporate machine learning predictions while continuously scanning data streams of PV installations. Based on these data updates, the algorithm can predict the likelihood of any given PV cell experiencing a fault or an imminent fault. Such models can capture outlier data points that might predict the onset of a fault condition, so the operating firm can address maintenance issues at the onset of or prior to failure. Furthermore, reinforcement learning (RL) can be implemented to dynamically update model weights and parameters based on real-time data and additional manual labeling to allow the model to shift its predictive capabilities over time to reflect the most recent pool of data on PV systems.

This adaptive element provided by reinforcement learning serves as a contrast to the traditional static anomaly detection systems. RL models learn from new datasets encoding novel or shifting relationships between features and classifications to improve decision making over time. These models do so by optimizing to receive an encoded reward by scoring algorithms if classifications are made accurately on new sets of data. This approach has been applied effectively in environments like industrial water treatment facilities, showcasing its potential in practical applications. These positive outlooks of RL in anomaly detection, however, are not without challenges: "The fully labeled data needed for RL training is a challenge unto itself, but more than that, one still needs to design and develop a robust model that can work effectively on real-world complexities". However, RL's pliant nature enables the continuous model improvement and adaptation that might bring about a solution to these limitations over larger time ranges. RL has gained support in its implementation for the detection of anomalies in view of network security and industrial systems. For example, experiments carried out on datasets like the SWaT Dataset from the Singapore University of Technology and Design have demonstrated the potential of RL in extracting anomalies from complex, multivariate time-series data.

#### 6. Future Directions and Challenges

Despite these notable advances in machine learning approaches for PV system monitoring and fault detection, several challenges remain. First, the computational cost of large data sets is of big concern and raises a question, especially for deep learning models. Future research should work towards optimizing these algorithms for faster processing times and greater accuracy within the consideration of real-world application constraints. The development of standardized datasets for training and testing algorithms could also enhance model reliability across different geographic locations and conditions. The further integration of internet-of-things (IoT) devices, smart sensors within the PV systems, and human factors will open the doors to increasingly dynamic and responsive monitoring solutions [2] [3] [7] [16].

# 7. Conclusion

In summary, this study has reviewed the significant development of machine learning methods for the detection of faults in photovoltaic (PV) panels, from the traditional manual methods to the array of advanced machine learning and deep learning techniques currently available. The integration of machine learning has, therefore, proven to be a very important tool in the analysis of electrical parameters, such as the I-V curves. Such models are capable of increasingly accurate, more efficient, and real-time fault detection. In addition, the advancement of deep learning, especially through the application of convolutional neural networks like EfficientNet-B0, has furthered the ability of identifying subtle faults on the solar panels using infrared imaging as an alternative model input. These technological improvements are more precisely able to detect faults and have the added advantage of making PV systems sustainable and durable by reducing energy wastage and enabling timely maintenance.

# **Conflicts of Interest**

The author declares no conflicts of interest regarding the publication of this paper.

## References

- Islam, M., Rashel, M.R., Ahmed, T., *et al.* (2023) Artificial Intelligence in Photovoltaic Fault Identification and Diagnosis: A Systematic Review. *Energies*, 16, 7417. <u>https://doi.org/10.3390/en16217417</u>
- [2] Duranay, Z.B. (2023) Fault Detection in Solar Energy Systems: A Deep Learning Approach. *Electronics*, **12**, 4397. <u>https://doi.org/10.3390/electronics12214397</u>
- [3] El Banby, G., Moawad, N.M., Abouzalm, B.A., et al. (2023) Photovoltaic System Fault Detection Techniques: A Review. Neural Computing and Applications, 35, 24829-24842.
- [4] Jamuna, V., Muniraj, C. and Periasamy, P.S. (2023) Fault Detection for Photovoltaic Panels in Solar Power Plants by Using Linear Iterative Fault Diagnosis (LIFD) Technique Based on Thermal Imaging Systems. *Journal of Electrical Engineering & Technology*, 18, 3091-3103.
- [5] Sundaram, K.M., Padmanaban, S., Holm-Nielsen, J.B. and Pandiyan, P. (2023) Photovoltaic Systems: Artificial Intelligence-Based Fault Diagnosis Approaches. *Renewable Energy*, **49**, 213-219.
- [6] Jumaboev, S., Jurakuziev, D. and Lee, M. (2022) Photovoltaics Plant Fault Detection Using Deep Learning Techniques. *Remote Sensing*, 14, 3728. <u>https://doi.org/10.3390/rs14153728</u>
- [7] Pierdicca, R., Paolanti, M., Felicetti, A., et al. (2020) Automatic Faults Detection of

Photovoltaic Farms: solAIr, a Deep Learning-Based System for Thermal Images. *Energies*, **13**, 6469. <u>https://doi.org/10.3390/en13246496</u>

- [8] Prabhakaran, S., Uthra, A. and Preetharoselyn, J. (2023) Comprehensive Analysis of Defect Detection through Image Processing Techniques. *IEEE Transactions on Industrial Informatics*, 14, 3621-3632.
- [9] Sonawane, P., Jog, P. and Shete, S. (2019) A Comprehensive Review of Fault Detection & Diagnosis in Photovoltaic Systems. *IOSR Journal of Electronics and Communication Engineering*, 14, 31-43.
- [10] Mohanapriya, V., Sharmila, B. and Manimegalai, V. (2021) Classification and Detection Techniques of Fault in Solar PV Systems. *IEEE Transactions on Sustainable Energy*, 7, 841-850.
- [11] Joseph, E., Singh, B.S.M. and Ching, D.L.C. (2023) Developing a Simple Algorithm for Photovoltaic Array Fault Detection Using MATLAB/Simulink Simulation. *International Journal of Energy Production and Management*, 8, 235-240. <u>https://doi.org/10.18280/ijepm.080405</u>
- [12] Lu, F., Niu, R., Zhang, Z., *et al.* (2022) A Generative Adversarial Network-Based Fault Detection Approach for Photovoltaic Panel. *Energies*, **12**, 1789. <u>https://doi.org/10.3390/app12041789</u>
- [13] Kaack, L.H. (2021) Machine Learning Enables Global Solar-Panel Detection. Nature, 586, 213-218. <u>https://doi.org/10.1038/d41586-021-02875-y</u>
- [14] Chen, K., Huang, C. and He, J. (2016) Fault Detection, Classification, and Location for Transmission Lines and Distribution Systems: A Review on the Methods. *High Voltage*, 1, 25-33.
- [15] Dhanraj, J.A., Mostafaeipour, A., Velmurugan, K., *et al.* (2021) An Effective Evaluation on Fault Detection in Solar Panels. *Energies*, **14**, 7770.
- [16] Voutsinas, S., Karolidis, D., Voyiatzis, I., *et al.* (2023) Development of a Machine-Learning-Based Method for Early Fault Detection in Photovoltaic Systems. *Journal of Engineering and Applied Science*, **70**, Article No. 27. https://doi.org/10.1186/s44147-023-00200-0