International Journal of Engineering Trends and Technology
ISSN: 2231–5381 / https://doi.org/10.14445/22315381/IJETT-V72I8P103

Volume 72 Issue 8, 18-26, August 2024

Original Article

Solar Panel Fault Detection Using Low Complex Convolution Neural Network Deep Learning Model

P. Sampurna Lakshmi[1], S. Sivagamasundari[2], Manjula Sri Rayudu[3]

[1,2]Department of Electronics and Instrumentation Engineering, Annamalai University, Chidambaram, Tamil Nadu, India.
[3]Department of Electronics and Instrumentation Engineering, Vallurupalli Nageswara Rao Vignana Jyothi Institute of Engineering &Technology (VNRVJIET), Hyderabad, Telangana, India.

[1]Corresponding Author : sampurnalakshmip7@gmail.com

Received: 06 January 2024 Revised: 03 May 2024 Accepted: 31 July 2024 Published: 28 August 2024

Abstract - *The use of Photovoltaic (PV) systems to collect energy from the sun has emerged as a viable option for meeting the world's increasing energy demands while reducing dependency on fossil fuels. At the core of these systems are solar panels, which convert sunlight into power. However, like other technical equipment, solar panels are susceptible to defects and failures. Recently, solar panel defect detection has become essential for ensuring the effective and reliable operation of PV systems. This paper presents a solar panel fault detection model using deep learning. We propose a low-complexity Convolutional Neural Network (CNN) consisting of Convolution 1D, activation, max pooling, and dense layers. The 1D CNNs automatically extract relevant features from the input data, detecting patterns in various positions of the input sequence. Low-complexity CNNs have fewer parameters and memory requirements, which is crucial for devices with limited resources. The proposed model achieved a fault detection accuracy of 98%.*

Keywords - *Convolutional Neural Network (CNN), Deep Learning, Photovoltaic (PV) Systems, Low complex, Solar panel defect detection, Solar panels.*

1. Introduction

Solar panels have long been acknowledged for their importance as a fundamental component in the development of renewable energy due to their enormous power output capacity [1]. The fact that sunlight may be passively used to collect solar energy, which can then be transformed into electricity, demonstrates the ever-present relevance of solar energy in the energy landscape. Recent developments in solar panel technology, which are being driven by continual breakthroughs and inventions, have resulted in an increase in the panels' efficiency as well as an expansion of the applications for which they may be used. The last few years have seen a proliferation of game-changing advancements in the field of solar panel technology. These developments include innovations in terms of materials, manufacturing procedures, and design approaches. Solar panels have had their energy conversion efficiency greatly enhanced as a result of these advancements, which has increased their desirability and made them more cost-effective for a wider variety of applications [2]. In recent years, a wide variety of solar cell varieties have been available, some of which include monocrystalline solar cells, polycrystalline solar cells, thin-film solar cells, and tandem solar cells. Solar panels are able to more effectively adapt to a wide variety of geographical regions and installation settings because of the fact that each of these cell types provides a distinct set of benefits. Installing Photovoltaic (PV) systems on a worldwide scale has been propelled in part by the fast uptake of solar energy as a source of electricity that is both clean and renewable [3]. Solar panels have a well-deserved reputation for requiring little in the way of upkeep, but that does not mean they are immune to flaws or failures. The performance of solar panels may gradually deteriorate over time due to a number of causes, including exposure to adverse weather, flaws in the manufacturing process, and general wear and tear [4]. These difficulties may vary from slight decreases in energy output to full panel failure, and they have the potential to result in considerable economic losses for the owners of PV systems as well as a decline in the sustainability and dependability of the generation of solar energy.

The identification of faults in solar panels [5] is an essential component in finding solutions to these problems. Researchers and industry experts want to quickly and effectively discover and diagnose problems with solar panels by making use of cutting-edge technology and methodologies, such as data analytics, thermal imaging, and sophisticated monitoring systems. This preventative strategy not only guarantees that Photovoltaic (PV) systems continue to function without interruption, but it also helps increase energy output, prolong the lifetime of solar panels, and cut down on the expenses of maintenance. The use of machine learning is one of the most promising approaches that may be taken to improve the identification of faults in solar panels [6]. When trained on enormous datasets of solar panel performance and defect data, machine learning algorithms may become skilled at recognizing subtle patterns and abnormalities that may evade human detection. This is because humans tend to focus on more obvious patterns and anomalies. They are able to do real-time analysis of the data that is gathered from a wide variety of sensors and monitoring equipment, which enables them to quickly locate issues such as hotspots, cracks, and electrical mismatches. Machine learning models have the potential to enhance their accuracy and efficiency in problem detection over time if they continually learn and adapt to their environments [7].

In addition, machine learning may make it possible to implement predictive maintenance procedures, which include the identification of prospective defects before these errors create severe disruptions. These models may predict when specific components of a solar panel system are likely to decay or fail by examining previous data and taking into consideration a variety of environmental parameters. This enables timely maintenance and prevents expensive downtime. The application of machine learning to the problem of defect detection in solar panels has a great deal of potential for the realization of the objective of improving the dependability and efficiency of photovoltaic systems [8]. Not only can we more precisely diagnose flaws by utilizing the power of data-driven algorithms and real-time monitoring, but we can also maximize energy output, increase the lifetime of solar panels, decrease maintenance costs, and contribute to a more sustainable and dependable energy future. As we work our way through the complexity of this important topic, it is becoming more clear that problem detection in solar panels using machine learning is not simply a technical effort but rather an essential step towards a greener and more reliable energy environment.

2. Literature Survey

Mahmoud Dhimish et al. [9] introduced a unique approach for detecting faults in photovoltaic (PV) bypass diodes. The algorithm has three primary stages. Initially, the threshold voltage of the current-voltage (I-V) curve is determined by evaluating various failure bypass diode situations. Furthermore, the identification of defective areas in bypass diodes is determined by the examination of voltage drop inside the current-voltage (I-V) characteristic, together with the voltage at the highest power point. Hosna Momeni et al. [10] present a complete approach for the identification, classification, localization, and rectification of defects. The approach under consideration is evaluated by expanding the diagnostic space of the graph-based semi-supervised learning algorithm and using a larger set of class labels. Once the kind and location of a defect have been identified, the system proceeds to temporarily isolate the issue in order to continue functioning without interruption until it is completely rectified. The issue pertaining to the overlapping of cell data in both normal and fault-prone modes may be effectively addressed by the use of distinct normalization techniques.

Manju Santhakumari et al. [11] offered a thorough examination of the impact that environmental conditions have on the many elements of solar systems. The study places significant attention on environmental conditions, including dust accumulation, ambient temperature, wind velocity, humidity levels, snowfall, hailstorms, and sandstorms. These elements have been shown to have a detrimental impact on the energy efficiency of solar plants. Additionally, the study examines the many failure mechanisms of solar panels that may be attributed to these environmental factors. V S Bharath Kurukuru et al. [12] investigate the challenges associated with current and voltage measurements in the context of online monitoring. Additionally, it proposes a series of suggestions for the implementation of quantitative model-based monitoring systems that rely on electrical measurements of current, voltage, and power. This study centers on the characterization of errors in voltage and current measurements of Photovoltaic (PV) cells and investigates their influence on the temperature-dependent output power, voltage, and current of the cells.

Siva Rama Krishna Madeti et al. [13] created A cost-effective and all-encompassing wireless monitoring system using Zigbee technology to facilitate the online monitoring of various configurations of Photovoltaic (PV) arrays. The system incorporates a defect detection approach to enhance its functionality. Various electrical parametric features were used to examine the vulnerability of different frequently employed Photovoltaic (PV) array topologies to partial shading and electrical faults. A predictive model is used to calculate various characteristics based on a specified combination of operational factors, namely solar irradiance and photovoltaic module temperature. The anticipated characteristics are then contrasted with those obtained from field measurements, leading to the detection of possibly faulty operational states. Furthermore, the development of a user-friendly online application is now underway with the aim of providing expedient internet accessibility to monitored data.

Jingyue Wang et al. [14] presented a unique approach to defect detection of solar modules using heterogeneous ensemble learning techniques. The proposed method utilizes current-voltage characteristic curves and ambient variables as key inputs for the diagnostic process. Furthermore, a thorough selection technique is used to screen base learners in order to achieve better diagnostic performance by taking into account both accuracy and variety. The elements of the optimum integration are included by using the probabilistic method and stacking algorithm, respectively. To test the efficacy of the suggested methodology, two sets of data were acquired from a laboratory experiment platform and its matching simulation model, respectively.

Imran Hussain et al. [15] presented a fuzzy logic algorithm that introduces an innovative approach for the identification and categorization of problems in solar photovoltaic (PV) systems. In addition, the approach being presented incorporates fault indexing as a performance metric, which quantifies the extent of divergence from the typical operational state of the solar system. The current-voltage trajectories and their derived properties serve as distinct indicators for each fault situation.Zhicong Chen et al. [16] presented a novel approach for intelligent fault detection and diagnosis in photovoltaic arrays. The method utilizes a newly developed deep residual network model, trained using the adaptive moment estimation deep learning algorithm. This model is capable of automatically extracting features from raw current-voltage curves, as well as ambient irradiance and temperature data. By employing a deeper network architecture, the proposed method effectively enhances the performance of fault detection and diagnosis in photovoltaic arrays.

3. Proposed Methodology

The use of machine and deep learning methods for the identification of faults in solar panels using V-I (Voltage-Current) data is a significant and pioneering endeavour within the realm of renewable energy. The V-I dataset, which encompasses the voltage and current attributes of a solar panel, comprises significant insights into the panel's operational efficiency. Through the use of machine and deep learning algorithms, it becomes feasible to identify diverse forms of flaws or irregularities inside the solar panels. These defects include a variety of difficulties, including partial shade, soiling, electrical faults, and probable hardware malfunctions. The algorithms undergo training using a varied dataset of V-I curves, whereby each curve reflects the performance of the panel under varying situations.The model acquires knowledge of the typical operational characteristics, so facilitating the identification of deviations from the anticipated V-I curve patterns, which might potentially indicate the presence of a problem. The use of a proactive fault detection strategy may assist operators and maintenance teams of solar plants in rapidly identifying difficulties, hence resulting in enhanced energy output, minimized periods of inactivity, and financial savings.One notable benefit of using V-I data and machine/deep learning techniques for solar panel problem identification is its capacity to discern subtle and intricate anomalies that may not be readily discernible via conventional inspection methodologies. Continuous monitoring of solar panels using V-I data analysis has the potential to facilitate early identification and timely repair, so assuring the attainment of optimum energy production and the long-term durability of the panels. Furthermore, the incorporation of these approaches into real-time monitoring systems facilitates the implementation of automated warnings and actions in response to the identification of defects. The increasing usage of solar energy has prompted the exploration of machine and deep learning techniques in the analysis of voltage-current (V-I) data. These advanced methods are expected to have a significant impact on enhancing the dependability and effectiveness of solar power production systems.

3.1. CNN

A Convolutional Neural Network (CNN) is a specialized kind of artificial neural network that has been specifically developed to handle the processing and analysis of data. The model in question is a deep learning architecture designed to replicate the functionality of the human visual system. It does this by using a hierarchical arrangement of linked layers, which enables it to autonomously acquire and extract intricate characteristics and patterns from unprocessed pixel input. Convolutional Neural Networks (CNNs) use convolutional layers to perform scanning and extraction of local features, while simultaneously maintaining the spatial correlations within the input. Additionally, pooling layers are employed to downsample the data and effectively decrease its dimensionality. This architectural design has shown significant efficacy in many tasks including as classification, object identification, and picture segmentation, Described here in Figure 1.

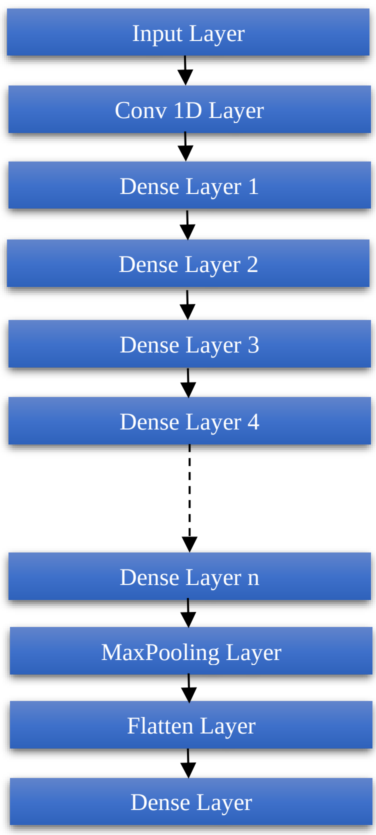

Fig. 1 Proposed CNN model structure

3.1.1. Input Layer

The input layer of a neural network functions as the primary interface for interacting with external data. The system comprises a collection of neurons, with each neuron being associated with a distinct characteristic or input dimension. The main function of the input layer is to receive and send the unprocessed data or characteristics from the dataset that is being processed. These features have the ability to represent a wide range of qualities, including but not limited to pixels in an image, words in a text document, or other pertinent data points. The neurons in the input layer transmit the received data to the successive layers of the neural network, where intricate calculations and transformations are conducted to extract significant patterns and facilitate predictions or classifications. The dimensionality of the input data determines the size of the input layer, which is a crucial factor in shaping the design and operation of the neural network. The input layer establishes the foundation for the flow and processing of information throughout the model.

3.1.2. Convolutional Layer

Convolutional Neural Networks (CNNs) are constructed using convolutional layers, which serve as the essential components. The entity in question consists of a collection of kernels or filters that have the potential to be acquired via the process of learning. The filters in question are small windows operating in two dimensions, which undergo a convolution process over the input image in order to extract localized features. The convolution technique is very advantageous in capturing patterns across many spatial scales, including edges, textures, and shapes.

3.1.3. Dense Layer

The Dense layer, sometimes referred to as a fully linked layer, is a crucial constituent of artificial neural networks. The layer is composed of several neurons, with each neuron establishing connections to all neurons in the preceding layer, resulting in a highly interconnected layer. Within this particular layer, every individual neuron carries out a computation involving the weighted summation of the inputs it receives, followed by the application of an activation function in order to generate an output. The cumulative contributions of the outputs from all neurons in the Dense layer determine the final prediction or output of the network. Dense layers play a vital role in acquiring detailed patterns and representations within data, and they are often used in the latter phases of neural network structures, such as feedforward neural networks, to capture and model complex interactions among elements effectively.

3.1.4. Max Pooling Layer

The Max Pooling layer is an essential element inside Convolutional Neural Networks (CNNs) that serves the purpose of extracting features and reducing dimensionality. The operation is performed by traversing a tiny window, often with dimensions of 2x2 or 3x3, over the input data, which is often feature maps generated by preceding convolutional layers. Within each window, the largest value is chosen. This procedure effectively decreases the spatial dimensions of the data, facilitating the extraction of the most significant characteristics while concurrently reducing computing complexity. The use of Max Pooling contributes to the attainment of translational invariance, a property that enables the identification of an object's existence in various places within the input as the same feature. This characteristic renders Max Pooling an essential component in Convolutional Neural Network (CNN) designs, particularly for tasks such as image recognition, where the preservation of spatial hierarchies of features has significant importance.

3.1.5. Flatten Layer

The Flatten Layer has significant importance in neural networks, particularly in the domains of deep learning and Convolutional Neural Networks (CNNs). The main purpose of this function is to convert input data with several dimensions into a one-dimensional array or vector. The smooth integration of convolutional layers, which handle spatial data like pictures, with fully connected layers, which need one-dimensional input, is of utmost importance. The Flatten Layer is a component that efficiently restructures the data by transforming the feature maps generated by convolutional layers, or any input with many dimensions, into a linear arrangement. This procedure facilitates the retrieval of spatial and hierarchical data from the given input, hence facilitating further processing and learning in future layers. The Flatten Layer is a crucial component in simplifying the transmission of information between different segments of a neural network, hence aiding in the extraction of features and patterns from intricate data structures.

3.2. Network Pruning

Network pruning is crucial to the proposed CNN model. Systematically removing network connections or neuron units improves model efficiency and reduces processing needs. The proposed CNN model optimizes its design by network pruning. Simplifying the model by removing less important connections or neurons speeds up inference and reduces memory use. Network pruning reduces overfitting and improves model generalization. Pruning connections or neurons is usually based on their weight magnitudes or model performance. During the pruning process, the network is iteratively evaluated to identify and delete redundant or less useful parts, producing a more compact, efficient neural architecture.

The proposed CNN model relies on network pruning to balance computational efficiency and model accuracy. It streamlines and simplifies deep learning model deployment in diverse applications by removing duplication and fine-tuning the architecture, improving performance and saving computational costs. The framework of network pruning in the proposed CNN model is depicted in Figure 2.

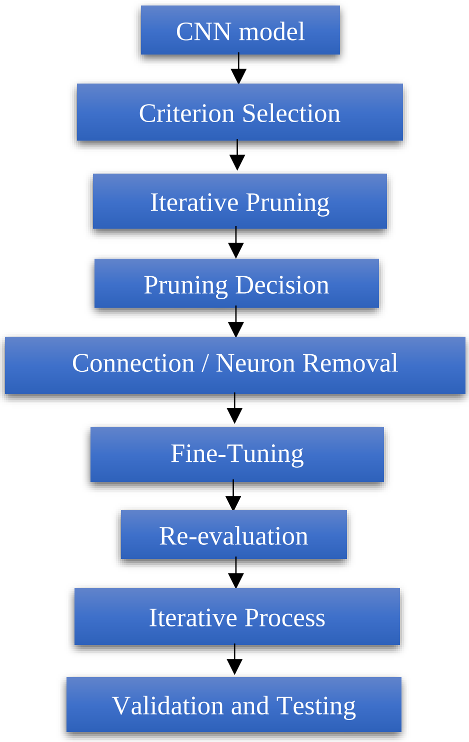

Fig. 2 Network pruning in proposed CNN model

Network pruning in a CNN model reduces network size while preserving performance to enhance efficiency. The following steps are involved in Network pruning:

1. CNN Architecture: Begin with a CNN model that has either already been trained or is currently being trained. This model may include a high number of connections and parameters. The performance of this model is often satisfactory, although it may be resource-intensive to run.
2. Criteria Selection: - Choose a criterion, or a combination of criteria, to use in order to figure out which connections or neurons need to be pruned. Criteria such as link weights, activation levels, and the significance of each component to the overall performance of the network are examples of common criteria. The amount of the weight being carried is often the major consideration.
3. Iterative Pruning: The process of pruning is an iterative one. At first, none of the connections between neurons or other cells are lost. Assess the significance of each connection and neuron in each iteration by judging them against the criteria that have been chosen. When ranking connections or neurons, the significance score is taken into consideration.
4. Pruning Decision: Decisions are based on significance ratings. Low-scoring connections or neurons should be removed. Pruning thresholds are defined by desired sparsity and size reduction vs. performance retention.
5. Connection or Neuron Removal: Connections or neurons that have been determined to be unnecessary are pruned from the network. Deactivating connections with masks might be permanent or temporary.
6. Fine-tuning: The model must be fine-tuned to regain performance lost via pruning. To adapt, the pruned architecture has the model trained for a few extra epochs. This fine-tuning may avoid major weight fluctuations by adopting a lower learning rate.
7. Reevaluation: The model that has been pruned is reevaluated to check that it continues to retain the required level of performance. If this is the case, further modifications, such as modifying the pruning criteria or thresholds, could be implemented.
8. Iterative Process: Steps 3 to 7 are iterated until the required amount of sparsity is reached while conserving performance, and then the process moves on to step 8. This iterative strategy enables a balance to be struck between the decrease in model size and the maintenance of accuracy.
9. Validation and Testing: Once the appropriate amount of sparsity has been achieved, the pruned model is verified and tested with independent datasets to verify that it continues to perform well in real-world circumstances.

Network pruning improves deep learning models for many purposes. It balances computational efficiency and accuracy by eliminating less important connections or neurons.

3.3. Hyperparameter Tuning

The process of tuning the hyperparameters of a proposed CNN model is an essential step that has a substantial influence on the model's overall performance and efficiency. The value of hyperparameter tuning lies in optimizing the parameters that are not learnt during training but greatly impact the network's capacity to learn and generalize. This is where the importance of hyperparameter tuning resides. The performance of the model is heavily dependent on the CNN hyperparameters, which include the learning rate, batch size, number of layers, and kernel size. Accuracy enhancement and acceleration of the convergence process are both attainable goals that may be accomplished by finding the ideal combination of hyperparameters.In the proposed CNN model, hyper parameter known as "sparse categorical cross entropy" is something that has a significant impact on the way in which neural networks are trained. This particular hyperparameter is most often connected with classification jobs, especially when dealing with issues involving many classes. It performs the role of the loss function, which is an essential part of the process of optimizing neural networks.A distinction is made using the "sparse categorical cross entropy" hyperparameter, which is meant to calculate the cross-entropy or log loss between the predicted class probabilities and the actual class labels. "sparse categorical cross entropy" is compatible with integer labels, as opposed to "categorical cross entropy," which demands the target labels to be one-hot encoded vectors. This is because "sparse categorical cross entropy" assumes that each label reflects the actual class for a particular data point.

In actual practice, this hyperparameter is used if one is working on classification projects that include a significant number of different categories. It makes the procedure easier to complete by doing away with the need for one-hot encoding of the target labels, which frees up both memory and computational resources. When designing a neural network, it is essential to choose the right loss function, such as "sparse categorical cross entropy," as one example. It has an immediate and tangible impact on the model's capacity to learn and provide accurate forecasts. Choosing the appropriate loss function may have a major influence on the training dynamics of the network as well as its overall performance.

3.4. Advantages of CNN in Solar Panel Fault Detection

The use of Convolutional Neural Networks (CNNs) in the context of solar panel defect detection, specifically in relation to V-I (Voltage-Current) patterns, has several benefits:

3.4.1. Effective Feature Extraction

The topic of effective feature extraction is being discussed. Convolutional Neural Networks (CNNs) have exceptional proficiency in autonomously acquiring and extracting hierarchical features from many types of input. Within the framework of V-I patterns, this particular skill enables the identification of intricate and nuanced patterns that serve as indicators for various categories of flaws seen in solar panels. The extraction of features plays a critical role in ensuring the precision of defect detection.

3.4.2. Spatial Sensitivity

Spatial sensitivity refers to the ability of a system or model to capture and represent spatial variations or patterns in data accurately. It involves the recognition and consideration of the spatial context, and Convolutional neural networks (CNNs) are specifically designed to collect and analyze spatial correlations within datasets effectively. Spatial patterns in voltage and current levels within V-I patterns often exhibit variances, which may provide significant diagnostic insights. Convolutional Neural Networks (CNNs) have shown a high degree of effectiveness in capturing these patterns and using them for the purpose of defect identification.

3.4.3. Robustness to Noise

The V-I data obtained from solar panels may be subject to noise stemming from various environmental conditions. Convolutional Neural Networks (CNNs) have gained recognition for their capacity to effectively handle noise and distinguish between genuine fault patterns and noise, hence improving the dependability of fault detection.

3.4.4. Scalability

Scalability refers to the capacity of a system or process to handle an increasing workload or accommodate a growing number of users without experiencing Convolutional Neural Networks (CNNs) has the capability to be effectively expanded in order to accommodate extensive datasets, rendering them well-suited for the analysis of voltage-current (V-I) patterns derived from a multitude of solar panels inside a solar farm. The capacity to scale is crucial for the real-time monitoring of the health of many panels.

3.4.5. Minimized Human Intervention

The use of Convolutional Neural Networks (CNNs) for automating the defect detection process leads to a decreased dependence on manual inspection, resulting in reduced human labor and mitigating the possibility of human mistakes.

3.4.6. Early Detection

The concept of early detection refers to the identification and diagnosis of a condition or disease at its earliest stages, often before symptoms become apparent. This proactive approach allows for timely intervention. Convolutional Neural Networks (CNNs) have the capability to identify and recognize defects in their early stages, hence preventing their progression into more severe and consequential problems. Implementing a proactive strategy may effectively reduce the expenses associated with the maintenance or replacement of solar panels.

3.4.7. Adaptability

The concept of adaptability refers to the capacity of an individual or system to adjust and respond effectively to changes in their environment or circumstances. Convolutional Neural Networks (CNNs) provide the capability to undergo fine-tuning or retraining processes in response to the availability of fresh data or changes in the features of solar panels. This ensures that the model remains adaptable to changing circumstances throughout time.

4. Experimental Results

This section is a description of the results obtained from the simulations conducted using the proposed methodology. The dataset used in this study was customized. The dataset underwent processing using the specified technique. The dataset contains several parameters. They are I1, I2, I1MAX, I1MIN, I2MAX, I2MIN, I3, I4, I3max, I3min, I4max, I4min, I5, I6, Itotal1, Vdcmean1, Vdcmax1, Vdcmin1, Pdcmean1, IR, Range1 and Range2. Figure 3 shows the sample data of the data set.

I2	I1MAX	I1MIN	I2MAX	I2MIN	I3	I4	I3max	I3min	...	I5	I6	Itotal
4132	3.776515	3.433125	3.776515	3.433125	3.755050	3.755050	3.782292	3.745653	...	3.755050	3.755050	426.8767
4766	2.611714	2.210978	2.611714	2.210978	2.572122	2.572122	2.614618	2.553622	...	2.572122	2.572122	426.4354
3360	4.101282	3.854098	4.101282	3.854098	4.094577	4.094577	4.106800	4.088955	...	4.094577	4.094577	426.9477
5389	3.285793	2.758149	3.285793	2.758149	3.247759	3.247759	3.289337	3.233321	...	3.247759	3.247759	426.7386
3798	5.455667	0.141519	5.455667	0.141519	5.465640	5.465640	5.472442	5.457335	...	5.465640	5.465640	497.6608
...	...	...	...	...	...	...	...	...	...	...	...	
4300	3.616700	2.806100	3.713500	3.618800	3.571800	3.571800	3.618700	3.560500	...	3.571800	3.571800	426.6700
3200	1.875500	1.563700	1.949200	1.877600	1.873500	1.873500	1.876900	1.871600	...	1.873500	1.873500	426.9400
9700	5.128300	4.992000	5.237100	5.130300	5.129800	5.129800	5.138400	5.121200	...	5.129800	5.129800	468.7800
3400	1.283700	0.151170	1.334700	1.285800	1.277000	1.277000	1.284800	1.274400	...	1.277000	1.277000	426.9500
4700	4.926600	4.222600	5.060500	4.928800	4.874700	4.874700	4.930400	4.858600	...	4.874700	4.874700	446.7200

Fig. 3 Sample data from the dataset

Table 1. Confusion matrix

Fault type	Fault-Free system-Class-0	String-Fault Class-1	String-to-Ground-Fault Class-2	String-to-String-Fault Class-3
Fault-Free system-Class-0	19	1	0	0
String-Fault Class-1	1	30	0	0
String-to-Ground-Fault Class-2	0	0	30	0
String-to-String-Fault Class-3	0	0	0	39

Table 2. Classification report

	Precision	Recall	F1-score
Fault-Free system-Class-0	0.95	0.95	0.95
String-Fault Class-1	0.97	0.97	0.97
String-to-Ground-Fault Class-2	1.00	1.00	1.00
String-to-String-Fault Class-3	1.00	1.00	1.00

Table 1, titled "Confusion Matrix," provides a comprehensive representation of a classification model's performance in categorizing instances across four different classes: Fault-Free System-Class-0, String-Fault Class-1, String-to-Ground-Fault Class-2, and String-to-String-Fault Class-3.The table displays the count of instances in the actual classes versus the predicted classes. For instance, along the diagonal elements, it shows the correctly classified instances, indicating that 19 instances of Fault-Free systems were accurately identified as such, 30 instances of String-Fault Class-1 were correctly classified, and so on. The off-diagonal elements reveal misclassifications, with the intersections between the actual and predicted classes indicating the extent of errors. This confusion matrix is a valuable tool for assessing the model's accuracy and performance in distinguishing between different classes, aiding in the evaluation of its strengths and weaknesses in classification tasks.

Table 2 presents a thorough classification report, which is often used in the assessment of machine learning models, specifically in tasks pertaining to categorization. The table is partitioned into four rows, with each row corresponding to a distinct class or category. Additionally, there are three columns in the table, which reflect precision, recall, and F1-score. The evaluation of these metrics is essential in order to gauge the effectiveness of a classifier. In the first row, denoted as "Fault-Free System-Class-0," the precision, recall, and F1-score exhibit a value of 0.95. Precision is a metric that quantifies the proportion of accurate positive predictions relative to the total number of positive predictions. In the present context, it reveals that 95% of the predictions made for Class-0 were correct.

In contrast, recall is a metric that measures the proportion of correctly identified positive cases out of the total number of real positive occurrences. In this case, the model accurately identified 95% of Class-0 instances. The F1-score is a statistic that quantifies the performance of a classifier by taking into account both precision and recall and is calculated as the harmonic mean of these two measures. This approach ensures a fair assessment of the classifier's performance. An F1-score of 0.95 indicates that the model demonstrates efficient classification of Class-0 examples, achieving a favorable trade-off between accuracy and recall.

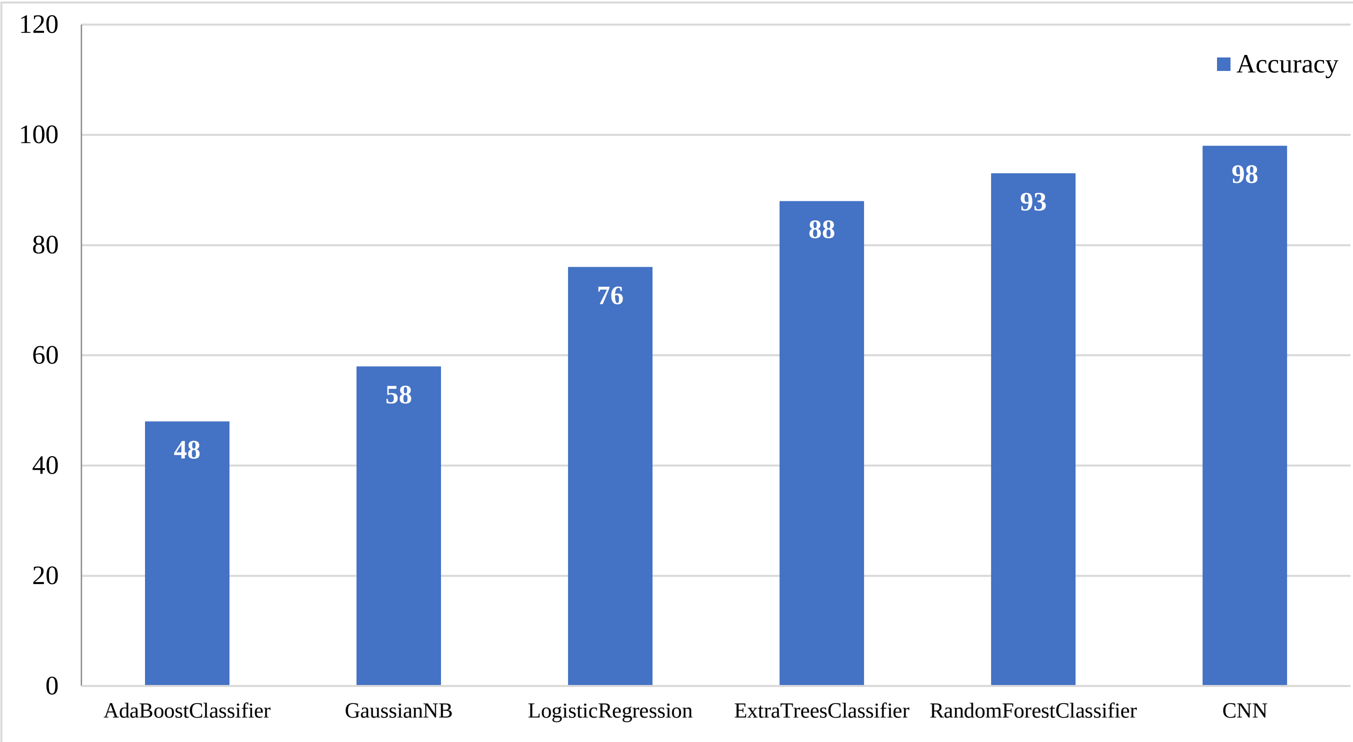

Fig. 4 Comparative analysis

Table 3. Comparative analysis

Method	Accuracy (%)
Ada Boost Classifier	48
Gaussian NB	58
Logistic Regression	76
Extra Trees Classifier	88
Random Forest Classifier	93
Proposed low complex CNN	98

The remaining rows have a similar structure, delineating distinct fault categories. The table presents noteworthy results, indicating good accuracy, recall, and F1-scores of 0.97 or 1.00 for the remaining classes. This suggests that the classifier is exhibiting high performance across all categories, with a notable proficiency in identifying String-Fault Class-1 as well as the more critical fault types such as String-to-Ground-Fault Class-2 and String-to-String-Fault Class-3. In general, Table 2 provides a comprehensive analysis of the classification performance of the model, demonstrating its robust capability to effectively discern various fault classes inside a given system. Table 3, entitled "Comparative Analysis," presents a comprehensive comparison of the performance of several categorization techniques. Each approach is assessed based on its accuracy. The evaluation of classification algorithms often involves the use of accuracy as a widely used statistic. Accuracy measures the percentage of properly categorized occurrences in a given dataset. This table presents an evaluation of six distinct methodologies.

Figure 4 shows the visual representation of comparative analysis. The Ada Boost Classifier had the lowest accuracy, attaining a score of 48, hence showing its proper classification of just 48% of the cases. The Gaussian Naive Bayes (Gaussian NB) algorithm had a somewhat superior performance, with an accuracy rate of 58%. This result suggests that Gaussian NB has a modest level of efficacy in the context of classification problems. The Logistic Regression model had a superior accuracy rate of 76%, indicating a more favourable performance when compared to the two preceding techniques.

The succeeding methodologies, namely Extra Trees Classifier, Random Forest Classifier, and CNN (Convolutional Neural Network) demonstrated superior levels of accuracy, hence suggesting their appropriateness for applications that need precise categorization. The Extra Trees Classifier achieved an accuracy rate of 88%, whilst the Random Forest Classifier exhibited superior performance with an accuracy rate of 93%. The Convolutional Neural Network (CNN) demonstrated the best level of accuracy among the methodologies outlined, with an outstanding rate of 98%.

5. Conclusion

In recent times, there has been a notable increase in the focus on the detection of defects in solar panels. This emphasis is driven by the recognition of its importance in maintaining the dependable and effective functioning of photovoltaic (PV) systems.

This paper presents an innovative methodology for identifying faults in solar panels by employing deep learning methodologies. The proposed strategy utilizes a low-complexity convolutional neural network (CNN) that is made up of dense layers, max-pooling layers, activation functions, and convolution 1D layers.

Convolutional neural networks (CNNs) have the ability to automatically extract relevant features from input data, enabling the detection of patterns at different places within the input sequence. It is essential to balance the benefits of a low-complexity CNN with the task's specific requirements and desired performance.

The results demonstrate the effectiveness of the proposed model, achieving an impressive fault detection accuracy of 98%. This high level of accuracy signifies the potential for the application of deep learning and low-complexity CNNs in addressing the challenges of solar panel defect detection, ensuring the continued reliability and performance of PV systems.

Acknowledgements
Author 1: Draft manuscript preparation, study conception and design, data collection.
Author 2: Study conception and design, data collection.
Author 3: Study conception and design, data collection.

References
[1] Xiang Chen et al., "Research on Real-Time Identification Method of Model Parameters for the Photovoltaic Array," *Applied Energy*, vol. 342, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[2] Chenxi Li et al., "A Fast MPPT-Based Anomaly Detection and Accurate Fault Diagnosis Technique for PV Arrays," *Energy Conversion and Management*, vol. 234, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[3] Yasemin Onal, and Umit Cigdem Turhal, "Discriminative Common Vector in Sufficient Data Case: A Fault Detection and Classification Application On Photovoltaic Arrays," *Engineering Science and Technology, an International Journal*, vol. 24, no. 5, pp. 1168-1179, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[4] Wenchao Miao, K.H. Lam, and Philip W.T. Pong, "A String-Current Behavior and Current Sensing-Based Technique for Line–Line Fault Detection in Photovoltaic Systems," *IEEE Transactions on Magnetics*, vol. 57, no. 2, pp. 1-6, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[5] Heidi Kalliojärvi-Viljakainen, Kari Lappalainen, and Seppo Valkealahti, "A Novel Procedure for Identifying the Parameters of the Single-Diode Model and the Operating Conditions of a Photovoltaic Module from Measured Current–Voltage Curves," *Energy Reports*, vol. 8, pp. 4633-4640, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[6] Baojie Li et al., "Fault Diagnosis of Photovoltaic Panels using Full I–V Characteristics and Machine Learning Techniques," *Energy Conversion and Management*, vol. 248, pp. 1-21, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[7] Zhicong Chen et al., "Rapid and Accurate Modeling of PV Modules Based on Extreme Learning Machine and Large Datasets of I-V Curves," *Applied Energy*, vol. 292, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[8] Mohammed Wadi, "Fault Detection in Power Grids Based on Improved Supervised Machine Learning Binary Classification," *Journal of Electrical Engineering*, vol. 72, no. 5, pp. 315-322, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[9] Mahmoud Dhimish, and Zhicong Chen, "Novel Open-Circuit Photovoltaic Bypass Diode Fault Detection Algorithm," *IEEE Journal of Photovoltaics*, vol. 6, pp. 1819-1827, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[10] Hosna Momeni et al., "Fault Diagnosis in Photovoltaic Arrays Using GBSSL Method and Proposing a Fault Correction System," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 8, pp. 5300-5308, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[11] Manju Santhakumari, and Netramani Sagar, "A Review of the Environmental Factors Degrading the Performance of Silicon Wafer-Based Photovoltaic Modules: Failure Detection Methods and Essential Mitigation Techniques," *Renewable and Sustainable Energy Reviews*, vol. 110, pp. 83-100, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[12] V.S Bharath Kurukuru, and Ahteshamul Haque, *Photovoltaic Module Fault, Part 2: Detection with Quantitative-Model Approach*, Chapter 4, Fault Analysis and its Impact on Grid-connected Photovoltaic Systems Performance, pp. 111-148, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[13] Siva Rama Krishna Madeti, "A Monitoring System for Online Fault Detection in Multiple Photovoltaic Arrays," *Renewable Energy Focus*, vol. 41, pp. 160-178, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[14] Jingyue Wang et al., "Novel Application of Heterogeneous Ensemble Learning in Fault Diagnosis of Photovoltaic Modules," *2021 International Conference on Smart-Green Technology in Electrical and Information Systems*, Sanur, Bali, Indonesia, pp. 118-124, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[15] Imran Hussain et al., "Unified Fuzzy Logic Based Approach for Detection and Classification of PV Faults Using I-V Trend Line," *Energies*, vol. 15, no. 14, pp. 1-14, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[16] Zhicong Chen et al., "Deep Residual Network-based Fault Detection and Diagnosis of Photovoltaic Arrays Using Current-Voltage Curves and Ambient Conditions," *Energy Conversion and Management*, vol. 198, 2019. [CrossRef] [Google Scholar] [Publisher Link]