IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received February 4, 2020, accepted February 22, 2020, date of publication February 28, 2020, date of current version March 11, 2020.

*Digital Object Identifier 10.1109/ACCESS.2020.2977116*

# A Novel Convolutional Neural Network-Based Approach for Fault Classification in Photovoltaic Arrays

**FARKHANDA AZIZ[1], AZHAR UL HAQ[1], SHAHZOR AHMAD[1], YOUSEF MAHMOUD[2], (Senior Member, IEEE), MARIUM JALAL[3,4], AND USMAN ALI[1]**

[1]Department of Electrical Engineering, College of Electrical and Mechanical Engineering, National University of Sciences and Technology (NUST), Islamabad 44000, Pakistan
[2]Department of Electrical and Computer Engineering, Worcester Polytechnic Institute (WPI), Worcester, MA 01609, USA
[3]Department of Electronic Engineering, Fatima Jinnah Women University, Rawalpindi 46000, Pakistan
[4]Department of Electrical Engineering, Lahore College for Women University, Lahore 54000, Pakistan

Corresponding author: Azhar Ul Haq (azhar.ulhaq@ceme.nust.edu.pk)

**ABSTRACT** Fault diagnosis in photovoltaic (PV) arrays is essential in enhancing power output as well as the useful life span of a PV system. Severe faults such as Partial Shading (PS) and high impedance faults, low location mismatch, and the presence of Maximum Power Point Tracking (MPPT) make fault detection challenging in harsh environmental conditions. In this regard, there have been several attempts made by various researchers to identify PV array faults. However, most of the previous work has focused on fault detection and classification in only a few faulty scenarios. This paper presents a novel approach that utilizes deep two-dimensional (2-D) Convolutional Neural Networks (CNN) to extract features from 2-D scalograms generated from PV system data in order to effectively detect and classify PV system faults. An in-depth quantitative evaluation of the proposed approach is presented and compared with previous classification methods for PV array faults – both classical machine learning based and deep learning based. Unlike contemporary work, five different faulty cases (including faults in PS – on which no work has been done before in the machine learning domain) have been considered in our study, along with the incorporation of MPPT. We generate a consistent dataset over which to compare ours and previous approaches, to make for the first (to the best of our knowledge) comprehensive and meaningful comparative evaluation of fault diagnosis. It is observed that the proposed method involving fine-tuned pre-trained CNN outperforms existing techniques, achieving a high fault detection accuracy of 73.53%. Our study also highlights the importance of representative and discriminative features to classify faults (as opposed to the use of raw data), especially in the noisy scenario, where our method achieves the best performance of 70.45%. We believe that our work will serve to guide future research in PV system fault diagnosis.

**INDEX TERMS** Photovoltaic array, maximum power point tracking, fault classification, convolutional neural network, scalograms, transfer learning.

## I. INTRODUCTION

The photovoltaic (PV) industry has garnered prominence in recent years due to the economic and environmental benefits of freely available solar energy. From 2017 to 2022, the total installed PV capacity is expected to rise up to 438 GW [1]. Despite its free availability and other desirable characteristics, the PV industry is still facing challenges such as reliability, reduction in output power, initial cost, vulnerability to faults and its colossal dependency on environmental conditions [2]. As PV systems are exposed to harsh outdoor environment, they are susceptible to several faults and anomalies such as line-to-line (LL), line-to-ground (LG), open-circuit (OC), Hot spot (HS), environmental effects (uniform and non-uniform shading, humidity, snow and dust accumulation), wiring losses and malfunctioning of power conditioning units. These faults may reduce the energy conversion efficiency and lifetime of PV arrays and are reported to be the major reason behind their catastrophic failure [3], [4]. A survey study conducted in 2010 showed that such faults can

The associate editor coordinating the review of this manuscript and approving it for publication was Pierluigi Siano.

reduce the generated power of photovoltaic systems annually by about 18.9% [5].

To investigate and mitigate the aforementioned faults, article 690 in National Electric Code (NEC) recommends the use of Ground Fault Protection Device (GFPD), Over Current Protection Device (OCPD) and arc Fault Circuit Interrupter (AFCI) to detect LL, LG and arc faults respectively. However, the incompatibility of such protection devices is portrayed in comprehensive studies such as [6]–[9]. Specifically, the nonlinear characteristics of PV arrays, low irradiance, maximum power point tracker (MPPT), faults impedance, degradation and presence of blocking diodes are some factors that prevent protection devices to trip under certain conditions. Hence, sometime faults may go undetected for a long time, showing practical limitations of conventional protection schemes in PV arrays.

In addition to conventional protection schemes, a number of advanced fault diagnosis and classification methods have been proposed in existing literature to provide reliable protection. The effectiveness of advanced protection methods for PV array fault detection have been discussed in detail in [10]. The fault diagnosis and classification schemes proposed in [9], [11] and [12] reveal that fault detection and classification is challenging in PV systems due to weather conditions, high impedance faults and when MPPT is in operation. Differentiating faults from PS condition is also essential for an efficient performance of PV system. Moreover, classification methods for PV array faults have been demonstrated for only a few fault cases, as in [13] and [14]. Specifically, [13] have worked on only one case of fault i.e. Line-Line (LL) fault, while only two cases of PV array faults (LL fault and hotspot fault) were considered in [14], with both works not incorporating MPPT. Another study [16] also includes only two scenarios of PV array faults, i.e., LL fault and open circuit (OC) fault.

To address the aforementioned challenges and shortcomings of existing work in fault diagnosis, this paper presents a **novel approach** that utilizes deep two-dimensional (2-D) Convolutional Neural Network (CNN) to extract features from 2-D scalograms generated from PV system data, in order to effectively detect and classify PV system faults in severe conditions. The proposed approach is further divided into two configurations – one in which the last few layers of a pre-trained AlexNet CNN [15] are fine-tuned to yield a 6-way classifier, and another in which features are obtained from a certain layer (fc7) of a pre-trained AlexNet and then used in conjunction with classical classifiers. An in-depth quantitative evaluation of the proposed approach is presented and compared with previous feature extraction and classification methods for PV array faults – both classical machine learning (ML) based and deep learning (DL) based. Unlike contemporary work, five different faulty cases (i.e., partial shading (PS), line to line fault (LL), open-circuit fault (OC), high-impedance series / arc fault, and faults in PS – on which no work has been done before in the machine learning domain), as well as the no fault case, have been considered in our study, along with the incorporation of MPPT. We also employ a consistent dataset over which to compare ours and previous approaches, to make for the first (to the best of our knowledge) a comprehensive and meaningful comparative evaluation of fault diagnosis. It is observed that the proposed method involving fine-tuned pre-trained CNN outperforms other techniques, achieving high fault detection accuracies for both noiseless and noisy data. As experimentally validated, our study also highlights the importance of representative and discriminative features to classify faults (as opposed to the use of raw data), especially in the noisy scenario. At the same time, deep learning based automatic feature extraction is found to be superior as opposed to classical handcrafted feature extraction.

The remaining sections of the paper are organized as follows. Section II presents related work, while Section III provides the description of the proposed methodology. Section IV gives a comprehensive quantitative evaluation and analysis of the proposed and existing classification methods for PV array faults. Finally, Section V presents our conclusions and future directions.

## II. LITERATURE SURVEY

In recent years, a number of online fault diagnosis and classification approaches have been devised to counteract the aforementioned protection challenges. These detection and classification methods may be divided up into three categories namely; signal processing, performance comparison and machine learning techniques [16]. In signal processing based methods, faults are detected on the basis of reflected signals. Using the time specific characteristic of reflected signals, Time Domain Reflectometry (TDR) detects and locate faults. In TDR, an impedance variation in the reflected signal is examined by injecting an external voltage signal into system. An experimental analysis was perform on 1 MW PV plant using TDR to detect and locate faults in [17]. Another paper [18] explains spread spectrum time domain reflectometry for the identification of LG faults. Nevertheless, this method detects LG faults of any mismatch level but its accuracy is greatly affected by the distance between the device and fault location. Moreover, this method requires high speed sampling, external function generator and baseline for comparison.

The second domain makes fault detection by using performance comparison techniques. Jenitha and Immanuel Selvakumar [19] explained the fault detection method by calculating the power loss between the measured and predicted AC power production. In [2], a fault detection scheme was proposed by employing MPPT algorithm to estimate the expected maximum power, which is then compared with the meter-read power. In [20], an entropy-based fault detection scheme was proposed to detect LL, LG and OC faults under low irradiances and in the presence of blocking diodes. The proposed scheme could distinguish the faults from PS. String to string (SS), string to ground (SG) and OC faults were identified in [1]. Hariharan *et al.* have presented a fault

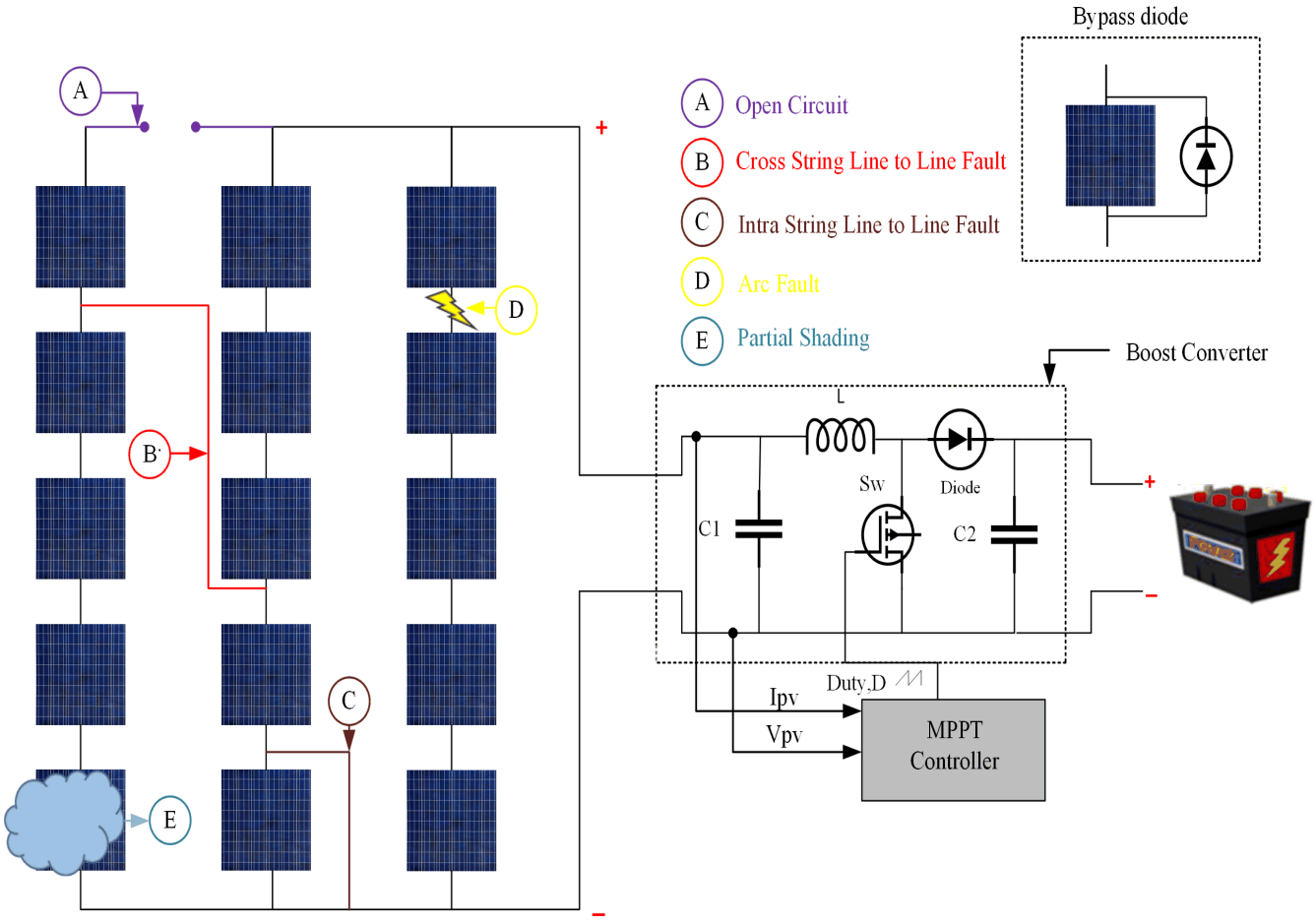


**FIGURE 1.** A typical configuration of PV system consisting of 5 × 3 PV array and a boost converter programmed with an MPPT algorithm to operate the PV module at their maximum power point (MPP).

diagnosis scheme with an ability to distinguish LL faults from normal conditions and impermanent faults such as PS in [21]. However, the proposed technique is not appropriate for large PV plants where the modules have different irradiance values. The accuracy of the aforementioned techniques depends on the quality of threshold limit. Even though the performance comparison techniques are comparatively simple to design but may not be effective due to following reasons stated in [9]: (1) They lack the ability to detect the type of faults, (2) dependence on MPPT operation, so MPPT failure may affect the diagnosis performance, (3) Model needs to be updated on regular basis, since PV parameters greatly depend upon seasonal conditions.

Machine learning techniques having an ability to classify faults according to their fault type have been propounded for fault detection in [13], [14] and [21]–[25]. Yi and Etemadi have proposed a support vector machine (SVM) based fault detection scheme for LL faults in [13]. The method suffers from the need for a large number of filters which makes the system expensive. Recently, deep learning based fault detection has been proposed in [14] for the identification of LL and HS faults. Although the proposed algorithm has achieved an impressive accuracy, MPPT was not deployed in the PV system as fault detection then becomes very challenging due to the presence of MPPT. Another research [22], utilizes Random Forest (RF) classifier for PV array faults classification, but the presented framework does not include high impedance LL faults. In [23], [24], a graph based semi supervised learning technique (GBSSL) was presented to detect and classify PV array faults. A probabilistic neural network-based fault detection scheme in [25] discusses the effects of LL and LG faults in the presence of blocking diodes by performing analysis of DC circuit of PV array. The proposed scheme, however, has not considered high impedance faults. Moreover, its feasibility in real PV arrays is not validated and it is complex in nature with high implementation cost. The existing literature shows that the detection of faults is difficult due to the presence of an MPPT, high fault impedance, low location mismatch, PS and due to degradation. It is very important to distinguish the faults from PS to avoid the false tripping of system.

### A. FAULTS IN PV ARRAYS

PV systems are frequently challenged by the occurrence of a number of electrical faults that may evolve due to several abnormalities in internal configuration [10]. Fig. 1 depicts

some of faults that have been analyzed in this study to perform the quantitative evaluation of classical ML and DL based fault detection schemes.

#### 1) LL FAULT

LL faults occur due to an unexpected short-circuiting between two points in PV array which are at different potential levels [26]. It may occur within same string or between two adjacent strings [10]. LL faults with low level of mismatch fault under low irradiances are difficult to detect because the fault current have low magnitude which remain undetected [13]. This mismatch level indicates the number of PV modules involved in the fault [26], [27]. The severity of LL fault usually depends on the impedance of fault path and level of mismatch. If the impedance is high and mismatch level is low, the fault current will be of small magnitude [13], [21]. Moreover, presence of MPPT optimizes the output power of PV array under various operating conditions which further makes LL fault detection challenging [8].

#### 2) OC FAULT

An OC fault in PV arrays is an accidental disconnection problem within a string or between two adjacent strings [28]. It may occur due to a number of reasons like breakage of the cable that connects two strings, any object falling on panels, loose connection between two points or an accidental disconnection at a current carrying conductor [25]. OC fault detection is studied in [20], [21] and [24].

#### 3) PARTIAL SHADING

Partial shading is a situation in which some modules are partially shaded while others are uniformly shaded. It is an impermanent condition which results in transitory reduction in output power [9]. Shading is categorized into two types [29]. One is static shading and the other is dynamic shading. Accumulation of dust, leaves and bird droppings on glass causes static shading, whereas dynamic shading occurs due to temporary shadow caused by nearby buildings or trees. In dynamic shading, shaded part changes with time so there is continuous variation in output power. Partial shading leads to the presence of multiple peaks in PV characteristics. Global maximum control helps to address this problem with the help of array reconfiguration method [30], [31]. In array reconfiguration, switches are involved to automatically disconnect the shaded module and reset the configuration in such a way that it can harvest maximum power from a PV array. The effects of PS on PV arrays are described in [21].

#### 4) ARC FAULT

In the ideal scenario, PV arrays have nearly zero impedance between module interconnections [32]. Arc faults occur due to loose connection at a conductor joint or cable insulation failure. There are two types of arc faults, namely series and parallel arc faults [33], [34]. Series arcs can often be observed within a module when there is a loose wired connection between modules or at the junction point. When two parallel conductors with different potential are placed close to each other parallel arcs are caused. Electrical simulations of series arc faults have been performed in [32].

### B. CONVOLUTIONAL NEURAL NETWORKS (CNNs)

A CNN is a special class of DL algorithms [35] commonly used for image recognition in computer vision [39]–[45]. A CNN architecture includes (besides various other parameters and special-purpose layers) three major layers, namely the convolutional layer, pooling layer and fully connected layer. The 2D convolution operation produces an optimal feature map of input data with the use of a number of learned filter kernels (as opposed to handcrafted or engineered filters typical in classical signal or image processing). The pooling layer serves as a down sampling layer in order to reduce the feature dimensions of data. Finally, fully connected layer is employed for the classification [36]. As the convolution and pooling operations are responsible for the feature extraction, it helps to learn patterns from the time-frequency representation and eliminates the requirement of hand-crafted expert features [37].

CNNs are learned using error-backpropagation, an optimization technique that allows a CNN to minimize classification error over the network parameters (i.e., filter kernels / weights). In other words, we are training the network to learn / mine discriminative features from the underlying data that would subsequently be conducive to classification of this data using a classifier like SVM or the CNN's softmax layer.

Training a CNN from scratch requires huge datasets to train and hence needs sufficiently large number of labeled data samples. Therefore, employing a pre-trained network has been recommended and successfully demonstrated by various studies [37], [38], [42]. In other words, a CNN network pre-trained on a large benchmark dataset that may be easily available (e.g., the ImageNet dataset for object recognition), can potentially be applied as an off-the-shelf feature extractor for completely different targeting a different domain (such as PV system faults, where a huge number of training samples may not be available). A number of well-known CNN architectures have been proposed in the computer vision community, such as the AlexNet [15], VGGNet [39], GoogLeNet [40], etc. and their pre-trained versions made publicly available.

#### 1) ALEXNET CNN

In this study, a pre-trained AlexNet CNN has been utilized as the feature extraction and classification algorithm. AlexNet is a CNN model presented by the seminal work of Alex Krizhevesky *et al.* in [15] that significantly outperformed the classical state-of-the art methods on the ImageNet Large Scale Visual Recognition Challenge ILSVRC), and went on to usher an era of deep learning trends in computer vision and related fields, supported in part by the widespread availability of massively parallel- processing hardware such as distributed clusters and GPUs. AlexNet was trained on 1.2 million high-resolution images depicting 1000 different

object categories. Fig. 2 depicts the AlexNet architecture along with fine-tuned AlexNet CNN. The AlexNet accepts an image of 227 × 227 as an input. Three operations namely, convolution, max pooling and Local Response Normalization (LRN) are performed in the first convolutional layer with the use of 96 different receptive filters with size of 11 × 11 which reduces image size to 55 × 55. In second layer, the image size reduces to 27 × 27 by applying 256 filters of 3 × 3 size when max pooling operation is applied. In third, fourth and fifth convolutional layers, ReLU is used and the image size is further reduced to 13 × 13. At the end, two fully connected layers having 4096 outputs with a dropout operation are utilized followed by a softmax layer [41]. ReLU non-linearity is applied after every convolutional and fully connected layer. ReLU is the most commonly used activation function in deep learning due to an ability to solve the nonlinear solving problems [42]. In [15], deep convolution neural networks with ReLUs trains much faster than the networks employed with standard *tanh* activation function. Furthermore, deep convolutional neural network with ReLU shows better performance with fast convergence [43]. Dropout is applied before the first and the second fully connected layer to reduce the risk of overfitting. For classification, softmax layer generates a distribution over 1000 class labels.

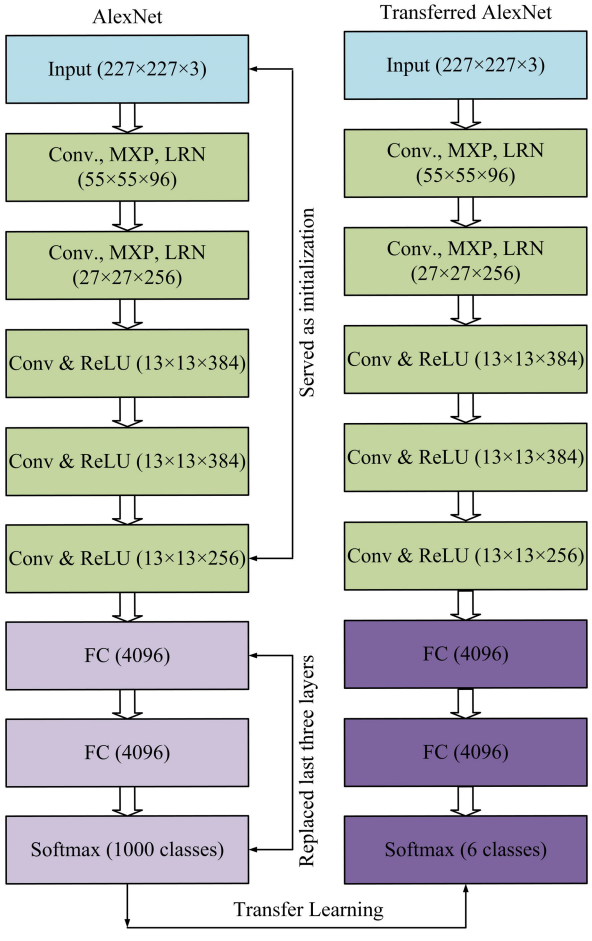


**FIGURE 2.** The AlexNet architecture and fine-tuned AlexNet CNN.

For classification, CNNs are mainly trained by employing three major methods, namely: training the CNN from scratch, using a pre-trained CNN as an off-the-shelf feature extractor followed by a separate classifier, and transfer learning, i.e., fine-tuning the pre-trained CNN model for the application under consideration [44] and [45]. Transfer learning is capable of solving the classification problems with small amount of data.

## III. PROPOSED METHODOLOGY

Fig. 3 represents the flow chart of the proposed technique (as well as other existing methods for fault diagnosis in PV arrays). First, collection of a dataset is performed, consisting of a total of 3456 samples with 576 samples per class, as discussed in Section III-B below. This dataset serves as a consistent test bench for the evaluation of ours as well existing approaches in our evaluation presented later in Section IV, to make for a meaningful comparison. We then process this 1D data using continuous wavelet transform (CWT) to generate 2D scalograms that efficiently capture the time-frequency characteristics of the PV system's 1D data (Section III-C). The resulting 2-D data is now in a form that can be readily used in conjunction with a pre-trained 2D convolutional neural network (CNN) for feature extraction (we used the AlexNet [15], but this proposed approach is independent of the exact CNN employed) and subsequent classification via classical techniques i.e., SVM and random forests (we used the AlexNet [15], but this proposed approach is independent of the exact CNN employed) and subsequent classification via classical techniques i.e., SVM and random forests. We also fine-tune the pre-trained AlexNet to yield a 6-way classifier that extracts features as well as performs classification in a joint framework. The description of each stage is discussed in detail in the following subsections.

### A. SIMULATED PV SYSTEM

Fig. 1 illustrates the PV System configuration considered in the proposed framework. It consists of 5×3 PV array, bypass diodes, a DC-DC boost converter with an MPPT programmed with perturb and observe (PO) algorithm. The studied PV array includes three parallel strings and each string is composed of five series modules. To build our PV array in Simulink, we use the widely referenced single diode model of [47], whose equivalent circuit (adopted from [46]) is shown in Fig. 4. The circuit consists of a current source $I_{ph}$ that represents the cell photocurrent, one Shockley diode represents PN junction of solar cell, $R_s$ and $R_{sh}$ represent the intrinsic series and shunt resistances of the cell respectively. $G$ is irradiance in W/m$^2$, $T$ is temperature in °C.

Our resulting Simulink model is presented in Fig. 5. It mathematically models the PV array in Fig. 1 using the terminal current and voltage relationship of the solar cell given by Eq. (1) in [3]. $K$ is Boltzmann constant ($1.3806503 \times 10^{-23} \times 10^{-23}$ J/K), $q$ is the electron charge ($1.60217646 \times 10^{-19} 17646 \times 10^{-19}$C) and $I_o$ is the saturation current. The

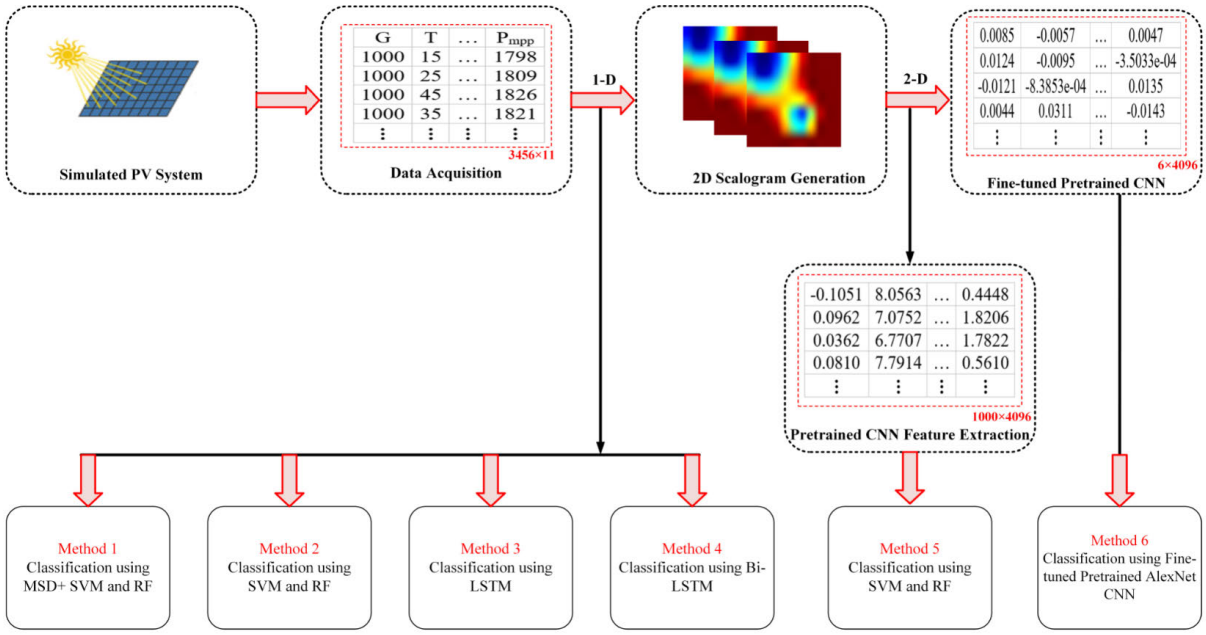


**FIGURE 3.** Flowchart of proposed PV array fault diagnosis method and existing methods.

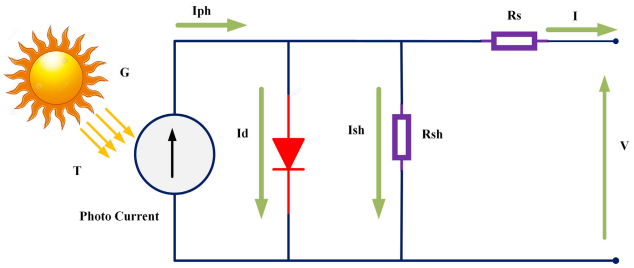


**FIGURE 4.** Single diode model (SDM) of PV cell [46].

**TABLE 1.** PV module specification at STC.

| Parameters | PV module specification |
|---|---|
| Maximum Power ($P_{max}$) | 49 Watt |
| Maximum Voltage ($V_{max}$) | 17 V |
| Maximum Current ($I_{max}$) | 2.88 A |
| Open circuit voltage ($V_{oc}$) | 21.8 V |
| Short circuit current ($I_{sc}$) | 3.11 A |

PV cell output current is $I$:

$$I = I_{ph} - I_o \left[ exp\left( \frac{q(V + R_s I)}{n \times KT} \right) - 1 \right] - \frac{V + R_s I}{R_{sh}} \quad (1)$$

Table 1 shows the standard test condition (STC) parameter specification of PV module. The irradiance and temperature fluctuations greatly affect the maximum power point (MPP) of PV array characteristic curves [48]. Figs. 6, 7 and 8 respectively describe the current-voltage (I-V) characteristic curves for the normal, LL fault and PS conditions of the PV array.

Fig. 6 depicts the output current of the PV array reduces as irradiance decreases and vice versa. It shows that PV array fault occurring at low irradiances will result in lower magnitude of fault current. Fig. 7 presents the LL fault at different fault impedances. We can observe that current remains unaffected as LL fault occurs with low mismatch level and at fault high impedance. From Fig. 8, we can examine that PV array under PS and faults in PS lead towards multiple peaks in characteristic curve.

### B. DATA ACQUISITION

To investigate the methodology proposed in Fig. 3, we have employed our Simulink model presented in Section III-A to collect a dataset under faulty and non-faulty operating conditions.

No fault, LL, OC, PS, fault in PS and series arc fault were the six cases considered for the evaluation of proposed method. Data attributes such as irradiance (G), temperature (T), short circuit current ($I_{sc}$), open circuit voltage ($V_{oc}$), photovoltaic current ($I_{pv}$), MPP current ($I_{mp}$), MPP voltage ($V_{mp}$) and MPP power ($P_{mp}$) were the eight values from the PV array output and three values from the boost converter such as maximum current, voltage and power were selected as data samples shown in Table 2. The reason for using them as data samples is that these values are directly or indirectly affected by the occurrence of faults in PV system. Total 3456 data samples and labels were collected. Each class (No Fault, LL, OC, PS, Fault in PS and arc fault) has 576 instances. For the simulation of arc fault, approached adopted in [32] is utilized in this study. Two faults namely, LL and OC have been generated for faults in PS case. These 3456 samples were collected by performing PV array

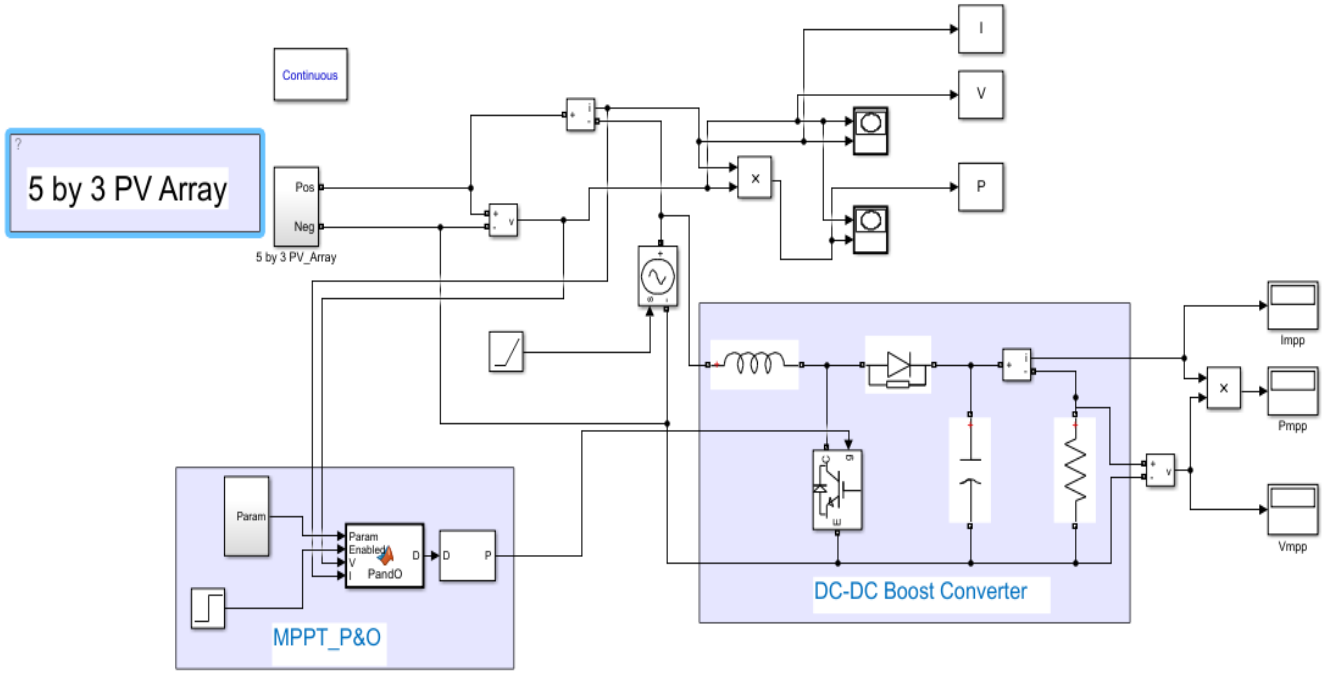


**FIGURE 5.** Simulink model of whole PV system.

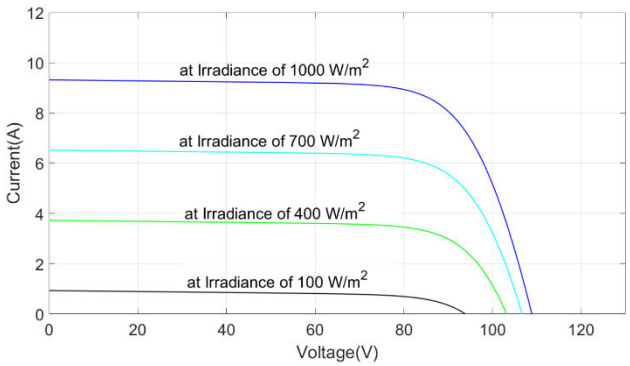


**FIGURE 6.** I-V Characteristics of PV array with ambient temperature of 25°C.

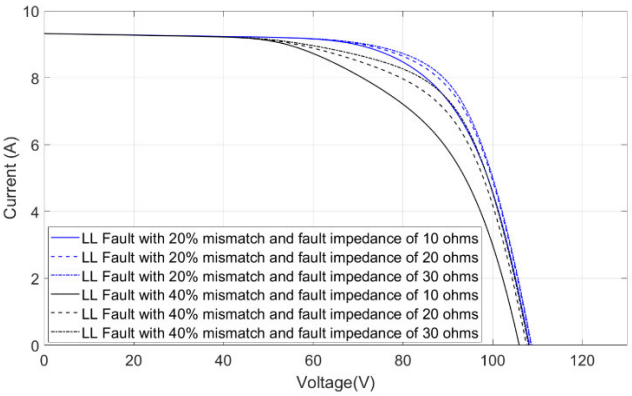


**FIGURE 7.** I-V Characteristics of PV array under LL fault.

simulation for the six cases with multiple combinations of following situations depicted in Table 3. Note that the high impedance parameter is employed only to generate instances

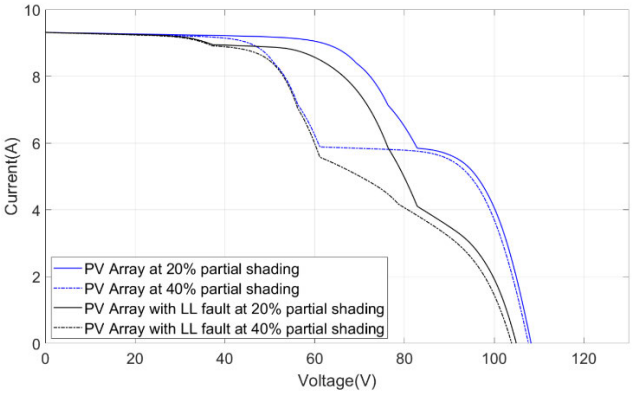


**FIGURE 8.** I-V Characteristic for PS and LL faults under PS.

**TABLE 2.** Data attributes selected for dataset collection.

| Sub-System | Data Attributes |
|---|---|
| PV Array | G, T, $I_{sc}$, $V_{oc}$, $I_{pv}$, $I_{mp}$, $V_{mp}$ and $P_{mp}$ |
| Boost Converter | $I_{max}$, $V_{max}$ and $P_{max}$ |

for the arc fault. The nature of collected data is 1-D. This 1-D data is passed to four existing methods of classification, some of which have been proposed previously for PV system fault diagnosis, while some we have adopted from the machine learning community.

The idea is to quantitatively evaluate all of these methods on a consistent dataset (as defined by Table 3) for a meaningful comparison. To the best of our knowledge, this kind of consistent comparative evaluation has not been done in past literature on fault diagnosis. Instead, authors have chosen

**TABLE 3. Combinations of parameter values adopted for dataset collection.**

| Parameters | Targeted values |
|---|---|
| Irradiance | 100 to 1200 W/m$^2$, at a step change of 25. |
| Temperature | 0 to 55 ℃, at a step change of 5. |
| Fault Impedance | 0 Ω to 30 Ω, at a step change of 10 |
| Percentage Mismatch | 20% to 80%, at a step change of 20 |
| Percentage of Partial Shading | 20% to 80%, at a step change of 20 |
| High Impedances values | 20, 40, 60, 80, 100 and 200 Ω |

to apply their techniques on their own datasets, collected with respect to different environmental conditions and system configuration (c.f. Table 3), differing array sizes, differing parameters, usually just two to three fault classes, and typically in the absence of MPPT. Thus, insofar as the previous literature on fault diagnosis is concerned, it is effectively meaningless to quote the performance presented in one work vs. that presented in some other paper. We briefly describe here the configurations employed for the various methods we have compared with our approach. In the first method, various features were extracted from Multiresolution Signal Decomposition (MSD) and classification is performed using SVM and RF, following [13].

Raw data without any feature extraction is directly passed to SVM and RF in the second method. In the third and fourth methods, this 1-D data is applied on DL algorithms LSTM (as proposed in [14]) and Bi-LSTM respectively for feature extraction and PV array fault classification.

In contrast, our approach first computes a 2D scalogram as discussed in the upcoming subsection, before it is used with pre-trained or fine-tuned CNNs (subsections III- D, E).

### C. 2-D SCALOGRAM GENERATION

CNNs work effectively by learning patterns from images. Nowadays, CNNs are widely used in biomedical domain by transforming signals from time domain (1-D) to frequency domain (2-D) using wavelet transform [36], [49]. As AlexNet CNN accepts only RGB images as input, we apply continuous wavelet transform (CWT) to data under analysis to generate scalograms (2-D image) presented in Fig. 3 and 9 respectively. The mathematical expression of CWT is given in (2) where $\psi(t)$ is a wavelet prototype provided in (3). $\psi(t)$ is shifted by $b$ and dilated by a factor $a$ before its product with $X(t)$ which is a time varying signal [50]. A Scalogram is a visual representation of signals based on time-frequency representation using wavelet transform (WT). Having created the scalogram plots for the data samples in Matlab, all scalogram images are cropped to a standard size of 227 × 227 to meet the requirement of input image layer of pre-trained AlexNet.

$$CWT\left\{x\left(t\right);a,b\right\} = \int x(t)\psi^{*}_{a,b}(t)dt \tag{2}$$

$$\psi_{a,b}\left(t\right) = \frac{1}{\sqrt{|a|}}\psi\left(\frac{t-b}{a}\right) \tag{3}$$

### D. PRE-TRAINED CNN FEATURE EXTRACTION

An important attribute of CNNs is that they automatically learn class-discriminative features from labeled training data, that go on to facilitate classification. As reviewed in Section II-B, re-training a CNN from scratch is not only computationally expensive, it also demands large amounts of data. Nevertheless, various works in the machine learning community have successfully demonstrated the generalization power of deep nets pre-trained on one, huge dataset performing well with regards to classification on other datasets, even from differing domains. Motivated by these observations, we apply (rather successfully as we shall see in Section IV) an AlexNet [15] pre-trained on a large-scale image classification dataset to our task of fault classification in PV arrays. Specifically, we use activations from the neurons on the second fully connected layer 'fc7' shown in Fig. 9. These extracted features from pre-trained AlexNet CNN were then passed to the fifth method of classification using SVM and RF depicted in Fig. 3 and 9, respectively.

### E. FINE-TUNED ALEXNET CNN

As a second configuration of our proposed method, we perform transfer learning by fine-tuning the last three layers (fully connected layer fc7, a softmax layer and a classification output layer) of a pre-trained AlexNet CNN (which is originally configured for 1000 classes) to finally yield a 6-dimensional output (which is the number of our classes) as depicted in Fig. 2.

The softmax layer calculates the probability distribution over each possible class and classifies the data according to most probable class. At the classification layer, the output with six classes was obtained as depicted in Fig. 9.

## IV. QUANTITATIVE EVALUATION

### A. METHODS COMPARED

As depicted in Fig. 3, we have considered six methods for PV array fault classification. We briefly describe them here. The **first two** are essentially our proposed techniques i.e., fine-tuned AlexNet (which jointly extracts features and performs classification), and pre-trained AlexNet (which is used to extract features that are subsequently classified using either SVMs or random forests, RF). As a **third** method, we employ SVMs and RF classification in conjunction with the raw 1D data as obtained from the PV system (i.e., the 11 dimensional parameter set described in Section III-B). The **fourth** method is the long-short term memory (LSTM) network, a deep learning paradigm originally proposed for fault diagnosis in [14]. Note however that we employ 11 features instead of the 3 originally used by this work, and that we apply it for the 6 fault cases in our standardized testbed (Section III-B) instead of the 2 faults considered in [14]. The **fifth** method is the Bi-directional LSTM (or Bi-LSTM), which trains the data both from front to back and as well as from back to

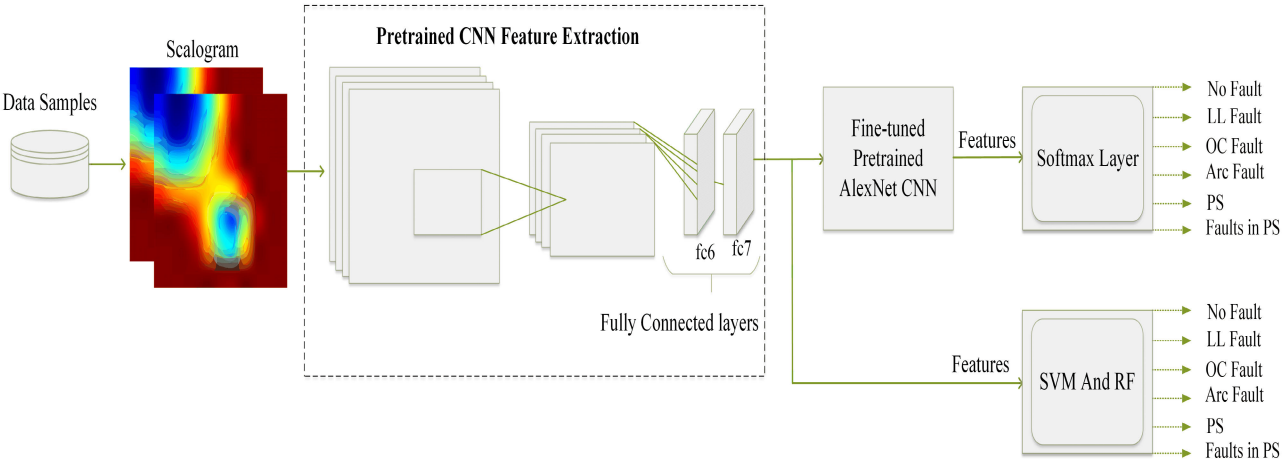


**FIGURE 9.** Framework of proposed 2-D CNN approach.

**TABLE 4.** Fault classification performance (averaged over four independent runs) of proposed vs. existing methods based on noiseless data (top two accuracies emboldened).

| Techniques | Average accuracy | STD |
|---|---|---|
| Fine-tuned AlexNet (proposed) | 73.53% | 1.03 |
| Pretrained AlexNet + SVM (proposed) | 70.09% | 0.844 |
| Pretrained AlexNet + RF | 64.31% | 1.127 |
| Raw + SVM | **82.22%** | 2.023 |
| Raw + RF | **78.47%** | 0.855 |
| LSTM **[14]** | 64.45% | 1.8 |
| BiLSTM | 62.75% | 1.23 |
| MSD + SVM **[13]** | 30.29% | 5.47 |
| MSD + RF | 49.23% | 6.406 |

front [56]. The **sixth** and last method is the Multiresolution Signal Decomposition (MSD) of [13], with classification performed by SVM. We also evaluate MSD with RF classification. Again, we have used our standardized testbed of 11 parameters and 6 fault classes, unlike the original work in [13]. The results for these methods and their various possible configuration are summarized in Tables 4 and 5 for noiseless and noisy data, respectively.

### B. COMPUTER SYSTEM CONFIGURATION

The work is carried out in Matlab 2018a, running on an Intel Core i7-7500U processor at 2.70GHz with 8GB RAM. The model of proposed framework is implemented in Matlab using the Deep Learning Toolbox and the Signal Processing Toolbox. The employed Simulink model of the PV system has already been presented in Section III-A (Fig.5).

**TABLE 5.** Fault classification performance (averaged over four independent runs) of proposed vs. existing methods based on noisy data (top two accuracies emboldened).

| Techniques | Average accuracy | STD |
|---|---|---|
| Fine-tuned AlexNet **(proposed)** | **70.45%** | 1.75 |
| Pretrained AlexNet + SVM **(proposed)** | **69.39%** | 0.635 |
| Pretrained AlexNet + RF | 63.68% | 1.96 |
| Raw + SVM | 67.86% | 3.058 |
| Raw + RF | 68.09% | 1.456 |
| LSTM **[14]** | 63.45% | 4.42 |
| BiLSTM | 61.00% | 2.54 |
| MSD + SVM **[13]** | 20.37% | 4.4 |
| MSD + RF | 22.69% | 4.869 |

### C. NOISY DATA

The evaluation is performed on both noiseless and noisy data. For noisy data, white noise is added to data samples in order to reflect real world values of electrical sensors and measuring devices. Electrical sensors have thermal noise which is equivalent to white noise. Studies conducted in [53] and [54] used white noise in their framework. White noise is a random signal having equal intensity over the full spectrum. The approach used for noise generation in this study is expressed in (4).

$$Noisy\ signal = Noiselesss\ signal + wgn(m, n, p) \quad (4)$$

where $wgn$ () is a predefined Matlab function for white noise which creates $m \times n$ matrix of white Gaussian noise, and $p$ is the power of signal in decibels.

### D. EVALUATION METRICS

The performance is quantified using accuracy (also called sensitivity or recall), averaged over four independent runs

(during which the dataset is randomly split into 70-30 to give training and testing samples). The standard deviation over these runs is also reported.

While accuracy is the typically employed evaluation metric in classification literature (as well as in previous works on fault diagnosis such as [13], [14]), we also briefly describe here additional metrics that may be obtained from the presented confusion matrices to judge a method's performance. Denote the number of true positives, false positives, true negatives and false negatives as TP, FP, TN and FN, respectively. We then define **accuracy** (also called **sensitivity** or **recall**, i.e. true positive rate) as TP / (TP+FN), i.e., the proportion of relevant examples that are correctly classified as relevant. Next, **precision** is defined as TP / (TP+FP), i.e., the proportion of classified examples that are actually relevant. Finally, **specificity** (also called true negative rate) is defined as TN / (TN+FP), i.e., the proportion of irrelevant examples that are correctly classified as irrelevant.

### E. CHOICE OF CNN IN PROPOSED APPROACH

We have performed an experiment to compare the performance in terms of accuracy and execution time of two other CNN architectures besides AlexNet [15] for our approach. Table 6 depicts the results showing that AlexNet not only has roughly half the execution time, but provides comparable, if not better, accuracy vs. GoogLeNet [40] and ResNet [52]. Similar observations are seen in other studies (e.g., [51]). Therefore, we choose to use AlexNet for all subsequent experiments involving our proposed approach. We believe the reason may have to do with the fact that AlexNet is a considerably simpler (shallower) architecture compared with the other two, more sophisticated, architectures and, since our PV system fault data is not as feature-rich and complex as natural images, AlexNet performs quite satisfactorily here.

**TABLE 6.** Fault classification accuracies of pretrained models of CNNs.

| CNN pretrained models | Fault classification accuracy | Execution Time (train + test) in minutes |
|---|---|---|
| AlexNet | 74.6% | 208.06 |
| GoogLeNet | 73.5% | 359.37 |
| ResNet | 72.4% | 369.76 |

### F. TRAINING AND HYPER-PARAMETERS

Performance of DL algorithms can be improved by tuning the hyperparameters. In this study, several combinations of hyperparameter values were tried and the best combination was adopted, as described in the following. The minimum batch size was set to 5 to boost the training accuracy for both the algorithms. The batch size refers to the number of training samples used for one iteration. Stochastic gradient descent with momentum set to 0.9 was used for fine-tuning our 2D CNN, while adaptive moment estimation (Adam) optimization algorithm was used to learn the LSTM. The learning rate chosen for both the algorithms was 0.0001. Fine-tuned CNN was trained with 70% randomly selected data with 100 epochs (an epoch is one complete round of training over the entire dataset), and maximum iterations of 48300 per epoch. Accuracy and loss were the two parameters utilized for performance evaluation. Training progress plotted in Fig. 10 shows the accuracy and losses over time based on noiseless data. Initially training accuracy is observed to be low due to small batch size, and it gradually increases and reaches around 98.67%. By inspection, it can be observed that loss slowly decreases and converges to around 0.092.

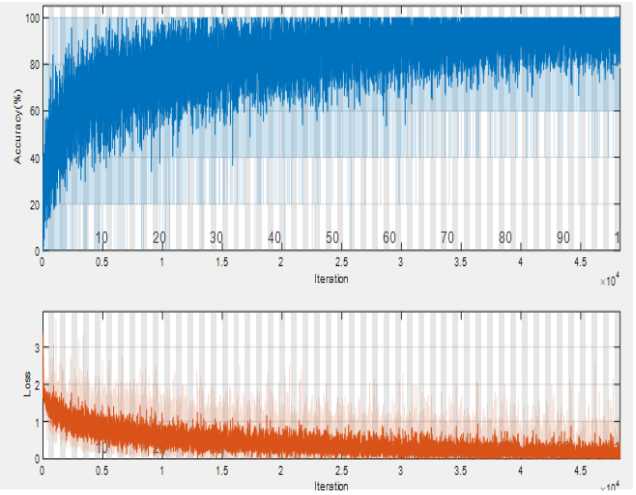


**FIGURE 10.** Convergence curves of training process.

### G. TESTING AND EVALUATION

Testing was performed on the 30% held-out data and best average accuracy and loss (over four random runs), averaged over the six classes (No fault, LL, OC, PS, fault in PS and arc fault) are observed to be 74.6% and 25.4% respectively. The confusion matrix is presented in Fig. 11, which portrays the per-class accuracy (sensitivity) along the bottom-most row, and the per-class precision along the right-most column. The bottom right cell indicates the overall accuracy, averaged over all the 6 classes. The worst accuracy obtained from AlexNet CNN is 72.1% with a loss of 27.9%, as depicted in the confusion matrix given in Fig. 12. In similar vein, the performance of LSTM [14] and BiLSTM are captured in the confusion matrix presented in Fig. 13 and 14. Accuracies of the various methods (averaged over the four independent runs as well as over all six classes) are presented in Table 4. Interestingly, the best performance is achieved when raw data (i.e., our 11 set of parameters) are directly used with either SVM (85.23%) or RF (79.55%) classifier. Our proposed methods closely follow behind in performance. But their real strength (and that of the other DL technique i.e. LSTM) will become

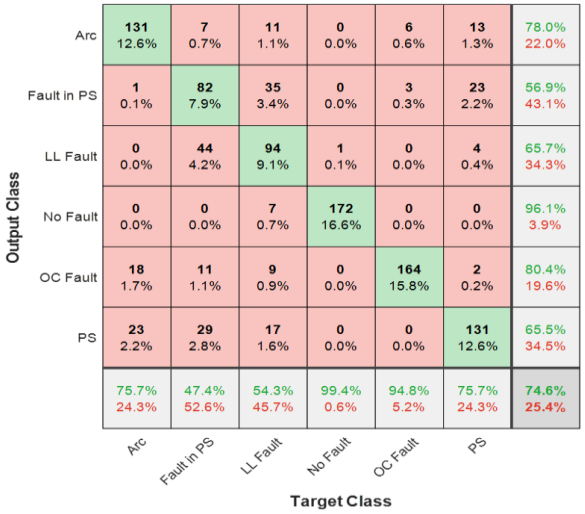


**FIGURE 11.** Best case of fault classification results for fine-tuned pre-trained AlexNet CNN based on noiseless data.

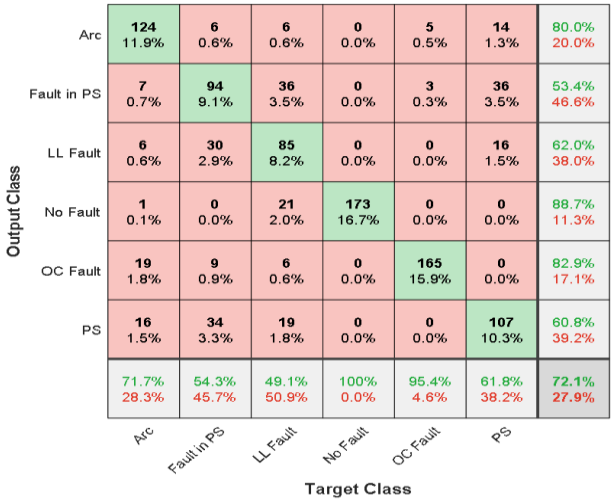


**FIGURE 12.** Worst case of fault classification results for fine-tuned pre-trained AlexNet CNN based on noiseless data.

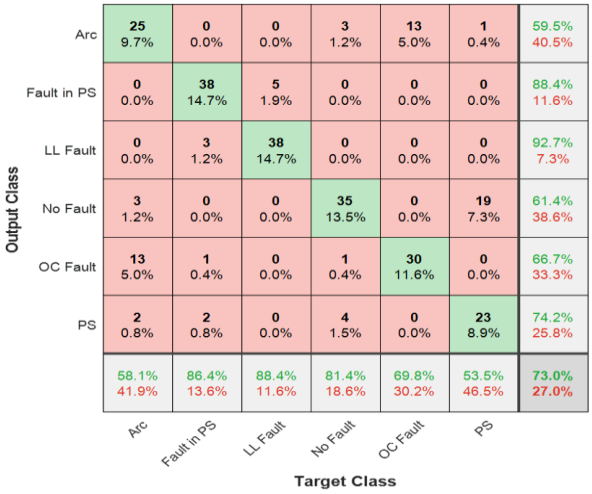


**FIGURE 13.** Fault classification results for LSTM based on noiseless data.

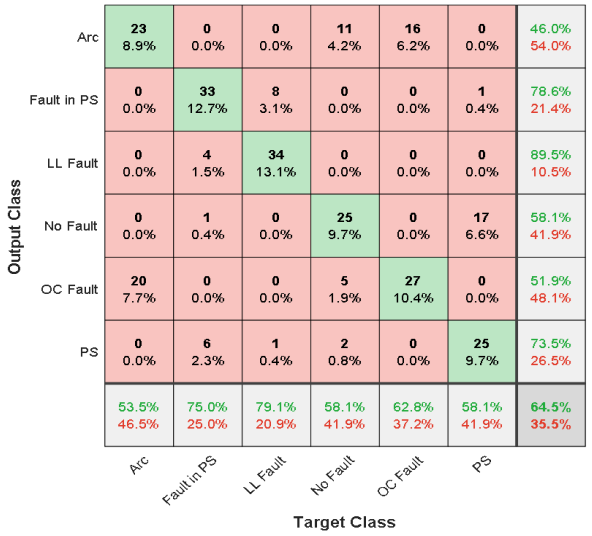


**FIGURE 14.** Fault classification results for Bi-LSTM based on noiseless data.

evident shortly in Section IV-J, where noisy data is used to more accurately simulate sensor data).

To compute accuracy (sensitivity), precision or specificity from a confusion matrix, we must first compute the four parameters TP, FP, TN and FN from it with respect to a certain class. For instance, consider the matrix presented in Fig. 11, and assume the analysis is for arc fault. **TP** (number of arc fault instances correctly classified as arc fault) is found to be 131. **FP** (number of irrelevant instances mistakenly classified as arc fault) is found by summing up the numbers in the first row (except in the first column), i.e., 37. **TN** (number of irrelevant instances correctly classified as irrelevant, i.e., not belonging to arc fault) is found as total number of irrelevant test instances minus FP, i.e., (5 × 173 - 37), where 5 is the number of irrelevant classes and 173 is the number of test instances for each of those classes. This comes out to be 828. **FN** (number of arc fault instances mistakenly classified as some other fault) is found by summing up the numbers in the first column (except in the first row), i.e., 42. Note that TP + FN is the total number of relevant instances (i.e., the number of test instances of class arc fault), and is a constant, i.e., 173.

Based on the above discussion, **accuracy** or **sensitivity** for class arc fault is TP / (TP+FN) = 131 / (131+42) = 0.76 (bottom row), **precision** is TP / (TP+FP) = 131 / (131+37) = 0.78 (rightmost column), while **specificity** is TN / (TN+FP) = 828 / (828 + 37) = 0.96. Note the values for accuracy presented in the Tables 4 and 5 are based on averaging the results over all the 6 classes.

### *H. INCREASING FAULT CLASSES*

Performance of the proposed method (fine-tuned AlexNet) as the number of classes is increased can be analyzed in Fig. 15. This shows that faults with low magnitude of fault current

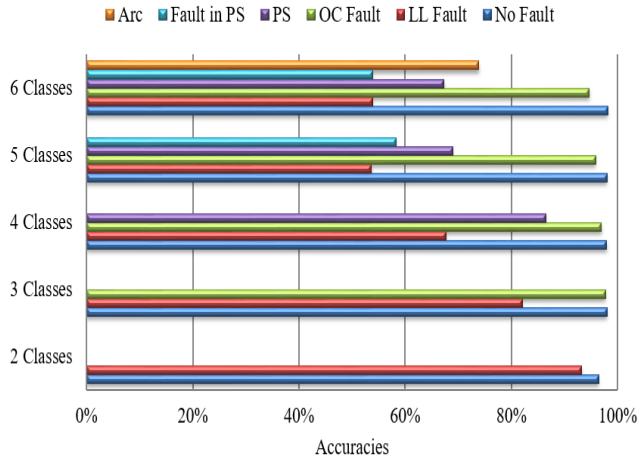


**FIGURE 15.** Performance as number of classes is increased (fine-tuned pre-trained CNN).

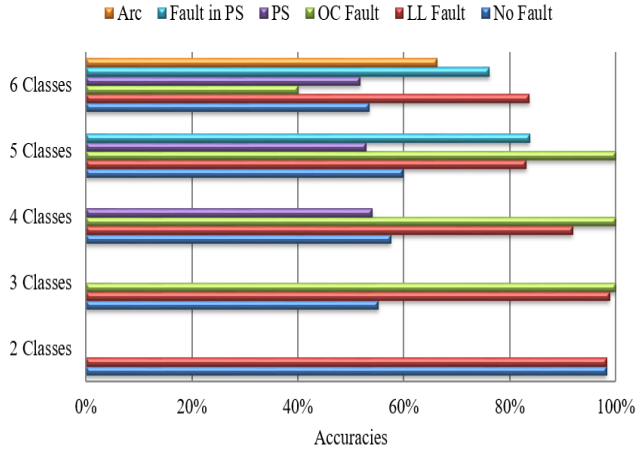


**FIGURE 16.** Performance as number of classes is increased (LSTM).

like LL faults are difficult to distinguish from other faults as the number of classes increase and for the severe conditions depicted in section IV-B (Table 3). The accuracy of LL fault falls to a considerable extent when PS is added as a fourth class. It may be noted that with two-class classification, individual accuracies for the two classes are more than 90%, but performance reduces as number of classes increases. Fig. 16 and 17 respectively give the analysis for LSTM [14] and Pretrained AlexNet + SVM.

### I. INCREASING TRAINING DATA SIZE

The performance of proposed algorithm has also been examined by fixing test data to 300 samples and increasing training samples to 286 samples. The results are depicted in Fig. 18, illustrating that as the number of training samples increases, accuracies are improved. Similar analysis is given in Fig. 19 for LSTM [14].

### J. NOISY DATA

The performance of proposed method (fine-tuned AlexNet) with noisy data with a power level of -3.01db is summarized in the confusion matrix (best performing average over the

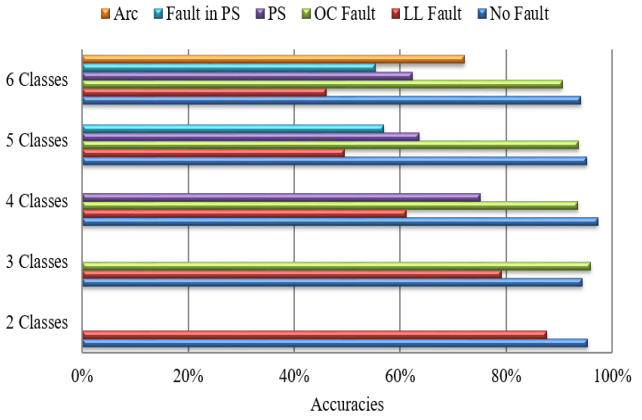


**FIGURE 17.** Performance as number of classes is increased (pretrained AlexNet + SVM).

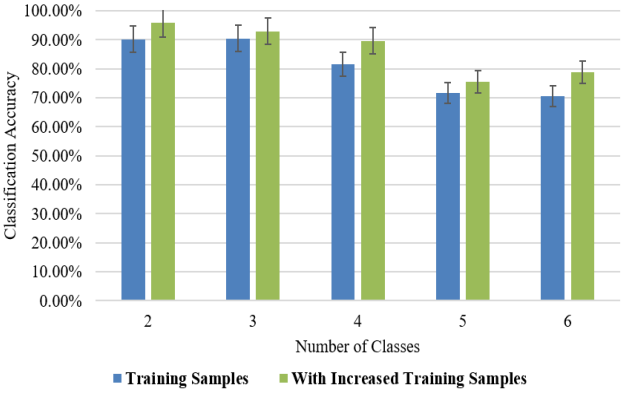


**FIGURE 18.** Evaluation of fault classification results for fine-tuned CNN based on increasing training samples.

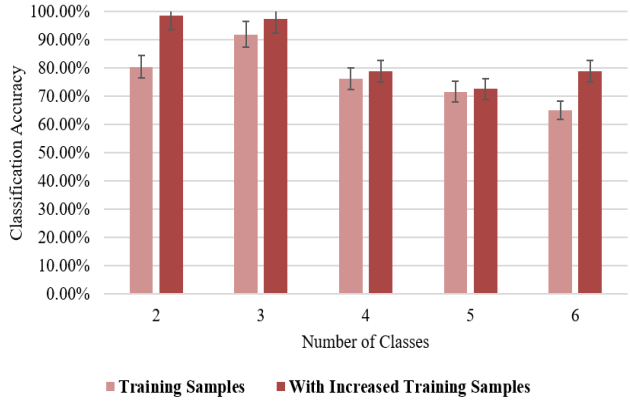


**FIGURE 19.** Evaluation of fault classification results for LSTM based on increasing training samples.

four independent runs) presented in Fig. 20. Performance for different noise levels is depicted in Fig. 21 for deep learning methods i.e., pre-trained AlexNet (proposed) vs. LSTM [14] and Bi-LSTM. Similarly, Fig. 22 gives the performance for classical machine learning methods. We observe that the proposed technique demonstrates resilience to noise at all

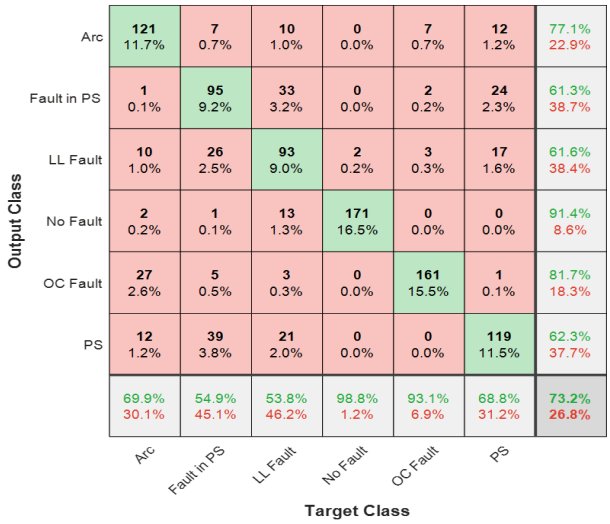


**FIGURE 20.** Fault classification results for fine-tuned pre-trained AlexNet CNN based on noisy data.

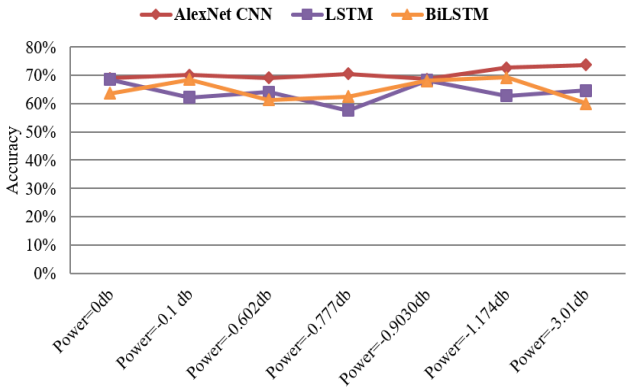


**FIGURE 21.** Fault classification results for deep learning algorithms based on noisy data.

power levels tested, and maintains a higher accuracy than the other two compared methods.

The average accuracies (over all classes and all runs) for the various methods are given in Table 5. We note that, in the presence of noise, the two configurations of our proposed approach yield the best fault classification performance of 70.45% and 69.39% among all tested methods. Another interesting conclusion is that for deep learning techniques (proposed, LSTM, Bi-LSTM), performance drops in the presence of noise is not as pronounced compared to approaches which either use raw data in conjunction with SVM / RF classifier, or handcrafted feature extraction i.e., MSD [13]. In other words, data-driven "learning" of discriminative features is indispensable for robust classification in the face of noise.

### *K. DISCUSSION*

In Table 3, we have utilized 200Ω high impedance value for the arc fault. Since such high impedance values would indeed

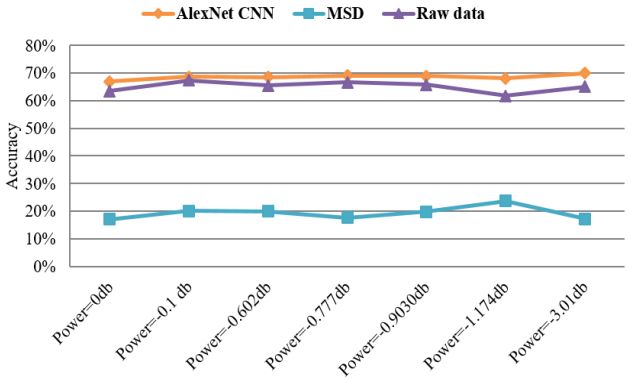


**FIGURE 22.** Fault classification results for classical machine learning algorithms based on noisy data.

not arise in practice. To ascertain the effect on the performance due to such samples, we first identify the samples of the arc fault class involving the 200Ω value. These are found to be 96 of a total of 576 arc fault class samples (note that training is undertaken on a randomly selected subset of 70% of the 576 samples, so it is difficult to exactly ascertain how many of these 96 samples were used for training and how many for testing for the originally reported results in Table 4). We re-run the experiment and find that while the original accuracy for arc fault class is 75.7% (c.f Fig. 11, bottom, left-most cell), the accuracy when removing the aforementioned 96 samples is interestingly higher, i.e., 83.3%. Nevertheless, the effect on the overall performance (averaged over all the six classes) is minimal, i.e., 75.4% as opposed to the originally reported 74.6%.

With regards to exactly what features are extracted by a CNN, there does exist some literature in the computer vision community. One such work is by Girshick *et al.* [57]. These works attempt to visualize the various kernels learned at different layers of a deep convolutional neural network.

Corroborating the findings in the studies of biological networks in the neuroscience community, these works suggest that the earlier layers of the network typically learn low-level image features such as oriented edges and bars, whereas subsequent layers learn to capture higher-level information such as shapes, and even higher layers learn to discern contextual constrains and semantic meaning from an image.

However, as far as our work is concerned, we note that the aforementioned 1D data consisting of only 11 features lacks the high dimensionality, as well as the semantic meaning and structure inherent in natural images. Therefore, visualization of the weights and kernels at various layers of a deep CNN would not be intuitive for human interpretation. Nevertheless, we have performed some analysis wherein we have computed the average of the 576 samples per class and visualized the 2D scalogram generated for that class. Fig. 23 reproduced below depicts these "average scalograms" for our various classes. One may observe the subtle differences between the images, and it is this typical difference which is probably learned by the 2D CNN in order to distinguish between the classes.

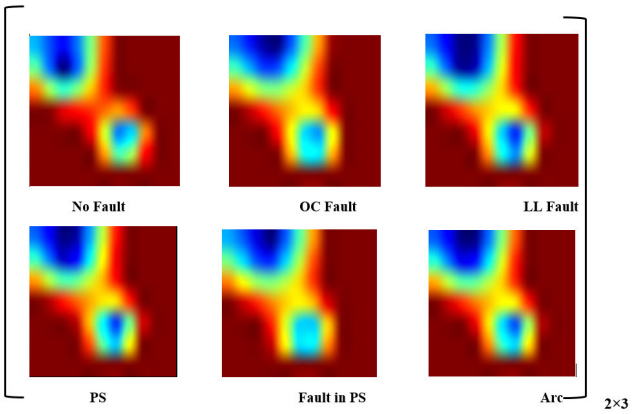


**FIGURE 23.** Visualization of "average scalograms" for our six classes.

To understand the limitation of our method in a scenario where MPPT may fail (and assuming the system can automatically receive a notification about its failure in a real setting), we make use of only 8 features (i.e., remove the three values obtained from the boost converter). Thus, training and testing system is performed on this feature-reduced dataset of 576 samples per class, each having only 8 instead of 11 features. The test results are presented in the Table 7.

**TABLE 7.** Fault classification accuracies with MPPT failure.

| Techniques | Fault classification accuracy (without MPPT) | Fault classification accuracy ( with MPPT) |
|---|---|---|
| Fine-tuned AlexNet | 45.35% | 73.53% |
| Pretrained AlexNet + SVM | 45.47% | 70.09% |

We observe that the performance of our approaches drops significantly due to the failure of MPPT. This is because 3 out of 11 data values are altogether lost, and now the method must work with a limited data per sample. The reason can be attributed to the requirement of 2D CNNs to be fed high-dimensional, structured data upon which they can mine discriminative and representative patterns and class typicalities. The deeper the network, the higher the dimensionality it subsequently expects to give good performance. We note however that a PV system is by nature a simple power generation system which typically may be monitored only based on its electrical output and meteorological data, as done in this paper (no further parameters are available due to the absence of turbines or generators or other moving mechanical parts). Hence, while our work has certainly shed light on the favorable potential of 2D CNNs for fault diagnosis, we believe that similar potential in 1D CNNs can be tapped to possibly yield better performance with limited monitored parameters. In similar vein, a custom 2D CNN design, with fewer parameters could also be investigated, instead of adopting large networks from the computer vision community, as done in this pioneering work. Furthermore, utilizing larger data sets for training CNNs from scratch could also improve the performance.

## V. CONCLUSION AND FUTURE WORK

In this paper, two convolutional neural network based configurations have been proposed for PV array fault classification. One in which the last few layers of a pre-trained AlexNet are fine-tuned to yield 6-way output, and another in which the features are obtained from fc7 layer of a pre-trained AlexNet and then used in conjunction with classical ML methods for classification. The proposed method measures the basic characteristics of a PV array to develop the fault classification algorithm without using any high end equipment (Thermal imagers). Moreover, a quantitative evaluation of PV array faults based on feature extraction and classification methods for six cases (No fault, LL fault, OC fault, PS, fault in PS and arc fault) under severe conditions (in which faults have remained undetected in past work) has been performed. Following **contributions** in this research have been made: **firstly**, a novel approach to PV array fault classification using 2D scalogram generation based on PV system data, followed by a 2D CNN, giving a high fault detection accuracy of 73.53%. **Secondly**, our study also highlights the importance of representative and discriminative features to classify faults (as opposed to the use of raw data), especially in the noisy scenario, where our method achieves the best performance of 70.45%. **Thirdly**, unlike contemporary work, five different faulty cases (along with the no fault case) have been considered in our study, along with a consistent dataset over which to compare ours and previous approaches, to make for the first (to the best of our knowledge) comprehensive and meaningful comparative evaluation of fault diagnosis. Further work is certainly needed in this seminal direction to systematically ascertain the strengths and weaknesses of deep learning paradigms such as CNNs and LSTMs in the context of fault diagnosis for efficiency enhancement of PV systems. We also intend to perform lab experiments on physical PV arrays to explore the techniques presented in this paper.

## REFERENCES

[1] A. Khoshnami and I. Sadeghkhani, "Fault detection for PV systems using Teager–Kaiser energy operator," *Electron. Lett.*, vol. 54, no. 23, pp. 1342–1344, Nov. 2018.

[2] C.-L. Kuo, J.-L. Chen, S.-J. Chen, C.-C. Kao, H.-T. Yau, and C.-H. Lin, "Photovoltaic energy conversion system fault detection using fractional-order color relation classifier in microdistribution systems," *IEEE Trans. Smart Grid*, vol. 8, no. 3, pp. 1163–1172, May 2017.

[3] Z. Chen, L. Wu, S. Cheng, P. Lin, Y. Wu, and W. Lin, "Intelligent fault diagnosis of photovoltaic arrays based on optimized kernel extreme learning machine and I-V characteristics," *Appl. Energy*, vol. 204, pp. 912–931, Oct. 2017.

[4] M. K. Alam, F. Khan, J. Johnson, and J. Flicker, "A comprehensive review of catastrophic faults in PV arrays: Types, detection, and mitigation techniques," *IEEE J. Photovolt.*, vol. 5, no. 3, pp. 982–997, May 2015.

[5] S. K. Firth, K. J. Lomas, and S. J. Rees, "A simple model of PV system performance and its use in fault detection," *Sol. Energy*, vol. 84, no. 4, pp. 624–635, Apr. 2010.

[6] D. S. Pillai and N. Rajasekar, "An MPPt-based sensorless line–line and line–ground fault detection technique for PV systems," *IEEE Trans. Power Electron.*, vol. 34, no. 9, pp. 8646–8659, Sep. 2019.

[7] D. S. Pillai and N. Rajasekar, "A comprehensive review on protection challenges and fault diagnosis in PV systems," *Renew. Sustain. Energy Rev.*, vol. 91, pp. 18–40, Aug. 2018.

[8] Y. Zhao, J.-F. De Palma, J. Mosesian, R. Lyons, and B. Lehman, "Line–line fault analysis and protection challenges in solar photovoltaic arrays," *IEEE Trans. Ind. Electron.*, vol. 60, no. 9, pp. 3784–3795, May 2013.

[9] B. P. Kumar, G. S. Ilango, M. J. B. Reddy, and N. Chilakapati, "Online fault detection and diagnosis in photovoltaic systems using wavelet packets," *IEEE J. Photovolt.*, vol. 8, no. 1, pp. 257–265, Jan. 2018.

[10] D. S. Pillai, F. Blaabjerg, and N. Rajasekar, "A comparative evaluation of advanced fault detection approaches for PV systems," *IEEE J. Photovolt.*, vol. 9, no. 2, pp. 513–527, Mar. 2019.

[11] Z. Yi and A. H. Etemadi, "Fault detection for photovoltaic systems based on multi-resolution signal decomposition and fuzzy inference systems," *IEEE Trans. Smart Grid*, vol. 8, no. 3, pp. 1274–1283, May 2017.

[12] Y. Hu, W. Cao, J. Wu, B. Ji, and D. Holliday, "Thermography-based virtual MPPT scheme for improving PV energy efficiency under partial shading conditions," *IEEE Trans. Power Electron.*, vol. 29, no. 11, pp. 5667–5672, Nov. 2014.

[13] Z. Yi and A. H. Etemadi, "Line-to-Line fault detection for photovoltaic arrays based on multiresolution signal decomposition and two-stage support vector machine," *IEEE Trans. Ind. Electron.*, vol. 64, no. 11, pp. 8546–8556, Nov. 2017.

[14] A. Y. Appiah, X. Zhang, B. B. K. Ayawli, and F. Kyeremeh, "Long short-term memory networks based automatic feature extraction for photovoltaic array fault diagnosis," *IEEE Access*, vol. 7, pp. 30089–30101, 2019.

[15] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," *Commun. ACM*, vol. 60, no. 6, pp. 84–90, May 2017.

[16] Y. Zhao, *Fault Detection, Classification and Protection in Solar Photovoltaic Arrays*. Boston, MA, USA: Northeastern University, 2015.

[17] L. Schirone, F. P. Califano, U. Moschella, and U. Rocca, "Fault finding in a 1 MW photovoltaic plant by reflectometry," in *Proc. IEEE Photovolt. Spec. Conf.*, vol. 1, 1994, pp. 846–849.

[18] S. Roy, M. K. Alam, F. Khan, J. Johnson, and J. Flicker, "An irradiance-independent, robust ground-fault detection scheme for PV arrays based on spread spectrum time-domain reflectometry (SSTDR)," *IEEE Trans. Power Electron.*, vol. 33, no. 8, pp. 7046–7057, Aug. 2018.

[19] P. Jenitha and A. Immanuel Selvakumar, "Fault detection in PV systems," *Appl. Sol. Energy*, vol. 53, no. 3, pp. 229–237, Oct. 2017.

[20] A. Khoshnami and I. Sadeghkhani, "Sample entropy-based fault detection for photovoltaic arrays," *IET Renew. Power Gener.*, vol. 12, no. 16, pp. 1966–1976, Dec. 2018.

[21] R. Hariharan, M. Chakkarapani, G. Saravana Ilango, and C. Nagamani, "A method to detect photovoltaic array faults and partial shading in PV systems," *IEEE J. Photovolt.*, vol. 6, no. 5, pp. 1278–1285, Sep. 2016.

[22] Z. Chen, F. Han, L. Wu, J. Yu, S. Cheng, P. Lin, and H. Chen, "Random forest based intelligent fault diagnosis for PV arrays using array voltage and string currents," *Energy Convers. Manage.*, vol. 178, pp. 250–264, Dec. 2018.

[23] H. Momeni, N. Sadoogi, M. Farrokhifar, and H. Farhadi Gharibeh, "Fault diagnosis in photovoltaic arrays using GBSSL method and proposing a fault correction system," *IEEE Trans Ind. Informat.*, to be published.

[24] Y. Zhao, R. Ball, J. Mosesian, J.-F. de Palma, and B. Lehman, "Graph-based semi-supervised learning for fault detection and classification in solar photovoltaic arrays," *IEEE Trans. Power Electron.*, vol. 30, no. 5, pp. 2848–2858, May 2015.

[25] M. N. Akram and S. Lotfifard, "Modeling and health monitoring of DC side of photovoltaic array," *IEEE Trans. Sustain. Energy*, vol. 6, no. 4, pp. 1245–1253, Oct. 2015.

[26] L. Chen, S. Li, and X. Wang, "Quickest fault detection in photovoltaic systems," *IEEE Trans. Smart Grid*, vol. 9, no. 3, pp. 1835–1847, Aug. 2018.

[27] K. AbdulMawjood, S. S. Refaat, and W. G. Morsi, "Detection and prediction of faults in photovoltaic arrays: A review," in *Proc. IEEE 12th Int. Conf. Compat., Power Electron. Power Eng. (CPE-POWERENG )*, Apr. 2018, pp. 1–8.

[28] R. Kase and S. Nishikawa, "Fault detection of bypass circuit of PV module—Detection technology of open circuit fault location," in *Proc. 19th Int. Conf. Electr. Mach. Syst. (ICEMS)*, Nov. 2016, pp. 1–4.

[29] P. Bulanyi and R. Zhang, "Shading analysis & improvement for distributed residential grid-connected photovoltaics systems," in *Proc. 52nd Annu. Conf. Aust. Sol. Counc.*, vol. 2, May 2014, pp. 1–12.

[30] H. Patel and V. Agarwal, "Maximum power point tracking scheme for PV systems operating under partially shaded conditions," *IEEE Trans. Ind. Electron.*, vol. 55, no. 4, pp. 1689–1698, Apr. 2008.

[31] M. Boztepe, F. Guinjoan, G. Velasco-Quesada, S. Silvestre, A. Chouder, and E. Karatepe, "Global MPPT scheme for photovoltaic string inverters based on restricted voltage window search algorithm," *IEEE Trans. Ind. Electron.*, vol. 61, no. 7, pp. 3302–3312, Jul. 2014.

[32] J. Flicker and J. Johnson, "Electrical simulations of series and parallel PV arc-faults," in *Proc. IEEE 39th Photovoltaic Spec. Conf. (PVSC)*, Jun. 2013, pp. 3165–3172.

[33] P. C. Hatton, M. Bathaniah, Z. Wang, and R. S. Balog, "Arc generator for photovoltaic arc fault detector testing," in *Proc. IEEE 43rd Photovoltaic Spec. Conf. (PVSC)*, Jun. 2016, pp. 1–6.

[34] Q. Xiong, S. Ji, L. Zhu, L. Zhong, and Y. Liu, "A novel DC arc fault detection method based on electromagnetic radiation signal," *IEEE Trans. Plasma Sci.*, vol. 45, no. 3, pp. 472–478, Mar. 2017.

[35] Q. V. Le, "A tutorial on deep learning Part 2: Autoencoders, convolutional neural networks and recurrent neural networks," *Tutorial*, vol. 1, pp. 1–20, Oct. 2015.

[36] D. Gurve and S. Krishnan, "Deep learning of EEG time–frequency representations for identifying eye states," *Adv. Data Sci. Adapt. Anal.*, vol. 10, no. 02, Aug. 2018, Art. no. 1840006.

[37] M. Papadomanolaki, M. Vakalopoulou, S. Zagoruyko, and K. Karantzalos, "Benchmarking deep learning frameworks for the classification of very high resolution satellite multispectral data," *ISPRS Ann. Photogram., Remote Sens. Spatial Inf. Sci.*, vols. 7, pp. 83–88, Jun. 2016.

[38] S. Lu, Z. Lu, and Y.-D. Zhang, "Pathological brain detection based on AlexNet and transfer learning," *J. Comput. Sci.*, vol. 30, pp. 41–47, Jan. 2019.

[39] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," 2014, *arXiv:1409.1556*. [Online]. Available: https://arxiv.org/abs/1409.1556

[40] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2015, pp. 1–9.

[41] Z. Alom, T. M. Taha, C. Yakopcic, S. Westberg, P. Sidike, M. S. Nasrin, B. C. Van Esesn, A. A. S. Awwal, and V. K. Asari, "The history began from AlexNet: A comprehensive survey on deep learning approaches," 2018, *arXiv:1803.01164*. [Online]. Available: https://arxiv.org/abs/1803.01164

[42] M. J. Brown, L. A. Hutchinson, M. J. Rainbow, K. J. Deluzio, and A. R. De Asha, "A comparison of self-selected walking speeds and walking speed variability when data are collected during repeated discrete trials and during continuous walking," *J. Appl. Biomech.*, vol. 33, no. 5, pp. 384–387, Oct. 2017.

[43] Q. Li, W. Cai, X. Wang, Y. Zhou, D. D. Feng, and M. Chen, "Medical image classification with convolutional neural network," in *Proc. 13th Int. Conf. Control Autom. Robot. Vis. (ICARCV)*, Dec. 2014, pp. 844–848.

[44] X. Han, Y. Zhong, L. Cao, and L. Zhang, "Pre-trained AlexNet architecture with pyramid pooling and supervision for high spatial resolution remote sensing image scene classification," *Remote Sens.*, vol. 9, no. 8, p. 848, Aug. 2017.

[45] H.-C. Shin, "Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning," *IEEE Trans. Med. Imag.*, vol. 35, no. 5, pp. 1285–1298, May 2016.

[46] F. Harrou, B. Taghezouit, and Y. Sun, "Improved $k$NN-based monitoring schemes for detecting faults in PV systems," *IEEE J. Photovolt.*, vol. 9, no. 3, pp. 811–821, May 2019.

[47] H. G. Abugoukh, "Mathematical modeling of photovoltaic module with simulink," *Lighting Eng. Power Eng.*, vol. 1, no. 54, pp. 47–54, 2019.

[48] A. Ahmed, L. Ran, S. Moon, and J.-H. Park, "A fast PV power tracking control algorithm with reduced power mode," *IEEE Trans. Energy Convers.*, vol. 28, no. 3, pp. 565–575, Sep. 2013.

[49] S. Guo, Q. Lv, B. Liu, Y. Lin, and R. Li, "Deep convolutional neural networks for electrocardiogram classification," *Lect. Notes Electr. Eng.*, vol. 536, no. 6, pp. 57–69, 2019.

[50] R. Rodriguez, S. Shrestha, R. Pascone, K. Lynch, E. Voudouri, G. Livanos, M. Zervakis, A. Deshpande, C. Narayan, Y. Na, and G. C. Giakos, "Low-order statistical analysis of 1-D diffuse reflectance signals from cancer cells using 2-D scalogram images," in *Proc. IEEE Int. Conf. Imag. Syst. Techn. (IST)*, Oct. 2016, pp. 171–176.

[51] R. K. Remya and M. Wilscy, "Pretrained convolutional neural networks as feature extractor for image splicing detection," in *Proc. Int. Conf. Circuits Syst. Digit. Enterprise Technol. (ICCSDET)*, Dec. 2018, pp. 1–5.
[52] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 770–778.
[53] J. Lai, T. Perazzo, Z. Shi, and A. Majumdar, "Optimization and performance of high-resolution micro-optomechanical thermal sensors," *Sens. Actuators A, Phys.*, vol. 58, no. 2, pp. 113–119, Feb. 1997.
[54] A. Kawamura, T. Haneyoshi, and R. G. Hoft, "Deadbeat controlled PWM inverter with parameter estimation using only voltage sensor," *IEEE Trans. Power Electron.*, vol. 3, no. 2, pp. 118–125, Apr. 1988.
[55] Y. Shima, "Image augmentation for object image classification based on combination of pre-trained CNN and SVM," *J. Phys., Conf. Ser.*, vol. 1004, Apr. 2018, Art. no. 012001.
[56] T. Le, M. T. Vo, B. Vo, E. Hwang, S. Rho, and S. W. Baik, "Improving electric energy consumption prediction using CNN and Bi-LSTM," *Appl. Sci.*, vol. 9, no. 20, p. 4237, 2019.
[57] R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2014, pp. 580–587.

**FARKHANDA AZIZ** received the B.E. degree in electronics engineering from the Mehran University of Engineering and Technology (MUET), Jamshoro, Pakistan. She is currently pursuing the M.S. degree in electrical engineering with specialization in control and power from the Electrical Engineering Department, College of Electrical and Mechanical Engineering, National University of Sciences and Technology, Pakistan. Her research interests include photovoltaic modeling, power electronics, and machine learning.

**AZHAR UL HAQ** received the Ph.D. degree from the University of L'Aquila, Italy, and University of Waterloo, Waterloo, ON, Canada, in 2016. He is currently working as an Assistant Professor with NUST, Islamabad, Pakistan. He is author of various journal articles and conference papers. His research interest includes power systems and operation. He was a recipient of several research and travel grants from HEC, Pakistan.

**SHAHZOR AHMAD** received the B.E. degree in electronics engineering from the National University of Sciences and Technology (NUST), Pakistan, in 2008, the M.Sc. degree in signal processing from Nanyang Technological University (NTU), Singapore, in 2011, and the Ph.D. degree from the National University of Singapore (NUS), in 2016, focusing on computer vision. He was a Teaching Assistant with NUS, from 2014 to 2016. He is currently an Assistant Professor with the Department of Electrical Engineering, NUST. His current research interests include applying computer vision and machine learning to engineering problems in intelligent transportation systems, autonomous vehicles, robotics, and power systems.

**YOUSEF MAHMOUD** (Senior Member, IEEE) received the B.Sc. degree in electrical power engineering from Albalqa'a Applied University, Amman, Jordan, in 2009, the M.Sc. degree in electrical power engineering from the Masdar Institute of Science and Technology, Abu Dhabi, United Arab Emirates, in 2012, and the Ph.D. degree in electrical and computer engineering from the University of Waterloo, Waterloo, ON, Canada, in 2016. He is currently an Assistant Professor with the Worcester Polytechnic Institute, Worcester, MA, USA. His research interests mainly include control and operation of grid-connected photovoltaic power systems and their power electronic converters, including PV modeling, maximum power point tracking, and partial shading impact.

**MARIUM JALAL** received the bachelor's and the master's degrees in electrical and telecommunications engineering from Pakistan, in 2006 and 2009, respectively, and the Ph.D. degree in electrical and information engineering from the University of L'Aquila, L'Aquila, Italy, in 2015. Her research interest includes cooperative wireless communications. She was awarded the Gold Medal for the Best Master's thesis.

**USMAN ALI** received the B.E., M.S., and Ph.D. degrees in electrical engineering from the College of Electrical and Mechanical Engineering (E&ME), National University of Sciences and Technology (NUST), Pakistan, in 2005, 2007, and 2012, respectively. He is currently an Assistant Professor with the Department of Electrical Engineering, College of E&ME, NUST. His research focuses on design and development of embedded systems (SoC) for computer vision, communication, and biomedical algorithms. The research area also includes renewable and sustainable energy systems.