Fault Detection in Photovoltaic Arrays via Sparse Representation Classifier

Heybet Kiliç*, Behnam Khaki†, Bilal Gumuş‡, Musa Yilmaz§ and Peter Palensky*
* TU Delft, † University of California Los Angeles, ‡ Dicle University, § Batman University,

***Abstract*—In recent years, there has been an increasing interest in the integration of photovoltaic (PV) systems in the power grids. Although PV systems provide the grid with clean and renewable energy, their unsafe and inefficient operation can affect the grid reliability. Early stage fault detection plays a crucial role in reducing the operation and maintenance costs and provides a long lifespan for PV arrays. PV Fault detection, however, is challenging especially when DC short circuit occurs under the low irradiance conditions while the arrays are equipped with an active maximum power point tracking (MPPT) mechanism. In this case, the efficiency and power output of a PV array decrease significantly under hard-to-detect faults such as active MPPT and low irradiance. If the hard-to-detect faults are not detected effectively, they will lead to PV array damage and potential fire hazard. To address this issue, in this paper we propose a new sparse representation classifier (SRC) based on feature extraction to effectively detect DC short circuit faults of PV array. The proposed compressive sensing method requires only the output current and voltage of PV arrays. To verify the effectiveness of the proposed SRC fault detection method, we use numerical simulation and compare its performance with the artificial neural network (ANN) based fault detection.**

***Index Terms*—Compressive sensing, Photovoltaic array fault detection, sparse representation.**

I. Introduction

In recent years, the electrical power generation form of renewable energy resources has become popular due to the significant increase in the energy demand and environmental issue relating to the fossil fuel energy. Among others, solar energy is reliable, broadly available, sustainable, and inexhaustible [1]. Several countries, such as USA, Germany, China, Japan, Italy and India have made tremendous effort and investment to meet their energy demand by increasing the penetration of renewable energy resources in the electrical power grids. As a case in point, India has an ambitious plan to increase its PV installation capacity to 34GW by 2022 [2]. By increasing the integration of PV systems in power grids, however, several technical issues emerge which should be effectively addressed. One of those issues is the electrical fault in PV arrays, especially DC side line-to-line (L-L) and line-to-ground (L-G) faults. Efficient, reliable, and safe operation of electrical systems with PV arrays requires the short circuit faults to be detected by a proper protection mechanism [3]. Undetected faults can cause PV array failure, fire hazard risk, as well as unexpected loss of energy economic benefits. The DC-side faults are hardly detectable under two conditions: (i) a low current fault occurs between points that have close electric potential, i.e. low mismatch faults or under low irradiance condition faults in the morning, evening, and cloudy days; and (ii) the PV array system is equipped with the maximum power point tracking (MPPT) algorithm [4].

In the literature, various methods have been proposed for PV fault detection. A Fractional-order colour relation classifier method is proposed in [5] to monitor mismatch, L-G, open-circuit, and bridge faults in a PV array using output power degradation. A practical fault detection is proposed in [6] which compares the predicted and measured AC power of the PV system. In [7], the authors utilize offline descriptive inferential statistical procedure and online inferential algorithm to monitor and supervise the PV system operation. The authors in [8] report a sensor suited approach comparing the measured and estimated voltage and current based on iMPPT algorithm presented in [9]. To detect the efficiency loss in the PV arrays, a multisensory approach proposed in [10] which is equipped with the inertial, temperature, irradiance, current, and voltage sensors. Despite the effectiveness and efficiency of the fault detection methods which are based on the comparison of measured and estimated parameters, they suffer from the following shortcomings: (1) they are costly due to the need for considerable number of sensors; (2) power losses due to the fault are quite similar under small mismatch and low irradiance conditions, thus it cannot work properly under those conditions; (3) they are not practical as their low fault detection accuracy results in a difference between the simulated and actual cases; and (4) due to non-linear operation of PV array, their performance depends on the PV array performance which is influenced under different environmental conditions.

In addition to the methods discussed above, which compare the estimated with measured parameters to detect the faults, there are several methods based solely on the measured values. A two-sectional fault detection method based on optimized voltage sensor locations is introduced in [11]. Although the method can locate the faulty string in small-scale arrays, it needs a large number of costly sensors for the large-scale arrays, so it is not economical for practical implementation. The authors in [12] propose a method to detected fault based on power loss, ratio of irradiance and DC power to detect partial shading and mismatch. The method, however, shows low accuracy due to the variation of irradiance over large-scale PV array. A quick fault detection scheme using multiple meters is proposed in [13] to measure the different outputs under the sequential change detection framework. However, its performance is degraded under irradiance fluctuations. To detect short and open circuit faults in a PV string, a probabilistic neural network method is presented in [14]. This method is valid only for high current faults, it suffers from

long testing time for each sample, and it is impractical for large-scale big data cases

According to the literature review, there are a number of researches on the DC-side short circuit faults such as L-L with low-mismatch percentage and L-G with high mismatch percentage under low irradiance and high-fault impedance. In addition, the impact of MPPT algorithm has not been taken into account. Although the algorithms in [15], [16] attempt to evaluate the impact of small mismatch, high-impedance, and low irradiance, their accuracy is not satisfactory. Moreover, they require large datasets of irradiance and temperature to train the algorithm. Studies in [17] and [18] demonstrate a time domain reflectometry approach to detect the impedance change by injecting a signal and observing its response. Previous studies such as [8], [10], and [19] focus on the detection of short circuit faults, but they have a low accuracy for low mismatch and high impedance under low irradiance condition.

Fuzzy Logic, Support Vector Machine (SVM), and Artificial Neural Network (ANN) are also deployed to detect faults in PV arrays. An unsupervised one class SVM is proposed for detecting anomalies in photovoltaic systems. Reference [20] presents a wavelet and ANN based L-L fault detection for ungrounded PV systems. However, the performance of these methods significantly depends on the choice of feature and tuning option, and they have high computational cost. In [21], a sample entropy method is used to detect L-L, L-G and short circuits (S-C) faults, but it requires a long computational time and is impractical for real-time fault detection.

To address the shortcomings in the literature, this study presents a new methodology, called spare representation (SR), to detect short-circuits faults in a PV array based on feature extraction and classification. The aim of the proposed method is to detect the faults under low-mismatch, low-irradiance and high-impedance conditions. The SR method consists of a dictionary of feature vectors obtained from measurements and an SR vector which includes non-zero (for a specific element) and negligible or zero elements for other vectors. To classify the inputs, the proposed method used a SR vector which consists of non-zero elements. The main contributions of this paper are:

- Although there are many methods in the field of fault-diagnosis such as multilayer perception (MLP) network,Probabilistic neural network (PNN),Decision tree (DT)Support vector machine (SVM). They suffer slow learning speed, limited to applications with relatively small data sets, large network structures. However our proposed SRC does not suffer parameter tuning due to iteratively training pro-poses, Hence it provides high learning speed.
- The computational complexity and storage space are crucial in many applications, especially in the case of dealing with a large dataset and requiring a fast response. The lower computational cost provides a quick response for a test sample. The sparse properties of our proposed method ensure low computational cost and low sample storage.
- The proposed method creates an overcomplete dictionary consisting of the training samples themselves. Due to including a small part of the overall training dataset, this representation proved an automatically discriminated nature. Therefore, our method is effective under hard-to-detect fault conditions such as partial shading, low-mismatch, high fault impedance and low irradiance with blocking diode and active MPPT cases.

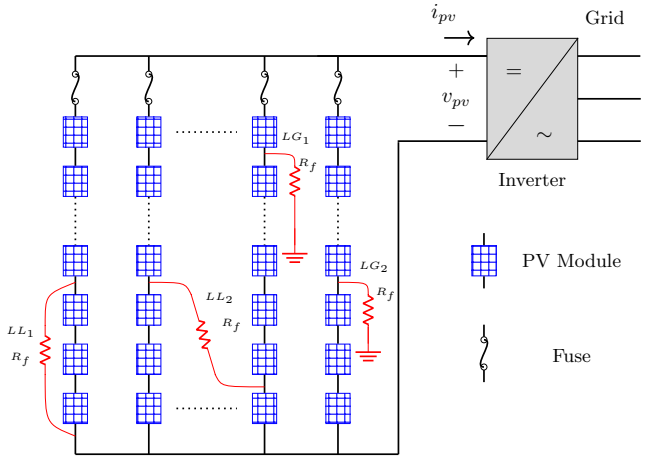

Fig. 1. A PV array configuration with potential short-circuit faults.

The remaining part of the paper proceeds as follows: in Section II, the PV array configurations, DC-side short-circuit faults, and their challenges are presented; in Section III, SRC method for PV fault detection and norm minimization method for feature extraction are discussed; in section IV, the simulation results for PV array fault detection using SR are given and followed by section V for discussion on the findings; and finally the paper is concluded in Section VI.

II. PV System Model and Fault Detection

A. PV Array Configuration

A generic PV systems structure, which consists of PV modules, DC-DC converters with MPPT algorithm, Pulse-Width-Modulation (PWM) inverters, electrical connections and wiring, and ground-protection devices, is shown in Fig. 1. Multiple PV modules are connected in series to construct a PV string. Then multiple PV string are connected in parallel to form a PV array. To maximize the output power of PV array which is influenced by the temperature and the irradiance, the PV array should be equipped with MPPT providing maximum PV array output power under different weather conditions.

Grounding of PV system in practice depends on the size of the solar plant, operating voltage, geographic location, local standards, and the type of installation (i.e. building mounted, ground-mount, roof-top, etc.). Generally, in the U.S. system grounding is defined as an electrical connection between ground and current carrying conductors (CCCs) through a ground-fault detection and interruption (GFDI) fuse. On the other hand, different grounding systems, such as DC-insulation resistance (Riso) measurement and the residual current monitoring devices (RCDs), are used outside the U.S, and there is no connection between ground and CCC [3].

B. PV Array Faults and Detection Challenges

The common faults in PV systems are short and open circuit faults as well as panel mismatch and module failures [3]. In this study, the L-L and L-G faults which are shown by LL and LG in Fig. 1 are studied. Those faults can happen spontaneously due to a connection between one PV and the ground (L-G faults) or two PV modules (L-L faults).

The short-circuit between CCCs is due to the DC junction box corrosion, animal chewing, mechanical damages, and water ingress [4]. The occurrence of short-circuit faults is not predictable, therefore, in order to prevent fire hazard and power loss, a protection scheme should be designed for early-stage fault detection. According to the U.S. National Electrical Code (NEC), the non-current-carrying (NCC) metal parts (e.g. chassis, mounting racks, panel frames, etc.) must be connected to the equipment grounding conductor (EGC) to prevent electric shocks. The connection between the NCC metals and EGC results in an L-G fault in PV arrays [22].

Short-circuit faults can be described as the percentage of mismatch which shows the number of PV panels influenced by the fault directly. The PV array in Fig. 1 consists of 10 panels linked in series and 10 strings linked in parallel. For instance, LL_1 and LL_2 represent 30% mismatch, while LG_1 and LG_2 represent 90% and 30% L-L faults, respectively. In general, L-G faults with high mismatch percentage and L-L faults with low mismatch percentage are considered as the hard-to-detect faults since they have a smaller current than the other faults. When a short-circuit fault occurs, the voltage of faulty string suddenly decreases and receives a back-feeding current from healthy strings. Thus, an over-current protection device (OCPD) or ground fault protection devices (GFPD) blow to prevent the back-feeding current. According to the UL standard $2579-7$ and U.S. NEC requirements, the threshold value for a fuse should be greater than $2.1I_{SC}$, where I_{SC} is the PV string's short-circuit current under standard testing condition (STC, irradiance = 1000 [W/m^2], temperature = 25 [^{o}C] [4], [22]. Under lower irradiance (less than 500 [W/m^2], the induced current is smaller, and therefore the back-feeding current induced by faults may not be large enough to blow the fuse, so it leads to undetectable faults. This case also causes losses, damage to PV panels, and fire hazard.

MPPT algorithm causes challenges for fault detection in PV array. Under low-irradiance condition, MPPT may hide the DC-side faults. If a fault occurs, the normal operating point of the PV array shifts to a different curve which has a lower open-circuit voltage. In Fig. 2, it is shown that the MPPT algorithm provides a new MPP to optimize the system output under the fault which diminishes the PV array's operating voltage and the string' back-feeding current of the string. In Fig. 3, the impact of MPPT algorithm on the back-feeding current of a PV string is demonstrated. In the case of 60% mismatch and without MPPT, the fuse blows to prevent the PV system from a back-feeding current. Also, in the case of 20%, 40%, and 60% mismatch with MPPT, the back-feeding current of string is less than $2.1I_{SC}$ which causes an inadequate fuse melting time.

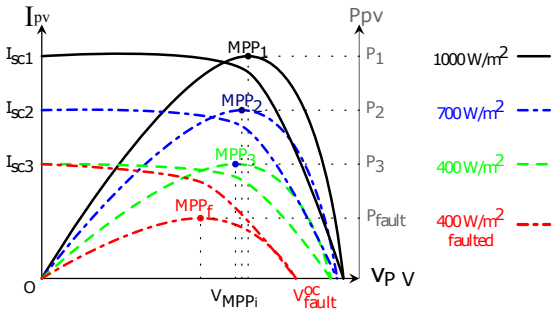

Fig. 2. Normalized I-V and P-V curves of a PV array for several irradiance levels.

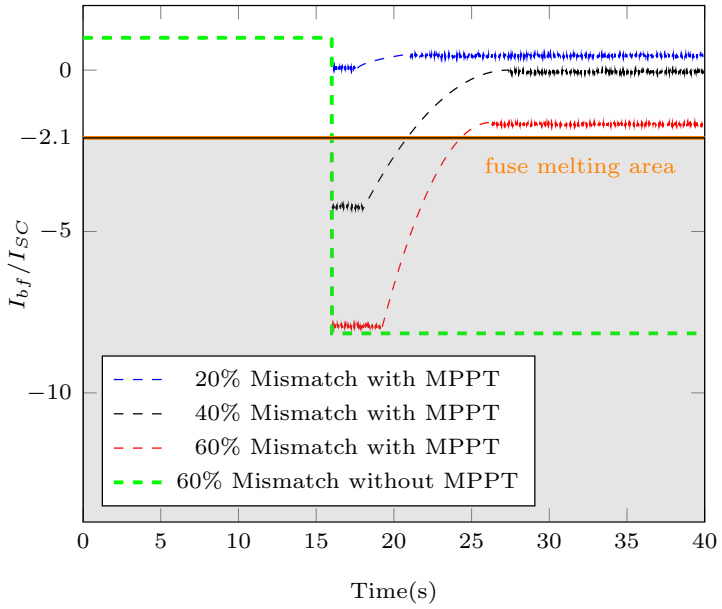

Fig. 3. Back-feeding current with different mismatch.

III. Feature Extraction and Sparse Representation

A. Data Acquisition

As discussed in the preceding section, DC-side short circuit faults detection is challenging. Our purpose is to design a fault detection method based on SR and evaluate its performance under the following conditions.

- Ambient temperature [^{o}C]: 10, 20, 30, 40, and 50.
- Irradiance [W/m^2]: 200, 400, 600, 800, and 1000.
- Mismatch percentage (%): 10, 20, 30, 40, 50, and 60.
- Fault impedance [Ω]: 0, 5, 15, 25.

By combining the above conditions, 2250 cases for L-L faults and 1500 cases for L-G faults are obtained. Faults include the grid phase-A voltage's negative and positive peaks and its zero-crossing.

B. Feature Extraction

To detect a fault with high accuracy, eight new features are created and defined as follows:

1) Features 1 and 2: To distinguish faults and make algorithm scalable, the normalized version of V_{MPP} and I_{MPP} are constructed.

$$f_1 = \frac{V_{MPP}}{N_{mod} V_{OC}} \tag{1}$$

$$f_2 = \frac{I_{MPP}}{N_{str} I_{SC}} \tag{2}$$

Where N_{mod} is the number of module in a string, while N_{str} is the number of string in PV array. V_{OC} and I_{SC} are open circuit voltage and short-circuit current of a PV module, respectively.

2) Feature 3: is the difference between maximum power and actual power, which is presented as

$$f_3 = V_{MPP} I_{MPP} \frac{I_{irr}}{I_ref} - V_{PV} I_{PV} \tag{3}$$

Where I_{irr} is the instantaneous irradiance and I_{ref} is $1000W/m^2$. These feature increase the efficiency of the algorithm under uniformly distributed irradance.

3) Features 4 and 5: These features are fill factor (FF) and Area under I-V curve which are fault indicator and sensitive to voltage and current change due to faults and shading conditions.

$$f_4 = \frac{V_{MPP} I_{MPP}}{V_{OC} I_{SC}} \tag{4}$$

$$f_5 = \int_0^{V_{OC}} V(I) dI \tag{5}$$

4) Features 6 and 7: These features are defined as the rate of change of resistance at the middle point between OC and maximum power points and created to distinguish SC faults under low mismatch and high impedance.

$$f_6 = \frac{dV}{dI}|_{V_{OC}} \tag{6}$$

$$f_7 = \frac{dV}{dI}|_{mV_{OC}} \tag{7}$$

5) Feature 8: It is defined as thermal voltage which helps to distinguish faults occurs due to low mismatch, and high fault impedance and which are very close to normal operation of low irradiance and high temperature.

$$f_8 = \frac{(2V_{MPP} - V_{OC})(I_{SC} - I_{MPP})}{I_{MPP} - (I_{SC} - I_{MPP}) \ln(\frac{I_{MPP - I_{SC}}}{I_{MPP}})} \tag{8}$$

6) Features 9, 10 and 11: Now, $l_1 - norm$, $l_2 - norm$ and $l_\infty - norm$ to extract new features and reduce dimension of feature space $f = \log[f1 \quad f2 \quad ... \quad f8]$. $l_1 - norm$, $l_2 - norm$, and $l_\infty - norm$ of original feature can be obtained as follows:

$$f_9 = ||f_i||_1 = \sum_{i=1}^{8} |f^i| \tag{9}$$

$$f_{10} = ||f_i||_2 = \sum_{i=1}^{8} |f^i|^{0.5}, \tag{10}$$

$$f_{11} = ||f_i||_\infty = \max\{|f^i|\}, \ j \in \{1, \cdots, 8\}. \tag{11}$$

C. Sparse Representation

Before explaining the SR method, it is worth mentioning that SR has been used successfully for face recognition [23], speaker verification [24], partial discharge detection [25], and fault diagnosis and location in power grids [26]. SR is a research area of interest within the field of compressive sensing (CS), and it can be defined as a feature vectors' dictionary extracted from measured data. The new input is classified according to the feature vectors' dictionary. The SR vector can be non-zero for a specific vector, and negligible (or zero) for other vector. The non-zero elements determine the class of input signal. The main goal is to characterize the input signal as a linear combination of feature vectors in the dictionary.

We choose *M* classes each of which has $n_i, \ i \in [1, \cdots, M]$ feature vectors including *k* subjects. Then, feature matrix can be defined as:

$$\mathbf{F}_i = [f_{i1}, f_{i2}, \cdots, f_{in_i}] \in \mathbb{R}^{k \times n_i}, \ i \in [1, \cdots, M], \tag{12}$$

where f_{ij} indicates the *j*th feature vector of the *i*th class. The feature vector of $y \in \mathbb{R}^p$ is presented as:

$$y = \mathbf{A}x. \tag{13}$$

In the case of $k > p$, (13) is over-determined so it has a unique solution. When $k < p$, (13) is under-determined, and there are multiple solutions for x. In our application, (13) can be rearranged as follows:

$$\begin{bmatrix} f_1 \\ \vdots \\ f_{11} \end{bmatrix} = \begin{bmatrix} f_{1,1} & \cdots & f_{1,M} \\ \vdots & \ddots & \vdots \\ f_{11,1} & \cdots & f_{11,M} \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_M \end{bmatrix} \tag{14}$$

where $y = [f_1, \cdots, f_{11}]^T \in \mathbb{R}^{11 \times 1}$ is the feature vector for test. The vectors $F_1 = [f_{1,1}, \cdots, f_{11,1}]^T \in \mathbb{R}^{2 \times 1}$ and $F_M = [f_{1,M}, \cdots, f_{11,M}]^T \in \mathbb{R}^{2 \times 1}$ are feature vectors for the 1st and the *M*th samples, respectively. The X_1 corresponds to the 1st sample, and X_M corresponds to the *M*th sample.

To find the sparsest solution for x, $l_1 - norm$ minimization can be used, as it is already utilized in many areas such as computer vision, sensor networks, data compression, and image processing. Many optimization methods such as augmented Lagrange multipliers, greedy algorithm, homotopy, iterative

shrinkage-thresholding, Bregman iterative algorithm, subspace pursuit, CoSaMP, linear programming, proximal gradient and and gradient projection (GP) are reported for $l_1 - norm$ in the literature [27]. The performance of these algorithms varies in terms of accuracy. For our case, the following algorithms are applicable: basis pursuit de-noise [28], and primal-dual interior point [29].

$l_0 - norm$ minimization, written as follows, provides a spare solution which finds the non-zero entries of x in (13):

$$\begin{aligned} \tilde{x}_0 &= \operatorname{argmin} \|x\|_0 \\ y &= \mathbf{A}x, \end{aligned} \tag{15}$$

which is known as an NP-hard problem [27]. The $l_1 - norm$ minimization in (16) provides the sparsest solution for x with mild conditions and high probability:

$$\begin{aligned} \tilde{x}_1 &= \operatorname{argmin} \|x\|_1 \\ y &= \mathbf{A}x. \end{aligned} \tag{16}$$

As it is difficult to use (15) in practice due to the noisy measurement data, we rearrange (15) as the stable $l_1 - norm$ minimization which is subject to an $l_2 - norm$ inequality as follows:

$$\tilde{x}_1 = \operatorname{argmin} \|x\|_1 \quad \text{s.t.} \quad \|Ax - y\| \leq \varepsilon \tag{17}$$

(17) indicates that the noise energy is restricted by ε. In the literature, there are several algorithms to solve (16) and (17), most of them, however, are not appropriate to provide a sufficient solution for a sparse x [30]. Some of those algorithms needs information regarding the sparsity of the vector x, and some others require an orthogonal basis set. As, none of them are applicable to our cases, we use the Primal-Dual interior point (PDIP) and Basis Pursuit De-noise (BPDN) algorithms to solve (16) and (17), respectively.

The solution obtained for $\tilde{x}_1$, which includes the non-zero terms associated with some classes, may not be related to the feature vector of one specific class. To avoid such a situation, a residual function can be used to minimize $l_2 - norm$ of the difference between the obtained and the actual signal feature vectors as follows:

$$r_i = \|y - A\tilde{x}_1(i)\|_2, \quad i = 1, ..., M, \tag{18}$$

where $\tilde{x}_1 \in R^{px1}$ is similar to $\tilde{x}_1$ except the n_i entry corresponding to the i^{th} class set, which is equal to zero. The residual of (18) is calculated for the test sample y that corresponds to the class with minimum residual and all M classes. Algorithm 1 presents the overall SRC based fault detection procedure.

Algorithm 1: Sparse Representation Classifier.

Inputs: Matrix of features(training) samples. $\mathbf{F}_i = [f_1, f_2, \cdots, f_n] \in \mathbb{R}^{k \times n},\ i \in [1, \cdots, M]$, $y \in \mathbb{R}^p$, and an error $\epsilon > 0$.
Normalize $\mathbf{F}$ to obtain unit $l_2 - norm$.
Solve the l_1 or l_2 -minimization problem:
$\tilde{x}_1 = \operatorname{argmin} \|x\|_1 \quad \text{s.t.} \quad y = \mathbf{A}x.$
$\tilde{x}_1 = \operatorname{argmin} \|x\|_1 \quad \text{s.t.} \quad \|Ax - y\| \leq \varepsilon$
for *i=1 to M* **do**
  Compute the residuals $r_i = \|y - A\tilde{x}_1(i)\|_2$
end
Output: $y = \operatorname{argmin}_i r_i$.

IV. Simulation Results

To verify and evaluate the effectiveness of the proposed method, the introduced test cases are simulated in PSCAD, and MATLAB is used to implement SR algorithm. To further validate our proposed method, a comparison is performed between Artificial Neural Network (ANN) and SR. An electrical power grid, is modelled, which is supplied by 100 PV panels equipped with bypass diode in parallel for each panel. It is worthwhile to notice that the system is not grounded. The challenges to detect DC side faults of PV arrays are already discussed in Section II. The output voltage and current of the PV are collected for several short-circuit faults under different irradiance levels to evaluate the performance of the proposed method.

Results obtained from the proposed method in the cases of L-L and L-G faults are shown in Table I. Results clarify the high accuracy of SR method in detection of L-Land L-G faults for PV array. Further analysis shows that when the mismatch percentage is greater than 30% for L-L faults with any fault impedance, the method provides $96,2\%$ detecting accuracy. Although SR has a high accuracy for fault detection in the case of high mismatch percentage, because of the minimal effect of shorted modules on PV system, under the conditions of low mismatch percentage and high fault impedance the accuracy of detecting fault decreases. Nevertheless, SR still has an accuracy rate higher than 72% and 65.5% under 20% and 10% mismatch percentage, respectively. In the case of L-G faults, the method also provides an accurate fault detection. In cases where mismatch percentage is lower than 40%, the accuracy is higher than 99.6%, while it will be more than 85% if mismatch is higher than 40%.

TABLE I
Result of Simulation Cases Under Various Conditions

Fault Types	Mismatch Percentage	Fault Impedance 0	5	15	25	Average Accuracy
L-L Faults	60	100%	100%	100%	100%	100%
	50	100%	100%	100%	100%	100%
	40	100%	100%	95.2%	90.3%	96.7%
	30	99.1%	93.6%	84.7%	75.4%	88.2%
	20	94.7%	70.5%	63.6%	59.2%	72%
	10	82.3%	68.7%	59.8%	49.8%	65.2%
L-G Faults	10	100%	100%	100%	100%	100%
	20	100%	100%	100%	100%	100%
	30	100%	100%	100%	95.3%	98.8%
	40	100%	100%	91.5%	85.4%	94.5%
	50	100%	95.2%	89.6%	75.5%	90.5%
	60	100%	98.1%	65.7%	61.2%	81.3%

The simulation also includes L-L and L-G fault detection in PV arrays with a blocking diode, and the results are shown in Table II. The result verify the effectiveness of our proposed method in PV arrays equipped with blocking diode, although it is designed to detect the short-circuit faults in a PV array without blocking diodes. The proposed method is successful in detecting faults with mismatch percentage over 30% for L-

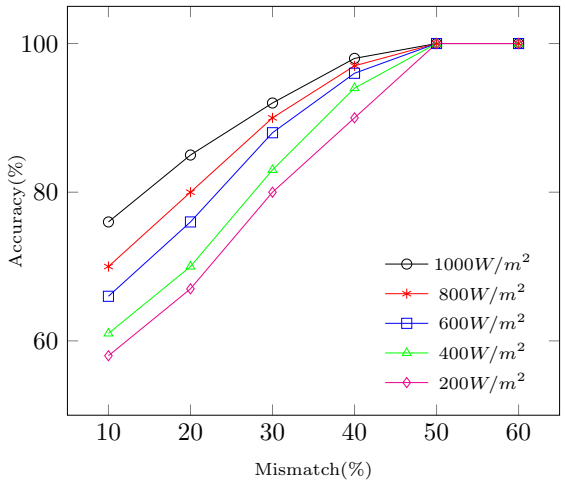

Fig. 4. Detection accuracy for L-L short-circuit faults of different mismatch percentage.

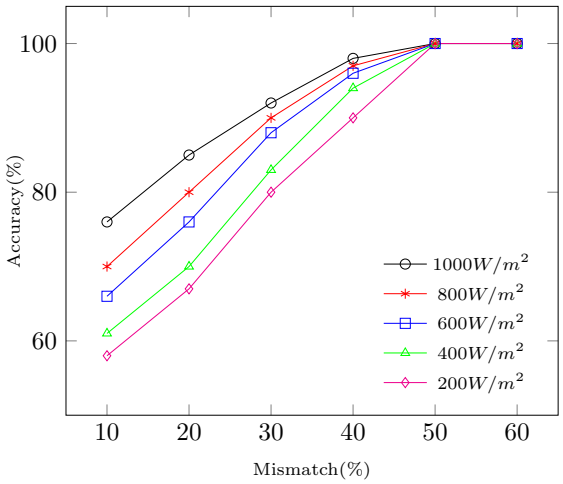

Fig. 5. Detection accuracy for L-G short-circuit faults of different mismatch percentage.

L faults and with mismatch below 40% in case each string has blocking diodes.Moreover, Fig. 4 and Fig. 5 include the results which indicate that the accuracy of different mismatch percentage under various irradiance levels.

TABLE II
RESULT OF SIMULATION CASES WITH BLOCKING DIODE UNDER VARIOUS CONDITIONS

Fault Types	Mismatch Percentage	Fault Impedance 0	5	15	25
L-L Faults	60	✓	✓	✓	✓
	50	✓	✓	✓	✓
	40	✓	✓	✓	×
	30	✓	✓	×	×
	20	✓	×	×	×
	10	✓	×	×	×
L-L Faults	10	✓	✓	✓	✓
	20	✓	✓	✓	✓
	30	✓	✓	✓	✓
	40	✓	✓	✓	✓
	50	✓	✓	×	×
	60	✓	×	×	×

The Feed Forward Back Propagation Neural Network (FFBP-NN) multi-layer preceptor (MLP) method is used to verify the validity of the proposed method to detect the short-circuit faults. ANN consists of three layers called as input, hidden and output layers. The input signal is connected to the hidden layer's neurons through the input layer neurons. The output of each neuron is associated with a Gaussian radial, hyperbolic tangent sigmoid, or logistic sigmoid based function of a weighted sum of inputs. The hidden layer's neurons are similar to those of the output layer. In general, BP algorithm is used to train the MLP. The weights of the connections between hidden and input layers' neurons change iteratively with a specified learning rate.

F_i matrices are the inputs of ANN. There are many options to select the transfer function of the FFBP networks, learning and training. The results show that a combination of logistic sigmoid transfer function, gradient descent weight, Levenberg-Marquardt training and bias learning, and mean-square normalized error performance provides better results. Initially, the weights of ANN are selected at random, which changes the simulation results for each case. Therefore, in order to provide the best results, we repeat the simulations several times. As a consequence, FFBP-NN provides similar results with SR in the case of high-mismatch percentage and low-fault impedance for L-L faults, and in the case of low-mismatch and low-fault impedance for L-G faults. In such cases as low-mismatch and high-fault impedance, ANN performance is slightly better than SR's. The results for ANN are shown in Table III. Nonetheless, ANN requires to simulate each case many times for various combinations of the tuning parameters, which means its performance depends on the tuning parameter, while SR does not suffer from this practical issue.

TABLE III
RESULT OF SIMULATION CASES UNDER VARIOUS CONDITIONS

Fault Types	Mismatch Percentage	Fault Impedance 0	5	15	25	Average Accuracy
L-L	60	100%	100%	100%	100%	100%
	10	86.5%	73.8%	66.2%	58.2%	71.2%
L-G	10	100%	100%	100%	100%	100%
	60	100%	100%	76.2%	69.8%	86.5%

V. DISCUSSIONS

Results reveals that the detection accuracy of the proposed method for L-L faults decreases as mismatch percentage decreases and fault impedance increases, and vice versa. Results also demonstrates that the fault detection accuracy decreases for low-irradiance level cases.

The performance of the SR-based fault detection method was verified in various scenarios, we trained FFBP-NN to detect L-L and L-G faults and compare its results with SR's. SR method and FFBP-NN show quite similar accuracy, especially in the case of high-mismatch percentage for L-L faults and low-mismatch percentage for L-G faults with low short-circuit impedance. Although the performance of FFBP-NN is slightly better than SR's in the case of low-mismatch percentage for L-L faults and high-mismatch percentage for L-G faults with high short-circuit impedance, SR method has noticeable and practical benefits: (i) the implementation of SR

is easier than FFBP-NN; (ii) SR does not need training to tune several parameters such as learning selection and number of neurons to obtain acceptable fault detection accuracy; and (iii) the convergence of SR method is always consistent and guaranteed, while it depends on the weights for training in FFBP-NN. It is also worth noting that to obtain the best results by FFBP-NN, we re-tune several parameters that are computationally inefficient.

VI. ConclusionFuture Work

This study proposed a method to detect short-circuits faults in PV systems based on feature extraction and classification. More specifically, our proposed method uses SR for feature extraction for measured signals. The detection accuracy of our method was verified for short-circuit faults which occur under challenging conditions such as high impedance, low irradiance, and the effects of MPPT algorithm, numerous scenarios were investigated. Several compressive sensing algorithms were examined for l-norm minimization, which revealed that BPDN and PDIP algorithms ensure better performance for extracting feature and effective computation. Compared to the present method in the literature, our proposed method reduces computational cost and number of sensors, and it benefits from a fast response to the faults.

As our future direction, it is also required to set-up an experiment for PV array and real cases to verify fault detection accuracy. A SR classifier can also be considered to train support vector machine or ANN for future research.

References

[1] F. E. Sahin and M. Yılmaz, "High concentration photovoltaics (hcpv) with diffractive secondary optical elements," in *Photonics*, vol. 6, no. 2. Multidisciplinary Digital Publishing Institute, 2019, p. 68.

[2] H. S. Sahu, S. K. Nayak, and S. Mishra, "Maximizing the power generation of a partially shaded pv array," *IEEE Journal of Emerging and Selected Topics in Power Electronics*, vol. 4, no. 2, pp. 626–637, June 2016.

[3] 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 Journal of Photovoltaics*, vol. 5, no. 3, pp. 982–997, May 2015.

[4] Y. Zhao, J. de Palma, J. Mosesian, R. Lyons, and B. Lehman, "Line–line fault analysis and protection challenges in solar photovoltaic arrays," *IEEE Transactions on Industrial Electronics*, vol. 60, no. 9, pp. 3784–3795, Sep. 2013.

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

[6] R. Platon, J. Martel, N. Woodruff, and T. Y. Chau, "Online fault detection in pv systems," *IEEE Transactions on Sustainable Energy*, vol. 6, no. 4, pp. 1200–1207, Oct 2015.

[7] S. Vergura, G. Acciani, V. Amoruso, G. E. Patrono, and F. Vacca, "Descriptive and inferential statistics for supervising and monitoring the operation of pv plants," *IEEE Transactions on Industrial Electronics*, vol. 56, no. 11, pp. 4456–4464, Nov 2009.

[8] P. Guerriero, F. Di Napoli, G. Vallone, V. d'Alessandro, and S. Daliento, "Monitoring and diagnostics of pv plants by a wireless self-powered sensor for individual panels," *IEEE Journal of Photovoltaics*, vol. 6, no. 1, pp. 286–294, Jan 2016.

[9] P. Guerriero, G. Vallone, V. d'Alessandro, and S. Daliento, "Innovative algorithm for true maximum detection based on individual pv panel sensor network," in *2013 International Conference on Clean Electrical Power (ICCEP)*, June 2013, pp. 42–47.

[10] B. Andò, S. Baglio, A. Pistorio, G. M. Tina, and C. Ventura, "Sentinella: Smart monitoring of photovoltaic systems at panel level," *IEEE Transactions on Instrumentation and Measurement*, vol. 64, no. 8, pp. 2188–2199, Aug 2015.

[11] Y. Hu, J. Zhang, W. Cao, J. Wu, G. Y. Tian, S. J. Finney, and J. L. Kirtley, "Online two-section pv array fault diagnosis with optimized voltage sensor locations," *IEEE Transactions on Industrial Electronics*, vol. 62, no. 11, pp. 7237–7246, Nov 2015.

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

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

[14] M. N. Akram and S. Lotfifard, "Modeling and health monitoring of dc side of photovoltaic array," *IEEE Transactions on Sustainable Energy*, vol. 6, no. 4, pp. 1245–1253, Oct 2015.

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

[16] Y. Zhao, B. Lehman, R. Ball, J. Mosesian, and J. de Palma, "Outlier detection rules for fault detection in solar photovoltaic arrays," in *2013 Twenty-Eighth Annual IEEE Applied Power Electronics Conference and Exposition (APEC)*, March 2013, pp. 2913–2920.

[17] M. K. Alam, F. Khan, J. Johnson, and J. Flicker, "Pv ground-fault detection using spread spectrum time domain reflectometry (sstdr)," in *2013 IEEE Energy Conversion Congress and Exposition*, Sep. 2013, pp. 1015–102.

[18] T. Takashima, J. Yamaguchi, and M. Ishida, "Fault detection by signal response in pv module strings," in *2008 33rd IEEE Photovoltaic Specialists Conference*, May 2008, pp. 1–5.

[19] Y. Hu, W. Cao, J. Ma, S. J. Finney, and D. Li, "Identifying pv module mismatch faults by a thermography-based temperature distribution analysis," *IEEE Transactions on Device and Materials Reliability*, vol. 14, no. 4, pp. 951–960, Dec 2014.

[20] F. Harrou, A. Dairi, B. Taghezouit, and Y. Sun, "An unsupervised monitoring procedure for detecting anomalies in photovoltaic systems using a one-class support vector machine," *Solar Energy*, vol. 179, pp. 48 – 58, 2019. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0038092X18312209

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

[22] J. C. Wiles, "Photovoltaic system grounding," *Southwest Technology Development Institute College of Engineering New Mexico State University October*, 2012.

[23] J. Wang, C. Lu, M. Wang, P. Li, S. Yan, and X. Hu, "Robust face recognition via adaptive sparse representation," *IEEE Transactions on Cybernetics*, vol. 44, no. 12, pp. 2368–2378, Dec 2014.

[24] J. M. K. Kua, J. Epps, and E. Ambikairajah, "i-vector with sparse representation classification for speaker verification," *Speech Communication*, vol. 55, no. 5, pp. 707–720, 2013.

[25] M. Majidi, M. S. Fadali, M. Etezadi-Amoli, and M. Oskuoee, "Partial discharge pattern recognition via sparse representation and ann," *IEEE Transactions on Dielectrics and Electrical Insulation*, vol. 22, no. 2, pp. 1061–1070, April 2015.

[26] K. Jia, C. Gu, L. Li, Z. Xuan, T. Bi, and D. Thomas, "Sparse voltage amplitude measurement based fault location in large-scale photovoltaic power plants," *Applied energy*, vol. 211, pp. 568–581, 2018.

[27] A. Y. Yang, Z. Zhou, A. G. Balasubramanian, S. S. Sastry, and Y. Ma, "Fast ℓ_1-minimization algorithms for robust face recognition," *IEEE Transactions on Image Processing*, vol. 22, no. 8, pp. 3234–3246, Aug 2013.

[28] E. Candes and J. Romberg, "l1-magic: Recovery of sparse signals via convex programming," *URL: www. acm. caltech. edu/l1magic/downloads/l1magic. pdf*, vol. 4, p. 14, 2005.

[29] E. J. Candes and T. Tao, "Near-optimal signal recovery from random projections: Universal encoding strategies?" *IEEE Transactions on Information Theory*, vol. 52, no. 12, pp. 5406–5425, Dec 2006.

[30] J. Yang and Y. Zhang, "Alternating direction algorithms for \ell_1-problems in compressive sensing," *SIAM journal on scientific computing*, vol. 33, no. 1, pp. 250–278, 2011.