Iraqi Journal for Computer Science and Mathematics

Journal Homepage: http://journal.esj.edu.iq/index.php/IJCM

e-ISSN: 2788-7421 p-ISSN: 2958-0544

Development of an Intelligent Detection Method of DC Series Arc Fault in Photovoltaic System Using Multilayer Perceptron and Bi-Directional Long Short-Term Memory

Alaa Hamza Omran[1,*], Dalila Mat Said[2], Siti Maherah Hussin[2], Sadiq H. Abdulhussein[3], Nasarudin Ahmad[2], Haidar Samet[4]

[1]Centre of Electrical Energy Systems (CEES), School of Electrical Engineering, Universiti Teknologi Malaysia (UTM), Malaysia. And university of Information Technology and Communications, Iraq. e-mail: alaa1990@graduate.utm.my

[2]Centre of Electrical Energy Systems (CEES), School of Electrical Engineering, Universiti Teknologi Malaysia (UTM), Malaysia (e-mail: dalila@utm.my, sitimaherah@utm.my, e-nasar@utm.my)

[3]Department of Computer Engineering, University of Baghdad, Baghdad, Iraq (e-mail: sadiqhabeeb@coeng.uobaghdad.edu.iq)

[4]School of Electrical and Computer Engineering, Shiraz University, Iran; and Department of Electrical Engineering, Eindhoven University of Technology, Netherlands (e-mail: samet@shirazu.ac.ir)

*Corresponding Author: Alaa Hamza Omran

DOI: https://doi.org/10.52866/ijcsm.2023.02.03.014
Received June 2023; Accepted August 2023; Available online August 2023

ABSTRACT: A DC series arc fault is one of the significant sources of electrical wiring fires in residential buildings. The production of extremely high temperatures may lead to the ignition of nearby combustible materials. The applications of arc fault diagnosis based machine learning are a global interest due to the immense challenge to create an accurate and efficient detection method. In this paper, a detection and classification method using a multilayer perceptron incorporated with Bi-Directional Long short-term Memory (MLP-BiLSTM) is proposed. In order to achieve this goal, nine series arc fault models are used in conjunction with data from real-world observations for simulation purposes using Power System Computer Aided Design (PSCAD) software. The simulation and experimental results confirm that the accuracy of the proposed detection and classification method reaches 99%, which results in that the methodology is believed to be accurate for DC series arc fault detection and classification in the PV system with relatively high accuracy.

Keywords: DC Fault, Series Arc Fault, Bi-LSTM, CNN, Fault detection, MLP, PV system.

1. INTRODUCTION

THere are many reasons which can cause electrical fires such as leakage, over-current, arc fault, or electrical appliances overheating. Residential power lines are often subject to arc fault due to several reasons such as virtual contacts, loss of electrical connections, or aging cables. These accidents can lead to extremely high temperatures over 20,000k [1], which results in fire ignition. In the case where an effective method of arc fault identification to perform the necessary interruption is absent, electrical fires or even explosions may occur [2]. In 2018, electrical fires were ranked first among all the other categories of fires, according to the Fire Rescue Bureau of China Emergency Management Department on China in terms of fire and alarm [3], wwhere the proportion of electrical fires caused by arc fault reached up to 34.6%. Therefore, the world should be more attentive towards avoiding electrical fires that result from the arc fault [4, 5]. The Under Laboratories (UL) Standard UL1699 [6] defines arcing as a continuous luminous that results from the electricity discharge past an insulating medium. Often, the arcing is accompanied by electrodes partial volatilization, which is an extremely complicated process of electromagnetic reaction[2]. It can be classified into three categories according to the

position of the fault occurrence, as illustrated in Figure 1. The series arc is the most common of these types of faults [7–9].

The other two types of faults are the parallel arc fault, which develops between dissimilar poles like positive and negative, or between distinct branches, and the ground arc fault, which can develop between any part of the system and the ground.

The stability and operational safety of the PV system are crucial concerns for the electrical grid and consumers [10–12]. ground and Parallel arc faults, on the other hand, are easier to identify with the over-current approach [13]. Because to the large current that typically accompanies them. Due to a little drop in line current following arc initiation [14].identifying a series arc fault is more difficult. Arc faults and their accompanying dangers can be avoided if they are discovered as soon as feasible. As a result, many scientists have spent the last few years looking for effective techniques to identify several types of arc fault [15–17]. Many strategies for dealing with this issue are discussed.

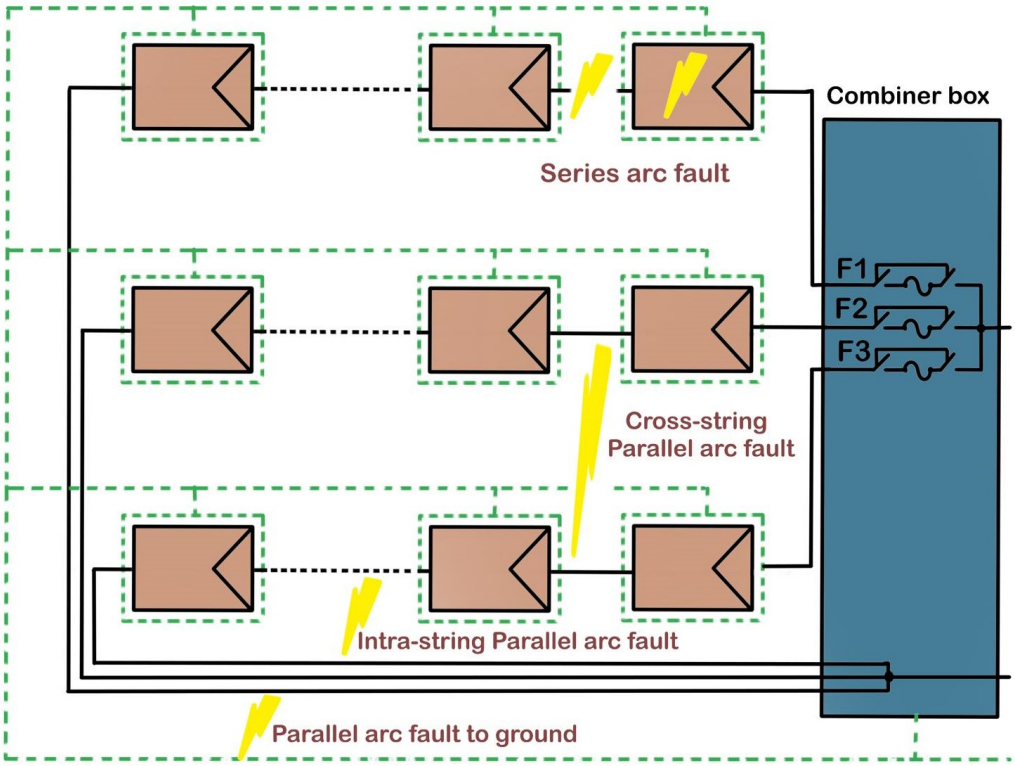


FIGURE 1. Examples of different DC arc faults in a PV array.

1.1 RELATED WORKS

It has become more clear that the use of intelligent system design and implementation is essential to the creation of new and improved products across a wide range of technological domains [18–23]. Automatic defect detection [24], uses artificial intelligence, namely an artificial network (ANN) in combination with more traditional analytic methods. Module temperature and irradiance are used as inputs to a two-layer artificial neural network (ANN) that calculates an estimated output power. After that, the fault class and occurrence are determined by comparing the estimated power to the measured power. The analytical Equation is then used to compare the open-circuit voltage and short-circuit current measurements, with the resulting values being integrated. This Equation is also characterized by its high precision, its speedy identification response, and its compact construction. There is no need for advanced methods or knowledge of systems. Bayesian Belief Networks (BBN) are the second AI method mentioned in [25] where they are said to be used. The author's technique to tracing the origins of flaws was conceptually similar to that described in [26] In this study, BBN was created to show how measuring factors, which are symbols for specific gadgets, depend on one another. BBN's measurements lead to an instantaneous determination about the causes of a malfunction. [27–29]. fuzzy-logic-based algorithms were first introduced. In particular, another study presented [27] method based on the notions of fuzzy mathematics and evidence. In the proposed research, a novel and exact framework of fault detection structure was created, and then the differences between the predicted and determined values were consolidated using a data fusion methodology and fuzzy mathematics.

According to the research, large-scale PV arrays can benefit from data fusion's improved ability to localize defects and manage uncertainty resulting from inter-array interactions [30], the authors propose an additional system of experts for PV monitoring that queries the system database to find the energy loss caused by inverter malfunctions. This method involves distinguishing between the value recorded in the database and the real-time output in a fault-free manner. Notably, this method requires monthly updates and cannot be used for real-time tracking. Therefore, the system does not attribute the power savings to the existence of shade. This method, like the shading strategy, takes into account the current status of the monitoring system. Even though it needs a wide variety of measuring instruments, it nonetheless simplifies operations and protects PV plants. Database requirements are used, data collected from the PV system is analyzed, and the plant's health is confirmed. Automatic error diagnosis is achieved via the application of artificial intelligence techniques. In [31], traditional analytic techniques were integrated with ANN. The power output was predicted using a two-layer ANN, which took into account both module irradiance and temperature. The defect and its nature were determined by comparing the predicted power with the actual energy. In order to find the source of the problem, the measured power was compared to the estimated power. In [32], researchers devised an artificial neural network–based method for identifying PV panel failure. A technique for detecting series arc defect using support vector machines was presented in [33] with better results than using artificial neural networks. For series arc defect identification, ref. [34] developed a technique using fuzzy mathematics and evidence theories. Data fusion and fuzzy mathematics were used to integrate the difference between anticipated and calculated values, resulting in a novel framework for accurate defect diagnosis. Large-scale applications of fault localization and PV arrays benefit greatly from data fusion's ability to manage uncertainty instances arising from their interaction. When applied to extremely transient and non-stationary data, the hidden Markov model (HMM) proved effective for detecting series arc failure in PV systems [17]; Instead of utilizing actual records from experience or a predetermined arc fault model, the authors of [35], offer a GAN-Deep approach for identification utilizing deep learning technology; this method generates multiple records of arc fault signals. The suggested approach, however, only identified fake records and did not pick up on series arc defect. While [36], proposes a method for detecting series arc faults, it is plagued by the disappearing problem.

1.2 CHALLENGES AND LIMITATIONS

The defect and its nature were determined by comparing the predicted power with the actual energy. In order to find the source of the problem, the measured power was compared to the estimated power. In [32], researchers devised an artificial neural network–based method for identifying PV panel failure. Ref. Since the PV system can continue to feed electricity despite an instance of arc fault as daylight, this increases the risk of fire and even electrical hazard of shock, making reliable identification of a serial arc fault in DC a challenging operation. Because the PV system module is a current-limited device, its behavior during a malfunction may be quite similar to that during normal operation. Furthermore, the DC waveform does not have a zero-crossing point, but the AC waveform does, making series arc fault identification simpler than it is with the DC waveform. Finally, the measurement of the PV system determines the number of conductor joints present, and a series arc failure might happen at any of them. There are many challenges for the detection and protection system design comes from many reasons such as [37]:

1. Many factors, including electrode materials and geome- tries, levels of voltage and current, type of load, and type of source, would influence the intensity and variation of arc noise. For example, when the non-arc loop current still the same, the discharge of arc will be highly stable voltage levels, leading to lower standard deviation of fault current (less variance of current). Furthermore, because the arc is symmetrical formed via the radially center of electrodes, the components of fault current frequency with rounded tip electrodes are visibly lower than as compared with those with flat tip electrodes.

2. The long cables of PV would behave as an antenna, where Radio Frequency (RF) noise will be picking up Mainly in the 100 kHz to 500 MHz frequency range. RF responsiveness is also affected by varied system topologies and sizes, which may change the arc noise profile. Cables of PV could also include an inductive component that works as a low pass filter, reducing the noise generated by arc faults.

3. high frequency electromagnetic interference noise which will be introduced as the existing of the electronic loads including a DC/DC converter and a DC/AC inverter into the circuit, potentially causing a nuisance trip.

4. Unwanted trip may be produced as a result of the Crosstalk effect.

5. In some PV system applications, the inverters of trans- formerless are used, leading to an The AC side of the circuit (the fundamental component of the 50 Hz grid and its harmonics) will inject noise.

6. Arc faults behave similarly to step changes caused by load shifting (turning off the inverter or converter), mechanical vibration caused by wind, fast moving clouds, shutdown of the system, operation of partial shading, and power adaption by inverter.

7. The detection of the arc fault has been widely investigated for AC electric power systems. These methods, however, cannot be directly applied to DC systems because AC electric systems totally differ from DC systems in terms of sources of voltage, loadings, topologies of system, and the kind and number of power electronics circuits. And this is essential in developing method for detecting and distinguishing arc faults from normal operational conditions.

Furthermore, the method based Artificial Intelligence may encounter a variety of difficulties:

a. Many AI systems suffer from a decline in detection performance.

b. Issues with over-fitting and under-fitting lower detection accuracy.

c. Extremely complex calculations

d. The Disappearing Issue

Therefore, several critical issues should be taken into consideration during the development of the finding technique based on AI, such as (accuracy, fault classification ability, reliability, safety, and Computational Complexity /Effort). These issues can present a series of difficulties that can affect the operation of the detection method and degradation in the performance.

1.3 CONTRIBUTION

The suggested method's key contribution is the development of a proactive and smart detection and classification approach that can accurately distinguish, diagnose, and classify the DC series arc fault-based AI with simplified and highly accurate of 99%, here Multilayer Perceptron (MLP) incorporated with Bidirectional Long Short-Term Memory (Bi-LSTM) is used. Due to the improved speed and accuracy in detecting the DC series arc fault, the suggested technique improved both the safety and reliability of DC network protection. At last, we have a detection algorithm that accurately distinguishes between certain series arc fault, the other PV system failure that can happen, and the background noise. We can summarize the contribution and novelty of this paper as follows :

1. To best of our knowledge, In no other research have two distinct AI algorithms been used in a two-stage process to detect and classify DC series arc faults. A novel approach to detecting shifts of DC series arc fault is developed to detect any abnormal signals and send a trigger to the CNN along with the analysis and classification capabilities of the Bi-LSTM model.

2. In this work, a combining of two models to create a model that uses the strengths of CNN and Bi-LSTM in order to capture the features extracted using CNN and use them as Bi-LSTM input. The strength of this model is that it uses CNN convolution layers to extract as much information as possible from the data. Then, this output will become the input to the Bi-LSTM; this feature will allow the data to be kept in a chronological order in the both directions.

3. In order to create the required dataset and produce massive dataset records of distinct scenarios, nine various arc models are created.

4. The effectiveness of the detection approach is further demonstrated using a real-world dataset.

2. UTILIZED METHODS FUNDAMENTALS

2.1 MULTILAYER PERCEPTRON (MLP)

MLP is considered as one of the most popular artificial neural network (ANN) algorithms. It is used to tackle the problem of supervised learning because of its ability to model complicated datasets despite the existence of noises or outliers. This mechanism has abounded directed of the acyclic network that consists of different layers with many neurons in each one of them. These layers are an input layer, an output layer, and one or more hidden layers. The aid of weights internally connects these neurons. A nonlinear mapping exists between the input and output time series. The neurons are structured as follows: In all MLP layers, as shown in Figure 2, each neuron is assigned to an activation function, which is

implemented in every neuron for the weighted sum of input data [38]. The neuron output can be formulated as in Equation (1)

$$u_j = f(b_0 x_0 + \sum_{i=1}^{n} b_i x_i) \tag{1}$$

where b_0 is a constant, and x_0 is always one. x_i b_i represents the values of input and bias weights, respectively. Sigmoid function, which is illustrated in Equation (2), is utilized in the nonlinear mapping between the input and output time series

$$f(u) = \frac{1}{1 + exp(-u)} \tag{2}$$

The input vector of the next layer neurons is the output of each neuron. The propagation of the transmission signal remains until it reaches the last layer. The optimal output can be computed through the repetitively adjusted network biases until a specific criterion is obtained. The received error in each iteration in the network is back-propagated, which leads to it being considerably minimized. The accurate relationship between the input and output data time series can be determined through the procedure of training that involves achieving suitable values for each weight.

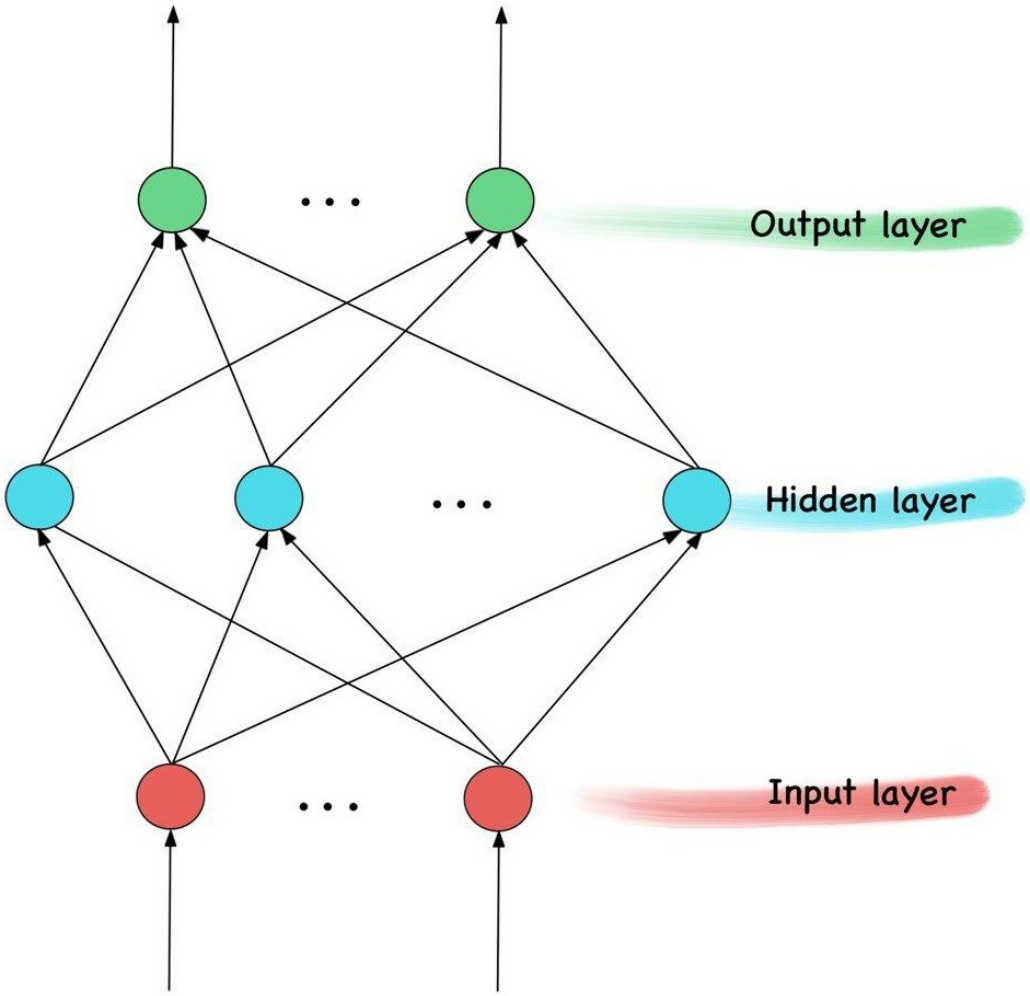


FIGURE 2. **MLP Network Architecture [38].**

2.2 BI-DIRECTIONAL LONG SHORT MEMORY (BI-LSTM)

Bi-LSTM is composed of Bi-directional Recurrent Networks (BiRNN) and Long Short-Term Memory (LSTM) [39],where Recurrent Neural Network (RNN) can be considered as a distinctive evolution of Artificial Neural Networks (ANN) that is used for sequences and time-series data processing. Also, it can encode the dependencies between inputs. Nev- ertheless,

RNN is not suitable for a long sequence of data because a problem of exploding and vanishing state against its gradient may occur. Therefore, Long Short-Term Memory (LSTM) is formed to eliminate the difficulties of the long- term in RNN. It consists of several different gates, which are the input gate for the input layer, and forget and output gates for the output layer [40]. However, the use of LSTM and RNN methods can only obtain the data from its prior state; therefore, another enhancement is made by using the Bidirectional Recurrent Neural Network (Bi-RNN) method. The information in the Bi-RNN method is handled in both forward and backward manners [41]. Therefore, combining the two methods, Bi-RNN and LSTM, formed the Bi-LSTM method, as shown in Figure 3. The new method, Bi-LSTM, inherits the advantages of the two previous methods, where the ability of storage in cell memory in the LSTM and the capability of processing data in two directions in the Bi-RNN are used in the Bi-LSTM [42]. Moreover, the Bi-LSTM can process the information with dependence on the long-range. Equation (3) and (4) illustrates the forward function using L units in inputs and H of hidden units.

$$a_h^t = \sum_{l=1}^{L} x_l^t w_{lh} + \sum_{h=\grave{l},t>0}^{H} b_{\grave{h}}^{t-1} w_{\grave{h}h} \tag{3}$$

$$b_h^t = \Theta_h(a_h^t) \tag{4}$$

where the sequence input is represented by x_t, and the LSTM network input is represented by a_h^t at with h unit at time t, b_h^t represents the activation function of h at time t. The weight of the input l towards h is symbolized as w_{lh}. The weight of the h hidden units towards the $\grave{h}$ hidden unit is represented by $w_(hh')$ the activation function of h hidden unit is symbolized by Θ_h. The calculations of the Bi-LSTM backward are described by the Equation of (5) and (6).

$$\frac{\delta O}{\delta w_{hk}} = \sum_{t=1}^{T} \frac{\delta O}{\delta a_h^t} b_h^t \tag{5}$$

$$\frac{\delta O}{\delta a_h^t} = \grave{\Theta}_h(a_h^t)\left(\sum_{k=1}^{K} \frac{\delta O}{\delta a_h^t} w_{hk} + \sum_{h=l,t>0}^{H} \frac{\delta O}{\delta a_h^{t+1}} w_{\grave{h}h}\right) \tag{6}$$

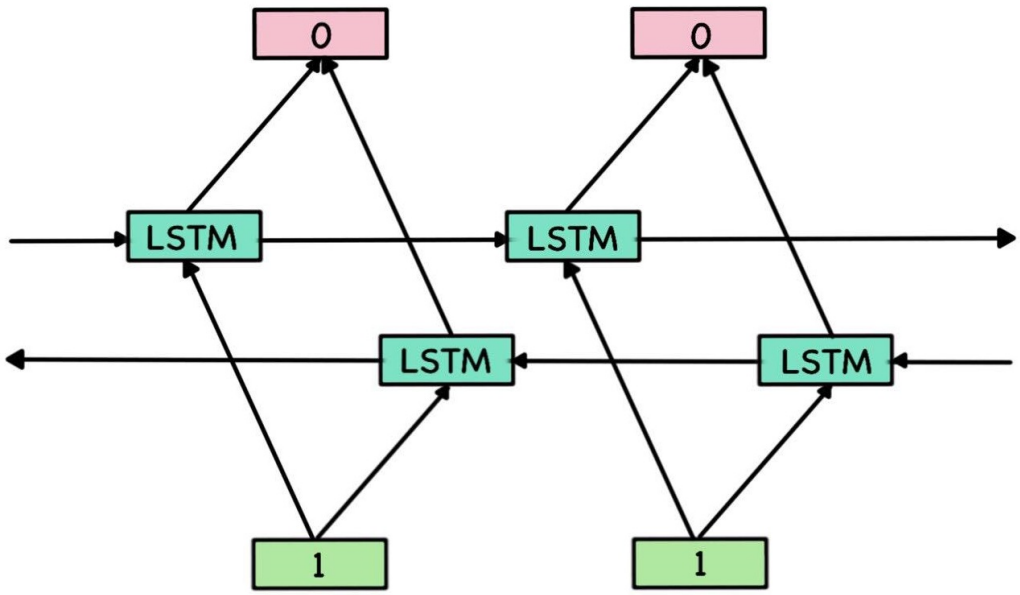


FIGURE 3. **General Bi-LSTM [43].**

3. METHODOLOGY

This section discusses the methodology of the research work; the general block diagram of Using the DC serial arc fault as an example, Figure 4 shows the suggested approach for detecting a failure in the PV system. An intelligent fault classification and detection model is also included in the mix, along with a PV model and nine distinct DC serial arc

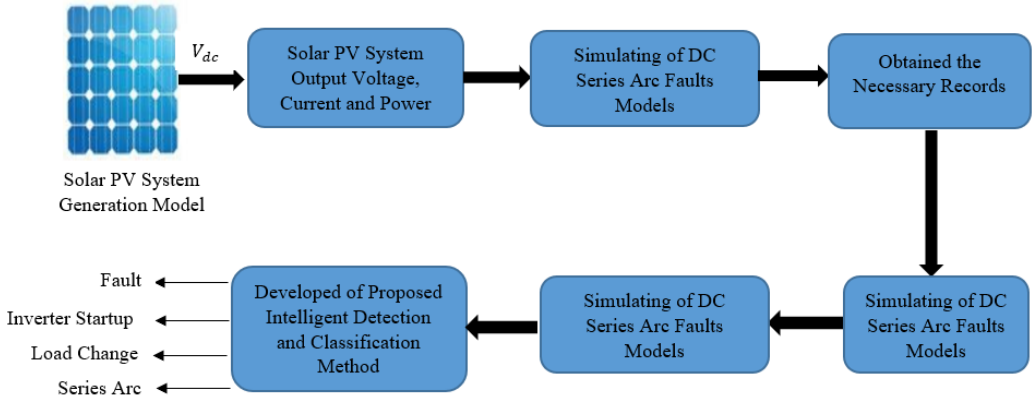


FIGURE 4. Block Diagram of Methodology.

Table 1. MODULES CHARACTERISTICS

Model	Polycrystalline HN-20/36
Maximum Power	20 W
Voltage at Maximum Power (V)	17.5 V
Current at Maximum Power (A)	1.16 A
Open Circuit Voltage (Voc)	21 V
Weight (kg)	2.2
Dimensions	350 × 25

fault models, respectively. The design procedures for this research are carefully discussed in more detail. In more details, Figure 5 presents the power system and the proposed method (detection and classification) in (a) and (b) respectively. Table I shows the various features of the PV modules that were employed.

3.1 UTILIZED ARC MODELS

Prefixes The design specifications of the PV system and its array components should be determined before proceeding with the main design steps. This gives an important indicator of how these proposed components worked well in their minimum requirements Nine different arc fault models are simulated, these models feature a number of fixed parameters that allow for the simulation of many scenarios. The mathematical models and their respective parameters are detailed below Guidelines for Graphics Preparation and Submission.

3.1.1. Ayrton Arc Model

Ayrton Arc Model: can be written as Equation (7). By adjusting its settings, this model may be used to mimic a variety of scenarios. The decline in the voltage (A) at the electrodes is one of the parameters, the voltage gradient (B), and length (L) of the arc being measured. C and D represent the nonlinear nature of the arc, distinguishing constants[44].

$$V_{arc} = A + BL + \frac{C + DL}{I_{arc}} \tag{7}$$

3.1.2. Steinmetz arc model

Figures The V–I Equation is a semi-empirical expression that is given in (8) and it is dependent on magnetite and carbon as its base. These constants are denoted by the letters A, C, and D. These factors are determined by the material that the electrodes are made of, as well as their arc length, and shape (Le) [45].

$$V_{arc} = A + \frac{C(Le + D)}{I_{arc}^{0.5}} \tag{8}$$

3.1.3. Nottingham arc model

The arcing phenomena may be modeled using this straightforward model's help, thanks to its inverse feature. The total length of the arcs and the substance of the electrode are both relevant criteria for determining the values of the two

constants A and B. In Equation (9) the uses is depending on the electrode material and arc length, the inverse current has the power of n. To what power (n)? a value close to 0.67 for arcing copper electrodes size in the range of 1-10 mm. Case studies have different parameters [46].

$$V_{arc} = A + \frac{B}{I_{arc}^{n}} \tag{9}$$

3.1.4. Hyperbolic approximation arc model

Equations (10-14) give a mathematical description of the features of the hyperbolic approximations arc model. The voltage source in this schematic is expressed by a (v_{gap})-dependent equivalent pulse force (e_{gap}), gap current (i_{gap}), (x_{gap}), (R_{gap}), the limit of the quenching phases, and the arc's burning intensity are all parameters of an electric arc (x_{crit}), e_{gap} slope λ, Dc voltage v_q before the arc (V_{dc}), and a specific variable used to control the v_q slope(α)[11].

$$e_{gap} = \frac{1}{2}(a + bx_{gap})(tanh(\lambda q) - tanh(\lambda(q-1))) \tag{10}$$

$$v_q = V_{dc}\left(\frac{1}{2} + \frac{1}{2}(tanh(\alpha(q-1)))\right) \tag{11}$$

$$v_{gap} = v_q + e_{gap} \tag{12}$$

$$q = \frac{x_{crit}}{x_{gap}} \tag{13}$$

$$R_{gap} = \frac{v_q}{i_{gap}} \approx \frac{v_{dc}}{I_{Load}} e^{2\alpha(q-1)} \tag{14}$$

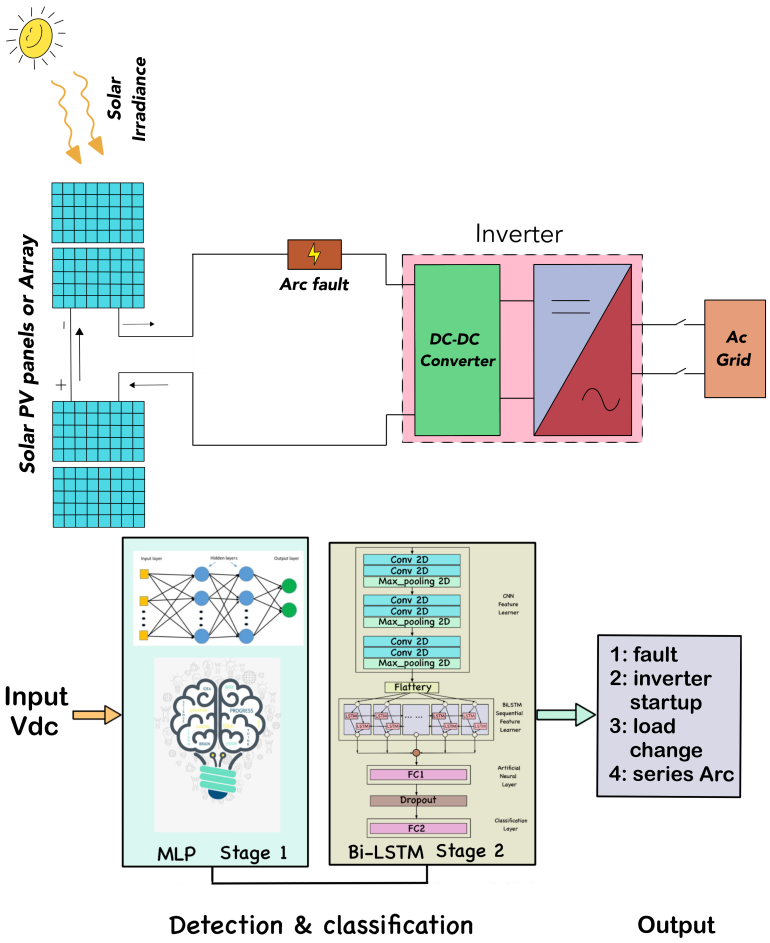


FIGURE 5. (a) Power System, and (b) Detection and Classification Method.

3.1.5. Cassie arc model

Cassie developed a formula to determine the arc's conductivity and

where u and i are voltages and currents across the gap, respectively, and g is the gap resistance. U_c and τ is the time constant, and the arc voltage. The combination of these two variables allows the simulation of a wide variety of arc situations [47].

$$\frac{1}{g}\frac{dg}{dt} = \frac{1}{\tau}(\frac{u^2}{U_c^2} - 1) \tag{15}$$

3.1.6. Mayr arc model

Mayr found new evidence suggesting that arc incidence is not the only cause of power losses in this investigation. He explained it in terms of heat transfer at extremely low levels of current. In order to demonstrate his concept, he created a mathematical equation circuit demonstrates the potential influence of cross-sectional area on heat conductivity. This impact may be seen in Eq. (16) [48, 49]. Arc conductance (g), arc current (i), arc voltage (u), arc time constant t, and arc cooling power τ are represented by the variables in this equation:

$$\frac{1}{g}\frac{dg}{dt} = \frac{1}{\tau}(\frac{ui}{P} - 1) \tag{16}$$

3.1.7. Habedank arc model

When the Cassie and Mayr forms are connected in sequence, as illustrated in Equation, the resulting model is a unique combination of the two (17-19). For more information. For more information [50].

$$\frac{dg_c}{dt} = \frac{1}{\tau_c}(\frac{u^2g^2}{U_c^2 g_c} - g_c) \tag{17}$$

$$\frac{dg_m}{dt} = \frac{1}{\tau_m}(\frac{u^2g^2}{P_0} - g_m) \tag{18}$$

$$\frac{1}{g} = \frac{1}{g_c} + \frac{1}{g_m} \tag{19}$$

3.1.8. KEMA arc model

The KEMA arc model, as demonstrated by Equation (20-25), closely matches the one given by Mayr. This model was constructed by modifying three Mayr models and connecting them in sequence. This model's predictions are more realistic since they are so near to the actual data [51].

$$\frac{dg_1}{dt} = \frac{g_1^{\lambda_1}u_1^2}{\pi_1\tau_1} - \frac{1}{\tau_1}g_1 \tag{20}$$

$$\frac{dg_2}{dt} = \frac{g_2^{\lambda_2}u_2^2}{\pi_2\tau_2} - \frac{1}{\tau_2}g_2 \tag{21}$$

$$\frac{dg_3}{dt} = \frac{g_3^{\lambda_3}u_3^2}{\pi_3\tau_3} - \frac{1}{\tau_3}g_3 \tag{22}$$

$$\frac{1}{g} = \frac{1}{g_1} + \frac{1}{g_2} + \frac{1}{g_3} \tag{23}$$

$$u = u_1 + u_2 + u_3 \tag{24}$$

$$i = gu \tag{25}$$

Table 2. CHARACTERISTICS OF PV MODULES THAT USED IN SIMULATION

Model	Parameters and Associated Values					
Ayrton	A	B	C	D		
	37	1.1	14.8	7.88		
Steinmetz	A	C	D			
	36	130	0.33			
Nottingham	A	B	n			
	27.5	44	0.67			
Hyperbolic	a	bx_{gap}	α	λ	q	V_c
	0.01	0.005	4	100	$0 < q < 1$	260
Cassie	U_c^2	τ_c				
	6400	$1.2e^{-5}$				
Mayer	P	τ				
	1	$1.2e^{-5}$				
Habedank	U_c^2	P_0	τ_c	τ_m		
	6400	1	$8e^{-5}$	$8e^{-5}$		
KEMA	u_i^2	τ				
	6400	$1.2e^{-5}$				
Schwarz	a	b	P	τ		
	-0.9	1.2	11000	$1.2e^{-5}$		

3.1.9. Schwarz Arc Model

The arc model initially proposed by Mayr underwent modifications, as depicted in Equation (26), to incorporate the dependence of cooling power and time constant on the conductance of the arc. The aforementioned alteration was devised by Schwarz, therefore leading to the nomenclature of this type being derived from his surname. For a more comprehensive understanding, readers are encouraged to consult the source cited as [51]. In summary, this model was characterized by multiple configurable characteristics, including the time constant associated with the arc. τ, constant of cooling (P), and another two parameters that have an impact on the conductance dependency, τ (a) and P (b).

$$\frac{1}{g}\frac{dg}{dt} = \frac{1}{\tau g^a}\left(\frac{u_i}{Pg^b} - 1\right) \tag{26}$$

Finally, the introduced models are used to accomplish a vast data set in order to evaluate the presented algorithm. PSCAD is used for simulating these models, and the effect of their parameters is studied and presented in different cases. Different parameters associated with their numerical values for each arc model are used to generate and gather the different cases of the arc signals as shown in Table II.

3.2 THE PROPOSED DETECTION AND CLASSIFICATION MODELS

The suggested technique has two primary components: the first is to identify any modification in the signal, and the second is to categorize the nature of the change; thus, this part is consisting mostly of two parts: classification models and change detection techniques. 1) Change Detection Model:

3.2.1. Change Detection Model

Here, we suggest a change detection approach for spotting unexpected behavior in the output signal, and we show how a Multilayer Perceptron (MLP) may be utilized to construct the network required for this purpose. In order to train the network, it is fed a massive dataset of signals as the dataset to feed the proposed MLP network during the process of network training. These signals are classified into five categories, which are fault, signal integrity during inverter start-up, load changes, a series arc, and at normal operating levels. It's easily compatible with existing classification methods, as shown in Figure 6. In addition, the change detection technique may be applied at each branch of the circuit to identify signal distortion based on the master/slave operating concept. The incoming signal will be analyzed, but the final choice will be dependent on two factors. The first one, if there is any noise or distortion in the signal, then the decision will be a trigger will be sent to the master detector, which is the Bi-LSTM model to analyse and classify this signal according to which category it belongs. In the second scenario, If the signal is clean and uncorrupted, the operation will continue without interruption. Figure 7 illustrates the procedure workflow for the enhanced method of change detection. Using a change detection approach will make the detection system more efficient and trustworthy. because instead of classifying

Table 3. MLP ARCHITECTURE.

No.	Layer Type	No. of Kernel	Dropout	Output	
1	Input_layer	-	-	-	200 × 100×3
2	Denese 1	tanh	256	-	200 × 100 × 256
3	MLP1	tanh	64	-	200 × 100 × 64
4	MLP2	tanh	32	-	200 × 100 × 32
5	MLP3	tanh	16	-	200 × 100 × 16
6	Dropout	-	-	0.5	200 × 100 × 16
7	MLP2	tanh	8	-	200 × 100 × 8
8	Denese2	SoftMax	2	-	No. of Classes

Table 4. BI-LSTM TRAINING PROCESS[52]

Sequential Forward	Sequential Backward
$i_t = \sigma(W_{ix}x_t + W_{ih}h_(t-1) + b_i)$	$_t = \sigma(W_{ix}x_t + W_{ih}h_(t+1) + b_i)$
$f_t = \sigma(W_{fx}x_t + W_{fh}h_(t-1) + b_f)$	$f_t = \sigma(W_{fx}x_t + W_{fh}h_(t+1) + b_f)$
$o_t = \sigma(W_{ox}x_t + W_{oh}h_(t-1) + b_o)$	$o_t = \sigma(W_{ox}x_t + W_{oh}h_(t+1) + b_o)$
$c_t = f_t c_{t-1} + i_t * tanh(W_{cx}x_t + W_{ch}h_(t-1) + b_c)$	$c_t = f_t c_{t+1} + i_t * tanh(W_{cx}x_t + W_{ch}h_(t+1) + b_c)$
$h_t = o_t * tanh(c_t)$	$h_t = o_t * tanh(c_t)$

all the coming signal through the Bi- LSTM classification model, only the signal with distortion or noise will be classified to specify exactly the type of noise using the classification model. The architecture and the configuration of the proposed detection method (MLP) is shown in Table III.

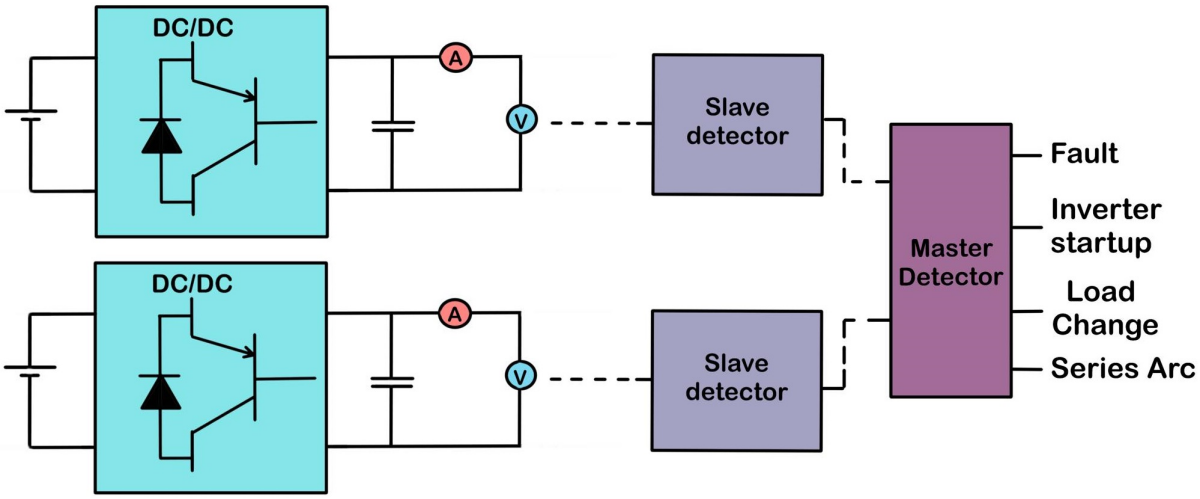


FIGURE 6. The Principles of Proposed Change Detection Method.

3.2.2. Classification Model

Bi-LSTM method make the pro- cess of training into a sequential form; in other words, the process will be forward and backward instead of forward-only, as shown in Table IV and Figure 8.

The suggested classification model employs the Bi-LSTM technique because it can isolate the series arc fault from any other disturbances present in the signal. Nine arc models are developed and utilized to create the signal of the series arc, as well as additional problems that can occur in the PV system, including the normal activity of the inverter starting up and the load changing, as described in section (III.A). The proposed classification model trained on these types of signals to precisely determine the series arc upon occurrence. Furthermore, the strength of this model is that it uses CNN convolution layers to extract as much information as possible from the data. Then, this output will become the input to the Bi-LSTM; this feature will allow the data to be kept in a chronological order in the both directions Table V and Figure 8 illustrate the structure and layer details of the proposed classification model. Because of each model has its own architecture and strengths, combining CNN and Bi-LSTM models necessitates a unique design: • CNN is well-known for extracting as

Table 5. CNN-Bi-LSTM ARCHITECTURE

No.	Layer Type	Activation	No. of Kernel	Dropout	Output (Activations)
1	Input_layer	-	-		200 × 100×3
2	Time_distributed_1(Conv. 1)	ReLU	32	-	200 × 100×32
3	Time_distributed_2(Conv. 2)	ReLU	64	-	200 × 100×64
4	Max_Pooling	-	-	-	100 × 50×64
5	Time_distributed_3(Conv. 3)	ReLU	128	-	100 × 50×128
6	Time_distributed_4(Conv. 4)	ReLU	128	-	100 × 50 ×128
7	Max_Pooling	-	-	-	50 × 25 × 128
8	Time_distributed_5(Conv. 5)	ReLU	256	-	50 × 25 ×256
9	Time_distributed_6(Conv. 6)	ReLU	256	-	50 × 25 ×256
10	Max_Pooling	-	-	-	25 × 12 ×256
11	Time_distributed_10(Flatten)	-	1	-	1 × 1 × 256
12	Bidirectional_LSTM1	ReLU	512	-	1 × 1 × 512
13	Bdirectional_LSTM2	ReLU	512	-	1 × 1 × 512
14	Denese1	ReLU	128	0.5	1 × 1 × 128
15	Denese2	Softmax	3	-	No. of Classes

many features and attributes from the documents.

• Because of Bi-LSTM maintains the chronological order which it can ignore the unnecessary words using the delete gate. The main objective of integrating the two models is to create a model that uses the strengths of CNN and Bi-LSTM in order to capture the features extracted using CNN and use them as a Bi-LSTM input. As a consequence, a model is created that achieves this goal, in which the vectors created and used as convolutional neural network input. Then, different size of filters is applied; the output of all max pooling layers is then concatenated to form the Bi-LSTM input, which employs a Bi-LSTM layer to filter the data using its three gates. This step's output is used to feed the fully connected layer, which connects each part of input information with a part of output information.Finally, in order to achieve the desired output, the SoftMax function is used as an activation function to assign the classes to the type of fault. As a result, the architecture of the proposed model, which is divided into three parts as follows:

• Pre-processing Step: At this stage, data is normalized and pre-processing is performed. It is then used to prepare data for convolutional layer. The resultant vector is used as an input to to the next stage.

• Convolution Step: Convolution and max pooling layers are used for feature extraction at this stage to obtain high-level features. This stage output is the next stages input.

• Bi-LSTM/ Fully-connected Step: Bi-LSTM and fully connected layers are used for the process of feature selection at this stage. The final classification of the data is the result of this stage (as fault, inverter startup, load change and series arc).

Adam optimizer was used in the model to obtain useful and accurate results. Moreover, ReLU and Softmax, are selected and used as an activation function in the hidden and output layers, respectively. Combining the CNN Model with Bi-LSTM will reduce the spatial dimension very fast and reduces the computational cost very much by adopting stride=2 in the first conv. layer and then max pool. Therefore, the spatial feature dimension becomes 1/4 of the input in the third layer. This will make the model more fast as well as the combination will increase the accuracy of the proposed model as it will integrate the strength features of the both models (CNN Model with Bi-LSTM).

4. RESULTS AND EVALUATION

In the first of this section's three main sections, we introduce the system that was simulated by showing examples of the data that was simulated and the various sorts of records that were generated, and the simulation results of the classified model are clarified in the second part. And the third part present the comparison with state-art algorithms.

4.1 SIMULATED SYSTEM OF DC SERIES ARC FAULT

In this part, various arc models are simulated via PSCAD to acquire different dataset records; Figure 9 shows a block diagram of the PSCAD simulation system and the model circuit presented in Figure 10. Moreover, Table VI shows the characteristics of the Simulation PV System.

The generation of arc fault models is simulated using PSCAD software. Nine different arc models, as previously mentioned in section (III.A), are used to generate the necessary records that take into consideration many different cases through changing various factors such as the length of the arc, irradiance, the frequency of inverter switching, the load

Table 6. CHARACTERISTICS OF SIMULATION AND EXPERIMENTAL PV SYSTEM

Model	Simulation System	Experimental
DC terminal voltage (v)	260	100
Number of Modules	18	5
Number of Strings	10	2
Ambient temperature (°C)	25	25
Irradiation (W/m2)	1000	1000
Load Current (A)	1, 5, or 10	1
Sampling Frequency (kHz)	100	100
Arc Length (mm)	1, 1.5, 2, 2.5	1, 1.5, 2, 2.5

current, and temperature. These criteria were adopted to obtain better results and to produce a large number of records that covers all the predicted fault cases that can potentially be faced by the PV system. Moreover, the produced records can be classified into:
I. A set of files specific to the inverter power-on procedure
II. Changes to the load configuration are represented by a collection of records.
III. Inferring a sequence of arc fault from a set of records.
IV. A collection of data pointing to the second arc fault.
About 800 recordings totaling 6000 data points were created. Following normalization of samples according to a predetermined window length, training can begin on the newly created records data using the proposed approach to begin developing a detection strategy.

4.1.1. Dataset of Inverter Start-up and Load setup change

This dataset is a subset of the produced dataset chosen to illustrate two key scenarios that constitute typical PV system operation. The inverter's first power-up is the first. Changing the load is the second typical action of a PV system. The PV system's output voltage changes because of events like the inverter starting up and/or a change in the load. Several examples of inverter startup and load change are displayed in Figure 11.

4.1.2. Dataset of the other Fault

: These logs pertain to additional PV system problems, including the short circuit fault, which might occur in the system. These logs are created to improve the detecting method's efficiency via the capacity to tell out these faults from the series arc, as seen in Figure 12.

4.1.3. Dataset of Series Arc Fault

This dataset represents the series arc fault, as shown in Figure 13. It was generated from the used circuit of the arc fault model. However, in reality, several reasons lead to the occurrence of this type of faults, such as incorrect installation, irregular maintenance, and some environmental effects. The proposed method was trained to precisely detect it Using a variety of arc models with the different specifications for each model, as discussed above and mentioned in Table 8, have two main benefits that are generating a massive dataset of different types of records, which results in increasing the accuracy of the proposed model. Also, training the proposed models on different types of arc fault signal leads to an increase in the precision of the model and to be more efficient.

4.2 SIMULATION RESULTS

In this section, the results of the two levels of the proposed method, which are change and classification level of the LSTM models are presented and discussed. For the first level, MLP is used to detect any abnormal behaviour of the input signal. Figure 14 shows the proposed network associated with training results. For the second level, Python is used to simulate the suggested schema for each model where the $Keras_{T}ensorflow_{G}oogle$ was used. Moreover, the results for each model are shown in Figures 15 and 16 in terms of model accuracy and model loss, respectively.

For evaluating the performance of the suggested classification model using Bi-LSTM, a set of critical metrics have been utilizing as follows:

1. Accuracy metric: It is the right classification ratio in which the proportion of cases that have been rightly classified (99%) as shown in Figure 15.

2. Loss metric: it is the opposite of accuracy metric, refer to the proportion of cases that have been not rightly classified

Table 7. A SUMMARY OF ARC FAULT MODELS [37]

Arc model	Current level	The gap width between the two electrodes (mm)	Model Type	High-frequency variation
Ayrton	Low	1–10 mm	V-I characteristic-based arc model	Gaussian/Pink noise
Cassie	High	Depends	Physics-based arc model	Gaussian/Pink noise
Habedank	Low/High	Depends	Physics-based arc model	Additional Equation
Hyperbolic	Low/High	Depends	Heuristic arc	Additional Equation
Kema	High	Depends	Physics-based arc model	Gaussian/Pink noise
Nottingham	Low (below 10 A)	1–10 mm	V-I characteristic-based arc model	Gaussian/Pink noise
Schwarz	Low	Depends	Physics-based arc model	Gaussian/Pink noise
Steinmetz	Low	Depends	V-I characteristic-based arc model	Gaussian/Pink noise
Mayr	Low	Depends	Physics-based arc model	Gaussian/Pink noise

(0.01%), as shown in Figure 16.

3. Precision metric: It is the proportion of cases rightly predicted true noticing to the full predicted true noticing; the value of this metric is 100% for Suggested Schema which can be calculated in equation (27).

$$Pr = T/(T + F) \times 100\% \tag{27}$$

4. Recall metric: It is the proportion of cases rightly predicted true noticing to the full noticing in actual class; the value of this metric is 98% for Suggested Schema. which can be calculated in equation (28).

$$Re = TR/(TR + FR) \times 100\% \tag{28}$$

5. F1-score metric: It is the regular rate of Recall and Precision; the value of this metric is 98% for Suggested Schema. which can be calculated in equation (29).

$$F1 - Score = 2 * ((precision * recall)/(precision + recall)) \tag{29}$$

Furthermore, obtaining the arc fault signal through the use of a real photovoltaic system is expensive and sophisticated due to the variation in the fault location and the levels of current. Therefore, to develop and evaluate the proposed arc fault classification method, nine different models of arc fault are simulated and used. The classification accuracy for each model, as well as the overall, was 99%. However, these models are used because each one of them has its own specific criteria. In general, these models are classified into three main groups standard V-I empirical models obtained from measurement data, physical principles-based model, and heuristic models. In more detail, the Ayrton model covers the only low level of the current, which belongs to the group of the V-I characteristic. In contrast, the cassie model belongs to the physical-based group that works with a high level of the current and depending on the gap between the two electrodes. Also, the two other models, which are the Habedank and Hyperbolic, cover both levels of the current, low, and high. However, similar to the cassie model that is depending on the gap between the two electrodes; although, they belong to a different group of arc model-based. Habdenk is a physics-based arc model, while hyperbolic is a heuristic arc model. The other models, Schwarz, Steinmetz, and Mayr, have covered the low level of the current and similar to the previously mentioned models that depend on the electrodes gap. Schwarz and Mayr belong to the family of physics-based, while Steinmetz is V-I characteristic based. A Summary of these Arc Fault Models can be seen in the Table VII [37].

4.3 COMPARISON WITH OTHER METHODS

According to recent publications, the features and benchmarks of Artificial Intelligence-based detection methods are as follows:

1. Accuracy
2. Fault Classification
3. Over-Fitting
4. Vanishing

5. Computational/ Simplicity

The aforementioned factors must to be taken into account when formulating prospective models. The forthcoming section will provide a comprehensive analysis of these constraints in order to elucidate how the suggested methodology would address the limitations encountered in prior approaches. 1. The initial and foremost criterion to be addressed in the design of the suggested approach is accuracy. The proposed methodology must exhibit a high level of precision in order to effectively detect DC series arc faults. This requirement is fulfilled by the implementation of the four models that have been put out. Various types and quantities of layers are employed in the construction of effective models. The function of fault classification criteria is effectively demonstrated by the utilization of a robust approach known as Convolutional Neural Network (CNN). The utilization of this strategy in the second stage of the proposed approach can yield a robust classification function. The suggested models address the issue of overfitting by incorporating a dropout layer, which effectively mitigates the occurrence of intricate co-adaptations on the training data. As additional layers with specific activation functions are incorporated into neural networks, the gradients of the loss function tend to diminish, hence rendering the network challenging to train. Hence, the suggested models incorporate batch normalization layers to address the challenge of disappearing gradients.Convolutional neural networks (CNNs) are employed for the purpose of mitigating complexity. The reduction in computational complexity can be attributed to the parameter sharing capability of Convolutional Neural Networks (CNNs). Furthermore, these networks possess the ability to autonomously extract valuable features from unprocessed material. In order to effectively utilize such networks, it is necessary to capture sensor-acquired data in the form of sequence data. The suggested methodology was assessed by conducting a comparative analysis with previously established methods that utilized artificial intelligence techniques. The evaluation focused on the overall accuracy of the method, as well as its capacity to classify normal situations and accurately identify series arc faults from other fault cases. In addition, it is important to examine the issues of over-fitting and Vanishing difficulties, since these might result in a decline in the system's performance. Moreover, the role of computing complexity is crucial in probabilistic model-based approaches, such as Hidden Markov Models (HMMs), necessitating additional work [37]. Table VIII presents a summary of the comparison between the discussed approach and other existing methods, based on the specified criteria. The utilization of artificial neural networks (ANN) is observed to be susceptible to the issue of vanishing difficulties. In addition, it is crucial to exercise caution while selecting the learning rate in order to mitigate the issue of overfitting. Furthermore, it is worth noting that a moderate level of computational complexity and a high level of effort were associated with a rather low accuracy rate of 73%. Various Artificial Neural Networks (ANNs) exhibit different performance outcomes due to the random initialization of weight parameters in each network. Consequently, identifying the most optimal performing ANN becomes questionable, while utilizing the identical training data. Bi-directional Long Short-Term Memory (Bi-LSTM) models address the problem by employing a distinctive additive gradient structure. This structure incorporates direct access to the activations of the forget gate, allowing the network to promote desired behavior through frequent updates of the gates at each time step during the learning process. Hence, the Bi-LSTM method can be regarded as a dependable approach. In contrast, the support vector machine (SVM) method demonstrates a notable accuracy rate of 99%. Nonetheless, it is important to note that the issue of vanishing problem persists. Additionally, the software lacks support for the categorization feature, and it exhibits a medium level of computational complexity that requires a significant amount of effort to obtain results. In addition, the Fuzzy based method has a high level of accuracy, namely 95.8%. Moreover, it exhibits superior capability in classifying series arc faults compared to other fault signals, while maintaining a low computational complexity and moderate effort. However, it fails to address the issues of overfitting and disappearance. The HMM model, another approach employed in this study, shares similarities with the fuzzy-based method in terms of its inability to address overfitting and vanishing issues. However, it should be noted that the HMM model outperforms the fuzzy model in terms of accuracy, achieving a notable 98.3%. The Convolutional Neural Network-Generative Adversarial Network (CNN-GAN) The utilization of a deep learning model for generating multiple instances of the arc fault signal, as opposed to relying on real-world data or a specific arc fault model, may introduce inaccuracies in the detection process. The proposed method aims to address this issue by focusing on distinguishing between these artificially generated records, rather than detecting the actual occurrence of a series arc fault. Consequently, any errors present in these generated records can lead to incorrect detections. Despite achieving a commendable model accuracy of 98.5%, there are still notable limitations in terms of addressing overfitting and vanishing issues, as well as the absence of support for classification functionality. The subsequent CNN-based model exhibits a notable accuracy of 98.9% and possesses the capability to differentiate the series arc across various fault signals. Moreover, this method can be regarded as a dependable solution that effectively addresses the issue of overfitting by employing appropriate Dropout settings. Nevertheless, it continues to grapple with the issue of disappearance. The suggested method demonstrated exceptional performance, achieving a remarkable accuracy rate of 99%. It effectively classified series arc faults and was specifically taught to differentiate between routine scenarios such as inverter startup and load changes, as well as other types of faults. Furthermore, the approach employed in this study was deemed dependable and effectively addressed the issue of

Table 8. A COMPARISON WITH PRIOR METHODS

System	Methodology	Accuracy	Fault Classification	Over-fitting Solution	Vanishing Solution	Computational Complexity /Effort
James et al. [32]	ANN	73%	×	×	×	Medium high /
Zhan et al. [33]	SVM	99%	×	✓	×	Medium high /
Benjamin et al. [34]	Fuzzy	95.8 %	✓	×	×	Low Medium /
Rory et al. [17]	HMM	98.3%	✓	×	×	High
Shibo et al.[35]	CNN-GAN	98.5	×	×	×	Very High
Alaa et al.[36]	CNN	98.9%	✓	✓	×	Low
Proposed Model	MLP & Bi-LSTM	99%	✓	✓	✓	Low

over-fitting by implementing suitable Dropout values. Additionally, the vanishing problem was successfully mitigated through the utilization of the Bi-LSTM model. Furthermore, the method under consideration was simulated with reduced computational effort and without the requirement of integrating it with an auxiliary methodology.

5. OFFLINE EEXPERIMENTAL RESULTS

Assessing the efficacy of the model is crucial in conjunction with simulation. An evaluation was conducted using a small-scale test bench system. Two distinct categories of solar panels were employed. The photovoltaic (PV) system incorporates a solar tracker known as the Sun-Tracking highest Power Point Tracking (SR-MT) system. This tracking mechanism is responsible for identifying and directing the points of highest power generated by the PV system to the load. The electricity is then sent to the load through the use of an inverter, specifically the Cotek No. SK350-224 inverter. The electric arc was intentionally created to be connected in series with the photovoltaic (PV) terminal by the utilization of a custom-made arc generator. The movement of the two electrodes of the arc generator along the x-axis was regulated by a motor equipped with a stopper mechanism. The data collection process involved the utilization of the GPS-1204C oscilloscope, which is a digital storage oscilloscope from the series. The device operated within a bandwidth of 200 MHz. A voltage signal was obtained using a sampling frequency of 10 kHz for a length of 4.5 seconds. The data collected from the experiment were converted into comma-separated values and subsequently analyzed using the proposed model. Furthermore, the measurement of sun irradiation was conducted using the MIC-98206 pyranometer. Various scenarios were employed in the practical implementation to simulate the effects of different environmental variables, such as load current and the distance between the electrodes. The trials were conducted at varying degrees of radiation intensity and ambient conditions. The outcomes of the suggested strategy are depicted in Figure 17, with the practical data that was obtained. The findings validate the efficacy of the suggested method in detecting series arc faults, as well as its applicability in detecting other faults under normal settings.

5.1 CASE STUDY: AN OFFLINE VALIDATION RESULTS

To demonstrate and evaluate the performance of the proposed MLP-BiLSTM methodology, Table IX presents a brief comparison among several prior experimental methods and the proposed method such as follows:

1. A Deep neural network trained on source-domain and target-domain normal data for domain Adaptation in Fault Diagnosis (DAFD). As suggested in [53], SVM is used as the classifier. The system was tested experimentally and an accuracy of 89.79% was achieved which indicating that the percentage of loss is approximately 10.21%; while the precision (sensibility) of the system is 99.04%. The system Recall (safety) is 83.48%, leading to a proportion of 90.5% as the regular rate of recall and precision (F1-score) according to equation (28).

2. Proposed DA-DCGAN punishment term without MMD trained on source-domain data, normal data, and dummy arcing data from the target domain[35]. The result of the system was verified experimentally presenting an accuracy of 95.76%. The PV emulator is replaced for target-domain data collection and real-time testing by a rooftop PV string made up of four JINKO JKM350M-72 monocrystalline PV panels. This model's precision, is 99.35%. Moreover, the achieved percentage of recall for this model is 92.65%. Therefore, the regular rate of recall and precision for this model is 95.8%.

Table 9. A comparison with other experimental systems

System	Methodology	Accuracy	Precision (Sensibility)	Recall (Safety)	F1-Score
1	DNN with SVM	89.79%	99.04%	83.48%	90.5%
2	CNN-GAN	95.76%	99.35%	92.65%	95.8%
3	ANN	93.4%	98.49%	76.94%	86.4%
4	LSSVM	97.5%	99.22%	93.31%	96.1%
5	ELM	97.5%	99.54%	94.69%	97%
Proposed system	MLP & BiLSTM	99%	100%	98%	98%

3. To develop the fault classifier, the extracted features are trained using an Artificial Neural Network classifier [54]. The experimental analysis is carried out on a capacity grid-connected roof top solar PV system with 169.92kW (320W X 531 modules arranges in 27 strings) at Faculty of Engineering and Technology, Jamia Millia Islamia, New Delhi, India. Texture feature analysis is used to investigate the characteristics of various faulty panel thermal images, the developed algorithm depicted an accuracy of 93.4% where the losses of the system is 6.06%. The precision of this model is 98.49% by a recall of 76.94%. The proportion of the F1-score metric, which is the regular rate of recall and precision, is 98% for this model.

4. To diagnose photovoltaic array fault, a fault detection method based on the Least Squares Support Vector Machine (LSSVM) in the Bayesian framework is proposed [55]. Moreover, a 5×3 photovoltaic array and a reference photovoltaic string are established and experimentally tested in order to validate the performance of the proposed method where an accuracy of 97.5% and losses of 2.5% achieved. Where the precision (sensibility) of the system is 99.22%. The system Recall (safety) is 93.31%, leading to a proportion of 96.1% as the regular rate of recall and precision (F1-score) according to equation (28).

5. In this study [56], an online monitoring system for PV arrays based on IoT is designed. The system consists of a data gateway, data acquisition and a website for the PV monitoring center (PVMC). The results of the experimental demonstrate that the proposed monitoring system can effectively monitor the PV array in actual environments, and the fault diagnosis method depicts an accuracy of 97.5% and a losses of 2.5%. This model's precision, is 99.54%; moreover, the achieved percentage of recall for this model is 94.69%. Therefore, the regular rate of recall and precision for this model is 97%.

6. Proposed (MLP-BiLSTM) A detection and classification method using a multilayer perceptron incorporated with Bi-Directional Long short-term Memory (MLP-BiLSTM) is proposed. Two different types of solar panels were used. The PV system involves a SR-MT for tracking the points of maximum power and sending them to the load by the invertor (Cotek No. SK350-224). The experimental results confirm that the accuracy of the proposed detection and classification method reaches 99% with losses of 1%. The precision of this model is 100% by a recall of 98%. The proportion of the F1-score metric, which is the regular rate of recall and precision, is 98% for this model. The results of the proposed method are believed to be accurate for DC series arc fault detection and classification in the PV system with relatively high accuracy.

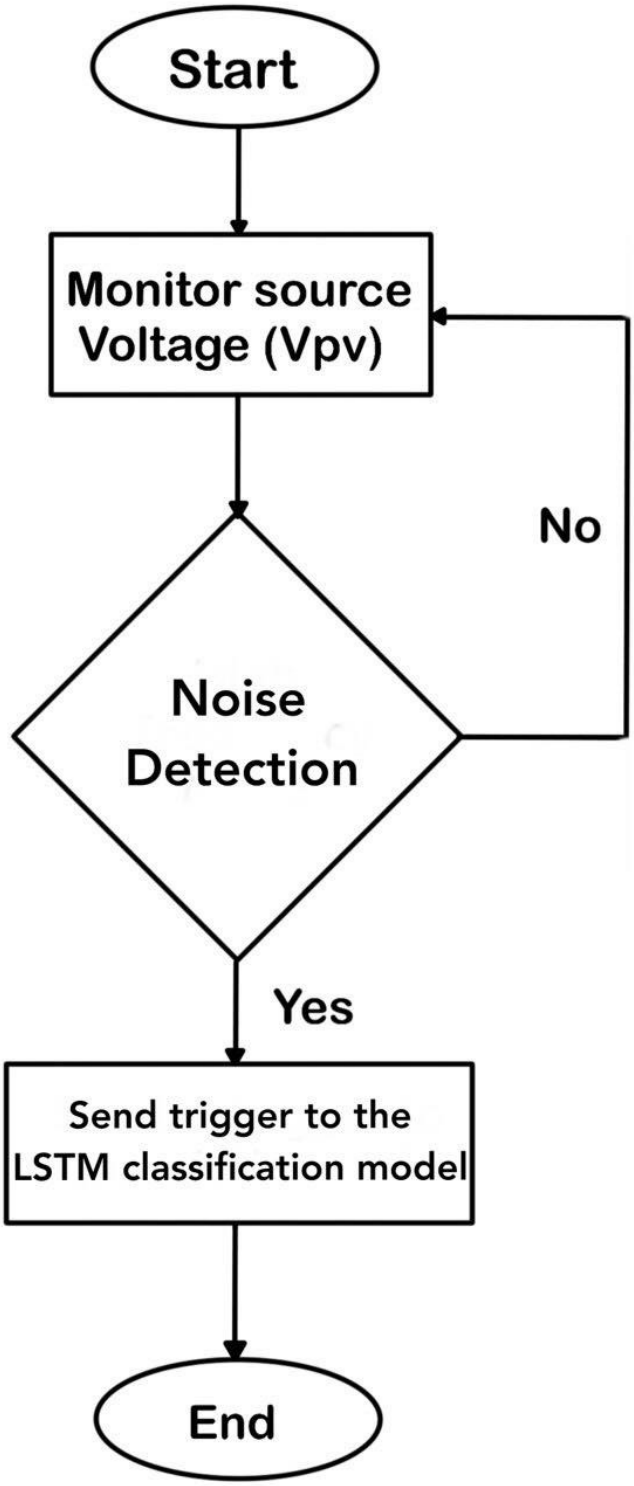


FIGURE 7. Flowchart of the change detection method operation.

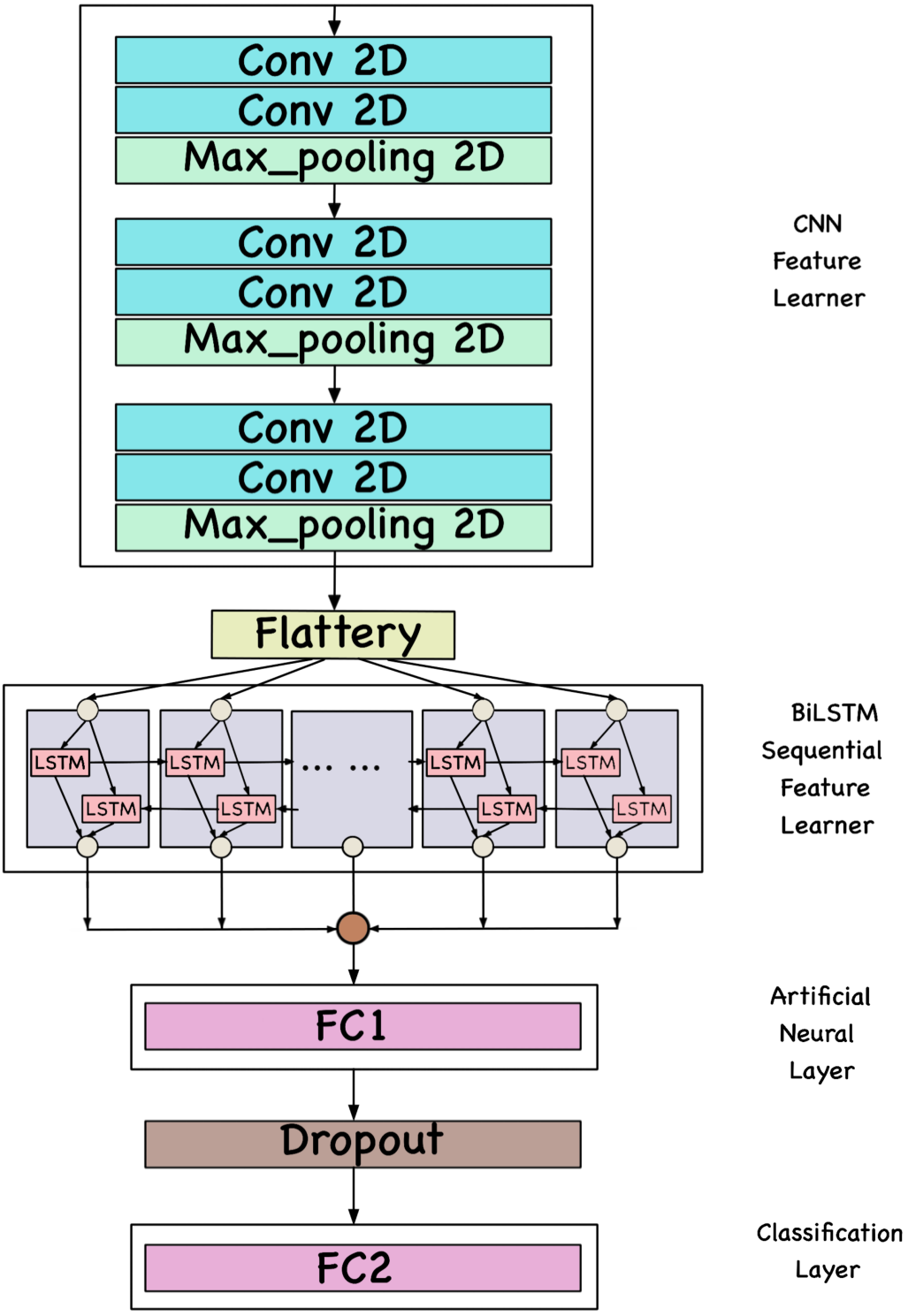


FIGURE 8. **Proposed Classification Model Structure.**

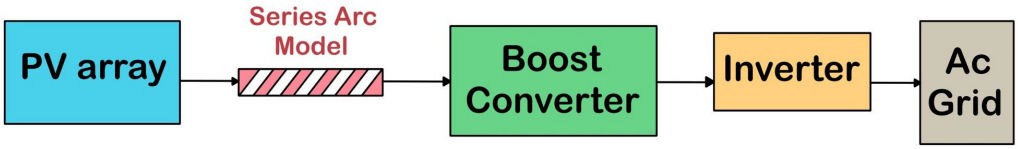


FIGURE 9. **Simulation System Block Diagram.**

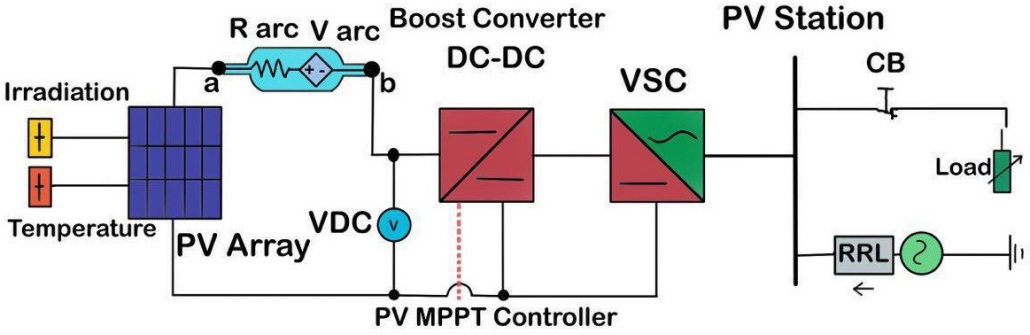


FIGURE 10. **PSCAD DC Series Arc Fault Model Circuit.**

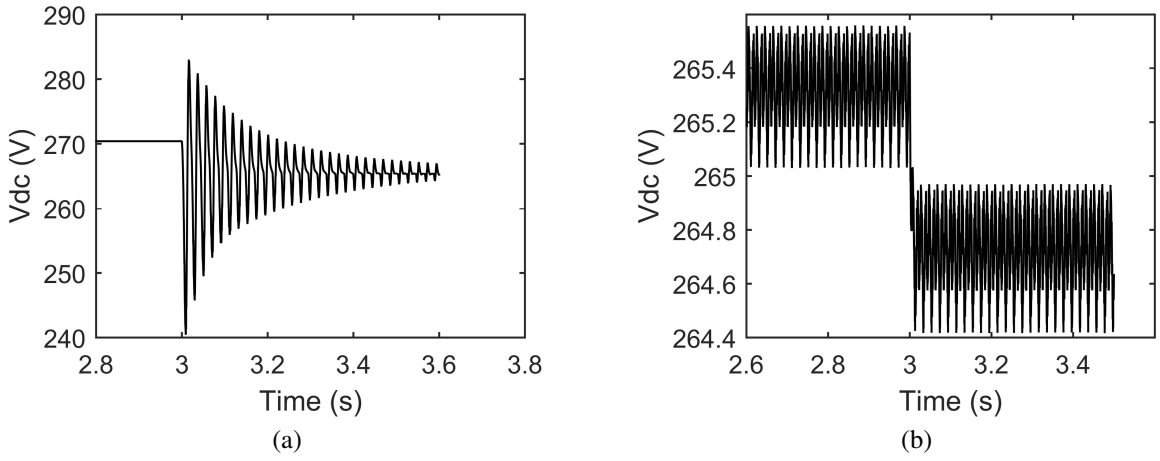


FIGURE 11. **Samples of the Normal Case: (a) Sample of Inverter Startup; (b) Sample of Load Change**

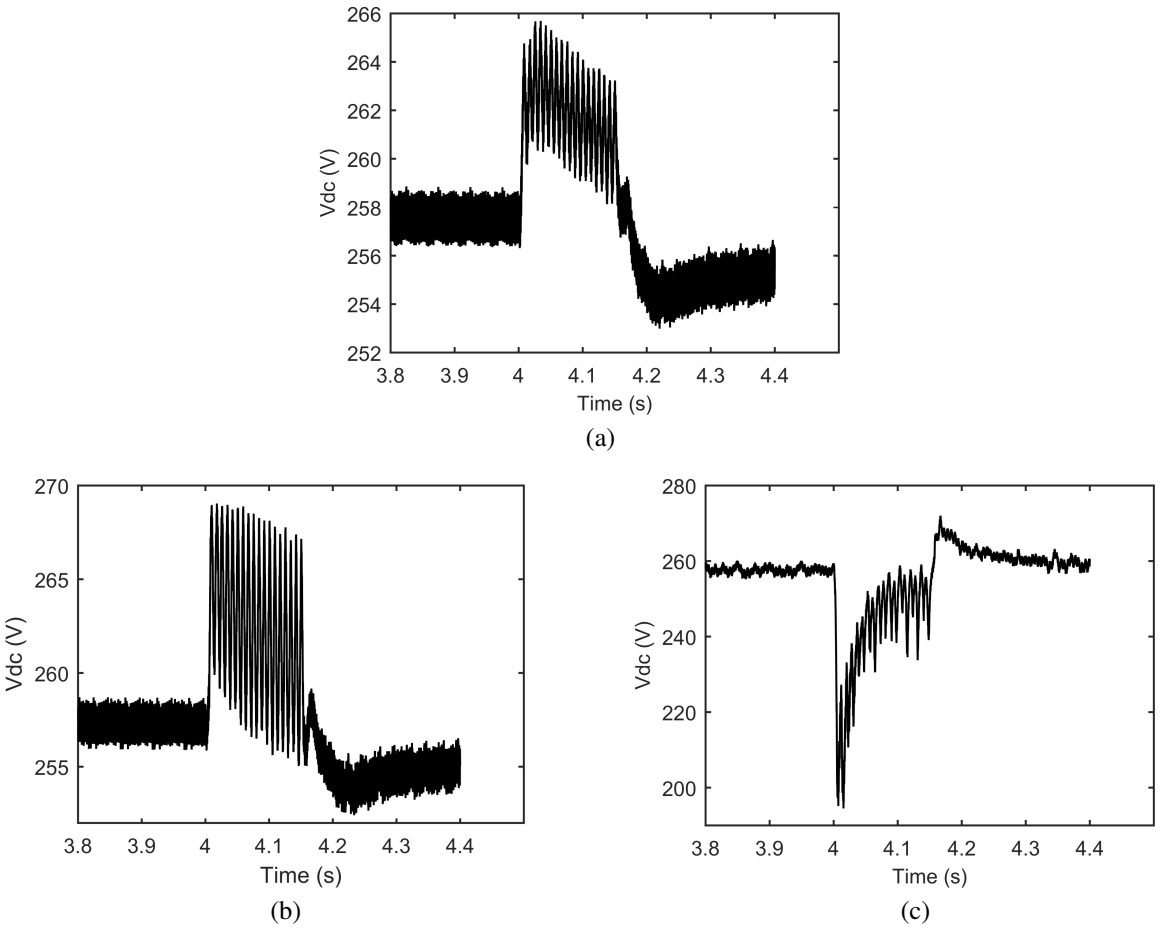


FIGURE 12. Different Samples of faults

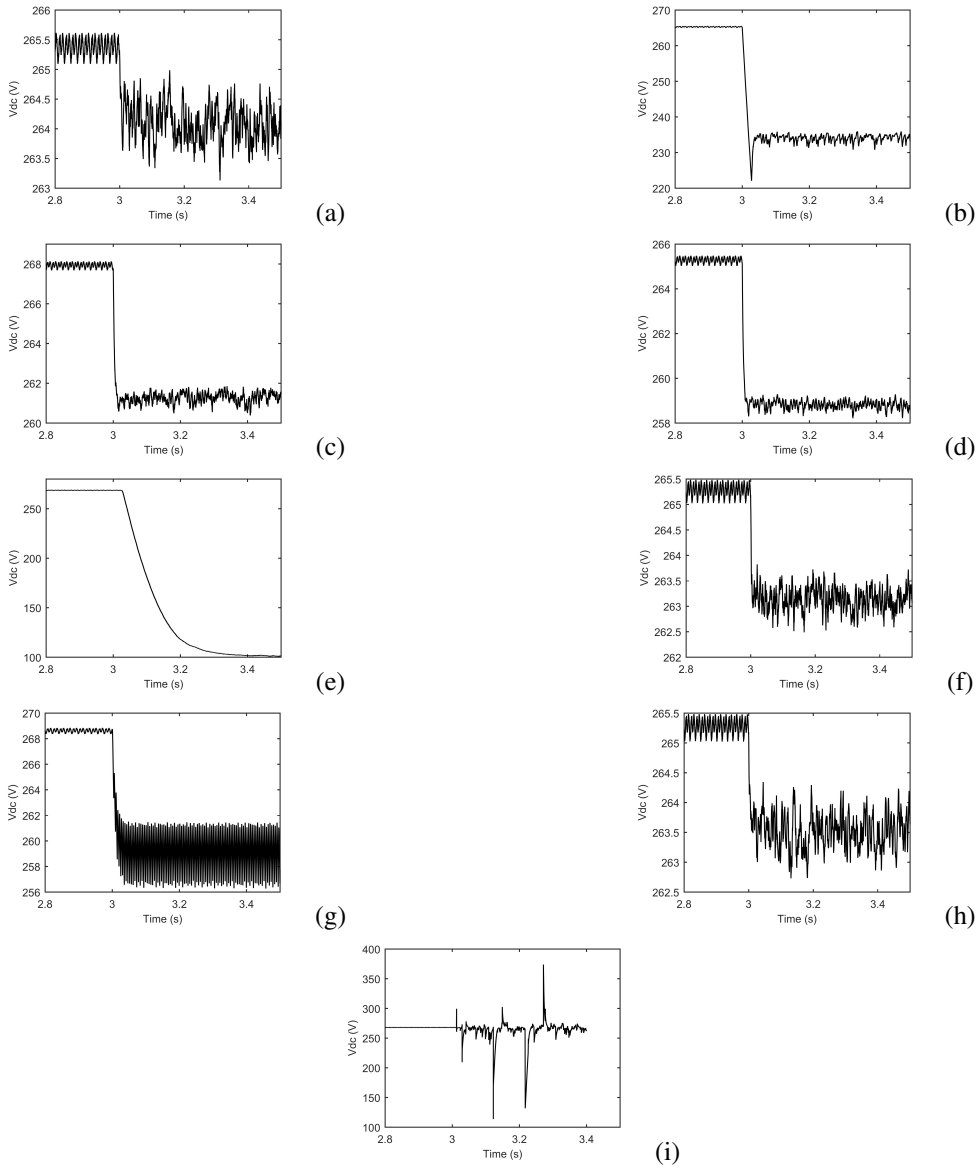
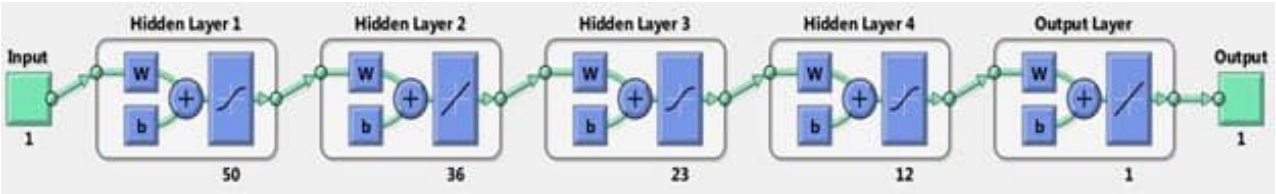
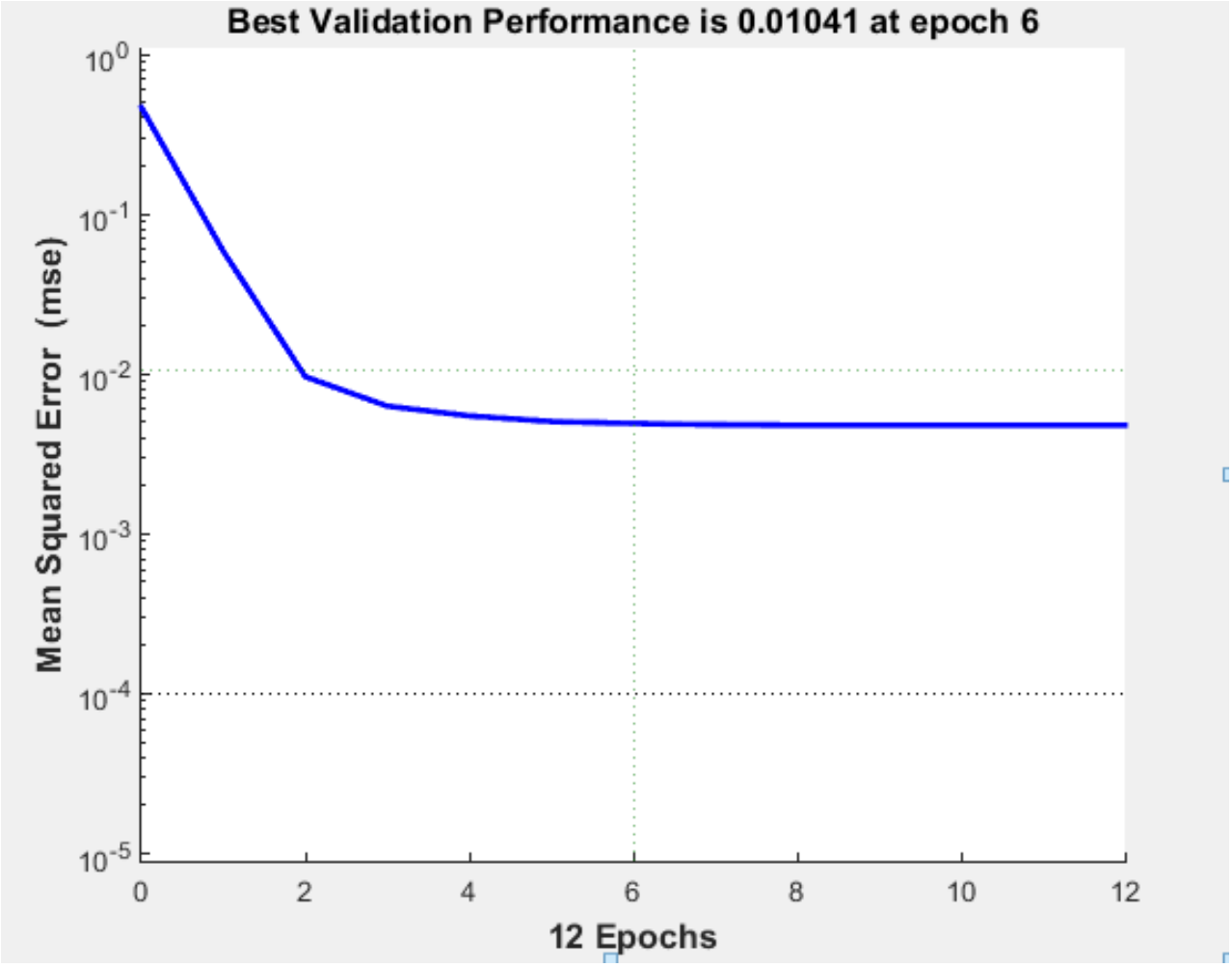


FIGURE 13. Series arc fault waveforms of different Models (a) Ayrton (b)Cassie (c) Habedank (d)Hyperbolic (e) Kema (f) Nottingham (g) Schwarz (h) Steinmetz (i) Myer.



(a)



(b)

FIGURE 14. Change Detection Level (a) the Proposed Network (b) Results of the Training.

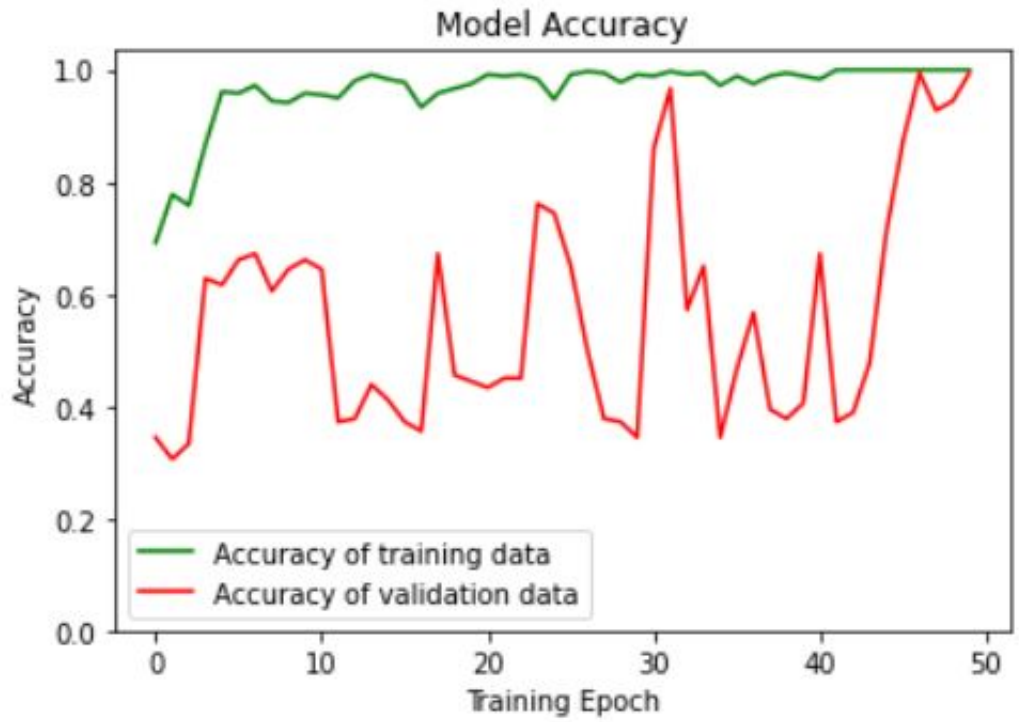


FIGURE 15. **Model Accuracy for Suggested Schema.**

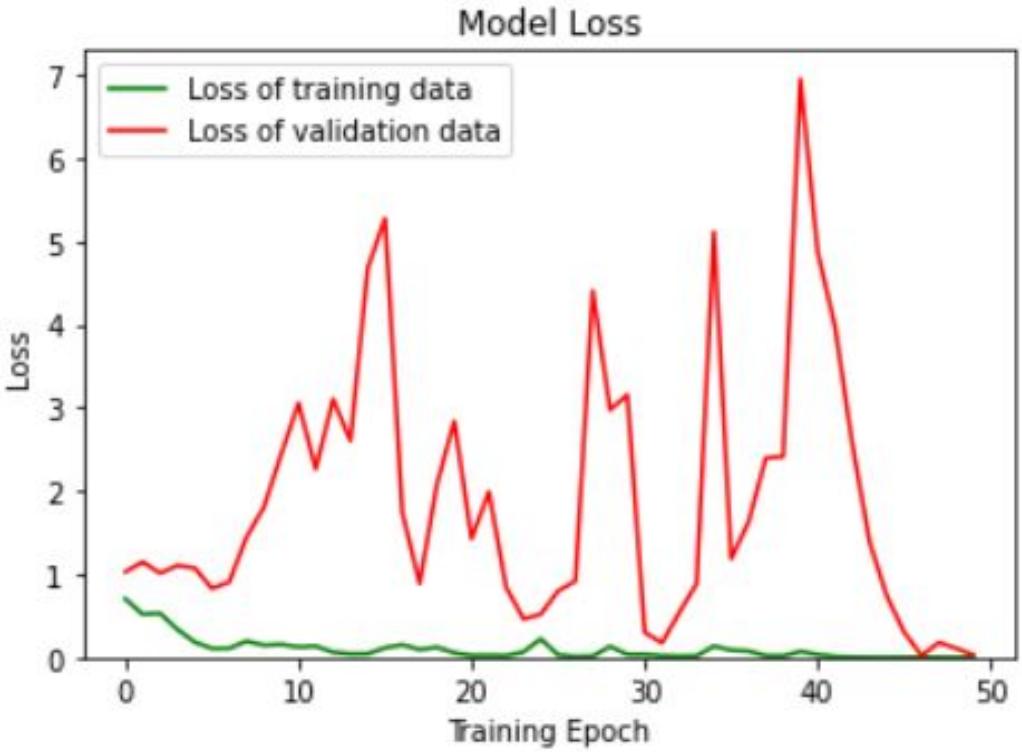
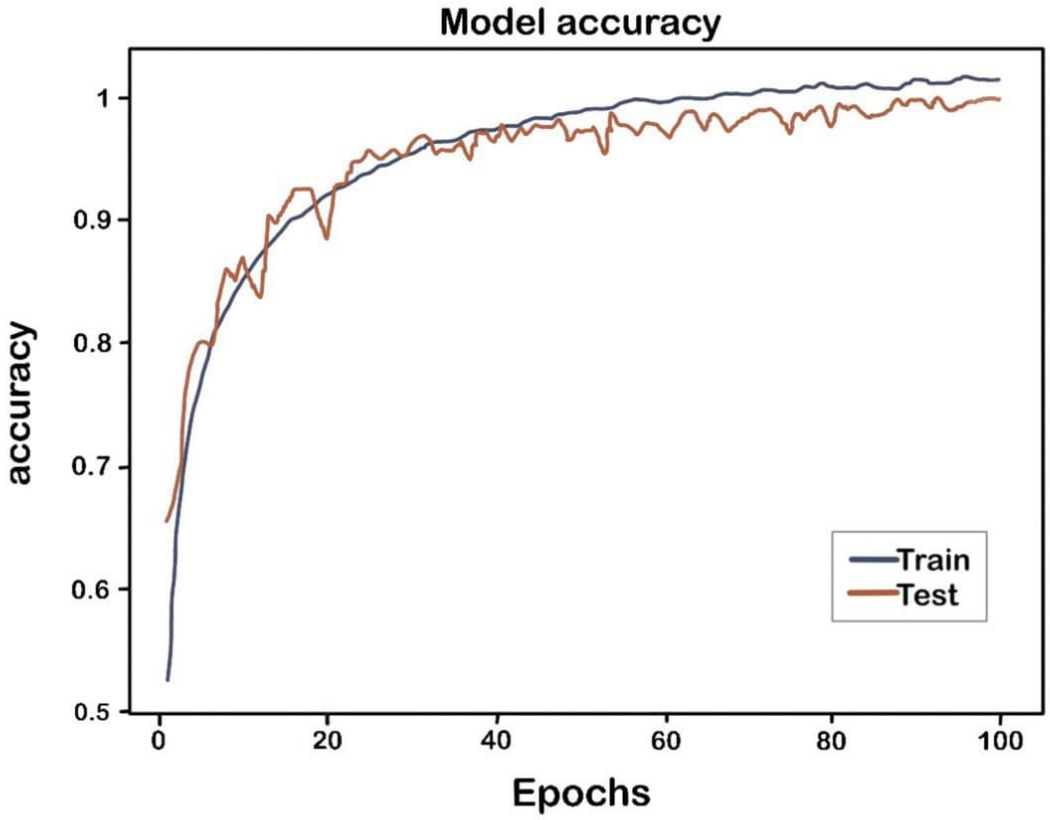
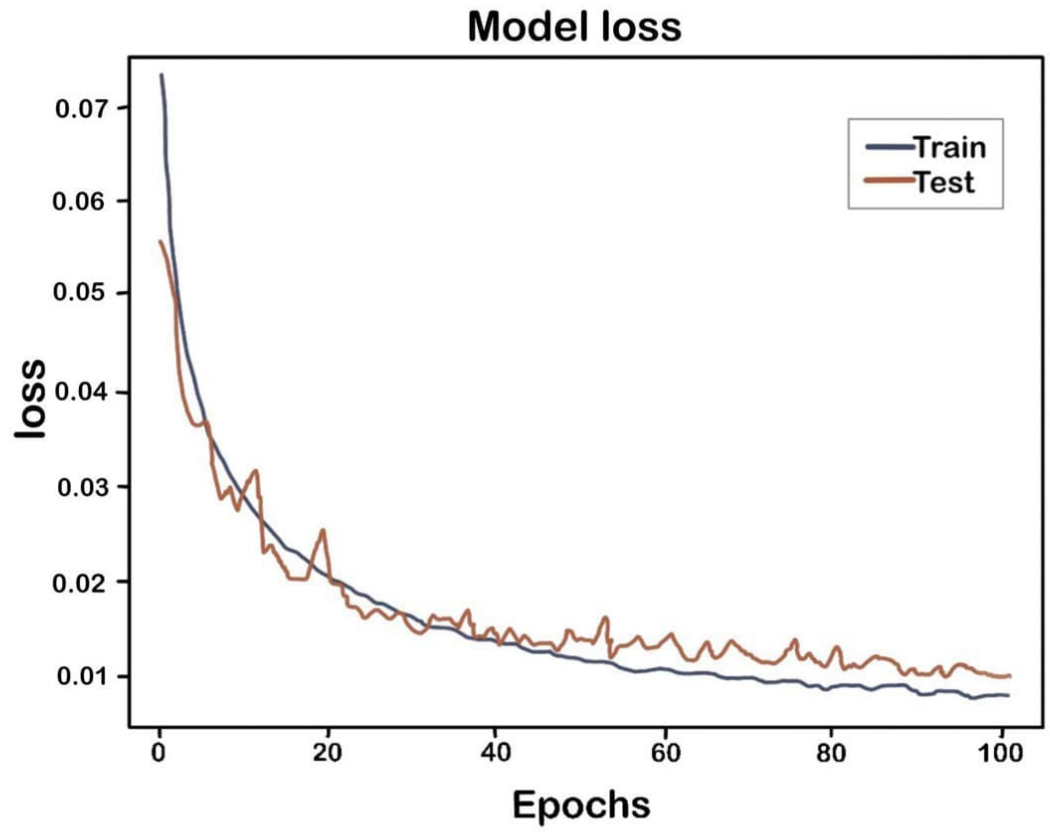


FIGURE 16. **AModel Loss for Suggested Schema.**



(a)



(b)

FIGURE 17. Results of Using Practical Dataset

6. CONCLUSION

This paper presents a new intelligent detection and classification method of series arc fault in photovoltaic systems. Two levels of detection and classification method based two different techniques of artificial intelligence are used. This method is based on extracting high frequency components from the DC voltage of PV systems. The primary issue with integrating new technology, such as series arc fault detection in PV arrays, is that the immature products will be susceptible to nuisance trips or alerts. To eliminate a such problem, voltage components associated with their energy variance is used to aid in the reduction of nuisance maloperation of the method; therefore, the components of the high-frequency signal of the voltage were used in this method. Furthermore, to evaluate the proposed method's performance, studies have been conducted in both simulation and practice, considering various possible scenarios which may emerge. To such an aim, alongside the records from actual measurements in practice, nine models of SAF are also employed for simulation, as well as a comprehensive comparison with related methods was presented. The results of the proposed method successfully demonstrate that it has the ability to detect the series arc with a very high accuracy of 99%. Although the proposed method has achieved a high accuracy, it can be improved by employing more models with the presented model and consider the highest frequency of voting. As the future work for the proposed method, it can be implemented in ASIC (Application-Specic Integrated Circuit) or FPGA (Field programmable Gate Arrays) to improve system reliability, increase speed of detection and reduce the cost.

FUNDING

None.

ACKNOWLEDGMENT

The authors would like to express the appreciation to the Ministry of Higher Education Malaysia (MOHE), the support of the sponsors [Vot Number = Q.J130000.3551.07G53 and R.J130000.4J347] and also to the Universiti Teknologi Malaysia (UTM) for providing the best education and research facilities to achieve the aims and goals in research studies and works.

CONFLICTS OF INTEREST

The author declares no conflict of interest.

REFERENCES

[1] R. H. Lee, "The other electrical hazard: Electric arc blast burns," *IEEE Transactions on Industry Applications*, no. 3, pp. 246–251, 1982.
[2] J.-H. Du, R. Tu, Y. Zeng, L. Pan, and R.-C. Zhang, "An experimental study on the thermal characteristics and heating effect of arc-fault from cu core in residential electrical wiring fires," *PloS one*, vol. 12, no. 8, p. e0182811, 2017.
[3] K. Yang, R. Chu, R. Zhang, J. Xiao, and R. Tu, "A novel methodology for series arc fault detection by temporal domain visualization and convolutional neural network," *Sensors*, vol. 20, no. 1, p. 162, 2020.
[4] Y. Wang, F. Zhang, and S. Zhang, "A new methodology for identifying arc fault by sparse representation and neural network," *IEEE Transactions on Instrumentation and Measurement*, vol. 67, no. 11, pp. 2526–2537, 2018.
[5] H.-K. Ji, G. Wang, W.-H. Kim, and G.-S. Kil, "Optimal design of a band pass filter and an algorithm for series arc detection," *Energies*, vol. 11, no. 4, p. 992, 2018.
[6] G. D. Gregory, K. Wong, and R. F. Dvorak, "More about arc-fault circuit interrupters," *IEEE Transactions on industry applications*, vol. 40, no. 4, pp. 1006–1011, 2004.
[7] Y.-H. Lin, C.-W. Liu, and C.-S. Chen, "A new pmu-based fault detection/location technique for transmission lines with consideration of arcing fault discrimination-part i: theory and algorithms," *IEEE Transactions on power delivery*, vol. 19, no. 4, pp. 1587–1593, 2004.
[8] M. Naidu, T. J. Schoepf, and S. Gopalakrishnan, "Arc fault detection scheme for 42-v automotive dc networks using current shunt," *IEEE Transactions on power electronics*, vol. 21, no. 3, pp. 633–639, 2006.
[9] G. D. Gregory, K. Wong, and R. Dvorak, "More about arc-fault circuit interrupters," in *38th IAS Annual Meeting on Conference Record of the Industry Applications Conference, 2003.*, vol. 2, pp. 1306–1313, IEEE, 2003.
[10] M. J. Albers and G. Ball, "Comparative evaluation of dc fault-mitigation techniques in large pv systems," *IEEE Journal of Photovoltaics*, vol. 5, no. 4, pp. 1169–1174, 2015.
[11] F. M. Uriarte, A. L. Gattozzi, J. D. Herbst, H. B. Estes, T. J. Hotz, A. Kwasinski, and R. E. Hebner, "A dc arc model for series faults in low voltage microgrids," *IEEE Transactions on smart grid*, vol. 3, no. 4, pp. 2063–2070, 2012.
[12] P. C. Hatton, M. Bathaniah, Z. Wang, and R. S. Balog, "Arc generator for photovoltaic arc fault detector testing," in *2016 IEEE 43rd Photovoltaic Specialists Conference (PVSC)*, pp. 1702–1707, IEEE, 2016.
[13] J. Johnson, M. Montoya, S. McCalmont, G. Katzir, F. Fuks, J. Earle, A. Fresquez, S. Gonzalez, and J. Granata, "Differentiating series and parallel photovoltaic arc-faults," in *2012 38th IEEE Photovoltaic Specialists Conference*, pp. 000720–000726, IEEE, 2012.
[14] L. Mackay, A. Shekhar, B. Roodenburg, L. Ramírez-Elizondo, and P. Bauer, "Series arc extinction in dc microgrids using load side voltage drop detection," in *2015 IEEE First International Conference on DC Microgrids (ICDCM)*, pp. 239–244, IEEE, 2015.

[15] A. Shekhar, L. Ramírez-Elizondo, S. Bandyopadhyay, L. Mackay, and P. Bauera, "Detection of series arcs using load side voltage drop for protection of low voltage dc systems," *IEEE Transactions on Smart Grid*, vol. 9, no. 6, pp. 6288–6297, 2017.

[16] H. Sagha, G. Mokhtari, A. Arefi, G. Nourbakhsh, G. Ledwich, and A. Ghosh, "A new approach to improve pv power injection in lv electrical systems using dvr," *IEEE Systems Journal*, vol. 12, no. 4, pp. 3324–3333, 2017.

[17] R. D. Telford, S. Galloway, B. Stephen, and I. Elders, "Diagnosis of series dc arc faults—a machine learning approach," *IEEE Transactions on Industrial Informatics*, vol. 13, no. 4, pp. 1598–1609, 2016.

[18] S. H. Abdulhussain, B. M. Mahmmod, M. I. Saripan, S. Al-Haddad, T. Baker, W. N. Flayyih, W. A. Jassim, *et al.*, "A fast feature extraction algorithm for image and video processing," in *2019 international joint conference on neural networks (IJCNN)*, pp. 1–8, IEEE, 2019.

[19] S. H. Abdulhussain, S. A. R. Al-Haddad, M. I. Saripan, B. M. Mahmmod, and A. Hussien, "Fast temporal video segmentation based on krawtchouk-tchebichef moments," *IEEE Access*, vol. 8, pp. 72347–72359, 2020.

[20] A. H. Omran, Y. M. Abid, and H. Kadhim, "Design of artificial neural networks system for intelligent chessboard," in *2017 4th IEEE International Conference on Engineering Technologies and Applied Sciences (ICETAS)*, pp. 1–7, IEEE, 2017.

[21] A. H. Omran, Y. M. Abid, A. S. Ahmed, H. Kadhim, and R. Jwad, "Maximizing the power of solar cells by using intelligent solar tracking system based on fpga," in *2018 Advances in Science and Engineering Technology International Conferences (ASET)*, pp. 1–5, IEEE, 2018.

[22] H. Samet, D. K. Asl, T. Ghanbari, and A. H. Omran, "Optimal number and location of the required measurment units for fault detection of pv arrays," in *2018 IEEE International Conference on Environment and Electrical Engineering and 2018 IEEE Industrial and Commercial Power Systems Europe (EEEIC/I&CPS Europe)*, pp. 1–5, IEEE, 2018.

[23] A. Aljuboori, A. Hamza, and Y. Alasady, "Novel intelligent traffic light system using pso and ann," *J. Adv. Res. Dyn. Control Syst*, vol. 11, pp. 1528–1539, 2019.

[24] L. L. Jiang and D. L. Maskell, "Automatic fault detection and diagnosis for photovoltaic systems using combined artificial neural network and analytical based methods," in *2015 International Joint Conference on Neural Networks (IJCNN)*, pp. 1–8, IEEE, 2015.

[25] A. Coleman and J. Zalewski, "Intelligent fault detection and diagnostics in solar plants," in *Proceedings of the 6th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems*, vol. 2, pp. 948–953, IEEE, 2011.

[26] A. Drews, A. De Keizer, H. G. Beyer, E. Lorenz, J. Betcke, W. Van Sark, W. Heydenreich, E. Wiemken, S. Stettler, P. Toggweiler, *et al.*, "Monitoring and remote failure detection of grid-connected pv systems based on satellite observations," *Solar energy*, vol. 81, no. 4, pp. 548–564, 2007.

[27] Z. Cheng, D. Zhong, B. Li, and Y. Liu, "Research on fault detection of pv array based on data fusion and fuzzy mathematics," in *2011 Asia-Pacific Power and Energy Engineering Conference*, pp. 1–4, IEEE, 2011.

[28] P. Ducange, M. Fazzolari, B. Lazzerini, and F. Marcelloni, "An intelligent system for detecting faults in photovoltaic fields," in *2011 11th International Conference on Intelligent Systems Design and Applications*, pp. 1341–1346, IEEE, 2011.

[29] L. Bonsignore, M. Davarifar, A. Rabhi, G. M. Tina, and A. Elhajjaji, "Neuro-fuzzy fault detection method for photovoltaic systems," *Energy Procedia*, vol. 62, pp. 431–441, 2014.

[30] Y. Yagi, H. Kishi, R. Hagihara, T. Tanaka, S. Kozuma, T. Ishida, M. Waki, M. Tanaka, and S. Kiyama, "Diagnostic technology and an expert system for photovoltaic systems using the learning method," *Solar Energy Materials and Solar Cells*, vol. 75, no. 3-4, pp. 655–663, 2003.

[31] A. A. A. M. Amiruddin, H. Zabiri, S. A. A. Taqvi, and L. D. Tufa, "Neural network applications in fault diagnosis and detection: an overview of implementations in engineering-related systems," *Neural Computing and Applications*, vol. 32, no. 2, pp. 447–472, 2020.

[32] J. A. Momoh and R. Button, "Design and analysis of aerospace dc arcing faults using fast fourier transformation and artificial neural network," in *2003 IEEE Power Engineering Society General Meeting (IEEE Cat. No. 03CH37491)*, vol. 2, pp. 788–793, IEEE, 2003.

[33] Z. Wang and R. S. Balog, "Arc fault and flash detection in photovoltaic systems using wavelet transform and support vector machines," in *2016 IEEE 43rd Photovoltaic Specialists Conference (PVSC)*, pp. 3275–3280, IEEE, 2016.

[34] B. Grichting, J. Goette, and M. Jacomet, "Cascaded fuzzy logic based arc fault detection in photovoltaic applications," in *2015 International Conference on Clean Electrical Power (ICCEP)*, pp. 178–183, IEEE, 2015.

[35] S. Lu, T. Sirojan, B. Phung, D. Zhang, and E. Ambikairajah, "Da-dcgan: An effective methodology for dc series arc fault diagnosis in photovoltaic systems," *IEEE Access*, vol. 7, pp. 45831–45840, 2019.

[36] A. H. Omran, D. M. Said, S. M. Hussin, N. Ahmad, and H. Samet, "A novel intelligent detection schema of series arc fault in photovoltaic (pv) system based convolutional neural network," *Periodicals of Engineering and Natural Sciences (PEN)*, vol. 8, no. 3, pp. 1641–1653, 2020.

[37] S. Lu, B. Phung, and D. Zhang, "A comprehensive review on dc arc faults and their diagnosis methods in photovoltaic systems," *Renewable and Sustainable Energy Reviews*, vol. 89, pp. 88–98, 2018.

[38] N. A. M. Isa, F. R. Hashim, F. W. Mei, D. A. Ramli, W. M. W. Omar, and K. Z. Zamli, "Predicting quality of river's water based on algae composition using artificial neural network," in *2006 4th IEEE International Conference on Industrial Informatics*, pp. 1340–1345, IEEE, 2006.

[39] R. Pascanu, T. Mikolov, and Y. Bengio, "On the difficulty of training recurrent neural networks," in *International conference on machine learning*, pp. 1310–1318, PMLR, 2013.

[40] C. An, J. Huang, S. Chang, and Z. Huang, "Question similarity modeling with bidirectional long short-term memory neural network," in *2016 IEEE First International Conference on Data Science in Cyberspace (DSC)*, pp. 318–322, IEEE, 2016.

[41] X. Li, H. Xianyu, J. Tian, W. Chen, F. Meng, M. Xu, and L. Cai, "A deep bidirectional long short-term memory based multi-scale approach for music dynamic emotion prediction," in *2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 544–548, 2016.

[42] D. Tao, Y. Wen, and R. Hong, "Multicolumn bidirectional long short-term memory for mobile devices-based human activity recognition," *IEEE Internet of Things Journal*, vol. 3, no. 6, pp. 1124–1134, 2016.

[43] I. N. Yulita, M. I. Fanany, and A. M. Arymuthy, "Bi-directional long short-term memory using quantized data of deep belief networks for sleep stage classification," *Procedia Computer Science*, vol. 116, pp. 530–538, 2017. Discovery and innovation of computer science technology in artificial intelligence era: The 2nd International Conference on Computer Science and Computational Intelligence (ICCSCI 2017).

[44] H. Ayrton, *The electric arc*. Cambridge University Press, 2012.

[45] C. P. Steinmetz, "Transformation of electric power into light," *Proceedings of the American Institute of Electrical Engineers*, vol. 25, no. 11, pp. 755–779, 1906.

[46] T. Gammon, W.-J. Lee, Z. Zhang, and B. C. Johnson, "A review of commonly used dc arc models," *IEEE Transactions on Industry Applications*, vol. 51, no. 2, pp. 1398–1407, 2015.

[47] A. Cassie, "Theorie nouvelle des arcs de rupture et de la rigidité des circuits," *Cigre, Report*, vol. 102, pp. 588–608, 1939.

[48] O. Mayr, "Beiträge zur theorie des statischen und des dynamischen lichtbogens," *Archiv für Elektrotechnik*, vol. 37, pp. 588–608, 1943.

[49] O. Mayr, "Über die theorie des lichtbogens und seiner löschung," *Elektrotechnische Zeitschrift*, vol. 64, no. 49/50, pp. 645–652, 1943.

[50] U. Habedank, "On the mathematical description of arc behaviour in the vicinity of current zero," *etz. Archiv*, vol. 10, pp. 339–343, 1988.

[51] R. Smeets and V. Kertesz, "Evaluation of hv-circuit breaker performance with a validated arc model," *Generation, Transmission and Distribution, IEE Proceedings-*, vol. 147, pp. 121 – 125, 04 2000.
[52] A. Hu and K. Zhang, "Using bidirectional long short-term memory method for the height of f2 peak forecasting from ionosonde measurements in the australian region," *Remote Sensing*, vol. 10, no. 10, 2018.
[53] W. Lu, B. Liang, Y. Cheng, D. Meng, J. Yang, and T. Zhang, "Deep model based domain adaptation for fault diagnosis," *IEEE Transactions on Industrial Electronics*, vol. 64, no. 3, pp. 2296–2305, 2017.
[54] A. Haque, K. V. S. Bharath, M. A. Khan, I. Khan, and Z. A. Jaffery, "Fault diagnosis of photovoltaic modules," *Energy Science & Engineering*, vol. 7, no. 3, pp. 622–644, 2019.
[55] J. Sun, F. Sun, J. Fan, and Y. Liang, "Fault diagnosis model of photovoltaic array based on least squares support vector machine in bayesian framework," *Applied Sciences*, vol. 7, no. 11, 2017.
[56] Y. F. Li, P. J. Lin, H. F. Zhou, Z. C. Chen, L. J. Wu, S. Y. Cheng, and F. P. Su, "On-line monitoring system of PV array based on internet of things technology," *IOP Conference Series: Earth and Environmental Science*, vol. 93, p. 012078, nov 2017.