IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received 29 June 2022, accepted 15 July 2022, date of publication 20 July 2022, date of current version 26 July 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3192517

RESEARCH ARTICLE

Identifying DC Series and Parallel Arcs Based on Deep Learning Algorithms

HOANG-LONG DANG[1], **SANGSHIN KWAK**[1], **(Member, IEEE), AND SEUNGDEOG CHOI**[2], **(Senior Member, IEEE)**
[1]School of Electrical and Electronics Engineering, Chung-Ang University, Seoul 06974, South Korea
[2]Department of Electrical and Computer Engineering, Mississippi State University, Starkville, MS 39762, USA

Corresponding authors: Sangshin Kwak (sskwak@cau.ac.kr) and Seungdeog Choi (seungdeog@ece.msstate.edu)

This work was supported in part by the Technology Development Program to Solve Climate Changes through the National Research Foundation of Korea (NRF) through the Ministry of Science, ICT, under Grant 2021M1A2A2060313; and in part by the Korea Electric Power Corporation under Grant R21XA01-3.

ABSTRACT Arc phenomena are usually related to the undesired disengagement of two electrical connections. The emission power discharge from the failure arc may damage wiring and can present a fire hazard. Numerous studies have been proposed to detect arc events and quickly isolate them from an electrical system. DC arc faults are often sorted into two types: series and parallel arcs. A series arc may be the outcome of discharging links in electrical wiring. By contrast, the parallel arc occurs between two electric wires, or between a link and a ground, owing to contamination or poor isolation. The currents in a system with an arc fault are considerably greater when the arc parallel in nature than when the arc is series in nature. In this paper, the electric activities of a network are investigated for the duration of series and parallel arc failures in both the time and frequency domains. The arcing behavior investigated is selected to allow for the identification of series and parallel arcs. The sorting of electrical arcs in an accurate and reliable manner is useful for electrical protection schemes. The identification process used here is based on data related to different domains, such as load current and voltage. In this study, eight learning techniques are investigated with the aim of detecting series and parallel arc faults. The arc behaviors were studied in the various domains. We used the load current and voltage characteristics as an statistic for categorizing a given arc failure. This study could be beneficial to enhance the stability and reliability of arc-fault detectors.

INDEX TERMS Arc diagnosis, artificial intelligence, DC arc failure, identifying arc fault.

I. INTRODUCTION

The arc hazard posed in DC systems has become well known over the past decade due to the widespread application of DC networks in aircraft, the solar industry, data management, electric transmission, and numerous fields related to facilitated usage and high efficiency. An arcing event is a hazardous incident that must be seriously considered in any power network [1]. Steps should be taken to avoid system failure, especially arc faults, in utility DC systems. In DC systems, arc failures are categorized as series and parallel failures. A series type can be caused by discharging electrical wiring links [2]. An impermanent short circuit typically triggers the series fault. Potential causes include the releasing of wiring contacts resulting in feeble connections [3]. If the range of an arc current is within the range of the implemented safety devices, the arc will not be detected. Suppose the range of the arc currents is two to five times the safety device limit. In that case, the arc can scorch the conductors in an excessively extended time before the safety means disconnecting the error from the functioning system. Parallel arcs represent discharge events between two points that exhibit a voltage difference. This kind of fault may occur due to damaged insulation [4]. There is a possibility of fire

The associate editor coordinating the review of this manuscript and approving it for publication was Ton Duc Do.

when the arc current is smaller than the rated current of the safety gadgets implemented in the network in the case of a parallel arc. Furthermore, the parallel arc current increases when an arc occurs; thus, parallel arcs are potentially more dangerous than series arcs. The rise in current amplitude and heat in a short time period in the case of a parallel arc fault can induce a large flare, with the potential to destroy conductors and electrical wiring [5]. Therefore, detecting arc failures timely and correctly is a critical assignment in ensuring the safety and reliability of electrical networks. When the DC arc faults are initiated, several abnormal behaviors can be used to diagnose an arc event, such as current fluctuations and rapid changes in light and/or heat output. These abnormal phenomena can be used to detect DC arcs [6]–[11]. However, the investigation of parallel arc in DC systems is still at a primitive step [12]–[15]. The characteristic differences between arc types in a DC network are demonstrated in Ref. [16]. Recently, advanced techniques have received increased research attention due to their flexible capabilities for the variety of purposes. Artificial intelligence (AI) models have been efficaciously utilized in numerous fields. They supply effective approaches for the identification of failures in many applications. Learning algorithms have been effectively adapted for discovering arc failures, and talented results have been obtained, for instance, various features, such as high-frequency components and alterations of current, can be obtained, and the weighted least squares SVM algorithm can be used to diagnose series arcs [17]. The sparse coding characteristics combined with an artificial neural network for arc fault diagnosis were proposed in Ref. [18]. In Refs. [19]–[22], numerous AI models were employed to diagnose series arcing events using different characteristics as inputs. The adoption of AI algorithms for parallel arc diagnosis was proposed in [23]. Previous studies have illustrated performance comparisons among various AI algorithms in DC networks [24]. Commonly, mentioned investigations concentrate simply on series or parallel types, while the implementation of AI to identify arcing types has not been completely explored. When an arc failure occurs in practical systems, the type and location are unknown. Usually, one approach is adopted to detect series fault type and another for parallel type. For example, there are various types of faults in the DC networks. Each type of failure needs a different detecting approach; this could increase the additional hardware, complexity, and cost of the systems. Suppose it is possible to identify the type of arc fault using a universal approach. In that case, the proper solutions can be applied promptly to maintain the system's safety, stability, and economy. However, the potential capability of AI techniques is not fully utilized. This study aims to provide an insight into a potentially universal approach that is the capability to detect various types of faults, especially arc faults. In this study, the AI training models are obtained by using the data in certain operating conditions. If the operating conditions are changed, the AI models must be trained again to learn new characteristics of the input signals. All the AI techniques in this research belong to supervised AI algorithms. Therefore, the trained models need adjustments from supervisors for any changes in working conditions to maintain high performance.

This paper implements eight AI algorithms with the aim of detecting arc events and identifying the arc type. Furthermore, this study recommends the most suitable input that returns the best result for the different AI algorithms and arc types. In addition, the operation of the AI models and inputs is compared and discussed. This research is arranged as follows: Part 2 describes the experimental setup and the characteristics of arcing and normal states of both parallel and series arcs. Part 3 describes the learning algorithms and inputs utilized for arc recognition. Part 4 compares the identification performances of the eight algorithms using different combinations of features for arc events that occur in various operating situations. Finally, in Part 5, we discuss recommendations for arc failure detection and the identification diagnosis rates.

II. PROPERTIES OF DC ARCING FAILURES

The arc setup was formed to gather arc data conforming to the UL1699B standard; the setup used is shown in Figure 1. The arc generation setup includes a DC source, an arc generator, and an inverter load [3]. The amplitude of the supply voltage used in the arc experiment was 300 V for both the arc failures. The power source used in the parallel and series arc experiments was the KEYSIGHT N8741A (maximum current: 11 A, maximum voltage: 300 V, and maximum power: 3.3 kW). The parameter i_{arc} represents the arc current flowing through the bars. The step motor slowly splits the bars to make the arcing failure safely. The gap was monitored with an electrical meter, which was set up parallel to the bars. An inductor and a resistor of 10 mH and 10 Ω, respectively, were behaved as the loads of the inverter. The parallel and series arc circuits used the same load parameters. In the parallel arc experiment, the arc generator was connected in series with the resistor R_{limit} to ensure the safe operation of the circuit. This was done because the amount of source current, i_s, increases rapidly when an arc is generated. Table 1 shows the values of the parameters that characterize the parallel and series arc faults.

The experimental process used was the same for both the parallel and series arcs. The load voltage (v_L), load current (i_L), and arc current were stored using an oscilloscope with a sample rate of 250 kHz. In practical applications, the measurement of the arc current, especially in the case of a parallel arc fault, is unachievable because the position of the fault is unidentified. Thus, the arc current is not utilized in the identification procedure of this study. Instead, the load current and the load voltage could be utilized for arc identification. For the various applications, the load current is measured for controlled purposes. Therefore, the current sensors are already implemented. The demand for additional sensors is neglected using the load current as a parameter. When the

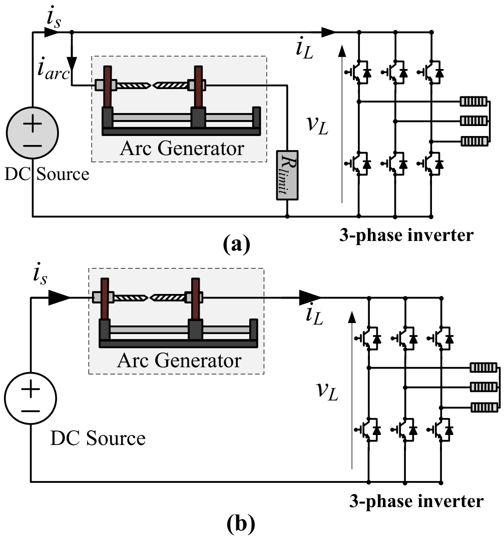

FIGURE 1. DC arc setup. The circuit used for (a) parallel and (b) series arc faults.

TABLE 1. Characterization of the experiments.

Arc type	Load & modulation technique	Load current (i_L)	Limit resistor (R_{limit})	Arc current (i_{arc})	Switching frequency
Parallel	Three-phase inverter and space vector modulation	3 A	600 Ω	0.5 A	5 kHz
					15 kHz
			300 Ω	1 A	5 kHz
					15 kHz
		5 A	600 Ω	0.5 A	5 kHz
					15 kHz
			300 Ω	1 A	5 kHz
					15 kHz
Series	Three-phase inverter and space vector modulation	5 A	N/A	5 A	5 kHz
					10 kHz
					15 kHz
					20 kHz
		8 A	N/A	8 A	5 kHz
					10 kHz
					15 kHz
					20 kHz

amplitudes of load current are changed, the training stage must be repeated for the new data. So that AI models can learn the new characteristic to maintain high performance. The identification procedure was then performed using codes written in MATLAB. The range of sample rate (250 kHz) was chosen regarding recent studies of arc faults in DC networks [25]–[29], thus, the number of data in 2 ms period is 500. A higher sampling rate could result in more data in every duration. Nevertheless, it might grow the performance time and computation weight; one of the highest primacies of arc diagnosis is identifying faults in a timely manner in order to quickly isolate the error from the network. Hence, the sample rate of 250 kHz was considered to be sufficiently high to ensure a balance between efficiency and execution time. Additionally, the recent arc fault research selected similar data durations (arrangements of 2, 3, or 4 ms periods) [21]–[25]. Therefore, this study chooses the window duration at a 2 ms period. The experimental waveforms were split into altered arrangements of 2 ms periods; these signals were then used for test and training procedures in the AI methods. This research utilized space vector modulation to control the inverter load. Figure 2 shows the experimental signals corresponding to the arcing and normal states of the series and parallel arcs under different settings (load voltage: 300 V, load current: 5 A, arc current: 1 A, and switching frequency: 15 kHz in the case of the parallel arc; load voltage: 300 V, 8 A load and arc current, and switching frequency: 15 kHz in the case of the series arc). As shown in Figure 2, the shapes of waveforms were steady and alike before the arcs were induced. When the arcs were instigated, however, various unusual characteristics were induced in the waveforms, such as harmonic elements, distortions, and the fluctuations of the waveforms. These unusual behaviors led to the initiation of enormous negative oscillations in the parallel types [Figure 2(a)]. Some over-current detection schemes could detect the enormous oscillations in load current and voltage of parallel arc fault. However, there were also slight decreases in the load current and the voltage in the case of the series arc [Figure 2(b)]. These slight decreases in the load current and voltage of series cases can not be detected using the same parallel case detection schemes. Using different detection schemes for different arc-type faults could increase the additional hardware, complexity, and cost of the systems. Therefore, a universal scheme is more beneficial in identifying the types of arc faults. This research aims to provide an insight into the universal approach that is the capability to detect various types of arc faults. Thus, the influences of different load types are omitted in this study. These influences will be the future work of this research because each load type shows different characteristics. When applying the new load types, the data of the new load type must be trained so that the AI models can learn the unique features of each load type. If the data of load type are not trained, AI models' performance could degrade considerably.

III. ARTIFICIAL LEARNING ALGORITHMS AND FEATURE ANALYSIS

A. ARTIFICIAL LEARNING PRINCIPLES

Figure 3 demonstrates the concept and construction of several AI models. The intention of SVM model is to discover the greatest boundary hyperplane, which is then utilized to categorize the elements into different classes [30]. The K-nearest neighbor (KNN) method presupposes the same objects located in adjacent familiarity. In other words, the same objects encircle their relatives [31]. A decision tree (DT) structure represents the workflow; this workflow considers all possible decisions and the outcomes of those decisions. This structure shows the predictions resulting from different nodes. The root node represents the start of the DT, and the structure is terminated with the conclusion made at the "leaves" [32]. The random forest (RF) technique contains numerous separate DTs; this method works as an ensemble of

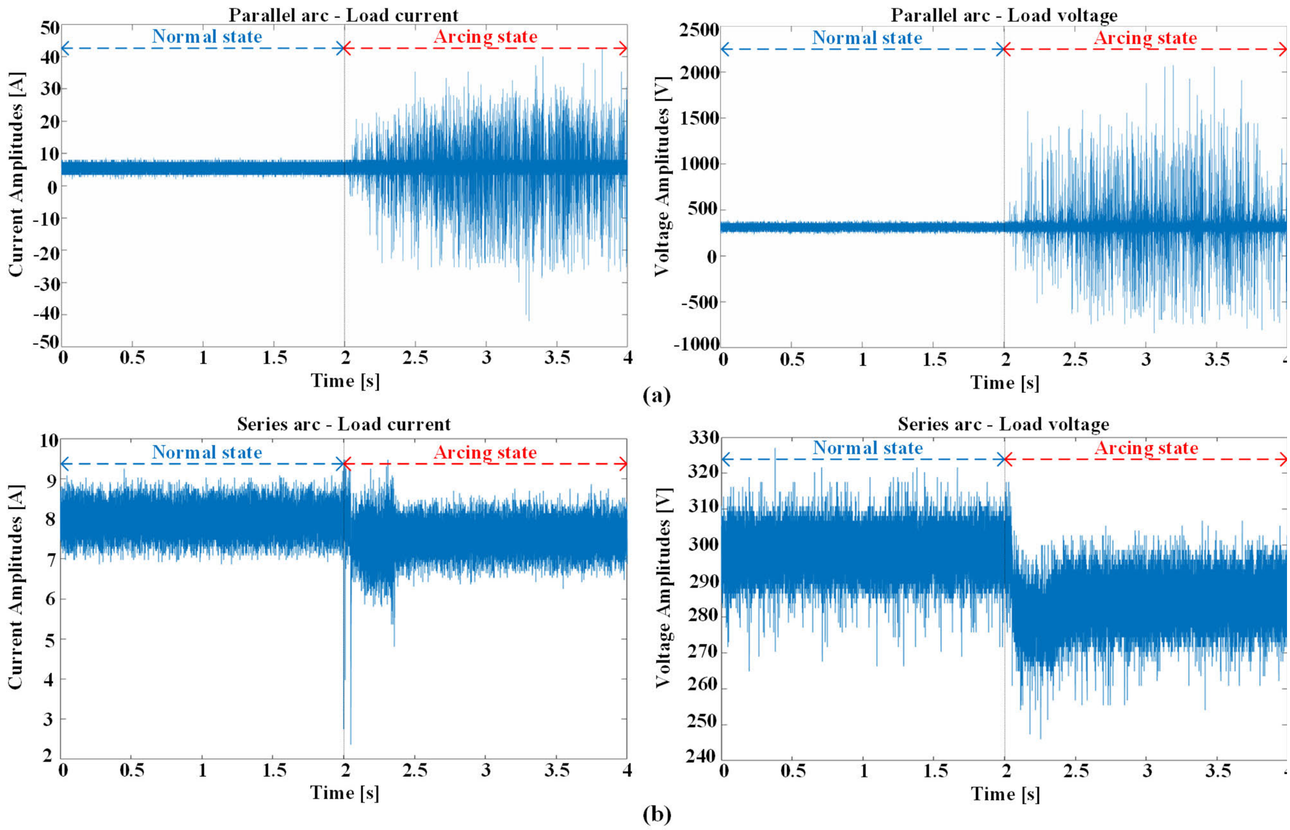

FIGURE 2. The load current and voltage waveforms in the case of parallel and series arcs. In the case of (a) a parallel arc with a 5 A load current and a 300 V load voltage and (b) a series arc with an 8 A load current and a 300 V load voltage.

DTs. Each DT in a RF yields a prediction, and the prediction generated by the RF is given by the most popular prediction from each of the individual DTs [33]. The Naïve Bayes (NB) models are a family of procedures based on Bayes' statement; all of the classifiers considered are based on a mutual assumption. In order to reduce the computational burden, this theorem assumes all features are separate from others [34]. On the other hand, deep learning (DL) trains processores to mimic the natural computational ability of humans, including learning, by example. In DL, a processing model is adopted to train and classify elements based on various types of data, such as text, sound, or images. DL models can achieve superior accuracy and exceed human-level performance in some tasks. A large data set is usually employed to train DL models. One of the most widely used DL models is the neural network (NN) model. There are various layers in NN's structure, and they are the input layer, hidden layers, and output layer. The hidden structure consists of several layers. There are various neurons in each layer. The input of a given neuron in the n^{th} layer could be the output of a neuron in the $n\text{-}1^{th}$ layer [35]. Table 2 lists the hidden structures of DL methods (Gated Recurrent Unit (GRU), Long-Short Term Memory (LSTM) and Deep Neural Network (DNN)) utilized in this study. The trial and error method chooses the hidden layer arrangements. The selected structures of the DL methods provided the highest efficiency across several arrangements. However, other possible arrangements, which may also be appropriate, exist [36].

TABLE 2. Hidden layers in deep learning models.

DL models	1st layer/ neurons	2nd layer/ neurons	3rd layer/ neurons	4th layer/ neurons	5th layer/ neurons
LSTM	Fully Connected /100	LSTM/ 16	Fully Connected /8	LSTM/8	Fully Connected /2
GRU	Fully Connected /100	GRU/16	Fully Connected /8	GRU/8	Fully Connected /2
DNN	Fully Connected /100	Fully Connected /50	Fully Connected /2	N/A	N/A

B. FEATURE ANALYSIS

Features can be considered in the frequency or time domains. However, they require significant computational effort and high sample rates to use feature analysis in the frequency domain effectively. For example, the feature could be directly extracted using the time domain signal. In contrast, it must be transformed to the frequency domain before extracting features. Generally, the data for the training stage of AI models should be large to obtain high performance. The

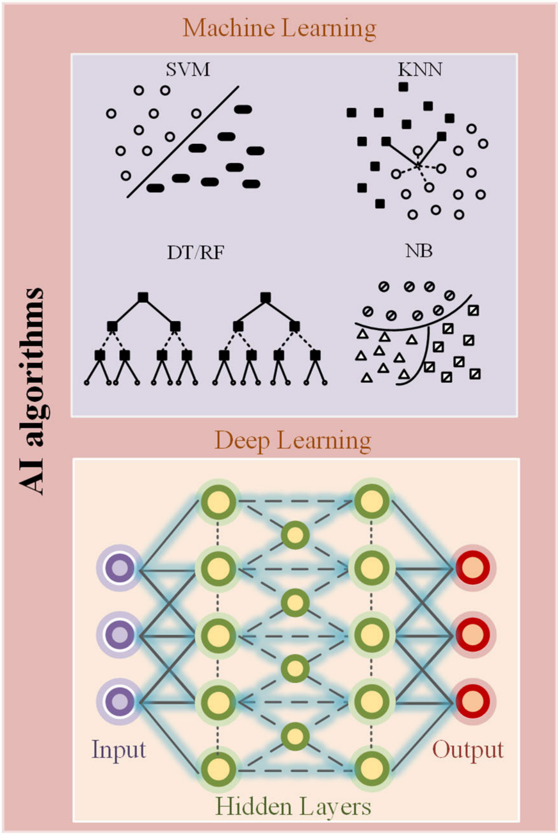

FIGURE 3. Concepts and configurations of the AI models.

time for transforming data from the time domain to the frequency domain increase with the increase of data quantities. Therefore, the additional frequency-domain transform would increase the performance time and adversely affect performance when a fault occurs in actual systems. This study employed time-domain inputs for the identification of parallel and series arcs. Figure 4 displays the feature investigation of the parallel and series arc load currents under the operating conditions represented in Figure 2. Figure 4 shows that the feature waveforms of both parallel and series arcs were different in the arcing states. When the arc fault was generated, many changes between the normal and fault states could be observed in the characteristics of signals recorded here. These abnormal characteristics are well suited to arc fault identification. In addition, we note that the parallel arc amplifies the load current when the arc happens. As a result, the distortions in the failure state are greater than those observed in the normal state. By contrast, the load current in a series arc tends to decrease when an arc event occurs. Thus, the harmonics in the series arc are much smaller than those observed in the parallel arc. The data were first stored at a sample rate of 250 kHz in this work. After that, the stored data were split into altered data sets of 2 ms in duration. Then, the features for each data set were obtained to construct a five-feature group. The raw signals and feature analysis were then used as inputs for the eight advanced learning algorithms to identify the arc type.

IV. ARTIFICIAL LEARNING ALGORITHMS IN DC ARC DIAGNOSIS

Figure 5 shows the diagnosis scheme for identifying the arc type. First, the load current and voltage data are collected and processed to obtain the features. There are three cases with three different inputs: features, raw data, and features and raw data combined. These inputs are then used in the DL and ML techniques to identify the arc type. Case 1 uses only the features obtained from the experimental data as input, and case 2 uses the corresponding raw data. In the third case, both the obtained features and the raw data are input into the learning algorithms. The training set is constructed from 1,000 data sets from each experimental case, and there are 16,000 data set in each training case (cases 1, 2, and 3). Similarly, the test set is constructed from 800 data sets for each experimental case, and there are 12,800 data set in each test case (including cases 1, 2, and 3). The same data sets do not appear in both the test and training data. Additionally, eight experimental cases for each parallel or series arc exist in Table 1. Therefore, there exist 64 possible pairs between parallel and series arcs in each case for the test process. Because the number of test pairs is large, the performance of the test pairs is given by the average performance of the test pairs with the same load and arc currents with different switching frequencies, as shown in Figure 6. The ratio data corresponding to an arc state (either parallel or series) and a normal state is 1:1 in both the training and test processes.

The accuracy metric used here is adapted to measure the proficiency of the identification performance. The identification judgment rate is given by the proportion of the data sets that are correctly predicted and the total number of test data sets. The effectiveness of parallel arc identification is expressed as:

$$Parallel\ Acc. = \frac{\#\ of\ correctly\ predicted\ parallel\ data\ sets}{total\ \#\ of\ parallel\ data\ sets} \tag{1}$$

$$Seires\ Acc. = \frac{\#\ of\ wrong\ predicted\ parallel\ data\ sets\ as\ series\ arc}{total\ \#\ of\ parallel\ data\ sets} \tag{2}$$

$$Normal\ Acc. = \frac{\#\ of\ wrong\ predicted\ parallel\ data\ sets\ as\ normal}{total\ \#\ of\ parallel\ data\ sets} \tag{3}$$

Using the above similar manners, the effectiveness identifications of series and normal states could be obtained. The best learning algorithm is that which has the highest identification and lowest wrong prediction rates.

Figure 7 demonstrates the average identification performance of the various AI models in case 1 with the test pair of parallel arcs with a 3 A load current, a 0.5 A arc

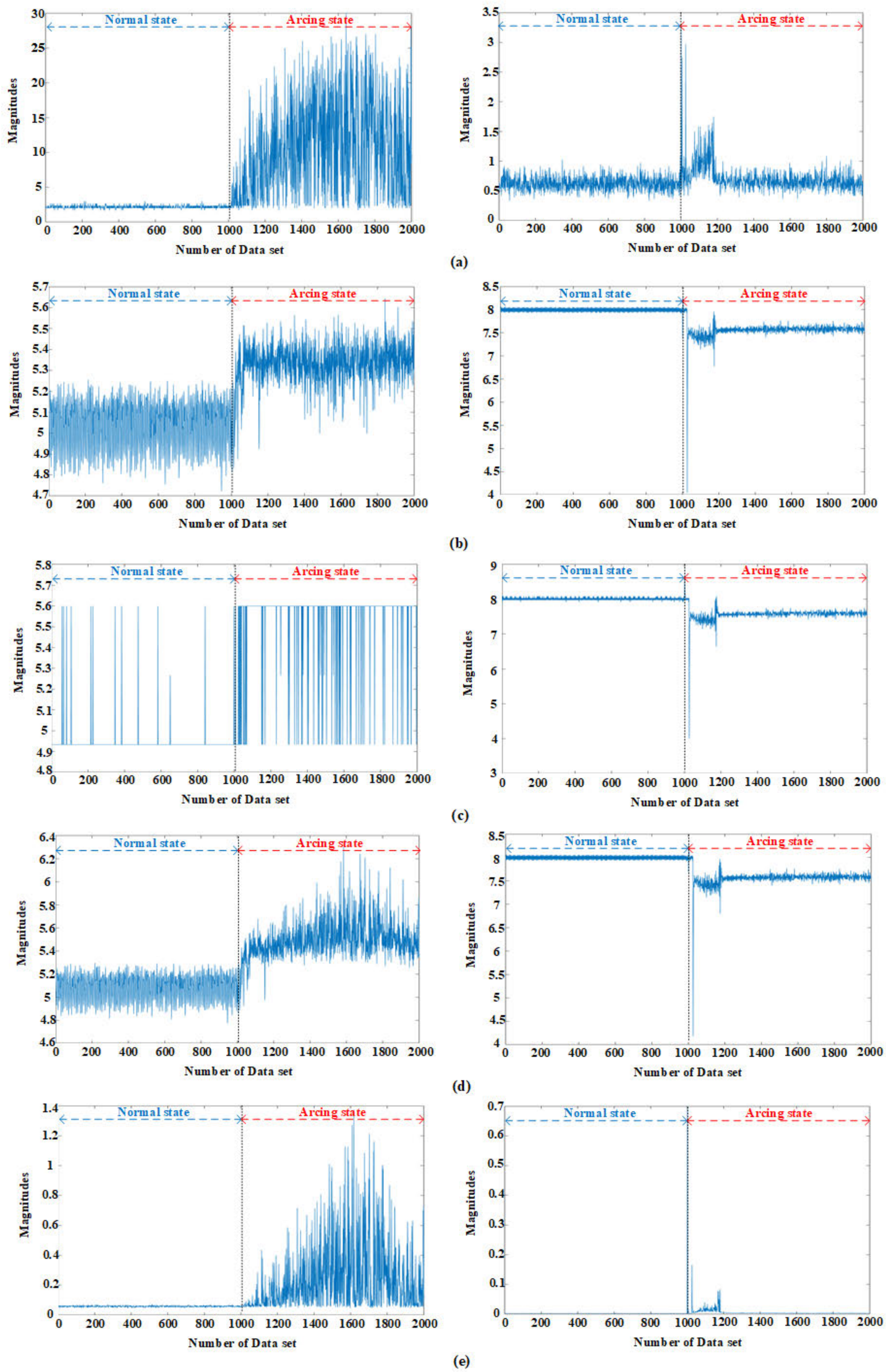

FIGURE 4. Feature analysis from load current related to parallel (left) and series (right) arcs. These features are (a) Peak-to-peak values, (b) average values, (c) median values, (d) rms values, and (e) variance values.

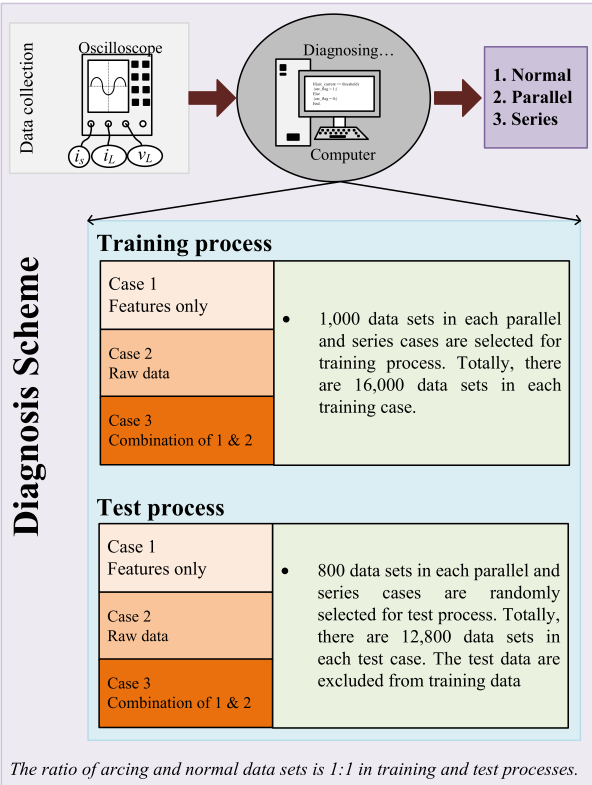

FIGURE 5. Fraction of data.

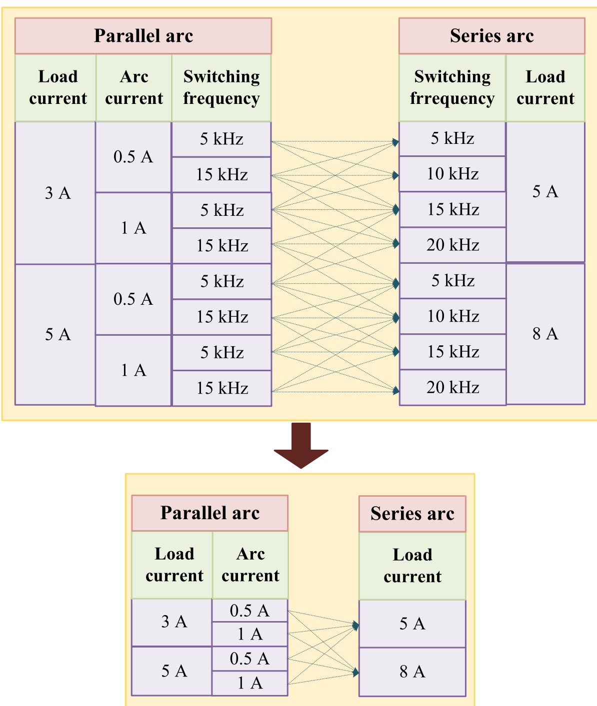

FIGURE 6. Possible test pairs involving parallel and series arcs.

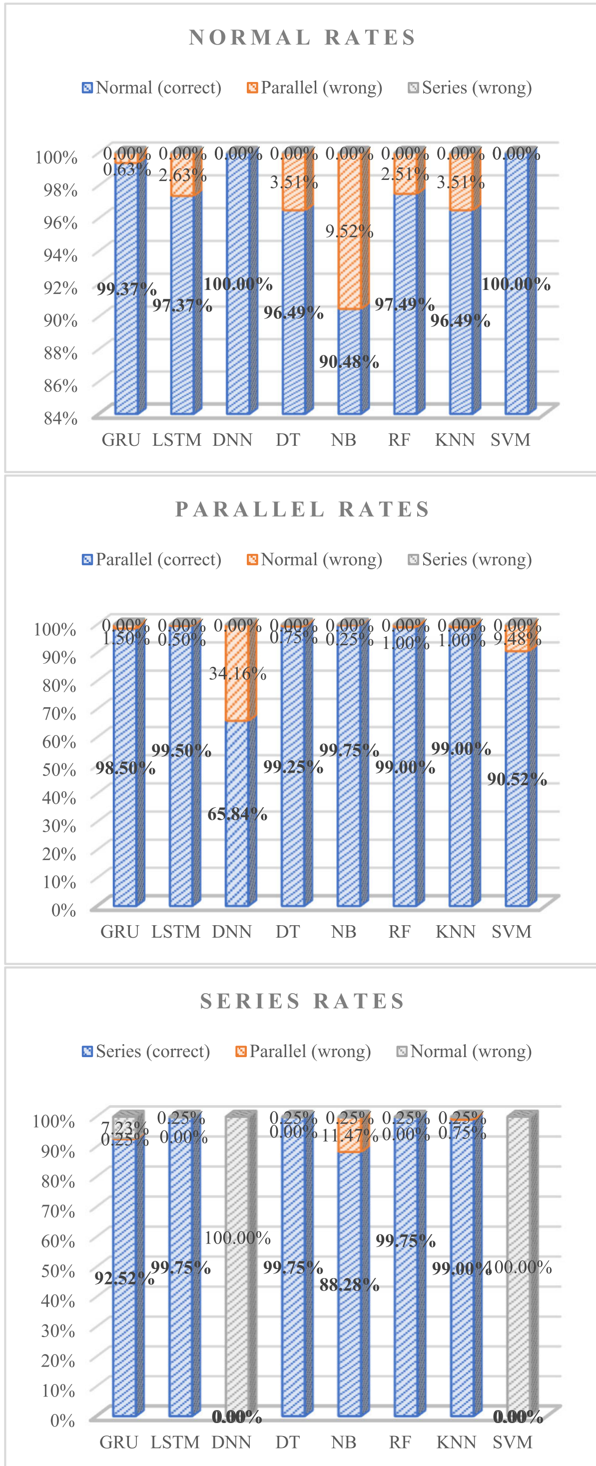

FIGURE 7. Identification performance of AI models in case 1 with the test pair of parallel arc at 3 A load current, 0.5 A arc current, and series arc at 5 A load current.

current, and a series arc at 5 A load current. All AI techniques show high performance in the identification of the normal, parallel, and series arcing states, with the exception of the DNN and SVM algorithms. The RF, LSTM, DT, and KNN algorithms were found to have superior performances than the alternative techniques. Their accuracies were above 96% for all the considered states. The performances of the GRU and NB algorithms were high (above 90%). Figure 8 shows

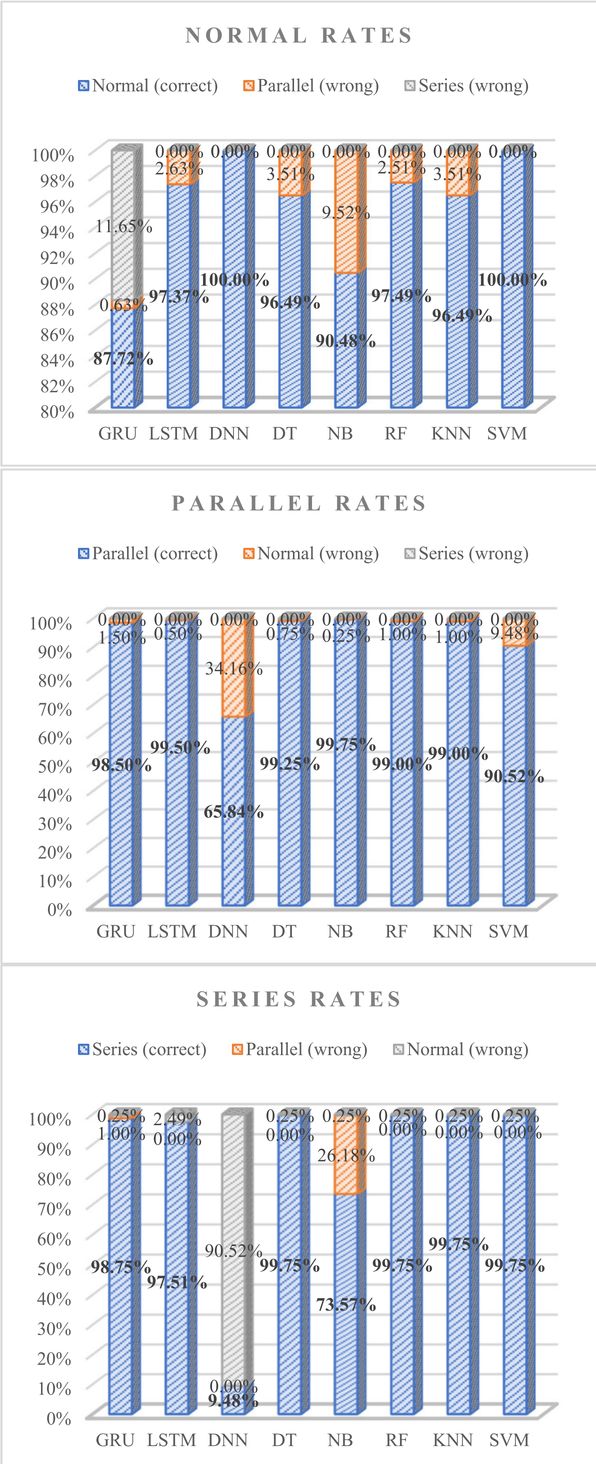

FIGURE 8. Identification performance of the AI models in case 1 with the test pair of parallel arc at 3 A load current, 0.5 A arc current, and series arc at 8 A load current.

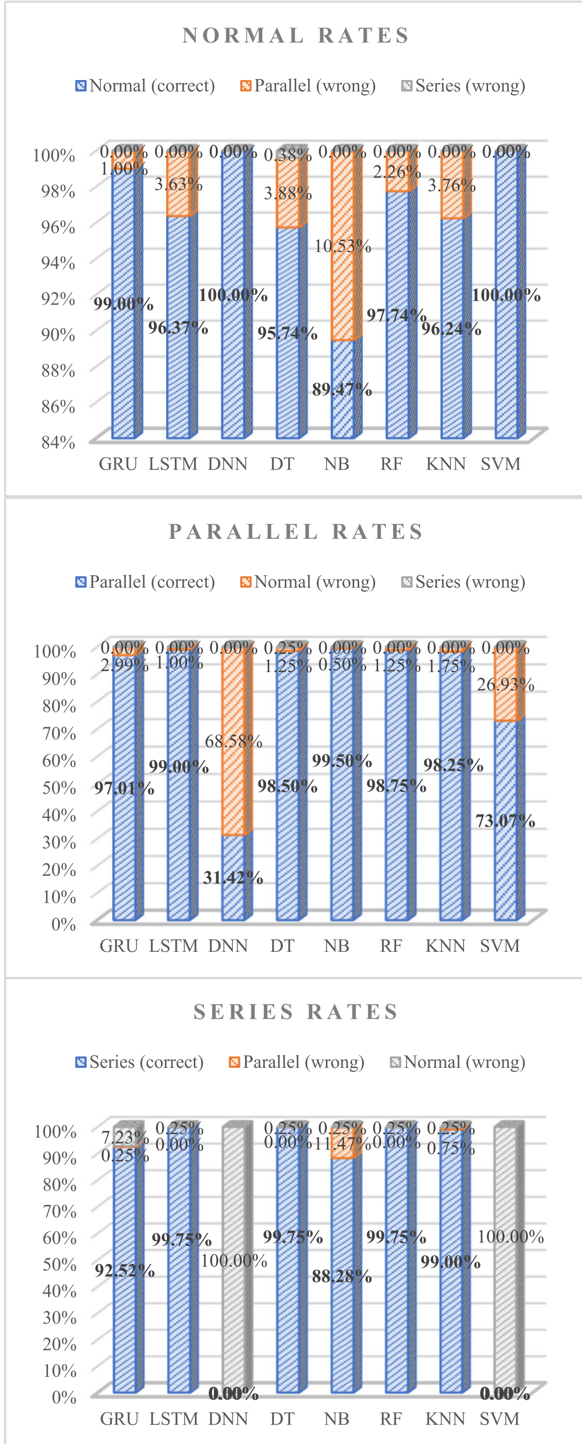

FIGURE 9. Identification performance of AI models in case 1 with the test pair of parallel arc at 3 A load current, 1 A arc current, and series arc at 5 A load current.

results obtained considering the same values of load and arc currents of a parallel arc as were used to obtain the data represented in Figure 7, whereas the load current of the series arc was increased to 8 A. The DT, KNN, RF, and LSTM algorithms were found to detect the states with the highest grades (above 96%) compared with the other AI models considered here. The identifying accuracies of the SVM and GRU methods were also high (above 87%). The NB and DNN algorithms were found to have mediocre performance; the identification accuracies of the DNN and NB methods were the lowest for detecting series arcing state.

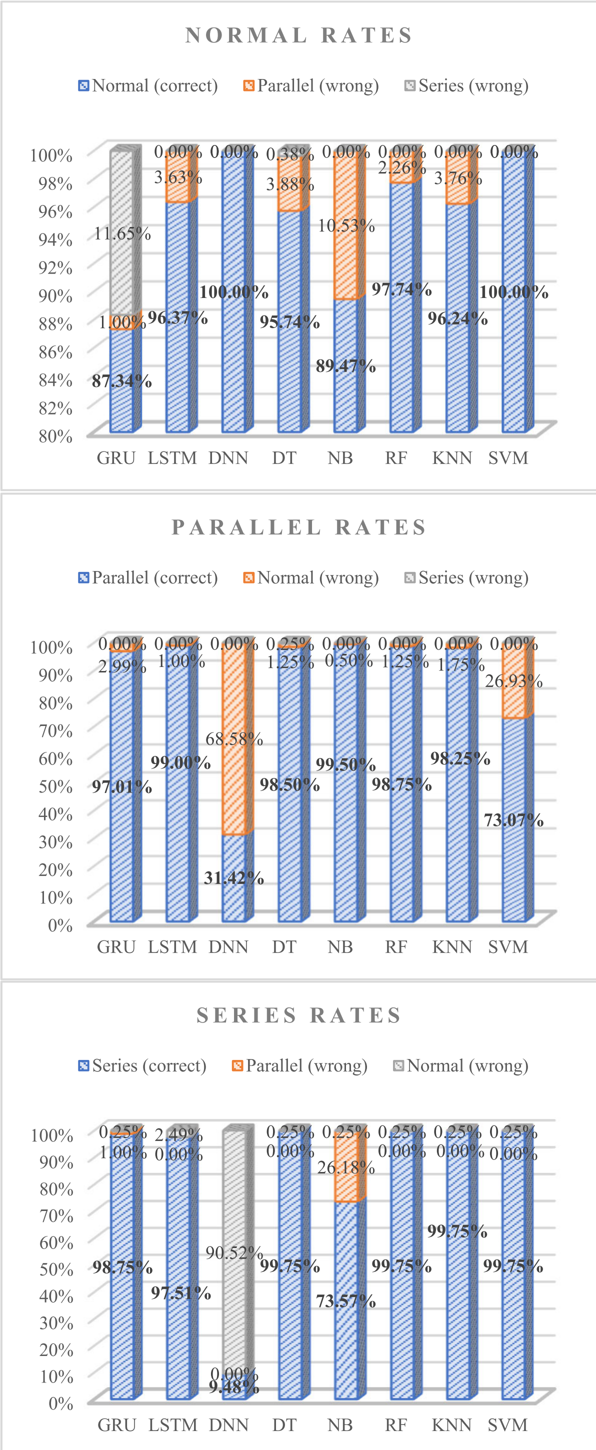

FIGURE 10. Identification performance of AI models in case 1 with the test pair of parallel arc at 3 A load current, 1 A arc current, and series arc at 8 A load current.

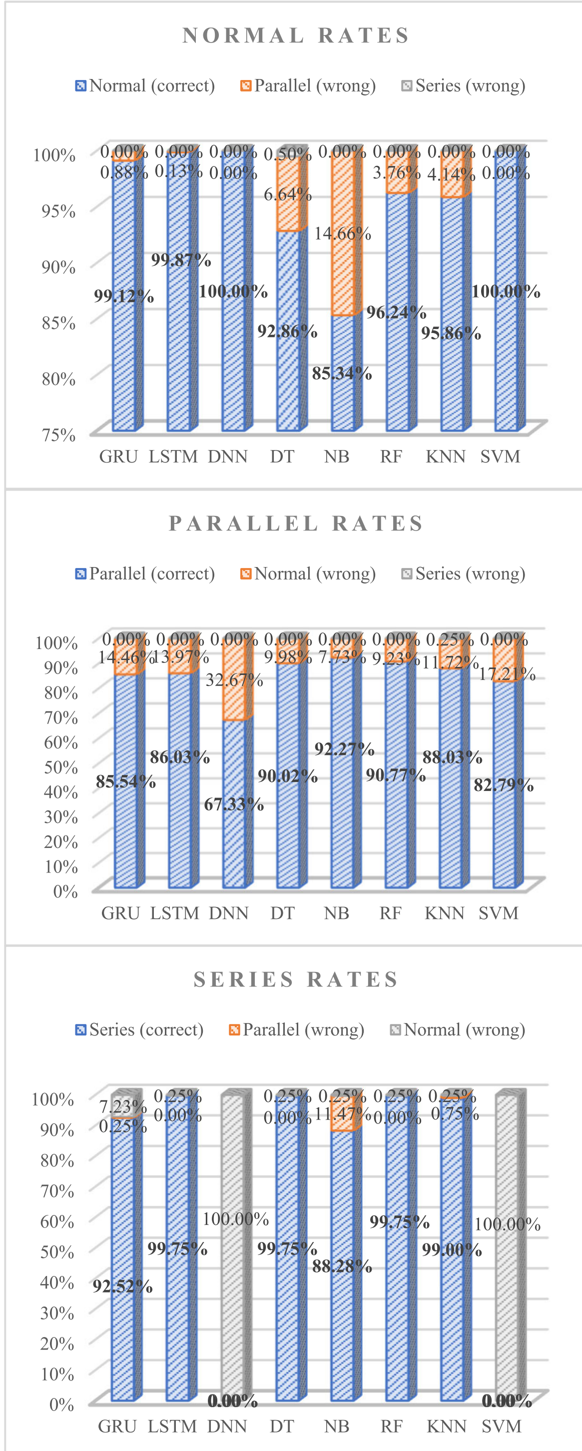

FIGURE 11. Identification performance of AI models in case 1 with the test pair of parallel arc at 5 A load current, 0.5 A arc current, and series arc at 5 A load current.

In the following, we report results obtained for a parallel arc with an increased arc current of 1 A; the load currents of parallel and series arc were equal to those used to generate the data shown in Figure 7. The identification accuracy rates of the eight learning techniques are presented in Figure 9. The RF, LSTM, KNN, and DT techniques were found to have the best diagnosis rates (above 96%). The performance of the NB algorithm was found to be mediocre, and the accuracies of the SVM and DNN techniques were the lowest of the techniques considered here. Next, we consider results

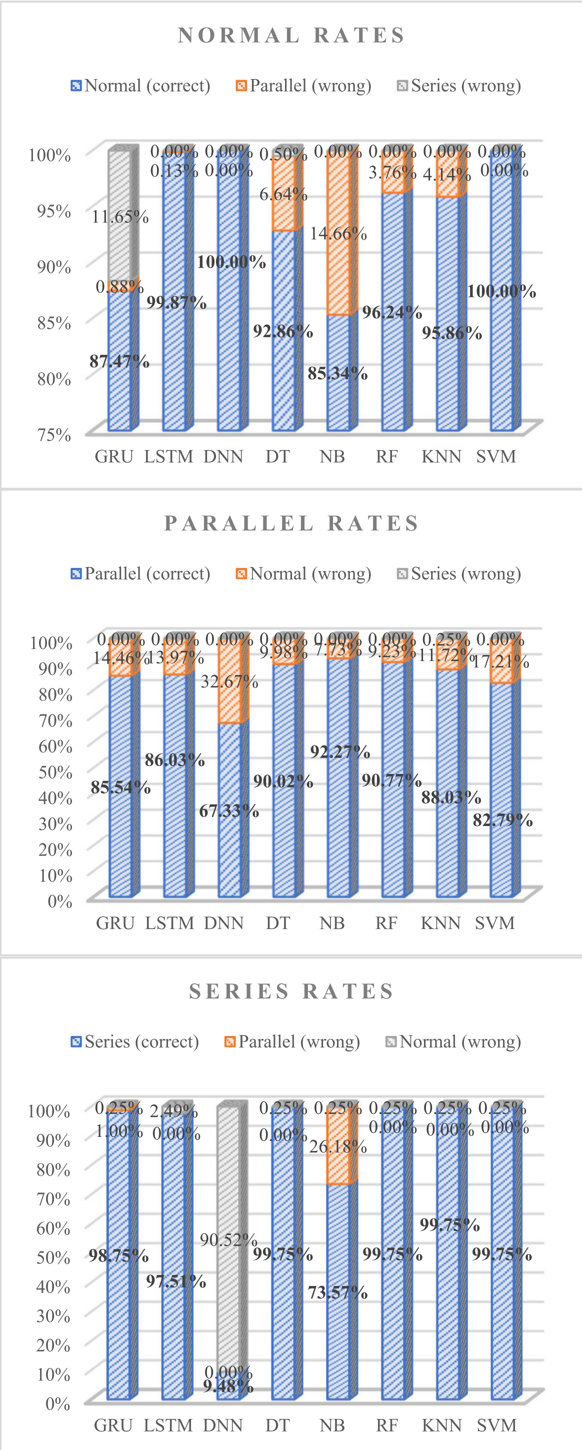

FIGURE 12. **Identification performance of AI models in case 1 with the test pair of parallel arc at 5 A load current, 0.5 A arc current, and series arc at 8 A load current.**

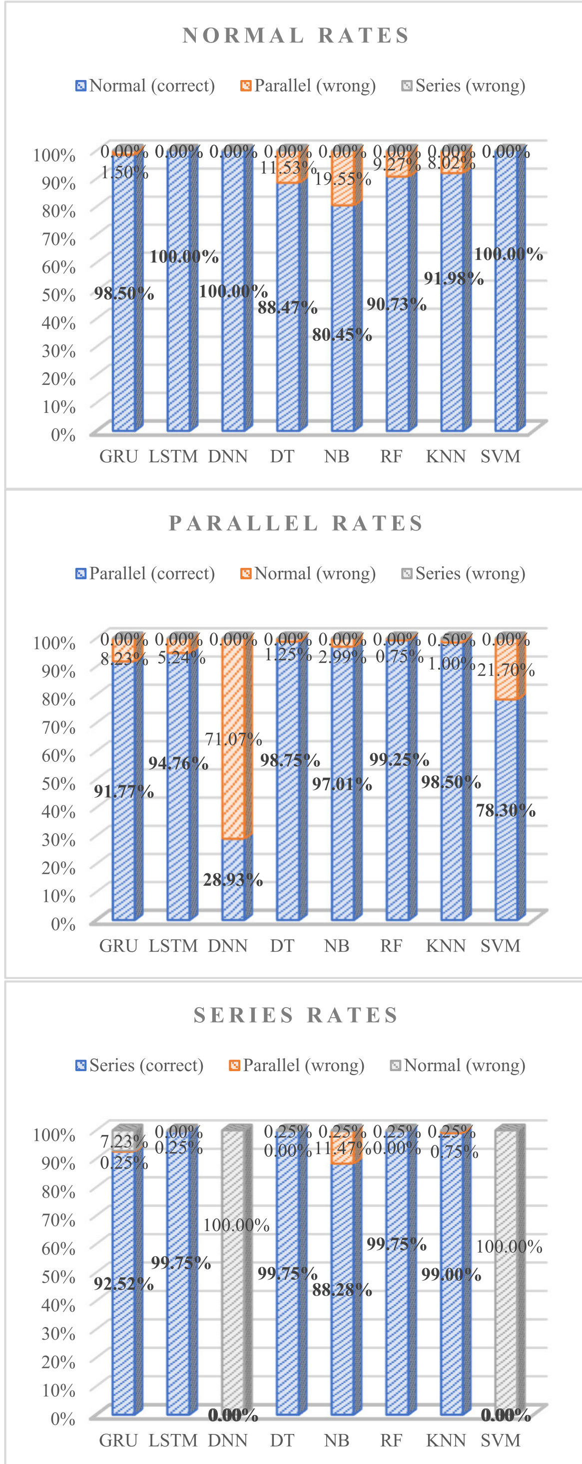

FIGURE 13. **Identification performance of AI models in case 1 with the test pair of parallel arc at 5 A load current, 1 A arc current, and series arc at 5 A load current.**

corresponding to a series arc with an increased load current of 8 A; the condition of the parallel arc was the same as that in Figure 9. The identifying accuracy rates of eight learning techniques are presented in Figure 10. The RF, KNN, LSTM, and DT algorithms were found to have excellent identification rates for all states, whereas the NB, SVM, and GRU algorithm identification performances were mediocre, and the identification accuracy of the DNN technique was the lowest of the techniques considered here. Figure 11 shows the average identification performance of the AI algorithms using the data of case 1 with the test pair comprising a

FIGURE 14. Identification performance of AI models in case 1 with the test pair of parallel arc at 5 A load current, 1 A arc current, and series arc at 8 A load current.

parallel arc with a 5 A load current and a 0.5 A arc current and series arc with a 5 A load current. The RF, DT, and KNN algorithms detected the states with the highest accuracy. The identification accuracies of the NB, LSTM, and GRU algorithms were also high. The identification accuracies of the DNN and SVM techniques were lowest for the detection of series arcing states. Figure 12 shows the identification accuracies of the techniques for a load current in the series arc of 8 A; the load and arc currents of the parallel arc used to generate the data in Figure 12 are equal to those considered in Figure 11. The identification accuracy rates of the eight learning techniques are presented in Figure 12. The RF and DT algorithms were found to have the best diagnosis rates. The identification accuracies of the KNN, LSTM, SVM, and GRU techniques were also high. The performance of the NB method was mediocre, and the accuracy of the DNN method was the lowest of the techniques considered here. For the generation of the data represented in Figure 13, the load currents of both arcs were 5 A, whereas the arc current of the parallel arc was increased to 1 A. The GRU, KNN, RF, and LSTM methods detected the states with the extraordinary accuracy rates (above 90 %). The DT and NB algorithms exhibited high and mediocre performances, respectively. The identifying accuracies obtained using the DNN and SVM methods were lowest for the detection of series arcing states. Figure 14 shows the average identification performance of the AI algorithms for case 1 with the test pair of parallel arcs with a 5 A load current and a 1 A arc current, and series arc with an 8 A load current. The RF, KNN, and LSTM algorithms show excellent identification rates for all states. The identification accuracies of the DT and GRU methods were also high, whereas the identification performances of the NB, SVM, and GRU algorithms were found to be mediocre. The identifying accuracies of the DNN algorithm were the lowest of the techniques considered here. Table 3 summarizes the average identification accuracies of all the AI techniques for data related to case 1. The SVM, DNN, and LSTM algorithms were found to be the best three techniques in terms of normal detection rates, and the LSTM algorithm also showed excellent identification rates for parallel and series arcs. By contrast, the SVM and DNN algorithms showed poor performance in the identification of other states. The RF, DT, and NB methods detected the parallel arcs with excellent accuracy. In addition, the RF and DT techniques could identify the normal and series arc states with low error rates, whereas the NB algorithm showed mediocre performance in identifying other states. The remaining techniques (KNN and GRU) detected the states with a high accuracy. Generally, using feature analysis, ML (RF, KNN, DT), and DL (GRU, LSTM) techniques were able to identify the states with high performance in the case of data corresponding to case 1. Table 4 shows the average identification precisions of all the AI models for data related to case 2. The RF, KNN, and LSTM techniques were found to be the best three techniques for detection of the normal state, and the LSTM algorithm also showed excellent identification rates of the arc faults. In addition, the RF technique showed high detection rates in the identification of the arc type, whereas the KNN algorithm showed poor and high performances in identifying parallel and series arcs, respectively. The DT and GRU algorithms detected the parallel arcs with high accuracy.

TABLE 3. The average identification rates of DC arc faults in case 1.

Detection rates		GRU	LSTM	DNN	DT	NB	RF	KNN	SVM
Normal detection rates	Normal (correct)	93%	98%	100%	93%	86%	96%	95%	100%
	Parallel (wrong)	1%	2%	0%	6%	14%	4%	5%	0%
	Series (wrong)	6%	0%	0%	0%	0%	0%	0%	0%
Parallel detection rates	Parallel (correct)	93%	95%	48%	97%	97%	97%	96%	81%
	Normal (wrong)	7%	5%	52%	3%	3%	3%	4%	19%
	Series (wrong)	0%	0%	0%	0%	0%	0%	0%	0%
Series detection rates	Series (correct)	96%	99%	5%	100%	81%	100%	99%	50%
	Parallel (wrong)	1%	0%	0%	0%	19%	0%	0%	0%
	Normal (wrong)	4%	1%	95%	0%	0%	0%	0%	50%

TABLE 4. The average identification rates of DC arc faults in case 2.

Detection rates		GRU	LSTM	DNN	DT	NB	RF	KNN	SVM
Normal detection rates	Normal (correct)	96%	100%	51%	90%	50%	98%	98%	92%
	Parallel (wrong)	3%	0%	12%	10%	6%	2%	2%	8%
	Series (wrong)	1%	0%	37%	0%	44%	0%	0%	0%
Parallel detection rates	Parallel (correct)	95%	98%	22%	92%	75%	91%	66%	27%
	Normal (wrong)	5%	1%	47%	5%	25%	9%	18%	72%
	Series (wrong)	0%	1%	31%	3%	0%	0%	16%	1%
Series detection rates	Series (correct)	98%	100%	50%	75%	99%	100%	100%	50%
	Parallel (wrong)	0%	0%	2%	24%	1%	0%	0%	0%
	Normal (wrong)	2%	0%	48%	0%	0%	0%	0%	50%

In addition, the GRU algorithm identified the normal and series arc states with a low error rate, whereas the DT method showed mediocre performance in identifying other states. The remaining techniques (SVM, NB, and DNN) detected the states with high rates. Generally, using the raw data, the performance of ML (SVM, KNN, NB, and DT) techniques were low, except for the RF algorithm; DL (GRU, LSTM) techniques, however, were able to identify the states with excellent accuracy. Table 5 illustrates the average identification accuracy of all the AI techniques using the data of case 3. The RF, NB, and LSTM techniques were found to be the best three techniques for the detection of normal states. The RF and LSTM algorithms also showed high identification rates for parallel and series arcs. By contrast, the NB technique showed poor performance when detecting the other states. The RF, DT, and GRU algorithms detected the parallel arc with high accuracy. In addition, the DT algorithm was able to identify normal and

TABLE 5. The average identification rates of DC arc faults in case 3.

Detection rates		GRU	LSTM	DNN	DT	NB	RF	KNN	SVM
Normal detection rates	Normal (correct)	**86%**	**98%**	**56%**	**94%**	**100%**	**98%**	**97%**	**92%**
	Parallel (wrong)	12%	1%	32%	6%	0%	2%	3%	7%
	Series (wrong)	2%	1%	13%	0%	0%	0%	0%	1%
Parallel detection rates	Parallel (correct)	**97%**	**96%**	**92%**	**96%**	**0%**	**96%**	**93%**	**84%**
	Normal (wrong)	3%	2%	6%	4%	100%	4%	6%	15%
	Series (wrong)	0%	4%	2%	0%	0%	0%	1%	1%
Series detection rates	Series (correct)	**97%**	**95%**	**17%**	**99%**	**0%**	**100%**	**100%**	**43%**
	Parallel (wrong)	0%	4%	4%	0%	0%	0%	0%	0%
	Normal (wrong)	3%	1%	78%	1%	100%	0%	0%	57%

series arc states with low errors, whereas the GRU technique showed mediocre and high performances in the identification of normal and series arc states, respectively. The remaining techniques (SVM, DNN) detected the states with high error rates with the exception of the KNN technique. Generally, when using both the feature and raw data, the accuracy of the ML (RF, KNN, and DT) and DL (LSTM and GRU) techniques were found to be high with the exception of the SVM, DNN, and NB methods, which showed mediocre and poor performance. This research focuses on identifying the types of arc faults. Thus, the comparison for detection time is omitted. However, the expected detection time of all AI models should be lower than conventional techniques because the AI techniques determine the arcing or normal events based on the training models, which were already trained. The training stage is time-consuming, but the decision stage only takes a few milliseconds.

V. CONCLUSION

We demonstrated the identification of DC arcs using different combinations of different input and advanced ML and DL algorithms. DL algorithms were found to be more effective when raw signal inputs were used compared with ML methods for the detection of any given state. On the other hand, the identification rates of ML methods were found to be higher than those obtained using DL techniques when feature analysis inputs were used. The combination of inputs yielded a balance in the proficiency of DL and ML methods for all states. In general, the RF and LSTM algorithms were found to be the are the best AI techniques to identify the normal or failed states in DC systems. These algorithms provided excellent performances with various input parameters. This research offers insight into the failure diagnosis of DC arcs. Nevertheless, this research was executed in a research laboratory and therefore appropriate testing and adaptation must be undertaken before applying these techniques to practical systems or applications. One of the limitations of this work is that this paper did not study the hyperparameter adjustments of the AI models. As a result, the same AI model might deliver altered precision grades for the same data set with different quantities of hyperparameters. The hyperparameters in all learning algorithms were chosen based on the trial and error method. Numerous examinations are mandatory to discover the finest hyperparameters. Nevertheless, there is no way to ensure that the selected hyperparameters return the best performance for all cases. Additionally, the performance of all AI algorithms varies with the operative conditions (load type, current amplitudes, and switching frequency). This means that the optimum hyperparameters in one specific condition are not optimal for others. From the above diagnosis rates, some learning techniques offer high performance in several cases, while their performance was mediocre or poor in other cases. Another limitation of this work is that it did not investigate feature analysis in the frequency domain. The reason for this omission is that such analysis entails high computational cost and sample rates. These shortcomings would extend the performance time and reduce the precision in the identification of a fault in actual systems. The time-domain signals could be handled using a lower sample rate than the frequency domain signals, which results in a fast algorithm for a given computational cost.

The above results indicate that machine learning techniques should use feature extractions in order to sustain a high diagnosis rate. Alternatively, DL methods need an enormous data set and high computational effort owing to deep hidden structure compared with ML techniques. The detection rates show that DL algorithms performed better than the other AI techniques when the raw data in the time domain was considered. By contrast, their performance was reduced when the raw signals were removed from the input data.

This study aims to classify the arc type when an arc event occurs. Because the protection scheme for each type of arc fault is different, if the type of fault is not classified, the proper protection acts may not be applied correctly. Furthermore, there are various parallel and series arc combinations in practical applications. Therefore, the training stage is vital when the new combinations are integrated into the system. For the practical systems, if there is any change, such as load types and current amplitudes, training new data is vital to maintain the high performance of AI models. Otherwise, the detection rates might be poor or false detections could be made. This study provides insight into different AI methods. This research may be of interest when selecting AI techniques, input types, and feature extraction methods; this work may therefore contribute to constructing safer and more stable systems involving arc fault recognition schemes related to altered priorities, such as robustness, reliability, execution time, and costs.

REFERENCES

[1] W. P. Mendenhall, "DC arc hazard mitigation design at a nuclear research facility," in *Proc. IEEE IAS Electr. Saf. Workshop*, Feb. 2014, pp. 1–4.

[2] J. Kim, S. Kwak, and S. Choi, "DC series arc detection algorithm based on adaptive moving average technique," *IEEE Access*, vol. 9, pp. 94426–94437, 2021.

[3] J.-C. Kim and S.-S. Kwak, "Frequency-domain characteristics of series DC arcs in photovoltaic systems with voltage-source inverters," *Appl. Sci.*, vol. 10, no. 22, p. 8042, Nov. 2020.

[4] R. Grassetti, R. Ottoboni, M. Rossi, S. J. I. I. Toscani, and M. Magazine, "A low cost arc fault detection in aerospace applications," *IEEE Instrum. Meas. Mag.*, vol. 16, no. 5, pp. 37–42, Oct. 2013.

[5] M. Naidu, T. J. Schoepf, and S. Gopalakrishnan, "Arc fault detection scheme for 42-V automotive DC networks using current shunt," *IEEE Trans. Power Electron.*, vol. 21, no. 3, pp. 633–639, May 2006.

[6] Q. Xiong, X. Feng, A. L. Gattozzi, X. Liu, L. Zheng, L. Zhu, S. Ji, and R. E. Hebner, "Series arc fault detection and localization in DC distribution system," *IEEE Trans. Instrum. Meas.*, vol. 69, no. 1, pp. 122–134, Jan. 2019.

[7] J.-C. Gu, D.-S. Lai, J.-M. Wang, J.-J. Huang, and M.-T. Yang, "Design of a DC series arc fault detector for photovoltaic system protection," *IEEE Trans. Ind. Appl.*, vol. 55, no. 3, pp. 2464–2471, May 2019.

[8] M. Ahmadi, H. Samet, and T. Ghanbari, "A new method for detecting series arc fault in photovoltaic systems based on the blind-source separation," *IEEE Trans. Ind. Electron.*, vol. 67, no. 6, pp. 5041–5049, Jun. 2020.

[9] R. Jiang and Y. Zheng, "Series arc fault detection using regular signals and time-series reconstruction," *IEEE Trans. Ind. Electron.*, early access, Apr. 12, 2022, doi: 1109/TIE.2022.3165260.

[10] A. Amiri, H. Samet, and T. Ghanbari, "Recurrence plots based method for detecting series arc faults in photovoltaic systems," *IEEE Trans. Ind. Electron.*, vol. 69, no. 6, pp. 6308–6315, Jun. 2022.

[11] R. Jiang, G. Bao, Q. Hong, and C. D. Booth, "A coupling method for identifying arc faults based on short-observation-window SVDR," *IEEE Trans. Instrum. Meas.*, vol. 70, pp. 1–10, 2021.

[12] C. He, L. Mu, and Y. Wang, "The detection of parallel arc fault in photovoltaic systems based on a mixed criterion," *IEEE J. Photovolt.*, vol. 7, no. 6, pp. 1717–1724, Nov. 2017.

[13] M. Murtadho, E. Prasetyono, and D. O. Anggriawan, "Detection of parallel arc fault on photovoltaic system based on fast Fourier transform," in *Proc. Int. Electron. Symp. (IES)*, Sep. 2020, pp. 21–25.

[14] R. Grassetti, R. Ottoboni, and M. Rossi, "A novel algorithm for the parallel arc fault identification in DC aircraft power plants," in *Proc. IEEE Int. Instrum. Meas. Technol. Conf.*, May 2012, pp. 148–153.

[15] C. Strobl and P. Meckler, "Arc faults in photovoltaic systems," in *Proc. 56th IEEE Holm Conf. Electr. Contacts*, Oct. 2010, pp. 1–7.

[16] J. Johnson, "Differentiating series and parallel photovoltaic arc-faults," in *Proc. 38th IEEE Photovoltaic Spec. Conf.*, Jun. 2012, pp. 720–726.

[17] K. Yang, R. Zhang, J. Yang, C. Liu, S. Chen, and F. Zhang, "A novel arc fault detector for early detection of electrical fires," *Sensors*, vol. 16, no. 4, p. 500, 2016.

[18] Y. Wang, F. Zhang, and S. Zhang, "A new methodology for identifying arc fault by sparse representation and neural network," *IEEE Trans. Instrum. Meas.*, vol. 67, no. 11, pp. 2526–2537, Nov. 2018.

[19] H.-L. Dang, J. Kim, S. Kwak, and S. Choi, "Series DC arc fault detection using machine learning algorithms," *IEEE Access*, vol. 9, pp. 133346–133364, 2021.

[20] H.-L. Dang, S. Kwak, and S. Choi, "Different domains based machine and deep learning diagnosis for DC series arc failure," *IEEE Access*, vol. 9, pp. 166249–166261, 2021.

[21] J.-Y. Jeong, J.-C. Kim, and S. Kwak, "DC series arc diagnosis based on deep-learning algorithm with frequency-domain characteristics," *J. Power Electron.*, vol. 21, pp. 1900–1909, Nov. 2021.

[22] R. Jiang, G. Bao, Q. Hong, and C. Booth, "Machine learning approach to detect arc faults based on regular coupling features," *IEEE Trans. Ind. Informat.*, early access, Feb. 23, 2022, doi: 10.1109/TII.2022.3153333.

[23] H.-L. Dang, S. Kwak, and S. Choi, "Parallel DC arc failure detecting methods based on artificial intelligent techniques," *IEEE Access*, vol. 10, pp. 26058–26067, 2022.

[24] S. Lu, A. Sahoo, R. Ma, and B. T. Phung, "DC series arc fault detection using machine learning in photovoltaic systems: Recent developments and challenges," in *Proc. 8th Int. Conf. Condition Monitor. Diagnosis (CMD)*, Oct. 2020, pp. 416–421.

[25] K. Xia, B. Liu, X. Fu, H. Guo, S. He, W. Yu, J. Xu, and H. Dong, "Wavelet entropy analysis and machine learning classification model of DC serial arc fault in electric vehicle power system," *IET Power Electron.*, vol. 12, no. 15, pp. 3998–4004, Dec. 2019.

[26] W. Miao, Q. Xu, K. H. Lam, P. W. T. Pong, and H. V. Poor, "DC arc-fault detection based on empirical mode decomposition of arc signatures and support vector machine," *IEEE Sensors J.*, vol. 21, no. 5, pp. 7024–7033, Mar. 2021.

[27] S. Lu, T. Sirojan, B. T. 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.

[28] X. Yao, L. Herrera, S. Ji, K. Zou, and J. Wang, "Characteristic study and time-domain discrete-wavelet-transform based hybrid detection of series DC arc faults," *IEEE Trans. Power Electron.*, vol. 29, no. 6, pp. 3103–3115, Jun. 2014.

[29] K. Xia, H. Guo, S. He, W. Yu, J. Xu, and H. Dong, "Binary classification model based on machine learning algorithm for the DC serial arc detection in electric vehicle battery system," *IET Power Electron.*, vol. 12, no. 1, pp. 112–119, Jan. 2019.

[30] B. E. Boser, I. M. Guyon, and V. N. Vapnik, "A training algorithm for optimal margin classifiers," in *Proc. 5th Annu. Workshop Comput. Learn. Theory (COLT)*, 1992, pp. 144–152.

[31] T. Cover and P. Hart, "Nearest neighbor pattern classification," *IEEE Trans. Inf. Theory*, vol. IT-13, no. 1, pp. 21–27, Jan. 1967.

[32] L. Breiman, J. H. Friedman, R. A. Olshen, and C. J. Stone, *Classification and Regression Trees*. Evanston, IL, USA: Routledge, 2017.

[33] L. Breiman, "Random forests," *Machine Learn.*, vol. 45, no. 1, pp. 5–32, 2001.

[34] P. Langley, W. Iba, and K. Thompson, "An analysis of Bayesian classifiers," in *Proc. AAAI*, vol. 90. Princeton, NJ, USA: Citeseer, 1992, pp. 223–228.
[35] C.-J. Park, H.-L. Dang, S. Kwak, and S. Choi, "Deep learning-based series AC arc detection algorithms," *J. Power Electron.*, vol. 21, no. 10, pp. 1621–1631, Oct. 2021.
[36] C.-J. Park, H.-L. Dang, S. Kwak, and S. Choi, "Detection algorithms of parallel arc fault on AC power lines based on deep learning techniques," *J. Electr. Eng. Technol.*, vol. 17, pp. 1195–1202, Jan. 2022.

HOANG-LONG DANG received the B.S. degree in electrical and electronics engineering from the Ho Chi Minh City University of Technology, Vietnam, in 2015. He is currently pursuing the combined M.S. and Ph.D. degrees in electrical and electronics engineering with Chung-Ang University, Seoul, South Korea. His research interests include matrix converters, fault detections, and artificial intelligences.

SANGSHIN KWAK (Member, IEEE) received the Ph.D. degree in electrical engineering from Texas A&M University, College Station, TX, USA, in 2005. From 2007 to 2010, he was an Assistant Professor at Daegu University, Gyeongsan, South Korea. Since 2010, he has been working at Chung-Ang University, Seoul, South Korea, where he is currently a Professor. His current research interests include the design, modeling, control, and analysis of power converters for electric vehicles and renewable energy systems as well as the prognosis and fault tolerant control of power electronics systems.

SEUNGDEOG CHOI (Senior Member, IEEE) received the B.S. degree in electrical and computer engineering from Chung-Ang University, Seoul, South Korea, in 2004, the M.S. degree in electrical and computer engineering from Seoul National University, Seoul, in 2006, and the Ph.D. degree in electric power and power electronics from Texas A&M University, College Station, TX, USA, in 2010. From 2006 to 2007, he was a Research Engineer at LG Electronics, Seoul. From 2009 to 2012, he was a Research Engineer at Toshiba International Corporation, Houston, TX, USA. From 2012 to 2018, he was an Assistant Professor at The University of Akron, Akron, OH, USA. Since 2018, he has been working as an Associate Professor at Mississippi State University, Starkville, MS, USA. His current research interests include degradation modeling, fault tolerant control, and fault tolerant design of electric machines, power electronics, batteries, solar panels, and wider vehicular/aircraft microgrid systems.