Parallel DC Arc Failure Detecting Methods Based on Artificial Intelligent Techniques

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)

ABSTRACT The unwanted electric discharge usually relates to arc phenomena between two connectors. The energy from an arc might fuse the electric wiring and be responsible for a fire. Various researches have been investigated for safety operations to improve detected techniques for arc diagnosis. There are two types of arc faults: parallel and series arcs. A parallel arc happens among two electrical lines, or line and ground, due to degrading insulation or contamination. On the other hand, a series arc might result from releasing connections in the wiring. The system's current can be significantly increased by parallel arc fault compared with the series arc. In this work, the electrical behavior of the system is investigated during parallel arc faults to understand the arcing characteristics from different cases, identify electrical characteristics that are useful and reliable for the diagnosis process, and determine the location of the fault based on current or voltage of the faulted system. Eight learning techniques are adopted to detect arc fault in this study. Parallel arc signals were analyzed in the time and frequency domains, and unique characteristics of the current are extracted using Fourier analysis as an indicator for diagnosing an arc fault. This research can be used to improve arc-fault detector reliability and robustness.

INDEX TERMS Artificial intelligence, fault diagnosis, DC parallel arc.

I. INTRODUCTION

DC networks are widely used in aerospace, photovoltaic systems, data storage, electric transportation, and various areas. However, the increasing applications of DC networks will certainly create more and more potential failures. Arc failure is a dangerous event that cannot be ignored in individual local power networks. Utility the DC systems as the source should pay attention to prevent the system from failure, especially arc faults. The arc failure is sorted as parallel and series in the DC network. A series arc might result from releasing connections in the wiring [1]. A temporary short circuit usually causes this type of arc fault. Possible reasons for series arc are loosening busbars or cable connections with poor contact. The connected loads limit the series arc current. If the arc current is in the safety device rated current range, the failure might not be identified in time. Otherwise, if the magnitude of arc current

is two to five times of rated current of safety devices, the arc would be burned in a too long time before the protection device separates the fault from the operating network. On the other hand, a parallel arc is an event of discharge situation between two points with voltage differences. It can occur owing to wrecked or scratched insulation [2]. Therefore, there is a chance to burn with the arc current lower than the rated current of protection devices for the parallel arc. The parallel arc fault can be more dangerous than the series arc because it can increase the current in the system. The rise of arc current and the increasing heat for the period of the parallel arc fault could make the flare to be larger and scorching, with the major harms being destroying and vanishing the conductors and wirings [3]. Furthermore, it could lead to physical losses worse than the series ones.

Research on parallel arc fault detection in the DC system is still in the initial stage, and a complete protection scheme has not been formed. Nevertheless, many pieces of literature have demonstrated how to detect the arc failure

event. The differences of the parallel and series arcs in the DC system is illustrated in [4]. The frequency- and time-domain characteristics of the arc failure in DC systems are illustrated in [5]. The large fluctuations of the current are adopted to detect the parallel arc event in [6]. Moreover, one of many works of has utilized the PSIM program to simulate and explain the arc event in [7].

The evolution of advanced techniques is gained more consideration from scientists due to its flexible capability in different applications. Artificial intelligence (AI) or Advanced learning techniques have been successfully employed in various areas. They provide powerful approaches for identifying failure in different applications, such as failure detection in electrical machines [8], fault diagnosis in medium voltage networks based on high impedance [9], and fault diagnosis of track circuits in railway systems [10]. Scientists have successfully adopted these learning techniques for detecting arc failure and reached promising outcomes, such as the combination of the wavelet packet decomposition (WPD) and support vector machine (SVM) algorithm in DC system [11], and the use of a cascaded fuzzy logic system for series arc diagnosis [12]. Several characteristics, such as the high-frequency signals and current variations, are extracted and adapted for training models centered on weighted least squares SVM algorithms to diagnose series arc [13]. In addition, an attractor matrix, which is constructed from singular value decomposition and current signals, is employed to obtain features in [14]. The combination of an artificial neural network and sparse coding characteristics for arc detection was proposed in [15]. The adoption of a neural network for arc failure detection was presented in [16]. In [17] and [18], several AI algorithms were adopted to detect DC series arc fault. In addition, five features in the time domain were utilized as inputs of learning algorithms. These features were chosen to prevent the overlapped feature between arcing and normal states. Studies illustrate comparisons of performance between various AI methods in DC systems in [19]. On the other hand, a short-observation-window singular value decomposition and reconstruction algorithm are proposed to identify AC series arc fault [20]. Although this method guarantees a high diagnosis rate with different load types, the complexity and the need for additional hardware are the limitations of this proposed method. Generally, these studies focus only on series arc fault, whereas the application of AI for the parallel arc is not thoroughly investigated. Therefore, there is a need for a study with various operating conditions for a parallel arc fault.

In this paper, eight AI algorithms have been implemented to detect the parallel arc event and compare the performance between techniques. Furthermore, finding the proper input for the best result because the arc current of parallel arc is not measurable directly like series arc fault. Comparing and discussing the performance of AI techniques and input combinations in both frequency and time domains is presented. This paper is organized as follows. Section 2 specifics the arc-generation hardware and the current phenomena in arcing
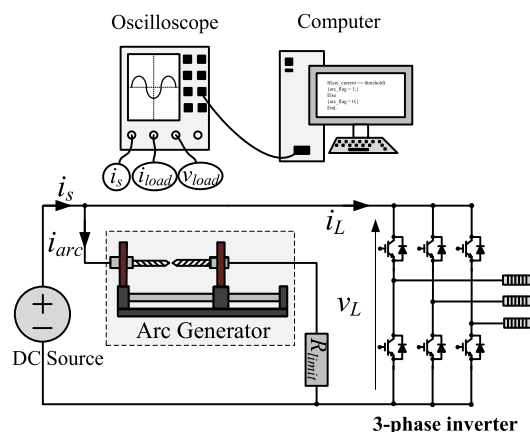

FIGURE 1. Parallel DC arc setup.

and normal states. Section 3 describes the learning techniques used for arc diagnosis, and research analyzes input extractions. Section 4 compares detection performances using different combinations from six features parameters and eight learning techniques when an arcing event occurs in altered operating conditions. In conclusion, the discussion and recommendation of arc failure detection regarding the diagnosis rates of various combinations between feature parameters and advanced learning techniques are desmonstrated in Section 5.

II. CHARACTERISTICS OF PARALLEL DC ARC

The arc generator and experimental hardware were configured to collect arc data, as shown in Figure 1 by referring to UL1699B standard. DC source represents the DC supply voltage, and its magnitude in the arc experiment is 300V. The DC supply used in the experiment is KEYSIGHT N8741A (maximum voltage 300V, maximum current 11A, maximum power 3.3kW). i_{arc} is the arc current passing the arc rods. The step motor separates the arc rods to generate arc events safely. In addition, the separated gap of the rods is checked with an electric ruler installed parallel to the rods. A resistance and an inductance of 10 Ω and 10 mH are used as the loads for the three-phase inverter. The experimental setup consists of a DC supply source, arc generator, and load (three-phase inverter)[21]. The resistor R_{limit} is inserted in series in the arc generator to limit the arc current for safety because when generating a parallel arc using an arc generator, the amount of source current i_s increases rapidly. Table 1 shows the specifications for the parallel arc fault.

First, a DC voltage is applied to drive the inverter load. Then, the arc rods are separated by the step motor connected to the rods with a speed of 2.5 mm/s. The arc current before and after the separation of arc rods is collected by an oscilloscope with a sampling frequency of 250 kHz. When the arc is initiated, the added arc current noise results in large fluctuations. Using the collected data, the diagnosis process is executed by using MATLAB. The selection of sampling frequency is based on several recent research about arc fault in DC systems [22]–[26]. Using a higher sampling
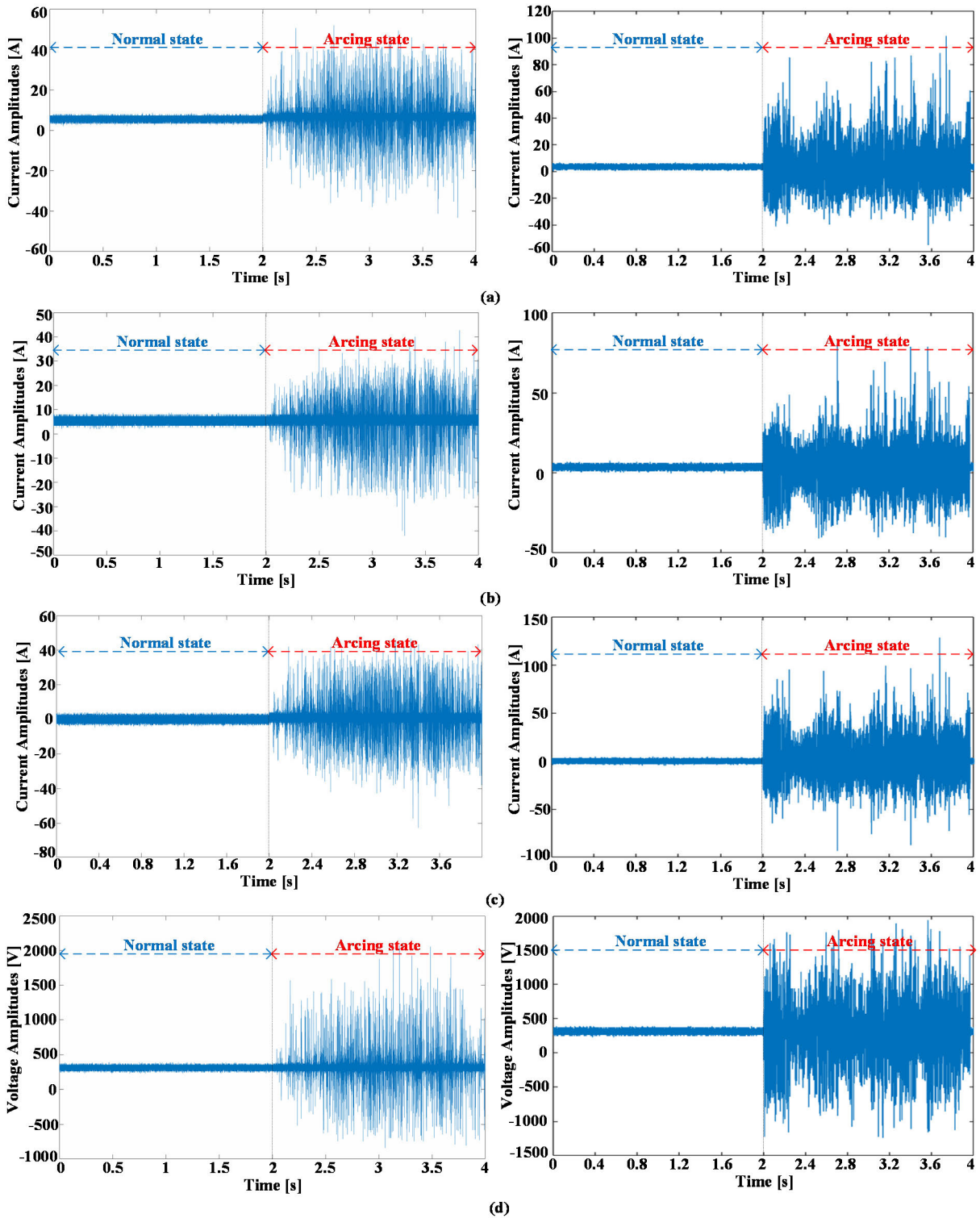
FIGURE 2. The normal state and the arcing state waveforms for 5 A load current, 1 A arc current at 15 kHz (left) and 3 A load current, 0.5 A arc current at 5 kHz (right). (a) Source current, (b) load current, (c) arc current, and (d) load voltage.

frequency oscilloscope could result in more information in each signal. However, it could increase the execution time and computation burden, whereas one of the most priorities

is detecting arc fault in time to separate the fault from the system. Therefore, the sampling rate of 250 kHz is high enough to balance the efficiency and execution time. Similar

TABLE 1. Specifications of loads used in the experiment.

Load type	Modulation technique	Load current (i_L)	Limit resistor (R_{limit})	Arc current (i_{arc})	Switching frequency
Three-phase inverter	Space vector modulation	3 A	600 Ω	0.5 A	5 kHz
				0.5 A	15 kHz
			300 Ω	1 A	5 kHz
				1 A	15 kHz
		5 A	600 Ω	0.5 A	5 kHz
				0.5 A	15 kHz
			300 Ω	1 A	5 kHz
				1 A	15 kHz

FIGURE 3. Principles and structures of artificial intelligence algorithms.

TABLE 2. Hidden stuctures of deep learning techniques for DC parallel arc diagnosis.

Structures	1st layer/ neurons	2nd layer/ neurons	3rd layer/ neurons	4th layer/ neurons	5th layer/ neurons
DNN	FC/100	FC/50	FC/2	N/A	N/A
LSTM	FC/100	LSTM/16	FC/8	LSTM/8	FC/2
GRU	FC/100	GRU/16	FC/8	GRU/8	FC/2

to the window length, the longer window results in more information, which could increase the diagnosis accuracy. However, the increasing amount of information could affect the processing speed and calculation resources. Therefore, the selected window length is a 2 ms period. The collected signals were divided into different sets of 2 ms periods for testing and training processes using AI techniques. This study employed space vector modulation (SVPWM) to control the three-phase inverter. The objective of SVPWM was to utility the desired voltage and modulated switches to imitate the three-phase waveforms with sinusoidal form, whereas amplitude and frequency were designable. Figure 2 shows the experimental arcing and normal state signals at various setting conditions (3 and 5 A load current, 0.5 and 1 A arc current, 5 and 15 kHz switching frequency). As shown in the figure, all the forms of signals or the waveform shapes were stable and similar before arcing points. However, when a failure event was initiated, numerous abnormal behaviors were added to the signals, such as harmonic components in the arc signals, the distortion of the waveforms, and the fluctuations in the signal amplitude. This led to the generation of large negative fluctuations in the observed signals. The abovementioned unusual activities could be useful and potentially adopted for diagnosing the fault event. The arc current is obtained by using the relationship between the source and load currents. In practical systems, the measurement of arc current is not possible because the location of an arc event is unknown. Therefore, the arc current is not used for the fault diagnosis process in this study. It is only used for setting the working conditions in each case of the experiment.

III. ADVANCED LEARNING ALGORITHMS

A. ADVANCED LEARNING ALGORITHMS STRUCTURES

Figure 3 illustrates the principle and structure of various advanced algorithms. The objective of the support vector machine (SVM) algorithm is to locate a hyperplane with the largest margin. Then, this hyperplane is used to classify the data of one class from another class [27]. The K-Nearest neighbor (KNN) technique assumes that similar things locate in neighboring closeness. In another way, similar things are close to their family [28]. A decision tree (DT) can be used for classification as well as regression problems. The name itself
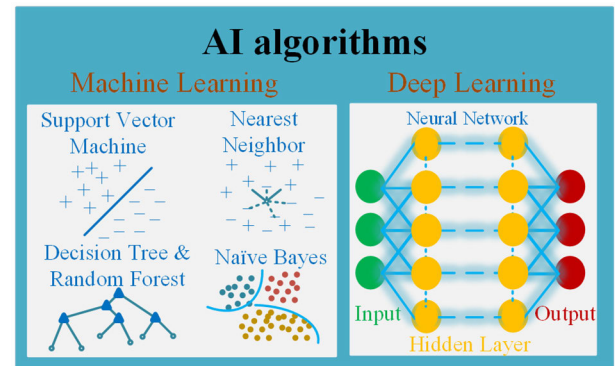
suggests that it uses a flowchart like a tree structure to show the predictions that result from a series of feature-based splits. It starts with a root node and ends with a decision made by leaves [29]. Random forest (RF) consists of many individual decision trees that operate as an ensemble. Each tree in the random forest returns an independent prediction, and the class with the most votes becomes the model's prediction [30]. Naïve Bayes (NB) classifiers are classification algorithms based on Bayes' theorem. It is a family of algorithms that share a common principle. That every pair of features being classified is independent of each other [31]. Unlike machine learning, deep learning (DL) teaches computers to do what comes naturally to humans: learn by example. In deep learning, a computer model performs classification tasks directly from images, text, or sound. Deep learning models can achieve high accuracy, sometimes exceeding human ability. Models are trained using a large set of labeled data and neural network architectures containing many layers, such as input, hidden, and output layers. Each layer contains various neurons; the output of one neuron in the nth layer is the input of another neuron in the $n+1^{th}$ layer [32]. The hidden configurations of DL algorithms (deep neural network (DNN), gated recurrent unit (GRU), and long-short term memory (LSTM)) are listed in Table 2.

There are three layers in deep neural network, and five layers in LSTM and GRU structures. The DNN's layers are fully connected (FC) type, and the neuron amounts of layers 1, 2, and 3 were 100, 50, and 2, respectively. The numeral neurons in GRU and LSTM layers 1, 2, 3, 4, and 5 were 100, 16, 8, 8,
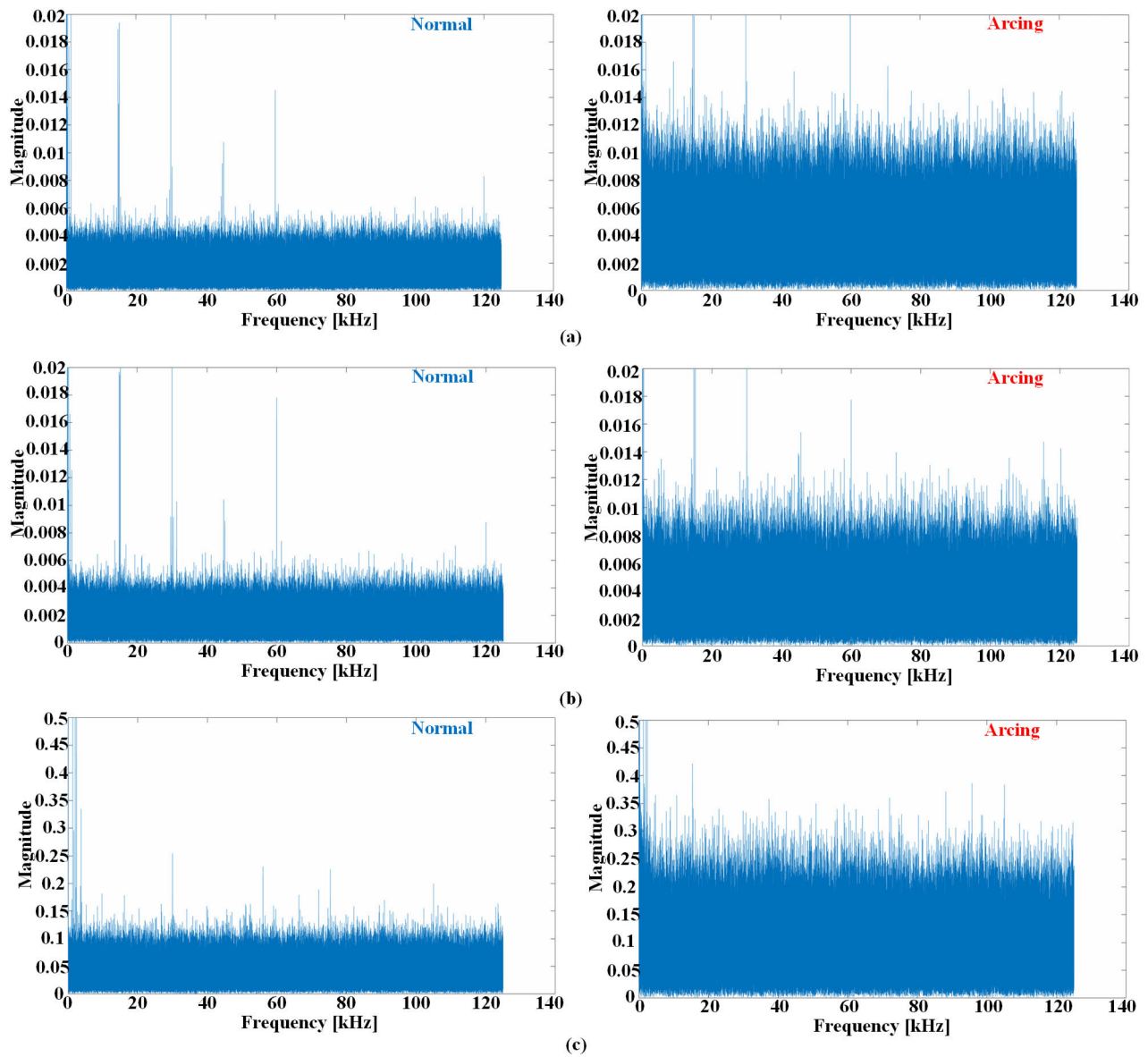
FIGURE 4. FFT analysis. (a) Source current, (b) load current, and (c) load voltage.

and 2, respectively. The difference between GRU and LSTM structures is that the second and fourth layers of the GRU structure are the GRU type, whereas they are LSTM type in the LSTM structure. The trial and error method chooses the specifications of hidden layers. The selected layer structures of DL methods provided the highest effectiveness among numerous configurations. However, there are other possible structures, which are also appropriate.

B. INPUT ANALYSIS
Features from the signals could be extracted by using several techniques, such as wavelet transformation and fast Fourier transform (FFT). Figure 4 shows the FFT analysis of source current, load current, and load voltage for 5 A load current,

1 A arc current for 15 kHz switching frequency. The larger harmonic components around 15, 30, 45, 60 kHz are the harmonics content concentrated around multiples of the switching frequency (15 kHz) owing to the utilization of SVPWM for controlling the three-phase inverter load. This technique utilizes three switching vectors in one sampling frequency. Therefore, the switching frequency is constant and results in the larger harmonics content around multiples of switching frequency. In addition, the parallel arc tends to increase the current when it occurs. Therefore, the harmonics in the arcing state are much larger than that of in the normal state. As shown in the figure, the high order harmonics are added in the signals after the arc event is initial. However, these feature analyses belong to the frequency domain, and their

TABLE 3. Possible cases of input combinations for DC parallel arc fault diagnosis.

Signals		Cases								
		1	2	3	4	5	6	7	8	9
Time domain	Load voltage ($v_{load}(t)$)	■	⊗	■	■	⊗	■	⊗	⊗	■
	Load current ($i_{load}(t)$)	■	⊗	■	⊗	■	⊗	■	⊗	■
	Source current ($is(t)$)	■	⊗	⊗	■	■	⊗	⊗	■	■
Frequency domain	Load voltage ($v_{load}(fft)$)	⊗	■	⊗	⊗	■	⊗	■	■	■
	Load current ($i_{load}(fft)$)	⊗	■	⊗	■	⊗	■	⊗	■	■
	Source current ($is(fft)$)	⊗	■	■	⊗	⊗	■	■	⊗	■

■: data used as input of AI algorithms; ⊗: data not used for AI algorithms.
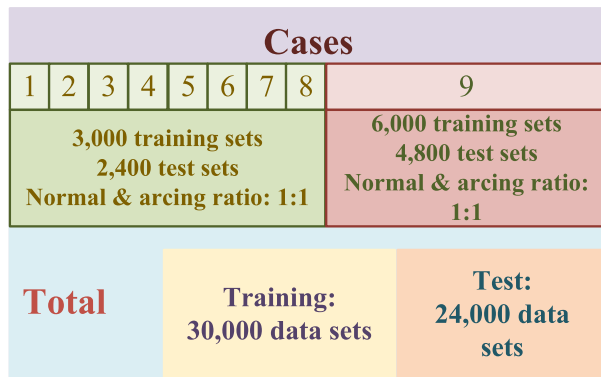

FIGURE 5. Proportion of test and training data.

analysis needs high computational resources and sampling rates. These disadvantages could suspend execution time and disturb precision when a failure happens in real applications. On the other hand, the signal in the time domain could be processed with a low sampling rate, which offers a fast computation effort. This study utilized both time- and frequency-domain inputs for parallel arc diagnosis. First, the signals were collected at a sampling rate of 250 kHz. Next, the collected signals were divided into different sets of 2 ms periods for testing and training processes. The FFT technique is adopted for each set of data to obtain the frequency-domain feature. After that, time and frequency signals were used as input for eight advanced algorithms to diagnosis the parallel arc event. Table 3 presents the possible cases of different combinations between time and frequency signals. Each case has at least one load voltage, one load current, and one source current, whether they belong to time or frequency domains. In the last case, all time and frequency domains signals are employed as inputs for learning algorithms.

IV. THE PERFORMANCE OF ADVANCED LEARNING ALGORITHMS IN PARALLEL ARC DIAGNOSIS

The proportion of test and training data is illustrated in Figure 5. There are nine possible cases for time and frequency domain combination. There are 3,000 training sets and 2,400 test sets in each case, from cases one to eight. In case nine, all time and frequency signals are employed as inputs of AI techniques; thus, there are 6,000 training sets and 4,800 test sets in this case. Totally, 30,000 data sets are used for training, and 24,000 data sets are used for the test process. The proportion of arcing and normal sets is 1:1 in all cases. The accuracy metric is adopted to value the effectiveness of the eight advanced algorithms. The accuracy detection rate is the proportion of correctly predicted data sets to the total number of test data sets. It is expressed as

$$\% \text{ of Total Acc.} = \frac{\# \text{ of correctly predicted data sets}}{\text{total } \# \text{ of test data sets}}. \quad (1)$$

The best advanced algorithm is the algorithm with the highest accuracy.

Figure 6 demonstrates the performances of AI algorithms in case 1. When the load current and the arc current are set at 3 and 0.5 A, respectively, RF and DNN hit the maximum diagnosis rate (100%) for both 5 and 15 kHz switching frequency. SVM and the other two DL techniques also show superior performances. Their accuracies are above 97.75 % at 5 kHz and higher at 15 kHz. The performances of KNN and DT are high (about 90-95%). The accuracies of NB are lowest even though its performance is improved when increasing switching frequencies. The load current has remained the same for the following condition, whereas the arc current is increased to 1 A. RF, LSTM, and GRU detect the arc event with the highest rates (above 97.5 %) compared with other learning techniques for both 5 and 15 kHz switching frequencies. The diagnosis accuracy of DNN is also high (around 96%) and increases with the increase of switching frequency. SVM and KNN show mediocre performance; the accuracies of DT and NB are lowest at 5 and 15 kHz, respectively. In the next condition, the load current is increased to 5 A, and the arc current is reduced to 0.5 A. RF, NB, and GRU show the best diagnosis rates (above 99%) at 5 and 15 kHz; LSTM also shows superior performances. DNN, KNN, and SVM show high performance at 5 kHz; however, their accuracy decreases significantly when the switching frequency increases. The performance of DT is mediocre, and its accuracies are lowest compared with other techniques. The accuracy of all techniques declines with the rise of the switching frequency in this condition. Next, the arc current is increased to 1 A. NB, KNN, and GRU show superior detected rates at 5 kHz, whereas their detected rates are declined at 15 kHz. There are two trends in this condition, the accuracies of AI techniques such as GRU, NB, KNN, and RF declined when the switching frequency increased, whereas the detected rates of SVM, DT, and DNN rise with the increase of switching rate. However, the increased accuracies of SVM, DT, DNN are still lower than the declined accuracies of other learning techniques. Table 4 presents the average accuracies of all AI techniques in case 1. The average accuracy of each learning technique considers all the mentioned working conditions. The best three diagnosis rates are highlighted for

TABLE 4. The comparison of diagnosis rates in case one for DC parallel arc fault.

AI algorithms	Operating conditions								
	3 A load current & 0.5 A arc current at 5 kHz	3 A load current & 0.5 A arc current at 15 kHz	3 A load current & 1 A arc current at 5 kHz	3 A load current & 1 A arc current at 15 kHz	5 A load current & 0.5 A arc current at 5 kHz	5 A load current & 0.5 A arc current at 15 kHz	5 A load current & 1 A arc current at 5 kHz	5 A load current & 1 A arc current at 15 kHz	Average accuracy
SVM	**99.75 %**	**99.8 %**	87.1 %	91.15 %	96.5 %	88 %	67 %	72.8 %	88 %
KNN	90.85 %	94.85 %	90.6 %	82.5 %	98.9 %	96.35 %	**99.15 %**	89 %	93 %
RF	**100 %**	**100 %**	**98 %**	**100 %**	**99.9 %**	**99.9 %**	97.3 %	**93.05 %**	**99 %**
NB	59.35 %	64.5 %	74 %	78.25 %	**100 %**	**100 %**	**100 %**	**93.05 %**	84 %
DT	92.9 %	90.65 %	71.35 %	81.6 %	83.85 %	81.85 %	65.2 %	73.55 %	80 %
DNN	**100 %**	**100 %**	95.75 %	97.65 %	98.4 %	88.85 %	61.5 %	79.15 %	90 %
LSTM	97.75 %	99 %	**97.5 %**	**98.15 %**	98 %	**99.9 %**	92.15 %	**92.75 %**	**97 %**
GRU	**99.75 %**	98.75 %	**99.15 %**	**99.9 %**	**100 %**	**99.25 %**	**99.9 %**	92.4 %	**99 %**

TABLE 5. The performance summary of nine cases for DC parallel arc fault diagnosis.

AI algorithms	Cases								
	1	2	3	4	5	6	7	8	9
SVM	88 %	92 %	90 %	89 %	89 %	83 %	94 %	90 %	88 %
KNN	93 %	**98 %**	**99 %**	89 %	97 %	**97 %**	**98 %**	95 %	**98 %**
RF	**99 %**	**96 %**	**100 %**	**98 %**	**99 %**	96 %	**99 %**	**99 %**	**99%**
NB	84 %	82 %	**99 %**	82 %	83 %	**98 %**	**99 %**	92 %	**98 %**
DT	80 %	**99 %**	**99 %**	84 %	85 %	**99 %**	**99 %**	85 %	**99 %**
DNN	90 %	91 %	90 %	91 %	87 %	90 %	89 %	76 %	91 %
LSTM	**97 %**	94 %	91 %	**92 %**	**99 %**	92 %	92 %	**97 %**	93 %
GRU	**99 %**	93 %	93 %	**98 %**	**98 %**	91 %	93 %	**96 %**	91 %

each condition at a specific switching frequency. Then, the average accuracies are obtained. RF, GRU, and LSTM are the best three techniques for DC parallel arc diagnosis in case 1. Their performances are superior in all conditions. In case 1, only the time-domain signals are adopted as the inputs of all learning algorithms. The advantage of DL algorithms is clearly demonstrated; they do not need any feature parameter analysis to achieve high accurateness, whereas other ML techniques show mediocre performances. The accuracy of all learning algorithms usually increases with the rise of the switching frequency. When the switching rate increases,

there might be more useful evidence in each divided set. Thus, the diagnosis accuracies could be enhanced. Similar evaluations are replayed for each case to obtain the average accuracies. Table 5 illustrates the performance summary of all learning techniques in nine cases. The condition for nine cases is the same for all setting parameters, rated voltage, and current. The only difference is the use of input signals; each case uses different input combinations in time and frequency domains. This study aims to find the suitable input combination for each learning algorithm to achieve the highest performance. First, the process in case 1 is repeated

FIGURE 6. Detection rates for different conditions in case 1.

for the other eight cases with different input signals regarding Table 3. Then, Table 5 summarizes the average diagnosis rate in each case. The best three AI algorithms are RF, KNN, and GRU, whereas LSTM shows superior performance (average 94%). The other techniques show high (DT, NB) and mediocre (SVM, DNN) performances. ML algorithms need feature extractions to maintain the high detection rates. Thus, their accuracy increases when a frequency-domain signal is applied, and their performances are greatly improved with the presence of two or three frequency-domain signals as the

input. In cases 2, 3, 6, 7, and 9, when the input signals are analyzed using the FFT technique, the average accuracies of all ML techniques increase compared with that of in other cases. Comparing cases 3, 4, and 5, all ML techniques use only one processed signal, but the accuracies of ML techniques in case 3 are higher than in cases 4, and 5. Similarly, comparing cases 6, 7, and 8, all ML methods use two processed signals, but the accuracies of ML techniques in cases 6, and 7 are higher than in case 8. It can be seen that the processed source current signal is the critical signal for the high

performance of ML techniques. In cases 1, 4, 5, and 8, the accuracies of LSTM and GRU are higher than that of in other cases (especially case 3). Case 3 uses the raw signals of load voltage and current for arc diagnosis, the accuracies of LSTM and GRU in case 3 are lower than in cases 4, and 5; which are also use two raw signals (one of them is source current signal). Similarly, case 8 uses only one raw source current signal, and the average accuracies of LSTM and GRU in this case are still higher than cases 6, and 7. The raw source current signal is the key input for the high accuracy of DL approaches. Some useful evidence might vanish due to the source current's FFT analysis, reducing the deep learning technique accuracy. The results show that the diagnosis techniques should be chosen depending on the input signal. When the raw source current in the time domain is adopted as input, the DL techniques should be used for parallel arc detection. Otherwise, the ML should be adopted with the presence of the processed signal, such as the use of two or three frequency-domain signals as the input. Among six inputs (three inputs in time domain and three inputs in frequency domain), the source current in the time and frequency domains relates to the high performance of DL and ML techniques, respectively. The study focuses on finding the suitable input signals returning the highest general diagnosis rate. The feature extractions of the suitable input signal will result in higher accuracy if adopted. Therefore, finding the critical input signal is essential before features are extracted. In addition, several learning techniques, such as DL, have not required feature extraction for obtaining high accuracy rates. Therefore, feature extraction might degrade the performance of DL detection owing to the loss of useful information during the extraction progress.

V. CONCLUSION

Eight advanced algorithms combined with various input features in both frequency and time domains to detect parallel arc in this study under UL1699B standard. Generally, when the switching frequency increases, the diagnosis accuracy increases whether the inputs belong to the frequency or time domain. Furthermore, when the switching frequency to be used for controlling the 3-phase inverter increases, the number of sampling periods in each data increases. Therefore, the valuable characteristics in each data set are obviously increased, which results in the improvement of the accuracy rates. In addition, if the switching frequency is increased, the distortion of the signals will be lower. Thus, signals become smooth, and the difference between arcing and normal states becomes clear to detect. Furthermore, using a higher sampling frequency oscilloscope could result in more information in each signal. However, it could increase the execution time and computation burden, whereas one of the most priorities is detecting arc fault in time to separate the fault from the system.

The training process is needed when a new operating condition is applied, such as the change of switching frequency, load types, current and voltage amplitudes, or the weather. For example, when a cloud or rain cuts off the sunlight, the current and voltage characteristics in both arcing and normal states might be changed. Thus, the data, when the current and voltage changed, should be trained to enhance the diagnosis results. Otherwise, the accuracy rates might be degraded. The FFT analysis is employed to obtain the frequency-domain inputs. However, these inputs need higher computational resources and sampling rates than time-domain input. Thus, using inputs in the frequency domain could consume more processing time and hardware resources. This study offers a specific view and helpful information for arc failure diagnosis. However, this study is implemented in a laboratory environment and needs proper adjustments before applying to practical systems or applications. Another limitation is that the authors did not consider the hyper-parameter modifications of different AI techniques. It has been argued that the same AI technique could provide different accuracy rates for the same data set with different values of hyper-parameters. Thus, this study mainly focuses on finding the most suitable inputs returning the highest general diagnosis rate for ML and DL algorithms.

Machine learning techniques need feature extraction to maintain high detection rates. The source current in the frequency domain is the key to ML techniques' high performance. On the other hand, the high performance of deep learning approaches requires an extensive data set without feature analysis and high computational cost owing to their deeper structures compared with that of ML approaches. The hidden layer configurations in deep learning algorithms (LSTM, GRU, and DNN) were chosen based on the trial and error method. Therefore, many tests are required to find the most optimal performance. RF is the best diagnosis technique for DC parallel arc detection among eight learning algorithms. The detection rates of RF are greater than 96 % in all cases, and it offers high performance for both raw and processed input signals.

The diagnosis results prove that the source current in the time and frequency domains significantly relates to the high performance of DL and ML techniques, respectively. Thus, the diagnosis techniques should be chosen depending on the input signal. When the raw source current in the time domain is adopted as input, the DL techniques should be used for parallel arc detection. Otherwise, the ML should be adopted with the presence of the processed source current signal. This study offers a specific view of different learning techniques and input types. This might be helpful research for selecting the combinations between learning techniques, input types, feature extraction methods, which can support in building more reliable and robust systems when implementing an arc fault detection system regarding different priorities.

REFERENCES

[1] 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.

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

[3] 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.

[4] J. Johnson, M. Montoya, and S. McCalmont, "Differentiating series and parallel photovoltaic arc-faults," in *Proc. 38th Photovolt. Spec. Conf.*, 2012, pp. 720–726.

[5] M. Zhen, W. Li, and S. Qiangang, "The characteristics of DC arc faults current," in *Proc. 15th Eur. Conf. Power Electron. Appl. (EPE)*, Sep. 2013, pp. 1–9.

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

[7] Z. Li, Z. Ye, C. Wu, and W. Xu, "Modeling and simulation study of photovoltaic DC arc faults," in *Advanced Computational Methods in Energy, Power, Electric Vehicles, and Their Integration*. Springer, 2017, pp. 137–146.

[8] R. Liu, G. Meng, B. Yang, C. Sun, and X. Chen, "Dislocated time series convolutional neural architecture: An intelligent fault diagnosis approach for electric machine," *IEEE Trans. Ind. Inf.*, vol. 13, no. 3, pp. 1310–1320, Jun. 2017.

[9] T. Sirojan, S. Lu, B. Phung, D. Zhang, and E. Ambikairajah, "Sustainable deep learning at grid edge for real-time high impedance fault detection," *IEEE Trans. Sustain. Comput.*, early access, Nov. 7, 2018, doi: 10.1109/TSUSC.2018.2879960.

[10] T. de Bruin, K. Verbert, and R. Babuska, "Railway track circuit fault diagnosis using recurrent neural networks," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 28, no. 3, pp. 523–533, Mar. 2017.

[11] K. Xia, S. He, Y. Tan, Q. Jiang, J. Xu, and W. Yu, "Wavelet packet and support vector machine analysis of series DC ARC fault detection in photovoltaic system," *IEEJ Trans. Elect. Electron. Eng.*, vol. 14, no. 2, pp. 192–200, Feb. 2019.

[12] B. Grichting, J. Goette, and M. Jacomet, "Cascaded fuzzy logic based arc fault detection in photovoltaic applications," in *Proc. Int. Conf. Clean Electr. Power (ICCEP)*, 2015, pp. 178–183.

[13] 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.

[14] H. Gao, X. Wang, T. Nguyen, F. Guo, Z. Wang, J. You, and Y. Deng, "Research on feature of series arc fault based on improved SVD," in *Proc. IEEE Holm Conf. Electr. Contacts*, Sep. 2017, pp. 325–331.

[15] 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.

[16] 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, Dec. 2021.

[17] 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.

[18] 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.

[19] 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.

[20] 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.

[21] 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.

[22] 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.

[23] 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.

[24] 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.

[25] 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.

[26] 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.

[27] 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*, 1992, pp. 144–152.

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

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

[30] L. Breiman, "Random forests," *Mach. Learn.*, vol. 1, pp. 5–32, Oct. 2001.

[31] P. Langley, W. Iba, and K. Thompson, "An analysis of Bayesian classifiers," in *Proc. AAAI*, vol. 90, vol. 1992, pp. 223–228.

[32] 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.

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

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 with 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 and 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 with 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.

• • •