Journal of Power Electronics
https://doi.org/10.1007/s43236-021-00299-5

ORIGINAL ARTICLE

Deep learning-based series AC arc detection algorithms

Chang-Ju Park[1] · Hoang-Long Dang[1] · Sangshin Kwak[1] · Seungdeog Choi[2]

Received: 21 June 2021 / Revised: 3 August 2021 / Accepted: 4 August 2021

Abstract
Various studies on arc detection methods are described. Series AC arc is detected based on the characteristics extracted from arc voltage, frequency, and time domain of the current. Methods of arc detection using artificial intelligence have been studied previously. In the present study, the performance of multiple methods is analyzed by comparing different input parameters and artificial neural networks. In addition to the input parameters presented in the literature, the performance is compared and analyzed using the following parameters: zero-crossing period, frequency average, instantaneous frequency, entropy, combination of fast Fourier transform (FFT) and maximum slip difference, and combination of FFT and frequency average. These parameters and different neural networks are studied in the bounded and unbounded case, and the performance is compared. For different combinations of neural networks and input parameters, another research question is to identify the input parameters to be used if the number of training data is limited. Moreover, this study investigates the change in detection rate depending on the number of training samples. As a result, the minimum dataset size required to obtain the final detection rate is identified.

Keywords Arc detection · Deep learning · Neural network · Series arc

1 Introduction

Arcs may cause electrical fires in power systems. An arc is caused by the heat of an insulation breakdown, which carbonizes the line and results in a gap. The sparks generated in an arc fault potentially lead to fire and may cause damage to property and human life. The two main types of arcs are parallel and series arcs. Parallel arcs occur when an insulation breakdown between two or more lines occurs because of an external force or heat, a gap is generated, a line is formed, and the current flows through it. Series arc is a similar phenomenon that occurs in a single line. An insulation breakdown is also caused by an external force or heat, a gap appears in the line, and the current flows through the gap.

Both series and parallel arcs are difficult to detect with conventional detection equipment. To prevent such an accident, the series arc should be detected rapidly. Extensive literature on this topic is available. For example, many studies have investigated the detection of shouldering phenomena in AC currents. Several representative methods are as follows. A series arc was detected by passing through a harmonic filter when an AC current arc occurred [1] or by analyzing the harmonic components of the current fundamental wave using the chirp Z-transform [2]. Series and parallel arcs were identified by measuring the current when an arc occurred and detecting nonstationary components of the phase [3]. Series, parallel, and ground arcs were detected using wavelet transform [4–6]. Moreover, arcs were detected using the current characteristics [7–11]. A study compared the current and previous values to find the moving average of the current and identify a stationary frequency component [12]. Other studies calculated the RMS of the current [13] or used statistical characteristics [14]. Some other arc detection methods were based on previous methods and used signal processing and artificial neural networks. A study identified whether the heart rate was standard by classifying the electrocardiogram using long short-term memory (LSTM), which is a type of recurrent neural network (RNN) [16]. RNNs were also used in the research on speech recognition [17].

✉ Sangshin Kwak
sskwak@cau.ac.kr

[1] School of Electrical and Electronics Engineering, Chung-Ang University, Seoul, Korea

[2] Department of Electrical and Computer Engineering, Mississippi State University, Starkville, MS, USA

Published online: 17 August 2021

Many studies investigated techniques using a multilayer perceptron (MLP). A method of detecting a series arc used the zero-crossing period (ZCP) of the current as a parameter in a deep neural network (DNN), a type of MLP. In [18–20] series arcs were detected using the current ZCP, maximum slip difference (MSD), and fast Fourier transform (FFT) as input parameters in a DNN [21]. A technique for detecting a series arc using an RNN has been studied using time–domain current data as input parameters without modification [22]. Another DNN-based method for detecting series arcs used the current data whose fundamental wave component was attenuated by passing the time–domain current data through the high pass filter (HPF) as an input parameter [23]. The occurrence of a series arc was detected in a DC converter [24]. Support vector machines (SVMs) were also used [25].

We compared arc detection for different types of input variables and artificial neural networks. The following ten types of input parameters were included: (1) instantaneous frequency, (2) signal entropy, (3) MSD, (4) FFT, (5) frequency average, (6) ZCP, (7) FFT + MSD, (8) FFT + frequency average, (9) FFT + MSD + ZCP, and (10) time–domain current. We compared the performance of different combinations of neural networks and input parameters.

The remainder of this paper is organized as follows. "Series AC arc characteristics" describes the shape of the voltage and the current of each part of the circuit, which changes in the time and frequency domains when a series arc occurs owing to a line failure. "Arc detection using neural network" shows the types of artificial neural networks and the parameters used as inputs. "Series arc detection" describes the detection of bounded and unbounded cases of series arcs. Moreover, it presents the change in accuracy depending on the number of training data. Finally, "Conclusion" explains the suitability of parameter types depending on artificial neural networks and summarizes the performance for different types of input variables and networks.

2 Series AC arc characteristics

Figure 1 shows a circuit diagram used to obtain the actual data of the series arc. When an arc occurs in a circuit, it has the same effect as the impedance inserted into the circuit. Accordingly, the voltage across the load and the size of $i_s(t)$ decreases. Therefore, some power is lost in the arc, and the power delivered to the load is reduced.

In the series arc experiment, as shown in Fig. 1, the arc generation experiment setup is composed of an AC supply, arc generator, and load. An arc fault of the arc generator occurs when the arc rods are slowly separated and an arc is generated. An oscilloscope is used to collect and transfer data to a computer for arc detection. Figure 2 shows the

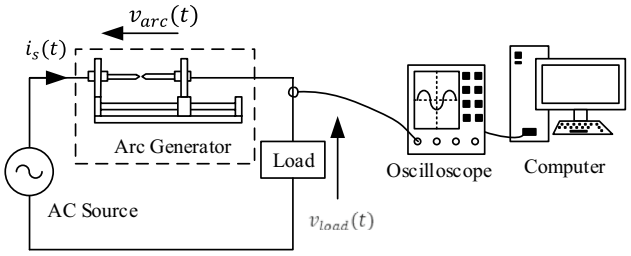

Fig. 1 Series arc schematic

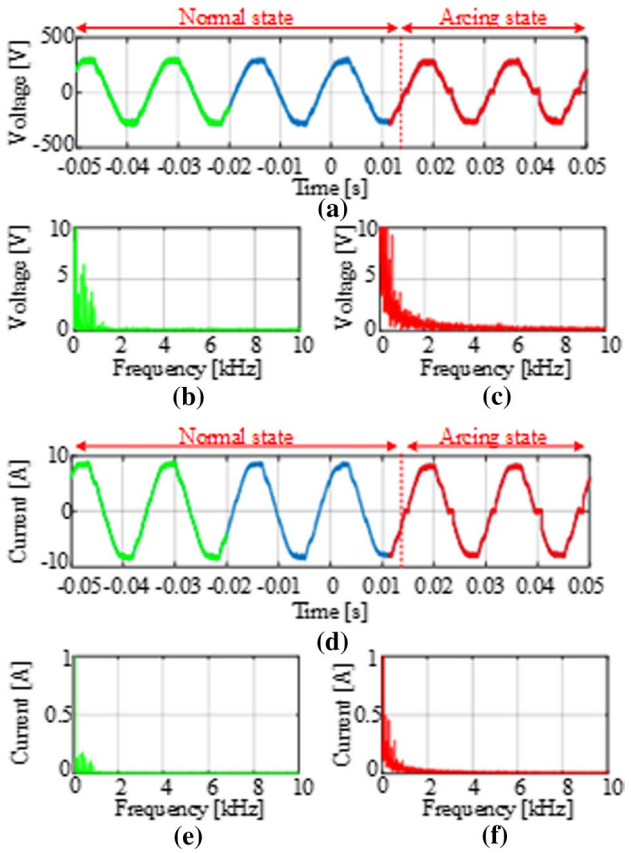

Fig. 2 Characteristics of the load voltage and the current: **a** load voltage, **b**, **c** FFT of the normal-state and the arc-state load voltage. **d** Load current. **e, f** FFT of the normal-state and the arc-state load current

waveforms of the load voltage, load current at 8 A, and their FFT analysis before and after arcing. The distortion occurred when an arc appeared in the waveform of the load voltage and the load current; it took the form of a sine wave in the normal state. When the arc rods are separated, an arc is generated, and harmonic components are added to the load voltage and load current. The length of the period when the voltage passes through the zero points is increased. Similarly, current $i_s(t)$ has a longer period passing through the zero points during the arc generation. The phenomenon of

lengthening the section passing through zero points is called shouldering. This phenomenon occurs due to the influence of the voltage $v_{arc}(t)$ given that the arc generated from the arc rod acts as an impedance. FFT was applied to determine the frequency changes before and after arcing. The data before and after arcing are marked in green and red, respectively. The multiples of the switching frequency are smaller before arcing and larger after arcing.

3 Arc detection using neural network

3.1 Neural network structures

DNN and RNN were used as artificial neural networks, and the ZCP and frequency average were used as parameters. In addition, FFT was applied, and the performance was compared using MSD, instantaneous frequency, and entropy [18–20].

Figure 3 shows the basic structure of an MLP: it accepts n parameters as inputs, passes them through a neural network composed of N layers, and derives results. The state of first layer was as follows:

$$\boldsymbol{h}_1 = f_1\left(\boldsymbol{W}_1^T\boldsymbol{X} + \boldsymbol{b}_1\right), \tag{1}$$

where $\mathbf{X}$ is the input parameter, $\boldsymbol{W}_1^T$ is the weight of the first layer, and $\boldsymbol{b}_1$ is the bias in the first layer. $\boldsymbol{h}_1$ is the output of the first layer, which is transferred to the second layer. This process was repeated N times. The basic method of training artificial neural networks is to perform one epoch, and then use the error to update weight and bias in the direction of reducing the error in each layer. This method is called error backpropagation. Figure 4 shows the structure of a single neuron in a fully connected layer (FC). This simplest structure has an important role in simulating a function: it

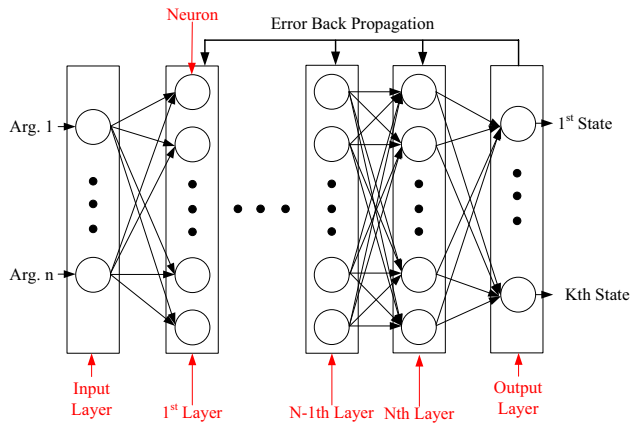

Fig. 3 Basic structure of artificial neural networks

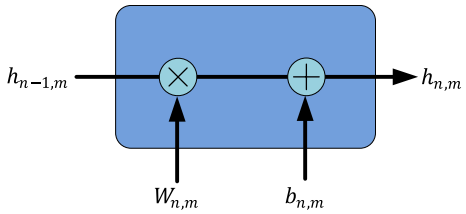

Fig. 4 Fully connected layer structure of DNN

connects all cells located before and after it. If the neurons of all layers are composed of FCs, then the corresponding neural network is called a DNN. $h_{n,m}$ is the output of the mth neuron of the nth layer. In addition, $W_{n,m}$ and $b_{n,m}$ represent the weight and bias of the mth neuron of the nth layer, respectively. Figure 5 shows the structure of a single neuron in LSTM. The LSTM structure can achieve efficient feedback of adequate information through selective forgetting and memory mechanisms. Thus, the network achieves the approximation of complex time-varying nonlinear functions better. In the case of LSTM, long-term and short-term memory are managed separately. h_{t-1} and h_t are the short-term memory states at the previous moment and at present, respectively. C_{t-1} and C_t are the long-term memory states at the previous moment and at present, respectively.W_0 and b_o are the current weight and bias of the current LSTM cell, respectively. x_t represents the input data, which come from another LSTM cell. y_t represents the output data, which are sent to another LSTM cell.

$$y_t = \left(\sigma\left(h_{t-1}x_t + W_0h_{t-1} + C_tC_{t-1} + b_o\right)\right)\tanh\left(c_t\right), \tag{2}$$

where σ is a variable that determines how much the weight and bias values are changed for the data received in one iteration as the learning rate. Figure 6 shows the structure of a gated recurrent unit (GRU) neuron. Different from LSTM, GRU is characterized by combining the long-term and short-term memory of the LSTM. h_{t-1} and h_t are the memory states

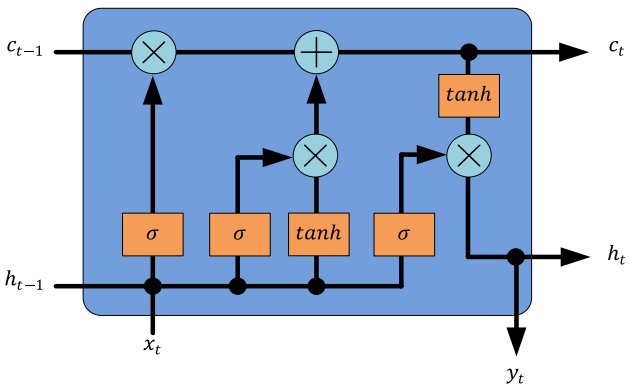

Fig. 5 Structure of LSTM

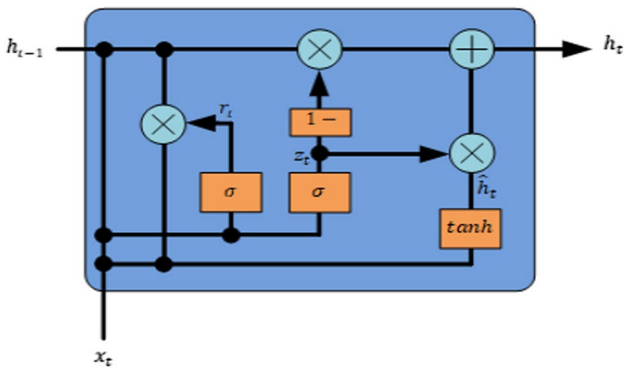

Fig. 6 Structure of a gated recurrent unit (GRU)

at the previous and present time, respectively. z_t and $\hat{h}_t$ are the update gate vector and candidate activation vector, and x_t is the current input. Equation (2) shows the output of each GRU neuron.

$$h_t = z_t \hat{h}_t + (1 - z_t) h_{t-1} \tag{3}$$

In this study, four artificial neural networks were used to verify the performance. As shown in Table 1, the structures are DNN (100, 50, 2), DNN (4, 5, 5, 2), LSTM (16, 16, 8, 8, 2), and GRU (16, 16, 8, 8, 2). Here, DNN (100, 50, 2) represents a three-layer DNN structure, and each layer contains 100, 50, and 2 neurons. A similar explanation is provided for the three other network structures. We denote DNN (100, 50, 2) as DNN1, DNN (4, 5, 5, 2) as DNN2, LSTM (16, 16, 8, 8, 2) as LSTM, and GRU (16, 16, 8, 8, 2) as GRU to simplify the notation. The depth and width of the hidden layers were selected by trial and error. The chosen layers showed the best performance among various structures. Interestingly, other suitable layer configurations exist.

Table 1 Structures of artificial neural networks used for arc detection

	DNN1	DNN2	LSTM	GRU
1st layer/ # of neurons	FC/ 100	FC/ 4	FC/ 16	FC/ 16
2nd layer/ # of neurons	FC/ 50	FC/ 5	LSTM/ 16	GRU/ 16
3rd layer/ # of neurons	FC/ 2	FC/ 5	FC/ 8	FC/ 8
4th layer/ # of neurons	N/A	FC/ 2	LSTM/ 8	GRU/ 8
5th layer/ # of neurons	N/A	N/A	FC/ 2	FC/ 2

3.2 Input parameters

The power supply current is measured to detect the series arc. When a series arc occurs, the ZCP tends to increase because of the shouldering in the power supply current. As a result, many parameters can be used to detect a series arc. The input parameters for series arc detection were as follows: time–domain current data, ZCP and FFT, frequency average, MSD, instantaneous frequency, and entropy.

Figure 7 shows the signals of the input parameters used for arc detection. The current data in the time domain have a $1 \times N$ array structure. N is 4165, which is the number of data per period (16.66 ms) of the current measured as 250,000 samples per second. The ZCP represents the time close to the current zero point, which indicates the time when the current passes ± 1/10 of the maximum value in one cycle. The ZCP of

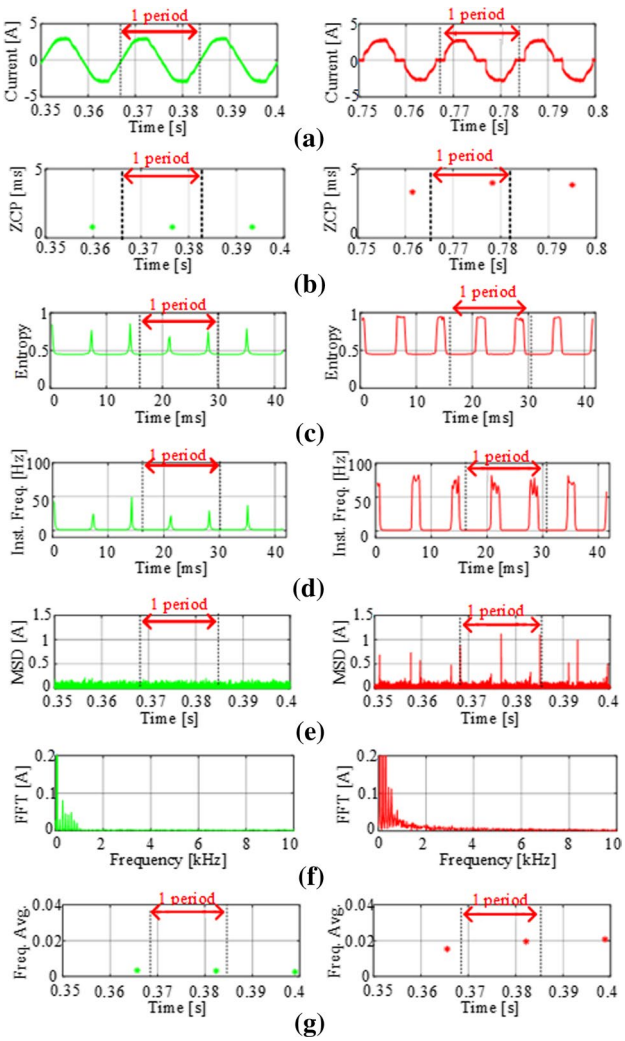

Fig. 7 Input parameters in normal (left) and arc (right) states: **a** current waveform; **b** ZCP of current; **c** entropy of current; **d** instantaneous frequency of current; **e** MSD of current; **f** FFT of current; **g** frequency average of current

the normal-state current is maintained at approximately 1 ms, whereas the ZCP of the current is maintained up to approximately 4–5 ms when an arc occurs. The size of the ZCP array is 1×1.

Entropy is an indicator used in the information and communication theory to represent the uncertainty of a signal. When the probability of a specific event i is p_i, it can represent the expected value of a specific event. Entropy tends to increase as it deviates from its expected value. In other words, it tends to increase as it deviates from the average [17]. This can be represented by the following equation:

$$H = -\sum_i p_i \log_2 p_i. \tag{4}$$

Instantaneous frequency is a concept from signal theory, and it is the amount of phase change of a function in the complex coordinate plane [7]. In (8), when the phase of an arbitrary function over time is ϑ(t), it can be obtained by differentiating it at a discrete time, as in (3).

$$\vartheta(\mathrm{t}) = \omega t + \theta \tag{5}$$

$$instFreq = \frac{1}{2\pi}\frac{\Delta\vartheta}{\Delta t} = \omega \tag{6}$$

Comparing Fig. 7d, we conclude that the component of the instantaneous frequency fluctuates significantly in the shouldering section when an arc occurs.

The MSD can check the degree of a sudden change in the current using the arc detection method suggested in [19]. In addition, this parameter allows the observation of a sudden change in the current before and after shouldering owing to high-order harmonics or shouldering. MSD can be calculated as follows:

$$MSD = \max\left|\frac{\sum_i^{i+4} I_i - \sum_{i+5}^{i+9} I_i}{5}\right|, \tag{7}$$

where I_i is the ith current value. The value of the MSD of the normal-state current is maintained below 0.5 in all sections. However, in the section, where the shouldering of the arc current occurs, the MSD has a shape similar to a pulse exceeding 0.5. The size of the MSD is an array of 1×4165.

FFT is a technique applied to analyze the frequency band of a signal. The analyzed signal has only a maximum of N frequency components, and the spectrum is limited to a 2π section and is repeated periodically [15, 26].

$$X[k] = \sum_{n=0}^{N-1} x[n] e^{-j2\pi kn/N}, 0 \le k \le N-1, \tag{8}$$

where $x[n]$ is a discretized current signal with a $1 \times N$ size, N is the size (4165) of one cycle of the current, and k represents the frequency band.

The frequency average is obtained using (9), as follows:

$$\mathrm{I}_{\mathrm{FreqAvg}} = \sqrt{\sum_{\mathrm{k=1}}^{\mathrm{N}} \frac{\mathrm{I}_{\mathrm{Freq}}(\mathrm{k})^2}{\mathrm{N}-1}} \tag{9}$$

First, the current of the nth cycle is collected. Then, FFT is performed to extract only the 1–3 kHz band and obtain the frequency average $I_{FreqAvg}$. This value is compared with the threshold value Ith. If it is higher than the threshold, then it is judged as an arc [27, 28]. The frequency average of the normal-state current is maintained at approximately 0.005, whereas the frequency average of the current reaches approximately 0.02 when an arc occurs. The size of the frequency average is 1×1.

Table 2 presents the order of time required to process one cycle of the current as an input parameter. The fastest time to process the power current waveform, which includes raw data, was 0.2 ms. On the contrary, the slowest time to obtain the instantaneous frequency was 40 ms.

4 Series arc detection

4.1 Bounded cases

In arc detection using an artificial neural network, two types of data can be distinguished, namely, (1) data from a category that has already been learned and (2) data from a category that has not been learned. They are referred to as the bounded and unbounded cases.

Figure 8a shows the distribution of training and test data for detecting the series arc in the bounded case. When the

Table 2 Order of input parameters used in series arc detection versus time complexity

Order	Input arguments	Time for calculation [ms]
1	Current @time domain	0.2
2	FFT	2.2
3	Frequency average	2.9
4	FFT + frequency average	3.1
5	ZCP	4
6	MSD	6.3
7	MSD + FFT	8.6
8	MSD + ZCP + FFT	12.8
9	Entropy	17
10	Instantaneous frequency	40

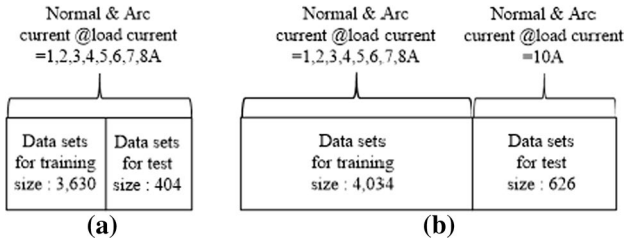

Fig. 8 Classification of training and testing data: **a** bounded case; **b** unbounded case

load current was 1, 2, 3, 4, 5, 6, 7, and 8 A, 90% of the normal-state and arc-state current data were used as a training set, and the remaining 10% was used as a test set. Figure 8b shows the distribution of the training and test data for the series arc detection in the unbounded case. In the unbounded case, the training set included the data for the cases when the load current was 1, 2, 3, 4, 5, 6, 7, and 8 A. On the contrary, the test set included the normal-state and arc-state current data when the load current was 10 A. As shown in Fig. 8a, the detection rate was compared for different input parameters and neural networks in the bounded cases. A total of 3630 data sets were used for training, and 404 data sets for testing.

Figure 9 shows the detection results of series bounded cases. The parameter with the highest detection rate for DNN1 is the frequency average and the combination of the FFT and frequency average, with a detection rate of 99%, outperforming the raw current presented in this study by 37%. The parameter indicating the highest detection rate for DNN2 is a combination of FFT and frequency average, and FFT shows the highest detection rate of 98%. This result shows a 42% higher detection performance than the raw current presented in this study. Compared with the case of using the time–domain current as the parameter of LSTM, the detection rate is 100%, which is 1% higher than the raw current used in this study, when the combination of the FFT, ZCP, and MSD is used as parameters. Compared with the time–domain current as a parameter in the GRU structure, the detection rate is 100%, which is 2% higher than that of the raw current, when the combination of the FFT, ZCP, and MSD is used as the parameter. LSTM and GRU have high detection rate for a combination of parameters as well as the raw current, but the disadvantage is that they require a longer time for training. On the contrary, DNN1 and DNN2 required less time to learn the same training data.

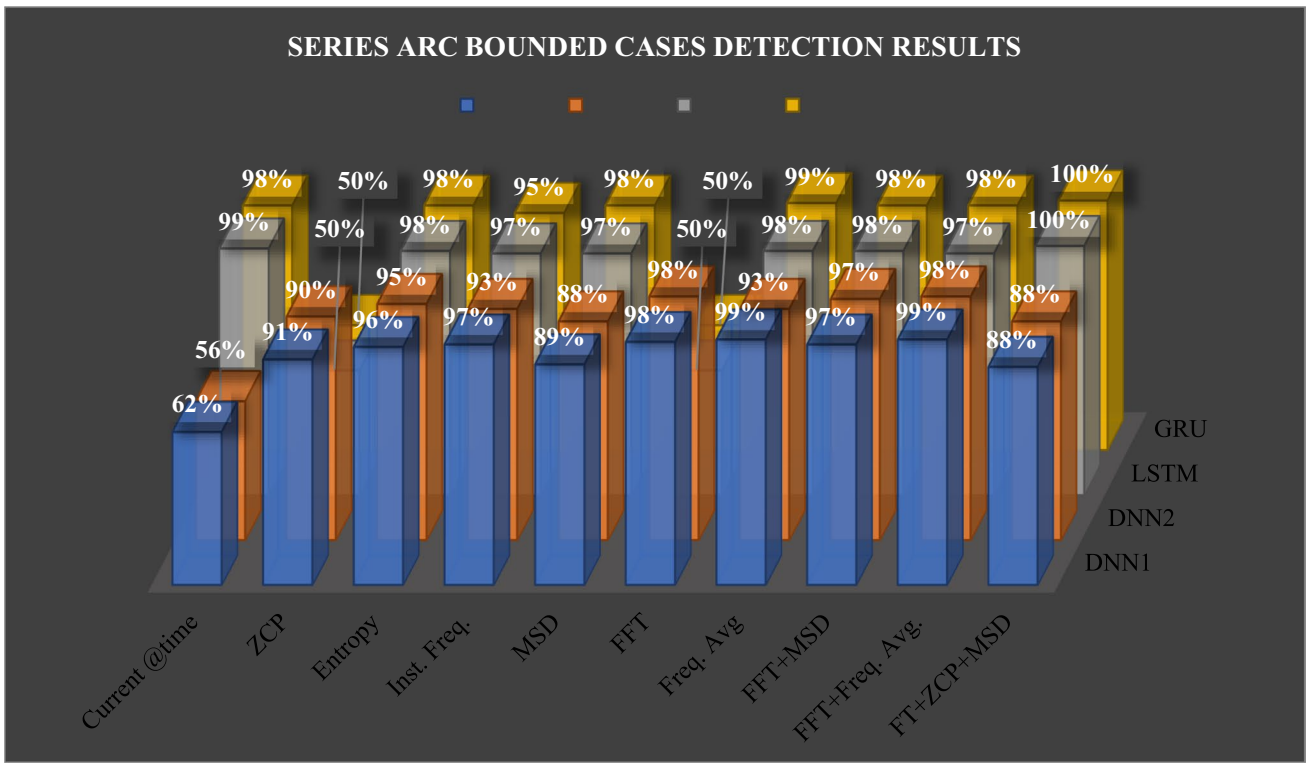

Fig. 9 Detection results of series arc bounded cases

4.2 Unbounded cases

This section compares the detection rate for different input parameters and neural networks in unbounded cases. A total of 4034 data sets are used for training used, and 626 data sets are used for testing.

Figure 10 shows the detection results of series unbounded case. For DNN1, the parameters showing the highest detection rate (100%) were the FFT and the MSD, the combination of FFT and frequency average, and the combination of FFT and MSD. This finding is 38% better detection performance than the time–domain current presented in this study. The highest detection rate for DNN2 was achieved by a combination of FFT and MSD and by MSD at 100%. This result has 44% higher detection performance than the time–domain current presented in this study. Compared with the case of using the time–domain current as the parameter for the LSTM structure, the detection rate is 6% higher when the FFT and MSD combination was used as a parameter. Compared with the time–domain current as the parameter of GRU, when the combination of entropy and instantaneous frequency, FFT, and MSD was used as parameters, the detection rate is the same as in the combination of FFT and MSD at 98%. Similarly, in the bounded cases, LSTM and GRU have high detection rate, but the disadvantage is that they require longer time for training. On the contrary, DNN1 and DNN2 require less time to learn the same training data. The input parameters (time–domain current, entropy, instantaneous frequency, MSD, and frequency average) with continuity over time show high detection rates for LSTM and GRU. In addition, the input parameters (ZCP, MSD, instantaneous frequency, FFT, and frequency average) with prominent features show a high detection rate for the DNN structures. In particular, the performance of combining two or more input parameters was better than that of one input parameter.

4.3 Detection rate considering the number of training data

This section discusses the change in the detection rate depending on the number of training data used to train the artificial neural network. The number of training data used to detect series arcs in the bounded case was 3630, the number of test data was 404, and the change from 0.5 to 100% in the number of training data was examined. In the unbounded case, the number of data for learning is 4034, and the number of test data is 626. The detection rate of training data varied from 0.5 to 100% as the bounded case was executed.

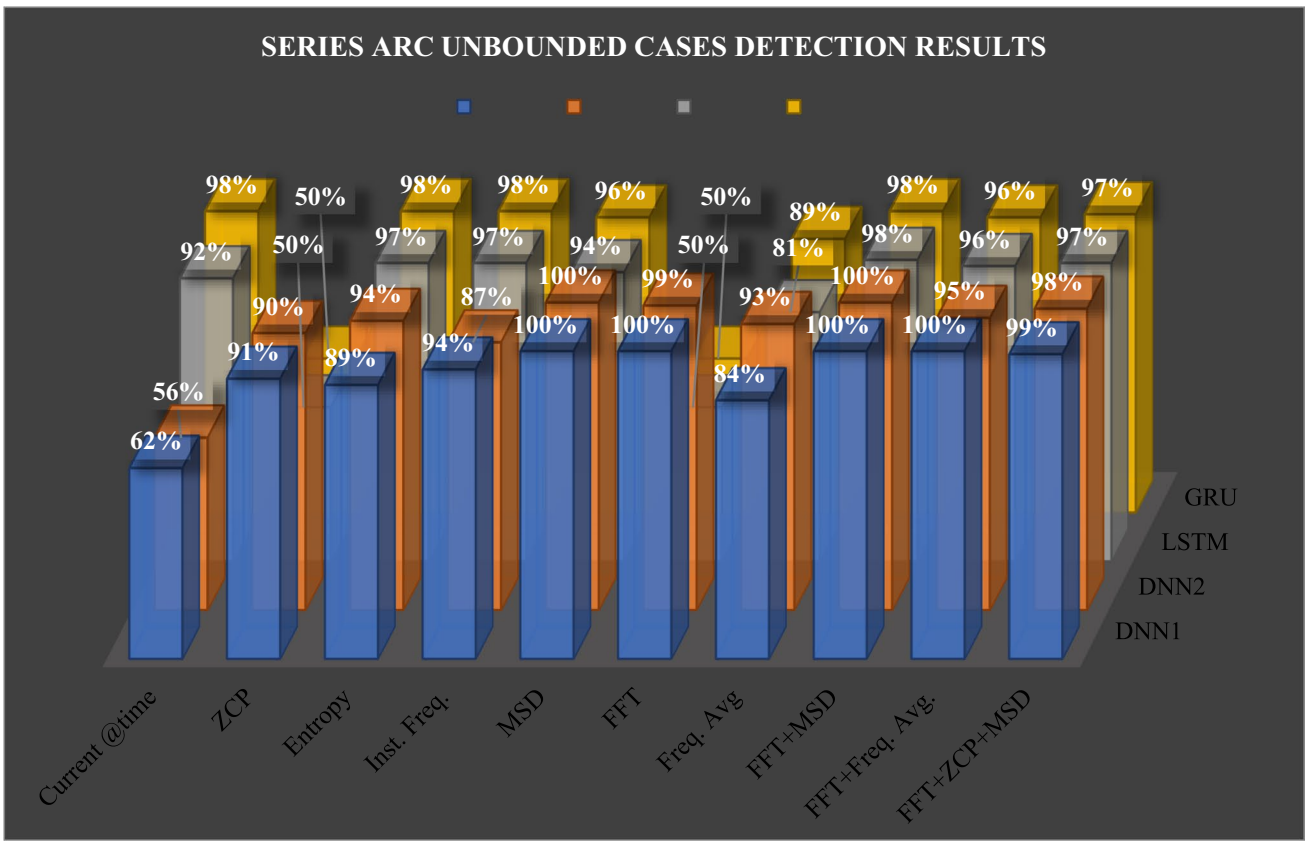

Fig. 10 Detection results of series arc unbounded cases

Figure 11 shows a plot of the relative accuracy band (RAB), indicating the extent to which the final detection rate was reached. The RAB indicates a band in the range of ±5% of the final accuracy of the final detection rate.

In Fig. 11, the accuracy remains within this band for the minimal dataset size of 2.5%. This dataset is called the minimum range of the dataset required to reach the RAB.

$$\Delta Accuracy = \max(accuracy) - \min(accuracy) \tag{10}$$

In addition, according to (10), the detection rate fluctuation $\Delta Accuracy$ is the maximum value of the different detection rates after the minimum range of the dataset that reached the RAB (hereinafter, the minimum range of the dataset) between the maximum value of the detection rate max(*accuracy*) and the minimum value of the detection rate min(*accuracy*). As shown in Fig. 11, in four cases, the detection rate is located outside the RAB after the minimum range of the dataset.

$$\%\ \text{of data out of band} = \frac{\#\ \text{of data out of band}}{(\#\ \text{of data in the band} + \#\ \text{of data out of band})} \tag{11}$$

In addition, the number of samples out of the band after the minimum range of the dataset that reached the accuracy band and the data beyond the range (the union of the datasets located within and outside the accuracy band) are expressed as a percentage, and the percentage of data outside the band can be determined by (11).

Table 3 presents the variation in the detection rate of the relative accuracy band depending on the input parameters of the bounded case using various neural network types. The combination of the FFT and MSD shows the lowest detection rate fluctuation for DNN1, and the percentage of samples out of the RAB is 0%. Furthermore, the minimum range of dataset is only 10%, which is lower than the other input parameters. For DNN2, the FFT shows the best performance with the lowest detection rate fluctuation, and the percentage of samples out of the RAB is 0%. Interestingly, the combination of the FFT and frequency average also shows high performance with low detection-rate fluctuation at 4.48%, which is only 1% higher than that of FFT. The instantaneous frequency provides the lowest detection rate fluctuation for LSTM. However, the percentage of minimum range of the dataset is 60%. Thus, the combination of the FFT and MSD shows the best overall performance. Similarly, the ZCP and the FFT offer high performance for GRU. Table 3 shows that the combination of FFT and MSD and the combination of FFT and frequency average show high performance for DNN1, DNN2, and LSTM. However, the ZCP and the FFT show high performance for GRU. The variation in the detection rate of the relative accuracy band depends on the input parameters of the unbounded case, and the neural network type is also presented in Table 3. According to the table, the combination of FFT and MSD shows the high performance for DNN1 and LSTM. The FFT and combination of FFT and frequency average show high detection rate for DNN2, and the percentage of samples out of the band is 0%. In addition, the ZCP shows a low detection rate fluctuation for LSTM and GRU.

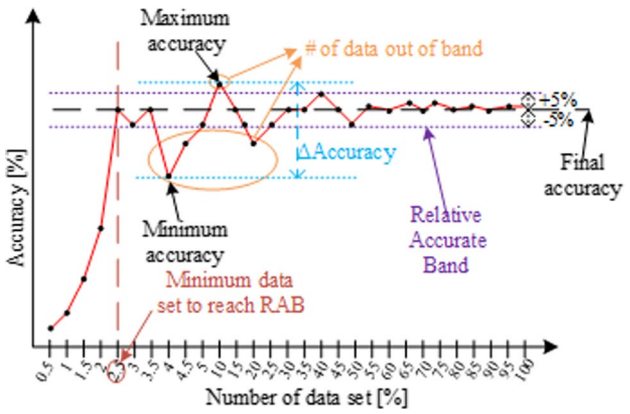

Fig. 11 Relative accuracy band

5 Conclusion

This study investigated the performance of various methods for detecting a series arc. Combinations of four types of artificial neural networks and ten input parameters were examined. In the case of a bounded series arc, the combination of FFT and MSD and the combination of FFT and the frequency average had high detection rate. In the unbounded case, the combination of FFT and MSD and the combination of FFT and frequency average, FFT, and MSD had high detection rates regardless of the neural network structure. The study also examined the arc detection rate depending on the number of training data. The results showed that the combination of FFT and MSD and the combination of FFT and frequency average can be used as input parameters in the training data regardless of the network structure. For the DNN structures, the use of the current waveform as an input parameter resulted in poor performance for the bounded case, and the detection accuracy was high with a small number of unbounded cases. Moreover, the use of ZCP and FFT for LSTM and GRU structures resulted in the lowest accuracy among the ten input parameters for the bounded and unbounded cases. In addition, the combination of FFT and

Table 3 Summary of detection rate fluctuations in relative accuracy bands according to the series bounded and unbounded cases, input parameters, and neural network type

AI Algorithms		Input parameters																		
		Current @Time		ZCP		Entropy		Instantaneous frequency		MSD		FFT		Frequency average		FFT + MSD		FFT + Frequency Average		FFT + ZCP + MSD	
		1	2	1	2	1	2	1	2	1	2	1	2	1	2	1	2	1	2	1	2
DNN1	Mi	1	0.5	50	0.5	30	0.5	30	0.5	40	0.5	4	0.5	50	2	10	0.5	20	0.5	40	0.5
	Acc	15.42	20.9	9	49.04	50.75	6.23	5.47	14.7	6.02	3.67	47.26	2.72	4.48	16.23	3.48	0	7.46	4.63	4.99	53.04
	Out	35	20	5	45	55	0	0	5	5	0	15	0	0	30	0	0	0	0	0	30
DNN2	Mi	0.5	1	40	0.5	30	0.5	3	0.5	60	0.5	10	0.5	20	0.5	20	0.5	10	0.5	20	0.5
	Acc	20.9	15.42	8.51	42.01	6.97	15.18	33.83	22.84	4.49	13.94	3.48	4.95	6.01	10.06	15.94	7.83	4.48	4.95	6.5	45.89
	Out	20	40	0	10	0	40	40	20	0	40	0	0	0	5	5	10	0	0	15	30
LSTM	Mi	30	1.5	5	0.5	80	60	60	20	30	60	5	0.5	50	0.5	3.5	0.5	1.5	0.5	1	0.5
	Acc	6.51	29.23	5.47	0	4.99	1	1.98	4.28	5.08	4.35	5.47	7.99	5.49	22.56	4	1.92	6.47	0.64	9.96	4.68
	Out	10	40	50	0	0	0	0	0	20	0	45	5	5	45	0	0	0	0	10	0
GRU	Mi	20	0.5	0.5	0.5	20	40	30	2.5	30	50	0.5	0.5	70	2	2	0.5	1	0.5	0.5	0.5
	Acc	11.44	5.91	5.47	0	7.46	9.33	19.4	26.12	4.47	7.47	5.47	7.99	0.52	16.07	6.47	4.95	17.42	12.94	17.91	5
	Out	10	5	10	0	5	5	5	45	0	10	10	5	0	10	5	0	35	5	15	0

1 bounded case, *2* unbounded case, *Mi* minimum range of dataset [%], *Acc* accuracy fluctuation [%], *Out* dataset out of band [%]

MSD and the combination of FFT and frequency average showed lower detection rate fluctuations than the remaining input parameters. Therefore, when detecting an arc using artificial neural networks, the accuracy detection was higher in the case of the combination of MSD and FFT and the combination of FFT and the frequency average for all network structures. The input parameters (time domain current, entropy, instantaneous frequency, MSD, and frequency average) with continuity over time show high detection rates for LSTM and GRU. In addition, the input parameters (ZCP, MSD, instantaneous frequency, FFT, and frequency average) with prominent features show high detection rate for the DNN structures. In particular, the performance achieved by combining two or more input parameters was higher than that of one input parameter. The structure of a neural network was essential. The GRU required the smallest amount of training samples to reach the final detection rate. However, LSTM and GRU generally required longer time than DNNs to train the same data samples. Assuming that all input parameters show equivalent performance, the DNNs can deliver the results faster than LSTM and GRU.

Generally, the use of a combination of FFT and MSD or a combination of FFT and the frequency average for all neural networks owing to their high performance is reasonable.

Acknowledgements This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korean government (MSIT) (No. 2020R1A2C1013413) and the Korea Electric Power Corporation (Grant number: R21XA01-3).

References

1. Bao, G., Jiang, R., Gao, X.: Novel series arc fault detector using high-frequency coupling analysis and multi-indicator algorithm. IEEE Access. **7**, 92161–92170 (2019)
2. Artale, G., Cataliotti, A., Cosentino, V., Di Cara, D., Nuccio, S., Tinè, G.: Arc Fault detection method based on CZT low-frequency harmonic current analysis. IEEE Trans. Instrum. Meas. **66**(5), 888–896 (2017)
3. Saleh, S.A., Aljankawey, A.S., Errouissi, R., Rahman, M.A.: Phase-based digital protection for arc flash faults. IEEE Trans. Ind. Appl. **52**(3), 2110–2121 (2016)
4. Saleh, S.A., Rahman, M.A.: Modeling and protection of a three-phase power transformer using wavelet packet transform. IEEE Trans. Power Delivery **20**(2), 1273–1282 (2005)
5. Kim, C.-H., Kim, H., Ko, Y.-H., Byun, S.-H., Aggarwal, R.K., Johns, A.T.: A novel fault-detection technique of high-impedance arcing faults in transmission lines using the wavelet transform. IEEE Trans. Power Delivery **17**(4), 921–929 (2002)
6. Lim, J., Runolfsson, T.: Improvement of the voltage difference method to detect arcing faults within unfused grounded-Wye 22.9-kV shunt capacitor bank. IEEE Trans. Power Delivery. **22**(1), 95–100 (2007)
7. Radojevic, Z.M., Terzija, V.V., Djuric, N.B.: Numerical algorithm for overhead lines arcing faults detection and distance and directional protection. IEEE Trans. Power Delivery **15**(1), 31–37 (2000)
8. Wang, B., Geng, J., Dong, X.: High-impedance fault detection based on nonlinear voltage-current characteristic profile identification. IEEE Trans. Smart Grid. **9**(4), 3783–3791 (2018)
9. Djuric, M.B., Radojevic, Z.M., Terzija, V.V.: Time-domain solution of fault distance estimation and arcing faults detection on overhead lines. IEEE Trans. Power Delivery **14**(1), 60–67 (1999)
10. Lee, C.J., Park, J.B., Shin, J.R., Radojevie, Z.M.: A new two-terminal numerical algorithm for fault location, distance protection, and arcing fault recognition. IEEE Trans. Power Syst. **21**(3), 1460–1462 (2006)
11. Russell, B.D., Benner, C.L.: Arcing fault detection for distribution feeders: security assessment in long term field trials. IEEE Trans. Power Delivery **10**(2), 676–683 (1995)
12. Saleh, S.A., Valdes, M.E., Mardegan, C.S., Alsayid, B.: The state-of-the-art methods for digital detection and identification of arcing current faults. IEEE Trans. Ind. Appl. **55**(5), 4536–4550 (2019)
13. Charytoniuk, W., Lee, W.-J., Chen, M.-S., Cultrera, J., Maffetone, T.: Arcing fault detection in underground distribution networks-feasibility study. IEEE Trans. Ind. Appl. **36**(6), 1756–1761 (2000)
14. Qu, N., Wang, J., Liu, J.: An arc fault detection method based on current amplitude spectrum and sparse representation. IEEE Trans. Instrum. Meas. **68**(10), 3785–3792 (2019)
15. Goldberger, A.L., Amaral, L.A.N., Glass, L., Hausdorff, J.M., Ivanov, PCh., Mark, R.G., Mietus, J.E., Moody, G.B., Peng, C.-K., Stanley, H.E.: PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals. Circulation **101**(23), 215–220 (2000)
16. Li, Z., et al.: E-RNN: Design Optimization for Efficient Recurrent Neural Networks in FPGAs. 2019 IEEE International Symposium on High Performance Computer Architecture (HPCA), Washington, DC, USA, 69–80 (2019)
17. Şentürk, Ü., Yücedağ, I. and Polat, K.: Repetitive neural network (RNN) based blood pressure estimation using PPG and ECG signals. 2018 2nd International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT), Ankara, 1–4 (2018).
18. Wang, Y., Zhang, F., Zhang, X., Zhang, S.: Series AC arc fault detection method based on hybrid time and frequency analysis and fully connected neural network. IEEE Trans. Industr. Inf. **15**(12), 6210–6219 (2019)
19. Wang, Y., Hou, L., Paul, K. C., Ban, Y., Chen, C. and Zhao T.: Series AC arc fault detection based on raw current and convolutional neural network. IEEE Transactions on Industrial Informatics (2021)
20. Abdulrachman, M. A., Prasetyono, E., Anggriawan, D. O. and Tjahjono A.: Smart detection of AC series arc fault on home voltage line based on fast Fourier transform and artificial neural network. 2019 International Electronics Symposium (IES). 439–445 (2019)
21. Jiang, J., et al.: Series arc detection and complex load recognition based on principal component analysis and support vector machine. IEEE Access. **7**, 47221–47229 (2019)
22. Li, W., Liu, Y., Li, Y., Guo, F.: Series arc fault diagnosis and line selection method based on recurrent neural network. IEEE Access. **8**, 177815–177822 (2020)
23. Wang, Y., Zhang, F., Zhang, S.: A new methodology for identifying arc fault by sparse representation and neural network. IEEE Trans. Instrum. Meas. **67**(11), 2526–2537 (2018)
24. Le, V., Yao, X., Miller, C., Tsao, B.: Series DC arc fault detection based on ensemble machine learning. IEEE Trans. Power Electron. **35**(8), 7826–7839 (2020)
25. Gao, H., et al.: Research on feature of series arc fault based on improved SVD. 2017 IEEE Holm Conference on Electrical Contacts, Denver, CO, 325–331 (2017)
26. Satpathi, K., Yeap, Y.M., Ukil, A., Geddada, N.: Short-time Fourier transform based transient analysis of VSC interfaced

point-to-point DC system. IEEE Trans. Industr. Electron. **65**(5), 4080–4091 (2018)
27. Gu, J., Lai, D., Wang, J., Huang, J., Yang, M.: Design of a DC series arc fault detector for photovoltaic system protection. IEEE Trans. Ind. Appl. **55**(3), 2464–2471 (2019)
28. Borges, F.A.S., Fernandes, R.A.S., Silva, I.N., Silva, C.B.S.: Feature extraction and power quality disturbances classification using smart meters signals. IEEE Trans. Industr. Inf. **12**(2), 824–833 (2016)
29. Park, C.-J.: Research on series and parallel arc detection technology on AC line. Master thesis, Chung-Ang University (2021)

Chang-Ju Park received his B.S. and M.S degrees in the Department of Electrical Engineering and School of Electrical and Electronics Engineering from Kwang Woon University and Chung-Ang University, Seoul, in 2019 and 2021, respectively. His current research interests include the design, model, control, and analysis of power system for energy storage system.

Hoang-Long Dang received his B.S. degree in electrical and electronics engineering from the Ho Chi Minh University of Technology, Viet Nam, in 2015. He is currently pursuing 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 received his 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, Korea. Since 2010, he has been working at Chung-Ang University, Seoul, Korea, where he is presently 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 received his B.S. degree in Electrical and Computer Engineering from Chung-Ang University, Seoul, Korea, in 2004; his M.S. degree in Electrical and Computer Engineering from Seoul National University, Seoul, Korea, in 2006; his Ph.D. degree in Electric Power and Power Electronics at Texas A&M University, College Station, TX, USA, in 2010. From 2006 to 2007, he was a Research Engineer at LG Electronics, Seoul, Korea. From 2009 to 2012, he was a Research Engineer at Toshiba International Corp., 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.