IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received November 24, 2021, accepted December 11, 2021, date of publication December 14, 2021,
date of current version December 24, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3135526

Different Domains Based Machine and Deep Learning Diagnosis for DC Series Arc Failure

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 National Research Foundation of Korea (NRF) funded by the Korean Government (MSIT) under Grant 2020R1A2C1013413, and in part by the Korea Electric Power Corporation under Grant R21XA01-3.

ABSTRACT Series arc faults are becoming more dangerous in DC systems. Without detecting in time and separation correctly, these fault events can cause electrical fires or explosions, creating a massive threat to people's safety and properties. This paper presents an analysis and comparison of DC series arc fault detection using various artificial intelligence (AI) algorithms in DC systems. The combinations of six feature parameters in both time and frequency domains with various AI techniques are recommended to detect DC series arc fault effectively. The performance and effectiveness of different combinations between feature parameters and learning techniques are summarized and discussed. Finally, practical challenges are identified, and suitable combinations of feature parameters and learning techniques are recommended for different operation conditions.

INDEX TERMS Fault diagnosis, DC series arc, machine learning, different domains.

I. INTRODUCTION

Recently, with the rising of pollution concerns on the global environment, green energies have become a potential candidate to replace traditional energies that come from fossil fuels. One of the most popular green energies is solar power. Photovoltaic (PV) generation on the rooftop or a solar power farm could be essential in supplying power and supporting different loads and micro-grids. However, the operation of PV systems can create DC arc fault [1]. An arc fault is a dangerous event that must be detected quickly. The arc fault can lead to fire or explosion and harm human safety and properties [2]. There are three types of DC arc fault, and they are grounding arc, serial arc, and parallel arc. Generally, a fracture of wiring or a loosening connection at the join junction could cause the serial arc fault. On the other hand, the parallel arc happens among the electrical lines, whereas the grounding arc is a short-circuit event among the line and the ground. The damage of wire insulation mainly causes these two arcs (parallel and grounding). Then, the magnitude of current will level up significantly when the parallel or grounding arcs happen. Conventional protective devices could be helpful to maintain the stability and safety of DC systems. However, the current will decrease when the serial arc occurs. Then, conventional protective devices will not be triggered, and the network is in a great dangerous situation [3]. Therefore, detecting serial arc fault correctly and timely is a vital task in maintaining the stable operation of DC systems. Several abnormal behaviors could be used when the serial arc is initial, such as the increase of heat, arc light, fluctuations of current, and electromagnetic interference to detect an arc fault. These visible behaviors can be adopted to judgment the serial arc [4]–[8]. Although specific physical behaviors could detect DC arc faults, some of these techniques are challenging to detect arc faults in practical applications because of the installation location of the sensors at unknown points.

The development of information technology is drawn more and more attention from researchers owing to their flexible capability in different applications. Advanced algorithms, such as Artificial intelligence (AI) techniques, have been successfully applied in different fields. They offer helpful methodologies for detecting faults in different fields, such as fault diagnosis in medium voltage networks based high impedance [9], failure detection in electrical machines [10], and fault diagnosis of track circuits in railway systems [11].

The associate editor coordinating the review of this manuscript and approving it for publication was Zhixiang Zou.

Researchers have successfully applied these advanced techniques for DC arc fault diagnosis and achieved several positive results, such as the combination of the wavelet packet decomposition and support vector machine (SVM) algorithm for DC arc fault diagnosis [12], the hidden Markov model (HMM) was adopted for correctly detecting series arc faults using the maximum likelihood [13]. Several characteristics, such as the variations of current and high-frequency components, are obtained and used for training models centered on weighted least squares SVM techniques to detect series arc [14]. In addition, constructing an attractor matrix from current signals and singular value decomposition is adopted to extract features in [15], the combination of sparse coding characteristics and an artificial neural network for arc fault diagnosis was proposed in [16], and the domain adaptation combining with a deep convolutional generative adversarial network was presented in [17]. A study illustrates a comparison between different AI methods in DC systems [18]. In general, mentioned researchers mainly concentrate on one specific working condition, such as one control technique, one input type, or a particular switching rate. In contrast, the effectiveness of learning techniques is critically dependent on the working conditions. Different operation conditions that could lead to inaccuracy on arc detection are not thoroughly investigated. There is a necessity for summary research with different loads, modulation algorithms, feature parameters, switching rates. In this paper, six feature parameters and eight learning techniques have been joined into different combinations; they are used for detecting arc faults and the performances are compared between different combinations. Two types of feature parameters, which are belonged to time and frequency domains, such as the peak-to-peak value, average value, RMS value, variance value, median value [19], and the fast Fourier transform (FFT). There are three groups of inputs, and they are time-domain input (average, variance, median, rms, and peak-to-peak values), frequency-domain input (FFT), and combined input (all features). These features are adopted as inputs of AI algorithms. Three types of loads were used for arc experiments: resistor, single-phase inverter, and three-phase inverter loads. The space vector modulation and model predictive control were adopted to control the three-phase converter, whereas sinusoidal pulse width modulation was employed to regulate the single-phase inverter. The switching frequencies and current amplitudes vary from 5 to 20 kHz and from 3 to 8 A, respectively. The performance of different combinations between AI algorithms and input types is compared and discussed in different working conditions. The rest of this study is structured as follows. Section 2 details the arc-generation hardware and the characteristics of current in normal and arcing states when an arc happens, and feature extractions are analyzed for arc diagnosis in this study. Section 3 presents the learning techniques used for arc fault diagnosis. Section 4 discusses diagnosis performances using different combination from eight learning techniques and six features parameters when an arc fault happens in altered operation conditions. In conclusion, the summary of the arc fault diagnosis regarding the performance of different combinations between learning techniques and feature parameters is illustrated in Section 5.

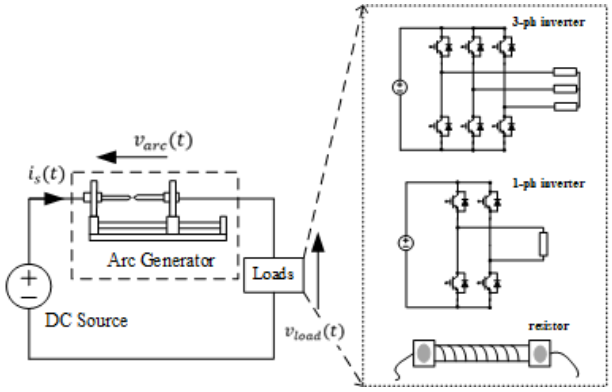

FIGURE 1. Experimental hardware setup.

TABLE 1. Parameters of nonlinear loads in the arc-generation circuit.

Switching frequencies	Nonlinear loads		
	Three-phase inverter		Single-phase inverter
	Space vector pulse width modulation	Model predictive control	Sinusoidal pulse width modulation
5 kHz	Current amplitudes: 3, 5, and 8 A	Current amplitudes: 5, and 8 A	Current amplitude: 8 A
10 kHz			
15 kHz			
20 kHz			

II. ARC CURRENT CHARACTERISTICS AND FEATURE EXTRACTIONS

The block diagram in Figure 1 illustrates the hardware setup for collecting arc data. The arc circuit was built regarding UL1699B standard [20], and [21]to obtain the data. The experimental setup consists of a DC source, arc generator, and loads (three- and single-phase inverters, resistors) [22]. DC supply source model in the experiment is N8741A (Keysight Technologies, USA). The nonlinear load specifications are illustrated in Table 1. The inverters (three- and single-phases) were assembled by using several transistor modules (insulated gate bipolar transistor) [23]. The switching rate was varied when the model predictive control (MPC) was adopted. Therefore, the switching rate of MPC in this study was the average rate acquired from the ratio between the turned on and off number with the specific operating period of the switching. Different combinations were used to compare the effectiveness of arc diagnosis for different working conditions. Then, discussion on which combination yields the best performance for arc fault detection in different conditions.

Figure 2 shows the waveforms of different features for resistor load in normal and arcing states. As shown in the figure, the feature waveforms before and after arcing

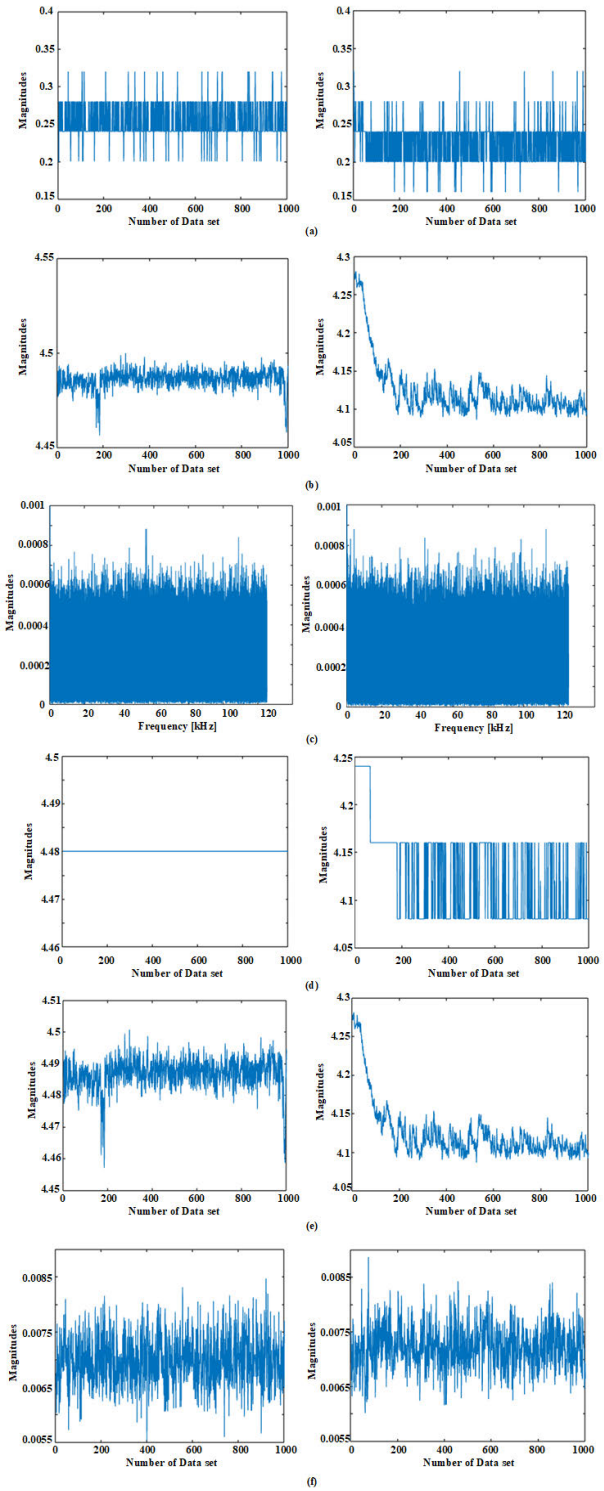

FIGURE 2. Feature extractions for resistor load at 5 A in normal (left) and arc (right) states. (a) Peak-to-peak, (b) Average, (c) FFT, (d) Median, (e) RMS, and (f) Variance.

were different. When an arc was initial, there were many differences in the feature waveforms before and after arcing point. These abnormal differences could be usefully used for arc fault diagnosis. First, the arc signal was collected

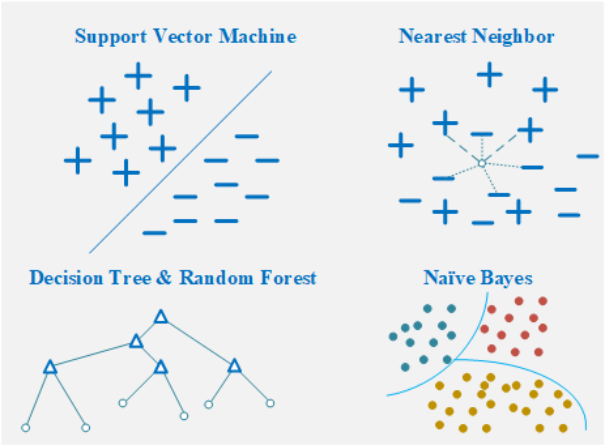

FIGURE 3. Principles of machine learning algorithms.

at 250 kHz sampling rate. After that, the sampled data were divided into 2 ms data sets for training and testing processes. Then, the feature extraction is processed for each data set to gain a set of features. Each set contains six different feature parameters, and they were entered as inputs of AI algorithms to detect the arc fault. They were average, FFT, variance, median, rms, and peak-to-peak values. For the frequency-domain feature, time-domain data sets with a 2 ms period were transformed into a frequency-domain signal using the FFT technique. The size of each FFT signal is an array of 500 $\times$ 1. Then, these FFT signals were used as input of AI algorithms. Obtaining features is an essential process for learning algorithm implementation. A feature can represent one aspect of the original data but not fully demonstrate the original signal. Therefore, a set of features increases the effectiveness of ML algorithms. For example, suppose only FFT (Figure 2(c)) or variance (Figure 2(f)) is used as an input for artificial learning algorithms. In that case, the diagnosis performance could be degraded cause the differences between normal and arcing states are unclear in some sectors. Furthermore, when different working conditions are applied, using one or some but not all features could lead to the poor accuracy of arc fault detection.

III. ARTIFICIAL LEARNING TECHNIQUES

Figures 3 and 4 present the principle of five machine learning (ML) and three deep learning (DL) algorithms used in this study. They are support vector machine (SVM), K-nearest neighbor (KNN), decision tree (DT), Random Forest (RF), Naive Bayes (NB), deep neural network (DNN), long short-term memory (LSTM), and gated recurrent unit (GRU). The objective of the SVM algorithm is to find a hyperplane that, to the best degree possible, separates data points of one class from those of another class [24]. KNN algorithm assumes that similar things exist in close proximity. In other words, similar things are near to each other [25]. The goal of DT is to generate a model that could predict the target value by learning simple decision rules indirectly

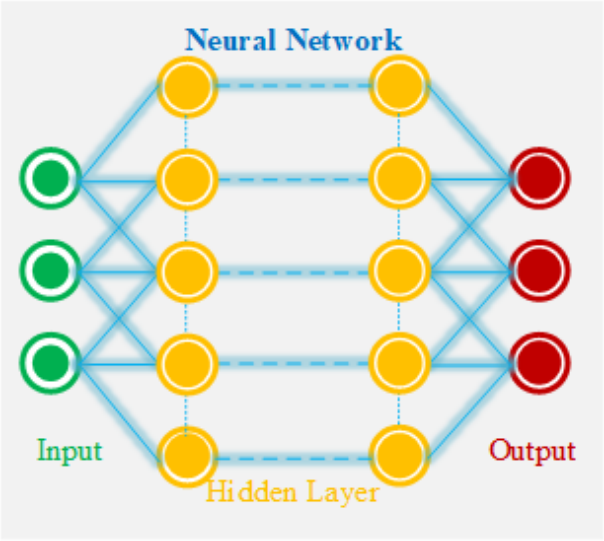

FIGURE 4. Structure of deep learning algorithms.

TABLE 2. Layer configurations of deep learning algorithms for arc fault diagnosis.

Layers		DNN	LSTM	GRU
1st	Layer type	FC	FC	FC
	Number of neurons	500	16	16
2nd	Layer type	FC	LSTM	GRU
	Number of neurons	412	16	16
3rd	Layer type	FC	FC	FC
	Number of neurons	256	8	8
4th	Layer type	FC	LSTM	GRU
	Number of neurons	128	8	8
5th	Layer type	FC	FC	FC
	Number of neurons	2	2	2

from the data features. A tree can be seen as a piecewise constant approximation [26]. RF is a meta estimator that fits a number of decision tree classifiers on various sub-samples of the dataset and uses averaging to improve the predictive accuracy and control over-fitting [27]. NB methods are a set of supervised learning algorithms based on applying Bayes' theorem with the ''naive'' assumption of conditional independence between every pair of features given the value of the class variable [28]. Unlike ML, DL algorithms are based on the brain of a human, copying how biological neurons communicate with other ones. DL techniques depend on training models to learn and improve their accuracy over time. The structure of DL algorithms comprises multiple layers, containing an input layer, several hidden layers, and an output layer. Each layer consists of several neurons; the output of one neuron in the n^{th} layer is the input of another neuron in the $n+1^{th}$ layer [29].

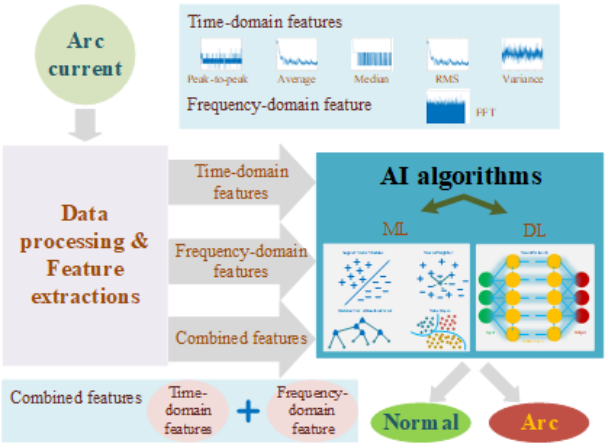

FIGURE 5. Block diagram of arc fault detection.

FIGURE 6. The contribution sizes of training and test data.

The structure configurations of DL techniques (DNN, LSTM, and GRU) are shown in Table 2. There are five layers in the DNN, LSTM, GRU structures, and the layers in DNN are fully connected (FC) layers. The neuron numbers of DNN layers 1, 2, 3, 4, and 5 were 500, 412, 256, 128, and 2, respectively. The neuron numbers of LSTM and GRU layers 1, 2, 3, 4, and 5 were 16, 16, 8, 8, and 2, respectively. The second and fourth layers of the LSTM structure are the LSTM layers, whereas they are GRU layers in the GRU structure. The trial and error method chooses the number of layers and neurons of hidden layers. The selected layer configurations of DL techniques offered the best effectiveness among numerous configurations. However, there are other layer structures, which are also suitable.

IV. ARC FAULT DIAGNOSIS BASED MACHINE LEARNING TECHNIQUES

In this study, the accuracy metric evaluates the effectiveness of the learning techniques. The accuracy rate is obtained by dividing the correctly predicted data sets and the total data

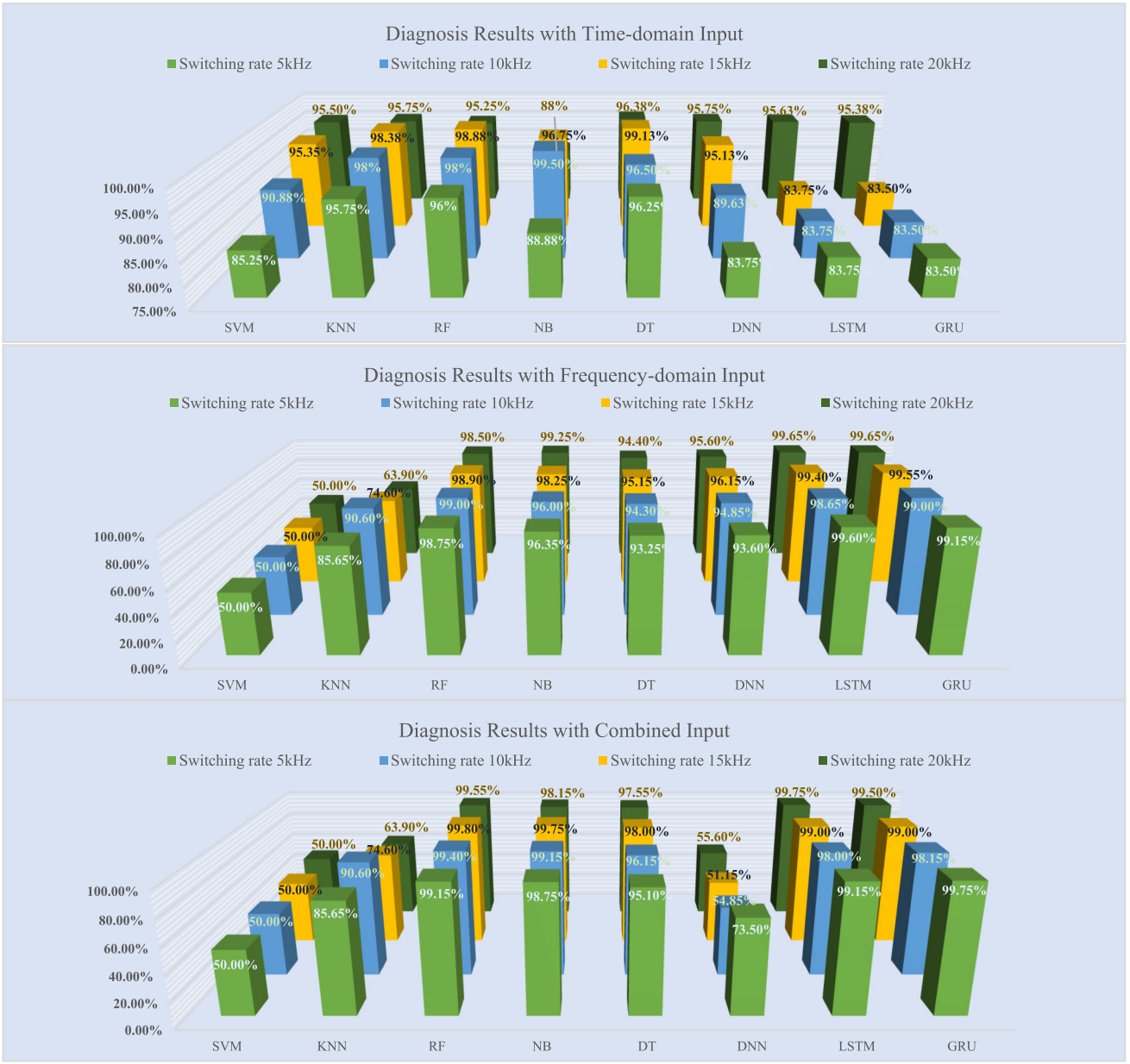

FIGURE 7. Arc diagnosis results for the three-phase inverter load with SVPWM (3PSVM) at 3 A under different input types.

sets in the test group. It is expressed as

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

Therefore, the technique with the higher accuracy rate is the better detection technique. Figure 5 illustrates the block diagram of arc fault diagnosis. The current data is sampled and processed to extract the features. There are three types of input features: time-domain, frequency-domain, and combined features, which is a combination of time-domain and frequency-domain features. Then, three input types of input are used for ML and DL algorithms to detect arc fault. The size ratio between the training set and test set is shown in Figure 6. There are two groups of data, the training group and the test group. The current amplitudes are 3, 5, and 8 A for a three-phase inverter load with SVPWM (3PSVM); 5 and 8 A for a three-phase inverter load with MPC (3PMPC) and resistor; and 5 A for a single-phase inverter load with SPWM (1PSPWM). The switching rates are set at 5, 10, 15, and 20 kHz, respectively. As shown in Figure 6, there are 26,000 sets of the arcing data and normal data in the training group, whereas 20,800 data sets of arcing and normal states are allocated in the test group. The training data were omitted from the test data. The number of the normal and arcing sets is equal in both training and test groups.

Figure 7 presents accuracy rates for 3PSVM at 3 A current amplitude when different input types were employed. In the

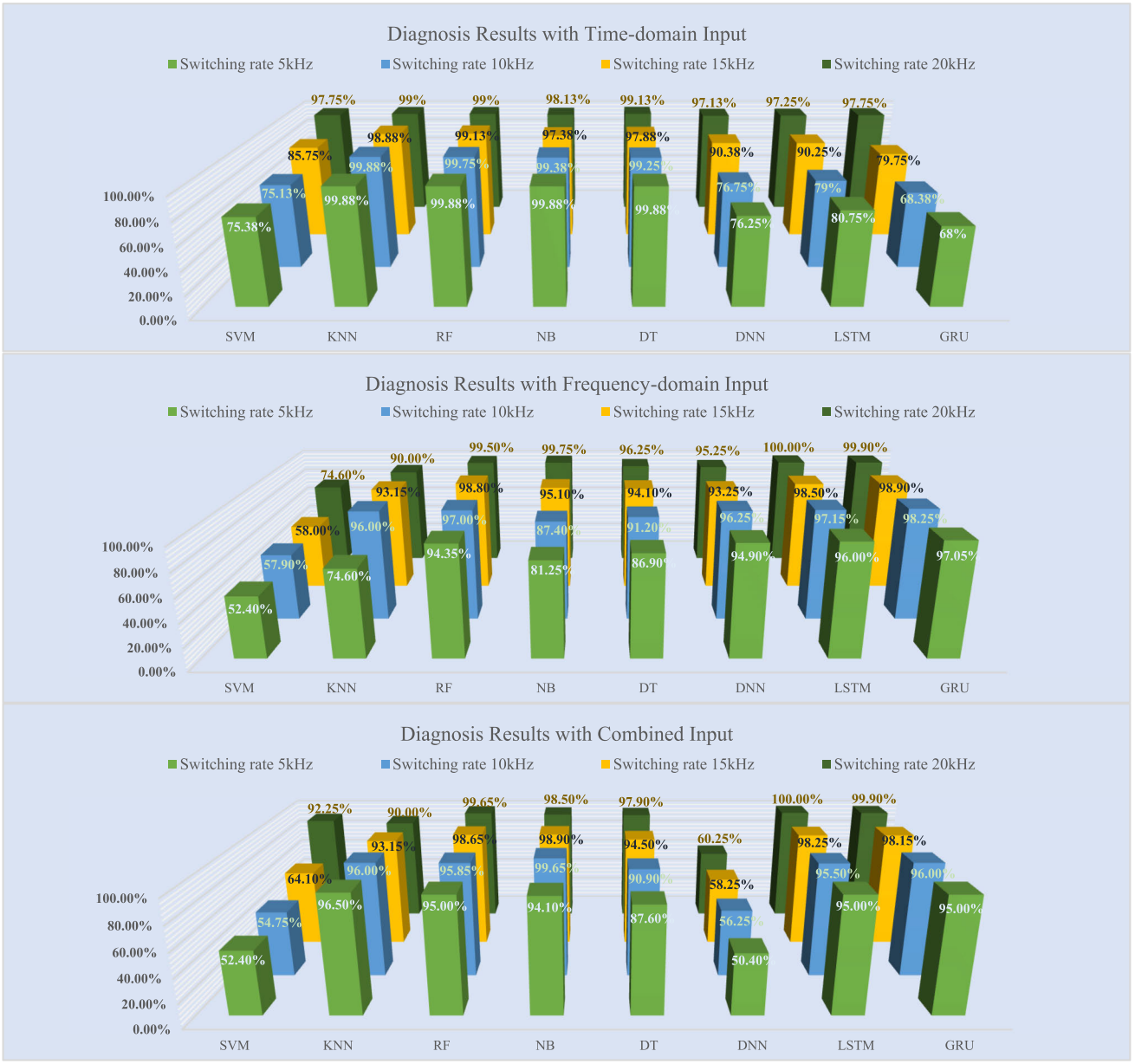

FIGURE 8. Arc diagnosis results for 3PSVM at 5 A under different input types.

first case, in which the time-domain feature was adopted as input of AI algorithms, DT achieved the best accuracy rate at switching frequencies 5 and 15 kHz. The performance of NB was the best at 10 kHz. GRU performed the lowest diagnosis rate at switching frequencies 5, 10, and 15 kHz. At 20 kHz, the best detection technique was DT with 96.38% accuracy, and the technique with the lowest accuracy was NB. The overall accuracy of DT was highest in all switching bands, and the performance of SVM had the lowest accuracy compared with other ML techniques. The best performance between DL techniques was DNN, and GRU was the lowest diagnosis rate. The ML algorithms' detection rates increased when the switching rates rose from 5 to 15 kHz, and the best performance was achieved at switching rate of 15 kHz. In contrast, the performance of DL techniques was highest at 20 kHz. When the frequency-domain feature was used, the accuracy rate of LSTM was highest at 5 kHz. On the other hand, RF and GRU techniques offered the highest diagnosis rate at 10 kHz. In addition, the accuracy of GRU was higher than the other techniques at switching frequency 15 kHz and 20 kHz with LSTM. The performances of SVM and KNN were poor or mediocre in all switching bands. Other learning techniques achieved high performance with minor differences in all frequency ranges. When time- and frequency-domains inputs were combined, GRU, LSTM, DT, NB, and RF displayed high performance (above 95%) in all switching rates, whereas KNN and DNN displayed mediocre performance. The diagnosis results for 3PSVM at 5 A current

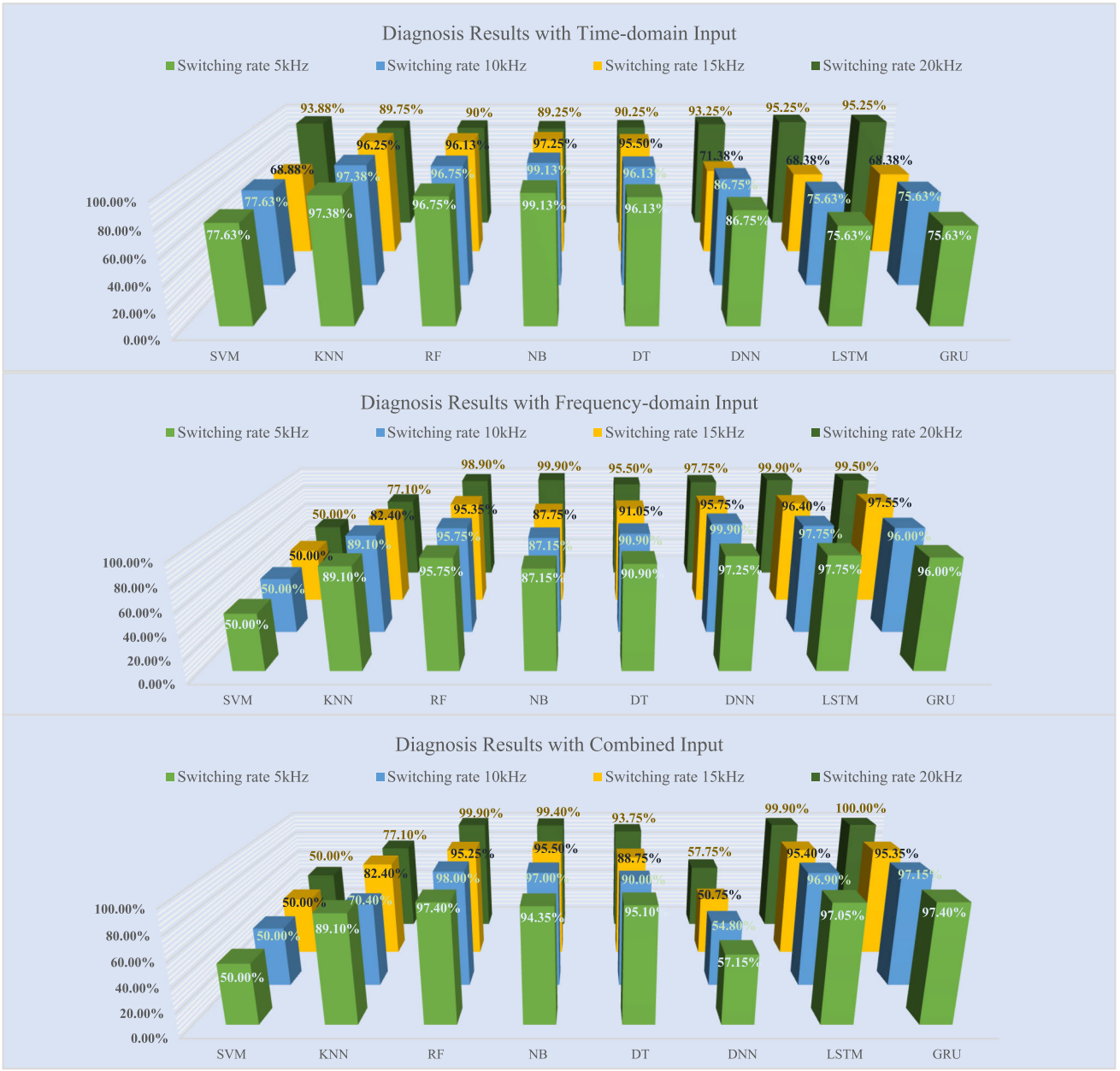

FIGURE 9. Arc diagnosis results for 3PSVM at 8 A under different input types.

amplitude are illustrated in Figure 8. NB, RF, DT, and KNN shared first place in accuracy at 5 kHz when time-domain input was employed. On the other hand, RF, DT, and KNN achieved the best performance at switching rates 10, 15, and 20 kHz, respectively. In contrast, the accuracy of GRU was the lowest in the switching rates 5, 10, and 15 kHz, and DNN was the lowest diagnosis rate at 20 kHz. The best average accuracy between ML techniques was RF, whereas the accuracy of SVM was the lowest in all frequency bands; however, the accuracy differences of DT, NB, RF, and KNN were pretty slight. The overall performance of LSTM was the best, whereas GRU had the lowest rate compared with DNN and LSTM techniques. All learning techniques showed similarities in accuracy and achieved high performance at 20 kHz. In the case of frequency-domain input, GRU offered the highest performance at 5, 10, and 15 kHz switching rates, whereas LSTM took first place at 20 kHz. SVM showed poor performance. Other learning algorithms showed mediocre diagnosis results at low switching rates (5 and 10 kHz), and the accuracy was improved at high switching frequencies (15 and 20 kHz) except SVM. When the time- and frequency-domain inputs were combined, the performance of KNN, RF, NB, DT was improved compared with that of frequency-domain input, whereas the accuracy of GRU and LSTM was increased with that of time-domain input. SVM and DNN performances decreased compared with that of time-domain

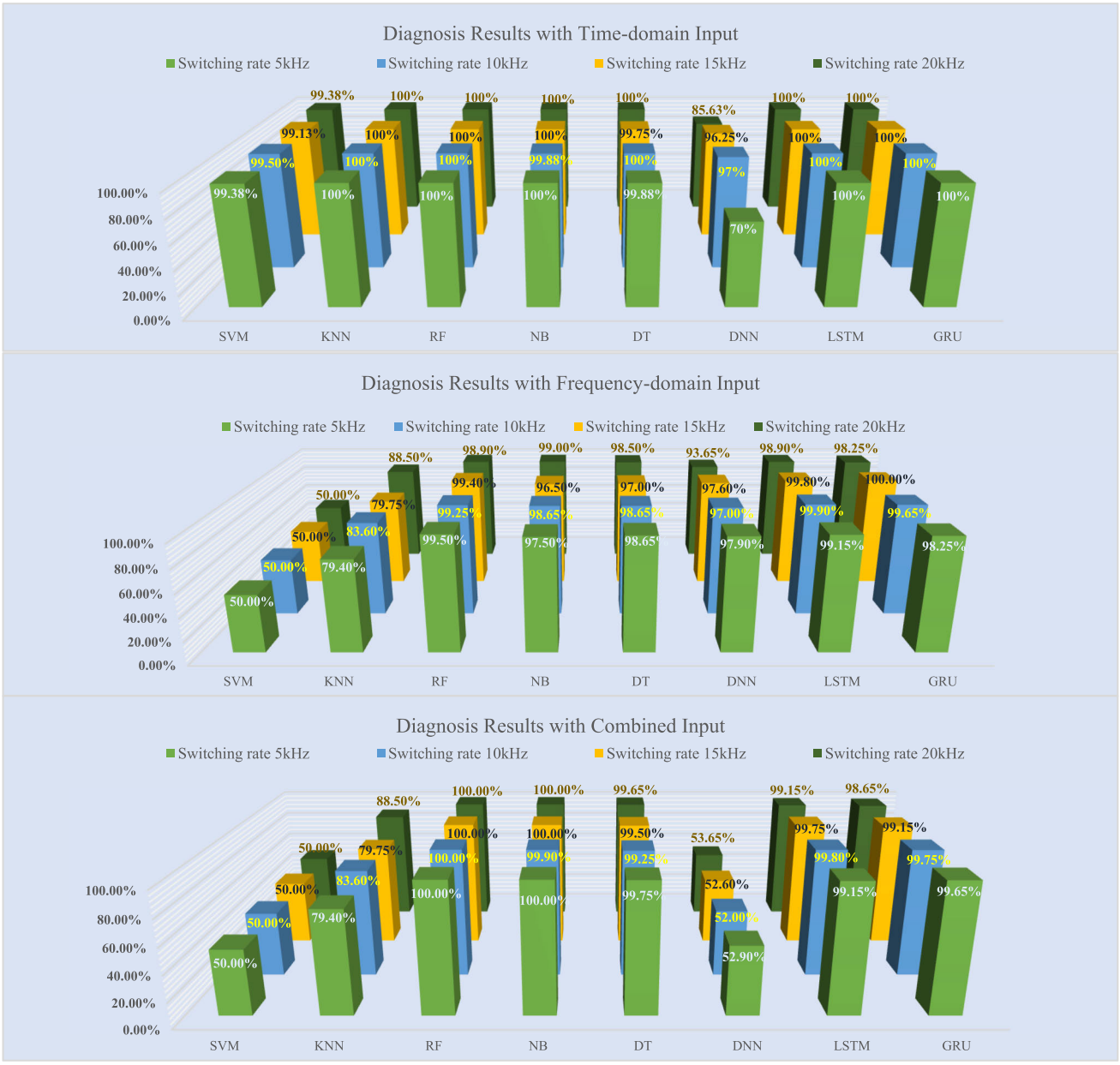

FIGURE 10. Arc diagnosis results for the three-phase inverter load with MPC (3PMPC) at 5 A under different input types.

and frequency-domain inputs, respectively. The accuracies of GRU, LSTM, DT, NB, RF, and KNN were similar with time- or frequency-domain inputs at the high switching frequency. The diagnosis results at 8 A current amplitude are presented in Figure 9. NB showed the best detection rates at 5, 10, and 15 kHz switching rates; and at 20 kHz, GRU and LSTM shared first place in terms of accuracy when the time-domain input was applied. In contrast, the performance of GRU and LSTM was poor at 5, 10, and 15 kHz; and the diagnosis rate of NB was the lowest at 20 kHz. NB showed the best average accuracy, and the diagnosis performance of SVM was lowest compared with other ML techniques in all frequency bands. However, the performance differences between DT, NB, RF, and KNN were expressively minor. The detection rate of DNN was the best at 5, 10, and 15 kHz compared with other DL algorithms, and GRU and LSTM achieved the highest performance at 20 kHz. The detection rates of three DL algorithms and SVM increased, and those of DT, NB, RF, and KNN decreased with the increase of switching rate. When the frequency-domain feature was used, three DL algorithms (GRU, DNN, and LSTM) offered the best diagnosis at 5, 10, and 15 kHz, respectively. In addition, NB and LSTM shared first place in terms of performance at 20 kHz. The performance of SVM was poor in all switching bands. In the case of combined input, RF and GRU shared first place at 5 kHz. GRU, NB, and RF offered the best diagnosis results at

FIGURE 11. Arc diagnosis results for 3PMPC at 8 A under different input types.

10, 15, and 20 kHz, respectively. The performances of SVM and DNN were poor in all switching rates.

For the nonlinear load 3PMPC, Figure 10 presents diagnosis performances at 5 A current amplitude. When the inputs in the time domain were adopted, all advanced learning algorithms, except for DNN, achieved superior performance (above 99%) in all frequency bands. The accuracy rate of SVM was the lowest among them; however, the difference in accuracy rates between all learning techniques was minimal. When the frequency feature was in use, the diagnosis rates of GRU, LSTM, DNN, DT, NB, and RF were high in all frequency bands with slight differences. In contrast, KNN and SVM showed mediocre and poor performances. When the combined input was employed, there was a similar trend as time-domain and frequency-domain inputs. The accuracy differences between GRU, LSTM, DT, NB, and RF were significantly slight. In contrast, the performances of DNN and SVM were poor in all frequency bands. Similar to Figure 10 for 3PMPC load, Figure 11 shows accuracy rates at 8 A. In the case of time-domain input, the accuracy of GRU was highest at the switching rate of 5 kHz; and at 10, 15, and 20 kHz switching frequencies, DT took first place in terms of accuracy. The performance of DNN was lowest in all switching rates. LSTM, NB, RF, and KNN's diagnosis rates were superior (upper 91%) in all switching bands. The performance of SVM was mediocre compared

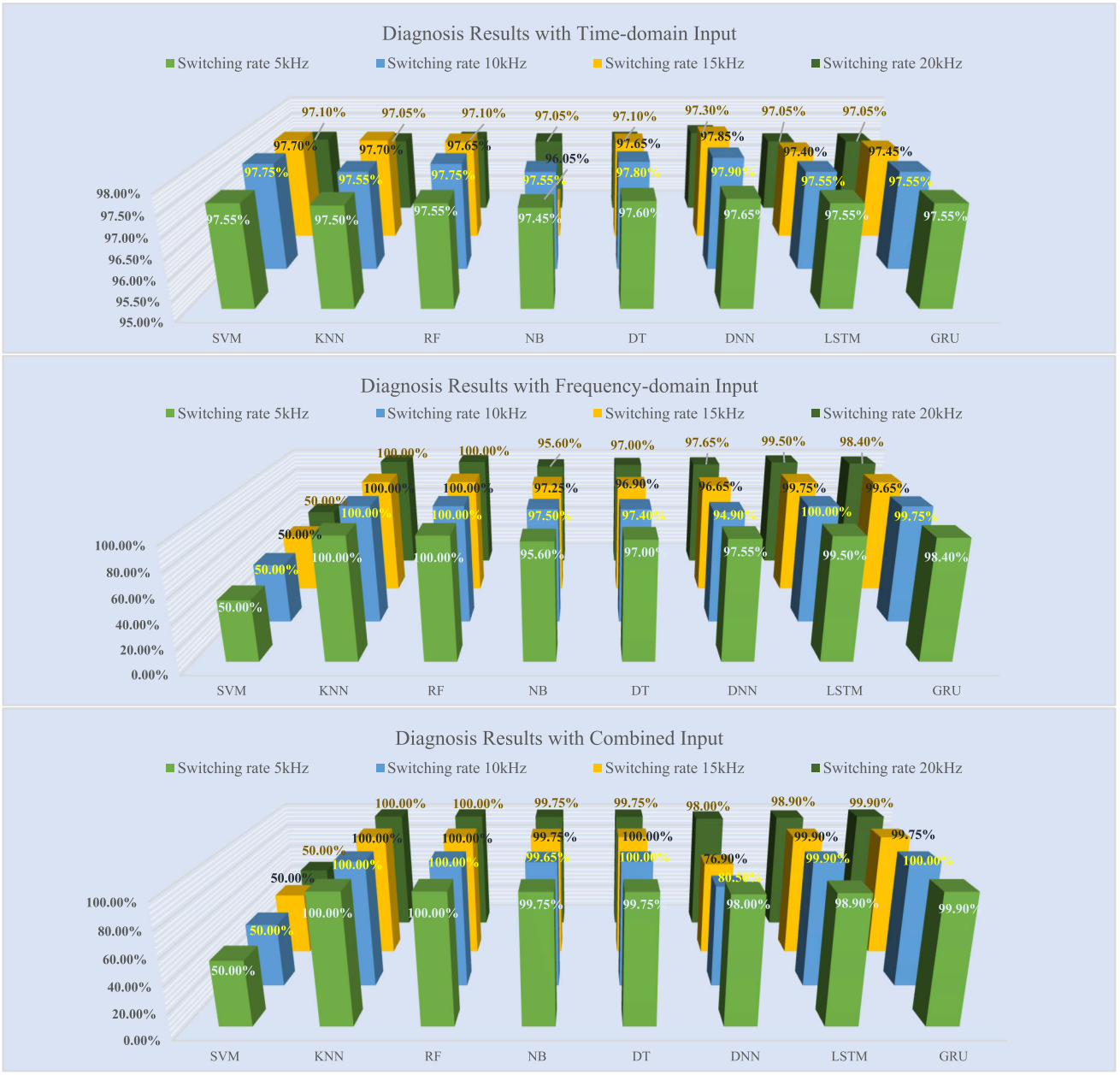

FIGURE 12. Arc diagnosis results for the single-phase inverter with SPWM at 5 A under different input types.

with those of other learning algorithms at switching rates of 5 and 10 kHz. NB, RF, KNN, and SVM showed improvements in accuracy with the increase of switching rates. In contrast, GRU, LSTM, DNN, and DT accuracies increased when the switching rate increased from 5 to 15 kHz and marginally declined at 20 kHz. DNN offered the most optimal diagnosis rates at 5 and 10 kHz when the frequency input was implemented, whereas RF took first place at 15 and 20 kHz. The performance of SVM was poor in all switching rates. When combined input was in service, DT hit the highest rate at 5 kHz, whereas the performance of RF was the best at 10, 15, and 20 kHz switching frequencies. SVM and DNN presented poor detection rates in all switching rates.

Figure 12 illustrates diagnosis results for 1PSPWM at 5 A current amplitude. When the time features were examined, all techniques achieved superior performance (upper 97%) in all ranges. The performance of DNN was the best, and the detection rate of NB was the lowest among all learning techniques. However, the difference in detection rates of all learning algorithms was insignificant. In the case of frequency-domain input, all learning techniques showed high performance with slight differences in accuracy except SVM. KNN and RF shared first place in all frequency range with maximum detection rates. A similar trend was observed with the combined input in use. The difference was that DNN showed great diagnosis results

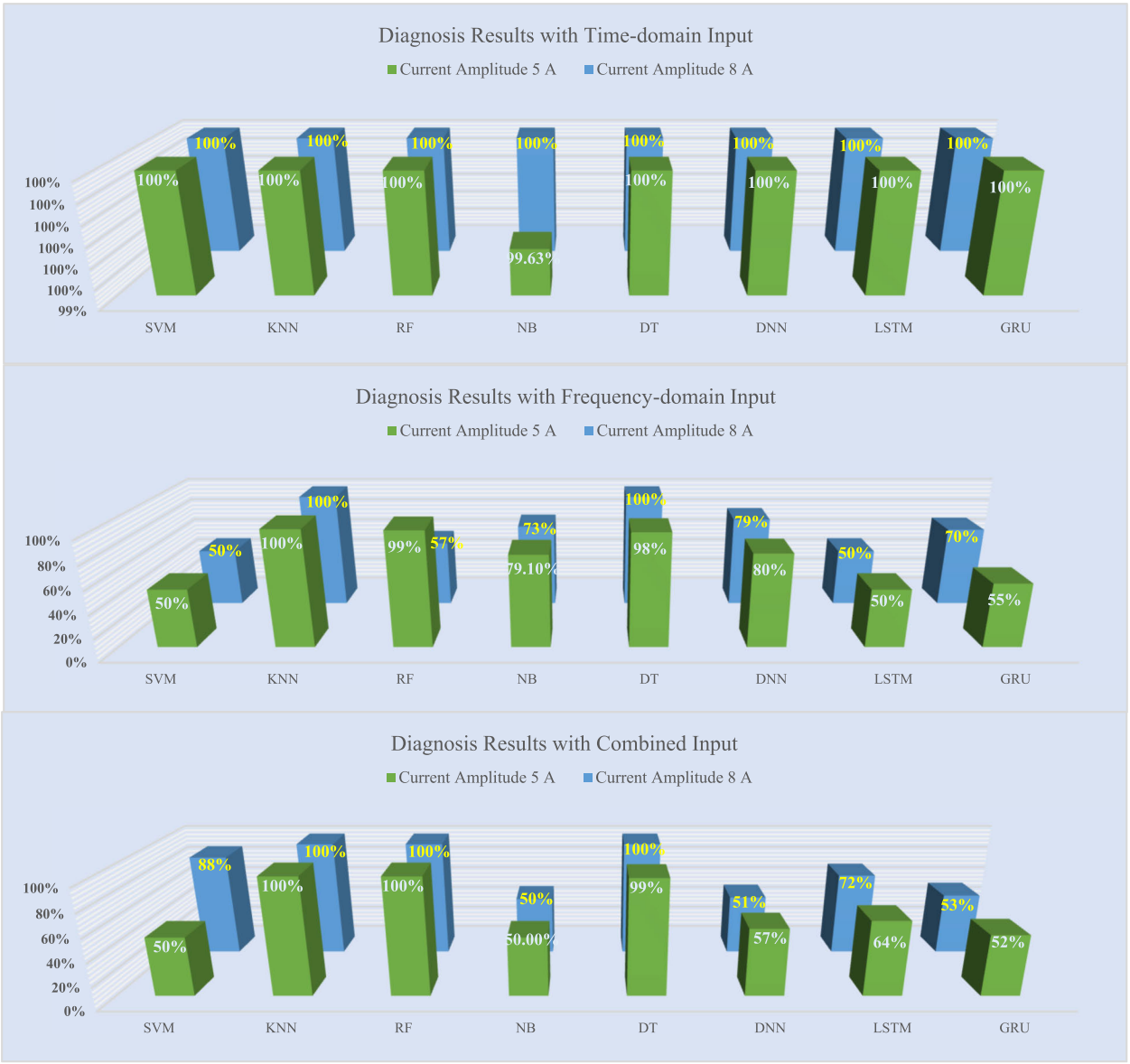

FIGURE 13. Arc diagnosis results for the resistor load at 5 and 8 A under different input types.

at 5 and 20 kHz, whereas the detection rates were lower at 10 and 15 kHz.

The accuracy rates for the resistor load at 5 and 8 A amplitudes were demonstrated in Figure 13. In the case of time-domain input in service, all AI algorithms achieved absolute detection rates (nearly 100%), and there is no missed diagnosis or wrong decisions, except for NB at 5 A. When the frequency input was entered, KNN and DT offered the best diagnosis results at 5 and 8 A. In contrast, RF only achieved a high diagnosis rate at 5A, and other learning techniques showed mediocre or poor performances. Similar performance of frequency input was observed for the combined input. Only DT, RF, and KNN offered excellent detection rates at 5 and 8 A current amplitudes.

Generally, ML techniques show high performance with time-domain input, whereas DL algorithms detect arc fault more accurately when using frequency-domain input. Some useful data could have vanished during the feature mining leading to the low performance of DL algorithms. In addition, DL algorithms contain numerous neurons and layers, which can increase the execution time and the computational cost. This could be beneficial for real-world systems, which is a priority for reliability, robustness, and cost. When the resistor is connected, the difference in FFT analysis between normal and arc states is unclear, as shown in Figure 2(c), resulting in poor performance when the frequency-domain input is applied. The advantage of ML techniques is the high performance at small data sets and simple setup. However, their limit is the requirement for the presence of features to achieve high diagnosis rates. Alternatively, DL algorithms need an extensive training set and high computational cost owing to the deeper configurations compared with

ML techniques. Furthermore, the configurations (number of neurons and layers) in DL algorithms were chosen based on the trial and error method. Numerous simulations were performed to find the best finest configurations. However, there is no manner to assurance that the chosen configurations return the best diagnosis results for all operating conditions. For example, the finest configurations of layers could be altered if the trial number is changed. Furthermore, the performance of DL algorithms differs with the working circumstances (frequency, load type, and current amplitude). This means that the optimum configurations in one case might not be ideal for other conditions. As shown in the above diagnosis analysis, DL algorithms presented better diagnosis rates compared with the other AI techniques in several conditions, whereas their diagnosis rates were mediocre or even poor in other conditions. The combined input offers a balance performance between ML and DL algorithms in all load types among three types of inputs.

V. CONCLUSION

Various possible combinations for arc fault diagnosis were tested using eight AI techniques and various frequency-domain and time-domain input parameters. DL techniques are more successful when frequency-domain input is applied than ML algorithms in all switching frequency ranges. In contrast, ML techniques are more accurate than DL algorithms when time-domain input is employed. The combined input offers a balance performance between ML and DL algorithms in all load types. Generally, the diagnosis results of all learning algorithms using time-domain or frequency-domain inputs increase with frequency increase.

Frequency-domain input is obtained from the input data using the FFT analysis. However, this feature locates in the frequency domain; and needs a high sampling rate and resource to obtain. In addition, there are many neurons and layers in deep learning structures, which could increase the computational effort and processing time compared with ML techniques. In actual applications, the computational time and accuracy could be significantly affected when an arc happens. In comparison, time-domain features could be obtained at a lower rate, resulting in fast execution. This study provides a specific view and helpful information for implementing arc fault detect mechanisms for practical systems, prioritized for various goals, such as costs, execution time, robustness, and reliability.

It is noticed that there is no learning algorithm, which achieved a high detection rate in all cases. Each algorithm reached high accuracies in different working conditions; otherwise, its effectiveness was mediocre or poor. Therefore, several learning algorithms should be combined to optimize the detection rate and maintain high accuracy in various working situations for arc fault diagnosis. Depending on the input types, the combinations of AI techniques should be different. For the time-domain input type, the ML techniques should be chosen; whereas, several DL techniques are selected for the frequency-domain input. This study provides a detailed vision of altered AI techniques. This could be a helpful investigation for choosing AI techniques, input types, feature extraction methods, which can contribute to the construction of more reliable and robust systems when executing an arc fault diagnosis system regarding altered primacies.

REFERENCES

[1] S. Lu, B. T. Phung, and D. Zhang, "A comprehensive review on DC arc faults and their diagnosis methods in photovoltaic systems," *Renew. Sustain. Energy Rev.*, vol. 89, pp. 88–98, Jun. 2018.
[2] S. Barmada, M. Raugi, M. Tucci, and F. Romano, "Arc detection in pantograph-catenary systems by the use of support vector machines-based classification," *IET Electr. Syst. Transp.*, vol. 4, no. 2, pp. 45–52, Jun. 2014.
[3] 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.
[4] Q. Xiong, S. Ji, L. Zhu, L. Zhong, and Y. Liu, "A novel DC arc fault detection method based on electromagnetic radiation signal," *IEEE Trans. Plasma Sci.*, vol. 45, no. 3, pp. 472–478, Mar. 2017.
[5] G.-S. Seo, J.-I. Ha, B.-H. Cho, and K.-C. Lee, "Series arc fault detection method based on statistical analysis for DC microgrids," in *Proc. IEEE Appl. Power Electron. Conf. Expo. (APEC)*, Mar. 2016, pp. 487–492.
[6] K.-J. Lee, G.-S. Seo, and B.-H. Cho, "DC arc fault detection method for DC microgrid using branch monitoring," in *Proc. 9th Int. Conf. Power Electron. ECCE Asia (ICPE-ECCE Asia)*, Jun. 2015, pp. 2079–2084.
[7] S. Dhar and P. K. Dash, "Differential current-based fault protection with adaptive threshold for multiple PV-based DC microgrid," *IET Renew. Power Gener.*, vol. 11, no. 6, pp. 778–790, May 2017.
[8] S. Dhar, R. K. Patnaik, and P. K. Dash, "Fault detection and location of photovoltaic based DC microgrid using differential protection strategy," *IEEE Trans. Smart Grid*, vol. 9, no. 5, pp. 4303–4312, Sep. 2018.
[9] S. Tharmakulasingam, S. Lu, B. T. 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] 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.
[11] 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.
[12] 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. Electr. Electron. Eng.*, vol. 14, no. 2, pp. 192–200, Feb. 2019.
[13] R. D. Telford, S. Galloway, B. Stephen, and I. Elders, "Diagnosis of series DC arc faults—A machine learning approach," *IEEE Trans. Ind. Informat.*, vol. 13, no. 4, pp. 1598–1609, Aug. 2017.
[14] 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.
[15] 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.
[16] 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.
[17] 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.
[18] 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 Monit. Diagnosis (CMD)*, Oct. 2020, pp. 416–421.
[19] V. Le, X. Yao, C. Miller, and B.-H. Tsao, "Series DC arc fault detection based on ensemble machine learning," *IEEE Trans. Power Electron.*, vol. 35, no. 8, pp. 7826–7839, Aug. 2020.

[20] *Outline of Investigation for Photovoltaic (PV) DC Arc-Fault Circuit Protection*, Standard UL 1699B, 2018.
[21] 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.
[22] 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.
[23] 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.
[24] 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, no. 12, pp. 1900–1909, Dec. 2021.
[25] T. Cover and P. Hart, "Nearest neighbor pattern classification," *IEEE Trans. Inf. Theory*, vol. IT-13, no. 1, pp. 21–27, Jan. 1967.
[26] L. Breiman, J. H. Friedman, R. A. Olshen, and C. J. Stone, *Classification and Regression Trees* London, U.K.: Routledge, 2017.
[27] L. Breiman, "Random forests," *Mach. Learn.*, vol. 45, no. 1, pp. 5–32, 2001.
[28] P. Langley, W. Iba, and K. Thompson, "An analysis of Bayesian classifiers," in *Proc. AAAI*. Princeton, NJ, USA: Citeseer, vol. 90, 1992, pp. 223–228.
[29] 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 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 with 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.