IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 26 August 2022, accepted 8 September 2022, date of publication 14 September 2022, date of current version 21 September 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3206547

RESEARCH ARTICLE

A Novel Series Arc Fault Detection Method Based on Mel-Frequency Cepstral Coefficients and Fully Connected Neural Network

YAO WANG[1,4], (Member, IEEE), DEJIE SHENG[1], HAIQING HU[2], KEFAN HAN[1], JIAWANG ZHOU[1], AND LINMING HOU[3]

[1]School of Electrical Engineering, Hebei University of Technology, Tianjin 300400, China
[2]National Low Voltage Electrical Apparatus Quality Supervision and Inspection Center, Wenzhou 325600, China
[3]Zhejiang Testing and Inspection Institute for Mechanical and Electrical Products Quality Company Ltd., Hangzhou 310051, China
[4]Yangtze Delta Region Center of Electrical Engineer Innovation, Wenzhou 325600, China

Corresponding author: Linming Hou (holm2020@163.com)

This work was supported in part by the Natural Science Foundation of Hebei Province under Grant E2020202204, in part by the Central Funds Guiding the Local Science and Technology Development under Grant 226Z2102G, and in part by the Zhejiang Provincial Natural Science Foundation of China under Grant LGG20E070002.

ABSTRACT Arc faults pose challenges to electric safety, which can cause serious fire hazards. However, the commonly used arc fault detection method is prone to nuisance tripping. This paper proposed a hybrid arc fault detection method based on the improved Mel-Frequency Ceptral Coefficients (MFCC) for preprocessing and a neural network model for arc identification called ARC_MFCC. As per the IEC 62606, twelve different loads/scenarios are considered for this research. An arc tangent-based core filter is employed to improve the MFCC to enhance the arc features within a bandwidth of 3 kHz to 7 kHz. A lightweight neural network model of fully connected cascaded with the MFCC-based preprocessing, which can distinguish the arc fault with normal operation under different test conditions. As a verification result, the ARC_ MFCC can achieve an accuracy of 99.34%. Moreover, the proposed method is implemented by Raspberry pie 4B. Test results show an average running time of about 4.2 ms per sample, which Ensures that the tripping time can meet IEC 62606.

INDEX TERMS Series arc fault protection, improved Mel-Frequency Ceptral Coefficients (MFCC), artificial intelligence, fully connected neural network.

I. INTRODUCTION

According to the IEA [1], global power demand growth of more than 6% is reported in 2021 and the report forecasts an average annual power demand growth of 2.7% from 2022-2024. The increasing demand for electricity has also brought more safety risks. Arc faults can occur due to aging of transmission lines, loosening of wire connection terminals, and substandard insulation materials. Arc faults can raise the local temperature up to 5000°C through gas discharge, resulting in severe electrical fires. In order to avoid arc fault, arc fault circuit breakers were required to be used in residential electrical installations in the United States in 2002 [2]. In 2013, the International Electrotechnical Commission (IEC) also proposed a protection technology called arc fault detection device (AFDD) [3].

Arc faults are generally divided into series arc faults and parallel arc faults. Parallel arc faults are easy to detect because of their high current. In series arc fault, the arc current is less than normal, and Load current especially nonlinear load current covers part of the current arc characteristics, making series arc faults more challenging to detect than parallel arc faults. Traditional AFCI has only a 61.67% success rate in series arc fault detection [4]. Therefore, scholars worldwide have done much research on series arc faults.

Traditional methods generally obtain arc characteristics by analyzing voltage and current signals in the time-frequency domain and then setting the threshold to achieve the purpose

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

of protection. Scholars such as Artale used low-frequency harmonics after Chirp Z-Transform (CZT) to identify arc faults [5]. Kim detected the series AC arc voltage waveform and then analyzed the unique symmetric energy distribution generated by the arc fault to identify the arc fault [6]. However, the installation position of the acquisition device severely limits the collection of voltage waveforms. Recently, some novel techniques such as neural network algorithms have been applied. Kai Yang proposed an arc fault identification method based on a time-domain optical convolutional neural network, using the grayscale image corresponding to the arc current's half-wave signal as the model input [7]. Wenchu Li proposed a series arc fault diagnosis and line selection method based on a recursive neural network, which achieved 98.7% diagnosis and line selection accuracy [8]. Wang Yao proposed an arc detection model ArcNet based on a convolutional neural network (CNN), which uses 10 kHz of original current data to achieve a maximum arc detection accuracy of 99.47% [9]. Joshua E. Siegel develops Deep Neural Networks (DNNs) taking Fourier coefficients, Mel-Frequency Cepstrum data, and Wavelet features as input for differentiating normal from malignant current measurements [10]. HOANG-LONG DANG chose average value, median value, variance value, RMS value, and distance of the maximum and minimum values of time-domain parameters of current for arc fault detection [11]. Na Qu uses Fast Fourier Transform (FFT) to extract ten current features in the frequency domain. The support vector machine optimized by particle swarm optimization (PSO-SVM) is designed to detect the arc fault [12]. The algorithm proposed by Edwin is based on a Kalman filter, used for identifying fault symptoms and a decision block, which confirms the presence of a series arc fault to activate a tripping signal [13]. Na Qu proposed an Online Adjustment Method of regular order p that can improve the sparsity and accuracy of the sparse representation algorithm [14].

The method based on artificial intelligence has excellent accuracy. However, implementing artificial intelligence methods in industrial embedded microcontrollers is challenging because of the complicated model, which leads to complexity of pressure calculation and requires high computing power equipment to preprocess the data in advance. In addition, other methods have used less load types than the standard requirements [3].

This paper proposes an algorithm based on the improved MFCC and neural network, which can reduce the computational pressure while ensuring high precision. The proposed method is also verified on the industrial embedded platform.

The main contributions of this paper are as follows:

1) This paper relied on an advanced arc current data acquisition device to collect and analyze the data of 12 loads. We selected the influential frequency band for arc fault detection to highlight data discrimination with or without arc current.

2) An improved MFCC arc current feature extraction method is proposed. The arctan filter bank is used to replace the original Mel filter bank, which significantly enhances the pertinence of the original method to arc fault detection.

3) An arc fault detection method ARC_MFCC is proposed by combining the artificial intelligence method with the improved MFCC method. We made various datasets for comparison to select the optimal data processing format.

4) The memory space occupied by the method layout and real-time computing pressure is reduced by transforming and compressing the model. Our method is successfully implemented on two commonly used industrial embedded microprocessors, and experimental results prove the practicability of the method.

II. DATA COLLECTION AND ANALYSIS

A. ARC FAULT EXPERIMENT PLATFORM

In practice, arc faults occur randomly, making it challenging to pinpoint their exact timing and gather the necessary data in a timely manner. Therefore, it is necessary to produce arc under various fundamental load types using a particular test platform. The arc fault experiment platform is created in accordance with the IEC 62606 standard, as seen in Fig.1.

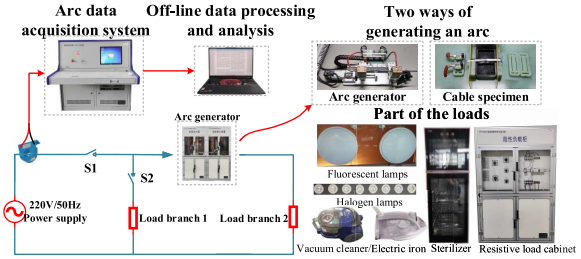


FIGURE 1. Arc data collection process and circuit arrangement.

Arcs can be generated in two different ways – one using an arc generator and the other using a cable sample. We employed both techniques to come as near to the real scenario as possible. The arc generator can replicate the actual arc generated during a slack cable connection, and the arc generated by the cable sample can simulate the arc caused by insulation breakdown in the circuit. The input of this device is powered by a 220 V, 50 Hz power supply. The two load positions can be combined with different loads according to the various experimental types in the standard to build experimental circuits. A current transformer (CT) is used to sense the current in the line. In addition, the arc signal acquisition system can automatically record the data in the experimental process and generate data files for PC offline data processing and analysis.

B. TYPES OF LOADS AND VISIBLE ARC FEATURES

As arc fault protection is mainly used in residential areas, different household appliances were selected as experimental loads. The arc current and normal load current under 12 different loads were collected in this experiment. The types of loads included are shown in the TABLE 1. The specifications of the load comply with the IEC 62606 standard [3]. The arc

TABLE 1. Load type and number.

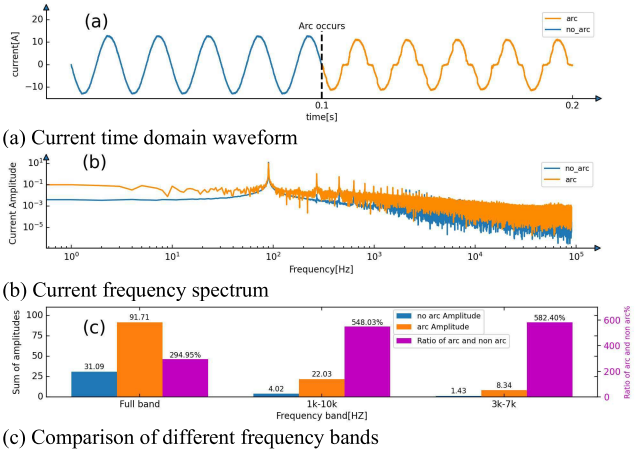
number	The load type	number	The load type
01	Resistive load	07	electric hand tool (drill)
02	air compressor	08	Electric iron
03	Vacuum cleaner	09	Electric kettle
04	Fluorescent lamp	10	air conditioner
05	Dimming light	11	Disinfection cabinet
06	electronic switch mode power supply	12	Halogen lamp

current and normal load current come from a single load and different loads connected in parallel.

Regardless of the linear or nonlinear load, the non-arc current waveform is comparatively constant over a brief observation period. However, the arc current lacks significant stability. Arc discharge will be impacted by a number of elements including load, electrode material, etc. because arc discharge is a very complex physical and chemical process in and of itself. Even in adjacent periods, the current waveform is significantly different, and no stable periodicity was even observed. As shown in Fig.3(a).

In addition, the arc discharge process changes abruptly. If the judgment time is too long, the protection may fail. Nowadays, many common arc faults protection methods take the alternating current period as the calculation unit [9], it is not easy to find the dynamic characteristics of the arc current in one AC cycle, so a method that can detect the current arc characteristics effectively and detect arc faults faster with smaller calculation units is needed.

Due to irregular arc combustion, from the perspective of the frequency domain of Fig.3(b), arc current signals are distributed within 0 to 10 kHz and the high-frequency component of the resistive load is small. In addition, the high-frequency component of the switching power supply load is prominent in the middle part of the frequency band, At the low-frequency and higher frequency bands, the spectrum amplitude of arc and non-arc signals overlap. Therefore,

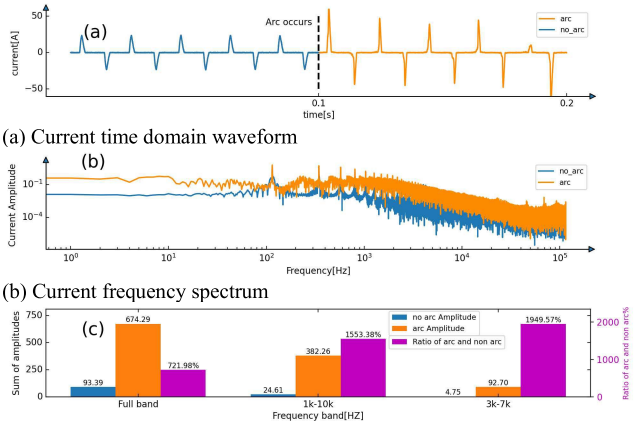


(a) Current time domain waveform

(b) Current frequency spectrum

(c) Comparison of different frequency bands

FIGURE 2. Resistive load.



(a) Current time domain waveform

(b) Current frequency spectrum

(c) Comparison of different frequency bands

FIGURE 3. Electronic switch mode power supply.

the frequency domain characteristics of the arc should be selected to exclude the influence of other load harmonics.

In order to differentiate the arc-free current spectral amplitudes of signals are added in different bands to calculate the total spectrum amplitude ratio k of arc and non-arc state. As can be seen from the calculation results in Fig.2(c) and Fig.3(c). when the resistive load is in the 1 kHz-10 kHz band, $k = 548.03\%$, much higher than 294.95% in the whole band. Under electronic switch mode power supply, In 1 kHz to 10 kHz band, $k = 1533.38\%$, and in 3 k to 7 kHz band, $k = 1949.57\%$, the distinction between arc and non-arc current is more pronounced than resistive load. 1 k to 10 kHz can cover most of the frequency domain characteristics of arc current without shielding the characteristics of an arc fault.

III. ARC CURRENT FEATURE EXTRACTION BASED ON IMPROVED MFCC

A. MFCC

MFCC is one of the most widely used speech feature parameters in speech recognition, it can effectively reflect the speech signal energy in the different frequency ranges. According to the above study, arc fault current signal has obvious frequency distribution characteristics, and analysis of characteristic frequency signals can be more conducive to arc fault feature extraction. The speech signal can be regarded as a random signal in a stationary signal in speech recognition. In arc fault detection, the normal current signal of the linear and nonlinear loads is a periodic stationary signal, while the arc fault is a sudden random signal. The arc fault is consistent with the randomness of the speech signal. In addition, the MFCC method reduces the calculation dimension and the amount of data, and these advantages also play a positive role in the extraction of arc fault features. Therefore, MFCC is improved to find a suitable method for arc fault feature extraction.

To intensify the high-frequency signal, MFCC first pre-weighted the target signal. The sound signal is divided into data of the same duration for periodic processing in the

second step of frame splitting and window processing. Then, to improve the continuity between frames, a hamming window is added. The third stage involves applying the Fourier transform to the signal to determine its energy distribution on the spectrum, followed by the modulo square of the speech signal's spectrum to determine its energy spectrum. The energy spectrum signal is then passed through the triangular Mel filter bank based on the Mel scale value in the fourth stage. Since the high-frequency range is closer to the auditory mechanism of the human ear, the Mel value increases slowly in the high-frequency range while increasing quickly in the low-frequency range [15], [16]. The energy spectrum signal from the filter bank is subjected to a logarithmic operation in the fifth step. In order to see the signal masked by the low amplitude noise, the lower amplitude components must be dragged higher than the larger amplitude component. After the discrete cosine transform, the typical MFCC characteristic parameters can be derived.

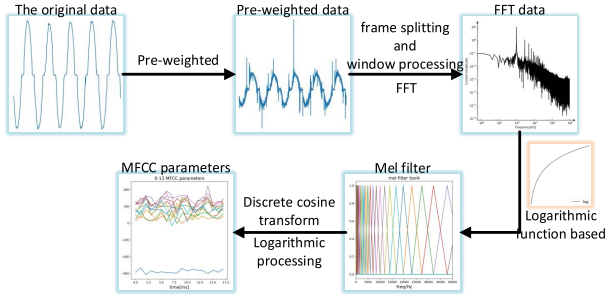


FIGURE 4. MFCC feature extraction process.

B. IMPROVED MFCC

Although MFCC and arc fault feature extraction share some similarities. But the direct application is not recommended because the inherent characteristics of arc fault (characteristic frequency band) are not exactly same as speech recognition. Therefore, it is necessary to improve the Mel filter for the MFCC core and adjust other steps appropriately to increase the applicability of the MFCC for arc fault detection. MFCC is optimized in accordance with the criteria for dependability and real-time arc fault detection. It also makes a suggestion for a faster and more accurate arc fault detection system by further utilizing its advantages of effective feature extraction and reduced data volume.

The original MFCC method uses log function mapping, as shown in Fig.5(a) Due to the human ear auditory mechanism and masking effect, the original MFCC method pays more attention to the low-frequency part of the speech signal. In the low-frequency band of regular frequency, The Mel value at low regular frequency increases much faster than that at high, so the processing of low frequency signal is emphasized. in Fig.5(b), the slope of the logarithm mapping function decreases significantly from regular low frequency to high. The MFCC method pays more attention to the signal at the frequency band with the higher slope of the mapping function. Mel frequency mapping is as follows:

$$mel = m \times lg(1 + \frac{f}{700}) \tag{1}$$

Generally, $m = 2595$.

The characteristic frequency band of an arc fault is 3 K to 7 kHz, and the frequency band of 1 K to 10 kHz can cover most characteristic signals of an arc fault. In order to match the inherent characteristics of arc faults, The signal on the characteristic frequency band shall be mainly processed. The low frequency and high frequency signals outside the feature band will shield the current arc feature and disturb the feature extraction. Consider the case of Mel mapping, the mapping function with the highest slope at the characteristic frequency band should be selected, and the slope should decrease as it gets closer to the interference frequency band. Common functions such as Arctan trigonometric functions, Sigmoid, and Tanh follow this rule. Therefore, function mapping relations based on Arctan, Sigmoid, and Tanh functions are constructed as follows:

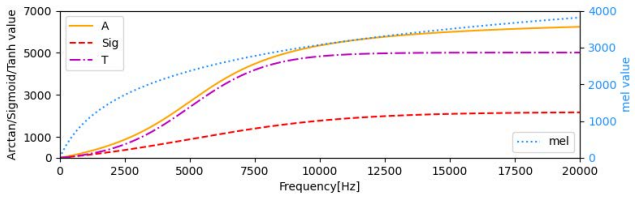
$$A = m \times (arctan\left(\frac{(f - f_0)}{S}\right) + D_a) \tag{2}$$

$$Sig = m \times (\frac{1}{1 + e^{\frac{-(f-f_0)}{s}}} + D_{Sig}) \tag{3}$$

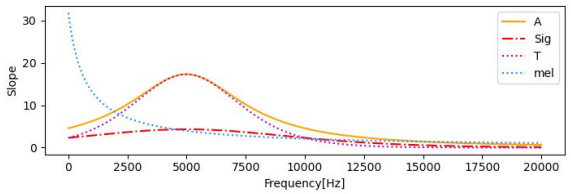
$$T = m \times (\frac{e^{\frac{(f-f_0)}{s}} - e^{\frac{-(f-f_0)}{s}}}{e^{\frac{(f-f_0)}{s}} + e^{\frac{-(f-f_0)}{s}}} + D_T) \tag{4}$$

A, Sig and T are the mapping values under the three functions, respectively. m is the magnification is to enlarge the three mappings to the amplification multiple taken under the same order of magnitude as Mel frequency. f is the regular frequency value, f_0 is the regular frequency amplification center value; S controls the change rate of the slope of the mapping function. In order to ensure the validity of the comparison, $m = 2595$, $f_0 = 5000$, and $S = 3000$ are all taken in the three mappings. D_a, D_{Sig} and D_T are offsets to ensure that the mapping value grows from zero regular frequency. These values do not affect the critical slope parameter of the mapping relation. The three mappings are shown in Fig.5(a).

In Fig.5(b), arctan, Sigmoid and Tanh functions all have the maximum slope at a certain point and maintain a large slope near this point. The farther away from this point, the slope decreases significantly, matching the inherent characteristic of the feature frequency band for arc fault feature extraction. However, it can be seen from the fig.5(b) that Sig based on Sigmoid has a slow convergence [17], resulting in a gentle slope change, and its effect of highlighting the magnification is much weaker than other mapping relationships. In addition, a large number of operations were carried out on the three mapping relations, respectively, and then the average time of multiple operations was obtained. This calculation is based on PC. The operation time of Sig is 2.6 times that of A. The operation time of T is 10.32 times that of A and 3.96 times that of Sig. The reason for Sig's long operation time is due to the complex exponential operation of natural constant e in



(a) Correspondence between frequency and mapping value



(b) Correspondence between frequency and the slope of the waveform in (a)

FIGURE 5. Comparison of Mel and arctan/Sigmoid/Tanh.

Sig. There are more such operations in T, resulting in a longer operation time. If the operation time is not considered, T can achieve similar effects to A, which is also a good choice. If we consider the real-time computing effect of the method and its implementation on the real-time computing platform with low computational power, there is no doubt that A is the most competitive. The specific situation is shown in TABLE 2.

TABLE 2. Arctan/Sigmoid/Tanh contrast.

The basic mapping function	The rate of convergence	Average time per 100,000 operations[ns]	Out reason
Arctan	Fast	912500	-
Sigmoid	Slow	2375700	Convergence is slow and takes a long time
Tanh	Fastest	9418700	Time is too long

After determining the mapping relation of the Arctan function, the functions of each parameter of A are further introduced in detail:

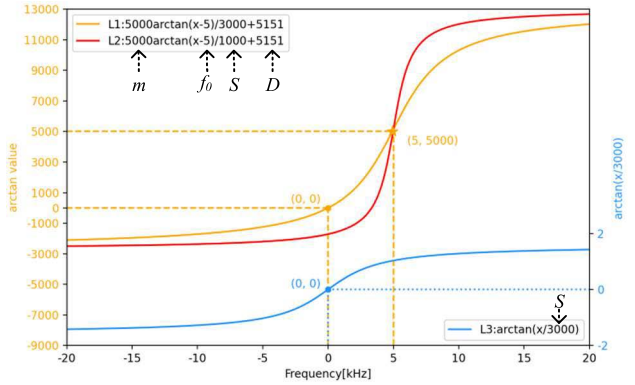


FIGURE 6. Effect of arctan function mapping parameters.

Where f_0 determines the regular frequency center band position of amplification in the arctan domain. The frequency band near the center corresponds to a broader Arctan domain. The characteristic frequency signal can be effectively amplified, and the non-characteristic frequency signal can be suppressed. The excessively high-frequency and low-frequency bands can shield and interfere with arc fault feature extraction, even though their frequency range is wide, their corresponding range in the arctan domain is narrow. Therefore, f_0 should be selected in the center of the characteristic frequency band of the arc fault. Because the characteristic frequency band is 3 kHz-7 kHz, 5 kHz can be selected as the frequency center value.

By controlling S, the change speed of arctan value can be changed. When S increases, the Arctan curve flattens, and the effective amplification frequency band becomes broader. The smaller the S is, the steeper the Arctan curve becomes, and the effective amplification frequency band becomes narrower, but the band's resolution near the center frequency is enhanced. If we want to carry out feature extraction on the central frequency band more centrally and reduce the influence of the interference frequency bands at both ends, the S value can be reduced. Generally, if we focus on the current data of the 1 kHz-10 kHz frequency band for arc feature extraction, S is about 3000.

In order to make the arctan value grow from zero with the regular frequency, eliminate the negative value and simplify the calculation, D_a is computed as follows:

$$D_a = -arctan\left(\frac{0 - f_0}{s}\right) \tag{5}$$

m is used to magnify arctan data and align them with regular frequency values at the center frequency for easy comparison and reference. Its value changes with the change of the central frequency. Refer to the value of f_0, $m = 5000$. In Fig.6, the functions of f_0, D_a, and m are reflected by L1 and L2, and L1 and L3 reflect the functions of S.

With the above correspondence, the arctan filter needs to be constructed, just like the Mel filter of MFCC. It needs to take a group of equidistant arctan points between the highest and lowest arctan values in the arctan domain. Where the lowest arctan value is set to 0, The highest arctan value:

$$H_{arctan} = m \times (arctan\left(\frac{\left(SR/2 - f_0\right)}{S}\right) + D) \tag{6}$$

where SR is the signal sampling rate.

The regular frequency points corresponding to arctan points can be obtained through the inverse transform between arctan value and regular frequency value:

$$f = S \times tan\left(\frac{a}{m} - D\right) + f_0 \tag{7}$$

Draw the arctan filter in two domains respectively, as follows:

As shown in fig.7(b), the arctan filter in the regular frequency domain is a set of filters with the center frequency

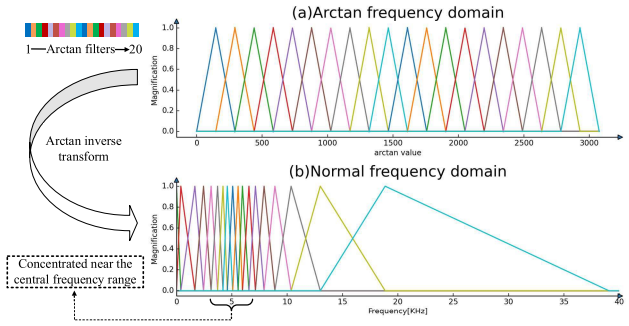


FIGURE 7. Arctan filters distribution in different domains.

(5 kHz) as the center, and the width of the filter gradually widens with the increase or decrease of the frequency.

The Arctan filter can smooth the spectrum, eliminate the high harmonics, and highlight the formant of the signal. Therefore, the amplitude of a signal is not be presented within the improved MFCC parameters, which the arc fault feature extraction characterized by the improved MFCC is not be affected by the change in the input signal amplitude. There is also a significant reduction in the amount of computation.

Since each filter can produce one MFCC parameter, the standard MFCC comprises multiple MFCC parameters. From the perspective of regular frequency, more MFCC parameters are generated in the central frequency band, so the part in the central frequency band of the signal is considered emphatically. The influence on the frequency band that interfere with the feature extraction of the arc fault is weakened, which is in line with the inherent characteristics of the arc fault.

Arc fault protection standards IEC 62606 require the real-time detection of arc fault (minimum detection time of series arc 0.12 s). Given this, the following improvements to the original MFCC method are proposed.

Low voltage AC signal has a natural frequency cycle (50 Hz/20 ms), some existing AC arc fault detection methods are conducted using the AC period as the data process time unit [9]. In order to better embody the change of current arc characteristics within a period, we choose the 0.1 AC frequency period (2 ms) as the window length. The window slides with the minimum indexing value of the AC frequency period (1 ms), the data overlap rate is 50%, as shown in Fig.8.

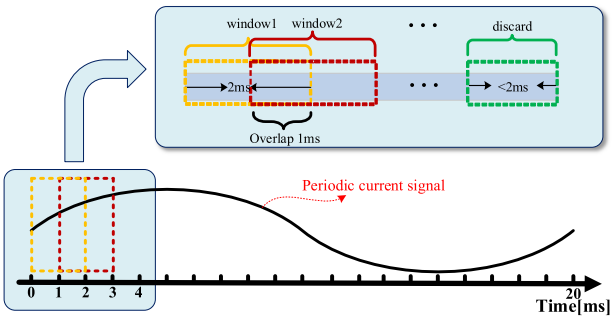


FIGURE 8. Data frame and window processing.

Redundant data will not be used, and there won't be a lot of calculating strain(pressure).

Standard MFCC parameters can only reflect the static characteristics of the signal. Arc faults are highly random, and their dynamic characteristics change dramatically due to the influence of various factors on the combustion state [18]. For random signals, the combination of dynamic and static characteristics can effectively improve the recognition effect. The combination of dynamic and static characteristics can effectively improve the recognition effect [19]. The difference calculation can be performed using the following formula:

$$d_t \begin{cases} C_{t+1} - C_t, & t \le K \\ \frac{\sum_{k=1}^{K} k(C_{t+k} - C_{t-k})}{\sqrt{2\sum_{k=1}^{K} k^2}}, & others \\ C_t - C_{t-1}, & t > Q - K \end{cases} \tag{8}$$

where t represents the serial number of dynamic parameters, dt and C_t represent the first-order difference and the Ceptral coefficient when the serial number is t. Q represents the order of the Ceptral coefficient. K represents the time difference of the derivative, which can be taken as 1 or 2.

IV. METHODOLOGY: ARC_MFCC

Arc fault detection has high real-time, high reliability, and fulfill the distribution requirements. In order to meet the requirements, the proposed method must be implemented on Low computing power platform, otherwise it will not have significant impact on the practical applications. The improved MFCC can extract the characteristics of arc faults according to their inherent characteristics. However, MFCC is generally composed of multiple parameters, it is difficult for human eyes to find the law of its characteristics, so it is difficult to set the judgment standard artificially, and the accuracy is also difficult to guarantee.

A new method is needed to use the improved MFCC parameters to achieve arc fault detection. At present, the most advanced algorithm uses a deep neural network to identify series fault arcs. Deep learning is a new research direction in machine learning. Deep learning has excellent adaptive characteristics and unparalleled data processing power, suitable for processing improved MFCC parameters. Therefore, neural network is selected to cooperate with the improved MFCC to identify and judge the arc fault features.

A. DATA PREPROCESS AND MAKE THE ARC_MFCC DATASET

The 100 kHz raw data collected by the arc data acquisition system is considered as a first-level data set. The most fundamental prerequisite for using deep learning algorithms for training and recognition is that all input data must be consistent. As a result, the original data must be divided into equal-length segments and processed. The inherent power frequency period (20 ms) of ac electrical signals makes them a good choice.

The MFCC approach can attempt to overcome the circumstance where the power frequency period is the smallest Data

processing unit because it relies on the spectrum analysis of the signal but disregards its time-domain waveform. After the arc fault occurs, some phenomena such as flat shoulder and high current slope at zero crossing point can occur[citation]. The 20 ms window must include all the characteristics of an arc fault, including two flat shoulder phenomena. The 10 ms window guarantees that it includes a complete flat shoulder or part of each of the two flat shoulders [20]. The duration of the two flat shoulder parts is still a complete flat shoulder phenomenon duration. There are still enough visual arc features in the signal segment. It is worth mentioning that 10 ms is the minimum window length to ensure that the window covers a complete flat shoulder phenomenon duration. Therefore, the secondary data set was made by data segmentation of 10 ms.

The improved MFCC method is used to process the data of the secondary data set. The final ARC_MFCC data set is obtained. The final dataset consists of 56,420 samples, including 29,045 arc samples and 27,375 non-arc samples.

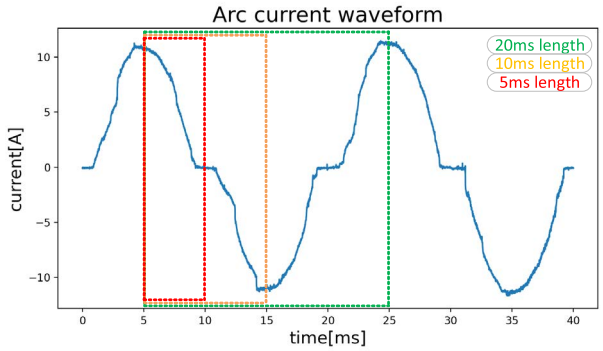


FIGURE 9. Comparison of different data segmentation duration.

B. NEURAL NETWORK MODEL

The improved MFCC increases data processing steps, which will consume extra computing time [21]. However, the arc fault features are effectively extracted from the data processed by the improved MFCC, it is not necessary to entirely rely on the neural network model to extract features, so a relatively simple network model can be adopted. Convolutional pooling and other operations can significantly increase the amount of computation. If such steps can be omitted in the simplified network model, the computing time can be effectively reduced [10], [22].

A deep learning neural network model is proposed based on the fully connected network. The improved MFCC parameters are used as the neural network input to carry out adaptive recognition of arc fault features. In order to enable the data to be input to the fully connected layer, a layer of flattened processing is required for the data. Then the model uses a fully connected 4-layer neural network, in which the first three layers with 128, 64, and 32 neurons, respectively. The fourth fully connected layer uses the Softmax function to achieve multi-state classification. The input signal passes through the influence of the corresponding weights and Relu function of the first three fully connected networks and then passes to the next layer. The TensorFlow framework implements the arc fault detection model.

This section describes experimental results to evaluate the performance of the proposed ARC_MFCC method. Databases containing both Arc and non-arc samples were used in all experiments. Samples from the test dataset do not appear in the training or validation dataset.

ARC_MFCC model is trained with the ARC_MFCC dataset. Although the amount of data is relatively large, it can be automatically marked with and without an arc. The data set was randomly split, 75% of the data was used for training, 15% for testing, and 10% for validation. The number of labels corresponding to the load and their respective training, verification, and test samples is shown in TABLE 3.

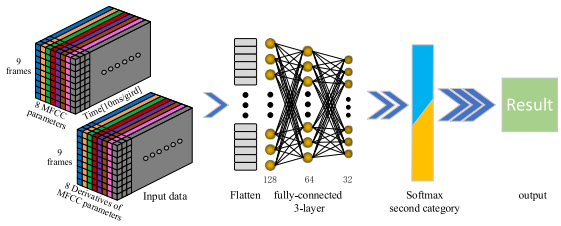


FIGURE 10. The network architecture of the ARC_MFCC.

The neural network model of ARC_MFCC is implemented on the TensorFlow backend using Keras. The model uses an adaptive learning rate strategy to train 200 iterations at a batch size of 100 and monitor ''Validation loss'' at an adaptive learning rate of 10 patience. The initial learning rate is set to 0.0001, and the minimum learning rate is 0.00001. The training and validation accuracy diagrams are shown in Fig11. According to the arc and normal load current classification results, the maximum arc fault detection accuracy of ARC_MFCC is 99.34%. Fig.12. provides the confusion matrix of ARC_MFCC for arc fault detection results.

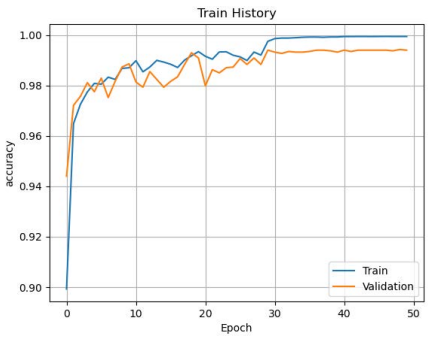


FIGURE 11. Training accuracy chart.

The number in the square represents the proportion of each part. 0.90% of the samples with arc were judged to have no arc, and 0.40% of the samples without arc were judged to have an arc. The color of the square represents the number of

TABLE 3. ARC_MFCC dataset.

Load type	Signal types	The label	Training Samples	Test Samples	Validation Samples	Total Samples
Resistive load	arc	0	3436	687	458	4581
	normal	1	3461	692	462	4615
air compressor	arc	0	1876	375	250	2501
	normal	1	1733	347	231	2311
Vacuum cleaner	arc	0	1171	234	156	1561
	normal	1	1962	392	262	2616
Fluorescent lamp	arc	0	2639	528	352	3518
	normal	1	1823	365	243	2431
Dimming light	arc	0	818	164	109	1091
	normal	1	1793	359	239	2391
electronic switch mode power supply	arc	0	4343	869	579	5791
	normal	1	1628	326	217	2171
electric hand tool (drill)	arc	0	750	150	100	1000
	normal	1	1756	351	234	2341
Electric iron	arc	0	1460	292	195	1946
	normal	1	1456	291	194	1941
Electric kettle	arc	0	1517	303	202	2023
	normal	1	1010	202	135	1346
air conditioner	arc	0	1500	300	200	2000
	normal	1	1808	362	241	2411
Disinfection cabinet	arc	0	518	104	69	691
	normal	1	346	69	46	461
Halogen lamp	arc	0	1756	351	234	2341
	normal	1	1756	351	234	2341
Total number of samples=			42316	8464	5642	56420

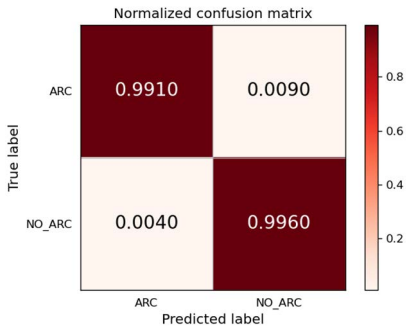


FIGURE 12. Confusion matrix.

TABLE 4. Precision, recall, & accuracy of ARC_MFCC.

		Predicted Class		Total
		Arc	Normal	
Actual Class	Arc	2879	26	2905
	Normal	11	2726	2737
	Total	3477	2882	5642
Precision=		99.62%		
Recall=		99.10%		
Accuracy=		99.34%		

samples in proportion to the population. The darker the color, the bigger the proportion.

In addition, models can be evaluated using performance metrics such as accuracy, precision, recall, and Fβ scores in machine learning-based classification systems. These performance indicators were calculated using the parameters shown in TABLE 4. The accuracy of ARC_MFCC detection is 99.62%, the Recall is 99.10%, and the accuracy of arc detection is 99.34%. The F0.5 score of the model was 99.52%, and the F0.95 score was 99.37%. The high accuracy, recall rate, and accuracy values verify that ARC_MFCC is a good model for arc fault detection.

C. COMPARISON OF DIFFERENT DATASETS

According to the above section, the shorter the data segmentation length is, the fewer arc fault features are included. In addition, when the segmentation length is less than 10 ms, any part of the flat shoulder phenomenon may not be included in the window due to the location. However, a shorter data segmentation means faster arc fault response and more arc fault judgment times. Five data sets are created, as shown in TABLE 4, to test the effects of various data segmentation lengths and the presence or absence of dynamic parameters on ARC MFCC arc fault detection methods.

Four different data sets based on the improved MFCC method are obtained through different data segmentation and whether dynamic parameters are added. Data sets processed by the original MFCC method are also made for comparison. The models are retrained for different datasets, and the hyper-parameters are fine-tuned to suit each dataset.

As the length of data segmentation decreases, the decrease in accuracy is evident and expected. Because the decrease of data segmentation length leads to the loss of some critical arc features in the window, which affects feature extraction and makes the arc fault features less evident, it is easy to confuse with non-arc samples. Secondly, from the perspective of identification accuracy, the addition of dynamic parameters also improves the method's accuracy, which also conforms to the randomness of arc faults. In TABLE 5, the decrease in accuracy is generally accompanied by a decrease in the amount of data per frame, which also means that the model can calculate results faster and the calculation pressure is low. Since complete identification accuracy has already been attained in the case of 10 ms, the accuracy does not considerably increase even though the amount of data in each window grows with the addition of 20 ms. The parameter composition of the standard enhanced MFCC parameters combined with dynamic parameters under the 10 ms is best appropriate for the ARC MFCC method when accuracy and calculation pressure are taken into account holistically.

TABLE 5. Accuracy under different datasets.

Data segmentation length	kinds of dataset	Data points per sample	Recognition accuracy
5 ms	Standard Improved MFCC	32	90.97%
	Standard Improved MFCC + first derivative	64	92.29%
10 ms	Standard Improved MFCC	72	93.29%
	Standard Improved MFCC + first derivative	144	99.34%
	Original MFCC	72	74.12%
20 ms	Standard Improved MFCC	152	98.08%
	Standard Improved MFCC + first derivative	304	99.53%

V. IMPLEMENTATION ON REAL-TIME COMPUTING PLATFORM

Arc faults may occur in any position of the power grid at any time, and their randomness is extreme. Therefore, if we want to effectively reduce arc fault damage, we need to widely install AFDD, which called the distributed nature of arc fault protection. Since it needs to be used in large quantities, the cost of a single AFDD can't be too high. Therefore, the arc fault protection method should realize accurate arc fault detection on an embedded real-time computing platform with low cost, low computation power, and small volume. If the method can't meet the above requirements, its practical application value will be significantly reduced.

At the beginning of the design, the ARC_MFCC method takes reducing the calculation pressure and ensuring accuracy as the basic requirements. In order to demonstrate its practicability, the implementation scheme based on the raspberry PI 4B embedded platform are proposed. The above experiments show that the 10 ms segmentation length can ensure high accuracy of arc fault detection and will not lead to excessive data volume. Therefore, the method verification based on the embedded platform adopts the 10 ms data segmentation length.

A. IMPLEMENTED ON RASPBERRY PI 4B

The raspberry PI 4B features a high-performance 64-bit quad-core processor, up to 4 GB of RAM, and a wealth of IO ports, enabling it to quickly and flexibly process data and control peripheral devices.

The standard neural network model is complicated and takes a long time. TensorFlow Lite provides support for running the TensorFlow model on a variety of devices. TensorFlow Lite can be used to convert the model to a particular format that reduces model file size and introduces optimizations without compromising accuracy, thus reducing model computation time without compromising accuracy.

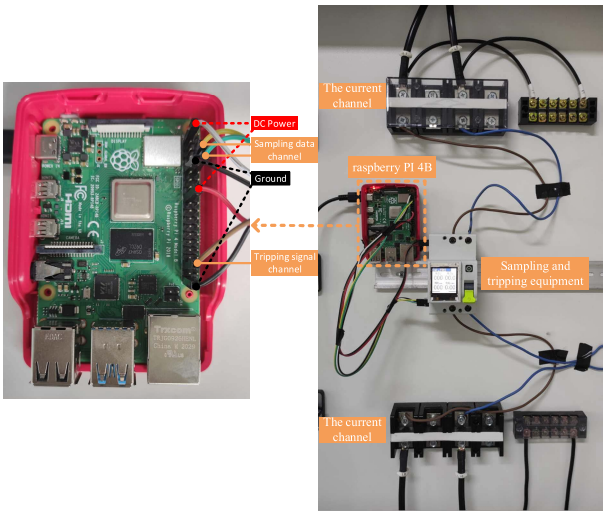
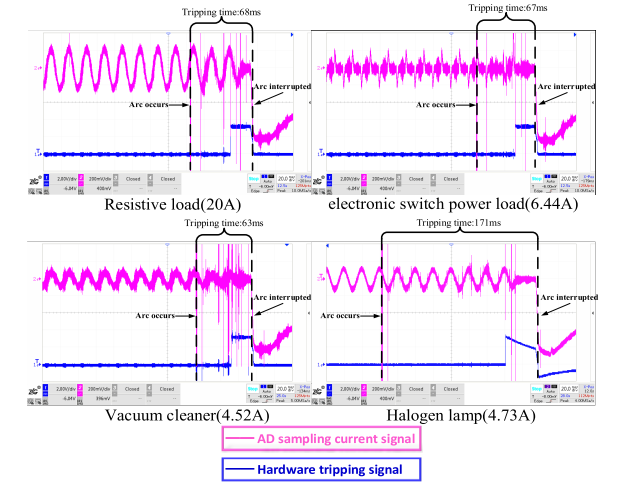
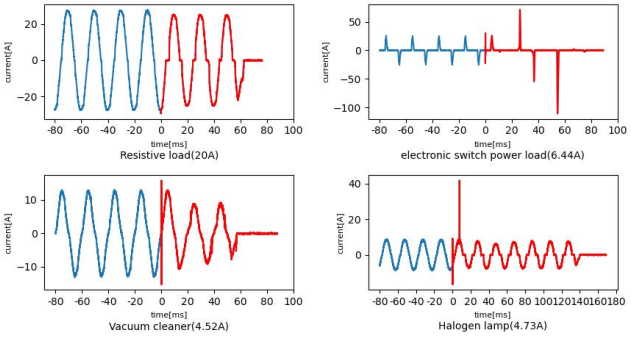


FIGURE 13. Hardware wiring diagram based on raspberry PI 4B.

In reality, series fault arc is usually caused by cable insulation aging damage, poor contact, or broken line [23], and carbonized cables can better mimic the situation of arc fault in the actual process [24]. Based on using the arc generator, we also used the cable equipment to generate arc to verify our hardware equipment, which can be more suitable for the actual arc fault scenario. When the raspberry PI 4B detects an arc fault, it can send a trip signal to the external tripper. Each sample takes an average of 4.2 ms to run the test on the raspberry PI 4B, indicating that it can process all received data in time without causing data loss. The twelve loads in the



(a) Current signal and trip signal sampling and tripping equipment



(b) Current waveform acquired on the arc data acquisition system

FIGURE 14. Actual arc fault test diagram.

TABLE 6 represent four typical load types (resistible load, Power electronics load, motor type load, and gas discharge load), all of which can trip within the IEC 62606 standard time requirements.

B. COMPARISON OF ARC_MFCC ATTRIBUTES WITH PRIOR METHODS

The ARC_MFCC method is compared with the existing methods from the following aspects:

TABLE 6. Trip times for different loads.

number	Load type	Current [A]	Tripping time[s] 1	2	3
01	Resistive load	2.61	0.541	0.288	0.405
		5.03	0.257	0.330	0.343
		17.13	0.044	0.042	0.045
		30.46	0.033	0.032	0.039
02	air compressor	5.51	0.070	0.068	0.090
03	Vacuum cleaner	4.68	0.055	0.062	0.057
04	Fluorescent lamp	3.26	0.141	0.146	0.119
05	Dimming light	2.66	0.163	0.169	0.128
06	electronic switch mode power supply	6.69	0.068	0.056	0.062
07	electric hand tool (drill)	5.99	0.048	0.045	0.054
08	Electric iron	4.56	0.040	0.045	0.043
09	Electric kettle	8.20	0.042	0.055	0.044
10	air conditioner	6.82	0.070	0.081	0.078
11	Disinfection cabinet	3.65	0.045	0.046	0.053
12	Halogen lamp	5.02	0.146	0.129	0.142

1) Can the detection method detect arc faults under more loads? Is high accuracy achieved even under high-frequency signal interference?
2) Whether the layout of the method on the embedded platform is realized, and whether the arc fault can be detected and the arc current can be cut off in time under the actual situation? Furthermore, offer more than one option?
3) Is the method simple in practical application? Do additional hardware support and algorithms need to be implemented?

The respective characteristics of the ARC_MFCC method and other methods are illustrated in TABLE 7. The ARC_MFCC method still guarantees an accuracy of 99.34% even with 12 loads from standard requirements and common household appliances, providing more excellent protection and closer to the actual situation than other methods. Thanks

TABLE 7. Comparison of attributes with prior methods.

Particulars	Wenchu Li *et al.* [8]	Kai Yang *et al.* [7]	Joshua E. Siegel *et al.* [10]	ArcNet (previous work) [9]	ARC_MFCC (proposed)
Method framework	Based RNN proposed a series arc fault diagnosis and line selection method	TDV-CNN (temporal domain visualization convolutional neural network methodology)	Deep Neural Networks (DNNs) taking DWT, MFCC, and DFT as input	Applied CNN based network directly to the raw data without processing it	Improved MFCC combined with neural network
Max. detection accuracy	98.7%	98.7%	99.95%	99.47%	99.34%
Number of types of loads	1(Mainly for motor)	5	4	8(in 4 groups)	12
Suitable for industrial MCU	No	No	Raspberry Pi 3B	Raspberry Pi 3B	Raspberry Pi 4B
Runtime on industrial MCU	-	-	18 ms (200 ms sampling time)	31 ms (20 ms sampling time)	4.2 ms (10 ms sampling time)

to the lower computational pressure, we can successfully layout the ARC_MFCC method on the real-time computing platforms. The arc current can be faster detected and cut off in actual application, which provides a foundation for others to apply this method or make improvements. Compared with our previous work (ArcNet), this method does not need the zero-crossing reference signal of the current cycle. It does not need to mark the load type of samples, so it is more convenient for making data sets and practical applications.

VI. CONCLUSION

In this paper, a hybrid arc fault detection method is proposed, which is based on the improved MFCC, and a fully-connected neural network model called ARC_MFCC. Arc current signals are analyzed in time and frequency domains under twelve test conditions according to IEC 62606. As a result, the current components within a bandwidth of 3 kHz to 7 kHz are used for arc fault detection. An arc tangent-based core filter is employed for the MFCC to make it focus on the arc features. Moreover, a data segmentation window of 10 ms and first-order derivative processing is applied for the MFCC-based algorithm, which can improve the computation efficiency. A lightweight fully-connected neural network of three layers is followed to distinguish the arc fault from normal operation. As a verification result, the ARC_MFCC can achieve an accuracy of 99.34%. Finally, the proposed method is implemented by Raspberry pie 4B. Test results show an average running time of about 4.2 ms, which can meet the protection requirements of IEC 62606.

REFERENCES

[1] IEA. (2022). *Electricity Market Report—July 2022*. Paris. [Online]. Available: https://www.iea.org/reports/electricity-market-report-july-2022

[2] K. J. Lippert and T. Domitrovich, "AFCIs—From a standards perspective," *IEEE Trans. Ind. Appl.*, vol. 50, no. 2, pp. 1478–1482, Mar./Apr. 2014.

[3] *General Requirements for Arc Fault Detection Devices*, Standard IEC 62606, 2013.

[4] P. Qi, S. Jovanovic, J. Lezama, and P. Schweitzer, "Discrete wavelet transform optimal parameters estimation for arc fault detection in low-voltage residential power networks," *Electr. Power Syst. Res.*, vol. 143, pp. 130–139, Feb. 2017.

[5] G. Artale, A. Cataliotti, V. Cosentino, D. Di Cara, S. Nuccio, and G. Tinè, "Arc fault detection method based on CZT low-frequency harmonic current analysis," *IEEE Trans. Instrum. Meas.*, vol. 66, no. 5, pp. 888–896, May 2017.

[6] J. C. Kim, D. O. Neacsu, B. Lehman, and R. Ball, "Series AC arc fault detection using only voltage waveforms," in *Proc. IEEE Appl. Power Electron. Conf. Expo. (APEC)*, Mar. 2019, pp. 2385–2389.

[7] K. Yang, R. Chu, R. Zhang, J. Xiao, and R. Tu, "A novel methodology for series arc fault detection by temporal domain visualization and convolutional neural network," *Sensors*, vol. 20, no. 1, p. 162, Dec. 2019.

[8] W. Li, Y. Liu, Y. Li, and F. Guo, "Series arc fault diagnosis and line selection method based on recurrent neural network," *IEEE Access*, vol. 8, pp. 177815–177822, 2020.

[9] Y. Wang, L. Hou, K. C. Paul, Y. Ban, C. Chen, and T. Zhao, "ArcNet: Series AC arc fault detection based on raw current and convolutional neural network," *IEEE Trans. Ind. Informat.*, vol. 18, no. 1, pp. 77–86, Jan. 2022.

[10] J. E. Siegel, S. Pratt, Y. Sun, and S. E. Sarma, "Real-time deep neural networks for internet-enabled arc-fault detection," *Eng. Appl. Artif. Intell.*, vol. 74, pp. 35–42, Sep. 2018, doi: 10.1016/J.ENGAPPAI.2018.05.009.

[11] 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, doi: 10.1109/ACCESS.2021.3115512.

[12] N. Qu, J. Zuo, J. Chen, and Z. Li, "Series arc fault detection of indoor power distribution system based on LVQ-NN and PSO-SVM," *IEEE Access*, vol. 7, pp. 184020–184028, 2019, doi: 10.1109/ACCESS.2019.2960512.

[13] E. Calderon-Mendoza, P. Schweitzer, and S. Weber, "Kalman filter and a fuzzy logic processor for series arcing fault detection in a home electrical network," *Int. J. Elect. Power Energy Syst.*, vol. 107, pp. 251–263, May 2019, doi: 10.1016/J.IJEPES.2018.11.002.

[14] N. Qu, J. Wang, and J. Liu, "An arc fault detection method based on current amplitude spectrum and sparse representation," *IEEE Trans. Instrum. Meas.*, vol. 68, no. 10, pp. 3785–3792, Oct. 2019, doi: 10.1109/TIM.2018.2880939.

[15] A. Winursito, R. Hidayat, and A. Bejo, "Improvement of MFCC feature extraction accuracy using PCA in Indonesian speech recognition," in *Proc. Int. Conf. Inf. Commun. Technol. (ICOIACT)*, Mar. 2018, pp. 379–383, doi: 10.1109/ICOIACT.2018.8350748.

[16] R. Hidayat, A. Bejo, S. Sumaryono, and A. Winursito, "Denoising speech for MFCC feature extraction using wavelet transformation in speech recognition system," in *Proc. 10th Int. Conf. Inf. Technol. Electr. Eng. (ICITEE)*, 2018, pp. 280–284.

[17] P. W. Zaki, A. M. Hashem, E. A. Fahim, M. A. Masnour, S. M. ElGenk, M. Mashaly, and S. M. Ismail, "A novel sigmoid function approximation suitable for neural networks on FPGA," in *Proc. 15th Int. Comput. Eng. Conf. (ICENCO)*, 2019, pp. 95–99.

[18] Z. Chunyan, H. Junjia, H. Weichao, Z. Wei, W. Lizhong, and L. Chengyan, "Investigation on change laws of arc burning time of the arc plasma," in *Proc. IEEE Conf. Electr. Insul. Dielectr. Phenomena*, Aug. 2009, pp. 580–583.

[19] A. Chaudhari, A. Rahulkar, and S. B. Dhonde, "Combining dynamic features with MFCC for text-independent speaker identification," in *Proc. Int. Conf. Inf. Process. (ICIP)*, 2015, pp. 160–164.

[20] 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, doi: 10.1109/TIM.2018.2826878.

[21] Y. Wang, F. Zhang, X. Zhang, and S. Zhang, "Series AC arc fault detection method based on hybrid time and frequency analysis and fully connected neural network," *IEEE Trans. Ind. Informat.*, vol. 15, no. 12, pp. 6210–6219, Dec. 2019.

[22] 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.

[23] N. Xu, Y. Yang, Y. Jin, and J. He, "Identification of series fault arc of low-voltage power cables in substation based on wavelet transform," in *Proc. IEEE 5th Int. Conf. Integr. Circuits Microsyst. (ICICM)*, Oct. 2020, pp. 188–192, doi: 10.1109/ICICM50929.2020.9292218.

[24] Q. Zibo, G. Wei, G. Chen, and X. Kaili, "The development of AC arc fault simulation test device with arc breaking function," in *Proc. 7 IEEE Int. Conf. Comput. Sci. Eng. (CSE) IEEE Int. Conf. Embedded Ubiquitous Comput. (EUC)*, Jul. 2017, pp. 224–227, doi: 10.1109/CSE-EUC.2017.47.

YAO WANG (Member, IEEE) was born in Hebei, China, in 1981. He received the B.S., M.S., and Ph.D. degrees in electrical engineering from the Hebei University of Technology, Tianjin, China, in 2006, 2009, and 2012, respectively. He is currently an Associate Professor with the State Key Laboratory of Reliability and Intelligence of Electrical Equipment, School of Electrical Engineering, Hebei University of Technology. He has authored or coauthored more than 20 technical articles. His current research interests include emerging technology in AI-based fault diagnosis and intellectualizing of electrical apparatus.

DEJIE SHENG was born in Weifang, Shandong, China, in 1998. He received the B.S. degree from the Jiangsu University of Science and Technology, Zhenjiang, China, in 2020. He is currently pursuing the M.S. degree in electrical engineering with the Hebei University of Technology, Tianjin, China. His current research interests include electrical contact theory and arc.

HAIQING HU was born in Wenzhou, Zhejiang, China, in 1972. He received the B.S. degree from the Hunan Institute of Technology, Hengyang, China.

He is currently the Deputy Director of Zhejiang High and Low Voltage Electrical Product Quality Inspection Center. His current research interests include low-voltage electrical apparatus testing equipment and methods, and the research of low-voltage electrical apparatus structure and standardization.

KEFAN HAN was born in Weifang, Shandong, China, in 1998. He received the B.S. degree from the Qingdao University of Technology, Qingdao, China, in 2021, where he is currently pursuing the M. S. degree. His current research interests include load identification and low voltage circuit breaker.

JIAWANG ZHOU was born in Hebei, China, in 1998. He received the B.S degree in electrical engineering and automation from the Hebei University of Technology, Tianjin, China, in 2021, where he is currently pursuing the M.S. degree in electrical engineering. His current research interest includes arc fault diagnosis based on artificial intelligence.

LINMING HOU was born in Shandong, China, in 1995. He received the B.S. degree in electrical engineering and automation from the Qingdao University of Science and Technology, Qingdao, China, in 2018, the M.S. degree in electrical engineering from the Hebei University of Technology of Technology, Tianjin, China, in 2021. He is currently works at Zhejiang Testing and Inspection Institute for Mechanical and Electrical Products Quality Company Ltd., Hangzhou, China. His current research interests include low voltage appliance detection methods and A.C. and D.C. arc fault detection technologies.