

RESEARCH ARTICLE

Explainability Approach-Based Series Arc Fault Detection Method for Photovoltaic Systems

YAO WANG^{1,2}, (Member, IEEE), JIAWANG ZHOU^{1,2},
KAMAL CHANDRA PAUL^{1,3}, (Graduate Student Member, IEEE),
TIEFU ZHAO^{1,3}, (Senior Member, IEEE), AND
DEJIE SHENG^{1,2}, (Student Member, IEEE)

¹School of Electrical Engineering, Hebei University of Technology, Tianjin 300400, China

²State Key Laboratory of Reliability and Intelligence of Electrical Equipment, Hebei University of Technology, Tianjin 300132, China

³Department of Electrical and Computer Engineering, The University of North Carolina at Charlotte, Charlotte, NC 28223, USA

Corresponding author: Dejie Sheng (202311401004@stu.hebut.edu.cn)

This work was supported in part by the Science and Technology Research Project for Colleges and Universities in Hebei Province under Grant CXY2023006, in part by the Central Funds Guiding the Local Science and Technology Development under Grant 226Z2102G, and in part by Zhejiang Provincial Natural Science Foundation of China under Grant LTGG23E070001.

ABSTRACT Arc fault detection devices are mandatory worldwide for mitigating DC series arc faults in photovoltaic systems. However, they are prone to nuisance tripping. Artificial intelligence-based approaches can be a solution, but they are “black boxes” and challenging to modify. This paper proposes an explainability and attention-based method to investigate the intensive details of such an algorithm. The contributions of an arc feature to the proposed model can be visualized by the proposed interpretable methodology so that insensitive arc features can be removed to reduce the quantity of input data. Additionally, the structure of the proposed model can be optimized by cutting the redundant layers. Thus, an accuracy of 99.63% is achieved with only 48.48% of the parameters compared to the original model. Finally, the optimized model is implemented by a Cortex M7-based microprocessor with a runtime of only 7.8 ms, making it ready for industrial application.

INDEX TERMS Arc discharge, artificial intelligence, deep learning, discrete Fourier transforms, electrical fault detection, electrical safety, fault diagnosis, machine learning, photovoltaic systems, proactive detection.

I. INTRODUCTION

As the hazards such as rising sea levels caused by global warming become increasingly significant, carbon peaking and carbon neutrality have become global development goals. People need to wean themselves from fossil fuels to reduce carbon dioxide emissions. The energy of photovoltaic (PV) power generation is derived from solar energy, which will not cause environmental pollution during power generation. Additionally, the power generation is considerable, making it essential to distributed power systems [1], [2]. Because PV power generation is easy to install, it is widely installed in places such as home roofs. However, an arc fault in a

PV system generates a high temperature of 20,000 K, which causes electrical fires, making arc faults a major threat to PV system safety [3], [4].

PV system arc faults are divided into series, parallel, and ground-arc faults. Parallel arc faults and ground-arc faults are easily cut off by overload protection circuit breakers due to their large currents [5], [6]. However, due to the load limitation of series arc faults, the current generated is generally smaller than the normal current; thus, protection by traditional protection devices such as traditional overload protection circuit breakers is challenging. Therefore, series arc faults are the focus of arc fault detection.

The threshold detection method, which uses current characteristics, is the most widely used arc fault detection method. Ahmadi et al. [7] decomposed the information

The associate editor coordinating the review of this manuscript and approving it for publication was Gerard-Andre Capolino.

matrix composed of voltage and current to obtain arc characteristics. Chae et al. [8] simultaneously detected load current characteristics in time and frequency domains based on relative amplitude comparisons. Ahn et al. [9] used a discrete wavelet transform to obtain the current frequency domain signal and then used the zero range density as the arc fault detection criterion to achieve effective arc faults detection. Wang et al. [10] combined the current waveform similarity and the voltage signal sag criterion and used the Hausdorff distance algorithm to calculate the degree of load-side voltage waveform distortion to identify arc faults.

However, under different system operating conditions, the variation range of the time-frequency domain characteristic quantity of the arc fault is different. Additionally, it is difficult to set the fault identification threshold accurately based on engineering experience. In particular, the inverter in a PV system is a power electronic device prone to inevitable conduction interference in the bus current [11], [12]; its frequency band overlaps with the arc current frequency band, which seriously affects arc fault identification accuracy.

In recent years, many scholars have applied artificial intelligence methods to solve the arc fault identification problem [13], [14]. Wang et al. [15] used the raw current signal of an AC system as the input of a convolutional neural network to detect arc faults. The proposed detection model has a computing time of 30 ms for embedded devices. Xing et al. [16] constructed a network structure of CNN+LSTM to detect arc faults and predict arc fault occurrence. Cai and Wai et al. [17] used optimized variational modal decomposition to extract arc features and utilized support vector machine(SVM) as a classifier to identify arc features. Similarly, Jiang et al. [18] also employ frequency domain analysis methods to decompose current signals, using the resulting current frequency domain features as input for the SVM algorithm to achieve arc fault detection. Yan et al. [19] built a sequential convolutional network to extract DC arc current signal characteristics. Tang et al. [4] and Wang et al. [20] use 1D convolution to extract arc features of the original current signal and a fully connected neural network to classify circuit states. Likewise, Jiang et al. [21] also use a 1D convolution to extract features of high-frequency components in current signals to identify arc faults. Zhang et al. [22] integrates load classification with arc fault detection by first utilizing event classification to determine the type of load; it then selects the appropriate arc fault detection algorithm based on the load type; ultimately, the corresponding K-Nearest Neighbor algorithm applicable to the identified load type is employed for arc detection. Shen and Xu [23] analyze signals through empirical wavelet transform, extracting frequency domain energy metrics of different frequency bands, and ultimately input the extracted frequency domain features into a fully connected neural network to achieve arc fault detection.

The artificial intelligence method does not need to set the recognition threshold manually to identify the fault arc. However, the current method based on machine learning needs to extract arc features before training the network model [7]. Previous arc detection research has developed from the time domain characteristics of arcs to high-order cumulants [24] and singular values [25], which reflects the increasing emphasis on arc feature selection. Arc feature extraction needs to relies on continuous trial and error or adopting an optimization method independent of the arc recognition model [16], [26]. Because the deep neural network model is a “black box”, its internal behavior characteristics are unknown, which may lead to a mismatch between the arc features extracted by this method and those extracted by the machine learning model. In addition, the current machine learning algorithm has many network model parameters and computations, making it impossible to apply to arc fault circuit breakers using embedded microprocessors [27], [28]. Since the time from fault arc occurrence to fire is less than 2.5 seconds, the network model must be directly deployed to the user and be lightweight enough to meet the requirements of embedded microprocessors. Therefore, it is necessary to determine the critical characteristics of fault arcs according to the network model and optimize the network structure and parameters to meet the application requirements of embedded devices.

The attention mechanism can help the model select effective and appropriately scaled features so that the model can complete the task effectively. Embedding attention mechanisms into other models is widely used in deep learning [29], [30]. In addition to using the attention mechanism to help models select valuable features, the interpretability property of attention weights can also be used to visualize data relationships. Song et al. [31] combined an attention mechanism and long short-term memory (LSTM) to analyze the dependencies of time series data in the clinical domain. Lee et al. [32] proposed a model based on the attention mechanism, which can select words with semantic importance and emotional classification meaning from documents without word- or phrase-level emotional polarity information. Tran et al. [33] combined the idea of bilinear projection with an attention mechanism. The proposed architecture can detect and focus on crucial temporal information and highlight the contribution and importance of each temporal instance. We use the interpretability of the attention mechanism to visualize the importance of arc features and use this information as a guide to select key feature frequency bands and optimize the network structure. This process improves the reliability of the network model and reduces the number of parameters and computations. The main contributions of this paper are as follows:

- 1) To detect arc faults in PV systems, an experimental platform was built to analyze arc faults at different current levels and locations. The study also designed an experiment to evaluate the robustness of the arc fault detection device

(AFDD) in the UL 1699B standard against light mutation interference caused by PV inverter start-up and shadow occlusion. Simultaneously, to meet the development requirements of higher power capacity of PV systems, in addition to the 3 A, 8 A, 8.5 A, and 16 A current level experiments specified in UL1699B, two larger current level arc tests at 20 A and 25 A were added;

2) Due to the complexity and interference of PV system working conditions, extracting critical fault arc features for accurate detection is crucial. This study analyses the power spectrum characteristics of the DC series arc current signal in a PV system. It extracts critical characteristic frequency bands in a targeted manner using the interpretability of the attention mechanism. The results show that this approach effectively distinguishes arc fault states from normal working states and reduces the amount of misoperation caused by interference, improving fault detection accuracy;

3) To meet the requirements of edge computing with embedded microprocessors, the network model needs to be lightweight and accurate enough to be deployed directly to the user end. This study optimized the network model based on the critical characteristics of the arc fault to reduce the number of model parameters and calculations while ensuring accuracy. The experimental results show that the proposed lightweight algorithm can be implemented by a Cortex M7-based microprocessor with a runtime of only 7.8 ms, outperforming other algorithms.

The paper is structured as follows: Section II introduces arc data analysis, processing methods, and the creation of an arc database. Section III outlines the proposed arc fault detection model based on deep learning interpretability. In Section IV, we discuss the arc feature selection process and optimization of the arc identification model based on the interpretable principle of the attention mechanism. We also present the results from immunity tests and actual tests of the optimized arc identification model. Finally, Section V contains the conclusions of the paper. The overall flow of the article is shown in Fig. 1.

II. DATA COLLECTION AND ANALYSIS

This section describes the acquisition and processing of the PV arc data used and details the number of samples in the dataset.

A. DATA COLLECTION

A test bench is constructed for UL1699B for data collection and arc feature investigation, as shown in Fig. 2. The test bench comprises a PV simulator, an arc generator, a PV inverter, a decoupling network, and a module/conductor impedance network. The PV simulator of Itech IT6018C is used, which is a programmable power source capable of simulating different outputs under different conditions, such as a solar array, such as load changes, switch operations, and shadow influences. The arc generator can trigger an

TABLE 1. Technical specifications of the experimental test bench for series DC arc faults in PV systems.

Component	Value	Component	Value
C1	20 μ F	L2, L3	60 μ H
C2, C3	22 nF	L4, L5	50 μ H
C4	10 μ F, 150 nF	R3	1 Ω
		R4	1 Ω
C5, C6	1 nF	R5, R6	1 Ω
L1	12 mH	L6, L7	70 μ H

arc by separating the moving electrode from the stationary electrode at a speed of approximately 5 mm/s. Moreover, the module and conductor impedances are considered since the distributed capacitance of the DC bus can attenuate the high-frequency components of the arc current. The decoupling network is used for isolating noise from the PV simulator. The parameters for the impedance network are shown in Table 1. These parameters should be tuned according to the length of the DC bus, whose typical value is 80 m.

The arc test was carried out at the beginning and end of the PV module (the string head and the string tail). The DC arc current signal in the solid conductor may be attenuated by impedance. To reduce the impact of this situation on DC arc fault detection, a test is performed by adding conductor impedance before and after the arc current signal acquisition location in addition to the UL1699B standard required arc current signal collection location. The four kinds of fault arc current signal collection positions in the experiments are shown in A1, A2, A3, and A4 in Fig. 2. Simultaneously, to meet the development requirements of higher power capacity PV systems, in addition to the 3 A, 8 A, 8.5 A, and 16 A current level experiments specified in UL1699B, two larger current level arc tests were performed at 20 A and 25 A.

According to the Nyquist theorem, the sampled signal can completely retain the original signal's information when the sampling rate is greater than twice the frequency of the target acquisition signal. The PV DC series arc fault signal is generally distributed in the 3 kHz-120 kHz frequency band [27]. To collect as much of a complete DC arc fault signal as possible and to reflect the high-frequency characteristics more clearly, the sampling rate was set to 250 kHz in the test.

B. CHARACTERISTIC ANALYSIS OF THE PV ARC

Because of the unpredictable location of the arc and the difficulty in accurately measuring the arc voltage, the current signal of the arc fault is usually used to identify the arc fault. Since PV power generation is influenced by light intensity and temperature, the current is more random than that in an AC system, making it difficult to distinguish the arc current signal from the normal current signal using only time domain characteristics. Therefore, the frequency domain feature of the DC arc fault current is analyzed. The

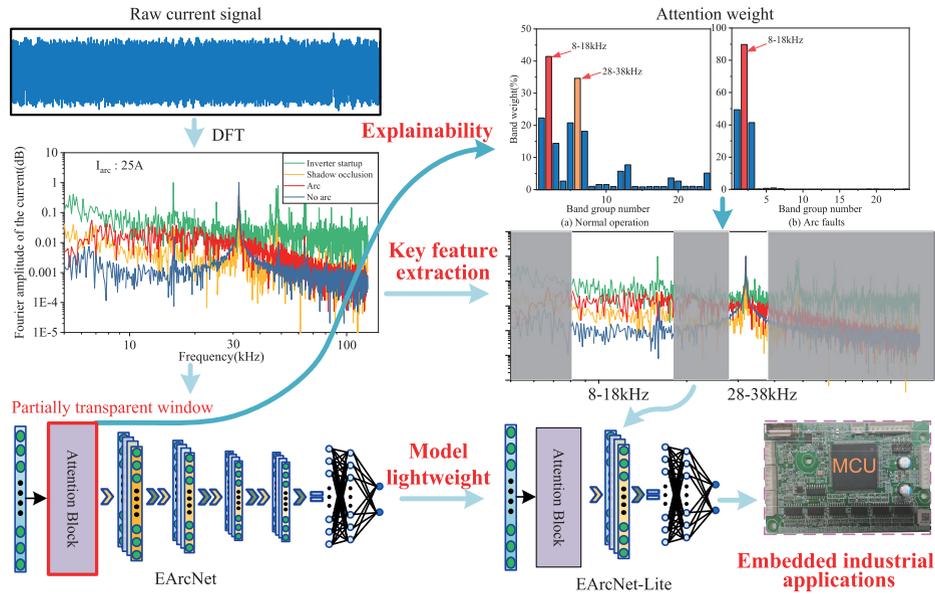


FIGURE 1. Block diagram of the proposed method for series arc fault detection in PV systems based on an explainability approach.

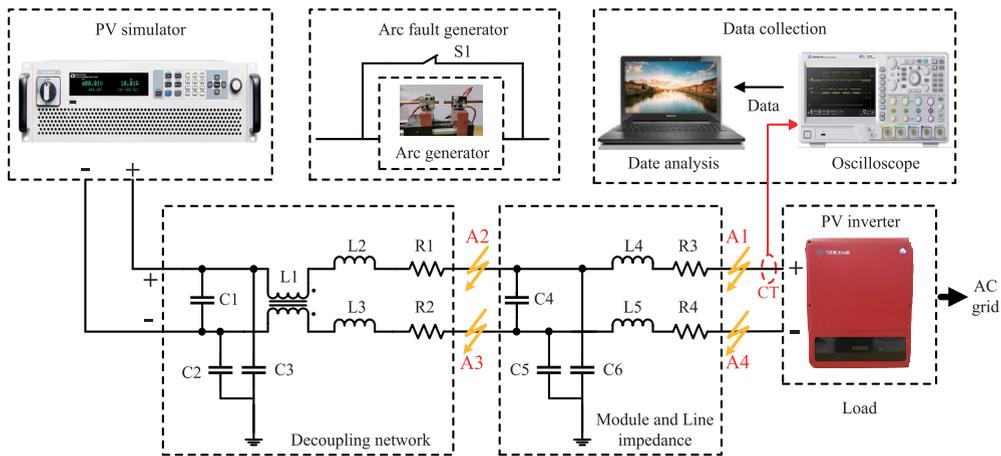


FIGURE 2. Schematic diagram of the test circuit layout, potential arc position, and data collection of the experimental test bench.

discrete fourier transform (DFT) is a widely used signal analysis method that can discretize the DC arc fault current signal in the frequency domain [34], [35]. By dividing the DC arc fault signal into several small signals of equal length through the time window, the DFT can effectively distinguish arc and nonarc features. Although other methods, such as short-time Fourier transform and wavelet transform, can distinguish arc and normal state features, DFT is preferred due to its simple calculation, low application threshold, and ability to arbitrarily select the number of transformation points [7]. Therefore, this paper chooses DFT to extract the characteristics of the arc fault current signal.

Using the DFT to analyze the signal $x(n)$, $X(k)$ is still a finite-length sequence of length N , and the transformation

process can be expressed as

$$\begin{aligned}
 X(k) &= \text{DFT}[x(n)] = \sum_{n=0}^{N-1} x(n)e^{-j\frac{2\pi}{N}kn} \\
 &= \sum_{n=0}^{N-1} x(n)W_N^{kn}, k \in [0, N-1]. \quad (1)
 \end{aligned}$$

To accurately analyze the DC arc fault current signal, determining an appropriate time window is necessary before performing characteristic analysis. The smaller the time window is, the better the temporal resolution of the spectrum results, but this may overload the computing power of the hardware processor. A time window that is too large will put considerable pressure on the processor and affect its

TABLE 2. Summary of experimental dataset categories and features used for the design and validation of DC arc fault detection algorithm in PV systems.

Dataset	Label	Sample number	Total number	Test condition
Training set	Arc	6700	13500	Arc fault detection test includes 6 current levels: 3A, 8A, 8.5A, 16A, 20A, 25A.
	No Arc	6800		
Validation set	Arc	900	1800	Immunity test includes 2 disturbances: Inverter start and Shadow occlusion.
	No Arc	1000		
Test set	Arc	1300	2700	
	No Arc	1400		

resolution in the time domain, which may reduce the real-time performance of the detection scheme. To balance the real-time requirements of the arc detection task and the resolution of the arc characteristics in the frequency domain, this paper selects a time window of 10 ms based on the AR model proposed in [27]. This time window size is feasible for the detection scheme and can provide effective features to for distinguishing arc faults from normal states.

C. DATA PREPROCESSING AND CREATION OF AN ARC DATABASE

For embedded practical applications, current signals are collected by current transformers, and both normal state data and arc data are collected under all experimental conditions. A DFT is performed on the divided small signal segments. The arc characteristics are mainly after 3 kHz, and there is considerable low-frequency harmonic interference in the frequency band below 3 kHz, which easily affects arc fault identification. Therefore, the normalized spectrum data of the 3-125 kHz frequency band are used as the input of EArcNet.

Finally, the set-aside method is used to randomly divide the dataset into 75% of the training set, 10% of the validation set, and 15% of the test set. The number of divided datasets is shown in Table 2.

III. EARCNET ANALYSIS

A convolutional neural network has excellent one-dimensional sequence data classification capabilities and can precisely extract the characteristics of the arc current and normal operating current [15]. A series of deep learning algorithms such as convolutional neural networks have high arc fault identification accuracy. However, their powerful learning ability comes at the cost of a large number of network parameters and computations, making the model size and computations far more than the memory capacity and computing capabilities of the embedded microprocessor. Therefore, it is necessary to apply the interpretability technology of deep learning to arc fault identification, extract key arc features, optimize the network structure, and reduce the number of network parameters and calculations so that the arc fault identification algorithm based on deep learning can be applied to the AFDD with an embedded microprocessor as the main application scenario.

TABLE 3. Comparison of the advantages and disadvantages of scaled dot-product attention, additive attention, and convolutional attention.

	Advantage	Disadvantage
Scaled dot-product attention	Simple computation and high efficiency	For longer sequences, efficiency may decrease
Additive attention	Adaptable to inputs of varying lengths	Computation is relatively complex, requiring the configuration of additional layers and parameters
Convolutional attention	Capable of capturing both local and global dependencies	High computational complexity

The proposed EArcNet modifies the structural framework of [15] and incorporates an attention mechanism. The attention mechanism can enable the neural network to focus on important information, thereby improving the network's ability to identify arc faults. The interpretability of the attention mechanism enables one to grasp the decision-making behavior of the network model [37], [38]: the key feature frequency bands of PV arc faults are extracted by visualizing feature weights, which improves the reliability of the network.

In the current research, the attention mechanisms predominantly employed encompass scaled dot-product attention, additive attention, and convolutional attention. Table 3 provides a comparative analysis of the performance metrics associated with these three attention mechanisms. Table 3 shows that while additive attention exhibits enhanced handling capabilities for inputs of varying lengths, and convolutional attention can capture both local and global dependencies, the scaled dot-product attention mechanism boasts superior computational efficiency. Given that the task of arc fault detection places a significant emphasis on the algorithm's real-time performance, the scaled dot-product attention mechanism has been selected due to its higher computational efficiency. The schematic diagram of the calculation process of the scaled dot-product attention is illustrated in Fig. 3. In Fig. 3, Q (query), K (key), and V (value) are derived from the following equations:

$$\begin{cases} Q = W_Q X_i \\ K = W_K X_i \\ V = W_V X_i \end{cases} \quad (2)$$

where X_i represents the input data, while W_Q , W_K , and W_V are the weight matrices corresponding to the query, key, and value. W_Q , W_K , and W_V are each a fully connected layer containing 16 neurons.

The scaled dot-product attention computes attention scores by calculating the dot product between the query and the key. Subsequently, these scores are scaled by dividing by the square root of the dimensionality of the key. The softmax function is then applied to obtain the attention weights. The output of the attention mechanism is calculated by performing a dot product between the attention weights and the value. The mathematical expression for scaled

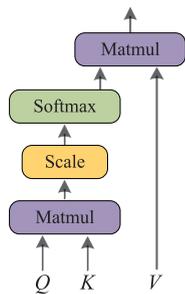


FIGURE 3. Schematic diagram of the calculation process of the scaled dot-product Attention.

dot-product attention is:

$$\text{Attention}(Q, K, V) = \text{Softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (3)$$

where d_k denotes the dimensionality of the key.

In scaled dot-product attention, sinusoidal position encoding is employed. This encoding method does not require learning through training. By applying sine and cosine functions of different frequencies to each position, a unique encoding can be generated for each position. The calculation formula for Sinusoidal Position Encoding is:

$$\begin{cases} PE(pos, 2i) = \sin\left(\frac{pos}{10000^{2i/d_{model}}}\right) \\ PE(pos, 2i + 1) = \cos\left(\frac{pos}{10000^{2i/d_{model}}}\right) \end{cases} \quad (4)$$

where $PE(pos, 2i)$ and $PE(pos, 2i + 1)$ respectively denote the encodings at the positions of the $2i$ and $2i + 1$ columns in the pos row of the encoding matrix, while d_{model} represents the dimensionality of embeddings within the model.

EArcNet uses 1D convolution to extract arc features, and the feature extraction layer composed of multilayer convolution optimizes the network parameters by repeated back-propagation and forward-propagation. If the length of the 1-D convolution vector f is l and the length of the convolution kernel k is r , the result $(f \times k)$ of the j th convolution kernel in the i th convolution can be expressed as

$$(f \times k)(i) = \sum_{j=1}^r k(j)f\left(i-j + \frac{r}{2}\right). \quad (5)$$

The rectified linear unit(ReLU) activation function ReLU can be expressed as

$$\text{ReLU}(x) = \text{Max}(0, x). \quad (6)$$

The cross-entropy loss function can be expressed as

$$\text{Categorical Cross Entropy} = -\frac{1}{N} \sum_j y_j \log(\hat{y}_j) \quad (7)$$

where N is the total number of training data, y is the true label of the training sample, and \hat{y} is the predicted label.

With each 10 ms window, 2500 current signal points can be collected at a sampling rate of 250 kHz. After

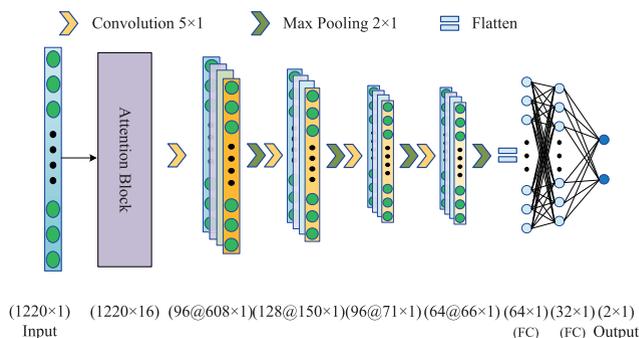


FIGURE 4. Network architecture of the proposed EArcNet model (FC stands for Fully Connected layer of the network).

performing the DFT transformation, 1250 energy spectrum amplitudes are obtained. These energy spectrum amplitudes are evenly distributed at 3-125 kHz, and the frequency coordinates between every two energy spectrum amplitude coordinates have a difference of 100 Hz. The normalized energy spectrum amplitude is taken as the input data of EArcNet; thus, the input size of EArcNet is 1220×1 , and the network structure of EArcNet is shown in Fig. 4. EArcNet enters the input data into the attention module and then inputs the features processed by the attention module into a 1-dimensional convolution, maximum pooling, and fully connected layer. The network has an attention module, 4 1D convolutional layers, and 4 max-pooling layers followed by three fully connected layers. The data output by the attention mechanism is two-dimensional; it requires a dimensionality increase operation before input into the convolutional layers. Each convolutional layer is followed by a ReLU operation and a max pooling layer. Each layer of the first and third convolutional layers has 96 filters, the second convolutional layer has 128 filters, and the fourth convolutional layer has 64 filters. All the filters in the convolutional layers have the same kernel size of 5×1 . The maximum pooling layer size is 2×1 , which can effectively reduce the feature map size and improve computational efficiency. To accommodate the length of the input data, the stride of the first and second convolutional layers in EArcNet is set to 2, while the stride of the third and fourth convolutional layers is set to 1. Similarly, the stride of the first and second pooling layers is set to 2, and the stride of the third and fourth pooling layers is set to 1. After the last max pooling layer, the flatten operation is performed to reduce the dimensionality of the data, and three fully connected layers with 64, 32, and 2 neurons are followed by the flatten operation. The output layer is the classification layer. The softmax function is used to convert the scores into the probability that the sum is 1, and the class with the highest probability is taken as the final classification state.

The EArcNet model is implemented in Keras using a TensorFlow backend. During training, EArcNet employs an adaptive learning rate technique with a batch size of 100. The adaptive learning rate monitors “validation_loss” with

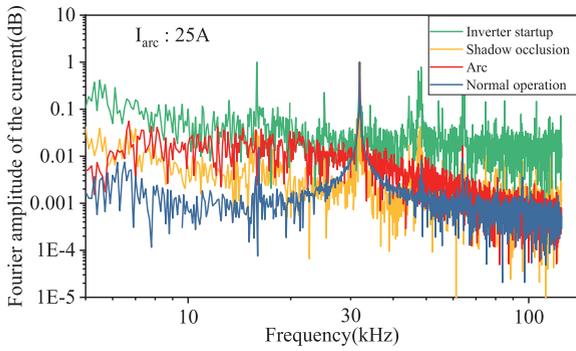


FIGURE 5. Spectrum diagram of the measured current under the conditions of normal operation, arc fault, inverter startup, and shadow occlusion in PV systems.

a tolerance of 10. The initial learning rate is set to 0.001, and the minimum learning rate is set to 0.00001.

IV. MODEL OPTIMIZATION BASED ON DEEP LEARNING INTERPRETABILITY

This section analyses the experimental results of EArCNet. Using deep learning explainability to extract key features of the PV system arc energy spectrum and optimize the EArCNet network structure. The optimized EArCNet is cross-validated and experimentally verified, and an arc detection algorithm based on an embedded microprocessor is proposed and verified.

A. EXPERIMENTAL RESULTS AND ANALYSIS

When training EArCNet, we used an adaptive learning rate strategy and trained for 120 epochs with a batch size of 100 and an initial learning rate of 0.00001. Finally, the accuracy of EArCNet in identifying arc faults is 99.78%, which verifies that EArCNet is competent for arc fault identification tasks.

B. KEY FREQUENCY BAND SELECTION BASED ON DEEP LEARNING INTERPRETABILITY

Selecting the characteristic frequency band for a PV arc is crucial for arc fault identification. Fig. 5 shows that the PV system arc fault feature is mainly distributed in the 3-125 kHz frequency range. However, in the frequency range after 20 kHz, the energy spectra of the arc state, normal operation state, inverter startup, and shadow occlusion overlap in some frequency bands. Energy spectra of overlapping frequency bands will not be helpful for arc fault identification and may even reduce the model’s accuracy. Therefore, key arc features beneficial for arc fault identification must be extracted from the 3-125 kHz arc characteristics.

The attention mechanism embedded in EArCNet allows us to directly inspect the inner workings of deep learning architectures [38], [39], making EArCNet interpretable. Attention mechanisms improve the model’s local interpretability to some extent. They act like transparent windows, allowing us to see the importance of arc characteristics,

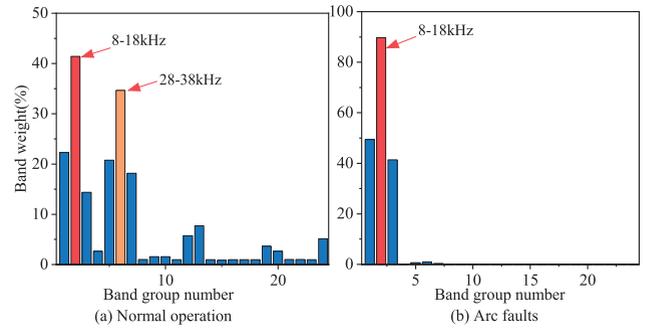


FIGURE 6. Attention weights of the proposed EArCNet model for each current spectrum subband: (a) normal operation and (b) arc faults.

which is crucial in arc fault identification, especially in arc feature selection. According to the attention weight of the arc characteristic frequency band, EArCNet can check its recognition of arc characteristics, visualize the importance of arc characteristics, and extract key arc fault characteristics.

Fig. 6(a) shows the attention weights of EArCNet when identifying nonarcs and Fig. 6(b) shows the arc data. To determine the importance of feature bands, we extract the attention weights of EArCNet when identifying the arcs and normal states under all working conditions and calculate their average values. We resampled the 3-125 kHz band with a step size of 5 kHz and a width of 10 kHz, resulting in 24 subbands of 10 kHz each, and calculated the sum of the attention weights within each subband. We found that EArCNet mainly focused on the energy spectrum features of the 8-18 kHz sub-band when identifying arc data, while it also paid attention to the energy spectrum features of the 28-38 kHz subband when identifying nonarc data. Moreover, from Fig. 6 we can observe that the critical features of arc data are mainly concentrated below 23 kHz, while those of nonarc data are distributed in a broader range. To reduce the input dimension and computations of the neural network and meet the real-time and memory requirements of the embedded microprocessor, we extracted the joint energy spectrum feature band consisting of the highest weight sub-band 8-18 kHz for arc data and important weight sub-band 28-38 kHz for non-arc data and retrained EArCNet.

C. NETWORK OPTIMIZATION BASED ON DEEP LEARNING INTERPRETABILITY

Although EArCNet achieved satisfactory arc fault detection accuracy, it had four convolutional layers and three fully connected layers, which exceeded the memory capacity and computational power of the microcontroller. To reduce the input feature dimension, we use the joint energy spectrum feature band as input to EArCNet and we simplify the network structure by removing convolutional, pooling, and fully connected layers. We refer to the optimized version of EArCNet as EArCNet-Lite, its structure is illustrated in Fig. 7. After the attention mechanism, EArCNet-Lite had only one convolutional layer with 64 filters and one

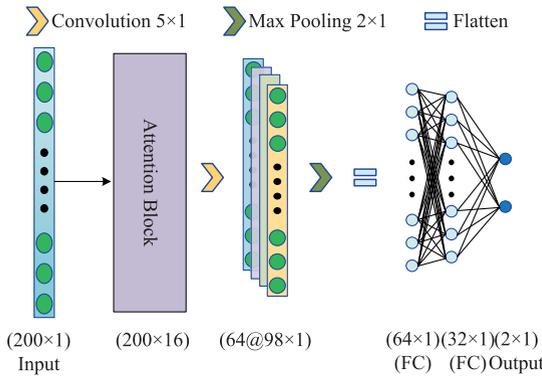


FIGURE 7. Network architecture of the proposed EArcNet-Lite model (FC stands for Fully Connected layer of the network).

TABLE 4. Confusion matrix, precision, recall, and overall accuracy of the proposed EArcNet-Lite model.

		Predicted Class		
		Arc	Normal	Total
Actual Class	Arc	1297	3	1300
	Normal	7	1393	1400
	Total	1304	1396	2700
Precision =		99.46%		
Recall =		99.77%		
Accuracy =		99.63%		

TABLE 5. Comparison of total parameters and arc fault detection accuracies between EArcNet-light and EArcNet models.

	Total parameters	Accuracy
EArcNet	430,930	99.78%
EArcNet-Lite	208,914	99.63%

max pooling layer of size 2×1 . After the max pooling layer, we perform a flattening operation to reduce the data dimension and add three fully connected layers with 64, 32, and 2 neurons. A softmax function was used as the output layer following the last FC layer. We trained the model using the extracted joint energy spectrum feature band and fine-tuned the hyperparameters to maximize its arc fault detection accuracy.

Table 4 shows the confusion matrix for EArcNet-Lite training. The precision, recall, and precision were assessed with respect to the Table 4 confusion matrix to evaluate model performance. The EArcNet-Lite model has an accuracy of 99.63%, a recall of 99.77%, and precision of 99.46%. Both the accuracy and recall rates reach more than 99%, which verifies that EArcNet-Lite is competent for the task of arc fault identification.

Table 5 compares the characteristics and precision of EArcNet with those of EArcNet-Lite. The experimental results show that when the critical energy spectral features of arcs extracted based on deep learning interpretability are used as arc detection features, EArcNet-Lite can still maintain 99.63% accuracy even if the optimized network model parameters are only 48.48% of those of EArcNet.

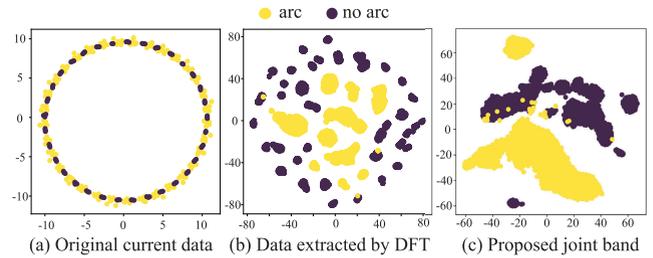


FIGURE 8. t-SNE visualization results for each dataset: (a) original current data, (b) data extracted by the DFT-based method, and (c) data extracted by the proposed joint subband-based method.

Since the network needs to learn fewer features, the extracted joint frequency bands are more representative of PV arc faults. The experimental findings indicate that the characteristic frequency bands of the arc critical energy spectrum that were extracted based on interpretability can accurately represent the arc characteristics and reduce the number of network parameters and computations while maintaining network identification accuracy.

D. T-SNE DATA DISTRIBUTION VISUALIZATION

To further demonstrate the ability of our attention mechanism to extract key arc fault features, we used t-SNE [40], a dimensionality reduction technique in manifold learning, to visualize and compare the original current data, the energy spectrum data extracted by DFT, and the key arc fault features extracted by the attention mechanism. The experimental results are shown in Fig. 8. Fig. 8(a) shows the distribution of the original current data, for which distinguishing between the arc fault and normal state is difficult due to randomness and redundancy. Fig. 8(b) shows the energy spectrum data extracted by the DFT, which have slight overlaps but are still not well clustered. Fig. 8(c) shows the arc fault key features extracted by the attention mechanism, which have large clusters for both arc fault and normal state data, providing good knowledge extraction for the subsequent network model. The t-SNE visualization results validate the significance of our attention-based arc fault key feature extraction approach.

E. ABLATION EXPERIMENT

An ablation experiment is designed to explore the influence of the key arc characteristics and network model optimization proposed in this paper on algorithm performance. We trained and tested the EArcNet and EArcNet-Lite models with two different input datasets to compare their arc fault detection performances. The introduction is as follows:

EArcNet (3-125 kHz): Using the EArcNet model, the input data are the current energy spectrum amplitude of 3-125 kHz, so the input dimension is 1220×1 .

EArcNet (joint frequency band): Using the EArcNet model, the input data is the optimized key arc features, so the input dimensions are 200×1 .

TABLE 6. Comparison of total parameters, floating points operations per second(FLOPs), and arc fault detection accuracy of the models from ablation experiments.

	Accuracy	Total parameters
EArcNet(3-125 kHz)	99.78%	430,930
EArcNet(Joint frequency band)	99.90%	303,954
EArcNet-Lite(3-125 kHz)	97.38%	1,253,394
EArcNet-Lite(Joint frequency band)	99.63%	208,914

EArcNet-Lite (3-125 kHz): Using the EArcNet-Lite model, the input data are the current energy spectrum amplitude of 3-125 kHz, so the input dimension is 1220×1 .

EArcNet-Lite (joint frequency band): Using the EArcNet-Lite model, the input data are the optimized key arc features, so the input dimension is 200×1 .

EArcNet (3-125 kHz) and EArcNet (joint frequency band) have the same number of network layers and convolution kernel sizes except for different input dimensions. We fine-tuned the hyperparameters for each scheme to achieve optimal accuracy. Table 6 shows the comparison results of the four schemes in the ablation experiment.

Compared with EArcNet-Lite, EArcNet has deeper network layers and can learn more knowledge. Compared with 3-125 kHz, the key characteristic frequency band of the arc extracted through interpretability analysis avoids the influence of interference in the system on arc fault identification. Therefore, EArcNet (joint frequency band) has the highest accuracy rate. However, the deeper network structure of EArcNet (joint frequency band) also makes it more computationally intensive, as it outperforms the RAM of the i.MX RT1064 used in this article, and ultimately cannot run in the embedded microprocessor. Therefore, it is necessary to optimize EArcNet. EArcNet-Lite (3-125 kHz) only optimizes EArcNet but does not use interpretable analysis to extract key arc characteristic frequency bands. A reduction in the network depth reduces the fitting ability of the network, and 3-125 kHz frequencies include interference frequency bands that are likely to interfere with arc fault detection, so the network depth has the lowest recognition accuracy. Although EArcNet-Lite sacrifices part of its fitting ability, EArcNet-Lite (joint frequency band) extracts key characteristic frequency bands that can characterize arc faults based on interpretability analysis. Therefore, an accuracy rate of 99.63% can be maintained, and the calculation time of the i.MX RT1064 embedded microprocessor is only 7.8 ms, which meets the requirements of outstanding real-time performance and high arc fault detection accuracy.

In summary, introducing interpretability provides evident benefits, which manifest in two main aspects:

1) Resolution of key feature extraction in arc fault identification: We effectively extract critical arc fault frequency bands through interpretability analysis, avoiding interference bands and significantly enhancing arc fault identification accuracy. Specifically, we use the visualization of the attention weight distribution to identify the importance of arc

TABLE 7. Performance evaluation of EArcNet and EArcNet-Lite under different testing conditions.

Model	Test item category	Sample number	Test accuracy
EArcNet-Lite	Inverter startup	60	100%
	Shadow occlusion	60	100%
EArcNet	Inverter startup	60	95%
	Shadow occlusion	60	100%

features and select critical feature bands in the photovoltaic system's arc energy spectrum. This interpretable critical feature extraction method enables the model to maintain high accuracy while reducing the number of input dimensions and avoiding frequency bands with interfering features.

2) The feature extraction pressure of the EArcNet-lite model is reduced. This method also helps to reduce the feature extraction pressure on EArcNet-Lite, by reducing its input data volume, parameter volume, and calculation volume, allowing EArcNet-Lite to be easily deployed on embedded microprocessors, promoting the industrial application of artificial intelligence technology in arc fault identification.

F. IMMUNITY TEST

The shading of PV panels and the startup process of inverters interfere with PV arc fault detection. To test the anti-interference performance, 60 sets of samples were collected for each situation to form a dataset. Table 7 summarizes the results of the anti-interference tests. EArcNet-Lite, based on key arc features, can eliminate both interferences. EArcNet-Lite uses attention weights to extract feature bands that contribute the most to arc identification and avoid interference bands, thus improving the anti-interference ability of the detection scheme. Although EArcNet has a slightly higher accuracy in the validation set, it has two misjudgments in the inverter startup interference test, resulting in an accuracy decrease to 95% because both inverter startup and shading occlusion change the current magnitude of the PV system. The transient change in the time domain also affects frequency domain features, causing interference features to overlap with arc features, leading to network model misjudgment. In contrast, the current changes faster during inverter startup, and power electronic devices in the inverter also generate switching noise to interfere with arc fault identification, decreasing the resistance of the original scheme to interference during the inverter startup process.

G. HARDWARE VERIFICATION OF ARC DETECTION ALGORITHMS

To evaluate the performance of EArcNet-Lite, we separately validate the EArcNet-Lite model and the entire arc detection algorithm on i.MX RT1064. i.MX RT1064 With a 600 MHz ARM Cortex-M7 CPU and 1M capacity SRAM. In i.MX RT1064, the prestored data are used to verify the EArcNet-Lite model. First, the current data collected by the transformer are prestored in i.MX RT1064. Then each data sample is subjected to data preprocessing, and the processed joint

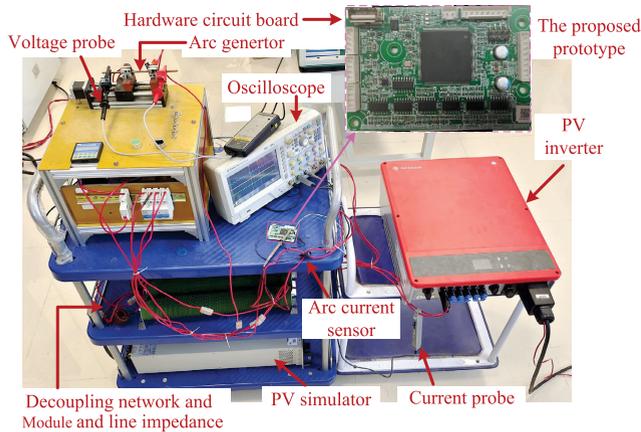


FIGURE 9. Prototype and experimental test bench for arc fault detection and performance evaluation.

energy spectrum data are input into the EArcNet-Lite model for testing. Finally, the average running time of all test samples is calculated. The final running time of the EArcNet-Lite model and data preprocessing on the i.MX RT1064 is 7.8 ms, which can meet the real-time requirements of photovoltaic arc detection.

The accuracy of the arc detection model is not 100%, which may cause the AFDD to malfunction. As a result, the running status cannot be obtained from detecting only a single piece of data. Therefore, we make the AFDD generate an action signal only when three arc fault signals are detected consecutively. Finally, we use the apparent diffusion coefficient(ADC) and timer on an i.MX RT1064 to collect the data and preprocess the collected data to test the entire arc detection algorithm. The verification platform of the arc fault detection algorithm is shown in Fig. 9, and the test results are shown in Fig. 10. According to the experimental results, the time required for AFDD to detect arc faults via the EArcNet-Lite detection algorithm is 43 ms, because the third sampling window is generated in the fourth sampling window. The test results show that the algorithm proposed in this paper can meet the real-time requirements of arc fault detection in PV systems.

Validating the interference resistance of EArcNet-Lite using the i.MX RT1064 microprocessor is equally crucial. Fig. 11 presents the interference test during the inverter startup process. Fig. 12 demonstrates the disturbance test conditions caused by intermittent weather factors such as shading on the PV system’s current variation. In these interference tests, the copper rod electrode maintains close contact without generating an electric arc, effectively creating short-circuit conditions; hence, the arc voltage remains consistent at zero. The results reveal that, regardless of whether the inverter is started or due to current variations induced by intermittent weather factors, EArcNet-Lite does not experience any false detections. In experimental testing, AFDD did not cause misjudgments or missed judgments in arc fault detection or anti-interference testing. Through the experimental results of the AFDD, we verify the reliability

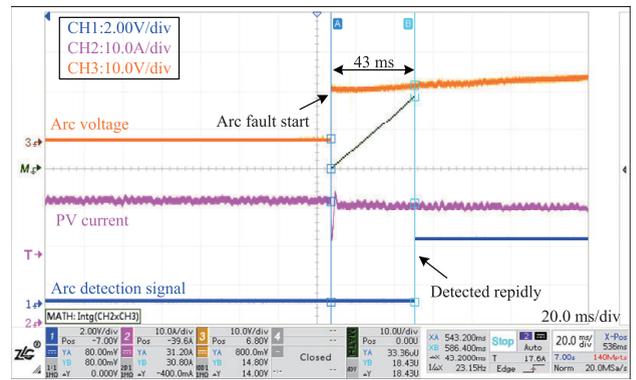


FIGURE 10. Response time of the proposed model for series DC arc fault detection in PV systems.

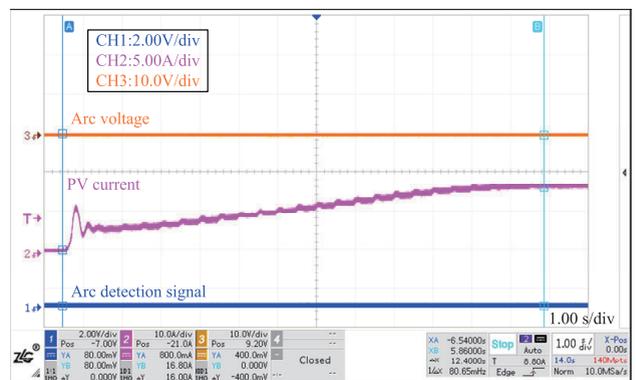


FIGURE 11. Test result of the proposed algorithm under inverter startup test conditions (no unwanted tripping).

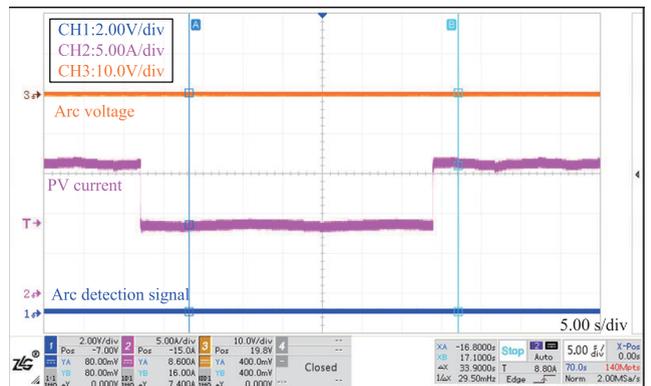


FIGURE 12. Test result of the proposed algorithm under intermittent weather test conditions of shadow occlusion (no unwanted tripping).

of EArcNet-Lite in operating on an arc database with an accuracy of 99.63%, further confirming that EArcNet-Lite not only possesses high computational speed but also exhibits remarkable resistance to interference.

H. COMPARISON AND DISCUSSION OF DIFFERENT METHODS

To better illustrate the effectiveness of the EArcNet-Lite algorithm, a comparison is made with other PV arc

TABLE 8. Comparative analysis of arc fault detection capabilities and operating environments: EArcNet vs. contemporary advanced methods.

Algorithm	Proposed method	A [16]	B [17]	C [43]	D [19]	E [28]	F [7]
classifier	Attention+CNN	CNN+LSTM	PSO-SVM	CNN	TCN	Random Forests	Threshold value
Sampling frequency	250 kHz	500 kHz	250 kHz	20 kHz	250 kHz	40 kHz	100 kHz
Accuracy	99.63%	98.43%	99.20%	99.4%	99.88%	98.93%	76.06% [17]
Platform	i.MX RT1064	High-Performance Computing Platform	i7-7700 CPU RTX2070GPU	cRIO-9030	Jetson Nano ARM-A57	ARM Cortex-A72	TMS320F28335
Run time	7.8 ms	4.3 ms	138 ms	60 ms	150 ms	250 ms	12 ms
Are different fault positions considered?	Yes	No	Yes	Yes	No	No	Yes
Is the inverter startup considered?	Yes	No	Yes	Yes	Yes	No	Yes
Is shadow occlusion considered?	Yes	No	Yes	Yes	No	No	Yes

detection algorithms. This comparison covers a variety of factors, including the type of classifier used, sampling rate, accuracy rate, verification platform, average running time, whether different arc occurrence conditions are considered, whether inverter startup interference can be avoided and whether shadow occlusion interference can be prevented. The algorithm verification platform and operation time in Table 8 are from the original paper.

Algorithm A [32] uses a convolutional neural network to extract static features and a long-short-term memory network to capture dynamic time-varying features. Algorithm B [17] employs variational modal decomposition to extract fault information from the current signal and a particle swarm optimization-based SVM for classification. Algorithm C [41] arranges the PV loop currents into a two-dimensional matrix for DC series arc fault diagnosis. Algorithm D [19] employs a temporal convolution network to extract features from current waveforms. Algorithm E [28] utilizes random forests to extract characteristics of electric arcs. Algorithm F [14] constructs three continuous windows of DC bus terminal voltage and current into an information matrix, which principal component analysis (PCA) decomposes.

Table 8 demonstrates the superiority of the EArcNet-Lite algorithm in terms of detection accuracy, speed, required hardware and computing resources. Although Algorithm A has a lower calculation time of 4.3 ms, its dependence on a high-performance computer platform limits its practical application. Additionally, Algorithm A does not consider shadow interference or different conditions of arc occurrence, which limits its application scope. The input feature selection of Algorithms A, B, and C relies on trial and error, while the EArcNet-Lite algorithm utilizes deep learning interpretability technology to select reliable arc features. While Algorithms D and E resolved the issue of extracting features from electric arc faults, they entail a substantial computational load, with processing times exceeding 150 ms. This is markedly greater than the 7.8 ms computational time achieved by the algorithm presented in this paper. Although Algorithm F requires the least computational resources and has a fast computation time, its dependence on a fixed threshold makes it unsuitable for application in power supply and distribution environments with considerable interference.

In summary, the EArcNet-Lite algorithm uses an attention mechanism to extract key features of fault arcs and uses a neural network to establish arc fault identification conditions, considering various experimental conditions and interference in UL1699B. In addition, larger current levels of 20 A and 25 A are considered according to the development trend of PV systems. It can be used in embedded microprocessors and has good running speed and detection accuracy, making it superior to other methods.

V. CONCLUSION

An attention-based approach called EArcNet is proposed for series arc fault detection in PV systems. The database used for model training included the test currents of 3 A, 8 A, 8.5 A, and 16 A, as specified by the UL1699B standard. In addition, larger current levels of 20 A and 25 A are considered according to the development trend of PV systems. The decision-making behavior of EArcNet is visualized using the interpretable principle of the attention mechanism, which can be helpful for generating lightweight input data and algorithm structures. Therefore, compared with the original model, EArcNet-Lite can achieve a detection accuracy of 99.63% with only 48.48% of the parameters. The data of two disturbance conditions, shadow occlusion and inverter startup, are collected for algorithm verification. Although the absolute accuracy of this method is not the highest compared with other methods, the proposed method is lightweight enough to be implemented by the Cortex M7-based microprocessor i.MX RT1064 as a commercial AFDD. In addition, the single runtime of the algorithm only needs 7.8 ms, which can fully meet the requirement of a UL1699B of less than 2500 ms.

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] H.-L. Dang, S. Kwak, and S. Choi, "DC series arc failure diagnosis using artificial machine learning with switching frequency component elimination technique," *IEEE Access*, vol. 11, pp. 119584–119595, 2023.
- [3] J. Kim and S. Kwak, "Detection and identification technique for series and parallel DC arc faults," *IEEE Access*, vol. 10, pp. 60474–60485, 2022.
- [4] A. Tang, Z. Wang, S. Tian, H. Gao, Y. Gao, and F. Guo, "Series arc fault identification method based on lightweight convolutional neural network," *IEEE Access*, vol. 12, pp. 5851–5863, 2024.

- [5] V. Psaras, Y. Seferi, M. H. Syed, R. Munro, P. J. Norman, G. M. Burt, R. Compton, K. Grover, and J. Collins, "Review of DC series arc fault testing methods and capability assessment of test platforms for more-electric aircraft," *IEEE Trans. Transport. Electrification*, vol. 8, no. 4, pp. 4654–4667, Dec. 2022.
- [6] Q. Xiong, X. Feng, A. L. Gattozzi, X. Liu, L. Zheng, L. Zhu, S. Ji, and R. E. Hebner, "Series arc fault detection and localization in DC distribution system," *IEEE Trans. Instrum. Meas.*, vol. 69, no. 1, pp. 122–134, Jan. 2020.
- [7] M. Ahmadi, H. Samet, and T. Ghanbari, "A new method for detecting series arc fault in photovoltaic systems based on the blind-source separation," *IEEE Trans. Ind. Electron.*, vol. 67, no. 6, pp. 5041–5049, Jun. 2020.
- [8] S. Chae, J. Park, and S. Oh, "Series DC arc fault detection algorithm for DC microgrids using relative magnitude comparison," *IEEE J. Emerg. Sel. Topics Power Electron.*, vol. 4, no. 4, pp. 1270–1278, Dec. 2016.
- [9] J.-B. Ahn, H.-B. Jo, and H.-J. Ryoo, "Real-time DC series arc fault detection based on noise pattern analysis in photovoltaic system," *IEEE Trans. Ind. Electron.*, vol. 70, no. 10, pp. 10680–10689, Oct. 2023.
- [10] K. Wang, X. Lin, F. Wei, and H. Wei, "Detection of series arcs in low-voltage AC power lines using load side voltage sag and waveform similarity comparison," *IEEE Access*, vol. 12, pp. 1483–1496, 2024.
- [11] X. Liu, "A series arc fault location method for DC distribution system using time lag of parallel capacitor current pulses," in *Proc. IEEE Int. Power Modulator High Voltage Conf. (IPMHVC)*, Jun. 2018, pp. 218–222.
- [12] X. Zhang, H. Wang, V. Stojanovic, P. Cheng, S. He, X. Luan, and F. Liu, "Asynchronous fault detection for interval type-2 fuzzy nonhomogeneous higher level Markov jump systems with uncertain transition probabilities," *IEEE Trans. Fuzzy Syst.*, vol. 30, no. 7, pp. 2487–2499, Jul. 2022.
- [13] Y. Wang, D. Sheng, H. Hu, K. Han, J. Zhou, and L. Hou, "A novel series arc fault detection method based on mel-frequency cepstral coefficients and fully connected neural network," *IEEE Access*, vol. 10, pp. 97983–97994, 2022.
- [14] Y. A. M. Alsumaidae, S. P. Koh, C. T. Yaw, S. K. Tiong, C. P. Chen, T. Yusaf, F. Benedict, K. Kardigama, T. C. Hong, and A. N. Abdalla, "Fault detection for medium voltage switchgear using a deep learning hybrid 1D-CNN-LSTM model," *IEEE Access*, vol. 11, pp. 97574–97589, 2023.
- [15] 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.
- [16] L. Xing, Y. Wen, S. Xiao, D. Zhang, and J. Zhang, "A deep learning approach for series DC arc fault diagnosing and real-time circuit behavior predicting," *IEEE Trans. Electromagn. Compat.*, vol. 64, no. 2, pp. 569–579, Apr. 2022.
- [17] X. Cai and R.-J. Wai, "Intelligent DC arc-fault detection of solar PV power generation system via optimized VMD-based signal processing and PSO-SVM classifier," *IEEE J. Photovolt.*, vol. 12, no. 4, pp. 1058–1077, Jul. 2022.
- [18] R. Jiang, G. Bao, Q. Hong, and C. D. Booth, "Machine learning approach to detect arc faults based on regular coupling features," *IEEE Trans. Ind. Informat.*, vol. 19, no. 3, pp. 2761–2771, Mar. 2023.
- [19] J. Yan, Q. Li, and S. Duan, "A simplified current feature extraction and deployment method for DC series arc fault detection," *IEEE Trans. Ind. Electron.*, vol. 71, no. 1, pp. 625–634, Jan. 2024.
- [20] Z. Wang, S. Tian, H. Gao, C. Han, and F. Guo, "An on-line detection method and device of series arc fault based on lightweight CNN," *IEEE Trans. Ind. Informat.*, vol. 19, no. 10, pp. 9991–10003, Oct. 2023.
- [21] R. Jiang, Y. Wang, X. Gao, G. Bao, Q. Hong, and C. D. Booth, "AC series arc fault detection based on RLC arc model and convolutional neural network," *IEEE Sensors J.*, vol. 23, no. 13, pp. 14618–14627, Jul. 2023.
- [22] Y. Zhang, H. C. Chen, Z. Li, C. Jia, Y. Du, K. Zhang, and H. Wei, "Lightweight AC arc fault detection method by integration of event-based load classification," *IEEE Trans. Ind. Electron.*, vol. 71, no. 4, pp. 4130–4140, Apr. 2024.
- [23] Y.-L. Shen and Z. Xu, "Fast detection of weak arc faults based on progressive singular-value-decomposition and empirical analyses," *IEEE Access*, vol. 10, pp. 130586–130601, 2022.
- [24] G. Bao, R. Jiang, and D. Liu, "Research on series arc fault detection based on higher-order cumulants," *IEEE Access*, vol. 7, pp. 4586–4597, 2019.
- [25] R. Jiang, G. Bao, Q. Hong, and C. D. Booth, "A coupling method for identifying arc faults based on short-observation-window SVDR," *IEEE Trans. Instrum. Meas.*, vol. 70, pp. 1–10, 2021.
- [26] G. Zou, G. Fu, B. Han, W. Wang, and C. Liu, "Series arc fault detection based on dual filtering feature selection and improved hierarchical clustering sensitive component selection," *IEEE Sensors J.*, vol. 23, no. 6, pp. 6050–6060, Mar. 2023.
- [27] Y. Wang, C. Bai, X. Qian, W. Liu, C. Zhu, and L. Ge, "A DC series arc fault detection method based on a lightweight convolutional neural network used in photovoltaic system," *Energies*, vol. 15, no. 8, p. 2877, Apr. 2022.
- [28] Z. Yin, L. Wang, S. Yang, and Y. Gao, "A series arc fault diagnosis method in DC distribution systems based on multiscale features and random forests," *IEEE Sensors J.*, vol. 23, no. 3, pp. 2605–2617, Feb. 2023.
- [29] Z. Yang, D. Yang, C. Dyer, X. He, A. Smola, and E. Hovy, "Hierarchical attention networks for document classification," in *Proc. Conf. North Amer. Chapter Assoc. Comput. Linguistics, Human Lang. Technol.*, 2016, pp. 1480–1489.
- [30] O. Vinyals, L. Kaiser, T. Koo, S. Petrov, I. Sutskever, and G. Hinton, "Grammar as a foreign Language," in *Proc. NIPS*, 2015, pp. 2773–2781.
- [31] H. Song, D. Rajan, J. J. Thiagarajan, and A. Spanias, "Attend and diagnose: Clinical time series analysis using attention models," 2017, *arXiv:1711.03905*.
- [32] G. Lee, J. Jeong, S. Seo, C. Kim, and P. Kang, "Sentiment classification with word localization based on weakly supervised learning with a convolutional neural network," *Knowl.-Based Syst.*, vol. 152, pp. 70–82, Jul. 2018.
- [33] D. T. Tran, A. Iosifidis, J. Kannianen, and M. Gabbouj, "Temporal attention-augmented bilinear network for financial time-series data analysis," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 30, no. 5, pp. 1407–1418, May 2019.
- [34] X. Yao, L. Herrera, S. Ji, K. Zou, and J. Wang, "Characteristic study and time-domain discrete-wavelet-transform based hybrid detection of series DC arc faults," *IEEE Trans. Power Electron.*, vol. 29, no. 6, pp. 3103–3115, Jun. 2014.
- [35] B. Jo and S. Beack, "Representations of the complex-valued frequency-domain LPC for audio coding," *IEEE Signal Process. Lett.*, vol. 31, pp. 361–365, 2024.
- [36] Z. Wang and R. S. Balog, "Arc fault and flash signal analysis in DC distribution systems using wavelet transformation," *IEEE Trans. Smart Grid*, vol. 6, no. 4, pp. 1955–1963, Jul. 2015.
- [37] S. Wiegrefe and Y. Pinter, "Attention is not an explanation," in *Proc. Conf. Empirical Methods Natural Lang. Process. 9th Int. Joint Conf. Natural Lang. Process. (EMNLP-IJCNLP)*, 2019, pp. 11–20.
- [38] A. Galassi, M. Lippi, and P. Torrioni, "Attention in natural language processing," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 32, no. 10, pp. 4291–4308, Oct. 2021.
- [39] Y. Chai, Y. Zhou, W. Li, and Y. Jiang, "An explainable multi-modal hierarchical attention model for developing phishing threat intelligence," *IEEE Trans. Dependable Secure Comput.*, vol. 19, no. 2, pp. 790–803, Mar. 2022.
- [40] L. Van der Maaten and G. Hinton, "Visualizing data using t-SNE," *J. Mach. Learn. Res.*, vol. 9, no. 11, pp. 1–27, 2008.
- [41] 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.



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 Hebei University of Technology, Tianjin, China, in 2006, 2009, and 2012, respectively.

He is currently a 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.



JIawang ZHOU was born in Hebei, China, in 1998. He received the B.S. degree in electrical engineering and automation from Hebei University of Technology, Tianjin, China, in 2021, where he is currently pursuing the M.S. degree in electrical engineering.

His current research interests include emerging technology in AI-based fault diagnosis and intellectualizing of electrical apparatus.



KAMAL CHANDRA PAUL (Graduate Student Member, IEEE) received the bachelor's degree in electrical and electronic engineering (EEE) from Khulna University of Engineering and Technology (KUET), Khulna, Bangladesh, in 2009, and the M.S. degree in electrical engineering from the Ingram School of Engineering, Texas State University, San Marcos, TX, USA, in 2018. He is currently pursuing the Ph.D. degree in electrical engineering with the Department of Electrical and

Computer Engineering, The University of North Carolina at Charlotte, Charlotte, NC, USA.



TIEFU ZHAO (Senior Member, IEEE) received the B.S. and M.S. degrees from Tsinghua University, Beijing, China, in 2003 and 2005, respectively, and the Ph.D. degree from North Carolina State University, Raleigh, NC, USA, in 2010, all in electrical engineering.

He was a recipient of the 2015 IAS Andrew W. Smith Outstanding Young Member Award and the 2015 STEM Forward Young Engineer of the Year Award. He served as the Organizing Committee Member on several international conferences, including ECCE 2016–2019 and PEDG 2018. He is the Founding Chair of the IEEE Charlotte Section PELS/IAS/IES Chapter. He served as an Associate Editor for the IEEE JOURNAL OF EMERGING AND SELECTED TOPICS IN POWER ELECTRONICS.



DEJIE SHENG (Student Member, IEEE) was born in Weifang, Shandong, China, in 1998. He received the B.S. degree from Jiangsu University of Science and Technology, Zhenjiang, China, in 2020. He is currently pursuing the Ph.D. degree in electrical engineering with Hebei University of Technology, Tianjin, China. His current research interests include arc fault detection and electrical fire prevention.

...