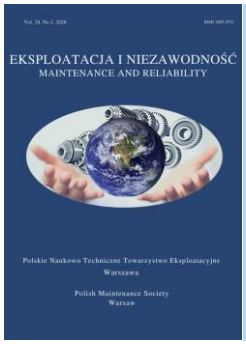


Eksploatacja i Niezawodnosc – Maintenance and Reliability
Volume 28 (2026), Issue 2
journal homepage: http://www.ein.org.pl

Article citation info:
Yu Q, Tan W, Wu Q, The Research on Low Voltage Series Fault Arc Diagnosis Based on Synchrosqueezing Wavelet Transform and Convolutional Neural Network, Eksploatacja i Niezawodnosc – Maintenance and Reliability 2026: 28(2) http:/10.17531/ein/213710

The Research on Low Voltage Series Fault Arc Diagnosis Based on Synchrosqueezing Wavelet Transform and Convolutional Neural Network

Indexed by: Web of Science Group

Qiongfang Yu[a,b,*], Wenxin Tan[a], Qiong Wu[a]

[a] School of Electrical Engineering and Automation, Henan Polytechnic University, Jiaozuo 454003, China
[b] Henan Key Laboratory of Intelligent Detection and Control of Coal Mine Equipment, Jiaozuo 454003, China

Highlights

- Proposed CASCNet, a ShuffleNetV2-CatBoost hybrid for improved arc fault classification.
- First proposed SWT-based fault arc feature extraction, enhancing accuracy significantly.
- Designed a multi-branch arc detection platform collecting multi-branch arc data.

Abstract

The operational dynamics of residential appliances impede arc fault identification, particularly in multi-branch circuits where fault signatures are obscured by load interference, increasing diagnostic complexity. To address this challenge, we develop the CASCNet diagnostic framework through synergistic integration of ShuffleNetV2's hierarchical feature learning and CatBoost's categorical processing. Firstly, synchrosqueezing wavelet transform（SWT） converts arc signals into 2D time-frequency images to enhance clarity and reduce noise. Next, ShuffleNetV2 efficiently extracts features, optimized with a CA attention mechanism to reduce complexity and improve channel interaction. Finally, CatBoost's robust generalization and noise resistance classify features. Experimental results show that the CASCNet method accurately identifies series arc faults and load types, maintaining good performance under noisy conditions. On the NVIDIA Jetson Orin Nano platform, CASCNet achieves an inference time of 4.2ms and an average accuracy of 99% for six-branch detection.

Keywords
Arc Fault Detection, Synchrosqueezing Wavelet Transform, ShuffleNetV2, CatBoost

1. Introduction

In recent years, with the rapid development of society and the ongoing rise in electrification levels, the incidence of electrical fires has markedly increased. Electrical fires have various causes, with arc faults in low-voltage lines being a key factor1.When a series fault arc occurs, the current change is not significant, and traditional electrical fault protection equipment is not suitable for such situations. Therefore, it is particularly important to develop a high-precision series arc fault (SAF) classification method in order to issue timely alarms when arc faults occur and ensure electrical safety2-3.

The series fault arc detection methods are mainly divided into three categories: Physical signature-based approaches utilizing multi-modal sensing (acoustic, optical, magnetic).The method is limited by the installation position of measurement sensors, coupled with the randomness and difficulty in locating fault arcs, and is commonly used to detect fault arcs in specific equipment such as switchgear and distribution cabinets4. (2) The detection method based on arc mathematical model requires a large number of parameters and is only applicable to specific circuits, which is still at the level of simulation

(*) Corresponding autho
E-mail addresses: Q. Yu (ORCID: 0009-0004-6525-5220) yuqf@hpu.edu.cn, W. Tan (ORCID: 0009-0004-1218-7156) twx001122@163.com, Q. Wu (ORCID: 0009-0003-6538-5749) 2603629422@qq.com

analysis5. (3) Extracting voltage and current features based on deep learning algorithms for arc fault detection has become the mainstream trend in low-voltage SAF research6-7.

The uncertainty of arc location means that the voltage-based series arc detection method requires data acquisition devices at multiple points, making on-site application more difficult8-9. At present, research mainly focuses on arc current signals, as they are easier to collect and analyze while solving the problems caused by the randomness of faulty arcs. Reference 10 extracts current signals and inputs them into a convolutional network for arc fault recognition. Reference 11 employs raw data input with a lightweight residual network (LRN) for arc fault detection. While temporal features offer extraction convenience, their instability and susceptibility to outliers limit reliability. To enhance diagnostic precision, we propose a hybrid method integrating frequency-domain and time-frequency analysis of current signals. Methods like Fast Fourier Transform (FFT) 12-13, wavelet transform14, Empirical Mode Decomposition (EMD) 15-16, and Hilbert Huang Transform (HHT) 17-18 effectively extract time-varying frequency components of current signals, facilitating detection of high-frequency arc fault traits.FFT is suitable for analyzing stationary signals and is computationally efficient, but it lacks temporal resolution and is difficult to capture time-varying characteristics. DWT is effective for non-stationary signals and can capture local features, but it requires a large amount of computation and is affected by the selection of fundamental waves. EMD can decompose nonlinear signals and extract local features, but it is susceptible to mode mixing and boundary effects. HHT provides instantaneous frequency, suitable for arc signals, but it is computationally complex and subject to boundary effects interference.

Reference 19 developed an arc fault detection method that integrates Variational Mode Decomposition (VMD), enhanced Multi-scale Fuzzy Entropy (IMFE), and Support Vector Machine (SVM). Reference 20 used wavelet transform, Fourier transform, periodic difference, and periodic similarity methods to extract features, and combined them with backpropagation (BP) neural network for fault arc detection. Reference 21 proposed an enhanced MFCC-FCNN method for series arc fault detection, combining Mel-frequency cepstral coefficients with a fully-connected neural network. However, MFCC-based approaches exhibit limitations such as noise susceptibility, limited temporal resolution, and high computational complexity.

The application of neural networks in fault diagnosis has significantly increased in recent years22. Reference 23 proposes a bearing fault diagnosis method based on multi-channel Transformer and transfer learning. Reference 24 proposes the AT-ICNN model, which combines CNN and hybrid attention mechanisms, significantly improving the accuracy of mechanical fault diagnosis. Scholars have mapped signals to high-dimensional space for fault diagnosis, further improving diagnostic performance. Reference 25 uses Short-Time Fourier Transform (STFT) for feature extraction and applies a SAGPSO-optimized CNN to diagnose short-circuit faults in medium-voltage direct current shipboard power systems. However, the fixed frequency resolution of STFT limits detailed signal analysis. Reference 26 extracted the 5-layer wavelet detail coefficients from the current signal and fused them into a matrix. This matrix was subsequently transformed into an RGB image via color mapping, and a ResNet was employed for arc detection. Reference 27 employed discrete wavelet transform with color mapping analysis to extract arc signal features, then utilized a Deep Residual Network (DRN) for fault detection. Discrete wavelet transform is susceptible to noise interference and has limited feature extraction capabilities. Reference 28 performed continuous wavelet transform on current signals and converted the results into grayscale images, using Deep Residual Shrinkage Network (DRSN) for SAF recognition.

In the study of two-dimensional inputs, image quality directly affects the feature extraction and classification performance of deep learning. Therefore, it is crucial for improving model accuracy to establish an effective signal conversion mechanism. Synchrosqueezing wavelet transform (SWT) leverages the time-frequency analysis strength of wavelet methods to reconstruct and compress frequency components, thereby enhancing frequency resolution. This method can effectively remove cross term interference, reduce energy diffusion, and present clearer time-frequency images29. This technology has demonstrated its advantages in processing non-stationary signals in fields such as mechanics30, medicine31, earth sciences32-33 and power systems34.

Most methods effectively detect single-load arc faults, but

the complexity of low-voltage distribution systems with multiple parallel branches reduces detection accuracy for multi-branch arc faults. To address this, a low-voltage multi-branch arc fault data collection platform was developed to study multi-branch arc fault characteristics. A detection method employing synchrosqueezing wavelet transform (SWT) and CASCNet was proposed. Current signals are processed into 2D time-frequency maps using SWT to improve time-frequency resolution. A ShuffleNetV2-based CNN extracts image features coupled with CatBoost classification for precise series arc fault detection.

The core innovations of this study include:

(1) A new fault arc classification model, CASCNet, has been proposed, which integrates the robust feature extraction ability of ShuffleNetV2 with the strengths of CatBoost in handling categorical features. Optimizing the structure of ShuffleNetV2 and introducing collaborative attention mechanism (CA) enhance the perception ability of key features, reduce computational costs, and improve feature expression accuracy.

(2) For the first time, synchrosqueezing wavelet transform is proposed for feature extraction of fault arcs. SWT operates as a wavelet-based signal rearrangement algorithm utilizing time-frequency analysis, which generates localized energy distributions to extract critical features and enhance fault detection accuracy.

(3) Per the UL1699-2011 standard, a multi-branch experimental platform for arc fault detection is developed. This platform simulates diverse on-site conditions and gathers various experimental data, enhancing the database's comprehensiveness and representativeness.

2. Network Model

This study proposes a hybrid model CASCNet that combines ShuffleNetV2 and CatBoost for fault arc classification. Figure 1 depicts the CASCNet architecture, which integrates a ShuffleNetV2-based feature extraction module with a CatBoost classification layer.

ShuffleNetV2 is a lightweight and efficient architecture that reduces computational and storage costs while maintaining high accuracy, widely used in image classification and object detection. Compared to traditional CNNs, it further minimizes computation while preserving accuracy through deep convolution, channel shuffling, and residual connections35-36.

However, its architecture lacks an efficient attention mechanism for boosting feature learning capability. To address this, we integrate Coordinate Attention (CA) into the last block of each stage, enabling the model to capture both channel and spatial information without significantly increasing complexity. Additionally, to optimize computational efficiency, we reduce the number of stacked basic units to 2 and 3 per stage and eliminate the third stage entirely. Finally, the features extracted by ShuffleNetV2 are classified using CatBoost, a gradient-boosting decision tree algorithm, to achieve accurate fault detection.

The input is a 64×64×1 grayscale image, and after 3×3 convolution and 3×3 max pooling, the output feature map size becomes 16×16×48. Next, after ShuffleV2Block unit processing, each stage is followed by a CA attention mechanism, which does not change the number of channels. The output of Stage 1 is 8×8×96, The output of Stage 2 is 8×8×96, and then the number of channels is expanded to 1024 through a 1x1 convolution. Convert to a vector of size 1024 through global average pooling. The feature vector is input into CatBoost classifier to output category prediction. Finally, the model performance is evaluated by accuracy.

To ensure consistency and reproducibility in the training process, Table 1 lists the hyperparameters, while Table 2 presents the configuration of the CNN model architecture. The style and format of Table 2 follow the original ShuffleNetV2 paper to maintain consistency, transparency, and reproducibility in the experimental design and results.

Table 1. Hyperparameter Settings

Hyperparameter	Value
Batch Size	32
Epochs	20
Learning Rate	0.0001
CatBoost Iterations	100
CatBoost Depth	6

Table 2. CNN Model Architecture and Parameters.

Layer	Output Size	Kernel Size	Stride	Repeat
Input	64×64×1	-	-	-
Conv1	32×32×24	3×3	2	1
Max Pooling	16×16×24	3×3	2	1
Stage 1	8×8×48	-	2 1	1 2
Stage 2	4×4×96	-	2 1	1 3
Conv5	4×4×1024	1×1	1	1
Global Average Pooling	1×1×1024	-	-	-

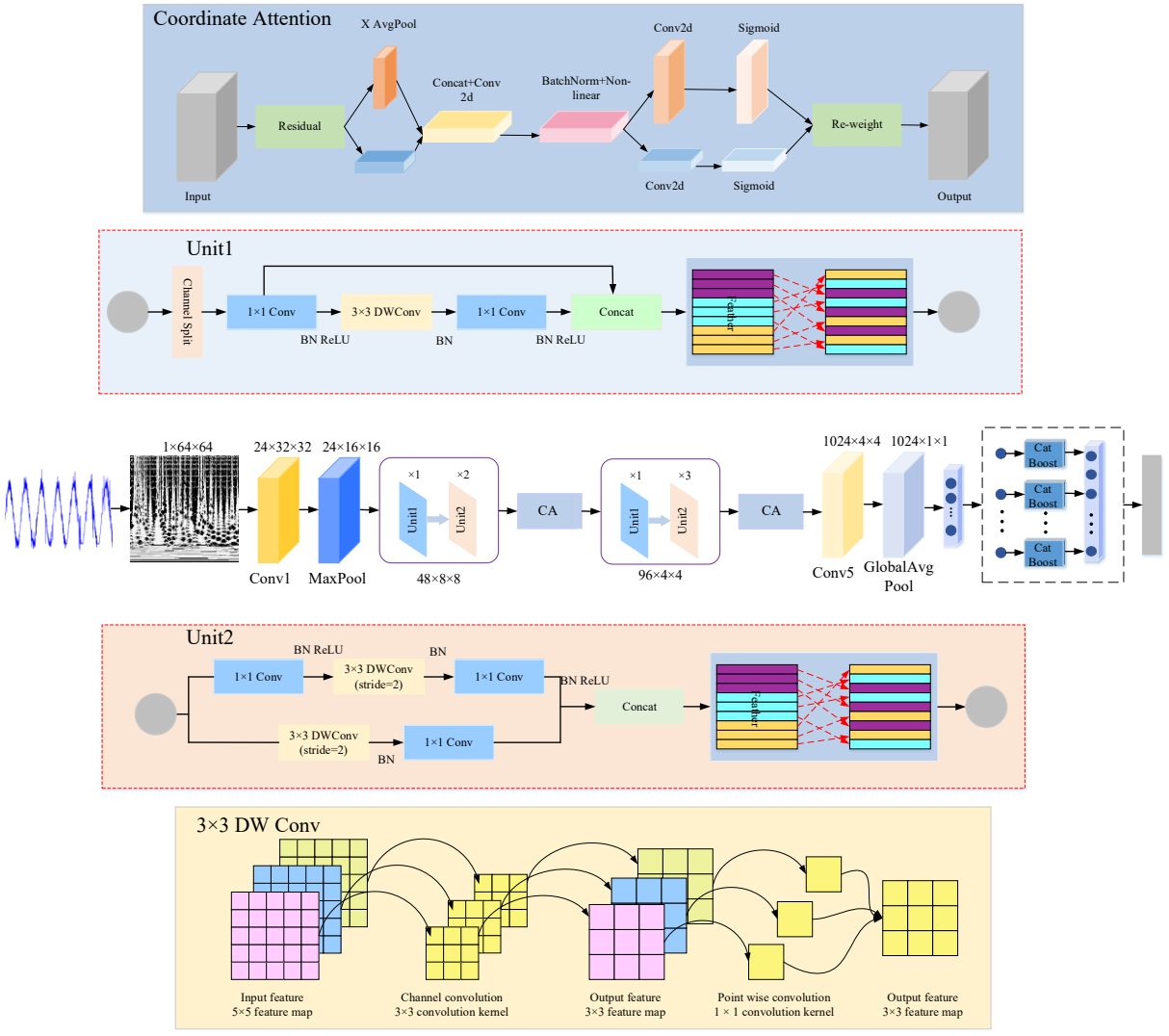


Fig. 1. CASCNet network mode.

2.1. Synchrosqueezing Wavelet Transform

The synchrosqueezing wavelet transform (SWT), proposed by Daubechies et al. (2011), enhances time-frequency resolution via wavelet coefficient compression in the frequency domain, focusing signal energy near true frequencies. The method's implementation involves three key phases:

(1)The Continuous Wavelet Transform (CWT) of signal $s(t)$ is calculated as:

$$W_s(a,b) = \frac{1}{\sqrt{a}} \int s(t) \psi^* \left(\frac{t-b}{a} \right) dt \quad (1)$$

Here, a denotes the scale factor, b represents the time shift factor, and ψ^* indicates the wavelet basis's complex conjugate. The output is then transformed into the frequency domain using a Fourier transform.

$$W_s(a,b)_f = \frac{1}{2\pi}\int \hat{s}(\varepsilon)a^{-\frac{1}{2}}\hat{\psi}^*(a\varepsilon)e^{ib\varepsilon}d\varepsilon \quad (2)$$

Here, ε denotes the angular frequency, while $\hat{s}(\varepsilon)$ and $\hat{\psi}^*(\varepsilon)$ represent the Fourier transforms of $s(t)$ and $\psi^*(t)$, respectively.

(2)The instantaneous frequency $W_s(a,b)$ is obtained from the signal's phase stability. Due to wavelet coefficient dispersion in the time-scale domain and the scale a invariance of signal phase, the phase derivative provides:

$$\omega_s(a,b) = \begin{cases} \frac{-i\,\partial_b W_s(a,b)}{W_s(a,b)} & |W_s(a,b)| > 0 \\ \infty & |W_s(a,b)| = 0 \end{cases} \quad (3)$$

(3) Wavelet coefficients are reassigned from the time-scale domain $W_s(a,b)$ to the time-frequency domain $W_s[\omega_s(a,b),b]$, with subsequent frequency-directional compression. For discrete application, the synchronized compression variable is specified as:

$$T_s(\omega_l,b) = (\Delta\omega_l)^{-1}\sum_{a_k:|\omega(a_k,b)-\omega_l|\leqslant\Delta\omega/2} W_s(a_k,b)a_k^{-\frac{3}{2}}\Delta a_k \quad (4)$$

In the formula, ω_l denotes the l-th discrete frequency value, a_k represents the kth discrete scale factor, $\Delta a_k = a_k - a_{k-1}$,$\Delta\omega_l = \omega_l - \omega_{l-1}$.

In the case of continuous variables, it becomes:

$$T_s(\omega,b) = \int_{A(b)} W_s(a,b)a^{-3/2}\delta(W(a,b)-\omega)da \quad (5)$$

Select Morlet wavelets as the wavelet basis function for synchrosqueezing wavelet transform by using the ssq_cwt function in the ssqueezepy library in Python, and then perform time-frequency analysis. After 6 layers of Morlet wavelet decomposition, each signal generates a time-frequency map Tx with the shape of (50,10000), representing the time-frequency information of 50 frequencies and 10000 time points. Finally, the time-frequency image is transformed into a (64,64) grayscale image and stored as a standard PNG file, with labels, as training data for the deep learning model.

2.2. CA Attention Mechanism

When conducting low-voltage series arc fault detection, the ShuffleNetV2 model faces issues with channel-independent processing and feature redundancy, impacting feature extraction efficiency. To tackle these challenges, this paper introduces the Coordinate Attention (CA) mechanism, depicted in Figure 2.

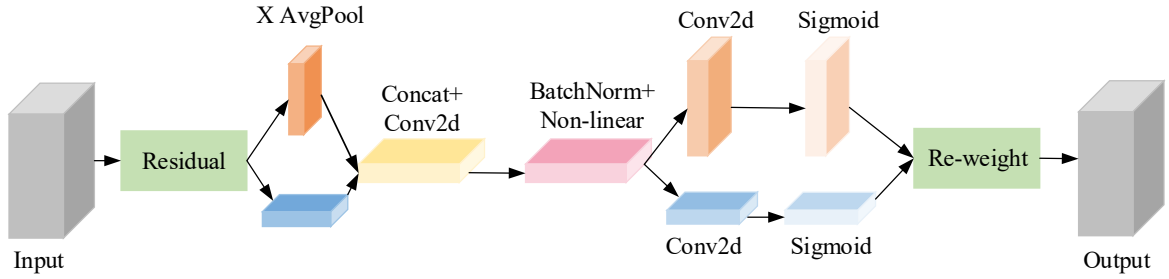


Fig. 2. CA attention module.

The CA mechanism accounts for both channel data and directional positional information. The formula for the CA mechanism is as follows:

Global average pooling in the height direction:

$$Z_c^h(h) = \frac{1}{W}\sum_{0\leq i<W} W_c(h,i) \quad (6)$$

Global average pooling in the width direction:

$$Z_c^w(w) = \frac{1}{H}\sum_{0\leq i<H} W_c(j,w) \quad (7)$$

After obtaining feature maps in two directions, the width and height feature information are transposed to the same dimension and stacked to obtain a feature layer of C×(H+W) ×1. Then, the F1 operation is performed:

$$F = \delta(F_1([Z^h,Z^w])) \quad (8)$$

In the formula, [,]represents the concatenation operation of spatial dimensions, and δ is a nonlinear activation function.

Subsequently, f is decomposed into fh and fw, corresponding to the vertical/horizontal axes of the feature map. A 1×1 convolution kernel is used to enhance dimensionality, with Sigmoid activation yielding the final attention vector:

$$g^{\square} = \sigma(F^{\square}(f^{\square})) \quad (9)$$

$$g^w = \sigma(F^w(f^w)) \quad (10)$$

$g^h$and g^w represent the final attention vectors in the height and width directions, respectively, σ represents the Sigmoid activation function.

The output formula after CA attention module:

$$W_c(i,j) = W_c(i,j)\times g^h(i)\times g^w(j) \quad (11)$$

2.3. CatBoost

CatBoost, a 2017 Yandex GBDT model, uses ordered boosting to cut gradients and bias, boosting generalization and category feature handling. It avoids one-hot encoding's high-dimensional issues, reducing overfitting and increasing accuracy. Its symmetric tree structure enhances efficiency, eases hyperparameter tuning, and fits large-scale data.

In arc fault diagnosis, CatBoost manages nonlinearity, noise, and complex arc signal features effectively. Its anti-noise capability ensures stability, while automated feature selection enhances fault recognition efficiency. It minimizes overfitting, improving model accuracy and reliability.

CatBoost processes training data for each permutation p by building n decision trees T1, T2, ..., Tn. T1 uses initial samples, with later trees trained on smaller sets to correct errors. The final prediction aggregates all tree outputs, averaging results across permutations to improve accuracy and reduce overfitting.

Multiple data permutations train the model, with predictions averaged for the final result.

For any arrangement $\delta = (\delta_1, \delta_2 \ldots \delta_N)$, there is:

$$x_{si,m} = \frac{\sum_{i-1}^{j=1} [x_{\delta i,m} = x_{\delta j,m}] y_{\delta j} + d \times p}{\sum_{i-1}^{j=1} [x_{\delta i,m} = x_{\delta j,m}] + d} \tag{12}$$

Among them, $p$is the prior value, d is the prior weight.

3. Experimental Design and Data Collection

3.1. Arc Fault Experimental Platform

Based on UL1699-2011 and GB14287.4-2014, a low-voltage arc fault data platform was built for single and multi-load parallel circuits (Fig. 3). The fault device uses a 6mm carbon rod (static) and copper rod (moving), adjusted by a stepper motor to trigger arcs. Powered by 220V/50Hz AC, 4mm copper wires ensure safety and signal stability. Two breakers control the circuit, with a disconnection point for protection. Six loads—heater, kettle, fan, AC, TV, PC—are paralleled, each with a switch to simulate branch faults. A JLT4 current transformer and TiePieSCOPE HS801 analyzer capture main current signals.

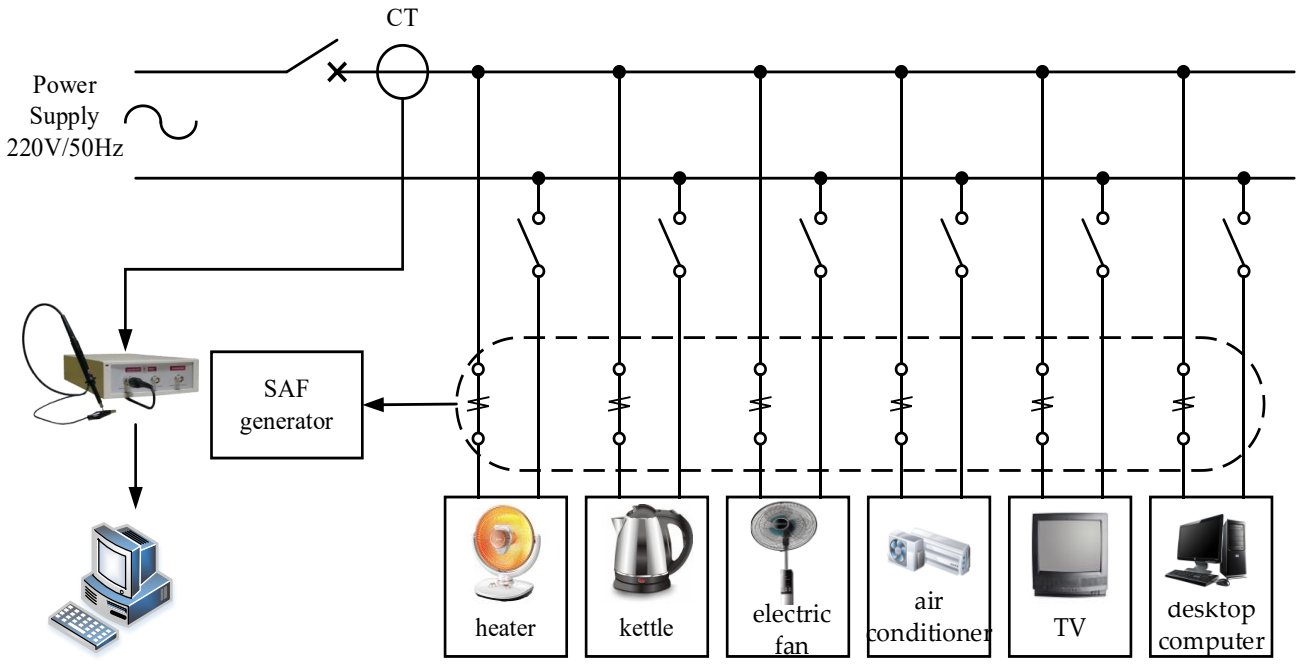


Fig. 3. Physical simulation diagram of the experimental platform.

3.2. Arc Fault Experimental Platform

The multiple sets of complex branch parallel experiments were designed in this paper to simulate the changing load environment in actual power systems. The experiment is divided into six categories, ranging from single branch to six branch loads. Each branch is configured with different types of loads, including resistive, inductive, and nonlinear loads, covering low-power to high-power devices to ensure the authenticity and reliability of experimental results. The attributes and parameters of the load are detailed in Table 3.

Table 3. Load properties and parameters.

Load name	Load	Load type	Rated power /W
heater	Load 1	Resistive load	400
kettle	Load 2		1500
electric fan	Load 3	Inductive load	60
air conditioner	Load 4		1500
TV	Load 5	Nonlinear load	120
desktop computer	Load 6		350

This paper established six experimental categories comprising 18 test cases to identify arc faults in electrical branches. Each type of experiment adopts a random load combination, covering resistive, inductive, and nonlinear loads, simulating various electrical devices in actual circuits. Six loads are connected in parallel on the main circuit, where switch S controls arc fault simulation. Specifically: when S is open, the arc generator activates to create a fault arc in the corresponding branch; closing S restores normal branch operation. As per Table 4, each experiment ensures only one branch emulates a fault while others function normally, with main circuit current signals acquired via a transformer. 3900 samples per category (normal/faulty) are collected and split into training/testing sets at a 10:3 ratio. In the 10 parts of the training set, we further divided it into 7:3 sections specifically for training and validation to ensure effective training and robust evaluation of the model.

Table 4. Experimental plan.

Group	Number of branches	Load numbers	Fault branch	Load attribute of the fault that occurred	Normal dataset	Fault dataset
1		2	2	resistive	1300	1300
2	1	4	4	inductive	1300	1300
3		5	5	nonlinear	1300	1300
4		1、4	4	inductive	1300	1300
5	2	3、5	5	nonlinear	1300	1300
6		2、5	2	resistive	1300	1300
7		1、3、5	5	nonlinear	1300	1300
8	3	1、4、5	4	inductive	1300	1300
9		2、4、6	2	resistive	1300	1300
10		1、3、5、6	5	nonlinear	1300	1300
11	4	1、3、4、6	4	inductive	1300	1300
12		1、3、5、2	2	resistive	1300	1300
13		1、3、5、4、6	4	inductive	1300	1300
14	5	1、2、3、5、6	2	resistive	1300	1300
15		1、2、3、5、6	5	nonlinear	1300	1300
16		1、2、3、4、5、6	2	resistive	1300	1300
17	6	1、2、3、4、5、6	4	inductive	1300	1300
18		1、2、3、4、5、6	5	nonlinear	1300	1300

To better classify arc fault load types, all detected fault samples were categorized into three groups based on inherent characteristics: resistive, inductive, and nonlinear loads. Normal operation samples formed a fourth category. Each category consists of 7,800 datasets, which are split into training and testing sets at a 10:3 ratio. The training set is further divided into training and validation sets with a 7:3 ratio.

3.3. Waveform Analysis

The experiment setup used a 50 kHz sampling rate (10,000 points/0.2s) to capture 10-cycle current waveforms. Signal duration refers to the total duration of the current waveform,

which is 0.2 seconds. Arc duration refers to the time when fault features appear in the waveform. If ≥8 cycles within the 10 cycles show fault features, it is identified as a fault signal. The experimental waveforms for three load types in a single branch are presented in Figure 4. The television, classified as a nonlinear load, shows minor waveform distortion under normal conditions while retaining periodicity. During an arc fault, the waveform displays peak pulses and abrupt changes. The resistive-load kettle maintained sinusoidal waveforms normally, developing zero-crossing "flat shoulders" with harmonic interference during arc faults. The inductive-load air conditioner displayed phase-lagged sine waves under normal operation, transitioning to distorted waveforms with both "flat shoulders" and sharp peak pulses when faulted.

The current waveforms obtained from the experiment of multiple branches with the same load (taking the kettle as the load in single, three, and six branches as an example) are shown in Figure 5. In normal operation, the current waveforms present an approximate sine wave shape. When an arc fault occurs in this branch, this stable state will be broken, and the waveforms will undergo significant distortion, manifested as abrupt peak pulses in the waveforms and a "flat shoulder" phenomenon at the zero crossing point. As the circuit structure becomes more complex, the operational stability of the circuit has been significantly improved. However, this stability also brings certain challenges, that is, the characteristics of the fault arc may become less obvious and more difficult to detect in complex circuit environments, thereby increasing the difficulty of accurately identifying arc faults.

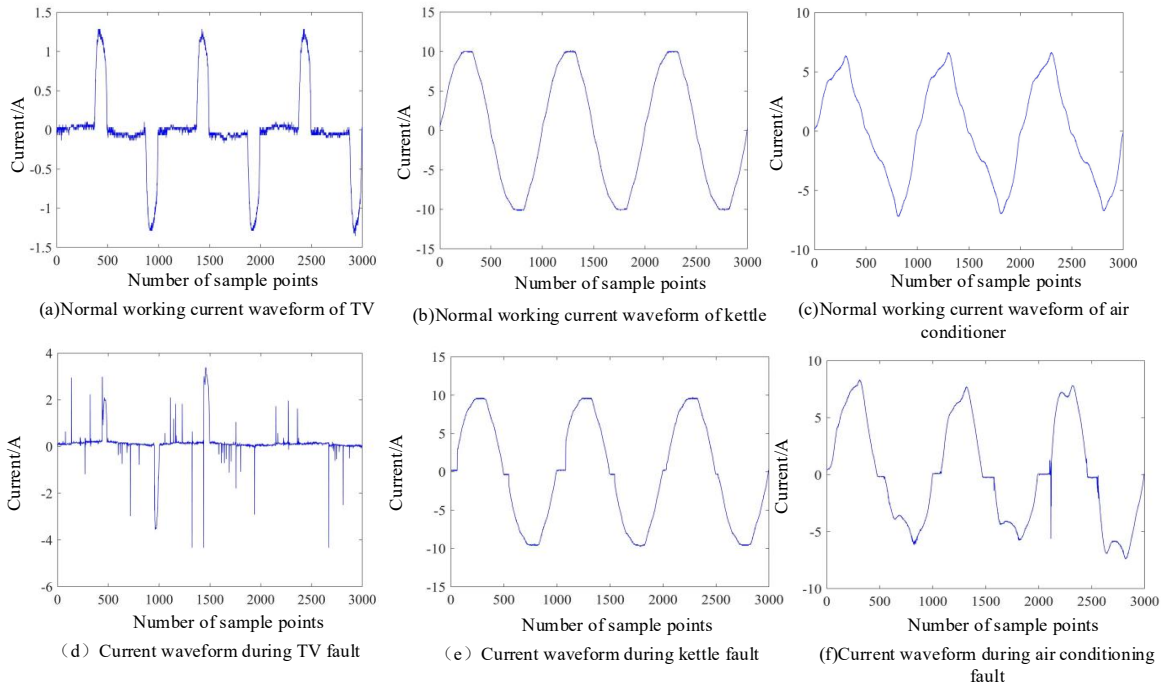
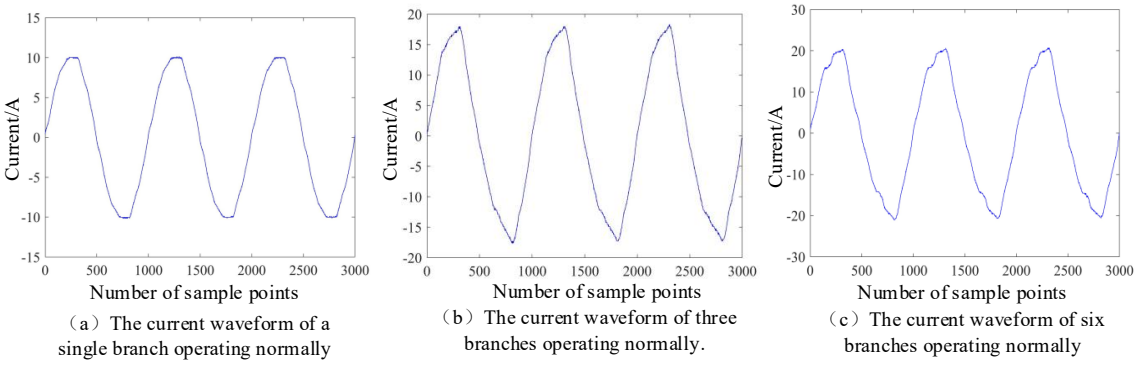
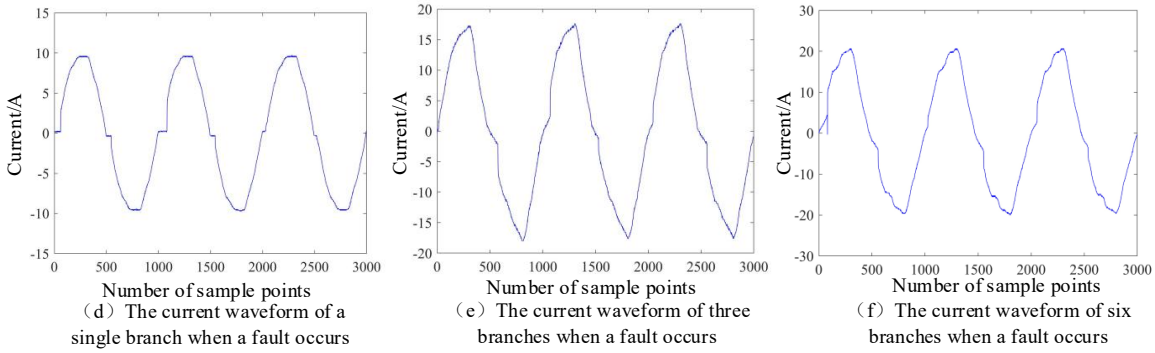


Fig.4. Single branch current waveforms with different loads.





（d）The current waveform of a single branch when a fault occurs

（e）The current waveform of three branches when a fault occurs

（f）The current waveform of six branches when a fault occurs

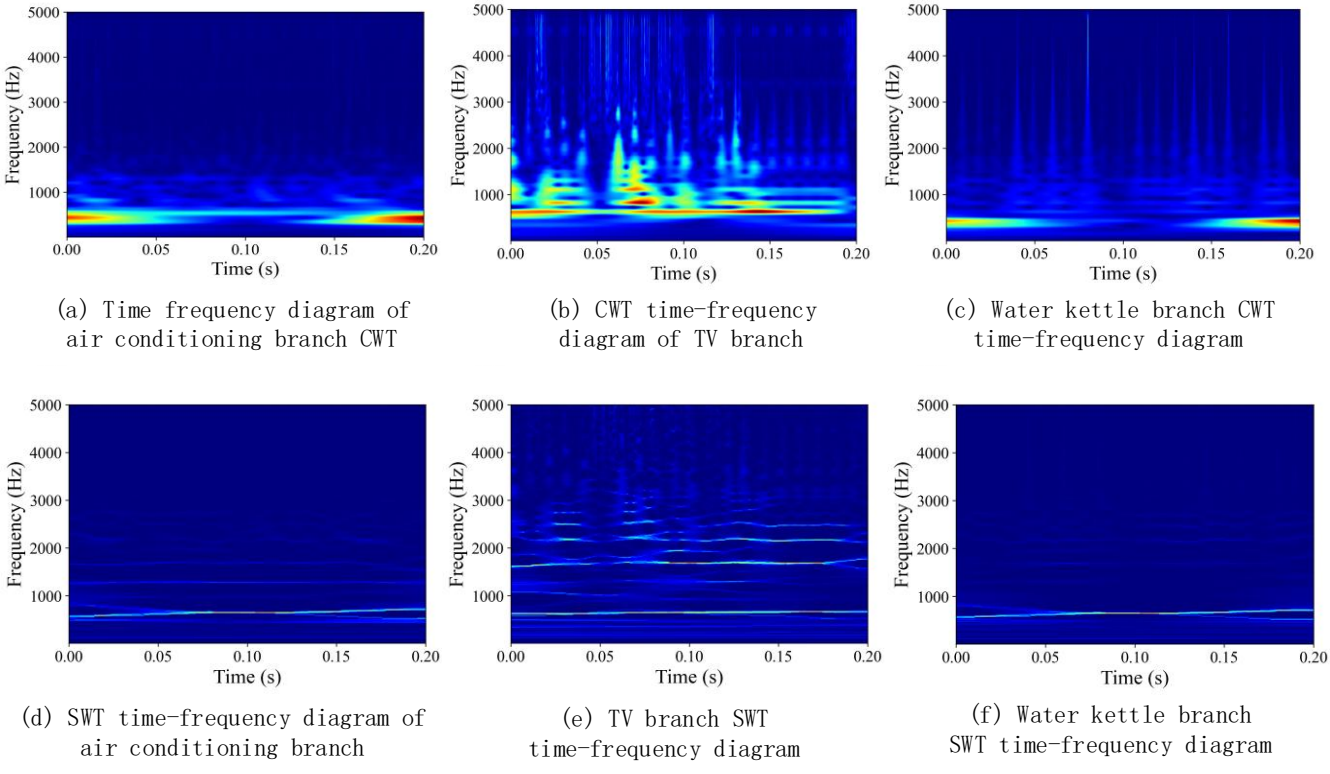
Fig.5. Current waveforms of the same load in different branches.

3.4. Comparison of Time-frequency Analysis Methods

SWT and CWT are applied to a signal duration of 0.2 seconds, extracting time-frequency features within the 0.2-second window. Figure 6 displays time-frequency images of current signals processed by CWT and SWT, with Rényi entropies calculated, as shown in Table5. Compared to CWT, SWT better captures local signal features with more concentrated energy. Table 3 indicates lower Rényi entropy for SWT images, reflecting higher time-frequency resolution and clustering. SWT employs phase-guided synchrosqueezing to reassign wavelet coefficients to true instantaneous frequencies, producing clear, compact images, whereas CWT yields blurred features due to coefficient dispersion, reducing extraction efficiency. Additionally, SWT's phase reassignment suppresses high-frequency noise, improving robustness in low-voltage systems. This high-resolution capability enhances SWT's ability to detect transient arc faults and extract key features efficiently. Hence, SWT images are selected as inputs for the convolutional neural network.

Table 5. The Rényi entropy of CWT and SWT time-frequency images.

Time-frequency map	Air conditioner	**TV**	Heate
CWT	12.40	13.48	12.93
SWT	11.06	11.95	11.31



(a) Time frequency diagram of air conditioning branch CWT

(b) CWT time-frequency diagram of TV branch

(c) Water kettle branch CWT time-frequency diagram

(d) SWT time-frequency diagram of air conditioning branch

(e) TV branch SWT time-frequency diagram

(f) Water kettle branch SWT time-frequency diagram

Fig.6. Different time-frequency analysis methods for time-frequency analysis.

4. Experimental Design and Data Collection

The experimental setup uses an Intel (R) Core (TM) I7-7700HQ processor with 16.00GB of RAM. After training the CASCNet model using PyCharm Community 2021, the converged model was quantized and transferred to the NVIDIA Jetson Orin Nano embedded development kit for deployment and inference testing. This was done to assess the real-time performance and robustness of the model in a resource-constrained environment. The appearance of the hardware platform is shown in Figure 7.

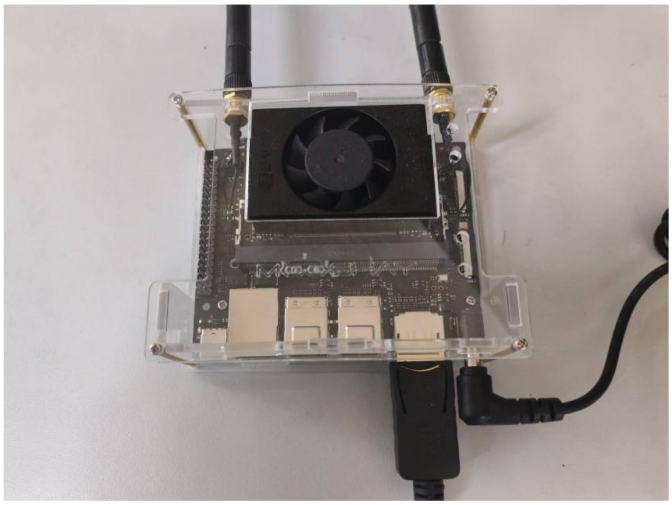


Fig.7. NVIDIA Jetson Orin Nano Developer Kit.

To assess the precision of network classification predictions, this paper adopts accuracy as the model evaluation metric, enabling clear observation of network performance.

Accuracy denotes the proportion of correct predictions within the total sample.

$$Accuracy = \frac{TP+TN}{TP+FN+FP+TN} \tag{13}$$

Among them, TP represents the actual count of correct categories predicted as true categories, TN indicates the actual count of incorrect categories classified as false, FP signifies the count of incorrect categories misclassified as true, and FN represents the actual number of correct categories predicted as false categories.

The dataset uses binary classification: "without arc" (normal, label 0) and "with arc" (fault, label 1). Samples with the same label are shuffled and then divided into training, validation, and testing sets.

4.1. Arc Fault Identification

During the training process, ShuffleNetV2 was first trained through multiple iterations using 6000 sets of normal and incorrect training data, without the involvement of CatBoost. After training, the feature tensors of the training data were extracted from the global average pooling layer and input into CatBoost for training. As training batches grow, the model's precision on the test set steadily rises. Taking a single branch as an example, the accuracy increases from about 0.7 to about 0.97. Figure 8 depicts the model's precision on the testing set, showing that it effectively captures the data's features and patterns.

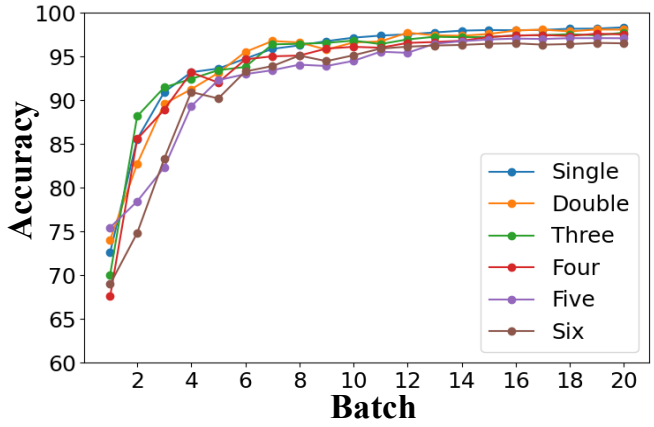


Fig.8. Training accuracy of different branches.

The processed training data from branches 2 to 6 were fed into the model for training, and the resulting model was assessed using these data. The CatBoost classifier has undergone random search optimization on key hyperparameters. To minimize randomness, the program was executed 5 times, and the average was calculated to determine the accuracy of each branch, as presented in Table 6. The results showed that CASCNet had better accuracy than ShuffleNetV2 in all branches, indicating that the model has stronger generalization and prediction abilities.

Table 6. Accuracy of different branch tests.

Method	ShuffleNetV2	CASCNet
Number of branches	Accuracy	Accuracy
1	97.39	99.16
2	97.11	99.05
3	96.94	98.94
4	96.72	98.83
5	96.38	98.78
6	96.16	98.68

When identifying the load type experiencing a fault, the "with arc" sample data are categorized into three types based on the load where the arc fault occurs: resistive, inductive, and nonlinear. Four categories are labeled as 0 (normal), 1 (resistive), 2 (inductive), and 3 (nonlinear), each containing 7,800 data points, which are split into training and testing sets at a 10:3

ratio. The training set is further divided into training and validation sets at a 7:3 ratio. The test set is then fed into the trained model to evaluate classification accuracy, as detailed in Table 7.

Table 7. Results for identifying load types.

Form	Accuracy/%	Overall accuracy/%
Normal	98.77	98.93
Resistance	99.06	
Inductive	99.00	
Nonlinear	98.87	

To visually assess the model's performance more effectively, we have drawn three confusion matrices, as shown in Figure 9. These matrices are used to represent arc fault detection in branches one through six and the identification of the faulty load type. When assessing the presence of an arc fault, the horizontal axis indicates the model's diagnostic outcomes, the vertical axis depicts the actual conditions, 0 represents normal operation, and 1 represents the occurrence of a fault. When identifying the type of fault load, the horizontal axis represents the model prediction result, the vertical axis represents the actual load type, 0 is normal operation, 1 is resistive load fault, 2 is inductive load fault, and 3 is nonlinear load fault. The numbers on the diagonal indicate diagnostic accuracy, and darker colors indicate higher accuracy. From the confusion matrixes, it can be seen that the misdiagnosis samples are fewer and overall accuracy is high using CASCNet to determine faults and their load types results.

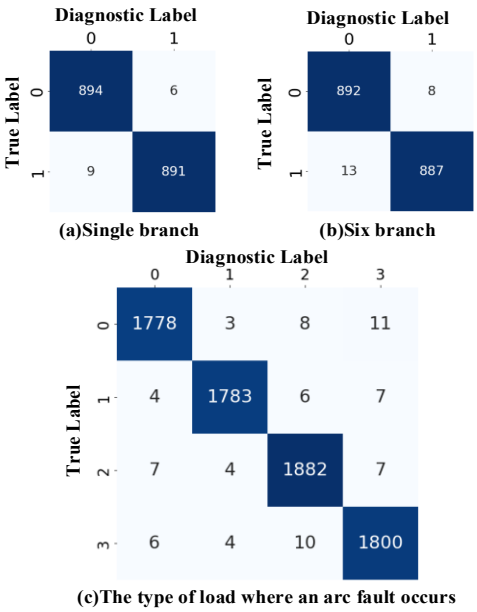


Fig.9. Confusion Matrix.

4.2. Ablation Experiment

In order to verify the contribution of each component to performance improvement, a formal ablation experiment was designed in this study. The experiment further validated the effectiveness of the hybrid architecture by systematically quantifying the role of each module. The experimental setup includes: (1) using only ShuffleNetV2 (2) ShuffleNetV2_Structural Adjustment (SA); (3) ShuffleNetV2_ (SA) _CA; (4) The CASCNet. The experiment was conducted on single branch and six branch datasets, with accuracy as the evaluation metric. The experimental results are shown in Table 8.

Table 8. CASCNet ablation experiment results.

Method	One branch accuracy/%	Six branch accuracy/%
ShuffleNetV2	96.28	95.56
ShuffleNetV2_(SA)	96.33	95.22
ShuffleNetV2 _ (SA) _CA	97.39	96.16
Method	One branch accuracy/%	Six branch accuracy/%
CASCNet	99.16	98.68

These experimental data effectively demonstrate the performance improvement effect of each module in the CASCNet model. Especially through CA attention mechanism, structural adjustment of ShuffleNetV2, and integration of CatBoost, the accuracy of the model is improved.

4.3. Model Anti Noise Capability Verification

To thoroughly validate the proposed method's enhanced noise resistance, a series of comparative experiments were designed in this study. Firstly, in response to the problem of a large amount of high-frequency electromagnetic noise introduced during the operation of nonlinear loads widely used in low-

voltage systems, such as computers and power adapters, these noises can cause interference components to be mixed in the circuit current waveform. To simulate this actual situation, Gaussian noise with a signal-to-noise ratio of 50dB was added to the current signals in the experiment, and the proposed model was used for fault diagnosis and performance analysis. To simulate this actual situation, Gaussian noise with a signal-to-noise ratio of 50dB was added to the current signals in the experiment, and the proposed model was used for fault diagnosis and performance analysis.

On this basis, further comparison and evaluation of the performance of our method are conducted from the following three aspects: (1) comparison of different input methods (such as CWT-CASCNet), (2) Comparison of different classifiers (such as SWT ShuffleNetV2, SWT ShuffleNetV2-RF), (3) Comparison with traditional rolling arc fault diagnosis methods (such as SWT-VGG16 and SWT AlexNet). The experimental results are shown in Table 9.

Table 9. Different models' classification results on Gaussian noise faults.

Method	One branch accuracy/%	Six branch accuracy/%	Determination of load type accuracy/%
CWT-ShuffleNetV2-CatBoost	96.11	95.42	95.56
SWT-ShuffleNetV2-Softmax	96.94	95.89	96.27
SWT-ShuffleNetV2-RF	97.50	96.27	97.06
SWT-VGG16	95.83	94.67	95.27
SWT-ResNet18	96.38	95.61	96.16
Proposed	98.89	97.94	98.50

The initial training model maintains high recognition accuracy despite Gaussian noise interference. The findings show that this approach exhibits strong noise resistance.

Under noise interference, when using CWT time-frequency images as input, the model's diagnostic precision is notably inferior to the approach presented in this paper. In contrast, SWT can more stably extract multi-scale signal characteristics, particularly under high noise, and it can more effectively preserve useful information of the original signals.After analysis, the main reason is that compared to synchronously compressed wavelet time-frequency maps, CWT is more susceptible to noise interference when processing noise. CWT may introduce high-frequency noise into time-frequency maps in multi-scale analysis, resulting in blurry signal features and decreased diagnostic accuracy. SWT, through its synchronous compression mechanism, is more effective in separating signals from noise, especially in the case of high noise, and can extract effective information of signals more stably at various scales, reducing the impact of noise on time-frequency images.

4.4. The Influence of Different Mother Wavelets on Accuracy

In this experiment, we selected three common wavelet bases (Morlet wavelet, Mexican Hat wavelet, and Haar wavelet), and tested them under three branch conditions in noise free and 50dB Gaussian white noise environments. The experimental results show that Morlet wavelets perform well under both signal-to-noise ratio conditions. Under 50dB Gaussian white noise, the accuracy of Morlet wavelet is 98.61%, while under noise free the accuracy is 98.94%. The specific results are shown in Table 10.

Table 10. The influence of different mother wavelets on accuracy.

Mother wavelet	The accuracy under noise free/%	The accuracy under 50dB Gaussian white noise/%
Morlet	98.94	98.61
Mexican Hat	97.39	96.38
Haar	97.83	96.89

4.5. Comparative Study of SWT and Traditional Time-Frequency Methods

In order to verify the advantages of SWT over traditional time-frequency analysis methods (such as CWT, STFT, and HHT), the accuracy of four methods under three branch conditions was compared in the experiment, and 50 dB of Gaussian white noise, harmonic noise, and switch transient noise were added to the testing of each method. The results are shown in Table 11.

Table 11. Comparison of accuracy of different time frequency analysis methods.

Method	The accuracy under noise free/%	The accuracy under 50dB noise/%
HHT	94.67	94.11
STFT	94.28	93.22
CWT	95.94	95.22
SWT	98.94	98.50

The experimental results show that the synchronized wavelet transform (SWT) exhibits high accuracy under both noisy and noiseless conditions. Under 50dB noise, SWT can still maintain an accuracy of 98.50%, which is superior to other methods. The accuracy of CWT, HHT, and STFT decreases significantly under noise interference, especially STFT, which is significantly affected by noise.

SWT improves resolution and effectively suppresses high-frequency noise through synchronous compression, making it ideal for capturing the transient changes of arc faults. In contrast, STFT struggles with non-stationary signals due to its fixed frequency resolution. CWT faces issues with resolution blurring when dealing with non-stationary signals, due to scale selection and redundancy. HHT is prone to modal aliasing, which reduces accuracy. Thus, SWT's high resolution and noise resistance outperform STFT, CWT, and HHT, making it particularly effective in arc fault detection.

In summary, the noise-related experiments validate the noise resilience and generalization capability of CASCNet combined with SWT in noisy environments. SWT-CASCNet effectively extracts signal features through the synchronous compression mechanism, significantly outperforming CWT, STFT, and HHT, demonstrating strong stability and adaptability. Additionally, compared to other neural networks, CASCNet shows higher accuracy and stability in handling noise interference. The specific experimental results are as follows:

(1) SWT-CASCNet maintains high accuracy under noise interference, outperforming both CWT and other neural networks.

(2) The Morlet wavelet shows stable performance in noisy environments, improving model stability.

(3) SWT effectively preserves signal information under noise interference, achieving higher accuracy than CWT, STFT, and HHT.

4.6. Comparative Experiment

In order to comprehensively evaluate the performance of CASCNet in practical applications, we conducted experiments on the case of three branches and compared the VGG16 and ResNet18 models. In this experiment, we focused on four indicators: accuracy across three branches, parameter count, FLOPs, and inference time. These indicators demonstrate the model's advantages in accuracy and computational efficiency. The experimental results are shown in Table 12.

Table 12. Comparison of model performance under three branch setting.

Method	The accuracy/%	Parameter count/M	FLOPs/G	Inference Time/ms
VGG16	95.55	39.93	1.273	1.53
ResNet18	96.06	11.19	2.227	1.97
CASCNet	98.94	3.22	0.246	1.18

The results of this experiment show that under the three branch setting, CASCNet not only has significant advantages in accuracy compared to VGG16 and ResNet18, but also performs more efficiently in terms of computational resource consumption.

4.7. Hardware Performance Evaluation

To verify the applicability of the model in resource constrained environments, we quantified the trained CASCNet model and deployed it in the NVIDIA Jetson Orin Nano development kit (see Figure 7). In the six branch detection task, CASCNet achieved an average accuracy of 99% and inference time of only 4.2 milliseconds. The generation time of SWT time-frequency graph on NVIDIA Jetson Orin Nano platform is 0.74 seconds, which fully proves the efficiency and reliability of the model in embedded systems.

5. Conclusion

This paper presents a method for arc fault detection based on SWT and CASCNet. By monitoring only the main circuit current, SAF fault detection can be implemented in multi-branch circuits and efficiently run on embedded

microprocessors. Experimental results show that when CASCNet runs on embedded devices, the average inference time per sample is 4.2 ms. Moreover, the method maintains high accuracy in a 50 dB Gaussian white noise environment, demonstrating strong noise resistance and confirming its potential for practical deployment on embedded microprocessors.

Although the Jetson Orin Nano platform outperforms other embedded devices (e.g., Raspberry Pi) in terms of performance, its higher cost limits its use in home arc fault circuit interrupters (AFCI or AFDD). Therefore, future work will focus on optimizing the SWT and CASCNet models to reduce hardware costs while maintaining detection performance, aiming to expand the application of home electrical fault monitoring systems.

Additionally, the variation in sampling frequencies caused by different data acquisition cards, a common issue in engineering practice, has not been fully addressed. Future work should investigate its impact on time-frequency images and fault detection, with an emphasis on improving the model's robustness under different sampling conditions.

Acknowledgments

This work is supported by the National Natural Science Foundation of China (61601172) and the China Postdoctoral Science Foundation Program (2018M641287)

References

1. Chen.Characteristics and fire-inducing risk analyses of arc faults in low-voltage electrical systems. Electric Power Systems Research 2025 ; 111199, https://doi.org/10.1016/j.epsr.2024.111199.
2. 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[J]. Elect. Power Syst. Res., 2017, 143: 130-139. https://doi.org/10.1016/j.epsr.2016.10.008.
3. L. Du, Z. Xu, H. Chen and D. Chen. Feature Selection-Based Low-Voltage AC Arc Fault Diagnosis Method [J]. IEEE Transactions on Instrumentation and Measurement, 2023, 72: 1-12. https://doi.org/10.1109/TIM.2023.3322486.
4. P. H. Schavemaker and L. van der Slui. An improved Mayr-type arc model based on current-zero measurements [circuit breakers][J]. IEEE Transactions on Power Delivery, 2000, 15(2): 580-584. https://doi.org/10.1109/61.852988.
5. W. B. Nottingham. A new equation for the static characteristic of the normal electric arc[J]. Journal of the American Institute of Electrical Engineers, 1923, 42(1): 12-19. https://doi.org/10.1109/JoAIEE.1923.6591851.
6. Z. Ming, Y. Tian and F. Zhang. Design of arc fault detection system based on CAN bus[J]. International Conference on Applied Superconductivity and Electromagnetic Devices, Chengdu, China, 2009. https://doi.org/10.1109/ASEMD.2009.5306631.
7. Huaijun Zhao, Jinpeng Liu, Junchao Lou. Series arc fault detection based on current fluctuation and zero-current features[J]. Electric Power Systems Research, 2022, 202:107626. https://doi.org/10.1016/j.epsr.2021.107626.
8. Kim, J.C., Neacşu, D.O., Lehman, B., Ball, R. Series AC Arc Fault Detection Using Only Voltage Waveforms[J]. Proceedings of the 2019 IEEE Applied Power Electronics Conference and Exposition (APEC), Anaheim, CA, USA, 17–21 March 2019: 2385–2389. https://doi.org/10.1109/APEC.2019.8721892.
9. Ke, Y., Zhang, W., Suo, C., Wang, Y., & Ren, Y. Research on Low-Voltage AC Series Arc-Fault Detection Method Based on Electromagnetic Radiation Characteristics[J]. Energies, 2022, 15(5): 1829. https://doi.org/10.3390/en15051829.
10. A. Tang, Z. Wang, S. Tian, H. Gao, Y. Gao and F. Guo, "Series Arc Fault Identification Method Based on Lightweight Convolutional Neural Network," in IEEE Access, vol. 12, pp. 5851-5863, 2024. https://doi.org/10.1109/ACCESS.2024.3350644.
11. Gao, X., Zhou, G., Zhang, J., Zeng, Y., Feng, Y., Liu, Y. Fault Arc Detection Based on Channel Attention Mechanism and Lightweight Residual Network[J]. Energies, 2023, 16: 4954. https://doi.org/10.3390/en16134954.
12. Gao, Yang et al. "Fault Diagnosis of Centrifugal fan Bearings Based on I-CNN and JMMD in the Context of Sample Imbalance." Eksploatacja I Niezawodnosc-maintenance And Reliability, vol. 26, no. 4, 2024. doi:10.17531/ein/191459.
13. J. E. Siegel, S. Pratt, Y. Sun and S. E. Sarma. Real-time deep neural networks for Internet-enabled arc-fault detection[J]. Eng. Appl. Artif. Intell., 2018, 74: 35-42. https://doi.org/10.1016/j.engappai.2018.05.009.
14. Z. Wang and R. S. Balog. Arc fault and flash signal analysis in DC distribution systems using wavelet transformation[J]. IEEE Trans.

Smart Grid, 2015, 6(4): 1955-1963. https://doi.org/10.1109/TSG.2015.2407868.

15. Jinde Zheng, Miaoxian Su, Wanming Ying, Jinyu Tong, Ziwei Pan. Improved uniform phase empirical mode decomposition and its application in machinery fault diagnosis[J]. Measurement, 2021, 179: 109425. https://doi.org/10.1016/j.measurement.2021.109425.
16. Diao, Y., Jia, D., Liu, G. et al. Structural damage identification using modified Hilbert-Huang transform and support vector machine[J]. J Civil Struct Health Monit 11, 2021: 1155-1174. https://doi.org/10.1007/s13349-021-00509-5.
17. Chunlin Xia, Yangfang Wu, Qianqian Lu, Bingfeng Ju. Surface characteristic profile extraction based on Hilbert–Huang transform[J], Measurement, 2014, 47: 306-313. https://doi.org/10.1016/j.measurement.2013.08.066.
18. Wang L, Qiu H, Yang P et al. Arc fault detection algorithm based on variational mode decomposition and improved multi-scale fuzzy entropy[J]. Energies, 14(14): 4137. https://doi.org/10.3390/en14144137.
19. GONG Q Y,PENG K,WANG W, et al. Series arc fault identification method based on multi-feature fusion [J]. Frontiers in Energy Research, 2022, 9: 824414. https://doi.org/10.3389/fenrg.2021.824414.
20. 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[J]. IEEE Access, 2022, 10: 97983-9799. https://doi.org/10.1109/ACCESS.2022.3206547.
21. Karabacak, Yunus. "Deep learning-based CNC milling tool wear stage estimation with multi-signal analysis." Eksploatacja I Niezawodnosc-maintenance And Reliability, vol. 25, no. 3, 2023. doi:10.17531/ein/168082.
22. Li X., Xiao S., Li Q., Zhu L., Wang T., Chu F. The bearing multi-sensor fault diagnosis method based on a multi-branch parallel perception network and feature fusion strategy[J]. Reliability Engineering & System Safety, 2025, 261: 111122. https://doi.org/10.1016/j.ress.2025.111122.
23. S. Zhi, K. Su, J. Yu, X. Li, and H. Shen. An unsupervised transfer learning bearing fault diagnosis method based on multi-channel calibrated Transformer with shiftable window[J]. Structural Health Monitoring, 2025, 24(1). https://doi.org/10.1177/14759217251324671.
24. Li X., Xiao S., Zhang F., Huang J., Xie Z., Kong X. A fault diagnosis method with AT-ICNN based on a hybrid attention mechanism and improved convolutional layers[J]. Applied Acoustics, 2024, 225: 110191. https://doi.org/10.1016/j.apacoust.2024.110191
25. J. Mou, Q. Zhu, Y. Tian, and W. Xu, An SAGPSO-CNN approach for supervised fault classification in MVdc shipboard power systems, IEEE Sensors J., vol. 25, no. 12, pp. 22956-22966, Jun. 15, 2025, doi: 10.1109/JSEN.2025.3565747.
26. Park C J,Dang H L,Kwak S, et al. Detection algorithms of parallel arc fault on AC power lines based on deep learning techniques [J]. Journal of Electrical Engineering & Technology, 2022, 17(2): 1195-1205. https://doi.org/10.1007/s42835-021-00976-2.
27. Zhang, S., Qu, N., Zheng, T., Hu, C. Series Arc Fault Detection Based on Wavelet Compression Reconstruction Data Enhancement and Deep Residual Network[J]. IEEE Trans. Instrum. Meas. 2022, 71: 1-9. https://doi.org/10.1109/TIM.2022.3158990.
28. Hu C., Qu N., Zhang S. Series arc fault detection based on continuous wavelet transform and DRSN-CW with limited source data[J]. Sci Rep 12, 2022: 12809. https://doi.org/10.1038/s41598-022-17235-7.
29. Daubechies I, Lu J, Wu H T. Synchrosqueezed wavelet transforms：An empirical mode decomposition-like tool [J]. Applied and computational harmonic analysis, 2011, 30(2): 243-261. https://doi.org/10.1016/j.acha.2010.08.002.
30. S. Wang, X. Chen, C. Tong and Z. Zhao. Matching Synchrosqueezing Wavelet Transform and Application to Aeroengine Vibration Monitoring[J]. IEEE Transactions on Instrumentation and Measurement, 2017, 66(2): 360-372. https://doi.org/10.1109/TIM.2016.2613359.
31. Kabir MM, Tereshchenko LG. Development of Analytical Approach for an AutomDevelopment of Analytical Approach for an Automated Analysis of Continuousated Analysis of Continuous Long-Term Single Lead ECG for Diagnosis of Paroxysmal Atrioventricular Block[J]. Comput Cardiol (2010), 2014, 41: 913-916.
32. P. Wang, J. Gao and Z. Wang, Time-Frequency Analysis of Seismic Data Using Synchrosqueezing Transform[J]. IEEE Geoscience and Remote Sensing Letters, 2014, 11(12): 2042-2044. https://doi.org/10.1109/LGRS.2014.2317578.
33. Ashtari Jafari, Mohammad. Comparative Application of Time-Frequency Methods on Strong Motion Signals[J]. Advances in Civil Engineering, 2021. https://doi.org/10.1155/2021/9933078.
34. T. Jiang, B. Liu, G. Liu, B. Wang, X. Li and J. Zhang. Forced Oscillation Source Location of Bulk Power Systems Using Synchrosqueezing Wavelet Transform[J]. IEEE Transactions on Power Systems, 2024, 39(5): 6689-6701. https://doi.org/10.1109/TPWRS.2024.3351915.
35. G. Chen et al. Research on Transmission Line Hardware Identification Based on Improved YOLOv5 and DeblurGANv2[J]. IEEE Access, 2023, 11: 133351-133362. https://doi.org/10.1109/ACCESS.2023.3336905.

36. L. -L. Zhang, Y. Jiang, Y. -P. Sun, Y. Zhang and Z. Wang. Improvements Based on ShuffleNetV2 Model for Bird Identification[J]. IEEE Access, 2023, 11: 101823-101832. https://doi.org/10.1109/ACCESS.2023.3314676.