International Core Journal of Engineering Volume 12 Issue 2, 2026
ISSN: 2414-1895 DOI: 10.6919/ICJE.202602_12(2).0013

A Single-phase Grounding Fault Location Method for Active Distribution Networks based on CPCA-CNN

Tengfei Cha

School of Automation and Electrical Engineering, University of Jinan, Jinan 250022, China

Abstract

Aiming at the problems of weak fault features and localization difficulties in active distribution networks, this paper proposes a fault location method that integrates Channel Prior Convolutional Attention (CPCA) and Convolutional Neural Networks (CNN). A Butterworth band-pass filter is utilized to preprocess zero-sequence currents, filtering out power frequency components and high-frequency noise. The one-dimensional transient signals are then transformed into high-dimensional time-frequency images through Transient Extraction Transform (TET). A CNN-based model is constructed to perform deep learning and feature extraction from these images. Generalization tests verify the effectiveness of the proposed method, achieving an overall accuracy of 99.98%.

Keywords

Active Distribution Network; Fault Location; Transient Extraction Transform; Attention Mechanism.

1. Introduction

Single-phase grounding faults account for more than 80% of all faults in distribution networks. New technical guidelines have revised the handling protocol for single-phase grounding faults from "permitting operation with faults for 1–2 hours" to "safe arc suppression for transient faults and rapid isolation for permanent faults"[1]. Therefore, rapid fault localization is of great significance for enhancing power supply reliability and ensuring the safety of both equipment and personnel.

To address the challenges of single-phase grounding fault location in active distribution networks, extensive research has been conducted in academia. Based on the differences in the types of signals utilized and the processing methods employed, existing fault location methods can be primarily categorized into two major categories: traditional localization methods and artificial intelligence methods.

Traditional localization methods mainly include the fault feature method [2], the impedance method [3], and the signal injection method [4]. In the fault feature method, the steady-state component method locates the fault by analyzing the differences in amplitude and phase of zero-sequence current, voltage, or power; however, it is prone to failure under the over-compensation condition of the arc suppression coil. The transient quantity method utilizes the abundant transient information generated by the fault for identification through waveform similarity or energy distribution; however, in cases of high-resistance grounding or small initial phase angles, the feature signals are weak and require extremely high sampling frequencies. Although the impedance method has a simple principle, it is significantly affected by transition resistance and changes in system operation modes, making it difficult to guarantee localization accuracy. While the signal injection method can enhance fault features by actively changing system parameters, it requires additional equipment, which not only involves high costs but may also introduce new safety risks.

ISSN: 2414-1895 DOI: 10.6919/ICJE.202602_12(2).0013

In recent years, with the large-scale integration of Distributed Generation (DG) [5], distribution networks have gradually evolved into multi-source mesh structures with bidirectional power flow and dynamic network topologies. This evolution leads to severe distortion of fault signals, further complicating fault location [6]. In this context, artificial intelligence technology has demonstrated remarkable advantages in the field of fault diagnosis. Traditional machine learning methods, such as Support Vector Machines (SVM) [7] and Bayesian networks [8], rely heavily on manual extraction of time-frequency features; consequently, their generalization ability and robustness are often constrained by the quality of feature engineering. In contrast, deep learning methods can construct deep neural networks to adaptively mine high-dimensional fault features from raw data [9], effectively addressing the poor robustness inherent in manual feature extraction. To further exploit fault information, recent studies have attempted to transform fault waveforms collected at various measurement points into time-frequency images [10], subsequently utilizing deep learning models to extract texture and energy distribution features to achieve high-precision fault location.

This paper proposes a fault location method for active distribution networks integrating Channel Prior Convolutional Attention (CPCA) and Convolutional Neural Networks (CNN). First, a Butterworth band-pass filter is used to preprocess the zero-sequence current, filtering out power frequency components and high-frequency noise interference. Second, the Transient Extraction Transform (TET) is employed to transform the one-dimensional zero-sequence current data into high-resolution time-frequency images. Subsequently, the CPCA attention mechanism is introduced into the CNN architecture to adaptively focus the model on the feature channels containing key fault information by dynamically adjusting channel weights. Finally, the effectiveness and superiority of the proposed method are verified through generalization performance tests and comparative experiments.

2. Fault Location Method based on CPCA-CNN

2.1 TET Time-Frequency Analysis Method

Transient Extraction Transform possesses the advantage of high time-frequency resolution. It was initially applied to vibration signal processing in mechanical fault diagnosis and has recently been successfully utilized for fault characteristic characterization [11]. Its basic principles are as follows.

The Short-Time Fourier Transform (STFT) result of signal $s(t)$ is expressed as:

$$\begin{aligned} G(t,\omega) &= \int_{-\infty}^{+\infty} g_w(u-t)\cdot s(u)\cdot \mathrm{e}^{-\mathrm{j}\omega u}\mathrm{d}u \\ &= \int_{-\infty}^{+\infty} g_w(u-t)\cdot A\cdot \delta(u-t_\mathrm{b})\cdot \mathrm{e}^{-\mathrm{j}\omega u}\mathrm{d}u \\ &= A\cdot g_w(t_\mathrm{b}-t)\cdot \mathrm{e}^{-\mathrm{j}\omega t_\mathrm{b}} \end{aligned} \tag{1}$$

Where $G(t,\omega)$ is a bivariate function of t and ω; u is the integration variable; j is the imaginary unit; $s(u)$ is the result of replacing u with t in the analyzed signal $s(t)$; $g_w(t)$ is the Gaussian window, defined as $g_w(t)=\mathrm{e}^{-t^2/2}$; and $g_w(u-t)$ is the result of replacing $(u-t)$ with t in the Gaussian window $g_w(t)$.

The TET transformation result $T_e(t,\omega)$ of signal $s(t)$ is given by:

$$T_e(t,\omega) = G(t,\omega)\delta(t-t_\mathrm{L}(t,\omega)) \tag{2}$$

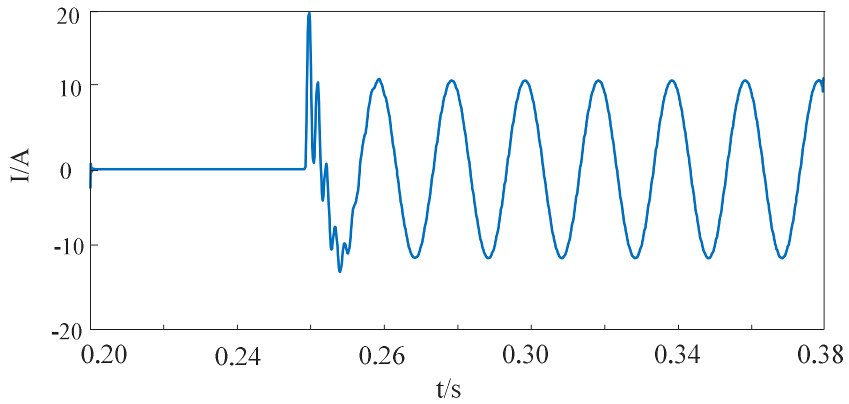
Where $T_e(t,\omega)$ is a bivariate function of t and ω; $\delta(t-t_\mathrm{L}(t,\omega))$ represents the TET transformation operator. The definition of $t_\mathrm{L}(t,\omega)$ is as follows:

ISSN: 2414-1895 DOI: 10.6919/ICJE.202602_12(2).0013

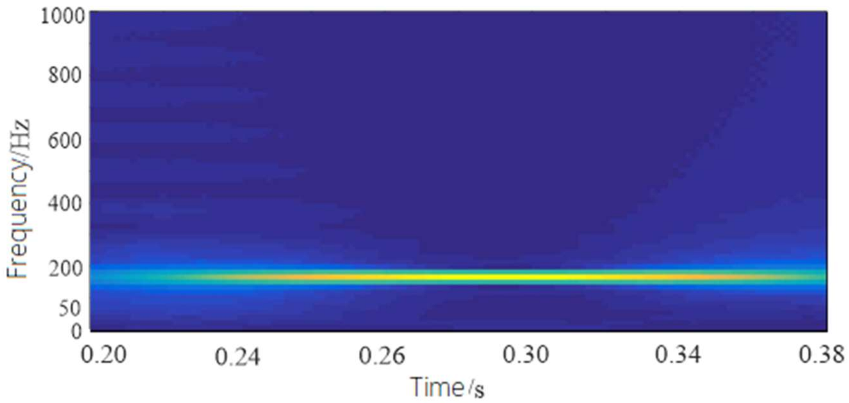
$$t_{\mathrm{L}}(t,\omega) = \mathrm{j}\frac{\partial_{\omega}G(t,\omega)}{G(t,\omega)} \tag{3}$$

Where $\partial_{\omega}G(t,\omega) = \partial G(t,\omega)/\partial\omega$.

From the above calculation process, it can be seen that TET is a post-processing procedure of STFT. In STFT, the window function leads to energy leakage into frequency bands adjacent to the instantaneous frequency ω_0, which in turn causes frequency divergence issues. By introducing the synchronization extraction operator shown in Eq. (3), TET can accurately characterize the time-frequency fault features of the zero-sequence current. The figure below shows a comparison between the zero-sequence current waveform and the TET time-frequency representation.



(a) Zero-sequence current waveform



(b) TET time-frequency image

Fig. 1 Comparison of zero-sequence current waveform and TET time-frequency representation

As shown in Fig. 1(a), although the zero-sequence current waveform can intuitively depict the instantaneous moment of fault occurrence and the subsequent transient oscillation process, it is difficult to effectively strip away and present the frequency and energy distribution characteristics inherent in the signal. This limitation is unfavorable for deep learning models to perform deep feature mining. In contrast, the TET time-frequency representation shown in Fig. 1(b) demonstrates excellent energy concentration. This high-definition time-frequency expression significantly enhances the discriminability of fault features, transforming complex electrical signal fluctuations into image features with strong spatial correlation. This is more conducive to the CPCA-CNN model's adaptive focus on key fault information, thereby improving fault localization accuracy.

ISSN: 2414-1895 DOI: 10.6919/ICJE.202602_12(2).0013

2.2 CPCA-CNN

To achieve precise identification of fault segments, a Convolutional Neural Network model integrating Channel Prior Convolutional Attention was constructed. The model takes the preprocessed TET time-frequency images as input, extracts high-dimensional spatial features through the CNN, and utilizes the CPCA mechanism to dynamically adjust the weights of feature channels, thereby enhancing the model's focus on critical fault information.

2.2.1 CNN

CNN is a deep learning model capable of adaptively extracting local features from images, consisting of convolutional layers, pooling layers, and fully connected layers. The convolutional layer performs sliding calculations on the TET time-frequency image via convolutional kernels to capture the texture and edge features of fault signals in the time-frequency domain.

The convolution operation is the core of the CNN, which extracts local features by sliding convolutional kernels across the input image. The calculation formula for convolution is given by:

$$X_j^l = f\left(\sum_{i\in M_J} X_I^{l-1} * k_{ij}^l + b_j^l\right) \tag{4}$$

Where X_j^l represents the j feature map output by the l layer; k_{ij}^l is the weight matrix connecting the i feature map of the l-1 layer with the j feature map of the l layer; b_j^l is the bias term; $*$ denotes the convolution operation; and $f(\cdot)$ is the ReLU activation function.

The features extracted via convolution then enter the pooling layer, where the maximum pooling operation is employed to reduce feature dimensionality. This process decreases the computational load while preserving the most prominent feature responses, thereby preventing model overfitting. Finally, the fully connected layer maps the extracted features into fault segment categories to achieve the ultimate fault classification.

2.2.2 CPCA

To further enhance the model's capability to extract key fault features, a CPCA attention module is introduced after the convolutional layers [12]. By applying weighting to the input features, the module concentrates the model's attention on the most critical regions. When training the CNN model integrated with the attention mechanism, the CNN focuses more intently on features relevant to the fault segments, thereby improving the accuracy of fault localization.

CPCA first performs global spatial aggregation on the input feature map $X \in R^{C\times H\times W}$, extracting channel attention features through parallel global average pooling and global maximum pooling. The calculation results are given by:

$$M_c(X) = \sigma(\mathrm{MLP}(\mathrm{AvgPool}(X)) + \sigma(\mathrm{MLP}(\mathrm{MaxPool}(X)) \tag{5}$$

Where $M_c(X) \in R^{C\times 1\times 1}$ represents the generated channel attention weight vector; σ is the Sigmoid activation function used to map weights to the interval [0, 1]; MLP denotes a multi-layer perceptron used to learn the non-linear dependencies between channels.

The generated attention weights are utilized to weight the original feature map, obtaining the weighted feature representation X':

$$X' = M_c(X) \otimes X \tag{6}$$

Where $\otimes$ denotes element-wise multiplication.

The different channels of the TET time-frequency representation correspond to time-frequency feature responses across various frequency bands. Based on the prior knowledge acquired during the training process, the CPCA mechanism adaptively increases the gain of channels containing critical transient components while suppressing ineffective channels heavily influenced by distributed generation or measurement noise. This feature filtering mechanism significantly enhances the model's robustness against complex operating conditions, such as varying transition resistances and initial phase angles, providing highly discriminative inputs for the subsequent precise classification by the fully connected layer.

In summary, a CPCA-CNN model for fault localization has been constructed, with its specific architecture illustrated in Fig. 2. Taking the zero-sequence current time-frequency representation as input, the network achieves fault localization by learning the mapping relationship between the fault features within the time-frequency representation and the corresponding fault segments.

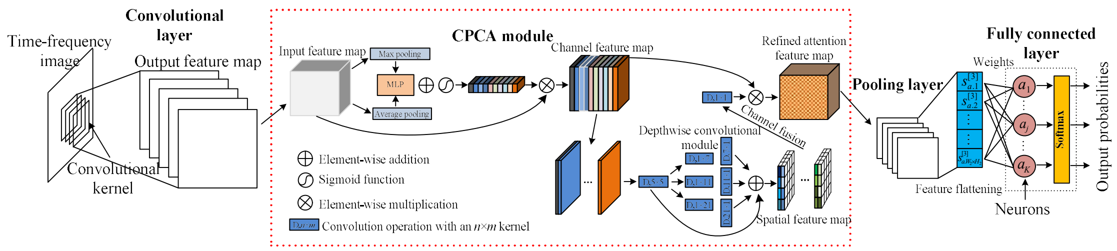


Fig. 2 Architecture of the proposed CPCA-CNN model

3. Fault Localization Methodology Flowchart

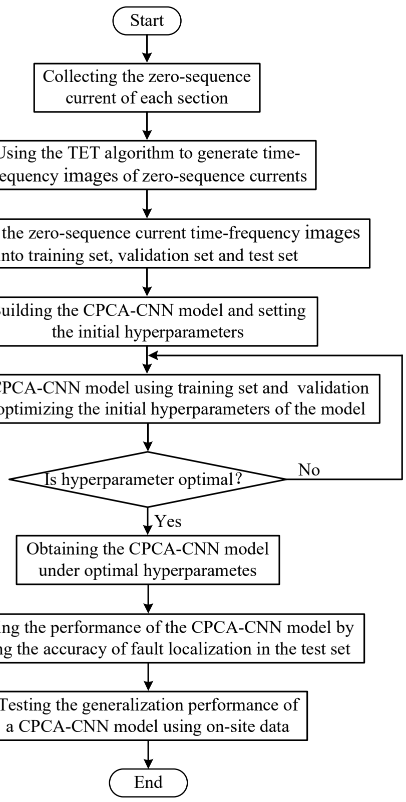


Fig. 3 Flowchart of fault location method

ISSN: 2414-1895
DOI: 10.6919/ICJE.202602_12(2).0013

The complete flowchart of the proposed fault localization method is illustrated in Fig. 3.

The specific steps are as follows:

1) Acquisition and processing of fault data

Fault data from various segments of the faulty lines are collected within the simulation software. Since high-frequency noise signals can cause distortion in the zero-sequence current waveforms, which hindering the model from accurately extracting fault features. The 50 Hz low-frequency signal may mask critical fault characteristics such as energy distribution, a Butterworth bandpass filter is employed to filter out the power frequency components and high-frequency interference from the zero-sequence current fault data.

Then, the TET is applied to generate time-frequency representations. The TET images from different segments are stacked vertically to form a single composite image. Finally, the processed time-frequency images are labeled with their corresponding segment tags and partitioned into a training set, a validation set, and a test set.

2) Training the TET-CNN fault localization model

The fault localization model is trained using the k-fold cross-validation method. The principle of its hyperparameter optimization is illustrated in Fig. 4.

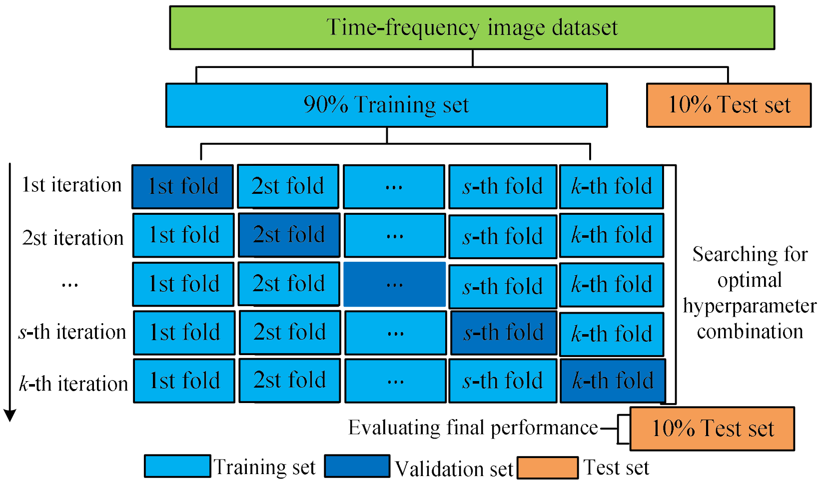


Fig. 4 Optimization principle of k-fold cross-validation

3) Online localization

The generalization performance of the trained model is evaluated using the fault data from the test set.

4. Simulation Verification and Analysis

4.1 Generation of Sample Set

An active distribution network simulation model simulating a substation scenario is constructed. The model consists of six lines, incorporating various topologies such as overhead lines, cables, and hybrid overhead-cable lines. Regarding segment partitioning, each line is divided into five independent fault segments based on 20% of its total length. To ensure the comprehensiveness of the samples, fault points are positioned at 5%, 25%, 45%, 65%, and 85% of the total length within each segment.

The simulation conditions are as follows: the transition resistances are set to1Ω,10Ω,30Ω,50Ω,100Ω,1000Ω,1300Ω,1600Ω and 2000Ω; three neutral grounding modes are included; the initial phase angles range from 0-360° with a step of 30°; the fault types cover single-phase-to-ground faults for phases A, B, and C; and a total of five fault locations are utilized. Based

ISSN: 2414-1895 DOI: 10.6919/ICJE.202602_12(2).0013

on these conditions, a total of 9×3×12×3×5=4860 fault recording datasets were established. Subsequently, 4,860 TET time-frequency representations were generated using the Transient Extraction Transform algorithm, establishing a solid foundation for model training.

4.2 Validity Verification Experiment

To verify the effectiveness of the CPCA-CNN model in the fault localization task, the training process of the model was analyzed. The curves illustrating the changes in the loss function and accuracy with respect to the number of iterations during the training process are shown in Fig. 5.

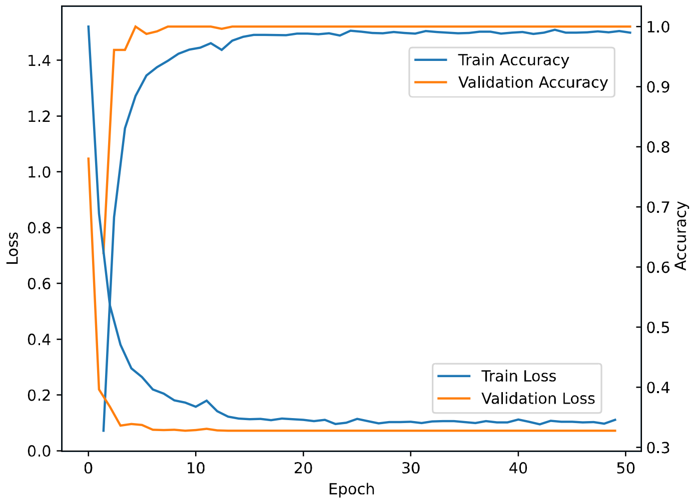


Fig. 5 Training process curve of the CPCA-CNN model

The analysis results indicate that in the initial stage of training, the loss values of both the training and validation sets decrease rapidly, while the accuracy increases significantly, which demonstrates that the model possesses strong learning capability and a fast convergence speed. When the number of iterations exceeds 20, the training and validation curves tend to stabilize and remain closely aligned. Ultimately, the training set accuracy stabilizes above 99%. Furthermore, the training loss value eventually stabilizes below 0.1, proving that the introduction of the CPCA module effectively suppresses the overfitting problem of the model. In summary, the proposed model exhibits excellent classification performance and robust generalization capability when processing complex TET time-frequency image features, enabling precise fault localization.

4.3 Comparative Experiments of Various Deep Learning Models

To further verify the superior performance of the proposed CPCA-CNN model in fault localization, a comparative analysis was conducted against the traditional CNN, Diffusion Model, and Transformer model.

Table 1. Model comparison experiments

Model	Validation set localization accuracy(%)	Loss function value
CNN	97.52	0.0552
CPCA-CNN	99.98	0.0005
Diffusion Model	98.20	0.0352
Transformer	99.02	0.0111

Analysis of the experimental results reveals that when processing complex time-frequency representations containing interference from distributed generation, the traditional CNN is susceptible to noise from ineffective channels, which limits its localization accuracy. Although the

Transformer model possesses global modeling capabilities, its sensitivity to local transient abrupt features is still slightly inferior to that of an attention-enhanced convolutional architecture. By integrating dual channel and spatial attention mechanisms, the CPCA-CNN model can adaptively amplify the weights of critical transient components while suppressing background noise. Consequently, it achieves an exceptional fault localization performance, with a validation set accuracy of 99.98%, which is higher than that of the other models.

5. Conclusion

This paper proposes a fault localization method for distribution networks combining TET and CPCA-CNN. The key conclusions are as follows:

(1) By utilizing bandpass filtering and high-resolution TET, signal interference is effectively eliminated and fault transient features are strengthened. This provides highly discriminative time-frequency inputs, significantly enhancing localization accuracy.

(2) The integration of the CPCA hybrid attention mechanism allows the model to adaptively focus on critical channels and spatial locations within feature maps. This enables the precise capture of subtle differences in zero-sequence currents before and after a fault.

(3) The model demonstrates exceptional robustness across various complex operating conditions, achieving a validation accuracy of 99.98%. It significantly outperforms traditional models such as CNN and Transformer in terms of both precision and convergence.

References

[1] Mirshekali H, Dashti R, Keshavarz A, et al. A novel fault location methodology for smart distribution networks[J]. IEEE Transactions on Smart Grid, 2021, 12(2): 1277-1288.

[2] Yao Z, Liu Y, Chen J, et al. Active high-impedance fault detection method for resonant grounding distribution networks[J]. IEEE Access, 2024, 12: 10932-10945.

[3] Dashti R, Daisy M, Shaker H R, et al. Impedance-based fault location method for four-wire power distribution networks[J]. Ieee Access, 2017, 6: 1342-1349.

[4] Jiang K, Wang H, Shahidehpour M, et al. Block-sparse Bayesian learning method for fault location in active distribution networks with limited synchronized measurements[J]. IEEE Transactions on Power Systems, 2021, 36(4): 3189-3203.

[5] Lin J, Guo M, Zheng Z. Active location method for single-line-to-ground fault of flexible grounding distribution networks[J]. IEEE Transactions on Instrumentation and Measurement, 2023, 72: 1-12.

[6] Zhang Z, Shi H, Li G, et al. A Method for Identifying The Medium-Voltage Phase-to-Phase Short-Circuit Fault Section in Active Distribution Networks Based on the Equivalent Impedance Ratio of Low-Voltage Monitoring Nodes[J]. IEEE Transactions on Smart Grid, 2025, 16(6): 4534-4548.

[7] Catucuamba J, Téllez A A. Design of a Generic Fault Diagnosis Model for Electrical Distribution Networks Using a Support Vector Machine (SVM) Algorithm[J]. IEEE Access, 2025, 13: 160175-160192.

[8] Jiang K, Wang H, Shahidehpour M, et al. Block-sparse Bayesian learning method for fault location in active distribution networks with limited synchronized measurements[J]. IEEE Transactions on Power Systems, 2021, 36(4): 3189-3203.

[9] Guo M F, Gao J H, Shao X, et al. Location of single-line-to-ground fault using 1-D convolutional neural network and waveform concatenation in resonant grounding distribution systems[J]. IEEE Transactions on Instrumentation and Measurement, 2020, 70: 1-9.

[10] Xu Y, Zhao C, Xie S, et al. Novel fault location for high permeability active distribution networks based on improved VMD and S-transform[J]. IEEE Access, 2021, 9: 17662-17671.

[11] Ma Y, Yu G, Lin T, et al. Synchro-transient-extracting transform for the analysis of signals with both harmonic and impulsive components[J]. IEEE Transactions on Industrial Electronics, 2024, 71(10): 13020-13030.

[12] Huang H, Chen Z, Zou Y, et al. Channel prior convolutional attention for medical image segmentation[J]. Computers in Biology and Medicine, 2024, 178: 108784.