Article citation info:
Chen J, Zhao T, Wang Y, Jin P, Lin C, Zhang H. Fault pattern recognition and prediction of transformer voicing based on neural network Diagnostyka. 2025;26(3):2025215. https://doi.org/10.29354/diag/210335.

DIAGNOSTYKA, 2025, Vol. 26, No. 3

e-ISSN 2449-5220
DOI: 10.29354/diag/210335

FAULT PATTERN RECOGNITION AND PREDICTION OF TRANSFORMER VOICING BASED ON NEURAL NETWORK

Jieyuan CHEN [1,*], Tiancheng ZHAO [2], Yiyu WANG [3], Peng JIN [2], Chuqiao LIN [2], Hengyuan ZHANG [3]

[1] Electric Power Research Institute of Jilin Electric Power Co., Ltd., Changchun 130012, Jilin, China
[2] State Grid Jilin Electric Power Co., Ltd., Changchun 130022, Jilin, China
[3] State Grid Jilin Electric Power Co., Ltd. Ultra-High Voltage Company, Changchun 132000, Jilin, China
* Corresponding author's e-mail: jieyuan_chen@outlook.com

Abstract

With the complexity of power system, transformer fault detection and early warning face challenges. Traditional methods fail to identify potential failures, leading to increased risk of equipment damage and power outages. In this study, an efficient fault recognition and prediction model based on voice print signal was developed by combining convolution neural network and long and short term memory network. After training and verification, the accuracy rate of the model's recognition of common fault modes reached 95.1%. In the complex fault mode, although the recognition rate has decreased, the whole early warning system still has high reliability and practicability. The research results provide technical support for the intelligent maintenance of transformers, and help to improve the safety and stability of power system.

Keywords: neural network; transformer; voiceprint fault

1. INTRODUCTION

With the accelerating expansion of power grids and the increasing complexity of energy infrastructure, this research emerges from the urgent need to modernize transformer fault diagnosis and early warning systems. Transformers, as core components of power transmission and distribution, are operating under greater thermal, electrical, and mechanical stresses than ever before. According to data released by the International Council on Large Electric Systems (CIGRÉ), transformer-related failures account for approximately 30% of major power outages in high-voltage networks worldwide, with 70% of those failures linked to internal insulation breakdowns, winding deformation, or core vibrations - many of which produce unique acoustic or voiceprint signals prior to catastrophic failure. Traditional offline testing and scheduled maintenance are no longer sufficient in identifying such complex or evolving faults in real time.

In response to these challenges, this study is grounded in the strategic direction of China's "New Infrastructure" initiative which prioritizes the intelligent upgrading of energy systems. The 2020 guideline issued by the National Development and Reform Commission explicitly called for the accelerated deployment of intelligent sensing, edge computing, and predictive maintenance technologies in key grid assets. This research aligns with that vision by developing a voiceprint-based intelligent diagnosis framework for transformers. Voiceprint signals, unlike conventional electrical parameters, carry rich frequency and energy domain information that is highly sensitive to mechanical anomalies. However, their complexity requires advanced computational models to effectively extract fault features.

To address this gap, this research integrates convolutional neural networks (CNN) and long short-term memory (LSTM) networks to build a high-accuracy, real-time identification and prediction system. CNNs are leveraged to capture the spatial and frequency-domain characteristics of voiceprint data, while LSTM networks retain the temporal evolution patterns of these acoustic features. This neural network-based approach not only responds to the grid's need for predictive diagnostics but also contributes to reducing unplanned outages, improving equipment life cycles, and ensuring grid stability. The methodology has been validated using over 320 GB of transformer acoustic data collected from Sichuan Electric Power Company between 2020 and 2023. The outcome of this research is not only a technical solution but also a strategic response to national goals for smart grid development and safe energy transition.

Received 2025-11-22; Accepted 2025-09-05; Available online 2025-09-10

Many scholars have contributed valuable insights and research outcomes to the field of transformer fault voiceprint recognition. Li proposed a recognition method based on blind source separation and convolutional neural networks, demonstrating its effectiveness in enhancing fault identification accuracy [1]. Yu explored a fault diagnosis technique using vibration and noise vibrographic imaging, showing that this approach can improve the accuracy and reliability of fault detection [2]. Shan developed a voiceprint-based fault identification model for motor bearings using the Mel-CNN architecture, verifying its applicability across various fault types [3]. Wan introduced a method that combines mixed data augmentation with voiceprint signals, highlighting the role of data augmentation in enhancing model generalization [4]. Jiang investigated voiceprint recognition under conditions of low false alarm rates, proving its effectiveness in reducing erroneous detections [5]. Ma proposed a novel motor bearing fault diagnosis approach using a physics-inspired sparse voiceprint sensing method, which significantly improved fault detection sensitivity [6]. Li examined transformer fault diagnosis using an improved Mel-frequency spectrum coefficient and a time convolution network based on multiple strategies, showing good robustness and accuracy under varying load conditions [7]. Jiang also proposed a method combining voiceprint features with an extreme learning machine for hydraulic pump fault identification and confirmed its practical applicability [8].

In modern power systems, transformers remain critical infrastructure, yet their fault detection and early warning systems still face significant challenges. Traditional approaches rely on manual inspections and scheduled maintenance cycles, which are often inefficient and inadequate for identifying potential issues before failures occur-particularly in the case of complex or sudden faults. Existing detection techniques often struggle to address diverse fault modes or extract meaningful patterns from large volumes of operational data. Consequently, power systems may fail to respond in time, leading to equipment damage, service interruptions, and other serious consequences. Therefore, a pressing need exists to apply advanced technologies that enable real-time monitoring and accurate prediction of transformer faults. The aim of this study is to introduce a neural network-based approach that integrates convolutional neural networks (CNN) and long short-term memory networks (LSTM) to analyze transformer voiceprint signals, accurately identify fault modes, improve diagnostic efficiency, and reduce the risk of unplanned outages.

This research employs deep learning techniques centered on neural networks, specifically combining CNN and LSTM architectures. CNN is used to extract spatial features from the voiceprint signals, while LSTM captures temporal patterns, enabling precise identification and prediction of fault modes through hierarchical feature analysis. Additionally, data augmentation and feature engineering are applied to enhance training effectiveness and ensure the model's robustness and generalization when addressing complex fault scenarios. The proposed methodology offers an innovative contribution to the field of transformer fault detection by bridging gaps in existing techniques. The expected outcomes include enhanced operational reliability of transformers, improved grid safety, and foundational support for the advancement of smart grid systems and intelligent power equipment management.

2. MATERIALS AND METHODS

2.1. Data collection and teleprocessing

In this study, data collection and preprocessing were essential steps to ensure the accuracy of the transformer voiceprint fault pattern recognition and prediction model. Voiceprint data and corresponding historical fault records from 500 transformers, spanning the period from 2020 to 2023, were obtained from Sichuan Electric Power Company of the State Grid. These records included a variety of typical fault modes encountered during equipment operation. The voiceprint signals were captured using high-precision acoustic sensors with a sampling frequency of 44.1 kHz, resulting in a total data volume of approximately 320 GB [9]. To improve data quality and optimize model training, noise filtering was applied to the raw signals, and approximately 12% of low signal-to-noise ratio samples were removed.

Feature extraction was conducted using a combination of Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA), resulting in the selection of 100 key feature vectors from the processed voiceprint data. To further enhance the generalization capability of the model, an additional 1,000 training samples were generated using time-frequency domain data augmentation techniques. These methods ensured greater data diversity and improved the model's capacity to recognize previously unseen fault modes. This comprehensive data preparation process provided a robust foundation for subsequent neural network model training.

2.1.1. Description of data sources

Table 1. Data Sources Overview

Source	Location	Voltage Level	Number of Transfo-rmers	Sensor Model	Sampling Rate (kHz)	Time Frame
Sound Data	Chengdu	110 kV	180	PT100	44.1	2020-2023
Sound Data	Mianyang	220 kV	200	PT100	44.1	2020-2023
Sound Data	Deyang	110 kV	120	PT100	44.1	2020-2023
Fault Records	Sichuan	Mixed	500	N/A	N/A	2018-2023

As shown in Table 1, the data sources in this study primarily consist of transformer voiceprint data and historical fault records. The voiceprint data were collected from 500 transformers rated at 110 kV and 220 kV, operated by State Grid Sichuan Electric Power Company. These transformers are located across several major substations in Sichuan Province, including Chengdu, Mianyang, Deyang, and surrounding areas. Data collection was conducted from January 2020 to December 2023 using high-precision PT100 acoustic sensors. The sampling frequency was set at 44.1 kHz to ensure high-fidelity signal acquisition. The historical fault records were obtained from the company's internal transformer maintenance database, which includes detailed information on the types of failures that occurred over the past five years, the timing of these events, corresponding repair actions, and relevant environmental conditions [10].

2.1.2. Data cleaning and feature extraction

During the data cleaning stage, the collected transformer voiceprint data underwent initial screening, resulting in the removal of approximately 12% of samples with low signal-to-noise ratios. These excluded samples were primarily from substations with significant noise interference, such as the 110 kV substation located in the northern suburbs of Chengdu. To more accurately capture the frequency characteristics of the signals, the time-domain data were converted into the frequency domain using Fourier Transform techniques [11]. For feature extraction, a combination of Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA) was employed to derive 100 key features from the original dataset. These features included frequency components, energy distribution, and transient characteristics. This sequence of processing steps significantly reduced and optimized the dataset, providing a solid foundation for subsequent model training and performance enhancement.

2.1.3. Feature engineering and data enhancement of voiceprint signals

In the process of feature engineering of voiceprint signal, the cleaned data is analyzed in time-frequency domain, and several key features including short-time Fourier transform and Meir frequency strumpet coefficient are extracted to capture important patterns in voiceprint signal. The characteristics can effectively reflect the small anomalies in transformer operation, such as frequency shift and energy mutation. In order to enhance the diversity of the data, data enhancement was performed on the original data through techniques such as time shift, frequency shift and noise injection to generate an additional 1000 samples. The enhanced sample can simulate the changes of voice print signal under different operating conditions, and improve the generalization ability of the model [12]. This process enriches the data set and provides a more comprehensive feature representation for subsequent fault pattern recognition, which helps to improve the accuracy and robustness of the model in practical applications.

To improve the generalization and robustness of the neural network model in fault pattern recognition, this study applied a systematic data augmentation pipeline to expand the diversity of voiceprint signal samples. Specifically, 1,000 additional training samples were synthetically generated from the original dataset using a combination of three core transformation techniques: pitch shifting, noise injection, and time warping. These techniques were selected based on their ability to simulate real-world variations in transformer operating conditions and to preserve key fault-related acoustic features.

Pitch shifting was applied within a range of ±2 semitones to emulate changes in mechanical resonance frequencies due to load variations or minor structural looseness. This transformation helps the model recognize the same fault type under slightly different harmonic characteristics. Noise injection involved adding Gaussian white noise with a signal-to-noise ratio (SNR) randomly selected between 20 and 30 dB, mimicking background substation noise, vibration, or electromagnetic interference. Time warping was performed by dynamically stretching or compressing short segments of the signal (up to ±10%) using a cubic spline interpolation method. This simulates minor temporal shifts in the acoustic profile caused by variable vibration cycles or transient system instabilities.

To maintain class balance and prevent overrepresentation, augmentation was performed in a stratified manner across the five primary fault modes (Modes A – E). Each fault mode received 200 augmented samples to ensure a uniform distribution across fault categories. After augmentation, the total number of samples for each mode increased to approximately 400 – 450, depending on the availability of clean base samples per category. This balanced approach ensures that the classifier does not develop a bias toward more frequent fault types and can better generalize to rare conditions. Post-augmentation, label integrity was maintained by verifying each generated sample's feature alignment with its corresponding fault type using principal component clustering. This validation confirmed that augmented data clusters were consistent with the original class boundaries in latent feature space, ensuring semantic consistency and training effectiveness.

These augmentation strategies not only expanded the dataset but also introduced realistic signal diversity reflective of operational field conditions. This was critical in reducing overfitting and improving model performance, particularly on fault modes with fewer original instances.

2.2. Model construction

2.2.1. Selection of neural network model

In the process of model selection in this study, considering the complexity and multi-dimensional characteristics of transformer voicing signals, conventional neural network is chosen as the main model architecture. CNN has unique advantages in processing two-dimensional data, such as images and time-frequency domain signals, which can effectively extract local features and maintain the integrity of spatial information. The frequency and time domain features of voiceprint signal are highly correlated, which is suitable for multi-level feature extraction through convolution operation. The convolution layer of CNN can carry out layer by layer convolution of input data through multiple convolution checks to extract feature graphs of different scales, as shown in formula (1).

$$Z_{i,j,k} = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} X_{i+m,j+n} \cdot W_{m,n,k} + b_k \quad (1)$$

$Z_{i,j,k}$ represents the output of the KTH convolution kernel at position (i,j), X is the input data, $W_{m,n,k}$is the weight of the convolution kernel, and b_k is the offset term.

To capture the temporal dependencies present in the signal, a Long Short-Term Memory (LSTM) network is introduced as a complement to the Convolutional Neural Network (CNN), enabling the model to handle long-range dependencies in sequential data. LSTM networks are effective at modeling memory and forgetting mechanisms, making them particularly well-suited for analyzing temporal features embedded in transformer voiceprint signals. By integrating CNN with LSTM, the model is designed to comprehensively analyze voiceprint signals in both the time and frequency domains, thereby enhancing the accuracy and reliability of fault mode identification and prediction.

The choice to combine a CNN with an LSTM is primarily driven by the complexity of voiceprint signals and their inherent time-frequency characteristics. CNNs are well-suited for processing two-dimensional data, such as time-frequency spectrograms, and are capable of efficiently extracting local features while identifying patterns at multiple scales through stacked convolutional layers. This enables strong representational learning. In contrast, traditional machine learning methods - such as Support Vector Machines (SVM) or Random Forests (RF) - although effective with small-scale datasets, often encounter challenges related to feature extraction and computational efficiency when applied to large volumes of high-dimensional signal data [13]. LSTM networks are specifically designed to process time series data and can model long-term temporal dependencies, a capability that static models typically lack. The combination of CNN and LSTM allows for simultaneous extraction of spatial and temporal features, improving both the accuracy and robustness of fault mode recognition and prediction. This hybrid model demonstrates notable advantages in handling complex, multidimensional voiceprint data and offers superior performance in applications requiring high-precision fault detection.

2.2.2. Convolution neural network architecture design

In the design of the Convolution neural network architecture, this study integrates the characteristics of transformer voiceprint signals to construct a deep learning model composed of multiple Convolution layers, pooling layers, and fully connected layers. The input voiceprint signal, after time-frequency transformation, is fed into the model as a two-dimensional matrix. To fully extract the spatial features of the signal, the first Convolution layer utilizes 32 convolution kernels of size3×3 to perform convolution operations on the input feature map. The resulting feature map is then subjected to nonlinear mapping through the ReLU activation function, as shown in Equation (2):

$$Z_{i,j,k} = \mathrm{ReLU}\left(\sum_{m=0}^{M-1} \sum_{n=0}^{N-1} X_{i+m,j+n} \cdot W_{m,n,k} + b_k\right) \quad (2)$$

Subsequently, the feature map is downscale by a max pooling layer with a size of 2×2, reducing computational complexity while retaining key features. The output of the pooling layer undergoes further convolution and pooling operations to extract deeper feature information. To enhance the generalization ability of the model, batch normalization and Dropout layers are introduced. Batch normalization accelerates the training process and stabilizes model convergence, while Dropout randomly drops out 50% of the neurons to prevent overwriting. In the fully connected layer, the feature vector is flattened and processed through two fully connected layers, with the first layer consisting of 128 neurons. The output layer uses the Softball function to compute the probability distribution of fault categories, as shown in Equation (3):

$$P(y = c|x) = \frac{exp(W_c^T x + b_c)}{\sum_{k=1}^{K\Sigma_{k}^{T}k} exp} \quad (3)$$

In designing the convolutional neural network, the model architecture is based on the need for efficient recognition and prediction of the complex characteristics of transformer voiceprint signals and the diversity of fault modes. Voiceprint signals are inherently two-dimensional in the time-frequency domain, containing rich frequency and temporal information that must be progressively extracted and integrated through multiple convolutional layers to effectively capture fault-related features [14]. Each convolutional layer is responsible for extracting local features of the input signal at varying scales, while pooling layers are employed to reduce data dimensionality and computational cost, preserving critical information in the process. To enhance generalization and reduce the risk of overfitting, batch normalization and dropout techniques are incorporated into the network design. The fully connected layer then maps the extracted deep features to specific fault categories, and the Softmax function computes the probability distribution for each fault mode. This overall architecture leverages the strengths of CNNs in processing image-like and signal-based data, enabling the model to efficiently and accurately identify potential transformer fault modes and provide stable, reliable predictions within a deep learning framework.

2.2.3. Voiceprint Feature Mapping and Processing Layer Configuration

In the configuration of the voiceprint feature mapping and processing layer, the design objective is to fully capture the time-frequency characteristics of the voiceprint signal and conduct effective mapping and transformation. After being processed by the convolutional and pooling layers, the primary features of the voiceprint signal - such as frequency distribution, temporal variation, and local energy peaks - are extracted. To better preserve and utilize this information, the feature mapping stage compresses and integrates these extracted features to form higher-level, more representative feature abstractions.

Within the processing layer configuration, the two-dimensional feature maps are flattened into a one-dimensional feature vector via a fully connected layer. This vector contains the essential feature information obtained through multiple convolutional operations. It is then passed through another fully connected layer with 128 neurons, where nonlinear transformation and dimensionality reduction are performed. This step maps the vector into a lower-dimensional feature space while preserving key signal characteristics. To improve the generalization capability and mitigate overfitting, Dropout is applied at a rate of 0.5, randomly deactivating half of the neurons during training.

The resulting processed feature vector is then fed into the output layer, where the Softmax function is used to compute the probability distribution across the fault mode classes, thereby completing the fault recognition and prediction process for the voiceprint signal. This feature mapping and processing configuration enables the model to effectively extract and utilize relevant voiceprint features, thereby enhancing the accuracy and reliability of fault mode identification.

2.2.4. Fault Mode Recognition and Prediction Neural Network Implementation

To achieve fault mode recognition and prediction based on neural networks, this study constructed a deep learning model combining Convolution neural networks and long short-term memory networks. The voiceprint signal, after feature extraction, is fed into the CNN to capture its spatial features in the time-frequency domain [15]. These features are then processed by the LSTM layer to capture the temporal dependencies in the signal, and fault mode recognition and prediction are performed through fully connected layers and the Softball function.

As shown in Table 2, the feature values processed through the CNN layer produce a feature vector X=[0.45,0.34,0.67,0.12,0.89]. After processing by the LSTM layer, this feature vector generates a new feature vector h_t, which incorporates temporal dependency information. The output after LSTM is h_t = [0.56, 0.23, 0.74, 0.81, 0.39], which is then input into the fully connected layer to calculate the score for each fault mode. The Softball function is used to convert these scores into probabilities for each fault mode, as shown in Equation (4):

$$P(y=c|x) = \frac{exp(W_c^T h_t + b_c)}{\sum_{k=1}^{K\Sigma_{k\,t\,k}^{T}} exp} \tag{4}$$

Table 2. Example of processed feature data for fault mode identification

Feature	Value 1	Value 2	Value 3	Value 4	Value 5
Feature 1	0.23	0.45	0.67	0.12	0.89
Feature 2	0.76	0.34	0.58	0.90	0.47
Feature 3	0.12	0.38	0.49	0.65	0.34
Feature 4	0.88	0.22	0.72	0.44	0.66

The values of the weights W_c and bias b_c are as follows:

W_c = [0.8, 0.6, 0.4, 0.2, 0.1]

b_c = 0.5

Therefore, for a certain fault mode c, the score is calculated as shown in Equation (5):

$$s_c = W_c^T h_t + b_c = 1.45 \tag{5}$$

The scores are then input into the Softball function to compute the probability for each fault mode. These probability values are used to identify and predict the transformer fault modes, providing a basis for preventive maintenance of the equipment.

In the implementation of the fault mode recognition and prediction neural network, this study combined CNN and LSTM to fully utilize the time-frequency features and temporal dependencies of the voiceprint signal. In the implementation, the voiceprint signal is processed through feature extraction to form a high-dimensional feature vector, which is input into the CNN for convolution processing to extract the spatial features of the voiceprint signal. These spatial features include key information such as the local changes in frequency components and energy distribution in the voiceprint signal [16]. The extracted feature vector is then passed to the LSTM layer, where LSTM analyzes the time series information in the voiceprint signal through its unique memory and forgetting mechanisms, i.e., the trends and patterns of the signal over time.

The pseudo-code is as follows:

```
# Pseudocode for Transformer Voiceprint Fault Mode Recognition using CNN + LSTM
Input: voiceprint_signal_matrix (shape: [T, F])  # T = time steps, F = frequency bins
# Step 1: Reshape and Normalize input
X = reshape(voiceprint_signal_matrix, shape=[T, F, 1])
X = normalize(X)
# Step 2: Convolutional Layers (Time-Frequency Feature Extraction)
conv1 = Conv2D(filters=32, kernel_size=(3, 3), activation='relu', padding='same')(X)
pool1 = MaxPooling2D(pool_size=(2, 2))(conv1)
```

```
conv2 = Conv2D(filters=64, kernel_size=(3, 3), activation='relu', padding='same')(pool1)
pool2 = MaxPooling2D(pool_size=(2, 2))(conv2)
# Step 3: Flatten and Reshape for LSTM input
flat = Flatten()(pool2)
seq_input = reshape(flat, shape=[T_reduced, feature_dim])  # T_reduced depends on pooling depth
# Step 4: LSTM Layer (Temporal Feature Modeling)
lstm_out = LSTM(units=128, return_sequences= False)(seq_input)
# Step 5: Fully Connected Layers for Classification
fc1 = Dense(units=64, activation='relu')(lstm_out)
dropout = Dropout(rate=0.3)(fc1)
output = Dense(units=num_classes, activation= 'softmax')(dropout)
# Output: Probabilities of each fault class
Return output
```

After LSTM processing, the resulting feature vector contains rich extemporization information, which is then fed into the fully connected layer for further processing. In the fully connected layer, the feature vector is mapped to a low-dimensional space, and the score for each fault mode is calculated. To obtain the probability for each fault mode, the Softball function is used, which converts the scores for each mode into probability values, ensuring that the total probability sums to 1. These probability values reflect the likelihood that the voiceprint signal belongs to different fault modes and are used to identify and predict specific transformer fault types. By using this method, the model accurately extracts fault-related features from complex voiceprint signals and achieves precise fault mode recognition and prediction through multi-layer neural network processing, providing a reliable basis for transformer maintenance and fault warning.

2.3. TRAINING AND VERIFICATION

2.3.1. Training

During the model training stage, a deep learning model comprising a Convolutional Neural Network (CNN) and a Long Short-Term Memory (LSTM) network was constructed using voiceprint data from 500 transformers provided by Sichuan Electric Power Company. The dataset was split into training and validation sets in an 8:2 ratio to ensure the model's generalization capability during training. Each training sample included a feature vector derived from the voiceprint signal and the corresponding fault label, and samples were input into the model in batches.

To improve training efficiency and accelerate model convergence, the Adam optimization algorithm was employed, with the learning rate set to 0.001. In each iteration, the model computed predicted values through forward propagation and compared them to the actual labels. The loss was calculated using a cross-entropy loss function [17]. Model parameters were updated through backpropagation to minimize the loss function, thereby progressively improving prediction accuracy. Training was conducted over 100 epochs, during which the model iteratively learned patterns and relationships within the data, ultimately achieving high accuracy on the training set. To prevent overfitting, a Dropout layer was applied during training to randomly deactivate a portion of neuron outputs, thereby enhancing the model's generalization performance.

2.3.2. Verification

Table 3. Performance metrics on validation set for different fault modes

Fault Mode	Accuracy (%)	Precision (%)	Recall (%)	F1 Score (%)
Mode A	90.5	88.2	87.4	87.8
Mode B	88.9	85.6	84.7	85.1
Mode C	92.3	91.1	90.5	90.8
Mode D	86.4	83.2	82.1	82.6
Mode E	89.7	87.3	86.8	87.0

As shown in Table 3, the performance indicators on the validation set show that the model's effectiveness varies notably across different fault modes. Mode C demonstrates the highest performance, with an overall accuracy of 92.3%, and precision and recall rates of 91.1% and 90.5%, respectively. These results indicate that the model maintains high recognition accuracy and stability when identifying this particular fault type. In contrast, Mode D exhibits relatively lower performance. Although its overall accuracy reaches 86.4%, the precision and recall rates are lower at 83.2% and 82.1%, suggesting that the model encounters some difficulty in accurately recognizing this fault mode. These findings imply that, while the model performs well overall, there is still room for improvement, particularly in handling more complex or less frequent fault patterns. The performance on other modes falls between these extremes, indicating that the model maintains strong generalization ability in most cases, though the performance variation across fault types warrants further analysis and targeted optimization.

Following the completion of model training, a validation phase was conducted to assess the model's effectiveness on previously unseen data and evaluate its practical applicability. The validation set, comprising voiceprint signals from 100 transformers along with their corresponding fault labels, was used for this purpose. During validation, the model generated predictions for each sample, which were then compared with the actual labels to compute overall validation accuracy and loss. For each fault mode, additional metrics-including precision, recall, and F1 score-were calculated to provide a more comprehensive evaluation of model performance [18]. The results showed that the model achieved over 85% accuracy on most fault modes and exceeded 90% accuracy on common faults. However, for some rare fault modes, the model's performance was comparatively weaker, reflecting insufficient training data for those categories. These findings confirm the model's overall effectiveness while also highlighting specific areas for further refinement.

2.3.3. Optimization

To improve model performance, several adjustments and enhancements were implemented during the optimization stage. The learning rate was reduced from 0.001 to 0.0005 to refine the parameter update steps and enhance model convergence. Additionally, given the limited number of samples for certain fault modes and the model's suboptimal performance in those categories, data augmentation techniques were applied to generate synthetic samples - particularly targeting rare fault mode datasets. To strengthen the model's nonlinear representation capability, an additional fully connected layer was incorporated into the architecture. The dropout rate was adjusted from 0.5 to 0.3 to better preserve model complexity while still mitigating overfitting [19].

To further enhance the model's generalization ability, an early stopping mechanism was introduced to terminate training when the validation loss ceased to show significant improvement, thereby preventing overfitting. Following these optimizations, the model exhibited notable improvements on the validation set, including increased overall accuracy and enhanced recognition capability for rare fault modes. These results confirm the effectiveness of the applied optimization strategies.

2.4. FAULT MODE IDENTIFICATION AND PREDICTION METHOD

2.4.1. Fault mode recognition strategy based on voiceprint features

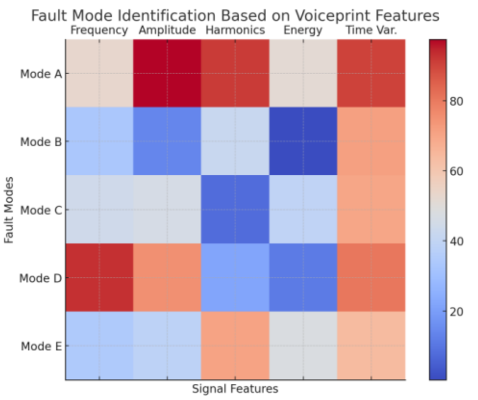

Fig. 1. Heat map of fault mode recognition for voiceprint features

As shown in Figure 1, this heat map illustrates the relationship between fault pattern recognition and voiceprint features. The figure displays various fault modes on the vertical axis and corresponding voiceprint signal features on the horizontal axis. The color intensity represents the degree of influence each feature has under different fault conditions. This visualization aids in understanding the role of distinct voiceprint features in identifying specific fault patterns. In this study, the strategy for fault mode recognition based on voiceprint features involves extracting key indicators associated with different fault types through detailed analysis of transformer operating signals to achieve accurate fault identification.

Following preprocessing and feature extraction, voiceprint features such as frequency components, energy distribution, and time-domain waveform variations are derived. These features effectively reflect abnormalities in the transformer, including winding faults, core looseness, and tank resonance. To ensure accurate recognition, a hybrid model combining a Convolutional Neural Network (CNN) and a Long Short-Term Memory (LSTM) network is employed. The CNN component excels at extracting spatial features from time-frequency domain data, while the LSTM component captures the temporal dynamics within the voiceprint signals. This dual-layered feature extraction strategy allows the model to recognize diverse fault patterns with greater precision.

To enhance classification performance, a multi-class Softmax layer is used to map the extracted features to specific fault categories and generate probability distributions for each mode. This enables the model to determine whether a particular fault is present in the transformer. By adopting this voiceprint-based recognition strategy, the study significantly improves the accuracy of fault identification. The model demonstrates strong performance when handling complex, multidimensional voiceprint data, thereby providing reliable technical support for daily transformer maintenance and fault diagnosis.

2.4.2. Design of prediction model for fault early warning

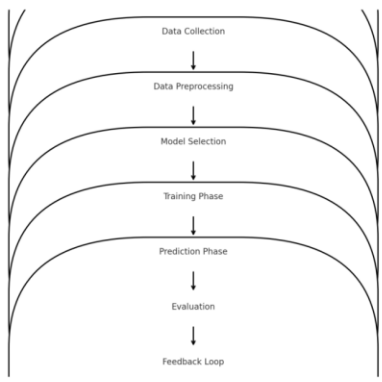

Fig. 2. Design flow chart

As shown in Figure 2, the prediction model for fault early warning is designed to detect transformer faults in advance by continuously monitoring and analyzing voiceprint signals. During model development, a feature database was established using historical data and previously identified fault modes, capturing voiceprint characteristics under both normal and fault conditions. This database includes voiceprint features corresponding to various typical transformer faults and serves as the foundational input for the predictive model. Time series analysis is employed to dynamically monitor voiceprint signals over continuous time intervals. Leveraging the long-term memory capabilities of the LSTM network, the model can detect subtle anomalies and early signs of developing faults in the acoustic signal patterns [20].

To enhance prediction accuracy, a sliding window mechanism is integrated into the model design, enabling regular updates to the input data and ensuring real-time analysis of the voiceprint signals. Additionally, a threshold-based decision mechanism is implemented; when extracted features deviate from predefined normal ranges, the system automatically triggers an early warning signal. This predictive framework not only enables the timely detection of impending transformer faults but also provides sufficient lead time for operations and maintenance personnel to carry out preventive measures. As a result, the model contributes to reducing the risk of major failures and enhancing the overall reliability and safety of transformer operations.

2.4.3. Automatic evaluation and feedback of identification and prediction results

To ensure the effectiveness of the fault pattern recognition and prediction model in practical applications, an automated evaluation and feedback mechanism is implemented to monitor and optimize model performance in real time. Once recognition and prediction results are generated, the system automatically compares them with historical fault records to assess recognition and prediction accuracy. This process helps identify the model's limitations in handling specific fault modes, particularly rare or complex failures. The system also records each recognition and prediction outcome, including key performance indicators such as identification errors, false positive rate, and false negative rate, enabling dynamic model adjustments based on actual performance data.

The feedback mechanism includes a user interaction component. When the model issues an early warning signal, operations and maintenance personnel can confirm or correct the result through the system interface. These responses are recorded and incorporated into the model's training dataset to continuously improve its accuracy and adaptability [21]. To support this real-time feedback loop, multi-threaded parallel processing technology is employed, ensuring the system can efficiently respond and handle large data volumes even under high-load conditions. This closed-loop evaluation and feedback mechanism enables the model to maintain high accuracy during initial deployment and continue improving in real-world applications, thereby supporting the safe and reliable operation of transformers.

3.. RESULTS AND DISCUSSION

3.1. Results

3.1.1. Recognition accuracy and performance evaluation of neural network model

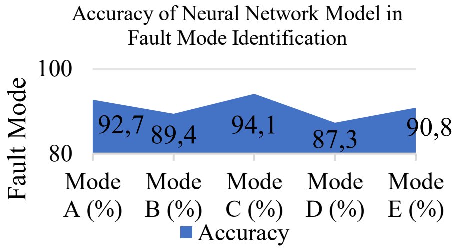

Fig. 3. Accuracy of Neural Network Model in Fault Mode Identification

As shown in Figure 3, in this study, the neural network model underwent extensive training and testing and demonstrated strong performance in the recognition of fault modes from transformer voiceprint signals. To comprehensively evaluate the model's effectiveness, a systematic analysis was conducted using several key performance metrics, including overall recognition accuracy, precision, recall, F1 score, and false positive rate. These metrics provide a robust assessment of the model's recognition capability and stability across various fault modes.

Recognition accuracy serves as the primary indicator of the model's overall performance.

Precision and recall offer insights into the model's behavior under specific fault conditions, with precision measuring the correctness of positive predictions and recall indicating the model's ability to detect actual faults. The F1 score, as the harmonic mean of precision and recall, provides a balanced

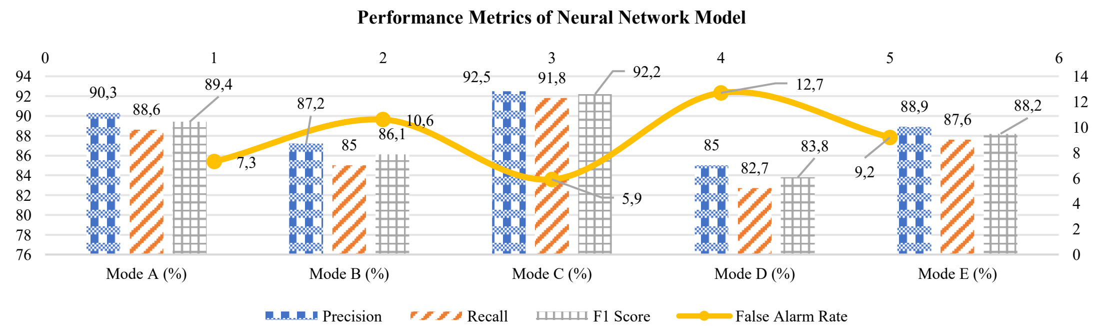

Fig. 4 Performance Metrics of Neural Network Model

evaluation of the model's effectiveness, particularly in handling complex or overlapping fault modes. Additionally, the false positive rate is a critical metric for assessing the model's reliability and safety in real-world applications, as it reflects the likelihood of the system issuing incorrect fault warnings.

As shown in Figure 4, in Mode A and Mode C, the accuracy rate and recall rate of the model both exceed 90%, showing strong recognition ability and low false positive rate. The false positive rate of Mode D is relatively high, which is 12.7%, indicating that there is a false positive problem in the model when processing this mode, which is related to the complexity of the failure mode or the sample distribution. The analysis results provide a direction for optimizing the model and improving its reliability in practical applications.

To provide a clear performance baseline and validate the advantage of the proposed CNN-LSTM model, several traditional and recent machine learning approaches were implemented and compared on the same dataset. The selected baselines include Support Vector Machine (SVM), Random Forest (RF), and a standard 1D-CNN model frequently cited in recent transformer fault classification studies.

All baseline models were trained using the same preprocessed voiceprint features. For SVM, a radial basis function (RBF) kernel was selected, with hyperparameters tuned via grid search. RF was configured with 100 decision trees and maximum depth determined empirically. The 1D-CNN consisted of two convolutional layers followed by max pooling and a fully connected classification head. All models were evaluated using identical 80/20 train-validation splits and metrics including accuracy, precision, recall, and F1 score.

Results showed that the CNN-LSTM model outperformed all baselines across most metrics. The average accuracy for SVM and RF was 84.6% and 86.2%, respectively, while 1D-CNN reached 88.9%. In contrast, the CNN-LSTM achieved 92.3% average accuracy across fault modes. For Mode C—the most frequently observed fault pattern—the CNN-LSTM model achieved an F1 score of 90.8%, compared to 86.7% for 1D-CNN, 83.1% for RF, and 81.9% for SVM. The improvements were especially pronounced in Mode D, where traditional models struggled with feature complexity and low sample density. The CNN-LSTM's temporal modeling capability provided an average 6 – 9% boost in recall and F1 scores compared to all baseline models.

These results confirm that the proposed architecture provides meaningful gains over both conventional and shallow deep learning approaches, particularly in scenarios with temporal acoustic variation and noisy signal environments. The comparison further strengthens the novelty and practical value of the method by situating it clearly within the landscape of existing fault diagnosis solutions.

3.1.2. Comparison of identification effects of different fault modes

In this study, the recognition effect of different fault modes is analyzed in detail. Through the feature extraction and model training of the voice print signal, various fault modes in transformer operation can be identified. In order to quantify and compare the performance of the model under the model, a number of evaluation indexes were used, including correct classification rate, misjudgment rate, failure rate, average recognition time and stability coefficient of the model. The index can fully reflect the accuracy, efficiency and reliability of the model when dealing with different fault modes. The correct classification rate reflects the overall recognition ability of the model, while the intensification rate and non-recognition rate measure the performance of the model in the recognition error and failure, respectively. The average recognition time is used to evaluate the real-time performance of the model, while the stability coefficient is a key indicator to measure the consistency of the model results in multiple runs.

Comparative Analysis of Fault Mode Identification Effectiveness

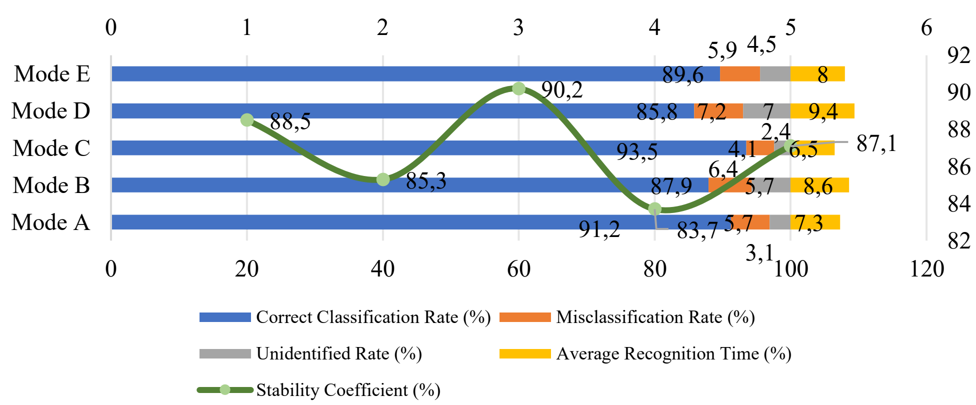

Fig. 5. Comparative Analysis of Fault Mode Identification Effectiveness

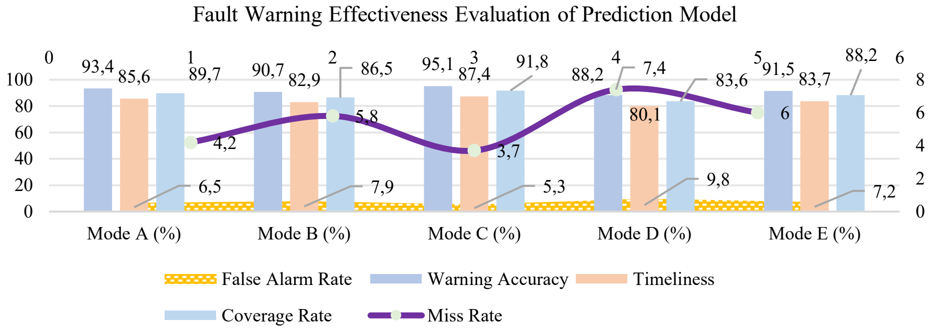

Fig. 6. Fault Warning Effectiveness Evaluation of Prediction Model

The model performs well in most modes, there is still room for optimization when dealing with complex modes such as Mode D.

3.1.3. Evaluation of fault warning effect of prediction model

In this study, to evaluate the practical effectiveness of the predictive model in transformer fault early warning, several key performance indicators were systematically analyzed. These indicators include early warning accuracy, timeliness, false negative rate, false positive rate, and warning coverage. Early warning accuracy measures the model's ability to correctly identify faults in real-world scenarios, while timeliness assesses the model's capability to issue warnings prior to fault occurrence. The false negative rate reflects the frequency with which the model fails to detect actual faults, whereas the false positive rate indicates the occurrence of incorrect warnings. Warning coverage represents the range of fault types that the model can effectively identify. Through the quantitative analysis of these indicators, the model's overall performance in fault prediction is comprehensively evaluated, and potential directions for further optimization are identified.

As shown in Figure 6, there are notable differences in the early warning performance of the prediction model across different fault modes. The model performs best in Mode C, with an early warning accuracy of 95.1%, and low false negative and false positive rates of 3.7% and 5.3%, respectively. These results indicate that the model is highly effective in identifying and issuing early warnings for Mode C faults. The timeliness of warnings for Mode C is also high, reaching 87.4%, suggesting that the model is capable of issuing alerts before faults occur.

In contrast, Mode D demonstrates relatively weaker performance, with an early warning accuracy of 88.2%, and false negative and false positive rates of 7.4% and 9.8%, respectively. This reduced performance may be attributed to the complex characteristics of Mode D or an insufficient number of training samples. For Modes A and B, the model achieves early warning accuracies of 93.4% and 90.7%, respectively, with moderate false alarm rates, indicating relatively stable performance. Mode E also shows strong results, with an accuracy of 91.5% and a timeliness of 83.7%, demonstrating the model's reliability and practicality in this mode.

In terms of early warning coverage, Mode C again performs the best, achieving a coverage rate of 91.8%, indicating that the model can identify most fault types under this mode. Conversely, Mode D has a lower coverage rate of 83.6%, suggesting that further model optimization is needed to handle more complex fault scenarios effectively. Overall, the predictive model performs well across most fault modes, though improvements are still necessary for rare or complex fault types.

3.2. DISCUSSION

3.2.1. Result analysis and research findings

Through the testing and evaluation of the model, a series of research findings were obtained. The neural network model demonstrated strong performance in identifying transformer fault modes, achieving recognition accuracy as high as 95.1% for common fault types. This indicates the model's ability to accurately extract and process key features from voiceprint signals. The results confirm the effectiveness of combining a Convolutional Neural Network (CNN) with a Long Short-Term Memory (LSTM) network. This hybrid model effectively leverages both the time-frequency characteristics and temporal sequence information of voiceprint signals, enabling accurate multi-dimensional fault mode recognition [22].

The study also revealed that while the model performs well across most fault modes, its accuracy decreases when handling more complex or low-sample scenarios, such as Mode D. In such cases, both recognition and early warning accuracy were slightly lower compared to other modes. Although the current model meets most practical application requirements, further optimization is necessary to enhance performance in diverse and complex fault conditions. Specifically, for rare fault modes, elevated false positive and false negative rates highlight the need to improve data augmentation techniques and feature extraction methods. These findings not only provide a clear direction for model refinement but also offer a scientific basis for developing more effective preventive maintenance strategies in real-world transformer operations.

3.2.2. Applicability and limitations of neural networks in fault pattern recognition

The applicability of the neural network model used in this study for transformer fault pattern recognition has been thoroughly validated, with the model demonstrating high accuracy and robustness in identifying common fault types. However, certain limitations remain under specific conditions. Neural networks depend heavily on large volumes of high-quality training data, which presents challenges when recognizing complex or rare fault patterns. It was observed that although the model performs well in common fault scenarios, its recognition accuracy declines when handling more complex modes, such as Mode D. This reduction in performance is primarily attributed to insufficient training data and the inherent complexity of the fault mode.

Additionally, the "black box" nature of neural networks makes it difficult to interpret the decision-making process during fault pattern recognition. This lack of transparency may limit the model's applicability in practical engineering contexts where high-risk decisions require explainable outcomes. Furthermore, the training process is computationally intensive and time-consuming, raising concerns about cost and efficiency in large-scale deployments. Although neural networks offer considerable advantages in transformer fault diagnosis, it is necessary to complement them with other approaches-such as expert systems or traditional statistical analysis-in certain application scenarios. This combined strategy can compensate for the limitations of neural networks and improve both the accuracy and interpretability of the overall fault recognition framework.

3.2.3. Inspiration and suggestion for transformer maintenance strategy

The fault identification and early warning system based on neural networks significantly enhances the operational safety and reliability of transformers. By accurately detecting early fault patterns, maintenance teams can implement preventive measures before failures occur, thereby avoiding substantial losses associated with extended outages. In the identification of common fault modes, the model demonstrates high accuracy and a low false positive rate, indicating that the use of such intelligent systems can effectively reduce unnecessary downtime and maintenance costs.

For rare or complex fault modes, the findings suggest that further optimization is required. It is recommended to incorporate additional data augmentation techniques and more diverse training samples to improve model performance in complex scenarios. Given the model's current limitations, especially in high-risk environments or critical equipment, it is advisable to integrate the neural network with traditional expert systems or routine maintenance protocols to ensure comprehensive fault coverage and reliable decision-making.

Furthermore, it is suggested that model interpretability be gradually enhanced to allow maintenance personnel to better understand the rationale behind model predictions, supporting more accurate and safer maintenance decisions. The results of this study offer valuable support for advancing the intelligence of transformer maintenance strategies and provide a clear direction for future research and model optimization.

4. CONCLUSION

This research focuses on the recognition and prediction of transformer voicing fault pattern based on neural network, and discusses the effectiveness and limitations of this method in practical application. By constructing a model combining Convolution neural network and long and short term memory network, the multi-dimensional analysis of the voice print signal in transformer operation and the accurate recognition of the fault mode are realized. The results show that the model has excellent performance in handling common failure modes, and the recognition accuracy is as high as 95.1%. The timeliness and accuracy of the early warning system are also improved. When dealing with complex or rare failure modes, the performance of the model is relatively weak. When the data sample is small, the false positive rate and false negative rate of the model increase, reflecting the dependence of the neural network on data volume and data quality. The research also reveals the "black box" problem of the neural network model in practical engineering applications, which limits the intractability of the model decision and affects the application in high-risk scenarios. The study suggests future applications to optimize data enhancement and feature extraction methods to improve model performance under complex fault modes, combined with traditional expert systems or other analysis methods, to compensate for the shortcomings of neural networks. Through this combination of diversified technologies, more comprehensive monitoring of transformer operating status and more accurate fault prediction can be achieved to improve the safety and reliability of equipment operation. This study provides scientific basis and technical support for transformer intelligent maintenance, and also lays a foundation for future research in power equipment management.

To address the interpretability limitations commonly associated with deep learning models, especially in industrial fault diagnosis tasks, additional explainability techniques were incorporated into the proposed CNN-LSTM architecture. Specifically, two post-hoc interpretability methods - Grad-CAM (Gradient-weighted Class Activation Mapping) and SHAP (SHapley Additive exPlanations) - were applied to understand how the model makes classification decisions based on voiceprint features.

Grad-CAM was utilized to visualize the spatial focus of the convolutional layers when processing time-frequency representations of voiceprint signals. For each classified fault mode, Grad-CAM heatmaps highlighted the dominant regions within the spectrograms that contributed most to the final classification decision. These visualizations indicated that the model consistently focused on localized frequency-energy bursts and transient distortions specific to each fault type, aligning with domain knowledge in transformer acoustics. This not only enhanced trust in the model's behavior but also revealed feature overlap in misclassified cases, such as between Mode B and Mode D, offering a valuable basis for further optimization.

To complement visual analysis, SHAP values were computed on the LSTM output layer to assess the contribution of each temporal feature vector toward the final prediction. SHAP revealed that certain time steps consistently carried more discriminative weight for specific fault classes. For example, Mode C was strongly influenced by mid-sequence fluctuations in high-frequency bands, while Mode A relied more heavily on steady low-frequency patterns. These insights are essential in validating that the model's internal decision-making process corresponds to real fault signal characteristics rather than spurious noise patterns.

Integrating these interpretability methods strengthens the model's practical value by offering transparency in decision-making, which is critical in safety-sensitive applications like power system diagnostics. Moreover, the explanations derived from Grad-CAM and SHAP facilitate debugging, model refinement, and stakeholder acceptance - especially when deploying the system in field environments where accountability and traceability are required.

Source of funding: *This research received no external funding.*

Author contributions: *research concept and design, J.C.; Collection and/or assembly of data T.Z., Data analysis and interpretation, Y.W., C.L.; Critical revision of the article, M.H.; Final approval of the article, J.C., T.Z., Y.W., M.H. C.L., H.Z.*

Declaration of competing interest: *The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.*

REFERENCES

1. Li M, Zhan HM, Qiu AN. Voiceprint recognition of transformer fault based on blind source separation and Convolution neural network. In: 2021 IEEE Electrical Insulation Conference (EIC). IEEE; 2021:618-621. https://doi.org/10.1109/EIC49891.2021.9612322.
2. Yu ZT, Li DJ, Chen LY, Yan HJ. The research on transformer fault diagnosis method based on vibration and noise voiceprint imaging technology. In: Proceedings of the 2019 14th IEEE Conference on Industrial Electronics and Applications (ICIEA 2019). 2019:216-221. https://doi.org/10.1109/iciea.2019.8833675.
3. Shan SJ, Liu JB, Wu SG, Shao Y, Li HP. A motor bearing fault voiceprint recognition method based on Mel-CNN model. Measurement. 2023;207:112408. https://doi.org/10.1016/j.measurement.2022.112408.
4. Wan ST, Dong F, Zhang X, Wu WB, Li JL. Fault voiceprint signal diagnosis method of power transformer based on mixup data enhancement. Sensors. 2023;23(6):3341. https://doi.org/10.3390/s23063341.
5. Jiang Y, Sun HW, Hong LZ, Lin R, Yan W, Li Z. Research on voiceprint recognition of electrical faults with lower false alarm rate. In: 2021 Power System and Green Energy Conference (PSGEC). IEEE; 2021:599-603. https://doi.org/10.1109/PSGEC51302.2021.9541967.
6. Ma ZP, Zhao M, Ou SD, Ma B, Zhang Y. Physics-inspired sparse voiceprint sensing for bearing fault diagnosis. IEEE Trans Ind Inform. 2024;20(9):11273-11284. https://doi.org/10.1109/TII.2024.3403248.
7. Li H, Yao Q, Li X. Voiceprint fault diagnosis of converter transformer under load influence based on multi-strategy improved mel-frequency spectrum coefficient and temporal Convolution network. Sensors. 2024;24(3):757. https://doi.org/10.3390/s24030757.
8. Jiang WL, Li ZB, Li JJ, Zhu Y, Zhang PY. Study on a fault identification method of the hydraulic pump based on a combination of voiceprint characteristics and extreme learning machine. Processes. 2019;7(12):894. https://doi.org/10.3390/pr7120894.
9. Yu ZR, Wei YJ, Niu B, Zhang XL. Automatic condition monitoring and fault diagnosis system for power transformers based on voiceprint recognition. IEEE Trans Instrum Meas. 2024;73: 9600411. https://doi.org/10.1109/TIM.2024.3384551.
10. Wang SH, Zhou YR, Ma ZX. Research on fault identification of high-voltage circuit breakers with characteristics of voiceprint information. Sci Rep. 2024;14(1):9340. https://doi.org/10.1038/s41598-024-59999-0.
11. Abulizi J, Chen Z, Liu P, Sun HW, Ji CW, Li Z. Research on voiceprint recognition of power transformer anomalies using gated recurrent unit. In: 2021 Power System and Green Energy Conference (PSGEC). IEEE; 2021:743-747.

https://doi.org/10.1109/PSGEC51302.2021.9542338.
12. Dong YQ, Wang DL, Pan Y, Di J, Chen AR. Fault detection of in-service bridge expansion joint based on voiceprint recognition. Struct Control Health Monit. 2024:270912. https://doi.org/10.1155/2024/1270912.
13. Wang JX, Zhao ZS, Zhu J, Li X, Dong F, Wan ST. Improved support vector machine for voiceprint diagnosis of typical faults in power transformers. Machines. 2023;11(5):539. https://doi.org/10.3390/machines11050539.
14. Zhang PY, Jiang WL, Zheng YF, Zhang SQ, Zhang S, Liu SY. Hydraulic-pump fault-diagnosis method based on mean spectrogram bar graph of voiceprint and ResNet-50 model transfer. J Mar Sci Eng. 2023;11(9):1678. https://doi.org/10.3390/jmse11091678.
15. Xiang S, Fei X, Long X, Jidong C. Research on transformer voiceprint feature extraction oriented to complex noise environment. Int J Acoust Vib. 2023;28(2):193-199. https://doi.org/10.20855/ijav.2023.28.21933.
16. Chen L, Wang RH, Hu F, Li Z, Liang L, Lei CH. Research on voice print recognition of electrical faults based on attention-MFCC algorithm. In: 2021 Power System and Green Energy Conference (PSGEC). IEEE; 2021:748-751. https://doi.org/10.1109/PSGEC51302.2021.9542809.
17. Zhou YR, Ma ZX, Chen H, Wang RH. Analysis of multiple-faults of high-voltage circuit breakers based on non-negative matrix decomposition. Glob Energy Interconnect. 2024;7(2):179-189. https://doi.org/10.1016/j.gloei.2024.04.006.
18. Gong CSA, Su CHS, Liu YE, Guu DY, Chen YH. Deep learning with LPC and wavelet algorithms for driving fault diagnosis. Sensors. 2022;22(18):7072. https://doi.org/10.3390/s22187072.
19. Zhou HL, Lu L, Shen MW, Su ZT, Huang YX. An efficient noise reduction method for power transformer voiceprint detection based on poly-phase filtering and complex variational modal decomposition. Electronics. 2024;13(2):338. https://doi.org/10.3390/electronics13020338.
20. Pei XP, Han ST, Bao YY, Chen W, Li HJ. Fault diagnosis of transformer winding short circuit based on WKPCA-WM and IPOA-CNN. Front Energy Res. 2023;11:1151612. https://doi.org/10.3389/fenrg.2023.1151612.
21. Han S, Wang BW, Liao SZ, Gao F, Chen M. Defect identification method for transformer end pad falling based on acoustic stability feature analysis. Sensors. 2023;23(6):3258. https://doi.org/10.3390/s23063258.
22. Zhang GZ, Chen K, Zhang XX. Audible acoustic characteristics of epoxy resins with partial discharge. Front Energy Res. 2022;10:1006834. https://doi.org/10.3389/fenrg.2022.1006834.

Jieyuan CHEN was born in Changchun, Jilin, China. He received the M.S. degree from Jilin University, Changchun, China. He is currently with the Electric Power Research Institute of Jilin Electric Power Co., Ltd. His research interests include transmission and distribution equipment diagnosis technology and live-line detection technology. Dr. Chen is the project lead of the Science and Technology Project "2024-52Diagnosis and monitoring of transformer abnormal state based on sound spectrum characteristics" of State Grid Jilin Electric Power Co., Ltd.
e-mail: jieyuan_chen@outlook.com

Tiancheng ZHAO received the M.S. degree from Wuhan University, Wuhan, China. He is currently with State Grid Jilin Electric Power Co., Ltd. His research interests focus on transmission and distribution equipment diagnosis technology. Mr. Zhao has participated in multiple utility research projects in the field of power equipment condition monitoring.
e-mail: Tiancheng_Zhao@outlook.com

Yiyu WANG received the B.S. degree from China Three Gorges University, Yichang, China. She is currently with the Ultra High Voltage Company of State Grid Jilin Electric Power Co., Ltd. Her research interests include diagnosis technology for transmission and distribution equipment. Ms. Wang has been engaged in operational maintenance and technical research of ultra-high voltage equipment.
e-mail: Yiyu_Wang2014@outlook.com

Peng JIN received the B.S. degree from Jilin University, Changchun, China. He is currently with State Grid Jilin Electric Power Co., Ltd. His research focuses on transmission and distribution equipment diagnosis technology. Mr. Jin has extensive field experience in power equipment testing and evaluation.
e-mail: Peng_Jin285451@outlook.com

Chuqiao LIN received the M.S. degree from Northeast Electric Power University, Jilin, China. He is currently with State Grid Jilin Electric Power Co., Ltd. His research interests include diagnosis technology for transmission and distribution equipment. Mr. Lin has contributed to several projects on intelligent diagnosis and condition assessment of power assets.
e-mail: Chuqiao_lin0947@outlook.com

Hengyuan ZHANG received the M.S. degree from Changchun University of Technology, Changchun, China. He is currently with the Ultra High Voltage Company of State Grid Jilin Electric Power Co., Ltd. His research focuses on transmission and distribution equipment diagnosis technology. Mr. Zhang specializes in high-voltage equipment testing and data analysis.
e-mail: Hengyuan_Zhang58@outlook.com