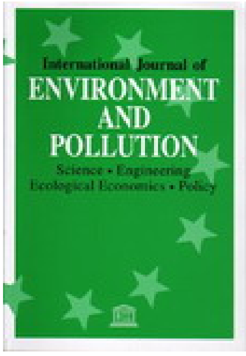
INDERSCIENCE PUBLISHERS
Linking academia, business and industry through research



International Journal of Environment and Pollution

ISSN online: 1741-5101 - ISSN print: 0957-4352
https://www.inderscience.com/ijep

Research on distribution network fault identification based on active transfer learning and autoencoder

Youzhuo Zheng, Jiang Lin, Kun Zhou, Hengrong Zhang, Kailei Chen

DOI: 10.1504/IJEP.2026.10076030

Article History:

Received:	03 April 2025
Last revised:	26 June 2025
Accepted:	24 October 2025
Published online:	18 February 2026

Copyright © 2026 Inderscience Enterprises Ltd.

Research on distribution network fault identification based on active transfer learning and autoencoder

Youzhuo Zheng

Electric Power Research Institute of Guizhou Power Grid Co., Ltd.,
Guiyang, 550002, China
Email: danis032@163.com

Jiang Lin*

Nanjing DeRuan Information Technology Development Co., Ltd.,
Nanjing, 210000, China
Email: jiffer826@126.com
*Corresponding author

Kun Zhou, Hengrong Zhang and Kailei Chen

Electric Power Research Institute of Guizhou Power Grid Co., Ltd.,
Guiyang, 550002, China
Email: kunz02@163.com
Email: zhanghr91@126.com
Email: Ckail021@163.com

Abstract: The stable operation of distribution networks is crucial for economic stability and daily life, yet frequent faults arise due to long transmission lines and wide coverage. Conventional fault diagnosis based solely on voltage and current data is limited by load fluctuations and external interference. To address this, this paper introduces time-frequency entropy (TFE) features into an autoencoder model to characterise signal dynamics and energy distribution for improved fault identification. Furthermore, an active transfer learning strategy combined with a self-attention-based autoencoder is developed to enhance cross-domain adaptability between source and target networks. Experiments on the IEEE 33-node distribution network show that the proposed method achieves 99.49% accuracy while significantly reducing training time, demonstrating its effectiveness and practical value for cross-scenario fault diagnosis in distribution networks.

Keywords: distribution network; fault identification; active transfer learning; autoencoder; TTE; time-frequency entropy; transformer; cross-domain fault diagnosis.

Reference to this paper should be made as follows: Zheng, Y., Lin, J., Zhou, K., Zhang, H. and Chen, K. (2026) 'Research on distribution network fault identification based on active transfer learning and autoencoder', *Int. J. Environment and Pollution*, Vol. 76, Nos. 1/2, pp.140–157.

Biographical notes: Youzhuo Zheng holds a Master of Engineering degree and is a Senior Engineer. He serves as a Specialist at the Electric Power Research Institute of Guizhou Power Grid Co., Ltd. and a Level-2 Leading Professional

Copyright © The Author(s) 2026. Published by Inderscience Publishers Ltd. This is an Open Access Article distributed under the CC BY license. (http://creativecommons.org/licenses/by/4.0/)

and Technical Expert of Guizhou Power Grid. His main research directions include smart distribution networks and new-type power systems.

Jiang Lin holds a Master's degree and is a Senior Electrical Engineer. Employed at Nanjing Deruan Information Technology Development Co., Ltd., his research focuses on new-type power systems and condition monitoring and diagnosis of electrical equipment. With expertise in smart grid optimisation and advanced equipment diagnostics, he leads technical innovation projects in power system reliability and renewable energy integration.

Kun Zhou holds a Master of Engineering degree and is a Senior Engineer. He serves as the Leader of the Distribution Network Automation Section at the Dispatch Center of Bijie Power Supply Bureau, Guizhou Power Grid Co., Ltd., and is a Level-1 Top-notch Professional and Technical Skill Expert of Guizhou Power Grid. His main research directions include distribution network dispatch automation, smart distribution grids, and new-type power systems.

Hengrong Zhang holds a Master of Engineering Degree and is a Mid-Level Engineer. Serving as an Assistant Specialist at the Electric Power Research Institute (EPRI) of Guizhou Power Grid Co., Ltd., he is also a Level-3 Top-notch Professional and Technical Expert of Guizhou Power Grid. His research focuses on Smart Distribution Grids (aligned with IEEE 2030.5 standards) and New-Type Power Systems (as defined in China's national energy policies), with expertise in distributed energy resource integration and grid-edge intelligence.

Kailei Chen holds a Bachelor of Engineering degree and is an Engineer. He serves as a Specialist at the Operation Branch of Guizhou Power Grid Co., Ltd. and a Level-3 Top-notch Professional and Technical Expert of Guizhou Power Grid. His main research directions include smart distribution networks and new-type power systems.

1 Introduction

The stable and reliable operation of distribution networks is closely connected to economic stability and daily life. Because of long transmission lines and wide coverage, faults in distribution lines happen often (Stefanidou-Voziki et al., 2022). As new power systems keep developing, fault recorders are now widely used to monitor and log the status of distribution networks, helping identify faults. Common distribution network faults include single-phase-to-ground faults, phase-to-ground short circuits, and phase-to-phase short circuits. The voltage and current waveforms change for different fault types; therefore, accurately identifying fault types based on waveform details recorded by fault recorders is key to lowering fault impacts (Khavari et al., 2020).

The core of fault type identification involves extracting fault features. How to effectively capture the key information in voltage and current waveforms has become a research focus, attracting widespread attention from researchers. Traditional fault identification methods determine fault types by manually analysing waveform characteristics of different faults based on expert experience (Dai et al., 2024). However, manual identification makes it challenging to ensure real-time performance and objectivity. With the emergence of machine learning technology, extensive research has been conducted on extracting critical fault feature information.

In this study, we propose a distribution network fault identification method that combines active transfer learning with a self-attention-based autoencoder. Recognising that traditional approaches relying solely on voltage and current signals are susceptible to disturbances from load fluctuations and external interference, we incorporate time-frequency entropy (TTE) to better capture signal dynamics and complexity. To address domain discrepancies between labelled source data and unlabelled target data, we employ a maximum mean discrepancy (MMD)-based loss function, enabling efficient knowledge transfer without the need for extensive labelled datasets.

The key contributions of this study are summarised as follows:

1 An autoencoder model integrating TTE features and self-attention mechanism is developed to enhance the extraction of complex fault characteristics from distribution network signals.

2 An active transfer learning framework incorporating a MMD-based adaptation layer is proposed, enabling effective cross-domain knowledge transfer and significantly reducing the demand for labelled data in target domains.

3 Extensive experiments on the IEEE 33-node system demonstrate that the proposed method achieves superior fault identification accuracy and computational efficiency across multiple fault scenarios, validating its practicality and generalisation ability in distribution network fault diagnosis.

The remainder of this paper is organised as follows. Section 2 reviews related work on fault diagnosis methods. Section 3 introduces the Autoencoder Model with TTE. Section 4 presents the Distribution Network Fault Identification Based on Active Transfer Learning. Section 5 provides simulation setup and method validation. Finally, Section 6 concludes the study.

2 Related work

Traditional fault identification methods determine fault types by analysing waveform features based on signal processing techniques. Methods such as the Fourier transform, wavelet transform, and S-transform have been widely applied. Guo et al. (2020) converted fault recording data into matrix modulus through the S-transform to extract features. Khavari et al. (2020) applied wavelet transform combined with overcurrent relays to detect fault characteristics. Hassan et al. (2024) and Cano et al. (2024) further improved the extraction process by integrating wavelet decomposition and neural networks. Touati et al. (2023) introduced a hybrid Fast Fourier Transform and Fuzzy Logic approach to improve reliability in distinguishing faults from disturbances. Guillen et al. (2021) utilised Taylor-Fourier transform-based filtering for fault diagnosis. However, these methods often rely heavily on prior knowledge and are susceptible to human subjectivity, limiting their generalisation in complex scenarios.

With the development of machine learning, neural networks have become widely applied to fault identification. Saad et al. (2022), Alhanaf et al. (2023), and Mo et al. (2023) applied convolutional neural networks (CNNs) for automatic feature extraction.

Zou et al. (2022) proposed a dual-branch CNN framework. Liu et al. (2022) utilised dual-channel CNN models on realistic waveform measurements. Dai et al. (2021), Veerasamy et al. (2021), and Moradzadeh et al. (2021) employed recurrent neural networks (RNN) and long short-term memory (LSTM) networks to extract sequential temporal dependencies. While deep learning reduces manual feature engineering, CNNs often struggle with capturing long-term dependencies, and RNN-based models may encounter vanishing gradient problems in long sequences.

To better handle sequential dependencies, transformer-based models incorporating self-attention mechanisms have recently emerged. Huang et al. (2023a) developed transformer-based approaches for fault classification and energy forecasting. Although transformer models demonstrate superior performance in capturing temporal correlations, they require large labelled datasets for effective training, which are often difficult to obtain in practical power systems.

Transfer learning has become a promising solution to overcome limited labelled data. Li et al. (2023) developed transfer learning models for load forecasting; Zhong et al. (2024) applied transformer-based transfer learning for fault detection; Zhang et al. (2023) proposed unsupervised transfer learning integrating dynamic graph attention networks; Yu et al. (2023) employed transfer learning in resonant grounding networks. However, many transfer learning approaches still face challenges in domain discrepancy adaptation, leading to potential negative transfer effects.

3 Autoencoder model with time-frequency entropy

3.1 Time-frequency entropy

During the operation of the distribution network, although the fault recorder can collect voltage and current time-series data in real-time, the voltage and current signals are easily affected by load fluctuations, external interference, and other factors. Relying only on voltage and current signals for fault identification still has certain limitations, as these signals are susceptible to disturbances. For this reason, when constructing the autoencoder model, this paper additionally introduces TFE features in the input layer (Sucic et al., 2014). TFE can effectively reflect the complexity and dynamic changes of the signal, quantify the energy distribution across different frequency bands, and provide a favourable basis for fault identification. The calculation method of TFE is shown as follows:

1 *Time–frequency domain signal decomposition*

In this paper, the signal is processed using the short-time Fourier transform (STFT), which decomposes it into both time and frequency domains. This conversion results in a time–frequency distribution matrix, where each element represents the amplitude or energy at a specific frequency for a given time.

$$X(t,f)=\int e^{-j2\pi f\tau}x(\tau)w(\tau-t)\mathrm{d}\tau \tag{1}$$

$$w(n)=\frac{1}{2}\left(1-\cos\left(\frac{2\pi n}{N-1}\right)\right),\quad n=0,1,\ldots,N-1 \tag{2}$$

where $X(t,f)$ represents the signal components at time t and frequency f .The original recorded waveform signal is denoted as $x(\tau)$ .The window function $w(\tau - t)$ controls time-frequency resolution and minimises spectral leakage. The window length N determines the balance between time and frequency resolution. The short-time shifting step n defines the overlap between adjacent windows.

2 *Calculation of power distribution*

Based on the time-frequency distribution matrix $X(t,f)$ obtained from STFT, the energy density distribution within the signal is further calculated, quantifying the signal's energy distribution of the signal across different times and frequencies.

$$P(t,f) = \frac{|X(t,f)|^2}{\sum_{t,f} |X(t,f)|^2} \tag{3}$$

where $P(t,f)$ represents the probability distribution on the time-frequency plane.

3 *Calculation of TFE*

TFE is used to quantify the complexity of energy distribution in the time-frequency domain. TFE quantifies the complexity and randomness of how energy is distributed across the time–frequency plane. A more uniform distribution results in higher entropy, while a more concentrated distribution indicates lower entropy. It is computed as follows:

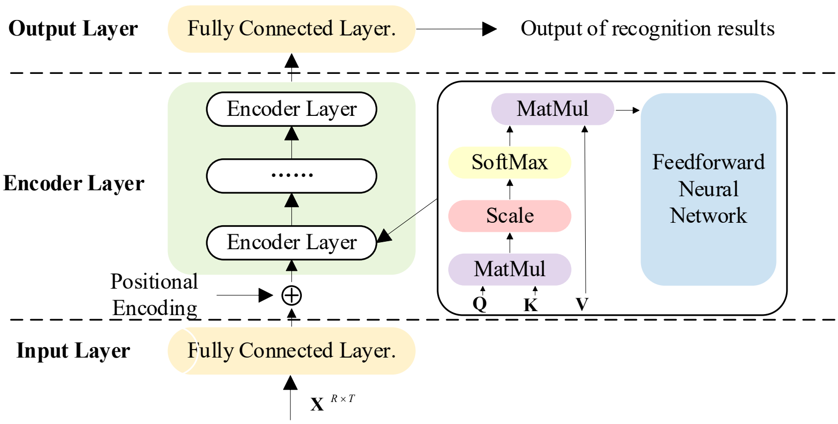
$$H = -\sum_{t,f} P(t,f) \log P(t,f) \tag{4}$$

where H represents the TFE value.

3.2 Autoencoder model structure

The model based on the Transformer demonstrates significant advantages in time-series processing (Huang et al., 2023b). The traditional self-attention mechanism aggregates information point by point, making it difficult to capture reliable temporal dependencies in complex fault signals. To address this challenge, this study constructs an autoencoder model that integrates the self-attention mechanism into the Transformer encoder, achieving more effective information aggregation. Specifically, the self-correlation mechanism discovers periodic dependencies by computing the autocorrelation of the sequence and merges similar subsequences through time-delay aggregation to enhance information utilisation. The architecture of the model constructed in this study is illustrated in Figure 1. $\mathbf{X}^{R\times T}$ is the input sequence. R is the input sequence feature number. T is the number of time steps. $\mathbf{W}_Q$ 、 $\mathbf{W}_K$ 、 $\mathbf{W}_V$ are learnable parameter matrices.

Figure 1 Diagram of autoencoder model (see online version for colours)



The self-correlation mechanism discovers periodic dependencies by computing the autocorrelation of the sequence and merges similar subsequences through time-delay aggregation to enhance information utilisation. The calculation method of the autocorrelation coefficient is shown in equation (5).

$$R_{xx}(\tau) = \lim_{L \to \infty} \frac{1}{L} \sum_{t=1}^{L} x_t x_{t-\tau} \tag{5}$$

where x_t represents a feature sequence. $R_{xx}(\tau)$ denotes the similarity between x_t and its τ delayed counterpart $x_{t-\tau}$. L represents the duration of the time-series sequence. The autocorrelation coefficient $R_{xx}(\tau)$ is used as an unnormalised confidence measure for estimating the period length.

For time-series data, after projection, the query vector $\mathbf{Q}$, key vector K, and value vector $\mathbf{V}$ are obtained. The calculation formulas for the self-correlation mechanism are shown in equations (6)–(9).

$$\tau_1, \tau_2, ..., \tau_k = \arg \mathrm{Topk}\left(R_{Q,K}(\tau)\right) \tag{6}$$

$$\hat{R}_{Q,K}(\tau_1), \hat{R}_{Q,K}(\tau_2), \ldots, \hat{R}_{Q,K}(\tau_k) = \mathrm{SoftMax}\left(R_{Q,K}(\tau_1), R_{Q,K}(\tau_2), \cdots, R_{Q,K}(\tau_i)\right) \tag{7}$$

$$S_{\mathrm{Auto-Corr}}(\mathbf{Q}, K, \mathbf{V}) = \sum_{i=1}^{k} \mathrm{Roll}(V, \tau_i)\, \hat{R}_{Q,K}(\tau_i) \tag{8}$$

$$k = \lfloor c \log L \rfloor \tag{9}$$

where $\arg \mathrm{Topk}(\ )$ represents the k most probable values of τ, denoted as $\tau_1, \tau_2, \ldots, \tau_k$. $\lfloor \cdot \rfloor$ is a rounding function, c being a hyperparameter. $R_{Q,K}(\tau_i)$ denotes the self-correlation relationship between $\mathbf{Q}$ and $\mathbf{K}$ at different time lags τ_i. $\hat{R}_{Q,K}(\tau_i)$ represents the self-correlation relationship after SoftMax normalisation of different $R_{Q,K}(\tau_i)$. $S_{\mathrm{Auto-Corr}}(\mathbf{Q}, \mathbf{K}, \mathbf{V})$ represents the computed self-correlation value, which measures the relationship between $\mathbf{Q}$, $\mathbf{K}$ and $\mathbf{V}$, $\mathrm{Roll}(V, \tau_i)$ denotes the lag operation on $\mathbf{V}$ by τ_i, where the element removed from the first position is reintroduced at the last position.

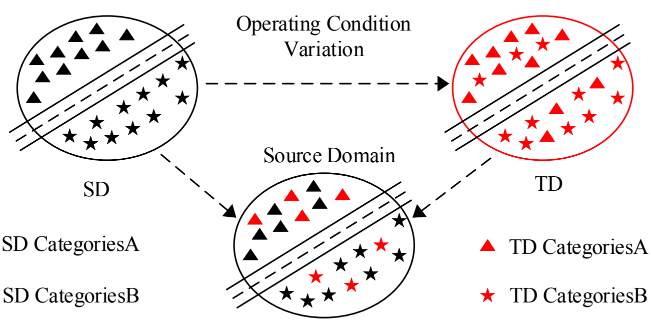
The extracted fault features obtained from the autoencoder model introduced in this section serve as inputs for the active transfer learning approach detailed in Section 3, thereby facilitating efficient knowledge transfer across different operational scenarios.

4 Distribution network fault identification based on active transfer learning

4.1 *Overview of cross-domain fault diagnosis*

The challenge of cross-domain fault diagnosis lies in training model parameters using annotated samples from the source domain (SD), while data in the target domain (TD) remains unlabelled, achieved by extracting common features across domains. The SD data and TD data can be operational data from different nodes within the same distribution network, where fault characteristics exhibit both commonality and uniqueness. By learning the shared features between SD and TD, fault types can be accurately identified in the target domain, achieving the goal of unsupervised fault diagnosis for TD data, as shown in Figure 2.

Figure 2 Cross-domain fault diagnosis principle (see online version for colours)



4.2 *Maximum mean discrepancy (MMD)*

To measure the difference between SD and TD, this paper employs MMD to quantify the distance between them. MMD is a non-parametric measure that maps data into a reproducing kernel Hilbert space to evaluate the discrepancy between datasets. The calculation method is shown in equation (10).

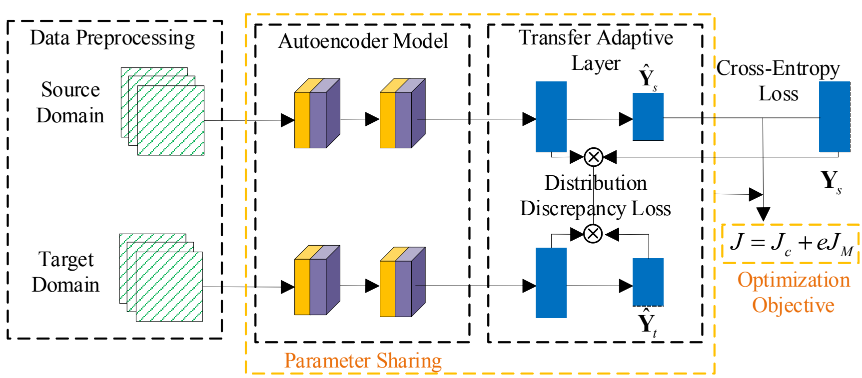
$$\mathrm{MMD}_h^2\left(D_S^k, D_T^l\right) = \left\| \frac{1}{N_S^k} \sum_{i=1}^{N_S^k} \phi\left(x_i^{D_S^k}\right) - \frac{1}{N_T^l} \sum_{j=1}^{N_T^l} \phi\left(x_j^{D_T^l}\right) \right\|_h^2 \tag{10}$$

where D_S^k represents the k th SD, D_T^l represents the l th TD, $x_i^{D_S^k}$ denotes the i th sample in the k th SD, $x_j^{D_T^l}$ denotes the j th sample in the l th TD, N_S^k is the total number of samples in the k th SD, N_T^l is the total number of samples in the l th TD, $\phi(\ )$ is the mapping function that projects data into the Hilbert space. h represents the reproducing kernel Hilbert space, $\|\ \|_h$ denotes the norm in the reproducing kernel Hilbert space.

4.3 *Transfer model structure*

In the distribution network, different nodes operate under different conditions, resulting in fault characteristics with varying distributions, while the TD data samples remain unlabelled. To achieve this, this paper designs an appropriate network structure and loss function to enable the effective transfer of knowledge between the SD and TD. The fault diagnosis model constructed in this study is shown in Figure 3, where the shared common features across the SD and TD are mainly extracted by the autoencoder, while the domain-specific feature differences are primarily reflected in the transfer adaptation layer. Here, $\hat{\mathbf{Y}}_s$ represents the predicted labels for the SD, $\hat{\mathbf{Y}}_t$ represents the predicted labels for the TD, $\mathbf{Y}_s$ denotes the true labels for the SD.

Figure 3 Active transfer learning-based fault identification model for distribution networks (see online version for colours)



At the data preprocessing level, the voltage and current data collected by the fault recorder are normalised, and the TFE is calculated to construct the input set. Then, an autoencoder architecture is employed to derive fault representations for different fault types. Unlike the previous approach, a pair of dense layers is appended following the autoencoder as a transfer adaptation layer. The activation function outputs the predicted fault category labels, and cross-domain loss, along with cross-entropy loss, is calculated using the true labels $\mathbf{Y}_s$ to facilitate transfer feature learning. This approach eliminates the need for retraining the model and improves computational efficiency.

4.4 Design of loss functions

To address the discrepancy in data distributions across the SD and TD in cross-domain fault diagnosis, the model in this paper employs two key loss functions for optimisation: multi-class cross-entropy loss and MMD loss.

Minimising the cross-entropy loss over the SD dataset guarantees the fault classification accuracy of the fault diagnosis model on the SD data. The cross-entropy calculation method is shown in equation (11).

$$J_{\mathrm{C}} = -\frac{1}{a}\sum_{b=1}^{a}\sum_{g=1}^{d} u_{bg} \lg P_{bg} \tag{11}$$

where a represents the total number of samples, d represents the total number of classification categories, u_{bg} denotes the encoding of the b th category in the g th sample, P_{bg} represents the predicted probability of the b th category in the g th sample.

The MMD loss function is used to quantify and minimise the domain shift between the SD and the TD, thereby enhancing the model's adaptability to new domains. The overall objective function of the model is given in equation (12).

$$J = J_c + eJ_M \tag{12}$$

where e represents the balancing hyperparameter, J_M denotes the MMD loss function.

To address the critical issue of determining when to stop the training process during cross-domain adaptation, we adopt a monitoring strategy based on the stability of the combined loss function. Specifically, the training process is terminated when the total loss (a weighted sum of the cross-entropy loss and MMD loss) shows no significant improvement over a preset number of consecutive validation epochs (set as 10 in our experiments). This early stopping strategy helps to prevent overfitting to noisy pseudo-labels in the target domain and avoids potential negative transfer effects. Furthermore, we observe that as domain alignment progresses, the MMD loss gradually stabilises, providing an additional indicator of sufficient domain adaptation. This approach ensures that the model achieves a good trade-off between adequate feature alignment and generalisation performance across domains.

5 Simulation and method validation

5.1 Experimental setup and dataset

To evaluate the performance and reliability of the proposed method, this paper selects an IEEE 33-bus benchmark system from several standard feeder networks published by the IEEE Power and Energy Society (Baran and Wu, 1989) and employs PSCAD/EMTDC for simulation, enabling comparative analysis with existing research. Figure 4 illustrates the structural layout of the distribution network. In this network, four fault points are set, with only one type of fault occurring at each fault point per simulation. Each fault simulation varies fault conditions, such as type, location, grounding resistance, and initial phase angle of the fault. The specific fault simulation parameters are listed in Table 1. For each fault simulation, 0.1 s of fault signals are collected with a sampling step of

100 µs. Fault simulations are conducted under 10 different operating conditions, and each node obtains a dataset of 4900 three-phase voltage and current fault samples. To emulate the practical conditions where voltage and current measurements are often affected by sensor noise and external interferences, additive Gaussian noise was injected into the simulated voltage and current waveforms. The noise level was set to achieve a signal-to-noise ratio (SNR) of 30 dB, simulating typical measurement disturbances in real-world distribution networks. The resulting noise-augmented dataset was then used for model training and evaluation to verify the robustness of the proposed method under noisy conditions.

Figure 4 IEEE 33-node system diagram (see online version for colours)

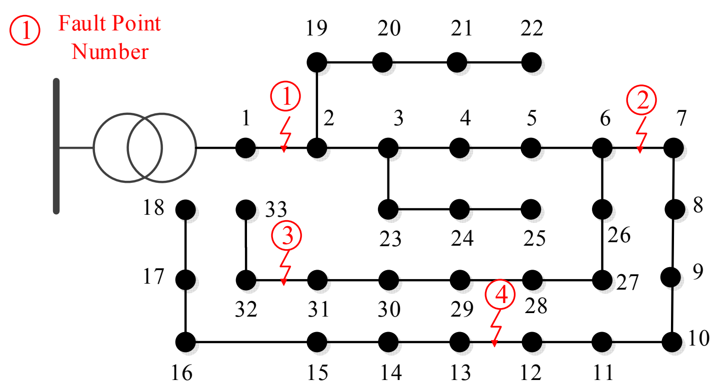


Table 1 Different fault parameter table

Fault nodes	*1, 2, 3, 4*
Fault resistance/Ω	1,10,20,50,100,200,500
Fault initial phase angle/°	0,10,20,50,100,200,300
Fault types	AG, BG, CG, ABG, ACG, BCG, AB, AC, BC, ABCG

The training data uses a batch size of 64, the sampling rate is configured at 16 kHz, and the Adam optimisation algorithm is employed over 300 epochs to iteratively refine the model parameters. The learning rate and other hyperparameters are set using a step-down approach, with an initial learning rate of 0.001. Since the features extracted by the transfer learning model are unstable during the initial training phase, the entire model is pretrained on the SD data for 50 epochs prior to initiating the transfer learning process. This enables the cross-domain model to acquire a certain diagnostic capability for distribution network fault characteristics, thereby improving prediction accuracy.

5.2 Fault identification results

The confusion matrix is effective for evaluating the performance of a classification model, making it convenient to compare predicted labels with the actual outcomes. The basic metrics of the confusion matrix are presented in Table 2. This paper uses the confusion matrix to present the distribution network fault identification results. The rows represent the actual types of faults occurring in the distribution network, while the columns indicate the fault classification results predicted by the proposed model.

Table 2 Basic metrics table of the confusion matrix

	Positive	*Negative*
Actual positive	TP	FN
Actual negative	FP	TN

The metrics derived from the confusion matrix include accuracy, precision, recall, and F1 score. Based on these derived metrics, the fault identification results can be further evaluated. The calculation methods for these derived metrics are given in equations (13)–(16).

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

$$\text{Precision } = \frac{TP}{TP+FP} \tag{14}$$

$$\text{Recall } = \frac{TP}{TP+FN} \tag{15}$$

$$\text{Specificity } = \frac{TN}{TN+FP} \tag{16}$$

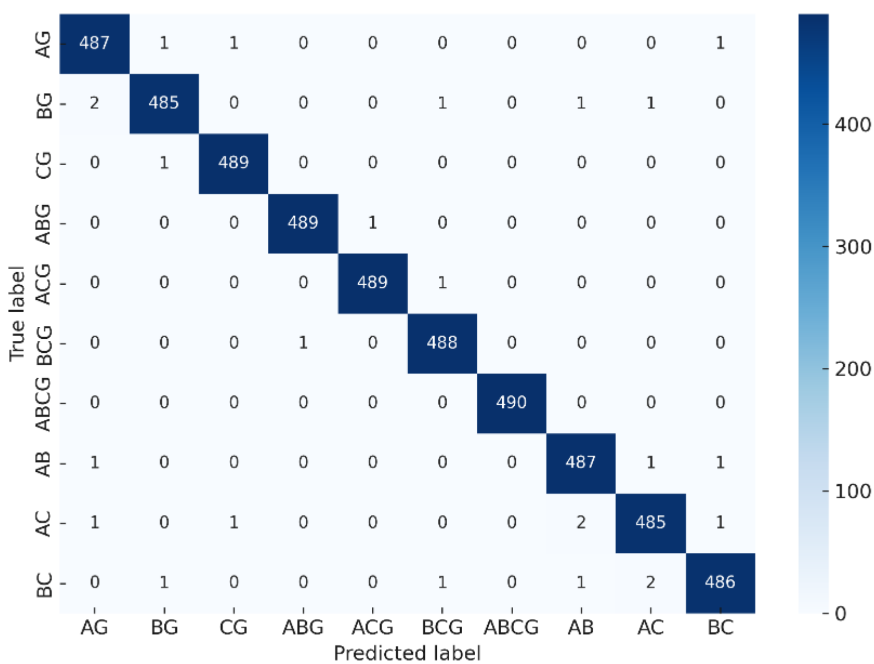
When the learning rate is set to 0.01, the batch size is 128, and the dropout rate is 0.2, the evaluation metrics for Fault Point 1 are shown in Table 3. The model is capable of identifying most types of distribution network faults. The evaluation accuracy reached 99.49%, as shown in Figure 5. The model is generally accurate in identifying fault types; however, there is some deviation in determining the fault phase.

To further verify the advantages of the autoencoder-based identification model used in this paper for distribution network fault identification, the constructed dataset is compared with other related research methods, including support vector machine (SVM) (Manohar and Koley, 2017), Decision Tree (DT) (Mishra et al., 2015), and CNN (Ash et al., 2019). The comparison results are presented in Table 4. The table shows that SVM has a shorter training time but the longest testing time, making it difficult to use for real-time computation and rapid fault diagnosis. DT has both shorter training and testing times, but its identification accuracy is relatively low. The CNN algorithm requires a longer time to train the network, but both its accuracy and testing time are satisfactory. The approach introduced in this paper can effectively capture the temporal characteristics of the training set. Although it achieves the highest accuracy, its training time is significantly longer compared to other methods. It is worth mentioning that, compared to the simulation without the TFE indicator, the method used in this paper reduced model training time by 23% while also improving accuracy, further demonstrating the effectiveness and reliability of the proposed model. The above experiments fully illustrate that the approach introduced in this paper demonstrates superior fault identification performance for distribution networks.

Table 3 Evaluation metrics table for fault point 1

Fault type	*Precision*	*Recall*	*Specificity*
AG	0.992	0.994	0.999
BG	0.994	0.990	0.999
CG	0.996	0.998	1.000
ABG	0.998	0.998	1.000
ACG	0.998	0.998	1.000
BCG	0.996	0.998	1.000
ABCG	1.000	1.000	1.000
AB	0.994	0.994	0.999
AC	0.994	0.990	0.999
BC	0.996	0.992	0.999

Figure 5 Fault identification confusion matrix (see online version for colours)



To further evaluate the training behaviour of the proposed model, we analysed the relationship between the number of training epochs and performance metrics such as accuracy and loss. Figure 6 shows the training curve for the IEEE 33-node dataset. As illustrated, the model exhibits rapid improvement in both accuracy and loss during the initial 50 epochs. Following this phase, the model continues to optimise, with performance gradually stabilising after approximately 150 epochs. Minor fluctuations are observed throughout the training process due to stochastic mini-batch gradient updates, which are typical in deep learning optimisation. The stable convergence of both metrics confirms that the model effectively captures critical fault features while maintaining robustness during training. This also supports the choice of 300 training epochs for the final model configuration.

Figure 6 Training curve with realistic fluctuations (see online version for colours)

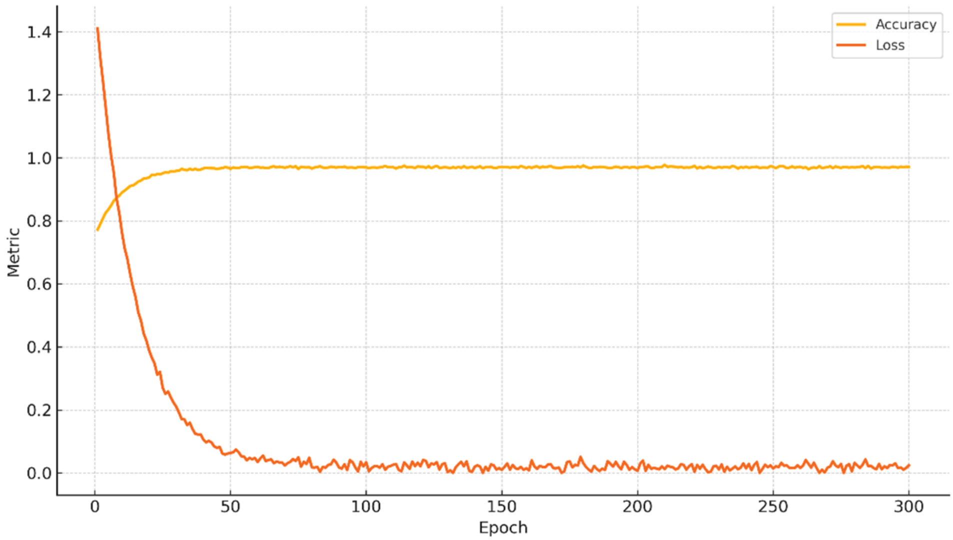


Table 4 Comparison table of different studies for fault point 1

Fault type	*Accuracy* (%)	*Training time* (s)	*Testing time* (s)
SVM	96.62	121.67	5.32
DT	97.23	12.42	0.67
CNN	98.82	872.28	0.21
Proposed method (without TFE)	99.22	1321.53	0.12
Proposed method	99.49	1013.17	0.11

5.3 *Cross-domain task identification comparison*

Although the autoencoder used in this paper achieves high accuracy in fault classification, the model's training time is too long. Relying solely on the autoencoder model for distribution network fault classification requires retraining whenever the fault node or operating conditions change, which is highly time-consuming and inefficient. Further integrating the proposed active transfer model, the fault identification accuracy for Fault Points 2, 3, and 4 is calculated. The fault diagnosis accuracy and computation time under different learning rates are shown in Table 5. From the table, it can be observed that transferring the learning model from Fault Point 1 to other nodes allows for rapid training while maintaining a high fault identification rate. Additionally, the table shows that by properly setting parameters, the interference caused by incorrect pseudo-labels in the TD during early training can be mitigated, thereby improving the accuracy of all transfer tasks.

Table 5 Transfer learning results table under different learning rates

		Learning rate		
		0.001	*0.0001*	*0.00001*
Fault Point 2	Accuracy (%)	99.86	99.83	99.78
	Training time (s)	9.23	15.32	11.25
	Testing time(s)	0.11	0.18	0.13
Fault Point 3	Accuracy (%)	99.62	99.45	99.49
	Training time (s)	9.76	16.43	12.15
	Testing time(s)	0.08	0.13	0.10
Fault Point 4	Accuracy (%)	99.53	99.43	99.68
	Training time (s)	11.23	19.11	14.79
	Testing time(s)	0.13	0.19	0.15
Average accuracy (%)		99.70	99.57	99.64

To demonstrate the superiority of the proposed method, three advanced active transfer learning methods were also introduced for comparison: BADGE (Su, 2020), AADA (Huang et al., 2018), and TQS. Table 6 presents the classification accuracy and duration of each method. The proposed method outperforms the compared methods in classification accuracy, and actively learning labelled TD data further enhances classification performance. In terms of time consumption, the proposed method is similar to AADA and outperforms BADGE, which requires computations in the gradient space, as well as TQS, which involves training multiple classifiers. This demonstrates that the proposed method not only achieves better classification accuracy but also has an advantage in training efficiency.

Table 6 Transfer learning results table for fault points 2, 3, and 4

Methods	*Fault Point 2*		*Fault Point 3*		*Fault Point 4*		*Average Accuracy* (%)	*Average Time* (s)
	Accuracy (%)	*Time* (s)	*Accuracy* (%)	*Time* (s)	*Accuracy* (%)	*Time* (s)		
BADGE	99.39	16.21	99.28	17.91	99.12	19.06	99.70	17.06
AADA	99.42	15.38	99.32	17.00	99.21	18.19	99.57	16.19
TQS	99.53	31.68	99.37	35.02	99.35	35.35	99.64	33.35
Proposed method	99.86	9.23	99.62	9.76	99.53	11.23	99.70	10.07

This study evaluated the proposed model based on the IEEE 33-node system and a synthetic dataset, demonstrating its effectiveness in fault diagnosis and classification. Further validation is needed to assess its applicability in real-world distribution networks, as the complexity of the grid structure and the presence of actual noise may significantly impact the model's performance. Future work should focus on applying this model to real distribution systems with actual noisy datasets to further validate its robustness and generalisation ability.

An important aspect to be explored is the relationship between the number of autoencoder layers and fault identification accuracy. Although the current model demonstrates strong feature extraction capabilities, optimising the depth of the autoencoder can enhance its ability to capture complex fault patterns, especially in highly dynamic environments. Studying this relationship can provide valuable insights for further improving the model's performance.

By extending the application scope to real-world scenarios and optimising the model architecture, future research will help make the proposed method more practical and reliable for large-scale deployment in modern distribution networks.

6 Conclusions

In this study, we propose a novel approach for fault diagnosis in distribution networks based on active transfer learning combined with a self-attention-based autoencoder. Specifically, TFE features are integrated into the input layer of the autoencoder to capture complex dynamic behaviours of voltage and current signals effectively, significantly enhancing the extraction of discriminative fault characteristics. An active transfer learning framework incorporating a MMD-based adaptation layer is constructed, effectively addressing the problem of domain differences between labelled source data and unlabelled target data.

Extensive simulation experiments conducted on the IEEE 33-node benchmark distribution network demonstrate that the proposed method achieves a high fault classification accuracy of 99.49%, significantly outperforming traditional methods such as SVM, decision trees, and CNNs in terms of both accuracy and computational efficiency. The analysis of training dynamics further confirms the robustness and stability of our method, showing rapid convergence and stable performance despite the presence of artificially introduced noise designed to emulate practical measurement disturbances.

Overall, the proposed active transfer learning method exhibits strong cross-domain generalisation capabilities and robustness under noisy measurement conditions, making it particularly suitable for real-world deployment in modern distribution networks with increasing penetration of distributed energy resources. Future research directions include further validation using actual noisy datasets from operational distribution grids, as well as optimisation of the network architecture to enhance performance and practical applicability.

Funding

This research was funded by the Science and Technology Project of Guizhou Power Grid Co., Ltd., grant number GZKJXM20232292. The APC was funded by Guizhou Power Grid Co., Ltd.

Conflicts of interest

All authors declare that they have no conflicts of interest.

Author contributions

Conceptualisation, Youzhuo Zheng; methodology, Youzhuo Zheng; software, Youzhuo Zheng and Hengrong Zhang; validation, Youzhuo Zheng, Jiang Lin, and Kailei Chen; formal analysis, Youzhuo Zheng; investigation, Kun Zhou; resources, Jiang Lin; data curation, Kun Zhou and Kailei Chen; writing – original draft preparation, Jiang Lin and Hengrong Zhang; writing – review and editing, Youzhuo Zheng and Kailei Chen; visualisation, Kailei Chen; supervision, Kailei Chen; project administration, Youzhuo Zheng; funding acquisition, Jiang Lin. All authors have read and agreed to the published version of the manuscript.

Data availability statement

Not applicable.

References

Alhanaf, A.S., Balik, H.H. and Farsadi, M. (2023) 'Intelligent fault detection and classification schemes for smart grids based on deep neural networks', *Energies*, Vol. 16, No. 22, p.7680, DOI: 10.3390/en16227680.

Ash, J.T., Zhang, C., Krishnamurthy, Langford, J. and Agarwal (2020) 'Deep batch active learning by diverse, uncertain gradient lower bounds', *International Conference on Learning Representations (ICLR 2020)*, Addis Ababa, Ethiopia, DOI: 10.48550/arXiv.1906.03671.

Baran, M.E. and Wu, F.F. (1989) 'Network reconfiguration in distribution systems for loss reduction and load balancing', *IEEE Trans. Power Deliv.*, Vol. 4, No. 2, pp.1401–1407, DOI: 10.1109/61.25627.

Cano, A., Arévalo, P., Benavides, D. and Jurado, F. (2024) 'Integrating discrete wavelet transform with neural networks and machine learning for fault detection in microgrids', *Int. J. Electr. Power Energy Syst.*, Vol. 155, No. B, p.109616, DOI: 10.1016/j. ijepes.2023.109616.

Dai, Q., Wang, Y., Han, X., Gao, J., Zhang, Y., Li, K. and Jiang, Q. (2024) 'Analysis of protection misoperation caused by two-point grounding of current transformer secondary circuit', *Power Systems and Big Data*, Vol. 27, No. 12, pp.63–74, DOI: 11.19317/J.CNKI.1008-083X. 2024.12.008.

Dai, Z., Wang, B., Zhou, T. and Wang, K. (2021) 'False-tolerant fault location method for active distribution network based on LSTM', *Power Systems and Big Data*, Vol. 24, No. 7, pp.40–48, DOI: 10.19317/J. CNKI.1008-083X.2021.07.006.

Guillen, D., Serna, J.A.D., Zamora-méndez, A., Paternina, M.R.A. and Salinas, F. (2021) 'Taylor-Fourier filter-bank implemented with O-splines for the detection and classification of faults', *IEEE Trans. Ind. Informat.*, Vol. 17, No. 5, pp.3079–3089, DOI: 10.1109/TII.2020.3010268.

Guo, S.Y., Miao, S.H., Zhao, H.P., Yin, H.R. and Wang, Z.X. (2020) 'A novel fault location method of a 35-kV high-reliability distribution network using wavelet filter-s transform', *Energies*, Vol. 13, No. 19, p.5118, DOI: 10.3390/en13195118.

Hassan, M., Biswas, S.P., Chowdhury, S.A.H., Mondal, S., Islam, M.R. and Shah, R. (2024) 'Fault classification and localization of multi-machine-based IEEE benchmark test case power transmission lines using optimizable weighted extreme learning machine', *Electr. Power Syst. Res.*, Vol. 235, p.110857, DOI: 10.1016/j. epsr.2024.110857.

Huang, S.J., Zhao, J.W. and Liu, Z.Y. (2018) 'Cost-effective training of deep CNNs with active model adaptation', *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. Association for Computing Machinery,* London, UK, pp.1580–1588.

Huang, S., Yan, C. and Qu, Y. (2023a) 'Deep learning model-transformer based wind power forecasting approach', *Front. Energy Res*., Vol. 10, p.1055683, DOI: 10.3389/FENRG.2022. 1055683.

Huang, Z., Chen, Y., Liu, J., Ni, H. and Zhang, D. (2023b) 'Rolling bearing fault diagnosis based on the enhanced channel attention network', *Cyber-Physical Systems*, Vol. 10, No. 2, pp.197–213, https://doi.org/10.1080/23335777.2022.2163705

Khavari, S., Dashti, R., Shaker, H.R. and Santos, A. (2020) 'High impedance fault detection and location in combined overhead line and underground cable distribution networks equipped with data loggers', *Energies*, Vol. 13, No. 9, p.2331, DOI: 10.3390/en13092331.

Li, S., Wu, H., Wang, X., Xu, B., Yang, L. and Bi, R. (2023) 'Short-term load forecasting based on AM-CIF-LSTM method adopting transfer learning', *Front. Energy Res*., Vol. 11, p.1162040, DOI: 10.3389/FENRG.2023.1162040.

Liu, H., Liu, S., Zhao, J.B., Bi, T.S. and Yu, X.J. (2022) 'Dual-channel convolutional network-based fault cause identification for active distribution system using realistic waveform measurements', *IEEE Trans. Smart Grid*, Vol. 13, No. 6, pp.4899–4908, DOI: 10.1109/TSG. 2022.3182787.

Manohar, M. and Koley, E. (2017) 'SVM-based protection scheme for microgrid', *Proceedings of the 2017 International Conference on Intelligent Computing, Instrumentation and Control Technologies (ICICICT)*, Kerala, India, pp.429–432, DOI: 10.1109/ICICICT1.2017.8342600.

Mishra, D.P., Samantaray, S.R. and Joos, G. (2015) 'A combined wavelet and data-mining based intelligent protection scheme for microgrid', *IEEE Trans. Smart Grid*, Vol. 7, pp.2295–2304.

Mo, H., Peng, Y., Wei, W., Xi, W. and Cai, T. (2023) 'SR-GNN based fault classification and location in power distribution network', *Energies*, Vol. 16, No. 1, p.433, DOI: 10.3390/en16 010433.

Moradzadeh, A., Teimourzadeh, H., Mohammadi-Ivatloo, B. and Pourhossein, K. (2022) 'Hybrid CNN-LSTM approaches for identification of type and locations of transmission line faults', *Int. J. Electr. Power Energy Syst.*, Vol. 135, p.107563, DOI: 10.1016/j. ijepes.2021.107563.

Saad, M., Kim, C.H. and Munir, N. (2022) 'Single-phase auto-reclosing scheme using particle filter and convolutional neural network', *IEEE Trans. Power Deliv.*, Vol. 37, No. 6, pp.4775–4785, DOI: 10.1109/TPWRD.2022.3159256.

Stefanidou-Voziki, P., Sapountzoglou, N., Raison, B. and Dominguez-Garcia, J.L. (2022) 'A review of fault location and classification methods in distribution grids', *Electr. Power Syst. Res.*, Vol. 209, p.108031, https://doi.org/10.1016/j.epsr.2022.108031.April

Su, J.C., Tsai, Y.H., Sohn, K., Liu, B., Maji, S. and Chandraker, M. (2020) 'Active adversarial domain adaptation', *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, IEEE, Snowmass, USA, pp.739–748.

Sucic, V., Saulig, N. and Boashash, B. (2014) 'Analysis of local time-frequency entropy features for nonstationary signal components time supports detection', *Digit. Signal Process*., Vol. 34, pp.56–66, DOI: 10.1016/j. dsp.2014.07.013.

Touati, K.O.M., Boudiaf, M., Mazouz, L. and Cherroun, L. (2023) 'Efficient hybrid strategy based on FFT and fuzzy logic techniques applied to fault diagnosis in power transmission line', *Soft Comput.*, Vol. 28, pp. 8911–8930, DOI: 10.1007/s00500-023-09089-6.

Veerasamy, V., Wahab, N.I.A., Othman, M.L., Padmanaban, S., Sekar, K., Ramachandran, R., Hizam, H., Vinayagam, A. and Islam, M.Z. (2021) 'LSTM recurrent neural network classifier for high impedance fault detection in solar PV integrated power system', *IEEE Access*, Vol. 9, pp.32672–32687, DOI: 10.1109/ACCESS.2021.3060800.

Yu, X.Y., Cao, J., Fan, Z., Xu, M.M. and Xiao, L.Y. (2023) 'Faulty feeder identification in resonant grounding distribution networks based on deep learning and transfer learning', *CSEE Journal of Power and Energy Systems*, Vol. 9, No. 6, pp.2168–2178, https://doi.org/10.17775/CSEEJPES.2020.05450

Zhang, R.F., Yao, W., Shi, Z.T., Ai, X.M., Tang, Y. and Wen, J.Y. (2023) 'Towards multi-scenario power system stability analysis: an unsupervised transfer learning method combining DGAT and data augmentation', *IEEE Transactions on Power Systems*, Vol. 38, No. 6, pp.5367–5380, https://doi.org/10.1109/TPWRS.2022.3220569

Zhong, J., Chen, Y., Gao, J. and Lv, X. (2024) 'Drone image recognition and intelligent power distribution network equipment fault detection based on the transformer model and transfer learning', *Front. Energy Res.*, Vol. 12, p.1364445, DOI: 10.3389/FENRG.2024.1364445.

Zou, M., Zhao, Y., Yan, D., Tang, X.L., Duan, P. and Liu, S.W. (2022) 'Double convolutional neural network for fault identification of power distribution network', *Electr. Power Syst. Res.*, Vol. 210, p.108085, DOI: 10.1016/j. epsr.2022.108085.