Article citation info:
Wang D, Zhou L, Xie L, Dong S, Liu J. High-voltage circuit breaker fault voiceprint recognition based on prototype similar domain adaptive spectral morphological neural network. Diagnostyka. 2025;26(3):2025301. https://doi.org/10.29354/diag/206079.

DIAGNOSTYKA, 2025, Vol. 26, No. 3

e-ISSN 2449-5220
DOI: 10.29354/diag/206079

HIGH-VOLTAGE CIRCUIT BREAKER FAULT VOICEPRINT RECOGNITION BASED ON PROTOTYPE SIMILAR DOMAIN ADAPTIVE SPECTRAL MORPHOLOGICAL NEURAL NETWORK

Dong WANG, Lubo ZHOU *, Liyun XIE, Xipu LIU, Shiqi DONG, Junhua LIU

State Grid Shanghai Ultra High Voltage Company, Shanghai 200122, China
* Corresponding author, e-mail: bravezlb123@163.com

Abstract

High-voltage circuit breakers will emit continuous vibration signals during operation. The signals contain a large number of pulses and fluctuations caused by faults. They are the main data source for evaluating the functioning condition for high-tension breaker switches. To aim for examine its vocal features in vibration acoustic signals of high-tension breaker switches in different operating states, a prototype similarity domain adaptive spectral morphological neural network (PSDA-SMNN) was proposed. First, the vibration signal is denoised using spectral morphology variational mode decomposition combined with fast singular value decomposition method. Secondly, the labeled data of a certain operational state serves utilized for this information origin area, while this untagged information from different operational states serves utilized for this training objective area, and the prototype network distance similarity is used to align the feature distribution between the domains; then, meta-training is used to the domain network undergoes internal supervised training, and the network in the target domain undergoes external unsupervised training using the virtual label backpropagation algorithm. Through internal and external loop training, the difference in feature distribution between domains is reduced, and unlabeled faults of high-tension breaker switches in varying operational states are recognized. Test outcomes indicate which the suggested framework is able to precisely identify the malfunction operational condition for high-tension breaker switches and detect common issues for high-tension breaker switches under various interference settings, having a classification precision of approximately 95%.

Keywords: high-voltage circuit breaker, voiceprint recognition, prototype similarity domain, deep transfer learning, spectral morphology variational mode decomposition

1. INTRODUCTION

As a critical apparatus in the electrical network, the dependability and steadiness of the high-tension breaker's operational mechanism are directly linked to the secure operation of the whole electrical network. With the increase of electrical demand and the advancement of technology, it is especially crucial to perform continuous surveillance and malfunction diagnosis of the operational mechanism of high-tension breakers. Traditional diagnostic methods mainly rely on manual inspection and regular maintenance. This method is not only inefficient, b ut also difficult to detect potential faults in time, increasing the risk of accidents. Therefore, developing an efficient, accurate and automated diagnostic technology has become an important issue in modern power system maintenance [1-3].

As a non-contact detection method, voiceprint recognition has shown great potential within this area for machinery anomaly detection. It analyzes the sound signals generated by mechanical equipment during work, extracts feature information, and combines it with machine learning algorithms to achieve fault classification and prediction. Li et al. and Xie et al. [4-5] proposed a method based on short-time Fourier transform (STFT), which can convert the sound signal within this temporal range to a frequency map to capture different frequency components changes. However, STFT has the problem of low resolution when processing non-stationary signals, which limits its application in complex working conditions. In order to overcome this limitation, Wang et al. introduced the wavelet transform (WT) technology, which can decompose signals at multiple scales and improve the ability to capture transient features. Nonetheless, WT is more sensitive to noise and can easily lead to misjudgment [6-8]. In addition, Chen et al. applied the long short-term memory network (LSTM) to voiceprint recognition, which effectively solved the memory problem of temporal sequence information and enhanced the construction of long-term dependencies. model ability [9-11]. However,

Received 2025-02-08; Accepted 2025-06-09; Available online 2025-07-02

LSTM still has the problem of gradient disappearance when faced with high-frequency changing voiceprint features, which affects the convergence speed of the model [12-14]. Hu et al. proposed an improved LSTM model combined with an attention mechanism. By introducing an attention weight allocation mechanism, the model can focus on important feature areas and improve the discrimination of different types of faults. At the same time, Zhao et al. applied an autoencoder for unsupervised feature learning, which can complete preliminary feature extraction in the absence of label data, further improving the model's general performance [15-17] capabilities.

The domain adaptation method is also an extremely effective method for fault diagnosis. By adding virtual pseudo-labels to the target domain data based on the self-training method, and improving the credibility of the virtual pseudo-labels through backpropagation, label-free recognition of the target domain is achieved [18-19]. Sun et al. proposed a reverse pseudo-label optimal transmission method to calculate the distribution differences in different domains through Wasserstein and Gromov-Wasserstein clustering [20-21]. Han et al. [22] proposed a deep migration method, using CNN to establish a feature extractor and using JDA to align features in different domains, and used virtual pseudo labels to reversely update the feature extractor parameters to complete the label-free state. Identify. Yang et al. [23] proposed a deep transfer learning method of polynomial kernel-induced MMD, which uses virtual labels to learn the conditional distribution shifts of different domain features. Feng et al. [24] proposed a semi-supervised meta-learning network that squeezes incentive attention, and achieved label-free fault identification of bearings under varying working conditions through prototype distance similarity.

This paper proposes a prototype similarity domain adaptive spectral morphological neural network (PSDA-SMNN), aiming to solve the limitations of existing technologies for anomaly detection in high-tension breaker switches. Specifically, we first use spectral morphology variational mode decomposition combined with fast singular value decomposition method to denoise the vibration signal. Secondly, the labeled data of a certain operational state serves utilized for this information origin area, while this untagged information from different operational states serves utilized for this training objective area. The prototype network distance similarity is used to align the feature distribution between the domains. Then, the source domain network is internally supervised and trained using meta-training, and the target domain network is externally unsupervised using the virtual label back propagation algorithm. The difference in feature distribution between domains is reduced by internal and external loop training, and unlabeled faults of high-tension breaker switches in varying operational states are recognized.

The primary achievements from the document are as listed:

Innovative denoising method: A denoising method combining spectral morphological variational mode decomposition with fast singular value decomposition is proposed, which significantly improves the signal quality.

Prototype similarity domain adaptation framework: A new prototype similarity domain adaptation framework is constructed, which aligns the characteristic spread in the origin area and the destination area through the prototype network distance similarity, and enhances the generalization ability of the model.

Efficient internal and external loop training strategy: An efficient and stable training process is achieved by combining meta-training and virtual label back propagation.

Experimental verification and performance comparison: Extensive experimental verification was conducted through the acoustic dataset collection from the high-tension breaker switch operating mechanism under four different working conditions. The outcomes indicate that the suggested framework possesses a diagnostic precision of 95% for typical faults in different noise environments. %, which is better than several existing mainstream models.

2. PROTOTYPE SIMILARITY DOMAIN ADAPTIVE SPECTRAL MORPHOLOGY NEURAL NETWORK

2.1. Spectral morphology variational mode decomposition denoising process

In order to accurately extract the voiceprint information contained in the oscillation waveform from the electrical switch and eliminate the adverse effects of noise, the vibration signal is subjected to noise reduction processing. Perform spectral morphological variational mode decomposition on the original vibration signal. First perform variational mode decomposition, and then determine the value of the decomposition layer number K: initialize the intrinsic type operation, central pitch, Lagrangian operator, and the count of cycles n=0, the number of decomposition layers K=2, the process is as follows:

$$\frac{\hat{x}(\omega)-\sum_{i<k}^{K} \hat{\mu}_i^{n+1}(\omega)-\sum_{i>k}^{K} \hat{\mu}_i^{n}(\omega)+\frac{\hat{\lambda}^n(\omega)}{2}}{1+2\alpha(\omega-\omega_k^n)^2} \rightarrow \hat{\mu}_k^{n+1}(\omega) \tag{1}$$

$$\frac{\int_0^\infty \omega \,|\, \hat{\mu}_k^{n+1}(\omega) \,|^2 \, d\omega}{\int_0^\infty |\, \hat{\mu}_k^{n+1}(\omega) \,|^2 \, d\omega} \rightarrow \omega_k^{n+1} \tag{2}$$

$$\hat{\lambda}^n(\omega) + \tau(\hat{x}(\omega) - \sum_{k=1}^{K} \hat{\mu}_k^{n+1}(\omega)) \rightarrow \hat{\lambda}^{n+1}(\omega) \tag{3}$$

In the formula: x is the input signal, k=1,2..., K; α is the penalization factor, τ is the time step, and "→" means update.

According to the above formula, update. The iteration stop condition is:

$$\sum_k \frac{||\hat{\mu}_k^{n+1}-\hat{\mu}_k^n||^2}{||\hat{\mu}_k^n||_2^2} < \varepsilon \tag{4}$$

Perform fast singular value decomposition on the determined eigenmode function:

Assume that the determined intrinsic mode function is $\hat{\mu}_k = [y_1, y_2, \dots y_N]$, N is the length of the signal, then the Hankel matrix can be constructed.

$$A = \begin{bmatrix} y_1 & y_2 & \cdots & y_n \\ y_2 & y_3 & \cdots & y_{n+1} \\ \vdots & \vdots & \ddots & \vdots \\ y_m & y_{m+1} & \cdots & y_N \end{bmatrix} \tag{5}$$

Where: $N = m + n - 1$, select half of the length of the oscillation waveform as the best count of rows m of the Hankel matrix. Matrix A can be decomposed into:

$$A = U\Sigma V^T \tag{6}$$

In U represents the left principal matrix, V represents the right principal matrix; the square matrix $\Sigma = \begin{bmatrix} \Sigma_1 & 0 \\ 0 & 0 \end{bmatrix}$, and $\Sigma_1 = diag(\lambda_1, \lambda_2, \cdots, \lambda_r)$, are the singular values $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_r > 0$ of matrix A.

Perform signal reconstruction and determine the vibration signal after noise reduction. The signal after noise reduction is shown in Fig. 1. After noise reduction processing, the white noise interference in the environment is filtered out.

2.2. Prototype similarity domain adaptive network

The architecture framework configuration developed in this document is presented in Fig. 2. The framework comprises source domain feature extractor (Gf1), fault classifier (Gd1), target domain feature extractor (Gf2), and fault classifier (Gd2). During the training process, internal loop training is first performed on Gf1 and Gf2, so that the source domain network can learn the status knowledge of the circuit breaker, and then the network parameters of Gd1 and Gd2 are updated through external loop training. During the entire process, the network parameters within the source region and target zone are uniformly disseminated.

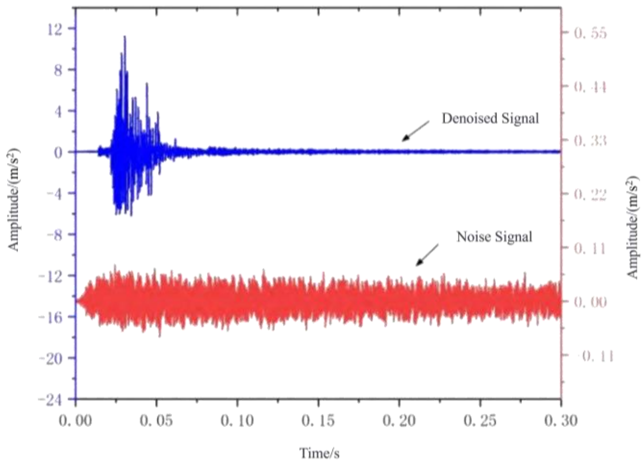

Fig. 1. Noise-reduced signal of normal closing operation for high-voltage circuit breaker

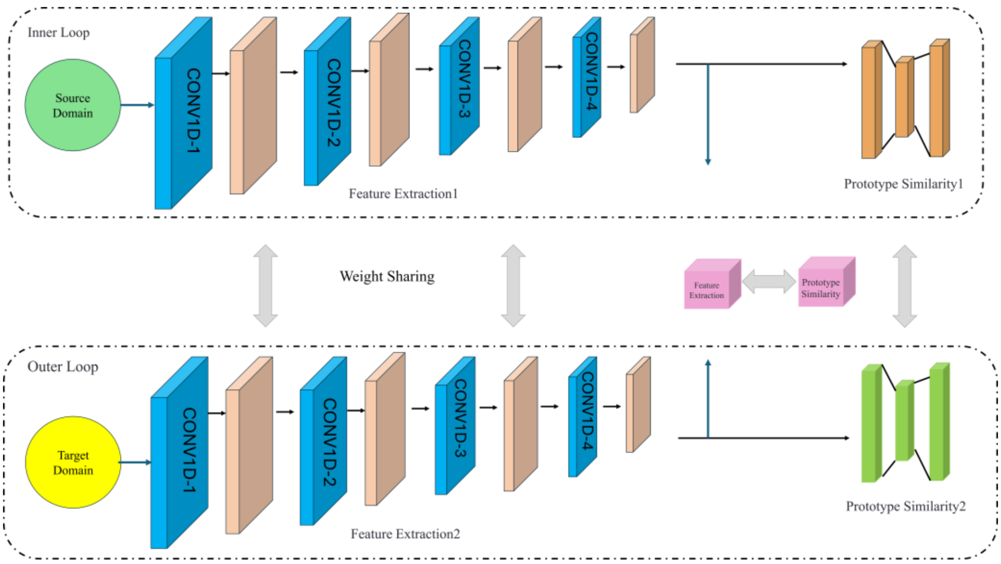

Fig. 2. Network model architecture diagram

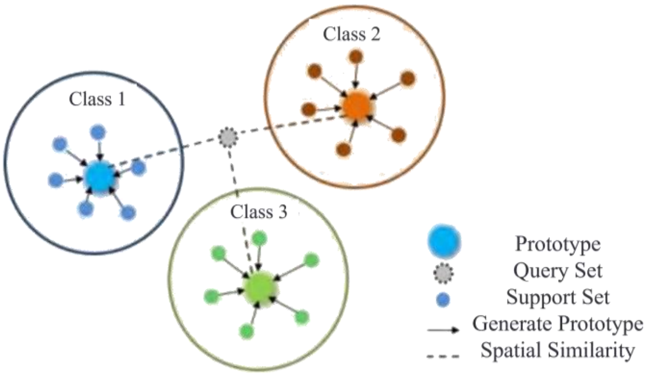

Fig. 3. Schematic diagram of prototype network

To purpose for precisely determine the current condition of the circuit breaker under changing operating conditions, this paper proposes to use a prototype network and virtual pseudo-label strategy to mine the invariant characteristics of domain faults and diagnose unlabeled faults of the electrical switch in varying operational scenarios.

The network assumes that in a multi-category space, each category has a prototype point. The closer the sample in the space is to the prototype point, the greater the probability that the sample belongs to the prototype point label. Fig. 3 shows the corresponding 3-category sample. Prototype point. The prototype network maps input features to M-dimensional feature space through a deep neural network. By calculating the mean of the sample in M-dimensional space, k is the sample label; represents the number of samples belonging to the kth category.

Take the data from the query set to verify the prototype network, that is, calculate the distance from the prototype point and normalize the distance:

$$D_\phi(y=k\mid x)=\frac{exp(-d(f_\phi(x),C_k)}{\sum_{k'} exp(-d(f_\phi(x),C_{k'})} \tag{7}$$

In the above formula, d is the elastic distance; $k' \in \{1,2,\dots,J\}$ is the sample type. The cost operation throughout the learning procedure is as outlined:

$$L_\phi=-\frac{1}{n_q}\sum_{(\tilde{x_i},k_i')\in Q} \quad log(D_\phi(y=k'|\tilde{x_i})) \tag{8}$$

The target domain virtual loss function constructed in this article is as follows:

$$F=\begin{bmatrix} f_{1,1} & f_{1,2} & \cdots & f_{1,m} \\ f_{2,1} & f_{2,2} & \cdots & f_{2,m} \\ \vdots & \vdots & \ddots & \vdots \\ f_{n,1} & f_{n,2} & \cdots & f_{n,m} \end{bmatrix} \tag{9}$$

$$P=Onchot(max(Softmax(F))) \tag{10}$$

$$P=\begin{bmatrix} p_1 & 0 & \cdots & 0 \\ p_2 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & p_m \end{bmatrix} \tag{11}$$

$$L_s=-\sum_{i=1}^{n} \quad \hat{p}_i g\, log(f_i), \hat{p}_i=\{p_i,\dots,0\}, \tag{12}$$

In the above formula: F is the output feature of the fault classification model in the target domain; P is the virtual label obtained by processing the output feature using the Softmax function combined with Onehot encoding technology; is the cross-entropy loss function, used to calculate the error between F and P to update the entire Network parameters. The calculation process of virtual tags is shown in Fig. 4.

The diagnosis process is: perform supervised inner-loop training on the source domain model so that the model can learn the state knowledge of the source domain. Perform outer loop training on the target domain network, first update the feature extraction model parameters of the origin realm and destination realm grounded upon the gap among the query set and the prototype point, and then use virtual labels to update the relevant model parameters of the target domain. Through internal and external training, the invariant features of domain faults are mined to complete the transfer of

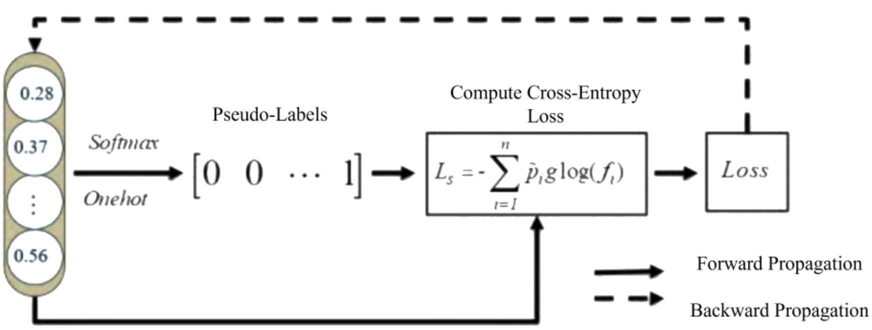

Fig. 4. The process of computing Pseudo-Labels

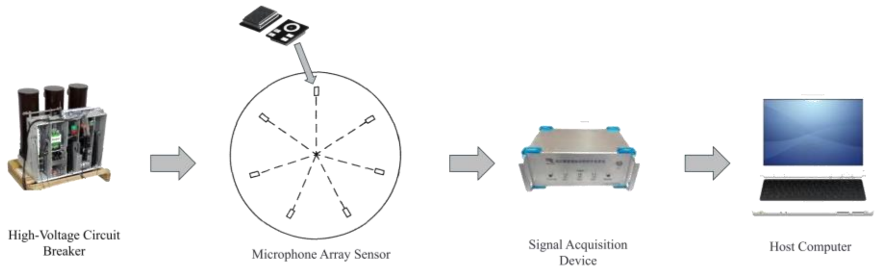

Fig. 5. Voiceprint acquisition system

source domain knowledge. Test model: Model testing is divided into origin realm and destination realm fault testing. The source domain testing is to prove the feature extraction capability of the feature extractor, and the target domain testing is to prove that the method can diagnose the variable working condition unlabeled fault of the circuit breaker.

3. EXPERIMENTAL ANALYSIS

3.1. Experimental device

A high-voltage circuit breaker in northwest China was used as the experimental object for voiceprint collection. Fig. 5 shows the voiceprint collection system of the operating mechanism regarding the high-tension electrical switch. In accordance with the system collection front end, the CHZ-213+YG-201 sound pressure sensor was selected, and its nominal sensitivity is 50mV. /Pa, the frequency range is 20Hz-20kHz, and the signal is sent to the host computer using a signal acquisition device. According to the sampling theorem and combined with engineering experience, the sampling rate of each channel of the acquisition system is set to 48kHz. Arrange the sensor array. The sound pressure sensor is arranged on the circular outline of the circuit breaker (1.5m away from the reference emission surface of the circuit breaker) with a height of 0.75m. The distance between sensors on the same side is d=0.5m. All sensors have capacitive poles. The head is facing the circuit breaker, and the 7-channel sensors are M1~M7 in order3.2. PSDA-SMNN model verification analysis

In this section, several transfer learning tasks with multiple source domains are designed using multiple datasets. Such endeavors intend to ascertain the efficacy of the suggested approach. This document employs three dataset collections to ascertain the efficacy of the suggested approach.

The voiceprint data set 1 (D1), data set 2 (D2), and data set 3 (D3) of the high-voltage circuit breaker operating mechanism under different environmental working conditions were selected respectively. In this study, D1, D2, and D3 are composed of four status modes, namely normal closing (N),

The shaft pin falls off (F1), spring fatigue (F2), and the iron core is stuck (F3). Each failure mode contains 150,000 voiceprint data, which are organized into 1000 samples; each sample contains 150 voiceprint data. The dataset collection is partitioned into learning collection and evaluation collection in accordance with the proportion of 7:3.

The D1-D2, D1-D3, and D2-D3 diagnostic tasks were designed using the data sets D1, D2, and D3. Transferring tasks D1-D2 signifies that the framework is educated upon the annotated origin dataset gathered under operational scenario D1 and conveyed to the unannotated destination dataset gathered under operational scenario D2. In the pre-education procedure, the acquisition velocity is established to {0.001, 0.002 0.003, 0.004, 0.005}, the value range of the hyperparameter is {0.1, 0.2 0.3, 0.4, 0.5}, and Adam is used to optimize the controller, model accuracy and learning the relationship between rate and hyperparameters is shown in Fig. 6.

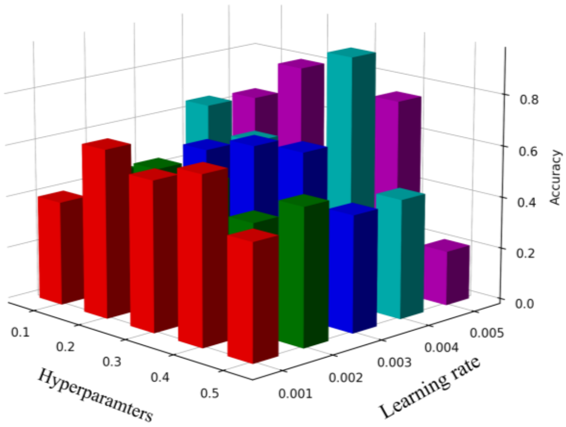

Fig. 6. 3D Visualization of learning rate, hyperparameters, and model accuracy

To purpose for ascertain the efficacy of the framework, juxtapose it with the efficacy of the subsequent neural network: MMD-DANN (Maximum Mean Discrepancy Domain Adversarial Neural Network) is a domain adaptation framework that combines maximum mean difference (MMD) and realm antagonistic instruction. By minimizing the distance between the feature distributions of the source domain and the target domain, and the adversarial training mechanism is used to make the model have better generalization capabilities for data in different fields. It is especially suitable for tasks that require precise alignment of feature distributions in two fields to ensure that the model can perform well in new environments. The performance is stable and reliable.

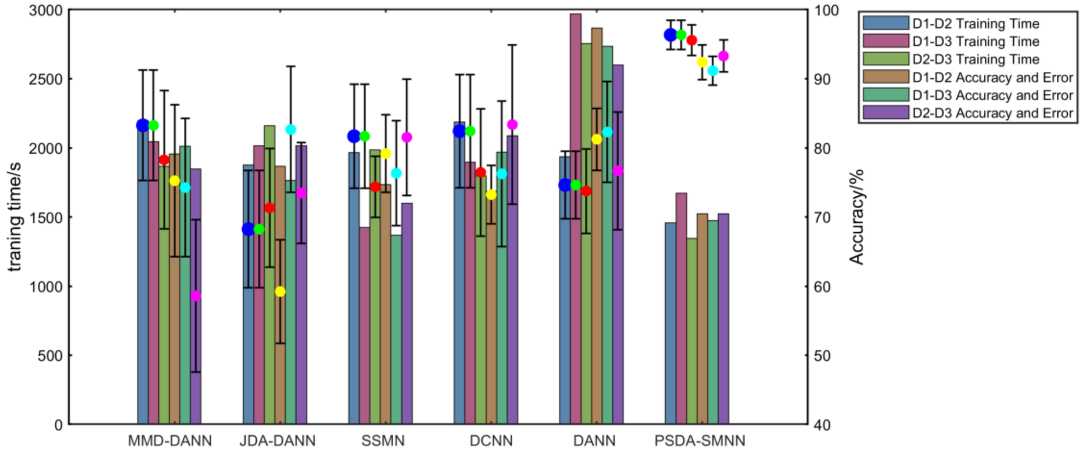

Fig. 7. Training accuracy charts for different networks on various tasks

JDA-DANN (Joint Distribution Adaptation Domain Adversarial Neural Network) introduces Joint Distribution Adaptation (JDA), which not only considers the alignment of marginal distributions in the feature space, but also considers the alignment of conditional distributions, through concurrently harmonizing the entry characteristics of the origin realm and the destination realm. and label distribution, which improves the model's adaptability to complex and changeable data distribution. It is especially suitable for handling situations where there are significant differences among the origin realm and the destination realm, and can effectively improve the model's performance in cross-domain tasks. SSMN (Self-Supervised Meta Network) uses self-supervised learning and meta-learning mechanisms to gradually adjust model parameters to adapt to the data distribution in new fields. It can achieve effective unsupervised domain adaptation without accessing the target domain labels, and is particularly suitable for processing For large-scale or complex data sets, the learning of target domain samples is self-enhanced by constructing pseudo-labels, which enhances the model's ability to understand unlabeled data, thereby improving the overall generalization performance, especially in a few-shot learning environment. Excellent. DCNN is a deep convolutional neural network that is widely used in numerous areas like visual identification and linguistic speech analysis. Although it is not specifically designed for domain adaptation, it can be used through appropriate architectural adjustments and technical means (such as Adversarial training, feature alignment, etc.), can be effectively applied to domain adaptation tasks, and its powerful feature extraction capabilities make it perform well in processing many types of input data, especially in high-dimensional data processing. DANN (Domain Adversarial Neural Network) is a classic domain adaptation method that reduces the gap among the origin realm and the destination realm by introducing a domain classifier (Domain Classifier) and conducting confrontation training with the main task classifier (Task Classifier). Difference, simple and effective, enables the model to learn domain-independent feature representations through the adversarial training mechanism, improving the model's generalization ability in the target domain. It is especially suitable for tasks that require rapid adaptation to new fields and can effectively handle complex tasks. of multimodal data (see Fig. 7It is observed evident within the information in Diagram 7, the suggested model PSDA-SMNN achieved the highest accuracy (95% on average) and the fastest diagnosis speed (average processing time of 1514 seconds) in each diagnosis task. Compared with several existing mainstream models, the proposed model exhibits significant advantages. MMD-DANN has a slow convergence speed in complex and changeable data environments, resulting in longer training time. In contrast, PSDA-SMNN significantly shortens the training time while maintaining high accuracy by introducing fast singular value decomposition and internal and external loop training mechanisms. In the D1-D2 migration diagnosis task, the accuracy of MMD-DANN was 82.27%, while PSDA-SMNN reached 96.45%; in addition, the training time of MMD-DANN was approximately 2249 seconds, while PSDA-SMNN only took 1498 seconds. The JDA-DANN adversarial training mechanism increases the computational burden and affects the training efficiency. PSDA-SMNN aligns the feature distribution through the distance similarity of the prototype network, simplifying the training process and reducing the consumption of computing resources. In the D1-D3 tasks, JDA -The accuracy of DANN is 79.54%, while PSDA-SMNN reaches 95.5%, and the training time is shortened by about 20%. SSMN relies on a large number of pre-training steps, and the overall training time is long. PSDA-SMNN combines internal and external loop training with virtual label backpropagation, which not only speeds up the training speed, but also improves the generalization ability of the model. In D2-D3 tasks Among them, the accuracy of SSMN is 81.8%, while PSDA-SMNN reaches 95.2%, and the training time is also shorter. When DCNN is applied to domain adaptation tasks, additional architectural adjustments and technical means are required. PSDA-SMNN integrates a variety of advanced technologies and is directly applied to high-voltage circuit breaker fault diagnosis without complicated adjustments, showing higher flexibility and practicality. The average accuracy of DCNN is 83.57%, while PSDA-SMNN reaches 95.27%. DANN: Its simple structure has limited performance when processing complex multi-modal data. On the basis of inheriting the advantages of DANN, PSDA-SMNN further optimizes the feature alignment and unsupervised training mechanism, significantly improving the accuracy of the model under different working conditions. and stability. Across all tasks, the average accuracy of DANN was 83.72%, while PSDA-SMNN reached 95.27%.

To sum up, PSDA-SMNN neither merely attains the supreme precision and swiftest diagnostic velocity in diverse diagnostic tasks, yet too attains the supreme precision and swiftest diagnostic velocity through integrating a variety of advanced technologies, such as fast singular value decomposition, internal and external loop training mechanism, and prototype network distance similarity alignment. And virtual label backpropagation, which greatly reduces the training time and computing resource consumption, improves the model's generalization ability and its flexibility and practicality in practical applications.

In order to intuitively demonstrate the superior performance of PSDA-SMNN in feature representation and domain adaptation, we used the t-SNE (t-Distributed Stochastic Neighbor Embedding) diagram for visual analysis, as shown in Fig. 8, starting west toward east, starting apex toward base, they are the t-SNE diagrams of DCNN, SSMN and PSDA-SMNN. t-SNE is a powerful dimensionality reduction technique that can map high-dimensional features into two- or three-dimensional space, making the distribution between different categories more clearly visible. By comparing the t-SNE plots of DCNN, SSMN and PSDA-SMNN, we can clearly see the significant advantages of PSDA-SMNN in feature representation: The t-SNE plot of PSDA-SMNN demonstrates a clearer separation between source and target domain features, even in complex and changing data environments. In contrast, the feature distribution of DCNN is relatively mixed and it is difficult to distinguish samples of different categories; while SSMN has certain improvements, there is still a certain overlap when processing large-scale unlabeled data. PSDA-SMNN effectively diminishes the characteristic dispersion disparity among the origin realm and the destination realm through internal and external loop training combined with virtual label backpropagation. It can be seen from the t-SNE diagram that the feature points of PSDA-SMNN form closer and consistent clusters between the two domains, indicating that it has stronger adaptability and generalization ability to data under different working conditions. The t-SNE plot also reveals the high-accuracy performance of PSDA-SMNN in classification tasks. Fault samples of different categories are clearly distinguished in the feature space of PSDA-SMNN with almost no crossover phenomena. This not only verifies the efficient learning ability of the model during the training process, but also proves its reliability and stability in practical applications.

In summary, through the visual analysis of t-SNE diagrams, PSDA-SMNN has demonstrated significant advantages in feature representation, domain adaptability and accuracy, providing a more efficient and reliable solution for high-voltage circuit breaker fault diagnosis.

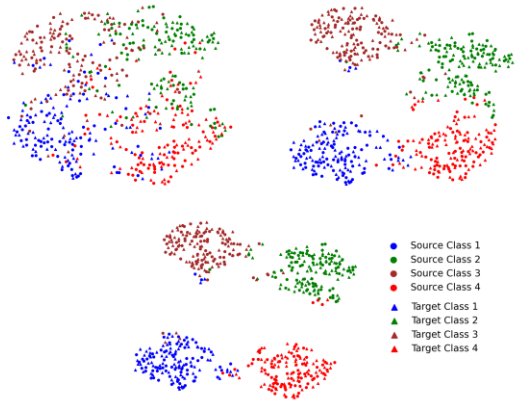

Fig. 8. T-SNE plots of DCNN, SSMN, and PSDA-SMNN

4. CONCLUSION

This paper proposes a prototype similarity domain adaptive spectral morphological neural network (PSDA-SMNN), aiming to solve the limitations of existing technologies for anomaly identification in high-tension loop interrupters. Specifically, we first use spectral morphology variational mode decomposition combined with fast singular value decomposition method to denoise the vibration signal. Secondly, the labeled data of a specific operational state serves as the information origin realm, the untagged information from alternative operational states serves as the instructional destination realm, and the prototype network distance similarity is used to align the feature distribution between the domains. Then, the source domain network is internally supervised using meta-training, and the target domain network is externally unsupervised using the virtual label backpropagation algorithm. Through internal and external loop training, the difference in feature distribution between domains is reduced, and unmarked anomalies in high-tension loop interrupters amid diverse operational states are recognized.

This paper conducts extensive experimental verification through t the acoustic dataset collection from the high-tension loop interrupter operational system in four diverse operational states. The outcomes indicate that the suggested framework possesses an diagnostic precision of about 95% for typical faults in different noise environments, which is better than Several existing mainstream models.

Source of funding: *This study was funded by the Science and Technology Project of State Grid Shanghai Municipal Electric Power Company (520950240009).*

Author contributions: *Dong Wang Conceptualization and Data Curation. Lubo Zhou Formal Analysis and Investigation. Liyun Xie Methodology as well as Resources. Xipu Liu data Interpretation and Paper revision. Shiqi Dong Validation and Visualization. Junhua Liu Writing – Original Draft Preparation.*

Declaration of competing interest: *The authors declare that they have no conflict of interest.*

REFERENCES

1. Zang CY, Zeng J, Li P, et al. Intelligent diagnosis model of mechanical fault for power transformer based on SVM algorithm. High Voltage Apparatus (China) 2023; 59(12): 216-222+229. https://doi.org/10.1109/ICSRS.2018.8688719
2. Shao Y, Wu JW, Ma SL, et al. Integrated SVM method with AM-ReliefF feature selection for mechanical fault diagnosis of high voltage circuit breakers. Proceedings of the CSEE (China) 2021; 41(08): 2890-2901.
3. Wan ST, Ma XD, Chen L, et al. State evaluation and fault diagnosis of high-voltage circuit breaker based on short-time energy entropy ratio of vibration signal and DTW. High Voltage Engineering (China) 2020; 46(12): 4249-4257.
4. Wang S, Zhou Y, & Ma Z. Research on fault identification of high-voltage circuit breakers with characteristics of voiceprint information. Sci Rep. 2024; 14: 9340. https://doi.org/10.1038/s41598-024-59999-0.
5. Chen H, Han C, Zhang Y, et al. Investigation on the fault monitoring of high-voltage circuit breaker using improved deep learning. PLOS ONE 2023; 18(12):e0295278. https://doi.org/10.1371/journal.pone.0295278.
6. Ye XY, Yan J, Wang YX, et al. A novel capsule convolutional neural network with attention mechanism for high-voltage circuit breaker fault diagnosis. Electric Power Systems Research 2022; 209:108003. https://doi.org/10.1016/j.epsr.2022.108003.
7. Ye X, Yan J, Wang Y, et al. A novel U-Net and capsule network for few-shot high-voltage circuit breaker mechanical fault diagnosis. Measurement, 2022;199:111527. https://doi.org/10.1016/j.measurement.2022.111527,
8. Zhang D, Wang Y, Li Z, et al. Analysis of pre-breakdown arcing time in vacuum circuit breakers based on empirical wavelet transform. Transactions of China Electrotechnical Society 2024; 1-12.
9. Chen H, Li Z, Zhang S, et al. Short-term transformer error prediction based on MHA-CNN-SLSTM and error compensation. Power System Protection and Control 2024; 52(24), 74-84.
10. Ruan J, Wang X, Zhou T, et al. Fault identification of high voltage circuit breaker trip mechanism based on PSR and SVM. IET Generation, Transmission & Distribution. 2023;17(6):1179-1189. https://doi.org/10.1049/gtd2.12725.
11. Yang Q, Liao Y, Li J, et al. Cross-domain zero-shot learning for enhanced fault diagnosis in high-voltage circuit breakers. Neural Networks 2024; 180: 106681. https://doi.org/10.1016/j.neunet.2024.106681.
12. Li T, Xia Y, Pang X, et al. Mechanical fault diagnosis of high voltage circuit breaker using multimodal data fusion. PeerJ Computer Science 2024; 10: e2248. https://doi.org/10.7717/peerj-cs.2248.
13. Yan J, Wang Y, Yang Z, et al. Few-shot mechanical fault diagnosis for a high-voltage circuit breaker via a transformer–convolutional neural network and metric meta-learning. In IEEE Transactions on Instrumentation and Measurement 2023; 72: 1-11. https://doi.org/10.1109/TIM.2023.3309387.
14. Moradzadeh A, Pourhossein K, Mohammadi-Ivatloo B, Mohammadi F. Locating inter-turn faults in transformer windings using isometric feature mapping of frequency response traces. IEEE Trans. Ind. Informat. 2021; 17(10): 6962–6970. https://doi.org/10.1109/TII.2020.3016966
15. Zhao Y, Xu H, Tian Y, et al. Fault detection of distribution network photovoltaic inverters based on a fusion attention mechanism autoencoder. Smart Power 2024; 52(11): 1-7.
16. Chen Y, Duan W, He Y, et al. State of health estimation for batteries using denoising autoencoder and gated recurrent hybrid neural network. Transactions of China Electrotechnical Society 2024; 39(24): 7933-7949.
17. Luo G, Yang X, Shang B, et al. High-impedance ground fault detection in distribution networks based on an improved stacked denoising autoencoder. Power System Protection and Control 2024; 52(24): 149-160.
18. Wang J, Yan X, Ye Q, et al. Few-shot transfer learning with attention mechanism for high-voltage circuit breaker fault diagnosis. In IEEE Transactions on Industry Applications 2022; 58(3): 3353-336. https://doi.org/10.1109/TIA.2022.3159617.
19. Vu-Cong T, Dalstein M, Toigo C, Jacquier F, Girodet A. Partial discharge measurement in DC GIS: Comparison between conventional and UHF methods. in Proc. IEEE Conf. Electr. Insul. Dielectric Phenomena (CEIDP) 2021; 607–610. https://doi.org/10.1109/CEIDP50766.2021.9705420
20. Chen L, Wan S, Dou L. Improving diagnostic performance of high-voltage circuit breakers on imbalanced data using an oversampling method. In IEEE Transactions on Power Delivery 2022; 37(4): 2704-2716. https://doi.org/10.1109/TPWRD.2021.3114547.
21. Qin Y, Qian Q, Wang Y, et al. Intermediate distribution alignment and its application into mechanical fault transfer diagnosis. IEEE Transactions on Industrial Informatics 2022; 18(10): 7305-7315. https://doi.org/10.1109/TII.2022.3149934.
22. Han T, Liu C, Yang WG, et al. Deep transfer network with joint distribution adaptation: A new intelligent fault diagnosis framework for industry application. ISA Transactions 2020; 97: 269-281. https://doi.org/10.1016/j.isatra.2019.08.012
23. Xu H. A hybrid knowledge-based and data-driven method for aging-dependent reliability evaluation of high-voltage circuit breaker. In IEEE Transactions on Power Delivery 2023; 38(6): 4384-4396. https://doi.org/10.1109/TIE.2019.2953010.

24. Feng Y, Chen JL, Zhang TC, et al. Semi supervised meta-learning networks with squeeze-and excitation attention for few-shot fault diagnosis. ISA Transactions 2022; 120: 383-401. https://doi.org/10.1016/j.isatra.2021.03.013

Dong WANG received the Master's degree from the Shanghai Jiaotong University. He is currently working in State Grid Shanghai Ultra High Voltage Company. His current research interests including high voltage and insulation technology and substation maintenance.
e-mail: attribe@hotmail.com

Lubo ZHOU received the Master's degree from the Shanghai Jiaotong University. He is currently working in State Grid Shanghai Ultra High Voltage Company. His current research interests including high voltage and insulation technology and electrical maintenance.
e-mail: bravezlb123@163.com

Liyun XIE received the Bachelor's degree from the Shanghai University of Electric Power. She is currently working in State Grid Shanghai Ultra High Voltage Company. Her current research direct is high voltage and insulation technology.
e-mail: 17574212@qq.com

Xiepu Liu received the Master's degree from the Southeast University. He is currently working in State Grid Shanghai Ultra High Voltage Company. His main research direct is power electronics.
e-mail: shengming781118@163.com

Shiqi Dong received the Bachelor's degree from the Wuhan University. She is currently working in State Grid Shanghai Ultra High Voltage Company. Her current research direct is high voltage and insulation technology.
e-mail: 1123960516@qq.com

Junhua Liu received the Ph.D degree from the Shanghai Jiaotong University. He is currently working in State Grid Shanghai Ultra High Voltage Company. His main research direct is high voltage and insulation technology.
e-mail: drunhua@163.com