1

REINFORCEMENT-ENHANCED DEEP TRANSFER FUSION NETWORK FOR MULTI-DOMAIN TRANSFORMER FAULT DIAGNOSIS VIA ACOUSTIC SIGNATURES

Mingliang MU * , **Jianfei NIU** , **Guangqian GUO** , **Maoxiang TIAN** , **Jinlin LI** , **Jingyi LI**

Binzhou Power Supply Company, State Grid Shandong Electric Power Company, Binzhou 256600, China
* Corresponding author, e-mail: bzgdmm88@163.com

Abstract

During operation, power transformers generate continuous vibrational signatures bearing mechanical fault-induced impulses, which constitute the fundamental evidence for equipment condition assessment. To investigate acoustic-fingerprint characteristics under varying operating conditions, this paper proposes a Reinforced Integrated Deep Transfer Learning Network (REDTLN) for multi-fault-domain diagnostics. The methodology first constructs multiple specialized Deep Transfer Learning Networks (DTLNs) using novel kernel Maximum Mean Discrepancy (kMMD) variants, enabling source-specific adaptation to enrich transferable feature representations. Subsequently, a unified unsupervised ensemble framework integrates multi-metric divergence measures, employing a reinforcement-guided combinatorial search algorithm to discover optimal DTLN integration rules. This intelligent fusion mechanism significantly enhances multi-source transfer capability, improving diagnostic accuracy and robustness in dynamic noise environments and complex operational scenarios. Experimental results confirm the model's efficacy in precisely identifying abnormal states while maintaining sustained >95% accuracy for representative faults under diverse acoustic-interference conditions.

Keywords: power transformer, voiceprint recognition, ensemble learning, deep transfer learning, target domain adaptation

1. INTRODUCTION

Power converters are the cornerstone for safe transfer and allocation of electric energy. Real-time and accurate monitoring of their operating conditions is crucial for guaranteeing the robustness and secure operation of the network [1-2].

In order to detect these faults timely and accurately, and even predict some dangerous operating conditions, it is necessary to study online status monitoring and fault diagnosis methods of transformers. According to the working principle of a power transformer, when the transformer is working, the vibration of the transformer core and windings will emit continuous and stable sound signals, which are tightly associated with the operational state of the equipment [3-4]. Alternatively, the condition monitoring technique grounded in the voiceprint characteristics of the sound signal has the advantages of being non-destructive and having low requirements on the environment and detection equipment. It has been applied in many fields, especially the operating condition monitoring of transformers combined with deep learning [5-6]. Related studies have been published in recent years. Wu et al. [7] analyzed the noise propagation mode and parameters affecting performance, and elaborated on the importance of voiceprint recognition for Transformer status monitoring and fault diagnosis; Rodriguez-Serna et al. [8] utilized The threshold value obtained by vibration information fusion is used to judge the DC bias fault of 500 kV auto-transformer; in [9], the Mel frequency representation of the acoustic signal is input into the deep learning model to achieve voiceprint-based Transformer operating status judgment; Hou and Bergmann [10] used the visual geometric group network to classify the gamma cepstral coefficient characteristic parameters of the sound wave and distinguish the diverse operational states of the 10 kV dry Transformer. However, various random environmental noises can cause differences between training signals and actual signals, thereby reducing fault diagnosis accuracy.

Collective training is a beneficial AI technique that seeks to enhance overall effectiveness by merging various distinct predictors through efficient guidelines [11]. Hence, collective training serves as an admirable option for executing multi-origin field adaptation learning. Via collective training, valuable insights embedded within numerous origin fields can be thoroughly harnessed. For guided collective

training, integration principles are typically crafted based on precision or a compromise between precision and variety [12-13]. Nonetheless, in this examination, the destination field datasets lack labels, which will render it considerably more challenging to achieve a potent integration principle.

Strengthening education is an additional potent artificial intelligence technique. For example, Zoph et colleagues crafted an artificial network architecture hunt strengthening education technique founded on strategy slope and realized commendable outcomes. Wang et associates utilized it for roller defect detection [14-15]. Motivated by such, our team investigates employing strengthening education for ideal amalgamation guideline creation for collective training paradigms. A strengthening education system comprises an orchestrator that fabricates maneuvers (amalgamation guidelines) and a supervisor that acquires incentives. The orchestrator comprises a Recurrent Neural Network (RNN) strategy slope employed to revise settings.

In this examination, a Reinforcement Ensemble Deep Transfer Learning Network (REDTLN) is suggested to execute the migration education challenge in multiple-origin zones. Its central concepts encompass three principal components: single-origin single-goal zone adaptation, uniform gauge blueprint for unsupervised collective training, and multi-origin and multi-paradigm strengthening amalgamation. Initially, a sequence of profound transfer learning networks (DTLN) with diverse novel kernel MMDs (kMMD) are established to study various transferable characteristics. This investigation employs some innovative core operations for draft distinct kMMDs and assemble diverse DTLNs. In such manner, the efficiency of diverse core operations might be confirmed and the range of separate paradigms might be augmented. The poly-origin zone and destination zone data collections are arranged into numerous pairs, and every duo is input to the DTLN for individual origin-singular destination zone adjustment. Subsequently, a fresh uniform unsupervised collective training performance assessment indicator is devised based on MMD and diversity metrics. Ultimately, a uniform gauge is utilized as an incentive to construct a strengthening education technique to seek effective amalgamation regulations of these DTLNs and attain multi-origin and multi-paradigm strengthening amalgamation.

2. REINFORCEMENT ENSEMBLE DEEP TRANSFER LEARNING NETWORK

In this segment, a strengthening collective profound adaptation education network is suggested to address the adaptation education challenge in multi-origin zones. It is partitioned into the following five sections, namely: the foundational theory of zone adjustment, single-origin and single-destination zone adjustment, uniform metric blueprint for unguided collective training, and multi-origin and multi-paradigm enhancement integration.

2.1. Basic theory of domain adaptation

Over the past few years, profound learning approaches for defect identification have realized significant triumphs [16-17]. On most occasions, they tackle the issue with sufficient tagged training information [18-23]. Training and examination datasets must originate from an identical distribution [24]. However, occasionally it is challenging to fulfill all aforementioned criteria. If no tagged data are accessible, current profound learning techniques might not be sensible. Consequently, adaptation learning was devised to address this predicament by utilizing tagged data from some differing yet associated fields [25].
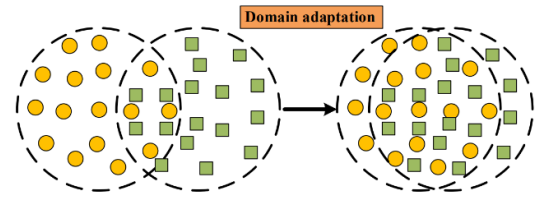

Fig. 1. Schematic diagram of domain adaptation

As depicted in Fig. 1, various fields frequently generate field discrepancies, so merely training with tagged data from an alternative field may lead to inferior performance on the target field. To this aim, adjustment is introduced to diminish distribution variations. Presently, MMD (Maximum Mean Discrepancy) is the most prevalent distance metric to lessen the variations between distinct fields [26-28]. TCA, JDA DDC, and DAN are representative MMD-based adjustment techniques.

To address the challenge of defect identification with untagged data, this document employs an adjustment technique utilizing kMMD. Suppose there exists a source field D_S and a destination field D_T , $X = \{\hat{y}^i | X_i, i = 1,2,\dots,s\}$ denotes the origin field data set, where s indicates the quantity of samples; the sample i is represented by X_i, and $\hat{y}^i$ represents its label; $X = \{Z_j, j = 1,2,\dots,t\}$ represents the untagged destination field data set; t denotes the sample count; Z_j represents the jth sample. The kMMD between the source zone D_S and the target zone D_T can be determined as Eq. (1):

$$M_k(X,Z) = \frac{1}{s^2} \sum_{i,j=1}^{s} k(X_i, X_j)$$
$$+ \frac{1}{t^2} \sum_{i,j=1}^{t} k(Z_i, Z_j) \qquad (1)$$
$$- \frac{2}{st} \sum_{i,j=1}^{s,t} k(X_i, Z_j)$$

Which $k(x,y)$ indicates the core function, like straight core, Gaussian core, etc.

2.2. Single source single target domain adaptation

His examination employs kMMD to assemble an individual-origin individual-destination zone adaptive DTLN. As depicted in Fig. 2, the assembled DTLN comprises three ConvM, a flatten layer (FL), two fully linked tiers (FC) and a probability sorter (C). Every filtering unit sequentially includes a filter tier (Conv), a batch standardization tier (BN), and a peak pooling tier (MPL).

First ConvM Unit: Employs a 1D convolutional structure with an input channel count of 1, output channel count (number of filters) of 256, kernel size of 1×1, stride of 1, no padding, and no bias.

Second ConvM Unit: Input channel count of 256, output channel count of 512, kernel size maintained at 1×1, stride of 1, no padding, and no bias.

Third ConvM Unit: Configuration identical to the second unit, with input channel count of 512, output channel count of 512, kernel size of 1×1, stride of 1, and no bias.

The dimensionality design of the fully connected layers (FC) is:

First FC Layer: Takes flattened features from the convolutional output, reducing dimensionality from 512 to 128.

Second FC Layer: Further compresses the 128-dimensional features to the dimension corresponding to the number of fault categories.

The input of the model in this paper is the time-domain signals of acoustic fingerprints collected directly from acoustic/vibration sensors. This model is designed to analyze and extract features from these acoustic fingerprint time-domain signals.
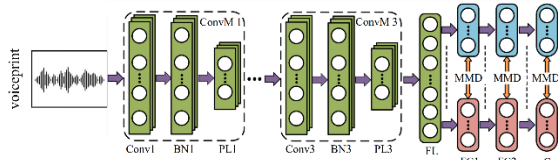

Fig. 2. DTLN architecture

In this examination, merely the count of ConvM1 sieves is modifiable to lessen the complexity of hyperparameter choice. Following determination of the hyperparameters, DTLN undergoes pre-training employing tagged source information before zone adaptation. Throughout the preliminary education procedure, the Adam method is utilized to refine the later target metric, as shown in Eq. (2):

$$J_E = -\frac{1}{s}\sum_{i=1}^{s}(\hat{y}^i log y^i) \quad (2)$$

in which s denotes the count of instances; y^i and $\hat{y}^i$ represent the predicted tag and actual tag of the ith instance, respectively. The research results show that high-level features are more susceptible to domain differences than basic-level characteristics. As illustrated in Fig. 2, to derive field-consistent traits and discriminators, we utilize many-tier field

adjustment via shell MMD in wholly linked tiers FC1, FC2 and classifiers.

Primary aim within DTLN largely comprises pair divisions. Initial division represents information expense derived input domain, utilized toward directed education, whereas secondary division encompasses densely linked layer alongside core MMD pertaining model discrimination, used for domain adaptation, as shown in Eq. (3):

$$J_{DTLN} = J_E + \alpha \left[\sum_{i \in L} M_k(F_{D_S,i}, F_{D_T,i}) + M_k(\hat{y}_{D_S}, y_{D_T}) \right] \quad (3)$$

Among them, J_E represents the cross-entropy cost on the origin field, $L = \{FC1, FC2\}$, $F_{D_S,i}, F_{D_T,i}$ represents the origin field and destination field characteristics of the ith layer respectively, $\hat{y}_{D_S}$ signifies the true classification for the source domain, y_{D_T} signifies the forecasted classification of the destination domain, and α serves as the limiting factor.

To acquire portable characteristics, DTLN is utilized to execute single-origin single-destination zone adaptation. Multi-origin zone data and destination zone data are arranged into pairs. Each pair comprises an origin zone dataset and a destination zone dataset. To boost variety, six DTLNs featuring six unique kMMDs were compiled. As shown in Table 2, the cores of kMMD include Gaussian core alongside other innovative cores, such as exponential core, Cauchy core, hyperbolic tangent core, polynomial core, and logarithmic core. Hence, should there be N source areas, one would generate N sets of data collections. Every set of data collections is input into six varieties of DTLN for uni-source uni-target area adjustment. An aggregate of 6N domain acclimation models shall result.

2.3. A novel integrated evaluation framework for unsupervised collective education

Many of the current collective education techniques handle situations with annotated information, that is, guided collective education. Guided collective education typically creates integration guidelines via precision or a compromise of precision and variety metrics. Certain variety metrics utilized in guided collective education also necessitate the accuracy of an individual foundational framework. Nevertheless, limited endeavors have been undertaken regarding unguided collective education. Within this investigation, the destination field datasets are unannotated, thus the precision is not obtainable. Absent annotation details, it grows more challenging to formulate efficient integration guidelines. Consequently, there exists an immediate necessity to devise a novel approach to produce combination principles appropriate for unguided collective education. Within this investigation, an integrated evaluation criterion is structured to assess the effectiveness of collective

education and direct the system to produce efficient integration guidelines. As no annotation details are accessible, an integrated evaluation criterion is structured relying on alternative viable options. The initial option is the MMD among the forecasted annotation of M_ST collective education and the annotations of numerous origin fields, calculable as Eq. (4):

$$M_{ST} = \sum_{i=1}^{N} M_K(\hat{y}_{D_S^i}, y_{D_T^e}) \qquad (4)$$

Within these, N signifies the count of origin fields; the true annotation of the origin field is denoted as $\hat{y}_{D_S^i}$; *the forecasted annotation of collective education is denoted as* $y_{D_T^e}$. This criterion can diminish the discrepancy in annotation distribution between the origin field and the destination field. Hence, the lesser M_{ST}, the superior the effectiveness. Variety is crucial to the triumph of collective education and is considered in crafting integrated metrics. Variety metrics encompass paired metrics and unpaired metrics. Paired metrics traditionally evaluate the disparity between two foundational learners and achieve the ultimate overall variety by averaging all paired metrics. Non-paired metrics largely directly gauge the variety of the collective instead of computing the average of paired metrics. This research examines both paired and unpaired metrics to acquire varied frameworks. The discrepancy metric is utilized to execute the paired metric and is computed as Eqs. (5) and (6):

$$
\begin{aligned}
&div\_p \\
&= \frac{2}{n(n-1)} \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} dis(y_{D_T(i)}, y_{D_T(j)})
\end{aligned} \qquad (5)
$$

$$
\begin{aligned}
&dis(y_{D_T(i)}, y_{D_T(j)}) \\
&= \frac{1}{t} \sum_{k=1}^{t} \mathrm{II}(y_{D_T(i)}^k \neq y_{D_T(j)}^k)
\end{aligned} \qquad (6)
$$

Within these, div_p signifies the total variety, $dis(y_{D_T(i)}, y_{D_T(j)})$ signifies the coupled variety between the primary learner $y_{D_T(i)}$ and $y_{D_T(j)}$; m≤6N denotes the quantity of chosen primary learners, II() signifies the flag function; when the forecasted annotation of the ith primary framework is $y_{D_T(i)}^k$ discordant with the forecasted annotation of the jth primary framework $y_{D_T(j)}^k$, the output is 1, otherwise it is 0. The div_p v score span is from 0 to 1. The higher the score, the more significant the variety.

Shannon entropy serves to assess uncoupled variety, computed as Eqs. (7) and (8):

$$
\begin{aligned}
div\_np &= \frac{1}{t} \sum_{k=1}^{t} \sum_{i=1}^{c} (-H(y_{D_T}^k \\
&= i|Z_k) log H(y_{D_T}^k \\
&= i|Z_k))
\end{aligned} \qquad (7)
$$

$$H(y_{D_T}^k = i|Z_k) = \frac{1}{n} \sum_{j=1}^{1} \pi(y_{D_T(j)}^k = i) \qquad (8)$$

where c denotes the count of classes. Likewise, the higher the score, the more significant the variety. Ultimately, these three evaluations are integrated into a comprehensive evaluation to assess the effectiveness of collective education more efficiently. The comprehensive evaluation E is presented in Eq. (9):

$$S = \mu \frac{1}{M_{ST}} + \beta div\_p + \omega div\_np \qquad (9)$$

Within these, μ , β and ω denote the balance factors for three distinct evaluations; to ensure the coherence of the evaluations, we adopt the inverse of M_{ST}. The greater the integrated evaluation S is, the superior the collective education effectiveness is.

These balancing factors are automatically learned through a reinforcement learning-based adaptive optimization framework. Specifically, during the training process of REDTLN, the weighting parameters are modeled as trainable variables. They undergo joint optimization with the feature transfer objectives via policy gradient algorithms, PPO algorithms, enabling them to dynamically adapt to the distribution discrepancy characteristics across different source-target domain combination.

2.4. Multi-source and multi-model enhanced integration

Within this chapter, multi-origin and multi-framework enhancement integration shall be attained by merging collective education and enhancement learning. Collective education is utilized to integrate the outcomes of 6P DTLNs to derive collective outcomes. Enhancement learning is formulated to produce appropriate integration guidelines for the collective.

Typically, the integration guidelines of collective education encompass mean voting and weighted voting. Mean voting allocates identical weight to every framework for the collective. Weighted voting assigns varying weights to every framework according to some criterion (such as precision). Rational allocation of voting weights serves as an assurance to secure effective integration. Toward this aim, an enhancement learning approach is formulated to explore the optimal integration of guidelines for DTLNs. As depicted in Fig. 3, the enhancement collective model comprises four components: regulator, operation, sub-collective model, and incentive.
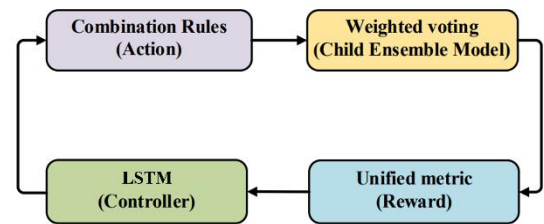

Fig. 3. Enhanced integrated model

As depicted in Fig. 4, the regulator is constructed via a long short-term memory network (LSTM) and a softselect classifier. The regulator emits the likelihood of the selection operation.
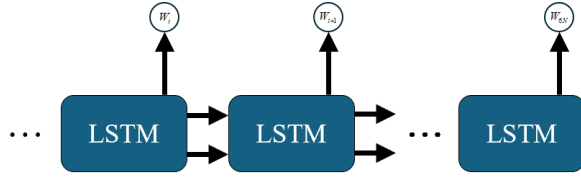

Fig. 4. LSTM stack structure

LSTM is introduced as an exceptional kind of recursive neural web (RNN) for handling temporal sequence progressive data. The framework includes self-links and three components, specifically entry valve, exit valve, and reset valve. These valves enable LBMN storage units to retain and retrieve information across extended durations, alleviating the diminishing slope issue. The computation procedure is illustrated in Eqs. (10)-(14):

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \qquad (10)$$
$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \qquad (11)$$
$$C_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \qquad (12)$$
$$C_t = f_t * C_{t-1} + i_t * C_t \qquad (13)$$
$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \qquad (14)$$

Within these; f_t, i_t, o_t individually signify a regulation factors for a reset valve; entry valve or exit valve within LSTM: C_{t-1} denotes an quantity for a storage cell within LSTM at a prior instant (temporal point): Storage cells serve crucial elements in LSTM frameworks for retaining and transmitting data: An C_t signifies a fresh input obtained by LSTM at a present instant: That might be novel information from an input series, utilized to refresh the storage cell: Denotes the storage cell condition of LSTM at a present instant: Storage cells are locations where data is preserved within the framework and can hold data from past instants as well as the effect of fresh inputs: An v_z signifies the result of LSTM at a present instant, too termed the concealed state: That is the data delivered by the LBMN framework to the top or bottom layer, typically employed for forecasting or categorization of tasks: W and b: signify the mass array and offset array accordingly: $\sigma(x)$ is an activation function and its formula is $\sigma(x) = \frac{1}{1+e^{-x}}$; $Tanh(x)$ is a activation function and its formula is $Tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$: Presuming that the parameters of the regulator are denoted as θ_c, the goal function is computed as Eq. (15):

$$J_c(\theta_c) = -\frac{1}{6N} \sum_{i=1}^{6N} (\hat{P}_i log P_i + \lambda \sum \theta_c^2) \qquad (15)$$

Within these, the initial segment denotes the disorder cost; the following section signifies the mass reduction clause; λ represents the associated mass reduction coefficient. The operation within enhancement education serves the integration guideline, the selection mass for each DTLN, which forms a chain of 6N items. Operations are articulated as $A = w = [w_1, w_2, \dots, w_{6N}]$ produced with likelihood P.

Upon an integration guideline being secured, a sub-collective framework may be assembled via weighted selection. The outcomes from collective education might be achieved. Presume that the integration guideline emitted by the regulator is denoted as Eq. (16):

$$Score_{D_T(e)}^k = \sum_{i=1}^{6N} (w_i y_{D_T(i)}^k) \qquad (16)$$

In accordance with the ultimate selection outcome $Score_{D_T(e)}^k$, the classification with the peak rating shall be termed the collective's forecasted identifier, symbolized as $y_{D_T}^e$. For executing strategy slope proficiently, incentives are the paramount component. Within this examination, we employ a harmonized criterion S to specify incentives R, as depicted in Eq. (17):

$$R_q = S_q \qquad (17)$$

Within these, S_q and R_q are the quantities of the harmonized metric and incentive in the qth interval respectively, wherein $S_0 = 0$ .. Within this examination, the condition domain is partitioned into a sequence of distinct masses, articulated as $S = \{0, 0.05, 0.1, \dots, 0.95, 1\}$. The execution of the enhancement education collective is illustrated beneath. The regulator can emit the likelihood P_i of a operation A_i. Based on P_i, the forecasted quantity W_i that might be acquired. Subsequently transmits the forecasted quantity into the succeeding temporal phase to forecast the quantity of W_{i+1} Upon acquiring W_{6N}, the consolidation outcome might be derived based on Eq. (16). The incentive might be computed based on Eq. (17). To locate ideal masses, the anticipated incentive necessitates optimization. Consequently, the goal function is articulated as Eq. (18):

$$L(\theta_C) = E_{P(A_1:A_{6N};\theta_C)}[R] \qquad (18)$$

Subsequently, the strategy slope technique shall be executed to refresh the factors of the regulator, computed as Eq. (19):

$$\nabla_{\theta_C} L(\theta_C) = E_{P(A_1:A_{6N};\theta_C)}[\nabla_{\theta_C} J_C(\theta_C) R] \qquad (19)$$

Ordinarily, Eq. (20) is utilized to estimate Eq. (19) to compute the inclination.

$$\nabla_{\theta_C} L(\theta_C) = \sum_{q=1}^{Q} \nabla_{\theta_C} J_C(\theta_C) R_q \qquad (20)$$

wherein Q denotes the count of collective sub-models within a batch. The preceding formula constitutes an impartial estimation of the inclination. This estimation exhibits significant variability. Consequently, to diminish this variability, the baseline function S' is incorporated into the evaluation, as depicted in Eqs. (21)-(23):

$$\nabla_{\theta_t} L(\theta_C) = \sum_{i=1}^{o} \nabla_e J_C(\theta_{(\theta)}) R'_a \qquad (21)$$

$$R'_q = (S_q - S') \tag{22}$$
$$S' = \gamma S' + (1 - \gamma S_{q-1}) \tag{23}$$

The S_q depends solely on the operations prior to the present phase, ensuring that Eq. (21) stays an impartial estimation of the inclination. Ultimately, based on the acquired inclination, the factors of the regulator might be refreshed iteratively as illustrated in Eq. (24):

$$\theta_C = \theta_C - \eta \nabla_{\theta_C} L(\theta_C) \tag{24}$$

The REDTLN structure is exhibited in Fig. 5. Arrange N origin realm data and destination realm data into N pairs of individual origin-individual destination data collections. 6N DPLMs are established using diverse kernel MMDs and pre-conditioned using N origin realm data. Feed each pair of data into the corresponding DPLM to attain individual origin-individual destination realm adaptation. Ultimately, a multi-origin multi-model enhancement collective education model is formulated.
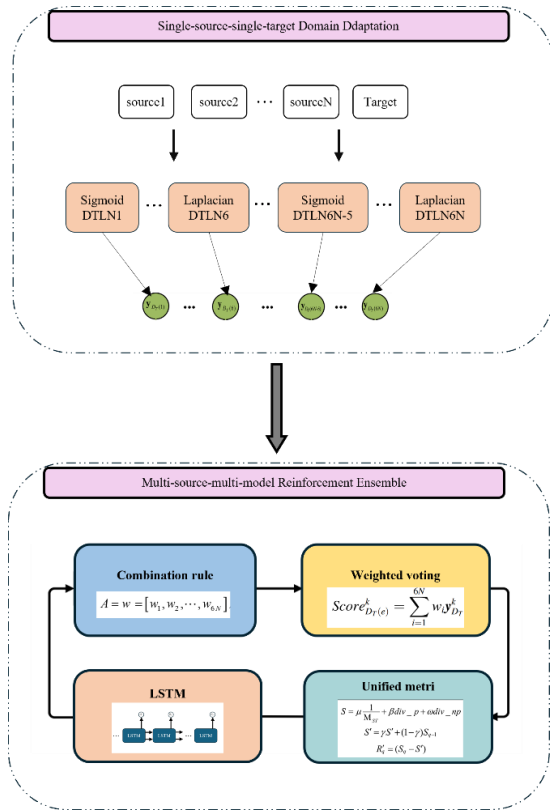

Fig. 5. REDTLN architecture

3. EXPERIMENTAL ANALYSIS

3.1. MFD-REDTLN model verification analysis

Within this chapter, numerous knowledge migration exercises involving various origin areas are designed using multiple datasets. These tasks are designed to demonstrate the efficacy of the presented solution. This paper applies three data sets to confirm the effectiveness of the proposed approach.

The voiceprint data collection 1 (D1), data collection 2 (D2), and data collection 3 (D3) of power transformers under different environmental working conditions were selected respectively. In this study, D1, D2, and D3 each consist of six failure modes, which are partial discharge (PD), secondary circuit, short-circuit impact, loose components, severe overload, and cooler noise. Each failure framework comprises 120,000 information records, which are arranged into 100 examples; each example includes 1200 data points.
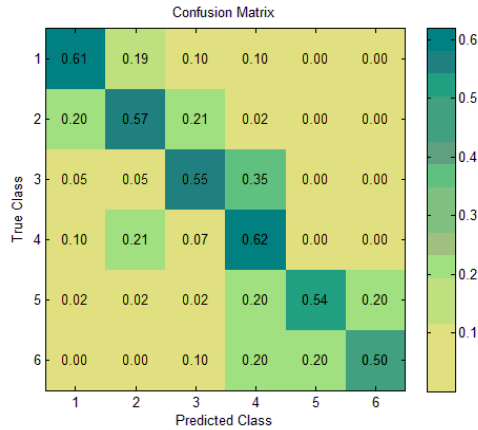
Based on the multi-source domain fault identification experiments designed in this study, data partitioning strictly adheres to the domain isolation principle to prevent data leakage. For each experimental scenario, the source domain data is merged and uniformly divided as follows: Training set: 70% – Used for training the feature extractor and classifier of the DTLN network. Validation set: 15% – Employed for hyperparameter tuning and early stopping strategies. Test set: 15% – Utilized to evaluate model generalization capability. The entire target domain dataset exclusively serves as the test set for the final assessment of the REDTLN model's cross-domain diagnostic performance. This partitioning ensures that the model accesses only source domain data during training and validation, while the target domain data remains completely independent of the training process. Furthermore, the source and target domains are completely mutually exclusive in terms of specific transformer units or operating conditions, with no data overlap existing between them.

By adjusting equipment load conditions-including no-load, rated-load, and overload states - we simulated operational processes under varying working conditions. Simultaneously, background noise sources were introduced, such as electromagnetic whine from motors, cavitation noise phenomena in fluid pipelines, and ambient acoustic interference. During the data collection phase, high signal-to-noise ratio (SNR) microphone arrays were employed to enhance the signal fidelity.
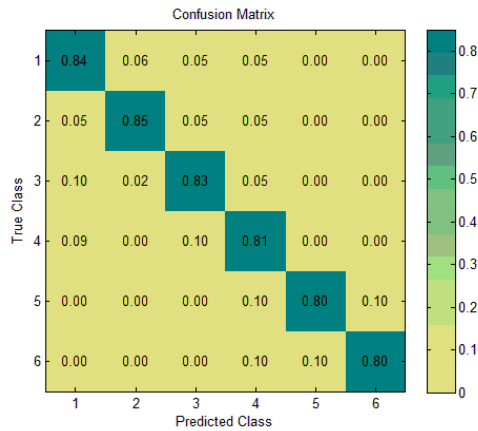
Three multi-origin zone defect identification experiments, namely (D1, D2)-D3, (D1, D3)-D2 and (D2, D3)-D1, are designed using data collections D1, D2 and D3 to validate the efficacy of the suggested approach. Effectiveness. Taking (D1, D2)-D3 for instance, this involves implementing varied-region adjustment from origin zones D1 and D2 to destination zone D3. During the pre-training procedure of DTLN, the number of cores m of ConvM1 is selected from 10 to 20, the corresponding training times are selected as 10, and the education pace is fixed at 0.001. Throughout the process of uni-origin uni-destination area adjustment training, the training times of each DTLN are established from 10 to 30, the education pace is fixed from 0.0001 to 0.005, and α is established from 0.05 to 2.5. Within the diverse-input varied-framework reinforcement combination procedure, the count of LSTM cells within the regulator is chosen from 20 to 100, λ is fixed to 0.001, the education pace is fixed at 0.001, and Adam is used to improve the regulator, and γ is

DIAGNOSTYKA, Vol. 26, No. 3 (2025)
Mu M, Niu J, Guo G, Tian M, Li J, Li J. *Reinforcement-enhanced deep transfer fusion network for ...*

7

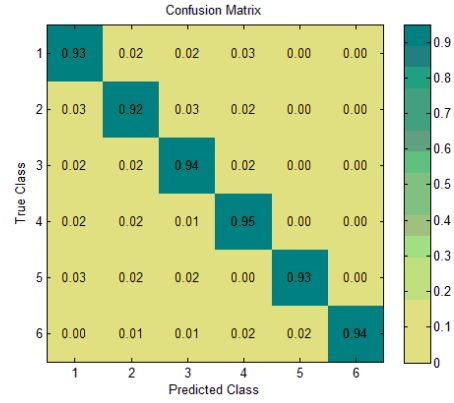picked from 0.5 to 0.5. 0.95, and the number of iterations is fixed at 100.

A single DTLN achieves uni-origin uni-destination area adjustment, while MFD-REDTLN achieves multi-source and multi-model enhanced integration. Separate DTLNs leveraging diverse kMMDs display varied effectiveness, several notably distinct. Upon introducing an isolated DTLN to various initial regions, it similarly demonstrates differing knowledge migration efficiencies. During trials, MFD-REDTLN surpasses lone DTLN. Regarding destination region D1 paired with starting region D2, DTLN1 reached the top accuracy (61.18%); utilizing D3 as baseline area info, DTLN8 attained maximum precision (84.67%). Employing D2 and D3 as base field details, MFD-REDTLN performance stands at 99.50%, significantly exceeding all standalone DTLNs. The confusion matrix for the superior DTLN and MFD-REDTLN across diverse settings is depicted in Fig. 6. Through multi-input multi-framework reinforced collective learning, MFD-REDTLN attains greater accuracy compared to the leading DTLN in each malfunction condition. These results highlight the efficiency of the introduced technique.

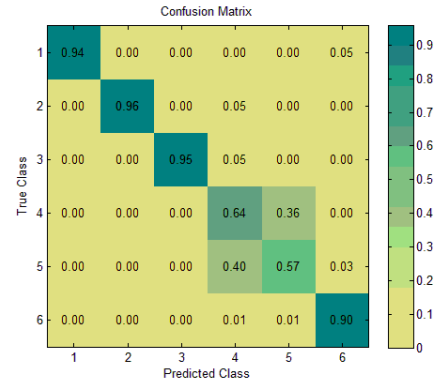a) DTLN1 confusion matrix with source field D2

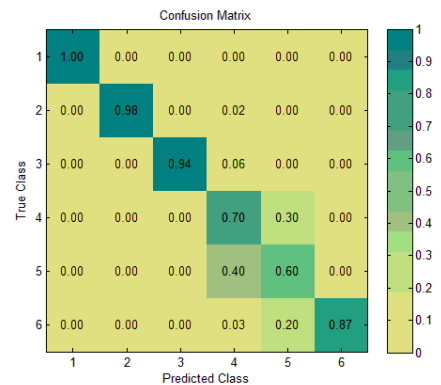b) DTLN8 confusion matrix with source field D3

c) MFD-REDTLN confusion matrix
Fig. 6. Error classification tables for best-performing DTLN and MFD-REDTLN across varied situations within destination area D1
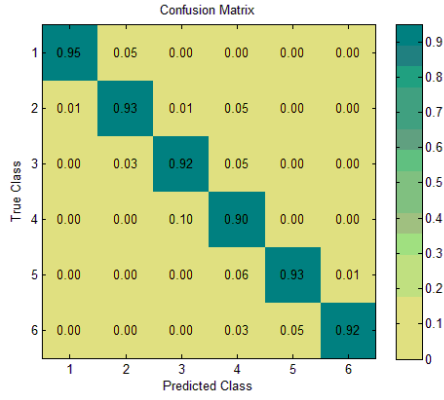
Likewise, for the target zone D2 and origin zone D1, DTLN1 reached the top reliability (82.67%). When D3 serves as source field data, DTLN7 obtains the highest precision (84.83%). Upon utilizing D2 and D3 as origin area information, the precision of MFD-REDTLN reaches 94.83%, marking it as the peak performance among all individual DTLNs. The error classification table for the top-performing DTLN and MFD-REDTLN across varied scenarios is shown in Fig. 7. MFD-REDTLN mostly achieves superior accuracy compared to the top DTLN in each failure mode. These results further confirm the effectiveness of the suggested approach.

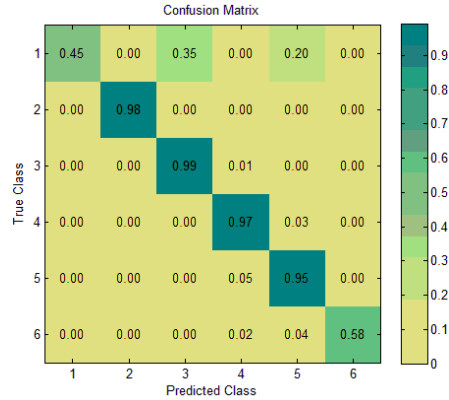
a) DTLN1 confusion matrix with source field D1

b) DTLN7 confusion matrix with source field D2

DIAGNOSTYKA, Vol. 26, No. 3 (2025)
Mu M, Niu J, Guo G, Tian M, Li J, Li J. *Reinforcement-enhanced deep transfer fusion network for …*
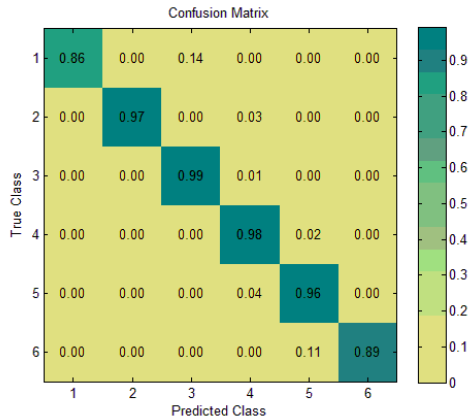
8

c) MFD-REDTLN confusion matrix
Fig. 7. Error classification tables for top-performing
DTLN and MFD-REDTLN across various settings
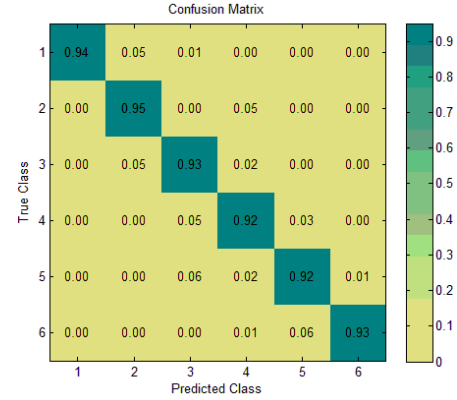within destination region D2

For target zone D3, DTLN3 reached the top accuracy (82.00%) when D1 served as source domain data. When D3 functions as source field data, DTLN11 achieves the peak precision (94.16%). Upon applying D2 and D3 as origin area information, the precision of MFD-REDTLN reaches 98.67%, marking it as the peak performance among all individual DTLNs. The confusion matrix of the optimal DTLN and MFD-REDTLN in different scenarios is shown in Fig. 8.

a) DTLN3 confusion matrix with source field D1

b) DTLN11 confusion matrix with source field D3

c) MFD-REDTLN confusion matrix
Fig. 8. Error classification tables for top-performing
DTLN and MFD-REDTLN across various settings
within destination region D3

These findings show that the effectiveness of separate DTLNs using various kMMD and origin fields differs greatly and that they lack sufficient accuracy and stability. There is presently no reliable approach to assist in selecting appropriate kMMD and origin fields. MFD-REDTLN combines single-origin single-destination zone adaptation with multi-origin multi-paradigm enhanced integration. It eliminates the necessity to choose a particular kMMD or origin field, but typically delivers better outcomes than an individual DTLN. This underscores the efficiency of MFD-REDTLN.

3.2. Analysis of the search process of reinforcement learning

Fig. 9, Fig. 10 and Fig. 11 illustrate the unified metric and related ensemble learning accuracy as a function of the iteration count for experiments (D2, D3)-D1, (D1, D3)-D2, and (D1, D2)-D3 accordingly. chart. The x-axis indicates the cycle number, the left section shows the standardized measure, scaled by dividing by its peak value, and the right panel displays reinforcement learning ensemble accuracy.
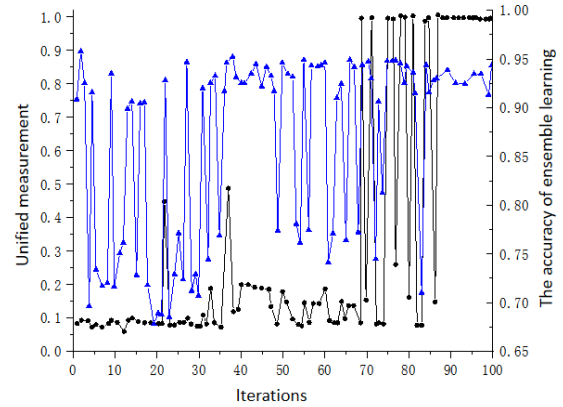

Fig. 9. Changes of the unified metric and ensemble learning accuracy curves with the number of iterations in Experiment (D2, D3)-D1
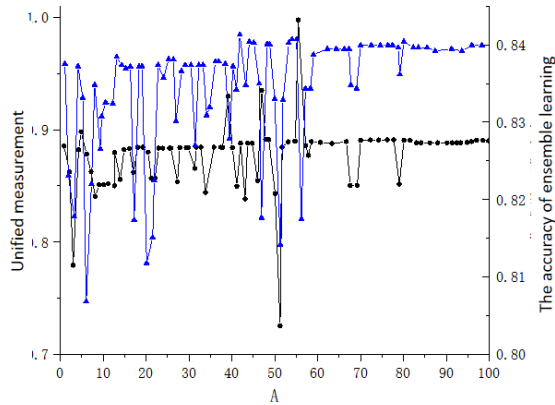
Fig. 10. Experiment (D1, D3)-Unified metric
and ensemble learning accuracy curves in D2
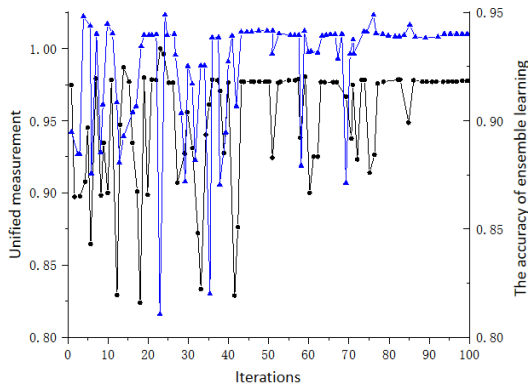change with the number of iterations

Fig. 11. Experiment (D1, D2)-Unified metric
and ensemble learning accuracy curves in D3
change with the number of iterations

Test outcomes indicate that the integrated metric generally stabilizes with the number of iterations. The unified measure curve and the ensemble learning precision curve exhibit good alignment. Hence, the integrated measure can serve as an alternative performance indicator to assess collective learning effectiveness without correct data. It is fitting to use unified metrics as rewards for enhancement learning models to achieve multi-origin and multi-paradigm integration. Nevertheless, the final integrated measure and ensemble learning effectiveness determined through reinforced learning models are occasionally not globally optimal. This signifies an inherent limitation of reinforced learning. Not all cycle values of the integrated measure accurately correspond to the ensemble learning performance metrics. The reason is that the label information from the destination region is not employed in crafting the integrated measure, and the relationship between the diversity indicator and ensemble learning outcomes is not entirely accurate.

3.3. Model comparison analysis

MFD-REDTLN is contrasted with several traditional single transfer education techniques,

encompassing CNN, DDC, D-CORAL, DANN, TCA and JDA. CNN is a genuine profound network with identical structure and hyperparameters as singular DTLN, yet does not employ MMD. DDC is a realm adaptation network akin to DTLN2, but solely utilizes Gaussian kernel MMD to adapt to the characteristics of FC2. D-CORAL also possesses an identical structure as DTLN but employs CORAL loss to adapt the characteristics of FC1 and FC2. DANN has an identical structure as DTLN, but includes an additional realm classifier for realm adaptation; the characteristics of FC are fed into the realm classifier. TCA and JDA are shallow realm adaptation methods that utilize Gaussian kernel MMD. The experimental outcomes are exhibited in Table 1, Table 2 and Table 3.

Table 1. Comparison of MFD-REDTLN and other models in different scenarios in target domain D1

Model	Source domain - target domain	Accuracy±standard deviation/%
DTLN1	D2-D1	61.18±1.07
CNN	D2-D1	59.24±5.28
DDC	D2-D1	52.54±4.16
D-CORAL	D2-D1	53.15±4.02
DANN	D2-D1	57.62±5.95
TCA	D2-D1	32.54
JDA	D2-D1	48.75
DTLN8	D2-D1	84.67±1.05
CNN	D3-D1	78.21±3.78
DDC	D3-D1	69.54±4.68
D-CORAL	D3-D1	72.14±5.76
DANN	D3-D1	68.76±4.23
TCA	D3-D1	38.21
JDA	D3-D1	39.54
MFD-REDTLN	(D2,D3)-D1	95.21±0.48

Profound transfer education models DDC, D-CORAL, and DANN generally exhibit superior performance compared to shallow models TCA and JDA, and within these experiments, MFD-REDTLN on average achieves higher precision outcomes than existing classical transfer education models. Specifically, for the destination realm D1, the accuracy of DTLN1 utilizing origin realm data D2 is 61.18%, and the accuracy of DTLN8 utilizing origin realm data D3 is 84.67%, both of which are the highest relative to the single model, when using origin realm data D2 the variability of DDC, D-CORAL and DANN all exceed 3.5%. The variabilities of DDC, D-CORAL and DANN using origin realm data D3 are 3.78%, 4.68%, 5.76% and 4.23% respectively, indicating their stability. The accuracy is inadequate. The accuracy of MFD-REDTLN is 95.21%, the variability is merely 4.77%, and the effect is consistent when the destination realm is D2 and D3. This demonstrates that it not only identifies failure patterns more accurately but also exhibits greater stability.

Therefore, MFD-REDTLN performs superiorly. Within these experiments, the proposed MFD-REDTLN not only identifies malfunctions more

accurately but also exhibits greater robustness compared to the individual transfer education model. This further illustrates the superiority of MFD-REDTLN.

Table 2. Comparison between MFD-REDTLN and other models in different scenarios in target domain D2

Model	Source domain - target domain	Accuracy±standard deviation/%
DTLN1	D1-D2	82.67±1.21
CNN	D1-D2	75.12±4.21
DDC	D1-D2	76.25±3.98
D-CORAL	D1-D2	58.24±5.67
DANN	D1-D2	68.36±4.68
TCA	D1-D2	48.21
JDA	D1-D2	57.36
DTLN7	D1-D2	84.83±1.09
CNN	D3-D2	72.65±2.98
DDC	D3-D2	57.24±3.45
D-CORAL	D3-D2	69.78±3.65
DANN	D3-D2	65.32±2.98
TCA	D3-D2	43.26
JDA	D3-D2	59.34
MFD-REDTLN	(D1, D3)-D2	94.83±0.59

Table 3. Comparison between MFD-REDTLN and other models in different scenarios in target domain D3

Model	Source domain - target domain	Accuracy±standard deviation/%
DTLN3	D1-D3	82.00±1.49
CNN	D1-D3	68.42±2.69
DDC	D1-D3	75.62±7.85
D-CORAL	D1-D3	65.32±4.63
DANN	D1-D3	70.14±5.36
TCA	D1-D3	42.12
JDA	D1-D3	57.36
DTLN11	D1-D3	91.16±1.36
CNN	D2-D3	57.62±2.98
DDC	D2-D3	68.24±3.47
D-CORAL	D2-D3	78.21±2.89
DANN	D2-D3	65.47±3.36
TCA	D2-D3	45.36
JDA	D2-D3	52.14
MFD-REDTLN	(D1, D2)-D3	94.67±0.75

4. CONCLUSIONS

This document presents MFD-REDTLNfor fault-diverse area voiceprint recognition. Its core concepts include three primary components: origin-focused goal-oriented area adjustment, unified metric design for unsupervised collective learning, and diversified framework-reinforced integration. Initially, various novel kMMD are employed to construct several DTLN for origin-specific destination-oriented area modification, which can render the DTLN transferable attributes with diverse traits. Next, a novel unified unsupervised group learning is crafted based on MMD and diversity indicators, and by employing this unified metric as incentive, a reinforcement education technique is applied constructed to search for effective

combinations of these DTLNs Rules to achieve enhanced integration of multi-source deep transfer learning networks. Improved accuracy and robustness in dynamic noise environments and complex operating conditions. Test results indicate that the proposed framework can quickly identify the fault operating status of the transformer and diagnose six typical faults of the transformer in different noise environments. The diagnosis accuracy reaches more than 95%, which improves the model's general performance under different working conditions. ization ability. And compared with the classic individual transfer learning methods of DTLN, DDC, D-CORAL, DANN, TCA and JDA, the results show that MFD-REDTLN performs better and the recognition accuracy is above 95%. Experimental results show that the MFD-REDTLN model can not only identify faults more accurately and improve the accuracy of fault diagnosis, but is also more robust than other transfer learning models.

Author contributions: *Mingliang Mu: mathematical model and the simulation techniques. Jianfei Niu: spelling and grammar checking as well as virtual validation, Guangqian Guo: Software, Maoxiang Tian: Writing – Original Draft Preparation, Jinlin Li: Resources, Jingyi Li: experimental validation.*

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

REFERENCES

1. Tao LY, Yang XH, Zhou XC, et al. A novel transformers fault diagnosis method based on probabilistic neural network and bio-inspired optimizer. Sensors. 2021;21(11):3623. https://doi.org/10.3390/s21113623.
2. Abbasi AR, Mohammad RM, Mohammad MA. Transformer winding faults detection based on time series analysis. IEEE T. Instrum. Meas. 2021;70:1–10. https://doi.org/10.1109/TIM.2021.3076835.
3. Liu JN, Yao CG, Yu L, et al. Using MLP to locate transformer winding fault based on digital twin. Front. Energy Res. 2023;11:1175808. https://doi.org/10.3389/fenrg.2023.1175808.
4. Zeng B, Guo J, Zhu WQ, et al. A transformer fault diagnosis model based on hybrid grey wolf optimizer and LS-SVM. Energies. 2019;12(21):4170. https://doi.org/10.3390/en12214170.
5. Peng XS, Yang F, Wang GJ, et al. A convolutional neural network-based deep learning methodology for recognition of partial discharge patterns from high-voltage cables. IEEE T. Power Deliver. 2019;34(4): 1460–1469. https://doi.org/10.1109/TPWRD.2019.2906086.
6. Besharatifard H, Hasanzadeh S, Heydarian-Forushani E, et al. Acoustic based localization of partial discharge inside oil-fflled transformers, IEEE Access. 2022;10:55288–55297. https://doi.org/10.1109/ACCESS.2022.3177603.

7. Wu Y, Sun X, Zhang Y, et al. A power transformer fault diagnosis method-based hybrid improved seagull optimization algorithm and support vector machine. IEEE Access. 2022;10:17268–17286. https://doi.org/10.1109/ACCESS.2021.3127164.

8. Rodriguez-Serna JM, Albarracin-Sanchez R, Garnacho F, et al. Partial discharges measurements for condition monitoring and diagnosis of power transformers: A review. Proc. 6th Int. Adv. Res. Workshop Transformers (ARWtr). 2019:83–88. https://doi.org/10.23919/ARWtr.2019.8930183.

9. Liu M, Wang P, Zhou Q, et al. Remote fault diagnosis system for power transformer insulation based on RIA model. Proc. IEEE Int. Conf. High Voltage Eng. Appl. (ICHVE). 2016:1–4. https://doi.org/10.1109/ICHVE.2016.7800679.

10. Hou L, Bergmann NW. Novel industrial wireless sensor networks for machine condition monitoring and fault diagnosis, IEEE Trans. Instrum. Meas. 2012;61(10):2787–2798. https://doi.org/10.1109/TIM.2012.2200817.

11. Bagheri M, Naderi MS, Blackburn T. Advanced transformer winding deformation diagnosis: Moving from off-line to on-line, IEEE Trans. Dielectr. Electr. Insul. 2012;19(6):1860–18702012. https://doi.org/10.1109/TDEI.2012.6396941.

12. Portilla WH, Mayor GA, Guerra JP, et al. Detection of transformer faults using frequency-response traces in the low-frequency bandwidth, IEEE Trans. Ind. Electron. 2014;61(9):4971–4978. https://doi.org/10.1109/TIE.2013.2282605.

13. Guo H, Hou B, Jing Y, et al. Method for balancing short circuit impedance of multi-split rectiffer transformer, in Proc. IEEE Int. Conf. Appl. Supercond. Electromagn. Devices (ASEMD). 2020:1–2. https://doi.org/10.1109/ASEMD49065.2020.9276100.

14. Moradzadeh A, Pourhossein K, Mohammadi-Ivatloo B, et al. Locating inter-turn faults in transformer windings using isometric feature mapping of frequency response traces. IEEE Trans. Ind. Informat. 2021;17(1):6962–69702021. https://doi.org/10.1109/TII.2020.3016966.

15. Nurmanova V, Bagheri M, Zollanvari A, et al. A new transformer FRA measurement technique to reach smart interpretation for inter-disk faults. IEEE Trans. Power Del. 2019;34(4):1508–1519. https://doi.org/10.1109/TPWRD.2019.2909144.

16. Gomez-Luna E, Aponte Mayor G, Gonzalez-Garcia C, et al. Current status and future trends in frequency-response analysis with a transformer in service. IEEE Trans. Power Del. 2013;28(2):1024–1031. https://doi.org/10.1109/TPWRD.2012.2226495

17. Tenbohlen S, Beura CP, Siegel M. Condition assessment of power transformers by UHF PD measurements, in Proc. 9th Int. Conf. Condition Monitor. Diagnosis (CMD). 2022:5–10. https://doi.org/10.23919/CMD54214.2022.9991308.

18. Liao R, Guo C, Wang K, et al. Adaptive optimal kernel time-frequency representation technique for partial discharge ultra-high-frequency signals classiffcation, Electric Power Compon. Syst. 2015;43(4):449–460. https://doi.org/10.1080/15325008.2014.986775.

19. Vu-Cong T, Dalstein M, Toigo C, et al. 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. Faiz J, Soleimani M. Assessment of computational intelligence and conventional dissolved gas analysis methods for transformer fault diagnosis, IEEE Trans. Dielectr. Electr. Insul. 2018;25(5):1798–1806. https://doi.org/10.1109/TDEI.2018.007191.

21. Bacha K, Souahlia S, Gossa M. Power transformer fault diagnosis based on dissolved gas analysis by support vector machine, Electric Power Syst. Res. 2012;83(1):73–79. https://doi.org/10.1016/j.epsr.2011.09.012.

22. Yan R, Gao RX. Hilbert–Huang transform-based vibration signal analysis for machine health monitoring, IEEE Trans. Instrum. Meas. 2006;55(6):2320–2329. https://doi.org/10.1109/TIM.2006.887042.

23. Li R, He D. Rotational machine health monitoring and fault detection using EMD-based acoustic emission feature quantiffcation, IEEE Trans. Instrum. Meas. 2012;61(4):990–1001. https://doi.org/10.1109/TIM.2011.2179819.

24. Sun M, Bai X, Zhang W, et al. Jointed task of multi-scale CNN based power transformer fault diagnosis with vibration and sound signals, in Proc. Panda Forum Power Energy (PandaFPE). 2023:1761–1765. https://doi.org/10.1109/PandaFPE57779.2023.10140472

25. Dang XJ, Wang FH, Zhang X. Cochlear fflter cepstral coefffcients of acoustic signals for mechanical faults identiffcation of power transformer, in Proc. IEEE Sustain. Power Energy Conf. (iSPEC). 2019:1195–1200. https://doi.org/10.1109/iSPEC48194.2019.8975291.

26. Han S, Wan X, Lai Y, et al. DC bias diagnosis of power transformer with soundprint features, in Proc. 7th Int. Conf. Cyber Secur. Inf. Eng. 2022:213–217. https://doi.org/10.1145/3558819.3565080.

27. Li X, Liao X, Lu D, et al. A robust classification method for power transformer status recognition based on sound signals, in Proc. 5th Int. Conf. Power Energy Eng. (ICPEE). 2021:153–157. https://doi.org/10.1109/ICPEE54380.2021.9662622.

28. Stone GC, Stranges MK, Dunn DG. Common questions on partial discharge testing: A review of recent developments in IEEE and IEC standards for offline and online testing of motor and generator stator windings, IEEE Ind. Appl. Mag. 2016;22(1):14–19 https://doi.org/10.1109/MIAS.2015.2458337.

Mingliang MU obtained a Master's degree in Engineering from the School of Information and Electrical Engineering at China University of Mining and Technology in Xuzhou, China in 2015. Currently, I work at Binzhou Power Supply Company of State Grid Shandong Electric Power Company, with research interests in substation operation and maintenance, substation equipment operation, and smart grid.
e-mail: bzgdmm88@163.com

Jianfei NIU working at Binzhou Power Supply Company of State Grid Shandong Electric Power Company. I was born in May 1983 and hold a bachelor's degree. I am a senior engineer from Weifang, Shandong, and my research focuses on power system and substation operation and maintenance.
e-mail: jianfeiNiu12@163.com

Guangqian GUO working at Binzhou Power Supply Company of State Grid Shandong Electric Power Company, born in April 1992. I hold a bachelor's degree and am an intermediate engineer from Weifang, Shandong.
e-mail: GuangqianGuo888@126.com

Maoxiang TIAN working at State Grid Shandong Electric Power Company Binzhou Power Supply Company, born in May 1987, bachelor's degree, senior engineer, born in Binzhou, Shandong.
e-mail: tmxwenwen11@126.com

Jinlin LI working at State Grid Shandong Electric Power Company Binzhou Power Supply Company, born in March 1971, bachelor's degree, intermediate engineer, born in Binzhou, Shandong, specializing in power system and substation operation and maintenance.
e-mail: jinlinlixx@163.com

Jingyi LI working at State Grid Shandong Electric Power Company Binzhou Power Supply Company, born in April 2000 in Binzhou, Shandong, holds a bachelor's degree and is an intermediate engineer.
e-mail: lijingyi200602@163.com