Parallel Relation Network for Intelligent Fault Detection and Localization of Train Transmission Systems with Zero-fault Sample

Zhixu Duan
School of Mechanical and Electrical Engineering
University of Electronic Science and Technology of China
Chengdu, China
2023040906024@uestc.edu.cn

Ruoxi Liu
School of Mechanical and Electrical Engineering
University of Electronic Science and Technology of China
Chengdu, China
2023040903004@uestc.edu.cn

Zuoyi Chen
School of Mechanical and Electrical Engineering
University of Electronic Science and Technology of China
Chengdu, China
zuoyichen@uestc.edu.cn

Hong-Zhong Huang
School of Mechanical and Electrical Engineering
University of Electronic Science and Technology of China
Chengdu, China
hzhuang@uestc.edu.cn

***Abstract*—In recent years, deep learning methods have empowered the fault detection of train transmission systems (TTS). However, the existing methods require a large number of fault samples, which are often difficult to obtain in TTS. In order to address the challenge, we propose a parallel relation network (PRN) for fault detection of TTS. The PRN method uses residual shrinkage network to extract the effective features from input samples and then form feature pairs to describe the relationship between the samples. Finally, Kolmogorov-Arnold networks (KAN) is used to assess the health states of the samples. In addition, we generate different auxiliary sample libraries to enhance the model's ability to extract features inherent to the health state. In experiments, we verified the feasibility of the PRN method, which achieved more than 98% accuracy and recall in different fault states of TTS.**

Keywords-fault detection; train transmission systems; residual shrinkage network; Kolmogorov-Arnold networks; relation network

I. INTRODUCTION

The train transmission System (TTS) plays an important role in the proper operation of a train, and have a direct impact on a train's performance, efficiency and reliability. TTS often works in harsh and challenging environments characterized by high temperatures, high humidity and dusty tunnel interiors. As a result, the risk of unforeseen system failures increases. Therefore, fault detection techniques are instrumental in ensuring the reliability and safety of TTS operations.

Recently, deep learning (DL) is widely applied for fault detection due to its capacity of autonomously extracting pertinent information from input samples [1-3]. A significant advantage of the fault detection methods based on DL is that they have low requirements on the fault occurrence mechanism, and they can be effective as long as the amount of data is sufficient. Hence, these methods have garnered considerable interest due to the growing availability of data [4-7]. However, an obvious disadvantage of DL-based fault detection methods is that they require a large amount of fault data. TTS operate in a healthy state the vast majority of the time because they are regularly maintained. Thus, obtaining fault samples are extremely difficult. This severely limits the application of DL fault detection methods to TTS.

To address the above issue, research scholars have attempted to apply transfer learning (TL) to the fault detection of mechanical equipment. TL trains the model using the source domain data and transfers the model to the target domain data. The fault data in the source domain is easy to obtain, which overcomes the limitation that fault samples in the target domain are difficult to obtain [8-10]. Guo et al. [11] combined TL with a domain classifier, which significantly improved the fault detection accuracy of certain mechanical equipment. Although TL-based fault detection methods have showed excellent results in the case of insufficient fault samples, they exhibit the following limitations.

1) TL-based fault detection methods require high correlation of fault data in the source and target domains. However, in practical application environments, it is often difficult to obtain such demanding source domain data.

2) While these methods do not need the labels of fault samples in the target domain, the fault data in the source domain must have the same fault types as the data in the target domain. So, this method is not essentially a direct solution to the lack of fault samples, but a transformation of the domain.

3) Existing fault detection methods can just determine whether the system is malfunctioning or not, but cannot further locate where the fault occurred.

To address these challenges, we propose a novel method called the parallel relation network (PRN) for fault detection and localization of TTS with zero-fault samples. In this method, we first utilize residual shrinkage networks (RSN) to reduce the noise in the input signal and mine meaningful features. Then, feature pairs are created to capture the relationship of the health and other states. Finally, Kolmogorov-Arnold relation networks (KARN) are employed to assess relations among these pairs, facilitating type evaluation. In addition, to enhance the robustness and reliability of health feature extraction, we constructed an auxiliary sample library (ASL). This PRN method effectively detects and locates faults in the system's state-to-state configuration even there are no faulty samples, which reduces the reliance on existing detection methods' dependency on fault data.

The structure of this paper is as follows. Section I serves as an introduction. Section II provides a comprehensive description of the proposed method, PRN. Section III shows the experimental outcomes and associated discussions.

II. METHODOLOGY

A. Overview of the proposed method

For more accurate detection and precise positioning, we constructed three parallel models to detect faults at each of the three components of the TTS. The PRN method aims to achieve fault detection and positioning by acquiring data from healthy instance samples, denoted as D_H. It is presumed that the TTS encompasses three pivotal monitoring sites, each equipped with respective sensors to track their status. The dataset for training encompasses the healthy instance samples D_H of the TTS and the ASL D_A. The D_H (D_H^1, D_H^2, D_H^3) consist of data gathered from three distinct sensors. The ASL D_A consists of two parts of data. One part of ASL comes from other devices' fault data by Paderborn University [12], in which $D_d^1, D_d^2, \dots D_d^n$ indicate n different fault types. This part of ASL will be used in all components. The other part of the sample library was generated by the soft Brownian offset method [13] based on the three components corresponding to each of them. D_o^1, D_o^2, D_o^3 represent the sample libraries generated according to different components. In the training stage, the different components only need to be trained using their respective generated D_o^n. We showed the architecture of PRN in Fig.1, which consists training and testing stages.

During the training stage, through CWT (continuous wavelet transform), health samples D_H are change into time frequency images. Images are inputted into the module RSN θ to mine useful features in three channels for different components. RSN outputs both inputs D_H and D_A as features $\tilde{g}_h(\tilde{g}_h^1, \tilde{g}_h^2, \tilde{g}_h^3)$ and $\tilde{g}_a(\tilde{g}_a^1, \tilde{g}_a^2, \tilde{g}_a^3)$. Next, health and anomaly samples $\tilde{g}_a^n$ and $\tilde{g}_h^n$ are spliced into feature pairs g_a^n to obtain unique characteristics of health state. Healthy features are also spliced into g_h^n to describe common characteristics of health state. At last, through the KARN G_θ to which our feature pairs transmit, we get relation scores matrix R_P^1, R_P^2, R_P^3 in the range of 0 to 1 to describe the similarities between health and fault states. Higher scores mean more relevance. We calculate three losses by aiming for true relation scores R_T^1, R_T^2, R_T^3 of which the value is 0 or 1 on different components. The metric function $F(x, D_H^n, \theta)$ used in training are defined as following:

$$F(x, D_H;\ \theta) = \begin{cases} F(x,\ X_H^1)^1 \\ F(x,\ X_H^2)^2 \\ F(x,\ X_H^3)^3 \end{cases} \rightarrow \begin{cases} r_{x,X_H^1} \\ r_{x,X_H^2} \\ r_{x,X_H^3} \end{cases} \rightarrow \begin{cases} (C_H\ or\ C_A)^1 \\ (C_H\ or\ C_A)^2 \\ (C_H\ or\ C_A)^3 \end{cases} \quad (1)$$

where x and X_H^n form the feature pairs. θ is the parameter to be learned in the PRN. $(C_H\ or\ C_A)^n$ indicates the health state of the n-th component. C_H is health state and C_A is fault state.

Then, in the testing stage, historical health data and online monitoring data $D_{online}(D_{online}^1, D_{online}^2, D_{online}^3)$ from three sensors are used to evaluate models' performance. With the three models already trained, we obtain relation scores $(r_{X_{online}^n, X_H^n})$ for each online monitoring sample between online and historical health samples. The type of sample is determined by health boundary score (HB) as following expressions:

$$\begin{cases} \text{Health} & r_{X_{online}^n, X_H^n} \geq HB \\ \text{Fault} & r_{X_{online}^n, X_H^n} < HB \end{cases} \quad (2)$$

B. PRN structure

The PRN structure comprises of four modules: CWT module, RSN $\boldsymbol{F}_\vartheta$ module, feature concatenation module and KARN $\boldsymbol{G}_\theta$ module.

Firstly, the raw data is converted to 2-dimensional time-frequency images by CWT in order to get multi-scale information about the health feature [14]. The formula of wavelet transform is defined as:

$$WT_f(w, \tau) = |w|^{-1/2} \int_{-\infty}^{+\infty} f(t)\, \varphi\left(\frac{t-\tau}{w}\right) dt \quad (3)$$

where w and τ represent the scale and offset. $f(t)$ denotes the vibration signal, $\varphi(\cdot)$ denotes a wavelet basis function.

Then, the RSN $\boldsymbol{F}_\vartheta$ is designed to extraction the high-quality features by itself from time-frequency images[15]. Unlike ResNets, RSN $\boldsymbol{F}_\vartheta$ adds a soft threshold as a nonlinear adjustment layer. This layer effectively removes noise-related information, thereby extracting meaningful features. The method of the soft threshold is provided by

$$y = \begin{cases} x - \tau & x > \tau \\ 0 & -\tau \leq x \geq \tau \\ x + \tau & x < \tau \end{cases} \quad (4)$$

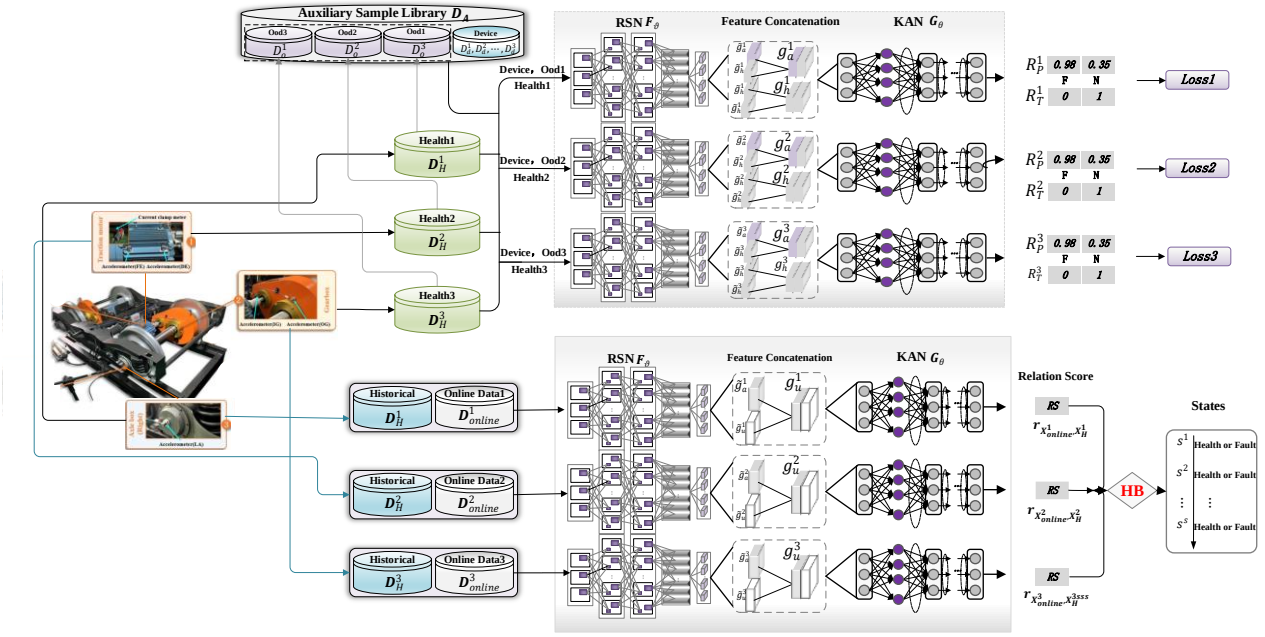

Figure 1. The architecture of the PRN method.

where x represents input features and y represents output features, τ represents the threshold value. The soft threshold ensures that negative features are set to zero, preserving valuable features and attenuating unhelpful ones. This method effectively prevents gradient vanishing and exploding issues, as the derivative of the output is constrained to values of either 1 or 0.

Next, in order to describe the similarity of features between samples, we established the concept of sample pairs. In the feature concatenation module, features from different source samples are concatenated together as sample pairs of which the expression is:

$$\begin{cases} g_a^n = [\tilde{\mathrm{g}}_h^n, \tilde{\mathrm{g}}_a^n] \\ g_h^n = [\tilde{\mathrm{g}}_h^n, \tilde{\mathrm{g}}_h^n] \end{cases} \tag{5}$$

where $[\cdot,\cdot]$ means the operation of concatenation. Health and anomaly samples are used to learn the unique characteristics while health samples are used to learn common characteristics. By analyzing the relationship between sample pairs, we can quantify the similarity between known and unknown samples. This helps to improve the accuracy of our fault detection under zero-sample conditions.

Finally, in the KARN module, we use Kolmogorov-Arnold Networks (KAN) to serve as relation networks. KAN is an innovative neural network architecture that proposes an alternative to the traditional multilayer perceptron (MLP). The core feature of KAN is that it places learnable activation functions at the edges of the network instead of at the traditional nodes, and each weight parameter is replaced by a univariate function, usually parameterized in the form of a spline function. This design not only improves the expressive power of the model, but also enhances its interpretability. KAN demonstrates excellent performance in tasks such as data fitting and partial differential equation solving, outperforming MLPs in accuracy and interpretability. In addition, KAN's neural scaling law shows the advantage of a faster decline of the test loss as the parameter is increased. The KAN's localized spline function also helps to avoid the catastrophic loss that occurs in the process of continuous learning. learning process, where catastrophic forgetting occurs [16]. These features make KAN have great potential for application in scientific research and provide a new direction for the development of fault detection techniques.

III. Experiment and Discussion

A. Dataset

The datasets used in our experiment are provided by the State Key Laboratory of Advanced Rail Autonomous Operation of Beijing Jiaotong University[17]. Fig.2 shows the test bed. The test bed is based on a real metro bogie scaled down and adjusted to a 1:2 ratio. Two types of signals, three-way vibration and three-phase current, were acquired with a sampling frequency of 64 kHz for each channel. The length of each sample is 64000. We selected three components (motor, gearbox, left axle box) to be tested for our model and have three fault types for each location. The detailed fault types are shown in Table Ⅰ. We obtained 2000 samples for each component in health state and 1000 samples for each fault. Every sample included 2048 data points. Training dataset and validation dataset include 800 and 200 health samples in training stage. Testing dataset includes 1000 health samples and 1000 fault samples.

Figure 2. Test bed of train transmission systems.

TABLE I. HEALTH STATE IN COMPONENTS

Component	Label	Health State
Motor	M0	Normal condition
	M1	Short circuit
	M2	Broken rotor bar
	M3	Bearing fault
Gearbox	G0	Normal condition
	G5	Bearing inner race fault
	G6	Bearing outer race fault
	G7	Bearing rolling element fault
Left axle box	LA0	Normal condition
	LA1	Bearing inner race fault
	LA2	Bearing outer race fault
	LA3	Bearing rolling element fault

B. *Results and discussion*

We trained the PRN model by Adam optimizer. The learning rate and batch size were set to 0.001 and 8, respectively.

Fig. 3 visualizes the trend of losses during training. We recorded the loss changes of the three parallel models over 1000 epochs and found that the PRN converged extremely fast, to a very low level around the 100th epoch and has remained low since then. However, at the same time, we also found that despite using parallel models, the training process at different locations showed large differences. The motor has a very smooth training process, but the rest of the components, especially the left axle box, show more severe loss oscillations throughout the training process. This suggests that for different objects, our model still needs to improve its generalization ability to make the training smoother and more reliable.

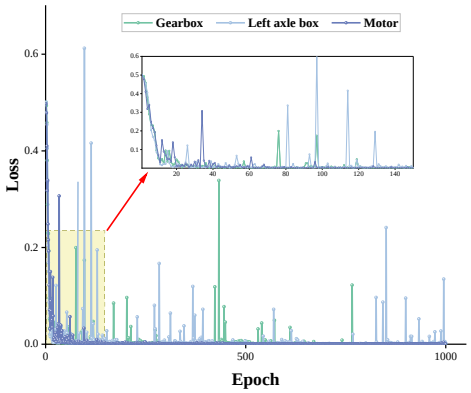

Figure 3. The training process of the PRN model.

Fig.4 shows the test results of the effect of HB on accuracy and standard deviation (Std). The HB ranges from 0.9 to 0.99 and is determined using the validation dataset. Ten trials are conducted for each HB. We find that when the HB reaches 0.98, the accuracy of all three components decreases substantially. Two of the components, the motor and the gearbox, both exhibit over 99% accuracy from 0.9-0.98, while the left axle box component shows a decreasing trend with increasing HB. std overall rises with decreasing accuracy, but it also largely stays at a lower level. Through testing it was concluded that 0.9-0.93 was the optimal range for HB.

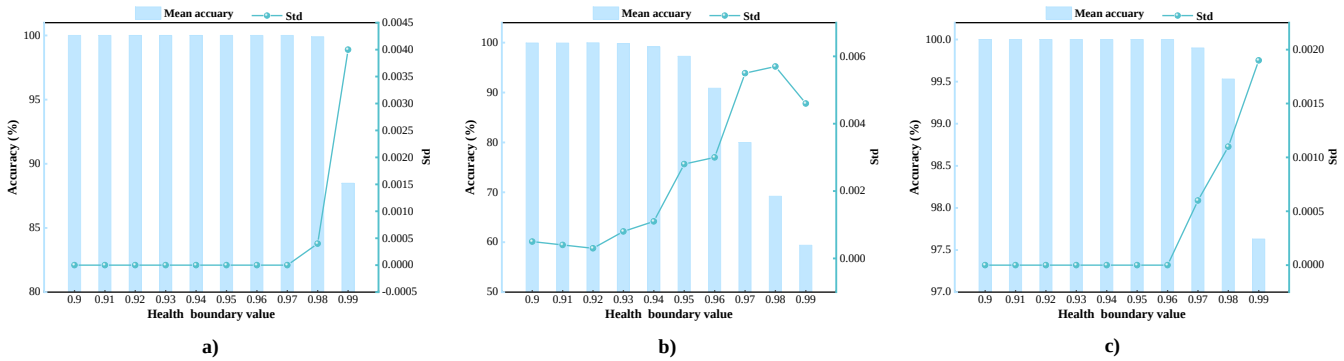

Figure 4. Fault detection accuracy under different HB scores, a) gearbox; b) left axle box: c) motor.

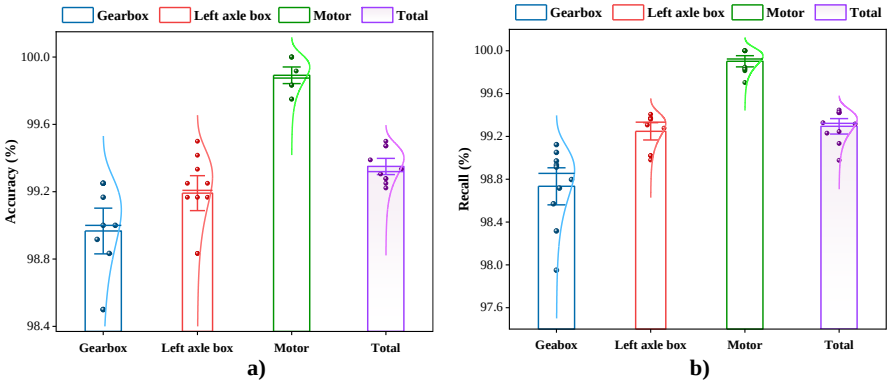

Figure 5. Fault detection results of the RRN. a) accuracy; b) recall.

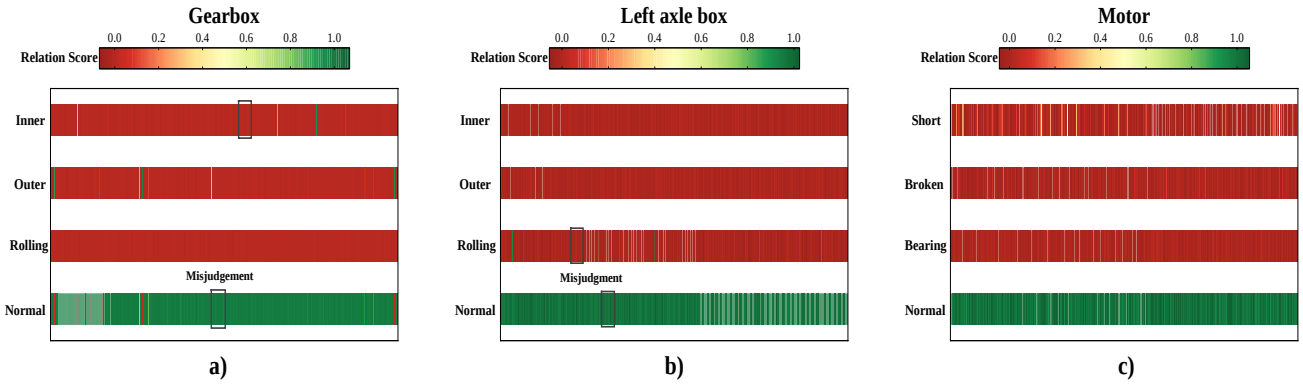

Figure 6. Relation scores for different fault types. a) gearbox; b) left axle box; c) motor.

Both still show surprising results in terms of recall and accuracy. In Fig.5, although the performance of the three positions varies slightly, they all basically stay above the 98% level. The overall accuracy and recall at the motor are even higher at over 99.8%. This also fully demonstrates the feasibility of the PRN model. There is basically no difference in this accuracy and recall images, which also demonstrates the superior performance of the PRN model in any health state.

Fig.6 shows the relation scores output by the three parallel models under different fault types tested. The distribution of similarity scores is evaluated by observing the degree of shading transformation of the colors using color as a judgement criterion. In Fig.6, the bands in each health state are basically solid colors with no change. It clearly shows that the overall performance of the PRN in recognizing all fault types of the three components is superior, which indicates that the proposed PRN achieves excellent test results in TTS fault detection. For example, in the normal state of the gearbox, the relation scores are mostly output near 1, while the misjudgment tend to be near 0. This indicates that this model has a high level of discrimination for features.

IV. Conclusion

In this paper, a novel PRN-based TTS detection method is proposed. In the proposed method, targeted auxiliary sample libraries are constructed for each of the different components. RSN was constructed to extract useful features from health samples and ASL. Sample pairs are constructed to achieve the learning of similar relations. KAN was used as a relation network to determine the health state by outputting relation scores. The goal of the PRN method is to achieve accurate fault detection for the target without fault samples. We designed several detection tasks in the dataset under multiple types of faults in the TTS. Experimental results validate the effectiveness of the proposed PRN method.

Acknowledgment

This work was supported in part by the National Science and Technology Major Project under the grant number J2019-IV-0001-0068.

References

[1] Z. Y. Chen, Z. R. Li, J. Wu, C. Deng, and W. Dai, “Deep residual shrinkage relation network for anomaly detection of rotating machines,” J Manuf Syst, vol. 65, pp. 579-590, Oct 2022.

[2] K. Hu, Y. W. Cheng, J. Wu, H. P. Zhu, and X. Y. Shao, “Deep bidirectional recurrent neural networks ensemble for remaining useful life prediction of aircraft engine,” IEEE T Cybernetics, vol. 53, pp. 2531-2543, Apr 2023.

[3] Y. W. Cheng, M. X. Lin, J. Wu, H. P. Zhu, and X. Y. Shao, “Intelligent fault diagnosis of rotating machinery based on continuous wavelet transform-local binary convolutional neural network,” Knowl-Based Syst, vol. 216, Mar 15 2021.

[4] R. X. Wang, H. K. Jiang, K. Zhu, Y. F. Wang, and C. Q. Liu, “A deep feature enhanced reinforcement learning method for rolling bearing fault diagnosis,” Adv Eng Inform, vol. 54, Oct 2022 .

[5] H. D. Shao, J. S. Cheng, H. K. Jiang, Y. Yang, and Z. T. Wu, “Enhanced deep gated recurrent unit and complex wavelet packet energy moment entropy for early fault prognosis of bearing,” Knowl-Based Syst, vol. 188, Jan 5 2020.

[6] Y. W. Cheng, C. Wang, J. Wu, H. P. Zhu, and C. K. M. Lee, “Multi-dimensional recurrent neural network for remaining useful life prediction under variable operating conditions and multiple fault modes,” Applied Soft Computing, vol. 118, 2022.

[7] Z. Y. Chen, J. Wu, C. Deng, X. Q. Wang, Y. H. Wang, “Deep attention relation network: a zero-shot learning method for bearing fault diagnosis under unknown domains,” IEEE T Reliab, vol. 72, no. 1, pp. 79-89, March 2023.

[8] Z. H. An, X. X. Jiang, J. Cao, R. Yang, X. G. Li, “Self-learning transferable neural network for intelligent fault diagnosis of rotating machinery with unlabeled and imbalanced data,” Knowl-Based Syst, vol. 230, Oct 2021.

[9] Z. Y. Chen, J. Wu, C. Deng, C. Wang, Y. H. Wang, “Residual deep subdomain adaptation network: A new method for intelligent fault diagnosis of bearings across multiple domains,” Mech Mach Theory, vol. 169, Mar 2022.

[10] X. Liu, Y. G. Li, Q. L. Meng, G. X. Chen, “Deep transfer learning for conditional shift in regression,” Knowl-Based Syst, vol. 227, Sept 2021.

[11] L. Guo, Y. G. Lei, S. B. Xing, T. Yan, N. P. Li, “Deep Convolutional Transfer Learning Network: A New Method for Intelligent Fault Diagnosis of Machines With Unlabeled Data,” IEEE T Ind Electron, vol. 66, no. 9, pp. 7316-7325, Sept. 2019.

[12] C. Lessmeier, J. K. Kimotho, D. Zimmer, W. Sextro. “Condition monitoring of bearing damage in electromechanical drive systems by using motor current signals of electric motors: A benchmark data set for data-driven classification,” 2016.

[13] F. Möller, D. Botache, D. Huseljic, F. Heidecker, M. Bieshaar, & B. Sick. (2021). “Out-of-distribution detection and generation using soft Brownian offset sampling and autoencoder,” 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 46-55.

[14] X. N. Liu, X. Z. Zhao, X. A. Liu, W. B. Shangguan, “A load spectrum editing method of time-frequency for rubber isolators based on the continuous wavelet transform,” Measurement, vol. 198, July 2022.

[15] H. Hu, X. Ma, Y. Z. Shang, “A novel method for transformer fault diagnosis based on refined deep residual shrinkage network,” IET Electr Power App, vol. 16, Oct. 2021.
[16] Z. M. Liu, Y. X. Wang, S. Vaidya, F. Ruehle, J. Halverson, M. Soljacic, Hou, T.Y., & Tegmark, M. (2024). “KAN: Kolmogorov-Arnold networks.” ArXiv, abs/2404.19756.
[17] A. Ding, Y. Oin, B. Wang, L. Guo, L. Jia, and X. Cheng, “Evolvable graph neural network for system-levelincremental fault diagnosis of train transmission systems,” Mechanical Systems and Signal Processing, vol. 210, March 2024.