Academic Journal of Emerging Technologies
ISSN Print: 3104-4417
ISSN Online: 3104-4425
CD ISSN: 3104-4433
https://doi.org/10.63313/AJET.9077

Intelligent Fault Diagnosis of High-Speed Train Bearings Based on Wide-Kernel CNN and CORAL Transfer Learning

Jincheng Li*. Xiye Luo. Taoyu Wang. Zhichao Hu. Yuhao Ma and Jie Xu
Tianjin University of Technology and Education, Tianjin, 300222, China
Email: 19902182847@163.com

How to cite this paper: Li, J., Luo, X., Wang, T., Hu, Z., Ma, Y., & Xu, J. (2026). Intelligent fault diagnosis of high-speed train bearings based on wide-kernel CNN and CORAL transfer learning. Academic Journal of Emerging Technologies, 4(1), 66–75. ISSN Print: 3104-4417, ISSN Online: 3104-4425.
https://doi.org/10.63313/AJET.9077
Published: 2026-11-18

Copyright © 2026 by author(s) and Erytis Publishing Limited.
This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).
http://creativecommons.org/licenses/by/4.0/



Abstract

High-speed train bearings operate under strong noise with extremely scarce fault labels, whereas laboratory testbeds provide abundant labeled data but with a severe domain shift. This paper proposes an interpretable fault diagnosis framework integrating mechanism-based feature extraction, a wide first-layer kernel CNN (WDCNN), and CORAL transfer learning. A 15-dimensional feature set (11 time-domain indicators and 4 fault-characteristic-frequency band energies) is constructed from bearing kinematics and impact dynamics, and a three-stage scheme (common-feature extraction, CORAL alignment, target-domain fine-tuning) bridges the testbed-to-train gap. The model achieves source-domain precisions above 0.94, outperforming SVM, RF, GBDT, and a standard DNN, and all predictions on 16 unlabeled target samples reach confidence no lower than 0.78. SHAP and UMAP analyses confirm that the decisions are consistent with bearing fault physics, providing an accurate, robust, and interpretable solution for train bearing monitoring.

Keywords

Bearing Fault Diagnosis; Wide First-Layer Kernel Convolutional Neural Network; CORAL Transfer Learning; Feature Extraction; Interpretability; High-Speed Train

1. Introduction

High-speed trains have become a major part of China's transportation system, with an operating mileage exceeding 42,000 km [1]. Bearings are core components of the running gear; under high speed, heavy load, and complex conditions, faults such as raceway spalling and rolling-element damage readily occur, and once initiated they can propagate through the entire system, causing train delays or even derailments. Intelligent bearing fault diagnosis is therefore essential for predictive maintenance.

Mechanical fault diagnosis has developed along two lines: signal processing (hybrid deep convolutional networks with extreme learning machines [3] and

sparse-signal-representation methods [4]) and data-driven learning, in which deep learning is now mainstream [5]–[7] and transfer learning bridges gaps across operating conditions and machines [8]–[10]; public benchmarks such as the CWRU bearing data center [2] support standardized evaluation.

Despite these advances, real train bearings face three bottlenecks: strong noise and non-stationarity (target-domain SNR about 40% lower than the source-domain drive-end data), a severe domain shift (about 1772 rpm → 600 rpm and 12 kHz → 32 kHz), and extremely scarce target labels (16 unlabeled samples), while safety-critical applications demand physically interpretable decisions.

To address these issues, this paper proposes a hybrid framework combining mechanism-driven feature engineering, a WDCNN tailored to one-dimensional vibration signals, and CORAL-based unsupervised domain adaptation. The main contributions are:

(1) A physics-guided 15-dimensional feature set covering the impact and frequency-modulation characteristics of bearing faults, with a multi-sensor priority strategy based on transfer-path attenuation analysis.

(2) A WDCNN with a wide first-layer kernel that adaptively learns short-time fault impact waveforms, guided by a receptive-field criterion that guarantees full coverage of the impact period.

(3) A three-stage transfer framework (common-feature extraction, CORAL domain adaptation, target-domain fine-tuning) with progressive easy-to-hard alignment and adaptive batch-normalization statistics, achieving confident diagnosis (confidence ≥ 0.78) on only 16 unlabeled target samples.

(4) A "physics + model + transfer + interpretation" explainability framework combining pre-hoc mechanism-driven design, in-process CORAL alignment analysis, and post-hoc SHAP/UMAP validation.

2. Proposed Methodology

2.1. Overall Framework

The proposed framework consists of four stages: (i) preprocessing and mechanism-driven 15-dimensional feature extraction; (ii) WDCNN training on the source domain to identify ball (B), inner-race (IR), outer-race (OR), and normal (N) states; (iii) CORAL alignment of the second-order statistics of the source and target distributions, followed by fine-tuning with layered learning rates and early stopping; and (iv) three-level interpretation (pre-hoc mechanism-driven design, in-process CORAL alignment logic, and post-hoc SHAP attribution and UMAP visualization).

2.2. Data Acquisition, Source-Domain Screening, and Preprocessing

The source-domain data are obtained from the CWRU bearing data center [2], which provides drive-end (DE), fan-end (FE), and base (BA) acceleration signals of SKF6205 deep-groove ball bearings. Faulty bearings were manufactured by

electro-discharge machining, covering three defect locations (ball, inner race, and outer race) with defect diameters of 0.007, 0.014, and 0.021 inches; the sampling frequency is 12 kHz, and each analysis sample contains 2048 data points.

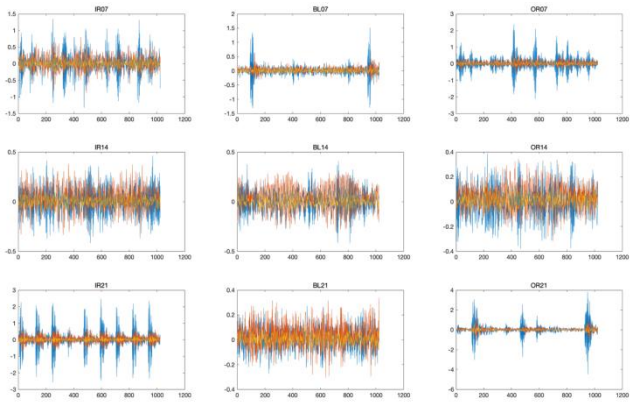


Figure 1. Time-domain waveforms of raw vibration signals under different bearing health conditions.

Because the target-domain data are limited, the source-domain subset is screened to reduce the distribution gap: it covers all four states and all fault diameters, preferentially uses the 12 kHz drive-end records whose fault features are most significant, and selects the 1772 rpm records, the closest CWRU operating points to the 600 rpm target bearing; the full source samples are finally used to ensure coverage of all fault types, fault sizes, and load conditions.

In preprocessing, the raw signals are converted to .csv format, denoised to suppress environmental, sensor, and other-component noise, and segmented with 2048 points per segment and a 50% overlap rate (a step size of 512).

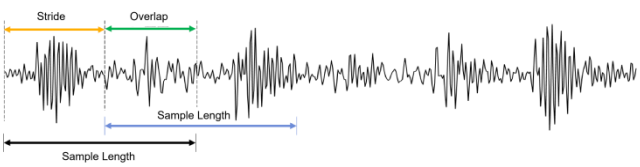


Figure 2. Schematic diagram of overlapping segmentation for vibration signal preprocessing.

Each segment is then standardized by removing its mean and dividing by its standard deviation.

2.3. Fault Mechanism Modeling and Feature Extraction

Analysis is performed from the time, frequency, and time-frequency perspectives: time-domain analysis extracts dimensionless impact indicators (e.g., kurtosis and crest factor) that are highly sensitive to early-stage faults; frequency-domain analysis computes the energy around the fault characteristic frequencies and their harmonics; and time-frequency analysis with the continuous wavelet transform serves as an auxiliary verification tool rather than as input features. Based on the signal-to-noise ratio maximization principle, the drive-end (DE) acceleration signal is selected as the main analysis object.

The fault characteristic frequencies are computed from the bearing geometry and rotational speed as

$$\mathrm{BPFO} = \frac{n}{2} f_r \left(1 - \frac{d}{D}\cos\alpha\right) \tag{1}$$

$$\mathrm{BPFI} = \frac{n}{2} f_r \left(1 + \frac{d}{D}\cos\alpha\right) \tag{2}$$

$$\mathrm{BSF} = \frac{D}{2d} f_r \left[1 - \left(\frac{d}{D}\cos\alpha\right)^2\right] \tag{3}$$

$$\mathrm{FTF} = \frac{1}{2} f_r \left(1 - \frac{d}{D}\cos\alpha\right) \tag{4}$$

When a rolling element passes a fault point, a transient impact is generated whose response is modulated by the damping ratio and the natural frequency; for inner-race and rolling-element faults the amplitude is additionally modulated by the shaft rotation or cage revolution frequency, and the sensor signal is the superposition of the impact response, the transfer-path effect, and Gaussian white noise.

To characterize bearing faults, 11 time-domain statistics sensitive to impacts (including dimensionless indicators such as kurtosis and crest factor) and 4 band-energy indicators (the energy integrals within the BPFO, BPFI, BSF, and FTF bands) are selected; candidates that are highly correlated or weakly fault-related are removed by mutual-information screening, yielding the final 15-dimensional feature vector, which is averaged per sample and standardized before partitioning.

2.4. Wide First-Layer Kernel Deep Convolutional Neural Network (WDCNN)

To match the bearing impact response, this paper adopts a WDCNN with a wide first-layer kernel (64 points) and small 3×1 kernels in the subsequent layers. The wide kernel extracts short-time impact features in a manner comparable to the short-time Fourier transform, but with weights learned adaptively, so that discriminative fault-related features are autonomously mined while irrelevant information is filtered out; the small-kernel layers then extract deep features with fewer parameters, suppressing overfitting. A batch normalization layer after each convolution layer and a 2×1 max-pooling operation accelerate training and improve stability. The overall architecture is illustrated in Figure 3.

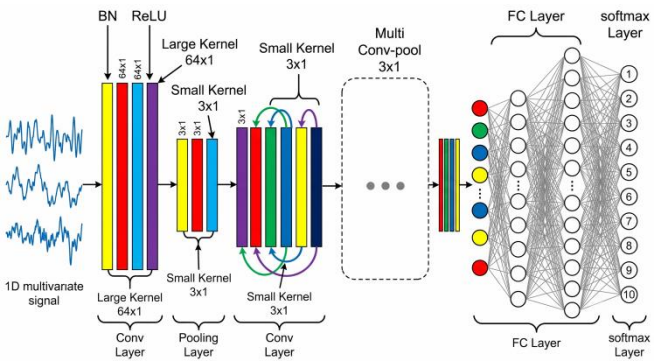


Figure 3. Architecture of the proposed Wide First-Layer Kernel Deep Convolutional Neural Network (WDCNN).

For the WDCNN structure with $S_l = 1$, $K_l = 3$ for $l > 1$, and pooling size 2, the receptive field $R(0)$ of a neuron in the last pooling layer on the input signal is

$$R(0) = 2^n - 1 + \sum_{i=1}^{n} 2^{i-1} K_{n-i+1} \tag{5}$$

Together with the design criterion $T \le R(0) \le L$, where T is the impact period and L is the input length, and the condition that $S(1)$ must divide L, the final design criterion is

$$T \le 2^n - 1 + \sum_{i=1}^{n} 2^{i-1} K_{n-i+1} \le L, S(1)|L \tag{6}$$

This criterion ensures that the receptive field covers the complete fault impact period and that the network is structurally transparent before training, providing a stable feature-representation space for the subsequent transfer learning.

2.5. CORAL-Based Transfer Learning

The source and target domains differ in bearing type, rotational speed, and sampling frequency. The target train bearing runs at a rated speed of about 600 rpm, significantly lower than the source-domain operating points of about 1772 rpm, so the fault characteristic frequencies shift systematically and the center of the vibration energy distribution moves accordingly. The sensor mounting position in the target domain is farther from the bearing, so high-frequency components are severely attenuated and the target SNR is about 40% lower than that of the source-domain drive-end data; the train environment additionally introduces track vibration and gearbox noise with spectral characteristics different from laboratory noise. Since the target domain contains only 16 unlabeled samples, independent model training is impossible, and the source-domain knowledge must be effectively reused through transfer learning.

CORAL is a feature-based unsupervised transfer method that reduces the domain distribution difference by aligning the second-order statistics (covariance matrices) of the source and target features. Let Xs (ns samples, d dimensions) and Xt (nt samples) denote the source and target features; CORAL seeks a linear transformation matrix A such that the covariance matrix of the transformed source features approaches that of the target features:

$$\min_A \frac{1}{4d^2} \|\mathrm{Cov}(X_s A) - \mathrm{Cov}(X_t)\|_F^2 \tag{7}$$

where $\mathrm{Cov}(\cdot)$ is the covariance matrix and $\|\cdot\|F$ is the Frobenius norm.

Domain alignment follows a progressive easy-to-hard strategy: impact features with high generalization (kurtosis, impulse factor) are aligned first, then energy features affected by transfer-path attenuation (RMS, standard deviation), and finally fault-frequency features are aligned via band-energy integration instead of absolute frequency values, avoiding speed-dependent shift errors and reducing the risk of negative transfer.

The transfer procedure follows a three-stage framework: common-feature extraction, domain adaptation, and target-domain fine-tuning. In the first stage, the 12 kHz drive-end data with the most significant fault features are selected and the cross-domain stable 15-dimensional feature system is constructed; the WDCNN further extracts multi-resolution impact features from the raw segments, enhancing the physical interpretability of the representation. In the fine-tuning stage, layered learning rates are adopted, with a small rate preserving the bottom-layer common features and a large rate adapting the top layers to the target domain, and early stopping prevents overfitting.

For the target domain, the batch normalization statistics are additionally adapted so that the network adjusts to the target-domain data distribution while keeping the trained weights.

2.6. Interpretability Framework

The interpretability of the transfer diagnosis is elaborated at three levels. **Pre-hoc:** the 15-dimensional features directly encode bearing kinematics — the time-domain statistics capture the impact response of local damage and the envelope-spectrum energies of BPFO, BPFI, BSF, and FTF encode the fault-frequency knowledge — and the receptive-field criterion guarantees coverage of the complete impact period.

In-process: CORAL minimizes the covariance difference between the two domains, equivalently rotating and scaling the source feature distribution toward the target distribution so that the energy patterns of the fault characteristic frequencies remain recognizable, and each progressive-alignment step corresponds to an explicit physical objective.

Post-hoc: SHAP converts the nonlinear mapping of the WDCNN into feature-contribution explanations based on Shapley values, and UMAP projects the high-dimensional feature structure into two dimensions to inspect the separability and clustering of the fault types, cross-checked by PCA and spectral analysis.

3. Experiments and Results

3.1. Experimental Setup

The source domain consists of the CWRU drive-end bearing data (SKF6205/SKF6203) covering ball, inner-race, and outer-race faults and the normal state, sampled at 12 kHz with 2048 points per sample and recorded at 1772 rpm; the target domain consists of 16 unlabeled samples of an unknown train bearing collected at about 600 rpm with a sampling frequency of 32 kHz. A hold-out validation is adopted, dividing the data into mutually exclusive training, validation, and test sets, and precision, recall, and F1-score are reported for each fault class.

3.2. Source-Domain Diagnosis Results and Comparison with Baselines

On the source domain, the proposed WDCNN achieves precisions of 0.95, 0.96, and 0.97 for the ball, inner-race, and outer-race faults, respectively, with recall and

F1-score no lower than 0.94 and 0.95, indicating that the model learns both the impact and the modulation characteristics of the different fault mechanisms.

The diagnostic performance is compared with SVM, RF, GBDT, and a standard DNN (Figure 4), and the per-class F1-scores are summarized in Table 1. As shown in Table 1, WDCNN achieves F1-scores no lower than 0.95 for every class, clearly superior to SVM (average F1 0.54) and the standard DNN (average F1 0.65); the shallow ensembles RF and GBDT also remain below 0.81 in every class. This comparison also serves as an ablation analysis: replacing the WDCNN feature extractor with shallow classifiers removes the transfer-capable representation, and replacing the wide-kernel front-end with a plain fully connected DNN drops the F1-scores from 0.96/0.95/0.96 to 0.67/0.55/0.72, confirming that both the architecture and the mechanism-driven features contribute to the final accuracy.

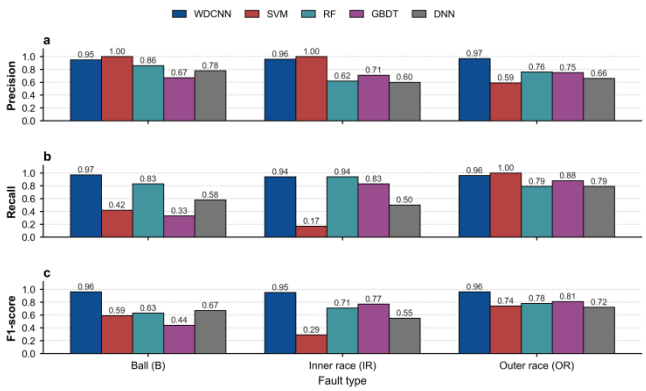


Figure 4. Performance comparison of WDCNN and baseline models on Precision, Recall, and F1-score.

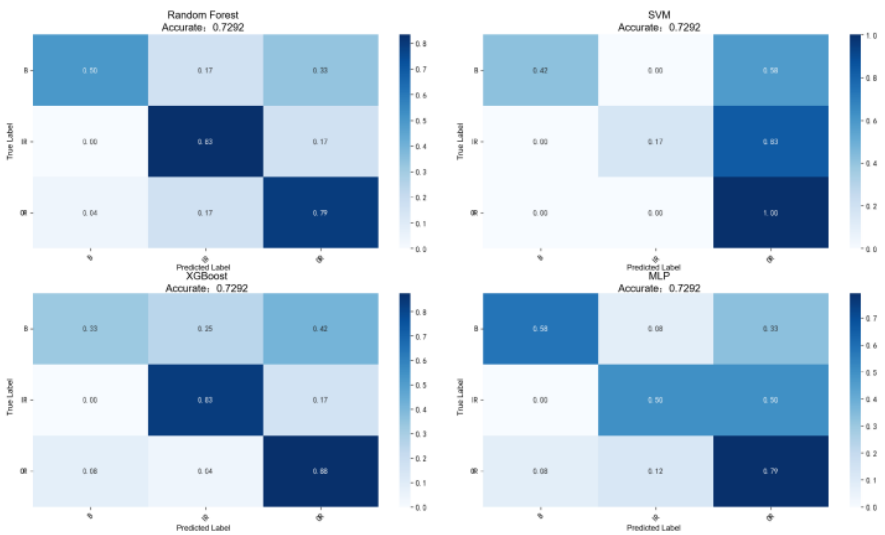


Figure 5. Confusion matrices of baseline machine learning models on the source domain test set.

Table 1. F1-score comparison of WDCNN and baseline models on the source domain.

Method	Ball-F1	IR-F1	OR-F1	Average F1
WDCNN (proposed)	0.96	0.95	0.96	0.96
SVM	0.59	0.29	0.74	0.54
RF	0.63	0.71	0.78	0.71
GBDT	0.44	0.77	0.81	0.67
DNN	0.67	0.55	0.72	0.65
Deep CORAL [12]	—	—	—	—
DANN [11]	—	—	—	—

Deep CORAL [12] and DANN [11] are discussed qualitatively; quantitative results are not yet available because the 16 target-domain samples are unlabeled and accuracy-based metrics cannot be computed without ground-truth labels.

Compared with the mainstream unsupervised domain adaptation methods Deep CORAL [12] and DANN [11], the proposed framework avoids the training instability of adversarial min-max optimization while maintaining the source-domain classification accuracy; a quantitative head-to-head comparison is left for future work.

3.3. Transfer Diagnosis Results on the Target Domain

The CORAL-adapted and fine-tuned model is applied to the 16 unlabeled target-domain samples. All predictions reach confidence ≥ 0.78 (the minimum is 0.78 for sample G.mat), with 10 of the 16 samples above 0.90; the predicted types cover 7 outer-race, 5 inner-race, and 3 ball faults plus 1 normal sample (L.mat, confidence 0.98). Although the target domain contains only 16 samples, the model predictions are all confident and their physical rationality is verified by the spectral analysis of the target-domain features.

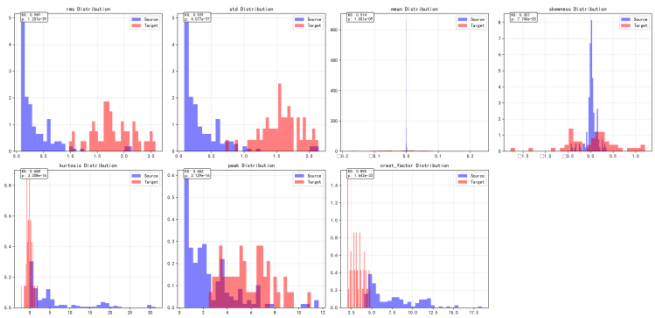


Figure 6. Histograms with KS test results illustrating the feature distribution gap before domain adaptation.

3.4. Visualization and Interpretability Analysis

The kurtosis and impulse factor distributions of the source and target domains show high-value overlap (kurtosis significantly above 3 and impulse factor above 5 in both domains), indicating that the target signals retain the core impact features of bearing faults and that these dimensionless indicators can be transferred directly across operating conditions.

RMS and standard deviation differ in absolute value between the domains (e.g., normal-sample RMS about 0.2 versus 0.3) but preserve a consistent relative law — fault samples have higher energy and fluctuation, increasing with fault severity — so only the absolute values need to be adjusted by domain alignment (Figure 7).

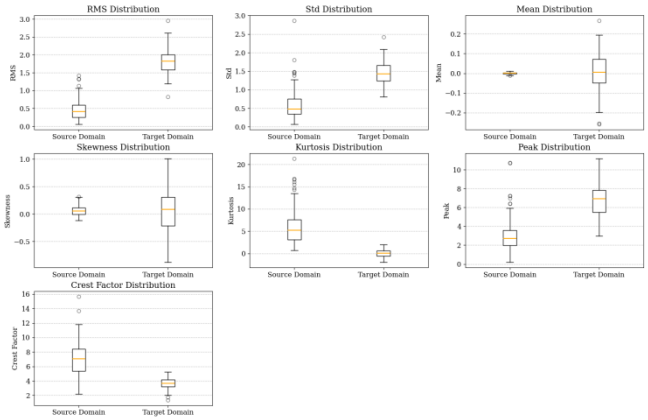


Figure 7. Boxplots comparing the statistical distributions of the 7 engineered features between the source and target domains.

As shown in Figure 8, the UMAP visualization exhibits clear clustering of the fault types: the normal-state samples form a compact, well-separated cluster; the outer-race cluster is the most compact among the fault types because the outer race is stationary and unmodulated; and the inner-race and ball clusters partially overlap as both are modulated and produce similar broadband impact signatures. SHAP further attributes the decisions to the mechanism-driven features (kurtosis, impulse factor, and fault-frequency energies) rather than noise-related artifacts, confirming that the model decisions are consistent with the bearing fault physics.

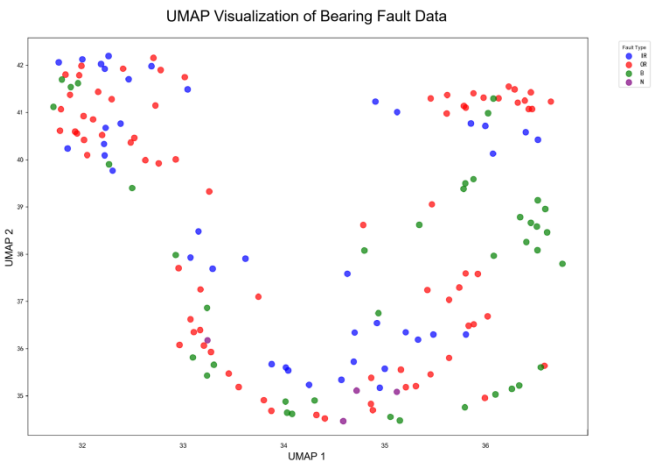


Figure 8. UMAP visualization of the high-dimensional feature space for different bearing fault types.

4. Conclusion

This paper presents an interpretable fault diagnosis framework for high-speed train bearings integrating mechanism-driven features, a WDCNN, and CORAL transfer learning. A 15-dimensional feature set with a receptive-field criterion ensures impact-period coverage, and a three-stage transfer framework with CORAL alignment and layered learning rates achieves confident diagnosis (≥ 0.78) on 16 unlabeled target samples with source precisions above 0.94. Interpretability analyses (SHAP, UMAP, PCA, spectral) verify consistency with bearing fault physics.

Limitations include static feature snapshots, missing variable-speed/load validation, and reduced weak-fault sensitivity. Future work will address condition-adaptive alignment, degradation modeling, and edge deployment. The framework also extends to other rotating machinery by replacing equipment-specific parameters.

References

[1] Lu, F., Tong, Q., Jiang, X., et al. (2024) An Interpretable Method for Mechanical Fault Diagnosis Based on Conditional Metric Transfer Learning. Chinese Journal of Scientific Instrument, 45, 250-262.

[2] Case Western Reserve University (2025) Bearing Data Center. https://engineering.case.edu/bearingdatacenter/download-data-file

[3] Chen, Z., Gryllias, K. and Li, W. (2019) Mechanical Fault Diagnosis Using Convolutional Neural Networks and Extreme Learning Machine. Mechanical Systems and Signal Processing, 133, Article ID: 106272.

[4] Qiao, Z., Zhao, Z., Cao, D., et al. (2024) Sparse Signal Representation and Its Applications in Mechanical Fault Diagnosis: A Review. Mechanical Systems and Signal Processing, 208, Article ID: 110948.

[5] Zhang, W., Peng, G., Li, C., Chen, Y. and Zhang, Z. (2017) A New Deep Learning Model for Fault Diagnosis with Good Anti-Noise and Domain Adaptation Ability on Raw Vibration Signals. Sensors, 17, Article ID: 425.

[6] Lei, Y., Yang, B., Jiang, X., Jia, F., Li, N. and Nandi, A.K. (2020) Applications of Machine Learning to Machine Fault Diagnosis: A Review and Roadmap. Mechanical Systems and Signal Processing, 138, Article ID: 106587.

[7] Zhao, Z., Li, T., Wu, J., Sun, C., Wang, S., Yan, R. and Chen, X. (2020) Deep Learning Algorithms for Rotating Machinery Intelligent Diagnosis: An Open Source Benchmark Study. ISA Transactions, 107, 224-255.

[8] Guo, L., Lei, Y., Li, N., Yan, T. and Xing, S. (2019) Deep Convolution Transfer Learning Network: A New Method for Intelligent Fault Diagnosis of Machines with Unlabeled Data. IEEE Transactions on Industrial Electronics, 66, 7316-7325.

[9] Li, X., Zhang, W., Ding, Q. and Sun, J.-Q. (2019) Multi-Layer Domain Adaptation Method for Rolling Bearing Fault Diagnosis. Signal Processing, 157, 180-197.

[10] Yang, B., Lei, Y., Jia, F. and Xing, S. (2019) An Intelligent Fault Diagnosis Approach Based on Transfer Learning from Laboratory Bearings to Locomotive Bearings. Mechanical Systems and Signal Processing, 122, 692-706.

[11] Ganin, Y., Ustinova, E., Ajakan, H., et al. (2016) Domain-Adversarial Training of Neural Networks. Journal of Machine Learning Research, 17, 1-35.

[12] Sun, B. and Saenko, K. (2016) Deep CORAL: Correlation Alignment for Deep Domain Adaptation. In: Proceedings of the European Conference on Computer Vision Workshops (ECCVW), Amsterdam, The Netherlands, 443-450.