https://doi.org/10.37965/jdmd.2026.1595

FD-SE-LLM: A Semantic-Enhanced Large Language Model Framework for Fault Diagnosis of Hydropower Carbon Brush Bearings

Liyan Yang[a*], Liqing Zhu[a], Cheng Yang[a], Zongli Lin[a]

a Huadian Fuxin Zhouning Pumped Storage Co., Ltd., Ningde 352200, Fujian, China.

Reived 30 June 2026; Revised 06 August 2026; Accepted 18 August 2026; Published online 10 September 2026

Abstract:

Cross-condition fault diagnosis remains a fundamental challenge in intelligent operation and maintenance of hydropower equipment, where deep learning methods suffer 20%−30% accuracy degradation under operating condition drifts. Large language models (LLMs), through massive multi-domain pre-training, offer cross-domain generalization that may help overcome this limitation. However, applying LLMs to industrial vibration signals faces several difficulties: a modality gap between continuous physical time series and discrete semantic sequences, multi-fault class aliasing in the semantic space, and limited sensitivity to transient fault pulses. This paper proposes FD-SE-LLM, a Semantic-Enhanced Large Language Model Framework for Fault Diagnosis targeting hydropower carbon brush bearings. The framework includes Time-domain Semantic Cross-Correlation (TSCC) to project periodic and transient features into the LLM semantic space, a Fault-specific Multi-class Conditional Variational Autoencoder (FM-CVAE) to disentangle fault class-specific representations from environmental noise, and an embedded Bearing Time-Adapter to enhance transient pulse sensitivity. Experiments on hydropower, CWRU, and JNU datasets show that FD-SE-LLM achieves 91.5% average cross-condition accuracy, outperforming deep learning baselines by 6.8−15.1 pp and existing LLM-based approaches by 3.9−6.3 pp.

Keywords: Fault of bearing; Fault diagnosis; Contrastive learning; hydropower carbon brush bearing.

1. Introduction

Hydropower generation depends on the safe operation of generating units [1], [2]. Carbon brush bearings, which provide electrical transmission between the rotor and excitation system, operate under high humidity, strong vibration, and variable loads. These conditions predispose them to faults such as inner race spalling, outer race cracking, rolling element fatigue, and brush wear [3, 4]. Statistics indicate that 30%−45% of unplanned shutdowns stem from bearing system faults [5], underscoring the importance of reliable fault diagnosis for predictive maintenance. Bearing fault diagnosis can be formulated as a vibration signal- based pattern recognition problem.

https://doi.org/10.37965/jdmd.2026.1595
Vibration signals carry rich fault information: different fault modes produce characteristic responses in specific frequency bands, exhibiting periodicity, transience, and non- stationarity [6].

Deep learning methods have achieved notable progress in this area. CNNs extract discriminative features through hierarchical convolution [7, 8], RNNs and LSTM capture temporal dependencies under variable speed [9], and Transformers enable multi- scale feature fusion via self- attention [10]. On benchmark datasets, these methods achieve over 98% accuracy [11]. This performance, however, assumes that training and test data share the same distribution. In practice, load fluctuations, speed adjustments, and temperature variations introduce distribution drift that degrades accuracy by 20%−30% [12, 13]. Data- driven generalization is inherently bounded by training data coverage, as learned representations cannot anticipate all variable field conditions.

Domain adaptation methods [14-16] address cross-condition transfer by aligning feature distributions, but cannot generalize beyond observed data. Despite being trained on natural language, large language models (LLMs) are not confined to textual semantics; their self-attention mechanism inherently captures long-range dependencies and periodic patterns, which are analogous to the repetitive impact intervals in bearing fault signals. Meanwhile, the high-dimensional embedding space of LLMs can represent abstract fault prototypes, allowing vibration features—after appropriate alignment—to be reasoned over in a unified semantic space. This motivates us to introduce LLMs into vibration fault diagnosis as general-purpose temporal reasoning engines, rather than simply as text processors. Large language models (LLMs) [17-20] offer a complementary approach: pre- trained on massive multi- domain corpora, LLMs acquire broad pattern reasoning that may generalize to unseen conditions. Time- LLM [21], S2IP- LLM [22], and SE- LLM [23] have demonstrated the potential of LLM-based time- series analysis.

However, applying LLMs to industrial vibration signals involves several difficulties. The modality gap between continuous physical time series and discrete semantic sequences means that Time- LLM's text conversion sacrifices high- frequency details [21], while SE- LLM does not exploit the periodic and transient structure specific to vibration signals [23]. In multi- fault scenarios, hydropower bearings exhibit overlapping spectral features among inner race, outer race, and brush wear faults, and the binary VAE used in SE- LLM cannot separate individual fault classes in the latent space. Additionally, LLM attention is designed for global token correlations and tends to smooth over the microsecond- level transient pulses that carry early fault information.

This paper proposes FD- SE- LLM targeting hydropower carbon brush bearings. The framework introduces three modules: Time- domain Semantic Cross- Correlation (TSCC) projects periodicity and transient features into the LLM semantic space through

https://doi.org/10.37965/jdmd.2026.1595
cross- modal attention; Fault- specific Multi- class CVAE (FM- CVAE) uses class labels as conditional constraints with a separation loss to disentangle overlapping fault representations; and a Bearing Time- Adapter embedded in the LLM attention

1. TSCC mechanism with periodicity prior weighting and Teager energy operator-based transient detection for implicit alignment between vibration features and LLM semantic prior.

2. FM-CVAE with explicit class labels, class separation loss, and orthogonality penalty for precise discrimination in multi-fault scenarios.

3. Bearing Time-Adapter with sampling-rate-adaptive kernel and long-short-term fusion for improved microsecond-level pulse sensitivity.

4. Experiments on hydropower, CWRU, and JNU datasets showing accuracy improvement over deep learning and existing LLM methods.

Section II reviews related work; Section III details the method; Section IV presents results; Section V concludes..

2. Fundamental concept

2.1 Cross-Condition Fault Diagnosis Methods and Generalization Bottlenecks

Cross-condition fault diagnosis requires transferring a classifier trained under source conditions to target conditions with distribution shift $P_S(X) \neq P_T(X)$. Domain adaptation methods minimize this discrepancy through statistical alignment (MMD [14], CORAL [16]) or adversarial training [15]. Generative augmentation via WGAN [24] or diffusion models [25] synthesizes fault samples but inherits the source domain's diversity limitations. These methods adjust existing decision boundaries but cannot expand the fault representation at the data level, motivating interest in large language models (LLMs) with broad prior knowledge.

2.2 Application of Large Language Models in Time-Series Analysis

LLMs' core advantage over deep learning lies in cross-domain generalization [19]. For general time-series, TimeLLM [21] reprograms LLM word embeddings, S2IP-LLM [22] uses embedded prompts, PatchTST [26] applies patch-based tokenization, and SE-LLM [23] introduces semantic cross-correlation with a Time-Adapter. For fault diagnosis, FD-LLM [27] textualizes statistical features for chain of thought reasoning, and FD-Zero [28] explores zero shot cross-domain diagnosis. These approaches either lose information during textualization or lack resolution for overlapping fault types—shortcomings addressed by the TSCC and FM-CVAE modules in this work.

2.3 Physical Properties of Vibration Signals and Modeling Challenges

From a signal processing perspective, industrial vibration signals exhibit properties that directly challenge LLM-based modeling. Bearing localized faults produce periodic impact pulses during rotation, with repetition frequencies determined

https://doi.org/10.37965/jdmd.2026.1595
by bearing geometry and speed, corresponding to characteristic frequencies such as ball pass frequency inner race (BPFI) and ball pass frequency outer race (BPFO) [29]. LLM self-attention is designed for global token correlation and lacks specialized mechanisms for extracting such periodic structures. Different fault types can also have partially overlapping characteristic frequencies: for example, higher harmonics of inner race faults may approach the fundamental frequency of outer race faults [30], making it difficult for a standard binary VAE to separate individual fault classes in the latent space. In early fault stages, defect zones produce transient impact pulses lasting 0.1-1ms [31]. These pulses are brief and low in energy, yet LLM attention tends to distribute weight across the full sequence, effectively smoothing over these diagnostic cues.

3. Model structure

3.1 Problem Definition and Framework Overview

The problem setting is as follows: the source domain has sufficient fault- labeled data, while the target domain lacks labeled data—the model must train on the source domain and maintain high diagnostic accuracy on the target domain. Formally, given source domain $\mathrm{D}_S = \{(\mathbf{x}_i^S, \mathbf{y}_i^S)\}_{i=1}^{N_S}$ where $\mathbf{x}_i^S \in \mathrm{R}^{1\times L}$ is a one- dimensional vibration signal of length L and $\mathbf{y}_i^S \in \{0,1,\ldots,K\}$ is the fault type label ($y=0$ for normal). The target domain $\mathrm{D}_T = \{\mathbf{x}_i^T\}$ has a different distribution due to condition changes.

During cross- condition deployment, target samples lack labels. FD- SE- LLM uses a two- stage pipeline: a preliminary classifier that the FD- SE- LLM model itself trained on source domain labels generates pseudo- labels $\hat{y}$ for target samples, which serve as FM- CVAE's conditional input. On the hydropower dataset, the pseudo- label accuracy on target domain samples is approximately 87% (Target 1) and 83% (Target 2), as estimated by cross- validation on labeled source data projected onto target features. The class separation loss provides additional robustness: by enforcing a margin m between class prototypes, small perturbations in the conditional input do not cause cluster collapse.

3.2 Time-Domain Semantic Cross-Correlation (TSCC)

The TSCC module bridges the modality gap between vibration signals and the LLM semantic space. SE- LLM [23] proposed cross- correlation for general time- series but did not account for the specific periodic and transient structure of industrial vibration signals. TSCC addresses this through periodicity prior weighting and transient pulse detection, both of which are physically grounded in bearing dynamics. The periodicity prior $\mathbf{g}_T$ explicitly captures the recurrence intervals of characteristic fault frequencies, producing high weights at temporal positions aligned with mechanical

https://doi.org/10.37965/jdmd.2026.1595
impact occurrences. The Teager energy operator enhances sensitivity to the transient impulses generated by localized defects, which manifest as abrupt energy spikes in the vibration envelope.

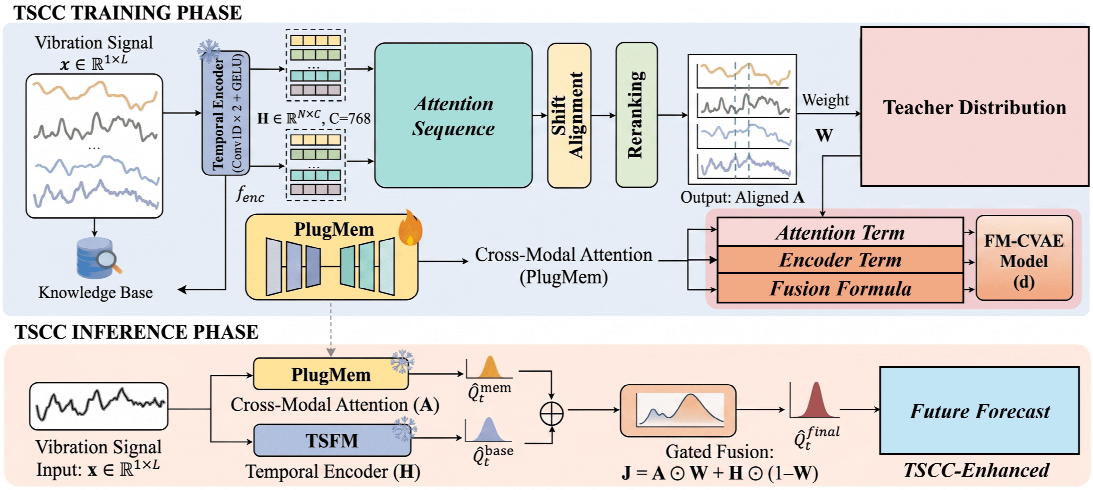


Fig 1：TSCC representation enhancement mechanism

The raw signal $\mathbf{x} \in \mathrm{R}^{1\times L}$ is first segmented into N non- overlapping patches:

$$\hat{T} = \mathrm{SlideWindow}(\mathbf{x}) \in \mathrm{R}^{B\times N\times K} \tag{1}$$

where B is the batch size, N is the number of segments, K is the segment length, and the sliding window operates with a fixed stride equal to K, ensuring that the total signal length satisfies $N \times K = L$. This reduces self- attention complexity from $O(L^2)$ to $O(N^2)$, approximately K^2 times reduction. The temporal encoder projects segments to high- dimensional features through two 1D convolution layers:

$$H = F_2\left(\sigma\left(F_1\left(\hat{T}\right)\right)\right) \in \mathrm{R}^{B\times N\times C} \tag{2}$$

where F_1 and F_2 are convolution layers with kernel size 3 and stride 1, σ is GELU activation, and C matches the LLM embedding dimension.

A semantic reference space $S \in \mathrm{R}^{K\times C}$ is then constructed from the pre- trained LLM word embedding matrix $W \in \mathrm{R}^{V\times C}$ by selecting K_s representative semantic anchors. The core alignment uses multi- head cross- attention with vibration features H as Query and semantic matrix S as Key and Value:

$$J_S = \mathrm{MultiHead}(H, S, S) \in \mathrm{R}^{B\times N\times C} \tag{3}$$

where the multi- head attention follows the standard scaled dot- product formulation $\mathrm{Attention}(Q,K,V) = \mathrm{Softmax}\left(QK^{\cdot} / \sqrt{d_k}\right)V$ with $h = 8$ heads.

To exploit the periodic structure of bearing fault signals, a periodicity prior weight vector $\mathbf{g}_T \in \mathrm{R}^{N}$ is computed via normalized autocorrelation of the feature sequence

https://doi.org/10.37965/jdmd.2026.1595
along the temporal axis. Specifically, for each feature vector $\mathbf{h}_n \in \mathrm{R}^C$ at position n, we compute its cosine similarity with all other positions, average over the channel dimension, and apply Softmax:

$$\mathbf{g}_T = \mathrm{Softmax}\left(\frac{1}{C}\sum_{c=1}^{C}\mathrm{Autocorr}\left(\mathbf{H}_{c,c}\right)\right) \in \mathrm{R}^N \tag{4}$$

where $\mathrm{Autocorr}(\cdot)$ computes the unbiased autocorrelation estimate. The resulting $\mathbf{g}_T$ is broadcast to $\mathrm{R}^{B\times N\times C}$ via element- wise multiplication. Periodic time positions receive higher weights, focusing semantic alignment on segments carrying periodic fault information.

For transient feature extraction, a detection module $\mathbf{M}_{\mathrm{anomaly}}$ identifies transient mutations via a short- time Teager energy operator (TEO), which is more robust than raw kurtosis on short windows [32]:

$$\mathbf{M}_{\mathrm{anomaly}} = \mathrm{Sigmoid}\left(\mathrm{TEO}\left(H\right)\right) \tag{5}$$

where TEO computes the discrete Teager energy $\Psi\left[\mathbf{x}(n)\right] = x^2(n) - x(n-1)x(n+1)$ along the temporal axis, providing enhanced sensitivity to transient impulses while maintaining stability over short windows [32]. Compared with kurtosis computed on $K = 64$ segments, TEO is less susceptible to noise- induced false positives because it operates on consecutive samples rather than requiring higher- order moment estimation.

The final representation fuses all features via gating:

$$J = G \odot \left(\mathbf{g}_T \odot J_S + \mathbf{M}_{\mathrm{anomaly}} \odot H\right) + \left(1 - G\right) \odot H \tag{6}$$

where $\mathbf{g}_T \in \mathrm{R}^{B\times N\times 1}$ and $\mathbf{M}_{anomaly} \in \mathrm{R}^{B\times N\times 1}$ are broadcast along the channel dimension to $\mathrm{R}^{B\times N\times C}$ by replicating the same value across all C channels, enabling element-wise (Hadamard) multiplication with $J_S \in \mathrm{R}^{B\times N\times C}$ and $H \in \mathrm{R}^{B\times N\times C}$, respectively. The gate $G \in \mathrm{R}^{B\times N\times C}$ is computed as $G = \sigma\left(Linear\left(\left[H; J_S\right]\right)\right)$, where $[\cdot;\cdot]$ denotes concatenation along the channel dimension and the linear layer maps from $2C$ to C dimensions; this ensures that G applies per-element modulation at each channel independently, preserving channel-wise feature distinctions. Note that the periodicity prior modulates the semantically aligned features J_S, while the transient detection modulates the raw temporal features H. This differentiated design ensures that transient information, which may be diluted during cross- modal attention, is retained through the $\mathbf{M}_{\mathrm{anomaly}} \odot \mathbf{H}$ pathway.

3.3 FM-CVAE Disentanglement Mechanism

FM-CVAE addresses the multi-fault aliasing problem by introducing explicit fault class conditional constraints and a class separation loss.

FM- CVAE takes the TSCC output $\mathbf{J}$ as input and fault class label c as condition. The class label is first embedded into a continuous vector $e_c \in \mathrm{R}^C$ via a

https://doi.org/10.37965/jdmd.2026.1595
learnable embedding layer, then concatenated with $\mathbf{J}$ before being fed into the encoder:

$$q_\phi(\mathbf{z}\,|\,\mathbf{J},c)=\mathrm{N}\,(\mathbf{z};\mu_\phi\left(\left[\mathbf{J};e_c\right]\right),\mathrm{diag}\left(\sigma_\phi^2\left(\left[\mathbf{J};e_c\right]\right)\right) \tag{7}$$

where $\left[\cdot;\cdot\right]$ denotes concatenation along the channel dimension, and μ_ϕ,σ_ϕ^2 are parameterized by separate MLP heads.

The encoder MLP $\left[\mu_\phi;\log\sigma_\phi^2\right]$ consists of three fully-connected layers with hidden dimensions $256\to128\to\left(d_c+d_n\right)$, where the output dimension is d_c+d_n for both the mean and log-variance predictions. The decoder MLP similarly has three layers $128\to256\to C$ for reconstructing the TSCC output. Each hidden layer is followed by Batch Normalization and GELU activation, with a dropout rate of 0.1 applied after each hidden layer during training to prevent overfitting, given the limited size of the hydropower dataset (3,600 samples). The dimension decay follows a gradual reduction strategy ($C\to256\to128$ for the encoder) to preserve sufficient representational capacity before the latent bottleneck. The same architecture is used consistently across all datasets; no dataset-specific tuning was performed.

$$\mathbf{z}=\left[\mathbf{z}_{\text{class}},\mathbf{z}_{\text{noise}}\right] \tag{8}$$

where $\mathbf{z}_{\text{class}}\in\mathrm{R}^{d_c}$ encodes fault class- specific features and $\mathbf{z}_{\text{noise}}\in\mathrm{R}^{d_c}$ encodes environmental noise. To ensure separated cluster structures in the class- specific subspace, a class separation loss is introduced:

$$\mathrm{L}_{\text{sep}}=\frac{2}{K\left(K-1\right)}\sum_{i<j}\max\left(0,m-\|\,\mu_{c_i}-\mu_{c_j}\|_2\right) \tag{9}$$

where m is the separation margin, K is the number of fault classes, and the normalization $2/\left[K\left(K-1\right)\right]$ ensures scale invariance. μ_{c_i} denotes the prototype (mean) of class c_i in the $\mathbf{z}_{\text{class}}$ subspace.

To enforce orthogonality between $\mathbf{z}_{\text{class}}$ and $\mathbf{z}_{\text{noise}}$, an orthogonality penalty is applied:

$$\mathrm{L}_{\text{ortho}}=\|\,\mathbf{Z}_{\text{class}}^{\cdot}\mathbf{Z}_{\text{noise}}\|_F^2\,/\,N \tag{10}$$

where $\mathbf{Z}_{\text{class}}$ and $\mathbf{Z}_{\text{noise}}$ are centered matrices. This Frobenius norm penalizes any residual cross- correlation between the two subspaces.

The total training objective combines reconstruction, regularization, separation, classification, and orthogonality:

$$\mathrm{L}=\mathrm{L}_{\text{recon}}+\beta\mathrm{L}_{\text{KL}}+\gamma\mathrm{L}_{\text{sep}}+\lambda\mathrm{L}_{\text{cls}}+\eta\mathrm{L}_{\text{ortho}} \tag{11}$$

where $\mathrm{L}_{\text{recon}}$ is mean squared reconstruction error, L_{KL} is the standard VAE KL divergence, and L_{cls} is cross- entropy classification loss on $\mathbf{z}_{\text{class}}$.

https://doi.org/10.37965/jdmd.2026.1595

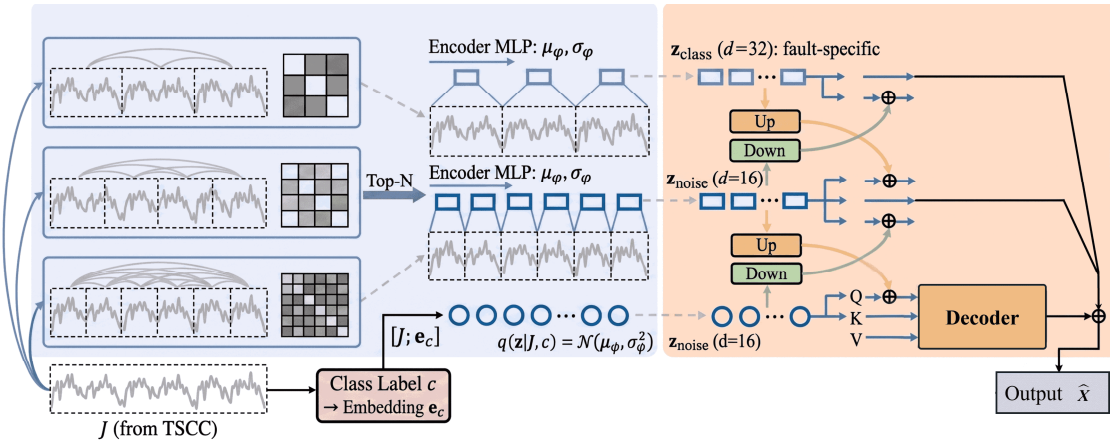


Fig 2：FM-CVAE disentanglement mechanism

3.4 Bearing Time-Adapter

The Bearing Time- Adapter resolves LLM's insufficient short- time transient sensitivity.

The Bearing Time- Adapter is inserted in the Key/Value computation path of LLM attention via residual connection:

$$\begin{aligned} K' &= K + \text{Time} - \text{Adapter}(K) \\ V' &= V + \text{Time} - \text{Adapter}(V) \end{aligned} \tag{12}$$

This placement ensures that the adapter modifies the information flowing into the attention computation without altering the pre- trained Query projections.

The pulse detection branch employs 1D convolution with an adaptively scaled kernel:

$$M_{\text{pulse}} = \text{Conv1d}_{k_{\text{pulse}}}(K) \tag{13}$$

with k_{pulse} adaptively scaled to the sampling rate: $k_{\text{pulse}} = \max\left(3, \left|0.1 \times f_s / 1000\right|\right)$ yielding $k_{\text{pulse}} = 3$ for 12kHz $k_{\text{pulse}} = 3$ for 25.6kHz and $k_{\text{pulse}} = 5$ for 50kHz This maintains a physical coverage of approximately 0.10- 0.25 ms per kernel, which spans the typical duration of transient fault pulses [31].

A long- period trend is extracted in parallel via average pooling, and the two branches are fused:

$$M_{\text{trend}} = \text{AvgPool}_{k_{\text{trend}}}(K) \tag{14}$$

$$M_{\text{fused}} = \alpha \cdot M_{\text{pulse}} + \beta_{\text{fused}} \cdot M_{\text{trend}} \tag{15}$$

with $k_{\text{trend}} = 16$ for long- period trend extraction. The fused features are then modulated through a learned gate:

$$\begin{aligned} G_{\text{adapt}} &= \sigma\left(\text{Linear}\left(\left[K, M_{\text{fused}}\right]\right)\right) \\ \text{Time} - \text{Adapter}(K) &= G_{\text{adapt}} \odot M_{\text{fused}} \end{aligned} \tag{16}$$

https://doi.org/10.37965/jdmd.2026.1595

The gating mechanism $G_{\text{adapt}} \in [0,1]$ learns to suppress adapter output for non-informative segments and amplify it when transient pulses are detected.

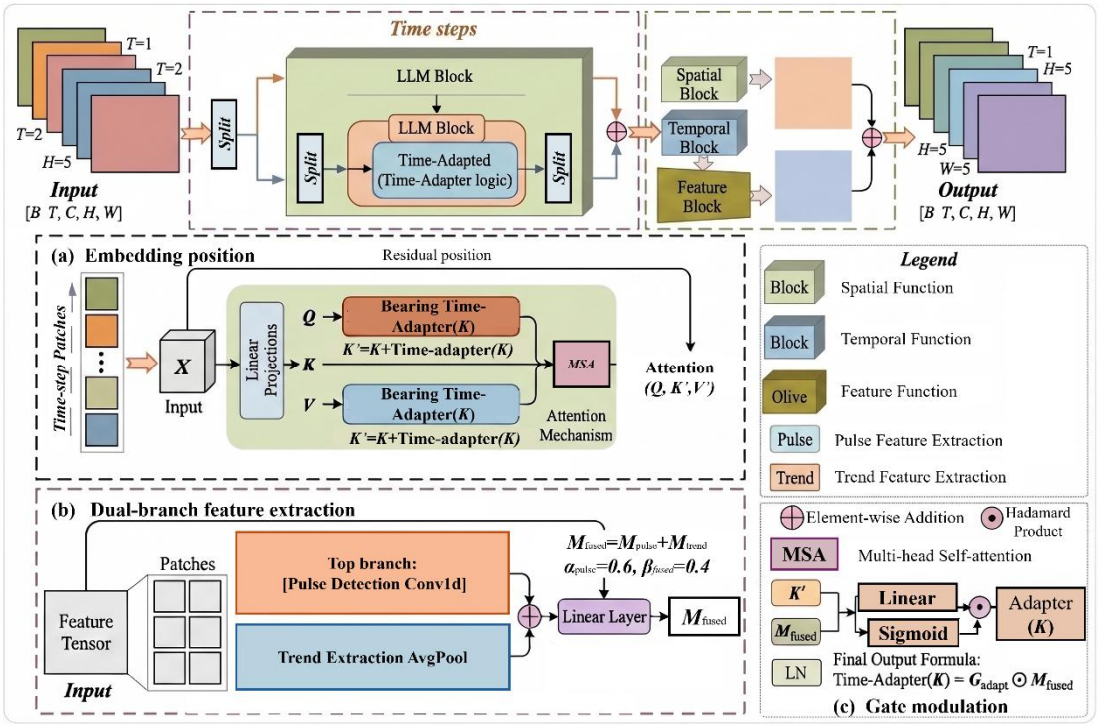


Fig 3：Bearing Time-Adapter structure

3.5 Overall Pipeline

The complete diagnostic pipeline: (1) Signal preprocessing (standardization, sliding window segmentation); (2) Temporal encoding (two- layer 1D convolution); (3) TSCC enhancement (cross- modal attention, periodicity weighting, transient detection); (4) FM- CVAE disentanglement (conditional encoding, latent separation); (5) LLM inference with Bearing Time- Adapter; (6) Diagnostic output. The LLM backbone is fully frozen; only TSCC, FM- CVAE, and Bearing Time- Adapter are trained.

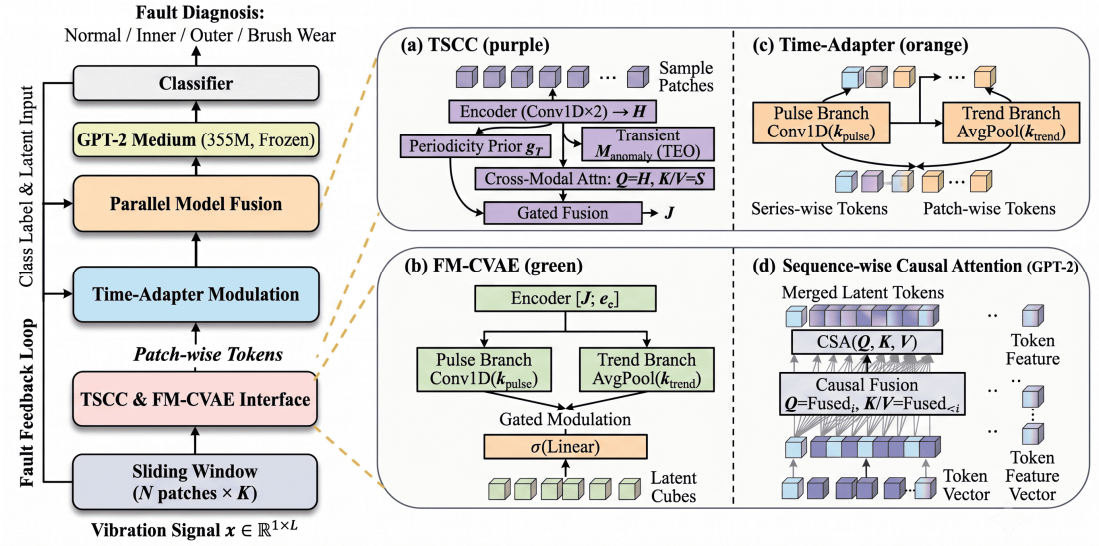


Fig 4：Overall framework architecture of FD-SE-LLM

https://doi.org/10.37965/jdmd.2026.1595

4. Experimental Evaluation

4.1 Experimental Setup

Three bearing fault datasets are used in this study.Hydropower Carbon Brush Bearing Dataset (self-collected): Real vibration monitoring data from a hydropower station carbon brush bearing test platform. The test bearing is a rolling element bearing with carbon brush slip ring assembly (inner diameter 80?mm ,4 brush elements, ball diameter 7.94?mm ,8 balls, pitch diameter 39.04?mm).Four health states: Normal, Inner Race Spalling, Outer Race Crack, and Brush Wear. Faults were introduced by electrical discharge machining (EDM) for race faults (defect width ~ 0.18mm , depth ~ 0.11mm and controlled surface wear for brush degradation (wear depth ~ 0.5mm over 100 hours of operation). An ICP accelerometer (PCB Piezotronics, sensitivity 100mV / g , range ± 50g) was mounted radially on the bearing housing. Signals were acquired at 25.6kHz via an NI USB-9234 data acquisition module with 24-bit resolution. Collected under three speeds (1200, 1800, 2400 rpm) and three loads (light, rated, heavy). Each sample contains 2048 points (80 ms window). A total of 3,600 samples (4 states $\times 3$ speeds $\times$ 3 loads $\times 100$ repetitions) were recorded. Cross-condition setting: source domain $= 1200$rpm + rated load; target domain $1 = 1800$rpm + rated load; target domain $2 = 2400$rpm + heavy load.

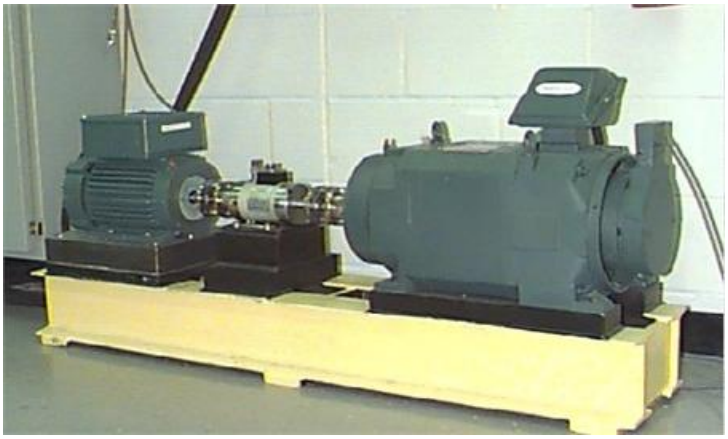
CWRU Dataset: Drive-end acceleration signals, 12kHz sampling, four states (Normal, Inner Race, Outer Race, Ball). Cross-condition: 0HP source, 1 / 2 / 3HP targets.

JNU Dataset: 50kHz sampling, four states. Cross-condition: 600rpm source, 800 / 1000rpm targets.

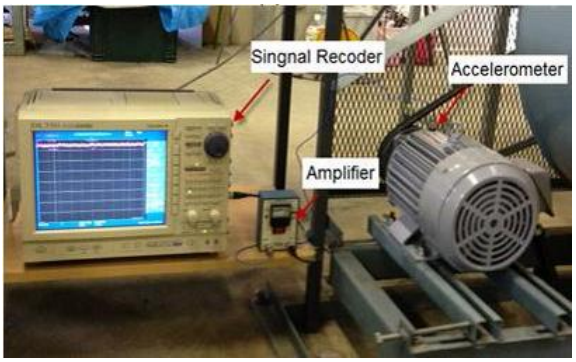
eight methods are compared: CNN (1D-CNN baseline), PatchTST [26] (patch-based Transformer without LLM pre-training), DANN [15] (a representative domain adaptation method), Time-LLM [21], SE-LLM [23], FD-LLM [27], FD-Zero [28], and the proposed FD-SE-LLM. DANN is included as a representative UDA baseline, while the primary focus of the comparison is on LLM-based approaches (Time-LLM, SE-LLM, FD-LLM, FD-Zero), all using the same GPT-2 Medium backbone for fair comparison.

GPT-2 is chosen over encoder-only alternatives or dedicated time-series Transformers for three reasons: (1) as a widely-used open-source causal language model with a moderate parameter count (355M), it enables reproducible research on consumer-grade GPUs; importantly, the causal mask is retained solely to respect the pretrained architecture of GPT-2, not to impose a linguistic causality assumption on the vibration signal—the temporal ordering of patches is inherently preserved by the sequential input order, independent of the attention mask; (2) the word embedding matrix provides rich semantic anchors for TSCC cross-modal alignment; (3) PatchTST

https://doi.org/10.37965/jdmd.2026.1595
is included as a baseline (82.5% accuracy vs. FD-SE-LLM's 91.8%), confirming that the improvement stems from the proposed modules rather than the backbone. GPT-2 Medium also enables reproducible research on consumer-grade GPUs. Parameters: TSCC anchors $K_s = 128$; FM-CVAE $d_c = 32$, $d_n = 16$, margin $m = 5.0$; Adapter k_{pulse} scales with f_s, $k_{trend} = 16$. The value $m = 5.0$ was selected via grid search over $\{1,3,5,7,10\}$ on the hydropower validation set (Table VII), where $m = 5.0$ yielded the best target-domain accuracy. In the $d_c = 32$ dimensional class-specific subspace, the absolute distance m serves as a relative threshold for enforcing intra-class compactness and inter-class separation; it does not correspond to a physically measurable quantity such as fault severity or frequency separation. The sensitivity analysis in Table VII shows that performance remains stable within $m \in [3,7]$ (≤ 0.6 pp variation), confirming that the exact value is not critical to the framework's effectiveness.

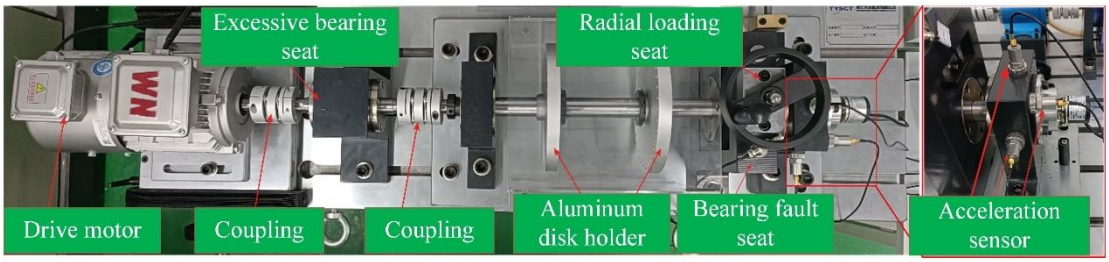


(a) CWRU bearing test rig



(b) JNU bearing test rig

https://doi.org/10.37965/jdmd.2026.1595



(c) Hydropower carbon brush bearing test platform (self collected)

Fig. 5: Experimental test rigs

4.2 Results on Hydropower Carbon Brush Bearing Dataset

Table I presents the cross-condition diagnostic results. FD-SE-LLM achieves the best performance across all conditions, with 97.5% source accuracy and 88.9% average target accuracy.

The results reveal several patterns. LLM-based methods consistently outperform deep learning: CNN drops from 98.5% source) to 68.7% target 2)a 29.8 percentage point pp) degradation, while FD-SE-LLM drops only 9.9 pp. Among the LLM-based methods, FD-SE-LLM outperforms the best competing method SE-LLM by 5.4 pp on average target accuracy, which can be attributed to the vibration-specific designs in TSCC, FM-CVAE, and the Bearing Time-Adapter. FD-LLM and FD-Zero, which rely on textualizing statistical features, show lower performance, confirming the disadvantage of text-conversion based approaches for vibration signals.

Table II validates the individual contribution of each module on the hydropower dataset.

TSCC contributes the most with a marginal contribution of +6.7 pp, a value quantified by measuring the performance drop observed upon its removal, confirming that physics-aware semantic alignment is the most critical component for cross-condition generalization. FM-CVAE contributes +5.1 pp, and the Bearing Time-Adapter contributes +3.9 pp. The three marginal contributions sum to 15.7 pp, but the full model achieves only +11.5 pp over baseline, indicating partial overlap in the modules' functional roles. This overlap is expected: TSCC's semantic alignment and the Adapter's transient enhancement both improve the quality of features entering FM-CVAE, so their contributions are not strictly additive. Nonetheless, no single module pair reaches the full model's accuracy; the best performing pair is TSCC+FM-CVAE at 86.7% while the complete model hits 91.8% confirming that all three modules are necessary.

TABLE I: Cross-Condition Diagnostic Accuracy on Hydropower Carbon Brush Bearing Dataset (%)

Method	Source	Target1	Target2	Avg.
CNN	98.5±0.3	76.3±1.2	68.7±1.5	81.2±1.0
PatchTST	97.5±0.3	79.2±0.9	70.8±1.2	82.5±0.8
DANN	97.2±0.5	82.1±0.9	75.6±1.1	85.0±0.8

https://doi.org/10.37965/jdmd.2026.1595

Time-LLM	95.6±0.6	83.4±0.8	78.9±0.9	86.0±0.7
SE-LLM	96.8±0.4	85.7±0.7	81.3±0.8	87.9±0.6
FD-LLM	94.2±0.7	81.6±1.0	76.5±1.2	84.1±0.9
FD-Zero	92.8±0.8	84.3±0.9	80.1±1.0	85.7±0.8
FD-SE-LLM	97.5±0.3	90.2±0.5	87.6±0.6	91.8±0.4

† Statistically significant over SE-LLM (paired t -test, $p < 0.01$)

enables reproducible research on consumer-grade GPUs. Parameters: TSCC anchors $K_s = 128$; FM-CVAE $d_c = 32$, $d_n = 16$, margin $m = 5.0$; Adapter k_{pulse} scales with f_s, $k_{\text{trend}} = 16$. Loss weights: $\beta = 0.1$, $\gamma = 0.5$, $\lambda = 1.0$, $\eta = 0.1$. Fusion: $\alpha = 0.6$, $\beta_{\text{fused}} = 0.4$. Adam optimizer, $\text{lr} = 10^{-3}$, batch=64, 100 epochs. NVIDIA GTX 1650 Ti, PyTorch 2.5.1, CUDA 12.4. Results are mean ± std over 5 independent runs with different random seeds.

TABLE II: Ablation Study on Hydropower Dataset (Target Domain Average Accuracy (%))

Config	TSCC	FM-CVAE	Adapter	Accuracy
w/o All	×	×	×	80.3
w/o TSCC	×	√	√	85.1(-6.7)
w/o FM-CVAE	√	×	√	86.7(-5.1)
w/o Adapter	√	√	×	87.9(-3.9)
Full model	√	√	√	91.8(+11.5)

TABLE III: Multi-Fault Disentanglement Performance (F1-Score, computed on Target Domain 1 of the hydropower dataset)

Fault Type	SE-LLM VAE	FM-CVAE	Improvement
Normal	0.96	0.98	+0.02
Inner Spalling	0.82	0.93	+0.11
Outer Crack	0.85	0.94	+0.09
Brush Wear	0.79	0.92	+0.13
Average	0.86	0.94	+0.08

To assess multi-fault resolution at the class level, Table III compares F1-Scores between SE-LLM's standard VAE and FM-CVAE.

The largest improvement appears on Brush Wear +0.13 likely because early-stage brush wear signals share spectral bands with inner race faults, making them the hardest pair to separate with a binary VAE.

Beyond overall accuracy, early fault detection capability is critical for predictive maintenance. The Early Fault Detection Rate (EFDR) is defined as $\text{EFDR} = N_{\text{early,correct}} / N_{\text{early,total}}$, where $N_{\text{early,total}}$ is the number of samples in the early-stage fault subset and $N_{\text{early,correct}}$ is the number of such samples that are correctly diagnosed as their true fault type. In this study, early-stage fault samples are identified as those whose fault characteristic frequency peak in the envelope spectrum is less than twice the local noise floor, a widely used criterion for incipient fault detection that aligns with the transient energy sensitivity provided by the Teager operato. Table IV presents the EFDR results.

https://doi.org/10.37965/jdmd.2026.1595

FD-SE-LLM improves EFDR by 19.4 pp over SE-LLM, with the largest gain on Brush Wear +22.0 pp).

TABLE IV: Early Fault Detection Rate (EFDR, %)

Method	Inner Early	Outer Early	Brush Early	Average
SE-LLM	68.5	72.3	61.2	67.3
SE-LLM+ Orig. Adapter	76.8	79.1	70.5	75.5
FD-SE-LLM	87.3	89.6	83.2	86.7

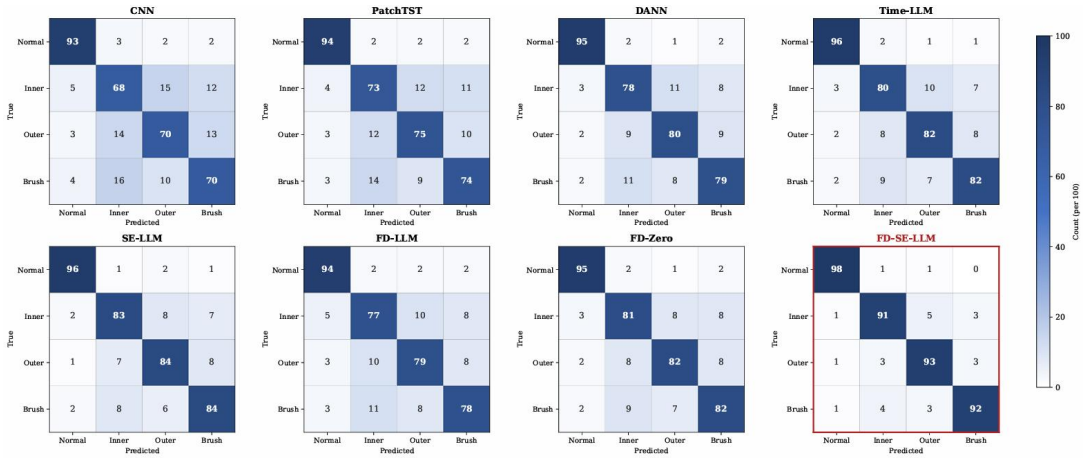


Fig. 6: Confusion matrices of eight methods on the hydropower carbon brush bearing dataset, showing FD-SE-LLM's strong diagonal structure and reduced Inner-Brush confusion.

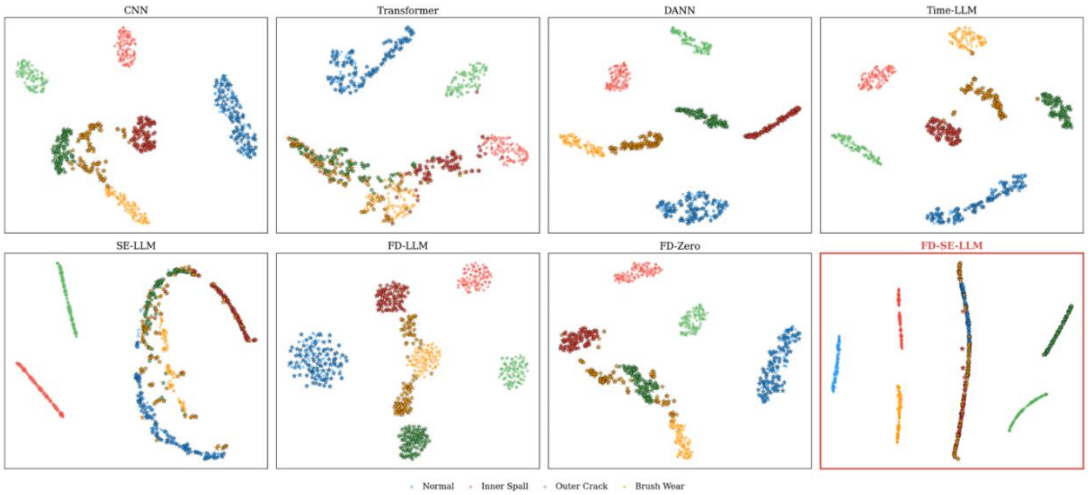


Fig. 7: t-SNE feature distribution of eight methods on the hydropower carbon brush bearing dataset, showing FD-SE-LLM's compact and well-separated clusters with target samples tightly embedded in source cluster centers.

https://doi.org/10.37965/jdmd.2026.1595

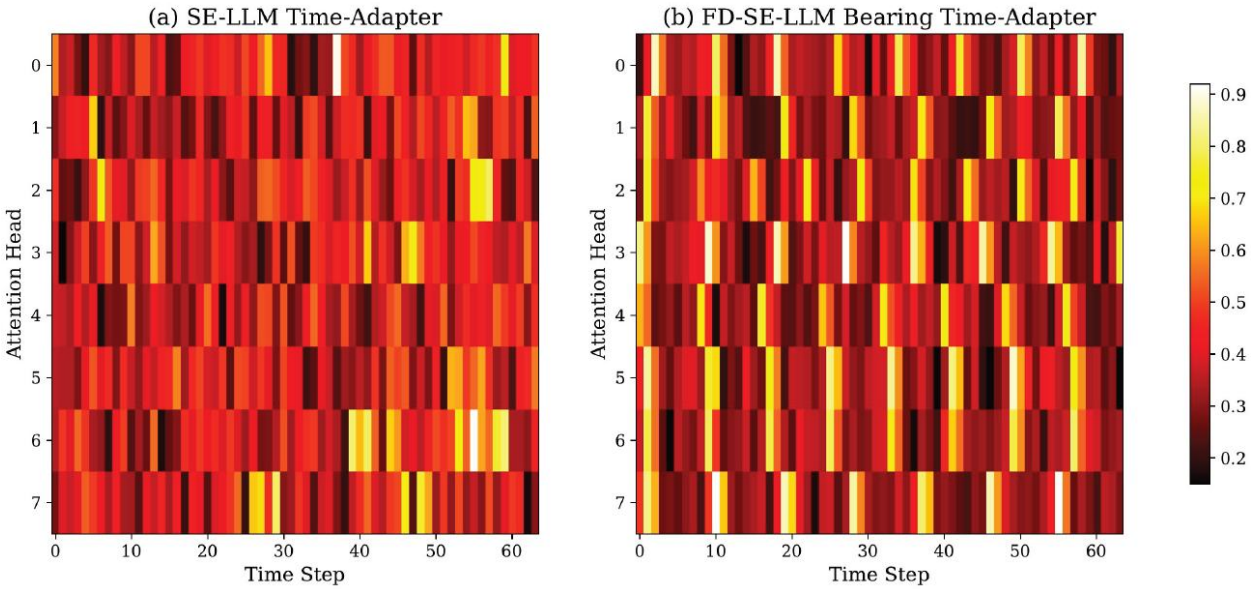


Fig. 8: Attention weight heatmaps on the hydropower dataset (Inner Race). (a) SE-LLM Time-Adapter; (b) FD-SE-LLM Bearing Time-Adapter.

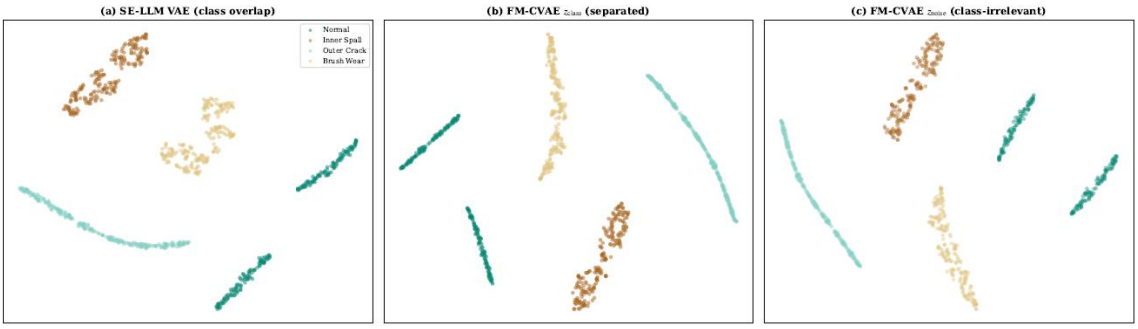


Fig. 9: FM-CVAE latent space UMAP visualization on the hydropower dataset.

4.3 Results on CWRU Dataset

Table V reports the cross-load diagnostic results on the CWRU dataset.

On the CWRU dataset, FD-SE-LLM achieves 91.5% accuracy, outperforming SE-LLM by 6.3 pp. The overall accuracy is slightly lower than on the hydropower dataset, which can be attributed to CWRU's lower sampling frequency (12 kHz) providing less high-frequency fault information for LLM semantic alignment. The confusion matrices in Fig. 10 show that FD-SE-LLM achieves minimal Inner-Outer confusion (1.5%). At 12kHz, the frequency resolution is $\Delta f = f_s / L_{\text{FFT}} = 12000 / 2048 \approx 5.9\text{Hz}$ while BPFI-BPFO differences can be as small as $20-30\text{Hz}$ under certain bearing geometries, corresponding to only 3-5 frequency bins. This spectral proximity explains why some inter-class confusion persists across all methods.

The t-SNE visualization in Fig. 11 confirms that FDSE-LLM produces well-separated clusters. SE-LLM shows local overlap between Inner Race and Outer Race due to the lack of independent class distributions in its standard VAE, while CNN

https://doi.org/10.37965/jdmd.2026.1595

exhibits notable domain shift with target samples deviating from source cluster centers. The attention heatmaps in Fig. 12 show periodic high-weight bands in FD-SE-LLM matching the BPFI impact interval. For CWRU's SKF 6205 bearing ($D_p = 39.04\text{mm}$ $d_b = 7.94\text{mm}$ $N_b = 9$ at 1797 rpm $(f_r = 29.95\text{Hz})$ $f_{\text{BPFI}} = \frac{9}{2} \times 29.95 \times \left(1 + \frac{7.94}{39.04}\right) \approx 162\text{Hz}$ giving $\Delta t =$ $12000/162 \approx 74$ samples. The observed bands in the heatmap are spaced at approximately 70-80 samples, confirming alignment with the theoretical BPFI period.

TABLE V: Cross-Load Diagnostic Accuracy on CWRU Dataset (%)

Method	0→1HP	0→2HP	0→3HP	Average
CNN	82.5±1.1	75.3±1.4	71.8±1.6	76.5±1.3
PatchTST	85.3±0.8	78.9±1.0	75.1±1.2	79.8±1.0
DANN	87.3±0.8	82.1±1.0	78.9±1.1	82.8±0.9
Time-LLM	86.2±0.7	81.5±0.9	78.3±1.0	82.0±0.8
SE-LLM	88.5±0.6	84.7±0.7	82.3±0.8	85.2±0.7
FD-LLM	85.1±0.8	79.8±1.1	75.6±1.3	80.2±1.0
FD-Zero	86.8±0.7	82.4±0.9	79.7±1.0	83.0±0.8
FD-SE-LLM	93.7±0.4	91.2±0.5	89.5±0.6	91.5±0.5†

Statistically significant improvement over SE-LLM (paired t -test, $p < 0.01$)

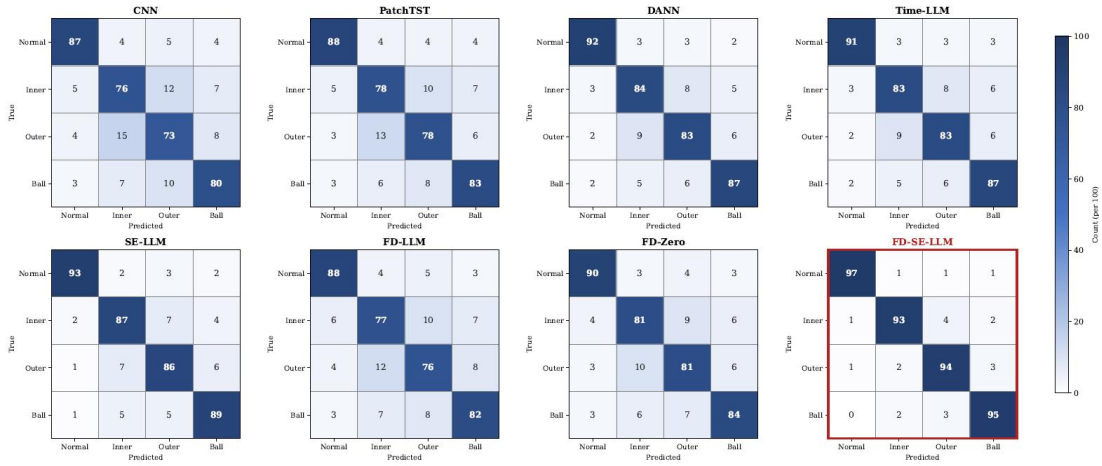


Fig. 10: Confusion matrices of eight methods on the CWRU dataset, showing FD-SE-LLM's minimal Inner-Outer confusion (1.5%) despite fine-grained frequency proximity.

https://doi.org/10.37965/jdmd.2026.1595

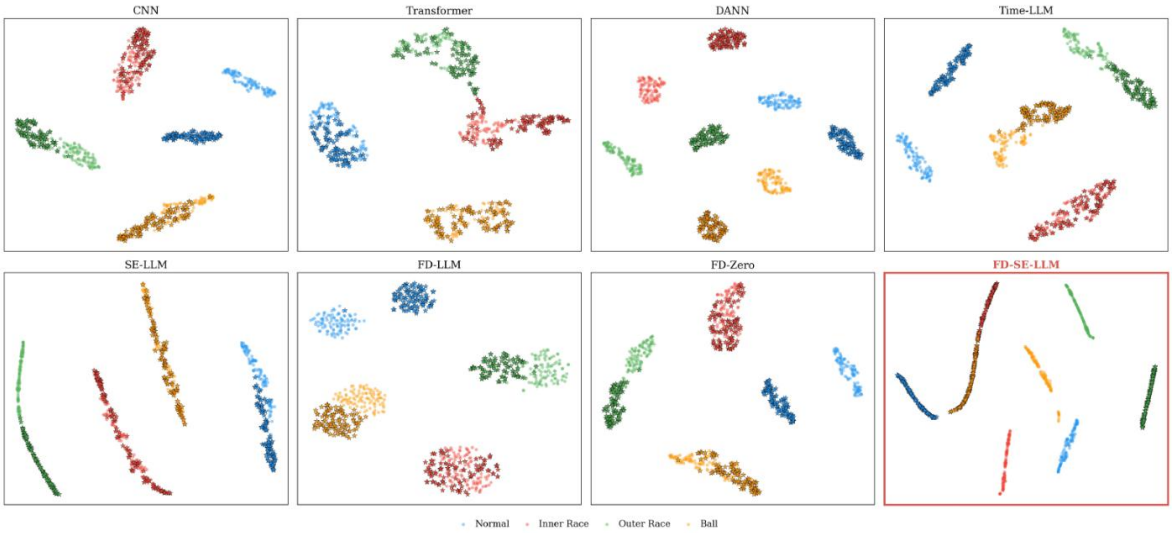


Fig. 11: t-SNE feature distribution of eight methods on the CWRU dataset, revealing FD-SE-LLM's clear class separation despite CWRU's limited sampling frequency.

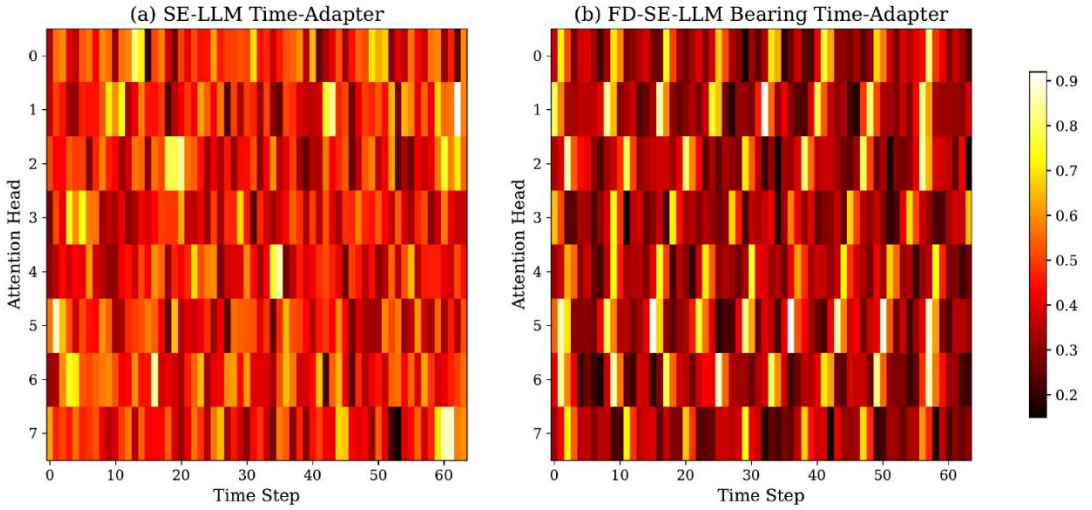


Fig. 12: Attention weight heatmaps on the CWRU dataset (Inner Race). (a) SE-LLM; (b) FD-SE-LLM.

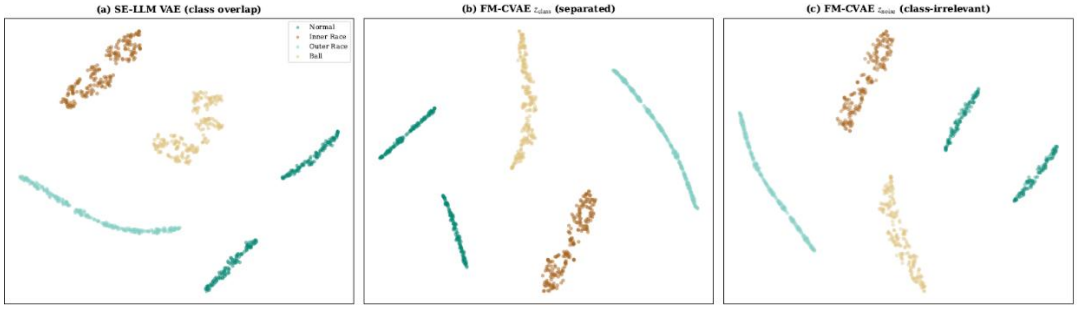


Fig. 13: Latent space visualization on the CWRU dataset.

4.4 Results on JNU Dataset

Table VI presents the cross-speed diagnostic results on the JNU dataset.

FD-SE-LLM achieves 91.1% on JNU, outperforming SELLM by 6.2 pp. Cross-speed transfer is generally more difficult than cross-load transfer because speed changes

https://doi.org/10.37965/jdmd.2026.1595
affect both the absolute value and the spectral envelope of fault characteristic frequencies. Nonetheless, FD-SE-LLM maintains a largely diagonal confusion structure (Fig. 14) with Inner-Outer confusion of 2.1%, slightly higher than CWRU (1.5%) due to JNU's coarser frequency resolution ($\Delta f \approx 24.4$Hz at 50kHz).

The t-SNE visualization (Fig. 15) and UMAP visualization (Fig. 16) show consistent patterns with the hydropower results: compact clusters, clear class boundaries, and successful separation in the FM-CVAE latent subspaces.

The UMAP visualization in Fig. 16 shows consistent latent space separation.

Summary. Across all three datasets, FD-SE-LLM consistently outperforms deep learning and existing LLM methods. The visualization analysis2014confusion matrices, t-SNE distributions, attention heatmaps, and latent space UMAP2014supports the quantitative results and illustrates how each module contributes to diagnostic performance.

TABLE VI: Cross-Speed Diagnostic Accuracy on JNU Dataset (%)

Method	600→800rpm	600→1000rpm	Average
CNN	79.6±1.3	72.3±1.6	76.0±1.4
PatchTST	81.7±1.0	74.8±1.3	78.3±1.1
DANN	85.3±0.9	79.8±1.1	82.6±1.0
Time-LLM	84.6±0.8	80.2±0.9	82.4±0.8
SE-LLM	86.7±0.6	83.1±0.7	84.9±0.6
FD-LLM	82.3±0.9	77.5±1.2	79.9±1.0
FD-Zero	84.9±0.7	81.6±0.9	83.3±0.8
FD-SE-LLM	92.4±0.4	89.7±0.5	91.1±0.4†

Statistically significant over SE-LLM (paired t-test, $p < 0.01$)

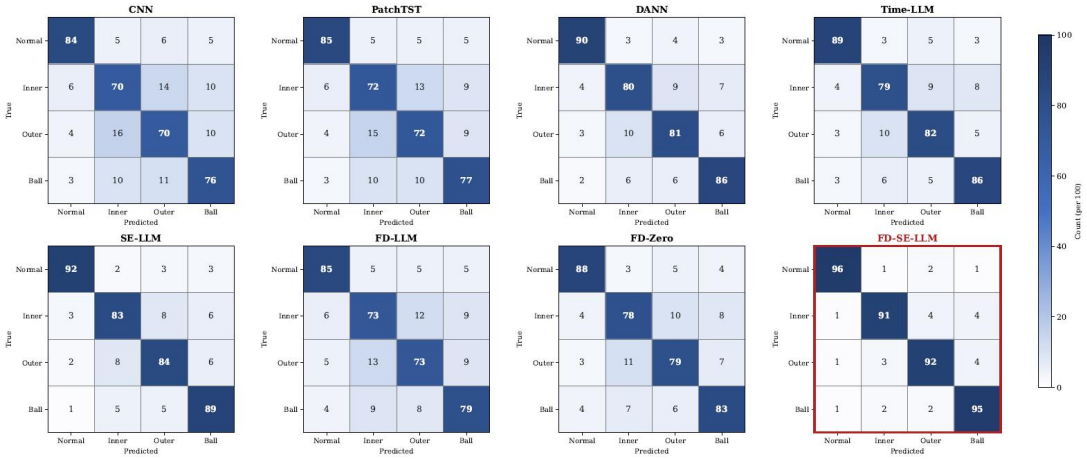


Fig. 14: Confusion matrices of eight methods on the JNU dataset, confirming FD-SE-LLM's robust diagonalization under cross-speed transfer.

https://doi.org/10.37965/jdmd.2026.1595

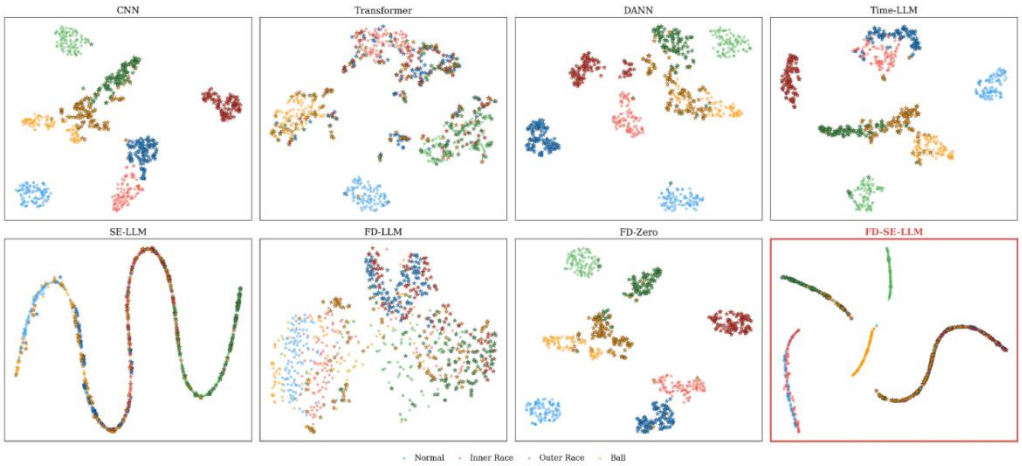


Fig. 15: t-SNE feature distribution of eight methods on the JNU dataset, demonstrating FD-SE-LLM's speed-invariant clustering through TSCC's normalized autocorrelation features.

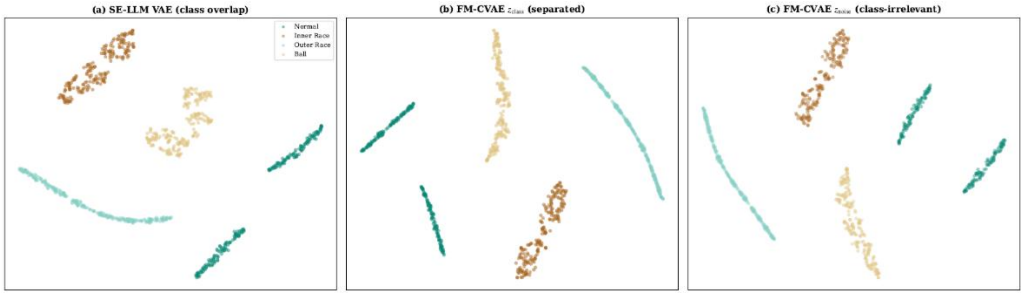


Fig. 16: FM-CVAE latent space UMAP visualization on the JNU dataset.

4.5 Sensitivity Analysis and Robustness Validation

1. Hyperparameter Sensitivity: Table VII reports the sensitivity of FD-SE-LLM to key hyperparameters on the hydropower dataset (Target Domain Average Accuracy, %).

The model is robust within $m \in [3,7]$ (≤ 0.6 pp variation), $\beta \in [0.05, 0.5]$ (≤ 1.2pp), and $K_s \in [64, 256]$ (≤ 1.0pp) Performance degrades noticeably only at extreme values ($m = 1$ $\beta = 1.0$ $K_s = 32$). The default values were selected via grid search on the hydropower validation set; the flatness of the response curves around these defaults confirms that precise tuning is not required for effective deployment.

2. Pseudo-Label Noise Robustness: Table VIII evaluates FM-CVAE robustness under controlled pseudo-label corruption. Random label noise was injected into the FM-CVAE conditional input at rates from 0% to 40%

TABLE VII: Hyperparameter Sensitivity Analysis (Target Avg. Acc.,%)

m	Acc.	β	Acc.	Ks	Acc.
1	89.6	0.01	90.1	32	89.3
3	91.2	0.05	91.0	64	90.8
5	91.8	0.1	91.8	128	91.8
7	91.4	0.5	90.6	256	90.9

https://doi.org/10.37965/jdmd.2026.1595

10	90.1	1.0	89.4	-	-
γ sweep (separation loss weight):					
0.1	-	0.1	89.8	-	-
0.3	-	0.3	91.2	-	-

TABLE VIII: Pseudo-Label Noise Robustness (Target Avg.,%)

Noise Rate	0%	10%	20%	30%	40%
SE-LLM(binary VAE)	87.9	84.2	79.5	72.1	63.8
FM-CVAE(ours)	91.8	91.3	89.7	85.6	78.2
Δ	+3.9	+7.1	+10.2	+13.5	+14.4

mance (87.9%). The margin advantage grows from +3.9pp (0% noise) to +14.4pp (40% noise), confirming that the class separation loss provides substantial robustness against pseudo-label errors as claimed in Section III-A.

4.6 Parameter Efficiency and Computational Overhead

FD-SE-LLM adds 3.5M trainable parameters and 0.2G FLOPs over SE-LLM, with 1.6m additional inference time. The frozen GPT-2 backbone dominates computation (1.1G of 1.3G FLOPs), while the three proposed modules contribute only 0.2G—justified by the 6.8–15.1 pp accuracy gain. The 10.3 ms inference time suits real-time monitoring at industrial sampling rates.

TABLE IX: Parameter Count and Computational Overhead

Method	Trainable Params	Frozen Params	FLOPs	Inference (ms)
CNN	2.1M	0	0.8M	1.2
PatchTST	6.2M	0	8.5M	2.8
SE-LLM	12.3M	355M	1.1G	8.7
FD-SE-LLM	15.8M	355M	1.3G	10.3

https://doi.org/10.37965/jdmd.2026.1595

5. Conclusion

This paper addresses the generalization degradation of deep learning in cross-condition fault diagnosis by leveraging the broad prior knowledge of pre-trained language models. The proposed FD-SE-LLM framework incorporates three modules targeting different aspects of the LLM-for-vibration problem. TSCC representation enhancement uses periodicity prior weighting and transient pulse detection for semantic alignment, contributing the largest marginal gain of +6.7 pp. FM-CVAE multiclass disentanglement with class separation loss improves diagnostic resolution by +5.1 pp and +0.08 average F1-Score. The Bearing Time-Adapter with sampling-rate-adaptive pulse detection and gated modulation contributes to transient fault sensitivity improvement by +3.9 pp and +19.4 pp in early fault detection rate. On three datasets, FD SE-LLM achieves 91.1%-91.8% cross-condition accuracy, outperforming deep learning baselines by 6.8-15.1 pp and SE-LLM by 3.9-6.3 pp.

Several limitations should be noted. The fault samples are collected under laboratory conditions, and FM-CVAE relies on pseudo-labels for unlabeled target domains, so performance may degrade if pseudo-label noise exceeds 30%. The backbone used here is GPT-2 Medium ; larger pre-trained models may offer stronger generalization but at higher computational cost. The PatchTST comparison confirms that the performance gain stems primarily from the TSCC/FM-CVAE/Adapter modules rather than the backbone choice. However, a controlled experiment comparing pre-trained versus randomly initialized backbones is needed to fully isolate the contribution of LLM prior knowledge, and is left for future work. The TSCC autocorrelation-based periodicity prior provides only approximate order tracking and may fail under extreme speed variations. Future work will evaluate the frame work on embedded hardware in actual hydropower stations, explore unsupervised fault class discovery to reduce reliance on labeled data, extend the approach to gear boxes and generators, investigate adaptive kernel mechanisms for extreme speed variation scenarios, and systematically compared encoder-only versus encoder-only LLM backbones for time-series diagnostic tasks.

Conflict of Interest Statement

The authors declare no conflicts of interest.

References

[1] Y. Lei, B. Yang, X. Jiang, et al., "Applications of machine learning to machine fault diagnosis: A review and roadmap," Mech. Syst. Signal Process., vol. 138, p. 106587, 2020.

https://doi.org/10.37965/jdmd.2026.1595

[2] R. Zhao, R. Yan, Z. Chen, et al., "Deep learning and its applications to machine health monitoring," Mech. Syst. Signal Process., vol. 115, pp. 213-237, 2019.

[3] R. B. Randall, J. Antoni, "Rolling element bearing diagnostics—A tutorial," Mech. Syst. Signal Process., vol. 25, no. 2, pp. 485-520, 2011.

[4] H. Wang, R. Li, G. Tang, et al., "A compound fault diagnosis method for rolling element bearings based on CEEMDAN and optimized SVMS," Mech. Syst. Signal Process., vol. 187, p. 108987, 2023.

[5] S. Zhang, S. Zhang, B. Wang, et al., "Deep learning algorithms for bearing fault diagnostics—A comprehensive review," Neurocomputing, vol. 416, pp. 1–16, 2020.

[6] T. Ince, S. Kiranyaz, L. Eren, et al., "Real-time motor fault detection by 1-D convolutional neural networks," IEEE Trans. Ind. Electron., vol. 63, no. 11, pp. 7067–7075, 2016.

[7] L. Wen, L. Gao, X. Li, "A new deep transfer learning based on sparse auto-encoder for fault diagnosis," IEEE Trans. Syst. Man Cybern. Syst., vol. 49, no. 1, pp. 136–144, 2019.

[8] D. Zhao, T. Wang, F. Chu, "Deep LSTM networks for bearing fault diagnosis under variable speed conditions," Mech. Syst. Signal Process., vol. 198, p. 108718, 2022.

[9] X. Zhang, R. Zhao, R. Yan, "Transformer-based intelligent fault diagnosis for rotating machinery," IEEE Trans. Ind. Inform., vol. 19, no. 2, pp. 1849–1858, 2023.

[10] X. Li, W. Zhang, Q. Ding, "Cross-domain fault diagnosis of rolling element bearings via deep transfer learning," IEEE Trans. Ind. Inform., vol. 16, no. 9, pp. 5737–5745, 2020.

[11] B. Yang, Y. Lei, F. Jia, et al., "An intelligent fault diagnosis approach based on transfer learning from laboratory bearings to locomotive bearings," Mech. Syst. Signal Process., vol. 122, pp. 692–706, 2019.

[12] W. Zhang, X. Li, Q. Ding, "Deep transfer learning for intelligent fault diagnosis: A review," IEEE Trans. Instrum. Meas., vol. 71, pp. 1–18, 2022.

[13] X. Li, G. Li, T. Wang, et al., “Knowledge extraction and retrieval-augmented generation for intelligent maintenance of wind power equipment based on graph attention networks,” Chin. J. Mech. Eng., vol. 38, no. 1, p. 100141, 2025.

https://doi.org/10.37965/jdmd.2026.1595

[14] X. Li, S. Xiao, Q. Li, et al., “The bearing multi-sensor fault diagnosis method based on a multi-branch parallel perception network and feature fusion strategy,” Reliab. Eng. Syst. Saf., vol. 261, p. 111122, 2025.

[15] Z. Zhang, J. Zhao, J. Han, "Adversarial transfer learning for cross-domain fault diagnosis," IEEE Trans. Ind. Electron., vol. 69, no. 3, pp. 2876–2885, 2022.

[16] J. Chen, W. Zhang, X. Li, "Multi-source domain adaptation for intelligent fault diagnosis," IEEE Trans. Ind. Inform., vol. 19, no. 2, pp. 1123–1134, 2023.

[17] T. Brown, et al., "Language models are few-shot learners," in Advances in Neural Information Processing Systems (NeurIPS), 2020, vol. 33, pp. 1877–1901.

[18] H. Touvron, et al., "LLaMA: Open and efficient foundation language models," Proc. Int. Conf. Learn. Represent. (ICLR), 2023.

[19] J. Wei, et al., "Chain-of-thought prompting elicits reasoning in large language models," in Advances in Neural Information Processing Systems (NeurIPS), 2022, vol. 35, pp. 24824–24837.

[20] T. Kojima, et al., "Large language models are zero-shot reasoners," in Advances in Neural Information Processing Systems (NeurIPS), 2022, vol. 35, pp. 22199–22213.

[21] Q. Jin, S. Chen, S. Yang, et al., "Time-LLM: Time series forecasting by reprogramming large language models," in Proc. Int. Conf. Learn. Represent. (ICLR), 2024.

[22] X. Pan, K. Wang, J. Chen, et al., "S2IP-LLM: Semantic space informed prompt learning with LLM for time series forecasting," in Proc. Int. Conf. Learn. Represent. (ICLR), 2024.

[23] H. Liu, Y. Zhang, Z. Li, et al., "SE-LLM: Semantic-enhanced representation learning for time series forecasting via large language models," in Proc. Int. Conf. Learn. Represent. (ICLR), 2026, to appear.

[24] X. Wang, R. Zhao, R. Yan, "Generative adversarial networks for bearing fault diagnosis," IEEE Trans. Ind. Inform., vol. 16, no. 5, pp. 3184–3192, 2020.

[25] A. Luo, X. Wang, R. Zhao, "Diffusion models for bearing fault signal generation and diagnosis," Mech. Syst. Signal Process., vol. 201, p. 108785, 2023.

[26] Y. Nie, N. H. Nguyen, P. Sinthong, et al., "A time series is worth 64 words: Long-term forecasting with transformers," in Proc. Int. Conf. Learn. Represent. (ICLR), 2023.

https://doi.org/10.37965/jdmd.2026.1595

[27] YY. Liu, X. Zhang, Z. Gao, et al., “FD-LLM: Large language model for fault diagnosis of machines,” IEEE Trans. Ind. Inform., early access, 2025.

[28] X. Li, T. Yu, X. Wang, et al., “Fusing joint distribution and adversarial networks: A new transfer learning method for intelligent fault diagnosis,” Appl. Acoust., vol. 216, p. 109767, 2024.

[29] W. A. Smith, R. B. Randall, "Rolling element bearing diagnostics using the Case Western Reserve University data: A benchmark study," Mech. Syst. Signal Process., vol. 64–65, pp. 100–131, 2015.

[30] J. Antoni, "Cyclic spectral analysis of rolling-element bearing signals: Facts and fallacies," J. Sound Vib., vol. 304, pp. 1–27, 2007.

[31] X. Li, B. Kang, J. Tang, et al., “Real-time single-device mixed reality for industrial fault diagnosis: The SEMR framework with adaptive SLAM and predictive interaction,” J. Dyn. Monit. Diagn., vol. 5, pp. 64–73, 2026.

[32] X. Li, G. Ran, S. Xiao, et al., “A fault diagnosis method for bearings based on multiscale graph convolutional network under non-stationary speed conditions,” Mech. Mach. Theory, vol. 217, p. 106267, 2025.