

DOI: 10.19666/j.rlfd.202605041

基于原型度量与混合状态空间模型的 锅炉跨工况少样本故障诊断

马良玉, 孟文龙, 薛博元, 张驰

(华北电力大学控制与计算机工程学院, 河北 保定 071003)

[摘要] 【目的】鉴于火电机组深度调峰变工况下, 锅炉早期微弱泄漏诊断面临特征分布漂移与样本极度匮乏的双重挑战, 而现有深度学习诊断模型在跨域迁移场景中常因特征流形坍塌与严重过拟合而失效, 为此提出一种融合时序多尺度解耦与原型度量的跨工况少样本故障诊断方法 Proto-MixMamba。【方法】基于某 600 MW 超临界机组高精度仿真数据构建变负荷工况数据集, 改变传统的单尺度序列建模与参数固化分类范式, 建立适配多变量时间序列的 1D-PAM 与 Mamba 双支路骨干网络。该网络结合一维位置注意力机制与选择性状态空间模型, 在保留多传感器物理耦合的同时, 以线性复杂度高效解耦并捕获泄漏信号的局部微弱瞬态冲击与全局低频演化特征。为解决跨工况时参数幅值基准偏移带来的特征混淆, 引入 N-way K-shot 情景式元学习策略, 并采用原型度量层替代传统全连接分类器, 利用相对欧氏距离构建类原型, 解决变负荷掩盖下物理位置对称泄漏等细粒度相似故障决策边界模糊的问题。【结果】实验结果表明: 该方法有效克服了左右侧省煤器等对称位置泄漏的误判混淆; 在 600 MW 至 480 MW 深度调峰跨域任务且仅有 1-shot 极限条件下, 诊断准确率达到 88.4%, 显著优于 LSTM 等传统基准模型, 有效避免了小样本导致的严重过拟合。【结论】所提模型能够有效提取抗变工况干扰的不变故障特征, 为工业现场少标签环境下的锅炉“四管”故障精准诊断提供了一条可行的技术路线。

[关键词] 故障诊断; 跨工况; 少样本学习; Mamba; 原型度量

[引用本文格式] (本段作者勿动) 作者姓名. 中文标题[J]. 热力发电, 2026, 55(1): 起始页码-终止页码. 作者姓名. 英文标题[J]. Thermal Power Generation, 2026, 55(1): 起始页码-终止页码.

Cross-condition few-shot fault diagnosis of boilers based on prototypical metric and mixed state-space model

Ma Liangyu, Meng Wenlong, Xue Boyuan, Zhang Chi

(School of Control and Computer Engineering North China Electric Power University, Baoding 071003, China)

Abstract: [Objective] With the increasing integration of intermittent renewable energy sources into the power grid, large-capacity thermal power units are frequently required to participate in deep peak-shaving operations. This operational transition leads to frequent and drastic changes in boiler load, which significantly intensifies the thermal stress on heating surfaces and increases the risk of boiler "four-tube" leakages. However, early-stage weak leakage diagnosis faces two critical challenges: feature distribution drift caused by variable-load conditions and an extreme scarcity of fault samples in industrial sites. Traditional deep learning-based diagnostic models, which rely on fixed classification heads, often suffer from feature manifold collapse and severe overfitting when transferred from base-load to deep peak-shaving scenarios. To address these issues, this paper proposes a cross-condition few-shot fault diagnosis framework named Proto-MixMamba, which integrates time-series multi-scale decoupling

收稿日期: 2026-05-15 修回日期: 2026-06-15 接受日期: 2026-06-16 网络首发日期: XXXX-XX-XX

基金项目: 国家自然科学基金项目 (61963117); 河北省中央引导地方科技发展资金项目 (226Z2103G)

Supported by: National Natural Science Foundation of China (61963117); Central Government Guides Local Science and Technology Development Fund of Hebei Province (226Z2103G)

第一作者简介: 马良玉 (1972), 男, 博士, 教授, 主要研究方向为智能技术在火电机组和新能源系统建模、优化控制及故障诊断中的应用, maliangyu@ncepu.edu.cn.

通信作者简介: 孟文龙 (2002), 男, 硕士研究生, 主要研究方向为发电机组故障诊断, mwlong6620@outlook.com.

with prototypical metric learning. [Methods] The research utilizes high-precision simulation data from a 600 MW supercritical unit to construct comprehensive datasets covering various load conditions and leakage severities. To move beyond traditional single-scale sequence modeling and parameter-solidified classification paradigms, a dual-branch backbone network comprising a One-Dimensional Position Attention Module (1D-PAM) and a Mamba State Space Model is established. Specifically, the 1D-PAM branch is designed to capture local weak transient shocks and high-frequency fluctuations in thermal parameters, while the Mamba branch leverages its selective scanning mechanism to model global low-frequency evolution trends and long-term dependencies with linear computational complexity. These multi-scale features are adaptively integrated through a gating fusion mechanism that balances the importance of local and global information. To resolve the feature confusion caused by amplitude baseline shifts during load variations, the framework introduces an N-way K-shot episodic meta-learning strategy. A prototypical metric layer is employed to replace the conventional fully connected classifier. By calculating the relative Euclidean distance between query samples and class prototypes in a high-dimensional embedding space, the model constructs a robust decision boundary that is invariant to absolute parameter magnitudes. [Results] Extensive experimental evaluations demonstrate the superiority of the Proto-MixMamba model. The results indicate that the proposed method effectively overcomes the long-standing industrial challenge of misjudging physically symmetrical leakage locations, such as the left and right economizers, which exhibit highly similar thermal responses under variable-load interference. In the challenging cross-condition diagnosis task transitioning from 600 MW to 480 MW (deep peak-shaving), and under the extreme constraint of the 1-shot setting (only one reference sample per category), the proposed model achieves a diagnostic accuracy of 88.4%. This performance significantly outperforms traditional baseline models such as Long Short-Term Memory (LSTM) and ResNet, whose accuracies drop below 50% due to severe overfitting and domain shift. Visualization via t-SNE confirms that the prototypical metric mechanism successfully maintains high intra-class compactness and inter-class separability in the target domain.[Conclusion] The proposed Proto-MixMamba framework demonstrates exceptional robustness and generalization capabilities in cross-condition few-shot scenarios. By successfully extracting invariant fault features and establishing a reliable metric space, this research provides a technically feasible and highly effective route for the accurate early warning and diagnosis of boiler "four-tube" leakages in label-scarce and operationally complex industrial environments.

Key words: fault diagnosis; cross-condition; few-shot learning; Mamba; prototypical metric

随着新能源电力系统的飞速发展,风能、太阳能等具有间歇性、不稳定性的电源容量在电网中的比重持续攀升,大容量火电机组越来越承担着保障电网安全稳定运行的关键任务,参与电网调频及深度调峰已成为常态^[1]。频繁的深度变负荷运行加剧了锅炉受热面的热应力交变,使得水冷壁、过热器、再热器及省煤器(简称“四管”)等泄漏故障发生概率大大增加。若能利用某一工况下积累的少样本数据,实现跨工况场景下的四管泄漏故障识别与精确定位,对于避免泄漏事故扩大,降低由此带来的巨大经济损失和安全隐患,具有重要的工程现实意义^[2]。

随着工业大数据不断积累和机器学习、深度学习等人工智能技术的突飞猛进,基于数据驱动的故障诊断技术得到进一步发展。雷亚国等人^[3]提出的基于深度归一化卷积神经网络(deep normalized convolutional neural network, DNCNN)的方法,通过引入归一化层加速模型收敛,成功应用到行星齿轮箱的健康监测当中,实现了复杂工况下的特征自适应提取。Ramezani等人^[4]提出了基于深度神经网络的锅炉声发射信号检测模型,有效实现了管壁泄

漏的自动识别。Khalid等人^[5]设计了一种混合深度学习框架,验证了对多变量热工参数的建模能力,显著增强了模型的特征提取效率。但上述方法需要大量样本数据进行模型训练,而限于故障的复杂性和偶发性,实际生产过程中往往难以获取足量的故障样本数据,严重限制了数据驱动故障诊断方法在现场的实际应用^[6]。

为克服故障样本极度匮乏的困境,少样本学习(Few-Shot Learning)已成为跨工况故障诊断领域的研究热点^[7-8]。现有研究主要从数据增强与特征分布对齐两个维度展开^[9]。在数据扩充方面,马良玉等人^[10]提出一种基于改进生成对抗网络的故障样本扩充方法,解决传统GAN训练不稳与生成质量差的问题,显著提升诊断精度。在特征对齐方面,Li等人^[11]构建了无源渐进式域自适应网络;Hu等人^[12]提出了基于元学习与域对抗图卷积的跨域故障诊断模型;Wang等人^[13]则构建了基于自监督信号表示学习的少样本诊断框架,该框架通过深度挖掘时序信号的内在结构表征,有效降低了模型对大规模先验标签的依赖,显著提升了变工况下向目标域的知识迁移与泛化能。然而,面对火电机组多变

量、强耦合的热工时间序列，上述方法难以有效解耦其中复杂的物理机理，导致跨工况诊断性能严重退化。针对跨工况特征漂移这一瓶颈，原型网络（Prototypical Networks）作为一种高效的度量学习新范式脱颖而出^[14-15]。该方法摒弃了复杂的分布对齐或伪样本生成策略，假设同类样本在嵌入空间中应紧密聚集，通过计算支持集各类样本的特征均值动态构造类原型，并依据查询样本与各原型间的距离分布实现分类。在此基础上，An 等人^[16]通过引入原型对比学习机制，进一步优化了少样本跨工况场景下的特征边界。这种原型度量机制能够在深层流形空间中有效滤除变工况引入的幅值漂移噪声，同时施加“同类紧致、异类排斥”的局部拓扑约束。因此，若能将这一强大的原型度量思想与具备全局长程依赖捕获能力的深层不变特征提取技术相融合，将为突破火电机组跨工况细粒度故障辨识难题提供一条极具潜力的解决路径。

在序列建模领域，为了降低 Transformer 的二次计算复杂度，Gu 等人^[17]提出的 Mamba 架构凭借线性计算复杂度与卓越的全局建模能力，为处理长序列工业信号提供了新范式。基于此，Than 等人^[18]率先尝试将 Mamba 引入轴承小样本故障诊断，提出了 MixMamba 架构，利用混合注意力机制在轴承故障诊断中取得了较好的效果。但是，将 MixMamba 方法直接应用于锅炉四管泄漏故障诊断仍然存在严峻挑战，问题在于数据模态不匹配以及物理耦合缺失。由于 MixMamba 是为 2D 时频图像打造的，而四管泄漏监测的数据是包含温度、压力、流量等诸多测点的多变量时间序列。对其强行进行图像转换，不仅导致计算冗余，还使得多传感器之间原本固有的物理耦合性以及时间同步性遭到破坏。在四管泄漏的情况下，比如高温过热器与高温再热器这类相邻受热面^[19]，其故障特征呈现出高度相似的状态，故障的区分存在困难，并且该方法用于跨工况诊断会产生严重的特征混淆。

考虑 Mamba 在长序列建模上的底层优势，并为了克服现有 MixMamba 在多变量跨域下的局限性，本文提出一种基于原型度量与改进 MixMamba 的小样本故障诊断方法（Prototypical Mixed Mamba, Proto-MixMamba），并应用于锅炉四管泄漏故障诊断。实验表明，该方法可大幅提升模型在少样本跨工况条件下对四管泄漏故障的辨识精度。

1 算法原理

针对火电机组深度调峰工况下故障样本稀缺、特征分布漂移显著的跨工况少样本诊断难题，本文提出了 Proto-MixMamba 诊断框架，该框架摒弃了将一维信号转换为二维图像的传统模式，构建了适配一维多变量时间序列的 1D-PAM-Mamba 骨干网络，并通过自适应门控融合机制和原型度量网络，该模型能够在有限样本条件下，实现对各类故障的高精度辨识。

1.1 状态空间模型与 Mamba 机制

Mamba 的核心基础是连续状态空间模型（state space model, SSM）^[20]。该模型通过隐状态变量 $h(t)$ 将一维输入 $x(t)$ 映射为输出 $y(t)$ 。为适配火电机组离散时间序列，引入步长参数 Δ 并采用零阶保持（Zero-Order Hold, ZOH）技术进行离散化，SSM 转化为可通过线性递推高效计算的形式：

$$h_t = \bar{A}h_{t-1} + \bar{B}x_t, \quad y_t = Ch_t \quad (1)$$

式中： $C \in \mathbb{R}^{1 \times N}$ 为输出投影矩阵； $\bar{A}$ 与 $\bar{B}$ 为离散化状态转移矩阵与输入投影矩阵，计算公式为 $\bar{A} = \exp(\Delta A)$ ， $\bar{B} = (\exp(\Delta A) - I)A^{-1}B$ ，其中 A, B 为连续参数矩阵， Δ 为时间步长， I 为单位矩阵。

传统 SSM 受限于线性时不变特性，参数在所有时间步保持固定。为此，Mamba 引入选择性扫描（Selective Scan）机制，将核心参数转化为输入 x_t 的函数：

$$B_t = W_B x_t, \quad C_t = W_C x_t \quad (2)$$

$$\Delta_t = \text{Softplus}(b_\Delta + W_\Delta x_t) \quad (3)$$

式中： B_t, C_t, Δ_t 为输入依赖的动态参数； W_B, W_C, W_Δ 为可学习投影权重矩阵； b_Δ 为可学习的步长偏置。输入依赖的步长 Δ_t 直接控制动态离散化矩阵 $\bar{A}_t$ 与 $\bar{B}_t$ ：当系统演化发生显著变化时，较大的 Δ_t 驱动模型刷新隐状态以锁定关键信息；而在稳态噪声下，较小的 Δ_t 使系统维持历史状态。该机制在保持线性计算复杂度的前提下，实现了对热工长序列中退化趋势的高效捕获^[21]。

1.2 一维并行注意力机制（1D-PAM）

尽管 Mamba 机制能够高效捕获工况的全局长期演化趋势，但在火电机组四管泄漏初期，故障征兆通常表现为极短促且微弱的局部瞬态冲击。为弥补 SSM 在处理高频瞬态突变时的平滑效应，实现对微弱异常的高灵敏度捕捉，本文进一步引入以多头自注意力为核心的一维并行注意力机制（1D-PAM）作为互补特征提取支路^[22]。

给定输入特征序列 $X \in \mathbb{R}^{T \times D}$ ，首先叠加位置嵌入矩阵 W_{pos} 注入绝对时序信息，得到 $X_{\text{pos}} = X + W_{\text{pos}}$ 。随后，通过 h 个并行的注意力头，将输入特征线性投影为查询（ Q ）、键（ K ）和值（ V ）并计算缩放点积注意力，从而在不同的特征子空间中联合捕获瞬态冲击：

$$H_{\text{MHA}}(X_{\text{pos}}) = \text{Concat}(H_{\text{head},1}, \dots, H_{\text{head},h})W_O \quad (4)$$

式中： X_{pos} 为注入位置信息后的特征张量； $H_{\text{MHA}}(X_{\text{pos}})$ 为多头注意力机制的整体输出特征张量； $\text{Concat}(\cdot)$ 为张量拼接操作函数； W_O 输出线性映射权重矩阵； $H_{\text{head},h}$ 为第 h 个单头注意力的输出特征张量，其中第 i 个单头注意力的计算方式为：

$$H_{\text{head},i} = \text{softmax}(Q_i K_i^T / \sqrt{d_k}) V_i \quad (5)$$

式中： Q_i, K_i, V_i 分别为第 i 个注意力子空间的查询、键和值张量； d_k 为注意力子空间的缩放因子维度。

最终，多头自注意力的输出经多层感知机进一步提取深层抽象特征，并辅以残差连接，得到该支路最终的局部聚焦特征 F_p ：

$$F_p = f_{\text{MLP}}(H_{\text{MHA}}(X_{\text{pos}})) + H_{\text{MHA}}(X_{\text{pos}}) \quad (6)$$

式中： $f_{\text{MLP}}(\cdot)$ 表示包含两层线性映射与非线性激活函数的多层感知机。1D-PAM 能够在网络浅层自适应地为微弱泄漏冲击所在的时间步分配较高的权重，为后续跨域度量提供局部表征。

1.3 原型度量网络优化目标

在借助 Mamba 与 PAM 双支路分别获取全局趋势与局部突变联合特征后，为解决少样本跨工况特征混淆问题，本文引入原型网络作为分类器与优化目标。少样本学习采用元学习训练范式，设特征提取网络的非线性映射函数为 $f_\phi(\cdot)$ 。对于类别 k ，其类原型 c_k 定义为支持集 S_k 内所有同类样本特征的均值向量：

$$c_k = \frac{1}{|S_k|} \sum_{(x_i, y_i) \in S_k} f_\phi(x_i) \quad (7)$$

给定查询样本 x ，其与原型的距离采用平方欧氏距离度量：

$$d(f_\phi(x), c_k) = \|f_\phi(x) - c_k\|_2^2 \quad (8)$$

查询样本属于类别 k 的预测概率由距离分布经 Softmax 函数归一化求得，整体优化目标是最小化负对数似然损失：

$$\mathcal{L} = -\log p(y = k | x) = d(f_\phi(x), c_k) + \log \sum_{k'} \exp(-d(f_\phi(x), c_{k'})) \quad (9)$$

式中： $\mathcal{L}$ 为模型整体优化的负对数似然损失； x 为给定的查询样本， c_k 为该查询样本真实类别 k 的类原型中心向量， $d(\cdot, \cdot)$ 为高维嵌入空间中的平方欧氏距离， k' 为元学习任务中所涵盖的对比目标类别。该损失函数在深层流形空间内构建了同类紧致、异类排斥的几何约束。其旨在缩小查询样本与真实类原型的空间偏差，促使同类特征加速聚拢；还起到推远异类原型的作用，进而拓宽不同类别间的决策边界^[15]。由于该度量机制的判决逻辑建立在特征的相对拓扑距离之上，彻底剥离了对数据绝对幅值的依赖，因而能够有效抵御由工况转换引发的特征尺度漂移，赋予模型极强的跨域鲁棒性。

2 Proto-MixMamba 诊断模型构建

针对深度调峰工况下故障样本极少、特征漂移显著的难题，本文构建了端到端的 Proto-MixMamba 诊断框架。本节将阐述其总体框架、双支路提取机制及跨域度量策略。

2.1 总体诊断框架

图 1 为 Proto-MixMamba 模型的总体结构，涵盖预处理、混合块堆叠、特征压缩及跨域分类度量。

在数据预处理阶段，针对输入数据，采用无重叠采样策略还原泄漏演化的真实物理过程，避免特征互扰。数据经 Z-Score 归一化后，通过一维卷积与线性映射进行特征升维。预处理后的张量依次穿过多个串联混合块，提取深层时序特征。在网络末端，特征序列经自适应平均池化、展平与层归一化压缩后，输入至分类决策层。

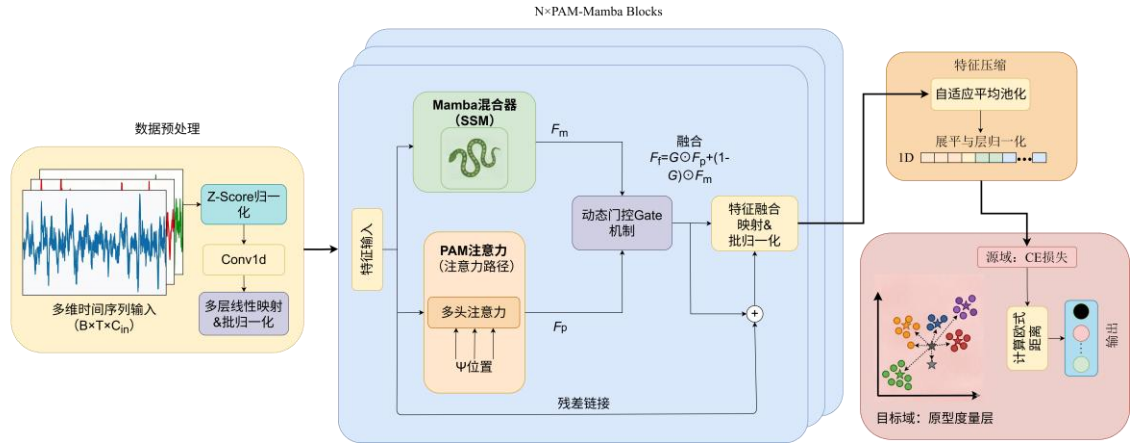


图 1 Proto-MixMamba 总体框架图
Fig.1 Overall framework diagram of Proto-MixMamba

2.2 1D-PAM-Mamba 并行特征提取机制

为克服单一网络在处理高频瞬态冲击与低频稳态演化时的局限，本文设计了双支路并行结构，如图 2 所示。

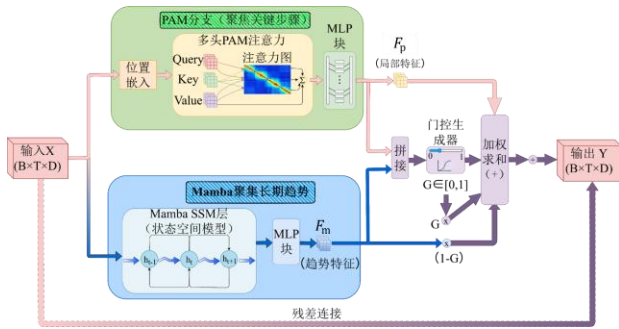


图 2 1D-PAM-Mamba 模块详解图
Fig.2 Detailed Diagram of the 1D-PAM-Mamba Module

在每个混合块内，输入张量 X 被同步送入两条互补支路。在 PAM 分支中，网络融合位置嵌入与多头注意力，为包含瞬态泄漏冲击的时间步分配高响应权重，提取局部突变特征 F_p 。在 Mamba 分支中，利用选择性扫描机制聚合长期工况演化规律，并滤除高频工况噪声，输出全局趋势特征 F_m 。两条支路并行运作，实现了高频局部特征与低频全局特征的解耦建模。

2.3 自适应门控融合与跨域分类策略

考虑到不同负荷工况下局部冲击与全局趋势的贡献度动态变化，本文引入自适应门控融合机制，让网络根据输入内容自主权衡支路权重。将 F_p 与 F_m 拼接后，经由含 Sigmoid 激活函数的多层感知机映射，生成动态门控权重矩阵 $G \in (0,1)$ ，并通过下式实现特征的自适应软切换：

$$Y = G \odot F_p + (1 - G) \odot F_m + X \quad (10)$$

该设计完成了多尺度信息的融合，其余残差连接确保了梯度传播的稳定性。

完成特征压缩后，网络进入跨域度量学习阶段。本文采用“源域预训练—目标域元学习”的两阶段训练策略：在源域上，首先利用交叉熵损失对网络进行充分预训练，使其初步具备多变量特征的解耦与模式提取能力；当模型迁移至目标域时，则主动舍弃线性分类头，转而接入原型度量层。在这一度量空间中，分类不再依赖固定的超平面划分，而是通过动态计算支持集的原型中心，并以查询样本与各原型之间的平方欧氏距离作为相似度判据来完成。该策略有效规避了因全连接层参数固化所导致的跨工况幅值尺度漂移问题，使模型在少样本条件下仍能对样本实现精准辨识，赋予整个诊断框架良好的跨域鲁棒性。

3 实验验证与分析

3.1 实验数据集

3.1.1 仿真数据采集与样本构建

本文实验数据采集自某 600 MW 超临界火电机组全范围仿真机，该仿真机经现场测试验收，具有较高的静态精度和逼真的动态特性，可准确模拟机组正常启动、停机和不同方式的变负荷过程，并对机组的各种故障进行全过程仿真，设定仿真机分别在 600 MW、540 MW 以及 480 MW 负荷下处于协调控制稳定运行模式。在上述三种典型工况下模拟了正常运行状态及锅炉不同受热面在不同泄漏程度下的故障，共设置 8 种状态类别（标记为 M0~M7）。以右侧省煤器泄漏（M1）为例，图 3

展示了 600 MW 负荷下 4 个关键参数在重度、中度以及轻度泄漏程度下的动态变化趋势。

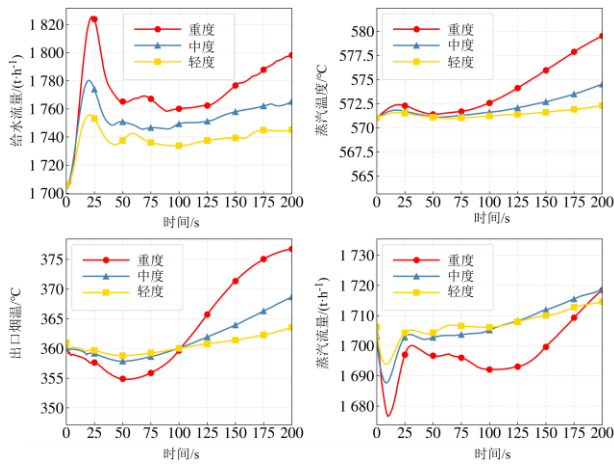


图 3 600 MW 负荷下右侧省煤器不同泄漏程度的关键参数动态响应

Fig.3 Dynamic responses of key parameters under different leakage severities of the right economizer at 600 MW load

在数据提取阶段，系统在正常稳定运行一段时间后注入不同泄漏程度故障。每次故障截取注入开始到状态稳定的 200 个连续数据点，采样步长设定为 1 s。详细的故障类别与标签映射见表 1。

通过上述独立采样机制，本文在每种负荷工况下，针对每一类故障均提取了 30 个独立样本，分别构建了数据集 A(600 MW)，数据集 B(540 MW)和数据集 C(480 MW)。各数据集包含相同的故障类别且各类别样本数量相等，但由于运行机理的变动，同一故障在不同工况下的特征分布存在显著差异，这为后续的少样本变工况诊断评估提供了数据基础。

表 1 故障类别描述与标签定义

Tab.1 Description of fault categories and label definitions

故障标签	故障类别
M0	正常工况
M1	右侧省煤器泄漏
M2	右侧高温再热器泄漏
M3	右侧高温过热器泄漏
M4	左侧省煤器泄漏
M5	左侧高温再热器泄漏
M6	左侧高温过热器泄漏
M7	水冷壁泄漏

基于此，本文严格采用情景式元学习采样策略来划分支持集与查询集。在每次构建独立的元学习任务时，对于任意设定的 K -shot 实验，模型从每个状态类别的 30 个样本中无放回地随机抽取 K 个

样本构成支持集，专门用于在深层度量空间中动态计算该类的原型中心。同时，该类别剩余的样本全部划入查询集，用于计算负对数似然损失并客观评估模型当前的跨工况诊断精度。这一动态随机抽样机制保证了支持集与查询集在样本层面的绝对正交与隔离，从根源上彻底杜绝了数据泄露问题。

3.1.2 基于物理机制的特征提取

针对火电厂分布式控制系统 (distributed control system, DCS) 海量监测数据中存在的信息冗余与高维特征问题，若直接将全部测点输入诊断模型，不仅会显著增加模型的计算复杂度，还可能因无关特征的干扰影响模型准确率。因此，基于锅炉汽水系统的物理运行机理进行特征筛选至关重要。

机组负荷与主汽压力作为表征机组能量平衡的核心状态变量，在发生受热面泄漏时，工质的质量流失必然打破原有的热力学平衡，导致上述全局参数出现宏观偏离。除了宏观指标外，泄漏引发的局部流场与温度场畸变会以多物理场耦合的形式映射至相关子系统。具体而言，泄漏点下游的受热面管壁温度及出口汽温会因工质流量的突变而产生非线性异常波动；与此同时，给水流量与主蒸汽流量之间的质量不平衡信号，构成了判断泄漏发生的直接物理依据。此外，考虑到闭环控制系统的调节作用，为了维持参数稳定，减温水喷淋量的开度变化往往隐含着被掩盖的故障演化早期征兆。基于上述机理分析，本文最终筛选出 20 个关键热工参数构建模型的输入特征空间^[19]，针对不同传感器的数据量纲不同，采用 Z-Score 来消除量纲影响。具体清单如表 2 所示。

表 2 特征选择结果

Tab.2 Feature selection results

3.2 对比实验分析

为了全面验证本文提出的 Proto-MixMamba 模型在小样本跨工况条件下的优越性，选取了涵盖传统深度模型 ResNet1D(1D Residual Network)、LSTM(Long Short-Term Memory)、Vanilla Transformer，经典少样本方法

故障特征参数	
机组实际负荷/MW	左高温再热器出口蒸汽温度/°C
主汽压力/MPa	右高温再热器出口蒸汽温度/°C
给水流量/(t·h ⁻¹)	左高温再热器出口蒸汽压力/MPa
蒸汽流量/(t·h ⁻¹)	右高温再热器出口蒸汽压力/MPa
左高温过热器出口蒸汽温度/°C	左再热器减温水喷水量/(t·h ⁻¹)
右高温过热器出口蒸汽温度/°C	右再热器减温水喷水量/(t·h ⁻¹)
左高温过热器出口蒸汽压力/MPa	左右省煤器出口烟温差/°C
右高温过热器出口蒸汽压力/MPa	省煤器出口给水温度/°C
左二级减温水喷水量/(t·h ⁻¹)	水冷壁出口集箱温度/°C
右二级减温水喷水量/(t·h ⁻¹)	汽水分离器出口蒸汽压力/MPa

Siamese-WDCNN(Siamese-Deep Convolutional Neural Networks with Wide First-layer Kernels), 以及前沿改进模型 Ensembling Transformer 在内的多组基准进行对比^[23-25]。在具体参数配置上, 本文所提模型的 Mamba 分支状态维度、一维卷积核大小与扩展因子分别设为 16、4 与 2, 1D-PAM 分支多头注意力头数设为 4。在模型训练方面, 以上基准模型均通过网络搜索进行超参数寻优, 并且采用了与本文一致的训练策略: 选用 AdamW 优化器, 初始学习率设为 0.001, 批次大小设为 32, 权重衰减系数设为 1×10^{-4} , 并统一引入耐心值为 20 的早停机制以防止少样本条件下的过拟合。为消除抽样带来的偶然性误差, 本文所有定量实验均独立重复运行 10 次。

图 4 展示了各模型在 600 MW $\rightarrow$ 540 MW 及 600 MW $\rightarrow$ 480 MW 跨工况任务下的诊断准确率。从实验结果可见, 在 10-shot 相对充裕的样本条件下, 多数改进模型均能达到较高的诊断精度, 但是当训练样本数量渐趋减少时, 模型间的性能差距显著拉大, 特别是在极具挑战性的 1-shot 设定下, 性能分层现象呈现得格外明显。特别是在跨度更大

的 600 MW $\rightarrow$ 480 MW 深度调峰任务中, 由于剧烈的变工况特征漂移, LSTM 的准确率跌破 50%, 即便是针对少样本设计的 Siamese-WDCNN 也仅勉强维持在 66.8% 左右, 陷入决策失效。相比之下, Proto-MixMamba 在上述两项任务的 1-shot 条件下, 依然分别保持了约 90.3% 与 88.4% 的较高准确率, 展现出极强的跨域鲁棒性。这一结果充分证实了 1D-PAM 与 Mamba 的联合架构成功提取了抗工况干扰的不变特征, 而原型的度量机制彻底规避了全连接层的幅值漂移, 确保了模型在少样本条件下的统泛化能力。

此外, 在相同硬件环境下对计算成本进行了简要评估。在当前序列长度 $N=200$ 下, Proto-MixMamba 的平均单次迭代训练耗时约为 22.8 ms, 与 Vanilla Transformer 约 24.5 ms 基本持平。这表明所提模型在序列长度较短的前提下, 大幅提升了少样本跨域诊断的精度, 且得益于底层状态空间模型的线性计算复杂度, 其在应对未来更长的时间序列数据时, 能有效避免自注意力机制带来的耗时激增, 具备良好的工程延展性。

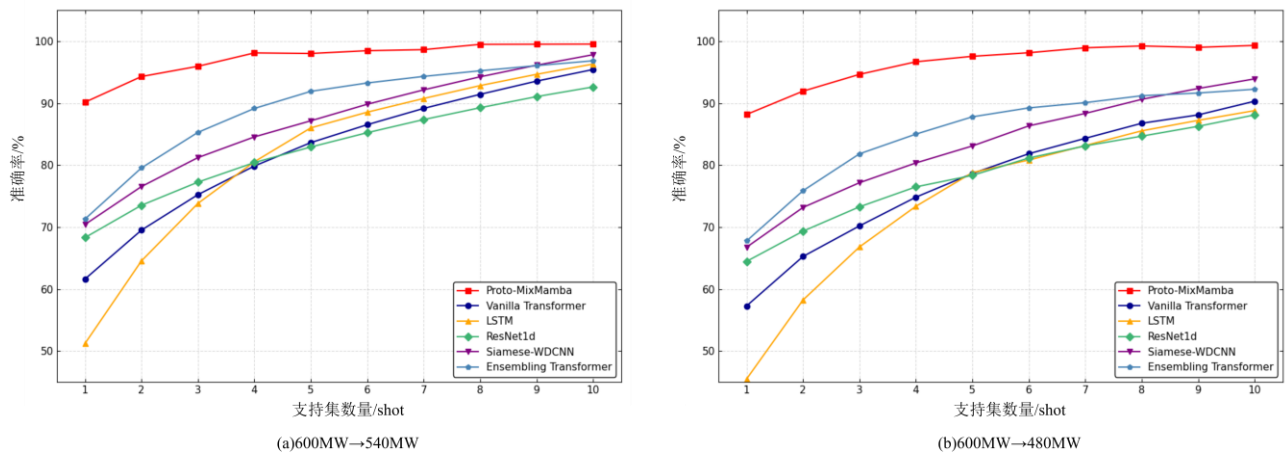


图 4 各模型在不同跨工况少样本任务下的诊断准确率对比

Fig.4 Comparison of diagnostic accuracy among different models under cross-condition few-shot tasks

3.3 模型组件消融实验

为了验证 Proto-MixMamba 架构中各关键组件的有效性, 设计了多组变体模型进行消融实验研究, 结果如表 3 所示。

由表 3 可见, 完整模型均取得了最优性能, 说明各组件协同工作的必要性。移除 1D-Mamba 分支对性能的负面影响最为显著, 其准确率在 3-shot 下大幅衰减, 表明 Mamba 提供的全局长程依赖建模

对长序列中的故障演化规律的捕捉具有决定性作用。与此同时, 移除 PAM 分支也导致了准确率的明显下降, 说明缺失局部多尺度特征会削弱模型对瞬态冲击信号的捕捉能力。此外, 去除自适应门控融合并改为简单相加后, 模型性能也有所回落, 这进一步验证了在处理非平稳故障信号时, 动态调节局部与全局特征权重的机制是不可或缺的。

表 3 不同组件对模型性能影响的消融实验结果
Tab.3 Ablation study results of different components on

model performance

模型配置	3-shot/%	5-shot/%	10-shot/%
完整模型	94.64±0.65	97.55±0.42	99.31±0.18
移除 PAM	80.18±1.42	90.55±0.95	96.04±0.52
移除 1D-Mamba	70.73±1.78	87.16±1.15	93.90±0.68
移除 Adaptive Gated Fusion	82.01±1.18	89.77±0.84	96.34±0.45

为了更直观地验证双支路架构的分工机制, 本文通过改变样本窗口长度进行了动态消融实验, 结果如图 5 所示。在样本窗口长度较短时, 系统主要体现为初期的瞬态冲击特性, 此时仅保留 PAM 支路的模型准确率上升更快, 这验证了 1D-PAM 对早期突变特征的高灵敏性。而当窗口拉长到 150 s~300 s 后, 系统的长程迟滞特性逐渐显露, 这时只保留 Mamba 支路的模型则展现出更强的长程建模能力, 准确率稳步反超。

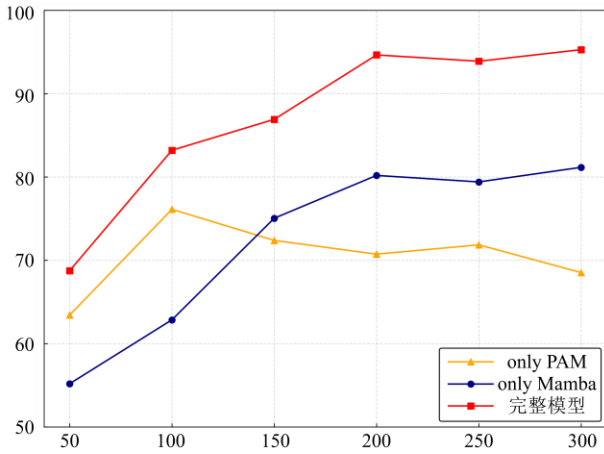


图 5.不同窗口长度下各消融模型的诊断准确率对比
Fig.5 Comparison of diagnostic accuracy of ablated models under different window lengths

3.4 可视化分析

为了进一步揭示 Proto-MixMamba 模型在克服变工况与少样本瓶颈时的内部运作机理, 本文针对 600 MW → 480 MW 跨域任务 (3-shot 设定), 对其诊断结果进行了深度的可视化剖析。

图 6 展示了所提 Proto-MixMamba 模型在该任务下的分类混淆矩阵。由图可知, 该矩阵呈现出极高的对角线纯度, 模型对正常状态 (M0) 的识别率高达 99.3%。值得特别关注的是, 在 M1 (右侧省煤器泄漏) 与 M4 (左侧省煤器泄漏) 这组空间耦合性故障辨识上, 所提模型取得了 90.1% 与 92.9% 的高准确率。在实际运行中, 由于左右侧省煤器物理位置对称且热力学响应特征高度相似, 在 480 MW 低负荷运行引发的全局参数漂移干扰下, 传统基线模型 (如 ResNet1D) 往往会出现严重的决策混乱, 极易将 M1 样本错判为 M4, 其对正常状态的

辨识度也极低。而 Proto-MixMamba 能够彻底攻克这一难题, 直观证明了 1D-PAM 模块的引入使得模型能够精准捕捉到左右侧泄漏在流场局部产生的微弱非对称突变, 而不再仅仅依赖那些容易被变工况掩盖的宏观迟滞参数。

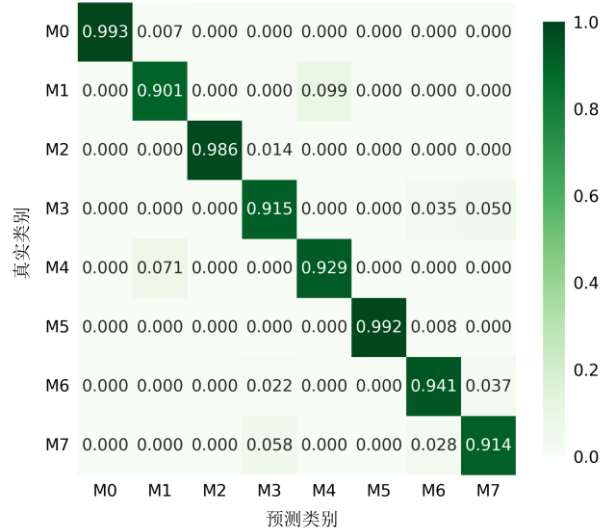


图 6 600 MW → 480 MW 任务下 Proto-MixMamba 在 3-shot 设定下的混淆矩阵

Fig.6 Confusion matrix of the proposed Proto-MixMamba under the 3-shot setting in the 600 MW → 480 MW task

为进一步从特征空间维度验证上述结论, 图 7 给出了该跨域任务下不同模型的 t-SNE 可视化对比结果。观察图 7a) 可知, 由于缺乏对变负荷漂移的抵抗力, 传统模型提取的特征在目标域内发生了严重的畸变与重叠, 类间边界极度模糊, 尤其是类别 M3、M6 与 M7 高度混杂。在这种特征坍塌的状态下, 仅依靠 3 个支持集样本根本无法划定有效的分类超平面。

反观图 7d) 中 Proto-MixMamba 的特征可视化结果, 8 种状态类别在二维空间中形成了清晰、独立且紧凑的“岛屿状”聚类结构。这完美印证了本文情景式元学习与原型度量损失的有效性: 即使在源域和目标域数据分布差异巨大的情况下, 基于欧氏距离的度量机制依然能够强制网络在深层空间中形成“同类样本紧密聚集、异类样本彼此远离”的分布。这种优异的特征表达能力, 保障了模型在仅有 3 个参考样本时, 依然能动态生成高精度的类原型, 从而稳健地完成跨工况少样本诊断。

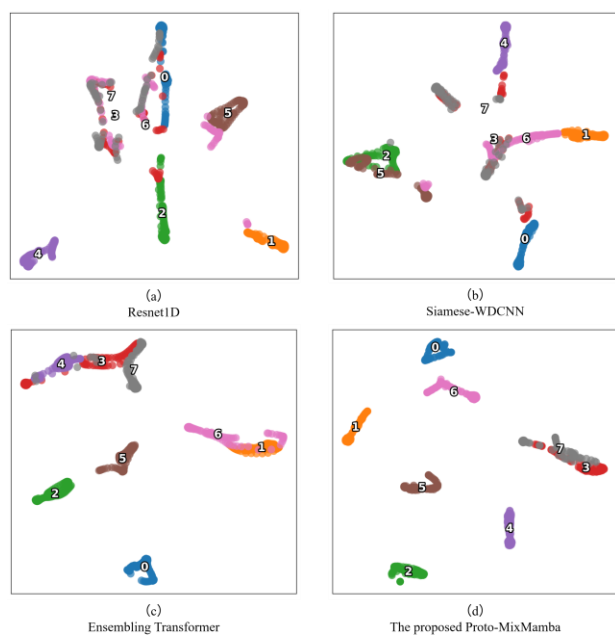


图 7 600 MW → 480 MW 任务下目标域特征流形空间的 t-SNE 可视化

Fig.7 t-SNE visualization of the target domain feature manifold space in the 600 MW → 480 MW task

4 结论与展望

针对火电机组在深度调峰变工况下,因特征分布漂移与故障样本极度匮乏导致传统数据驱动模型泛化性能严重下降的问题,本文提出了一种融合时序多尺度解耦与原型度量的跨工况少样本故障诊断框架——Proto-MixMamba 模型。基于高精度仿真数据的严格评估,主要得到以下结论:

1) 针对热工参数非线性突变与长程迟滞并存的演化特性,所设计的 1D-PAM 与 Mamba 双支路特征提取架构,实现了对局部瞬态波动和全局低频趋势的提取。结果表明,该机制有效提升了变负荷干扰下物理对称泄漏位置(如左/右侧省煤器泄漏)的辨识准确率,改善了传统深度模型在该场景下的误判问题。

2) 针对变工况引起的特征漂移,本文摒弃了全连接层分类方式,引入了 N-way K-shot 情景式元学习与原型度量层。可视化结果表明,基于相对欧氏距离的度量计算缓解了变负荷带来的参数幅值偏移影响,使目标域特征呈现出良好的类内聚集与类间可分性,为少样本分类构建了清晰的特征空间。

3) 在 600 MW → 540 MW 及更具挑战性的 600 MW → 480 MW 跨工况任务中,Proto-MixMamba 展现出良好的泛化能力。在 1-shot 的极限条件下,诊断准确率分别保持在 90.3% 和 88.4% 左右,有效抑

制了过拟合,为工业现场少标签环境下的早期预警提供了一种可行的技术路线。

尽管本文模型在跨工况少样本诊断方面取得了明显进展,但距离工业现场全面部署仍存在优化空间。未来研究将从以下两个方面展开:

1) 本文目前的验证主要基于较为理想的仿真数据开展,未显式叠加高斯白噪声等模拟现场环境的测量噪声。考虑到实际火电机组 DCS 系统中,热工信号不可避免地会受到高频测量噪声与电磁干扰的影响,未来研究将在仿真环节系统性地注入多尺度高斯白噪声及脉冲噪声。在此基础上,将进一步探索并验证 Proto-MixMamba 模型在强背景噪声下的抗噪鲁棒性,以提升模型在复杂工业现场的泛化与抗干扰能力。

2) 现阶段的诊断任务主要针对单一受热面发生的独立故障。考虑到机组在实际长期运行中,可能会出现多管线并发泄漏等复合故障情况,后续研究将结合多标签分类等机制,探讨对叠加故障特征的有效分离方法,以实现对多重故障的准确识别。

【参考文献】

- [1] 杜小泽, 张璇, 庞力平, 等. 燃煤发电机组灵活调峰下机炉安全状态监测与调控研究综述[J]. 中国电机工程学报, 2024, 44(18): 7178.
Du X Z, Zhang X, Pang L P, et al. Review on the monitoring and regulation of boiler and turbine safety states in coal-fired power generation units under flexible peak shaving[J]. Proceedings of the CSEE, 2024, 44(18): 7178.
- [2] Zhou C, Huang J, Li L P, et al. Failure analysis of leakage in high-temperature superheater tubes of a deep peak regulation supercritical boiler TP347H[J]. Engineering Failure Analysis, 2025, 173: 109443.
- [3] Lei Y G, Jia F, Lin J, et al. An intelligent fault diagnosis method using unsupervised feature learning towards mechanical big data[J]. IEEE Transactions on Industrial Electronics, 2016, 63(5): 3137.
- [4] Ramezani M G, Hasanian M, Golchinfar B, et al. Automatic boiler tube leak detection with deep bidirectional LSTM neural networks of acoustic emission signals[C]//Sensors and Smart Structures Technologies for Civil, Mechanical, and Aerospace Systems 2020. Bellingham: SPIE, 2020: 113791I.
- [5] Khalid S, Azad M M, Kim H S. A generalized autonomous power plant fault detection model using deep feature extraction and ensemble machine learning[J]. Mathematics, 2025, 13(3): 342.
- [6] Cai X Q, Wang J J, Ye Y C, et al. An early warning method of pipeline leakage monitoring with limited leakage samples[J]. Measurement, 2025, 242(Part C): 116013.
- [7] Xue Y L, Li D L, Song Y, et al. Novel deep learning framework for evaporator tube leakage estimation in supercharged boiler[J]. Energies, 2025, 18(15): 3986.
- [8] Li G L, Li Y J, Su J, et al. Fault diagnosis for supercharged boiler based on self-improving few-shot learning[J]. Energy, 2025, 316: 134507.
- [9] 陈启泰, 翟春华, 黄庆, 等. 基于 TimeGAN 的燃气轮

- 机气路故障 数据增强方法[J/OL]. 热力发电, 2025-05-26. <https://doi.org/10.19666/j.rld.202501003>.
- Chen Q T, Zhai C H, Huang Q, et al. Gas path fault data augmentation method for gas turbine based on TimeGAN[J/OL]. Thermal Power Generation, 2025-05-26. <https://doi.org/10.19666/j.rld.202501003>.
- [10] 马良玉, 黄日灏, 段晓冲, 等. 基于改进生成对抗网络的滚动轴承故障样本扩充及诊断方法[J]. 华北电力大学学报(自然科学版), 2026, 53(3): 129.
- Ma L Y, Huang R H, Duan X C, et al. Fault sample expansion and diagnosis method of rolling bearing based on improved generative adversarial networks[J]. Journal of North China Electric Power University (Natural Science Edition), 2026, 53(3): 129.
- [11] Li J P, Yue K, Wu Z Q, et al. Source-free progressive domain adaptation network for universal cross-domain fault diagnosis of industrial equipment[J]. IEEE Sensors Journal, 2025, 25(5): 8067.
- [12] Hu J W, Li W G, Zhang Y, et al. Cross-domain few-shot fault diagnosis based on meta-learning and domain adversarial graph convolutional network[J]. Engineering Applications of Artificial Intelligence, 2024, 136(Part B): 108970.
- [13] Wang H, Wang X D, Yang Y Z, et al. A few-shot machinery fault diagnosis framework based on self-supervised signal representation learning[J]. IEEE Transactions on Instrumentation and Measurement, 2024, 73: 3509114.
- [14] Snell J, Swersky K, Zemel R. Prototypical networks for few-shot learning[C]//Proceedings of the 31st International Conference on Neural Information Processing Systems. Red Hook: Curran Associates Inc., 2017: 4080.
- [15] Yue K, Li J P, Deng S H, et al. A relationship-aware calibrated prototypical network for fault incremental diagnosis of electric motors without reserved samples[J]. Reliability Engineering & System Safety, 2024, 252: 110429.
- [16] An Y Y, Li Z F, Li Y Y, et al. Few-shot learning-based fault diagnosis using prototypical contrastive-based domain adaptation under variable working conditions[J]. IEEE Sensors Journal, 2024, 24(15): 25019.
- [17] Gu A, Dao T. Mamba: linear-time sequence modeling with selective state spaces[PP/OL]. V2. (2024-05-31)[2026-05-15]. <https://doi.org/10.48550/arXiv.2312.00752>.
- [18] Than X, Liu Y, Wang Z. MixMamba: A novel Mamba architecture for few-shot fault diagnosis of bearings[J]. Measurement, 2024, 224: 113942.
- [19] 马良玉, 韩立凯. 融合混合专家及多目标域对抗迁移网络的锅炉汽水系统跨工况故障诊断[J]. 中国电机工程学报, 2026, 46(11): 4644.
- Ma L Y, Han L K. Cross-condition fault diagnosis of boiler steam-water system based on mixture of experts and multi-target domain adversarial migration network[J]. Proceedings of the CSEE, 2026, 46(11): 4644.
- [20] Gu A, Goel K, Ré C. Efficiently modeling long sequences with structured state spaces[PP/OL]. V3. (2022-08-05)[2026-05-15]. <https://doi.org/10.48550/arXiv.2111.00396>.
- [21] Liu K, Li M X, Shen Q Q, et al. M-TSINR: multiscale implicit neural representations via mamba encoder for time series anomaly detection[C]//2025 IEEE International Conference on Knowledge Graph (ICKG). Limassol: IEEE, 2025: 229.
- [22] Vaswani A, Shazeer N, Parmar N, et al. Attention is all you need[C]//Proceedings of the 31st International Conference on Neural Information Processing System. Red Hook: Curran Associates Inc., 2017: 6000.
- [23] He K M, Zhang X Y, Ren S Q, et al. Deep residual learning for image recognition[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas: IEEE, 2016: 770.
- [24] Zhang L Y, Dong H H, Jia L M, et al. Multi-scale Siamese network for few-shot fault diagnosis of bogie component[C]//Gong M, Jia L M, Qin Y, et al. Proceedings of the 6th International Conference on Electrical Engineering and Information Technologies for Rail Transportation (EITRT) 2023. Singapore: Springer, 2024: 638.
- [25] Xu B, Li H P, Ding R C, et al. Fault diagnosis in electric motors using multi-mode time series and ensemble transformers network[J]. Scientific Reports, 2025, 15(1): 7834.

(责任编辑 李园)