

DOI:10.19651/j.cnki.emt.2519440

基于 DE 优化 VMD 与 CNN-BiGRU-Attention 结合的滚动轴承故障诊断*

邵鲁川 赵兵 康旭涛

(青海大学机械工程学院 西宁 810016)

摘要: 针对滚动轴承的振动信号易受到噪声干扰、故障特征提取困难、诊断精度低的问题,提出一种基于 DE 优化 VMD 与 CNN-BiGRU-Attention 结合的模型进行故障识别与诊断的方法。根据最小化包络熵原则,用 DE 优化 VMD,得到最佳分解层数和惩罚因子,计算最佳分解参数组合下的各个 IMF 的相关系数,根据设定的阈值筛选有用信号完成信号重构。首先将识别模型与常用分类模型进行对比了实验,本文所提方法的 accuracy、precision、recall 和 F1 score 均高于其他方法,故障诊断准确率达到 99.17%。其次将原始信号和重构信号特征输入 CNN-BiGRU-Attention 模型中,原始信号在本文所提模型的故障准确率为 89.58%,低于信号降噪后的准确率。最后将 CNN-BiGRU-Attention 与 CNN-BiLSTM-Attention 模型对比,CNN-BiGRU-Attention 比 CNN-BiLSTM-Attention 模型准确率高 1.25%、GPU 占用率降低 21%、CPU 占用率降低 23%,训练时间快 35s,实验研究结果可为现有滚动轴承故障诊断技术提供一种有效的改进方法。

关键词: 滚动轴承;信号降噪;变分模态分解;诊断模型;故障诊断

中图分类号: TH133.33;TN911.72 **文献标识码:** A **国家标准学科分类代码:** 460.2020

Rolling bearing fault diagnosis based on DE optimization of VMD combined with CNN-BiGRU-Attention

Shao Luchuan Zhao Bing Kang Xutao

(School of Mechanical Engineering, Qinghai University, Xining 810016, China)

Abstract: To address the challenges of noise interference, difficult fault feature extraction, and low diagnostic accuracy in rolling bearing vibration signals, this study proposes a fault diagnosis method based on differential evolution (DE)-optimized variational mode decomposition (VMD) combined with a comprehensive model integrating convolutional neural networks (CNN)-bidirectional gated recurrent unit (BiGRU)-Attention. Following the minimum envelope entropy principle, DE is employed to optimize VMD, obtaining the optimal decomposition layer number and penalty factor. The correlation coefficients of intrinsic mode functions (IMFs) under the optimal parameter combination are calculated, and useful signals are reconstructed based on a predetermined threshold. Comparative experiments with conventional classification models demonstrate that the proposed method achieves superior performance in accuracy, precision, recall and F1 score, with a fault diagnosis accuracy of 99.17%. When comparing the diagnostic results between raw and reconstructed signals in the CNN-BiGRU-Attention model, the accuracy for raw signals is 89.58%, lower than that of the denoised signals. Finally, CNN-BiGRU-Attention was compared with the CNN-bidirectional long short-term memory (BiLSTM)-Attention model. The CNN-BiGRU-Attention model showed a 1.25% higher accuracy, a 21% reduction in graphic processing unit (GPU) usage, a 23% reduction in central processing unit (CPU) usage, and a training time that was 35 seconds faster. These experimental results can provide an effective improvement method for existing rolling bearing fault diagnosis technologies.

Keywords: rolling bearing; signal denoising; variational mode decomposition; diagnostic model; fault diagnosis

0 引言

滚动轴承作为机械设备中常用的关键部件^[1],轴承的

健康状态对整个设备的运行状况有着直接的关系^[2]。据统计,在使用滚动轴承的旋转机械设备中,约有 30% 的机械故障与轴承损坏有关,轴承故障的高占比反映了轴承故障

收稿日期:2025-07-24

* 基金项目:青海省 2025 年重点研发与转化计划项目(2025-QY-209)资助

诊断的重要性和必要性^[3]。因此,研究高精度的轴承故障诊断方法,实现对轴承的实时健康监测和故障诊断,对降低机械设备维护成本、减少事故的发生具有重要意义。

轴承故障检测的信息主要来自于振动信号,振动信号是一种时间序列数据,包含了各种状态信息^[4]。当前,轴承故障诊断流程主要涵盖3个环节:振动信号采集、特征提取以及故障类型识别^[5]。在振动信号采集过程中,由于轴承的运行通常是在变化和恶劣的条件下进行的,有强烈的噪声干扰^[6]。为了从采集到的振动信号中获得真实的故障特征信息,需要对实测振动信号进行降噪,完成信号重构工作。各种信号降噪技术均已被探索,包括经验模式分解(empirical mode decomposition, EMD)、变分模态分解(variational mode decomposition, VMD)、小波分析(wavelet transform, WT)和奇异值分解(singular value decomposition, SVD)^[7]。在EMD的基础上,刘金锁等^[8]将集成经验模态分解(ensemble empirical mode decomposition, EEMD)应用到岩性识别技术问题中,有效解决了EMD存在的模态混叠现象,但EEMD需预设两个关键参数:添加噪声的幅值和集成次数,参数选择不当会导致降噪失败。周顺勇等^[9]对低信噪比环境下调制信号识别精度受限的问题,引入SVD算法对接收的I/Q信号进行去噪。SVD降噪需将一维信号构造为Hankel矩阵或Toeplitz矩阵,其行数 m 直接影响分解效果,不合理的构造参数会导致信号成分与噪声无法有效分离。

VMD算法抛弃了循环递归筛选的处理方式,通过在变分框架范围内求解最优变分模型,获得模态分量,该方法不同于传统递归算法,具有良好的鲁棒性和严谨的数学理论支撑^[10]。VMD因其理论扎实、鲁棒性强、抗混叠性强等优点,在旋转机械故障诊断领域得到广泛应用^[11]。Gu等^[12]针对植物电信号中噪声过大的问题,提出了一种将VMD与经验小波变换(empirical wavelet transform, EWT)相结合的去噪方法。其中VMD的分解性能受分解层数 K 和惩罚因子 α 影响, $[K, \alpha]$ 往往依赖于工程师和研究人员的经验,导致主观性对分析结果影响极大^[13]。为此,相关研究者又以引力搜索算法、蝗虫优化算法、麻雀搜索算法等智能优化算法为手段,通过设置目标函数如最小化包络熵等,对VMD参数进行优化,但上述智能优化算法均在一定程度上存在有陷入局部最优后无法跳出、后期收敛速度慢等问题^[14]。而差分进化算法(differential evolution, DE)是一种基于群体的进化算法,具有寻优能力强、收敛速度快、精度高、结构简单和控制参数少等特点,被广泛应用于机械优化设计中^[15]。

在信号特征提取和故障类型识别上用传统诊断模型如卷积神经网络(convolutional neural networks, CNN)、支持向量机(support vector machine, SVM)存在特征难提取,所提取特征维度大,导致诊断模型效果差的问题^[16]。Ni等^[17]构建了一个双流卷积神经网络故障诊断模型用来识

别轴承故障,其中第一个流处理一维振动信号谱,第二个流处理来自相同信号的二维时频图。CNN通过卷积层提取局部特征,这可能在某些故障诊断场景中不足以捕捉全局特征或复杂的时空关系。沈伟杰等^[18]将提取到的轴承信号时频图作为输入,利用CNN深层次学习故障特征信息。最后,采用SVM对输出结果进行分类,以实现精确的故障类型识别。其中SVM的性能很大程度上取决于所选的核函数(如线性、多项式、径向基函数等)以及相关的参数(如正则化参数 C 和核函数参数),选择不当可能导致模型性能不佳。Fu等^[19]提出了一种基于连续小波变换(continue wavelet transform, CWT)、带通道注意的CNN、双向长短时记忆网络(bidirectional long short-term memory, BiLSTM)和残差模块的故障诊断模型,利用BiLSTM提取信号的时间特征作为一条路径,利用具有高效信道注意的CNN提取信号的空间特征作为另一条路径,这种并行神经网络有助于更好地进行特征提取,但由于BiLSTM的参数量大,存在计算复杂度高,训练速度慢的问题。门控循环单元(gate recurrent unit, GRU)是长短时记忆网络(long short-term memory, LSTM)的一个变体,GRU比LSTM结构更简单,更容易训练,简化了单元结构,减少了所需的训练参数^[20]。

针对上述问题,本文提出了一种基于DE优化VMD信号降噪重构和CNN-双向门控循环单元(bidirectional gated recurrent unit, BiGRU)-Attention相结合的滚动轴承故障诊断方法。用DE算法优化VMD,解决其他算法优化VMD参数中易陷入局部最优后无法跳出和对多目标混合整数优化需要进行离散化处理等问题;CNN-BiGRU联合提取局部和全局时序特征,克服传统神经网络仅能捕捉局部特征和忽略时序依赖的问题。Attention强化故障特征,提升分类置信度。CNN-BiGRU-Attention形成“空间特征捕获、长时依赖建模和关键信息聚焦”的流水线式深度处理链条。基于DE优化VMD与CNN-BiGRU-Attention的紧密结合,形成了从信号预处理源头优化到预测模型深度强化的完整步骤。实验结果表明,本文所提方法具有良好的诊断性能,能为滚动轴承故障诊断提供一种新的研究思路。

1 理论基础

1.1 DE算法原理

差分进化算法由Storn和Price^[21]提出,具有结构简单、易于实施、收敛速度快、鲁棒性好等优点。该算法包括初始化种群、变异、交叉、选择和终止条件5个步骤:

1)初始化种群:在 n 维的搜索空间中随机生成一个初始种群 $P = \{X_1, X_2, \dots, X_{N_p}\}$,其中 N_p 表示种群规模。

2)变异:对每个目标个体 $X_i^{(g)}$ (第 g 代),生成变异向量 $V_i^{(g)}$:

$$V_i^{(g)} = X_{r_1}^{(g)} + F \cdot (X_{r_2}^{(g)} - X_{r_3}^{(g)}) \quad (1)$$

式中: r_1, r_2, r_3 为互不相同的随机索引(且 $\neq i$), $F \in [0, 2]$ 为缩放因子(控制差分步长)。

3)交叉:通过二项交叉生成试验向量 $U_i^{(g)}$:

$$U_{i,j}^{(g)} = \begin{cases} V_{i,j}^{(g)}, & \text{rand}(0,1) \leq CR \text{ or } j = j_{\text{rand}} \\ X_{i,j}^{(g)}, & \text{其他} \end{cases} \quad (2)$$

式中: $CR \in [0,1]$ 为交叉概率, j_{rand} 为随机选择的维度索引(确保至少有一个维度来自变异向量)。

4)选择:比较试验向量 $U_i^{(g)}$ 和原个体 $X_i^{(g)}$, 保留更优者进入下一代:

$$X_i^{(g+1)} = \begin{cases} U_i^{(g)}, & f(U_i^{(g)}) \leq f(X_i^{(g)}) \\ X_i^{(g)}, & \text{其他} \end{cases} \quad (3)$$

5)终止条件:达到最大迭代次数或收敛值。

1.2 VMD 原理

VMD 通过构造约束优化问题将信号分解为 K 个本征模态函数^[22], 其核心公式为:

$$\left\{ \begin{array}{l} \min \sum_{k=1}^K \left\| \partial_t \left(\frac{\delta(t)}{2\pi} * u_k(t) e^{-j\omega_k t} \right) \right\|_2^2 \\ \text{subject to } \sum_{k=1}^K u_k(t) = x(t) \end{array} \right. \quad (4)$$

其中, ∂_t 表示对时间 t 的偏导数, $\delta(t)$ 是狄拉克函数, $*$ 表示卷积函数, j 是虚数单位。所有模态叠加后需严格等于原始信号 $x(t)$ 。通过增广拉格朗日法将约束优化转化为无约束问题, 引入拉格朗日乘子 $\lambda(t)$ 和二次惩罚项 α , 得到增广拉格朗日函数表达式:

$$\mathcal{L}(\{u_k\}, \{\omega_k\}, \lambda) = \alpha \sum_{k=1}^K \left\| \partial_t \left(\frac{\delta(t)}{2\pi} * u_k(t) e^{-j\omega_k t} \right) \right\|_2^2 + \left\| x(t) - \sum_{k=1}^k u_k(t) \right\|_2^2 + \langle \lambda(t), x(t) - \sum_{k=1}^k u_k(t) \rangle \quad (5)$$

从 VMD 分解过程可以看到, K 和 α 是控制 VMD 结果质量的 2 个重要参数。 K 设置得不合适可能会造成过分解或者欠分解, α 设置得不合适可能引起信息的丢失或者冗余。 K 必须是整数, α 可以为小数, 这是一个多目标混合整数优化问题, DE 算法对整数和连续混合变量具有天然兼容性, DE 算法中的变异操作, 通过差分向量生成新解时, 可直接对 K 取整, 无需额外编码。交叉操作保持 α 的连续性, 无需离散化处理。像粒子群优化算法 (particle swarm optimization, PSO)、鲸鱼优化算法 (whale optimization algorithm, WOA) 和北方苍鹰优化算法 (northern goshawk optimization, NGO) 等优化算法通常需要对整数变量特殊处理(如取整或二进制编码), 可能破坏搜索方向。 PSO 易陷入局部最优, WOA 和 NGO 收敛速度慢。

1.3 CNN-BiGRU 原理

CNN 是一个多层神经网络, 由输入层、卷积层、激活函数层、池化层和全连接层组成。 CONV 是 CNN 的核心部分, 其主要功能是对数据的卷积核进行卷积运算, 并将其输出到网络的下一层。卷积运算如下:

$$h_i^l = f \times (W_i^l * X^{l-1} + B_i^l) \quad (6)$$

式中: h_i^l 是层 l 的第 i 个特征, f 是激活函数, W_i^l 是层 l 的第 i 个卷积核权重矩阵, $*$ 表示卷积操作, X^{l-1} 则表示来自第 $(l-1)$ 层的输出, 而 B_i^l 则是第 l 层的第 i 个偏置项。

每个 GRU 模块都包含一个更新门和一个复位门来调节信息流^[23]。每个 GRU 模块分为两个段, 一个段从前到后处理数据, 另一个段从后到前处理数据^[24]。由于其双向架构, 该模型可以接收和处理过去和未来的信息, 从而更全面地理解振动序列数据。具体模型表示为:

$$z_t = \sigma(W_{xz}x_t + W_{sz}S_{t-1} + b_z) \quad (7)$$

$$r_t = \sigma(W_{xr}x_t + W_{sr}S_{t-1} + b_r) \quad (8)$$

$$\hat{S}_t = \tanh(W_{xs}x_t + W_{ss}(r_t \odot S_{t-1}) + b_h) \quad (9)$$

$$S_t = (1 - z_t) \odot S_{t-1} + z_t \odot \hat{S}_t \quad (10)$$

式中: t 表示时间, z_t 和 r_t 分别为更新门和复位门; \hat{S}_t 为候选隐藏层状态; x_t 为输入向量; $W_{xz}, W_{sz}, W_{xr}, W_{sr}, W_{ss}$ 和 W_{ss} 为权重矩阵; b_z, b_r 和 b_h 为偏差向量; $\sigma(\cdot)$ 为 sigmoid 激活函数; \odot 表示点积。 BiGRU 的网络结构如图 1 所示。

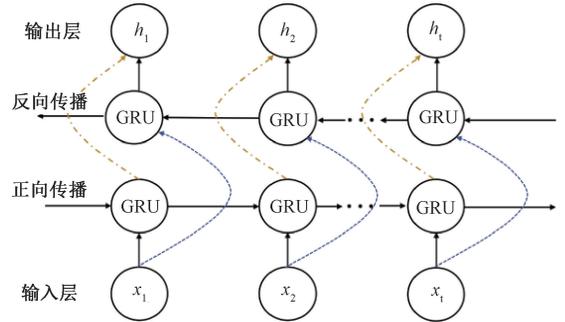


图 1 BiGRU 的网络结构图

Fig. 1 Network architecture diagram of BiGRU

1.4 Attention 原理

注意力机制是一种用于资源分配的策略, 模拟人类的注意力, 将关注集中在相关区域, 减少对无关信息的关注。通过将不同权重的隐式状态分配给 BiGRU 的输出, 结合 Attention 来增强重要信息的影响^[25]。同时, 通过映射权值和学习参数, 可强化重要信息的影响。 Attention 基本网络结构如图 2 所示。

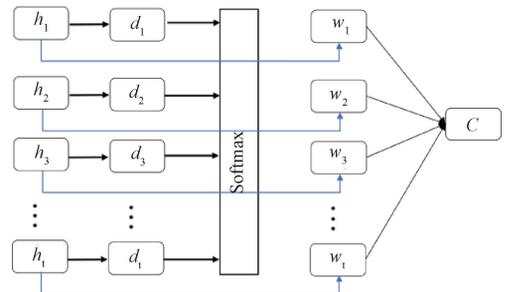


图 2 注意力机制的基本结构

Fig. 2 The basic structure of the attention mechanism

本文通过增加关键特征对 BiGRU 输出变量的影响来提高模型的预测精度, Attention 的计算公式为:

$$d_t = u \tanh(nh_t + b) \tag{11}$$

$$w_t = \frac{\exp(d_t)}{\sum_{t=1}^T \exp(d_t)} \tag{12}$$

$$C = \sum_{t=1}^T w_t h_t \tag{13}$$

式中: t 表示时间, h_t 为 BiGRU 在时刻 t 的隐藏状态输出, d_t 为 h_t 在时刻 t 的注意概念分布; u 和 n 为权重系数; b 为偏置系数; C 为权重特征。

2 基于 DE-VMD 与相关系数筛选的信号重构

本文根据最小化包络熵原则, 用 DE 优化 VMD, 得到最佳 $[K, \alpha]$ 参数组合。包络熵反映了信号的稀疏性。轴承故障信号经 VMD 算法处理后, 如果所得本征模态函数 (intrinsic mode function, IMF) 中含噪声较多, 与故障相关的周期性冲击特征不明显, 则分量信号的稀疏性较弱, 包络熵较大; 如果 IMF 分量中包含的故障特征信息较多, 波形中出现规律性冲击脉冲, 则信号将呈现出较强的稀疏特性, 包络熵值较小。包络熵 E_p 数学定义为:

$$\left\{ \begin{aligned} E_p &= - \sum_{j=1}^m p_j \lg p_j \\ p_j &= a(j) / \sum_{j=1}^m a(j) \\ a(j) &= \sqrt{[x(j)]^2 + \{H[x(j)]\}^2} \end{aligned} \right. \tag{14}$$

式中: $a(j)$ 表示为信号 $x(j)$ ($j = 1, 2, \dots, m$) 通过希尔伯特解调后所得的包络信号序列, H 表示信号的 Hilbert 变换, p_j 是 $a(j)$ 的归一化形式。

在得到最佳分解层数 K 后, 部分 IMF 分量因与原信号相关性较差, 为无用 IMF 分量, 若参与信号重构则影响故障特征信息, 因此计算在最佳 $[K, \alpha]$ 参数组合下的 IMF 分量与原信号的相关系数, 定义公式如下:

$$\rho_{XY} = \frac{E[(X - \mu_X)(Y - \mu_Y)]}{\sigma_X \sigma_Y} \tag{15}$$

式中: ρ_{XY} 是信号 X 和 Y 的相关系数, $E[(X - \mu_X)(Y - \mu_Y)]$ 是协方差, μ_X 和 μ_Y 分别表示信号 X 和 Y 的均值, σ_X 和 σ_Y 分别表示信号 X 和 Y 的标准差。 ρ_{XY} 值趋近于 1 说明两个变量之间存在较强的正相关关系, 保留了大部分原始信号故障特征信息, 而 ρ_{XY} 值越趋近于 0 则表明二者之间的相关性较弱, 即几乎没保留原信号的主要故障特征信息。白毅翔等^[26]将 $\rho_{XY} < 0.1$ 的 IMF 分量认为其为无用信号不参与下一步信号重构。目对相关系数的阈值没有绝对统一的标准, Yang 等^[27]认为相关系数小于 0.2 即为无用信号, 因此本文定义当 $\rho_{XY} < 0.2$ 即为无用信号, 不参与下一步信号重构。所提出的基于 DE 优化 VMD 与相关系数筛选的信号重构的流程图如图 3 所示。

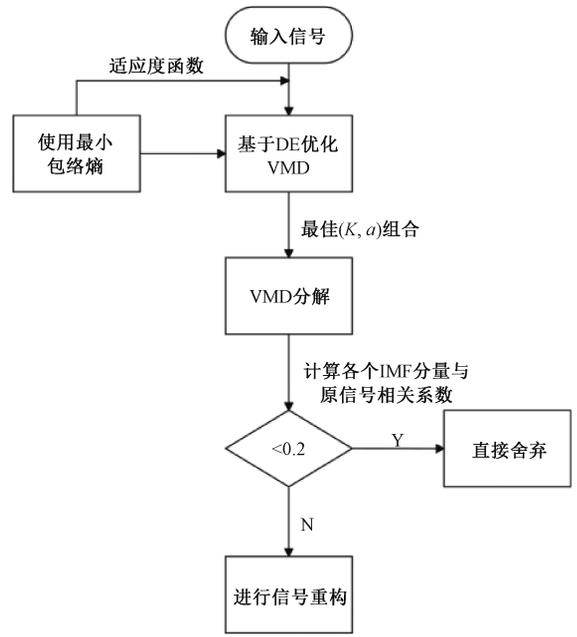


图 3 信号重构流程图

Fig. 3 Signal reconstruction flowchart

3 轴承故障诊断模型构建

本节将详细介绍基于 VMD, CNN-BiGRU 和 Attention 的滚动轴承故障诊断方法。该方法综合利用了 VMD 的信号分解能力、CNN-BiGRU 的时间序列特征提取能力, 以及 Attention 的加权聚焦故障特征能力, 从而实现了对轴承故障的高效、准确诊断。基于 DE 优化 VMD 和 CNN-BiGRU-Attention 模型结合的轴承故障诊断方法的整体流程如图 4 所示。

第 1 步: 采集轴承不同状态下的振动信号;

第 2 步: DE 参数初始化, 分解层数 K 和惩罚因子 α 取值范围为 $K = [2, 10]$, $\alpha = [100, 5\ 000]$ 。将振动信号首先进行 VMD 分解得到不同的 IMF 分量, 计算每个 IMF 分量的与原信号的相关系数值, 设定相关系数 0.2 为阈值, 小于阈值 IMF 分量舍弃, 大于阈值的完成信号重构;

第 3 步: 制作数据集。将重构信号后进行特征提取, 以 512 个数据为一个样本进行特征提取。轴承工况有内圈故障、保持架故障、外圈故障、正常轴承 4 种, 按工况情况分别打上 0、1、2、3 标签, 每种轴承工况有 200 个样本, 将数据集随机分割为训练集和测试集, 训练集和测试集划分比例为 7 : 3;

第 4 步: 对振动信号样本进行标准分数 (z-score) 标准化处理, 以消除不同特征量纲的影响和抑制异常值干扰。z-score 标准化处理数学定义为: 对于单个振动信号样本 $x = [x_1, x_2, \dots, x_N]$, 其标准化公式为:

$$\hat{x}_i = \frac{x_i - \mu_x}{\sigma_x}, i = 1, 2, \dots, N \tag{16}$$

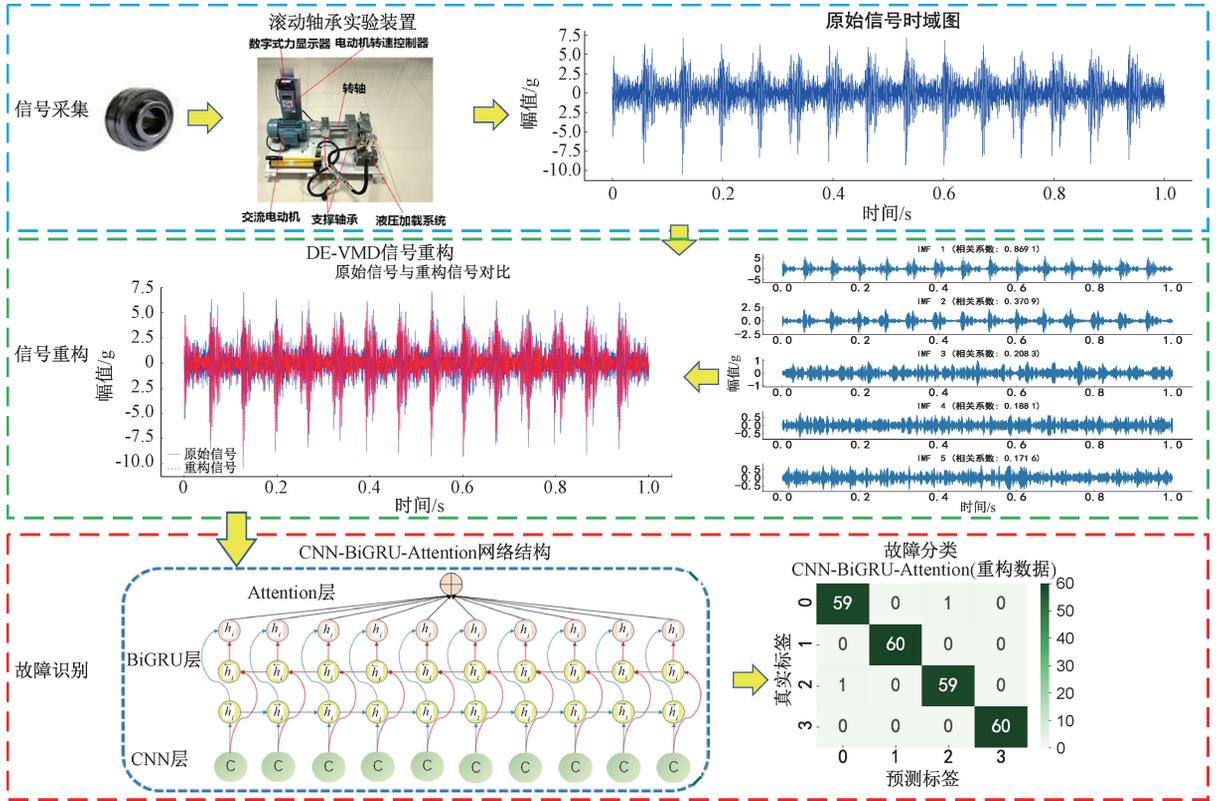


图 4 故障诊断方法流程图

Fig. 4 Fault diagnosis methodology flowchart

式中: μ_x 为当前样本的均值, σ_x 为当前样本的标准差;

第 5 步:将数据输入到 CNN-BiGRU-Attention 模型中进行训练,提取特征信息并学习故障特征;

第 6 步:用测试集测试训练模型的性能,计算评价指标,输出分类标签,预测滚动轴承故障类型。

4 实验结果与分析

4.1 数据集介绍

XJTU-SY 轴承数据集^[28]由西安交通大学设计科学与基础部件研究所提供。有两个 PCB 352C33 单向加速度传感器分别通过磁座固定于测试轴承的水平和竖直方向上。有 4 种模式,包括正常状态,内圈损坏,外圈损坏和保持架损坏。数据集包含 3 种运行条件,第 1 组转速 2 100 rpm,径向力 12 kN,第 2 组转速 2 250 rpm,径向力 11 kN,第 3 组转速 2 400 rpm,径向力 10 kN。该轴承故障诊断试验台如图 5 所示。本文采用 25.6 kHz 的采样频率,转速 2 250 rpm,径向力 11 kN 下的滚动轴承试验数据。

4.2 信号重构

根据最小包络熵适应度函数,用 DE 优化下的 VMD 方法对内圈故障振动信号进行分解,适应度函数值随迭代次数的增加而降低直至达到收敛状态,如图 6 所示。由图 6 可知,适应度值迭代到 24 次时达到最小,此时对应的 $[K, a]=[5, 4 199.67]$ 。

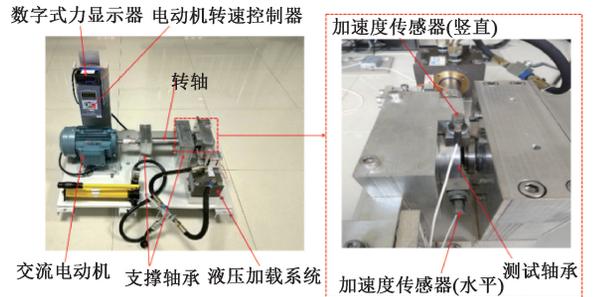


图 5 滚动轴承试验台

Fig. 5 Rolling bearing test rig

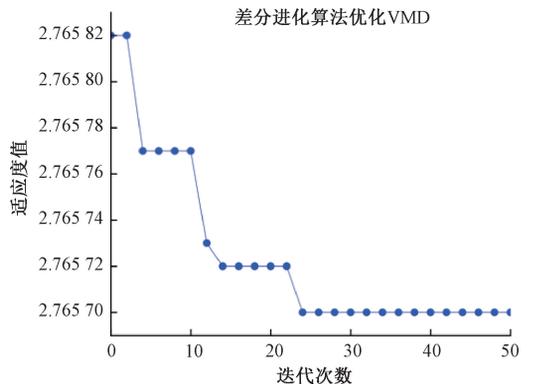


图 6 适应度值收敛曲线

Fig. 6 Fitness value convergence curve

依次用 DE 优化下的 VMD 方法对滚动体故障、内外圈故障和正常轴承信号进行分解得到每种轴承状态下 VMD 分解中的最优分解层数 K 和惩罚因子 α , 结果如表 1 所示。

计算各个工况在最佳 $[K, \alpha]$ 参数组合下, VMD 分解得到的 IMF 分量的相关系数值, 结果如表 2 所示。

按照本文方法, 依次将各种轴承工况下的 IMF 分量进行筛选完成信号重构, 图 7 为内圈故障信号 IMF 分量图, 图 8 为内圈故障信号重构时域图, 可以从图中看出, 该方法降噪效果明显。

表 1 VMD 分解的最佳层数和惩罚因子

Table 1 The optimal number of decomposition layers and penalty factor in VMD

标签	轴承工况	K	α
0	内圈故障	5	4 199.67
1	保持架故障	4	1 533.78
2	外圈故障	6	4 962.54
3	正常轴承	7	3 117.57

表 2 各个 IMF 分量的相关系数值

Table 2 Correlation coefficient values of individual IMF components

标签	轴承工况	IMF1	IMF2	IMF3	IMF4	IMF5	IMF6	IMF7
0	内圈故障	0.869 1	0.370 9	0.208 3	0.188 1	0.171 6	—	—
1	保持架故障	0.668 4	0.600 0	0.350 0	0.306 0	—	—	—
2	外圈故障	0.563 5	0.683 0	0.245 5	0.221 2	0.150 8	0.228 3	—
3	正常轴承	0.584 1	0.473 0	0.368 3	0.261 6	0.369 7	0.294 9	0.232 7

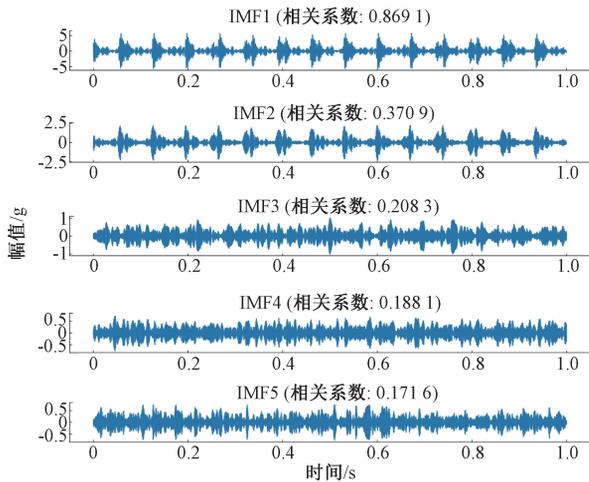


图 7 VMD 分解图

Fig. 7 VMD decomposition diagram

有正样本数的比例。精确率(Precision)是模型正确预测为正类的样本数占模型预测为正类的总样本数的比例。F1 评分(F1 Score)是精确率和召回率的调和平均数, 用于综合评估模型的性能。其数学表达式定义为:

$$\left. \begin{aligned} Accuracy &= \frac{TP + TN}{TP + TN + FN + FP} \\ Precision &= \frac{TP}{TP + FP} \\ Recall &= \frac{TP}{TP + FN} \\ F1\ Score &= 2 \times \frac{Precision \times Recall}{Precision + Recall} \end{aligned} \right\} \quad (17)$$

实验设置两组对比实验, 第一组实验将经过 DE-VMD 信重构后的信号输入到基础模型中的 SVM、CNN、BiGRU、CNN-BiGRU 及目前先进方法中的深度残差网络(deep residual network, ResNet)和基于 CNN-Transformer 模型进行对比, 以验证 CNN-BiGRU-Attention 模型的优越性。第二组实验将原始信号输入到 CNN-BiGRU-Attention 模型中, 用来验证本文降噪方法的可行性。得到它们的测试集测试结果如图 9、图 10 和表 3 所示。

从图 9~10 和表 3 可知, CNN-BiGRU-Attention 相对于其他的模型, 在准确率、精确率、召回率和 F1 评分中均取得了最佳性能。CNN-BiGRU-Attention(重构数据)与单独的 CNN 相比, 在准确率提高了 17.92%, F1 评分提高了 18.11%。CNN-BiGRU-Attention(重构数据)与 BiGRU 模型相比, 准确率提高了 18.75%, F1 分数提高了 19.31%。CNN-BiGRU-Attention(重构数据)与 CNN-BiGRU 模型相比, 准确率提高了 7.5%。经典模型中的 SVM, 准确率仅为 77.92%, 效果较差。ResNET 和 CNN-

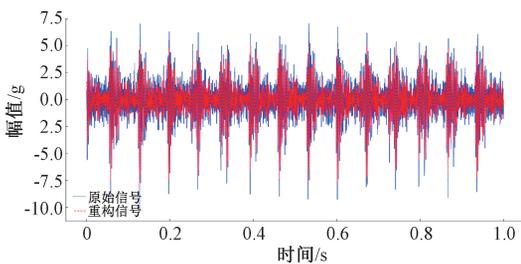


图 8 原始信号和重构信号对比

Fig. 8 Comparison between original and reconstructed signals

4.3 故障识别

在故障诊断中, 准确率(Accuracy)是模型预测正确的样本数占总样本数的比例, 反映了模型的整体预测能力。召回率(Recall), 是模型正确预测为正类的正样本数所占

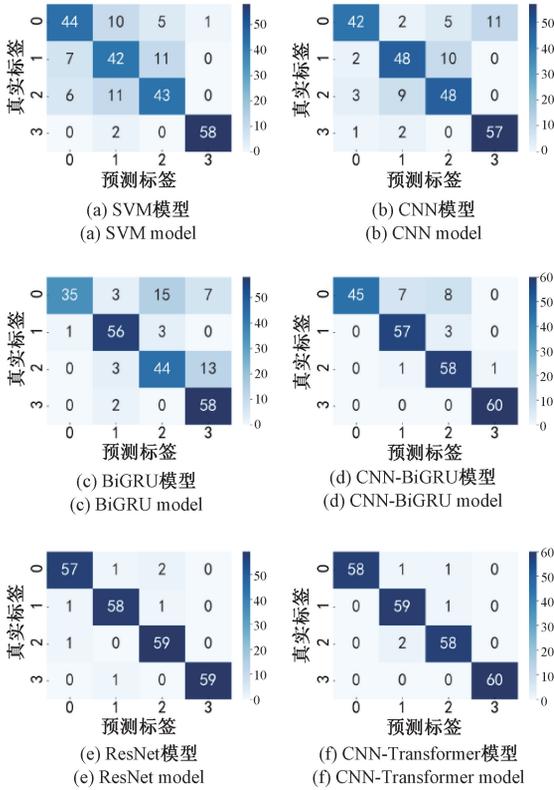


图 9 常用和先进模型故障分类混淆矩阵图

Fig. 9 Confusion matrices of conventional and advanced fault classification models

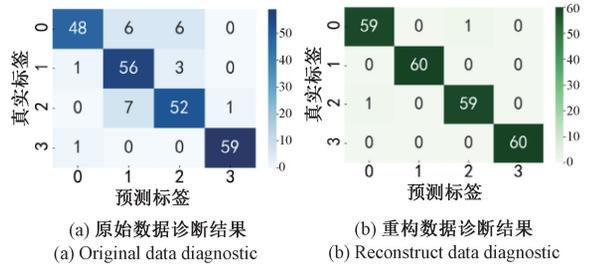


图 10 本文所提模型故障分类混淆矩阵图

Fig. 10 Confusion matrix of fault classification using the proposed model

Transformer 准确率为 97.08% 和 97.92%，而 CNN-BiGRU-Attention 模型在准确率上分别实现了 2.09% 和 1.25% 的提升。这得益于 CNN-BiGRU-Attention 模型通过卷积局部感知、双向时序建模和注意力聚焦的三阶段协同，在故障诊断中实现了更高精度的要求。DE 优化的 VMD 为模型输入提供了最优化的、任务相关的信号重构基础；而 CNN-BiGRU-Attention 模型则在此最优基础上，进行了深度、自适应的特征提取与预测。原始数据在没有经过 DE-VMD 减噪完成信号重构，直接输入 CNN-BiGRU-Attention 模型方法中准确率为 89.58%，低于重构后的 99.17% 的准确率，这也证明了所提出的 DE-VMD 信号降噪重构方法的必要性。

其中为了防止实验结果具有偶然性，用 CNN-BiGRU-

表 3 不同模型的故障诊断结果

Table 3 Fault diagnosis results comparison of different models

模型	Accuracy	Precision	Recall	F1 Score
SVM	77.92	78.25	77.92	78.04
CNN	81.25	81.55	81.25	81.06
BiGRU	80.42	82.51	80.42	79.86
CNN-BiGRU	91.67	92.53	91.67	91.50
ResNet	97.08	97.11	97.08	97.09
CNN-Transformer	97.92	97.96	97.92	97.92
CNN-BiGRU-Attention(原始数据)	89.58	90.18	89.58	89.59
CNN-BiGRU-Attention(重构数据)	99.17	99.17	99.17	99.17

Attention 模型对实验数据进行了 10 次验证，得到 10 次诊断结果的平均值和波动情况如表 4 所示。

表 4 10 次实验的平均故障诊断结果

Table 4 Average fault diagnosis results from 10 experimental trials

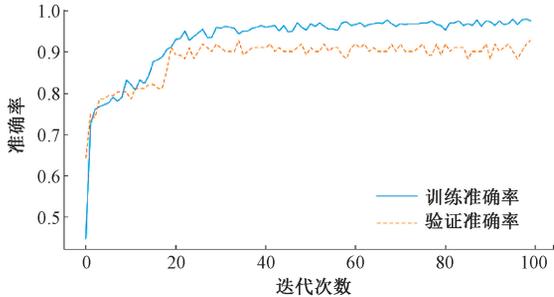
故障诊断指标	平均值±方差
准确率	99.17±0.09
精确率	99.17±0.08
回归率	99.17±0.09
F1 评分	99.17±0.08

根据表 4 结果可知，CNN-BiGRU-Attention 模型在 10 次实验中，准确率平均结果达到了 99.17%，具有良好的诊断效果。

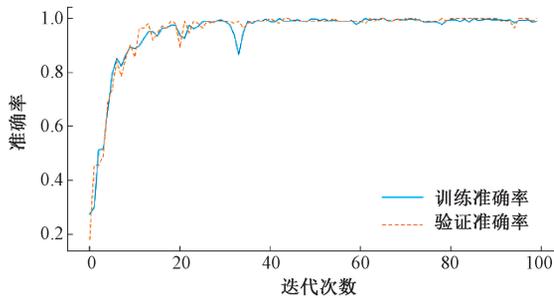
本文将 CNN-BiGRU-Attention 与 CNN-BiLSTM-Attention 模型进行对比，二者都是处理序列数据的双向循环神经网络。LSTM 由 3 个功能门（输入门、遗忘门和输出门）和 1 个候选记忆计算组成，而 GRU 的核心门控由 2 个功能门（更新门和重置门）及候选隐藏状态计算组成。LSTM 和 GRU 的总参数量计算公式为：

$$\begin{cases} \text{LSTM} = 4 \times (d_h + d_{in}) \times d_h \\ \text{GRU} = 3 \times (d_h + d_{in}) \times d_h \end{cases} \quad (18)$$

式中: d_m 为输入特征的维度, d_h 为隐藏状态的维度。GRU简化了LSTM的结构少一个门控, 参数量减少25%, GRU的更新门机制能更快过滤无关历史信息, LSTM的遗忘门则需要更多迭代学习长期依赖。同等情况下, 使用GRU模型会比LSTM模型在算力资源和训练速度上更有优势。为了验证BiGRU比BiLSTM模型简单高效, 本文进行对比实验, 其中实验电脑配置中央处理(central processing unit, CPU)为AMD5600GT, 图形处理器(graphic processing unit, GPU)为NVIDIA GeForce RTX4060, 实验数据不变, 依次用CNN-BiLSTM-Attention与CNN-BiGRU-Attention模型进行验证, 得到结果如图11和表5所示, 可以看出BiGRU模型各方面性能更好, CNN-BiGRU-Attention比CNN-BiLSTM-Attention模型准确率高1.25%, GPU占用率降低21%, CPU占用率降低23%, 训练时间快35s, 结果验证了BiGRU模型更简单高效。



(a) CNN-BiLSTM-Attention模型训练结果
(a) CNN-BiLSTM-Attention model training results



(b) CNN-BiGRU-Attention模型训练结果
(b) CNN-BiGRU-Attention model training results
图11 模型训练准确率迭代图

Fig. 11 Model training accuracy iteration chart

表5 模型性能对比表

Table 5 Model performance comparison table

模型	Accuracy	GPU 占用率	CPU 占用率	训练时 间/s
CNN-BiLSTM-Attention	97.92%	44%	48%	130
CNN-BiGRU-Attention	99.17%	23%	25%	95

5 结 论

提出一种基于DE优化VMD与CNN-BiGRU-Attention模型相结合的滚动轴承故障诊断方法。利用DE

优化VMD进行信号重构, 解决了VMD在信号降噪过程中人为选择分解参数的难题, 降低了噪声对特征信号的干扰。CNN-BiGRU-Attention模型与SVM、CNN、BiGRU、CNN-BiGRU、ResNet和CNN-Transformer模型相比, 具有更强的提取信号中的故障特征信息能力, 故障分类准确率更高。此外, 还进行了CNN-BiGRU-Attention与CNN-BiLSTM-Attention的对比实验, 证明同为双向循环神经网络, BiGRU比BiLSTM模型简单高效, 训练速度快, 在算力资源有限情况下, 优先选择BiGRU模型。本文只对单一轴承故障工况进行了实验验证, 在未来的工作中, 考虑建立新的故障识别模型, 对轴承复合故障工况进行故障检测。

参考文献

- [1] 刘馨雅, 马超, 黄民, 等. 变转速工况下基于角度重采样与PCA-XGBoost轴承故障诊断方法研究[J]. 电子测量与仪器学报, 2024, 38(3): 45-54.
LIU X Y, MA CH, HUANG M, et al. Research on angle resampling and PCA-XGBoost bearing fault diagnosis method under variable speed working condition[J]. Journal of Electronic Measurement and Instrumentation, 2024, 38(3): 45-54.
- [2] 赵颖超, 张菡, 岳新宇, 等. 基于AELSTM模型迁移学习的滚动轴承剩余寿命预测[J]. 国外电子测量技术, 2024, 43(2): 43-50.
ZHAO Y CH, ZHANG W, YUE X Y, et al. Rolling bearing remaining life prediction based on AELSTM and model transfer learning[J]. Foreign Electronic Measurement Technology, 2024, 43(2): 43-50.
- [3] LI N, LEI Y, LIN J, et al. An improved exponential model for predicting remaining useful life of rolling element bearings[J]. IEEE Transactions Industrial Electronics, 2015, 62(12): 7762-7773.
- [4] ZENG M, ZHANG W M, CHEN Z. Group-based K-SVD denoising for bearing fault diagnosis[J]. IEEE Sensors Journal, 2019, 19(15): 6335-6343.
- [5] 李灿, 王广斌, 赵树标, 等. 基于多重流形标签传播的滚动轴承故障诊断方法[J]. 振动与冲击, 2025, 44(1): 121-133.
LI C, WANG G B, ZHAO SH B, et al. Rolling bearing fault diagnosis method based on multi-manifold label propagation[J]. Journal of Vibration and Shock, 2025, 44(1): 121-133.
- [6] 王爽, 丁传仓, 曹懿, 等. 加权多尺度卷积稀疏表示及其在滚动轴承复合故障诊断中的应用[J]. 仪器仪表学报, 2024, 45(5): 197-207.
WANG SH, DING CH C, CAO Y, et al. Weighted multiscale convolutional sparse representation and its application in rolling bearing compound fault

- diagnosis[J]. Chinese Journal of Scientific Instrument, 2024, 45(5):197-207.
- [7] LU H F, ZHOU K D, HE L. Bearing fault vibration signal denoising based on adaptive denoising autoencoder [J]. Electronics, 2024, 13 (12): 2403-2403.
- [8] 刘金锁, 张平松, 刘盛东, 等. 基于 EEMD 和 DE-PNN 的随钻测震岩性识别技术研究[J]. 煤炭学报, 2025, 50(12):5182-5192.
LIU J S, ZHANG P S, LIU SH D, et al. Research on lithology identification technology of seismic measurement while drilling based on EEMD and DE-PNN[J]. Journal of China Coal Society, 2025, 50(12): 5182-5192.
- [9] 周顺勇, 陆欢, 胡琴, 等. 基于 SVD 与混合神经网络模型的自动调制识别[J]. 电子测量技术, 2024, 47(21): 111-121.
ZHOU SH Y, LU H, HU Q, et al. Automatic modulation recognition based on SVD and hybrid neural network model [J]. Electronic Measurement Technology, 2024, 47(21):111-121.
- [10] 李可, 牛园园, 宿磊, 等. 参数优化 VMD 的滚动轴承故障诊断方法[J]. 振动工程学报, 2023, 36(1):280-287.
LI K, NIU Y Y, SU L, et al. Rolling bearing fault diagnosis method based on parameter optimization VMD[J]. Journal of Vibration Engineering, 2023, 36(1):280-287.
- [11] GU J, PENG Y X, LU H, et al. A novel fault diagnosis method of rotating machinery via VMD, CWT and improved CNN[J]. Measurement, 2022, 200: 111635.
- [12] GU J M, TIAN F M, SHI J X, et al. Noise reduction and analysis of leaf electrical signals of strap-leaved plants based on VMD-EWT [J]. Computers and Electronics in Agriculture, 2024, 226: 109441.
- [13] 马荣增, 杜泽辉, 李文, 等. 基于 OOA-VMD 的隧道爆破振动信号降噪方法[J]. 爆破, 2025, 42(1):175-182.
MA R Z, DU Z H, LI W, et al. A noise reduction method for tunnel blasting vibration signals based on OOA-VMD[J]. Blasting, 2025, 42(1):175-182.
- [14] 孙姿姣, 周湘贞, 李松洋. 鹈鹕算法参数优化 VMD 联合 SVDS 的电机轴承故障诊断[J]. 机械设计, 2025, 42(4):150-155.
SUN Z J, ZHOU X ZH, LI S Y. Bearing fault diagnosis of electric machinery based on VMD with parameters optimized by pelican optimization algorithm (POA) and singular value difference spectrum (SVDS) [J]. Journal of Machine Design, 2025, 42(4):150-155.
- [15] 文世坤, 吉爱红, LEE Heow Pueh, 等. 面向并联机构正运动学方程多根问题的改进差分进化算法[J]. 机械工程学报, 2025, 61(19):43-53.
WEN SH K, JI AI H, LEE H P, et al. Improvement of differential evolution algorithm for solving the multi-root in forward kinematic equations of parallel mechanisms[J]. Journal of Mechanical Engineering, 2025, 61(19):43-53.
- [16] 栾孝驰, 汤捷中, 沙云东. 基于蜣螂算法优化深度极限学习机的中介轴承故障诊断方法[J]. 振动与冲击, 2024, 43(21):96-106, 127.
LUAN X CH, TANG J ZH, SHA Y D. Inter-shaft fault diagnosis method based on deep extreme learning machine optimized with dung beetle optimizer [J]. Journal of Vibration and Shock, 2024, 43(21):96-106, 127.
- [17] NI Z, TONG Y F, SONG Y X, et al. Enhanced bearing fault diagnosis in nc machine tools using dual-stream CNN with vibration signal analysis [J]. Processes, 2024, 12(9):1951.
- [18] 沈伟杰, 肖茂华, 宋新民, 等. 基于连续小波变换的 CNN—SVM 农机滚动轴承故障诊断[J]. 中国农机化学报, 2025, 46(4):254-264.
SHEN W J, XIAO M H, SONG X M, et al. Fault diagnosis of agricultural machinery rolling bearing based on CNN—SVM by continuous wavelet transform [J]. Journal of Chinese Agricultural Mechanization, 2025, 46(4):254-264.
- [19] FU G H, WEI Q J, YANG Y S, et al. Bearing fault diagnosis based on CNN-BiLSTM and residual module[J]. Measurement Science and Technology, 2023, 34(12), DOI: 10.1088/1361-6501/acf598.
- [20] HOU D, CHE X X, LI F F, et al. Prediction of anaerobic digestion performance by quantum convolutional reconstruction gated recurrent neural network [J]. Physica Scripta, 2024, 99 (5), DOI: 10.1088/1402-4896/ad3e3c.
- [21] STORN R, PRICE K. Differential evolution-A simple and efficient heuristic for global optimization over continuous spaces[J]. Journal of Global Optimization, 1997, 11(4): 341-359.
- [22] 李浩然, 刘德平. 基于参数优化 VMD 和改进 GoogLeNet 的滚动轴承故障诊断[J]. 机械传动, 2025, 49(1):163-170.
LI H R, LIU D P. Rolling bearing fault diagnosis based on parameter optimization VMD and improved

- GoogLeNet[J]. Journal of Mechanical Transmission, 2025, 49(1):163-170.
- [23] 全睿,程功,周宇龙,等.基于EWOA优化CNN-BiGRU-AT模型的燃料电池衰退预测[J/OL]. 电工技术学报, 1-17 [2025-09-06]. <https://doi.org/10.19595/j.cnki.1000-6573.tce.s.241835>.
- QUAN R, CHENG G, ZHOU Y L, et al. Optimization of CNN-BiGRU-AT model based on EWOA Fuel cell decay prediction [J/OL]. Transactions of China Electrotechnical Society, 1-17 [2025-09-06]. <https://doi.org/10.19595/j.cnki.1000-6573.tces.241835>.
- [24] WANG F K, AMOGNE Z E, TSENG C, et al. A hybrid method for online cycle life prediction of lithium-ion batteries [J]. International Journal of Energy Research, 2022, 46(7): 9080-9096.
- [25] 赵洪山,吴雨晨,温开云,等.基于时空注意力机制的台区多用户短期负荷预测[J]. 电工技术学报, 2024, 39(7):2104-2115.
- ZHAO H SH, WU Y CH, WEN K Y, et al. Multi-user short-term load forecasting based on spatio-temporal attention mechanism [J]. Transactions of China Electrotechnical Society, 2024, 39 (7): 2104-2115.
- [26] 白毅翔,陈志英,张修伦,等.基于改进VMD-WT的油浸式变压器局部放电超声信号去噪方法[J]. 电子测量与仪器学报, 2024, 38(12):237-249.
- BAI Y X, CHEN ZH Y, ZHANG X L, et al. Ultrasonic signal denoising method for partial discharge of oil-immersed transformer based on improved VMD-WT [J]. Journal of Electronic Measurement and Instrumentation, 2024, 38 (12): 237-249.
- [27] YANG J, BAI Y P, CHENG Y Y, et al. A new model for bearing fault diagnosis based on optimized variational mode decomposition correlation coefficient weight threshold denoising and entropy feature fusion [J]. Nonlinear Dynamics, 2023, 111(18): 17337-17367.
- [28] WANG B, LEI Y G, LI N P, et al. A hybrid prognostics approach for estimating remaining useful life of rolling element bearings[J]. IEEE Transactions on Reliability, 2020, 69(1): 401-412.

作者简介

邵鲁川,硕士研究生,主要研究方向为信号处理和故障诊断。

E-mail:luchuanshao@qhu.edu.cn

赵兵(通信作者),博士,副教授,主要研究方向为装配工艺基础理论、数字化与智能化制造相关研究工作。

E-mail:zzipwe@126.com

康旭涛,硕士研究生,主要研究方向为深度学习和故障诊断。

E-mail:kxt121412@163.com