

DOI:10.3969/j.issn.1001-4551.2023.11.008

# 基于 ICEEMDAN 和 IMWPE-LDA-BOA-SVM 的齿轮箱 损伤识别模型\*

王 洪<sup>1</sup>, 张锐丽<sup>1</sup>, 吴 凯<sup>2</sup>

(1. 宜宾职业技术学院 智能制造学院, 四川 宜宾 644003; 2. 成都理工大学 机电工程学院, 四川 成都 610059)

**摘要:**针对齿轮箱振动信号中的背景噪声过大影响故障特征质量,进而降低故障识别准确率的问题,提出了一种基于改进自适应噪声完备集成经验模态分解(ICEEMDAN)、改进多尺度加权排列熵(IMWPE)、利用线性判别分析(LDA)、蝴蝶优化算法(BOA)优化支持向量机(SVM)的齿轮箱故障诊断方法(ICEEMDAN-IMWPE-LDA-BOA-SVM)。首先,采用ICEEMDAN对齿轮箱振动信号进行了分解,生成了一系列从低频到高频分布的本征模态函数分量;接着,基于相关系数筛选出包含主要故障信息本征模态函数分量,进行了信号重构,降低了信号的噪声;随后,提出了改进多尺度加权排列熵的非线性动力学指标,并利用其提取了重构信号的故障特征,以构建反映齿轮箱故障特性的故障特征;然后,利用线性判别分析(LDA)对原始故障特征进行了压缩,以构建低维的故障特征向量;最后,采用蝴蝶优化算法(BOA)对支持向量机(SVM)的惩罚系数和核函数参数进行了优化,以构建参数最优的故障分类器,对齿轮箱的故障进行了识别;基于齿轮箱复合故障数据集对ICEEMDAN-IMWPE-BOA-SVM方法进行了实验和对比分析。研究表明:该方法能够较为准确地识别齿轮箱的不同故障类型,准确率达到99.33%,诊断时间只需5.31 s,在多个方面都优于其他对比方法,在齿轮箱的故障诊断中更具有应用潜力。

**关键词:**故障特征提取;信号分解及信号重构;特征降维;改进自适应噪声完备集成经验模态分解;改进多尺度加权排列熵;线性判别分析;蝴蝶优化算法;支持向量机

中图分类号:TH132.41

文献标识码:A

文章编号:1001-4551(2023)11-1709-09

## Gear box damage identification model based on ICEEMDAN and IMWPE-LDA-BOA-SVM

WANG Hong<sup>1</sup>, ZHANG Ruili<sup>1</sup>, WU Kai<sup>2</sup>

(1. Intelligent Manufacturing College, Yibin Vocational and Technical College, Yibin 644003, China;

2. School of Electromechanical Engineering, Chengdu University of Technology, Chengdu 610059, China)

**Abstract:** Aiming to address the issue of excessive background noise in gearbox vibration signals affecting the quality of fault features and thereby reducing the accuracy of fault identification, a gearbox fault diagnosis method (ICEEMDAN-IMWPE-LDA-BOA-SVM) based on improved adaptive noise complete ensemble empirical mode decomposition (ICEEMDAN), improved multi-scale weighted permutation entropy (IMWPE), linear discriminant analysis (LDA), butterfly optimization algorithm (BOA), and support vector machine (SVM) optimization was proposed. Firstly, ICEEMDAN was used to decompose the gearbox vibration signal and generate a series of intrinsic mode function components distributed from low frequency to high frequency. Next, based on the correlation coefficient, the intrinsic mode function components containing the main fault information were selected for signal reconstruction to reduce signal noise. Subsequently, a nonlinear dynamic index for improving multi-scale weighted permutation entropy was proposed, and it was used to extract fault features of the reconstructed signal to construct fault features that reflect the fault characteristics of the gearbox. Then, linear discriminant analysis (LDA) was used to compress the original fault features to construct a low dimensional fault feature vector. Finally, the butterfly optimization algorithm (BOA) was used to optimize the penalty coefficients and kernel function parameters of the support vector machine (SVM), in order to construct a fault classifier with the optimal parameters, achieving fault identification of the gearbox. Experimental and comparative studies

收稿日期:2023-03-13

基金项目:国家自然科学基金资助项目(51875479)

作者简介:王洪(1981-),男,四川宜宾人,讲师,主要从事机械设计与自动化方面的研究。E-mail:rvz4u6d@163.com

were conducted on the ICEEMDAN-IMWPE-BOA-SVM method based on the gearbox composite fault dataset. The research results show that the method can accurately identify different fault types of gearboxes, with an accuracy rate of 99.33% and a diagnosis time of only 5.31 s. It is superior to other comparative methods in multiple aspects and has more potential for application in fault diagnosis of gearboxes.

**Key words:** fault feature extraction; signal decomposition and signal reconstruction; feature dimension reduction; improved complete ensemble empirical mode decomposition with adaptive noise (ICEEMDAN); improved multi-scale weighted permutation entropy (IMWPE); linear discriminant analysis (LDA); butterfly optimization algorithm (BOA); support vector machine (SVM)

## 0 引 言

齿轮箱是一种常用的传动系统部件,目前已被广泛应用于汽车、航空航天、化工等领域。由于齿轮箱的结构复杂,并且运行工况不稳定,导致齿轮箱内部的零部件极易发生疲劳失效,导致整个设备的振动,不利于设备的安全运行<sup>[1]</sup>。

齿轮箱发生故障后,其振动信号中包含大量与故障特征无关的谐波信息,干扰其故障特征的提取<sup>[2]</sup>。因此,有必要采取信号分解方法对源信号进行处理,以降低噪声的干扰<sup>[3]</sup>。

高淑芝等人<sup>[4]</sup>提出了一种基于集成经验模态分解(ensemble empirical mode decomposition, EEMD)的故障诊断方法,采用该方法准确地识别出了滚动轴承的故障信息;然而 EEMD 的运行效率极为低下,且引入噪声进行辅助分解,会导致分解分量残留部分噪声。为此,肖俊青等人<sup>[5]</sup>采用自适应噪声完备集成经验模态分解方法,对故障信号进行了处理,结果表明了其效果优于 EEMD,故障信号分解得更为精确;然而其分解的信号依然残留少量噪声。针对该问题,李铭等人<sup>[6]</sup>将改进自适应噪声完备集成经验模态分解(ICEEMDAN)用于处理滚动轴承振动信号,并进行了实验,结果验证了 ICEEMDAN 的优越性,其能够更好地去除信号中的噪声。

基于此,笔者采用 ICEEMDAN 来剔除齿轮箱振动信号中的噪声,以减少噪声对后续分析的干扰。

齿轮箱振动信号为非线性的一维数据,通常需要采用基于熵值的非线性分析方法进行处理<sup>[7]</sup>。常见的熵值分析方法包括样本熵、排列熵、多尺度熵等。排列熵(permutation entropy, PE)具有计算效率高、抗噪性强和性能优异等特点,被广泛用于故障诊断的研究中。

顾云青等人<sup>[8]</sup>提出了一种基于 ICEEMDAN 和 PE 的滚动轴承故障诊断方法,并将其用于轴承故障诊断,诊断结果验证了排列熵的性能;然而 PE 忽略了信号的幅值信息。为此,ZHOU Shen-han 等人<sup>[9]</sup>将加权排列熵(weighted permutation entropy, WPE)用于表征滚动轴承的健康状态,准确检测和识别了滚动轴承的故

障;然而 WPE 只是进行了单一尺度的分析。YUAN Xu-yi<sup>[10]</sup>提出了多尺度加权排列熵(multiscale weighted permutation entropy, MWPE),采用 MWPE 有效提取出了隔膜泵的故障特征;然而 MWPE 的粗粒化处理依赖于分析数据的长度,造成熵值偏差随着尺度的增加而变大。为此,王振亚等人<sup>[11]</sup>对粗粒化方式进行了优化,提出了改进多尺度加权排列熵,并将其应用于故障诊断,结果验证了该算法的有效性;但是在分析某些信号时,基于方差改进的加权排列熵可能会失去有效性<sup>[12]</sup>。

针对上述问题,笔者采用均方根对改进多尺度加权排列熵进行改进,提出另一种形式的改进加权排列熵(IMWPE),以进一步增强算法的特征提取能力;在此基础上,提出一种基于 ICEEMDAN 信号重构、IMWPE 故障特征提取、线性判别分析(LDA)特征降维和蝴蝶优化算法优化支持向量机(BOA-SVM)的损伤识别模型(ICEEMDAN-IMWPE-LDA-BOA-SVM)。

首先,采用 ICEEMDAN 对样本进行处理,筛选出核心组件,进行故障信号重构;随后,利用 IMWPE 提取重构信号的故障特征,并采用 LDA 进行特征的降维;最后,采用基于蝴蝶算法优化的支持向量机 SVM 进行故障识别。

## 1 ICEEMDAN 信号重构

### 1.1 ICEEMDAN 算法原理

ICEEMDAN 利用数据的局部均值来获取  $k$  阶模态,能够进一步消除数据中的残留噪声和伪模态,使得分解得到的本征模态函数(intrinsic mode function, IMF)分量具有更清晰的物理意义。ICEEMDAN 对信号处理的详细过程可参考逢英等人的研究<sup>[13]</sup>。

### 1.2 分量选择原则

相关系数能够准确地刻画 IMF 分量与原始数据之间的紧密程度,根据相关系数,能够据此保留信息量更多的 IMF 分量<sup>[14]</sup>。首先,笔者计算了归一化处理后的原始数据自相关函数,以及各阶 IMF 分量自相关函数之间的相关系数。

相关系数的定义如下:

$$p_{xy} = \frac{\sum_{i=1}^n (x_i - E(x))(y_i - E(y))}{\sqrt{\sum_{i=1}^n (x_i - E(x))^2 (y_i - E(y))^2}} \quad (1)$$

式中: $E(x)$ 为数据序列  $x_i$  的均值; $E(y)$ 为数据序列  $y_i$  的均值; $n$  为数据的采样点数。

随后,定义分量的选择阈值。假设各样本经过 ICEEMDAN 分解后得到的最少 IMF 分量数量为  $m$ ,则阈值  $\delta$  定义为:

$$\delta = \frac{1}{m} \sum_{i=1}^m P_{xy}^i \quad (2)$$

式中: $P_{xy}^i$ 为各样本的第  $i$  个 IMF 分量与原始数据的相关系数。

当 IMF 分量的相关系数大于阈值时,则将其视为有效分量;否则,将其剔除。

## 2 改进多尺度加权排列熵特征提取

### 2.1 加权排列熵算法

排列熵对时间序列的幅值不敏感,对幅值相等的时间序列无法进行有效的区分<sup>[15]</sup>。

受到基于方差的加权排列熵的启发,笔者采用均方根对排列熵进行改进,提出了加权排列熵的另外一种形式,其原理如下:

1)对信号  $X = \{x_1, x_2, \dots, x_N\}$  进行  $m$  维的相空间重构,生成若干个子信号  $X_i^{(m)}$ :

$$X_i^{(m)} = (x_i, x_{i+\tau}, \dots, x_{i+(m-1)\tau}) \quad (3)$$

式中: $m$  为嵌入维数; $\tau$  为时间延迟。

2)定义各子信号的权重  $\omega_i$ :

$$\omega_i = \sqrt{\frac{\sum_{k=1}^m x_{i+(k-1)\tau}^2}{m}} \quad (4)$$

3)任意  $X_i^{(m)}$  的特征信息通过权重  $\omega_i$  和排序模式  $\pi_k$  进行表征。对于信号  $X$  共有  $K$  种排序模式,每种形式  $\pi_k$  的加权概率定义如下:

$$P_\omega(\pi_k) = \frac{\sum \left\{ \omega_i \mid 1 \leq i \leq N - (m-1)\tau, i \in Z^+, N(X_i^{(m)}) \right\}}{\sum \omega_i} \quad (5)$$

式中: $N(X_i^{(m)})$  为  $X_i^{(m)}$  的排序形式  $\pi_k$ 。

4)求解信号  $X$  的 WPE 值为:

$$WPE(X, m, \tau) = - \sum_{k=1}^K P_\omega(\pi_k) \ln P_\omega(\pi_k) \quad (6)$$

### 2.2 多尺度加权排列熵算法

多尺度加权排列熵能够有效地从多个尺度来表征

非线性数据的复杂性。对于信号  $X = \{x_1, x_2, \dots, x_N\}$ , 多尺度加权排列熵的具体计算步骤如下:

1)对信号进行粗粒化处理,生成粗粒化序列  $Y^s = \{y^s(j)\}$ 。

$$y^s(j) = \frac{1}{s} \sum_{i=(j-1)s+1}^{js} x(i), 1 \leq j \leq \frac{N}{s} \quad (7)$$

式中: $s$  为尺度因子。

2)提取不同尺度因子  $s$  下,各粗粒化序列  $Y^s$  的加权排列熵值,则得到 MWPE 如下:

$$MWPE(X, m, \tau, s) = WPE(Y^s, m, \tau) \quad (8)$$

### 2.3 改进多尺度加权排列熵

MWPE 避免了 WPE 只能进行单尺度分析的不足,然而原始的粗粒化处理会造成熵值误差随着尺度因子的增加而增大。因此,笔者平均同一尺度下全部粗粒序列的 WPE 熵值,以避免由于粗粒序列变短而造成 WPE 熵值突变现象的出现,使得复杂性分析更加准确<sup>[16]</sup>。

对于信号  $X = \{x_1, x_2, \dots, x_N\}$ ,改进多尺度加权排列熵的原理如下:

1)对信号  $X$  进行改进的粗粒化处理,生成尺度因子  $s$  下的若干个粗粒序列  $Y_k^s = \{y_{k,j_1}^s, y_{k,j_2}^s, \dots, y_{k,j_s}^s\}$ :

$$y_{k,j}^s = \frac{1}{s} \sum_{i=(j-1)s+k}^{js+k-1} x(i), 1 \leq j \leq \frac{N}{s}, 1 \leq k \leq s \quad (9)$$

2)通过平均该尺度因子  $s$  下的多个粗粒序列  $Y_k^s$  ( $k=1, \dots, s$ ) 的 WPE 值,生成 IMWPE 值,如下所示:

$$IMWPE(X, s, m, \tau) = \frac{1}{s} \sum_{k=1}^s WPE(Y_k^s, m, \tau) \quad (10)$$

当尺度因子为  $s=3$  时,传统的粗粒化处理和改进的粗粒化处理如图 1 所示。

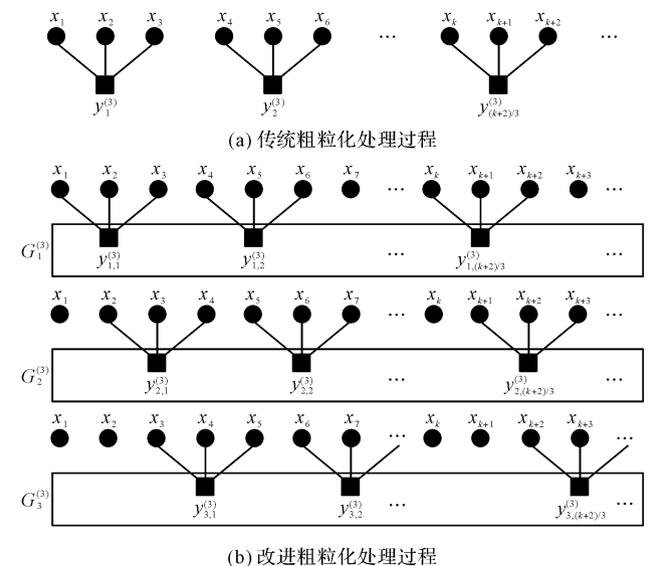


图 1 尺度因子  $s=3$  时的传统粗粒化和改进粗粒化过程  
Fig. 1 Traditional coarse grained and improved coarse-grained processing when scale factor  $s=3$

由图 1 可以发现:在传统的粗粒化过程中,粗粒序列中的元素数量会随着尺度的增加而变少,导致熵值的不稳定性。此外,传统的粗粒化处理还忽视了部分相邻样本点之间的动态关系,如在图 1(a)中, $x_3$  和  $x_4$  之间的信息就被忽略了;反之,在改进的粗粒化处理过程中,在相同尺度因子  $s$  下,改进的粗粒化处理能够获得  $s$  组粗粒序列,对这些序列进行平均处理,能够有效解决熵值不稳定的问题。

而且改进粗粒化处理通过滑动操作,充分考虑了相邻样本点之间的关系,数据的利用率更高。

IMWPE 方法的性能与嵌入维数  $m$ 、时间延迟  $\tau$  和尺度因子  $s$  有关。 $m$  选择得过小,重构向量的长度变短,算法无法有效探测信号的动力学突变;相反, $m$  设置过大,不仅降低计算效率,而且无法有效表征信号的微弱波动。因此,通常设置嵌入维数的取值范围在  $[4,7]$ ,笔者设置为  $m=5$ 。

尺度因子  $s$  的设置没有严格的要求,通常设置为  $s \geq 10$ ,笔者设置为  $s=25$ 。

时间延迟  $\tau$  对 IMWPE 的影响几乎可以忽略不计,笔者设置为  $\tau=1$ 。

### 3 LDA 特征降维和 BOA-SVM 分类识别

#### 3.1 LDA 算法原理

LDA 算法基于 Fisher 判别定理,定义子空间  $J$ ,并将数据向子空间  $J$  做投影,生成低维样本作为新的特征样本,进而实现降维的目的<sup>[17]</sup>。

假定有  $C$  个样本类别,而数据  $\mathbf{X} = \{x_i \in R^n, i=1, \dots, N\}$  则是  $N$  个  $n$  维的训练样本,则基于 Fisher 定理对空间  $J$  进行定义为:

$$J(\mathbf{W}) = \operatorname{argmax} \frac{|\mathbf{W}^T \mathbf{S}_b \mathbf{W}|}{|\mathbf{W}^T \mathbf{S}_w \mathbf{W}|} \quad (11)$$

$$\mathbf{S}_w = \sum_{j=1}^c \sum_{i=1}^{n_j} (\mathbf{x}_i^j - m_j)(\mathbf{x}_i^j - m_j)^T, \quad (12)$$

$$m_j = \frac{1}{n_j} \sum_{i=1}^{n_j} \mathbf{x}_i^j \quad (12)$$

$$\mathbf{S}_b = \sum_{j=1}^c (m_j - m)(m_j - m)^T \quad (13)$$

式中: $\mathbf{W}$  为最优投影矩阵; $\mathbf{x}_i^j$  为第  $j$  类的第  $i$  个向量; $n_j$  为第  $j$  类的向量数量; $m_j$  为第  $j$  类的均值; $m$  为全部类的均值。

当  $\mathbf{S}_w$  非奇异时,广义特征方程的  $d(d \leq C-1)$  个最大的特征值所对应的特征向量  $\mathbf{w}_i$  即为最佳投影矩阵  $\mathbf{W}$  的列向量:

$$\mathbf{S}_b \mathbf{w}_i = \lambda \mathbf{S}_w \mathbf{w}_i \quad (14)$$

笔者将数据  $x_i$  向子空间  $J$  进行投影,即  $y_i = \mathbf{W}^T x_i$ ,则  $y_i$  对应低维特征向量。

#### 3.2 蝴蝶优化算法

蝴蝶在飞行时,会在周围散发一种气味,其他蝴蝶闻到该气味后,会基于香味的浓度而被吸引。当某只蝴蝶闻到最佳香味时,会朝着散发气味的蝴蝶靠近,这个过程称为全局搜索阶段;而当蝴蝶无法闻到任何蝴蝶气味时,将在全局搜索范围内随机选择移动位置,这个过程称为局部搜索阶段<sup>[18]</sup>。

蝴蝶会在飞行时散发气味,定义气味浓度为:

$$f = cI^a \quad (15)$$

式中: $I$  为刺激强度; $c$  为感知形态; $a$  为激励系数。

蝴蝶的运动模型包含两个阶段:

##### 1) 全局搜索阶段

蝴蝶在飞行中会散发气味,其他蝴蝶会按照闻到的气味浓度来搜索目标,模型定义如下:

$$x_i(t+1) = x_i(t) + (r_1^2 \cdot g^* - x_i(t)) \cdot f_i \quad (16)$$

式中: $x_i(t)$  为迭代  $t$  时蝴蝶  $i$  的位置; $g^*$  为全局最佳解; $r_1$  为  $[0,1]$  之间的随机数; $f_i$  为蝴蝶  $i$  的气味浓度。

##### 2) 局部搜索阶段

若蝴蝶缺乏感知其他蝴蝶气味的能力,则它将在搜索范围内随机移动,模型定义为:

$$x_i(t+1) = x_i(t) + (r_2^2 \cdot x_j(t) - x_k(t)) \cdot f_i \quad (17)$$

式中: $x_j(t), x_k(t)$  为任意选择的蝴蝶; $r_2$  为  $[0,1]$  之间的随机数。

BOA 进行全局寻优或局部寻优的概率由阈值  $P$  决定,在每次迭代时会随机生成随机数  $r_3$ ,并与  $P$  进行对比,定义为:

$$x_i(t+1) = \begin{cases} x_i(t) + (r_1^2 \cdot g^* - x_i(t)) \cdot f_i, r_3 \leq P \\ x_i(t) + (r_2^2 \cdot x_j(t) - x_k(t)) \cdot f_i, r_3 > P \end{cases} \quad (18)$$

### 4 ICEEMDAN-IMWPE-LDA-BOA-SVM 方法

#### 4.1 故障诊断流程

基于上述理论介绍,笔者提出了基于 ICEEMDAN 信号重构、IMWPE 特征提取、LDA 特征降维、BOA-SVM 模式识别的齿轮箱故障诊断方法。

该方法的主要诊断步骤如下:

1) 采集齿轮箱在多个常见工况下的振动信号,并将其分割为长度为 2 048 的多组样本;

2) 采用 ICEEMDAN 对振动信号进行分解,基于 IMF 分量选择原则,筛选得到包含主要信息的 IMF 分量,并进行信号的重构;

3) 利用 IMWPE 提取重构信号的故障特征, 并采用 LDA 进行特征降维, 得到低维敏感特征, 并将其分为训练样本和测试样本;

4) 利用 BOA 算法对 SVM 模型中的参数进行优化。其中, 对于蝴蝶种群, 适应度设置为 SVM 对训练样本的识别精度;

5) 判断是否符合迭代终止条件, 若符合则输出最佳参数的  $(C, g)$  组合, 并将其代入至 SVM 模型中;

6) 将测试样本输入至优化后的 SVM 模型中, 实现故障的识别。

### 4.2 故障诊断实验

#### 4.2.1 实验平台及数据

为了检验 ICEEMDAN-IMWPE-LDA-BOA-SVM 方法, 笔者采用东南大学的齿轮箱故障数据进行实验。

齿轮箱故障实验平台如图 2 所示<sup>[19]</sup>。

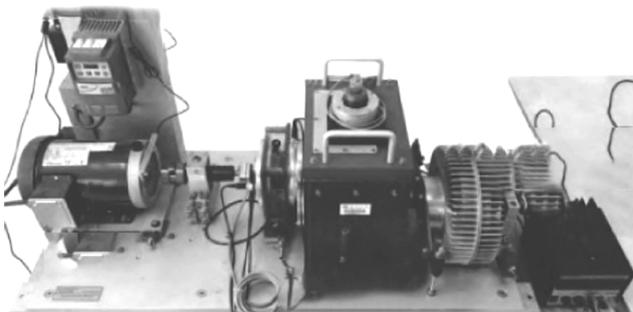


图 2 齿轮箱实验平台

Fig. 2 Gearbox experimental platform

该平台由驱动电机、电机控制器、行星齿轮箱、平行齿轮箱和制动器等组成。

振动信号由布置在行星齿轮箱表面的振动加速度计以 5 120 Hz 的频率进行采集。转速和负载分 2 两种, 分别是 20 Hz/0 V 和 30 Hz/2 V, 笔者选择 30 Hz 的振动数据进行分析; 每种工况均截取 60 个长度 2 048 的样本, 其中 30 组用于训练, 剩余 30 组用于测试。

齿轮箱故障及样本的详细信息如表 1 所示。

表 1 齿轮箱样本的详细信息

Table 1 Detailed information on gearbox samples

故障部位	故障描述	分类标签
齿轮箱	健康齿轮	1
齿轮箱	齿轮脚存在裂纹	2
齿轮箱	齿轮脚破损	3
齿轮箱	齿轮齿根存在裂纹	4
齿轮箱	齿轮表面存在磨损	5

齿轮箱振动信号的波形如图 3 所示。

#### 4.2.2 结果与分析

首先, 为了消除齿轮箱振动信号中的噪声, 笔者利用

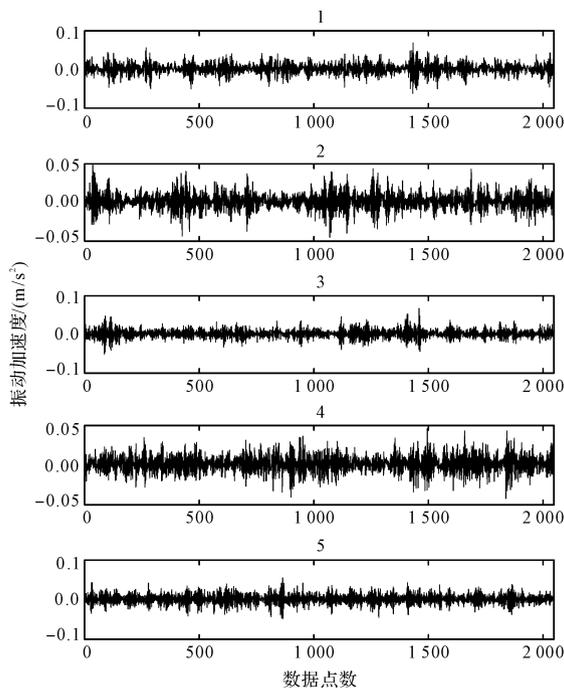


图 3 齿轮箱振动信号波形

Fig. 3 Waveform of gearbox vibration signal

ICEEMDAN 对振动信号进行分解, 并进行信号的重构。

此处以健康齿轮振动信号为例, 其 ICEEMDAN 的分解结果如图 4 所示。

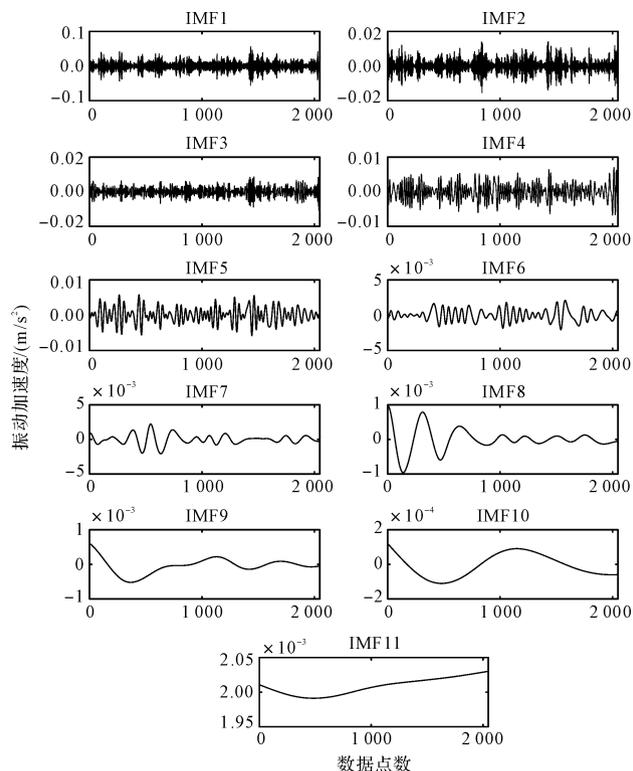


图 4 健康齿轮的 ICEEMDAN 分解波形

Fig. 4 ICEEMDAN decomposition results of healthy gears

从图 4 可以发现: 对于健康齿轮的振动信号而言, 在经过 ICEEMDAN 分解后, 不同频率成分的特征信息实现了分离, 但是高阶 IMF 分量包含的信息较少, 更

多的是噪声成分,有必要进行剔除。

为此,笔者基于相关系数,对 IMF 分量进行重要性评估。

所有样本的相关系数如图 5 所示。

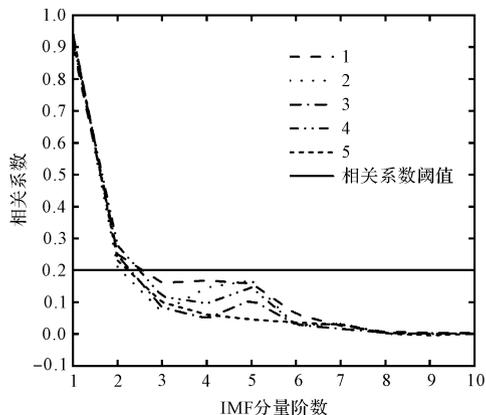


图 5 IMF 分量与原始信号的相关系数

Fig. 5 Correlation coefficient between IMF component and raw signal

经过计算,5 种工况样本的相关系数阈值均近似为 0.2,因此,笔者根据该阈值选择 IMF 分量进行信号重构。

从图 5 可以发现:前两阶 IMF 分量的相关系数均大于阈值,即这两个分量包含了原始信号的大部分故障信息,因此,笔者选择这两阶分量进行信号重构。

重构信号如图 6 所示。

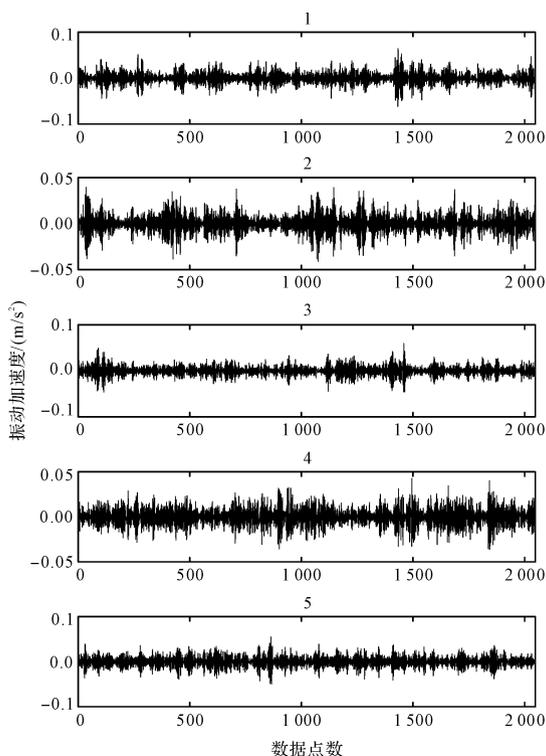


图 6 齿轮箱重构信号

Fig. 6 Gearbox reconstruction signal

从图 6 可以发现:经过信号重构后,部分的冲击成

分更加明显(这是因为剔除的 IMF 分量中包含许多噪声),可见,信号重构有利于突出核心的故障成分。

随后,笔者利用 IMWPE 提取齿轮箱重构信号的故障特征,得到 5 种状态下的 IMWPE 值,如图 7 所示。

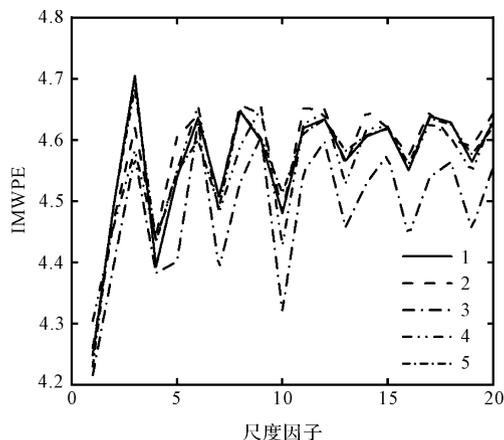


图 7 5 种状态下的 IMWPE 值

Fig. 7 IMWPE values in five states

从图 7 可知:5 种状态下的 IMWPE 值具有类似的趋势,即熵值随着尺度因子的变化而存在较显著的波动,且在不同尺度下熵值存在明显的交叉重叠,故不能直接将原始的故障特征输入至分类器进行识别。

为了提高特征的准确率,笔者对 IMWPE 特征进行优化,利用 LDA 对特征进行可视化降维,提取前 2 个故障特征进行分类识别。

5 种工况样本的 LDA 降维结果如图 8 所示。

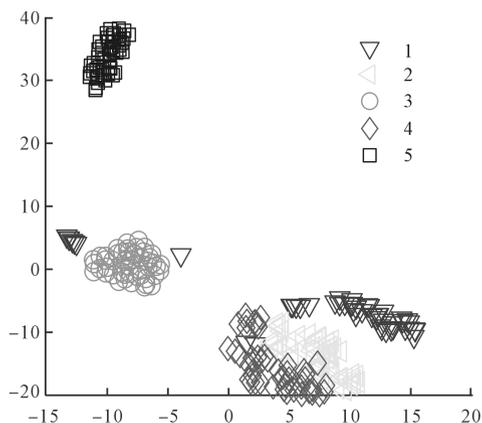


图 8 IMWPE 特征的可视化

Fig. 8 Visualization of IMWPE features

从图 8 可知:基于 ICEEMDAN-IMWPE 提取的故障特征具有比较好的质量,可视化的结果表明,样本 5 能够实现完全的分,而样本 4、样本 2、样本 1 发生了较明显的混叠,证明这 3 种样本可能会出现错误识别的样本。

笔者利用 BOA 方法,对支持向量机中的参数进行

优化,随后对测试样本进行识别。

ICEEMDAN-IMWPE-LDA-BOA-SVM 的故障识别结果,如图 9 所示。

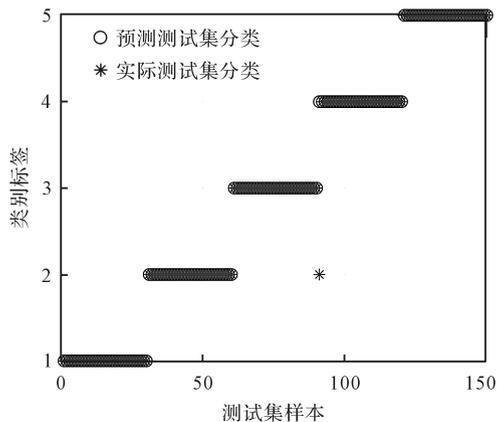


图 9 ICEEMDAN-IMWPE-LDA-BOA-SVM 的故障识别结果

Fig.9 Fault identification results of ICEEMDAN-IMWPE-BOA-SVM

从图 9 可知:ICEEMDAN-IMWPE-LDA-LDA-BOA-SVM 故障诊断方法仅错误识别了一个样本,即将样本 3 错误识别为样本 2,总的准确率为 99.33%。

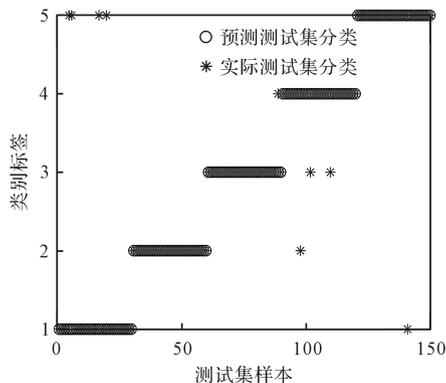
该结果证明,该方法能够较为有效地识别齿轮箱的故障类型,具有一定的应用潜力。

为了验证 ICEEMDAN-IMWPE-LDA-BOA-SVM 故障诊断方法相对于其他故障诊断方法的优劣,笔者从 3 个维度开展对比分析。

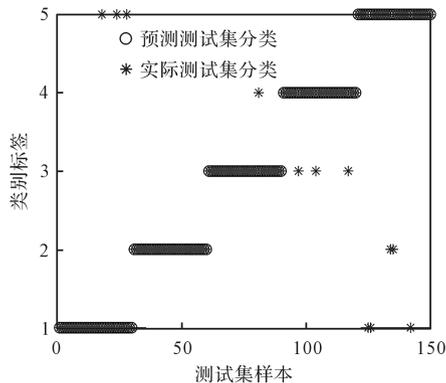
1) 特征提取维度对比

笔者将 ICEEMDAN-IMWPE-LDA 故障特征提取方法分别与 EEMD-IMWPE-LDA 故障特征提取方法、IMWPE-LDA 故障特征提取方法和 ICEEMDAN-IMWPE 故障特征提取方法进行比较,利用上述几种方法进行故障特征提取。

笔者将特征输入到 BOA-SVM 中进行故障识别,识别结果如图 10 所示。



(b) IMWPE-LDA 诊断结果



(c) ICEEMDAN-IMWPE 诊断结果

图 10 不同故障特征提取方法的诊断结果

Fig.10 Diagnostic results of different fault feature extraction methods

4 种方法的详细诊断结果如表 2 所示。

表 2 4 种特征提取方法的故障诊断结果

Table 2 Fault diagnosis results of four feature extraction methods

特征提取算法	诊断准确率/%
ICEEMDAN-IMWPE-LDA	99.33
EEMD-IMWPE-LDA	96
IMWPE-LDA	94
ICEEMDAN-IMWPE	92

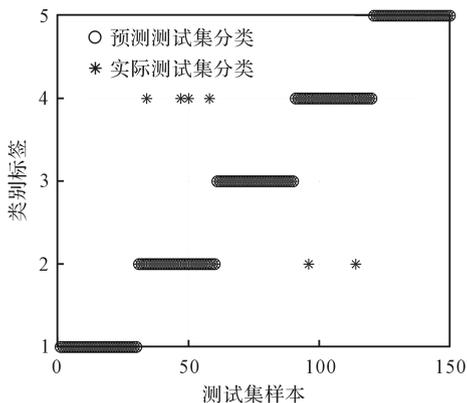
由表 2 可知:经 ICEEMDAN-IMWPE-LDA 故障特征识别后的准确率高高于 EEMD-IMWPE-LDA 特征提取方法,ICEEMDAN 算法在特征提取效果方面优于 EEMD 的原因是,EEMD 进行信号重构时无法完全去除信号中的噪声和干扰。

ICEEMDAN-IMWPE-LDA 和 EEMD-IMWPE-LDA 的准确率高高于 IMWPE-LDA,即对信号进行分解和重构的准确率高高于对原始信号进行分析的准确率,这是因为原始信号中的噪声会对分析结果造成极大影响。

而 ICEEMDAN-IMWPE-LDA 的结果高高于 ICEEMDAN-IMWPE 方法,这是因为 LDA 特征降维有效去除了冗余信息,突出了核心故障信息,从而提高了识别准确率。

2) 熵值方法对比

为了证明 IMWPE 方法在特征提取中的优越性,



(a) EEMD-IMWPE-LDA 诊断结果

笔者利用改进多尺度排列熵(improved multiscale permutation entropy, IMPE)、MPE、多尺度加权排列熵(MWPE)、基于方差的多尺度加权排列熵(multiscale weighted permutation entropy based on variance, MWPE\_rms)进行对比,并将采用上述 5 种方法提取的故障特征经过 LDA 降维,然后将其输入至 BOA-SVM 中,进行故障的识别。

笔者重复 10 次上述操作,得到了故障诊断结果,如表 3 所示。

表 3 故障诊断结果对比

Table 3 Comparison of fault diagnosis results

故障诊断方法	故障诊断准确率/%		
	最大	最小	平均
IMWPE	100	96.67	98.47
IMPE	98	93.33	95.73
MPE	85.33	76.67	81.47
MWPE	84.67	78.67	81.33
MWPE_rms	84.67	80	82.27

由表 3 可知:基于 IMWPE 的故障诊断方法实现了 98.47% 的平均识别准确率,高于其他 4 种诊断方法;与其他 4 方法相比,该方法多次实验的平均识别精度分别高了 2.74%、17%、17.14% 和 16.20%,证明了其具有较好的稳定性;同时,其最小准确率为 96.67%,表明其能够稳定可靠地识别齿轮箱的不同故障类型。

### 3) 分类模型对比

为了验证 BOA-SVM 的可行性,笔者采用常见的遗传算法优化支持向量机(genetic algorithm optimized support vector machine, GA\_SVM)、粒子群算法优化支持向量机(particle swarm optimization optimized support vector machine, PSO\_SVM)和差分算法优化支持向量机(differential evolution optimized support vector machine, DE\_SVM)进行故障诊断对比,得到的诊断结果如表 4 所示<sup>[20-22]</sup>。

表 4 故障诊断结果对比

Table 4 Comparison of fault diagnosis results

分类模型	模型耗时/s	诊断精度/%
BOA_SVM	5.31	99.33
PSO_SVM	10.26	97.33
GA_SVM	17.44	96.67
DE_SVM	9.06	98.67

由表 4 可知:相较于其他 3 种故障诊断模型,基于 BOA 优化的 SVM 取得了最高的故障诊断准确率和效率,诊断精度为 99.33%,用时 5.31 s,比 GA\_SVM 缩短了三分之二的时间,比 PSO\_SVM 缩短了近一半的时间,说明 BOA 优化得到的参数最优,同时耗时最少,在故障诊断中最具有优势。

笔者基于多个维度开展了对比实验,结果表明 ICEEMDAN-IMWPE-LDA-BOA-SVM 方法在效率和准确率方面均具有一定的优越性。

## 5 结束语

综上所述,笔者提出了一种基于 ICEEMDAN 信号重构、IMWPE 特征提取、LDA 特征降维和 BOA-SVM 故障识别的齿轮箱故障诊断方法,并利用东南大学齿轮箱数据集对其进行了实验,以验证该方法的有效性。

研究结论如下:

1) 采用 ICEEMDAN 对信号进行处理,结合相关系数分量筛选准则能够实现信号的降噪和重构。利用齿轮箱数据集对 ICEEMDAN-IMWPE-LDA-BOA-SVM 方法进行了验证,得到的故障识别准确率为 99.33%,证明了该方法的有效性;

2) 在提取故障特征方面,基于均方根的 IMWPE 特征提取方法优于 IMPE、MWPE、MPE 和 MWPE\_rms 方法,多次实验的平均识别精度也得到了不同程度的提高;

3) 采用 BOA 对 SVM 的惩罚系数和核函数参数进行了优化,减小了分类识别的不确定性,增加了分类精度,同时降低了模型的训练时间;与其他 3 种分类模型相比较,其效率最高,诊断时间只需 5.31 s,诊断精度为 99.33%,在故障诊断中更具比较优势。

虽然该方法取得了不错的损伤识别结果,但最低识别准确率为 96.67%,证明诊断性能存在提升空间。后续,笔者将进一步降低信号中的噪声,提升故障识别精度。

## 参考文献 (References):

- [1] 戚晓利,崔创创,杨艳,等.基于 RTSMFE、M-KRCDA 与 COA-SVM 的行星齿轮箱故障诊断[J].振动与冲击,2022,41(21):109-120.  
QI Xiao-li, CUI Chuang-chuang, YANG Yan, et al. Planetary gearbox fault diagnosis based on RTSMFE, M-KRCDA and COA-SVM[J]. Journal of Vibration and Shock, 2022, 41(21): 109-120.
- [2] 齐咏生,单成成,贾舜宇,等.一种基于增强型最大二阶循环平稳盲解卷积的齿轮箱复合故障诊断[J].中国机械工程,2022,33(24):2927-2941,2952.  
QI Yong-sheng, SHAN Cheng-cheng, JIA Shun-yu, et al. A gearbox composite fault diagnosis based on enhanced CYCBD [J]. China Mechanical Engineering, 2022, 33(24): 2927-2941, 2952.
- [3] 张玉兰,张宏伟,王新环.基于 CEEMDAN-TSMPE-PCA 风力发电机齿轮箱高速轴承故障诊断技术[J].噪声与振动控制,2022,42(4):158-164,234.  
ZHANG Yu-lan, ZHANG Hong-wei, WANG Xin-huan. fault diagnosis technology for wind turbine bearings based on CEEMDAN-TSMPE-PCA[J]. Noise and Vibration Control, 2022, 42(4): 158-164, 234.
- [4] 高淑芝,王拳,张义民.EEMD 熵特征和 t-SNE 相结合的滚动轴承故障诊断[J].机械设计与制造,2023,43

- (6):229-233.  
GAO Shu-Zhi, WANG Quan, ZHANG Yi-min. Rolling bearing fault diagnosis combining EEMD entropy feature and t-SNE[J]. Machinery Design & Manufacture, 2023, 43(6): 229-233.
- [5] 肖俊青,金江涛,李 春,等. 基于 CEEMDAN 模糊熵 CNN 轴承故障诊断研究[J]. 机械强度, 2023, 45(1): 26-33.  
XIAO Jun-qing, JIN Jiang-tao, LI Chun, et al. Research on bearing fault diagnosis based on CEEMDAN fuzzy entropy and conventional neural network[J]. Journal of Mechanical Strength, 2023, 45(1): 26-33.
- [6] 李 铭,何毅斌,马 东,等. 基于 ICEEMDAN-MPE 和 AO-LSSVM 的滚动轴承故障诊断[J]. 电子测量技术, 2022, 45(23): 66-71.  
LI Ming, HE Yi-bin, MA Dong, et al. Rolling bearing fault diagnosis based on ICEEMDAN-MPE and AO-LSSVM[J]. Electronic Measurement Technology, 2022, 45(23): 66-71.
- [7] LI Yong-bo, WANG Xian-zhi, LIU Zhen-bao, et al. The entropy algorithm and its variants in the fault diagnosis of rotating machinery: a review [J]. IEEE. Access, 2018, 6(11): 66723-66741.
- [8] 顾云青,苏玉香,沈晓群,等. 基于改进的 CEEMDAN 排列熵和 GWO-SVM 的滚动轴承故障诊断[J]. 组合机床与自动化加工技术, 2022, 62(8): 62-66.  
GU Yun-qing, SU Yu-xiang, SHEN Xiao-qun, et al. Rolling bearing fault diagnosis based on improved CEEMDAN permutation entropy and GWO-SVM [J]. Modular Machine Tool & Automatic Manufacturing Technique, 2022, 62(8): 62-66.
- [9] ZHOU Sheng-han, QIAN Si-lin, CHANG Wen-bing, et al. A novel bearing multi-fault diagnosis approach based on weighted permutation entropy and an improved SVM ensemble classifier[J]. Sensors, 2018, 18(6): 1934.
- [10] YUAN Xu-yi, FAN Yu-gang, ZHOU Cheng-jiang, et al. Research on twin extreme learning fault diagnosis method based on multi-scale weighted permutation entropy [J]. Entropy, 2022, 24(9): 1181.
- [11] 王振亚,姚立纲,蔡永武,等. 基于熵-流特征和樽海鞘群优化支持向量机的故障诊断方法[J]. 振动与冲击, 2021, 40(6): 107-114.  
WANG Zhen-ya, YAO Li-gang, CAI Yong-wu, et al. Fault diagnosis method based on the entropy-manifold feature and SSO-SVM[J]. Journal of Vibration and Shock, 2021, 40(6): 107-114.
- [12] 葛红平,刘晓波,黄朝晖,等. 粒化幅值感知排列熵和 WOA-SVM 的滚动轴承故障诊断[J]. 噪声与振动控制, 2021, 41(4): 101-108, 160.  
GE Hong-ping, LIU Xiao-Bo, HUANG Chao-qui, et al. Fault diagnosis method of rolling bearings based on granulation amplitude aware permutation entropy and WOA-SVM[J]. Noise and Vibration Control, 2021, 41(4): 101-108, 160.
- [13] 逢 英,高军伟. 基于 ICEEMDAN 能量矩和 MFOA-PNN 的轴承故障诊断[J]. 现代制造工程, 2022, 43(3): 122-126, 153.  
PANG Ying, GAO Jun-wei. Fault diagnosis of bearing based on ICEEMDAN energy moment and MFOA-PNN[J]. Modern Manufacturing Engineering, 2022, 43(3): 122-126, 153.
- [14] LIU Xiao-ming, SHU Ling. Fault detection of rotating machinery based on marine predator algorithm optimized resonance-based sparse signal decomposition and refined composite multiscale fluctuation dispersion entropy [J]. Review of Scientific Instruments, 2022, 93(11): 114703.
- [15] XUE Shao-hua, TAN Jian-ping, SHI Li-xiang, et al. Rope tension fault diagnosis in hoisting systems based on vibration signals using EEMD, improved permutation entropy, and PSO-SVM[J]. Entropy, 2020, 22(2): 209.
- [16] 李永健,宋 浩,刘吉华,等. 基于改进多尺度排列熵的列车轴箱轴承诊断方法研究[J]. 铁道学报, 2020, 42(1): 33-39.  
LI Yong-jian, SONG Hao, LIU Ji-hua, et al. A study on fault diagnosis method for train axle box bearing based on modified multiscale permutation entropy [J]. Journal of the China Railway Society, 2020, 42(1): 33-39.
- [17] 李慧梅,封会娟,张 坚,等. 基于 LDA 和 HMM 的变速箱故障诊断方法 [J]. 军事交通学院学报, 2020, 22(10): 33-37.  
LI Hui-mei, FENG Hui-juan, ZHANG Jian, et al. Fault diagnosis method for gearbox based on LDA and HMM [J]. Journal of Military Transportation University, 2020, 22(10): 33-37.
- [18] 王仲英,刘秋菊. 改进蝴蝶算法优化支持向量机的土壤含水量预测模型[J]. 计算机工程与设计, 2023, 44(2): 612-621.  
WANG Zhong-ying, LIU Qiu-ju. Soil water content prediction model based on improved butterfly algorithm optimizing support vector machine [J]. Computer Engineering and Design, 2023, 44(2): 612-621.
- [19] ZHENG Li-kang, HE Ye, CHEN Xiao-an. Research on a fault diagnosis method for rolling bearing based on improved multiscale range entropy and hierarchical prototype [J]. Measurement Science and Technology, 2021, 32(9): 095003.
- [20] 高坤明,张 衡,张振京,等. 基于时频分析和 SVM 的发动机齿轮箱故障诊断[J]. 机电工程技术, 2022, 51(3): 287-291.  
GAO Kun-ming, ZHANG Heng, ZHANG Zhen-jing, et al. Engine gearbox fault diagnosis based on time-frequency analysis and SVM [J]. Mechanical & Electrical Engineering Technology, 2022, 51(3): 287-291.
- [21] 董 营,赵 磊,田家彬,等. 风电齿轮箱塔上更换工装设计研究及应用[J]. 机械设计, 2022, 39(5): 116-121.  
DONG Ying, ZHAO Lei, TIAN Jia-bin, et al. Design and application of replacement tooling for wind power gearbox towers [J]. Journal of Machine Design, 2022, 39(5): 116-121.
- [22] 卢占标,黄细霞,姬 克,等. 基于多通道输入 CNN-SVM 的齿轮箱故障诊断[J]. 传感器与微系统, 2021, 40(6): 138-141.  
LU Zhan-biao, HUANG Xi-xia, JI Ke, et al. Fault diagnosis of gearbox based on multi-channel input CNN-SVM [J]. Transducer and Microsystem Technologies, 2021, 40(6): 138-141.

#### 本文引用格式:

王 洪,张锐丽,吴 凯. 基于 ICEEMDAN 和 IMWPE-LDA-BOA-SVM 的齿轮箱损伤识别模型[J]. 机电工程, 2023, 40(11): 1709 - 1717.

WANG Hong, ZHANG Ruili, WU Kai. Gear box damage identification model based on ICEEMDAN and IMWPE-LDA-BOA-SVM [J]. Journal of Mechanical & Electrical Engineering, 2023, 40(11): 1709 - 1717.