

结合振动特征优选和 GWOA-XGBoost 的电机轴承故障诊断*

于飞,樊清川,宣敏

(海军工程大学电气工程学院,湖北武汉 430033)

摘要:为解决电机轴承故障状态难以识别,从而造成诊断精度不高的情况,提出了一种基于信号特征提取与极限梯度提升算法(extreme gradient boosting, XGBoost)结合的电机轴承故障诊断模型。使用优化的变分模态分解获得振动信号的固有模态函数(intrinsic mode function, IMF)分量,再基于多尺度熵理论计算各 IMF 分量的多尺度熵值进行特征重构。在鲸鱼优化算法(whale optimization algorithm, WOA)中引入遗传算法的选择、交叉、变异操作对 WOA 进行改进。用改进的 WOA 算法对 XGBoost 的超参数进行寻优,获得了帮助 XGBoost 取得最优分类效果的超参数组合,将 7 种不同故障类型的振动信号进行重构后输入优化的 XGBoost 模型进行故障诊断。实验结果表明,所提 GWOA-XGBoost 模型的电机轴承故障诊断精度能够达到 97.14%,相较于传统诊断方法,性能提升效果显著。

关键词:电机轴承;故障诊断;变分模态分解;鲸鱼优化算法;极限梯度提升

中图分类号:TH133;TH165.3

开放科学(资源服务)标识码(OSID):

文献标志码:A 文章编号:1001-2486(2023)03-099-09



听语音
与作者互动
聊科研

Motor bearing fault diagnosis by combining vibration feature optimization and GWOA-XGBoost

YU Fei, FAN Qingchuan, XUAN Min

(College of Electrical Engineering, Naval University of Engineering, Wuhan 430033, China)

Abstract: In order to solve the problem that the fault state of motor bearings is difficult to identify, which results in low diagnosis accuracy, a motor bearing fault diagnosis model was proposed by combining the signal feature extraction and XGBoost (extreme gradient boosting algorithm). The optimized variational mode decomposition was used to obtain the IMF (intrinsic mode function) components of the vibration signal, and the multi scale entropy value of each IMF component was calculated on the basis of the multi-scale entropy theory for feature reconstruction. The WOA (whale optimization algorithm) was improved by introducing the selection, crossover and mutation operation of genetic algorithm. The improved WOA algorithm was used to optimize the hyperparameters of XGBoost, and obtain the hyperparameter combination that helped the XGBoost achieve the best classification effect. The vibration signals of 7 different types of faults were reconstructed and input into the optimized XGBoost model for fault diagnosis. Experimental results show that the motor bearing fault diagnosis accuracy of the proposed GWOA-XGBoost model can reach 97.14%, which significantly improves the performance compared with the traditional diagnosis method.

Keywords: motor bearings; fault diagnosis; variational modal decomposition; whale optimization algorithm; extreme gradient boosting

轴承在电机的运行中起着支撑和引导旋转的作用,据统计,电机故障有 40% 左右是由轴承故障引发^[1]。轴承是电机健康监测的重要对象,直接影响到电机运转的整体性能。因此,开展电机轴承的故障诊断研究意义重大。

目前基于振动信号的轴承故障诊断是最为常见的方法,当电机轴承出现故障时振动信号会出现明显的冲击波动,若对振动信号的时频信号进行有效的提取,就能为电机轴承的故障诊断提供有利的数据基础。文献[2]利用经验模态分解获

得振动信号的固有模态函数(intrinsic mode function, IMF),再提取 IMF 分量的奇异值以获得降维后的新特征向量,最后利用模糊神经网络有效地评估了电机轴承的运行状况。文献[3]提取了振动信号的能量特征,然后利用隐马尔可夫模型建立分类模型,能够有效区分不同类型故障的序列。文献[4]将卷积神经网络(convolutional neural networks, CNN)应用于电机轴承的故障诊断中,同时还使用了传统的快速傅里叶变换进行振动信号特征提取,但无法充分发挥 CNN 的特征

* 收稿日期:2021-07-29

基金项目:国家自然科学基金资助项目(51877212)

作者简介:于飞(1974—),男,辽宁大连人,副教授,博士,硕士生导师,E-mail:gfnu999@163.com

提取优势,从而造成诊断效果欠佳。文献[5]利用集成经验模态分解(ensemble empirical mode decomposition, EEMD)提取 IMF 分量,并通过 Hilbert 变换获得包络谱,最后利用优化过的深度信念网络(deep belief network, DBN)进行故障分类,结果较好,但在识别滚动体故障时准确率较低。文献[6]采用连续小波变换将振动信号转换为二维灰度图像,将 CNN 从灰度图像中提取的特征输入射频,有效诊断滚动轴承故障。文献[7]提出了一种广义多尺度动态时间弯曲算法,用于从风力发电机组齿轮箱振动信号中提取故障特征,采用拉普拉斯评分法选择敏感特征构造特征向量,采用随机森林进行故障状态分类。结果表明,该方法不仅能准确、高效地识别出不同的故障状态,而且与其他故障状态分类方法相比,准确性更高。

针对电机轴承振动信号,本文提出利用变分模态分解(variational mode decomposition, VMD)获得振动信号的 IMF 分量,并计算各 IMF 分量的多尺度熵重构特征向量,再利用改进的 WOA 算法对 XGBoost 模型进行超参数寻优,最后利用 GWOA-XGBoost 模型对重构的特征向量进行故障诊断,并通过实验案例证明所提模型的可行性。

1 基于优化 VMD 的特征提取基本原理

1.1 VMD 分解

VMD 是在维纳滤波、Hilbert 变换以及外差调解的基础上建立的一种新的非线性信号分解方法,避免了 EMD 在处理信号时可能出现端点效应或模态混叠的情况。VMD 算法预先确定模态分量个数 K ,保证它们的带宽之和最小,每个 IMF 分量被新定义成一个调幅-调频信号,表示如下:

$$u_k(t) = A_k(t) \cos(\phi_k(t)) \quad (1)$$

其中: $A_k(t)$ 为 $\cos(\phi_k(t))$ 的幅值; $\phi_k(t)$ 为 $u_k(t)$ 的相位, $\omega_k(t) = \phi_k'(t)$ 为 $u_k(t)$ 的瞬时频率。具体分解步骤如下:

1)利用 Hilbert 变换求出每个 IMF 分量的解析信号,其单边频谱表示如下:

$$\left(\delta(t) + \frac{j}{\pi t}\right)u_k(t) \quad (2)$$

2)估算 K 个 IMF 分量的中心频率,引入指数项 $e^{-j\omega_k t}$,调制各 IMF 分量的频谱到其所属基频带:

$$\left[\left(\delta(t) + \frac{j}{\pi t}\right)u_k(t)\right]e^{-j\omega_k t} \quad (3)$$

3)求出调节信号的梯度平方范数,估算每个

IMF 分量的带宽。最后基于约束条件,得到变分约束模型:

$$\begin{cases} \min_{\{u_k\}, \{\omega_k\}} \left\{ \sum_{k=1}^K \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) u_k(t) \right] e^{-j\omega_k t} \right\|_2^2 \right\} \\ \text{s. t. } \sum_k u_k(t) = x(t) \end{cases} \quad (4)$$

4)将二次罚参数 α 和 Lagrange 乘数 λ 引入,使变分约束模型转换成变分非约束模型:

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

基于交替方向乘子法迭代更新增广 Lagrange 表达式的“鞍点”,直至满足 $\sum_{k=1}^K (\|\hat{u}_k^{n+1}(\omega) - \hat{u}_k^n(\omega)\|_2^2 / \|\hat{u}_k^n(\omega)\|_2^2) < \varepsilon$ 时停止迭代,获得变分约束模型的最优解,即输出 K 个有限带宽的 IMF 分量^[8]。

1.2 VMD 的参数优化方法

在使用 VMD 进行分解时,分解层数 K 和惩罚因子 α 决定着信号处理的效果。当 K 值较小时,会存在分解不完全造成原始信号信息丢失的情况;当 K 值较大时,相邻 IMF 分量的频率中心间距较小,可能出现模态混叠的现象。惩罚因子 α 影响着分解过程的降噪性能以及细节保留度。本文利用粒子群优化算法(particle swarm optimization, PSO)获得最优 K 与 α 的组合。

局部极小熵的值反映了概率分布特性,可将其作为判断信号稀疏度的标准。引入局部极小熵作为 PSO 的适应度函数,迭代寻优步骤如下:

步骤 1:确定 PSO 的适应度函数并初始化其关键参数。

步骤 2:将待优化参数(K, α)作为个体粒子坐标,随机生成每个个体的初始坐标与初始速度。

步骤 3:每次迭代更新,使用 VMD 对每个个体进行分解,计算其对应的局部极小熵值,将所有个体进行比较,更新局部极值以及全局极值。

步骤 4:更新个体的速度和位置,跳至步骤 3,当迭代进行到预设最大次数或满足预设适应度值条件时停止,最优个体即为所需最佳 VMD 参数。

1.3 多尺度熵理论

基于样本熵概念 Costa 等^[9]提出了多尺度

熵(multi scale entropy, MSE)理论,从不同的尺度定义了时间序列的不规则变化,对原始信号进行粗粒化计算,得到样本序列在多个尺度下的样本熵。样本熵反映了信号在不同尺度下的复杂特性,进而可区分信号间的差异性,计算方法如下:

步骤1: 设一组数据 $X = \{X_1, X_2, X_3, \dots, X_n\}$, 对其进行粗粒化加工,目的在于改变原始数据的尺度,组成新的向量,即:

$$y_i^\tau = \frac{1}{\tau} \sum_{j=i}^{i+\tau-1} x_j, 1 \leq j \leq \frac{N}{\tau} \quad (6)$$

其中: τ 是尺度因子; y_i^τ 是生成的新粗粒化向量。当 $\tau > 1$ 时,生成的新向量的长度为 $\frac{N}{\tau}$ 。粗粒化就是

用宽度为 τ 的平滑窗不重叠滑动,求取窗内数列的平均值。

步骤2: 设置相似限度 r 以及嵌入维度 m , 粗粒化生成 m 维矢量:

$$Y(i) = \{Y_i, Y_{i+1}, Y_{i+2}, \dots, Y_{i+m-1}\} \quad (7)$$

其中, $i = 1, 2, 3, \dots, n - m$ 。

步骤3: $Y(i)$ 与 $Y(j)$ 两个样本之间最大距离的绝对值记为 $D(i, j)$:

$$D(i, j) = |Y(i) - Y(j)| = \max |X(i+k) - X(j+k)| \quad (8)$$

其中: $i, j = 1, 2, 3, \dots, n - m; k = 1, 2, 3, \dots, m - 1$ 。

步骤4: 记录 $D(i, j) < r$ 的个数 H , 并计算 H 与距离总数 $N - m - 1$ 的比值, 记作 $B_i^m(r)$, 其平均值记为 $B^m(r)$:

$$B^m(r) = \frac{1}{N - m} \sum_{i=1}^{N-m} B_i^m(r) \quad (9)$$

步骤5: 若重构数据有 $m + 1$ 个时, 重复步骤 1~4 得到 $B^{m+1}(r)$ 。则数列 $X = \{X_1, X_2, X_3, \dots, X_n\}$ 的样本熵为:

$$S(m, r) = \lim_{N \rightarrow \infty} \left(-\ln \frac{B^m(r)}{B^{m+1}(r)} \right) \quad (10)$$

当 N 为有限数时, 样本熵表示为:

$$S(m, r, N) = -\ln \frac{B^m(r)}{B^{m+1}(r)} \quad (11)$$

步骤6: 不同尺度 τ 的样本熵共同表示多尺度熵:

$$M(\tau | S(m, r, N)) = \left\{ \tau \left| -\ln \frac{B^m(r)}{B^{m+1}(r)} \right. \right\} \quad (12)$$

计算多尺度熵时需预先设置嵌入维度 m 、相似限度 r 以及尺度因子 τ , 基于经验设置 $m = 2$, r 为 $0.1\delta \sim 0.25\delta$ (δ 表示数据序列标准差), $\tau \leq 20^{[10]}$ 。本文选择 $m = 2, r = 0.2\delta, \tau = 15$ 。

2 基于优化 XGBoost 的故障诊断

2.1 XGBoost 模型介绍

XGBoost 模型是基于树模型集成的一种新型机器学习算法, 它主要是将多个分类精度不高的分类回归树(classification and regression tree, CART)模型结合, 搭建出一个精度较高的模型。模型精度迭代增加, 每一轮迭代使用一个新的 CART 树模型去拟合前一个树模型的残差, 该过程叫作梯度提升。XGBoost 模型构建如下。

对于某一数据集 $D = \{(x_i, y_i)\} (x_i \in \mathbf{R}^m, y_i \in \mathbf{R})$ 包含 n 个样本、 m 维特征, K 棵树模型最终输出为:

$$\hat{y}_i = \sum_{k=1}^K f_k(x_i), f_k \in F \quad (13)$$

$$F = \{f(x) = \omega_{q(x)}\}, q: \mathbf{R}^m \rightarrow T, \omega \in \mathbf{R}^T \quad (14)$$

其中: 每个 f_k 对应着一个单独的树模型; q 表示一个从样本指向对应叶子标签的结构; ω 表示树模型的权重, 即第 i 个节点的权重为 ω_i ; T 表示叶子节点的数量; F 表示树模型组成的集合; $\omega_{q(x)}$ 表示对样本 x 的打分, 即模型输出值。

XGBoost 的目标函数正则化可以表示为:

$$L(\phi) = \sum_i (\hat{y}_i, y_i) + \sum_k \Omega(f_k) \quad (15)$$

$$\Omega(f) = \gamma T + \frac{1}{2} \lambda \omega^2 \quad (16)$$

其中, 式(15)是可微的损失函数, 表示输出值和设置标签的差值; $\sum_k \Omega(f_k)$ 表示正则项, 用来惩罚模型的复杂性, 对叶子节点进行平滑处理, 防止其过拟合; γ, λ 是模型的自定义参数, 数值越小, 模型约束越容易出现过拟合现象。当 γ, λ 为 0 时, 目标函数就退化为梯度提升决策树了。传统方法无法求解式(15), 故利用贪婪学习方法使目标函数 f_k 最小: 第 i 个树模型第 t 次迭代的输出值记为 \hat{y}_i , 每一轮迭代引入一个最优新函数 f_t , 即利用加性模型以及向前分步算法, 最大可能减小目标函数。

2.2 改进的 WOA 算法

WOA^[11]是在 2016 年提出的一种元启发式群体智能优化算法, 该算法是受鲸鱼的气泡网摄食行为启发而来的。如图 1(a)所示, 鲸鱼在猎物的正下方, 沿着螺旋状路径, 使用收缩包围的方法逐步逼近猎物, 在逼近的同时, 不断向上喷吐气泡, 猎物无法游过气泡网, 与此同时鲸鱼加速上升, 吃掉猎物, 完成狩猎。WOA 的参数优化如图 1(b)所示, 包含三个机制: 环绕包围、

螺旋更新和全局搜索。

$$a = \left(1 - \frac{t}{t_{\max}}\right) \quad (20)$$

其中, t_{\max} 是最大迭代次数。

2.2.2 螺旋更新

螺旋更新机制的数学模型为:

$$X(t+1) = X_p(t) + D' \cdot e^{bl} \cdot \cos(2\pi l) \quad (21)$$

$$D' = |X_p(t) - X(t)| \quad (22)$$

其中: D' 表示搜索目标和搜索个体之间的距离; b 是形状参数用来辅助定义图 1 (b) 中的对数螺旋线, l 是 $[-1, 1]$ 之间的随机数。

但要注意, 捕猎过程中, 鲸鱼收缩环形路径的过程与鲸鱼沿环形路径前进的过程是同步进行的, 因此, 在优化时, 群体进行环绕包围和螺旋更新的概率 p 是相同的, 均为 0.5, 如下所示:

$$X(t+1) = \begin{cases} X_p(t) - A \cdot |B \cdot X_p(t) - X(t)|, & p \leq 0.5 \\ X_p(t) + D' \cdot e^{bl} \cos(2\pi l), & p > 0.5 \end{cases} \quad (23)$$

其中, p 是 $[0, 1]$ 之间随机生成的一个数。

2.2.3 全局搜索

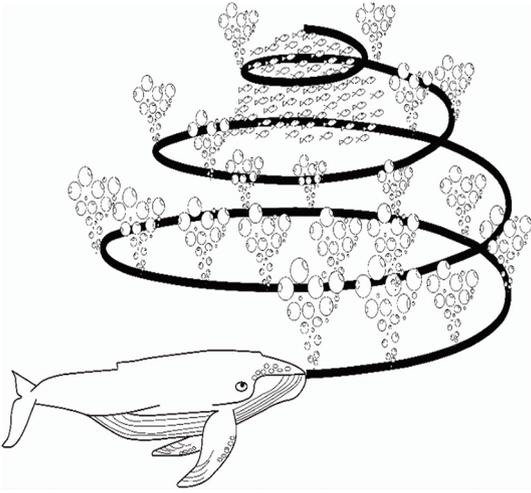
在该阶段, 对鲸鱼搜寻猎物的行为进行建模。搜索个体的位置是根据种群中随机选取的一个搜索个体进行更新, 而非当前的最优个体。该阶段其他个体将远离随机选取的搜索个体, 以执行全局搜索:

$$X(t+1) = X_{\text{rand}}(t) - A \cdot |B \cdot X_{\text{rand}}(t) - X(t)| \quad (24)$$

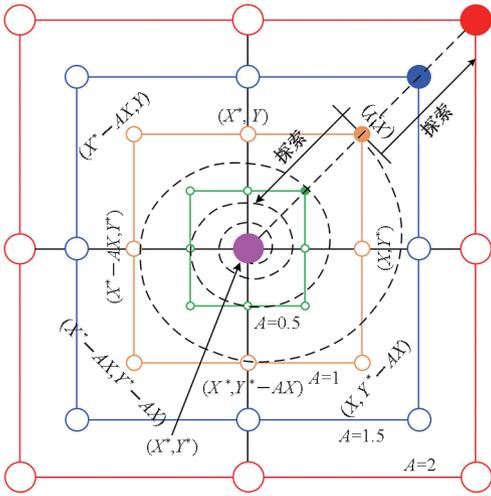
其中, $X_{\text{rand}}(t)$ 为种群中随机选取的寻优个体位置。

WOA 往往存在过早陷入局部最优解的现象, 于是在 WOA 中引入 GA 的选择、交叉、变异操作 (获得 GWOA 算法), 能够使得个体跳出局部最优解而获得全局最优解, 其基本模型参见文献 [13]。

在进行选择操作时常用“轮盘赌法”, 该操作能够增加种群的平均适应度, 但难以产生更优的个体, 会导致种群多样性下降, 造成部分有意义的点丢失, 所以在进行选择时增加一定的选择比例既可提高平均适应度, 又可保证种群基因的多样性。对于选择操作, 在整个迭代过程的早期, 使用较大的交叉概率尽可能获得更优的个体; 在算法迭代的后期, 个体间的差异较小, 为避免优秀个体被破坏, 需降低交叉概率。同理, 当进行变异操作时, 前期使用较小的变异概率促进种群前期的进化, 后期增加变异概率, 避免陷入局部最优。改进的 WOA 算法流程如



(a) 捕食行为
(a) Predation



(b) 参数优化机制
(b) Parameter optimization mechanism

图 1 鲸鱼优化模型

Fig. 1 Whale optimization model

2.2.1 环绕收缩

在该阶段, 设定种群中的最优单位作为目标, 其他个体更新自己的位置并向目标猎物逼近, 其数学模型为:

$$X(t+1) = X_p(t) - A \cdot D \quad (17)$$

$$D = |B \cdot X_p(t) - X(t)| \quad (18)$$

其中: $X(t)$ 是鲸鱼个体的位置向量; t 为当前的迭代次数; $X_p(t)$ 是猎物的位置向量^[12]; A 和 B 表示系数向量, 表达式如式 (19) 所示。

$$\begin{cases} A = 2a \cdot r_1 - a \\ B = 2r_2 \end{cases} \quad (19)$$

其中: r_1 和 r_2 是 $[0, 1]$ 内的随机向量; a 表示调控参数, 随着迭代的进行, a 从 2 递减至 0, 即

图 2 所示。借用 Srinivas 等^[14]所提自适应遗传算法里的思想得到的自适应交叉概率和变异概率的计算如下:

$$p_m = \begin{cases} 0.85 - \frac{(0.85 - 0.55)(f' - f_{\text{mean}})}{f_{\text{max}} - f_{\text{mean}}} & f' \geq f_{\text{mean}} \\ 0.85 & f' < f_{\text{mean}} \end{cases} \quad (25)$$

$$p_n = \begin{cases} 0.1 - \frac{(0.1 - 0.01)(f_{\text{max}} - f)}{f_{\text{max}} - f_{\text{mean}}} & f \geq f_{\text{mean}} \\ 0.1 & f < f_{\text{mean}} \end{cases} \quad (26)$$

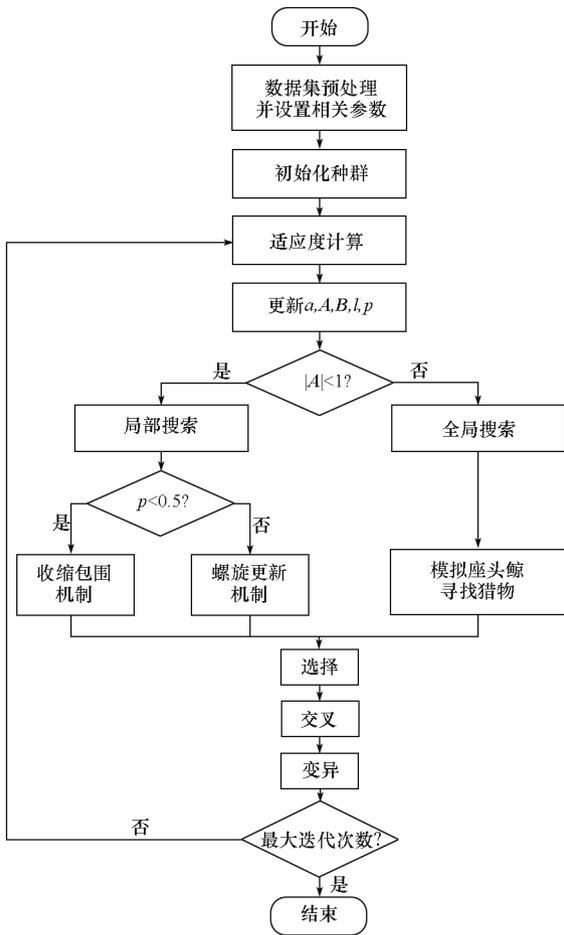


图 2 改进 WOA 算法流程

Fig. 2 Improved WOA algorithm flow

2.3 GWOA 优化的 XGBoost 模型

利用 XGBoost 模型进行分类训练时,迭代次数 n 过少存在欠拟合现象,导致模型无法充分训练,当 n 过大,则存在过拟合现象,导致模型对数据训练过于极限而泛化能力下降;学习率 η 过小时,模型训练过慢,过大时可能忽略全局最优解;决策树最大深度 d_{max} 、随机样本比 $r_{\text{subsample}}$ 、特征占比 $r_{\text{colsample}}$ 以及决策树分裂点指标 γ_{split} 等共同影响模型的分类能力和预测精度。利用改进的 WOA

算法对这 6 个参数进行迭代寻优,以增强模型的故障诊断性能。基于 GWOA 优化的 XGBoost 电机轴承故障诊断流程如图 3 所示。

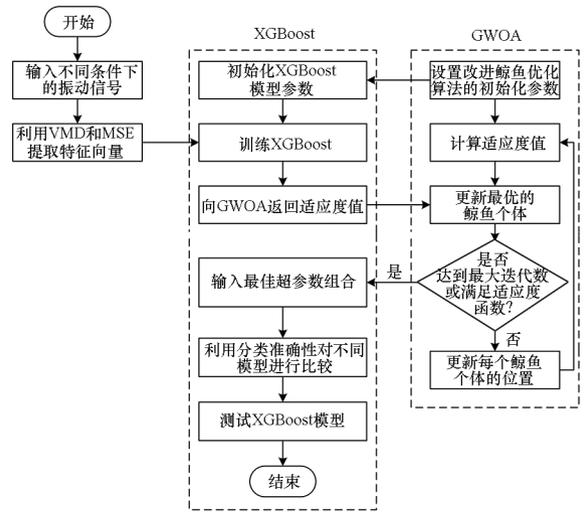


图 3 电机轴承故障诊断流程

Fig. 3 Motor bearing fault diagnosis process

3 算例分析

本文所使用的实验数据来源于美国凯斯西储大学^[15],数据集包含电机轴承的内圈、外圈、滚动体的轻微故障以及严重故障再加上正常运行状态共七类数据。样本运行状态编码如表 1 所示。

表 1 电机轴承故障编码

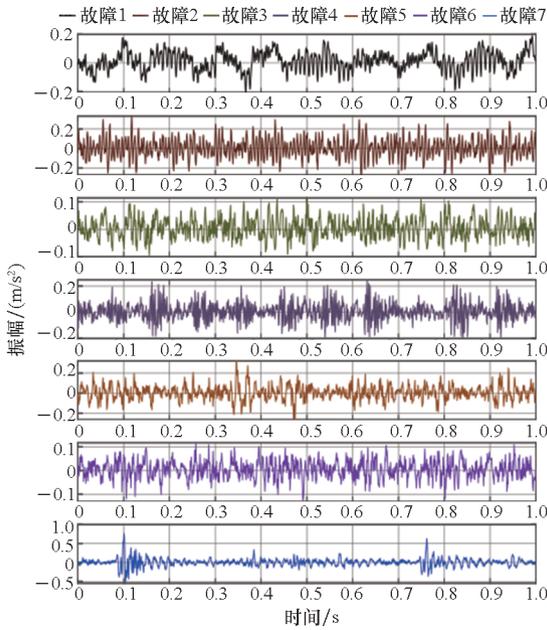
Tab. 1 Motor bearing fault code

故障类别	电机轴承运行状态	状态编码
1	正常状态	[1,0,0,0,0,0,0]
2	内圈轻微裂纹	[0,1,0,0,0,0,0]
3	滚动体轻微裂纹	[0,0,1,0,0,0,0]
4	外圈轻微裂纹	[0,0,0,1,0,0,0]
5	内圈严重故障	[0,0,0,0,1,0,0]
6	滚动体严重故障	[0,0,0,0,0,1,0]
7	外圈严重故障	[0,0,0,0,0,0,1]

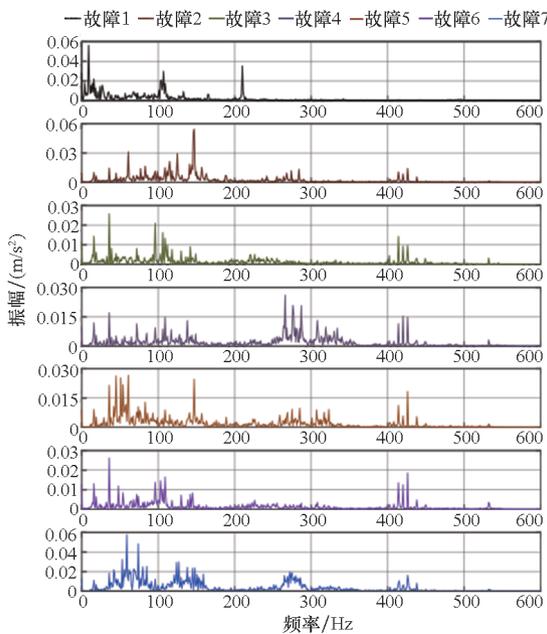
选择电机运行工况为 735 W、转速 1 772 r/min、采样频率 12 kHz 的电机轴承故障数据集,将数据集切分成长度为 1 200 个点的 700 组数据来验证所提模型的可行性,不同故障状态的时频波形如图 4 所示。

3.1 基于优化 VMD 分解的特征提取

利用 PSO 算法训练 VMD 以获得最优超参数组合,将局部极小熵值作为适应度,适应度变化曲线如图 5 所示。



(a) 时域
(a) Time domain



(b) 频域
(b) Frequency domain

图 4 不同故障状态时频波形

Fig. 4 Time-frequency waveforms for different fault states

由图 5 看出,在 8 次迭代后适应度值基本不变,进行 10 次计算取其平均值,获得最优超参数组合 $(K, \alpha) = (4, 134.7)$,此时的局部最小包络熵值最小。使用优化后的 VMD 对轴承故障数据进行分解,由于篇幅有限,此处仅列出外圈严重故障分解的时频波形,如图 6 所示。

正交性指数(index of orthogonality, IO)可以用来表征各 IMF 分量的整体正交性,IO 值越小,则各 IMF 分量的正交性越强;能量保存度(index

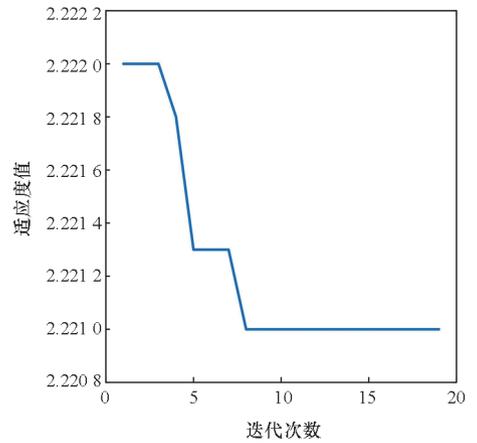


图 5 适应度变化曲线

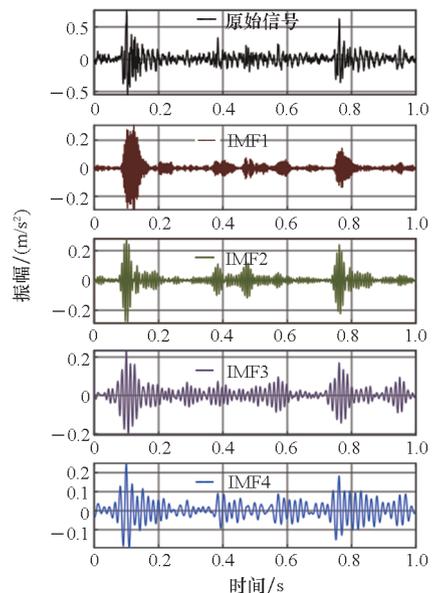
Fig. 5 Fitness curve

of energy conservation, IEC) 表示分解前后能量对比量,IEC 的值与 1 越接近,则分解时保留下的能量越多,分解稳定性越好^[16]。优化前后 VMD 分解效果对比见表 2,可以看出优化后的 VMD 分解正交性与稳定性都更优异。

使用优化后的 VMD 对电机轴承振动信号分解后获得多个 IMF 分量,由于 MSE 能够反映信号的复杂性,于是再进行 MSE 计算重构特征向量。

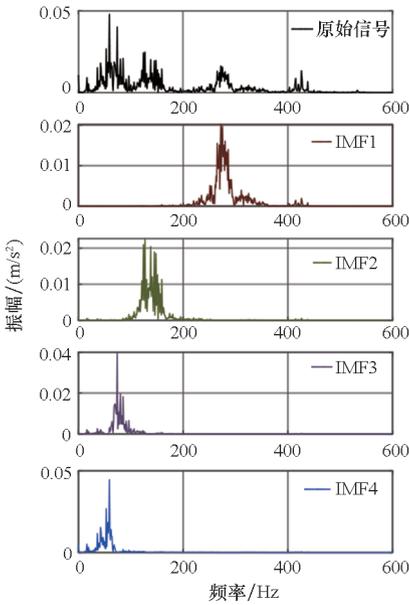
3.2 基于 GWOA-XGBoost 模型的故障诊断

将处理好的 700 组数据随机抽取 90% 作为训练集,剩下 10% 作为测试集。利用改进 WOA 算法优化 XGBoost 参数,设置 WOA 寻优个体数量为 20,维数为 7,最大迭代次数为 100,将训练集数据输入 GWOA-XGBoost 模型训练,得最优超



(a) 时域

(a) Time domain



(b) 频域

(b) Frequency domain

图 6 各 IMF 时频波形

Fig. 6 Time-frequency waveform of each IMF

表 2 VMD 优化前后效果对比

Tab. 2 Effect comparison before and after VMD optimization

参数	VMD	优化 VMD
IO	0.052 3	0.035 8
IEC	0.874 7	0.925 8

参数如表 3 所示。

表 3 最优超参数

Tab. 3 Optimal super parameter

超参数	最优取值
n	810
η	0.57
d_{max}	7
$r_{subsample}$	0.62
$r_{colsample}$	0.72
γ_{split}	0.31

将训练好的 XGBoost 分类器应用于测试集,故障诊断结果如图 7 所示。为验证 GWOA-XGBoost 模型对电机轴承故障诊断的提升效果,分别利用 WOA-XGBoost、PSO-XGBoost、GA-XGBoost、XGBoost 对相同的训练集和测试集进行故障诊断,故障诊断结果如图 8 ~ 11 所示。可以看出,未经优化的 XGBoost 模型的故障诊断

效果是很差的,而 GWOA-XGBoost 模型的故障诊断效果极好,仅在滚动体严重故障中出现了 2 处错误诊断,诊断准确率高达 97.14%,完全满足工程需求。出现误诊的原因是采集振动信号时,传感器离内圈和外圈更近,能够更清晰地采集到内外圈的振动特性,同时滚动体在运转时除了随着内外圈进行公转还存在自传现象,这也为振动信号的采集造成了干扰,因此滚动体的故障诊断更容易出现误诊现象。

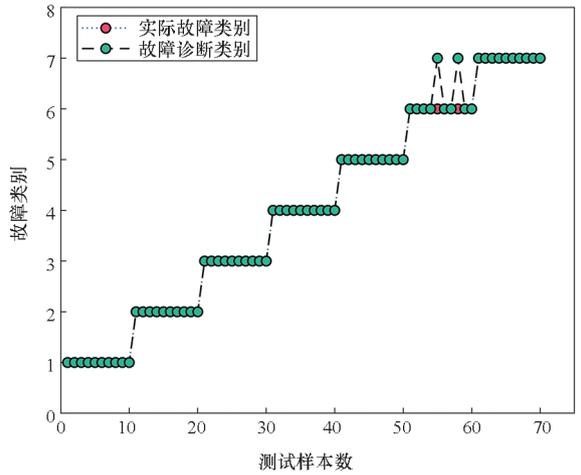


图 7 GWOA-XGBoost 故障诊断结果

Fig. 7 GWOA-XGBoost fault diagnosis results

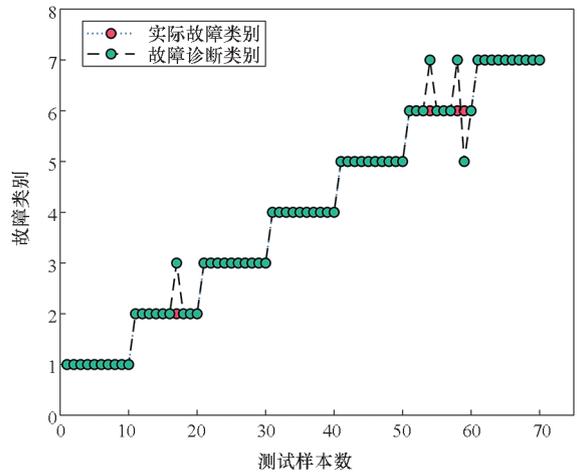


图 8 WOA-XGBoost 故障诊断结果

Fig. 8 WOA-XGBoost fault diagnosis results

为了进一步说明本文建立的电机轴承故障诊断模型的可靠性与准确性,将所建模型与常用的故障诊断模型如 BP 神经网络、支持向量机 (support vector machine, SVM)、极限学习机^[17-19] (extreme learning machine, ELM) 进行对比,结果如表 4 所示。从表中可以看出,本文建立的 GWOA-XGBoost 模型的故障诊断准确率均高于其他算法,准确率高达 97.14%。

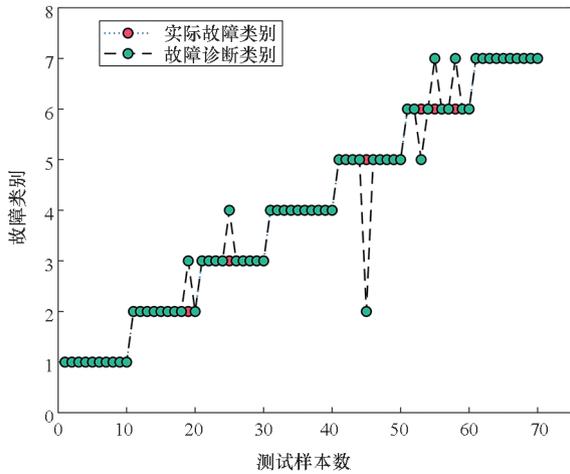


图 9 PSO-XGBoost 故障诊断结果

Fig. 9 PSO-XGBoost fault diagnosis results

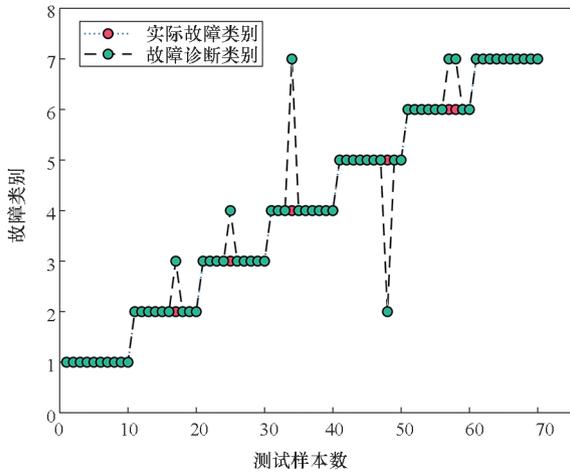


图 10 GA-XGBoost 故障诊断结构

Fig. 10 GA-XGBoost fault diagnosis results

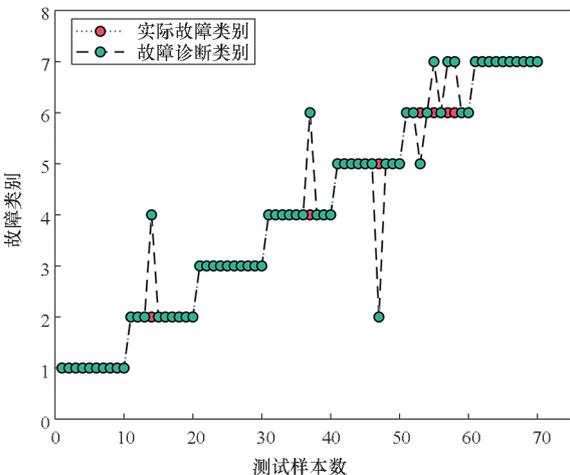


图 11 XGBoost 故障诊断结构

Fig. 11 XGBoost fault diagnosis results

4 结论

针对电机轴承故障的多样性和模糊性造成诊断精度较低的问题,建立了一种基于 XGBoost 的

表 4 不同模型的故障诊断准确率

Tab. 4 Fault diagnosis accuracy of different models

故障诊断模型	误诊数	准确率/%
BP	9	87.14
SVM	13	81.43
ELM	6	91.43
XGBoost	7	90.00
GWOA-XGBoost	2	97.14

电机轴承故障诊断模型,结合算例分析,所得结论如下:

1) 针对轴承振动信号的非线性和非平稳性特点,利用优化的 VMD 算法对振动信号进行预处理,结合多尺度熵理论提高信号特征提取的能力,有助于深度挖掘振动信号与故障状态的内在联系。

2) 在 WOA 算法中引入遗传算法的选择、交叉、变异操作,有效地帮助 WOA 跳出局部最优解,再利用改进后的 WOA 算法优化 XGBoost 模型。所提模型与传统算法相比在电机轴承的故障诊断中具有更高的诊断精度。

参考文献 (References)

[1] 安国庆,秦程,郭立炜,等. 峭度滤波器用于电机轴承早期故障特征提取[J]. 电机与控制学报, 2014, 18(6): 55-60.
AN G Q, QIN C, GUO L W, et al. Kurtosis filter in feature extraction for incipient motor bearing fault diagnosis [J]. Electric Machines and Control, 2014, 18(6): 55-60. (in Chinese)

[2] GAI J B, HU Y F, SHEN J X. A bearing performance degradation modeling method based on EMD-SVD and fuzzy neural network [J]. Shock and Vibration, 2019, 2019: 5738465.

[3] 王国锋,李玉波,秦旭达,等. 基于 TVAR-HMM 的滚动轴承故障诊断[J]. 天津大学学报, 2010, 43(2): 168-173.
WANG G F, LI Y B, QIN X D, et al. Fault diagnosis of rolling bearing based on TVAR and HMM [J]. Journal of Tianjin University, 2010, 43(2): 168-173. (in Chinese)

[4] 李恒,张氢,秦仙蓉,等. 基于短时傅里叶变换和卷积神经网络的轴承故障诊断方法[J]. 振动与冲击, 2018, 37(19): 124-131.
LI H, ZHANG Q, QIN X R, et al. Fault diagnosis method for rolling bearings based on short-time Fourier transform and convolution neural network [J]. Journal of Vibration and Shock, 2018, 37(19): 124-131. (in Chinese)

[5] 王玉静,那晓栋,康守强,等. 基于 EEMD-Hilbert 包络谱和 DBN 的变负载下滚动轴承状态识别方法[J]. 中国电机工程学报, 2017, 37(23): 6943-6950, 7085.
WANG Y J, NA X D, KANG S Q, et al. State recognition method of a rolling bearing based on EEMD-Hilbert envelope spectrum and DBN under variable load [J]. Proceedings of

- the CSEE, 2017, 37 (23): 6943 - 6950, 7085. (in Chinese)
- [6] XU G W, LIU M, JIANG Z F, et al. Bearing fault diagnosis method based on deep convolutional neural network and random forest ensemble learning [J]. *Sensors (Basel, Switzerland)*, 2019, 19(5): 1088.
- [7] PANG B, TIAN T, TANG G J. Fault state recognition of wind turbine gearbox based on generalized multi-scale dynamic time warping [J]. *Structural Health Monitoring*, 2021, 20(6): 3007 - 3023.
- [8] 江志农, 魏东海, 张进杰, 等. 基于 VMD 和 SVD 的柴油机气门间隙异常特征提取研究[J]. *振动与冲击*, 2020, 39(16): 23 - 30.
- JIANG Z N, WEI D H, ZHANG J J, et al. A study on valve clearance anomaly feature extraction of diesel engines based on VMD and SVD [J]. *Journal of Vibration and Shock*, 2020, 39(16): 23 - 30. (in Chinese)
- [9] COSTA M, GOLDBERGER A L, PENG C K. Multiscale entropy analysis of complex physiologic time series [J]. *Physical Review Letters*, 2002, 89(6): 068102.
- [10] 何葵东, 陈伽, 金艳, 等. EEMD 多尺度熵和 ELM 在水电机组振动信号特征提取中的应用[J]. *中国农村水利水电*, 2021(5): 176 - 182, 187.
- HE K D, CHEN J, JIN Y, et al. Application of EEMD multi-scale entropy and ELM in feature extraction of vibration signal of hydropower unit [J]. *China Rural Water and Hydropower*, 2021(5): 176 - 182, 187. (in Chinese)
- [11] MIRJALILI S, LEWIS A. The whale optimization algorithm [J]. *Advances in Engineering Software*, 2016, 95: 51 - 67.
- [12] 朱志青. 改进鲸鱼算法及其在认知无线电频谱分配中的应用[D]. 重庆: 重庆大学, 2018.
- ZHU Z Q. A modified whale optimization algorithm and its application in cognitive radio spectrum allocation [D]. Chongqing: Chongqing University, 2018. (in Chinese)
- [13] 黄新波, 王玉鑫, 朱永灿, 等. 基于遗传算法与模糊逻辑融合的线路覆冰预测[J]. *高电压技术*, 2016, 42(4): 1228 - 1235.
- HUANG X B, WANG Y X, ZHU Y C, et al. Icing forecast of transmission line based on genetic algorithm and fuzzy logic [J]. *High Voltage Engineering*, 2016, 42(4): 1228 - 1235. (in Chinese)
- [14] SRINIVAS M, PATNAIK L M. Adaptive probabilities of crossover and mutation in genetic algorithms [J]. *IEEE Transactions on Systems, Man, and Cybernetics*, 1994, 24(4): 656 - 667.
- [15] Case Western Reserve University. Rolling bearing vibration signal dataset [DS/OL]. [2021 - 06 - 01]. <https://engineering.case.edu/bearingdatacenter/download-data-file>.
- [16] 葛兴来, 张鑫. 采用奇异能量谱与改进 ELM 的轴承故障诊断方法[J]. *电机与控制学报*, 2021, 25(5): 80 - 87.
- GE X L, ZHANG X. Bearing fault diagnosis method using singular energy spectrum and improved ELM [J]. *Electric Machines and Control*, 2021, 25(5): 80 - 87. (in Chinese)
- [17] 皮骏, 刘鹏, 马圣, 等. 基于 MGA-BP 网络的航空轴承故障诊断[J]. *振动·测试与诊断*, 2020, 40(2): 381 - 388, 423.
- PI J, LIU P, MA S, et al. Aero-engine bearing fault diagnosis based on MGA-BP neural network [J]. *Journal of Vibration, Measurement & Diagnosis*, 2020, 40(2): 381 - 388, 423. (in Chinese)
- [18] 吴春光, 王建朝, 化麒. 基于 NSP 和 SVM 的滚动轴承故障诊断方法[J]. *轴承*, 2016(12): 39 - 42, 55.
- WU C G, WANG J C, HUA Q. Fault diagnosis of rolling bearings based on null-space pursuit algorithm and support vector machine [J]. *Bearing*, 2016(12): 39 - 42, 55. (in Chinese)
- [19] 秦波, 孙国栋, 王建国. 基于 VMD 与分层极限学习机的滚动轴承故障诊断方法研究[J]. *组合机床与自动化加工技术*, 2017(4): 107 - 110.
- QIN B, SUN G D, WANG J G. Research on rolling bearing fault diagnosis method based on VMD and hierarchical extreme learning machine [J]. *Modular Machine Tool & Automatic Manufacturing Technique*, 2017(4): 107 - 110. (in Chinese)