

基于 FCBF-IPSO 的旋转机械故障诊断方法^{*}

尹海涛¹, 赵荣珍^{1,2}, 马 驰^{1,3}, 邓林峰¹

(1. 兰州理工大学机电工程学院 兰州, 730050)

(2. 广州理工学院智能制造与电气工程学院 广州, 510540)

(3. 酒泉职业技术学院机电工程学院 酒泉, 735000)

摘要 针对旋转机械高维故障数据集存在着冗余特征导致分类困难和故障识别率偏低的问题,提出一种将快速相关过滤(fast correlation-based filter, 简称 FCBF)算法和改进粒子群优化(improved particle swarm optimization, 简称 IPSO)算法相结合的故障敏感特征选择方法。首先,利用 FCBF 算法和数据集值域状况初步筛选特征,剔除与类别信息不相关的特征和冗余特征;其次,使用 IPSO 算法对筛选后的特征子集进行二次筛选,进一步剔除其中的冗余特征,得到有利于分类运算的低维敏感特征子集;最后,通过转子故障模拟数据集进行了实验验证。结果表明:该方法可有效剔除故障数据集中的不相关特征和冗余特征,利用 IPSO 算法对支持向量机(support vector machine, 简称 SVM)参数 C 和 σ 进行的优化,达到了显著提高分类器识别精度和运行效率的效果。本研究方法为降低旋转机械故障数据资源的规模提供了一种敏感特征筛选策略,并丰富了特征选择的基础理论。

关键词 故障诊断;特征选择;快速相关过滤算法;粒子群优化算法

中图分类号 TP18;TH165

引 言

流程工业的旋转机械设备结构复杂且运行周期长,为确保此类设备健康平稳、优质高效运行,对其进行故障诊断和寿命预测至关重要^[1]。机械故障诊断的核心在于模式识别,故障特征提取技术是影响模式识别决策质量的关键^[2]。为尽可能全面表征设备状态信息,通常从时域、频域及时频域等角度提取特征并构建故障特征数据集^[3]。在这种多域的高维非线性特征数据集中,包含的大量冗余特征和噪声会造成“维数灾难”问题^[4]。因此,设计有效的降维方法以获取表征故障本质的敏感特征子集,是克服“维数灾难”从而提升故障辨识质量的关键对策之一。

特征选择是维数约简即降维的关键技术^[5-6]。目前,特征选择算法基本上可分为过滤式(Filter)和封装式(Wrapper)2 大类^[7]。其中,过滤式算法是根据特征与类别间的相关性,通过排序筛选出敏感特征。典型的如 FCBF 算法^[8]通过引入对称不确定性指标(symmetrical uncertainty, 简称 SU)来评估这种相关性,达到同时去除不相关和冗余特征的效果。具有较高分类准确率的封装式算法是利用分类结果

评估特征子集的优劣,例如,将粒子群优化算法(particle swarm optimization, 简称 PSO)与 SVM 封装后进行的特征选择。但是,上述 2 类算法仍存在一些缺陷。过滤式算法虽然运算速率快,但其特征子集的分类准确率偏低;封装式算法的计算量偏大。为此,将 2 类算法结合以进行更有效的特征选择尝试引起了关注。Zhou 等^[9]提出的一种两级粒子协同特征选择方法,通过二进制粒子群优化两级粒子的合作策略有效减少了特征数量。肖艳等^[10]采用 PSO-SVM 算法进行特征选择,获得了较优的特征子集。薛瑞等^[11]利用 ReliefF 算法快速筛选出重要特征,并通过量子粒子群算法(quantum particle swarm optimization, 简称 QPSO)寻找最优特征子集,实现了提高分类精度的效果。上述方法均是借助 PSO 算法的优良搜索能力,提高了模型的分类精度。但是,PSO 算法目前还存在易陷入局部最优解、全局与局部搜索不平衡以及迭代后期收敛速度慢等不足。

针对去除故障数据集的冗余特征以降低分类运算难度、提高故障辨识精度的需求,本研究将过滤式和封装式 2 类算法相结合,提出了一种具有两级筛

^{*} 国家自然科学基金资助项目(51675253, 62241308)

收稿日期:2023-07-01;修回日期:2023-10-23

选策略的FCBF-IPSO故障敏感特征选择方法。该方法首先利用FCBF算法剔除数据集的不相关特征和冗余特征,然后通过IPSO算法去除残留的冗余特征,得到低维敏感特征子集。本研究可为获得高质量旋转机械故障数据集提供理论参考依据。

1 相关原理简介

1.1 FCBF算法

FCBF^[8]属于过滤式算法范畴,其采用后向顺序搜索策略,先剔除与类别信息不相关或弱相关特征,然后每次从剩余特征中选取排序首位的特征并剔除所有与其冗余的特征,以此快速筛选出低维特征子集。

FCBF算法使用SU指标评价特征与类别、特征与特征之间的相关性。SU的计算公式为

$$SU(X, Y) = 2 \left[\frac{H(X) - H(X|Y)}{H(X) + H(Y)} \right] \quad (1)$$

其中: $H(X)$ 为随机变量 X 的熵; $H(X|Y)$ 为条件熵,即在已知随机变量 Y 的条件下,随机变量 X 的不确定性; $SU(X, Y) \in [0, 1]$,取值为0,表示两特征完全独立;取值为1,表示两特征完全相关。

设故障数据集为 $D = \{(x_i, y_i) | i = 1, 2, \dots, m\} \in R^{m \times n}$,其中包含故障特征集 F 和类别标签 L ,即 $D = \{F, L\}$, $F = \{f_1, \dots, f_i, \dots, f_{n-1}\}$, $L = (l_i)_{m \times 1}$ 。

FCBF算法的步骤如下:

- 1) 输入故障特征集 F 、类别标签 L 和阈值 δ ,初始化集合 F_{best} ;
- 2) 对于每个特征 $f_i \in F (i = 1, 2, \dots, n-1)$,计算其与类别向量之间的 $SU(f_i, L)$ 值;
- 3) 选出 F 中 $SU(f_i, L) > \delta$ 的特征 f_i ,并将特征 f_i 按照 $SU(f_i, L)$ 值降序排列存入集合 S 中,由此剔除与类别信息不相关和弱相关特征;
- 4) 将 S 中排序第1的特征 f_j 作为主特征存入 F_{best} 中,然后从 S 中剔除特征 f_j ;
- 5) 对 S 中剩余的第 i 个特征 f_i ,计算其与主特征 f_j 的 $SU(f_j, f_i)$ 值。若 $SU(f_j, f_i) \geq SU(f_i, L)$,则认为 f_i 是与主特征 f_j 冗余的特征,并将其从 S 中剔除;
- 6) 重复步骤4和5,直到 S 为空集;
- 7) 输出精简后的特征子集 F_{best} 。

1.2 粒子群优化算法

PSO^[12]是一种启发式全局搜索算法。假设种群中存在 m 个粒子,每个粒子包含 d 维特征,设 $v_i =$

$(v_{i1}, v_{i2}, \dots, v_{id})$, $x_i = (x_{i1}, x_{i2}, \dots, x_{id})$ 分别为第 i 个粒子的速度向量和位置向量, $i = 1, 2, \dots, m$;粒子最优位置向量为 $P_{\text{best}i} = (p_{i1}, p_{i2}, \dots, p_{id})$,群体最优位置向量为 $g_{\text{best}i} = (g_{i1}, g_{i2}, \dots, g_{id})$ 。第 i 个粒子的速度和位置的更新表达式分别为

$$v_{id}^{k+1} = \omega v_{id}^k + c_1 r_1 (p_{\text{best}id}^k - x_{id}^k) + c_2 r_2 (g_{\text{best}id}^k - x_{id}^k) \quad (2)$$

$$x_{id}^{k+1} = x_{id}^k + v_{id}^{k+1} \Delta t \quad (3)$$

其中: v_{id}^k 和 x_{id}^k 分别为第 i 个粒子在 k 次迭代后的速度和位置; k 为当前迭代次数; Δt 为2次迭代之间的时间间隔,此处设为1; ω 为惯性权值; c_1 和 c_2 为加速因子; r_1 和 r_2 为 $[0, 1]$ 范围内的随机数。

1.3 支持向量机

SVM^[13]是一种有监督的机器学习算法。该算法在处理小样本、非线性和高维数据集的模式识别问题中具有良好的分类性能。其基本思想是寻找一个最优分类超平面,实现让不同类别数据子集间的间隔最大化。

给定训练样本集 $H = \{(x_i, y_i) | i = 1, 2, \dots, m\}$,其中: $x_i \in R^{n-1}$; $y_i \in \{-1, +1\}$ 。设超平面方程为 $\omega^T x + b = 0$,其中: ω 为法向量; b 为位移项。对于线性不可分样本,满足约束条件的目标函数为

$$\begin{aligned} \min_{\omega, b, \xi} & \left(\frac{1}{2} \|\omega\|^2 + C \sum_{i=1}^m \xi_i \right) \\ \text{s.t.} & \begin{cases} y_i (\omega^T x_i + b) \geq 1 - \xi_i \\ \xi_i \geq 0 \quad (i = 1, 2, \dots, m) \end{cases} \end{aligned} \quad (4)$$

其中: $C > 0$,为惩罚系数; $\xi_i \geq 0$,为松弛变量。

对于非线性分类问题,SVM通过核函数将样本映射到线性可分的空间。本研究采用的径向基核函数为

$$K(x_i, x_j) = \exp \left(- \frac{\|x_i - x_j\|^2}{2\sigma^2} \right) \quad (5)$$

其中: x_i, x_j 为代表2个任意样本的特征向量; $\sigma > 0$,为核函数参数。

由于惩罚系数 C 与核函数参数 σ 决定着SVM的分类性能,故在分类模型建立前选取合适的 C 和 σ ,将能够显著提高分类器的辨识精度。本研究拟通过IPSO算法对SVM的参数 C 和 σ 进行优化。

FCBF算法具有运行效率高、适用于剔除数据集集中的不相关特征和冗余特征,但仍存在不能完全去除冗余特征的缺陷^[13]。PSO算法虽然具有较强的全局搜索能力、可通过与SVM封装进一步去除子集中的冗余特征,但仍存在易陷入局部最优值、全

局与局部搜索不平衡和迭代后期收敛速度慢等不足。SVM具有优良的分类性能和计算效率,但其分类性能依赖模型参数。因此,本研究对PSO算法的不足予以改进,以提升收敛速度与寻优性能,并将其与SVM进行封装,实现对特征数据集的二次特征筛选和对SVM的参数优化。

2 改进的粒子群优化算法

2.1 IPSO算法的3种改进措施

2.1.1 采用 Tent 混沌映射初始化种群

在PSO算法中,一个好的种群初始位置可以加速算法收敛并且会提高最终解的质量^[14]。由于 Tent 混沌映射^[15]具有随机性、遍历性和均匀的概率密度等特性,故在此采用了式(6)所示的 Tent 混沌映射来代替随机初始化种群的方式,即

$$z_{k+1} = \begin{cases} z_k/\alpha & (0 < z_k \leq \alpha) \\ (1 - z_k)/(1 - \alpha) & (\alpha < z_k \leq 1) \end{cases} \quad (6)$$

其中: $\alpha \in (0, 1)$,在此取 $\alpha = 0.4$ ^[15]; z_k 为 Tent 映射第 k 次迭代生成的混沌值,用于赋值给粒子某维度的初始位置,且 $z_k \in (0, 1)$

图1为单个维度上 Tent 映射混沌序列的迭代轨迹。可见, Tent 映射生成的混沌值在搜索区间内均匀遍历,无明显聚集区域,基于该序列生成的初始种群能够均匀覆盖整个搜索空间。由于粒子能够从多个不同区域开始搜索,故更易遍历到全局最优解的位置,这将会有效降低陷入局部最优的风险。

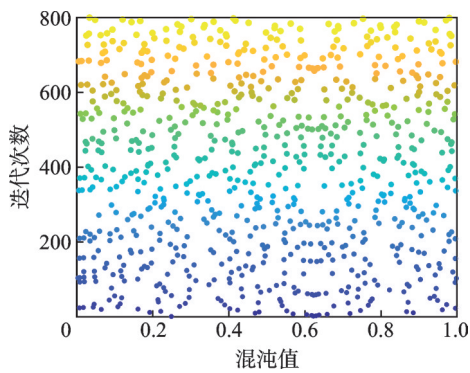


图1 Tent映射混沌序列的迭代轨迹

Fig.1 Iterative trajectory of the Tent map chaotic sequence

2.1.2 引入 Logistic 混沌映射更新惯性权值

较大的惯性权值 ω 可增强算法的全局搜索能力,使粒子不易陷入局部最优,反之则有助于提升局部搜索性能并提高解的精度^[16]。为了避免算法陷入局部最优并平衡全局与局部搜索性能,在此引入 Logistic 混沌映射^[15]来更新 ω 值,其结果是该映射可在一定范围内随机遍历取值,即使在迭代后期也能

够得到较大的 ω 值,使粒子跳出局部最优位置。

依据文献[17],设 ω 的取值范围为 $[0.4, 0.9]$, 其计算公式为

$$\omega = z_{k+1} = \mu z_k (1 - z_k) \quad (k = 0, 1, \dots) \quad (7)$$

其中: k 为当前迭代次数; z_k 为第 k 次迭代的混沌序列值, $z_0 \in [0.4, 0.9]$; 因控制参量 $\mu \in (0, 4]$, 故选取 $\mu = 4$ ^[15]。

基于 Logistic 混沌映射的惯性权值变化如图2所示。由图2可知, ω 值随着迭代次数在 $[0.4, 0.9]$ 区间内动态变化,在有效调节搜索状态的同时,还可避免算法陷入局部最优解,由此能够提升搜索过程的多样性。

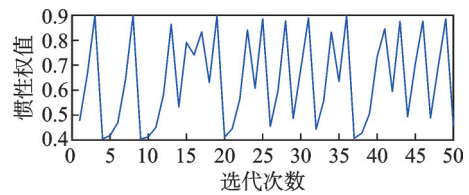


图2 基于 Logistic 混沌映射的惯性权值变化图

Fig.2 Inertia weights based on Logistic map changes

2.1.3 设计的基于正弦函数变化的加速因子更新策略

式(2)的加速因子 c_1 与 c_2 分别被称为认知成分和社会成分。当 c_1 较大时,算法的全局搜索能力强;当 c_2 较大时,其局部搜索能力强^[17]。据此原则,笔者提出了一种基于正弦函数变化的 c_1 和 c_2 更新策略,以增强算法前期的全局搜索性能和迭代后期的收敛速度。 c_1, c_2 的取值分别为 $[2.5, 0.5]$ 和 $[0.5, 2.5]$ ^[17], 更新公式分别为

$$c_1 = A \sin\left(\left(1 - \frac{k}{K}\right) \frac{\pi}{2}\right) + B \quad (8)$$

$$c_2 = -A \sin\left(\left(1 - \frac{k}{K}\right) \frac{\pi}{2}\right) + C \quad (9)$$

其中: $A = 2$; $B = 0.5$; $C = 2.5$; k 为当前迭代次数; K 为最大迭代次数。

加速因子的变化趋势如图3所示。

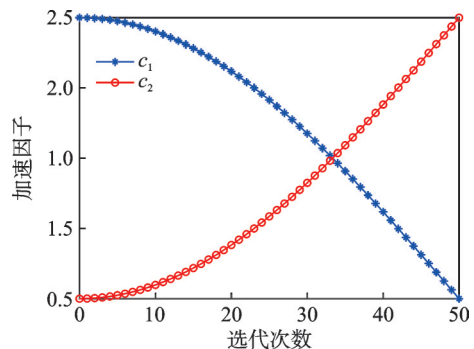


图3 加速因子的变化趋势

Fig.3 Variation trend of acceleration factor

2.2 粒子的表征形式及特征选择过程

在IPSO算法中,每个粒子可表示为 $1 \times D$ 维的行向量,其中 D 为特征维度、且每个维度的取值均应在 $0 \sim 1$ 之内。通过设定阈值 T 判定是否选中该维度的特征, T 为 $0.2 \sim 0.8$ 的随机数。如果该维度其值大于 T ,则选中该维度特征并将该维度的值赋为1;反之,则不被选中且赋值为0。通过上述方法即可将粒子的位置向量转换为已选取的特征向量。假定 $T=0.5$,则粒子的表征形式及特征选择过程如图4所示。



图 4 粒子的表征形式及特征选择过程

Fig.4 Particle representation form and feature selection process

2.3 IPSO 算法的适应度函数

在进行特征选择时,IPSO算法通过适应度函数来评价所选子集的优劣。参照文献[18],本研究构建的适应度函数为

$$\text{fitness}(x_i) = rE + (1 - r) \frac{F_{\text{selected}}}{F_{\text{all}}} \quad (10)$$

其中: E 为分类错误率; $r = 0.9^{[18]}$; F_{selected} 为已选择的特征数量; F_{all} 为特征总数。

2.4 IPSO 算法的收敛性能分析

设定 PSO 和 IPSO 这 2 个算法的初始参数相同,对比分析二者的收敛性能。图 5 为 2 种算法在不同适应度值的变化对比。图 5 表明,相对于 PSO 算法,IPSO 算法不仅收敛速度快,且收敛时的适应度值更低,这意味着特征子集的分类准确率更高。

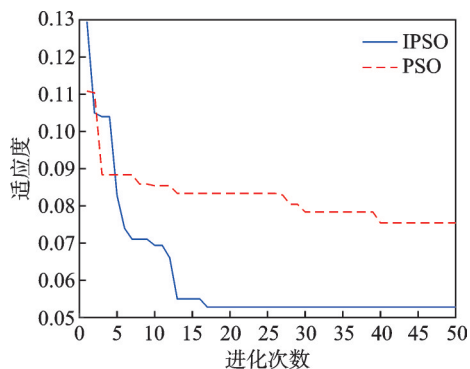


图 5 2 种算法在不同适应度值的变化对比

Fig.5 Comparison of different fitness values of 2 algorithms

这是因为 IPSO 算法采用了混沌初始化种群,使初始粒子的位置分布更加均匀,大大增强了算法的全局搜索能力,提高了算法的收敛性能。

2.5 IPSO 算法的寻优性能分析

为了验证 IPSO 算法的寻优能力,本研究使用同一数据集在相同搜索区域内对比了网格搜索、PSO 和 IPSO 算法在 SVM 中的参数寻优性能。SVM 最佳参数值及分类准确率如 1 所示。

表 1 SVM 最佳参数值及分类准确率

Tab.1 SVM best parameter value and classification accuracy

寻优算法	RBF 核函数	最佳 C 最佳 σ		运行	分类
				时间/s	准确率/%
网格搜索	(C, σ)	0.09	0.37	90.03	87.58
PSO	(C, σ)	1.10	0.48	41.15	87.78
IPSO	(C, σ)	0.44	0.24	34.94	87.78

由表 1 可知:①经 PSO 和 IPSO 寻优后的测试集,其分类准确率可以达到与使用全局搜索算法(网格搜索)相当的水平,但这 2 种算法的运行时间明显低于网格搜索算法;②与 PSO 算法相比,IPSO 算法的运行时间显著降低,即其寻优性能更为优异。这得益于混沌映射动态调整的惯性权值与基于正弦函数自适应变化的加速因子,二者协同改善了算法的搜索特性,使得该算法不仅能有效规避局部最优解,还能在较短时间内快速收敛到全局最优。

3 基于 FCBF-IPSO 的转子故障诊断流程

本研究设计的基于 FCBF-IPSO 的故障诊断流程如图 6 所示,对应的故障诊断步骤如下:

- 1) 对消噪后的振动信号提取时域、频域及时频域特征,构建初始故障特征数据集并进行归一化处理;
- 2) 设定阈值 δ ,利用 FCBF 算法剔除数据集中的不相关特征和冗余特征,得到特征子集;
- 3) 将特征子集按比例划分训练与测试数据集;
- 4) 通过 IPSO 算法进行二次特征筛选,进一步去除冗余特征,再将每次选择的训练集输入到 SVM 中进行训练并优化参数 C 和 σ ;
- 5) 设定最大迭代次数 K ,并将其设为算法的终止条件,迭代完成即输出优化的特征子集及参数,并完成 SVM 的训练;否则,返回步骤 4 继续循环。

6) 使用优化的特征子集和最佳参数在训练集上重新训练SVM分类器,随后,将测试集输入该模型完成故障辨识。

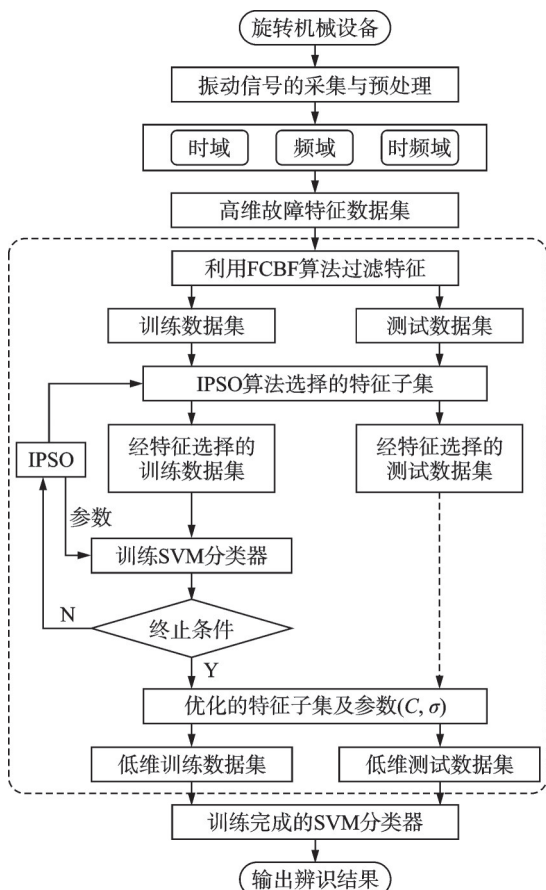


图6 基于FCBF-IPSO的故障诊断流程图

Fig.6 Fault diagnosis flow chart based on FCBF-IPSO

4 实验及结果分析

4.1 转子故障模拟实验

为了验证提出的FCBF-IPSO算法性能,本研究用图7所示的转子故障模拟实验台上的故障模拟信号集合对其可行性进行验证。该实验用6个型号为WT01的电涡流传感器,分别采集了2个轴承端面处转子的垂直方位信号、质量盘截面的垂直与水平方位信号。在采样频率为10 kHz、转速为3 000 r/min的工况下,模拟6类故障类型,分别为转子不对中、质量不平衡、动静碰磨、气流扰动、轴承座松动以及正常。各类故障下采集的100组样本按照6:4比例,划分成训练集与测试集。表2^[3]为单个通道提取的特征,其中:特征1~16为时域特征;17~22为频域特征;23~38为时频域特征。对各通道信号均提取了如表2^[3]所示的38个特征,6个通道信号

共228个特征构成了数据样本。

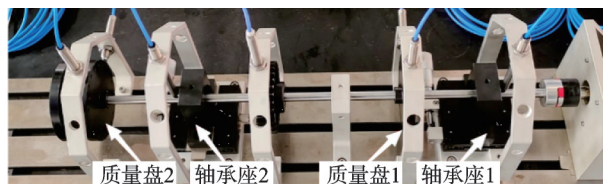


图7 转子故障模拟实验台

Fig.7 Rotor failure simulation test bench

表2 单个通道提取的特征^[3]

Tab.2 Features extracted from a single channel

序号	特征	序号	特征	序号	特征
1	最大值	9	方根幅值	17	均值频率
2	最小值	10	均方根值	18	频谱二阶矩
3	峰峰值	11	绝对均值	19	标准偏差率
4	方差	12	波形指标	20	峭度频率
5	均值	13	峰值指标	21	均方根频率
6	歪度	14	脉冲指标	22	中心频率
7	峭度	15	裕度指标	4层小波包 23~38 分解频带能量特征	
8	均方值	16	峭度指标		

4.2 算法的参数设定

图6中需设定的参数包括:FCBF算法的阈值 δ ,其需在保证FCBF算法第3步流程之后得到的特征子集分类准确率不低于80%的前提下设定;种群数 N 一般取20~50、迭代次数 K 常取50~100^[12],故在本研究中 $N=30$ 、 $K=50$ 。

4.3 实验结果分析

4.3.1 选择的特征子集与分类效果分析

为验证FCBF-IPSO算法的降维效果,本研究选择FCBF、ReliefF、IPSO、FCBF-PSO等算法进行对比分析。利用t-SNE算法^[19]分别对所选低维测试数据集的分类效果进行可视化呈现,如图8所示。

为进一步评估各特征选择算法对故障辨识精度的影响,将各算法所选特征子集输入到SVM中进行故障辨识,特征子集维数及不同状态识别率如表3所示。

将图8和表3进行对比可知:①经FCBF算法筛选出的19维特征子集,同类样本的聚集程度较好,但其中有2类故障样本仍然混叠在一起,因此故障辨识的准确率不高;②与FCBF算法相比,通过ReliefF算法仅保留权重排名前19位的特征时,有2类故障样本的混叠现象更严重,故障辨识的精度最

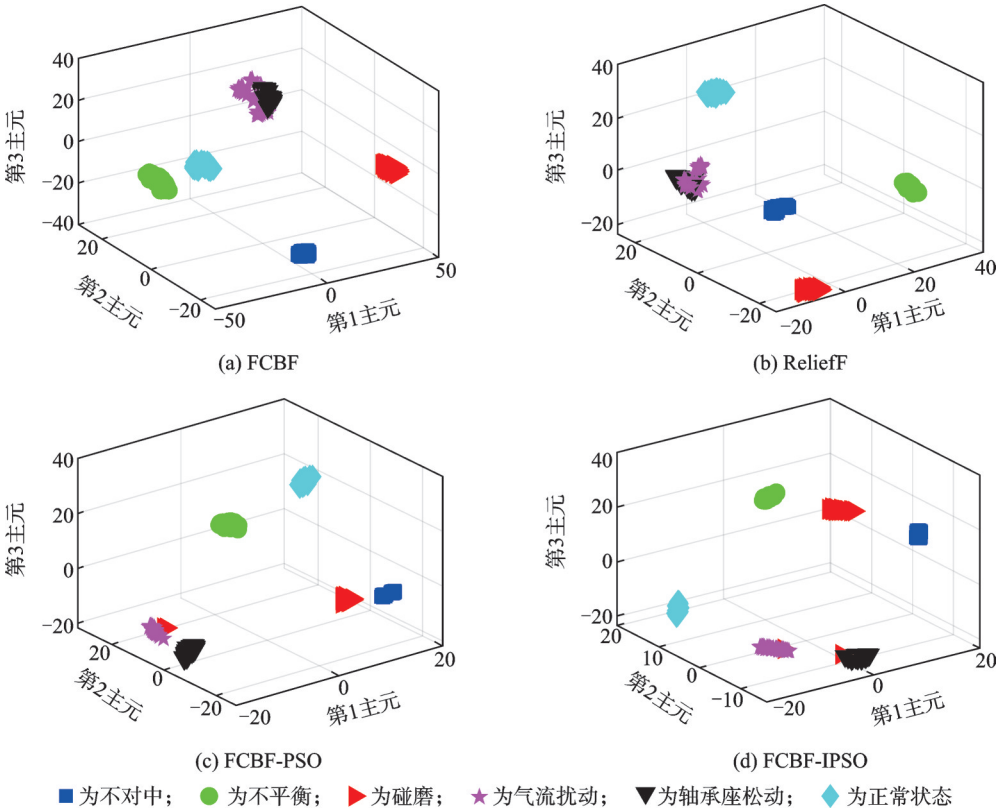


图 8 测试数据集分类效果的可视化结果

Fig.8 Test data set classification effect visualization

表 3 特征子集维数及不同状态识别率

Tab.3 Subset dimension and different state recognition rate

特征选择算法	子集维数	不同状态下的故障识别准确率/%						平均识别准确率/%
		不对中	不平衡	碰磨	气流扰动	松动	正常	
FCBF	19	100.0	100.0	100.0	37.5	100.0	100.0	89.6
Relieff	19	100.0	100.0	100.0	7.5	100.0	100.0	84.6
IPSO	11	100.0	100.0	100.0	60.0	100.0	100.0	93.3
FCBF-PSO	9	100.0	100.0	95.0	97.5	100.0	100.0	98.8
FCBF-IPSO	5	100.0	100.0	95.0	100.0	100.0	100.0	99.2

低。此结果表明,FCBF 算法能够剔除不相关特征和去除大部分冗余特征,表明该算法的特征筛选性能优于Relieff 算法。此外,笔者还尝试了单独使用IPSO 算法的效果,虽然其也能够取得较好的辨识精度,但计算量大、运行效率低,若想进一步提高辨识精度,就必须尝试融入过滤式特征选择方法。

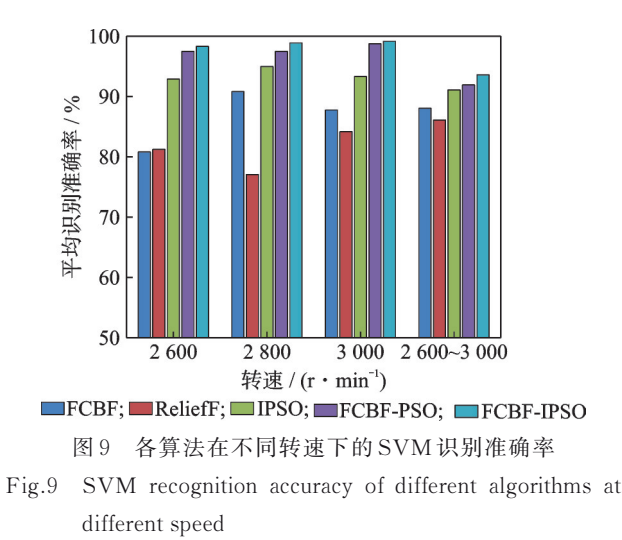
由图 8(d)可知,本研究提出的 FCBF-IPSO 算法的测试集分类效果最好,同类样本聚集程度高、不同类别样本之间的间隔大,只有个别样本被错分,平均识别准确率达到了 99.2%。进一步对比发现,FCBF-PSO 算法的分类性能和故障识别精度虽与 FCBF-IPSO 算法相近,但后者的改进机制使其不仅能够求得全局最优解、且所选的特征子集维数更低,

仅为前者的一半。这说明本研究提出的 FCBF-IPSO 算法筛选出的特征子集维度最低,表明本算法拥有最优分类性能和最高的故障识别准确率。

4.3.2 变工况实验

本研究选取 2 600、2 800 和 3 000 r/min 这 3 种工况下的故障数据集对 FCBF-ISPO 算法的变工况故障辨识性能进行了验证。将 3 个转速下的故障数据集混合后重新划分训练集与测试集,仍然选取 4 种特征选择算法与之进行对比。图 9 为各算法在不同转速下的 SVM 识别准确率。

图 9 表明,在不同转速下本研究所提算法的识别准确率均高于其他算法。虽然不同转速下的故障数据之间存在一些差异,特征之间的关联性也有变



化,这些差异导致了故障识别的准确率有所下降,但本算法仍展现出较强的抗转速波动性能和工况适

应性。

4.3.3 算法泛化性实验

为了验证 FCBF-IPSO 算法的通用性,采用文献[3]一套双跨转子的故障模拟信号集合对其性能进行了验证。该信号集合由 5 种故障状态组成,包括正常、轴承座松动、质量不平衡、转子不对中和动静碰磨。验证期间按 4:6 比例将每种状态的 100 组样本划分成训练集与测试集。每个通道信号仍然是按表 2^[3]提取 38 个特征参数,12 个通道信号共得到 456 个特征参数组成故障样本。

使用不同算法对故障数据集进行降维,并将筛选得到的低维特征子集输入到 SVM 中进行故障辨识,得到不同算法的子集维数及识别准确率如表 4 所示。

表 4 不同算法的子集维数及识别准确率							
Tab.4 The subset dimension and recognition accuracy of different algorithms							
特征选择算法	子集维数	不同状态下的故障识别准确率/%					平均识别准确率/%
		正常	松动	不平衡	不对中	碰磨	
FCBF	26	88.3	100.0	88.3	88.3	98.3	92.6
ReliefF	26	88.3	98.3	83.3	75.0	98.3	88.6
IPSO	15	100.0	93.3	100.0	100.0	73.3	93.3
FCBF-PSO	10	100.0	98.3	100.0	100.0	91.7	98.0
FCBF-IPSO	7	100.0	100.0	100.0	100.0	98.3	99.7

表 4 表明,本研究算法选择的特征子集维数最小且故障识别准确率最高。这表明本算法能够有效筛选出利于实施分类运算的低维特征子集,其故障辨识精度与泛化性能均表现优异。

5 结 论

- 1) 将过滤式和封装式特征选择算法进行融合,可有效剔除故障数据集中的不相关和冗余特征,得到一种低维敏感特征子集,由此可显著降低故障分类的难度。
- 2) 改进的 IPSO 算法收敛速度快、寻优能力强,利用其对 SVM 的参数 C 和 σ 进行优化,显著提高了分类器的辨识精度。
- 3) 本研究构建的 FCBF-IPSO 算法展现出更优的降维性能与更高的故障识别准确率,且在变工况环境下保持了良好的稳定性和泛化性能。研究表明,本研究能够为降低故障数据集的维度,继而提高其知识价值密度提供了一种数据科学依据。

参 考 文 献

[1] 原健辉,赵荣珍,邓林峰.半监督超边界 Fisher 分析的转子故障数据降维算法[J].振动、测试与诊断,2025,45(3): 543-550.

YUAN Jianhui, ZHAO Rongzhen, DENG Linfeng. Dimensionality reduction algorithm of rotor fault dataset based on semi-supervised hyper-marginal Fisher analysis [J]. Journal of Vibration, Measurement & Diagnosis, 2025, 45(3): 543-550. (in Chinese)

[2] 雷亚国,贾峰,孔德同,等.大数据下机械智能故障诊断的机遇与挑战[J].机械工程学报,2018,54(5): 94-104.

LEI Yaguo, JIA Feng, KONG Detong, et al. Opportunities and challenges of machinery intelligent fault diagnosis in big data era[J]. Journal of Mechanical Engineering, 2018, 54(5): 94-104. (in Chinese)

[3] 董晓鑫,赵荣珍.基于 SILPDA 的旋转机械故障诊断方法[J].振动与冲击,2023,42(2): 16-22.

DONG Xiaoxin, ZHAO Rongzhen. Fault diagnosis method of rotating machinery based on SILPDA [J]. Journal of Vibration and Shock, 2023, 42(2): 16-22.

- (in Chinese)
- [4] 戚晓利, 杨艳, 崔创创, 等. 基于RTSWMFE, IS-GSE与COOT-SVM的行星齿轮箱故障诊断[J]. 振动、测试与诊断, 2025, 45(1): 132-139.
QI Xiaoli, YANG Yan, CUI Chuangchuang, et al. Fault diagnosis of planetary gearbox based on RTSWMFE, IS-GSE and COOT-SVM[J]. Journal of Vibration, Measurement & Diagnosis, 2025, 45(1): 132-139. (in Chinese)
- [5] 李郅琴, 杜建强, 聂斌, 等. 特征选择方法综述[J]. 计算机工程与应用, 2019, 55(24): 10-19.
LI Zhiqin, DU Jianqiang, NIE Bin, et al. Summary of feature selection methods[J]. Computer Engineering and Applications, 2019, 55(24): 10-19. (in Chinese)
- [6] WANG Y, LI X, WANG J. A neurodynamic optimization approach to supervised feature selection via fractional programming[J]. Neural Networks, 2021, 136(5): 194-206.
- [7] LIU Z, CHANG B, CHENG F. An interactive filter-wrapper multi-objective evolutionary algorithm for feature selection[J]. Swarm and Evolutionary Computation, 2021, 65(4): 100925.
- [8] YU L, LIU H. Efficient feature selection via analysis of relevance and redundancy[J]. The Journal of Machine Learning Research, 2004, 5(12): 1205-1224.
- [9] ZHOU Y, KANG J, GUO H. Many-objective optimization of feature selection based on two-level particle cooperation[J]. Information Sciences, 2020, 532(1): 91-109.
- [10] 肖艳, 姜琦刚, 王斌, 等. 基于ReliefF和PSO混合特征选择的面向对象土地利用分类[J]. 农业工程学报, 2016, 32(4): 211-216.
XIAO Yan, JIANG Qigang, WANG Bin, et al. Object based land-use classification based on hybrid feature selection method of combining ReliefF and PSO[J]. Transactions of the Chinese Society of Agricultural Engineering, 2016, 32(4): 211-216. (in Chinese)
- [11] 薛瑞, 赵荣珍. ReliefF与QPSO结合的故障特征选择算法[J]. 振动与冲击, 2020, 39(11): 171-176, 208.
XUE Rui, ZHAO Rongzhen. The fault feature selection algorithm of combination of ReliefF and QPSO[J]. Journal of Vibration and Shock, 2020, 39(11): 171-176, 208. (in Chinese)
- [12] WANG D, TAN D, LEIL. Particle swarm optimization algorithm; an overview[J]. Soft Computing, 2018, 22(2): 387-408.
- [13] 叶明全, 高凌云, 伍长荣, 等. 基于对称不确定性和SVM递归特征消除的信息基因选择方法[J]. 模式识别与人工智能, 2017, 30(5): 429-438.
YE Mingquan, GAO Lingyun, WU Changrong, et al. Informative gene selection method based on symmetric uncertainty and SVM recursive feature elimination[J]. Pattern Recognition and Artificial Intelligence, 2017, 30(5): 429-438. (in Chinese)
- [14] GAO W F, LIU S Y. A modified artificial bee colony algorithm[J]. Computers & Operations Research, 2012, 39(3): 687-697.
- [15] YANG Y, GAO S C, SHI C, et al. CBSO: a memetic brain storm optimization with chaotic local search[J]. Memetic Computing, 2018, 10(4): 353-367.
- [16] TOO J, ABDULLAH A R, SAAD N M. A new co-evolution binary particle swarm optimization with multiple inertia weight strategy for feature selection[J]. Informatics, 2019, 6(2): 1-14.
- [17] SHAMI T M, EL-SALEH A A, ALSWAITTI M, et al. Particle swarm optimization: a comprehensive survey [J]. IEEE Access, 2022, 10: 10031-10061.
- [18] NGUYEN B H, XUE B, ANDREAE P, et al. A new binary particle swarm optimization approach: momentum and dynamic balance between exploration and exploitation[J]. IEEE Transactions on Cybernetics, 2021, 51(2): 589-603.
- [19] KIMURA M. Generalized t-SNE through the lens of information geometry[J]. IEEE Access, 2021, 9: 129619-129625.



第一作者简介:尹海涛,男,1997年11月生,硕士生。主要研究方向为旋转机械智能故障诊断技术。

E-mail: yinhaitao69@163.com

通信作者简介:赵荣珍,女,1960年12月生,博士、教授、博士生导师。主要研究方向为旋转机械故障诊断与机械系统动力学。

E-mail: zhaorongzhen@lut.edu.cn