

基于白鲸优化 VMD 算法和 CNN-BiLSTM 的滚动轴承故障诊断研究

黄 伟¹ 曾德良¹

HUANG Wei ZENG Deliang

摘 要

传统的滚动轴承故障诊断方法在从振动信号中提取特征信息时存在一定局限性, 从而影响故障诊断的准确性。文章通过结合白鲸算法优化的变分模态分解(VMD)、卷积神经网络(CNN)和双向长短期记忆(BiLSTM)方法研究了一种新的滚动轴承故障诊断技术。经过多组对比实验, 研究结果表明该方法在滚动轴承故障诊断方面表现出良好的效果, 与支持向量机(SVM)和核极限学习机(KELM)相比, VMD-CNN-BiLSTM 提升了 40.33% 的识别准确度, 达到 99.67%。结果表明: 该算法在滚动轴承故障诊断中具有显著优势, 为滚动轴承故障诊断提供了一种新的借鉴方法。

关键词

滚动轴承; 变分模态分解; 振动信号; 信号识别; 故障诊断

doi: 10.3969/j.issn.1672-9528.2025.01.025

0 引言

滚动轴承是机械设备中的关键部件, 支撑旋转载荷并影响设备稳定性。然而, 在复杂的运行环境下, 滚动轴承易出现故障, 且这些故障隐蔽性强, 识别难度颇高, 一旦发生, 小则导致设备停机, 造成生产中断, 大则可能给机械设备带来严重损害。目前, 振动信号分析是滚动轴承故障诊断最常用且行之有效的手段, 故而精准地对故障信号进行提取与诊断, 对于保障机械设备的正常运转、降低故障损失具有极为重要的意义^[1]。

为此, 研究者们开展了大量轴承故障诊断研究。Abboud 等人^[2]通过改进平方包络谱法, 提高变转速条件检测能力; Guo 等人^[3]结合 Hilbert 包络分析与 SVM, 实现精确诊断; 谢锋云等人^[4]结合时域特征、小波包变换与 LSSVM, 解决了多故障判别问题。

传统诊断方法包括时域与频域分析, 如均值、方差、峭度、FFT 等, 但在非线性、非平稳信号中表现有限。为此, 时频分析技术融合时域与频域成为重点。唐先广等人^[5]基于 STFT 提出改进包络分析。Qiu 等人^[6]通过优化小波变换提升准确度, 但计算成本较高。杨建华等人^[7]引入随机共振系统作为 EMD 预处理, 提升分解速度但噪声敏感。为此, Wu 等人^[8]提出 EEMD 抑制模式混叠, 但存在处理成本高等问题。

为有效提取故障特征, 本文采用 VMD, 其自适应性避免了递归方法误差, 但需优化分量数和惩罚参数。邢燕好等人^[9]

采用粒子群算法优化 VMD 参数, 用于电磁超声数据处理。冯寅等人^[10]提出基于 CNN 与 LSTM 的算法, 用原始数据识别故障, 但网络复杂性可能导致训练困难和性能退化问题。

本文采用白鲸优化算法(BWO)以最小包络熵为适应度函数优化 VMD 参数组 $[K, \alpha]$, 提取最佳 IMF 分量, 并基于 9 个指标(均值、方差、峰值、峭度等)构造特征向量, 有效避免模态混叠和端点效应。结合 CNN 和 BiLSTM, 利用 CNN 的局部特征提取能力和 BiLSTM 的时序处理优势, 最大程度保留全局和局部特征, 提升故障诊断的准确性。仿真结果表明, 该方法在准确性、灵敏度和特异性方面优于单一模型, 为滚动轴承故障诊断提供了有效参考。

1 白鲸优化算法

白鲸优化算法(BWO)由勘探、开发和鲸坠 3 个计算阶段组成, 其中自适应的鲸落概率和平衡因子在探索与开发能力的调控中至关重要。基于种群机制, 白鲸被视为搜索代理, 每只白鲸作为优化过程中的候选解, 其位置会持续更新。白鲸的种群和适应度表示为:

$$X = \begin{bmatrix} x_{1,1} & x_{1,2} & x_{1,3} & \cdots & x_{1,d} \\ x_{2,1} & x_{2,2} & x_{2,3} & \cdots & x_{2,d} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_{n,1} & x_{n,2} & x_{n,3} & \cdots & x_{n,d} \end{bmatrix}, \quad (1)$$

$$Y_X = \begin{bmatrix} f(x_{1,1}, x_{1,2}, x_{1,3}, \cdots, x_{1,d}) \\ f(x_{2,1}, x_{2,2}, x_{2,3}, \cdots, x_{2,d}) \\ \vdots \\ f(x_{n,1}, x_{n,2}, x_{n,3}, \cdots, x_{n,d}) \end{bmatrix}$$

1. 华北电力大学控制与计算机工程学院 北京 102206

式中：假设 n 代表白鲸种群数量； d 代表问题的维度； X 代表白鲸种群； Y 代表适应度。

BWO 算法通过平衡因子 B_f 实现了种群从探索到开发的平稳过渡。其计算公式为：

$$B_f = B_0 \cdot \left(1 - \frac{t}{2T}\right) \quad (2)$$

式中：在每次迭代中，随机生成一个介于 0 和 1 之间的值 B_0 ，并根据当前迭代次数 t 和总迭代次数 T 来调整 B_0 的取值。当平衡因子 B_f 大于 0.5 时，处于探索阶段；当平衡因子 B_f 小于等于 0.5 时，处于开发阶段。

在探索阶段，白鲸种群的位置用公式表示为：

$$\begin{cases} x_{i,j}^{t+1} = x_{i,p}^t + (x_{r,p}^t - x_{i,p}^t) \times \\ (1+r_1) \cdot \sin(2\pi r_2) \quad j=2n \\ x_{i,j}^{t+1} = x_{i,p}^t + (x_{r,p}^t - x_{i,p}^t) \times \\ (1+r_1) \cdot \cos(2\pi r_2) \quad j=2n+1 \end{cases} \quad (3)$$

式中：在下一个迭代中，第 i 个个体在第 j 维上的位置可表示为 $x_{i,j}^{t+1}$ 。其中， p 是一个随机整数，表示在问题维度为 D 的情况下，取值范围为 $[1, D]$ 。因此， $x_{i,j}^t$ 表示当前迭代下第 i 个个体在随机维度 p 上的位置。

另外， r 也是一个随机整数，表示种群规模为 N 时的范围为 $[1, N]$ 。因此， $x_{r,p}^t$ 表示当前迭代下，随机选取的个体 r 在随机维度 p 上的位置。而 r_1 和 r_2 均为介于 0 和 1 之间的随机数。采用正弦和余弦函数来模拟白鲸在水面上的镜像运动。位置信息的更新通过选择奇数和偶数维度进行，从而有效模拟两只白鲸在游泳及跃出水面时的同步和反射行为。

在开发阶段时，为了增强算法的收敛性，种群使用了 Levy 飞行策略。因而白鲸的种群位置更新为：

$$x_i^{t+1} = r_3 \cdot x_{\text{best}}^t - r_4 \cdot x_i^t + C_1 \cdot L_F \cdot (x_r^t - x_i^t) \quad (4)$$

$$C_1 = 2 \cdot r_4 \cdot \left(1 - \frac{t}{T}\right) \quad (5)$$

$$L_F = 0.05 \cdot \frac{\mu \cdot \sigma}{|\nu| \cdot \frac{1}{\beta}} \quad (6)$$

$$\sigma = \left[\frac{\Gamma(1+\beta)}{\beta \cdot \Gamma((1+\beta)/2)} \times \frac{\sin(\pi\beta/2)}{2^{(\beta-1)/2}} \right]^{1/\beta} \quad (7)$$

式中： r_3 与 r_4 是值在 0 到 1 的随机数； x_{best}^{t+1} 与 x_i^t 表示第 i 条白鲸在更新后和当前的位置； x_r^t 表示第 r 条白鲸当前的位置； x_{best}^t 表示种群中最好位置； $\Gamma(\cdot)$ 表示 gamma 函数。其中 μ 和 ν 是正态分布的随机数， β 是默认为 1.5 的常数。

种群在鲸落阶段时，白鲸的种群位置更新为：

$$X_i^{t+1} = r_5 X_i^t - r_6 X_i^t + r_7 X_{\text{step}} \quad (8)$$

式中： r_5 、 r_6 、 r_7 都是 (0, 1) 的随机数； X_{step} 表示鲸鱼坠落步长。

$$X_{\text{step}} = \exp\left(-\frac{t \cdot C_2}{T}\right) \cdot (U_b - L_b) \quad (9)$$

式中： U_b 和 L_b 表示搜索空间的上下界限； C_2 表示阶跃因子，表达式为 $C_2 = 2W_f \times N$ ，其值与鲸鱼掉落概率 W_f 和种群规模 N 有关。

$$W_f = 0.1 - 0.05 \times \frac{t}{T} \quad (10)$$

2 变分模态分解

VMD 算法的核心思想是建立一个变分问题，并解决这个变分模型。约束变分模型表示为：

$$\begin{cases} \min \left\{ \sum_k \left\| \partial_t \left[(\delta(t) + \frac{j}{\pi t}) \times u_k(t) \right] \times e^{-j\omega_k t} \right\|_2^2 \right\} \\ s.t. \sum_k u_k = f. \end{cases} \quad (11)$$

式中： $\{u_k\} = \{u_1, u_2, \dots, u_K\}$ 代表经过变分分解后的 K 个模态分量 IMF； $\{\omega_k\} = \{\omega_1, \omega_2, \dots, \omega_K\}$ 表示模态分量的中心频率； $f(t)$ 是原始信号； $\delta(t)$ 是狄拉克函数； ∂_t 表示对时间 t 求偏导数， $(\delta(t) + j/\pi t) \times u_k(t)$ 表示 $u_k(t)$ 的 Hilbert 谱。

借助二次平衡参数 α 和拉格朗日乘子 λ ，受约束的变分问题可以通过适当的方法被转化为无约束的变分问题，以便于求解：

$$\begin{aligned} L(u_k, \omega_k, \lambda) = & \alpha \sum_{k=1}^K \left\| \partial_t \left[(\delta(t) + \frac{j}{\pi t}) \right. \right. \\ & \times u_k(t) \left. \left. \right] e^{-j\omega_k t} \right\|_2^2 + \left\| f(t) - \sum_{k=1}^K u_k(t) \right\|_2^2 \\ & + \langle \lambda(t), f(t) - \sum_{k=1}^K u_k(t) \rangle \end{aligned} \quad (12)$$

用交替方向乘子法来解决上述问题，每次迭代时分别更新 ω_k 和 μ_k 的值，即

$$\begin{aligned} \hat{u}_k^{n+1}(\omega) = & \frac{\hat{f}(\omega) - \sum_{i \neq k} \hat{u}_i^n(\omega) + \frac{\hat{\lambda}^n(\omega)}{2}}{1 + 2\alpha(\omega - \omega_k^n)^2} \\ \omega_k^{n+1} = & \frac{\int_0^\infty \omega |\hat{u}_k^{n+1}(\omega)|^2 d\omega}{\int_0^\infty |\hat{u}_k^{n+1}(\omega)|^2 d\omega} \end{aligned} \quad (13)$$

式中： n 为迭代次数， $\hat{u}_k^{n+1}(\omega)$ 、 $\hat{f}(\omega)$ 、 $\hat{\lambda}^n(\omega)$ 分别表示为 $u_k^{n+1}(t)$ 、 $f(t)$ 、 $\lambda(t)$ 的傅里叶变换。继续迭代直至符合式 (14) 收敛条件。

$$\sum_{k=1}^K \left\| \hat{u}_k^{n+1} - \hat{u}_k^n \right\|_2^2 / \left\| \hat{u}_k^n \right\|_2^2 < \gamma \quad (14)$$

式中： γ 是收敛误差。

3 网络模型简介

CNN 受动物视觉皮层研究启发，模拟人类视觉系统，由输入层、卷积层、池化层、全连接层和 Softmax 层组成。前向传播提取特征，反向传播优化参数。

BiLSTM 是处理序列数据的深度学习模型，结合正向和反向 LSTM 捕捉上下文信息。正向 LSTM 按时间顺序处理数据，反向 LSTM 逆序处理，两者结合利用过去和未来信息，通过输出层完成预测。

传统 CNN 通过卷积核学习局部特征，难以捕捉全局信息。

因此，本研究在池化层与全连接层之间添加了 BiLSTM 层，以捕获音频信号中的长期依赖。特征通过输入层进入网络，卷积层提取空间特征，ReLU 函数增强非线性性和稀疏性。池化层减少特征维度，降低过拟合风险。BiLSTM 层提取序列的深层特征，全连接层整合特征并通过 Softmax 分类器进行分类。

4 本文算法步骤

基于滚动轴承振动信号的故障诊断方法，往往采用对振动信号进行特征提取，再利用信号以及特征向量和各类算法实现轴承故障诊断，本文将白鲸优化 VMD 算法和 CNN-BiLSTM 应用到轴承故障诊断中，采用白鲸优化算法（BWO）对 VMD 参数进行寻优，找到 VMD 最优的分解层数和惩罚因子，并利用优化后的 VMD 对轴承原始信号进行分解。通过所得 IMF 分量中的最小包络熵原则，在每个样本中选择最佳的 IMF 分量，然后对其 9 个指标进行分别计算，使用 9 个指标构建特征向量作为模型输入，特征向量被分割训练集和测试集为 3:1，用于构建滚动轴承故障诊断模型。训练集被输入到 CNN-BiLSTM 模型中进行分类训练，而测试集则用来验证模型的准确性。本文的故障诊断模型如图 1 所示。

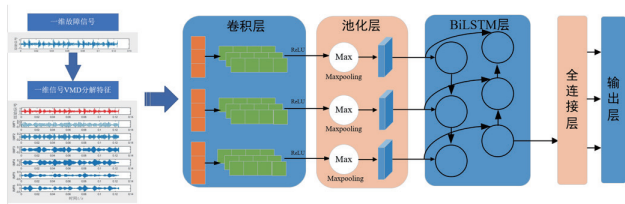


图 1 白鲸优化 VMD-CNN-BiLSTM 故障诊断模型

5 实验验证

5.1 数据集介绍

本文采用 CWRU 滚动轴承振动数据集，包含多工况下轴承正常及故障状态（内圈、外圈、滚动体故障）的振动信号。数据通过故障诊断实验台采集，使用加速度计固定在驱动端和风扇端，记录多种负载和转速条件下轴承信号。故障由电火花加工技术制造，实验设备包括 2 马力电机、扭矩传感器和加载电机。该数据集广泛用于轴承故障诊断及预测技术验证。如图 2 所示。

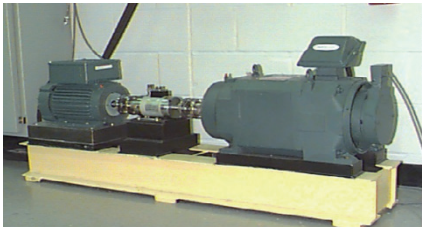


图 2 滚动轴承试验台

采样频率为 12 kHz，负荷大小为 2 HP。所用轴承型号为

SKF6205，涉及 3 种不同的故障直径（0.007 英寸、0.014 英寸、0.021 英寸），转速为 1750 r/min，采样位置和故障位置分别位于驱动端的内圈故障（IRF）、外圈故障（ORF）、滚珠故障（BF）以及正常状态信号。数据集包括原始时间序列数据，这些数据可以用于直接的时间域分析，或者转换到频域进行分析。依据轴承尺寸参数，得到故障特征频率如表 1 所示。

表 1 SKF6205 轴承故障特征频率

故障类型	特征频率 f_i / Hz
内圈故障	161.73
外圈故障	107.82
滚动体故障	61.99

5.2 数据处理

本文在美国西储大学轴承数据集中选取的驱动端数据均为 MATLAB 格式。其中 97.mat 表示的是正常运行数据，107.mat、120.mat、132.mat 分别代表同等直径尺寸 0.007 英寸，转速同为 1750 时的 3 种故障类型。该数据集包含 3 种轴承故障类型，分别为滚动体点蚀、内圈点蚀和外圈点蚀。对于 10 种类型，每种收集 120 组样本，滑动窗口 w 的大小设置为 1000，样本故障点 m 取 2048，收集好的数据集后用作特征提取。

5.3 故障特征的提取

本文共涉及 10 种故障类型，每种故障抽取了 120 组样本，每组样本长度为 2048 个数据点。除正常信号外，每种故障的 120 组样本几乎包含了所有数据点。从每个类别标签的 120 组样本中，随机选择 90 组样本作为训练样本，另外 30 组样本作为测试样本。设置优化变量数目为 2，惩罚参数 α 的上下限为 [100, 2500]，分量的数量上下限为 [3,10]。最大迭代数目 T 为 50，种群规模 N 为 10。图 3 展示了实测获取的 3 种故障状态及正常状态下的原始振动信号时域波形图。

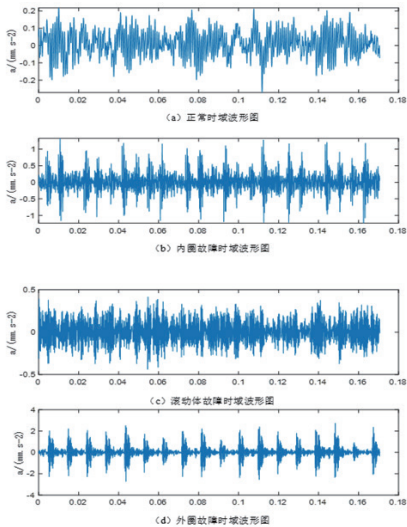


图 3 原始振动信号波形图

本文利用白鲸优化算法（BWO）自主优化变分模态分解

（VMD）的分解层数 K 和惩罚因子 α ，以最小包络熵为适应度函数。实验选用 9 种故障类型和 1 种正常类型数据，共 1200 组样本，每组包含 2048 个数据点。结果如图 4、图 5 所示，轴承振动信号被分解为 6 个瞬时模态函数（IMF），各 IMF 中心频率分别约为 600 Hz、1500 Hz、2600 Hz、2800 Hz、3400 Hz 和 3700 Hz，其频谱独立，避免模态混叠问题，可重构原始信号频谱。与传统极值包络递归提取方法相比，BWO 优化的 VMD 使分解过程更简洁，模态分量纯净度更高。

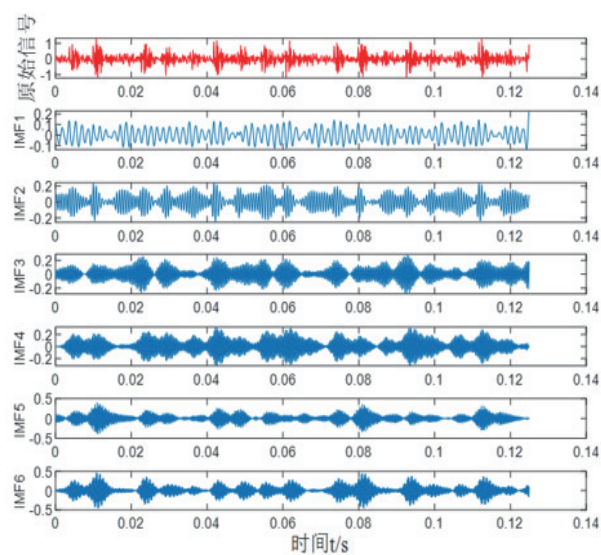


图 4 原始信号经 VMD 分解后的各 IMF 图

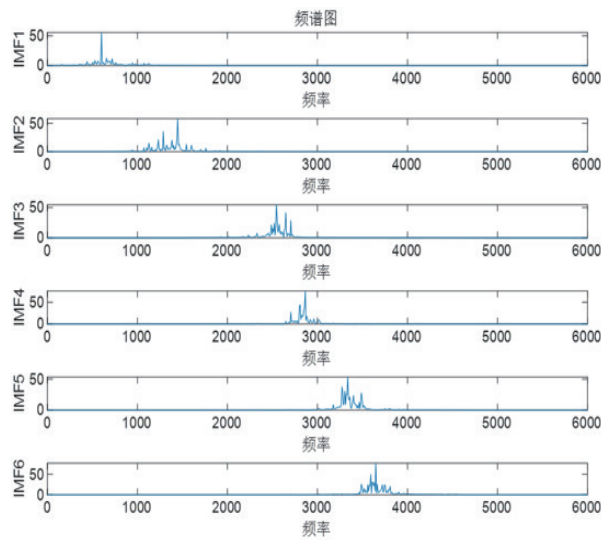


图 5 各 IMF 分量的中心频率

5.4 故障诊断与算法对比

5.4.1 与同类算法对比

接着以最小包络熵对应的 IMF 对应分量求取均值、方差、峰值、峭度、有效值、峰值因子、脉冲因子、波形因子、裕度因子^[11]，以特征构造特征向量为基础，研究中将每类故障

的前 90 组样本的特征向量数据输入至 CNN-LiLSTM 模型进行训练，而后将剩余 30 组特征向量数据输入 CNN-BiLSTM 模型进行预测。训练和预测的结果如图 6 所示。由图 6 可知该算法的最终识别率达到了 99.67%。为了展示本方法在故障诊断中准确性和效率上的优势，将优化后的算法的诊断效果与未经过优化的 VMD 提取的 CNN-BiLSTM 诊断模型、CNN-LSTM 诊断模型以及经过 VMD 优化的 CNN-LSTM 诊断模型对比。

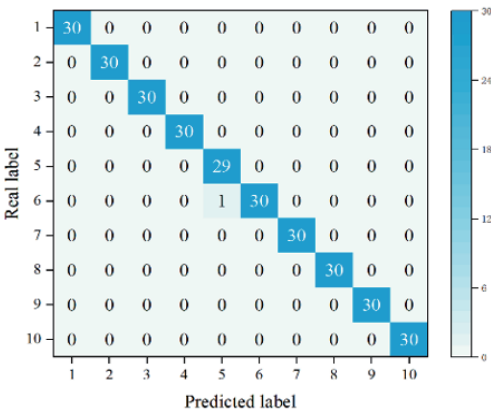


图 6 白鲸优化 VMD-CNN-BiLSTM 诊断结果

图 7 展示了 CNN-BiLSTM 诊断模型的结果，识别率达到了 88%。

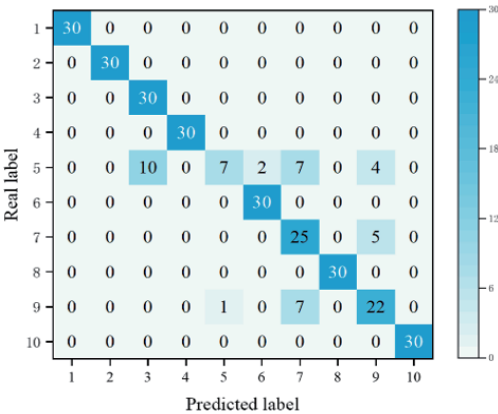


图 7 CNN-BiLSTM 诊断结果

就准确率而言，VMD-CNN-BiLSTM 算法在测试集上的识别率高达 99.67%，并比其他算法高出 16.33%。在保证高准确率的同时，该方法的优化时间也比其他算法更短。4 种算法故障诊断结果对比如表 2 所示。综上所述，本文提出的方法在故障诊断方面表现出更优异的性能。

表 2 故障诊断结果对比

所用算法	准确率 /%	时间 /s
VMD-CNN-BiLSTM	99.67	9.09
VMD-CNN-LSTM	99.00	10.75
CNN-BiLSTM	88.00	220.12
CNN-LSTM	83.33	157.34

5.4.2 与其他算法对比

用未经过白鲸优化 VMD 算法特征提取过得原始数据, 数据量为 1200×2048 , 输入到常用分类诊断模型中, 如支持向量机 (SVM) 和核极限学习机 (KELM)。所得诊断效果很差, 其中 KELM 诊断结果如图 8 所示, 故障识别准确率仅为 59.33%, 说明原始数据采取特征提取和特征向量构造步骤的重要性。原始数据经过 SVM、KELM 和减法平均优化器 (SABO) 优化核极限学习机 (KELM) 之后的故障识别诊断结果如表 3 所示。KELM 的表现较差, 原因可能为未经特征提取、核函数选择不够理想或模型参数未调整至以适应数据集。SVM 这一常用强大分类器在诊断结果中表现极差, 进一步说明了特征提取及特征工程的重要性。

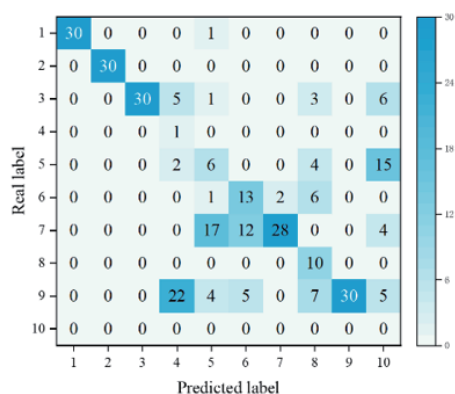


图 8 原始数据经过 SABO 优化 KELM 诊断分类结果混淆矩阵

表 3 其他分类算法故障诊断结果对比

所用算法	准确率 /%
SABO-KELM	59.33
KELM	24.00
SVM	32.67

6 结论

为了实现轴承早期故障预警、健康状态评估以及优化维护策略, 本文提出了一种结合白鲸优化算法优化的 VMD 特征提取方法和 CNN-BiLSTM 算法的诊断模型。实验结果表明: (1) 通过 VMD 分解原信号, 并以最小包络熵为自适应函数, 提取最小包络熵对应的 IMF 分量, 再提取这些分量的 9 种时域指标作为特征向量, 能够有效保留故障信息, 实现准确的故障特征提取; (2) VMD-CNN-BiLSTM 算法在故障识别中表现出高准确率和快速诊断的优势, 验证了方法的有效性, 并为类似转轮和轴承故障诊断提供了有效解决方案。

参考文献:

[1] 张琛, 赵荣珍, 邓林峰. 基于 EEMD 奇异值熵的滚动轴承故障诊断方法 [J]. 振动. 测试与诊断, 2019, 39(2):353-

358+446-447.

- [2] ABBOUD D, ANTONI J, SIEG-ZIEBA S, et al. Envelope analysis of rotating machine vibrations in variable speed conditions: a comprehensive treatment[J]. Mechanical systems and signal processing, 2017, 84(2): 200-226.
- [3] GUO L, CHEN J, LI X L. Rolling Bearing Fault Classification Based on Envelope Spectrum and Support Vector Machine[J]. Journal of vibration and control, 2009, 15(9): 1349-1363.
- [4] 谢锋云, 符羽, 王二化, 等. 基于多特征提取和 LSSVM 的轴承故障诊断 [J]. 机床与液压, 2020, 48(17): 188-190+157.
- [5] 唐先广, 郭瑜, 丁彦春, 等. 基于短时傅里叶变换和独立分量分析的滚动轴承包络分析 [J]. 机械强度, 2012, 34(1): 1-5.
- [6] QIU H, LEE L J, LIN J, et al. Wavelet filter-based weak signature detection method and its application on rolling element bearing prognostics[J]. Journal of sound and vibration, 2006, 289(4-5): 1066-1090.
- [7] 杨建华, 韩帅, 张帅, 等. 强噪声背景下滚动轴承微弱故障特征信号的经验模态分解 [J]. 振动工程学报, 2020, 33(3): 582-589.
- [8] WU Z H, HUANG N E. Ensemble empirical mode decomposition: a noise-assisted data analysis method[J]. Advances in adaptive data analysis, 2009, 1(1): 1-41.
- [9] 邢燕好, 于昊, 张佳, 等. 基于粒子群参数优化的 O-VMD 数据处理方法研究 [J]. 仪器仪表学报, 2023, 44(4): 304-313.
- [10] 冯寅, 贾宏涛, 杨震强, 等. 基于 CNN-LSTM 的输电线路故障诊断方法研究 [J]. 电网与清洁能源, 2023, 39(11): 59-65.
- [11] 杨森, 王恒迪, 崔永存, 等. 基于改进 AFSA 的参数优化 VMD 和 ELM 的轴承故障诊断 [J]. 组合机床与自动化加工技术, 2023(4): 67-70.

【作者简介】

黄伟(1997—), 男, 江西九江人, 硕士研究生, 研究方向: 发电机组故障诊断与信号处理。

曾德良(1969—), 男, 四川成都人, 博士, 教授, 研究方向: 发电过程建模与优化控制和智能发电。

(收稿日期: 2024-09-23)