

# 基于数据层特征融合的滚动轴承故障诊断\*

张 雄<sup>1,2</sup>, 李嘉禄<sup>2,3</sup>, 武文博<sup>2</sup>, 董 帆<sup>2</sup>

(1. 河北省电力机械装备健康维护与失效预防重点实验室 保定, 071003)

(2. 华北电力大学机械工程系 保定, 071003)

(3. 天津津铁电子科技有限公司 天津, 300380)

**摘要** 针对传统时域、频域统计指标构建的特征矩阵在处理复杂非线性数据时的局限性, 提出一种基于数据层特征融合和卷积神经网络(convolutional neural network, 简称 CNN)相结合的轴承故障诊断方法。首先, 将采集到的信号进行变分模态分解(variational mode decomposition, 简称 VMD), 选择峭度值大于阈值的分量进行重构; 其次, 从时域、频域、能量和稳定性等方面计算构建重构信号的多维度复合特征矩阵, 将数据特征进行融合; 然后, 对特征矩阵进行核主成分分析(kernel principal component analysis, 简称 KPCA)降维处理, 去除冗余信息; 最后, 将得到的低维矩阵输入到批量归一化(batch normalization, 简称 BN)层优化的 CNN 模型中进行故障识别与分类, 并通过辛辛那提大学智能维护系统(intelligent maintenance systems, 简称 IMS)、旋转机械振动分析与故障诊断实验平台(QPZZ-II 型)2 组实验数据进行数据验证。结果表明: 所提方法对轴承故障分类具有较好的处理效果和稳定性。

**关键词** 滚动轴承; 特征矩阵; 降维; 卷积神经网络

**中图分类号** TH133; TH17

## 引 言

滚动轴承是机器运行的关键部件, 若发生故障会影响机器运行的稳定性, 严重时甚至发生安全事故<sup>[1-3]</sup>。因此, 对滚动轴承进行有效故障识别与诊断具有极其重要的意义。在以往研究中, 轴承研究方向多以传统方法为主, 如对原始故障信号进行频谱分析、包络谱分析以及利用模态分解算法对信号进行分解等。然而, 传统方法较为依赖专家经验, 过多的人工干预难免会对诊断结果产生一定影响。随着轴承故障情况的复杂化, 单纯使用传统诊断方法已不能满足轴承故障诊断的需要。

随着智能化水平的提高, 各种轴承数据的收集速率与数量级均得到广泛提升, 这也为深度学习方法进入轴承故障诊断领域奠定了良好基础<sup>[4-6]</sup>。深度学习作为一种当前常用的智能化研究方法, 在对轴承数据的分析过程中可避免人工干预, 并减少由经验带来的误差, 因而越来越广泛地应用于轴承故障诊断领域。张继旺等<sup>[7]</sup>将 VMD 与 CNN 相结合, 对轴承早期微弱故障进行诊断。李思琦等<sup>[8]</sup>利用经验模态分解(empirical mode decomposition, 简称

EMD)算法分解信号, 选取有用分量重构信号, 计算一系列指标后输入 CNN 进行故障诊断。姚峰林等<sup>[9]</sup>先用小波包变换对信号进行消噪处理, 再提取消噪信号的时域特征值后, 输入极限学习机进行故障分类与诊断。何江江等<sup>[10]</sup>将支持向量机和集合 EMD 算法结合进行信号分解, 对分解的分量进行选择后再分解, 并构建能量向量, 最后导入 CNN 中进行验证。李子国等<sup>[11]</sup>用哈里斯鹰算法优化 VMD 相关参数, 用优化后的 VMD 分解原始振动信号, 并选择最优模态分量后, 输入一维 CNN 模型进行故障诊断。

上述方法对轴承的故障诊断均取得了较好的效果, 但均为利用时域或频域的特征指标对数据进行的分析。由于轴承振动数据的非平稳性和非线性, 时域特征指标和频域特征指标可能受到非周期性瞬态冲击和谐波成分的干扰。此外, 时域特征指标对于冲击特性比较敏感, 信号中的瞬态冲击成分可能会被错误诊断为轴承故障; 而谐波信号在频谱中与冲击成分类似, 用频域特征指标进行诊断时也难免出现误差, 因而增大了诊断的难度。

针对上述问题, 本研究提出一种基于数据层特征融合和 CNN 相结合的诊断方法。首先, 对信号进

\* 国家自然科学基金资助项目(52105098); 河北省自然科学基金资助项目(E2024502052, E2021502038); 中央高校基本科研业务费专项资金资助项目(2025MS137)

收稿日期: 2022-12-26; 修回日期: 2023-02-14

行VMD分解并重构,摒弃干扰信息,最大限度保留故障特征;其次,从时域、频域、能量和稳定性等方面对重构信号构建多维度复合特征矩阵,将数据特征进行融合;然后,利用KPCA对特征矩阵进行降维,简化矩阵;最后,用经过BN层优化的CNN模型对低维特征矩阵进行识别与诊断,并分别以IMS和QPZZ-II实验数据验证所提方法对轴承多故障分类的处理效果和稳定性。

## 1 VMD算法

VMD算法是一种自适应的模式变分和信号处理方法<sup>[12]</sup>,其主要思想是通过控制带宽来避免模式混叠现象。VMD可自适应地实现信号的频域部分和各分量的有效分离<sup>[13]</sup>,相比于传统的EMD方法有着极大改进。VMD分解原理如下。

1) 将原始信号分解为 $m$ 个本征模态分量,其受约束的变分表达式为

$$\min_{\{u_m\}, \{\omega_m\}} \left\{ \sum_m \left\| \partial_t \left[ \left( \delta(t) + \frac{j}{\pi t} \right) u_m(t) \right] e^{-j\omega_m t} \right\|_2^2 \right\} \quad (1)$$

s.t.  $\sum_m u_m = f$

其中: $u_m$ 和 $\omega_m$ 分别为模态分量和中心频率。

2) 引入惩罚因子 $\alpha$ ,构造增广拉格朗日函数,从而求出各个本征模态函数(intrinsic mode function, 简称IMF)分量,即变分模型的最优解

$$L(\{u_m\}, \{\omega_m\}, \lambda) = \alpha \sum_m \left\| \partial_t \left[ \left( \delta(t) + \frac{j}{\pi t} \right) u_m(t) \right] e^{-j\omega_m t} \right\|_2^2 + \left\| f(t) - \sum_m u_m(t) \right\| + \left\langle \lambda(t), f(t) - \sum_m u_m(t) \right\rangle \quad (2)$$

3) 对拉格朗日函数进行时频域变换并对其求极值,得到 $u_m$ 和 $\omega_m$ 的频域表达式为

$$\hat{u}_m^{n+1}(\omega) = \frac{\hat{f}(\omega) - \sum_{i \neq m} \hat{u}_i^{n+1}(\omega) + \frac{\hat{\lambda}(\omega)}{2}}{1 + 2\alpha(\omega - \omega_m)^2} \quad (3)$$

$$\omega_m^{n+1} = \frac{\int_0^\infty \omega |\hat{u}_m(\omega)|^2 d\omega}{\int_0^\infty |\hat{u}_m(\omega)|^2 d\omega} \quad (4)$$

4) 采用交替方向乘子算法得到约束变分模型的最优解。

在VMD算法的多个参数中,模态分量个数 $m$ 对分解结果影响较大。若 $m$ 值过大,则信号被过度分

解,会出现模式混叠现象,影响分解的准确率<sup>[14]</sup>;若 $m$ 值过小,则信号欠分解,各阶模态中心频率的收敛速度变慢,影响分解效率。在进行实验前,本研究对 $m$ 取不同值的情况进行验证,发现当 $m=4$ 时,虽然未出现模式混叠现象,但各阶模态中心频率收敛速度较慢,分解效率较低;当 $m=6$ 时,有模式混叠现象发生,且收敛速度较慢;当 $m=5$ 时效果最好,各阶分量中心频率波动较小,收敛速度较快,无模式混叠现象。因此,在实验中将原始信号分解为5个模态分量。

## 2 特征矩阵构建

轴承振动信号的时频域特征值可反应轴承的总体状态,是衡量信号特征的重要指标。常用的时频域特征有峭度、均方根、峰峰值、偏度、裕度因子、波形因子、峰值因子、脉冲因子、重心频率、平均频率、频率标准差和均方根频率等。但考虑到有些时频域特征之间具有相关性和重复性,以及时频域特征可能受到瞬态冲击和谐波成分等无用信息的干扰,在进行故障特征提取时仅凭时频域特征指标难以有效反映样本的有用信息,因此需要引入其他特征指标进行补充。

除时频域特征外,能量法和熵值法也是故障诊断中常用的方法。由于不同故障信号所含的能量不同,其复杂程度和稳定性也不同,因此根据信号的能量和熵值也可区分不同的信号。计算信号的多种特征值并构建多维度复合特征矩阵,可以提高特征提取效果,达到从多方面反映样本有效信息的目的。本研究中,选取上述的12种时频域特征和短时能量、Teager能量算子2个能量值以及近似熵、样本熵2个熵值共16个特征进行特征值的计算,并将计算结果组合构建成16维复合特征矩阵,利用特征矩阵进一步分析和处理信号。

## 3 KPCA降维

特征降维是一种常用的数据处理方法,特征矩阵降维后可以得到优化,数据分布也可以进行可视化展示,便于观察数据。KPCA是一种非线性主元分析降维方法,其主要思想是先将非线性的样本映射到高维空间,再在高维空间中使用线性降维。其算法原理如下。

1) 利用核函数把待降维数据映射到高维空间,即

$$X = \begin{bmatrix} x_{11} & \cdots & x_{1n} \\ \vdots & & \vdots \\ x_{n1} & \cdots & x_{nm} \end{bmatrix} \Rightarrow \varphi(X) = \varphi \left( \begin{bmatrix} x_{11} & \cdots & x_{1n} \\ \vdots & & \vdots \\ x_{n1} & \cdots & x_{nm} \end{bmatrix} \right) \quad (5)$$

2) 计算高维空间中数据的协方差矩阵,进一步计算特征值与特征向量,即

$$\varphi(X)\varphi(X)^T w_l = \lambda w_l \quad (6)$$

3) 使用高维样本点将高维特征空间的投影向量以线性表示为

$$\begin{cases} \varphi(X)\varphi(X)^T w_l = \lambda w_l \\ w_l = \sum_{i=1}^m \varphi(x_i) \alpha_i = \varphi(X) \alpha \end{cases} \quad (7)$$

上述推导过程为  $\varphi(X)\varphi(X)^T \varphi(X) \alpha = \lambda \varphi(X) \alpha \Rightarrow \varphi(X)^T \varphi(X) \varphi(X)^T \varphi(X) \alpha = \lambda \varphi(X)^T \varphi(X) \alpha \Rightarrow N^2 \alpha = \lambda N \alpha \Rightarrow N \alpha = \lambda \alpha$

4) 进行核矩阵计算,矩阵  $N$  计算式为

$$N = \varphi(X)^T \varphi(X) = \begin{bmatrix} \kappa(x^{(1)}, x^{(1)}) & \kappa(x^{(1)}, x^{(2)}) & \cdots & \kappa(x^{(1)}, x^{(n)}) \\ \kappa(x^{(2)}, x^{(1)}) & \kappa(x^{(2)}, x^{(2)}) & \cdots & \kappa(x^{(2)}, x^{(n)}) \\ \vdots & \vdots & & \vdots \\ \kappa(x^{(n)}, x^{(1)}) & \kappa(x^{(n)}, x^{(2)}) & \cdots & \kappa(x^{(n)}, x^{(n)}) \end{bmatrix} \quad (8)$$

通过变换使得核矩阵  $N$  更为聚集(中心化),即

$$N' = N - l_n N - N l_n + l_n N l_n \quad (9)$$

5) 求解矩阵  $N$  的特征值和特征向量,取最大的  $d$  个特征值对应的特征向量,以列向量组成向量矩阵,即为降维后的数据集。

## 4 CNN

CNN 是当前较为常用的一种神经网络,最早主要用于图像处理。随着研究发展,多数数据也逐渐开始采用 CNN 进行处理与分析。典型的 CNN 结构由输入层、卷积层、池化层、全连接层以及输出层组成<sup>[15-17]</sup>。CNN 结构模型如图 1 所示。输入层的作用是接收传入神经网络的信号,本研究是将经 KPCA 降维后的低维矩阵作为信号输入;卷积层的作用是对输入的信号进行局部感知,提取重要特征;池化层一般用于对卷积层提取的特征进行降维,增加模型的抗畸变能力;全连接层负责将前面一系列处理后输出的二维特征矩阵转化成一个一维向量;输出层位于全连接层之后,对从全连接层得到的特征进行分类输出。

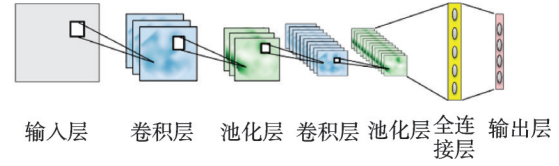


图1 CNN结构模型

Fig.1 Structure model of CNN

CNN 的优势在于:①泛化能力较强,优于其他方法,已广泛应用于各个领域;②有良好的容错能力,可处理各种复杂的信息,包括有缺损、畸变的信息;③具有极强的适应性,善于提取数据的关键特征,分类效果较好。

对于传统神经网络,随着模型复杂程度的增加,训练难度同步增大,甚至出现过拟合现象,严重影响训练效果。本研究对传统 CNN 进行改进,在卷积后引入 BN 层,可起到加快训练和收敛的速度以及防止过拟合的作用。本研究在经过多次实验对比后,确定 CNN 与 BN 算法的联合使用可大幅提高诊断结果的准确率,相较于单一 CNN 的效果有极大改进。CNN 参数设置如表 1 所示。

表 1 CNN 参数设置

Tab.1 Parameter settings of CNN

CNN 网络结构	参数设置
卷积层 $C_1$	$10 \times 10 \times 4$
批量标准层 $B_1$	—
最大池化层 $M_1$	池化窗口 $1 \times 1$ , 步长 1
卷积层 $C_2$	$5 \times 5 \times 4$
批量标准层 $B_2$	—
最大池化层 $M_2$	池化窗口 $1 \times 1$ , 步长 1
卷积层 $C_3$	$3 \times 3 \times 8$
批量标准层 $B_3$	—
卷积层 $C_4$	$3 \times 3 \times 8$
批量标准层 $B_4$	—
卷积层 $C_5$	$3 \times 3 \times 8$
批量标准层 $B_5$	—
最大池化层 $M_3$	池化窗口 $1 \times 1$ , 步长 1
扁平层 $F_1$	—
全连接层 $D_1$	$64 \times 1$
全连接层 $D_2$	$4 \times 1$

## 5 故障诊断流程

故障诊断流程如图 2 所示。该流程主要分为 4 部分:①对原始信号进行 VMD 分解,利用峭度准则筛选分量并重构信号;②计算重构信号的多种特征值,将特征进行融合,构建信号的多维度复合特征矩阵;③利用 KPCA 方法对特征矩阵进行降维,去除



冗余特征;④将得到的低维特征矩阵输入CNN模型中进行训练,对轴承故障进行识别与诊断。

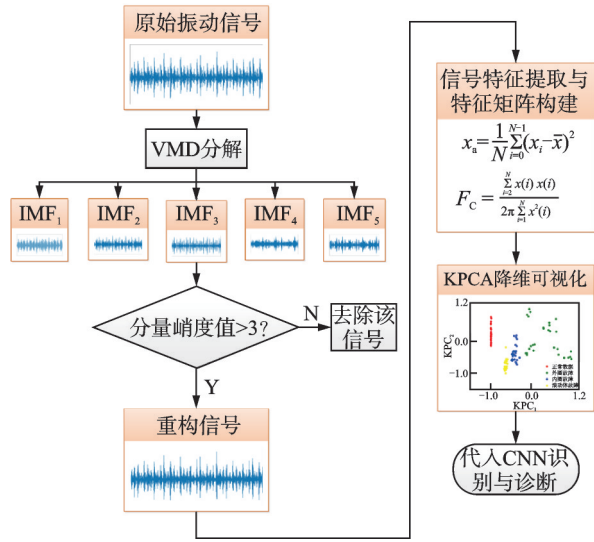


图2 故障诊断流程

Fig.2 Fault diagnosis process

## 6 IMS轴承数据实验分析

### 6.1 数据预处理

IMS轴承实验台如图3所示,本研究实验数据来源于该实验台,实验平台由1个转速为2 000 r/min的电动机、4个轴承和振动传感器组成。

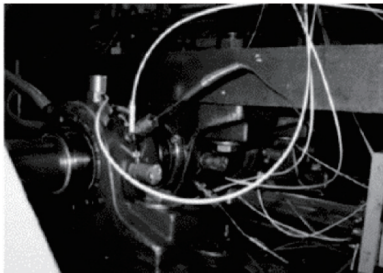


图3 IMS轴承实验台

Fig.3 IMS bearing test bench

IMS轴承数据共包含3组实验数据集,本实验采用数据集1中的内圈故障数据和滚动体故障数据、数据集2中的外圈故障数据以及数据集3中的正常数据进行实验验证。将4类数据分为30组,每组4 000个采样点。其中25组投入训练,剩下5组进行测试。IMS各状态信号时域波形图如图4所示。

以轴承内圈故障为例进行信号分解与重构,获得VMD分解后的5个分量时域图。IMS轴承内圈故障信号分解分量如图5所示。对各个分量进行峭度计算,得知5个分量的峭度值均大于3,所以该信

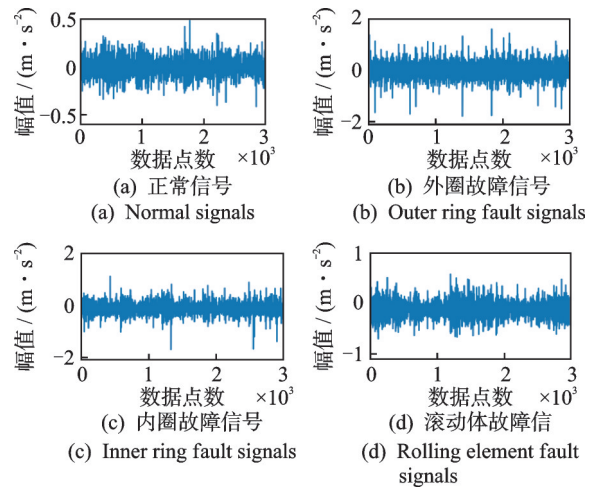


图4 IMS各状态信号时域波形图

Fig.4 Time domain waveform of each state signal in IMS

号无需重构,直接分析原始信号即可。

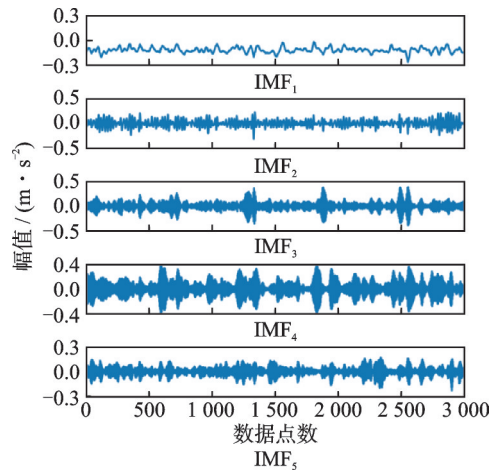


图5 IMS轴承内圈故障信号分解分量

Fig.5 Decomposition components of fault signals for bearing inner ring in IMS

计算重构信号的多个特征值,构建16维复合特征矩阵。对特征矩阵进行KPCA降维处理,再次提取信号关键特征,得到新的二维特征矩阵。IMS-KPCA降维分类可视化二维图如图6所示。

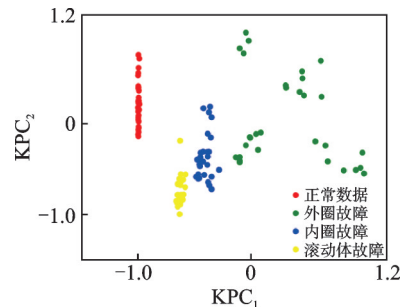


图6 IMS-KPCA降维分类可视化二维图

Fig.6 Two-dimensional diagram of IMS-KPCA dimensional reduction classification visualization

6.2 CNN 模型

实验采用的深度学习框架为 Tensorflow,硬件配置为 Core(TM) i5-8265U CPU 处理器和 NVIDIA GeForce MX230 显卡。

将得到的二维特征矩阵放入 CNN 模型中进行训练。模型采用 Adam 优化器自动优化学习率,并在传统 CNN 的基础上加入 BN 层进行改进,以避免过拟合现象,从而使得结果更加精确。IMS 训练和测试准确率曲线如图 7 所示。IMS 训练和测试损失曲线如图 8 所示。由图 7、8 可知,训练集的准确率在迭代 30 次左右时已完全收敛,达到 100%;损失随着迭代次数的增加而迅速下降,在 200 次左右时完全收敛,无限接近于 0。测试集的准确率在迭代 150 次左右时已完全收敛,达到 100%,损失随着迭代次数增加而逐渐下降,在 300 次左右时完全收敛,无限接近于为 0。

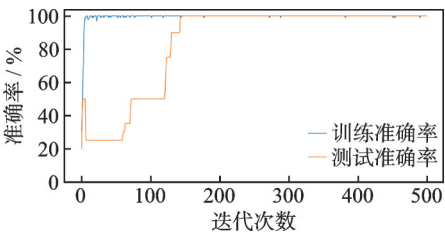


图 7 IMS 训练和测试准确率曲线

Fig.7 Curves of training and testing accuracy in IMS

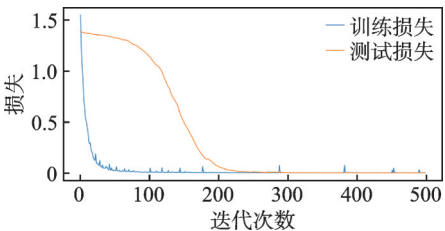


图 8 IMS 训练和测试损失曲线

Fig.8 Curves of training and testing loss in IMS

构建混淆矩阵对最终结果进行分析,IMS 混淆矩阵如图 9 所示。颜色标尺代表预测概率的强度(或置信度),数值范围为 0~1.0,数值越大,颜色越深,表明模型对该预测类别的置信程度越高。由图可知,模型在每种数据类型上的识别准确率均达到 100%,可准确诊断出轴承故障。采用 KPCA 算法分别对经过 3 个最大池化层和最后的全连接层之后的数据分布进行可视化表示,IMS-CNN 可视化效果如图 10 所示,KPC<sub>1</sub> 和 KPC<sub>2</sub> 分别代表通过核主成分分析得到的第 1 主成分和第 2 主成分。由图可知,经过 3 个卷积层和池化层之后,模型提取到的特征逐渐呈现出聚集的趋势,并在全连接层之后达到优异的分类效果。与特征矩阵降维时的可视化图对

比可知,经过神经网络训练后,同类型的特征数据更加聚集,不同类型的特征数据更加分散,说明训练后的分类效果更好。同时,也证明了所提方法对于滚动轴承的故障诊断具有较好的效果。

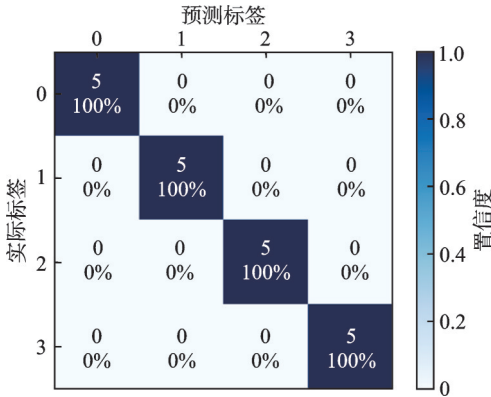


图 9 IMS 混淆矩阵

Fig.9 Confusion matrix in IMS

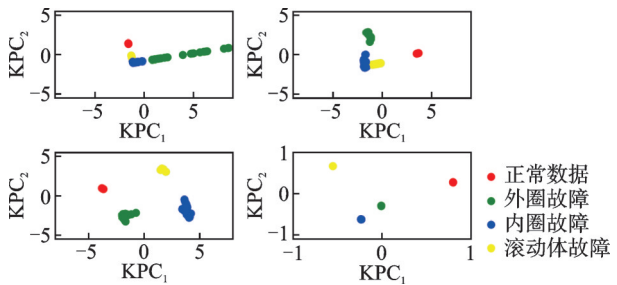


图 10 IMS-CNN 可视化效果

Fig.10 Visualization effect of IMS-CNN

6.3 不同故障诊断方法对比

为验证笔者所提方法的优越性,与长短期记忆(long short term memory,简称 LSTM)、简单循环神经网络(simple recurrent neural network,简称 SimpleRNN)、多层感知机(multilayer perceptron,简称 MLP)和 K 近邻算法(K-nearest neighbors,简称 KNN)等传统典型诊断方法进行对比。在对比过程中,传统典型诊断方法均先采用与本研究相同的预处理方法对原始数据进行处理,再用诊断模型进行分析。对每种方法进行 5 次实验并取平均值,不同诊断方法对比如表 2 所示。

表 2 不同诊断方法对比

Tab.2 Comparison of different diagnostic methods %

诊断方法	训练准确率	测试准确率
所提方法	100	100
LSTM	100	96
SimpleRNN	100	69
MLP	100	47
KNN	28	25

由表可知,所提方法的效果最好,且效果稳定,训练和测试每次均能达到 100% 的准确率;LSTM 算法的平均训练准确率为 100%,个别测试准确率稍低,但平均测试准确率也达到 96%,对故障分类的效果较好,但由于直接对大量的原始数据进行分析,其训练时间较长;SimpleRNN 算法的训练准确率较高,但测试准确率较差,平均值仅为 69%,诊断能力相比所提方法较差;MLP 和 KNN 算法效果均不佳,测试准确率较低,难以达到故障诊断的要求。

### 7 QPZZ-Ⅱ 实验台轴承数据分析

本研究中所用轴承为 6205 型深沟球轴承,实验选取其中的正常数据、内圈外圈混合故障数据、加载情况下滚动体故障数据和加载情况下外圈故障数据进行分析。实验轴承故障程度设置如表 3 所示。实验采样频率为 12 800 Hz。

表 3 实验轴承故障程度设置

Tab.3 Settings of failure degree of experimental bearing	
故障类型	故障程度
内圈故障	切深为 1.5 mm,宽为 0.2 mm
外圈故障	切深为 2 mm,宽为 0.2 mm
滚动体故障	切深为 1 mm,宽为 0.4 mm

实验同样将 4 种状态下的数据分别分为 30 组,每组 4 000 个采样点。其中 25 组投入训练,剩下 5 组进行测试。轴承各状态信号时域波形如图 11 所示。

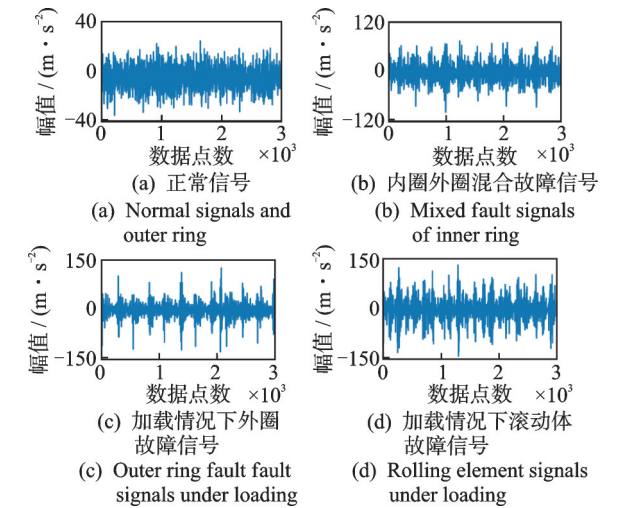


图 11 QPZZ-Ⅱ 轴承各状态信号时域波形图  
Fig.11 Time domain waveform of various state signals of bearing in QPZZ-Ⅱ

图 12 为加载情况下 QPZZ-Ⅱ 轴承外圈故障信号分解分量,展示了以加载情况下的外圈故障为例的信号分解与重构。对各个分量进行峭度计算,得

知 5 个分量的峭度值均大于 3,因而该信号无需重构,直接分析原始信号即可。

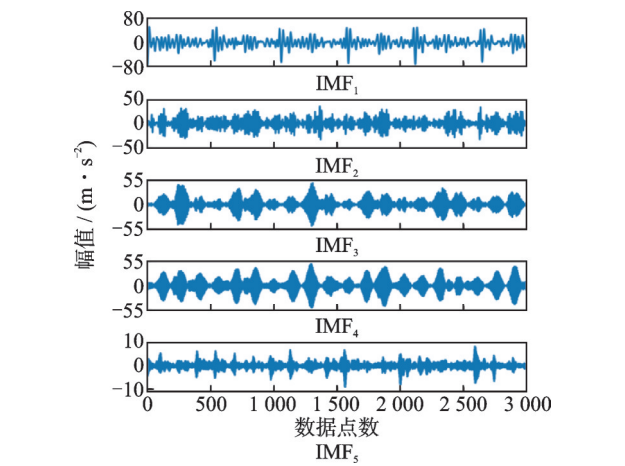


图 12 加载情况下 QPZZ-Ⅱ 轴承外圈故障信号分解分量  
Fig.12 Decomposition components of fault signals for bearing outer ring under load and QPZZ-Ⅱ

计算特征值,构建复合特征矩阵并进行 KPCA 降维,得到新的二维特征矩阵。QPZZ-Ⅱ-KPCA 降维可视化二维图如图 13 所示。

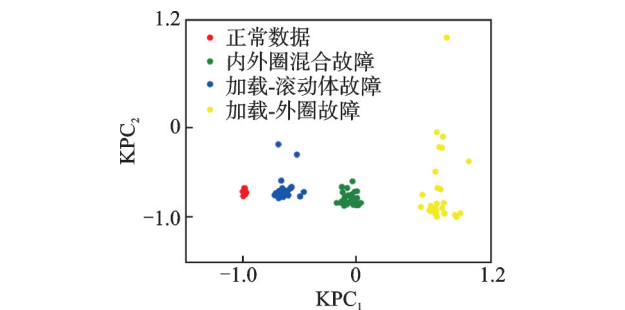


图 13 QPZZ-Ⅱ-KPCA 降维可视化二维图  
Fig.13 Two-dimensional diagram of QPZZ-Ⅱ-KPCA dimensionality reduction visualization

将二维特征矩阵代入 CNN 进行识别与诊断,得到 QPZZ-Ⅱ 训练与测试准确率曲线如图 14 所示, QPZZ-Ⅱ 训练和测试损失曲线如图 15 所示。由图 14、15 可知,训练集的准确率在迭代 15 次左右时已完全收敛,达到 100%,损失随着迭代次数增加而迅速下降,在 200 次左右时完全收敛,无限接近于 0;测试集的准确率在迭代 200 次左右时已完全收敛,达到 100%,损失随着迭代逐渐下降,在 300 次左右时完全收敛,无限接近于 0。

构建混淆矩阵对最终结果进行分析,QPZZ-Ⅱ 混淆矩阵如图 16 所示。由图可知,模型在每种数据类别上的识别准确率均达到 100%,可准确诊断出轴承故障。采用 KPCA 算法分别对经过 3 个最大池化层和最后的全连接层之后的数据分布进行可视化表



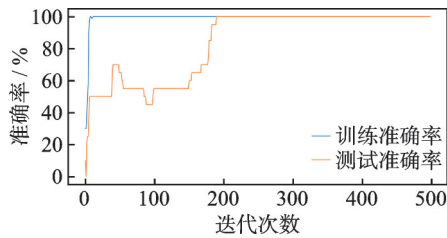


图 14 QPZZ-II 训练和测试准确率曲线

Fig.14 Curves of training and testing accuracy in QPZZ-II

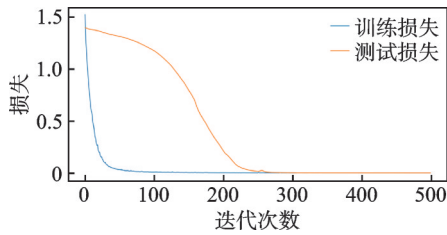


图 15 QPZZ-II 训练和测试损失曲线

Fig.15 Curves of training and testing loss in QPZZ-II

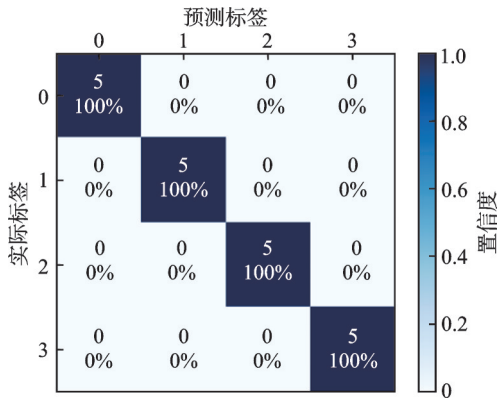


图 16 QPZZ-II 混淆矩阵

Fig.16 Confusion matrix in QPZZ-II

示, QPZZ-II-CNN 可视化效果如图 17 所示。由图可知, 经过 3 个卷积层和池化层之后, 模型提取到的特征逐渐呈现出聚集的趋势, 并在全连接层之后达到极高的分类效果。与前文 KPCA 降维图对比可知, 经过 CNN 训练后, 同类型的特征数据更加聚集, 不同类型的特征数据更加分散, 说明训练过后的分类效果更好。同时, 再次证明了所提方法对滚动轴承的故障诊断具有较好的效果以及较高的准确率。

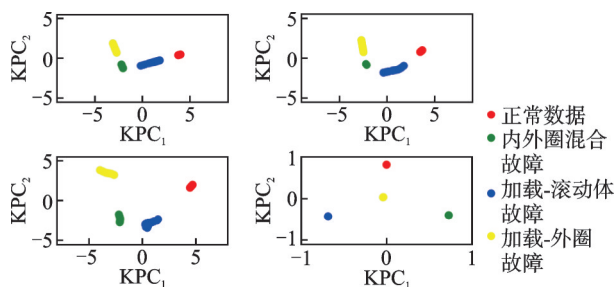


图 17 QPZZ-II-CNN 可视化效果图

Fig.17 Visualization effect of QPZZ-II-CNN

## 8 结 论

1) 针对传统时域、频域统计指标构建的特征矩阵在处理复杂非线性数据时的局限性, 提出了一种基于数据层特征融合和 CNN 模型相结合的诊断方法。

2) 通过 VMD 分解和峭度准则筛选分量, 可以摒弃干扰成分; 从时域、频域、能量和熵值等方面构建特征矩阵, 对于故障信号分析更加全面, 对特征提取能力更强。

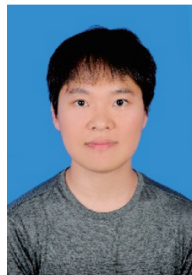
3) 利用 KPCA 方法降维, 提取重要特征, 减少输入神经网络的数据量, 减轻网络的运算压力; 利用 BN 层对神经网络模型进行优化, 加快训练和收敛的速度, 增加模型的稳定性, 防止过拟合。

4) 分别用 IMS 轴承实验数据和 QPZZ-II 实验台轴承数据进行实验, 均取得较好的分类效果和极高的诊断准确率, 证明了所提方法的有效性和优越性。

## 参 考 文 献

- [1] 董绍江, 裴雪武, 吴文亮, 等. 改进抗干扰 CNN 的变负载滚动轴承损伤程度识别[J]. 振动、测试与诊断, 2021, 41(4): 715-722.  
DONG Shaojiang, PEI Xuewu, WU Wenliang, et al. Damage degree identification of rolling bearings under variable load with improved anti-interference CNN[J]. Journal of Vibration, Measurement & Diagnosis, 2021, 41(4): 715-722.(in Chinese)
- [2] 赵小强, 张青青. 改进 Alexnet 的滚动轴承变工况故障诊断方法[J]. 振动、测试与诊断, 2020, 40(3): 472-480.  
ZHAO Xiaoqiang, ZHANG Qingqing. Improved alexnet based fault diagnosis method for rolling bearing under variable conditions[J]. Journal of Vibration, Measurement & Diagnosis, 2020, 40(3): 472-480.(in Chinese)
- [3] 谷玉海, 朱腾腾, 饶文军, 等. 基于 EMD 二值化图像和 CNN 的滚动轴承故障诊断[J]. 振动、测试与诊断, 2021, 41(1): 105-113.  
GU Yuhai, ZHU Tengting, RAO Wenjun, et al. Fault diagnosis for rolling bearing based on EMD binarization image and CNN[J]. Journal of Vibration, Measurement & Diagnosis, 2021, 41(1): 105-113.(in Chinese)
- [4] 金江涛, 许子非, 李春, 等. 基于卷积双向长短期记忆网络与混沌理论的滚动轴承故障诊断[J]. 振动与冲击, 2022, 41(17): 160-169.  
JIN Jiangtao, XU Zifei, LI Chun, et al. Fault diagnosis of rolling bearing based on CCNN-BiLSTMN method[J]. Journal of Vibration and Shock, 2022, 41(17): 160-169.(in Chinese)
- [5] 童靳于, 罗金, 郑近德. 基于增强深度自编码网络的

- 滚动轴承故障诊断方法[J]. 中国机械工程, 2021, 32(21): 2617-2624.
- TONG Jinyu, LUO Jin, ZHENG Jinde. Rolling bearing fault diagnosis method based on enhanced deep auto-encoder network [J]. China Mechanical Engineering, 2021, 32(21): 2617-2624.(in Chinese)
- [6] 游达章, 陈林波, 张业鹏, 等. 基于PCA-CNN的滚动轴承故障诊断方法研究[J]. 机床与液压, 2021, 49(19): 172-177.
- YOU Dazhang, CHEN Linbo, ZHANG Yepeng, et al. Study on fault diagnosis methods of rolling bearing based on principal component analysis and convolutional neural network[J]. Machine Tool & Hydraulics, 2021, 49(19): 172-177.(in Chinese)
- [7] 张继旺, 丁克勤, 王洪柱. 基于VMD-CNN的滚动轴承早期微弱故障智能诊断方法[J]. 组合机床与自动化加工技术, 2020(11): 15-19.
- ZHANG Jiwang, DING Keqin, WANG Hongzhu. Rolling bearing early weak fault intelligent diagnosis based on VMD and CNN[J]. Modular Machine Tool & Automatic Manufacturing Technique, 2020(11): 15-19.(in Chinese)
- [8] 李思琦, 蒋志坚. 基于EEMD-CNN的滚动轴承故障诊断方法[J]. 机械强度, 2020, 42(5): 1033-1038.
- LI Siqi, JIANG Zhijian. Fault diagnosis method of rolling bearing based on EEMD-CNN[J]. Journal of Mechanical Strength, 2020, 42(5): 1033-1038.(in Chinese)
- [9] 姚峰林, 谢长开, 吕世宁, 等. 基于小波包变换和ELM的滚动轴承故障诊断研究[J]. 安全与环境学报, 2021, 21(6): 2466-2472.
- YAO Fenglin, XIE Changkai, LÜ Shining, et al. Research on fault diagnosis of rolling bearings based on wavelet packet transform and ELM[J]. Journal of Safety and Environment, 2021, 21(6): 2466-2472.(in Chinese)
- [10] 何江江, 李孝全, 赵玉伟, 等. 基于改进EEMD的卷积神经网络滚动轴承故障诊断[J]. 重庆大学学报, 2020, 43(1): 82-89.
- HE Jiangjiang, LI Xiaoquan, ZHAO Yuwei, et al. Fault diagnosis of rolling bearing based on improved EEMD and convolutional neural network[J]. Journal of Chongqing University, 2020, 43(1): 82-89.(in Chinese)
- [11] 李子国, 石晴, 刘继超, 等. 基于参数优化VMD和1D-CNN的滚动轴承故障诊断[J]. 现代信息科技, 2022, 6(16): 66-70.
- LI Zigu, SHI Qing, LIU Jichao, et al. Rolling bearing fault diagnosis based on parameter optimization VMD and 1D-CNN[J]. Modern Information Technology, 2022, 6(16): 66-70.(in Chinese)
- [12] 吕阳, 廖与禾, 王报祥, 等. 基于VMD和CNN的滚动轴承故障定量诊断方法[J]. 中国科技论文, 2020, 15(7): 735-742.
- LÜ Yang, LIAO Yuhe, WANG Baoxiang, et al. Quantitative diagnosis method for rolling bearing faults based on VMD and CNN[J]. China Sciencepaper, 2020, 15(7): 735-742.(in Chinese)
- [13] 余志锋, 熊邦书, 熊天旻, 等. 基于VMD-CWT和改进CNN的直升机轴承故障诊断[J]. 航空动力学报, 2021, 36(5): 948-958.
- YU Zhifeng, XIONG Bangshu, XIONG Tianyang, et al. Fault diagnosis of helicopter bearing based on VMD-CWT and improved CNN[J]. Journal of Aerospace Power, 2021, 36(5): 948-958.(in Chinese)
- [14] 廖晨, 王艳丰, 李舜韶, 等. 基于VMD-GST和AMCNN相结合的滚动轴承故障诊断方法[J]. 轴承, 2023(2): 82-88, 96.
- LIAO Chen, WANG Yanfeng, LI Shunming, et al. Fault diagnosis method for rolling bearings based on VMD-GST and AMCNN[J]. Bearing, 2023(2): 82-88, 96.(in Chinese)
- [15] 朵慕社, 纪国宜, 朱海龙, 等. 基于VMD降噪和CNN的轴承故障诊断[J]. 噪声与振动控制, 2021, 41(5): 155-160.
- DUO Mushe, JI Guoyi, ZHU Hailong, et al. Bearing fault diagnosis based on VMD noise reduction and CNN [J]. Noise and Vibration Control, 2021, 41(5): 155-160.(in Chinese)
- [16] 吴东升, 杨青, 张继云, 等. 基于VMD-LMD-CNN的集合型故障诊断方法[J]. 轴承, 2020(10): 57-63.
- WU Dongsheng, YANG Qing, ZHANG Jiyun, et al. Ensemble fault diagnosis method based on VMD-LMD-CNN[J]. Bearing, 2020(10): 57-63.(in Chinese)
- [17] 吴楠, 王喆. 基于一维CNN与BiLSTM结合的轴承故障诊断[J]. 组合机床与自动化加工技术, 2021(9): 38-41, 45.
- WU Nan, WANG Zhe. Bearing fault diagnosis based on the combination of one-dimensional CNN and BiLSTM[J]. Modular Machine Tool & Automatic Manufacturing Technique, 2021(9): 38-41, 45.(in Chinese)



**第一作者简介:**张雄,男,1990年5月生,博士、副教授、硕士生导师。主要研究方向为旋转机械状态监测与故障诊断、振动/声学信号分析与处理。曾发表《A bearing fault diagnosis method based on multi-scale dispersion entropy and GG clustering》(《Measurement》2021, Vol. 185)等论文。

E-mail: hdjxzx@ncepu.edu.cn

**通信作者简介:**李嘉禄,男,1997年7月生,硕士生。主要研究方向为滚动轴承故障诊断。

E-mail: ljl2481390881@163.com