

基于改进 CNN 和 Kmeans 的双转子轴承半监督故障诊断*

崔锦森¹, 贺雅^{1,2}, 冯坤^{1,2}

(1.北京化工大学发动机健康监控及网络化教育部重点实验室 北京,100029)

(2.北京化工大学高端机械装备健康监控与自愈化北京市重点实验室 北京,100029)

摘要 针对双转子轴承复杂信号故障特征难提取、工程中某些类型的故障数据缺乏时卷积神经网络(convolution neural network,简称 CNN)难训练的问题,提出一种基于改进 CNN 和 Kmeans 的双转子轴承半监督故障诊断方法。首先,利用自回归(autoregressive,简称 AR)模型对双转子轴承信号去噪,并基于傅里叶变换得到信号频谱作为 CNN 输入;其次,以 Morlet 小波基频域函数作为激活函数构建 CNN,结合 Softmax 损失和提出的权重内积最小化损失在少类别训练数据下训练 CNN;最后,基于 Kmeans 聚类算法分析 CNN 线性输出确定无标签数据伪标签,并结合半监督学习中自训练思想迭代 CNN 更新伪标签,继而依据伪标签划分双转子轴承正常、已知故障和未知故障状态。利用双转子轴承故障模拟试验台数据进行验证,结果表明,在少类别训练数据下,其诊断效果相较于人工神经网络(artificial neural network,简称 ANN)等方法更佳,诊断准确率达到 100%,验证了所提方法的有效性。

关键词 双转子轴承;故障诊断;卷积神经网络;自回归模型;Kmeans 聚类;半监督学习

中图分类号 TH133.33;TP277

引言

双转子轴承是一种特殊的滚动轴承,其内外圈均与转子相连,以减轻重量获得紧凑结构,被广泛用于燃气轮机等双转子设备中^[1]。作为双转子设备关键支撑部件之一,双转子轴承的可靠性直接影响设备的稳定健康运行,因此对其进行状态监测和故障诊断十分必要,其中振动监测是最常用的手段^[2-3]。

双转子轴承内外圈均转动,无固定轴承座,振动信号只能在设备机匣间接采集,因此其振动信号相较于普通滚动轴承更为复杂。艾延廷等^[4]利用局部包络谱峰值因子优化 Morlet 复小波,然后对双转子轴承振动信号滤波,提取双转子轴承微弱故障。Jiang 等^[5]利用 AR 对双转子轴承振动信号进行去噪,基于快速谱峭度方法确定中心频率和带宽,并对信号再次滤波,共振解调得到包络信号,最后对包络信号角域重采样,提取双转子轴承阶次特征。上述方法虽然在一定程度上增强了双转子轴承微弱故障特征,但严重依赖专家诊断知识,智能化水平低。为了减轻故障诊断对信号处理技术及专家诊断知识的依赖,提高故障诊断智能化水平,缩短诊断维修周期, CNN 等端到端的深度学习技术逐渐应用于故障诊断领域^[6]。王奉涛等^[7]提出一种基于 CNN 的双转子轴承故障诊断方法,对轴承振动信号进行灰度变换

后输入 CNN 进行诊断。王丽华等^[8]以短时傅里叶变换后的时频谱图作为 CNN 输入来进行诊断故障。

基于 CNN 的有监督诊断方法,需要一定数量的各类轴承故障数据。工程实际中,轴承某些故障类型数据不易获取,限制了上述方法的工程应用。半监督深度学习利用标签数据和无标签数据共同训练网络,一定程度上减弱了网络训练对标签数据的依赖。Li 等^[9]提出基于数据增强和深度稀疏自动编码器的半监督故障诊断方法,减小了网络训练对有标签数据数量的要求,但是网络微调仍需各类别的故障数据支撑。Lee^[10]提出伪标签自训练的半监督深度学习方法,在一定程度上摆脱了测试集类别数不能大于训练集类别数的限制。

笔者提出一种基于改进 CNN 和 Kmeans 的双转子轴承半监督故障诊断方法。该方法基于 AR 预处理双转子轴承振动信号,以频谱幅值序列作为改进 CNN 输入,基于自训练思想在少类别训练数据下诊断双转子轴承故障,减小了网络训练对有标签故障数据的依赖,工程应用价值更高。

1 方法原理

1.1 自回归模型

AR 是一种常见的滚动轴承振动信号去噪预处理

* NSFC-辽宁联合基金重点资助项目(U1708257);国家重点研发计划资助项目(2017YFC0805702)

收稿日期:2021-03-23;修回日期:2021-05-30

理方法^[5],可预测振动信号中平稳信号成分。从原始振动信号中剔除平稳信号成分即可获得滚动轴承非平稳振动信号,其原理如下。

通过 t 时刻前 p 个数据预测 $x(t)$ 数据值,即

$$x'(t) = - \sum_{i=1}^p a_i x(t-i) \quad (1)$$

其中: $x(t-i)$ 为 t 时刻前第 i 个数据值; a_i 为 $x(t-i)$ 对应的系数; $x'(t)$ 为 t 时刻预测值。

$x'(t)$ 为平稳信号成分,从原始信号中剔除平稳信号成分即可获得轴承非平稳冲击信号

$$e(t) = x(t) - \hat{x}(t) = x(t) + \sum_{i=1}^p a_i x(t-i) \quad (2)$$

展开式(2)可得

$$e(t) = x(t) + a_1 x(t-1) + a_2 x(t-2) + \dots + a_p x(t-p) \quad (3)$$

$e(t)$ 可以看成滤波器系数 $a(p) = \{1, a_1, \dots, a_p\}$ 在时间序列 $x(t) = \{x(t), x(t-1), \dots, x(t-p)\}$ 上的卷积,其中 AR 滤波器系数 $a(p)$ 可通过 Yule-Walker 方程求解^[5]。

1.2 权重归一化卷积神经网络

卷积神经网络是一种集特征提取与分类于一体的深度学习方法,其兴起于计算机视觉和语音识别领域,并逐步应用在设备智能故障诊断中^[11]。卷积神经网络随输入数据维度可分为 1DCNN 和 2DCNN,其中 1DCNN 更适用于振动波形和频谱等 1 维输入数据序列。典型的 1DCNN 结构示意图如图 1 所示,包括卷积层、池化层和全连接层。为加速网络收敛,Salimans 等^[12]提出权重归一化方法加速网络训练过程中权重调整。Jia 等^[13]将权重归一化卷积神经网络应用于故障诊断领域,进一步验证该方法的有效性。权重归一化卷积神经网络通过约束卷积核和全连接层权重模长,加速随机梯度下降算法对卷积核和全连接层权重在空间方向上的调整,从而加速网络收敛,其原理如下。

1) 卷积层:利用卷积核在输入上滑动卷积自动

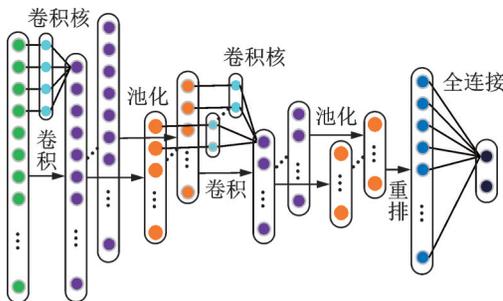


图1 典型的1DCNN结构示意图

Fig.1 Structure of typical one dimensional CNN

提取数据特征,通过局部连接实现权值共享,相较于全连接结构大大减小了网络参数数量。其中,一个卷积核作用于输入得到的特征称为一个特征图,其前向计算公式为

$$y_i^k = f_c(\gamma^k \sum_{j=1}^n \sum_{l=1}^m \frac{\omega_{j,l}^k}{\|\omega^k\|_2} x_{j,l+\text{step}*(i-1)} + b^k) \quad (4)$$

其中:step为卷积层的步长; ω^k 为第 k 个 $n \times m$ 的卷积核; $\|\omega^k\|_2$ 为第 k 个卷积核的欧几里得范数; γ^k 为卷积核 ω^k 对应的偏移系数; n 为卷积层输入特征图的个数(第 1 层卷积层中 $n=1$); m 为卷积核长度; b^k 为第 k 个卷积核对应的偏置; $f_c(\cdot)$ 为卷积层的激活函数; y_i^k 为第 k 个特征图中第 i 个节点的输出。

2) 池化层:通过下采样约减数据特征维度,提高网络的鲁棒性。常用的池化方式有最大池化和平均池化,本研究选取的平均池化前向计算公式为

$$y_i^k = \sum_{r=1}^{\text{step}} x_{k,r+\text{step}*(i-1)} / \text{step} \quad (5)$$

其中:step为池化层的步长; x_k 为池化层第 k 个输入特征图; y_i^k 为池化层第 k 个特征图第 i 个节点的输出。

池化层输入特征图个数等于输出特征图个数。

3) 全连接层:通过全连接结构将网络学习的分布式特征映射到样本标记空间,起到分类器的作用,其前向计算公式为

$$y_i = f_f(\gamma_i \sum_{j=1}^m \frac{\omega_{i,j}}{\|\omega_i\|_2} x_j + b_i) \quad (6)$$

其中: $\omega_{i,j}$ 为全连接层第 i 个输出节点与第 j 个输入节点的连接权重; ω_i 为所有输入节点与第 i 个输出节点的权重向量; $\|\omega_i\|_2$ 为 ω_i 的欧几里得范数; γ_i 为 ω_i 对应的偏移系数; x_j 为第 j 个输入节点值; m 为输入节点个数; b_i 为偏置; $f_f(\cdot)$ 为激活函数; y_i 为第 i 个节点输出值。

1.3 基于自训练的半监督深度学习

故障诊断领域故障数据获取困难,半监督深度学习相较于有监督学习对标签训练数据依赖更小。基于自训练的半监督深度学习示意如图 2 所示,其基于少量有标签数据预训练网络,利用大量无标签数据进行测试,并生成无标签数据的伪标签,结合标签和伪标签数据迭代网络,摆脱了测试集类别数不能大于训练集类别数的限制,应用价值更高^[10,14]。如何确定无标签数据的伪标签是算法关键,聚类分析是常用方法之一。Kmeans 算法是一种典型的利用统计方法来分析数据分布特性的无监督学习算法,通过 Kmeans 分析 CNN 线性输出,可初步判断

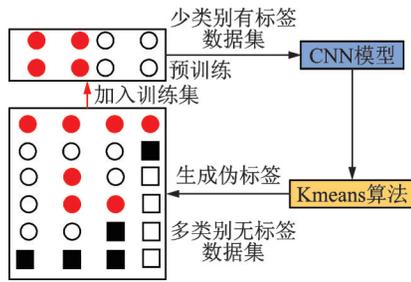


图 2 基于自训练的半监督深度学习示意图

Fig.2 Schematic diagram of semi-supervised deep learning based on self-training

无标签数据对应的伪标签,其聚类数 K 可通过轮廓系数 (silhouette coefficient, 简称 SC) 来确定^[15]。

SC 是衡量聚类效果的重要指标之一,单个样本的 SC 计算公式为

$$sc_i = (a_i - b_i) / \max(a_i, b_i) \quad (7)$$

$$b_i = \frac{1}{N_k - 1} \sum_{j=1}^{N_k-1} \|x_i - x_j\| \quad (x_i \in c_k, x_j \in c_k) \quad (8)$$

$$a_i = \frac{1}{M - N_k} \sum_{j=1}^{M-N_k} \|x_i - x_j\| \quad (x_i \in c_k, x_j \notin c_k) \quad (9)$$

其中: M 为样本总数; b_i 为第 i 个样本与同类样本的平均距离; a_i 为第 i 个样本与不同类样本的平均距离。

样本集合的 SC 为所有样本 SC 的均值,即

$$SC = \frac{1}{M} \sum_{i=1}^M sc_i \quad (10)$$

SC 越大,则表示数据类内紧凑,类间距离大,聚类效果越好。通过比较各 K 值下 SC 的大小,以 SC 随 K 变化的转折点作为最佳聚类数 K 值。

2 本研究提出的方法

2.1 改进 CNN

为进一步提升 CNN 从双转子轴承数据中提取复杂非线性特征的能力,笔者从激活函数和损失函数两方面改进 CNN。从故障诊断角度选取非线性程度高且与 CNN 输入更相关的函数作为激活函数,从 CNN 内积运算的数学角度约束权重向量以增加权重多样性。

2.1.1 Morlet 小波频域激活函数

激活函数通过非线性激励神经网络,可拟合出任何复杂的非线性关系。各激活函数及导数如图 3 所示,其中 Sigmoid 和 Tanh 因导数计算简单被广泛应用,但其响应区间小,输出范围窄,仅在 $[0, 1]$ 和 $[-1, 1]$ 范围内输出,且随着网络层数的加深易造成梯度弥散,增加了网络训练难度^[16]。Relu 和 Leaky-

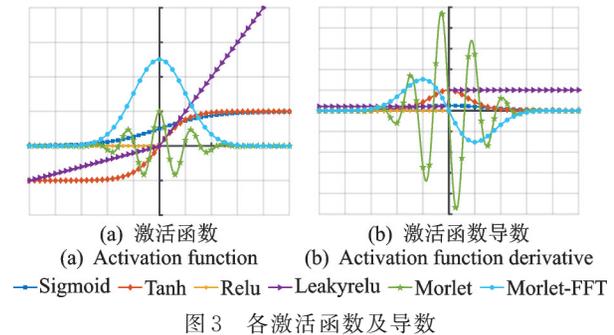


图 3 各激活函数及导数

Fig.3 Activation function and derivative curve

relu 激活函数在一定程度上解决了 Sigmoid 和 Tanh 的梯度弥散问题,但非线性程度低,仅在正负半轴具备不同斜率,不利于提取双转子轴承振动信号中的复杂非线性特征。文献[17]在以时域波形作为输入时,以非线性更强的 Morlet 小波时域函数作为激活函数,增强了网络非线性特征提取能力,提升了故障诊断效果。

笔者在以频谱幅值序列作为 CNN 输入时,相应的以 Morlet 小波频域函数 (Morlet wavelet frequency domain function, 简称 Moret-FDF) 作为激活函数。

1) 频谱中包含有部件故障特征频率及特征频段等相关信息,相较于时域波形故障特征更直观,且数据内存小 (一般情况下只考虑幅值谱,为时域的一半)。Xu 等^[18]将图像变换至频域,以频域特征作为输入,相较于直接以空间域图像作为输入,在相同精度下大幅降低计算成本,验证了频谱特征的有效性。

2) Morlet 及 Morlet-FDF 激活函数计算公式如表 1 所示。 $a=1, b=0$ 时的激活曲线及导数见图 3。相较于传统激活函数,该函数具有以下优点:①是 Morlet 小波的频域响应,非线性程度更高,且曲线为单瓣,与冲击信号频谱最相似,频谱中主要信息位于窗函数效应的主瓣中,该种响应曲线可能更契合频谱输入;②是可变尺度自适应激活函数,自适应平移系数 b 通过调节激活函数的响应区间保留频谱有效频段特征,自适应缩放因子 a 通过调节激活函数响应宽度及平滑度放大有效频段特征,更有利于 CNN

表 1 常用激活函数计算公式

Tab.1 Calculation formula of activation function and its derivative

激活函数	计算式
Morlet	公式 $f(x) = \cos(f_0 x) e^{-\frac{x^2}{2}}$
	导数 $f'(x) = -(f_0 \sin(f_0 x) + x \cos(f_0 x)) e^{-\frac{x^2}{2}}$
Morlet-FDF	公式 $f(x) = \sqrt{2\pi} e^{-\frac{(x-b)^2}{2a}}$
	导数 $f'(x) = -\sqrt{2\pi} \left(\frac{x-b}{a}\right) e^{-\frac{(x-b)^2}{2a}}$

自提取故障相关特征。

2.1.2 权重内积最小化损失

Ayinde 等^[19]通过设定卷积核间余弦相似度阈值,建立卷积核间相似度惩罚机制,提高网络各层特征提取的多样性,但卷积核间余弦相似度惩罚阈值难以设定,阈值过大达不到效果提升,阈值过小则易欠拟合。笔者结合网络权重归一化约束^[12-13],在文献[19]基础上提出网络权重内积最小化损失。权重归一化约束通过规范权重的模长加速单位权重向量在空间方向上的调整,权重内积最小化损失则通过惩罚单位权重向量内积,增大各单位权重向量在空间上的角度,使不同权重向量提取的特征方向尽可能不同,以增强 CNN 特征提取的多样性。

卷积神经网络权重内积最小化损失计算公式为

$$L_{\omega} = \frac{1}{2} \sum_l L^l = \frac{1}{2} \sum_l \left(\sum_{i=1}^N \sum_{j=1}^N L^l_{i,j,i \neq j} \right) \quad (11)$$

其中: l 为卷积层或全连接层层数; N 为 l 层中卷积核或权重向量个数; $L^l_{i,j}$ 为 l 层第 i 和第 j 个卷积核或权重向量间内积; L^l 为 l 层权重内积损失。

对于卷积层则有

$$L^l_{i,j,i \neq j} = \left\langle \frac{\mathbf{w}^i}{\|\mathbf{w}^i\|_2} \cdot \frac{\mathbf{w}^j}{\|\mathbf{w}^j\|_2} \right\rangle \quad (12)$$

$$\frac{\partial L_{\omega}}{\partial \mathbf{w}^i} = \frac{\partial L^l}{\partial \mathbf{w}^i} = \sum_{j=1, j \neq i}^N \frac{\mathbf{w}^j}{\|\mathbf{w}^i\| \|\mathbf{w}^j\|} - \frac{\mathbf{w}^i}{\|\mathbf{w}^i\|^2} L^l_{i,j} \quad (13)$$

对于全连接层则有

$$L^l_{i,j,i \neq j} = \left\langle \frac{\mathbf{w}^i}{\|\mathbf{w}^i\|_2} \cdot \frac{\mathbf{w}^j}{\|\mathbf{w}^j\|_2} \right\rangle \quad (14)$$

$$\frac{\partial L_{\omega}}{\partial \mathbf{w}_i} = \frac{\partial L^l}{\partial \mathbf{w}_i} = \sum_{j=1, j \neq i}^N \frac{\mathbf{w}_j}{\|\mathbf{w}_i\| \|\mathbf{w}_j\|} - \frac{\mathbf{w}_i}{\|\mathbf{w}_i\|^2} L^l_{i,j} \quad (15)$$

选取 Softmax 损失函数,网络最终损失函数为

$$L = L_s + \lambda L_{\omega} = - \sum_{i=1}^M t_i \ln(y_i) + \frac{\lambda}{2} \sum_l \left(\sum_{i=1}^N \sum_{j=1}^N L^l_{i,j,i \neq j} \right) \quad (16)$$

其中: λ 为 L_s, L_{ω} 损失间的系数; M 为训练样本数。 λ 可自适应给定,即

$$\lambda_l = \frac{\sum |d^l_{L_s}|}{\sum |d^l_{L_{\omega}}|} \quad (17)$$

其中: λ_l 为第 l 层 λ 参数; $\sum |d^l_{L_s}|$ 为第 l 层 L_s 损失对应梯度绝对值之和; $\sum |d^l_{L_{\omega}}|$ 为第 l 层 L_{ω} 损失对应梯度绝对值之和。

通过系数 λ_l 保证第 l 层 L_s 损失和 L_{ω} 损失梯度数量级接近,从而保证网络收敛。

2.2 基于改进 CNN 和 Kmeans 的双转子轴承半监督故障诊断方法

本研究以双转子轴承为研究对象,提出一种基于改进 CNN 和 Kmeans 的半监督故障诊断方法,如

图 4 所示,其基本步骤如下。

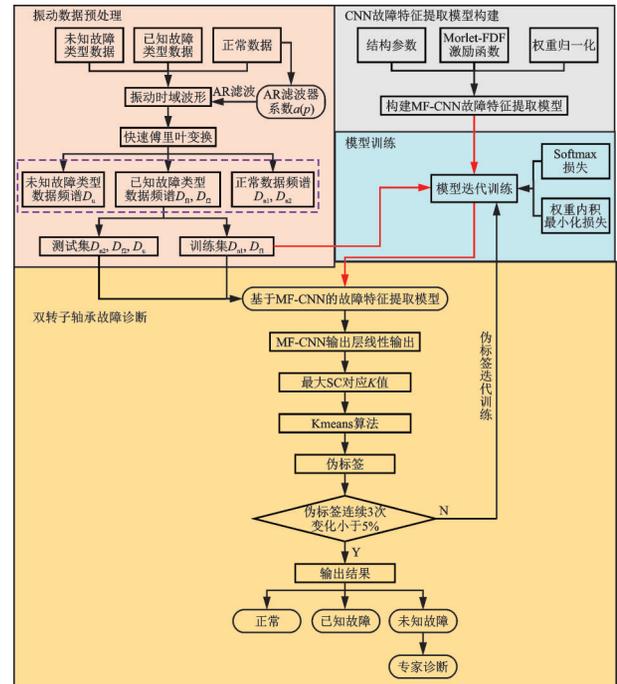


图 4 基于改进 CNN 和 Kmeans 的半监督故障诊断方法

Fig.4 The semi-supervised fault diagnosis method based on improved CNN and Kmeans

1) 振动数据预处理:利用双转子轴承正常状态下振动数据求解 AR 滤波器系数 $a(p)$,基于 $a(p)$ 对所有振动数据滤波去噪。对滤波后的时域波形进行傅里叶变换,得到 1 维频谱序列,并将数据划分为训练集和测试集。其中:训练集包含有标签正常数据 D_{n1} 和有标签故障数据 D_{n1} ;测试集包含正常数据 D_{n2} 、已知故障类型数据 D_{t2} 和未知故障类型数据 D_u 。

2) CNN 故障特征提取模型构建:根据样本频谱序列长度确定 1DCNN 结构参数,基于提出的 Morlet-FDF 激活函数及权重归一化方法构建 MF-CNN 故障特征提取模型。

3) 模型预训练:基于有标签的训练集数据 D_{n1} 和 D_{n1} ,结合提出的权重内积最小化和 Softmax 损失,迭代训练 MF-CNN 故障特征提取模型。

4) 双转子轴承故障诊断:将训练集数据 D_{n1}, D_{n1} 和测试集数据 D_{n2}, D_{t2}, D_u 输入 MF-CNN,得到 MF-CNN 线性输出特征。基于 Kmeans 算法对 MF-CNN 输出特征聚类,计算各聚类数 $K(K \geq k, k$ 为训练集数据类别)下的 SC 值,选取最大 SC 值对应的 K 值作为最终聚类数,得到最终聚类结果,即无标签数据伪标签。利用有标签训练数据和伪标签测试数据共同迭代网络,直至伪标签连续 3 次变化小于 5%,即认为伪标签最吻合测试集真实标签,从而判别测试集样本状态。

当模型诊断出未知故障时,通过专家诊断确定故障类别后,可对未知故障数据重新标记以扩充故障数据库。

3 试验验证

3.1 试验台介绍

双转子轴承故障模拟试验台如图 5 所示,利用相关数据验证所提方法的有效性。试验台主要由电主轴(外)、联轴器、轴承座、双转子轴承、普通滚动轴承、轴承座、联轴器及电主轴(内)组成。由于双转子轴承内外圈不固定,无法直接安装传感器,因此只能在临近的#2轴承座上安装振动加速度传感器。振动信号通过转子传递至#2轴承,然后传递至 2#轴承座,传递路径相较于普通滚动轴承更为复杂。

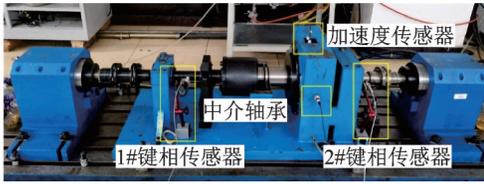


图 5 双转子轴承故障模拟试验台

Fig.5 The fault simulation test rig of the dual-rotor bearing

以 LMS Test. Lab 振动测试系统采集双转子轴承在正常(N)、外圈故障(OF)、内圈故障(IF)及滚动体故障(RF)下轴承座处外侧垂直方向的加速度信号,其中加速度传感器型号为 BK4519,采样频率为 25.6 kHz。选取内圈转速为 900 r/min、外圈转速为 1 500 r/min 工况下的轴承数据组成验证数据集,为模拟实际中有标签故障数据类别缺乏现象,选取正常数据、某一类轴承故障数据或干扰噪声数据(IN)构成训练集,正常数据和所有轴承故障数据构成测试集,数据集组成如表 2 所示。

表 2 数据集组成

Tab.2 Composition of data

训练集数据类型	训练集数据样本量/组	测试集数据类型	测试集数据样本量/组
N+IN	20+20	N+IN+	20+20+
N+OF		OF+IF+RF	20+20+20
N+IF		N+OF+	20+20+
N+RF		IF+RF	20+20

3.2 数据去噪预处理

冲击调制是滚动轴承发生局部故障的重要特征,滚动体转过滚动轴承缺陷处会产生瞬时冲击,同

时激起轴承固有振动产生幅值调制现象,将低频滚动轴承故障特征频率调制至高频段,因此故障特征主要集中在加速度频谱高频段。为验证 AR 预处理方法去除双转子轴承中低频干扰噪声的有效性,以高频段能量占频谱全频带的大小作为评价指标,对比原始信号、AR 处理、小波软阈值降噪处理以及经验模态分解(empirical mode decomposition,简称 EMD)处理后频谱中高频段能量占比。其中高频段能量占比计算公式为

$$\eta = \sum_{f_i \in (f_c, +\infty)} A_{f_i}^2 / \sum A_f^2 \quad (18)$$

其中: A_f 为信号频谱幅值序列; f_i 为高频段截止频率。

处理后频谱中高频能量占比见图 6,处理过程如下:①依据 2.2 节中步骤 1 进行 AR 处理,其中预测阶次为 100^[5];②小波软阈值降噪,分解层数为 4 层,小波基取与冲击信号最相似的 sym8 小波^[20];③进行 EMD 分解,取与原始信号相关系数最大的本征模态分量分析^[21]。取 $f_i=8$ kHz,可以看出,AR 处理后故障数据高频能量均占比相较于其他处理方法增大最明显,且故障数据间高频段能量占比差异更大,但外圈和滚动体故障间差异仍不明显,再次证明了双转子轴承信号的强噪声干扰性。

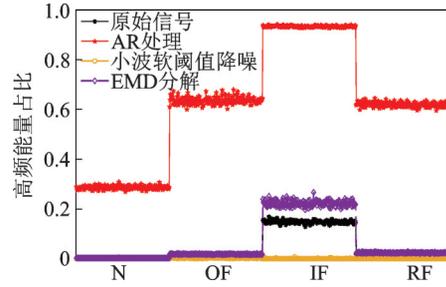


图 6 高频能量占比

Fig.6 Proportion of high frequency energy in spectrum

图 7 为内外圈转速为 900 和 1 500 r/min 时双转子轴承各状态原始数据及经 AR 滤波处理后数据的频谱,可以看出,AR 滤波处理抑制了故障数据低频平稳干扰信号,突出了高频成分,增强了双转子轴承微弱故障特征。

3.3 CNN 模型构建与训练

为使各样本中充分包含故障信息,本研究截取 0.64 s(时域波形为 16 384 点,频谱序列为 8 192 点)振动数据时域波形作为一个样本,以频谱序列作为 CNN 输入,并结合图 1 所示的网络结构构建 CNN,网络结构参数见表 3,其中 c 为训练数据类别数。

从激活函数、CNN 输入及损失函数 3 个方面对比验证本研究所提方法有效性。首先,以 Sigmoid, Tanh, Relu, Leakyrelu, Morlet(取 $f_0=5$ ^[17])及 Morlet-

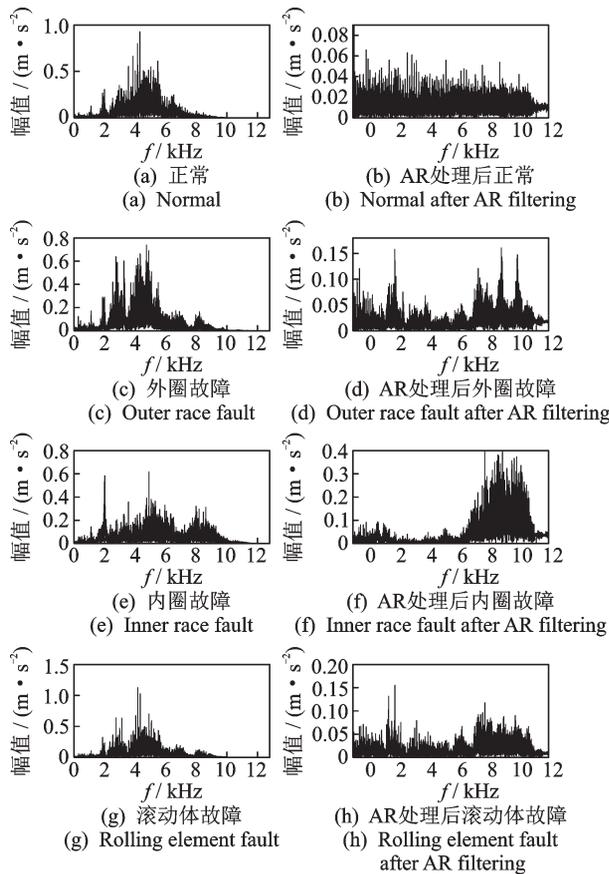


图 7 双转子轴承频谱
Fig.7 Data spectrum of dual-rotor bearing

表 3 网络结构参数

Tab.3 Parameter of CNN

层数	名称	核尺寸	步长	输出尺寸
L_1	输入层	-	-	8192×1
L_2	卷积层	$93 \times 1 \times 6$	1	$8100 \times 1 \times 6$
L_3	池化层	4×1	4	$2025 \times 1 \times 6$
L_4	卷积层	$76 \times 6 \times 8$	1	$1950 \times 1 \times 8$
L_5	池化层	2×1	2	$975 \times 1 \times 8$
L_6	卷积层	$50 \times 8 \times 10$	1	$926 \times 1 \times 10$
L_7	池化层	2×1	2	$463 \times 1 \times 10$
L_8	卷积层	$40 \times 10 \times 12$	1	$424 \times 1 \times 12$
L_9	池化层	2×1	2	$212 \times 1 \times 12$
L_{10}	卷积层	$1 \times 12 \times 1$	1	$212 \times 1 \times 1$
L_{11}	全连接层	$212 \times c$	-	$c \times 1$

FDF 激活函数构建 S-CNN, T-CNN, R-CNN, L-CNN, M-CNN 及 MF-CNN 网络模型;其次,分别以原始频谱数据和经 AR 处理后的频谱数据作为 CNN 输入,利用表 2 中的训练集数据在 Softmax 损失 (J_1 损失)、结合权重内积最小化的 Softmax 损失 (J_2 损失)下预训练各网络模型;最后,基于训练数据和测试数据,结合自训练思想迭代训练网络并测试网络效果。网络迭代参数如表 4 所示。

表 4 网络迭代参数

Tab.4 Iteration parameter of CNN

批量数/组	学习率	训练终止误差	梯度下降算法
5	1×10^{-4}	0.1	Adam

3.4 结果分析

除构建 T-CNN, L-CNN, M-CNN 及 MF-CNN 网络外(S-CNN, R-CNN 出现梯度弥散不收敛),本研究还依据表 3 中 MF-CNN 结构构建卷积自动编码器(convolutional autoencoder, 简称 CAE)、网络层数更多的深度卷积神经网络(deep convolution neural network, 简称 DCNN)及 ANN,并将频谱划分为 11 段提取频谱能量作为 ANN 的输入。各网络模型以 N+IN(IN 为正常数据频谱和 AR 处理后频谱之差,保证噪声数据也具备频谱特征)、N+OF, N+IF 及 N+RF 等少类别标签数据作为训练集,各模型故障识别效果如表 5 所示,可以看出:

1) MF-CNN 较 T-CNN, L-CNN 及 M-CNN, 预训练时故障识别准确率更高,说明 Morlet-FDF 激活函数优于 Tanh, Leakyrelu 和 Morlet 激活函数;

2) 基于 J_2 损失训练的 MF-CNN 较基于 J_1 损失训练的 MF-CNN, 3 次迭代后故障识别准确率更高,说明权重内积最小化损伤有助于网络提取多样性特征,提升了故障识别准确率;

3) 以 AR 处理后的频谱数据作为输入,较以原始频谱数据作为输入,3 次迭代后故障识别准确率更高,进一步说明 AR 模型对双转子轴承数据进行了特征增强;

4) 以 AR 处理后的频谱数据作为输入, J_2 损失训练的 MF-CNN 网络输出特征聚类时, SC 值随聚类数 K 的变化趋势如图 8 所示。可以看出, SC 值转

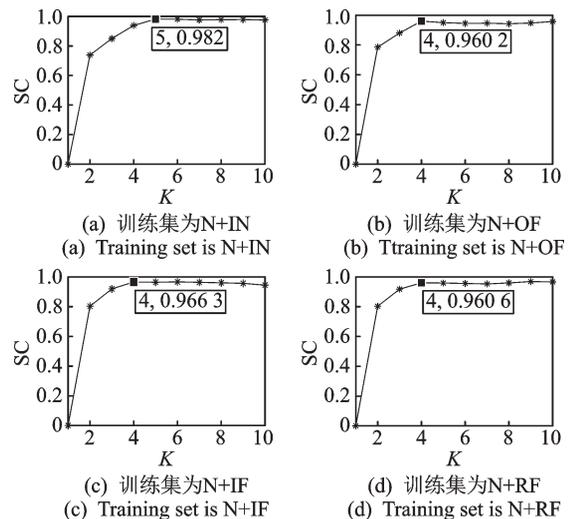


图 8 SC 随 K 值变化趋势

Fig.8 Variation of SC with K value

表5 各模型故障识别效果
Tab.5 Fault identification results of each model

网络	训练数据 类型	训练 时间/s	识别准确率/%			网络	训练数据 类型	训练 时间/s	识别准确率/%				
			预训练	迭代 1次	迭代 2次				迭代 3次	预训练	迭代 1次	迭代 2次	迭代 3次
T-CNN (J_1 损失)	N+IN	28	60	80	76	100	N+IN	32	100	100	100	100	
	N+OF	28	100	100	100	100	MF-CNN	N+OF	105	100	100	100	
	N+IF	28	75	75	75	75	(AR+ J_1 损失)	N+IF	57	75	100	100	100
	N+RF	28	75	75	75	75	N+RF	365	98.75	100	100	100	
L-CNN (J_1 损失)	N+IN	24	80	100	100	100	N+IN	35	100	100	100	100	
	N+OF	34	100	100	100	100	MF-CNN	N+OF	66	100	100	100	100
	N+IF	28	75	75	75	75	(AR+ J_2 损失)	N+IF	33	100	100	100	100
	N+RF	29	92.5	75	100	100	N+RF	69	100	100	100	100	
M-CNN (J_1 损失)	N+IN	118	80	100	100	100	N+IN	45	100	100	100	100	
	N+OF	98	100	100	100	100	DCNN	N+OF	148	100	100	100	100
	N+IF	70	75	75	75	75	(AR+ J_2 损失)	N+IF	42	100	100	100	100
	N+RF	117	62.5	75	100	100	N+RF	46	75	100	100	100	
MF-CNN (J_1 损失)	N+IN	364	80	100	100	100	N+IN	8 400	80	80	80	80	
	N+OF	73	100	100	100	100	CAE	N+OF	7 500	75	75	75	75
	N+IF	24	75	75	75	75	(AR+ J_2 损失)	N+IF	8 000	75	75	75	75
	N+RF	32	100	100	100	100	N+RF	7 600	75	75	75	75	
MF-CNN (J_2 损失)	N+IN	372	74	100	100	100	N+IN	17	60	60	60	60	
	N+OF	29	100	100	100	100	ANN	N+OF	20	75	75	75	75
	N+IF	32	50	75	100	100	(AR+ J_2 损失)	N+IF	10	75	75	75	75
	N+RF	29	100	100	100	100	N+RF	14	75	75	75	75	

折点对应的 K 值即为聚类数 $K=5$ 或 4 , 与实际相符。

综上,通过分析各模型故障识别准确率及模型 SC 随 K 变化曲线,验证了本研究提出的 AR 信号去噪方法、Morlet-FFT 激活函数及权重内积最小化损失的有效性,以及基于改进 CNN 和 Kmeans 的双转子轴承半监督故障诊断方法在少类别故障数据训练时进行故障诊断的有效性。

4 结 论

- 1) 基于 AR 对振动信号预处理,并进行快速傅里叶变换获取频谱输入,减小了 CNN 特征提取复杂度。
- 2) 从激活函数、损失函数角度改进 CNN,增强了 CNN 特征提取能力。
- 3) 根据自训练思想,在故障数据缺乏情况下,基于少类别训练数据结合 Kmeans 算法迭代训练网络,减小网络训练对有标签故障数据的依赖。
- 4) 通过双转子轴承故障模拟试验台数据,验证了所提方法在故障数据缺乏时进行双转子轴承故障

诊断的有效性,双转子轴承正常、已知故障和未知故障状态识别率达到 100%,减小了双转子轴承智能诊断方法对故障训练样本类别的依赖,提高了工程应用价值。

参 考 文 献

- [1] 廖明夫,刘永泉,王四季,等.中介轴承对双转子振动的影响[J].机械科学与技术,2013,32(5):641-646.
LIAO Mingfu, LIU Yongquan, WANG Siji, et al. The vibration features of a twin spool rotor system with an inter-bearing[J]. Mechanical Science and Technology for Aerospace Engineering, 2013, 32(5): 641-646. (in Chinese)
- [2] EL-THALJI I, JANTUNEN E. A summary of fault modelling and predictive health monitoring of rolling element bearings[J]. Mechanical Systems and Signal Processing, 2015(60/61): 252-272.
- [3] MISHRA R K, MUDULI S K, SRINIVASAN K, et al. Failure analysis of an inter-shaft bearing of an aero gas turbine engine[J]. Journal of Failure Analysis and Prevention, 2015, 15(2): 205-210.
- [4] 艾延廷,田博文,田晶,等.Morlet 复小波频带优化及其

- 在中介轴承故障诊断中的应用[J]. 航空动力学报, 2020, 35(1):153-161.
- AI Yanting, TIAN Bowen, TIAN Jing, et al. Frequency band optimization of Morlet complex wavelet and its application in fault diagnosis of inter-shaft bearing [J]. Journal of Aerospace Power, 2020, 35(1):153-161. (in Chinese)
- [5] JIANG Z, HU M, FENG K, et al. Weak fault feature extraction scheme for intershaft bearings based on linear prediction and order tracking in the rotation speed difference domain[J]. Applied Sciences, 2017, 7(9):1-25.
- [6] 雷亚国, 贾峰, 周昕, 等. 基于深度学习理论的机械装备大数据健康监测方法[J]. 机械工程学报, 2015, 51(21):49-56.
- LEI Yaguo, JIA Feng, ZHOU Xin, et al. Deep learning-based method for machinery health monitoring with big data[J]. Journal of Mechanical Engineering, 2015, 51(21):49-56. (in Chinese)
- [7] 王奉涛, 薛宇航, 王洪涛, 等. GLT-CNN方法及其在航空发动机中介轴承故障诊断中的应用[J]. 振动工程学报, 2019, 32(6):1077-1083.
- WANG Fengtao, XUE Yuhang, WANG Hongtao, et al. GLT-CNN and its application of aero-engine intermediary bearing fault diagnosis [J]. Journal of Vibration Engineering, 2019, 32(6):1077-1083. (in Chinese)
- [8] 王丽华, 谢阳阳, 周子贤, 等. 基于卷积神经网络的异步电机故障诊断[J]. 振动、测试与诊断, 2017, 37(6):1208-1215.
- WANG Lihua, XIE Yangyang, ZHOU Zixian, et al. Motor fault diagnosis based on convolution neural networks[J]. Journal of Vibration, Measurement & Diagnosis, 2017, 37(6):1208-1215. (in Chinese)
- [9] LI X, LI J, QU Y, et al. Semi-supervised gear fault diagnosis using raw vibration signal based on deep learning[J]. Chinese Journal of Aeronautics, 2020, 33(2):418-426.
- [10] LEE D H. Pseudo-Label: the simple and efficient semi-supervised learning method for deep neural networks [C]//Workshop on Challenges in Representation Learning. Atlanta, GA, USA: ICML, 2013:2-7.
- [11] LEI Y, YANG B, JIANG X, et al. Applications of machine learning to machine fault diagnosis: a review and roadmap[J]. Mechanical Systems and Signal Processing, 2020, 138:106587.
- [12] SALIMANS T, KINGMA D P. Weight normalization: a simple reparameterization to accelerate training of deep neural networks [C]//Proceedings of the 30th International Conference on Neural Information Processing Systems. Barcelona, Spain: Curran Associates Inc, 2016: 901-909.
- [13] JIA F, LEI Y, LU N, et al. Deep normalized convolutional neural network for imbalanced fault classification of machinery and its understanding via visualization[J]. Mechanical Systems and Signal Processing, 2018, 110:349-367.
- [14] JING L, TIAN Y. Self-supervised visual feature learning with deep neural networks: a survey[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2021, 43(11):4037-4058.
- [15] 张靖, 段富. 优化初始聚类中心的改进k-means算法[J]. 计算机工程与设计, 2013, 34(5):1691-1694, 1699.
- ZHANG Jing, DUAN Fu. Improved k-means algorithm with meliorated initial centers[J]. Computer Engineering and Design, 2013, 34(5):1691-1694, 1699. (in Chinese)
- [16] YOU W, SHEN C, WANG D, et al. An intelligent deep feature learning method with improved activation functions for machine fault diagnosis[J]. IEEE Access, 2019, 8:1975-1985.
- [17] SHAO H, JIANG H, LI X, et al. Intelligent fault diagnosis of rolling bearing using deep wavelet auto-encoder with extreme learning machine[J]. Knowledge-Based Systems, 2018, 140:1-14.
- [18] XU K, QIN M, SUN F, et al. Learning in the frequency domain[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. [S.l.]: The Computer Vision Foundation, 2020:1740-1749.
- [19] AYINDE B O, INANC T, ZURADA J M. Regularizing deep neural networks by enhancing diversity in feature extraction[J]. IEEE Transactions on Neural Networks and Learning Systems, 2019, 30(9):2650-2661.
- [20] MIN H, LIU Y, XI J, et al. Noise smoothing for nonlinear time series using wavelet soft threshold [J]. IEEE Signal Processing Letters, 2007, 14(1):62-65.
- [21] DONG H, QI K, CHEN X et al. Sifting process of EMD and its application in rolling element bearing fault diagnosis[J]. Journal of Mechanical Science & Technology, 2009, 23(8):2000-2007.



第一作者简介: 崔锦森, 男, 1995年11月生, 硕士生。主要研究方向为旋转机械智能故障诊断。
E-mail: oooomiaol@163.com