

DOI: 10.13382/j.jemi.B2407341

基于时频滤波器和偏移注意神经网络的 轴承故障诊断*

赵运基 危思成 许孝卓

(河南理工大学电气工程与自动化学院 焦作 454003)

摘要:针对轴承故障数据分布不一致导致特征偏移、标志性特征提取困难的问题,提出一种基于时频滤波器和偏移注意神经网络的轴承故障诊断方法,从离线和在线两个方向分别对故障信号进行处理。在离线部分提出时频滤波器,分别从时域和频域提取故障信号的标志性特征;提出兼顾全局特征和局部特征的空间采样方法。在线部分提出偏移注意神经网络,与自注意相比,偏移注意更有利于偏移特征的提取,从而降低数据分布不一致造成的影响。在西安交通大学和凯斯西储大学的轴承数据集上进行实验,达到了100%的精度,证明了所提方法能够很好的提取故障信号的标志性特征,并且能够有效抑制特征偏移对故障识别精度的影响。而在凯斯西储大学轴承数据集上的对比实验则证明了所提方法的优越性。除此之外,还在工业现场采集的燃气轮机主轴承数据集上进行了实验,结果证明所提方法具有实际应用意义。

关键词: 故障诊断; 特征偏移; 标志性特征; 时频滤波器; 偏移注意

中图分类号: TN911.7; TH133.33

文献标识码: A

国家标准学科分类代码: 460.4020

Bearing fault diagnosis based on time-frequency filter and offset attention neural network

Zhao Yunji Wei Sicheng Xu Xiaozhuo

(School of Electrical Engineering and Automation, Henan Polytechnic University, Jiaozuo 454003, China)

Abstract: To address the inconsistent bearing fault data distribution that leads to the difficulty of feature offset and distinctive feature extraction, a bearing fault diagnosis method based on time-frequency filter and offset attention neural network is proposed, which processes the fault signal from offline and online parts. In the offline part, a time-frequency filter is proposed to extract the distinctive features from time domain and frequency domain; A spatial sampling method considering both global and local features is proposed. In the online part, an offset attention neural network is proposed. Compared with self attention, offset attention is more conducive to the extraction of offset features, so as to reduce the impact caused by inconsistent data distribution. Experiments on the bearing datasets of Xi'an Jiaotong University (XJTU) and Case Western Reserve University (CWRU) have achieved 100% accuracy, which proves that the proposed method can efficiently extract the distinctive features of fault signals, and effectively suppress the influence of feature offset. The comparative experiment on the bearing dataset of CWRU proves the superiority of the proposed method. In addition, experiments are also carried out on the dataset of gas turbine main bearing collected in the industrial field, and the results show that the proposed method has practical significance.

Keywords: fault diagnosis; feature offset; distinctive feature; time-frequency filter; offset attention

0 引言

旋转机械广泛应用于工业生产,如风力发电机组、水轮发电机组和计算机数控机床^[1]。作为其重要组成部分

的轴承,因为工作环境的复杂多变而面临许多故障触发因素。而旋转设备的运行状况主要取决于轴承的状况^[2]。当轴承出现故障时容易引发安全问题,造成损失,影响生产安全和经济效益^[3]。因此,对轴承的故障状态进行检测有着重要的意义。

收稿日期: 2024-03-15 Received Date: 2024-03-15

* 基金项目: 国家自然科学基金(61973105, 61573130, 52177039)、河南省科技攻关项目(242102221034, 232102240096)资助

近年来,振动信号因其易于获取和分析而被广泛用于轴承故障诊断。基于数据驱动的故障诊断方法主要有两大步骤,特征提取和识别分类。在对信号进行特征提取时,主要从时域、频域或时频域进行。在时域提取故障信号特征的方法有中值滤波、数字滤波等。在频域有快速傅里叶变换(fast Fourier transform, FFT)、离散傅里叶变换等。在时频域,有小波变换(wavelet transform, WT)^[4]、经验模态分解(empirical mode decomposition, EMD)^[5]和希尔伯特-黄变换^[6]等。然而在出现数据分布移位的情况下,上述方法会因为特征偏移而最终影响分类精度。除此之外,如何高效提取故障信号的全局特征与局部特征同样也是实现高精度故障诊断的关键。

另一重要步骤为故障信号的识别与分类。近年来,机器学习方法被广泛应用于轴承故障诊断。而机器学习中的深度学习可以有效地提取振动信号的深层特征,能够高效的对故障信号进行识别并完成分类。Zhao 等^[7]根据不同的输入大小,自适应调整可变形卷积网络(deformable convolutional network, DCN)中卷积核的大小和形状。通过颜色空间模型(color names, CN)与主成分分析(principal component analysis, PCA)的结合(即空间映射),将原始数据投影到更具判别性的空间中。最后,通过 DCN 对故障完成分类。Xu 等^[8]将注意力机制与多尺度卷积神经网络相结合,由此提升神经网络的泛化能力,最终提升故障诊断精度。Zhou 等^[9]提出了一种频率注意机制,并将其与卷积神经网络相结合,最终实现对含噪信号的故障诊断。针对特征偏移问题,Liu 等^[10]利用平移不变字典来对局部故障引起的冲击信号进行稀疏表示,并根据循环结构对其进行傅里叶对角化。Xing 等^[11]提出了一种基于麻雀搜索算法的分层平移不变脉冲字典匹配追踪方法,该方法采用层次化的思想,利用字典的平移不变结构来提取每个故障分量。Hei 等^[12]将注意力机制与平移不变字典学习法相结合,取得了良好的分类精度。然而平移不变字典的创建需要依赖于复杂的先验知识。为充分发挥注意力机制提取区分度特征的能力,同时有效解决特征偏移问题,提出了偏移注意,并将其与神经网络相结合。

鉴于上述分析,提出了一种基于时频滤波器和偏移注意神经网络的轴承故障诊断方法。分别在离线和在线两部分提取故障信号的特征。在离线状态下,为降低特征偏移对识别精度的影响,有效提取对序列分布偏移具有鲁棒性的标志性特征,采用梅尔频率倒谱系数(Mel frequency cepstrum coefficients, MFCC)将时域信号转换为频域信号。采用最远点采样和最近邻采样获取全局特征,采用邻域嵌入获取局部特征。在在线状态下,为进一步提取故障信号的偏移特征,提出了一种偏移注意神经网络。与普通注意机制相比,偏移注意神经网络对特征

偏移具有更强的鲁棒性。

1 方法与模型

1.1 时频滤波器

时频滤波器对信号的处理流程可以分为以下几步:首先,对原始故障信号进行预加重。预加重可对输入信号的高频分量进行补偿,从而使输入信号的波形不失真。其次是分帧与加窗,分帧将信号划分为多个片段,由此来保留时间序列信息,使用汉明窗来消除信号的不连续性。最后,对信号进行快速傅里叶变换,将所得的频率映射到梅尔频率,映射函数如式(1)所示。

$$f_{mel} = 2595 \lg(1 + \frac{f}{700}) \quad (1)$$

其中, f_{mel} 表示梅尔频率, f 是 FFT 后获得的频率。每个梅尔滤波器的频率响应如式(2)和(3)所示。

$$f_u(m-1) = f_c(m) = f_l(m+1) \quad (2)$$

$$H_m(k) = \begin{cases} 0, k < f_c(m-1) \\ \frac{k - f_c(m-1)}{f_c(m) - f_c(m-1)}, f_c(m-1) \leq k \leq f_c(m) \\ \frac{f_c(m+1) - k}{f_c(m+1) - f_c(m)}, f_c(m) < k \leq f_c(m+1) \\ 0, k > f_c(m+1) \end{cases} \quad (3)$$

其中, $H_m(k)$ 表示第 m 个梅尔滤波器。 f_c, f_l, f_u 是梅尔滤波器的中心频率、下截止频率和上截止频率。在获得梅尔频率之后,依次进行对数运算和离散余弦变换(discrete cosine transform, DCT)。最后对信号进行增强,增强函数如式(4)所示。

$$f_{lifting}(g) = \left(1 + \frac{G}{2} \sin \frac{\pi g}{G}\right) f(g) \quad (4)$$

其中, $f_{lifting}(g)$ 是增强后的 MFCC, g 是 MFCC 的阶数, $f(g)$ 为输入信号。信号增强对不同阶数的特征进行加权,以增强低阶和中阶特征^[13]。

MFCC 最初应用于音频提取算法,使用梅尔滤波器处理声音信号,可以抵抗背景噪声的干扰^[14]。受此启发,将 MFCC 应用于滚动轴承故障信号的特征提取。与时域中的故障信号相比,频域信号的变化更为显著,易于分析和观察,并且频域中的故障信号对特征偏移具有鲁棒性。

1.2 邻域嵌入

最远点采样的核心原理是通过优化算法使得所有采样点在特征空间中相互尽可能远离,从而实现对数据的离散和均匀采样。该方法能够有效地进行欠采样,从而解决样本数据中存在的冗余等问题。为进一步提取全局

特征,使用最近邻采样法来获取样本中的显著特征。

虽然同时使用最远点采样和最近邻采样能够很好的提取全局特征,但往往会忽略局部信息。然而,对神经网络的学习而言,局部信息同样至关重要。此外,虽然注意力机制在全局特征的获取上表现出一定的有效性,但通常会忽略局部信息的贡献。为解决以上问题,引入邻域嵌入策略来优化数据点的表示。与单个数据点的局部特征相比,邻域嵌入包含邻域内所有数据点的特征信息,有助于注意力模块更好地提取特征。

邻域嵌入模块包括两个线性层(linear layers)、批量归一化层(batch normalization layers)和渗漏整流线性单元(leaky rectified linear unit)。以上各层构成 LBR 层。除 LBR 层外,邻域嵌入还包含两个采样和分组(sampling and grouping, SG)层^[15]。LBR 层作为基础数据点嵌入。而 SG 层则用于在特征聚合过程中扩大感受野。在采样过程中,SG 层通过对每个数据点进行 k 近邻搜索来聚合局部邻域的特征。邻域嵌入的输出

$F_s(p)$ 如式(5)所示。

$$\Delta F(p) = \text{concat}_{q \in \text{knn}(p, P)} (F(q) - F(p))$$

$$\tilde{F}(p) = \text{concat}(\Delta F(p), RP(F(p), k))$$

$$F_s(p) = MP(LBR(LBR(\tilde{F}(p)))) \quad (5)$$

其中, $F(p)$ 为点 p 的输入特征, $F_s(p)$ 为采样点 p 的输出特征, MP 为最大池化算子, $RP(x, k)$ 为运算符, 代表重复 x 向量 k 次来生成一个矩阵。

最终,最远点采样、最近邻采样和领域嵌入共同构成了空间采样方法,能够同时兼顾数据中的全局特征与局部特征。

1.3 偏移注意

为更好地解决数据分布偏移的问题,提出偏移注意(offset attention)来替代自注意(self attention)。其原理是将注意力特征替换为自注意模块输入与注意特征之间的偏移量。偏移注意的结构如图 1 所示。

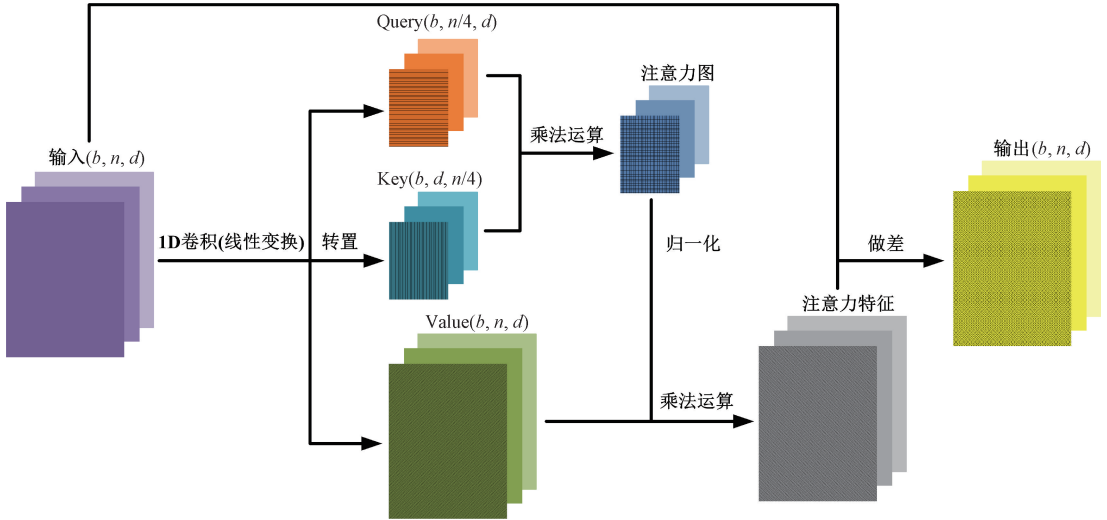


图 1 偏移注意

Fig. 1 Offset attention

入维数为 (b, n, d) , 经过一维卷积, 分别得到查询向量(query vector)、键向量(key vector)和值向量(value vector)。这 3 个键值与 Transformer^[16] 中的键值相似。其中查询向量大小为 $(b, n/4, d)$, 键向量大小为 $(b, d, n/4)$, 值向量大小为 (b, n, d) 。将查询向量与键向量相乘, 得到注意力图。对注意力图进行归一化:

$$\bar{\alpha}_{i,j} = \text{softmax}(\tilde{\alpha}_{i,j}) = \frac{\exp(\tilde{\alpha}_{i,j})}{\sum_k \exp(\tilde{\alpha}_{k,j})}$$

$$\alpha_{i,j} = \frac{\bar{\alpha}_{i,j}}{\sum_k \bar{\alpha}_{i,k}} \quad (6)$$

对第 1 个维度使用 softmax 算子, 在第 2 个维度上使用 $L1$ 范数来规范化。将归一化后的注意力图与值向量相乘, 得到注意力特征。最后, 计算开始时的输入和注意力特征之间的差值, 得到偏移注意。偏移注意力能够增强注意力权重并减少噪声的影响, 有利于后续对故障信号的分类。

2 整体框架

传感器测得的振动信号如图 2 所示。若样本 1 为训练样本(点线框), 样本 2 为测试样本(虚线框), 二者的数据分布存在差异。数据分布不一致导致的特征偏移会

影响分类精度。如果信号具有相同的内部结构, 总体显示出周期性递归, 即这些内部结构在信号中的不同位置相互位移, 这被称为位移不变量^[17]。

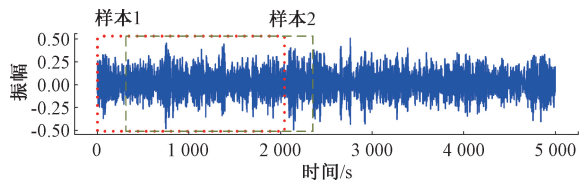


图 2 轴承振动信号

Fig. 2 Bearing vibration signal

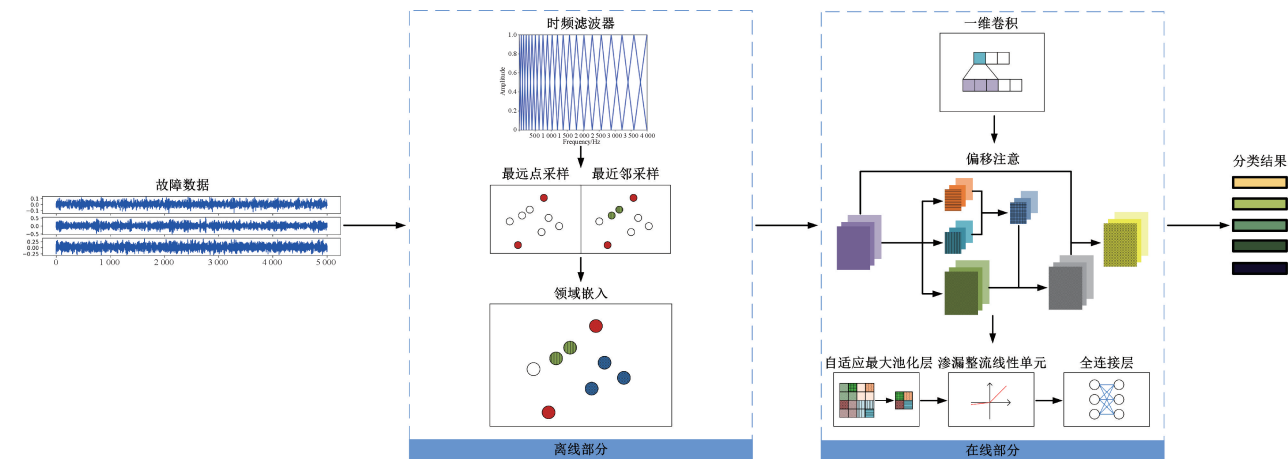


图 3 基于时频滤波器和偏移注意神经网络的轴承故障诊断整体框架

Fig. 3 The whole framework of the bearing fault diagnosis based on time-frequency filter and offset attention neural network

离线部分使用时频滤波器将 2048×3 的单个信号样本转换为 600×3 (3 代表 3 个数据通道)。对处理后的信号采用最远点采样和最近邻采样, 提取全局特征。利用邻域嵌入方法获取局部特征。邻域嵌入模块包括 LBR 层和两个 SG 层。LBR 层作为数据点嵌入的基础。在特征聚合过程中, SG 层用于扩大感受野。SG 层利用欧氏距离对每个搜索组进行特征聚合。在线部分使用偏移注意力神经网络对偏移特征进行提取和分类。

在获得偏移注意力后, 将其与输入特征融合。经过一维自适应最大池化层、渗漏整流线性单元和全连接层后, 最终输出分类结果。自适应最大池化与普通最大池化的区别在于其输出特征大小仅与输出大小相关。而普通最大池化的输出大小需要由核大小、填充和其他参数共同计算得出。通过设置自适应最大池化的参数, 可使池化层完全平移不变, 最终实现特征对齐^[18]。

渗漏整流线性单元是一种非线性函数, 在网络中作为激活函数, 可提高模型的表达性, 使网络更好地拟合数据。作为标准整流线性单元的一个变体, 它包含了一个新的参数 η , 它允许负值传递, 而不是直接将值设置为零。因此, 渗漏整流线性单元激活函数使动态范围更宽, 解决了如梯度消失等深度学习问题。此外, 渗漏整流线

性单元激活函数可以更好地拟合模型, 具有更快的收敛速度并且更容易激活神经元^[19]。

为消除特征偏移对分类精度的影响, 提出了基于时频滤波器和偏移注意神经网络的轴承故障诊断方法。所提出方法的总体框架如图 3 所示。提出的方法分为离线部分和在线部分, 其中离线部分主要包含时频滤波器, 在线部分为偏移注意神经网络。对原始故障信号分别采用固定间隔采样和随机采样, 由此得到数据集。为减少网络训练在线部分的计算量, 在离线状态下对数据进行预处理。由此网络能够更好地提取滚动轴承故障信号中的偏移特征。

3 实验验证

3.1 实验数据

为验证提出方法的有效性, 分别在 CWRU 和 XJTU 数据集上进行了实验。CWRU 数据集中共有 10 个故障类别, XJTU 数据集中共有 15 个故障类别。

1) CWRU 数据集

CWRU 数据集由凯斯西储大学轴承数据中心提供^[20]。CWRU 数据集是最经典的公共数据集之一, 已成为验证各种故障诊断方法的标准参考。CWRU 实验平台如图 4 所示。

实验中使用的三通道数据指的是 CWRU 中的驱动端加速度数据 (DE)、风扇端加速度数据 (FE) 和基座加速度数据 (BA)。所选择的采样频率为 12 kHz。使用的 10 种故障类型如表 1 所示, 包括故障直径分别为 0.177 8 mm 和 0.533 4 mm 的滚珠轴承故障、内圈故障以及 3 点钟、6 点钟和 12 点钟方向的外圈故障。

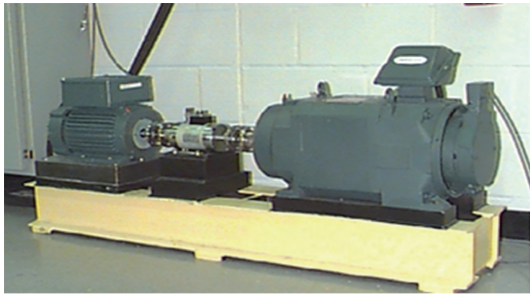


图 4 CWRU 实验平台

Fig. 4 CWRU experimental platform

表 1 CWRU 数据集的十类故障

Table 1 Ten classes of faults in CWRU

类别标签	故障类别	故障直径/mm	单个样本长度
0	滚珠轴承故障	0.177 8	121 556
1	内圈故障	0.177 8	121 917
2	3 点钟方向外圈故障	0.177 8	122 281
3	6 点钟方向外圈故障	0.177 8	122 571
4	12 点钟方向外圈故障	0.177 8	122 136
5	滚珠轴承故障	0.533 4	122 136
6	内圈故障	0.533 4	121 991
7	3 点钟方向外圈故障	0.533 4	122 136
8	6 点钟方向外圈故障	0.533 4	121 991
9	12 点钟方向外圈故障	0.533 4	121 846

数据的离线处理以及训练和测试数据集的划分如图 5 所示。图中以滚珠轴承故障为例,在一个数据文件中,故障信号大小为 121 556×3。滚珠轴承故障数据集的大小为 1 200×2 048×3。共有 10 类数据,最终的数据集大小为 12 000×2 048×3。训练测试比为 2 : 8,即训练数据大小为 2 400×2 048×3,测试数据大小为 9 600×2 048×3。为验证数据特征偏移对识别精度的影响,训练数据采用顺序截取,而测试数据采用随机截取。

2) XJTU 数据集

XJTU 数据集由西安交通大学提供^[21]。数据集中共有 3 种振动频率:35 Hz(2 100 r/min,12 kN)、37.5 Hz(2 250 r/min,11 kN)和 40 Hz(2 400 r/min,10 kN)。每个频率包含 5 种故障类型。因此,该数据集共包含 15 种滚动轴承故障数据。XJTU 轴承数据集有两个通道:垂直振动和水平振动。

XJTU 数据集的获取方式与 CWRU 类似,不同之处在于 XJTU 只有两个数据通道。为了适应提出的故障诊断方法,增加数据的通道数,第 3 条通道的数据为垂直振动通道数据和水平振动通道数据的算术平均值。

3) 燃气轮机主轴承数据集

为进一步验证所提方法的有效性,将算法应用于工业现场采集的数据集,具体为燃气轮机主轴承数据集。该燃气轮机主要用于发电。在运行过程中,使用传感器

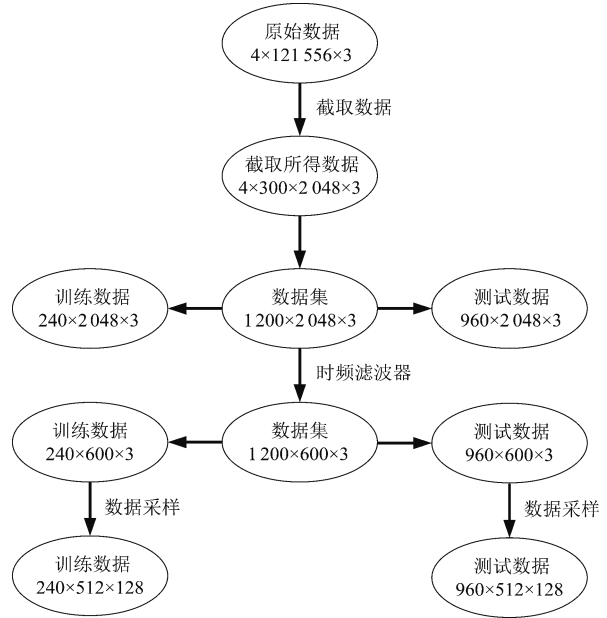


图 5 离线处理数据过程

Fig. 5 Offline processing of the data

收集主轴承的振动信号。装置整体如图 6 所示,实验中使用的数据是在两个月内所测量的 X 轴、Y 轴和 Z 轴的振动信号。

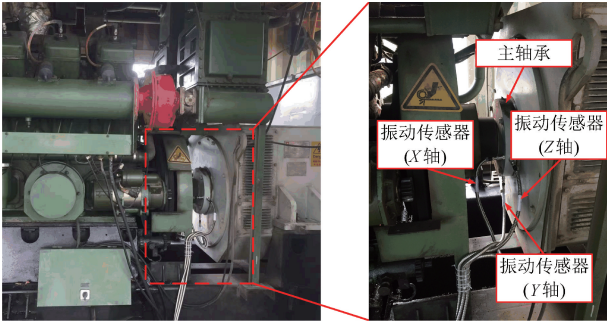


图 6 燃气轮机主轴承振动信号的采集

Fig. 6 Vibration signal acquisition of gas turbine main bearing

传感器以 8 kHz 的频率对振动信号进行采样,每个采样周期共持续 12 s。因此,在每个周期中收集的振动信号的长度为 96 000。燃气轮机的主轴承大约每两个月需要更换一次,在此期间进行了 305 次测量。根据使用时间对轴承的损耗划分等级。在一定程度上,均方根 (root mean square, RMS) 曲线可以反映轴承的健康状况。RMS 曲线表示传感器在 X、Y 和 Z 这 3 个方向上的测量值的均方根。其具体计算如式 (7) 所示。数据集的 RMS 曲线如图 7 所示。

$$f_{\text{RMS}} = \sqrt{\frac{\sum_{i=1}^n x_i^2}{n}} \quad (7)$$

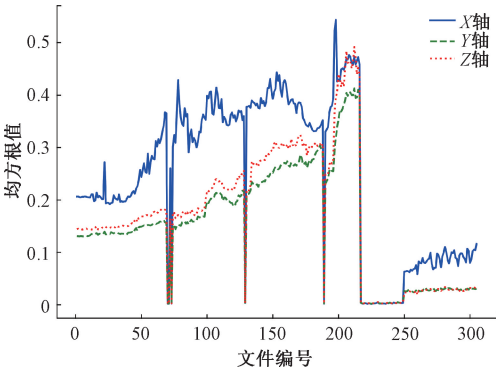


图 7 均方根曲线

Fig. 7 Root mean square curve

在实际运行时间等于 1 063 h (对应第 215 个数据文件)时,轴承被更换为新的。此时 3 个方向上的传感器值对应于 RMS 曲线的值基本上为 0。指定该时间段故障等级为 0,即为正常状态;实际运行时间等于 1 203 h 到实际运行时间等于 1 579 h 的故障等级为 1;实际运行时间等于 0 h 到实际运行时间等于 195 h 的故障等级为 2;实际运行时间等于 411 h 到实际运行时间等于 715 h 的故障等级为 3;实际运行时间等于 895 h 到实际运行时间等于 1 063 h 的故障等级为 4。具体如表 2 所示。

表 2 燃气轮机主轴承数据集的故障类别

Table 2 Category of gas turbine main bearing dataset

数据文件编号	实际运行时间/h	故障等级	标签
0~50	0~195	Fault-2	2
90~140	411~715	Fault-3	3
185~215	895~1 063	Fault-4	4
215~250	1 063~1 203	Normal	0
250~300	1 203~1 579	Fault-1	1

表 2 中数据文件编号表示每次采样对应的数据,而实际运行时间是指轴承从数据采集开始到结束的实际时间,采集开始时的时间记为 0,时间计量单位为 h。采样的时间间隔是非连续的。与 CWRU、XJTU 数据类似,从数据文件中截取数据作为样本,每个样本长度为 2 048,共有 3 个维度(X 轴、Y 轴和 Z 轴)。每类数据共截取 1 200 个,共有 5 类数据,最终的数据集大小为 6 000×2 048×3。训练测试比为 2 : 8,即训练数据大小为 1 200×2 048×3,测试数据大小为 4 800×2 048×3。

3.2 实验设置

验证算法的硬件条件是 Intel i9 处理器和 RTX 3080 显卡。PyCharm 被用作集成开发环境。在分类器训练过程中,将学习率调整为 0.1,并选择随机梯度下降作为优化器。将批量大小设置为 32。3 个数据集的训练数据和测试数据的比例都为 2 : 8。训练步长都为 100。消融实验采用控制变量法进行,实验分为多组。训练数据通过

固定间隔采样获得,测试数据通过随机采样获得。通过分组和控制变量法,分别比较了时频滤波器和偏移注意力网络的有效性和可靠性。

除此之外,为证明所提出方法的优越性,将提出的方法与其他方法进行对比。对比实验分为模型对比与总体方法对比。使用的对比模型有 Resnet50^[22]、ViT^[23]、3DCNN^[24]和 WDCNN^[25]。模型对比实验中数据都是经过时频滤波器处理后的。方法对比中有 1D-ViT-1Depth^[26]、MAB-DrNet^[27]、Lightweight + MSF^[28]和 SVAMN^[29]。对比实验是在 CWRU 数据集上进行的,为确保实验的公平,各实验参数设置都是相同的。

3.3 实验结果

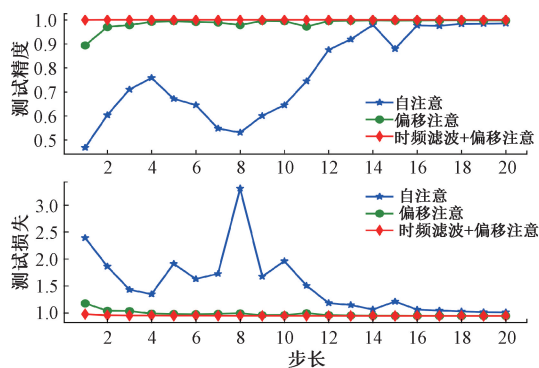
为验证数据偏移对分类精度的影响,训练数据采用固定间隔采样的方式获取,测试数据采用随机采样的方式获取。实验中步长为 100,每 5 步取平均值,最后得到 20 个值。实验分为 3 组:一组自注意,一组偏移注意,一组时频滤波与偏移注意。3 组实验在 CWRU 数据集上的测试精度与损失如图 8(a)所示,在 XJTU 数据集上的测试精度与损失如图 8(b)所示,在燃气轮机主轴承数据集上的测试精度与损失如图 8(c)所示。

由图 9 可以看出,在 CWRU 和 XJTU 数据集上,时频滤波器和偏移注意力的组合都有着更好的表现。在燃气轮机主轴承数据集上,虽然 3 条曲线的最终结果接近,但是时频滤波器和偏移注意力的组合有着更快的收敛速度,且较于其他两条曲线更加的稳定。除此之外,相较于自注意,在精度还是损失上偏移注意的故障诊断效果都更为优秀。结合以上实验结果,可以说明提出的方法具有较好的泛化性和较高的鲁棒性。

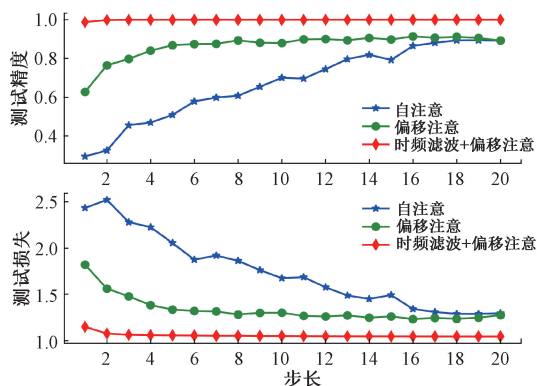
为更好地观察模型对各类故障类型的诊断结果,利用混淆矩阵对其进行可视化,3 个数据集的测试集混淆矩阵如图 9 所示。混淆矩阵中横坐标为预测类别,纵坐标为实际类别,故对角线上的值表示该故障类型的识别率。提出的方法在对 CWRU 数据集和 XJTU 数据集的类别诊断中,对每类别故障的诊断准确率都能达到 100%,如图 9(a)和(b)所示;而在工业现场采集的燃气轮机主轴承数据集上也能达到较高的分类精度,如图 9(c)所示。说明该模型对各类故障类型具有较好的故障识别准确率,且有着很好的泛化能力。

最后,为了直观的得到提出方法的分类表现,在燃气轮机主轴承数据集的测试集上,绘制测试样本特征在度量空间中的 t 分布随机近邻嵌入(t-distributed stochastic neighbor embedding, t-SNE)降维可视化,如图 10 所示。每类故障都由不同的颜色、不同形状的图形来表示。

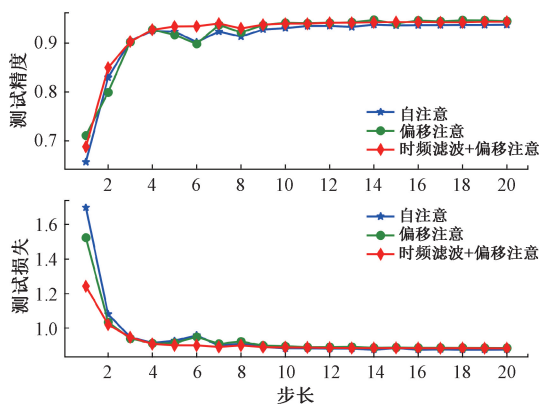
由图 10 可以看出,使用自注意的神经网络对故障进行分类,故障特征的分布具有明显的聚类现象,但样本依旧存在混淆堆叠的现象,如图 10(a)所示。基于偏移注



(a) CWRU测试结果
(a) CWRU dataset test result



(b) XJTU测试结果
(b) XJTU dataset test result

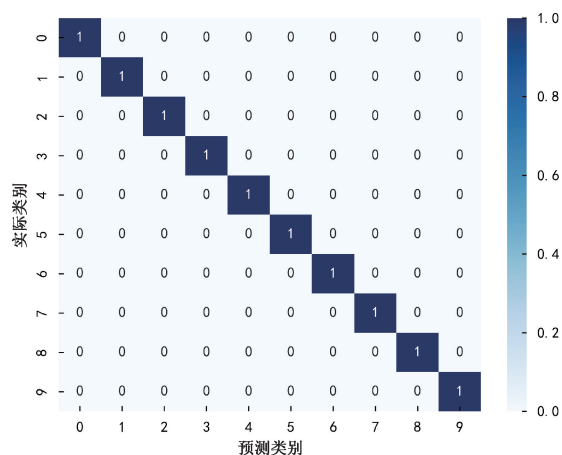


(c) 燃气轮机主轴承数据集测试结果
(c) Gas turbine main bearing dataset test result

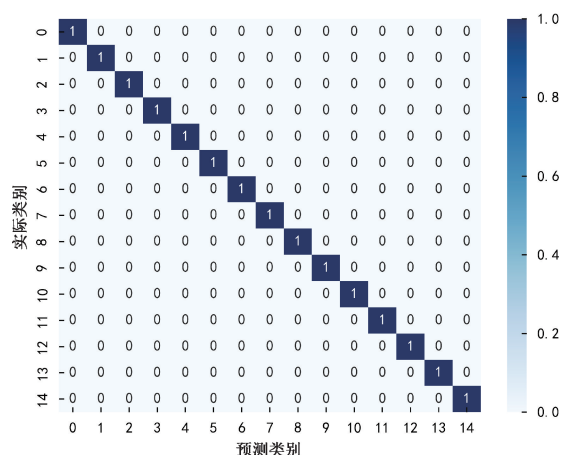
图 8 分组实验的测试结果

Fig. 8 Group experiment test results

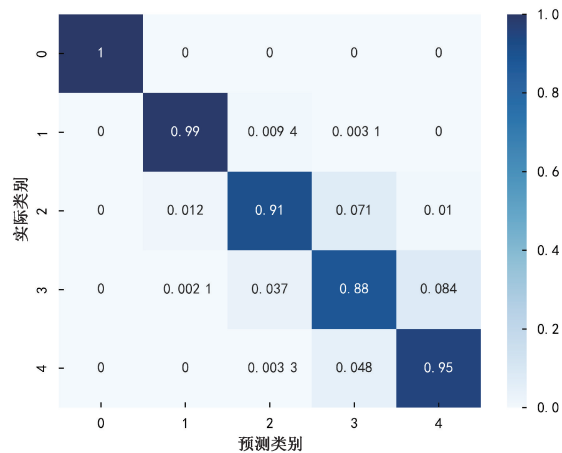
意的神经网络,能够更好的获取故障特征,因此特征分布具有更好的聚类效果,如图 10(b) 所示。经过时频滤波器处理的信号,更好的提取了故障信号的偏移特征,使得各个故障类别的分布更加聚集,如图 10(c) 所示。时频滤波和偏移注意神经网络可以很好的提取故障信号的标志性特征和偏移特征。最终使得类内间距更加集中,类间间距更加分离,即便对难以区分类别也有着明显的分离作用,各个故障类型的特征有着较好的聚类分布。



(a) CWRU测试集混淆矩阵
(a) CWRU test set confusion matrix



(b) XJTU测试集混淆矩阵
(b) XJTU test set confusion matrix



(c) 燃气轮机主轴承测试集混淆矩阵
(c) Gas turbine main bearing test set confusion matrix

图 9 测试集混淆矩阵

Fig. 9 Test set confusion matrix

对比实验结果如表 3 和 4 所示,其中表 3 为模型对比,表 4 为方法对比。对比实验都是在 CWRU 数据集上进行的,且为确保实验的公平,各实验参数设置都是相同的。

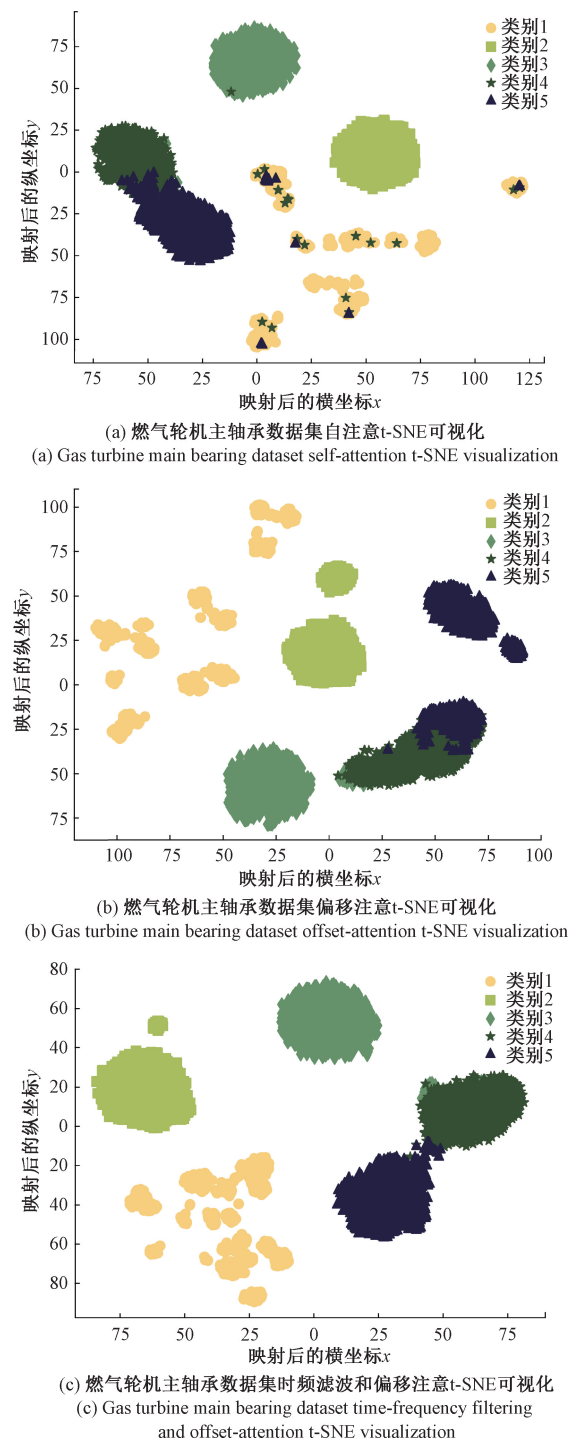


图 10 t-SNE 可视化

Fig. 10 Visualization obtained by t-SNE

表 3 可以看出,在数据都经过时频滤波处理的前提下,偏移注意神经网络有着更高的测试精度。由此体现了偏移注意神经网络的优越性。

表 4 可以看出,时频滤波与偏移注意神经网络的组合方法有着更高的测试精度。同时表 4 中 ViT 测试精度为 98.52%,而在表 3 中使用时频滤波处理后的数据测试

精度可以达到 99.59%,这体现了时频滤波器能够很好的提取标志性特征和偏移特征。

Table 3 Results of model comparison	
模型	测试精度/%
Resnet50	99.22
3DCNN	99.25
WDCNN	99.39
ViT	99.59
偏移注意神经网络	100

Table 4 Results of method comparison	
方法	测试精度/%
1D-ViT-1Depth	98.52
MAB-DrNet	99.33
Lightweight+MSF	99.50
SVAMN	99.98
时频滤波+偏移注意神经网络	100

4 结 论

在故障诊断中,训练数据与测试数据分布不一致时,会导致特征偏移,进而影响故障识别精度。针对这一问题,提出了一种时频滤波器与偏移注意神经网络相结合的故障诊断方法。通过时频滤波器对原始故障信号进行处理,提取标志性特征和偏移特征。构建偏移注意神经网络来进一步提取故障信号的偏移特征。实验结果表明,所提方法能够很好的解决特征偏移问题。然而,提出的方法仅考虑了在同一工况下的故障诊断,在实际工业环境中,轴承的工况通常是多变的,在不同工况下的数据存在分布不一致的问题。未来将在元学习、迁移学习等领域进行研究,以实现跨工况下的故障诊断。除此之外,实际应用中,还存在样本数量稀少、样本中含有噪声等情况。未来将结合数据增强、数据生成方法来实现小样本的故障诊断。针对信号中的噪声,未来将对时频滤波器进行改进,如在滤波器中加入自适应权重等。

参考文献

[1] 王鸽,吴国新,刘秀丽. 基于 MADCNN 的故障诊断方法研究 [J]. 电子测量与仪器学报, 2023, 37 (3): 187-193.

WHANG G, WU G X, LIU X L. Research on fault diagnosis method based on MADCNN [J]. Journal of Electronic Measurement and Instrumentation, 2023, 37(3): 187-193.

[2] 栾孝驰,郝冠丞,沙云东,等. 基于局部能量密度的中介轴承故障特征提取与诊断方法 [J]. 仪器仪表学报,

- 2024, 45(5): 239-250.
- LUAN X CH, HAO G CH, SHA Y D. Local energy density-based method for intermediary bearing fault feature extraction and diagnosis[J]. Chinese Journal of Scientific Instrument, 2024, 45(5): 239-250.
- [3] ZHAO Y J, XU J. A small sample bearing fault diagnosis method based on novel Zernike moment feature attention convolutional neural network[J]. Measurement Science and Technology, 2024, 35(6): 066208.
- [4] 张焱,何姝钡,韩延,等. 多小波系数增强动态聚合联邦深度网络的多工况故障诊断[J]. 电子测量与仪器学报, 2023, 37(5): 68-78.
- ZHANG Y, HE SH B, HAN Y. Multi-wavelet coefficients enhanced dynamic aggregation federal deep network for fault diagnosis under multiple conditions[J]. Journal of Electronic Measurement and Instrumentation, 2023, 37(5): 68-78.
- [5] 马超,王少红,徐小力. 基于 EEMD 的声阵列滚动轴承故障诊断[J]. 电子测量与仪器学报, 2017, 31(9): 1379-1384.
- MA CH, WANG SH H, XU X L. Fault diagnosis for rolling bearing by using acoustic array based on EEMD [J]. Journal of Electronic Measurement and Instrumentation, 2017, 31(9): 1379-1384.
- [6] ALICANDO M A R, RAMOS G M, OSTIA C F. Bearing fault detection of a single-phase induction motor using acoustic and vibration analysis through hilbert-huang transform[C]. 2021 IEEE 13th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment, and Management (HNICEM). IEEE, 2021: 1-6.
- [7] ZHAO Y J, ZHOU M L, XU X Z, et al. Fault diagnosis based on space mapping and deformable convolution networks[J]. IEEE Access, 2020, 8: 212599-212607.
- [8] XU Z F, LI CH, YANG Y. Fault diagnosis of rolling bearings using an improved multi-scale convolutional neural network with feature attention mechanism[J]. ISA Transactions, 2021, 110: 379-393.
- [9] ZHOU H, LIU R, LI Y, et al. A rolling bearing fault diagnosis method based on a convolutional neural network with frequency attention mechanism [J]. Structural Health Monitoring, 2024, 23(4): 2475-2495.
- [10] LIU ZH Z, DING K, LIN H B, et al. Sparse coefficient fast solution algorithm based on the circulant structure of a shift-invariant dictionary and its applications for machine fault diagnosis [J]. Measurement, 2022, 203: 111943.
- [11] XING ZH, YI C, LIN J H, et al. Multi-component fault diagnosis of wheelset-bearing using shift-invariant impulsive dictionary matching pursuit and sparrow search algorithm[J]. Measurement, 2021, 178: 109375.
- [12] HEI ZH D, SHI Q, FAN X F, et al. Distance-guided domain adaptation for bearing fault diagnosis under variable operating conditions[J]. Measurement Science and Technology, 2024, 35(8): 086128.
- [13] YAO P, WANG J X, ZHANG F Y, et al. Intelligent rolling bearing imbalanced fault diagnosis based on mel-frequency cepstrum coefficient and convolutional neural networks[J]. Measurement, 2022, 205: 112143.
- [14] QI M S, ZHOU R, ZHANG Q Q, et al. Feature classification method of frequency cepstrum coefficient based on weighted extreme gradient boosting[J]. IEEE Access, 2021, 9: 72691-72701.
- [15] GUO M H, CAI J X, LIU ZH N, et al. Pct: Point cloud transformer[J]. Computational Visual Media, 2021, 7: 187-199.
- [16] FANG H R, DENG J, BAI Y X, et al. CLFormer: A lightweight transformer based on convolutional embedding and linear self-attention with strong robustness for bearing fault diagnosis under limited sample conditions[J]. IEEE Transactions on Instrumentation and Measurement, 2021, 71: 1-8.
- [17] ZHANG ZH H, WANG P, DING J M. Fault detection and analysis for wheelset bearings via improved explicit shift-invariant dictionary learning[J]. ISA Transactions, 2023, 136: 468-482.
- [18] CHEN J B, HUANG R Y, ZHAO K, et al. Multiscale convolutional neural network with feature alignment for bearing fault diagnosis [J]. IEEE Transactions on Instrumentation and Measurement, 2021, 70: 1-10.
- [19] XU J, LI Z SH, DU B W, et al. Reluplex made more practical: Leaky ReLU [C]. 2020 IEEE Symposium on Computers and Communications (ISCC). IEEE, 2020: 1-7.
- [20] SMITH W A, RANDALL R B. Rolling element bearing diagnostics using the case western reserve university data: a benchmark study[J]. Mechanical Systems and Signal Processing, 2015, 64: 100-131.
- [21] WANG B, LEI Y G, LI N P, et al. A hybrid prognostics approach for estimating remaining useful life of rolling element bearings[J]. IEEE Transactions on Reliability, 2018;69(1): 401-412.
- [22] KHARE S K, BAJAJ V. Time-frequency representation and convolutional neural network-based emotion recognition[J]. IEEE Transactions on Neural Networks and Learning Systems, 2020, 32(7): 2901-2909.

[23] ALEXEY D. An image is worth 16×16 words: Transformers for image recognition at scale[J]. ArXiv preprint arXiv: 2010.11929, 2020.

[24] CHEN Y SH, JIANG H L, LI CH Y, et al. Deep feature extraction and classification of hyperspectral images based on convolutional neural networks[J]. IEEE Transactions on Geoscience and Remote Sensing, 2016, 54 (10): 6232-6251.

[25] ZHANG AN S, LI SH B, CUI Y X, et al. Limited data rolling bearing fault diagnosis with few-shot learning[J]. IEEE Access, 2019, 7: 110895-110904.

[26] XU P H, ZHANG L Y. A fault diagnosis method for rolling bearing based on 1D-ViT model [J]. IEEE Access, 2023, 11: 39664-39674.

[27] ZHANG F Q, YIN ZH Y, XU F L, et al. MAB-DrNet: Bearing fault diagnosis method based on an improved dilated convolutional neural network[J]. Sensors, 2023, 23(12): 5532.

[28] MENG Z, LUO C, LI J, et al. Research on fault

diagnosis of rolling bearing based on lightweight model with multiscale features [J]. IEEE Sensors Journal, 2023, 23(12): 13236-13247.

[29] ZHAO Y J, BAO W M, XU X ZH. Rolling bearing fault diagnosis with distribution shift data using improved spatial distribution filters and constraint feature extraction [J]. Journal of Electrical Engineering & Technology, 2024: 1-15.

作者简介



赵运基 (通信作者), 2012 年于华南理工大学获得博士学位, 现为河南理工大学副教授, 主要研究方向为模式识别、信号处理与故障诊断。
E-mail: auyjz@hpu.edu.cn

Zhao Yunji (Corresponding author) received his Ph. D. degree from South China University of Technology in 2012. Now he is an associate professor in Henan Polytechnic University. His main research interests include pattern recognition, signal processing and bearing fault diagnosis.