

采用递归图编码技术与残差网络的 滚动轴承故障诊断

张龙, 胡燕青, 赵丽娟, 张号, 王晓博, 文培田

(华东交通大学轨道交通基础设施性能监测与保障国家重点实验室, 330013, 南昌)

摘要: 针对传统故障诊断方法未充分挖掘故障信号的时间序列间关联性特征的问题, 将递归图编码技术引入故障诊断领域, 提出了递归图编码技术与残差网络的滚动轴承故障诊断模型。采用递归图编码方式将振动信号转换为增强信号特征的二维纹理图像; 将这些特征图像输入残差网络中, 结合残差网络对二维图像数据优秀的自适应特征提取能力, 对滚动轴承进行故障诊断。使用凯斯西储大学轴承数据集和某局机务段采集的真实机车轴承数据进行试验验证, 结果表明: 所提模型对轴承故障诊断的识别准确率为 99.99% 和 99.83%; 在输入不同的数据长度和变工况的试验中, 所提模型均保持了良好的故障诊断效果; 对比其他常见的故障诊断方法, 所提模型拥有更好的泛化性能和识别准确率。

关键词: 滚动轴承; 故障诊断; 时间序列; 递归图; 残差网络

中图分类号: TH13; TH17 **文献标志码:** A

DOI: 10.7652/xjtuxb202302012 **文章编号:** 0253-987X(2023)02-0110-11

Fault Diagnosis of Rolling Bearings Using Recurrence Plot Coding Technique and Residual Network

ZHANG Long, HU Yanqing, ZHAO Lijuan, ZHANG Hao, WANG Xiaobo, WEN Peitian

(State Key Laboratory of Performance Monitoring and Protecting of Rail Transit Infrastructure,
East China Jiaotong University, Nanchang 330013, China)

Abstract: Traditional fault diagnosis methods do not fully exploit the correlation characteristics between the time series of fault signals. To address this problem, recurrence plot coding technique is introduced into fault diagnosis, and a rolling bearing fault diagnosis model using this technique and residual network is proposed. Recurrence plots coding converts vibration signals into 2D texture images with enhanced signal features. With these feature images imported, the residual network will exert its excellent capacity to extract adaptive features of 2D image data and carry out fault diagnosis on rolling bearings. Tests are conducted using the Case Western Reserve University bearing dataset and real locomotive bearing data collected from a bureau locomotive depot for validation. In tests, the proposed model yields accuracy of 99.99% and 99.83% in bearing fault detection and delivers good fault diagnosis results with different data lengths and variable operating conditions input. To sum up, the proposed model has better generalization per-

收稿日期: 2022-06-21。 作者简介: 张龙(1980—), 男, 副教授, 硕士生导师。 基金项目: 国家自然科学基金资助项目(51665013); 江西省自然科学基金资助项目(20212BAB204007); 江西省研究生创新资金资助项目(YC2020-S335; YC2021-S42)。

网络出版时间: 2022-09-02

网络出版地址: <https://kns.cnki.net/kcms/detail/61.1069.T.20220901.1806.006.html>

formance and detection accuracy than other common fault diagnosis methods.

Keywords: rolling bearing; fault diagnosis; time series; recurrence plots; residual networks

滚动轴承是旋转机构的关键元件,其健康状态会对机械设备的性能、稳定性和寿命产生重大影响^[1-2]。对于滚动轴承不同种类的故障,其振动信号有明显的差异,通过对振动信号进行实时监测能够有效避免重大事故的发生。因此,建立一种实用的方法来准确、可靠地诊断原始振动信号与滚动轴承故障模式之间复杂的不确定相关性具有重要的现实意义。

机器学习方法被广泛应用于滚动轴承的故障诊断。李兵等^[3]提出了一种改进随机森林的故障诊断模型,分别改进了随机森林方法中的决策树方法及集成投票方法,能够对轴承故障状态进行有效识别。田晶等^[4]提出了一种将集成经验模态分解(EEMD)和空域相关降噪相结合的滚动轴承故障诊断方法,实现了对滚动轴承故障的精确诊断。何大伟等^[5]建立了基于代价敏感支持向量机(CS-SVM)的故障诊断模型,采用改进 FOA 方法(IFOA)提升了对滚动轴承故障种类的识别率。

然而,传统的机器学习方法在故障诊断前需要人工提取浅层信号特征^[6],不能自动生成原始数据的特征向量。特征提取作为一项烦琐的工作,严重依赖专家经验。将人工提取的特征手动输入分类器进行故障诊断,可能无法充分反映故障模式,导致故障诊断模型准确度低、识别效果差。这些在特征提取时面临的问题极大地影响了机器学习方法的结果^[7]。

深度学习采用端对端的方式将特征提取与故障诊断结合在一个模型中^[8-12],避免了传统机器学习方法的缺点。考虑到振动信号是一维的时间序列信号,因此张龙等^[13]将一维的信号直接输入卷积神经网络(CNN),实现对滚动轴承的精确诊断。此外,利用双通道的 CNN 对轴承进行故障特征融合识别也取得了优秀的诊断效果^[14]。由于 CNN 的卷积层具有权值共享和局部连接的特点,因此对二维图像进行特征提取时能获得优秀的效果。所以,将原始一维振动信号转换成二维图像数据并进行故障诊断引起了广泛的研究。刘炳集等^[15]利用傅里叶变换将信号构造为二维时频图,输入 CNN 完成故障分类;Wen 等^[16]提出了一种将信号转换为灰度图的方法,然后将灰度图作为 CNN 的输入向量进行故障诊断;全钰等^[17]对一维振动信号使用格拉姆角场(GAF)变换,然后利用 CNN 对生成的图像进行识别分类。

尽管上述研究取得了良好的诊断结果,但 CNN 网络模型存在网络层数较浅,无法提取更深层的隐含特征,这限制了故障诊断模型的精度。因此,赵小强等^[18]使用残差网络(ResNet)对滚动轴承进行故障诊断,并取得了优秀的诊断效果。但是,其仅将轴承的一维原始时域信号直接作为输入,并没有考虑 ResNet 网络在处理图像数据时拥有强大的自适应特征提取能力。

为充分利用时间序列数据点具有关联性的特征,本文使用递归图(RP)^[19]对轴承信号进行时间序列数据的特征分析,并结合 ResNet 对图像强大的自适应特征提取能力,提出了一种新的递归图编码技术与残差网络(RP-ResNet)模型。本文的主要贡献总结如下:

(1)将 RP 引入故障诊断领域,RP 图像编码方式增强了一维时域信号的特征信息,同时避免了故障信号的微弱信息丢失;

(2)利用 ResNet 对二维图像数据强大的自适应特征提取能力,充分提取 RP 二维纹理图像中的多层次的隐含特征。

本文使用凯斯西储大学(CWRU)的轴承数据集和某局机务段采集的真实机车轴承数据进行试验,RP-ResNet 的识别准确率达到 99.9% 和 99.83%。结果表明:在输入不同的数据长度和变工况的试验中,所提模型均保持了良好的故障诊断效果;对比其他常见的故障诊断方法,RP-ResNet 拥有更好的泛化性能和识别准确率。

1 准备知识

1.1 递归图

RP 是由 Eckmann 等在庞加莱定理的基础上提出的一种回复性可视化的方法。RP 可以将非线性动力系统的行为可视化。它从时间序列的内部结构中获得先验知识,解释了时间序列的相似性、信息量,并能够对信号进行可预测性分析。RP 是分析时间序列周期性、混沌性以及非平稳性的一种重要方法。RP 的本质是一种时-时的信号处理方法,被用于将时间序列中出现的循环行为编码成二维图像进行表示。时间序列数据可以使用独特的重复性行为(如周期性和不规则周期性方面)进行分类。此外,时间序列数据是以状态重复的形式生成的,这是

不断变化的不规则系统或随机过程的正常现象。RP 作为一种工具,用于通过重复出现的二维表示可视化和研究多维相空间轨迹^[20]。RP 的主要思想是揭示从当前状态到前一状态的轨迹运动,其可表述为

$$R_{i,j} = \theta(\epsilon - \|S_i - S_m\|),$$

$$S \in \mathbf{R}^n, l, m = 1, 2, \dots, K \quad (1)$$

其中: K 是 S 的状态数; ϵ 是距离的阈值; $\|\cdot\|$ 是范数; $\theta(\cdot)$ 是 Heaviside 函数。

为了在递归图像中找到可区分的模式, Dias 等^[21]研究使用了一种编码时间序列 T_x 状态的循环程度,有效地进行了预处理,其实现如下。

设 T 是一个有 n 个点的时间序列。定义 T 的 RP 表达式为 R_T , 计算式为 $R_{i,j} = f(T_i, T_j)$, 其中, $i, j \in \{1, 2, \dots, n\}$, f 是距离函数,用于编码时间序列 T_i 状态的循环程度。函数 f 的 3 种不同功能实现如下

$$R_{i,j} = |T_i - T_j| \quad (2)$$

$$R_{i,j} = \frac{T_i}{T_j} \quad (3)$$

$$R_{i,j} = T_i T_j \quad (4)$$

式中: T_i 是具有 n 个观测值的时间序列,并且 $i, j = 1, \dots, N$ 。式(2)~(4)提供具有 $n \times n$ 维的图像(每个方程一个),在本研究中采用了式(4)的图像编码计算方式,编码后的特征图像如图 1 所示。

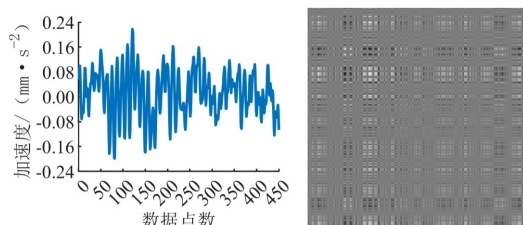


图 1 时域波形图和使用递归图像编码方式的特征可视化图像

Fig. 1 Time domain waveforms and feature visualizations using the RP image encoding method

1.2 残差网络

作为一种深度学习网络, CNN 的层数对学习效果具有关键作用。CNN 的网络层数加深意味着可以更充分地提取不同层次的特征,网络越深入,提取到的特征维度越高,获得的信息量也就越多。但是,简单粗暴地加深 CNN 的网络层数会导致梯度弥散或者梯度爆炸,从而导致模型训练准确率下降。

残差网络^[22]的出现极大程度地消除了这个现象。这归功于 ResNet 引入了一种全新的结构:每两层之间增加一个短路机制,形成了一个残差块,如图 2 所示。

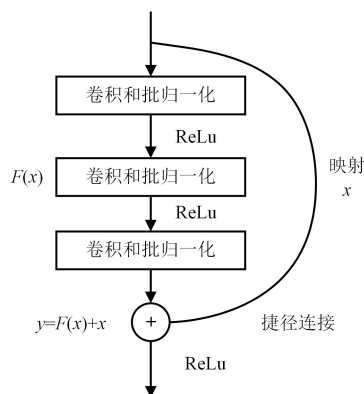


图 2 残差块结构

Fig. 2 Residual block structure diagram

残差块的优势在于提出了两种编排组织方式(mapping):一种是残差映射(residual mapping),指图 2 中直线连接的部分;另一种是恒等映射(identity mapping),指图 2 中的旁路连接,旁路连接会在神经网络中形成一种捷径连接(shortcut connection)。残差映射将得到输出 $F(x)$,而捷径连接得到的输出为 x ,所以该残差块最后输出值为 $y = F(x) + x$ 。

假设输入是一个图片 X_0 ,经过 L 层的神经网络,其中第 i 层的非线性变换记为 H_i , H_i 允许是各种函数的叠加运算,如卷积(Conv)、池化(Pooling)、独立同分布(Batch Normalization)、Sigmoid、ReLU 等激活函数。经过各种函数的累加运算,将第 i 层的特征输出记作 X_i 。

传统的卷积神经网络使用第 i 层的输出 X_i 作为第 $i+1$ 层的输入,即 $X_i = H_i(X_{i-1})$,而 ResNet 增加了旁路连接,输出为

$$X_i = H_i(X_{i-1}) + X_{i-1} \quad (5)$$

残差块的作用是当网络达到最优状态时,通过进一步深化网络结构将残差映射输出值压缩为 0,只留下恒等映射。恒等映射的编排方式允许梯度传播到前面的层。它能够识别每一层的输入,学习并形成残差函数。这种残差函数不仅加深了网络的层数,能够提取更丰富的特征集,而且也更容易优化,防止梯度弥散或者梯度爆炸的发生。理论上,网络将保持在最佳状况,随着深度的增加,在性能方面却不会出现负优化。残差网络-18 的结构参数如表 1 所示。

表 1 残差网络-18 的结构

Table 1 The structure of residual network-18

网络结构	参数	输出尺寸
卷积层 1	$7 \times 7, 64$, 步长 2 最大池化 3×3 , 步长 2	112×112
卷积层 2	$\begin{bmatrix} 3 \times 3, & 64 \\ 3 \times 3, & 64 \end{bmatrix} \times 2$	56×56
卷积层 3	$\begin{bmatrix} 3 \times 3, & 128 \\ 3 \times 3, & 128 \end{bmatrix} \times 2$	28×28
卷积层 4	$\begin{bmatrix} 3 \times 3, & 256 \\ 3 \times 3, & 256 \end{bmatrix} \times 2$	14×14
卷积层 5	$\begin{bmatrix} 3 \times 3, & 215 \\ 3 \times 3, & 512 \end{bmatrix} \times 2$	7×7
输出层	平均池化, 全连接层, SoftMax 分类器	1×4

2 RP-ResNet 的滚动轴承故障诊断模型

本文方法结合了递归图编码技术可以充分挖掘故障信号的时间序列间关联性特征和残差网络在处理图像数据时拥有的自适应特征提取能力的优势,其中,递归图编码技术与残差网络诊断方法流程如图 3 所示。

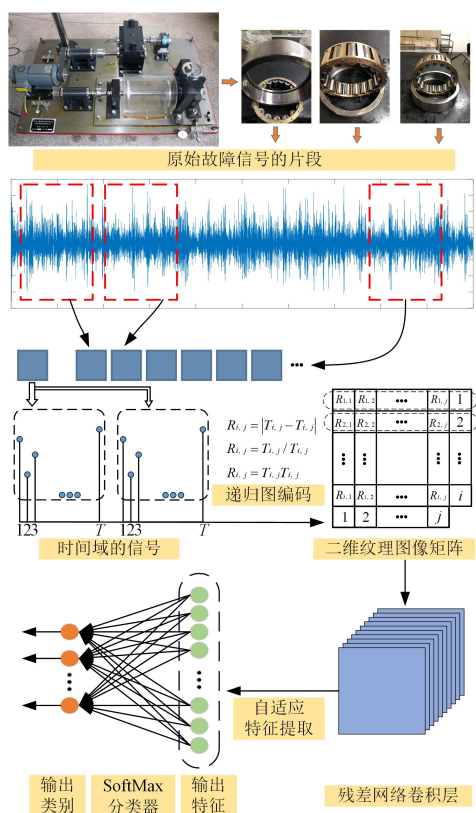


图 3 递归图编码技术与残差网络诊断方法流程

Fig. 3 Diagnostic flow chart using recurrence plot coding technique and residual network

首先,将采集到的原始故障信号使用无重叠的方式构造样本,按式(4)将一维的故障信号转换成增强了轴承故障特征的二维纹理图像矩阵,并输入 ResNet 网络的卷积层中进行自适应特征提取,得到不同故障类别的滚动轴承特征;然后,利用 SoftMax 分类器建立它们和对应故障类别之间的映射关系,从而完成故障诊断。

图 4 为 RP-ResNet 模型训练流程。首先,将采集到的轴承原始振动信号使用无重叠的方式构造样本;然后,使用 RP 图像编码方式转换成增强轴承故障特征的二维纹理图像矩阵,按照设定的 8:1:1 划分为训练集、验证集和测试集。随后,利用训练集和验证集对 ResNet 网络进行训练,获取最优参数,并保存模型。最后,将测试集数据输入到训练好的 ResNet 模型中进行故障诊断,并获取轴承故障分类结果和识别准确率。

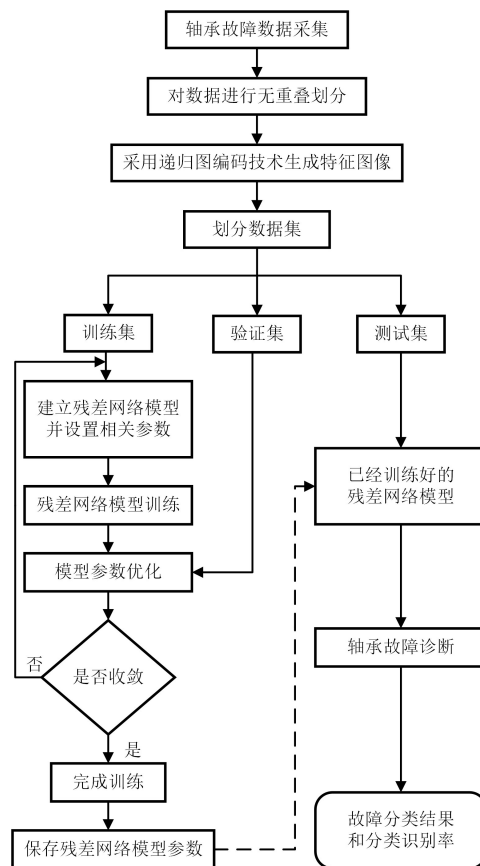


图 4 RP-ResNet 模型训练流程

Fig. 4 Flowchart of RP-ResNet model training

3 案例研究与试验结果分析

本节对 CWRU 的轴承数据集和某局机务段的 JL-501 机车真实轴承数据集进行了故障诊断分析。

试验运行的软件环境为 Windows10 64 位操作系统,运行内存为 16 GB,在 PyCharm 平台上使用 pytorch 框架,硬件环境为 AMD Ryzen 7 5800H, RTX 3070 Laptop GPU 显卡。使用残差网络-18 结构,结构参数如表 1,其他设置如下:

(1)使用 Adam 优化器^[23],初始学习率设置为默认值 0.001;

(2)采用学习率衰减机制^[24],设置模型每训练 10 代,学习率乘以衰减系数 0.9,以此解决一些不稳定情况,并提升模型的稳定性。

为避免偶然性结果,每种数据长度均试验 100 次,取平均值作为最终试验结果。

3.1 案例 1:凯斯西储轴承数据集故障诊断

为了衡量所提出的 RP-ResNet 模型性能,在故障试验中使用了 CWRU 的轴承故障数据^[25]。

CWRU 的轴承故障模拟试验台如图 5 所示,主要由一个 2×745.7 W 的电机、一个扭转编码器、一个功率计和控制电子单元组成。使用 16 通道采样仪器采集滚动轴承的振动信号,频率为 12 kHz。选用轴承驱动端的数据进行试验,轴承故障采用电火花加工,故障深度均为 0.28 mm,故障直径分别为 0.18、0.36、0.54 mm,共 9 种故障类型。

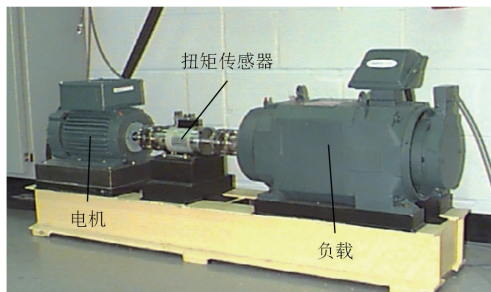


图 5 CWRU 轴承故障试验台

Fig. 5 CWRU bearing failure test bench

(1)不同数据长度下模型性能分析。本文将一维振动信号使用非重叠的方式截取成相同长度的数据片段。因此,在试验前需确定不同数据长度对 RP-ResNet 模型的故障诊断效果的影响。每个数据片段的长度参考以下两个标准进行选取:①数据长度应足以捕捉信号的局部特征;②为了减少模型的训练时间,应尽量缩短样本的数据长度。

当电机载荷为 2×745.7 W 时,电机的转速为 1 750 r/min,由此可计算出轴承的每个周期约包含 411 个信号点。按照 8:1:1 的比例将所有样本划分为训练集、验证集与测试集,10 类不同状态的试验样本构造如表 2 所示(以样本长度设置为 400 个

信号点为例)。

表 2 滚动轴承试验样本构造

Table 2 Rolling bearing sample structure

样本类型	样本长度	样本数	标签
正常	450	265	C0
内圈故障 0.18 mm	450	265	C1
内圈故障 0.36 mm	450	265	C2
内圈故障 0.54 mm	450	265	C3
外圈故障 0.18 mm	450	265	C4
外圈故障 0.36 mm	450	265	C5
外圈故障 0.54 mm	450	265	C6
滚动体故障 0.18 mm	450	265	C7
滚动体故障 0.36 mm	450	265	C8
滚动体故障 0.54 mm	450	265	C9

图 6 是数据长度对模型诊断精度的影响。可以看出,数据长度对模型的准确率影响较小,这是因为不同故障类型的轴承,其振动信号的特征存在一定的细微差异。同时,RP 编码方式作为一种时-时信号处理方法,能够揭示信号从当前状态到前一状态的轨迹运动,对数据有一定的分析预测作用,并且能增强不同轴承故障类型之间的差异。

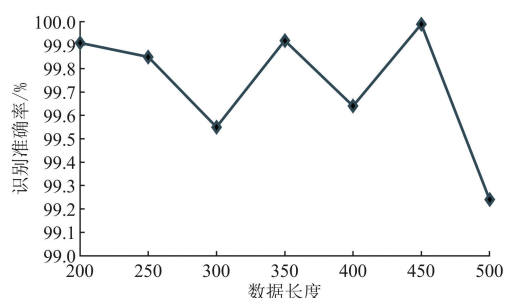


图 6 案例 1 中数据长度对模型诊断精度的影响

Fig. 6 The effect of data length on model diagnostic accuracy in case 1

图 7 是对数据长度为 200 的样本使用 RP 编码所获得的二维图像。可以发现,RP 编码方式能有效地对时间序列的内部特征进行分析预测,并编码为特征具有明显区别的二维纹理图像。在合理的数据长度范围内,模型识别准确率均能达到 99.5% 以上;当数据长度为 450 点时,该模型的平均识别准确率达到 99.99%;当数据长度增长到 500 点时,识别准确率下降。这是因为使用 RP 编码方式所获得的二维图像中冗余信息增多,不能很好地反映数据的内在特征联系。同时,随着二维图像尺寸增大,导致 ResNet 网络的卷积层提取到的图像特征维度降低,进而造成识别准确率下降。

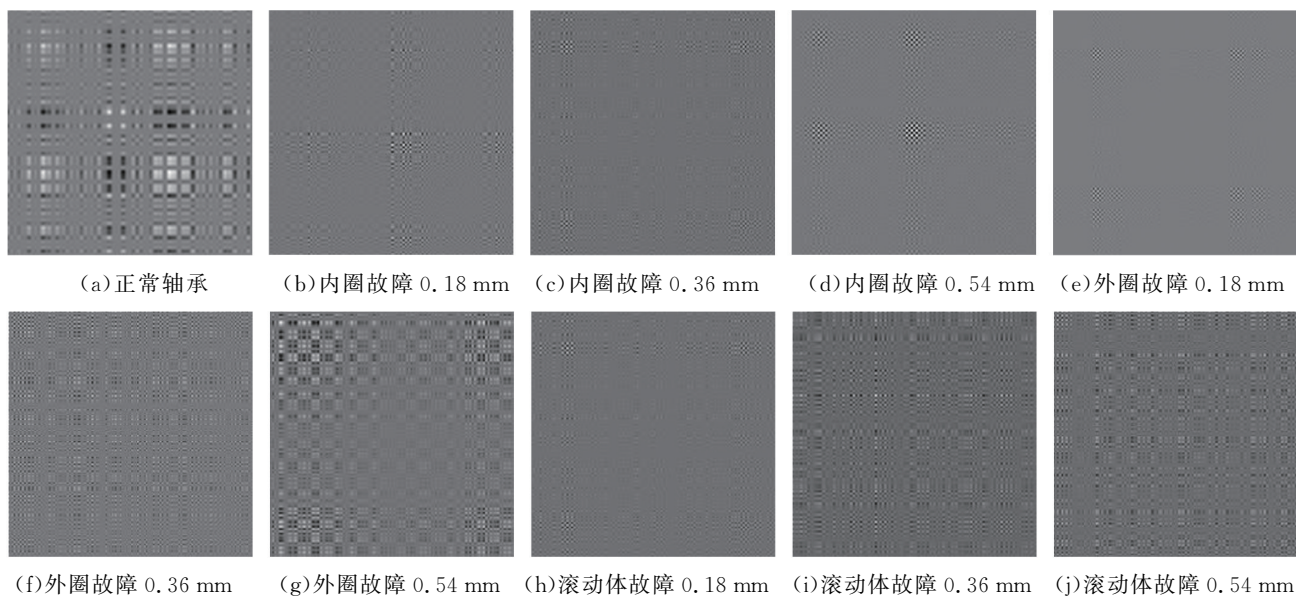
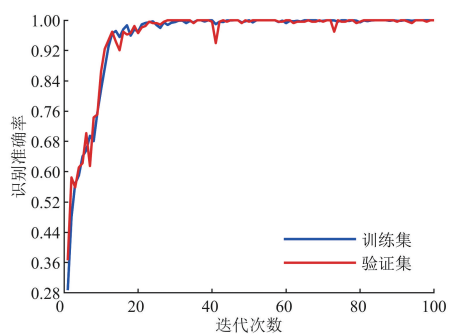


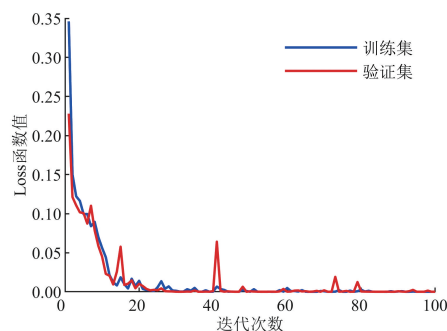
图7 使用递归图编码技术所得轴承故障特征图

Fig. 7 Bearing fault characteristic map using recurrence plot coding technique

图8是使用表2的滚动轴承试验数据,输入RP-ResNet模型中经过100轮训练后得到的识别准确率和Loss函数曲线。可以看出,当模型第80次训练时,训练集和验证集的精度为100%,Loss函数值减小到0.01以下,之后准确率和Loss函数曲线均逐步平缓,模型达到稳定收敛状态。本文RP-ResNet故障诊断模型在该试验数据中达到了较为满意的分类效果。



(a) 识别准确率变化曲线



(b) Loss 函数变化曲线

图8 案例1中模型诊断曲线

Fig. 8 Model diagnostic curve in case 1

为了进一步考察 RP-ResNet 模型对轴承故障识别的能力,引入混淆矩阵对诊断结果进行量化分析。混淆矩阵详细地反映了滚动轴承不同故障类型下的误判数量和误判类型,混淆矩阵分类结果可视化如图9所示。使用表2试验数据进行了100次重复试验,总计2650个试验样本。试验结果显示,仅3个0.18 mm外圈故障样本被误分为0.56 mm外圈故障,故障误诊类型为轻度故障,误判为同类型故障的重度故障,具有一定的预警作用。该模型在滚动轴承的其他类故障诊断中均获得了较高的识别准确率,由此说明本文所提出的 RP-ResNet 故障诊断模型在滚动轴承中具有优秀的故障诊断效果。

C0	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
C1	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
C2	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
C3	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00
C4	0.00	0.00	0.00	0.00	0.99	0.00	0.01	0.00	0.00	0.00
C5	0.00	0.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00
C6	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00
C7	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.00	0.00
C8	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.00
C9	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00
	C0	C1	C2	C3	C4	C5	C6	C7	C8	C9

图9 滚动轴承分类混淆矩阵

Fig. 9 Rolling bearing classification confusion matrix

为了更直观地展现模型的特征提取能力,使用 t -sne 进行特征可视化,结果如图 10 所示。可以看出,测试集训练前各种故障的特征分类并不明显,有较大的重叠区域。

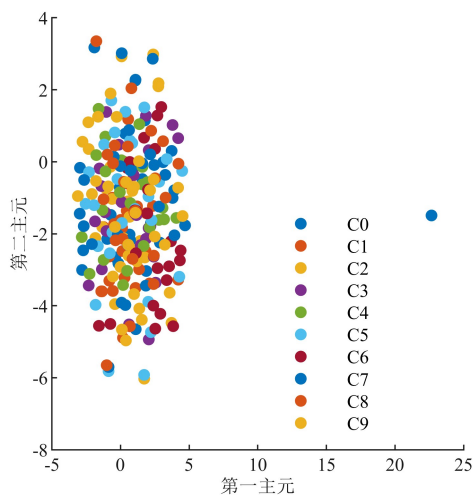


图 10 数据自适应特征提取前的降维图

Fig. 10 Dimensionality reduction map before data adaptive feature extraction

图 11 为测试集使用本文所提 RP-ResNet 故障诊断模型进行训练后,进行自适应特征提取所获得的输出特征的特征降维图。可以看出,各种故障类型没有重叠区域,特征分类明显,区分度有明显的提高。

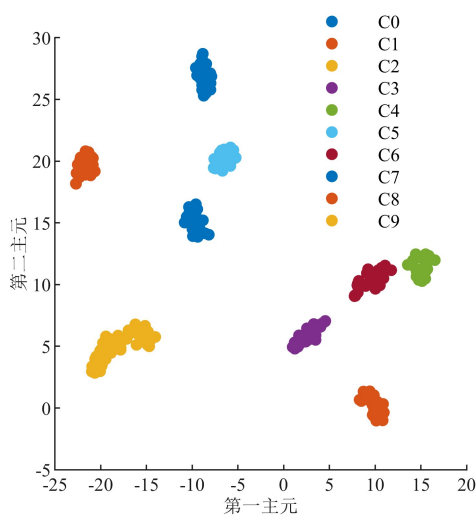


图 11 数据自适应特征提取后的输出特征降维图

Fig. 11 Reduced dimensional map of output features after data adaptive feature extraction

综合图 10 与图 11 可以看出,本文所提 RP-ResNet 故障诊断模型能够将测试集中滚动轴承的 10 种不同状态的特征正确分离,从而得到较高的识别准确率。

(2)变工况下的模型性能分析。为进一步验证本文所提 RP-ResNet 故障诊断模型的效果,将 CWRU 轴承故障数据按照电机负载 $0 \sim 3 \times 745.7 \text{ W}$ (电机转速为 $1\,797 \sim 1\,730 \text{ r/min}$) 分为 4 组数据。为验证不同数据长度对 RP-ResNet 模型的影响,同时兼顾模型的训练速度,对 CWRU 的 4 组电机负载轴承故障数据每段数据长度取 200 个数据点进行试验,总共 600 个样本,其中训练、验证、测试集按照 8:1:1 划分,样本个数分别为 480、60、60。不同电机负载的试验样本标签与样本类型对应关系与表 2 一致。使用所提的 RP-ResNet 模型进行试验,得到变工况下故障诊断识别率曲线,如图 12 所示。

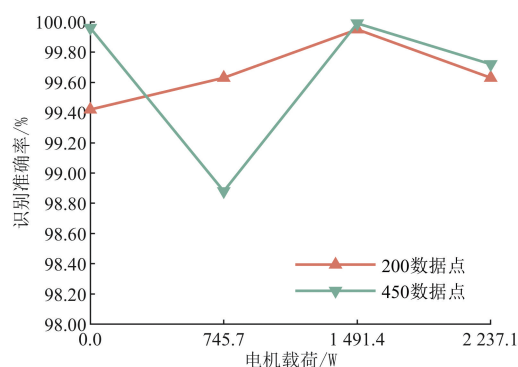


图 12 变工况下模型的准确率

Fig. 12 Accuracy of the model under different working conditions

从图 12 可知,在数据长度分别选取 200 与 450 时,RP-ResNet 模型依旧有良好的诊断稳定性。使用两种不同数据长度的编码图在不同电机载荷下进行试验,故障识别准确率保持了较高的水准,平均准确率均达到 98.85% 以上。450 的数据长度样本在 $1 \times 745.7 \text{ W}$ 电机负载工况下的识别率表现不佳,为 98.88%,可能是因为该工况条件复杂,振动信号的噪声过多。同时,较长数据长度的样本使用 RP 编码形成二维图像中冗余信息较多,这影响了 RP-ResNet 模型故障诊断的效果。综合以上试验,可以说明所提出的 RP-ResNet 模型在不同的数据长度、不同的工况下诊断都是可靠的、有效的。

(3)不同特征图像对诊断结果的影响。为进一步说明本文方法的优越性,验证 ResNet 对 RP 编码图的特征学习和故障分类能力,参考其他几种图像编码方式(时频图^[14]、灰度图^[15]、GAF^[17]),并将其作为深度残差网络的输入向量进行试验。使用表 2 中的数据分别进行编码处理,并输入 ResNet 网络。为了避免干扰,取 10 次诊断的平均结果作为滚

动轴承最终准确率,结果如下:时频图+ResNet、GAF+ResNet、灰度图+ResNet、RP+ResNet 的识别准确率分别为 98.81%、98.85%、97.85%、99.99%。可以看出:灰度图的编码方式简单,平均识别率略低于其他方法;时频图的高识别准确率建立在需要人工设置采样点数并且控制好输出图片尺寸的基础上,如果图片尺寸不合适,则识别准确率会大幅下降;GAF 的图像编码方式能够通过角度透视识别出不同时间间隔内的时间相关性,但其准确率略低于本文方法。与其他几种常见的图像编码方式对比,RP 图像编码方式结合 ResNet 网络能够获取更优秀的滚动轴承故障识别效果。

(4)不同分类器方法对诊断结果的影响。为了验证 RP 与 ResNet 结合的优势,使用朴素贝叶斯(NBM)、支持向量机(SVM)、多层感知器(MLP)、BP 神经网络、循环神经网络(RNN)、经典卷积神经网络 LeNet-5 以及 VGG-16 等分类方法进行试验对比,结果如表 3 所示。为了避免偶然误差,准确率取 10 次试验的平均值。

可以看出:MLP 在试验数据集上获得的准确率最低,其主要原因是多层感知器作为一种早期人工神经网络,使用的隐层层数较少,无法对 RP 编码图进行深层次的特征识别提取;随着网络结构中隐层的数目增加,BP 神经网络对 RP 编码图的特征识别效果也略优于 MLP。对比上述两种方法,RNN 作为一种递归神经网络,对 RP 的编码图分类识别有一定的优势,分类效果略优于上述两种,但是受限于网络层数的原因,不能够进一步识别更深层次的特征。贝叶斯方法的优点是,当数据量很大时,可以展现出较高的分类精度,但是当特征数量较多或特征间的相关性较大时,分类准确率会出现下滑。SVM 方法的优势在于使用核函数减少了大量的映射计算,但是当图像的特征过多时,会导致分类效果不佳。

表 3 不同分类器方法的模型识别准确率

Table 3 Model accuracy of different classifier methods

分类方法	平均识别准确率/%
RP+MLP	84.53
RP+BP	88.30
RP+RNN	88.68
RP+NBM	92.00
RP+SVM	92.00
RP-LeNet5	97.36
RP+VGG16	98.53
RP+ResNet	99.99

将 RP 编码获得的特征图像作为输入,使用卷积神经网络进行故障识别,准确率均能够达到 97.5%以上。同时发现,随着网络层数的依次加深(LeNet5、VGG、ResNet),网络对 RP 特征图像也会有更高的识别准确率。综上所述,与其他几种传统的机器学习方法相比,RP 图像编码方式结合 ResNet 方法网络能够获取更优秀的滚动轴承故障识别效果。

3.2 案例 2:机车轴承故障诊断

(1)不同数据长度下模型性能分析。试验机车轴承故障数据^[13]来自于某局机务段的 JL-501 机车轴承试验台,如图 13 所示。机车轴承试验台主要由主轴箱、液压系统、电气系统和台身组成。在主轴箱上对待测轴承进行检测安装、旋转、径向加载等流程,液压系统则实现对待测轴承进行径向加载和卸载。电源使用 380 V、50 Hz 三相四线交流电;电机驱动总功率为 5 kW,试验转速设置为 500 r/min。试验所用机车轴承为 NJ2232WB 系列圆柱滚子轴承,轴承内径为 160 mm,外径为 290 mm。在 20 kHz 的采样频率下,使用三通道(加速度传感器 A、B、C)采集机车轴承的振动信号。

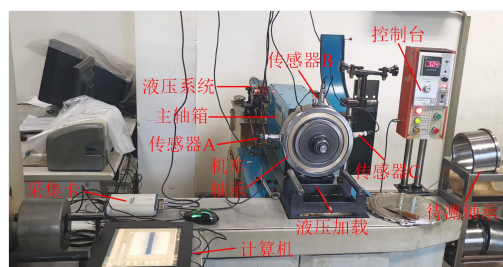


图 13 JL-501 机车轴承故障诊断试验台

Fig. 13 JL-501 locomotive bearing fault diagnosis test bench

对传感器 B 采集的 6 种机车轴承振动信号数据——正常轴承、内圈故障、外圈故障、滚动体故障、保持架故障、保持架滚动体复合故障依次编号为 D1~D6,并使用无重叠的采样方式进行样本构造,样本数量均为 300 个。

参考案例 1 的试验,并结合文献 [17],可以得出:图像样本构造的数据长度未包含轴承旋转一圈的数据点,识别准确率也有可能为试验最佳值。针对机车轴承数据,在数据长度为 200~500 时,进行样本构造并试验,结果如图 14 所示。可以看出,不同数据长度的划分对模型的准确率影响较小,与案例 1 结论一致。当数据长度为 250~500 时,模型识别准确率均能达到 98.5%以上;当数据长度为 400 和 500 时,模型识别准确率达到 99.83%和 99.78%。当数据长度为 200 时,数据长度较短,所含信息量较少,导致 RP

无法对信号进行较好的分析预测,因此编码获得的特征图不能充分反映数据的内在特征联系,所以输入 ResNet 进行诊断时,识别准确率仅为 97.45%。

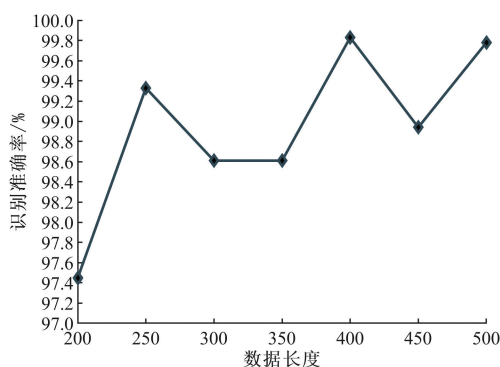
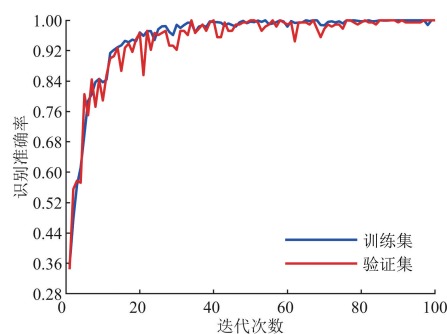


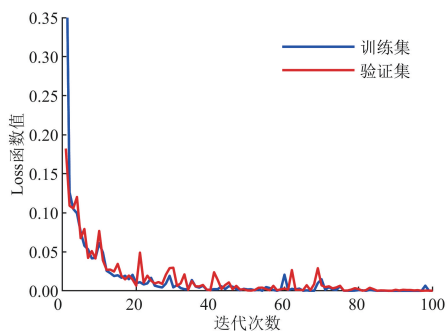
图 14 案例 2 中数据长度对模型诊断精度的影响

Fig. 14 The effect of data length on model diagnostic accuracy in case 2

图 15 是使用案例 2 的机车轴承试验数据,输入 RP-ResNet 模型中经过 100 轮训练后得到的准确率和 Loss 函数曲线。可以看出,在第 80 次训练时,训练集和验证集的精度已经高达 100%, Loss 函数减小到 0.01 以下,之后的准确率和 Loss 函数曲线均逐步平缓,模型达到稳定收敛状态。由此说明,本文 RP-ResNet 故障诊断模型在该试验数据中达到了较为满意的分类效果。



(a) 识别准确率变化曲线



(b) Loss 函数变化曲线

图 15 案例 2 中模型诊断曲线

Fig. 15 Model diagnostic curves in case 2

机车轴承故障诊断试验的三维混淆矩阵如图 16 所示。分析可知,在 300 个测试样本中,有 5 个保持架和滚动体的复合故障被误分为内圈故障。这是因为真实场景中机车轴承运行环境复杂,保持架和滚动体的动态特性更为复杂,因此导致保持架和滚动体的复合故障类型易发生误判。但是,该模型在机车轴承的其他类故障诊断中均获得了较高的识别准确率,说明 RP-ResNet 模型在真实机车轴承故障诊断中具有优异的故障识别能力。

D1	1.00	0.00	0.00	0.00	0.00	0.00
D2	0.00	0.99	0.00	0.00	0.00	0.01
D3	0.00	0.00	1.00	0.00	0.00	0.00
D4	0.00	0.00	0.00	1.00	0.00	0.00
D5	0.00	0.00	0.00	0.00	1.00	0.00
D6	0.00	0.00	0.00	0.00	0.00	1.00
	D1	D2	D3	D4	D5	D6

图 16 案例 2 中三维混淆矩阵

Fig. 16 Three-dimensional confusion matrix in case 2

(2)不同特征图像对诊断结果的影响。使用案例 2 机车轴承数据,采用不同图像编码方式结合 ResNet 网络进行故障诊断,结果如下:时频图 + ResNet 的识别准确率为 97.56%, GAF + ResNet 的准确率为 97.72%,灰度图 + ResNet 的识别准确率为 97.56%, RP + ResNet 的识别准确率最优,为 99.83%。

(3)不同分类器方法对诊断结果的影响。为了验证 RP 与 ResNet 结合的优势,使用案例 2 的机车轴承数据,将 RP 编码后的二维图像数据输入传统的机器学习方法和简单的深度学习网络中。为了避免偶然误差,准确率取 10 次试验的平均值,试验结果如下: RP + MLP、RP + BP、RP + RNN、RP + NBM、RP + SVM、RP-LeNet5、RP + VGG16、RP + ResNet 的识别准确率分别为 71.67%、76.11%、72.78%、67.00%、74.00%、96.11%、95.94%、99.83%。ResNet 结合 RP 图像编码方式的识别准确率最高。

4 结 论

本文针对传统故障诊断方法未充分挖掘故障信

号的时间序列间关联性特征的问题,将 RP 图像编码方式引入滚动轴承故障诊断中,提出了一种 RP-ResNet 的滚动轴承故障诊断方法。使用 RP 图像编码方式将原始的一维振动信号数据转换成二维图像特征数据,避免了故障信号中所包含的微弱特征信息丢失的问题。同时,结合 ResNet 对图像数据优秀的自适应特征提取能力,实现了滚动轴承的故障诊断。本文结论如下。

(1) 本文将 RP 引入故障诊断领域,提出了 RP-ResNet 的滚动轴承故障诊断模型。RP 编码方式利用时间序列数据点具有关联性的特征,从时间序列的内部结构中获得特征信息,能对时间序列信号进行分析预测。结合 ResNet 对图像数据的优秀自适应特征提取能力,对滚动轴承进行故障诊断。在 CWRU 的轴承数据集和某局机务段采集的真实机车轴承数据集上试验,识别准确率分别为 99.99% 和 99.83%。

(2) 本文模型在不同的数据长度下均能够取得较高的识别准确率,这是因为不同故障类型的轴承振动信号特征存在一定的细微差异,RP 编码方式则增强了信号中包含的微弱特征。这在很大程度上消除了人工特征提取时专家经验可能会带来的误差。同时,在变工况和不同数据集的试验中均获得了较高的诊断准确率,有效地验证了模型的泛化能力。

(3) 与不同的图像编码方式和不同的分类器方法进行试验对比,结果表明:使用 RP 的图像编码转换方式识别准确率高于其他图像编码方法;与多种分类器方法的试验对比发现,ResNet 对使用 RP 图像编码所获得的二维特征图像能够进行更深层次的学习,从而获得更高的分类识别准确率。

(4) ResNet 结构的层数较多、网络较深。为了获得相对稳定、优异的诊断结果,需要较长时间进行模型训练。接下来,考虑引入迁移学习加快方法模型的训练。

参考文献:

- [1] MA Zhonghai, WANG Shaoping, SHI Jian, et al. Fault diagnosis of an intelligent hydraulic pump based on a nonlinear unknown input observer [J]. Chinese Journal of Aeronautics, 2018, 31(2): 385-394.
- [2] HOANG D T, KANG H J. A survey on deep learning based bearing fault diagnosis [J]. Neurocomputing, 2019, 335: 327-335.
- [3] 李兵, 韩睿, 何怡刚, 等. 改进随机森林方法在电机轴承故障诊断中的应用 [J]. 中国电机工程学报, 2020, 40(4): 1310-1319.
- [4] LI Bing, HAN Rui, HE Yigang, et al. Applications of the improved random forest algorithm in fault diagnosis of motor bearings [J]. Proceedings of the CSEE, 2020, 40(4): 1310-1319.
- [5] 田晶, 王英杰, 王志, 等. 基于 EEMD 与空域相关降噪的滚动轴承故障诊断方法 [J]. 仪器仪表学报, 2018, 39(7): 144-151.
- [6] TIAN Jing, WANG Yingjie, WANG Zhi, et al. Fault diagnosis for rolling bearing based on EEMD and spatial correlation denoising [J]. Chinese Journal of Scientific Instrument, 2018, 39(7): 144-151.
- [7] 何大伟, 彭靖波, 胡金海, 等. 基于改进 FOA 优化的 CS-SVM 轴承故障诊断研究 [J]. 振动与冲击, 2018, 37(18): 108-114.
- [8] HE Dawei, PENG Jingbo, HU Jinhai, et al. Bearing fault diagnosis based on a modified CS-SVM model optimized by an improved FOA algorithm [J]. Journal of Vibration and Shock, 2018, 37(18): 108-114.
- [9] 张西宁, 李霖, 刘书语, 等. 基于能量峰定位的经验小波变换及在轴承微弱故障诊断中的应用 [J]. 西安交通大学学报, 2021, 55(8): 1-8.
- [10] ZHANG Xining, LI Lin, LIU Shuyu, et al. Empirical wavelet transform based on energy peak location with applications to bearing weak fault diagnosis [J]. Journal of Xi'an Jiaotong University, 2021, 55(8): 1-8.
- [11] WANG Jinjiang, MA Yulin, ZHANG Laibin, et al. Deep learning for smart manufacturing: methods and applications [J]. Journal of Manufacturing Systems, 2018, 48, Part C: 144-156.
- [12] LECUN Y, BENGIO Y, HINTON G. Deep learning [J]. Nature, 2015, 521(7553): 436-444.
- [13] ZHAO Rui, YAN Ruqiang, CHEN Zhenghua, et al. Deep learning and its applications to machine health monitoring [J]. Mechanical Systems and Signal Processing, 2019, 115: 213-237.
- [14] 李益兵, 王磊, 江丽. 基于 PSO 改进深度置信网络的滚动轴承故障诊断 [J]. 振动与冲击, 2020, 39(5): 89-96.
- [15] LI Yibing, WANG Lei, JIANG Li. Rolling bearing fault diagnosis based on DBN algorithm improved with PSO [J]. Journal of Vibration and Shock, 2020, 39(5): 89-96.
- [16] 温江涛, 闫常弘, 孙洁娣, 等. 基于压缩采集与深度学习的轴承故障诊断方法 [J]. 仪器仪表学报, 2018, 39(1): 171-179.
- [17] WEN Jiangtao, YAN Changhong, SUN Jiedi, et al.

- Bearing fault diagnosis method based on compressed acquisition and deep learning [J]. Chinese Journal of Scientific Instrument, 2018, 39(1): 171-179.
- [12] 周奇才, 沈鹤鸿, 赵炯, 等. 基于改进堆叠式循环神经网络的轴承故障诊断 [J]. 同济大学学报(自然科学版), 2019, 47(10): 1500-1507.
ZHOU Qicai, SHEN Hehong, ZHAO Jiong, et al. Bearing fault diagnosis based on improved stacked recurrent neural network [J]. Journal of Tongji University (Natural Science), 2019, 47(10): 1500-1507.
- [13] 张龙, 徐天鹏, 王朝兵, 等. 基于改进谱峭度与一维卷积神经网络的轴承故障诊断 [J]. 机械设计与研究, 2021, 37(4): 99-105.
ZHANG Long, XU Tianpeng, WANG Chaobing, et al. Bearing fault diagnosis based on an improved kurtogram kurtosis and one-dimensional convolution neural network [J]. Machine Design & Research, 2021, 37(4): 99-105.
- [14] 张龙, 甄灿灿, 熊国良, 等. 基于深度时频特征的机车轴承故障诊断 [J]. 交通运输工程学报, 2021, 21(6): 247-258.
ZHANG Long, ZHEN Canzhuang, XIONG Guoliang, et al. Locomotive bearing fault diagnosis based on deep time-frequency features [J]. Journal of Traffic and Transportation Engineering, 2021, 21(6): 247-258.
- [15] 刘炳集, 熊邦书, 欧巧凤, 等. 基于时频图和CNN的滚动轴承故障诊断 [J]. 南昌航空大学学报(自然科学版), 2018, 32(2): 86-91.
LIU Bingji, XIONG Bangshu, OU Qiaofeng, et al. Fault diagnosis of rolling bearing based on time-frequency representations and CNN [J]. Journal of Nanchang Hangkong University (Natural Sciences), 2018, 32(2): 86-91.
- [16] WEN Long, LI Xinyu, GAO Liang, et al. A new convolutional neural network-based data-driven fault diagnosis method [J]. IEEE Transactions on Industrial Electronics, 2018, 65(7): 5990-5998.
- [17] 仝钰, 庞新宇, 魏子涵. 基于GADF-CNN的滚动轴承故障诊断方法 [J]. 振动与冲击, 2021, 40(5): 247-253, 260.
TONG Yu, PANG Xinyu, WEI Zihan. Fault diagnosis method of rolling bearing based on GADF-CNN [J]. Journal of Vibration and Shock, 2021, 40(5): 247-253, 260.
- [18] 赵小强, 梁浩鹏. 使用改进残差神经网络的滚动轴承变工况故障诊断方法 [J]. 西安交通大学学报, 2020, 54(9): 23-31.
ZHAO Xiaoqiang, LIANG Haopeng. Fault diagnosis method for rolling bearing under variable working conditions using improved residual neural network [J]. Journal of Xi'an Jiaotong University, 2020, 54(9): 23-31.
- [19] ECKMANN J P, KAMPHORST S O, RUELE D. Recurrence plots of dynamical systems [J]. Europhysics Letters (EPL), 1987, 4(9): 973-977.
- [20] DIAS D, PINTO A, DIAS U, et al. A multi-representational fusion of time series for pixelwise classification [J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2020, 13: 4399-4409.
- [21] DIAS D, DIAS U, MENINI N, et al. Pixelwise remote sensing image classification based on recurrence plot deep features [C]//IGARSS 2019: 2019 IEEE International Geoscience and Remote Sensing Symposium. Piscataway, NJ, USA: IEEE, 2019: 1310-1313.
- [22] HE Kaiming, ZHANG Xiangyu, REN Shaoqing, et al. Deep residual learning for image recognition [C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ, USA: IEEE, 2016: 770-778.
- [23] 张西宁, 刘书语, 余迪, 等. 改进深度卷积神经网络及其在变工况滚动轴承故障诊断中的应用 [J]. 西安交通大学学报, 2021, 55(6): 1-8.
ZHANG Xining, LIU Shuyu, YU Di, et al. Improved deep convolutional neural network with applications to bearing fault diagnosis under variable conditions [J]. Journal of Xi'an Jiaotong University, 2021, 55(6): 1-8.
- [24] 蒋文斌, 彭晶, 叶阁焰. 深度学习自适应学习率方法研究 [J]. 华中科技大学学报(自然科学版), 2019, 47(5): 79-83.
JIANG Wenbin, PENG Jing, YE Geyan. Research on adaptive learning rate algorithm in deep learning [J]. Journal of Huazhong University of Science and Technology (Natural Science Edition), 2019, 47(5): 79-83.
- [25] Case Western Reserve University Bearing Data Center. Bearing data file [DS/OL]. [2022-06-01]. <https://csegroups.case.edu/bearingdatacenter/pages/download-data-file>.

(编辑 陶晴 杜秀杰)