

基于 GAF-MCNN 的轴承智能故障诊断方法研究*

张超^{1,2} 房颖涛^{1,3} 冯建睿⁴ 杨柯⁵ 何世烈^{1,6} 董志杰⁷

- (1.西北工业大学航空学院 西安 710072;2.飞行器基础布局全国重点实验室 西安 710072;
3.中国人民解放军 77110 部队 德阳 618000;4.宁夏回族自治区中卫市人民政府办公室 中卫 755099;
5.中国运载火箭技术研究院 北京 100076;6.工业和信息化部电子第五研究所 广州 510510;
7.中国电子信息产业集团有限公司第六研究所 北京 100083)

摘要:针对轴承微小故障信号非平稳非线性且易受背景噪声干扰的特点,提出了一种基于格拉姆角场和多尺度卷积神经网络(Gramian angular field and multi-scale convolutional neural network, GAF-MCNN)的智能故障诊断方法。首先,利用分段聚合近似算法对原始振动信号进行压缩降维预处理,以减少数据存储空间和提升计算效率;然后,利用格拉姆角场算法将一维序列信号转换为二维矩阵热图,二维化后的矩阵加强了原始振动信号间的时间关系,将时间维度编码到了矩阵结构中;最后,设计了基于多尺度卷积神经网络对故障进行高效快速智能诊断。实验结果表明,GAF-MCNN 诊断方法不仅克服了传统卷积神经网络诊断方法存在的计算效率较低的问题,而且诊断准确率优于单尺度卷积神经网络方法,具有较强的工程实用性。

关键词:分段聚合近似;格拉姆角场;卷积神经网络;故障诊断

中图分类号: TP182;TN06 **文献标识码:** A **国家标准学科分类代码:** 520.20

New fault diagnosis method of bearing based on GAF-MCNN

Zhang Chao^{1,2} Fang Yingtao^{1,3} Feng Jianrui⁴ Yang Ke⁵ He Shilie^{1,6} Dong Zhijie⁷

- (1. School of Aeronautics, Northwestern Polytechnical University, Xi'an 710072, China; 2. National Key Laboratory of Aircraft Design, Xi'an 710072, China; 3. Unit 77110 of PLA, Deyang 618000, China;
4. Office of Zhongwei Municipal People's Government in Ningxia Hui Autonomous Region, Zhongwei 755099, China; 5. China Academy of Launch Vehicle Technology, Beijing 100076, China;
6. China Electronic Product Reliability and Environmental Testing Institute, Guangzhou 510610, China;
7. The 6th Research Institute of China Electronics Corporation, Beijing 100083, China)

Abstract: An intelligent fault diagnosis method based on Gramian angular field and multi-scale convolutional neural network (GAF-MCNN) was proposed to solve the non-stationary, nonlinear and easily disturbed background noise of motor bearing micro-fault signals. Firstly, the piecewise aggregation approximation algorithm is used to compress and reduce the dimension of the original vibration signals to reduce the data storage space and improve the computational efficiency. Then, the one-dimensional sequence signals are converted into two-dimensional matrix heat maps using the gramian angular field algorithm. The two-dimensional matrix strengthens the time relationship between the original vibration signals and encodes the time dimension into the matrix structure. Finally, a multi-scale convolutional neural network is designed to diagnose the fault efficiently and quickly. An example of motor bearing fault diagnosis shows that GAF-MCNN method not only overcomes the problem of low computational efficiency of traditional convolutional neural network diagnosis methods, but also has better diagnostic accuracy than single-scale convolutional neural network method, and has strong engineering practicability.

Keywords: piecewise aggregate approximation; gramian angular field; convolutional neural network; fault diagnosis

收稿日期:2024-07-30

* 基金项目:国家级重点科研项目(JSZL2022607B002, JSZL202160113001, JCKY2021608B018)、国家重点研发计划项目(2023YFF0719100)、工业和信息化部项目(CEIEC-2022-ZM02-0249)资助

0 引言

轴承是当代机械设备中一种重要的零部件^[1],主要功能是支撑机械旋转体,降低旋转机械在运动过程中的摩擦系数,并保证它们的回转精度。一旦轴承失效,可能会危及设备的正常运行并造成经济损失,有时甚至导致严重的安全事故^[2]。因此,滚动轴承设备的故障诊断对于维持设备的正常运行和安全具有重要的意义^[3]。轴承故障诊断算法的核心在于特征提取和故障识别。特征提取算法包括快速傅里叶变换(FFT)^[4]、小波变换^[5]、经验模式分解^[6]、变异模式分解^[7]等。可用于故障识别的算法包括支持向量机(support vector machine, SVM)^[8]、反向传播神经网络^[9]、贝叶斯分类器^[10]、K近邻法^[11]等。

随着人工智能技术的不断发展,以深度学习为代表的驱动模型被广泛应用于故障诊断、机械设备的健康监测等领域中^[12-19]。深度学习作为机器学习的重要分支之一,由于其具有较强的非线性特征提取能力和出色的推广性能,加速了其在智能故障诊断中的应用。常用的深度学习学习方法包括基于自动编码器^[20]的模型、基于深度信念网络^[21]的模型、基于卷积神经网络(convolutional neural network, CNN)^[22]的模型、基于循环神经网络^[23]的模型和基于残差网络^[24]的模型。深度学习技术及其强大的自动特征学习能力为轴承的故障诊断提供有效的解决方案。深度学习的核心是特征学习,其目的是通过分层网络获取分层特征信息,从而解决了过去需要手动提取特征的重要问题。CNN是一种强大的深度学习方法,已成功应用于各种领域,例如计算机视觉、语音识别和自然语言处理等一维信号处理。但传统的CNN,其训练参数量大,因此训练耗时较长,网络计算效率较低、参数调试需人工经验指导。Jiang等^[25]提出了一种新的多层卷积神经网络识别滚动轴承故障的方法,以提高滚动轴承的故障识别精度。在没有任何手动特征提取的情况下,该方法的性能优于SVM和人工神经网络(artificial neural network, ANN)法。Pan等^[26]修改了标准CNN,根据嘈杂的振动数据识别电机轴承的故障。Jiao等^[27]提出了一个CNN网络,通过分析横向振动和逆向振动来检测行星变速箱故障。Li-ang等^[28]采用一维时域信号作为输入,将锯齿形扩展卷积引入到CNN中,可以有效地提高卷积层的感受野。Wang等^[29]根据故障可视化特征和特征自动提取的特征,提出了一种将对称点模式表示与支持挤压激励的CNN模型结合起来的轴承故障智能诊断方法,降低了模型冗余信息。但上述方法并未考虑在受到载荷变化、噪声干扰等影响时,振动信号在多个时间尺度上具有复杂的特征^[30-31]。

为了提取不同时间尺度的特征,近年来应用多尺度卷积神经网络(multi-scale convolutional neural network, MCNN)进行轴承故障诊断得到广泛研究,MCNN利用不同大小的卷积提取轴承振动信号中的不同尺度的特征,特征提取能力较单一尺度卷积提取特征更具鲁棒性。文

献^[32-36]利用MCNN提取特征,能在振动信号中提取更多有效特征,提高了模型的抗噪能力。但直接使用原始振动信号开展网络训练,会因数据量大而导致训练时间长,因此需要在保留信号整体信息的情况下压缩信号长度。

综上,针对传统CNN的不足,面向轴承微小故障信号非平稳非线性且易受背景噪声干扰的特点,提出了一种基于格拉姆角场和多尺度卷积神经网络(Gramian angular field and multi-scale convolutional neural network, GAF-MCNN)的智能故障诊断方法,并进行了实验验证。

1 GAF-MCNN 故障诊断新方法

基于格拉姆角场算法的优点和多尺度卷积神经网络强大的特征提取能力,提出一种故障诊断新方法,如图1所示。该方法首先使用分段聚合近似(piecewise aggregate approximation, PAA)算法^[37]对振动信号进行压缩,接着使用格拉姆角场算法^[38]对压缩后的信号进行编码,将一维原始信号转换为二维矩阵形式的方法。经过该方法处理后的二维矩阵加强了原始振动信号间的时间关系,将时间维度编码到了矩阵结构中。同时提出了一个具有多尺度卷积核的故障诊断模型,多尺度卷积核捕获的特征更为全面,并且利用批量归一化算法加快模型的训练速度。

诊断总体步骤如图2所示,流程如下。

步骤1)利用轴承数据采集平台对6种状态信号进行

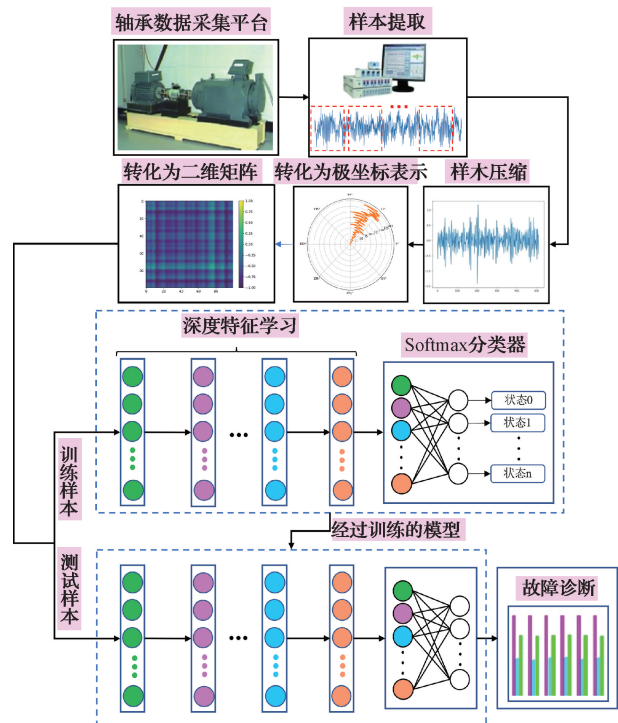


图1 GAF-MCNN 故障诊断框架

Fig. 1 The flowchart of the fault diagnosis method

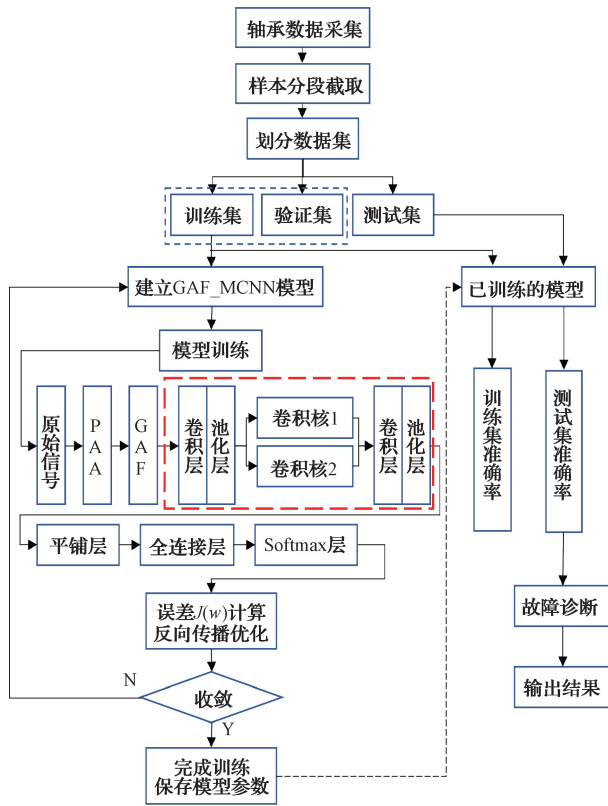


图2 GAF-MCNN故障诊断流程

Fig.2 GAF-MCNN fault diagnosis block diagram

采集归类。

步骤2)对采集到的信号进行截取、打标签形成数据集。

步骤3)按照一定比例将数据集划分为训练集和测试集。

步骤4)建立GAF-MCNN模型。

步骤5)利用PAA算法对信号进行压缩,随后使用GAF算法将压缩后的一维信号转化为二维矩阵。

步骤6)将训练集中处理后的二维矩阵输入到模型进行训练。若模型收敛,保存模型参数,执行下一步操作;若模型不收敛,返回步骤4)重新建立新的模型。

步骤7)将测试集中的样本输入到训练好的模型,得到测试集准确率。

1.1 信号处理

1)格拉姆角场算法

首先,对信号 $X = \{x_1, x_2, \dots, x_n\}$ 通过式(1)或式(2)进行标准化,将序列 X 中的值缩放到区间 $[-1, 1]$ 或者 $[0, 1]$ 中,缩放后的序列记为 \tilde{X} 。

$$\tilde{x}_{-1}^i = \frac{(x_i - \max(X)) + (x_i - \min(X))}{\max(X) - \min(X)} \quad (1)$$

$$\tilde{x}_0^i = \frac{x_i - \min(X)}{\max(X) - \min(X)} \quad (2)$$

式中: \tilde{x}_{-1}^i 代表将 x_i 缩放到区间 $[-1, 1]$; \tilde{x}_0^i 代表将 x_i 缩放到区间 $[0, 1]$ 。

然后,将缩放后的时间序列 \tilde{X} 转换到极坐标中,如下所示:

$$\begin{cases} \phi_i = \arccos(\tilde{x}_i), & 0 \leq \tilde{x}_i \leq 1, \tilde{x}_i \in \tilde{X} \\ r_i = \frac{t_i}{N}, & i \in 1, 2, \dots, N \end{cases} \quad (3)$$

式中: ϕ_i 为极坐标中的角坐标,由经过缩放后的时间序列值 \tilde{x}_i 经反余弦函数得出; N 为时间序列的长度; r_i 为极坐标中的半径,将区间 $[0, 1]$ 分成 N 等份。

以上为时间序列的编码过程,上述过程不仅可以独一无二的将每个时间序列中值编码到极坐标中,而且这个过程是可逆的。当完成时间序列极坐标编码后,因为极坐标保留了绝对的时间关系,所以就可以利用角度的相加,来定义任意两两时间点间的时序关系:

$$GASF = [\cos(\phi_i + \phi_j)], \quad i, j \in 1, 2, \dots, N \quad (4)$$

$$G = \begin{pmatrix} \cos(\phi_1 + \phi_1) & \cos(\phi_1 + \phi_2) & \dots & \cos(\phi_1 + \phi_n) \\ \cos(\phi_2 + \phi_1) & \cos(\phi_2 + \phi_2) & \dots & \cos(\phi_2 + \phi_n) \\ \vdots & \vdots & \ddots & \vdots \\ \cos(\phi_n + \phi_1) & \cos(\phi_n + \phi_2) & \dots & \cos(\phi_n + \phi_n) \end{pmatrix} \quad (5)$$

转换为类Gram矩阵的二维化矩阵有如下4个优点。

(1)保存时间序列中的信号点之间的对时间的依赖性。因为随着 i 的增加,时间序列中的 x_i 对应矩阵中的位置会从左上向右下推移,保持着时间序列的连续性,让我们在二维化过程中不失原本时间序列的特性。(2)保留了时间序列的相关性,因为 $G_{(i,j)||i-j|=k}$ 代表的是在时间差距 k 之下,信号点 x_i 和 x_{i+k} 的相关性。(3)在二维化后矩阵中的对角线上还保留着时间序列的原始值的顺序性,也可以通过 \cos 函数将经过缩放后的时间序列 \tilde{X} 还原。(4)具有稀疏表示的特点,很容易与高斯噪声区分开。 $\cos(\phi_i + \phi_j)$ 过程在笛卡尔坐标系下表示为:

$$\cos(\phi_i + \phi_j) = \cos(\arccos x + \arccos y) = x \cdot y - \sqrt{1-x^2} \cdot \sqrt{1-y^2} \quad (6)$$

从式(6)可以看出,对 x 和 y 的内积增加了惩罚项 $-\sqrt{1-x^2} \cdot \sqrt{1-y^2}$, 信号 x 和信号 y 越接近0,表明他们更接近高斯噪声,所以增大惩罚值,使得有用信号和高斯噪声更容易区分开。

2)PAA算法

虽然GAF算法可以将时序信号 $X = \{x_1, x_2, \dots, x_n\}$ 编码到二维矩阵 $G = [m_{i,j}]$ 中,并且保留了信号之间复杂的时域关系。但是,转换后的二维矩阵 G 的维度是 $n \times n$ 的。在滚动轴承信号故障诊断领域中,为了更好的检测出轴承的工作状态,通常传感器以很高的采样频率进行信号采样,因此时序信号 X 长度会很大。直接使用GAF算法对原始信号序列进行转换,将会产生超大型矩阵。不仅在

转换过程中耗费大量时间,而且在 CNN 模型学习过程中要耗费大量计算资源,很难满足实时故障诊断的需求。因此,不能直接使用原始信号,需要对滚动轴承信号进行压缩。

时间序列数据压缩通常采用将原始序列进行转化或者特征提取的方式,来达到将原始信号压缩降低维度的目的。PAA 算法是一种有效的数据降维方法。该方法将时间序列信号分割成多个子序列,每个子序列的值使用原始序列的分割区间的均值来表示。给定长度为 n 的序列 $S = (s_1, s_2, \dots, s_n)$, 把它转换成另一条长度为 m 的序列 $Q = (q_1, q_2, \dots, q_m)$, 其中 $n > m$ 。那么,该方式是一个压缩比为 $k = n/m$ 的时间序列数据降维的过程。新序列中任意元素 q_i 满足式(7)。

$$q_i = \frac{1}{k} \sum_{j=k \cdot (i-1) + 1}^{k \cdot i} S_j, 1 \leq i \leq m \quad (7)$$

对时间序列 S 进行分段后求出每段的均值,每段的均值形成序列 Q , 进而达到序列降维的目的。负载 1 下长度为 400 的标签为 0 的信号通过 PAA 降维后与降维前的结果对比如图 3 所示。

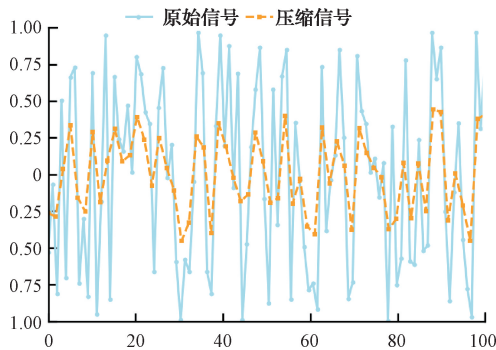


图 3 原始信号与压缩后信号的对比

Fig. 3 Comparison of original signal and compressed signal

PAA 算法通过分段后的子序列均值来表示序列片段特征,这种做法虽然会忽略数据的局部变化的情况。但是,在滚动轴承信号故障诊断领域,信号的局部变化不是研究的关键,信号序列的整体形态的变换才是重点关注的特征与研究的关键。从图 3 可以看出,经过 PAA 算法压缩后得到的信号能够准确地反映长时间序列数据形态的整体变换趋势,并且可以起到减小数据存储空间和提高计算效率的作用。

1.2 多尺度卷积神经网络

CNN 是多级神经网络,是一种专门用来处理具有类似网格结构的数据的神经网络。图 4 所示为二维 CNN 的结构,典型的 CNN 由 3 种网络结构组成,分别为卷积层、池化层和全连接层。卷积层通常伴有激活函数,用于从输入数据中自动提取特征。池化层将提取到的特征进行下采样操作,相当于数据压缩。经过多层卷积与池化操作,

最终将提取到的高维特征送入全连接层进行分类。

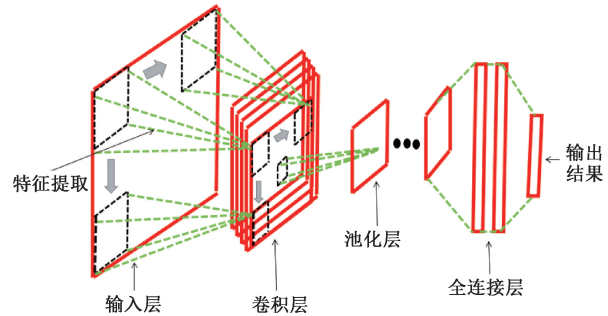


图 4 卷积神经网络结构

Fig. 4 Convolutional neural network structure

在 CNN 中有一个概念叫做感受野,用来表示网络内部的不同位置的神经元在原图像上的感受范围的大小。卷积核越大,表明该层神经元的感受野也越大,即获得的特征更为全局。然而,大卷积核使得模型在浅层网络就获得较大感受野,不利于模型深度的增加,同时会导致计算量暴增,给模型的性能带来瓶颈。目前流行的 VGG、Inception 网络结构中,为了避免大卷积核带来的缺陷,又要保证感受野足够大,因此使用多个 3×3 的小卷积进行堆叠来替代单层使用大的卷积核。并且,在实际应用中,小卷积核堆叠组合取得了较好的效果,被广泛应用在各种模型中。

振动信号本身就蕴含着各种不平衡和不对称响应的振动信息,以及传递过程中附加的整个机械系统产生的多种振动干扰。尤其是处于故障状态下的机械设备,故障特征信号不会以稳定的周期出现,其会伴随着时间增加,而产生波动与偏移。因此,当振动信号转换到二维矩阵后,关键特征的占比将会不同,用单一尺度的卷积核只会关注一类关键特征,忽略掉其他尺度特征,从而影响故障诊断算法的性能。本文将会在新的模型中增加一层使用多尺度卷积进行特征提取的结构,采用 3×3 和 5×5 两种尺度的卷积核分别对前一层得到的特征图进行特征提取,然后将各自得到的特征图进行合并,达到捕获更全面的特征的目的。

1.3 Dropout 算法

过拟合是机器学习和深度学习在模型训练过程中常见的问题,比如模型在训练数据集上的损失函数值较小、准确率较高。但在测试数据集上应用训练好的模型时,损失函数值较大、准确率较低。为解决过拟合问题,Hinton 提出了 Dropout 算法,它是一种计算复杂度低,但功能强大的正则化模型算法,是深度学习中防止过拟合提高模型效果的利器。Dropout 可以被认为是集成了大量深层神经网络的实用 Bagging 方法,集成包括所有从基础网络除去非输出单元后形成的子网络。在训练过程中以一定的概率随机地“临时丢弃”一部神经元节点,相当于每次迭代都在训练不同结构的神经网络。标准网络与带有 Dropout 网络操作对比如图 5 所示。

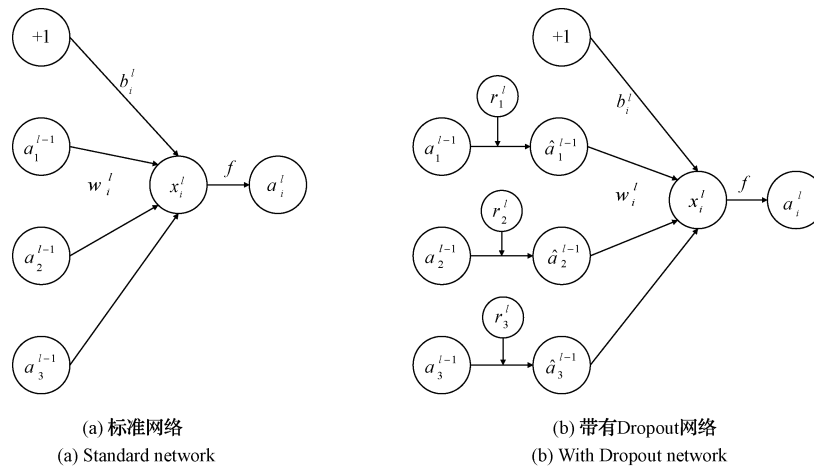


图5 标准网络与带有 Dropout 网络对比

Fig. 5 Comparison of standard network and network with Dropout

1.4 批量归一化

在深度学习中,往往会使用较深层数的网络进行训练,以达到对训练数据集的复杂线性关系与非线性关系的拟合,这样会导致深层神经网络的训练调参更为困难且复杂,其主要操作步骤如下。

1)如果是全连接层输入,则分别统计 mini-batch 下各个特征的均值和标准差,用全连接层输入减去均值再除以其标准差。如果是多通道卷积层输入,将这些通道视为不同特征,分别统计 mini-batch 下各个通道的均值和标准差,进行标准化操作。以上做法使得每一层网络的输入数据分布近似。

2)由于步骤 1)改变了网络每一层分布,因而改变了网络层中本身数据的表达能力。因此,将标准化后的值重新乘以一个缩放因子 γ 并加上偏置量 β ,用于增强表达。

1.5 RMSprop 优化算法

对于浅层的神经网络,通常采用随机梯度下降(stochastic gradient descent, SGD)算法来寻找最优点,其中对于超参数学习率要进行反复调试来达到最优。学习率对神经网络模型的性能有显著的影响,损失值通常高度敏感于参数空间中的某些方向,以相同的学习率去更新参数的各个分量会不合适,动量算法在一定程度上缓解这些问题。但是这样做的代价会引入了一个新的超参数,并没有实质的解决学习率设置难的问题。RMSprop 优化算法由 Hinton 提出,在 AdaGrad 自适应学习率算法的基础上,改变了梯度积累方式,使用指数加权的移动平均。AdaGrad 算法根据平方梯度的整个历史收缩学习率,而 RMSProp 算法使用指数衰减平均以丢弃遥远过去的历史,在找到目标函数的凸结构后,可以快速收敛。

2 实例分析

2.1 数据集描述

本文试验数据选自凯斯西储大学(CWRU)滚动轴承数据中心,采用电火花加工将单点故障引入试验轴承,传感器的采样频率为 12 kHz。被诊断的轴承一共有 3 种缺陷位置,滚动体损伤、外圈损伤和内圈损伤,其中外圈损伤又分为轴承外圈 3 点钟方向,6 点钟方向和 12 点钟方向的 3 种外圈损伤故障。选择轴承损伤直径大小为 0.007 in (1 in=25.4 mm) 的损伤状态,共 5 种损伤状态。试验数据集共有 A、B、C 3 个分别对应负载 1、2 和 3 hp (1 hp \approx 0.746 kW),每个数据集中,每种类型包含 300 个样本,共计 1 800 个样本,测试样本占 10%,如表 1 所示。在负载 1 hp 下,6 种轴承状态原始振动信号、压缩后并归一化的信号、转到极坐标下的信号、二维矩阵热图如图 6 所示。

表 1 数据集描述

Table 1 Dataset description

损伤位置	标签	数据集 A、B、C 样本数量	
		训练集	测试集
正常	0	270	30
滚动体	1	270	30
内圈	2	270	30
3 点钟	3	270	30
外圈 6 点钟	4	270	30
12 点钟	5	270	30

2.2 基于 GAF 的二维 CNN 故障诊断模型

1) 网络结构设计

本文使用多个堆叠小卷积核。同时,使用多尺度卷积核去取代单一尺度卷积核,多尺度卷积核结构如图 7

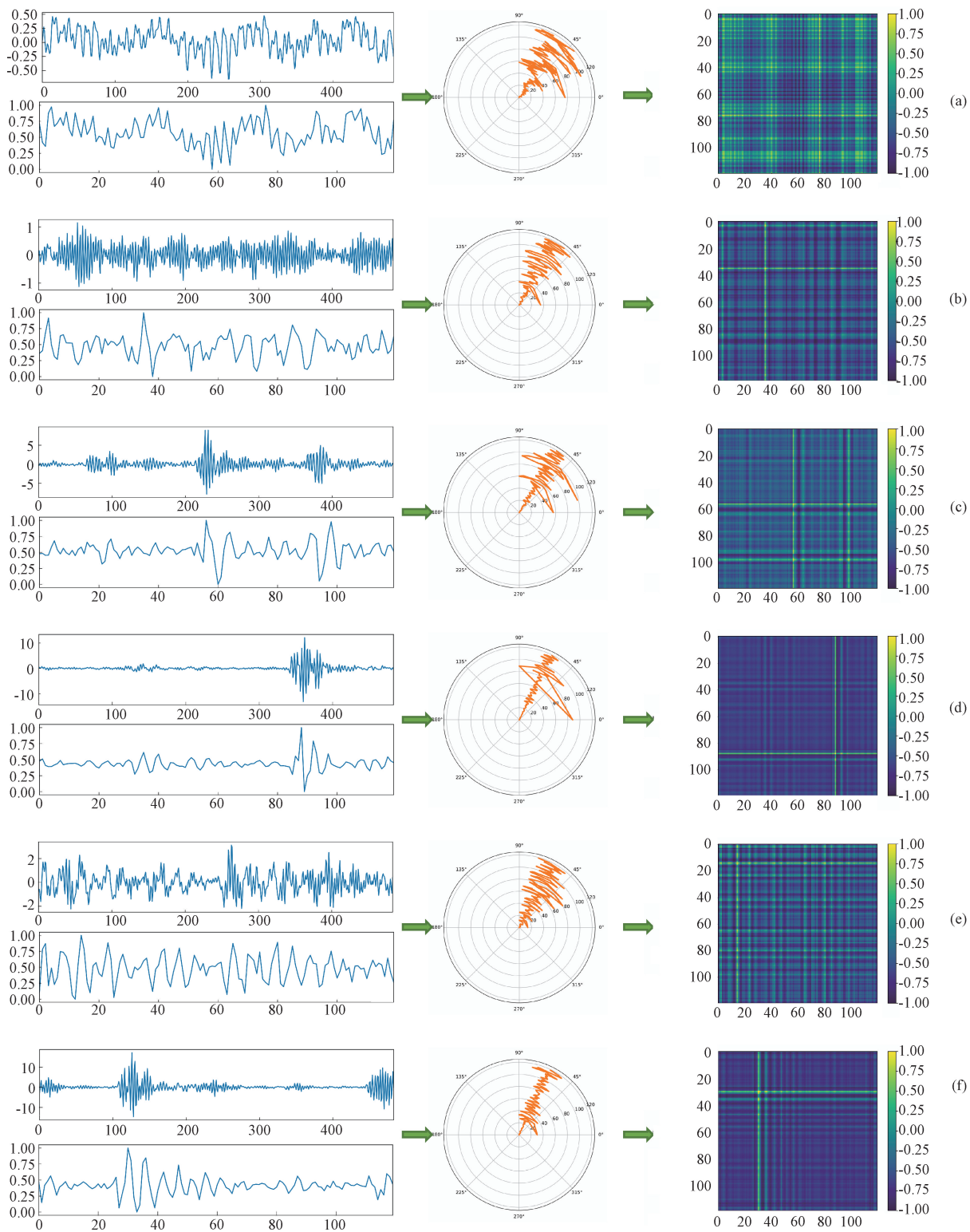


图6 负载1 hp下6种轴承状态波形

Fig. 6 Waveform diagrams of six bearing states under a load of 1 hp

所示。最后,在卷积层之间加入批量归一化操作和在全连接层加入 Dropout 操作。为了减少引入多个尺寸的卷

积核带来大量额外参数的问题,使用大小为 1×1 的特殊卷积核。它的计算发生在通道维度上,起到保持特征图

的高与宽两个维度尺寸不变的作用,相当于卷积层中的全连接层,通过调整特征图上的通道数来降低模型的复杂性。

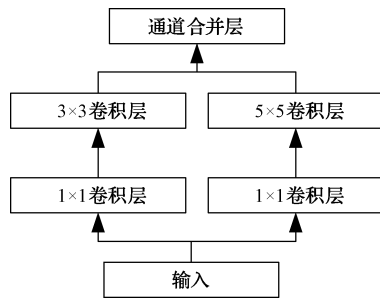


图7 多尺度卷积核结构

Fig. 7 Multiscale convolution kernel structure

2) 网络结构主要参数设置

试验使用的二维 CNN 模型参数如表 2 所示。其中第 1 层和第 2 层都使用 5×5 的步长为 1 的卷积核进行堆叠,他们组合的视野域与 9×9 的卷积核相同。第 3 层使用多尺度卷积核结构,利用 3×3 和 5×5 的两种不同的尺度卷积进行特征提取,在通道维度进行合并。为了增强网络的表达能力,其余卷积层的卷积核大小均为 3×3 。这样有利于加深网络,同时可以抑制过拟合。

表 2 GAF-MCNN 模型参数

Table 2 GAF-MCNN model parameters

编号	网络层	卷积核大小/步长	零补
1	卷积 1	$5 \times 5 / 1 \times 1$	是
2	卷积 2	$5 \times 5 / 1 \times 1$	否
3	池化 1	$3 \times 3 / 1 \times 1$	否
4	多尺度卷积	$1 \times 1 - > 3 \times 3 / 1 \times 1$ $1 \times 1 - > 5 \times 5 / 1 \times 1$	是
5	池化 2	$3 \times 3 / 1 \times 1$	否
6	卷积 3	$3 \times 3 / 1 \times 1$	是
7	池化 3	$2 \times 2 / 2 \times 2$	否
8	卷积 4	$3 \times 3 / 1 \times 1$	是
9	池化 4	$2 \times 2 / 2 \times 2$	否
10	全连接 1	128	否
11	全连接 2	64	否
12	Softmax	6	否

每层卷积操作之后均进行批量归一化,然后进行 2×2 的最大池化。第 1 个隐藏层神经元个数为 64,第 2 个隐藏层神经元个数为 32,最后 Softmax 层共有 6 个输出,对应 6 种轴承状态。

3) 网络搭建和训练流程

本文在 Keras 框架下进行搭建和训练网络模型。Keras 是一个具有高度模块化特点的神经网络库,它是由

Python 编写并基于 Tensorflow、Theano 以及 CNTK 作为后端的高级库,本文使用 Tensorflow 作为后端。其模块简洁,并且可以随意组合,在研究领域的应用率很高,是一个自由度非常高的框架。在 Keras 框架下,输入尺寸为 $80 \times 80 \times 1$ 的数据在 GAF-MCNN 模型中各个模块的输入输出尺寸结果如图 8 所示。在训练时,每个 mini-batch 的大小设置为 128,RMSProp 算法的初始学习率设置为 0.001。

2.3 数据压缩比例研究

在负载 1 下,超参数(数据压缩比)对 GAF-MCNN 模型性能的影响,并确定最优的数据压缩比。试验中使用 400 个采样点作为一个样本。分别使用 $k=4,5,8,10,20$ 的压缩比例,对长度为 400 的样本点进行压缩。对应试验数据如表 3 所示。为了验证算法稳定性,每个试验重复 5 次,每次试验结果如图 8 所示。

表 3 数据压缩比试验组

Table 3 Data compression ratio test group

试验编号	压缩比例 (k)	原始信号长度 (L)	压缩后信号长度 (L)
1	20	400	20
2	10	400	40
3	8	400	50
4	5	400	80
5	4	400	100

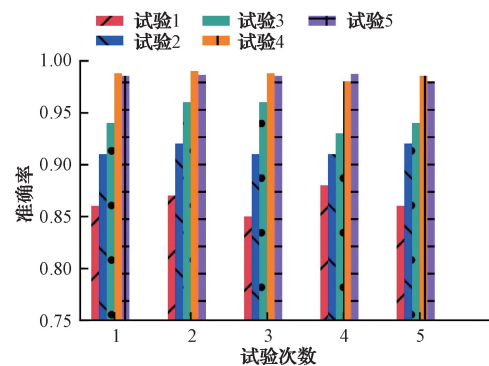


图 8 5 种压缩比例试验结果

Fig. 8 Test results of five compression ratios

由图 8 可以看出,GAF-MCNN 算法很稳定,5 种压缩比例下的 5 次重复试验结果波动较小。随着压缩比例的降低,故障识别准确率逐步增加。当压缩比例为 5 时,即压缩后的信号长度为 80 时,平均准确率最高为 98.84%,其次是压缩比例为 4 时,平均准确率为 98.57%。压缩比例为 20 时,平均准确率最低为 86.07%,其余试验的平均准确率均达到 90% 以上。

不同压缩比例下,算法迭代一个周期所需要的时间如

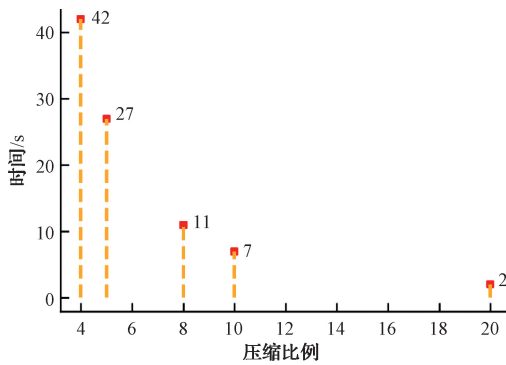


图9 训练一个周期所用的时间
Fig. 9 Training time for one cycle

图9所示。从图9可以看出,虽然压缩比例为5时的平均准确率与压缩比例为4时的平均准确率相近,但是压缩比例为5时,训练一个周期所用的时间要远远小于压缩比例为4时,从准确率和训练时间两种因素综合考量下,最终选定压缩比例5作为后续试验参数。

2.4 消融实验

为了更好说明提出的方法在轴承故障诊断任务上的有效性,构造3个新的网络模型进行消融实验对比。GAF-CNN1,使用 5×5 的卷积替代多尺度卷积层;GAF-CNN2,使用 3×3 的卷积替代多尺度卷积层;MCNN,不利用格拉姆角场算法将信号转换为二维矩阵形式。GAF-CNN1、GAF-CNN2和MCNN其余参数配置均GAF-MCNN相同。4种网络分别进行5次重复试验,平均准确率如图10所示。从图10看出,GAF-MCNN准确率最高,平均准确率为98.84%,其次是不采用格拉姆角场算法的多尺度卷积核网络MCNN,平均准确率为95.93%,而后是卷积核为 5×5 的GAF-CNN1,平均准确率为94.53%,卷积核为 3×3 的GAF-CNN2平均准确率最低,为91.54%。由试验结果可以得出,基于GAF-MCNN的智能故障诊断方法对故障特征的提取效果优于单一尺度卷积核和未经格拉姆角场算法转换方法的效果,可以显著地提高网络的故障诊断识别能力。

2.5 不同轴承故障诊断方法分类性能比较

为了更好解释GAF-MCNN用于轴承故障诊断的准确性及优势,将本文提出方法与当前具有代表性的智能故障诊断方法进行对比分析^[39-40]。选用负载1下的训练数据集和测试数据集作为研究对象,在样本长度为400, PAA压缩比例为5的条件下进行研究。在采用相同数据预处理方式条件下对比一下4种方法:方法1为本文提出的基于GAF-MCNN的智能故障诊断方法;方法2为基于CNN的智能故障诊断方法;方法3为结合CNN和长短时记忆网络(long short-term memory, LSTM)的智能故障诊断方法;方法4为采用自适应多尺度卷积核(multi-scale adaptive selective convolutional neural network, MSASC-

NN)的智能故障诊断方法。4种方法均经过5次重复试验,平均准确率和标准差如表4所示。

由表4可知,GAF-MCNN的平均诊断准确率高于其他3类轴承故障诊断方法。相较于CNN、CNN+LSTM和MSASCNN平均准确率分别高出7.14%、4.96%和1.31%,表明GAF-MCNN在轴承故障诊断中具有较高的准确性。

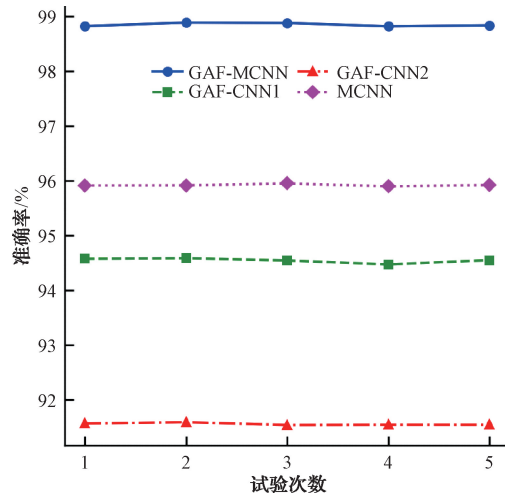


图10 消融试验结果对比
Fig. 10 Comparison of ablation test results

表4 4种轴承故障诊断方法测试结果

方法名称	平均准确率	标准差
GAF-MCNN	98.84	2.91
CNN	91.70	2.63
CNN+LSTM	93.88	1.78
MSASCNN	97.53	2.32

引入混淆矩阵对结果进行对比分析,结果如图11所示。混淆矩阵也称误差矩阵,以 N 行 N 列的矩阵形式来表示。其中, N 代表数据集中的类别数目,每一行代表了数据的真实归属类别,每一列代表了算法预测类别。从图11(a)可以得出,GAF-MCNN预测错误集中在类别4和类别5,其余类别准确率均达到100%,效果最佳。从图11(b)看出,使用CNN预测也有同样的问题,只有类别1能100%预测准确,其余类别均会出现误判,类别0和类别4分类效果最差,发生误判次数最多。图11(c)为CNN+LSTM的混淆矩阵,不仅无法准确的分辨出类别4和类别5,而且在类别0也出现错误。从图11(d)看出,使用MSASCNN只有类别4和5会出现错误,其余均能100%预测准确。此外,对比4类方法的混淆矩阵可知,GAF-MCNN具有较高的稳定性和可靠性,且健康工况、滚动体和内圈故障诊断效果相对较好。

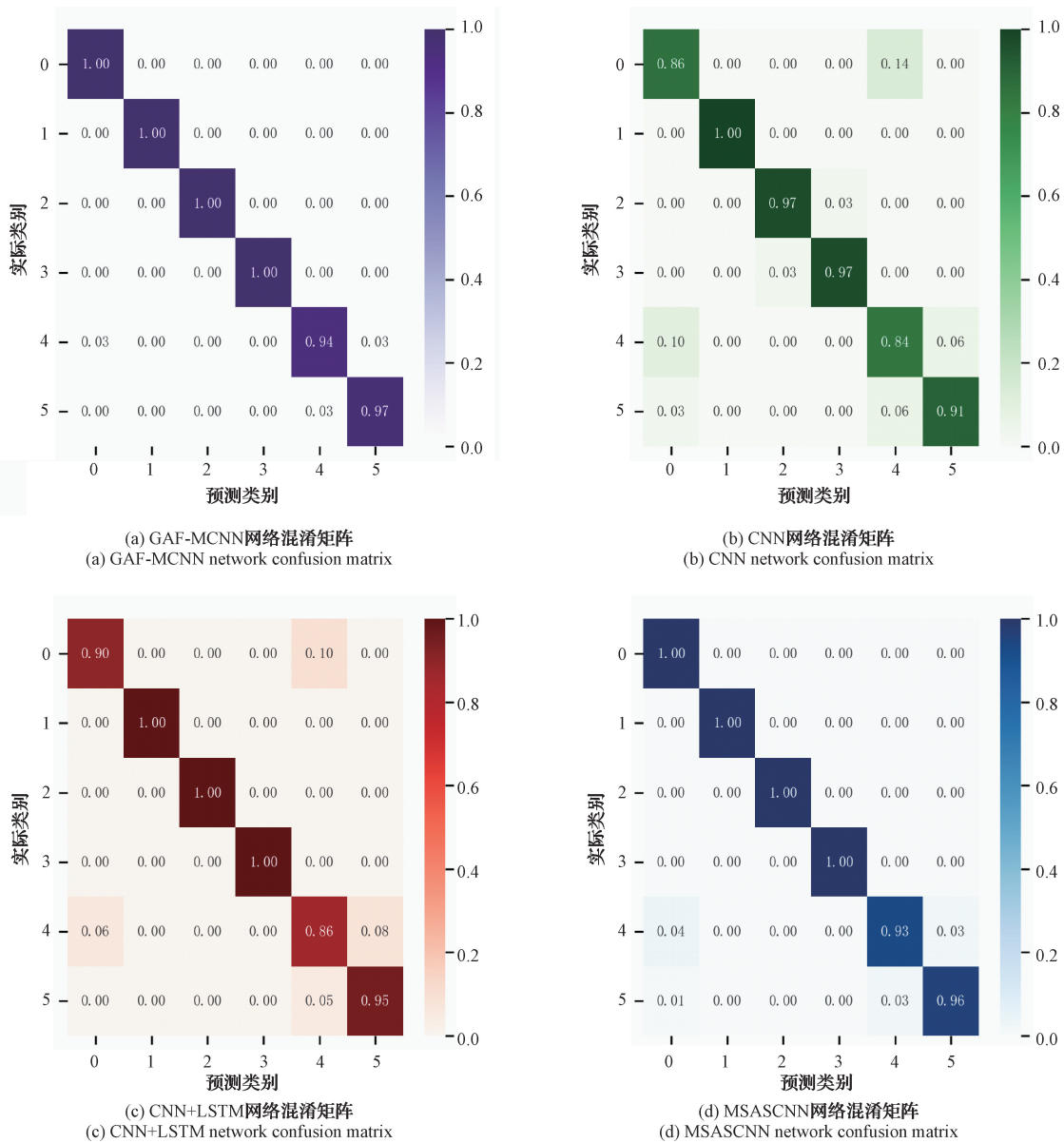


图 11 4 种网络结构试验结果混淆矩阵

Fig. 11 Confusion matrix of four network structures test results

3 结 论

本文提出了一种基于 GAF-MCNN 模型的智能故障诊断新方法。在数据处理方面,首先使用 PAA 算法将一维轴承信号进行降维,然后通过 GAF 算法将降维后的信号转换为二维矩阵形式,不仅保留了信号的时序关系,还有消除高斯白噪声的效果,增强了信号之间的联系。在网络结构方面,引入了多尺度卷积核,用于捕获不同尺度的特征,达到了增强网络的非线性表达能力。GAF-MCNN 算法准确率超过 98%。

参 考 文 献

[1] DAS O, DAS D B, BIRANT D. Machine learning for

fault analysis in rotating machinery: A comprehensive review[J]. HELIYON, 2023, 9(6): 17584-17618.

[2] AGRAWAL A K, CHATTOPADHYAYA S, MURTHY V M S R, et al. A novel method of laser coating process on worn-out cutter rings of tunnel boring machine for eco-friendly reuse[J]. Symmetry, 2020, 12(3): 471-471.

[3] MIAN Z B, DENG X F, DONG X H, et al. A literature review of fault diagnosis based on ensemble learning [J]. Engineering Applications of Artificial Intelligence, 2024, 127(B): 107357-107385.

[4] ZHENG J D, CAO S J, PAN H Y, et al. Spectral

- envelope-based adaptive empirical Fourier decomposition method and its application to rolling bearing fault diagnosis[J]. *ISA Transactions*, 2022, 129(B): 476-492.
- [5] WANG D, KWOK L T, QIN Y. Optimization of segmentation fragments in empirical wavelet transform and its applications to extracting industrial bearing fault features[J]. *Measurement*, 2019, 133: 328-340.
- [6] SUN Y J, LI S H, WANG X H. Bearing fault diagnosis based on EMD and improved Chebyshev distance in SDP image[J]. *Measurement*, 2021, 176: 109100-109113.
- [7] GAI J B, SHEN J X, HU Y F, et al. An integrated method based on hybrid grey wolf optimizer improved variational mode decomposition and deep neural network for fault diagnosis of rolling bearing [J]. *Measurement*, 2020, 162: 107901-107913.
- [8] ORHAN Y, FERHAT Y, AYHAN A. A fault detection method based on embedded feature extraction and SVM classification for UAV motors [J]. *Microprocessors and Microsystems*, 2022, 94: 104683-104693.
- [9] QIN Y, YANG R, HE B, et al. Multi-layer convolutional dictionary learning network for signal denoising and its application to explainable rolling bearing fault diagnosis[J]. *ISA Transactions*, 2024, 147: 55-70.
- [10] SUBRATA M, VIKASH K, SOMNATH S, et al. Gearbox fault diagnosis using advanced computational intelligence[J]. *Procedia Computer Science*, 2020, 167: 1594-1603.
- [11] FRANCISCO J G G, VÍCTOR M M, RUBEN R G, et al. Vibration-based monitoring of agro-industrial machinery using a K-Nearest Neighbors (KNN) classifier with a Harmony Search (HS) frequency selector algorithm[J]. *Computers and Electronics in Agriculture*, 2024, 217: 108556-108570.
- [12] WANG Z Y, YAO L G, CAI Y W. Rolling bearing fault diagnosis using generalized refined composite multiscale sample entropy and optimized support vector machine [J]. *Measurement*, 2020, 156: 107574-107589.
- [13] GAI J B, SHEN J X, HU Y F. An integrated method based on hybrid grey wolf optimizer improved variational mode decomposition and deep neural network for fault diagnosis of rolling bearing [J]. *Measurement*, 2020, 162: 107901-107913.
- [14] CAI B P, SHAO X Y, LIU Y H, et al. Remaining useful life estimation of structure systems under the influence of multiple causes: subsea pipelines as a case study[J]. *IEEE Trans. Ind. Electron.*, 2020, 67(7):5737-5747.
- [15] ZHAO R, YAN R Q, CHEN Z H, et al. Deep learning and its applications to machine health monitoring[J]. *Mech. Syst. Sig. Process.*, 2019, 115: 213-237.
- [16] 郝洪涛, 王凯, 张炳建, 等. 多尺度特征自适应融合的气动控制阀故障诊断[J]. *仪器仪表学报*, 2023, 44(10): 167-178.
- HAO H T, WANG K, ZHANG B J, et al. Fault diagnosis of pneumatic control valves with multi-scale features adaptive fusion [J]. *Chinese Journal of Scientific Instrument*, 2023, 44(10): 167-178.
- [17] 杨东儒, 魏建文, 林雄威, 等. 基于自注意力机制的深度学习模拟电路故障诊断[J]. *仪器仪表学报*, 2023, 44(3): 128-136.
- YANG D R, WEI J W, LIN X W, et al. A fault diagnosis algorithm for analog circuits based on self-attention mechanism deep learning [J]. *Chinese Journal of Scientific Instrument*, 2023, 44(3): 128-136.
- [18] FU Z Y, LIU Z, PING S R, et al. TRA-ACGAN: A motor bearing fault diagnosis model based on an auxiliary classifier generative adversarial network and transformer network[C]. *ISA Transactions*, 2024.
- [19] DONG Y T, JIANG H K, YAO R H, et al. Rolling bearing intelligent fault diagnosis towards variable speed and imbalanced samples using multiscale dynamic supervised contrast learning[J]. *Reliability Engineering & System Safety*, 2024, 243: 109805-109819.
- [20] HOANG D T, KANG H J. A survey on deep Learning based bearing fault diagnosis [J]. *Neurocomputing*, 2019, 335: 327-335.
- [21] XU Z F, LI C, YANG Y. Fault diagnosis of rolling bearing of wind turbines based on the variational mode decomposition and deep convolutional neural networks[J]. *Applied Soft Computing*, 2020, 95: 106515-106530.

- [22] CHE C C, WANG H W, NI X M, et al. Domain adaptive deep belief network for rolling bearing fault diagnosis[J]. *Computers & Industrial Engineering*, 2020, 143:106427-106436.
- [23] LI C, BAO Z T, LI L H, et al. Exploring temporal representations by leveraging attention-based bidirectional LSTM-RNNs for multi-modal emotion recognition [J]. *Information Processing & Management*, 2020, 57: 102185-102193.
- [24] 康守强, 杨加伟, 王玉静, 等. 基于联邦多表示域适应的不同工况下滚动轴承故障诊断方法[J]. *仪器仪表学报*, 2023, 44(6): 165-176.
- KANG SH Q, YANG J W, WANG Y J, et al. Fault diagnosis method of rolling bearings under different working conditions based on federated multi-representation domain adaptation[J]. *Chinese Journal of Scientific Instrument*, 2023, 44(6): 165-176.
- [25] JIANG H K, WANG F A, SHAO H D, et al. Rolling bearing fault identification using multilayer deep learning convolutional neural network [J]. *Journal of Vibroengineering*, 2017(1): 138-149.
- [26] PAN J, ZI Y Y, CHEN J L, et al. LiftingNet: A novel deep learning network with layerwise feature learning from noisy mechanical data for fault classification [J]. *IEEE Transactions on Industrial Electronics*, 2018, 65(6): 4973-4982.
- [27] JIAO J Y, ZHAO M, LIN J, et al. Deep coupled dense convolutional network with complementary data for intelligent fault diagnosis[J]. *IEEE Transactions on Industrial Electronics*, 2019, 66(12): 9858-9867.
- [28] LIANG H P, ZHAO X Q. Rolling bearing fault diagnosis based on one dimensional dilated convolution network with residual connection [J]. *IEEE Access*, 2021(9): 31078-31091.
- [29] WANG H, XU J W, YAN R Q, et al. A new intelligent bearing fault diagnosis method using SDP representation and SE-CNN[J]. *IEEE Transactions on Instrumentation and Measurement*, 2020, 69(5): 2377-2389.
- [30] LIANG H P, CAO J, ZHAO X Q. Multi-scale dynamic adaptive residual network for fault diagnosis [J]. *Measurement*, 2022, 188: 110397.
- [31] YAN S, SHAO H D, XIAO Y M, et al. Hybrid robust convolutional autoencoder for unsupervised anomaly detection of machine tools under noises[J]. *Robotics and Computer-Integrated Manufacturing*, 2023, 79: 102441.
- [32] HUANG W Y, CHENG J S, YANG Y, et al. An improved deep convolutional neural network with multi-scale information for bearing fault diagnosis[J]. *Neurocomputing*, 2019, 359: 77-92.
- [33] QIAO H H, WANG T Y, WANG P, et al. An adaptive weighted multiscale convolutional neural network for rotating machinery fault diagnosis under variable operating conditions [J]. *IEEE Access*, 2019, 7: 118954-118964.
- [34] CHEN X H, ZHANG B K, GAO D. Bearing fault diagnosis base on multi-scale CNN and LSTM model[J]. *Journal of Intelligent Manufacturing*, 2021, 32(4): 971-987.
- [35] JIA L S, CHOW T W S, WANG Y, et al. Multiscale residual attention convolutional neural network for bearing fault diagnosis [J]. *IEEE Transactions on Instrumentation and Measurement*, 2022, 71: 1-13.
- [36] HUANG Y J, LIAO A H, HU D Y, et al. Multi-scale convolutional network with channel attention mechanism for rolling bearing fault diagnosis [J]. *Measurement*, 2022, 203: 111935.
- [37] WANG Y, WANG J H, FAN Y R, et al. Interference fading suppression for distributed acoustic sensor using frequency-shifted delay loop[J]. *Optics & Laser Technology*, 2024, 171: 110441-110453.
- [38] CHENG Y N, LU M D, GAI X Y, et al. Research on multi-signal milling tool wear prediction method based on GAF-ResNext[J]. *Robotics and Computer-Integrated Manufacturing*, 2024, 85: 102634-102649.
- [39] 张玺君, 尚继洋. 采用多尺度自适应选择卷积神经网络的轴承故障诊断研究[J]. *西安交通大学学报*, 2024, 58(2): 127-135.
- ZHANG X J, SHANG J Y. Bearing fault diagnosis based on multi-scale adaptive selective convolutional neural network [J]. *Journal of Xi'an Jiaotong University*, 2024, 58(2): 127-135.
- [40] 宁方立, 王珂, 郝明阳. 融合 CNN 和 ViT 的声信号轴承故障诊断方法[J]. *振动与冲击*, 2024, 43(3): 158-163,170.
- NING F L, WANG K, HAO M Y. Fault diagnosis method for bearing based on fusing CNN and ViT[J].

Journal of Vibration and Shock, 2024, 43(3): 158-163,170.

作者简介

张超(通信作者),博士,教授,博士生导师,主要研究方向为测试性设计、故障诊断、容错控制。

E-mail:caec_zc@nwpu.edu.cn

房颖涛,硕士研究生,主要研究方向为测试性设计。

E-mail:fangyingtao@yeah.net

冯建睿,硕士,主要研究方向为故障诊断及容错控制。

E-mail:fjr123321@163.com

杨柯,硕士,工程师,主要研究方向为深度学习及故障诊断。

E-mail:yk1994@mail.nwpu.edu.cn

何世烈,硕士,工程师,主要研究方向为装备故障预测与健康管理技术及应用。

E-mail:hysl@ceprei.com

董志杰,硕士,工程师,主要研究方向为信号处理。

E-mail:zhjdong0036@hotmail.com