

基于混合域残差注意力网络的滚动轴承智能故障诊断方法^①

贾立新^② 陈永毅 倪洪杰^③ 张 丹

(浙江工业大学信息工程学院 杭州 310023)

摘 要 机械设备正朝着大型化、精密化和自动化的方向发展,机械系统也因此变得越来越复杂。考虑到机械系统可能会发生无特征的灾难性故障,因此机械故障的自动检测是一个巨大的挑战。然而,现有的故障检测方法在对高度复杂的工业系统进行故障类型识别时,误诊率较高,无法给出准确的故障诊断结果。针对这一问题,本文以滚动轴承这一机械设备关键部件作为研究对象,提出一种基于混合域残差注意力网络的故障诊断方法,旨在结合深度卷积神经网络自动学习表示的优点,并配合通道注意力机制和空间注意力机制的关键特征提取能力,提高故障检测性能。实验结果表明,所提出的方法能够准确地检测轴承故障类型,在准确度指标方面优于其他方法。

关键词 故障诊断;滚动轴承;通道注意力机制;空间注意力机制;卷积神经网络(CNN)

随着工业生产和生活对机械装备功能要求越来越高,工业装备的内部组件变得越来越复杂,在系统运行过程中极易发生故障。只有构建安全可靠的工业装备与系统,才能满足工业生产的需要。滚动轴承是旋转机械的重要组成部分,其健康状况对工业机械的实用性和可用性有重大影响。同时,轴承也是旋转机械中最脆弱的部件之一,轴承故障是导致机械故障的主要原因^[1]。轴承故障可能会导致机器的严重损坏,进而迫使重要机械的不可用,导致经济损失和严重的安全隐患。因此,滚动轴承的早期故障诊断是保证机器不间断运行的关键。为了管理机械设备的潜在故障,轴承的故障检测和诊断已经引起了研究者的极大关注。研究人员引入了不同的方法来诊断滚动轴承的缺陷,其中时域、频域和时频域的振动信号分析方法应用最为广泛^[2-3]。然而,这是一个面向专家的任务,在快速响应大量数据方面,人工检测通常不是最有效的手段。为了减少对维护人员的依赖,智能状态监测技术正受到越来越

多研究人员的关注。

在过去的 20 年中,研究人员提出了许多与人工智能技术或传统的机器学习技术^[4]相关的新方法,如使用人工神经网络(artificial neural network, ANN)进行滚动轴承故障诊断^[5]、基于主成分分析(principal component analysis, PCA)的机器缺陷分类特征选择方案^[6]、基于支持向量机(support vector machine, SVM)的机器状态监测和故障诊断方法^[7]和基于神经网络和 SVM 的轴承缺陷诊断方法^[8]。近年来,也有研究者将机器学习算法与遗传算法相结合^[9],实现对轴承的智能故障诊断。例如,文献[10]采集了不同工况下的振动和电流数据,首先从振动和电流信号中提取时域、频域和时频域特征;并采用自动编码器进行特征学习;最后通过特征融合集成不同领域的信息,利用 softmax 层进行最终分类。文献[11]提出了一种基于变分自编码器和随机森林方法的多源信息融合算法,用于有限标注情况下的轴承故障诊断。文献[12]提出了一种基于电机电

① 国家重点研发计划(2018YFB1403702),宁波市科技创新 2025 重大专项(2019B1003)和国家自然科学基金(61873237)资助。

② 男,1966 年生,副教授;研究方向:工业过程监测装备控制;E-mail: jlx@zjut.edu.cn。

③ 通信作者, E-mail: zdfynhj@zjut.edu.cn。

(收稿日期:2022-05-19)

流信号的轴承故障诊断方法,利用卷积神经网络(convolutional neural networks, CNN)和信息融合技术从当前信号的灰度图像中自动提取特征。在应用信息融合技术后,采用经典的机器学习算法进行分类,能够准确定位故障位置。尽管故障诊断方法在轴承预测性维护方面取得了重大进展,但仍然存在一些局限性。一方面,基于机器学习的经典方法对复杂系统的表征能力有限,性能较差^[13-14];另一方面,这些方法高度依赖输入数据的故障特征,特征提取能力较弱。

为了解决上述故障诊断的挑战,深度学习(deep learning, DL)技术在近几年被广泛应用于机电系统的故障诊断^[15-16]。DL 通过使用具有多层的神经网络和非线性变换^[17],从数据中提取复杂的关系。与基于机器学习算法^[18]的手工特征提取过程不同,DL 以有效和自适应的方式深化输入数据特征的表征,能够显著提高故障诊断的准确率。卷积神经网络^[19]、长短期记忆网络(long short term memory, LSTM)^[20]以及深度自编码器(deep-autoencoders, DAE)^[21]在故障诊断中得到了广泛的应用。文献[22]通过 CNN 提取混沌序列中的高级抽象特征,并将提取的特征输入到 LSTM 作进一步处理,最后采用灰狼算法优化的 SVM 进行故障诊断,在噪声条件下取得了较高的准确率。文献[23]设计了一种基于双向 LSTM 的故障诊断模型,相比于 CNN,能够提取更多的故障特征信息,实现了端到端的故障诊断。

传统深度学习网络虽然能够准确诊断轴承故障类型,但容易受到噪声干扰。为了进一步提高故障诊断网络的鲁棒性,研究人员开始将计算机视觉中的注意力机制引入到故障诊断领域。文献[24]将一维振动信号时间序列转化为二维图像,并将图像数据作为网络的输入传入到所构建的多尺度注意力卷积神经网络模型中,通过注意力机制对不同的故障特征赋予不同的权重,使模型更关注于最具类别区分度的区域,从而提高模型的特征学习能力,在强噪声环境下表现出良好的鲁棒性。文献[25]将加权密集连接网络和注意力机制相结合,将不同空间级别的特征进行融合,使故障诊断模型具有较高的鲁棒性。文献[26]在 CNN 提取特征的基础上,采

用带有注意力机制的双向长短期记忆网络单元给 CNN 提取的特征赋予不同的权重,稳定性优于其他基于深度学习的诊断模型。文献[27]针对滚动轴承在强噪声环境下故障诊断效果不佳的问题,设计了一种具有通道注意力机制的多尺度卷积网络,在不同的噪声环境下均取得了较高的准确率。

虽然基于深度注意力机制网络的故障诊断方法在噪声条件下的鲁棒性得到了大幅度增强,但随着噪声的进一步增大,精度开始明显减小。为了更加充分地提取振动信号中的故障特征,本文提出了基于混合域残差注意力网络(mixed domain attention residual network, MDARN)的滚动轴承智能故障诊断方法。该网络是由 3 个注意力模块堆叠而成,堆叠方式采用的是残差网络的堆叠方式,能够避免直接叠加导致模型性能明显下降的问题。同时,采用最大池化层和上采样层模拟自下而上的快速前馈过程和自上而下的注意力反馈过程,能够显著提高故障诊断网络的鲁棒性。最后在滚动轴承故障诊断试验台上进行测试,验证了所提方法相比于现有方法的优越性。

1 残差网络

1.1 残差网络基本组件

残差网络具有与传统深度卷积网络相同的一些基本组件,包括卷积层、整流线性单元(rectified linear units, ReLU)激活函数、批归一化(batch normalization, BN)层、全局平均池化(global average pooling, GAP)层和交叉熵误差函数。下面对这些基本组件的概念作进一步详细介绍。

卷积层是使 CNN 区别于传统的全连接(fully connected, FC)神经网络的关键部分。由于卷积层通过卷积操作进行数据特征提取,而不是采用矩阵乘法进行运算,因此卷积层可以大大减少需要训练的参数量。此外,由于可训练参数较少,深度学习方法发生过拟合的可能性较小,因此可以更容易地对测试数据集产生较高的精度。输入特征映射与卷积核的卷积,加上偏置项可以表示为

$$y_j = \sum_{i \in M_j} x_i \times k_{ij} + b_j \quad (1)$$

式中, x_i 为输入特征图的第 i 个通道, y_j 为输出特征图的第 j 个通道, k 为卷积核, b 为偏置, M_j 为计算输出特征图的第 j 个通道的通道集合。

特征的分布往往在训练迭代过程中不断变化。在这种情况下,卷积层中的参数必须不断更新以适应变化的分布,这增加了训练的难度。BN 是一种特征规范化技术,作为可训练过程被插入到深度学习体系结构中。BN 的目的是减少内部协变位移,BN 首先将特征归一化为平均值为 0、标准差为 1 的固定分布,然后将特征调整为在训练过程中学习到的理想分布。BN 的过程可表示为

$$\mu = \frac{1}{N_{\text{batch}}} \sum_{n=1}^{N_{\text{batch}}} x_n \quad (2)$$

$$\sigma^2 = \frac{1}{N_{\text{batch}}} \sum_{n=1}^{N_{\text{batch}}} (x_n - \mu)^2 \quad (3)$$

$$\hat{x}_n = \frac{x_n - \mu}{\sqrt{\sigma^2 + \varepsilon}} \quad (4)$$

$$y_n = \gamma \hat{x}_n + \beta \quad (5)$$

其中, x_n 和 y_n 分别表示小批量第 n 个观测的输入和输出特征; μ 和 σ^2 是小批量的均值和方差, N_{batch} 是批量大小; γ 和 β 是 2 个可训练的参数来缩放和移动分布; ε 是常数,接近于 0; $\hat{x}_n$ 表示批归一化后的 x_n 。

在卷积层和 BN 之后一般紧跟激活函数,对提取到的特征进行非线性变换。激活函数通常是神经网络的一个重要组成部分,用于非线性变换。在过去的几十年里,已经开发了各种激活函数,包括 sigmoid、tanh 和 ReLU。ReLU 激活函数由于能有效地防止梯度消失,近年来受到广泛关注。ReLU 激活函数在负数部分的值恒为 0,此时导数为 0;而在正数部分的值为 x ,导数为 1,这有助于保持特征在网络层间流动时的范围不变。ReLU 激活函数如式(6)所示。

$$y = \max(x, 0) \quad (6)$$

其中, x 和 y 分别为 ReLU 激活函数的输入和输出。

1.2 残差网络设计

残差网络是近年来备受关注的—种新兴的深度学习方法,其中残差块是残差网络基本组成部分。如图 1 所示,残差块由 2 个 BN、2 个 ReLU、2 个卷

积层和 1 个跳跃连接组成。跳跃连接是残差网络优于一般深度卷积网络的部分。随着神经网络变得越来越复杂,网络深度越来越深。随着网络层数的逐渐增加,传统的 CNN 会产生梯度消失和梯度爆炸现象,严重影响网络的学习。

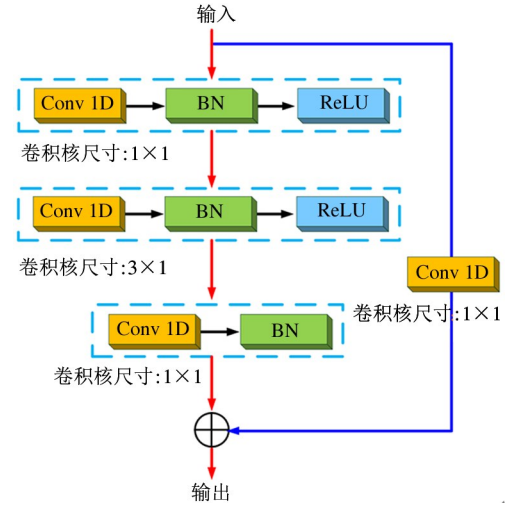


图 1 残差块

为了使深度神经网络能够更好地学习,残差块添加了跳跃连接,使输入特征能够直接传输到输出端,防止信息的快速流失。在使用了跳跃连接以后,梯度可以有效地流向较早的层,这些层接近输入层,因此可以更有效地更新参数,有效解决了梯度消失和梯度爆炸问题。

2 混合域残差注意力网络

2.1 注意力机制

振动信号中含有大量与故障不相关的信息,这些信息会严重干扰故障诊断网络的学习,使得故障诊断网络无法作出准确的判断。为了提高网络的抗干扰能力,研究者们开始将研究重心集中在注意力网络上。CNN 的典型注意力机制主要包括通道注意力和空间注意力。与传统的 CNN 不同的是,在注意力机制增强的 CNN 中,前一个卷积层输出的特征图在进入下一个卷积层之前会根据注意力系数进行加权,而不是直接输入到下一层。在通道注意力方面,首先采用全局平均池化(GAP)和全局最大池化

(global maximum pooling, GMP) 对通道维度上的特征映射进行压缩;然后将得到的特征映射送入一个权值共享的多层感知器(multilayer perceptron, MLP)中生成两组权值;最后,将两组权重之和作为注意力权重。在空间注意力方面,采用平均池化和最大池化的方法沿着空间方向压缩特征图。将得到的特征映射叠加成双通道格式后,由辅助卷积层生成代表不同区域重要性的空间注意力权值。注意力机制的引入改变了原来的等权模式,对网络的不同通道或空间区域赋予不同的权值,能够使网络更加关注对预测任务重要的特征,进而降低了噪声信息对预测任务的干扰。

本文在残差网络基础上,构建如图 2 所示的基于混合域残差注意力网络的故障诊断模型。该模型对振动数据采用掩码方式进行特征提取,削弱噪声信息对特征提取的干扰,提高网络的诊断精度。混合域残差注意力网络是通过堆叠多个注意力模块进行构建^[28]。每个注意力模块整体结构类似于残差块,分为 2 个路径:软掩码路径和主路径。主路径由 2 个残差块组成,负责对输入数据进行特征提取。软掩码路径模拟自底向上和自顶向下的结构学习相同大小的掩码,以此实现快速前馈和反馈的注意力过程。

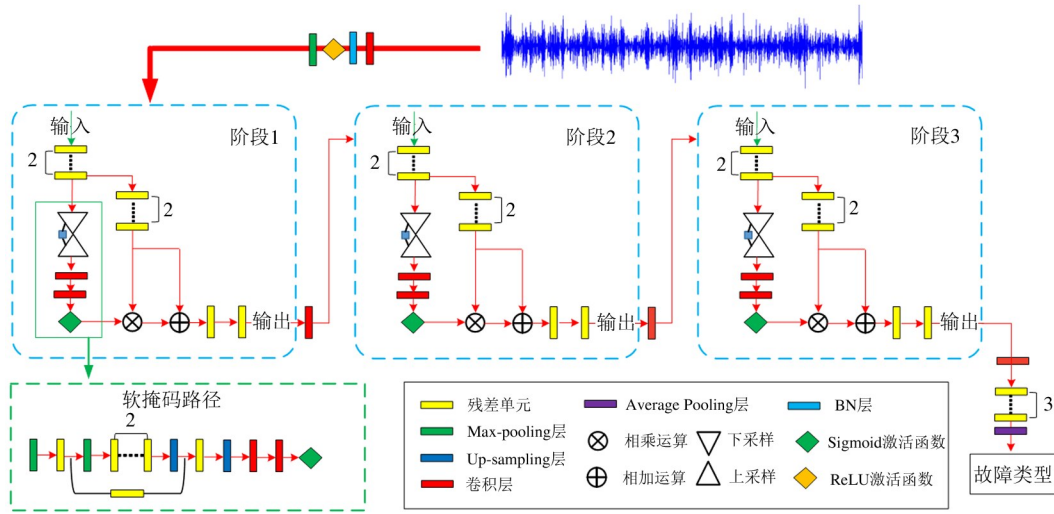


图 2 混合域残差注意力网络

2.2 软掩码路径

注意力网络通过软掩码路径求取注意力系数,整个软掩码路径包含快速前馈扫描和自顶向下的反馈步骤^[28]。前者快速扫描整个特征图的全局特征,后者则用来整合全局信息和原始特征图。

软掩码路径首先将输入数据 x 输入到连续 2 个最大池化层和残差块中,进行下采样操作,能够迅速增大网络层的感受野,实现对振动信号长期依赖信息的捕获。输入数据 x 在残差块中的处理过程如下:

$$y_r^1 = \text{Conv}_{1 \times 1}^{1 \times 1}(\delta(\text{BN}(x))) \quad (7)$$

$$y_r^2 = \text{Conv}_{3 \times 1}^{2 \times 1}(\delta(\text{BN}(y_r^1))) \quad (8)$$

$$y_r^3 = \text{Conv}_{1 \times 1}^{1 \times 1}(\delta(\text{BN}(y_r^2))) \quad (9)$$

$$y_{sc} = \text{Conv}_{1 \times 1}^{2 \times 1}(x) \quad (10)$$

$$y_{out}^r = y_r^3 \oplus y_{sc} \quad (11)$$

其中, y_r^1 、 y_r^2 、 y_r^3 、 y_{sc} 和 y_{out}^r 表示残差块中连续的卷积层的输出特征, $\text{Conv}_{v \times 1}^{s \times 1}$ 是步长为 $s \times 1$, 卷积核大小为 $v \times 1$ 的卷积层, δ 是 ReLU 激活函数, BN 是批归一化层, $\oplus$ 是逐元素相加。

在感受野足够大后,再对称于下采样过程,连续堆叠 2 个残差块和上采样层,对提取的特征进行上采样。全局信息通过对称的自顶向下架构进行扩展,以指导每个位置的输入特征。从图 2 中的软掩码中可以看到,在自下而上和自上而下之间添加了跳跃连接,以此来捕获来自不同尺度的信息,使捕获的信息更加丰富^[28]。针对每个通道和空间位置上的特征 F_M , 通过 Sigmoid 激活函数求取混合域注意力系数 M (通道域和空间域),将注意力系数限制在

$[0,1]$ 之间,以此防止出现过大的注意力系数,导致网络训练的崩溃。

$$M = \frac{e^{-F_M}}{(1 + e^{-F_M})^2} \quad (12)$$

2.3 混合域残差注意力模块

在获得注意力系数 M 之后,通过将主路径的特征 F_A 与注意力系数 M 进行逐元素相乘,便可得到输入数据 x 的注意力特征 F_{att} 。

$$F_{att} = F_A \otimes M \quad (13)$$

其中, M 是软掩码的注意力参数,取值范围是 $[0,1]$ 。

随着注意力模块的叠加,输出的注意力特征图会越来越小,有可能打破网络原有的特性,使网络性能降低。为了解决这个问题,求取的注意力特征 F_{att} 参照 ResNet 恒等映射的方法,将求取的注意力特征与主路径提取的特征进行相加,在确保注意力特征不损坏的前提下,保证了注意力模块输出的特征不会太小:

$$H = F_{att} \oplus F_A \quad (14)$$

其中, H 是注意力模块的输出。这就构成了混合域残差注意力模块,能将振动数据的特征和加强注意力之后的特征一同输入到下一模块中。注意力模块通过主路径和软掩码路径,既能获取输入数据中的最具类别区分度的关键特征,又能使其绕过软掩码路径,保持原始数据的良好特征,削弱软掩码路径的特征选择能力,使注意力模块的输出特征具有较强的抗干扰能力。

通过上述操作,网络在不损失振动信号原有特征的基础上,确保了重要特征能在所有特征中占据更大的比重,进而削弱噪声信息对故障诊断模型的干扰,增大不同故障数据间的可区分度。

2.4 混合域残差注意力网络设计

混合域残差注意力网络由3个连续的混合域残差注意力模块堆叠而成。通过堆叠注意力模块可以逐渐细化特征图。如图2所示,输入数据首先经过具有大尺寸卷积核的卷积层,用于削弱振动信号中的高频噪声;随后将压缩后的特征输入到连续的注意力模块和残差块中。随着网络深度的加深,特征变得越来越清晰,可以持续提高残差注意力网络的性能。网络的具体参数如表1所示。

表1 MDARN 的网络参数

网络类型	尺寸	数量	步长
输入			
卷积层	64×1	16	8×1
最大池化层	2×1	16	1×1
阶段1			
残差块1	-	16	-
注意力模块1	-	32	-
卷积层	1×1	32	-
阶段2			
残差块2	-	32	2×1
注意力模块2	-	64	-
卷积层	1×1	64	-
阶段3			
残差块3	-	64	2×1
注意力模块3	-	128	-
卷积层	1×1	128	-
输出			
残差块4	-	128	-
残差块5	-	256	-
残差块6	-	256	-
平均池化层	-	-	4×1
展平层	-	-	-
全连接层	-	10	-

3 实验及结果分析

3.1 实验数据采集

如图3所示,采用滚动轴承故障测试实验平台进行实验数据采集。该实验台主要由电动机、液态水冷器、盘式涡流制动器和转速调节器组成。无线振动传感器安装在电动机风扇端的上方,用于对滚动



图3 滚动轴承故障测试实验平台

动轴承振动信号进行采集。实验过程中的采样频率为 16 000 Hz。如表 2 所示,本文考虑了电动机滚动轴承的 10 种健康状况,包括 1 种健康状况,3 种不同损伤程度的轴承故障,损伤程度分别为 0.1 mm、0.3 mm 和 0.5 mm。

表 2 故障类型

故障类型	故障信息	样本长度	样本数目	简称
滚动体故障	0.1 mm	4 096	1 500	Ball _ L
滚动体故障	0.3 mm	4 096	1 500	Ball _ M
滚动体故障	0.5 mm	4 096	1 500	Ball _ H
内圈故障	0.1 mm	4 096	1 500	Inner _ L
内圈故障	0.3 mm	4 096	1 500	Inner _ M
内圈故障	0.5 mm	4 096	1 500	Inner _ H
外圈故障	0.1 mm	4 096	1 500	Outer _ L
外圈故障	0.3 mm	4 096	1 500	Outer _ M
外圈故障	0.5 mm	4 096	1 500	Outer _ H
正常	无	4 096	1 500	Normal

对于每种健康状态,实验中考虑了如表 3 所示的不同转速(20 Hz 和 30 Hz)和负载(1 A 和 2 A)的运行工况,共采集了 4 种运行工况的数据集。在每个特定转速和负载下,收集不同健康状态下的振动信号。同时,采用滑动窗口的方法将振动信号分割为长度为 4 096 的数据样本,并按照 7 : 3 的比例分为训练集和测试集。值得注意的是,为了验证所提出方法的有效性,使用这种短信号进行故障诊断任务更具挑战性。在实际应用中,可以使用数据点较多的长信号。为了验证所开发的 MDARN 在不同背景噪声下诊断机器故障的有效性,首先将高斯白噪声加入到每个信号中,其信噪比分别为 -2 ~ 10 dB。具体地说,对原始振动信号进行噪声相加。然后,在深度学习模型优化过程中保持噪声振动信号不

表 3 实验数据参数

数据集	转速/Hz	负载/A
A1	20	1
A2	20	1
A3	30	2
A4	30	2

变。其中,每个噪声都是独立产生的,因此对振动信号所添加的噪声是不同的。

3.2 实验流程

本文所提出的故障诊断算法整体流程(图 4)如下。

步骤 1 通过更换电机故障轴承模拟不同健康状况下的轴承故障。待电机平稳运行后开始接收无线传感器采集到的振动信号。

步骤 2 将采集到的振动信号进行分割,获得适用于网络输入的数据样本,并将每种故障类型的数据分为训练集和测试集。

步骤 3 将训练样本输入到 MDARN 中进行网络训练。训练集按照批处理大小为 10 依次输入到网络中,先后通过 3 个混合域残差注意力模块,提取振动信号中的故障特征。

步骤 4 将提取的故障特征最终输入到具有 softmax 激活函数的全连接层中进行故障类型预测。

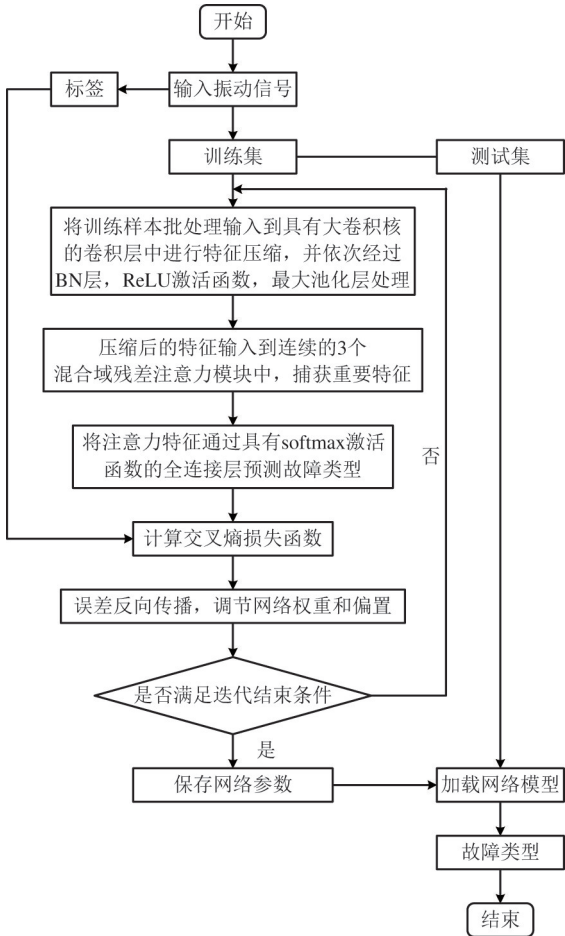


图 4 基于混合域残差注意力网络的滚动轴承故障诊断算法流程图

将预测的结果与真实标签一同输入到交叉熵损失函数中进行损失计算,并将损失值通过反向传播调节各层网络参数。

步骤 5 迭代结束,保存网络参数。

步骤 6 将保存的网络参数加载到 MDARN 中,并将测试集输入到 MDARN 中进行模型性能测试。

3.3 实验结果分析和比较

将 MDARN 与现有先进故障诊断方法进行比较,以进一步说明所提方法的优越性。为了在接下来的实验中使每种方法得到公平的比较,每种诊断方法对每个故障诊断任务运行 10 次,取 10 次实验的平均结果作为最终的实验结果。对比方法的架构基准和超参数设置与 MDARN 相同。一方面,为了验证混合域残差注意力网络的优点,本文使用了 2 个相似的注意力网络 SEnet^[29] 和 CBAM^[30] 进行比较。SEnet 仅使用通道域注意力对输入数据进行特

征加强,而 CBAM 则同时考虑了通道域和空间域注意力。此外,采用经典的基于 CNN (ResNet、WDCNN^[31]) 的故障诊断网络进一步加强实验的说服力。

从表 4 可以看出,首先,MDARN 在不同信噪比 (signal noise ratio, SNR) 噪声影响下都取得了较好的结果,优于其他网络。与经典的通道注意力网络 (SEnet) 和混合域注意力网络 (CBAM) 相比,自下而上的快速前馈过程和自上而下的注意力反馈过程能够更好地捕捉到振动信号的长期依赖信息,从而使网络能够捕获到更多的关键特征,这在 SNR 较低情况下尤为明显。对于不使用注意力机制网络的故障诊断模型 (ResNet、WDCNN),虽然采用 CNN 在无噪声条件下取得了较高的准确率,但随着噪声的逐渐增强,普通的卷积神经网络将无法从损坏的振动信号中提取出故障信息,使得不同类的故障数据无法有效分离。

表 4 各种方法的故障诊断准确率/%

故障诊断方法	SNR								平均准确率
	-2	0	2	4	6	8	10	无	
SEnet	96.66	96.90	98.56	99.65	99.67	100.00	100.00	100.00	98.93
CBAM	97.56	97.56	98.92	99.74	99.56	99.92	100.00	100.00	99.16
ResNet	90.12	92.23	93.13	93.56	94.01	94.99	94.99	98.24	93.91
WDCNN	97.32	98.61	99.66	99.61	99.72	99.91	99.95	100.00	99.35
MDARN	98.56	98.64	99.67	99.89	100.00	100.00	100.00	100.00	99.60

此外,引入混淆矩阵来评价每种故障的诊断性能,4 种工况下的故障诊断结果的混淆矩阵如图 5 所示。从图 5 中可以看到,MDARN 在数据集 A1 中正常轴承数据会被识别为外圈高度损坏,内圈低度损坏会被识别为滚动体中度损坏和内圈中度损坏,滚动体低度损坏则会被误判为内圈低度损坏和滚动体中度损坏。而在数据集 A2、A3 和 A4 中,各个类别的故障诊断准确率都取得了不错的结果,误判率较低。同时,MDARN 在 4 种工况下均没有出现将故障数据误判为正常数据的情况,能够实现轴承故障的准确预警。

为了更好地评价 MDARN 提取的特征是有效的,采用主成分分析的方法将特征层的特征维数降至二维。数据集 A2 的二维特征如图 6 所示。通过

分析图 6 可以得出结论,同一故障类别的特征紧紧聚集在一起,不同故障类别的特征能够有效地分离。这意味着 MDARN 能够有效识别每种类型的轴承故障,展现出了 MDARN 优越的故障诊断性能。

为了进一步验证所提方法的性能,本文在一个更大的数据集——凯斯西储大学 (Case Western Reserve University, CWRU) 轴承故障数据集上^[32] 对 MDARN 进行实验验证。在 CWRU 数据集中,采样频率为 12 000 Hz。共设置了 4 种不同的运行工况:数据集 A (1 797 rpm, 0 hp)、数据集 B (1 772 rpm, 1 hp)、数据集 C (1 750 rpm, 2 hp) 和数据集 D (1 730 rpm, 3 hp)。每个数据集包括正常数据、内圈故障、外圈故障和滚动体故障,故障直径分别为 0.18 mm、0.36 mm 和 0.54 mm。MDARN 的实验结果如表 5 所示,可以

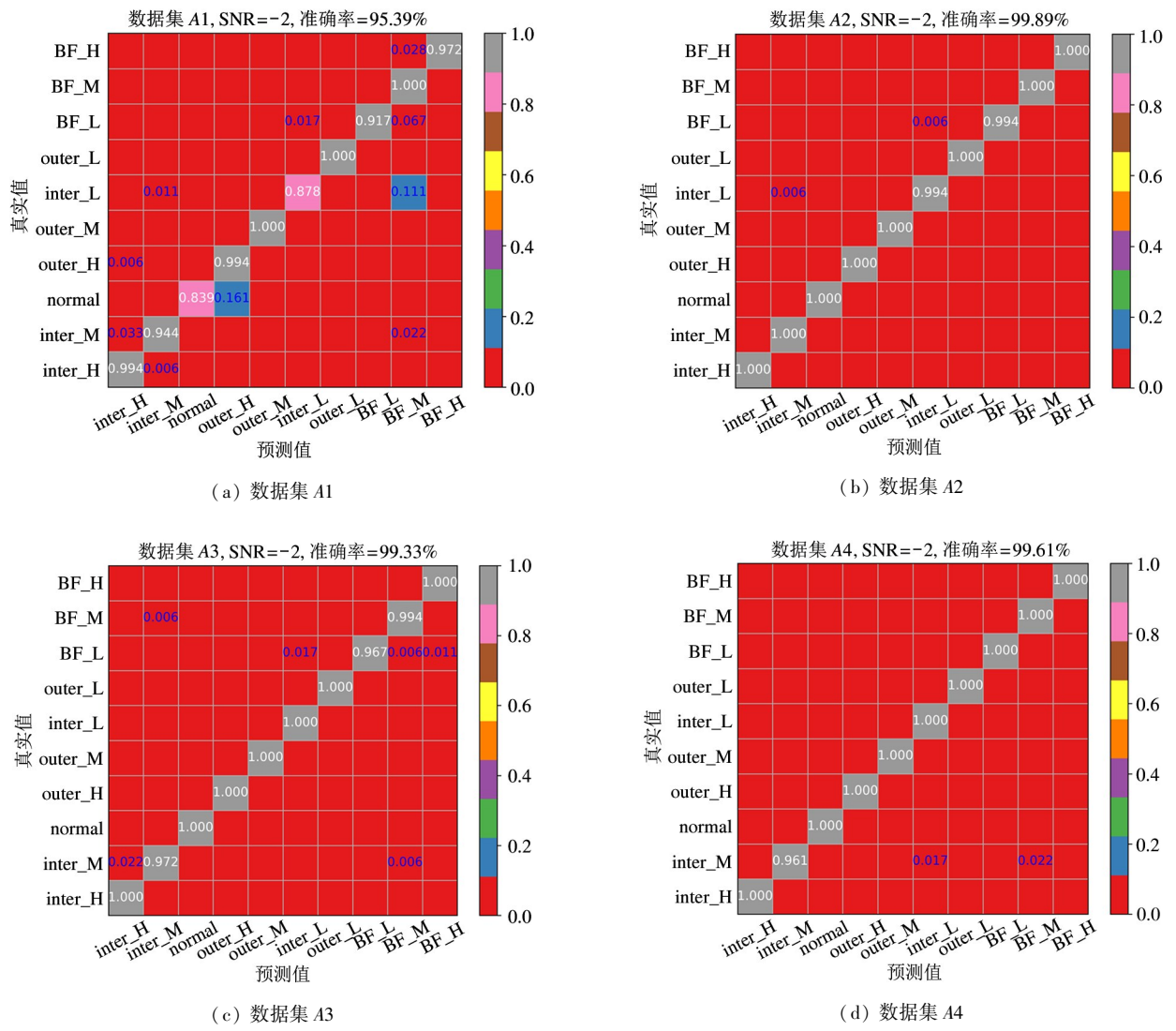


图 5 混淆矩阵

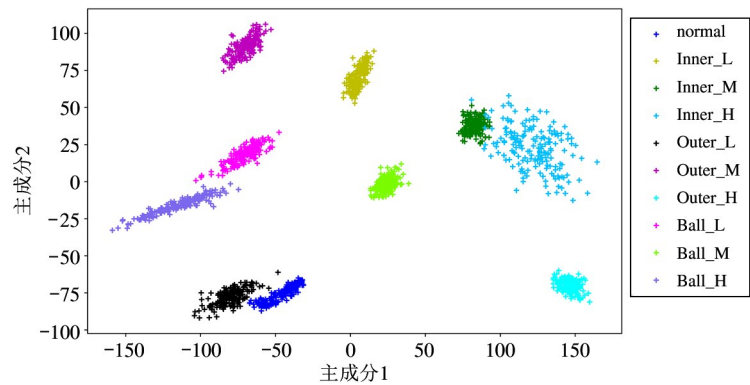


图 6 MDARN 的特征可视化

看出,随着噪声强度的增加,MDARN 仍能保持较高的故障诊断精度,验证了所提方法的有效性。

4 结 论

为了克服滚动轴承故障诊断中的抗干扰能力弱

表5 故障诊断准确率

SNR/dB	平均准确率/%
-2	97.39
0	99.80
2	100.00
4	99.59
6	100.00
8	100.00
10	100.00
无	100.00

的问题,构建了一种基于混合域残差注意力网络的故障诊断模型。首先,采用自下而上和自上而下的结构捕获振动信号的长期依赖信息,并根据残差网络的结构特点,构建了注意力模块。其次,为了加深网络,提高网络的泛化能力,连续堆叠了3个注意力模块。最后,将提取的注意力特征输入到3个连续的残差块中作进一步细化,实现对轴承故障类型的准确识别。对比结果表明,与其他基于深度学习的故障诊断模型相比,MDARN 具有更好的诊断性能和有效性。

参考文献

- [1] 姚明海, 袁惠. 基于优化卷积神经网络的表面缺陷检测[J]. 高技术通讯, 2019, 29(6): 564-569.
- [2] KIRAL Z, KARAGÜLLE H. Simulation and analysis of vibration signals generated by rolling element bearing with defects[J]. Tribology International, 2003, 36(9): 667-678.
- [3] 李倩, 刘惠康, 皮瑶, 等. 基于深度置信网络的制动器故障诊断方法[J]. 高技术通讯, 2021, 31(10): 1075-1080.
- [4] CHANG C W, LEE H W, LIU C H. A review of artificial intelligence algorithms used for smart machine tools[J]. Inventions, 2018, 3(3): 41-69.
- [5] LI B, CHOW M Y, TIPSUWAN W, et al. Neural-network-based motor rolling bearing fault diagnosis[J]. IEEE Transactions on Industrial Electronics, 2000, 47(5): 1060-1069.
- [6] MALHI A, GAO R X. PCA-based feature selection scheme for machine defect classification[J]. IEEE Transactions on Instrumentation and Measurement, 2004, 53(6): 1517-1525.
- [7] SU H, CHONG K T. Induction machine condition monitoring using neural network modeling[J]. IEEE Transactions on Industrial Electronics, 2007, 54(1): 241-249.
- [8] 陈永毅, 张丹, 薛焕新. 基于集合经验模态分解的滚动轴承智能故障诊断方法[J]. 高技术通讯, 2021, 31(5): 559-568.
- [9] TAYYAB S M, ASGHAR E, PENNACCHI P, et al. Intelligent fault diagnosis of rotating machine elements using machine learning through optimal features extraction and selection[J]. Procedia Manufacturing, 2020, 51(1): 266-273.
- [10] SAUCEDO-DORANTES J J, ARELLANO-ESPITIA F, DELGADO-PRIETO M, et al. Diagnosis methodology based on deep feature learning for fault identification in metallic, hybrid and ceramic bearings[J]. Sensors, 2021, 21: 5832.
- [11] MA J, LI C, ZHANG G. rolling bearing fault diagnosis based on deep learning and auto-encoder information fusion[J]. Symmetry, 2022, 14: 13.
- [12] HOANG D T, KANG H J. A motor current signal based bearing fault diagnosis using deep learning and information fusion[J]. IEEE Transactions on Instrumentation and Measurement, 2019, 69(6): 3325-3333.
- [13] BENGIO Y, COURVILLE A, VINCENT P. Representation learning: a review and new perspectives[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2013, 35(8): 1798-1828.
- [14] RUI Z, YAN R, CHEN Z, et al. Deep learning and its applications to machine health monitoring[J]. Mechanical Systems and Signal Processing, 2019, 115: 213-237.
- [15] TANG S, YUAN S, ZHU Y. Convolutional neural network in intelligent fault diagnosis toward rotatory machinery[J]. IEEE Access, 2020, 8: 86510-86519.
- [16] JIANG G Q, YAN H B, XIE J, et al. Multiscale convolutional neural networks for fault diagnosis of wind turbine gearbox[J]. IEEE Transactions on Industrial Electronics, 2018, 66: 3196-3207.
- [17] CHANG C H. Deep and shallow architecture of multilayer neural networks[J]. IEEE Transactions on Neural Networks and Learning Systems, 2015, 26(10): 2477-2486.
- [18] 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.
- [19] JING L Y, ZHAO M. A convolutional neural network

- based feature learning and fault diagnosis method for the condition monitoring of gearbox [J]. Measurement, 2017,111:1-10.
- [20] ZHAO H, SUN S, BO J. Sequential fault diagnosis based on LSTM neural network [J]. IEEE Access, 2018,6(99):12929-12939.
- [21] LI X, JIA X, ZHANG W, et al. Intelligent cross-machine fault diagnosis approach with deep auto-encoder and domain adaptation[J]. Neurocomputing, 2020,383(28):235-247.
- [22] 金江涛, 许子非, 李春, 等. 基于深度学习与混沌特征融合的滚动轴承故障诊断[J]. 控制理论与应用, 2022,39(1):109-116.
- [23] 赵志宏, 赵敬娇, 魏子洋. 基于 BiLSTM 的滚动轴承故障诊断研究[J]. 振动与冲击, 2021,40(1):95-101.
- [24] 丁雪, 邓艾东, 李晶, 等. 基于多尺度和注意力机制的滚动轴承故障诊断[J]. 东南大学学报(自然科学版), 2022,52(1):172-178.
- [25] 赵一瑾. 基于加权密集连接网络和注意力机制的滚动轴承故障诊断[J]. 计算机测量与控制, 2021,29(9):23-28.
- [26] 金余丰, 姚美常, 刘晓锋, 等. 残差网络和注意力机制相结合的滚动轴承故障诊断模型[J]. 机械科学与技术, 2020,39(6):919-925.
- [27] 赵小强, 张亚洲. 利用改进卷积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, 2021,55(12):108-118.
- [28] WANG F, JIANG M, CHEN Q, et al. Residual attention network for image classification[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, USA: IEEE, 2017:6450-6458.
- [29] HU J, SHEN L, SUN G, et al. Squeeze-and-excitation networks[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, USA: IEEE, 2018:7132-7141.
- [30] WOO S, PARK J, LEE J Y, et al. CBAM: convolutional block attention module[C]//2018 European Conference on Computer Vision. Munich, Germany: ECCV, 2018:3-19.
- [31] ZHANG W, PENG G, LI C, et al. A new deep learning model for fault diagnosis with good anti-noise and domain adaptation ability on raw vibration signals[J]. Sensors, 2017,17:425.
- [32] Bearing Data Center. Case Western Reserve University, Cleveland, OH[EB/OL]. [2022-12-03]. <http://www.eecs.case.edu/laboratory/bearing>.

Intelligent fault diagnosis method of rolling bearing based on mixed domain residual attention network

JIA Lixin, CHEN Yongyi, NI Hongjie, ZHANG Dan

(College of Information Engineering, Zhejiang University of Technology, Hangzhou 310023)

Abstract

Mechanical equipment is developing in the direction of large-scale, precision and automation, and mechanical systems are becoming increasingly complex. The automatic detection of mechanical faults is a great challenge considering that those mechanical systems may suffer from featureless catastrophic failures. However, existing fault detection methods have a high misdiagnosis rate when identifying fault types in highly complex industrial systems, and cannot produce accurate fault diagnosis results. To solve this problem, a new fault diagnosis method based on mixed domain residual attention network is proposed in this paper, which takes rolling bearing as the research object as it is the key component of mechanical equipment. This paper aims to improve the performance of fault detection by combining the advantages of automatic learning representation of deep convolutional neural network and key feature extraction ability of channel attention mechanism and spatial attention mechanism. Experimental results show that the proposed method can accurately detect bearing faults and is superior to existing state-of-the-art methods.

Key words: fault diagnosis, rolling bearing, channel attention mechanism, spatial attention mechanism, convolutional neural network(CNN)