

基于 DACNN 的电机滚动轴承故障诊断方法^{*}

贾朱植¹ 刘凯² 刘佳鑫² 祝洪宇² 宋向金³

(1. 辽宁科技大学应用技术学院 鞍山 114051; 2. 辽宁科技大学电子与信息工程学院 鞍山 114051;

3. 江苏大学电气信息工程学院 镇江 212013)

摘要:针对强噪声、跨工况场景下数据分布差异导致传统卷积神经网络(CNN)模型泛化性能低、诊断能力不足的问题,提出一种基于并行卷积核和通道注意力机制的滚动轴承故障诊断方法。构造了带有不同尺度卷积核的并行网络结构,可以在抑制噪声干扰的同时有效提取出数据中的故障特征信息;融合通道注意力机制对卷积层特征提取能力进行增强,提升模型抗噪性能以及跨工况负载下的自适应诊断能力。利用凯斯西储大学轴承数据集训练并测试诊断效果,将该方法与其他方法进行了性能对比。结果表明,在跨工况不同负载情况下,所提方法的诊断平均准确率为 97.3%,在不同信噪比噪声干扰情况下的诊断精度平均达 93.8%,均高于其他比较方法,所提出的方法在复杂多变工况下具有良好的抗噪性能和泛化能力。

关键词:电机;轴承故障诊断;卷积神经网络;注意力机制

中图分类号: TH17 **文献标识码:** A **国家标准学科分类代码:** 460.99

DACNN based fault diagnosis of rolling bearing in motor

Jia Zhuzhi¹ Liu Kai² Liu Jiaxin² Zhu Hongyu² Song Xiangjin³

(1. School of Applied Technology, University of Science and Technology Liaoning, Anshan 114051, China; 2. School of Electronic and Information Engineering, University of Science and Technology Liaoning, Anshan 114051, China;

3. School of Electrical and Information Engineering, Jiangsu University, Zhenjiang 212013, China)

Abstract: In view of the problems of poor generalization ability and insufficient diagnostic capability of traditional convolutional neural network (CNN) model due to the data distribution discrepancy in strong noise environment and across working conditions, a fault diagnosis method for rolling bearings based on parallel convolution kernel and channel attention mechanism is proposed. Using this method, a parallel network structure with different convolution kernel scales was designed to effectively extract feature information from data while suppressing noise interference. Meanwhile, channel attention mechanism was added to enhance the feature extraction capability of the convolutional layer, and improve the anti-noise performance of the model and the adaptive ability in across working conditions. Diagnosis effects were trained and tested by using bearing data set of Case Western Reserve University. The proposed method was compared with peer approaches under different signal-to-noise ratio (SNR) cases and across working conditions, it was shown that the proposed method achieves an average diagnosis accuracy rate of 97.3% in across working conditions and in the variable noise experiment on the bearing dataset from Case Western Reserve University the diagnostic accuracy rate is beyond 93.8%, which are obviously higher than the competing methods; the proposed method have better noise resistance and generalization ability under complex and variable working conditions.

Keywords: motor; bearing fault diagnosis; convolutional neural network; attention mechanism

收稿日期:2024-01-17

^{*} 基金项目:国家自然科学基金(52007078)、辽宁省教育厅基本科研项目(JYTMS20230946)资助

0 引言

滚动轴承摩擦系数小、运转精度高,良好的性能优势使其在异步电机中获得了广泛应用。在电机长时间运行过程中,受时变转速、复杂载荷以及大冲击扰动等恶劣工况影响,滚动轴承的内圈、外圈、滚道及滚动体(球体)极易出现裂纹、点蚀等疲劳损伤或磨损损伤,不健康的工作状态不仅直接影响电机的工作进程,还有可能造成潜在的经济损失和人员伤亡。开展滚动轴承早期故障监测诊断对保证电机安全可靠运行具有重要意义^[1-3]。

以振动、电流等信号的时域、频域或时频域分析为代表的信号处理方法是轴承故障诊断经典方法。这些方法的故障特征提取和识别过程受限于信号处理技术,诊断结果分析依赖于先验知识和专家经验^[3-5]。随着工业大数据时代的到来,海量状态监测数据被不断收集和存储,通过非线性结构学习数据内在规律、构建样本和结果之间映射关系的深度学习方法逐渐成为故障诊断领域的研究热点^[6-10]。

深度学习中的卷积神经网络(convolutional neural networks, CNN)是有监督的深度学习算法,无需对原始数据预处理就能实现端到端的滚动轴承故障诊断^[6-7]。曲建岭等^[11]建立了基于一维 CNN 的自适应故障诊断模型,解决了传统 CNN 模型自适应性差的问题,由于该模型采用的是单尺度的卷积核,无法实现微弱故障特征信息的有效提取。Ye 等^[12]注意到卷积核对 CNN 模型学习特性的影响并提出一种自适应核稀疏网络,但是该网络从信号中识别特征,并没有充分利用卷积核提取特征的能力。在强噪声、跨工况的工业背景下,基于单尺度卷积核的传统 CNN 模型难以同时兼顾运算速度和微小故障特征信息的全面挖掘,诊断性能受到很大限制。作为改进,全大英等^[13]提出双通道特征融合 CNN,孟良等^[14]提出一种大卷积层和并行卷积层相结合的网络结构 PCNN,该网络更有利于降低计算复杂度和特征的快速提取。实践表明,上述通过模型结构复杂化来改善 CNN 诊断性能的思路在提高网络学习效率方面的作用是有限的,不仅容易导致网络性能退化也会限制工况变化时的网络泛化能力。学者们注意到不同的工况会导致数据中故障特征分布发生变化,因此提出了多尺度 CNN 来提高网络学习能力^[15-18]。多尺度 CNN 通过设置多个卷积通路来分别获取不同尺度下的信息,需要充分考虑不同尺度间信息差异并对有用信息予以格外关注,否则极易受到无关分量和冗余信息的干扰^[18-19]。为此,唐斯等^[16]提出了加入通道注意力的多尺度 CNN 抗噪模型。该模型在提取不同尺度的特征时利用通道注意力自适应地选择包含故障特征的通道来提高模型抗噪能力。康涛等^[19]设计了融合多注意力机制的 CNN 自适应抗噪模型,该模型利用通道和时间的复合注意力机制优化学习过程,从不同角度抑制噪声及干扰信息的影响,自适应地增强故障特征。赵小强等^[17]设计了双路并

行多尺度神经网络,通过通道注意力机制的引入解决数据突变性和差异性的问题。研究表明,传统 CNN 模型在强噪声、跨工况条件下的诊断性能存在着特征提取能力不足、准确率不高的缺陷,将改进 CNN 网络结构和优化学习机制结合起来提高 CNN 模型的诊断能力是全新的解决思路。

基于此,本文在传统 CNN 模型基础上,提出一种带有并行卷积核和注意力机制的卷积神经网络(convolutional neural networks with double first-layer and attention model, DACNN)模型。通过设计带有不同卷积核尺度的并行网络结构,该模型能够自适应地提取数据样本中的不同层级特征,在抑制噪声干扰的同时,能够有效挖掘数据中的关键故障特征信息;通过融合通道注意力机制对卷积层特征提取能力进行增强,针对大卷积核的特征感受范围过大无法捕捉到小范围内重要特征的情况,在第 1 层大卷积核之后加入 SE-Block (squeeze-and-excitation block) 模块,在第 1 层小卷积核之后加入 AMM(attention model based on global-max-pooling)模块,分别采用全局平均池化和全局最大池化方法,以最大限度保留大范围内重要信息的同时突出小范围感受野中的重要特征,提升模型抗噪性能以及跨工况、不同负载下的自适应诊断能力。

1 CNN 原理

1.1 CNN

CNN 由多个具有过滤作用的神经网络构成,如图 1 所示,可分成滤波器与分类器两个部分。其中,滤波器包括卷积层、池化层与激活层 3 个基本单元,用来提取输入信号的特征;分类器包括全连接隐含层和 Softmax 层,前者将滤波器部分学习到的特征展开铺平,后者完成在此基础之上的特征分类。

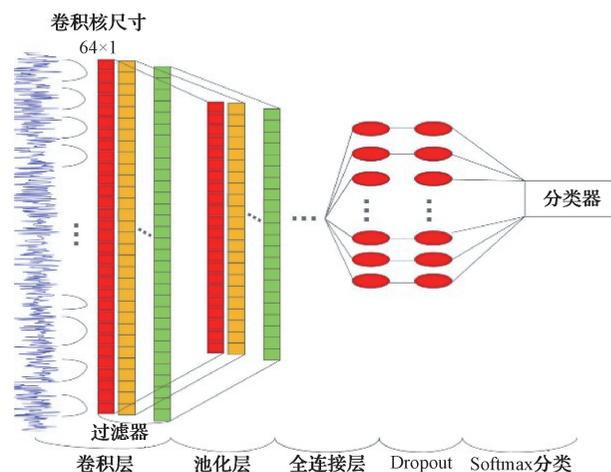


图 1 卷积神经网络原理结构

Fig. 1 Structural diagram of convolutional neural network

卷积层将输入的局部区域与滤波器的卷积核进行卷积操作,实现局部特征提取。卷积层采用权重比参数共享

的方式使每个卷积层都有固定的权重比,大幅减少网络训练参数,降低网络的系统内耗。卷积层中,每个卷积层的卷积核大小不尽相同,卷积层内通过固定步长对输入进行遍历。假设第 j 层是卷积层,则卷积过程如下:

$$y_i^j(k) = c_i^j \cdot x_j(k) \quad (1)$$

式中: c_i^j 表示第 j 层的第 i 个卷积核的权重比和偏置; $x^j(k)$ 表示第 j 层的第 k 个局部区域。

卷积层运算使得原本分布散乱的输入数据具有线性可分离特性。为增强网络模型对数据特征线性可分性,采用激活函数对卷积运算得到的输出值进行非线性变换。CNN 中常用的激活函数包括 Sigmoid 函数、双曲正切 Tanh 函数以及线性整流单元 ReLU(rectified linear unit) 函数。不同的激活函数对应着不同的函数表达式和非线性变换特性。ReLU 函数表达式如下:

$$f(y_i^j(k)) = \max\{0, y_i^j(k)\} \quad (2)$$

式中: $y_i^j(k)$ 代表卷积层的输出; $f(y_i^j(k))$ 是 $y_i^j(k)$ 经过激活函数处理之后的输出值。

数据经过激活函数处理后,在池化层进行降采样操作以降低数据宽度、减少神经网络参数,进而缩短网络运行时间,防止过拟合现象。常用的池化函数包括均值池化和最大值池化。前者将感知域的神经元均值作为输出值,后者将感知域中的最大值作为输出值。最大值池化如下:

$$p^j(k) = \max_{(k-1)W+1 \leq l \leq kW} \{f(y_i^j(k))\} \quad (3)$$

式中: $p^j(k)$ 为池化层的输出值; $f(y_i^j(k))$ 为激活层的输出值; W 为池化层的宽度。

原始输入数据经过多层卷积、激活和池化操作后,大部分重要特征被筛选出来。全连接层将最后一层池化层输出的特征展开铺平,形成一维特征向量,进行最后的特征筛选,再通过 Softmax 分类器对筛选后的特征分类,实现对原始输入数据分类的目的。

2 基于 DACNN 的轴承故障诊断方法

采用大卷积核的 CNN 模型计算能力有限,且由于特征提取时无法考虑到不同特征的映射程度,信息处理效率和特征识别准确性受到很大限制^[6]。相比之下,并行卷积核一维卷积神经网络模型(DCNN)在第 1 层卷积层使用两个并行的卷积核(图 2)。这种模型结构通过并行计算减少了数据读取次数,降低了数据频繁传输造成的时延和功耗,但是由于每个卷积核只能提取到固定的特征,卷积核捕捉高低频特征的能力仍然不足,无法满足强噪声、跨工况下的轴承故障诊断要求。为解决上述问题,本文在对 DCNN 模型改进基础上,提出一种带有并行卷积核和通道注意力机制的一维卷积神经网络模型 DACNN。模型结构如图 3 所示。

2.1 DACNN 模型描述

本文提出的 DACNN 模型共包含 5 个卷积层,5 个池化层,一个全连接隐含层和一个 Softmax 层。除第 1 层的

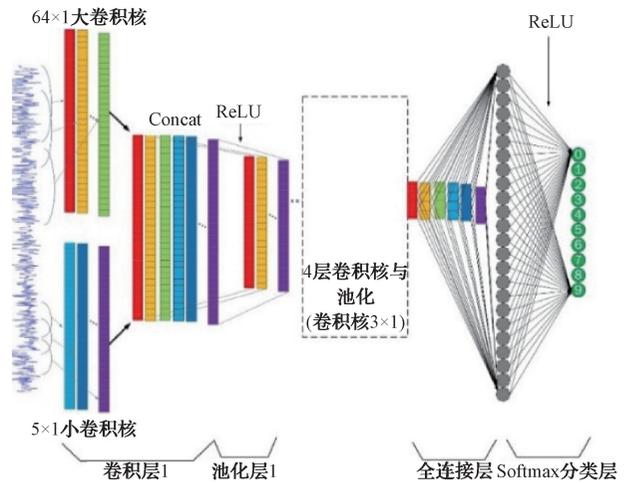


图 2 DCNN 模型结构

Fig. 2 Structure diagram of DCNN

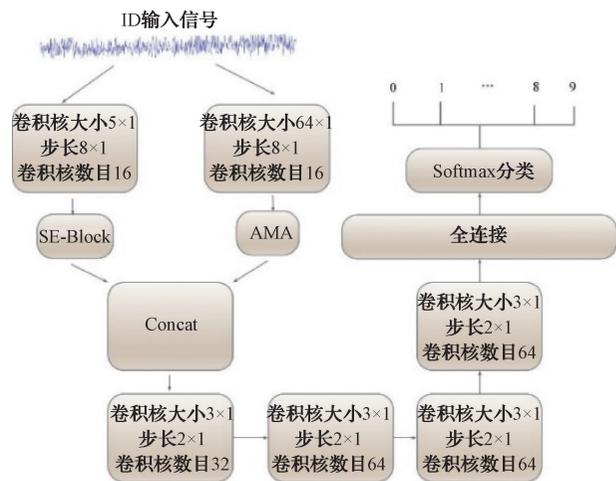


图 3 DACNN 模型结构

Fig. 3 Structure diagram of DACNN

两个卷积核大小分别为 64×1 和 5×1 之外,其他卷积层的卷积核大小均为 3×1 。这样的卷积核参数少,利于加深网络和抑制过拟合。每层卷积操作后均进行批量归一化处理(batch normalization, BN)。在 DACNN 模型中,第 1 层卷积层使用了两个并行的卷积核,大卷积核用以提取局部大范围重要特征,小卷积核用以提取小范围内的重要特征。为了抑制无用信息,提高网络处理效率和准确性,在第 1 层大卷积核模块后引入通道注意力机制 SE-Block 模块,并在 SE-Block 中使用全局平均池化策略以最大限度地保留大范围感受野内的重要信息;在第 1 层小卷积核模块后加入 AMM 模块,以获取小卷积核的小范围感受野中的局部重要特征。SE-Block 和 AMM 模块结构分别如图 4 和 5 所示。除第 1 层之外,从第 2 层开始,均采用最大值池化策略。为解决梯度弥散问题,采用 ReLU 函数作为卷积网络的激活函数。

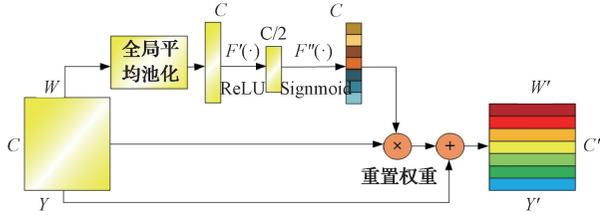


图4 SE-Block 模块结构

Fig. 4 Structure diagram of SE-Block module

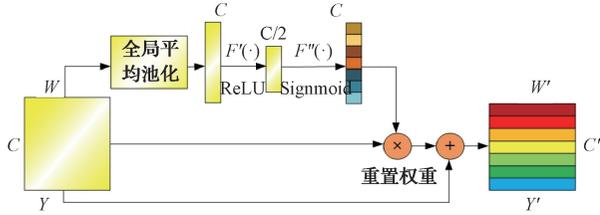


图5 AMM 模块结构

Fig. 5 Structure diagram of AMM module

DACNN 模型基本工作过程是:首先将诊断信号输入第 1 个卷积层,经过两个卷积核及 ReLU 激活层提取特征图,然后进行 Concat 连接操作,再通过 4 层 3×1 的卷积核进行计算,将最后一个池化层的特征图与全连接隐含层相连,经过 ReLU 激活后传递到最后的 Softmax 层完成特征分类。

2.2 卷积神经网络的权值优化机制

DACNN 模型需要优化的参数主要包括卷积核参数 k 、下采样层权值 β 、全连接层网络权值 W 及各层偏置 b 。本文选用如式(4)所示的交叉熵损失函数 $J(\theta)$ 作为目标函数,对每层网络输出值与目标值的一致性进行评价。

$$J(\theta) = -\frac{1}{m} \sum_{i=1}^m (y^{(i)} \log(h\theta(x^{(i)})) + (1 - y^{(i)}) \log(1 - h\theta(x^{(i)}))) \quad (4)$$

式中: m 为输入时的批量大小; $h\theta(x^{(i)})$ 为网络计算出的预测值; $y^{(i)}$ 为目标值; $J(\theta)$ 为损失函数输出值。

确定目标函数权重导数值的方法是通过链式法则,利用误差反向传播算法从全连接层开始由后往前逐层计算池化层和卷积层目标函数关于所有权值的导数值,然后采用随机梯度下降优化算法(stochastic gradient descent, SGD)进行更新权值,求解出使目标函数最小时的权值。当目标函数取得最小值时,对应的权重值达到最优,优化过程停止。上述过程公式如下:

$$g = \frac{1}{m} \nabla \theta \sum_{i=1}^m L(f(x_i; \theta), y_i) \quad (5)$$

$$v = \alpha v - (1 - \alpha)g \quad (6)$$

$$\theta^* = \theta + v \quad (7)$$

式中: $L(\cdot)$ 、 $f(\cdot)$ 分别为目标函数值和输出值; θ 为卷积神经网络中的参数; m 为批量大小; x_i 为网络的输入; y_i

为预测值; v 为抑制 SGD 振荡加入的动量; θ^* 为优化后的网络参数。

2.3 通道注意力机制

为提高 DACNN 模型故障特征提取能力,本文在 DC-NN 网络基础上引入通道注意力机制 SE-Block 模块,旨在通过原有数据的自动学习来增强网络对重要特征的关注度并抑制无用信息。SE-Block 操作过程分为挤压(squeeze)和激励(excitation)两个步骤,前者用来获取每个通道的全局特征向量,后者用来学习每个通道的特征权值。图 6 所示为 SE-Block 模块在 DACNN 中的结构。图 6 中, X 为卷积层输入, F_{tr} 为卷积操作, U 为卷积层输出, \tilde{X} 中不同颜色的特征向量代表 U 中的特征向量经 SE-Block 操作后的结果。卷积层输出 U 的计算如式(8)所示。

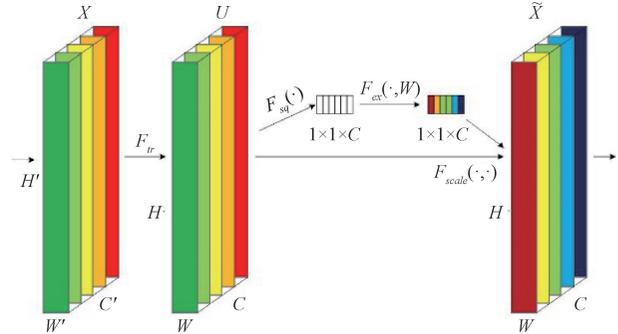


图6 SE-Block 在 DACNN 中的结构

Fig. 6 Structure diagram SE-Block module in DACNN

$$u_c = v_c * X = \sum_{s=1}^{c'} v_c^s * X^s \quad (8)$$

式中: $*$ 表示卷积操作; v_c 为卷积第 c 个过滤器内核集参数; u_c 为卷积操作后第 c 个过滤器的输出。

挤压过程 F_{sq} 的计算如式(9)所示。

$$z_c = F_{sq}(u_c) = \frac{1}{h \times w} \sum_{i=1}^h \sum_{j=1}^w u_c(i, j) \quad (9)$$

式中: z_c 为挤压输出结果; h 、 w 分别为通道长宽。 F_{sq} 通过全局平均池化将通道内的全部信息关联起来,获得整个通道上响应的全局分布,使靠近输入层的卷积层获得基于全局的感受野,实现对整个通道的作用。

激励过程 F_{ex} 的计算如式(10)所示。

$$s = F_{ex}(z, W) = \sigma(g(z, W)) = \sigma(W_2 \delta(W_1 z)) \quad (10)$$

式中: s 为激励输出结果; δ 为 ReLU 函数; σ 为 Sigmoid 函数; W 生成每个通道的权重, W_1 和 W_2 操作限制模型的复杂性使之一般化。 F_{ex} 过程由两层全连接层实现,通过 Sigmoid 函数将结果限制于 $[0, 1]$ 之间,再将所得值与 U 相乘,所得结果作为下一层的输入。

SE-Block 最后进行权重更新操作,操作计算如式(11)所示。

$$\tilde{x}_c = F_{scale}(u_c, s_c) = s_c u_c \quad (11)$$

式中： \tilde{x}_c 为 SE-Block 的输出； s_c 和 u_c 分别为激励输出和卷积输出结果。权重更新操作将激励操作 F_{ex} 中获得的权重与卷积层的输出 U 相乘，实现权重更新。在该操作后，卷积层输出中的重要特征得到强化，无关噪声得到抑制。

2.4 一维卷积神经网络超参数设计

学习率影响卷积神经网络模型的收敛速度、收敛程度和模型精度。通过分析学习率与损失值之间的关系，找出最优学习率，以此为基础，对网络参数进行调试，得到效果良好的 DACNN 网络模型。

1) 学习率寻优

学习率决定了参数更新过程中模块参数的调整程度，学习率太小，收敛速度变慢，学习率太大则损失值过大甚至振荡。本文设置批量尺寸 (Batch_size) 为 128，采用倍增方式寻找最优学习率。学习率初始值设定为 10^{-5} ，每 5 次迭代后将学习率减小 5 倍，计算不同学习率对应的损失值。随着迭代次数增加，损失值逐渐下降并趋于稳定。经过 50 次迭代，当学习率的对数为 -1.8563 时，即学习率

为 0.031 2 时，损失值达到最小，此时的模型性能达到最好。

2) 卷积神经网络结构参数调试

学习率固定后，通过反复实验选取网络结构的超参数。实验在 Tensorflow 框架上进行，使用 Python3.8 版本，计算机配置为 Intel(R) Core(TM) i7-7500UCPU@2.70 GHz 2 904 MHz，运行内存为 16 GB。每次参数调试实验均经过 20 次迭代，调试过程及结果如表 1 所示。表 1 中，不同的实验序号对应着不同的参数调试过程。实验 1~7 和实验 8~11 分别用于分析卷积核大小和批量大小对模型诊断精度的影响；实验 12、13，实验 14、15 和实验 16、17 分别用于分析 Dropout、批量正则化和标准化对模型的影响；实验 18 用于分析卷积核正则化的影响。经过 18 次调试发现，实验 6 对应的验证集准确率达到 99.4%，表明与该实验对应的网络结构参数可使模型诊断性能达到最好。在 2.1 节提出的 5 层网络结构基础上，经过数次参数调整，得到 DACNN 网络模型详细参数如表 2 所示。

表 1 DACNN 参数调试过程

Table 1 Parameter debugging procedure of DACNN

| 实验 | 第 1 层大卷积核 | 第 1 层小卷积核 | 卷积核步长 | 标准化 | Batch size | Batch Norm | Dropout | 卷积核正则化 | 验证集准确率/% |
|----|-----------|-----------|-------|-----|------------|------------|---------|--------|----------|
| 1 | 32 | 5 | 8 | 否 | 128 | 是 | 0 | 否 | 99.0 |
| 2 | 64 | 7 | 8 | 否 | 128 | 是 | 0 | 否 | 98.4 |
| 3 | 64 | 3 | 8 | 否 | 128 | 是 | 0 | 否 | 98.8 |
| 4 | 32 | 7 | 8 | 否 | 128 | 是 | 0 | 否 | 98.8 |
| 5 | 128 | 5 | 8 | 否 | 128 | 是 | 0 | 否 | 99.1 |
| 6 | 64 | 5 | 8 | 否 | 128 | 是 | 0 | 否 | 99.4 |
| 7 | 128 | 5 | 8 | 否 | 128 | 是 | 0 | 否 | 99.1 |
| 8 | 64 | 5 | 8 | 否 | 32 | 是 | 0 | 否 | 96.5 |
| 9 | 64 | 5 | 8 | 否 | 64 | 是 | 0 | 否 | 98.8 |
| 10 | 64 | 5 | 8 | 否 | 256 | 是 | 0 | 否 | 98.8 |
| 11 | 128 | 5 | 8 | 否 | 512 | 是 | 0 | 否 | 67.1 |
| 12 | 64 | 5 | 8 | 否 | 64 | 是 | 0.1 | 否 | 98.5 |
| 13 | 64 | 5 | 8 | 否 | 128 | 是 | 0.1 | 否 | 97.9 |
| 14 | 64 | 5 | 8 | 否 | 128 | 否 | 0 | 否 | 98.7 |
| 15 | 64 | 5 | 8 | 否 | 128 | 否 | 0 | 否 | 98.4 |
| 16 | 64 | 5 | 8 | 是 | 64 | 是 | 0 | 否 | 97.6 |
| 17 | 64 | 5 | 8 | 是 | 128 | 是 | 0 | 否 | 97.2 |
| 18 | 64 | 5 | 8 | 否 | 128 | 是 | 0 | 是 | 98.2 |

3 实验结果及分析

3.1 CWRU 数据集

采用凯斯西储大学 (Case Western Reserve University, CWRU) 滚动轴承中心的公开数据集验证 DACNN 模型诊断性能。CWRU 实验平台如图 7 所示。实验轴承为 SKF 深沟球轴承，型号为 6205-2RS JEM。对轴承球体、

内圈、外圈进行钻孔处理，以模拟 9 种不同损伤程度的故障，孔洞直径分别为 0.007、0.014 和 0.021 in (1 in = 25.4 mm)。在轴承驱动端和风扇端采集加速度信号，采样频率为 12 kHz。实验在 0~3 HP (1 HP ≈ 746 W) 负载状态下进行，4 种不同负载情况对应的转速为 1 797、1 772、1 750 和 1 730 r/min。

实际采集到的原始信号是具有周期性特点的一维时

表 2 DACNN 网络结构参数

Table 2 Structural parameters of DACNN

| 编号 | 网络层 | 卷积核大小 | 卷积核数目 | 输出大小 | 零补 |
|----|---------|----------|-------|--------|----|
| 1 | 卷积 1_1 | 64×1/8×1 | 16 | 256×16 | 否 |
| 2 | 卷积 1_2 | 5×1/8×1 | 16 | 256×16 | 否 |
| 3 | Concat | — | 32 | 256×32 | 否 |
| 4 | 池化 1 | 2×1/2×1 | 16 | 128×32 | 否 |
| 5 | 卷积 2 | 3×1/1×1 | 32 | 128×32 | 否 |
| 6 | 池化 2 | 2×1/2×1 | 32 | 64×32 | 否 |
| 7 | 卷积 3 | 3×1/1×1 | 64 | 64×64 | 否 |
| 8 | 池化 3 | 2×1/2×1 | 64 | 32×64 | 否 |
| 9 | 卷积 4 | 3×1/1×1 | 64 | 30×64 | 是 |
| 10 | 池化 4 | 2×1/2×1 | 64 | 15×64 | 否 |
| 11 | 卷积 5 | 3×1/1×1 | 64 | 13×64 | 是 |
| 12 | 池化 5 | 2×1/2×1 | 64 | 6×64 | 否 |
| 13 | 全连接 | 128 | 1 | 128×1 | — |
| 14 | Softmax | 10 | 1 | 10 | — |

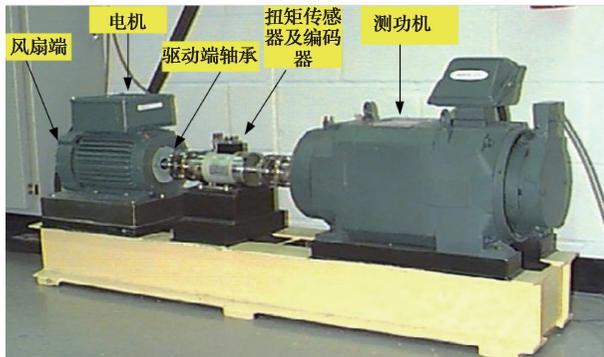


图 7 CWRU 滚动轴承系统平台

Fig. 7 CWRU rolling bearing system platform

序信号。对于这种类型的数据,为防止网络训练时出现过拟合现象,提高 DACNN 模型泛化能力,采用重叠采样技术对训练数据进行增强处理,增加数据多样性,操作如图 8 所示。在原始信号输入时,将信号分成若干数据段,每段信号与相邻段信号之间有部分重叠,再对训练样本重叠采样。每次采样的数据长度为 2 048,偏移量为 28,以此增加训练样本数量,满足卷积神经网络的数据需求,提高泛化性能。

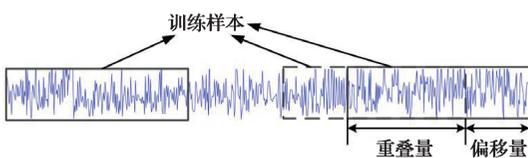


图 8 重叠采样示意图

Fig. 8 Illustration of overlapping sampling

根据负载状态将实验数据分成 A、B、C、D 4 个数据

集,分别对应 0、1、2 和 3 HP 4 个不同的负载状态,每个数据集包括 3 种故障类型(球体、内圈、外圈)和 4 种故障严重程度(轻微、中度、重度、无故障)的数据,共 10 种数据,如表 3 所示。每个数据集包含 7 000 个训练样本、2 000 个验证样本和 1 000 个测试样本。实验过程中,验证集和测试集不使用重叠采样技术,以确保性能验证时所使用的数据具有独特性。

表 3 实验数据集描述

Table 3 Description of experimental data set

| 标签 | 故障描述 | 故障直径/in | 负载情况/HP | | | |
|----|--------|---------|---------|-------|-------|-------|
| | | | 1 797 | 1 772 | 1 750 | 1 730 |
| 0 | 球体轻微故障 | 0.007 | 0 | 1 | 2 | 3 |
| 1 | 球体中度故障 | 0.014 | 0 | 1 | 2 | 3 |
| 2 | 球体重度故障 | 0.021 | 0 | 1 | 2 | 3 |
| 3 | 内圈轻微故障 | 0.007 | 0 | 1 | 2 | 3 |
| 4 | 内圈中度故障 | 0.014 | 0 | 1 | 2 | 3 |
| 5 | 内圈重度故障 | 0.021 | 0 | 1 | 2 | 3 |
| 6 | 外圈轻微故障 | 0.007 | 0 | 1 | 2 | 3 |
| 7 | 外圈中度故障 | 0.014 | 0 | 1 | 2 | 3 |
| 8 | 外圈重度故障 | 0.021 | 0 | 1 | 2 | 3 |
| 9 | 无故障 | 0 | 0 | 1 | 2 | 3 |

3.2 消融实验

随机选择 1 HP 负载、信噪比(SNR)为 10 条件下的数据集进行消融实验,分析 SE-Block 或 AMM 模块加入前后的网络性能变化,评价指标包括准确率、精确率和召回率。实验结果如表 4 所示。从表 4 可知,DCNN 的初始性能明显优于支持向量机(SVM)网络,准确率、精确率和召回率 3 项指标分别高出 5.65%、11.9% 和 12%。在 DCNN 中加入 AMM 或 SE-Block 注意力机制模块之后,3 项指标均有不同程度的提升,表明两类模块都具有优化网络性能的作用,但是在改进程度方面两者差别不大。当 DCNN 同时搭载 AMM 和 SE-Block 模块时,DACNN 模型的准确率、精确率和召回率分别高达 97.4%、95.75% 和 95.75%,相比于未加注意力机制的 DCNN 模型,3 项指标分别提高了 2.85%、2.60% 和 1.85%,体现出 AMM 和 SE-Block 模块在增强可用特征信息提取能力、实现通道自

表 4 不同模型的性能指标对比

Table 4 Comparison of performance metrics of different models

| 网络结构 | models | | (%) | | |
|----------|--------|----------|-------|-------|-------|
| | AMM | SE-Block | 准确率 | 精确率 | 召回率 |
| SVM | — | — | 88.9 | 81.25 | 81.9 |
| DCNN | — | — | 94.55 | 93.15 | 93.9 |
| AMM-DCNN | ✓ | — | 96.4 | 94.5 | 94.3 |
| SE-DCNN | — | ✓ | 96.45 | 93.72 | 95.1 |
| DACNN | ✓ | ✓ | 97.4 | 95.75 | 95.75 |

适应校准方面的优势,也表明 DACNN 模型具有良好的特征识别性能。

3.3 跨工况不同负载情况下的故障诊断性能

1) 数据集构建

随机选择 A、B、D 3 种负载状态下的实验数据作为训练集对网络模型进行训练,使用 B、D 单一负载状态下的实验数据、以及不同负载状态对应的实验数据组合集 ABC、ABCD 作为测试集进行测试。训练集和测试集数据的不同组合方式如表 5 所示。

表 5 训练集和测试集数据的不同组合

Table 5 Different combination of data set for training and testing

| 训练集 | A | A | B | D | B | B |
|-----|---|---|---|---|-----|------|
| 验证集 | B | D | D | B | ABC | ABCD |

2) 诊断准确性分析

在相同实验条件下,对 SVM、WDCNN、DCNN 和 DACNN 4 种模型的故障诊断准确率进行比较,结果如图 9 和表 6 所示。其中,SVM 采用径向基核函数。图 9 和表 6 中的数据均为 8 次迭代实验的平均值。通过比较可以看出,SVM、WDCNN、DCNN 算法的诊断平均准确率分别为 84.7%、91%和 95%,而 DACNN 的诊断准确率高达 97.3%,表明所提出的 DACNN 模型有良好的变负载自适应能力。

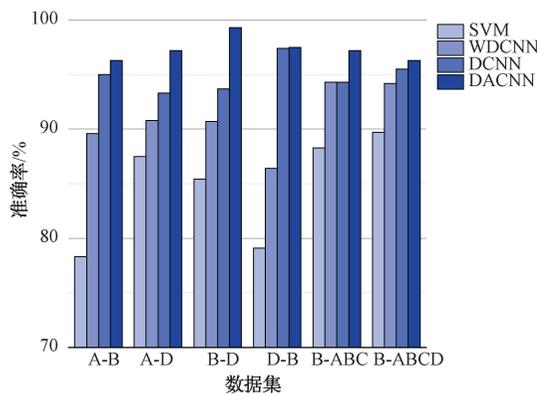
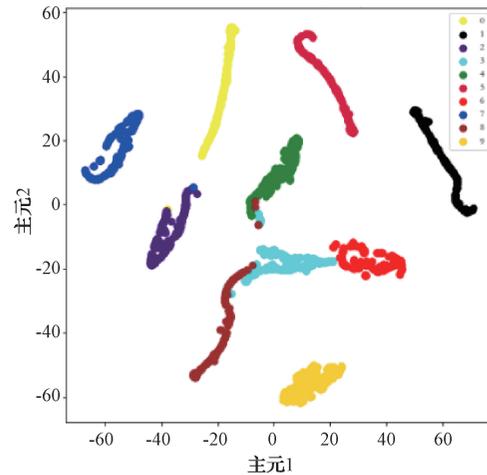


图 9 跨工况情况下 DACNN 与其他模型诊断准确率比较
Fig. 9 Comparison of diagnostic accuracy between DACNN and other models under across working conditions

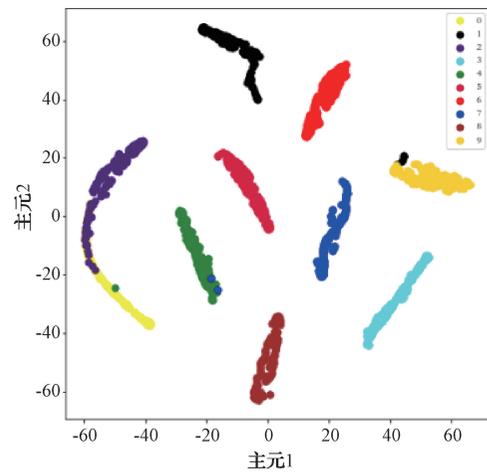
表 6 跨工况情况下 DACNN 与其他模型的诊断准确率
Table 6 Diagnostic accuracy of DACNN and other models under across working conditions (%)

| 数据集 | A-B | A-D | B-D | D-B | B-ABC | B-ABCD |
|-------|------|------|------|------|-------|--------|
| SVM | 78.3 | 87.5 | 85.4 | 79.1 | 88.3 | 89.7 |
| WDCNN | 89.6 | 90.8 | 90.7 | 86.4 | 94.3 | 94.2 |
| DCNN | 95.0 | 93.3 | 93.7 | 97.4 | 94.3 | 95.5 |
| DACNN | 96.3 | 97.2 | 99.3 | 97.5 | 97.2 | 96.3 |

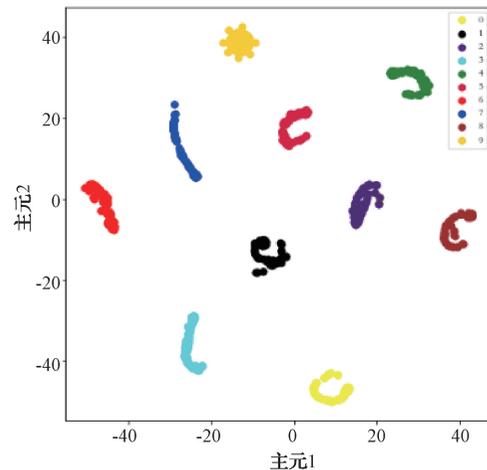
为了直观地比较 DACNN 与其他模型的诊断准确率,采用可视化技术呈现不同模型分类过程中的聚类效果。图 10 所示为 WDCNN、DCNN 与 DACNN 模型在数据集为 B-D 组合情况下最后一层隐藏层的特征聚类效果



(a) WDCNN



(b) DCNN



(c) DACNN

图 10 DACNN 和其他模型的可视化输出
Fig. 10 Visual output of DACNN and other models

果。输出特征经过 *t*-SNE 技术处理后降至 2 维,横纵坐标分别代表信号中蕴含信息最多的两个主元成分。由图 10 可知,DCNN 模型具有良好的聚类效果,但仍有少许特征未完全分开,WDCNN 模型聚类效果比 DCNN 模型稍差,不同故障特征未完全分开,使得诊断出现偏差。相比之下,本文提出的 DACNN 模型聚类效果优秀,该模型将故障特征聚集在更小区域,且不同故障特征之间距离间隔较大,表明 DACNN 模型具有更好的变负载自适应能力。

3.4 噪声环境下的故障诊断性能分析

在 CWRU 数据集上加入高斯白噪声,以模拟电机实际运行时的强噪声环境工作状态,对模型性能进行评估,分析模型抗噪能力。采用 SNR 作为噪声强弱的评估标准。SNR 的公式如下:

$$SNR(dB) = 10\lg\left(\frac{P_s}{P_n}\right) \quad (12)$$

式中: P_s 和 P_n 分别代表信号和噪声的有效功率。SNR 越高,说明混在信号中的噪声越小,SNR=0 时,说明信号与噪声的有效功率相当。

以表 5 数据为基础,在 B-ABCD 数据集中加入高斯白噪声以测试各模型的抗噪性能。实验数据集包含 700 个训练样本,200 个验证样本,100 个测试样本,样本长度为 2 048,训练样本采用重叠采样技术,偏移量为 28。实验结果如图 11 和表 7 所示。由图 11 可知,在变负载情况下,SVM 的性能最差,WDCNN 与 DCNN 模型的准确率接近,DACNN 与 SVM、DCNN、WDCNN 3 种模型相比更优越。由表 7 可知,在加入信噪比为 -2 或更低噪声时,SVM 模型的诊断精度均在 65% 以下。DCNN 模型与 WDCNN 模型的表现平均在 83% 左右,DACNN 的平均表现在 94% 左右,在信噪比为 -10 的强噪声情况下,SVM 的准确率仅 40.2%,无法完成诊断任务;WDCNN 的诊断精度较低。分别为 55.9% 和 64.9%。相比之下,DACNN 的诊断精度为 86.5%,比其他 3 种模型至少高出 20%,抗噪性能远好于其他 3 种模型。

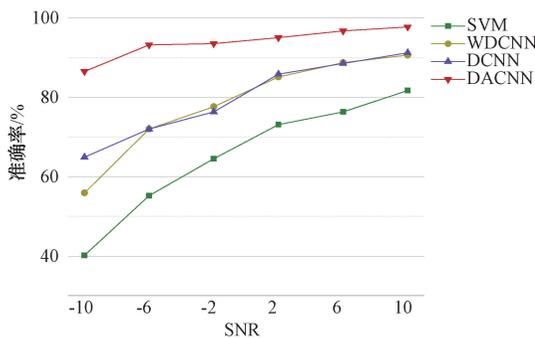


图 11 噪声环境下 DACNN 与其他模型诊断准确率比较
Fig. 11 Comparison of diagnostic accuracy of DACNN with other models in noisy environment

表 7 不同信噪比条件下 DACNN 与其他模型的诊断准确率

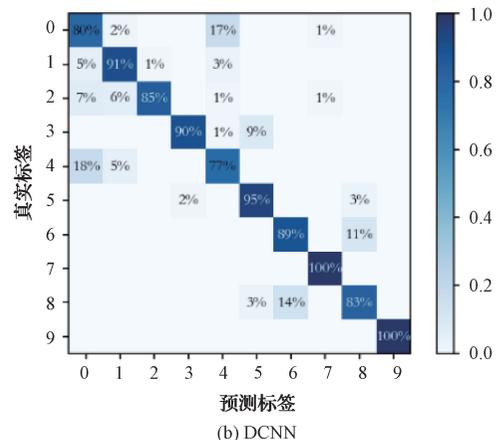
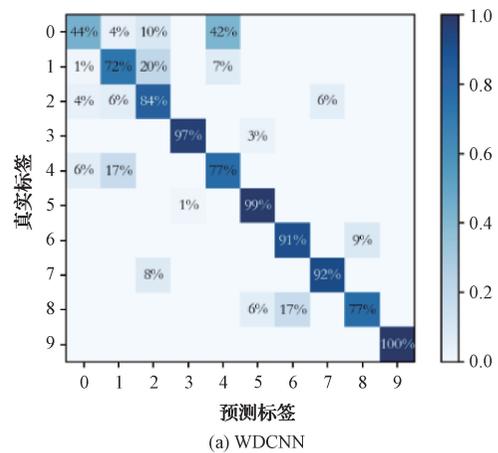
Table 7 Diagnostic accuracy of DACNN and other models under different SNR conditions (%)

| 模型 | SNR=-10 | SNR=-6 | SNR=-2 | SNR=2 | SNR=6 | SNR=10 |
|-------|---------|--------|--------|-------|-------|--------|
| SVM | 40.2 | 55.2 | 64.5 | 73.1 | 76.3 | 81.7 |
| WDCNN | 55.9 | 72.0 | 77.6 | 85.1 | 88.7 | 90.6 |
| DCNN | 64.9 | 72.0 | 76.3 | 85.8 | 88.5 | 91.2 |
| DACNN | 86.5 | 93.2 | 93.5 | 95.0 | 96.7 | 97.7 |

3.5 模型可靠性验证

为了考察 DACNN 模型对故障误判情况,对测试结果做混淆矩阵实验。图 12 所示为 WDCNN、DCNN 与 DACNN 3 种网络模型在 B-A 数据集加入 SNR=0 噪声情况下的混淆矩阵。

由图 12 可知,WDCNN 模型在滚动体轻度、中度、重度故障,以及内圈中度故障、外圈重度故障等 5 种情况的诊断精度均在 85% 以下,对滚动体轻度故障的诊断准确率仅达 44%,有 42% 被误诊断为内圈中度故障。DCNN 模型在滚动体轻度故障、内圈中度故障、外圈重度故障等 3 种故障情况的诊断精度小于 85%,其中滚动体轻度故障有 17% 被误诊为内圈中度故障,内圈中度故障有 18% 被误诊断为滚动体轻度故障。相比之下,本文提出的 DAC-



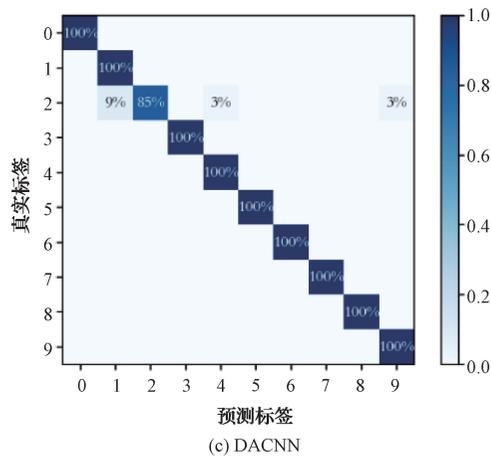


图 12 不同模型的混淆矩阵

Fig. 12 Confusion Matrix of different models

NN 模型仅在滚动体重度故障情况下出现了误诊断,误诊率均在 9%以下。可见,即使在噪声环境下 DACNN 模型依然比其他两种模型具有更高的可靠性。

3.6 模型可视化

1) 神经元可视化

不同故障类型在 DACNN 模型上的第 1 层神经元可视化如图 13 所示。图 13(a)~(d) 分别对应无故障、滚动体重度故障、内圈中度故障、外圈轻度故障状态。其中,红色为未激活,蓝色为最大激活,神经元的激活程度与图中颜色的亮度正相关。由图 13 可知,不同故障状态对应着不同的神经元激活情况,表明该种状态的重要特征已被网络提取出来。由 13(c) 可知,第 1 层网络中有 3 处明显发亮,表明这 3 处的神经元被激活,蕴含在信号中的内圈中度故障特征信息得到了有效提取。

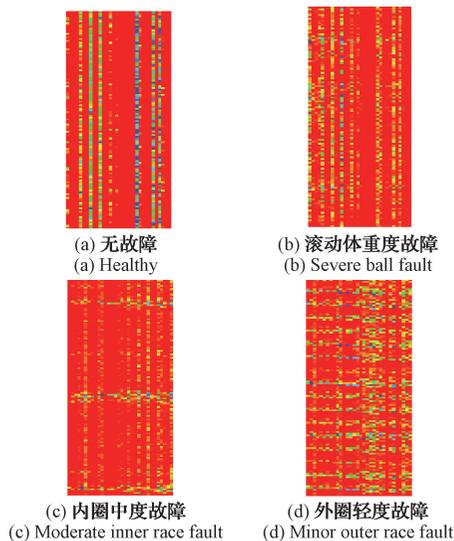


图 13 不同故障类型第 1 层神经元激活状态可视化
Fig. 13 Visualization of neuron activated at first level of DACNN for different fault

卷积激活层及全连接层的神经元激活如图 14 所示。由图 14 可知,原始故障信号经第 1 层激活层后,有 7 处明显发亮,表明信号中蕴含的故障特征信息在神经元中得以保留。后续的第 2~4 层继续对上一卷积层传递过来的重要特征进行筛选,未被激活的神经元被滤除,重要的特征更加突出。随着层数的增加,被激活的位置渐趋一致,在第 5 层卷积激活层中已经可以明确得出外圈轻度故障的重要特征。上述过程表明 DACNN 模型具有优秀的特征提取能力。

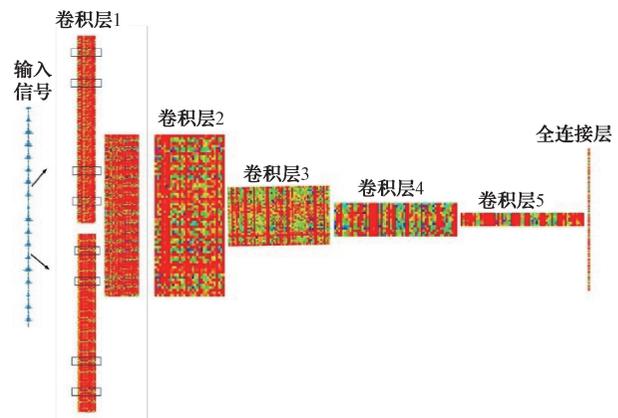
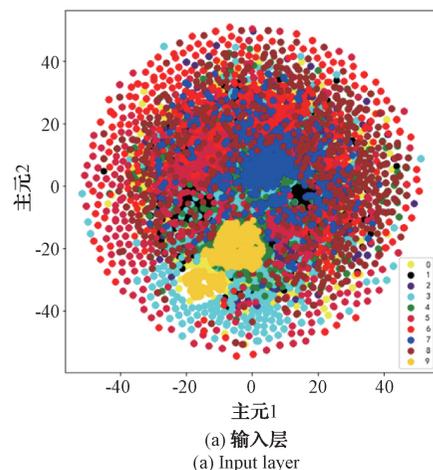


图 14 DACNN 模型神经元激活状态可视化

Fig. 14 Visualization of neuron activated at each level of DACNN

2) 分类过程可视化

以 B-ABCD 测试集特征提取过程为例呈现模型分类过程中不同卷积单元的特征聚类效果,聚类效果如图 15 所示。通过比较可以看出,随着卷积层数的增加,代表不同类型故障的特征得到了有效区分,且可分性越来越明显,表明 DACNN 模型能够在强噪声和变工况不同负载环境下准确完成轴承故障诊断任务,所提出的 DACNN 模型网络结构、学习机制设计和池化策略选择是有效的、合理的。在图 15(g) 中,标签 1 和 2 所代表的特征出现了部分重叠,表明模型的诊断性能还有改进和提升的空间。



(a) 输入层
(a) Input layer

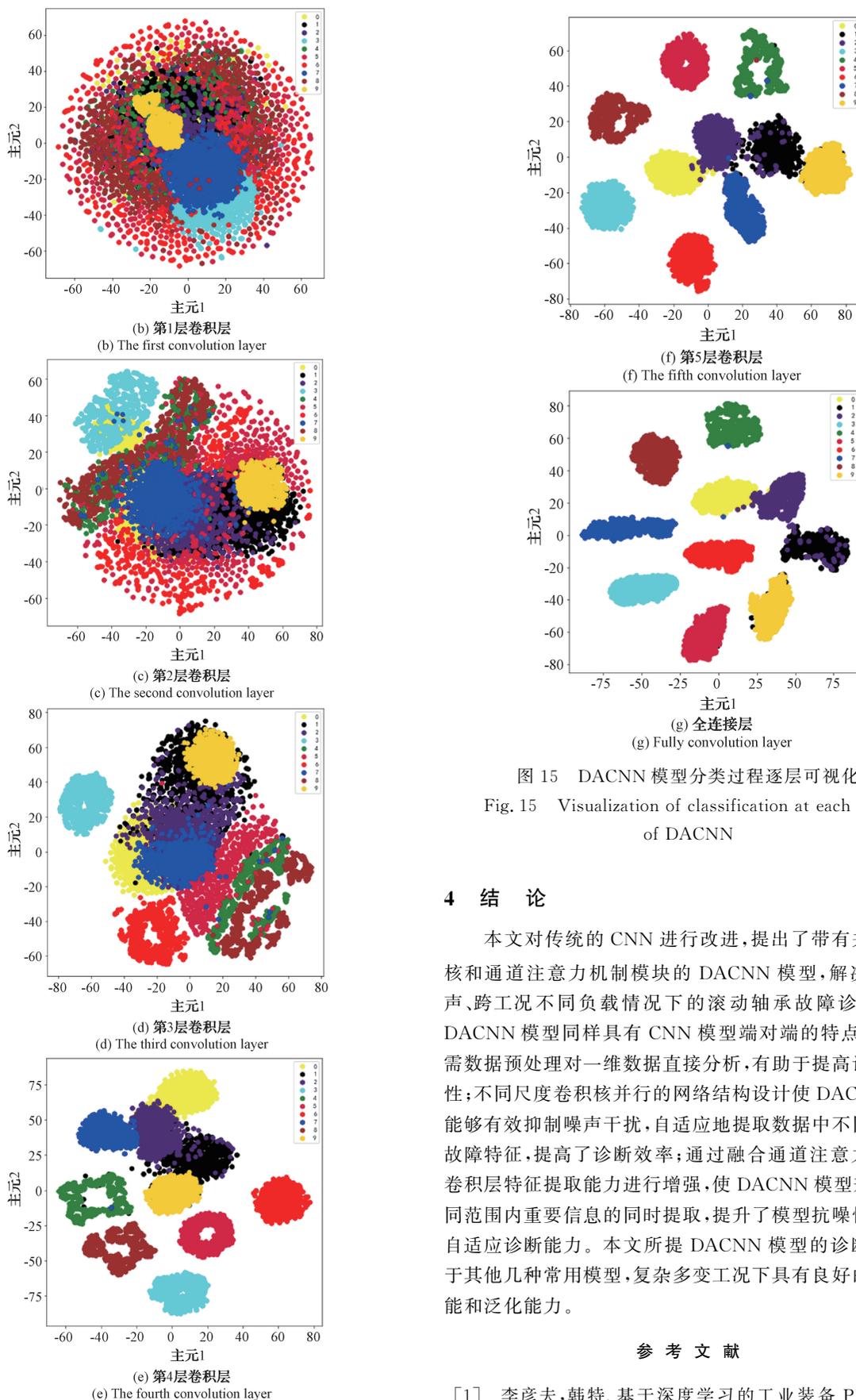


图 15 DACNN 模型分类过程逐层可视化

Fig. 15 Visualization of classification at each level of DACNN

4 结 论

本文对传统的 CNN 进行改进,提出了带有并行卷积核和通道注意力机制模块的 DACNN 模型,解决了强噪声、跨工况不同负载情况下的滚动轴承故障诊断问题。DACNN 模型同样具有 CNN 模型端对端的特点,可以无需数据预处理对一维数据直接分析,有助于提高诊断实时性;不同尺度卷积核并行的网络结构设计使 DACNN 模型能够有效抑制噪声干扰,自适应地提取数据中不同层级的故障特征,提高了诊断效率;通过融合通道注意力机制对卷积层特征提取能力进行增强,使 DACNN 模型兼顾了不同范围内重要信息的同时提取,提升了模型抗噪性能以及自适应诊断能力。本文所提 DACNN 模型的诊断性能优于其他几种常用模型,复杂多变工况下具有良好的抗噪性能和泛化能力。

参 考 文 献

[1] 李彦夫,韩特. 基于深度学习的工业装备 PHM 研究

- 综述[J]. 振动测试与诊断, 2022, 42(5): 835-847, 1029.
- LI Y F, HAN T. Deep learning based industrial equipment prognostics and health management: A review[J]. Journal of Vibration, Measurement & Diagnosis, 2022, 42(5): 835-847, 1029.
- [2] WANG X Y, XU H X, YUAN L M, et al. A lightweight and stochastic depth residual attention network for remote sensing scene classification [J]. IET Image Processing, 2023, 17(11): 3106-3126.
- [3] 陈是扞, 彭志科, 周鹏. 信号分解及其在机械故障诊断中的应用研究综述[J]. 机械工程学报, 2020, 56(17): 91-107.
- CHEN SH Q, PENG ZH K, ZHOU P. Review of signal decomposition theory and its applications in machine fault diagnosis [J]. Journal of Mechanical Engineering, 2020, 56(17): 91-107.
- [4] 刘方, 翟涛涛, 侯超强, 等. 基于多普勒调制时移 Laplace 小波的列车轴承故障声信号瞬态成分快速提取方法[J]. 仪器仪表学报, 2022, 43(3): 40-48.
- LIU F, ZHAI T T, HOU CH Q, et al. A fast transient component extraction method of train bearing fault acoustic signal based on Doppler modulated time-shifting [J]. Chinese Journal of Scientific Instrument, 2022, 43(3): 40-48.
- [5] 王骁贤, 陆思良, 何清波, 等. 变转速工况下基于多传感器信号深度特征融合的电机故障诊断研究[J]. 仪器仪表学报, 2022, 43(3): 59-67.
- WANG X X, LU S L, HE Q B, et al. Motor fault diagnosis based on deep feature fusion of multi-sensor data under variable speed condition [J]. Chinese Journal of Scientific Instrument, 2022, 43(3): 59-67.
- [6] HE M, HE D. Deep learning based approach for bearing fault diagnosis [J]. IEEE Transactions on Industry Applications. 2017, 53(3): 3057-3065.
- [7] JANSSENS O, SLAVKOVIKJ V, VERVISCH B, et al. Convolutional neural network based fault detection for rotating machinery [J]. Journal of Sound and Vibration, 2016, 377: 331-345.
- [8] 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.
- [9] LIANG H P, CAO J, ZHAO X Q. Multi-scale dynamic adaptive residual network for fault diagnosis[J]. Measurement, 2022, 188: 110397.
- [10] 杨春柳. 基于卷积神经网络的多层域自适应滚动轴承故障诊断方法[J]. 电子测量与仪器学报, 2021, 35(2): 122-128.
- YANG CH L. Multi-domain adaptive rolling bearing fault diagnosis based on convolutional neural network[J]. Journal of Electronic Measurement and Instrumentation, 2021, 35(2): 122-128.
- [11] 曲建岭, 余路, 袁涛, 等. 基于一维卷积神经网络的滚动轴承自适应故障诊断算法[J]. 仪器仪表学报, 2018, 39(7): 134-143.
- QU J L, YU L, YUAN T, et al. An adaptive fault diagnosis algorithm for rolling bearings based on one-dimensional convolutional neural network[J]. Chinese Journal of Scientific Instrument, 2018, 39(7): 134-143.
- [12] YE Z, YU J B. A novel convolutional neural network with adaptive kernel width and sparse regularization for machinery fault diagnosis [J]. Journal of Manufacturing Systems, 2021, 59: 467-480.
- [13] 全大英, 陈, 贇, 唐泽雨, 等. 基于双通道卷积神经网络的雷达信号识别[J]. 上海交通大学学报, 2022, 56(7): 877-885.
- QUAN D Y, CHEN Y, TANG Z Y, et al. Radar signal recognition based on dual channel convolutional neural network [J]. Journal of Shanghai Jiaotong University, 2022, 56(7): 877-885.
- [14] 孟良, 苏元浩, 许同乐, 等. 并行卷积神经网络的风电机组故障诊断方法[J]. 太阳能学报, 2023, 44(5): 449-486.
- MENG L, SU Y H, XU T L, et al. Wind turbine fault diagnosis method based on parallel convolutional neural network [J]. Acta Energetica Solaris Sinica, 2023, 44(5): 449-486.
- [15] 卢浩龙, 朱彦敏. 基于多尺度图 Transformer 的滚动轴承故障诊断[J]. 国外电子测量技术, 2023, 42(12): 186-194.
- LU H L, ZHU Y M. Fault diagnosis of rolling bearing based on multi-scale graphTransformer [J]. Foreign Electronic Measurement Technology, 2023, 42(12): 86-194.
- [16] 唐斯, 陈新楚, 郑松. 基于注意力与多尺度卷积神经网络的电机轴承故障诊断[J]. 电气技术, 2020, 21(11): 32-38.
- TANG S, CHEN X H, ZHENG S. Fault diagnosis method of motor bearing based on attention and multi-scale convolution neural network[J]. Electrical Engineering, 2020, 21(11): 32-38.
- [17] 赵小强, 张毓春. 基于双路并行多尺度 ResNet 的滚动轴承故障诊断方法[J]. 振动与冲击, 2023, 42(3): 199-208.
- ZHAO X Q, ZHANG Y CH. Fault diagnosis method of rolling bearing based on dual-path parallel multi-

- scale ResNet method[J]. Journal of Vibration and Shock, 2023, 42(3):199-208.
- [18] 张瑞博,李凌均. 基于注意力机制与 Inception-ResNet 的轴承故障诊断方法[J]. 电子测量技术, 2023, 46(21):107-113.
ZHANG R B, LI L J. Bearing fault diagnosis method based on attention mechanism and Inception-ResNet[J]. Electronic Measurement Technology, 2023, 46(21):107-113.
- [19] 康涛,段蓉凯,杨磊,等. 融合多注意力机制的卷积神经网络轴承故障诊断方法[J]. 西安交通大学学报, 2022, 56(12): 68-77.
KANG T, DUAN R K, YANG L, et al. Bearing fault diagnosis using convolutional neural network based on a multi-attention mechanism[J]. Journal of

Xi'an Jiaotong University, 2022, 56(12): 68-77.

作者简介

贾朱植, 博士, 副教授, 主要研究方向为鼠笼电机故障诊断方法。

E-mail: jiazhu zhi@163.com

刘凯, 硕士研究生, 主要研究方向为鼠笼电机轴承故障智能诊断方法。

刘佳鑫, 硕士研究生, 主要研究方向为鼠笼电机轴承故障智能诊断方法。

祝洪宇(通信作者), 博士, 教授, 主要研究方向为设备在线监测及故障诊断。

E-mail: zhuhongyu71@sina.cn