

滚动轴承域适应迁移学习故障识别方法

申 炜, 黄晋英, 范振芳, 王宇轩

(中北大学 机械工程学院, 山西 太原 030051)

摘要: 针对实际工程中轴承数据包含噪声, 且难以获取标签数据, 导致难以进行故障识别的问题, 提出了一种基于小波包一维卷积神经网络与复合域适应损失函数的迁移学习故障识别方法。所提出的小波包分支注意力卷积神经网络(Wavelet Packet Branch Attention Convolutional Neural Network, WPBA-CNN)综合小波包方法与注意力机制对数据特征进行提取, 并针对多尺度分支结构特点提出了分支最大均值差异(Branch Maximum Mean Discrepancy, BMMD)损失函数, 结合交叉熵损失函数与快速批量核范数最大化(Fast Batch Nuclear-norm Maximization, FBNM)方法, 构建了一种新颖的域适应复合损失函数(Domain Adaptation Compound Loss, DACL)进行迁移学习故障识别。结果表明, 在-4 dB噪声数据集实验中, WPBA-CNN-DACL的准确率较具有训练干扰的卷积神经网络(Convolution Neural Networks with Training Interference, TICNN)提升了16个百分点, 其BMMD组件的准确率较传统MMD提高了3.3个百分点, 20组迁移任务的平均准确率达98.24%。这些实验结果验证了本文方法在噪声抑制与跨域适应中的协同优势, 该方法可以作为无标签轴承故障诊断的有效解决方案。

关键词: 卷积神经网络; 迁移学习; 域适应; 复合损失函数

中图分类号: TH133.3 **文献标识码:** A **doi:** 10.62756/jnuc.issn.1673-3193.2025.04.0008

引用格式: 申炜, 黄晋英, 范振芳, 等. 滚动轴承域适应迁移学习故障识别方法[J]. 中北大学学报(自然科学版), 2025, 46(5): 632-640.

SHEN Wei, HUANG Jinying, FAN Zhenfang, et al. Domain-adaptive transfer learning method for rolling bearing fault identification[J]. Journal of North University of China(Natural Science Edition), 2025, 46(5): 632-640.

Domain-Adaptive Transfer Learning Method for Rolling Bearing Fault Identification

SHEN Wei, HUANG Jinying, FAN Zhenfang, WANG Yuxuan

(School of Mechanical Engineering, North University of China, Taiyuan 030051, China)

Abstract: To address challenges in practical engineering where bearing data contains noise and labeled data is scarce, this study proposed a transfer learning method for fault identification using a wavelet packet 1D-CNN and compound domain adaptation loss. The developed wavelet packet branch attention CNN (WPBA-CNN) integrated wavelet packet analysis and attention mechanisms for noise-resistant feature extraction. A branch maximum mean discrepancy (BMMD) loss was designed for multi-scale branches, and combined cross-entropy loss with fast batch nuclear-norm maximization (FBNM) method to form the domain adaptation compound loss (DACL). Experimental results demonstrate that the accuracy of the WPBA-CNN-DACL method increases by 16 percentage points compared to the TICNN method, and the accuracy of the BMMD component increases by 3.3 percentage points compared to the traditional MMD.

收稿日期: 2025-04-14

基金项目: 山西省基础研究计划资助项目(202203021211096); 中国博士后科学基金面上资助项目(2024M752992)

作者简介: 申炜(2000—), 男, 硕士生, 主要从事故障诊断与健康管理的研究。

通信作者: 黄晋英(1971—), 女, 教授, 博士, 主要从事结构动力学分析、检测与控制、故障诊断与健康管理的研究。E-mail: jyhuang@nuc.edu.cn。

The average accuracy rate of the 20 migration tasks reaches 98.24%. These experimental results validate the synergistic advantages of our method in noise suppression and cross domain adaptation, and this method can serve as an effective solution for unlabeled bearing fault diagnosis.

Key words: convolutional neural network; transfer learning; domain adaptation; composite loss function

0 引言

在当代工业体系中,确保机械设备故障诊断的精准性已成为保障安全生产不可或缺的一环,轴承是旋转机械的关键部件,其运行状态的健康与否直接关系到整个机械系统能否安全稳定运行。一旦滚动轴承发生故障,很可能对设备造成重大损坏,甚至造成人员伤亡^[1]。目前,通过振动信号进行故障识别是一种常用的故障诊断方法^[2],传统的轴承故障诊断大多是利用信号分析的方法^[3]。HUANG等^[4]提出的经验模态分解(Empirical Mode Decomposition, EMD)方法是信号分析方法发展的一个重要的里程碑。利用傅里叶变换等方法对频域信号进行分析的故障诊断方法效果也很好^[5]。随着计算机科学的发展,人工智能、模式识别、机器学习等智能方法也被用于了滚动轴承故障诊断^[6-7]。

卷积神经网络作为深度学习方法,具有强大的非线性提取特征的能力^[8]。Zhang等^[9]将一维原始信号输入到卷积神经网络中,通过首层宽卷积扩大模型感受视野,并结合防止过拟合技巧提出了具有训练干扰的卷积神经网络(Convolution Neural Networks with Training Interference, TICNN)。Zheng等^[10]提出的具有多尺度分支结构的结构优化深度卷积神经网络(Structure Optimized Deep Convolutional Neural Networks, SOCNN),在噪声环境下的轴承迁移学习任务中表现出了较好的鲁棒性与抗噪性。Woo等^[11]基于注意力机制和已有研究提出了卷积块的注意模型(Convolutional Block Attention Module, CBAM),通过同时聚焦通道与空间,增强了模型对关键信息识别的能力。

在实际工程中,任务目标往往没有足够的标签数据,需要迁移学习来解决这一问题。迁移学习是利用在一个任务上学习到的知识来帮助解决另一相关任务^[12],域适应是迁移学习的一个子类,是在源域和目标域任务相同但数据分布不同的情况下,通过调整模型使得在源域上训练的模型能够在目标域上取得良好性能的过程^[13]。最大均值化差异(Maximum Mean Discrepancy, MMD)作为衡量两组

数据分布相似度的度量,在域适应问题中发挥着重要作用。通过最小化MMD损失函数,可以引导模型提取源域和目标域的共有特征或映射到一个共享的特征空间,从而实现源域到目标域的迁移^[14]。为了解决域适应边界上的可判别性问题,Cui等^[15]提出了批量核范数最大化(Batch Nuclear-norm Maximization, BNM),在不增加额外参数的情况下,提高了源域的认识准确率。

综上所述,为了更好的面对轴承故障迁移学习任务,解决工业运行环境中监测目标域无标签的实际工程问题,避免进行不必要的人工信号处理,本研究提出了小波包分支注意力卷积神经网络(Wavelet Packet Branch Attention Convolutional Neural Network, WPBA-CNN)与域适应复合损失(Domain Adaptation Compound Loss, DAACL)函数。

1 特征提取模型

1.1 小波包多尺度卷积神经网络

本研究以一维卷积神经网络为基础,将小波包分解与首层宽卷积和多尺度卷积相结合,构建了小波包分支注意力卷积神经网络(WPBA-CNN),模型整体结构如图1所示。



图1 小波包多尺度卷积神经网络结构

Fig. 1 Multiscale convolutional neural networks structure of wavelet packet

整个模型分为三部分,前两部分为模型特征提取部分,分别为小波包宽卷积层与多尺度卷积层,核心为基于CBAM^[11]改进的分支卷积块注意力模块(Branch Convolutional Block Attention Module, BCBAM)。

小波包宽卷积层结合了小波包分解与首层宽核卷积方法,将输入数据按照频率成分进行初步的特征提取,并通过BCBAM进行特征融合;随后通过多尺度卷积层对特征进行细致的提取并经BCBAM融合后输出;最后的全连接分类层通过特征展平和两层全连接层将提取到的特征进行分

类,并通过Softmax函数输出为概率预测矩阵,这样就完成了诊断。

1.2 分支卷积块注意力机制

随着卷积神经网络的发展,模型优化不再局限于增加网络深度,也可以通过构建多尺度并行分支的横向扩展结构来提升性能。为此,基于CBAM^[11]

构建分支卷积块注意力机制(BCBAM),其核心是在传统通道注意力前加入分支注意力机制(Branch Attention Module, BAM)。BCBAM先对各分支特征进行加权融合,再执行CBAM注意力机制,该结构通过动态分配分支权重,实现更有效的特征融合,从而优化特征利用效率。

BCBAM的计算流程如图2所示。

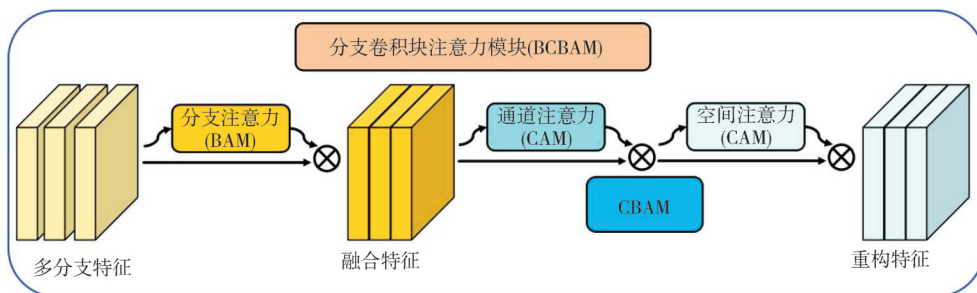


图2 分支卷积块注意力结构图

Fig. 2 Attention structure diagram of branch convolutional block

具体流程如下:

- 1) 特征打包: 将输入各分支的特征沿新增的分支维度堆叠,形成多分支特征集合;
- 2) 分支展平: 将每个分支的特征展开为“超级通道”,即合并一维特征与通道维度;
- 3) 分支加权: 计算各“超级通道”的CAM权重,反映分支重要性,并通过加权调整多尺度特征的贡献;
- 4) 特征融合与注意力学习: 恢复原始分支结构,移除分支维度后输入标准CBAM模块,进行通道与空间注意力学习。

1.3 小波包宽卷积层

为解决人工经验不足情况下的噪声振动信号识别问题,同时保留信号小波包分解方法处理信号噪声的优点,本文提出一种将小波包分解集成

在神经网络前端的小波包宽卷积层方法。

本文方法将原始信号输入模型,免去对信号进行预处理的步骤,在模型内部进行信号分解,并对原信号及其分解信号同步实施宽核卷积,通过注意力机制聚焦关键分量。这种方法不仅摆脱了传统方法中对于人工先验知识的依赖,而且通过注意力机制的自适应权重分配,能够更为精准地识别和保留那些与噪声混杂的关键特征,从而有效避免了人工选择故障信号分量的遗漏。具体步骤如下:

- 1) 在应用离散小波包进行分解时, Daubechies小波族中的db 4小波基对于旋转机械振动信号的识别效果突出^[16]。因此,本文选择db 4小波基对原始输入信号进行三重分解,以获取三层信号的小波分量。在每一层分解过程中,输入信号被分为两个长度减半的分量,从而降低计算冗余。

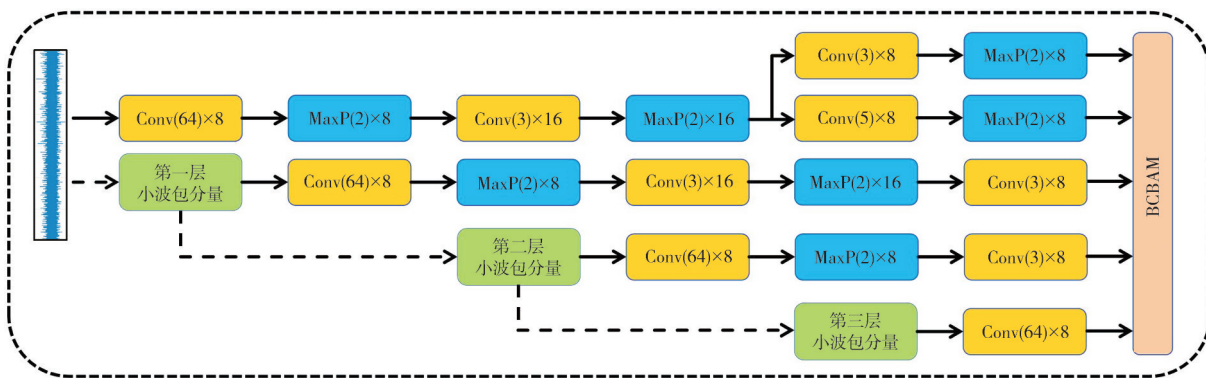


图3 小波包宽卷积层

Fig. 3 Wide convolutional layer of wavelet packet

- 2) 对原始信号以及每一层分解得到的信号分量,使用长度为64的宽卷积核进行卷积操作。模

型在第一层卷积层使用较宽的卷积核,旨在提高模型对一维振动信号中空间特征信息的提取能

力,尤其是在浅层网络中,能够捕捉到更全面的信号特征。随后,对所有信号传递分支进行正常的卷积和池化操作,并对原始信号进行双重尺度的卷积,在最大化保留故障特征的同时,确保所有信号分量的输出尺寸保持一致。

3) 通过注意力机制赋予各信号分支权重并整合关键信息,使关键的信号分量能在诊断中对结果有更大的影响力。小波包宽卷积层的网络结构如图 3 所示,方块中的参数为操作名称(卷积核尺寸)×输出通道数。

1.4 多尺度卷积层

多尺度卷积神经网络是指使用多个不同大小的卷积核进行特征提取。这些卷积核具有不同的感受野,能够在不同尺度上捕获图像特征^[16]。

在该结构中,较短的卷积核专注于提取序列中的短时依赖性和高频特征,如快速变化的趋势或短暂的信号模式;较长的卷积核则能够捕捉到更长时间的依赖性和低频特征,如周期性变化或长期趋势。多尺度卷积方法通过组合使用这些不同长度的卷积核,使卷积模型能够在多个时间尺度上生成丰富的特征表示,使其能够更好地适应各种复杂和多样化的应用场景。这些卷积核能够覆盖从局部细节到全局趋势的不同范围,从而增强模型对信号特征的多尺度理解能力^[17]。模型的具体架构如图 4 所示。

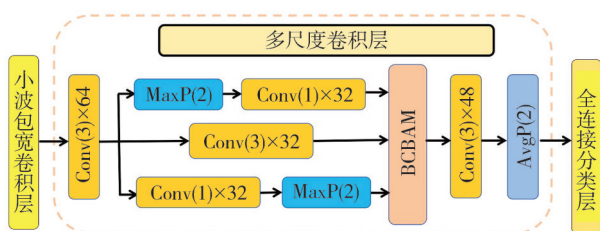


图 4 多尺度卷积层

Fig. 4 Multi-scale convolutional layer

2 域适应迁移学习损失函数

2.1 交叉熵损失函数

交叉熵损失(Cross-Entropy Loss, CEL)是机器学习领域,特别是在分类问题中广泛使用的一种损失函数,用于衡量模型输出与真实标签之间的差异,并试图通过最小化这个差异来优化模型。

在多分类问题中,模型输出的通常是每个类别的预测概率,这些概率通过 Softmax 函数得到,确保所有类别的预测概率之和为 1。真实标签 y_i

是一个独热编码的向量,其中只有一个元素为 1,表示该样本属于的类别,其余元素为 0。交叉熵损失函数可以表示为

$$Loss_{CEL} = - \sum_{i=1}^C y_i \log(\hat{y}_i), \quad (1)$$

式中: \hat{y}_i 为概率预测矩阵; C 为样本种类数量。

本研究针对源域有标签的数据,采用交叉熵损失作为源域训练的损失函数。

2.2 分支最大均值差异

在迁移学习中,域适应方法是解决目标域无标签问题,提高模型在目标域中的泛化能力,从而实现迁移学习的有效方法。最大均值化差异(MMD)损失函数作为经典域适应方法,在机器学习和迁移学习中常用于度量两个不同但相关的分布之间的距离。MMD 方法通过比较两个分布的特征映射后的均值差异来评估它们之间的差异,MMD 损失函数的公式可以表示为

$$Loss_{MMD} = \left\| \frac{1}{n} \sum_{i=1}^n \phi(x_i) - \frac{1}{m} \sum_{j=1}^m \phi(y_j) \right\|_2^2 = \frac{1}{n^2} \sum_{i,j=1}^n K(x_i, x_j) + \frac{1}{m^2} \sum_{i,j=1}^m K(y_i, y_j) - \frac{2}{nm} \sum_{i=1}^n \sum_{j=1}^m K(x_i, y_j), \quad (2)$$

式中: x_i 与 y_j 为两个不同分布的样本中的元素; $\phi(\cdot)$ 为映射函数; n 与 m 为集合的大小; $K(x, y)$ 为高斯核函数,即 RBF 核函数。

$$K(x, y) = \exp\left(-\frac{|x - y|^2}{2\sigma^2}\right), \quad (3)$$

式中: σ 为带宽参数,用来控制核函数的平滑程度。

为了更好地适应所提出的 WPBA-CNN 模型中的分支结构,提出了分支最大均值差异(Branch Maximum Mean Discrepancy, BMMD)方法,在 BMMD 框架下,在每个尺度上独立计算 MMD 损失,以精确评估源域与目标域在该尺度特征空间上的分布差异。

具体来说,对于每个尺度,首先通过相应的特征提取分支获取特征表示,然后计算源域和目标域特征之间的 MMD 距离,不同尺度的 MMD 损失不仅反映了各自尺度下分布差异的精细程度,还相互补充,共同刻画了整体数据的复杂性和多样性。

为了综合不同尺度的 MMD 损失,采用加权求和的方式得到 BMMD 损失函数的总损失,以平衡不

同尺度特征的重要程度。BMMD损失函数公式可表示为多个MMD损失函数通过权重分配后相加的形式。

$$Loss_{BMMD} = \sum_{i=1}^B \frac{1}{B} Loss_{MMD_i}, \quad (4)$$

式中： B 为所求不同尺度的MMD损失的尺度数量； $Loss_{MMD}$ 为单个分支的MMD损失。

域适应损失函数通常与分类损失函数结合使用，形成网络总体损失。本研究将CEL和BMMD损失结合构建了复合损失函数进行域适应迁移学习，以实现轴承目标域无监督故障识别。

2.3 批量核范数最大化

在域适应迁移学习任务中，由于目标域数据缺乏标签信息，会在决策边界附近产生大量混淆数据。为了应对这一挑战，Cui等^[15]提出了批量核范数最大化(Batch Nuclear-norm Maximization, BNM)，在不增加网络模型额外参数的前提下，从可判别性和多样性两个维度对迁移学习过程中的域适应进行了优化。

通过核范数最大化可以提高模型的可判别性和多样性，可以用奇异值 σ_i 计算 F 范数与秩，因此需要计算矩阵 \hat{y} 中的所有奇异值，为了快速、有效地计算，采用矩阵 \hat{y} 中各类别响应的组合来近似估计，最大奇异值 σ_j 可以近似为

$$\sigma_j \approx \text{top} \left(\sum_{i=1}^m \hat{y}_{j,i}^2, i \right) \quad \forall j \in 0, \dots, \min(m, n)。 \quad (5)$$

通过对奇异值的近似，可以在不影响性能的情况下，以更快的方式计算批核范数，即使用快速批量核范数最大化(Fast Batch Nuclear-norm Maximization, FBNM)，公式为

$$\|\hat{y}\|_* = \sum_{j=1}^{\min(m, n)} \sqrt{\sigma_j}。 \quad (6)$$

由上可知，FBNM值越大，预测矩阵的可判别性与多样性越好，又由奇异值不为负数，可定义FBNM损失函数。

$$Loss_{FBNM} = -\|\hat{y}\|_*。 \quad (7)$$

2.4 域适应复合损失函数

在基于域适应的迁移学习中，复合损失函数是一种强大的工具。在域适应迁移学习的框架下，为了最大程度地促进源域知识向目标域有效迁移，有效整合不同损失函数的优势并克服单一损失函数的局限性，本研究通过综合源域的交叉

熵损失函数(CEL)、源域和目标域之间的分支最大均值差异损失函数(BMMD)，以及提高目标域预测可判别性和多样性的核范数最大化损失函数(FBNM)，提出了一种域适应复合损失函数(DACL)，其值随epoch的变化而改变，公式为

$$L_{DAFL} = \begin{cases} L_{CEL} + 0.5 \times L_{BMMD}, & epoch < 100, \\ L_{CEL} + L_{BMMD}, & 100 \leq epoch < 150, \\ L_{CEL} + L_{BMMD} + L_{FBNM}, & epoch \geq 150. \end{cases} \quad (8)$$

这种复合损失函数结合了多个不同的损失函数，通过改变损失函数权重，最大程度地优化源域的分类性能、源域与目标域之间的分布对齐，以及目标域的控制边界稳定性，可以为实现高效的知識迁移提供有力支持。该函数不仅丰富了域适应迁移学习的理论体系，也为实际应用中的域适应问题提供了新的解决思路。

3 迁移学习实验验证

3.1 训练方法

WPBA-CNN-DAFL是一种创新的域适应迁移学习方法，通过将小波包分支注意力卷积神经网络(WPBA-CNN)与域适应复合损失函数(DACL)有机结合，显著提升了跨域知识迁移的性能。该方法的核心创新体现在两个方面：

在特征提取方面，WPBA-CNN采用小波包分解技术对输入数据进行多尺度分析，并构建并行卷积分支结构。每个分支配备注意力机制，能够动态调整对不同特征表示的关注权重，从而更有效地提取对分类任务最具判别性的特征。这种设计不仅增强了模型处理复杂数据的能力，还显著提高了系统的抗噪性和鲁棒性。

在域适应优化方面，DAFL损失函数整合了3种关键损失，分阶段地对目标进行优化：交叉熵损失(CEL)确保源域分类性能，分支最大均值差异损失(BMMD)最小化域间分布差异，快速批量核范数最大化损失(FBNM)增强目标域特征判别性。这种复合的分阶段优化策略实现了源域知识向目标域的有效迁移。

实验采用Python 3.9环境和PyTorch框架实现，设置batch_size为50，学习率lr为0.001，使用Adam优化器进行250个epoch的训练。

3.2 实验数据集

本研究实验数据选用德国帕德伯恩大学(Paderborn University, PU)的滚动轴承数据集^[18], 选用了数据集中轴承的振动信号, 轴承实验装置如图 5 所示。

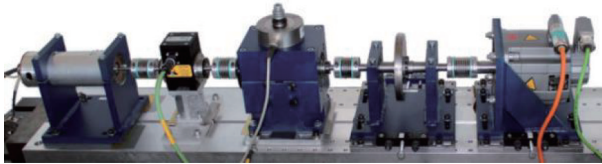


图 5 轴承数据实验台

Fig. 5 Bearing data test bench

为了进行不同工况下的轴承故障诊断研究, 本研究采用了帕德伯恩大学轴承实验装置在 3 种不同工况下的参数设置, 具体实验环境参数如表 1 所示。

表 1 实验数据采集工况

Tab. 1 Experimental data collection conditions

编号	转速/(r·min ⁻¹)	扭矩/(N·m)	径向力/N
A	1 500	0.7	1 000
B	1 500	0.7	400
C	1 500	0.1	1 000

在这 3 种工况下采集的轴承原始振动信号构成了本研究的实验数据集。根据故障的严重等级和形成方式的不同, 选取了 1 种正常类型、5 种外圈故障类型和 3 种内圈故障类型, 共计 9 种故障类型, 每种类型各具特点。每个实验环境下的轴承故障数据集数据种类的详细信息如表 2 所示。

表 2 实验数据种类

Tab. 2 Experimental data types

编号	故障等级	故障位置	故障形成方式
0	健康	—	—
1	1 级	外圈	电火花(裂纹)
2	1 级	外圈	电雕刻(点蚀)
3	2 级	外圈	电雕刻(点蚀)
4	1 级	外圈	钻孔(剥落)
5	2 级	外圈	钻孔(剥落)
6	1 级	内圈	电火花(裂纹)
7	1 级	内圈	电雕刻(点蚀)
8	2 级	内圈	电雕刻(点蚀)

上述每种工况的每种实验数据均划分出 500 条数据, 每条数据包含 2 048 个采样数据点, 数据集按照训练集与测试集的比例为 9:1 进行随机乱序划分, 即每种工况包含 4 050 条训练数据, 450 条测试数据。

为了检验模型的抗噪性能, 在原始振动信号中分别加入 -4, -2, 0, 2, 4 和 6 dB 的高斯白噪声, 在实际工业运行环境中, 噪声是由多种来源、不同概

率分布的独立随机变量叠加形成的复杂复合体。基于中心极限定理, 当噪声源数量增多时, 这些变量的归一化和会逼近高斯分布。因此, 采用高斯白噪声对其综合噪声进行模拟具有合理性和有效性^[19]。

3.3 对比试验设置

为了验证所提出的 WPBA-CNN-DACL 方法模型在噪声环境下的鲁棒性和优越性, 本研究设置了两种对比试验。

1) 基准模型对比实验。

WPBA-CNN(本研究提出的新型分支注意力网络): 其创新的小波分支注意力结构赋予模型卓越的抗噪能力。

TICNN(传统干扰训练 CNN)^[9]: 代表传统卷积神经网络在任务中的基准表现。

SOCNN(最新结构优化深度网络)^[10]: 作为前沿研究成果, 对模型进行了结构优化。

2) 损失函数对比实验。

WPBA-CNN: 采用交叉熵损失训练;

WPBA-CNN-MMD^[14]: 基于最大均值差异的方法, 进行传统的域适应迁移学习。

WPBA-CNN-BMMD: 基于所提出的分支结构, 使用改进后的 BMMD 损失函数进行域适应迁移学习。

WPBA-CNN-DACL: 本研究提出的复合损失函数, 采用了分段训练思想。

上述对比实验结果不仅可以验证 WPBA-CNN 模型在噪声环境下的鲁棒性和优越性, 还揭示了 DACL 域适应策略对其性能提升的重要作用。

3.4 结果分析

根据上述对比实验设置, 对所提出的模型与方法进行训练测试, 并对迁移学习目标域的准确率进行统计, 每组实验进行 5 次并取平均值, 具体情况如表 3 所示。

在噪声环境下测试 3 种基模型, 通过准确率的对比分析可知, WPBA-CNN 在多个迁移场景中展现出显著优势。以噪声最严重的 -4 dB 环境为例: 在 C→B 迁移任务中, WPBA-CNN 目标域的准确率达到 85.33%, 较 TICNN 方法提升了 7.78 百分点, 较 SOCNN 方法提升了 5.11 百分点; 在 B→A 迁移中, WPBA-CNN 方法的准确率(83.52%)较 TICNN 提升了 5.84 百分点, 较 SOCNN 提升了 4.98 百分点。这主要得益于 WPBA-CNN 的小波包分支注意力结构能够有效

抑制噪声干扰,提升模型的抗噪性,尽管SOCNN 波包宽卷积层在高噪声环境下性能突出,验证了采用多尺度卷积后的性能优于TICNN方法,但小 WPBA-CNN的优越性。

表3 各噪声下不同方法在目标域的准确率

Tab.3 Accuracy of different methods in target domain under each noise

迁移方向	噪声/dB	目标域的准确率/%					
		TICNN	SOCNN	WPBA-CNN	WPBA-CNN-MMD	WPBA-CNN-BMMD	WPBA-CNN-DACL
A→B	4	90.12	93.42	94.67	99.05	99.23	99.57
A→B	2	85.76	88.17	92.14	95.63	97.34	98.07
A→B	0	82.25	83.12	86.04	90.42	95.78	96.87
A→B	-2	81.22	82.45	85.33	89.36	95.28	97.56
A→B	-4	79.85	82.56	86.53	91.01	93.12	96.58
A→C	4	99.12	98.89	99.78	99.94	100.00	100.00
A→C	2	98.16	97.25	99.85	99.72	100.00	100.00
A→C	0	98.91	98.04	99.64	99.72	99.81	99.78
A→C	-2	97.15	97.42	98.75	99.21	99.14	99.33
A→C	-4	94.55	96.28	98.53	97.12	98.96	99.12
B→A	4	91.23	92.55	94.25	98.56	99.78	99.89
B→A	2	83.24	87.22	91.22	93.54	96.12	97.75
B→A	0	80.56	81.25	85.11	89.63	93.22	97.89
B→A	-2	81.22	82.17	86.14	88.67	94.17	95.18
B→A	-4	77.68	78.54	83.52	87.56	91.24	95.32
B→C	4	84.35	85.89	91.12	94.94	98.24	99.89
B→C	2	80.16	81.56	88.14	92.54	99.03	99.78
B→C	0	79.54	82.01	87.52	89.78	96.89	97.78
B→C	-2	80.25	80.51	85.16	90.15	95.12	96.00
B→C	-4	74.28	77.21	82.42	87.56	91.22	94.44
C→A	4	99.25	99.17	99.78	99.89	100.00	100.00
C→A	2	98.96	98.78	99.81	99.75	100.00	100.00
C→A	0	99.15	99.25	99.78	99.62	99.85	100.00
C→A	-2	96.25	96.31	98.56	99.34	99.52	99.82
C→A	-4	93.28	95.21	97.56	98.05	99.12	99.32
C→B	4	85.26	87.25	90.44	95.24	97.67	98.54
C→B	2	79.96	82.76	87.58	93.55	96.25	97.89
C→B	0	80.56	83.05	88.22	94.22	98.75	98.89
C→B	-2	81.26	81.05	87.24	92.16	94.24	95.12
C→B	-4	77.55	80.22	85.33	89.68	91.56	93.56

针对域适应方法的有效性分析,在C→B迁移方向-4 dB极端噪声下,BMMD(91.56%)较MMD(89.68%)提升了1.88个百分点,证明基于分支结构的BMMD损失函数能更有效地对齐跨域特征分布。在A→B迁移方向-4 dB噪声条件下,WPBA-CNN-DACL的准确率为96.58%,较仅用交叉熵损失的基线模型提升了10.05个百分点,较传统MMD方法提升了5.57个百分点,充分体现了分阶段优化策略的优势。特别是在极端噪声条件下(如C→B迁移-4 dB),DACL(93.56%)仍保持良好的稳定性能,较MMD方法提升了3.88个百分点,较BMMD(91.56%)提升了2.00个百分点。

为了更直观地对比域适应结果,绘制如图6所示的各模型卷积特征层后的t-SNE特征降维

图,通过降维特征之间的距离来分析模型在目标域对数据的提取结果。

由图6的特征降维图可视化分析发现,WPBA-CNN在3种基模型中的分类效果最优,但不使用域适应训练方法的模型,特征降维图比较分散,经过MMD方法后特征更加集中,多个特征已完全分开,BMMD方法利用多分支使其特征点距离凝聚,但边界稍有混淆,而所提出的DACL方法通过分阶段应用BMMD和FBNM方法,在特征降维图中的数据边界最为清晰,故障识别效果最优。WPBA-CNN-DACL作为复合损失函数,通过分阶段进行优化,不仅减少了域间差异,还优化了域边界表示,在所测试的噪声情况下的效果最佳,展示了其在噪声和迁移学习环境下的鲁棒性和优越性。

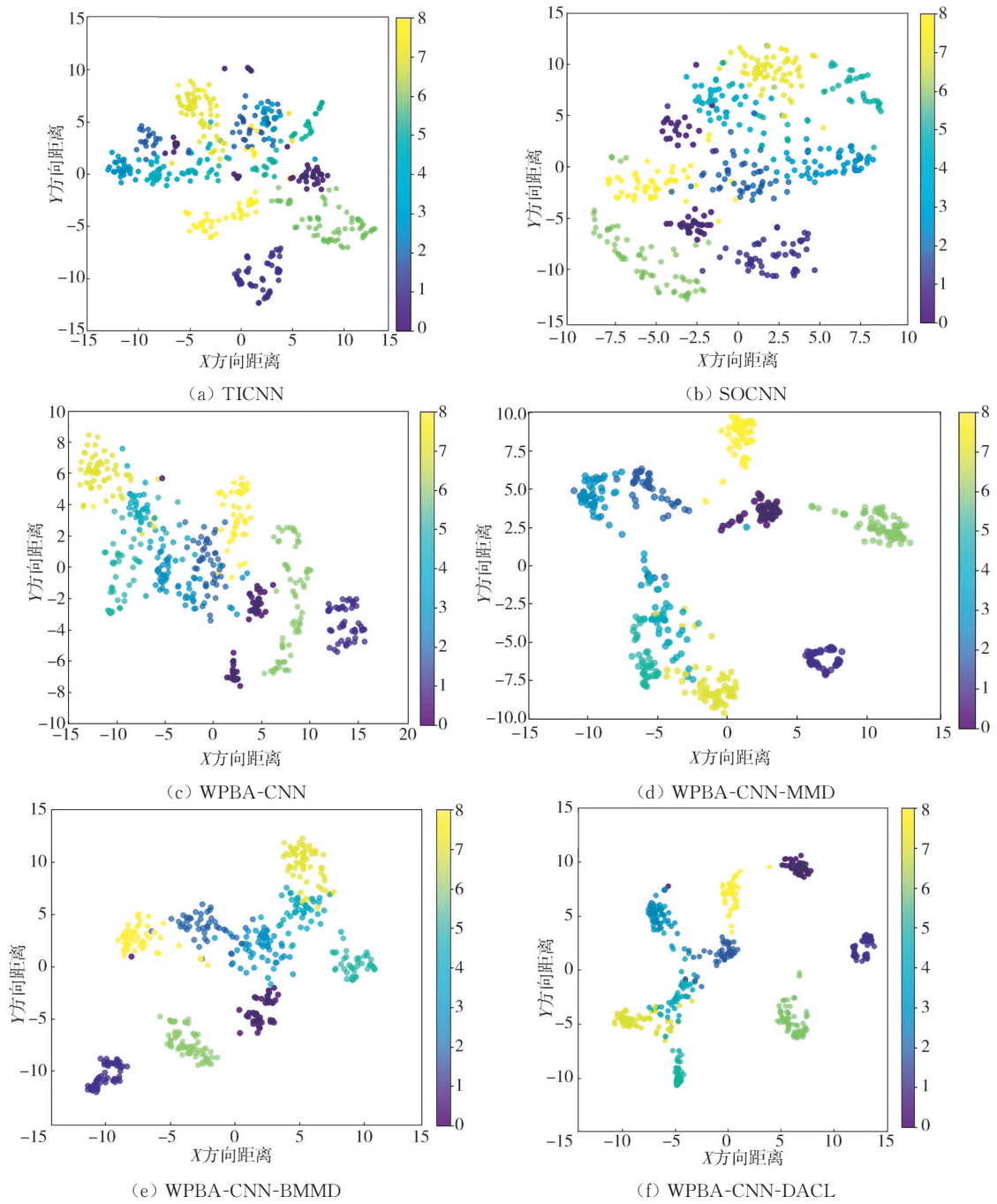


图 6 各方法在-4 dB噪声下C→B目标域的特征降维图

Fig. 6 Characteristic downscaling map of C→B target domain in -4 dB noise for different methods

4 结束语

本文提出了一种基于小波包的一维卷积神经网络的分支注意力卷积神经网络(WPBA-CNN),通过其分支结构构建了分支最大均值差异(BMMD)方法,并基于源域 CEL 损失函数和 FBNM 方法共同构建了域适应复合损失函数 DACL。该方法输入为一维振动信号,无需对信号进行预处理,且目标域数据无标签,更加符合

实际工程中数据缺少标签的真实情况。

在基于PU轴承数据集的迁移任务对比试验中,具体数据分析表明:在-4 dB 高斯白噪声条件下,所提出的 WPBA-CNN 模型在 C→B 迁移任务中目标域准确率达 85.33%,较传统 TICNN(77.55%)和前沿 SOCNN(80.22%)分别提升 7.78 个百分点和 5.11 个百分点,验证了小波包分支注意力结构的抗噪优势。同时,WPBA-CNN-DACL 方法在相同噪声场景下的准确率提升至 93.56%,较单一 MMD 方法

(89.68%)和BMMD方法(91.56%)分别提升了3.88个百分点和2.00个百分点,体现了复合损失函数的协同作用。此外,BMMD方法在B→C迁移方向4 dB噪声下的准确率达98.24%,较传统MMD(94.94%)提升了3.3个百分点,证明分支结构能更精准地对齐跨域特征分布。通过t-SNE可视化分析发现,DACL方法边界最为清晰,特征最为凝聚,在各方法中的效果最优。

参考文献:

- [1] LI J, LUO W, BAI M. Review of research on signal decomposition and fault diagnosis of rolling bearing based on vibration signal [J]. *Measurement Science and Technology*, 2024, 35 (9): 092001.
- [2] WU G G, YAN T Y, YANG G L, et al. A review on rolling bearing fault signal detection methods based on different sensors[J]. *Sensors*, 2022, 22(21): 8330-8330.
- [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 Shiqian, PENG Zhike, ZHOU Peng. Review of signal decomposition theory and its applications in machine fault diagnosis [J]. *Journal of Mechanical Engineering*, 2020, 56(17): 91-107. (in Chinese)
- [4] HUANG N E, SHEN Z, LONG S R, et al. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis [J]. *Proceedings of the Royal Society of London Series A: Mathematical, Physical and Engineering Sciences*, 1998, 454(1971): 903-995.
- [5] ALTHUBAITI A, ELASHA F, TEIXEIRA J A. Fault diagnosis and health management of bearings in rotating equipment based on vibration analysis-a review [J]. *Journal of Vibroengineering*, 2022, 24 (1): 46-74.
- [6] LEI Y G, JIA F, KONG D T, et al. Opportunities and challenges of machinery intelligent fault diagnosis in big data era[J]. *Journal of Mechanical Engineering*, 2018, 54(5): 94-104.
- [7] 司伟伟, 岑健, 伍银波, 等. 小样本轴承故障诊断研究综述[J]. *计算机工程与应用*, 2023, 59(6): 45-56.
SI Weiwei, CEN Jian, WU Yinbo, et al. Review of research on bearing fault diagnosis with small samples [J]. *Computer Engineering and Applications*, 2023, 59(6): 45-56. (in Chinese)
- [8] LI X, MA Z, YUAN Z, et al. A review on convolutional neural network in rolling bearing fault diagnosis [J]. *Measurement Science and Technology*, 2024, 35 (7): 072002.
- [9] ZHANG W, LI C H, PENG G L, et al. A deep convolutional neural network with new training methods for bearing fault diagnosis under noisy environment and different working load [J]. *Mechanical Systems and Signal Processing*, 2018, 100: 439-453.
- [10] ZHENG B, HUANG J H, MA X, et al. An unsupervised transfer learning method based on SOCNN and FBNN and its application on bearing fault diagnosis[J]. *Mechanical Systems and Signal Processing*, 2024, 208: 111047.
- [11] WOO S, PARK J, LEE J Y, et al. CBAM: Convolutional block attention module [C]//*Computer Vision-ECCV 2018*. Cham: Springer, 2018: 3-19.
- [12] MISBAHI, LEE C K M, KEUNG K L. Fault diagnosis in rotating machines based on transfer learning: Literature review [J]. *Knowledge-Based Systems*, 2024, 283: 111158.
- [13] ZHANG S Y, SU L, GU J F, et al. Rotating machinery fault detection and diagnosis based on deep domain adaptation: A survey [J]. *Chinese Journal of Aeronautics*, 2023, 36 (1): 45-74.
- [14] LI X Y, YUAN P, SU K Y, et al. Innovative integration of multi-scale residual networks and MK-MMD for enhanced feature representation in fault diagnosis [J]. *Measurement Science and Technology*, 2024, 35 (8): 086108.
- [15] CUI S H, WANG S H, ZHUO J B, et al. Towards discriminability and diversity: batch nuclear-norm maximization under label insufficient situations [C]//*2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2020: 3940-3949.
- [16] 郑旭. 基于离散小波变换的特征提取和故障分类方法研究[D]. 北京: 北京化工大学, 2017.
- [17] LI J M, MAO W L, YANG B X, et al. RUL prediction of rolling bearings across working conditions based on multi-scale convolutional parallel memory domain adaptation network [J]. *Reliability Engineering and System Safety*, 2024, 243: 109854.
- [18] LESSMEIER C, KIMOTHO J K, ZIMMER D, et al. Condition monitoring of bearing damage in electromechanical drive systems by using motor current signals of electric motors: A benchmark data set for data-driven classification [J]. *PHM Society European Conference*, 2016: 3(1): 1577.
- [19] 高越. 强噪声条件下基于对抗学习与域迁移的齿轮箱故障诊断方法[D]. 北京: 北京化工大学, 2024.