

# 基于加权对抗域自适应的旋转机械开放集跨域故障诊断

陆碧良, 张英杰<sup>†</sup>, 孙庆帅, 李蒲德

(湖南大学 信息科学与工程学院, 长沙 410082)

**摘要:** 传统深度学习故障诊断假设训练集与测试集独立同分布且标签空间一致, 然而, 在实际工业中, 常出现新故障类型和数据“分布漂移”, 即开放集跨域识别, 导致模型泛化能力差、诊断精度下降. 针对开放集跨域故障诊断问题, 提出一种基于深度一维卷积神经网络 (1D-CNN) 的改进对抗域自适应的故障诊断方法. 首先, 利用 1D-CNN 从多领域的输入样本中提取代表性特征; 然后, 使用一个特征细粒度分类器来区分多领域中的共享和离群特征; 接着, 采用加权模块指导域判别器实现域不变特征学习; 最后, 域判别器利用域不变特征指导特征分类器的分类, 最小化全局损失, 实现开放集跨域故障诊断. 实验结果表明, 该模型在美国凯斯西储大学 (CWRU) 数据集和齿轮箱数据集上的跨域故障诊断精度高于其他方法, 特别是在开放集数据中优势明显.

**关键词:** 旋转机械; 故障诊断; 深度学习; 开放集跨域诊断; 对抗域自适应

中图分类号: TP273 文献标志码: A

DOI: 10.13195/j.kzyjc.2024.1038

引用格式: 陆碧良, 张英杰, 孙庆帅, 等. 基于加权对抗域自适应的旋转机械开放集跨域故障诊断 [J]. 控制与决策, 2025, 40(10): 3136-3144.

## Open-set cross-domain fault diagnosis of rotating machinery based on weighted adversarial domain adaptation

LU Bi-liang, ZHANG Ying-jie<sup>†</sup>, SUN Qing-shuai, LI Pu-de

(College of Information Science and Engineering, Hunan University, Changsha 410082, China)

**Abstract:** Traditional deep learning-based fault diagnosis assumes that the training set and test set are independently and identically distributed with consistent label spaces. However, in actual industrial scenarios, new fault types and "distribution shifts" in data frequently occur, necessitating open-set cross-domain recognition, which results in poor model generalization and reduced diagnostic accuracy. To address the issue of open-set cross-domain fault diagnosis, we propose an weighted adversarial domain adaptation method based on a deep one-dimensional convolutional neural network (1D-CNN). Firstly, the 1D-CNN is employed to extract representative features from input samples across multiple domains. Then, a fine-grained feature classifier is utilized to distinguish between shared and outlier features across these domains. Subsequently, a weighting module guides the domain discriminator to achieve domain-invariant feature learning. Finally, the domain discriminator uses these domain-invariant features to guide the feature classifier in minimizing global loss, thereby achieving open-set cross-domain fault diagnosis. Experimental results demonstrate that the proposed model achieves higher cross-domain fault diagnosis accuracy compared to other methods on the Case Western Reserve University (CWRU) dataset and the gearbox dataset, particularly showing significant advantages in open-set data scenarios.

**Keywords:** rotating machinery; fault diagnosis; deep learning; open-set cross-domain diagnosis; adversarial domain adaptation

## 0 引言

随着工业 4.0 的推进, 工业系统逐渐向数字化、信息化、自动化和智能化方向发展, 这对设备的高可靠性提出了更高的要求<sup>[1]</sup>. 旋转机械在工业中至关重

要, 通过在动力传输与转换、制造与加工、流体传输与控制等方面的应用, 确保了工业系统的高效运行. 然而, 一旦轴承、齿轮箱<sup>[2]</sup>等关键部件在运行中出现故障, 将会导致严重的安全事故和经济损失. 因此,

收稿日期: 2024-08-30; 录用日期: 2025-02-25.

基金项目: 海上风电机组智能监测与故障诊断技术研究项目 (2024A1515240036); 面向智能运维的风力机关键部件全域感知、信息集成与协同诊断技术项目 (2019YFE0105300).

责任编辑: 叶丹.

<sup>†</sup>通信作者. E-mail: zhangyj@hnu.edu.cn.

旋转机械的智能故障诊断技术一直是学术界研究的前沿和热点。

近年来,深度学习(DL)<sup>[3]</sup>算法由于强大的特征提取能力以及适应复杂、非线性系统,被广泛应用于智能故障诊断领域<sup>[4]</sup>。文献[5]使用经验模态分解提取故障特征,并提出了一种基于改进的双径向基神经网络故障诊断方法,获得了良好的故障诊断性能;文献[6]结合贝叶斯和随机变分推理,提出了一种基于随机变分推理贝叶斯神经网络故障诊断方法,在小样本故障诊断场景中表现良好;文献[7]提出了一种基于改进的胶囊网络的智能故障诊断方法,可实现精确的复合故障诊断。但是,这些诊断方法均存在一定局限性,DL假设训练数据和测试数据均来自同一分布,且样本间相互独立,同时,DL模型通常需要大量的标签数据进行训练。然而,在实际工业中,旋转机械面临时变的工作环境和复杂的耦合作用,致使训练集与测试集间出现“分布漂移”,且故障标签样本的获取困难和昂贵,因此,基于DL的智能故障诊断方法的基本假设难以满足,这使得模型泛化能力差、故障诊断性能下降。

迁移学习(TL)<sup>[8]</sup>为解决训练集与测试集间的“分布漂移”以及少标签样本提供了新思路,其旨在利用已有的知识或模型,从相似源领域(SD)中学习共享知识,提升在目标领域(TD)任务中的诊断效果,解决了SD与TD间的“分布漂移”和标签数据少的问题,其中基于域自适应(DA)<sup>[9-10]</sup>的智能故障诊断方法通过解决数据分布差异、标注需求高、环境变化等问题,极大地增强了故障诊断系统的鲁棒性和实用性,逐渐成为研究的新热点<sup>[11-12]</sup>。文献[13]提出了一种基于多源领域适应的深度迁移学习,通过多对抗学习策略来获取特征表征,获取多源领域中的共享知识,采用深度卷积神经网络实现了滚动轴承故障诊断的改进;文献[14]提出了一种基于知识映射的对抗性域适应方法,实现了目标数据映射到源特征空间以探索域不变知识,并具有良好的泛化性能和较高的故障检出率;文献[15]提出了一种基于最优联合分布的深度域自适应方法,用于不同工况下机电传动系统的故障诊断,对于齿轮箱和滚动轴承故障诊断性能提升明显。

上述DA方法均假设SD与TD标签空间同构,然后强制对齐SD与TD间的分布差异,最小化二者特征空间差异,从而利用共享知识来实现TD任务的故障分类。然而,在实际工业中,旋转机械故障模式种类繁多且复杂,新故障模式可能会随时出现,这些未知的故障模式在训练数据中没有体现,导致闭集空间假设无法满足实际需求。这需要构建开放集

跨域故障诊断模型,在面对未知故障模式和变化的工作条件时,依然能够保持高效、准确的故障检测和诊断能力。

针对开放集跨域故障诊断问题,本文提出一种加权对抗域自适应学习网络(WADA)来处理旋转机械的数据标签空间异构故障诊断问题。首先,构建深度一维卷积神经网络来实现数据的特征提取;然后,引入特征细粒度分类器来区分多领域中的共享和离群特征;接着,采用加权模块指导域判别器实现域不变特征学习,筛选出SD与TD间的强相关共享特征;最后,域判别器利用域不变特征指导特征分类器的分类,判别TD任务中的未知故障种类,实现对开放集跨域任务下的精准智能诊断。本文主要内容如下:

1) 为解决旋转机械的开放集跨域故障诊断问题,提出一种新的WADA网络,通过引入细粒度分类信息来区分多领域中的共享和离群特征,使得模型能够区分TD任务中的未知故障种类。

2) 设计一种新的加权模块,结合域判别器,实现域不变特征的学习。加权模块指导域判别器,通过最小化全局损失,实现源域与目标域特征分布的对齐。域判别器利用域不变特征进一步优化特征分类器的分类性能。

3) 在美国凯斯西储大学(CWRU)滚动轴承数据集和齿轮箱数据集上进行测试,实验结果表明所提出方法的故障诊断效果优于传统的域自适应与深度学习方法,特别是在开放集数据中优势明显。

## 1 理论基础

### 1.1 开放集跨域故障诊断原理

对于开放集跨域故障诊断问题,给定有标签源领域 $\mathcal{D}_s = \{\mathcal{X}, \mathcal{Y}, p_s(x, y)\}$ 及其故障诊断任务 $\mathcal{T}^s = \{\mathcal{Y}^s, f^s(\cdot)\}$ 、无标签目标领域 $\mathcal{D}^t = \{\mathcal{X}^t, p(X^t)\}$ 及其故障诊断任务 $\mathcal{T}^t = \{\mathcal{Y}^t, f^t(\cdot)\}$ 。根据基本假设条件,有标签源领域与无标签目标领域数据分布不同,故障标签空间异构,即 $p(X^s) \neq p(X^t)$ ,  $\mathcal{Y}^s \neq \mathcal{Y}^t$ ,  $\mathcal{Y}^t \in \mathcal{Y}^s$ 。开放集跨域故障诊断是使得在标签源领域数据集 $\{X^s, Y^s\} = \{(x_i^s, y_i^s)\}_{i=1}^{n^s}$ 训练的模型能够完成无标签目标领域数据集 $\{X^t\} = \{(x_i^t)\}_{i=1}^{n^t}$ 故障诊断,构建从源领域到目标领域的有效映射函数 $f^t(\cdot)$ 。

### 1.2 对抗域自适应网络

对抗域自适应网络(ADAN)是通过对抗学习的机制来实现源域与目标域间的特征对齐,从而提高模型在目标域的泛化能力。其核心思想是引入一个对抗性训练过程,使得源域与目标域的特征分布在特征空间中尽可能接近,从而实现跨域适应。ADAN的主要组成部分及其工作原理如下。

如图1所示, ADAN 主要分为3个部分: 特征提取器 $G_f$ , 领域判别器 $G_d$ 和任务分类器 $G_y$ .  $G_f$ 用于从输入数据中提取高维特征表示, 可采用深度卷积神经网络(DCNN)或其他适合特定任务的神经网络架构;  $G_d$ 是一个二分类器, 旨在区分输入特征来自源域还是目标域, 通过对抗训练,  $G_d$ 的目标是最大化源域与目标域特征的可区分性;  $G_y$ 用于执行目标任务(如故障分类), 根据提取的特征进行标签预测,  $G_y$ 的训练主要是在源域数据上进行, 但是也会结合目标域数据来提升泛化性能. ADAN 的损失函数定义如下所示:

$$\begin{aligned} \mathbb{E}(\theta_f, \theta_y, \theta_d) = & \sum_{x_i \in \mathcal{D}_s} \mathcal{L}_y(G_y(G_f(x_i; \theta_f); \theta_y), y_i) - \\ & \lambda \sum_{x_i \in \mathcal{D}_s \cup \mathcal{D}_t} \mathcal{L}_d(G_d(G_f(x_i; \theta_f); \theta_d), d_i). \end{aligned} \quad (1)$$

其中:  $x_i$ 为输入的训练向量;  $\mathcal{L}_y$ 为标签分类器 $G_f$ 的损失;  $\mathcal{L}_d$ 为域判别器 $G_d$ 的损失;  $y_i$ 为 $x_i$ 的标签集合;  $d_i$ 为第 $i$ 个样本的二进制变量(域标签), 它表示 $i$ -th是来自源分布(若 $d_i = 0$ , 则 $x_i \in \mathcal{D}_s$ )还是来自目标分布(若 $d_i = 1$ , 则 $x_i \in \mathcal{D}_t$ );  $\lambda$ 为一个超参数. 基本的对抗性域适应追求SD与TD间的特征分布匹配, 假设SD与TD间的标签空间同构, 难以满足实际应用中故障诊断需求.

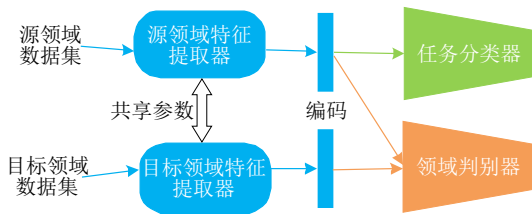


图1 对抗性领域适应框架

## 2 加权对抗域自适应故障诊断模型

所提出加权模块结合复杂多变的工业环境, 在

以下几方面具有优势: 动态、数据驱动的加权策略, 增强了模型对于未知故障类型的敏感性; 与对抗域自适应框架的深度融合, 提升了模型在开放集条件下的域适应能力; 生成器与分类器的互学习机制, 提高了对不同故障类型的细粒度区分能力.

### 2.1 模型的提出

本节结合一维卷积神经网络与对抗域自适应框架提出一种加权的对抗域自适应网络(WADA), 以实现旋转机械关键部件的开放集跨域智能故障诊断, WADA结构如图2所示. 在WADA中, 给定一组带标签的源域数据 $\mathcal{D}_s = \{\mathcal{X}, \mathcal{Y}, p_s(x, y)\}$ 和无标签的目标域数据 $\mathcal{D}^T = \{\mathcal{X}^T, p(X^T)\}$ , 其中 $\mathcal{X}_s \subset \mathcal{X}_T$ . 设计以下生成器和两个鉴别器: 特征生成网络 $G$ (使用1DCNN网络提取特征)和细粒度分类器 $C_{\text{extend}}$ , 后者通过特征生成网络 $G$ 从数据中获取特征, 并将其分为若干类别 $\mathcal{K} + 1$ , 这里 $\mathcal{K}$ 为已知类别的数量, 即 $\mathcal{K}$ 表示两个域中常见样本类型, 而 $\mathcal{K} + 1$ 类型在TD中是异常. 因此, 鉴别器能够评估每个样本属于常见类别 $\mathcal{K}$ 和异常类别 $\mathcal{K} + 1$ 的概率, 并将其输出 $\{\mathcal{P}_1, \mathcal{P}_2, \dots, \mathcal{P}_k, \mathcal{P}_{k+1}\}$ . 此外, 对抗训练方法能够使得WADA实现域适应.

WADA的目标是将未标记的常见TD样本准确分类到相应的已知类别中 $\mathcal{X}_T^{i=1,2,\dots,k} \subset \{1, 2, \dots, \mathcal{K}\}$ , 并将异常TD样本识别为未知, 但是不知道TD的任何标签信息. 因此, WADA使用一个弱训练分类器将TD样本识别为异常类别 $\mathcal{K} + 1$ , 并为异常建立一个伪决策边界. 然后, 对抗训练的方式允许训练特征生成器 $G$ 欺骗分类器, 生成器 $G$ 的最终目标是使得两个域的分布遵循相同的分布. 因此, 生成器将尝试增加 $\{\mathcal{K} + 1\}$ 类别的概率值, 这意味着可以区分异常.

为了在TD中区分异常数据与已知数据, 需要训练分类器和生成器正确分类SD样本. 因此, 首先扩展细粒度分类器对SD的交叉熵损失, 如下所示:

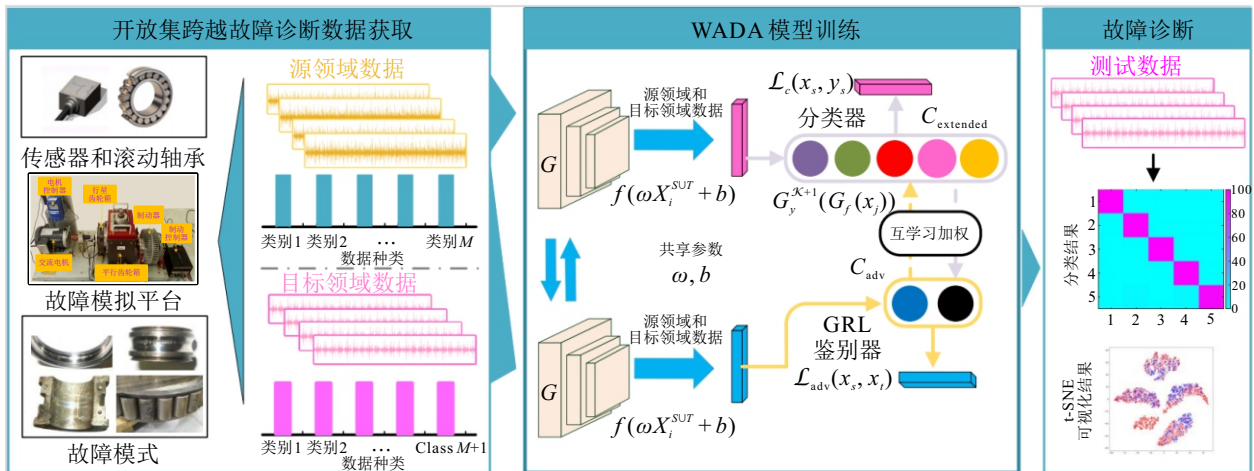


图2 加权对抗域自适应故障诊断框架

$$\begin{aligned} \mathcal{L}_{\text{cls}}^s &= \frac{1}{n_s} \sum_{x_i \in \mathcal{X}_s} \mathcal{L}_y(G_y^{1:\mathcal{K}}(G_f(x_i)), y_i) = \\ & \frac{1}{n_s} \sum_{i=1}^n y_i \log(G_y^{1:\mathcal{K}}(G_f(x_i))) + \\ & (1 - y_i) \log(1 - G_y^{1:\mathcal{K}}(G_f(x_i))). \end{aligned} \quad (2)$$

其中:  $\mathcal{L}_y$  为交叉熵损失;  $G_y$  是针对类别  $\{\mathcal{K} + 1\}$  的特征分类器;  $1 : \mathcal{K}$  为源域 (SD) 中的已知类别, 表示样本在源域相对于已知类别的概率. 所设计网络模型应与 SD 和 TD 数据在共享标签空间中的分布对齐. 同时, 分类器需要能够在异常值与已知类别样本间形成决策边界. 为实现这些目标, 使用 softmax 函数构建分类器, 相关输出用作软实例权重, 即  $\omega_j = G_y^{\mathcal{K}+1}(G_f(x_j))$ ,  $\omega_j$  值越大, 该样本来自异常值的可能性越高. 最后使用二元交叉熵损失来训练网络模型, 具体方法如下所示:

$$\begin{aligned} \mathcal{L}_d &= \\ & \frac{1}{n_s} \sum_{x_i \in \mathcal{X}_s} \mathcal{L}_{\text{bce}}(G_d(G_f(x_i)), d_i) + \\ & \frac{1}{\sum_{x_i \in \mathcal{X}_T} (1 - \omega_j)} \sum_{x_i \in \mathcal{X}_T} (1 - \omega_j) \mathcal{L}_{\text{bce}}(G_d(G_f(x_j)), d_j). \end{aligned} \quad (3)$$

这里:  $\mathcal{L}_{\text{bce}}$  为二元交叉熵损失,  $G_d$  为域鉴别器的输出,  $d_i$  和  $d_j$  表示样本被标记为已知 ( $d_{i,j} = 0$ ) 或未知 ( $d_{i,j} = 1$ ). 此外, 模型还需要从未标记的目标域中获取异常值, 以改变生成器的参数, 为额外的“未知”类别. 然后, 生成器  $G$  可以最大化分类器  $C$  的错误, 以达到拒绝异常值的目的. 基于分类器  $C$  的已知类别和异常类别权重  $\omega_j$ , 可以定义一个加权损失将“异常”类别分为

$$\mathcal{L}_{\text{cls}}^t = \frac{1}{\mathcal{K}} \frac{1}{\sum_{x_j \in \mathcal{X}_T} \omega_j} \sum_{x_j \in \mathcal{X}_T} \omega_j \mathcal{L}_y(G_y^{1+\mathcal{K}}(G_f(x_j)), 1 + \mathcal{K}). \quad (4)$$

### 2.2 构建模型损失函数

所提出 WADA 网络的主要核心是通过细粒度分类器加权筛选出共享特征, 并指导领域判别器区分领域特征, 同时, 用共享特征指导特征分类器以实现未知类识别, 二者在训练过程中相互学习. 网络整体损失函数为对抗域适应损失和分类损失, 可表示为

$$\mathcal{L}_{\text{total}} = \mathcal{L}_d + \mathcal{L}_{\text{cls}}^s + \mathcal{L}_{\text{cls}}^t. \quad (5)$$

结合对抗性领域自适应技术来匹配 SD 和 TD 中已知样本的分布, 并通过生成器  $G$  来实现拒绝异

常类别的效果. 使用  $\theta_d$ 、 $\theta_f$  和  $\theta_y$  表示  $C_{\text{adv}}$ 、 $G$  和  $C_{\text{extend}}$  的权重和偏置. 然后, 可通过以下方式获得最优参数:

$$(\tilde{\theta}_y, \tilde{\theta}_d) = \arg \min_{\theta_y, \theta_d} \mathcal{L}_{\text{cls}}^s + \mathcal{L}_{\text{cls}}^t + \mathcal{L}_d, \quad (6)$$

$$(\tilde{\theta}_f) = \arg \min_{\theta_f} \mathcal{L}_{\text{cls}}^s + \mathcal{L}_{\text{cls}}^t - \mathcal{L}_d. \quad (7)$$

在整个训练过程中, 使用以下规则更新参数:

$$\theta_f \leftarrow \theta_f - \mu \left( \frac{\partial \mathcal{L}_{\text{cls}}^s}{\partial \theta_f} + \frac{\partial \mathcal{L}_{\text{cls}}^t}{\partial \theta_f} - \lambda \frac{\partial \mathcal{L}_d}{\partial \theta_f} \right), \quad (8)$$

$$\theta_y \leftarrow \theta_y - \mu \left( \frac{\partial \mathcal{L}_{\text{cls}}^s}{\partial \theta_y} + \frac{\partial \mathcal{L}_{\text{cls}}^t}{\partial \theta_y} \right), \quad (9)$$

$$\theta_d \leftarrow \theta_d - \mu \frac{\partial \mathcal{L}_d}{\partial \theta_d}. \quad (10)$$

其中:  $\mu$  为学习率,  $\frac{\partial \mathcal{L}_{\text{cls}}^s}{\partial \theta_y}$ 、 $\frac{\partial \mathcal{L}_{\text{cls}}^t}{\partial \theta_f}$  和  $\frac{\partial \mathcal{L}_d}{\partial \theta_d}$  为损失函数相对于  $\tilde{\theta}_f$ 、 $\tilde{\theta}_y$  和  $\tilde{\theta}_d$  参数的偏导数, 符号  $-\lambda$  表示梯度的反转.

### 2.3 故障诊断流程

WADA 故障诊断流程通过结合加权机制和对抗域自适应技术, 旨在有效解决“分布漂移”问题, 整体提高了模型的鲁棒性和泛化能力, 适用于复杂多变的实际工业环境. 算法流程如图 3 所示, 具体步骤如下.

step 1: 构建数据集. 将各标签源域和目标域数

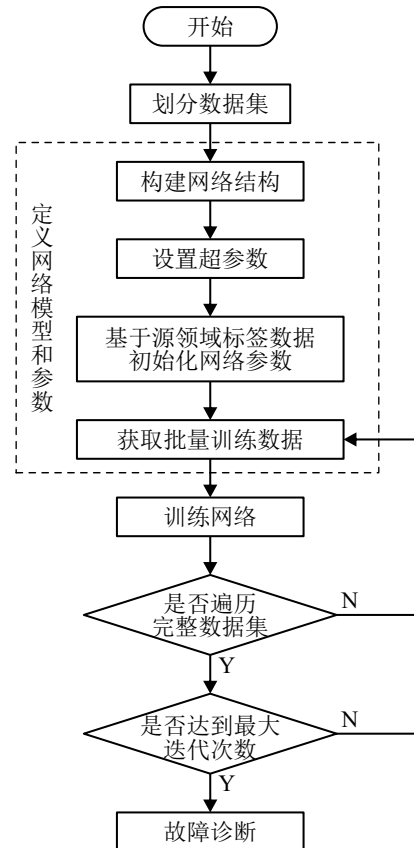


图3 加权对抗性领域自适应网络故障诊断流程

据进行样本分割和标准化, 并对样本进行标注, 获取标签源域数据集; 同时, 获取目标域训练数据集和测试数据集, 并对测试数据集样本进行标注, 获取标签测试数据集.

step 2: 确定网络结构和设置超参数. 确定 WADA 的网络结构, 并设置网络中的超参数.

step 3: 网络预训练. 利用标签源域数据集和目标域训练数据集进行预训练.

step 4: 获取批量数据. 从各标签源域数据集和无标签目标域训练数据集中获取批次的训练数据和测试数据.

step 5: 训练判别器、特征提取器和分类器. 利用源领域数据和无标签目标领域数据训练特征提取器和判别器; 利用源领域数据训练分类器.

step 6: 重复 step 4 和 step 5, 直至遍历所有源域数据和无标签目标领域数据集.

step 7: 重复 step 6, 直至达到最大迭代次数. 然后, 将训练好的特征提取器和分类器级联成最后的故障诊断模型, 利用测试数据集验证模型的有效性.

### 3 实验与结果分析

为验证 WADA 的半监督泛化故障诊断性能, 本节选择行星齿轮箱数据集和 CWRU 滚动轴承数据集来验证算法, 表 1 为两个数据集的故障标签信息. 同时, 选择 5 种先进的智能诊断方法与 WADA 进行比较. 所有实验均在 Python 平台上进行, 所有算法的具体细节将在下文给出.

表1 实验数据集标签信息

标签	1	2	3	4	5	6	7	8	9	10
CWRU滚动轴承数据集										
位置	NR	IF			BF			OF		
尺寸	0	7	14	21	7	14	21	7	14	21
DDS齿轮箱数据集										
位置	NR	CF	MF	RF	SF	—	—	—	—	—

#### 3.1 数据集

1) 滚动轴承测试数据集: 该数据集由 CWRU 轴承数据中心<sup>[16]</sup>提供. 本文使用的实验数据是在采样频率为 12 kHz 的情况下, 从驱动端轴承收集的, 包括正常和 3 种类型的故障, 即内圈 (IF)、球体 (BF) 和外圈 (OF). 记录了 1 797 RPM/0 马力 ~ 1 720 RPM/3 马力电机负载下的振动数据. 此外, 每种故障类型有 3 种故障直径, 分别为 0.007 英寸、0.014 英寸和 0.021 英寸. 因此, 每个域中有 10 个类别, 每个类别有 1 000 个样本, 每个样本包含 1 024 个数据点. 根据上述信息, 可用于实验验证的开放集跨域故障诊断场

景共有 6 个 (名为  $R_1 \sim R_6$ ), 如表 2 所示.

表2 开放集跨域故障诊断任务信息

任务	领域偏移/hp	源领域 标签空间	目标领域 标签空间	离群标签 样本
$R_1$	0 → 1	1, 2, 3, 4	1, 2, 3, 4, 5	5
$R_2$	0 → 2	1, 2, 3, 4	1, 2, 3, 4, 6	6
$R_3$	0 → 3	1, 2, 3, 4	1, 2, 3, 4, 7	7
$R_4$	1 → 0	1, 2, 3, 4	1, 2, 3, 4, 8	8
$R_5$	1 → 2	1, 2, 3, 4	1, 2, 3, 4, 9	9
$R_6$	1 → 3	1, 2, 3, 4	1, 2, 3, 4, 10	10
$A \rightarrow W$	0 → 2	1, 2, 3, 4	1, 2, 3, 4, 5	5
$W \rightarrow A$	2 → 0	1, 2, 3, 4	1, 2, 3, 4, 5	5

2) 齿轮箱测试数据集: 数据来自传动系统动态模拟器 (DDS)<sup>[17]</sup>, 实验设备如图 4 所示. 设备包括行星齿轮箱、驱动电机、制动器和制动控制器等主要组件. 对于齿轮箱, 有两种工作条件 (域), 随着电机速度-负载配置的变化, 这些场景可构建为  $A \rightarrow W$  和  $W \rightarrow A$ . 此外, 每个域包含正常状态和 4 种类型的故障, 即齿轮断裂 (CF)、齿轮缺失 (MF)、齿根断裂 (RF)、齿面损坏 (SF).

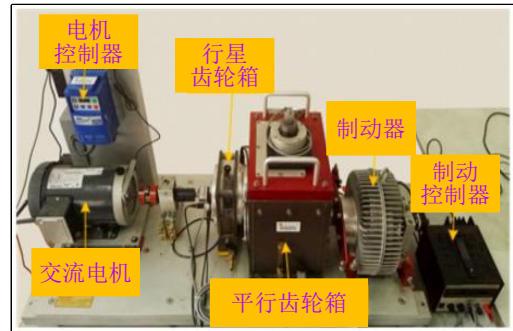


图4 齿轮箱数据实验台

#### 3.2 比较方法和应用细节

为验证 WADA 算法的性能, 使用上述两个数据集设计了多种开放集跨域诊断场景, 这些场景在表 2 中进行了总结. 5 种比较方法包括支持向量机 (SVM)<sup>[18]</sup>、深度卷积神经网络 (DCNN)<sup>[19]</sup>、基于深度模型的领域适应故障诊断 (DAFD)<sup>[20]</sup>、神经网络的领域对抗训练 (DANN)<sup>[21]</sup>、通过反向传播的开放集领域适应 (OSBP)<sup>[22]</sup>. 同时, 1D-CNN 作为网络基础架构, 对特征提取效果起到关键性作用, 对于高频采样数据, 较小的接收域通常更有效, 因为它们能够捕获到高频下的细微变化. 结合经验性分析, 对 1D-CNN 架构进行了优化, 具体操作为调整接收域大小, 根据分类准确率确定最优接收域值, 具体的实验参数如表 3 所示. 对比方法以及 WADA 的细节总结如下:

1) SVM: 使用高斯核函数, 实验通过 scikit-learn

表3 1D-CNN 架构详细配置

连接层	类别	R-C-S-P	输出
输入层	数据	—	1024×1
模块1-卷积1	卷积操作	64×1-16-1-32	1024×16
模块1-池化1	最大池化	2×1-1-2	512×16
模块1-正则化1	批正则化	—	512×16
模块1-ReLU 1	ReLU	—	512×16
模块2-卷积2	卷积操作	5×1-32-1-2	512×32
模块2-池化2	最大池化	2×1-1-2	256×32
模块2-正则化2	批正则化	—	256×32
模块2-ReLU2	ReLU	—	256×32
模块3-卷积3	卷积操作	5×1-64-1-2	256×64
模块3-池化3	最大池化	2×1-1-2	128×64
模块3-正则化3	批正则化	—	128×64
模块3-ReLU 3	ReLU	—	128×64
模块4-卷积4	卷积操作	5×1-64-1-2	128×64
模块4-池化4	最大池化	2×1-1-2	64×64
模块4-正则化4	批正则化	—	64×64
模块4-ReLU 4	ReLU	—	64×64
全连接层1	全连接	1/1024	1024
全连接层2	全连接	1/1024	1024
输出层	全连接	1/C	C

注: R-C-S-P表示接收域大小-通道数-步长-填充.

= 0.24.1 进行, 对于所有可调参数, 使用模块推荐的默认值.

2) DCNN: 这是 WADA 的基本卷积神经网络, 它有一个与 WADA 一致的网络框架, 以验证所提出方法的有效性. 具体的详细参数如表 3 所示.

3) DAFD: 这是一个基于最大均值差异 (MMD) 和堆叠自编码器的流行比较方法, 有几个可调节的超参数, 包括隐藏层的数量、学习率、权重因子等, 本文选择了原文中优化的参数.

4) DANN: 致力于通过 SD 与 TD 间相同的映射实现分布对齐. 具体而言, 通过对抗训练方法, 领域鉴别器不能区分 SD 与 TD, 使得 SD 与 TD 在分布上对齐. 其特征提取器与 WADA 一致.

5) OSBP: 这是由 Saito 等<sup>[22]</sup>提出的经典开放集

DA 方法, 与所提出方法进行比较, 所有主要优化参数在原文中给出.

6) WADA: 主要结构参数的细节显示在表 3 中. 此外, 初始化阶段的迭代次数和小批量设置分别为 10 和 64. 初始化后, 模型的学习率和迭代次数调整为 0.001 和 2000. 所有参数如表 4 所示.

表4 实验任务中使用的参数

参数	值	参数	值
迭代次数	2000	$N_{input}$	256
批量大小	256	动量	0.9
学习率 $\alpha$	1e-3	权重衰减系数	0.0005

所有实验测试均在配备 32 GB RAM 的 Windows 台式机上进行, Python 版本为 3.6, Pytorch 版本为 1.7.1, 中央处理器和图形处理单元分别为 i9-10900K 和 GTX 3090, 24 GB VRAM. 此外, 计算统一设备架构 (CUDA) 选为 11.0, 因为低版本的 CUDA 不支持 GTX 3090 的计算.

评估指标: 本文对开放集跨域故障诊断采用两种评估指标: OS: 所有类别 (包括未知类别) 的平均准确率; OS\*: 仅在已知类别上的平均准确率.

### 3.3 滚动轴承故障诊断实验

本节使用滚动轴承的数据构建了 6 个开放集跨域故障诊断场景, 并在表 5 中呈现了实验结果. 结果清楚地显示, 与其他 5 种比较方法相比, WADA 实现了显著更高的平均准确率. 在这些方法中, SVM 在准确性方面表现最差, 还观察到 OS 与 OS\* 分类准确率间的明显差异, 这种现象可归因于 SVM 在处理 TD 中异常样本时的局限性. 作为一种监督学习方法, SVM 错误地将异常样本归类为已知类别, 且其难以从 SD 和 TD 中提取共同特征. DCNN 的分类准确率显著优于 SVM, 但是其不能处理 TD 中的异常样本; DAFD、DANN 也有同样的问题. 对于 OSBP, 它实现了 10% 以上的平均性能提升. 然而, 它的表现不如 WADA 模型一致, 这表明虽然 OSBP 模型准

表5 开放集跨域故障诊断任务信息

任务	方法											
	SVM		DCNN		DAFD		DANN		OSBP		WADA	
	OS	OS*	OS	OS*	OS	OS*	OS	OS*	OS	OS*	OS	OS*
$R_1$	46.70	58.37	81.25	<b>100.0</b>	77.00	93.86	79.09	97.14	90.50	96.50	95.74	99.57
$R_2$	39.92	49.90	76.56	96.32	79.09	96.15	79.11	98.06	91.00	95.00	<b>96.18</b>	<b>98.35</b>
$R_3$	26.42	33.03	68.75	88.67	76.66	91.47	79.01	96.32	94.25	96.00	<b>99.26</b>	99.35
$R_4$	45.18	56.48	82.81	<b>100.0</b>	78.72	95.42	79.03	94.84	92.40	95.90	<b>97.34</b>	99.30
$R_5$	54.60	68.25	85.94	<b>100.0</b>	65.55	74.36	78.98	95.10	94.80	97.00	<b>99.20</b>	99.75
$R_6$	33.48	41.85	79.56	96.47	78.84	94.53	79.11	98.21	93.90	96.50	<b>97.80</b>	99.88
平均值	41.05	51.31	79.15	96.91	75.98	90.96	79.06	96.61	92.81	96.15	<b>97.59</b>	<b>99.37</b>

确,但是不如 WADA 模型有效. 最后,所提出方法不仅可以提高模型的泛化能力,还可以在开放集跨域故障诊断场景中拒绝 TD 的异常样本.

为进一步分析比较方法和 WADA 在开放集跨域故障诊断场景中的域移动适应和异常拒绝能力,使用 t 分布随机邻域嵌入 (t-SNE) 可视化 SVM、

DCNN、DANN、DAFD、OSBP 和 WADA 的特征,同时,由于 SVM 本身不具备特征提取的功能,其使用的特征直接对应于原始数据. 通过分析支持向量机的特征可视化图,可以有效评估其他方法在开放集跨域故障诊断任务中的性能表现,具体结果如图 5 所示.

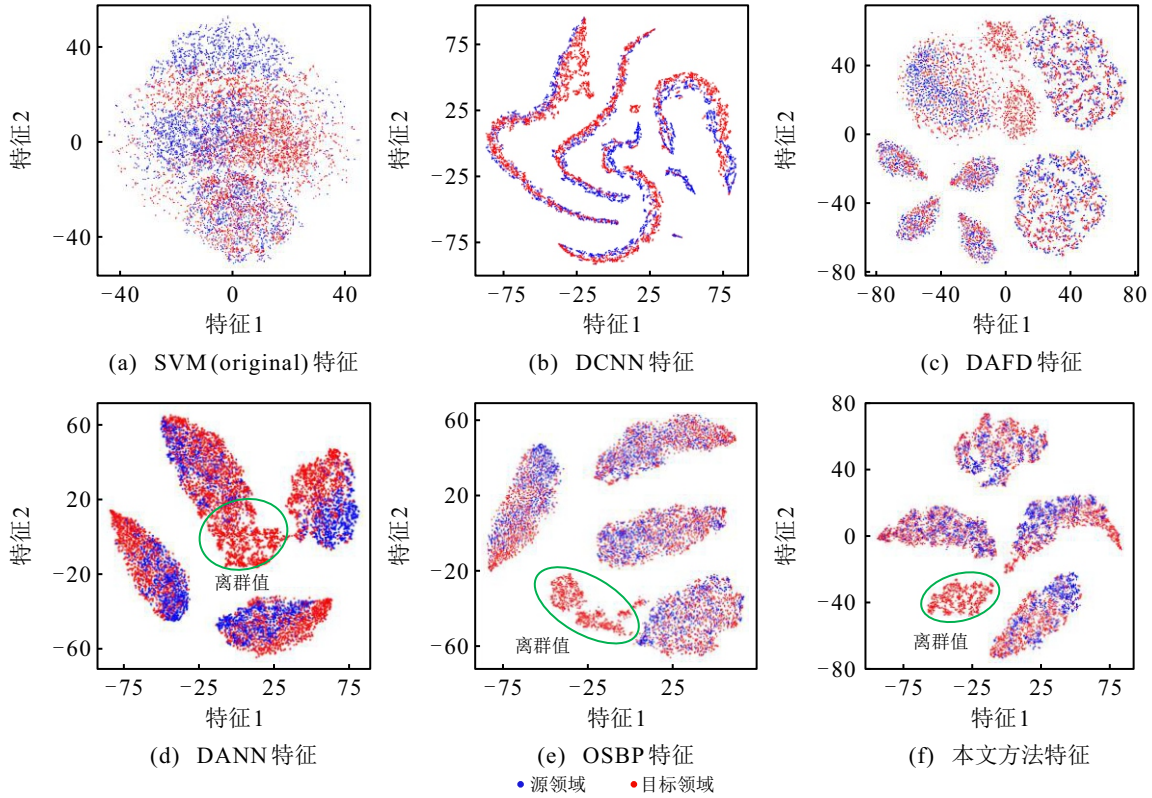


图5 t-SNE降维显示的特征

由图 5(a) 所示的可视化分析中明显观察到,所有已知类型的特征与异常特征在特征空间中高度混合,4 种类型间未显示出明显的区分界限. 这一现象反映了原始数据集中各类别特征的高度重叠性. 因此, SVM 在此数据集上难以生成有效的类别间决策边界,导致其在分类各类型时面临挑战. 与此相对照,图 5(b) 和图 5(c) 中的可视化结果突出显示了通过 DCNN 和 DANN 获得的特征明显区分了 4 种不同的类型. 这些方法借助多层特征提取架构增强了泛化能力,有效减少了 SD 与 TD 间的域偏移. 然而,这两种方法存在的一个显著缺陷是它们倾向于将异常特征纳入 SD 与 TD 的共同特征集中,从而引入了负迁移效应,导致在开放集跨域故障诊断场景中产生次优结果. 相比之下,针对开放集跨域故障诊断的 OSBP,其结果显示在图 5(e) 中,可以清楚地看出源域与目标域间的公共特征在减少域偏移方面实现了更好的匹配. 此外, OSBP 在保证对齐公共类别的同时,实现了对异常特征的有效区分,展现出在处理开

放集跨域故障诊断问题时的优越性. 由图 5(f) 的分析中可以观察到,情况与之前有显著不同. bSD 和 TD 中,4 种已知类别间的域差异显著减少. 更为重要的是,异常样本也成功地从公共特征空间中区分出来,有效避免了负迁移现象. 这些观察结果表明 WADA 在开放集跨域故障诊断任务中,不仅实现了域间的有效匹配,且能够有效拒绝异常类别的干扰. 这一性能表明, WADA 能够处理复杂的域偏差问题,且在保持类别间清晰分界的同时,能够提高系统的整体诊断准确性.

表6 开放集跨域故障诊断任务信息

方法	$A \rightarrow W$		$W \rightarrow A$		平均值
	OS	OS*	OS	OS*	
SVM	42.35	52.94	44.70	55.87	48.97
DCNN	79.48	<b>99.35</b>	79.96	<b>99.95</b>	89.69
DAFD	79.02	95.00	78.11	95.30	86.86
DANN	78.74	96.63	78.25	96.98	87.65
OSBP	87.80	97.34	90.50	95.16	92.70
WADA	<b>96.87</b>	99.16	<b>97.34</b>	99.54	<b>98.23</b>

### 3.4 齿轮箱故障诊断实验

表6为各种方法在两个开放集跨域故障诊断场景中的诊断性能. 由表6可见: 与其他方法相比, SVM显示出最低的平均准确率(48.97%); DCNN、DAFD、DANN仍然显著优于SVM; 对于开放集DA方法, OSBP在其性能中弱于WADA; WADA平均准确率为98.23%. 这些结果表明, 所提出方法可在开放集跨域故障诊断中很好地区分异常样本.

### 3.5 收敛性能与实时性分析

在使用基于深度学习的故障诊断技术时, 需要考虑两个重要因素: 模型收敛性能和训练时间. 这些模型的复杂性, 表现在它们对于深度网络和数据集的依赖, 常常导致训练时间延长. 为确保全面地分析, 必须比较所有考虑中方法的训练时间. 首先, 算法的运行时间均有可能通过使用更强大的CPU或GPU而减少, 因为硬件性能可以影响计算速度. 此外, 本研究主要考虑了3个参数: 训练时间、分类精度和迭代次数, 以评估模型的收敛性能. 以CWRU滚动轴承 $R_2$ 任务下的训练信息为可视化对象, 对于训练时间, 表7提供了所有详细信息. SVM有最短的训练时间, 因为它不需要迭代计算. WADA和CIDAA需要最长的训练时间, 但是这在可接受范围内.

表7 所有实验模型 $R_2$ 任务下的训练时间

所有实验方法	训练时间/s	迭代次数
SVM	3.25	0
DCNN	92.87	1000
DAFD	220.12	1000
DANN	249.56	1000
OSBP	246.80	1000
WADA	225.41	1000

图6为WADA的收敛性能. 由图6可见, 验证了WADA在有限次迭代中的快速收敛, 尽管训练与测试数据集内的分布不同, 以及类别与域间的标签空间不一致, 依然呈现出良好的故障诊断结果. 最后, 在可视化域鉴别器损失曲线时, 选取每个batch\_size下的平均值. 由于域鉴别器的损失在训练过程中存

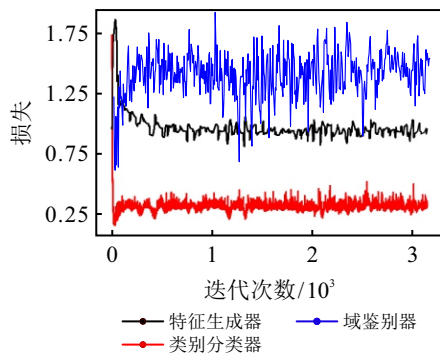


图6 WADA收敛性能曲线

在较大的波动性, 可视化所有批次的点可能会导致曲线过于嘈杂, 难以清晰反映损失的整体变化趋势. 类别分类器损失在后期的训练中趋于稳定, 为了更清晰地展示训练过程中损失的变化情况, 可视化所有batch\_size下的样本损失点.

### 3.6 噪声分析

在实际工业场景中, 噪声对于模型性能的影响可能会显著降低其可靠性和适用性. 噪声的种类、来源以及强度决定了对数据质量和模型表现的具体干扰程度, 本节以滚动轴承任务 $R_3$ 作为无噪声数据, 选取高斯噪声、椒盐噪声和标签噪声对模型的泛化性能进行分析, 同时, 也考虑了不同噪声比例下的实验结果, 表8为WADA的抗噪结果, 从中分析可得到如下结论: 1) 高斯噪声会扰乱全局特征分布, 椒盐噪声破坏局部特征提取, 而标签噪声直接影响分类边界学习. 不同类型的噪声会导致分类准确率、特征提取能力和离群样本检测能力下降. 2) 轻度噪声(如0.1%)对于模型性能影响有限, 但是, 中度至重度噪声(如10%或更高)会降低分类准确率, 标签噪声对性能的破坏尤为显著, 导致模型泛化能力和鲁棒性下降.

表8 不同噪声和比例下的故障诊断结果

噪声类型	噪声比例/%	分类准确率/%
无噪声	0	99.26
	0.1	98.85
	5	97.53
	10	92.67
	20	89.48
高斯噪声	0.1	98.91
	5	96.83
	10	92.32
	20	87.47
	椒盐噪声	0.1
5		94.68
10		89.71
20		81.59
标签噪声		0.1
	5	94.68
	10	89.71
	20	81.59

## 4 结论

本文提出了一个符合旋转机械实际工作条件的领域适应场景, 称为开放集跨域故障诊断. 针对这一问题, 本文设计了一种新颖的双分类器互学习的域自适应网络WADA, 并将其应用于旋转机械的实验数据(特别是滚动轴承和齿轮箱). 全面的实验和分析结果验证了所提出方法在开放集跨域故障诊断场景中的有效性. 尽管WADA在开放集跨域故障诊断方面表现出显著的有效性, 但是在处理高度异构的异常数据时, 识别精度和稳定性仍然有待提高. 异常类别的复杂性和多样性可能会影响模型的准确诊断.

同时,工业环境中噪声可以掩盖或模拟为故障信号,导致误诊或漏诊。因此,未来的工作可以研究异常数据的内在特征以及系统地评估模型在不同类型和强度的噪声条件下的表现,确保模型在实际应用中的鲁棒性。

### 参考文献 (References)

- [1] 刘小峰, 金燕, 柏林. 从仿真到现实的多层级虚实域适应的滚动轴承故障诊断方法[J]. 控制与决策, 2025, 40(3): 889-898.  
(Liu X F, Jin Y, Bo L. Simulation-to-reality fault diagnosis based on multi-level joint adaptive network in virtual and real domains of rolling bearing[J]. Control and Decision, 2025, 40(3): 889-898.)
- [2] He M, He D. Deep learning based approach for bearing fault diagnosis[J]. IEEE Transactions on Industry Applications, 2017, 53(3): 3057-3065.
- [3] LeCun Y, Bengio Y, Hinton G. Deep learning[J]. Nature, 2015, 521(7553): 436-444.
- [4] Zhuang F Z, Qi Z Y, Duan K Y, et al. A comprehensive survey on transfer learning[J]. Proceedings of the IEEE, 2021, 109(1): 43-76.
- [5] 薛萍, 郝鹏, 王宏民. 基于径向基神经网络的新型齿轮故障诊断方法[J]. 控制与决策, 2022, 37(2): 409-416.  
(Xue P, Hao P, Wang H M. Novel gear fault diagnosis method based on RBF neural network[J]. Control and Decision, 2022, 37(2): 409-416.)
- [6] 王进花, 岳亮辉, 曹洁, 等. 基于随机变分推理贝叶斯神经网络的发电机轴承故障诊断[J]. 控制与决策, 2023, 38(4): 1015-1021.  
(Wang J H, Yue L H, Cao J, et al. Fault diagnosis of generator bearing based on stochastic variational inference Bayesian neural network[J]. Control and Decision, 2023, 38(4): 1015-1021.)
- [7] 黎德才, 张冕, 王科盛, 等. 基于改进的胶囊网络的行星齿轮箱故障诊断方法[J]. 控制与决策, 2023, 38(3): 661-669.  
(Li D C, Zhang M, Wang K S, et al. Fault diagnosis method of planetary gearbox based on enhanced capsule network[J]. Control and Decision, 2023, 38(3): 661-669.)
- [8] 李晨昀, 景旭文, 李炳强, 等. 基于深度多源子域适应网络的滚动轴承跨域故障诊断[J]. 控制与决策, 2024, 39(3): 877-884.  
(Li C Y, Jing X W, Li B Q, et al. Cross-domain fault diagnosis of rolling bearings based on deep multi-source sub-domain adaptation networks[J]. Control and Decision, 2024, 39(3): 877-884.)
- [9] Weiss K, Khoshgoftaar T M, Wang D D. A survey of transfer learning[J]. Journal of Big Data, 2016, 3(1): 9.
- [10] Yang S K, Kong X G, Wang Q B, et al. A multi-source ensemble domain adaptation method for rotary machine fault diagnosis[J]. Measurement, 2021, 186: 110213.
- [11] Wei D D, Han T, Chu F L, et al. Weighted domain adaptation networks for machinery fault diagnosis[J]. Mechanical Systems and Signal Processing, 2021, 158: 107744.
- [12] Lu B L, Dibaj A, Gao Z, et al. A class-imbalance-aware domain adaptation framework for fault diagnosis of wind turbine drivetrains under different environmental conditions[J]. Ocean Engineering, 2024, 296: 116902.
- [13] Zhu J, Chen N, Shen C Q. A new multiple source domain adaptation fault diagnosis method between different rotating machines[J]. IEEE Transactions on Industrial Informatics, 2021, 17(7): 4788-4797.
- [14] Li Q, Shen C Q, Chen L, et al. Knowledge mapping-based adversarial domain adaptation: A novel fault diagnosis method with high generalizability under variable working conditions[J]. Mechanical Systems and Signal Processing, 2021, 147: 107095.
- [15] Liu Z H, Lu B L, Wei H L, et al. Fault diagnosis for electromechanical drivetrains using a joint distribution optimal deep domain adaptation approach[J]. IEEE Sensors Journal, 2019, 19(24): 12261-12270.
- [16] Smith W A, Randall R B. Rolling element bearing diagnostics using the Case Western Reserve University data: A benchmark study[J]. Mechanical Systems and Signal Processing, 2015, 64: 100-131.
- [17] Shao S Y, McAleer S, Yan R Q, et al. Highly accurate machine fault diagnosis using deep transfer learning[J]. IEEE Transactions on Industrial Informatics, 2019, 15(4): 2446-2455.
- [18] Gu Y K, Zhou X Q, Yu D P, et al. Fault diagnosis method of rolling bearing using principal component analysis and support vector machine[J]. Journal of Mechanical Science and Technology, 2018, 32(11): 5079-5088.
- [19] 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.
- [20] Lu W N, Liang B, Cheng Y, et al. Deep model based domain adaptation for fault diagnosis[J]. IEEE Transactions on Industrial Electronics, 2017, 64(3): 2296-2305.
- [21] Ganin Y, Ustinova E, Ajakan H, et al. Domain-adversarial training of neural networks[J]. Journal of Machine Learning Research, 2016, 17(1): 2096-2030.
- [22] Saito K, Yamamoto S, Ushiku Y, et al. Open set domain adaptation by backpropagation[C]. Computer Vision — ECCV 2018. Munich, 2018: 156-171.

### 作者简介

陆碧良 (1995-), 男, 博士生, 主要研究方向为旋转机械故障诊断、深度学习、不平衡迁移学习, E-mail: [lubiliang@hnu.edu.cn](mailto:lubiliang@hnu.edu.cn);

张英杰 (1970-), 男, 教授, 博士, 主要研究方向为人工智能、智能控制、故障诊断、节能控制, E-mail: [zhangyj@hnu.edu.cn](mailto:zhangyj@hnu.edu.cn);

孙庆帅 (1994-), 男, 博士生, 主要研究方向为时间序列分析和预测、故障诊断, E-mail: [sunqs@hnu.edu.cn](mailto:sunqs@hnu.edu.cn);

李蒲德 (1991-), 男, 博士生, 主要研究方向为风电机组传动链故障诊断、智能控制, [lipude@hnu.edu.cn](mailto:lipude@hnu.edu.cn).