

# 基于多头自注意力和动态对齐的轴承故障诊断<sup>\*</sup>

李杰 刘天羽

(上海电机学院电气学院 上海 201306)

**摘要:**为解决滚动轴承在不同工况下故障诊断过程中源域与目标域之间分布差异导致的诊断性能下降问题,本文提出了一种融合多头自注意力机制和动态联合分布自适应的轴承故障诊断方法。首先,在特征提取模块中引入多头自注意力机制,从原始振动信号中提取更具判别性和域不变特征。其次,分别采用最大均值差异与局部最大均值差异对齐边缘分布及条件分布,从而缩小源域与目标域之间的分布差异。最后,设计一种新的动态权重因子自适应地调整边缘分布与条件分布的权重,提升跨域故障诊断的鲁棒性与泛化能力。实验结果表明,所提出的方法在两种公开数据集上分类准确率分别达到了99.84%和98.97%,显著优于其他方法。同时在强噪声干扰下仍表现出良好的稳定性与鲁棒性,为滚动轴承故障诊断提供了一种有效的解决方案。

**关键词:**轴承故障诊断;领域自适应;多头自注意力机制;动态联合分布自适应;迁移学习

**中图分类号:**TH133.3; TN911   **文献标识码:**A   **国家标准学科分类代码:**460.2020

## Bearing fault diagnosis method based on multi-head self-attention and dynamic alignment

Li Jie Liu Tianyu

(School of Electrical Engineering, Shanghai Dianji University, Shanghai 201306, China)

**Abstract:** To address the issue of performance degradation in rolling bearing fault diagnosis under varying operating conditions caused by distribution discrepancies between the source and target domains, this paper proposes a novel fault diagnosis method that integrates a multi-head self-attention mechanism with dynamic joint distribution adaptation. Firstly, a multi-head self-attention mechanism is incorporated into the feature extraction module to extract more discriminative and domain-invariant features from raw vibration signals. Secondly, maximum mean discrepancy and local maximum mean discrepancy are employed to align the marginal and conditional distributions, thereby reducing the distribution difference between the source and target domains. Finally, a dynamic weighting factor is designed to adaptively adjust the importance of marginal and conditional distribution alignment, enhancing the robustness and generalization ability of cross-domain fault diagnosis. The experimental results demonstrate that the proposed method achieved classification accuracies of 99.84% and 98.97% on two public datasets, significantly outperforming other approaches. Moreover, it maintained strong stability and robustness under severe noise interference, providing an effective solution for rolling bearing fault diagnosis.

**Keywords:** bearing fault diagnosis; domain adaptation; multi-head self-attention mechanism; dynamic joint distribution adaptation; transfer learning

## 0 引言

旋转机械设备广泛用于工业生产、交通运输、电力等领域,其运行状态直接影响机械的性能与安全<sup>[1]</sup>。然而,这些设备大多处于复杂且恶劣的运行环境中,其故障极易引发设备损坏乃至安全事故,给生产系统带来巨

大的经济损失<sup>[2]</sup>。因此,如何实现对轴承健康状态的准确识别,已成为保障设备稳定运行和提升生产效率的核心问题<sup>[3]</sup>。

振动信号由于能够连续监测轴承的工作状态,已被广泛应用于轴承健康状态的诊断<sup>[4]</sup>。传统诊断方法依赖人工特征提取和先验知识,存在泛化性差、特征不稳定等问题。

因此,深度学习因其强大的特征提取能力和端到端的诊断优势,逐渐成为轴承故障诊断领域的研究热点<sup>[5]</sup>。例如,Zhang等<sup>[6]</sup>应用卷积神经网络构建了多尺度残差注意力网络,在有限的标记数据下提高了故障诊断的性能。Xue等<sup>[7]</sup>通过多域特征融合和注意力机制进行电机轴承故障诊断,有效提高了模型的诊断性能。

尽管深度学习在故障诊断中表现优异,但其通常假设训练数据与测试数据来自相同分布,这一假设在实际中难以成立<sup>[8]</sup>。由于不同设备、负载和工况下的信号存在明显差异,导致模型在跨域场景中性能下降<sup>[9]</sup>。为了解决这一问题,迁移学习<sup>[10]</sup>为机械设备故障诊断提供了一种可行的解决方案。其通过在源域中学习已标注的知识,并迁移至目标域未标注样本,有效提升了模型的泛化能力。

无监督域适应是迁移学习的一个重要分支,旨在通过利用源域上带标签的数据和目标域上未带标签的数据,解决源域和目标域之间分布差异的问题<sup>[11-12]</sup>。例如,Sun等<sup>[13]</sup>提出了深度相关对齐(deep coral, DCORAL)方法,通过线性变换对源域和目标域分布的二阶统计量进行对齐,有效缓解了域偏移问题。Cao等<sup>[14]</sup>提出了一种基于无偏估计的柯西核诱导的最大均值差异(maximum mean discrepancy, MMD)算法,从而实现全局特征对齐。然而,上述方法忽略了源域与目标域之间条件分布差异,从而限制了迁移性能。为进一步提升迁移效果,Wang等<sup>[15]</sup>提出了一种深度子域迁移学习网络,利用局部最大均值差异(local maximum mean discrepancy, LMMD)来协调源域和目标域的子域分布,从而实现更有效的特征迁移。

尽管上述方法在一定程度上解决了领域自适应问题,但大多数方法只关注边缘分布对齐或者条件分布对齐,未能同时关注边缘分布和条件分布之间的差异,导致其在复杂工况下的性能受到限制。为了解决这一问题,潘晓博等<sup>[16]</sup>采用改进联合分布适应方法对源域和目标域特征集进行分布适应处理,降低域间分布差异。Qian等<sup>[17]</sup>提出一种深度判别迁移学习网络(deep discriminative transfer learning network, DDTLN),结合改进的联合分布自适应来增强领域混淆。然而,现有研究大多假设边缘分布和条件分布的同等重要,或手动调整其权重,这种做法在某些情况下可能不够灵活,从而降低了域适应性能。

针对上述问题,本文提出了一种融合多头自注意力机制和动态联合分布自适应(multi-head self-attention mechanism and dynamic joint distribution adaptation, MHSA-DJDA)的轴承故障诊断方法。首先,在特征提取模块中引入多头自注意力(multi-head self-attention, MHSA)机制,以便从原始振动信号中提取出不同故障模式下的代表性特征。其次,分别采用MMD与LMMD对边缘分布与条件分布进行特征对齐,减少源域与目标域之间的差异。最后,设计一种动态权重因子,根据实际工况自适应地调整边缘分布和条件分布的权重,从而进一步提升特征对齐的

效果。实验结果表明,该方法不仅获得了较高的准确率,还增强了在噪声环境下诊断的性能。为工业场景下的轴承故障诊断提供了一种有效的解决方案。

## 1 理论背景

### 1.1 无监督领域自适应

在无监督领域自适应中,假设给定一个源域数据集 $D_s = \{\mathbf{x}_i^s, y_i^s\}_{i=1}^{n_s}$ ,其中 $\mathbf{x}_i^s$ 是源域的第*i*个样本, $y_i^s$ 为其对应的标签,且 $n_s$ 为源域样本的总数。源域样本服从边缘概率分布 $P(\mathbf{x}_s)$ 和条件概率分布 $P(y_s | \mathbf{x}_s)$ 。同时,目标域数据集定义为 $D_t = \{\mathbf{x}_j^t\}_{j=1}^{n_t}$ ,其中 $\mathbf{x}_j^t$ 是目标域的第*j*个样本, $n_t$ 是目标域样本的总数。目标域样本服从边缘概率分布 $P(\mathbf{x}_t)$ 和条件概率分布 $P(y_t | \mathbf{x}_t)$ 。然而,源域和目标域之间存在差异,即: $P(\mathbf{x}_s) \neq P(\mathbf{x}_t)$ 和 $P(y_s | \mathbf{x}_s) \neq P(y_t | \mathbf{x}_t)$ 。

### 1.2 最大均值差异和局部最大均值差异

MMD作为一种非参数度量方法,已广泛应用于多个迁移学习算法中,用于衡量源域和目标域特征分布之间的差异。MMD的定义如下:

$$MMD(X_s, X_t) = \left\| \frac{1}{n_s} \sum_{i=1}^{n_s} \phi(\mathbf{x}_i^s) - \frac{1}{n_t} \sum_{j=1}^{n_t} \phi(\mathbf{x}_j^t) \right\|_{\mathfrak{H}}^2 \quad (1)$$

其中, $n_s$ 和 $n_t$ 分别表示源域和目标域的样本数量; $\phi(\cdot)$ 是表示从原始空间到再生核希尔伯特空间的映射, $\|\cdot\|_{\mathfrak{H}}^2$ 表示该空间的范数。为了简化计算,通常使用核函数代替特征映射函数,MMD的公式可以进一步简化为:

$$MMD^2(X_s, X_t) = \frac{1}{n_s^2} \sum_{i=1}^{n_s} \sum_{j=1}^{n_s} k(\mathbf{x}_i^s, \mathbf{x}_j^s) + \frac{1}{n_t^2} \sum_{i=1}^{n_t} \sum_{j=1}^{n_t} k(\mathbf{x}_i^t, \mathbf{x}_j^t) - \frac{2}{n_s n_t} \sum_{i=1}^{n_s} \sum_{j=1}^{n_t} k(\mathbf{x}_i^s, \mathbf{x}_j^t) \quad (2)$$

其中, $k(\mathbf{x}_i^s, \mathbf{x}_j^t)$ 是核函数,通常由多个高斯核函数的线性组合构成。

为了强调同一类别内源域和目标域之间的关系,LMMD<sup>[18]</sup>被提出,用于度量两个子域之间的差异。LMMD的定义如下:

$$d_{\mathfrak{H}(p,q)} \triangleq E_c \left\| E_{p^{(c)}} [\phi(\mathbf{x}^s)] - E_{q^{(c)}} [\phi(\mathbf{x}^t)] \right\|_{\mathfrak{H}}^2 \quad (3)$$

其中, $\mathbf{x}^s$ 和 $\mathbf{x}^t$ 分别表示源域 $D_s$ 和目标域 $D_t$ 中的样本, $p^{(c)}$ 和 $q^{(c)}$ 分别表示源域和目标域中属于类别 $c$ 的概率分布。每个样本的权重根据类计算。因此,LMMD可以进一步化简为:

$$\hat{d}_{\mathfrak{H}(p,q)} = \frac{1}{C} \sum_{c=1}^C \left\| \sum_{\mathbf{x}_i^s \in D_s} \omega_i^{sc} \phi(\mathbf{x}_i^s) - \sum_{\mathbf{x}_j^t \in D_t} \omega_j^{tc} \phi(\mathbf{x}_j^t) \right\|_{\mathfrak{H}}^2 \quad (4)$$

其中, $\omega_i^{sc}$ 和 $\omega_j^{tc}$ 表示属于类别 $c$ 的源域和目标域样本

的权重;  $\sum_{x_i \in D} \omega_i^c \phi(x)$  是类别  $c$  的加权和。权重的计算方法参考文献[18]。

目标域是无标签的,因此直接计算权重是不可行的。因此,利用深度学习模型预测的伪标签,计算权重。改写后的 LMMD 方程为:

$$\hat{d}_{l(p,q)} = \frac{1}{C} \sum_{c=1}^C \left[ \sum_{i=1}^{n_s} \sum_{j=1}^{n_s} \omega_i^c \omega_j^c k(z_i^{sl}, z_j^{sl}) + \sum_{i=1}^{n_t} \sum_{j=1}^{n_t} \omega_i^{tc} \omega_j^{tc} k(z_i^{tl}, z_j^{tl}) - 2 \sum_{i=1}^{n_t} \sum_{j=1}^{n_t} \omega_i^{sc} \omega_j^{tc} k(z_i^{sl}, z_j^{tl}) \right] \quad (5)$$

其中,  $z^l$  为第  $l$  层的激活。

## 2 基于多头自注意力和动态联合分布自适应的轴承故障诊断方法

### 2.1 基于多头自注意力机制的特征提取模块

自注意力机制能够有效捕捉时序数据中的长期依赖关系,进而增强特征表达。其数学表达式如下:

$$Q = F * W^Q, K = F * W^K, V = F * W^V \quad (6)$$

$$\text{Attention}(Q, K, V) = \text{soft max}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (7)$$

其中,  $Q, K$  和  $V$  分别表示查询矩阵、键矩阵和值矩阵。 $F \in \mathbb{R}^{L \times d_k}$  表示输入的振动数据经过卷积后的输出矩阵,其中  $L$  是输入序列的长度,  $d_k$  表示查询、键和值的特征维度。 $W^Q, W^K, W^V \in \mathbb{R}^{d_m \times d_k}$  是对应的权重矩阵,它们随着模型的训练不断更新。 $d_m$  表示输入的特征维度。

MHSA 是由多个并行执行的自注意力机制组成,每个头都有独立的权重矩阵,从而能够捕捉数据中的不同特征子空间的依赖关系。MHSA 的数学表达式如下:

$$\text{Head}_i = \text{Attention}(QW_i^Q, KW_i^K, VW_i^V), i = 1, 2, \dots, h \quad (8)$$

$$\text{MHSA}(Q, K, V) = \text{concat}(\text{head}_1, \text{head}_2, \dots, \text{head}_h)W^V \quad (9)$$

其中,  $W_i^Q, W_i^K, W_i^V \in \mathbb{R}^{d_m \times d_k}$  分别表示第  $i$  个头的查询、键和值的权重矩阵,  $h$  表示头的数量,  $W^V \in \mathbb{R}^{d_m \times d_m}$  是一个可训练的权重矩阵。特征提取网络如图 1 所示。在 MHSA 模块中,输入特征维度设定为 512。该维度足以提供丰富的信息表征能力,使模型能够捕捉到信号中的复杂模式和潜在关联。为了捕捉多样化特征的同时平衡计算复杂度,本文采用了 8 个注意力头。在完成多头自注意力机制的计算后,输出特征维度与输入维度保持一致,模块的输出直接作为后续卷积层的输入。基于对任务特性和模型结构的深入分析,在 MHSA 模块之后,本文选择不进行残差连接和层归一化。

### 2.2 改进的动态联合分布自适应

联合分布自适应的目标是减少源域和目标域之间的边缘分布和条件分布的差异。其目标函数可以定义为:

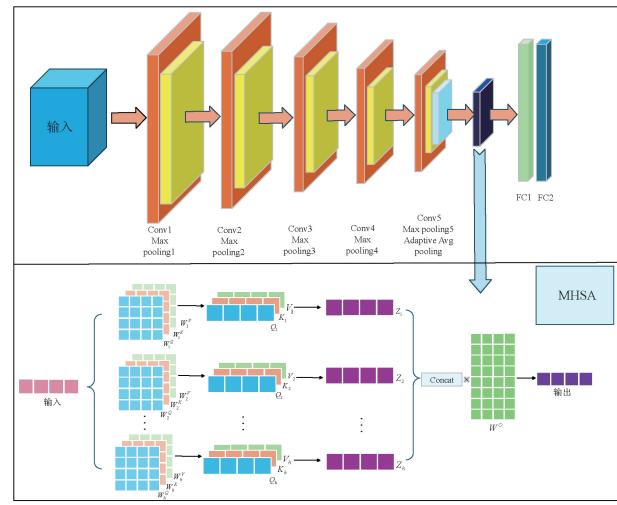


图 1 基于多头自注意力的特征提取模块

Fig. 1 Feature extraction module based on multi-head self-attention

$$L_{JDA} = \|E_{P(x_s)}[T(x_s)] - E_{P(x_t)}[T(x_t)]\|^2 + \sum_{c=1}^C \|E_{P(y_s=c|x_s)}[y_s=c | T(x_s)] - E_{P(y_t=c|x_t)}[y_t=c | T(x_t)]\|^2 \quad (10)$$

其中,  $T(\cdot)$  为图 1 中的特征提取器,由于目标域没有标签,无法直接计算条件概率分布  $P(y_t | x_t)$ 。因此,采用类条件概率分布  $P(x_t | y_t)$  近似条件概率分布  $P(y_t | x_t)$ 。根据贝叶斯定理,可以将其转化为类条件概率的形式:

$$P(y=c | x) = \frac{P(y=c)P(x | y=c)}{P(x)} \quad (11)$$

其中,类条件概率分布  $P(x | y=c)$  可以根据伪标签计算,类先验分布  $P(y=c)$  可按以下方式估算:

$$P(y_s=c) = \frac{n_s^c}{\sum_i^n n_s^i}, P(y_t=c) = \frac{n_t^c}{\sum_j^n n_t^j} \quad (12)$$

其中,  $c$  表示类别数,  $n_s^c$  和  $n_t^c$  分别表示源域和目标域中第  $c$  类的数量,同时  $\sum_c n_s^c = n_s$  和  $\sum_c n_t^c = n_t$ 。

因此,条件分布可以进一步表示为:

$$L_{CDA} = \sum_{c=1}^C \|E_{P(x_s|y_s=c)}[T(x_s) | y_s=c]P(y_s=c) - E_{P(x_t|y_t=c)}[T(x_t) | y_t=c]P(y_t=c)\|^2 \quad (13)$$

以往大多数联合分布自适应方法假设边缘分布和条件分布具有同等重要性。然而,当源域和目标域相似度较高时,条件分布对齐更为重要;而当两域差异较大时,边缘分布对齐变得更加关键。因此,提出了一个动态平衡因子  $\mu$ ,动态调整边缘分布和条件分布的权重。最终的动态联合分布机制可以定义为:

$$L_{DJDA} = \mu \|E_{P(x_s)}[T(x_s)] - E_{P(x_t)}[T(x_t)]\|^2 + (1-\mu) \sum_{c=1}^C \|E_{P(x_s|y_s=c)}[T(x_s) | y_s=c]P(y_s=c) - E_{P(x_t|y_t=c)}[T(x_t) | y_t=c]P(y_t=c)\|^2 \quad (14)$$

其中,  $\mu \in [0, 1]$ 。当  $\mu \rightarrow 0$  时, 表示源域和目标域数据集具有较高的相似性, 条件分布对齐更为重要; 当  $\mu \rightarrow 1$  时, 边缘分布对齐变得更加关键。

边缘分布对齐使用 MMD 计算, 同时为了更好的对齐同一类别在不同域中的子域, 并捕捉更多细粒度信息, 条件分布对齐使用 LMMD 来评估分布差异。最终的动态联合分布损失函数可以改写为:

$$\begin{aligned} L_{DJDA} &= \mu MMD(T(\mathbf{x}_s), T(\mathbf{x}_t)) + \\ &(1-\mu) \sum_{c=1}^C LMMD[(T(\mathbf{x}_s) | y_s = c) P(y_s = c), \\ &(T(\mathbf{x}_t) | y_t = c) P(y_t = c)] \end{aligned} \quad (15)$$

由于目标域是无标签的, 交叉验证方法将存在一定的局限性。因此, 本文使用边缘分布损失和条件分布损失来计算动态平衡因子  $\mu$ , 计算方法为:

$$\mu = \frac{loss_{MDA}}{loss_{MDA} + loss_{CDA}} \quad (16)$$

其中,  $loss_{MDA}$  是边缘分布损失,  $loss_{CDA}$  是条件分布损失。

### 2.3 总损失函数

为了更好地学习源域上的知识, 使用交叉熵损失函数。分类损失可以定义为:

$$L_c = L_s + \beta L_t \quad (17)$$

其中,  $L_s$  和  $L_t$  分别表示源域和目标域分类损失; 参数  $\beta$  为权衡参数。

根据改进的动态联合分布自适应, 整个网络的损失函数可以定义为:

$$L_{att} = L_c + \lambda L_{DJDA} \quad (18)$$

其中,  $L_{DJDA}$  为动态联合分布损失,  $\lambda$  是权衡因子。在整个训练过程中, 采用 Adam 优化算法进行优化。为了增强模型特征的可迁移性,  $\lambda$  的权衡因子设计为:

$$\lambda = \frac{2}{1 + \exp\left(\frac{-10epoch}{max\_epoch}\right)} - 1 \quad (19)$$

在训练初期,  $\lambda$  的值为 0, 有助于更好地学习基础的故障特征知识。随着训练次数的增加,  $\lambda$  的值单调增加, 从而逐渐增强动态联合分布损失, 更好地对齐源域和目标域的特征。

如图 2 所示, 网络由特征提取模块、基于 MMD 和 LMMD 的动态联合分布损失模块和分类模块构成。

## 3 实验和结果

### 3.1 实验数据

本文使用两个不同的滚动轴承数据集进行诊断实验。以下是数据集的详细信息。

1) 凯斯西储大学(Case Western Reserve University, CWRU)数据集<sup>[19]</sup>: 数据采集系统由三大部分组成: 电机、扭矩传感器和测功机。本文使用的数据来源于电机驱动端的振动信号, 采样频率为 48 kHz。数据集包含 4 种不同工况下的原始振动信号, 分别为 0 HP、1 HP、2 HP 和 3 HP。在每种工况下, 数据包括 4 种状态: 正常状态、内圈故障、外圈故障和滚动体故障。每种故障类型对应 3 个不同的故障程度, 分别为 0.007 inch、0.014 inch 和 0.021 inch。

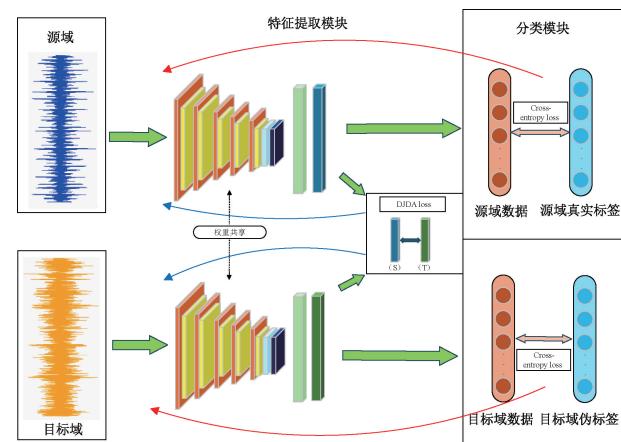


图 2 MHSAs-DJDA 框架

Fig. 2 MHSAs-DJDA framework

况下的原始振动信号, 分别为 0 HP、1 HP、2 HP 和 3 HP。在每种工况下, 数据包括 4 种状态: 正常状态、内圈故障、外圈故障和滚动体故障。每种故障类型对应 3 个不同的故障程度, 分别为 0.007 inch、0.014 inch 和 0.021 inch。

2) 江南大学数据集<sup>[20]</sup>: 该数据由江南大学在 Spectra Quest 转子实验平台上采集的, 包含 3 种故障类型: 内圈故障, 外圈故障滚动体故障。数据集包含 3 种不同转速振动信号, 转速分别为 600、800 和 1 000 rpm, 采样频率为 50 kHz。

### 3.2 实验细节

为了评估 MHSAs-DJDA 的性能, 将本文方法与几种知名的领域自适应方法进行了比较, 包括深度域混淆(deep domain confusion, DDC)<sup>[21]</sup>, DCORAL<sup>[13]</sup>, 深度子域自适应网络(deep subdomain adaptation network, DSAN)<sup>[18]</sup>, 域对抗神经网络(domain adversarial neural network, DANN)<sup>[22]</sup> 和 DDTLN<sup>[17]</sup>。

为验证所提出的 MHSAs-DJDA 方法, 每个迁移任务均重复执行 5 次。诊断性能通过准确率、精确率、召回率和 F1 分数等指标进行综合评估。

在每个故障诊断任务中, 每个类别的样本数均为 1 000, 实验中将每类样本按 7:3 的比例随机划分为训练集和测试集。由于实际采集的原始数据量有限, 本文采用滑动采样技术进行数据增强, 窗口长度设为 3 072, 步长为 256。为提升模型的训练效率和性能对输入数据进行标准化处理。

考虑到伪标签和真实标签可能存在差异, 将参数  $\beta$  设置为 0.1, 参数  $\mu$  根据计算的损失动态调整。学习率设置为 0.001, 迭代次数为 100, 批次大小设置为 128。

### 3.3 不同工况下的迁移实验

为了评估 MHSAs-DJDA 在不同工况下诊断轴承故障的性能, 本文在 CWRU 数据集上分别建立了数据集 A、B、C 和 D, 用于迁移学习实验。在江南大学轴承数据集上根

据不同的工况建立了数据集 E、F 和 G, 表 1 给出了两种数据集的详细信息。每个数据集包含 4 种故障类型: 正常状态(NC)、内圈故障(IF)、外圈故障(OF)和滚动体故障(BF)。例如,以 A → B 为例,A 表示为带标签的源域,B 则为未带标签的目标域。实验结果表 2 和表 3 所示。

在 CWRU 数据集上, MHSA-DJDA 在所有任务中的平均准确率为 99.84%, 优于其他方法。其他方法的平均准确率分别为 DDC(94.86%)、DCORAL(96.85%)、DSAN(98.85%)、DANN(92.61%) 和 DDTLN(99.53%)。这表明 MHSA-DJDA 能够有效地应对不同跨域任务, 通过动态平衡不同数据集的边缘分布和条件分布差异, 提升了诊断准确性。在标准偏差方面, MHSA-DJDA 的标准偏差

表 1 CWRU 数据集和江南大学数据集

Table 1 CWRU dataset and JNU dataset

数据集	工况	速度/	负载/	故障尺寸/	故障类型
		(r·min <sup>-1</sup> )	HP	inch	
CWRU	A	1 797	0	0.021	正常状态
	B	1 772	1	0.021	
	C	1 750	2	0.021	
	D	1 730	3	0.021	
JNU	E	600	—	—	内圈故障
	F	800	—	—	
	G	1 000	—	—	

表 2 在 CWRU 数据集上不同方法的准确率

Table 2 Accuracy of different methods on CWRU dataset

任务	模型性能/%					
	DDC	DCORAL	DSAN	DANN	DDTLN	MHSA-DJDA
A → B	93.90±4.19	98.26±1.63	98.95±1.80	98.43±2.26	99.78±0.26	99.82±0.32
A → C	94.65±3.09	94.91±1.15	96.32±2.78	92.40±3.77	98.28±0.33	99.75±0.26
A → D	96.05±1.72	95.91±0.82	96.13±3.02	78.38±4.26	99.34±0.30	99.68±0.26
B → A	86.23±3.39	97.30±5.45	99.32±1.51	98.61±1.80	99.93±0.11	99.87±0.30
B → C	95.42±4.90	96.15±1.30	99.53±0.35	95.35±3.03	99.24±0.64	99.88±0.22
B → D	98.68±0.75	95.48±4.18	99.57±0.44	87.98±4.39	99.43±0.18	99.83±0.10
C → A	90.82±6.49	98.32±3.31	99.93±0.07	97.22±3.01	99.51±0.35	99.78±0.49
C → B	96.72±2.57	98.05±3.90	99.72±0.55	97.12±3.04	99.90±0.18	99.95±0.11
C → D	94.01±3.98	97.18±4.86	99.81±0.16	98.38±3.35	99.81±0.13	99.85±0.29
D → A	96.67±3.31	96.22±5.21	97.95±3.72	91.01±4.55	99.79±0.21	99.88±0.26
D → B	98.81±1.00	97.01±3.00	99.57±0.57	76.72±2.59	99.89±0.11	99.82±0.41
D → C	96.35±3.86	97.40±2.02	99.38±0.70	99.68±0.33	99.42±0.74	99.93±0.11
平均值	94.86±3.27	96.85±3.07	98.85±1.31	92.61±3.03	99.53±0.29	99.84±0.26

表 3 在江南大学数据集上不同方法的准确率

Table 3 Accuracy of different methods on JNU dataset

任务	模型性能/%					
	DDC	DCORAL	DSAN	DANN	DDTLN	MHSA-DJDA
E → F	86.73±5.65	94.38±4.17	99.73±0.30	86.95±4.05	95.73±1.04	99.88±0.09
E → G	98.96±0.83	93.30±4.05	98.56±2.60	89.15±4.02	98.44±0.62	99.26±0.54
F → E	95.43±1.97	91.18±4.91	98.03±1.98	86.20±3.23	94.22±2.95	99.88±0.14
F → G	98.18±2.65	98.41±1.96	97.78±2.41	96.16±1.54	99.13±0.33	99.60±0.44
G → E	91.45±3.33	76.73±4.60	78.61±2.59	77.06±4.60	93.44±1.42	96.58±1.74
G → F	96.78±2.02	86.70±3.41	98.53±1.44	96.18±3.39	98.38±0.68	98.65±0.98
平均值	94.59±2.74	90.12±3.85	95.21±1.86	88.62±3.47	96.56±1.17	98.97±0.66

为 0.26%, 低于其他方法, 如 DDC(3.27%)、DCORAL(3.07%)、DSAN(1.31%)、DANN(3.03%) 和 DDTLN(0.29%)。低标准偏差表明该方法具有更好的稳定性, 能够在不同任务间保持一致的表现。

在江南大学数据集上, MHSA-DJDA 在所有任务中的平均准确率为 98.97%。其他方法的平均准确率分别为 DDC(94.59%)、DCORAL(90.12%)、DSAN(95.21%)、DANN(88.62%) 和 DDTLN(96.56%)。进一步证实,

MHSA-DJDA 在不同的迁移任务中提供了稳定且高效的故障诊断。

在稳定性方面, MHSA-DJDA 记录的平均标准偏差为 0.66%, 是所有方法中最低的, 这表明提出的模型不仅实现了高精度, 而且在一系列诊断任务中保持了一致的性能。

对比其他方法, 本文提出的 MHSA-DJDA 方法在精确率、召回率和 F1 分数指标上均表现出更优越的性能, 实验结果如表 4 所示。精确率为 98.46%, 召回率为 98.42%, F1 分数为 98.40%, 在所有对比方法中均为最高。这进一步验证了该方法在保持高准确率的同时, 能够稳定识别各类样本, 具备更强的鲁棒性与泛化能力。

表 4 不同模型的实验结果

Table 4 Experimental results of different models %

模型	精确率	召回率	F1
DDC	95.25	93.99	94.11
DCORAL	79.25	79.08	78.31
DSAN	85.37	84.67	84.46
DANN	84.28	83.07	82.27
DDTLN	95.75	95.566	95.51
MHSA-DJDA	98.46	98.42	98.40

为了进一步说明 MHSA-DJDA 方法在跨领域任务中对齐同一类样本的能力, 对  $G \rightarrow E$  进行特征可视化。利用 t 分布随机邻域嵌入 (t-distribution stochastic neighbor embedding, t-SNE) 对训练数据特征进行可视化。结果如图 3 所示, 每种颜色代表一个故障类别。

从图 3 中可以看出, MHSA-DJDA 方法在不同迁移任务中表现出最小的错误聚类数, 这表明本文的方法能够更清晰地将目标域中的样本进行区分。相比之下, 其他方法均有不少的错误分类。

DSAN 主要关注源域和目标域的子领域对齐, 忽略了边缘分布对齐。DDC 和 DCORAL 仅关注源域和目标域的边缘分布和二阶统计特性, 因此它们的对齐效果较差。DDTLN 通过同时对齐边缘分布和条件分布来提升模型性

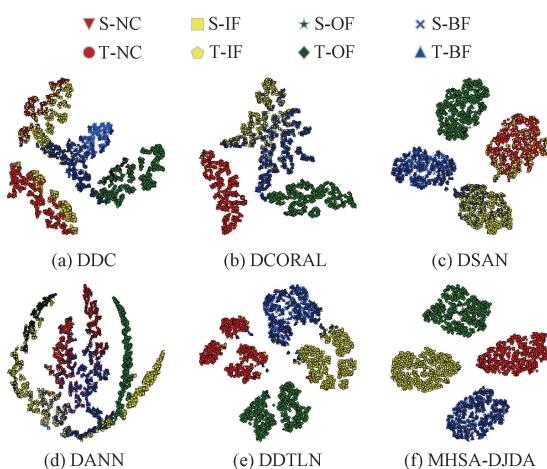


图 3  $G \rightarrow E$  任务 t-SNE 可视化

Fig. 3 t-SNE visualization of task  $G \rightarrow E$

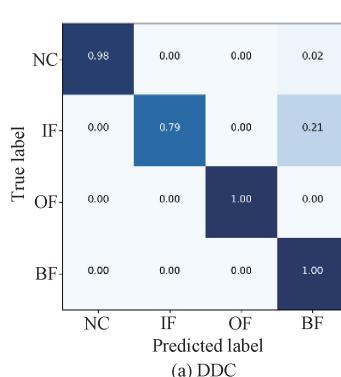
能, 尽管任务中也表现出了较好的性能。但在某些类别的边界处仍存在混叠现象, 导致对齐效果不理想。DANN 由于采用了对抗训练机制来混淆源域和目标域, 导致其对齐效果在所有方法中最差。

进一步, 通过混淆矩阵分析分类结果, 结果如图 4 所示。可以清楚地看出, MHSA-DJDA 方法的表现尤为突出, 能够正确地分类大部分故障样本, 而其他方法在分类任务中表现出不同程度的误分类, 且性能波动较大。

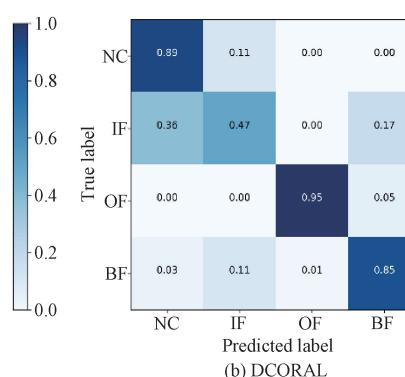
### 3.4 不同故障尺寸的迁移实验

为了进一步验证所提出方法的鲁棒性和泛化性, 在 CWRU 数据集中选择了不同故障尺寸的数据进行实验, 分别命名为 R0、R1 和 R2。数据集的详细信息如表 5 所示, 实验对比结果如表 6 所示。

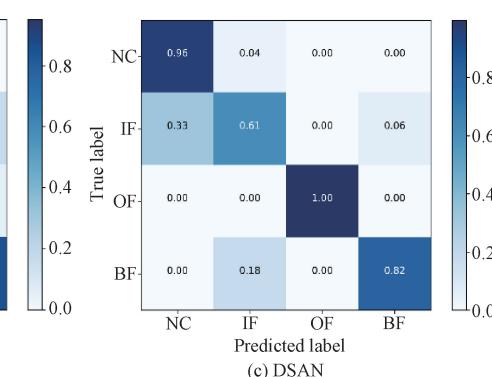
从结果可以看出, MHSA-DJDA 方法仍然取得了最高的平均准确率, 同时标准偏差也达到了最低值。相比之下, 基于对抗学习的方法 DANN 在实验中表现出最低的平均准确率, 说明其在不同故障直径下泛化性较差。DDC 和 DCORAL 虽然对齐源域和目标域的边缘分布来减小领域间的差异, 但由于未能充分考虑不同故障类型的特征差异, 在执行不同严重程度的轴承故障诊断时, 其准确性相



(a) DDC



(b) DCORAL



(c) DSAN

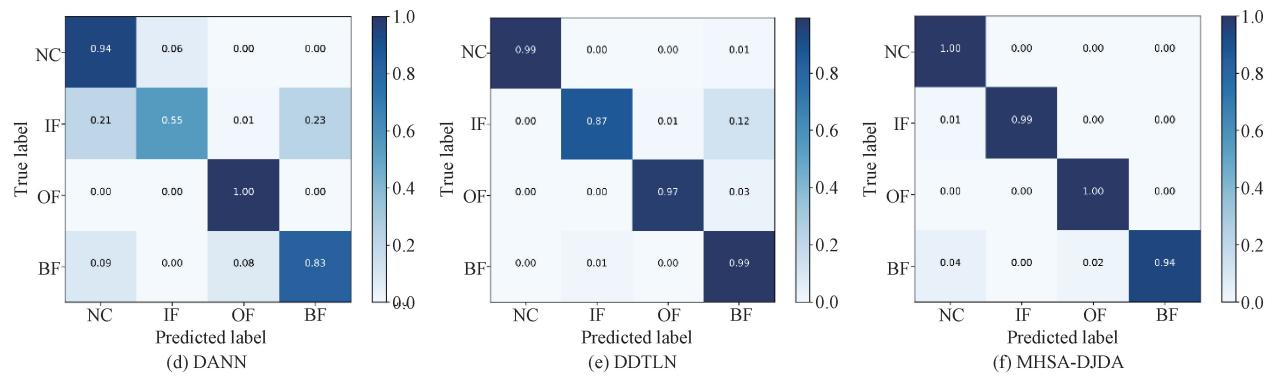


图 4 G → E 任务混淆矩阵

Fig. 4 Confusion Matrix of task G → E

表 5 CWRU 不同故障尺寸数据集

Table 5 CWRU different fault size dataset

数据集	工况	负载 /	故障尺寸 /	故障类型
		HP	inch	
CWRU	R0	2	0.007	正常状态
	R1	2	0.014	内圈故障
	R2	2	0.021	外圈故障 滚动体故障

较于其他方法较低。DSAN 未能关注边缘分布,因此其性能有所限制。DDTLN 方法虽然结合了边缘分布和条件分布,但由于缺乏动态平衡机制,其准确率相对较低,未能有效地提升模型的泛化性。相比于其他域适应方法,MHSA-DJDA 在不同故障直径下依然能够获得较好的跨域故障诊断能力,展现其出色的鲁棒性和泛化性能。

### 3.5 消融实验

为了进一步验证本文方法的性能,并评估各个模块的重要性,随机选取任务 F → E 和 R1 → R0 进行消融实验。

表 6 不同故障尺寸的对比实验结果

Table 6 Comparative experimental results for different fault size

任务	模型性能/%					
	DDC	DCORAL	DSAN	DANN	DDTLN	MHSA-DJDA
R0 → R1	77.88 ± 7.77	80.67 ± 2.16	78.58 ± 5.18	51.85 ± 3.70	62.10 ± 0.90	89.23 ± 1.42
R0 → R2	66.36 ± 3.27	81.45 ± 2.76	94.15 ± 10.3	69.97 ± 4.35	97.72 ± 1.40	99.57 ± 0.60
R1 → R0	94.25 ± 8.49	77.22 ± 7.17	83.07 ± 8.27	70.65 ± 9.39	62.38 ± 3.65	95.59 ± 1.43
R1 → R2	50.57 ± 0.60	91.83 ± 2.42	94.82 ± 2.48	63.17 ± 8.66	99.40 ± 0.47	98.80 ± 0.80
R2 → R0	71.17 ± 5.99	84.05 ± 8.07	97.08 ± 6.20	71.06 ± 5.70	98.57 ± 1.14	99.73 ± 0.55
R2 → R1	81.77 ± 5.87	67.58 ± 4.75	92.47 ± 5.78	55.85 ± 5.92	94.65 ± 2.62	99.07 ± 1.13
平均值	73.67 ± 5.33	80.47 ± 4.56	90.03 ± 6.37	63.76 ± 6.29	85.80 ± 1.70	96.99 ± 0.99

首先为了验证 MHSA 在特征提取中的有效性,设计了实验 DJDA,仅保留了动态联合分布自适应模块。接着,为了验证动态权重因子在平衡边缘分布和条件分布中的作用,设计了消融实验 MHSA-JDA。实验组合和结果如表 7 所示。

从实验结果可以观察到,本文所提出的方法在两个任务上分别达到了 99.88% 和 95.59% 的准确率,显著高于其他模型。这一结果进一步证明了本文方法可以实现有效的轴承故障诊断。

以任务 F → E 为例,各消融实验模型在训练过程中的总损失变化如图 5 所示。可以发现 MHSA-DJDA 的总损失下降速度显著快于其他模型,进一步验证了其在域适应

表 7 消融实验模型性能对比

Table 7 Comparison of ablation experiment model performance

模型	MHSA	动态因子 $\mu$	F → E 准确率/%	R1 → R0 准确率/%
JDA			94.20	91.78
DJDA		✓	95.08	93.82
MHSA-JDA	✓		95.65	93.40
MHSA-DJDA	✓	✓	99.88	95.59

任务中的优越性能。

由于滚动轴承时长运行在不同强度噪声的环境中,外

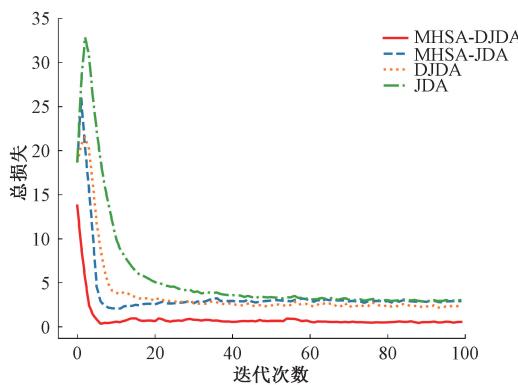


图 5 不同模型的总损失

Fig. 5 Total losses for different models

部噪声会影响模型的诊断效果。由于实验数据集不含噪声,因此,为了评估模型在噪声干扰下的稳定性和实用性,本文在目标域信号中人工添加不同信噪比(signal-to-noise ratio, SNR)的高斯白噪声,以模拟真实工业环境中的噪声扰动。信噪比计算公式为:

$$SNR_{dB} = 10\lg\left(\frac{P_{signal}}{P_{noise}}\right) \quad (20)$$

其中,  $P_{signal}$  表示信号功率,  $P_{noise}$  是噪声功率。信噪比的范围从 -4 dB~10 dB, 实验结果如图 6 所示。随着信噪比的不断增大,信号受到噪声的干扰越小,MHSAs-DJDA 方法的诊断准确率显著提升。即便在信噪比较低( $< 0$  dB)时,依然取得了较高的准确率,展现了极强的抗噪能力。相比之下,DDTLN 模型在低信噪比下表现不佳,但其准确率随信噪比的上升提升显著,表明其对环境噪声敏感,鲁棒性较差。

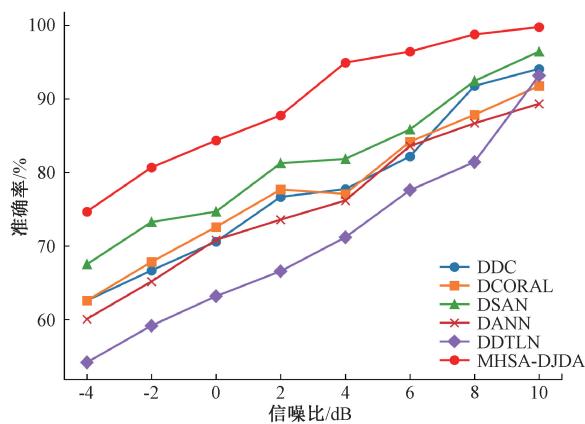


图 6 不同信噪比信号的实验结果

Fig. 6 Experimental results of signals with different SNRs

### 3.6 诊断训练时间对比分析

为全面评估各模型的计算复杂度,选取任务  $F \rightarrow E$ , 将 DDC、DCORAL、DSAN、DANN、DDTLN 以及本文提出的 MHSAs-DJDA 模型,在同一 GPU 环境下进行训练与推理,并从参数量、FLOPs、训练时间和推理时间 4 个方面进行

了对比,结果如表 8 所示。

表 8 不同模型的训练性能对比

Table 8 Comparison of training performance of different models

模型	参数/ ( $\times 10^6$ )	FLOPs/ ( $\times 10^6$ )	训练 时间/s	推理 时间/ms
DDC	3.41	164.29	125.88	121.97
DCORAL	3.40	162.82	110.73	124.77
DSAN	3.97	525.07	475.85	148.39
DANN	0.66	101.74	82.53	133.17
DDTLN	0.79	9.66	399.43	127.66
MHSAs-DJDA	1.71	10.57	328.38	128.51

从表 8 中可以看出, MHSAs-DJDA 在引入 MHSAs 以提升特征表达能力的同时,仍有效控制了模型的参数规模与 FLOPs 数量, 显著低于 DSAN 等复杂模型。推理时间方面, MHSAs-DJDA 保持了与其他方法相近的响应速度(128.51 ms), 仅略高于参数更少的 DANN(133.17 ms)与 DDTLN(127.66 ms)。虽然其训练时间(328.38 s)略高于部分轻量模型,但在保持较低计算开销的同时获得了更强的模型性能和更高的准确率,充分验证了模型的实用性和高效性。

## 4 结论

本文提出了一种基于多头自注意力和动态联合分布自适应的轴承故障诊断方法。与传统的域适应方法相比,本文方法在跨领域故障诊断中展现出更强的鲁棒性和泛化性。首先,本文设计了一种基于多头自注意力机制的特征提取模块,能够有效捕捉复杂振动信号中的动态特征,显著增强了模型对不同故障模式的诊断能力。该模块通过并行处理多个特征子空间,提升了特征提取的效率和准确性。其次,本文通过 MMD 和 LMMD 分别对齐源域和目标域的边缘分布与条件分布,并设计了一种新的动态权重因子,能够根据具体工况自适应地调整边缘分布和条件分布的重要性,从而有效减少领域之间的分布差异。最后,通过在数据集上的对比实验,验证了所提方法的有效性和优越性。尽管所提出的方法取得了令人满意的结果,但本文仍存在一定的局限性。本文假设源域和目标域具有相同的故障数据。然而,在实际工业生产中,故障数据通常相对稀缺且难以获取。因此,未来工作需要针对目标域样本不平衡的问题进行更深入的研究。

## 参考文献

- [1] 马萍,梁城,王聪,等. 小样本下基于原型域增强的 Meta-DAE 故障诊断[J]. 华南理工大学学报(自然科学版),2025,53(1):62-73.  
MA P, LIANG CH, WANG C, et al. Meta-DAE

- fault diagnosis based on prototype domain enhancement in few-shot[J]. Journal of South China University of Technology (Natural Science Edition), 2025, 53(1): 62-73.
- [2] 车鲁阳,高军伟,付惠琛.基于多通道融合的滚动轴承剩余寿命预测[J].电子测量与仪器学报,2023,37(12):225-233.
- CHE L Y, GAO J W, FU H CH. Residual life prediction of rolling bearings based on multi-feature fusion[J]. Journal of Electronic Measurement and Instrumentation, 2023, 37(12): 225-233.
- [3] 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.
- [4] BAI Y H, CHENG W D, WEN W G, et al. Application of time-frequency analysis in rotating machinery fault diagnosis[J]. Shock and Vibration, 2023, 2023(1): 9878228.
- JIA B H, SU J CH, GAO Y. Gearbox fault diagnosis across different operating conditions based on improved domain-versarial network [J]. Electronic Measurement Technology, 2025, 48(3): 83-91.
- [5] 贾宝惠,苏家成,高源.基于改进域对抗网络的齿轮箱跨工况故障诊断[J].电子测量技术,2025,48(3): 83-91.
- JIA B H, SU J CH, GAO Y. Gearbox fault diagnosis across different operating conditions based on improved domain-versarial network [J]. Electronic Measurement Technology, 2025, 48(3): 83-91.
- [6] ZHANG W W, CHEN D J, XIAO Y. A novel fault diagnosis method based on semisupervised contrast learning [C]. IEEE International Conference on Prognostics and Health Management (ICPHM). IEEE, 2022: 82-87.
- XUE Y P, WEN CH B, WANG Z D, et al. A novel framework for motor bearing fault diagnosis based on multi-transformation domain and multi-source data [J]. Knowledge-Based Systems, 2024, 283: 111205.
- [7] 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.
- 范永胜,丁雪,邓艾东.基于深度条件子域自适应网络的轴承跨域故障诊断研究[J].动力工程学报,2024,44(1):62-67,83.
- FAN Y SH, DING X, DENG AI D. Research on cross-domain fault diagnosis of bearings based on deep conditional subdomain adaptive network[J]. Journal of Power Engineering, 2024, 44(1): 62-67,83.
- [10] 胡春生,李国利,赵勇,等.变工况滚动轴承故障诊断方法综述[J].计算机工程与应用,2022,58(18):26-42.
- HU CH SH, LI G L, ZHAO Y, et al. Summary of fault diagnosis methods for rolling bearings under variable working conditions[J]. Computer Engineering and Applications, 2022, 58(18): 26-42.
- [11] 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.
- [12] 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 bearings based on multi-scale graph Transformer[J]. Foreign Electronic Measurement Technology, 2023, 42(12): 186-194.
- [13] SUN B CH, SAENKO K. Deep coral: Correlation alignment for deep domain adaptation[C]. Computer Vision-ECCV 2016 Workshops, 2016: 443-450.
- [14] CAO H R, SHAO H D, ZHONG X, et al. Unsupervised domain-share CNN for machine fault transfer diagnosis from steady speeds to time-varying speeds[J]. Journal of Manufacturing Systems, 2022, 62: 186-198.
- [15] WANG J J, YANG SH P, LIU Y Q, et al. Deep subdomain transfer learning with spatial attention ConvLSTM network for fault diagnosis of wheelset bearing in high-speed trains[J]. Machines, 2023, 11(2): 304.
- [16] 潘晓博,葛鲲鹏,钱孟浩,等.基于改进联合分布适应的轴承智能故障诊断方法[J].机电工程,2023,40(9): 1354-1362.
- PAN X B, GE K P, QIAN M H, et al. Bearing intelligent fault diagnosis based on improved joint distribution adaptation[J]. Journal of Mechanical & Electrical Engineering, 2023, 40(9): 1354-1362.
- [17] QIAN Q, QIN Y, LUO J, et al. Deep discriminative transfer learning network for cross-machine fault diagnosis [J]. Mechanical Systems and Signal Processing, 2023, 186: 109884.
- [18] ZHU Y CH, ZHUANG F ZH, WANG J D, et al. Deep subdomain adaptation network for image classification [J]. IEEE Transactions on Neural Networks and Learning Systems, 2020, 32 (4): 1713-1722.
- [19] 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.
- [20] ZHAO ZH B, LI T F, WU J Y, et al. Deep learning algorithms for rotating machinery intelligent diagnosis: An open source benchmark study[J]. ISA Transactions, 2020, 107: 224-255.
- [21] 朱继扬,孙虎儿,张天源,等.基于多表示动态自适应的不同工况下滚动轴承故障诊断[J].机电工程,2023,40(2):178-185,203.
- ZHU J Y, SUN H ER, ZHANG T Y, et al. Rolling bearing fault diagnosis under different working conditions based on MRDA[J]. Journal of Mechanical and Engineering, 2023, 40(2): 178-185,203.
- [22] GANIN Y, LEMPITSKY V. Unsupervised domain adaptation by backpropagation [C]. Proceedings of Machine Learning Research, 2015, 37: 1180-1189.

### 作者简介

李杰,硕士研究生,主要研究方向为故障诊断。

E-mail:23610001030604@st.sjtu.edu.cn

刘天羽(通信作者),博士,教授,主要研究方向为控制理论与控制工程、智能控制、模式识别和风力发电机组故障诊断。

E-mail:liuty@sjtu.edu.cn