

# 控制与决策

Control and Decision

## 基于可信多源领域自适应的跨域滚动轴承故障诊断

周旷, 覃文博, 孙天宇

引用本文:

周旷, 覃文博, 孙天宇. 基于可信多源领域自适应的跨域滚动轴承故障诊断[J]. *控制与决策*, 2025, 40(7): 2251-2260.

在线阅读 View online: <https://doi.org/10.13195/j.kzyjc.2024.1430>

---

## 您可能感兴趣的其他文章

### Articles you may be interested in

#### 基于度量学习和典型相关分析的亲缘关系识别网络

Kinship relationship recognition network based on metric learning and canonical correlation analysis

控制与决策. 2021, 36(8): 1977-1983 <https://doi.org/10.13195/j.kzyjc.2019.1798>

#### 基于数据分布特性的代价敏感宽度学习系统

[Data distribution-based cost-sensitive broad learning system](#)

控制与决策. 2021, 36(7): 1686-1692 <https://doi.org/10.13195/j.kzyjc.2019.1484>

#### 基于DLSR的归纳式迁移学习

DLSR based inductive transfer learning method

控制与决策. 2021, 36(12): 2982-2990 <https://doi.org/10.13195/j.kzyjc.2020.0703>

#### $l_p$ -范数约束下MKL-OC-ELM的装备故障检测

MKL and OC-ELM fault detection based on  $l_p$ -norm constraint

控制与决策. 2021, 36(10): 2379-2388 <https://doi.org/10.13195/j.kzyjc.2020.0443>

#### 改进集成深层自编码器在轴承故障诊断中的应用

Application of improved ensemble deep auto-encoder in bearing fault diagnosis

控制与决策. 2021, 36(1): 135-142 <https://doi.org/10.13195/j.kzyjc.2019.0270>

# 基于可信多源领域自适应的跨域滚动轴承故障诊断

周旷<sup>1†</sup>, 覃文博<sup>1</sup>, 孙天宇<sup>2</sup>

(1. 西北工业大学 数学与统计学院, 西安 710129; 2. 西安现代控制技术研究, 西安 710065)

**摘要:** 针对现有多源领域自适应故障诊断方法未充分考虑不同源域信息质量差异的问题, 提出一种基于可信多源领域自适应的跨域故障诊断模型 (TMDAFD). 首先, 构建源域与目标域的最优公共嵌入空间, 以实现特征有效对齐; 然后, 基于证据深度学习理论, 设计具有不确定度量的领域自适应模型, 量化不同源域支持下目标域诊断结果的不确定性, 为多源决策融合提供可信依据; 最后, 结合所得到的不确定度指标设计信息融合策略, 有效整合多源决策信息, 进一步提升跨域故障诊断模型的准确性. 在凯斯西储大学 (CWRU)、江南大学 (JNU) 和渥太华大学 (UO) 三个滚动轴承故障数据集上进行实验验证. 实验结果表明, TMDAFD 模型显著提升了跨域故障诊断任务的性能, 验证了其在复杂多源场景下的有效性.

**关键词:** 多源领域自适应; 迁移学习; 不确定度量; 故障诊断; 滚动轴承; 证据深度学习

中图分类号: TP391 文献标志码: A

DOI: 10.13195/j.kzyjc.2024.1430

引用格式: 周旷, 覃文博, 孙天宇. 基于可信多源领域自适应的跨域滚动轴承故障诊断 [J]. 控制与决策, 2025, 40(7): 2251-2260.

## Cross-domain fault diagnosis of rolling bearings based on trusted multi-source domain adaptation

ZHOU Kuang<sup>1†</sup>, QIN Wen-bo<sup>1</sup>, SUN Tian-yu<sup>2</sup>

(1. School of Mathematics and Statistics, Northwestern Polytechnical University, Xi'an 710129, China; 2. Xi'an Modern Control Technology Research Institute, Xi'an 710165, China)

**Abstract:** To address the limitations of existing multi-source domain adaptation (MSDA) methods for fault diagnosis, which fail to adequately account for variations in source domain data quality, this paper proposes a trustworthy multi-source domain adaptation-based fault diagnosis (TMDAFD) model. First, an optimal common embedding space is constructed between the source and target domains to achieve effective feature alignment. Second, based on evidential deep learning theory, a domain adaptation model with uncertainty quantification is designed to measure the reliability of diagnostic results in the target domain under different source domain supports, providing a trustworthy basis for multi-source decision fusion. Finally, an information fusion strategy is designed based on the obtained uncertainty measure, which effectively integrates decision information from multiple source domains and consequently enhances the accuracy of cross-domain diagnosis. Experimental evaluations on rolling bearing fault datasets from Case Western Reserve University (CWRU), Jiangnan University (JNU), and the University of Ottawa (UO) demonstrate that the TMDAFD model significantly enhances the performance of cross-domain fault diagnosis tasks, validating its effectiveness in complex multi-source scenarios.

**Keywords:** multi-source domain adaptation; transfer learning; uncertainty quantification; fault diagnosis; rolling bearings; evidential deep learning

## 0 引言

滚动轴承是旋转机械设备的核心支撑部件, 其受恶劣工作环境和重负荷、连续运转等工作状态影响, 极易出现故障<sup>[1]</sup>. 因此, 对滚动轴承的健康状况进

行诊断, 是保障设备安全运行的重要环节. 近年来, 随着机器学习和人工智能技术的迅速发展, 基于深度学习的故障诊断方法已得到了广泛应用<sup>[2-3]</sup>. 然而, 这些方法的性能高度依赖于充足的训练样本. 在现

收稿日期: 2024-12-09; 录用日期: 2025-04-08.

基金项目: 国家社会科学基金项目 (23BTJ053); 航空科学基金项目 (20182053023); 国家自然科学基金项目 (92371101, 61701409).

†通信作者. E-mail: kzhoumath@nwpu.edu.cn.

实应用场景中, 受限于时间和成本因素, 获取足够的训练样本通常十分困难, 这严重制约了基于深度学习的故障诊断方法的性能。

领域自适应 (DA) 作为一种重要的迁移学习方法, 是解决小样本问题的有效途径。其核心思想是通过利用相似领域 (源域) 的历史标注数据, 辅助当前领域 (目标域) 的预测任务, 以提升小样本数据学习模型的性能<sup>[4-5]</sup>。近些年, 基于领域自适应的故障诊断方法得到了长足的发展。李晨昀等<sup>[6]</sup> 使用了局部最大均值差异进行特征映射, 对齐每个源域与目标域的子领域; Chen 等<sup>[7]</sup> 提出了同时考虑边缘分布和条件分布的方法, 有效减少了不同域间的分布差异; Yin 等<sup>[8]</sup> 提出了协调产出最大均值差异, 能够有效减少源域与目标域间的分布差异, 简化了对齐过程, 显著提升了故障诊断的精度和迁移学习模型的性能; Chi 等<sup>[9]</sup> 提出了基于对抗迁移学习的异常检测算法, 将鲁棒自适应约束集成到改进的领域对抗神经网络中; 文利燕等<sup>[10]</sup> 提出了利用对抗域自适应训练来提取源域和目标域的深层不变特征, 并通过最小化源域与目标域间的类原型距离, 实现域与类的联合对齐。

上述基于领域自适应的故障诊断方法主要是使用单一源域数据来辅助目标域的诊断任务, 属于单源领域自适应 (SDA) 方法。实际中, 可能会存在多个可用于辅助诊断目标域样本健康状况的源域数据。相比于单一源域数据, 多源域数据蕴含了更为丰富的可迁移信息<sup>[11]</sup>。然而, 上述方法不能直接处理多源域数据, 难以对多源信息有效融合利用, 限制了跨域迁移学习模型的性能。虽然可通过改进单源领域自适应中的特征学习策略, 构建所有源域与目标域的最优嵌入空间来实现跨域特征对齐, 但是, 其关键在于缺乏对不同源域数据质量差异的评估机制, 在诊断决策中将所有源域平等对待, 限制了跨域迁移学习模型的性能。

鉴于此, 本文提出基于可信多源领域自适应的故障诊断模型 (TMDAFD), 用于解决跨域滚动轴承数据下故障诊断问题。证据深度学习 (EDL) 利用主观逻辑 (SL) 理论, 能够实现深度学习决策的不确定度量<sup>[12]</sup>, 在各领域得到了广泛应用。本文将 EDL 应用于多源领域自适应网络, 构建具有决策不确定度量的损失函数, 能够利用不同源域数据训练得到关于目标域样本的诊断决策, 并给出决策可信度, 进而设计多源决策融合策略, 从而获得更加可靠的诊断结果。相比于单源领域自适应方法, TMDAFD 模型能够利用多个领域知识, 丰富训练特征; 相比于现

有多源领域自适应方法, TMDAFD 模型能够通过不确定度量刻画源域为目标域提供迁移信息的质量差异, 以实现多源信息的有效融合利用。本文的主要内容如下。

1) 针对滚动轴承跨域故障诊断中多源信息难以有效融合利用的问题, 提出一种可信多源领域自适应方法, 建立不确定度驱动的多源数据融合决策方法, 能够提升故障诊断模型准确率。

2) 面向跨域故障诊断领域自适应深度网络, 构建包含决策不确定度量的损失函数, 刻画源域对目标域决策的可信度, 为实现在诊断决策中对不同源域信息的差异化利用提供依据。

3) 在实际滚动轴承故障数据集中进行实验验证, 实验结果表明, 所提出 TMDAFD 模型相较于对比方法, 诊断准确率可得到显著提升。消融实验的结果可进一步验证所设计决策不确定度量化以及融合策略的有效性。

## 1 国内外研究现状

本文提出一种具有不确定度量的多源领域自适应跨域故障诊断模型。下面分别对当前国内外对于基于多源领域自适应的故障诊断方法和深度学习不确定度量方法的研究进展进行讨论和分析。

### 1.1 基于多源领域自适应的故障诊断方法

基于深度学习的故障诊断方法通过提取设备的状态监测信号或历史数据中的特征信息, 对样本的健康状况进行预测, 具有更强的通用性和更广泛的适用范围。然而, 在实际应用中, 可能会存在历史故障数据较少、可利用的标签数据不多等问题。此时, 直接利用小样本数据进行训练, 会使得模型无法学到足够多的特征, 进而影响模型的诊断性能。领域自适应方法可以解决小样本数据在训练时存在的局限性, 其通过将相关领域的数据或知识迁移至目标领域, 能够有效解决目标领域样本缺乏标签等问题。根据源域数量的多少, 可将领域自适应方法分为单源领域自适应和多源领域自适应。

在现实生活中, 与目标领域相似的源域的数量往往不止一个。因此, 基于多源领域自适应的故障诊断方法得到了广泛关注。这些方法将源域和目标域映射至潜空间, 在潜空间中对齐源域与目标域的分布。根据对齐方式的不同, 可将这些方法分为基于分布距离度量的方法和基于对抗的方法。

在基于分布距离度量的方法中: Tian 等<sup>[13]</sup> 利用多分支网络结构匹配每个源域与目标域的特征空间分布, 利用 MMD (maximum mean discrepancy) 距离

将细粒度局部对齐, 实现了设备状态监测; Wen 等<sup>[14]</sup>通过提出一种在潜在特征空间中对齐源域与目标域分布并最小化互信息的方法, 将所有源域合并, 并和带有伪标签的目标域样本一起训练, 解决了多源领域自适应中联合分布差异的问题; Liang 等<sup>[15]</sup>提出用于提取丰富迁移特征的多源领域适配网络, 并从多个源域中获取互补信息, 模型通过采用伪边界向量处理目标域中的未知故障, 通过向传统最大均值差异中添加权重, 设计了新的损失函数, 使得公共标签集更加兼容, 实现了滚动轴承的精确故障诊断。

在基于对抗的方法中: Liu 等<sup>[16]</sup>提出了深度多源对抗性差异匹配适应网络, 利用多个领域学到的泛化知识诊断未知任务, 并用对抗方法进行训练, 提高了跨域诊断的准确性; Shi 等<sup>[17]</sup>提出了对抗性多源数据子域适应模型, 该模型封装了共享特征提取器、标签预测器和域鉴别器, 通过对抗的方法获得了多源异构数据的域不变特征; Gao 等<sup>[18]</sup>利用对抗性迁移网络进行了特征学习和知识迁移, 并通过联合分布域适应机制捕捉域不变特征, 最后, 通过类相关的决策融合策略构建了决策空间。然而, 不同源域与目标域间往往存在显著分布差异, 导致各源域向目标域迁移的知识质量参差不齐: 与目标域分布更接近的源域能够提供更可靠的迁移信息, 而分布差异较大的源域可能会引入噪声干扰。上述方法将各源域信息平等对待, 既无法实现迁移知识的动态择优, 又难以抑制低质量迁移源的负面影响, 限制了基于多源信息的故障诊断模型的性能。

## 1.2 深度学习不确定度量方法

迁移学习可有效利用已有领域的知识, 解决深度学习由于目标领域样本缺少标注数据导致模型训练困难的问题。然而, 在现实中, 源域与目标域样本具有分布差异, 决策结果存在不确定性。若缺少对决策结果的不确定度量, 无法体现源域对目标域决策的可信程度, 进而降低了迁移学习模型性能。

目前, 已有面向故障诊断的深度学习不确定度量方法。Xu 等<sup>[19]</sup>提出了一种基于主观逻辑的分层系统诊断融合方法, 能够处理在故障诊断问题中由于缺乏证据带来的不确定性的问题; Zhou 等<sup>[20]</sup>开发了一种集成证据理论的新型卷积神经网络, 以实现具有预测不确定性估计的故障分类; Xiao 等<sup>[21]</sup>利用贝叶斯变分学习, 将不确定性引入 Transformer 的注意力权重中, 构建了可信赖的贝叶斯 Transformer 模型, 增强了模型的可解释性和可信度; Zhou 等<sup>[22]</sup>提出了一种新的概率贝叶斯卷积神经网络来量化预测不确

定性, 并设计了可信赖的故障诊断框架。上述方法主要是在利用单一领域数据进行训练, 在多源领域自适应故障诊断领域中, 由于源域与目标域存在分布差异, 导致源域为目标域提供的故障诊断结果的不确定程度各异, 这便需要考虑不同源域下决策结果的可信度, 从而实现对多源信息更为有效的融合利用。

目前, 已有对多源领域自适应不确定度量方法的讨论。Zhang 等<sup>[23]</sup>提出了多源主动领域迁移学习方法, 利用证据深度学习度量领域不确定性和预测不确定性, 能够自适应学习适用于多个源域的模型参数; Pei 等<sup>[24]</sup>在多源领域自适应中设计了证据预测不确定性和证据邻接一致性不确定性度量, 量化了目标样本在源域模型上的不确定性, 从而确定目标样本对于不同源模型的偏好。

本文提出一种具有不确定度量的多源领域自适应故障诊断模型 TMDAFD。该模型基于证据深度学习方法, 度量由于分布差异造成的诊断结果的不确定性, 刻画不同源域支持下目标域决策的可信度, 并据此设计多源决策融合准则, 提升多源故障诊断结果的可靠性和准确性。

## 2 基于可信多源领域自适应的故障诊断

假设有  $R$  个源域  $\mathcal{D}_s^r = \{\mathbf{X}_s^r, \mathbf{Y}_s^r\}$  和一个目标域  $\mathcal{D}_t = \{\mathbf{X}_t\}$ 。其中:  $\mathbf{X}_s^r \in \mathbb{R}^{N_s^r \times d}$  ( $r = 1, 2, \dots, R$ ) 和  $\mathbf{X}_t \in \mathbb{R}^{N_t \times d}$  分别为第  $r$  个源域和目标域样本集合,  $\mathbf{Y}_s^r \in \mathbb{R}^{N_s^r \times 1}$  为第  $r$  个源域的标签集合,  $N_s^r$  和  $N_t$  分别为第  $r$  个源域和目标域样本的数量,  $d$  为样本维度。TMDAFD 模型主要包括具有不确定度量的多源领域自适应网络和多源决策融合两部分, 整体框架如图 1 所示。下面对两个模块分别进行详细介绍。

### 2.1 具有不确定度量的多源领域自适应网络框架

具有不确定度量的多源领域自适应网络框架包括特征构建、故障分类器和不确定度量 3 部分。利用第  $r$  个源域训练的分类器的目标函数如下所示:

$$\mathcal{L}^r = \lambda_1 \mathcal{L}_{\text{MMD}} + \sum_{\mathbf{x} \in \mathcal{D}_s^r} [\mathcal{L}_{\text{Cls}}^r(\mathbf{x}) + \lambda_2 \mathcal{L}_{\text{KL}}^r(\mathbf{x})], \quad (1)$$

其中  $\lambda_1$  和  $\lambda_2$  分别为 MMD 距离损失  $\mathcal{L}_{\text{MMD}}$  和 KL (Kullback-Leibler) 散度正则化项  $\mathcal{L}_{\text{KL}}^r$  的平衡参数。目标函数 (1) 中的第 1 项为 MMD 距离损失, 该项是迁移学习中常见的分布距离度量函数, 通过最小化每个源域  $\mathcal{D}_s^r$  与目标域  $\mathcal{D}_t$  分布间的 MMD 距离, 使得源域和目标域的特征分布在嵌入空间中足够接近, 进而提高模型的训练效果; 目标函数 (1) 中的第 2 项损失  $\mathcal{L}_{\text{Cls}}^r$  为具有不确定度量的交叉熵分类损失, 通过将

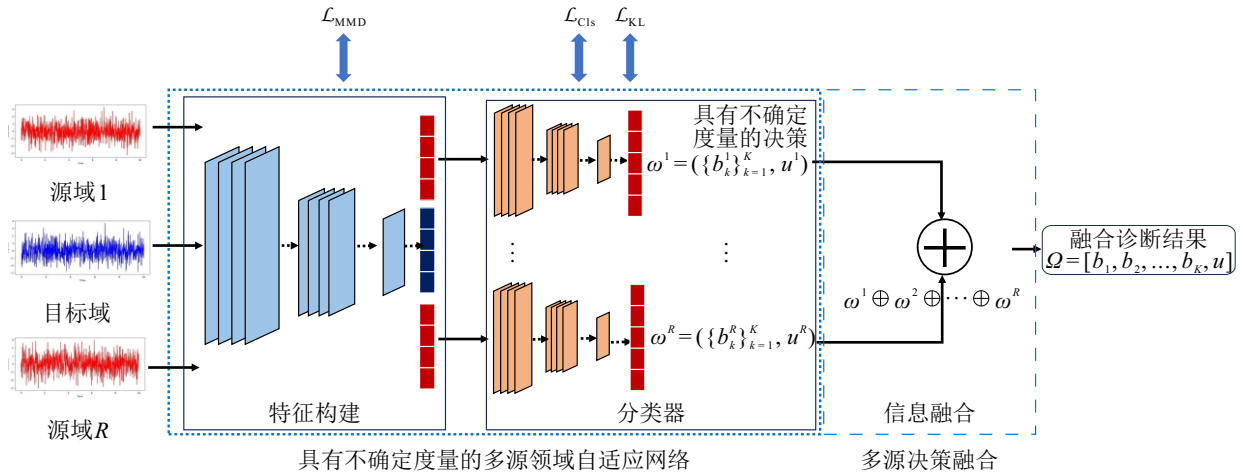


图1 TMDAFD 模型框架

交叉熵分类损失与 Dirichlet 分布相结合, 组成贝叶斯风险损失, 能够计算样本在不确定下的分类损失; 目标函数 (1) 中的第 3 项为正则化项  $\mathcal{L}_{\text{KL}}^r$ , 其能够通过惩罚来降低非正确类别的支持度, 防止模型错误预测. 下面对 TMDAFD 目标函数的每一项进行详细介绍.

### 2.1.1 MMD 距离损失

MMD 距离是迁移学习中常见的度量源域与目标域分布差异的方法. 在损失函数中加入 MMD 距离, 可实现源域与目标域分布的对齐. 首先, 将源域和目标域数据通过一维 CNN 网络, 提取更为丰富的故障特征<sup>[25]</sup>. 将 CNN 提取的特征表示为  $f(\cdot)$ . 然后, 利用核函数  $\phi(\cdot)$ , 将特征映射至无穷维再生希尔伯特空间  $\mathcal{H}$ , 并计算两个特征在  $\mathcal{H}$  中的 MMD 距离. MMD 距离损失为

$$\mathcal{L}_{\text{MMD}} = \sum_{r=1}^R \left\| \frac{\sum_{\mathbf{x}_i \in \mathcal{D}_s^r} \phi(f(\mathbf{x}_i))}{|\mathcal{D}_s^r|} - \frac{\sum_{\mathbf{x}_j \in \mathcal{D}_t} \phi(f(\mathbf{x}_j))}{|\mathcal{D}_t|} \right\|_{\mathcal{H}}^2. \quad (2)$$

$\phi(\cdot)$  将自变量映射至无穷维空间, 式 (2) 无法显式求得, 可将其展开, 变为求取两个向量的内积. 展开后得到 MMD 距离公式为

$$\mathcal{L}_{\text{MMD}} = \sum_{r=1}^R \text{tr} \left( \begin{bmatrix} \phi_{s,s} & \phi_{s,t} \\ \phi_{t,s} & \phi_{t,t} \end{bmatrix} \mathbf{M}^r \right). \quad (3)$$

其中:  $\text{tr}$  为矩阵的迹.  $\mathbf{M}^r$  各元素具体定义为

$$M_{ij}^r = \begin{cases} \frac{1}{|\mathcal{D}_s^r|^2}, & \mathbf{x}_i, \mathbf{x}_j \in \mathcal{D}_s^r; \\ \frac{1}{|\mathcal{D}_t|^2}, & \mathbf{x}_i, \mathbf{x}_j \in \mathcal{D}_t; \\ -\frac{1}{|\mathcal{D}_s^r||\mathcal{D}_t|}, & \text{otherwise.} \end{cases} \quad (4)$$

高斯核函数将数据映射至无穷维空间, 能够很好地捕捉数据中的局部差异和非线性结构, 因此, 一

般采用高斯核函数映射数据, 具体公式如下所示:

$$\phi(\mathbf{x}_i, \mathbf{x}_j) = \exp \left( -\frac{\|\mathbf{x}_i - \mathbf{x}_j\|^2}{2\sigma^2} \right). \quad (5)$$

### 2.1.2 具有不确定度量的交叉熵损失

在提取源域和目标域特征后, 需要输入至分类器中进行训练. 采用交叉熵损失来衡量分类器的预测类别与实际类别间的差异. 假设共有  $K$  个类别, 第  $r$  个源域的交叉熵损失函数公式为

$$\mathcal{L}_{\text{Cross}}^r(\mathbf{x}) = \sum_{k=1}^K -y_k^r \log p_k^r, \quad (6)$$

其中  $y_k^r \in \{0, 1\}$  和  $p_k^r$  分别为第  $r$  个源域样本  $\mathbf{x}$  的第  $k$  个类别和分类器对第  $k$  个类别的预测概率. 然而, 交叉熵损失并不能反映由于源域对目标域的不确定带来的损失, 因此, 需要设计新的损失函数, 将不确定与分类损失相结合. 首先, 根据基于主观逻辑的深度学习方法<sup>[12]</sup>, 将分类器的激活函数由 Softmax 替换为 ReLU, 将分类器的输出视为证据值  $\mathbf{e}^r = [e_1^r, e_2^r, \dots, e_K^r]$ ; 然后, 将证据值转化为 Dirichlet 分布参数  $\boldsymbol{\alpha}^r = [\alpha_1^r, \alpha_2^r, \dots, \alpha_K^r]$ . 具体转化公式如下所示:

$$\mathbf{e}^r = \boldsymbol{\alpha}^r + \mathbf{1}. \quad (7)$$

在 Dirichlet 分布中, 参数  $\alpha_k^r$  可以反映第  $r$  个分类器对样本第  $k$  个类别的支持程度. 当  $\boldsymbol{\alpha}^r$  中每个  $\alpha_k^r$  的值均接近于 1 时, Dirichlet 分布更接近均匀分布, 表示分类器对于每个类别的偏好程度相似, 分类决策的不确定程度较高; 当  $\boldsymbol{\alpha}^r$  中某个  $\alpha_k^r$  的值相对较大时, 分类器对  $\alpha_k^r$  值较高的类别支持度高, 决策的不确定性较小<sup>[12]</sup>. 将交叉熵损失与 Dirichlet 分布相结合, 组成具有不确定度量的交叉熵损失, 在衡量模型预测结果的同时, 也考虑了不确定对于预测结果的影响. 具有不确定度量的交叉熵损失的公式如下所示:

$$\mathcal{L}_{\text{cls}}^r(\mathbf{x}) = \int \left[ \sum_{k=1}^K -y_k^r \log(p_k^r) \right] D(\mathbf{p}^r | \boldsymbol{\alpha}^r) d\mathbf{p}^r = \sum_{k=1}^K y_k^r (\psi(S^r) - \psi(\alpha_k^r)). \quad (8)$$

其中:  $\psi(x) \equiv \frac{d}{dx} \ln \Gamma(x)$  为 digamma 函数,  $p_k^r$  为  $\mathbf{x}$  属于第  $k$  类的概率,  $S = K + \sum_{k=1}^K e_k^r = \sum_{k=1}^K \alpha_k^r$  为分类器对所有类别支持程度的和.

在样本输入至  $R$  个分类器后, 分类器会输出  $R$  个证据值  $e^1, e^2, \dots, e^R$ . 然后, 通过主观逻辑的方法, 证据值被转化为分类器对样本分类结果的信度和不确定, 具体公式如下所示:

$$\begin{cases} b^r = \frac{e^r}{K + \sum_{k=1}^K e_k^r} = \frac{e^r}{S}, \\ u^r = \frac{K}{K + \sum_{k=1}^K e_k^r} = \frac{K}{S}. \end{cases} \quad (9)$$

其中:  $\mathbf{b}^r = [b_1^r, b_2^r, \dots, b_K^r]$  为信度, 表示分类器对每个类别的信度分配;  $u^r$  为不确定, 表示分类器对于目标域分类结果的不确定度. 由式 (9), 可得到  $\mathbf{b}^r$  与  $u^r$  间的关系以及概率向量  $\mathbf{p}^r$  的计算公式, 如下所示:

$$u^r + \sum_{k=1}^K b_k^r = 1, \quad p_k^r = b_k^r + a_k^r u^r, \quad (10)$$

这里  $\mathbf{a}^r = [a_1^r, a_2^r, \dots, a_K^r]$  为先验概率. 在通常情况下, 假设先验概率服从均匀分布. 目标域样本在输入至  $R$  个分类器后, 获得了  $R$  个包含信度和不确定的结果  $\omega^1, \omega^2, \dots, \omega^R$ .

### 2.1.3 正则化损失

式 (8) 保证每个样本的正确类别标签会比其他类别产生更多的证据, 使得最终结果支持正确类别. 然而, 其仅对正确类别的证据增长提供约束, 而对于错误类别的证据缺乏明确的抑制机制, 无法保证错误标签会产生更少的证据. 为弥补这一不足, 引入正则化项, 对错误标签的证据施加约束, 进一步提升决策的可信程度. 可通过 KL 散度控制正确类别与错误类别间的分布差异, 使得模型能够更加严格地区分正确的类别与错误的类别. KL 正则化项具体如下所示:

$$\mathcal{L}_{\text{KL}}^r(\mathbf{x}) = \text{KL}[D(\mathbf{p}^r | \tilde{\boldsymbol{\alpha}}^r) \| D(\mathbf{p}^r | \mathbf{1})] =$$

$$\log \left( \frac{\Gamma \left( \sum_{k=1}^K \tilde{\alpha}_k^r \right)}{\Gamma(K) \prod_{k=1}^K \Gamma(\tilde{\alpha}_k^r)} \right) + \sum_{k=1}^K (\tilde{\alpha}_k^r - 1) \left[ \psi(\tilde{\alpha}_k^r) - \psi \left( \sum_{j=1}^K \tilde{\alpha}_j^r \right) \right]. \quad (11)$$

其中:  $D(\mathbf{p}^r | \mathbf{1})$  表示均匀分布,  $\tilde{\alpha}_i = y_i + (1 - y_i) \odot \alpha_i$  为去除非误导证据后的狄利克雷参数. KL 散度将错误类别的 Dirichlet 分布参数惩罚为 1, 使得错误类别的证据值降为 0, 以减少模型在训练时对不正确类别的支持程度. 在训练初期, 若对 KL 散度赋予较高的权重, 则由式 (11) 可知, 模型会过早地与均匀分布对齐, 导致模型欠拟合. 若在训练初期将  $\lambda_2$  设置为较小的值, 则能够令模型首先聚焦于数据特征的学习, 在训练后期再逐步引入正则化效应, 使得训练数据的分布与均匀分布对齐. 模拟退火方法能够使得模型在训练前期专注于数据的拟合, 并在训练后期逐步增加 KL 散度对于数据的影响, 从而鼓励模型学习输出正确的类别, 既避免了训练初期的噪声干扰, 又能够在后期有效地正则化模型, 从而提升对测试集的泛化能力. 模拟退火的具体公式为

$$\lambda_2 = \min \left( 1, \frac{\text{epoch\_num}}{\text{annealing\_step}} \right). \quad (12)$$

这里: epoch\_num 为当前模型已经训练的轮数; 而 annealing\_step 为预设的退火步长, 用于控制  $\lambda_2$  的增速. 该参数通过一个单调递增的方式逐渐从 0 增加至 1. 当 epoch\_num < annealing\_step 时,  $\lambda_2 = \frac{\text{epoch\_num}}{\text{annealing\_step}}$ ; 当 epoch\_num  $\geq$  annealing\_step 时,  $\lambda_2 = 1$ . 通过这种方式实现了模拟退火的效果.

### 2.2 多源决策结果融合

根据第 2.1 节内容, 获得了目标域样本在  $R$  个源域支持下的决策结果. 接下来将利用信息融合方法, 综合  $R$  个决策结果, 获得可信的故障诊断结果.

令  $\omega^r = (\{b_k^r\}_{k=1}^K, u^r)$  和  $\omega^s = (\{b_k^s\}_{k=1}^K, u^s)$  分别表示第  $r$  个和第  $s$  个源域对目标域样本的决策结果, 则融合的公式为

$$\Omega = \omega^r \oplus \omega^s, \quad (13)$$

其中  $\oplus$  为融合算子. 具体融合公式为

$$b_k = \frac{b_k^r b_k^s + b_k^r u^s + b_k^s u^r}{1 - C}, \quad u = \frac{u^r u^s}{1 - C}, \quad (14)$$

这里  $C = \sum_{i \neq j} b_i^r b_j^s$  度量了两个不同来源的信度冲突. 通过两两融合, 可得到  $R$  个源域融合的结果为

$$\Omega = \omega^1 \oplus \omega^2 \oplus \dots \oplus \omega^R. \quad (15)$$

TMDAFD 模型算法具体流程如算法 1 所示.

算法 1 TMDAFD 模型.

input: 源域  $\mathcal{D}_s^r = \{\mathbf{X}_s^r, \mathbf{Y}_s^r\}$ , 目标域  $\mathcal{D}_t = \{\mathbf{X}_t\}$ ; 平衡参数  $\lambda_1, \lambda_2$ ; 训练次数  $T$ ; CNN 网络结构参数.

output: 目标域样本的诊断结果  $\mathbf{Y}_t$ .

1. for  $t \leq T$  do
2. 利用 1D-CNN 提取所有源域和目标域的特征;
3. 由式 (2), 计算  $R$  个源域与目标域间的 MMD 损失  $\mathcal{L}_{\text{MMD}}$  并更新 CNN 网络结构参数;
4. for  $r \leq R$  do
5. 由式 (7) 和 (10), 计算  $\mathcal{L}_{\text{cls}}$  和  $\mathcal{L}_{\text{KL}}$ , 更新 CNN 网络结构的参数;
6. end for
7. end for
8. for  $r \leq R$  do
9. 获得目标域样本在嵌入空间的表征, 并输入至第  $r$  个分类器中, 获得证据  $e^r$ , 并由式 (9) 计算信度和不确定;
10. end for
11. 由式 (14) 和 (15), 融合  $R$  个决策结果;
12. return  $\mathbf{Y}_t$

### 3 实验和结果分析

为验证所提出跨域故障诊断方法 TMDAFD 的优越性, 本节将采用真实滚动轴承数据集进行实验验证.

#### 3.1 数据集与对比方法

实验采用如下 3 个数据集: 凯斯西储大学 (CWRU) 轴承故障数据集、江南大学 (JNU) 轴承故障数据集和渥太华大学 (UO) 轴承故障数据集. 在实验中, 每个数据集选择一个域为目标域, 其余域的数据为源域, 作为一个跨域故障诊断任务. 实验所用的数据集的详细信息如表 1 所示.

表1 数据集详细信息

| 数据集  | 工况 |          | 健康状况 |         | 样本数量    | 类别个数 | 实验所用样本数  |
|------|----|----------|------|---------|---------|------|----------|
|      | 域  | 转速       | 健康类型 | 标签      |         |      |          |
| CWRU | A  | 1720 rpm | H    | 0       | 473     | 10   | 10 × 118 |
|      | B  | 1750 rpm | 3O   | 1, 2, 3 | 3 × 118 |      |          |
|      | C  | 1772 rpm | 3I   | 4, 5, 6 | 3 × 118 |      |          |
|      | D  | 1797 rpm | 3R   | 7, 8, 9 | 3 × 118 |      |          |
| JNU  | A  | 600 rpm  | H    | 0       | 1466    | 4    | 4 × 128  |
|      | B  | 800 rpm  | O    | 1       | 488     |      |          |
|      |    |          | I    | 2       | 488     |      |          |
|      | C  | 1000 rpm | R    | 3       | 488     |      |          |
| UO   | A  | 加速       | H    | 0       | 1953    | 3    | 3 × 256  |
|      | B  | 减速       | O    | 1       | 1953    |      |          |
|      | C  | 先加速再减速   | O    | 1       | 1953    |      |          |
|      | D  | 先减速再加速   | I    | 2       | 1953    |      |          |

数据均为电机采集的信号数据, 因此, 需要对数据进行预处理. 具体包括: 首先, 使用长度为 1024 的步长对信号进行采样; 然后, 使用快速傅立叶变换 (FFT), 将采样后的时域信号转为频域信号, 作为模型的输入. 实验采用了 6 种对比方法, 具体如下.

1) DAN-optimal: DAN 是一种单源领域自适应方法, 其通过最大均值差异来衡量源域与目标域间的边缘概率分布差异<sup>[26]</sup>. 基于 DAN-optimal 方法的实验分别在不同源域数据和目标域数据上实施 DAN, 并将所有源域作用下的最优值作为诊断决策.

2) DAN-combine: 该方法在合并过程中不考虑域标签, 直接将多源域样本简单融合为一个源域, 然后利用合并的源域训练分类器来诊断故障.

3) DANN-optimal: DANN 方法通过利用源域有标签数据和目标域无标签数据, 使得神经网络能够在跨域学习中高效迁移<sup>[27]</sup>. DANN-optimal 方法分别利用不同源域数据和目标域数据开展实验, 并将所有源域作用下的最优值作为诊断决策.

4) DANN-combine: 通过将多源域样本简单合并为一个源域, 再实施 DANN 方法, 以实现利用多源域数据的故障诊断任务.

5) MFSAN<sup>[28]</sup>: 该方法基于多源无监督域适应 (MUDA) 框架, 其核心思想是分两步进行特征对齐. 首先, 在多个特定特征空间中分别对齐每个源域与目标域的分布; 然后, 通过域特定决策边界对分类器的输出进行对齐. 该方法能够有效应对多源域间的分布差异所带来的挑战.

6) MSSA<sup>[13]</sup>: 该方法构建了一个具有共享子网络和多个源特定子网络的集成网络模型, 并将每对源域和目标域分别映射至特定的特征空间来消除域偏移, 通过多分支网络更准确地对齐每个源域与目标域间的特征, 从而提升诊断的精度.

#### 3.2 实验设置

实验在 Pytorch 深度学习框架下进行, GPU 版本为 GeForce RTX3060, 显存为 16 GB.

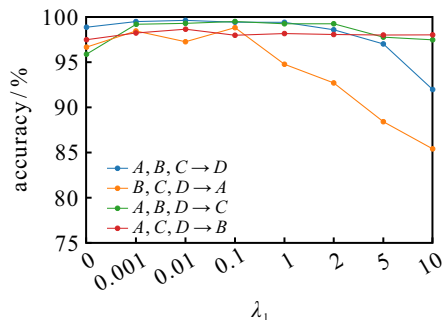
在 TMDAFD 模型中: 首先, 基于一维 CNN 网络提取输入数据的特征. 该网络包括 1 个输入层, 由 6 个卷积层、6 个池化层组成的隐藏层以及 1 个输出层. 在隐藏层中, 每层的神经元个数分别为 4、8、16、24、32 和 32. 类似地, 分类器基于一维 CNN 网络搭建, 包括 1 个输入层, 由 2 个卷积层、2 个池化层、2 个全连接层组成的隐藏层以及 1 个输出层. 其中: 卷积层的权重初始化采用 Kaiming 初始化方法, 全连接层的权重初始化采用 Xavier 方法.

6种对比方法的参数均设置为默认参数. 为保障实验的公平性, 6种对比方法均采用同样的数据预处理方式并采用一维 CNN 网络对特征进行提取. 同时, 为减少随机性对实验结果的影响, 将所有方法在每个任务中重复执行 10 次, 并报告 10 次实验下的平均准确率.

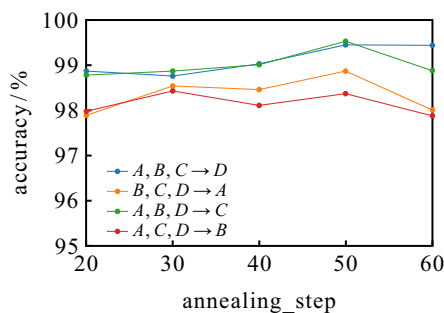
### 3.3 参数敏感性分析

由式 (1) 可见, 在所构造的损失函数中, 有两个权重参数, 分别为 MMD 距离损失的平衡因子  $\lambda_1$  和 KL 散度项的平衡因子  $\lambda_2$ ,  $\lambda_2$  的值由退火步长确定. 本节将对两个平衡参数进行敏感性实验, 分析参数变化对于模型性能的影响. 实验分为以下两部分.

首先, 将退火步长 `annealing_step` 固定为 50, 考察  $\lambda_1 \in \{0, 0.001, 0.01, 0.1, 1, 2, 5, 10\}$  参数空间对于目标域故障诊断精度的影响, 结果如图 2(a) 所示. 由图 2(a) 可见: 随着  $\lambda_1$  值的增加, 诊断准确率呈现出先上升再下降的趋势; 模型的性能在  $\lambda_1 = 0.01$  时达到最优, 之后  $\lambda_1$  的增加对准确率的影响变化不显著, 而在  $\lambda_1$  值超过一定阈值后性能逐渐下降. 然后, 将  $\lambda_1$  的值固定为 0.01, 分别令退火步长 `annealing_step` 在  $\{20, 30, 40, 50, 60\}$  内变化, 得到不同退火步长下的故障诊断准确率结果如图 2(b) 所示. 由图 2(b) 可见, 随着退火步长值的增加, 诊断准确率同样出现先上升再下降的变化.



(a) 参数  $\lambda_1$  取不同值时的分类准确率



(b) 参数  $\lambda_2$  中退火步长取不同值时的分类准确率

图2 参数敏感性分析

由敏感性分析结果可见, MMD 距离损失  $\mathcal{L}_{\text{MMD}}$

和正则化项  $\mathcal{L}_{\text{KL}}$  项对于模型性能均有促进作用. 当其权重参数取值在一定范围内时, 能够提升模型性能. 根据敏感性分析实验的结果, 将目标函数中 MMD 损失项的平衡参数  $\lambda_1$  设置为 0.01, 将决定正则化损失项的平衡参数  $\lambda_2$  的退火步长参数设为 50.

### 3.4 结果分析

本节将对实验结果进行分析, 在 CWRU、JNU 和 UO 等典型跨域故障诊断数据集上, 通过平均诊断准确率验证方法有效性, 具体结果如表 2 ~ 表 4 所示. 下面分别将 TMDAFD 与单源领域自适应以及多源领域自适应方法进行对比分析, 通过综合评估揭示 TMDAFD 在不同域适应场景下的方法优势.

#### 1) 与传统单源领域自适应方法对比.

由表 2 ~ 表 4 的结果可见, TMDAFD 方法相较于单源域最优方法以及源域合并方法, 均表现出更优的性能. 相比于最优的对比方法, TMDAFD 方法在 3 个数据集下的平均诊断准确率分别提高了 3.25%、4.76% 和 4.13%, 这表明 TMDAFD 作为多源领域自适应方法, 能够更有效地整合和利用多个源域的数据, 克服了单一源域方法的局限性.

#### 2) 与多源领域自适应方法的对比.

由表 2 ~ 表 4 的结果可见, 与其他两种多源领域自适应方法 (MFSAN / MSSA) 对比的结果显示, TMDAFD 模型在 3 个数据集上均取得了显著的性能提升, 平均分类准确率相比于对比方法分别提高了 4.53% / 2.13%、4.75% / 6.07% 和 2.70% / 2.47%. TMDAFD 与 MFSAN、MSSA 均是基于多源域信息实现跨域故障诊断的, 而 TMDAFD 方法通过量化分析各源域支持下目标域诊断结果的信度水平, 在决策过程中建立源域质量差异敏感的加权融合模型, 有效克服了传统多源域方法因忽略域间分布偏移程度差异所导致的负迁移问题, 提升了多源跨域故障诊断模型的准确率.

### 3.5 消融实验

不确定度量和融合方法是 TMDAFD 模型的关键. 本节将分别对上述两个模块设计消融实验, 验证这两个模块在模型中的作用.

TMDAFD 模型的主要贡献在于通过证据深度学习方法, 建立了包含不确定度量的交叉熵损失, 为每个源域下决策结果提供了可信用度量, 并据此设计了融合规则. 在关于不确定度量的消融实验中, 将对对比方法的损失函数替换为普通的交叉熵分类损失, 并将融合方法替换为简单的少数服从多数的投票法. 在关于融合方法的消融实验中, 将对对比方法的不同

表2 CWRU 数据集上的诊断准确率

| 方法           | 任务名                     |                         |                         |                         | 均值           |
|--------------|-------------------------|-------------------------|-------------------------|-------------------------|--------------|
|              | $A, B, C \rightarrow D$ | $A, B, D \rightarrow C$ | $A, C, D \rightarrow B$ | $B, C, D \rightarrow A$ |              |
| DAN-optimal  | 94.38 ± 0.02            | 94.54 ± 0.03            | 92.17 ± 0.03            | 91.34 ± 0.04            | 93.11        |
| DANN-optimal | 94.64 ± 0.02            | 92.33 ± 0.03            | 93.84 ± 0.02            | 92.75 ± 0.03            | 93.39        |
| DAN-combine  | 96.86 ± 0.02            | 95.49 ± 0.03            | 95.57 ± 0.03            | 94.95 ± 0.05            | 95.72        |
| DANN-combine | 96.75 ± 0.02            | 95.21 ± 0.03            | 93.83 ± 0.02            | 93.05 ± 0.03            | 94.71        |
| MSSA         | 95.49 ± 0.02            | 94.28 ± 0.02            | 94.69 ± 0.03            | 93.30 ± 0.02            | 94.44        |
| MFSAN        | 97.82 ± 0.02            | 97.29 ± 0.02            | 96.85 ± 0.03            | 95.41 ± 0.02            | 96.84        |
| TMDAFD       | <u>99.45</u> ± 0.01     | <u>99.53</u> ± 0.01     | <u>98.01</u> ± 0.01     | <u>98.87</u> ± 0.01     | <u>98.97</u> |

表3 JNU 数据集上的诊断准确率

| 方法           | 任务名                  |                      |                      | 均值           |
|--------------|----------------------|----------------------|----------------------|--------------|
|              | $A, B \rightarrow C$ | $A, C \rightarrow B$ | $B, C \rightarrow A$ |              |
| DAN-optimal  | 91.04 ± 0.01         | 92.32 ± 0.01         | 87.44 ± 0.02         | 90.27        |
| DANN-optimal | 91.76 ± 0.02         | 93.41 ± 0.01         | 90.46 ± 0.03         | 91.88        |
| DAN-combine  | 92.07 ± 0.01         | 92.97 ± 0.01         | 94.68 ± 0.02         | 93.24        |
| DANN-combine | 92.56 ± 0.02         | 93.75 ± 0.01         | 92.13 ± 0.03         | 92.81        |
| MSSA         | 92.52 ± 0.01         | 93.69 ± 0.01         | 93.55 ± 0.03         | 93.25        |
| MFSAN        | 90.21 ± 0.01         | 92.54 ± 0.01         | 93.04 ± 0.02         | 91.93        |
| TMDAFD       | <u>97.78</u> ± 0.02  | <u>98.29</u> ± 0.02  | <u>97.94</u> ± 0.01  | <u>98.00</u> |

表4 UO 数据集上的诊断准确率

| 方法           | 任务名                     |                         |                         |                         | 均值           |
|--------------|-------------------------|-------------------------|-------------------------|-------------------------|--------------|
|              | $A, B, C \rightarrow D$ | $A, B, D \rightarrow C$ | $A, C, D \rightarrow B$ | $B, C, D \rightarrow A$ |              |
| DAN-optimal  | 93.28 ± 0.03            | 94.48 ± 0.02            | 92.37 ± 0.02            | 91.57 ± 0.02            | 92.93        |
| DANN-optimal | 92.56 ± 0.02            | 91.86 ± 0.02            | 90.23 ± 0.02            | 92.75 ± 0.04            | 91.85        |
| DAN-combine  | 95.66 ± 0.01            | 95.28 ± 0.02            | 94.38 ± 0.03            | 93.31 ± 0.03            | 94.66        |
| DANN-combine | 95.56 ± 0.01            | 94.38 ± 0.03            | 96.48 ± 0.03            | 94.58 ± 0.03            | 95.25        |
| MSSA         | 97.03 ± 0.01            | 96.88 ± 0.02            | 96.23 ± 0.02            | 96.56 ± 0.02            | 96.68        |
| MFSAN        | 97.32 ± 0.02            | 96.45 ± 0.02            | 95.56 ± 0.03            | 98.29 ± 0.02            | 96.91        |
| TMDAFD       | <u>99.28</u> ± 0.01     | <u>99.44</u> ± 0.01     | <u>99.28</u> ± 0.01     | <u>99.51</u> ± 0.01     | <u>99.38</u> |

源域样本合并为一个源域,提取该域与目标域的特征并在嵌入空间中对齐,通过一个分类器输出最后的诊断结果.修改后的对比方法分别记为

TMDAFD\_w/o\_Uncertainty、TMDAFD\_w/o\_Fusion,其与TMDAFD方法在3个滚动轴承数据集上的实验结果如图3所示.

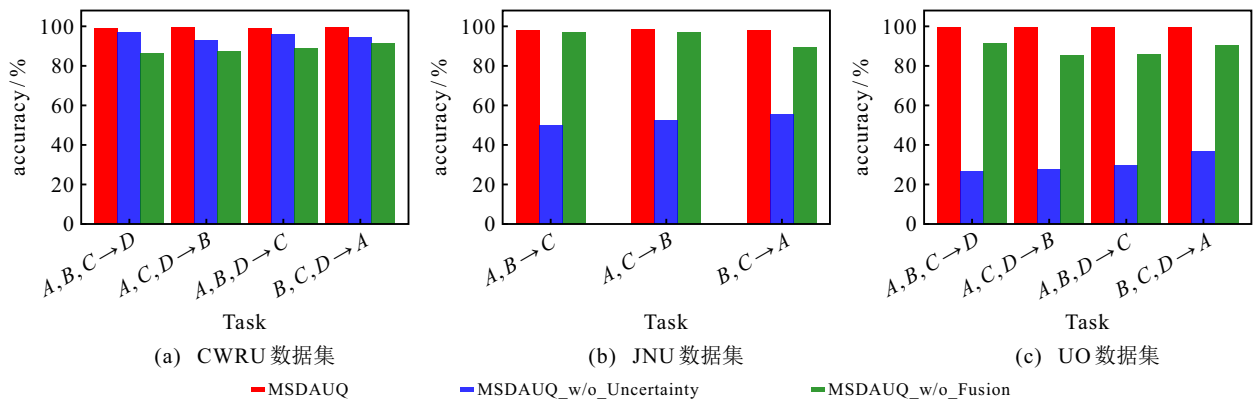


图3 不确定度量方法与融合方法的消融实验结果

由图3可知,TMDAFD在3个数据集上的准确率均优于TMDAFD\_w/o\_Uncertainty方法与TMDAFD\_w/o\_Fusion方法.上述结果表明,所提出

不确定度量方法能够有效量化源域训练分类器对目标域诊断结果的不确定性.通过加权融合机制,实现了多分类器决策信息的差异化集成.

## 4 结 论

针对现有多源领域自适应故障诊断方法中未充分考虑不同源域信息质量差异的问题, 本文提出了一种基于可信多源领域自适应的跨域滚动轴承故障诊断模型 TMDAFD. 该模型首先利用 MMD 距离, 构建了源域与目标域的最优公共嵌入空间, 实现了特征有效对齐, 降低了域间分布差异; 然后, 设计了基于主观逻辑的深度神经网络不确定性度量方法, 度量不同源域对于目标域诊断决策的可信度, 刻画了不同源域样本在目标域故障诊断任务中的作用, 为信息融合提供了可靠依据; 接着, 根据所得到的可信度量, 提出了多源决策信息融合策略, 有效利用多个源域知识, 提升了目标域诊断决策的准确性; 最后, 分别在凯斯西储大学、江南大学和渥太华大学的滚动轴承故障数据集上开展了实验验证. 实验结果表明, TMDAFD 方法在跨域故障诊断任务中取得了显著的性能提升.

### 参考文献 (References)

- [1] 高佳, 王艳红, 张俊, 等. 基于多元信号自适应分解的轴承故障诊断方法[J]. 控制与决策, DOI: [10.13195/j.kzyjc.2024.0758](https://doi.org/10.13195/j.kzyjc.2024.0758). (Gao J, Wang Y H, Zhang J, et al. Bearing fault diagnosis method based on adaptive decomposition of multivariate signals[J]. Control and Decision, DOI: [10.13195/j.kzyjc.2024.0758](https://doi.org/10.13195/j.kzyjc.2024.0758).)
- [2] Xiao Y M, Shao H D, Han S Y, et al. Novel joint transfer network for unsupervised bearing fault diagnosis from simulation domain to experimental domain[J]. *IEEE/ASME Transactions on Mechatronics*, 2022, 27(6): 5254-5263.
- [3] 毛文涛, 施华东, 张艳娜, 等. 轴承在线早期故障检测的无监督张量深度迁移学习方法[J]. 控制与决策, 2024, 39(3): 867-876. (Mao W T, Shi H D, Zhang Y N, et al. Unsupervised tensor-based deep transfer learning for online early fault detection of bearing[J]. Control and Decision, 2024, 39(3): 867-876.)
- [4] Tang S N, Ma J T, Yan Z Q, et al. Deep transfer learning strategy in intelligent fault diagnosis of rotating machinery[J]. *Engineering Applications of Artificial Intelligence*, 2024, 134: 108678.
- [5] Jiang M, Zhou K, Gao J H, et al. Integrating causal representations with domain adaptation for fault diagnosis[J]. *Reliability Engineering & System Safety*, 2025, 260: 110999.
- [6] 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.)
- [7] Chen G P, Xiang D, Liu T T, et al. Deep conditional adversarial subdomain adaptation network for unsupervised mechanical fault diagnosis[J]. *Knowledge-Based Systems*, 2024, 300: 112180.
- [8] Yin H, Chen Q T, Chen L, et al. Cross-attention transformer-based domain adaptation: A novel method for fault diagnosis of rotating machinery with high generalizability and alignment capability[J]. *IEEE Sensors Journal*, 2024, 24(23): 40049-40058.
- [9] Chi J K, Mao Z Z. Deep domain-adversarial anomaly detection with robust one-class transfer learning[J]. *Knowledge-Based Systems*, 2024, 300: 112225.
- [10] 文利燕, 陈金陵, 姜斌, 等. 基于联合对抗域自适应网络的跨工况故障诊断方法[J]. 控制与决策, 2025, 40(5): 1503-1511. (Wen L Y, Chen J L, Jiang B, et al. A joint adversarial domain adaptive network based cross working conditions fault diagnosis method[J]. Control and Decision, 2025, 40(5): 1503-1511.)
- [11] 史玉皓, 田建艳, 刘军军, 等. 基于流形结构的多源自适应迁移学习算法及应用研究[J]. 控制与决策, 2023, 38(3): 797-804. (Shi Y H, Tian J Y, Liu J J, et al. Research on multi-domain adaptation-manifold regularization and application[J]. Control and Decision, 2023, 38(3): 797-804.)
- [12] Sensoy M, Kaplan L, Kandemir M. Evidential deep learning to quantify classification uncertainty[C]. Proceedings of the 32nd International Conference on Neural Information Processing Systems. Montréal, 2018: 3183-3193.
- [13] Tian J H, Han D Y, Li M D, et al. A multi-source information transfer learning method with subdomain adaptation for cross-domain fault diagnosis[J]. *Knowledge-Based Systems*, 2022, 243: 108466.
- [14] Wen L S, Chen S T, Xie M Y, et al. Training multi-source domain adaptation network by mutual information estimation and minimization[J]. *Neural Networks*, 2024, 171: 353-361.
- [15] Liang P F, Tian J Y, Wang S Y, et al. Multi-source information joint transfer diagnosis for rolling bearing with unknown faults via wavelet transform and an improved domain adaptation network[J]. *Reliability Engineering & System Safety*, 2024, 242: 109788.
- [16] Liu S W, Jiang H K, Wu Z H, et al. Intelligent fault diagnosis of rotating machinery using a multi-source domain adaptation network with adversarial discrepancy matching[J]. *Reliability Engineering & System Safety*, 2023, 231: 109036.
- [17] Shi J C, Wang X L, Lu S L, et al. An adversarial multisource data subdomain adaptation model: A promising tool for fault diagnosis of induction motor under cross-operating conditions[J]. *IEEE Transactions on Instrumentation and Measurement*, 2023, 72:

- 3519014.
- [18] Gao T Y, Yang J L, Tang Q. A multi-source domain information fusion network for rotating machinery fault diagnosis under variable operating conditions[J]. *Information Fusion*, 2024, 106: 102278.
- [19] Xu P, Cho J H, Salado A. Expert opinion fusion framework using subjective logic for fault diagnosis[J]. *IEEE Transactions on Cybernetics*, 2022, 52(6): 4300-4311.
- [20] Zhou H T, Chen W H, Cheng L S, et al. Trustworthy fault diagnosis with uncertainty estimation through evidential convolutional neural networks[J]. *IEEE Transactions on Industrial Informatics*, 2023, 19(11): 10842-10852.
- [21] Xiao Y M, Shao H D, Feng M J, et al. Towards trustworthy rotating machinery fault diagnosis via attention uncertainty in transformer[J]. *Journal of Manufacturing Systems*, 2023, 70: 186-201.
- [22] Zhou T T, Zhang L B, Han T, et al. An uncertainty-informed framework for trustworthy fault diagnosis in safety-critical applications[J]. *Reliability Engineering & System Safety*, 2023, 229: 108865.
- [23] Zhang W Q, Lv Z Q. Revisiting the domain shift and sample uncertainty in multi-source active domain transfer[C]. *IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Seattle, 2024: 16751-16761.
- [24] Pei J B, Men A D, Liu Y, et al. Evidential multi-source-free unsupervised domain adaptation[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2024, 46(8): 5288-5305.
- [25] Kiranyaz S, Avci O, Abdeljaber O, et al. 1D convolutional neural networks and applications: A survey[J]. *Mechanical Systems and Signal Processing*, 2021, 151: 107398.
- [26] Long M S, Cao Y, Wang J M, et al. Learning transferable features with deep adaptation networks[C]. *Proceedings of the 32nd International Conference on Machine Learning*. Lille, 2015: 97-105.
- [27] Ganin Y, Ustinova E, Ajakan H, et al. Domain-adversarial training of neural networks[J]. *Journal of Machine Learning Research*, 2016, 17(59): 1-35.
- [28] Zhu Y C, Zhuang F Z, Wang D Q. Aligning domain-specific distribution and classifier for cross-domain classification from multiple sources[J]. *Proceedings of the AAAI Conference on Artificial Intelligence*, 2019, 33(1): 5989-5996.

### 作者简介

周旷 (1987-), 男, 副教授, 博士, 主要研究方向为多源信息融合、因果推断与复杂网络, E-mail: [kzhoumath@nwpu.edu.cn](mailto:kzhoumath@nwpu.edu.cn);

覃文博 (2000-), 男, 硕士生, 主要研究方向为领域自适应与不确定量化, E-mail: [qinwenbo@mail.nwpu.edu.cn](mailto:qinwenbo@mail.nwpu.edu.cn);

孙天宇 (1982-), 男, 研究员, 博士, 主要研究方向为通用质量特性系统工程、可靠性统计, E-mail: [sunono3997@163.com](mailto:sunono3997@163.com).