



基于联合对抗域自适应网络的跨工况故障诊断方法

文利燕, 陈金陵, 姜斌, 马亚杰, 冯靖智

引用本文:

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.

在线阅读 View online: <https://doi.org/10.13195/j.kzyjc.2024.0767>

您可能感兴趣的其他文章

Articles you may be interested in

基于Hessian正则的自适应损失半监督特征选择

Adaptive loss semi-supervised feature selection based on Hessian regularization

控制与决策. 2021, 36(8): 1862–1870 <https://doi.org/10.13195/j.kzyjc.2019.1510>

基于广义主成分分析的重构故障子空间建模方法

Reconstructed fault subspace modelling method based on generalized principal component analysis

控制与决策. 2021, 36(4): 808–814 <https://doi.org/10.13195/j.kzyjc.2019.0818>

基于双权重多邻域保持嵌入的间歇过程故障检测

Fault detection of batch process based on double weight and multiple neighborhoods preserving embedding

控制与决策. 2021, 36(12): 3023–3030 <https://doi.org/10.13195/j.kzyjc.2020.0659>

高超声速飞行器间歇故障改进自适应容错控制

Improved adaptive fault-tolerant control of intermittent faults in hypersonic flight vehicle

控制与决策. 2021, 36(11): 2627–2636 <https://doi.org/10.13195/j.kzyjc.2020.0483>

考虑加速度约束的终端角度约束滑模制导律设计

Sliding-mode guidance law with acceleration and angle constraints

控制与决策. 2021, 36(10): 2511–2516 <https://doi.org/10.13195/j.kzyjc.2020.0284>

基于联合对抗域自适应网络的跨工况故障诊断方法

文利燕, 陈金陵, 姜斌[†], 马亚杰, 冯靖智

(南京航空航天大学 自动化学院, 南京 211106)

摘要: 针对机床零部件在实际场景中难以获得带标签的故障数据, 且训练数据与测试数据分布不一致导致诊断模型不适用的问题, 提出一种基于联合对抗域自适应网络 (JADAN) 的跨工况故障诊断方法. 首先, 利用对抗域自适应训练来提取源域和目标域的深层域不变特征, 以提高诊断模型在目标域的泛化能力; 其次, 提出一种基于 Softmax 预测和结构化预测的伪标签策略, 使模型能够为无标签的目标域数据生成伪标签; 同时, 加入类对齐模块, 最小化源域和目标域之间的类原型距离, 实现域与类的联合对齐, 有效减少决策边界附近样本被错误分类的概率; 然后, 在域判别器中引入源域样本的权重分配机制, 为每个源域样本自适应地分配权重, 有效解决模型训练过程中的负迁移问题, 提升模型的鲁棒性; 最后的实验结果表明, 所提出的方法能够更有效地解决跨工况故障诊断问题.

关键词: 机床零部件; 对抗域自适应; 类对齐; 权重分配机制; 跨工况故障诊断

中图分类号: TP183

文献标志码: A

DOI: 10.13195/j.kzyjc.2024.0767

引用格式: 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.

A joint adversarial domain adaptive network based cross working conditions fault diagnosis method

WEN Li-yan, CHEN Jin-ling, JIANG Bin[†], MA Ya-jie, FENG Jing-zhi

(College of Automation Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing 211106, China)

Abstract: A joint adversarial domain adaptive network (JADAN) based cross working conditions fault diagnosis method is proposed to address the challenges of acquiring labeled fault data for machine tool components in real-world scenarios and the inconsistencies between training and test data distributions that render diagnostic models ineffective. Initially, adversarial domain adaptation is employed to extract deep domain-invariant features from both source and target domains, enhancing the generalization capability of the diagnostic model in the target domain. A Softmax and structural prediction-based pseudo-label strategy is proposed, enabling the model to generate pseudo-labels for unlabeled target domain data. Simultaneously, a class alignment module is incorporated to minimize the distance between class prototypes of the source and target domains, achieving joint alignment of domain and class. This effectively reduces the probability of samples near the decision boundary being misclassified. Additionally, a weight assignment mechanism for source domain samples is integrated into the domain discriminator, enabling adaptive weight assignment for each source sample, which effectively mitigates the negative transfer issue during model training and enhances model robustness. Experimental results demonstrate that the proposed method can effectively address the problem of cross working conditions fault diagnosis.

Keywords: machine tool components; adversarial domain adaptation; class alignment; weight assignment mechanism; cross working conditions fault diagnosis

0 引言

随着高科技产品在航空、航天等行业中的广泛应用^[1-3], 机械加工制造业在其制造过程中扮演了重要角色. 其中, 数控机床以其加工精密组件的能力成

为该领域的核心. 然而, 恶劣的加工环境使得机床零部件容易出现故障, 对其故障诊断变得尤为重要^[4].

深度学习技术的快速发展为故障诊断提供了新的解决方案^[5]. 然而, 深度学习的有效性依赖于带标

收稿日期: 2024-06-29; 录用日期: 2024-10-28.

基金项目: 国家重点研发计划项目 (2021YFB3301300).

[†]通信作者. E-mail: binjiang@nuaa.edu.cn.

签的数据和训练与测试数据分布的一致性. 然而, 在实际应用中, 由于样本标注成本高和安全性限制, 大多数故障数据无标签^[6-7], 且不同任务下的机床工况变化导致数据分布存在差异. 这种跨工况问题严重影响了诊断模型的性能, 因此, 开展跨工况故障诊断具有重要意义^[8].

域适应作为解决跨工况问题的一种有效迁移学习方法, 能够利用源域数据或模型适应目标域数据分布, 从而降低泛化误差并提高模型鲁棒性. 文献[9]提出一种基于最大均值差异 (MMD) 和生成对抗网络 (GAN) 组合的故障诊断模型, 该方法通过卷积神经网络提取数据特征, 然后结合 MMD 和 GAN 来减小源域与目标域数据集之间的分布差异. 然而, 在实践中 MMD 忽略了标签信息与数据的空间结构关系. 针对该问题, 文献[10]提出一种新的度量结构深度迁移学习故障诊断方法, 在 MMD 的基础上学习两个转移域的公共度量矩阵, 设计了一种新的域适应模块, 可以使两个转移域的分布在均值差异和度量结构上相似.

随着 GAN 模型应用在故障诊断领域后, 其对抗博弈的思想也被引进了迁移学习中, 于是出现了一种基于对抗学习的域对抗神经网络 (DANN), 已经成为主流的跨工况故障诊断方法^[11]. 文献[12]提出了一种深度混合域自适应网络框架, 在 DANN 基础上结合再平衡混合训练技术, 确保基于数据不可知的可靠分布匹配; 采用强弱学习框架自动学习具有代表性的特征, 挖掘源域与目标域共享的隐藏信息, 以增强特征提取器的能力. 文献[13]在 DANN 的域判别自适应模块中加入了 MMD, 通过 MMD 损失和域判别损失共同帮助特征提取器学习域不变特征. 更进一步, 文献[14]提出了一种深度卷积多对抗域自适应模型, 使用改进的深度残差网络提取可转移特征; 通过多核 MMD 和多域判别器调整边缘与条件分布, 并设计自适应因子动态度量其相对重要性.

尽管 DANN 在改善跨域诊断方面存在较大潜力, 但仍存在一些局限性. 上述研究重点关注的是域对齐, 而对类别信息考虑得不够充分, 如果此时类间距不大, 则类边界周围样本的类别会被混淆. 此外, 在进行域适应时, 源域中的每个样本都被赋予相同的权重. 但有些源域样本的特征分布与目标域有较大区别, 它们也被赋予同样的重要性就会导致负迁移, 从而影响诊断准确率. 因此, 迫切需要开发一种更全面的跨工况诊断框架.

为了提高跨工况故障诊断的准确率, 本文提出一种基于联合对抗域自适应网络 (JADAN) 的跨工

况故障诊断方法, 主要工作为:

1) 针对跨工况下决策边界附近的样本容易被错误分类的问题, 引入类对齐模块, 实现领域与类的联合对齐, 以最小化源域和目标域的原型之间的距离, 提取到更有效的域不变特征, 进一步提升模型的泛化性.

2) 针对目标域标签难获取的问题, 设计 Softmax 预测和结构化预测相结合的伪标签策略, 通过特征提取器提取目标域样本特征后, 利用 Softmax 预测和结构化预测, 分别计算目标域样本属于各类别的条件概率, 取最大条件概率对应的类别作为目标域样本的伪标签.

3) 考虑到源域与目标域样本特征分布不同, 提出一种基于域相似度的权值分配机制, 为每个源域样本分配不同的权值, 抑制了模型诊断过程中出现的负迁移现象.

1 问题描述

1.1 跨工况故障诊断问题

机床在不同加工任务下运行工况会变化, 导致采集到的信号特征分布发生改变, 使在工况 A 上训练的诊断模型在工况 B 上效果变差. 为此, 故障诊断模型需提取出不同工况的共同特征. 传统机器学习模型在测试集数据分布变化时易出现域漂移, 缺乏泛化性, 导致误诊断, 这就是跨工况故障诊断问题.

为了使问题更加清晰直白, 给出一些假设和基本定义来介绍跨工况故障诊断任务. 首先, 将数据集分为源域数据集和目标域数据集, 假设源域数据集 $X_s = \{x_i^s, y_i^s\}_{i=1}^{n_s}$ 处于工况 Z_s , 目标域数据集 $X_t = \{x_i^t\}_{i=1}^{n_t}$ 处于工况 Z_t , 其中 n_s 和 n_t 分别表示源域和目标域样本数量, 且 $n_s = n_t$. 这两个域的标签空间是一致的, 即 $Y_s = Y_t$. 但由于它们的工况 $Z_s \neq Z_t$, 这两个域的边缘概率分布为 $P_s(x^s) \neq P_t(x^t)$, 条件概率分布为 $Q_s(y^s|x^s) \neq Q_t(y^t|x^t)$. 跨工况故障诊断的目的是根据有标签的源域样本和无标签的目标域样本训练一个函数模型, 使它可以识别目标域样本. 图 1

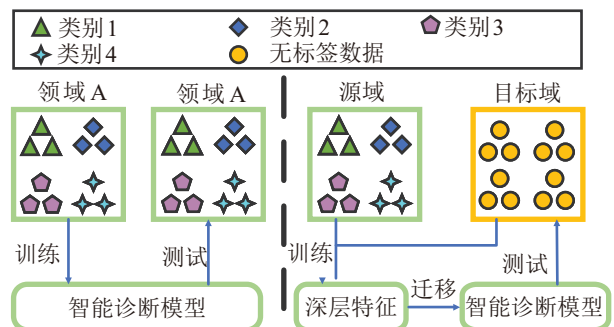


图1 传统故障诊断与跨工况故障诊断的区别

展示了跨工况故障诊断与传统故障诊断的区别, 左边是传统故障诊断, 右边是跨工况诊断。

本文所要解决的主要问题是针对跨工况故障诊断情形, 实现源域和目标域的深层相同特征提取。

1.2 技术基础: 域对抗神经网络

DANN 是一种对抗式迁移学习的经典模型, 由特征提取器、标签分类器和域判别器构成。其中特征提取器用于提取域的深层不变特征, 标签分类器对这些特征进行分类, 而域判别器用于判断特征提取器提取的特征属于源域还是目标域。此外, 在特征提取器与域判别器之间还有一个梯度反转层 (GRL) 用于对抗训练。

DANN 首先对标签分类器 G_y 进行训练, 通过最小化源域样本的标签分类损失 L_y , 可以优化特征提取器参数 θ_f 和标签分类器参数 θ_y 。采用多分类交叉熵损失作为 G_y 的损失函数, L_y 定义如下:

$$L_y(\theta_f, \theta_y) = -\frac{1}{n_s} \sum_{i=1}^{n_s} \sum_{k=1}^{K_s} I[y_i^s = k] \log(G_y^k(G_f(x_i^s))). \quad (1)$$

其中: K_s 为标签类别数; I 为指示函数, 参数为真时返回值为 1; G_y^k 表示标签分类器 G_y 输出第 k 类的概率; G_f 为特征提取器。

接下来, 需要对域判别器 G_d 进行训练, 通过最小化域判别损失 L_d 来更新域判别器参数 θ_d , 使其具有分辨领域类别的能力。采用二元交叉熵损失作为 G_d 的损失函数, L_d 定义如下:

$$L_d(\theta_f, \theta_d) = -\frac{1}{n_s} \sum_{i=1}^{n_s} \log(G_d(G_f(x_i^s))) - \frac{1}{n_t} \sum_{j=1}^{n_t} \log(1 - G_d(G_f(x_j^t))). \quad (2)$$

与此同时, 进行对抗网络的训练, 通过最大化域判别损失 L_d 来更新特征提取器参数 θ_f , 使其提取到的特征是域不变特征。可以看出, L_d 既要被最大化也要被最小化, 形成了对抗训练的情况。此外, 标签分类损失 L_y 要被最小化, 最终得到整个模型的损失函数

$$L(\theta_f, \theta_y, \theta_d) = L_y(\theta_f, \theta_y) - \lambda L_d(\theta_f, \theta_d), \quad (3)$$

其中 λ 是 GRL 系数。

2 联合对抗域自适应网络故障诊断方法

2.1 联合对抗域自适应模型

随着 DANN 的发展, 凭借其强大的域适应能力, 可以有效地对无标签数据进行诊断。该方法采用对抗性训练策略, 在新的空间中提取源域和目标域数

据的共同特征, 从而混淆两个域的数据分布。但是它只考虑了源域与目标域的域间距离, 没有考虑两个域之间同一类别的类间距离, 这将导致标签分类器的决策边界附近的样本被错误分类, 如图 2 所示。

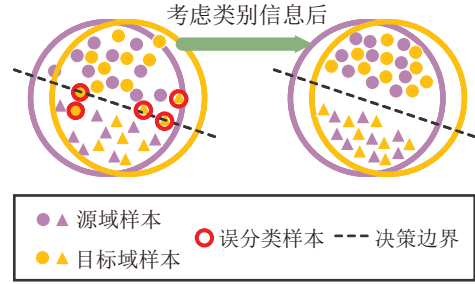


图2 考虑类别信息前后分类对比

此外, 现有的 DANN 方法在训练域判别器时, 没有考虑每个源域样本的重要性, 但是每个源域样本和每个目标域样本的差异性是不同的。如果不考虑这些差异性就更新参数, 则会导致负迁移问题。针对 DANN 存在的局限性, 本文提出了 JADAN 模型, 整体架构如图 3 所示。

该模型在 DANN 的基础上, 加入了一种伪标签策略来生成目标域样本的伪标签, 通过伪标签引入类对齐模块, 充分考虑了同一类别之间的类间距离。同时, 引入一种基于领域相似度的权重分配策略, 在域判别损失中根据领域相似度重新为每个源域样本分配权重。

1) 类对齐模块。

在传统的跨域迁移学习方法中, 域对齐通常只关注源域和目标域的整体特征分布 (即全局对齐), 而忽略了源域与目标域在类别级别上的分布差异。这样的全局对齐方式可能导致不同类别的样本在特征空间中混淆, 特别是在决策边界附近, 容易出现类别混淆的情况。这种现象在跨工况故障诊断任务中尤为显著, 因为不同工况下, 故障类别的分布差异较大, 仅依靠全局对齐难以满足诊断精度的要求。为了应对这一问题, 本文提出了自适应的类对齐模块, 通过最小化源域与目标域相同类别的类原型距离, 确保类别级别的特征对齐, 从而减少了分类错误。

源域和目标域样本的类标签是计算类原型所必须的条件, 由于目标域没有类标签, 需要生成目标域的类标签。但是, 如果直接利用标签分类器预测出目标域的类标签, 则极有可能不准确。根据错误的类标签计算出的类原型肯定也是错误的, 从而会引发错误的层层积累, 因此提出一种伪标签策略来计算目标域的类原型。相较于传统方法, 单一的伪标签生成方法可能在不同域分布下表现不佳, 为此, 提出一种

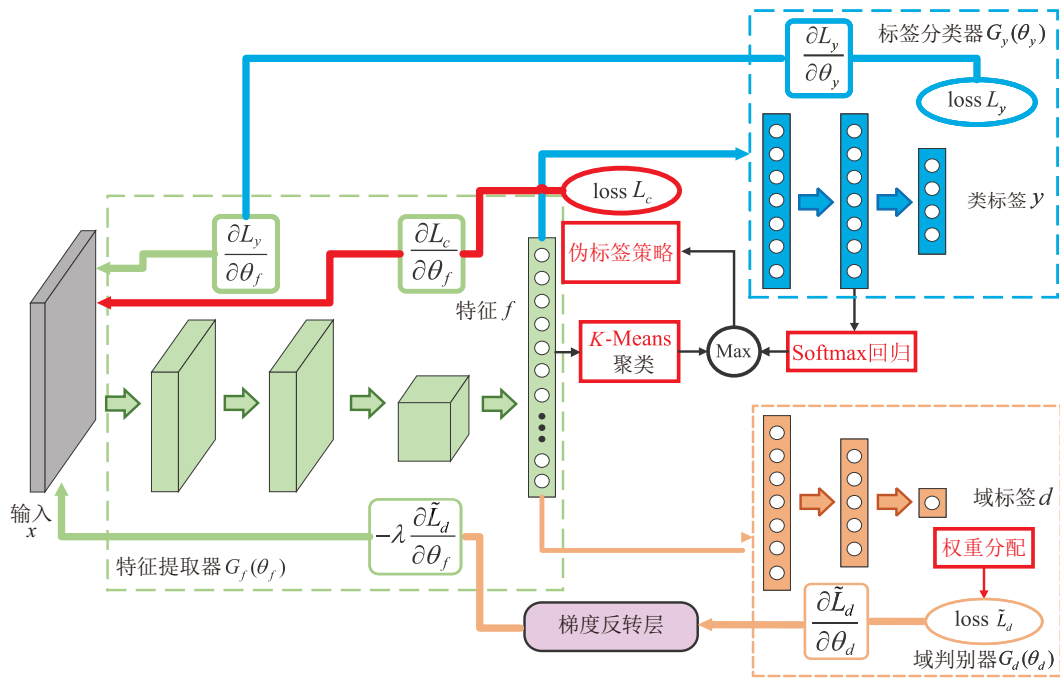


图3 JADAN 整体架构

基于 Softmax 预测和结构化预测联合的伪标签生成策略,其分别从概率角度和空间分布角度为目标域样本生成伪标签.这种双重策略能够确保伪标签的可靠性和稳定性,特别是当目标域样本与源域样本的分布较接近时,Softmax 预测能够提供准确的类别预测;而当目标域样本与源域样本差异较大时,结构化预测可通过相似性分析帮助生成更可信的标签.该伪标签生成策略考虑了多种因素,在计算时通过不同的策略结合,能够避免单一伪标签生成方法带来的偏差.

首先通过特征提取器提取到目标域样本的特征,然后利用标签分类器对目标域样本进行 Softmax 预测获得备选伪标签.目标域样本 x_i^t 属于备选伪标签 y 的条件概率为

$$p_1(y|x_i^t) = \frac{1}{\sum_{l=1}^k \exp(\theta_l^T f_i^t)} \begin{bmatrix} \exp(\theta_1^T f_i^t) \\ \exp(\theta_2^T f_i^t) \\ \vdots \\ \exp(\theta_k^T f_i^t) \end{bmatrix}. \quad (4)$$

其中: k 为类别数, f_i^t 为目标域样本的特征, θ_l^T 为与类别 l 相关的模型参数的转置向量. Softmax 预测能够基于源域分类器的输出直接为目标域样本分配类别.这种方式简单有效,特别适用于目标域与源域分布较为接近的情形.

当目标域与源域样本分布差异较大时,直接使用 Softmax 可能会导致较低的准确率,为应对这一情况,本文引入结构化预测方法.结构化预测获得备选伪标签的方法是采用 K-Means 聚类算法,通过不

断迭代来计算并矫正目标域样本 x_i^t 属于标签 y 的条件概率.如图4所示,主要步骤如下:

- step 1: 随机选取 k 个样本作为聚类中心.
- step 2: 计算其他样本与初始聚类中心的距离,然后将这些样本与他们各自最近的初始聚类中心归为一类.
- step 3: 计算每一类样本的平均值作为新的聚类中心.
- step 4: 判断聚类中心是否发生变化,如果改变,则回到 step 2 继续迭代;如果不变,则达到迭代次数后输出聚类结果.

将 K-Means 算法的最大迭代次数设置为 300,每批目标域样本经过 300 次迭代后获得聚类中心.由于源域样本有标签,可将同一批次中对应的源域样本的类原型作为目标域样本的初始聚类中心,有

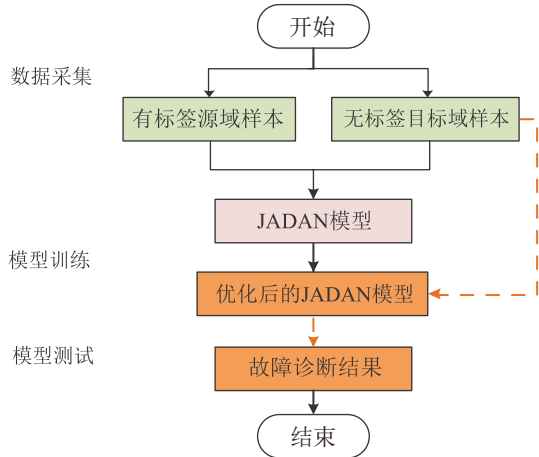


图4 跨工况故障诊断流程

$$\bar{f}_y^s = \frac{\sum_{i=1}^{n_s} f_i^s \delta(y, y_i^s)}{\sum_{i=1}^{n_s} \delta(y, y_i^s)}. \quad (5)$$

其中: f_i^s 为源域样本的特征; 若 $y = y_i^s$, 则 $\delta(y, y_i^s)$ 的值是 1, 否则是 0. 然后通过迭代获得目标域样本的聚类中心 $\bar{f}_y^t$, 目标域样本 x_i^t 属于标签 y 的条件概率为

$$p_2(y|x_i^t) = \frac{\exp(-\|f_i^t - \bar{f}_y^t\|)}{\sum_{l=1}^k \exp(-\|f_i^t - \bar{f}_l^t\|)}. \quad (6)$$

至此, 共得到了两个备选伪标签的条件概率. 通过 Softmax 预测的方法能够给予与源域样本分布差异较小的样本更高的概率, 而通过结构化预测的方法能够给予接近目标域聚类中心的样本更高的概率. 结合两种方法, 可以得到如下的条件概率:

$$p(y|x_i^t) = \max\{p_1(y|x_i^t), p_2(y|x_i^t)\}. \quad (7)$$

于是, 可以得到最终的目标域样本伪标签为

$$\hat{y}_i^t = \arg \max p(y|x_i^t). \quad (8)$$

与源域样本的类原型计算类似, 目标域样本的类原型为

$$\bar{f}_y^t = \frac{\sum_{j=1}^{n_t} f_j^t \delta(y, \hat{y}_j^t)}{\sum_{j=1}^{n_t} \delta(y, \hat{y}_j^t)}. \quad (9)$$

得到伪标签之后即可通过欧式距离定义源域与目标域的类原型之间的距离, 最小化源域与目标域的类原型距离, 类对齐损失能够增强分类决策的稳定性, 从而得到类对齐损失

$$L_c = \frac{1}{k} \sum_{y=1}^k \|\bar{f}_y^s - \bar{f}_y^t\|. \quad (10)$$

类对齐模块的引入进一步弥补了传统域对齐方法的不足. 通过最小化源域与目标域类原型之间的距离, 模型不仅实现了全局的域对齐, 还通过更细粒度的类对齐减少了决策边界附近的误分类问题. 该机制显著提升了模型在跨工况故障诊断任务中的分类准确性和泛化能力, 解决了传统域适应方法在处理复杂类间分布时的不足.

2) 源域样本权重分配机制.

经典的 DANN 方法在进行域对抗训练时没有考虑每个源域样本的权重分配问题, 导致与目标域样本分布差异不同的源域样本都拥有相同的权重, 这会引起诊断模型的负迁移问题. 因此, 基于不同领域之间样本的相似性, 在域判别损失中加入了源域

样本权重分配机制, 赋予了每个源域样本单独的权重来抑制负迁移.

考虑到域判别器对源域与目标域之间相似度较高的样本更难区分, 而对差异性较大的样本更容易区别, 于是以域判别器的域判别误差作为基础, 计算每个源域样本所分配的权重

$$\omega_i^s = -\log(G_d(G_f(x_i^s))), \quad (11)$$

其中 ω_i^s 表示源域中第 i 个样本的权重. 然后, 使用最小最大值归一化, 得到归一化后的权重

$$\tilde{\omega}_i^s = \frac{\omega_i^s - \min(\omega_i^s)}{\max(\omega_i^s) - \min(\omega_i^s)}. \quad (12)$$

源域样本权重分配机制的设计是以域判别器的域判别误差作为基础, 计算每个源域样本所分配的权重, 其反映了源域样本与目标域样本在特征分布上的相似性. 分类误差越小, 说明该源域样本与目标域的分布更为接近, 因此应给予更高的权重, 以增强其在域对齐过程中的影响力. 通过使用域判别器的分类误差来计算源域样本的权重, 能够有效地区分源域样本对目标域迁移的贡献. 与传统方法对所有源域样本赋予相同权重不同, 本文的权重分配机制通过自适应调整源域样本的权重, 能够有效避免负迁移现象, 即避免那些与目标域差异较大的源域样本对模型学习产生干扰. 负迁移在跨域任务中特别常见, 而该机制通过对相似度的度量, 确保了源域样本对目标域特征分布的积极贡献, 从而提高了跨域对齐的效果. 此外, 通过引入域判别器进行动态权重调整, 能够在训练过程中实时优化源域样本的影响力, 提高模型的泛化能力, 确保源域样本对目标域学习的贡献度能够随训练过程不断调整, 减少由于特征分布差异而带来的模型训练不稳定性.

将归一化后的权重代入式 (2), 可以得到新的域判别损失

$$\begin{aligned} \tilde{L}_d(\theta_f, \theta_d) = & -\frac{1}{n_s} \sum_i \tilde{\omega}_i^s \log(G_d(G_f(x_i^s))) - \\ & \frac{1}{n_t} \sum_j \log(1 - G_d(G_f(x_j^t))). \end{aligned} \quad (13)$$

本文将源域样本的权重引入域判别损失, 使得每个源域样本对域对齐损失的贡献与其权重直接相关. 损失优化过程中会更加关注那些与目标域特征分布相似的源域样本, 提升域对齐的效率, 有效减少与目标域分布差异较大的样本对模型的不利影响, 使得对抗训练中, 源域样本的选择更具针对性. 传统的 DANN 模型在损失优化时没有考虑源域样本之

间的差异性,导致模型容易受负迁移的影响,从而降低跨域任务的诊断准确性.本文的损失函数通过权重分配机制,能够针对不同工况下的样本差异进行自适应调整,显著提升跨工况诊断的准确性和泛化性能.

2.2 模型优化

结合类对齐模块和源域样本权重分配机制,JADAN 模型的最终优化目标函数为

$$L_J(\theta_f, \theta_y, \theta_d) = L_y(\theta_f, \theta_y) - \lambda \tilde{L}_d(\theta_f, \theta_d) + \eta L_c(\theta_f), \quad (14)$$

其中 η 表示类对齐损失的正则化参数.

在 JADAN 模型的训练过程中,通过最小化标签分类损失 L_y 和类对齐损失 L_c ,同时最大化域判别损失 L_d 来优化特征提取器的参数;通过最小化标签分类损失 L_y 来优化标签分类器的参数;通过最小化域判别损失 L_d 来优化域判别器的参数.因此,模型的参数优化目标为

$$(\hat{\theta}_f, \hat{\theta}_y) = \arg \min_{\theta_f, \theta_y} L_J(\theta_f, \theta_y, \hat{\theta}_d), \quad (15)$$

$$\hat{\theta}_d = \arg \max_{\theta_d} L_J(\hat{\theta}_f, \hat{\theta}_y, \theta_d), \quad (16)$$

其中 $\hat{\theta}_f$ 、 $\hat{\theta}_y$ 和 $\hat{\theta}_d$ 为最佳参数.

对所提出的模型使用随机梯度下降算法进行梯度更新,参数 θ_f 、 θ_y 和 θ_d 将更新如下:

$$\theta_f = \theta_f - \mu \left(\frac{\partial L_y}{\partial \theta_f} - \lambda \frac{\partial \tilde{L}_d}{\partial \theta_f} + \eta \frac{\partial L_c}{\partial \theta_f} \right), \quad (17)$$

其中 μ 为学习率.

2.3 模型结构

在 JADAN 模型中,特征提取器的具体结构是 5 个一维卷积层和 Flatten 层.其中第 1 个卷积层是大卷积核,目的是提取短时特征,其余的卷积层都是 3×1 的小卷积核.每个卷积层后都会经过批量归一化(BN)层,然后使用 Relu 激活函数,再进行最大池化操作,最后通过 Flatten 层展开提取特征.标签分类器和域判别器的具体结构都是全连接层,区别是前者有 3 个全连接层,而后者增加了 1 个全连接层,且第 3 个全连接层后使用 LeakyRelu 激活函数代替 Relu 激活函数.

2.4 故障诊断方法流程

为了解决源域和目标域样本工况不一致导致诊断模型不适用的问题,提出一种基于 JADAN 模型的故障诊断方法.整个方法主要由 3 个部分组成,分别是数据采集、模型训练和模型测试.主要步骤如下:

step 1: 数据采集.确定采集的机床零部件对象,分别在不同工况的情形下采集其振动信号,并将振

动信号分为源域和目标域样本.

step 2: 模型训练.将有标签源域样本和无标签目标域样本都输入 JADAN 模型中,根据第 2.2 节所介绍的方法进行模型参数的优化.

step 3: 模型测试.将目标域样本输入已经优化参数后的 JADAN 模型中进行分类,最终输出故障诊断结果.

3 实验验证

3.1 数据集介绍

1) CWRU 轴承数据集.

该数据集来自 CWRU 的轴承数据中心^[15-16],已被广泛应用于验证故障诊断方法.在实验中,通过电火花加工在 SKF6205 深沟球轴承中引入内圈故障、外圈故障和滚动体故障,故障尺寸分别为 0.177 8 mm、0.355 6 mm 和 0.533 4 mm.使用连接在驱动端电机和风扇端电机上的加速度传感器来收集振动数据,用于分析的数据来自驱动端的轴承,以 12 kHz 的采样频率采集轴承数据.将正常轴承归为一类,并为每类故障轴承定义 3 个标签,共 10 个标签.

本实验 4 种负载情况下使用的数据集如表 1 所示.其中:工况 A 表示负载为 0 HP,工况 B 表示负载为 1 HP,工况 C 表示负载为 2 HP,工况 D 表示负载为 3 HP.每个状态样本长度为 1 024,每类状态样本数量为 180,其中训练集数量为 100,测试集数量为 80.

表1 工况 A、B、C、D 下的轴承数据集

状态	故障直径/mm	标签	数量
正常	0	0	180
球故障	0.1778	1	180
球故障	0.3556	2	180
球故障	0.5334	3	180
内圈故障	0.1778	4	180
内圈故障	0.3556	5	180
内圈故障	0.5334	6	180
外圈故障	0.1778	7	180
外圈故障	0.3556	8	180
外圈故障	0.5334	9	180

2) 东南大学齿轮箱数据集.

东南大学齿轮箱数据集通过实验台使用 7 种型号为 608A11 的振动传感器分别以 5 120 Hz 的采样频率,采集行星齿轮箱和减速齿轮箱的 x 、 y 、 z 三轴以及电机 z 轴上的振动信号.齿轮共有 5 种状态,分别为正常、断齿、缺齿、齿根裂纹以及齿轮表面磨损.

本实验选取行星齿轮箱 x 轴上的振动信号作为实验对象,数据集如表 2 所示,共研究了两种工况,转速系统负载设置为 20 Hz-0 V 和 30 Hz-2 V.其中

工况 E 表示负载为 20 Hz-0 V, 工况 F 表示负载为 30 Hz-2 V.

表2 工况 E、F 下的齿轮箱数据集		
状态	标签	数量
正常	0	180
断齿	1	180
缺齿	2	180
齿根裂纹	3	180
表面磨损	4	180

3.2 CRWU 轴承数据集结果与分析

为了验证本文所提出 JADAN 模型的有效性, 将所提出模型与以下模型进行对比:

- 1) 卷积神经网络 (CNN): 基准模型, 不涉及迁移学习的内容.
- 2) 自适应批量归一化 (AdaBN)^[17]: 将 CNN 模型中的 BN 层替换为 AdaBN 层, 减小源与目标域的距离.
- 3) 深度域混淆 (DDC)^[18]: 利用 MMD 和源域上的分类损失来学习域不变特征.
- 4) 深度适应网络 (DAN)^[19]: 在 DDC 基础上适配了多层网络, 并将 MMD 替换为多核 MMD.
- 5) DANN: 利用对抗的思想训练特征提取器, 使其提取域不变特征.

为保证对比实验的公平, 这 5 种模型的结构和参数都是相同的, 均在 Pytorch 深度学习框架下实现, 共进行 12 组跨工况任务. 实验结果都是 10 次实验的平均值, 表 3 展示了跨工况下不同模型的诊断结果. 可以看出, 本文所提出的 JADAN 模型诊断准

确率最高, 平均值达到了 98.17 %, 显示了 JADAN 模型的优越性. 通过结果对比可知, 没有加入迁移学习思想的 CNN 模型准确率最低, 只有 77.31 %, 而迁移学习模型比 CNN 模型在跨工况诊断中的效果更好; 进一步分析结果, 除了工况 A→D 和工况 B→D, DANN 模型的准确率都比 AdaBN、DDC 和 DAN 模型有所提高, 说明在迁移学习中引入对抗机制可以进一步加强跨工况诊断的能力.

表3 跨工况下不同模型在轴承数据集的诊断结果							%
任务	CNN	AdaBN	DDC	DAN	DANN	JADAN	
工况A→B	83.38	87.21	87.66	93.08	95.52	98.88	
工况A→C	76.62	85.45	78.98	92.12	92.72	97.75	
工况A→D	71.43	83.28	74.28	91.42	90.53	96.25	
工况B→A	79.88	84.01	86.03	92.62	93.64	96.72	
工况B→C	80.97	88.17	87.52	94.18	94.32	99.88	
工况B→D	75.38	83.29	82.25	93.54	92.28	98.36	
工况C→A	74.02	84.41	79.22	91.74	93.62	97.28	
工况C→B	76.98	87.88	86.75	93.18	95.86	98.78	
工况C→D	75.38	86.91	85.24	92.56	93.88	98.12	
工况D→A	75.66	87.62	80.45	94.56	96.58	97.54	
工况D→B	78.32	89.28	84.88	96.77	97.34	98.76	
工况D→C	79.64	91.12	87.54	97.13	97.98	99.75	
均值	77.31	86.55	83.4	93.56	94.52	98.17	

为进一步体现 JADAN 模型的优势, 采用 t 分布随机邻居嵌入 (t -SNE) 技术对特征提取器最后一层的输出进行降维, 以工况 B→C 为例, 将源域和目标域的数据特征降到二维进行可视化. 图 5 展示了不同

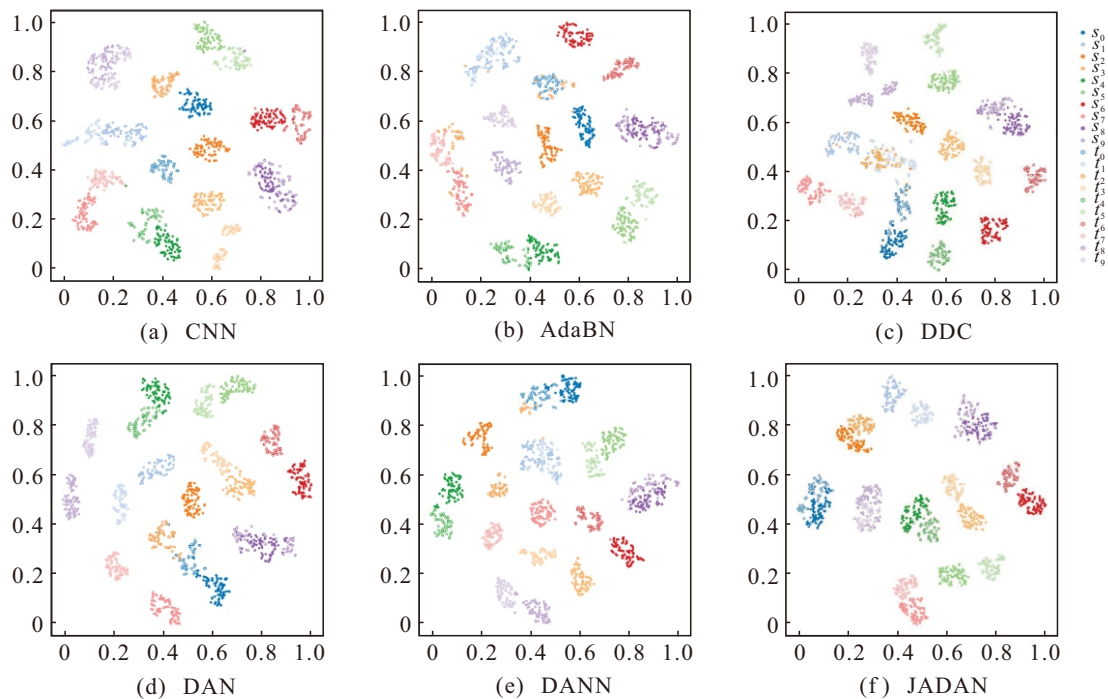


图5 不同模型在轴承数据集的特征可视化

对比模型和本文所提出 JADAN 模型的可视化结果. 从图 5(a)、5(b) 和 5(c) 中可以看出, 有些源域和目标域的特征存在相互重叠的现象, 说明特征提取器并不能完全将源域和目标域的共同特征提取出来. 而从图 5(d) 和 5(e) 中可以看出, 虽然仍存在两个域特征重叠的现象, 但明显少了很多, 而且有些源域和目标域特征已经融合在一起. 进一步观察图 5(f), 可以看到两个域的特征基本完全融合在一起, 并且没有出现特征重叠的现象, 相同类别的源域与目标域之间特征的间距也非常小, 体现了加入类对齐模块和源域样本权重分配机制的重要性.

为了评估 JADAN 模型中改进的两个模块的重要性, 通过在 JADAN 模型中分别移除这两个组件进行消融实验. 实验分别在工况 A→B、工况 A→C 和工况 A→D 中进行, 分别移除类对齐模块和权重分配模块, 诊断准确率如图 6 所示. 可以看出, 当移除类对齐模块或权重分配机制后, 诊断准确率明显下降, 在移除前者时下降得更多, 这也充分说明了这两个组件的重要性, 同时类对齐模块更加重要.

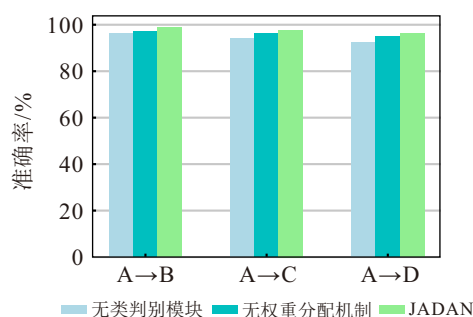


图6 轴承数据集消融实验结果

3.3 东南大学齿轮箱数据集结果与分析

为了验证 JADAN 模型的泛化性, 在东南大学齿轮箱数据集上进行实验^[16]. 对比方法与上一小节的 5 种方法相同, 由于该数据集只有两种工况, 因此共进行两组跨工况任务, 重复 10 次取其平均值作为每次任务的实验结果, 表 4 展示了跨工况下不同方法的诊断结果.

表4 跨工况下不同模型在齿轮箱数据集的诊断结果

任务	CNN	AdaBN	DDC	DAN	DANN	JADAN
工况E→F	53.42	77.82	73.53	89.75	91.74	99.98
工况F→E	50.25	71.75	68.49	85.42	87.75	96.82
均值	51.84	74.79	71.01	87.59	89.75	98.4

从表 4 中可以看出, 在这两种工况的任务下, 所有使用了迁移学习思想的方法诊断准确率都比 CNN 模型高, 这也验证了迁移学习在跨工况问题中的必要性. 同时, 在迁移学习的方法中, DAN 和 DANN

模型整体比 AdaBN 和 DDC 模型的诊断效果好, 在两种工况的平均准确率分别达到了 87.59 % 和 89.75 %. 特别地, 本文所提出的 JADAN 模型准确率最高, 这说明了其可以很好地解决跨工况问题.

4 结论

本文提出了一种基于 JADAN 的跨工况故障诊断方法, 通过对抗训练提取源域和目标域的不变特征, 解决了实际工程中常见的跨工况问题. 通过在 DANN 模型中引入类对齐模块, 最小化源域和目标域同一类别的差异, 同时实现了域对齐和类对齐. 此外, 在域判别损失函数中加入了源域样本权重分配机制, 有效地抑制了负迁移问题. 实验结果表明了本文所提出的方法在跨工况诊断中的有效性, 并且通过消融实验验证了模型中所加入两个模块的重要性. 由于本文假设源域与目标域的标签是一致的, 但经常会出现目标域标签增加的情况; 因此, 在今后工作中需要在这种情况下设计新的故障诊断模型.

参考文献 (References)

- [1] 周文惠, 齐瑞云, 姜斌. 面向突发故障的分布式多无人机任务重规划方法[J]. 控制与决策, 2023, 38(5): 1373-1385.
(Zhou W H, Qi R Y, Jiang B. Mission replanning method of distributed multiple unmanned aerial vehicles for pop-up faults[J]. Control and Decision, 2023, 38(5): 1373-1385.)
- [2] 姜斌, 吴云凯, 陆宁云, 等. 高速列车牵引系统故障诊断与预测技术综述[J]. 控制与决策, 2018, 33(5): 841-855.
(Jiang B, Wu Y K, Lu N Y, et al. Review of fault diagnosis and prognosis techniques for high-speed railway traction system[J]. Control and Decision, 2018, 33(5): 841-855.)
- [3] 姜斌, 孟庆开, 杨浩. 航天器姿轨控制研究综述: 微分几何控制方法[J]. 控制与决策, 2023, 38(8): 2079-2092.
(Jiang B, Meng Q K, Yang H. A survey on spacecraft attitude and orbit control: Differential geometric control approaches[J]. Control and Decision, 2023, 38(8): 2079-2092.)
- [4] Xue R J, Zhang P S, Huang Z G, et al. Digital twin-driven fault diagnosis for CNC machine tool[J]. The International Journal of Advanced Manufacturing Technology, 2024, 131(11): 5457-5470.
- [5] 任浩, 屈剑锋, 柴毅, 等. 深度学习在故障诊断领域中的研究现状与挑战[J]. 控制与决策, 2017, 32(8): 1345-1358.
(Ren H, Qu J F, Chai Y, et al. Deep learning for fault diagnosis: The state of the art and challenge[J]. Control and Decision, 2017, 32(8): 1345-1358.)
- [6] 王村松, 陆宁云, 程月华, 等. 基于无标签、不均衡、初值不确定数据的设备健康评估方法[J]. 控制与决策,

- 2020, 35(11): 2687-2695.
(Wang C S, Lu N Y, Cheng Y H, et al. Equipment health risk assessment based on unlabeled, unbalanced data under uncertain initial condition[J]. Control and Decision, 2020, 35(11): 2687-2695.)
- [7] Shao L X, Lu N Y, Jiang B, et al. Improved generative adversarial networks with filtering mechanism for fault data augmentation[J]. *IEEE Sensors Journal*, 2023, 23(13): 15176-15187.
- [8] Zhao J T, Huang W T. Transfer learning method for rolling bearing fault diagnosis under different working conditions based on CycleGAN[J]. *Measurement Science and Technology*, 2022, 33(2): 025003.
- [9] Li Z M, Wang X H, Yang R. Fault diagnosis of bearings under different working conditions based on MMD-GAN[C]. The 33rd Chinese Control and Decision Conference. Kunming, 2021: 2906-2911.
- [10] Xiao Y Q, Wang J Q, He Z M, et al. Deep transfer learning with metric structure for fault diagnosis[J]. *Knowledge-Based Systems*, 2022, 256: 109826.
- [11] Liu C Y, Gryllias K. Simulation-driven domain adaptation for rolling element bearing fault diagnosis[J]. *IEEE Transactions on Industrial Informatics*, 2022, 18(9): 5760-5770.
- [12] Tan Y W, Guo L, Gao H L, et al. MiDAN: A framework for cross-domain intelligent fault diagnosis with imbalanced datasets[J]. *Measurement*, 2021, 183: 109834.
- [13] Guo L, Lei Y G, Xing S B, et al. Deep convolutional transfer learning network: A new method for intelligent fault diagnosis of machines with unlabeled data[J]. *IEEE Transactions on Industrial Electronics*, 2019, 66(9): 7316-7325.
- [14] Wan L J, Li Y Y, Chen K Y, et al. A novel deep convolution multi-adversarial domain adaptation model for rolling bearing fault diagnosis[J]. *Measurement*, 2022, 191: 110752.
- [15] 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.
- [16] 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.
- [17] Zhai X D, Qiao F, Ma Y M, et al. A novel fault diagnosis method under dynamic working conditions based on a CNN with an adaptive learning rate[J]. *IEEE Transactions on Instrumentation and Measurement*, 2022, 71: 3177233.
- [18] Zhao B X, Cheng C M, Zhang G Z, et al. An instance and feature-based hybrid transfer model for fault diagnosis of rotating machinery with different speeds[J]. *IEEE Transactions on Instrumentation and Measurement*, 2022, 71: 3210965.
- [19] 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.

作者简介

文利燕 (1987-), 女, 副教授, 博士生导师, 主要研究方向为自适应控制及其应用、故障诊断与容错控制, E-mail: wenliyan8710@163.com;

陈金陵 (1997-), 男, 硕士, 主要研究方向为故障诊断与容错控制, E-mail: jlchen17@nuaa.edu.cn;

姜斌 (1966-), 男, 教授, 博士生导师, 主要研究方向为多智能体系统智能故障诊断与容错控制技术、飞控与导航系统故障检测技术, E-mail: binjiang@nuaa.edu.cn;

马亚杰 (1987-), 男, 教授, 博士生导师, 主要研究方向为基于数据驱动的故障检测与健康管理、飞控与导航系统故障检测技术, E-mail: yajiema@nuaa.edu.cn;

冯靖智 (2000-), 男, 硕士生, 主要研究方向为故障诊断与容错控制, E-mail: fengjingzhi@nuaa.edu.cn.