



非平衡数据下基于自适应迁移与宽度学习的多工况故障诊断

姜午恺¹, 周平^{1,2*}

1. 东北大学流程工业综合自动化国家重点实验室, 沈阳 110819

2. 中国矿业大学煤炭加工与高效洁净利用教育部重点实验室, 徐州 221116

* 通信作者. E-mail: zhouping@mail.neu.edu.cn

收稿日期: 2022-03-24; 修回日期: 2022-05-17; 接受日期: 2022-06-27; 网络出版日期: 2023-05-12

国家自然科学基金 (批准号: 61890934, U22A2049, 61790572, 61991400) 和兴辽英才项目 (批准号: XLYC1907132) 资助

摘要 针对高维不平衡数据的多工况故障诊断难题, 提出一种非平衡数据条件下基于动态自适应迁移的宽度学习方法 (TL-BLS), 并应用于污水处理过程的多工况故障诊断. 首先, 针对工业过程中不同工况下存在数据分布差异过大的问题, 采用迁移学习流形变换方法拉近源域数据和目标域数据的距离; 其次, 针对后续自适应迁移中需要伪标签的条件, 提出考虑数据不平衡的两阶段目标域伪标签预测方法, 第一阶段使用基于类间关系的伪标签方法得到可靠度高的伪标签, 第二阶段使用考虑子集差异的集成方法来获得可靠度更高且兼顾数据不平衡问题的伪标签; 最终, 建立带流形正则化项和具有动态平衡因子的自适应迁移与宽度学习方法, 并在宽度学习系统和自适应迁移中引入权重矩阵来解决数据非平衡的问题, 实现多工况过程下的故障诊断. 污水处理过程数据实验及其比较分析表明所提方法具有更优的故障识别效果和更好的实用性.

关键词 故障诊断, 迁移学习, 宽度学习系统, 流形学习, 集成学习, 污水处理过程, 非平衡数据

1 引言

现代工业已经迈入高质量、低碳发展阶段, 保证工业生产长期稳定顺行是实现工业高质量可持续发展的前提条件. 日常操作不当、天气变化、原(燃)料波动等众多不确定因素的影响, 使得工业生产往往出现各种各样的异常工况或者故障, 影响系统运行性能和安全. 因此, 能够及时发现生产异常的故障诊断技术是现代工业过程安全稳定运行和高质量低碳生产不可或缺的组成部分^[1,2]. 近年来, 随着计算机技术与人工智能技术的蓬勃发展以及工业互联网与新型传感技术在工业领域的广泛应用, 工业数据的规模和质量迅速提高, 越来越多的研究都开始探讨基于数据驱动的故障诊断方法. 相对于传

引用格式: 姜午恺, 周平. 非平衡数据下基于自适应迁移与宽度学习的多工况故障诊断. 中国科学: 信息科学, 2023, 53: 946–971, doi: 10.1360/SSI-2022-0118
Jiang W K, Zhou P. Multi-condition fault diagnosis based on adaptive transfer and broad learning under imbalanced data (in Chinese). Sci Sin Inform, 2023, 53: 946–971, doi: 10.1360/SSI-2022-0118

统基于模型的故障诊断方法, 基于数据驱动的故障诊断方法不需要建立生产过程复杂机理的精确数学模型, 因此更具有实用性, 目前已经成为学术界和工业界的研究热点^[3~6].

常见的数据驱动故障诊断方法有机器学习类方法、信号处理类方法以及多元统计分析类方法等. 相对于其他两类较早出现的方法, 机器学习类方法是近年随着人工智能技术发展起来的一类新型故障诊断方法, 主要采用各种机器学习算法对被诊断动态系统进行数据分析、处理和建模, 从而实现数据驱动的故障诊断, 该类故障诊断方法发展极其迅速, 并广泛应用于各个工业领域. 如文献 [7] 提出一种基于饱和铁芯超导故障限流器控制系统模型的支持向量机方法, 在传感器故障诊断上取得了良好效果. 文献 [8] 提出一种基于敏感性分析的新型加权随机神经网络算法, 并在大型旋转机械实验平台上进行故障诊断的实验验证. 文献 [9] 应用最近邻算法与支持向量机对感应电机进行电气和机械的故障诊断, 并得到了良好的诊断效果.

随着现代大规模工业生产对低碳高质量发展的迫切需求, 对故障诊断的准确度和性能要求也不断提升, 因而应用深度神经网络解决故障诊断问题也越来越受关注, 成为了机器学习领域一大研究热点. 如文献 [10] 提出一种基于对抗网络的深度学习方法对恶劣条件下的工作设备进行故障诊断. 文献 [11] 将多标签分类的概念引入故障诊断, 并提出一种基于多标签卷积神经网络的深度多标签学习框架. 文献 [12] 提出一种基于张量分解与深度神经网络模型相结合的自适应顺序故障诊断方法, 对工业生产过程进行监控. 这些基于深度学习的故障诊断方法是针对大规模复杂工业系统的运行数据维度较高, 常规浅层神经网络的故障分类结果精度不够, 无法对多特征复杂数据进行准确处理等问题而提出的. 然而, 深度神经网络所构建的模型结构复杂, 不仅在训练过程中常常面临训练时间过长的的问题, 而且在网络层数等各种超参数设置上也没有严格的理论依据, 严重依赖操作者的经验, 因而有着可解释性差的缺点. 为解决这些问题, Chen 等^[13] 提出了宽度学习系统 (broad learning system, BLS) 作为深度神经网络的替代. 宽度学习系统作为一种神经网络横向扩展的高效增量型学习系统, 可有效解决深度学习权值更新慢, 无监督、半监督网络分类成功率低的问题. 宽度学习系统以随机神经网络为映射特征, 其本质结构为单隐层前馈型神经网络, 通过神经网络增强节点映射特征并将增强节点直接连接到输出端. 众多研究表明: 相对于深度学习网络, 宽度学习神经网络的模型结构更加简单, 性能更加完善, 不仅保留了深度学习精度高、泛化性能好的优点, 而且弥补了深度学习反向传播耗时长、收敛慢的缺陷^[14]. 因此, 宽度学习一经提出, 就得到了广泛关注和应用. 如文献 [15] 提出在线半监督的宽度学习算法, 对空气压缩机工作过程进行故障诊断, 文献 [16] 构建了基于改进主成分分析与宽度学习系统的增强型故障诊断架构, 并在高速列车牵引平台上验证了算法的有效性.

由于实际工业生产过程处于长时间平稳运行的状态, 包含故障的样本在数量级上难以与正常运行数据相匹配, 从而使得用于诊断建模的过程数据存在类别不平衡的问题^[17, 18]. 在应用常规机器学习方法进行故障诊断时, 一般会给每个样本相同的权重, 这样会导致在诊断建模时为了获得更高的准确率, 而忽略少数样本所对应的类别^[19, 20]. 所以, 为了提高少数类样本的建模精度, 需要首先解决数据样本不平衡问题. 实际上, 正常数据样本数与故障数据样本数的不平衡是数据不平衡学习问题的典型情况, 也是机器学习领域的又一研究热点^[21, 22]. 通常, 解决非平衡数据分类问题可以从算法和数据两方面入手: 算法方面, 可以通过相关技术来修改已有算法, 以适应数据样本不平衡的问题, 如代价敏感学习^[23]、单类学习^[24] 和加权集成学习^[25]; 而从数据层面, 可以通过改变原始数据的分布来解决数据不平衡的问题, 改变数据分布的方法主要为各种抽样方法, 如重采样、欠采样等^[26, 27]. 此外, 集成学习也经常与抽样方法结合来处理数据不平衡的问题^[28].

实际工业系统的运行环境和外部干扰动态多变, 会造成运行工况复杂迥异, 即实际工业过程具有多工况特性. 不同工况下的数据在分布上会存在一定差异, 有时会造成数据的统计特性明显不同. 对

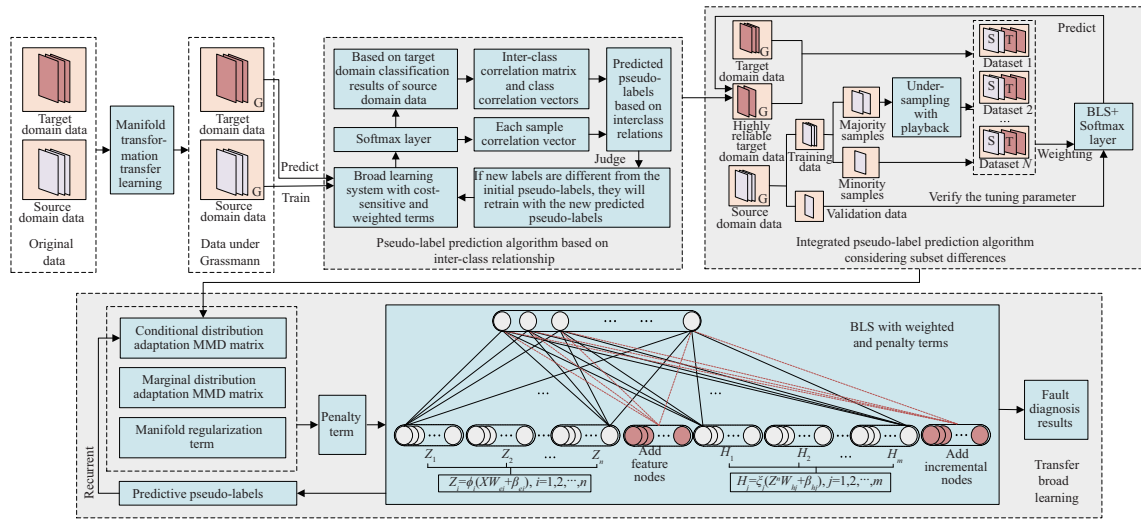


图 1 (网络版彩图) 所提算法结构图

Figure 1 (Color online) Structure diagram of the proposed algorithm

于数据驱动故障诊断方法来说, 不同工况下数据的差异性意味着即使故障诊断模型在某种工况下经过了充分训练, 但是很可能在另一种工况下的诊断效果会很差 [29~32]。针对这一问题, 如何建立具有良好泛化能力的诊断方法有着十分重要的实际工程意义。最简单的方法就是针对不同的工况建立不同的诊断模型, 但这显然会显著加大诊断建模的工作量和时效。目前, 应用迁移学习对不同工况下的数据进行分析是一种更为有效的方法。迁移学习作为一种跨领域、跨任务的学习方法, 能够复用旧数据解决不同场景下的新问题, 已在多工况故障诊断领域得到一些应用。如文献 [33] 提出一种结合迁移稀疏编码 (transfer sparse coding, TSC) 和联合几何和统计对齐 (joint geometric and statistical alignment, JGSA) 特点的迁移学习故障诊断方法, 并对变工况下滚动轴承的故障诊断进行了研究。文献 [34] 提出一种具有源域和目标域加权机制的混合方法, 通过使用额外的异常值标识符, 可以实现自动识别未知故障模式。此外, Dong 等 [35] 利用统计分布和几何空间联合调整, 适配两域分布, 并考虑了样本的类内距和类间距, 实现了变工况下的故障诊断。由于数据中的流形结构蕴含潜在的判别信息, 在适配两域分布的同时对复杂数据分布流形结构进行挖掘成为应对不同工况下复杂分布差异的关键。流形正则化能够通过无标记数据挖掘样本点间的流形几何结构, 考虑数据点间的几何结构和数据潜在的判别信息 [36]。因此, 自适应正则化迁移学习 (adaptation regularization transfer learning, ARTL) 将联合分布适配和流形正则化嵌入结构风险最小化框架中 [37, 38], 可构建具有强鲁棒性的自适应诊断模型, 从而有效挖掘复杂数据分布流形隐含的信息, 为多工况复杂系统数据分布差异的故障诊断提供了一种新的思路。

综上, 对上述问题, 本文提出一种非平衡数据条件下基于迁移学习与宽度学习的新型多工况故障诊断方法, 具体算法框架如图 1 所示。

(1) 针对工业过程运行存在的多工况问题, 采用迁移学习流形变换的方法对源域和目标域数据进行 Grassmann 高维流形下域间距离的拉近, 为后续的迁移方法和故障分类提供便利和基础。

(2) 针对流形变换之后构建条件分布适配矩阵需要目标域伪标签的问题, 提出了不平衡数据条件下两阶段的目标域伪标签预测算法, 第一阶段使用基于类间关系的伪标签预测算法, 得到高可靠性的伪标签, 第二阶段在第一阶段的基础上使用考虑子集差异的集成方法来获得可靠度更高且兼顾数据不

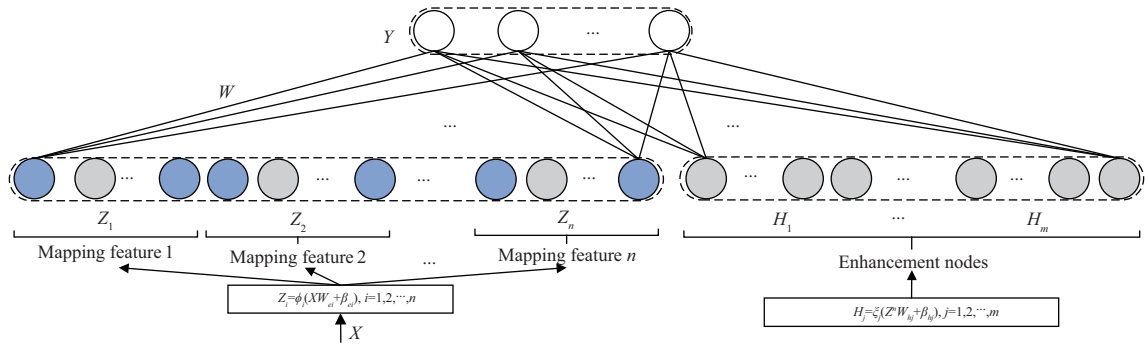


图 2 (网络版彩图) 宽度学习系统结构图

Figure 2 (Color online) Structure diagram of broad learning system

平衡问题的伪标签, 得到可靠度高的伪标签的同时避免了不平衡问题, 提高了最终伪标签的预测精度.

(3) 为了同时利用源域和目标域数据中所含的流形信息和域差异信息, 建立带流形正则化项和动态分布适配惩罚的宽度学习算法, 并使用加权方法来解决分类时数据不平衡的问题. 改进后的迁移宽度学习算法能够同时应用流形空间与域迁移方面的信息, 因而可达到更好的效果, 实现高精度故障诊断.

(4) 将所提方法应用到多工况污水处理过程中, 通过充分的数据实验和比较分析验证了所提新型故障诊断方法的有效性、先进性和实用性.

2 预备知识

2.1 宽度学习系统

与深度学习结构不同, 宽度学习系统是在随机权神经网络 (random vector functional link network, RVFLN) 的基础上发展起来的单隐层网节点横向扩展模型^[13]. BLS 网络结构如图 2 所示, 传统 RVFLN 直接将样本数据和生成的增强节点作为网络的输入特征, 而 BLS 对原始数据进行一系列特征提取, 并将映射的特征作为一层, 称为特征层. BLS 将特征层增强为一个随机生成的权重增强节点, 称为增强层. 然后将特征层的所有特征节点和增强层的增强节点直接连接到输出端, 称为输出层. 假设学习的训练样本为 $\{x_i | x_i \in \mathbb{R}^M, i = 1, \dots, N\}$, 相应标签为 $\{y_i | y_i \in 1, \dots, C\}$, 其中, N 表示训练样本个数, C 表示样本种类个数. 由 n 组特征节点和 m 组增强节点组成的 BLS 构造如图 2 所示.

首先, BLS 通过稀疏特征映射得到 i 组特征节点共同构成特征层, 如下所示:

$$Z_i = \phi_i(XW_{ei} + \beta_{ei}), \quad i = 1, 2, \dots, n, \quad (1)$$

其中, W_{ei}, β_{ei} 是通过随机生成的权重和偏差, ϕ_i 是非线性激活函数, ei 表示特征层节点数, n 为特征窗口数.

接着, 将这些随机生成的映射特征组收集在一起, 记为 $Z^n = [Z_1, Z_2, \dots, Z_n]$. BLS 将这些组映射特征组成的 Z^n 通过非线性激活函数扩展增强层, 如下所示:

$$H_j = \xi_j(Z^n W_{hj} + \beta_{hj}), \quad j = 1, 2, \dots, m, \quad (2)$$

其中, $\xi_j(\cdot)$ 为非线性激活函数. 在 BLS 学习的第二阶段, 计算隐含层和输出层之间的权值. 增强节点的输出可以表示为 $H^m = [H_1, H_2, \dots, H_m]$. 将映射特征节点和增强节点连接成一个矩阵, 最终 BLS

可以表示为

$$Y = [Z_1, Z_2, \dots, Z_n | H_1 \cdots H_m] W^m = [Z^n | H^m] W^m. \quad (3)$$

输出层权值可以快速计算为

$$W^m = [Z^n | H^m]^+ Y. \quad (4)$$

然而, 由于训练数据的维数和计算速度问题, 采用正交投影、迭代和奇异值分解等标准方法来计算广义逆的代价太大, 因此采用如下方法来求解伪逆:

$$\arg \min_W : \|AW - Y\|_v^{\sigma_1} + \lambda \|W\|_u^{\sigma_2}, \quad (5)$$

其中, $\sigma_1 > 0$, $\sigma_2 > 0$, v 和 u 是正则化范数. 该方法在原广义逆估计 $A = [Z^n | H^m]$ 的病态条件下, 在原广义逆估计的最小二乘估计中加入约束项 λ , 使原广义逆估计的伪逆得以求出. 当 $\lambda = 0$ 时, 逆问题退化为最小二乘问题, 并得到原伪逆问题的解. 取 $\sigma_1 = \sigma_2 = v = u = 2$, 将上述最优问题设置为岭回归学习算法. 因此, BLS 的连接权重可以近似为

$$W^m = (\lambda I + AA^T)^{-1} A^T Y, \quad (6)$$

$$A^+ = \lim_{\lambda \rightarrow 0} (\lambda I + AA^T)^{-1} A^T. \quad (7)$$

2.2 集成学习

集成学习是通过集成多个弱分类器来完成学习训练任务的一种常用的机器学习方法^[39]. 该方法将多个分类器结合, 可以提高模型的泛化性能. 本文引入 EasyEnsemble 集成学习方法^[40] 对多数类样本进行欠采样, 从而解决数据不平衡的问题, 并集成多个子分类器作为最终的强分类器, 以充分挖掘不平衡数据集的潜在有用信息, 算法结构如图 3 所示. 考虑应用集成学习解决多分类问题, 记含有 K 类的不平衡数据集类别为 $1, 2, \dots, K$, 样本数目最少的一类为样本集 N , 样本数目为 $|N|$, 其余多数类样本集为 N_1, N_2, \dots, N_{K-1} . 对所有多数类样本集重复 i 次有放回的随机独立欠采样, 第 i 个多数类样本集 N_i 在第 t 次独立重复采样得到子集 N_i^t , 且满足样本数量 $|N_i^t| = |N|$, 其中 $i = 1, 2, \dots, K-1$, $t = 1, 2, \dots, T$.

在上述采样过程中, 多数类样本 N_i 中的任意样本 $(x_j^i, y_j^i) \in N_i$ 至少一次被采集到子集 N_i^t 中的概率为 $1 - (1 - 1/|N_i|)^{|N|}$, 总体而言, 其在 T 个训练子集中至少出现一次的可能性 P_{one} 和全部出现的可能性 P_{all} 分别为

$$P_{\text{one}} = 1 - \left(1 - \frac{1}{|N_i|}\right)^{|N|T}, \quad (8)$$

$$P_{\text{all}} = \left(1 - \left(1 - \frac{1}{|N_i|}\right)^{|N|T}\right)^T. \quad (9)$$

在保证最少类数目不变的条件下, 采样次数越多, 非最少类的样本在全体训练子集中至少出现一次的可能性 P_{one} 就越大, 从而可保证绝大部分样本被采样到训练集中, 防止欠采样过程中缺少和损失部分样本. 同时, 在存在部分较少类样本集样本数与 $|N|$ 数量相差不大的情况下, 根据式 (9), 较少类样本 (x_j^i, y_j^i) 在全体训练子集中出现的可能性 P_{all} 相对于相差倍数更大的多数类样本而言更高, 也能获得更高的关注权重, 可以保证数目较少的多数类样本的采样质量.

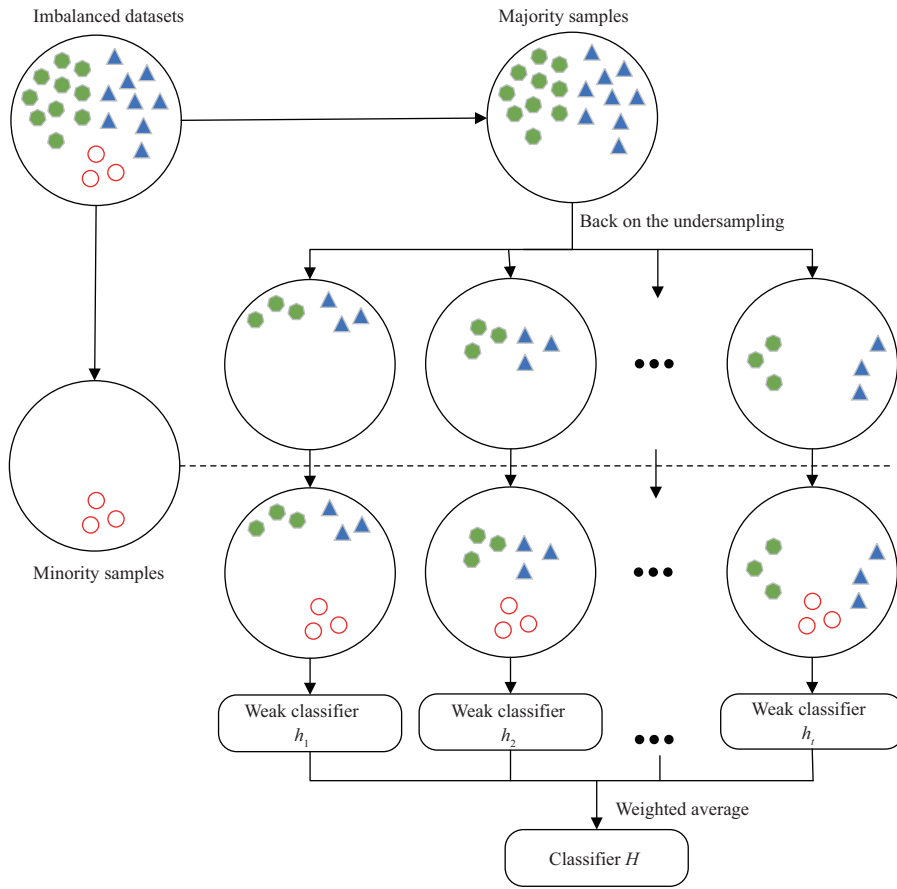


图 3 (网络版彩图) EasyEnsemble 算法结构图

Figure 3 (Color online) Structure diagram of EasyEnsemble algorithm

将第 t 次欠采样过程中得到的多数类样本子集 $N_t = \{N_1^t, N_2^t, \dots, N_{K-1}^t\}$ 与少数类样本集 N 组合为弱分类器 h_t 的训练子集 D_t , 记为 $D_t = N_t \cup N$. 此时训练子集每个类别的数目相同, 即为平衡数据集, 以训练子集 D_t 作为输入, 对每一弱分类器 h_t 进行次数为 S_i 的迭代训练. EasyEnsemble 中每一弱分类器的输出可表示为

$$h_t(x) = \arg \max_k \sum_{d=1}^{s_i} \left(\frac{1 - \varepsilon_t}{\varepsilon_t} \right) \|h_{td}(x) = y\|, \quad (10)$$

其中, y 为输入 x 的真实标签, $h_{td}(x)$ 为训练迭代次数为 d 时的输出, ε_t 为弱分类器 h_t 的误差函数, $k = 1, 2, \dots, K$ 表示数据集中的类别.

需要注意的是, EasyEnsemble 的最终分类器是对所有弱分类器 h_t 的参数进行集成, 而不是对弱分类器的输出 $h_t(x)$ 进行结果投票获得最终决策, 最终分类器的输出可表示为

$$H(x) = \arg \max_k \sum_{t=1}^T \sum_{d=1}^{s_i} \left(\frac{1 - \varepsilon_t}{\varepsilon_t} \right) \|h_{td}(x) = y\|. \quad (11)$$

2.3 迁移学习

迁移学习目前主流的方法是数据分布自适应方法, 分为边缘分布自适应和条件分布自适应. 边缘

分布自适应方法的目标是减小源域和目标域的边缘概率分布的距离, 从而完成迁移学习. 从形式上来说, 边缘分布自适应方法是用 $P(x_s)$ 和 $P(x_t)$ 之间的距离来近似两个领域之间的差异的, 用 X_s, X_t 表示源域数据和目标域数据, 即

$$\text{dis}(X_s, X_t) \approx \|P(x_s) - P(x_t)\|. \quad (12)$$

通过引入最大均值差异 (maximum mean discrepancy, MMD) 将式 (12) 转化为 MMD 距离:

$$\text{dis}(X_s, X_t) = \left\| \frac{1}{n_1} \sum_{i=1}^{n_1} \phi(x_i) - \frac{1}{n_2} \sum_{j=1}^{n_2} \phi(x_j) \right\|. \quad (13)$$

根据文献 [41], 通过引入核矩阵以及低维矩阵 W , 同时保证迁移前后数据散度不变, 得到边缘分布自适应的最终优化目标为

$$\begin{aligned} \min_w \text{tr}(W^T K L K W) + \mu \text{tr}(W^T W) \\ \text{s.t. } W^T K H K W = I_m, \end{aligned} \quad (14)$$

其中, $K = \begin{bmatrix} K_{s,s} & K_{s,t} \\ K_{t,s} & K_{t,t} \end{bmatrix}$, H 是中心矩阵, 可以表示为 $H = I_{n_1+n_2} - 1/(n_1+n_2)\mathbf{1}\mathbf{1}^T$.

与边缘分布自适应类似, 条件分布自适应是通过减小源域和目标域的概率分布的距离而完成迁移的, 其目标函数与式 (12) 类似:

$$\text{dis}(X_s, X_t) \approx \|P(y_s|x_s) - P(y_t|x_t)\|. \quad (15)$$

应用贝叶斯 (Bayes) 公式, 并忽略 $P(x_t)$, 可以得到

$$\text{dis}(X_s, X_t) \approx \|P(x_s|y_s) - P(x_t|y_t)\|. \quad (16)$$

从式 (16) 可以得到, 对源域目标域进行条件分布适配距离的计算时需要目标域的标签分布, 而一般目标域没有标签数据, 需要用伪标签来进行代替.

具体到每一对源域目标域而言, 其边缘分布适配和条件分布适配的重要性并不相同, 需要对每对源域目标域分别求其各部分权重, 能够将分布距离表示为

$$\text{dis}(X_s, X_t) \approx (1 - \mu) \|P(x_s) - P(x_t)\| + \mu \|P(y_s|x_s) - P(y_t|x_t)\|, \quad (17)$$

其中, μ 为平衡因子, 为了计算平衡因子, 引入文献 [42] 中的 A 距离, 定义 A 距离为建立一个二分类器进行两个不同领域的分类得出的误差, 公式为

$$d_A(\Omega_s, \Omega_t) \approx 2(1 - 2\varepsilon(h)), \quad (18)$$

其中, $\varepsilon(h)$ 是用线性分类器区分领域 Ω_s, Ω_t 的误差, 应用式 (18) 来计算, 可得平衡因子 μ 为

$$\mu = 1 - \frac{d_M}{d_M + \sum_{c=1}^C d_c}, \quad (19)$$

其中, d_M 表示边缘分布距离, d_c 表示类别 c 的条件分布距离.

3 所提算法

宽度学习系统作为一种代替深度网络结构进行高性能学习的模型, 因为具有结构灵活反应迅速等优势而正在被广泛使用. 为了整体迁移效果和域间距离的拉近, 首先应用流形变换迁移学习的方法得到差异更小的源域和目标域数据. 域适应作为解决子集与目标域之间存在的分布差异问题的迁移学习方法, 也被应用在宽度学习中. 文献 [43] 创新性地提出了同时利用流形正则化和分布适配惩罚的宽度迁移学习算法, 但只考虑了域适应迁移学习中边缘分布适配的情况. 根据相关工作对域适应迁移学习的描述 (见 2.3 小节), 分布适配可以分为两部分: 一部分是边缘分布适配, 另一部分是条件分布适配, 二者在不同的数据域下比重不同, 仅考虑某一种分布适配惩罚是不全面的, 因此本文在文献 [43] 的基础上采用含动态平衡因子的分布自适应方法来进行迁移, 并应用考虑先验概率的加权条件分布适配矩阵替代一般的条件分布适配矩阵, 来弥补故障数据的不平衡特性. 此外, 为了计算条件分布惩罚, 首先要给出目标域的伪标签, 然而由于故障诊断中故障数据和正常数据的类别不平衡性, 因此在计算伪标签时可能会出现伪标签缺类少类的问题. 针对此问题, 提出了基于类间关系和考虑子集差异集成的两阶段伪标签预测算法, 阶段一算法得到高可靠伪标签, 阶段二考虑了阶段一伪标签预测中的数据不平衡问题, 进而得到可靠度更高且兼具平衡性的伪标签. 在上述两阶段伪标签算法的基础上, 本文提出了基于流形正则化和动态分布适配惩罚的自适应迁移与宽度学习算法来针对复杂工业过程进行故障诊断, 具体算法如下.

3.1 流形变换迁移学习算法

应用域适应迁移方法对不同数据域进行迁移时, 常常由于域间差异过大使得最终迁移效果不佳, 可以应用流形迁移学习的方法提取数据中的可迁移特征 [44]. 本文使用测地线流式核方法 (geodesic flow kernel, GFK) 来消除特征退化的问题, 流形空间下的部分几何结构还能够避免特征在原始空间产生扭曲. 设 $P_s, P_t \in \mathbb{R}^{D \times d}$ 是属于源域和目标域的两个基础空间分布, 设 $R_s \in \mathbb{R}^{D \times (D-d)}$ 为 P_s 的正交补, 即 $R_s^T P_s = 0$, 使用黎曼 (Riemann) 流形的标准欧几里得 (Euclid) 度量, 测地线流可以被参数化为

$$\Phi : t \in [0, 1] \rightarrow \Phi(t) \in G(d, D). \quad (20)$$

特别地, 当 $t = 0, 1$ 时, $\Phi(0) = P_s, \Phi(1) = P_t$; 而当 $t \neq 0, 1$ 时, $\Phi(t) = P_s U_1 \Gamma(t) - R_s U_2 \Sigma(t)$, 其中 $U_1 \in \mathbb{R}^{d \times d}$ 和 $U_2 \in \mathbb{R}^{(D-d) \times d}$ 都是从下式中得到的标准正交矩阵:

$$P_s^T P_t = U_1 \Gamma V^T, \quad R_s^T P_t = -U_2 \Sigma V^T, \quad (21)$$

其中, Γ, Σ 是 d 维方阵, 对角线元素为 $\cos \theta_i, \sin \theta_i$, θ_i 是 P_s, P_t 之间的主角:

$$0 \leq \theta_1 \leq \theta_2 \leq \dots \leq \theta_d \leq \frac{1}{2}\pi. \quad (22)$$

主角可以衡量子空间之间的重叠的程度, $\Gamma(t), \Sigma(t)$ 是 t 距离上的对角矩阵, 对角元素为 $\cos t\theta_i, \sin t\theta_i$. 为了使两域数据能从 $\Phi(0)$ 转移到 $\Phi(1)$, 对该路径上所有 $t \in [0, 1]$ 进行积分作为一种从 $\Phi(0)$ 转移到 $\Phi(1)$ 的增量式办法. 用 x_i, x_j 来表示原来 D 维特征向量, 可将其计算为 $\Phi(t)$, 将变换后的特征设为无限维向量 z_i^∞, z_j^∞ , 测地流核定义为

$$\langle z_i^\infty, z_j^\infty \rangle = \int_0^1 (\Phi(t)^T x_i)^T (\Phi(t)^T x_j) dt = x_i^T G x_j, \quad (23)$$

其中, $G \in \mathbb{R}^{D \times D}$ 是半正定矩阵, 近似于核特征, 可以从之前的定义中表示出来:

$$G = [P_S U_1 \quad R_S U_2] \begin{bmatrix} \Lambda_1 & \Lambda_2 \\ \Lambda_2 & \Lambda_3 \end{bmatrix} \begin{bmatrix} U_1^T P_S^T \\ U_2^T R_S^T \end{bmatrix}. \quad (24)$$

至此, 流形迁移方法的目标完成, 获得了转移到 Grassmann 空间中的特征, 新源域和目标域可以表示为 $X_s^g = \sqrt{G}X_s$, $X_t^g = \sqrt{G}X_t$.

3.2 两阶段伪标签预测算法

在搭建自适应分布迁移学习框架之前, 首先要解决的是条件分布适配中所需的目标域伪标签问题. 为了解决该问题, 文献 [34] 直接使用源域数据去训练分类器并用训练之后的分类器进行预测, 但域漂移现象的存在, 导致其准确率不高. 而文献 [45, 46] 采用 softmax 来获得含有类间关系的软标签, 并通过 KL 散度考虑类间的关系. 但由于故障诊断问题中类别较少, KL 散度在低维下效果不好, 导致其在故障诊断问题中难以应用. 同时样本内部不平衡数据的限制, 导致在考虑输出关系时不可避免地受到样本不平衡的影响. 基于上述问题, 本小节提出两阶段伪标签预测算法来获得可靠度更高且兼具平衡性的伪标签.

阶段一: 将 GFK 方法处理后的数据带入添加了 softmax 的考虑代价敏感函数和加权项的宽度学习系统 (cost sensitive-weighted broad learning system, Cs-WBLS) 来获得分布概率, 即应用源域数据进行训练得到模型, 并用该模型预测目标域样本各类别的概率. 应用预测的概率来计算 Wasserstein 距离获得基准类别相关矩阵, 通过基准相关矩阵和样本的类间相关向量作比较, 得到迭代伪标签并重复上述步骤, 当两次得到的迭代伪标签一致时, 得到阶段一的基于类间关系的伪标签预测结果 \tilde{Y}_1 .

阶段二: 根据阶段一获得的基于类间关系的伪标签 \tilde{Y}_1 , 选取其中排名靠前的 10% 伪标签作为更高可靠性的伪标签, 将其对应的数据与源域数据混合成为新的训练集, 应用 EasyEnsemble 方法将训练集数据分为多个平衡子集, 并对每一个子集应用一个分类器进行训练, 输出根据每一个子集与目标域的 MMD 距离进行加权得到输出结果作为新的伪标签, 重复上述步骤, 循环终止时得到伪标签预测结果.

3.2.1 阶段一: 基于类间关系伪标签预测算法

为了解决伪标签预测问题中存在的样本不平衡问题, 在宽度学习中加入代价敏感函数以及加权项来修改多数类和少数类的权重, 将式 (5) 改写为

$$\min \frac{\theta}{2} \|W\|^2 + \frac{Q}{2} \|AW - Y\|^2, \quad (25)$$

其中, Q 表示误差加权项, 定义为 $Q = TC$, 其中 T 是加权矩阵, C 是代价敏感矩阵且定义为

$$C = \begin{bmatrix} 0 & C_{12} & \cdots & C_{1L} \\ C_{21} & 0 & \cdots & C_{2L} \\ \vdots & \vdots & \ddots & \vdots \\ C_{L1} & C_{L2} & \cdots & 0 \end{bmatrix}, \quad (26)$$

其中, C_{ij} 表示将样本 x_i 误判为样本 x_j 的代价. 通常少数类误判成多数类的代价高于多数类误判为少数类的代价. 而为了防止代价敏感函数使得分类边界迁移, 引入加权矩阵, 由经过流形变换之后的

源域产生, 对多数类样本权重取 0.8, 少数类样本权重取 1. 应用源域数据进行训练, 输入目标域数据来得到每个的输出, 输出结果时, 叠加一个全连接层, 得到基于概率的目标域伪标签. 定义通过 softmax 层之后得到的第 i 个样本输出分量为

$$f_{\text{cls}}^{(i)} = \frac{e^{w_{i\text{softmax}}(a(x)w_{\text{bls}})}}{\sum_{j=1}^C e^{w_{j\text{softmax}}(a(x)w_{\text{bls}})}} = p_i(x), \quad (27)$$

其中, $a(x), w_{\text{bls}}$ 分别对应式 (25) 中 A, W 的分量, 预测结束后, 得到各样本的预测分布集合 $\{p_1, \dots, p_s\}$, 将所有预测输出按类别求均值, 可以得到标签分布的类别均值集合 $\{\bar{p}^{(1)}, \dots, \bar{p}^{(C)}\}$, 其中 $\bar{p}^{(C)}$ 表示类别 C 的平均标签分布. 应用 Wasserstein 距离来度量各类别平均标签分布的关系, 得到平均标签分布的类别相关性矩阵, 称作基准类间相关矩阵 K_{wr} :

$$K_{wr} = \begin{bmatrix} W(\bar{p}^{(1)}, \bar{p}^{(1)}) & W(\bar{p}^{(1)}, \bar{p}^{(2)}) & \dots & W(\bar{p}^{(1)}, \bar{p}^{(C)}) \\ W(\bar{p}^{(2)}, \bar{p}^{(1)}) & W(\bar{p}^{(2)}, \bar{p}^{(2)}) & \dots & W(\bar{p}^{(2)}, \bar{p}^{(C)}) \\ \dots & \dots & \dots & \dots \\ W(\bar{p}^{(C)}, \bar{p}^{(1)}) & W(\bar{p}^{(C)}, \bar{p}^{(2)}) & \dots & W(\bar{p}^{(C)}, \bar{p}^{(C)}) \end{bmatrix}, \quad (28)$$

其中, $W(\bar{p}^{(1)}, \bar{p}^{(1)})$ 表示归一化之后的 Wasserstein 距离, 记 K_{wr} 的第 i 行为 $k_{wr}^{(i)}$, 称为类间相关向量, 而对于第 x 个样本而言, 其类间相关向量可以表示为

$$k_{wr}^{(x)} = \{w(p_x, \bar{p}^{(1)}), \dots, w(p_x, \bar{p}^{(C)})\}. \quad (29)$$

通过比较 $k_{wr}^{(x)}$ 与 $k_{wr}^{(i)}$ 之间的距离, 将最小距离所对应的类别作为该样本的类别预测标签, 同时应用新伪标签重新进行训练, 并重复上述过程, 最终根据类别预测值得到基于类间关系的伪标签 \tilde{Y}_1 . 具体实现算法如算法 1 所示.

算法 1 Inter-class relationship pseudo-label prediction algorithm

Input: Source domain data and target domain data after GFK method X_s^g, X_t^g ;

Output: Target domain pseudo-label based on inter-class relations \tilde{Y}_1 ;

- 1: Use the source domain data X_s^g to train the broad learning system with cost-sensitive and weighted terms, and substitute the target domain data X_t^g to get the output of each sample. The sample output components $p_i(x)$ and the benchmark category correlation matrix K_{wr} are obtained through the softmax layer, (27), and (28);
 - 2: Update the iterative pseudo-label;
 - 3: **while** the iterative pseudo-label \neq the last pseudo-label **do**
 - 4: Calculate the inter-class benchmark correlation vector of each sample according to (29), and compare to obtain the iterative pseudo-label;
 - 5: If the obtained iterative pseudo-label is inconsistent with the last pseudo-label, retrain with the new pseudo-label; otherwise output the pseudo-label;
 - 6: **end while**
-

3.2.2 阶段二: 考虑子集差异的集成伪标签预测算法

为了减少直接预测伪标签带来的误差, 使用阶段一得到的基于类间关系目标域伪标签 \tilde{Y}_1 作为本小节方法第一步使用的伪标签, 同时为了解决生成伪标签时数据不平衡的问题, 采用 EasyEnsemble 算法进行求解. 选取 \tilde{Y}_1 中排名前 10% 的伪标签对应的数据与源域数据混合成为新的训练集 X_{train} , 将

训练集中样本数量最少的类型作为少数类样本, 其余类别作为多数类样本进行重复欠采样, 得到 T_1 个新的训练子集. 应用 BLS 方法作为 EasyEnsemble 中对应每一个子集的基分类器, 同时加入 softmax 层以保证输出的概率. 由于在经过流形变换之后, 两域仍存在分布差异, 在重复欠采样之后, 新子集与源域也会存在差别. 为了尽可能减小该差异, 应用式 (13) 计算 MMD 距离下的两域差异, 两数据域之间的 MMD 距离越大表示其联系越小, 具有较低差异的子集应该获得更大的权重, 这将在最终加权求和时得以体现. 用 $X_{s_i}^g$ 表示第 i 个子集, 第 i 个子集与目标域的相似程度表示为

$$d(X_{s_i}^g, X_t^g) = \frac{1}{\text{MMD}(X_{s_i}^g, X_t^g)}. \quad (30)$$

定义第 i 个子集与目标域的权重为

$$w_i = \frac{d(X_{s_i}^g, X_t^g)}{\sum_j^{T_1} d(X_{s_j}^g, X_t^g)}. \quad (31)$$

用 p_i 来表示第 i 个子集分类的结果, 加权后分类结果可以表示为 $Y = \arg \max \sum_i^{T_1} w_i p_i$, 将该分类结果取前 10% 作为可靠度高的伪标签, 将与之对应的数据和之前的训练集作为新的训练集并重复上述步骤, 最终得到不平衡数据条件下考虑子集差异的集成伪标签预测结果, 具体实现算法如算法 2 所示.

算法 2 A pseudo-label prediction ensemble algorithm considering subset differences

Input: Source domain data and target domain data after GFK method X_s^g, X_t^g ; target domain pseudo-label based on the inter-class relationship generated in the first stage algorithm \tilde{Y}_1 ;

Output: Target domain pseudo-label considering subset differences \tilde{Y} ;

- 1: Use the data of top 10% pseudo-labels in the target domain based on the inter-class relationship \tilde{Y}_1 generated in stage 1, together with X_s^g , as new X_{train} . Apply the EasyEnsemble undersampling method to sample X_{train} , get T_1 balanced subsets and calculate the MMD distances between each subset and the target domain according to (30);
 - 2: Apply BLS with a fully connected layer to train each subset to get T_1 classifiers;
 - 3: Apply the obtained classifiers to predict the target domain, and apply (31) to calculate the weight of each subset to obtain the classification result;
 - 4: Resample the corresponding data and training set X_{train} of the top 10% ranking labels of the classification results, and repeat steps 2 and 3 to get the final result.
-

3.3 基于流形正则化和动态分布适配惩罚的宽度学习算法

本节应用 3.1 小节中 GFK 后的源域目标域数据进行训练, 通过 3.2 小节中给出的伪标签预测算法来估计条件分布适配矩阵, 并通过基于流形正则化和动态分布适配惩罚的宽度学习算法进行故障诊断, 具体算法如下.

3.3.1 具有动态平衡因子的概率分布自适应惩罚算法

由于对于不同数据域而言, 两种分布适配惩罚的重要性不一定相同, 所以一般采用含动态平衡因子的分布自适应方法来进行迁移, 现在考虑将带动态平衡因子的 MMD 距离当作惩罚给入后续的宽度学习系统. 首先计算动态分布适配惩罚, 定义为

$$\bar{D}_f(D_s, D_t) = (1 - \mu)D_f(P_s, P_t) + \mu \sum_{c=1}^C D_f^{(c)}(Q_s, Q_t), \quad (32)$$

其中, $\mu \in [0, 1]$ 是平衡因子, $c = 1, \dots, C$ 是类指标, $D_f(P_s, P_t)$ 表示边缘分布适配, $\sum_{c=1}^C D_f^{(c)}(Q_s, Q_t)$ 表示类的条件分布适配. 当 $\mu \rightarrow 0$ 时表示其相关性差, 距离远, $\mu \rightarrow 1$ 时表示分布差距较小. 应用 MMD 距离对式 (32) 进行计算, 得到边缘分布适配、条件分布适配:

$$D_f(P_s, P_t) = \|E[f(x_s)] - E[f(x_t)]\|_{H_K}^2, \quad (33)$$

$$D_f^{(c)}(Q_s, Q_t) = \|E[f(x_s^{(c)})] - E[f(x_t^{(c)})]\|_{H_K}^2. \quad (34)$$

综合式 (32)~(34), 最终得到

$$\bar{D}_f(D_s, D_t) = (1 - \mu)\|E[f(x_s)] - E[f(x_t)]\|_{H_K}^2 + \mu \sum_{c=1}^C \|E[f(x_s^{(c)})] - E[f(x_t^{(c)})]\|_{H_K}^2. \quad (35)$$

3.3.2 具体方法

为了更好地利用源域和目标域数据, 将动态平衡因子惩罚与流形框架结合起来, 同时考虑应用加权的方法来处理分类中数据不平衡问题. 定义输入数据为经过 GFK 方法的源域数据目标域数据 X_s^g, X_t^g , 应用前文式 (1)~(3) 能够得到

$$A_n^m = [Z^n | H^m] = [a_1(x_1), \dots, a_s(x_s), \dots, a_{s+t}(x_{s+t})]^T, \quad (36)$$

其中, $A_n^m \in \mathbb{R}^{(s+t) \times (n \times N_1 \times m)}$, n 和 m 分别表示特征节点和增强节点数, N_1 表示一个特征节点中包含的节点数量, 将输出 Y 表示为

$$Y = [a_1(x_1), \dots, a_{s+t}(x_{s+t})]^T [w_1, \dots, w_{s+t}], \quad (37)$$

其中, $y = a_i(x_i)w_i$ 为定义的针对某一样本的输出, 下面对所提算法最小化目标进行定义:

$$\min \frac{\theta}{2} \|W\|^2 + \frac{Q}{2} \|AW - Y\|^2 + \frac{\lambda}{2} \bar{D}_f(X_s^g, X_t^g) + \frac{\rho}{2} R_f(X_s^g, X_t^g), \quad (38)$$

其中, θ, λ, ρ 是正则化项参数, Q 为误差加权项, $\bar{D}_f(X_s^g, X_t^g)$ 是前文 3.3.1 小节中所描述具有动态平衡因子的分布自适应惩罚项, 即式 (32), $R_f(X_s^g, X_t^g)$ 为流形正则化项. 误差加权项 Q 是 $\mathbb{R}^{(s+t) \times (s+t)}$ 的矩阵, 定义为

$$Q = \begin{bmatrix} Q_{1s \times s} & 0_{s \times t} \\ 0_{t \times s} & 0_{t \times t} \end{bmatrix}, \quad (39)$$

其中, $Q_{1s \times s} = Q_1 T_c$, T_c 是加权矩阵, 由流形变换之后的源域产生, 可以降低多数类样本的权重, Q_1 是基底正则化项, 可以给算法目标一定的惩罚.

此外, 式 (38) 中, $\bar{D}_f(X_s^g, X_t^g)$ 是含动态因子概率分布自适应惩罚项. 通过式 (35) 可以将二者间距离以 MMD 形式表示, 带入宽度学习系统, 可以得到

$$\begin{aligned} \bar{D}_f = (1 - \mu) & \left\| \frac{1}{n_s} \sum_{i=1}^s [a_i(x_i)w_i] - \frac{1}{n_t} \sum_{j=1}^t [a_{s+j}(x_{s+j})w_{s+j}] \right\|_{H_K}^2 \\ & + \mu \sum_{c=1}^C \left\| \frac{1}{n_s^{(c)}} \sum_{i=1}^{s^{(c)}} [a_i(x_i^{(c)})w_i] - \frac{1}{n_t^{(c)}} \sum_{j=1}^{t^{(c)}} [a_{t+j}(x_{t+j}^{(c)})w_{t+j}] \right\|_{H_K}^2. \end{aligned} \quad (40)$$

式 (40) 可以通过定义 MMD 矩阵的方法转化为

$$\bar{D}_f(X_s^g, X_t^g) = (1 - \mu)\text{tr}(W^T A^T(X) M_0 A(X) W) + \mu \sum_{c=1}^C \text{tr}(W^T A^T(X) M_c A(X) W), \quad (41)$$

其中, 定义 M_0 为边缘分布适配的 MMD 矩阵:

$$M_0(i, j) = \begin{cases} \frac{1}{n^2}, & x_i, x_j \in X_s^g, \\ \frac{1}{m^2}, & x_i, x_j \in X_t^g, \\ -\frac{1}{mn}, & \text{otherwise.} \end{cases} \quad (42)$$

定义 M_c 为条件分布适配的 MMD 矩阵, 为了解决不平衡迁移的问题, 选择考虑先验概率的加权的条件分布适配矩阵, 如下所示:

$$M_c(i, j) = \begin{cases} \frac{P(y_s^c)}{n_c^2}, & x_i, x_j \in (X_s^g)^{(c)}, \\ \frac{P(y_t^c)}{m_c^2}, & x_i, x_j \in (X_t^g)^{(c)}, \\ -\frac{\sqrt{P(y_s^c)P(y_t^c)}}{m_c n_c}, & \begin{cases} x_i \in (X_s^g)^{(c)}, x_j \in (X_t^g)^{(c)}, \\ x_i \in (X_t^g)^{(c)}, x_j \in (X_s^g)^{(c)}, \end{cases} \\ 0, & \text{otherwise,} \end{cases} \quad (43)$$

其中, $P(y_s^c)$ 和 $P(y_t^c)$ 表示源域和目标域中 c 类的先验概率, 对式 (41) 进行简化, 令 $M = (1 - \mu)M_0 + \mu \sum_{c=1}^C M_c$, 能够得到

$$\bar{D}_f(X_s^g, X_t^g) = \text{tr}(W^T A^T M A W). \quad (44)$$

流形正则化项 $R_f(X_s^g, X_t^g)$ 中, 拉普拉斯 (Laplace) 矩阵一般可表示为

$$L = \frac{1}{2} \sum_i^{s+t} \sum_j^{s+t} w_{ij} \|P(y|x_i) - P(y|x_j)\|^2. \quad (45)$$

由于条件概率很难计算, 所以使用样本的输出来进行近似:

$$L = \frac{1}{2} \sum_i^{s+t} \sum_j^{s+t} w_{ij} \|\hat{y}_i - \hat{y}_j\|^2, \quad (46)$$

其中, w_{ij} 表示样本 x_i 与 x_j 之间的成对相似性, 可以通过高斯 (Gauss) 核函数来进行求解:

$$w_{ij} = \begin{cases} \exp(-\frac{\|x_i - x_j\|^2}{2\sigma^2}), & x_i \in N_k(x_j) \text{ or } x_j \in N_k(x_i), \\ 0, & \text{otherwise,} \end{cases} \quad (47)$$

其中, $N_k(x)$ 表示 x 的 k 个近邻集合, 所有样本产生的 w_{ij} 可以构成邻接矩阵 W , 将邻接矩阵的每列求和可以得到对角元素, 即

$$D_{ii} = \sum_{j=1}^{s+t} w_{ij}, \quad (48)$$

其中, 所有对角元素 D_{ii} 构成了度矩阵 D , 拉普拉斯矩阵就可定义为 $L = D - W$, 并可通过 $L = D^{-\frac{1}{2}}(D - W)D^{\frac{1}{2}}$ 进行标准化.

综合式 (38) 以及式 (45)~(48) 与拉普拉斯矩阵, 可以得到宽度学习框架下的流形正则化项为

$$R_f(X_s^g, X_t^g) = \text{tr}(W^T A^T L A W), \quad (49)$$

其中, $L \in \mathbb{R}^{(s+t) \times (s+t)}$, 定义为

$$L = \begin{bmatrix} 0_{s \times s} & 0_{s \times t} \\ 0_{t \times s} & L_{t \times t} \end{bmatrix}. \quad (50)$$

根据式 (36), (37), (44) 和 (49), 将式 (38) 转化为

$$\min \frac{\theta}{2} \|W\|^2 + \frac{Q}{2} \|AW - Y\|^2 + \frac{\lambda}{2} \text{tr}(W^T A^T M A W) + \frac{\rho}{2} \text{tr}(W^T A^T L A W). \quad (51)$$

将上式对 W 求导可得

$$\theta W + A^T Q (AW - Y) + \lambda A^T M A W + \rho A^T L A W = 0. \quad (52)$$

最终得到宽度学习系统的输出权重为

$$W = (\theta I + A^T (Q + \lambda M + \rho L) A^{-1}) A^T Q Y. \quad (53)$$

至此, 算法主体部分完成, 得到权重矩阵 W 后对目标域数据进行预测, 并将宽度迁移学习系统得到的结果作为目标域新伪标签重新计算 MMD 矩阵, 重复上述实验步骤直至满足终止条件, 算法整体实现步骤如算法 3 所示.

算法 3 Transfer broad learning algorithm based on imbalanced data

Input: Source domain subset and target domain data X_s, X_t ;

Output: Fault diagnosis classification results;

- 1: Apply the manifold transformation transfer learning algorithm in Subsection 3.1 to narrow the distance between X_s, X_t , and obtain the data X_s^g, X_t^g under the Grassmann manifold;
 - 2: Apply Algorithms 1 and 2 in Subsection 3.2 to calculate high-reliability pseudo-labels, and use (41)–(44) to calculate the MMD penalty term;
 - 3: Apply (45)–(49) to calculate the manifold regularization term to obtain the implicit information under the manifold of the target domain;
 - 4: Use (51)–(53) to obtain the output weight matrix, and use this matrix to obtain the target domain prediction result;
 - 5: Recalculate the MMD penalty term using the prediction result as a pseudo-label and (41)–(44), and repeat steps 2–4 to obtain the final prediction result.
-

4 污水处理过程数据实验

4.1 实验数据来源

污水处理是我国水资源可持续发展的关键, 其根本目的就是将城市生活、工业生产等产生的污水经过污水处理的各道工序后达到国家规定的出水指标. 目前, 应用最广泛的污水处理工艺是活性污泥法^[47], 如图 4 所示. 具体过程通常分三级: 一级处理通过洗砂间、初沉池等步骤初步处理掉塑料、石子等常见固体废物; 二级处理主要是通过厌氧、好氧等生化反应, 完成有机污染物的分解; 三级处理通过加入氯气、活性炭等进一步分解二级处理未能完全去除的磷、有机物污染等, 最后将其排入到大

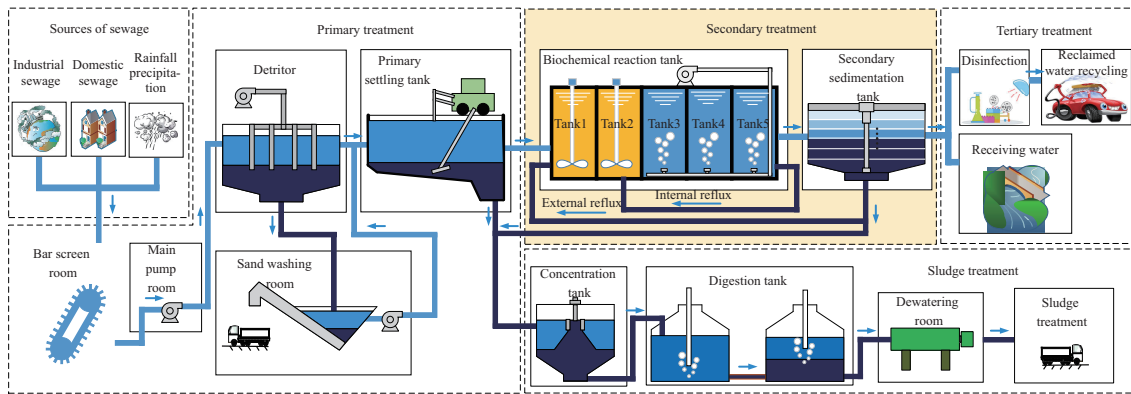


图 4 (网络版彩图) 污水处理工艺流程示意图

Figure 4 (Color online) Schematic diagram of sewage treatment process

自然中, 污水处理出水质量指标主要包括生化需氧量 (biochemical oxygen demand, BOD)、化学需氧量 (chemistry oxygen demand, COD)、悬浮物 (suspended solid, SS)、总磷 (total phosphorous, TP)、总氮 (total nitrogen, TN) 等. 污水处理过程是一个多变量、强耦合、大时滞、工况非平稳波动的复杂生化反应过程^[48]. 污水处理的效果与被动接受的进水水量与水质息息相关, 一旦进水量和水质产生变化, 就会造成污水处理中工况变化, 整个污水处理系统产生动态波动, 在不同工况下如果污水处理系统发生故障, 就会最终导致能耗高、出水水质下降等一系列问题. 为此, 采用所提方法对污水处理过程异常工况进行诊断, 应用由国际水质协会和欧盟科学技术与合作组织共同开发的前置反硝化工艺污水处理基准平台 (benchmark simulation model 1, BSM1) 进行实验. BSM1 是评价污水处理过程控制方法的主流平台. BSM1 设置了 3 种不同的运行工况, 分别对应晴天、连绵细雨和暴雨情况的进水信息.

首先, 对本文实验所用污水处理过程数据进行描述. 所选数据包含表 1 中 28 个过程变量数据与对应 BOD, COD, TN, TP 4 个出水水质指标变量的数据, 本实验将污水处理过程数据对应工况的正常与否情况分为 3 类: (1) 正常 (Normal)、(2) 毒性冲击 (Fault1)、(3) 泡沫 (Fault2). 实验中, 将晴天数据作为源域数据, 将连绵细雨天气和暴雨天气作为目标域数据, 源域数据共 2686 组, 其中正常样本 2166 组, 毒性冲击数据 270 组, 泡沫数据 250 组; 目标域数据共 2002 组, 其中正常样本 1529 组, 毒性冲击数据 252 组, 泡沫数据 221 组.

4.2 实验与分析

4.2.1 故障诊断结果和对比实验

将污水处理过程数据应用到本文所提算法中, 为从多个角度对 TL-BLS 算法性能进行评估, 使用 K 近邻 (K-nearest neighbor, KNN)、核支持向量机 (kernel support vector machine, KSVM)、加权极限学习机 (weighted extreme learning machine, WELM) 等作为不同迁移学习方法下的分类方法, 同时为了对比深度学习和宽度学习的区别, 还使用了一维卷积神经网络 (convolutional neural networks, CNN) 作为分类器. 为了从整体上比较所提方法的优越性, 使用成体系的迁移学习正则化框架方法 ARTL^[37] 进行迁移学习整体方法的比较, 根据文献 [38] 使用结构风险最小化的分类方式作为 ARTL 中的分类器. 使用 3 种域适应方法进行不同迁移方法下的对比, 应用迁移成分分析法 (transfer component analysis, TCA)、联合分布自适应方法 (joint distribution adaptation, JDA)、带动态平衡因子的自适应方法 (balance distribution adaptation, BDA). 由于迁移对象是非平衡数据域, 根据式 (43) 计算加权后

表 1 影响污水处理过程出水水质的主要过程变量

Table 1 Main process variables affecting effluent quality in sewage treatment process

No.	Symbol	Physical meaning of variables	No.	Symbol	Physical meaning of variables
1	Q_{in}	Inlet water flow	15	$S_{S,3}$	Readily biodegradable organic substrate amount of tank 3
2	$S_{NH,in}$	Influent ammonia concentration	16	$S_{ALK,3}$	Alkalinity of tank 3
3	$X_{BH,1}$	Active heterotrophic bacteria biomass of tank 1	17	$X_{BH,4}$	Active heterotrophic bacteria biomass of tank 4
4	$S_{NO,1}$	Nitrate nitrogen concentration of tank 1	18	$X_{BA,4}$	Active autotrophic bacteria biomass of tank 4
5	$S_{S,1}$	Readily biodegradable organic substrate amount of tank 1	19	$S_{O,4}$	Dissolved oxygen concentration of tank 4
6	$S_{ALK,1}$	Alkalinity of tank 1	20	$S_{NH,4}$	Ammonia concentration of tank 4
7	$X_{BH,2}$	Active heterotrophic bacteria biomass of tank 2	21	$S_{S,4}$	Readily biodegradable organic substrate amount of tank 4
8	$S_{NO,2}$	Nitrate nitrogen concentration of tank 2	22	$S_{ALK,4}$	Alkalinity of tank 4
9	$S_{S,2}$	Readily biodegradable organic substrate amount of tank 2	23	$X_{BH,5}$	Active heterotrophic bacteria biomass of tank 5
10	$S_{ALK,2}$	Alkalinity of tank 2	24	$X_{BA,5}$	Active autotrophic bacteria biomass of tank 5
11	$X_{BH,3}$	Active heterotrophic bacteria biomass of tank 3	25	$S_{O,5}$	Dissolved oxygen concentration of tank 5
12	$X_{BA,3}$	Active autotrophic bacteria biomass of tank 3	26	$S_{NH,5}$	Ammonia concentration of tank 5
13	$S_{O,3}$	Dissolved oxygen concentration of tank 3	27	$S_{S,5}$	Readily biodegradable organic substrate amount of tank 5
14	$S_{NH,3}$	Ammonia concentration of tank 3	28	$S_{ALK,5}$	Alkalinity of tank 5

的条件分布适配矩阵, 代替一般的条件分布适配矩阵带入 JDA 和 BDA 方法中。

对于所提 TL-BLS 方法, 两阶段伪标签预测算法中选择子集数 $T_1 = 21$, TL-BLS 中惩罚项和正则化参数 θ, λ, ρ 分别选择为 $10^{-4}, 10^{-3}, 10^{-3}$, 对角矩阵 Q_1 选择 10^{-1} 作为加权项的基底, 对多数类权重取 0.8, 少数类权重取 1, 用 $N_1^{TL}, N_2^{TL}, N_3^{TL}$ 表示 TL-BLS 算法中每个特征窗口下的特征节点数、特征窗口数和增强节点数, 由于 BLS 分类器效果主要受特征节点总数 $N_1^{TL} \times N_2^{TL}$ 和增强节点数 N_3^{TL} 的影响, 为方便选取参数, 固定选择特征节点窗口 $N_2^{TL} = 15$ 时, 不同节点数下 TL-BLS 模型在污水处理过程中的故障诊断结果如表 2 所示。从表中可以看出, 固定特征节点窗口数时, 随着增强节点数 N_3^{TL} 的增加, 分类准确率会出现先增高后降低的现象, 增加特征节点数会导致训练时间的增长, 同时适度增加也会增加训练精度, 由于污水处理过程故障诊断中对时间的需求没有对准确率的需求高, 所以在实验中选择最高准确率较为合适, 当特征节点窗口数固定为 $N_2^{TL} = 15$ 时, 特征节点 $N_1^{TL} = 50$, 增强节点 $N_3^{TL} = 2000$ 时精度较高, 所以将 TL-BLS 的结构定为 50-15-2000。

应用上述 6 种方法和 3 种域适应迁移学习方法进行对比。KNN 方法中, 近邻数选择 Neighbor = 10, 使用两阶段伪标签预测法获得的伪标签以及域适应方法后得到的新源域数据进行训练; KSVM 方法中, 核函数选择高斯核, 使用两阶段伪标签预测法获得的伪标签以及域适应方法后得到的新源域数据进行训练; WELM 方法中, 节点个数为 2500, 加权项选择多类权重为 0.8, 少类权重为 1, 使用两阶

表 2 不同节点数下 TL-BLS 的实验结果

Table 2 Experimental results of TL-BLS with different node numbers

Feature nodes N_1^{TL}	Enhancement nodes N_3^{TL}	TL-BLS	
		ACC (%)	Time (s)
50	500	81.87	21.54
50	1000	92.56	25.63
50	1500	93.71	30.08
50	1600	93.51	32.32
50	1700	93.00	35.28
50	1800	93.21	36.09
50	1900	93.61	37.34
50	2000	93.96	39.56
50	2100	93.46	41.29
50	2200	93.36	43.24
50	2300	92.91	46.62
50	2400	92.65	47.55
50	2500	93.86	49.23
100	1500	80.52	32.45
100	2000	90.56	42.78
100	2500	90.61	56.11
150	1500	80.87	36.86
150	2000	81.21	49.42
150	2500	90.11	61.50

段伪标签预测法获得的伪标签以及域适应方法后得到的新源域数据进行训练; ARTL 方法中, 使用文献 [38] 中的参数和获得伪标签的方法以及域适应方法后得到的平衡因子和 MMD 惩罚项进行训练; 深度方法选择 3 层一维 CNN 方法, 卷积核选择 3, 通道数选择 100, 使用两阶段伪标签预测法获得的伪标签以及域适应方法后得到的新源域数据进行训练; TL-BLS 参数设置与上文保持一致. 为了证明本文方法在处理不平衡问题上的优越性, 对比指标选择准确率和宏观的 F1 值, 即 Macro-F1, 得到对比结果如表 3 所示, 同时为了具体体现 TL-BLS 算法在不平衡故障诊断方面的优越性, 绘制了图 5 所示的混淆矩阵来表示各单一故障类型的诊断准确性, 图 5 中 Normal 表示正常数据, Fault1 表示毒性冲击故障, 而 Fault2 表示泡沫故障. 从图 5 可以看出, 故障 1 的准确率有 90.1%, 故障 2 的诊断准确率为 88.2%, 所提诊断方法对每个故障的分类精度均达到 85% 以上, 说明该方法对非平衡数据的适配能力较强, 表明本文域适应迁移学习方法中使用的考虑先验概率的条件分布适配矩阵和加权的迁移宽度学习算法在解决非平衡数据的问题上有着良好的性能. 从表 3 中可以看出, 除 ARTL 外, 在 3 种域适应的迁移学习方法中, 应用 BDA 方法和 JDA 方法准确率均高于 TCA 方法, 且 BDA 方法在准确率上略优于 JDA 方法, 是因为 BDA 方法应用了其他两种方法没有的动态平衡因子, 能更准确地表示源域和目标域的差异. 同为正则化框架的 ARTL 方法由于并未使用本文所提的伪标签方法进行条件分布概率计算, 所以 JDA 和 BDA 方法在最终准确率上不如仅考虑边缘分布适应的 TCA 方法, 也从侧面反映出使用高精度伪标签的重要性. 将 CNN 方法和本文 TL-BLS 方法进行比较, CNN 方法能在一定程度上解决不平衡的故障诊断问题, 但网络结构、网络参数等问题, 使其与本文所提 TL-BLS 方法

表 3 不同方法故障诊断对比结果

Table 3 Comparative results of fault diagnosis by different methods

	TCA		JDA		BDA	
	ACC (%)	Macro-F1	ACC (%)	Macro-F1	ACC (%)	Macro-F1
KSVM	83.27	0.6011	86.16	0.7128	86.46	0.7130
WELM	85.96	0.6969	86.66	0.7171	87.41	0.7309
KNN	83.57	0.6583	85.31	0.6923	86.36	0.7135
CNN	77.72	0.5809	83.52	0.6522	84.52	0.7339
ARTL	87.71	0.7411	85.61	0.6931	85.71	0.6939
TL-BLS	91.81	0.8576	93.86	0.9038	93.96	0.9042

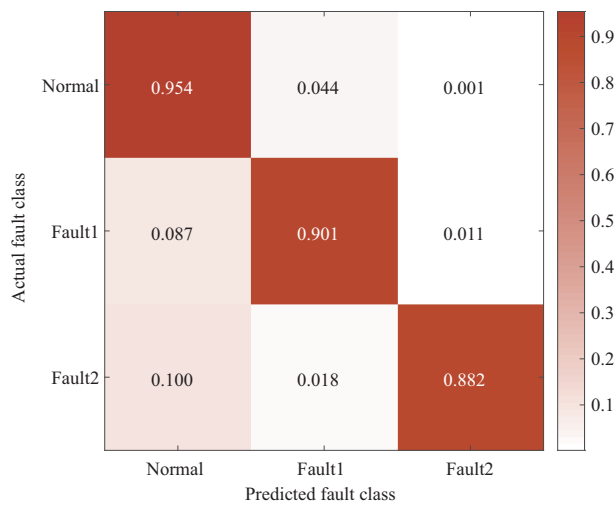


图 5 (网络版彩图) 故障类型诊断混淆矩阵

Figure 5 (Color online) Confusion matrix of fault diagnosis

有一定差距. 结合图 5 和表 3 能够看出, 本文所提的 TL-BLS 算法在不平衡的故障诊断问题中, 准确率和 Macro-F1 的表现均好于其他方法, 准确率提高了 6.25%, Macro-F1 值提高了 0.1631. 所以本文提出的 TL-BLS 算法在非平衡数据的条件下更加适用, 表现更加良好, 总的来说, 本文提出的 TL-BLS 故障诊断方法能够实现更加准确、更加均衡的诊断.

4.2.2 两阶段伪标签预测算法实验与分析

(1) 一阶段伪标签预测算法实验分析. 为了满足计算条件分布适配需要目标域标签的条件, 同时在不平衡数据的条件下, 防止预测伪标签出现缺类少类的情况, 应用本文 3.2 小节中算法来实现对伪标签进行预测. 在算法 1 基于类间关系的伪标签方法中, 需要先应用 Cs-WBLS 给出基于概率的伪标签, 首先对该 BLS 进行节点数选择. 用 N_1^1, N_2^1, N_3^1 表示每个特征窗口下的特征节点数、特征窗口数和增强节点数, 固定选择特征节点窗口数 $N_2^1 = 15$, 使用 $N_1^1 = 50, N_3^1 = 600$ 作为初始值, N_3^1 每次增加 25 进行参数的选择, 不同节点数下的效果如图 6 所示, 根据训练集准确度曲线显示, 选择 $N_1^1 = 50, N_3^1 = 800$ 能够同时兼顾训练时间和训练精度, 在选择结束后, 用目标域数据进行预测, 得到伪标签准确度曲线, 从图中能够得到在 $N_1^1 = 50, N_3^1 = 800$ 时, 伪标签的训练精度也较高, 所以选择此时 Cs-WBLS 的参数

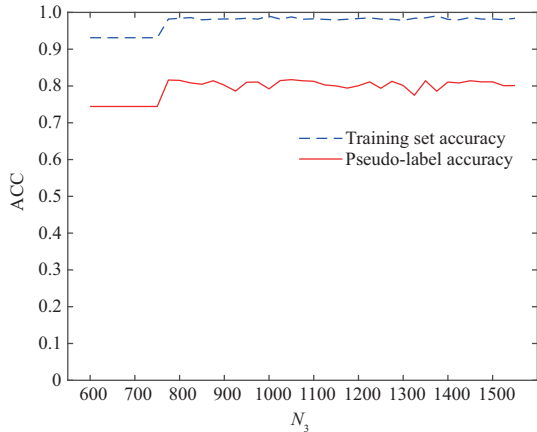


图 6 (网络版彩图) 预测准确率随 BLS 增强节点数变化图

Figure 6 (Color online) Prediction accuracy varying with BLS enhancement node numbers

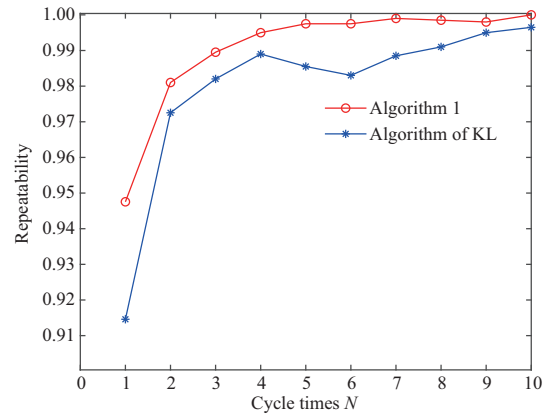


图 7 (网络版彩图) KL 散度方法和算法 1 收敛速度对比

Figure 7 (Color online) Convergence speed comparison of KL algorithm and Algorithm 1

表 4 伪标签初始值预测准确率

Table 4 Pseudo-label initial value prediction accuracy

	Source domain direct prediction	KL divergence prediction	Proposed algorithm 1
ACC (%)	81.52	82.27	83.56
Time (s)	—	41.7	21.6

为 $N_1^1 = 50, N_2^1 = 15, N_3^1 = 800$.

为了验证一阶段伪标签方法的准确性和有效性, 使用上述参数进行实验. 将 KL 散度预测方法和所提算法 1 的收敛速度进行对比, 得到图 7, 同时为表示一阶段伪标签预测算法 1 的优越性, 应用源域直接预测法、KL 散度预测法, 以及伪标签预测第一阶段的算法 1 进行对比, 结果如表 4 所示. 由图 7 可以看出所提方法得到的伪标签在准确度上与直接使用源域训练好的分类器相比要更高, 再结合表 4 可以看出, 在收敛速度上, 一阶段伪标签预测算法 1 可快速收敛到重复率 99% 以上, 同时准确率也高于 KL 预测算法, 所以相比源域直接预测法和 KL 散度预测法, 一阶段伪标签预测算法 1 能够提供可靠性更强的伪标签, 供算法 2 进行筛选和使用.

(2) 二阶段伪标签预测算法实验分析. 在应用伪标签预测第一阶段得到的可靠性高的基于类间关系的伪标签和算法 2 时, 需要对算法 2 中两个部分进行参数选择. 首先, 由于考虑不平衡数据的原因, 重复欠采样生成的新子集样本数量和原来 BLS 的训练样本数量不同, 为了同时兼顾准确率和平衡性, 固定特征节点窗口数 $N_2^2 = 15$ 不变时, 选择 $N_1^2 = 50, N_3^2 = 1500$ 作为此时使用的 BLS 节点数. 其次, 在重复欠采样时, 子集个数的不同对伪标签预测的结果有一定的影响, 选取不同子集个数对伪标签结果进行比较, 用 T_1 表示子集数量, 取 $T_1 \in \{6, 9, 12, 15, 18, 21, 24, 27, 30, 33\}$, 可得不同子集数对应的伪标签预测平均准确率如表 5 所示. 从表中可知得到的伪标签算法获得的子集最高平均准确率是 89.25%. 但由于采样的不确定性, 重复多次实验并绘制图 8. 可以看出, 当子集个数 $T_1 > 18$ 时, 生成伪标签的平均准确率提升较小, 波动率趋于不变, $T_1 = 21, T_1 = 24$ 和 $T_1 = 30$ 时平均准确率均为 89.25%, 虽然 $T_1 = 21$ 并不是最高准确度伪标签对应的子集个数, 但 $T_1 = 21$ 时有较高的平均子集

表 5 不同子集个数的伪标签预测平均准确率

Table 5 Average accuracy of pseudo-label prediction with different numbers of training subsets

	$T_1 = 6$	$T_1 = 9$	$T_1 = 12$	$T_1 = 15$	$T_1 = 18$	$T_1 = 21$	$T_1 = 24$	$T_1 = 27$	$T_1 = 30$	$T_1 = 33$
ACC (%)	89.08	89.09	89.18	89.22	89.14	89.25	89.25	89.21	89.25	89.19
Time (s)	42.72	46.81	63.61	77.55	92.52	108.99	124.12	140.61	151.83	167.97

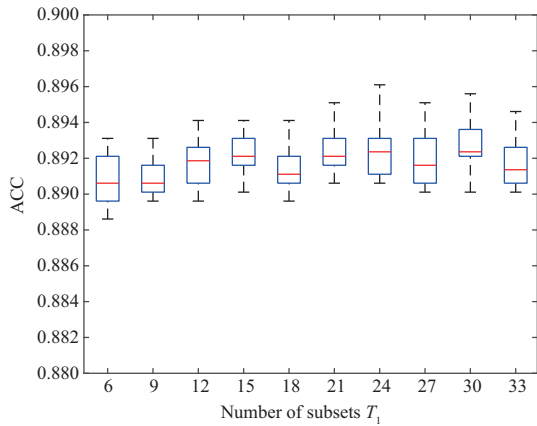


图 8 (网络版彩图) 不同子集数对应的预测准确率

Figure 8 (Color online) Prediction accuracy with different numbers of training subsets

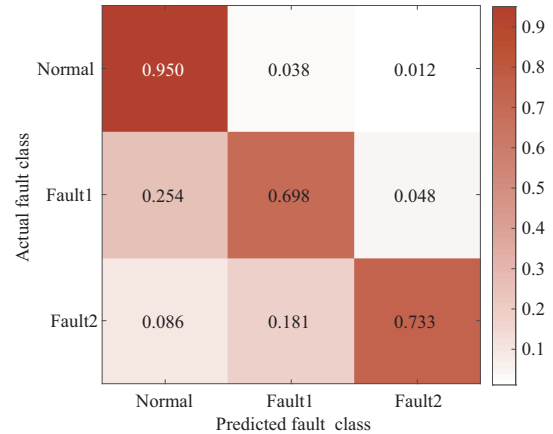


图 9 (网络版彩图) 伪标签预测混淆矩阵

Figure 9 (Color online) Confusion matrix of pseudo-label prediction

分类准确率和较低的波动, 但由于 $T_1 = 24$ 和 $T_1 = 30$ 时其运算时间过长, 训练代价过高, 最终选择波动更小, 训练更快的 $T_1 = 21$ 作为集成学习的子集个数.

为了证明算法 2 的处理非平衡问题的优越性, 使用上述参数下的算法 2 模型得到的伪标签结果绘制混淆矩阵, 得到图 9. 可以看出两阶段伪标签预测算法能够有效解决伪标签预测中的数据不平衡问题, 并避免由于不平衡问题的存在出现的缺类少类现象. 应用本文提出的两阶段伪标签预测算法能够保证输入迁移宽度学习的伪标签少类精度达到 69% 以上, 虽然目前正常类与 Fault1 类有部分混淆现象, 但也能够保证其大部分的数据分类正确, 错误部分也能够后续通过域适应的加权条件分布适配矩阵和迁移宽度学习中的加权项进行调整, 为最终结果的准确性提供了保证. 总之, 本文提出的两阶段伪标签预测算法能够对非平衡数据进行处理, 能够产生较高准确率和更加平衡的伪标签.

4.2.3 消融实验

(1) **GFK 算法实验结果.** 通过前文 3.1 小节所述 GFK 方法对源域与目标域数据进行拉近, 以减少两域数据距离, 为了能直观地说明流形迁移学习的作用, 选用 t-SNE 方法对经过 GFK 方法处理前后的源域和目标域数据进行可视化处理, 用 S 表示源域数据, 用 T 表示目标域数据, 可视化结果如图 10(a) 和 (b) 所示. 从图 10(a) 中可以看出未经流形迁移学习方法处理的数据分布散乱, 两域数据具有差异, 从图 10(b) 中可以看出二者虽然仍有分布上的差异, 但经过 GFK 方法处理后的不同域的数据分布更加集中, 有利于域的迁移.

为了进一步体现 GFK 方法对域间距离的拉近, 从数据上对应用 GFK 方法前后的效果进行对比, 使用 KNN 方法、KSVM 方法、WELM 方法和本文 TL-BLS 方法进行比较分析. 进行消融实验时, 对

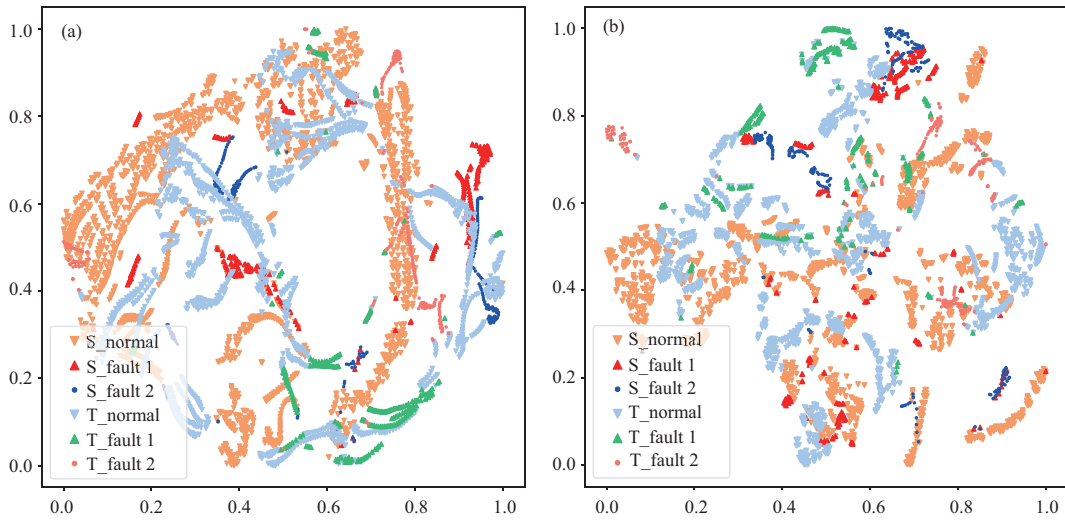


图 10 (网络版彩图) (a) 未经过 GFK 处理和 (b) 经过 GFK 处理的数据 t-SNE 可视化图

Figure 10 (Color online) t-SNE visualization of data (a) without and (b) with GFK processing

表 6 使用 GFK 方法后故障诊断结果

Table 6 Fault diagnosis results using the GFK method

Data class	KNN		KSVM		WELM		TL-BLS	
	P (%)	R (%)	P (%)	R (%)	P (%)	R (%)	P (%)	R (%)
Normal	85.4	98	83.9	99.3	80.5	95.2	97.1	95.4
Toxic shock	42.6	19.4	76.8	21	22.2	12.7	75.9	90.1
Foam	79.0	47.5	83.6	46.2	84.0	19.0	97.5	88.2

于 KNN 方法、KSVM 方法和 WELM 方法, 应用上述参数, 采用 GFK 方法处理前后的源域数据进行训练. 对于所提 TL-BLS 方法, 采用 GFK 处理前后源域数据和目标域进行实验, 并选用带加权后的条件分布适配矩阵的 BDA 方法进行迁移, 表 6 为应用 GFK 方法之后的详细诊断结果, 应用 GFK 方法前后对比结果如图 11 所示.

根据表 6, 能够看出所提 TL-BLS 方法在使用 GFK 方法后对故障诊断模型中的少数类故障有较高分类准确率, 同时结合图 11, 能够看出通过 GFK 方法对污水处理过程数据进行迁移效果明显, 可提升分类准确率达到 5% 以上, 表明了 GFK 方法具有拉近域距离的良好特性和有效性.

(2) 两阶段伪标签预测算法实验结果. 为了验证伪标签方法对不平衡问题处理的有效性以及预测的准确性, 对本文提出的两阶段伪标签预测方法进行消融实验, 选择本文提出的 TL-BLS 方法以及 WELM 方法进行对比. 实验中, (1) 使用 KNN 方法直接预测伪标签; (2) 仅用一阶段伪标签预测算法 1 进行伪标签预测; (3) 应用两阶段伪标签预测算法 1 和 2 进行伪标签预测, 上述方法均使用动态的平衡因子方法进行域适应的迁移. WELM 使用不同伪标签算法得到的动态平衡因子域适应迁移学习方法后得到的数据进行分类, TL-BLS 方法使用不同伪标签方法得到的动态平衡因子和 MMD 矩阵惩罚项进行分类, 对比结果如图 12 所示. 从图 12 可以看出本文提出的算法 1 在伪标签预测结果上相比仅对数据使用 GFK 方法后进行分类准确度更高, 而算法 1 和 2 结合使用后的效果提升更大. 在考虑了类间关系和不平衡数据的情况下, 提出的伪标签算法获得了更高的准确率, 与此同时还兼顾了算

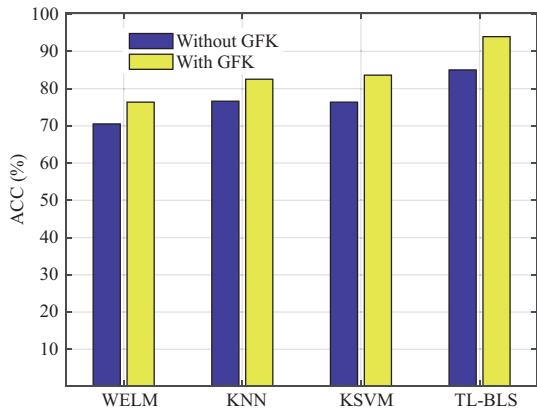


图 11 (网络版彩图) 应用 GFK 方法前后对比结果

Figure 11 (Color online) Accuracy comparison of the methods with and without GFK

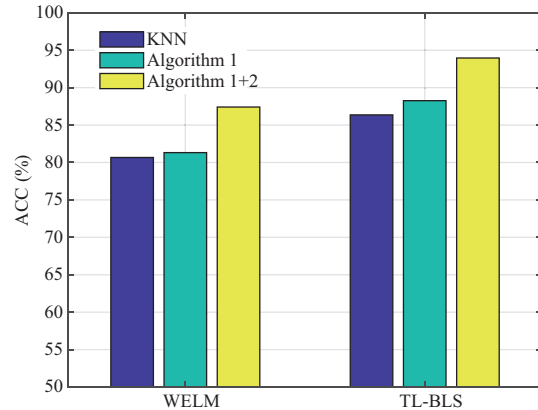


图 12 (网络版彩图) 伪标签消融实验结果

Figure 12 (Color online) Ablation experiment results

法 3 中所要求的伪标签不能缺类少类的基本条件. 总的来说, 提出的伪标签算法在提高准确率和防止伪标签缺类少类的问题上有显著的效果.

4.2.4 参数敏感性分析

本文迁移宽度学习算法主要包含 4 个可调参数, 分别是算法正则化系数 θ , MMD 矩阵正则化系数 λ , 流形正则化系数 ρ , 以及对角矩阵加权项的基底 Q_1 , 为了寻找上述参数的可行范围, 对仅改变上述参数进行参数敏感性实验. 由于加权项在给定权重上已有策略, 固定加权基底 Q_1 , 限制其等于 10^{-1} . 对于其他参数, 在离散集 $\theta \in \{10^{-2}, 10^{-3}, 10^{-4}, 10^{-5}\}$; $\lambda \in \{10^0, 0.5, 0.2, 10^{-1}, 10^{-2}, 10^{-3}, 10^{-4}\}$; $\rho \in \{10^{-1}, 0.05, 0.02, 10^{-2}, 10^{-3}, 10^{-4}, 10^{-5}\}$ 中进行取值, 并在前置算法设置的参数上进行实验, 实验结果如图 13(a)~(d) 所示, 可以看出加权基底 Q_1 限定在 10^{-1} 的条件下时, 算法正则化系数 θ 取值范围更适合在 $\theta \in \{10^{-3}, 10^{-4}\}$ 中取值, MMD 矩阵正则化系数 λ 的取值范围在 $[0.2, 10^{-4}]$ 之间; 流形正则化系数 ρ 在 $[10^{-2}, 10^{-4}]$ 内取值有较好的效果, 由于更小的取值对算法的影响力较小, 所以参数敏感性实验中并未对更小的取值进行讨论. 相比于流形正则化系数 ρ , MMD 矩阵正则化系数 λ 可供选择的区间更多, 而流形正则化系数 ρ 对准确率的影响更大, 总之, 以上参数在上述给定的取值范围内, 都能够取得超过一般方法的准确率.

5 结论

针对工业过程多工况特性以及数据不平衡的故障诊断问题, 提出基于自适应迁移与宽度学习的多工况故障诊断方法, 同时使用数据中的流形信息以避免数据降维损失. 所提方法首先在高维空间中寻找测地线流式核模型来缩小目标域和源域之间的距离. 然后采用基于类间关系和考虑子集差异集成的两阶段伪标签预测算法来给出伪标签, 解决了伪标签出现缺类少类的现象, 为应用灵活分布自适应方法提供基础. 之后, 将动态的自适应 MMD 惩罚项和流形正则化项融入 BLS 模型中, 最终建立了迁移宽度学习算法的诊断方法. 最后, 采用污水处理过程数据进行实验, 本文提出的 TL-BLS 算法在污水处理过程故障诊断实验中能够达到 93.96% 的准确率, 同时还能够保证少数类诊断准确率超过 85%. 通过与 WELM, KNN, KSVM, CNN 以及 ARTL 算法的多方面比较表明: 本文所提 TL-BLS 方法具有

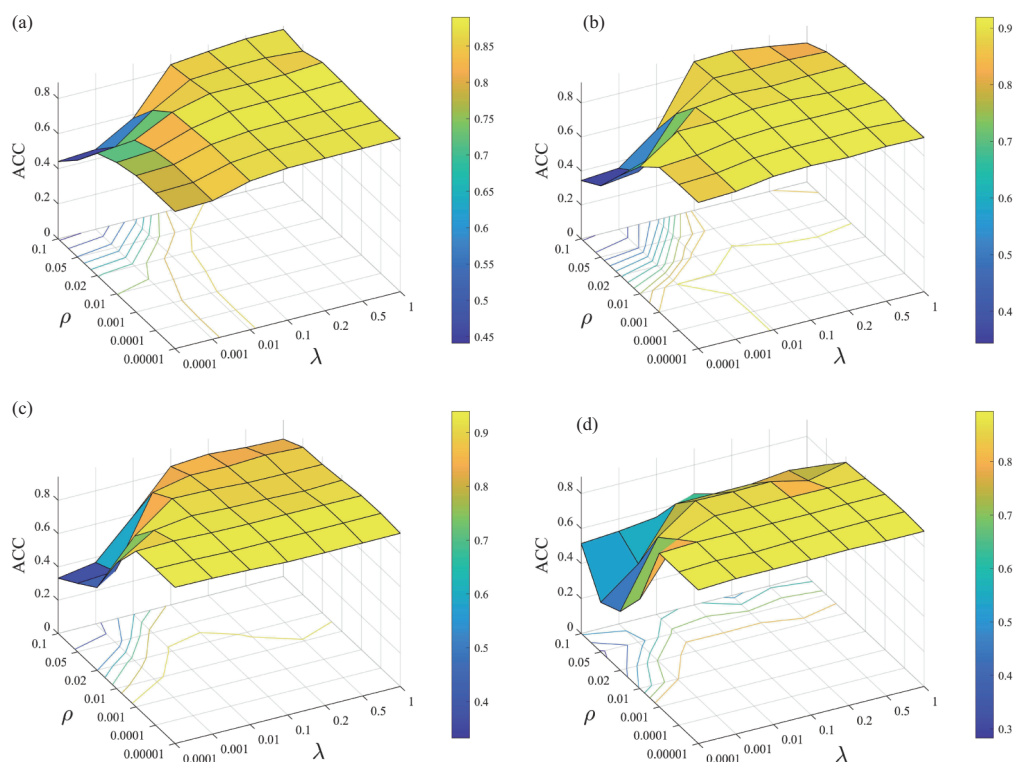


图 13 (网络版彩图) 参数敏感性实验. (a) $\theta = 10^{-2}$; (b) $\theta = 10^{-3}$; (c) $\theta = 10^{-4}$; (d) $\theta = 10^{-5}$

Figure 13 (Color online) Parameter sensitivity experiment. (a) $\theta = 10^{-2}$; (b) $\theta = 10^{-3}$; (c) $\theta = 10^{-4}$; (d) $\theta = 10^{-5}$

良好的分类能力和跨域能力. 因此, 所提方法能够有效解决非平衡数据下多工况的故障诊断问题, 具有很好的实用性.

参考文献

- 1 Guo L, Li N, Jia F, et al. A recurrent neural network based health indicator for remaining useful life prediction of bearings. *Neurocomputing*, 2017, 240: 98–109
- 2 Li X, Zhang W, Ding Q, et al. Diagnosing rotating machines with weakly supervised data using deep transfer learning. *IEEE Trans Ind Inf*, 2020, 16: 1688–1697
- 3 Li J, Wang Y, Zi Y, et al. A local weighted multi-instance multi-label network for fault diagnosis of rolling bearings using encoder signal. *IEEE Trans Instrum Meas*, 2020, 69: 8580–8589
- 4 Mao Z H, Gu Y X, Jiang B, et al. Incipient fault diagnosis for high-speed train traction systems via improved LSTM. *Sci Sin Inform*, 2021, 51: 997–1012 [冒泽慧, 顾彧行, 姜斌, 等. 基于改进 LSTM 的高速列车牵引系统微小渐变故障诊断. *中国科学: 信息科学*, 2021, 51: 997–1012]
- 5 Liu Q, Zhan Z Q, Wang S, et al. Data-driven multimodal operation monitoring and fault diagnosis of high-speed train bearings. *Sci Sin Inform*, 2020, 50: 527–539 [刘强, 詹志强, 王硕, 等. 数据驱动的高速列车轴承多模态运行监控与故障诊断. *中国科学: 信息科学*, 2020, 50: 527–539]
- 6 Jia F, Li S, Zuo H, et al. Deep neural network ensemble for the intelligent fault diagnosis of machines under imbalanced data. *IEEE Access*, 2020, 8: 120974–120982
- 7 He Y, Du C Y, Li C B, et al. Sensor fault diagnosis of superconducting fault current limiter with saturated iron core based on SVM. *IEEE Trans Appl Supercond*, 2014, 24: 1–5
- 8 Zhang J, Zhang Q, He X, et al. Compound-fault diagnosis of rotating machinery: a fused imbalance learning method.

- IEEE Trans Contr Syst Technol, 2020, 29: 1462–1474
- 9 Ali M Z, Shabbir M N S K, Liang X, et al. Machine learning-based fault diagnosis for single- and multi-faults in induction motors using measured stator currents and vibration signals. *IEEE Trans Ind Applicat*, 2019, 55: 2378–2391
 - 10 Liu Z H, Lu B L, Wei H L, et al. Deep adversarial domain adaptation model for bearing fault diagnosis. *IEEE Trans Syst Man Cybern Syst*, 2019, 51: 4217–4226
 - 11 Shen J, Li S, Jia F, et al. A deep multi-label learning framework for the intelligent fault diagnosis of machines. *IEEE Access*, 2020, 8: 113557–113566
 - 12 Luo L, Xie L, Su H. Deep learning with tensor factorization layers for sequential fault diagnosis and industrial process monitoring. *IEEE Access*, 2020, 8: 105494–105506
 - 13 Chen C L P, Liu Z. Broad learning system: an effective and efficient incremental learning system without the need for deep architecture. *IEEE Trans Neural Netw Learn Syst*, 2018, 29: 10–24
 - 14 Chen C L P, Liu Z. Broad learning system: a new learning paradigm and system without going deep. In: *Proceedings of the 32nd Youth Academic Annual Conference of Chinese Association of Automation (YAC)*, 2017. 1271–1276
 - 15 Pu X, Li C. Online semisupervised broad learning system for industrial fault diagnosis. *IEEE Trans Ind Inf*, 2021, 17: 6644–6654
 - 16 Cheng C, Wang W J, Chen H T, et al. Enhanced fault diagnosis using broad learning for traction systems in high-speed trains. *IEEE Trans Power Electron*, 2021, 36: 7461–7469
 - 17 Jia F, Lei Y, Lu N, et al. Deep normalized convolutional neural network for imbalanced fault classification of machinery and its understanding via visualization. *Mech Syst Signal Process*, 2018, 110: 349–367
 - 18 Feng Y, Wang T, Hu B, et al. An integrated method for high-dimensional imbalanced assembly quality prediction supported by edge computing. *IEEE Access*, 2020, 8: 71279–71290
 - 19 Yu L, Zhou R, Tang L, et al. A DBN-based resampling SVM ensemble learning paradigm for credit classification with imbalanced data. *Appl Soft Computing*, 2018, 69: 192–202
 - 20 Liu X Y, Wang S T, Zhang M L. Transfer synthetic over-sampling for class-imbalance learning with limited minority class data. *Front Comput Sci*, 2019, 13: 996–1009
 - 21 Liu Y, Wang Y, Ren X, et al. A classification method based on feature selection for imbalanced data. *IEEE Access*, 2019, 7: 81794–81807
 - 22 Zou Q, Xie S, Lin Z, et al. Finding the best classification threshold in imbalanced classification. *Big Data Res*, 2016, 5: 2–8
 - 23 Dhar S, Cherkassky V. Development and evaluation of cost-sensitive universum-SVM. *IEEE Trans Cybern*, 2017, 45: 806–818
 - 24 Dufrenois F. A one-class kernel fisher criterion for outlier detection. *IEEE Trans Neural Netw Learn Syst*, 2015, 26: 982–994
 - 25 Galar M, Fernandez A, Barrenechea E, et al. A review on ensembles for the class imbalance problem: bagging-, boosting-, and hybrid-based approaches. *IEEE Trans Syst Man Cybern C*, 2012, 42: 463–484
 - 26 Wan C, Cao Z J, Lee W J, et al. An adaptive ensemble data driven approach for nonparametric probabilistic forecasting of electricity load. *IEEE Trans Smart Grid*, 2021, 12: 5396–5408
 - 27 Wang Z, Cao C, Zhu Y. Entropy and confidence-based undersampling boosting random forests for imbalanced problems. *IEEE Trans Neural Netw Learn Syst*, 2020, 31: 5178–5191
 - 28 Zhang H, Liu W, Wang S, et al. Resample-based ensemble framework for drifting imbalanced data streams. *IEEE Access*, 2019, 7: 65103–65115
 - 29 Zhu J, Chen N, Shen C Q. A new deep transfer learning method for bearing fault diagnosis under different working conditions. *IEEE Sens J*, 2020, 20: 8394–8402
 - 30 Pan S J, Yang Q. A survey on transfer learning. *IEEE Trans Knowl Data Eng*, 2010, 22: 1345–1359
 - 31 Zhou Y X, Dong Y N, Zhou H K, et al. Deep dynamic adaptive transfer network for rolling bearing fault diagnosis with considering cross-machine instance. *IEEE Trans Instrum Meas*, 2021, 70: 1–11
 - 32 Li Y B, Ren Y, Zheng H L, et al. A novel cross-domain intelligent fault diagnosis method based on entropy features and transfer learning. *IEEE Trans Instrum Meas*, 2021, 70: 36293–36312
 - 33 Yu Y, Zhang C, Li Y, et al. A new transfer learning fault diagnosis method using TSC and JGSA under variable condition. *IEEE Access*, 2020, 8: 177287–177295

- 34 Zhang W, Li X, Ma H, et al. Universal domain adaptation in fault diagnostics with hybrid weighted deep adversarial learning. *IEEE Trans Ind Inf*, 2021, 17: 7957–7967
- 35 Dong S, He K, Tang B. The fault diagnosis method of rolling bearing under variable working conditions based on deep transfer learning. *J Braz Soc Mech Sci Eng*, 2020, 42: 585
- 36 Belkin M, Niyogi P, Sindhvani V. Manifold regularization: a geometric framework for learning from labeled and unlabeled examples. *J Mach Learn Res*, 2006, 7: 2399–2434
- 37 Long M, Wang J, Ding G, et al. Adaptation regularization: a general framework for transfer learning. *IEEE Trans Knowl Data Eng*, 2014, 26: 1076–1089
- 38 Wang J D, Feng W J, Chen Y Q, et al. Visual domain adaptation with manifold embedded distribution alignment. In: *Proceedings of the 26th ACM Multimedia Conference, Seoul, 2018*. 402–410
- 39 Zhu Z, Wang Z, Li D, et al. Geometric structural ensemble learning for imbalanced problems. *IEEE Trans Cybern*, 2020, 50: 1617–1629
- 40 Li Q Q, Liu X Y. Multi-class category imbalance learning algorithm: EasyEnsemble.M. *Patt Recognit Artif Intell*, 2014, 27: 187–192 [李倩倩, 刘胥影. 多类别不平衡学习算法: EasyEnsemble.M. *模式识别与人工智能*, 2014, 27: 187–192]
- 41 Pan S J, Tsang I W, Kwok J T, et al. Domain adaptation via transfer component analysis. *IEEE Trans Neural Netw*, 2011, 22: 199–210
- 42 Wang J D, Chen Y Q, Feng W J, et al. Transfer learning with dynamic distribution adaptation. *ACM Trans Intell Syst Technol*, 2020, 11: 1–25
- 43 Zhao H M, Zheng J J, Guo C, et al. Domain adaptive BLS model based on manifold regularization framework and MMD. *Act Autom Sin*, 2021. doi: 10.16383/j.aas.c210009 [赵慧敏, 郑建杰, 郭晨, 等. 基于流形正则化框架和 MMD 的域自适应 BLS 模型. *自动化学报*, 2021. doi: 10.16383/j.aas.c210009]
- 44 Ham J, Lee D D. Grassmann discriminant analysis: a unifying view on subspace-based learning. In: *Proceedings of the 25th International Conference Machine Learning (ICML)*, 2008. 376–383
- 45 Tzeng E, Hoffman J, Darrell T, et al. Simultaneous deep transfer across domains and tasks. In: *Proceedings of the 15th IEEE International Conference on Computer Vision (ICCV)*, Chile, 2015. 4068–4076
- 46 Chen Y, Yang C L, Zhang Y, et al. Deep conditional adaptation networks and label correlation transfer for unsupervised domain adaptation. *Pattern Recognit*, 2020, 98: 107072
- 47 Meng X, Qiao J F, Han H G. Soft measurement of key effluent parameters in wastewater treatment process using brain-like modular neural networks. *Act Autom Sin*, 2019, 45: 906–919 [蒙西, 乔俊飞, 韩红桂. 基于类脑模块化神经网络的污水处理过程关键出水参数软测量. *自动化学报*, 2019, 45: 906–919]
- 48 Han H G, Wu X L, Zhang L, et al. Identification and suppression of abnormal conditions in municipal wastewater treatment process. *Act Autom Sin*, 2018, 44: 1971–1984 [韩红桂, 伍小龙, 张璐, 等. 城市污水处理过程异常工况识别和抑制研究. *自动化学报*, 2018, 44: 1971–1984]

Multi-condition fault diagnosis based on adaptive transfer and broad learning under imbalanced data

Wukai JIANG¹ & Ping ZHOU^{1,2*}

1. *State Key Laboratory of Synthetical Automation for Process Industries, Northeastern University, Shenyang 110819, China;*

2. *Key Laboratory of Coal Processing and Efficient Utilization, Ministry of Education, China University of Mining and Technology, Xuzhou 221116, China*

* Corresponding author. E-mail: zhouping@mail.neu.edu.cn

Abstract This paper addresses the problem of multi-condition fault diagnosis under high-dimensional imbalanced data by proposing a dynamic adaptive transfer and broad learning method (TL-BLS) under the condition of imbalanced data and applies it to the multi-condition fault diagnosis of the sewage treatment process. First, given the problem of excessive differences under different working conditions in the industrial process, the transfer learning manifold transformation method is used to shorten the distance between the source and target domain data. Then, a two-stage, target domain pseudo-label prediction method considering data imbalance is proposed. The first stage uses a pseudo-label method based on inter-class relationships to obtain pseudo-labels with high reliability. The second stage uses an ensemble method considering subset differences to obtain a more reliable pseudo-label that also considers the data imbalance problem. Finally, an adaptive transfer and broad learning method is established, with a manifold regularization term and dynamic balance factor. A weight matrix is also introduced into the broad learning system and the adaptive transfer method to solve the problem of data imbalance and realize fault diagnosis under a multi-condition process. The data experiment and comparative analysis of the sewage treatment process show that the proposed method has a better fault identification capability and practicability.

Keywords fault diagnosis, transfer learning, broad learning system, manifold learning, ensemble learning, sewage treatment process, imbalanced data