

中文引用格式:陈岑,纪育博,王欢,等. 基于迁移学习的燃气管网泄漏定位方法[J]. 中国安全科学学报,2025,35(3):212-220.

英文引用格式:CHEN Cen,JI Yubo,WANG Huan,et al. Gas pipeline network leak localization method based on transfer learning[J]. China Safety Science Journal,2025,35(3):212-220.

基于迁移学习的燃气管网泄漏定位方法*

陈岑¹,纪育博²高级工程师,王欢²高级工程师,聂荣山^{3,4},梁晓瑜^{**1,3}教授

(1 中国计量大学 计量测试与仪器学院,浙江 杭州 310018;2 宁波华润兴光燃气有限公司,

浙江 宁波 315010;3 中国计量大学 能源环境与安全工程学院,浙江 杭州 310018;

4 中国安全生产科学研究院 交通安全研究所,北京 100012)

中图分类号:X928.03

文献标志码:A

DOI: 10.16265/j.cnki.issn1003-3033.2025.03.0957

基金项目:国家市场监督管理总局科技计划项目(2023MK230);国家自然科学基金面上项目资助(51871206)。

【摘要】 为增强燃气管网运行的可靠性与安全性,提高燃气管网泄漏故障的诊断能力,解决真实燃气管网泄漏数据样本稀缺及工况差异影响问题,提出基于迁移学习的燃气管网泄漏定位方法。首先,采用随机森林特征重要性排序方法,选取出 TGNET 仿真管网的 5 个压力监测点;然后,将 3 种不同压力工况下的压力监测点数据分别作为源域和目标域,输入特征,改进迁移学习传统联合概率分布适应(JDA)方法,以减小源域与目标域特征距离;最后,采用布谷鸟搜索(CS)算法,优化改进迁移学习算法的参数(映射后维度 d' 和学习率 λ),实现无标签目标域泄漏管段的诊断。结果表明:所提复杂燃气管网泄漏定位方法可以有效提高无标签燃气管网泄漏识别效果,相比传统联合概率分布适应有更高的准确率。

【关键词】 迁移学习; 燃气管网; 泄漏定位; 随机森林; 布谷鸟搜索(CS)算法; 联合分布自适应(JDA)

Gas pipeline network leak localization method based on transfer learning

CHEN Cen¹, JI Yubo², WANG Huan², NIE Rongshan^{3,4}, LIANG Xiaoyu^{1,3}

(1 College of Metrology and Instrument, China Jiliang University, Hangzhou Zhejiang 310018, China;

2 Ningbo China Resources Xingguang Gas Co., Ltd., Ningbo Zhejiang 315010, China; 3 College of

Energy Environment and Safety Engineering, China Jiliang University, Hangzhou Zhejiang 310018, China;

4 Traffic Safety Research Institute, China Academy of Safety Science and Technology, Beijing 100012, China)

Abstract: In order to enhance the reliability and safety of gas network operations and improve the fault diagnosis capabilities for gas network leaks, while addressing issues such as the scarcity of real gas network leak data samples and variations in operating conditions, a gas network leak localization method based on transfer learning was proposed. Firstly, the Random Forest feature importance ranking method was used to select five pressure monitoring points in the TGNET simulation network. Subsequently, pressure monitoring point data under three different pressure conditions were respectively used as the source domain and target

* 文章编号:1003-3033(2025)03-0212-09; 收稿日期:2024-10-19; 修稿日期:2024-12-20

** 通信作者:梁晓瑜(1975—),男,安徽宿州人,博士,教授,博士生导师,主要从事燃气安全运维与智慧计量等方面的研究。E-mail:xyliang@cjl.edu.cn。

domain input features. The traditional JDA method of transfer learning was improved to reduce the feature distance between the source domain and the target domain. Furthermore, the CS algorithm was employed to optimize the dimensionality after mapping d' and the learning rate λ parameters of the improved transfer learning algorithm, ultimately achieving the diagnosis of unlabeled target domain leak segments. The results indicated that the proposed leak localization method for complex gas networks can effectively improve the localization accuracy of unlabeled gas network leaks, achieving higher accuracy compared to traditional.

Keywords: transfer learning; gas pipeline network; leak localization; random forest; cuckoo search(CS) algorithm; joint distribution adaptation(JDA)

0 引言

天然气管网作为重要的能源输送途径,一旦发生泄漏事故可能会导致重大财产损失和人员伤亡^[1-2]。因此,有必要监控燃气管道的运行情况,并采取预防措施避免或减少泄漏^[3]。

目前,燃气管道运行过程监测方法主要分为基于硬件、基于解析模型和基于数据驱动的方法^[4]。近年来,随着大数据技术和机器学习的发展,基于数据驱动的方法逐渐成为燃气管道监测的主流选择。机器学习算法主要通过对采集的管网运行压力、流量等信号进行实时特征值提取与分析来实现管网泄漏的检测与位置识别。GUPTA 等^[5]采用贝叶斯概率方法基于管网流量和压力数据,实现了低压燃气管网的泄漏检测和定位。贾文龙等^[6]基于支持向量机建立管道泄漏信号识别模型,实现了对输气管网压降速率信号的泄漏特征识别。基于传统机器学习理论对燃气管道异常状况的监测与分类依赖于大量带标签的数据,而在实际工况下燃气管道泄漏数据较少,而且对所有工况的泄漏情况进行标注是不实际的,导致传统机器学习方法在不同工况间识别效果不理想。YANG Tao 等^[7]将迁移学习和深度学习相结合,应用于图像分类和时间序列预测任务,以处理没有源域数据的迁移场景;朱良玉等^[8]将改进的果蝇优化算法和主题相关性分析结合,解决了不同工况下轴承振动数据分布不一致且目标工况轴承诊断数据不足的难题;JIAO Jinyang 等^[9]针对工业生产中数据分布差异,提出了残差关节自适应对抗网络迁移学习框架,对轴承和齿轮箱故障进行了智能诊断。迁移学习在故障诊断领域表现出良好的效果,它通过在不同任务之间共享知识和特征,可以利用已有的数据,在新的任务上进行快速学习和调整,实现不同工况的故障诊断识别。然而,将迁移学习应用于管网泄漏定位的研究较少,且传统迁移学习

方法在源域与目标域的距离衡量和迁移学习参数选择上仍存在局限性。

鉴于此,笔者拟改进迁移学习传统的联合分布自适应(Joint Distribution Adaptation, JDA)方法,首先,采用随机森林算法选取特征数据作为源域和目标域数据;然后,采用能量距离方法计算平衡因子 μ ,以实现更准确的均值差异(Maximum Mean Discrepancy, MMD)计算,从而减小源域和目标域数据的映射距离;最后,运用布谷鸟搜索(Cuckoo Search, CS)算法优化改进联合概率分布参数,提高分类准确率,以期实现不同工况下异常管段的监测识别。

1 基础理论

1.1 随机森林

随机森林算法属于集成学习方法 bagging 类,基于决策树算法,通过对样本进行有放回的随机抽样,生成多个样本子集,每个子集用于构建一棵决策树,经过多次抽样,形成多棵决策树,组成随机森林^[10]。通过计算每个特征在随机森林中的平均不纯度减少量(Mean Decrease Impurity, MDI)方法,可实现各特征对目标变量预测的相对重要性衡量与排序。MDI方法衡量的是每个特征对于降低决策树节点的不纯度贡献的程度,贡献越大的特征,其重要性越高^[11]。

1.2 迁移学习

迁移学习方法不需要满足测试数据和训练数据是独立分布^[12],在不同工况下燃气管网的泄漏识别中展现出了优势。运用基于特征的迁移方法^[13],将源域和目标域的数据特征变换到统一特征空间中,并利用机器学习进行识别分类,判断目标域的数据分类,达到迁移的目的。基于特征的迁移方法结构如图1所示。

对于管网泄漏定位,不同的气源压力会造成源

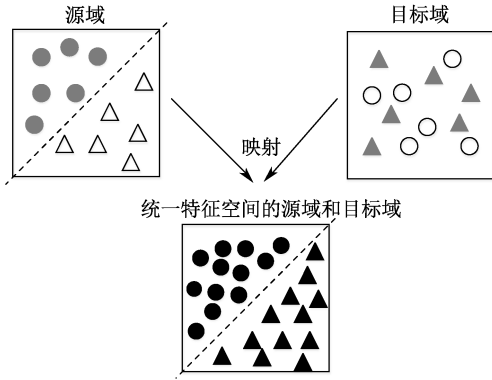


图1 基于特征的迁移方法结构

Fig.1 Feature transfer method structure

域与目标域数据的不同分布,因此,采用 JDA,拉近源域和目标域的距离。JDA 结合了条件分布 $P(y|x)$ 和边缘分布 $P(x)$ ^[14],计算 2 种分布距离之和作为源域和目标域的距离,通过最小化该距离,实现自适应学习,其中, x 和 y 为随机变量, $P(\cdot)$ 为概率分布。JDA 的距离衡量表示为:

$$S(D_s, D_t) \approx \| P(x_s) - P(x_t) \| + \| P(y_s | x_s) - P(y_t | x_t) \| \quad (1)$$

式中: $S(D_s, D_t)$ 为源域 D_s 和目标域 D_t 的距离; x_s 和 x_t 分别为源域的输入数据和目标域输入数据; y_s 和 y_t 分别为源域数据标签和目标域数据标签。

使用特征映射函数 φ 将输入特征映射到一个共享的特征空间,使得在特征空间中的源域和目标域分布尽可能接近,实现源域和目标域特征对齐。使用 MMD^[15] 度量 2 个分布之间的差异, MMD 公式为:

$$MMD(x_s, x_t) = \left\| \frac{1}{n_1} \sum_{i=1}^{n_1} \varphi(x_i) - \frac{1}{n_2} \sum_{j=1}^{n_2} \varphi(x_j) \right\|_H \quad (2)$$

式中: n_1, n_2 分别为源域和目标域的样本数量; $\varphi(x_i), \varphi(x_j)$ 为样本经过映射后的值; $\|\cdot\|_H$ 为再生核希尔伯特空间。通过最小化 MMD 距离,让源域和目标域的特征分布在特征空间中尽可能接近,从而实现迁移学习。

JDA 方法将源域和目标域之间的边缘分布和条件分布直接相加,忽略了实际数据中边缘分布自适应和条件分布自适应重要性差异,影响 MMD 距离计算的准确性。因此,改进传统 JDA 方法,引入平衡因子 μ ,对于不同的分类任务,分别计算数据边缘分布和条件分布的距离占比,得到平衡因子 μ ,从而根据数据特点自适应地调整边缘分布和条件分布在分布适配过程中的重要性。改进后的距离衡量表示

公式如下:

$$S(D_s, D_t) \approx (1 - \mu)S(P(x_s), P(x_t)) + \mu S(P(y_s | x_s), P(y_t | x_t)) \quad (3)$$

平衡因子 μ 采用能量距离方法进行近似计算。应用 K 最近邻算法求出能量距离值,并以整体域分布的能量距离值与整体和局部能量距离总和的比值来计算 μ ,具体计算公式如下:

$$\mu = \frac{E(\varphi(x_s), \varphi(x_t))}{\frac{1}{c} \sum_{i=1}^c E(x_{s_i}, x_{t_i}) + E(\varphi(x_s), \varphi(x_t))} \quad (4)$$

式中: $E(\cdot)$ 为能量距离; c 为样本类别总数; x_{s_i} 和 x_{t_i} 分别为源域和目标域各类别的样本特征; i 为源域和目标域类别数。改进后的 MMD 矩阵表示如下:

$$MMD = (1 - \mu)MMD(x_s, x_t) + \mu \sum_{i=1}^c MMD(x_{s_i}, x_{t_i}) \quad (5)$$

1.3 CS 算法

CS 算法是一种启发式算法,它以布谷鸟育雏的寄生行为和果蝇的莱维飞行行为为启发来解决复杂的优化问题^[16],具有较强的全局搜索能力且简单易实现,表现出良好的实际效果。CS 算法首先随机生成若干初始解 X_0 ,然后利用莱维飞行生成新的解 X_t 。计算新解的适应度,如果新解的适应度优于现有解,则替换,选择出最佳解。莱维飞行采用步长服从莱维分布的随机游走策略,如下式^[17]:

$$X_{t+1,e} = X_{t,e} + \alpha_0 \frac{\sigma u}{|v|^{1/\beta}} (X_{t,e} - X_t^*) \quad (6)$$

式中: $X_{t,e}$ 为 t 代的第 e 个解; X_t^* 为第 t 代最优解; α_0 为步长常数; u 和 v 为正态分布随机变量; β 为方向参数,取 1.5; σ 为尺度, σ 表达式为:

$$\sigma = \left(\frac{\Gamma(1 + \beta) \sin(\pi\beta/2)}{\Gamma\left(\left(\frac{1 + \beta}{2}\right)\beta 2^{(\beta-1)/2}\right)} \right)^{1/\beta} \quad (7)$$

式中 Γ 为伽马函数。

2 燃气管网泄漏定位方法

2.1 随机森林选取泄漏监测特征点

燃气管网存在多个节点可作为压力监测点监测压力数据,从而判断管网泄漏情况,而过多的数据可能会存在相关特征,产生冗余,增加数据处理量,因此,需要对数据特征点即压力监测点进行选择。运用随机森林的特征重要性排序方法选出合理的压力

监测点。具体步骤如下:

1) 采用不同管段泄漏状态和正常状态下的各节点压力数据构建数据集,将是否泄漏和泄漏管段位置作为分类依据对数据标记标签。

2) 将带标签数据集作为输入数据,使用随机森林分类器对输入数据进行拟合训练,建立分类模型。

3) 通过计算每个特征在随机森林中的 MDI 来确定特征的重要性,计算返回并输出各特征点的重要性评分和排序。

4) 根据返回各监测点特征重要性评分,测试不同数量监测点分类准确率,得到压力监测点。

2.2 基于 CS 算法的迁移学习识别泄漏点

通过对 JDA 的改进,可以更准确地估计出源域与目标域之间的分布差异,从而达到减小两域分布差异,并用源域数据标签对目标域进行泄漏定位的目的。由于迁移映射后维度 d' 和学习率 λ 这 2 个参数会对迁移学习效果产生影响,因此,采用 CS 算法优化 d' 和 λ 的值,以获得更高的准确率。该方法的计算流程如图 2 所示,具体步骤如下:

1) 采集压力监测点的数据,选择其中一种工况下正常状态及各管段泄漏状态时各监测点压力数据,并标记标签,作为源域数据集,将其余不同压力工况下各点压力数据作为目标域数据。

2) 初始化 CS 算法和迁移学习参数,以分类准确率的负数 $-acc$ 作为适应度值。

3) 运用能量距离方法计算两域的边缘分布和条件分布平衡因子 μ , 计算 MMD 矩阵,返回经过转换后的源域和目标域数据,用于后续的分类任务。

4) 基于 K 最近邻的方法将转换后的源域和目标域数据,在目标域上进行泄漏分类,得到目标域的伪标签。

5) 将伪标签代入 MMD 矩阵进行计算,重复步骤 3) 和 4),直到达到迁移学习迭代次数,输出管路泄漏状况诊断结果和准确率。

6) 利用莱维飞行公式生成 d' 和 λ 的新解,并进行适应度值评估和更新最优解。

7) 重复步骤 3—步骤 6,直到达到 CS 算法的迭代次数,输出 d' 和 λ 的最优值、诊断结果和准确率。

3 数据处理及泄漏定位效果分析

3.1 燃气管网泄漏数据来源

搭建管网泄漏监测试验平台,如图 3 所示。试验平台主要由气源模块、试验管网和数据采集模块

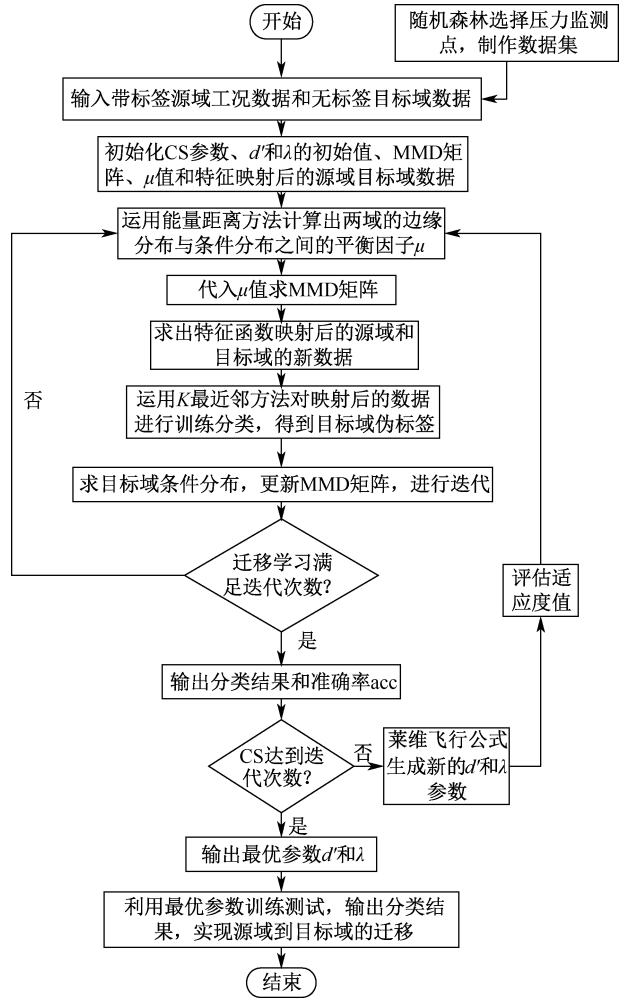


图 2 基于 CS 的改进 JDA 方法流程

Fig. 2 Improved JDA method flow based on CS

组成。气源模块由空气压缩机和气罐构成;考虑实际燃气管网的拓扑结构及运行参数,根据相似性理论,试验管网设计了包含多分支节点、分支回路、枝状管道、网状管道多种管路类型,采用外径 50 mm,内径为 3.7 mm 的聚氯乙烯管道,在 15 处管段中间设置带刻度的球阀以模拟 15 处泄漏(L1—L15);数据采集模块主要由流量计、压力变送器、数据采集卡和上位机组成,压力变送器分别设置在 P1—P11 共 11 个管网节点附近,采集管网节点压力数据,并通过数据采集卡将采集的数据传送至上位机。

图 4 为 0.3 MPa 工况下,管网由正常运行状态到泄漏模拟点 L3 发生泄漏,直至管网压力重新达到平稳状态,各压力变送器测量的压力试验数据经过平滑去噪处理后图像。管网发生泄漏,各压力监测点压力值由平稳状态开始下降,并最终在较低的压力值达到新的平衡,由于各压力监测点与泄漏点位置距离不同,各压力监测点压力下降值存在差异,因

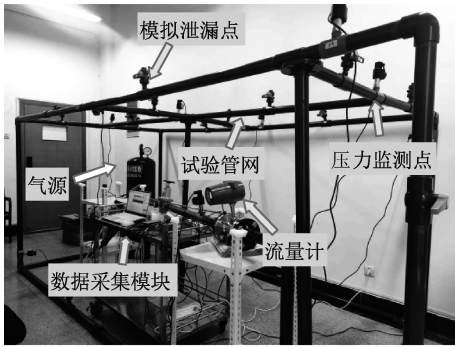


图3 管网泄漏监测试验平台

Fig. 3 Pipeline leakage monitoring experimental platform

此,可以利用泄漏之后管网各节点的压力值对泄漏点进行定位。

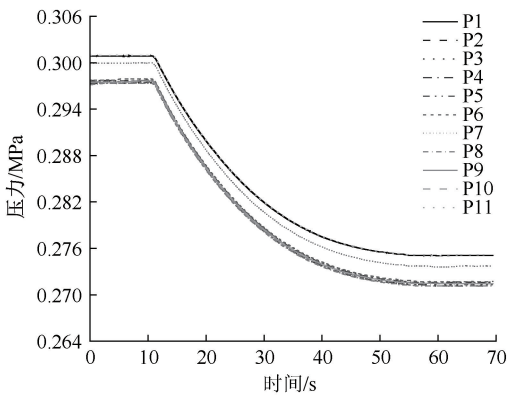


图4 平滑去噪后压力监测数据

Fig. 4 Smoothed and denoised pressure monitoring data

TGNET 是天然气工业中常用的输气管网仿真模拟软件,可对复杂燃气管网进行稳态和瞬态模拟,精确模拟燃气管网在不同工况下的运行状况,被运用于管网调整优化、输气调度方案制定等方面,其有效性得到广泛实践验证^[18]。

文中使用 TGNET 软件对管网泄漏监测试验平台的管网结构进行仿真模拟,管网模拟结构如图 5 所示。将试验管网基本属性数据和管网运行的监测数据代入仿真模型,并利用试验数据对仿真模型进行校核。经过校核后,更新的管网模型再次进行数据仿真,以获得不同工况下的中低压配气管网运行状况数据。将试验数据与 TGNET 仿真数据进行比较,结果显示,试验值与仿真值的相对误差小于 5%,表明仿真数据与实际情况能够吻合,仿真数据具有合理性和有效性。

在 TGNET 模拟过程中,分别调整气源点的压力值为 0.4、0.3 和 0.2 MPa,模拟燃气管网不同的压力工况,数据采集的间隔时间为 3.6 s,通过调整不同时刻的用气点流量值,模拟不同气体流量下管网

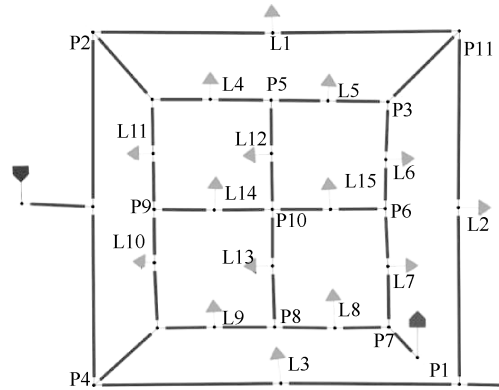


图5 管网模拟结构

Fig. 5 Pipeline simulation structure

运行状况。为评估管网不同管段泄漏识别和定位有效性,分别设置管网正常运行状态和 15 个管段中点分别泄漏,共 16 种情况,每种运行状态分别采集 150 组数据。

采集 16 种运行状态下 14 个管网节点的压力监测数据,每组压力数据形式为 $X_k = (x_1, x_2, \dots, x_{14})$, k 为每个不同的采样时刻, $x_1 \sim x_{14}$ 是 14 个压力监测点在时刻 k 的监测数据。将数据和对应的标签代入 2.1 节所述随机森林方法,选取特征重要性评分大于 0.05 的 10 个压力监测点,按照特征重要性排名高到低的次序依次增加压力监测点数量,分别制作 1~10 个压力特征的数据集,运用支持向量机 (Support Vector Machine, SVM) 方法对 3 种压力工况下不同压力监测点数量的数据集进行分类,分类结果如图 6 所示。由此可知:在 3 种压力工况下,当压力监测点数为 5 时,分类准确率接近 100%,且压力监测点数大于 5 时,增加压力监测点对分类结果不产生影响。因此,选用特征重要性排名前 5 的压力监测点数据进行试验,每组压力数据形式为 $X_k = (x_1, x_2, \dots, x_5)$ 。

3.2 数据处理结果

为体现数据的时间序列特性,增加数据的特征维度,对采集的数据按时间序列进行滑动窗口处理,设置窗口大小为 10,移动步长为 1,每移动一个窗口,即每采样 10 组数据,将 10 个采样点的各监测点压力值作为数据特征值,每组压力数据形式为 $X_k = (x_{1,1}, x_{1,2}, \dots, x_{1,5}, x_{2,1}, x_{2,2}, \dots, x_{2,5}, \dots, x_{5,1}, \dots, x_{5,10})$,其中, $x_{5,10}$ 代表第 5 个压力监测点的第 10 时刻压力值。

通过对数据标注标签,标签内容为泄漏管段序号,其中 1—15 代表 15 个管段分别泄漏,16 代表未发生泄漏,最终得到的数据集见表 1。每种压力工况最

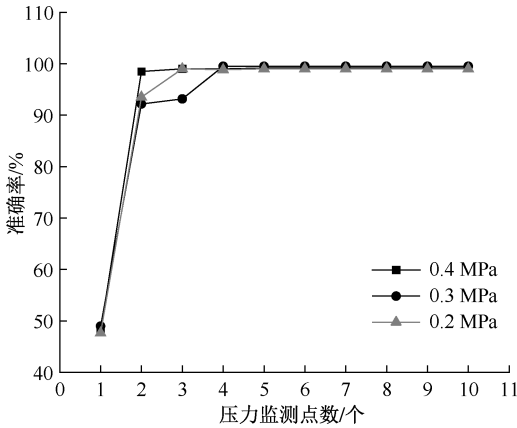


图6 压力监测点数量与准确率的关系

Fig. 6 Relationship between the number of pressure monitoring points and accuracy

终得到 2 391 组样本数据,每组数据样本维度为 50。

表1 压力监测数据集

Table 1 Pressure monitoring dataset

气源压力/MPa	泄漏标签	样本数
0.4	1-16	2 391
0.3	1-16	2 391
0.2	1-16	2 391

将以上 3 种压力工况下的 3 组数据集分别作为源域,其余 2 种压力工况作为目标域,以此验证迁移学习对泄漏管段的分类效果,可以进行 6 组试验。试验选用的核函数为 primal,运用 2.2 节所述基于 CS 的迁移学习方法,在 python3.9 平台对数据进行训练,迁移学习映射后维度 d' 和正则化强度 λ 的参数设置范围和最优解见表 2,基于 CS 的迁移学习泄漏定位效果见表 3。

表2 优化参数范围和结果

Table 2 Optimize parameter range and results

参数	下限	上限	最优解
d'	10	100	16
λ	0.001	1	0.597

表3 迁移学习不同工况泄漏定位效果

Table 3 Leakage localization effect under different working conditions in transfer learning

源域工况压力/MPa	目标域工况压力/MPa	准确率/%	平均准确率/%
0.4	0.3	99.85	97.79
	0.2	95.73	
0.3	0.2	99.97	99.97
	0.4	99.96	
0.2	0.3	99.92	93.08
	0.4	86.24	

由以上分析结果可知:经过文中所述迁移学习方法之后,源域数据对目标域数据的识别准确率呈现良好的效果,3 种工况迁移的平均准确率都达到 90% 以上。这表明所采用的迁移学习方法有效地利用了源域知识,提升了目标域数据的分类性能。在不同工况下的试验结果中,压力相差较小的工况迁移效果显著优于压力相差更大的工况,分析原因为压力工况相差较大,导致源域和目标域之间的样本分布差异更为显著,较难以识别数据间的关联性与共性,进而降低了识别准确率。

利用 t-分布随机邻域嵌入 (t-Distributed Stochastic Neighbor Embedding, t-SNE) 图^[19],将迁移前后数据降维到二维空间进行数据可视化,直观体现所述方法进行迁移学习的效果。3 种工况数据迁移前后的 t-SNE 图如图 7 所示。图 7 中,不同灰度的散点代表不同泄漏类别,其中图 7a—图 7c 分别是目标域 0.2、0.3 和 0.4 MPa 的初始数据二维空间表示。由图 7a—图 7c 可知:初始数据不同灰度的数据点混杂在一起,除少数分散点外,其余点呈现聚集性,但是聚集点类别多样,无明显规律,在无目标域标签的情况下难以将各类进行分离。图 7d—图 7f 和图 7g—图 7i 分别是以其余 2 种工况压力数据作为源域对第一行初始目标域数据进行迁移分类后的结果。由此可知:运用源域数据和标签对目标域数据进行迁移之后,数据按类别聚集,能够明显看出 16 种类别之间的类间距,可以实现较好的分类效果。

为验证文中提出的迁移学习方法相比于传统机器学习方法的优势,将该方法与其他传统机器学习方法进行对比。对比的机器学习方法包括 SVM、随机森林和 K 最近邻方法。对于传统机器学习方法,使用 0.4、0.3、0.2 MPa 等 3 种不同压力工况数据作为训练集。测试集和验证集数据则由 3 种工况数据按比例混合而成。训练集、验证集和测试集数据按 2 : 1 : 1 比例进行划分并随机打乱。分类结果见表 4。由表 4 中不同方法进行的管道泄漏诊断准确率结果可知:传统机器学习方法对不同工况管道泄漏的诊断识别能力明显弱于迁移学习方法,该现象归因于传统机器学习假设训练集和测试集是独立同分布,从而限制了它们在跨领域应用中的表现。而迁移学习能够学习数据高层次特征,增强了模型在不同工况下的泛化能力,在特征提取方面表现出色。因此在目标域工况数据标签不足时,通过迁移学习方法,用源域数据和标签对目标域工况泄漏进行定

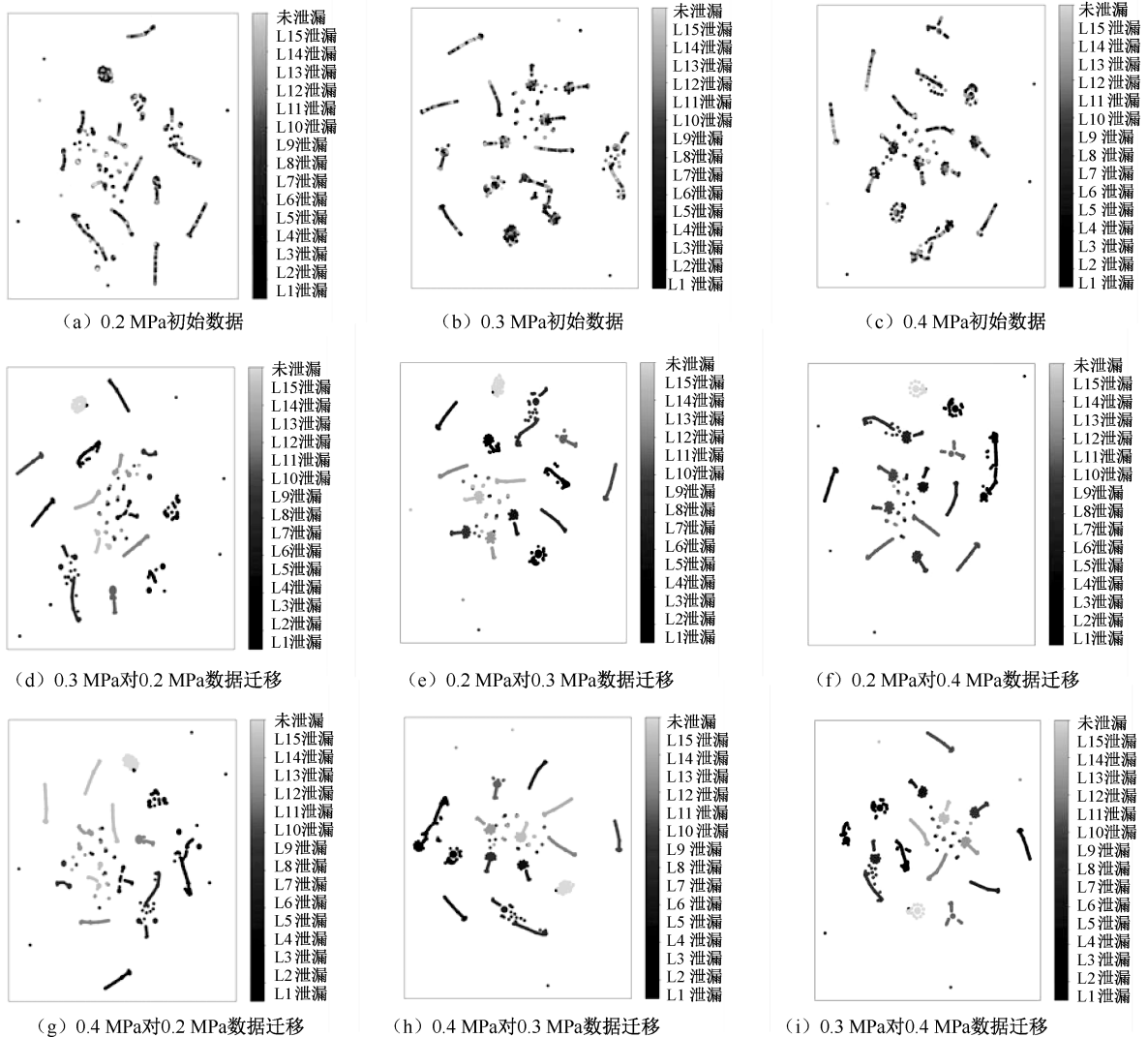


图7 数据迁移效果 t-SNE 图

Fig.7 The t-SNE plot of data migration effect

位识别可以替代传统机器学习方法,并取得更高的准确率。试验结果证明了文中提出的迁移学习方法对比传统机器学习方法的巨大优势。

表4 不同算法的泄漏定位准确率

Table 4 Leakage localization accuracy of different algorithms

任务	SVM	随机森林	K 最近邻	文中方法
0.4 MPa→0.3 MPa	36.96	37.68	35.98	99.85
0.4 MPa→0.2 MPa				95.73
0.3 MPa→0.2 MPa	36.99	37.29	36.07	99.97
0.3 MPa→0.4 MPa				99.96
0.2 MPa→0.3 MPa	37.02	37.57	36.43	99.92
0.2 MPa→0.4 MPa				86.24

为验证采用随机森林算法挑选压力监测点,引入平衡因子和能量距离计算方法改进 JDA 的域距

离计算,并用 CS 算法优化改进算法的参数 d' 和 λ 方法各步骤的有效性,通过消融试验对比改进前 JDA 算法、未优化参数、未进行压力监测点挑选和用本文方法进行迁移任务的准确率,分析结果如图 8 所示。由图 8 可知:提出的迁移学习方法相比传统 JDA 方法有更好的分类效果,并且未挑选压力监测点和未对迁移学习参数进行优化都对分类的准确率产生了消极影响。

4 结论

1) 提出改进的迁移学习方法,该方法对 3 种不同压力工况下管网泄漏的 6 个迁移任务的平均准确率超过 90%,可实现良好的分类效果,且分类准确率远大于传统机器学习方法。未带标签的管网压力数据可以用不同工况下带标签的管网压力数据进行

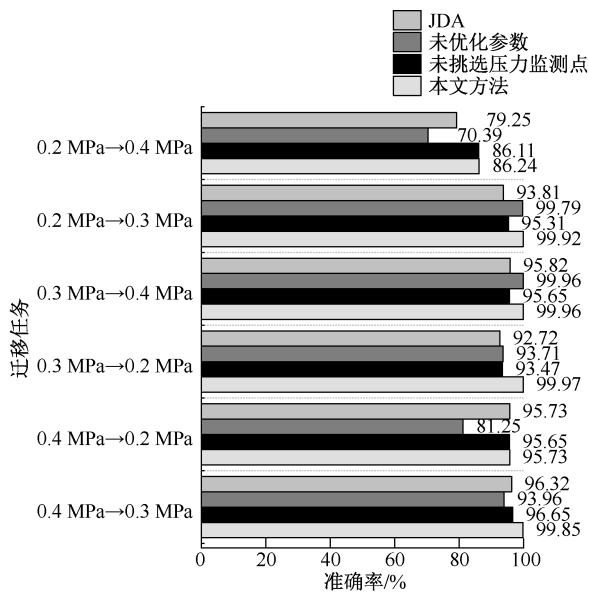


图8 消融试验结果

Fig. 8 Ablation experiment results

迁移分类,从而识别泄漏。

2) 消融试验结果证明:挑选压力监测点、改进JDA及优化迁移学习参数,均对泄漏分类准确率产生积极作用。改进的迁移学习方法准确率高于传统JDA方法。

3) 通过对6个迁移任务准确率的对比可知:工况压力值更接近的数据,迁移分类效果优于压力值差距更大的数据,因此,在实际对目标域的迁移训练任务中,可以选择工况更接近目标域的数据作为源域。

4) 文中提出的管网泄漏定位方法虽能有效地在不同工况下准确定位泄漏管段,但尚无法精确确定泄漏点具体位置。未来研究可以通过扩充数据集,并将深度学习方法与迁移学习相结合,实现不同工况下管网泄漏点的具体定位。

参考文献

- [1] LIU Yiming, BAO Yi. Review on automated condition assessment of pipelines with machine learning[J]. Advanced Engineering Informatics, 2022, 53: 1-5.
- [2] 向鹏成,张子薇,李卉,等. 城市地下燃气管道泄漏事故致因因素系统分析[J]. 中国安全科学学报, 2023, 33(12): 140-147.
XIANG Pengcheng, ZHANG Ziwei, LI Hui, et al. Systematic analysis on causative factors of gas pipelines leakage accidents [J]. China Safety Science Journal, 2023, 33(12): 140-147.
- [3] KORLAPATI N V S, KHAN F, NOOR Q, et al. Review and analysis of pipeline leak detection methods[J]. Journal of Pipeline Science and Engineering, 2022, 2(4): 1-17.
- [4] 康俊鹏. 基于压力信号处理的油气管道泄漏检测技术[J]. 石化技术, 2015, 22(11): 93.
KANG Junpeng. Oil and gas pipelines leakage detection based on pressure signal processing [J]. Petrochemical Industry Technology, 2015, 22(11): 93.
- [5] GUOTA P, GOYAL A, DAUWELS J, et al. Bayesian detection of leaks in gas distribution networks[C]. IECON 42nd Annual Conference of the IEEE Industrial Electronics Society, 2016: 855-860.
- [6] 贾文龙,孙溢彬,汤丁,等. 基于支持向量机的输气管道泄漏压降信号智能识别方法[J]. 化工进展, 2022, 41(9): 4 713-4 722.
JIA Wenlong, SUN Yibin, TANG Ding, et al. Intelligent identification method of pressure drop signals for gas pipeline leakage based on support vector machine [J]. Chemical Industry Progress, 2022, 41(9): 4 713-4 722.
- [7] YANG Tao, YU Xia, MA Ning, et al. Deep representation-based transfer learning for deep neural networks[J]. Knowledge-Based Systems, 2022, 253: 1-6.
- [8] 朱良玉,陶林,胡超凡,等. 基于 IFOA-TCA 的跨工况轴承故障诊断[J]. 轴承, 2023(4): 73-79.
ZHU Liangyu, TAO Lin, HU Chaofan, et al. Fault diagnosis for bearings under crossed operating conditions based on IFOA-TCA[J]. Bearing, 2023(4): 73-79.
- [9] JIAO Jinyang, ZHAO Ming, LIN Jing, et al. Residual joint adaptation adversarial network for intelligent transfer fault diagnosis[J]. Mechanical Systems and Signal Processing, 2020, 145: 1-4.
- [10] 郭慧,刘明艳. 基于遗传算法和随机森林的入侵检测方法研究[J]. 计算机应用与软件, 2024, 41(1): 304-309, 314.
GUO Hui, LIU Mingyan. Intrusion detection method based on genetic algorithm and random forest [J]. Computer Applications and Software, 2024, 41(1): 304-309, 314.
- [11] 李红霞,徐浩冉,田水承. 基于随机森林的矿工不安全行为预测预警模型[J]. 中国安全科学学报, 2022, 32(12):

10-18.

LI Hongxia, XU Haoran, TIAN Shuicheng. A prediction and early warning model of miners unsafe behavior based on random forest [J]. China Safety Science Journal, 2022, 32(12): 10-18.

- [12] 姜宇迪. 基于无监督深度迁移学习的电梯制动器实时状态监测和预警功能安全系统研究[D]. 上海: 上海交通大学, 2020.
- JIANG Yudi. Functional safety system of elevator brake real-time status monitoring and prognostics based on unsupervised deep transfer learning[D]. Shanghai: Shanghai Jiaotong University, 2020.
- [13] ANVAR A A T, MOHAMMADI H. A novel application of deep transfer learning with audio pre-trained models in pump audio fault detection[J]. Computers in Industry, 2023, 147: 1-5.
- [14] WANG Pengju, LU Luxi, LI Jian, et al. Transfer learning with joint distribution adaptation and maximum margin criterion[J]. Journal of Physics: Conference Series, 2019, 1169(1): 1-9.
- [15] LI Yibai, WANG Haoyun, ZHANG Yuzhuo, et al. Inversion of the optical properties of apples based on the convolutional neural network and transfer learning methods[J]. Applied Engineering in Agriculture, 2022, 38(6): 931-939.
- [16] XING Bo, GAO Wenjing, MARWLA T. An overview of cuckoo-inspired intelligent algorithms and their applications[C]. 2013 IEEE Symposium on Swarm Intelligence, 2013: 85-89.
- [17] WANG Lijin, ZHONG Yiwen. Cuckoo search algorithm with chaotic maps[J]. Mathematical Problems in Engineering, 2015, 2015(1): 1-14.
- [18] 李兴泉, 马玉兴. 基于 TGNET 软件的天然气管网动静态模拟分析[J]. 节能, 2023, 42(4): 19-21.
- LI Xingquan, MA Yuxing. Dynamic and static simulation analysis of natural gas pipeline networks based on TGNET software[J]. Energy Conservation, 2023, 42(4): 19-21.
- [19] LAURENS V D M, HINTON G. Visualizing data using t-SNE[J]. Journal of Machine Learning Research, 2008, 9: 2579-2605.

作者简介: 陈岑 (2000—), 女, 江西井冈山人, 硕士研究生, 研究方向为燃气管网异常诊断。E-mail: 1621179538@qq.com。

