

DOI: 10.19650/j.cnki.cjsi.J2209864

基于选择性加权适配网络的多域新故障识别方法*

陈祝云^{1,2,3}, 林慧斌¹, 夏景演¹, 晋刚¹, 李巍华^{1,2,4}

(1. 华南理工大学机械与汽车工程学院 广州 511442; 2. 琶洲实验室 广州 510005; 3. 北京信息科技大学重点科研机构 北京 100192; 4. 华南理工大学吴贤铭智能工程学院 广州 511442)

摘要:针对深度迁移学习诊断方法要求机械设备训练数据与测试数据具有相同类别空间,同时难以有效识别新故障的问题,提出了一种基于选择性加权适配网络的多域新故障识别方法。所提方法利用一维卷积神经网络提取源域与目标域深度判别特征,并集成领域判别器与多分类器结构,构建源域与目标域权重函数,自适应度量源域与目标域类别的相似程度;从而利用对抗学习策略来有效减少源域与目标域共享类数据的分布差异;最后利用高斯分布拟合方法自动判别权重阈值,实现对目标域已知故障和新故障的有效诊断。在齿轮箱变工况迁移诊断任务上对所提方法进行分析与应用验证,并与现有的其它方法进行比较,所提方法在所有任务上的调和平均值(E-score)达到0.8以上,验证了所提方法的有效性与优越性。

关键词:故障诊断;对抗学习;新故障;深度学习;迁移学习

中图分类号: TH17 **文献标识码:** A **国家标准学科分类代码:** 510.40

Selective weighted adaptive network for multi-domain emerging fault identification

Chen Zhuyun^{1,2,3}, Lin Huibin¹, Xia Jingyan¹, Jin Gang¹, Li Weihua^{1,2,4}

(1. School of Mechanical and Automotive Engineering, South China University of Technology, Guangzhou 511442, China; 2. Pazhou Lab, Guangzhou 510005, China; 3. Beijing Key Laboratory of Measurement Control of Mechanical and Electrical System Technology, Beijing Information Science Technology University, Beijing 100192, China; 4. Shien-Ming Wu School of Intelligent Engineering, South China University of Technology, Guangzhou 511442, China)

Abstract: The existing deep transfer learning-based diagnosis methods usually require that the same fault class space is shared by training and test data, which fail to effectively identify new faults. Thus, a multi-domain emerging fault identification method based on selective weighted adaptive network is proposed. Firstly, a one-dimensional convolutional neural network is adopted to extract depth discriminative features across domains. Then, a domain discriminator and multi-classifier structures are integrated to construct weight functions of source and target domains to adaptively measure the similarity across different categories. The adversarial learning strategy is utilized to effectively reduce the distribution differences of shared classes across domains. Finally, the Gaussian distribution-based fitting method is adopted to automatically discriminate weight thresholds to realize effective fault diagnosis of known faults and emerging faults in the target domain. Experiments are conducted on a gearbox transmission test rig, where the transfer diagnosis tasks under variable operation conditions are designed. The proposed method obtains 0.8 E-score in various tasks. The effectiveness and the superiority of the proposed method are fully validated in comparison with other existing methods.

Keywords: fault diagnosis; adversarial learning; emerging fault; deep learning; transfer learning

收稿日期: 2022-05-27 Received Date: 2022-05-27

*基金项目: 国家自然科学基金(52205101, 51875208, 52275111)、广东省基础与应用基础研究基金区域联合基金青年基金(2021A1515110708)、广州市基础研究计划基础与应用基础研究项目(202201010615)、北京信息科技大学重点科研机构(KF20212223204)项目资助

0 引言

随着科学技术的发展和现代机械制造技术水平的提升,机械设备的发也来越趋向于大型化、复杂化、精密化和智能化。旋转机械作为机械设备的组成部分,广泛应用在车辆、船舶、航空航天等领域。由于长期工作在交变负荷下,旋转机械的核心零部件,如齿轮、轴承和转轴等在实际运行过程中不可避免的会出现损伤,进而引发设备故障,带来潜在的经济损失和人员伤亡^[1]。因此,开展机械设备的智能故障诊断研究,对提高设备的运行效率、降低故障的影响,保证安全有序生产至关重要。

基于机器学习,尤其是基于深度学习的智能故障诊断方法近年来获得了极大的关注。许多深度学习模型,如深度自编码网络、卷积神经网络等用于机械设备智能故障诊断,并取得了良好的效果^[2-3]。然而其性能的发挥通常需依托大量训练数据,同时要求训练数据和测试数据满足独立同分布,因而并不能很好的适用于训练数据量少,场景变化和任务变化等领域^[4]。

迁移学习,作为一种新的机器学习范式,可从一个或者多个相关场景中提取知识以帮助提升目标域场景的学习性能,其不仅减少了数据量的要求,同时放松了独立同分布的假设,为弥补上述不足提供了一种新的解决思路^[5-6]。Li等^[7]利用对抗训练机制对不同传感器数据分布差异进行域适配,解决了机械设备不同位置数据的故障诊断问题。Xia等^[8]提出了一种基于惯性匹配的多源域适配网络,实现了变载荷下的轴承故障诊断。Zhang等^[9]提出了一种自适应卷积神经网络,用于不同工况下的滚动轴承故障诊断。Han等^[10]提出一种深度对抗卷积神经网络,在变工况诊断任务下,获得了比传统方法更优的诊断性能。Chen等^[11]开发了一种域对抗迁移网络,通过构建两个独立的编码网络进行对抗学习,实现了跨工况下的齿轮和轴承故障诊断。Guo等^[12]开发了一种新颖的深度迁移诊断方法,实现了跨机械设备下的滚动轴承迁移诊断。尽管上述方法在解决域迁移诊断问题上获得很好的分类性能,但其通常要求训练数据和测试数据具有相同的标签空间,导致已训练的诊断模型只能识别源域空间中已有的故障类别,难以用于诊断测试数据中存在的新故障。

为了实现变工况下的新故障迁移诊断任务,Li等^[13]提出了一种深度对抗迁移学习网络,通过构建一个多标签分类器进行对抗训练,有效实现目标域任务中新故障的诊断。此外,Li等^[14]进一步开发了两阶段迁移对抗网络,通过引入无监督卷积自编码和轮廓系数,进一步对多个未知故障进行了聚类识别。Zhao等^[15]提出了一个结

合域对抗网络和多标签分类器的对抗网络,用于实现未知域的故障诊断。Zhang等^[16]提出了一种基于类别层加权的深度神经网络,在滚动轴承新故障诊断任务中获得了优异的性能。然而,上述所提方法尽管可以实现新故障的诊断,但所提方法要求源域类别数必须要小于目标域类别数,即源域故障类别是目标域故障类别的一个子集,难以适用于多域下的新故障诊断。

在实际工业环境中,目标域不仅可能存在源域未知的故障类型,同时源域也可能存在目标域中没有的故障这是一个更符合实际同时也是更具挑战性的多域新故障诊断问题。为了解决多域下的新故障诊断,Yu等^[17]提出了一种深度学习的开放集诊断模型,首先利用深度网络进行特征提取,并进一步利用极限值理论模型实现多域下的新故障诊断。但所提方法需分别独立执行特征提取和故障识别两步分类任务,无法实现端到端的故障诊断,因而削弱了其模型的学习和诊断性能。Zhang等^[18]构建了一种混合加权的端到端通用域故障诊断网络,可同时诊断包含多域新故障等不同场景下诊断任务,然而所提方法聚焦于通用域任务,没有充分利用故障类别的监督信息以及分类器的不确定信息,同时依赖于手动阈值进行新故障判别,因而在多域新故障诊断任务上导致性能下降。

因此,基于现有诊断方法的不足,提出一种基于选择性加权适配网络(selective weighted adaptive network, S-WAN)的多域新故障识别方法。所提方法以机械设备振动信号作为输入,通过充分利用深度学习的特征提取能力、领域判别器的信息相似性评估与多分类器的信息不确定性评估能力,结合对抗学习技术对网络进行端到端训练,最后利用高斯拟合分布自适应选取阈值,实现多变工况下的已知故障和新故障识别,为机械智能故障诊断提供一种新的方法与思路。

本文的主要创新点如下:

1) 提出了一种基于选择性加权适配网络的机械设备多域新故障识别方法。所提方法聚焦于源域与目标域不仅包含共享类,同时还各包含私有类的多域新故障诊断问题,通过利用加权学习和对抗学习策略,直接对原始输入数据进行端到端训练,有效实现已知故障和新故障的智能诊断。

2) 综合考虑类别信息的相似性和分类器的不确定性,所提方法充分集成领域判别器的相似性评估能力与多分类器结构的不确定性评估能力,构建源域与目标域权重函数,可实现对不同域中共享类与私有类的有效度量,促进共享类诊断知识的有效迁移。

3) 相比于依赖人工经验手动构建阈值,所提方法通过综合分析目标域类别中的共享类(已知类)和私有类(新故障)权重度量差异,利用双高斯函数来拟合权重分

布,自适应确定共享类故障和新故障的判断阈值,从而提高诊断的智能型和可靠性。

1 研究背景

迁移学习的域适配方法致力于在领域之间迁移共享知识。传统域适配方法通常需在源域与目标域类别一致的情况下进行迁移学习,即 $C_s = C_t$;当源域与目标域类别空间不一致时,即目标域中含有“新故障类别”时,此时目标域类别大于源域类别,即 $C_s \in C_t$,网络需进一步识别目标域中的新故障;然而,在实际应用中,由于机械传动系统故障具有突发性,采集到的故障数据种类较少且可能出现新故障类别,导致目标域所含类别无法提前预知,而源域的带标签数据通常是在已知故障类别情况下进行采集的,这将导致目标域类别空间与源域不一致。因而,容易出现源域与目标域均包含私有类,且两域的公共标签集 C 未知的情况,即 $C_s \neq C_t, C_s \cap C_t = C$,这是一种更符合实际的多域新故障识别问题,给现有诊断方法带来了很大挑战。

对于源域及目标域中都包含共享类和私有类的问题,给定的带标签的源域 $D_s = \{(x_i^s, y_i^s)\}_{i=1}^{n_s}$, 概率分布为 P , 类别标签集为 C_s ;对无标签的目标域 $D_t = \{x_j^t\}_{j=1}^{n_t}$, 概率分布为 Q , 类别标签集为 C_t 。 $C = C_s \cap C_t$ 表示两域共享的公共标签集, $\bar{C}_t = C_t \setminus C$ 表示目标域的私有类别, $\bar{C}_s = C_s \setminus C$ 表示源域的私有类别。 P_{C_s} 表示源域样本 $x_s \in C_s$ 的数据分布, P_C 表示源域样本 $x_s \in C$ 的数据分布; Q_{C_t} 表示目标域样本 $x_t \in C_t$ 的数据分布, Q_C 表示目标域样本 $x_t \in C$ 的数据分布。

在多域新故障诊断任务中,由于源域与目标域采集自不同工况,其概率分布 $P(x_s) \neq Q(x_t)$, 并且两域类别空间不对齐 $C_s \neq C_t$, 公共标签集 C 未知。由于数据分布及类别标签都不相同,因而直接采用传统的域适配技术,由于源域和目标域私有类别的干扰,会导致类别的误匹配,造成负迁移问题。因此,需综合考虑和判别源域的私有类与目标域的私有类别带来的影响,减少两者的干扰,以实现最终新故障的识别。

2 所提的选择性加权适配网络

本文所提选择性加权适配网络的多域新故障识别方法,需要迁移源域任务诊断知识(故障标签信息已知)到目标域故障诊断任务(故障标签信息未知),其核心难点在于,在执行域适配过程中,如何有效实现源域与目标域共享类的对齐(特征对齐)以及避免私有类的误匹配(负迁移)。本质上,如果能够直接知道源域和目标域中的共

享类和私有类,则可以直接利用传统的域适配方法执行故障诊断。然而,由于目标域诊断任务的故障类别是未知的,因而不可能提前知道源域和目标域中共享类与私有类。因此,对于源域和目标域都包含了私有类别的迁移诊断任务,提出一种选择性加权适配网络用于齿轮箱智能故障诊断。所提方法的核心在于通过构建源域与目标域的权重函数,来量化每个样本的可迁移性,并将每个样本数据的域相似度和预测不确定性整合到了网络的加权机制中,从而实现共享类样本的适配与新故障的识别。其诊断流程框架及加权机制如图1所示。

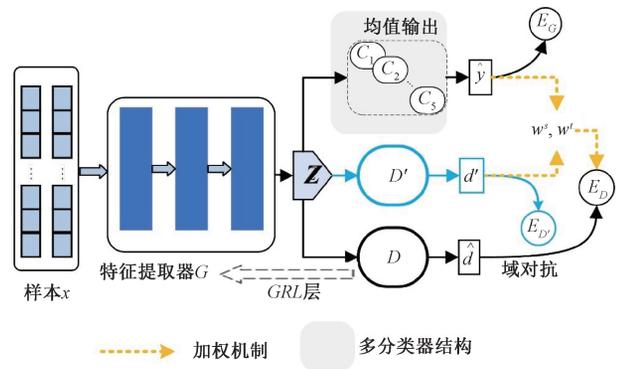


图1 所提选择性加权适配网络故障诊断方法

Fig. 1 The proposed diagnosis scheme of the selective weighted adaptive network

2.1 加权适配网络构建

在所提的网络结构中,首先构建一个特征提取器 G , 将样本 x 输入到特征提取器 G 学习数据的高维特征表示 $Z = G(x)$ 。由于卷积神经网络具有极好的特征学习能力,因此构建了一种一维的卷积神经网络作为特征提取器 G , 具体结构如表1所示。网络在第1层采用大的一维卷积核结构,进行特征提取,以增强其对输入抗噪性能,滤波器大小设置为64,滤波器数目为20,并利用最大池化结构,减少网络参数运算量;接下来采用小的卷积核结构,提取高维特征,以提高其特征提取能力。

然后将提取的高维表示特征 Z 分别输入到标签分类器 C 、非对抗领域判别器 D' 和领域对抗判别器 D 中。其中标签分类器 C 由全连接层(full connection, FC)和softmax分类器组成,用于对提取的高维特征进行判别,以获得源域标签集 C_s 上 x 的概率分布。

考虑分类器对不同样本预测的不确定程度,同时增强分类器对目标域预测样本的信心,利用集成学习可以组合多个弱监督模型以得到一个更好更全面的强监督模型的思想^[19],进一步构建多个相同结构的标签分类器 $\{C_k\}_{k=1}^K$, 对不同样本的概率进行预测, K 为分类器的数目;由于不同分类器具有不同的初始参数和优化结构,

表 1 特征提取器结构

Table 1 Structure of feature extractor

层类型	激活函数	数目	核尺寸×步长	输出尺寸
输入层	-	-	-	(2 048, 1)
Conv_1	ReLU	20	64×16	(20, 125, 1)
BN_1	-	-	-	(20, 125, 1)
MaxPool_1	-	20	2×2	(20, 62, 1)
Conv_2	ReLU	20	3×1	(20, 62, 1)
BN_2	-	-	-	(20, 62, 1)
MaxPool_2	-	20	2×2	(20, 31, 1)
Conv_3	ReLU	40	3×1	(40, 31, 1)
BN_3	-	-	-	(40, 31, 1)
MaxPool_3	-	40	2×2	(40, 31, 1)
Conv_4	ReLU	40	3×1	(40, 15, 1)
BN_4	-	-	-	(40, 15, 1)
MaxPool_4	-	40	2×2	(40, 7, 1)

因此多分类器输出均值的预测鲁棒性更强,在实验中,考虑计算时间和精度,设置 $K=5$,进行集成学习分类,并将最终预测结果,通过集成学习器投票法,将多个集成分类器预测结果的平均值作为最终的预测结果: $\hat{y} = \sum_{k=1}^K C_k$ 。

由于源域数据为带标签样本,对特征提取器 G 进行优化训练,其本质上是一个有监督训练,损失函数 E_G 如式(1)所示。

$$E_G = \sum_{k=1}^K E_{(x,y)-p} L(y, C_k(G(x))) \quad (1)$$

其中, L 为标准的交叉熵损失函数, y 为源域样本标签, $C_k(G(x))$ 为第 k 个分类器 C_k 输出的标签类别。

进一步将提取的特征输入到非对抗领域判别器 D' , 通过输入特征分布 Z 获得域相似度 $\hat{d}' = D'(Z)$, 目的是要使得源域的样本预测为 1, 目标域的样本预测为 0, 量化每个样本的域相似度。对于源域样本, \hat{d}' 越小, 说明此样本接近目标域; 对于目标域样本, \hat{d}' 越大, 说明此样本与源域越相似。损失函数 $E_{D'}$ 如式(2)所示。

$$E_{D'} = -E_{x-p} \log D'(G(x)) - E_{x-q} \log(1 - D'(G(x))) \quad (2)$$

领域对抗判别器 D 的目标是区分源域和目标域数据是属于公共标签集 C 还是私有类, 而后通过对此部分数据样本加权计算, 使得在公共标签集 C 中源域与目标域样本能最大程度匹配。其损失函数 E_D 如式(3)所示。

$$E_D = -E_{x-p}(x) w^s \log D(G(x)) - E_{x-q} w^t(x) \log(1 - D(G(x))) \quad (3)$$

其中, $w^s(x)$ 与 $w^t(x)$ 分别为源域和目标域权重度量函数, 表示源域与目标域样本属于公共标签集 C 的概

率。领域对抗判别器 D 通过 $w^s(x)$, $w^t(x)$ 的权重值来度量源域和目标域数据属于公共标签集 C 还是属于其各自私有类数据的概率。

在 S-WAN 中, 特征提取器 G 提取源域与目标域数据的特征分布, 用于混淆领域对抗判别器 D , 从而在公共标签集 C 中获得两域的领域不变特征, 促进源域与目标域的特征分布匹配。使得利用源域样本数据训练所得的标签分类器 C_k 能较好的应用于目标域数据判别。其本质上是一个极大-极小的博弈过程, 如下式所示。

$$\max_D \min_{G, C_F} (E_G - E_D) \quad (4)$$

$$\min_D E_{D'} \quad (5)$$

所提方法通过在特征提取器 G 和领域对抗判别器 D 之间添加梯度反转层 (GRL), 以实现诊断网络所有参数的端对端训练与参数更新。

2.2 权重函数构建

针对源域权重度量函数 $w^s(x)$ 和目标域权重度量函数 $w^t(x)$, 将对标签分类器 $\{C_k\}_{k=1}^K$ 的预测不确定性 $H(\hat{y}) = -\hat{y} \log(\hat{y})$, 和非对抗领域判别器 D' 输出的域相似度 \hat{d}' 进行计算。

1) 预测的不确定性

其核心思想在于引入熵 (entropy) 来度量网络的标签分类器 $C_k(G(x))$ 对每个样本预测值 ($\hat{y} = \sum_{k=1}^K C_k$) 的不确定性程度。熵值 $H(\hat{y})$ 越大, 则网络预测输出的不确定性程度越大。如果样本来源于源域, 且其熵值 $H(\hat{y})$ 比较大, 则可以预期, 网络的标签分类器 $\{C_k\}_{k=1}^K$ 更容易将其预测为源域和目标域的共享类, 而不是源域私有类。如果样本来源于目标域, 且其熵值 $H(\hat{y})$ 比较大, 则可以预期, 网络的标签分类器 $\{C_k\}_{k=1}^K$ 更容易将其预测为目标域的私有类, 而不是源域和目标域的共享类, 反之亦然。

具体的, 对于源域标签集 C_s , 目标域共享类标签集 C 及私有标签集 \bar{C}_i 预测的不确定性为 $E_{x-Q_{\bar{C}_i}} H(\hat{y}) > E_{x-Q_C} H(\hat{y}) > E_{x-P_C} H(\hat{y})$ 。对于源域样本, 由于源域与目标域包含公共标签集 C , 若 $x_s, x_t \in C$, ($x_s \sim P_C, x_t \sim Q_C$), 数据分布 P_C, Q_C 在一定程度上具有相似性, 在训练过程中, 网络对目标域样本预测的不确定性较高, 则其熵值较大。由于源域私有标签集数据与目标域数据没有交集, 即 $\bar{C}_s \cap C_t = 0$, 因此 $x_s \sim P_{\bar{C}_s}$ 不受目标域数据的影响, 其可靠性较高。故可得 $E_{x-P_C} H(\hat{y}) > E_{x-P_{\bar{C}_s}} H(\hat{y})$ 。

对于源域与目标域标签预测的不确定性, 其熵值大小如式(6)所示。

$$E_{x \sim Q_{\bar{C}_i}} H(\hat{y}) > E_{x \sim Q_C} H(\hat{y}) > E_{x \sim P_C} H(\hat{y}) > E_{x \sim P_{\bar{C}_s}} H(\hat{y}) \quad (6)$$

2) 域相似度

其核心思想在于构建一个非对抗领域判别器 D' , 通过监督训练一个二分类网络来对输入的源域与目标域类别进行预测(训练时源域数据标签为 1, 目标域数据标签为 0)。可以预期, 当网络训练好以后, 针对源域数据输出值 \hat{d}' , 其值越大, 则该输入数据越有可能属于源域私有类, 而不是源域与目标域的共享类。针对目标域数据输出值 \hat{d}' , 其值越大, 则该输入数据越有可能属于源域与目标域的共享类, 而不是目标域的私有类, 反之亦然。

具体的, 根据式(6)可知, 对于源域标签集 C_s , 目标域共享类标签集 C 及私有标签集的域相似度大小为 $E_{x-P_{C_s}}(\hat{d}') > E_{x-Q_C}(\hat{d}') > E_{x-Q_{\bar{C}_i}}(\hat{d}')$ 。由于 $x_s, x_t \in C$, ($x_s \sim P_c, x_t \sim Q_c$), 两者共用相同的类别标签集, $x_s \sim P_c$ 相比于 $x_s \sim P_{\bar{C}_s}$ 在一定程度上更接近于 $x_t \sim Q_c$, 而由于 \bar{C}_s 为源域私有类, $x_s \sim P_{\bar{C}_s}$ 不受目标域数据的影响。故可得 $E_{x-P_{\bar{C}_s}}(\hat{d}') > E_{x-P_c}(\hat{d}')$ 。

对于源域与目标域样本的域相似度, 其判别关系如下式:

$$E_{x-P_{\bar{C}_s}}(\hat{d}') > E_{x-P_c}(\hat{d}') > E_{x-Q_C}(\hat{d}') > E_{x-Q_{\bar{C}_i}}(\hat{d}') \quad (7)$$

3) 加权计算策略

其核心思想在于, 希望能构建一个源域权重函数 $w^s(x)$, 来度量源域样本 x 属于共享类 C 与源域私有类 \bar{C}_s 的概率; 同时构建一个目标域权重函数 $w^t(x)$, 来度量目标域样本 x 属于共享类 C 与目标域私有类 \bar{C}_t 的概率。因而, 在对抗学习过程中, 通过利用多域加权学习策略, 使得来自源域和目标域中的共享类样本可获得更大权重参与网络训练, 而对私有类样本可获得小的权重参与训练, 减少其对域适配的影响, 从而有效提升网络的学习质量, 减少误匹配的影响。

具体的, 根据预测的不确定性及域相似度的分析, 借鉴通用域学习的思路^[20], 源域样本的加权函数计算与目标域的加权函数计算为:

$$w^s(x) = \frac{H(\hat{y})}{\log |C_s|} - \hat{d}'(x) \quad (8)$$

$$w^t(x) = \hat{d}'(x) - \frac{H(\hat{y})}{\log |C_s|} \quad (9)$$

源域与目标域权重函数通过最大值归一化 ($\log |C_s|$) 将数值转换至 $[0, 1]$ 的范围中。对于 $w^s(x)$, 由于 $x_s, x_t \in C$, ($x_s \sim P_c, x_t \sim Q_c$), 两者标签类别空间一致, 在一定程度上存在相似性。对于源域样本, 若 $x_s \in C$, 由于受到 $x_t \sim Q_c$ 的影响, 其共享类权重比较大, 即 $w_c^s(x) > w_{\bar{c}_s}^s(x)$, 在进行加权训练时, 源域中属于公共标签集 C 的样本将被赋予较大的权重值, 源域私有类别样

本将被赋予较小值, 从而减小源域私有类别的影响。对于目标域样本, 若 $x_t \in C$, 由于受到 $x_s \sim Q_c$ 的影响, 其共享类权重较大, 即 $w_c^t(x) > w_{\bar{c}_t}^t(x)$, 在进行加权训练时, 将目标域中属于公共标签集 C 的样本赋予较大的权重值, 将目标域私有类别样本赋予较小值, 从而减小目标域私有类别的影响。最终, 确定的权重阈值, 施加到域对抗判别器 D 的损失函数中, 并引入梯度反转层 (GRL)^[21], 通过 Adam 算法更新网络中每个模块参数。

2.3 诊断测试

在网络训练测试过程中, 对每个输入样本都进行了权重计算, 最终通过阈值进行类别判定。因此需要设定合理的阈值, 才能使得网络在测试过程中, 能够较好的区分目标域中“新故障类别”, 从而使得网络同时识别共享类与新类故障。相比于依赖人工经验构建阈值, 本文考虑权重函数分为私有类与共享类, 而私有类的权重值比共享类权重值更低, 因此, 利用已训练网络获取目标域训练数据的权重值, 其分别对目标域共享类和私有类的输出可以假设服从两个不同的分布, 对私有类和共享类权重采用双高斯分布进行拟合, 找寻两峰高斯分布的均值 μ_1, μ_2 , 通过试验, 将求得均值 μ_1, μ_2 中较小值设定为区分私有类与共享类的权重阈值, 即 $w^0 = \min(\mu_1, \mu_2)$ 。在测试阶段, 当输入样本权重 $w^i < w^0$, 此时认定该样本为新故障类别, 反之则为已知类别, 网络通过多分类器对该样本进行进一步预测, 通过投票原则, 获得样本的具体故障类别, 有效实现对已知类故障的识别以及新故障的检测。

3 故障诊断试验

3.1 数据采集与描述

为了对所提的网络模型进行评估, 采集不同工况下的齿轮与轴承故障数据进行验证。齿轮和轴承故障数据采集自汽车传动试验台的五档变速箱, 被试齿轮箱为 SG135-2 汽车五档变速器, 试验平台如图 2 所示。振动响应信号采集自安装于壳体输出轴承座上垂直方向的加速度传感器, 设置采样频率为 24 kHz, 输入转速分别为 1 000 和 1 250 rpm, 输出轴负载情况分别为: 无负载 (0 N·m)、50 N·m。齿轮箱故障使用线切割机器加工, 共设置加工了多种不同的健康状况。齿轮故障分别设置为正常、五档完全断齿、二档剥落, 同时加工了 0.2 mm 的轴承内圈故障, 可同时模拟齿轮和轴承的单一故障类型以及齿轮和轴承的复合故障。

3.2 数据采集与描述

为验证所提 S-WAN 方法的有效性和优越性, 将其与其他几种域适配迁移学习方法进行比较。

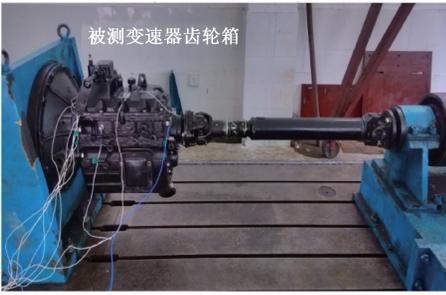


图 2 传动试验台

Fig. 2 The transmission platform

1) MLDANN^[7]。网络采用了最大均值差异函数来度量源域与目标域的数据分布差异,并进一步对深度网络的不同层特征执行联合域适配,在不同设备的轴承迁移诊断任务上获得了优异的性能。

2) DACNN^[10]。网络利用一维卷积神经网络进行特征提取,并构建分类模块与对抗学习模块,通过引入梯度反转层(GRL)和对抗学习思路,减少源域与目标域数据的分布差异,在变工况下的机械故障诊断任务上取得了良好的诊断结果。

3) CMU^[22]。CMU 是由 Fu 等^[22]提出的一种通用域适配方法,通用域适配(universal domain adaptation, UniDA)可在领域之间传递知识,对两域的标签集无任何约束;在 UniDA 中,源标签集和目标域标签集可含有私有类,也可不含私有类别,因而 CMU 可以适用于多域新故障下的迁移诊断任务。

4) UAN^[19]。通用域适配网络 UAN 的核心思想在于利用领域判别器相似性和分类器度量策略来构建源域与目标域的权重函数,从而通过加权对抗思想对多种迁移任务提供通用解决方案,相关策略已经用于通用域(包含多域新故障识别)下的迁移诊断任务^[18]。该网络与 S-WAN 的核心区别在于所提 S-WAN 方法进一步构建了多分类器结构与阈值自适应学习模块。

5) S-WAN(wo)是在所提 S-WAN 方法基础上,去掉了多分类器学习模块,只采用单一分类器,并进一步执行目标域下的域适配任务,因而可以有效验证所提选择性加权学习策略的有效性。

为了对上述所提方法进行公平比较,所有网络都利用构建的源域数据与目标域数据参与对抗训练。所采用的特征提取器都采用与 S-WAN 一样结构的一维卷积神经网络,所涉及到其他模块采用的全连接结构等与 S-WAN 保持一致。并在优化过程中,采用相同的 Adam 优化函数和迭代次数,从而使得网络都能充分训练。

3.3 评价指标

1) 领域标签集相似性指标

为定义源域与目标域的相似程度,即为了表示两域

的共享类别的占比,通过杰卡德相似系数(Jaccard similarity coefficient)来表示领域类别空间的重合度,如下式所示。

$$\xi_c = \frac{|C_s \cap C_t|}{|C_s \cup C_t|} \quad (10)$$

其中, ξ_c 表示领域标签集的相似性,若 $\xi_c = 1$,表示两域类别空间一致,即为闭集域适配问题;若 $\xi_c = 0$,表示两域类别空间无交集,则此时两域之间无法进行迁移学习。对于相似性指标, ξ_c 越小,表示两域的公共标签集越小,两域共享的知识越小,则两域的域适配难度越大。

2) 性能评估指标

前期研究更多通过网络预测输出的类别标签与真实标签的比值计算分类精度。然而,针对目标域中含有私有类(新故障)的情形,共享类与新故障对网络的诊断精度具有相同的贡献。但常规的评估方法,在共享类数据量较大、新故障数据量较小的情况下,新故障分类正确与否对整体的诊断精度影响小。为了构建更加合理的网络性能评估指标,除了分类准确率以外,根据 F1-score 指标如下式,构建网络性能评估指标。

$$F_1 = 2 \cdot \frac{\text{精确率} \cdot \text{召回率}}{\text{精确率} + \text{召回率}} \quad (11)$$

F1-score 本质上是分类问题的衡量指标,它表示精确率和召回率的调和平均值,同时考虑了分类模型的精确率和召回率,其最大值为 1,最小值为 0。因此,根据 F1-score 的思想,构建网络性能评估指标 E-score,它表示共享类与新故障诊断准确度的调和平均值。

$$E\text{-score} = 2 \cdot \frac{A_c \cdot A_{\bar{c}_i}}{A_c + A_{\bar{c}_i}} \quad (12)$$

其中, A_c 表示目标域中共享类的分类准确度; $A_{\bar{c}_i}$ 表示目标域中新故障类别的分类准确度,其最大值为 1,最小值为 0。若 A_c 和 $A_{\bar{c}_i}$ 中有一个值为 0,则评估指标 E-score 值也为 0;只有当两者的分类精度都较高时,评估指标 E-score 值才会高。

3.4 跨转速故障诊断试验

1) 迁移诊断任务介绍

齿轮箱故障数据集如表 2 所示,共包含了齿轮与轴承 6 种健康状况。根据获取的齿轮箱数据中所含故障类别,对原始振动信号进行分段,每段包含样本点数 2 048 点,每个类别各获取 500 个信号样本,并进一步执行快速傅里叶变换,获取其频谱信息,作为网络输入。

通过上述获取的齿轮箱故障数据集,考虑不同的运行工况和故障类型,共设计了 3 个跨转速下的迁移诊断任务,用于对所提 S-WAN 故障诊断方法的有效验证,具体研究方案如表 3 所示。

表2 齿轮箱数据集描述

Table 2 Dataset description of Gearbox

类别标签	故障类型	样本量	转速/rpm
1	齿轮正常_内圈正常	500 & 500	1 000 & 1 250
2	齿轮正常_内圈故障	500 & 500	
3	五档断齿_内圈故障	500 & 500	
4	五档断齿_内圈正常	500 & 500	
5	二档剥落_内圈故障	500 & 500	
6	二档剥落_内圈正常	500 & 500	

表3 齿轮箱跨转速迁移任务描述

Table 3 Cross-speed tasks of gearbox dataset

试验	迁移任务	数据集	转速/负载	故障类别	ξ
-	-	S_1	1 000/0	1/2/3/4	-
试验1	$S_1 \rightarrow T_1$	T_1	1 250/0	1/3/4/5	0.6
试验2	$S_1 \rightarrow T_2$	T_2	1 250/0	1/3/4/6	0.6
实验3	$S_1 \rightarrow T_3$	T_3	1 250/0	1/3/4/5/6	0.5

以试验1中的迁移任务 S_1-T_1 为例, S_1 为源域数据集, 包含4种齿轮健康状态(对应类别标签为1、2、3和4), 其振动信号采集自转速1 000 rpm, 负载为0 N·m的工况; 而 T_1 为目标域数据集, 包含4种齿轮状态(即类别标签为1、3、4和5)。源域与目标域的公共标签集 $C = (1, 3, 4)$, 源域私有类为 $\bar{C}_s = (2)$, 目标域的私有类为 $\bar{C}_t = (5)$, 两域的相似性指数为 $\xi = 0.6$ 。在迁移阶段, 随机选取源域与目标域的1/2数据参与训练, 剩下的目标域数据用于进行网络测试。

2) 迁移诊断结果

诊断结果如表4所示。在3个迁移试验中, 综合考虑总体分类准确率及 E-score 值, 所提方法 S-WAN 取得了最好的结果, 3个迁移任务的总体准确率均高于90%, 同时其 E-score 值也达到了最高, 平均 E-score 值达到了85%。再进一步观察对比方法, 前两种对比方法 MLDAN 与 DACNN 由于只能诊断已知故障, 而无法识别新故障, 因而其诊断准确率总体偏低, 尤其是在相似度指标更低的 $S_1 \rightarrow T_3$, 结果最差。而另外3种对比方法 CMU、WAN 与 S-WAN(wo), 相比传统的域适配迁移诊断方法获得了明显更高的精度, 但总体上仍然低于所提 S-WAN 方法。

此外, 为了更好的评估4种新故障检测方法(CMU、UAN、S-WAN(wo)与S-WAN), 进一步计算出每个任务的共享类识别准确率与私有类识别准确率, 如图3所示。可以看到, 所提方法 S-WAN 具有较高的总体分类准确率和 E-score 的同时, 对于目标域共享类及私有类的分类

表4 跨转速下诊断结果

Table 4 Diagnosis result under cross-speed tasks

方法	S_1-T_1		S_1-T_2		S_1-T_3	
	准确率	E-score	准确率	E-score	准确率	E-score
MLDAN	74.48	-	74.82	-	59.34	-
DACNN	74.06	-	73.80	-	59.72	-
CMU	82.31	0.817	77.65	0.784	80.91	0.671
UAN	90.15	0.888	89.93	0.803	87.71	0.769
S-WAN(wo)	86.78	0.792	84.49	0.684	89.36	0.704
S-WAN	94.77	0.893	91.37	0.807	93.32	0.848

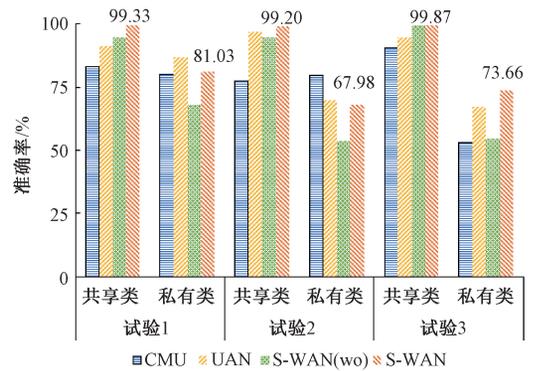


图3 跨转速下已知类与新故障类别诊断精度

Fig. 3 Diagnosis accuracy of known and unknown classes under cross-speed tasks

精度也最高。UAN 方法, 相比 S-WAN 呈现了一定的竞争力, 但在试验3任务上诊断能力较弱。

S-WAN 由于未将源域的私有类纳入考虑, 虽然其总体诊断精度高于 CMU, 但其对于目标域私有类判别效果较差, 这可能是由于在进行分类过程中, 源域私有类对领域适配产生影响, 导致两域数据分布匹配效果较差, 使得判定目标域私有类的权值时产生混淆, 因而 CMU 的总体诊断精度高于 WAN。对于跨转速的迁移试验, 所提 S-WAN 网络在所有任务上都具有最佳的诊断性能, 且其优势较为明显, 表明 S-WAN 在进行两域类别交叉的迁移任务时, 其采用对抗学习、权值判别, 可以有效的减小源域中不相关类别的影响, 判别目标域中的私有类别, 从而提高了各类别的诊断精度。

3) 跨转速与跨载荷下故障诊断试验

考虑不同载荷与不同转速下的诊断问题, 根据齿轮箱数据集设计了3个跨转速跨载荷下的迁移任务, 即源域工况为转速1 250 rpm, 载荷50 N·m, 目标域工况为转速1 000 rpm, 载荷0 N·m, 源域与目标域数据来源于不同的转速和载荷, 构建的迁移任务如表5所示。

表 5 齿轮箱跨转速与跨载荷迁移任务描述

Table 5 Cross-speed and load tasks of gearbox dataset

试验	迁移任务	数据集	转速/负载	故障类别	ξ
-	-	S_1	1 250/0	1/2/3/4	-
试验 1	$S_1 \rightarrow T_1$	T_1	1 000/50	1/3/4/5	0.6
试验 2	$S_1 \rightarrow T_2$	T_2	1 000/50	1/3/4/6	0.6
试验 3	$S_1 \rightarrow T_3$	T_3	1 000/50	1/3/4/5/6	0.5

由于所构建的源域与目标域数据集采自不同转速、不同负载的工况下,两域的数据分布差异性大于单一跨转速诊断任务。因此,通过不同工况下两域的故障诊断试验,能够更好地分析网络的有效性及泛化性。针对构建的 3 个不同工况(源域与目标域转速不同,负载不同)的迁移任务,将提出的 S-WAN 与其他 5 种域适配诊断方法进行对比,其具体结果如表 6 所示。

表 6 跨转速与跨载荷下诊断结果

Table 6 Diagnosis result under cross-speed and load tasks

方法	试验 1		试验 2		试验 3	
	准确率	E-score	准确率	E-score	准确率	E-score
MLDAN	73.66	-	73.50	-	56.17	-
DACNN	75.18	-	73.20	-	56.89	-
CMU	74.24	0.776	87.17	0.871	85.43	0.793
UAN	91.91	0.861	90.55	0.928	82.95	0.756
S-WAN(wo)	76.24	0.757	78.53	0.815	76.94	0.718
S-WAN	90.22	0.908	90.97	0.926	88.91	0.799

在 3 个迁移任务中,与跨转速试验类似,由于缺少对新故障的识别能力,传统 MLDAN 与 DACNN 方法,总体识别精度最低。而另外 3 种比较方法 CMU、UAN 与 S-WAN(wo)呈现了明显更好的识别精度,尤其是在试验 3 上,获得了与 S-WAN 竞争性的 E-score 识别精度。但综合 3 个迁移诊断任务,所提方法 S-WAN 的表现最好,试验 1、试验 2 的总体诊断精度及 E-score 值均达到了 90% 以上;对于试验 3,虽然 S-WAN 的总体准确率为 88.91%, E-score 值为 0.799,但与 CMU、UAN 与 S-WAN(wo)相比,后 3 者的评估指标 E-score 的均值分别为 0.793、0.756 和 0.718 仍然具有一定的优势。

此外,4 种新故障检测方法的共享类与私有类识别精度如图 4 所示。可发现,尽管 CMU 在试验 2 共享类与试验 3 私有类任务,呈现了与 S-WAN 竞争性的结果,但综合评估,所提方法依然呈现了最优的性能。

因此,通过不同工况下的诊断试验可知,所提出的 S-WAN 在 3 个迁移任务中都具有最佳的诊断性能,并且

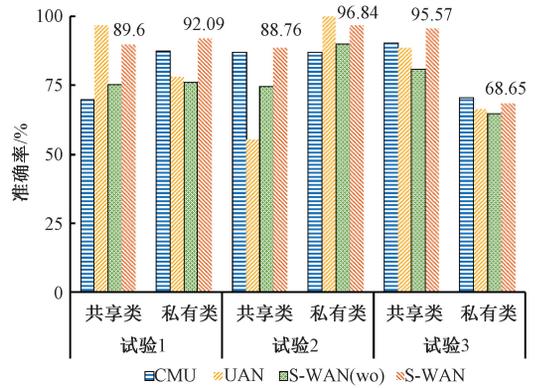


图 4 跨转速与跨载荷下已知类与新故障类别诊断精度
Fig. 4 Diagnosis accuracy of known and unknown classes under cross-speed and load tasks

在试验 1、试验 2 中其优势更加明显,表明所提的选择性加权迁移诊断方案的可行性,同时相比其他方法在分类能力和泛化性能的优越性。

4 结 论

针对源域和目标域数据分布不一致,且两域类别空间不对齐下的迁移诊断问题,提出了选择性加权适配网络用于多域新故障诊断。所提方法构建了源域和目标域样本的加权机制,有效减小了两域私有类对领域适配的影响。并利用高斯分布拟合方法自动判别权重的阈值,克服了人工阈值选取的局限性。通过在齿轮箱数据下的跨转速以及跨转速与跨载荷下的迁移诊断试验对所提方法进行验证,并与其他方法进行比较比较,验证了所提方法的准确率及泛化性,为解决多域下的新故障检测提供了有效的解决思路。

需要指出的是,在实际工业应用中,齿轮和轴承等还可能发生不同严重程度故障,如何将所提方法应用于不同故障严重程度评估和跨工况下的迁移诊断也是未来研究的重点。

参考文献

[1] LI W, HUANG R, LI J, et al. A perspective survey on deep transfer learning for fault diagnosis in industrial scenarios: Theories, applications and challenges [J]. Mechanical Systems and Signal Processing, 2022, 167: 108487.

[2] 李巍华, 何琛, 陈祝云, 等. 基于对称式对比学习的齿轮箱无监督故障诊断方法 [J]. 仪器仪表学报, 2022, 43(3): 121-131.

LI W H, HE CH, CHEN ZH Y, et al. Unsupervised

- fault diagnosis of gearbox based on symmetrical contrast learning[J]. Chinese Journal of Scientific Instrument, 2022, 43(3): 121-131.
- [3] 董勋, 郭亮, 高宏力, 等. 代价敏感卷积神经网络: 一种机械故障数据不平衡分类方法[J]. 仪器仪表学报, 2019, 40(12): 205-213.
- DONG X, GUO L, GAO H L, et al. Cost sensitive convolutional neural network: A classification method for imbalanced data of mechanical fault[J]. Chinese Journal of Scientific Instrument, 2019, 40(12): 205-213.
- [4] 唐竞鹏, 王红军, 钟建琳, 等. 基于 WDCNN-SVM 深度迁移学习的燃气轮机转子故障诊断方法[J]. 电子测量与仪器学报, 2021, 35(11): 115-123.
- TANG J P, WANG H J, ZHONG J L, et al. Gas turbine rotor fault diagnosis method based on WDCNN-SVM deep transfer learning[J]. Journal of Electronic Measurement and Instrumentation, 2021, 35(11): 115-123.
- [5] 陈仁祥, 朱玉清, 胡小林, 等. 自适应正则化迁移学习的不同工况下滚动轴承故障诊断[J]. 仪器仪表学报, 2021, 41(8): 95-103.
- CHEN R X, ZHU Y Q, HU X L, et al. Fault diagnosis of rolling bearing under different working conditions using adaptation regularization based transfer learning [J]. Chinese Journal of Scientific Instrument, 2021, 41(8): 95-103.
- [6] 陈祝云, 钟琪, 黄如意, 等. 基于增强迁移卷积神经网络的机械智能故障诊断[J]. 机械工程学报, 2021, 57(21): 96-105.
- CHEN ZH Y, ZHONG Q, HUANG R Y, et al. Intelligent fault diagnosis for machinery based on enhanced transfer convolutional neural network [J]. Journal of Mechanical Engineering, 2021, 57(21): 96-105.
- [7] LI X, ZHANG W, DING Q, et al. Multi-layer domain adaptation method for rolling bearing fault diagnosis[J]. Signal Process, 2019, 157: 180-197.
- [8] XIA Y, SHEN C, WANG D, et al. Moment matching-based intraclass multisource domain adaptation network for bearing fault diagnosis[J]. Mechanical Systems and Signal Processing, 2022, 168, 108697.
- [9] ZHANG B, LI W, NG S. Intelligent fault diagnosis under varying working conditions based on domain adaptive convolutional neural networks [J]. IEEE Access, 2018, 6: 66367-66384.
- [10] HAN T, LIU C, YANG W, et al. A novel adversarial learning framework in deep convolutional neural network for intelligent diagnosis of mechanical faults[J]. Knowledge-Based Syst, 2019, 165(1): 474-487.
- [11] CHEN Z, HE G, LI J, et al. Domain adversarial transfer network for cross-domain fault diagnosis of rotary machinery [J]. IEEE Transactions on Instrumentation and Measurement, 2020, 69(11): 8702-8712.
- [12] GUO L, LEI Y, XING S, 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.
- [13] LI J, HUANG R, HE G S, et al. A deep adversarial transfer learning network for machinery emerging fault detection [J]. IEEE Sensors Journal, 2020, 20(15): 8413-8422.
- [14] LI J, HUANG R, HE G, et al. A two-stage transfer adversarial network for intelligent fault diagnosis of rotating machinery with multiple new faults [J]. IEEE/ASME Transactions on Mechatronics, 2020, 26(3): 1591-1601.
- [15] ZHAO C, SHEN W. Dual adversarial network for cross-domain open set fault diagnosis [J]. Reliability Engineering & System Safety, 2022, 221, 108358.
- [16] ZHANG W, LI X, MA H, et al. Open set domain adaptation in machinery fault diagnostics using instance-level weighted adversarial learning [J]. IEEE Transactions on Industrial Informatics, 2021, 17(11): 7445-7455.
- [17] YU X, ZHAO Z, ZHANG X, et al. Deep learning-based open set fault diagnosis by extreme value theory [J]. IEEE Transactions on Industrial Informatics, 2021, 18(1): 185-196.
- [18] ZHANG W, LI X, MA H, et al. Universal domain adaptation in fault diagnostics with hybrid weighted deep adversarial learning [J]. IEEE Transactions on Industrial Informatics, 2021, 17(12): 7957-7967.
- [19] ZHOU Z. Ensemble learning. In Machine learning [J]. Singapore, 2021: 181-210.

- [20] YOU K, LONG M, CAO Z, et al. Universal domain adaptation [J]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019: 2720-2729.
- [21] GANIN Y, LEMPITSKY V. Unsupervised domain adaptation by backpropagation [J]. International Conference on Machine Learning, PMLR, 2015: 1180-1189.
- [22] FU B, CAO Z, LONG M, et al. Learning to detect open classes for universal domain adaptation [J]. In European Conference on Computer Vision, Springer, Cham, 2020: 567-583.

作者简介



陈祝云, 2013 年于南京农业大学获得学士学位, 2020 年于华南理工大学获得博士学位(硕博连读), 现为华南理工大学博士后/助理研究员, 主要研究方向为工业大数据处理、装备智能运维与健康管理。

E-mail: mezychen@scut.edu.cn

Chen Zhu Yun received his B. Sc. degree from Nanjing

Agricultural University of Technology in 2013, and received his Ph. D. degree with successive master & doctoral program from South China University of Technology in 2020. He is currently a postdoctoral fellow at South China University of Technology. His main research interests include industrial big data processing and intelligent maintenance and health management of equipment.



李巍华(通信作者), 1995 年于太原理工大学获得学士学位, 1998 年于太原理工大学获得硕士学位, 2003 年于华中科技大学获得博士学位, 现为华南理工大学教授, 主要研究方向为装备智能运维、汽车智能驾驶等。

E-mail: whlee@scut.edu.cn

Li Weihua (Corresponding author) received his B. Sc. and M. Sc. degrees both from Taiyuan University of Technology in 1995 and 1998, and received his Ph. D. degree from Huazhong University of Science and Technology in 2003. He is currently a professor at South China University of Technology. His main research interests include intelligent maintenance and health management and autonomous vehicle.