

DOI: 10.19650/j.cnki.cjsi.J2310975

基于双识别器对抗的开放域自适应故障诊断方法^{*}

余博¹, 梁伟阁¹, 秦奋起², 董海迪¹

(1. 海军工程大学兵器工程学院 武汉 430000; 2. 中国船舶重工集团公司第七一三研究所 郑州 450000)

摘要: 领域自适应问题在机械设备故障诊断领域已被广泛研究, 当前大多数封闭域自适应方法通常都假设源域和目标域共享相同的标签类型空间。然而, 这不完全符合机械设备真实的诊断需求, 实际上会出现新的故障类型, 因而传统依据边缘分布对齐的方法难以处理此开放域问题, 不能正确辨识出已存在的样本类型和新出现的类型。针对源域和目标域标签类型空间部分重叠的这另一典型开放域诊断问题, 提出了一种基于双识别器对抗的开放域自适应故障诊断方法。两个结构相同的神经网络被引入进行对抗性训练, 以增强模型对已知类型辨识的领域自适应性能, 利用源域与目标域熵最大与最小化策略, 以及目标域样本输出的二元交叉方案用以建立分界线隔离新出现的未知类型。利用轴承数据和自吸式离心泵数据进行分析验证, 实验结果表明: 所提方法相对于典型的封闭域和开放域模型, 能更准确的判定机械设备已存在的故障类型和新出现的未知故障类型, 在各诊断任务中, 均能达到90%以上的正确率。

关键词: 封闭域; 开放域; 双识别器; 熵; 故障诊断

中图分类号: TH165.3 **文献标识码:** A **国家标准学科分类代码:** 520.20

Open set domain adaptation method based on adversarial dual classifiers for fault diagnosis

She Bo¹, Liang Weige¹, Qin Fenqi², Dong Haidi¹

(1. College of Weaponry Engineering, Naval University of Engineering, Wuhan 430000, China;

2. Research Institute of China Shipbuilding, Zhengzhou 450000, China)

Abstract: The domain adaptation problem has been widely studied in the field of mechanical equipment fault diagnosis. At present, the most closed set domain adaptation methods generally assume the source domain and target domain share the same label space, which is not practical in real application. This can be called open set domain, because novel fault classes may actually emerge, these conventional methods which only rely on marginal distribution alignment are difficult to separate the new emerging classes and known classes. One of the typical open set domain adaptation problems is that the label spaces of source domain and target domain are partially overlapped. In this article, a novel open set domain adaptation method based on adversarial dual classifiers (OSDA-ADC) is proposed to address this issue. Two neural networks with the same structure are introduced for adversarial training to enhance the domain adaptive performance of the model for known classes identification. The maximization and minimization entropy strategies of source domain and target domain, as well as the binary cross scheme of target domain sample output are used to establish a boundary to isolate unknown classes. In addition, the bearing data set and the self-priming centrifugal pump are selected to evaluate the effectiveness of the proposed method. The experimental results show that the proposed method can more accurately identify the existing known fault classes and new emerging unknown fault classes of mechanical equipment than the typical closed set domain and open set domain models. In each diagnosis task, the proposed method can achieve more than 90% accuracy.

Keywords: close set domain; open set domain; dual classifiers; entropy; fault diagnosis

收稿日期: 2023-01-11 Received Date: 2023-01-11

^{*} 基金项目: 国家自然科学基金(62101579)项目资助

0 引言

机械设备在工业、交通、国防等领域是重要的基础设施,在寿命周期内运行,力学环境的变化会导致机械零部件受力情况与正常状态下产生偏差,零部件表面易出现磨损、裂纹等变化,一旦疲劳恶化,可能造成机械设备的停机。监测、诊断、预测、数据孪生等技术的发展,是推进传统工业制造向智慧车间、智能制造方向达成的有力措施^[1-3]。

近年来,基于数据驱动的故障诊断技术得以长足发展,特别是深度学习方法在故障诊断领域已经被广泛研究。相对于传统的浅层机器学习方法,卷积神经网络、稀疏自编码、深度残差网络等深度学习方法在辨识性特征提取方面更具优势。杨蕊等^[4]提取原始信号的频谱的相关峭度值作为新的样本,结合深度自编码网络实现了轴承的智能诊断;Kong等^[5]提出了采用多个不同激活函数的堆栈自编码网络,来提取差异性的隐藏特征;Plakias等^[6]提出了一种注意力机制的卷积神经网络,来克服依据人工经验主观性特征提取进行轴承诊断的不足。这些研究的共同之处在于机械设备的工况是恒定的,比如转速、负载不变等。然而,在实际的工业制造中,设备运行工况的变化更符合真实情况,这也使得诊断模型所采用的源域和目标域数据分布存在差异,传统的诊断模型缺乏领域适应性,难以满足诊断需求。

为解决这一问题,领域自适应性技术被用于提高诊断模型的泛化性。其中,一种方案是将源域学习的知识迁移到目标域,减弱样本数据分布差异带来的影响。陈仁祥等^[7]提出了在注意力卷积神经网络中添加领域适配层,完成不同转速下滚动轴承深度特征的迁移适配;罗宏林等^[8]提出了将不同工况下特征值集,利用有监督学习模式构造的自编码器,向参考工况下特征集做映射迁移。解决跨域诊断问题的另一种方案是最小化不同领域间的偏差来对齐源域与目标域的分布。Mao等^[9]提出了在特征层和模型层构建具有联合对抗训练的深度神经网络,获取域不变特征;Ma等^[10]采用最大平均差异(maximum mean discrepancy, MMD)度量来评估源域和目标域之间的距离,以提取域不变特征;Saito等^[11]提出了最大识别器差异(maximum classifier discrepancy, MCD)度量算法来建立决策分界线,以对齐源域与目标域分布。

然而,大部分领域适应性研究方法均假设源域和目标域共享相同的标签类型空间,也就是目标域中包含的机械设备样本状态类型与源域中样本状态类型一致。很明显,这样的假设在实际诊断场景应用中存在很大的局限性。例如,存在目标域样本类型空间为源域样本类型空间子集的情况,即目标域样本类型小于源域样本类型,

为解决此类部分域转移学习的问题^[12],Li等^[13]提出了一种类权重对抗性神经网络,Zhang等^[14]也提出了一种类加权对齐的迁移网络,这两种方法均激励知识在共享已知类之间的正向迁移,并减弱源域中独有类型的样本对知识迁移的负影响;雷亚国等^[15]提出了机械故障靶向迁移诊断方法,解决跨设备目标域样本标签缺失的问题。然而,这些研究尚忽视了另一类诊断场景,也就是目标域会存在新出现的故障类型,即源域与目标域有共同的已知类故障类型,而目标域还有源域中不存在的未知类故障。因而,开放域自适应方法被提出用于解决此类问题^[16-19]。Wang等^[20]提出了一种新的深度度量学习模型,用于非平衡故障诊断以及新出现故障类型的判定;Li等^[21]构造了一个深度对抗迁移网络来识别目标域中与源域共同的已知故障类型,另选择一个硬阈值建立分界线来判定新出现故障类型,但是该方法只能隔离一种新出现的故障,而实际场景中可能出现多类未知的故障类型;Zhang等^[22]提出了基于对抗学习策略的3个分类模块,以识别共同的已知类故障类型,同样未知类故障利用阈值进行界定;Mao等^[23]利用1个封闭域对抗网络和1个开放域对抗网络来辨识共同的已知类故障和未知类故障。当前开放域诊断的研究,只是考虑了源域标签类型空间为目标域标签类型空间子集的情况,另一种实际的诊断需求则是源域标签类型空间与目标域标签类型空间存在交叉重叠的情况,也即源域与目标域均有共同的已知类故障,并且目标域存在源域没有的未知类故障,而源域同样也存在目标域没有的未知类故障。

基于此,本文提出一种基于双识别器对抗的开放域自适应(open set domain adaptation approach based on adversarial dual classifiers, OSDA-ADC)故障诊断方法,来解决源域与目标域标签类型空间部分重叠的开放域诊断问题。通过构造两个结构相同的卷积神经网络,基于最大识别器差异度量算法,融入样本权重,来获取域不变特征,用于辨识目标域中的已知类故障;引入熵和目标域样本输出的二元交叉方案,增强目标域未知类样本与已知类样本的差异,无需建立阈值判定目标域中的未知类型。

1 领域部分重叠的开放域问题

领域部分重叠的开放域问题,如图1所示。假设 $S = \{x_i^s, y_i^s\}_{i=1}^{n_s}$ 和 $T = \{x_i^t\}_{i=1}^{n_t}$ 分别表示从机械设备运行中采集的源域与目标域数据集,包含 n_s 个有标签的源域样本,以及 n_t 个无标签的目标域样本。其中,源域样本集 S 和目标域样本集 T 分别服从边缘分布 P_s 和 P_t , Y_s 和 Y_t 分别表示源域和目标域标签类型空间。在领域部分重叠的开放域问题中,假设 $Y^c = Y_s \cap Y_t$, Y^c 表示源域 S 和目标域 T 数据集中相同的已知类型; $Y_s^u = Y_s \setminus Y_t$, Y_s^u 表示源域

中的未知故障类型; $Y_t^U = Y_s/Y_t$, Y_t^U 表示目标域中的未知故障类型; 那么, 数据集边缘分布 $P_s \neq P_t$, 并且标签空间 $Y_s \neq Y_t$ 。

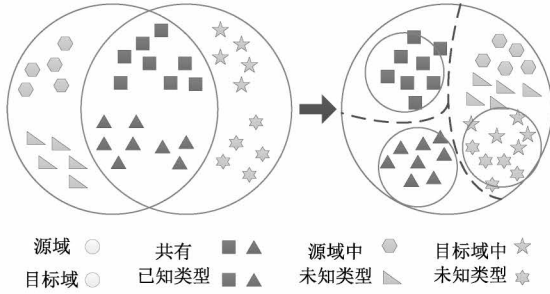


图1 源域与目标域标签类型空间部分重叠的开放域诊断
Fig. 1 Open set domain adaptation of the source domain and target domain with label class

2 最大识别器差异度量算法

大多数基于对抗性学习的领域自适应方法存在着共同的问题: 领域识别器只用于判定提取的特征是否来自源域或是目标域, 不考虑类间的分界线, 导致特征生成器在不同类附近产生难以界定的模糊特征, 影响模型的泛化性。

MCD 度量是一种新的领域自适应方法, 可弥补上述对抗性学习存在的不足。其主要思路是通过模型的训练建立类分界线, 先最大化两个识别器 f_1 和 f_2 输出的差异, 以检测出远离两个识别器支持的目标域样本; 然后, 最小化两个识别器 f_1 和 f_2 输出的差异, 以训练特征生成器 g 用于欺骗识别器, 使得远离识别器支持的目标域样本进

入类分界线内; 最后, 迭代对抗性训练后, 同类型的目标域数据分布与源域数据分布对齐。MCD 的训练步骤如下^[11]:

1) 训练两个识别器和一个特征生成器, 使其能正确的辨识源域样本类型, 目标函数采用最小化 softmax 交叉熵损失函数。

$$\min_{g, f_1, f_2} L(X_s, Y_s) \quad (1)$$

$$L(X_s, Y_s) = -E_{(x_s, y_s) \sim (X_s, Y_s)} \sum_{k=1}^C 1\{k = y_s\} \log p(y|x_s) \quad (2)$$

2) 训练识别器 f_1 和 f_2 , 最大化二者输出的差异。

$$\min_{f_1, f_2} L(X_s, Y_s) - L_{adv}(X_t) \quad (3)$$

$$L_{adv}(X_t) = E_{x_t \sim X_t} [d(p_1(y|x_t), p_2(y|x_t))] \quad (4)$$

$$d(p_1, p_2) = \frac{1}{C} \sum_{k=1}^C |p_{1k} - p_{2k}| \quad (5)$$

式中: $p_1(y|x_t)$ 表示目标域样本 x_t 输入识别器 f_1 的概率输出矩阵, p_{1k} 表示 $p_1(y|x_t)$ 的第 k 维元素, $d(p_1, p_2)$ 表示 L1 范数。

3) 训练生成器 g , 最小化识别器输出的差异。

$$\min_g L_{adv}(X_t) \quad (6)$$

3 基于双识别器对抗的开放域诊断模型

3.1 模型结构

OSDA-ADC 故障诊断模型结构图 2 所示, 包括一个特征生成器 $g(x; \theta_g)$ 、两个结构相同的识别器 $f_1^k(g(x); \theta_{f_1})$ 和 $f_2^k(g(x); \theta_{f_2})$, 网络结构参数如表 1 所示。其中,

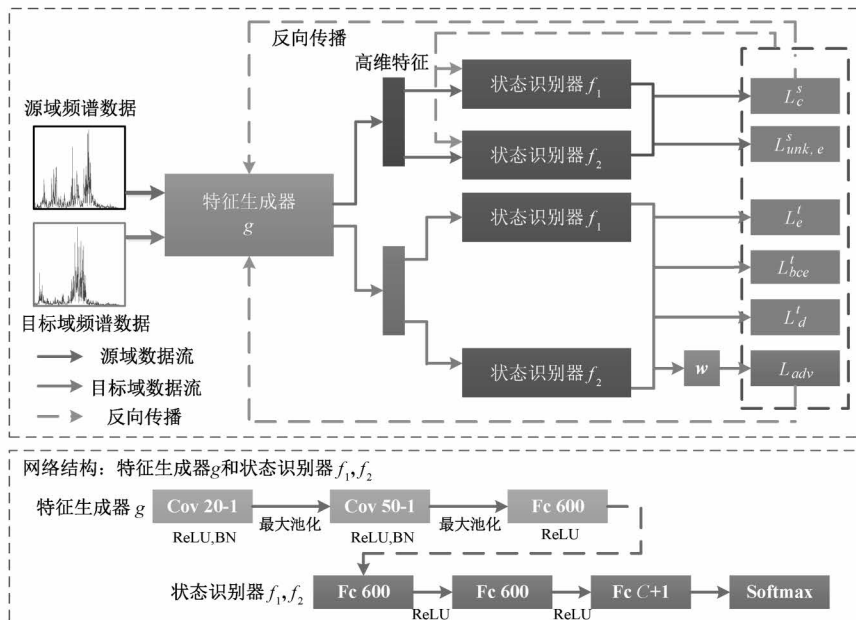


图2 基于双识别器对抗的开放域自适应故障诊断模型结构

Fig. 2 Architecture of the proposed OSDA-ADC

θ_g, θ_{f_1} 和 θ_{f_2} 是网络参数。特征生成器 g 由 2 个卷积层、2 个池化层和 1 个全连接层组成, 两层卷积核分别为 20 和 50, 最大池化区域均为 2×1 , 全连接层神经元数量为 600; 两个识别器 f_1^k 和 f_2^k 在特征生成器 g 网络结构的基础上, 增加 2 个全连接层, 神经元数量分别为 100 和 $C + 1$ 。判定目标域数据类型的流程为: 目标域 x_i 作为特征生成器 g 的输入, 获取的特征 $g(x; \theta_g)$ 经两个识别器 f_1^k 和 f_2^k 处理输出 $C + 1$ 维, 前 C 维表示辨识为已知类型的概率, 最后 1 维表示判定为新出现未知故障类型的概率。

表 1 网络结构参数

Table 1 Parameters of network architecture

序号	网络模块	网络层	参数
1	特征生成器 g	卷积层	卷积核大小: 20×1 ; 激活函数: ReLU; BN 归一化
2		池化层	最大池化: 2×1
3		卷积层	卷积核大小: 50×1 激活函数: ReLU; BN 归一化
4		池化层	最大池化: 2×1
5	状态识别器 f_1 和 f_2	全连接层	神经元数量: 600; 激活函数: ReLU
6		全连接层	神经元数量: 600; 激活函数: ReLU
7		全连接层	神经元数量: $C + 1$
8		输出层	Softmax: $C + 1$ 维

3.2 基于熵的权重

在开放域问题解决方案中, 通常引入熵。例如基于熵最小化的开放域自适应模型 (open set domain adaptation with entropy minimization, OSDA-EM)^[24] 采用熵最小化来实现无监督的跨领域识别; 基于已知类感知自集成的开放域自适应模型 (known-class aware self-ensemble for open set domain adaptation, KASE)^[25] 利用未知类型具有较大的熵来增强对目标域中未知类的判定; 基于未知类软隔离的开放域自适应模型 (open set domain adaptation, OSDA)^[26] 在构建损失函数中使目标域样本熵最小化。

本文也采用熵来表征样本是否属于已知类或者未知类的不确定性。熵越小, 目标域样本更大概率属于已知类; 反之, 越大概率属于新出现的未知类。在最大识别器差异度量算法的基础上, 为了更有效地检测远离两个识别器支持的样本, 识别器对目标域样本的输出, 与源域特征分布对齐这一过程, 给具有较大熵值的目标域样本赋予较小的权重^[26]。基于熵的权重计算如下:

$$w(x_i^t) = \exp\left(-\sum_{k=1}^{C+1} p_k(x_i^t) \log p_k(x_i^t)\right) \quad (7)$$

$$p_k(x_i^t) = \frac{1}{2}(f_1^k(x_i^t) + f_2^k(x_i^t)) \quad (8)$$

式中: $p_k(x_i^t)$ 表示目标域样本 x_i^t 属于第 k 类的概率, $f_1^k(x_i^t)$ 和 $f_2^k(x_i^t)$ 分别表示识别器 f_1 和 f_2 对样本 x_i^t 概率输出矩阵的第 k 维元素。

3.3 已知类和未知类的辨识

最大识别器差异度量算法建立类间分界线的关键是正确识别源域样本类型, 对源域样本的监督训练, 采用最小化交叉熵损失函数 L_c^s , 如下:

$$L_c^s = \frac{1}{n_s} \sum_{i=1}^{n_s} \sum_{k=1}^{C+1} 1[y_i^s = k] \log p_k(x_i^s) \quad (9)$$

源域中也包含目标域不存在的未知类型, 为了增强未知类型与已知类型样本之间的差异, 在源域未知类型样本训练中, 激励其获取较大的熵值, 最大化其损失函数 $L_{unk,e}^s$, 如下:

$$L_{unk,e}^s = - \sum_{i=1}^{n_u} \sum_{k=1}^{C+1} p_k(x_i^{s,u}) \log p_k(x_i^{s,u}) \quad (10)$$

式中: n_u 表示源域中未知类型样本 $x_i^{s,u}$ 的数量。

将基于熵的权重融入最大识别器差异度量算法, 实际上目标域样本权重一定程度上也可以反映出其属于已知类或者未知类的概率。先最大化两个识别器的差异度量损失函数, 检测出远离两个识别器支持的样本; 然后, 最小化两个识别器的差异度量损失函数, 将属于已知类但远离识别器支持的样本归入识别器的支持, 以此判定目标域已知类。损失函数 L_{ade}^t 如下:

$$L_{ade}^t = \frac{1}{n_t} \sum_{i=1}^{n_t} w(x_i^t) |f_1^k(x_i^t) - f_2^k(x_i^t)| \quad (11)$$

假设识别器对目标域样本 x_i^t 的概率输出 $p_k(x_i^t)$ 的各维均在 $1/(C + 1)$ 左右, 那么判定为第 k 类的可信度较低。为使对目标域样本辨识的不确定度低、可信度高, 最小化目标域样本的熵值损失函数 L_e^t , 如下:

$$L_e^t = - \sum_{i=1}^{n_t} \sum_{k=1}^{C+1} p_k(x_i^t) \log p_k(x_i^t) \quad (12)$$

本文提出的方法不利用阈值判定目标域未知类, 而是依据 $p_k(x_i^t)$ 的第 $C + 1$ 维。因此, 利用两个二元交叉方案, 其一是使目标域样本概率输出的第 $C + 1$ 维的值趋近于 t , 这里 $t = 0.5$, 最小化二元交叉熵损失函数 L_{bce}^t , 如下:

$$L_{bce}^t = \frac{1}{n_t} \sum_{i=1}^{n_t} t \log(p(y = C + 1 | x_i^t)) - \frac{1}{n_t} \sum_{i=1}^{n_t} (1 - t) \log(1 - p(y = C + 1 | x_i^t)) \quad (13)$$

其二是为避免将目标域中所有已知类、未知类样本都判定为未知类^[24], 在 $p_k(x_i^t)$ 的前 C 维与第 $C + 1$ 之间建立差异损失函数, 最小化二元交叉熵损失函数 L_d^t ,

如下:

$$L_d^t = q_{C+1}^t \log q_{C+1}^t + (1 - q_{C+1}^t) \log(1 - q_{C+1}^t) \quad (14)$$

$$q_{C+1}^t = \frac{1}{n_t} \sum_{i=1}^{n_t} p(y = C + 1 | x_i^t) \quad (15)$$

$$1 - q_{C+1}^t = \frac{1}{n_t} \sum_{i=1}^{n_t} \sum_{k=1}^C p(y = k | x_i^t) \quad (16)$$

3.4 损失函数的训练

所提出的 OSDA-ADC 模型训练分为 3 步,需要说明的是,不存在有标签的目标域样本参与模型的训练。

1) 利用源域样本及其标签对特征生成器 g 和两个识别器 f_1, f_2 进行监督学习,更新网络参数 θ_g, θ_{f_1} 和 θ_{f_2} , 损失函数如下:

$$\min_{\theta_g, \theta_{f_1}, \theta_{f_2}} L_c^s \quad (17)$$

2) 训练两个识别器 f_1, f_2 , 增加二者之间输出的差异,更新网络参数 θ_{f_1} 和 θ_{f_2} , 损失函数如下:

$$\min_{\theta_{f_1}, \theta_{f_2}} L_c^s - L_{unk,e}^s + \lambda_b L_{bce}^t + L_e^t + \lambda_d L_d^t - L_{adv}^t \quad (18)$$

式中: $\lambda_b = 0.2, \lambda_d = 0.3$ 。

3) 训练特征生成器 g , 减小两个识别器输出的差异,更新网络参数 θ_g , 损失函数如下:

$$\min_{\theta_g} L_{adv}^t \quad (19)$$

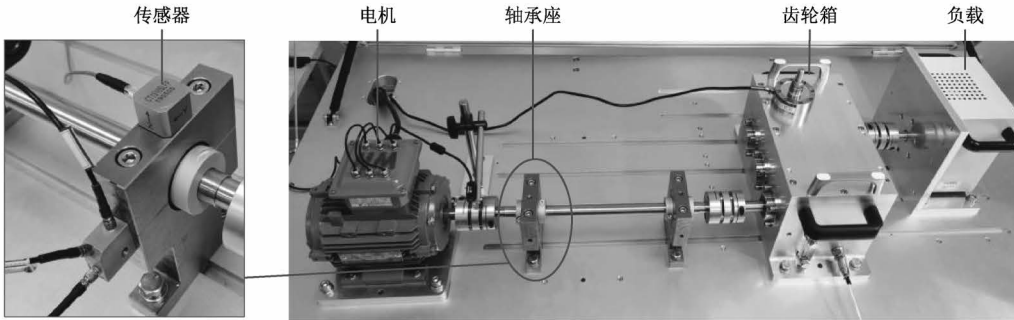


图 3 轴承实验台

Fig. 3 Bearing testing rig

实验数据集如表 2 所示,本实验数据的采样频率为 200 kHz,并且实验过程中电机的转速分别设定 1 000、1 250、1 750 rpm,可获取不同工况下轴承不同状态下的振动信号。此外,构造源域和目标域样本集,其中,在源域中,为每类轴承状态构造 60 个样本,在目标域中,为每类轴承状态构造 90 个样本。

为了检验所提出的模型在变工况条件下,解决开放域诊断问题的能力,本节在此轴承数据集的基础上,构造了 24 个开放域自适应诊断子任务,如表 3 所示。其中,任务 J 包括 6 个子任务 $J_1 \sim J_6$,各子任务中目标域与源域的类型用表 1 中轴承的状态类型 H, C, I-O-B 表示;任务 L 包括子任务 $L_1 \sim L_6$,各子任务中目标域与源域的类型用 H, I, O-B-C 表示;任务 M 包括子任务 $M_1 \sim M_6$,各

4 实验验证与分析

本节采用两个数据集对所提出的方法进行验证,分别为:轴承损伤数据集和自吸式离心泵损伤数据集。需要说明的是,在本节中不论是源域中的未知类型,或是目标域中的未知类型,都只归为 1 类,统称为未知类型。为了验证所提出模型的有效性,采用以下典型识别器进行对比分析,将原始时域样本进行傅里叶变换,得到的频域数据作为各识别器模型的输入。

1) 卷积神经网络 (convolutional neural network, CNN): CNN 的网络结构与所提方法中 f_1 的网络结构一致,在最后一层连接 Softmax 函数层。

2) 领域自适应识别模型: MCD 度量方法^[11]。

3) 开放域自适应识别模型: OSDA-EM^[24]、KASE^[25]、OSDA^[26]。

4.1 轴承损伤实验分析

如图 3 所示,轴承故障综合实验台由电机、轴承座、齿轮箱和负载组成,加速度传感器安装在轴承座上以采集实验台运行中的振动信号。轴承状态共有 5 类,包括:健康状态、混合损伤、内圈损伤、外圈损伤和滚动体损伤。

表 2 实验数据集

Table 2 Data set of the experiment

损伤位置	状态类型	标签
轴承	健康状态 (H)	1
	混合损伤 (C)	2
	混合损伤包括: 内圈、外圈、滚动体裂纹	
	内圈损伤 (I)	3
	外圈损伤 (O)	4
	滚动体损伤 (B)	5

子任务中目标域与源域的类型用 H, C-I, O-B 表示;任务 S 包括子任务 $S_1 \sim S_6$,各子任务中目标域与源域的类型

型用 H, B-C, I-O 表示。此外,“1 000→1 250”表示源域与目标域的数据采集工况,源域数据采集于 1 000 rpm 的转速下,目标域数据采集于 1 250 rpm 的转速。以任务 J 为例,各开放域诊断任务组成如下:

任务 J: 包括子任务 $J_1 \sim J_6$, 该任务中轴承状态有 4 类, 由 3 类已知类和 1 类未知类组成, 源域与目标域用 H, C, I-O-B 表示。其中, H, C, I 表示源域和目标域中共同存在的 3 类轴承状态, 将其设定为 3 类已知类, 分别为: 健康状态、混合损伤和内圈损伤; O 表示源域中未知的轴承状态, 为外圈损伤, 并将其设定为 1 类未知类; B 表示目标域中未知的轴承状态, 为滚动体损伤, 也将其设定为 1 类未知类。在各子任务中, 源域样本共有 240 个, 目标域共有 360 个样本。

表 3 轴承数据开放域诊断任务
Table 3 Details of the open set domain adaptation tasks on the bearing dataset

序号	源域→目标域	诊断任务			
1	1 000→1 250	J_1	L_1	M_1	S_1
2	1 000→1 750	J_2	L_2	M_2	S_2
3	1 250→1 000	J_3	L_3	M_3	S_3
4	1 250→1 750	J_4	L_4	M_4	S_4
5	1 750→1 000	J_5	L_5	M_5	S_5
6	1 750→1 250	J_6	L_6	M_6	S_6

将所提出方法以及对比的各模型训练的 Epochs 取 30, 批处理 Batch-size 取 20。采用不同的识别模型对变工况下 6 个开放域任务 ($J_1 \sim J_6$) 的诊断正确率如图 4 所示。可以发现, 适用于解决传统封闭域故障诊断问题的模型, 如 CNN 和 MCD 方法, 在大部分子任务中诊断正确率均低于 OSDA-EM、KASE、OSDA 开放域诊断模型, 主要原因是封闭域模型假设源域与目标域样本对齐分布, 而各开放域模型都考虑了新出现的未知故障类型对源域和目标域边缘分布的影响, 以增大未知类型与已知类型之间的差异。此外, 相对于本文所提出的 OSDA-ADC 方法, 在大部分诊断子任务中, 此 3 种开放域模型获取的正确率都低于 OSDA-ADC 模型。

针对不同的开放域诊断场景 J、L、M 和 S, 以及 $J_1 \sim J_6$ 、 $L_1 \sim L_6$ 、 $M_1 \sim M_6$ 、 $S_1 \sim S_6$ 这 24 个开放域子任务, 各模型的诊断效果如图 5 所示。可知, 在不同的开放域场景中, OSDA-EM 和所提出的 OSDA-ADC 模型的平均诊断正确率波动较小, 相对较稳定。此外, 各场景中 OSDA-ADC 模型的平均诊断正确率也高于 CNN、MCD、KASE、OSDA 和 OSDA-EM 模型。

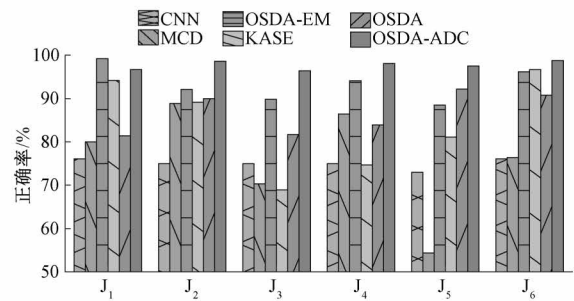


图 4 不同方法在开放域任务下 ($J_1 \sim J_6$) 的诊断正确率

Fig. 4 Testing accuracies in the open set domain adaptation tasks ($J_1 \sim J_6$) by different methods

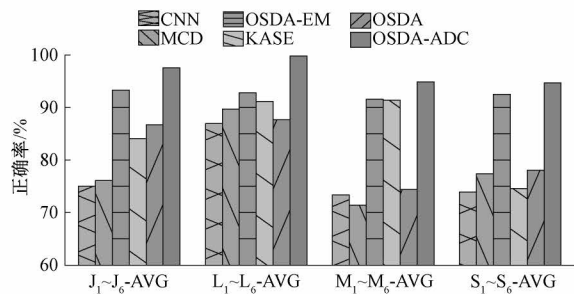


图 5 不同方法在开放域任务下的平均诊断正确率

Fig. 5 Average testing accuracies in the open set domain adaptation tasks by different methods

为了进一步检验各模型对目标域已知类样本和未知类样本的诊断效果, 以开放域诊断场景 J 为例, 列举各模型对 6 个子任务 $J_1 \sim J_6$ 的诊断正确率, 如图 6 所示。可知, 在此诊断任务中, 封闭域诊断模型 CNN 几乎无法正确判定目标域中的未知类故障, MCD 正确分离未知类型的能力也较弱; OSDA-EM 和 KASE 模型能较准确地诊断出目标域中的未知类, 而对目标域中的已知类故障类型, 变工况下诊断能力有所欠缺; OSDA 模型能较好的识别已知类故障, 但是, 也存在较大误诊概率将未知类划分为已知类, 导致其准确判定未知类的能力相对较弱; 然而, 所提出的 OSDA-ADC 模型不仅能较准确地隔离目标域中的未知类, 同时, 变工况下也具有较强的领域自适应性来判定已知故障类型。

4.2 自吸式离心泵数据

实验数据来源于自吸式离心泵^[27]数据采集系统, 如图 7 所示。加速度传感器安装在电机外壳上方, 轴转速为 2 900 rpm, 采样频率为 10 240 Hz。

实验数据包括离心泵的 5 类状态: 正常状态(N)、轴承内圈磨损(I)、轴承外圈磨损(O)、轴承滚动体磨损(R)和叶轮磨损(P), 如表 4 所示。在源域中, 为每类离心泵状态构造 45 个样本, 在目标域中, 为每类离心泵状态构造 30 个

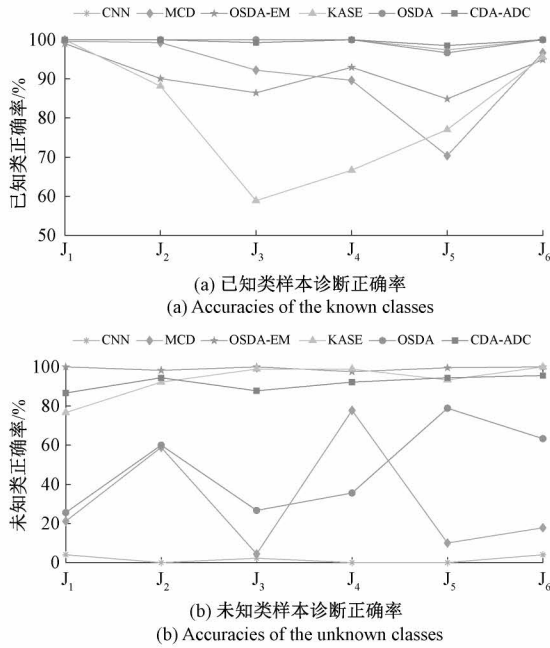


图6 不同方法在开放域任务下($J_1 \sim J_6$)对已知类和未知类样本的诊断正确率

Fig. 6 Testing accuracies of the known and unknown classes in the open set domain adaptation tasks ($J_1 \sim J_6$) by different methods

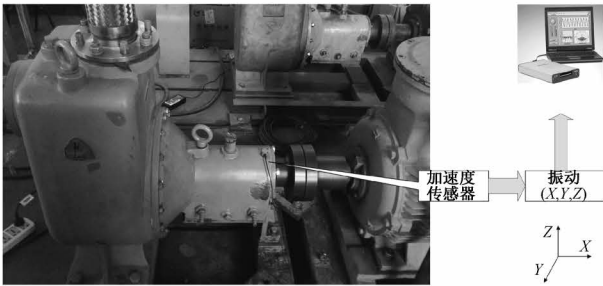


图7 自吸式离心泵数据采集系统

Fig. 7 Self-priming centrifugal pump data acquisition system

样本,各样本由1 024个采样点组成。本节在离心泵数据集的基础上,构造了8个开放域自适应诊断子任务。

任务D:包括4个子任务($D_1 \sim D_4$),该任务中离心泵状态有4类,由3类已知类和1类未知类组成。子任务 D_1 的源域与目标域用N, I, O-R-P表示,子任务 D_2 的源域与目标域用N, P, I-O-R表示,子任务 D_3 的源域与目标域用N, R, P-I-O表示,子任务 D_4 的源域与目标域用N, O, R-P-I表示。以子任务 D_1 为例,N, I, O表示源域和目标域中共同存在的3类离心泵状态,将其设定为3类已知类,分别为:正常状态、内圈损伤、外圈损伤;R表示源域中未知的离心泵状态,为滚动体损伤,并将其设定为1类未知类;P表示目标域中未知的离心泵状态,为叶轮磨损,

表4 实验数据集

Table 4 Data set of the experiment

损伤位置	状态类型	标签
离心泵	正常状态(N)	1
	内圈损伤(I)	2
	外圈损伤(O)	3
	滚动体损伤(R)	4
	叶轮磨损(P)	5

也将其设定为1类未知类。在各子任务中,源域样本共有180个,目标域共有120个样本。

任务E:包括4个子任务($E_1 \sim E_4$),该任务中离心泵状态有3类,由2类已知类和1类未知类组成。子任务 E_1 的源域与目标域用N, I-O, R-P表示,子任务 E_2 的源域与目标域用N, P-I, O-R表示,子任务 E_3 的源域与目标域用N, R-P, I-O表示,子任务 E_4 的源域与目标域用N, O-R, P-I表示。以子任务 E_1 为例,N, I表示源域和目标域中共同存在的2类已知类,分别为:正常状态、内圈损伤;O, R表示源域中未知的离心泵状态,为外圈和滚动体损伤,并将其设定为1类未知类;P表示目标域中未知的离心泵状态,为叶轮磨损,也将其设定为1类未知类。在各子任务中,源域样本共有180个,目标域共有90个样本。

为了验证所提模型的有效性,也采用CNN、MCD、KASE、OSDA-EM、OSDA模型进行对比分析,将采集的离心泵时域振动信号进行傅里叶变换,各类型频谱数据作为各识别器模型的输入,各模型训练的Epochs取30,批处理Batch-size取12。

各模型对开放域诊断子任务 D_1, D_2, D_3, D_4 的诊断正确率、对任务D中4个子任务 $D_1 \sim D_4$ 的平均诊断正确率以及对任务E中4个子任务 $E_1 \sim E_4$ 的平均诊断正确率结果如图8所示。可以看到,利用自吸式离心泵在工业生产中产生的实际振动数据构建的开放域诊断任务中,所提出的OSDA-ADC模型的诊断正确率高于CNN、MCD、KASE、OSDA和OSDA-EM模型。此外,在处理不同的子任务 D_1, D_2, D_3, D_4 中,OSDA-ADC模型的诊断正确率依次为100%、99.1%、100%、100%,波动较小。然而,KASE、OSDA和OSDA-EM开放域模型的诊断正确率波动较大。

为验证各模型提取辨识性特征的能力,以子任务 J_5 为例,将各模型最后一层全连接层的隐藏性特征进行提取,并利用t分布随机近邻嵌入(t-distributed stochastic neighbor embedding, t-SNE)流形学习方法^[28]将提取的高维特征降至2维,散点分布如图9所示。S-1, S-2, S-3和T-1, T-2, T-3分别表示源域和目标域中已知类(类别1, 2, 3)的特征,S-unknown表示源域中的未知类(类别4)的特征,T-unknown表示目标域中的未知类

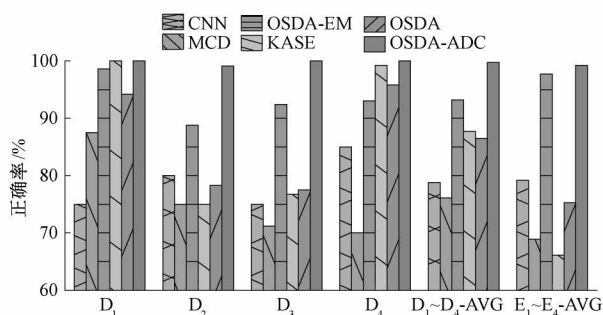
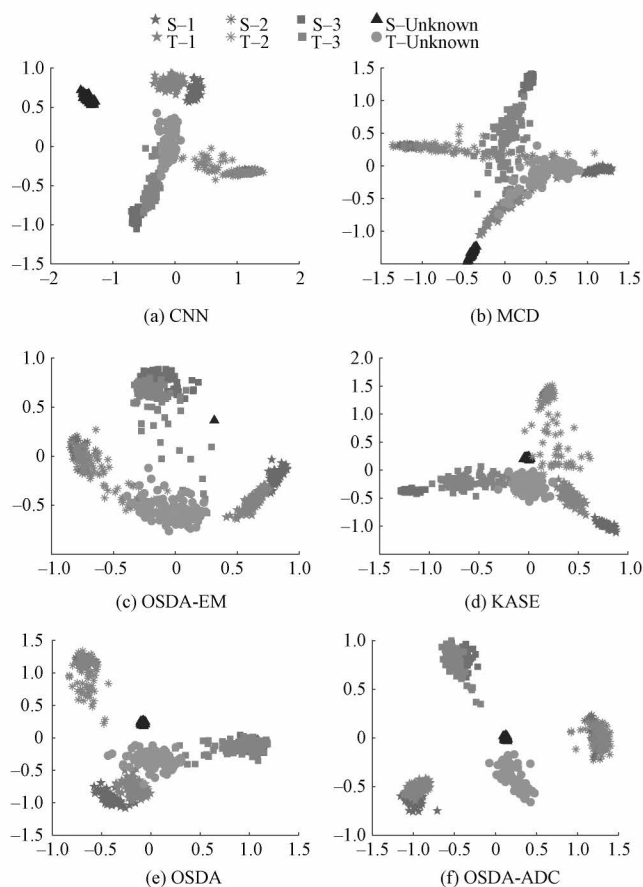


图8 不同方法在开放域任务下(D和E)的诊断正确率

Fig. 8 Testing accuracies in the open set domain adaptation tasks (D and E) by different methods

图9 在开放域任务下(J₅)不同方法的隐藏特征可视化图Fig. 9 Feature visualizations of different methods for the open set domain adaptation task (J₅)

(类别5)的特征。可知,在此诊断子任务中,利用 CNN、MCD 模型提取的目标域未知类特征,在空间上与已知类的特征更为接近,并不利于对目标域未知类的判定。采用所提出方法 OSDA-ADC 提取的目标域未知类特征,相对于 OSDA-EM、KASE、OSDA 开放域模型提取的特征,不

仅在空间上更加聚集,并且远离各已知类,这将有利于分离目标域中的未知类。在子任务 J₅ 中,OSDA-ADC 模型诊断已知类故障的正确率为 98.5%,判定未知类的正确率为 94.4%;OSDA 模型诊断已知类故障的正确率为 96.6%,而判定未知类的正确率为 78.8%,特征分布图中显示部分目标域未知类的特征接近已知类,也易引起未知类误诊断为已知类,导致诊断未知类的正确率相对较低。

5 结 论

本文提出了一种基于双识别器对抗的开放域自适应故障诊断方法,针对源域与目标域标签空间部分交叉重叠的诊断需求,采用实验室变工况下轴承数据与工业上自吸式离心泵损伤数据集进行验证,所提方法不仅能辨识出目标域中与源域相同的已知故障类型,也能判定目标域中新出现的未知故障类型。主要结论如下:

1) 采用加权的最大识别器差异度量算法,通过迭代的最大化与最小化双识别器输出差异,可使识别器对目标域中已知类型样本的输出能与源域相同已知类型样本分布对齐,并减弱源域未知类型对辨识目标域已知类型的影响,提升模型领域自适应性。

2) 采用熵值表征样本是否属于已知类或者未知类的不确定性,并结合二元交叉方案,增强未知类型样本与已知类型样本的差异,提升模型判定目标域未知故障类型的准确性。

3) 利用双识别器输出的 $C+1$ 维,第 $C+1$ 维数值最大即判定为未知类型,无需依赖人工经验建立软硬阈值辅助判定。

参考文献

- [1] BRITO L C, SUSTO G A, BRITO J N, et al. An explainable artificial intelligence approach for unsupervised fault detection and diagnosis in rotating machinery [J]. Mechanical Systems and Signal Processing, 2022, 163: 1-21.
- [2] 梁思远,周金滢,高占宝,等. 机电系统健康状态预测和维修决策的双向优化方法[J]. 仪器仪表学报, 2023, 44(1):131-141.
- LIANG S Y, ZHOU J H, GAO ZH B, et al. Bi-directional optimization method for health state prediction and maintenance decision-making of electromechanical systems[J]. Chinese Journal of Scientific Instrument, 2023, 44(1):131-141.

- [3] 陶飞, 张辰源, 张贺, 等. 未来装备探索: 数字孪生装备[J]. 计算机集成制造系统, 2022, 28(1): 1-16.
TAO F, ZHANG CH Y, ZHANG H, et al. Future equipment exploration: Digital twin equipment [J]. Computer Integrated Manufacturing Systems, 2022, 28(1): 1-16.
- [4] 杨蕊, 李宏坤, 王朝阁, 等. 利用 FCKT 以及深度自编码神经网络的滚动轴承故障智能诊断[J]. 机械工程学报, 2019, 55(7): 65-72.
YANG R, LI H K, WANG CH G, et al. Intelligent fault detection for rolling element bearing based on FCKT and deep auto-coding neural network [J]. Journal of Mechanical Engineering, 2019, 55(7): 65-72.
- [5] KONG X G, MAO G, WANG Q B, et al. A multi-ensemble method based on deep auto-encoders for fault diagnosis of rolling bearings [J]. Measurement, 2020, 151: 1-15.
- [6] PLAKIAS S, BOUTALIS Y S. Fault detection and identification of rolling element bearings with attentive dense CNN[J]. Neurocomputing, 2020, 405: 208-217.
- [7] 陈仁祥, 唐林林, 胡小林, 等. 不同转速下基于深度注意力迁移学习的滚动轴承故障诊断方法[J]. 振动与冲击, 2022, 41(12): 95-101.
CHEN R X, TANG L L, HU X L, et al. A rolling bearing fault diagnosis method based on deep attention transfer learning at different rotations [J]. Journal of Vibration and Shock, 2022, 41(12): 95-101.
- [8] 罗宏林, 柏林, 侯东明, 等. 有限变工况特征迁移学习方法及其在高速列车轴箱轴承故障诊断中的应用[J]. 仪器仪表学报, 2022, 43(3): 132-145.
LUO H L, BAI L, HOU D M, et al. A transfer learning method for bearing fault diagnosis under finite variable working conditions and its application in train axle-box bearings fault diagnosis [J]. Chinese Journal of Scientific Instrument, 2022, 43(3): 132-145.
- [9] MAO W T, DING L, LIU Y M, et al. A new deep domain adaptation method with joint adversarial training for online detection of bearing early fault [J]. ISA Transactions, 2020, 122: 444-458.
- [10] MA P, ZHANG H L, FAN W H, et al. A diagnosis framework based on domain adaptation for bearing fault diagnosis across diverse domains [J]. ISA Transactions, 2020, 99: 465-478.
- [11] SAITO K, WATANABE K, USHIKU Y, et al. Maximum classifier discrepancy for unsupervised domain adaptation [C]. In The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2018.
- [12] WANG Q, MICHAU C, FINK O. Missing-class-robust domain adaptation by unilateral alignment [J]. IEEE Transactions on Industrial Electronics, 2021, 68(1): 663-671.
- [13] LI X, ZHANG W, MA H, et al. Partial transfer learning in machinery cross-domain fault diagnostics using class-weighted adversarial network [J]. Neural Networks, 2020, 129: 313-322.
- [14] ZHANG X, WANG J R, JIA S X, et al. Partial domain adaptation method based on class-weighted alignment for fault diagnosis of rotating machinery [J]. IEEE Transactions on Instrumentation and Measurement, 2022, 71: 1-14.
- [15] 雷亚国, 杨彬, 李乃鹏, 等. 跨设备的机械故障靶向迁移诊断方法[J]. 机械工程学报, 2022, 58(12): 1-9.
LEI Y G, YANG B, LI N P, et al. Targeted transfer diagnosis method across different machines [J]. Journal of Mechanical Engineering, 2022, 58(12): 1-9.
- [16] SAITO K, YAMAMOTO S, USHIKU Y, et al. Open set domain adaptation by backpropagation [C]. In The European Conference on Computer Vision (ECCV), 2018.
- [17] GAO Y, MA A J, GAO Y, et al. Adversarial open set domain adaptation via progressive selection of transferable target samples [J]. Neurocomputing, 2020, 410: 174-184.
- [18] LI Z, FANG Z, LIU F, et al. Bridging the theoretical bound and deep algorithms for open set domain adaptation [J]. IEEE Transactions on Neural Networks and Learning Systems, 2021: 1-15.
- [19] CHEN Z Y, LIAO Y X, LI J P, et al. A multi-source weighted deep transfer network for open-set fault diagnosis of rotary machinery [J]. IEEE Transactions on Cybernetics, 2022: 1-12.
- [20] WANG C J, XIN C, XU Z L. A novel deep metric learning model for imbalanced fault diagnosis and toward open-set classification [J]. Knowledge-Based Systems, 2021, 220(23): 1-14.

- [21] LI J P, HUANG R Y, HE G L, et al. A deep adversarial transfer learning network for machinery emerging fault detection [J]. IEEE Sensors Journal, 2020, 20(15): 8413-8422.
- [22] 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.
- [23] MAO G, LI Y B, JIA S X, et al. Interactive dual adversarial neural network framework: An open-set domain adaptation intelligent fault diagnosis method of rotating machinery[J]. Measurement, 2022, 195: 1-11.
- [24] WU X F, CHEN L, ZHANG S F. Open set domain adaptation with entropy minimization [C]. In The Conference on Pattern Recognition and Computer Vision (PRCV), 2020: 29-41.
- [25] LIAN Q, LI W, CHEN L, et al. Known-class aware self-ensemble for open set domain adaptation [J]. ArXiv Preprint, 2019, ArXiv:1905.01068v1.
- [26] XU Y M, CHEN L, DUAN L X, et al. Open set domain adaptation with soft unknown-class rejection [J]. IEEE Transactions on Neural Networks and Learning Systems, 2021: 1-12.
- [27] LU C, WANG Y, RAGULSKIS M, et al. Fault diagnosis for rotating machinery: A method based on image processing[J]. Plos One, 2016, 11(10): 1-22.
- [28] MAATEN L V D, HINTON G. Visualizing data using t-SNE [J]. Journal of Machine Learning Research, 2008, 9(11): 2579-2605.

作者简介



余博,2012年于北京理工大学获得学士学位,2015年于海军大连舰艇学院获得硕士学位,2018年于海军工程大学获得博士学位,现为海军工程大学讲师,主要研究方向为机械设备健康管理技术。

E-mail: she1611@126.com

She Bo received his B.Sc. degree from Beijing Institute of Technology in 2012, received his M.Sc. degree from Dalian Nave Academy in 2015, and received his Ph.D. degree from Naval University of Engineering in 2018. He is currently a lecture at Naval University of Engineering. His main research interest includes machinery health management technology.



梁伟阁(通信作者),分别在2003年、2009年、2013年于海军工程大学获得学士、硕士和博士学位,现为海军工程大学讲师,主要研究方向为武器系统故障诊断、力学环境建模与评估。

E-mail: 1312021010@nue.edu.cn

Liang Weige (Corresponding author) received his B.Sc., M.Sc., and Ph.D. degrees all from Naval University of Engineering in 2003, 2009, and 2013, respectively. He is currently a lecture at Naval University of Engineering. His main research interests include weapon system fault diagnosis, modeling and evaluation of mechanical environment.