



中国科技期刊卓越行动计划项目入选期刊

控制与决策

CONTROL AND DECISION



面向少样本故障诊断的知识自监督深度表征学习方法

姚家琪, 宋鹏宇, 沈萌, 赵春晖, 王文海

引用本文:

姚家琪, 宋鹏宇, 沈萌, 赵春晖, 王文海. 面向少样本故障诊断的知识自监督深度表征学习方法[J]. *控制与决策*, 2024, 39(10): 3357–3365.

在线阅读 View online: <https://doi.org/10.13195/j.kzyjc.2023.0334>

您可能感兴趣的其他文章

Articles you may be interested in

基于广义主成分分析的重构故障子空间建模方法

Reconstructed fault subspace modelling method based on generalized principal component analysis

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

标签Petri网的路径信息在故障诊断中的应用

Application of path information of labeled Petri nets in fault diagnosis

控制与决策. 2021, 36(2): 325–334 <https://doi.org/10.13195/j.kzyjc.2019.0698>

基于核典型相关性-熵成分分析的工业过程质量监测方法

A quality monitoring method for industrial process based on kernel canonical correlation–entropy component analysis

控制与决策. 2021, 36(12): 2999–3006 <https://doi.org/10.13195/j.kzyjc.2020.0548>

基于DLSR的归纳式迁移学习

DLSR based inductive transfer learning method

控制与决策. 2021, 36(12): 2982–2990 <https://doi.org/10.13195/j.kzyjc.2020.0703>

改进集成深层自编码器在轴承故障诊断中的应用

Application of improved ensemble deep auto–encoder in bearing fault diagnosis

控制与决策. 2021, 36(1): 135–142 <https://doi.org/10.13195/j.kzyjc.2019.0270>

面向少样本故障诊断的知识自监督深度表征学习方法

姚家琪¹, 宋鹏宇¹, 沈萌¹, 赵春晖^{1,2†}, 王文海¹

(1. 浙江大学控制科学与工程学院, 杭州 310027;

2. 深圳职业技术学院智能科学与工程研究院, 广东深圳 518055)

摘要: 目前机器学习技术已被广泛应用于工业智能故障诊断中,但其成功应用的前提条件是能够获取到充足的含标签故障数据以对机器学习模型进行训练. 实际工业场景中,设备常运行于正常状态,故障数据的获取与标注成本巨大,无法达到模型训练的需求. 现有少样本诊断方法通常依赖于额外的有标签数据,无法克服工业场景下的数据采集与标注局限. 对此,挖掘关键时序依赖特征以及专家先验知识与故障诊断任务之间的内在关联,提出一种知识自监督深度表征学习方法以实现少样本故障诊断. 该方法设计了一个掩码信号重构与先验特征预测多任务联合的模型预训练策略,利用与目标设备类似设备积累的海量历史无标签数据对工业智能故障诊断模型中的特征提取器模型进行预训练,使得模型在无需额外标注数据的前提下,提取具有专家先验知识引导的时序变化模式,从而获取高泛化故障表征能力. 通过上述基于知识自监督表征的预训练方法,在诊断过程中仅需利用目标设备的少量有标签故障样本对模型的全局参数进行微调,从而克服模型对有标签样本的依赖性难题. 最后,通过一个跨数据集的故障诊断实验来模拟跨设备的少样本故障诊断场景,验证所提出方法在少样本场景下的有效性.

关键词: 少样本学习; 故障诊断; 自监督学习; 表征学习

中图分类号: TP277

文献标志码: A

DOI: 10.13195/j.kzyjc.2023.0334

引用格式: 姚家琪, 宋鹏宇, 沈萌, 等. 面向少样本故障诊断的知识自监督深度表征学习方法[J]. 控制与决策, 2024, 39(10): 3357-3365.

Knowledge-aided self-supervised deep representation learning method for few-shot fault diagnosis

YAO Jia-qi¹, SONG Peng-yu¹, SHEN Meng¹, ZHAO Chun-hui^{1,2†}, WANG Wen-hai¹

(1. College of Control Science and Engineering, Zhejiang University, Hangzhou 310027, China; 2. Institute of Intelligence Science and Engineering, Shenzhen Polytechnic University, Shenzhen 518055, China)

Abstract: Machine learning technology has been widely used in industrial intelligent fault diagnosis, but the prerequisite for its successful application is to obtain enough labeled fault data to train the machine learning model. In actual industrial scenarios, equipment often works in normal state, and the cost of obtaining fault data and marking fault labels is huge. Therefore, the training requirements of machine learning model cannot be guaranteed, and then it is difficult to apply the model to the target equipment. To solve this problem, this paper explores the intrinsic relationship between key temporal dependent features, expert prior knowledge and fault diagnosis tasks, and proposes a knowledge self-supervised deep representation learning method for few-shot fault diagnosis. In this method, a model pre-training strategy combining prior feature prediction and mask signal reconstruction is designed. The feature extractor model of the industrial intelligent fault diagnosis model is pre-trained by using massive historical unlabeled data accumulated by similar equipments. This pre-training strategy can make the model add knowledge guidance of artificial prior features on the basis of mining temporal dependent features of samples, so as to obtain high generalization fault representation ability. After training the industrial fault diagnosis model based on the knowledge self-supervised representation learning method, only a few labeled fault samples of the target device are used to fine-tune the global parameters of the model in the diagnosis process, and the problem of model dependence on a large number of labeled samples will be solved. Finally, a cross-dataset fault diagnosis experiment is conducted to simulate the cross-equipment fault diagnosis scenario with few samples, and the effectiveness of the proposed method in the small-sample scenario is verified.

Keywords: few-shot learning; fault diagnosis; self-supervised learning; representation learning

收稿日期: 2023-03-22; 录用日期: 2023-08-22.

基金项目: 国家自然科学基金杰出青年基金项目 (62125306); 广东省基础与应用基础研究基金项目 (2022A1515240003).

责任编辑: 方京华.

†通讯作者. E-mail: chhzhao@zju.edu.cn.

0 引言

现代工业设备逐渐趋于大型化、复杂化、规模化,如果未能及时发现并定位设备中存在的故障,将会造成巨大的经济损失甚至人员伤亡^[1].为了保障现代工业设备安全平稳运行,基于数据驱动的机器学习方法因其能够自动从数据中挖掘潜在故障模式^[2-3],在工业智能故障诊断领域得到了广泛的关注与应用^[4-8].然而,此类工业智能故障诊断方法通常需要大量的有标签故障样本进行训练^[9].在实际工业场景中,设备在绝大部分时间处于正常工作状态,仅能采集到极少量的故障样本.另一方面,对这些故障样本进行标注往往需要经过停机检查与专家分析,这个过程会耗费巨大的成本;因此,实际工业场景中通常难以收集到充足的含标签故障样本用于机器学习模型训练^[10].为了解决这个问题,少样本故障诊断方法近年来得到了越来越多的关注,即利用少量的来自目标设备的含标签故障样本,训练一个具有泛化能力的故障诊断模型,使之在测试与应用阶段仍能保持较高的诊断精度.现有的解决少样本故障诊断问题的方法包括两类,基于数据增强的方法和基于元学习的方法.

基于数据增强的方法是指利用现有的少量故障数据生成更多的虚拟样本供诊断模型训练.如文献^[11]提出了一种利用多模块梯度惩罚生成对抗网络来生成更多故障样本,文献^[12]对条件变分自编码器进行改进以用于生成更多更加真实的虚拟机械故障样本.然而,生成模型的质量依然依赖于一定规模的训练样本,当真实样本过少时,此类基于数据增强的方法难以保证生成数据的正确性和多样性^[13],进而难以保证通过虚拟样本训练得到的模型能够具有稳定的诊断精度.

基于元学习的方法将模型训练分为了元训练与元测试两个阶段,在元训练阶段引入一个含故障类别标签的额外数据集,该数据集与来自目标设备的数据集具有一定的相似性(如均为轴承故障诊断数据集),通过在该数据集上采样与目标任务相关的额外学习任务并训练,让模型学会能够在新任务上快速泛化的学习经验,然后在元测试阶段利用少量的故障分类标签的目标设备样本对模型进行微调,实现在目标设备的小样本故障诊断^[14],例如,文献^[15]采用基于度量的元学习方法来解决少样本故障诊断问题,利用大规模的额外有标签数据集训练特征距离度量器,然后度量目标样本与各类已有的少量故障样本的平均特征距离,选择最近邻的故障类别作为诊断结果.模型无关的元学习算法(model-agnostic meta-learning, MAML)也被引入用于解决少样本故障诊断

问题.文献^[16]将MAML应用于滚动轴承少样本故障诊断,该方法通过在额外含故障标签数据集上采样额外的元训练任务,让模型寻找任务敏感的初始化参数,可以使得模型在少样本情况下也能在目标任务上优化得到最优参数.然而,目前基于元学习的方法需要引入大量额外的有标签样本集来进行元训练^[14],即便这些额外样本并非来自目标设备,但在工业场景下收集这些额外样本并进行标注的成本仍然较高,为其实际应用带来限制.

目前用于解决少样本故障诊断的方法仍对有标签故障样本存在较大的依赖,这与工业场景下的高标注成本相矛盾,依然难以解决工业场景下的少样本问题.虽然在工业场景中获取充分的有标签故障样本难度较大,但是在大量工业设备服役的过程中依然能够积累海量的无标签数据,这些无标签数据包含了关键的时序依赖特征,使得模型可以利用这些海量无标签数据学习得到关键时序依赖特征提取能力,这些关键时序依赖特征可以反映设备所属的故障状态,具有较好的泛化性能.此外,经典的故障诊断方法常依靠专家在实践中获得的经验,如通过设备运行和故障发生机理,结合信号分析与统计特征提取以及相关的诊断规则来判断故障设备属于何种异常状态.本文将这些诊断经验称为先验知识.研究表明,先验知识有助于模型学习更泛化的故障特征,并有助于提高模型的可解释性^[17].目前应用于工业故障诊断的深度学习方法很少考虑到将这些重要的先验知识融入智能故障诊断模型,若能有效利用这些专家先验知识,则会进一步提高模型的特征提取的泛化能力.总之,样本所包含的关键时序依赖特征与专家所积累的先验知识均与工业故障诊断任务存在一定的内在关联性,通过挖掘这种内在关联性,可实现在无需额外标注数据的前提下解决少样本故障诊断问题.

基于以上认知,本文提出一种面向少样本故障诊断的知识自监督深度表征学习方法.该方法利用采集自与目标设备相似设备的海量无标签数据对模型进行知识自监督表征训练.具体而言,设计了掩码信号重构与先验特征预测两种自监督学习方法进行协同训练.其中掩码信号重构考虑到了工业传感器数据信息密度低^[18]的特点,对输入样本以样本子片段为单位进行大规模掩码并进行自监督重构,使模型获得关键时序依赖特征提取能力.在此基础上,进一步考虑到了专家经验知识可以用于指导模型特征提取,设计先验特征预测任务,以故障诊断中传感器信号的典型时域统计特征作为先验知识的载体与监督信息,引导特征提取器实现泛化特征提取.后续在诊断过

程中仅需对模型全局参数在具有少量有标签故障样本的目标设备的故障数据集上进行微调,从而克服模型对有标签样本的依赖性挑战. 本文的主要工作总结如下:

1) 与现有依赖有标签样本的少样本诊断方法不同,本文面向充分的无标签样本,提出一种全新的少样本故障诊断范式. 通过挖掘关键时序依赖特征以及专家先验知识与故障诊断任务的内在关联性,在无需额外有标签数据的前提下使工业智能故障诊断模型获取少样本故障诊断所要求的泛化特征提取能力.

2) 根据工业数据样本信息密度低的特点,设计大规模掩码信号自重构的表征学习任务,在此基础上进一步设计具有专家先验知识引导的先验特征预测自监督任务,从而形成多任务联合优化,实现在专家先验知识引导下的关键时序依赖特征提取.

1 少样本故障诊断问题描述

在以往的少样本故障诊断的研究中,通常有两种规则来对少样本场景进行定义. 第 1 种规则是限制标签样本数与总样本数的比值,例如,当有标签的训练样本不超过总样本的 10% 时,则定义为少样本场景^[19]. 第 2 种方法是限制有标签的训练样本的绝对数量,例如,当每类训练样本的数量不超过 20 个时,则定义为少样本场景^[1]. 如果仅限制有标签样本比例,则当总体样本规模较大时,仍有充分的有标签故障样本可供训练,因此本文遵循第 2 条少样本定义规则,即当每类有标签故障样本少于或等于 30 时,则视为少样本场景. 为简化说明,使用 K -shot 来描述具体的少样本场景, K 表示每类故障中可用于训练的有标签样本数. 少样本故障诊断要求利用这些少量的有标签样本构建高可靠的工业智能故障诊断模型. 同时,不允许引入额外的有标签故障数据来辅助模型训练,包括采集自其他设备或同设备其他工况的有标签数据,因为这样依然会带来较高的标注成本,而少样本故障诊断应当尽量减少标签依赖. 本文旨在提出一种知识自监督深度表征学习方法,利用易获取且无需标注的无标签数据来帮助提高工业通过故障诊断模型泛化特征的提取能力,以实现少样本故障诊断.

2 知识自监督深度表征学习方法

本节首先介绍所提出方法的总体框架,然后分别对所设计的掩码信号重构任务、先验特征预测任务以及所使用的智能故障诊断模型结构进行介绍.

2.1 知识自监督深度表征学习框架

本文所提出的知识自监督深度表征学习方法框架如图 1 所示. 知识自监督深度表征学习方法包括自

监督预训练与微调两个步骤.

通常一个基于深度学习的智能故障诊断模型可以表示为 $f_d(\mathbf{x}) = h_{\theta_c}(g_{\theta_e}(\mathbf{x}))$. 其中: g_{θ_e} 为特征提取器, h_{θ_c} 为分类头, θ_e 与 θ_c 分别为特征提取器与分类头的可训练参数, \mathbf{x} 为待诊断的样本. 本文主要通过知识自监督表征学习方法使得 g_{θ_e} 能够从海量无标签数据中自动挖掘诊断知识,获得泛化特征提取能力,实现少样本故障诊断.

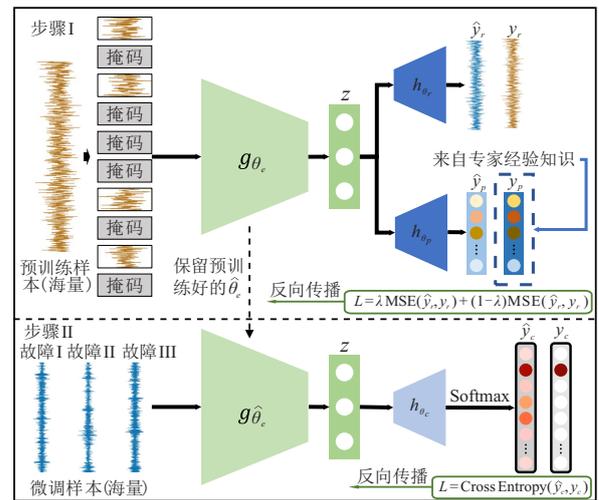


图 1 知识自监督深度表征学习框架

在自监督预训练阶段,收集与目标设备相似设备在历史服役阶段积累的无标签数据制作为预训练样本集 $\{\mathbf{x}_i^s\}_{i=1}^{n_s}$. 其中: s 指的是该样本来自源设备,即与目标设备相似的设备,而 n_s 是样本数量. 本文采用代理任务对 g_{θ_e} 进行自监督预训练. 预训练样本集本身没有故障标签,因此依据代理任务为样本集构建代理标签,生成预训练集 $\{\mathbf{x}_i^s, \mathbf{y}_i^{s1}, \mathbf{y}_i^{s2}, \dots, \mathbf{y}_i^{sm}\}_{i=1}^{n_s}$. 其中: \mathbf{y}_i^{sm} 表示针对样本 \mathbf{x}_i^s 与代理任务 T^m 构建的代理标签, m 表示第 m 个代理任务,即在预训练阶段可采用多个代理任务对 g_{θ_e} 进行联合训练,不同的代理任务可以从不同的角度帮助特征提取器挖掘诊断知识,提升特征提取器的泛化性能. 针对 m 个代理任务,分别对应 m 个损失函数,因此有损失函数集 $\{\mathcal{L}_i\}_{i=1}^m$; 对于第 i 个代理任务,使用相应的代理任务头 $h_{\theta_s^i}$ 代替分类头 h_{θ_c} , θ_s^i 为第 i 个任务头的可训练参数. 这些代理任务头共享同一个特征提取器 g_{θ_e} , 形成一个由特征提取器为主干,多任务头为分支的预训练模型. 在自监督预训练阶段的优化目标为

$$\hat{\theta}_e, \Theta = \arg \min_{\theta_e, \Theta} \sum_{i=1}^m \lambda_i \mathcal{L}_i(\mathbf{y}^i, h_{\theta_s^i}(g_{\theta_e}(\mathbf{x}))). \quad (1)$$

其中: λ_i 为第 i 个代理任务的损失权重,用于调节不同任务对模型预训练的影响程度; Θ 为代理任务头网络的参数集; \mathbf{x} 为预训练样本. 需要指出的是,本文共设计了 2 个代理任务,分别为掩码信号重构任务与先

验特征预测任务. 当预训练完成后, 得到训练好的特征提取器 $g_{\hat{\theta}_e}$, 该特征提取器通过自监督学习从海量的无标签数据中挖掘诊断知识, 具有较好的泛化特征提取能力. 所挖掘到的诊断知识由模型参数承载, 因此保留预训练参数 $\hat{\theta}_e$. 同时, 将分类头替换回原网络, 得到待微调模型 $f_d(\mathbf{x}) = h_{\theta_c}(g_{\hat{\theta}_e}(\mathbf{x}))$, 然后采用目标设备的少样本集 $\{\mathbf{x}_i^t, \mathbf{y}_i^t\}_{i=1}^{n_t}$ 对 f_d 进行全局参数微调. 在微调过程中, 对于特征提取器参数采用较小的学习率, 对分类头参数采用较大的学习率进行优化. 微调所使用的损失函数为交叉熵损失函数, 其计算公式为

$$\mathcal{L}_{\text{CE}} = - \sum_{i=1}^K y_i \log(\hat{y}_i). \quad (2)$$

其中: K 表示类别数; y_i 表示该样本是否属于第 i 个类别, 若是则为 1, 否则为 0; \hat{y}_i 表示模型对样本属于该类别的预测概率. 值得注意的是, 微调训练完成的模型即为最终可用的工业智能故障诊断模型, 即使该模型只经过了少量目标设备的有标签样本微调, 也未引入更多的有标签数据, 但由于模型在预训练阶段已获得从海量无标签数据中挖掘的诊断知识, 具有泛化特征提取能力, 也能具有较好的诊断精度.

2.2 用于时序依赖特征提取的掩码信号重构任务

掩码信号重构任务通过训练模型使其具有编码还原能力来使其获取关键时序依赖特征提取功能, 因为这些特征反映了样本由故障引起的时序变化模式, 只有模型能够较好地提取样本的关键时序依赖特征时, 才能够使模型很好地还原其原始样本. 该方法要求模型具有一个特征提取器网络 g_{θ} 与一个重构器网络 h_{θ_r} , 重构器网络即前文所述的代理任务头, 其输出维度与信号长度一致. 在掩码信号重构中, 将特征提取器与重构头视作一个非对称的自编码器. 首先对输入样本 \mathbf{x} 进行大规模掩码操作. 具体实现方式为: 将输入样本 \mathbf{x} 平均分割为 n 段子样本, 以掩码率 R 随机对 m 个子样本进行掩码, 即全部置 0, 其中 $m = R \times n$, 且 R 应大于 50%, 而分段个数 n 视传感器采样速率、样本长度与数据特性而酌情设定. 将掩码后的子样本与未被掩码的子样本按照原顺序进行拼接, 得到掩码样本 \mathbf{x}_m . 将掩码样本输入特征提取器 g_{θ} 得到隐变量 \mathbf{z} , \mathbf{z} 也可以称为表征或特征. 将隐变量 \mathbf{z} 输入重构器 h_{θ_r} , 得到重构样本 $\hat{\mathbf{x}}_m$. 整个过程可以用下式表示:

$$\hat{\mathbf{x}}_m = h_{\theta_r}(g_{\theta}(\mathbf{x}_m)). \quad (3)$$

然后使用均方误差损失函数对模型进行训练, 损失函数为

$$\mathcal{L}_{\text{MSE}}^r = \frac{1}{n} \sum_{i=1}^n (\mathbf{x} - \hat{\mathbf{x}}_m)^2, \quad (4)$$

其中 n 表示样本个数.

该任务以子片段为单位进行掩码, 可避免模型仅利用相邻样本点进行插值填补而实现重构. 结合大规模的掩码率, 使模型更倾向学习通过少量剩余未掩码样本提供的局部片段信息与上下文时序依赖信息对样本全局进行重构, 使得模型获取更加关键的时序依赖表征能力, 有利于其在少样本场景下的微调应用.

2.3 专家经验知识引导的先验特征预测任务

为了有效利用专家经验知识, 本文进一步设计了基于先验特征预测的自监督学习方法. 在该方法中, 将人工总结的统计特征作为专家先验知识的载体. 本文使用了工业故障诊断领域常用的 15 个时域统计特征作为先验特征^[8,20-21]. 具体如表 1 所示, 其中 $x(n)$ 表示时间序列, n 表示第 n 个样本点.

表 1 先验统计特征表

特征名称	表达式
均值	$\bar{x} = \frac{1}{N} \sum_{n=1}^N x(n)$
方差	$\delta = \frac{1}{N} \sum_{n=1}^N x_n^2$
方根幅值	$x_r = \left(\frac{1}{N} \sum_{n=1}^N \sqrt{ x(n) } \right)^2$
峰值	$x_p = \max x(n) $
最小值	$x_{\min} = \min(x_n)$
峰值指标	$C = \frac{x_p}{x_{\text{rms}}}$
裕度指标	$L = \frac{x_p}{x_r}$
峭度	$K = \frac{\sum_{n=1}^N [x(n) - \bar{x}]^4}{(N-1)\sigma_x^4}$
绝对平均值	$ \bar{x} = \frac{1}{N} \sum_{n=1}^N x_n $
均方根值	$x_{\text{rms}} = \sqrt{\frac{1}{N} \sum_{n=1}^N x^2(n)}$
标准差	$\sigma_x = \sqrt{\delta}$
最大值	$x_{\max} = \max(x_n)$
波形指标	$W = \frac{x_{\text{rms}}}{\bar{x}}$
脉冲指标	$I = \frac{x_p}{\bar{x}}$
偏斜度	$S = \frac{\sum_{n=1}^N [x(n) - \bar{x}]^3}{(N-1)\sigma_x^3}$

对于一个预训练样本 \mathbf{x} , 利用计算机计算该样本的这 15 个先验统计特征, 并将这 15 个先验特征拼接为先验特征向量 \mathbf{y}_p , 作为预训练样本在先验特征任务下的代理标签. 先验特征预测任务与掩码信号重

构任务共享了特征提取器网络 g_θ ,并新添了一个预测器网络 h_{θ_p} ,预测器网络即前文所述的代理任务头,其输出维度与先验特征向量长度一致.同样对预训练样本按照掩码信号重构任务下的掩码操作对预训练样本 \mathbf{x} 进行掩码,得到掩码样本 \mathbf{x}_m ,掩码的目的是迫使模型学会如何利用剩余未掩码片段所提供的局部信息和上下文信息来预测信号特征,以强化先验特征的引导作用.

将掩码样本输入特征提取器 g_θ ,得到隐变量 \mathbf{z} , \mathbf{z} 也可以称为表征或特征.将隐变量 \mathbf{z} 输入预测器 h_{θ_p} ,得到预测向量 $\hat{\mathbf{y}}_p$.整个过程可以用下式表示:

$$\hat{\mathbf{y}}_p = h_{\theta_p}(g_\theta(\mathbf{x}_m)). \quad (5)$$

同样使用均方误差损失函数对模型进行训练,损失函数为

$$\mathcal{L}_{\text{MSE}}^p = \frac{1}{n} \sum_{i=1}^n (\mathbf{y}_p - \hat{\mathbf{y}}_p)^2. \quad (6)$$

通过训练使得模型输出的预测特征逼近具有专家诊断知识的先验统计特征,可以实现利用知识引导模型特征提取过程,使其具有泛化性能.同时,由于这些专家知识具有一定的机理基础,还可以提高模型特征提取的可靠性.

本文结合了上述两种自监督学习任务的优势,使用两种代理任务联合训练网络的特征提取器,以尽可能使其具有泛化特征提取的能力.在预训练阶段最终的优化目标为

$$\hat{\theta}_e, \Theta = \arg \min_{\theta_e, \Theta} [\lambda \mathcal{L}_{\text{MSE}}^r + (1 - \lambda) \mathcal{L}_{\text{MSE}}^p]. \quad (7)$$

其中: λ 为超参数,用于调节两个任务对模型训练的影响程度,本文设 $\lambda = 0.5$.

2.4 智能故障诊断模型结构

本文所采用的智能故障诊断模型由一个特征提取器与一个分类头构成,同时在预训练阶段还使用了多个代理任务头来代替分类头以实现自监督预训练.本文有2个代理任务头,即2.2节所述的重构器与2.3节所述的预测器.

对于特征提取器,采用WDCNN^[22]作为特征提取器网络,这是一个第1层卷积核尺寸较大的卷积神经网络,该网络已成为多种智能故障诊断算法中的基础模型^[22-23].针对工业传感器数据,尤其是高频采样的振动信号,第1层卷积核尺寸较大的卷积层可以较好地抵抗工业场景中的高频噪声干扰,为模型特征提取提供更好的鲁棒性^[22].在每个卷积层后均连接了一个最大池化层,池化长度为2,池化后均采用激活函数进行非线性激活.激活函数采用高斯误差线性单

元(Gaussian error linear unit, GELU)进行激活.

3 案例验证

本节将以XJTU-SY滚动轴承加速寿命试验数据集^[24]作为预训练集,以CWRU滚动轴承故障诊断数据集^[25]作为目标集,验证本文所提出方法的有效性.

3.1 数据集介绍与处理

本节将对所使用的预训练集与目标集进行详细的描述,并介绍将其用于本文实验的数据处理过程.

3.1.1 预训练集

本文以XJTU-SY滚动轴承加速寿命试验数据集作为预训练集.该数据集中的轴承故障由加速寿命实验制作,采集滚动轴承从开始实验到轴承劣化至严重故障时的表面振动信号作为全寿命周期数据,实验轴承型号均为LDK UER204.

该数据集中滚动轴承的全寿命周期数据可以模拟与目标设备相似的设备在服役过程中历史积累的海量无标签数据,因此使用该数据集制作为预训练集.本文选择XJTU-SY数据集中同工况下5个轴承的全寿命周期数据进行处理,制作为预训练集,所用轴承的故障状态包括不同受损程度下的轴承外圈故障、内圈故障与保持架故障.

3.1.2 目标集

微调集与测试集均采集自目标设备,因此统称为目标集.本文所有目标集均通过凯斯西储大学(case western reserve university, CWRU)滚动轴承数据中心开源的滚动轴承故障数据集制作.该中心通过对测试台上的滚动轴承施加电火花来进行故障植入,分别在滚动轴承的内圈、外圈以及滚动体表面施加了损伤,损伤直径均包括7 mil、14 mil以及21 mil三种.依据不同损伤位置与损伤直径进行划分^[26],该数据共提供了9种故障状态的数据与正常状态的数据,基于CWRU数据集的故障诊断任务可以视作一个10分类任务^[22].本文所使用的数据由加速度计采集,数据采集频率为12 kHz,受损轴承型号为SKF6205.

预训练集与目标集所对应的设备皆为滚动轴承,但是其运行工况、轴承型号、传感器型号、传感器采样速率、装配方式、采样位置以及故障成因均不相同;并且预训练集中缺少轴承滚动体故障的样本,相比目标集多了保持架故障的样本,因此更加符合历史积累数据无标签情况下可能导致的预训练集与目标集标签空间不一致的场景.

3.1.3 数据集处理

对于预训练集,XJTU-SY数据集在轴承的X轴与Y轴方向同时采集其振动信号,将每间隔1 min采

样0.1 s的信号作为一个样本. 本文对其每个 x 轴与 y 轴方向的原始样本随机采样2个预训练样本, 共获取到了62 640个预训练样本, 每个样本长度为2 048个数据点, 对其进行标准化处理后依据2.3节所展示的先验特征预测方法, 计算预训练样本的先验统计特征, 作为预训练样本在预训练阶段的代理标签. 对这些计算得到的先验统计特征也进行标准化处理, 在分段掩码操作中, 分段个数为32.

而对于目标集, 在CWRU数据中, 对9类故障状态样本与1类正常状态样本, 每类分别随机采样50个样本用于训练, 150个样本用于测试, 并且训练样本与故障样本之间不存在重叠, 共有500个训练样本与1 500个测试样本. 在后续的测试中, 更少的每类训练样本均从这500个训练样本中采样得到. 对于目标集的样本, 不进行标准化或归一化处理, 样本长度与预训练样本长度一致.

3.2 训练设置

本文首先利用基于掩码信号重构-先验特征预测联合知识自监督学习任务对智能故障诊断模型进行预训练. 预训练阶段采用的优化器为Adam, 学习率设置为0.000 1, 训练迭代轮次为600. 而预训练完成后, 利用目标集中的训练集对模型进行微调. 微调过程中, 优化器选择为Adam, 对特征提取器参数的学习率为0.001, 对分类头参数的学习率为0.01, 训练轮数为20. 为避免初始化参数的随机性对实验结果的影响, 采用10个不同的随机数种子重复实验, 采取平均结果来评估诊断性能. 本实验使用的硬件训练平台为NVIDIA GeForce RTX 3090, 操作系统为Ubuntu20. 04, 所使用的深度学习框架为Pytorch.

3.3 结果分析

采用分类准确率来评估模型的实验中模型的性能, 即

$$\text{Accuracy} = \frac{\text{分类正确样本数}}{\text{测试样本总数}} \times 100\%. \quad (8)$$

本文采用不同样本数量的微调集对模型进行微调测试, 对比加入本文所提出的知识自监督学习方法前后的工业智能故障诊断模型的诊断性能, 以验证所提出方法的效果. 用BASE表示未经过知识自监督学习方法预训练的模型, 该模型仅通过微调极少量样本端到端训练得到. 用BASE+SS(R)表示经过知识自监督学习方法预训练后的模型, 其中 R 表示掩码率. 同时本文还将所提出方法与当下较为流行的少样本故障诊断方法进行了性能对比, 以展示所提出方法的优越性. 对比方法包括基于度量的元学

习的ProtoNet^[27]、RelationNet^[28]以及优化的元学习的MAML^[16]. 在进行对比实验时, 对比方法中所用到的特征提取器模型与所提出方法实验中设计的特征提取器模型结构与参数一致, 并且在CWRU数据集中额外采样了3 HP工况下每类100个含故障标签的样本支撑对比方法的元训练; 而在元测试阶段使用了与所提出方法一致的微调集与测试集来微调模型并测试验证性能. 在对比方法的元训练阶段, 采样了100个元训练任务进行元训练, 训练轮次为400, 同时采用了与所提出方法实验一致的10个随机数种子进行重复实验, 并取平均结果作为诊断性能.

图2展示了在不同有标签样本数训练下以所提出方法为基础模型带来的性能提升效果, 具体的各类方法性能指标如表2所示. 可以看出, 如果智能故障诊断模型仅利用少量的有标签样本进行训练, 则其在测试阶段精度将会呈现出诊断性能极差的情况. 而在经过本文所提出的知识自监督表征学习方法预训练后, 尤其是在超大规模掩码率($R = 87.5\%$)的预训练设置下, 在5-shot实验中, 预训练后的模型相较于端到端训练的基础模型, 其预训练后诊断准确率可以提高20.63%; 在10-shot实验中, 预训练后模型平均诊断准确率可达95.9%; 而当每类有标签故障样本数超过20时, 诊断准确率即可接近98%. 这验证了工业智能故障诊断模型在所设计的大规模掩码重构学习与专家先验知识的引导下, 实现了从海量无标签数据中挖掘诊断知识, 并以模型参数为载体将知识迁移至目标少样本场景, 成功实现少样本高精度故障诊断. 而经典的元学习方法在元训练阶段已经利用来自目标设备的样本及其故障类别标签进行学习, 在少样本场景下也表现出了较好的性能, 但是当样本数量规模增大时, 如每类含标签故障样本在30以上时, 诊断性能劣于所提出方法. 而本文所提出方法可以在各种样本数下均为基础模型带来较大的性能提升. 另一方面,

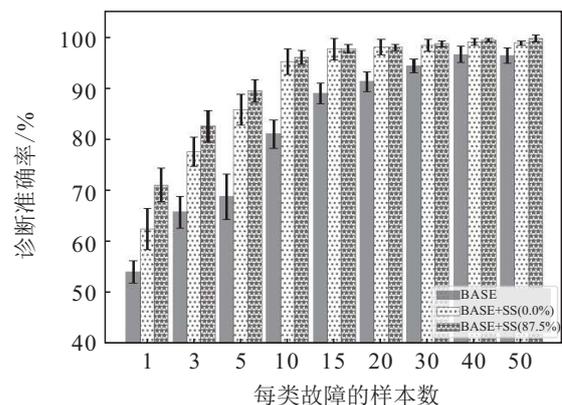


图2 不同有标签样本数训练下的模型性能

表 2 不同训练样本数下模型故障诊断准确率

模型	K-shot								
	1-shot	3-shot	5-shot	10-shot	15-shot	20-shot	30-shot	40-shot	50-shot
Base	53.98	65.69	68.71	80.93	88.92	91.16	94.22	96.5	96.25
ProtoNet ^[27]	86.00	93.87	96.44	97.56	97.93	98.24	98.49	98.59	98.80
RelationNet ^[28]	81.52	73.71	84.75	89.24	85.74	87.83	89.92	91.92	89.15
MAML ^[16]	72.00	91.33	72.4	96.2	95.47	95.20	96.6	96.50	96.98
Base + PF	69.34	81.63	83.05	89.53	94.96	97.22	98.67	98.9	99.30
Base + AEISS	64.66	74.42	75.59	84.58	89.08	90.01	92.04	93.15	94.66
Base + PFSS	75.66	84.22	86.39	94.60	95.34	97.64	98.22	98	98.86
Base + SS (0.0%)	62.35	77.47	85.69	95.14	97.59	97.92	98.23	98.91	98.73
Base + SS (87.5%)	71.01	82.49	89.34	95.9	97.59	97.92	98.59	99.31	99.60

本文未引入额外的有标签故障样本,更进一步降低了少样本学习对有标签故障样本的依赖,提高了实用性。

为了验证本文所设计的两个代理任务在自监督深度表征学习中的贡献,本文还进行了消融实验,实验结果在表 2 中列出。其中 Base + PF 表示将先验统计特征与特征提取器提取到的特征并行输入分类器以实现故障诊断,特征提取器与分类器均通过目标集的少量样本及其标签端到端训练得到,Base + AEISS 表示仅经过掩码信号重构任务预训练后的模型,Base + PFSS 表示仅经过先验特征预测任务预训练后的模型,并且预训练时掩码率 R 均设置为 87.5%。通过分析表 2 所展示的实验结果可以得出:如果仅将先验统计特征与特征提取器提取到的特征并行输入分类器,则由于先验统计特征本身具有一定的泛化性,可以帮助故障诊断模型在少样本场景下得到性能提升,但是先验统计特征没有优化特征提取器使其提取更具有泛化性的特征。因此,经过特征提取器得到的特征在少样本场景下仍存在过拟合风险,Base + PF 的故障诊断性能逊于所提出方法。通过观察 Base + AEISS 与 Base + PFSS 的性能可以知道:两种预训练任务均可有效提升基础模型在少样本场景下的诊断性能,尤其在 1-shot 与 3-shot 的极端少样本场景下,Base + PFSS 表现出了优异的性能。而本文所提出的多任务联合训练的知识自监督深度表征学习方法,结合了两个任务各自的优势,即在先验知识的引导下捕捉故障相关的关键时序依赖特征,使故障诊断模型不仅在极端少样本场景下具有良好的性能,而且在 10-shot 及其以上的场景中也表现出了比仅经过单一预训练任务得到的模型更好的诊断性能。

图 3 展示了在 1-shot 与 3-shot 实验中,不同预训练掩码率情况下模型性能。当掩码率不断升高时,模型性能变化的整体趋势为波动上升。在掩码率较低时,低掩码率无法有效屏蔽冗余信息,使得模型可能

利用冗余信息来进行重构,缺少对关键时序依赖关系的捕捉。因此,存在预训练不稳定的情况,无法得到较优的预训练模型。而当掩码率达到 87.5% 时,模型诊断精度得到了大幅的提升。这表明在大规模掩码的情况下,可以掩蔽掉更多的冗余信息,使得模型在预训练时专注更加关键的故障信息,从而更好地发挥本文所提出知识自监督表征学习方法的性能。

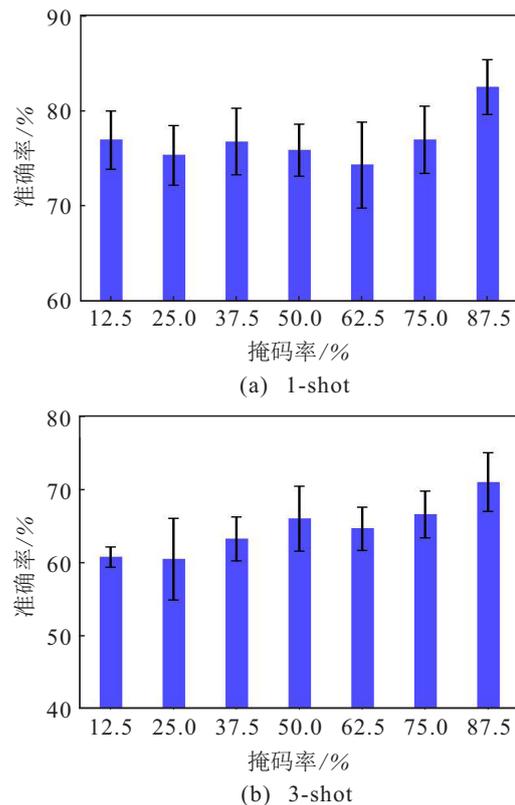


图 3 不同预训练掩码率情况下模型性能

为进一步验证本文的泛化特征提取效果,本文也利用 T-SNE 方法对未经过微调的预模型所提取到的特征的分布进行可视化,输入样本为测试集中的所有样本,可视化结果如图 4 所示,图中不同的颜色表示不同的故障标签。为了直观对比预训练模型的特征提取与经过端到端训练的基础模型的特征提取效果,图 4 还展示了在 30-shot 下 Base 方法的特征提取分布

可视化图. 观察图4可以得出, 经过本文所提出的知识自监督表征学习预训练后, 即使没有利用目标设备样本对特征提取器进行微调, 模型提取到的特征分布也已经显著呈现出类间分离的模式, 大部分类别的特征已经初步具备了可分性, 甚至部分类别样本的特征分离程度优于经过目标设备样本训练的Base模型中提取到的特征分布. 如Base模型中标签为1、2、7的样本特征存在混叠, 而在Base+SS (87.5%)模型中这3类样本的特征具有明显分离的模式, 这会有利于分类器进行故障分类, 也验证了本文所提出的知识自监督表征学习方法可以有效挖掘诊断知识, 提升智能故障诊断模型的泛化特征提取能力, 从而该模型在经过少样本微调后即可获得较高的故障诊断性能.

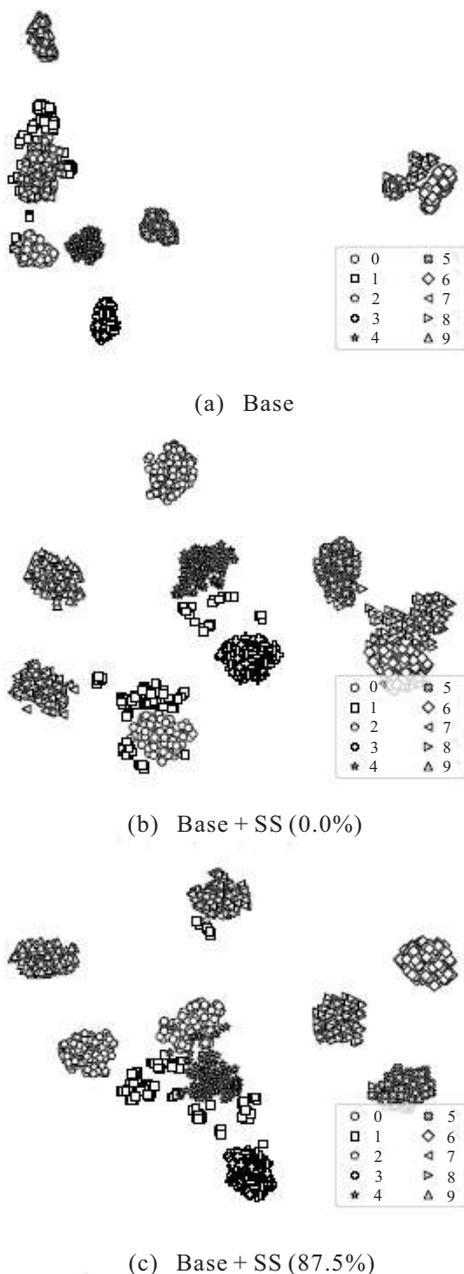


图4 模型特征提取分布可视化图

4 结论

本文通过挖掘关键时序依赖特征以及专家先验知识与故障诊断任务的内在关联性, 提出了一种无需额外标签数据的知识自监督深度表征学习方法, 克服了现有少样本故障诊断方法对额外有标签数据的依赖性挑战. 该方法利用历史积累的海量无标签数据对模型进行自监督训练, 设计了大规模掩码信号重构与先验特征预测联合预训练任务, 分别从关键时序依赖特征提取以及专家先验知识引导两个角度提升了模型特征提取的泛化性能. 通过在少量的目标场景的故障数据集上进行微调, 实现了不依赖额外有标签样本的少样本故障诊断. 本文通过从XJTU-SY数据集到CWRU数据集的跨数据集轴承故障诊断实验, 验证了本文所提出的方法在没有额外标签数据辅助的前提下, 能够有效提高少样本场景下的故障诊断精度. 在每类故障样本仅有5个的场景下, 本文所提出的自监督表征学习方法能够帮助原有基础模型提高20.63%的诊断精度, 当每类训练故障样本数量超过10时, 即可获得95%以上的故障诊断精度, 大大降低了智能故障诊断模型对有标签故障样本数量的要求.

参考文献(References)

- [1] 任浩, 屈剑锋, 柴毅, 等. 深度学习在故障诊断领域中的研究现状与挑战[J]. 控制与决策, 2017, 32(8): 1345-1358.
(Ren H, Qu J F, Chai Y, et al. Deep learning for fault diagnosis: The state of the art and challenge[J]. Control and Decision, 2017, 32(8): 1345-1358.)
- [2] Zhao C H, Chen J H, Jing H. Condition-driven data analytics and monitoring for wide-range nonstationary and transient continuous processes[J]. IEEE Transactions on Automation Science and Engineering, 2021, 18(4): 1563-1574.
- [3] 李晗, 萧德云. 基于数据驱动的故障诊断方法综述[J]. 控制与决策, 2011, 26(1): 1-9.
(Li H, Xiao D Y. Survey on data driven fault diagnosis methods[J]. Control and Decision, 2011, 26(1): 1-9.)
- [4] Song P Y, Zhao C H. Slow down to go better: A survey on slow feature analysis[J]. IEEE Transactions on Neural Networks and Learning Systems, DOI: 10.1109/TNNLS.2022.3201621.
- [5] 姚远, 佟佳蓉, 高军, 等. 基于分层分块DLPPCA-SVM的复杂工业过程监测与故障诊断方法[J]. 控制与决策, 2022, 37(5): 1402-1408.
(Yao Y, Tong J R, Gao J, et al. Monitoring and fault diagnosis method of complex industrial process based on DLPPCA-SVM[J]. Control and Decision, 2022, 37(5): 1402-1408.)
- [6] Song P Y, Zhao C H, Huang B. MPGE and RootRank: A sufficient root cause characterization and quantification

- framework for industrial process faults[J]. *Neural Networks*, 2023, 161: 397-417.
- [7] Chen X H, Zhang B K, Gao D. Bearing fault diagnosis base on multi-scale CNN and LSTM model[J]. *Journal of Intelligent Manufacturing*, 2021, 32(4): 971-987.
- [8] Yu W K, Zhao C H. Broad convolutional neural network based industrial process fault diagnosis with incremental learning capability[J]. *IEEE Transactions on Industrial Electronics*, 2020, 67(6): 5081-5091.
- [9] Lei Y G, Yang B, Jiang X W, et al. Applications of machine learning to machine fault diagnosis: A review and roadmap[J]. *Mechanical Systems and Signal Processing*, 2020, 138: 106587.
- [10] Feng L J, Zhao C H. Fault description based attribute transfer for zero-sample industrial fault diagnosis[J]. *IEEE Transactions on Industrial Informatics*, 2021, 17(3): 1852-1862.
- [11] Zhang T C, Chen J L, Li F D, et al. A small sample focused intelligent fault diagnosis scheme of machines via multimodules learning with gradient penalized generative adversarial networks[J]. *IEEE Transactions on Industrial Electronics*, 2021, 68(10): 10130-10141.
- [12] Dixit S, Verma N K. Intelligent condition-based monitoring of rotary machines with few samples[J]. *IEEE Sensors Journal*, 2020, 20(23): 14337-14346.
- [13] Xing S B, Lei Y G, Yang B, et al. Adaptive knowledge transfer by continual weighted updating of filter kernels for few-shot fault diagnosis of machines[J]. *IEEE Transactions on Industrial Electronics*, 2022, 69(2): 1968-1976.
- [14] Feng Y, Chen J L, Xie J S, et al. Meta-learning as a promising approach for few-shot cross-domain fault diagnosis: Algorithms, applications, and prospects[J]. *Knowledge-Based Systems*, 2022, 235: 107646.
- [15] Zhang A S, Li S B, Cui Y X, et al. Limited data rolling bearing fault diagnosis with few-shot learning[J]. *IEEE Access*, 2019, 7: 110895-110904.
- [16] Zhang S, Ye F, Wang B N, et al. Few-shot bearing fault diagnosis based on model-agnostic meta-learning[J]. *IEEE Transactions on Industry Applications*, 2021, 57(5): 4754-4764.
- [17] Zhang T C, Chen J L, He S L, et al. Prior knowledge-augmented self-supervised feature learning for few-shot intelligent fault diagnosis of machines[J]. *IEEE Transactions on Industrial Electronics*, 2022, 69(10): 10573-10584.
- [18] 雷亚国, 杨彬, 杜兆钧, 等. 大数据下机械装备故障的深度迁移诊断方法[J]. *机械工程学报*, 2019, 55(7): 1-8.
(Lei Y G, Yang B, Du Z J, et al. Deep transfer diagnosis method for machinery in big data era[J]. *Journal of Mechanical Engineering*, 2019, 55(7): 1-8.)
- [19] Ding Y, Ma L, Ma J, et al. A generative adversarial network-based intelligent fault diagnosis method for rotating machinery under small sample size conditions[J]. *IEEE Access*, 2019, 7: 149736-149749.
- [20] Lei Y G, He Z J, Zi Y Y, et al. Fault diagnosis of rotating machinery based on multiple ANFIS combination with GAs[J]. *Mechanical Systems and Signal Processing*, 2007, 21(5): 2280-2294.
- [21] Lei Y G, Zuo M J, He Z J, et al. A multidimensional hybrid intelligent method for gear fault diagnosis[J]. *Expert Systems with Applications*, 2010, 37(2): 1419-1430.
- [22] Zhang W, Peng G L, Li C H, et al. A new deep learning model for fault diagnosis with good anti-noise and domain adaptation ability on raw vibration signals[J]. *Sensors*, 2017, 17(2): 425.
- [23] Gao Y D, Kim C H, Kim J M. A novel hybrid deep learning method for fault diagnosis of rotating machinery based on extended WDCNN and long short-term memory[J]. *Sensors*, 2021, 21(19): 6614.
- [24] 雷亚国, 韩天宇, 王彪, 等. XJTU-SY滚动轴承加速寿命试验数据集解读[J]. *机械工程学报*, 2019, 55(16): 1-6.
(Lei Y G, Han T Y, Wang B, et al. XJTU-SY rolling element bearing accelerated life test datasets: A tutorial[J]. *Journal of Mechanical Engineering*, 2019, 55(16): 1-6.)
- [25] Hendriks J, Dumond P, Knox D A. Towards better benchmarking using the CWRU bearing fault dataset[J]. *Mechanical Systems and Signal Processing*, 2022, 169: 108732.
- [26] Chai Z, Zhao C H. Fault-prototypical adapted network for cross-domain industrial intelligent diagnosis[J]. *IEEE Transactions on Automation Science and Engineering*, 2022, 19(4): 3649-3658.
- [27] Snell J, Swersky K, Zemel R. Prototypical networks for few-shot learning[C]. *Proceedings of the 31st International Conference on Neural Information Processing Systems*. New York, 2017: 4080-4090.
- [28] Sung F, Yang Y X, Zhang L, et al. Learning to compare: Relation network for few-shot learning[C]. *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Salt Lake City, 2018: 1199-1208.

作者简介

姚家琪(1998—),男,硕士生,从事工业智能故障诊断、表征学习等研究, E-mail: yjiaqi@zju.edu.cn;

宋鹏宇(1998—),男,博士生,从事工业大数据分析、工业过程故障检测等研究, E-mail: pypsong@zju.edu.cn;

沈萌(2002—),女,本科生,从事变工况故障诊断、迁移学习等研究, E-mail: 3200103582@zju.edu.cn;

赵春晖(1979—),女,教授,博士生导师,从事工业大数据分析与应用(包括状态监测、故障诊断、软测量)等研究, E-mail: chhzha@zju.edu.cn;

王文海(1967—),男,教授,博士生导师,从事机器视觉、工业缺陷检测等研究, E-mail: zdzzlab@zju.edu.cn.