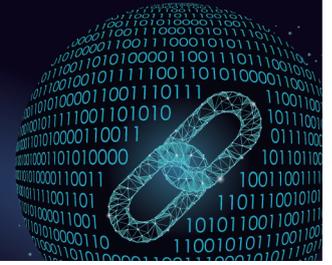




控制与决策

CONTROL AND DECISION



从零样本学习理论模型到工业应用——动机、演变与挑战

赵健程, 冯良骏, 岳嘉祺, 张堡霖, 赵春晖, 付永鹏, 王福利

引用本文:

赵健程,冯良骏,岳嘉祺,张堡霖,赵春晖,付永鹏,王福利. 从零样本学习理论模型到工业应用———动机、演变与挑战[J]. *控制与决策*, 2024, 39(9): 2833–2857.

在线阅读 View online: <https://doi.org/10.13195/j.kzyjc.2023.0924>

您可能感兴趣的其他文章

Articles you may be interested in

基于聚类簇结构特性的自适应综合采样法在入侵检测中的应用

Toward intrusion detection via cluster structure-based adaptive synthetic sampling approach

控制与决策. 2021, 36(8): 1920–1928 <https://doi.org/10.13195/j.kzyjc.2019.1672>

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>

基于改进卷积神经网络的动力下肢假肢运动意图识别

Intent recognition of power lower-limb prosthesis based on improved convolutional neural network

控制与决策. 2021, 36(12): 3031–3038 <https://doi.org/10.13195/j.kzyjc.2020.0326>

基于共享隐空间的多视角SVM

Multi view SVM based on common hidden space

控制与决策. 2021, 36(3): 534–542 <https://doi.org/10.13195/j.kzyjc.2019.0829>

不平衡数据分类方法综述

Review of imbalanced data classification methods

控制与决策. 2019, 34(4): 673–688 <https://doi.org/10.13195/j.kzyjc.2018.0865>

从零样本学习理论模型到工业应用 —— 动机、演变与挑战

赵健程^{1,2}, 冯良骏^{1,2}, 岳嘉祺^{1,2}, 张堡霖^{1,2}, 赵春晖^{1,2†}, 付永鹏^{1,2}, 王福利³

(1. 浙江大学 控制科学与工程学院, 杭州 310027; 2. 浙江大学 工业控制技术全国重点实验室, 杭州 310027; 3. 东北大学 信息科学与工程学院, 沈阳 110819)

摘要: 随着工业大数据技术的发展,应用于工业对象的有监督方法得到广泛研究. 真实数据往往遵循长尾类分布,导致传统有监督模型在实际应用过程中存在模型退化以及模型失效等问题. 零样本学习 (zero-shot learning, ZSL) 技术的提出为这一问题提供了一种新的解决思路. 零样本学习的目标是使用收集到的已见类别数据训练模型,使得训练好的模型对于收集不到数据的未见类别同样适用. 零样本学习通过将故障的文本描述等辅助知识引入到模型中,一定程度上缓解了模型在实际工业场景中对于训练数据收集的依赖,提高了模型的泛化性能. 然而,目前领域内尚缺乏对零样本学习在工业领域应用的系统梳理与讨论,而工业零样本学习在辅助知识的收集和处理、研究方法、应用场景等方面与其他领域的零样本学习相比也具有独特性. 鉴于零样本学习在工业领域潜在的巨大应用价值和未来的发展潜力,系统性梳理和展示了从零样本学习理论模型到工业应用的动机、演变与挑战. 首先,回顾零样本学习设定与相关方法的发展脉络,分析零样本学习与其他任务设定之间的关联,并指出本文与前人综述的区别. 接下来,回顾工业领域零样本学习的研究现状,介绍典型的工业零样本学习任务 and 辅助知识,分析工业零样本学习的特征和典型问题,梳理工业领域零样本任务中应用的现有方法. 此外,梳理工业零样本任务的基准数据集和开源工作. 最后,在现有研究的基础上总结工业零样本任务面临的问题与挑战,并对该领域的研究做出展望.

关键词: 零样本学习; 迁移学习; 工业人工智能; 机器学习; 故障诊断; 缺陷检测

中图分类号: TP277 文献标志码: A

DOI: 10.13195/j.kzyjc.2023.0924

引用格式: 赵健程,冯良骏,岳嘉祺,等. 从零样本学习理论模型到工业应用 —— 动机、演变与挑战[J]. 控制与决策, 2024, 39(9): 2833-2857.

From zero-shot learning theoretical model to its industrial application: Motivation, evolution and challenges

ZHAO Jian-cheng^{1,2}, FENG Liang-jun^{1,2}, YUE Jia-qi^{1,2}, ZHANG Bao-lin^{1,2}, ZHAO Chun-hui^{1,2†}, FU Yong-peng^{1,2}, WANG Fu-li³

(1. College of Control Science and Engineering, Zhejiang University, Hangzhou 310027, China; 2. State Key Laboratory of Industrial Control Technology, Zhejiang University, Hangzhou 310027, China; 3. College of Information Science and Engineering, Northeastern University, Shenyang 110819, China)

Abstract: With the advancement of industrial big data technology, supervised methods applied to industrial objects have been widely studied. However, real data often follow the long-tailed distribution phenomenon, which poses problems such as model degradation and failure in practical applications for traditional supervised models. The proposal of zero-shot learning (ZSL) technology provides a new approach to solve this problem. The objective of ZSL is to train the model using collected seen category data so that it can also be applied to unseen categories whose data cannot be collected. By incorporating auxiliary knowledge such as fault text description into the model, ZSL reduces the dependence of the model on training data collection in practical industrial scenarios and enhances its generalization performance. However, there is still a lack of systematic review and discussion on the application of ZSL in the industrial field. Compared with ZSL in other fields, industrial ZSL is unique in terms of auxiliary knowledge collection and processing, research methods, and application scenarios. Given the potential great application value and future development potential of ZSL in the industrial field, this paper systematically summarizes and presents the motivation, evolution, and challenges of ZSL theoretical

收稿日期: 2023-07-01; 录用日期: 2023-09-14.

基金项目: 国家自然科学基金杰出青年项目(62125306); 国家自然科学基金重点项目(62133003); 浙江大学工业控制技术全国重点实验室开放课题项目(ICT2023B01).

†通讯作者. E-mail: chhzhao@zju.edu.cn.

models for industrial applications. Firstly, this paper reviews the development of ZSL settings and related methods, analyzes their correlation with other task settings, and highlights differences between this paper's review and previous ones. Next, this paper reviews the current state of zero-shot learning research in the industrial field, introduces typical industrial zero-shot learning tasks and auxiliary knowledge, analyzes the features and typical problems of industrial zero-shot learning, and summarizes the existing methods used in industrial zero-shot tasks. In addition, this paper also presents the benchmark datasets and open-source works for industrial zero-shot tasks. Finally, based on the existing research, this paper summarizes the problems and challenges faced by industrial zero-shot tasks, and provides some prospects for the research in this field.

Keywords: zero-shot learning; transfer learning; industrial artificial intelligence; machine learning; fault diagnosis; defect inspection

0 引言

随着工业4.0时代的来临,学术界和产业界面向智能制造这一国家战略广泛开展了关于工业人工智能的深入研究与探索^[1]. 根据原理的不同,工业领域的人工智能可以大致划分为3类方法:基于机理的方法、基于知识的方法和基于数据驱动的方法. 传统的基于解析模型与专家先验知识的建模方法依赖于人工分析,且未必能够精准有效地应对复杂多变的实际工业对象. 随着物联网技术和大数据处理方法的快速发展,越来越多的工业数据能够被收集和和处理^[2],面向工业对象的数据驱动方法得到了更多的重视,在诸多实际场景得到了广泛的研究与应用,例如过程监测^[3-5]、故障诊断^[6-8]、软测量^[9-11]、缺陷检测^[12]、根因识别^[13]等. 传统的数据驱动方法往往使用有监督学习范式,即从标签化训练数据中推断出从输入到输出映射函数的机器学习范式. 自2012年在图像挑战赛上提出 AlexNet^[14]起,有监督学习范式开始蓬勃发展,目前基于有监督学习范式的方法在 ImageNet1K 物体识别任务^[15]中的 Top-5 准确率可达 98.7%. 有监督学习范式通常假设训练数据和测试数据是独立同分布的^[16]. 然而,一方面,在真实的工业场景中,由于生产条件的改变、外部干扰、设备老化等复杂因素^[17-19],可能导致训练与测试数据间具有显著分布差异;另一方面,真实的数据分布往往具有长尾效应,部分场景下某些类别的样本数目稀少甚至完全缺失^[20-21]. 在数据欠缺的场景下,有监督学习范式存在性能下降甚至无法建模的问题.

不同于有监督学习范式,本文关注一种常见却富有挑战性的学习任务,即尝试学习某些完全没有收集到训练数据的目标类别,这种学习范式称为零样本学习^[22]. 零样本学习旨在解决传统机器学习模型无法识别未见类别的问题,从而降低机器学习对数据标注的依赖. 在零样本学习任务中,已经收集到数据的类别称为已见类别;相应地,没有收集到数据的目

标类别称为未见类别;知识则指领域专家提供的辅助信息. 在很多实际场景下,数据收集成本高昂,难以实现,从而造成数据缺失的情况. 零样本学习通过引入辅助知识作为桥梁,将已见类别和未见类别联系起来,从而使模型具有识别未见类别的能力. 例如,对于机器视觉领域的零样本图片分类任务^[23-24],可通过人工标注的图像的属性(马的形状、老虎的斑纹、熊猫的黑白颜色),以及已见类别训练样本(马、老虎、熊猫等)的学习,对未见类别的图像(斑马,具有马形、斑纹、黑白颜色)进行分类. 这一过程模拟了人类基于过去学习的知识对新事物进行分析、想象、判别的过程,一定程度上缓解了数据驱动方法对于训练数据的强依赖. 最近,大语言模型引发了人们的广泛关注,部分大语言模型同样引入了人类定义的知识以改善模型性能. 例如,百度公司构建的文心一言模型使用了知识内化和知识外用等技术,以提高模型的训练效率和回答准确度,这也说明在模型的学习和推理过程中引入知识具有重要价值.

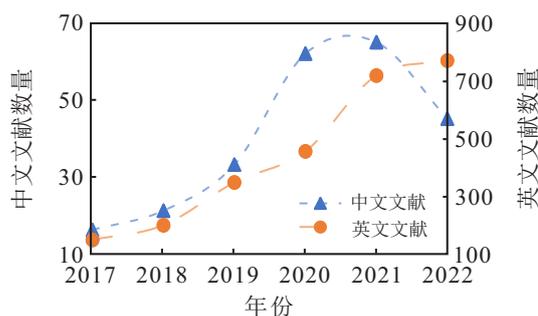


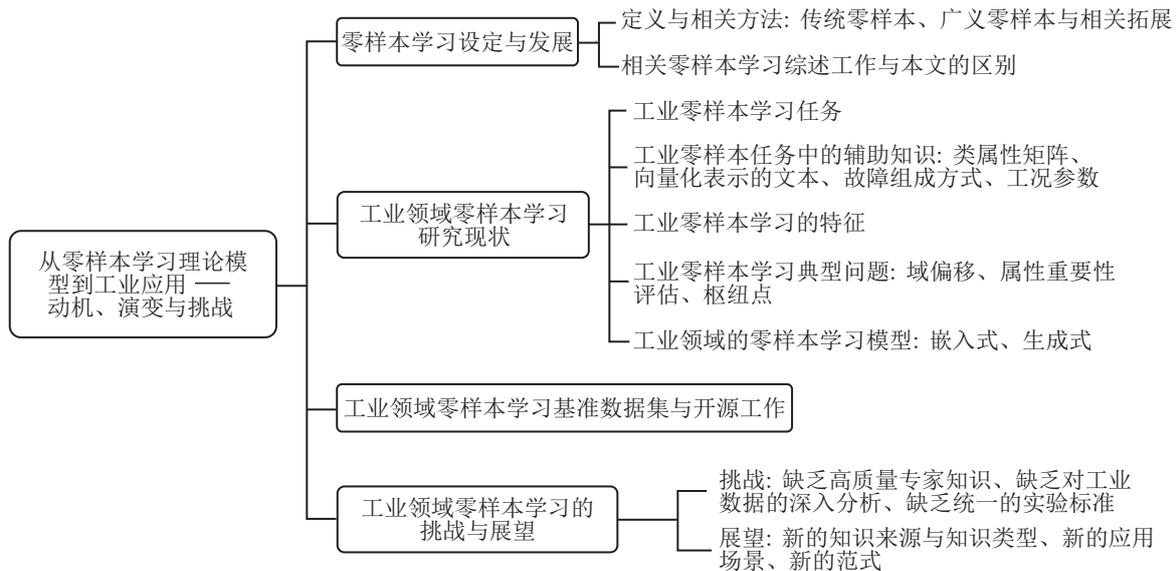
图1 本领域近6年文献数量发展趋势

近5年来,零样本学习任务因其重要的潜在应用价值得到了许多学者的关注. 在中国知网、Web of Science 数据库中分别以“零样本学习”“zero-shot” or “zero-sample”为关键词进行中、英文文献检索,得到相关领域各年份已发表文章数量的发展趋势. 如图1所示,2017年~2022年的6年中,本领域的相关研究得到了高速发展. 其中,每年发表的中文文献数量从2017年的16篇增长到2022年的45篇,英文文献数量

从145篇增长到769篇,这表明了相关学者对本领域的关注程度不断上升.近4年来,由于零样本学习有助于提升模型在数据类别长尾分布情况下的泛化能力,其在工业领域也开始得到一定程度的关注.目前,工业零样本的研究主要面向故障诊断^[22, 25-26]和表面缺陷检测^[12, 27]两方面应用.这两者本质上均为分类任务,通过自动辨识出故障或者缺陷的类别为后续的故障处理与设备或产品维护工作提供参考.在实际工业场景中,不同类别数据不平衡的问题广泛存在,特别是对于这两类任务而言,由于实际工业设备往往长期处于正常状态,且受限与收集、存储成本以及安全因素的考虑,很难收集到充足类别的故障或缺陷数据用于训练.因此,能够降低对数据收集依赖程度的零样本相关研究在工业领域具有重要的潜在应用价值.一些学者对零样本学习任务的设定、方法、应用进行了总结与回顾^[28-30],然而大多侧重于机器视觉、自然语言处理等热门领域,几乎没有涉及到零样本学习在工业领域的应用这一新兴但是十分有价值的场景.当前的工业零样本学习研究领域也缺乏对任务、

方法、问题、数据集、应用前景相关的系统性梳理与总结.因此,本文面向零样本学习的工业应用展开综述,在梳理现有工业零样本领域研究、总结其发展过程的同时,尝试指出本领域未来的发展方向,希望对后续的工业领域的零样本研究有所启发,并吸引更多的学者和工程技术人员关注这一研究方向.

本文结构如图2所示.首先介绍不同设定下零样本学习任务的发展过程与方法,包括传统零样本学习任务与广义零样本任务,并分析零样本学习与直推式学习、迁移学习、异常检测、开集分类间的联系;在此基础上,介绍工业零样本学习领域目前的两种典型任务,梳理类-属性矩阵等4种零样本学习过程使用的辅助知识,从任务、数据等角度分析工业领域零样本学习的特征和典型问题,并介绍现有的两大类工业零样本学习模型;接着,为推动本领域进一步发展,梳理目前本领域的基准数据集和开源工作;最后从专家知识、数据、标准3个方面分析本领域面临的挑战,并从知识、场景、方法3个角度指出零样本学习在工业领域潜在的研究方向.



1 零样本学习设定与发展

本节主要介绍零样本学习理论与任务设定的发展,并总结目前该领域已发表的综述类文章,基于此对比说明本文的贡献.

1.1 零样本学习定义与相关方法

零样本学习方法通过学习已见类别数据与其相应的辅助知识信息的关联,将从训练样本中学习到的知识由已见类别迁移到未见类别,从而实现对未见类别样本的判别.根据测试类别是否包含来自已见类别的样本,零样本学习任务的设定可被划分为传统

零样本学习和广义零样本学习 (generalized zero-shot learning, GZSL). 下文将对这两种设定进行具体介绍,本节所使用符号如表1所示.

1.1.1 传统零样本学习

在传统零样本学习^[31]任务中,训练集中样本的类别称为已见类别,记为 $\mathcal{S} = \{c_i^s | i = 1, 2, \dots, N_s\}$, 其中 N_s 为训练类别数. 训练集记为 $\mathcal{D}^t = \{\mathbf{x}^t \in \mathcal{X}_s, \mathbf{y}^t \in \mathcal{Y}_s\}$. 对于传统零样本学习任务,其训练集全部来自已见类别 \mathcal{S} , 测试集仅来自未见类别 $\mathcal{U} = \{c_i^u | i = 1, 2, \dots, N_u\}$, 其中 N_u 为测试类别数. 测试集

表1 本节符号说明

设定	符号	定义
零样本学习	\mathcal{S}	已见类别集合
	\mathcal{U}	未见类别集合
	\mathcal{A}	所有类别属性描述集合
	\mathcal{X}_s	已见类别样本集合
	\mathcal{Y}_s	已见类别标签集合
	\mathcal{X}_u	未见类别样本集合
	\mathcal{Y}_u	未见类别标签集合
	$\mathcal{X}_s^{\text{tr}}$	已见类别训练样本集合
	$\mathcal{Y}_s^{\text{tr}}$	已见类别训练标签集合
	$\mathcal{X}_s^{\text{te}}$	已见类别测试样本集合
$\mathcal{Y}_s^{\text{te}}$	已见类别测试标签集合	
迁移学习	\mathcal{X}^{sr}	源域样本集合
	\mathcal{Y}^{sr}	源域标签集合
	\mathcal{X}^{t}	目标域样本集合
	\mathcal{Y}^{t}	目标域标签集合

表示为 $\mathcal{D}^{\text{te}} = \{\mathbf{x}^{\text{te}} \in \mathcal{X}_u, \mathbf{y}^{\text{te}} \in \mathcal{Y}_u\}$. 已见类别与未见类别不重叠, 即 $\mathcal{S} \cap \mathcal{U} = \emptyset$. 零样本学习的目标是利用训练集 \mathcal{D}^{tr} 学习能够应用在测试集上的映射函数 $f(x) : \mathcal{X} \rightarrow \mathcal{U}$, 使误差 δ 最小, 即

$$f^* = \arg \min_f E_{(\mathbf{x}, \mathbf{y}) \in \mathcal{D}^{\text{te}}} \delta(f(\mathbf{x}), \mathbf{y}). \quad (1)$$

知识是零样本学习方法能够在已见域与未见域之间完成迁移的关键因素. 一方面, 知识应当具有一定的信息量, 且应该与相应的数据之间具有充分的相关性^[32]; 另一方面, 知识应具有一定范围内的通用性和可迁移性, 代表已见类别与未见类别间的共性信息. 在图像分类任务中, 知识常常表现为图片中物体的属性标注^[23, 33], 如大小、颜色等. 在属性空间中, 类别的属性记为 $\mathbf{a} \in \mathbb{R}^C$, C 为属性数量. 训练集的属性记为 $\mathcal{A}_s \in \mathbb{R}^{N_s \times C}$, 测试集的属性记为 $\mathcal{A}_u \in \mathbb{R}^{N_u \times C}$, 所有类别的属性集合为 $\mathcal{A} = \mathcal{A}_s \cup \mathcal{A}_u$.

对于传统零样本学习任务, 嵌入式方法是一种经典有效的解决方法. 嵌入式方法一般会利用训练集学习一个映射函数 $m : \mathcal{X} \rightarrow \mathcal{A}$, 该函数的作用是将原始数据或原始特征映射至语义空间. 由于语义空间内的属性信息由训练集和测试集共同拥有, 在测试阶段可以基于该映射函数预测出的属性通过最近邻等方法为测试样本匹配类别标签. 自2009年起, 学界发展了诸多嵌入式零样本学习方法, 其中心思想是设计不同的嵌入函数. 例如, Lampert等^[23]提出了直接属性迁移方法, 其嵌入函数是在已见类别上训练的属性判别器, 最早实现了零样本学习任务. Akata等^[34]、Frome等^[35]通过兼容度函数对数据特征空间和语义特征空间进行匹配, 从而实现对测试样本的分类. 相较于直接属性迁移方法, 兼容度函数方法迁移考虑属

性之间的相关关系, 提高了属性信息的利用率.

1.1.2 广义零样本学习

在传统零样本学习任务中, 只要求在已见类别数据上训练出的模型对于未见类别数据具有判别能力, 但在真实世界中往往会要求模型同时对已见类别样本和未见类别样本均具有识别能力^[36]. 因此, 广义零样本学习任务被提出^[33]. 对于广义零样本学习任务, 训练集和传统零样本任务的定义一致, 但是测试集中的类别同时来自已见类别 \mathcal{S} 和未见类别 \mathcal{U} . 广义零样本学习任务的测试集可记为 $\mathcal{D}^{\text{te}} = \{\mathbf{x}^{\text{te}} \in \mathcal{X}_u \cup \mathcal{X}_s^{\text{te}}, \mathbf{y}^{\text{te}} \in \mathcal{Y}_u \cup \mathcal{Y}_s^{\text{te}}\}$.

相较于传统零样本学习, 广义零样本学习的实用性更强, 但挑战性也更大. 目前, 广义零样本学习的相关方法主要分为嵌入式和生成式两种. 嵌入式方法在传统零样本学习中取得了较好的效果, 然而对于广义零样本学习, 嵌入式方法往往将未见类别测试样本映射为已见类别, 且过拟合现象严重, 效果不佳^[29]. 2018年后兴起的生成式方法^[37-38]以数据增强的手段一定程度上缓解了嵌入式方法存在的问题. 生成式方法旨在基于已见类别的样本和全部类别的语义知识, 生成未见类别样本, 从而将零样本学习任务转化为传统的有监督学习任务. 具体而言, 在训练阶段, 生成器使用已见类别的语义属性向量 \mathbf{a} 和高斯噪声 \mathbf{z} 生成伪造样本 $G : (\mathbf{a}, \mathbf{z}) \rightarrow \hat{\mathbf{x}}$, 并通过真实的已见类别样本评估生成质量, 从而训练生成器. 在测试阶段, 可以向生成器中输入未见类别的语义和高斯噪声, 即可获得大量未见类别的合成样本.

目前的生成式零样本学习工作大多基于生成式对抗网络 (generative adversarial networks, GAN) 或变分自编码器 (variational autoencoder, VAE). CVAE-ZSL (conditional variational autoencoders for zero-shot learning)^[37] 和 f-CLSWGAN (feature generating networks for zero-shot learning)^[38] 是生成式零样本学习的代表性方法, 分别利用变分自编码器、生成式对抗网络来生成未见类别的样本. f-CLSWGAN 利用在已见类别上预训练的分类网络帮助对抗生成网络生成可分性更高的虚拟特征, 从而获得比 CVAE-ZSL 方法更优的性能. Schönfeld等^[39]在 CVAE-ZSL 的基础上做出改进, 提出了 CADA-VAE (cross and distribution aligned variational autoencoder) 模型, 通过在变分自编码器的隐空间中将数据特征与语义特征对齐, 使模型能更好地完成语义知识的迁移. 相较于嵌入式方法, 生成式方法缓解了未见类别样本被误分类为已见类别的问题, 但其训练难度往往更大.

1.1.3 零样本学习的设定拓展

第1.1.1节和第1.1.2节对零样本学习的介绍均处于归纳式机器学习设定下,本小节首先介绍直推式设定下的零样本学习. 迁移学习与零样本学习的概念高度相关,本节也介绍了两者之间的联系与差异. 此外,介绍异常检测、开集分类、广义零样本分类3个相关任务之间的关系.

1) 零样本学习与直推式设定.

在归纳式训练设定中,使用已见类别的训练数据建立一个通用模型,对测试数据进行预测. 与归纳式设定不同,在直推式设定中,同时利用有标签的训练数据 and 无标签的测试数据来建立模型,因此直推式设定下得到的模型性能往往优于归纳式,但相对于归纳式而言应用场景受限. 在直推式设定下,传统与广义零样本学习的训练集变为 $\mathcal{D}^u = \{\mathbf{x}^u \in \mathcal{X}_s^u \cup \mathcal{X}_u, y^u \in \mathcal{Y}_s^u \cup \emptyset\}$, 测试集的定义与归纳式设定相同. 直推式零样本学习允许模型在训练时利用未见类别的无标签数据 \mathcal{X}_u 改善模型的迁移性能. 因此,直推式零样本学习可以被认为是半监督学习的特例,其与一般的半监督学习区别在于所利用的无标签样本只来自未见类别.

2) 零样本学习与迁移学习.

在迁移学习^[40-41]任务中,利用数据、任务或模型之间的相似性,通过源域与目标域之间知识的迁移使在源域训练而成的机器学习模型可以应用于目标域. 迁移学习的任务可以定义为:给定源域数据集 $\mathcal{D}^{sr} = \{\mathbf{x}^{sr} \in \mathcal{X}^{sr}, y^{sr} \in \mathcal{Y}^{sr}\}$ 和源域学习任务 T^{sr} , 以及目标域数据集 $\mathcal{D}^t = \{\mathbf{x}^t \in \mathcal{X}^t, y^t \in \mathcal{Y}^t\}$ 和目标域任务 T^t , 在 $\mathcal{D}^{sr} \neq \mathcal{D}^t$ 或 $T^{sr} \neq T^t$ 条件下,降低目标域预测模型 $f(\cdot)$ 的泛化误差. 目前已发表的迁移学习相关的研究可分为基于样本的迁移、基于特征的迁移、基于模型的迁移等^[42]. 零样本学习和迁移学习既有联

系又有不同. 从定义上而言,零样本学习可被认为是迁移学习的一种极端情况,即训练集(源域)与测试集(目标域)的样本类别完全不同. 由于零样本学习任务的特殊性,即训练集和测试集的标签空间不同,迁移学习的主流算法(如域自适应、域对抗)往往并不能直接适用于零样本学习任务.

3) 广义零样本分类与异常检测、开集分类.

异常检测任务、开集分类任务和广义零样本任务三者层层递进,分类粒度由粗至细. 广义零样本分类与异常检测、开集分类之间的关系如图3所示. 异常检测任务^[3,43]旨在利用正常数据训练模型,判断测试数据是否出现异常,是已见与未见之间的二分类任务;对于开集分类任务^[44-46],其对测试集中的已见类别样本给出具体的分类结果,识别出测试集中全部未见类别样本并归为一类;而广义零样本分类任务则在开集分类任务的基础上进一步要求对测试集中的未见类别样本,在辅助知识的帮助下实现细粒度的分类,即在开集分类的目标上增加了零样本类别分类的目标.

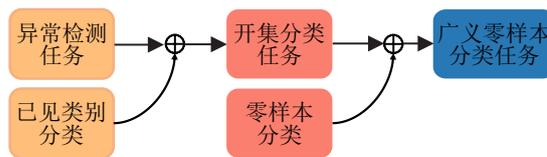


图3 异常检测、开集分类、广义零样本任务之间的联系

4) 小结.

本节介绍了零样本学习相关的多种任务设定,即迁移学习、传统零样本学习、广义零样本学习和直推式零样本学习. 表2对比了几种任务设定的训练集、测试集与任务难度. 对于零样本的各种变体而言,广义零样本学习难度最大,但其实用价值也最高;而直推式零样本学习由于提供了未见类别的无标签数据,其任务难度得到降低.

表2 各类零样本学习与迁移学习任务设定对比

任务设定	训练集	测试集	提出时间	任务难度
1 传统零样本学习 ^[23]	$\mathcal{X}_s^{tr} \cup \mathcal{Y}_s^{tr} \cup \mathcal{A}$	$\mathcal{X}_u \cup \mathcal{A}$	2009年	较难
2 广义零样本学习 ^[33]	$\mathcal{X}_s^{tr} \cup \mathcal{Y}_s^{tr} \cup \mathcal{A}$	$\mathcal{X}_u \cup \mathcal{X}_u^{te} \cup \mathcal{A}$	2018年	难
3 直推式零样本学习 ^[33]	$\mathcal{X}_s \cup \mathcal{Y}_s \cup \mathcal{X}_u \cup \mathcal{A}$	\mathcal{X}_u	2018年	易
4 迁移学习 ^[42]	$\mathcal{X}_s^{tr} \cup \mathcal{Y}_s^{tr}$	\mathcal{X}_t	1995年	—

1.2 相关综述工作与本文的区别

近年来,零样本学习任务得到了学者的广泛关注,一些学者对零样本学习设定、发展、知识、应用场景进行了梳理和总结,如表3所示. Wang等^[28]讨论了零样本学习中不同的语义空间,并基于构建方式的不同将其划分为工程化的语义空间(engineered semantic spaces)和学习到的语义空间

(learned semantic spaces)两类. 此外,基于在训练过程中是否使用未见类别的语义原型和样本,将零样本学习的设定划分为类归纳-样本归纳(class-inductive, instance-inductive)、类直推-样本归纳(class-transductive, instance-inductive)、类直推-样本直推(class-transductive, instance-transductive)3种类型. 冀中等^[47]以时间为线索,梳理了零样本学习近十年

来的各个发展阶段,同时在该研究中总结了零样本学习在机器视觉领域常用的数据集以及评价准则. Pourpanah 等^[29]首次聚焦于广义零样本学习范式,给出了广义零样本模型的分类,介绍了映射式和生成式两大范式中的代表性方法. 张鲁宁等^[48]较为详细地梳理了嵌入式零样本学习的相关分类,并且给出了各类模型在基准数据集下的实验结果. Cao 等^[49]讨论了零样本学习在除机器视觉外多种不同领域的应用,如自然语言处理、人类行为识别以及知识

图谱中的实体分类任务,并对零样本与其他技术(如强化学习)的结合进行了展望; Sun 等^[50]介绍了从充足样本、少样本到零样本学习的发展过程,并从视觉特征提取、语义的表示和视觉特征到语义的映射3个方面详细分析了零样本学习的关键技术; Chen 等^[51]介绍了零样本学习中应用的不同类型的辅助知识,如文本、属性、知识图谱以及实体与规则. 冯耀功等^[52]进一步梳理了零样本学习中“知识”这一概念所覆盖的范畴,共划分为初级知识、抽象知识和外部知识.

表3 零样本学习相关综述总结

作者	时间	任务设定	方法总结	知识类型	工业领域	贡献
Wang 等 ^[28]	2019年	✓	✓	×	×	1. 给出了零样本学习设定正式分类和定义 2. 对语义空间进行了详细划分
冀中等 ^[47]	2019年	✓	✓	✓	×	1. 以时间为线索梳理零样本学习十年发展 2. 总结了零样本学习常用的数据集和评价准则
Pourpanah 等 ^[29]	2020年	✓	✓	✓	×	1. 首次针对广义零样本学习进行了全面综述 2. 给出广义零样本模型的分类
张鲁宁等 ^[48]	2020年	✓	✓	✓	×	1. 着重介绍了嵌入式零样本方法的分类 2. 给出多个嵌入式模型在标准数据集下的结果
Cao 等 ^[49]	2020年	✓	×	×	×	1. 总结了零样本学习技术在除机器视觉外多个领域的应用 2. 对零样本学习和其他技术的结合进行展望
Sun 等 ^[50]	2021年	✓	✓	✓	×	1. 梳理了从充足样本、少样本再到零样本研究的演化过程 2. 从视觉特征提取、语义的表示和视觉特征到语义的映射3个方面进行详细分析
Chen 等 ^[51]	2021年	✓	✓	✓	×	1. 聚焦于目前零样本学习引入的辅助知识 2. 对各知识类型的特点和优缺点进行对比
冯耀功等 ^[52]	2021年	✓	✓	✓	×	1. 梳理了零样本领域中知识的范畴 2. 根据知识的分类总结零样本工作
本文	2023年	✓	✓	✓	✓	1. 首次基于实际的工业场景,探究工业零样本学习的任务设定以及方法 2. 指出现有工业零样本研究存在的不足,并且从知识、应用以及范式角度进行展望

然而,目前零样本学习的研究综述与应用大多集中在机器视觉以及自然语言处理领域,鲜有涉足工业领域. 本文第1次基于实际的工业场景,对工业零样本学习的特征、典型问题和现有模型进行详细阐述,重点在于工业领域零样本学习的应用. 详细介绍了工业零样本学习的两种典型任务、辅助知识的使用以及现有模型,以满足实际应用需求. 通过关注工业领域的具体问题和解决方案,本文对于那些在工业实践中寻求应用零样本学习的读者具有较为直接的实用价值. 在此基础上,提出零样本学习在工业领域潜在的研究方向,这些方向包括专家知识、数据和标准等方面的挑战,展现了本文对本领域新的见解,为读者提供进一步深入研究和启示,激发读者对于本领域更进一步探索的兴趣.

2 工业领域零样本学习研究现状

本节主要介绍工业领域零样本学习研究现状. 首先介绍工业场景下有代表性的零样本学习任务,即故障诊断和缺陷检测,然后介绍在工业零样本任务中使用的辅助知识. 同时,梳理工业领域零样本学习任务的特征和典型问题,并以此为中心介绍工业零样本学习任务所使用的模型.

2.1 工业零样本学习任务

在工业领域中,零样本学习任务与具体的应用场景紧密结合. 由于工业场景较为广泛,本小节将着重介绍工业零样本中两种最为典型的任务:故障诊断和缺陷检测. 这两类任务目前在工业零样本学习相关的研究中得到了较多的关注,无论在工业界还是在学术界都得到了广泛应用和较为深入的研究,积累的

专家经验知识较多,有助于零样本学习方法的实现.

故障诊断任务面向时序数据,是一种典型的分类任务^[53],其训练集中包含收集到的不同类别的故障数据,基于训练好的模型对在线发生的出现故障的样本进行分类,即诊断其所属的故障类别.对于零样本故障诊断任务而言,其零样本特性体现在训练集与测试集中的故障类别不一致;有些零样本故障诊断任务虽然训练集与测试集中类别一致,但故障发生时的工况参数(如系统负载、故障程度)不同^[31-32],即训练与测试时的运行工况不一致.此外,对于复合故障零样本诊断任务,其训练集中均为单个类型的故障,测试集中均为复合故障(两种及以上单故障同时发生).对于故障诊断任务而言,实际工业生产过程往往积累了许多由现场专工撰写的故障案例分析文档,为零样本方法的实施创造了条件.

缺陷检测任务面向图像数据,训练集中为不同类型的缺陷工件或产品的图像,如油污、磨损、划痕等,基于训练数据训练分类模型,对测试图像进行分类,判断其缺陷类别.由于不同的缺陷发生的频率不同,数据集中不同类别的样本数量分布往往极不均匀,类别收集不全面,不同类型的缺陷可能同时出现,且在实际过程中可能出现完全未见的类别.传统机器学习算法很难正确识别缺乏训练样本的类别,因此,零样本学习方法对于缺陷检测任务具有重要价值.零样本缺陷检测任务即训练集与测试集中的缺陷类别不同,需要借助辅助知识实现从已见类别到未见类别的迁移.

总体而言,故障诊断任务和缺陷检测任务本质上均为分类问题.事实上,分类问题也是零样本学习领域重点关注的一类问题.对于训练集中的样本 $x \in \mathcal{X}$ 及其对应的类别标签 $y \in \mathcal{Y}$,模型需要学习映射函数 $f: x_t \rightarrow y_t$,使其具有最小的分类损失 L ,如下所示:

$$L = \text{CLoss}(\mathcal{Y}, f(\mathcal{X})), \quad (2)$$

其中 CLoss 为任意的分类损失函数.工业零样本任务以领域专家提供的辅助知识为桥梁,使得所学习到的

映射函数能够从已见类别迁移到未见类别.

2.2 工业零样本任务中的辅助知识

在工业零样本学习任务中,知识的来源往往更加专业,高度依赖于工厂专工等工业领域专家的标注和相关的专业文档,如田纳西-伊斯曼工业过程对各个故障的介绍^[54].工业零样本任务中的知识可以被定义为能够在已见域与未见域之间迁移的、有助于故障分类或缺陷检测的信息,即提取零样本任务所应用的知识时,需要重点关注已见域与未见域之间的共同信息,而根据关注角度的不同,所利用的知识也各不相同.从单个类别的故障或缺陷而言,每种类型可通过专家标注获得其所具有的属性,且属性一般可以在不同类别之间共享.正因如此,类-属性矩阵^[22,31,55]是目前零样本任务中应用最为广泛的知识表示形式;从常用的类别描述的角度而言,已见域和未见域中不同类别的文字描述在语义上存在共通之处,因此使用文本向量化^[56]编码的类别描述也可以认为是工业零样本的一种知识来源.此外,面向具体的应用场景,往往需要针对性地设计辅助知识的类型.例如,对于复合型的分类任务(复合故障诊断,或多种缺陷同时发生),故障或缺陷的组成方式^[57-58]也可以作为一种知识,从而使得模型能够基于此对学到的单个类别信息进行组合,从而获得判断复合类别的能力.从运行机理的角度而言,设备的运行工况参数在已见域与未见域之间也存在共性,即两者可共享相同的运行参数,因此也可以基于工况参数^[31-32]完成工业零样本任务.通常而言,工业零样本任务大多都是分类任务,对未见类别的分类都需要基于辅助知识实现.此外,还有少数工业零样本任务不使用任何类型的辅助知识,这一类任务本质上是异常检测任务与聚类任务相结合^[25,27,59-60].由于没有辅助知识的辅助,在这些场景中,仅能区分出测试样本中属于未见类别的部分,或者进一步对测试样本进行聚类,但是无法将每个类别对应到具体的类别标签上,因为未见类别没有任何信息可供参考.总体而言,知识对于工业零样本学习

表4 不同形式的工业领域知识

知识类型	描述	知识来源	知识的获取难度
类-属性矩阵	使用布尔值或连续值描述类别属性,以矩阵形式表示	专家标注	难
向量化表示的文本	故障类别名称、描述等的向量化表示	故障记录及生产机理描述文档	较难 (通常需要预训练的语言模型)
故障组成方式	复合故障由哪些单故障组成	专家标注	容易
工况参数	不同类型故障的参数设置,如负载、工作频率等	工艺、工况设定	容易

任务是至关重要的.表4从知识描述、知识来源、知识获取难度3个角度对比了不同形式的工业领域知识.

2.2.1 类-属性矩阵

近年来,在工业零样本任务中,由于其知识表示形式高度精简且易于使用,类-属性矩阵这一知识表示形式近年来得到了一定程度的关注^[22,31,55].类-属性矩阵的构建过程如下:首先,基于预定义的类别和故障属性列表,由专家对每个类别进行标注,确定每个类别具有的属性;然后,将每个类别表示为一个行向量,同一行中的元素表示该类别是否具有相应的属性,对于全部类别便构成了类-属性矩阵.对于故障诊断任务,标注的属性如故障的位置、故障的原因、故障的类型等;对于缺陷检测任务,属性可以为缺陷的颜色、缺陷的部件、缺陷的原因等.类-属性矩阵提供了一种不同类别间基于属性的知识迁移方式,通过将不同类别具有的同种属性视作共性特征,对属性进行建模和辨识便能够实现不同类别间知识的迁移和辨识.从这个角度来讲,基于类-属性矩阵的零样本任务也可以看作是有监督的属性识别任务.任务的难点在于,一方面,训练集和测试集中不同属性的组合方式不同,即数据分布上的差异;另一方面,属性之间本身可能是高度相关与耦合的,在辨识属性的过程中不同的属性间存在一定程度的干扰.此外,任务的效果与专家的标注质量高度相关,专家标注的完整性与正确性极大地影响了任务的效果.

2.2.2 向量化表示的文本

在零样本学习任务中,文本的向量化表示也是在模型中引入知识的一种重要方式.通过将类别名称^[61]、类别的文字描述^[62]这一类非结构化信息编码为特征向量,可以将文本形式的知识引入到模型中,从而基于未见类别的文本知识实现对未见类别的表示和辨识.例如,在零样本气缸套表面缺陷检测任务中,Guo等^[56]使用GloVE^[63]模型对类别标签名称,如缸套裂纹、缸套磨损、缸套收缩,提取其相应的词向量构建语义空间,通过将语义空间中不同类别之间的相关性迁移到特征空间中实现对未见类别特征的生成,从而实现零样本缺陷检测.然而,文本向量化方法目前在工业零样本任务中应用很少,这可能是由于对于一般的图像零样本分类任务而言,有大量的开源语料可用于训练大规模自然语言模型.然而,这些开源语料库中包括的工业领域语料往往较少,特别是某些工业术语可能完全没有涉及,从而限制了文本向量化方法在工业领域的应用.随着工业信息化进程的推进,这一问题有望改善.

2.2.3 故障组成方式

故障组合方式是工业零样本故障诊断中一种特殊设计的知识类型,描述了每种复合故障由哪些单故障组成,在零样本复合故障故障诊断中具有重要的应用价值.零样本复合故障故障诊断任务中仅拥有多种类型的单故障数据,希望模型对复合故障(即多种单故障同时发生导致的故障)具有诊断能力,一方面,基于故障组成方式,可以使用单故障表征合成复合故障的表征^[58,64];另一方面,也可以将复合故障诊断问题转换为多标签分类问题加以解决^[57].除故障诊断任务外,这种类型的知识也可以应用到缺陷检测任务中,从而应对多种缺陷同时发生的情况.

2.2.4 工况参数

工况参数指的是设备在运行中设置的运行状态、负载、故障程度等以各种数值参数为表现形式的知识^[31-32].为便于后续处理,采取一些量化或标准化手段对工况参数值进行标准化.不同工况参数的组合方式对应不同类型的故障.在训练过程中,收集到全部类型的故障较为困难,通常只能收集到一部分工况参数下的故障数据,需要基于未见类别的工况参数生成未见类别的数据,从而实现未见类别的故障诊断.

2.3 工业零样本学习的特征

零样本学习起源于机器视觉领域的图像分类任务,大量的零样本学习相关研究围绕其展开^[29].然而,由于面向对象的不同,不能直接将机器视觉领域的零样本学习方法套用至工业领域中,即需要考虑工业零样本学习的特征.如表5所示,本节以对比的方式从典型任务、数据模态、数据收集方式、数据类别多样性、知识收集方式、知识类型出发,阐述工业零样本学习的6种典型特征.就典型任务而言,如前文所述,工业领域目前主要面向过程故障诊断和产品缺陷监测两方面任务.由于面向的研究领域不同,任务中面向的研究对象不同,工业对象和任务需求的个性化相叠加,导致所处理数据的模态、零样本任务中使用辅助知识的类型等方面存在诸多不同,体现了工业领域的特征,从而需要对工业领域中的零样本学习方法进行针对性设计,以更好地符合工业应用场景下的需求.

就数据的模态而言,工业领域和机器视觉领域面向的数据模态不同,机器视觉领域面向二维或三维的图像数据,其往往只存在空间上的相关性;而工业数据往往由大量部署在工业现场中的传感器测点收集得到,经由分布式控制系统(distributed control system, DCS)汇总到数据库中存储,其按时间顺序产生,带有时间戳,具有明显的时序特性^[65].相比于其

表5 工业领域零样本学习的特征

	机器视觉领域	工业领域	工业领域的特征
典型任务	图像分类	过程故障诊断、产品缺陷检测	工业对象与任务需求的个性化叠加
数据的模态	图像数据	既包括温度、振动、压力等时序数据,也存在图像数据	工业系统的复杂化引发测点的时空关联性更强
数据的收集方式	通过互联网收集,且有相对完善成熟的基准数据集	工业传感器现场采集,易受到电磁场、设备老化等干扰	工业现场外界干扰导致工业数据受噪声影响更强
数据的类别	由于收集难度相对较低,图片类别十分丰富	大多为正常状态采集到的数据,故障数据难以获取	工业数据的类别更加单一
知识的收集方式	通过互联网收集或通过众包方式获得	需要工厂专工提供专业性知识,分析并标注故障案例	工业知识的专业性门槛更高
知识的类型	基准数据集中的类-属性矩阵、文本描述等关于颜色、形状、环境等常识性知识	不同于生活常识,工业领域需要对工艺、操作流程等的专业认知	工业知识的专业性门槛更高

他领域数据,由于工业环境通常需要监测大量生产参数,以热轧钢产线为例,包括环境与设备温度、轧制压力、电机相关参数、钢板长度参数等几百个测点需要进行监测,导致工业领域中的时序数据具有高维度的特点^[66],且这些参数中涵盖了大量相互耦合的信息,具有时空强关联特性^[67],包含设备之间的相互作用、生产环境因素以及现场人员操作等因素带来的复杂影响,导致故障信息表征十分困难.此外,基于产量要求和原材料供给等要求,钢材生产过程往往具有多种不同的操作状态或模式,且这些模式之间可能存在转移或切换关系,与控制策略紧密相关,从而导致收集的数据具有多状态和强动态特点^[68-71].因此,对于工业零样本任务而言,在对数据和知识间的关系进行建模时处理数据中时序性带来的问题是一个有挑战性的任务.

就数据的收集方式而言,在机器视觉领域,一方面,可以通过互联网搜索引擎等手段获取到海量的训练数据;另一方面,领域内也有大量相对完善的开源数据集可供使用^[38].然而,工业数据往往来源于工业现场部署的大量传感器,许多传感器的工作条件往往比较恶劣,在数据采集、传输、存储的过程中,受到生产环境因素如温度、湿度、磁场、电场的影响,以及设备老化等因素,可能出现数据传输异常或者传感器自身发生漂移,导致数据的离群点和缺失值增多、数据分布发生漂移^[72-74].例如,工业电网、高频震动、厂房中的金属设备都会对工业信号的传输过程产生干扰.工业数据高噪声的特点给数据预处理以及后续建模带来挑战,增加了工业零样本任务中数据表征和知识迁移的难度^[75].此外,对于缺陷检测任务,与一般领域的图像数据相比,因为不同缺陷零件或产品的图像之间的差别往往较小^[50],有的缺陷甚至是肉眼难以分辨的,使得缺陷检测等任务的工业图像数据更加难以辨识.

就数据的类别多样性而言,在机器视觉领域,由于样本的收集难度较低,来源较多,导致能够收集到的图像类别十分丰富.然而,对于工业领域,工业装备在服役过程中往往长期处于正常工作状态,数据类别相对单一,一些发生频次比较低的故障可能从未收集到相关数据,数据类别的分布具有长尾效应,导致所收集到的监测数据信息重复性大,数据价值密度较低,需要采取手段进行筛选和提纯^[76].故障状态往往会降低生产产品的品质,或对生产设备产生危害,因此一般不会为了获得故障数据而主动在实际过程中模拟故障的产生^[77],从而导致异常故障样本“可遇而不可求”.且工业故障往往存在非线性、非高斯、时变与多变量耦合的复杂特性^[78],许多工业生产过程特别是一些复杂化工过程的机理尚不完全清晰,因此在计算机上对故障进行仿真常常难以实现.工厂的存储能力、存储成本以及数据管理水平也对故障数据的积累产生了限制.此外,由于各工厂生产工艺、配方的保密性要求以及政策监管要求,工业领域还存在数据孤岛现象^[79],数据的流动性、开放性较差,这进一步加剧了故障数据缺失的问题.这些问题导致提取泛化性好的表征并在零样本学习过程中实现不同类别间的迁移变得较为困难.

就知识的收集方式而言,对于机器视觉领域,图像的“颜色”“形状”等知识大多接近于生活常识,辅助知识可通过互联网收集相关资料或通过众包等方式由普通人标注获得;对于工业领域,不同于一般的生活常识,工业知识的专业性门槛更高,所需要的辅助知识往往是轴承异常振动的程度、温度异常上升的设备部件、某故障发生的原因等关于设备、工艺、操作流程等专业化知识,需要由工厂专工等工业领域专家总结或对相关故障案例进行标注,其收集和标注成本较为高昂.此外,对知识提供者的专业性要求也导致所收集的知识往往是相对稀少或不够全面的,从

而增加了工业零样本学习任务的难度。

就知识的类型而言,面向特定的工业应用场景,工业零样本的辅助知识类型还需要针对性的设计,如工况参数等反映运行状态的指标等。

2.4 工业零样本学习典型问题

工业领域零样本学习的价值在于减少了对训练样本收集的需求。然而,由于模型无法接触未见类别数据,已见域与未见域之间的偏移问题会较为突出。同时,由于许多现有的工业零样本学习方法大多利用语义属性作为辅助知识,如何恰当地评判不同属性的重要性也十分重要。此外,在数据空间至语义空间的映射中,部分使用最近邻搜索的分类方法可能受到“枢纽点”现象的干扰。本节将从域偏移^[29]、属性重要性评估^[80]和枢纽点^[81]3个角度介绍工业零样本学习任务面临的问题。

2.4.1 域偏移问题

零样本学习任务中广泛存在域偏移问题(domain shift),具体而言,由于已见类与未见类的数据样本相关性小且分布不同,仅使用已见类别数据训练的模型对未见类的泛化能力较差^[29]。在复杂的工业场景下,受到生产设备、环境或数据传输的影响,已见类别和未见类别的数据分布差距较大,域偏移问题更加显著。例如,在零样本工业故障检测中,若已见故障是关于轴承的,而未见故障是关于齿轮的,则模型较难实现从已见到未见的泛化。域偏移问题在广义零样本学习和归纳式设定下更具有挑战性,因为广义零样本学习的测试集同时包含已见类别和未见类别样本,且归纳式设定的训练集完全不包含未见类别样本的有关信息。在广义零样本学习任务中,由于域偏移问题,在测试阶段,模型往往会将未见类别样本误判为某个已见类别。遗憾的是,目前尚未出现能够以足够精度区分已见和未见类别的模型,这是由于模型只能接触到已见类别数据,无法接触到未见类别数据,故二者的分类边界难以划定。

2.4.2 属性重要性评估问题

在工业领域的零样本学习研究中,就类-属性矩阵这一知识表示形式而言,如何高效、高质量地构建类-属性矩阵并对其质量进行评估^[80]是值得关注的问题。在工业零样本学习任务中,常常利用具有若干属性的语义描述建立已见类别与未见类别的关联。一般情况下,语义描述中每个属性都被平等对待,即默认属性间是相互独立且重要性相当。然而,对于某个具体的故障或缺陷类别而言,并非所有的属性对其识别都是有帮助的,尤其在工业数据集中,尽

管属性大多来自专家的标注,但受限于时间和人力成本等因素,往往不会提供属性的重要程度这一信息。对于不同的故障而言,属性的重要性应该是不同的,且属性间可能相互关联。若简单地将属性视作同等地位,则可能影响零样本诊断模型的性能。更进一步,如何基于属性的特性对于不同模态的数据进行解耦表征学习^[82],以及如何考虑属性间的层级与耦合特性^[83-84],都值得进一步研究。

2.4.3 枢纽点问题

枢纽点现象^[85-86]是指将原始特征空间中的样本映射到用于分类的特征子空间时,如果在特征子空间中基于 K 近邻方法进行分类,则某些点会成为大多数点的最近邻点,导致测试样本被误分类,这是高维空间的固有问题。枢纽点现象可以使用 k -出现(k -occurrence)的概念进行描述,即一个数据点 x 作为其他数据点的 k 近邻的次数,可记为 $N_k(x)$ 。如果数据点 x 的 $N_k(x)$ 远高于全部数据点的平均水平,则数据点 x 是一个枢纽点。对于工业零样本学习中的嵌入式零样本学习方法而言,往往采用 K 近邻方法在高维空间中进行分类,由于工业数据高维度的特点,其会受到枢纽点现象的干扰而影响性能。

2.5 工业领域零样本学习模型

工业零样本学习模型主要分为嵌入式和生成式两种。嵌入式模型的目标是学习一个嵌入函数,将数据和辅助知识投影到公共空间,从而通过相似度的度量实现分类。对于未见类别,只需要具有其辅助知识即可实现分类。生成式模型则基于已见类别数据和辅助知识间的联系生成未见类别的样本,从而将零样本学习任务转化为传统的有监督学习任务^[87]。下文将对这两大类模型进行具体介绍。

2.5.1 嵌入式工业零样本学习模型

在实际工业场景下,由于故障的样本“可遇而不可求”,可用于训练的故障类型往往是稀缺的。与故障数据相比,专家多年积累的经验知识是相对全面的,从而可以基于专家知识尝试诊断未见类别的故障,其关键在于如何利用专家知识。一种直接有效的做法是基于已见故障建立数据和知识的映射,并将其迁移到未见类别实现零样本诊断,即嵌入式零样本学习模型。如图4所示,嵌入式零样本学习模型将原始数据映射至语义空间,或是将语义信息映射至数据或者特征空间,进而识别未见故障。根据所采用映射方式的不同,本节介绍了数据空间至语义空间的映射和语义空间至数据空间的映射两类模型。

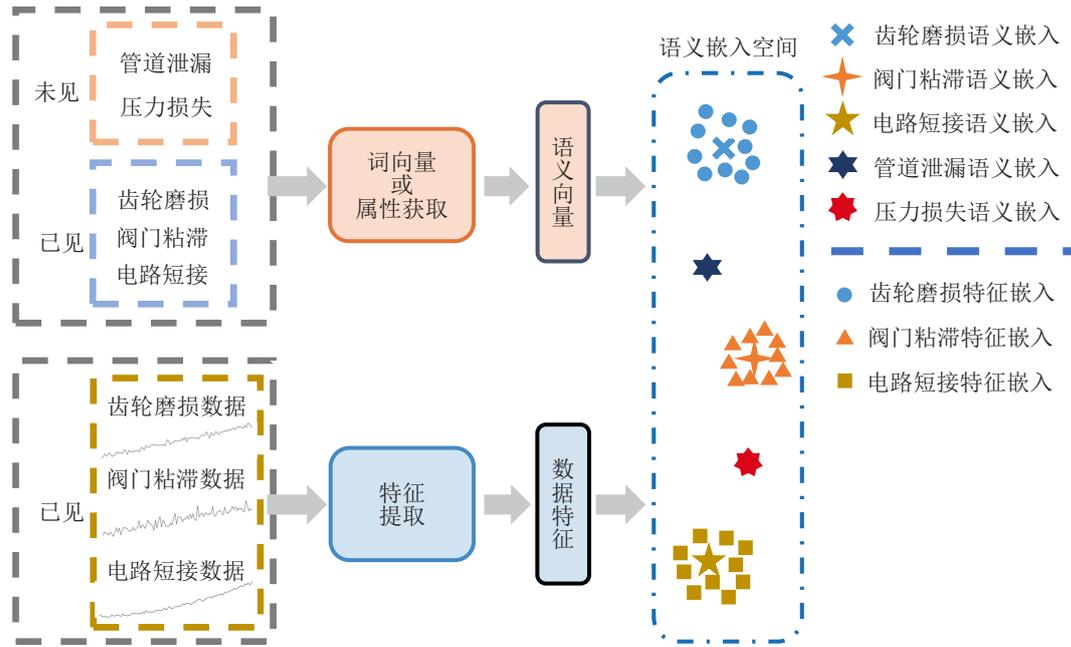


图4 嵌入式范式

1) 数据空间至语义空间的映射.

该类方法中最具代表性的便是直接属性预测 (direct attribute prediction, DAP)^[22], 如图5所示. DAP方法以类别具有的属性作为一种沟通已见与未见类别的桥梁, 这也是零样本学习领域早期的研究工作之一. 针对工业故障诊断任务, DAP方法首先对数据进行特征提取, 然后设置一组相互独立的属性二分类判别器 $\alpha = [\alpha_1, \alpha_2, \dots, \alpha_C]$. 其中: α_i 对应第 i 个属性, C 为属性个数. 通过具有属性标签的训练样本训练属性判别器, 从而可预测未见类别样本的属性. 具体而言, 通过转换映射 ϕ 从样本 x 中提取特征 b , 基于特征判断其所具有的故障属性, 即 $p(\mathbf{a}|\mathbf{b}) = \prod_{i=1}^C p(a_i|\mathbf{b})$, 其中 $\mathbf{a} \in \mathbb{R}^C$ 为故障描述矩阵 \mathbf{A} 中一种故障的属性描述. 这里, 第 t 类的属性向量 \mathbf{a} 表示为 $\mathbf{a}^t = [a_1^t, \dots, a_C^t]$. 基于贝叶斯规则, 推理过程可表示为

$$p(t|\mathbf{a}) = \frac{p(t)}{p(\mathbf{a}^t)} [\mathbf{a} = \mathbf{a}^t]. \quad (3)$$

如果 $\mathbf{a} = \mathbf{a}^t$, 则 $[\mathbf{a} = \mathbf{a}^t] = 1$, 否则为0. 在缺少明确信息的情况下, 各类别的类先验概率 $p(t)$ 被认为是等同的. 对于 $p(\mathbf{a}^t) = \prod_{i=1}^C p(a_i^t)$, 可使用训练集各个已见类别的经验均值 $p(a_i^t) = \frac{1}{N_s} \sum_{j=1}^{N_s} a_i^{c_j^s}$ 代替. 对于测试样本 x , 其故障类别后验概率如下:

$$p(t|x) = p(t|\mathbf{a})p(\mathbf{a}|\mathbf{b}) = \frac{p(t)}{p(\mathbf{a}^t)} \prod_{i=1}^C p(a_i^t|\mathbf{b}), \quad (4)$$

其中 $\mathbf{b} = \phi(\mathbf{x})$. 为了从所有的故障类别 t_1, \dots, t_{N_u} 中为一个测试样本挑选出最可能的故障类别, 可以使用最大概率估计的方法, 即

$$f(\mathbf{x}) = \arg \max_{j=1,2,\dots,N_u} \prod_{i=1}^C \frac{p(a_i^{t_j}|\phi(\mathbf{x}))}{p(a_i^{t_j})} = \arg \max_{j=1,2,\dots,N_u} \frac{p(\mathbf{a}^{t_j}|\mathbf{b})}{p(\mathbf{a}^{t_j})}. \quad (5)$$

其中: \mathbf{a}^{t_j} 为第 j 种目标故障的属性描述向量, $a_i^{t_j}$ 为 \mathbf{a}^{t_j} 的第 i 个元素^[22].

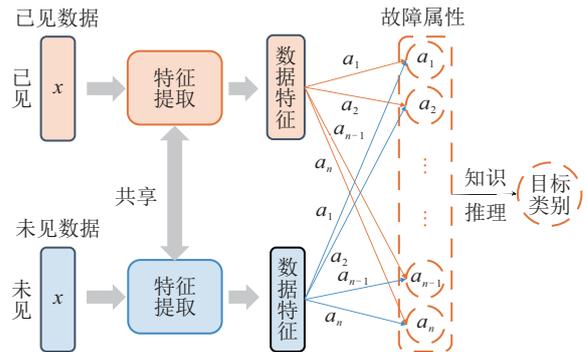


图5 直接属性预测方法

直接属性预测方法要求各属性之间相互独立, 这在工业过程中往往难以满足. 由于工业过程的强关联特性, 以及工业过程本身的机理因素, 描述生产过程的各个属性很可能相互关联、耦合. 同时, 当工业系统较为复杂时, 描述其的属性往往也较多, 此时若使用 DAP 模型则所需要的模型数量随属性数量增加, 导致实际应用上的不便. 因此, 一些工作^[56,58,88-89] 利用非线性映射 ϕ 对数据进行特征提取, 并利用兼容度函数完成语义和工业数据特征的匹配. 样本 x_i 与属

性向量 \mathbf{a}_j 的兼容度可表示为

$$C(\mathbf{x}_i, \mathbf{a}_j) = \phi(\mathbf{x}_i)^T W \mathbf{a}_j, \quad (6)$$

其中 W 为可学习的参数. 通过优化过程期望对于相匹配的数据特征和语义特征能够基于兼容度函数计算出较大的兼容度, 可规定该兼容度不应小于设定的超参数 λ , 其目标函数为

$$L = \frac{1}{N_{tr}} \frac{1}{N_s} \sum_{i=1}^{N_{tr}} \sum_{j=1}^{N_s} \max[0, \lambda - \Pi_{ij} \times C(\mathbf{x}_i, \mathbf{a}_j)]. \quad (7)$$

其中: N_{tr} 为训练样本数量; N_s 为已见类别数; 当 \mathbf{x}_i 和 \mathbf{a}_j 来自同一个类别时 Π_{ij} 取 1, 否则取 0.

基于这一类模型, 面向工业过程零样本故障诊断中的关键与挑战性问题, 本领域学者开展了一些探索性研究. Feng 等^[22] 首次将零样本学习应用于故障诊断领域, 基于 DAP 方法提出了基于故障描述的属性迁移方法, 以解决零样本故障诊断任务. 该方法通过人为定义的过程故障描述 (故障位置、故障原因等) 而不是收集的故障数据来确定故障类别, 每种故障均由 20 个专家标注的细粒度的二值属性进行描述. 基于此, 模型可以在已见故障数据中学习辨识故障属性的能力, 从而识别未见类别故障. Huang 等^[90] 在 DAP 模型基础上, 通过属性辨识结果判断测试样本是否属于已见故障, 从而实现广义零样本诊断. Lin 等^[89] 利用固定步长的滑动窗口在原数据集中采样, 通过非线性映射将样本特征投影到语义属性空间, 并在分类阶段通过计算样本属性与所有未见故障属性的相似度进行分类. Zhang 等^[91] 定义了轴承故障的多个属性, 包括故障位置、故障属性等, 利用 1DCNN 作为时序特征提取器完成了针对高频故障数据的零样本诊断. Guo 等^[56] 提出了潜特征指导的属性注意力网络 (latent feature guide attribute attention, LGFAA). 该网络包含两个分支, 一个分支使用全连接层学习视觉空间到语义空间和潜在属性空间的映射, 得到预测语义属性、潜在属性和基于目标的属性注意力; 另一分支使用语义嵌入模块学习视觉空间到语义空间的映射, 反馈机制将输出的语义向量反馈到公共属性空间, 得到修改的潜在属性, 这样可以在一定程度上考虑到属性的重要程度, 实现对属性的优化. 最后, 使用兼容度函数结合潜在属性空间和语义属性空间, 在气缸套缺陷数据集上实现了未见缺陷类别判断.

2) 语义空间到数据空间的映射.

在数据空间至语义空间的映射中, 数据特征被投影到语义空间, 并在高维空间中利用最近邻搜索确定类别. 这种情况下, 最近邻搜索易出现前述的枢纽

点问题, 即在高维空间中某些点会成为大多数点的最近邻点^[81,92], 影响分类性能. 但若采用逆映射策略^[12], 即从语义空间映射到数据空间, 则可以有效缓解枢纽点问题. 此类范式目前在工业缺陷检测任务中应用较多, 这是由于缺陷工件图像的特征维度通常较高, 若从数据空间映射到语义空间, 则空间萎缩的程度更高. 在训练阶段, 该方法学习一个映射函数 $f(\cdot)$, 将语义描述符 \mathbf{a}_i 映射到其对应的图像特征 $\phi(\mathbf{x}_i)$. 为了使语义描述与对应的图像特征更接近, 通常采用最小二乘损失函数来最小化差异, 有

$$L = \frac{1}{N_{tr}} \sum_{i=1}^{N_{tr}} \|f(\mathbf{a}_i) - \phi(\mathbf{x}_i)\|_2^2 + \lambda g(f). \quad (8)$$

其中: $g(\cdot)$ 为对模型参数的正则项, λ 为正则项的权重超参数. 在样本分类阶段, 对于每个测试样本 \mathbf{x} , 可在数据空间中采用最近邻方法预测类别, 即

$$\hat{y} = \arg \min \|\mathbf{x} - f(\mathbf{a}^c)\|_2, \quad c \in T. \quad (9)$$

其中: \hat{y} 为预测结果, T 为所有测试类别的集合.

基于这一范式, Sun 等^[12] 使用轮毂缺陷数据集, 利用卷积神经网络将缺陷图像映射到图像特征空间, 并采用多层感知器将缺陷语义描述也映射到该空间, 通过一一配对的方式将语义描述嵌入到相应的图像特征中. 在测试阶段, 则采用基于最近邻分类器的方法对未见轮毂故障进行分类. Hu 等^[93] 提出了一种语义一致性嵌入方法, 通过为故障数据和故障语义各设计一个自编码器, 不仅完成了二者的重构, 还要求二者的潜变量能够互相重构. 此外, 该工作还通过优化数据潜变量和语义潜变量的 Barlow 矩阵, 学习具有跨模态一致性的潜空间.

2.5.2 生成式工业零样本学习模型

如前文所述, 嵌入式零样本学习仍然面临着未见类别数据缺失的问题, 因此在广义零样本场景下, 嵌入式方法的域偏移现象十分严重. 而生成式工业零样本学习模型基于现有的故障数据和专家知识生成新的样本, 如图 6 示, 以缓解故障样本相对少且不全面的问题^[25,31-32,64,94], 从而减弱域偏移的负面影响. 生成式模型通过学习语义描述和真实特征分布的关系, 利用未见类的语义描述生成未见类的数据或特征, 从而将零样本学习问题转换为有监督学习问题. 根据所使用的生成方法的不同, 生成式工业零样本模型可以分为基于自编码器的生成和基于生成式对抗网络的生成. 需要指出的是, 基于 VAE 的生成和基于 GAN 的生成只是使用的生成模型不同, 两者所适用的任务和场景是基本相同的. 对于生成式方法而言, 通过生成未见类别的样本, 可以训练一个一般

的有监督分类器用于分类任务而不局限于 K -NN 分类器,从而回避了枢纽点问题。

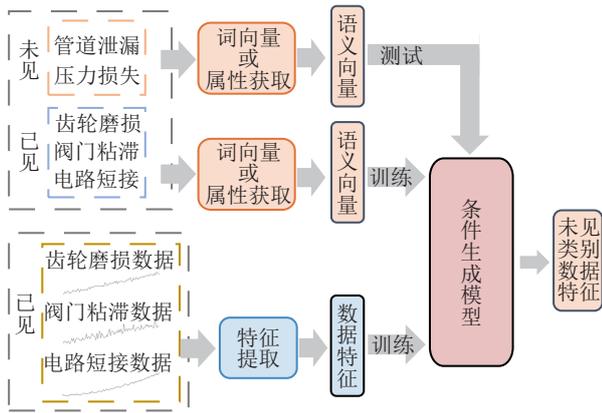


图6 生成式范式

1) 基于变分自编码器(VAE)的生成。

VAE 是一种有效的特征生成模型,包含编码器(encoder)和解码器(decoder)两部分^[95]。在训练阶段,VAE通过损失函数限制encoder编码的隐变量尽可能服从标准正态分布,并限制decoder尽可能重构输入。若要生成新的特征,则只需要从标准正态分布中采样出一个隐变量,再将隐变量输入decoder获得输出。在工业零样本学习领域,常采用条件变分自编码器(conditional variational autoencoder, CVAE)作为生成模型,CVAE将某些先验条件 a 与原数据组合作为模型输入,从而控制生成的类别,结构如图7所示。其中: N_s 为已见类别数, x 为某已见类别样本, a 为其对应语义属性描述, x 和 a 的组合作为CVAE的输入。编码器将 x 和 a 的组合转换为潜变量 z ,其服从均值为 μ 和方差为 σ 的高斯分布,然后将编码器的输出 z 和语义属性 a 的组合作为解码器输入,从而得到解码器的输出 \tilde{x} 。CVAE损失函数为

$$\begin{aligned} \mathcal{L}_{cvae} &= \mathcal{L}_{rec} + \text{KL}(\mathcal{N}(\mu_x, \sigma_x) \parallel \mathcal{N}(\mathbf{0}, \mathbf{I})), \\ \mathcal{L}_{rec} &= \frac{1}{2n} \sum_{i=1}^{N_s} \|x_i - \tilde{x}_i\|_{l_2}^2. \end{aligned} \quad (10)$$

其中: L_{rec} 为重建损失, KL 为 Kullback-Leibler 散度。可以看到,CVAE要求解码器输出要重构原输入,并希望编码器生成的潜变量 z 遵循高斯分布 $\mathcal{N}(\mathbf{0}, \mathbf{I})$ 。训练得到的解码器即为生成器,无论是对于故障诊断任务还是缺陷检测任务,根据其知识(语义)信息,生成器均可生成未见类别的故障特征或缺陷特征,从而将零样本学习问题转换为常见的有监督学习问题。令 z 为服从 $\mathcal{N}(\mathbf{0}, \mathbf{I})$ 的高斯噪声,可使用已见类别语义属性 $a \in \mathcal{A}_s$ 和 z 生成已见故障的伪样本,也可以根据未见的语义属性 $a \in \mathcal{A}_u$ 和 z 生成未见故障的伪样本。基于变分自编码器的生成式模型已在工

业领域得到应用,例如,Huang等^[55]针对广义零样本场景,利用Feng等^[22]所构造的语义属性作为VAE生成模型的输入生成未见类别的样本,由此训练一个二分类的门区分已见类与未见类。对于已见类使用监督学习方法分类,对未见类则采用Feng等^[22]提出的故障属性迁移方法识别。Li等^[79]在云边协同的场景下,设计了一种双向对齐网络,该网络使用两个耦合的变分自编码器实现数据与属性的融合,并利用生成重放机制整合每个客户端的知识,从而提高全局模型的泛化性,提升对未见故障的诊断能力。

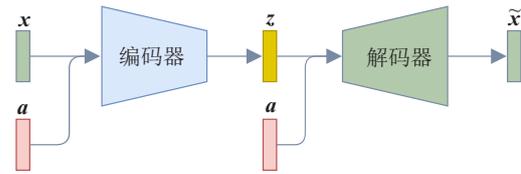


图7 CVAE结构

2) 基于生成式对抗网络(GAN)的生成。

生成式对抗网络是由Goodfellow等^[96]于2014年提出的通过对抗训练生成数据的方法。GAN通过模拟博弈过程训练模型,使网络学习到从正态分布到目标分布的映射,并具有良好的生成能力。Gao等^[97]指出,VAE对复杂数据分布的拟合能力偏弱,而GAN则能够更好地拟合数据的全局信息。但从训练过程的角度看,GAN的训练不稳定性较高。在工业零样本的场景中,常采用条件生成对抗网络(conditional generative adversarial network, CGAN)作为生成模型。CGAN利用已见类别的语义作为条件训练GAN,再使用GAN根据未见类别语义生成未见类别数据,从而完成已见域到未见域的跨越。CGAN结构如图8所示,由生成器 G 和判别器 D 组成,生成器输入为高斯噪声 z 和语义向量 a ,输出为 a 对应类别的合成数据 \tilde{x} 。判别器输入为真实数据 x 、合成数据 \tilde{x} 和语义向量 a ,输出为判定输入真假的分数。生成器 G 的目标函数为

$$L_G = \mathbf{E}[\log(1 - D(G(z, a)))] \quad (11)$$

判别器 D 的目标函数为

$$\begin{aligned} L_D &= -\mathbf{E}[\log D(x, a)] = \\ &= -\mathbf{E}[\log(1 - D(G(z, a), a))]. \end{aligned} \quad (12)$$

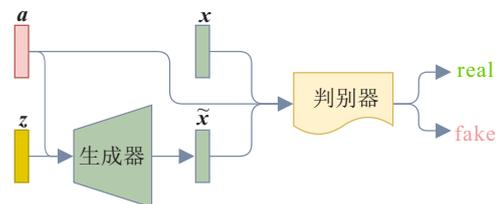


图8 CGAN结构

二者以博弈的形式协同训练,直至生成器生成的合成样本能够迷惑判别器为止,便可以认为生成器能够在语义信息的条件下拟合从高斯分布到目标类别数据的分布.由此,通过向生成器中输入未见类别语义描述 \mathbf{a}_u 和高斯噪声,即 $\tilde{\mathbf{x}}_u = G(\mathbf{z}, \mathbf{a}_u)$,即可生成未见类别的样本.利用这些样本可以训练出对未见类别具有识别能力的分类器,从而完成零样本学习任务.一些学者面向工业应用场景对基于生成自编码器的基本框架进行改进,以解决工业零样本诊断任务.例如,Zhuo等^[31]利用生成对抗网络,以语义属性作为生成模型输入,生成未见故障的伪样本,从而将零样本任务转化为有监督故障诊断任务,同时在模型中引入对比损失辅助区分不同故障的细节信

息,从而使模型能够根据有限的故障样本生成更可靠的虚拟样本.面向机械旋转设备的零样本故障诊断任务,Chen等^[25]基于生成对抗网络,用振动信号特有的统计属性作为知识,利用正常样本训练生成不可见的伪样本,最后利用深度神经网络进行故障识别.Xu等^[64]提出了一种复合故障零样本诊断方法,该方法将复合故障分解为若干个单一故障表示,将单一的故障属性向量作为工业知识,利用生成对抗网络生成复合故障的伪样本.

3 工业领域零样本学习数据集与开源工作

在本节中,为了便于本领域后续的研究开展,梳理了工业零样本学习的基准数据集,并介绍了本领域相关的开源工作,如表6所示.

表6 现有的工业零样本研究工作

范式	年份	主要模型	辅助知识	应用对象	数据集	广义零样本
嵌入式	2020	DAP ^[22]	故障属性矩阵	火电过程/化工过程	TPP/TEP	否/否
	2020	CSA ^[60]	工作状态参数	轴承设备	CWRU	否
	2021	NMC ^[89]	故障属性矩阵	化工过程	TEP	否
	2021	SF-LFGAA ^[56]	语义词向量	气缸套缺陷	CLSDD	否
	2021	ZSR ^[12]	故障属性矩阵	轮毂缺陷	WHD-9	否
	2022	LDS-IFD ^[58]	复合故障组成	轴承设备/行星齿轮箱	LB/DDS	是/是
	2022	ZLCFDM ^[88]	复合故障组成	轴承设备	自建平台	否
	2022	SCE ^[93]	故障属性矩阵	化工过程	TEP	否
	2022	ZS-CKG ^[98]	向量化文本	热轧钢带缺陷	NEU	否
	2022	SSB ^[91]	故障属性矩阵	轴承设备/化工过程	CWRU/TEP	否/否
2023	SGZSL ^[90]	故障属性矩阵	化工过程	TEP	是	
生成式	2019	CGAN ^[25]	无	轴承设备	CWRU/SQ	否/否
	2020	CADAE ^[32]	工作状态参数	轴承设备	CWRU/船舶天线/SQ	否/否/否
	2021	FAGAN ^[31]	故障属性矩阵/工作状态参数	化工过程/液压设备	TEP/HS	否/是
	2021	CVAE ^[55]	故障属性矩阵	化工过程	TEP	是
	2021	GAN ^[64]	复合故障组成	轴承设备	自建平台	否
	2023	BAN ^[79]	故障属性矩阵	火电过程	TPP	否

3.1 基准数据集

故障诊断和缺陷检测是当前工业零样本研究中重要且主要的应用场景.下文梳理了典型的工业零样本故障诊断以及工业零样本缺陷检测的数据集,包括化工过程的田纳西-伊斯曼数据集、机械旋转设备数据集、火力发电过程故障数据集以及产品缺陷检测数据集.表7总结现有的开源数据集,并从数据形式、故障类型、案例数量、变量数目等角度进行了对比.从数据形式上看,现有的开源工业零样本数据

集以时序数据为主,也存在图像数据集.其中时序数据所对应的对象有化工过程(低频)、液压设备(中频)以及轴承设备(高频),相对丰富,图像数据所对应的对象则为热轧钢带.从数据规模上看,大部分数据集规模较为适中.从数据特征上看,化工过程和液压设备的测点较多,而轴承设备采样频率较高,且包含的测点较少.时序数据所对应的语义一般为专家给出的故障属性矩阵或者工况参数,图像数据对应的语义则为其文本描述的向量化表示.

表7 开源数据集

	面向对象	数据形式	采样频率/Hz	故障类型数	案例数量	变量数目	语义知识
TEP ^[54]	化工过程	时序数据	5.56×10^{-3}	21	21	52	故障属性矩阵
CWRU ^[99]	轴承	时序数据	48 k/12 k	9	161	3	工况参数
SQ ^[32]	轴承	时序数据	25.6 k	7	84	2	工况参数
HS ^[100]	液压设备	时序数据	100/10/1	256	2205	17	故障属性矩阵
NEU ^[101]	热轧钢带	图像数据	—	6	1800	200×200 像素	向量化文本

1) 田纳西-伊斯曼过程数据集.

田纳西-伊斯曼过程 (Tennessee-Eastman process, TEP) 是基于 Tennessee Eastman 化学公司实际化工生产过程的仿真系统^[54], 具体流程如图9所示. TE 过程模拟了实际复杂工业过程系统的许多典型特征, 是化工领域故障诊断领域广泛应用的数据集, 有助于公平地展现零样本诊断模型的有效性. TE 过程由5个主

要子系统组成, 包括反应器、冷凝器、压缩机、气液分离器和汽提塔. 常使用的TEP数据集中包含21种故障, 每种故障包含41个测量变量和11个操作变量. 每种故障训练集包含480个样本, 测试集包含960个样本. 由于后6种故障在数据集中未提供相关描述, 无法提炼语义知识, 在零样本故障诊断任务中通常使用前15种故障类别.

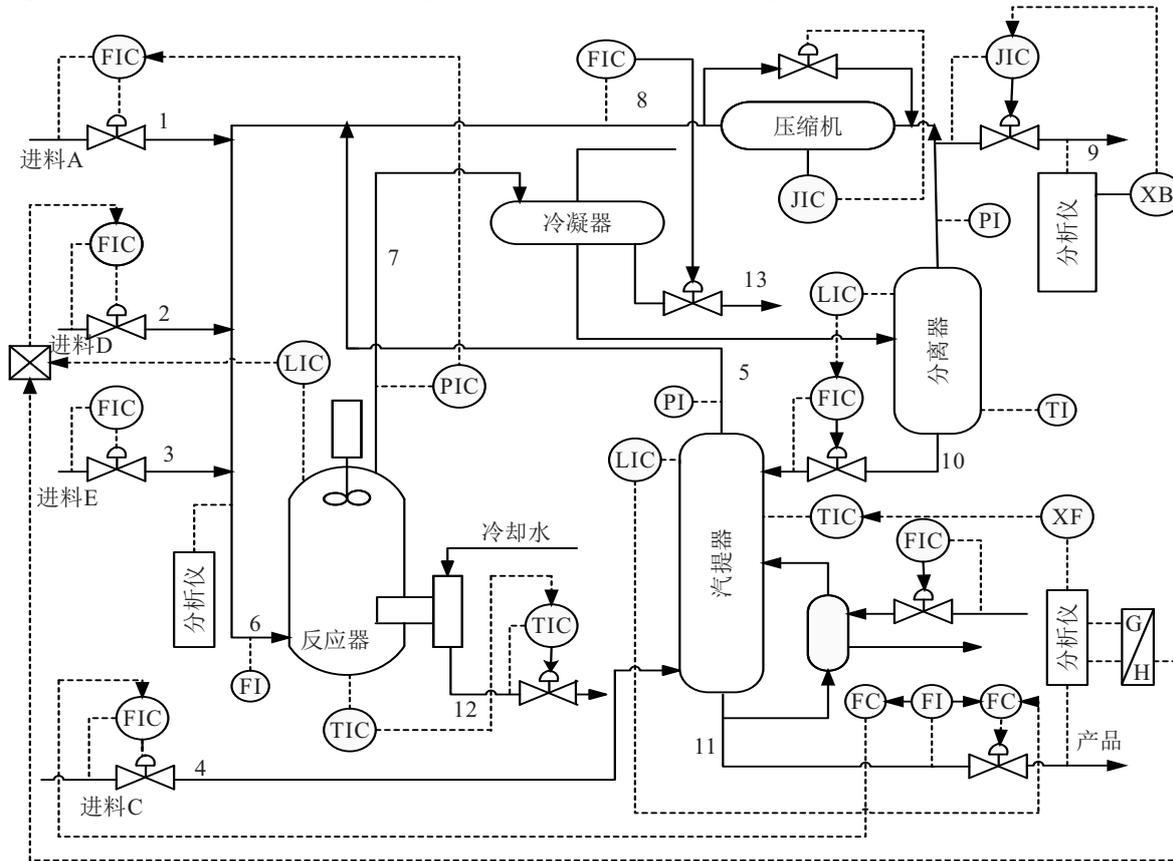


图9 Tennessee Eastman过程

2) 机械旋转设备数据集.

轴承作为以化工、能源领域为代表的制造业大型装备核心部件, 长期在高转速下运行极易故障, 因此对机械旋转设备的监测、故障诊断具有重要意义. 机械制造领域数据集较多, 如凯斯西储大学 (Case Western Reserve University, CWRU) 轴承数据集^[99]、辛辛那提轴承数据集^[102]、机车轴承 (locomotive bearing, LB) 数据集^[103]、动力传动系统模拟 (drive-train diagnostics simulator, DDS) 数据集^[104]、Spectra Quest (SQ) 机械故障模拟数据集^[32]、液压 (hydraulic system, HS) 故障数据集^[100]等, 其中以CWRU数据集应用最为广泛.

分别为7 mils、14 mils、21 mils、28 mils和40 mils. 数据集中包含4种不同的工况, 分别是正常工况、外圈故障、内圈故障和轧辊故障. 与TE过程不同的是, 机械设备的故障属性一般并不是由故障位置、部件等对故障的描述, 而往往是运行数据本身的一些统计特性, 例如利用振动信号本身的统计属性, 如均值、方差等构成^[25].

如图10所示, CWRU数据集由轴承测试台设备生成, 实验台由三相感应电机、轴和测功机组成. 为了模拟轴承故障, 通过电火花加工方法在轴承的内圈、滚动体和外圈引入直径不同的5种单点故障, 故障点

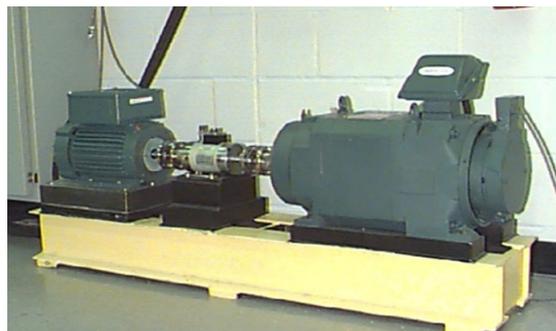


图10 CWRU数据集实验平台

3) 火力发电过程故障数据集.

火力发电(thermal power plant, TPP)过程故障数据集来自如图11所示的1000 MW超超临界火电机组的实际工业过程^[98]. 热过程由两个主要子系统组成,即锅炉系统和汽轮机系统. 发电过程中,锅炉系统

首先将水加热成高压高温蒸汽,然后蒸汽被输送到汽轮机系统以驱动发电机,整个动力单元实现了化学能向电能的转化. 由于火力发电是一个大规模过程,通过机组停机收集故障样本可能会造成巨大损失,因此在零样本设置下诊断某些故障类别具有重要意义.

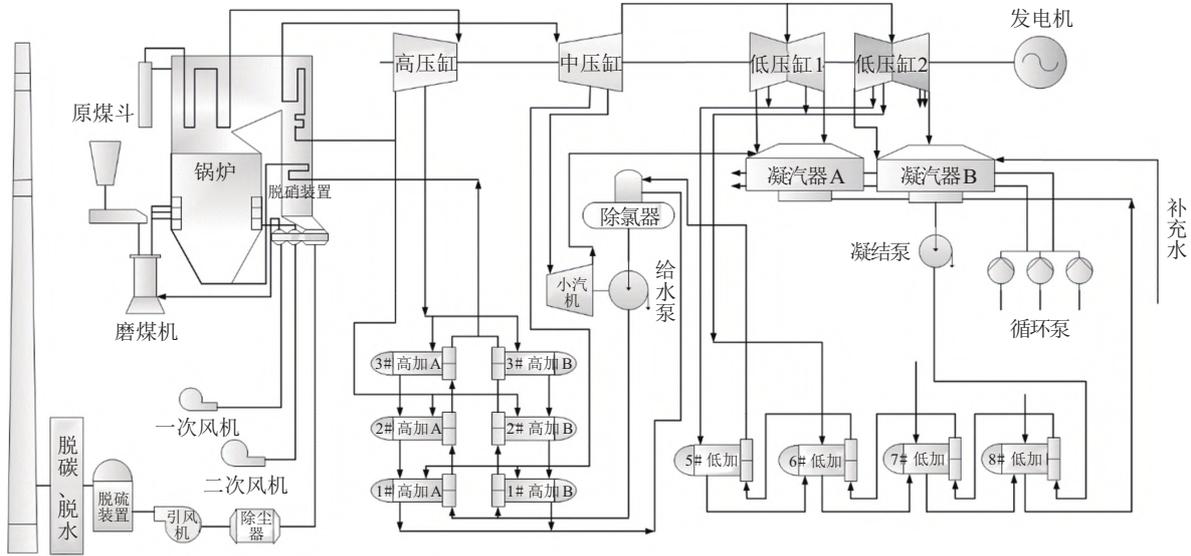
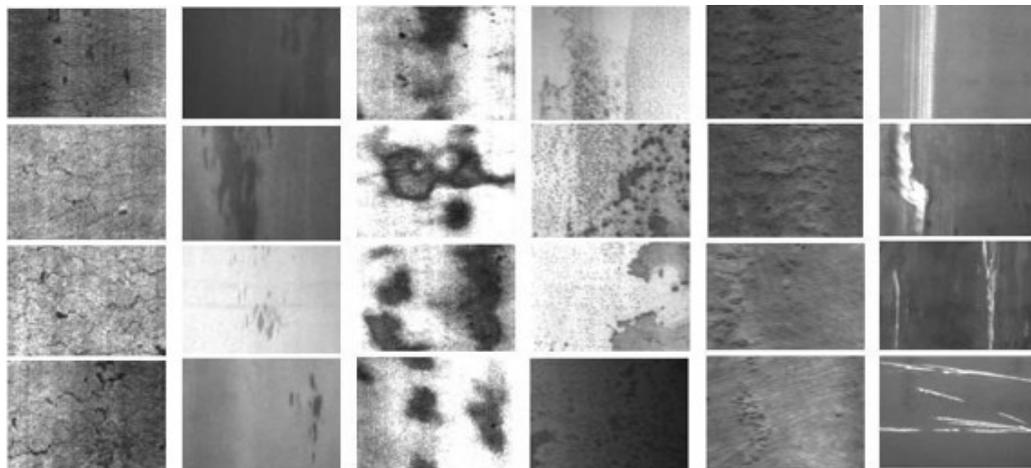


图 11 1000 MW 超超临界火电机组系统配置

4) 产品缺陷检测数据集.

气缸套缺陷数据集(cylinder liner surface defect dataset, CLSDD)^[56]由一家工业技术公司提供,包含5种缺陷类型数据(裂纹、磨损、凸面、收缩、气蚀)和正常类型数据共412张图片. 轮毂缺陷(wheel hub defects, WHD-9)数据集是Sun^[12]等与轮毂厂商合作,采集生产线上的轮毂缺陷图像数据并完成标注得到的数据集,包括9种轮毂缺陷(油污、磨齿、划痕、堵头、下垂、压痕、橘皮、变形、灰尘). 此外,东北大学(North Eastern University, NEU)表面缺陷数

据库^[101,105]和Xsteel表面缺陷数据集(Xsteel surface defect dataset, X-SDD)^[106]是钢表面缺陷检测场景下从热轧钢带上收集的两个常用数据集,其中NEU数据集包含6种不同表面缺陷类别(轧入氧化皮、斑块、龟裂、麻点表面、夹杂物、划痕)的1800张单通道图像,每个类别包含300个样本,如图12所示. X-SDD数据集包含7种不同缺陷类别(红铁皮、铁皮灰、夹杂物、表面划痕、精轧辊、板材系统氧化皮、温度系统氧化物),共1360张图像.



(a) 龟裂 (b) 夹杂物 (c) 斑块 (d) 麻点表面 (e) 轧入氧化皮 (f) 划痕

图 12 NEU数据集

3.2 开源工作

目前本领域已有部分研究人员将研究工作开源,这对于零样本研究的发展具有重大意义和价值。对于嵌入式零样本学习范式,Feng等^[107]发布了基于直接属性预测的工业零样本诊断代码,并在TEP数据集上进行验证。Paul^[108]复现并开源了5种经典的嵌入式零样本学习范式,包括语义自编码器(semantic autoencoder, SAE)、属性标签嵌入(attribute label embedding, ALE)、结构联合嵌入(structured joint embedding, SJE)、ESZSL(embarrassingly simple approach to zero-shot learning)和深度视觉-语义嵌入模型(deep visual-semantic embedding model, DeViSE)。Hu等^[109]开源了语义一致嵌入方法(semantic-consistent embedding for zero-shot fault diagnosis, SCE),同样在TEP数据集上进行了验证。对于生成式零样本学习范式,一些基础的模型代码同样也被开源,如基于生成对抗网络的f-CLSWGAN^[110]和基于变分自编码器的CADAVAE^[111]。最近,Feng等^[112]发布了10种生成式零样本的最新工作的复现模型。在这些开源工作的基础上,研究者能够更加方便地开展工业领域零样本学习的相关研究。

4 工业领域零样本学习的挑战与展望

自零样本学习引入工业领域至今已3年有余,前文已经对目前的研究做出了系统性的回顾,本节则总结梳理了工业零样本现存的重要挑战以及未来有潜力的诸多研究方向,以供参考。

4.1 工业领域零样本学习的挑战

虽然目前已经出现了多篇工业零样本领域的研究工作,但是由于该领域的复杂性仍然存在许多挑战。本节分析了工业零样本学习任务面临的3方面关键性挑战,即缺乏高质量的专家知识、缺乏对工业数据特性的深入分析、缺乏统一的实验标准。

4.1.1 缺乏高质量的专家知识

在当前的工业零样本学习研究中,所利用的知识大多都是专家标注后高度抽象化的知识,依赖于专家的手工标注,如类-属性矩阵等。这限制了工业零样本学习方法的推广应用,一方面,由于收集与提供工业领域知识的专业性要求,导致高质量的工业领域知识获取成本较高;另一方面,也缺乏合适的手段评估工业零样本学习任务中使用的专家知识的质量。辅助知识的完备性和准确性直接影响工业零样本学习模型的效果^[22],因此,目前的工业零样本工作大多在传统零样本场景中表现较好,而在更加困难的广义零样本场景下精度偏低,从而影响了实用价值。总体而言,

如何更高效地获取高质量的专家知识仍是工业零样本研究需要解决的重要问题。

4.1.2 缺乏对工业数据特性的深入分析

对工业数据进行建模和分析的过程需要基于对工业对象的深刻认知。例如:数据来自不同层级设备或者不同机组时需要分别建模;不同设备同一时刻的数据对应的不一定是同一批的产品,需要重新进行时序对齐等。现有的部分工业零样本相关研究对此重视不足,直接继承了机器视觉领域的相关方法。然而,图像数据和工业时序信号在数据结构、数据量和数据质量等方面具有明显的不同。事实上,面向前文所述工业时序数据的相关特性,应采用一些针对性方法,如统计机器学习方法中的慢特征分析^[113-115]、深度学习中的长短期记忆网络(long short-term memory network, LSTM)^[116]、Transformer^[117]等时序模型。此外,即使是与一般的零样本图像分类任务相近的工业异常检测领域,处理对象为缺陷零件的图像样本,其处理方式与一般的动植物分类也有所不同,因为不同缺陷零件图像之间的差别较小^[50],有的甚至肉眼难以分辨,给缺陷检测任务带来了挑战。目前的工业零样本研究中较少考虑到所研究工业对象的特性,并在方法中做出对应的创新,限制了工业零样本模型的效果。

4.1.3 缺乏统一的实验标准

在机器视觉领域,零样本学习经过10余年的发展,已经形成了领域内认可的标准数据集、公认的训练集与测试集划分和评估指标^[33]。然而,目前工业零样本相关工作尚未形成领域内普遍使用的数据集和辅助知识,训练集和测试集划分标准也各有不同,加上开源的工作较少,很难公平地对比不同方法的优劣。零样本任务的性能与训练集和测试集划分标准高度相关,在某一个固定标准上进行对比只能说明在该视角下的方法性能,更换另一种划分标准可能会带来完全不同的结果。采用多种数据集对工业零样本方法进行评估能够减弱这种非公平性。因此,目前工业零样本领域需要形成一批标准数据集、固定训练集测试集划分和准确度标准来促进和规范该研究领域的进一步发展。

4.2 工业领域零样本学习研究展望

对于工业领域的零样本学习应用而言,在未来还有大量方向值得探索。比如,零样本的核心在于知识的获取与利用,然而现有工作对于工业领域知识的来源与利用方式较为单一,限制了工业零样本方法的应用。此外,目前工业零样本的应用场景基本上集中在

故障诊断和缺陷检测,还需要在新的场景中进一步拓展.再者,工业零样本范式也值得进一步研究,针对新范式的研究不仅能够使零样本更加灵活,适应更多的工业对象,还可以解决零样本本身的一些难题(如迁移效果不理想等).因此,本文从知识、应用场景、方法范式3个角度进行研究展望,以期对未来的研究工作产生一定的启发.

4.2.1 新的知识来源与知识类型

从知识的角度而言,一方面,应该找到更加多样的知识来源,从而减少对专家手工标注的依赖;另一方面,应该考虑如何得到更加结构化的知识,从而对知识进行更高层次的抽象与更高效的利用.

1) 更加丰富的知识来源.

随着工业信息化过程的进行,在实际工业生产过程中积累了许多非结构化的专家经验与系统机理文本描述,例如故障诊断记录、设备操作手册、生产过程机理描述、设备维护日志等^[118].这些非结构化的文本知识与专家经验没有得到很好的利用.当然,这些文本知识的利用也存在许多难题需要解决,如文本本身的质量问题、非结构化文本的自动化处理问题、现有的在通用语料库中训练的自然语言模型能否适用于工业领域的问题.此外,除自然语言外,其他模态的知识也值得考虑,如时间序列数据中蕴含的工业系统的时序知识等.未来的改进目标可能包括开发新的自然语言处理算法,以更有效地从这些非结构化且多模态的信息中提取相关知识.

2) 知识的结构化.

近年来,如何将知识组织为信息密度更高、更加结构化、易于应用的形式得到了越来越多的重视,例如知识图谱、树状语义知识等.知识图谱这一概念由Google在2012年提出,它是由不同的结点(实体)和边(关系)以三元组的形式组成的语义网络,能够对多模态异构数据进行整合.工业知识图谱主要针对生产制造领域,以图结构的方式将多源异构数据中的制造知识组合起来.知识图谱作为一种图数据,可以结合多种图表征学习手段加以利用,如TransE^[119]、图神经网络^[120].对于故障诊断任务而言,工厂积累了大量的历史故障日志,这些非结构化文本难以直接利用.对此,可以通过构建故障诊断知识图谱,对文本中故障信息进行本体化,补充实体之间的关联,完成对故障信息的语义理解,从而应用于智能故障诊断.此外,具有深层结构的树状语义知识^[121]也具有一定的应用潜力,这种方法构建了从高层语义概念到低层语义概念间的属性关系,从而提高了零样本分类任

务的可解释性,对于故障诊断等实际任务而言很有价值.如何自动化地提高知识的结构化程度,以更高效地应用于工业零样本任务,仍是值得探索的研究方向.未来的研究可以关注如何设计新的知识抽象和表示方法,或者优化现有的知识图谱和树状语义知识等结构化知识表示方式.

4.2.2 新的应用场景

目前的工业领域零样本学习研究仍集中在故障诊断和缺陷检测两方面应用.然而,零样本学习作为一种提高模型泛化能力、缓解对标注数据依赖的研究,在多种工业领域任务中具有很大的潜在价值,本文尝试就生产事件检测、工业知识图谱的自动化构建、工业时间序列预测3个潜在应用场景进行展望.在这3类任务中,样本数据往往难以获取或者成本较高,且这3个潜在的应用场景具有专家知识积累较多的特性,有助于工业零样本方法的实现.

1) 生产事件检测.

在机器视觉领域,零样本动作识别^[122]、零样本事件检测^[123]任务得到了研究人员的关注.对于工业领域而言,在生产现场,同样部署了大量的摄像头对生产过程进行监测,希望发现人员和生产过程中的异常情况,并能够基于文本描述对监控视频进行检索.尽管为了实现生产事件的智能检测,工程师和专家通常会记录和分析生产过程中的各种事件和故障案例,形成丰富的知识库,例如工人的错误操作方式、设备的异常情况等等.然而,收集并标注全部类型的异常几乎是不可能实现的,且往往由于生产事件的发生频率较低,异常事件的样本数量有限.因此,在工业领域中引入零样本动作识别、零样本事件检测技术极具实际价值.

2) 工业知识图谱的自动化构建.

工业知识图谱的价值已经在实际应用中得到证实^[124].工业知识图谱的自动化构建可以整合和利用工业领域中丰富的知识资源,支持决策和问题解决.在工业知识图谱的自动化构建中,可以利用已有的工业文档、标准规范、专家知识等来构建知识图谱.然而,手工构建知识图谱往往需要大量的人力物力,特别是对于工业等专业领域而言,手工构建知识图谱会占用大量专家时间进行知识的整理和抽象.近年来,随着零样本学习技术与自然语言处理技术相结合,零样本命名实体识别^[125]、零样本关系抽取^[126]开始得到研究人员的关注,有助于降低知识图谱的构建成本、提高构建过程的自动化程度.

3) 工业时间序列预测.

时间序列预测任务在工业领域十分常见,例如生产负荷预测任务^[127]. 工业时间序列预测可以用于预测生产过程中的关键指标,如产量、能耗等,以优化生产计划和资源分配. 时间序列预测模型的建立通常需要大量的与目标任务高度相关的数据用于建模,工业时间序列数据的采集和记录也需要经验丰富的工程师和设备支持. 近年来,零样本时间序列预测任务被提出,通过提取多个时间序列预测任务的共享特性,即时间序列预测的一般知识,并将其迁移到未见类别的时间序列中,从而实现未见类别的零样本时间序列预测^[128],这有助于缓解时间序列预测任务中模型的冷启动问题,降低对建模数据的要求,特别是对于数据收集困难的场景,例如某些生产指标测量成本高昂的情况,是十分有意义的.

4.2.3 新的范式

第2.5节介绍了目前工业零样本学习的两大类模型,需要指出的是,虽然这两种模型都能够利用知识完成从已见域到未见域的迁移,但并不意味着能够完成迁移的方法只有这两类. 工业领域任务场景复杂多变,对象特性亦有诸多不同,因此在工业零样本学习领域中挖掘新的零样本学习范式是很有必要的. 结合机器学习领域的研究经验而言,图学习范式^[129]、元学习范式^[130]、预训练范式^[131]和云边协同范式^[79]是当下各种零样本范式中具有工业应用潜力的范式,对复杂工业场景的适应能力更强,因此本节将对其进行具体介绍.

1) 图学习范式.

图学习^[132](graph learning)是近年来机器学习研究的重点之一,能够以包含节点和边的图描述结构化的对象. 图学习的核心思想是利用结构化的图表示来捕捉数据之间的相互关系,以便更好地建模. 目前在工业领域中,图卷积网络^[133]作为一种融合了拓扑结构的深度学习方法,相比于其他方法在挖掘数据关联特性能力上更加强大,得到了广泛的应用. 在零样本学习中,图本身可以作为一种可迁移的知识利用,也可以利用图学习从数据中提取结构化、网络化的可迁移特征. 例如,可以通过图学习来表征工业过程数据中各个变量之间的相关关系或因果关系^[134],从而与语义知识相结合对新发故障的根因进行更深入的分析与诊断. 或者,将语义用图来表示也是未来的研究方向. 目前尚未有利用图学习的工业零样本工作发表,因此在未来的研究中可以考虑引入图学习范式以增强对工业数据、相关知识特性的精细化挖掘,

以促进工业零样本领域的发展.

2) 元学习范式.

元学习^[135](meta-learning)是机器学习领域的一个研究重点,旨在令模型“学会如何学习”. 相比于传统的机器学习,元学习希望模型从已有的一批任务中学习可迁移的特性,从而能够对新任务具有更强的适应能力,因此元学习的适用范围极广. 具体到零样本学习领域,基于元学习范式,可将训练集随机划分为多个支持集和查询集,分别对应已见类和未见类. 模型因此可在训练集中模拟完成多个零样本任务,这有助于从已见类别到未见类别的迁移^[136]. 在工业场景下,由于数据类别和模态较为复杂,工业零样本模型的泛化性值得关注,而元学习的强大迁移能力有可能帮助缓解这一难题.

3) 预训练范式.

预训练技术即在大规模数据集上对模型进行训练,然后在特定任务的小型数据集上进行微调,以缓解欠数据带来的过拟合问题. 由于模型在初期接触过大量数据,会具有较高的泛化能力. 如果这种泛化能力足够强大,模型则无需利用特定任务的小型数据集进一步训练,能够直接完成最终任务,这同样属于零样本学习的研究领域. 近年来,由于图像领域和自然语言处理领域出现了大规模的数据集,在这些领域已经出现了基于预训练的零样本模型,如CLIP(contrastive language-image pre-training)^[131]、GPT(generative pre-training)^[137]等. 然而,在工业领域中,由于数据收集难度大、质量较低、标注难度较高等原因,难以形成如同机器视觉、自然语言处理等计算机领域中类似的大规模数据集,以支持大规模预训练技术,进而使模型获得零样本能力. 这是预训练范式应用到工业对象上需要解决的问题. 后续研究者可探索利用机理仿真模型或者数字孪生技术来解决工业数据的规模问题.

4) 云边协同范式.

云边协同^[138-139]是一种融合物联网、边缘计算等技术的新兴智能制造范式. 通过云端服务器和边端各个节点的交互,能够在云端收集并组织大量数据信息,从而打破工业领域的数据孤岛困境. 对各边缘端而言,由于接触的数据种类相对不丰富,对无法收集到的类别存在零样本学习的需求. 对于统筹各个边端的云端而言,由于获得了多个边端的信息,对不同类别具有更好的泛化性,如果能保证各个边缘端隐私安全的前提下,使各个边端能够基于云端进行一定程度的知识共享,则能够在云边协同框架下完成零样

本应用. 这一方向已有前沿工作^[79]发表,但在隐私保护、云边知识共享等方面尚存问题. 后续研究者可考虑使用联邦学习的相关技术解决零样本学习可能涉及的隐私问题.

5 结论

工业领域的零样本学习研究方兴未艾,仍处于不断探索和发展的阶段,尚未在工业实践中得到广泛应用. 然而,考虑到工业现场高质量训练案例缺乏的困境,研究人员们正在积极探索零样本学习在工业领域的潜在应用. 随着工业智能化进程的推进,工业零样本学习技术必将引起人们越来越多的关注. 工业零样本学习技术在故障诊断、缺陷检测等典型任务中具有重要应用价值,能够有效应对数据收集标注不均衡的问题,降低收集标注数据的成本. 本文针对工业领域的零样本学习研究现状进行了较为全面的文献综述与梳理,介绍了典型的工业零样本学习任务和使用不同类型的辅助知识,并分析了工业领域中零样本学习任务的特征和典型问题. 梳理了工业零样本学习任务,介绍了嵌入式和生成式这两大类工业零样本学习模型及其应用,并梳理了本领域的基准数据集和开源工作. 此外,从知识、数据、标准等方面梳理了工业领域的零样本学习研究目前存在的挑战,并从知识、场景、范式3个方面对工业零样本学习的未来研究进行了展望. 本综述将有望启发工业领域零样本学习的创新思想,推动未来工业智能研究的蓬勃发展.

参考文献(References)

- [1] 柴天佑. 工业人工智能发展方向[J]. 自动化学报, 2020, 46(10): 2005-2012.
(Chai T Y. Development directions of industrial artificial intelligence[J]. Acta Automatica Sinica, 2020, 46(10): 2005-2012.)
- [2] 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.)
- [3] 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.
- [4] Yao Y, Gao F R. Phase and transition based batch process modeling and online monitoring[J]. Journal of Process Control, 2009, 19(5): 816-826.
- [5] Zhao C H, Gao F R. Fault-relevant principal component analysis(FPCA) method for multivariate statistical modeling and process monitoring[J]. Chemometrics and Intelligent Laboratory Systems, 2014, 133: 1-16.
- [6] Wu H, Zhao J S. Deep convolutional neural network model based chemical process fault diagnosis[J]. Computers & Chemical Engineering, 2018, 115: 185-197.
- [7] Yu W K, Zhao C H. Sparse exponential discriminant analysis and its application to fault diagnosis[J]. IEEE Transactions on Industrial Electronics, 2018, 65(7): 5931-5940.
- [8] 康守强, 刘旺辉, 王玉静, 等. 基于深度在线迁移的变负载下滚动轴承故障诊断方法[J]. 控制与决策, 2022, 37(6): 1521-1530.
(Kang S Q, Liu W H, Wang Y J, et al. Fault diagnosis method of rolling bearing under varying loads based on deep online transfer[J]. Control and Decision, 2022, 37(6): 1521-1530.)
- [9] Feng L J, Zhao C H, Sun Y X. Dual attention-based encoder-decoder: A customized sequence-to-sequence learning for soft sensor development[J]. IEEE Transactions on Neural Networks and Learning Systems, 2021, 32(8): 3306-3317.
- [10] Wang Y L, Wu D Z, Yuan X F. A two-layer ensemble learning framework for data-driven soft sensor of the diesel attributes in an industrial hydrocracking process[J]. Journal of Chemometrics, 2019, 33(12): e3185.
- [11] Feng L J, Zhao C H. Adversarial sample based semi-supervised learning for industrial soft sensor[J]. IFAC-Papers OnLine, 2020, 53(2): 11644-11649.
- [12] Sun X H, Gu J N, Wang M M, et al. Wheel hub defects image recognition base on zero-shot learning[J]. Applied Sciences, 2021, 11(4): 1529.
- [13] 赵春晖, 宋鹏宇. 从结构推断到根因识别——工业过程故障根因诊断研究综述[J]. 控制与决策, 2023, 38(8): 2130-2157.
(Zhao C H, Song P Y. From structure inference to root cause identification: A survey on root cause diagnosis of industrial process faults[J]. Control and Decision, 2023, 38(8): 2130-2157.)
- [14] Krizhevsky A, Sutskever I, Hinton G E. ImageNet classification with deep convolutional neural networks[J]. Communications of the ACM, 2017, 60(6): 84-90.
- [15] Deng J, Dong W, Socher R, et al. ImageNet: A large-scale hierarchical image database[C]. IEEE Conference on Computer Vision and Pattern Recognition. Miami, 2009: 248-255.
- [16] Liao Y X, Huang R Y, Li J P, et al. Deep semisupervised domain generalization network for rotary machinery fault diagnosis under variable speed[J]. IEEE Transactions on Instrumentation and Measurement, 2020, 69(10): 8064-8075.
- [17] Zhang S M, Zhao C H, Huang B. Simultaneous static and dynamic analysis for fine-scale identification of process operation statuses[J]. IEEE Transactions on Industrial Informatics, 2019, 15(9): 5320-5329.
- [18] Cao H R, Shao H D, Zhong X, et al. Unsupervised

- domain-share CNN for machine fault transfer diagnosis from steady speeds to time-varying speeds[J]. *Journal of Manufacturing Systems*, 2022, 62: 186-198.
- [19] Chai Z, Zhao C H. A fine-grained adversarial network method for cross-domain industrial fault diagnosis[J]. *IEEE Transactions on Automation Science and Engineering*, 2020, 17(3): 1432-1442.
- [20] Song Y S, Wang T, Cai P Y, et al. A comprehensive survey of few-shot learning: Evolution, applications, challenges, and opportunities[J]. *ACM Computing Surveys*, 2023, 55(13s): 1-40.
- [21] Chen Z H, Chen J L, Liu S, et al. Multi-channel calibrated transformer with shifted windows for few-shot fault diagnosis under sharp speed variation[J]. *ISA Transactions*, 2022, 131: 501-515.
- [22] 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.
- [23] Lampert C H, Nickisch H, Harmeling S. Learning to detect unseen object classes by between-class attribute transfer[C]. *IEEE Conference on Computer Vision and Pattern Recognition*. Miami, 2009: 951-958.
- [24] Feng L J, Zhao C H, Li X. Bias-eliminated semantic refinement for any-shot learning[J]. *IEEE Transactions on Image Processing*, 2022, 31: 2229-2244.
- [25] Chen J L, Pan T Y, Zhou Z T, et al. An adversarial learning framework for zero-shot fault recognition of mechanical systems[C]. *IEEE 17th International Conference on Industrial Informatics*. Helsinki, 2020: 1275-1278.
- [26] Chen X, Zhao C H, Ding J L. Pyramid-type zero-shot learning model with multi-granularity hierarchical attributes for industrial fault diagnosis[J]. *Reliability Engineering & System Safety*, 2023, 240: 109591.
- [27] Nagy A M, Czumi L. Zero-shot learning and classification of steel surface defects[C]. *The 14th International Conference on Machine Vision*. Rome: SPIE, 2022: 386-394.
- [28] Wang W, Zheng V W, Yu H, et al. A survey of zero-shot learning: Settings, methods, and applications[J]. *ACM Transactions on Intelligent Systems and Technology*, 2019, 10(2): 1-37.
- [29] Pourpanah F, Abdar M, Luo Y X, et al. A review of generalized zero-shot learning methods[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2023, 45(4): 4051-4070.
- [30] Rezaei M, Shahidi M. Zero-shot learning and its applications from autonomous vehicles to COVID-19 diagnosis: A review[J]. *Intelligence-Based Medicine*, 2020, 3/4: 100005.
- [31] Zhuo Y, Ge Z Q. Auxiliary information-guided industrial data augmentation for any-shot fault learning and diagnosis[J]. *IEEE Transactions on Industrial Informatics*, 2021, 17(11): 7535-7545.
- [32] Lv H X, Chen J L, Pan T Y, et al. Hybrid attribute conditional adversarial denoising autoencoder for zero-shot classification of mechanical intelligent fault diagnosis[J]. *Applied Soft Computing*, 2020, 95: 106577.
- [33] Xian Y Q, Lampert C H, Schiele B, et al. Zero-shot learning—a comprehensive evaluation of the good, the bad and the ugly[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2019, 41(9): 2251-2265.
- [34] Akata Z, Reed S, Walter D, et al. Evaluation of output embeddings for fine-grained image classification[C]. *IEEE Conference on Computer Vision and Pattern Recognition*. Boston, 2015: 2927-2936.
- [35] Frome A, Corrado G S, Shlens J, et al. Devise: A deep visual-semantic embedding model[C]. *Advances in Neural Information Processing Systems*. Lake Tahoe: Curran Associates, Inc., 2013: 1-9.
- [36] Feng L J, Zhao C H. Transfer increment for generalized zero-shot learning[J]. *IEEE Transactions on Neural Networks and Learning Systems*, 2021, 32(6): 2506-2520.
- [37] Mishra A, Reddy S K, Mittal A, et al. A generative model for zero shot learning using conditional variational autoencoders[C]. *IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*. Salt Lake City, 2018: 2188-2196.
- [38] Xian Y Q, Lorenz T, Schiele B, et al. Feature generating networks for zero-shot learning[C]. *IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Salt Lake City, 2018: 5542-5551.
- [39] Schönfeld E, Ebrahimi S, Sinha S, et al. Generalized zero- and few-shot learning via aligned variational autoencoders[C]. *IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Long Beach, 2020: 8239-8247.
- [40] Chai Z, Zhao C H, Huang B, et al. A deep probabilistic transfer learning framework for soft sensor modeling with missing data[J]. *IEEE Transactions on Neural Networks and Learning Systems*, 2022, 33(12): 7598-7609.
- [41] 柴铮, 汪嘉业, 赵春晖, 等. 面向工业监控典型监督任务的深度迁移学习方法: 现状、挑战与展望[J]. *中国科学: 信息科学*, 2023, 53(5): 821-840.
(Chai Z, Wang J Y, Zhao C H, et al. Deep transfer learning methods for typical supervised tasks in industrial monitoring: State-of-the-art, challenges, and perspectives[J]. *Scientia Sinica: Informationis*, 2023, 53(5): 821-840.)
- [42] Pan S J, Yang Q. A survey on transfer learning[J]. *IEEE Transactions on Knowledge and Data Engineering*, 2010, 22(10): 1345-1359.
- [43] Yu W K, Zhao C H. Recursive exponential slow feature analysis for fine-scale adaptive processes monitoring with comprehensive operation status identification[J]. *IEEE Transactions on Industrial Informatics*, 2019,

- 15(6): 3311-3323.
- [44] Chen J M, Wang G J, Lv J C, et al. Open-set classification for signal diagnosis of machinery sensor in industrial environment[J]. *IEEE Transactions on Industrial Informatics*, 2023, 19(3): 2574-2584.
- [45] Peng P, Lu J X, Xie T Y, et al. Open-set fault diagnosis via supervised contrastive learning with negative out-of-distribution data augmentation[J]. *IEEE Transactions on Industrial Informatics*, 2023, 19(3): 2463-2473.
- [46] 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.
- [47] 冀中, 汪浩然, 于云龙, 等. 零样本图像分类综述: 十年进展[J]. *中国科学: 信息科学*, 2019, 49(10): 1299-1320.
(Ji Z, Wang H R, Yu Y L, et al. Summary of zero sample image classification: Ten years' progress[J]. *Scientia Sinica: Informationis*, 2019, 49(10): 1299-1320.)
- [48] 张鲁宁, 左信, 刘建伟. 零样本学习研究进展[J]. *自动化学报*, 2020, 46(1): 1-23.
(Zhang L N, Zuo X, Liu J W. Research and development on zero-shot learning[J]. *Acta Automatica Sinica*, 2020, 46(1): 1-23.)
- [49] Cao W P, Zhou C, Wu Y H, et al. Research progress of zero-shot learning beyond computer vision[C]. *International Conference on Algorithms and Architectures for Parallel Processing*. Cham: Springer, 2020: 538-551.
- [50] Sun X H, Gu J N, Sun H Y. Research progress of zero-shot learning[J]. *Applied Intelligence*, 2021, 51(6): 3600-3614.
- [51] Chen J Y, Geng Y X, Chen Z, et al. Knowledge-aware zero-shot learning: Survey and perspective[C]. *Proceedings of the 20th International Joint Conference on Artificial Intelligence*. California, 2021: 4366-4373.
- [52] 冯耀功, 于剑, 桑基韬, 等. 基于知识的零样本视觉识别综述[J]. *软件学报*, 2021, 32(2): 370-405.
(Feng Y G, Yu J, Sang J T, et al. Survey on knowledge-based zero-shot visual recognition[J]. *Journal of Software*, 2021, 32(2): 370-405.)
- [53] Chen J H, Zhao C H, Ding J L. A flexible probabilistic framework with concurrent analysis of continuous and categorical data for industrial fault detection and diagnosis[J]. *IEEE Transactions on Industrial Informatics*, 2023, 19(10): 10578-10590.
- [54] Downs J J, Vogel E F. A plant-wide industrial process control problem[J]. *Computers & Chemical Engineering*, 1993, 17(3): 245-255.
- [55] Huang J C, Li Z X, Ye L J, et al. Fault classification of industrial processes based on generalized zero-shot learning[C]. *IEEE 10th Data Driven Control and Learning Systems Conference*. Suzhou, 2021: 887-892.
- [56] Guo Y B, Fan Y M, Xiang Z Y, et al. Zero-sample surface defect detection and classification based on semantic feedback neural network[J/OL]. 2021, arXiv: 2106.07959.
- [57] Ma X, Hu Y, Wang M H, et al. Degradation state partition and compound fault diagnosis of rolling bearing based on personalized multilabel learning[J]. *IEEE Transactions on Instrumentation and Measurement*, 2021, 70: 1-11.
- [58] Xing S B, Lei Y G, Wang S H, et al. A label description space embedded model for zero-shot intelligent diagnosis of mechanical compound faults[J]. *Mechanical Systems and Signal Processing*, 2022, 162: 108036.
- [59] Pan T Y, Chen J L, Xie J S, et al. Deep feature generating network: A new method for intelligent fault detection of mechanical systems under class imbalance[J]. *IEEE Transactions on Industrial Informatics*, 2021, 17(9): 6282-6293.
- [60] Gao Y P, Gao L, Li X Y, et al. A zero-shot learning method for fault diagnosis under unknown working loads[J]. *Journal of Intelligent Manufacturing*, 2020, 31(4): 899-909.
- [61] Ba J L, Swersky K, Fidler S, et al. Predicting deep zero-shot convolutional neural networks using textual descriptions[C]. *IEEE International Conference on Computer Vision*. Santiago, 2016: 4247-4255.
- [62] Elhoseiny M, Saleh B, Elgammal A. Write a classifier: Zero-shot learning using purely textual descriptions[C]. *IEEE International Conference on Computer Vision*. Sydney, 2014: 2584-2591.
- [63] Pennington J, Socher R, Manning C. Glove: Global vectors for word representation[C]. *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing*. Stroudsburg, 2014: 1532-1543.
- [64] Xu J, Li K. Generative zero-shot learning compound fault diagnosis of bearings[C]. *International Conference on Sensing, Measurement & Data Analytics in the era of Artificial Intelligence*. Nanjing, 2022: 1-7.
- [65] 姜娜, 严蜜, 李柠. 基于时序上采样卷积神经网络的风机叶片结冰检测[J]. *控制与决策*, 2022, 37(8): 2017-2025.
(Jiang N, Yan M, Li N. Icing detection of wind turbine blade based on the time-dimensional upsampling convolutional neural network[J]. *Control and Decision*, 2022, 37(8): 2017-2025.)
- [66] Jing H, Zhao C H, Gao F R. Non-stationary data reorganization for weighted wind turbine icing monitoring with Gaussian mixture model[J]. *Computers & Chemical Engineering*, 2021, 147: 107241.
- [67] Wei W W S, Chatfield C. The analysis of time series: An introduction[J]. *Journal of the American Statistical Association*, 1991, 86(413): 246-248.
- [68] Zheng J L, Zhao C H, Gao F R. Retrospective comparison of several typical linear dynamic latent variable models for industrial process monitoring[J]. *Computers & Chemical Engineering*, 2022, 157: 107587.
- [69] Lomov I, Lyubimov M, Makarov I, et al. Fault detection

- in tennessee eastman process with temporal deep learning models[J]. *Journal of Industrial Information Integration*, 2021, 23: 100216.
- [70] Zhao C H. Perspectives on nonstationary process monitoring in the era of industrial artificial intelligence[J]. *Journal of Process Control*, 2022, 116: 255-272.
- [71] Zurita D, Delgado M, Carino J A, et al. Multimodal forecasting methodology applied to industrial process monitoring[J]. *IEEE Transactions on Industrial Informatics*, 2018, 14(2): 494-503.
- [72] Al-Abassi A, Karimipour H, HaddadPajouh H, et al. Industrial big data analytics: Challenges and opportunities[M]. Cham: Springer, 2020: 37-61.
- [73] 乔俊飞, 黄卫民, 丁海旭, 等. 复杂工业过程特征建模方法及应用研究[J]. *控制与决策*, 2023, 38(8): 2063-2078.
(Qiao J F, Huang W M, Ding H X, et al. Research on feature modeling method for complex industrial process and its application[J]. *Control and Decision*, 2023, 38(8): 2063-2078.)
- [74] Zheng J L, Chen X, Zhao C H. Interval-aware probabilistic slow feature analysis for irregular dynamic process monitoring with missing data[J]. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 2023, 53(10): 6553-6564.
- [75] Zhu J L, Ge Z Q, Song Z H, et al. Review and big data perspectives on robust data mining approaches for industrial process modeling with outliers and missing data[J]. *Annual Reviews in Control*, 2018, 46: 107-133.
- [76] 雷亚国, 贾峰, 孔德同, 等. 大数据下机械智能故障诊断的机遇与挑战[J]. *机械工程学报*, 2018, 54(5): 94-104.
(Lei Y G, Jia F, Kong D T, et al. Opportunities and challenges of machinery intelligent fault diagnosis in big data era[J]. *Journal of Mechanical Engineering*, 2018, 54(5): 94-104.)
- [77] 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.
- [78] 刘强, 柴天佑, 秦泗钊, 等. 基于数据和知识的工业过程监视及故障诊断综述[J]. *控制与决策*, 2010, 25(6): 801-807.
(Liu Q, Chai T Y, Qin S Z, et al. Progress of data-driven and knowledge-driven process monitoring and fault diagnosis for industry process[J]. *Control and Decision*, 2010, 25(6): 801-807.)
- [79] Li B X, Zhao C H. Federated zero-shot industrial fault diagnosis with cloud-shared semantic knowledge base[J]. *IEEE Internet of Things Journal*, 2023, 10(13): 11619-11630.
- [80] Guo Y C, Ding G G, Han J G, et al. Zero-shot learning with attribute selection[C]. *Proceedings of the AAAI Conference on Artificial Intelligence*, New Orleans, 2018: 7212-7220.
- [81] Shigeto Y, Suzuki I, Hara K, et al. Ridge regression, hubness, and zero-shot learning[M]. Cham: Springer International Publishing, 2015: 135-151.
- [82] Ge Y H, Abu-El-Haija S, Xin G, et al. Zero-shot synthesis with group-supervised learning[J/OL]. 2021, arXiv: 2009.06586v3.
- [83] Li A X, Luo T G, Lu Z W, et al. Large-scale few-shot learning: Knowledge transfer with class hierarchy[C]. *IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Long Beach, 2020: 7205-7213.
- [84] Wang X D, Mercer R, Rudzicz F. KenMeSH: Knowledge-enhanced end-to-end biomedical text labelling[C]. *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics*. Dublin, 2022: 2941-2951.
- [85] Radovanović M, Nanopoulos A, Ivanović M. Nearest neighbors in high-dimensional data: The emergence and influence of hubs[C]. *Proceedings of the 26th Annual International Conference on Machine Learning*. Montreal, 2009: 865-872.
- [86] Zimek A, Schubert E, Kriegel H P. A survey on unsupervised outlier detection in high-dimensional numerical data[J]. *Statistical Analysis and Data Mining: The ASA Data Science Journal*, 2012, 5(5): 363-387.
- [87] Guo Y C, Ding G G, Han J G, et al. Synthesizing samples for zero-shot learning[C]. *Proceedings of the 26th International Joint Conference on Artificial Intelligence*. Melbourne, 2017: 1774-1780.
- [88] Xu J, Zhou L, Zhao W H, et al. Zero-shot learning for compound fault diagnosis of bearings[J]. *Expert Systems with Applications*, 2022, 190: 116197.
- [89] Lin J, Ye R, Su X L, et al. A novel nonlinear mapping based compatibility method for zero-shot classification in intelligent fault diagnosis[C]. *China Automation Congress*. Beijing, 2022: 5719-5724.
- [90] Huang J C, Li Z X, Zhou Z. A simple framework to generalized zero-shot learning for fault diagnosis of industrial processes[J]. *IEEE/CAA Journal of Automatica Sinica*, 2023, 10(6): 1504-1506.
- [91] Zhang S Y, Wei H L, Ding J L. An effective zero-shot learning approach for intelligent fault detection using 1D CNN[J]. *Applied Intelligence*, 2023, 53(12): 16041-16058.
- [92] Lazaridou A, Dinu G, Baroni M. Hubness and pollution: Delving into cross-space mapping for zero-shot learning[C]. *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing*. Beijing, 2015: 270-280.
- [93] Hu Z W, Zhao H T, Yao L J, et al. Semantic-consistent embedding for zero-shot fault diagnosis[J]. *IEEE Transactions on Industrial Informatics*, 2023, 19(5): 7022-7031.
- [94] Zhao J C, Yue J Q, Feng L J, et al. Addressing domain shift via knowledge space sharing for generalized

- zero-shot industrial fault diagnosis[J/OL]. 2023, arXiv: 2306.02359.
- [95] Lopez R, Regier J, Jordan M I, et al. Information constraints on auto-encoding variational Bayes[J/OL]. 2018, arXiv: 1805.08672.
- [96] Goodfellow I, Pouget-Abadie J, Mirza M, et al. Generative adversarial nets[J]. *Advances in Neural Information Processing Systems*, 2014, 27: 1-15.
- [97] Gao R, Hou X S, Qin J, et al. Zero-VAE-GAN: Generating unseen features for generalized and transductive zero-shot learning[J]. *IEEE Transactions on Image Processing*, 2020, 29: 3665-3680.
- [98] 赵春晖, 胡赞昀, 郑嘉乐, 等. 数据驱动的燃煤发电装备运行工况监控——现状与展望[J]. *自动化学报*, 2022, 48(11): 2611-2633.
(Zhao C H, Hu Y Y, Zheng J L, et al. Data-driven operating monitoring for coal-fired power generation equipment: The state of the art and challenge[J]. *Acta Automatica Sinica*, 2022, 48(11): 2611-2633.)
- [99] Smith W A, Randall R B. Rolling element bearing diagnostics using the Case Western Reserve University data: A benchmark study[J]. *Mechanical Systems and Signal Processing*, 2015, 64/65: 100-131.
- [100] Helwig N, Pignatelli E, Schütze A. Condition monitoring of a complex hydraulic system using multivariate statistics[C]. *IEEE International Instrumentation and Measurement Technology Conference Proceedings*. Pisa, 2015: 210-215.
- [101] Song K C, Yan Y H. A noise robust method based on completed local binary patterns for hot-rolled steel strip surface defects[J]. *Applied Surface Science*, 2013, 285: 858-864.
- [102] Nectoux P, Gouriveau R, Medjaher K, et al. An experimental platform for bearings accelerated degradation tests[C]. *Proceedings of the IEEE International Conference Prognostics Health Manage*. Detroit: IEEE, 2012: 1-8.
- [103] Lei Y G, He Z J, Zi Y Y. Application of a novel hybrid intelligent method to compound fault diagnosis of locomotive roller bearings[J]. *Journal of Vibration and Acoustics*, 2008, 130(3): 1.
- [104] Xing S B, Lei Y G, Wang S H, et al. Distribution-invariant deep belief network for intelligent fault diagnosis of machines under new working conditions[J]. *IEEE Transactions on Industrial Electronics*, 2021, 68(3): 2617-2625.
- [105] Li Z F, Gao L, Gao Y P, et al. Zero-shot surface defect recognition with class knowledge graph[J]. *Advanced Engineering Informatics*, 2022, 54: 101813.
- [106] Feng X L, Gao X W, Luo L. X-SDD: A new benchmark for hot rolled steel strip surface defects detection[J]. *Symmetry*, 2021, 13(4): 706.
- [107] Feng L J, Zhao C H. Industrial-ZSL[DB/OL]. (2021-07-29) [2023-05-10]. https://github.com/chunhuiz/Industrial_ZSL.
- [108] Paul S. Popular-ZSL-algorithms[DB/OL]. (2020-05-19) [2023-05-10]. <https://github.com/mvp18/Popular-ZSL-Algorithms>.
- [109] Hu Z W, Zhao H T, Yao L J, et al. Semantic-consistent embedding for zero-shot fault diagnosis[DB/OL]. (2021-12-22) [2023-05-10]. <https://github.com/htz-ecust/Semantic-consistent-Embedding>.
- [110] Paul A. Feature-generating-networks-for-ZSL[DB/OL]. (2018-12-07) [2023-05-10]. <https://github.com/akku1506/Feature-Generating-Networks-for-ZSL>.
- [111] Schonfeld E. CADA-VAE-PyTorch[DB/OL]. (2020-09-04) [2023-05-10]. <https://github.com/edgarschnfld/CADA-VAE-PyTorch>.
- [112] Feng L J, Zhao J C, Zhao C H. A systematic evaluation and benchmark for embedding-aware generative models: Features, models, and any-shot scenarios[J]. *IEEE Transactions on Image Processing*, 2020, 29: 3665-3680.
- [113] Wiskott L, Sejnowski T J. Slow feature analysis: Unsupervised learning of invariances[J]. *Neural Computation*, 2002, 14(4): 715-770.
- [114] 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.
- [115] Song P Y, Zhao C H, Huang B. SFNet: A slow feature extraction network for parallel linear and nonlinear dynamic process monitoring[J]. *Neurocomputing*, 2022, 488: 359-380.
- [116] Pan H H, He X X, Tang S, et al. An improved bearing fault diagnosis method using one-dimensional CNN and LSTM[J]. *Journal of Mechanical Engineering*, 2018, 64: 443-452.
- [117] Zhou H Y, Zhang S H, Peng J Q, et al. Informer: Beyond efficient transformer for long sequence time-series forecasting[J]. *Proceedings of the AAAI Conference on Artificial Intelligence*, 2021, 35(12): 11106-11115.
- [118] Zhao M X, Wang H, Guo J, et al. Construction of an industrial knowledge graph for unstructured Chinese text learning[J]. *Applied Sciences*, 2019, 9(13): 2720.
- [119] Bordes A, Usunier N, Garcia-Duran A, et al. Translating embeddings for modeling multi-relational data[J]. *Advances in Neural Information Processing Systems*, 2013, 26: 1-10.
- [120] Kampffmeyer M, Chen Y B, Liang X D, et al. Rethinking knowledge graph propagation for zero-shot learning[C]. *IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Long Beach, 2020: 11479-11488.
- [121] Liu H M, Wang R P, Shan S G, et al. What is a tabby? Interpretable model decisions by learning attribute-based classification criteria[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2021, 43(5): 1791-1807.
- [122] Mettes P, Snoek C G M. Spatial-aware object embeddings for zero-shot localization and classification of actions[C]. *IEEE International Conference on*

- Computer Vision. Venice, 2017: 4453-4462.
- [123] Li Z H, Yao L N, Chang X J, et al. Zero-shot event detection via event-adaptive concept relevance mining[J]. Pattern Recognition, 2019, 88: 595-603.
- [124] Hubauer T, Lamparter S, Haase P, et al. Use cases of the industrial knowledge graph at siemens[C]. ISWC (P&D/Industry/BlueSky). Monterey, 2018: 1-2.
- [125] Aly R, Vlachos A, McDonald R. Leveraging type descriptions for zero-shot named entity recognition and classification[C]. Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing. Stroudsburg, 2021: 1516-1528.
- [126] Levy O, Seo M, Choi E, et al. Zero-shot relation extraction via reading comprehension[C]. Proceedings of the 21st Conference on Computational Natural Language Learning. Vancouver, 2017: 333-342.
- [127] Hong T, Fan S. Probabilistic electric load forecasting: A tutorial review[J]. International Journal of Forecasting, 2016, 32(3): 914-938.
- [128] Oreshkin B N, Carпов D, Chapados N, et al. Meta-learning framework with applications to zero-shot time-series forecasting[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2021, 35(10): 9242-9250.
- [129] Feng L J, Zhao C H, Li Y L, et al. Multichannel diffusion graph convolutional network for the prediction of endpoint composition in the converter steelmaking process[J]. IEEE Transactions on Instrumentation and Measurement, 2021, 70: 1-13.
- [130] Verma V K, Brahma D, Rai P. Meta-learning for generalized zero-shot learning[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2020, 34(4): 6062-6069.
- [131] Radford A, Kim J W, Hallacy C, et al. Learning transferable visual models from natural language supervision[C]. Proceedings of the International Conference on Machine Learning. Vienna: PMLR, 2021: 8748-8763.
- [132] Xia F, Sun K, Yu S, et al. Graph learning: A survey[J]. IEEE Transactions on Artificial Intelligence, 2021, 2(2): 109-127.
- [133] Chen Z W, Xu J M, Peng T, et al. Graph convolutional network-based method for fault diagnosis using a hybrid of measurement and prior knowledge[J]. IEEE Transactions on Cybernetics, 2022, 52(9): 9157-9169.
- [134] 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.
- [135] Hospedales T, Antoniou A, Micaelli P, et al. Meta-learning in neural networks: A survey[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2021, 44(9): 5149-5169.
- [136] Liu Z, Li Y, Yao L N, et al. Task aligned generative meta-learning for zero-shot learning[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2021, 35(10): 8723-8731.
- [137] Brown T, Mann B, Ryder N. Language models are few-shot learners[J]. Advances in Neural Information Processing Systems, 2020, 33: 1877-1901.
- [138] Yao Z J, Song P Y, Zhao C H. Finding trustworthy neighbors: Graph aided federated learning for few-shot industrial fault diagnosis with data heterogeneity[J]. Journal of Process Control, 2023, 129: 103038.
- [139] 柴天佑, 程思宇, 李平, 等. 端边云协同的复杂工业过程运行控制智能系统[J]. 控制与决策, 2023, 38(8): 2051-2062.
- (Chai T Y, Cheng S Y, Li P, et al. Intelligent system for operational control of complex industrial process based on end-edge-cloud collaboration[J]. Control and Decision, 2023, 38(8): 2051-2062.)

作者简介

赵健程(1999—), 男, 博士生, 从事工业故障诊断、零样本学习等研究, E-mail: zhaojiancheng@zju.edu.cn;

冯良骏(1995—), 男, 博士, 从事机器学习、人工智能和模式识别、零样本学习等研究, E-mail: liangjunfeng@zju.edu.cn;

岳嘉祺(2000—), 男, 博士生, 从事工业故障诊断、零样本学习等研究, E-mail: 12232051@zju.edu.cn;

张堡霖(2000—), 男, 硕士生, 从事工业故障诊断、零样本学习等研究, E-mail: 22232061@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、状态监测、故障诊断、软测量等研究, E-mail: chhzhao@zju.edu.cn;

付永鹏(1997—), 男, 硕士生, 从事工业故障监测、零样本学习等研究, E-mail: fuyongpeng@zju.edu.cn;

王福利(1957—), 男, 教授, 博士生导师, 从事复杂工业过程智能控制、过程监测与故障诊断、过程建模与优化等研究, E-mail: wangfuli@mail.neu.edu.cn.