

DOI: 10.19650/j.cnki.cjsi.J2311177

不平衡分布的数据驱动故障诊断的研究进展*

李川^{1,2}, 伍依凡^{1,3}, 杨帅¹

(1. 重庆工商大学装备系统服役健康保障研究中心 重庆 400067; 2. 重庆工商大学管理科学与工程学院 重庆 400067; 3. 韦仕敦大学机械与材料工程系 伦敦 N6A 3K7)

摘要:故障诊断是工业系统健康监测的重要内容,现有的数据驱动故障诊断方法多是利用类别平衡的数据集进行建模的。但在实际应用中,工业系统往往产生大量类别不平衡的样本,给数据驱动故障诊断带来挑战。这一问题受到了学术界和工业界的广泛关注,围绕该方面的研究取得了丰硕的成果。但是,目前针对不平衡分布的数据驱动故障诊断的研究进展综述相对较少,因此无法明确现实的挑战以及未来的研究方向。本文针对不平衡分布的约束问题,从数据驱动诊断方法和诊断应用场景这两个方面综述了国内外的研究进展,并提出了面临的挑战及未来的展望,为故障诊断的研究与应用提供参考。

关键词:不平衡分布;数据驱动;故障诊断;智能;工业系统

中图分类号: TH11-2179 **文献标识码:** A **国家标准学科分类代码:** 410.75

Research progress of fault diagnostics driven by imbalanced data distribution

Li Chuan^{1,2}, Wu Yifan^{1,3}, Yang Shuai¹

(1. Research Center for Systems Health Maintenance, Chongqing Technology and Business University, Chongqing 400067, China;
2. College of Management Science and Engineering, Chongqing Technology and Business University, Chongqing 400067, China;
3. Department of Mechanical and Materials Engineering, Western University, London N6A 3K7, Canada)

Abstract: Fault diagnosis is an important part of industrial system health monitoring. Existing data-driven diagnosis methods often use balanced datasets for fault modelling. However, in practical applications, industrial systems often produce many samples with imbalanced distribution, which pose challenges to data-driven fault diagnostics. This issue receives extensive attention from the academic and industrial communities. Many results have been achieved in this area. However, there have been a few reviews on the imbalanced data-driven fault diagnosis. It is difficult to clarify the real challenges and future research directions. In response to this problem, a comprehensive review on the research progress in data-driven diagnostic methods and diagnostic application scenarios is provided. It proposes the challenges and future prospects facing the field, which could provide a reference for the research and application of the fault diagnostics.

Keywords: imbalanced distribution; data-driven model; fault diagnosis; intelligence; industrial systems

0 引言

工业设备在复杂的工况中长期连续工作,很容易发生故障,引起设备异常甚至发生严重的生产事故。因此,对工业设备进行故障诊断,是系统健康维护的基本要求,

同时也是成本节约、提升生产效率的有效手段^[1-2]。数据驱动故障诊断采用数据分析的方法,对故障模式进行识别,并从状态监测的传感器数据中提取有用的健康状态信息^[3-4]。当诊断模型的训练数据集内不同状态类别分布不平衡时,需要采取有效方法确保每个状态都能够得到有效诊断和识别,因此不平衡分布的数据驱动故障诊断的研究越来越受到重视。

收稿日期:2023-03-14 Received Date: 2023-03-14

* 基金项目:国家自然科学基金(52175080)、重庆市自然科学基金创新群体项目(cstc2019jcyj-cxttX0003)、重庆工商大学研究生创新型科研项目(yjsx2021-112-84)资助

随着现代工业的发展和传感技术的进步,故障诊断的传感器数据已经非常容易获得,但这进一步增加了故障诊断的复杂性。数据驱动诊断模型是建立在充分的传感数据分析基础上的。训练样本数据越充足,故障类型越丰富,越有利于提高故障诊断的准确性。然而在实际工程中,设备系统正常工作时间多,故障案例相对较少。即使传感器可以不断地在系统运转过程中采集数据,大部分采集到的数据也都是健康状态数据,而故障数据很少甚至缺失某些故障类别的数据。此外,工况的复杂性还会产生大量不知故障类别的数据。而若在实验室搭建故障模拟试验台或者使用仿真软件模拟系统故障,存在无法模拟复杂的工况以及和实际产生的数据区别大等问题。因此,不平衡分布样本的现实情况给数据驱动诊断的模型训练带来严峻挑战,是学术界和工业界亟需研究的重要方向。直接使用不平衡数据进行故障诊断,往往存在诊断精度低、泛化性能差的问题,难以满足故障诊断的要求。

虽然不平衡数据在许多领域都是普遍存在的,但是目前国内关于不平衡分布的相关系统分析仍然较少。周玉等^[5]对不平衡数据集分类方法进行了综述,李昂等^[6]和李蒙蒙等^[7]从算法原理出发,对不平衡数据算法进行了分类和阐释,Zhang等^[8]对不平衡数据驱动诊断算法进行了具体阐释。这些研究主要从算法层面面对各模型进行说明,但对这些方法的实际应用范围和效果的说明不多,缺少对不平衡数据方法在故障诊断中具体应用场景的详细梳理和系统总结。随着新一代人工智能的发展,不断有新技术被用于不平衡数据的研究中,在促进诊断智能化的同时,还提供了新的不平衡样本分布约束的解决思路。因此,对现有不平衡分布的数据驱动故障诊断的研究进展进行综述是非常必要的,亟需总结本领域现有的成果,并探索未来的发展方向。

鉴于此,本文面向工业系统故障诊断,通过回顾 10 多年来的相关文献,将不平衡数据驱动故障诊断方法进行了分类,并重点介绍不平衡数据驱动故障诊断研究最新成果。从不平衡分布自身的特点出发梳理研究进展,将不平衡数据驱动故障诊断划分为样本不均约束的故障诊断、样本稀少约束的故障诊断、仅有部分标签约束的故障诊断、新类识别约束的故障诊断等 4 个类别。这不仅使本综述具有更强的领域特征,对该领域的研究人员具有更大的启发意义,而且还明确了不平衡数据的研究范畴和应用场景,为不平衡分布的数据驱动故障诊断研究提供了未来的发展方向。

1 文献收集

为了全面了解不平衡样本分布下数据驱动故障诊断的研究进展,按照文献学的方法,分别从中国知网(China national knowledge internet, CNKI)、知识网(Web of Knowledge)、工程索引(engineering index)、美国电气与电子工程师学会数据库(the institute of electrical and electronics engineers, IEEE)和斯高帕斯(Scopus)数据库中检索 2012 年~2022 年发表的相关文献。

检索结果显示,自 2012 年以来,已有越来越多关于不平衡数据驱动故障诊断的论文发表。由于研究对象是工业系统的智能故障诊断,因此,通过人工检查剔除了文本等自然语言的故障诊断。在检索中,出现了不属于本文研究范围内的文献,也人工进行剔除。图 1 示出了 2012~2022 年发表的学术期刊论文和会议论文的时间分布。删除重复文献后,该研究领域在近 11 年共有 119 篇中文文献和 423 篇英文文献发表。

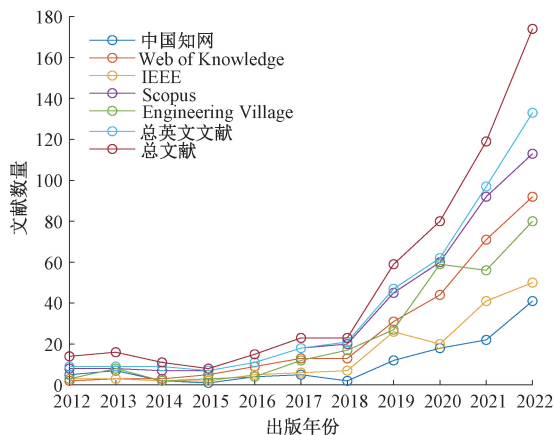


图 1 2012~2022 年发表的本领域论文的分布

Fig. 1 The distribution of the related articles published between 2012 and 2022

如图 1 所示,本领域的文献公开数量在 2018 年之前比较平稳,表明还没有引起故障诊断领域的充分重视。从 2019 年开始,不平衡分布的数据驱动故障诊断的研究论文数量大幅增加,而且中文文献与英文文献的增加趋势是一致的。这表明越来越多的领域专家开始关注不平衡数据分布对数据驱动故障诊断的约束作用,同时不断有较新的解决方案得以公开。尤其是近 3 年,每年在本领域发表的论文数量基本上以大约 50% 的速度增加,表明不平衡数据驱动故障诊断方法已经成为近年来智能故障诊断领域的研究热点。

不平衡数据驱动的故障诊断方法与工业应用结合紧密,典型的不平衡数据驱动故障诊断的工业应用场景如表 1 所示。由表中可以看出,不平衡分布的故障诊断在

机械系统、电气系统、精密仪器、其他系统等均有应用,表明不平衡分布的故障诊断问题是普遍存在、亟需解决的现实工程问题。

表 1 不平衡分布的数据驱动故障诊断典型工业应用
Table 1 Application of fault diagnostics driven by imbalanced data distribution

工业应用	典型文献
机械系统	风力涡轮机 ^[9-12] 、高速列车制动系统 ^[9,13] 、船舶柴油机 ^[14] 、动车组轴箱 ^[15] 、车辆传动系统 ^[16] 、全超导托卡马克分子泵 ^[17] 、炼油厂旋转机械 ^[18]
电气系统	变压器 ^[8,19-25] 、永磁同步电机 ^[26] 、电力线路 ^[27] 、高压电机 ^[28]
精密仪器	3D 打印机 ^[29-30] 、工业机器人 ^[31-32]
其他系统	冷水机组 ^[33-34] 、暖通系统 ^[35] 、空调 ^[36]

2 故障诊断的不平衡数据分布场景

故障诊断训练数据集由样本和标签组成,根据样本不平衡或者标签不平衡,可以将不平衡数据驱动故障诊断划分为样本不均约束的故障诊断、样本稀少约束的故障诊断、仅有部分标签约束的故障诊断和新类识别约束的故障诊断等 4 种类别。其中,样本不均和样本稀少涉及数据集内的样本不平衡,而仅有部分标签和新类识别问题涉及数据集中标签的不平衡。因此,本文总结的这 4 种类别在总体上具有显著的区分度。

一般来说,平衡数据集包含一类健康状态样本及 N 类故障样本,每类均有 M 个数据。因此,样本集合表示为 $\{S\} = \{S_0, S_1, S_2, \dots, S_N\}$ 类样本集可表示为:

$$S_j = \{(x_j^i, y_j) \mid i = 1, 2, \dots, M; j = 0, 1, 2, \dots, N\}$$
 (1)
式中: x_j^i 为第 j 类中第 i 组数据, y_j 是样本的类别标签。

受到部件运行状态、实际工况、传感器质量的影响,采集到的样本一般是不平衡的,第 1 个场景是样本数据不均,也就是是指不同类别的样本数量不均等的数据集,包括样本间极度不平衡的情况,可表示为:

$$\exists a < M,$$

$$\text{s. t. } S_j = \{(x_j^i, y_j) \mid i = 1, 2 \cdots, a; j = 0, 1, 2, \dots, N\}$$
 (2)

式中: a 是不平衡类别的样本数据数量。第 2 个场景是样本稀少,是指由于工业系统运转环境复杂,可用样本数据少,甚至某类样本仅有一个的情况,表示为:

$$\forall a \leq a_l$$

$$\text{s. t. } S_j = \{(x_j^i, y_j) \mid i = 1, 2 \cdots, a; j = 0, 1, 2, \dots, N\}$$
 (3)

式中: a_l 表示稀少的样本数据数量(例如故障样本数少于 30 个)。第 3 个场景是部分标签,指的是样本集中包

含了有标签样本和无标签样本,可表示为:

$$\exists y_j \in \emptyset,$$

$$\text{s. t. } S_j = \{(x_j^i) \mid i = 1, 2, \dots, a; j = 0, 1, 2, \dots, N\}$$
 (4)

式中: $\emptyset$ 表示部分空集。第 4 个场景是新类识别,指的是测试集中含有训练集中未包含的新类别样本,并且其不属于训练集中已知标签的任何一类,这是区别于训练集中出现的样本标签的一类新样本,可表示为:

$$\exists S_{M+1} \not\subset S,$$

$$\text{s. t. } S_{M+1} = \{(x_{M+1}^i, y_{M+1}) \mid i = 1, 2, \dots, M\}$$
 (5)

近年来报道的典型文献如表 2 所示,下面分别对这些不平衡数据故障诊断的场景进行介绍。

表 2 不平衡分布的数据驱动故障诊断典型文献
Table 2 Literatures of fault diagnostics driven by imbalanced data distribution

位置	具体不平衡场景	典型文献
样本不平衡	样本不均约束的故障诊断	文献[6-12, 14-15, 17, 19-20, 23-25, 27, 32-34, 37-54]
	样本稀少约束的故障诊断	文献[16, 20, 26, 29, 55-70]
标签不平衡	仅有部分标签约束的故障诊断	文献[22, 71-73]
	新类识别约束的故障诊断	文献[30-31, 74-78]

2.1 样本不均约束的故障诊断

张永宏等^[79]对滚动轴承进行故障诊断,设置了 3 种单一故障和 1 种复合故障,其中正常样本为 2 400 个,故障样本分别为 60 个,不平衡比例高达 1:40,样本集极度不平衡。张笑璐等^[20]在滚动轴承故障诊断中设置了 4 种单一故障,其中一种故障数量为 50 个,其余类别数量分别为 1 400 个。王艳等^[19]对变压器进行故障诊断,包括 6 类样本,其中正常样本为 400 个,故障样本最少为 84 个,最多为 144 个。栗磊等^[55]对变压器故障进行诊断,一般样本集数为 600,最少的样本集为 50。刘云鹏等^[21]对变压器进行故障诊断,样本集中包含 600 个正常数据,7 种故障数据,其中故障数据数量最低为 62、最高为 73。Qian 等^[9]对风力涡轮机叶片开裂故障和高速列车制动系统故障进行了诊断,其中风力涡轮机叶片数据不平衡比分别为 17.2:1、20.11:1 和 25.09:1,而高速列车制动数据不平衡比为 56.5:1。

2.2 样本稀少约束的故障诊断

张笑璐等^[29]对滚动轴承进行故障诊断,设置了 1 类正常样本和 4 类故障样本,每类样本仅 20 个进行了实验。又考虑了极端情况,即每类只有一个已知样本,进行

了第二次实验。李亮等^[56]对电力变压器故障进行诊断,训练集中包括 150 组正常样本、4 类故障样本,每类样本 20 个。Chen 等^[71]对轴承故障进行诊断,设置了不同工况和故障下多个数据集,其中健康状态样本数均为 60,故障状态样本数分别为 5、10 和 30。Li 等^[72]对 3D 打印机的重要零部件皮带和关节轴承进行故障诊断,将正常样本和每一个故障仅提供一个样本对模型进行训练,分别测试了 3 类皮带故障、4 类轴承故障和 7 类混合故障,取得了出色的效果。Xia 等^[22]对三缸泵进行故障诊断,使用源域数据训练模型后,仅用目标域中一个样本对模型进行微调,实现模型迁移学习。

2.3 仅有部分标签约束的故障诊断

在数据驱动故障诊断中,由于数据标签的获取代价相对昂贵,因此存在大量无标签的数据,这种利用少量已有标签数据识别大量无标签数据的情形也是不平衡数据集的一类。An 等^[74]使用迁移学习进行部分标签下轴承故障和齿轮箱故障的诊断,其中源域为平衡数量的标签样本,目标域训练集为不平衡数量的无标签样本,测试集为平衡数量的无标签样本。Zong 等^[75]在使用迁移学习对轴承数据集进行故障诊断时,采用两个不同的基准数据集,一个数据集作为源域,均为标签样本,另一个数据集作为目标域,包含少量标签样本和大量无标签样本。Yang 等^[76]选取基准齿轮箱数据集每类故障 30 个标签样本作为源域,以自建试验台的齿轮箱无标签样本作为目标域,对齿轮箱故障进行跨域迁移学习。

2.4 新类识别约束的故障诊断

Long 等^[31]对轴承及齿轮箱故障进行诊断,使用大量的标签样本对模型进行元学习,使用少量和前样本集不同类别的标签样本作为支持集,使用少量和前两个样本集不同类别的无标签样本作为测试集。赵晓平等^[77]对轴承故障进行诊断,通过对已知标签的样本学习,完成对未知标签样本的识别,其训练集和验证集的样本标签是完全不同的。张永宏等^[78]在 27 类基准轴承故障样本中随机选取的 21 类故障样本作为训练集,剩余 6 类故障样本作为测试集进行实验,共得到了 4 个数据集。Pu 等^[30]对工业机器人齿轮进行故障诊断,共有 3 种故障模式,分别使用仅健康状态样本和使用健康状态及其中一种故障样本对模型进行训练,输入所有样本进行故障识别。Xu 等^[26]使用 3 种有标签的单故障类型的轴承振动信号进行训练,以识别无标签的四种复合的故障。Xing 等^[57]训练单一故障来识别轴承、行星齿轮箱的复合故障。Li 等^[37]对 3D 打印机关节轴承和皮带的故障进行诊断,仅使用健康状态样本训练模型,每个故障类型选取 2 个样本作为验证集,每个故障类型选取 1 个样本作为测试集,即测试集中包括大量健康状态样本及极少数故障样本。

3 不平衡分布的数据驱动诊断算法

根据工业设备故障诊断的一般过程,不平衡数据驱动故障诊断算法总体上包括基于重采样的故障诊断方法、基于生成模型的故障诊断方法、基于特征学习的故障诊断方法和基于分类器优化的故障诊断方法。在模型构建层面,可以将不平衡分布的故障诊断算法归类为基于多模型学习的故障诊断方法,此外还有部分组合算法。这些算法的内在关系如图 2 所示,下面分别进行介绍。

3.1 重采样算法解决不平衡数据分布的问题

重采样算法是应对不平衡数据最为基础的方法,也是针对数据进行处理的方法。通过重采样使数据集达到平衡,再利用平衡后的数据集开展故障诊断训练建模。常用的重采样方法分为欠采样和过采样,以及欠采样与过采样相结合方法。

1) 欠采样故障诊断

欠采样通过某种方法删除多数类中的一些样本,使多数类样本规模减小,从而和少数类样本的数量达到平衡。在欠采样删除多数类样本方法中,最简单的是随机欠采样,通过随机丢弃多数类中的样本,使数据达到平衡。但这种方法可能会导致多数类样本的重要特征丢失,使得模型不能充分学习类别特征,降低了模型的识别准确性。因此,研究对在减少多数类样本的同时,保留其内在特征的方法进行了研究。王德成等^[38]提出了一种通过计算多数类样本到少数类样本中心的距离来筛选多数类样本的方法,该方法让远离最优分类超平面的样本被剔除,从而使样本间达到平衡。陶新民等^[23]先对正常样本进行聚类,从聚类结果中根据样本到聚类中心的距离选择合适的样本,剔除距离聚类中心远的样本,从而使不同类别的样本达到平衡。Qian 等^[9]使用聚类算法对不平衡数据集进行聚类,得到聚类误差,使用柔性致动评价(Soft actor-critic, SAC)强化学习法确定决策边界,优化采样过程,根据不平衡率确定采样权值。Shi 等^[24]使用一种基于线性判别分析(linear discriminant analysis, LDA)的欠采样方法,将实际距离投影到 LDA 的分类方向上,根据投影距离选取有代表性的多数类样本。Jang 等^[10]使用一种非结构化边界自组织映射模型,通过部分节点的连接确定整体数据的分布,从而提高边界数据的权重,突出边缘区域的数据。

2) 过采样故障诊断

过采样通过算法增加少数类中的样本,使少数类样本数量增加,从而和多数类样本的数量达到平衡,再进行故障诊断的方法。在过采样算法中,一个代表算法是合成少数类过采样技术(synthetic minority over-sampling technique, SMOTE)算法,它是在少数类相邻样本间通过

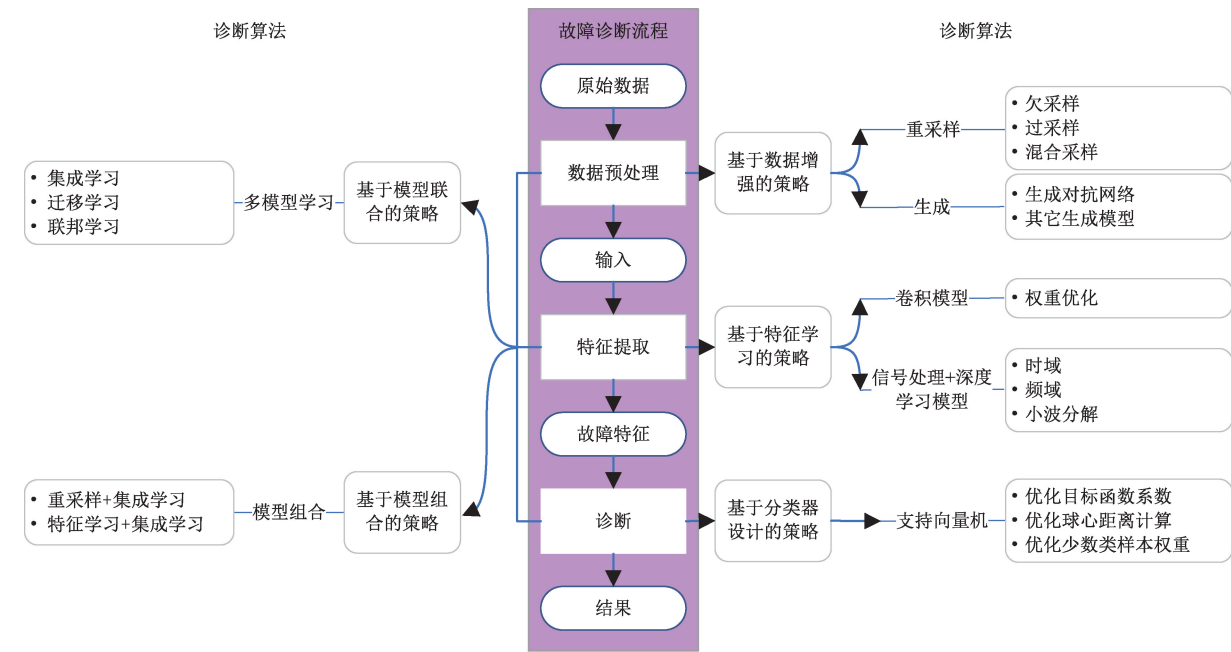


图 2 不平衡数据驱动故障诊断算法的基础类别

Fig. 2 Different classes of fault diagnostics driven by imbalanced data distribution

随机线性插值来合成一定数量的新样本,使少数类样本与多数类样本在数量达到平衡。

由于传统的 SMOTE 算法使用随机插值的方法,容易合成噪声样本,影响分类器学习,许多文献对 SMOTE 算法进行了改进。为了 SMOTE 算法尽可能少受噪声样本的影响,增加类内样本特征相似性和类间样本特征差异性,一些学者认为,分类器需要更好地学习边界,因此合成样本的数量和位置至关重要。余松等^[39]使用托梅克链接移除 (Tomek link removal, TLR) 和自适应样本合成方法,首先通过 TLR 方法对少数类样本去噪,然后将每个少数类样本找到其 K 近邻距离内的多数类样本,以确定需要合成的样本数量,再使用 SMOTE 法对少数类样本数量进行扩充,这种方法使少数类和多数类类别间的边界附近合成更多样本,而不是在少数类别内部。此外,刘云鹏等^[21]同样采取在类边界附近合成样本的理念,采用支持向量机作为决策机制,对少数类的支持向量样本选择 K 近邻,若 K 近邻都为多数类样本,则该少数类样本为噪声样本,否则对其进行插值,合成新样本。陶新民等^[80]提出了一种精简集约简上采样算法,同样采用支持向量机作为决策机制,针对支持向量机依赖边界样本的特征,对少数类的边界样本进行生成,其实验结果表明,相比于随机上取样方法、SMOTE 方法,在不平衡比例逐渐增大时能够获得更好的诊断效果。Jiang 等^[81]结合 SMOTE 设计了一种改进的过采样算法,考虑了数据过采样的时间依赖性和样本之间的关系。

3) 混合采样故障诊断

混合采样是一些将欠采样和过采样相结合方法。李亮等^[56]将 SMOTE 和欠采样方法结合起来,对少类样本使用 SMOTE 增加样本个数,对多数类样本使用最近邻欠采样减少样本个数,从而减少样本间的不平衡性。实验表明同时使用两种采样方式比只用一种采样方式,诊断的准确率得到了提升。孟宗等^[82]将随机欠采样用于多数类数据集、滑动窗口法用于少数类数据集以重构数据。Chen 等^[71]按重要性概率对原始数据集进行上下采样,上采样在同类别少数类样本距离间生成一个新样本,位置由两个父样本的接受率决定。通过下采样丢弃边界样本,保留接近类中心的样本。

4) 重采样算法的发展

回顾已有文献,从 2012 年起,重采样算法就被用于解决不平衡数据分布的故障诊断问题。在 2014 年以前,欠采样法作为主要方法用在不平衡分布的故障诊断研究中。2014 年,SMOTE 算法被用于不平衡数据下的故障诊断研究^[11,58]。此后,SMOTE 算法被广泛用于解决不平衡数据分布的故障诊断研究,直到 2022 年,仍有相当文献报道 SMOTE 算法解决不平衡诊断问题。

3.2 生成模型实现不同类别数据平衡化

近年来,以生成式对抗网络和变分自编码器为代表的生成模型受到广泛研究,并成功用于数据驱动的故障诊断中。生成模型是通过深度神经网络生成新样本,与原有样本进行对比学习,区别于传统的采样技术。其应用非常广泛,下面是典型的生成模型算法。

1) 生成对抗网络算法

生成式对抗网络是一种最常见的生成模型,在研究过程中衍生出许多变体,但其主要结构包括一个生成器和一个判别器,均由多层神经网络构成,具体如图3所示。

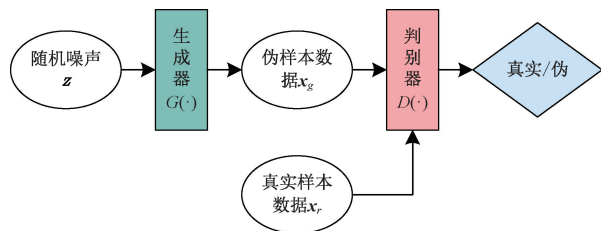


图3 生成对抗网络的基本结构

Fig. 3 The basic architecture of generative adversarial networks

由图3可知,生成器 $G(\cdot)$ 根据样本 X 的分布 p_x 取一组随机数据 z ,通过深度神经网络生成数据 x_g ;判别器 $D(\cdot)$ 通过输出一个概率值判断生成数据 x_g 和真实数据 x_r 。生成器目的是生成一个以假乱真的数据,而判别器的目的是分辨真实的数据和生成的数据,其目标函数表示如下:

$$\min_G \max_D W_{G,D} = E_{x \sim p_x} [\log D(x)] + E_{z \sim p_z} [\log(1 - D(G(z)))] \quad (6)$$

$$L_G = E_{z \sim p_z} [\log(1 - D(G(z)))] \quad (7)$$

$$L_D = -E_{x \sim p_x} [\log D(x)] - E_{z \sim p_z} [\log(1 - D(G(z)))] \quad (8)$$

其中, L_G 是生成器的目标函数, L_D 是判别器的目标函数。由于生成器和判别器在对抗过程中存在性能不匹配的问题,可能导致模型收敛速度慢或梯度消失,需要对生成式对抗网络进行优化。

常见的优化方法包括神经网络优化、损失函数优化以及目标函数优化。梯度惩罚瓦瑟斯坦(Wasserstein)生成对抗网络是一种对生成对抗网络的目标函数进行优化的模型。2019年生成对抗网络率先被报道用于解决不平衡分布的故障诊断问题,就采用了瓦瑟斯坦生成对抗网络^[59]。它是用瓦瑟斯坦距离替代詹森-香农(Jensen-Shannon)散度,并且通过引入一个梯度惩罚系数帮助判别器更加平滑,此方法可以解决生成器随判别器训练效果增加而梯度消失的问题,是一种有效的不平衡样本故障诊断生成式对抗网络^[25,32,40-42,60]。2022年开始,学者们加强对Wasserstein生成对抗网络研究。李慧芳等^[33]在Wasserstein生成对抗网络的基础上,针对样本不平衡的问题,使用委员会查询法和多样性评价指标对生成器生成的样本进行二次筛选,以增加某些类别的样本。Pu等^[61]在梯度惩罚Wasserstein生成对抗网络的基础上,采用基于Vapnik V矩阵的准则用于推广均方误差,并提出

了一种新的早期停止策略,在训练过程中跟踪最合适的模型。Yu等^[43]在在梯度惩罚Wasserstein生成对抗网络的基础上,提出了一种并行分类的架构,针对每个故障类别分别使用独立的梯度惩罚Wasserstein生成对抗网络,并使用皮尔逊(Pearson)损失函数和分离性(Separability)损失函数来改进了模型损失函数,提高故障诊断能力。

此外,为了应对不平衡分布的故障诊断,还有对神经网络和损失函数进行优化的生成对抗网络。张笑璐等^[29]将学习率设置为每500次训练衰减5%,以提高模型训练速度,并在模型中添加退出(Dropout)层防止过拟合。李冬冬等^[40]在神经网络中加入退出(Dropout)层并使用贝叶斯优化调整判别器超参数。高学金等^[44]使用最小二乘法替代传统的交叉熵法,减少类内距。杨光友等^[45]在生成式对抗网络的损失函数中引入自适应函数,通过衡量生成器和判别器的性能调整生成器的权重,使其性能相对稳定。

2) 其他生成模型故障诊断算法

变分自编码器也是一种常见的生成模型,它是从标准正态分布中抽取随机数,基于输入样本的均值和方差生成隐藏变量,再解码为输出样本,通过不断训练缩小输入样本和输出样本的差距。张永宏等^[79]将变分自编码器和生成式对抗网络相结合,将变分自编码器的隐藏变量放入生成器中生成新样本,再将其放入判别器中对抗训练。Wang等^[62]提出基于条件变分自编码网络-生成式对抗网络的方法,利用条件变分自编码器的编码器网络获取故障样本的分布,通过编码器网络生成故障样本后,利用对抗性学习机制对生成器、鉴别器和分类器的参数进行持续优化。Zhao等^[63]使用标准化条件变分自编码器增强模型特征学习能力,并使用一种自适应的Focal损失函数重点训练难分类的样本,以提高不平衡数据集的诊断能力。Zhao等^[83]提出将采集到的振动信号输入深度拉普拉斯自编码器算法进行分层特征提取,将提取的特征反向传播,输入分类器进行诊断,拉普拉斯正则化项用于平滑编码器中数据的流形结构。Li等^[84]提出了一种结合预测生成降噪自编码器(predictive generative denoising autoencoder, PGDAE)和深度珊瑚网络(deep coral network, DCN)的方法,其中PGDAE是由门控循环单元构成,用于生成少数类样本,DCN用于故障识别,由具有相关对齐的深度卷积神经网络构成。Liu等^[46]提出了一种基于深度遗憾分析的变分自编码器-生成式对抗网络数据增强方法。在生成式对抗网络中加入编码器提高样本生成质量,将判别器与深度遗憾分析相结合,通过对判别器施加梯度惩罚来避免模式崩溃。

3.3 特征学习解决数据分布不平衡问题

相比于采样法和生成模型法在数据层面增加或减少样本达到平衡效果,特征学习是从数据中隐藏特征的层

面提取和选择特征,使特征能够正确反应样本状况,从而提升诊断性能。2016年,特征学习被用于解决数据分布不平衡问题^[64]。2018年,深度学习被用于自动提取不平衡数据内在特征^[47]。

卷积神经网络是不平衡数据集故障诊断中特征提取的常用方法。董勋等^[48]、曹洁等^[49]、张洪等^[50]、杜小磊等^[51]和吴耀春等^[65]对卷积神经网络中的损失函数进行了优化,通过增加某些类样本的权重实现不平衡数据集故障诊断。韩淞宇等^[14]不仅优化了卷积神经网络中的损失函数,还提取了卷积神经网络中多尺度特征,并通过自适应权重单元对特征进一步选择。康守强等^[15]提出一种基于强化学习范式的残差网络提取振动信号深度特征。

2020年,Wang等^[85]将小波包分解与卷积神经网络相结合,杜小磊等^[51]将小波变换与分形网络相结合用于提取特征进行不平衡数据故障诊断,这是传统信号处理方法与深度学习相结合的特征提取方法的开始。此后,王瑞涵等^[66]提取均方值、裕度、K阶原点矩3个特征后,构造概率拓扑图输入到图卷积神经网络中预测标签。Wang等^[85]使用小波包分解在频域挖掘时间和频率尺度的信息,采用局部卷积和全卷积相结合的改进双层卷积神经网络进行特征提取和故障检测。

除了传统信号处理和深度学习方法以外,宋志坤等^[67]采用一种形状小波(Shapelets)的信号分类方法应用到不平衡数据集中,并结合机器学习方法优化目标函数加快模型收敛,其实验结果表明该方法拥有高的分类精度。

3.4 类别数据不平衡条件下的分类器优化

前面介绍了采样方法、生成模型方法和特征选择与评价方法,但在不平衡数据的故障诊断中,大部分属于分类任务。因此,本节介绍对分类器优化的研究进展。通过提高分类准确率,提升不平衡数据约束的诊断性能。在收集的文献中,相当部分研究使用支持向量机作为故障的分类器。即使在最早的文献中,支持向量机就用于数据不平衡的故障诊断^[16]。支持向量机是通过构造最优线性分类超平面来划分样本类别的分类方法。对于线性不可分的样本问题,支持向量机通过非线性变换把样本映射到高维空间,通过寻找最小超球体的方式确定最优分类超平面,如图4所示。

假设对于给定数据集 $\{x_i, y_i\}, i = 1, 2, \dots, N, y_i \in \{-1, 1\}, x_i \in \mathbb{R}^n$, 其目标函数可表示为:

$$\min \left(\frac{1}{2} \|w\|^2 + C \sum_{i=1}^N \xi_i \right) \quad (9)$$

s. t. $y_i(w \cdot \varphi(x_i) + b) \geq 1 - \xi_i, \xi_i \geq 0$

其中, C 为惩罚因子, ξ 为松弛变量, B 为分类阈值,

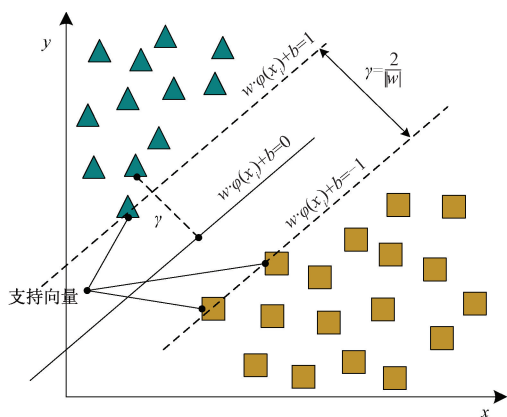


图4 支持向量机示意图

Fig. 4 The illustration of support vector machines

W 为权重张量, $\varphi(\cdot)$ 为非线性映射。

从2012—2022年间,支持向量机模型的发展主要包括两个方面,具体为对目标函数的优化和距离计算公式的改进。自2019年以来,主要在对目标函数的系数进行优化^[52,59,68],增加少数类样本的权重,减少数据不平衡带来的不利影响,从而提升模型的分类准确率。例如,位晶晶等^[68]除了引入惩罚因子项外,还为每个样本分配不同权重,以减少某些类样本被错误分类的概率,抵消不平衡数据引起的分类边界偏差。此外,还引入改进粒子群优化算法为支持向量机找到最佳惩罚因子及核函数参数。陶新民等^[17]在目标函数中样本到球心距离项增加了一个正则化项系数,防止分类边界向多数类偏移。

3.5 多模型学习解决数据不平衡约束

1) 集成学习算法

集成学习是不平衡数据故障诊断常见的方法,使用多个弱分类器对数据集进行诊断,然后使用某种策略将这些诊断结果组合起来作为最终的诊断结果。集成学习的思想是考虑到单一分类器可能鲁棒性不佳,故使用多个分类器共同帮助故障诊断。常见的集成学习有套袋(Bagging)学习、提升(Boosting)学习和堆积(Stacking)学习。

套袋(Bagging)学习是将数据集划分为若干个数据子集,每个数据子集对应一个弱分类器,各个弱分类器互不依赖,并行运行,如图5所示。其代表是随机森林算法和简易集成(Easy Ensemble)算法。

Bagging学习常常用于抽样策略的优化,例如针对极端不平衡数据集,Bagging学习在抽样时可能不会抽到某些类样本的问题。姜万录等^[86]采用了一种平衡随机森林算法,在模型训练时根据不同类样本的数量取相同数目的样本组成均衡的数据集,并在节点分裂时随机选取一个属性作为分裂变量,优化了随机森林的参数选择。

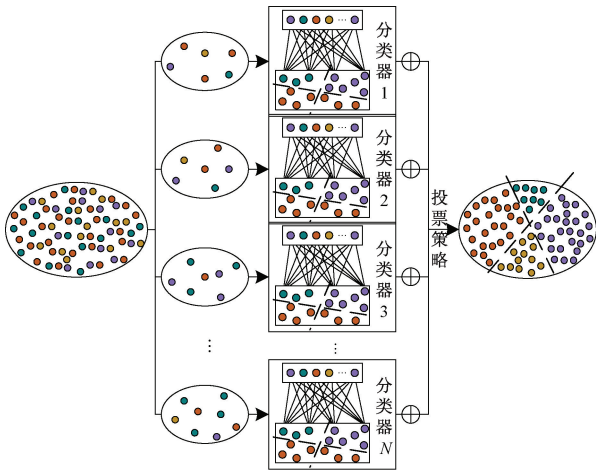


图5 Bagging 算法示意图

Fig. 5 The illustration of Bagging algorithm

其结果表明比随机森林算法、SMOTE、随机森林组合算法具有更高的分类精度。Bagging 学习也可以和重采样法相结合,对样本的选择进行优化。例如 Xu 等^[87]提取样本特征后,使用 Easy Ensemble 算法对多数类样本进行欠采样,构建平衡的多个训练子集分别训练,再将分类器集成成为最终的强分类器。

近年来,深度学习被广泛应用于不平衡数据分布的故障诊断研究中,Bagging 学习也与深度学习相结合,具体表现在使用深度网络作为基分类器。例如,深度学习拥有自助训练和有效处理高维数据的优点,但需要大量数据进行训练。张笑璐等^[20]针对这一问题,提出 Bagging 学习可以将在有限的数据下创造多个训练子集,可以解决深度学习的不足。将一维卷积神经网络作为弱分类器,使用 Bagging 思想进行集成学习,在不平衡数据集上得到 95% 的诊断结果。

在集成学习的 3 大基本模型中,提升 (Boosting) 学习在不平衡数据分布的故障诊断问题中应用最广泛。Boosting 学习是将数据集依次放入不同的弱分类器中进行迭代,每次迭代为样本和弱分类器增加不同的权重,最终将各个分类器按权重不同组合成一个强分类器进行诊断。如图 6 所示,Boosting 学习中的弱分类器是串行的,相互之间有依赖关系,其代表算法有自适应提升 (Adaboost) 和 lightGBM 算法。贾凯等^[53]对轻量梯度提升机 (light gradient boosting machine, lightGBM) 算法进行了优化,提出了一种基于代价敏感的 lightGBM 算法,在弱分类器中引入误分类代价指标,以平均总代价最小化为优化目标,在迭代中对不同类别赋予不同权重。Jiang 等^[88]使用不同权重的弱分类器,即对弱分类器中损失函数中少数类对应的项赋予较高的权重,以增强少数类样本的影响力,构建基于偏置权重的 Adaboost。

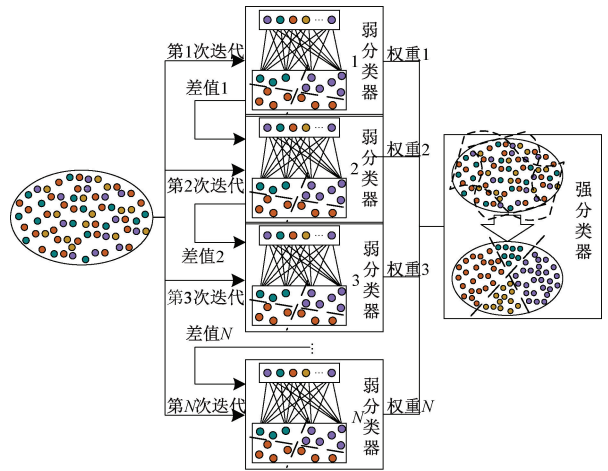


图6 Boosting 算法示意图

Fig. 6 The illustration of Boosting algorithm

2) 迁移学习和联邦学习等算法

迁移学习是利用已有的知识来解决目标领域中知识匮乏的问题,使用源域数据训练模型,再将模型迁移到目标域中进行测试。针对不平衡数据集,迁移学习通常是对有标签的数据集进行训练,再将模型参数迁移到无标签的数据集中进行测试,如图 7 所示。

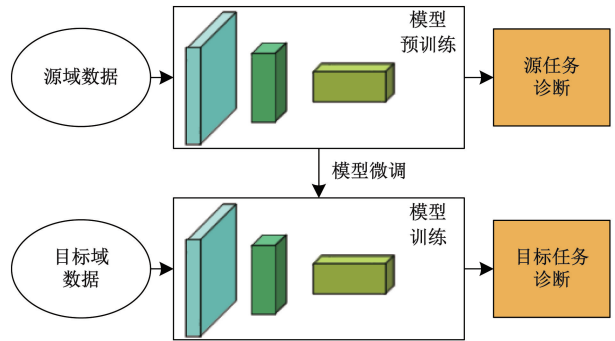


图7 迁移学习示意图

Fig. 7 The structure of Transfer Learning

迁移学习主要分为基于特征迁移、基于模型迁移和基于样例迁移这 3 个基本类别^[89],其中基于特征迁移和基于模型迁移在不平衡数据分布下故障诊断研究中应用较多。Zheng 等^[54]在数据增强阶段,采用多尺度递进生成对抗网络,通过迁移学习学习正常样本到故障样本的分布映射关系,通过递进对抗训练稳定生成不同尺度的故障样本,用于数据集增强。Liu 等^[27]为了减少跨域诊断困难,挖掘样本特征分布,构造了一种基于元数据的显式权重自分配策略,用于样本权重的再平衡,以防止大类别占主导地位和小类别过拟合。这两个应用都属于基于特征的迁移学习。Xu 等^[34]将特征注意力增强机制应用到迁移学习中,仅使用一种工况下不平衡样本进行训练,

增强不平衡样本中域不变特征,将训练好的模型直接用于其它工况的不平衡样本诊断,这是一种基于模型的迁移学习。

针对不平衡数据,联邦学习作为一种多模型学习将不同的本地数据集放入相同的模型中进行训练,再将训练好的模型参数整合成中央模型,更新后的模型参数再下发到本地进行训练。通过迭代,每个本地数据集都能获得较好的诊断结果,从而实现不平衡数据的故障诊断,其原理如图 8 所示。因为不涉及数据传输,联邦学习是解决数据隐私性的一种途径。

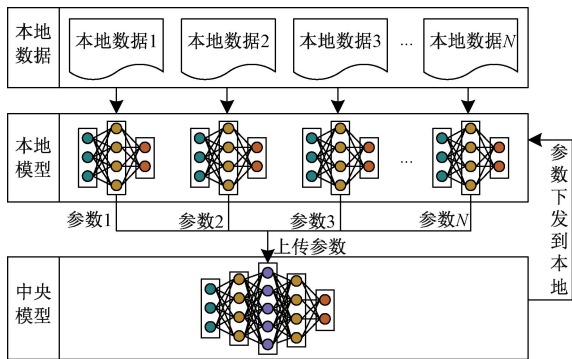


图 8 联邦学习算法示意图

Fig. 8 The schematic of federated learning

2022 年,这种多模型学习法被提出用于解决不平衡数据分布的故障诊断。郭方洪等^[12]将电力变压器视为不同本地数据集,使用联邦学习诊断故障,根据不同变压器数据差异设置比例以减少非独立同分布数据集间的不平衡性。Sima 等^[90]提出一种基于迁移和联邦学习的边缘云协作检测方法,该方法引入有标记数据集进行预训练,将预训练的模型部署到每个变电站,并以联邦学习方式进行微调,在云平台上进行全局聚合。

3.6 组合算法

除了上述方法以外,还有学者将上述两种或多种方法组合使用以解决不平衡数据的故障诊断。例如,将采样和集成学习相结合,先通过采样平衡数据集,再应用集成学习方法提高分类精度。潘进等^[91]考虑到不同类别的样本的分布密度不均,对合成少数类过采样技术(synthetic minority oversampling technique, SMOTE)进行了优化,使新合成的样本在密度集中的区域,避免合成噪音数据,再将极端梯度提升(eXtreme gradient boosting, XGBOOST)学习用于分类。还有将特征提取和集成学习相结合,黄子恒等^[92]将数据集小波变换重构去噪后,使用经验模态分解将非平稳信号分解为若干个平稳信号,提取时频特征后放入由 Focal 损失函数应用的 XGBOOST 算法和逻辑回归模型融合的极端梯度提升逻辑回归

(extreme gradient boosting-logistic regression, Xgblr) 模型中进行分类。

3.7 诊断算法的场景适应性发展脉络

将不平衡分布的数据驱动故障诊断算法与不平衡分布数据场景相对应,如表 3 所示。

表 3 不平衡数据驱动算法的场景适应性典型文献

Table 3 Literature summary of data-driven diagnosis algorithms and applications to different scenarios

算法模型	诊断场景	典型文献
多模型学习	部分标签	文献[22,71-72]
	新类识别	文献[74]
	样本不均	文献[7,17,27,52-54]
	样本稀少	文献[69]
分类器优化	样本不均	文献[37]
	样本稀少	文献[16,66-67]
生成模型	新类识别	文献[30-31,78]
	样本不均	文献[6,11,25,32-33,40-45]
	样本稀少	文献[29,55-56,58-63]
	新类识别	文献[75-77]
特征学习	样本不均	文献[14-15,46-51]
	样本稀少	文献[64-65]
	样本不均	文献[8-10,19-20,23-24,37-39]
重采样	样本稀少	文献[21,26,57,70]
	样本不均	文献[12,34]

由表 3 可以看出,基于重采样的故障诊断方法主要用于平衡样本不均的问题;而基于生成模型的故障诊断方法、基于特征学习的故障诊断方法和基于多模型学习的故障诊断方法用于解决标签的挑战。除此之外,基于多模型学习的故障诊断方法,特别是迁移学习,主要解决仅有部分标签约束的故障诊断问题。另外,基于重采样的故障诊断方法、基于特征学习的故障诊断方法、基于特征学习的故障诊断方法、基于生成模型的故障诊断方法和基于多模型学习的故障诊断方法都可用于解决样本不均的问题,而基于分类器优化的故障诊断方法主要解决样本稀少故障诊断问题。

回顾 2012—2022 年间不平衡分布的数据驱动故障诊断研究,总体上可以分为 3 个发展阶段:第 1 个阶段是以支持向量机、重采样等为代表的浅层学习驱动的不平衡诊断阶段;第 2 阶段是以自编码器、生成对抗网络等为代表的全监督深度学习驱动的不平衡诊断阶段;第 3 阶段是以新类识别为代表的研究标签缺失的半监督深度学习的不平衡诊断阶段。图 9 展示了不平衡分布的数据驱

动故障诊断研究发展脉络。在具体的不同发展阶段,Zhang 等^[47]于 2018 年在旋转机械不平衡数据故障诊断中,使用了自编码器自动学习特征。Cabrera 等^[59]和 Li 等^[93]于 2019 年将生成对抗网络应用到不平衡数据故障诊断研究中。同年,Xu 等^[28]使用迁移学习解决不平衡分布的故障诊断问题。2021 年,Li 等^[94]提出了故障诊断中新类识别的问题。

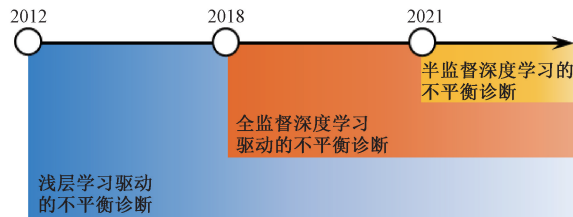


图 9 不平衡分布的数据驱动故障诊断研究发展脉络

Fig. 9 A timeline of fault diagnostics driven by imbalanced data distribution

4 主要的挑战

通过对不平衡数据驱动诊断场景及其算法的研究进展进行分析,在不平衡数据驱动故障诊断领域,目前还存在如下 5 个主要挑战。

1) 有效的诊断数据采集是不平衡数据驱动故障诊断研究的基础性挑战。在传统故障诊断数据采集的传感器、数据传输网络等基础设施挑战的基础上,不平衡故障诊断更多受到样本缺失或标签缺失的影响。目前大部分的研究中,虽然样本数量具有不平衡性,但样本标签基本上是已知的。然而在现实的故障诊断中,标签样本稀少且难以获得。不平衡数据往往带着样本标签的缺失,而样本标签缺失导致模型识别更困难,诊断精度也更低。

2) 诊断精度和效果仍是不平衡数据驱动故障诊断研究的主要挑战。当前大部分不平衡样本约束的故障诊断研究集中在样本不均上,甚至部分研究还停留在简单的二分类阶段。针对样本强不平衡甚至个别极端稀少的样本,例如只有一个样本的情况,当前的研究仍然较少。虽然这个挑战对其他故障诊断方法也很重要,但在数据驱动故障诊断领域尤为突出。

3) 诊断模型可靠性的评价指标是不平衡数据驱动故障诊断研究亟需解决的重点。在大多数研究中,虽然得出的实验结果精度较高,但是将数据集作为一个整体统一诊断,未对不平衡数据中各类别的诊断精度分别进行描述。在实际生产中,不平衡数据集中一些类别的诊断精度仍不高,这对平衡分布处理后的样本数据和样本标签的可靠性提出了严峻的挑战。

4) 不平衡分布的数据驱动故障诊断模型的泛化能

力仍然存在挑战。现有解决不平衡分布的诊断模型,主要在测试集上进行测试,而测试集的选择本身具有一定的局限性。对于所建立的数据驱动的不平衡故障诊断模型,由于缺乏与故障机理的有机联系,因此模型的可解释性不足,导致诊断模型的泛化能力不确定或不明确。

5) 通用化的不平衡数据故障诊断模型仍具有挑战性。随着工业应用中各种不平衡诊断场景的涌现,需要有效应对多模态、分布式、跨模型等需求。未来的研究可以更多地考虑通用化可裁剪的智能大模型,重点研究不平衡故障诊断大模型的建模、训练、优化、裁剪、部署等关键技术,进一步提升诊断通用性。

5 展 望

针对前述不平衡数据驱动故障诊断领域的主要挑战,下一步的研究展望归纳如图 10 所示,具体描述如下。

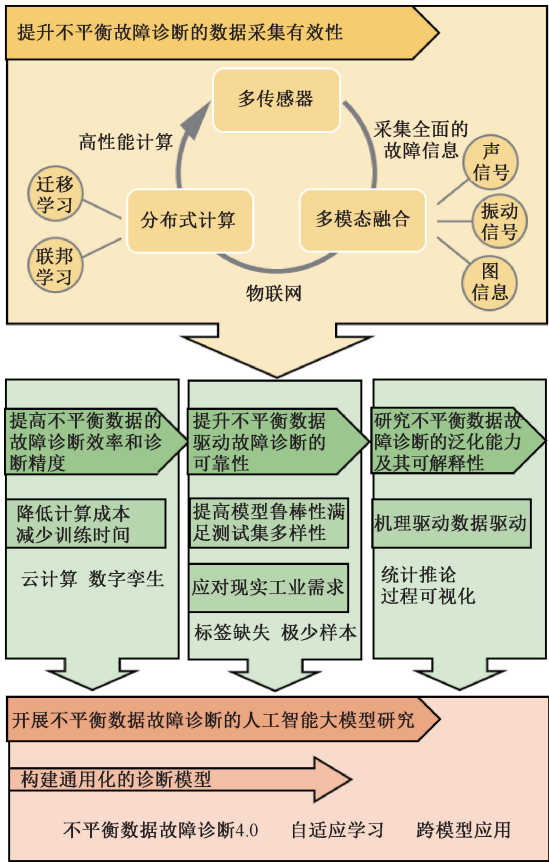


图 10 不平衡分布的数据驱动故障诊断的研究展望

Fig. 10 Perspectives of fault diagnostics driven by imbalanced distribution

1) 提升不平衡故障诊断的数据采集有效性。随着传感器技术的不断进步,可以利用多传感器数据采集更全面的故障信息。同时,为了解决不平衡数据带来的样

本不平衡和标签不平衡问题,通过多模态融合技术将不同类型的数据整合起来,以提供更充分的故障诊断信息。此外,采用分布式计算技术可以实现跨多个设备和系统的数据融合与处理,提高数据采集的效率和准确性。

2) 提高不平衡数据的故障诊断效率和诊断精度。需要关注不平衡分布下故障诊断模型预处理复杂、训练时间长、模型构建困难等问题,进一步研究,设计高效的算法和模型结构,以降低计算成本和训练时间,并提高实际应用中的效果。

3) 提升不平衡数据驱动故障诊断的可靠性。首先,模型要满足测试集的多样性,以覆盖各种真实工业场景,提高泛化能力。同时,要满足现实工业需求,将研究成果有效地应用到实际故障诊断系统中。此外,还需要应对标签缺失和极少样本的挑战,研究如何在这些情况下构建鲁棒性强的模型,以提高故障诊断的可靠性和鲁棒性。

4) 深入研究不平衡数据故障诊断的泛化能力及其可解释性。为了使模型的结果更容易理解和接受,增加人们对模型的信任,需要研究模型的可解释性。从此,人工智能进入3.0的时代。同样,需要深入研究不平衡数据故障诊断模型的可解释性。为了提高不平衡数据故障诊断的可解释性,需将数据驱动和机理驱动相结合,从而更好地理解 and 解释故障诊断的结果。同时,还应拓展泛化能力的研究,以应对新的故障类型和未知情况下的准确诊断。

5) 开展不平衡数据故障诊断的人工智能大模型研究,进入不平衡数据故障诊断4.0。为了进一步推动不平衡数据下故障诊断的研究,可以开展人工智能大模型的研究,将其应用于不平衡数据故障诊断领域。这需要构建通用化的诊断模型,解决内容生成模型在该领域的适用性问题。同时,还需要解决人工智能大模型的建模、训练以及多模态、分布式、跨模型应用等技术问题,以推动不平衡分布下数据驱动的故障诊断的进一步发展。

6 结 论

故障诊断是工业领域重要的研究内容,数据驱动的故障诊断模型是工业系统可靠性和健康管理的有效手段,随着科学技术的快速发展,不平衡数据驱动智能故障诊断具有更高的研究价值。通过对诊断算法和应用场景两个方面对不平衡数据驱动故障诊断研究进展进行系统的整理和总结,提出了目前研究面临的主要挑战,并对未来的研究方向进行了展望,对数据驱动故障诊断的研究与应用具有重要的参考价值。

参考文献

[1] 陈志强,陈旭东,OLIVIRA J V DE,等. 深度学习在设备故障预测与健康管理中的应用[J]. 仪器仪表学

报,2019,40(9):206-226.

CHEN ZH Q, CHEN X D, OLIVIRA J V DE, et al. Application of deep learning in equipment prognostics and health management [J]. Chinese Journal of Scientific Instrument, 2019,40(9):206-226.

[2] 伍济钢,文港,杨康. 改进一维卷积神经网络的航空发动机故障诊断方法[J]. 电子测量与仪器学报,2023,37(3):179-186.

WU J G, WEN G, YANG K. Improved one-dimensional convolutional neural network for aero-engine fault diagnosis [J]. Journal of Electronic Measurement and Instrumentation, 2023, 37(3): 179-186.

[3] 王鸽,吴国新,刘秀丽. 基于MADCNN的故障诊断方法研究[J]. 电子测量与仪器学报,2023,37(3):187-193.

WANG G, WU G X, LIU X L. Research on fault diagnosis method based on madcnn [J]. Journal of Electronic Measurement and Instrumentation, 2023, 37(3):187-193.

[4] 黄星华,吴天舒,杨龙玉,等. 一种面向旋转机械的基于Transformer特征提取的域自适应故障诊断[J]. 仪器仪表学报,2022,43(11):210-218.

HUANG X H, WU T SH, YANG L Y, et al. Domain adaptive fault diagnosis based on transformer feature extraction for rotating machinery[J]. Chinese Journal of Scientific Instrument, 2022,43(11):210-218.

[5] 周玉,孙红玉,房倩,等. 不平衡数据集分类方法研究综述[J]. 计算机应用研究,2022,39(6):1615-1621.

ZHOU Y, SUN H Y, FANG Q, et al. Review of imbalanced data classification methods [J]. Application Research of Computers(Natural Science Edition), 2022, 39(6):1615-1621.

[6] 李昂,韩萌,穆栋梁,等. 多类不平衡数据分类方法综述[J]. 计算机应用研究,2022,39(12):3534-3545.

LI ANG, HAN M, MU D L, et al. A survey of multi-class imbalanced data classification methods [J]. Journal of Intelligent & Fuzzy Systems, 2022, 39(12): 3534-3545.

[7] 李蒙蒙,刘艺,李庚松,等. 不平衡多分类算法综述[J]. 计算机应用,2022,42(11):3307-3321.

LI M M, LIU Y, LI G S, et al. Survey on imbalanced multi-class classification algorithms [J]. Journal of Computer Applications, 2022,42(11):3307-3321.

[8] ZHANG T, CHEN J, LI F, et al. Intelligent fault diagnosis of machines with small & imbalanced data: A state-of-the-art review and possible extensions [J]. Isa

- Transactions, 2022, 119:152-171.
- [9] QIAN M, LI Y F. A novel adaptive undersampling framework for class-imbalance fault detection[J]. IEEE Transactions on Reliability, 2022;1-15.
- [10] JANG J, KIM C O. Unstructured borderline self-organizing map: Learning highly imbalanced, high-dimensional datasets for fault detection [J]. Expert Systems with Applications, 2022, 188:116028.
- [11] BUSTILLO A, RODRIGUEZ J J. Online breakage detection of multitooth tools using classifier ensembles for imbalanced data [J]. International Journal of Systems Science, 2014, 45(12):2590-2602.
- [12] 郭方洪,刘师硕,吴祥,等. 基于联邦学习的含不平衡样本数据电力变压器故障诊断[J]. 电力系统自动化,2023,47(10):145-152.
- GUO F H, LIU SH SH, WU X, et al. Federated learning based fault diagnosis of power transformer with unbalanced sample data [J]. Automation of Electric Power Systems (Natural Science Edition), 2023, 47(10):145-152.
- [13] LIU J, LI Y, ZIO E. A SVM framework for fault detection of the braking system in a high speed train[J]. Mechanical Systems and Signal Processing, 2017, 87:401-409.
- [14] 韩淞宇,邵海东,姜洪开,等. 基于提升卷积神经网络的航空发动机高速轴承智能故障诊断[J]. 航空学报,2022,43(9):158-171.
- HAN S Y, SHAO H D, JIANG H K, et al. Intelligent fault diagnosis of aero-engine high-speed bearings using enhanced CNN [J]. Acta Aeronautica et Astronautica Sinica, 2022,43(9):158-171.
- [15] 康守强,刘哲,王玉静,等. 基于改进 DQN 网络的滚动轴承故障诊断方法[J]. 仪器仪表学报,2021,42(3):201-212.
- KANG SH Q, LIU ZH, WANG Y J, et al. A fault diagnosis method of rolling bearing based on the improved dqn network[J]. Chinese Journal of Scientific Instrument (Natural Science Edition), 2021,42(3):201-212.
- [16] 郝腾飞,陈果. 基于小球大间隔方法的机械故障检测[J]. 中国机械工程,2012,23(15):1765-1770.
- HAO T F, CHEN G. Machinery fault detection based on a small sphere and large margin approach [J]. China Mechanical Engineering, 2012,23(15):1765-1770.
- [17] 陶新民,李晨曦,李青,等. 不均衡最大软间隔 SVDD 轴承故障检测模型[J]. 振动工程学报,2019,32(4):718-729.
- TAO X M, LI CH X, LI Q, et al. Rolling bearings fault detection model using imbalanced maximum soft margin support vector domain description [J]. Journal of Vibration Engineering (Natural Science Edition), 2019, 32(4):718-729.
- [18] XU Q, LU SH, JIA W, et al. Imbalanced fault diagnosis of rotating machinery via multi-domain feature extraction and cost-sensitive learning [J]. Journal of Intelligent Manufacturing, 2020, 31(6):1467-1481.
- [19] 王艳,李伟,赵洪山,等. 基于融合少数类过采样均衡多分类数据的改进极限学习机的变压器故障诊断方法[J]. 电网技术,2023,47(9):3799-3807.
- WANG Y, LI W, ZHAO H SH, et al. Fusion synthetic minority over-sampling technique balanced multi-classification data for transformer fault diagnosis method based on improved extreme learning machine [J]. Power System Technology (Natural Science Edition), 2023, 47(9):3799-3807.
- [20] 张笑璐,邹益胜,张波,等. 基于 Bagging-MCNN 模型的不均衡样本轴承故障诊断方法[J]. 现代制造工程, 2022(1):104-112.
- ZHANG X L, ZOU Y SH, ZHANG B, et al. Unbalanced sample bearing fault diagnosis method based on Bagging-MCNN model [J]. Modern Manufacturing Engineering (Natural Science Edition), 2022(1):104-112.
- [21] 刘云鹏,和家慧,许自强,等. 基于 SVM SMOTE 的电力变压器故障样本均衡化方法[J]. 高电压技术, 2020,46(7):2522-2529.
- LIU Y P, HE J H, XU Z Q, et al. Equalization method of power transformer fault sample based on svm smote [J]. High Voltage Engineering (Natural Science Edition), 2020,46(7):2522-2529.
- [22] XIA M, SHAO H D, WILLIAMS D, et al. Intelligent fault diagnosis of machinery using digital twin-assisted deep transfer learning [J]. Reliability Engineering & System Safety, 2021, 215:107938.
- [23] 陶新民,张冬雪,郝思媛,等. 基于谱聚类下采样失衡数据下 SVM 故障检测 [J]. 振动与冲击, 2013, 32(16):30-36.
- TAO X M, ZHANG D X, HAO S Y, et al. Fault detection based on spectral clustering combined with under-sampling SVM under unbalanced datasets [J]. Journal of Vibration and Shock (Natural Science Edition), 2013,32(16):30-36.
- [24] SHI Q, ZHANG H. Fault diagnosis of an autonomous vehicle with an improved svm algorithm subject to

- unbalanced datasets[J]. IEEE Transactions on Industrial Electronics, 2021, 68(7):6248-6256.
- [25] 郭俊锋,王森生,孙磊,等. 基于生成对抗网络的滚动轴承不平衡数据集故障诊断新方法[J]. 计算机集成制造系统,2022,28(9):2825-2835.
- GUO J F, WANG M SH, SUN L, et al. New method of fault diagnosis for rolling bearing imbalance data set based on generative adversarial network[J]. Computer Integrated Manufacturing Systems (Natural Science Edition), 2022,28(9):2825-2835.
- [26] 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.
- [27] LIU X, CHEN J L, ZHANG K Y, et al. Cross-domain intelligent bearing fault diagnosis under class imbalanced samples via transfer residual network augmented with explicit weight self-assignment strategy based on meta data [J]. Knowledge-Based Systems, 2022, 251:109272.
- [28] XU K, LI SH M, WANG J R, et al. A novel convolutional transfer feature discrimination network for unbalanced fault diagnosis under variable rotational speeds [J]. Measurement Science and Technology, 2019, 30(10):105107.
- [29] 张笑璐,邹益胜,曾大懿,等. 样本不均衡下的 DCGAN 轴承故障诊断方法[J]. 机械科学与技术,2022, 41(1):9-15.
- ZHANG X L, ZOU Y SH, ZENG D Y, et al. DCGAN bearing fault diagnosis method under unbalanced samples[J]. Mechanical Science and Technology for Aerospace Engineering(Natural Science Edition), 2022, 41(1):9-15.
- [30] PU Z Q, CABRERA D, BAI Y, et al. A one-class generative adversarial detection framework for multifunctional fault diagnoses[J]. IEEE Transactions on Industrial Electronics, 2022, 69(8):8411-8419.
- [31] LONG J Y, ZHANG R X, YANG ZH, et al. Self-adaptation graph attention network via meta-learning for machinery fault diagnosis with few labeled data [J]. IEEE Transactions on Instrumentation and Measurement, 2022, 71:1-11.
- [32] 薛振泽,满君丰,彭成,等. 数据失衡下基于 WGAN 和 GAPCNN 的轴承故障诊断研究[J]. 计算机应用研究,2020,37(12):3681-3685.
- XUE ZH Z, MAN J F, PENG CH, et al. Research on bearing fault diagnosis based on wgan and gapcnn under imbalance of data [J]. Application Research of Computers (Natural Science Edition), 2020, 37(12):3681-3685.
- [33] 李慧芳,徐光浩,黄双喜. 基于主动生成式过采样和 DSN 的轴承故障诊断[J]. 计算机集成制造系统, 2023,29(1):146-159.
- LI H F, XU G H, HUANG SH X. Active generative oversampling and deep stacking network based bearing fault diagnosis approach [J]. Computer Integrated Manufacturing Systems(Natural Science Edition), 2023, 29(1):146-159.
- [34] XU K, LI SH M, JIANG X X, et al. A novel transfer diagnosis method under unbalanced sample based on discrete-peak joint attention enhancement mechanism[J]. Knowledge-Based Systems, 2021, 212:106645.
- [35] MATHEW J, PANG C K, LUO M, et al. Classification of imbalanced data by oversampling in kernel space of support vector machines [J]. IEEE Transactions on Neural Networks and Learning Systems, 2018, 29(9):4065-4076.
- [36] YAN K, HUANG J, SHEN W, et al. Unsupervised learning for fault detection and diagnosis of air handling units[J]. Energy and Buildings, 2020, 210:109689.
- [37] LI C, CABRERA D, SANCHO F, et al. Fusing convolutional generative adversarial encoders for 3D printer fault detection with only normal condition signals[J]. Mechanical Systems and Signal Processing, 2021, 147:107108.
- [38] 王德成,林辉. 一种 SVM 不平衡分类方法及在故障诊断的应用[J]. 电机与控制学报,2012,16(9):48-52.
- WANG D CH, LIN H. Imbalanced pattern classification method based on support vector machine and its application on fault diagnosis[J]. Electric Machines and Control(Natural Science Edition), 2012,16(9):48-52.
- [39] 余松,胡东,唐超,等. 基于 TLR-ADASYN 平衡化数据集的 MSSA-SVM 变压器故障诊断[J]. 高电压技术, 2021,47(11):3845-3853.
- YU S, HU D, TANG CH, et al. MSSA-SVM transformer fault diagnosis method based on tlr-adasyn balanced dataset[J]. High Voltage Engineering (Natural Science Edition), 2021,47(11):3845-3853.
- [40] 李东东,刘宇航,赵阳,等. 基于改进生成对抗网络的风机行星齿轮箱故障诊断方法[J]. 中国电机工程学报,2021,41(21):7496-7507.
- LI D D, LIU Y H, ZHAO Y, et al. Fault diagnosis method of wind turbine planetary gearbox based on

- improved generative adversarial network[J]. Proceedings of the CSEE(Natural Science Edition), 2021,41(21): 7496-7507.
- [41] 简献忠,张韬. 一种不平衡数据下的新型轻量级轴承故障诊断模型[J/OL]. 控制工程, 2023-08-12]. DOI:10.14107/j.cnki.kzgc.20211071.
JIAN X ZH, ZHANG T. A new lightweight bearing fault diagnosis model under unbalanced data[J/OL]. Control Engineering of China(Natural Science Edition), 2022: 1-9 [2023-08-12]. DOI: 10.14107/j.cnki.kzgc.20211071.
- [42] 李忠智,尹航,左剑凯,等. 不平衡训练数据下的基于生成对抗网络的轴承故障诊断[J]. 小型微型计算机系统, 2021,42(1):46-51.
LI ZH ZH, YIN H, ZUO J K, et al. Bearing fault diagnosis based on generative adversarial network on imbalanced data[J]. Journal of Chinese Computer Systems(Natural Science Edition), 2021,42(1):46-51.
- [43] YU Y X, GUO L, GAO H L, et al. PCWGAN-GP: A new method for imbalanced fault diagnosis of machines[J]. IEEE Transactions on Instrumentation and Measurement, 2022, 71:3515711.
- [44] 高学金,程琨,韩华云,等. 基于中心损失的条件生成式对抗网络的冷水机组故障诊断[J]. 化工学报, 2022,73(9):3950-3962.
GAO X J, CHENG K, HAN H Y, et al. Fault diagnosis of chillers using central loss conditional generative adversarial network[J]. CIESC Journal(Natural Science Edition), 2022,73(9):3950-3962.
- [45] 杨光友,刘浪,习晨博. 自适应辅助分类器生成式对抗网络样本生成模型及轴承故障诊断[J]. 中国机械工程, 2022,33(13):1613-1621.
YANG G Y, LIU L, XI CH B. Bearing fault diagnosis based on sa-acgan data generation model[J]. China Mechanical Engineering(Natural Science Edition), 2022,33(13):1613-1621.
- [46] LIU S W, JIANG H K, WU Z H, et al. Rolling bearing fault diagnosis using variational autoencoding generative adversarial networks with deep regret analysis[J]. Measurement, 2021, 168:108371.
- [47] ZHANG Y Y, LI X Y, GAO L, et al. Imbalanced data fault diagnosis of rotating machinery using synthetic oversampling and feature learning[J]. Journal of Manufacturing Systems, 2018, 48:34-50.
- [48] 董勋,郭亮,高宏力,等. 代价敏感卷积神经网络:一种机械故障数据不平衡分类方法[J]. 仪器仪表学报, 2019,40(12):205-213.
- DONG X, GUO L, GAO H L, et al. Cost sensitive convolutional neural network: A classification method for imbalanced data of mechanical fault[J]. Chinese Journal of Scientific Instrument, 2019,40(12):205-213.
- [49] 曹洁,何智栋,余萍,等. 数据不平衡分布下轴承故障诊断方法[J]. 吉林大学学报(工学版), 2022, 52(11):2523-2531.
CAO J, HE ZH D, YU P, et al. Bearing fault diagnosis method under unbalanced data distribution[J]. Journal of Jilin University(Engineering and Technology Edition), 2022,52(11):2523-2531.
- [50] 张洪,盛永健,黄子龙,等. 基于 W-DenseNet 的减压阀不平衡样本故障诊断模型[J]. 控制与决策, 2022, 37(6):1513-1520.
ZHANG H, SHENG Y J, HUANG Z L, et al. W-DenseNet-based fault diagnosis model of pressure-reducing valve with unbalanced samples[J]. Control and Decision, 2022,37(6):1513-1520.
- [51] 杜小磊,陈志刚,王衍学,等. 基于 IEWT 和 IFractalNet 的滚动轴承故障诊断[J]. 振动与冲击, 2020, 39(24):134-142.
DU X L, CHEN ZH G, WANG Y X, et al. Fault diagnosis of rolling bearings based on improved empirical wavelet transform and ifractalnet[J]. Journal of Vibration and Shock, 2020,39(24):134-142.
- [52] 马立玲,郭凯杰,王军政. 基于改进 SVM 的车辆传动系统故障诊断方法[J]. 北京理工大学学报, 2020, 40(8):856-860.
MA L L, GUO K J, WANG J ZH. A fault diagnosis method of vehicle transmission system based on improved SVM[J]. Transactions of Beijing Institute of Technology, 2020,40(8):856-860.
- [53] 贾凯,江明,袁啸林,等. 基于代价敏感型 LightGBM 的分子泵故障检测[J]. 电子测量与仪器学报, 2022, 36(10):55-64.
JIA K, JIANG M, YUAN X L, et al. Fault detection of molecular pump based on cost-sensitive lightgbm[J]. Journal of Electronic Measurement and Instrumentation(Natural Science Edition), 2022,36(10):55-64.
- [54] ZHENG M L, CHANG Q, MAN J F, et al. Two-stage multi-scale fault diagnosis method for rolling bearings with imbalanced data[J]. Machines, 2022, 10(5): 336.
- [55] 栗磊,王廷涛,赫嘉楠,等. 考虑过采样器与分类器参数优化的变压器故障诊断策略[J]. 电力自动化设

- 备,2023,43(1):209-217.
- LI L, WANG T T, HE J N, et al. Transformer fault diagnosis strategy considering parameter optimization of oversampler and classifier [J]. Electric Power Automation Equipment(Natural Science Edition), 2023, 43(1):209-217.
- [56] 李亮,范瑾,闫林,等. 基于混合采样和支持向量机的变压器故障诊断[J]. 中国电力, 2021, 54(12): 150-155.
- LI L, FAN J, YAN L, et al. Transformer fault diagnosis based on hybrid sampling and support vector machines[J]. Electric Power(Natural Science Edition), 2021, 54(12):150-155.
- [57] XING S B, LEI Y G, WANG SH 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.
- [58] 陈奎,韦晓广,陈景波,等. 基于样本数据处理和ADABOOST的小电流接地故障选线[J]. 中国电机工程学报,2014,34(34):6228-6237.
- CHEN K, WEI X G, CHEN J B, et al. Fault line detection using sampled data processing and ADABOOST for small current grounding system[J]. Proceedings of the CSEE, 2014, 34(34):6228-6237.
- [59] CABRERA D, SANCHO F, LONG J Y, et al. Generative adversarial networks selection approach for extremely imbalanced fault diagnosis of reciprocating machinery[J]. IEEE Access, 2019, 7:70643-70653.
- [60] 刘云鹏,许自强,和家慧,等. 基于条件式 Wasserstein 生成对抗网络的电力变压器故障样本增强技术[J]. 电网技术,2020,44(4):1505-1513.
- LIU Y P, XU Z Q, HE J H, et al. Data augmentation method for power transformer fault diagnosis based on conditional wasserstein generative adversarial network[J]. Power System Technology(Natural Science Edition), 2020, 44(4):1505-1513.
- [61] PU Z Q, CABRERA D, LI C, et al. VGAN: Generalizing mse gan and wgan-gp for robot fault diagnosis[J]. IEEE Intelligent Systems, 2022, 37(3): 65-75.
- [62] WANG Y R, SUN G D, JIN Q. Imbalanced sample fault diagnosis of rotating machinery using conditional variational auto-encoder generative adversarial network[J]. Applied Soft Computing, 2020, 92: 106333.
- [63] ZHAO X L, YAO J Y, DENG W X, et al. Normalized conditional variational auto-encoder with adaptive focal loss for imbalanced fault diagnosis of bearing-rotor system[J]. Mechanical Systems and Signal Processing, 2022, 170:108826.
- [64] 江丽,郭顺生. 基于无监督判别投影的滚动轴承故障诊断[J]. 中国机械工程,2016,27(16):2202-2206.
- JIANG L, GUO SH SH. Fault diagnosis of rolling bearings based on unsupervised discriminant projection[J]. China Mechanical Engineering(Natural Science Edition), 2016, 27(16):2202-2206.
- [65] 吴耀春,赵荣珍,靳伍银,等. 面向数据不平衡的卷积神经网络故障辨识方法[J]. 振动. 测试与诊断,2022, 42(2):299-307, 408.
- WU Y CH, ZHAO R ZH, JIN W Y, et al. Intelligent fault identification method based on convolutional neural network for imbalanced data[J]. Journal of Vibration, Measurement & Diagnosis, 2022, 42(2):299-307, 408.
- [66] 王瑞涵,陈辉,管聪,等. 基于图卷积网络的非均衡数据船舶柴油机故障诊断[J]. 中国舰船研究,2022, 17(5):289-300.
- WANG R H, CHEN H, GUAN C, et al. Fault diagnosis of marine diesel engines based on graph convolutional network under unbalanced datasets[J]. Chinese Journal of Ship Research, 2022, 17(5):289-300.
- [67] 宋志坤,徐立成,胡晓依,等. 基于改进型 shapelets 算法的动车组轴箱轴承故障诊断方法研究[J]. 仪器仪表学报,2021,42(2):66-74.
- SONG ZH K, XU L CH, HU X Y, et al. Research on fault diagnosis method of axle box bearing of emu based on improved shapelets algorithm[J]. Chinese Journal of Scientific Instrument, 2021, 42(2):66-74.
- [68] 位晶晶,刘勤明,叶春明,等. 基于 IPSO-BSVM 的小样本数据不均衡下的设备健康预测研究[J]. 计算机应用研究,2021,38(4):1119-1122,1127.
- WEI J J, LIU Q M, YE CH M, et al. Equipment health prognosis based on ipso-bsvm under small and imbalanced sample data[J]. Application Research of Computers(Natural Science Edition), 2021, 38(4): 1119-1122,1127.
- [69] 王劲锋,薛玉石,山春风. 模糊支持张量训练机及其在滚动轴承故障诊断中的应用[J]. 机电工程,2022, 39(10):1405-1411.
- WANG J F, XUE Y SH, SHAN CH F. Fuzzy support tensor train machine and its application in rolling bearing fault diagnosis[J]. Journal of Mechanical & Electrical

- Engineering (Natural Science Edition), 2022, 39(10): 1405-1411.
- [70] YAN K, JI Z W, LU H J, et al. Fast and accurate classification of time series data using extended elm: Application in fault diagnosis of air handling units[J]. IEEE Transactions on Systems Man Cybernetics-Systems, 2019, 49(7):1349-1356.
- [71] CHEN H, LI C S, YANG W X, et al. Deep balanced cascade forest: A novel fault diagnosis method for data imbalance[J]. ISA Transactions, 2022, 126:428-439.
- [72] LI C, CABRERA D, SANCHEZ F, et al. One-shot fault diagnosis of 3D printers through improved feature space learning [J]. IEEE Transactions on Industrial Electronics, 2021, 68(9):8768-8776.
- [73] 唐明珠,阳春华,桂卫华. 基于改进的 QBC 和 CS-SVM 的故障检测[J]. 控制与决策, 2012, 27(10): 1489-1493.
- TANG M ZH, YANG CH H, GUI W H. Fault detection based on modified QBC and CS-SVM[J]. Control and Decision, 2012, 27(10):1489-1493.
- [74] AN Z H, JIANG X X, CAO J, et al. Self-learning transferable neural network for intelligent fault diagnosis of rotating machinery with unlabeled and imbalanced data[J]. Knowledge-Based Systems, 2021, 230: 107374.
- [75] ZONG X, YANG R, WANG H S, et al. Semi-supervised transfer learning method for bearing fault diagnosis with imbalanced data[J]. Machines, 2022, 10(7):515.
- [76] YANG J L, WANG C D, WEI C A. A novel brownian correlation metric prototypical network for rotating machinery fault diagnosis with few and zero shot learners[J]. Advanced Engineering Informatics, 2022, 54:101815.
- [77] 赵晓平,吕凯扬,邵凡,等. 基于属性描述的零样本滚动轴承故障诊断[J]. 振动与冲击, 2022, 41(15): 105-115.
- ZHAO X P, LYU K Y, SHAO F, et al. Zero sample rolling bearing fault diagnosis based on attribute description[J]. Journal of Vibration and Shock (Natural Science Edition), 2022, 41(15):105-115.
- [78] 张永宏,邵凡,赵晓平,等. 基于多标签零样本学习的滚动轴承故障诊断[J]. 振动与冲击, 2022, 41(11): 55-64, 89.
- ZHANG Y H, SHAO F, ZHAO X P, et al. Rolling bearing fault diagnosis based on multi-label zero-shot learning[J]. Journal of Vibration and Shock (Natural Science Edition), 2022, 41(11):55-64, 89.
- [79] 张永宏,张中洋,赵晓平,等. 基于 VAE-GAN 和 FLCNN 的不均衡样本轴承故障诊断方法[J]. 振动与冲击, 2022, 41(9):199-209.
- ZHANG Y H, ZHANG ZH Y, ZHAO X P, et al. Bearing fault diagnosis method based on VAE-GAN and FLCNN unbalanced samples [J]. Journal of Vibration and Shock (Natural Science Edition), 2022, 41(9): 199-209.
- [80] 陶新民,李震,刘福荣,等. 基于精简集支持向量机的变压器故障检测方法[J]. 高电压技术, 2016, 42(10):3199-3206.
- TAO X M, LI ZH, LIU F R, et al. Fault detection method for power transformer based on svm using reduced vector set [J]. High Voltage Engineering (Natural Science Edition), 2016, 42(10):3199-3206.
- [81] JIANG N, LI N. A wind turbine frequent principal fault detection and localization approach with imbalanced data using an improved synthetic oversampling technique[J]. International Journal of Electrical Power & Energy Systems, 2021, 126:106595.
- [82] 孟宗,关阳,潘作舟,等. 基于二次数据增强和深度卷积的滚动轴承故障诊断研究[J]. 机械工程学报, 2021, 57(23):106-115.
- MENG Z, GUAN Y, PAN Z ZH, et al. Fault diagnosis of rolling bearing based on secondary data enhancement and deep convolutional network [J]. Journal of Mechanical Engineering (Natural Science Edition), 2021, 57(23):106-115.
- [83] ZHAO X L, JIA M P, LIN M Y. Deep laplacian auto-encoder and its application into imbalanced fault diagnosis of rotating machinery[J]. Measurement, 2020, 152:107320.
- [84] LI X Q, JIANG H K, LIU SH W, et al. A unified framework incorporating predictive generative denoising autoencoder and deep coral network for rolling bearing fault diagnosis with unbalanced data[J]. Measurement, 2021, 178:109345.
- [85] WANG J, ZHANG W Q, ZHOU J L. Fault detection with data imbalance conditions based on the improved bilayer convolutional neural network [J]. Industrial & Engineering Chemistry Research, 2020, 59(13): 5891-5904.
- [86] 姜万录,马敬宇,岳毅,等. 类间数据不平衡条件下基于平衡随机森林的轴向往柱塞泵故障诊断方法[J]. 液压与气动, 2022, 46(3):45-54.

- JIANG W L, MA X Y, YUE Y, et al. Fault diagnosis method of axial piston pump based on balanced random forest under imbalanced datasets[J]. Chinese Hydraulics & Pneumatics (Natural Science Edition), 2022, 46 (3): 45-54.
- [87] XU C, LI X L, WANG Z H, et al. Improved bls based transformer fault diagnosis considering imbalanced samples[J]. Energy Reports, 2022, 8:1446-1453.
- [88] JIANG X, XU Y, KE W, et al. An imbalanced multifault diagnosis method based on bias weights adaboost[J]. IEEE Transactions on Instrumentation and Measurement, 2022, 71:1-8.
- [89] 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 ZH, WANG J Y, ZHAO CH 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.
- [90] SIMA W, ZHANG H, YANG M. Edge-cloud collaboration detection approach for small-sample imbalanced faults in power lines[J]. IEEE Transactions on Instrumentation and Measurement, 2022, 71: 3513711.
- [91] 潘进,丁强,江爱朋,等. 基于 XGBoost 的冷水机组不平衡数据故障诊断[J]. 机械强度,2021,43(1): 27-33.
- PAN J, DING Q, JIANG AI P, et al. Fault diagnosis of unbalanced data of chillers based on xgboost[J]. Journal of Mechanical Strength(Natural Science Edition), 2021, 43(1): 27-33.
- [92] 黄子恒,许钊源,伍剑波,等. 基于优化模态分解和 Xgblr 的风机叶片故障诊断方法[J]. 机械设计,2022, 39(7):56-62.
- HUANG Z H, XU ZH Y, WU J B, et al. Fault diagnosis of wind turbine blade based on optimized mode decomposition and xgblr[J]. Journal of Machine Design, 2022,39(7):56-62.
- [93] LI Q, CHEN L, SHEN CH Q, et al. Enhanced generative adversarial networks for fault diagnosis of rotating machinery with imbalanced data [J]. Measurement Science and Technology, 2019, 30(11): 115005.
- [94] LI F D, CHEN J L, HE SH L, et al. Layer regeneration network with parameter transfer and knowledge distillation for intelligent fault diagnosis of bearing using class unbalanced sample [J]. IEEE Transactions on Instrumentation and Measurement, 2021, 70:1-10.

作者简介



李川 (通信作者), 2007 年获得重庆大学博士学位后,先后为加拿大渥太华大学博士后、韩国高丽大学研究教授、香港城市大学高级副研究者、厄瓜多尔慈幼理工大学普罗米修斯学者、重庆工商大学特聘教授,研究方向为设备智能运维。

E-mail: chuanli@21cn.com

Li Chuan (Corresponding author) received his Ph. D. degree from Chongqing University in 2007. He was a Postdoctoral Fellow with University of Ottawa, Canada, a research professor with Korea University, South Korea; a senior research associate with City University of Hong Kong, China; and a prometeo with Universidad Politecnica Salesiana, Ecuador. He is a distinguished professor with Chongqing Technology and Business University. His research interest is intelligent operation and maintenance of equipment.



伍依凡, 2017 年于吉林大学获得学士学位, 2023 年于重庆工商大学获得硕士学位, 现为韦仕敦大学博士生, 主要研究方向为数据分析, 故障诊断。

E-mail: wuyifanjl@163.com

Wu Yifan received her B.Sc. degree from Jilin University in 2017, and M.Sc. degree from Chongqing Technology and Business University in 2023. She is currently pursuing her Ph. D. degree at Western University. Her main research interests include data analysis and fault diagnostics.



杨帅, 2009 年于重庆大学获得学士学位, 2012 年于重庆大学获得硕士学位, 2017 年于渥太华大学获得博士学位, 现为重庆工商大学教授, 主要研究方向为振动控制, 故障诊断和制造系统健康管理。

E-mail: jerryang@ctbu.edu.cn

Yang Shuai received his B.Sc. degree and M.Sc. degree both from Chongqing University in 2009 and 2012, and Ph. D. degree from University of Ottawa in 2017. He is currently a professor at Chongqing Technology and Business University. His main research interests include vibration control, fault diagnosis and health management of manufacturing system.