



域泛化机械故障诊断研究进程与展望

邵海东*, 肖一鸣, 颜深, 王杰, 彭颖

湖南大学机械与运载工程学院, 长沙 410082

* E-mail: hdshao@hnu.edu.cn

收稿日期: 2024-06-14; 接受日期: 2024-09-10; 网络版发表日期: 2024-12-12

国家自然科学基金面上项目(批准号: 52275104)和湖南省科技创新计划(编号: 2023RC3097)资助

摘要 分布偏移阻碍了智能故障诊断模型在实际工程中的部署. 为缓解分布偏移, 基于迁移学习的智能故障诊断研究正如火如荼地开展中. 作为迁移学习的两种无监督变体, 域适应和域泛化尤其受到学者的关注, 其中域适应需利用目标域的无标签数据, 而域泛化解除了这一限制, 因此更具灵活性和泛用性. 自2020年域泛化首次被应用于故障诊断以来, 各类域泛化诊断技术层出不穷, 因此有必要对其进行全面调查. 本文首先按任务类型总体划分了现有域泛化故障诊断研究, 然后根据方法理论进一步细分了各任务下的研究, 总结了各方法理论的核心思想, 并回顾了基于各理论的故障诊断研究. 最后, 讨论了未来具有前景的研究方向.

关键词 分布偏移, 域适应, 域泛化, 机械故障诊断, 无监督迁移学习

1 引言

大数据时代下, 海量机械设备状态监测信号所蕴含的丰富信息为机械设备故障诊断带来了新机遇. 作为一种能自主挖掘监测信号中隐含的关键特征来建立原始数据与设备运行状态间映射关系的方法, 深度学习在迎接机遇的同时也带来了新的挑战^[1~3].

分布偏移, 即训练数据和测试数据服从相似但不同的分布, 是深度学习模型在实际工程中部署的一个重要挑战^[4,5]. 在机械故障诊断中, 由于训练数据和测试数据往往采集于机器的不同工况, 又或是源自相关但不同的机器, 分布偏移屡见不鲜^[6~8].

一种直观的解决方案是早期的监督迁移学习, 它使用少量的目标域有标签样本来微调在源域上训练的

模型^[9~13]. 然而, 标记样本往往费时费力, 因此催生了两种迁移学习的无监督变体, 域适应(Domain adaptation, DA)^[14~17]和域泛化(Domain generalization, DG)^[18~20]. 如图1所示, 域适应并不需要目标域的有标签样本, 但它仍然要求获取目标域的无标签样本用于训练. 考虑成本和安全等因素, 通常难以提前访问各类目标域数据, 因此域适应更适用于离线故障诊断. 相比之下, 域泛化解除了上述限制, 它仅使用源域数据训练模型, 而目标域数据仅出现在测试阶段. 域泛化这种“开箱即用”的属性使得其十分适用于机器的在线状态监测, 因此它已成为炙手可热的研究领域之一.

自2020年域泛化首次被应用于故障诊断以来, 学者们已相继开展了许多令人瞩目的域泛化故障诊断研究. 所涉及的任务类型已从最初的多源域泛化扩展至

引用格式: 邵海东, 肖一鸣, 颜深, 等. 域泛化机械故障诊断研究进程与展望. 中国科学: 技术科学, 2025, 55: 14–32

Shao H D, Xiao Y M, Yan S, et al. Progress and prospects of domain generalization mechanical fault diagnosis research (in Chinese). Sci Sin Tech, 2025, 55: 14–32, doi: [10.1360/SST-2024-0149](https://doi.org/10.1360/SST-2024-0149)

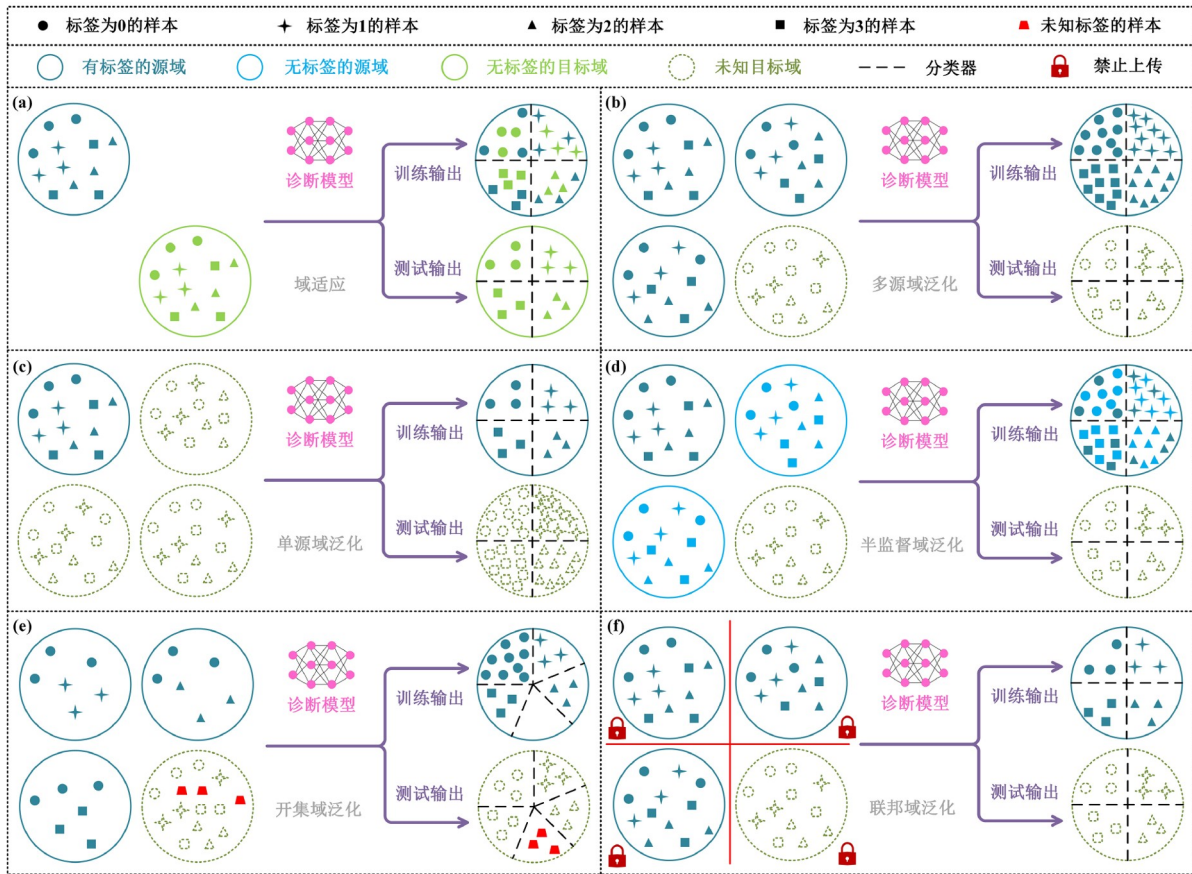


图 1 (网络版彩图)各任务的优化目标. (a) 域适应; (b) 多源域泛化; (c) 单源域泛化; (d) 半监督域泛化; (e) 开集域泛化; (f) 联邦域泛化

Figure 1 (Color online) Optimization objectives for each task. (a) Domain adaptation; (b) multi-source domain generalization; (c) single-source domain generalization; (d) semi-supervised domain generalization; (e) open domain generalization; (f) federated domain generalization.

单源域泛化, 以及半监督域泛化等. 同时, 所提出的方法设计策略也在不断发展, 从对齐源域以捕获域不变特征, 到利用元学习将模型置于模拟分布偏移, 再到通过样本生成增强源域等. 这些研究极大地丰富了域泛化故障诊断领域, 并促使本文对其进行全面调查.

自2019年以来, 学界已开展了一些与本文类似的调查研究^[21~28]. 然而, 现有研究^[21~25]通常更注重于对基于迁移学习的故障诊断领域做整体概述, 而未突出无监督迁移学习这一分支. 此外, 即使讨论了无监督迁移学习故障诊断, 其仍聚焦于域适应. 对于域泛化, 现有研究^[26,27]通常一笔带过甚至尚未提及. 2024年, Zhao等人^[28]首次以任务类型为导向回顾了现有域泛化故障诊断方法, 并在十个数据集上开展了对比实验, 作为后续研究的重要参考. 本文除类似地依据任务类

型对现有域泛化故障诊断方法进行总体划分外, 还额外基于方法的基本理论设计了分类法, 以进一步区分各任务类型下的研究. 此外, 对于每种域泛化任务, 本文并未局限于讨论现有方法, 而是提供了关于潜在解决方案的见解. 在计算机视觉领域所开展的域泛化调查研究中, 各类域泛化方法的基本理论已被初步介绍. 然而, 这些研究^[4,5]并未对不平衡域泛化、开集域泛化、联邦域泛化等新兴域泛化任务展开深入调研.

为填补上述研究的缺漏, 本文首先按任务类型总体划分了现有域泛化故障诊断方法, 然后根据方法的基本理论对它们进行进一步细分. 总结了各基本理论的核心思想, 并全面回顾和评价了各基本理论下的故障诊断研究. 最后, 在此基础上梳理了现存的挑战, 并提出了关于未来研究的设想.

2 研究背景

本节首先介绍一些通用的概念. 令 $\mathcal{X}$ 为输入空间, $\mathcal{Y}$ 为标签空间, 域可被定义为 $\mathcal{D} = \{(x, y)\} \sim p(X, Y)$, 其中, x 是一个采样自 $\mathcal{X}$ 的随机样本, y 是采样自 $\mathcal{Y}$ 的 x 的标签, X 是所有样本的集合, Y 是所有标签的集合, $p(X, Y)$ 是 X 和 Y 的联合分布. 该分布可被分解为 $p(X, Y) = p(X)p(Y|X)$, 其中, $p(X)$ 是 X 的边缘分布, $p(Y|X)$ 是 Y 条件于 X 的条件分布. 域泛化旨在仅使用源域, 学习一个在未知目标域上预测误差尽可能小的分类模型 $F \circ G: \mathcal{X} \rightarrow \mathcal{Y}$, 其中, F 是一个特征提取器, G 是一个分类器.

基于任务类型, 本文将现有的域泛化故障诊断研究划分为五类: 多源域泛化; 单源域泛化; 半监督域泛化; 开集域泛化; 联邦域泛化. 图1阐释了它们各自的优化目标, 表1对它们进行了区分. 此外, 图2展示了2020年至2024年(截至7月前)各类域泛化故障诊断研究数量的增长趋势, 图3则显示了各类研究的数量占总研究数量的比例. 该统计数据来自文献数据库Web of Science. 通过以关键词“Fault diagnosis”和“Domain generalization”进行搜索, 并将文献类型限制为“Article”, 共查找到70篇以轴承、齿轮箱等旋转机械为研究对象的域泛化故障诊断文献, 作为本文的主要调研对象. 此外, 还有1篇针对管道, 1篇针对高速列车刹车片, 以及1篇针对冷凝机的域泛化故障诊断文献被纳入本文的参考范围. 对于研究对象为交流电机、化工过程、核电站等的文献, 则不予考虑. 在文献检索过程中, 由于关键词不完整或不准确, 可能会存在遗漏. 例

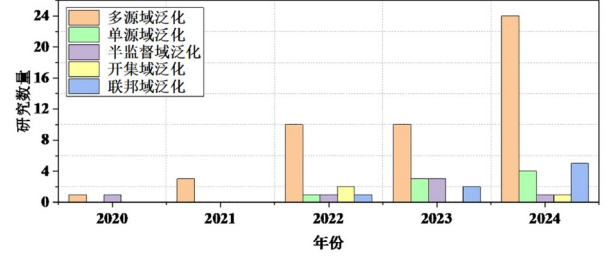


图2 (网络版彩图)近年域泛化故障诊断研究数量增长趋势
Figure 2 (Color online) Growth trend in the number of domain generalization fault diagnosis studies in recent years.

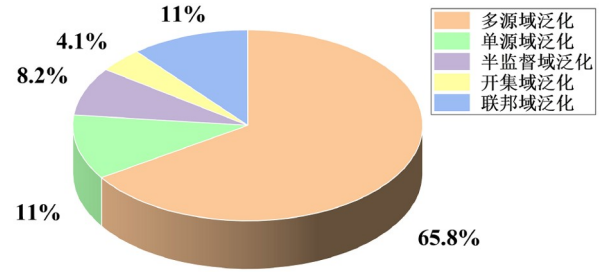


图3 (网络版彩图)各类研究的数量占总研究数量的比例
Figure 3 (Color online) Number of studies of each type as a percentage of total studies.

如, 一些学者会以“Cross domain”作为关键词来替代“Domain generalization”, 这是文献检索有效性的主要威胁. 从图2和3可以发现, 域泛化故障诊断研究, 尤其是针对新兴域泛化任务的研究, 其每年的研究数量正在逐年增长. 这表明域泛化故障诊断是一个重要的研究领域, 并将在未来几年持续受到关注.

(1) 多源域泛化. 如图1(b)所示, 多源域泛化是目前

表1 各域泛化任务的详细对比

Table 1 Detailed comparison of various domain generalization tasks

任务	源域	目标域	各域的标签集	数据存储
多源域泛化	$\mathcal{D}_S = \{\mathcal{D}_k = \{(x^{(k)}, y^{(k)})\}_{k=1}^K\}$	$\mathcal{D}_T = \{x^T\}$	$\mathcal{C}^{(1)} = \mathcal{C}^{(2)} = \dots = \mathcal{C}^{(K)} = \mathcal{C}^T$	中心存储
单源域泛化	$\mathcal{D}_S = \{\mathcal{D}_1 = \{(x^{(1)}, y^{(1)})\}\}$	$\mathcal{D}_T = \{x^T\}$	$\mathcal{C}^{(1)} = \mathcal{C}^T$	中心存储
半监督域泛化	$\mathcal{D}_S = \{\mathcal{D}_1 = \{(x^{(1)}, y^{(1)})\}, \mathcal{D}_2 = \{(x^{(2)}, y^{(2)})\}, \dots, \mathcal{D}_K = \{(x^{(K)}, y^{(K)})\}\}$	$\mathcal{D}_T = \{x^T\}$	$\mathcal{C}^{(1)} = \mathcal{C}^{(2)} = \dots = \mathcal{C}^{(K)} = \mathcal{C}^T$	中心存储
开集域泛化	$\mathcal{D}_S = \{\mathcal{D}_k = \{(x^{(k)}, y^{(k)})\}_{k=1}^K\}$	$\mathcal{D}_T = \{x^T\}$	$\mathcal{C}^{(1)} \neq \mathcal{C}^{(2)} \neq \dots \neq \mathcal{C}^{(K)} \subseteq \mathcal{C}^T$	中心存储
联邦域泛化	$\mathcal{D}_S = \{\mathcal{D}_k = \{(x^{(k)}, y^{(k)})\}_{k=1}^K\}$	$\mathcal{D}_T = \{x^T\}$	$\mathcal{C}^{(1)} = \mathcal{C}^{(2)} = \dots = \mathcal{C}^{(K)} = \mathcal{C}^T$	分布存储

域泛化故障诊断研究中最常用的任务设置. 在该任务中, 可获取到 K ($K > 1$) 个相关但不同的源域 $\mathcal{D}_S = \{\mathcal{D}_k = \{(x^{(k)}, y^{(k)})\}\}_{k=1}^K$, 每个源域均与一个联合分布 $p(X^{(k)}, Y^{(k)})$ 相关联, 其中, $(x^{(k)}, y^{(k)})$ 是一个采集自该分布的样本标签对. 对于任意的 $k \neq k'$ 且 $k, k' \in \{1, 2, \dots, K\}$, 均有 $p(X^{(k)}, Y^{(k)}) \neq p(X^{(k')}, Y^{(k')})$. 与目标域 $\mathcal{D}_T = \{x^T\}$ 相关联的联合分布标记为 $p(X^T, Y^T)$. 同样, 对于任意的 $k \in \{1, 2, \dots, K\}$, 有 $p(X^{(k)}, Y^{(k)}) \neq p(X^T, Y^T)$. 这种分布差异通常被归因于边缘分布 $p(X)$ 发生偏移, 而条件分布 $p(Y|X)$ 被认为是跨域稳定的. 在机械故障诊断的背景下, 造成 $p(X)$ 偏移的因素包括相同机器上转速和负载的变化, 以及不同机器的机械结构的差异等. 此外, 源域和目标域共享相同的标签集 $\mathcal{C} = \{0, 1, \dots, C-1\}$, 其中, C 是已知类别的总数, 标签 0 代表正常类, 而其他标签代表故障类. 这一任务设置较为直观, 因此受到了学者的青睐. 其他任务均可被视作此任务的拓展.

(2) 单源域泛化. 如图 1(c) 所示, 单源域泛化与多源域泛化的区别仅在于它假设只有一个源域可用, 即 $K=1$. 当有多个源域可用但域标签不可知时, 也可被视为是一种单源域泛化场景. 此外, 在机械故障诊断的背景下, 当用于训练的机器状态监测信号是在连续变化工况下采集的而难以被区分为多个域时, 也可被视作单源域泛化任务进行处理.

(3) 半监督域泛化. 半监督域泛化假设多个源域可获取, 但仅有一个源域的样本的标签可用, 即 $\mathcal{D}_S = \{\mathcal{D}_1 = \{(x^{(1)}, y^{(1)})\}, \mathcal{D}_2 = \{(x^{(2)})\}, \dots, \mathcal{D}_K = \{(x^{(K)})\}\}$. 需要指出的是, 半监督域泛化强调的是源域样本没有完全被标记, 这与无监督迁移学习所指代的目标域样本未被标记并不冲突. 如图 1(d) 所示, 半监督域泛化的目标是捕获标记源域和未标记源域之间的域不变特征.

(4) 开集域泛化. 开集域泛化是一项具有挑战性的任务, 它假设源域并不共享相同的标签集, 且目标域的标签集中包含所有源域中都不存在的新标签, 即 $\mathcal{C}^{(1)} \neq \mathcal{C}^{(2)} \neq \dots \neq \mathcal{C}^{(K)} \subseteq \mathcal{C}^T$. 这一设置由 Shu 等人^[19]首

次提出. 在机械故障诊断的背景下, 由于采集机器处于正常模式下的状态监测信号较为简便^[29,30], 因此每个源域应至少包含正常类的样本. 如图 1(e) 所示, 开集域泛化不仅要求模型在这些源域上学习可泛化表征来识别目标域中的已知类的样本, 还要求其检测并拒绝来自未知类的样本.

(5) 联邦域泛化. 如图 1(f) 所示, 联邦域泛化要求在联邦学习^[31~34]的框架中解决域泛化问题, 其中数据不再是中心存储而是分布在多个客户端中, 且每个客户端持有一个域. 源客户端(源域持有者)需要在本地数据不离开本地存储以保护数据隐私的前提下, 协作训练一个可供目标客户端(目标域持有者)使用的诊断模型. 由于传统的域泛化方法通常要求直接操纵数据或表征, 且分布偏移问题可能会造成在联邦学习中各客户端的局部优化目标远离全局优化目标^[35~38], 因此联邦域泛化对联邦学习研究和域泛化研究都提出了更高的挑战.

3 基于方法理论的分类法

本节根据方法的基本理论来进一步细分各任务下的研究, 同时总结各基本理论的核心思想并对各理论下的研究进行回顾及评价.

3.1 多源域泛化

2020 年, Zheng 等人^[39]利用局部 Fischer 判别分析, 将每个源域映射为格拉斯曼流形上的一个点, 构建了一种域泛化诊断方法. 这既是首个针对多源域泛化的故障诊断研究, 也拉开了域泛化在机械故障诊断中应用的序幕. 与其他任务相比, 多源域泛化故障诊断研究的数量较多, 所涉及的方法理论也涵盖了现有调查研究^[4,5]中所提到的大部分理论. 因此, 如图 4 所示, 本文参考现有调查研究^[4,5]来进一步划分多源域泛化故障诊断研究. 由于其他域泛化任务均可被视为多源域泛化的拓展, 因此针对它们的大多数研究都可被纳入图 4 中所示的类别. 现有域泛化故障诊断研究主要基于三种基本理论来构建其方法: (1) 数据增强; (2) 域对齐; (3) 特征解耦. 此外, 元学习和模型可解释性等新兴理论也逐渐受到了学者的重视. 需指出的是, 大多数方法都涉及上述多种理论. 本节将阐述每种理论的技术在域泛化任务中所能达成的目标, 以协助学者选择

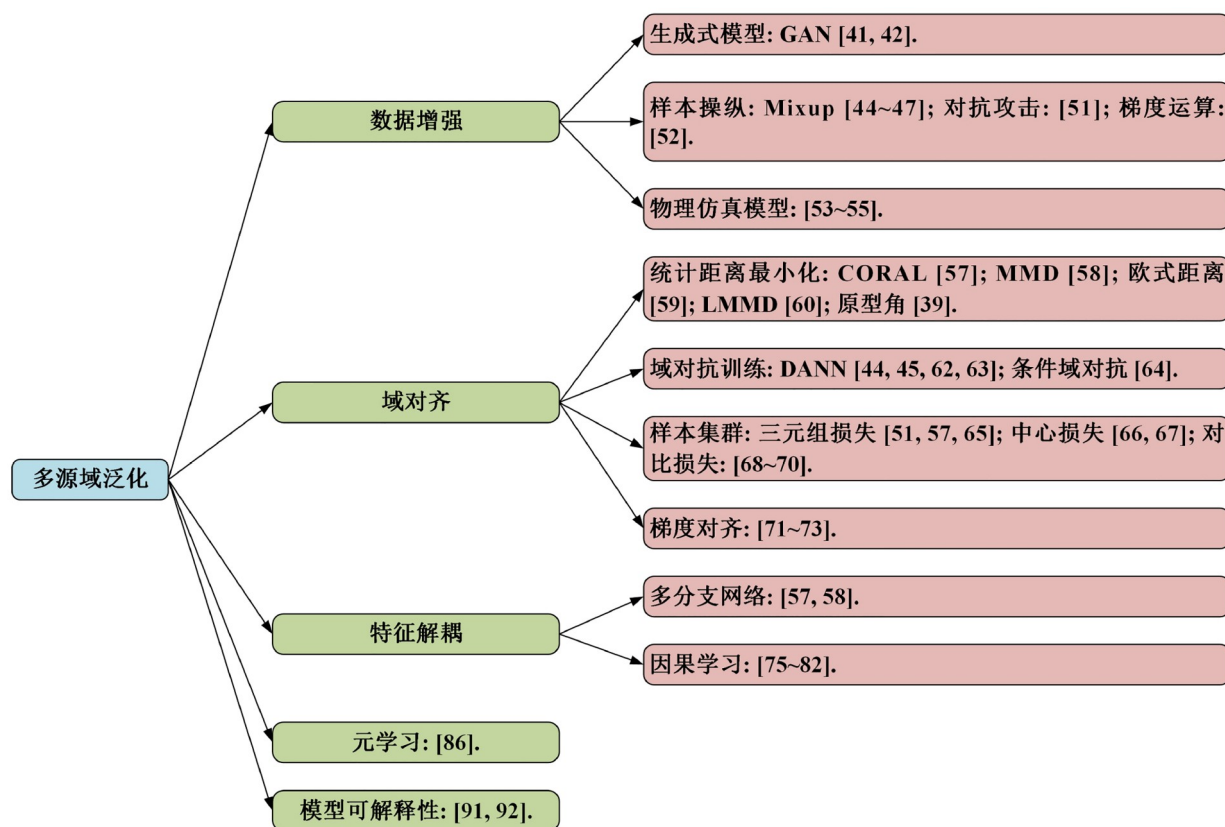


图4 (网络版彩图) 基于方法理论的多源域泛化故障诊断研究分类法

Figure 4 (Color online) Taxonomy of multi-source domain generalization fault diagnostic studies based on methodology.

和改进这些技术, 开发出适合其特定需求的域泛化故障诊断方法。

3.1.1 数据增强

该理论的技术通过处理原始样本来获得增强样本, 使得可用训练样本的数量和多样性都得到提升。这类技术可被归为三类, 分别是生成式模型, 样本操纵, 以及物理仿真模型。

(1) 生成式模型. 生成式模型可生成潜在域来扩充训练样本。在现有研究中, 最常用的生成式模型是生成对抗网络(Generative adversarial networks, GAN)及其变体^[40]。需注意的是, 在以往的GAN研究中, 一般要尽可能确保生成样本与原始样本的分布差异较小来补充原始样本。然而, 在如图5所示的基于GAN的域泛化故障诊断研究中, 其优化目标通常是最大化源域和生成的潜在域的分布差异, 以模拟真实的分布偏移并确保潜在域的多样性。2020年, Zhuang等人^[41]利用改进

GAN生成了多样化的潜在域, 并通过语义一致正则化缓解了源域和潜在域间的语义分歧, 实现了旋转机械的健康评估。2024年, Wang等人^[42]设计了一种攻防策略来增强管道故障诊断模型的泛化性, 模型需要区分生成的攻击样本和支持样本, 形成一个包含所有未知样本的新决策边界。更多基于生成式模型的域泛化故障诊断研究将在第3.2节单源域泛化中进行介绍。尽管构建生成式模型可获得多样化的训练样本, 但它们往往会引入对抗训练, 因此为模型优化带来了挑战。

(2) 样本操纵. 相较于生成式模型, 样本操纵是一类相对简便的数据增强方法, 其中最为流行的操纵方法是Mixup^[43]。无需复杂的模型构建和训练, 它直接将两个样本标签对以随机比例进行线性组合来生成增强的样本标签对。在多源域泛化故障诊断中, Mixup通常会被改进为: 随机组合来自不同源域的样本标签对来获取增强的样本标签对, 同时以相同比例组合它们的域标签来赋予增强样本域标签。具体定义如下:

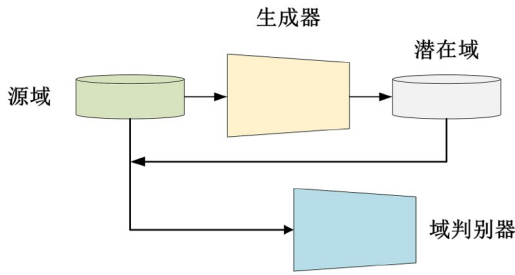


图 5 (网络版彩图)基于GAN的潜在域生成方法

Figure 5 (Color online) GAN-based methods for generating latent domains.

$$\begin{aligned} x' &= \sum_{k=1}^K \lambda^{(k)} x^{(k)}, \\ y' &= \sum_{k=1}^K \lambda^{(k)} y^{(k)}, \\ d' &= \sum_{k=1}^K \lambda^{(k)} d^{(k)}, \end{aligned} \quad (1)$$

式中, $d^{(k)}$ 为随机样本 $x^{(k)}$ 的域标签, x' 为生成的增强样本, y' 和 d' 分别为它的标签和域标签, $\lambda^{(k)}$ 是采样自预先定义的狄利克雷分布的随机变量. 这不仅补充了源域的数量和种类, 还平滑了域标签以更好地优化模型. 现有研究^[44~46]中所使用的样本操纵方法均遵循上述策略, 其中Li等人^[46]不仅将其应用于原始样本, 还使用该方法处理样本的表征. 2024年, Zhao和Shen^[47]提出了一种语义判别增强驱动的网络用于类别不平衡条件下的域泛化故障诊断, 其中具有语义正则化的Mixup被用于补偿少数类. 与前述基于Mixup的方法不同的是, 该方法仅线性混合相同类的两个样本来定向生成属于该类的增强样本.

此外, 有一种被称为对抗攻击的操纵方法, 它通过对源域样本施加微小扰动来生成攻击样本. 这些攻击样本会试图扰乱模型的训练方向, 而模型的训练目标即为抵御这些攻击, 从而提高其鲁棒性. 这类方法在计算机视觉领域备受关注, 并已开发出如FGSM^[48], DeepFool^[49]和JSMA^[50]等经典方法, 但在多源域泛化故障诊断中的应用较少. 2021年, Han等人^[51]通过对原始样本施加幅值偏移和高斯噪声来生成对抗样本, 并训练特征提取器将对抗样本与原始样本混淆, 以避免在处理噪声样本时出现过拟合.

除上述操纵方法外, 还有一种方法根据梯度来操纵样本的表征以生成增强表征. 2020年, Tang等人^[52]提出了一种表征梯度静音范式, 它通过丢弃表征中梯度较高的部分来迫使诊断模型关注剩余的且与标签相

关的部分, 从而提高其泛化性. 这种方法与dropout类似, 但不同之处在于它根据梯度来决定需要丢弃的部分, 而dropout则是随机丢弃.

(3) 物理仿真模型. 相较于前述两种方法, 物理仿真模型即使在源域样本不可获取的条件下也能生成大量样本, 且这些样本具备前者难以拥有的物理意义. 此外, 通过调节仿真模型的参数, 所生成的样本具有可控的工况条件及故障模式, 因此可灵活满足各类研究的需求, 并降低对模拟试验台的依赖. 2024年, Shi等人^[53]利用唯象模型生成了大量具有不同调制频率和信噪比的仿真数据, 为后续诊断模型训练提供支撑. 2024年, Gong等人^[54]构建了一种动力学仿真模型生成了充足的仿真数据, 在仅有少量真实数据的条件下捕获了域不变特征. 2024年, Pang等人^[55]设计了一种谱背景估计算法来估计真实系统中的潜在共振频率, 并用以指导物理仿真模型的参数设置, 从而缓解了仿真模型参数选择对经验的过度依赖. 尽管通过物理仿真模型生成数据相较于前述两种方法具备众多优势, 但需注意的是, 所构建的仿真模型必须能反映机械系统真实运行状况, 这就需要研究者掌握充足的先验物理知识. 此外, 仿真模型难以模拟设备真实运行状况下的背景噪声, 这可能会使生成的数据过于纯净, 无法为模型训练提供有力支持.

3.1.2 域对齐

该理论长期以来一直是解决分布偏移问题的重要依据, 并为域泛化的成功做出了显著贡献. 基于该理论的方法强调通过匹配所有可用域的表征分布来捕获可泛化至未知目标域的域不变表征. 根据域的匹配方式, 现有域对齐方法被划分为四类: 统计距离最小化; 域对抗训练; 样本集群; 梯度对齐.

(1) 统计距离最小化. 统计距离最小化会明确定义一个距离来衡量每两个域的表征分布间的差异, 同时在优化模型时不仅要使其在各个域上经验风险最小化, 还需使上述距离最小. 现有研究中使用最广泛的距离为最大均值差异(Maximum mean discrepancy, MMD), 其次是相关配准(Correlation alignment, CORAL). 此外, MMD的一种变体, 局部最大均值差异(local MMD, LMMD), 也受到了关注. 它们无需使用样本的标签信息, 因此起初被应用于域适应来测量源域和无标签目标域之间的差异. 需强调的是, MMD和

CORAL计算的是两个域之间的全局差异, 因此仅可用于在全局层面对齐域。相比之下, LMMD关注的是两个域的相关子域(具有相同标签的样本集)间的差异, 因此可用于在类层面对齐域。2022年, Zhao和Shen^[57]提出了一种兼顾域不变性和域特异性的域泛化故障诊断方法, 其中CORAL被用于全局域对齐。2023年, Wang等人^[58]构建了一个具有多个域特定分类器的域泛化诊断模型, 它使用MMD实现域对齐。尽管这两种方法在整体上属于特征解耦方法, 但它们在提取域不变特征时依然需要使用域对齐理论。2024年, Fan等人^[59]开发了一种用于故障诊断的深度混合域泛化网络, 该网络通过最小化不同域表征之间的欧氏距离来融合所有可用的域。2020年, Hu等人^[60]设计了一种用于高速列车刹车片健康监测的深度子域泛化网络, 它利用LMMD使具有相同标签的子域映射到合适的公共子空间。Zheng等人^[39]采用原型角计算两个子空间的距离, 并通过搜寻到所有域的距离之和最小的子空间, 得到它们的平均子空间。总体上, 统计距离最小化是一类相对易于实现的方法。

(2) 域对抗训练。域对抗训练并不预先定义一个距离来衡量两个域之间的差异, 而是使用域判别器来判断它们是否被混淆。特征提取器和域判别器会相互博弈, 前者尽可能使所有源域的特征与域无关, 而后者则试图确定这些特征来自哪个域。域判别器可类似如图6所示的经典域对抗神经网络(Domain adversarial neural networks, DANN)^[26]那样仅有一个分类模块, 它不利用类别信息因此在全局层面区分域, 其优化目标为

$$\begin{aligned} \min_{F,G,K} \frac{1}{K} \sum_{k=1}^K \mathbb{E}_{(x^{(k)}, y^{(k)}) \sim \mathcal{D}_k} \mathcal{L}(G(F(x^{(k)})), y^{(k)}) - \mathcal{L}(D(f(x^{(k)})), d^{(k)}), \\ \max_D \frac{1}{K} \sum_{k=1}^K \mathbb{E}_{(x^{(k)}, y^{(k)}) \sim \mathcal{D}_k} \mathcal{L}(G(F(x^{(k)})), y^{(k)}) - \mathcal{L}(D(f(x^{(k)})), d^{(k)}), \end{aligned} \quad (2)$$

式中, D 为判别器, $\mathcal{L}(\cdot, \cdot)$ 为交叉熵损失。此外, 也可如图7所示使用多个分类模块组成域判别器, 其中每个分类模块仅针对一个故障类别, 从而在类层面区分域, 这一方法也被称为条件域对抗^[61]。现有研究^[44,45,62,63]中所提的域对抗诊断模型多用于实现全局域对齐, 其中Shi等人^[44,63]还通过为源域自适应加权来降低负迁移的风险。相比之下, Zhang等人^[64]在2022年所提的条件对抗域泛化诊断模型则实现了类级域对齐。此外, 该方

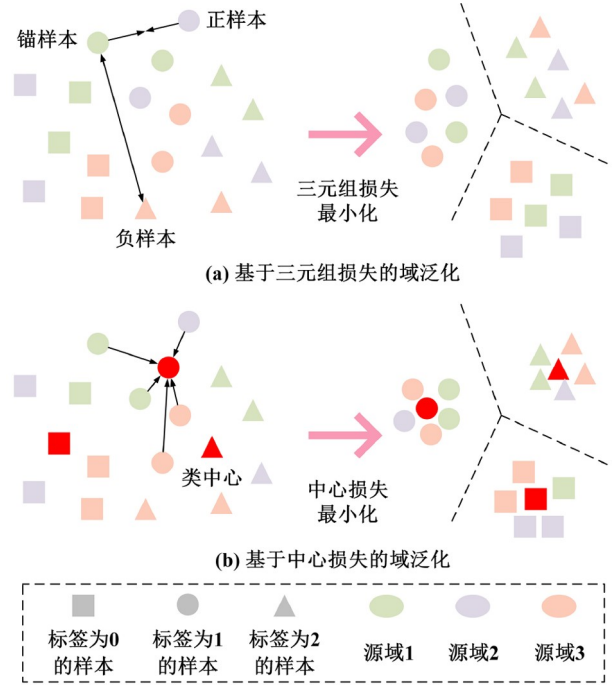


图6 (网络版彩图)经典DANN的模型架构

Figure 6 (Color online) Model architecture of classical DANN.

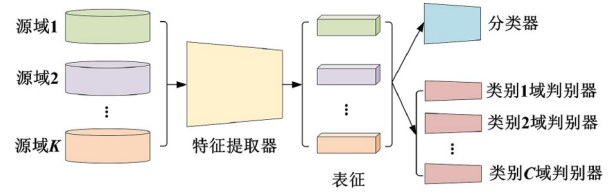


图7 (网络版彩图)条件域对抗网络的模型架构

Figure 7 (Color online) Model architecture of conditional domain adversarial network.

法还设计了一种新的对抗策略, 使得模型使用仅有一个分类模块的判别器即可达成条件域对抗, 从而显著降低了训练成本。与统计距离最小化相比, 域对抗训练在衡量分布差异方面更加灵活, 但也更难优化。

(3) 样本集群。样本集群并不像前两类方法那样关注两个域或两个子域的特征分布的差异, 而是明确要求聚集所有域中同一类别的特征, 同时不同类别的特征相互分离, 充分利用了多源域泛化中源域样本的标签可用这一条件。这使得不同域中的同类样本会被投射到一个共同子空间, 从而实现类级域对齐。这类方法通常涉及三元组损失和中心损失, 它们的工作原理如图8所示。2024年, Shen等人^[65]开发了一种新型的基

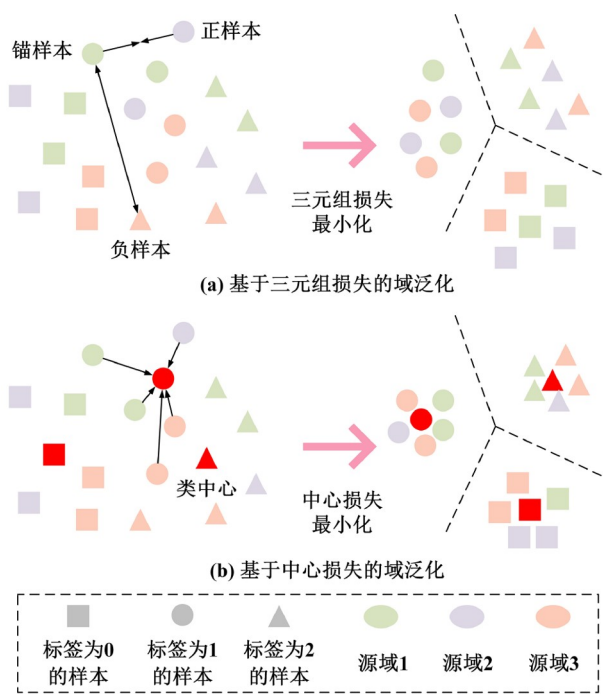


图 8 (网络版彩图)基于样本集群的域泛化方法的工作原理. (a) 三元组损失; (b) 中心损失

Figure 8 (Color online) Operating mechanism of the domain generalization methods based on sample cluster. (a) Triplet loss; (b) center loss.

于三元组损失的域泛化网络用于轴承故障诊断. Zhao和Shen^[57]使用CORAL和三元组损失来分别在全局层面和类层面对齐源域. Han等人^[51]利用三元组损失学习了源域的可泛化判别性结构. 2023年, Wang等人^[66]设计了一种用于轴承故障诊断的自适应类中心泛化网络, 其中心损失所使用的类中心可自适应更新. 2024年, Jia等人^[67]提出了一种动态平衡双原型域泛化网络用于跨机器故障诊断, 这一方法不仅会类似中心损失搜索类中心, 还会寻找域中心, 以弥补样本集群方法仅能学习域不变特征而难以抑制域特定特征的缺陷.

对比学习也被常用于开发样本集群方法. 2022年, Ragab等人^[68]提出了一种条件对比域泛化故障诊断方法, 它通过将所有域中相同类样本的互信息最大化, 而不同类样本的互信息最小化来实现类集群. 2023年, Shi等人^[69]开发了一种基于对比损失的跨机器故障诊断域泛化网络, 将来自不同域的两个同类样本视为一对正样本, 而不同类的样本为一对负样本来构建对比损失. 相比之下, 其他研究^[51,57,68]仅根据样本类别寻找

锚样本、正样本和负样本以构建三元组, 而不管它们源于哪个域. 这一差异使得Shi等人^[69]所提的方法需要获取域标签, 而其他研究^[51,57,68]则解除了这一限制. 2024年, He等人^[70]开发了一种基于时频自对比学习的域泛化方法, 用于工业电机的复合故障诊断. 与前述样本集群方法显著不同的是, 该方法利用信号处理技术获取单个样本的时域特征和频域特征来构建自对比. 因此, 它不依赖样本标签或域标签, 也不需要来自多个域的样本.

(4) 梯度对齐. 梯度对齐方法通过规范模型在各源域上的最优优化方向来寻找一个合适的公共方向, 以捕获域不变特征. 2024年, Ren等人^[71]提出了一种基于元学习的梯度对齐域泛化故障诊断框架, 其中元学习被用于规避计算梯度内积正则项时高阶导数难以求解的问题. 2024年, Ma等人^[72]开发了一种锐度感知梯度对齐方法用于标签含有噪声条件下的域泛化故障诊断, 该方法通过协调梯度从而在损失图上创建具有低损失值的平坦区域, 并在该区域上寻找平坦最小值来学习域不变性. 2023年, Ma等人^[73]设计了一种基于相互教学师生网络的梯度对齐域泛化诊断方法, 该方法利用细粒度梯度对齐来强调高阶领域一致性, 并与低阶分布统计数据协作以确保稳健的梯度对齐. 梯度对齐方法通常涉及求解高阶导数, 因此具有较高的优化难度.

3.1.3 特征解耦

该理论的方法与域对齐方法的主要区别在于: 它不仅关注域不变特征, 还强调提取域特定特征. 因此, 这类方法通常需要将特征分解为上述两种子特征来分别处理. 根据特征解耦的实现机制, 可将这类方法分为两类, 分别为多分支网络和因果学习.

(1) 多分支网络. 多分支网络可被视为一种集成学习方法, 后者已被广泛用于提升模型性能. 在域泛化中, 它们通常会如图9所示构建一个由域共享特征提取器和多个域特定分类器组成的多分支模型, 其中提取器处理域不变特征, 而每个分类器处理相应领域的域特定特征. 这类方法通常与域对齐方法高度相关. Zhao和Shen^[57]除使用CORAL和三元组损失实现域对齐外, 还使用多个域特定分类器保留域特定特征, 最终诊断结果由这些分类器共同决定. Wang等人^[58]首先使用多个域特定分类器学习域特定特征, 然后利用这些

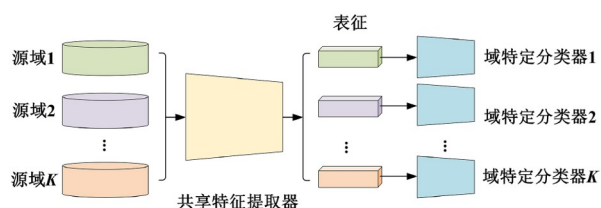


图9 (网络版彩图)基于多分支网络的域泛化方法的模型架构

Figure 9 (Color online) Model architecture of the domain generalization methods based on multi-branch networks.

分类器指导域共享特征提取器去除这些特征。需要注意的是, 尽管Zhao和Shen^[57]和Wang等人^[58]均使用多个分类器来提取域特定特征, 但Zhao和Shen^[57]保留了这些特征, 并利用投票机制将其用于决策, 而Wang等人^[58]则丢弃了这些特征, 以进一步提纯域不变特征。

(2) 因果学习。因果学习假定数据包含因果信息和非因果信息, 其中因果信息与标签相关且跨域不变, 而非因果信息与标签无关且与特定领域相关^[74]。开展因果学习通常首先就需要基于不同变量(噪声、负载、转速、机器等)的关系来构建可描述数据生成过程的结构因果模型。此外, 由于因果特征和非因果特征通常难以被观测和定义, 因此往往需要推测因果特征所具有的属性, 并在这些推论的基础上构建目标函数来对因果特征施加不变正则化, 从而重建因果机制。2022年, Li等人^[75]首次将因果学习应用于管道故障诊断, 将环境噪声视为背景, 而将速度、负载引起的特征视为域特定特征。同时, 使用层归一化来正则化噪声背景, 以及实例归一化来移除域特定特征。此外, Li等人^[75]还设计了一种因果损失用于捕获域不变特征。2023年, Guo等人^[76]提出了一种因果解释引导的域泛化模型, 不仅通过引导模型关注因果特征来增强模型的泛化能力, 还通过事后解释使用户得以进入诊断的决策环中。2024年, Jia等人^[77]开发了一种用于时间序列信号的因果解耦域泛化故障诊断方法, 其中因果聚合损失被用于分离故障特征中的因果信息和非因果信息, 并辅以重构损失确保所分离信息的完整性。因果学习也常用于跨机器故障诊断。2023年, Jia等人^[78]设计了一种深度因果图式分解网络来分离跨机器因果特征和非因果特征, 其中前者会最大化域判别器输出的熵, 而后者会最大化分类器输出的熵。2023年, Li等人^[79]开发了一种因果分解网络, 用于轴承连续退化模式的跨

机器知识泛化和诊断, 而不是像通常那样局限于固定级别的故障模式。2023年, Li等人^[80]提出了一种用于多机器协同域泛化故障诊断的因果一致性网络, 首次探讨了故障诊断中因果表征是否应该被施加不变正则化的问题。2024年, Cheng等人^[81]提出了一种名为对象条件域不变的学习规则赋能因果学习, 在源域数据是来自各种工况和机器的混合数据且无法访问域标签的条件下完成了轴承域泛化故障诊断。2024年, Zhu等人^[82]设计了一种因果和物理共同驱动的域泛化网络, 在挖掘故障因果关系的同时结合未知目标机器的物理先验知识, 实现了域不变特征提取和域特定知识嵌入。因果学习性能的关键在于能否正确地构建可描述当前任务的结构因果模型。

3.1.4 元学习

元学习, 也称学会学习, 旨在从多个相关任务中随机抽取的事件学习可用于未来学习的知识^[83,84]。基于元学习的域泛化方法并不明确要求模型学习某些表征以实现知识泛化, 而是将模型置于多个域偏移任务中, 并在这些任务中学习可用于应对未知域偏移任务的知识。通常, 源域会被随机划分为元训练域和元测试域, 元优化的目标是在元训练域上训练的模型能较好地泛化到元测试域上^[85]。元学习作为一种具有前景的域泛化方法备受计算机视觉领域学者的关注, 但在多源域泛化故障诊断领域中研究较少。2022年, Wang等人^[86]为域泛化诊断模型设计了模型不可知的学习程序, 通过最大化源域梯度间的点积, 寻找到域不变的梯度方向。

3.1.5 模型可解释性

模型可解释性已成为故障诊断领域一个炙手可热的研究方向^[87-90]。现有可解释故障诊断研究分为事前解释和事后解释两类, 前者通过在深层模型中嵌入可解释模块以赋予诊断结果物理意义; 后者则以诊断结果为基础反推模型做出决策的深层逻辑。由于事前解释方法通常认为在模型中嵌入物理知识可增强模型的泛化性, 因此十分适合域泛化故障诊断任务。当然, 利用事后解释方法帮助用户理解域泛化诊断模型并建立用户与模型间的依赖关系也同样重要。2022年, Liu等人^[91]基于Morlet小波变换设计了一种可解释卷积神经网络, 并证明了其提取的散射特征对线性时变系统的

不变性, 表明该模型可跨不同信号传输路径进行域泛化故障诊断. 2024年, Zhu等人^[92]提出了一种解耦可解释鲁棒域泛化网络实现了从人工数据到真实数据的跨域故障诊断, 其中, 神经基函数解耦模块被用于将信号分解为故障相关及故障无关基函数, 并辅以剪枝网络将后者去除.

3.2 单源域泛化

如图10所示, 现有单源域泛化故障诊断方法主要基于数据增强、域对齐以及模型可解释性理论.

由于单一源域提供的知识较少, 因此通常需要构建生成式模型来生成潜在域. 随后, 特征提取器需要匹配源域和潜在域. 这种数据增强和域对齐的组合模式受到了学者的青睐. 2023年, Zhao和Shen^[20]提出了一种对抗互信息指导的单源域泛化故障诊断网络, 其生成器通过最小化生成样本与源域样本间的互信息来生成多样的样本. 同时, 它的特征提取器利用对比损失来最大化它们间的互信息以实现类级域对齐. 2024年, Wang等人^[93]设计了一种基于多尺度风格生成的对抗对比模型, 其运行机制与Zhao和Shen^[20]所提的方法类似. 该方法^[93]相对前述方法^[20]的改进之处在于将生成器设计为基于风格学习的多分支网络. 2023年, Jiang等人^[94]设计了一种基于VIT的条件对抗域泛化网络进行冷凝机故障诊断, 其中协同条件域判别策略被用于增强模型对所提取的全局统计特征的分布判别能力, 进而对齐它们的条件分布. 2024年, Pu等人^[95]提出了一种单源增量泛化网络, 其特点是以增量的方式迭代生成潜在域. 此外, 该方法在最大化源域与潜在域的InfoNCE损失以确保潜在域有别于源域的同时, 还辅以循环一致性损失保持潜在域与源域语义一致. 需指出的是, 在先前所提的基于生成式模型的域泛化故障诊断研究^[41,42]中, 生成式模型和诊断模型是分离的, 进行对抗训练的是GAN中的生成器和域判别器. 然而, 上述单源域泛化研究^[20,93-95]中的生成式模型和诊断模型如图11所示被整合为一体, 其中诊断模型的特征提取器充当了先前GAN的判别器, 因此, 进行对抗训练的是生成器和特征提取器, 而域判别器仅仅发挥判断源域和潜在域的相似程度的作用.

在多源域泛化中, 不同域的样本以及它们的域标签可通过Mixup进行线性组合来生成具有新域标签的样本. 然而, 在仅有一个源域的单源域泛化任务中, 通

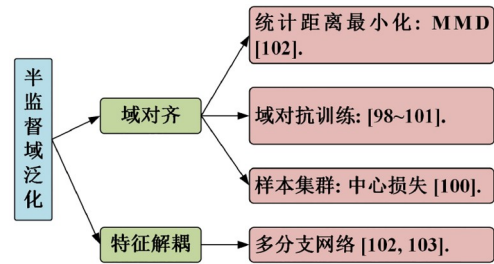


图 10 (网络版彩图)基于方法理论的单源域泛化故障诊断研究分类法

Figure 10 (Color online) Taxonomy of single-source domain generalization fault diagnosis studies based on methodology.

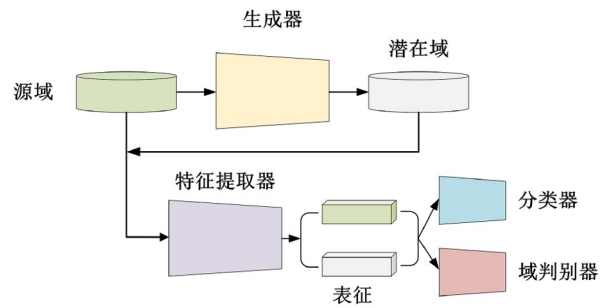


图 11 (网络版彩图)基于数据增强的对抗单源域泛化网络范式

Figure 11 (Color online) Adversarial single-source domain generalization network paradigm based on data augmentation.

过Mixup合成潜在域样本是一个具有挑战性的问题. 2024年, Tang等人^[96]提出了一种基于Mixup和直方图匹配的HmSeNet模型以实现单源域泛化, 其中Mixup用于将原始样本与直方图匹配生成的样本相结合来获取新样本. 这些样本与源域样本具有相同的语义, 但服从不同的分布.

除采用数据增强外, 在模型中嵌入先验物理知识也能缓解单源域信息不足的问题. 2024年, Kim等人^[97]提出了一种用于单域泛化的物理可解释轴承故障诊断框架, 利用信号预处理方法将先验知识——轴承信号中的故障特征是具有特征故障顺序的脉冲信号, 嵌入到诊断模型中, 从而提高了模型的泛化性.

总体上, 关于单源域泛化的故障诊断研究仍然相对较少. 这一领域的主要挑战是如何突破单源域信息不足的瓶颈. 构建一个嵌入物理知识的生成式模型可能是一个具有应用前景的方案. 此外, 不明确要求使用域标签的多源域泛化方法也可作为单源域泛化任务的潜在应对手段.

3.3 半监督域泛化

如图12所示, 现有半监督域泛化故障诊断研究主要集中在如何对齐有标签源域和无标签源域, 这与要求对齐有标签源域和无标签目标域的域适应类似. 两者的区别仅在于, 前者的目标域是未知的新域, 而后的目标域参与了训练. 尽管它们间存在此差异, 但由于这两种方法都旨在识别域不变特征, 而这些特征理应在不同域中通用, 因此半监督域泛化和域适应所使用的方法可相互借鉴. 此外, 无论是域适应还是半监督域泛化, 实现类级域对齐通常需要利用有标签样本的知识来赋予无标签样本伪标签, 从而补充无标签域中缺失的类信息.

2020年, Liao等人^[98]设计了一种基于Wasserstein GAN、梯度惩罚和伪标签学习的深度半监督域泛化诊断模型. 2022年, Li等人^[99]提出了一种用于轴承故障诊断的对抗域泛化网络, 其中域判别器由两个分类器组成, 它们的决策差异用于刻画不同域表征间的相似性. 2023年, Ren等人^[100]开发了一种深度模糊泛化网络, 利用对抗训练和基于伪标签的中心损失分别实现了全局域对齐和类级域对齐. 2023年, Ren等人^[101]利用两个网络分支分别提取判别特征和域不变特征, 并通过融合模块将其融合.

除上述域对齐方法外, 还有一些半监督域泛化故障诊断研究应用了特征解耦方法. 2023年, Zhao和Shen^[102]提出了一种多辅助半监督域泛化网络, 其中域对齐方法仅用于为无标签样本分配伪标签, 而低阶分

解用于指导主分支中域特定分类器和域不变分类器的训练. 2024年, Song等人^[103]提出了一种对比辅助域特定去除网络, 它利用域判别器指导域特定特征提取器学习有关域分类的信息, 同时令该提取器与故障分类器开展对抗训练来确保其所学特征与故障类别无关. 值得一提的是, 该研究利用域特定特征提取器来捕获域特定特征, 这有别于现有研究多使用域特定分类器来达成这一目标.

开展无监督域泛化故障诊断研究充分利用了大量无标签数据, 显著降低了标注成本. 它不仅使模型可泛化至未知的目标域, 还使其能像域适应模型那样处理来自无标签域的数据. 值得一提的是, 一个典型的域适应模型仅可用于应对单一的无标签域, 而一个半监督域泛化模型可同时处理多个无标签域. 为开发高性能的半监督域泛化模型, 一个关键挑战是如何为无标签样本分配高质量的伪标签. 此外, 还可进一步考虑如何在有标签域和无标签域不共享相同标签集的情况下匹配它们.

3.4 开集域泛化

现有开集域泛化故障诊断研究^[104~106]并未完全遵循第2节中所述的开集域泛化设置. 它们仅考虑目标域中存在未知故障类别的样本, 忽略了源域不共享相同标签集的情况. 此外, 在Lu等人^[104]所提的开集域泛化设置中, 源域存在目标域中不存在的私有类. 这并无实际意义, 其原因在于在域泛化中, 目标域仅作为测试集使用, 而不像域适应需要在训练阶段与源域适配.

检测目标域中未知类样本的任务类似于分布外(Out-of-distribution, OOD)检测^[107]这一主题, 其目的是仅通过学习分布内样本的知识来检测未知的分布外样本. 因此, 一些分布外检测方法, 如基于不确定性的方法, 基于重构的方法, 和基于softmax输出的方法可被应用于开集域泛化. 2022年, Ge等人^[105]提出了一种自适应诊断方法, 该方法不仅能检测目标域中的新类样本, 还能通过K-means算法对新类样本进一步分组. 2022年, Zhao和Shen^[106]设计了一种自适应开集域泛化网络来为每个已知类学习最优决策边界, 该边界会囊括相应类的样本, 并拒绝未知类的样本.

前述研究并未提及如何在第2节所述的源域不共享相同标签集的设置下处理源域样本. 在此, 本节提供一些潜在的解决方案. 由于源域间存在的严重类不匹

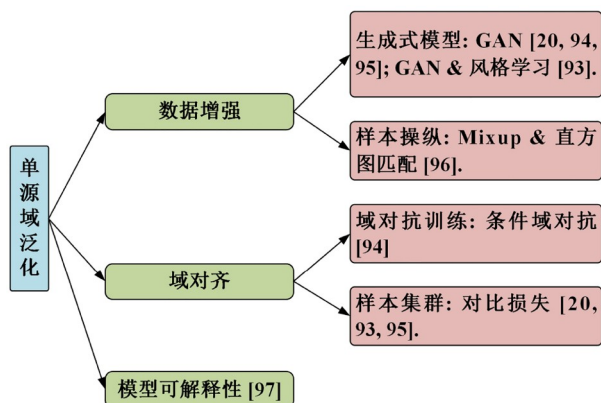


图12 (网络版彩图) 基于方法理论的半监督域泛化故障诊断研究分类法

Figure 12 (Color online) Taxonomy of semi-supervised domain generalization fault diagnosis studies based on methodology.

配使得它们的全局分布存在固有差异, 因此很难在全局层面对齐源域. 为此, 建议开发样本集群方法, 直接在类层面对齐源域. 其次, 不明确要求学习域不变表征的元学习也有望应用于开集域泛化. 此外, 考虑到对齐源域具有较高挑战, 一种有望可行的策略是用集成模型来重点学习域特定特征, 最终的诊断结果由所有集成成员共同决定. 除使用上述表征学习方法外, 还可利用数据增强方法, 如Mixup和CutMix^[108], 来补充源域中缺失的类别. 这些方法通过线性组合原始标签来生成新标签, 使得某些新标签即为所缺失的标签.

3.5 联邦域泛化

在联邦域泛化故障诊断中, 每个客户端仅能使用自己的私有数据, 而无法访问其他客户端持有的其他域的数据. 因此, 上述直接测量各域的表征分布差异或同类样本距离的方法难以适用. 此外, 在Cong等人^[109]所提的方法中, 客户端将本地训练学到的表征发送至中央服务器, 并由中央服务器计算表征分布间的差异. 这实际上是被禁止的, 其原因是表征包含数据的高维信息, 上传表征也会造成隐私泄露.

为以隐私保护的方式实现域间对齐, 一种可行的方法是构建生成器来生成一个参考分布, 同时所有客户端所学的表征分布都需要靠近此参考分布. 2022年, Wang等人^[110]提出了一种联邦对抗域泛化网络, 在本地训练阶段, 生成器生成的参考分布近似于本地表征分布, 而由这些本地生成器聚合而成的全局生成器则可生成所有客户端中心的参考分布, 用于指导模型实现域间对齐. 2024年, Li等人^[111]提出了一种融合共识知识的隐私保护联邦域泛化故障诊断方法, 其在云端的生成器以标签为输入进行逆映射来得到参考表征分布, 将集中式分布对齐问题转化为两两分布对齐问题. 2024年, Zhao和Shen^[112]开发了一种联邦蒸馏域泛化框架, 其中生成器会学习本地数据的分布. 与前述生成器生成参考分布来指导域对齐训练不同, 它直接生成虚拟数据来协助其余客户端, 而这些客户端会通过低秩分解方法对齐本地数据及虚拟数据以捕获域不变特征.

此外, 正如第3.4节所提及的, 既然对齐源域具有较高挑战, 不妨转而关注域特定表征. 2024年, Zhao和Shen^[113]开发了一个边云集成联邦学习框架, 其中各客

户端均有一个仅学习域特定知识的分类器. 在本地训练中, 特征提取器所提取的特征不仅需要对本地图像分类器有效, 而且对从未遇到本地数据的其他客户端上的分类器也需要同样有效. 这种方法确保了所学特征在不同领域具有通用性.

目前还有一些新颖的联邦域泛化故障诊断研究尝试解除传统联邦域泛化任务中所存在的一些限制. 例如, 联邦学习要求存在一个中心服务器用于处理客户端所上传的信息, 这可能会引起单点故障. 为此, 2024年, Xu等人^[114]提出了一种去中心化联邦域泛化故障诊断方法, 仅在相邻节点间交换模型权重, 并同时交换类集群统计信息以在类层面对齐相邻节点的表征分布. 2024年, Qian等人^[115]开展一种称为异构联邦域泛化的更为新颖的任务, 其中源客户端的数据与目标客户端的数据是异构的, 例如源客户端的是齿轮箱故障数据, 而目标客户端的是轴承故障数据.

4 研究展望

4.1 如何确保所生成潜在域的有效性

现有利用生成式模型生成潜在域以扩充训练样本的域泛化故障诊断研究多以所生成潜在域与源域的分布差异最大为目标来确保潜在域的多样性, 从而模拟真实的分布偏移. 然而, 仅仅以该目标训练生成式模型是否可确保所生成的潜在域能有效用于指导模型训练仍然存疑. 这是因为尽管在真实场景下各源域间存在分布偏移, 但它们也存在固有的相似性, 这种相似性是由故障生成机制所决定的. 因此, 并不是与源域存在分布差异的数据均可被认为是潜在域. 例如, 齿轮箱故障数据与轴承故障数据间存在明显的分布差异, 但轴承故障数据难以直接用于齿轮箱故障诊断模型的训练. 一味增大生成域与源域的差异可能会削弱潜在域的有效性, 反而有害于模型收敛. 因此, 有必要在生成式模型优化过程中添加更多约束, 在模拟分布偏移的同时也捕获域间固有的相似性.

4.2 仿真模型如何模拟实际工况下的背景噪声

利用仿真模型生成仿真数据对于提升智能诊断模型泛化性的确表现出巨大潜力. 通过调整仿真模型参数, 几乎可以获取任意机械设备在任意工作模式下的状态监测数据用于模型训练. 然而, 达成这一目标的

一个至关重要的前提是仿真模型精度必须足够高, 且能反映机械设备的真实运行状况。通常情况下, 机械设备在运行时所采集的信号伴随着严重的背景噪声, 而这是现有研究中难以进行模拟的。此外, 还有众多不确定性因素的影响, 例如边界条件、阻尼、零部件配合状况, 均使得仿真模型无法保证与实际物理系统之间完全匹配。为解决这一挑战, 除了必须进行完备的理论分析外, 可能还需要利用少量物理试验, 通过与仿真模型得到的动态响应进行匹配分析, 对仿真模型中敏感参数进行修正。

4.3 “域不变特征”是否具备域不变性

现有基于域对齐的域泛化故障诊断方法旨在将所有源域映射到一个共同的特征子空间, 并将以此捕获的特征视为可泛化至未知目标域的域不变特征。这种方法起初源于域适应, 其中源域与未标记目标域被映射至共同的子空间, 因此所提取的特征是跨源域与目标域不变的, 使得源域的标签知识可被用于处理未标记目标域。然而, 将这一做法沿用至域泛化似乎只是令所提取的特征跨各源域不变, 其对于未知目标域是否同样具有域不变性仍然存疑。为解决这一挑战, 有必要从故障生成机制出发寻找真正跨域不变的特征模式来开发域对齐方法。此外, 将域对齐方法与数据增强方法结合也是一种有效的策略, 这使得域对齐方法所捕获的域不变特征适用的潜在域更为多样, 甚至囊括各类未知目标域。

4.4 人机协同域泛化故障诊断

当前绝大部分域泛化故障诊断研究将模型准确率提升作为最终目标。尽管对于本质上为分类任务的故障诊断而言, 提升准确率十分重要, 但仅仅关注准确率似乎意味着将机器是否发生故障的决定权完全交付于智能模型。即使模型做出不可信的决策, 用户也无从知晓。相反, 一个值得信赖的诊断模型不仅需要足够准确, 还应在可能发生误判时进行预警。因此, 对诊断模型提供的结果进行不确定性估计和可信度分析就变得至关重要。若结果可信度高, 不确定性低, 则可作为用户做出决策的重要依据; 若结果可信度低, 不确定性高, 则用户应对机器的健康状况做进一步判断。不确定性估计和可信度分析可使用户进入决策环, 实现人机协同域泛化故障诊断^[116,117]。

4.5 持续域泛化

尽管域泛化模型“开箱即用”的特性使其无需像域适应那样为目标域重新训练一个模型, 但考虑到机器的状态监测信号通常会持续生成, 利用新数据训练域泛化模型来为其补充诊断知识也存在必要。将新数据与旧数据融合来重新训练模型的方法或许有效, 但有较高的计算代价。此外, 用新数据进一步训练模型可能会导致灾难性遗忘, 即所学的新知识会干扰旧知识, 甚至完全覆盖旧知识。因此, 开发持续学习(也称增量学习或终身学习)技术^[118]并应用于域泛化故障诊断具有重要意义。它使得模型可处理来自传感器的连续信号流, 赋予其在学习新知识的同时保留甚至优化旧知识的能力。

4.6 多模态数据融合

为尽可能提升模型泛化性, 有必要充分利用各类可用于指示机器健康状态的监测数据。这些信号可能是振动信号、声信号、电信号或热图像, 呈现出多模态的形式。它们之间必定会存在统计差异和特征冗余, 如何对齐这些多模态数据, 捕捉不同模态数据之间的互补特征并去除冗余特征具有挑战性。考虑到基于多模态数据(文本、图像和视频)的GPT模型在最近所取得的巨大成功, 使用多模态数据训练域泛化故障诊断模型具有前景。

5 总结

分布偏移已成为制约智能故障诊断模型在实际工程中部署的一个重要因素。为缓解分布偏移问题, 具有“开箱即用”优势的域泛化方法越来越受到学者的关注, 各类域泛化故障诊断方法也层出不穷。在此背景下, 本文首先按任务类型将现有域泛化故障诊断研究总体划分为多源域泛化、单源域泛化、半监督域泛化、开集域泛化, 以及联邦域泛化研究。然后, 根据方法的基本理论对它们进一步细分, 主要涉及数据增强、域对齐、特征解耦、元学习, 以及模型可解释性等理论。随后, 阐述了各理论的技术在域泛化任务中所能发挥的作用, 并回顾了它们的实际应用案例。最后, 指出了当前域泛化故障诊断研究仍面临的潜在挑战以及未来具有前景的发展方向。

参考文献

- 1 Lei Y G, Jia F, Kong D T, et al. Opportunities and challenges of machinery intelligent fault diagnosis in big data era (in Chinese). *J Mech Eng*, 2018, 54: 94–104 [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. 机械工程学报, 2018, 54: 94–104]
- 2 Lei Y G, Yang B, Du Z J, et al. Deep transfer diagnosis method for machinery in big data era (in Chinese). *J Mech Eng*, 2019, 55: 1–8 [雷亚国, 杨彬, 杜兆钧, 等. 大数据下机械装备故障的深度迁移诊断方法. 机械工程学报, 2019, 55: 1–8]
- 3 Jiang H K, Shao H D, Li X Q. Deep learning theory with application in intelligent fault diagnosis of aircraft (in Chinese). *J Mech Eng*, 2019, 55: 27–34 [姜洪开, 邵海东, 李兴球. 基于深度学习的飞行器智能故障诊断方法. 机械工程学报, 2019, 55: 27–34]
- 4 Zhou K, Liu Z, Qiao Y, et al. Domain generalization: A survey. *IEEE Trans Neural Netw Learn Syst*, 2023, 45: 4396–4415
- 5 Wang J, Lan C, Liu C, et al. Generalizing to unseen domains: A survey on domain generalization. *IEEE Trans Knowl Data Eng*, 2023, 35: 8052–8072
- 6 Xiao Y, Shao H, Wang J, et al. Bayesian variational transformer: A generalizable model for rotating machinery fault diagnosis. *Mech Syst Signal Processing*, 2024, 207: 110936
- 7 Shao H D, Xiao Y M, Deng Q W, et al. Trustworthy mechanical fault diagnosis using uncertainty-aware network (in Chinese). *J Mech Eng*, 2024, 60: 194–206 [邵海东, 肖一鸣, 邓乾旺, 等. 基于不确定性感知网络的可信机械故障诊断. 机械工程学报, 2024, 60: 194–206]
- 8 Yan S, Shao H, Wang J, et al. LiConvFormer: A lightweight fault diagnosis framework using separable multiscale convolution and broadcast self-attention. *Expert Syst Appl*, 2024, 237: 121338
- 9 Shao H D, Zhang X Y, Cheng J S, et al. Intelligent fault diagnosis of bearing using enhanced deep transfer auto-encode (in Chinese). *J Mech Eng*, 2020, 56: 84–90 [邵海东, 张笑阳, 程军圣, 等. 基于提升深度迁移自动编码器的轴承智能故障诊断. 机械工程学报, 2020, 56: 84–90]
- 10 Li X, Grandvalet Y, Davoine F. Explicit inductive bias for transfer learning with convolutional networks. In: *Proceedings of 35th International Conference on Machine Learning (ICML)*. Stockholm, 2018
- 11 Li X, Xiong H, Wang H, et al. DELTA: Deep learning transfer using feature map with attention for convolutional networks. In: *Proceedings of 7th International Conference on Learning Representations (ICLR)*. New Orleans, 2019
- 12 Liu B, Cai Y, Guo Y, et al. TransTailor: Pruning the pre-trained model for improved transfer learning. In: *Proceedings of 35th AAAI Conference on Artificial Intelligence/33rd Conference on Innovative Applications of Artificial Intelligence/11th Symposium on Educational Advances in Artificial Intelligence*. Electr Network, 2021. 35: 8627–8634
- 13 Shao S, McAleer S, Yan R, et al. Highly accurate machine fault diagnosis using deep transfer learning. *IEEE Trans Ind Inf*, 2019, 15: 2446–2455
- 14 Shao H D, Xiao Y M, Yan S, et al. Simulation data-driven enhanced unsupervised domain adaptation for bearing fault diagnosis (in Chinese). *J Mech Eng*, 2022, 59: 76–85 [邵海东, 肖一鸣, 颜深, 等. 仿真数据驱动的改进无监督域适应轴承故障诊断. 机械工程学报, 2022, 59: 76–85]
- 15 Lei Y G, Yang B, Li N P, et al. Targeted transfer diagnosis method across different machines (in Chinese). *J Mech Eng*, 2022, 58: 1–9 [雷亚国, 杨彬, 李乃鹏, 等. 跨设备的机械故障靶向迁移诊断方法. 机械工程学报, 2022, 58: 1–9]
- 16 Shao H D, Chen X K, Cao H R, et al. Internal adversarial guided unsupervised multi-domain adaptation network for collaborative fault diagnosis of bearing (in Chinese). *Sci Sin Tech*, 2023, 53: 1229–1240 [邵海东, 陈星恺, 曹鸿儒, 等. 内对抗指导无监督多域适配网络的轴承协同故障诊断. 中国科学: 技术科学, 2023, 53: 1229–1240]
- 17 Xiao Y, Shao H, Min Z, et al. Multiscale dilated convolutional subdomain adaptation network with attention for unsupervised fault diagnosis of rotating machinery cross operating conditions. *Measurement*, 2022, 204: 112146
- 18 Zhou K, Yang Y, Qiao Y, et al. Domain adaptive ensemble learning. *IEEE Trans Image Process*, 2021, 30: 8008–8018
- 19 Shu Y, Cao Z, Wang C, et al. Open domain generalization with domain-augmented meta-learning. In: *Proceedings of 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. Nashville, 2021. 9619–9628
- 20 Zhao C, Shen W. Adversarial mutual information-guided single domain generalization network for intelligent fault diagnosis. *IEEE Trans Ind Inf*, 2023, 19: 2909–2918
- 21 Hakim M, Omran A A B, Ahmed A N, et al. A systematic review of rolling bearing fault diagnoses based on deep learning and transfer learning: Taxonomy, overview, application, open challenges, weaknesses and recommendations. *Ain Shams Eng J*, 2023, 14: 101945
- 22 Chen X, Yang R, Xue Y, et al. Deep transfer learning for bearing fault diagnosis: A systematic review since 2016. *IEEE Trans Instrum Meas*, 2023, 72: 3508221

- 23 Misbah I, Lee C K M, Keung K L. Fault diagnosis in rotating machines based on transfer learning: Literature review. *Knowledge-Based Syst*, 2024, 283: 111158
- 24 Yan R, Shen F, Sun C, et al. Knowledge transfer for rotary machine fault diagnosis. *IEEE Sens J*, 2020, 20: 8374–8393
- 25 Yang D, Zhang W, Jiang Y. Mechanical fault diagnosis based on deep transfer learning: A review. *Meas Sci Technol*, 2023, 34: 112001
- 26 Zhao Z, Zhang Q, Yu X, et al. Applications of unsupervised deep transfer learning to intelligent fault diagnosis: A survey and comparative study. *IEEE Trans Instrum Meas*, 2021, 70: 3525828
- 27 Zhang S, Su L, Gu J, et al. Rotating machinery fault detection and diagnosis based on deep domain adaptation: A survey. *Chin J Aeronautics*, 2023, 36: 45–74
- 28 Zhao C, Zio E, Shen W. Domain generalization for cross-domain fault diagnosis: An application-oriented perspective and a benchmark study. *Reliability Eng Syst Saf*, 2024, 245: 109964
- 29 Han S Y, Shao H D, Jian H K, et al. Intelligent fault diagnosis of aero-engine high-speed bearings using enhanced CNN (in Chinese). *Chin J Aeronaut*, 2022, 43: 158–171 [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. *航空学报*, 2022, 43: 158–171]
- 30 Han S, Shao H, Huo Z, et al. End-to-end chiller fault diagnosis using fused attention mechanism and dynamic cross-entropy under imbalanced datasets. *Building Environ*, 2022, 212: 108821
- 31 Shao H D, Xiao Y M, Min Z S, et al. Blockchain and edge computing enabled federated learning fault diagnosis framework (in Chinese). *J Mech Eng*, 2023, 59: 283–292 [邵海东, 肖一鸣, 闵志闪, 等. 区块链和边缘计算赋能的联邦学习故障诊断框架. *机械工程学报*, 2023, 59: 283–292]
- 32 McMahan H, Moore E, Ramage D, et al. Communication-efficient learning of deep networks from decentralized data. In: *Proceedings of 19th International Conference on Artificial Intelligence and Statistics (AISTATS)*. Cadiz, 2016
- 33 Li T, Sahu A, Zaheer M, et al. Federated optimization in heterogeneous networks. In: *Proceedings of 3rd Conference on Machine Learning and Systems (MLSys)*. Austin, 2020
- 34 Ma Z, Ma J, Miao Y, et al. ShieldFL: Mitigating model poisoning attacks in privacy-preserving federated learning. *IEEE Trans Inform Forensic Secur*, 2022, 17: 1639–1654
- 35 Zhao K, Hu J, Shao H, et al. Federated multi-source domain adversarial adaptation framework for machinery fault diagnosis with data privacy. *Reliability Eng Syst Saf*, 2023, 236: 109246
- 36 Xu S, Ma J, Song D, et al. Open-set federated adversarial domain adaptation based cross-domain fault diagnosis. *Mech Syst Signal Process*, 2023, 34: 115004
- 37 Liu Q, Chen C, Qin J, et al. FedDG: Federated domain generalization on medical image segmentation via episodic learning in continuous frequency space. In: *Proceedings of 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. Nashville, 2021. 1013–1023
- 38 Zhang R, Xu Q, Yao J, et al. Federated domain generalization with generalization adjustment. In: *Proceedings of 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. Vancouver, 2023. 3954–3963
- 39 Zheng H, Wang R, Yang Y, et al. Intelligent fault identification based on multisource domain generalization towards actual diagnosis scenario. *IEEE Trans Ind Electron*, 2020, 67: 1293–1304
- 40 Shao H, Li W, Cai B, et al. Dual-threshold attention-guided GAN and limited infrared thermal images for rotating machinery fault diagnosis under speed fluctuation. *IEEE Trans Ind Inf*, 2023, 19: 9933–9942
- 41 Zhuang J, Jia M, Ding Y, et al. Health assessment of rotating equipment with unseen conditions using adversarial domain generalization toward self-supervised regularization learning. *IEEE ASME Trans Mechatron*, 2022, 27: 4675–4685
- 42 Wang C, Wang Z, Liu Q, et al. Support-sample-assisted domain generalization via attacks and defenses: Concepts, algorithms, and applications to pipeline fault diagnosis. *IEEE Trans Ind Inf*, 2024, 20: 6413–6423
- 43 Zhang H, Cisse M, Dauphin Y, et al. Mixup: Beyond empirical risk minimization. arXiv: [1710.09412](https://arxiv.org/abs/1710.09412), 2017
- 44 Shi Y, Deng A, Deng M, et al. Domain augmentation generalization network for real-time fault diagnosis under unseen working conditions. *Reliability Eng Syst Saf*, 2023, 235: 109188
- 45 Xiao Y, Shao H, Wang J, et al. Domain-augmented meta ensemble learning for mechanical fault diagnosis from heterogeneous source domains to unseen target domains. *Expert Syst Appl*, 2024, 259: 125345

- 46 Li Q, Chen L, Kong L, et al. Cross-domain augmentation diagnosis: An adversarial domain-augmented generalization method for fault diagnosis under unseen working conditions. *Reliability Eng Syst Saf*, 2023, 234: 109171
- 47 Zhao C, Shen W. Imbalanced domain generalization via Semantic-Discriminative augmentation for intelligent fault diagnosis. *Adv Eng Inf*, 2024, 59: 102262
- 48 Goodfellow I, Shlens J, Szegedy C. Explaining and harnessing adversarial examples. In: Proceedings of 3rd International Conference on Learning Representations (ICLR). San Diego, 2015
- 49 Moosavi-Dezfooli S M, Fawzi A, Frossard P. DeepFool: A simple and accurate method to fool deep neural networks. In: Proceedings of 2016 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas, 2016. 2574–2582
- 50 Papernot N, McDaniel P, Jha S, et al. The limitations of deep learning in adversarial settings. In: Proceedings of 2016 IEEE European Symposium on Security and Privacy (EuroS&P). Saarbruecken, 2016. 372–387
- 51 Han T, Li Y F, Qian M. A hybrid generalization network for intelligent fault diagnosis of rotating machinery under unseen working conditions. *IEEE Trans Instrum Meas*, 2021, 70: 3520011
- 52 Tang Z, Liu X, Sun K, et al. Fault diagnosis of rotating machinery under multiple operating conditions generalization: A representation gradient muting paradigm. *IEEE Trans Instrum Meas*, 2022, 71: 3525110
- 53 Shi Z, Chen J, Zi Y, et al. Semi-physical simulation-driven contrastive decoupling net for intelligent fault diagnosis of unseen machines under varying speed. *Meas Sci Technol*, 2024, 35: 076101
- 54 Gong F, Ma P, Wang N, et al. Cross-device fault diagnosis of rolling bearings using domain generalization and dynamic model. *J Vibration Contr*, 2024, doi: 10.1177/10775463241256253
- 55 Pang B, Liu Q, Xu Z, et al. Fault vibration model driven fault-aware domain generalization framework for bearing fault diagnosis. *Adv Eng Inf*, 2024, 62: 102620
- 56 Zhu Y, Zhuang F, Wang J, et al. Deep subdomain adaptation network for image classification. *IEEE Trans Neural Netw Learn Syst*, 2021, 32: 1713–1722
- 57 Zhao C, Shen W. A domain generalization network combing invariance and specificity towards real-time intelligent fault diagnosis. *Mech Syst Signal Processing*, 2022, 173: 108990
- 58 Wang R, Huang W, Lu Y, et al. A novel domain generalization network with multidomain specific auxiliary classifiers for machinery fault diagnosis under unseen working conditions. *Reliability Eng Syst Saf*, 2023, 238: 109463
- 59 Fan Z, Xu Q, Jiang C, et al. Deep mixed domain generalization network for intelligent fault diagnosis under unseen conditions. *IEEE Trans Ind Electron*, 2024, 71: 965–974
- 60 Hu R, Zhang M, Meng X, et al. Deep subdomain generalisation network for health monitoring of high-speed train brake pads. *Eng Appl Artif Intelligence*, 2022, 113: 104896
- 61 Li Y, Tian X, Gong M, et al. Deep domain generalization via conditional invariant adversarial networks. In: Proceedings of 15th European Conference on Computer Vision (ECCV). Munich, 2018. 624–639
- 62 Chen L, Li Q, Shen C, et al. Adversarial domain-invariant generalization: A generic domain-regressive framework for bearing fault diagnosis under unseen conditions. *IEEE Trans Ind Inf*, 2022, 18: 1790–1800
- 63 Shi Y, Deng A, Deng M, et al. Domain transferability-based deep domain generalization method towards actual fault diagnosis scenarios. *IEEE Trans Ind Inf*, 2023, 19: 7355–7366
- 64 Zhang Q, Zhao Z, Zhang X, et al. Conditional adversarial domain generalization with a single discriminator for bearing fault diagnosis. *IEEE Trans Instrum Meas*, 2021, 70: 3514515
- 65 Shen B, Zhang M, Yao L, et al. Novel triplet loss-based domain generalization network for bearing fault diagnosis with unseen load condition. *Processes*, 2024, 12: 882
- 66 Wang B, Wen L, Li X, et al. Adaptive class center generalization network: A sparse domain-regressive framework for bearing fault diagnosis under unknown working conditions. *IEEE Trans Instrum Meas*, 2023, 72: 3516511
- 67 Jia L, Chow T, Wang Y, et al. Dynamic balanced dual prototypical domain generalization for cross-machine fault diagnosis. *IEEE Trans Instrum Meas*, 2024, 73: 3516610
- 68 Ragab M, Chen Z, Zhang W, et al. Conditional contrastive domain generalization for fault diagnosis. *IEEE Trans Instrum Meas*, 2022, 71: 3506912

- 69 Shi Z, Chen J, Zhang X, et al. A reliable feature-assisted contrastive generalization net for intelligent fault diagnosis under unseen machines and working conditions. *Mech Syst Signal Processing*, 2023, 188: 110011
- 70 He Y, Zhao C, Shen W. Cross-domain compound fault diagnosis of machine-level motors via time–frequency self-contrastive learning. *IEEE Trans Ind Inf*, 2024, 20: 9692–9701
- 71 Ren L, Mo T, Cheng X. Meta-learning based domain generalization framework for fault diagnosis with gradient aligning and semantic matching. *IEEE Trans Ind Inf*, 2024, 20: 754–764
- 72 Ma Y, Yang J, Yan R. Sharpness-aware gradient alignment for domain generalization with noisy labels in intelligent fault diagnosis. *IEEE Trans Instrum Meas*, 2024, 73: 3523210
- 73 Ma Y, Yang J, Li L. Gradient aligned domain generalization with a mutual teaching teacher-student network for intelligent fault diagnosis. *Reliability Eng Syst Saf*, 2023, 239: 109516
- 74 Liu C, Sun X, Wang J, et al. Learning causal semantic representation for out-of-distribution prediction. arXiv: 2011.01681, 2020
- 75 Li J, Wang Y, Zi Y, et al. Whitening-Net: A generalized network to diagnose the faults among different machines and conditions. *IEEE Trans Neural Netw Learn Syst*, 2022, 33: 5845–5858
- 76 Guo C, Zhao Z, Ren J, et al. Causal explaining guided domain generalization for rotating machinery intelligent fault diagnosis. *Expert Syst Appl*, 2024, 243: 122806
- 77 Jia L, Chow T W S, Yuan Y. Causal disentanglement domain generalization for time-series signal fault diagnosis. *Neural Networks*, 2024, 172: 106099
- 78 Jia S, Li Y, Wang X, et al. Deep causal factorization network: A novel domain generalization method for cross-machine bearing fault diagnosis. *Mech Syst Signal Processing*, 2023, 192: 110228
- 79 Li J, Wang Y, Zi Y, et al. Causal disentanglement: A generalized bearing fault diagnostic framework in continuous degradation mode. *IEEE Trans Neural Netw Learn Syst*, 2023, 34: 6250–6262
- 80 Li J, Wang Y, Zi Y, et al. Causal consistency network: A collaborative multimachine generalization method for bearing fault diagnosis. *IEEE Trans Neural Netw Learn Syst*, 2023, 19: 5915–5924
- 81 Cheng L, Kong X, Zhang Y, et al. A novel causal feature learning-based domain generalization framework for bearing fault diagnosis with a mixture of data from multiple working conditions and machines. *Adv Eng Inf*, 2024, 62: 102622
- 82 Zhu Y, Zi Y, Li J, et al. PhysiCausalNet: A causal- and physics-driven domain generalization network for cross-machine fault diagnosis of unseen domain. *IEEE Trans Ind Inf*, 2024, 20: 8488–8498
- 83 Lin J, Shao H, Min Z, et al. Cross-domain fault diagnosis of bearing using improved semi-supervised meta-learning towards interference of out-of-distribution samples. *Knowledge-Based Syst*, 2022, 252: 109493
- 84 Feng Y, Chen J, Xie J, et al. Meta-learning as a promising approach for few-shot cross-domain fault diagnosis: Algorithms, applications, and prospects. *Knowledge-Based Syst*, 2022, 235: 107646
- 85 Li D, Yang Y, Song Y, et al. Learning to generalize: Meta-learning for domain generalization. arXiv: 1710.03463, 2017
- 86 Wang H, Bai X, Wang S, et al. Generalization on unseen domains via model-agnostic learning for intelligent fault diagnosis. *IEEE Trans Instrum Meas*, 2022, 71: 3506411
- 87 Li Y, Zhou Z, Sun C, et al. Variational attention-based interpretable transformer network for rotary machine fault diagnosis. *IEEE Trans Neural Netw Learn Syst*, 2024, 35: 6180–6193
- 88 Li T, Zhao Z, Sun C, et al. WaveletKernelNet: An interpretable deep neural network for industrial intelligent diagnosis. *IEEE Trans Syst Man Cybern Syst*, 2022, 52: 2302–2312
- 89 Shang Z, Zhao Z, Yan R. Denoising fault-aware wavelet network: A signal processing informed neural network for fault diagnosis. *Chin J Mech Eng*, 2023, 36: 9
- 90 Lin J, Jiao J Y. Research progress and challenges of interpretable mechanical intelligent diagnosis (in Chinese). *J Mech Eng*, 2023, 59: 215–224 [林京, 焦金阳. 可解释机械智能诊断技术的研究进展与挑战. *机械工程学报*, 2023, 59: 215–224]
- 91 Liu C, Ma X, Han T, et al. NTScatNet: An interpretable convolutional neural network for domain generalization diagnosis across different transmission paths. *Measurement*, 2022, 204: 112041
- 92 Zhu Q, Liu H, Bao C, et al. Decoupled interpretable robust domain generalization networks: A fault diagnosis approach across bearings, working conditions, and artificial-to-real scenarios. *Adv Eng Inf*, 2024, 61: 102445

- 93 Wang J, Ren H, Shen C, et al. Multi-scale style generative and adversarial contrastive networks for single domain generalization fault diagnosis. [Reliability Eng Syst Saf](#), 2024, 243: 109879
- 94 Jiang K, Gao X, Gao H, et al. VIT-GADG: A generative domain-generalized framework for chillers fault diagnosis under unseen working conditions. *IEEE Trans Instrum Meas*, 2023, 72: 3527413
- 95 Pu Y, Tang J, Li X, et al. Single-domain incremental generation network for machinery intelligent fault diagnosis under unknown working speeds. [Adv Eng Inf](#), 2024, 60: 102400
- 96 Tang J, Ding X, Wei C, et al. HmmSeNet: A novel single domain generalization equipment fault diagnosis under unknown working speed using histogram matching mixup. [IEEE Trans Ind Inf](#), 2024, 20: 7162–7172
- 97 Kim I, Wook Kim S, Kim J, et al. Single domain generalizable and physically interpretable bearing fault diagnosis for unseen working conditions. [Expert Syst Appl](#), 2024, 241: 122455
- 98 Liao Y, Huang R, Li J, et al. Deep semisupervised domain generalization network for rotary machinery fault diagnosis under variable speed. *IEEE Trans Instrum Meas*, 2020, 69: 8064–8075
- 99 Li J, Shen C, Kong L, et al. A new adversarial domain generalization network based on class boundary feature detection for bearing fault diagnosis. *IEEE Trans Instrum Meas*, 2022, 71: 2506909
- 100 Ren H, Wang J, Zhu Z, et al. Domain fuzzy generalization networks for semi-supervised intelligent fault diagnosis under unseen working conditions. [Mech Syst Signal Processing](#), 2023, 200: 110579
- 101 Ren H, Wang J, Huang W, et al. Domain-invariant feature fusion networks for semi-supervised generalization fault diagnosis. [Eng Appl Artif Intelligence](#), 2023, 126: 107117
- 102 Zhao C, Shen W. Mutual-assistance semisupervised domain generalization network for intelligent fault diagnosis under unseen working conditions. [Mech Syst Signal Processing](#), 2023, 189: 110074
- 103 Song Q, Jiang X, Liu J, et al. Contrast-assisted domain-specificity-removal network for semi-supervised generalization fault diagnosis. *IEEE Trans Neural Netw Learn Syst*, 2024, doi: 10.1109/TNNLS.2024.3383467
- 104 Lu B, Zhang Y, Sun Q, et al. A novel multidomain contrastive-coding-based open-set domain generalization framework for machinery fault diagnosis. [IEEE Trans Ind Inf](#), 2024, 20: 6369–6381
- 105 Ge Y, Zhang F, Ren Y. Adaptive fault diagnosis method for rotating machinery with unknown faults under multiple working conditions. [J Manufacturing Syst](#), 2022, 63: 177–184
- 106 Zhao C, Shen W. Adaptive open set domain generalization network: Learning to diagnose unknown faults under unknown working conditions. [Reliability Eng Syst Saf](#), 2022, 226: 108672
- 107 Han T, Li Y F. Out-of-distribution detection-assisted trustworthy machinery fault diagnosis approach with uncertainty-aware deep ensembles. [Reliability Eng Syst Saf](#), 2022, 226: 108648
- 108 Yun S, Han D, Chun S, et al. CutMix: Regularization strategy to train strong classifiers with localizable features. In: *Proceedings of 2019 IEEE/CVF International Conference on Computer Vision (ICCV)*. Seoul, 2019. 6022–6031
- 109 Cong X, Song Y, Li Y, et al. Federated domain generalization with global robust model aggregation strategy for bearing fault diagnosis. [Meas Sci Technol](#), 2023, 34: 115116
- 110 Wang R, Huang W, Shi M, et al. Federated adversarial domain generalization network: A novel machinery fault diagnosis method with data privacy. [Knowledge-Based Syst](#), 2022, 256: 109880
- 111 Li B, Song P, Zhao C. Fusing consensus knowledge: A federated learning method for fault diagnosis via privacy-preserving reference under domain shift. [Inf Fusion](#), 2024, 106: 102290
- 112 Zhao C, Shen W. A federated distillation domain generalization framework for machinery fault diagnosis with data privacy. [Eng Appl Artif Intelligence](#), 2024, 130: 107765
- 113 Zhao C, Shen W. Federated domain generalization: A secure and robust framework for intelligent fault diagnosis. [IEEE Trans Ind Inf](#), 2024, 20: 2662–2670
- 114 Xu D, Jia M, Chen T, et al. Decentralized federated domain generalization with cluster alignment for fault diagnosis. [Control Eng Pract](#), 2024, 148: 105951
- 115 Qian Q, Luo J, Qin Y. Heterogeneous federated domain generalization network with common representation learning for cross-load machinery fault diagnosis. *IEEE Trans Syst Man Cyber Syst*, 2024, 54: 5704–5716

- 116 Zhou T, Han T, Droguett EL. Towards trustworthy machine fault diagnosis: A probabilistic Bayesian deep learning framework. [Reliability Eng Syst Saf](#), 2022, 224: 108525
- 117 Xiao Y, Shao H, Feng M, et al. Towards trustworthy rotating machinery fault diagnosis via attention uncertainty in transformer. [J Manufacturing Syst](#), 2023, 70: 186–201
- 118 Lange M, Aljundi R, Masana M, et al. A continual learning survey: Defying forgetting in classification tasks. [IEEE Trans Pattern Anal Mach Intell](#), 2022, 44: 3366–3385

Progress and prospects of domain generalization mechanical fault diagnosis research

SHAO HaiDong, XIAO YiMing, YAN Shen, WANG Jie & PENG Ying

College of Mechanical and Vehicle Engineering, Hunan University, Changsha 410082, China

Distribution shifts pose challenges for deploying intelligent fault diagnosis models in real-world engineering applications. To address these shifts, research on intelligent fault diagnosis using transfer learning has been actively pursued. Among the unsupervised transfer learning methods, domain adaptation and domain generalization are particularly noteworthy. Domain adaptation requires unlabeled target domain data, while domain generalization eliminates this need, offering greater flexibility and versatility. Since domain generalization was first applied to fault diagnosis in 2020, numerous diagnostic techniques have emerged, necessitating a thorough review. This paper first classifies existing domain generalization fault diagnosis studies in general by task type and further refines them based on methodological theories. Subsequently, it summarizes the core concepts of each theory and reviews related studies in fault diagnosis. Finally, promising future research directions are discussed.

distribution shift, domain adaptation, domain generalization, mechanical fault diagnosis, unsupervised transfer learning

doi: [10.1360/SST-2024-0149](https://doi.org/10.1360/SST-2024-0149)