

文章编号: 1006-3080(2025)05-0693-11

DOI: 10.14135/j.cnki.1006-3080.20241216001

基于特征加权的化工过程中未见模式的故障诊断

易 璩, 侍洪波, 宋 冰, 陶 阳

(华东理工大学信息科学与工程学院, 上海 200237)

摘要:故障诊断是化工行业确保生产安全 and 产品质量的关键。现有的基于深度神经网络的故障诊断模型利用特定条件下的样本训练, 忽视了其领域泛化能力, 导致在复杂多变的场景中诊断性能出现明显下降。针对这一问题, 提出了一种基于加权的领域特定特征去除网络 (Weighted-Based Domain-Specific Feature Removal Network, WBDSFRN)。WBDSFRN 包含一个基于加权的领域特定特征去除模块, 在训练阶段区分领域不变特征和领域特定特征; 在测试阶段尽量将目标域特定特征去除, 从而减轻领域偏移的影响。最后, 利用田纳西-伊士曼工艺 (Tennessee-Eastman Process, TEP) 进行实验。结果表明, WBDSFRN 的性能优于现有方法, 在复杂多变的操作条件下也能表现出稳健的诊断性能。

关键词:故障诊断; 深度神经网络; 领域泛化; 领域偏移; 田纳西-伊士曼工艺

中图分类号: TP277

文献标志码: A

化学事故对生命安全和财产安全构成严重威胁^[1], 因此必须进行有效的过程监控, 故障检测与诊断是确保产品质量和生产安全的关键^[2-4]。然而, 随着现代工业规模和复杂性的不断增加, 暴露了传统故障检测与诊断方法的局限性, 引发了重大的安全问题^[5]。

工业自动化的兴起使得收集大量数据变得简单, 推动了深度神经网络 (Deep Neural Network, DNN) 在故障检测与诊断中的应用^[6-7]。DNN 能有效地从数据中提取与故障相关的特征, 从而获得较高的诊断准确率^[8]。孙俊静等^[9]使用残差模块和空洞卷积提高模型的特征提取能力, 并利用注意力机制提取关键特征, 进一步提高模型的性能。Song等^[10]将时间序列数据转换成矩阵图的形式, 并与多尺度卷积神经网络相结合, 改善了模型的诊断性能。Xu等^[11]利用图卷积网络捕获数据中的关联信息来完成故障诊断任务, 并提出了一种图优化策略来缓解过度平滑问题。邵浙梁等^[12]在卷积神经网络的基础上融合了一种改进的注意力机制, 并且提出了一

种基于注意力机制的软阈值激活函数, 提升了模型的诊断准确率和收敛性能。Zhang等^[13]提出了一种基于 transformer 模型的故障诊断方法, 并融合了图注意力神经网络使其能够处理结构化数据。然而, 这些模型通常假定训练数据和测试数据的分布完全相同^[14]。事实上, 由于运行条件不同, 这种情况很少发生, 从而限制了其适用性^[15]。为了解决这个问题, 人们越来越多地探索领域适应技术, 以便将诊断知识从源域迁移到目标域^[16-18]。Wang等^[19]提出了一种深度多对抗条件域自适应网络, 使用多个域鉴别器加权的方式将源域和目标域中的每个类别逐一配对, 提高了模型的迁移性能。Feng等^[20]将故障的属性知识进行迁移, 实现了目标域上的零样本故障诊断任务。Cao等^[21]提出了一种两阶段对齐方法, 该方法进一步缩小了源域和目标域之间的差异。Li等^[22]通过对每种故障类别设置阈值边界设计了一种更加通用的域自适应网络, 不仅能诊断已知故障类别, 还能将已知故障类别和未知故障类别进行区分。Chai等^[23]提出了多源域细致化的对抗适应策略和互

收稿日期: 2024-12-16

基金项目: 国家自然科学基金 (62073141, 62073140, 62103149)

作者简介: 易 璩 (2000—), 男, 江西九江人, 硕士生, 主要研究方向为故障诊断。E-mail: y30220959@mail.ecust.edu.cn

通信联系人: 侍洪波, E-mail: hbshi@ecust.edu.cn

引用本文: 易 璩, 侍洪波, 宋 冰, 等. 基于特征加权的化工过程中未见模式的故障诊断 [J]. 华东理工大学学报 (自然科学版), 2025, 51(5): 693-703.

Citation: YI Li, SHI Hongbo, SONG Bing, *et al.* Fault Diagnosis of Unseen Modes in Chemical Processes Based on Feature Weighted[J]. Journal of East China University of Science and Technology, 2025, 51(5): 693-703.

补分类器,解决了源域与源域之间和源域与目标域之间类别不一致的问题,放宽了模型的使用条件。Montesuma 等^[24]使用字典学习将多源域知识迁移至目标域上,有效地减轻了数据分布偏移。虽然这些方法在某些跨域故障诊断任务中取得了成功,但在复杂多变的场景中却面临挑战。因为目标域的数据通常无法获取,从而限制了其泛化能力^[25-26]。此外,在面对不断变化的目标域时,性能也会下降^[27],因此需要更通用的诊断模型。

领域泛化方法旨在提取多个源域的不变特征,并将其迁移到未见的目标域中,而无需在训练过程中使用目标域数据。最近的研究重点是利用领域泛化方法开发故障诊断模型^[28]。域对抗神经网络(Domain-Adversarial Neural Network, DANN)通常用于域适应任务,而由于其在跨域任务中表现出很好的性能,越来越多的域泛化任务也开始使用 DANN^[29]。Li 等^[30]利用域增强技术使源域多样化,并将 DANN 与度量学习相结合,以提高诊断性能。Han 等^[31]在 DANN 中引入了三重损失,以增强域内特征的可辨别性,同时保持跨域不变性。然而,现有的大多数方法仅挖掘源域之间的域不变特征而忽略了目标域的特异性。因此,即使对齐了源域的特征,但由于源域和目标域分布不一致,导致提取到的目标域特征仍存在一定程度的分布偏移,不利于模型完成域泛化任务。针对这一问题,提出了基于加权的领域特定特征去除网络(Weighted-Based Domain-Specific Feature Removal Network, WBDSFRN)用于未见工况下的故障诊断。WBDSFRN 包含一个由特定特征提取分支和加权机制组成的特定特征去除模块,用于过滤源域中的特定特征,增强不变特征的提取。同时,它也适用于未见的目标域,以减少目标域中特定特征的负面影响。

本文贡献主要如下:(1)设计了域特定特征提取分支,为模型学习领域无关的特征提供指导,这些特征可迁移至目标域。(2)引入了加权策略来补充域特定特征提取分支,有助于去除目标域的特定特征。这种方法可以减轻域特定特征的影响,从而增强模型的泛化能力。(3)基于田纳西-伊士曼数据集进行实验研究,验证了它在故障诊断任务中可靠的泛化能力。

1 研究方法

1.1 问题定义

将 $D^s = \{D_k^s | k = 1, 2, \dots, K\}$ 表示为多个源域, K

代表源域个数。每个源域 ($D_k^s = \{(x_{k,i}^s, y_{k,i}^s, d_{k,i}^s) | i = 1, 2, \dots, N_k^s\}$) 中包含 N_k^s 个样本,源域中 $x_{k,i}^s \in R^{n_{in}}$ 代表从不同的操作模式采集到的输入数据, n_{in} 代表输入数据的维度; $y_{k,i}^s \in R^{n_c}$ 表示样本 $x_{k,i}^s$ 对应的类别标签, n_c 代表控制系统健康状态的种类数; $d_{k,i}^s \in R^K$ 代表域标签。类似地, $D^T = \{(x_j^T, y_j^T) | j = 1, 2, \dots, N^T\}$ 表示未见的目标域, N^T 表示目标域中的样本数量。由于各源域数据是从不同的操作模式中采集得到的,因此它们之间存在分布差异。假设每个源域 D_k^s 分别服从 P_{XY}^{k1} 分布,则有 $P_{XY}^{k1} \neq P_{XY}^{k2}, 1 \leq k1 \neq k2 \leq K$ 。域泛化的目标是利用多源域数据学习一个预测函数 $f: X \rightarrow Y$,使其能够很好地完成未见目标域中的任务,目标域服从 P_{XY}^T 分布且有 $P_{XY}^T \neq P_{XY}^k, k = 1, 2, \dots, K$ 。

1.2 总体框架

WBDSFRN 总体框架如图 1 所示。WBDSFRN 将网络主要分为特征提取和领域特定特征去除两个阶段。在特征提取阶段,特征编码器 E_C 和 E_D 从处理好的时间窗数据中提取特征以供后续处理,其中 E_C 和 E_D 的结构相同,都是由 3 个因果卷积模块组成,具体结构如图 2 所示。在领域特定特征去除阶段,域特定特征编码器 E_p^k ($k = 1, 2, \dots, K$) 配合域分类器 F_D^1 提取领域特定特征;域分类器 F_D^2 对特征进行分类得到样本隶属于每个域的概率,把这个概率视为权重与 E_p^k 提取到的特征进行融合得到最终的域特定特征;再用 E_C 提取到的特征去除融合后的域特定特征,最后用故障分类器 F_C 进行分类得到最终的结果。网络中所用到的模块的具体结构和参数如表 1 所示。

1.3 设计方法

在迁移学习中,从各个领域中提取到的特征 F 通常可以分解为域不变特征和域特定特征:

$$F = C + P \quad (1)$$

其中, C 表示域不变特征, P 表示域特定特征。域泛化方法之所以能够将诊断知识从源域推广到未见的目标域,是因为目标域和源域中都包含同样的域不变特征。现有的大多数域泛化方法通常在训练阶段通过对齐不同源域的方式来提取域不变特征,并期望在推理阶段能直接提取目标域的域不变特征,从而完成在目标域上的故障诊断任务。然而,由于目标域和源域之间存在分布差异,用源域训练得到的模型提取到的目标域特征并非只有域不变特征,往往包含域特定特征,这使得模型并不能很好地完成目标域上的诊断任务,如图 3 所示。

因此,减小目标域特定特征的影响有利于模型

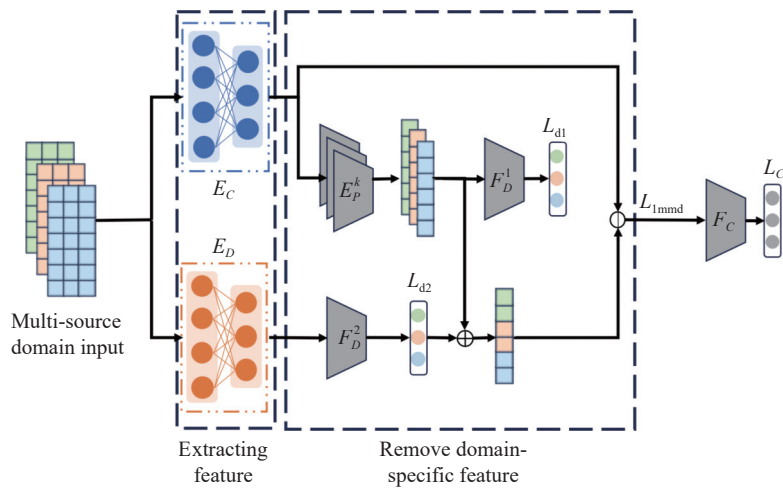


图 1 WBDSFRN 的总体框架

Fig. 1 Framework of WBDSFRN

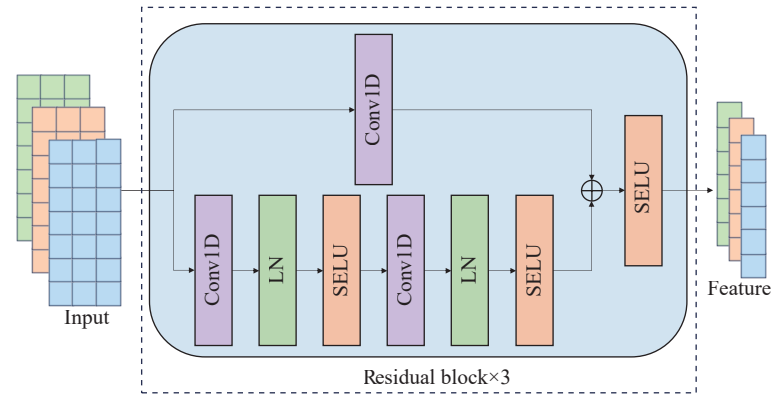


图 2 E_C 和 E_D 的结构

Fig. 2 Structure of E_C and E_D

表 1 WBDSFRN 的网络结构和参数

Table 1 Network architecture and parameters of the WBDSFRN

Module	Structure	Parameter
E_C/ E_D	BLOCK1	Kernel size 3×1 , stride 1, padding 2, dilation 1
	BLOCK2	Kernel size 5×1 , stride 1, padding 4, dilation 1
	BLOCK3	Kernel size 7×1 , stride 1, padding 6, dilation 1
E_P^k	Linear-LN-SELU	Linear layer weights 100×100
	Linear-LN-SELU	Linear layer weights 100×100
F_D^1/ F_D^2	Linear-LN-SELU	Linear layer weights 100×32
	Linear-LN-SELU	Linear layer weights 32×16
	Linear	Linear layer weights $16 \times \text{Domain number}$
F_C	Linear	Linear layer weights $100 \times \text{Class number}$

从源域推广到目标域。根据式 (1), 有:

$$C = F - P \tag{2}$$

然而, 由于目标域的不可见性, 无法直接提取到目标域的特定特征并去除。注意到源域的域特定特征可以通过提取来得到, 因此, 利用目标域与源域的

相似性可以近似得到目标域的域特定特征。具体地, 在训练阶段, 利用特征编码器 E_C 提取源域的特征 $f^S = E_C(x_t^S)$, 其包括域不变特征和域特定特征, 然后通过特定特征编码器 E_P^k 提取源域的域特定特征 $f_p^{S,k} = E_P^k(f^S)$, E_P^k 的训练由域分类器 F_D^1 完成:

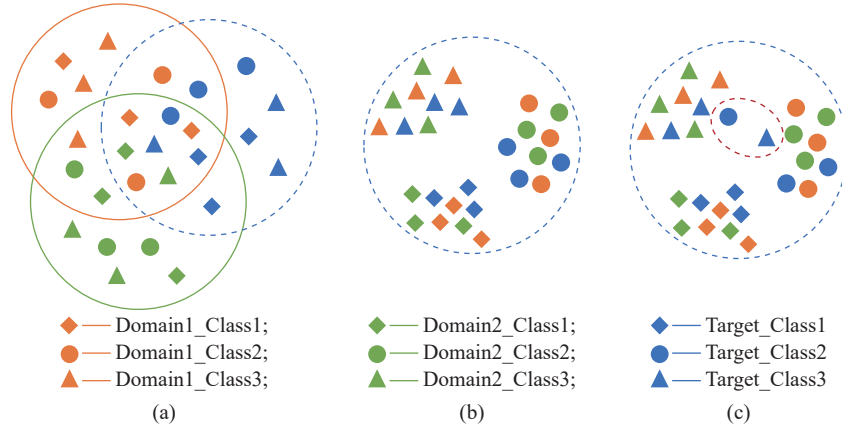


图 3 (a) 源域和目标域存在分布差异; (b) 理想情况下目标域和源域完全一致; (c) 实际上目标域仍存在域偏移, 导致误分类

Fig. 3 (a) Sources and target domains with distributional differences; (b) Target and source domains are perfectly aligned ideally; (c) There is still a domain shift in the target domain, leading to misclassification in reality

$$L_{d1} = \frac{1}{K} \sum_{k=1}^K \frac{1}{N_k^S} \sum_{i=1}^{N_k^S} L_{CE}(F_D^1(E_P^k(x_{k,i}^S)), d_{k,i}^S) \quad (3)$$

其中, K 表示源域数量, N_k^S 表示源域 k 的样本数量, $d_{k,i}^S$ 表示样本的域标签, L_{CE} 表示交叉熵损失。提取到源域的域特定特征后, 根据式 (1) 有:

$$f_{com}^S = f^S - \sum_{k=1}^K w_k^S f_p^{S,k} \quad (4)$$

其中, f_{com}^S 表示域不变特征。当域标签 d_k^S 等于 k 时, $w_k^S = 1$, 否则 $w_k^S = 0$ 。

为了使网络提取域不变特征, 对去除域特定特征之后的特征进行分布对齐。最大均值差异 (Maximum Mean Discrepancy, MMD) 方法能够估计两个分布之间的非参数距离, 因此已被广泛地应用在域适应任务中以减小源域和目标域之间的分布差异。然而, 基于 MMD 的方法关注的是两个域之间的全局分布, 并未考虑到两个域中同一类别之间的分布差异, 这意味着当两个域的全局分布对齐之后, 可能发生类分布未对齐的现象, 例如域 1 中的类别 1 可能会与域 2 中的类别 2 对齐。因此, Zhu 等^[32] 提出了局部最大均值差异 (Local Maximum Mean Discrepancy, LMMD) 方法, 将其应用在域泛化任务中:

$$L_{lmmd} = \sum_{k_1=1}^{K-1} \sum_{k_2=2}^K \frac{1}{N_C} \sum_{c=1}^{N_C} \left\| \sum_{x_i^S \in D_{k_1}^S} w_i^{sc} \phi(x_i^S) - \sum_{x_j^S \in D_{k_2}^S} w_j^{sc} \phi(x_j^S) \right\|_H^2 \quad (5)$$

其中, H 表示具有特征核 k 的再生核希尔伯特空间 (Reproducing Kernel Hilbert Space, RKHS), $\phi(\cdot)$ 表示能够将原始样本映射到 RKHS 的一个特征映射函数, N_C 表示类别数, X_i^S 和 X_j^S 表示源域样本, w_i^{sc} 和 w_j^{sc} 则表示属于类别 c 的样本的权重, 具体地, w_i^{sc} 的计算式为:

$$w_i^{sc} = \frac{y_{ic}}{\sum_{(x_j, y_j) \in D} y_{jc}} \quad (6)$$

其中, y_{ic} 是标签 y_i 的第 c 个元素, 即样本的真实类别。对源域而言, y_i 就是样本的真实标签 y_i^S 。此外, 由于无法直接计算 $\phi(\cdot)$, 需要使用核函数 $k(x_i^S, x_j^S) = \langle \phi(x_i^S), \phi(x_j^S) \rangle$, 其中 $\langle \cdot, \cdot \rangle$ 表示向量之间的内积, 公式 (5) 重新表示为:

$$L_{lmmd} = \sum_{k_1=1}^{K-1} \sum_{k_2=2}^K \frac{1}{N_C} \sum_{c=1}^{N_C} \left[\sum_{i=1}^{N_{k_1}^S} \sum_{j=1}^{N_{k_2}^S} w_i^{sc} w_j^{sc} k(z_i^S, z_j^S) + \sum_{i=1}^{N_{k_2}^S} \sum_{j=1}^{N_{k_1}^S} w_i^{sc} w_j^{sc} k(z_i^S, z_j^S) - 2 \sum_{i=1}^{N_{k_1}^S} \sum_{j=1}^{N_{k_2}^S} w_i^{sc} w_j^{sc} k(z_i^S, z_j^S) \right] \quad (7)$$

其中, z^S 表示中间层特征。将式 (7) 作为损失能够使不同源域按类别对齐。

在测试阶段, 将目标域样本输入到训练好的模型中可以得到 $f^T = E_C(x_i^T)$, $f_p^{T,k} = E_P^k(f^T)$, 类似地, 对于目标域的样本同样有:

$$f_{com}^T = f^T - \sum_{k=1}^K w_k^T f_p^{T,k} \quad (8)$$

$$w^T = F_D^2(E_D(x^T)) \quad (9)$$

其中, $w^T = [w_1^T, w_2^T, \dots, w_K^T]$ 并且 $\sum_{k=1}^K w_k^T = 1$ 。域相似度 w_k^T 表示目标样本 x^T 属于第 k 个源域的概率。通常, 较大的 w_k^T 表示目标域的样本 x^T 与第 k 个源域更相似。域特征编码器 E_D 和域分类器 F_D^2 由源域训练得到, 具体地, 域分类损失可以表示为:

$$L_{d2} = \frac{1}{K} \sum_{k=1}^K \frac{1}{N_k^S} \sum_{i=1}^{N_k^S} L_{CE}(F_D^2(E_P^k(x_{k,i}^S)), d_{k,i}^S) \quad (10)$$

通过式 (8) 和式 (9) 可以看出, 对于不同的目标

域有不同的 $f_p^{T,k}$ 和 w^T , 这使得目标域的特定特征能够随目标域变化而变化。这意味着网络始终能够根据不同的目标域得到几乎相同的 f_{com}^T , 并且有 $f_{\text{com}}^T \approx f_{\text{com}}^S$, 这是目标域样本能够利用源域样本训练的模型进行分类的关键。在提取到域不变特征 f_{com}^S 后, 利用交叉熵损失训练故障分类器 F_C , 因此, 分类损失 L_c 可以表示为:

$$L_c = \frac{1}{K} \sum_{k=1}^K \frac{1}{N_k^S} \sum_{i=1}^{N_k^S} L_{\text{CE}}(F_C(f_{\text{com}}^S), y_{k,i}^S) \quad (11)$$

其中, $y_{k,i}^S$ 表示样本的类别标签。最终, 所提方法的总体损失如式(12)所示:

$$L_{\text{obj}} = L_c + \alpha L_{d1} + \beta L_{\text{lmmd}} + L_{d2} \quad (12)$$

其中, α 和 β 是两个超参数, 用来平衡损失函数。使用 Adam 优化器对网络参数按如下方式进行更新:

$$\theta_{E_c}^{t+1} \leftarrow \theta_{E_c}^t - \mu \left(\frac{\partial L_c}{\partial \theta_{E_c}^t} + \alpha \frac{\partial L_{d1}}{\partial \theta_{E_c}^t} + \beta \frac{\partial L_{\text{lmmd}}}{\partial \theta_{E_c}^t} \right) \quad (13)$$

$$\theta_{E_p^k}^{t+1} \leftarrow \theta_{E_p^k}^t - \mu \left(\frac{\partial L_c}{\partial \theta_{E_p^k}^t} + \alpha \frac{\partial L_{d1}}{\partial \theta_{E_p^k}^t} + \beta \frac{\partial L_{\text{lmmd}}}{\partial \theta_{E_p^k}^t} \right), k = 1, 2, \dots, K \quad (14)$$

$$\theta_{F_D^1}^{t+1} \leftarrow \theta_{F_D^1}^t - \mu \left(\frac{\partial L_{d1}}{\partial \theta_{F_D^1}^t} \right) \quad (15)$$

$$\theta_{F_c}^{t+1} \leftarrow \theta_{F_c}^t - \mu \left(\frac{\partial L_c}{\partial \theta_{F_c}^t} + \alpha \frac{\partial L_{d1}}{\partial \theta_{F_c}^t} + \beta \frac{\partial L_{\text{lmmd}}}{\partial \theta_{F_c}^t} \right) \quad (16)$$

$$\theta_{E_D}^{t+1} \leftarrow \theta_{E_D}^t - \mu \left(\frac{\partial L_{d2}}{\partial \theta_{E_D}^t} \right) \quad (17)$$

$$\theta_{F_D^2}^{t+1} \leftarrow \theta_{F_D^2}^t - \mu \left(\frac{\partial L_{d2}}{\partial \theta_{F_D^2}^t} \right) \quad (18)$$

其中, θ 表示对应的结构参数, μ 表示学习速率, t 表示第 t 次迭代更新。所提方法的工作流程见表 2。

2 结果与讨论

2.1 田纳西-伊士曼工艺

田纳西-伊士曼工艺 (TEP)^[33] 是工厂级化学工艺的经典范例, 已被广泛应用于故障检测和诊断的相关研究中。该过程由 5 个主要单元组成: 反应器、汽提塔、冷凝器、再循环压缩机和分离器, 涉及 4 种气态反应物 A、C、D、E, 惰性物质 B, 副产品 F, 以及两种产品 G 和 H。根据产品 G 和 H 的质量比, TEP 设计了 6 种运行模式, 每种运行模式都有一个相应的设定值, 这导致了在不同运行模式下采集的数据存在差异。

Reinartz 等^[34] 对 TEP 进行了更深入的研究, 此后的相关研究大多以此为基础展开。本实验使用的 TEP 数据集来自 Reinartz 等提供的网站 (https://data.dtu.dk/articles/dataset/Tennessee_Eastman_Reference_Data_for_Fault-Detection_and_Decision_Support_Systems/)

表 2 WBDSFRN 工作流程

Table 2 Workflow of WBDSFRN

Algorithm 1: WBDSFRN

Training stage

Input: Multiple source dataset $D_k^S, k = 1, 2, \dots, K$, K is number of source domains

1. for epoch = 1 to epochs do

2. Randomly sample source data from $\{D_k^S\}_{k=1}^K$

3. Forward propagation to calculate Eq. (12)

4. Backward propagation to update $E_c, E_p^k, F_D^1, F_c, E_D, F_D^2$ by Eq. (13), (14), (15), (16), (17), (18)

5. end for

Return: The optimal $E_c, E_p^k, F_D^1, F_c, E_D, F_D^2$

Testing stage

Input: Unseen target dataset D^T

Model: WBDSFRN with optimal $E_c, E_p^k, F_D^1, F_c, E_D, F_D^2$

Process: Calculate the similarity w^T by Eq. (9) and obtain the domain-invariant features of the target domain by

Eq. (8), and then F_c is used to classify

Output: Final diagnosis decisions

13385936)。该数据集在每种运行模式下对每种故障模拟了 500 次,每次模拟持续时间为 100 h,采样率为每间隔 3 min 采样 1 次。在模拟的前 30 h,系统处于正常状态,而在后 70 h,系统处于故障状态。为确保类别平衡,从 1400 个故障样本中抽取 600 个样本,并使用大小为 15 的滑动窗口对样本进行处理。同时,文献 [34] 设计的故障中选取了 4 种常见的故障类型进行实验,包括阶跃故障(故障 1、2、4、5、6 和 7)、随机变量故障(故障 8、10、11 和 12)、慢漂移故障(故障 13)和黏滞故障(故障 14)。因此,本文用于实验的数据集由 6 种运行模式组成,每种运行模式包含 1 个正常状态和 12 个故障状态,每种模式共有 7 618 个样本。

2.2 实验比较

为了验证所提方法的有效性,使用了 2 种经典方法和 2 种更先进的方法,包括 DCORAL (Deep Correlation Alignment)^[35]、DANN^[29]、CCDG (Conditional Contrastive Domain Generalization)^[36]、CCN (Causal Consistency Network)^[37] 和 CDDG (Causal Disentanglement Domain Generalization)^[38]。为了确保公平性,本文方法和对比方法使用了相同的骨干网络。

TEP 共有 6 种不同的工作模式,分别命名为模式 1~模式 6。使用 t-SNE (t-Distributed Stochastic Neighbor Embedding)^[39] 对这 6 种模式的正常状态数据进行可视化处理,以展示这些模式之间的关系。如图 4 所示。在这 6 种模式中,模式 1、2 和 5 的分布相对接近,模式 3 和 6 的分布相对接近,而模式 4 则相对分散。基于这种关系,设计了以下几种不同的域泛化故障诊断任务,以充分验证所提方法的有效性:

任务 A: 模式 2、4、6 - 模式 1、3、5。在这项任

务中,源域是模式 2、4 和 6,目标域是模式 1、3 和 5,源域的分布包含目标域的所有分布,因此任务相对简单。

任务 B: 模式 1、2、3 - 模式 4、5、6。在这项任务中,模式 1、2、3 作为源域,模式 4、5、6 作为目标域。这种设置代表一种普遍情况,即某些模式如 5 和 6 的目标域和源域的分布较为相似,而模式 4 的分布与源域不同。

任务 C: 模式 1、2、5 - 模式 3、4、6。在这项任务中,模式 1、2 和 5 被用作源域,模式 3、4 和 6 被用作目标域。这种设置使源域和目标域之间的分布差异相对较大,用于模拟最坏情况。

任务 D: 模式 3、4、6 - 模式 1、2、5。该任务与任务 C 的设置相反。模式 3、4 和 6 被视为源域,模式 1、2 和 5 被视为目标域,也用于模拟最坏情况。

WBDSFRN 使用 Adam 优化器,所有网络层的学习率均设为 5×10^{-4} 。数据集的批次大小为 64,迭代次数为 20,滑动窗口设置为 15。式 (12) 中的参数 α 和 β 都设为 0.1。实验结果如表 3~表 6 以及图 5 所示。

对于任务 A(表 3),由于它模拟的是一种比较简单的情況,即所有目标域都能在训练数据中找到相

表 3 WBDSFRN 和其他方法在任务 A 上的故障诊断精度
Table 3 Fault diagnosis accuracy of WBDSFRN and other methods on task A

Model	Fault diagnosis accuracy/%			
	Target 1	Target 3	Target 5	Average
DCORAL	92.10	94.42	93.10	93.21
DANN	92.87	94.28	92.87	93.34
CCDG	92.82	93.96	89.59	92.12
CCN	93.74	94.97	93.62	94.02
CDDG	91.87	92.91	92.60	92.46
Ours	93.13	93.74	92.32	93.06

表 4 WBDSFRN 和其他方法在任务 B 上的故障诊断精度
Table 4 Fault diagnosis accuracy of WBDSFRN and other methods on task B

Model	Fault diagnosis accuracy/%			
	Target 4	Target 5	Target 6	Average
DCORAL	81.60	90.29	94.43	88.77
DANN	78.34	93.78	94.93	89.02
CCDG	73.63	93.31	95.46	87.47
CCN	80.00	94.57	95.25	89.94
CDDG	64.81	92.60	95.29	84.23
Ours	87.08	95.27	95.20	92.52

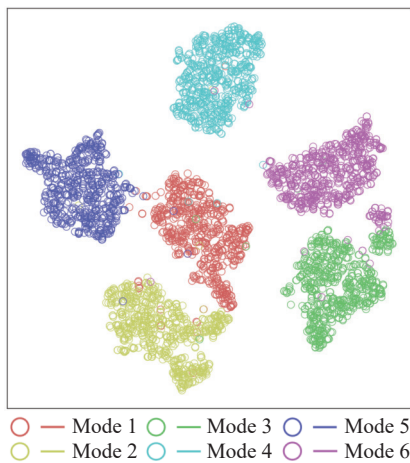


图 4 6 种模式的正常状态的 t-SNE 图

Fig. 4 Six modes of normal state of t-SNE embedding

表 5 WBDSFRN 和其他方法在任务 C 上的故障诊断精度

Table 5 Fault diagnosis accuracy of WBDSFRN and other methods on task C

Model	Fault diagnosis accuracy/%			
	Target 3	Target 4	Target 6	Average
DCORAL	75.91	66.07	72.51	71.50
DANN	75.77	60.62	71.28	69.22
CCDG	72.30	60.44	70.62	72.01
CCN	74.05	55.60	68.38	66.01
CDDG	74.42	50.39	71.88	65.56
Ours	77.33	78.28	75.27	76.96

似的源域,因此,虽然所提方法能够在训练阶段提取域不变特征,但在测试阶段,由于目标域和源域之间的分布差异较小,这削弱了特定特征去除模块的作用。由于此时模型可能将目标域视为某个源域,难以体现所提方法的优势,因此相对于其他方法,性能并没有提高。

对于任务 B(表 4),模型在目标域 5 上的提升非常小,而在目标域 6 上几乎没有改进,甚至有一点性能下降,这可能与任务 A 中的情况类似,目标域 5 与源域 1 和 2 比较相似,目标域 6 与源域 3 比较相似;而对于目标域 4,提出的方法有非常大的性能改进,本文方法的故障诊断精度比 DCORAL、DANN、CCDG、CCN、CDDG 分别高出 5.48%、8.74%、13.45%、7.08% 和 22.7%。这表明在测试阶段,域特定特征去除模块在很大程度上降低了目标域 4 中特定特征的影响,从而使其与源域很好地保持一致。

与前两个任务不同的是,任务 C 模拟的是所有目标域与源域分布差异相对较大的情况(表 5),这导致包括本文方法在内的所有模型都没有表现出与任务 A 和任务 B 相同的性能。但相对而言,本文方法在目标域 3、4 和 6 上的平均准确率比 DCORAL、DANN、CCDG、CCN、CDDG 分别高出 5.46%、7.74%、4.95%、10.95% 和 11.4%。这是因为在测试阶段,特征去除模块在一定程度上去除了目标域的特定特征,使处理后的目标域特征更接近源域特征,从而提高了模型的性能。

与任务 C 一样,训练数据中没有与目标域非常相似的源域,但在任务 D 中,源域的数据分布更加多样化(表 6)。在任务 C 中,源域由模式 1、2、5 组成,如前所述,这三个域的分布较为相似,而在任务 D 中,源域由模式 3 分别和模式 6 构成的一种分布以

表 6 WBDSFRN 和其他方法在任务 D 上的故障诊断精度

Table 6 Fault diagnosis accuracy of WBDSFRN and other methods on task D

Model	Fault diagnosis accuracy/%			
	Target 1	Target 2	Target 5	Average
DCORAL	89.94	87.82	84.30	87.35
DANN	84.88	86.51	83.79	85.06
CCDG	86.62	85.76	88.00	86.79
CCN	88.80	88.04	88.59	88.48
CDDG	88.93	89.49	86.13	88.18
Ours	91.44	90.25	92.31	91.33

及与模式 4 构成的另一种分布组成。这表明源域的多样性对模型的性能有很大影响,任务 A 与任务 B 的比较也说明了这一点,任务 A 中的所有方法都优于任务 B 中的方法。

2.3 消融实验

为了验证所提方法各部分的有效性,本节设计了消融实验。本文方法对泛化性能的提升主要来自于特征去除分支和加权策略,因此针对这两个部分设计了相应的模型: M1 模型删除了加权策略,因此 3 个私有特征提取器相应地变成了 1 个; M2 模型删除了特征去除分支和加权策略,即只保留主干网络、分类器和 LMMD 损失。

消融实验结果如表 7 所示。可以看出,当模型

表 7 目标域为模式 4,源域为模式 1、模式 2 和模式 3 时,消融实验的诊断精度

Table 7 Diagnosis accuracy of ablation experiments when mode 4 is target domain and source domains are mode 1, mode 2 and mode 3, respectively

Fault	Diagnosis accuracy/%		
	M1	M2	WBDSFRN
Normal	37.7	4.3	52.9
Fault 1	66.6	31.4	98.6
Fault 2	99.3	99.7	100.0
Fault 4	99.8	99.5	99.1
Fault 5	17.2	0.7	5.6
Fault 6	100.0	100.0	100.0
Fault 7	9.2	74.7	94.5
Fault 8	98.5	98.3	98.6
Fault 10	88.4	72.7	90.6
Fault 11	99.5	100.0	100.0
Fault 12	97.4	100.0	100.0
Fault 13	87.9	75.8	98.3
Fault 14	94.0	93.5	93.7
Average	76.6	73.1	87.1

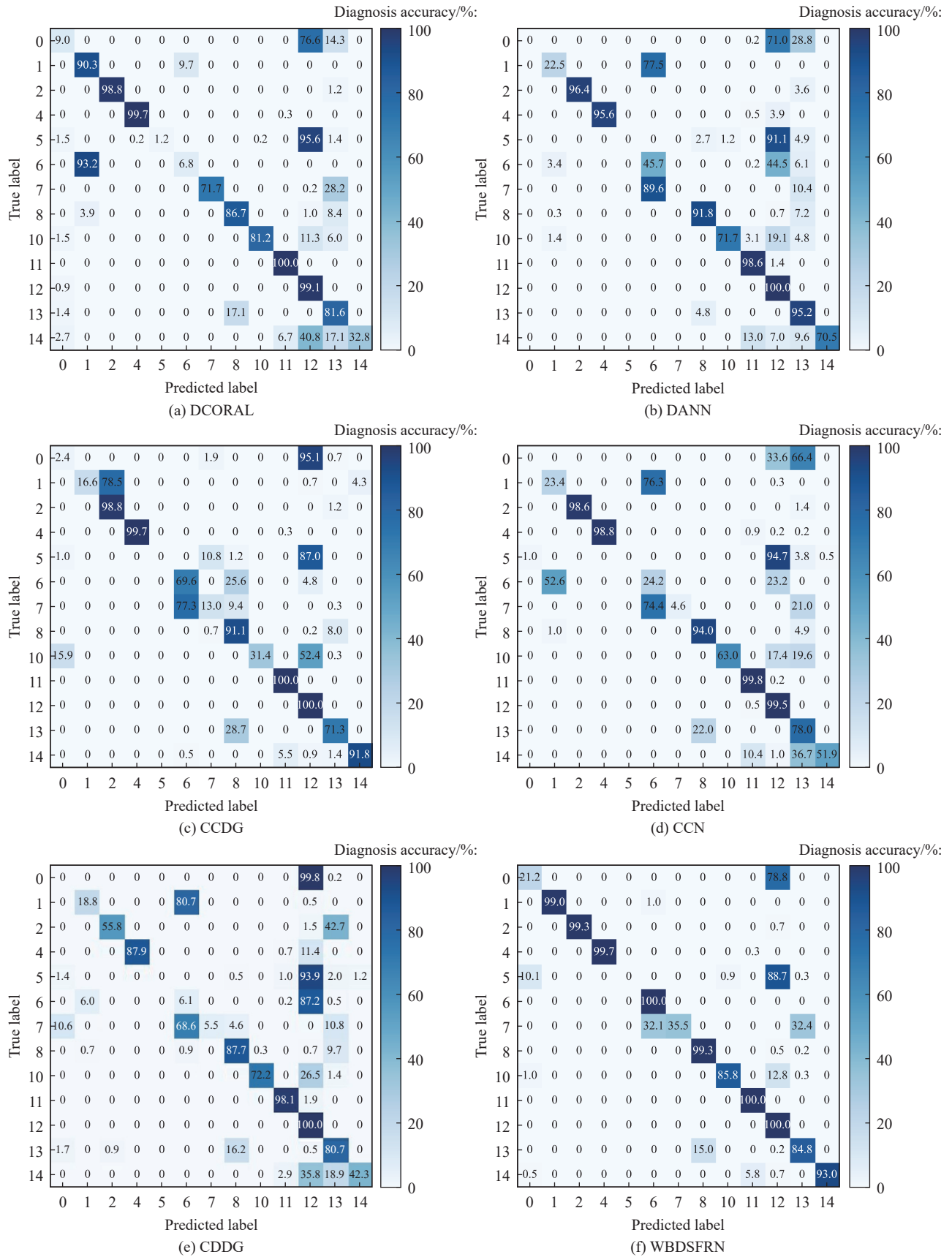


图 5 模式 4 为目标域, 源域为模式 1、模式 2 和模式 5 时, 不同方法的诊断结果

Fig. 5 Diagnosis result of different methods when mode 4 is target domain, and source domains are mode 1, mode 2 and mode 5, respectively

的两个核心部分都被移除时,性能会明显下降。对于 M1,当去掉加权策略时,可能会导致目标域中某些特定领域的特征没有被去掉,甚至可能导致目标域中某些有用的特征被去掉,进而影响模型的泛化性能。对于 M2,所提出的方法退化成为一种域通用故障诊断模型,仅使用 LMMD 来提取域不变特征。如 2.2 节所述,这种模型由于忽略了目标域的特定特征,导致从目标域提取的特征在测试阶段仍然存在很大程度的域偏差,从而降低了模型的诊断性能。

图 6 展示了 WBDSFRN 和两个消融模型在任务 B 上的特征分布,其中不同颜色代表不同的故障类别,不同形状表示不同域,图中 Domain1、Domain2 和 Domain3 分别表示模式 1、模式 2 和模式 3 这三个

源域, Domain4 表示模式 4 即目标域。从图中可以看出,相比于 M1 和 M2 两个模型,WBDSFRN 中的源域和目标域的所有相同故障类别的距离更近,这说明所提方法能提取到更好的域不变特征。另外,从目标域的特征分布上看, WBDSFRN 中目标域的故障类别只有很小一部分混淆,而 M1 和 M2 中均存在多种故障类别相互混淆,这说明利用特征加权的方式动态调整目标域特征分布是可行的。

3 结 论

提出一种新的领域泛化网络 WBDSFRN,能够通过去除领域特定特征来增强模型在故障诊断任务上的泛化能力。具体来说,在训练阶段,特征去除模块去除源域特定特征以获得域不变特征;而在测试阶段,目标域特定特征会被特征去除模块和相似性加权分支机制最大程度地去除,从而提高模型在未见目标域上的泛化能力。对比实验和消融实验的结果表明, WBDSFRN 的设计是有效的,表现出良好的诊断与泛化性能。

参考文献:

- [1] CHEN C, RENIERS G. Chemical industry in China: The current status, safety problems, and pathways for future sustainable development[J]. *Safety Science*, 2020, 128: 104741.
- [2] JI C, SUN W. A review on data-driven process monitoring methods: Characterization and mining of industrial data[J]. *Processes*, 2022, 10(2): 335.
- [3] HAN W X, LONG P, XU B. Fault detection for uncertain polynomial fuzzy systems using H-/L ∞ observer and ellipsoidal analysis[J]. *IEEE Transactions on Fuzzy Systems*, 2022, 31(4): 1113-1123.
- [4] CHEN Y, TONG C D, LAN T. Fault monitoring for chemical processes using neighborhood embedding discriminative analysis[J]. *Process Safety and Environmental Protection*, 2022, 164: 109-118.
- [5] XIAO Y T, SHI H B, WANG B Y, *et al.* Fault diagnosis of unseen modes in chemical processes based on labeling and class progressive adversarial learning[J]. *IEEE Transactions on Instrumentation and Measurement*, 2022, 72: 1-12.
- [6] WU H, ZHAO J S. Fault detection and diagnosis based on transfer learning for multimode chemical processes[J]. *Computers & Chemical Engineering*, 2020, 135: 106731.
- [7] CHANG P, MENG F C. Fault detection of urban wastewater treatment process based on combination of deep inform-

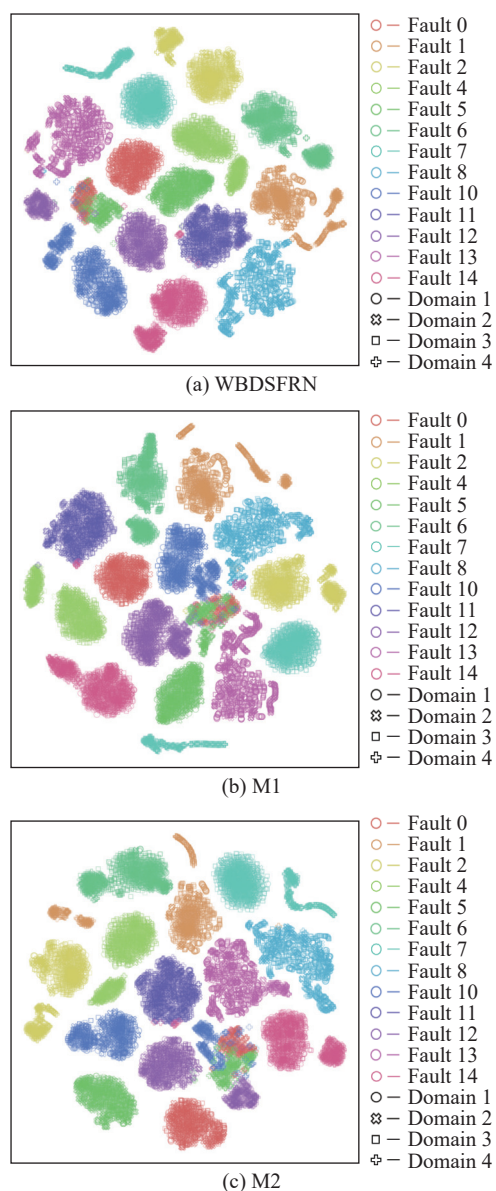


图 6 不同消融策略在任务 B 上的特征分布

Fig. 6 Distribution of features on task B for different ablation strategies

- ation and transformer network[J]. *IEEE Transactions on Neural Networks and Learning Systems*, 2024, 35(6): 8124-8133.
- [8] SHAO H D, LI W, CAI B P, *et al.* Dual-threshold attention-guided gan and limited infrared thermal images for rotating machinery fault diagnosis under speed fluctuation[J]. *IEEE Transactions on Industrial Informatics*, 2023, 19(9): 9933-9942.
- [9] 孙俊静, 顾幸生. 基于注意力机制多尺度卷积神经网络的轴承故障诊断 [J]. *华东理工大学学报 (自然科学版)*, 2024, 50(2): 247-256.
- [10] SONG Q S, JIANG P. A multi-scale convolutional neural network based fault diagnosis model for complex chemical processes[J]. *Process Safety and Environmental Protection*, 2022, 159: 575-584.
- [11] XU J M, KE H B, CHEN Z W, *et al.* OHCA-GCN: A novel graph convolutional network-based fault diagnosis method for complex systems via supervised graph construction and optimization[J]. *Advanced Engineering Informatics*, 2024, 61: 102548.
- [12] 邵浙梁, 戚知宽, 周邵萍. 基于改进注意力机制的 CNN 的齿轮箱故障诊断 [J]. *华东理工大学学报 (自然科学版)*, 2024, 50(6): 920-928.
- [13] ZHANG L, SONG Z H, ZHANG Q H, *et al.* Generalized transformer in fault diagnosis of tennessee eastman process[J]. *Neural Computing and Applications*, 2022, 34(11): 8575-8585.
- [14] WANG J D, LAN C L, LIU C, *et al.* Generalizing to unseen domains: A survey on domain generalization[J]. *IEEE Transactions on Knowledge and Data Engineering*, 2022, 35(8): 8052-8072.
- [15] WANG X G, LIU X Y, LI Y. An incremental model transfer method for complex process fault diagnosis[J]. *IEEE/CAA Journal of Automatica Sinica*, 2019, 6(5): 1268-1280.
- [16] WANG Y L, WU D Z, YUAN X F. LDA-based deep transfer learning for fault diagnosis in industrial chemical processes[J]. *Computers & Chemical Engineering*, 2020, 140: 106964.
- [17] SONG Q Y, JIANG X X, LIU J, *et al.* Contrast-assisted domain-specificity-removal network for semi-supervised generalization fault diagnosis[J]. *IEEE Transactions on Neural Networks and Learning Systems*, 2025, 36(3): 5403-5416.
- [18] 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.
- [19] WANG B S, BARALDI P, ZIO E. Deep multiadversarial conditional domain adaptation networks for fault diagnostics of industrial equipment[J]. *IEEE Transactions on Industrial Informatics*, 2022, 19(8): 8841-8851.
- [20] FENG L J, ZHAO C H. Fault description based attribute transfer for zero-sample industrial fault diagnosis[J]. *IEEE Transactions on Industrial Informatics*, 2021, 17(3): 1852-1862.
- [21] CAO W, MENG Z, SUN D Y, *et al.* A two-stage domain alignment method for multi-source domain fault diagnosis[J]. *Measurement*, 2023, 214: 112818.
- [22] LI J P, Zhang X G, YUE K, *et al.* An auto-regulated universal domain adaptation network for uncertain diagnostic scenarios of rotating machinery[J]. *Expert Systems with Applications*, 2024, 249: 123836.
- [23] CHAI Z, ZHAO C H, HUANG B. Multisource-refined transfer network for industrial fault diagnosis under domain and category inconsistencies[J]. *IEEE Transactions on Cybernetics*, 2022, 52(9): 9784-9796.
- [24] MONTESUMA E, MBOULA F M N, SOULOUMIAC A. Multi-source domain adaptation through dataset dictionary learning in wasserstein space[C]// 26th European Conference on Artificial Intelligence. Kraków, Poland: IOS Press, 2023: 1739-1746.
- [25] WANG H J, BAI X W, WANG S H, *et al.* Generalization on unseen domains via model-agnostic learning for intelligent fault diagnosis[J]. *IEEE Transactions on Instrumentation and Measurement*, 2022, 71: 1-11.
- [26] ZHAO C, SHEN W M. Adversarial mutual information-guided single domain generalization network for intelligent fault diagnosis[J]. *IEEE Transactions on Industrial Informatics*, 2022, 19(3): 2909-2918.
- [27] CHEN L, LI Q, SHEN C Q, *et al.* Adversarial domain-invariant generalization: A generic domain-regressive framework for bearing fault diagnosis under unseen conditions[J]. *IEEE Transactions on Industrial Informatics*, 2021, 18(3): 1790-1800.
- [28] ZHOU K, WANG R, TONG Y F, *et al.* Domain generalization of chemical process fault diagnosis by maximizing domain feature distribution alignment[J]. *Process Safety and Environmental Protection*, 2024, 185: 817-830.
- [29] GANIN Y, LEMPITSKY V. Unsupervised domain adaptation by backpropagation[C]//32th International Conference on Machine Learning. Iile, France: PMLR, 2015, 37: 1180-1189.
- [30] LI X, ZHANG W, MA H, *et al.* Domain generalization in rotating machinery fault diagnostics using deep neural networks[J]. *Neurocomputing*, 2020, 403: 409-420.
- [31] HAN T, LI Y F, QIAN M. A hybrid generalization network for intelligent fault diagnosis of rotating machinery under unseen working conditions[J]. *IEEE Transactions on Instrumentation and Measurement*, 2021, 70: 1-11.
- [32] ZHU Y C, ZHUANG F Z, WANG J D, *et al.* Deep subdo-

- main adaptation network for image classification[J]. IEEE Transactions on Neural Networks and Learning Systems, 2020, 32(4): 1713-1722.
- [33] DOWNS J J, VOGEL E F. A plant-wide industrial process control problem[J]. Computers & Chemical Engineering, 1993, 17(3): 245-255.
- [34] REINARTZ C, KULAHCI M, RAVN O. An extended tennessee eastman simulation dataset for fault-detection and decision support systems[J]. Computers & Chemical Engineering, 2021, 149: 107281.
- [35] SUN B C, SAENKO K. Deep coral: Correlation alignment for deep domain adaptation[C]//Computer Vision-ECCV 2016 Workshops: Amsterdam. Netherlands: Springer International Publishing, 2016, 9915: 443-450.
- [36] RAGAB M, CHEN Z H, ZHANG W Y, *et al.* Conditional contrastive domain generalization for fault diagnosis[J]. IEEE Transactions on Instrumentation and Measurement, 2022, 71: 1-12.
- [37] LI J, Wang Y, ZI Y Y, *et al.* Causal consistency network: A collaborative multimachine generalization method for bearing fault diagnosis[J]. IEEE Transactions on Industrial Informatics, 2022, 19(4): 5915-5924.
- [38] JIA L S, CHOW T W S, YUAN Y X. Causal disentanglement domain generalization for time-series signal fault diagnosis[J]. Neural Networks, 2024, 172: 106099.
- [39] Van der MAATEN L, HINTON G. Visualizing data using t-SNE[J]. Journal of Machine Learning Research, 2008, 9(86): 2579-2605.

Fault Diagnosis of Unseen Modes in Chemical Processes Based on Feature Weighted

YI Li, SHI Hongbo, SONG Bing, TAO Yang

(School of Information Science and Engineering, East China University of Science and Technology,
Shanghai 200237, China)

Abstract: Fault diagnosis is critical for ensuring production safety and product quality within the chemical industry. Existing fault diagnosis models that leverage deep neural networks are typically trained using samples from specific operating conditions, neglecting the domain generalization capabilities of these models. This leads to a significant decline in diagnostic performance in complex and variable scenarios. To address this issue, a weighted-based domain-specific feature removal network (WBDSFRN) is proposed. WBDSFRN incorporates a weighted-based domain-specific feature removal module, which effectively discriminates between domain-invariant and domain-specific features during the training phase. In the testing phase, the model largely eliminates domain-specific features of the target domain to mitigate domain shift effects. Finally, experiments are conducted using the Tennessee-Eastman process (TEP). The results demonstrate that WBDSFRN significantly outperforms existing approaches and exhibits robust diagnostic performance even under complex and variable operating conditions.

Key words: fault diagnosis; deep neural networks; domain generalization; domain shift; Tennessee-Eastman process

(责任编辑: 张欣)