

基于类感知对比学习的半监督故障诊断

金泽中, 叶春明

上海理工大学管理学院, 上海

收稿日期: 2024年1月6日; 录用日期: 2024年3月8日; 发布日期: 2024年3月15日

摘要

当前基于数据驱动的故障诊断方法依赖于标注完备的训练样本, 而在实际工程活动中标注足量故障样本需要耗费大量人力物力。对此提出一种基于类感知对比学习的半监督故障诊断方法以综合应用少量标注样本以及大量无标注样本进行训练, 减少模型训练对于标注样本的需求。首先根据模型最大概率预测值动态赋予无标注样本伪标签以参与模型训练, 并结合置信度筛选以减少伪标签中噪声标签所引起的负面影响, 同时引入一致性正则化, 增强模型对伪标签样本的特征表达能力, 构建更为完备的决策边界。随后设计类感知对比学习模块以确保模型特征空间中的各类故障样本间的类内一致性以及类间对比度, 实现判别能力增强。实验结果证实, 该方法能够在较少标签的条件下取得良好的诊断结果。

关键词

故障诊断, 对比学习, 半监督学习

Fault Diagnosis Based on Class-Aware Contrastive Semi-Supervised Learning

Zezhong Jin, Chunming Ye

Business School, University of Shanghai for Science and Technology, Shanghai

Received: Jan. 6th, 2024; accepted: Mar. 8th, 2024; published: Mar. 15th, 2024

Abstract

The current data-driven fault diagnosis methods rely on well-labeled training dataset, however those datasets in practical engineering activities require lots of resources. In this way, a semi supervised fault diagnosis method based on class-aware contrastive learning is proposed to comprehensively apply a small number of labeled samples and a large number of unlabeled samples for training, reducing the need for labeled samples in model training. Firstly, based on the maximum probability prediction value of the model, unlabeled samples are dynamically assigned pseudo

labels to participate in model training, and confidence screening is combined to reduce the negative impact caused by noisy labels in the pseudo labels. At the same time, consistency regularization is introduced to enhance the model's feature expression ability for pseudo label samples and construct a more complete decision boundary. Subsequently, a class-aware contrastive learning module is designed to ensure intra class consistency and inter class contrast among various fault samples in the feature space, achieving enhanced discriminative ability. The experimental results confirm that this method can maintain great fault diagnosis performance with few labels.

Keywords

Fault Diagnosis, Contrastive Learning, Semi-Supervised Learning

Copyright © 2024 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

轴承作为机械传动的关键零件, 被广泛应用于各种机械设备中, 其健康状况对于机械设备的安全性及稳定性具有重要影响。然而设备在一些恶劣环境下运行时, 轴承将会不可避免地发生退化, 产生裂纹、磨损等。一旦发生故障将直接影响整个设备的正常运行, 轻则给企业造成经济损失, 重则引发事故, 威胁生命安全[1]。因此, 为保证机械设备的正常运行, 对轴承的健康状况进行监测以及及时排除安全隐患具有重大的工程意义。

深度神经网络在工业故障诊断领域得到了广泛的研究, 以预防机械系统的突然故障以及临界损伤。Zhang 等[2]提出一种基于通道 - 空间注意力机制与特征融合的深度残差故障诊断网络, 诊断准确率可达 99.87%。Xu 等[3]针对复杂环境下系统故障诊断中多尺度模型外推效率低的问题, 提出权重软投票的多尺度决策加权融合模型, 该模型能够有效地捕获采集多尺度的时间与频率信息, 并分别在特征层与决策层考虑了不同尺度特征对故障信息的权重, 具有较为良好的泛化能力。然而上述故障诊断方法的性能很大程度上归功于标注完备的训练数据集。

在机械故障发生早期, 其故障特征微弱且状态处于不断发展变化中, 故障模式与故障表征间映射关系模糊, 导致早期故障信号分析和故障界定难度大, 更容易造成机械设备故障样本标记缺乏问题[4]。此外, 人工标注故障信息需投入大量人力物力成本, 在实际过程活动中难以实现。使得出现“大量无标记数据获取易, 大量有标记数据获取难”的问题。而仅采用无标注样本训练深度神经网络, 无法建立起准确可靠的设备故障诊断映射关系。因此, 如何训练深度神经网络充分提取无标注故障样本特征信息, 并结合少量标注样本促成模型构建完备的故障特征信息映射关系, 是当下故障诊断领域一个重要研究课题。

针对上述问题, 基于半监督学习的方法得到广泛研究。不同于监督学习与无监督学习, 半监督学习方法通过结合训练无标注以及少量标注训练样本信息, 利用无标注训练样本探索数据决策边界, 并通过标注样本引入显式监督信息, 构成更为完备的决策边界, 实现对故障信号的判别[5]。Yu 等[6]对无标注训练样本进行多次数据增强并迫使模型输出与原样本相同预测值以引入一致性正则化, 增强模型特征表达能力, 实现在训练样本有限情况下的故障诊断。Liao 等[7]采用对抗学习训练模型特征提取层, 并根据模型输出预测值赋予无标注样本伪标签以参与训练, 实现半监督故障诊断。然而当前半监督学习方法仅关注标签映射空间中监督映射关系联结, 而在特征空间缺乏显式监督, 限制了模型在低标签覆盖率情况

下的诊断精度。

为此, 本文提出一种基于类感知对比学习半监督故障诊断方法以增强模型在标签匮乏情况下的诊断精度。该方法首先采用伪标签以及一致性正则化策略, 根据模型最大概率预测值动态赋予无标注样本伪标签, 并对训练样本进行数据增强以提升模型特征表达能力。同时为充分表征样本特征间语义信息, 设计类感知对比学习模块以充分利用少量标记样本与大量无标记数据的潜在特征分布, 完善决策边界。此外, 为了减少模型伪标签错误引起的负面影响, 通过模型预测置信度量伪标签不确定性, 进而对比学习损失动态加权。在德国帕德博恩大学以及凯斯西储大学轴承数据集中进行实证分析, 结果表明本文所提方法能够在标注数据匮乏的情况下实现良好的诊断精度。

2. 问题描述

半监督学习问题可以定义为假定训练数据集中包含少量标注数据集 $D_l = \left\{ \left(x^{(i)}, y^{(i)} \right) \right\}_{i=1}^l$, 其中 $x^{(i)}$ 为其训练样本而 $y^{(i)} \in \mathcal{Y} = \{1, \dots, C\}$ 为相应的标注标签, C 为其故障类别数; 此外还包含大量无标注样本 $D_u = \left\{ \left(u^{(i)} \right) \right\}_{i=1}^u$ 。本文旨在设计半监督学习算法, 结合 D_l 以及 D_u 获取决策边界, 同时利用无标记样本探索特征空间中的数据分布情况, 引导模型的决策边界进行动态调整, 综合利用 D_l 以及 D_u 促成模型训练, 弥补仅采用 D_l 进行模型训练而引起的故障诊断信息不足问题[8]。

3. 基于类感知对比学习的故障诊断方法描述

半监督故障诊断分类模型由特征提取层 $F(\cdot)$ 以及故障分类层 $P_{cls}(\cdot)$ 组成。其中 $F(\cdot)$ 采用卷积神经网络提取故障信号特征, 对于故障信号样本 x , 特征提取层 $F(\cdot)$ 将其映射至深度特征嵌入空间, 即 $h = F(x)$, 其中 h 表示深度特征嵌入。随后采用由全连接层组成的故障分类层 $P_{cls}(\cdot)$ 将深度特征嵌入 h 映射至状态健康空间, 输出模型分类结果 p , 即 $p = P_{cls}(r)$, 同时为获取模型特征高维映射, 在特征提取层后引入投影层 $P(\cdot)$ 。方法整体框架如图 1 所示。

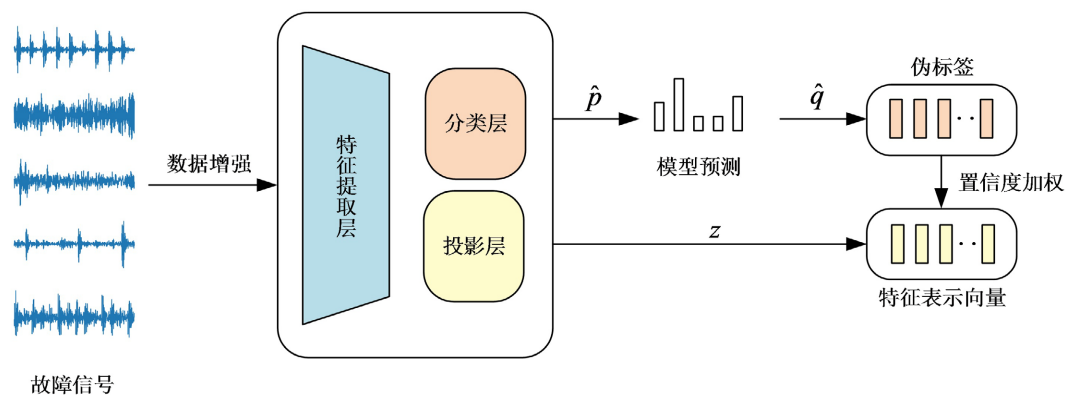


Figure 1. The framework for the proposed method
图 1. 所提方法框架图

3.1. 一致性正则化与伪标签策略

在模型训练过程中, 随机抽取 B 批次大小的标注故障信号样本 $\mathcal{X} = \left\{ \left(x^{(i)}, y^{(i)} \right) \right\}_{i=1}^B, i \in (1, \dots, B)$, $\mathcal{X} \in D_l$, 以及 μB 批次大小的无标注故障信号样本 $\mathcal{U} = \left\{ \left(u^{(i)} \right) \right\}_{i=1}^{\mu B}, i \in (1, \dots, \mu B)$, $\mathcal{U} \in D_u$, 其中 μ 为控制训练过程中

\mathcal{U} 与 \mathcal{X} 相对大小的比例的超参数。

对于标注样本 \mathcal{X} , 本文通过交叉熵损失表征模型预测与其真实标注标签差值, 并将其作为模型优化目标:

$$L_x = \frac{1}{B} \sum_{i=1}^B H(y^{(i)}, p^{(i)}) = -\frac{1}{B} \sum_{i=1}^B y^{(i)} \log(p^{(i)}) \quad (1)$$

而对于未标注样本 \mathcal{U} , 本文参照文献[9], 对于给定无标注样本 $u^{(i)}$, 获取其模型输出预测 $p_u^{(i)} = P_{cls}(F(u^{(i)}))$, 随后获取其伪标签 $\hat{q}^{(i)} = \arg \max(p_u^{(i)})$, 并对其施行以下一致性正则化:

$$L_u = \frac{1}{\mu B} \sum_{i=1}^{\mu B} \mathbb{1}(\max(p_u^{(i)}) \geq t) H(\hat{q}^{(i)}, P_{cls}(F(Aug(u^{(i)})))) \quad (2)$$

其中 t 为阈值常量; $\mathbb{1}$ 为指示函数, 仅当模型预测置信度 $\max(p_u^{(i)})$ 高于阈值 t 时置为 1, 即参与模型损失函数计算, 反之, 则置零; $Aug(\cdot)$ 表示对输入样本进行数据增强, 一致性正则化通过对输入样本施行数据增强并迫使模型输出相同目标分类值, 使其趋向于学习结构化特征, 实现判别能力增强。

3.2. 类感知对比学习模块

对比学习(Contrastive learning, CL)的提出, 主要是为了解决基于监督学习的深度学习依赖于大量标注完备数据的问题[10]。对比学习旨在构建映射函数, 将输入信息映射至超球体空间, 通过构建正负样本对以执行实例判别任务, 拉近正样本对在超球体空间中的映射距离、推远负样本特征距离, 优化特征空间以增强模型特征表达能力。SimCLR [11]作为一种经典对比学习算法, 通过对样本进行组合数据增强, 将同一样本经不同增强方式进行数据增强后的样本作为正对, 并最大化其互信息, 不断减小相似样本度量距离, 实现同类样本特征表示相似, 不同类特征表示互异。

在对比学习训练过程中, 首先随机抽样 N 批次大小的无标注故障信号样本 $\{x^{(i)}\}_{i=1}^N, i \in \{1, 2, \dots, N\}$, $N = \mu B$ 对于批次中每个样本实行两种不同数据增强方法 $Aug_a(\cdot)$ 以及 $Aug_b(\cdot)$ 以获得其相关实例 $\{x_a^{(1)}, \dots, x_a^{(N)}, x_b^{(1)}, \dots, x_b^{(N)}\}$ 。对于其中一个样本 $x_a^{(i)}$, 可以与其余 $2N - 1$ 个样本组成样本对, 其中 $(x_a^{(i)}, x_b^{(i)})$ 为正样本对, 相较于余下的 $2N - 2$ 样本为负样本。随后将其引入特征编码模块 $F(\cdot)$ 提取特征表示, 对于得到的特征向量: $h_a^{(i)} = F(x_a^{(i)})$, 随后参照文献[11]将特征向量 $h_a^{(i)}$ 引入投影层 $P(\cdot)$ 并对其进行 L2 正则化以将其映射至单位超球体空间, 获取其表示向量 $z_a^{(i)} = Norm(P(h_a^{(i)}))$, 其中 $Norm$ 表示 L2 正则化。可通过相似度矩阵 $S \in \mathbb{R}^{2N \times 2N}$ 表征各表示向量的相似度 $s_{ij} = \exp(z^{(i)} \cdot z^{(j)} / \tau)$, 其中 τ 为对比学习温度系数。并定义以下系数矩阵 $W_{con} \in \mathbb{R}^{2N \times 2N}$ 为:

$$w_{ij}^{con} = \begin{cases} 1 & i = j, \\ 1 & z^{(i)} \text{ 和 } z^{(j)} \text{ 为同一样本经不同增强方式得到} \\ 0 & \text{其他} \end{cases} \quad (3)$$

随后在对比学习训练过程, 为了识别整个数据集中的所有正对, 以获得更多实例之间的关系, 采用 InfoNCE 作为损失函数:

$$L_c = H(S, W_{con}) = \frac{1}{2N} \sum_{i=1}^{2N} L_i^{InfoNCE} \quad (4)$$

$$L_i^{InfoNCE} = -\log \frac{\exp(z^{(i)}, z_*^{(i)} / \tau)}{\sum_{t=1}^{2N} \mathbb{1}(t \neq i) \exp(z^{(i)}, z^{(t)} / \tau)} \quad (5)$$

其中 H 表示交叉熵损失函数, $z_*^{(i)}$ 为经过另一数据增强的相关实例正对样本 $z^{(i)}$ 。采用对比学习进行模型

训练虽能有效增强模型特征表示, 却并没有充分利用各样本间语义信息, 并不能充分地聚集同类样本。受文献[12]启发, 本文将模型对无标注样本所输出伪标签 \hat{q} 作为模型训练标签, 将对比系数矩阵 W_{con} 转化为监督对比系数矩阵 $W_{scon} \in \mathbb{R}^{2N \times 2N}$:

$$W_{ij}^{scon} = \begin{cases} 1 & i = j, \\ 1 & z^{(i)} \text{和} z^{(j)} \text{的伪标签为同一类别} \\ 0 & \text{其他} \end{cases} \quad (6)$$

在引入同类正对样本信息后, 能够有效拉近同类健康状态样本在特征空间中的嵌入距离, 增强模型特征判别能力, 实现更为紧凑的特征嵌入表示。同时考虑到在模型训练初期, 尚未完全拟合样本特征分布, 易对无标注样本产生错误伪标签, 使得模型过拟合于错误标签信息, 对此, 本文引入置信度加权策略, 对于监督对比系数矩阵 W_{scon} 动态赋予其标签样本以最大概率预测值 $p^{(i)}$, 减少因伪标签错误引起的负面影响, 得到最终的类感知对比系数矩阵 $W_{clacon} \in \mathbb{R}^{2N \times 2N}$:

$$W_{ij}^{clacon} = \begin{cases} p^{(i)} \cdot p^{(j)} \cdot w_{ij}^{scon} & i \neq j \\ w_{ij}^{scon} & \text{其他} \end{cases} \quad (7)$$

此时的训练目标为最小化表示向量的相似度矩阵 S 以及类感知对比系数矩阵 W_{clacon} 的差值, 其损失函数 L_c 为:

$$L_c = H(S, W_{clacon}) = \sum_{i=1}^{2N} \frac{1}{1 + |P(i)|} L_c^{(i)} \quad (8)$$

其中 $P(i)$ 表示来自同属于一类伪标签的索引。 $|P(i)|$ 表示其具有同种伪标签的样本数量, $|P(i)|+1$ 表示所有正对, $L_c^{(i)}$ 为:

$$L_c^{(i)} = -\log \frac{\exp(z^{(i)}, z_*^{(i)}/\tau)}{\sum_{t=1}^{2N} \mathbb{1}(t \neq i) \exp(z^{(i)}, z^{(t)}/\tau)} - \sum_{p \in P(i)} w_{ip} \cdot \log \frac{\exp(z^{(i)}, z^{(p)}/\tau)}{\sum_{t=1}^{2N} \mathbb{1}(t \neq i) \exp(z^{(i)}, z^{(t)}/\tau)} \quad (9)$$

总体训练目标为最小化以下综合损失:

$$L = L_x + L_u + L_c \quad (10)$$

4. 实验验证

本文在公共数据集上对所提方法进行验证以进一步证明方法有效性。首先对数据集进行系统性介绍, 然后提供模型框架及其相关参数设置, 最后阐述了本文算法与其他方法的比较实验结果, 并开展消融实验以验证所提不同模块有效性。

4.1. 数据集介绍

1) 德国帕德博恩大学(Paderborn University, PU)轴承数据集, 其实验台主要由轴承测试模块、扭矩测量轴以及电机组成。实验轴承为 6203 型球轴, 根据轴承故障位置及损坏尺度可将其划分为包括正常状态、内/外圈严重损伤、内/外圈轻微损伤在内的 5 类状态, 每类状态训练样本数为 800。

2) 凯斯西储大学(Case Western Reserve University, CWRU)轴承数据集, 实验台主要由电机、扭矩传感器以及控制电子设备组成。轴承型号为深沟球轴承 SKF6205, 采样频率为 12 kHz。实验中通过电火花加工技术模拟常见轴承故障类型, 每种故障类型的损伤直径分别为 0.18 mm、0.36 mm、0.54 mm 三种尺寸, 此外包括正常运行状态的轴承振动信号, 共计 10 种健康状态, 每类状态训练样本数为 100。

4.2. 模型参数设置

本文所提故障诊断特征提取网络由一维残差神经网络(PreAct Resnet)构成, 特征提取网络由两个残差块组成, 投影层由全连接层构成, 输出 128 维低维嵌入向量。同时为更好地提取故障信号中的时域信息, 本文采用经 Z-score 归一化的一维振动信号作为模型输入。并选取序列转置、加入高斯噪声、信号放缩、随机信号置零作为增强方式, 数据增强效果如图 2 所示。模型训练批量大小为 32, μ 为 5, 阈值常量 t 为 0.95, 对比学习温度系数 τ 设置为 0.2, 模型训练轮次为 100。

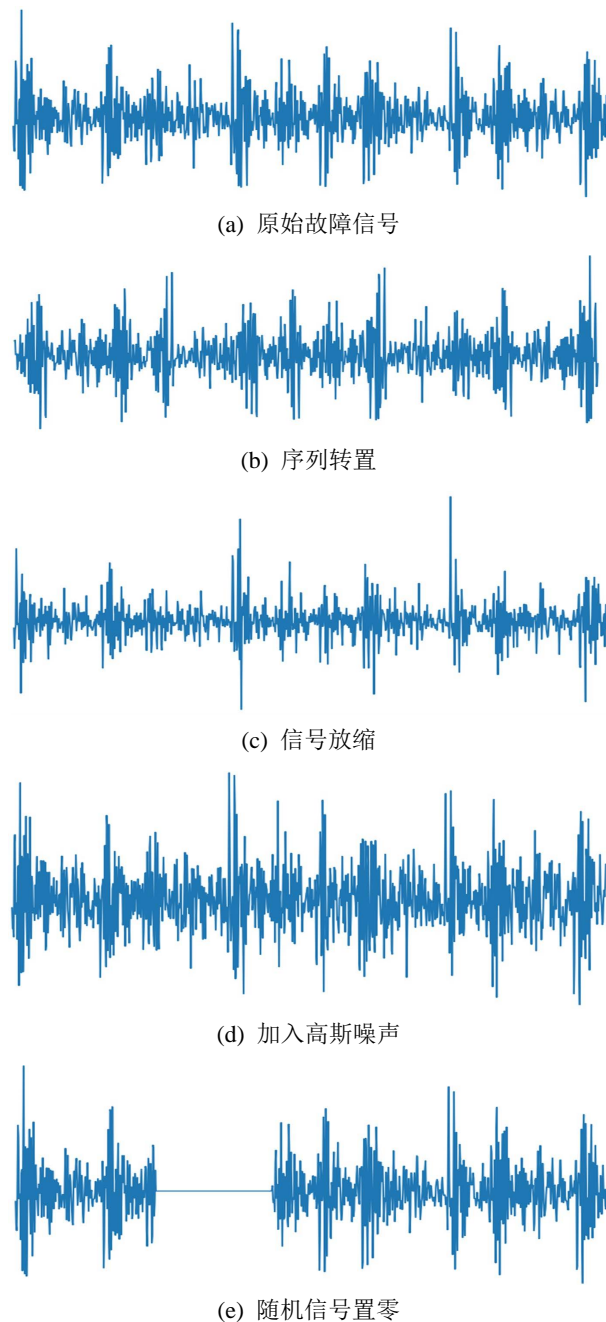


Figure 2. The ways of data augmentation
图 2. 数据增强方式

4.3. 实验结果与分析

本文将采用仅将标注数据集作为训练集的监督学习(Supervised Learning, SL)方法作为基准故障诊断方法, 并选取 PL [13]、Mixmatch [14]、SimCLR [11]作为对比方法, 在 3 种少标注样本情况下进行故障诊断实验, 实验结果如表 1 所示:

Table 1. Results of the training experiment

表 1. 试验结果数据

| 数据集 | PU | | | CWRU | | |
|----------|-------|-------|-------|-------|-------|-------|
| | 5 | 10 | 20 | 10 | 20 | 30 |
| SL | 50.38 | 63.28 | 79.35 | 42.45 | 58.00 | 71.42 |
| PL | 60.56 | 84.34 | 95.64 | 58.67 | 74.23 | 82.31 |
| Mixmatch | 65.25 | 87.75 | 98.62 | 65.57 | 79.36 | 90.43 |
| SimCLR | 77.75 | 97.37 | 99.34 | 75.35 | 81.65 | 89.21 |
| 本文方法 | 79.64 | 99.12 | 99.88 | 77.54 | 84.50 | 95.67 |

本文所提方法结合一致性正则化以及伪标签策略, 通过对输入样本施行数据增强并迫使模型输出相同目标分类值并动态筛选高置信度伪标签用于训练, 有效减缓了模型认知偏差问题, 在此基础上设计类感知对比学习模块, 在模型特征层面引入显示监督, 充分利用样本间语义信息, 实现标签匮乏情况下的故障诊断。

4.4. 消融实验

为进一步验证本文所提方法有效性, 通过设计以下六组对比实验以探究各模块效果, 消融实验结果如表 2 所示, w/o 代表在本文方法的基础上移除该模块, repl 表示被替换。

Table 2. Results of the ablation study

表 2. 消融实验结果

| 实验编号 | 数据集 | PU | | |
|------|------------------------|-------|-------|-------|
| | 标注数量 | 5 | 10 | 20 |
| 一 | SL | 50.38 | 63.28 | 79.35 |
| 二 | PL | 60.56 | 84.34 | 95.64 |
| 三 | w/o 类感知对比学习 | 62.34 | 98.53 | 99.56 |
| 四 | repl 类感知对比学习 w/对比学习 | 64.12 | 98.62 | 98.12 |
| 五 | w/o 置信度加权 | 46.25 | 71.04 | 80.17 |
| 六 | 本文方法 | 79.64 | 99.12 | 99.88 |

由实验一可知在标注样本量较少情况下采用监督学习方法无法使模型有效提取故障信息特征而导致泛化能力较差, 而实验二中在引入伪标签策略后, 无标注样本可以协同参与模型训练而扩充训练样本, 使得模型降低了过拟合于标注样本的风险, 然而其伪标签可靠性不佳限制了其诊断精度的进一步提升。

实验三结合了伪标签策略以及一致性正则化增强方法, 并采用了预测置信度筛选, 进一步提升伪标签质量与模型特征表达能力, 在各训练标注数量样本下均有有效提升。实验四、实验五和实验六均在实验三的基础上引入了对比学习策略, 进一步提升了模型特征表达能力。然而实验四中仅采用对比学习方法, 并没有充分利用样本间语义信息, 不能有效地聚集特征空间中同类样本, 诊断精度欠佳。而实验五中并没有引入置信度加权策略, 易使得模型过拟合于早期错误伪标签, 错误地拉近不同类样本在特征空间中的映射距离, 使得诊断精度不佳。本文方法结合伪标签以及一致性正则化方法, 在增强模型特征表达的同时动态筛选高置信度伪标签参与训练, 同时引入类感知对比学习模块, 利用伪标签动态引导模型拉近同类样本特征空间中映射距离, 优化决策边界, 在各标注情况下均能取得良好的诊断精度。

5. 结论

针对当前模型在标签匮乏情况下的诊断精度难以提升的问题, 本文提出一种基于类感知对比学习的半监督方法, 结合伪标签以及一致性正则化方法, 在扩充训练样本的同时增强模型对于无标注样本的特征表达能力, 并引入类感知对比学习利用伪标签动态优化特征空间映射关系, 从而提高模型泛化能力。此外, 通过实验表明了本文所提各方法的有效性。实现了故障诊断模型在仅采用少量标注样本以及大量无标注样本协同训练, 提升了其利用无标注样本获取特征信息的能力, 进一步证明了在特征空间中引入显式监督的有效性。然而, 由于机械传动系统的特性, 其无标注数据集中可能存在标注子集所未曾覆盖的故障类型, 使得标注子集与无标注样本的样本特征分布不一致[8]。一方面这会使得模型训练过程中影响决策边界的建立, 限制模型特征表达能力, 另一方面会妨碍对于新型故障类别的发现。在未来研究中会进一步探究如何减小模型训练过程中分布外样本的负面影响。

参考文献

- [1] 曹正志, 叶春明. 改进 CNN-LSTM 模型在滚动轴承故障诊断中的应用[J]. 计算机系统应用, 2021, 30(3): 126-133.
- [2] Zhang, S., Liu, Z., Chen, Y., et al. (2022) Selective Kernel Convolution Deep Residual Network Based on Channel-Spatial Attention Mechanism and Feature Fusion for Mechanical Fault Diagnosis. *ISA Transactions*, **133**, 369-383.
- [3] Xu, Z., Bashir, M., Zhang, W., et al. (2022) An Intelligent Fault Diagnosis for Machine Maintenance Using Weighted Soft-Voting Rule Based Multi-Attention Module with Multi-Scale Information Fusion. *Information Fusion*, **86-87**, 17-29. <https://doi.org/10.1016/j.inffus.2022.06.005>
- [4] Zhang, K., Tang, B., Deng, L., et al. (2021) A Fault Diagnosis Method for Wind Turbines Gearbox Based on Adaptive Loss Weighted Meta-ResNet under Noisy Labels. *Mechanical Systems and Signal Processing*, **161**, Article ID: 107963. <https://doi.org/10.1016/j.ymsp.2021.107963>
- [5] 韩特, 李彦夫, 雷亚国, 等. 融合图标签传播和判别特征增强的工业机器人关键部件半监督故障诊断方法[J]. 机械工程学报, 2022, 58(17): 116-124.
- [6] Yu, K., Ma, H., Lin, R., et al. (2020) A Consistency Regularization Based Semi-Supervised Learning Approach for Intelligent Fault Diagnosis of Rolling Bearing. *Measurement*, **165**, Article ID: 107987. <https://doi.org/10.1016/j.measurement.2020.107987>
- [7] Liao, Y., Huang, R., Li, J., et al. (2020) Deep Semi-Supervised Domain Generalization Network for Rotary Machinery Fault Diagnosis under Variable Speed. *IEEE Transactions on Instrumentation and Measurement*, **69**, 8064-8075. <https://doi.org/10.1109/TIM.2020.2992829>
- [8] Chen, Y., Mancini, M., Zhu, X., et al. (2022) Semi-Supervised and Unsupervised Deep Visual Learning: A Survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **6**, 302-314.
- [9] Sohn, K., Berthelot, D., et al. (2020) FixMatch: Simplifying Semi-Supervised Learning with Consistency and Confidence. *Proceedings of the 34th International Conference on Neural Information Processing Systems*, Vancouver, December 2020, 596-608.
- [10] Wang, T. and Isola, P. (2020) Understanding Contrastive Representation Learning through Alignment and Uniformity on the Hypersphere. *Proceedings of the 37th International Conference on Machine Learning*, 9871-9881. arXiv: 2005.10242v10.

- [11] Chen, T., Kornblith, S., *et al.* (2020) A Simple Framework for Contrastive Learning of Visual Representations. *Proceedings of the 37th International Conference on Machine Learning*, Vienna, 1597-1607.
- [12] Khosla, P., Teterwak, P., Wang, C., *et al.* (2020) Supervised Contrastive Learning. *Proceedings of the 34th International Conference on Neural Information Processing Systems*, December 2020, 18661-18673.
- [13] 吕枫, 王义, 阮胡林, 等. 深度嵌入关系空间下齿轮箱标记样本扩充及其半监督故障诊断方法[J]. 仪器仪表学报, 2021, 42(2): 55-65.
- [14] Berthelot, D., Carlini, N., Goodfellow, I., *et al.* (2019) MixMatch: A Holistic Approach to Semi-Supervised Learning. *Proceedings of the 33rd Advances in Neural Information Processing Systems*, 5603-5614. arXiv: 1905.02249v2.