

DOI: 10.19650/j.cnki.cjsi.J2311570

改进掩码自编码器的滚动轴承半监督故障诊断*

陈仁祥¹, 张 晓¹, 张 旭¹, 赵 玲¹, 夏 亮²

(1. 重庆交通大学交通工程应用机器人重庆市工程实验室 重庆 400074; 2. 重庆智能机器人研究院 重庆 400714)

摘要:针对滚动轴承在不同转速条件下数据分布不同以及实际工程应用中标签样本不足导致故障诊断精度低的问题,将领域适配模块融入掩码自编码器(MAE)中,提出了改进掩码自编码器(IMAE)的滚动轴承半监督故障诊断方法。首先,对滚动轴承振动信号进行连续小波变换(CWT)得到反应信号时频特征的二维时频图,然后对时频图随机掩码,利用无标签样本进行掩码自编码器预训练,获得数据中复杂的内在特征,减少对有标签样本的依赖;其次将领域适配模块引入到预训练后的编码器中,使用少量有标签源域数据对IMAE进行微调,在希尔伯特空间中利用最小化最大均值差异减小因转速不同造成的源域与目标域间数据分布差异;最后在Softmax分类层下实现滚动轴承半监督故障诊断。通过滚动轴承数据集实验验证,所提方法检测精度均达到94%以上,证明了该方法的可行性与有效性。

关键词:掩码自编码器;滚动轴承;不同转速;标签样本;半监督故障诊断

中图分类号: TH165+.3 TH17 **文献标识码:** A **国家标准学科分类代码:** 510.40

Improved semi-supervised fault diagnosis of rolling bearings with mask autoencoder

Chen Renxiang¹, Zhang Xiao¹, Zhang Xu¹, Zhao Ling¹, Xia Liang²

(1. Chongqing Engineering Laboratory for Transportation Engineering Application Robot, Chongqing Jiaotong University, Chongqing 400074, China; 2. Chongqing Robotics Institute, Chongqing 400714, China)

Abstract: To address the problems of different data distribution of rolling bearings under different speed conditions and low fault diagnosis accuracy caused by insufficient label samples in practical engineering applications, the domain adaptation modules are integrated into masked autoencoders (MAE). An improved masked autoencoders (IMAE) method for semi-supervised fault diagnosis of rolling bearings is proposed. Firstly, the two-dimensional time-frequency graph of the response signal is obtained by applying continuous wavelet transform (CWT) to the vibration signal of the rolling bearing. Then, the mask of the time-frequency graph is randomly masked, and the mask autoencoder is pre-trained with unlabeled samples to obtain the complex intrinsic features of the data. The reliance on labeled samples is reduced. Secondly, the domain adaptation module is introduced into the pre-trained encoder, and a small amount of labeled source domain data are used to fine-tune the IMAE, and the maximum mean difference is minimized in Hilbert space to reduce the data distribution difference between the source domain and the target domain caused by different rotational speeds. Finally, the semi-supervised fault diagnosis of rolling bearing is realized under the Softmax classification layer. Through the experimental evaluation of the rolling bearing data set, the detection accuracy of the proposed method is more than 94%, which proves the feasibility and effectiveness of the proposed method.

Keywords: mask autoencoder; rolling bearing; different speed; sample label; semi-supervised fault diagnosis

收稿日期: 2023-06-19 Received Date: 2023-06-19

* 基金项目: 国家自然科学基金(51975079, 62073051)、重庆市教委科学技术研究项目(KJZD-M202200701)、重庆市研究生联合培养基地项目(JDLHPYJD2021007)、重庆市自然科学基金创新发展联合基金(CSTB2023NSCQ-LZX0127)、重庆市专业学位研究生教学案例库(JDALK2022007)、重庆市研究生科研创新项目(2023S0072)资助

0 引言

滚动轴承作为广泛应用的基础零部件,其工作状态直接影响着整套设备的正常运行,在旋转机械中有举足轻重的作用^[1-2],对其进行故障诊断具有重要意义。

随着人工智能技术飞速发展和大数据时代来临,基于数据驱动的故障诊断方法在滚动轴承中得到了广泛应用。王鸽等^[3]提出了一种改进多尺度深度卷积神经网络,提高了模型的判别性特征提取能力。池福临等^[4]构建深度收缩残差网络与迁移学习的诊断模型,在转速变化和有噪声干扰时取得了较好效果。康守强等^[5]提出一种改进的深度Q网络进行数据不平衡和跨负载工况下滚动轴承故障识别。陈仁祥等^[6]提出基于深度注意力迁移学习的滚动轴承故障诊断方法,实现了不同转速下滚动轴承深度特征的迁移适配。尽管以上研究在故障诊断方面取得了较好效果,但均在有标记样本充足的数据集上进行实验验证,依赖于大量标记样本的监督训练。而实际工程中,设备长期处于正常工作状态,获取大量有标记故障样本十分耗时且昂贵,导致数据集含有大量无标注数据;此外,有标记样本稀缺而无法提供足够的监督训练时,深度学习模型容易过拟合,导致故障诊断效果较差。

近年来,有标签样本不足的故障诊断问题受到了许多学者的关注。目前用于解决标签样本不足的方法主要包括数据增强和迁移学习。数据增强通过对训练样本进行扩充以提高模型的泛化能力,黄南天等^[7]搭建了一种改进辅助分类生成对抗模型,并生成大量与真实数据分布近似的虚假样本。迁移学习可将源域标签数据下学习到的知识迁移应用到其他相关的领域,通过在其他域中获得的具有公共特征提取的模型来提高模型的泛化能力。张西宁等^[8]提出了基于迁移学习的小样本轴承故障诊断方法,在仅有少量目标域数据的情况下也能取得很高的诊断准确率。上述方法虽在样本不足时有较好表现,但其自身还有局限,如生成对抗网络训练困难,容易导致模型梯度消失,而迁移学习难以衡量源域和目标域之间的数据分布差异,且要求源域需要大量标签数据训练模型。

半监督学习在少量标签样本的基础上,配合大量无标签样本完成模型训练,既避免了样本量不足时深度学习精度不高的问题,又克服了标签样本匮乏的难题。为解决不同转速条件下数据分布不同以及标签样本不足导致滚动轴承故障诊断精度低的问题,提出改进掩码自编码器(improve masked autoencoders, IMAE)的半监督故障诊断方法。首先,利用连续小波变换(continuous wavelet transform, CWT)得到反应信号时频特征的二维时频图,然后对时频图随机掩码,利用无标签样本进行掩码自

编码器(masked autoencoders, MAE)预训练,获得数据中复杂的内在特征,减少对有标签样本的依赖;其次将领域适配模块引入到预训练后的编码器中,使用少量有标签源域数据对 IMAE 进行微调,在希尔伯特空间中利用最小化最大均值差异(max mean discrepancy, MMD)减小因转速不同造成源域与目标域间数据分布差异;最后在 Softmax 分类层下实现滚动轴承半监督故障诊断。

1 IMAE 原理

所建立 IMAE 半监督网络结构如图 1 所示,包含预训练网络和微调网络两部分。预训练网络包括多个 MAE,微调网络结构由训练完成的编码器、领域适配模块以及 Softmax 分类层构成。

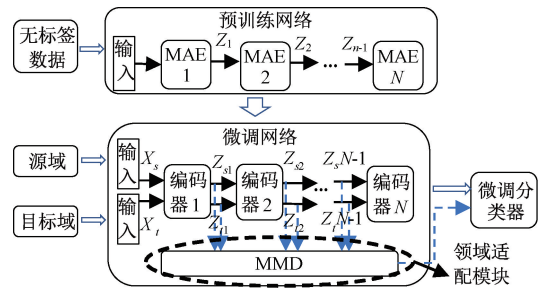


图 1 改进掩码自编码器网络结构

Fig. 1 Improved mask autoencoder network structure

1.1 MAE 原理

MAE 是一种非对称编码器-解码器结构,其中编码器只在未掩码的 patches 上工作,而解码器则负责结合掩码 patches 从隐藏空间重构图像^[9],如图 2 所示。在掩码处理层完成对样本的部分特征随机遮盖操作后,利用样本的部分特征预测被掩盖掉的特征以训练自编码器,使得自编码器尽量保留样本结构特征信息。

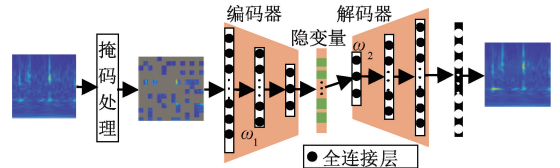


图 2 掩码自编码器网络结构

Fig. 2 Mask autoencoder network structure

1) 掩码处理

掩码处理(图 3)能把原始故障样本处理成更多的新训练样本,也能给正常样本添加额外的约束,从而更好获得数据本身的深层结构特征信息^[10]。在输入端,被掩码的数据和未被掩码的数据一起输入神经网络。在输出端,被掩码的原始信号作为解码器输出的监督信号。通

过这种简单的变换,故障数据可以成倍添加至训练集中,更好地完成模型的学习与训练。

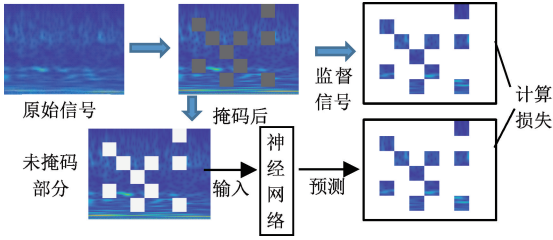


图3 掩码处理示意图

Fig. 3 Schematic diagram of mask

2) 编码器

编码器采用视觉自注意力模型(vision transformer, ViT)将输入信息映射到某个更低维度的空间,其仅对未屏蔽 patches 进行操作。编码器依据原始数据通过位置编码后的有效值提取特征,并输出相应的特征向量到解码器。其计算公式为:

$$\mathbf{z} = f_{enc}(\mathbf{x}) = f(\boldsymbol{\omega}_1 \mathbf{x} + \mathbf{b}_1) \quad (1)$$

式中: \mathbf{x} 是原始数据; $f_{enc}(\cdot)$ 为编码器模型; \mathbf{z} 是编码器网络输出的样本编码; $\boldsymbol{\omega}_1$ 是编码器的权重矩阵; \mathbf{b}_1 为编码器的偏置向量。

3) 解码器

解码器将特征编码重构为原输入,其最后一层是线性投影层,它输出的通道数等于图像的像素数。因此,解码器的输出将被进一步重塑为原始图像的形状。

$$\mathbf{x}' = f_{dec}(\mathbf{z}) = f(\boldsymbol{\omega}_2 \mathbf{z} + \mathbf{b}_2) \quad (2)$$

式中: \mathbf{x}' 是解码器输出的预测原本被淹没位置的部分信息特征; $f_{dec}(\cdot)$ 是解码器模型; $\boldsymbol{\omega}_2$ 是解码器权重矩阵; \mathbf{b}_2 为解码器偏置向量。

损失函数使用均方误差(mean square error, MSE)直接最小化重建图像 \mathbf{x}' 和原始图像 \mathbf{x} 之间的距离。

$$L(\mathbf{x}, \mathbf{x}') = \|\mathbf{x} - \mathbf{x}'\|_2^2 \quad (3)$$

1.2 微调网络

微调网络由预训练后的编码器、领域适配模块以及 Softmax 分类层构成。为减小因转速变化导致数据分布差异,引入领域适配模块利用源域与目标域特征之间的 MMD 距离作为两域之间损失,通过最小化 MMD 距离来减小源域与目标域之间的差异,其计算公式为:

$$MMD^2(D_s, D_t) \triangleq \left\| \frac{1}{n_1} \sum_{i=1}^{n_1} \phi(\mathbf{x}_i) - \frac{1}{n_2} \sum_{j=1}^{n_2} \phi(\mathbf{x}_j) \right\|_H^2 \quad (4)$$

式中: D_s 表示给定标记源域; D_t 表示给定未标记目标域; \mathbf{x}_i 和 \mathbf{x}_j 分别表示源域和目标域样本; n_1 和 n_2 分别表示源域和目标域的样本数量; H 是具有特征核 k 的再生核希尔伯特空间(RKHS)。其中, $\phi(\cdot)$ 表示将原始样本映射到

RKHS 的一些特征映射。

在微调过程中将目标域数据和源域数据共同输入到无解码器的网络中,以在编码器中学习通用特征表示。然后,从各层的输出中提取特征并计算 MMD 领域分布差异。将 MMD 加入到微调模型的目标损失函数中,通过同步训练更新权重,提高领域适配效率。

1.3 损失函数

损失函数由两部分构成,用 $L(\mathbf{x}, \boldsymbol{\omega})$ 表示对于数据 \mathbf{x} 关于参数 $\boldsymbol{\omega}$ 的损失函数,表达式为:

$$l(\mathbf{x}, \boldsymbol{\omega}) = \sum_{n=1}^N L_1(\mathbf{x}, \mathbf{x}'; \boldsymbol{\omega}_1, \boldsymbol{\omega}_2) + L_2(p(c|\mathbf{x}); \boldsymbol{\omega}_1, \boldsymbol{\omega}_3, \boldsymbol{\omega}_4) + L_3(\eta(c|\mathbf{x}); \boldsymbol{\omega}_1, \boldsymbol{\omega}_2, \boldsymbol{\omega}_3, \boldsymbol{\omega}_4) \quad (5)$$

式中: L_1 为重构阶段的损失函数; \mathbf{x} 表示原始数据; \mathbf{x}' 表示重构数据; L_2 为故障诊断网络中分类损失函数; $p(c|\mathbf{x})$ 表示数据掩码后编码器输出结果; $\boldsymbol{\omega}_1$ 、 $\boldsymbol{\omega}_3$ 和 $\boldsymbol{\omega}_4$ 分别为该网络中的权值矩阵; L_3 为 $\boldsymbol{\omega}_1$ 、 $\boldsymbol{\omega}_2$ 、 $\boldsymbol{\omega}_3$ 和 $\boldsymbol{\omega}_4$ 这 4 组参数联合构成的正则化损失函数; $\eta(c|\mathbf{x})$ 表示数据未掩码部分编码器输出结果。

$$L = L(\mathbf{x}, \mathbf{y}) + \gamma MMD^2(D_s, D_t) \quad (6)$$

式中: L 表示微调阶段用于监督模型训练的损失函数,采用交叉熵损失函数; γ 表示权重系数。通过最小化预测的掩码信息、原始样本掩码信息的距离、故障诊断网络的分类损失以及领域适配模块中 MMD 最小化,完成网络的训练。

2 诊断流程

改进掩码自编码器的滚动轴承半监督故障诊断分为 3 个主要步骤,信号预处理并划分数据集、模型设计与训练和故障分类识别,流程如图 4 所示。

1) 信号预处理并划分数据集。获取滚动轴承故障振动信号,利用 CWT 获得时频图,划分训练集(源域)和测试集(目标域),构建样本集。

2) 构建 IMAE 网络模型,设置网络初参数。通过批量的方式将无标签数据(预训练数据)输入掩码自编码器网络中,通过式(3)的目标损失函数对 MAE 模型的参数进行更新,实现 MAE 无监督预训练。

3) 重复步骤 2),直至网络收敛,完成 MAE 的预训练,得到训练完成的编码器。

4) 将具有少量标签的源域数据和目标域数据输入训练好的编码器,对预训练后的网络进行微调,通过领域适配模块中的 MMD 对源域与目标域特征进行适配,减小因转速不同导致的特征分布差异。

5) 运用训练完成的模型对测试集进行分类识别,最终得出故障诊断结果。

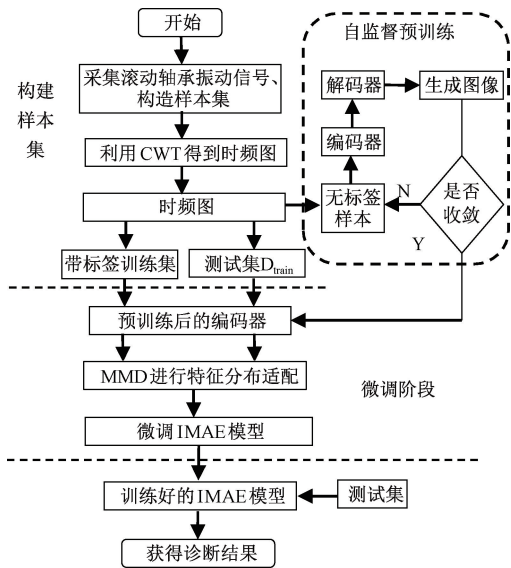


图 4 方法流程

Fig. 4 Flow chart of method

3 实验验证

3.1 实验数据

实验数据来自机械故障综合模拟实验台,由电机、联轴器、滚动轴承等组成,如图 5 所示。采用 er-16k 型深沟球轴承进行实验,轴承节径 1.516 in(1 in=25.5 mm),包含 9 个滚动体,接触角 9.08°。

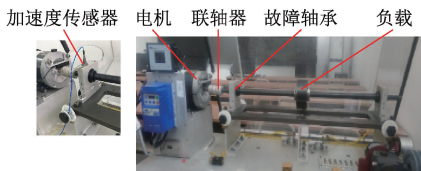


图 5 轴承故障模拟试验台

Fig. 5 Bearing fault test rig

采集时转速分别为 A 工况 1 000 r/min;B 工况 1 200 r/min;C 工况 1 400 r/min。加载负荷均为 5 kg,采样频率为 25.6 kHz。所设置的滚动轴承四种故障类型分别为内圈故障(I)、外圈故障(O)、滚动体故障(B)、正常(N)。

3.2 数据集构建

在机械装备实际运行中,历史数据中含有大量的无标签样本和少量的有标签样本^[11]。为验证所提方法在有标签训练标签样本不足、变转速下故障诊断效果,每种故障分别选取 30 个标签样本作为训练集,100 个样本作为测试集,每段采样点设置为 1 024 个。试验数据集分别如表 1 所示。

表 1 滚动轴承数据集

Table 1 Rolling bearing data set

故障类型	正常	内圈	外圈	滚动体	转速
标签	0	1	2	3	/(r·min ⁻¹)
A	训练集	30	30	30	1 000
	测试集	100	100	100	
B	训练集	30	30	30	1 200
	测试集	100	100	100	
C	训练集	30	30	30	1 400
	测试集	100	100	100	

4 实验结果及分析

试验基于 PyCharm2020.1.3 中的 pycorch 框架,硬件环境为 AMD Ryzen 7 4800 H with Radeon Graphics 2.90 GHz 双处理器和 NVIDIA Quadro P3060 显卡。训练过程采用 Adam 优化器对参数进行优化,学习率为 0.001,迭代次数为 200,Batch-size 设置为 16,为防止网络过拟合,设置 dropout 数值为 0.5,模型具体参数如表 2 所示。将所提方法 IMAE 与 MAE、深度网络自适应(deep domain confusion, DDC)^[12]、残差神经网络(residual networks, ResNet)^[13]、深度适配网络(deep adaptation network, DAN)^[14]、卷积神经网络+支持向量机(CNN+SVM)、深度置信网络(deep belief networks, DBN)^[15-16]和自监督预训练对比学习(self-supervised pretraining via contrast learning, SSPCL)^[17]进行对比,验证所提方法有效性。

表 2 模型参数设置

Table 2 Model parameters setting

模型结构	参数	值
Patch 嵌入	输入通道数	3
	输出通道数	1 024
	块大小	16×16
ViT 编码器	Transformer 层数	24
	多头自注意力头数	16
	MLP 扩展比例	4
	嵌入维度	1 024
ViT 解码器	Transformer 层数	8
	多头自注意力头数	16
	MLP 扩展比例	4
	嵌入维度	512

4.1 同转速不同有标签样本量下故障诊断结果与分析

为验证所提方法在标签数据稀少情况下故障诊断效果,与其他故障诊断方法进行对比。使用不同数量的有标记样本进行实验,训练集每类样本标记量分别为 8、10、15、20 和 30,测试集每类样本量均为 100,在多个转速下分别进行实验,结果如表 3 所示。

表 3 同转速下准确率
Table 3 Classification accuracy at the same speed %

不同转速	实验方法	标签样本数量				
		8 个	10 个	15 个	20 个	30 个
A→A	MAE	96.5	98	99.25	100	100
	ResNet	56.5	60	60.75	62.75	71
	CNN+SVM	86.28	88.58	89.23	89.75	90.75
	DDC	59.75	63	64.5	66.75	69.75
	DAN	67.5	73.5	75	75	85
	SSPCL	84	87.5	88	91.5	92.75
B→B	MAE	99.5	99.75	100	100	100
	ResNet	67	68.5	74.25	75	85.5
	CNN+SVM	84.25	86.75	88.35	89.05	90
	DDC	57.25	62.5	62.75	72.5	74.25
	DAN	71.75	72.75	80.5	85.5	88.75
	SSPCL	95	96.5	97.75	98	98
C→C	MAE	100	100	100	100	100
	ResNet	72.5	79.5	79	87.25	90.5
	CNN+SVM	84.15	85.23	87.86	89.85	91.25
	DDC	61.5	64.75	68.5	73.25	78
	DAN	73.25	74.5	82.25	90.5	92
	SSPCL	93.5	94.25	94.75	97.5	97.75

观察表 3 可知,工况 A,在标记样本量为 8 时,所提方法准确率已达到 96.5%,当样本量为 20 时,准确率均达到 100%;工况 B 在训练样本量为 8 时所提方法准确率高达 99.5%,当标签样本量为 15 时,准确率达到 100%,且处于稳定状态。在 3 种转速下,所提方法均能达到较好的准确率,其他方法在样本数量较少的情况下准确率较低,如工况 A 中 ResNet 在样本量为 8 时准确率仅 56.5%,当样本量达到 30 时准确率也仅 71%,可以看出 ResNet 可以识别深层故障特征但不能利用无标签数据实现网络训练。在样本量为 8 时,DDC 与 DAN 因含有适配层,准确率相较于 ResNet 分别提高了 3.25% 和 10%。SSPCL 利用无监督数据进行对比学习预训练,其准确率相较于 ResNet 提高了 27.5%。CNN+SVM 的识别准确率相较于 SSPCL 提高了 2.28%,这是因为该方法既可以利用 CNN 提取深层故障特

征,又可以利用 SVM 来解决训练样本不足的问题。但该方法低于 MAE 的识别准确率,说明解决有标签样本量不足时 MAE 方法优于 SVM。MAE 通过对编码器和解码器进行自监督学习,可将少量故障数据增强为数量足够大的自监督样本集合,因此其泛化能力相比于 DAN 得到进一步的提高,准确率相较于 DAN 提高了 29%。

进一步分析不同标记样本量下不同方法准确率变化情况,如图 6 所示,随着训练样本标签数量的增加,网络的过拟合得到缓解,每种方法准确率都有不同程度的提高。在样本数量为 150 的情况下,网络的过拟合问题已经得到了很大的缓解,各方法的准确率已经逐渐接近。结果说明其他方法需要数量足够的标记样本,所提方法对标记样本数量要求较少。

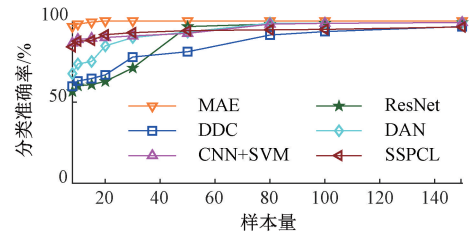


图 6 A→A 不同标签样本量下准确率对比

Fig. 6 A→A comparison of accuracy under different sample size of labels

为分析掩码自编码器数据扩充的可行性,对扩充后的图像进行分析,滚动轴承外圈故障时频图的掩码和重建如图 7 所示。对原始图像进行随机掩码处理(图 7(b)),将未掩码的 patches 输入编码器以学习其深层特征信息,并将得到的特征信息输入到解码器以重建一个完整的未掩码图像(图 7(d))。图 7(a)和(d)时频特征一致,说明使用掩码自编码器生成数据集可对无标签新数据进行扩展,可解决因训练数据不足导致网络模型过拟合问题。

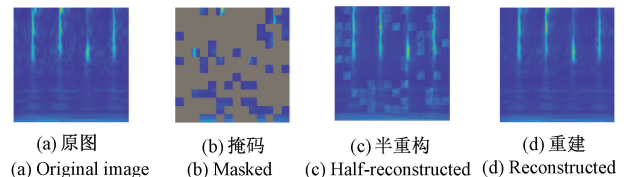


图 7 滚动轴承外圈故障时频图的掩码和重建

Fig. 7 Mask and reconstruction of fault time-frequency diagram of rolling bearing outer ring

4.2 不同转速小样本下故障诊断结果及分析

在实际工程中,滚动轴承通常在不同转速下运行,会导致数据分布差异,影响诊断结果。为验证所提方法在标签数据样本不足、转速变化时的故障诊断性能,分别使用包含 8、10、15、20、30 个标签数据样本的训练集,实验结果如表 4 所示(A→B 表示 A 工况为训练样本,B 工况

表 4 不同转速下的准确率

Table 4 Classification accuracy of different speeds %

样本数量	实验方法	不同转速						平均
		A→B	A→C	B→C	B→A	C→A	C→B	
8 个	DBN	65.35	59.21	58.63	57.42	58.63	60.26	59.92
	ResNet	55.75	53.50	69.00	51.75	49.75	68.00	57.96
	CNN+SVM	60.25	56.45	62.35	51.25	54.65	63.50	58.08
	MAE	88.75	87.00	86.75	84.00	81.25	85.75	85.58
	DDC	62.50	57.25	63.00	50.75	57.75	63.75	59.20
	DAN	67.50	62.51	67.75	63.25	53.50	73.00	64.59
	SSPCL	86.00	82.75	96.75	82.25	62.00	81.00	81.79
	所提方法	92.75	97.00	96.75	94.00	91.25	95.75	94.58
	10 个	DBN	67.25	60.14	59.63	58.75	60.14	62.50
ResNet		57.00	68.50	74.50	50.75	57.25	73.25	63.54
CNN+SVM		62.45	58.35	64.10	53.75	58.62	70.20	61.23
MAE		85.25	86.75	88.25	86.00	82.50	88.00	86.13
DDC		64.50	57.75	63.75	60.50	60.00	64.75	61.88
DAN		72.50	58.75	74.25	65.75	54.25	75.25	66.79
SSPCL		87.25	83.00	97.00	83.50	65.25	83.50	83.25
所提方法		95.25	96.75	98.25	96.00	92.50	98.00	96.13
15 个		DBN	68.15	61.79	62.15	59.85	61.2	63.71
	ResNet	58.75	59.25	74.25	51.75	74.75	55.5	62.38
	CNN+SVM	66.15	59.45	65.23	58.65	58.23	71.25	63.16
	MAE	87.00	86.75	90.75	87.00	92.75	93.70	89.66
	DDC	66.75	58.75	67.25	61.25	59.00	66.50	63.25
	DAN	74.50	64.50	78.00	68.50	58.50	82.75	71.13
	SSPCL	91.25	91.00	97.50	91.50	66.25	85.75	87.21
	所提方法	97.00	97.00	100.00	93.95	96.75	98.70	97.23
	20 个	DBN	70.21	62.14	63.74	60.78	62.10	64.75
ResNet		62.50	58.50	80.5	62.75	74.75	56.00	65.83
CNN+SVM		67.25	60.85	65.74	60.52	59.08	73.50	64.49
MAE		90.25	91.75	92.35	89.25	93.00	93.75	91.73
DDC		67.00	60.50	65.50	63.25	59.75	79.75	65.96
DAN		80.50	73.25	88.50	81.75	52.75	90.00	77.79
SSPCL		90.50	92.30	97.60	90.78	68.00	87.00	87.69
所提方法		98.25	97.75	100.00	94.00	98.50	99.25	97.96
30 个		DBN	70.58	63.32	65.86	61.75	62.78	66.12
	ResNet	66.75	64.00	86.00	65.00	77.75	77.25	72.79
	CNN+SVM	69.85	60.45	70.25	60.88	61.45	75.36	66.37
	MAE	91.75	93.00	94.50	90.25	93.00	93.65	92.69
	DDC	70.50	62.25	76.25	70.75	62.75	81.00	70.58
	DAN	85.75	75.25	92.25	83.00	63.00	92.50	81.96
	SSPCL	91.00	93.50	97.40	92.75	69.50	87.25	88.56
	所提方法	99.25	99.00	100.00	97.00	98.75	100.00	99.00

为测试样本, A→C 等同理)。

由表 4 可知,在标签样本数量为 8 时,所提方法平均准确率达到 94.58%,验证了所提方法在不同转速下对滚动轴承进行故障诊断的有效性。而 MAE 的平均识别精度仅为 85.58%,这是因为 MAE 在转速变化时不能减小源域与目标域之间的分布差异使得诊断效果不佳。SSPCL 通过无监督对比学习能够提取到数据中判别性特征,但由于没有适配层,平均精度仅为 81.79%。ResNet 同样不具备适配能力,准确率仅有 57.96%。DDC、DAN 能够通过分布适配拉近两域特征分布,平均准确率为 59.2%、64.59%,均高于 ResNet,说明不同转速下进行适配的必要性。CNN+SVM、DBN 也不具备适配能力,其平均识别准确率均低于 SSPCL。为进一步说明准确率与训练样本的关系,继续增加训练集有标签样本数量。给出不同标签样本量下 B→C 工况下不同方法的识别准确率。由表 4 以及图 8 可知,在样本标签数量不足(数目 8~30)的情况下,所提方法在标签样本不足和转速不同时均具有良好诊断效果,证明了所提方法的可行性与有效性。

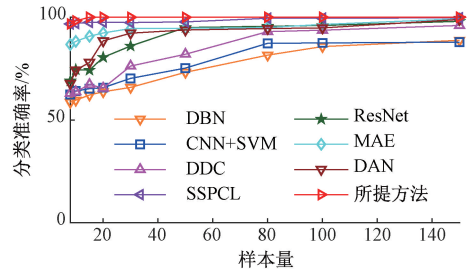
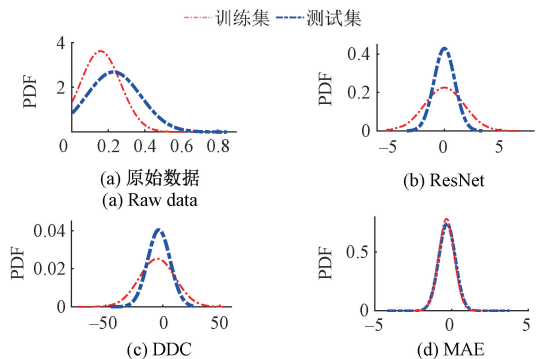


图 8 B→C 不同标签样本量下准确率对比准确率对比
Fig. 8 B→C comparison of accuracy under different sample size of labels

工况 C→B 在样本量为 8 时,不同方法概率密度如图 9 所示。观察图 9 可知,相对于原始数据,DDC、DAN 和所提方法的源域、目标域概率分布曲线有一定的拉近,这是因为 DDC、DAN 与所提方法都具备适配层,说明适配能够减小分布差异,且所提方法适配效果最好。ResNet 与 MAE 不具备分布适配能力,不能减小两域之间的分布差异。



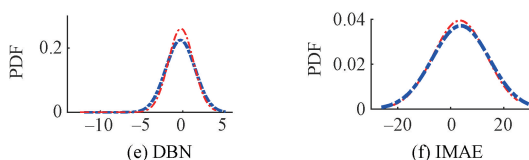


图9 不同方法概率密度

Fig.9 Probability density diagram of different methods

5 结 论

针对滚动轴承在标签样本不足、不同转速下诊断精度较低的问题,提出改进掩码自编码器的半监督故障诊断方法。该方法只需要少量标记样本和大量无标记样本对模型进行训练,就能够达到良好诊断效果。该方法具有如下优势。

1)所提方法使用无标签的源域数据对改进掩码自编码器进行预训练,然后将预训练模型参数用于目标域的领域自适应的微调过程,降低了算法的复杂度,避免了在目标域重新训练模型的需求。

2)所提方法在领域适配的过程中采用相同的预训练中隐藏层连接的模型进行源域训练和目标域训练,微调不需要解码过程,避免了解码过程的耗时和复杂度,同时降低了领域适配的复杂度。

参考文献

- [1] OJAGHI M, YAZDANDOOST N. Oil-whirl fault modeling, simulation, and detection in sleeve bearings of squirrel cage induction motors[J]. IEEE Transactions on Energy Conversion, 2015, 30(4): 1537-1545.
- [2] XIANG L, WEI H, QIAN D. Cross-domain fault diagnosis of rolling element bearings using deep generative neural networks [J]. IEEE Transactions on Industrial Electronics, 2019, 66(7): 5525-5534.
- [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] 池福临, 杨新宇, 邵思羽, 等. 基于深度收缩残差网络的轴承变工况故障诊断[J]. 计算机集成制造系统, 2023, 29(4): 1146-1156.
- CHI F L, YANG X Y, SHAO S Y, et al. Bearing off-design fault diagnosis based on deep shrinkage residual

network[J]. Computer Integrated Manufacturing System, 2023, 29(4): 1146-1156.

- [5] 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. Rolling bearing fault diagnosis method based on improved DQN network[J] Journal of Instrumentation, 2021, 42(3): 201-212.
- [6] 陈仁祥, 唐林林, 胡小林, 等. 不同转速下基于深度注意力迁移学习的滚动轴承故障诊断方法[J]. 振动与冲击, 2022, 41(12): 95-101, 195.
- CHEN R X, TANG L L, HU X L, et al. Rolling bearing fault diagnosis method based on deep attention transfer learning at different speed [J]. Journal of Vibration and Shock, 2022, 41(12): 95-101, 195.
- [7] 黄南天, 杨学航, 蔡国伟, 等. 采用非平衡小样本数据的风机主轴故障深度对抗诊断[J]. 中国电机工程学报, 2020, 40(2): 563-574.
- HUANG N T, YANG X H, CAI G W, et al. Fault depth countermeasure diagnosis of fan main bearing using unbalanced small sample data [J]. Chinese Journal of Electrical Engineering, 2020, 40(2): 563-574.
- [8] 张西宁, 余迪, 刘书语. 基于迁移学习的小样本轴承故障诊断方法研究[J]. 西安交通大学学报, 2021, 55(10): 30-37.
- ZHANG X N, YU D, LIU SH Y. Research on small sample bearing fault diagnosis method based on transfer learning [J]. Journal of Xi'an Jiaotong University, 2021, 55(10): 30-37.
- [9] HE K M, CHEN X L, XIE S N, et al. Masked autoencoders are scalable vision learners [C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2022.
- [10] 李刚, 徐长明, 龚翔, 等. 基于掩码自编码器的小样本深度学习道岔故障诊断模型[J]. 中国铁道科学, 2022, 43(6): 175-185.
- LI G, XU CH M, GONG X, et al. Turnout fault diagnosis model based on mask autoencoder based on small sample deep learning [J]. China Railway Science, 2022, 43(6): 175-185.
- [11] 江丽, 郭顺生. 基于半监督拉普拉斯特征映射的故障

- 诊断[J]. 中国机械工程, 2016, 27(14): 1911-1916.
- JIANG L, GUO SH SH. Fault diagnosis based on semi-supervised laplacian feature mapping [J]. China Mechanical Engineering, 2016, 27(14): 1911-1916.
- [12] TZENG E, HOFFMAN J, ZHANG N, et al. Deep domain confusion: Maximizing for domain invariance[J]. Computer Science, 2014, DOI: 10.48550/arXiv.1412.3474.
- [13] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition [J]. 2016 IEEE Conference on Computer Vision and Pattern Recognition, 2016.
- [14] LONG M SH, CAO Y, WANG J M, et al. Learning transferable features with deep adaptation networks[C]. International Conference on Machine Learning, 2015: 97-105.
- [15] 李巍华, 单外平, 曾雪琼. 基于深度信念网络的轴承故障分类识别[J]. 振动工程学报, 2016, 29(2): 340-347.
- LI W H, SHAN W P, ZENG X Q. Bearing fault classification and recognition based on deep belief network[J]. Journal of Vibration Engineering, 2016, 29(2): 340-347.
- [16] 陈仁祥, 杨星, 胡小林, 等. 深度置信网络迁移学习的行星齿轮箱故障诊断方法[J]. 振动与冲击, 2021, 40(1): 127-133.
- CHEN R X, YANG X, HU X L, et al. Fault diagnosis

method of planetary gearbox based on deep confidence network Transfer learning [J] Vibration and Shock, 2021, 40(1): 127-133.

- [17] DING Y, ZHUANG J, DING P, et al. Self-supervised pretraining via contrast learning for intelligent incipient fault detection of bearings[J]. Reliability Engineering & System Safety, 2022, 218: 108126.

作者简介



陈仁祥(通信作者), 分别在 2007 年和 2012 年于重庆大学获得学士学位和博士学位, 现为重庆交通大学教授、博士生导师, 主要研究方向为故障预测与健康管理的。

E-mail: manlou.yue@126.com

Chen Renxiang (Corresponding author) received his B. Sc. degree and Ph. D. degree both from Chongqing University in 2007 and 2012, respectively. He is currently a professor and a Ph. D. advisor at Chongqing Jiaotong University. His main research interests include prognostics and health management.



张晓, 2020 年于潍坊学院获得学士学位, 现为重庆交通大学硕士研究生, 主要研究方向为机电装备故障诊断。

E-mail: 1519835029@ qq.com

Zhang Xiao received her B. Sc. degree from Weifang University in 2020. She is currently a M. Sc. candidate at Chongqing Jiaotong University. Her main research interests include fault diagnosis of electromechanical equipment.