DOI: 10.13382/j.jemi.B2205664

基于改进域对抗网络的轴承故障迁移诊断方法*

张金汇 谢林柏
(江南大学物联网工程学院 无锡 214122)

摘 要:针对工业场景下复杂工况导致的轴承故障数据特征分布差异,以及难以获得大量有标签数据的问题,提出一种基于 Wasserstein 距离与局部最大平均偏差(LMMD)改进的一维卷积子域适应对抗迁移网络(SANN)。该网络首先构建 CNN 特征提取器进行预训练,学习领域特征表示,在对抗训练阶段,对抗层引入 Wasserstein 距离来度量源域与目标域的差异,实现边缘分布的对齐,固化训练结果。在特征提取层引入 LMMD 计算模块捕获每个类别的细粒度信息,实现条件分布的对齐。通过两种变工况下的轴承故障数据集对该模型性能进行验证。实验结果表明,无监督的条件下,本文所提方法在目标数据集上相较于基础域对抗网络分别提高了 5.0% 和 6.9% 的识别精度,性能优于现有的迁移算法。

关键词: 轴承故障;智能故障诊断;无标签数据;对抗迁移网络
中图分类号: TH17 **文献标识码**: A **国家标准学科分类代码**: 510.40

Bearing fault migration diagnosis method based on improved domain adversarial network

Zhang Jinhui Xie Linbo
(School of Internet of Things Engineering, Jiangnan University, Wuxi 214122, China)

Abstract: Aiming at the difference of bearing fault data feature distribution caused by complex working conditions in industrial scenes and the difficulty of obtaining a large number of labeled data, an one-dimensional convolution subdomain adaptive adversarial neural network (SANN) based on Wasserstein distance and local maximum mean discrepancy (LMMD) is proposed. Firstly, the network constructs a feature extractor based on CNN for pre-training and learning the domain feature representation. In the adversarial training stage, the adversarial layer introduces Wasserstein distance to measure the difference between the source domain and the target domain, realize the alignment of marginal distribution and solidify the training results. In the feature extraction layer, LMMD calculation module is introduced to capture the fine-grained information of each category to realize the alignment of conditional distribution. The performance of the model is verified by the bearing fault data sets under two different working conditions. The experimental results show that under unsupervised conditions, the proposed method improves the recognition accuracy by 5.0% and 6.9% respectively compared with the basic domain adversarial network on the target data set, and the performance is better than the existing migration algorithms.

Keywords: bearing fault; intelligent fault diagnosis; unlabeled data; adversarial network

0 引 言

滚动轴承在工业装备中扮演着重要的角色,同时也是最容易发生损伤的机械元件之一,随着使用需求的扩大和制造业的发展,其安全和可靠性也更受重视,如航空、化工和机械设备制造等对轴承的安全性具有极高的要求,因此轴承的健康状态管理和故障诊断逐渐成为工业界和学术界关注的热点问题[1]。

传统的轴承故障诊断方法一般需要人工对振动信号进行特征提取,利用一些经典的机器学习方法如人工神经网络(artificial neural networks, ANN)、贝叶斯网络(Bayes)、随机森林等构建分类器进行故障诊断识别。Cai 等[2]总结了贝叶斯网络在故障诊断领域的研究。Yin

收稿日期: 2022-07-07 Received Date: 2022-07-07
*基金项目:国家自然科学基金(61873112)项目资助

等[3]介绍了复杂环境下基于支持向量机的过程监测与故障诊断的研究进展。此类方法往往难以挖掘深层的故障特征,且模型结构简单,更适合小样本的情况。随着传感技术和深度学习技术的发展,深度学习不断进行跨域应用[4],轴承故障诊断领域也向深度学习进行了充分地借鉴和融合[5]。深度学习方法能够自适应提取信号特征,不再依赖于人工经验和领域内的专家知识,大大提高了网络适应性[6-7]。在实际工业场景下,由于机械设备在故障状态下运行的风险很高,因此很难获得大量的有标签的故障数据。另外,实验室条件下通过实验设备仿真机械运行状态得到的数据与实际工况下所采集的检测数据之间存在差异。上述因素均会导致实际工业场景中深度学习方法的性能发生退化,这也限制了智能诊断方法的工业化应用。

针对现有的问题,域适应(domain adaption,DA)算法可以提供一种有效的解决方法,通过拉近源域与目标域的特征分布,提取到域不变特征,以此提高目标域故障数据的识别精度。王廷轩等[8]采用卷积网络提取故障的特征,最小化多核最大均值差异(maximum mean discrepancy,MMD),以调整多层卷积模型参数,实现跨工况故障迁移诊断。2014 年 Goodfellow 等[9]提出生成对抗网络(generative adversarial networks,GANs)之后,具有博弈思想的对抗式训练逐渐被运用到迁移学习中。Ganin 等[10]首先在神经网络中加入了对抗机制,构建了域对抗网络(domain-adversarial neural network,DANN),该网络以特征提取器和判别器的博弈,在不断对抗中得到域不变特征,取得了不错的效果。Shao 等[11]采用短时傅里叶变换作数据的预处理,并应用 MMD 和对抗式的域混淆方法来获取域不变信息,从而完成轴承故障的迁移诊断。Chen 等[12]以对抗域适应为基本思想,提出了一种应对大跨度特征迁移的对抗式故障诊断模型,在旋转机械故障的诊断上获得了较好的效果。

虽然上述方法取得了不错的识别效果,但仍存在一些问题。比如,生成对抗网络在训练过程中需要最小化 KL(Kullback-Leibler)散度和 JS(Jensen-Shannon)散度的差值,这容易造成梯度消失和训练不稳定的问题[13],因 KL 散度、JS 散度在两个分布很远且不存在重叠部分的时候,不能反映两个分布之间的距离远近,同时也不能提供有效梯度。再如上述方法往往关注于边缘分布的差异,而忽略条件分布的对齐,从而丢失一些类别细粒度信息,导致分类准确率较低。

基于以上的分析,本文提出一种子域适应对抗迁移网络(subdomain adversarial neural network,SANN)轴承故障迁移诊断模型。首先在预训练中利用足够的源域数据训练一个基础 1DCNN 模型,完成原始振动信号到故障特征的深度映射。随后在对抗训练阶段建立一个改进域对抗模型来学习域不变特征。在对抗层引入 Wasserstein 距离对提取的源域数据和目标域数据的特征进行有效的距离度量,发挥 Wasserstein 距离在梯度方面的优势[13],以获得稳定的训练结果。在分类层引入局部最大平均偏差(local maximum mean discrepancy,LMMD)[14]来度量源域和目标域数据嵌入相关子域的差异,以捕获每个类别的细粒度信息,调整相关子域同一类别下的分布。通过最大化域判别损失,最小化分类器和特征提取器的损失,在不断的对抗中学习域不变特征。上述对抗学习过程,可将已标记的源域可迁移特征引入到新的且相关的目标域诊断任务中,而无需任何标记样本,实现端到端的无监督轴承故障迁移诊断。

1　基础知识

1.1　领域自适应

对于一个有标签的源域 $D_s=\{x_s,y_s\}$ 和一个无标签的目标域 $D_t=\{x_t\}$,假设它们共享特征空间,$X_s=X_t$,共享类别空间,$Y_s=Y_t$,条件概率分布相同,$P(y_s\mid x_s)=P(y_t\mid x_t)$。两域边缘概率分布不同,$P(x_s)=P(x_t)$,类内条件概率分布也可能不同,$P(x_s\mid y_s)=P(x_t\mid y_t)$。领域自适应的目标是利用数据信息去学习一个映射,获得源域与目标域的共同特征,来提高预测目标域的标签类别的精度。

根据上述定义,领域自适应的特性是源域数和目标域数据的概率分布不同,但分类任务相同,即分类标签一致,这种性质允许目标域数据可以没有标签。因此,无标签数据下的迁移学习的方法大多数都是聚焦于领域自适应方法的研究。

1.2　卷积神经网络

典型的 CNN 主要由两部分构成,一个是多层的降维提取结构,主要包含了卷积层和池化层,另一个是多层式的全连接神经网络结构。CNN 结构功能如下所述。

输入层,输入层为调整过的振动信号,像素矩阵等数据,一般数据输入前会进行预处理操作,如去噪、归一化等。

卷积层,卷积层是 CNN 的核心,用于提取输入的高维度特征,由多个卷积内核组成,当每个核在输入映射上滑动时,共享权重和偏置,通过将输入信号与卷积核进行卷积运算获得特征图(feature map)。卷积操作完成后,使用激活函数增加网络的非线性表达能力。卷积层的输出为:

$$C_n^l=f\left(\sum_{i=1}^{m}x_i^{l-1}*w_n^l+b_n^l\right) \tag{1}$$

$$f(x)=\max(0,x) \tag{2}$$

式中:$f(x)$ 为 ReLU 激活函数;x_i^{l-1} 为 $l-1$ 层输出的第 i

个特征图；w_n^l 为 l 层的第 n 个卷积核；b_n^l 为共享偏置；C_n^l 为 l 层的第 n 个特征图经过激活函数映射后的非线性输出。

池化层，池化层主要的作用是特征提取。通过去掉特征图中不重要的样本，进一步减少参数数量，也在一定程度上控制了过拟合。常见的池化层包括最大池化和平均池化等。

全连接层，前一部分的卷积和池化相当于特征工程，后一部分的全连接层则相当于特征加权，在整个卷积神经网络中起到了分类器的作用。

Softmax 层，经过 Softmax 层，可以得到当前样本中属于不同种类的概率分布情况。

1.3　沃瑟斯坦距离

沃瑟斯坦距离（Wasserstein distance，WD）可以衡量任意概率分布之间的差异，是一种平滑的距离度量。设 x_s 和 x_t 的联合概率分布 $\gamma(x_s,x_t)$ 为到 x_s 和 x_t 间的转移成本，则 WD 的定义为：

$$W(P_s,P_t)=\inf_{\gamma\in[P_s,P_t]}\iint\gamma(x_s,x_t)\parallel x_s-x_t\parallel \mathrm{d}x_s\mathrm{d}x_t \quad (3)$$

式(3)的形式在计算时非常困难，因此可以使用其对偶形式：

$$W(P_s,P_t)=\sup_{\|f\|_L\leqslant 1}E_{X_s\sim P_s}[f(x_s)]-E_{X_t\sim P_t}[f(x_t)] \quad (4)$$

式中：$f(\cdot)$ 需满足 1－Lipschitz 连续，即 $\|f(x_s)-f(x_t)\|\leqslant\|x_s-x_t\|$。

WD 的优越性在于它所度量的两个分布无论有没有重叠的部分，依然能反映这两个分布之间的远近，这是 KL 散度和 JS 散度不具备的特性。如图 1 所示，二维空间中的两个分布 P 和 Q，分别在线段 EF 和 CD 上均匀分布，两个分布的距离为参数 θ。

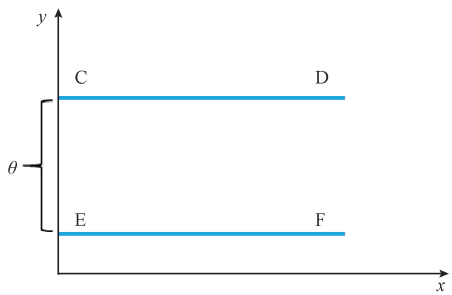

图 1　简单分布示意图

Fig. 1　Simple distribution diagram

分别计算两个分布的 KL 散度，JS 散度和 WD 如下：

$$JS(P\parallel Q)=\begin{cases}\log 2,\theta\neq 0\\0,\theta=0\end{cases} \quad (5)$$

$$KL(P\parallel Q)=KL(Q\parallel P)=\begin{cases}+\infty,\theta\neq 0\\0,\theta=0\end{cases} \quad (6)$$

$$W(P,Q)=|\theta| \quad (7)$$

观察上式发现，WD 是平滑的，在模型中使用 WD 作为损失函数时，通过梯度下降法优化参数，可以提供有效梯度值，使得网络训练收敛。相比而言，在高维空间中，两个分布不重叠或者重叠部分可以忽略时，KL 散度和 JS 散度不能反映分布之间的距离，也提供不了有意义的梯度。

2　子域适应对抗性网络

针对工业故障诊断中无监督变工况数据难以有效诊断的问题，构建一个如下的 SANN 故障迁移诊断框架。模型结构如图 2 所示。该模型分为 3 部分，包括特征提取器，域判别器和分类器。特征提取器由基础 CNN 模型构建，域判别器和分类器分别由 3 层和 2 层的浅层全连接网络构建。在特征提取部分，采用 1DCNN 模型自动从预处理过的信号中提取两域的高维深度特征，其参数由分类器，域判别器和引入的 LMMD 计算模块同时共享和更新。在域判别部分，引入 WD 最大化判别器误差。分类部分利用交叉熵损失（cross entropy loss）最小化源域分类器误差。训练过程需要以预训练为基础，使用源域数据训练特征提取器和分类器，以获取预训练参数，在正式训练时，将预训练参数加载到 SANN 模型中，完成模型参数的初始化。

2.1　特征提取

本文设计特征提取模块为 4 层的一维卷积神经网络，使用 ReLU 函数作为激活函数，对于一维振动信号，为了获取丰富的感受野，更好的提取信号的全局特征，第 1 层卷积层采用大尺寸卷积核[15]。在第 2 和第 4 个卷积层中分别加入最大池化层和自适应池化层进行采样操作，以达到缩小参数矩阵的尺寸的目的，其中自适应池化层输出维度为 4。

为了增强模型的泛化能力，减小模型复杂度，引入批归一层（batch normalization，BN）和 Dropout 层，分别在每个卷积层之后加入 BN 层[16]，可有效避免梯度消失的问题，提高模型训练速度。全连接层的 Dropout[17] 设为 0.5，有效提高模型泛化能力。具体的参数设置如表 1 所示。

表 1　1DCNN 网络结构

Table 1　1DCNN network structure

层类型	核尺寸及数量	步长
卷积层 1	16×15	1
卷积层 2	32×3	1
最大池化层	2	2
卷积层 3	64×3	1
卷积层 4	128×3	1
自适应池化层	—	—
全连接层	256	—

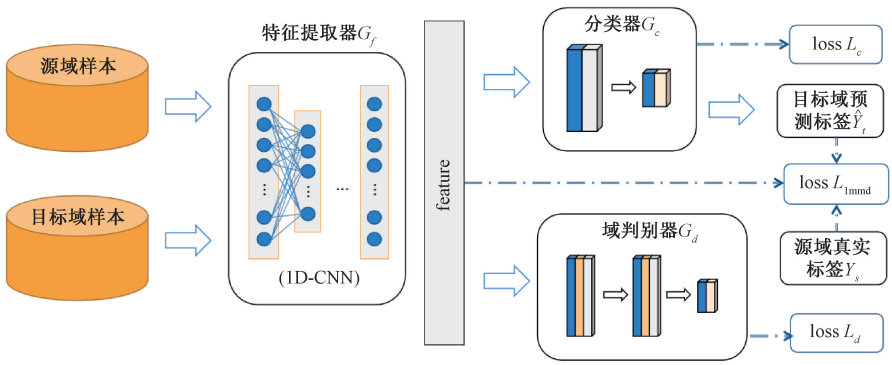

图2　子域适应对抗性网络

Fig. 2　Subdomain Adversarial Neural Network

2.2　基于WD的边缘分布自适应

设来自源域和目标域的样本为 X^s 和 X^t，将两域样本通过网络参数为 θ_f 的特征提取器 r_f 映射为高维特征，生成的源域特征和目标域特征分别为 $h_s = r_f(x_s)$ 和 $h_t = r_f(x_t)$。令 h_s 和 h_t 的边缘分布为 P_s 和 P_t。域判别器 r_d 将源域和目标域的特征映射得到 $r_d(h_s)$ 和 $r_d(h_t)$。两者均值作差如式(8)所示，得到基于WD度量的源域与目标域特征分布差异

$$W(P_s,P_t) = \sup_{\|r_d\|\leqslant 1} E_{h_s\sim P_s}[r_d(h_s)] - E_{h_t\sim P_t}[r_d(h_t)] \tag{8}$$

式中：sup 对所有满足 1-Lipschitz 连续的函数 r_d 有效。此时，WD的经验近似估计如下：

$$L_{wd} = \frac{1}{n_s}\sum_{x^s\in X^s} r_d(r_f(x^s)) - \frac{1}{n_t}\sum_{x^t\in X^t} r_d(r_f(x^t)) \tag{9}$$

训练过程中，r_d 的参数不断更新，以保证源域和目标域的特征的WD最大化，而 r_f 的参数优化方向则是最小化源域和目标域特征的WD，达到边缘分布自适应的目的。

为求解满足1-Lipschitz约束时最大化式(9)的优化问题。Arjovsky 等[18]提出在每次梯度更新后的权值裁剪方法，使参数压缩在一个紧凑的空间内。但是，当裁剪参数较大时，该方法耗时较长，当层数设置过大时，可能会导致梯度消失。因此 Miyato 等[19]使用谱归一化(sepctral normalization，SN)的方法来解决对抗训练的过程中判别器震荡的问题，即在判别器网络中每一层网络用谱范数来代替 BatchNorm 方法进行归一化。

2.3　LMMD计算模块

对抗域适应网络的主要关注点在于边缘分布对齐，而忽略了同一类别中两个子域之间的关系，事实上相同故障类别的数据具有更强的相关性，如图3所示，仅关注边缘对齐进行适配时，子域特征过于接近，而进行条件分布适配后，全局域和子域的特征都实现了分布对齐。

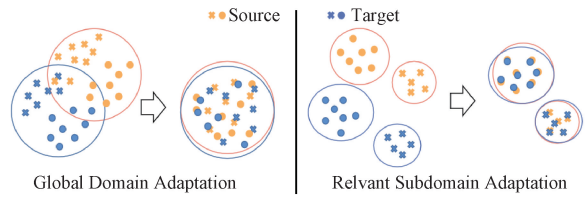

图3　领域适应

Fig. 3　Domain adaptation

本文引入LMMD以解决全局对齐下的细粒度问题，实现同一类别特征分布的对齐，公式可表示为：

$$L_c = \frac{1}{C}\sum_{c=1}^{C}\left\| \sum_{x_i^s\in D_S} w_i^{sc}\varphi(x_i^s) - \sum_{x_j^t\in D_t} w_j^{tc}\varphi(x_j^t)\right\|_{\mathrm{H}}^2 \tag{10}$$

$$w_i^c = \frac{y_{ic}}{\sum_{(x_j,y_j)\in D} y_{jc}} \tag{11}$$

其中，w_i^{sc} 和 w_j^{tc} 分别是属于类别 c 的 x_i^s 和 x_j^t 的权重；c 是故障类别的个数；对于源域中的样本，本文使用真实标签 y_i^s 的独热编码(one-hot)来计算每个样本属于类别 c 的权重 w_i^{sc}；y_{ic} 是独热向量的第 c 项；由于此研究下的目标域无可用标签，本文使用目标域经过分类器的 Softmax 输出作为伪标签，以此来计算 y_i 目标域每个样本属于类别 c 的权重 w_j^{tc}。

设源域和目标域中每个样本特征提取器的输出分别为 z_{si} 和 z_{ti}。利用核函数将 z_{si} 和 z_{ti} 映射到再生核希尔伯特空间，源域与目标域局部最大平均误差 LMMD 可表示为：

$$L_c = \frac{1}{C}\sum_{c=1}^{C}\left[\sum_{i=1}^{n_s}\sum_{j=1}^{n_s} w_i^{sc}w_j^{sc}k(z_i^s,z_j^s) + \sum_{i=1}^{n_t}\sum_{j=1}^{n_t} w_i^{tc}w_j^{tc}k(z_i^t,z_j^t) - 2\sum_{i=1}^{n_s}\sum_{j=1}^{n_t} w_i^{sc}w_j^{tc}k(z_i^s,z_j^t)\right] \tag{12}$$

2.4　训练策略

根据对抗的思想，域判别器 G_d 的目标是尽可能地区

分开特征提取器 G_f 所提取到的样本特征所属的领域,同时 G_f 所提取到的特征要尽量使判别器无法区分其类别,最终达到纳什均衡。在 SANN 模型的训练过程中,可分为两步。第 1 步,固定特征提取器的参数,通过最大化 WD 的距离损失,更新域判别器的参数,第 2 步则是首先固定域判别器的参数,通过最小化 WD 的距离损失,分类损失以及 LMMD 计算模块的类间距离损失,形成对抗局面。整个算法的总优化目标可定义如下:

$$\min_{\theta_c,\theta_f}\{l_c + \lambda_1 l_{lmmd} + \lambda_2 \max_{\theta_d}(l_{wd})\} \tag{13}$$

式中:λ_1 和 λ_2 为权重平衡参数。整个训练过程的伪代码如算法 1 所示。

算法 1 SANN 训练过程:

Require:目标域数据集:X^s 和 X^t;判别器学习率:α_1;分类器和特征提取器学习率:α_2;批次大小:m;判别器训练次数:C;平衡系数:λ_1 和 λ_2。

Training:while θ_f,θ_d 和 θ_c 不收敛

1　for $i=0,\cdots,C$

2　　求特征 $h_s \leftarrow r_f(X^s)$,$h_t \leftarrow r_f(X^t)$

3　　计算 $l_{wd}(X^s,X^t)$

4　　最大化 l_{wd},固定特征提取器参数 θ_f 和分类器参数 θ_c 更新 $\theta_d \leftarrow \theta_d + \alpha_1 \nabla[l_{wd}(X^s,X^t)]$

5　end for

6　　通过伪标签计算 $l_{lmmd} = l_{lmmd}(X^s,X^t,Y^s,\hat{Y}^t)$

7　　最小化 l_c 和 l_{lmmd},固定判别器参数 θ_d

8　　更新 $\theta_c \leftarrow \theta_c - \alpha_2 \nabla[l_c(X^s,X^s) + l_{lmmd}]$

9　　更新 $\theta_f \leftarrow \theta_f - \alpha_2 \nabla[l_c(X^s,X^s) + \lambda_1 l_{wd} + \lambda_2 l_{lmmd}]$

10　end while

3　实　验

本文采用凯斯西储大学(CWRU)和江南大学(JNU)公开轴承故障数据集,共设置 18 组迁移任务,在变负载和变转速的两种变工况条件下,对本文所提模型进行有效性验证。主要框架由 Pytorch 编写。所有实验均在一台装有 i7-9300H 中央处理器和 NVIDIA GeForce GTX 1050 GPU 的计算机上运行。

3.1　案例 1——CWRU 轴承故障数据集

1)数据集描述与预处理

凯斯西储大学(CWRU)轴承数据集是从图 4 实验平台收集而来的。

该实验使用的是型号为 6205-2RS JEM SKF 的齿轮箱高速轴承。该轴承实验分别搭载不同的负载,并在转速为 1 730(3HP)、1 750(2HP)、1 772(1HP)及 1 797(0HP)r/min 时分别采集滚动体、内圈及外圈的振动信号以及正常状态的信号如表 2 所示,损伤直径分别是 7、14 和 21 mils,故障为单点损伤,共 9 种故障状态和一种健康状态,共计 10 种类别。该实验采用加速度传感器对电机的驱动端和风扇端分别进行数据采集,采样频率为 f=12 kHz,每种类型的数据共有 121 991 个数据点。

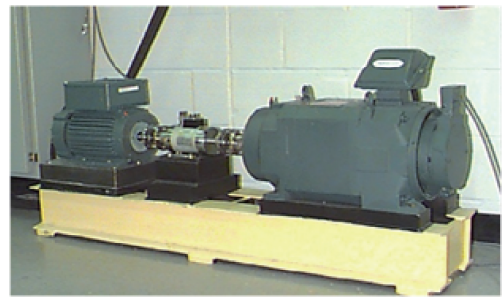

图 4　CWRU 轴承数据集实验台

Fig. 4　CWRU bearing dataset experimental device

表 2　CRWU 数据集工况信息

Table 2　CRWU dataset condition information

任务	0	1	2	3
负载/HP	0	1	2	3
转速/rpm	1 797	1 772	1 750	1 730

该实验中采用驱动端数据,每类状态取 120 000 个采样点,在每类状态中,将所有样本点等分为两份,前一部分用来划分训练样本,后一部分用来划分测试样本,划分训练样本时,为了增加样本数量,使用重叠方式的方式,将样本划分为 200 个样本,每个样本均由 1 024 个数据点构成。而在测试样本划分中不实施重叠,每类得到 50 个样本作为测试样本。根据实验搭载的不同负载划分为 4 种工况,每种工况有 2 000 个训练样本和 500 个测试样本,如表 3 所示。

表 3　CRWU 数据集状态描述

Table 3　CRWU dataset state description

健康状态	严重程度	样本总数	标签
正常状态	0	250	0
内滚道故障	0.007	250	1
滚珠故障	0.007	250	2
外滚珠故障	0.007	250	3
内滚道故障	0.014	250	4
滚珠故障	0.014	250	5
外滚珠故障	0.014	250	6
内滚道故障	0.021	250	7
滚珠故障	0.021	250	8
外滚珠故障	0.021	250	9

2)实验方案设定

本实验共制定了 12 个不同电机负载下的无监督迁移任务。以任务 0→1 为例,代指将模型从工况(0HP)下源域中学习的诊断知识迁移至工况(1HP)下的目标域进行诊断。

设定训练回合数(epochs)为100,对每种方法进行5次测验,记录5次实验平均值。训练时每批次样本数量(BatchSize)为64,其中,模型经多次实验分析,设置学习率 $lr=0.001$,超参数 $\lambda_1=0.1$, $\lambda_2=0.25$, $C=5$,使用RMS optimizer对模型进行优化。

3)迁移诊断结果分析

为了验证提出方法的有效性和泛化性,本实验采用条件域对抗网络(CDAN)[20]、多核最大平均差异(MK-MMD)[21]、域对抗网络(DANN)[11]等领域自适应方法进行比较实验,不同模型诊断精度如表4所示。以上方法均与本文所提模型采用一致的网络结构和参数,以保证对比实验的公平性,并加入1DCNN和WDANN模型进行对比实验,模型说明如下:

CNN(Basis),该模型为只使用有标签的源数据训练所得的模型,并直接在目标测试数据上进行测试。

WDANN,为了验证SAAN模型中组件的有效性,在域对抗模型的基础上,引入WD学习域不变特征,改进模型的边缘对齐。

表4　不同算法的迁移诊断精度

Table 4　Transfer diagnosis accuracy of different algorithms (%)

任务	CNN	CDAN	MK-MMD	DANN	WDANN	SANN
0→1	59.95	96.48	98.41	94.62	98.29	**99.75**
0→2	60.14	98.16	98.69	97.38	99.18	**100**
0→3	50.24	93.25	96.26	87.62	91.48	**98.48**
1→0	68.96	96.89	96.12	92.94	97.32	**100**
1→2	89.03	96.82	99.85	98.97	86.98	**100**
1→3	78.74	98.67	92.26	95.85	98.88	**100**
2→0	69.15	94.38	**97.70**	89.18	91.07	96.43
2→1	81.18	96.40	99.32	97.90	100	**100**
2→3	81.92	96.49	99.28	99.21	**100**	**100**
3→0	64.63	90.21	99.14	87.54	92.56	**99.32**
3→1	75.92	94.38	94.48	92.78	99.17	**99.17**
3→2	79.21	98.70	**100**	99.25	99.42	**100**
AVG	71.58	95.90	97.62	94.36	96.19	**99.42**

由表4可知,CNN的平均识别精度仅仅为71.58%,这是由于不同负载下的CWRU轴承数据集具有不同的数据分布所导致。各类迁移算法在诊断精度上都有了明显得提升,也说明了迁移算法的必要性。其中,MK-MMD和DANN两种算法考虑了全局域特征分布的对齐,平均诊断精度分别达到了97.62%和94.36%。CDAN考虑了特征表示和分类器之间的交叉方差进行特征对齐,平均诊断精度达到了95.90%。WDANN的平均诊断精度达到了96.19%,相较于DANN模型,各项迁移任务的精度有明显提升,说明了其在梯度特性的优越性使得模型更好地完成了分类任务。

SANN在12个迁移任务中都取得了优秀的诊断结果,平均诊断精度达到了99.42%。其中7个任务的诊断精度达到了100%,明显优于对比算法,表明其在边缘对齐的基础上进行条件分布对齐的有效性。且相较于基础域对抗模型DANN,其不仅在精度上提升了5.06%,训练过程中的预测结果的波动也远小于DANN模型,具有较高的鲁棒性。

以任务1→0训练过程为例,目标域数据在SANN、DANN和CNN模型上的测试结果如图5所示,由图可知,相较于DANN,改进后的域对抗模型SANN具有更快的收敛速度,更高的诊断精度以及更好的稳定性。

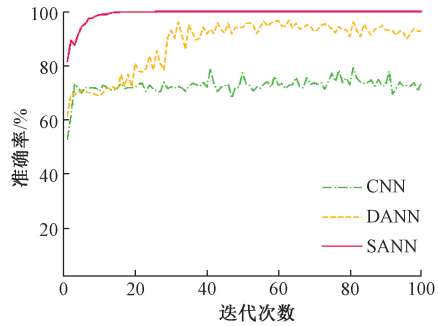

图5　训练过程中目标域测试结果

Fig. 5　Test results on the target domain during training

3.2　案例2——JNU轴承故障数据集

1)数据集描述与预处理

JNU轴承数据集是江南大学采集的一个轴承数据集,在3种不同转速(600、800和1 000 r/min)下,使用采用频率为50 kHz的的加速度计,进行了4种健康状态的实验,包含正常状态、内圈故障、外圈故障和滚动体故障。如表5所示,根据故障类型的不同,将故障数据分为3类。

表5　JNU数据集工况信息

Table 5　JNU dataset condition information

任务	A	B	C
转速/rpm	600	800	100

本实验对时域数据进行采样,在不使用重叠采样方法的情况下,每间隔1 024个点采样一次。其中100%的源域样本和80%的目标域样本作为训练集,目标域剩余的20%的样本作为测试集,表6为单一工况下4种故障状态的相关信息。

表6　JNU数据集状态描述

Table 6　JNU dataset state description

健康状态	样本总数	标签
正常状态	488	0
内圈故障	488	1
外圈故障	488	2
滚动体故障	488	3

2)实验方案设定

本实验共制定了 6 个不同电机负载下的无监督迁移任务。以任务 A→B 为例,代指将模型从工况(600 r/min)下源域中学习的诊断知识迁移至工况(800 r/min)下的目标域进行诊断。

设定训练回合数(epochs)为 300,对每种方法进行 5 次测验,记录 5 次实验平均值。训练时每批次样本数量(BatchSize)为 64,其中,模型经多次实验分析,设置学习率 $lr=0.001$,超参数 $\lambda_1=0.1$, $\lambda_2=0.5$, $C=5$,使用 RMS optimizer 对模型进行优化。

3)迁移诊断结果分析

为了验证提出方法的有效性和泛化性,本实验采用的对比算法与案例一保持一致,实验结果如表 7 所示。

表 7　不同算法的迁移诊断精度

Table 7　Transfer diagnosis accuracy of different algorithms (%)

任务	CNN	CDAN	MK-MMD	DANN	WDANN	SANN
A→B	70.47	86.37	89.38	84.50	85.39	**93.46**
A→C	58.00	90.82	87.40	86.80	**91.93**	91.07
B→A	56.86	92.31	92.42	87.24	91.48	**95.87**
B→C	91.49	97.09	96.85	94.19	97.32	**97.54**
C→A	35.17	85.72	91.23	81.63	86.98	**92.38**
C→B	79.24	97.64	98.78	93.49	98.88	**99.10**
AVG	65.21	91.65	92.67	87.97	91.99	**94.90**

根据表 7,相较于不进行迁移学习的基础 CNN 模型,各类迁移方法在诊断精度上都有了明显的提升,这是由于不同工况下故障数据具有不同的分布,导致域间特征差异明显,从而使得分类精度降低,说明了迁移算法的必要性。SANN 在不同的迁移任务中都取得了较好的诊断结果,平均准确率达到了 94.90%,明显优于其余对比算法,且所有任务的诊断结果都达到了 90%以上,具有较好的鲁棒性。

MK-MMD 在所有任务中也表现中了较为优越的性能,平均准确率达到了 92.67%,但仅考虑了边缘分布,CDAN 同时使用特征和标签进行分布匹配的方法,平均准确率达到了 91.65%。引入 WD 改进的 WDANN,相较于 DANN 模型,各项迁移任务的精度有大幅提升,说明了其在梯度特性的优越性使得模型更好地完成了分类任务。SANN 在边缘分布基础上进行了类内特征的对齐,实现了条件分布自适应,在识别精度上得到了进一步的提高。

为更详细的分析本文每种方法对每个子类别分类精度,以任务 B→A 为例,将目标域未知标签预测结果显示在混淆矩阵中,如图 6 所示,0~3 分别代表了轴承 4 种不同状态。由图 6 可知,本文所提 SANN 模型算法仅有 15 个错误样本,识别正确率达到 96.16%,且故障状态下仅将 2 个滚动体故障样本误识为正常状态,明显高于其他对比算法,说明了该方法的优越性。

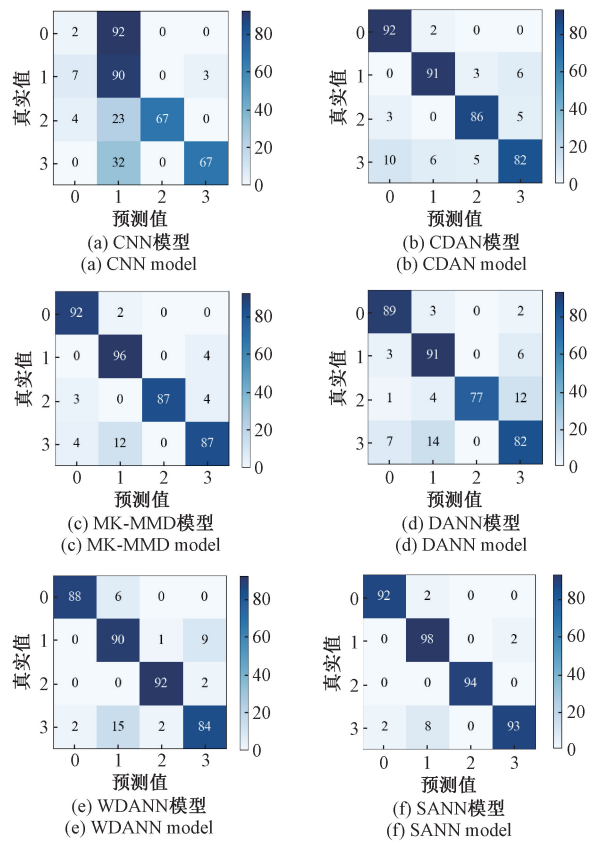

图 6　迁移任务 B→A 混淆矩阵

Fig. 6　Confusion matrix of transfer task

为更直观地分析迁移诊断的结果,采用 t 分布随机邻居嵌入算法(t-distributed stochastic neighbor embedding,t-SNE)算法[22],对任务 B→A 中源域和目标域的数据特征表示进行可视化分析,不同算法的特征降维表示如图 7 所示。

通过图 7(a)可知,由于源域与目标域之间存在的分布差异,使得目标域数据在仅使用源域数据训练的 CNN 模型上的特征映射结果与源域特征存在较大偏差,这也使得模型分类能力下降,识别精度降低。而其余五种算法都在不同程度上对两域特征进行了对齐处理。

由图 7(b)可知,CDAN 模型虽然通过对齐特征和类别的联合分布进行轴承故障健康状态分类,但是对齐效果比较一般。图 7(c)中 MK-MMD 模型在特征的边缘对齐上有明显效果,但各个类别特征都较为接近,特别是正常数据和滚动体故障数据的特征较为接近,使得分类难度增大。与其有同样问题的图 7(e)中的 WDANN 模型,除了正常数据以外,其 3 种故障数据的特征表示都较为接近。值得注意的是,相较于图 7(d)的 DANN 模型,通过引入 WD 度量两域对抗时的边缘分布差异的方式,使

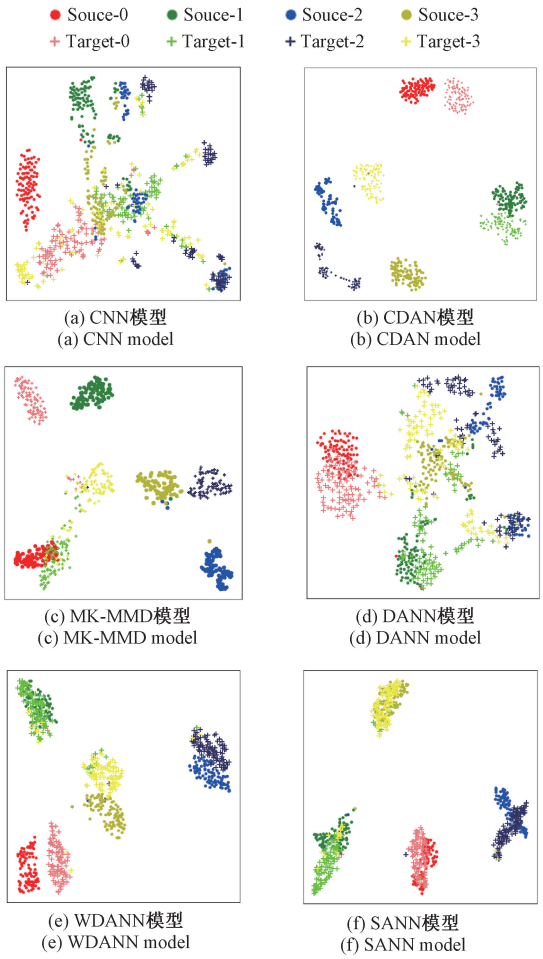

图7　迁移任务B→A特征可视化

Fig. 7　Feature visualization of transfer task

得特征对齐效果明显增强。而DANN模型在全局对齐上有一定效果,但存在如外圈故障特征分散,且各类数据特征过于接近等问题。

由图7(f)可知,本文所提算法SANN模型,在各类特征的边缘分布和条件分布的对齐上都取得了明显的效果,除部分类别混入少数几个其他类别的特征点外,各类别间有明显边际,这也使得在此任务下,SANN模型获得了最高的识别精度。

4　结　论

本文以轴承故障智能诊断的实际工程应用为目标,提出了一种端到端的改进域对抗网络故障迁移诊断模型,解决了实际生产中因缺乏标签数据且工况复杂而导致故障诊断困难,准确率下降的问题,并得如下结论:

1)通过在域对抗网络中引入WD学习特征表示,利用其在梯度特性上的优势和良好的泛化界,可有效地减少了域间特征差异,保证了域间特征的可转移性,提高了网络的鲁棒性,达到了边缘分布自适应的目的。

2)引入局部最大平均偏差LMMD,用于计算同一类别中特征的差异,可在很大程度上缩小特征在条件分布上的差异,采用t-SNE算法可直观的看出两域类间特征对齐的优异性。

3)将所提方法用于变转速、负载两种变工况的轴承故障诊断任务中,并与现有方法如MK-MMD、CDAN和DANN等算法进行对比,实验结果表明,本文所提的模型具有更高的诊断精度和更强的泛化性能。

参考文献

[1] 陈志强,陈旭东,JOSÉ VALENTE DE OLIVIRA,等. 深度学习在设备故障预测与健康管理中的应用[J]. 仪器仪表学报,2019,40(9):206-226.
CHEN ZH Q, CHEN X D, JOSÉ VALENTE DE OLIVIRA, et al. Application of deep learning in equipment prognostics and health management [J]. Chinese Journal of Scientific Instrument, 2019, 40(9): 206-226.

[2] CAI B, HUANG L, XIE M. Bayesian networks in fault diagnosis [J]. IEEE Transactions on Industrial Informatics, 2017, 13(5): 2227-2240.

[3] YIN Z, HOU J. Recent advances on SVM based fault diagnosis and process monitoring in complicated industrial processes[J]. Neurocomputing, 2016, 174: 643-650.

[4] LECUN Y, BENGIO Y, HINTON G. Deep learning[J]. Nature, 2015, 521(11): 436-444.

[5] 李云朋,熊柳景,牛刚. 嵌入式轴承故障诊断系统的设计与实现[J]. 国外电子测量技术,2019,38(11):147-152.
LI Y P, XIONG L J, NIU G. Design and implementation of embedded bearing fault diagnosis system[J]. Foreign Electronic Measurement Technology, 2019, 38(11): 147-152.

[6] 唐波,陈慎慎. 基于深度卷积神经网络的轴承故障诊断方法[J]. 电子测量与仪器学报,2020,34(3):88-93.
TANG B, CHEN SH SH. Method of bearing fault diagnosis based on deep convolutional neural network[J]. Journal of Electronic Measurement and Instrumentation, 2020, 34(3): 88-93.

[7] ZHANG W, LI C H, PENG G L, et al. A deep convolutional neural network with new training methods for bearing fault diagnosis under noisy environment and different working load [J]. Mechanical Systems and Signal Processing, 2018, 100: 439-453.

[8] 王廷轩,刘韬,王振亚,等. 融合迁移卷积神经网络的跨域滚动轴承故障诊断[J]. 电子测量技术,2021,

44(10):167-174.
WANG T X, LIU T, WANG ZH Y, et al. Cross domain rolling bearing fault diagnosis based on migration convolution neural network[J]. Electronic Measurement Technology, 2021,44 (10): 167-174.
[9] GOODFELLOW I J, POUGET-ABADIE J, MIRIRZA M, et al. Generative adversarial nets [J]. Advances in Neural Information Processing Systems, 2014, 27: 2672-2680.
[10] GANIN Y, USTINOVA E, AJAKAN H, et al. Domain-adversarial training of neural networks [J]. Journal of Machine Learning Research, 2017,17(1):2096-2030.
[11] SHAO J J, HUANG Z W, ZHU J M. Transfer learning method based on adversarial domain adaption for bearing fault diagnosis [J]. IEEE Access, 2020, 8: 119421-119430.
[12] CHEN Z Y, HE G L, LI J P, et al. Domain adversarial transfer network for cross-domain fault diagnosis of rotary machinery[J]. IEEE Transactions on Instrumentation and Measurement, 2020, 69(11): 8702-8712.
[13] GULRAJANI I, AHMED F, ARJOVSKY M, et al. Improved training of wasserstein gans[J]. Advances in Neural Information Processing Systems, 2017, 30.
[14] HU Y, ZHUANG F, WANG J, et al. Deep subdomain adaptation network for image classification[J]. IEEE Transactions on Neural Networks and Learning Systems, 2020 (99): 1-10.
[15] 张伟. 基于卷积神经网络的轴承故障诊断算法研究[D]. 哈尔滨:哈尔滨工业大学, 2017.
ZHANG W. Study on bearing fault diagnosis algorithm based on convolutional neural network [D]. Harbin : Harbin Institute of Technology, 2017.
[16] IOFFE S, SZEGEDY C. Batch normalization: Accelerating deep network training by reducing internal covariate shift [J]. arXiv preprint arXiv: 1502.03167, 2015.
[17] SRIVASTAVA N, HINTON G, KRIZHEVSKY A, et al. Dropout: A simple way to prevent neural networks from overfitting[J]. Journal of Machine Learning Research, 2014, 15(1): 1929-1958.
[18] ARJOVSKY M, BOTTOU L, et al. Towards principled methods for training generative adversarial networks[C]. arXiv preprint arXiv:1701.04862, 2017.
[19] MIYATO T, KATAOKA T, KOYAMA M, et al. Spectral normalization for generative adversarial networks [C]. arXiv preprint arXiv:1802.05957, 2018.
[20] LONG M, CAO Z, WANG J, et al. Conditional adversarial domain adaptation[C]. Advances in Neural Information Processing Systems, 2018,31: 1640-1650.
[21] LONG M, CAO Y, CAO Z, et al. Transferable representation learning with deep adaptation networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2019,41(12): 3071-3085.
[22] LAURENS V D M, HINTON G. Visualizing data using t-SNE[J]. Journal of Machine Learning Research, 2008, 2605(9): 2579-2605.

作者简介

张金汇,2020 年毕业于南京工程学院自动化专业,获得学士学位,现为江南大学电子信息专业硕士研究生,主要研究方向为深度学习与故障诊断。
E-mail: 1319101774@qq.com

Zhang Jinhui received a B. Sc. degree in automation from Nanjing Institute of Technology in 2020. Now he is a M. Sc. candidate in electronic information of Jiangnan University. His main research interests include deep learning and fault diagnosis.

谢林柏(通信作者),2004 年获华中科技大学控制理论与控制工程专业工学博士学位,现为江南大学物联网工程学院教授,主要研究方向为过程建模与控制、智能检测与系统安全性。
E-mail: xie_linbo@jiangnan.edu.cn

Xie Linbo (Corresponding author) received a B. Sc. degree in control theory and control engineering from Huazhong University of Science and Technology in 2004. Now he is a professor of the School of Internet of Things Engineering of Jiangnan University. His main research interests include process modeling and control, intelligent detection and system safety.