

11-30-2023

Rolling Bearing Fault Diagnosis Based on Weighted Domain Adaptive Convolutional Neural Network

Wenfeng Zhang

Engineering Research Center of Internet of Things Technology Applications Ministry of Education, Jiangnan University, Wuxi 214122, China, 6211924137@stu.jiangnan.edu.cn

Zhichao Zhu

Engineering Research Center of Internet of Things Technology Applications Ministry of Education, Jiangnan University, Wuxi 214122, China

Dinghui Wu

Engineering Research Center of Internet of Things Technology Applications Ministry of Education, Jiangnan University, Wuxi 214122, China, wdh123@jiangnan.edu.cn

Follow this and additional works at: <https://dc-china-simulation.researchcommons.org/journal>



Part of the Artificial Intelligence and Robotics Commons, Computer Engineering Commons, Numerical Analysis and Scientific Computing Commons, Operations Research, Systems Engineering and Industrial Engineering Commons, and the Systems Science Commons

This Paper is brought to you for free and open access by Journal of System Simulation. It has been accepted for inclusion in Journal of System Simulation by an authorized editor of Journal of System Simulation. For more information, please contact xtfzxb@126.com.

Rolling Bearing Fault Diagnosis Based on Weighted Domain Adaptive Convolutional Neural Network

Abstract

Abstract: A rolling bearing fault diagnosis method based on a weighted domain adaptive convolutional neural network (WDACNN) is proposed to solve the problem that the data distribution of vibration signals of rolling bearings changes due to workload changes, which leads to poor generalization of fault diagnosis algorithm. In this method, the domain adaptation algorithm is embedded in the convolutional neural network to make the classifier based on the source domain achieve excellent generalization in the target domain, and the weight coefficient is introduced to weight the samples in the source domain to reduce the influence of the class weight deviation. In the simulation experiment, six migration tasks are used to verify the effectiveness of the proposed method. The average fault diagnosis accuracy of the proposed method reaches 96.6%, which proves the effectiveness of the proposed method under different workload conditions.

Keywords

bearing, fault diagnosis, transfer learning, domain adaptation, weight coefficient

Recommended Citation

Zhang Wenfeng, Zhu Zhichao, Wu Dinghui. Rolling Bearing Fault Diagnosis Based on Weighted Domain Adaptive Convolutional Neural Network[J]. Journal of System Simulation, 2023, 35(11): 2445-2453.

基于加权域适应卷积神经网络的滚动轴承故障诊断

张文锋, 祝志超, 吴定会*

(江南大学 物联网应用技术教育部工程研究中心, 江苏 无锡 214122)

摘要: 针对工作负载变化引起滚动轴承振动信号的数据分布变化, 导致故障诊断算法的泛化性变差的问题, 提出一种基于加权域适应卷积神经网络 (weighted domain adaptive convolutional neural network, WDACNN) 的滚动轴承故障诊断方法。在卷积神经网络中嵌入域适应算法, 使基于源域的分类器在目标域上取得良好的泛化性, 并引入权重系数对源域样本进行加权, 降低类权重偏差的影响。采用 6 组迁移任务验证所提方法的有效性, 结果表明: 平均故障诊断精度达到 96.6%, 证明了本文方法在不同工作负载条件下的有效性。

关键词: 轴承; 故障诊断; 迁移学习; 域适应; 权重系数

中图分类号: TP391.9 文献标志码: A 文章编号: 1004-731X(2023)11-2445-09

DOI: 10.16182/j.issn1004731x.joss.22-0616

引用格式: 张文锋, 祝志超, 吴定会. 基于加权域适应卷积神经网络的滚动轴承故障诊断[J]. 系统仿真学报, 2023, 35(11): 2445-2453.

Reference format: Zhang Wenfeng, Zhu Zhichao, Wu Dinghui. Rolling Bearing Fault Diagnosis Based on Weighted Domain Adaptive Convolutional Neural Network[J]. Journal of System Simulation, 2023, 35(11): 2445-2453.

Rolling Bearing Fault Diagnosis Based on Weighted Domain Adaptive Convolutional Neural Network

Zhang Wenfeng, Zhu Zhichao, Wu Dinghui*

(Engineering Research Center of Internet of Things Technology Applications Ministry of Education, Jiangnan University, Wuxi 214122, China)

Abstract: A rolling bearing fault diagnosis method based on a weighted domain adaptive convolutional neural network (WDACNN) is proposed to solve the problem that the data distribution of vibration signals of rolling bearings changes due to workload changes, which leads to poor generalization of fault diagnosis algorithm. In this method, the domain adaptation algorithm is embedded in the convolutional neural network to make the classifier based on the source domain achieve excellent generalization in the target domain, and the weight coefficient is introduced to weight the samples in the source domain to reduce the influence of the class weight deviation. In the simulation experiment, six migration tasks are used to verify the effectiveness of the proposed method. The average fault diagnosis accuracy of the proposed method reaches 96.6%, which proves the effectiveness of the proposed method under different workload conditions.

Keywords: bearing; fault diagnosis; transfer learning; domain adaptation; weight coefficient

0 引言

滚动轴承是机械设备中重要的传动部件, 一

个微小的缺陷就会影响整个机器的性能, 甚至导致灾难的发生, 因此, 开发滚动轴承智能故障诊

收稿日期: 2022-06-06 修回日期: 2022-08-10

基金项目: 国家重点研发计划(2020YFB1711102)

第一作者: 张文锋(1998-), 男, 硕士生, 研究方向为智能故障诊断。E-mail: 6211924137@stu.jiangnan.edu.cn

通讯作者: 吴定会(1970-), 男, 教授, 博士, 研究方向为智能故障诊断。E-mail: wdh123@jiangnan.edu.cn

断工具对机械设备的稳定运行具有重要意义。

近年来,机器学习在故障诊断领域得到广泛应用,并取得了杰出的效果^[1-4]。其中,传统机器学习算法需要人为设计特征,易造成信息的冗余或丢失^[1]。为此,基于深度学习的故障诊断方法得到快速发展,如卷积神经网络(convolutional neural networks, CNN)^[2]、深度置信网络(deep belief networks, DBN)^[3]、长短时记忆(long short-term memory, LSTM)^[4],其能够自动提取振动信号的故障特征、增强非线性表达能力和泛化能力。

然而,深度学习通常假设源域和目标域具有相同的数据分布,这在实际中很难满足。一是实际工业中缺乏足够的用作测试集的故障数据;二是机械设备的工作负载多变,导致训练集和测试集数据分布不一致,影响模型的性能^[5]。

迁移学习^[6]能够将源域提取到的知识应用到目标域,有望解决上述问题。近年来,迁移学习方法被广泛应用于故障诊断任务^[7-9]。文献[7]构建了深度卷积迁移学习网络(deep convolutional transfer learning networks, DCTLN),解决在数据分布不同条件下的机械故障诊断问题。文献[8]提出了多层自适应卷积神经网络(multilayer adaptive convolutional neural network, MACNN),实现轴承在不同工作负荷下的故障诊断。文献[9]提出了一种基于条件数据对齐和无监督预测一致性的自适应局部域算法,解决机械故障诊断中存在的局部域适应问题。这些方法将深度学习网络与最大均值差异(maximum mean discrepancy, MMD)相结合提取域不变特征,但忽略了设备工作负载变化引起的数据类权重(即类先验分布)的变化,影响域适应效果。

为解决上述问题,本文提出基于加权域适应卷积神经网络(weighted domain adaptation convolutional neural network, WDACNN)的滚动轴承故障诊断方法。该方法通过最小化源域和目标域之间的加权最大均值差异(weighted maximum mean

discrepancy, WMMD)进行域适应,其中WMMD采用权重系数为源域样本进行加权,使加权后的源域和目标域保持相同的类权重,从而降低类权重偏差的影响。并将CNN提取到的特征输入到Softmax分类器进行分类,实现故障诊断。

1 域适应理论

1.1 域适应问题

域适应是一种基于特征的无监督迁移学习算法(图1),“域”的定义如下:给定样本空间 \mathcal{X} , $\mathbf{X}=\{x_1, x_2, \dots, x_n\} \in \mathcal{X}$ 为 \mathcal{X} 中的数据集, $p(\mathbf{X})$ 为 \mathbf{X} 的边缘概率分布,则域被定义为 $\mathcal{D}=\{\mathcal{X}, p(\mathbf{X})\}$ 。迁移学习中的域被分为源域和目标域,源域和目标域具有不同的数据分布,其中源域带标签,记为 $\mathcal{D}_s=\{\mathbf{x}_i^s, y_i^s\}_{i=1}^n$;而目标域不带标签,记为 $\mathcal{D}_t=\{\mathbf{x}_j^t\}_{j=1}^n$ 。未对源域和目标域进行域适应时,源域的分类器在源域上具有很好的分类效果,但在目标域上的分类效果不佳,如图1(a)所示。域适应算法可以为源域和目标域找到一个公共的特征提取器 $f(\cdot)$,在源域和目标域样本空间上得到特征空间 \mathcal{Z}^s 和 \mathcal{Z}^t ,使得源域分类器在目标域特征空间 \mathcal{Z}^t 能够取得良好的泛化性能。

1.2 最大均值差异

在域适应问题中,通常采用MMD度量源域和目标域特征之间的距离。

假设 \mathbf{X}^s 和 \mathbf{X}^t 分别为源域数据集和目标域数据集, s 和 t 分别为源域和目标域的数据分布,则 s 和 t 之间的最大均值差异可表示为

$$D_{\text{mmd}}(s, t) \triangleq \left\| E_s(\phi(\mathbf{X}^s)) - E_t(\phi(\mathbf{X}^t)) \right\|_{\mathcal{H}}^2 \quad (1)$$

式中: \mathcal{H} 为再生希尔伯特空间; $E_s(\cdot)$ 为数据分布的期望; $\phi(\cdot)$ 为将样本映射至与核 $k_\phi(\mathbf{X}^s, \mathbf{X}^t) = \langle \phi(\mathbf{X}^s), \phi(\mathbf{X}^t) \rangle$ 相关联的希尔伯特空间。

在样本量足够大的情况下,式(1)的经验估计公式可以表示为

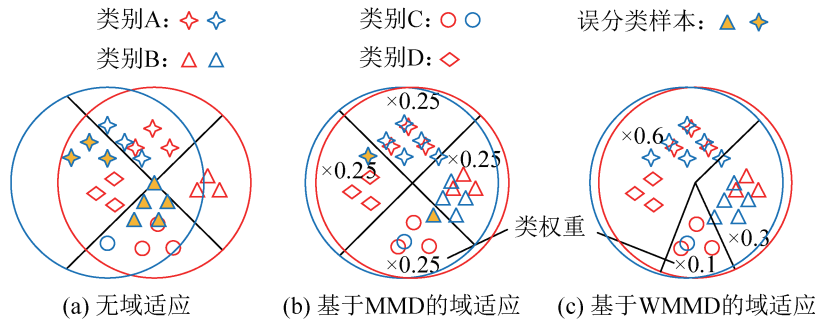


图 1 不同方法的迁移学习效果

Fig. 1 Transfer learning effect of different methods

$$\hat{D}_{\text{mmd}}(s, t) \triangleq \left\| \frac{1}{n_s} \sum_{i=1}^{n_s} \phi(x_i^s) - \frac{1}{n_t} \sum_{j=1}^{n_t} \phi(x_j^t) \right\|_{\mathcal{H}}^2 \quad (2)$$

式中: n_s 和 n_t 分别为源域和目标域的样本数量。

1.3 加权最大均值差异

在无监督迁移学习中源域和目标域类权重(即类先验分布)很难保持一致, 这种情况被称为类权重偏差^[10]。然而 MMD 往往忽视了类权重偏差的影响, 以图 1(b)为例, 图中 D 类样本在目标域中的类权重为 0, 因此从源域中 D 类样本学习到的知识对于目标域而言是无用的, 这部分知识会对目标域的特征提取造成干扰, 最终影响迁移学习的效果。

为解决这个问题, 本节引入加权最大均值差异, 其效果如图 1(c)所示。不同于 MMD 直接计算源域和目标域分布差异, WMMD 首先将源域映射为参考源分布 $p_{a,s}(X^s)$, 然后计算参考源分布与目标域之间的分布差异, WMMD 和 $p_{a,s}(X^s)$ 的定义分别为

$$D_{\text{wmmd}}(s, t) \triangleq \left\| E_{(a,s)}(\phi(X^s)) - E_t(\phi(X^t)) \right\|_{\mathcal{H}}^2 \quad (3)$$

$$p_{a,s}(X^s) = \sum_{c=1}^C w_c^t p_s(X^s | Y^s = c) = \sum_{c=1}^C \alpha_c w_c^s p_s(X^s | Y^s = c) \quad (4)$$

式中: $\alpha_c = w_c^t / w_c^s$ 为权重系数; C 为样本类别数目; c 为其中一个标签。

从式(3)~(4)可知, 利用权重系数 α_c 为源域样本加权, 得到的参考源分布能够和目标域保持相

同的类权重, 因此可以有效降低类权重偏差的影响。WMMD 的经验估计公式可写为

$$\hat{D}_{\text{wmmd}}(s, t) = \left\| \frac{1}{\sum_{i=1}^{n_s} \alpha_{y_i^s}} \sum_{i=1}^{n_s} \alpha_{y_i^s} \phi(x_i^s) - \frac{1}{n_t} \sum_{j=1}^{n_t} \phi(x_j^t) \right\|_{\mathcal{H}}^2 \quad (5)$$

式中: $y_i^s = 1, 2, \dots, i$ 。

接下来对权重系数进行估算。由于源域样本的标签已知, 可直接采用公式 $w_c^s = n_c^s / n_s$ 计算类先验概率, 其中 n_c^s 为源域中第 c 类样本的数量。然而目标域缺少真实标签, 无法直接计算类先验概率, 因此采用卷积神经网络输出的伪标签进行替代, 目标域的先验概率为

$$\hat{w}_c^t = \frac{1}{n_t} \sum_j 1_c(\hat{z}_j^t) \quad (6)$$

式中: n_t 为目标域样本数量; \hat{z}_j^t 为目标域样本 x_j^t 通过深度神经网络分类得出的伪标签; $1_c(\cdot)$ 为一个指示变量, 当输入变量为 c 时, 其值为 1, 其余情况为 0。

2 基于 WDACNN 的滚动轴承故障诊断

随机变化的风力使得风电机轴承工作负载发生变化, 并引起类权重偏差问题, 本文提出 WDACNN 的滚动轴承故障诊断方法, 在卷积神经网络中嵌入 WMMD 损失项, 并通过最小化该损失项来减小负载变化引起的数据分布差异, 实现域适应。

2.1 加权域适应卷积神经网络

基于 WDACNN 的故障诊断框架如图 2 所示，由数据预处理、特征提取、故障识别 3 个模块组成。数据预处理模块将滚动轴承的原始振动信号转换为小波包系数矩阵，特征提取模块由 2 个共享参数的 CNN 组成，用于提取两个域的域不变特征，故障识别模块由 Softmax 分类器实现，可将学习到的特征进行分类识别。

(1) 数据预处理模块: 由于滚动轴承在不同负载下的振动信号通常是非平稳信号，采用时频分析进行预处理将有利于 CNN 进行特征提取。在常用的几种时频分析方法中，小波包变换不仅可以分解低频信号，对于高频信号也具有较强的分析能力^[11]。本文采用小波包变换对原始振动信号进行预处理，以下对小波包原理进行介绍。

给定待分解振动序列 $s(t)$, $\{h_k\}_{k \in Z}$ 和 $\{g_k\}_{k \in Z}$ 分别为一组低通、高通共轭正交滤波器，则小波包系数的递推公式为

$$d_k^{j,2^{n-1}} = \sum_{k \in Z} h_{l-2k} d_l^{j-1,n} \tag{7}$$

$$d_k^{j,2^n} = \sum_{k \in Z} g_{l-2k} d_l^{j-1,n} \tag{8}$$

式中: $d_l^{j-1,n}$ 为第 $j-1$ 层小波包变换系数序列中的第 n 个系数; k 为第 k 个卷积核。

采用递推公式将振动序列分解为多个频带，每个频带具有多个小波包系数。设 u 为样本长度， m 为小波包变换深度，则分解后可得到 2^m 个频带，每个频带包含 $u/2^m$ 个小波包系数，将不同频带上的小波包系数组成小波包系数矩阵，作为下一个模块的输入。

(2) 特征提取模块: 小波包系数矩阵包含了大量的时频信息，但仍然属于浅层特征表示，需要利用 CNN 进一步提取深层特征。以下对 CNN 的基本原理进行介绍。

在卷积神经网络中，卷积层是特征提取的关键，其包含多个卷积核，它们以滑动窗口的方式从输入数据中提取特征^[12]。假设 CNN 网络第 l 层即视为卷积层，其后层作为池化层。第 l 层输出的第 d 个特征图的第 i 行、第 j 列的元素输出计算公式为

$$a_{i,j}^{d,l} = \sigma \left(\sum_{k=1}^K \sum_{d=0}^{D-1} \sum_{p,q=0}^{S-1} w^{k,l} \hat{x}_{i+p,j+q}^{d,l} + b^{k,l} \right) \tag{9}$$

式中: $\hat{x}_{i+p,j+q}^{d,l}$ 为第 l 层中输入的第 d 个特征图的第 $i+p$ 行、第 $j+q$ 列中的元素; $w^{k,l}$ 为第 l 层的第 k 个卷积核的权重; $b^{k,l}$ 为卷积核对应偏置项; D 为输入特征图的个数; S 为卷积核的尺寸; K 为卷积核的尺寸; $\sigma(\cdot)$ 为激活函数，本文采用 ReLU 作为卷积层的激活函数。

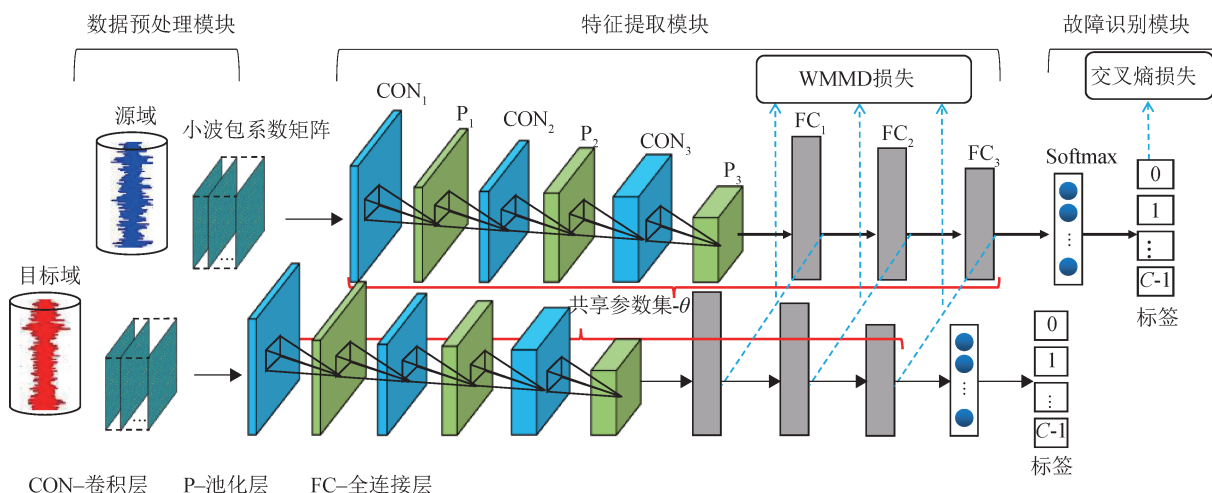


图2 基于 WDACNN 的故障诊断框架
Fig. 2 Fault diagnosis framework based on WDACNN

为了减少网络中参数的数量, 减少过拟合风险, 通常使用池化层对输入特征进行降采样^[13]。本文采用的采样方法是最大池化

$$y_{i,j} = \text{pool}_{(m,n) \in R_{i,j}}(\mathbf{x}_{m,n}) \quad (10)$$

式中: \mathbf{x} , \mathbf{y} 分别为池化前后的特征图; $R_{i,j}$ 为特征图中进行池化操作的区域。

经过若干个卷积层和池化层后, 连接3个全连接层, 全连接层的主要作用是非线性特征提取, 相邻全连接层之间的传递公式为

$$\mathbf{h} = \delta(\mathbf{w}_h \mathbf{x} + \mathbf{b}_h) \quad (11)$$

式中: \mathbf{x} , \mathbf{h} 分别为输入和输出的特征; \mathbf{w}_h , \mathbf{b}_h 分别为全连接层的权重和偏置; $\delta(\cdot)$ 为激活函数, 本文采用 Sigmoid 作为激活函数。

文献[14]指出, 随着网络层数的提升, 卷积神经网络学习到的特征具有更加明显的任务特性。因此较之层数较低的卷积层和池化层, 全连接层学习到的域特征更有利于分类器的学习。本文在全连接层中加入 WMDD 度量域特征的分布距离, 并通过最小化 WMMD 损失项使特征分布距离逐渐缩小, 从而实现域适应。由式(5)可知, 将 $\mathbf{s}_l, \mathbf{t}_l$ 代入至 WMMD 的经验估计公式, 第 l 层全连接层处的 WMMD 损失项可以表示为

$$\hat{D}_{\text{wmmd}}^l = \hat{D}_{\text{wmmd}}(\mathbf{s}_l, \mathbf{t}_l) \quad (12)$$

式中: $\mathbf{s}_l, \mathbf{t}_l$ 分别为源域和目标域在第 l 层全连接层输出的特征。

(3) 故障识别模块: 模块的主要作用是利用 Softmax 分类器对学习到的域特征进行分类^[15]。已经获得了源域和目标域的特征, 通过 softmax 对特征进行分类来估计轴承的健康状态, 样本的预测概率 y_i 的计算公式为

$$y_i = \frac{1}{\sum_{c=1}^C e^{h^{3,c}}} [e^{h^{3,1}} \quad e^{h^{3,2}} \quad \dots \quad e^{h^{3,C}}]^T \quad (13)$$

式中: $h^{3,c}$ 为卷积神经网络中的第3个全连接层在第 c 个神经元的输出; $c=1, 2, \dots, C$, C 为轴承健康状况类别的数量。

样本的伪标签为

$$\hat{z}_i = \arg \max_c (y_i) \quad (14)$$

2.2 目标函数及训练策略

WDACNN 在训练过程中主要通过最小化目标函数实现参数更新, 因此, 首先需要构建目标函数。为了能够正确识别源域样本的类别, 通常将源域的交叉熵损失项作为目标函数的一部分进行训练, 使输出值与真实值尽可能地接近^[16]。交叉熵损失函数的公式为

$$\mathcal{L}_c = -\frac{1}{n_s} \sum_{i=1}^{n_s} \sum_{c=1}^M r_{ic}^s \log(y_{ic}^s) \quad (15)$$

式中, n_s 为源域样本的数量; M 为类别的数量; r_{ic}^s 为一个符号函数; 当源域样本 \mathbf{x}_i^s 的标签为 c 时, r_{ic}^s 的值为 1, 其他情况为 0; y_{ic}^s 为样本 \mathbf{x}_i^s 属于类别 c 的概率。

为了提高 WDACNN 在负载变化下的故障诊断能力, 减小类权重偏差的影响, 需要最小化 WMDD 损失项来减小源域和目标域之间的特征分布差异, 以提取二者的域不变特征。因此将基于 WMMD 嵌入到全连接层, 得到如下 WMMD 损失项:

$$\mathcal{L}_d = \sum_{l=1}^{l_3} D_{\text{wmmd}}^l \quad (16)$$

将两个损失项进行合并, 得到最终的目标函数为

$$f_{\text{wmdd}} = \min_{\theta} \mathcal{L}_c + \lambda \mathcal{L}_d \quad (17)$$

式中: λ 为控制 \mathcal{L}_d 重要程度的超参数。

得到目标函数之后, 采用小批量随即梯度下降算法更新模型参数, 更新为

$$v \leftarrow \mu v + \left(\frac{\partial \mathcal{L}_c}{\partial \theta} + \lambda \frac{\partial \mathcal{L}_d}{\partial \theta} \right), \quad \theta \leftarrow \theta - \eta v$$

其中: μ , v 和 η 分别为动量, 速度和学习率; θ 为模型参数。

2.3 基于 WDACNN 的故障诊断步骤

本文提出的基于 WDACNN 的滚动轴承故障诊断步骤如图3所示。

(1) 将两种不同负载下采集到的滚动轴承振动信号划分为源域和目标域，其中源域的样本具有类别标签，目标域的样本没有类别标签；

(2) 按照数据预处理模块、特征提取模块、故障识别模块的顺序构建 WDACNN，并初始化网络中的参数；

(3) 将源域样本和目标域样本输入到 WDACNN 中进行训练，训练过程中 WDACNN 的参数不断更新，当循环迭代次数最大且目标函数收敛时训练完成；

(4) 将目标域样本输入到训练好的 WDACNN 中，利用其输出标签得到目标域样本的故障类别，实现故障诊断。

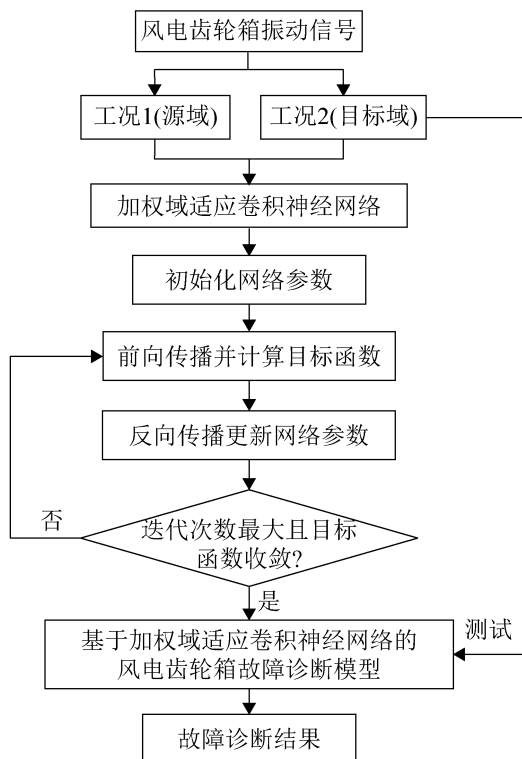


图 3 基于 WDACNN 的故障诊断步骤

Fig. 3 Troubleshooting steps based on WDACNN

3 仿真实验分析

为了验证 WDACNN 的有效性，将 WDACNN 与 5 种对比算法应用在滚动轴承振动数据集上进行仿真实验。

3.1 数据集描述

滚动轴承振动数据集：主要通过振动传感器采集不同位置处的振动信号，采样频率为 12 KHz。电机在 3 种不同的负载条件下(1 马力、2 马力、3 马力)运行，本文分别采用 A, B, C 来标记。

每种负载下包含 4 种健康状态的振动信号，即滚动体故障，内圈故障，外圈故障，正常。

为确保每个样本包含足够的故障信息，将样本长度设置为 4 096。每次选择一种负载下的数据集为源域，另一种负载下的数据集为目标域，可构成 6 个不同迁移故障诊断任务，分别标记为 A→B, A→C, B→A, B→C, C→A 和 C→B。箭头左边为源域，右边为目标域。为每个子数据集设置不同的类权重偏差，训练集和测试集描述如表 1 所示，其中训练样本均为 600 个。图 4 展示了轴承滚动体故障在不种负载条件下的振动信号波形。

表 1 轴承数据集描述

Table 1 Description of bearing data set			
健康状态	标签	工作负载	测试样本
滚动体故障	0	A	320
		B	300
		C	280
内圈故障	1	A	300
		B	280
		C	260
外圈故障	2	A	300
		B	280
		C	320
正常	3	A	260
		B	320
		C	300

3.2 实验设置

为了验证 WDACNN 的优越性，本文将 WDACNN 与其他 5 种算法进行比较。分别为误差反向传播学习神经网络(back propagation, BP)^[17]，LSTM，CNN，深度域适应网络(deep adaptation network, DAN)^[18]，MACNN^[19]。其中，BP、LSTM 和 CNN 未进行域适应，DAN 和 MACNN 均通过最小化 MMD 损失实现域适应。为保证对比实验

的公平性, 所有算法的输入信号均采用小波包变换进行预处理, 其中CNN, DNN, MACNN和WDACNN算法都采用表2所示的CNN结构, 最终的故障诊断精度取10次测试的平均值。

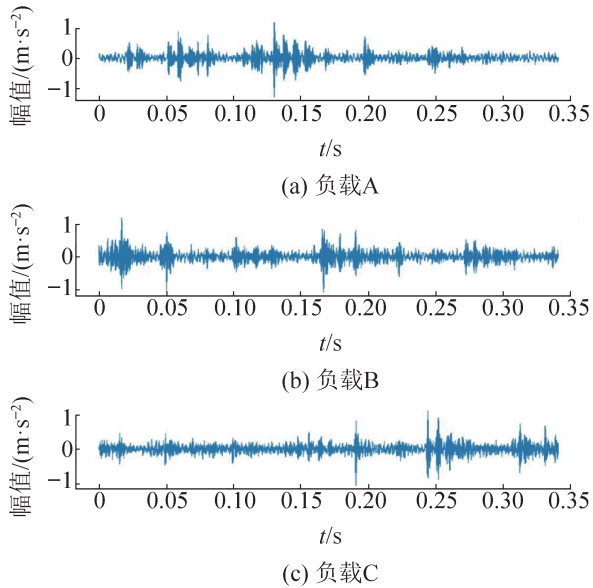


图4 轴承健康状态的振动信号波形

Fig. 4 Vibration signal waveform of bearing health status

表2 CNN网络结构
Table 2 CNN network structure

层数	类型	核数量	核尺寸	步长	激活函数
1	卷积层	1	3×3	2	ReLU
2	池化层	—	2×2	2	—
3	卷积层	8	3×3	1	ReLU
4	池化层	—	2×2	2	—
5	卷积层	16	3×3	1	ReLU
6	池化层	—	2×2	2	—
7	全连接层	1	800	—	Sigmoid
8	全连接层	1	100	—	Sigmoid
9	全连接层	1	20	—	Sigmoid
10	输出层	1	4	—	Softmax

注: “—”表示不存在此项。

所有任务均采用小批量随机梯度下降算法更新参数^[20], 动量参数设置为0.9, 学习率采用以下公式进行更新: $\eta = \eta_0 / (1 + \alpha r)^\beta$, 其中 r 随迭代次数从0到1线性增长, η_0 设为0.01, α 和 β 分别设置为10和0.75。

3.3 结果分析

滚动轴承振动数据集的故障诊断结果如表3所示。在所有方法中, BP, LSTM和CNN均没有采用迁移学习策略, 在测试集上的故障诊断精度都相对较低; WDACNN的故障诊断精度达到最高, 平均故障诊断精度为96.6%, 分别比BP, LSTM, CNN, DAN和MACNN高16.8%, 13.3%, 8.1%, 4.5%和2.0%。因此, 通过诊断结果表明, 与其他算法相比, WDACNN具有更好的迁移故障诊断性能。通过观察对比结果, 本文还发现以下两点信息:

(1) WMMD能够更好地学习域的域不变特征, 从而获取更高的故障诊断精度。在不同对比算法中, DAN采用MMD作为损失项, WDACNN采用WMMD作为损失项, 其余部分都相同, 但是DAN平均故障诊断精度明显低于WDACNN。

(2) 测试集上的故障诊断精度一定程度上取决于源域和目标域的分布差异大小。例如采用MACNN算法时, A→C和C→A的准确率分别为93.9%和91.5%, 在所有任务中的准确率中相对偏低。

表3 故障诊断结果
Table 3 Fault diagnosis results %

诊断任务	BP	LSTM	CNN	DAN	MACNN	WDACNN
AB	85.8	87.3	92.9	96.2	97.8	99.1
AC	81.5	82.3	88.2	91.3	93.9	96.5
BA	79.6	86.1	91.3	94.1	96.2	97.1
BC	77.9	81.8	86.1	92.1	95.4	98.9
CA	77.5	79.6	84.6	88.3	91.5	93.2
CB	76.7	82.6	88.1	89.9	92.9	95.3
AVG	79.8	83.3	88.5	92.1	94.6	96.6

注: “AB”表示“A→B”的迁移任务, “AVG”表示均值。

为了直观地显示源域和目标域之间的特征分布, 本文采用用 t 分布邻域嵌入(t-SNE)^[21]将最后一层全连接层学习到的特征进行可视化。

图5为不同算法在数据集任务B→C中的特征可视化图, 横纵坐标分别为特征向量的前两个主元特征。

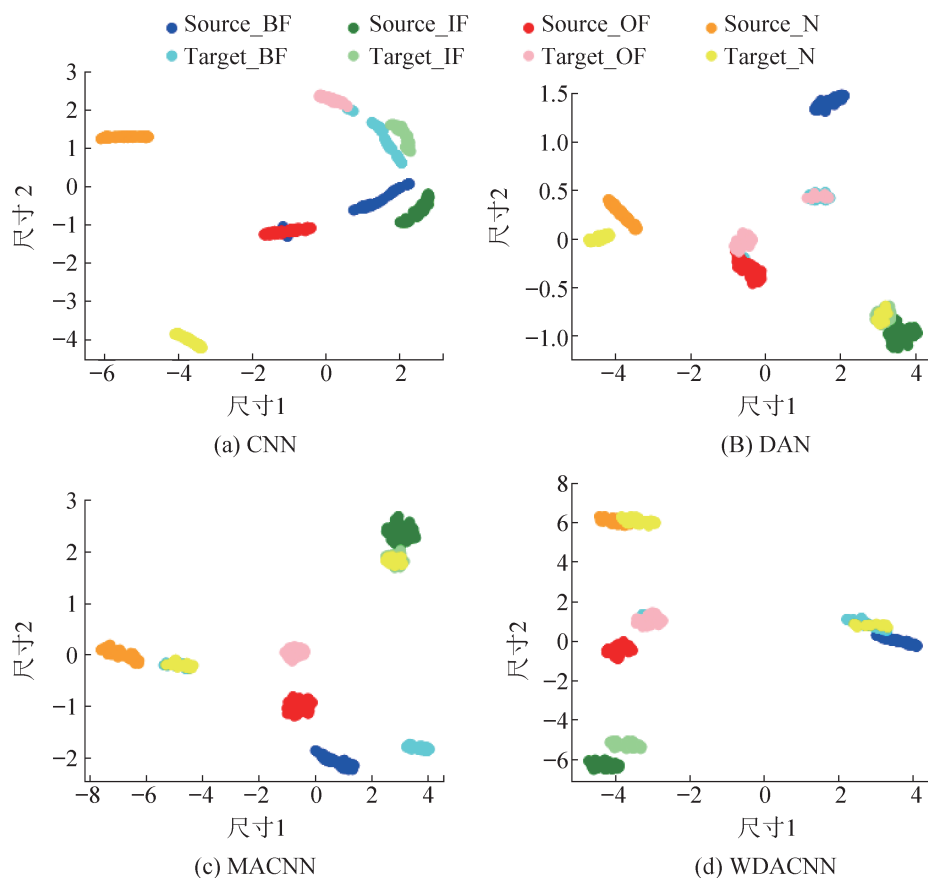


图5 不同算法的特征可视化图
Fig. 5 Feature visualization of different algorithms

可以发现，CNN算法学习到的源域和目标域的特征分布差异较大，这是由于CNN没有采用迁移学习策略，相比之下，DAN和MACNN学习到的源域和目标域之间的特征距离已经很接近，但特征对齐效果不佳。WDACNN中不同健康状态下的特征对齐效果较好，并且不同类的特征之间存在较高的区分度，这进一步验证了WDACNN在处理不同工作负载变化下的滚动轴承故障诊断时具有优越性。

4 结论

本文提出基于加权域适应卷积神经网络的滚动轴承故障诊断方法，该方法在深度学习框架中嵌入迁移学习算法，提取源域和目标域的域不变特征，使基于源域的分类器在目标域上取得良好的泛化性能，同时引入权重系数对源域样本进行

加权，以降低类权重偏差带来的影响。为了验证该方法的有效性，本文在滚动轴承振动数据集上进行迁移故障诊断任务，并与BP，LSTM，CNN，DAN，MACNN等多种故障诊断算法进行对比，实验结果表明，本文提出的算法能够有效地提取源域和目标域的域不变特征，在轴承数据集上的平均故障诊断精度达到了96.6%。

参考文献：

- [1] Wang Huaqing, Li Shi, Song Liuyang, et al. A Novel Convolutional Neural Network Based Fault Recognition Method via Image Fusion of Multi-vibration-signals[J]. Computers in Industry, 2019, 105: 182-190.
- [2] 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, 150.
Chen Renxiang, Yang Xing, Hu Xiaolin, et al. Planetary Gearbox Fault Diagnosis Method Based on Deep Belief Network Transfer Learning[J]. Journal of Vibration and

- Shock, 2021, 40(1): 127-133, 150.
- [3] 张群, 唐振浩, 王恭, 等. 基于长短时记忆网络的超短期风功率预测模型[J]. 太阳能学报, 2021, 42(10): 275-281.
- Zhang Qun, Tang Zhenhao, Wang Gong, et al. Ultra-short-term Wind Power Prediction Model Based on Long and Short Term Memory Network[J]. Acta Energetica Solaris Sinica, 2021, 42(10): 275-281.
- [4] Liu Han, Zhou Jianzhong, Zheng Yang, et al. Fault Diagnosis of Rolling Bearings with Recurrent Neural Network-based Autoencoders[J]. ISA Transactions, 2018, 77: 167-178.
- [5] Ben-David S, Blitzer J, Crammer K, et al. A Theory of Learning From Different Domains[J]. Machine Learning, 2010, 79(1): 151-175.
- [6] Guo Liang, Lei Yaguo, Xing Saibo, et al. Deep Convolutional Transfer Learning Network: A New Method for Intelligent Fault Diagnosis of Machines with Unlabeled Data[J]. IEEE Transactions on Industrial Electronics, 2019, 66(9): 7316-7325.
- [7] Jin Tongtong, Yan Chuliang, Chen Chuanhai, et al. New Domain Adaptation Method in Shallow and Deep Layers of the CNN for Bearing Fault Diagnosis Under Different Working Conditions[J]. The International Journal of Advanced Manufacturing Technology, 2023, 124(11): 3701-3712.
- [8] Li Xiang, Zhang Wei. Deep Learning-based Partial Domain Adaptation Method on Intelligent Machinery Fault Diagnostics[J]. IEEE Transactions on Industrial Electronics, 2021, 68(5): 4351-4361.
- [9] 孙琦钰, 赵超强, 唐漾, 等. 基于无监督域自适应的计算机视觉任务研究进展[J]. 中国科学(技术科学), 2022, 52(1): 26-54.
- Sun Qiyu, Zhao Chaoqiang, Tang Yang, et al. A Survey on Unsupervised Domain Adaptation in Computer Vision Tasks[J]. Scientia Sinica(Technologica), 2022, 52(1): 26-54.
- [10] 许亚雲, 严华. 无监督域适应的表示学习算法[J]. 哈尔滨工业大学学报, 2021, 53(2): 40-46.
- Xu Yayun, Yan Hua. Representation Learning for Unsupervised Domain Adaptation[J]. Journal of Harbin Institute of Technology, 2021, 53(2): 40-46.
- [11] Zhao Minghang, Kang M, Tang Baoping, et al. Deep Residual Networks with Dynamically Weighted Wavelet Coefficients for Fault Diagnosis of Planetary Gearboxes [J]. IEEE Transactions on Industrial Electronics, 2018, 65(5): 4290-4300.
- [12] Liang Mingxuan, Cao Pei, Tang J. Rolling Bearing Fault Diagnosis Based on Feature Fusion with Parallel Convolutional Neural Network[J]. The International Journal of Advanced Manufacturing Technology, 2021, 112(3): 819-831.
- [13] 崔石玉, 朱志宇. 基于参数迁移和一维卷积神经网络的海水泵故障诊断[J]. 振动与冲击, 2021, 40(24): 180-189.
- Cui Shiyu, Zhu Zhiyu. Seawater Pump Fault Diagnosis Based on Parameter Transfer and One-dimensional Convolutional Neural Network[J]. Journal of Vibration and Shock, 2021, 40(24): 180-189.
- [14] Kang Guoliang, Jiang Lu, Yang Yi, et al. Contrastive Adaptation Network for Unsupervised Domain Adaptation[C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ, USA: IEEE, 2019: 4888-4897.
- [15] 赵小强, 张亚洲. 利用改进卷积神经网络的滚动轴承变工况故障诊断方法[J]. 西安交通大学学报, 2021, 55(12): 108-118.
- Zhao Xiaoqiang, Zhang Yazhou. Improved CNN-based Fault Diagnosis Method for Rolling Bearings Under Variable Working Conditions[J]. Journal of Xi'an Jiaotong University, 2021, 55(12): 108-118.
- [16] 谢银成, 黎曦, 李天, 等. 基于改进ResNet和损失函数的表情识别[J]. 自动化与仪表, 2022, 37(4): 64-69.
- Xie Yincheng, Li Xi, Li Tian, et al. Expression Recognition Based on Improved ResNet and Loss Function[J]. Automation & Instrumentation, 2022, 37(4): 64-69.
- [17] 刘俊锋, 董宝营, 俞翔, 等. 基于FSC-MPE与BP神经网络的滚动轴承故障诊断方法[J]. 中国舰船研究, 2021, 16(6): 183-190.
- Liu Junfeng, Dong Baoying, Yu Xiang, et al. Rolling Bearing Fault Diagnosis Method Based on FSC-MPE and BP Neural Network[J]. Chinese Journal of Ship Research, 2021, 16(6): 183-190.
- [18] Long Mingsheng, Cao Yue, Cao Zhangjie, et al. Transferable Representation Learning with Deep Adaptation Networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2019, 41(12): 3071-3085.
- [19] Jin Tongtong, Yan Chuliang, Chen Chuanhai, et al. New Domain Adaptation Method in Shallow and Deep Layers of the CNN for Bearing Fault Diagnosis Under Different Working Conditions[J]. The International Journal of Advanced Manufacturing Technology, 2023, 124(11): 3701-3712.
- [20] Ganin Y, Lempitsky V. Unsupervised Domain Adaptation by Backpropagation[C]//Proceedings of the 32nd International Conference on Machine Learning. Chia Laguna Resort, Sardinia, Italy: PMLR, 2015: 1180-1189.
- [21] Laurens van der Maaten, Hinton G. Visualizing Data Using t-SNE[J]. Journal of Machine Learning Research, 2008, 9(86): 2579-2605.

<http://www.china-simulation.com>