

基于相似性度量迁移学习的轴承故障诊断

徐易芸¹，马健¹，陈良¹，沈长青²，李奇¹，孔林³

(1. 苏州大学 机电工程学院，苏州 215131；

2. 苏州大学 轨道交通学院，苏州 215131

3. 长光卫星技术有限公司，长春 130102)

摘要：滚动轴承工况多变，受负荷、转速等因素影响，故障信号的特征分布偏移会显著降低故障诊断模型的泛化能力。针对此问题，提出一种基于相似性度量迁移学习的轴承故障诊断方法。将迁移学习和相似性度量的思想结合，通过相关对齐损失计算变工况故障特征之间的相关性，最小化源域和目标域特征之间的分布差异。同时最大化输入特征与中心特征的相似性，利用目标域预测标签中包含的故障分类信息，提高故障特征聚类的准确性，来增加高相关性特征对模型的贡献度，减小非相关特征的影响。最后利用学习到的特征实现故障分类。在CWRU和自搭建实验平台上进行了对比实验，证明了所述方法能够更加准确地分类故障信号，更好解决不同工况下由特征分布偏移带来的故障诊断难点问题。

关键词：滚动轴承；故障诊断；迁移学习；相似性度量；特征分布

中图分类号：TP206+.3

文献标识码：A

A Bearing Fault Diagnosis Based on Similarity Measurement for Transfer Learning

Xu Yiyun¹, Ma Jian¹, Chen Liang¹, Shen ChangQing², Li Qi¹, Kong Lin³

(1. School Of Mechanical And Electrical Engineering, Soochow University Suzhou, 215131;

2. School of Rail Transportation, Soochow University Suzhou, 215131

3. Chang Guang Satellite Technology Co. Ltd., Changchun, 130102)

Abstract: According to the variable working conditions of rolling bearings due to the influence of load, speed and other factors, the feature distribution deviation of fault signals will significantly reduce the generalization ability of fault diagnosis model. To solve this problem, a bearing fault diagnosis method based on similarity measurement for transfer learning was proposed. The idea of transfer learning and similarity measurement were combined. The correlation between fault features under variable conditions is calculated through correlation alignment loss, and the distribution difference between source domain and target domain features is minimized. At the same time, the similarity between the input feature and the central feature is maximized, and the fault classification information contained in the label predicted by the target domain is used to improve the accuracy of fault feature clustering, so as to increase the contribution of highly correlated features to the model and reduce the influence of non-correlated features. Finally, the features learned are used to implement fault classification. Comparison experiments on CWRU and self-built experimental platform prove that the proposed method can classify fault signals more accurately and solve the difficult problems of fault diagnosis caused by feature distribution deviation under different working conditions.

Key words: rolling bearing; fault diagnosis; transfer learning; similarity measurement; feature distribution

旋转机械出现故障不仅会带来重大经济损失^[1]，而且可能危害到操作人员的安全，造成严重事故^[2]。轴承是一种应用广泛并且极其重要的旋转机械部件，针对轴承故障的诊断方法受到越来越多研究者的关注。利用轴承振动信号，文献上提出了较多故障诊断方法，包括多项逻辑回归^[3]，支持向量机（Support Vector Machine, SVM）^[4]和小波包变换^[5]等。经典方法如Kankar等^[6]将小波包变换运用到多逻辑回归算法中，建立了故障诊断模型，证明了能量和峰值可以作为用来分辨故障种类的两个特征。

为了获得更深层次的故障特征表示，基于深度学习的轴承故障智能诊断方法成为近年来的研究热点。通过卷积神经网络^[7]（Convolutional Neural Networks, CNN），自编码器^[8]（Autoencoder, AE），深度置信网络^[9]（Deep Belief Network, DBN），循环神经网络^[10]（Recurrent Neural Network, RNN），生成对抗网络^[11]（Generative Adversarial Networks, GAN）等新的机器学习模型能够自动从原始输入信号中挖掘故障特征并且进行诊断。Wang等^[12]提出了一种降噪自编码器和深度置信网络相融合模型来实现滚动轴承的故障诊断。Hong等^[13]在降噪自编码器的基础上，将改进的神经网络故障状态识别方法应用于航空发动机的故障诊断。以上研究表明，使用现有的深度学习故障诊断方法已

基金项目：国家自然科学基金（51875375）

第一作者 徐易芸 女，硕士研究生，1997年6月生

通信作者 陈良 男，博士，副教授，1981年8月生

经能够在一定程度上提升故障诊断性能。在实际工业场景中，工况往往复杂多变，且无法对所有工况数据进行有效标注。因此，基于已知域的有限有标签数据训练故障诊断模型，实现跨域场景下对其它异常工况的有效诊断，缩小由于工况变化和测试环境变化等因素带来的故障特征分布差异，是故障智能诊断领域所需要解决的关键问题。

迁移学习有助于解决这一问题。利用迁移学习从有标签的源域数据中学到知识，将这些知识迁移到无标签的目标域中完成诊断分类。Lu 等^[14]提出了一种深度域适应方法，通过使用最大平均差异（Maximum Mean Discrepancy, MMD）将数据映射到再生核希尔伯特空间来减小分布差异，同时通过权重正则化项来增强原始数据的代表性特征，从而在目标域内获得较高的分类精度。Guo 等^[15]提出了深度卷积迁移学习网络（Deep Convolutional Transfer Learning Network, DCTLN）。DCTLN 利用 CNN 自动学习特征，域自适应模块通过最大化域间判别误差，最小化域内概率分布距离帮助 CNN 学习域不变特征。Wen 等^[16]建立了一种新的深度迁移学习方法，利用三层稀疏自编码器来提取特征，并利用 MMD 来最小化训练数据和测试数据特征之间的分布差异。

使用迁移学习方法能够对不同工况下滚动轴承故障进行有效诊断。然而已有的迁移学习方法在模型训练过程中学习到的特征可能并不具有较高的特征相似性。由于异常工况的多样性、时变性和强非线性，低相关性特征往往更容易产生偏差，高相关性特征能更正确地表示对应的故障类型，由此，需要增加高相关性特征对模型的贡献度。针对此问题，本文提出一种基于相似性度量迁移学习的轴承故障诊断方法，采用滚动轴承的振动信号作为模型输入，通过 CNN 对故障特征自适应提取；结合迁移学习和相似性度量思想对损失函数改进，在域间相关对齐的同时减少类内的差异，利用相关对齐^[17]（Correlation Alignment, CORAL）损失最小化源域和目标域之间的分布差异。最大化输入特征与中心特征之间的相似性，用目标域预测标签中包含的故障分类信息作为特征聚类的划分中心，使模型能够学习与故障类型具有高相关性的特征，减小每个故障类别的类内距离，提升跨域的泛化能力，使用故障分类器实现多工况轴承故障诊断，结果表明，提出的方法具有更高的诊断精度。

1 迁移学习故障诊断基本思想

迁移学习为一类源域和目标域之间存在差异的问题提供了一种可行的解决方案^[18]。图1给出了迁移学习故障诊断的基本原理示意图。

在迁移学习中，给定源域 D_s 和任务 T_s ，目标域 D_t 和任务 T_t 。已知源域数据集 $D_s = [(x_i^s, y_i^s)]_{i=1}^{n_s}$ 的 n_s 个已

标记的样本和目标域数据集 $D_t = (x_i^t)_{i=1}^{n_t}$ 的 n_t 个未标记的样本。定义源域和目标域特征空间 X 和 Y ，对应的特征分布分别为 $P(X)$ 和 $Q(Y)$ ， $P \neq Q$ 。 D_s 和 D_t 分别从联合分布 $P(X, Y)$ 和 $Q(X, Y)$ 中采样。迁移学习的目的是构建一个深层神经网络，利用 D_s 有标签知识和 T_s 中的启发知识来提升目标预测函数 f_t 在目标域任务中的学习能力，减少跨域的分布差异，学习域不变特性并进行分类。本研究突破了故障诊断源域和目标域分布必须一致的假设，仅需要保证“源域和目标域相互关联，但分布可以不同；同时目标域的数据已知，但标签可以未知”，极大地放宽了故障诊断算法的约束条件，有助于实际工业场景的应用。

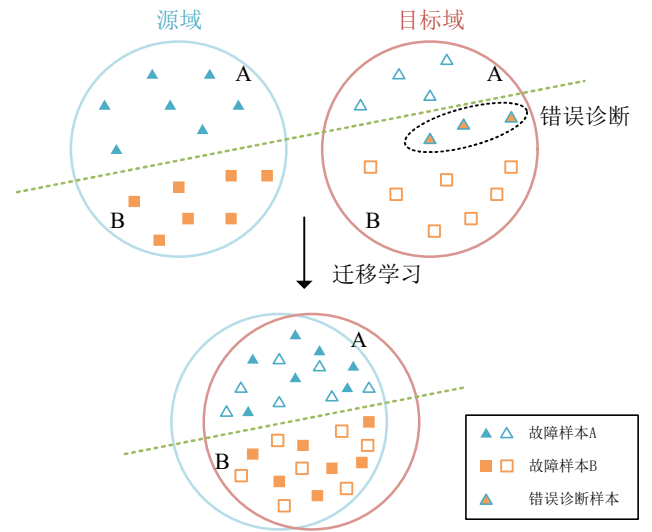


图1 迁移学习故障诊断基本原理示意图

Fig.1 Principle of transfer learning-based fault diagnosis

2 所述方法

2.1 故障诊断模型

本文提出的故障诊断模型由CNN构成。由于网络结构对模型的泛化能力有影响，针对不同问题对CNN结构参数可做适当优化，通过对CNN参数进行实验选择，考虑到计算量和效率，最终选定CNN模型如表1。

CNN模型包括特征提取器和故障分类器。特征提取器对故障特征自适应提取，使用二维卷积来保留数据的结构信息。输入由采集的轴承振动信号构建，输入尺寸的三个维度分别为通道数、特征图长与宽。特征提取器中共有5个卷积层，每一层卷积后都使用带泄露的线性修正单元（Leaky Rectified Linear Units, LeakyReLU）作为，它在训练过程中能够改善梯度消失或梯度爆炸的问题，使CNN获得更好的性能。卷积层特征计算公式为：

$$Out = LeakyReLU(W * X + b) \quad (1)$$

其中： Out 表示卷积层的输出， $*$ 是卷积操作， W 为卷积核， X 为输入特征， b 是对应的偏移量，LeakyReLU为激活函数。

表1 CNN结构详细信息

Tab.1 The details of CNN network

序号	结构	输入尺寸
1	输入	输入层
2	特征提取器	输入层
3	Conv1	卷积-LeakyReLU
4	Conv2	卷积-LeakyReLU
5	Conv3	卷积-LeakyReLU
6	Conv4	卷积-LeakyReLU
7	Conv5	卷积-LeakyReLU
8	故障分类器	卷积-LeakyReLU
9	P1	平均池化
10	FC1	全连接
11	FC2	全连接
12	FO	输出层
13		故障数

故障分类器基于特征提取器学习的特征来识别健康状况。在分类器中经过一层卷积Conv6、LeakyReLU整合提取到的特征。为了在保持重要空间信息的同时减小特征图的维数，P1层对特征平均池化。再将学习

到的特征平铺到两层全连接层FC1、FC2，最后连接输出层FO，给出故障诊断结果。

2.2 基于相似性度量迁移学习的轴承故障诊断方法

为了实现多工况轴承故障诊断，本文提出一种基于相似性度量迁移学习的轴承故障诊断方法。如图 2 所示，模型将特征分布差异度量和相似性度量与特征提取器连接，帮助 CNN 学习域不变特征。根据优化目标分为 3 个部分介绍：1) 源域数据的健康状况分类损失；2) 源域和目标域数据特征之间的 CORAL 损失；3) 源域与目标域之间的相似性度量。

1) 分类损失。采用交叉熵函数使源域训练数据的健康状况分类损失最小，定义如下：

$$L_c = -\frac{1}{N} \sum_{n=1}^N Y \ln \hat{Y} \quad (2)$$

其中：\$N\$ 为训练的样本数，\$Y\$ 为输入数据的标签向量，\$\hat{Y}\$ 为预测值向量。

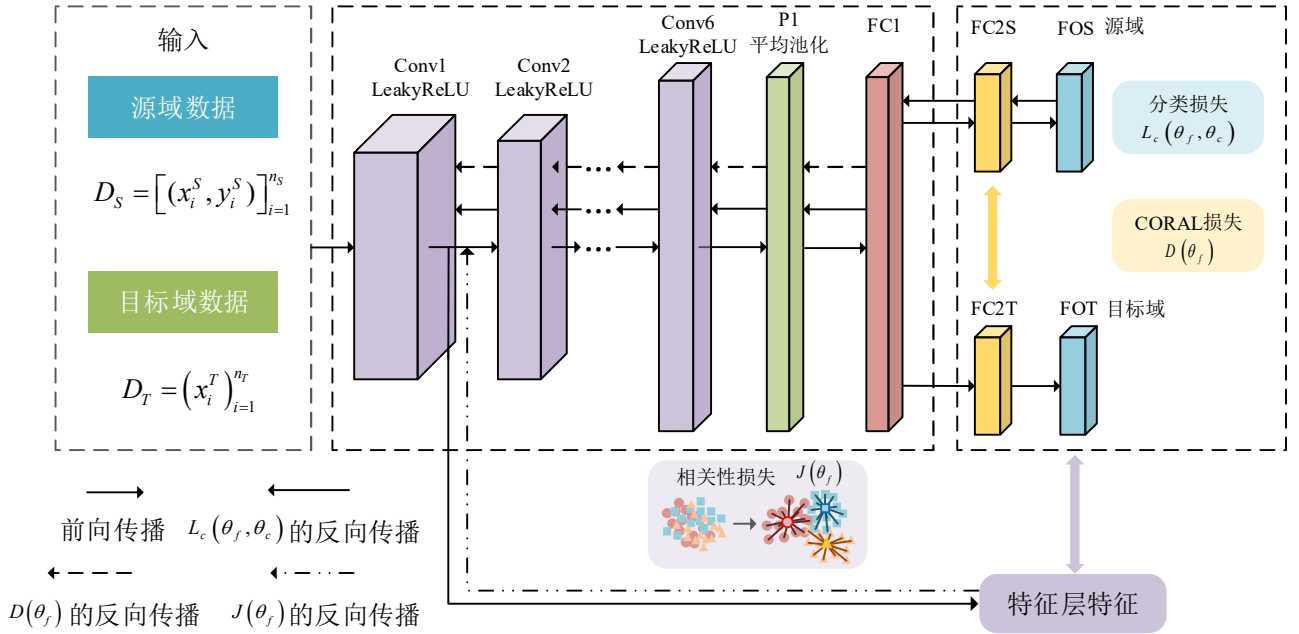


图2 故障诊断模型

Fig.2 Fault diagnosis model

得到 FOS 层输出的预测值和真实标签值的损失函数，将学习到的特征输入分类器按照源域已知故障标签的监督，通过前向和反向传播的计算和优化对模型进行更新，使得诊断模型对已知标签的各个领域数据进行有效识别。

2) CORAL 损失。使用 CORAL 损失最小化源域 FC2S 和目标域 FC2T 层之间的特征分布差异距离，计算源域和目标域特征的二阶统计量间的距离：

$$D = \frac{1}{4d^2} \|C_S - C_T\|_F^2 \quad (3)$$

其中：\$C_S\$ 和 \$C_T\$ 分别代表源域和目标域特征协方差矩阵，\$\|\cdot\|_F^2\$ 代表均方矩阵的 Frobenius 范数。

源域和目标域数据集的特征协方差矩阵定义为：

$$C_S = \frac{1}{n_s - 1} \left[D_S^T D_S - \frac{1}{n_s} (1^T D_S)^T (1^T D_S) \right] \quad (4)$$

$$C_T = \frac{1}{n_r - 1} \left[D_T^T D_T - \frac{1}{n_r} (1^T D_T)^T (1^T D_T) \right] \quad (5)$$

在处理协方差时，源域和目标域共享网络参数。

通过链式法则计算出 CORAL 损失相对于输入特征的梯度：

$$\frac{\partial D}{\partial D_s^{ij}} = \frac{1}{d^2(n_s - 1)} \left\{ \left[D_s^T - \frac{1}{n_s} (1^T D_s)^T 1^T \right]^T (C_s - C_T) \right\}^{ij} \quad (6)$$

$$\frac{\partial D}{\partial D_T^{ij}} = -\frac{1}{d^2(n_T - 1)} \left\{ \left[D_T^T - \frac{1}{n_T} (1^T D_T)^T 1^T \right]^T (C_s - C_T) \right\}^{ij} \quad (7)$$

3) 相似性度量。当数据来自不同域时, 一些特征与标签的关系会出现跨域变化^[19], 造成输入特征与中心特征之间的相似性减小。因此使用相似性度量来约束相似性之间的差异, 最大化类内的相似性度量, 即从数据中提取出与故障类型高度相关的特征。中心特征用数据标签提供的故障分类信息表示, 但由于迁移学习中目标域数据无标签, 无法在训练阶段提供故障信息, 因此使用模型预测的目标域标签, 使用预测标签中包含的分类信息为输入特征提供聚类中心, 将数据划分为不同的簇。在训练中通过前向传播计算和反向传播更新不断优化, 目标域预测的准确度逐步提高, 为特征聚类提供更准确的中心位置。

令 $f(x)$ 为模型的预测输出, 将中心特征与输入特征相似性最大化, 就是最小化特征的类内距离 J :

$$J = \frac{\beta}{k} \sum_{k=1}^k E_{x \sim \mathcal{D}(x|y=k)} \left[(f(x) - \mu_k)^2 \right] \quad (8)$$

其中: $\mu_k = E_{x \sim \mathcal{D}(x|y=k)} [f(x)]$ 表示预测输出的期望, β 为正则化参数, k 表示数据种类, \mathcal{D} 表示数据分布。

结合上述 3 个部分, 最终的损失函数写为:

$$L = L_C + \lambda D + \mu J \quad (9)$$

其中: 超参数 λ 和 μ 分别代表域适应和相似性度量的权重系数。

2.3 故障诊断步骤

本文提出的故障诊断方法流程如图 3 所示, 具体步骤如下:

- 1) 数据获取与预处理: 通过传感器采集各个领域的故障振动信号, 进行数据预处理, 如快速傅里叶变换转换成频域信号, 构建有标签的源域数据集和无标签的目标域数据集, 即构建 $D_s = [(x_i^s, y_i^s)]_{i=1}^{n_s}$ 和 $D_T = (x_i^T)_{i=1}^{n_T}$, 其中标签为故障标签。数据集构建后被划分为训练集和测试集, 训练集用于模型训练, 测试集用于验证模型的有效性。从每个样本数据中截取 1024 个时间点数据, 构建 32×32 的样本矩阵, 对数值作归一化处理, 用于模型的特征提取。
- 2) 模型训练: 读取源域和目标域的训练数据集, 建立故障诊断模型。初始化故障诊断模型的待训练参数, 设置学习率 ε 和超参数 λ 、 μ 。
前向传播: 将源域和目标域训练数据作为输入, 通过 CNN 提取特征。计算式(9)中源域和目标域训练数据特征的 CORAL 损失, 源域与目标域之间的特征相似度和源域数据的健康状况分

类损失, 输出最终的损失函数。

反向传播: 使用 Adam 优化算法^[20]反向逐层传递, 对模型进行更新。

- 3) 故障诊断: 将目标域测试数据输入最终训练好的模型, 面对不同领域的故障数据, 该模型能够有效地识别故障类型, 返回预测的健康状态。

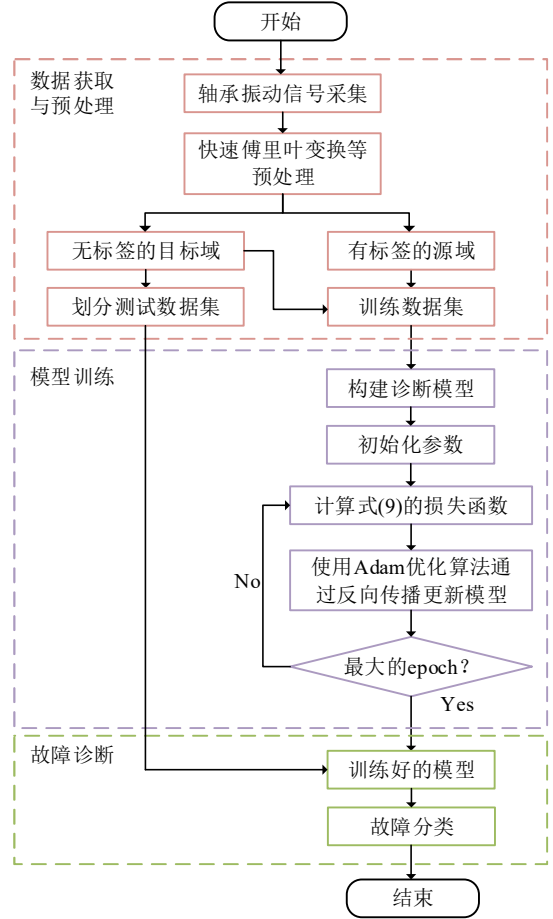


图 3 故障诊断方法流程图

Fig.3 The flow chart of diagnosis method

3 CWRU轴承实验台验证

3.1 数据集介绍与实验设计

本实验采用美国凯斯西储大学的 CWRU 轴承数据集^[21]。轴承型号为 SKF6205-2RS, 实验平台如图 1 所示。通过电火花加工模拟轴承的多种健康状况, 数据通过 16 通道的数据记录仪采集得到, 采样频率为 12kHz。数据集包含正常状态和外圈故障、内圈故障、滚动体故障三种故障类型, 轴承故障尺寸包含 7mils、14mils 和 21mils, 共 10 类不同的健康状况。

在数据采集实验中, 根据负载不同设置 4 种不同工况: 0hp、1hp、2hp 和 3hp, 一种工况对应一个数据集。每个数据集包括 4000 个训练样本和 2000 个测试样本。每次实验从 4 个数据集中选取 2 个数据集, 分别作为一个迁移任务的源域和目标域, 共 12 个不同的迁移任务。

实验中超参数 λ 和 μ 的选择经过网格搜索法进行

优化, 分别从 $\{0.001, 0.005, 0.01, 0.05, 0.1\}$ 和 $\{1, 5, 10, 15, 20\}$ 寻优范围中选择参数 λ 和 μ 。如图 5 所示, 为 0 到 1 迁移任务中选择不同 λ 和 μ 的诊断精度, 因此设置超参数 $\lambda=0.01$, $\mu=10$ 。每次循环批量大小 Batch size = 64, 学习率 $\varepsilon=0.0001$, 设定迭代次数 Epoch=100。

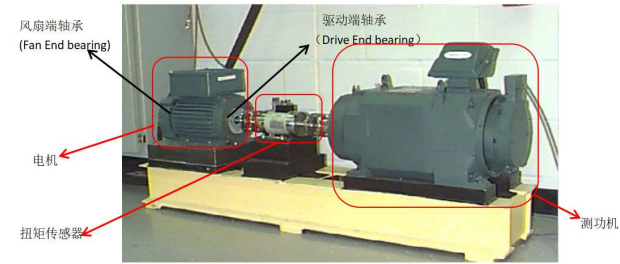


图 1 CWRU 轴承实验台
Fig.1 CWRU bearing experimental platform

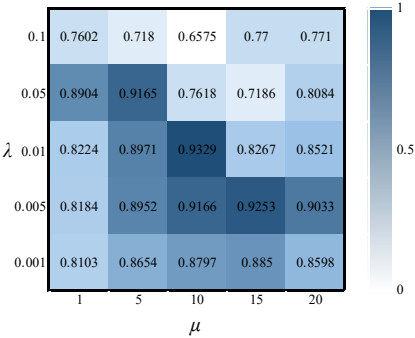


图 5 不同 λ 和 μ 的诊断精度
Fig.5 Diagnosis accuracy with varying λ and μ

3.2 对比方法

为验证本文方法的诊断性能, 与 5 种不同方法进行对比分析。根据比较目的不同, 将 5 种方法分为 3 种类型:

- 1) 传统迁移学习方法: 迁移成分分析 (Transfer Component Analysis, TCA) [22]、联合分布自适应 (Joint Distribution Adaptation, JDA) [23]。TCA 和 JDA 在源域和目标域处于不同数据分布时, 将两个领域的的数据一起映射到一个高维的再生核希尔伯特空间。在此空间中, 最小化源域和目标域数据的距离, 同时最大程度地保留它们各自的内部属性。
- 2) 未使用迁移学习的 CNN。对比实验 CNN 的网络与提出的架构相同, 其损失函数仅使用交叉熵计算分类损失, 即 $L = L_c$ 。
- 3) MMD 和 CORAL。未使用相似性度量, 分别利用 MMD 损失和 CORAL 损失降低领域间特征分布的差异。MMD 和 CORAL 与 CNN 的架构相同, 最小化 FC2 层上特征的分布差异, 即损失函数为 $L = L_c + \lambda D$, λ 为域适应权重系数。

3.3 实验结果分析

12 个迁移任务的故障诊断准确率如表 2 所示, 可以看出所提方法的诊断性能优于对比方法, 得到以下 3 个观察结论:

- 1) 对于不同工况下的故障诊断任务, 本文方法在 12 个迁移任务中准确率平均值为 86.81%, 优于对比方法。
- 2) 在故障诊断实验中, TCA 和 JDA 将振动数据映射到高维空间中最小化域间距离, CNN、MMD、CORAL 和本文方法直接从振动信号中学习特征, 能够防止丢弃有用的条件表示信息, 因此使用深度学习学习方法学习到的特征优于传统方法提取的特征。在相同的 CNN 架构中, 在 FC2 层使用 CORAL 减小分布差异的诊断结果优于 MMD。
- 3) 与迁移学习方法 CORAL 相比, 本文提出方法在

表2 CWRU实验台不同方法的故障诊断结果

Tab.2 Fault diagnosis results of different methods on CWRU experimental platform						
迁移任务	TCA	JDA	CNN	MMD	CORAL	Proposed
0→1	60.7%	61.48%	83.95%	85.39%	89.73%	93.29%
0→2	62.03%	57.40%	92.50%	91.35%	92.70%	95.60%
0→3	70.9%	73.00%	55.88%	59.26%	63.24%	78.55%
1→0	67.3%	85.58%	95.24%	81.40%	94.79%	96.15%
1→2	70.35%	86.85%	98.89%	93.78%	96.88%	98.24%
1→3	79.68%	78.03%	78.65%	72.40%	77.82%	83.45%
2→0	59.28%	60.43%	80.78%	86.59%	76.11%	80.75%
2→1	65.15%	86.73%	93.85%	95.17%	93.85%	92.65%
2→3	69.25%	80.00%	73.43%	82.71%	83.95%	85.11%
3→0	70.38%	70.38%	58.33%	59.01%	63.26%	71.60%
3→1	70.5%	77.35%	68.93%	75.09%	73.67%	79.01%
3→2	69.63%	79.73%	85.60%	80.42%	88.04%	90.43%
平均值	67.93%	74.75%	80.50%	80.21%	82.84%	86.81%

12 个迁移任务中获得了更高的健康状况, 本文方法比 CORAL 更有效地学习到具有高相关性的特征, 减少域间数据的分布差异。

4 自搭建轴承实验台验证

为进一步验证本方法性能, 在实验室自搭建轴承

数据采集实验平台上进行了验证，如图 6 所示。

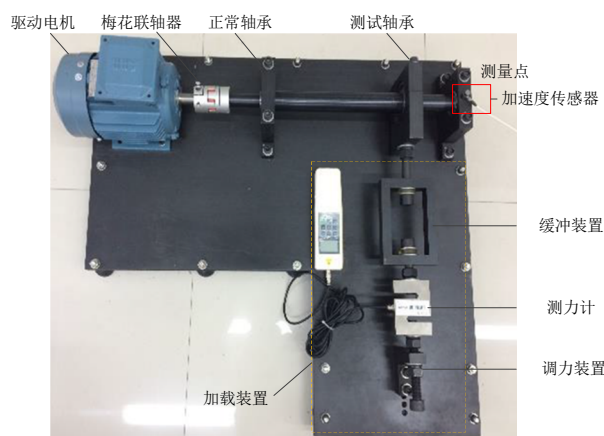


图 6 自搭建轴承实验台

Fig.6 Self-built bearing experimental platform

自搭建实验平台采集电机负载为 0kN、1kN、2kN 和 3kN 时不同健康状况的轴承振动信号，实验采用的轴承型号为 SKF6205-2RS，轴承健康状况分别为正常、外圈故障、内圈故障和滚动体故障，轴承故障尺寸为 0.2mm、0.4mm 和 0.6mm。如表 3 所示，将上述 10 种不同的健康状态分别设置为 0-9 类标签。其它实验设置与 CWRU 实验相同。

在本实验台上进行了 12 个迁移任务的实验，如图 7 所示。在不同的迁移任务中，本文方法平均诊断精度为 82.7%。与 CNN、MMD 和 CORAL 相比，本文方法的准确率高于对比方法，证明加入相似性度量后能够有效提高模型的诊断效果。

图 8 为本实验中 0 到 1 迁移任务结果的混淆矩阵。结果表明，本文方法在正常和 9 种故障状态下的平均测试精度为 94.7%，标准偏差为 5.3%，优于对比方法，证明了本文方法学习特征具有更高的故障相关性，在

不同工况下具有较好的区分度。实验结果进一步验证了本文方法的有效性和优越性。

表3 自搭建轴承故障数据

Tab.3 Fault data of self-built bearing

类别标签	故障大小(mm)	健康状态
0	无	正常
1	0.2	内圈故障
2		滚动体故障
3		外圈故障
4	0.4	内圈故障
5		滚动体故障
6		外圈故障
7	0.6	内圈故障
8		滚动体故障
9		外圈故障

5 结论

本文将相似性度量迁移学习应用于故障诊断领域，提出了一种轴承变工况跨域故障诊断新方法。将迁移学习和相似性度量结合，增加高相关性特征的贡献度，减小非相关特征的影响；通过相关对齐损失最小化源域和目标域特征之间的分布差异，最大化输入特征与中心特征之间的相似性。与未加入相似性度量的迁移学习诊断方法相比，本文所提模型经过训练后，能够有效地获得相关性较高的泛化特征，可以获得更高的轴承健康状态识别精度，从而有效解决特征分布偏移带来的故障诊断难点问题，具有较好的应用前景。

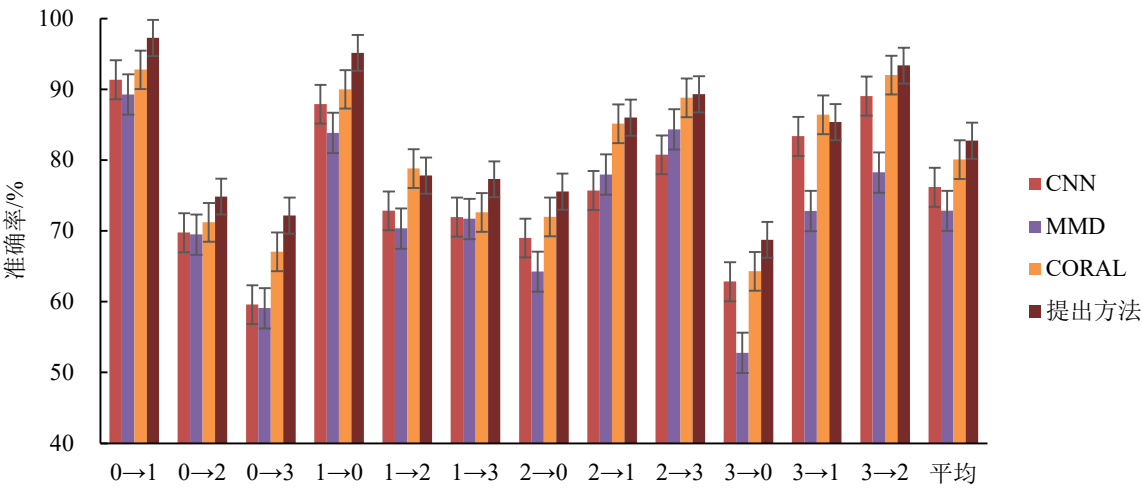


图 7 自搭建实验台故障诊断结果比较

Fig.7 Comparison of fault diagnosis results on self-built experimental platform

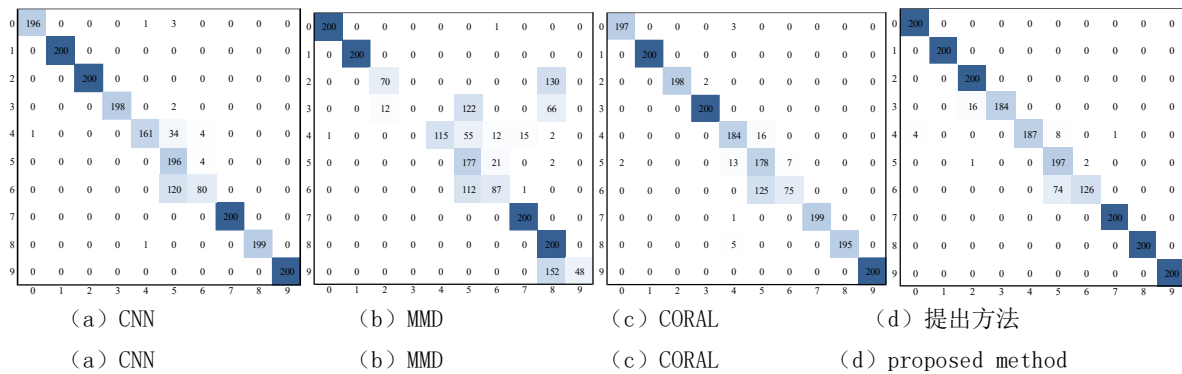


图 8 0→1 迁移任务测试数据分类的混淆矩阵

Fig.8 Confusion matrix for the classification of test data of 0→1 transfer task

参考文献 (小五黑体)

- [1] Ciabattini L, Ferracuti F, Freddi A, et al. Statistical Spectral Analysis for Fault Diagnosis of Rotating Machines [J]. IEEE Transactions on Industrial Electronics, 2018,65(5):4301-4310.
- [2] 雷亚国, 何正嘉. 混合智能故障诊断与预示技术的应用进展[J]. 振动与冲击, 2011, 30(9):129-135.
LEI Yaguo, HE Zhengjia. Advances in applications of hybrid intelligent fault diagnosis and prognosis technique[J]. Journal of Vibration and Shock, 2011, 30(9):129-135.
- [3] He Q. Time-frequency manifold for nonlinear feature extraction in machinery fault diagnosis[J]. Mechanical Systems and Signal Processing, 2013, 35(1-2):200-218.
- [4] Peng Z K, Chu F L, Tse P W. Detection of the rubbing-caused impacts for rotor-stator fault diagnosis using reassigned scalogram[J]. Mechanical Systems and Signal Processing, 2005, 19(2):391-409.
- [5] Yan R, Gao R X, Chen X. Wavelets for fault diagnosis of rotary machines: A review with applications[J]. Signal Processing, 2014, 96(PART A):1-15.
- [6] Kankar P K, Sharma S C, Harsha S P. Rolling element bearing fault diagnosis using wavelet transform[J]. Neurocomputing, 2011, 74(10):1638-1645.
- [7] Janssens O, Slavkovikj V, Vervisch B, et al. Convolutional Neural Network Based Fault Detection for Rotating Machinery[J]. Journal of Sound and Vibration, 2016, 377:331-345.
- [8] Li Y, Chen P, Wang K, et al. A deep convolutional autoencoder for wind turbine blades health condition monitoring based on FBG strain gauges measurement[C]// Proceedings of the 7th Asia-Pacific Workshop on Structural Health Monitoring, 2018: 973-981.
- [9] Chen Z, Li W. Multisensor feature fusion for bearing fault diagnosis using sparse autoencoder and deep belief network[J]. IEEE Transactions on Instrumentation and Measurement, 2017, 66(7):1693-1702.
- [10] Abed W, Sharma S, Sutton R, et al. Robust Bearing Fault Detection and Diagnosis Technique for Brushless DC Motors Under Non-stationary Operating Conditions[J]. Journal of Control, Automation and Electrical Systems, 2015, 26(3):241-254.
- [11] Liu H, Zhou J, Xu Y, et al. Unsupervised fault diagnosis of rolling bearings using a deep neural network based on generative adversarial networks[J]. Neurocomputing, 2018, 315:412-424.
- [12] 王应晨, 段修生. 深度学习融合模型在机械故障诊断中的应用[J]. 振动、测试与诊断, 2019, 39(6):1271-1276.
WANG Yingchen, DUAN Xiusheng. Improved Deep Learning Fusion Model in Fault Diagnosis[J]. Journal of Vibration, Measurement & Diagnosis, 2019, 39(6):1271-1276.
- [13] 洪骥宇, 王华伟, 车畅畅, 等. 改进降噪自编码的航空发动机气路故障诊断[J]. 振动、测试与诊断, 2019, 39(3):603-610.
HONG Jiyu, WANG Huawei, CHE Changchang, et al. Improved Deep Learning Fusion Model in Fault Diagnosis[J]. Journal of Vibration, Measurement & Diagnosis, 2019, 39(3):603-610.
- [14] Lu W, Liang B, Cheng Y, et al. Deep Model Based Domain Adaptation for Fault Diagnosis[J]. IEEE Transactions on Industrial Electronics, 2017, 64(3):2296-2305.
- [15] Guo L, Lei Y, Xing S, 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.
- [16] Wen L, Gao L, Li X. A new deep transfer learning based on sparse auto-encoder for fault diagnosis[J]. IEEE Transactions on Systems, Man, and Cybernetics: Systems, 2019, 49(1):136-144.
- [17] Wang X, He H, Li L. A hierarchical deep domain adaptation approach for fault diagnosis of power plant thermal system[J]. IEEE Transactions on Industrial Informatics, 2019, 15(9): 5139-5148.

- [18] 雷亚国, 杨彬, 杜兆钧, 等. 大数据下机械装备故障的深度迁移诊断方法[J]. 机械工程学报, 2019, 55(7):1-8.
LEI Yaguo, YANG Bin, DU Zhaojun, et al. Deep Transfer Diagnosis Method for Machinery in Big Data Era[J]. Journal of Mechanical Engineering, 2019, 55(7):1-8.
- [19] An Z, Li S, Wang J, et al. Generalization of deep neural network for bearing fault diagnosis under different working conditions using multiple kernel method[J]. Neurocomputing, 2019, 352:42–53.
- [20] Kingma D P, Ba J L. Adam: A method for stochastic optimization[J]. arXiv preprint arXiv:1412.6980, 2014.
- [21] Smith W A, Randall R B. Rolling element bearing diagnostics using the Case Western Reserve University data: A benchmark study[J]. Mechanical Systems and Signal Processing, 2015, 64-65:100-131.
- [22] Pan S J, Tsang I W, Kwok J T, et al. Domain adaptation via transfer component analysis[J]. IEEE Transactions on Neural Networks, 2011, 22(2):199–210.
- [23] Long M, Wang J, Ding G, et al. Transfer feature learning with joint distribution adaptation[J] IEEE International Conference on Computer Vision, 2013:2200–2207.