# 基于联合特征匹配的风电机组轴承故障诊断方法

王 娜[1,2]，王子从[1]，刘佳林[1]
(1. 天津工业大学 控制科学与工程学院， 天津 300387；2. 天津市电气装备智能控制重点实验室，天津 300387)

**摘 要**：针对风电机组中滚动轴承跨域诊断泛化性弱的问题，文章提出一种基于联合特征匹配的风电机组轴承故障诊断方法。首先，基于一维卷积神经网络设计了数据筛选方法，从源域中筛选出与目标域相似的数据组成新源域，重构训练集以减弱负迁移效应；其次，利用连续小波变换获取训练数据的二维时频特征，再通过残差神经网络获得一维特征向量作为迁移学习的初始特征，确保模型对训练数据特征的全面提取；再次，综合考虑源域和目标域的特征迁移损失，以及源域的分类损失，提出新的联合特征匹配损失函数，实现从源域到目标域特征的有效迁移；最后，利用所构建残差神经网络的 softmax 分类器实现故障的诊断分类。结果表明，通过多种源域与目标域的迁移任务仿真，验证了所提模型在跨域环境中具有较好的诊断性能。

**关键词**：风电机组；滚动轴承；跨域诊断；深度迁移学习；联合特征匹配
**中图分类号**：TK81 **文献标志码**：A **文章编号**：1671-5292(2025)12-1619-11

## 0 引言

风电机组的稳定运行对于保障能源供应和降低运营成本至关重要。而滚动轴承作为风电机组中关键的动力传输部件，长期运行于高负荷和复杂工况环境中，尤其在高湿、高腐蚀等恶劣条件下，轴承故障问题尤为突出。因此，及时准确地诊断轴承的故障状态，有效避免故障的扩大和非计划停机，对保障风电机组安全运行具有重要的现实意义[1,2]。

风电机组轴承的振动信号包含大量故障信息，且易于检测，故常采用数据驱动方法进行故障诊断。然而，传统的机器学习依赖人工经验，难以提取非平稳、非线性轴承信号的深层特征。深度学习因具有强大的深层特征提取能力，被广泛应用于故障诊断领域[3]。但深度学习通常需要大量标签数据进行训练，而轴承在故障状态下的标签数据极为稀缺，难以支撑模型的有效训练。此外，传统深度学习通常假设训练与测试数据独立分布，但现实中轴承工况多变，导致模型在不同工况下的泛化能力较差[4]。

对此，迁移学习(Transfer Learning，TL)以历史工况作为源域，待测工况作为目标域，在跨域学习中提高所建模型的泛化能力[5]。迁移学习在故障诊断中的应用主要分为三类：模型法、数据法和特征法。模型法通过源域数据预训练深度学习模型，再利用目标域数据微调模型参数，但微调层数的选取通常依赖试凑方式，主观性较强[6]；数据法则通过选取源域中与目标域相似的数据，并通过赋予其较高权重来训练深度学习模型，从而实现迁移，但未从根本上缩小源域与目标域的数据分布差异，易导致模型泛化能力受限[7]；特征法主要利用不同的方法对源域和目标域样本特征间的差异进行衡量，通过缩小差异对齐源域和目标域之间的特征，以实现特征的迁移。相较于数据法和模型法，特征法对源域与目标域之间的特征分布差异进行了具体的度量[8,9]，因此在训练数据与测试数据分布不同的情况下，能够表现出更优的诊断效果。文献[10]提出了一种基于几何空间变换的特征迁移方法，通过最大均值差异(Maximum Mean Discrepancy，MMD)度量准则寻找能最小化源域与目标域时域特征差异的变换矩阵，从而实现跨域诊断。文献[11]将 MMD 引入到卷积神经网络的损失函数中，通过最小化多层领域适配差异和伪标签学习的约束项，在网络优化过程中逐步缩小源域和目标域特征分布的差异，实现特征迁移。文献[12]提出了一种结合多尺度卷积神经网

**收稿日期**：2024-11-23。
**基金项目**：天津市自然科学基金重点项目(23JCZDJC01140)。
**作者简介**：王 娜(1977-)，女，博士，讲师，研究方向为智能故障诊断、态势感知。E-mail：wangna@tiangong.edu.cn

络与局部最大均值差异 (Local Maximum Mean Discrepancy, LMMD)的轴承跨域诊断方法,通过特征适配改善跨域诊断性能, 并在多个轴承数据集上进行了验证。

上述基于特征的迁移方法存在诸多不足。首先, 传统特征法忽视了源域中可能存在与目标域差异较大的数据, 这些数据会干扰网络的训练过程,甚至产生负迁移效应,降低模型的拟合效果。其次,传统特征法一般直接对原始信号进行操作,或仅采用时域、频域特征中的一种进行分析,存在难以全面提取信号中蕴含的复杂特征信息[13]的局限性。再次,传统特征法通常仅关注特征间的全局或局部差异中的一方面, 导致源域与目标域的特征分布未能充分对齐, 从而削弱了模型在迁移学习中的效果,影响了其泛化能力。

综上所述, 本文提出了一种基于联合特征匹配(Combined Features Adaptation, CFA)的跨域诊断方法。首先, 提出基于一维卷积神经网络(1D-Convolutional Neural Network, 1D-CNN)的数据筛选机制, 剔除与目标域差距较大的源域数据,减少该类数据对特征迁移的干扰,提高特征迁移学习的有效性和模型对目标域数据的泛化能力。其次, 利用连续小波变换生成二维时频特征图, 并将其作为残差神经网络 (Residual Neural Network, Resnet)的输入,以增强模型对轴承信号复杂非线性特征的表征能力。最后,设计新的联合特征匹配损失函数,兼顾域间差异和同类间差异,引入权重系数平衡新源域与目标域的特征迁移损失及新源域的分类损失, 实现特征的充分迁移与故障类别的精准划分。

## 1 基于联合特征匹配的风电机组轴承故障诊断

本文所提基于联合特征匹配的故障诊断方法的基本原理如图 1 所示。

首先, 利用 1D-CNN 对源域数据进行筛选,得到与目标域分布更相似的新源域数据集。其次,通过连续小波变换 (Continuous Wavelet Transform, CWT) 提取新源域与目标域的二维时频特征,将其输入初始 ResNet 网络映射中作为一维特征向量集。再次,计算新源域的分类损失,以及新源域与目标域特征向量间的迁移损失, 构建联合特征匹配损失函数,基于该损失函数对网络进行更新、迭代、调整网络参数,从而改变网络的映射空间,

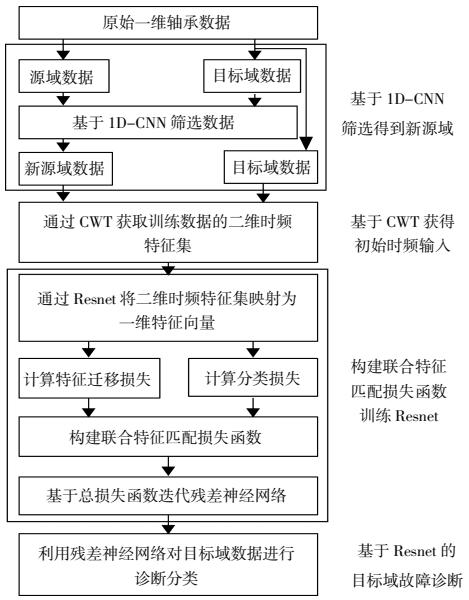



**图 1 基于联合特征匹配的风电机组轴承故障诊断基本原理**

Fig.1 Principle of wind turbine bearing fault diagnosis via combined features adaptation

提高特征分布的相似性。最终,获得残差神经网络模型,并得到目标域的故障诊断分类结果。

### 1.1 基于 1D-CNN 的新源域筛选

在迁移学习中, 若源域中存在与目标域差异较大的数据,将对迁移效果产生负面影响,从而降低模型在目标域上的表现,即负迁移效应。为此,本文提出了一种基于判别的数据筛选机制(Discrimination Data Screening, DDS)。在数据预处理阶段, 利用神经网络学习源域和目标域的显著特征, 再通过训练好的网络对源域数据进行判别。当源域数据被分类错误时,表明其特征分布更接近目标域;反之,则表示与目标域差异较大。通过此方法筛选出与目标域相似性更高的源域数据,以提升数据可迁移性,抑制负迁移效应。本文选用 1D-CNN 作为筛选网络, 其结构如图 2 所示。该网络通过多层一维卷积逐步提取深层特征信息,池化层在保留关键特征的同时进行降维,以减少计算量, 最终由全连接层将卷积输出映射为分类向量,并利用 Softmax 分类器完成判别。

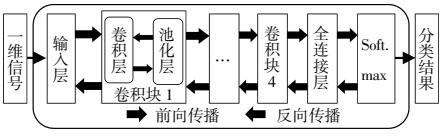



**图 2 1D-CNN 基本结构**

Fig.2 Basic structure of 1D-CNN

新源域的数据筛选过程如图 3 所示。将给定源域 $Z_s$ 划分为分类集 $Z_1$ 和训练集 $Z_2$，使用目标域 $Z_t$ 与 $Z_2$ 共同迭代训练 1D-CNN 网络。训练完成后，对 $Z_1$ 进行分类，得到正确类 $C_1$ 和错误类 $C_2$。从 $C_2$ 中每个类别下选取相同数量的数据,构成新源域 $Z_{ns}$。经过筛选得到的新源域与目标域之间的分布差异更小,提升了后续迁移效果,并有效抑制了负迁移效应对模型表现的影响。

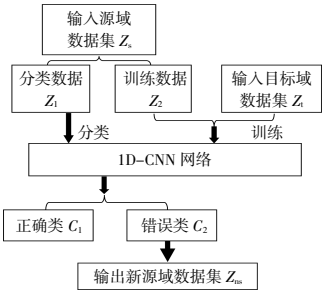



**图 3　筛选新源域**

Fig.3 Screening for the new source domain

### 1.2　基于连续小波变换的二维时频特征生成

通过连续小波变换获取新源域与目标域一维轴承信号的小波系数矩阵，再将其转化为二维图像，以获取适合 ResNet 网络输入的有效时频特征。小波系数 $W_f$ 的计算式为

$$W_f(a,b)=\frac{1}{\sqrt{a}}\int_{-\infty}^{\infty}f(t)\psi^*\left(\frac{t-b}{a}\right)\mathrm{d}t \tag{1}$$

$$\psi(t)=\mathrm{e}^{\omega_0 t\cdot i}\,\mathrm{e}^{-\frac{t^2}{2}} \tag{2}$$

$$a=(F_s\times F_c)/f \tag{3}$$

式中:$W_f$ 为小波系数;$a$ 为尺度参数;$b$ 为位移参数;* 为取共轭;$\psi(\cdot)$为母小波 $\psi(t)$在 $a$ 和 $b$ 不同取值下得到的一系列函数族,采用 Morlet 小波作为母小波;$\omega_0$ 为母小波的频率参数;$F_s$ 为采样频率;$F_c$ 为母小波的初始中心频率;$f$ 为不同尺度对应的实际频率。

通过上述 CWT 转换，将新源域和目标域中的一维轴承振动信号转换为二维时频特征图像,以 1×2 048 的一维振动信号为例,经过 CWT 转化为 224×224 的三通道时频图像,得到一维信号中的故障时频特征。

### 1.3　基于联合特征匹配函数的残差神经网络构建

经典迁移学习一般采用 MMD 作为特征迁移损失。MMD 是一种非参数距离估计指标,在度量两个分布的差异时，将数据映射到高维再生核希尔伯特空间再计算数据均值差异的平方，以此来评估分布的相似性。给定不同分布下的向量集合 $\boldsymbol{Z}^s=\{z_1^s,z_1^s,\cdots,z_n^s\}$,$\boldsymbol{Z}^t=\{z_1^t,z_1^t,\cdots,z_m^t\}$ 分别作为新源域数据集及目标域数据集，最大均值差异损失 $L_{MMD}$ 的计算式为

$$L_{MMD}=\left\|\frac{1}{n}\sum_{i=1}^{n}\phi(z_i^s)-\frac{1}{m}\sum_{j=1}^{m}\phi(z_j^t)\right\|_H^2=$$

$$\frac{1}{n^2}\sum_{i,j=1}^{n}\left[\phi(z_i^s),\phi(z_j^s)\right]+\frac{1}{m^2}\sum_{i,j=1}^{m}\left[\phi(z_i^t),\phi(z_j^t)\right]-$$

$$\frac{2}{nm}\sum_{i=1}^{n}\sum_{j=1}^{m}\left[\phi(z_i^s),\phi(z_j^t)\right]=\frac{1}{n^2}\sum_{i,j=1}^{n}k(z_i^s,z_j^s)+$$

$$\frac{1}{m^2}\sum_{i,j=1}^{n}k(z_i^t,z_j^t)-\frac{2}{nm}\sum_{i=1}^{n}\sum_{j=1}^{m}k(z_i^s,z_j^t) \tag{4}$$

$$k(x,y)=\exp\left(-\frac{\|x-y\|^2}{2\sigma^2}\right) \tag{5}$$

式中:$n,m$ 分别为新源域、目标域的数据量;$\phi(\cdot)$为非线性映射函数，作用是将数据映射到高维再生核希尔伯特空间;$H$ 为再生核希尔伯特空间;$[\cdot,\cdot]$为向量的内积运算;$k(\cdot,\cdot)$为高斯核函数;$z_i^s$ 为源域集合中第 $i$ 个向量;$z_j^s$ 为目标域中第 $j$ 个向量;$\|\cdot\|^2$ 表示再生核希尔伯特空间中的 2-范数运算;$\sigma$ 为高斯核函数的带宽参数。

MMD 中核函数的选取是固定的,可采用单一的高斯核函数或线性核函数，但无法确定是否为最优的核函数。多核最大均值差异(Multi-kernel Maximum Mean Discrepancy，MK-MMD) 是在 MMD 的基础上发展而来,MK-MMD 假设最优的核可以由多个核线性组合得到，因此采用了多核构造总核函数的形式。原因在于,单一核函数的表达能力有限,无法充分捕捉复杂分布之间的差异,而多个核的组合可以更灵活地适配分布特性不同的数据。通过优化核函数的权重组合,MK-MMD 能够在特征空间中更全面地捕捉分布信息，从而提高对复杂分布差异的衡量能力。多核最大均值差异损失 $L_{MK-MMD}$ 的计算式如式(6)所示。

相比于 $L_{MMD}$,$L_{MK-MMD}$ 中的单核函数进行了拓

$$L_{MK-MMD}=\sum_{u=1}^{k_n}\beta_u\left[\frac{1}{n^2}\sum_{i,j=1}^{n}k_u(z_i^s,z_j^s)+\frac{1}{m^2}\sum_{i,j=1}^{m}k_u(z_i^t,z_j^t)-\frac{2}{nm}\sum_{i=1}^{n}\sum_{j=1}^{m}k_u(z_i^s,z_j^t)\right] \tag{6}$$

展,由原本的单一高斯核改为高斯核的线性组合,高斯核函数中的带宽参数控制着核函数的相似度衰减速度,通过选择一系列不同的带宽参数,可以构成多种不同的高斯核函数,每个核函数对应不同的相似度尺度,以此建立高斯核的线性组合 $K$,其计算式为

$$K=\begin{cases}\sum_{u=1}^{n_k}\beta_u k_u \\ s.t.\ \sum_{u=1}^{n_k}\beta_u=1,\beta_u\geqslant 0\end{cases} \tag{7}$$

$$k_u(x,y)=\exp\left(-\frac{\|x-y\|^2}{2\sigma_u^2}\right) \tag{8}$$

$$L_{\mathrm{LMMD}}=\frac{1}{C}\sum_{c=1}^{C}\left\|\sum_{i=1}^{n}\omega_{i,c}^{s}\varphi\left(z_i^{s}\right)-\sum_{j=1}^{m}\omega_{j,c}^{t}\varphi\left(z_j^{t}\right)\right\|_H^2=$$
$$\frac{1}{C}\sum_{c=1}^{C}\left[\sum_{i,j=1}^{n}\omega_{i,c}^{s}\omega_{j,c}^{s}k\left(z_i^{s},z_j^{s}\right)+\sum_{i,j=1}^{m}\left[\omega_{i,c}^{t}\omega_{j,c}^{t}k\left(z_i^{t},z_j^{t}\right)\right]-2\sum_{i=1}^{n}\sum_{j=1}^{m}\omega_{i,c}^{s}\omega_{j,c}^{t}k\left(z_i^{s},z_j^{t}\right)\right] \tag{9}$$

式中:$C$ 为类别数;$\varphi(\cdot)$为映射至再生核希尔伯特空间中的非线性映射函数;$\omega_{i,c}^{s}$ 和 $\omega_{j,c}^{t}$ 分别为新源域向量 $z_i^{s}$ 和目标域向量 $z_j^{t}$ 属于第 $c$ 类的权重,其计算式分别为

$$\omega_{i,c}^{s}=\frac{y_{i,c}^{s}}{\sum_{l=1}^{n}y_{l,c}^{s}} \tag{10}$$

$$\omega_{j,c}^{t}=\frac{y_{j,c}^{t}}{\sum_{k=1}^{m}y_{k,c}^{t}} \tag{11}$$

式中:$y_{i,c}^{s}$ 为新源域中向量 $z_i^{s}$ 属于第 $c$ 类的标签向量;$y_{j,c}^{t}$ 为目标域中向量 $z_j^{t}$ 属于第 $c$ 类的标签向量;$n$ 为新源域数据量;$m$ 为目标域数据量。

新源域数据使用真实标签向量计算每个数据属于第 $c$ 类的权重,因目标域无标签而无法计算权重,考虑到神经网络的输出即是概率分布,将网络输出作为目标域的伪标签辅助计算权重。MMD 和 MK-MMD 主要从全局角度度量域间分布差异,LMMD 则通过关注局部结构衡量分布间的细微差异。为了充分发挥各自优势,本文引入联合分布适配思想[14],结合 MK-MMD 和 LMMD,综合考虑全局与局部差异,构建新的特征迁移函数,实现对新源域和目标域之间分布差异的全面度量。特征迁移损失 $L_{\mathrm{FT}}$ 的计算式为

$$L_{\mathrm{FT}}=(1-\mu)L_{\mathrm{MK-MMD}}+\mu\cdot L_{\mathrm{LMMD}} \tag{12}$$

式中:$\mu$ 为特征迁移权重系数,用来调整局部损失与全局损失的权重占比。

本文将 $L_{\mathrm{FT}}$ 嵌入至网络的损失函数计算层,将减小新源域和目标域的特征分布差异作为网络的寻优目标,以此在网络优化映射空间的过程中不断寻找能够使两域特征分布差异最小的映射空间,实现两域的特征对齐,原理如图 4 所示。

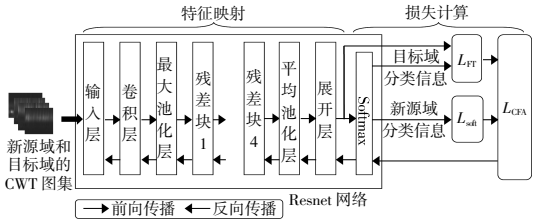



**图 4 构建联合特征匹配损失训练 Resnet**

Fig.4 Constructing combined features adaptation loss training Resnet

图 4 中,新源域与目标域的时频图像集通过 Resnet 网络的特征映射层得到对应的一维特征向量集,再计算新源域数据的分类损失 $L_{\mathrm{soft}}$,以及新源域与目标域特征向量间的特征迁移损失 $L_{\mathrm{FT}}$,以此构成联合特征匹配损失 $L_{\mathrm{CFA}}$。其中,$L_{\mathrm{soft}}$ 用于提升分类器对新源域数据的分类能力,而 $L_{\mathrm{FT}}$ 旨在

缩小源域与目标域的分布差异，从而确保分类器在源域上的分类能力能够同样作用在目标域上。将 $L_{CFA}$ 反馈回 Resnet 网络调整参数,并在网络的迭代学习中不断缩小两域间的特征分布差异,达到联合特征匹配的目标。最终,将目标域数据集输入至已训练的残差神经网络,完成分类诊断。

图 4 中的残差块由多个残差单元构成，参数设置在仿真部分说明,残差单元结构如图 5 所示。

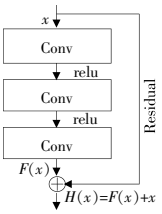



**图 5　残差单元的基本结构**

Fig.5 Structure of residual unit

图 4 中联合特征匹配损失 $L_{CFA}$ 的计算式为

$$L_{CFA}=\lambda L_{FT}+L_{soft} \tag{13}$$

式中:$L_{FT}$ 为特征迁移损失;$L_{soft}$ 为交叉熵分类损失;$\lambda$ 为总迁移权重系数,$\lambda\in(0,1)$,用来调整 $L_{FT}$ 所占比重。

$$\lambda=\frac{2}{1+e^{-\frac{10\cdot i}{p}}}-1 \tag{14}$$

式中:$i$ 为当前迭代次数;$p$ 为总迭代次数。

交叉熵分类损失 $L_{soft}$ 的计算式为

$$L_{soft}=-\sum_{i=1}^{n} y\lg\hat{y} \tag{15}$$

式中:$y$ 为新源域数据的标签值;$\hat{y}$为预测的概率值;$n$ 为新源域数据量。

### 1.4　方法流程

综上，基于联合特征匹配的风电机组轴承故障诊断方法步骤如下，对应的故障诊断流程如图 6 所示。

步骤 1：将采集的一维轴承信号按不同工况分为有标签的源域和无标签的目标域。

步骤 2：使用源域与目标域数据训练 1D-CNN 网络,计算 1D-CNN 网络的分类损失。

步骤 3:判断网络是否收敛,若是,则基于该网络将源域中分类错误的数据整合为新源域,若否,则返回步骤 2。

步骤 4：通过连续小波变换得到新源域和目标域的时频图像集合。

步骤 5：使用新源域和目标域的时频图像集训练 Resnet，计算特征迁移损失以及新源域的分类损失，构建联合特征匹配损失，反馈优化 Resnet。

步骤 6:判断是否达到最大迭代次数,若是,则保存模型参数，得到训练好的 Resnet 网络;若否,则返回步骤 5。

步骤 7：将目标域时频图像集输入已训练好的 Resnet 网络,获得目标域的诊断结果。

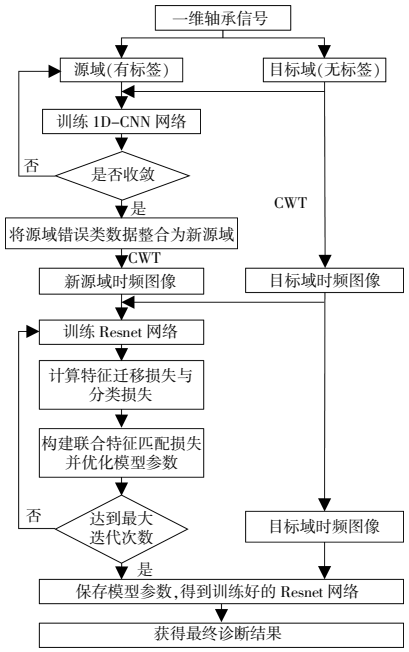



**图 6　基于联合特征匹配的风电机组轴承故障诊断算法流程**

Fig.6 Diagram of wind turbine bearing fault diagnosis method via combined features adaptation

## 2　仿真及结果分析

### 2.1　仿真任务介绍

选取由传动系统模拟实验平台收集的凯斯西储大学(Case Western Reserve University, CWRU)滚动轴承数据集模拟风电机组轴承,采样频率为 12 kHz 和 48 kHz,分别采集电机转速为1 797,1 772,1 750,1 730 r/min 4 种工况下风扇端与驱动端的振动加速度信号。每种工况对应的标签为 0,1,2,3,均包含 10 种状态,状态符号表示为 n(正常),I7(0.007 英寸内圈故障),B7(0.007 英寸滚动体故障),OR7 (0.007 英寸外圈故障),I14(0.014 英寸内圈故障),B14(0.014 英寸滚动体故障),OR14(0.014 英寸外圈故障),I21(0.021 英寸内圈故障),B21 (0.021 英寸滚动体故障),OR21

(0.021 英寸外圈故障)。为涵盖跨域的不同情况,仿真中设置了单源域、双源域和三源域 3 种跨域诊断任务,将数据维数设为 2 048。以单源域 0→1 任务为例,从 0 工况下每种轴承状态中随机采样 1 000 条数据,共计 10 000 条数据,形成源域数据集。对于目标域,选取工况 1,采用相同方法获取目标域数据集。通过类似操作,可分别构建双源域和三源域数据集。为确保数据集的均衡性,从不同源域中所示等比例选取数据用于诊断任务,具体如表 1 所示。

**表 1 CWRU 跨域诊断任务划分**

Table 1 CWRU cross-domain diagnostic task division

| 类型 | 源域工况 | 目标域工况 | 源域数据/条 | 新源域数据/条 | 目标域数据/条 | 任务标记 |
|---|---|---|---|---|---|---|
| 单源域 | 0 | 1 | 10 000 | 10 000 | | 0→1 |
| | | 2 | | 10 000 | | 0→2 |
| | | 3 | | 10 000 | | 0→3 |
| | 1 | 2 | 10 000 | 10 000 | 0:10 000 | 1→2 |
| | | 3 | | 10 000 | | 1→3 |
| | 2 | 3 | 10 000 | 10 000 | 1:10 000 | 2→3 |
| 双源域 | 01 | 2 | 10 000 | 10 000 | | 01→2 |
| | 01 | 3 | 10 000 | 10 000 | 2:10 000 | 01→3 |
| | 02 | 3 | 10 000 | 10 000 | | 02→3 |
| | 12 | 3 | 10 000 | 10 000 | 3:10 000 | 12→3 |
| 三源域 | 012 | 3 | 10 000 | 10 000 | | 012→3 |
| | 013 | 2 | 10 000 | 10 000 | | 013→2 |
| | 023 | 1 | 10 000 | 10 000 | | 023→1 |
| | 123 | 0 | 10 000 | 10 000 | | 123→0 |

针对不同诊断任务,对应14 个不同的新源域进行数据筛选,维度均与源域保持一致,使用 CWT 将所有新源域与目标域数据转化为 224×224 的二维时频图像。

### 2.2 仿真参数设置

仿真使用的深度学习框架为 Pytorch,编程语言为 Python,环境为 Windows11,显卡为 RTX3050Ti,1D-CNN 网络参数设置如表 2 所示。

采用 Adam 优化器,初始学习率为 0.001,最大迭代次数为 1 000,批处理样本数量为 30,激活函数采用 ReLU。

Resnet-50 参数设置如表 3 所示。同样选择 Adam 优化器,初始学习率为 0.001,批处理样本数量为 30,最大迭代次数为 100。其中每个 Block 都代表多个残差块,如 Block1 中含有 3 个结构相同的残差块,每个残差块有 3 层卷积层。

联合特征匹配损失中的参数 $\mu$ 通过本文所提方法对 0→1 任务进行测试确定,由于局部对齐对整个特征匹配过程起到主要影响,因此 $\mu$ 的取值初步选在[0.6,0.9]内,结果如图 7 所示。当 $\mu$=0.6 时,拟合曲线更平滑且最终得到的网络泛化性能更好,因此后续实验中 $\mu$ 均取值为 0.6。

**表 2 1D-CNN 参数设置**

Table 2 1D-CNN parameters setting

| 网络层 | 网络结构参数(卷积核尺寸,数量,步长) | 输出维度 | 补零 |
|---|---|---|---|
| Input | - | 1×2 048 | - |
| Conv 1 | 1×64,16,8 | 16×1×256 | 是 |
| Maxpool 1 | 1×2,-,2 | 16×1×128 | 否 |
| Conv 2 | 1×32,32,8 | 32×1×16 | 是 |
| Maxpool 2 | 1×2,-,2 | 32×1×8 | 否 |
| Conv 3 | 1×5,32,1 | 32×1×8 | 是 |
| Maxpool 3 | 1×2,-,2 | 32×1×4 | 否 |
| Conv 4 | 1×3,64,1 | 64×1×8 | 是 |
| Maxpool 4 | 1×2,-,2 | 64×1×2 | 否 |
| Flatten | 100 | 100×1 | - |

**表 3 Resnet-50 参数设置**

Table 3 Resnet-50 parameters setting

| 网络层 | 参数(尺寸,数量,步长) | 输出维度 | 补零 |
|---|---|---|---|
| Input | - | 3×224×224 | - |
| Conv 1 | 7×7,64,2 | 64×112×112 | 是 |
| MaxPooling | 3×3,-,2 | 64×56×56 | 是 |
| Block 1 | [1×1,64,1<br>3×3,64,1<br>1×1,256,1]×3 | 256×56×56 | 否<br>是<br>否 |
| Block 2 | [1×1,128,1<br>3×3,128,1<br>1×1,512,1]×4 | 512×28×28 | 否<br>是<br>否 |
| Block 3 | [1×1,256,1<br>3×3,256,1<br>1×1,1024,1]×6 | 1024×14×14 | 否<br>是<br>否 |
| Block 4 | [1×1,512,1<br>3×3,512,1<br>1×1,2048,1]×3 | 2048×7×7 | 否<br>是<br>否 |
| Gap | - | 2048×1×1 | 否 |
| Flatten | - | 2048×1 | - |

### 2.3 仿真分析

不同迁移任务下,不同方法的诊断准确率如表 4 所示。对比方法选择传统迁移学习与深度学习+迁移学习。其中:传统迁移学习包括迁移成分分析(Transfer Component Analysis, TCA)法和联合分布适配(Joint Distri-bution Adaptation, JDA)方法[15];深度学习+迁移学习(Deep Learning & Transfer Learning, DL+TL)包括深度域混淆网

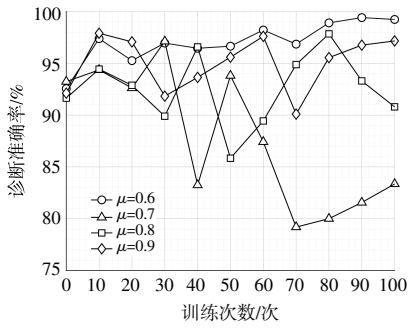



**图 7　模型在不同 $\mu$ 值下的准确率表现**

Fig.7 Model accuracy performance under different $\mu$ values

络(Deep Domain Confusion, DDC)模型[16]、域自适应网络(Domain Adaptive Networks, DAN)模型[17]、域对抗网络(Domain Adversarial Neural Networks, DANN)模型[18]和深度子领域迁移(Deep Subdomain Adaptation Network, DSAN)模型[19]。

本文方法具体架构:TCA 以 MMD 为准则,改变数据结构以缩小数据域间的分布差异;JDA 在 TCA 基础上引进联合分布准则;DDC 基于AlexNet 网络构建, 在第 7 层后加入 MMD 特征适配层; DAN 与 DDC 结构一致, 度量准则替换为MK-MMD, 适配层数由 1 层改为 3 层;DANN 采用 AlexNet 网络中前 5 层网络作为特征提取,后 3 层作为源域分类器, 再添加 3 层全连接层作为域分类器;DSAN 应用 LMMD 对齐子域分布。

**表 4　跨域场景下多种诊断方法的性能比较**

Table 4 Performance comparison of various diagnostic methods in cross-domain scenarios　%

| 迁移任务 | 迁移方法 | | | | | | |
|---|---|---|---|---|---|---|---|
| | TCA | JDA | DDC | DAN | DANN | DSAN | CFA |
| 0→1 | 74.83 | 78.46 | 87.17 | 91.61 | 95.43 | 96.63 | 98.53 |
| 0→2 | 71.40 | 74.36 | 88.09 | 90.19 | 93.90 | 95.28 | 99.94 |
| 0→3 | 72.52 | 76.01 | 89.93 | 92.66 | 94.82 | 97.48 | 99.71 |
| 1→2 | 69.40 | 72.66 | 85.83 | 93.77 | 94.56 | 96.79 | 99.77 |
| 1→3 | 73.33 | 75.09 | 86.26 | 90.76 | 93.68 | 96.12 | 99.94 |
| 2→3 | 75.53 | 77.43 | 87.94 | 93.74 | 94.72 | 97.35 | 99.90 |
| 01→2 | 66.57 | 74.33 | 84.46 | 87.30 | 87.80 | 94.62 | 99.92 |
| 01→3 | 70.60 | 74.03 | 83.46 | 90.79 | 91.54 | 93.39 | 99.53 |
| 02→3 | 65.27 | 70.83 | 85.49 | 88.54 | 89.98 | 91.23 | 100.00 |
| 12→3 | 85.17 | 87.70 | 85.92 | 89.42 | 90.91 | 91.10 | 100.00 |
| 012→3 | 73.93 | 68.80 | 80.27 | 81.50 | 90.22 | 90.66 | 100.00 |
| 013→2 | 78.77 | 76.37 | 83.76 | 85.38 | 87.98 | 90.83 | 100.00 |
| 023→1 | 75.13 | 82.87 | 82.73 | 82.50 | 89.71 | 89.18 | 99.33 |
| 123→0 | 82.00 | 85.30 | 81.91 | 83.49 | 88.71 | 92.45 | 98.20 |
| 平均诊断准确率 | 73.89 | 76.73 | 85.23 | 88.69 | 91.71 | 93.79 | 99.63 |

由表 4 可见, 深度迁移方法整体准确率高于传统迁移方法,原因在于 TCA 与 JDA 方法存在改变数据结构的问题, 且缺乏强大的特征提取能力。在深度迁移方法中,DDC 模型的平均准确率为 85.23%, 比 JDA 提升了 8.5%;DAN 相较 DDC 提升了 3.46%, 表明 MK-MMD 具有更强的分布描述能力, 且多层迁移能够提取出更具可分性的特征;DANN 相较于 DAN 提升了 3.02%, 说明对抗机制能够学习到更好的域不变特征,迁移效果优于基于 MMD 准则的方法; DSAN 的准确率比 DANN 提升了 2.08%, 因为该方法考虑了子域分布损失,LMMD 准则能够识别出更加显著的分布差异,从而实现了更有效的迁移学习。

所提 CFA 方法在不同任务下的平均准确率达到了 99.63%,诊断效果最佳,原因在于其他迁移方法仅关注单一准则,且未考虑负迁移效应。而本方法通过筛选新源域, 并综合考虑边缘分布与条件分布进行特征迁移, 获得了更好的共同特征域,提升了迁移效果。

为验证数据筛选机制的有效性, 删除此模块后在相同任务下进行消融实验,结果如表 5 所示。

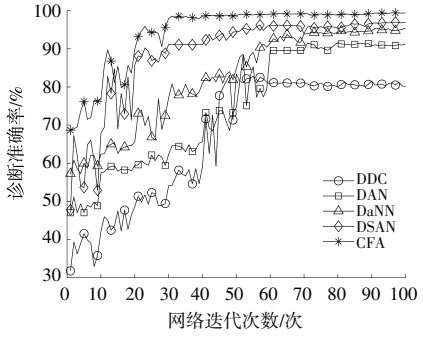
由表 5 可见,删除数据筛选后,本文方法在所有迁移任务上的诊断准确率均有所下降, 与表 4 相比,单源域、双源域及三源域的平均诊断准确率分别下降了 2.62%,4.86%,7.59%,表明在预处理

**表 5 本文方法在跨域无数据筛选条件下的诊断性能**
Table 5 Diagnostic performance of the proposed method under cross-domain conditions without data selection

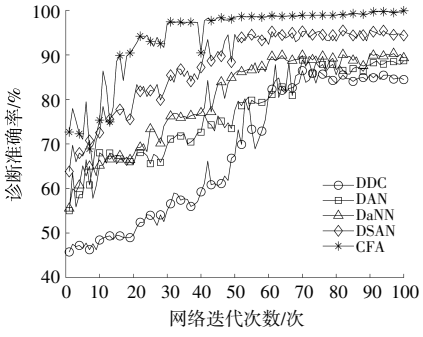
| 任务 | 诊断准确率/% | 平均诊断准确率/% |
|---|---|---|
| 0→1 | 97.10 | 单源域任务:97.01 |
| 0→2 | 96.54 | |
| 0→3 | 98.48 | |
| 1→2 | 95.60 | |
| 1→3 | 98.47 | |
| 2→3 | 95.89 | |
| 01→2 | 95.97 | 双源域任务:95.00 |
| 01→3 | 94.48 | |
| 02→3 | 95.17 | |
| 12→3 | 94.39 | |
| 012→3 | 90.32 | 三源域任务:91.79 |
| 013→2 | 93.01 | |
| 023→1 | 91.42 | |
| 123→0 | 92.42 | |
| 总平均诊断准确率 | 94.95 | |

阶段筛除与目标域差异较大的数据,可以有效提升迁移学习的泛化性能,并且随着源域的增多,效果愈发显著。

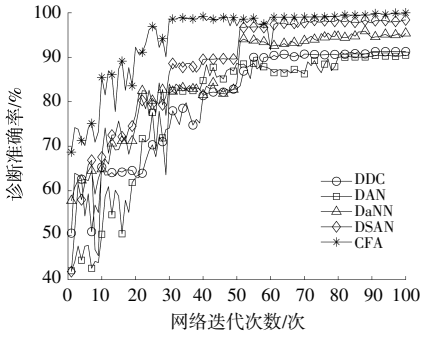
图 8 为深度学习+迁移学习类方法在 0→1,01→2,012→3 任务下的网络性能变化。与其他 4 种方法相比,本文方法在 40 次迭代后开始收敛,拥有更快的收敛速度,且在不同的迁移任务上均具有最高准确率,同时在收敛过程中,本文方法的




(a)0→1 中各方法的诊断性能比较




(b)01→2 中各方法的诊断性能比较



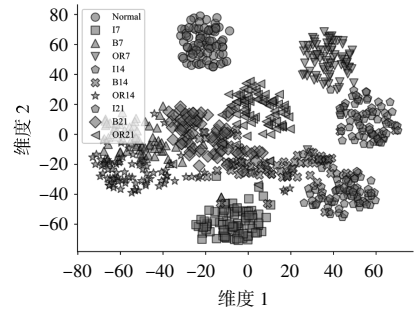

(c)012→3 中各方法的诊断性能比较

**图 8 DL+TL 方法在不同迁移任务下的诊断性能比较**
Fig.8 Comparison of diagnosis performace among DL+TL methods under different migration tasks

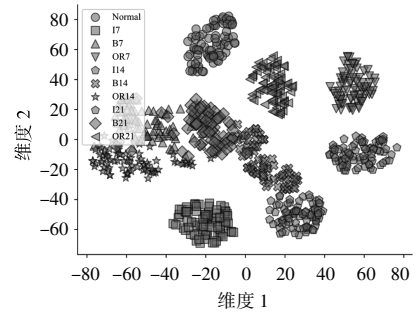
准确率曲线波动较小,说明所建模型的稳定性较好。

为进一步验证本文方法特征的显著性,在有/无迁移条件下,对不同方法得到的特征采用 T-分布随机领域嵌入方法 (T-distributed Stochastic Neighbor Embedding, TSNE)降维,结果如图 9 所示。

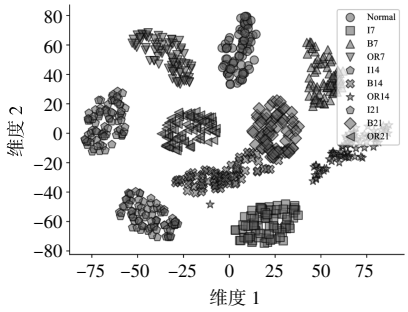
图 9(a)中因未进行特征迁移,Resnet 网络的映射特征很难有效区分相似的类,导致类间交叠明显;图 9(b)中,利用 MK-MMD 进行边缘分布对齐后,源域和目标域的映射特征分布差异明显减小,类间间距增大,但仍存在交叠;图 9(c)中,DSAN 经过 LMMD 条件分布对齐后,基本解决了



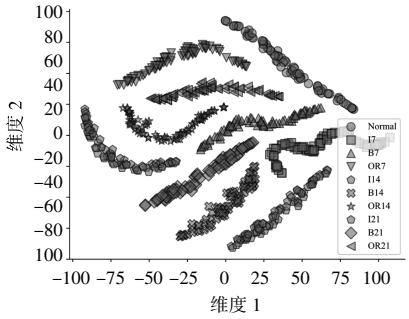

(a)DDS-Resnet 无迁移特征




(b)基于 MK-MMD 迁移的 DAN 特征




(c)基于 LMMD 迁移的 DSAN 特征




(d)基于 CFA 迁移的本方法特征

**图 9　012→3 任务中不同方法的可视化特征**

Fig.9 Visualization features of different methods in task 012→3

上述类间交叠的问题,可将各类明显区分开,但仍有少量特征点混杂其中,难以区分,影响模型的泛化性;图 9(d)中,CFA 从边缘分布与条件分布两种角度同时将特征匹配对齐后，各类故障特征未出现交叠,类间边界清晰。综上可知,本文方法所提损失函数可使网络提取出的映射特征能最大限度区分不同种类的故障，使最终诊断模型泛化性能更好。

### 2.4　泛化性分析

为验证本文方法在其他数据集上的泛化性,选取采集自机械故障模拟实验台的渥太华轴承数据集进行仿真实验。该数据集收集了实验轴承的转速信号及振动信号，本文使用振动信号作为仿真数据。采样频率为 200 kHz,时间为 10 s,包含升速(A)、降速(B)、先升速后降速(C)、先降速后升速(D)4 种工况,每种工况均包含正常(h)、滚动体故障(b)、内圈故障(i)、外圈故障(o)及复合故障(c)5 种类别。采用与 CWRU 数据集相同的数据划分方法以及训练参数设置，构建渥太华跨域诊断任务,具体如表 6 所示。

针对不同诊断任务，对应 8 个不同的新源域内数据筛选,使用 CWT 将新源域与目标域数据转化为224×224 的二维时频图像。利用本文方法对表 6 中的任务进行仿真,结果如表 7 所示。

**表 6　渥太华跨域诊断任务划分**

Table 6 Ottawa cross-domain diagnostic task division

| 类型 | 源域工况 | 目标域工况 | 源域数据/条 | 新源域数据/条 | 目标域数据/条 | 任务标记 |
|---|---|---|---|---|---|---|
| 单源域 | A | B | 5 000 | 5 000 | A:5 000 | A→B |
| | B | C | 5 000 | 5 000 | | B→C |
| | C | D | 5 000 | 5 000 | B:5 000 | C→D |
| | D | A | 5 000 | 5 000 | | D→A |
| 双源域 | AB | C | 5 000 | 5 000 | C:5 000 | AB→C |
| | AC | D | 5 000 | 5 000 | | AC→D |
| 三源域 | ABC | D | 5 000 | 5 000 | D:5 000 | ABC→D |
| | BCD | A | 5 000 | 5 000 | | BCD→A |

**表 7　本文方法在渥太华数据集中的诊断性能**

Table 7 The diagnostic performance of this method in the Ottawa dataset

| 任务 | 诊断准确率/% | 平均诊断准确率/% |
|---|---|---|
| A→B | 95.94 | 单源域任务:97.91 |
| B→C | 98.58 | |
| C→D | 99.00 | |
| D→A | 98.10 | |
| AB→C | 99.06 | 双源域任务:98.95 |
| AD→B | 98.84 | |
| ABC→D | 99.42 | 三源域任务:99.50 |
| BCD→A | 99.58 | |
| 平均诊断准确率 | 98.57 | |

由表 7 可以看出，本文所提方法在大多数跨域诊断任务中表现出较好的诊断效果，诊断结果较为准确。其中仅有单源域中的 A→B 任务诊断效果相对较差，这是由于 A 为升速工况,B 为降速工况所致。在升速过程中,轴承转速逐渐增加,振动频率随之升高,信号中的高频特征成分增多;而在降速过程中,转速逐渐减小,振动频率降低,低频成分更加突出,高频成分则逐渐衰减,导致源域与目标域之间的特征分布差异较大，影响了特征迁移的效果。而本文方法在其他跨域诊断任务中没有受到影响，是由于其他任务工况间的转速变化趋势有相似之处,信号特征的分布较为接近;且在多源跨域诊断任务中，多个源域的结合能够使模型学习到更多有助于分类的特征信息，从而实现更为准确的诊断。因此,在其他迁移任务中,本文方法也能取得较高的诊断准确率。

综上可知,在面对不同环境的数据集时,本文所提方法在跨域诊断任务中的优异表现充分证明了其拥有良好的泛化能力，能够适用于不同的跨

域诊断任务。

## 3 结论

本文针对风电机组中滚动轴承跨域诊断泛化性差的问题，提出了一种基于联合特征匹配的故障诊断方法,并通过 CWRU 与渥太华数据集进行了仿真验证。实验结果表明,所提方法在不同工况间的跨域迁移效果显著，特别是在多源域迁移的场景下,所提出的基于 1D-CNN 的数据筛选方法能够在预处理阶段筛选出源域中与目标域分布相近的数据,可以有效减弱负迁移效应,提升了模型在多源域下的迁移能力。同时,实验结果还证明了相较于传统度量准则，本文提出的联合特征匹配度量方法能够更有效地消除源域与目标域之间的分布差异,进而提高了跨域诊断的准确率。

综上所述，本文所提方法有效解决了传统方法在风电机组轴承跨域诊断中泛化性差的问题，具有较强的实际应用价值。

**参考文献:**


[1] 卢锦玲,张祥国,张伟,等.基于改进辅助分类生成对抗网络的风机主轴承故障诊断[J].电力系统自动化,2021,45(7):148-154.

[2] 江国乾,白佳荣,李文悦,等.基于卷积自编码器的风电机组变桨轴承故障预警与定位 [J]. 可再生能源,2023,41(11):1469-1475.

[3] 宋威,林建维,周方泽,等.基于改进降噪自编码器的风机轴承故障诊断方法 [J]. 电力系统保护与控制,2022,50(10):61-68.

[4] Shen C Q,Wang X,Wang D,et al. Dynamic joint distribution alignment network for bearing fault diagnosis under variable working conditions [J].IEEE Transactions on Instrumentation and Measurement,2021,70:1-13.

[5] 张雅洁,王罗,刘宇璐,等.基于平衡分布自适应迁移学习的多风电机组运行状态监测方法[J].可再生能源,2024,42(8):1068-1073.

[6] Pang S. Stacked maximum independence autoencoders: A domain generalization approach for fault diagnosis under various working conditions [J]. Mechanical Systems and Signal Processing,2024,208(7):111035.

[7] 贾峰,李世豪,沈建军,等.基于深度在线迁移的变负载下滚动轴承故障诊断方法[J].控制与决策,2022,37(6):1521-1530.

[8] Wang Y,Sun X J,Li J,et al. Intelligent fault diagnosis with deep adversarial domain adaptation [J]. IEEE Transactions on Instrumentation and Measurement,2021,70:1-9.

[9] 辛治铖,汪隆君,刘沈全,等.基于迁移学习的海上风电机组轴承早期故障预警策略[J].可再生能源,2024,42(7):915-922.

[10] Lu W N,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.

[11] Yang B,Lei YG,Jia F, et al.An intelligent fault diagnosis approach based on transfer learning from laboratory bearings to locomotive bearings [J]. Mechanical Systems and Signal Processing,2019,122:692-706.

[12] 宋向金,孙文举,刘国海,等.深度子领域自适应网络电机滚动轴承跨工况故障诊断 [J]. 电工技术学报,2024,39(1):182-193.

[13] Zhang P,Du Y,Habetler T G,et al. A survey of condition monitoring and protection methods for medium voltage induction motors [J].IEEE Transactions on Industry Applications,2021,47(1):34-46.

[14] Zhu Y,Zhuang F Z,Wang J D,et al. Deep subdomain adaptation network for image classification [J]. IEEE Transactions on Neural Networks and Learning Systems,2021,32(4):1713-1722.

[15] Zhao H M,Yang X X,Chen B J,et al. Bearing fault diagnosis using transfer learning and optimized deep belief network[J].Measurement Science and Technology,2022,33(6):065009.

[16] Yang B,Lei Y G,Li X,et al. Targeted transfer learning through distribution barycenter medium for intelligent fault diagnosis of machines with data decentralization[J]. Expert Systems with Applications,2024,244:1-15.

[17] Sun R X,Liu X Y,Liu S L,et al. A game theory enhanced domain adaptation network for mechanical fault diagnosis[J].Measurement Science and Technology,2022,33(11):115501.

[18] 胡若晖,张敏,许文鑫.基于 DCGAN 和 DANN 网络的滚动轴承跨域故障诊断[J].振动与冲击,2022,41(6):21-29.

[19] Luo J J, Shao H D, Cao H R, et al. Modified DSAN for unsupervised cross -domain fault diagnosis of bearing under speed fluctuation [J].Journal of Manufacturing Systems,2022,65(1):180-191.

# Wind turbine bearing fault diagnosis method via combined features adaptation

Wang Na[1,2], Wang Zicong[1], Liu Jialin[1]

(1.School of Control Science and Engineering, Tiangong University, Tianjin 300387, China; 2.Key Laboratory of Intelligent Control of Electrical Equipment, Tianjin 300387, China)

**Abstract**: In response to the weak generalization of cross-domain fault diagnosis for rolling bearings in wind turbine units, a Wind Turbine Bearing Fault Diagnosis Method via Combined Features Adaptation is proposed. Firstly, a data selection method based on one-dimensional convolutional neural networks is introduced to filter data from the source domain that is similar to the target domain, forming a new source domain to reconstruct the training dataset and mitigate the negative transfer effect. Secondly, the continuous wavelet transform is applied to obtain two-dimension time-frequency features of the training data, these features are afforded for the residual neural network, which generates one-dimensional feature vectors as the initial features for transfer learning, ensuring comprehensive extraction of the training data's features. Thirdly, a new combined feature adaptation loss function is presented. In it, not only the feature transfer loss of the source domain and the target domain, but also the classification loss of the source domain is considered. This enables effective feature transfer from the source domain to the target domain. Finally, the softmax classifier of the constructed residual neural network is used for the diagnosis classification of the fault. By the simulation under different kinds of migration tasks between source domain and target domain,the proposed model is of higher diagnosis performance in the cross-domain environment.

**Keywords**: wind turbine; rolling bearing; cross-domain diagnosis; deep transfer learning; combined features adaptation