

# 基于多任务学习的电机声信号域自适应故障诊断方法\*

王永淇<sup>1</sup> 肖登宇<sup>2</sup> 胡漫<sup>1</sup> 秦毅<sup>2</sup> 吴飞<sup>2</sup>

(1. 西南大学工程技术学院 重庆 400715; 2. 重庆大学机械与运载工程学院 重庆 400030)

**摘要:** 由于高质量的电机故障数据样本的采集和处理成本过高,新采集的数据样本存在无标注的情况,而域自适应可以借助现有数据对无标注的新数据进行处理识别,因而在故障诊断领域受到了广泛关注。在基于域自适应的电机故障诊断领域,存在两个问题,常用域自适应框架下会出现多任务梯度冲突。同时,现有方法极少研究复杂运行状态之间的迁移任务。因此本文提出了AMDA电机故障诊断方法以解决上述问题。AMDA方法利用多层一维卷积层、批量归一化层和池化层构成的特征提取器,提取源域和目标域的高阶特征;之后结合使用基于对抗的方法和基于分布差异度量的方法,减小源域和目标域数据特征的分布差异;最后引入基于梯度对齐的多任务学习方法,对故障分类器、域判别器和分布差异度量三个任务进行平衡和优化,减小多任务梯度之间的冲突,最终得到基于多任务学习的电机声信号的域自适应故障诊断模型。使用所提出的AMDA方法在多个试验设置下进行跨运行状态故障诊断试验,试验结果表明,AMDA方法在基于声信号的跨运行状态电机故障诊断试验中,完成了稳定运行状态(Stable)、启动运行状态(Start)和循环运行状态(NEDC)之间的迁移任务,最高诊断正确率可达91.47%。同时,AMDA方法在两个对比试验中,性能均显著高于其他方法,具有一定的研究价值和工程应用价值。

**关键词:** 电动机;声信号;故障诊断;域自适应;多任务学习

中图分类号: TH17; TP206<sup>+</sup>.3; TN06 文献标识码: A 国家标准学科分类代码: 460

## Domain-adaptation fault diagnosis method for motor acoustic signals based on multi-task learning

Wang Yongqi<sup>1</sup> Xiao Dengyu<sup>2</sup> Hu Man<sup>1</sup> Qin Yi<sup>2</sup> Wu Fei<sup>2</sup>

(1. College of Engineering and Technology, Southwest University, Chongqing 400715, China;

2. College of Mechanical and Vehicle Engineering, Chongqing University, Chongqing 400030, China)

**Abstract:** The high cost of collecting and processing high-quality motor fault data samples has resulted in the collection of newly unlabeled data samples. Domain adaptation has emerged as a promising approach to process and recognize new unlabeled data with the help of existing data. This has led to a surge of interest in domain adaptation in the field of fault diagnosis. In the field of electric machine fault diagnosis based on domain adaptation, two issues require attention. A conflict arises in the gradients of multiple tasks when employing the conventional domain adaptation framework. And, the existing methods rarely study the migration task between complex states. In light of the aforementioned issues, this paper puts forth AMDA motor fault diagnosis method based on multi-task alignment, with the aim of providing a solution to the aforementioned problems. The AMDA method employs a feature extractor comprising a multi-task one-dimensional convolutional layer, a batch normalization layer, and a pooling layer, to extract the higher-order features of the source and target domains. Subsequently, a combination of an adversarial-based approach and a distributional difference metric-based approach is utilized to reduce the distributional differences of data features. Finally, a multi-task learning approach based on gradient alignment is introduced to balance and optimize the three tasks: fault classifier, domain discriminator, and distributional difference metric. By reducing the conflicting gradients among the tasks, this approach ultimately enables the development of a domain adaptation fault diagnosis model for acoustic signals of electric motors based on multi-task learning. Cross-operational state fault diagnosis tests are conducted under multiple test setups using the proposed AMDA method, and the test results demonstrate that the AMDA method effectively accomplishes the migration task between stable operational state (Stable), start operational state (Start), and European driving cycle state (NEDC) in the acoustic signal. Based on cross-operational state electric motor fault diagnosis tests, the highest diagnosis accuracies reach 91.49%. Furthermore, the performance of AMDA method is significantly higher than that of other methods in the two comparison tests, which offer valuable insights for research and engineering applications.

**Keywords:** motor; acoustic signal; fault diagnosis; domain adaptation; multi-task learning

## 0 引言

随着“中国制造2025”战略不断推进,各类电机驱动的

新能源机械设备在各个领域的应用日益广泛,逐渐成为各个领域中不可或缺的重要组成部分。这些机械设备在运行过程中,由于机械设备的复杂性以及工作条件的多变性,其

中的电机可能会出现各种故障,造成严重后果。因此,及时对电机故障情况进行诊断非常重要。相对于传统基于振动信号的故障诊断技术,对噪声信号进行采集分析能够很好地避免实际作业中加速度传感器由于高温等异常环境而无法安装的问题。声信号的故障诊断方法在工程实际中具有重要的研究价值<sup>[1]</sup>。

随着深度学习领域的不断发展,基于深度学习的故障诊断方法的研究和应用已经兴起<sup>[2]</sup>,但基于机器学习和传统深度神经网络的故障诊断方法<sup>[3-6]</sup>若要获得良好分类效果往往需要大规模且高质量的标注样本。同时,基于传统深度神经网络的故障诊断在部署到工程中进行实际应用时,其可靠性远不如试验条件。此外机械的工作条件(如载荷和速度等)通常各不相同,这也要求使用一种工作条件数据集训练的模型能在另一种工作条件中有良好性能。再者,高质量的电机故障数据的采集成本和处理成本过高<sup>[7]</sup>,新采集的数据也存在少标注或者无标注的情况<sup>[8]</sup>。如何利用已有数据训练得到其他工况可用的故障诊断模型已逐渐成为研究热点。域自适应是迁移学习常用的一种方法,通过减小目标域数据和源域数据之间分布差异的方式,使在有标签的源域数据上训练的模型,在无标签的目标域数据上也能有良好的性能。目前,已有很多研究将域自适应应用在故障诊断领域<sup>[9-13]</sup>。

域自适应的方法可分为基于映射的域自适应方法和基于对抗的域自适应方法等。基于映射的域自适应方法<sup>[13]</sup>通过特征提取器将源域和目标域的实例映射到特征空间,并通过减小源域特征和目标域特征之间的距离的,来减少他们之间的分布差异。基于对抗的域自适应的方法<sup>[14]</sup>使用域鉴别器来减少由特征提取器产生的源域和目标域之间的特征分布差异的方法。目前在故障诊断领域的实际应用中,最常见的域自适应框架是将基于对抗的域自适应方法与基于映射方法结合使用。例如,Li等<sup>[15]</sup>分别使用MMD和领域对抗训练的两个特征提取器和分类器,同时,进一步利用集成学习得到最终结果;Jiao等<sup>[16]</sup>使用联合最大平均分布差异(joint maximum mean discrepancy, JMMD)与域对抗模块结合实现跨域迁移;Guo等<sup>[17]</sup>使用MMD和域对抗方法结合构成域自适应模块,实现了有效的跨域故障诊断。

但在常见的域自适应框架下,在模型训练的过程中需要同时完成最大化域鉴别器的分类误差,最小化特征空间中源域特征和目标域特征之间的距离和最小化源域故障分类器的分类误差等任务。因此需要在训练过程中寻求多个任务之间的平衡。在一般情况下,通过引入超参数来表示不同任务的权重,但这一方式需要大量的计算资源进行调参,同时直接优化跨任务平均目标可能会遇到梯度冲突问题。在常见的域自适应框架在进行故障诊断模型训练过程中梯度冲突问题明显,这样的梯度冲突会破坏训练过程的稳定性,降低整体性能<sup>[18]</sup>。

因此本文将多任务学习(multi-task learning, MTL)的思路引入域自适应问题中。目前的多任务学习方法往往通过优化梯度或损失的方法来实现多任务的梯度下降<sup>[19]</sup>。如:Sener等<sup>[20]</sup>提出将多任务学习看作多目标优化,在低计算资源损耗的情况下,获得了良好的性能。Yu等<sup>[18]</sup>提出了冲突梯度映射法(projecting conflicting gradients, PCGrad),当出现梯度冲突时,将每个任务的梯度投影到另一个任务梯度的法线平面上,以去除冲突成分。也有一些方法使用多任务损失平衡来缓解多任务冲突度问题,如:Liu等<sup>[21]</sup>提出了动态加权平衡(dynamic weight average, DWA)针对不同任务的损失变化率来对不同任务分配权重,来实现多任务平衡。在本文中,使用梯度对齐多任务学习算法(aligned-MTL)<sup>[22]</sup>,该方法基于梯度系统的条件数和稳定性之间的关系,通过对齐梯度矩阵的主成分来减小域自适应中多任务梯度之间的冲突,使训练过程更加稳定。

于此同时,现有基于域自适应的故障诊断方法的任务设置多为稳定运行状态下,不同载荷、不同转速之间的迁移任务<sup>[9-10,23]</sup>。而在工程实际中,电机的运行状态往往是复杂多样的,不仅会在稳定状态下运行,也会面对频繁启停,或者在以一定规律循环运行的情况。这类复杂运行状态之间的迁移任务很少被纳入故障诊断的研究范围。在本文的试验中,设置了稳定(Stable)运行状态,启动(Start)运行状态,以及模拟真实工作场景的欧洲驾驶循环(european driving cycle, NEDC)运行状态之间的迁移任务。运行状态之间的迁移任务不仅更接近实际情况,同时也对模型的泛化能力提出了更高的要求,更具有挑战性。

针对以上问题本文提出了一种基于多任务学习的电机声信号域自适应故障诊断方法,该方法的主要贡献有:

1)提出将基于梯度对齐的多任务学习方法引入域自适应故障诊断,提出基于多任务对齐的域自适应(aligned multi-task domain adaptation, AMDA)电机故障诊断方法,在解决域自适应中多任务的平衡问题,减小多任务梯度之间的冲突,使训练过程更加稳定、减少模型训练调参过程中的计算成本的同时,在基于声信号的电机域自适应故障诊断任务中有优秀的表现。

2)本文所提方法首次完成了稳定(Stable)运行状态,启动(Start)运行状态以及模拟真实工作场景的欧洲驾驶循环(NEDC)运行状态之间的迁移任务,并通过试验证明本文所提方法相较于其他方法能够在此类任务上取得优异性能。

## 1 域自适应和多任务优化理论

### 1.1 域自适应方法

域自适应(domain adaptation, DA)是一种用于减小源域和目标域之间特征分布差异的技术。通过域自适应,在源域上训练的模型也可以很好的应用在目标域上,以减少模型训练时对可以利用的已标注的目标域数据的依赖。域

自适应的方法按照具体实现方式可以分为: 基于映射的方法、基于对抗的方法等。

基于映射的域自适应方法通过将源域数据和目标域数据的特征映射到同一个特征空间, 并尽可能减小二者在该特征空间的分布差异。最大均值差异(maximum mean discrepancy, MMD)是一种经典的分布差异度量。对于源域数据  $D_S = \{(\mathbf{x}_i^S, \mathbf{y}_i^S)\}_{i=1}^{n_S}$  和目标数据  $D_T = \{(\mathbf{x}_j^T)\}_{j=1}^{n_T}$  (其中  $\mathbf{x}_i^S, \mathbf{x}_j^T$  表示源域和目标域样本,  $\mathbf{y}_i^S$  表示源域样本标签,  $n_S, n_T$  表示源域和目标域的样本数量)之间的 MMD 可以表示为:

$$MMD(D_S, D_T) = \left\| \frac{1}{n_S} \sum_{i=1}^{n_S} \phi(\mathbf{x}_i^S) - \frac{1}{n_T} \sum_{j=1}^{n_T} \phi(\mathbf{x}_j^T) \right\|_H^2 \quad (1)$$

式中:  $H$  为重构核希尔伯特空间(reproducing kernel Hilbert space, RKHS),  $\phi(\cdot)$  为将特征映射到  $H$  的非线性映射。

基于对抗的域自适应方法是指通过域鉴别器来减少由特征提取器产生的源域和目标域之间的特征分布差异的对抗性方法。具体而言, 域鉴别器的主要功能是区分来自源域和目标域的特征, 而特征提取器则是尽可能的提取无法被域鉴别器区分的特征, 特征提取器和域鉴别器形成对抗, 实现域混淆, 从而起到提取源域和目标域之间共有的特征<sup>[24]</sup>, 减少源域特征和目标域特征分布差异。

### 算法(1)梯度矩阵对齐

```
Require:  $\mathbf{g}_1, \mathbf{g}_2, \mathbf{g}_3 \dots \mathbf{g}_t \in \mathbb{R}^{|\theta| \times t}$ 
 $\mathbf{G} \leftarrow \{\mathbf{g}_1, \dots, \mathbf{g}_t\}$ 
/* 对梯度矩阵进行奇异值分解 */
 $\mathbf{M} \leftarrow \mathbf{G}^\top \mathbf{G}$ 
 $(\mathbf{\Lambda}, \mathbf{V}) \leftarrow \text{eigh}(\mathbf{M})$ 
 $\mathbf{\Sigma}^{-1} \leftarrow \text{diag}(\sqrt{\frac{1}{\lambda_1}}, \dots, \sqrt{\frac{1}{\lambda_R}})$ 
/* 计算平衡转化 */
 $\mathbf{\Omega} \leftarrow \sqrt{\lambda_R} \mathbf{V} \mathbf{\Sigma}^{-1} \mathbf{V}^\top$ 
/* 梯度对齐 */
 $\mathbf{g}_0 \leftarrow \mathbf{G} \mathbf{\Omega}$ 
return  $\mathbf{g}_0$ 
```

## 1.2 梯度对齐多任务学习

多任务学习的目标是在一个统一的模型中完成多个相关的任务的学习, 以实现任务之间知识共享的相互促进。由于这种学习范式不仅能显著减小模型大小, 提高推理速度, 而且还能提高性能。多任务模型的训练比单独训练的难度要大得多, 这是由于某些任务对模型参数具有支配性影响, 导致其他任务的表现不理想。这一现象与多任务学习中的任务平衡问题有关。

在数值分析中, 线性系统的稳定性可以用该系统矩阵

的条件数来衡量。在多任务学习中, 累积梯度可以看作是多个任务的线性组合:  $\mathbf{g}_0 = \mathbf{G} \mathbf{w}$ 。因此, 多任务学习优化过程中的稳定性, 可以使用梯度矩阵  $\mathbf{G}$  的条件数来衡量, 条件数越小, 稳定性越高。梯度矩阵  $\mathbf{G}$  的条件数等于矩阵最大奇异值和最小奇异值之比。 $\mathbf{G}$  为多个任务梯度结合而成的梯度矩阵  $\hat{\mathbf{G}} = \{\mathbf{g}_1, \dots, \mathbf{g}_t\}$ 。

$$\kappa(\mathbf{G}) = \frac{\sigma_{\max}}{\sigma_{\min}} \quad (2)$$

假设一个具有最小条件数的梯度系统, 根据奇异值分解定理, 其梯度矩阵  $\hat{\mathbf{G}}$  满足:

$$\hat{\mathbf{G}} = \mathbf{U} \mathbf{\Sigma} \mathbf{V}^\top \quad (3)$$

此时, 由于矩阵条件数  $\kappa(\hat{\mathbf{G}}) = 1$ , 可以得  $\mathbf{\Sigma} = \sigma \mathbf{I}$ 。此外, 由于  $\mathbf{U}, \mathbf{V}$  矩阵是标准正交的, 所以单个任务梯度范数等于  $\sigma$ 。因此, 最小化线性梯度系统的条件数可以减少系统内的优势和冲突。另一方面, 如果一个线性梯度系统没有定义, 那么减少梯度冲突和最小化条件数并不完全等价。当这两个问题同时解决, 并且梯度正交时, 条件数才能达到最小值。

结合上述理论, 梯度对齐多任务学习算法(aligned-MTL)<sup>[22]</sup>中, 将解决多任务梯度冲突优势问题等为一个优化问题。具体来说, 找到一个新累积梯度, 使得  $\|\mathbf{g}_0 - \hat{\mathbf{g}}_0\|_2^2$  最小, 而且线性系统是稳定的, 即  $\kappa(\hat{\mathbf{G}}) = 1$ 。对上述问题应用三角不等式, 得到:

$$\|\mathbf{g}_0 - \hat{\mathbf{g}}_0\|_2^2 \leq \|\mathbf{G} - \hat{\mathbf{G}}\|_F^2 \|\mathbf{w}\|_F^2 \quad (4)$$

进一步, 得到以下优化任务:

$$\min_{\hat{\mathbf{G}}} \|\mathbf{G} - \hat{\mathbf{G}}\|_F^2 \quad s.t. \quad \hat{\mathbf{G}}^\top \hat{\mathbf{G}} = \mathbf{I} \quad (5)$$

式中:  $\hat{\mathbf{G}}$  为对齐后的梯度矩阵,  $\mathbf{I}$  为单位矩阵。

式(5)可以被看作一个特殊的 Procrustes 问题, 该问题可以通过执行奇异值分解(singular value decomposition, SVD), 并重新缩放主成分对应的奇异值, 使它们等于最小的奇异值来获得封闭解。对齐后的梯度矩阵的奇异值分解为:

$$\hat{\mathbf{G}} = \mathbf{U} \mathbf{\Sigma} \mathbf{V}^\top \quad (6)$$

式中:  $\mathbf{U}, \mathbf{V}$  为  $\hat{\mathbf{G}}$  的奇异矩阵,  $\mathbf{\Sigma} = \sigma \mathbf{I}, \sigma$  为  $\hat{\mathbf{G}}$  的最小的奇异值。

由式(5)和式(6)进一步可得:

$$\hat{\mathbf{G}} = \sigma \mathbf{U} \mathbf{V}^\top = \sigma \mathbf{G} \mathbf{V} \mathbf{\Sigma}^{-1} \mathbf{V}^\top \quad (7)$$

在模型训练的每一个优化步骤中, 都会执行梯度对齐, 该方法的伪代码如算法(1)所示。

## 2 AMDA 电机故障诊断方法

### 2.1 AMDA 方法结构

为了解决电机域自适应故障诊断中, 跨运行状态迁移效果不佳以及常用域自适应框架下, 多任务梯度冲突问题,

本文提出 AMDA 方法。该方法可以利用已有运行状态下, 电机的声信号故障数据作为源域, 目标运行状态下的无标注声信号数据作为目标域。通过基于 CNN 的特征提取

器, 故障分类器, 域判别器, 分布差异度量和多任务优化器, 得到能够对目标运行状态下的电机故障诊断的模型。AMDA 方法的结构如图 1 所示。

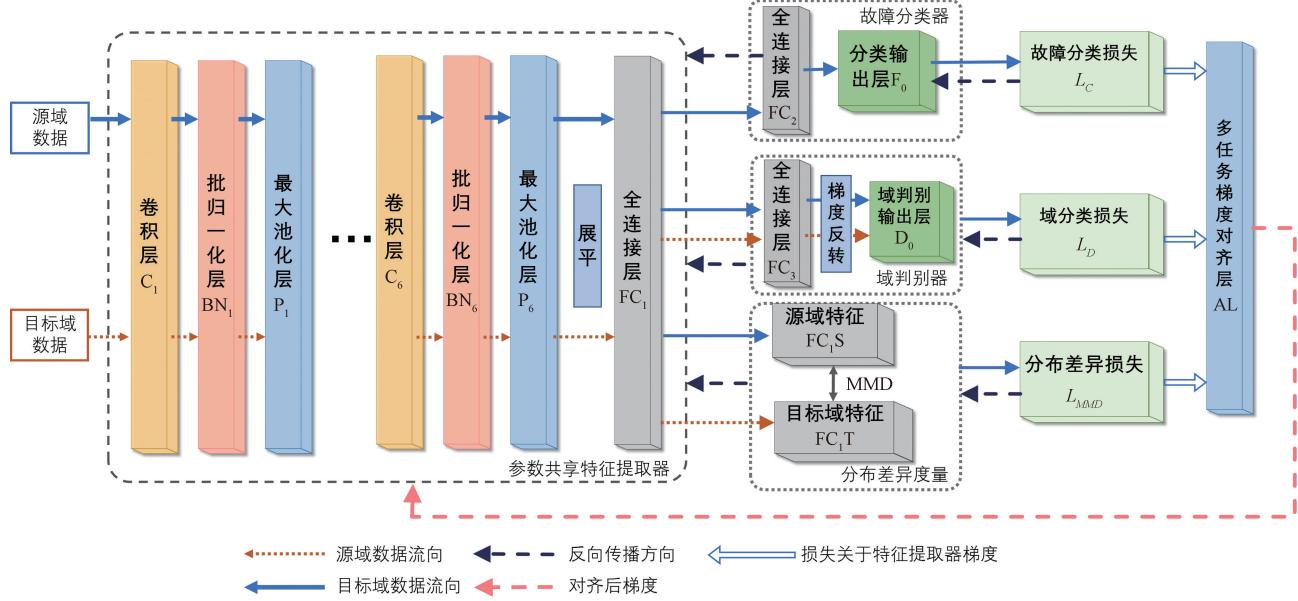


图 1 AMDA 方法模型结构图

Fig. 1 Structure of the AMDA mod

## 2.2 特征提取器

AMDA 模型中的特征提取器由 6 个卷积层-批量归一化层-池化层组的循环结构和 1 层全连接层构成。源域和目标域的输入数据为长度为  $L$  的一维声信号, 因此特征提取器中卷积全部为一维卷积。输入数据在卷积层与卷积核进行运算后得到卷积特征:

$$c_j = \text{Relu} \left( \sum_{i=1}^n w_c * x_{j-m+1:i} + b_c \right) \quad (8)$$

式中:  $*$  为一维卷积算子,  $x_{j-m+1:i}$  为第  $j$  段输入数据,  $w_c$  为卷积核,  $b_c$  为对应的偏置,  $n$  为卷积核的数量,  $c_j$  为卷积层的第  $j$  个输出点,  $\text{Relu}(\cdot)$  为激活函数。

在每一个卷积层后设置批量归一化层。批量归一化可以帮助网络更快地收敛, 缓解梯度消失和梯度爆炸的问题, 提高网络的稳定性。在卷积神经网络中, 卷积层后的批量归一化层在通道维度进行归一化操作, 如式(9)所示。

$$BN(x) = \gamma \frac{x - \mu_B}{\sigma_B} + \beta \quad (9)$$

式中:  $x$  为输入到批量归一化层的小批量数据,  $\mu_B, \sigma_B$  分别为小批量数据的均值和标准差,  $\gamma, \beta$  为批量归一化层的可学习参数, 将在模型训练过程中进行迭代。

在归一化层后设置池化层, 在池化层中进行池化操作, 降低卷积特征的维数, 使学习到的卷积特征具有平移不变性。本文使用最大池化函数, 返回某一子区域内的最大值如下:

$$p_j = \max\{x_{j \times k, (j+1) \times k}\} \quad (10)$$

式中:  $k$  为池化核长度,  $p_j$  为第  $j$  个池化输出。

在经过积层-批量归一化层-池化层的循环结构的输出进行平坦之后, 进入全连接层  $FC_1, FC_1$  的输出即为特征提取器的输出特征。

$$f = \sigma((w_f)^T x_f + b_f) \quad (11)$$

式中:  $w_f$  为两个全连接层之间的权重矩阵,  $b_f$  为相应的偏置,  $x_f$  为输入数据。

## 2.3 故障分类器

基于特征提取器的输出特征  $f$ , 在分类输出层  $F_0$  中, 通过  $\text{softmax}$  回归估计进行故障分类。

$$y = \frac{1}{\sum_{i=1}^K e^{((w_i)^T f + b)}} \begin{bmatrix} e^{((w_1)^T f + b)} \\ e^{((w_2)^T f + b)} \\ \vdots \\ e^{((w_K)^T f + b)} \end{bmatrix} \quad (12)$$

式中:  $w_i$  是第  $i$  个输出神经元的权重矩阵,  $b$  是相应的偏置向量,  $K$  为故障类别。

为了从源域中提取到可以识别故障类型的域不变特征, 故障分类器能够对源域数据的故障进行准确识别, 使源域数据上故障识别分类误差最小化, 对于具有  $K$  个故障类型的数据, 根据交叉熵损失函数, 故障分类损失可以定义为:

$$L_c = -\frac{1}{B} \left[ \sum_{i=1}^B \sum_{j=1}^K I[y_i = k] \log \frac{e^{((w_j)^T f + b)}}{\sum_{l=1}^K e^{((w_l)^T f + b)}} \right] \quad (13)$$

式中:  $B$  为训练样本的批量大小,  $I[\cdot]$  为输入样本的故障类型和输出之间的指示函数。

## 2.4 域判别器

域分类器同样接受来自特征提取器的输出特征,而且与故障分类器不同,域判别器同时接收来自源域和目标域的特征,并尝试对特征来源进行判别。如果域判别器无法区分一个特征的来源,则可以认为该特征是具有域不变性的<sup>[24]</sup>。可见,域判别器的目标是使源域和目标域的判别损失最大化。域判别可以看作一个二分类问题,域判别损失可以定义为:

$$L_d(\mathbf{x}_i) = \frac{1}{B} \sum_{i=1}^B (y_i^d \log d(\mathbf{x}_i) + (1-y_i^d) \log(1-d(\mathbf{x}_i))) \quad (14)$$

式中:  $y_i^d$  表示样本的真实域来源;  $d(\mathbf{x}_i)$  为域判别器的输出。

模型训练时,源域数据和目标域数据均参与域分类器的训练,因此域判别器的损失可以表示为源域数据样本和目标域数据样本的域判别损失相加。若有  $n_s$  个源域样本和  $n_t$  个目标域样本,则域判别器的损失为:

$$L_d = \frac{1}{n_s} \sum_{i=1}^{n_s} L_d(\mathbf{f}_i^{(S)}) + \frac{1}{n_t} \sum_{j=1}^{n_t} L_d(\mathbf{f}_j^{(T)}) \quad (15)$$

式中:  $\mathbf{f}_i^{(S)}, \mathbf{f}_j^{(T)}$  分别代表源域数据样本和目标域数据样本经过特征提取器的输出特征。

## 2.5 分布差异度量

通过特征提取器提取的高阶特征将用于故障分类器和域判别器,因此高阶特征直接关系到故障识别和域自适应的性能。为了减小从源域和目标域提取到的高阶特征之间的分布差异,同时确保故障分类器能够通过域不变的特征在源域和目标域上都有较好的性能,本文直接通过 MMD 测量源域数据样本和目标域数据样本经过特征提取器输出的高阶特征  $\mathbf{f}_i^{(S)}, \mathbf{f}_j^{(T)}$  之间的分布差异。通过高斯径向基函数(Gaussian radial basis function, RBF),作为核函数  $k(\mathbf{f}_i^{(S)}, \mathbf{f}_j^{(T)}) = \exp(-\|\mathbf{f}_i^{(S)} - \mathbf{f}_j^{(T)}\|^2/2\sigma^2)$ , 得到 MMD 的无偏估计  $\hat{D}$ 。

$$\hat{D} = \frac{1}{n_s^2} \sum_{i=1}^{n_s} \sum_{j=1}^{n_s} k(\mathbf{f}_i^{(S)}, \mathbf{f}_j^{(S)}) + \frac{1}{n_t^2} \sum_{i=1}^{n_t} \sum_{j=1}^{n_t} k(\mathbf{f}_i^{(T)}, \mathbf{f}_j^{(T)}) - \frac{2}{n_s n_t} \sum_{i=1}^{n_s} \sum_{j=1}^{n_t} k(\mathbf{f}_i^{(S)}, \mathbf{f}_j^{(T)}) \quad (16)$$

## 2.6 优化目标与多任务梯度对齐层

在 AMDA 方法中,为了实现域自适应,需要完成 3 个优化任务:1)最小化故障分类器的故障分类损失  $L_c$ 。2)最小化特征提取器输出的源域和目标域高阶特征之间的分布差异  $\hat{D}$ 。3)最大化域判别器的域判别损失  $L_d$ 。一般情况下,对于常用的随机梯度下降算法,需要将 3 个优化任务的损失项进行结合。

$$L = L_c - \lambda_d L_d + \lambda_{MMD} \hat{D} \quad (17)$$

设特征提取器的参数为  $\boldsymbol{\theta}_f$ , 故障分类器的参数为  $\boldsymbol{\theta}_c$ ,

域判别器的参数为  $\boldsymbol{\theta}_d$ , 可将式(15)写为:

$$L(\boldsymbol{\theta}_f^*, \boldsymbol{\theta}_c^*, \boldsymbol{\theta}_d^*) = \min_{\boldsymbol{\theta}_f, \boldsymbol{\theta}_c, \boldsymbol{\theta}_d} [L_c(\boldsymbol{\theta}_f, \boldsymbol{\theta}_c) - \lambda_d L_d(\boldsymbol{\theta}_f, \boldsymbol{\theta}_d) + \lambda_{MMD} \hat{D}(\boldsymbol{\theta}_f)] \quad (18)$$

根据式(18)和梯度下降算法,模型参数  $\boldsymbol{\theta}_f, \boldsymbol{\theta}_c$  和  $\boldsymbol{\theta}_d$  的迭代过程如下:

$$\begin{cases} \boldsymbol{\theta}_f \leftarrow \boldsymbol{\theta}_f - \epsilon \left( \frac{\partial L_c}{\partial \boldsymbol{\theta}_f} - \lambda_d \frac{\partial L_d}{\partial \boldsymbol{\theta}_f} + \lambda_{MMD} \frac{\partial \hat{D}}{\partial \boldsymbol{\theta}_f} \right) \\ \boldsymbol{\theta}_c \leftarrow \boldsymbol{\theta}_c - \epsilon \frac{\partial L_c}{\partial \boldsymbol{\theta}_c} \\ \boldsymbol{\theta}_d \leftarrow \boldsymbol{\theta}_d - \epsilon \frac{\partial L_d}{\partial \boldsymbol{\theta}_d} \end{cases} \quad (19)$$

式中:  $\epsilon$  为学习率。

在式(17)中,多个优化任务损失项的结合引入了表示多个任务之间权重的超参数  $\lambda_d$  和  $\lambda_{MMD}$ 。更多超参数的引入,导致需要耗费大量的计算资源来进行超参数调优。在一次训练中,超参数是固定的,不会随着训练的进行而调整。如果在训练过程中,多任务之间的梯度产生冲突或优势的情况,会导致模型难以收敛或者无法收敛到最优。因此引入多任务方法来解决域自适应问题的 3 个任务之间的冲突问题。在每一次参数优化前,分别计算特征提取器的模型参数对于每个任务损失的梯度。

$$\begin{cases} \mathbf{g}_c = \frac{\partial L_c}{\partial \boldsymbol{\theta}_f} \\ \mathbf{g}_d = \frac{\partial L_d}{\partial \boldsymbol{\theta}_f} \\ \mathbf{g}_{MMD} = -\frac{\partial \hat{D}}{\partial \boldsymbol{\theta}_f} \end{cases} \quad (20)$$

将式(20)中  $\mathbf{g}_c, \mathbf{g}_d, \mathbf{g}_{MMD}$  结合组成梯度矩阵  $\mathbf{G} = \{\mathbf{g}_c, \mathbf{g}_d, \mathbf{g}_{MMD}\}$ , 根据算法(1)进行梯度对齐,得到 3 个任务梯度的线性组合  $\mathbf{g}_0 = \mathbf{G}\boldsymbol{\omega}$ , 将  $\mathbf{g}_0$  作为特征提取器参数优化的依据。在这一过程中,通过算法(1)得到了优化后的梯度矩阵,不再需要计算  $L$  来对特征提取器进行优化,因此在缓解了梯度冲突的同时,也不需要再对超参数  $\lambda_d$  和  $\lambda_{MMD}$  进行调优。

因此 AMDA 方法中,模型参数  $\boldsymbol{\theta}_f, \boldsymbol{\theta}_c$  和  $\boldsymbol{\theta}_d$  的迭代过程变为:

$$\begin{cases} \boldsymbol{\theta}_f \leftarrow \boldsymbol{\theta}_f - \epsilon \mathbf{g}_0 \\ \boldsymbol{\theta}_c \leftarrow \boldsymbol{\theta}_c - \epsilon \frac{\partial L_c}{\partial \boldsymbol{\theta}_c} \\ \boldsymbol{\theta}_d \leftarrow \boldsymbol{\theta}_d - \epsilon \frac{\partial L_d}{\partial \boldsymbol{\theta}_d} \end{cases} \quad (21)$$

由此得到 AMDA 故障诊断流程,如图 2 所示。

## 3 试验

### 3.1 试验数据

本文试验数据由 Spectra Quest 公司的 Drivetrain

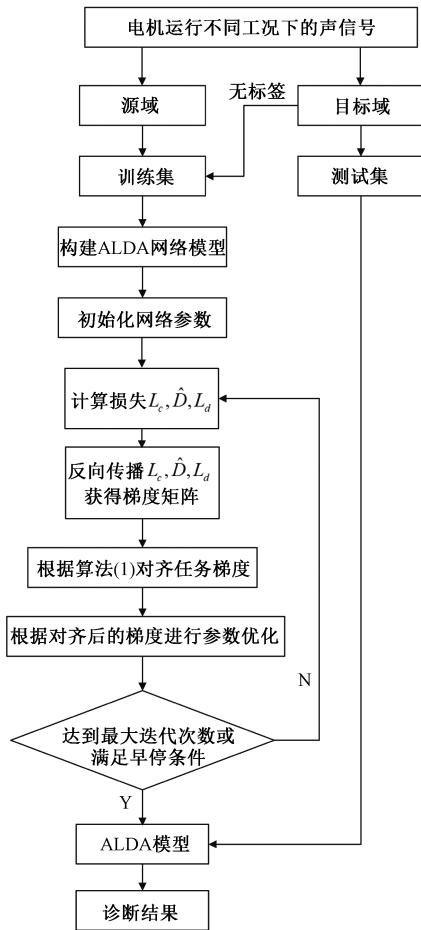


图 2 AMDA 方法诊断流程

Fig. 2 Diagnosis process of AMDA method

Dynamics Simulator 实验平台采集。如图 3 所示,实验台从左至右分别为实验电机,传感器,联轴器,齿轮箱及电磁制动器。采样率为 44.1 kHz。所使用的电动机为 Y 型连接的 3 马力的三相电动机,额定转速 2 850 rpm。测试电机的减速器为行星齿轮箱和正齿轮箱二级减速,行星齿轮箱的减速比为(100/28+1),正齿轮箱的减速比为(90/36×100/29),总减速比为 8 000/203≈39.409。电机定子共有 24 个定子槽,34 个转子导条。测试对象为正常电机和 5 种故障电机,故障类型以及在 NEDC 状态下,不同故障类型电机运行时采集的原始声信号如图 4 所示。

同时设置启动运行状态(Start)、稳定运行状态(Stable)、循环运行状态(NEDC)3 种运行状态。3 种运行状态在无故障下的电机运行时的声信号波形图如图 4 所示。

### 3.2 试验一: 域自适应方法对比

为了验证本文提出的 AMDA 方法的性能,进行了跨运行状态的域自适应试验。同时与同类型的其他域自适应方法进行对比,以验证所提出方法的优越性。

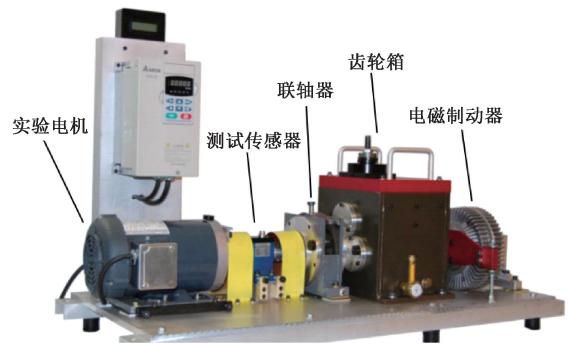


图 3 试验数据采集平台

Fig. 3 Experimental data collection platform

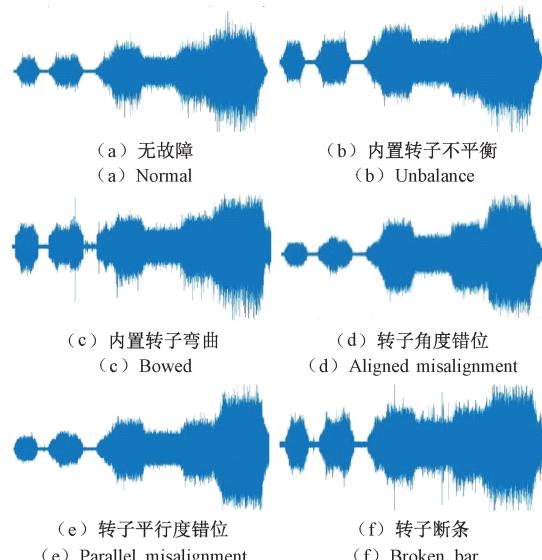


图 4 不同故障类型电机运行原始声信号

Fig. 4 Different fault motor operating original acoustic signal

用于对比试验的方法包括:

- 1) 神经网络的领域对抗训练<sup>[25]</sup> (domain adversarial training of neural networks, DANN)。
- 2) 深度适应网络<sup>[26]</sup> (deep adaptation networks, DAN)。
- 3) 深度卷积迁移学习网络<sup>[17]</sup> (deep convolutional transfer learning network, DCTLN)。

其中,DANN 方法是基于对抗的域自适应方法,通过对抗思想减小域之间的分布差异;DAN 方法使用 MMD 衡量源域和目标域之间的特征分布差异,通过减小 MMD 进行域自适应;DCTLN 方法可以看做 DAN 方法和 DANN 方法的结合,同时使用了基于对抗和基于分布差异度量的域自适应方法。

为了对比域自适应方法的性能,同时尽可能的控制试验中的变量,所有方法使用相同的特征提取器和故障分类器。AMDA,DCTLN 以及 DANN 方法中都使用了域判别器,因此所用的域判别器的结构相同。具体模型参数如

表 1 所示,表 1 中  $k$  表示类别数量。AMDA, DCTLN 以及 DAN 方法中使用的分布差异度量算法相同。根据试验数

据的 3 个运行状态,设计了 6 个跨运行状态的域自适应试验,如图 5 所示,图中箭头方向从源域指向目标域。

表 1 网络结构参数

Table 1 Network structure parameters

网络名称	编号	层类型	参数
特征提取器	$C_1$	卷积层	channels=16, kernel size=320, stride=8, padding=32
	$P_1$	池化层	kernel size=2, stride=2
	$C_2$	卷积层	channels=32, kernel size=160, padding=1
	$P_2$	池化层	kernel size=2, stride=2
	$C_3$	卷积层	channels=64, kernel size=32, padding=1
	$P_3$	池化层	kernel size=2, stride=2
	$C_4$	卷积层	channels=64, kernel size=16, padding=1
	$P_4$	池化层	kernel size=2, stride=2
	$C_5$	卷积层	channels=64, kernel size=16, padding=1
	$P_5$	池化层	kernel size=2, stride=2
故障分类器	FC <sub>1</sub>	全连接层	in features=960, out features=64
	FC <sub>2</sub>	全连接层	in features=64, out features=k
域判别器	FC <sub>3</sub>	全连接层	in features=64, out features=2

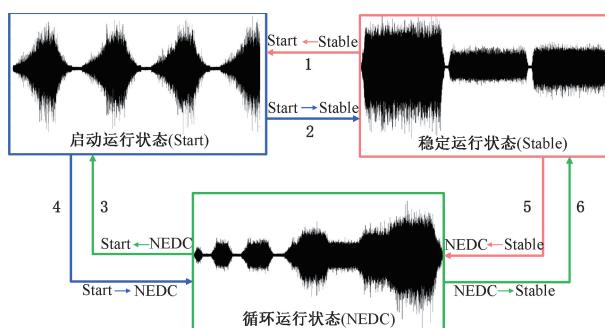


图 5 无故障电机不同运行状态下的声信号波形图与跨运行状态试验设置

Fig. 5 Waveforms of acoustic signals under different operating states of the normal motor and cross-operating state task settings

### 3.3 试验二:多任务方法对比

为了验证本文所用的多任务梯度对齐层的效果,进行了跨运行状态的域自适应试验,并与其它多任务学习方法

进行对比。用于对比试验的方法包括:

1) 动态加权平衡<sup>[21]</sup> (dynamic weight average, DWA): 以不同任务损失为依据,随着训练过程进行动态调整任务权重。

2) 随机损失权重<sup>[27]</sup> (random loss weighting, RLW): 通过在正态分布中采样损失权重,并最小化由损失权重加权的总损失。

3) 冲突梯度投影<sup>[18]</sup> (projecting conflicting gradients, PCGrad): 通过对多任务梯度进行投影避免任务梯度之间的消极影响。

其中, DWA 和 RLW 为基于损失的多任务方法, PCGrad 和本文所用的多任务梯度对齐层为基于梯度的多任务方法。

试验中,所有方法均使用如表 2 所示的模型网络结构,将多任务梯度对齐层替换为其他域自适应方法,与试验一相同,设置了 6 个跨运行状态的域自适应测试,如图 5 所示。

表 2 试验一:域自适应方法平均正确率对比结果表

Table 2 Experiment 1: Comparison of average accuracy of domain adaptation methods

%

任务编号	源域	目标域	AMDA	DCTLN	DANN	DAN
1	Stable	Start	<b><math>87.75 \pm 0.58</math></b>	$73.39 \pm 0.85$	$70.20 \pm 0.81$	$72.18 \pm 0.92$
2	Start	Stable	<b><math>89.55 \pm 0.40</math></b>	$76.87 \pm 0.24$	$75.24 \pm 0.37$	$78.28 \pm 0.89$
3	NEDC	Start	<b><math>82.88 \pm 1.03</math></b>	$73.10 \pm 0.70$	$71.72 \pm 0.72$	$71.57 \pm 0.88$
4	Start	NEDC	<b><math>82.13 \pm 0.93</math></b>	$73.29 \pm 0.50$	$75.54 \pm 0.53$	$75.22 \pm 0.73$
5	Stable	NEDC	<b><math>87.71 \pm 0.78</math></b>	$83.96 \pm 0.49$	$84.88 \pm 0.68$	$85.00 \pm 0.66$
6	NEDC	Stable	<b><math>91.47 \pm 0.76</math></b>	$88.51 \pm 0.20$	$88.03 \pm 0.48$	$85.90 \pm 0.50$

## 4 试验结果分析

根据第 3 节中的试验设置,每一个方法进行 10 次重复试验,根据源域验证集正确率设置训练早停以防止过拟合。本文试验在 Pytorch3.9 环境下,使用 NVIDIA GeForce RTX 3060 显卡进行模型训练。源域和目标域采用相同的数据处理,通过滑窗取样,每一个数据样本含有 0.1 s、采样率为 44.1 kHz 的声信号。使用 Adam 优化器

进行模型训练,试验前,在 Stable 工况下的训练集上,对学习率和批量大小进行超参数调优,确定批量大小为  $B = 64$ , 学习率为  $\epsilon = 5 \times 10^{-5}$ 。

模型训练完成后,在目标域测试集上验证模型性能,计算并统计 3 个对比方法和所提方法的目标域平均正确率和标准差统计如表 2、3 和图 6、7 所示。在每个任务上,根据 T 检验计算 AMDA 方法 10 次重复试验的结果与对比方法试验结果之间的显著性因子 P,如表 4 所示,取显著性水平。

表 3 试验二:多任务方法平均正确率对比结果表

Table 3 Experiment 2: Comparison of average accuracy of multi-task methods

%

任务编号	源域	目标域	AMDA	DWA	RLW	PCGrad
1	Stable	Start	<b>87.75 ± 0.58</b>	77.39 ± 0.14	78.39 ± 1.50	83.19 ± 0.48
2	Start	Stable	<b>89.55 ± 0.40</b>	82.01 ± 0.29	83.24 ± 1.11	81.31 ± 1.38
3	NEDC	Start	<b>82.88 ± 1.03</b>	75.46 ± 0.38	76.17 ± 0.64	79.87 ± 0.88
4	Start	NEDC	<b>82.13 ± 0.93</b>	74.85 ± 2.1	81.79 ± 0.92	83.05 ± 0.81
5	Stable	NEDC	<b>87.71 ± 0.78</b>	81.12 ± 1.6	85.38 ± 1.67	84.80 ± 0.59
6	NEDC	Stable	<b>91.47 ± 0.76</b>	86.64 ± 1.6	88.81 ± 0.46	89.79 ± 1.25

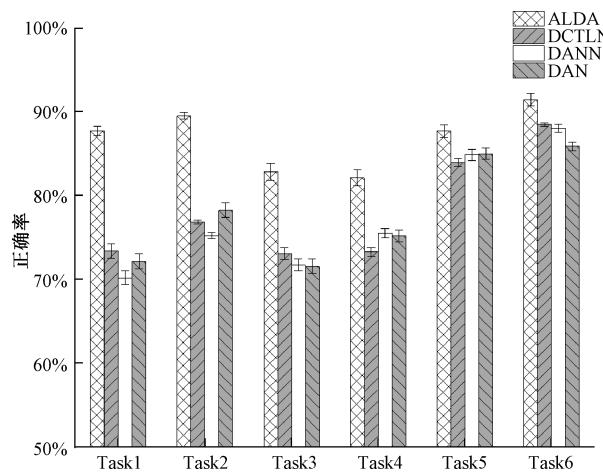


图 6 试验一结果图

Fig. 6 Result graph of experiment 1

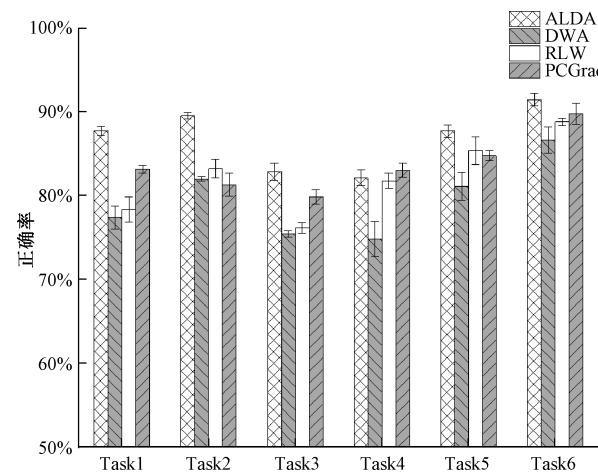


图 7 试验二结果图

Fig. 7 Result graph of experiment 2

表 4 ALDA 方法与对比方法之间的显著性因子  $P(\alpha=0.05)$

Table 4 P value significance factors between ALDA methods and comparison methods ( $\alpha=0.05$ )

任务编号	源域	目标域	DCTLN	DANN	DAN	DWA	RLW	PCGrad
1	Stable	Start	$2.26 \times 10^{-19}$	$3.46 \times 10^{-21}$	$1.42 \times 10^{-19}$	$4.45 \times 10^{-14}$	$9.34 \times 10^{-13}$	$5.23 \times 10^{-13}$
2	Start	Stable	$1.74 \times 10^{-24}$	$2.88 \times 10^{-24}$	$5.96 \times 10^{-18}$	$3.96 \times 10^{-20}$	$4.42 \times 10^{-12}$	$1.28 \times 10^{-12}$
3	NEDC	Start	$6.43 \times 10^{-15}$	$7.28 \times 10^{-16}$	$2.03 \times 10^{-15}$	$8.40 \times 10^{-14}$	$2.47 \times 10^{-12}$	$3.15 \times 10^{-6}$
4	Start	NEDC	$1.99 \times 10^{-15}$	$4.40 \times 10^{-13}$	$1.01 \times 10^{-12}$	$1.97 \times 10^{-8}$	$4.40 \times 10^{-2}$	<b><math>3.95 \times 10^{-2}</math></b>
5	Stable	NEDC	$4.10 \times 10^{-10}$	$1.98 \times 10^{-07}$	$2.85 \times 10^{-7}$	$2.17 \times 10^{-9}$	$1.36 \times 10^{-3}$	$5.62 \times 10^{-8}$
6	NEDC	Stable	$1.47 \times 10^{-9}$	$1.04 \times 10^{-09}$	$4.04 \times 10^{-13}$	$1.39 \times 10^{-7}$	$4.58 \times 10^{-8}$	$2.82 \times 10^{-3}$

结合试验结果和图 5 中的任务设置进行分析可知:

1) 根据试验结果,总体上看,AMDA 方法在除任务 4 以外的所有任务中都表现良好,且从表 4 中,AMDA 方法

的试验结果与对比方法的试验结果之间的显著性因子 P 值可以看出,所有任务上的 P 值都远小于显著性水平  $\alpha$ ,可以认为 AMDA 方法对与其他方法均有显著的优势,显

示出本文所提方法的可靠性。

2) 根据表 2 所示试验一结果, 将 AMDA 方法与其他域自适应方法进行了对比测试。AMDA 方法在 NEDC 工况和 Start 工况之间的迁移测试, 即任务 5、6 中, 取得了较高的平均正确率, 任务 6 中 AMDA 方法的正确率高达 91.47%。但在此任务 5、6 中, AMDA 方法与其他域自适应方法的差距不大, 与之相对应的, 任务 5、6 上的 P 值也相对其他任务较大。结合实际运行情况推测, NEDC 工况中包含了一定量的平稳运行过程, 域间差异相对较小, 导致此测试中 4 种域自适应方法都有较好的效果。

3) 根据试验一的结果, 在任务 1 由于作为源域的 Stable 工况并不包含作为目标域的 Start 工况的部分, 域间差异较大, 因而 AMDA 方法在该任务上领先其他方法幅度最大, 平均正确率分别高出 14.36%、17.55%、15.57%。同时在该任务上的上下限相较于其他方法也较小, 显示出本文所提方法的可靠性。

4) 根据试验一的结果, 通过任务 2、3、4 的试验结果, 结合了对抗和分布差异度量的 DCTLN 方法, 对比基于对抗的 DANN 和基于分布差异度量的 DAN 方法不仅没有显示出应有的优势, 反而在这些任务中的表现上有显著性落后(根据 T 检验结果, DCTLN 方法在任务 2 上, 性能劣于 DAN 方法的显著性因子为  $2.27 \times 10^{-4}$ ; 在任务 4 上, 性能劣于 DAN 和 DANN 的显著性因子分别为  $2.98 \times 10^{-8}$  和  $3.90 \times 10^{-6}$ )。在这些任务上, DCTLN 的故障分类器、域鉴别器、分布差异度量的优化任务之间出现了显著的冲突问题。而在这些任务中, 本文所提方法显著优于其他方法, 可以证明多任务梯度对齐在在训练过程中起到了促进作用。

5) 根据表 3 所示的试验二结果, 将 AMDA 方法中的多任务梯度对齐层替换为其他多任务方法, 并进行了对比测试。除了任务 4 的所有试验中, AMDA 方法的性能都明显优于其他方法, 仅在任务 4 中, 测试正确率落后 PCGrad 方法 0.92%, 从正确率上下限上看, AMDA 方法有更稳定的表现。且根据表 4, PCGrad 方法优于 AMDA 方法的显著性仅为 0.0395, 结合其他任务上的表现, PCGrad 方法的

对比 AMDA 方法的优势并不显著。

6) 根据试验二的结果, 在所有测试中, 基于损失的多任务方法(DWA, RLW), 在性能上落后于基于梯度的多任务方法(AMDA, PCGrad)。可见, 域自适应任务中的多任务冲突并不会单纯的体现在损失值的变化量上。必须要对多任务梯度进行分析操作, 才能有更好的效果。

7) 综合试验一和试验二的结果, 在域自适应任务中引入多任务方法, 能明显提高模型性能。在域间差异相对较大或多任务冲突显著的任务 1、2、3 和任务 4 中, 引入多任务方法后的优势更加明显。同时, 为验证本文所提出的 AMDA 方法的分类精度, 利用 t-SNE 技术<sup>[28]</sup>, 对任务 6 中, 4 种方法提取的目标域高阶特征进行降维和可视化, 如图 8 所示, 图中颜色代表故障类型, 数据点形状表示数据来源。从图中可以看出, 其他域自适应方法的目标域特征数据不同类别之间仍存在混淆, 且会出现多个类别混杂区域。而本文所提出的 AMDA 方法, 多类别混杂情况较少, 各类别分布边界清晰, 效果最好。

8) 为了进一步验证上述 5)、6)、7) 中关于 AMDA 方法的优势, 以任务 2 为例, 计算并记录 AMDA、PCGrad、DCTLN 方法在训练时, 任务梯度矩阵  $g_c, g_d$  之间余弦相似度的变化情况, 如图 9 所示。余弦相似度可以用来衡量任务梯度之间的冲突情况, 当余弦相似度小于 0 时, 认为任务梯度之间出现冲突。明显看出, DCTLN 方法进行模型训练时, 余弦相似度的值呈现不稳定的大幅度震荡, 且出现了较多余弦相似度小于 0 的情况, 看见 DCTLN 在训练时的梯度冲突情况较为严重, 且不稳定。与之相比, AMDA 和 PCGrad 作为基于梯度的多任务方法, 对于整个训练过程梯度冲突的优化非常明显, 不仅减弱了梯度冲突, 而且使训练过程相对稳定。

由图 9 所示, AMDA 方法在训练过程中, 能够将任务梯度之间的余弦相似度的值稳定在 0 附近。而 PCGrad 方法则能在训练初期任务梯度之间的余弦相似度提高, 但随着训练的进行, 余弦相似度又会下降到 0 值以下。从任务梯度之间的余弦相似度的变化反映到模型性能上, PCGrad 在训练初期的拟合速度较快, 但随着训练的进行, 模型由



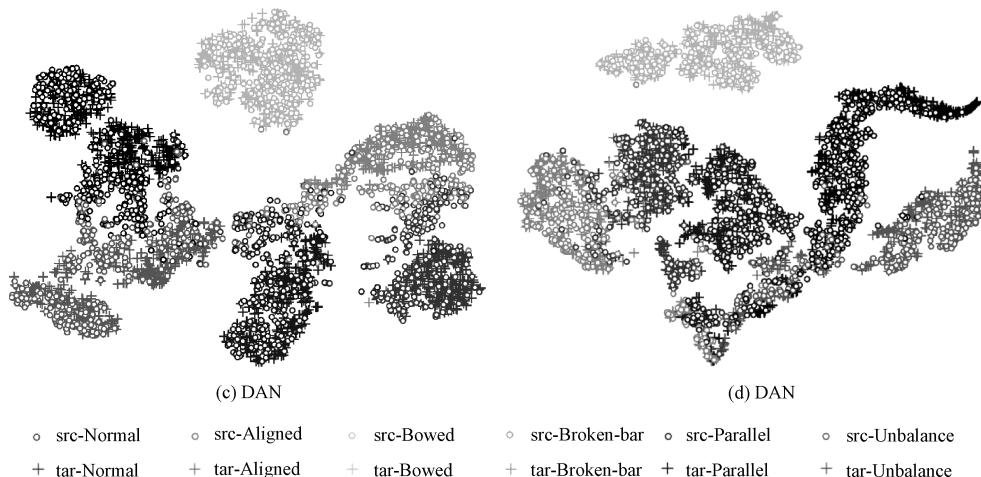


图8 高阶特征 t-SNE 可视化结果

Fig. 8 Higher-order feature t-SNE visualization results

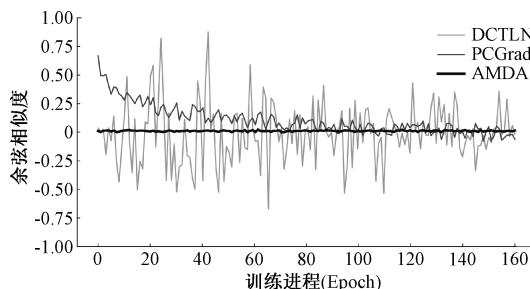


图9 训练过程中任务梯度矩阵余弦相似度

Fig. 9 Cosine similarity of tasks' gradient matrices during training

于梯度冲突的影响难以拟合到更优的位置;而 AMDA 方法则能在整个训练过程中保持梯度冲突水平的稳定,这就使得 AMDA 方法在大多数情况下得到的最终模型性能优于 PCGrad。

## 5 结 论

针对基于域自适应的电机故障诊断中存在的跨运行状态针对基于域自适应的电机故障诊断中存在的跨运行状态迁移效果不佳的问题,以及常用域自适应框架下多任务梯度冲突问题,本文提出了结合多任务学习的域自适应(AMDA)电机故障诊断方法。利用电机故障的声信号数据,将本方法与其他域自适应方法进行对比试验,主要结论如下:

所提方法在在训练过程中,在目标域没有可用数据标签的情况下,完成了启动、稳定、循环3种运行状态之间的域自适应任务。所提方法在此类任务上的最高正确率可达91.47%,显著高于其他与域自适应方法,同时所用多任务梯度对齐层的性能也高于在对比试验中的其他多任务方法。

所提方法在结构上加入到多任务梯度对齐层,虽然该模块的加入并不影响模型本身的复杂度,但是会略微影响

训练时长。AMDA方法相对与DCTLN方法,每一轮训练时长增加了3 s左右;另外,和其他域适应方法相同,AMDA方法仍需要目标域有一定量可用的无标签数据来进行模型训练,后续研究中将进一步摆脱对目标域数据的依赖,提高模型泛化能力。

通过与相同框架方法之间的对比可以看出,由于本文在域自适应任务中引入了多任务学习方法,在训练过程中,对域自适应问题的多个任务进行平衡和优化,减小多任务梯度之间的冲突,在获得更好的模型性能的同时,提高的模型的泛化能力和鲁棒性,验证了在域自适应故障诊断领域引入多任务学习方法的可行性和有效性,具有一定的工程应用前景。但在电机的运行使用过程中,可能面对多种使用场景,存在多种不同的背景噪声的干扰,或其他设备的工作噪声的干扰,在今后的研究中将进一步提升模型的抗干扰能力和鲁棒性,同时在跨运行状态的基础上,挑战跨机故障诊断任务,进一步提高方法的泛化能力。

## 参考文献

- [1] 陈剑,姜涛,陈品.基于声信号的离心泵故障诊断研究[J].电子测量与仪器学报,2024,38(5):169-177.  
CHEN J, JIANG T, CHEN P, et al. Research on fault diagnosis of centrifugal pump based on acoustic signal[J]. Journal of Electronic Measurement and Instrumentation, 2024, 38(5): 169-177.
- [2] 汪欣,毛东兴,李晓东.基于声信号和一维卷积神经网络的电机故障诊断研究[J].噪声与振动控制,2021,41(2):125-129.  
WANG X, MAO D X, LI X D. Motor fault diagnosis using microphones and one-dimensional convolutional neural network [J]. Noise and Vibration Control, 2021, 41(2): 125-129.
- [3] 宫文峰,陈辉,张美玲,等.基于深度学习的电机轴

- 承微小故障智能诊断方法[J]. 仪器仪表学报, 2020, 41(1): 195-205.
- GONG W F, CHEN H, ZHANG M L, et al. Intelligent diagnosis method for incipient fault of motor bearing based on deep learning [J]. Chinese Journal of Scientific Instrument, 2020, 41 (1): 195-205.
- [4] 谢佳琪, 尤伟, 沈长青, 等. 基于改进卷积深度置信网络的轴承故障诊断研究[J]. 电子测量与仪器学报, 2020, 34(2): 36-43.
- XIE J Q, YOU W, SHEN CH Q, et al. Bearing fault diagnosis based on improved convolution deep belief network[J]. Journal of Electronic Measurement and Instrumentation, 2020, 34(2): 36-43.
- [5] 范勇, 王鹏, 张宁超. 基于改进 EMD 与 SOM 神经网络的自动机驱动机构故障诊断研究[J]. 国外电子测量技术, 2020, 39(8): 5-10.
- FAN Y, WANG P, ZHANG N CH. Research on automaton driving mechanism fault diagnosis based on improved EMD and SOM neural network[J]. Foreign Electronic Measurement Technology, 2020, 39(8): 5-10.
- [6] 李永战, 谢磊, 夏政, 等. 基于 SVM 的风电机组偏航系统声学检测方法[J]. 电子测量技术, 2019, 42(17): 169-173.
- LI Y ZH, XIE L, XIA ZH, et al. Acoustic detection method of wind turbine yaw systems based on SVM[J]. Electronic Measurement Technology, 2019, 42(17): 169-173.
- [7] 宗子杨, 何军, 宦海, 等. 基于欧式距离对偶的对抗性无监督域适应算法研究[J]. 电子测量技术, 2023, 46(14): 95-101.
- ZONG Z Y, HE J, HUAN H, et al. Learning on the euclidean discrepancy dual for unsupervised domain adaptation[J]. Electronic Measurement Technology, 2023, 46(14): 95-101.
- [8] ZHAO ZH B, ZHANG Q Y, YU X L, et al. Applications of unsupervised deep transfer learning to intelligent fault diagnosis: A survey and comparative study[J]. IEEE Transactions on Instrumentation and Measurement, 2021, 70:1-28.
- [9] 王军辉, 雷文平, 刘华杰, 等. 基于深度动态域适应的轴承故障诊断研究[J]. 振动与冲击, 2023, 42: 245-250.
- WANG J H, LEI W P, LIU H J, et al. Bearing fault diagnosis based on deep dynamic domain adaptation[J]. Journal of Vibration and Shock, 2023, 42 (14): 245-250.
- [10] 张猛, 王波, 徐浩, 等. 伪标签驱动局部子空间对齐的跨域故障诊断方法[J]. 振动与冲击, 2023, 42(20): 105-113.
- ZHANG M, WANG B, XU H, et al. A cross-domain fault diagnosis method based on pseudo-label driving local subspace alignment[J]. Journal of Vibration and Shock, 2023, 42(20): 105-113.
- [11] 黄星华, 吴天舒, 杨龙玉, 等. 一种面向旋转机械的基于 Transformer 特征提取的域自适应故障诊断[J]. 仪器仪表学报, 2022, 43(11): 210-218.
- HUANG X H, WU T SH, YANG L Y, et al. Domain adaptive fault diagnosis based on Transformer feature extraction for rotating machinery[J]. Chinese Journal of Scientific Instrument, 2022, 43 (11): 210-218.
- [12] 余博, 梁伟阁, 秦奋起, 等. 基于双识别器对抗的开放域自适应故障诊断方法[J]. 仪器仪表学报, 2023, 44(7): 325-334.
- SHE B, LIANG W G, QIN F Q, et al. Open set domain adaptation method based on adversarial dual classifiers for fault diagnosis[J]. Chinese Journal of Scientific Instrument, 2023, 44(7): 325-334.
- [13] 雷杨博, 朱智勤, 柴毅, 等. 基于联合分布偏移差异的跨域滚动轴承故障诊断方法[J]. 电子测量与仪器学报, 2022, 36(10): 146-156.
- LEI Y B, ZHU ZH Q, CHAI Y, et al. Cross-domain fault diagnosis method of rolling bearings based on joint distribution offset difference [J]. Journal of Electronic Measurement and Instrumentation, 2022, 36(10): 146-156.
- [14] 张金汇, 谢林柏. 基于改进域对抗网络的轴承故障迁移诊断方法[J]. 电子测量与仪器学报, 2022, 36(11): 107-115.
- ZHANG J H, XIE L B, Bearing fault migration diagnosis method based on improved domain adversarial network[J]. Journal of Electronic Measurement and Instrumentation, 2022, 36(11): 107-115.
- [15] LI Y B, YAN S, LEI J, et al. Intelligent fault diagnosis by fusing domain adversarial training and maximum mean discrepancy via ensemble learning[J]. IEEE Transactions on Industrial Informatics, 2020, 17(4): 2833-2841.
- [16] JIAO J Y, ZHAO M, LIN J, et al. Residual joint adaptation adversarial network for intelligent transfer fault diagnosis[J]. Mechanical Systems and Signal Processing, 2020, 145: 106962.
- [17] GUO L, LEI Y G, XING S B, 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.
- [18] YU T H, KUMAR S, GUPTA A, et al. Gradient surgery for multi-task learning [J]. *Advances in Neural Information Processing Systems*, 2020, 33: 5824-5836.
- [19] ZHANG Y, YANG Q. A survey on multi-task learning [J]. *IEEE Transactions on Knowledge and Data Engineering*, 2021, 34(12): 5586-5609.
- [20] SENER O, KOLTUN V. Multi-task learning as multi-objective optimization [J]. *Advances in Neural Information Processing Systems*, 2018, 31, DOI: 10.48550/arXiv.1810.04650.
- [21] LIU S K, JOHNS E, DAVISON A J. End-to-end multi-task learning with attention [C]. *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2019: 1871-1880.
- [22] SENUSHKIN D, PATAKIN N, KUZNETSOV A, et al. Independent component alignment for multi-task learning [C]. *2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. Vancouver, BC, Canada: IEEE, 2023.
- [23] 范永胜, 丁雪, 邓艾东. 基于深度条件子域自适应网络的轴承跨域故障诊断研究 [J]. *动力工程学报*, 2024, 44(1): 62-67, 83.
- FAN Y SH, DING X, DENG AI D. Research on cross-domain fault diagnosis of bearings based on deep conditional subdomain adaptive network [J]. *Journal of Chinese Society of Power Engineering*, 2024, 44(1): 62-67, 83.
- [24] BEN-DAVID S, BLITZER J, CRAMMER K, et al. A theory of learning from different domains [J]. *Machine Learning*, 2010, 79: 151-175.
- [25] GANIN Y, LEMPITSKY V. Unsupervised domain adaptation by backpropagation [C]. *International Conference on Machine Learning*, PMLR, 2015: 1180-1189.
- [26] LONG M S, CAO Y, WANG J M, et al. Learning transferable features with deep adaptation networks [C]. *International Conference on Machine Learning*, PMLR, 2015: 97-105.
- [27] LIN B J, YE F Y, ZHANG Y, et al. Reasonable effectiveness of random weighting: A litmus test for multi-task learning [J]. *ArXiv preprint arXiv: 2111.10603*, 2021.
- [28] 陈剑, 程明. 基于 tSNE-ASC 特征选择和 DSmT 融合决策的滚动轴承声振信号故障诊断 [J]. *电子测量与仪器学报*, 2022, 36(4): 195-204.
- CHEN J, CHENG M. Fault diagnosis of rolling bearing acoustic vibration signal based on tSNE-ASC feature selection and DSmT evidence fusion [J]. *Journal of Electronic Measurement and Instrumentation*, 2022, 36(4): 195-204.

## 作者简介

王永淇,硕士研究生,主要研究方向为旋转机械智能故障诊断方法。

胡嫚(通信作者),副教授,硕士生导师,主要研究方向为智能监测与诊断系统、力学与计算方法。

E-mail:humanyyyes@swu.edu.cn