

模型无关元迁移学习用于空间滚动轴承 寿命阶段识别

李统一¹, 李 锋¹, 汤宝平², 汪永超¹

(1. 四川大学机械工程学院, 四川 成都 610065; 2. 重庆大学机械传动国家重点实验室, 重庆 400044)

摘要: 针对变工况下空间滚动轴承寿命阶段识别时因样本分布差异较大、可训练用寿命阶段样本较少以及不同寿命阶段样本数量不均等所造成的寿命阶段识别准确率较低的问题, 提出模型无关元迁移学习(Model-Agnostic Meta-Transfer Learning, MAMTL)用于空间滚动轴承寿命阶段识别。在MAMTL中, 将模型无关元学习和迁移学习相结合以实现多任务同步平行训练从而代替传统的迭代训练, 多个任务损失函数利用不同工况下无类标签样本和历史工况下少量有类标签样本共同更新MAMTL网络参数, 以寻求网络参数的全局最优解, 这使MAMTL具有更好的泛化能力, 因此MAMTL在较少历史工况有类标签训练样本情况下比传统迁移学习具有更好的域适配性; 在MAMTL中构建新型原型网络以将历史工况每一类别的样本表示为一个原型, 通过计算当前工况待测样本与原型的相似度完成当前工况待测样本分类, 且该分类过程无需参数学习, 因此可防止样本不均等情况下对于不同类别样本识别精度差距较大和在少量有类标签训练样本情况下网络出现过拟合的问题, 从而更好提高分类精度。MAMTL的以上优势使得它可利用空间滚动轴承历史工况下的少量、非均等已知寿命阶段的训练样本对当前工况待测样本进行较高精度的寿命阶段识别。空间滚动轴承寿命阶段识别实例验证了该方法的有效性。

关键词: 寿命阶段识别; 空间滚动轴承; 原型网络; 元学习; 迁移学习

中图分类号: TB114.33; TH133.33 **文献标志码:** A **文章编号:** 1004-4523(2023)05-1457-12

DOI: 10.16385/j.cnki.issn.1004-4523.2023.05.030

引 言

空间飞行器的机械部件能否正常运转、实现预定功能和达到预期服役寿命, 很大程度上取决于飞行器中各机械部件内滚动轴承的寿命和可靠性。由于空间在轨环境中难以获取轴承运行状态信息, 空间环境与地面常规环境之间存在较大区别, 处于空间环境下的滚动轴承要承受微重力、高真空、高能粒子辐照、高低温交变等极端恶劣环境的影响, 地面常规环境测试得到的滚动轴承运行特性并不能全面反映滚动轴承在空间环境下的运行性能^[1]。为保证安装到空间飞行器中的滚动轴承具有良好状态, 目前国内外相关研究机构包括欧洲航天局、美国国家航空航天局等通常采取在地面模拟空间环境开展空间滚动轴承寿命鉴定试验来识别空间滚动轴承的寿命阶段, 再对被识别为正常状态(即未产生精度失效)的空间轴承进行剩余寿命预测, 以实现从大量的候

选空间滚动轴承中筛选出最优寿命(即剩余寿命较长)轴承安装到空间飞行器中。

在地面模拟空间环境下, 空间滚动轴承的失效一般是由于固体润滑膜的磨损造成的精度失效, 相比于典型故障或早期故障状态, 其有效寿命阶段内的振动特征更为微弱。其次, 模拟空间环境的设备(如真空泵)空间狭小, 多组轴承同时试验, 导致空间轴承振动信号受到强烈的环境噪声干扰, 这些干扰成分与空间轴承真实振动信号的低频段分量混叠、耦合。空间滚动轴承一般工作在变工况(即变加速应力)下, 模拟空间环境下的滚动轴承往往采取径向无加载、轴向加载方式开展加速寿命试验, 轴承振动受轴向载荷大小影响明显, 随着轴承磨损, 轴承间隙改变, 轴向载荷将发生持续性的变化。同时, 由于真空室容量的限制, 难以实施电机转速的闭环控制, 轴承转速存在波动^[1-2]。以上地面模拟空间环境下空间滚动轴承加速寿命试验的特殊性使空间滚动轴承的运行、退化和失效过程与地面常规环境下的传统

收稿日期: 2021-10-30; **修订日期:** 2022-01-15

基金项目: 机械传动国家重点实验室开放基金资助项目(SKLM-T-KFKT-201718); 中国博士后科学基金第60批面上资助项目(2016M602685); 四川省中国制造2025四川行动资金项目计划(智能制造新模式应用)项目(2019CDYB-12)。

滚动轴承有明显差异,导致地面模拟空间环境下空间滚动轴承寿命阶段识别具有较大的挑战。

目前地面模拟空间环境下空间滚动轴承寿命阶段识别的研究刚刚起步,有为数不多的研究案例。如:陈仁祥等^[1]采用线性局部切空间排列算法(Linear Local Tangent Space Alignment, LLTSA)和最近邻分类器(K-Nearest Neighbors Classifier, KNNC)相结合的方法进行模拟空间环境下航天轴承不同寿命阶段的识别;Dong等^[2]结合沙利斯熵KPCA(Tsalis Entropy-KPCA, TE-KPCA)和优化模糊C均值模型(Optimized Fuzzy C-means Model, OFCM)进行模拟空间环境下空间滚动轴承寿命阶段识别;Miao等^[3]采用支持向量机(Support Vector Machine, SVM)和模糊逻辑推理(Fuzzy Logic Inference, FLI)来进行航空轴承寿命状态预测。然而,以上涉及到的机器学习方法都是基于概率分布一致性假设,而模拟空间环境下空间滚动轴承加速寿命试验是在变工况条件下进行的,基于分布一致性假设的机器学习方法在变工况条件下泛化能力较差,因此不是特别适用于地面模拟空间环境下的空间滚动轴承寿命阶段识别。以上所有机器学习方法都需要大量的有类标签的历史工况(即源域)数据来进行训练且要求各类训练样本数量必须均等,然而受试验成本和试验周期限制,地面模拟空间环境试验条件下往往仅能获得部分历史工况下的少量空间滚动轴承全寿命样本数据用于分类模型的训练,且空间滚动轴承不同寿命阶段的时间跨度的不均等因素往往造成各个寿命阶段的样本数量也不均等,以上复杂的地面模拟空间环境试验条件也决定了上述机器学习方法用于地面模拟空间环境下空间滚动轴承寿命阶段识别存在一定的局限性。

迁移学习(Transfer Learning, TL)^[4-5]为变工况下的空间滚动轴承寿命阶段识别提供了全新解决思路,然而传统迁移学习面对少样本和样本不均等情况时在泛化性能、分类性能等方面仍然存在不足。元学习(Meta Learning, ML)^[6-7]在解决少样本分类问题和提高网络泛化能力方面有很好的效果,在提高迁移学习网络的迁移性能方面具有潜力。原型网络(Prototype Network, PN)^[8]在小样本和不同类别样本不均等这两情况下对样本分类的精度都比一般神经网络更高。因此本文结合迁移学习、元学习和原型网络的各自优势,提出了一种新型无监督的迁移学习理论——模型无关元迁移学习(Model-Agnostic Meta-Transfer Learning, MAMTL)用于地面模拟空间环境下空间滚动轴承寿命阶段识别。

1 模型无关元迁移学习

MAMTL由内环平行网络、外环元学习网络和原型网络组成。内环平行网络和外环元学习网络都由相同的任一迁移学习网络构成。首先借助空间滚动轴承无类标签的源域和目标域样本对内环平行网络进行同步训练得到多任务所有的损失函数;然后通过多任务所有的损失函数的共同作用来训练外环元学习网络得到该网络参数的全局最优解作为该网络参数的初始值,这样使外环元学习网络具有更好的泛化能力;然后用空间滚动轴承目标域无类标签样本和少量的源域有类标签样本来参与训练外环元学习网络,以对外环元学习网络参数进行微调,使外环元学习网络具备小样本跨域迁移学习能力;最后用所构建的新型原型网络作为分类器,通过求空间滚动轴承目标域样本与每一类原型的相似度完成对目标域无类标签样本的类判别(即寿命阶段识别),MAMTL的结构框架如图1所示。

1.1 内环平行网络基础学习

记空间滚动轴承的源域数据集为 S_D ,目标域数据集为 T_D 。设置一次训练中内环平行网络有 N 个任务,记为 $T = \{T_i | i = 1, \dots, N\}$ 。每个任务 T_i 将数据分为支持集(记为 x_i)和查询集(记为 z_i),随机选取来自空间滚动轴承源域 S_D 的 K 个无类标签样本和来自目标域 T_D 的 K 个无类标签样本分别作为一个源域支持集 x_i^S 和一个目标域支持集 x_i^T ;同理,随机选取来自源域 S_D 的 K 个无类标签样本和来自目标域 T_D 的 K 个无类标签样本分别作为一个源域查询集 z_i^S 和一个目标域查询集 z_i^T ,即 $x_i^S = \{x_{ij}^S | j = 1, \dots, K\}$, $x_i^T = \{x_{ij}^T | j = 1, \dots, K\}$, $z_i^S = \{z_{ij}^S | j = 1, \dots, K\}$, $z_i^T = \{z_{ij}^T | j = 1, \dots, K\}$ 。

每个任务分别在具有相同初始值的内环平行网络中训练。每个内环平行网络由相同的任一迁移学习网络组成,令迁移学习网络的特征映射函数为 $F(\bullet)$,该网络的参数集为 θ ,该网络的分布差异度量函数为 $G(\bullet)$ 。假设第 m 次训练时外环元学习网络的参数初始值集合为 θ_m ,同时将其作为 N 个任务对应的内环平行网络的参数初始值集合。在任务 T_i 中,先将支持集样本 x_i^S, x_i^T 输入该任务所对应的迁移学习网络特征映射函数,以分别提取得到高维特征 $f_{x_i^S}, f_{x_i^T}$,该过程表达如下:

$$f_{x_i^S}(\theta_m) = F(x_i^S, \theta_m) \quad (1)$$

$$f_{x_i^T}(\theta_m) = F(x_i^T, \theta_m) \quad (2)$$

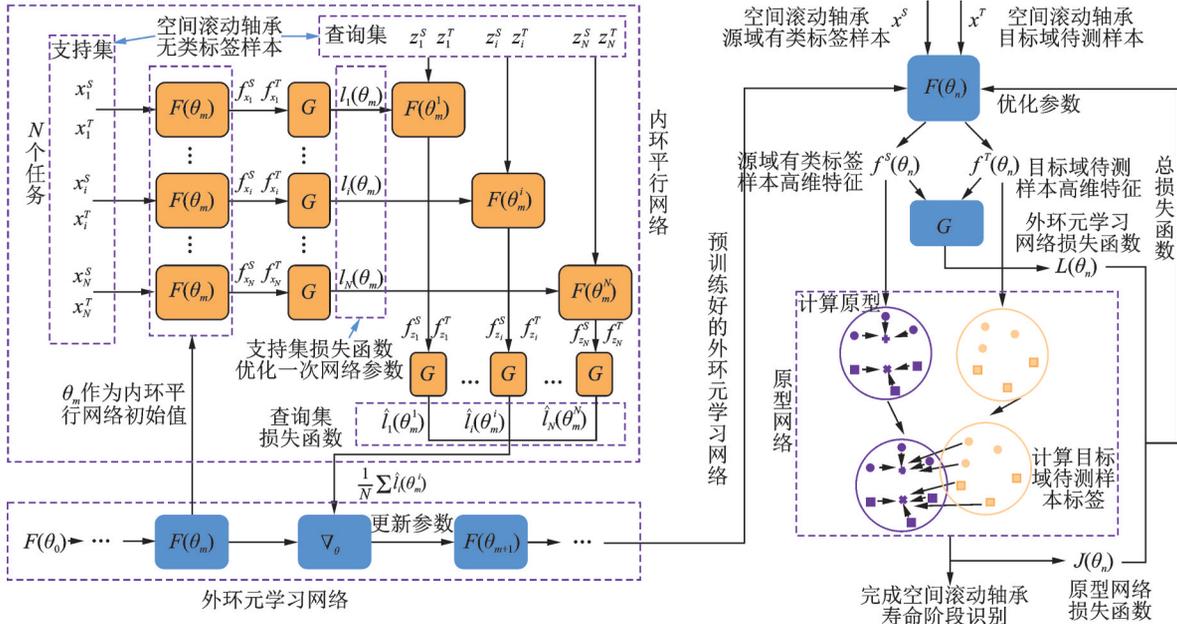


图1 MAMTL的结构框架图

Fig. 1 Structure frame diagram of MAMTL

提取到高维特征后,通过分布差异度量函数 $G(\bullet)$ 来构造如下支持集高维特征的损失函数:

$$l_i(\theta_m) = G(f_{x_i}^s(\theta_m), f_{x_i}^t(\theta_m)) \quad (3)$$

通过该损失函数优化迁移学习网络参数以实现源域样本高维特征和目标域样本高维特征分布差异的最小化,进而更好地实现从源域向目标域的跨域迁移。

得到支持集高维特征的损失函数后用随机梯度下降法对该迁移学习网络的参数 θ_m 进行一次更新,该更新过程如下:

$$\begin{aligned} \theta_m^i &= \theta_m - \alpha \nabla_{\theta} l_i(\theta_m) = \\ & \theta_m - \alpha \nabla_{\theta} G(f_{x_i}^s(\theta_m), f_{x_i}^t(\theta_m)) = \\ & \theta_m - \alpha \nabla_{\theta} G(F(x_i^s, \theta_m), F(x_i^t, \theta_m)) \end{aligned} \quad (4)$$

式中 α 为内环平行网络参数的学习率; ∇_{θ} 表示对参数集 θ 求偏导数。

于是,每个任务分别在对应的内环平行网络内依据式(1)~(4)来更新迁移学习网络参数,得到该任务更新后的内环平行网络参数集 θ_m^i 。

然后将每个任务的查询集 z_i^s, z_i^t 输入到对应更新好参数的内环平行网络中以提取出高维特征,该过程表达如下:

$$f_{z_i}^s(\theta_m^i) = F(z_i^s, \theta_m^i) \quad (5)$$

$$f_{z_i}^t(\theta_m^i) = F(z_i^t, \theta_m^i) \quad (6)$$

接下来,再用分布差异度量函数来构建查询集高维特征的损失函数,可以得到:

$$\begin{aligned} \hat{l}_i(\theta_m^i) &= G(f_{z_i}^s(\theta_m^i), f_{z_i}^t(\theta_m^i)) = \\ & G(F(z_i^s, \theta_m^i), F(z_i^t, \theta_m^i)) \end{aligned} \quad (7)$$

于是, N 个任务经过相对应的内环平行网络训

练后分别得到 N 个不同的查询集高维特征损失函数 $\hat{l}_1(\theta_m^1), \hat{l}_2(\theta_m^2), \dots, \hat{l}_i(\theta_m^i), \dots, \hat{l}_N(\theta_m^N)$ 。

1.2 外环元学习网络

将 N 个任务的查询集高维特征损失函数的平均值作为外环元学习网络的总损失函数,该过程可由下式表达:

$$\begin{aligned} L &= \frac{1}{N} \sum_{i=1}^N \hat{l}_i(\theta_m^i) = \\ & \frac{1}{N} \sum_{i=1}^N G(F(z_i^s, \theta_m^i), F(z_i^t, \theta_m^i)) \end{aligned} \quad (8)$$

用得到的总损失函数来优化外环元学习网络的参数集 θ_m ,完成一次外环元学习网络的参数更新,该参数优化过程可推导为:

$$\begin{aligned} \theta_{m+1} &= \theta_m - \beta \nabla_{\theta} L = \\ & \theta_m - \beta \nabla_{\theta} \frac{1}{N} \sum_{i=1}^N G(F(z_i^s, \theta_m^i), F(z_i^t, \theta_m^i)) \end{aligned} \quad (9)$$

式中 β 为外环元学习网络参数的学习率。优化得到的参数集 θ_{m+1} 作为下一次训练时内环平行网络的参数初始值集。

重复执行式(1)~(9)所示训练过程,直到将外环元学习网络参数训练至收敛,也就完成了外环元学习网络的预训练。由于外环元学习网络每一次更新的参数都是由内环平行网络中多个任务的所有损失函数共同作用得到的全局最优解,所以以最后更新好的全局最优解(即预训练好的外环元学习网络参数集 θ_n) 作为外环元(迁移)学习网络的起始点(即初始值)去学习新的任务时仅需要少量迭代次数就能使外环元学习网络达到收敛,即又好又快地适应

新的迁移学习任务,因此预训练好的外环元学习网络具有良好的泛化性能和域适配性。

1.3 MAMTL的原型网络

得到预训练好的外环元学习网络参数集 θ_n 后,分别取 K 个有类标签(即已知寿命阶段)的空间滚动轴承源域样本 $\mathbf{x}^s = \{\mathbf{x}_j^s | j \in 1, \dots, K\}$ (它们对应的类标签为 $\mathbf{y}^s = \{y_j^s | j \in 1, \dots, K\}$)和 K 个无类标签(即寿命阶段未知)目标域待测样本 $\mathbf{x}^t = \{\mathbf{x}_j^t | j \in 1, \dots, K\}$ 。将源域有类标签样本 \mathbf{x}^s 和目标域待测样本 \mathbf{x}^t 输入到预训练好的外环元学习网络中得到高维特征分别为 $\mathbf{f}^s(\theta_n), \mathbf{f}^t(\theta_n)$,该过程表达如下:

$$\mathbf{f}^s(\theta_n) = F(\mathbf{x}^s, \theta_n) \quad (10)$$

$$\mathbf{f}^t(\theta_n) = F(\mathbf{x}^t, \theta_n) \quad (11)$$

然后由以上两组高维特征值的分布差异度量函数来构造外环元学习网络的损失函数 $L(\theta_n)$:

$$L(\theta_n) = G(\mathbf{f}^s(\theta_n), \mathbf{f}^t(\theta_n)) = G(F(\mathbf{x}^s, \theta_n), F(\mathbf{x}^t, \theta_n)) \quad (12)$$

设空间滚动轴承源域和目标域全体样本一共有 Q 类(即 Q 个寿命阶段),令 S_q 表示属于第 q 类标签的样本,其中 $q \in 1, \dots, Q, n_q$ 表示属于第 q 类样本的数量。接下来,由源域有类标签样本的高维特征计算每一类的原型,该计算过程如下:

$$\mathbf{c}_q(\theta_n) = \frac{1}{n_q} \sum_{(\mathbf{f}_j^s, \mathbf{y}_j^s) \in S_q} \mathbf{f}_j^s(\theta_n) \quad (13)$$

计算目标域待测样本的高维特征与式(13)所示原型的相似度,并选择相似度最大的那一类原型所对应的类标签作为空间滚动轴承目标域待测样本的预测伪类标签,该过程表达如下:

$$\hat{y}_j^t = q_j = \arg \max_{q \in \{1, \dots, Q\}} d(\mathbf{c}_q, \mathbf{f}_j^t(\theta_n)) \quad (14)$$

式中 $d(\cdot)$ 表示两个向量之间的相似度函数。

计算该目标域待测样本属于伪类标签 q_j 的概率如下:

$$p(\hat{y}_j^t = q_j | \mathbf{f}_j^t(\theta_n)) = \frac{\exp(d(q_j, \mathbf{f}_j^t(\theta_n)))}{\sum_{q=1}^Q \exp(d(q, \mathbf{f}_j^t(\theta_n)))} \quad (15)$$

接下来,将所有目标域待测样本属于其对应的伪类标签概率的负对数之和作为原型网络的损失函数,该损失函数推导如下:

$$J(\theta_n) = - \sum_j^K \lg(p(\hat{y}_j^t = q_j | \mathbf{f}_j^t(\theta_n))) = - \sum_j^K \lg \frac{\exp(d(q_j, \mathbf{f}_j^t(\theta_n)))}{\sum_{q=1}^Q \exp(d(q, \mathbf{f}_j^t(\theta_n)))} \quad (16)$$

于是,整合外环元学习网络的损失函数 $L(\theta_n)$ 和原型网络的损失函数 $J(\theta_n)$ 来共同构建MAMTL的总损失函数如下:

$$L^*(\theta_n) = J(\theta_n) + \gamma \cdot L(\theta_n) = - \sum_j^K \lg \frac{\exp(d(q_j, \mathbf{f}_j^t(\theta_n)))}{\sum_{q=1}^Q \exp(d(q, \mathbf{f}_j^t(\theta_n)))} + \gamma \cdot G(F(\mathbf{x}^s, \theta_n), F(\mathbf{x}^t, \theta_n)) \quad (17)$$

式中 γ 为外环元学习网络的平衡约束参数,用于约束外环元学习网络局部寻优行为。使用随机梯度下降法将MAMTL的总损失函数训练至收敛,完成对外环元学习网络的参数微调,此时得到外环元学习网络对该任务的最优参数 θ^* ,也即完成对MAMTL的训练。

最后,再将空间滚动轴承源域有类标签样本和目标域待测样本输入训练好的MAMTL网络,计算出目标域待测样本的类标签 y_j^t (即寿命阶段),以完成元迁移学习全过程,该过程表达如下:

$$y_j^t = \arg \max_{q \in \{1, \dots, Q\}} d(\mathbf{f}_j^t(\theta^*), \mathbf{c}_q(\theta^*)) = \arg \max_{q \in \{1, \dots, Q\}} d(F(\mathbf{x}_j^t, \theta^*), \frac{1}{n_q} \sum_{(\mathbf{f}_j^s, \mathbf{y}_j^s) \in S_q} F(\mathbf{x}_j^s, \theta^*)) \quad (18)$$

由以上推导可知,为MAMTL构建的新型原型网络是将源域每一类别中所有的样本用一个原型表示,通过计算目标域待测样本与原型的相似度完成目标域待测样本的分类,这样可以避免因源域不同类别样本数量的差异而造成对不同类别样本分类精度差别过大(即对少样本类别的样本的分类精度过低)问题;同时,在计算目标域待测样本与不同原型的相似度时没有参数学习的过程,所以在训练小样本情况下不会出现过拟合现象。

2 基于MAMTL的空间滚动轴承寿命阶段识别

所提出的基于MAMTL的空间滚动轴承寿命阶段识别方法的实现过程如图2所示:

(1)选择空间滚动轴承无类标签的源域和目标域样本进行预训练,一次训练为内环平行网络设置 N 个学习任务,将每个任务的样本分为源域支持集、目标域支持集、源域查询集、目标域查询集。

(2)将每个任务的支持集样本输入与该任务对应的内环平行网络提取到高维特征,然后构建支持集高维特征的损失函数,以优化一次内环平行网络的参数。

(3)将每个任务的查询集输入更新参数后的与该任务对应的内环平行网络中提取高维特征,再将N个任务的查询集高维特征损失函数的平均值作为外环元学习网络的总损失函数用于优化外环元学习网络参数集,完成一次外环元学习网络的参数更新。重复(1)~(3)过程,直到外环元学习网络参数训练至收敛,完成外环元学习网络的预训练。

(4)将空间滚动轴承源域有类标签样本和目标域待测样本输入预训练好的外环元学习网络得到各自的高维特征,进而计算外环元学习网络的损失函

数;然后由源域有类标签样本的高维特征得出每个类别的原型;最后计算目标域样本高维特征与每个原型的相似度以得到目标域样本的伪类标签,进而得到原型网络的损失函数。

(5)由外环元学习网络的损失函数与原型网络的损失函数共同构建MAMTL的总损失函数,将该总损失函数训练至收敛,完成对外环元学习网络参数微调。

(6)用训练好的MAMTL完成对目标域待测样本的分类,即完成对空间滚动轴承的寿命阶段识别。

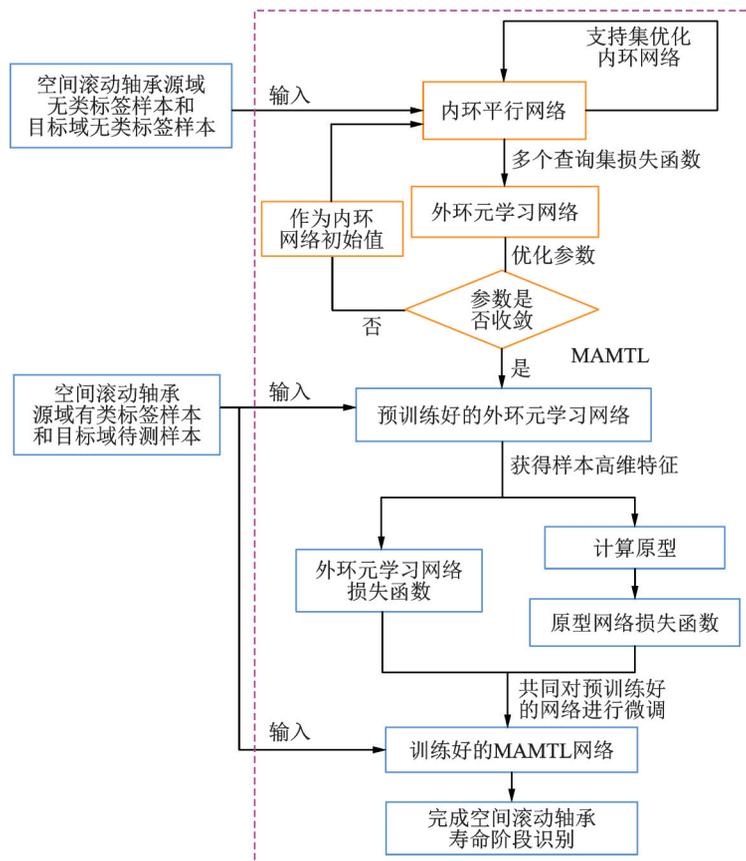


图 2 基于MAMTL的空间滚动轴承寿命阶段识别方法实现流程

Fig. 2 Implementation process of life stage identification method of space rolling bearings based on MAMTL

3 实例分析

3.1 实验装置

实验数据主要来自自主搭建的空间滚动轴承振动监测(即加速寿命试验)平台上采集的地面模拟真空环境下空间滚动轴承寿命试验数据。该平台如图3所示,主要由真空泵、轴承振动监测实验台架、压电式加速度传感器、双积分信号调理器、NI数据采集卡和计算机组成,将型号均为C36018的空间滚动轴承1和2安装于真空泵中的振动监测实验台架

上。在真空环境下轴承1,2均被加载7 kg的轴向预载(随着轴承磨损加剧,轴向载荷会发生持续性变化),并分别在约1000 r/min和3000 r/min的2种非平稳转速下运行(即表1中的非稳态工况C1和C2)。压电式加速度传感器对这两个空间滚动轴承进行振动监测并每隔2 h以25.6 kHz的采样频率采集一次它们的振动加速度信号,直至这两个轴承都完全止动失效。截取每1024个连续的振动加速度数据点作为一个样本,最终采集到这两个空间滚动轴承全寿命期的总样本个数均为744个。

实验中被标记为工况C3的实验数据来自Cin-

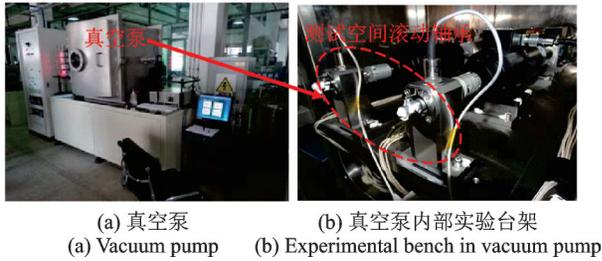


图3 空间滚动轴承振动监测平台

Fig. 3 Vibration monitoring platform for space rolling bearings

cinnati大学的滚动轴承加速寿命试验数据^[9]。如图4所示,将四个型号为ZA-2115双列滚子轴承安装在轴承试验台的旋转轴上,使用转速为2000 r/min的电机通过皮带驱动转轴,并通过弹簧机构在转轴和轴承上施加2721.55 kg (6000 lbs)的径向载荷,采样频率为20 kHz,每10 min采集一次轴承的振动加速度数据。对每次采集的加速度数据截取前1024个连续点作为一个样本,共获得全寿命期984个样本。

表1 实验工况表
Tab. 1 Detailed working conditions

工况	实验数据来源	转速/ (r·min ⁻¹)	载荷	监测对象
C1	自行搭建的模拟空间环境下空间	约1000	7 kg (轴向预载)	空间轴承 1
C2	滚动轴承振动监测平台	约3000	7 kg (轴向预载)	空间轴承 2
C3	Cincinnati大学	2000	2721.55 kg (6000 lbs) (径向载荷)	1号 滚动轴承

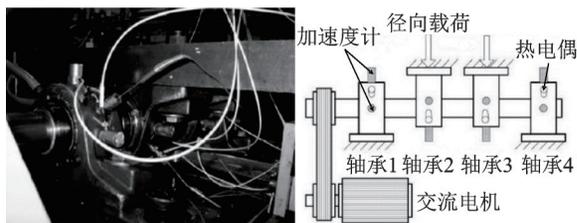


图4 Cincinnati大学滚动轴承加速寿命退化实验台
Fig. 4 Accelerated rolling bearings life cycle degradation test platform in University of Cincinnati

在寿命阶段识别实验开始前需要对三个工况下的全寿命数据进行寿命阶段划分。首先对每个样本提取来自时域、频域和时频域的27个特征^[10],然后用等度量映射(Isometric Mapping, Isomap)^[11]方法对提取的特征进行维数约简,获得一维的主特征,接着用威布尔分布模型^[12]对一维主特征构建可靠度评

估曲线,得到三个工况下的(空间)滚动轴承的可靠度评估曲线分别如图5~7所示。由于真空泵空间狭小,且两个空间轴承在真空泵内同时试验,导致空间轴承运行时的振动特征受到强烈的环境噪声干扰;同时,空间轴承所受轴向载荷随着轴承磨损和轴承间隙的改变而持续变化;且电机转速是非闭环控制,因而空间滚动轴承的转速具有一定波动性。以上原因造成空间滚动轴承的时域、频域和时频域振动特征的能量分布变得更为分散,进而Isomap对以上振动特征的维数约简结果也随之变得更为分散,导致图5和6所示空间滚动轴承的可靠性评估曲线较图7所示的普通滚动轴承波动更大。根据可靠度评估曲线,将全寿命数据划分为正常阶段、早期退化阶段、中期退化阶段、完全失效阶段这4个阶段:从空间滚动轴承运行安全性、稳妥性考虑,将第一次出现可靠度0.9对应的时间点作为划分正常阶段和早期退化阶段的时间点,该点也被视为空间滚动轴承精度失效阈值点^[13];将第一次出现可靠度0.5对应的时间点作为划分早期退化阶段和中期退化阶段的分界点^[14];将第一次出现可靠度0.1对应的时间点作为划分中期退化阶段和完全失效阶段的分界点^[14]。

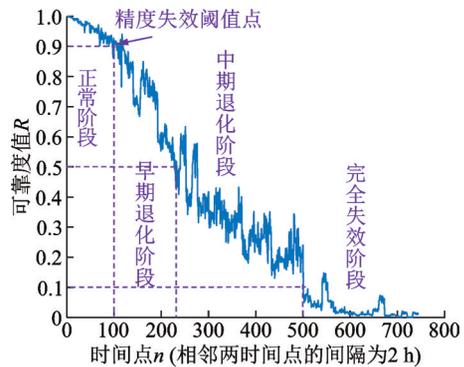


图5 空间滚动轴承1的可靠度评估曲线
Fig. 5 Reliability assessment curve of space rolling bearing 1

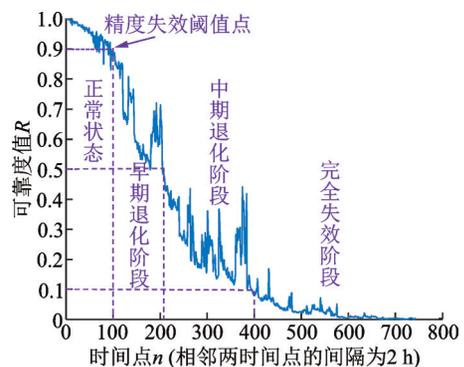


图6 空间滚动轴承2的可靠度评估曲线
Fig. 6 Reliability assessment curve of space rolling bearing 2

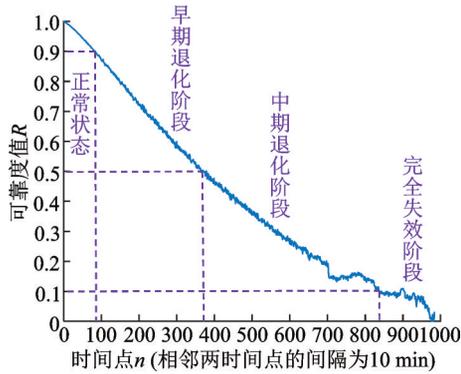


图7 来自 Cincinnati 大学的第二组实验中1号滚动轴承的可靠度评估曲线

Fig.7 Reliability assessment curve of rolling bearing 1 in the second set of experiments from University of Cincinnati

3.2 MAMTL网络结构设计和参数设置

MAMTL的网络结构设计如下:内环平行网络和外环元学习网络的特征映射函数都采用五层的卷积神经网络,该网络的具体配置如表2所示;分布差异度量函数采用联合分布适配函数^[15];原型网络中的相似度函数采用如下的余弦相似度:

$$\cos(a, b) = \frac{a \cdot b}{\|a\| \cdot \|b\|} \quad (19)$$

式中 a, b 表示向量。

表2 五层卷积神经网络结构表

Tab.2 Structure of five-layer convolutional neural network

层名	各层配置
卷积层1	卷积核尺寸为 3×3
最大池化层1	核尺寸为 2×2
卷积层2	卷积核尺寸为 3×3
最大池化层2	核尺寸为 2×2
卷积层3	卷积核尺寸为 3×3
最大池化层3	核尺寸为 2×2
卷积层4	卷积核尺寸为 3×3
最大池化层4	核尺寸为 2×2
卷积层5	卷积核尺寸为 3×3
最大池化层5	核尺寸为 2×2

MAMTL的参数设置如下:内环平行网络参数的学习率 $\alpha = 4 \times 10^{-2}$;外环元学习网络参数的学习率 $\beta = 2 \times 10^{-4}$;外环元学习网络的平衡约束参数 $\gamma = 0.5$;每一次训练内环平行网络的任务数 $N = 8$;每个任务中支持集样本和查询集样本的数量分别为 K (每个任务包含两个阶段:第一阶段是求出支持集高维特征的损失函数并更新一次内环平行网络的

参数;第二阶段是求出查询集高维特征的损失函数)。MAMTL的网络结构和参数设置好后,在以下所有实验中均维持不变。

3.3 实验1和分析对比

在本实验中,将工况C2下的正常状态阶段、早期退化阶段、中期退化阶段以及完全止动失效阶段的样本(即全寿命样本)作为源域样本来识别工况C1下的全寿命样本(即目标域样本)的寿命阶段(即:C2→C1)。实验前,分别对空间轴承2和空间轴承1的每一寿命阶段各随机取80个样本作为用于实验的源域各寿命阶段的总样本和目标域各寿命阶段的总样本,即用于实验的源域总样本数和目标域总样本数分别为320个。

(1)在源域按照1:1:1:1的比例为每一寿命阶段随机取 $K/4$ 个样本作为源域有类标签训练样本(即源域所有寿命阶段样本的总数为 $K, K \leq 320$),在目标域中也按每一寿命阶段1:1:1:1的比例随机取样作为目标域待测样本,待测样本总数也为 K 。将每一个样本按其元素的先后顺序分段重组为相应的 32×32 的矩阵作为MAMTL的一个输入样本,然后依据第2节所示的空间滚动轴承寿命阶段识别实现流程来用所提出的基于MAMTL的寿命阶段识别方法进行(工况C1下)空间滚动轴承1的寿命阶段识别。这里将本文所提出的方法对当前目标域待测样本的四种寿命阶段识别准确率及平均识别准确率与其他三种迁移学习方法,即:联合分布适配(JDA)^[15]、深度领域自适应(DDC)^[16]、和改进迁移联合匹配(ETJM)^[17]进行了比较。为了降低随机性带来的误差,每种方法取20次实验结果的平均值作为其最后实验结果(下同)。随着源域有类标签训练样本总数的减少,寿命阶段平均识别准确率对比结果如图8所示。在源域有类标签样本总数 $K = 16$ 时,所提出的方法和三种被比较的方法的寿命阶段识别准确率如图9所示。同时,为了可视化展示MAMTL的迁移和分类性能的优越性,利用t-分布随机邻域嵌入(t-distributed Stochastic Neighbor Embedding, t-SNE)算法^[18]将 $K = 16$ 时MAMTL和其他三种被对比方法提取到的高维特征降维到二维平面,并以散点图的形式呈现在图10~13中。

由图8和9对比结果可知,随着源域有类标签样本总数的减小,四种方法虽然由于训练都不充分导致它们的寿命阶段识别准确率都逐渐下降,但所提

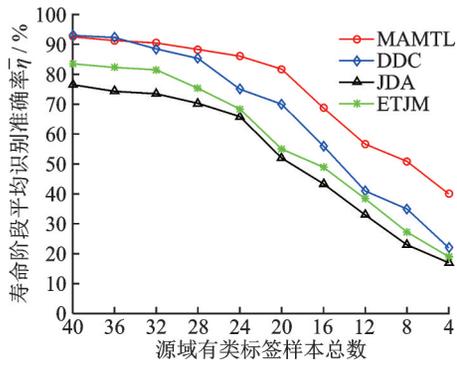


图8 空间滚动轴承1寿命阶段的平均识别准确率对比
Fig.8 Comparison of the average identification accuracy of the life stages of space rolling bearing 1

出的基于MAMTL的寿命阶段识别方法始终能得到比其他三种方法更高的寿命阶段识别精度。由图10~13的对比结果可知,所提出的MAMTL相比其他三种迁移学习方法能使得源域和目标域中相同类别的样本更好地聚合在一起,且两域中不同类别的

样本之间也相对更为分散,因此MAMTL的迁移和分类性能更好,基于MAMTL的寿命阶段识别方法的识别精度也就更高。

(2) 在源域中按照1:1:2:4的比例取正常阶段和早期退化阶段的样本数分别为 $K/8$,取中期退化阶段样本数为 $2K/8$,取完全失效阶段样本数为 $4K/8$ 用作源域有标签样本(即源域所有寿命阶段样本的总数为 $K, K \leq 160$),在目标域中也按每一寿命阶段1:1:2:4的比例随机取样作为目标域待测样本,待测样本总数也为 K 。将本文所提出的方法对当前目标域待测样本的四种寿命阶段识别准确率及平均识别准确率与其他三种迁移学习方法进行了比较。随着源域有类标签训练样本总数的减少,寿命阶段平均识别准确率对比结果如图14所示;在源域有类标签样本总数 $K=24$ 时,所提出的方法和三种被比较的方法的寿命阶段识别准确率如图15所示。

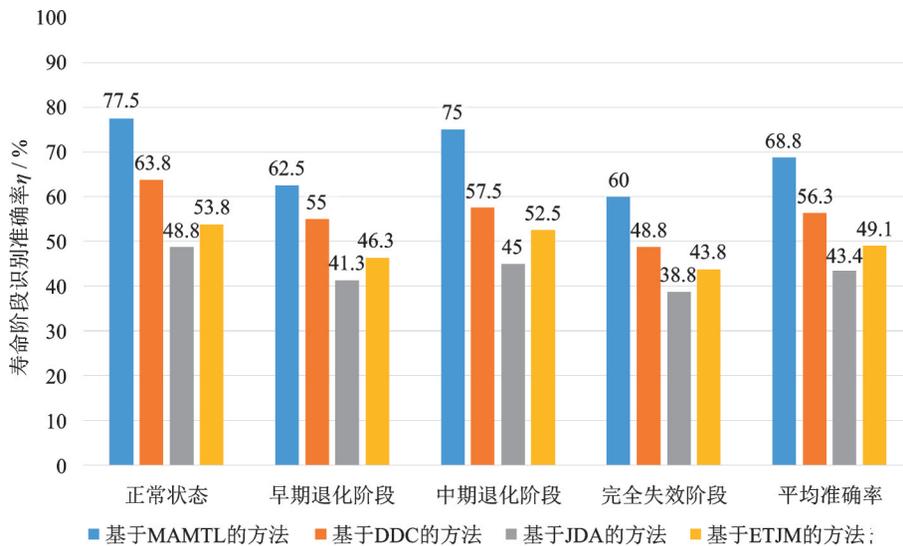


图9 源域有类标签样本总数 $K=16$ 时寿命阶段识别准确率
Fig.9 Life stage identification accuracy of space rolling bearing 1 when the total number of labeled samples in source domain is $K=16$

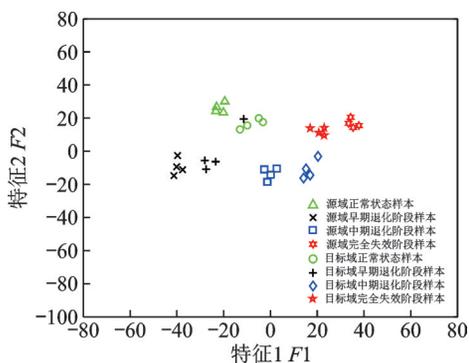


图10 MAMTL输出的高维特征经t-SNE降维后的散点图
Fig.10 Scatter plot of high-dimensional features output by MAMTL after dimension reduction by t-SNE

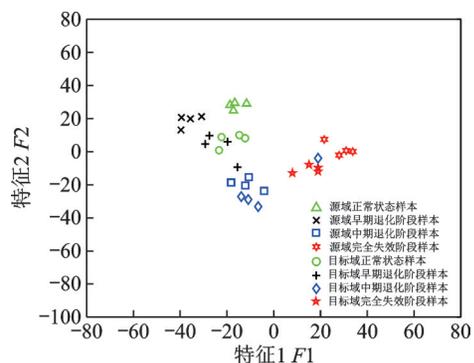


图11 DDC输出的高维特征经t-SNE降维后的散点图
Fig.11 Scatter plot of high-dimensional features output by DDC after dimension reduction by t-SNE

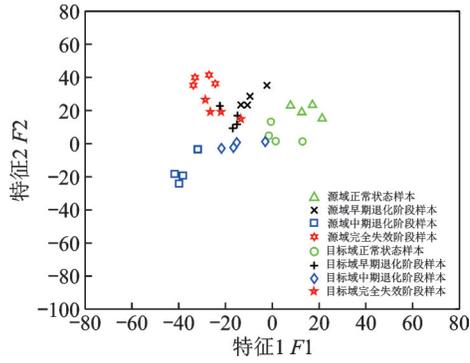


图 12 JDA 输出的高维特征经 t-SNE 降维后的散点图
Fig. 12 Scatter plot of high-dimensional features output by JDA after dimension reduction by t-SNE

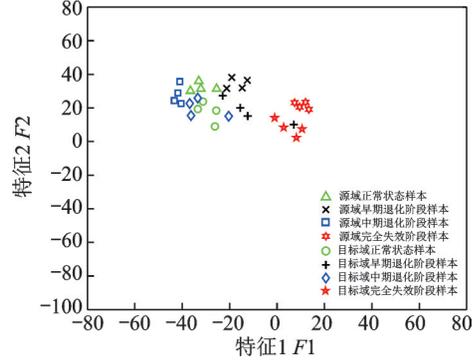


图 13 ETJM 输出的高维特征经 t-SNE 降维后的散点图
Fig. 13 Scatter plot of high-dimensional features output by ETJM after dimension reduction by t-SNE

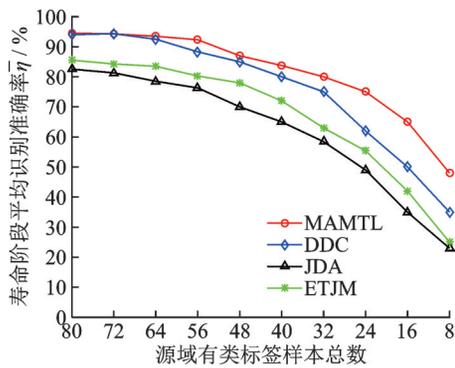


图 14 空间滚动轴承 1 寿命阶段的平均识别准确率对比
Fig. 14 Comparison of the average identification accuracy of the life stages of space rolling bearing 1

由图 14 和 15 的对比结果可知,随着源域有类标签样本总数的减小以及源域不同类标签样本数变得不均等,四种方法由于训练都不充分、不均衡导致它们对四种寿命阶段的识别准确率及平均识别准确率都逐渐下降,但所提出的基于 MAMTL 的寿命阶段

识别方法的识别准确率及平均识别准确率分别还是比其他三种方法更高。

3.4 实验 2 和分析对比

在本实验中,将工况 C3 下的正常状态阶段、早期退化阶段、中期退化阶段以及完全止动失效阶段的样本(即全寿命样本)作为源域样本来识别工况 C1 下的全寿命样本(即目标域样本)的寿命阶段(即: C3→C1)。实验之前,分别对工况 C3 下的 1 号滚动轴承和工况 C1 下的空间轴承 1 的每一寿命阶段各随机取 80 个样本作为用于实验的源域各寿命阶段的总样本和目标域各寿命阶段的总样本,即用于实验的源域总样本数和目标域总样本数分别为 320 个。

在源域中按照 4:3:2:1 的比例取正常阶段的样本数为 4K/10,取早期退化阶段的样本数为 3K/10,

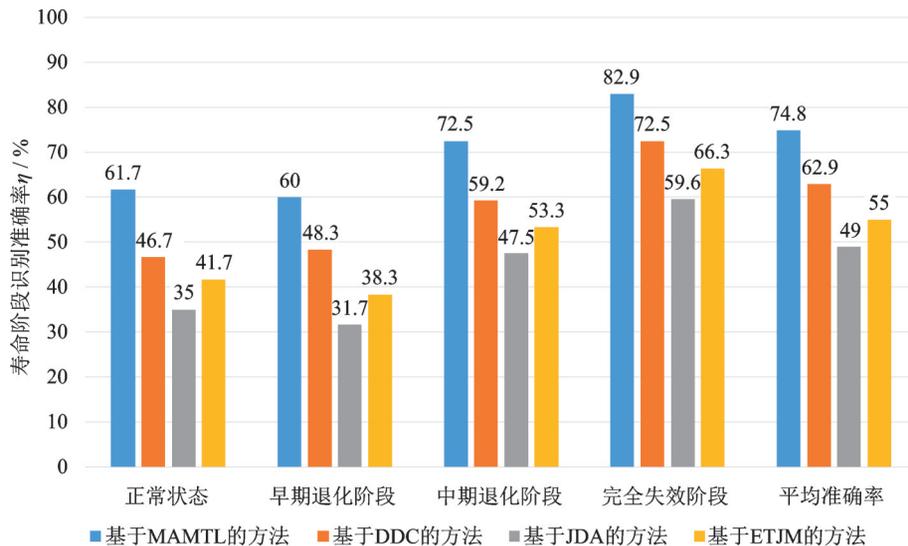


图 15 源域有类标签样本总数 K = 24 寿命阶段识别准确率

Fig. 15 Life stage identification accuracy of space rolling bearing 1 when the total number of labeled samples in source domain is K = 24

取中期退化阶段的样本数为 $2K/10$, 取完全失效阶段的样本数为 $K/10$ 用作源域有类标签样本 (即源域所有寿命阶段样本的总数为 $K, K \leq 200$), 在目标域中按照每一寿命阶段 2:4:1:3 的比例进行随机取样作为目标域当前待测样本, 待测样本总数也为 K 。将本文所提出的方法对目标域待测样本的四种寿命阶段识别准确率及平均识别准确率与其他三种迁移学习方法进行了比较。随着源域有类标签训练样本总数的减少, 寿命阶段平均识别准确率对比结果如图 16 所示。在源域有类标签样本总数 $K=30$ 时, 所提出的方法和三种被比较的方法的寿命阶段识别准确率如图 17 所示。

图 16 和 17 的对比结果表明, 即使源域有类标签样本总数逐渐减小以及源域不同类标签样本数变得

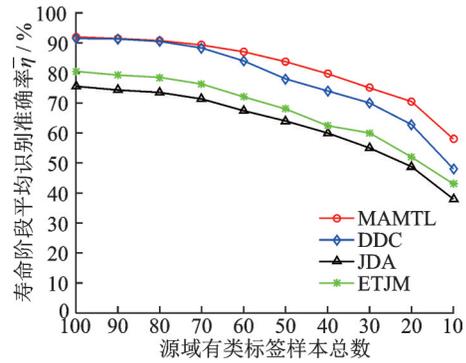


图 16 空间滚动轴承 1 寿命阶段的平均识别准确率对比
Fig. 16 Comparison of the average identification accuracy of the life stages of space rolling bearing 1

不均等, 所提出的方法对四种寿命阶段的识别准确率及平均识别准确率也分别比其他三种被对比方法更高。

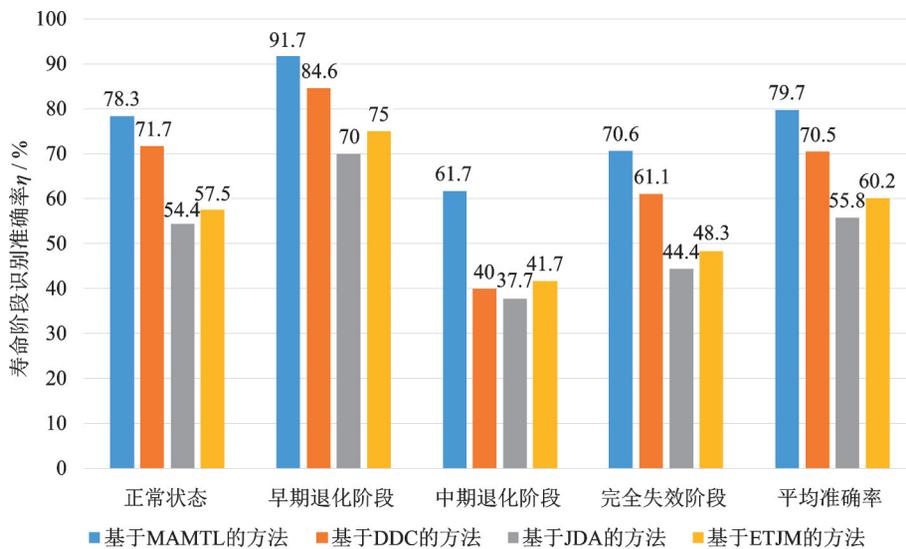


图 17 源域有类标签样本总数 $K=30$ 寿命阶段识别准确率

Fig. 17 Life stage identification accuracy of space rolling bearing 1 when the total number of labeled samples in source domain is $K=30$

4 结论

(1) 在所提出的 MAMTL 中, 将模型无关元学习和迁移学习相结合以实现多任务同步平行训练从而代替传统的迭代训练, 以改善 MAMTL 的泛化性能。具体而言, MAMTL 中的外环元学习网络每一次更新的参数都是由内环平行网络中多个任务的所有损失函数共同作用得到的全局最优解, 所以以这个全局最优解作为外环元学习网络的起始点去学习新的任务时仅需要少量迭代次数就能使外环元学习网络达到收敛, 即又好又快地适应新的迁移学习任

务, 因此 MAMTL 具有良好泛化性和域适配性。

(2) 在 MAMTL 中构建了新型原型网络作为分类器, 它将源域每一类别中所有的样本用一个原型表示, 通过计算目标域待测样本与原型的相似度完成目标域待测样本的分类, 这样避免了因源域不同类别样本数量的差异而造成对不同类别样本分类精度差别过大 (即对少样本类别的样本的分类精度过低) 问题; 同时, 在计算目标域待测样本与不同原型的相似度时没有参数学习的过程, 所以在训练小样本情况下不会出现过拟合现象。此外, MAMTL 的新型原型网络是与 MAMTL 匹配的框架性模型无关原型网络, 即该新型原型网络可以根据 MAM-

TL所采用的特征映射函数和分布差异度量函数来选择使分类精度最优的相似度函数,如余弦相似度、欧几里得距离、皮尔逊相关系数等。

(3) MAMTL的以上优势使得它可利用空间滚动轴承历史工况下的少量、非均等寿命阶段样本(即有类标签训练样本)来对当前待测样本进行较高精度的寿命阶段识别。

参考文献:

- [1] 陈仁祥, 陈思杨, 杨黎霞, 等. 基于振动敏感时频特征的航天轴承寿命状态识别方法[J]. 振动与冲击, 2016, 35(17): 134-139.
CHEN Renxiang, CHEN Siyang, YANG Lixia, et al. Life state recognition method for space bearings based on sensitive time-frequency features of vibration[J]. Journal of Vibration and Shock, 2016, 35 (17) : 134-139.
- [2] DONG S, SHENG J, TANG B, et al. Bearings in simulated space conditions running state detection based on Tsallis entropy-KPCA and optimized fuzzy c-means model[J]. Noise Control Engineering Journal, 2017, 65 (2): 62-70.
- [3] MIAO W, NIU C, YANG Y, et al. Grade-life prognostic model of aircraft engine bearing[J]. Transactions of Nanjing University of Aeronautics & Astronautics, 2012, 29(2): 171-178.
- [4] YANG B, LEI Y, JIA F, et al. A polynomial kernel induced distance metric to improve deep transfer learning for fault diagnosis of machines [J]. IEEE Transactions on Industrial Electronics, 2020, 67(11): 9747-9757.
- [5] RAZZAGHI P, RAZZAGHI P, ABBASI K. Transfer subspace learning via low-rank and discriminative reconstruction matrix[J]. Knowledge-Based Systems, 2019, 163: 174-185.
- [6] FENG Y, Chen J, Yang Z, et al. Similarity-based meta-learning network with adversarial domain adaptation for cross-domain fault identification[J]. Knowledge-Based Systems, 2021, 217: 106829.
- [7] WANG D, ZHANG M, XU Y, et al. Metric-based meta-learning model for few-shot fault diagnosis under multiple limited data conditions[J]. Mechanical Systems and Signal Processing, 2021, 155: 107510.
- [8] HUANG S, ZENG X, WU S, et al. Behavior regularized prototypical networks for semi-supervised few-shot image classification[J]. Pattern Recognition, 2021, 112: 107765.
- [9] QIU H, LEE J, LIN J, et al. Wavelet filter-based weak signature detection method and its application on rolling element bearing prognostics[J]. Journal of Sound and Vibration, 2006, 289(4-5): 1066-1090.
- [10] Helmi H, Forouzantabar A. Rolling bearing fault detection of electric motor using time domain and frequency domain features extraction and ANFIS[J]. IET Electric Power Applications, 2019, 13(5): 662-669.
- [11] Tasoulis S, Pavlidis N G, Roos T. Nonlinear dimensionality reduction for clustering[J]. Pattern Recognition, 2020, 107: 107508.
- [12] 丁锋, 何正嘉, 訾艳阳, 等. 基于设备状态振动特征的比例故障率模型可靠性评估[J]. 机械工程学报, 2009, 45(12): 89-94.
Ding Feng, He Zhengjia, Zi Yanyang, et al. Reliability assessment based on equipment condition vibration feature using proportional hazards model[J]. Journal of Mechanical Engineering, 2009, 45(12): 89-94.
- [13] 张焱, 汤宝平, 韩延, 等. 融合失效样本与截尾样本的滚动轴承寿命预测[J]. 振动与冲击, 2017, 36(23): 10-16.
ZHANG Yan, TANG Baoping, HAN Yan, et al. Life prediction for rolling bearings utilizing both failure and truncated samples[J]. Journal of Vibration and Shock, 2017, 36(23): 10-16.
- [14] 张小丽, 王保健, 马猛, 等. 滚动轴承寿命预测综述[J]. 机械设计与制造, 2015(10): 221-224.
ZHANG Xiaoli, WANG Baojian, MA Meng, et al. A review of life prediction for roller bearing[J]. Machinery Design & Manufacture, 2015(10): 221-224.
- [15] XUE X, ZHANG N, CAO S, et al. New state identification method for rotating machinery under variable load conditions based on hybrid entropy features and joint distribution adaptation[J]. Complexity, 2020, 2020: 7247195.
- [16] TZENG E, HOFFMAN J, ZHANG N, et al. Deep domain confusion: maximizing for domain invariance [J]. arXiv, 2014, doi: 10.48550/arXiv.1412.3474.
- [17] ZHANG Z, CHEN H, LI S, et al. Unsupervised domain adaptation via enhanced transfer joint matching for bearing fault diagnosis[J]. Measurement, 2020, 165: 108071.
- [18] CHEN P, LI Y, WANG K, et al. A novel knowledge transfer network with fluctuating operational condition adaptation for bearing fault pattern recognition[J]. Measurement, 2020, 158: 107739.

Model-agnostic meta-transfer learning for life stage identification of space rolling bearings

LI Tong-yi¹, LI Feng¹, TANG Bao-ping², WANG Yong-chao¹

(1.School of Mechanical Engineering, Sichuan University, Chengdu 610065, China;

2.The State Key Laboratory of Mechanical Transmissions, Chongqing University, Chongqing 400044, China)

Abstract: Considering the low accuracy of life stage identification caused by the large difference of sample distribution, the small number of life stage training samples and the unequal number of samples in different life stages, Model-Agnostic Meta-Transfer Learning (MAMTL) is proposed to identify the life stage of space rolling bearings. In MAMTL, model-agnostic meta-learning and transfer learning are combined to achieve multi-task synchronous parallel training instead of traditional iterative training. Multiple task loss functions in MAMTL use unlabeled samples under different working conditions and a small number of labeled samples under historical working conditions to jointly update the network parameters of MAMTL to seek global optimal solutions of them, which makes MAMTL have better generalization ability so that MAMTL has better domain adaptability than traditional transfer learnings when there are few labeled training samples under historical conditions. Moreover, a new prototype network is constructed in MAMTL to represent the samples of each class in historical working conditions as a prototype. Thus, the testing samples under current working conditions are classified by calculating the similarity between the testing samples and the prototypes, and the classification process does not need parameter learning, which can prevent the problem of large difference in identification accuracy of different classes under the unequal number of samples in different classes and the over fitting problem of network in the case of few labeled training samples, and then better improve the classification accuracy. The above advantages of MAMTL enable it to use few and unequal training samples known for the life stages under historical working conditions of space rolling bearings to perform high-precision life stage identification for the current testing samples. The life stage identification examples of a space rolling bearing demonstrate the effectiveness of the proposed MAMTL-based life stage identification method.

Key words: life stage identification; space rolling bearings; prototype network; meta learning; transfer learning

作者简介: 李统一(1998—),男,硕士研究生。电话: 13060050515; E-mail: 451628422@qq.com。

通讯作者: 李 锋(1982—),男,博士,副教授。电话: 18382385401; E-mail: lifeng19820501@163.com。