工业设备的健康状态评估和退化趋势预测联合研究

张永¹, 龚众望¹, 郑英², 谢林柏³, 张泽⁴, 刘振兴^{1*}

1. 武汉科技大学信息科学与工程学院, 武汉 430081;
2. 华中科技大学人工智能与自动化学院, 武汉 430074;
3. 江南大学物联网工程学院, 无锡 214122;
4. 亚马逊公司, 西雅图 98109, 美国

* E-mail: zhenxingliu@wust.edu.cn

收稿日期: 2021-08-01; 接受日期: 2021-10-08; 网络版发表日期: 2021-12-03

国家重点研发计划(编号: 2018YFB1701202)和国家自然科学基金(批准号: 61873197, 61873102)资助项目

摘要 工业设备作为工业互联网的基本组成部分, 其健康状态事关工业产品的质量水平和生产过程的稳定性以及流畅性. 因此, 对工业设备的健康状态进行评估和退化趋势预测具有重要的理论价值和工程实践意义. 本文基于深度学习方法, 拟构建双任务框架, 实现工业设备的健康状态评估和退化趋势预测联合研究. 具体来说, 首先对工业设备的监测信号进行时域和频域特征提取, 使用LightGBM选择强相关性的特征, 通过主成分分析降维建立健康指标, 并构建健康状态的类别标签; 接着发展领域自适应的迁移学习算法, 即流形空间分布对齐, 将源域和目标域在特征层面进行迁移, 完成源域和目标域的分布对齐; 最后, 融合因果膨胀卷积、双向门控循环单元以及注意力机制技术, 设计了双任务深度网络框架, 实现了工业设备的健康状态评估和退化趋势预测的双功能. 本文方法有效解决了单独研究工业设备的健康状态评估或退化趋势预测的缺陷, 可以同时和设备进行状态进行实时定性分析(健康状态)和定量分析(剩余使用寿命). 最后, 以工业关键部件的轴承和刀具为例, 通过和当前流行的深度学习方法比较, 实验结果验证了本文所提的方法能够实现双任务功能以及显著提升退化趋势预测的精度.

关键词 健康状态评估, 退化趋势预测, 流行空间分布对齐, 因果膨胀卷积, 注意力机制

1 引言

工业互联网依托其深度感知、智能分析、高效处理、集成互通等能力, 以数据分析决策为核心, 形成面向不同工业场景的智能化生产、网络化协同、规模化定制和服务化延伸等智能应用解决方案^[1]. 随着人工

智能、大数据以及5G等技术的应用和快速发展, 工业互联网有望成为推动制造业高质量发展, 促进先进制造业和现代服务业融合发展, 加快建设制造强国的具体实现方式和推进抓手. 李克强总理在《政府工作报告》提出要打造工业互联网平台, 拓展“智能+”, 为制造业转型升级赋能. 工业互联网平台是开发和运行各

引用格式: 张永, 龚众望, 郑英, 等. 工业设备的健康状态评估和退化趋势预测联合研究. 中国科学: 技术科学, 2022, 52: 180–197

Zhang Y, Gong Z W, Zheng Y, et al. Joint study on health state assessment and degradation trend prediction of industrial equipment (in Chinese). Sci Sin Tech, 2022, 52: 180–197, doi: 10.1360/SST-2021-0337

种工业互联网应用功能的平台, 主要包含边缘层、LaaS层、PaaS层和SaaS层等. 其中, 从设备的控制系统、传感器、可穿戴设备、摄像头和仪表等部分采集的设备监测数据是工业互联网平台的重要基础^[2].

无论工业互联网还是工业互联网平台, 工业设备都是基本部分. 工业关键设备在工业生产中担负着关键作用, 事关工业生产的顺利开展、工业系统的稳定运行、工业产品质量的水平以及生产厂家的竞争力. 随着设备使用时间的增加, 设备的运行经历了磨合期、健康生产期、磨损退化期甚至故障. 当设备发生退化时, 如果没有采取相应的维修措施, 退化发展到一定程度后将出现故障. 由于工业生产是一个复杂的耦合过程, 工业关键设备的退化将导致产品质量下降以及次品率增加, 影响整个生产系统的健康运转, 严重时引发重大安全事故. 因此, 在工业设备的退化阶段, 采用高效的健康监测技术和剩余使用寿命预测方法, 获得工业设备运行的相关信息和剩余使用寿命时间, 组织合理的运维, 才可以节省维护成本, 提高产品的质量, 避免事故的发生, 提高企业的利润^[2].

目前, 工业设备的健康状态评估(即健康阶段划分)和退化趋势预测(即剩余使用寿命预测, Remaining useful life, RUL)研究, 主要采用两种不同的方法来实现^[3,4]: 基于模型的方法^[5,6]和数据驱动的方法^[7]. 模型方法基于设备退化的物理规律或者经验知识, 常表示为微分方程^[5]或者差分方程^[6], 然后利用参数辨识方法来获得模型参数, 进而通过外推获得设备的剩余使用时间. 文献[8]利用Paris疲劳扩展模型预测轴承的缺陷增长率; 文献[9]采用双指数模型描述锂电池的退化过程, 结合自适应无迹卡尔曼滤波技术确定退化模型的参数进而预测电池的RUL; 文献[10]将指数模型和多项式模型进行集成模拟退化趋势并通过粒子滤波算法调整参数预测RUL. 然而, 获得精确的退化模型常常需要专业的经验知识和长时间的探索. 随着新材料和新技术的发展, 柔性制造和3D打印等先进制造模式的出现, 产品更新换代的加速, 以及高精尖制造过程各个组成部分的相互影响, 使得获得准确的退化模型变得越来越难以实现.

近年来, 随着大数据和人工智能技术的快速发展, 数据驱动的机器学习和深度学习方法引起了广泛的研究关注^[11], 数据驱动方法基于设备运行数据来挖掘工业设备的退化规律, 可以揭示其中蕴含的信息或退化

趋势, 在设备的评估和RUL预测研究中取得了许多非常有意义的成果^[12]. 其中, 常见的机器学习方法包括支持向量机、高斯过程回归以及浅层神经网络方法等, 有着数据处理能力强、训练快、易收敛等优点, 早期得到了广泛的应用. 比如文献[13]使用相空间重构寻找支持向量机模型最佳输入模式, 并且提出了误差补偿的思想进一步提高RUL预测精度. 文献[14]利用变分模态分解将退化数据平稳化, 然后利用高斯过程回归进行RUL预测. 文献[15]用PCA将轴承的特征进行融合, 进一步在单变量极限学习机的基础上构建多变量极限学习机模型来预测轴承剩余寿命. 然而浅层机器学习的特征表示能力相对有限, 所以不适合处理数据量较大和挖掘复杂数据中隐含的退化机理, 预测效果不太理想.

随着传感器技术的发展、5G网络的普及以及工业互联网的应用, 工业设备和产线运行的数据逐渐积累, 浅层机器学习在海量的数据上难以完成在少量数据上的精准分析, 并且浅层机器学习经常依赖较好的信号处理和特征工程技术. 深度学习作为浅层神经网络的发展, 以其强大的数据挖掘能力和高效的数据处理性能, 成为当前RUL预测研究的主要侧重点. 文献[16]利用集成经验模态分解进行特征趋势重构和健康指标构建, 并利用循环神经网络预测了涡轮发动机的剩余使用寿命; 文献[17]采用卷积神经网络和长短期记忆神经网络融合方法, 端对端地预测了锂电池的剩余使用寿命; 文献[18]提出谱主能量特征提取方法并使用深度卷积神经网络进行滚动轴承的RUL预测; 文献[19]利用集成经验模态方法构建轴承的退化趋势, 同时采用门控循环神经网络, 有效地检测了滚动轴承的退化点和剩余使用寿命. 尽管利用深度学习方法可以高效地分析工业设备的运行数据, 获得了经典方法无法获得的规律, 现阶段仍面临一系列问题, 比如为了获得预期的性能, 深度网络的调参花费了大量时间, 且分析结果鲁棒性不强, 随着工业设备运行场景的变化, 分析结果受到很大影响等, 限制了深度学习的进一步落地和产业化.

机器学习算法常常假设训练集和测试集数据满足独立同分布, 实际情况下数据的分布并不一定服从这种规律, 因而影响分析的结果. 迁移学习^[20]是一类解决这类问题的有效方法, 迁移学习主要利用数据、任务或模型之间的相似性, 通过将源域上学习到的知识

应用到目标域上来提高任务的性能, 已经在计算机视觉和自然语言处理等领域获得了成功的应用。目前, 在工业设备的安全性分析方面也引起了研究者的关注, 文献[20]提出了一种深度卷积迁移学习网络方法用于自动学习特征和识别设备, 然后基于域自适应模块最小化概率分布距离等技术实现了设备的故障诊断。文献[21]使用深度压缩降噪自编码器提取深度特征, 使用迁移成分分析在源域和目标域之间找到一个共同的子空间, 最后使用SVR实现了轴承的健康评估。文献[22]提出一种最小化序列相似度的领域自适应算法, 选取目标域与源域之间的公共敏感特征, 最后采用支持向量机构建预测模型。总的来说, 以上方面方法主要实现了特征或者模型上的迁移, 但是没有涉及到工业设备安全性分析中一类常见的数据“负迁移”问题。

对工业设备的健康监测来说, 主要分为异常检测、健康状态划分、剩余使用寿命预测以及故障诊断等方向^[23]。相对于事故发生后的故障诊断与溯源^[24], 异常检测、健康状态划分以及剩余使用寿命预测等研究, 主要关注事故发生前工业设备的运行状况, 更具有实际意义。基于设备评估的健康状态划分可以定性地展示设备处于健康、轻微退化、严重退化等健康状态, 健康信息简单明了实用; 设备的剩余使用寿命预测是对设备有效使用时间的一种精准预测, 进而可以获得比较精准的维护策略。比如, 文献[25]使用希尔伯特黄变换提取HI (Health indicator), 用支持向量机对退化信号进行分类划分退化阶段; 文献[26]运用对特征的无监督聚类来确定整体退化阶段的类别标签; 文献[27]将特征与威布尔分布拟合, 然后使用模糊自适应神经网络将轴承分为健康状态和退化的六个阶段。目前, 已有的成果主要基于机器学习方法单独地实现了工业设备的健康状态划分或剩余使用寿命预测功能, 而同时实现工业设备的健康状态划分和剩余使用寿命预测具有更好的工业实用价值。

基于研究现状以及工业设备健康监测现实, 本文将迁移学习与双任务框架相结合, 借助迁移学习提升预测模型的泛化性, 并利用双任务网络框架同时完成轴承的健康状态划分和RUL预测。具体来说, 首先对轴承的信号进行时域、频域的特征提取与选择, 接着利用主元分析建立健康指标, 构建RUL和退化阶段的标签。在特征处理方面, 利用迁移学习的MEDA (Manifold embedded distribution alignment)领域自适应^[28]的

思想, 将源域和目标域进行特征上的分布对齐, 同时完成对边缘分布和条件分布两种分布上的适配。最后, 将迁移之后的特征输入双任务模型, 输出RUL预测值和健康状态划分值。为了验证所提方法的有效性, 本文在IEEE PHM Challenge 2012轴承数据集和PHM2010高速机床铣削刀具数据集上进行了对比实验。本文的主要贡献可以概括为: (1) 使用LightGBM (Light gradient boosting machine)选择有强相关性的特征, 通过主成分分析降维建立HI; (2) 使用迁移学习领域自适应方法, 解决了特征迁移在特征层面解决数据分布不平衡的问题; (3) 设计了双任务网络框架, 能够并行地进行RUL预测和健康状态划分; (4) 通过对条件分布和边缘分布的适配, 解决负迁移产生的问题。

2 理论介绍

本节将详细阐述所提出的工业设备双任务分析框架, 图1给出了所提出方法的流程图。具体来说, 方法上包括四个步骤: 特征提取与选择、HI和标签的构建、MEDA领域自适应和双任务网络架构。各部分详细情况如下。

2.1 特征提取与选择

高效的特征提取和选择可以从工业设备运行数据中挖掘出反映设备运行规律的关键指标, 对工业设备的健康分析具有重要作用。本文采用的24个统计特征如表1所示, 包括16个时域特征和8个频域特征, 具体为最大值、最小值、均值、峰峰值、绝对平均值、均方根值、方根幅值、方差、峭度、偏度、平均幅值、裕度指标、波形指标、脉冲指标、峰值指标、峭度指标、频率中心、平均频率、频域均方根以及五个子频谱。如文献[29]所述, 振动信号是工业设备非常有效的特征, 通过经验提取的统计特征不一定适合需求, 需要对其进行重要度排序, 选择重要性更高的特征。树模型可以用来评估特征的重要性, 被称作LTK (Learn to rank)^[30], 特征被选为分割点的次数越多, 那么这个特征的重要性就越强, 其中LightGBM^[31]相比其他的树模型有着收敛速度快、精度高等优点, 因此, 本文使用其对提取的特征进行排序选择后并归一化。HI作为能直接反映轴承退化状态的指标, 本文采用降维方法进行构建。具体来说, 首先对特征进行提取和选择, 接

图 1 工业设备的双任务健康分析流程图

Figure 1 The flowchart of dual task health analysis for industrial equipment.

着采用主成分分析方法对特征进行降维, 取一维主成分并归一化, 作为HI. 同时, 由于工业设备的信号没有标签, 故设备在时间序列上的退化模式和程度是未知的. 因此, 它需要为健康状态划分的分类任务构造目标标签, 将归一化的HI按其变化速率进行健康阶段划分.

2.2 MEDA领域自适应算法

将迁移学习的源域记为 D_s , 目标域记为 D_t , 考虑到 D_s 和 D_t 的分布有差异, 包括在边缘分布上存在的差异 $P_s(X_s) \neq P_t(X_t)$ 和在条件分布上存在的差异 $Q_s(y_s|X_s) \neq Q_t(y_t|X_t)$. 传统的领域自适应的方法一般是通过构造子空间, 使得 D_s 和 D_t 的数据分布的距离更接近. 然而, 现存的方法大多只在原始的数据特征空间进行迁移, 由于原本空间下的特征通常存在扭曲, 因而会极大地减弱迁移的效果. 根据流形假设, 嵌入在流形空间中的数据点和与其相近的数据通常都有着相似的性质^[32]. 因此, 我们使用流形空间特征变换, 希望以此来减少原始特征空间下的失真, 从而来减小领域之间的数据发散^[33], 将源域和目标域共同映射到Grassman流形空间中, 将 S_s 和 S_t 分别表示其子空间, 则 G 可以视为所有的

维度子空间的集合. 每一个原始子空间都可以被看作 G 上的一个点. 因此, 在两点之间的测地线 $\{\Phi(t); 0 \leq t \leq 1\}$ 构成了两个子空间之间的一条路径, 令 $S_s = \Phi(0)$, $S_t = \Phi(1)$, 构造一条从 $\Phi(0)$ 到 $\Phi(1)$ 的测地流线, 相当于将原始特征变换到一个无穷维度的空间中, 最终减少特征的扭曲. 流形空间的特征可以表示为: $Z = \Phi(t)^T x$, 由文献^[32]可知, 变换后的特征 z_i 和 z_j 的内积定义一个侧地流核:

$$\langle z_i, z_j \rangle = \int_0^1 (\Phi(t)^T x_i)^T (\Phi(t)^T x_j) dt = x_i^T G x_j. \quad (1)$$

通过 $Z = \sqrt{G}X$, 原始特征变换到了Grassman流形空间中. 同时, 现存的分布适配方法通常假定边缘分布(P)和条件分布(Q)是同等重要的. 然而, 这种假设在应用时通常不会成立. 因此, 我们引入动态因子 μ 来动态衡量 P 和 Q 的重要程度, 动态分布对齐定义为

$$\bar{D}_f(D_s, D_t) = (1 - \mu)D_f(P_s, P_t) + \mu \sum_{c=1}^C D_f^{(c)}(Q_s, Q_t), \quad (2)$$

其中, P_s 和 P_t 分别代表源域和目标域的边缘分布, Q_s 和 Q_t 代表源域和目标域的条件分布. 用最大均值差异(Maximum mean discrepancy, MMD)来计算概率分布的距离, 表达式如下:

表 1 时频域特征

Table 1 Features of time-frequency domain

特征	公式	特征	公式
最大值	$\text{Max}(x)$	最小值	$\text{Min}(x)$
均值	$\frac{1}{n} \sum_{i=1}^n x_i$	峰峰值	$\text{Max}(x) - \text{Min}(x)$
绝对平均值	$\frac{1}{n} \sum_{i=1}^n x_i $	均方根值	$\sqrt{\frac{1}{n} \sum_{i=1}^n x_i^2}$
方根幅值	$\left(\frac{1}{n} \sqrt{\sum_{i=1}^n x_i } \right)^2$	方差	$\sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2}$
峭度	$\frac{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^4}{\left(\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 \right)^2}$	偏度	$\frac{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^4}{\left(\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 \right)^{\frac{3}{2}}}$
平均幅值	$\frac{1}{n} \sum_{i=1}^n x_i$	裕度指标	$\frac{\text{Max} x_i }{\left(\frac{1}{n} \sqrt{\sum_{i=1}^n x_i } \right)^2}$
波形指标	$\frac{\sqrt{\frac{1}{n} \sum_{i=1}^n x_i^2}}{\frac{1}{n} \sum_{i=1}^n x_i }$	脉冲指标	$\frac{\text{Max}(x)}{\frac{1}{n} \sum_{i=1}^n x_i }$
峰值指标	$\frac{\text{Max}(x)}{\sqrt{\frac{1}{n} \sum_{i=1}^n x_i^2}}$	峭度指标	$\frac{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2}{\left(\frac{1}{n} \sum_{i=1}^n x_i^2 \right)^2}$
频率中心	$\frac{\sum_{k=1}^K f_k^2 s(k)}{\sum_{k=1}^K s(k)}$	平均频率	$\frac{\sum_{k=1}^K f_k^2 s(k)}{K}$
频率均方根	$\sqrt{\frac{\sum_{k=1}^K f_k^2 s(k)}{\sum_{k=1}^K s(k)}}$	子频谱	$\sum_{k=\frac{m-1}{m}K}^{\frac{m}{m}K} f_k$

$$MMD(X, Y) = \left\| \sum_{i=1}^{n_1} \phi(x_i) - \sum_{j=1}^{n_2} \phi(y_j) \right\|_H^2. \quad (3)$$

MMD是一种非参数化的分布估计方法, 已经被广泛应用于多种迁移学习. 其中, $\phi(\cdot)$ 代表将原始值映射到再生希尔伯特空间的映射. MMD的具体计算方法按文献[34]中的构造核函数方法求得映射 $\phi(\cdot)$, 代入式(2)中计算. 通过求解半定规划优化问题, 可以得到最优权重矩阵和两个域的适应特征. μ 需要进行定量的估计, 我们采用A-distance来估计条件分布与边缘分布之间的距离.

$$d_A(D_s, D_t) = 2(1 - 2err(h)), \quad (4)$$

其中, A-distance被定义为建立一个线性分类器来区分

两个数据领域的hinge损失($err(h)$ 代表进行二类分类的hinge损失). 对边缘分布来说, 直接计算源域和目标域的A-distance, 记为 A_M . 对于条件分布来说, 利用2.1节中构建的类别标签, 对于两个域中来自同一个类别的数据, 计算A-distance. 将 A_C 记为所有类别之间A-distance的平均值, 最优适配因子 $\hat{\mu}$ 可以通过下式来估计:

$$\hat{\mu} = \frac{A_C}{A_M + A_C}. \quad (5)$$

MEDA领域自适应的具体算法如算法1.

2.3 双任务分析框架

基于深度学习的任务应用具有灵活性和高效性, 本文受多任务RNN^[35]的框架启发, 提出了一种联合分

算法1 MEDA序列领域自适应算法

Input: 经过选择后的源域的最优特征矩阵 $F_{\text{source}} = [F_{s1}, F_{s2}, \dots, F_{sn}]$ 和目标域最优的特征矩阵 $F_{\text{target}} = [F_{t1}, F_{t2}, \dots, F_{tm}]$.

Output: 迁移之后得到的新源域特征矩阵 $F_{\text{new_source}} = [\hat{F}_{s1}, \hat{F}_{s2}, \dots, \hat{F}_{sm}]$ 和新的目标域特征矩阵 $F_{\text{new_target}} = [\hat{F}_{t1}, \hat{F}_{t2}, \dots, \hat{F}_{tm}]$.

Step 1: 将 F_{source} 和 F_{target} 共同映射到Grassman流形空间中, 由公式(1)构造测地流核, 根据 $Z = \sqrt{G}X$ 得到新的流形特征 Z_s, Z_t .

Step 2: 通过公式(5)计算最优适配因子 $F_{\text{new_source}} = [\hat{F}_{s1}, \hat{F}_{s2}, \dots, \hat{F}_{sm}]$.

Step 3: 通过公式(3)利用核函数求适配后的MMD.

Step 4: 将流形特征 Z_s, Z_t 通过适配过条件分布和边缘分布后的MMD变换得到两个域自适应特征矩阵: 新源域特征矩阵 $F_{\text{new_source}} = [\hat{F}_{s1}, \hat{F}_{s2}, \dots, \hat{F}_{sm}]$ 和新的目标域特征矩阵 $F_{\text{new_target}} = [\hat{F}_{t1}, \hat{F}_{t2}, \dots, \hat{F}_{tm}]$.

类和回归方法. 已建立的端到端双任务网络的架构如图1所示, 这一框架包括三个部分: (1) 因果膨胀卷积; (2) 双层双向GRU的堆叠; (3) 注意力机制.

对于输入 $x=[x_0, x_1 \dots x_T]$ 的时序数据的预测问题, 可以考虑因果卷积^[36]是因为要求对 t 时刻的输出 y_t 只能使用 t 时刻之前的 $x_1 \dots x_{t-1}$ 数据进行判别, 因果卷积下一层 t 时刻的输出 y_t 可以表示为

$$y_t = \zeta(x_0, x_1 \dots x_{t-1}). \quad (6)$$

因果卷积只能以网络深度的线性大小回顾历史数据, 这使得因果卷积如果要想挖掘更深层次的特征, 需要线性地堆叠很多层. 为了降低计算量同时保证输出和输入大小一致, 提出了膨胀卷积^[37], 可以通过引入膨胀因子获取更大的感受野. 将因果卷积和膨胀卷积结合形成因果膨胀卷积^[38]能同时获得巨大的感受野和并行性. 在信号 x 的某元素 s 上进行膨胀卷积表达如下:

$$F(s) = (x *_{d} f)(s) = \sum_{i=0}^{k-1} f(i) \cdot x_{s-d \cdot i} \quad (7)$$

其中, k 是滤波器 f 的长度, d 为膨胀因子.

为了增加网络的感受野, 可以选择更大的 k 和 d , 其中 d 通常随网络深度指数增加, 对于第 θ 层网络:

$$d = 2^{\theta-1}. \quad (8)$$

相对于LSTM, GRU在结构上进行简化, 使得参数更少更容易收敛. GRU除了包括两个门控单元、更新门和重置门, 还包含隐藏状态和候选隐藏状态, z_t 表示 t 时刻更新门的输出, r_t 表示 t 时刻重置门的输出, h_t 和 $\tilde{h}_t$ 分别表示隐藏状态和候选隐藏状态的输出, GRU的表达式如下:

$$r_t = \sigma(W_r * [h_{t-1}, x_t] + b_r), \quad (9)$$

$$z_t = \sigma(W_z * [h_{t-1}, x_t] + b_z), \quad (10)$$

$$\tilde{h}_t = \tanh(W_{\tilde{h}} * [r_t \cdot h_{t-1}, x_t] + b_{\tilde{h}_t}), \quad (11)$$

$$h_t = (1 - z_t) \cdot h_{t-1} + z_t \cdot \tilde{h}_t. \quad (12)$$

经典的循环神经网络, 状态的传输是从前往后单向展开的. 某些设备运行数据处理时, 当前时刻的输出不仅和之前的状态有关系, 也和之后的状态相关. 为了有效地解决这类问题, 引入了双向RNN, 双向RNN是由两个RNN上下叠加在一起组成的. 输出由这两个RNN的状态共同决定. 作为双向RNN的改进版, BiGRU^[39]可以更好地挖掘设备运行的时间序列在退化过程中的建模能力.

注意力机制^[40]可以看作一种软寻址操作, 元素由键(Key)和值(Value)组成, 通过计算查询(Query)与Key值的相似性大小决定从对应的元素值中取出多少内容, 每个地址键对应的值都会被抽取后进行求和, 再加权求和得到最终的Value值, 也就是Attention值, 整个过程的详细计算表达如下.

首先根据Query和Key的相似性准则 G 计算二者的相似性, 得到注意力得分 β_i :

$$\beta_i = G(q, K_i). \quad (13)$$

然后利用Softmax函数对注意力得分进行转换, 归一化的同时获取对应的重要性权重 α_i :

$$\alpha_i = \text{Softmax}(\beta_i) = \frac{e^{\beta_i}}{\sum_{i=1}^N e^{\beta_i}}. \quad (14)$$

最后根据得到的权重系数 α_i 对Value进行加权求和, 得到Attention值:

$$\text{Attention}((K, V), q) = \sum_{i=1}^N \alpha_i v_i. \quad (15)$$

退化趋势预测在时刻 t 的训练损失定义为标注的真实值 l_t 与预测值 $\hat{l}_t$ 的误差, 以均方误差作为预测任务的损失函数:

$$L_{RUL} = \frac{1}{M} \sum_{t=0}^{M-1} (l_t - \hat{l}_t)^2. \quad (16)$$

设备的健康阶段划分与退化趋势预测共享的网络结构, 可以同时共享退化信息, 把健康状态划分看作一个多分类问题, 在 t 时刻损失函数定义为真实状态 $c_t^{(w)}$ 与预测状态 $\hat{c}_t^{(w)}$ 之间的交叉熵:

$$L_{DS} = -\frac{1}{M} \sum_{t=0}^{M-1} \sum_{w=0}^W c_t^{(w)} \cdot \ln \hat{c}_t^{(w)}, \quad (17)$$

其中, W 为退化状态的类别数, M 为工业设备的寿命。

本文设计的双任务网络每个任务都有自己的训练目标. 为了同时实现退化趋势预测和健康状态划分, 这两个相互关联且不同的任务被联合训练. 因此, 联合双任务的最终目标是最小化如下目标函数:

$$L = L_{RUL} + \gamma L_{DS}, \quad (18)$$

其中, γ 由二者在训练过程中损失函数下降的比值决定, 定为60.

至此, 通过特征提取和选择、MEDA特征迁移、因果膨胀卷积和注意力机制等模块, 已经建立了双任务网络框架, 可以实现工业设备的退化趋势预测和健康状态划分的双任务功能。

3 实验分析

轴承是支持机器正常运转的关键零部件, 刀具是机械加工的基础工艺装备, 这些关键零部件广泛应用于国民经济和国防生产的各个领域, 是代表性的工业关键设备和基础部件. 同时, 轴承和刀具的工作环境和场景复杂, 设备运行中的监测信号噪声和干扰多, 由于使用和负载的多样化, 使得健康状态划分和剩余使用寿命预测研究一直是工业设备健康分析的难点. 本文将以滚动轴承和铣削刀具为例, 验证所提出的MEDA迁移方法和双任务框架的有效性。

3.1 场景介绍

IEEE PHM 2012数据集来源于PRONOSTIA试验台^[41]. 该平台通过加速寿命退化实验提供滚动轴承从正常到故障的全寿命的实验数据. 该实验包含了三个不同工况下的轴承数据: 第一个工况下电机转速为1800 r/min, 负载为4000 N; 第二个工况下电机转速为1650 r/min, 负载为4200 N; 第三个工况下电机转速为1500 r/min, 负载为5000 N. 其中工况一和工况二各包含7个轴承, 工况三包含3个轴承. 本次实验中, 采用工况一下的Bearing1-1作为源域, 工况一下其他轴承作为目标域的待预测轴承。

刀具采用美国纽约预测与健康管理学会2010年高速数控机床刀具健康预测竞赛开放数据中的铣削实验数据^[42]进行实验. 整个数据采集过程中, 采集频率为50 kHz/通道, 铣削主轴速度为10400 r/min, 进给速度为1555 mm/min, Y轴径向铣削深度为0.125 mm, Z轴轴向铣削深度为0.2 mm. 共有六把铣刀在七个通道的数据, 每次铣削过程的磨损值分为flute1, flute2和flute3三个槽向的值. 本文采用c1, c4和c6三个具有磨损标签值的全寿命铣刀作为实验数据, 其中, c1作为源域, c4和c6作为目标域。

3.2 目标标签构建

对于滚动轴承运行数据集, 首先使用LightGBM排序方法对特征排序. 如图2所示, 选择靠前的12个特征. 使用PCA将最优特征集降至三维, 然后选取第一个主成分经过平滑后作为HI. 在划分退化阶段时, 考虑轴

图2 LightGBM特征选择
Figure 2 Feature selection of LightGBM.

承的数据量和轴承的健康程度,在整个轴承全寿命周期内,以Bearing1-1为例,从图3(a)可以看出,开始的大量时间,轴承的HI保持在较为稳定的范围,直到第1728个采样点HI的增长速度发生了突变。于是,把从0至1728个采样点设为阶段0,在阶段0内为轴承的正常状态。在状态0内退化过程极其缓慢,在1728个采样点后HI的斜率变快,增长变快,所以考虑从状态0以后开始设为轴承的退化状态并开始预测轴承的剩余寿命。由于阶段0占有HI的截距为0.2,所以接下来的三个阶段也按0.2的HI进行划分。在阶段1至阶段3,整个HI的变化速度越来越快,第1~3阶段的物理意义为轴承从轻微逐步到严重的退化阶段。最后剩下的为阶段4,此阶段轴承的退化速度已经达到最高,并相当接近轴承的故障水平。类似于文献[29],为RUL预测目标赋值最理想的方法是使RUL随时间线性减少至0,即在设备退化中蕴含的时间相关信息。然而,轴承在实际使用中,长期处于健康状态,随着使用时间的增加,退化将逐渐加速,在接近失效时,退化速度加剧。因此,将最大RUL限制为常量的分段线性RUL目标标签函数,可以更好地模拟RUL随时间的变化。将阶段0设为正常状态,阶段1~4设为退化状态。如图4所示,阶段0的最后一个采样点设为退化点,从退化点开始计算轴承的RUL,更加符合实际场景。表2为工况一场景下七个轴承的正常状态区间和退化状态区间。刀具特征提取与选择方法与轴承一致,在状态划分上如图3(b)所示,使用文献[43]的方法根据其刀具磨损值划分为初步磨损、正常磨损和急剧磨损三个阶段,对三个槽向的磨损值取其平均

值即平均磨损值作为最终的磨损标签。

3.3 退化趋势预测和健康状态划分

将源域轴承1-1的特征和目标域轴承1-3的特征使用MEDA领域自适应进行对齐,源域刀具c1的特征和目标域刀具c4的特征进行对齐。在迁移之前,使用PCA将最优特征集降至三维并取第一个主成分比较分布差异,从图5(a)和6(a)可以看出,源域和目标域的数据概率密度分布有较大差异。在MEDA领域自适应算法后,从图5(b)和6(b)可以看出,概率密度分布差异大大减少。表明通过MEDA领域自适应算法能够拉近源域和目标域的概率分布,有效将源域和目标域的特征进行对齐,解决了源域与目标域存在较大分布差异的问题,在机器学习独立同分布的假设前提下,更加有利于构建预测模型。

为了验证MEDA特征处理方法的有效性和双任务框架的功能,选择目标域中的轴承1-3,1-4,1-5和1-6以及刀具的c4和c6作为效果展示,此处未迁移方法即直接使用统计特征,使用双任务网络框架进行预测,从图7和8中可以明显看出,四个轴承和两把刀具经过迁移后的误差的面积均小于未迁移误差的面积,说明使用MEDA算法后预测效果更好。同时,为了更加完整地评估此方法的预测性能,选择工况一的六个轴承,用未迁移方法和迁移方法进行对比实验,以如下的退化阶段均方根误差和绝对百分比误差作为评价指标:

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2}, \quad (19)$$

图3 源域健康状态划分。(a) 轴承1-1; (b) 刀具c1

Figure 3 Health states division of source domain. (a) Bearing1-1; (b) tool c1.

图 4 构建轴承退化阶段标签
Figure 4 Label construction of degradation stage for bearing.

表 2 工况一下七个轴承的退化状态划分
Table 2 Degradation state of bearings under working condition one

	正常状态 (s)	退化状态 (s)
Bearing1-1	0~17280	17290~28030
Bearing1-2	0~6300	6310~8710
Bearing1-3	0~13650	13660~23750
Bearing1-4	0~4090	4100~14280
Bearing1-5	0~16900	16910~24630
Bearing1-6	0~15780	15790~24480
Bearing1-7	0~14920	14930~22590

图 5 轴承迁移前后的概率密度分布. (a) 迁移前; (b) 迁移后
Figure 5 Probability density distribution of bearings before and after transfer. (a) Before transfer; (b) after transfer.

图 6 刀具迁移前后的概率密度. (a) 迁移前; (b) 迁移后
Figure 6 Probability density of tools before and after tool transfer. (a) Before transfer; (b) after transfer.

图 7 轴承迁移前后退化趋势预测及其误差. (a) 轴承1-3; (b) 轴承1-4; (c) 轴承1-5; (d) 轴承1-6

Figure 7 Prediction of degradation trend and its error for bearings before and after transfer. (a) Bearing1-3; (b) bearing1-4; (c) bearing1-5; (d) bearing1-6.

图 8 迁移前后刀具磨损预测及其误差. (a) 刀具c4; (b) 刀具c6

Figure 8 Prediction of wear and its error for tool before and after transfer. (a) Tool c4; (b) tool c6.

$$\text{MAPE} = \frac{1}{n} \sum_{i=1}^n \left| \frac{\hat{y}_i - y_i}{y_i} \right|, \quad (20)$$

其中, y_i 表示每一点的真实值, $\hat{y}_i$ 表示相应的预测值.

通过表3中的误差结果可以再次确认我们的结论: 在使用MEDA迁移之后, 轴承的剩余使用寿命预测结果变得更加精确, 即MEDA领域自适应能够更加有效

地挖掘特征中隐藏的轴承退化的信息, 从而能够显著地提升RUL的预测效果.

双任务框架能够并行地进行轴承的退化趋势预测和健康状态划分, 图9和10分别展示了使用迁移后的轴承和刀具的健康状态划分, 图11和12分别展示了相应的混淆矩阵. 如图9~12所示, 双任务框架可以提供工业设备整个使用寿命期的阶段划分和其相应的退化趋势预测. 从获得的结果来看, 可以获得比较精确的健康状态划分结果, 偶尔存在的错误分类一般发生在两个相邻阶段. 由于工业设备的退化是一个连续不断的过程, 少许的预测误差仍然可以接受. 从应用上来说, 对退化阶段划分的评估, 更多的是反映工业设备自身的一个状态, 通过判断设备的健康状态, 设计运维方案,

安排人员对设备进行维护. 从这个意义上说, 虽然预测效果不是几乎绝对准确, 但是设备的健康状态划分的目标已经达到. 对于存在的退化趋势预测效果不佳的问题, 如图7(a)的起始预测阶段并不十分准确以及图9(d)的第二阶段产生的一些预测偏差, 此时可以利用预测值对轴承的退化状态做互补估计, 重复利用双任务的并行性, 实现决策上的互补.

为了更好地验证我们提出的方法, 便于比较, 对轴承使用Score, MAE和NRMSE作为评价指标, 其表达式如下:

$$\text{Score} = \frac{1}{N_{\tau}} \sum_{\tau=1}^{N_{\tau}} A_{\tau}, \quad (21)$$

表 3 目标域迁移的轴承剩余使用寿命预测RMSE和MAPE

Table 3 RMSE and MAPE of remaining useful life prediction for bearings in target domain transfer

		Bearing1-2	Bearing1-3	Bearing1-4	Bearing1-5	Bearing1-6	Bearing1-7	平均值
RMSE	未迁移方法	85.19	75.67	89.23	135.63	92.83	113.32	98.65
	迁移方法	61.24	57.56	73.71	31.39	70.03	39.83	52.63
MAPE	未迁移方法	0.21	0.38	0.28	0.19	0.42	0.26	0.29
	迁移方法	0.15	0.16	0.18	0.04	0.21	0.06	0.13

表 4 多种领域自适应方法的剩余使用寿命预测RMSE比较

Table 4 RMSE of bearing's RUL prediction with different domain adaptation methods

	No DA	TCA	GFK	SA	MEDA
Bearing1-2	85.19	121.32	75.88	185.21	61.24
Bearing1-3	75.67	68.45	92.84	188.95	57.56
Bearing1-4	89.23	97.60	80.59	175.32	73.71
Bearing1-5	135.63	67.54	77.25	235.21	31.39
Bearing1-6	92.83	107.51	103.22	197.21	70.03
Bearing1-7	113.32	140.68	57.74	301.94	39.83

表 5 多种领域自适应方法的剩余使用寿命预测MAPE

Table 5 MAPE of bearing's RUL prediction with different domain adaptation methods

	No DA	TCA	GFK	SA	MEDA
Bearing1-2	0.21	0.35	0.19	0.79	0.15
Bearing1-3	0.38	0.26	0.45	0.88	0.16
Bearing1-4	0.28	0.31	0.24	0.84	0.18
Bearing1-5	0.19	0.12	0.18	0.91	0.04
Bearing1-6	0.42	0.44	0.44	0.89	0.21
Bearing1-7	0.26	0.32	0.13	1.15	0.06

图 9 轴承退化状态预测. (a) 轴承1-3; (b) 轴承1-4; (c) 轴承1-5; (d) 轴承1-6
Figure 9 Degradation state prediction of bearings. (a) Bearing1-3; (b) bearing1-4; (c) bearing1-5; (d) bearing1-6.

图 10 刀具磨损状态预测. (a) 刀具c4; (b) 刀具c6
Figure 10 Prediction of tool wear state. (a) Tool c4; (b) tool c6.

图 11 轴承退化状态对应的混淆矩阵. (a) 轴承1-3; (b) 轴承1-4; (c) 轴承1-5; (d) 轴承1-6

Figure 11 Confusion matrix of degradation state for bearings. (a) Bearing1-3; (b) bearing1-4; (c) bearing1-5; (d) bearing1-6.

图 12 刀具磨损状态的混淆矩阵. (a) 刀具c4; (b) 刀具c6

Figure 12 Confusion matrix of tool wear state. (a) Tool c4; (b) tool c6.

$$A_{\tau} = \begin{cases} \exp(-\ln(0.5) * (E_{\tau} / 5)), E_{\tau} \leq 0, \\ \exp(\ln(0.5) * (E_{\tau} / 20)), E_{\tau} > 0, \end{cases} \quad (22)$$

$$E_{\tau} = \frac{RUL_{real}^{(\tau)} - RUL_{pred}^{(\tau)}}{RUL_{real}^{(\tau)}} * 100\%, \quad (23)$$

$$MAE = \frac{1}{N_{\tau}} \sum_{\tau=1}^{N_{\tau}} |RUL_{real}^{(\tau)} - RUL_{pred}^{(\tau)}|, \quad (24)$$

$$NRMSE = \frac{\sqrt{\frac{1}{N_{\tau}} \sum_{\tau=1}^{N_{\tau}} (RUL_{real}^{(\tau)} - RUL_{pred}^{(\tau)})^2}}{\frac{1}{N_{\tau}} \sum_{\tau=1}^{N_{\tau}} RUL_{real}^{(\tau)}}, \quad (25)$$

其中, N_{τ} 代表目标域轴承总数, E_{τ} 代表第 τ 个目标域轴承百分比误差, $RUL_{real}^{(\tau)}$ 和 $RUL_{pred}^{(\tau)}$ 分别代表第 τ 个目标域轴承的实际和预测的RUL值. Score越接近1, 预测效果越好, NRMSE和MAE如果越接近0, 则预测的准确性越高.

为了体现本文预测方法的优势, 对于轴承选择四种典型的方法进行对比, 以轴承1-3至轴承1-7比较, 结果见表6. 四种方法包括三种深度学习方法和一种物理模型的方法, 其中文献[44]使用冻结卷积激活记忆网络(Frozen convolution and activated memory network, FCAMN)的深度学习方法进行RUL预测; 文献[45]对时频域进行小波变换并对多尺度卷积神经网络(Multi-scale convolutional neural network, MSCNN)进行RUL预测; 文献[46]将特征以单调性进行选择后通过RNN构建HI并预测RUL; 文献[47]使用加权最小量化误差(Weighted minimum quantization error, WMQE)构建HI刻画轴承的退化过程, 结合粒子滤波(Particle filter, PF)进行RUL预测. 通过表6可以证实, 本文方法的预测误差均为最小值, Score高于其他方法, 同时在MAE和

NRMSE指标上也远远小于其他方法. 对于刀具也选择了四种方法进行比较, 结果见表7. 具体来说, 包括文献[48]提出的传统机器学习SVR和ISOMAP-SVR, 以及文献[49]中采用的CNN和BiLSTM两种深度学习方法进行刀具的磨损预测, 将本文的方法与这些方法进行比较可以看出, 本文提出方法的误差均小于其他方法, 再次说明了所提方法的有效性.

3.4 负迁移

使用迁移学习时, 经常会出现迁移之后效果反而变差的现象, 负迁移^[50]指的是在源域上学习到的知识, 对于目标域上的学习产生负面作用. 负迁移造成的原因本质上是没找准源域和目标域的相似处. 如图13(a)和(b)所示, 试图使用迁移成分分析(Transfer component analysis, TCA)领域自适应方法提升RUL预测精度时却收到了明显的反效果, 且在轴承1-3的RUL预测中尤其明显. 在多次使用领域自适应的方法中, 负迁移产生的频率非常高. 为了验证MEDA领域自适应在特征迁移中的效果以及其受负迁移影响的大小, 本文和三种代表性无监督的领域自适应方法SA^[51], GFK^[52]和TCA^[33]在轴承上进行对比. 为了保证实验的严谨性, 将这些领域自适应算法替换本文的MEDA领域自适应算法, 并且尽量保证参数大小一致, 对比结果使用目标域6个轴承的平均RUL预测误差, 具体如表4, 5以及图14的箱型图所示. 其中TCA与SA的平均RMSE和MAPE均大于不使用领域自适应的情况, 这表明使用这两种领域自适应方法会产生大量的负迁移现象. 使用MEDA领域自适应时, 误差小于GFK, 这两种方法均使用了Grassman流形空间变换, 但是MEDA领域自适应进行了边缘分布和条件分布的适配. 因此, 仿真结果表明, MEDA领域自适应的做法能够减少负迁移

表 6 多种剩余使用寿命预测方法的误差比较

Table 6 Comparison of RUL prediction error with different methods

Methods	Error (%)					Score	MAE	NRMSE
	Bearing1-3	Bearing1-4	Bearing1-5	Bearing1-6	Bearing1-7			
FCAMN ^[44]	-1.66	21.95	-15.22	-5.74	7.99	0.43	61.0	0.13
MSCNN ^[45]	17.43	10.69	148.20	-19.45	19.55	0.36	1091.8	0.35
RNN ^[46]	43.28	67.55	-22.98	21.23	17.83	0.27	1262	0.53
WMQE-PF ^[47]	-0.35	5.61	100	28.08	-19.55	0.45	707.8	0.32
本文方法	0.09	2.23	0.52	3.71	-3.04	0.88	41.37	0.02

表 7 多种刀具磨损预测方法的预测误差比较

Table 7 Prediction error of tool wear with different prediction methods

Method	MAE (10^{-3} mm)	RMSE (10^{-3} mm)
ISOMAP-SVR ^[48]	8.2567±0.3113	10.6625±0.5869
SVR ^[48]	9.3370±2.0422	11.968±3.3337
CNN ^[49]	11.0000±1.3000	14.0428±5.5588
BiLSTM ^[49]	10.6000±1.9000	13.6831±2.053
本文方法(未迁移)	26.6456±0.8052	28.5822±1.1791
本文方法(迁移)	7.1229±1.3355	8.9701±1.1949

产生的频率,即使用MEDA领域自适应对比传统领域自适应来说有着更好的迁移效果。

4 结论

在工业互联网框架下,随着大数据技术的发展,可以采集到大量的工业设备监测数据,使得数据驱动的健康状况分析成为可能。本文基于因果膨胀卷积和注意力融合的深度学习方法,建立了工业设备健康分析的双任务框架,可以同时获得设备的实时健康状态以及未来的设备退化趋势,解决了现有的单任务框架分析的局限性。在特征工程方面,迁移学习是一种有效的增强特征方法,本文所提的方法可有效解决源域和目标域数据分布差异过大的问题。通过对边缘分布和条件分布的适配减小了负迁移概率,提高了工业设备的退化趋势预测精度。

后继研究主要关注两个方面: (1) 将深入研究特征迁移相关问题,比如扩充源域的范围、使用多源迁移

图 13 迁移成分分析的轴承剩余使用寿命预测误差比较. (a) 轴承1-3; (b) 轴承1-4

Figure 13 Comparison of prediction error on remaining useful life for bearings by transfer component analysis. (a) Bearing1-3; (b) bearing1-4.

图 14 四类领域自适应方法的预测误差比较. (a) RMSE; (b) MAPE

Figure 14 Prediction error for bearing with four domain adaptation methods. (a) RMSE; (b) MAPE.

等, 同时也将定性分析负迁移产生的原因以及减少负迁移的办法, 并尝试动态迁移的思想; (2) 增强深度学习方法的双任务分析能力, 同时也关注分析过程和分析结果的可解释性, 以及双任务框架的鲁棒性等。

参考文献

- 1 Yu X H, Liu M, Jiang X H, et al. Industrial internet architecture 2.0 (in Chinese). *Comput Integr Manuf Syst*, 2019, 25: 2983–2996 [余晓晖, 刘默, 蒋昕昊, 等. 工业互联网体系架构2.0. 计算机集成制造系统, 2019, 25: 2983–2996]
- 2 Gao J J. Intelligent maintenance and autonomous health of equipments enabled by industrial internet (in Chinese). *Comput Integr Manuf Syst*, 2019, 25: 3013–3025 [高金吉. 工业互联网赋能装备智能运维与自主健康. 计算机集成制造系统, 2019, 25: 3013–3025]
- 3 Zhou D H, Wei M H, Si X S. A survey on anomaly detection, life prediction and maintenance decision for industrial processes (in Chinese). *Acta Automatica Sin*, 2013, 39: 711–722 [周东华, 魏慕恒, 司小胜. 工业过程异常检测、寿命预测与维修决策的研究进展. 自动化学报, 2013, 39: 711–722]
- 4 Lei Y, Li N, Guo L, et al. Machinery health prognostics: A systematic review from data acquisition to RUL prediction. *Mech Syst Signal Processing*, 2018, 104: 799–834
- 5 Si X S, Wang W, Hu C H, et al. Remaining useful life estimation—A review on the statistical data driven approaches. *Eur J Operational Res*, 2011, 213: 1–14
- 6 Cubillo A, Perinpanayagam S, Esperon-Miguez M. A review of physics-based models in prognostics: Application to gears and bearings of rotating machinery. *Adv Mech Eng*, 2016, 8: 1012–1024
- 7 Yuan Y, Ma G, Cheng C, et al. A general end-to-end diagnosis framework for manufacturing systems. *Natl Sci Rev*, 2020, 7: 418–429
- 8 Li Y, Kurfess T R, Liang S Y. Stochastic prognostics for rolling element bearings. *Mech Syst Signal Processing*, 2000, 14: 747–762
- 9 Xue Z, Zhang Y, Cheng C, et al. Remaining useful life prediction of lithium-ion batteries with adaptive unscented kalman filter and optimized support vector regression. *Neurocomputing*, 2020, 376: 95–102
- 10 Xing Y, Ma E W M, Tsui K L, et al. An ensemble model for predicting the remaining useful performance of lithium-ion batteries. *MicroElectron Reliability*, 2013, 53: 811–820
- 11 Yuan Y, Tang X, Li X, et al. Data driven discovery of cyber physical systems, *Nat Commun*, 2019, 10: 1–9
- 12 Pei H, Hu C H, Si X S, et al. Review of machine learning based remaining useful life prediction methods for equipment (in Chinese). *J Mech Eng*, 2019, 55: 1–13 [裴洪, 胡昌华, 司小胜, 等. 基于机器学习的设备剩余寿命预测方法综述. 机械工程学报, 2019, 55: 1–13]
- 13 Chen L, Zhang Y, Zheng Y, et al. Remaining useful life prediction of lithium-ion battery with optimal input sequence selection and error compensation. *Neurocomputing*, 2020, 414: 245–254
- 14 Wu Y, Wang Y R. Remaining useful life prediction of lithium-ion batteries based on VMD and GPR algorithm (in Chinese). *Comput Modern*, 2020, 294: 81–86 [吴伟, 王友仁. 基于变分模态分解和高斯过程回归的锂离子电池剩余寿命预测方法. 计算机与现代化, 2020, 294: 81–86]
- 15 He Q, Li L, Jiang G Q, et al. Residual life predictions for bearings based on PCA and MELM (in Chinese). *China Mech Eng*, 2014, 25: 984–989 [何群, 李磊, 江国乾, 等. 基于PCA和多变量极限学习机的轴承剩余寿命预测. 中国机械工程, 2014, 25: 984–989]
- 16 Zhao S, Zhang Y, Wang S, et al. A recurrent neural network approach for remaining useful life prediction utilizing a novel trend features construction method. *Measurement*, 2019, 146: 279–288
- 17 Ma G, Zhang Y, Cheng C, et al. Remaining useful life prediction of lithium-ion batteries based on false nearest neighbors and a hybrid neural network. *Appl Energy*, 2019, 253: 113626
- 18 Cheng C, Ma G, Zhang Y, et al. A deep learning-based remaining useful life prediction approach for bearings. *IEEE/ASME Trans Mechatron*, 2020, 25: 1243–1254
- 19 Xiao L, Liu Z, Zhang Y, et al. Degradation assessment of bearings with trend-reconstruct-based features selection and gated recurrent unit network. *Measurement*, 2020, 165: 108064
- 20 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. *IEEE Trans Ind Electron*, 2019, 66: 7316–7325
- 21 Mao W, He J, Zuo M J. Predicting remaining useful life of rolling bearings based on deep feature representation and transfer learning. *IEEE Trans Instrum Meas*, 2019, 69: 1594–1608

- 22 Chen J X, Mao W T, Liu J, et al. Remaining useful life prediction of bearing based on deep temporal feature transfer (in Chinese). *J Control Decis*, 2021, 36: 1699–1706 [陈佳鲜, 毛文涛, 刘京, 等. 基于深度时序特征迁移的轴承剩余寿命预测方法. *控制与决策*, 2021, 36: 1699–1706]
- 23 Yuan Y, Zhang Y, Ding H. Research on key technology of industrial artificial intelligence and its application in predictive maintenance (in Chinese). *Acta Automat Sin*, 2020, 46: 13–30 [袁烨, 张永, 丁汉. 工业人工智能的关键技术及其在预测性维护中的应用现状. *自动化学报*, 2020, 46: 13–30]
- 24 Zheng Y, Mao S, Liu S, et al. Normalized relative RBC-based minimum risk bayesian decision approach for fault diagnosis of industrial process. *IEEE Trans Ind Electron*, 2016, 63: 7723–7732
- 25 Soualhi A, Medjaher K, Zerhouni N. Bearing health monitoring based on Hilbert–Huang transform, support vector machine, and regression. *IEEE Trans Instrum Meas*, 2014, 64: 52–62
- 26 Scanlon P, Kavanagh D F, Boland F M. Residual life prediction of rotating machines using acoustic noise signals. *IEEE Trans Instrum Meas*, 2013, 62: 95–108
- 27 Ben Ali J, Chebel-Morello B, Saidi L, et al. Accurate bearing remaining useful life prediction based on weibull distribution and artificial neural network. *Mech Syst Signal Processing*, 2015, 56–57: 150–172
- 28 Wang J D, Feng W J, Chen Y Q, Visual domain adaptation with manifold embedded distribution alignment. In: *Proceedings of 2018 ACM Multimedia Conference on Multimedia Conference (ACM)*. Seoul, Republic of Korea: ACM, 2018. 402–410
- 29 Wang Y, Peng Y, Zi Y, et al. A two-stage data-driven-based prognostic approach for bearing degradation problem. *IEEE Trans Ind Inf*, 2016, 12: 924–932
- 30 Yazbek D, Sibindi J S, Zyl T V. Deep similarity learning for sports team ranking. arXiv: [2013.13736](https://arxiv.org/abs/2013.13736)
- 31 Gu T, Xu G L, Li W L, et al. Intelligent house price evaluation model based on ensemble LightGBM and bayesian optimization strategy (in Chinese). *J Comput Appl*, 2020, 361: 290–295 [顾桐, 许国良, 李万林, 等. 基于集成LightGBM和贝叶斯优化策略的房价智能评估模型. *计算机应用*, 2020, 361: 290–295]
- 32 Zhao L W, Luo S W, Zhao Y C, et al. Study on the low-dimensional embedding and the embedding dimensionality of manifold of high-dimensional data (in Chinese). *J Software*, 2005, 16: 1423–1430 [赵连伟, 罗四维, 赵艳敞, 等. 高维数据流形的低维嵌入及嵌入维数研究. *软件学报*, 2005, 16: 1423–1430]
- 33 Wang Y, Zhang S, Li Y. Fault detection for a class of non-linear networked control systems with data drift. *IET signal process*, 2015, 9: 120–129
- 34 Pan S J, Tsang I W, Kwok J T, et al. Domain adaptation via transfer component analysis. *IEEE Trans Neural Netw*, 2011, 22: 199–210
- 35 Miao H, Li B, Sun C, et al. Joint learning of degradation assessment and RUL prediction for aeroengines via dual-task deep LSTM networks. *IEEE Trans Ind Inf*, 2019, 15: 5023–5032
- 36 Brazil T J. Causal-convolution-a new method for the transient analysis of linear systems at microwave frequencies. *IEEE Trans Microwave Theor Techn*, 1995, 43: 315–323
- 37 Rhee J S. Certain radially dilated convolution and its application. *Honam Math J*, 2010, 32: 101–112
- 38 Bai S, Kolter J Z, Koltun V. Trellis networks for sequence modeling. arXiv: [1810.06682](https://arxiv.org/abs/1810.06682)
- 39 Fang Y, Yang S, Zhao B, et al. Cyberbullying detection in social networks using Bi-GRU with self-attention mechanism. *Information (Switzerland)*, 2021, 12: 171
- 40 Bricolo E, Ganesini T, Fanini A, et al. Serial attention mechanisms in visual search: A direct behavioral demonstration. *J Cognitive Neuroscience*, 2002, 14: 980–993
- 41 Nectoux P. PRONOSTIA: An experimental platform for bearings accelerated degradation tests. In: *Proceedings of 2012 IEEE Conference on Prognostics and Health Management (PHM)*. Beijing: IEEE, 2012. 1–8
- 42 Li X, Lim B S, Zhou J H, et al. Fuzzy neural network modelling for tool wear estimation in dry milling operation. In: *Proceedings of Annual Conference of the Prognostics and Health Management Society*. Basel, Switzerland, 2009. 1–11
- 43 Yu J Y, Zou Q. *Fundamentals of Mechanical Manufacturing Technology* (in Chinese). Beijing: Machinery Industry Press, 2009. 46–47 [于俊一, 邹青. *机械制造技术基础*第二版. 北京: 机械工业出版社, 2009. 46–47]
- 44 Chen Z, Tu X, Hu Y, et al. Real-time bearing remaining useful life estimation based on the frozen convolutional and activated memory neural network. *IEEE Access*, 2019, 7: 96583–96593
- 45 Zhu J, Chen N, Peng W. Estimation of bearing remaining useful life based on multiscale convolutional neural network. *IEEE Trans Ind Electron*,

2018, 66: 3208–3216

- 46 Guo L, Li N, Jia F, et al. A recurrent neural network based health indicator for remaining useful life prediction of bearings. *Neurocomputing*, 2017, 240: 98–109
- 47 Lei Y, Li N, Gontarz S, et al. A model-based method for remaining useful life prediction of machinery. *IEEE Trans Rel*, 2016, 65: 1314–1326
- 48 Wang J, Xie J, Zhao R, et al. Multisensory fusion based virtual tool wear sensing for ubiquitous manufacturing. *Robotics Comput-Integrated Manufacturing*, 2017, 45: 47–58
- 49 Zhao R, Yan R, Chen Z, et al. Deep learning and its applications to machine health monitoring. *Mech Syst Signal Processing*, 2019, 115: 213–237
- 50 Pan S J, Yang Q. A survey on transfer learning. *IEEE Trans Knowl Data Eng*, 2010, 22: 1345–1359
- 51 Fernando B, Habrard A, Sebban M, et al. Unsupervised visual domain adaptation using subspace alignment. In: *Proceedings of 2014 IEEE International Conference on Computer Vision (ICCV)*. Sydney, Australia: IEEE, 2014. 2960–2967
- 52 Gong B, Yuan S, Fei S, et al. Geodesic flow kernel for unsupervised domain adaptation. In: *Proceedings of 2012 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Providence, USA: IEEE, 2012. 2066–2073

Joint study on health state assessment and degradation trend prediction of industrial equipment

ZHANG Yong¹, GONG ZhongWang¹, ZHENG Ying², XIE LinBo³, ZHANG Ze⁴ & LIU ZhenXing¹

¹ School of Information Science and Engineering, Wuhan University of Science and Technology, Wuhan 430081, China;

² School of Artificial Intelligence and Automation, Huazhong University of Science and Technology, Wuhan 430074, China;

³ School of Internet of Things Engineering, Jiangnan University, Wuxi 214122, China;

⁴ Amazon, Seattle 98109, USA

As a basic component of industrial internet and industrial internet platform, the health state of industrial equipment (IE) is related to the stable operation of industrial systems and the quality level of industrial products. Therefore, both health status assessment (HSA) and degradation trend prediction (DTP) of IE have important theoretical value and engineering significance. This study proposes constructing a dual-task framework to realize the HSA and DTP of IE using deep learning techniques. First, monitoring signals of IE are extracted in the time and frequency domains so that strongly correlated features are selected by a light gradient boosting machine. After principal component analysis-based dimensionality reduction, both the health index and the category label of health state are constructed. Especially, an adaptive transfer learning algorithm, namely, manifold spatial distribution alignment, is used to migrate the source domain to the target domain at the feature level to complete distribution alignment. Finally, by integrating causal expansion convolution, a bidirectional gate recurrent unit, and an attention mechanism, a dual-task depth network framework is designed to realize the dual functions of HSA and DTP of IE. Experimental results show that the proposed method effectively addresses the limitation of existing methods, which is to investigate HSA and DTP alone. In addition, compared with typical deep learning methods, the presented methods significantly improve DTP accuracy and reduce the frequency of negative migration.

dual tasks framework, health status assessment, degrade trend prediction, causal expansion convolution, attention mechanism

doi: [10.1360/SST-2021-0337](https://doi.org/10.1360/SST-2021-0337)