基于多任务学习的装备智能诊断与寿命预测方法

黄如意^{1,2}, 李霖蒲², 王震², 夏景演², 陈祝云^{2,3}, 李巍华^{1,2,3*}

1. 华南理工大学吴贤铭智能工程学院, 广州 511442;

2. 华南理工大学机械与汽车工程学院, 广州 510641;

3. 人工智能与数字经济广东省实验室(广州), 广州 510335

* E-mail: whlee@scut.edu.cn

收稿日期: 2021-07-26; 接受日期: 2021-11-02; 网络版发表日期: 2021-11-26

国家重点研发计划项目(编号: 2018YFB1702400)、国家自然科学基金项目(批准号: 51875208)和广东省重点领域研发计划项目(编号: 2021B0101200004, 2019B010154002)资助

摘要 智能故障诊断与预测性维护技术是推动我国制造业向数字化、网络化、智能化方向发展的关键切入点和实现“弯道超车”的创新动力。得益于人工智能和机器学习技术的快速发展, 智能故障诊断与寿命预测方法在航空航天、石油化工、船舶等领域得到了广泛的研究与应用。然而, 传统智能方法为实现机械装备的初始异常检测、故障诊断与寿命预测等多种任务时, 均需逐一地训练检测、诊断与预测多个模型, 不仅增加了设备开发、部署与维护的成本, 也极大地限制了智能方法的实用性。为解决上述局限性, 基于多任务学习, 提出一种用于机械装备智能诊断(初始异常检测和故障诊断)与智能寿命预测等多种任务的健康监测方法。首先, 利用JS散度(Jensen-Shannon divergence)和相关分析实现设备初始异常检测; 其次, 结合多任务学习(Multi-task learning)的知识共享机制建立可同时实现故障诊断与寿命预测的智能模型, 仅需训练和部署一个模型即可同时获得设备的故障信息与剩余使用寿命信息; 最后, 利用轴承全寿命加速退化实验数据, 通过与领域内先进方法对比, 验证了所提方法在初始异常检测、故障诊断与寿命预测等多种任务的有效性和可靠性, 为工业设备的智能诊断与预测性维护提供重要的决策依据, 具有较强的工程应用价值。

关键词 故障诊断, 寿命预测, 初始异常检测, 多任务学习, 机械装备

1 引言

现代化机械装备(如大型舰船、高速列车、高精数控机床、航空航天设备及现代生产设备等)是我国现代制造业实现数字化、网络化和智能化的基础和立足点^[1]。现代化机械装备的服役环境恶劣多变, 高负荷率和高作业率使其关键零部件在长时间服役过程中不

可避免地出现疲劳失效、损坏等突发故障, 因而导致重大生产安全隐患, 甚至灾难性事故^[2]。若能对机械装备进行有效精确的诊断与预测, 及时发现故障并做出相应的维护措施, 将在最大程度上避免灾难性事故发生和减小经济损失, 因此, 智能诊断与预测性维护对保障现代化机械装备安全运行具有重大的意义, 也是企业的迫切需求^[3-6]。

引用格式: 黄如意, 李霖蒲, 王震, 等. 基于多任务学习的装备智能诊断与寿命预测方法. 中国科学: 技术科学, 2022, 52: 123–137

Huang R Y, Li J P, Wang Z, et al. Intelligent diagnostic and prognostic method based on multitask learning for industrial equipment (in Chinese). Sci Sin Tech, 2022, 52: 123–137, doi: [10.1360/SST-2021-0314](https://doi.org/10.1360/SST-2021-0314)

在智能故障诊断与预测性维护领域, 国内外学者和各产业界的研究人员都十分重视机械装备的故障诊断与寿命预测方法的应用研究. 如图1所示, 智能诊断与寿命方法通常包含以下四个步骤: 数据采集、数据处理、模型选择与训练、以及工程应用, 其中工程应用主要包含故障诊断与寿命预测. 在智能故障诊断方面, 严如强等人^[7,8]于2019年综述了深度学习方法及其在机械装备健康监测方面的应用与研究, 囊括了目前主流的智能诊断模型, 并整理开放了相应的数据和代码, 以便于研究人员复现或应用相关算法. Jiao等人^[9]根据输入数据维度(如: 一维振动信号和二维时频信号)的不同, 综述了一维和二维卷积神经网络在旋转机械智能故障诊断中的应用研究. 艾剑良和杨曦中^[10]利用相关分析方法和自适应神经网络, 提出了一种应用于航空发动机的故障诊断方法, 在只有少量训练样本的情况下, 依然保证有较高的诊断精度. 向家伟^[11]提出一种机械故障数值模拟驱动的生成式对抗网络的新型智能诊断方法, 并将其应用于轴承、齿轮和转子等部件的故障诊断, 为推广智能诊断的实用性提供了一种有效手段. 邵海东等人^[12,13]通过引入小波基作为自编码网络的激活函数, 提出了多种改进自编码网络, 成功应用于齿轮箱和电力机车轮对轴承的复合故障智能诊断, 显著提升了自编码网络对高噪、耦合、非线性振动信号的特征学习能力. 在智能寿命预测方面, Lei等人^[14]将机械装备智能寿命预测划分为四个阶段: 数据采集和预处理、健康指标构建、健康状态划分以

及剩余寿命预测, 并进行了详细的综述. Zhang等人^[15]提出了一种基于循环神经网络(Recurrent neural networks, RNN)的轴承退化评估方法, 通过提取波形熵作为轴承退化评估指标, 经过解调分析验证了其在故障程度识别方面的有效性, 并将多个时刻的振动特征顺序输入到循环神经网络中, 对轴承的剩余使用寿命进行预测. 秦庆斌等人^[16]对高速动车组S38C车轴疲劳强度及剩余寿命评价进行了研究和验证. Huang等人^[17]基于机床加工过程中采集到的振动信号, 利用时频分析提取多项特征指标, 并作为卷积神经网络(Convolutional neural networks, CNN)的输入, 实现了刀具磨损量的预测.

综上所述, 智能故障诊断和寿命预测方法已经取得了较大发展和显著成就, 但仍存在一定的局限性: (1) 为实现机械装备的故障诊断和寿命预测等多种任务, 均需逐一地训练智能诊断和智能寿命预测等多个模型, 无法利用同一智能模型实现多个任务, 这不仅增加了设备开发、部署与维护的成本, 也极大地限制了智能方法的实用性; (2) 初始异常检测作为实现寿命预测的关键步骤之一, 传统方法通过提取时频统计特征来建立健康指标存在阈值难选取、依赖人工经验等问题.

为解决上述局限性, 基于多任务学习, 提出一种用于机械装备智能诊断(初始异常检测和故障诊断)与智能寿命预测等多种任务的健康监测方法. 首先, 利用JS散度(Jensen-Shannon divergence)和相关分析实现设备

图1 (网络版彩图)智能故障诊断与寿命预测的常规技术路线

Figure 1 (Color online) General procedure of intelligent fault diagnostic and prognostic.

初始异常检测,可自动地提取设备初始故障点;其次,结合多任务学习(Multi-task learning)的知识共享机制,利用CNN和长短时记忆网络(Long short-term memory, LSTM)建立可同时实现故障诊断与寿命预测的智能模型,仅需训练和部署一个模型即可同时获得设备的故障信息与剩余使用寿命(Remaining useful life, RUL)信息;最后,利用轴承全寿命加速退化实验数据,通过与领域内先进方法对比,验证所提方法在初始异常检测、故障诊断与寿命预测等多种任务的有效性和可靠性,为工业设备的智能诊断与预测性维护提供重要的决策依据。

2 智能诊断与预测模型基础理论

2.1 卷积神经网络

卷积神经网络是一种利用卷积运算来学习或提取特征的深度前馈神经网络,通常由多个卷积层(Convolutional layer)和池化层(Pooling layer)交替堆叠、以及顶端少数的全连接层(Fully-connected layer)构成,具有稀疏交互(Sparse interactions)、权值共享(Parameter sharing)、等变表示(Equivariant representations)等特性^[18]。近年来,卷积神经网络在智能故障诊断领域“大放异彩”,成为了机械装备健康监测与故障诊断的关键工具,也是提升智能诊断模型准确性与可靠性的有效途径。

卷积层对输入的特征图进行非线性卷积操作。神经网络中的卷积核运算是一种类似信号分析中的卷积运算的互相关操作。不同于信号处理中的卷积运算,神经网络中的卷积核运算无需对原始数据进行反转,这样既可减小计算量,又可实现相同的特征提取能力,其计算过程可表述为

$$\mathbf{V} = \mathbf{W} * \mathbf{Z}, \quad (1)$$

式中, $\mathbf{W} \in \mathbb{R}^{m \times n}$ 为卷积核, $\mathbf{Z} \in \mathbb{R}^{M \times N}$ 为输入的特征表示, $\mathbf{V} \in \mathbb{R}^{(M-m+1) \times (N-n+1)}$ 为卷积形成新的特征表示。即 $\mathbf{V}$ 是对输入表示 $\mathbf{Z}$ 的高层表示, 包含有 $\mathbf{Z}$ 的局部特征信息。

池化层的功能是对卷积层输出进行下采样。通过池化层的下采样操作后, 输出特征的维度降低, 网络参数规模降低, 可以有效地减轻过拟合的风险。常用的池化方式包括以下两种。

(1) 最大值池化, 即将特征图的特定区域内的最大

值作为输出, 其数学模型为

$$\mathbf{Z}_{m,n}^d = \max_{i \in \mathcal{R}_{m,n}^d} \{\mathbf{x}_i\}. \quad (2)$$

(2) 均值池化, 即将特征图的特定区域内的平均值作为输出, 其数学模型为

$$\mathbf{Z}_{m,n}^d = \frac{1}{mn} \sum_{i \in \mathcal{R}_{m,n}^d} \{\mathbf{x}_i\}, \quad (3)$$

其中, $\mathbf{x}_i$ 为区域 $\mathcal{R}_{m,n}^d$ 内的神经元激活后的值, $\mathbf{Z}_{m,n}^d$ 为池化操作后的特征图。

卷积和池化后得到的特征图经特定的激活函数进行非线性激活, 引入非线性因素。如果深度特征不经非线性激活直接作为下一层卷积层的输入, 则网络无法拟合复杂的非线性映射。非线性激活的表达式为

$$\mathbf{Z}^l = \sigma(\mathbf{W}^l * \mathbf{Z}^{l-1} + \mathbf{b}^l), \quad (4)$$

其中, $\mathbf{Z}^l$ 为CNN第 l 层的输出, $\mathbf{Z}^{l-1}$ 为第 $l-1$ 层的输出, $\mathbf{b}^l$ 为网络偏置, $\sigma(\cdot)$ 为非线性激活函数, 如Sigmoid函数、ReLU函数和Tanh函数等。

2.2 长短时记忆网络

时序信息是指在时间维度上具有相互关联的信息, 如语音、文本等信息。传统的前馈神经网络无法建立不同时刻信息的关系, 因而在时间序列处理上表现不佳。长短时记忆网络是一种改进的基于门控的网络, 能够建立当前时刻状态与长期或短期历史时刻的联系, 对时序信息的提取能力较强。

长短时记忆网络由多个记忆单元构成, 记忆单元包括: 输入门、遗忘门、输出门, 通过细胞状态(Cell state)对内部信息进行循环传递, 结构如图2所示。 t 时刻输入 $\mathbf{x}_t$ 和 $t-1$ 时刻隐藏状态 $\mathbf{h}_{t-1}$ 同时进入遗忘门和输入门, 分别决定遗忘哪些信息和输入哪些信息, 并得到候选状态 $\tilde{C}_t$, 其计算过程可表示为

$$f_t = \sigma(\mathbf{W}_f \cdot [\mathbf{h}_{t-1}, \mathbf{x}_t] + \mathbf{b}_f), \quad (5)$$

$$i_t = \sigma(\mathbf{W}_i \cdot [\mathbf{h}_{t-1}, \mathbf{x}_t] + \mathbf{b}_i), \quad (6)$$

$$\tilde{C}_t = \tanh(\mathbf{W}_c \cdot [\mathbf{h}_{t-1}, \mathbf{x}_t] + \mathbf{b}_c), \quad (7)$$

再进行细胞状态更新, 由遗忘门的输出 f_t 、输入门的输出 i_t 、 $t-1$ 时刻细胞状态 $\mathbf{C}_{t-1}$ 和候选状态 $\tilde{C}_t$ 计算得到新的细胞状态 $\mathbf{C}_t$:

图2 (网络版彩图)长短时记忆网络单元结构示意图

Figure 2 (Color online) A cell unit of LSTM.

$$C_t = f_t \cdot C_{t-1} + i_t \tilde{C}_t, \quad (8)$$

最后经由输出门得到 t 时刻的隐藏状态 h_t , 传递到下一个单元:

$$O_t = \sigma(\mathbf{W}_o \cdot [h_{t-1}, \mathbf{x}_t] + b_o), \quad (9)$$

$$h_t = O_t \cdot \tanh(C_{t-1}). \quad (10)$$

在每个LSTM单元中, 输入门、遗忘门和输出门共同影响单元的状态, 动态地调整单元所记忆的信息.

2.3 JS散度

JS散度是KL散度(Kullback-Leibler divergence)的一种衍生度量准则, 是一种满足距离对称性的用来衡量两个分布(P_1 和 P_2)相似度的度量方式, 克服了KL散度距离不对称以及不具有三角不等式性质的局限性, 其定义为

$$D_{JS}(P_1, P_2) = \frac{1}{2}D_{KL}(P_1, m) + \frac{1}{2}D_{KL}(P_2, m), \quad (11)$$

式中, $m = \frac{P_1 + P_2}{2}$, $D_{KL}(\cdot)$ 为KL散度, 其定义为

$$D_{KL}(P_1, P_2) = \sum_{x \in X} P_1(x) \log \frac{P_1(x)}{P_2(x)}. \quad (12)$$

3 机械装备智能健康状态监测方法

基于上述基础理论, 构建涵盖初始异常检测、故障诊断与寿命预测等多种任务的机械装备新型智能健康监测方法.

3.1 初始异常检测方法

初始故障点(也称早期故障点)的识别是建立智能故障诊断与寿命预测模型的关键环节. 传统方法通过提取时频统计特征建立健康指标, 人为设定阈值准则(如 3σ 准则)实现初始故障点的识别, 其缺点在于: (1) 不同健康指标对退化的不同阶段敏感性不同, 例如RMS、峭度对退化初期的敏感性相差巨大, 以RMS、峭度作为健康指标, 会得出不同的状态划分结果; (2) 3σ 准则易误触发, 可能输出错误的结果; (3) 需要人为选取正常状态的部分数据来建立阈值, 人为主观因素的引入, 使得结果可信度降低.

为自动地实现机械装备的初始故障点识别, 提出一种基于JS散度和相关分析的初始异常检测方法. 所提初始异常检测方法旨在通过对设备运行状态数据的分析, 计算不同时刻设备状态的相似性来自动地检测与识别机械装备的初始故障点, 可避免传统方法存在的阈值难选取和依赖人工经验等问题. 如图3所示, 其主要包括以下步骤.

(1) 提取统计特征. 对采集到的监测信号提取多个时域、频域和时频域的统计特征, 表示为 $\mathbf{x} = [x_\mu, x_{rms}, \dots, x_{skew}]$, 不同的时频域统计特征具有不同的物理意义, 可反映设备在不同评价维度下的状态信息.

(2) 计算JS散度矩阵. 将提取的统计特征进行最大最小值归一化, 可得每个时刻设备的状态向量, 记为 $\mathbf{P} = [\mathbf{P}_1, \dots, \mathbf{P}_T]$, T 为设备失效时刻. 若将设备初始时刻的状态向量记为 $\mathbf{P}_1$, 第 t 时刻的状态向量记为 $\mathbf{P}_t$, 根据公式(11)可计算任意两两时刻的JS散度, 可得JS散度矩阵为

$$\mathbf{JS}_{matrix} = \begin{bmatrix} JS_{11} & JS_{12} & \cdots & JS_{1T} \\ JS_{21} & JS_{22} & \cdots & JS_{2T} \\ \vdots & \vdots & \ddots & \vdots \\ JS_{T1} & JS_{T2} & \cdots & JS_{TT} \end{bmatrix}, \quad (13)$$

式中, $\mathbf{JS}_{ij}$ 为第 i 时刻设备状态向量与第 j 时刻设备状态向量的JS散度.

(3) 计算相关系数矩阵. 基于JS散度矩阵计算线性相关系数矩阵, 可得相关系数矩阵为

$$\mathbf{P}_{matrix} = \begin{bmatrix} \rho_{11} & \rho_{12} & \cdots & \rho_{1T} \\ \rho_{21} & \rho_{22} & \cdots & \rho_{2T} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{T1} & \rho_{T2} & \cdots & \rho_{TT} \end{bmatrix}, \quad (14)$$

图 3 (网络版彩图)初始异常检测方法流程示意图

Figure 3 (Color online) Flowchart of infancy anomaly detection.

式中, ρ_{ij} 为JS散度矩阵中第 i 行与第 j 列的线性相关系数。

(4) 获取初始故障点. 对两两相邻时刻的相关系数计算平均值, 并沿主对角线滑动, 得平均相似性曲线:

$$CorLine = \left\{ \frac{\rho_{i,i} + \rho_{i,i+1} + \rho_{i+1,i} + \rho_{i+1,i+1}}{4} \right\}_{i=1}^{T-1}, \quad (15)$$

其最小值对应的时刻即为初始故障点:

$$\text{初始故障点} = \text{argmin}(CorLine). \quad (16)$$

3.2 基于多任务学习的智能故障诊断与剩余寿命预测方法

为使智能模型具有同时执行多个任务(如故障诊断和寿命预测)的能力, 提出基于多任务学习的智能故障诊断与寿命预测方法, 该方法的整体流程如图4所示. 利用卷积神经网络构建特征提取器同时为多个任务提取多源时序信息融合的多尺度共享特征知识, 再将提取到的共享特征分别输入到智能故障诊断和智能寿命预测两个分支网络中, 进一步学习特定任务的私有特征, 最终经过多任务联合训练, 同时实现智能故障诊断与寿命预测. 其具体原理和实施步骤如下.

(1) 收集数据. 通过布置不同的传感器来采集机械装备或关键零部件运行过程中的信号, 如振动、力、电流、声发射等, 收集多源传感器信息.

(2) 数据分析与预处理. 剔除因人为因素或环境因素导致的异常数据与“脏数据”.

(3) 退化评估与状态划分. 通过前述初始异常检测算法, 将设备历史数据划分为不同的健康状态, 作为训

练数据的故障标签与剩余寿命信息, 得到用于训练智能模型的数据集:

$$D = \{(\mathbf{X}, Y, RUL)\}, \quad (17)$$

式中, $\mathbf{X}$ 为采集到的监测数据, 即多任务智能模型的输入; Y 和 RUL 分别为对应的故障类型标签和剩余寿命信息.

(4) 构建含多个分支的多任务智能模型. 故障诊断任务选用基于交叉熵损失的分器, 剩余寿命预测任务选择SmoothL1回归, SmoothL1是Huber loss的一种, 其特点是对于噪声点、异常值不敏感, 具有较强的鲁棒性, 表达式为

$$J_{RUL} = \begin{cases} \frac{1}{2} \sum_{n=1}^N (y_{true}^n - y_{pred}^n)^2 / N, & |y_{true}^n - y_{pred}^n| < 1, \\ \sum_{n=1}^N (|y_{true}^n - y_{pred}^n| - 0.5) / N, & |y_{true}^n - y_{pred}^n| \geq 1, \end{cases} \quad (18)$$

式中, N 为样本量, y_{true}^n 为第 n 个样本的真实值, y_{pred}^n 表示第 n 个样本的预测值.

(5) 多任务智能模型的联合训练. 基于建立的数据集, 进行故障诊断与剩余使用寿命预测. 多任务联合损失函数可表示为

$$J(w) = \lambda_c J_c(f_c(D, w)) + \lambda_r J_r(f_r(D, w)), \quad (19)$$

式中, $f_c(\cdot)$ 和 $f_r(\cdot)$ 分别代表诊断和预测任务, $J_c(\cdot)$ 即为交叉熵损失, $J_r(\cdot)$ 为SmoothL1损失, λ_c 和 λ_r 分别代表故障诊断与寿命预测任务的权重, 本文中分别设置为0.2和0.8.

图 4 (网络版彩图)基于多任务学习的智能故障诊断与寿命预测方法

Figure 4 (Color online) Intelligent diagnostic and prognostic method based on multi-task learning.

(6) 测试与验证. 将训练得到的多任务智能模型应用于机械装备监测数据的测试并验证所提方法的有效性和可靠性.

4 实验设置与评价指标

4.1 实验平台简介

利用西安交通大学联合浙江长兴昇阳科技有限公司公开发布的轴承加速寿命实验数据集——XJTU-SY轴承数据集^[19]对所提方法的有效性和可靠性进行验证. 此数据集将滚动轴承这一机械装备中的关键部件作为试验对象进行加速寿命试验, 得到了不同工况下多个轴承的全寿命退化数据, 更加准确和真实地刻画了机械装备的运行退化过程.

轴承加速寿命实验平台如图5所示, 交流电动机通过转轴与测试轴承连接, 液压加载系统在水平方向施加径向负载, 转速控制器调节转速, 液压控制系统调节负载, 实现不同工况下的轴承退化. 测试轴承为滚动轴承, 具体参数见表1, 试验共包括3个工况, 具体工况参数见表2.

试验中, 分别在测试轴承的外壳体的竖直和水平方向布置单向加速度传感器, 采集振动加速度信号. 采样频率为25.6 kHz, 每分钟采样一次, 单次采样时长

为每分钟的前1.28 s. 试验停止条件为振动加速度的幅值大于健康状态运转时的最大幅值的10倍. 实验结束后, 拆取下故障轴承, 观察损坏情况并记录. 实验最终得到3个工况共计15个轴承的退化数据, 各轴承具体信息见表3.

4.2 评价指标

为了客观合理地评价不同方法的性能, 需要选取合理的评价指标. 分类任务常用的评价指标包括查准

图 5 (网络版彩图)轴承加速寿命实验平台^[19]

Figure 5 (Color online) Testbed of rolling element bearings [19].

表 1 实验轴承参数

Table 1 Parameters of the tested bearings

参数名称	数值	参数名称	数值
额定动载荷 (N)	12820	轴承中径 (mm)	34.55
接触角 (°)	0	滚动体个数	8
额定静载荷 (kN)	6.65	滚动体直径 (mm)	7.92

表 2 实验轴承加速寿命试验工况

Table 2 Operating conditions of the tested bearings

工况	1	2	3
转速 (r/min)	2100	2250	2400
径向力 (kN)	12	11	10

表 3 实验轴承数据集信息一览表

Table 3 XJTU-SY bearing datasets

工况	轴承	试验时长 (min)	失效类型
1	Bearing1_1	123	外圈
	Bearing1_2	161	外圈
	Bearing1_3	158	外圈
	Bearing1_4	122	保持架
	Bearing1_5	52	内圈、外圈
2	Bearing2_1	491	内圈
	Bearing2_2	161	外圈
	Bearing2_3	533	保持架
	Bearing2_4	42	外圈
	Bearing2_5	339	外圈
3	Bearing3_1	2538	外圈
	Bearing3_2	2496	内圈、滚动体、保持架、外圈
	Bearing3_3	371	内圈
	Bearing3_4	1515	内圈
	Bearing3_5	114	外圈

率和召回率. 查准率 *Precision*: 所有被预测为正样本中, 真实为正样本的比例:

$$Precision = TP / (TP + FP) \times 100\%, \quad (20)$$

召回率 *Recall*: 所有真实为正样本中, 被准确预测为正样本的比例:

$$Recall = TP / (TP + FN) \times 100\%. \quad (21)$$

在类别不均衡时, 单一的查准率或召回率指标都

会失真, 此时需要使用 F_1 分数这一综合指标来评价算法的性能. F_1 分数是查准率或召回率的数学调和平均值, 是一个综合指标, 越高代表算法性能越好, 且只有在查准率与召回率均较大的情况下, F_1 分数才会较高, 其表达式为

$$F_1 = \frac{2}{1/Precision + 1/Recall} = \frac{2 \times Precision \times Recall}{Precision + Recall} \times 100\%. \quad (22)$$

利用均方根误差 (Root mean square error, *RMSE*) 来评价剩余使用寿命预测的误差, 其数学表达式为

$$RMSE = \sqrt{\frac{1}{N} \sum_{n=1}^N (y_{true}^n - y_{pred}^n)^2}. \quad (23)$$

5 算法验证

5.1 初始异常检测方法验证

对采集到的两通道 (水平和垂直方向) 原始振动信号提取 16 种常用的时频统计指标 (均值、均方根值、方差、方根幅值、平均幅值、峰值、重心频率、峭度、斜度、峰值指标、波形指标、脉冲指标、裕度指标、峰峰值、均方频率、频率方差), 得到尺寸为 $T \times 32$ 的二维数据组, T 为设备的总寿命 (单位为 min). 按照前述步骤对所有实验轴承进行初始异常检测, 可得初始故障点, 其结果见表 4.

表 4 初始故障点识别结果

Table 4 Experimental results of infancy anomaly detection

工况	轴承	总实验时长 (min)	初始故障点 (min)
1	Bearing1_1	123	77
	Bearing1_2	161	35
	Bearing1_3	158	59
	Bearing1_5	52	33
2	Bearing2_1	491	454
	Bearing2_2	161	46
	Bearing2_3	533	325
	Bearing2_4	42	30
	Bearing2_5	339	122
3	Bearing3_1	2538	2400
	Bearing3_3	371	340
	Bearing3_4	1515	1418
	Bearing3_5	114	7

以Bearing1_1、Bearing2_1和Bearing3_1三个实验轴承为例, 将其中间结果进行可视化, 如图6所示. 图6(a)为Bearing1_1的JS散度矩阵, 图6(b)为对应线性相关系数矩阵. 可以清晰地看出不同状态的跃迁, 再通过滑动计算相邻时刻的平均相似性, 得到平均相似性曲线, 如图6(c)所示, 输出最小值对应的时刻(77 min), 即为初始故障点. 同理可得Bearing2_1和Bearing3_1的初始故障点分别为454和2400 min.

5.2 多任务智能模型训练数据集标定

基于初始异常检测方法识别的初始故障点, 对所采集的数据进行样本划分和标签构造. 在相邻时刻的多个采样信号中各选取长度为8192的信号段, 构造出长、宽、高为 $8192 \times 2 \times S$ 的三维时序样本, “2”为两个方向的传感器输入, “S”为时间步长(设置为3). 为获得足够的样本量, 在长度方向进行滑动, 构造更多的样本, 如图7所示.

图6 (网络版彩图)初始异常检测结果. JS散度矩阵: (a) Bearing1_1; (d) Bearing2_1; (g) Bearing3_1. 相关系数矩阵: (b) Bearing1_1; (e) Bearing2_1; (h) Bearing3_1. 平均相似性曲线: (c) Bearing1_1; (f) Bearing2_1; (i) Bearing3_1

Figure 6 (Color online) Experimental results of infancy anomaly detection. JS divergence matrix: (a) Bearing1_1; (d) Bearing2_1; (g) Bearing3_1. Correlation coefficient matrix: (b) Bearing1_1; (e) Bearing2_1; (h) Bearing3_1. Mean similarity curve: (c) Bearing1_1; (f) Bearing2_1; (i) Bearing3_1.

图 7 (网络版彩图) 样本构造示意图

Figure 7 (Color online) Segmentation of training and testing samples.

为了验证所提方法的性能, 选择各工况的第一个轴承作为测试轴承, 其余轴承数据用于训练智能模型. 由于 Bearing2_3 是唯一保持架故障轴承, 无其余轴承进行同类型故障测试, 故不参与训练; Bearing1_5 和 Bearing3_2 为混合故障, 亦不参与训练. 由于部分轴承的正常运转时间较长, 为便于对比分析, 统一选择初始故障点前 30 分钟视为正常样本, 其后为故障样本(当初始故障点小于 30 时, 则将初始故障点前所有时刻的数据作为正常样本).

将正常、外圈故障、内圈故障标签分别设置为 $[0, 1, 2]$. 将 RUL 归一化为 $0 \sim 1$ 的值, 即初始故障点之前 RUL 为 1, 失效点 RUL 为 0. 以 Bearing1_1 为例, 故障类型标签和 RUL 标签如图 8 所示. 最终得到用于训练多任务智能模型的数据集, 见表 5.

5.2 基于多任务学习的智能故障诊断与剩余寿命预测方法验证

将所提多任务智能模型与现有的故障诊断及剩余使用寿命预测方法进行对比. 多任务智能模型的

图 8 (网络版彩图) Bearing1_1 数据标定结果. (a) 故障标签; (b) 剩余使用寿命信息

Figure 8 (Color online) Annotation results of Bearing1_1. (a) Fault label; (b) RUL label.

网络结构与参数如图 9 所示, 其主要由特征提取器、故障诊断分支网络和寿命预测分支网络三个部分组成. 多任务模型的构建、训练与测试均在 Pytorch 深度学习框架下基于 Python 编程语言环境完成, 训练时批量大小为 256, 在训练数据集上迭代更新 30 次 (Epoch 设定为 30). 网络的重要超参数主要利用 Scikit-Learn 框架中网格搜索 (Grid search) 技术来选取或根据实验验证进行设定. 采用 Adam 优化算法来优化更新多任务网络模型参数^[20], 并使用权重衰减和学习率衰减技术来提升网络的泛化性能, 同时使网络的收敛过程更加稳定.

表 5 训练数据集

Table 5 Details of training and testing datasets

	实验轴承	样本量		
		正常	外圈故障	内圈故障
训练集	Bearing1_2, Bearing1_3, Bearing2_2, Bearing2_4, Bearing2_5, Bearing3_3, Bearing3_4, Bearing3_5	960	2680	512
测试集	Bearing1_1, Bearing2_1, Bearing3_1	60	245	37

图 9 (网络版彩图)多任务网络结构与参数

Figure 9 (Color online) Architecture and hyperparameters of the proposed method.

需特殊说明, 所提多任务方法采用多任务联合训练, 仅训练一个模型即可同时得出故障诊断和剩余使用寿命预测结果; 对比方法均为单任务训练, 单独训练两个模型, 分别进行故障诊断和剩余使用寿命预测, 对比方法如下。

(1) LSTM^[21]: 基于堆栈LSTM的时间序列预测方法, 由两层LSTM组成, 输入为时频统计特征。实验中, 分别用单任务实现故障诊断与RUL预测。

(2) WDCNN^[22]: 一种新型CNN结构, 由5层卷积层和全连接层组成, 特点在于第一层卷积层使用的卷积核为 64×1 , 步长为 16×1 。实验中, 分别用单任务实现故

障诊断与RUL预测, 训练超参数设置与所提方法一致。

(3) 对所提方法采用单任务训练, 即分别单独训练两个任务分支, 得到两个模型, 分别来实现故障诊断与RUL预测。

实验中, 对每个测试轴承均进行3次实验, 取平均值作为最终结果, 实验结果见表6, 将实验结果以柱状图进行可视化, 如图10所示。同时为了直观展示不同方法的性能, 对剩余使用寿命预测结果进行可视化, 如图11所示。由表6可以看出, 所提方法通过多任务联合训练不仅同时实现了故障诊断和剩余使用寿命预测, 且诊断的平均 F_1 分数为0.909, 剩余使用寿命预测的平均

表 6 故障诊断与寿命预测验证结果

Table 6 Validation results of intelligent fault diagnostic and prognostic

轴承编号	LSTM		WDCNN		所提方法单任务		所提方法多任务	
	F_1	$RMSE$	F_1	$RMSE$	F_1	$RMSE$	F_1	$RMSE$
Bearing1_1	0.306	0.211	0.953	0.310	0.958	0.188	0.958	0.155
Bearing2_1	0.467	0.101	0.606	0.121	0.790	0.083	0.932	0.120
Bearing3_1	0.867	0.219	0.805	0.294	0.931	0.242	0.837	0.147
平均值	0.547	0.177	0.788	0.242	0.893	0.171	0.909	0.141

图 10 (网络版彩图)故障诊断与寿命预测验证结果. (a) F_1 分数; (b) $RMSE$ Figure 10 (Color online) Validation results of intelligent fault diagnostic and prognostic. (a) F_1 score; (b) $RMSE$.

$RMSE$ 为0.141, 均优于对比方法, 具有优良的故障诊断和寿命预测性能.

智能故障诊断任务方面, 对比LSTM和WDCNN的实验结果可以看出, LSTM的诊断性能(平均 F_1 分数为0.547)明显劣于WDCNN的诊断性能(平均 F_1 分数为0.788), 是因为WDCNN具有更强的深度特征提取能力使得其提取的特征分辨性优于LSTM方法中人为提取的特征. 对比所提方法在单任务和多任务上的实验结果可知, 多任务模型仅在Bearing3_1测试时其诊断精度(F_1 分数为0.837)略逊于单任务模型的诊断精度(F_1 分数为0.931), 而对其余轴承测试的诊断精度, 多任务模型的性能(F_1 分数)均优于单任务模型. 总体来看, LSTM, WDCNN和单任务模型的平均 F_1 分数分别为0.547, 0.788和0.893, 而所提多任务方法的平均 F_1 分数为0.909, 实验结果证明所提多任务方法在智能故障诊断任务上总体性能优于其他对比方法. 而且, 多任务模型对不同测试轴承的 F_1 分数最小值为0.837, 表明所提方法可以在类别不均衡的情况下取得较好的故障诊断

效果, 并具有较优的准确性和泛化性.

智能寿命预测任务方面, 同样对比LSTM和WDCNN的实验结果可以看出, LSTM的预测性能($RMSE$ 分别为0.211, 0.101和0.219)明显优于WDCNN的预测性能($RMSE$ 分别为0.310, 0.121和0.294), 是因为LSTM在时序信息提取方面的优良特性使得其在智能寿命预测任务中取得较好的表现. 对比所提方法在单任务和多任务上的实验结果可知, 多任务模型仅在Bearing2_1测试时其预测性能($RMSE$ 为0.120)略逊于单任务模型的预测性能($RMSE$ 为0.083), 而对其余轴承测试的预测性能, 多任务模型($RMSE$ 值分别为0.155和0.147)均优于单任务模型($RMSE$ 值分别为0.188和0.242). 总体来看, LSTM, WDCNN和单任务模型的平均 $RMSE$ 为0.177, 0.242和0.171, 而所提多任务方法的平均 $RMSE$ 为0.141, 实验结果证明所提多任务方法在智能寿命预测任务上总体性能优于其他对比方法. 而且, 由图11可以看出, 所提方法预测的寿命曲线具有较好的阶段性和单调性, 即正常状态下剩余使用寿命

图 11 (网络版彩图)所提方法与对比方法的剩余使用寿命预测结果. LSTM方法: (a) Bearing1_1; (b) Bearing2_1; (c) Bearing3_1. WDCNN方法: (d) Bearing1_1; (e) Bearing2_1; (f) Bearing3_1. 单任务方法: (g) Bearing1_1; (h) Bearing2_1; (i) Bearing3_1. 多任务方法: (j) Bearing1_1; (k) Bearing2_1; (l) Bearing3_1

Figure 11 (Color online) Predicted RUL results of the compared methods and the proposed methods. LSTM: (a) Bearing1_1; (b) Bearing2_1; (c) Bearing3_1. WDCNN: (d) Bearing1_1; (e) Bearing2_1; (f) Bearing3_1. Proposed method with single task: (g) Bearing1_1; (h) Bearing2_1; (i) Bearing3_1. Proposed method with multi-task: (j) Bearing1_1; (k) Bearing2_1; (l) Bearing3_1.

保持在1.0附近, 当轴承开始退化, 剩余使用寿命曲线逐渐下降, 并与真实剩余使用寿命具有较好的贴合度, 表明所提方法可以实现准确的剩余使用寿命预测. 且对每个轴承的预测性能较均衡, 无严重过拟合现象, 多任务智能模型的泛化性较强.

通过对比所提方法的单任务模型和多任务模型可以看出, 虽然单任务训练所得到的模型其测试效果在个别轴承上优于多任务训练所得的模型, 但是单任务训练得到的模型平均 F_1 分数与 $RMSE$ 分别为0.893和0.171, 均劣于多任务训练所得模型的0.909和0.141, 即

表明所提方法采用多任务联合训练所得到的多任务模型具有更好的泛化性能.

综上所述可得, 所提多任务方法可以同时实现故障诊断和寿命预测, 并与其他三种方法相比, 具有更高的准确性和泛化性, 可为实际维修决策提供更为可靠的综合依据, 具有较强的实用价值.

为了进一步验证多任务模型在深度特征提取能力和泛化性能方面的优越性, 分析多任务学习中知识共享机制的作用, 分别将训练集输入到多任务模型、单任务诊断模型、单任务预测模型中, 可得特征提取器

输出的深度特征, 再利用t-SNE (t-distributed stochastic neighbor embedding)对深度特征降维后进行故障类别特征分布可视化和外圈故障的剩余使用寿命分布可视化, 其结果如图12所示。

智能故障诊断任务方面, 对比图12(a)~(c)可知, 多任务模型和单任务诊断模型提取的深度特征对于不同故障类型的样本均具有明显可分辨性, 表现为类内聚集度高, 类间差异性大, 即模型对不同故障的识别性能较好; 而单任务预测模型所提取到的不同故障特征的分辨性较差, 不同故障的特征相互重叠。智能寿命预测任务方面, 单任务诊断模型提取到的轴承全寿命周期样本深度特征的剩余使用寿命分布没有明显的趋势性, 如图12(e)所示; 单任务预测模型提取到的轴承全寿命周期样本深度特征的剩余使用寿命分布中心随着剩余使用寿命的减小而有规律偏移, 但趋势性不强, 如图12(f)所示; 而多任务模型提取到的轴承全寿命周期样本深度特征的剩余使用寿命分布中心随着剩余使用寿命的减小而向上有规律地偏移, 且趋势性较强, 如图12(d)所示。由此可知, 与单任务模型相比, 多任务模型所提取到的深度特征在故障诊断与寿命预测两个任务上均取得了最优的效果。

以上实验表明, 所提多任务方法通过多任务联合

训练, 可以实现不同任务的知识共享, 既能提取到类别信息, 又能提取到时序信息, 实现了多种信息的融合交互。不同任务间的共通性和差异性使得所提方法可以在故障诊断和故障预测上同时取得优异的表现, 具有良好的泛化性能。同时得益于多任务学习的知识共享机制, 多任务方法只需训练一个模型, 而单任务方法需要训练得到两个模型, 多任务模型的参数规模和运算时间复杂度大大降低, 具有更强的实际应用价值。

6 结论与展望

针对传统智能方法无法同时实现机械装备状态监测的多种任务这一局限性, 基于多任务学习, 提出一种用于机械装备智能诊断(初始异常检测和故障诊断)与智能寿命预测等多种任务的健康监测方法, 仅需训练和部署一个模型即可同时获得设备的故障信息与剩余使用寿命信息。通过轴承全寿命退化实验对所提方法进行验证, 并与多种现有智能方法进行对比, 主要结论如下。

(1) 基于JS散度和相关分析的初始异常检测方法, 解决了传统健康状态划分方法存在阈值难选取、依赖

图12 (网络版彩图)深度特征可视化。(a) 多任务模型故障类型分布; (b) 单任务诊断模型故障类型分布; (c) 单任务预测模型故障类型分布; (d) 多任务模型RUL分布; (e) 单任务诊断模型RUL分布; (f) 单任务预测模型RUL分布

Figure 12 (Color online) Visualization of deep features. (a) Multi-task model for fault classification; (b) diagnostic model for fault classification; (c) prognostic model for fault classification; (d) multi-task model for RUL prediction; (e) diagnostic model for RUL prediction; (f) prognostic model for RUL prediction.

人工经验等问题, 可自动提取机械装备的初始故障点。

(2) 所提集多任务的新型智能监测方法通过建立含多个分支的智能模型, 仅需训练和部署一个模型即可同时实现机械装备的故障诊断与寿命预测, 相较传统单任务方法, 大大降低了开发和部署成本, 为机械装备的智能诊断与预测性维护提供更全面的决策依据, 具有较强的工程应用价值。

(3) 所提方法通过多任务学习的知识共享机制, 使

不同任务均取得较好的表现, 通过实验对比验证, 所提方法具有优于传统方法的有效性和可靠性。

所提多任务方法相对于传统智能故障诊断或智能寿命预测方法取得了一定的进展, 但也存在有待重点关注并进一步深入完善的方面, 例如: 多任务模型的工程实际部署问题。今后将考虑利用在线学习或增量学习等技术, 进一步优化完善模型的在线实时更新性能, 以便满足实际工业场景的需求。

参考文献

- 1 Zhou J. Intelligent manufacturing is the main direction of “made in China 2025” (in Chinese). *China Mech Eng*, 2015, 26: 2273–2284 [周济. 智能制造——“中国制造2025”的主攻方向. *中国机械工程*, 2015, 26: 2273–2284]
- 2 Lin J, Zhao M. Dynamic signal analysis for speed-varying machinery: A review (in Chinese). *Sci Sin Tech*, 2015, 45: 669–686 [林京, 赵明. 变转速下机械设备动态信号分析方法的回顾与展望. *中国科学: 技术科学*, 2015, 45: 669–686]
- 3 Wang Z, Huang R Y, Li J P, et al. Multi-task feature sharing neural network used for fault diagnosis and prognosis (in Chinese). *Chin J Scient Instr*, 2019, 40: 169–177 [王震, 黄如意, 李霁蒲, 等. 一种用于故障分类与预测的多任务特征共享神经网络. *仪器仪表学报*, 2019, 40: 169–177]
- 4 Lei Y G, Jia F, Kong D T, et al. Opportunities and challenges of machinery intelligent fault diagnosis in big data era (in Chinese). *J Mech Eng*, 2018, 54: 94–104 [雷亚国, 贾峰, 孔德同, 等. 大数据下机械智能故障诊断的机遇与挑战. *机械工程学报*, 2018, 54: 94–104]
- 5 Tao F, Zhang H, Qi Q L, et al. Theory of digital twin model and its ten applications (in Chinese). *Comput Integ Manuf Syst*, 2021, 27: 1–15 [陶飞, 张贺, 戚庆林, 等. 数字孪生模型构建理论及应用. *计算机集成制造系统*, 2021, 27: 1–15]
- 6 Tao F, Qi Q L. Service-oriented smart manufacturing (in Chinese). *J Mech Eng*, 2018, 54: 11–23 [陶飞, 戚庆林. 面向服务的智能制造. *机械工程学报*, 2018, 54: 11–23]
- 7 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
- 8 Zhao Z, Li T, Wu J, et al. Deep learning algorithms for rotating machinery intelligent diagnosis: An open source benchmark study. *ISA Trans*, 2020, 107: 224–255
- 9 Jiao J, Zhao M, Lin J, et al. A comprehensive review on convolutional neural network in machine fault diagnosis. *Neurocomputing*, 2020, 417: 36–63
- 10 Ai J L, Yang X Z. Fault diagnosis of aero-engine based on self-adaptive neural network (in Chinese). *Sci Sin Tech*, 2018, 48: 326–335 [艾剑良, 杨曦中. 一种基于自适应神经网络的航空发动机故障诊断方法. *中国科学: 技术科学*, 2018, 48: 326–335]
- 11 Xiang J W. Numerical simulation driving generative adversarial networks in association with the artificial intelligence diagnostic principle to detect mechanical faults (in Chinese). *Sci Sin Tech*, 2021, 51: 341–355 [向家伟. 机械故障数值模拟驱动的生成式对抗网络及智能诊断原理. *中国科学: 技术科学*, 2021, 51: 341–355]
- 12 Shao H, Jiang H, Wang F, et al. An enhancement deep feature fusion method for rotating machinery fault diagnosis. *Knowledge-Based Syst*, 2017, 119: 200–220
- 13 Shao H D, Jiang H K, Zhao K, et al. A novel tracking deep wavelet auto-encoder method for intelligent fault diagnosis of electric locomotive bearings. *Mech Syst Signal Processing*, 2018, 110: 193–209
- 14 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
- 15 Zhang B, Zhang S, Li W. Bearing performance degradation assessment using long short-term memory recurrent network. *Comput Industry*, 2019, 106: 14–29
- 16 Qin Q B, Wu S C, Hu Y N, et al. Fatigue strength and residual life assessment of high-speed railway vehicle used S38C hollow axles (in Chinese). *Sci Sin Tech*, 2019, 49: 840–850 [秦庆斌, 吴圣川, 胡雅楠, 等. 高速动车组S38C车轴疲劳强度及剩余寿命评价. *中国科学: 技术科学*, 2019, 49: 840–850]

- 17 Huang Z, Zhu J, Lei J, et al. Tool wear predicting based on multi-domain feature fusion by deep convolutional neural network in milling operations. *J Intell Manuf*, 2020, 31: 953–966
- 18 Qiu X P. Neural Network and Deep Learning (in Chinese). Beijing: China Machine Press, 2020 [邱锡鹏, 神经网络与深度学习. 北京: 机械工业出版社, 2020]
- 19 Lei Y G, Han T Y, Wang B, et al. XJTU-SY rolling element bearing accelerated life test datasets: A tutorial (in Chinese). *J Mech Eng*, 2019, 55: 1–6 [雷亚国, 韩天宇, 王彪, 等. XJTU-SY滚动轴承加速寿命试验数据集解读. 机械工程学报, 2019, 55: 1–6]
- 20 Kingma D P, Ba J. Adam: A method for stochastic optimization. arXiv: [1412.6980](https://arxiv.org/abs/1412.6980), 2014
- 21 Zheng S, Ristovski K, Farahat A, et al. Long short-term memory network for remaining useful life estimation. In: Proceedings of 2017 IEEE International Conference on Prognostics and Health Management (ICPHM). IEEE, 2017. 88–95
- 22 Zhang W, Peng G, Li C, et al. A new deep learning model for fault diagnosis with good anti-noise and domain adaptation ability on raw vibration signals. *Sensors*, 2017, 17: 425

Intelligent diagnostic and prognostic method based on multitask learning for industrial equipment

HUANG RuYi^{1,2}, LI JiPu², WANG Zhen², XIA JingYan², CHEN ZhuYun^{2,3} & LI WeiHua^{1,2,3}

¹ Shien-Ming Wu School of Intelligent Engineering, South China University of Technology, Guangzhou 511442, China;

² School of Mechanical & Automotive Engineering, South China University of Technology, Guangzhou 510640, China;

³ Guangdong Artificial Intelligence and Digital Economy Laboratory (Guangzhou), Guangzhou 510335, China

Intelligent fault diagnostic and prognostic methods are significant tools that drive the manufacturing industry into the emerging trend of complex, digital and intelligent. With the rapid development of artificial intelligence and machine learning technology, intelligent fault diagnostic and prognostic methods have been widely investigated and applied in many industry scenarios, including aerospace, petrochemical, and shipping scenarios. However, as an obvious limitation, traditional methods construct and train only one model for a specific task (i.e., either fault diagnosis or remaining useful life prediction), which seriously restricts the practical application of intelligent methods. To address this issue, an intelligent fault diagnostic and prognostic method based on multitask learning is proposed herein for industrial equipment, which can implement the multitasking of infancy anomaly detection, fault diagnosis, and remaining useful life prediction. First, the infancy anomaly detection algorithm was developed based on the Jensen–Shannon divergence and correlation analysis. Second, an intelligent model that can simultaneously implement the tasks of fault diagnosis and remaining useful life prediction was constructed by introducing the multitask learning technology that can monitor the health condition of machines and at the same time obtain the remaining useful life. The experimental results based on the XJTU-SY Rolling Element Bearing Accelerated Life Test Datasets demonstrate that the proposed method is effective and reliable for the multitasking of infancy anomaly detection, fault diagnosis, and remaining useful life prediction, providing a promising solution for intelligent fault diagnostic and prognostic in practical industry applications.

fault diagnosis, remaining useful life prediction, infancy anomaly detection, multitask learning, industrial equipment

doi: [10.1360/SST-2021-0314](https://doi.org/10.1360/SST-2021-0314)