

DOI: 10.19650/j.cnki.cjsi.J2312235

多传感器融合和 MHA-LSTM 的电机轴承剩余寿命预测*

张 苑^{1,2}, 张泰禹¹, 贾民平³, 蔡 骏¹

(1. 南京信息工程大学自动化学院 南京 210044; 2. 江苏省大气环境与装备技术协同创新中心 南京 210044;
3. 东南大学机械工程学院 南京 211102)

摘 要: 轴承作为电机的核心部件, 主要起到支撑引导轴、减小设备摩擦、连接不同设备等作用, 其剩余寿命预测对系统健康管理起着十分重要的作用。针对单一传感器信号通常难以全面描述系统的潜在退化机制, 论文提出一种基于多头注意力机制和长短时记忆神经网络的电机轴承剩余寿命预测模型。首先, 基于马氏距离确定轴承性能退化起始点, 将滚动轴承全寿命周期分为正常阶段与退化阶段; 其次, 使用自编码器自动提取振动信号特征, 并将其与电机电流、轴承温度融合, 构成多源信息特征矩阵; 然后基于多头注意力机制和长短时记忆网络模型动态选择相关度较高的特征, 提高寿命预测的准确性。最后, 采用实验数据进行验证, 结果表明所提出的模型具有更高的准确性。

关键词: 电机轴承; 多传感器融合; 多头注意力机制; 长短期记忆网络; 剩余寿命预测

中图分类号: TH17 **文献标识码:** A **国家标准学科分类代码:** 460.99

Prediction of remaining life of motor bearings using multi-sensor fusion and MHA-LSTM

Zhang Wan^{1,2}, Zhang Taiyu¹, Jia Minping³, Cai Jun¹

(1. Department of Automation, Nanjing University of Information Science and Technology, Nanjing 210044, China;
2. Jiangsu Collaborative Innovation Center for Atmospheric Environment and Equipment Technology, Nanjing 210044, China; 3. School of Mechanical Engineering, Southeast University, Nanjing 211102, China)

Abstract: As a core component of motors, bearings primarily serve functions such as supporting and guiding shafts, reducing friction in equipment, and connecting different components. Predicting the remaining life of bearings is crucial for system health management. However, single sensor signals often fail to comprehensively describe the potential degradation mechanisms of the system. This paper proposes a novel approach for predicting the remaining life of motor bearings based on the multi-head attention mechanism and long short-term memory neural network. Firstly, Mahalanobis distance is used to determine the starting point of bearing performance degradation by dividing the entire life cycle of rolling bearings into normal and degradation phases. Secondly, an Autoencoder is employed to automatically extract vibration signal features, which are subsequently fused with motor current and bearing temperature signal to construct a multi-source information feature matrix. Subsequently, the multi-head attention mechanism and long short-term memory network dynamically select features with high relevance, thereby improving the accuracy of the remaining life prediction. Finally, the model is validated using experimental data, and the results show that the proposed model has higher accuracy.

Keywords: motor bearing; multi-sensor fusion; multi-head attention mechanism; long short-term memory network; remaining life prediction

收稿日期: 2023-12-05 Received Date: 2023-12-05

* 基金项目: 国家自然科学基金资助项目(52077105)、江苏省自然科学基金资助项目(BK20211285)、先进数控和伺服驱动技术安徽省重点实验室(安徽工程大学)开放基金资助项目(XJSK202105)资助

0 引 言

随着科学技术的发展和生产工艺的进步,当代设备日益朝着大型化、复杂化、自动化以及智能化方向发展,对设备的安全可靠运行提出了更高的要求^[1]。在实际工程中,电机是常见的重要生产设备,轴承作为其重要的零部件之一,主要起到支撑引导轴、减小设备摩擦、连接不同设备等作用,其性能状态是否正常对整个装置性能有很大影响,一旦轴承发生疲劳失效,会直接影响整个设备的正常运行,严重时会给工业生产及生命财产带来威胁^[2-3]。为确保设备安全可靠地运行,剩余寿命(remaining useful life, RUL)预测作为状态维护与健康管理的核心,能够为设备的主动维护提供理论支撑和实践依据^[4-5]。

滚动轴承的剩余寿命预测方法可分为基于物理模型驱动、基于数据驱动和两者融合的方法^[2]。基于物理模型的寿命预测方法主要是通过研究部件的损伤机理,探寻部件性能退化规律,进而对部件的剩余寿命进行预测^[6]。研究人员在基于模型驱动的 RUL 预测方面做出了许多相应的研究,Paris-Erdogan 模型是一种裂纹扩展方程,在早期的机械设备剩余寿命预测中被广泛应用^[7]。Qian 等^[8]结合了增强相空间与改进后的 Paris 裂纹生长模型,利用多时间尺度建模方法跟踪轴承的失效演化过程。Liao^[9]提出了一种改进的 Paris-Erdogan 模型,用遗传规划方法捕捉故障发展过程的深层特征实现 RUL 预测。Gupta 等^[10]提出了一种广义应力容量方程来预测滚动轴承的 RUL,对于点接触和线接触,用不同的几何材料参数表示,该模型可以有效地预测滚动轴承在不同工况下的疲劳寿命。由于系统的复杂性和过程的随机性,基于模型驱动的寿命预测方法也有一定的局限性,比如建立准确的机理模型在实际工程中是比较困难的,通常需要根据不同轴承的失效形式建立不同的机理模型进行分析。

基于数据驱动模型多是在获取大量数据的基础上经过训练和拟合形成自动化的决策模型。陈保家等^[11]提出了一种基于注意力机制和残差长短时记忆网络(long short-term memory, LSTM)的剩余使用寿命预测方法,并将该方法在航空发动机实验数据集上进行验证。Ding 等^[12]提出了一种基于深度传递自编码器方法,使用参数迁移和特征迁移来实现将源域中的特征提取模型和剩余寿命回归模型迁移到目标域中。Yang 等^[13]结合图卷积神经网络,从空间和时间两个角度建立门控循环单元实现轴承 RUL 预测。Wei 等^[14]引入 Siamese LSTM 网络对退化阶段进行分类,引入了自适应图卷积网络和自注意机制对轴承剩余寿命进行预测。She 等^[15]提出了提

一种基于 bootstrap 方法的双向门控循环单元用于对轴承剩余寿命的预测。Ren 等^[16]提出了一种基于时间注意力机制的多通道网络,用于工业健康指标的预测,该方法通过通道注意力机制权衡不同通道的贡献,避免时间信息的丢失。尽管基于数据驱动方法较好的预测部件的剩余寿命,但由于过度依赖单一传感器导致模型鲁棒性较差,当传感器信号掺杂大量噪声,信号将难以准确反映轴承现阶段的运行状态。此外,提取特征中会存在与剩余寿命相关度较低的冗余特征,也会降低剩余寿命预测的精度。

随着设备结构的日趋智能化和复杂化,设备经常工作在比较恶劣的环境,采用单一传感器数据通常不足以准确描述系统的潜在退化机制,从而导致 RUL 预测结果的不准确^[17]。为了充分利用多传感器信息,国内外学者在多传感器数据融合的剩余寿命预测方面也有了一定的研究。李乃鹏等^[18]提出了一种融合多传感器数据的数模联动的机械寿命预测方法,通过粒子滤波实现数据与模型的联合动态匹配,融合多传感器数据预测刀具剩余寿命。Ta 等^[19]提出了一种基于多传感器和多特征融合的自适应分阶段 RUL 预测方法。Wang 等^[20]提出了一个门控图卷积网络实现多传感器信号融合和剩余寿命预测。Li 等^[21]提出了分层注意力图卷积网络对传感器的空间依赖关系进行建模,采用双向长短期记忆网络对传感器测量的时间依赖关系进行建模。

上述研究中,更偏重多传感器的深层特征提取对剩余寿命预测的影响。在实际的多个传感器监测中,部分传感器的监测数据特征与设备的退化机制相关度高,在设备退化过程中退化趋势明显;而其他传感器的监测数据特征与设备的退化机制相关性并不高,退化趋势并不明显^[22]。因此,需要充分考虑不同的传感器数据特征对预测结果的影响,强化与设备退化机制相关度高的特征,抑制与设备退化机制相关度低的特征。

针对现有的寿命预测方法在处理多传感器融合预测中所存在的问题,论文提出一种基于多传感器融合的多头注意力(multi-head attention, MHA)和 LSTM 网络模型的电机轴承剩余寿命预测方法。首先利用马氏距离确定轴承性能退化的起始点,将滚动轴承全寿命周期分为正常运行阶段与退化阶段。其次采用自编码器自动提取振动信号特征,并将振动信号特征与电机电流、轴承温度信号融合,构成多源信息特征矩阵,并利用多头注意力机制动态选择相关性较高的特征,然后采用长短时记忆网络模型实现剩余寿命预测,最后通过滚动轴承全寿命试验数据对所提出的模型进行验证,结果表明,所提出的多传感器融合的 MHA-LSTM 模型可显著提高电机轴承 RUL 的预测准确度。

1 理论分析

1.1 马氏距离判定退化起始点原理

马氏距离是由印度统计学家 Mahalanobis 提出的,其通过计算不同数据间的协方差矩阵来表示点与一个分布之间的距离^[23]。

对于数据集 $\mathbf{X} = (\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n)$, $\mathbf{X}_n$ 表示第 n 个数据且每个数据维度为 m , 其中均值为 $\boldsymbol{\mu} = (\boldsymbol{\mu}_1, \boldsymbol{\mu}_2, \dots, \boldsymbol{\mu}_m)^T$ 。若协方差矩阵为 Σ , 数据 $\mathbf{x}$ 为 $\mathbf{x} = (x_1, x_2, \dots, x_m)^T$, 则其马氏距离如下所示:

$$D_M(\mathbf{x}) = \sqrt{(\mathbf{x} - \boldsymbol{\mu})^T \Sigma^{-1} (\mathbf{x} - \boldsymbol{\mu})} \quad (1)$$

式中: Σ^{-1} 为协方差矩阵 Σ 的逆矩阵。

马氏距离可以反映数据之间相关度,由于轴承自身的工业生产流程,使用信号采集设备所采集到的轴承全寿命周期数据中的各个特征具有很高的相关性,因此采用马氏距离用于轴承数据划分具有很高的可行性。对于轴承振动数据集 $\mathbf{X} = (\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n)$, 数据维度为 m 。其中正常数据集记为 $\mathbf{X}_N = (\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_j)$, 正常数据均值为 $\boldsymbol{\mu}_N = (\boldsymbol{\mu}_1, \boldsymbol{\mu}_2, \dots, \boldsymbol{\mu}_m)^T$, 协方差矩阵为 Σ_N , 由式(1)可知正常数据集中每个数据的马氏距离记为:

$$D_N = (D_1, D_2, \dots, D_j) \quad (2)$$

式中: D_j 表示第 j 个正常数据的马氏距离。由式(1)得到的数据 $\mathbf{X}_i$ 与正常数据集均值的马氏距离为:

$$D_M(\mathbf{X}_i) = \sqrt{(\mathbf{X}_i - \boldsymbol{\mu}_N)^T \Sigma_N^{-1} (\mathbf{X}_i - \boldsymbol{\mu}_N)} \quad (3)$$

基于统计学原理,正常数据 $\mathbf{X}_i$, 其马氏距离 $D_M(\mathbf{X}_i)$ 应符合正常数据马氏距离数据集 D_N 的统计分布;若 $\mathbf{X}_i$ 为异常数据, $D_M(\mathbf{X}_i)$ 将不符合 D_N 的统计分布。

若数据集 D_N 的均值为 $\boldsymbol{\mu}_{D_N}$, 标准差为 σ_{D_N} , 由 3σ 准则可知, 大部分正常数据的马氏距离分布于 $(\boldsymbol{\mu}_{D_N} - 3\sigma_{D_N}, \boldsymbol{\mu}_{D_N} + 3\sigma_{D_N})$ 区间中。为了确定退化起始点, 将退化起始点的阈值设定为:

$$Thr = k \cdot \boldsymbol{\mu}_{D_N} \quad (4)$$

式中: k 为阈值系数。

1.2 基于自编码器的特征提取

自编码器是深度学习网络中一种对称的网络结构, 是一种自动特征提取算法, 其结构包括具有编码和解码功能的两部分, 其结构如图1所示。

编码器输入层 x_i 与隐层 h_i 之间的编码关系可表示为:

$$h_i = \sigma_e(w_i x_i + b) \quad (5)$$

式中: w_i 与 b 为编码器的权重偏置, σ_e 为编码层激活函数。隐层表达经过解码后的输出 x'_i 可表示为:

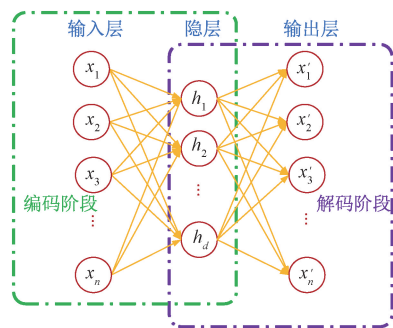


图1 自编码器结构图

Fig. 1 Autoencoder architecture diagram

$$x'_i = \sigma_d(w'_i h_i + b') \quad (6)$$

式中: w'_i 和 b' 分别为解码层的权重和偏置, σ_d 为解码层的激活函数。

通过增加编码器的层数以提高编码器信号压缩以及重构的能力, 对于第 m 层自编码器, 其编码和解码函数分别为下式:

$$h_m = \sigma_e(w_{m,i} h_{m,i} + b_{m,i}) \quad (7)$$

$$h'_m = \sigma_e(w_{m,i+1} h_{m,i+1} + b_{m,i+1}) \quad (8)$$

式中: $w_{m,i}$ 和 $w_{m,i+1}$ 分别为第 m 层 AE 编码器和解码器权重, $b_{m,i}$ 和 $b_{m,i+1}$ 分别为偏置。 h'_m 为第 m 层输入 h_m 的重构。

通过调节权重与偏置式重构误差最小, 及缩小 x'_i 与 x_i 的差距, 则对于整个训练集的代价函数为:

$$\theta = \underset{\theta}{\operatorname{argmin}} \frac{1}{n} \sum_{i=1}^n \mathcal{E}(x_i, x'_i) \quad (9)$$

1.3 多头注意力机制

注意力机制 (attention mechanism)^[24] 可以有选择地聚焦于与任务相关度更高的特征, 抑制无用特征的影响, 从而提高模型的性能。对于给定的 $\mathbf{q}$ 和 $\mathbf{X}$, 选择第 i 个输入信息的概率 α_i 为:

$$\alpha_i = p(z = i | \mathbf{X}, \mathbf{q}) = \operatorname{softmax}(s(x_i, \mathbf{q})) = \frac{e^{s(x_i, \mathbf{q})}}{\sum_{j=1}^N e^{s(x_j, \mathbf{q})}} \quad (10)$$

式中: z 表示索引位置, 输入信息 $\mathbf{X} = [x_1, x_2, \dots, x_N]$, 其维度为 N , $\mathbf{q}$ 为查询矩阵, $s(x_i, \mathbf{q})$ 为注意力评分函数:

$$s(x_i, \mathbf{q}) = \frac{\mathbf{x}_i^T \mathbf{q}}{\sqrt{d}} \quad (11)$$

式中: d 为输入信息的维度。使用的缩放点积注意力函数:

$$\operatorname{Attention}(\mathbf{Q}, \mathbf{K}, \mathbf{V}) = \operatorname{softmax}\left(\frac{\mathbf{QK}^T}{\sqrt{d_k}}\right) \mathbf{V} \quad (12)$$

式中: $\mathbf{Q}$ 为查询矩阵, $\mathbf{K}$ 为关键字矩阵, $\mathbf{V}$ 为关键字的数值矩阵。多头注意力机制是由多个自注意力机制结构组成

的,用于同时处理同一特征信息,其输出为多个自注意力的拼接,如图2所示。这种结构可以更好地捕捉不同特征之间的依赖关系,并进一步提高模型的表现,其表达式为:

$$\text{MultiHead}(\mathbf{Q}, \mathbf{K}, \mathbf{V}) = \text{Concat}(\text{head}_1, \dots, \text{head}_h) \mathbf{W}^o \quad (13)$$

式中: $\text{head}_i = \text{Attention}(\mathbf{Q}\mathbf{W}_i^Q, \mathbf{K}\mathbf{W}_i^K, \mathbf{V}\mathbf{W}_i^V)$, $\mathbf{W}_i^Q, \mathbf{W}_i^K, \mathbf{W}_i^V$ 是映射矩阵权重, $\mathbf{W}^o$ 是输出权重矩阵。

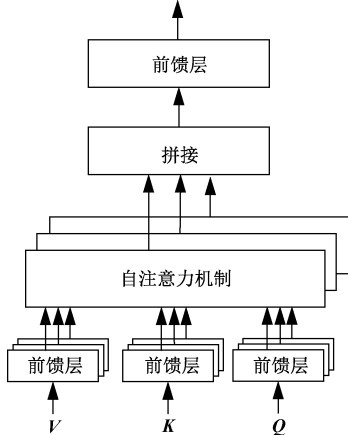


图2 多头注意力机制

Fig. 2 Multi-head attention mechanism

1.4 LSTM 基本原理

LSTM 理论框图如图3所示,其由输入门、遗忘门和输出门组成。LSTM 隐层状态值的更新依赖于前一时刻和当前时刻的隐层状态值,并通过循环反馈进行持续更新,因此 LSTM 具备了长时间保存传递有效信息的能力。

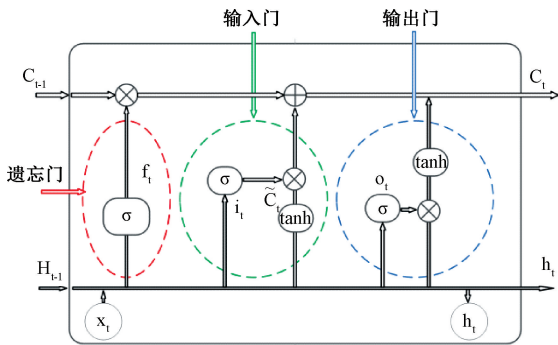


图3 LSTM 理论框

Fig. 3 Theoretical framework of LSTM

输入门决定单元状态中存储的信息,包括两个参数 i_t 和 $\bar{C}_t$ 。由 sigmoid 激活函数得到的 i_t 用于确定更新值;由 tanh 激活函数得到的 $\bar{C}_t$ 表示新的候选向量。其表达式如下:

$$i_t = \sigma(\omega_i \cdot [h_{t-1}, x_t] + b_i) \quad (14)$$

$$\bar{C}_t = \tanh(\omega_{\bar{C}} \cdot [h_{t-1}, x_t] + b_{\bar{C}}) \quad (15)$$

式中: i_t 为新信息的更新权重, ω_i 和 b_i 为输入门的权重和偏置, $\omega_{\bar{C}}$ 和 $b_{\bar{C}}$ 记忆细胞的权重和偏置。

遗忘门决定所记忆的信息是否被遗忘,然后更新信息,将前一时间步的隐层状态 h_{t-1} 与当前状态输入 x_t 送至 sigmoid 函数 σ 中得到遗忘概率 f_t , 如下式所示:

$$f_t = \sigma(\omega_f \cdot [h_{t-1}, x_t] + b_f) \quad (16)$$

式中: σ 为 sigmoid 函数; ω_f 为遗忘门权重; b_f 为遗忘门偏置。

当前细胞状态 $\bar{C}_t$ 由遗忘门、输入门、上一时刻隐层状态和临时记忆状态决定,如下式所示:

$$C_t = f_t \cdot C_{t-1} + i_t \cdot \bar{C}_t \quad (17)$$

输出门输出记忆单元信息:

$$o_t = \sigma(\omega_o \cdot [h_{t-1}, x_t] + b_o) \quad (18)$$

$$h_t = o_t \cdot \tanh(C_t) \quad (19)$$

式中: ω_o 和 b_o 为输出门权重和偏置, $\tanh$ 为激活函数。 h_t 为下一层的隐层状态, o_t 为输出内容。

2 多传感器融合与注意力机制的电机轴承剩余寿命预测

2.1 多传感器融合

多传感器融合主要是通过多传感器采集信息,能够充分收集研究对象状态信息并展现信息背后的深层意义,进而提升剩余寿命预测的精度。在滚动轴承剩余寿命预测应用中,与传统的单一传感器相比,多传感器融合可以提取更多关于对象的信息,更加精准的得到被测事物的有效信息,同时信息中还能包含各个传感器之间的相互关系,反映了系统的整体性能。

由于电机轴承工作环境的复杂多噪性,同时为了反应电机轴承的整体运行状态,将轴承振动信号特征与轴承温度、电机电流信息相融合,构造电机系统的状态特征矩阵 F :

$$F = [v, t, c] \quad (20)$$

式中: v 为归一化后的振动信号特征矩阵, t 和 c 为归一化后的轴承温度和电机电流特征。

2.2 数据切片

数据切片是将数据集按照特定的规则或条件分割成若干个子集的过程,每个子集都包含了数据集中特定的一部分数据。数据切片是为了更好地管理和处理大量的数据,可以将时间序列数据按照特定的时间段进行切片,更好地观察数据在不同时间段内的变化趋势。将数据从单一的点输入变成固定的段输入,能够对数据变化趋势进行更好的预测。在数据切片的过程中,需要考虑多个因素,来确定切片的窗口大小。这些因素的选择和调整通常需要根据具体的数据集和分析目的进行优化和调

整,以达到最佳的数据处理效率和结果。如果切片窗口过小,就无法很好的观测到数据的变化趋势;如果切片的窗口过大,则数据量会变得很庞大,切片之间的重合度太高,数据处理效率太低,预测结果也会变差。

2.3 MHA-LSTM 网络模型

本文提出一种基于 MHA-LSTM 模型的电机轴承剩余寿命预测,图 4 所示为 MHA-LSTM 网络模型结构。将输入信息通过全连接层 (fully-connected, FC) 提高维度,然后由 MHA 层强化关键特征权重,再由两层 LSTM 层深度挖掘隐藏信息,然后使用拼接 (Concatenate) 层将 MHA 与 LSTM 输出特征联合,确保各层之间信息流动最大,减小降维信息损失,最后由展平层 (Fatten) 和全连接层输出 RUL。

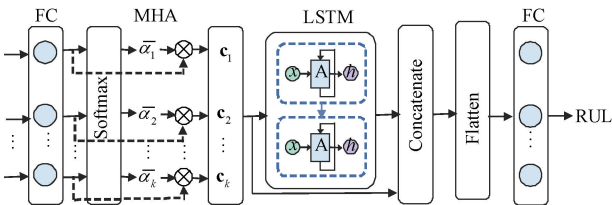


图 4 MHA-LSTM 结构

Fig. 4 MHA-LSTM architecture

2.3 所提出方法实施步骤

本文所提出的基于多传感器融合和 MHA-LSTM 模型的电机轴承剩余寿命预测方法的技术路线如图 5 所示,具体步骤如下:

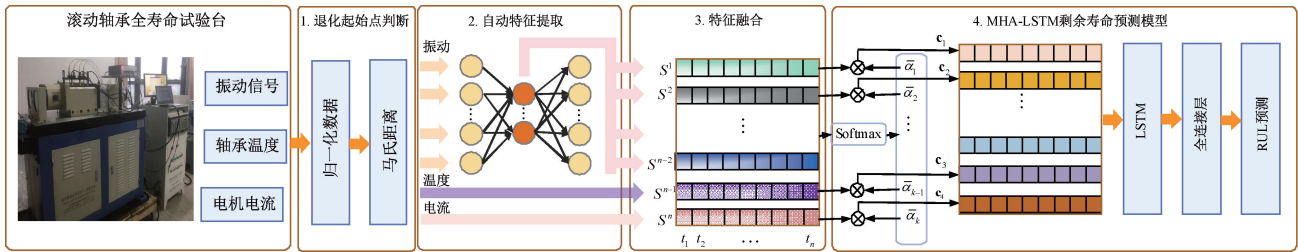
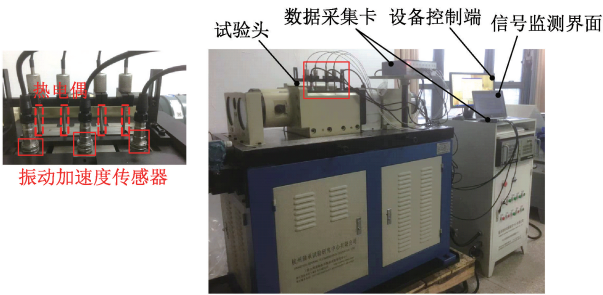


图 5 所提出方法的流程图

Fig. 5 Flowchart of the proposed method

- (1) 采集滚动轴承全寿命试验台中轴承的振动信号、轴承温度和电机电流信号,对三种传感器信号进行归一化预处理;再基于马氏距离的异常值判断方法对预处理后的数据进行退化起始点识别;
- (2) 利用自编码器提取振动信号特征,深度挖掘隐藏特征,减少手动特征提取中主观因素的影响;
- (3) 利用多传感器融合,将电机温度、电流信号与振动特征进行融合,构建全面反映系统运行状态的多尺度信息矩阵;
- (4) 搭建 MHA-LSTM 预测模型,对电机轴承剩余寿命进行预测。

试验台安装了 4 个温度传感器、3 个振动加速度传感器和电机电流传感器,通过上位机监测机器运行状态并保存数据。图 6(c) 是实验头内部结构,工位 1~4 分别对应图 6(b) 中轴承 1~4,可以看出工位 2 和 3 是刚性连接,

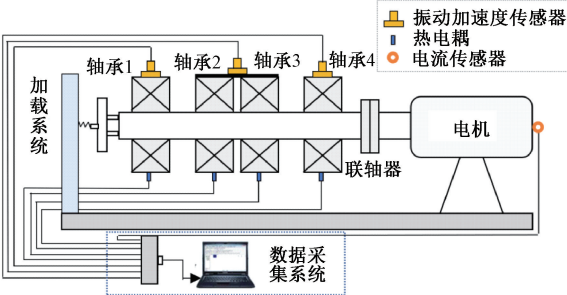


(a) 轴承强化疲劳试验台
(a) Bearing failure test bench

3 试验数据与评价指标

3.1 试验数据说明

为了验证所提出算法的有效性,在 ABLT-1A 轴承寿命强化试验台上开展电机轴承的全寿命试验,实验台如图 6(a) 所示,轴承寿命强化实验台是由实验头座、传动系统、加载系统、润滑系统、电气控制系统、计算机控制系统等部分组成。试验台由传动系统带动轴承转动,能够同时进行 4 个轴承的全寿命周期试验,此外通过加载系统对轴承径向施加一定当量载荷。试验台的示意图如图 6(b) 所示,从图中可以看出传感器的具体安装位置,



(b) 试验台的示意图
(b) Schematic of the test bench

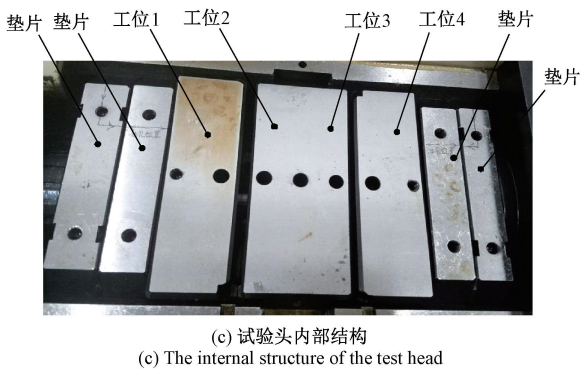


图 6 ALT-1A 全寿命实验台

Fig. 6 ALT-1A full-life cycle test rig

所以只安装了 1 个振动加速度传感器同时监测轴承 2 和轴承 3 的振动加速度信号。

本文所使用的轴承为 6008 单列深沟球轴承,开展全寿命实验的试验条件如表 1 所示。分别开展了两组全寿命实验,实验条件相同,两组轴承全寿命过程的时间如表 2 所示,从表中可以看出轴承 A 和轴承 B 分别运行了 765 min 和 1 018 min,由于振动达到失效阈值实验机停机。图 7 中(a)和(b)分别是采集的全寿命周期的振动加速度信号,从图中可以看出起始阶段振动加速度比较平稳,进入退化阶段后振动加速度幅值急剧增强,并且在较短时间内达到失效阈值。

表 1 试验条件	
Table 1 Experimental conditions	
实验条件	详细参数
轴承型号	6 008
实验转速	7 000 r/min
轴承实验数量	4 套
采样频率	25 600 Hz
数据保存间隔	1 min
每次保存数据长度	25 600
载荷	5 kN
润滑与冷却方式	脂润滑/32#机械油润滑

表 2 轴承 A 和轴承 B 的全寿命试验过程			
Table 2 The full-life tesing of bearing A and bearing B			
数据集	轴承	运行时间维度	运行总时长 /min
训练集	A	01. 10 23:44—01. 11 12:28	765
测试集	B	01. 27 12:43—01. 28 05:40	1 018

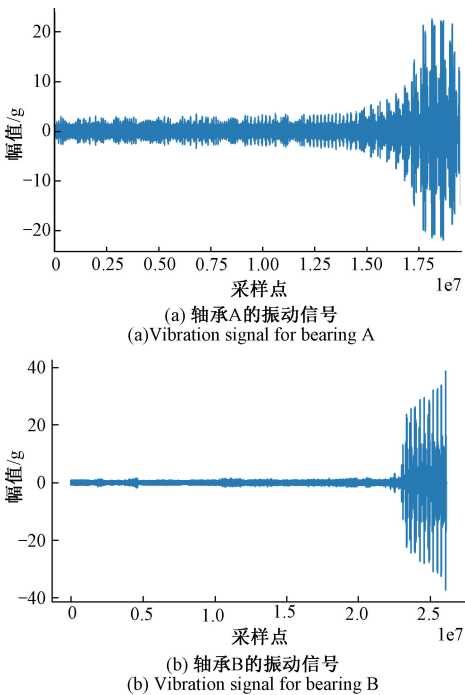


图 7 轴承 A 和轴承 B 全寿命周期的振动加速度信号

Fig. 7 The vibration acceleration signals collected under the full-life cycle of bearing A and bearing B

3.2 评价指标

为了评估预测模型的性能,本文选取 RMSE (root mean square error)、MAE (mean absolute error)、MSE (mean squared error) 和 R^2 四项评价指标。

均方根误差 RMSE 可以监测预测异常值,反映模型的稳定性,其公式如下所示:

$$RMSE = \sqrt{\frac{\sum_{i=1}^N |y_i - \hat{y}_i|^2}{n}} \tag{21}$$

式中: N 是样本数量, y_i 是实际目标值, $\hat{y}_i$ 是预测值。平均绝对误差 MAE 计算模型绝对误差的均值,其公式如下所示:

$$MAE = \frac{\sum_{i=1}^N |y_i - \hat{y}_i|}{n} \tag{22}$$

均方误差是用于衡量预测值与实际值之间差异的一种常用的回归性能评估指标。其计算公式为:

$$MSE = \frac{1}{n} \sum_{i=1}^N (y_i - \hat{y}_i)^2 \tag{23}$$

决定系数 R^2 衡量了模型对数据的拟合程度,其取值范围为 0 到 1。当 R^2 为 1 时,表示模型对数据的拟合程度最好;当 R^2 为 0 时,表示模型对数据的拟合程度最差,其公式如下所示:

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y}_i)^2}$$

(24)

式中： $\bar{y}_i$ 是目标值的平均值。

4 试验验证

4.1 退化起始点划分

取轴承全寿命周期振动数据的前 20% 为正常数据集,首先计算正常数据集的协方差矩阵及其逆矩阵,之后由式(1)计算全寿命数据中各个数据与正常数据集的马氏距离。再根据式(4)可以计算退化起点的域值,这里设定阈值系数 k 为 3。计算得到轴承 A 的退化阈值 $Thr_A = 4.63$, 对应的退化起始点为 542 min;轴承 B 的退化阈值 $Thr_B = 3.98$, 对应的退化起始点为 869 min;当马氏距离超过阈值后,轴承则进入退化阶段。轴承 A 和轴承 B 的马氏距离的曲线如图 8 所示,从图中可以看出随着轴承运行时间增加,距离起始的正常数据集的马氏距离逐渐增大。

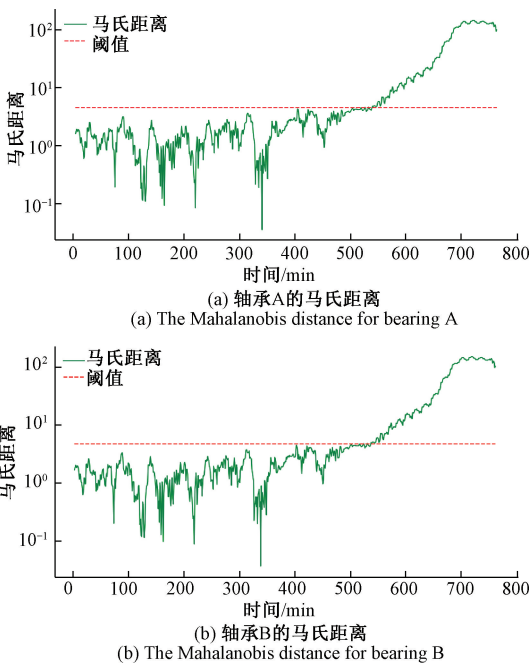


图 8 退化起始点的确定结果

Fig. 8 Determination results of degradation start point

4.2 基于自编码器的特征提取

由于多传感器融合需要考虑到信息的相关性,由振动传感器采集的数据为每隔 1 min 采集 1 s 的数据,每次采集 25 600 个数据点,而轴承全寿命试验台记录的电机电流、轴承温度信号为每分钟采样一个数据点,因此将同

1 min 的振动信号合并提取每分钟内轴承的运行情况特征,每组振动数据中提取 2 000 个数据点足以包含振动信号的绝大多数信息。

堆叠自编码器的深度由堆叠的层数决定,若网络过浅则无法提取具有代表性的特征,而过深则会增加参数量并导致训练速度下降及过拟合等问题。因此,选择合适的网络深度非常关键。本文将编码器层数设置为 5 层,经过 1 层隐层后由对称的解码器输出。采用轴承 A 为训练集,轴承 B 的数据为测试集,训练次数为 30 次,堆叠自编码器神经元划分如表 3 所示。

表 3 堆叠自编码器的参数设置
Table 3 Parameter settings for stacked autoencoder

层	神经元数量
编码器(5 层)	512-256-128-64-32
隐层(1 层)	10
解码器(5 层)	32-64-128-256-512

使用堆叠自编码器对振动信号的特征提取结果如图 9 所示,去除零特征后保留 6 维的特征,如图 9 所示的特征 0~5,可以看出所提取的特征可以反映图 7(a)和(b)原始振动信号的退化特性。

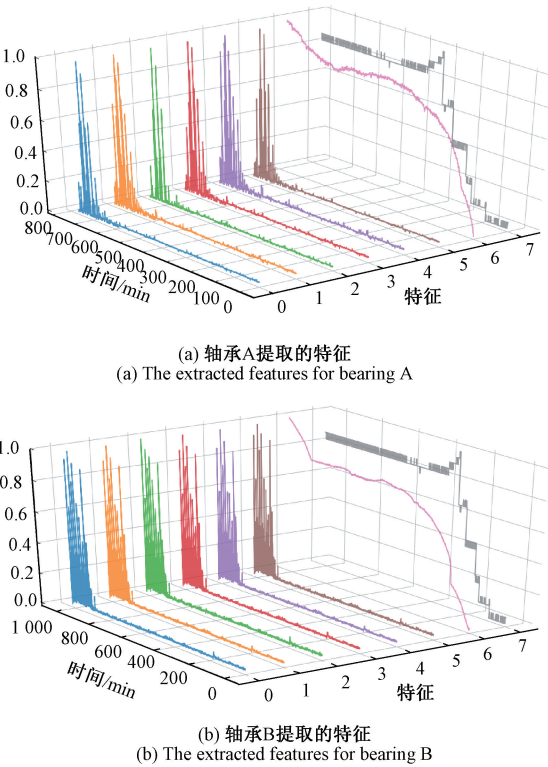


图 9 多传感器融合的特征

Fig. 9 Features from multi-sensor fusion

4.3 多传感器融合

为了完整的反应电机轴承系统运行状态,本文采用多传感器融合特征。将轴承振动信号进行特征提取与该时间段轴承温度、电机电流相融合得到多源信息特征如图9所示,图9(a)和(b)分别表示轴承A和轴承B的融合特征,其中特征0~5为振动信号经过自编码器提取的特征,特征6为轴承温度信号,特征7为电机电流信号,从图中可以看出轴承温度在进入退化阶段后上升速度加快。因此轴承温度和电流信号能在一定程度上反映轴承的性能退化过程。

4.4 基于MHA-LSTM的滚动轴承剩余寿命预测

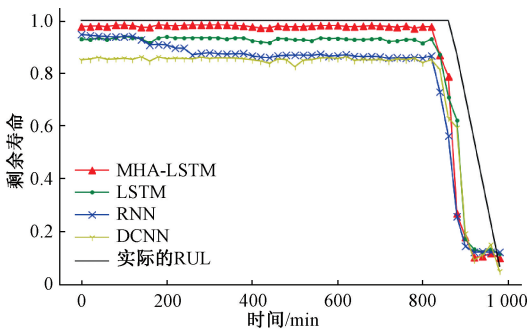
根据图5所示的MHA-LSTM模型对滚动轴承的剩余寿命进行预测,模型参数设置如表4所示。分别使用LSTM、RNN、DCNN^[25](Deep CNN, DCNN)与MHA-LSTM模型进行训练和预测。图10(a)所示为基于振动传感器的寿命预测结果,对比4种预测算法,结果表明MHA-LSTM的预测曲线最接近实际剩余寿命,比LSTM的预测效果更好,表明MHA-LSTM预测模型由于引入注意力机制加强重要特征,动态选择与剩余寿命相关度较高的特征,使得预测准确度有了一定的提升。图10(b)为基于温度信号的寿命预测,采用LSTM预测模型,从图中可以看出在起始和最后阶段温度的预测效果较好,但是中间过程预测的波动性较大,预测误差RMSE为0.174,MAE为0.133,MSE为0.0302,R²为0.209。图10(c)是基于多传感器融合的寿命预测结果,可以看出相对于基于单个传感器预测,多传感器融合的寿命预测可以提高各个模型剩余寿命预测准确度,且比基于温度预测效果更好。四种模型的预测误差如表5和6所示,表明基于多传感器融合数据较单一振动信息特征数据预测精度整体提高,这是因为轴承温度、电流电流特征与轴承剩余寿命有较高相关度,利用多传感器融合可以完整地反应轴承的整体运行状态,更精确地提供轴承退化状态信息,较单一振动传感器提高了模型的鲁棒性,因此具有更高的预测精度。

表4 MHA-LSTM模型参数

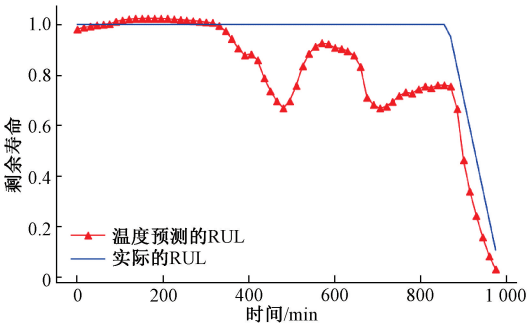
Table 4 Parameters of MHA-LSTM model				
预测模型	参数			
MHA-LSTM	注意力头数=4,输入序列长度=30,输出序列长度=1,迭代轮次=30,批大小=32			

表5 基于振动传感器的剩余寿命预测误差

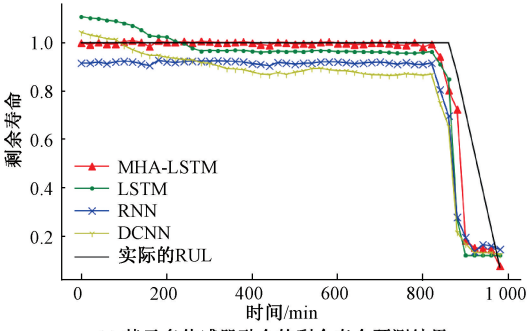
Table 5 Prediction error of RUL based on vibration sensor				
模型	RMSE	MAE	MSE	R ²
MHA-LSTM	0.144	0.066	0.021	0.447
LSTM	0.159	0.109	0.025	0.314
RNN	0.161	0.128	0.033	0.296
DCNN ^[25]	0.183	0.142	0.041	0.104



(a) 基于振动传感器的剩余寿命预测结果
(a) RUL prediction results based on vibration sensors



(b) 基于温度信号的剩余寿命预测结果
(b) RUL prediction results based on temperature signals



(c) 基于多传感器融合的剩余寿命预测结果
(c) RUL prediction results based on multi-sensor fusion

图10 剩余寿命预测结果对比

Fig. 10 Comparison of RUL prediction results

表6 基于多传感器融合的剩余寿命预测误差

Table 6 Prediction error of RUL based on multi-sensor fusion				
模型	RMSE	MAE	MSE	R ²
MHA-LSTM	0.135	0.049	0.018	0.509
LSTM	0.148	0.081	0.022	0.417
RNN	0.157	0.117	0.024	0.340
DCNN ^[25]	0.178	0.132	0.032	0.154

4.5 模型参数分析

由于MHA-LSTM预测模型中历史输入步长会影响网络的学习深度,步长过小模型难以学习到深层特征,过

大则会出现模型过拟合风险,也会影响模型的预测精度。为了确定合适的输入步长,计算输入长度分别为 10、20、30、40、50 情况下模型预测结果的 RMSE、MAE、MSE、 R^2 ,结果如表 7 所示。结果表明,当输入步长为 30 时,MHA-LSTM 预测误差均达到最小。

表 7 步长对 MHA-LSTM 预测精度影响

Table 7 Effect of step size on the prediction accuracy of MHA-LSTM

输入步长	RMSE	MAE	MSE	R^2
10	0.159	0.118	0.025	0.323
20	0.154	0.084	0.023	0.361
30	0.135	0.049	0.018	0.509
40	0.141	0.072	0.020	0.465
50	0.145	0.112	0.021	0.442

5 结 论

本文提出一种基于多传感器融合与 MHA-LSTM 的电机轴承剩余寿命预测方法。首先利用马氏距离确定轴承初始退化点,并针对传统的手动特征提取方法具有较高的人为因素影响,本文采用堆叠自编码器进行自动特征提取,可以挖掘振动信号中深层次特征。此外,针对传统单一传感器信号鲁棒性较差问题,采用多传感器融合技术,提取更加全面的反应系统运行状态的信息,利用多头注意力机制动态选择特征。最后,构建了 MHA-LSTM 预测模型实现对电机轴承的剩余寿命预测。具体结论如下:

- (1) 利用马氏距离表征数据点与正常数据间的间隔距离结合统计学分布规律可以较好的判断轴承退化起始点,为剩余寿命预测的起始点识别提供依据。
- (2) 针对滚动轴承恶劣环境下信息缺失等问题,将轴承温度、电机电流与振动信号特征融合,采用堆叠自编码器深度挖掘数据中的重要特征,以更全面、准确地反映系统整体运行状态,获取更加全面的反应系统性能退化信息。
- (3) 本文提出的 MHA-LSTM 剩余寿命预测模型是利用注意力机制自适应选择重要特征,强化与剩余寿命相关度较高特征的权重,在预测模型将注意力层与 LSTM 层拼接,以防止信息传递丢失。结果表明本文提出的剩余寿命预测模型预测误差均小于 LSTM、RNN、DCNN 三种模型,体现了所提出模型的优越性。另外,多传感器融合的预测模型与传统单一振动传感器预测模型相比,所有模型的预测准确度均有不同程度的提高,这也表明了多传感器融合在剩余寿命预测领域较单一传感器的优越性。

参考文献

[1] 雷亚国,贾峰,周昕,等. 基于深度学习理论的机械装备大数据健康监测方法[J]. 机械工程学报, 2015, 51(21): 49-56.
LEI Y G, JIA F, ZHOU X, et al. A deep learning-based method for machinery health monitoring with big data[J]. Journal of Mechanical Engineering, 2015, 51(21): 49-56.

[2] 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[J]. 机械工程学报, 2019, 55(8): 1-13.
PEI H, HU CH H, SI X SH, et al. Review of machine learning based remaining useful life prediction methods for equipment[J]. Journal of Mechanical Engineering, 2019, 55(8): 1-13.

[3] OMSHI E M, GRALL A, SHEMEHSAVAR S. A dynamic auto-adaptive predictive maintenance policy for degradation with unknown parameters [J]. European Journal of Operational Research, 2020, 282(1): 81-92.

[4] LEI Y G, LI N P, GUO L, et al. Machinery health prognostics: A systematic review from data acquisition to RUL prediction [J]. Mechanical Systems and Signal Processing, 2018, 104: 799-836.

[5] SI X SH, WANG W B, HU CH H, et al. Remaining useful life estimation-A review on the statistical data driven approaches[J]. European Journal of Operational Research, 2011, 213: 1-14.

[6] DAIGLE M J, GOEBEL K. Model-based prognostics with concurrent damage progression processes [J]. IEEE Transactions on Systems, Man, and Cybernetics: Systems, 2013, 43(3): 535-546.

[7] SHARIFF A A. A stochastic Paris-Erdogan model for fatigue crack growth using two-state model [J]. The Bulletin of the Malaysian Mathematical Society Series, Second Series, 2008, 31(1): 97-108.

[8] QIAN Y N, YAN R Q, GAO R X. A multi-time scale approach to remaining useful life prediction in rolling bearing[J]. Mechanical Systems and Signal Processing, 2017, 83: 549-567.

[9] LIAO L X. Discovering prognostic features using genetic programming in remaining useful life prediction [J]. IEEE Transactions on Industrial Electronics, 2013, 61(5): 2464-2472.

[10] GUPTA P K, ZARETSKY E V. New stress-based fatigue life models for ball and roller bearings [J]. Tribology Transactions, 2018, 61(2): 304-324.

[11] 陈保家,郭凯敏,陈法法,等. 基于残差 NLSTM 网络和注意力机制的航空发动机剩余使用寿命预测[J].

- 航空动力学报, 2023, 38(5): 1176-1184.
- CHEN B J, GUO K M, CHEN F F, et al. Prediction of remaining useful life of aero-engine based on residual NLSTM neural network and attention mechanism [J]. Journal of Aerospace Power, 2023, 38(5): 1176-1184.
- [12] DING Y F, DING P, JIA M P. A novel remaining useful life prediction method of rolling bearings based on deep transfer auto-encoder [J]. IEEE Transactions on Instrumentation and Measurement, 2021, 70: 1-12.
- [13] YANG X Y, ZHENG Y, ZHANG Y, et al. Bearing remaining useful life prediction based on regression shapalet and graph neural network [J]. IEEE Transactions on Instrumentation and Measurement, 2022, 71: 1-12.
- [14] WEI Y P, WU D ZH, TERPENNY J. Bearing remaining useful life prediction using self-adaptive graph convolutional networks with self-attention mechanism[J]. Mechanical Systems and Signal Processing, 2023, 188: 110010.
- [15] SHE D M, JIA M P. A BiGRU method for remaining useful life prediction of machinery [J]. Measurement, 2021, 167: 108277.
- [16] REN L, LIU Y X, HUANG D, et al. MCTAN: A novel multichannel temporal attention-based network for industrial health indicator prediction [J]. IEEE Transactions on Neural Networks and Learning Systems, 2022, 34(9): 6456-6467.
- [17] SAXENA A, GOEBEL K, SIMON D, et al. Damage propagation modeling for aircraft engine run-to-failure simulation [C]. IEEE International Conference on Prognostics and Health Management, 2008: 1-9.
- [18] 李乃鹏, 蔡潇, 雷亚国, 等. 一种融合多传感器数据的数模联动机械剩余寿命预测方法[J]. 机械工程学报, 2021, 57(20): 29-37.
- LI N P, CAI X, LEI Y G, et al. A model-data-fusion remaining useful life prediction method with multi-sensor fusion for machinery [J]. Journal of Mechanical Engineering, 2021, 57(20): 29-37.
- [19] TA Y T, LI Y F, CAI W A, et al. Adaptive staged remaining useful life prediction method based on multi-sensor and multi-feature fusion [J]. Reliability Engineering & System Safety, 2023, 231: 109033.
- [20] WANG L, CAO H, XU H, et al. A gated graph convolutional network with multi-sensor signals for remaining useful life prediction [J]. Knowledge-Based Systems, 2022, 252: 109340.
- [21] LI T F, ZHAO ZH B, SUN C, et al. Hierarchical attention graph convolutional network to fuse multi-sensor signals for remaining useful life prediction [J]. Reliability Engineering & System Safety, 2021, 215: 107878.
- [22] LIU K B, GEBRAEEL N Z, SHI J J. A data-level fusion model for developing composite health indices for degradation modeling and prognostic analysis[J]. IEEE Transactions on Automation Science and Engineering, 2013, 10(3): 652-664.
- [23] IMANI M. Difference-based target detection using Mahalanobis distance and spectral angle [J]. International Journal of Remote Sensing, 2019, 40(3): 811-831.
- [24] FIRAT O, CHO K, SANKARAN B, et al. Multi-way, multilingual neural machine translation [J]. Computer Speech & Language, 2017, 45: 236-252.
- [25] WANG B, LEI Y G, LI N P, et al. Deep separable convolutional network for remaining useful life prediction of machinery [J]. Mechanical Systems and Signal Processing, 2019, 134: 106330.

作者简介



张苑(通信作者), 2010年于安徽信息工程学院获得学士学位, 2013年于安徽工程大学获得硕士学位, 2018年于东南大学获得博士学位, 现为南京信息工程大学自动化学院讲师, 主要研究方向为机电系统状态监测和寿命预测。

E-mail: zhangwan@nuist.edu.cn

Zhang Wan (Corresponding author) received her B. Sc. degree from Anhui Institute of Information Technology in 2010, followed by an M. Sc. degree from Anhui Polytechnic University in 2013, and ultimately her Ph. D. degree from Southeast University in 2018. She is currently a lecturer at the School of Automation, Nanjing University of Information Science and Technology, focusing on research in the areas of condition monitoring and life prediction of electromechanical systems.



贾民平, 分别在1982年、1985年和1991年于东南大学获得学士、硕士和博士学位, 现为东南大学机械工程学院教授, 主要研究方向为机械装备健康管理与智能运维。

E-mail: mpjia@seu.edu.cn

Jia Mingping received his B. Sc., M. Sc., and Ph. D. degrees from Southeast University in 1982, 1985, and 1991, respectively. He is currently a professor at the School of Mechanical Engineering, Southeast University. His research mainly focus on health management, and intelligent operation and maintenance of mechanical equipment.