

基于 LSTM-ES-RVM 的滚动轴承剩余寿命预测方法

周圣文¹, 郭顺生^{1,2}, 杜百岗^{1,2}

(1. 武汉理工大学机电工程学院, 湖北 武汉 430070; 2. 数字制造湖北省重点实验室, 湖北 武汉 430070)

摘要: 为解决滚动轴承在寿命预测时精度不高, 且性能退化趋势及波动范围难以预测等问题, 提出了基于 LSTM-ES-RVM 的滚动轴承剩余寿命预测方法。在无先验知识或人工经验的干扰下, 利用长短期记忆(Long Short-Term Memory, LSTM)网络直接对频率数据进行特征提取, 构建退化过程的初步健康指标(Health Indicator, HI); 为了消除 HI 曲线的局部剧烈振荡, 提出了带斜率的极端拐点(Extreme Inflection Point with a Slope, ES)模型改善其整体单调性; 使用相关向量机(Relevance Vector Machine, RVM)模型对 HI 曲线进行趋势预测, 实现了滚动轴承的剩余寿命(Remaining Useful Life, RUL)预测。实验结果表明, 所提方法相较于对比方法具有较好的预测精度。

关键词: 剩余寿命预测; 滚动轴承; 长短记忆神经网络; 健康指标; 带斜率的极端拐点模型

中图分类号: TP114.33; TH133.33 **文献标志码:** A **文章编号:** 1004-4523(2023)06-1723-13

DOI: 10.16385/j.cnki.issn.1004-4523.2023.06.027

1 概述

滚动轴承作为旋转机械设备中的关键机械基础件之一, 因其退化周期长且失效具有随机性, 其安全性备受关注。相关统计数据表明, 超过 30% 的旋转机械设备故障是由滚动轴承故障引起的^[1]。

相比故障诊断, 寿命预测可以提前预测设备未来的健康状况, 具有更多的意义。设备 RUL 预测主要包括获取状态信息、提取性能退化 HI 曲线和 HI 曲线趋势预测三大部分, 其中提取性能退化 HI 曲线是一个重要的研究问题。基于此, 许多学者提出了多种方法来提取 HI 曲线, 并取得了一定的研究成果。Lei 等^[2]使用均方根(Root Mean Square, RMS)构建 HI 曲线。Antoni^[3]使用峰度(kurtosis)时域指标构建 HI 曲线。Rai 等^[4]采用经验模式分解(Empirical Mode Decomposition, EMD)方法将每个信号分解成多个固有模态函数(Inherent Mode Functions, IMFs)成分, 然后选择前两个 IMFs 作为输入, 进行奇异值分解(Singular Value Decomposition, SVD), 生成 HI 曲线。Zhang 等^[5]利用短时傅里叶变换和非负矩阵分解模型, 从时频分布中提取时频码(Time-Frequency Codes, TFCs), 随后通过自组织特征映射(Self-Organizing Feature Map,

SOM)神经网络来量化特征向量之间的相似性并生成 HI 曲线。

然而, 上述方法存在以下问题: (1) 上述 HI 曲线构建模型需要先验知识和人工经验, 例如, 选择至少一个或多个融合的时频域指标(RMS 和 kurtosis 等)进行组合; (2) 这些 HI 构建模型需要通过几个不同的模型进行融合, 过程比较繁琐。

LSTM 模型为多层网络结构, 通过建立先前信息和当前环境之间的时间相关性, 能有效地降低对专家经验的依赖性。基于此, 众多学者采用 LSTM 网络来提取 HI 曲线。Cheng 等^[6]提出了一种基于 LSTM 的融合方法, 该方法包括时频域指标、核谱聚类 and LSTM 三大部分, 以识别设备的退化趋势。申彦斌等^[7]提出了一种基于双向长短时记忆(Bi-directional LSTM)网络的循环神经网络结构, 提取轴承的 HI 曲线。黄宇等^[8]提出一种结合双向长短时记忆网络与注意力机制的神经网络模型构建 HI 曲线。

HI 曲线的单调性是影响设备 RUL 预测结果的关键因素。虽然 LSTM 在一定程度上已经应用于提取 HI 曲线, 然而通过 LSTM 提取的 HI 曲线中存剧烈振荡。如图 1(b) 所示, 通过 LSTM 构建的 LSTM-HI 曲线在第 250 个数据点之前具有明显的振荡, 而从第 250 个点附近开始微弱增长。因此, 有必要找到一种合适的方法来处理图 1(b) 中的红色

虚线矩形中的剧烈振荡区域并改善其单调性。

指数加权移动平均(Exponentially Weighted Moving-Average, EWMA)、异常检测(Outlier Detection)和基于均值指数函数(Exponent Function based on the Mean)是三种常用来处理HI振荡的模型。其中,EWMA模型通过对时间序列数据的加权平均来控制预测过程中的均值漂移,可有效平滑时间序列数据。Tse等^[9]使用EWMA模型平滑提取油砂泵的HI曲线,并取得了良好的结果。异常值检测模型则利用数据的差值和标准差来检测异常数据点,以帮助识别异常数据并剔除其影响。谢雨洁等^[10]引入了基于多尺度局部核回归的异常值检测方法,提高了轴承退化阶段识别的准确性。基于均值指数函数则使用起始时间到当前时间期间数据点的平均值来减弱剧烈振荡,并利用指数函数本身的单调递增特性增强HI曲线的单调性。Tse等^[11]利用均值指数函数对HI曲线进行平滑处理,并获得了良好的效果。

为了消除LSTM-HI曲线的剧烈振荡区域并改善其单调性,本文分别利用EWMA、异常值检测和基于均值指数函数对LSTM-HI曲线的振荡区域进行平滑处理,如图1(c),(d),(f)所示。其中,基于均值指数函数提取的HI(LSTM-EHI)的计算公式如下:

$$\text{LSTM-EHI}(i) = e^{\text{mean}\{HI_1, HI_2, \dots, HI_i\}},$$

$$i = 1, 2, \dots, S \quad (1)$$

式中 HI_i 表示 i 时刻的 LSTM-HI 值; $\text{mean}\{HI_1, HI_2, \dots, HI_i\}$ 为从 HI_1 到 HI_i 的平均值。

然而,上述方法存在以下问题:

(1)相较于图1(b),图1(c)通过EWMA平滑方法所得到的曲线并没有明显改善HI曲线的平滑度和单调性。尤其是在红色虚线矩形区域内,曲线仍然出现了剧烈振荡。此外,曲线在红色虚线区域内的削弱,使得整个缓慢增长趋势变得不明显,并且在蓝色虚线椭圆区域内的整体单调性也被削弱,这将导致预测RUL的效果较差。此外,EWMA方法需要具备先验知识才能设置指数系数和滑动窗口的数量。

(2)图1(d)中仅检测到一个异常点,因此无法生成异常点区域。而在图1(e)中,异常点区域需要一定数量的异常点来确定。然而在实际工程中,并不是所有的数据集都总是能够生成异常区域。此外,异常点检测只能提高局部异常点区域的单调性,而不能提高HI曲线整体的平滑性和单调性。

(3)图1(f)为通过LSTM-EHI方法提取的HI曲线,与图1(c)(EWMA)相比,LSTM-EHI曲线更

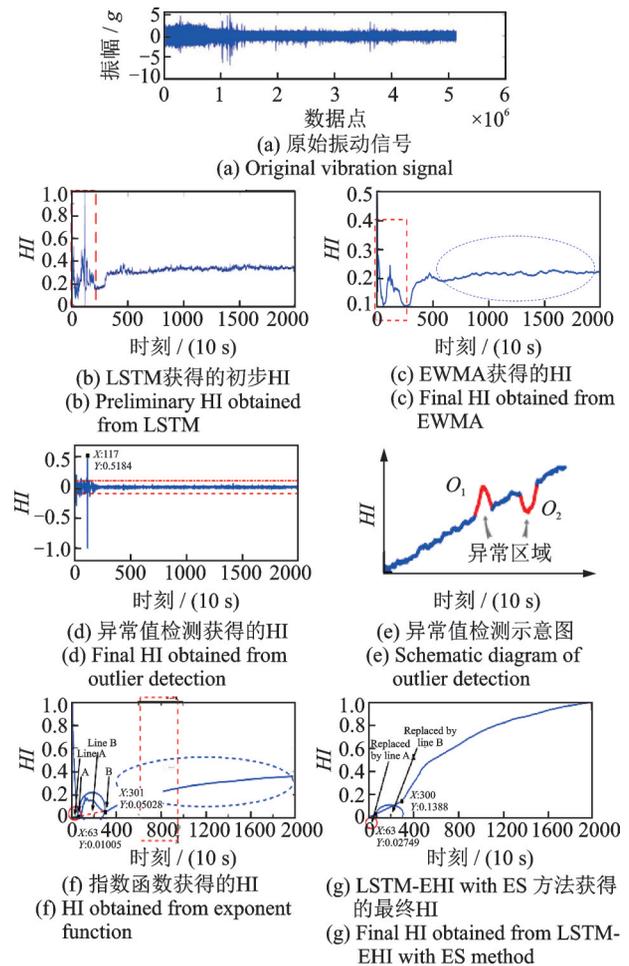


图1 使用不同方法处理LSTM-HI曲线剧烈振荡区域并改善单调性对比图

Fig. 1 Different methods are used to deal with the violent oscillation region of LSTM-HI curve and to improve the monotony contrast graph

具平滑性和单调性。然而,红色虚线区域中的剧烈振荡曲线降低了整体的平滑性和单调性。出现这种情况的原因在于,轴承或设备在系统运行的开始阶段未达到稳定状态,从而导致振动信号的略微波动,进而导致红色虚线区域的HI值不稳定。因此,需要对这部分曲线进行处理和替换。

此外,近年来随着人工智能技术的发展,相关学者已经提出了多种趋势预测方法,主要包括误差反向传输(Back Propagation, BP)^[12]神经网络、粒子群优化算法(Particle Swarm Optimization, PSO)^[13]以及支持向量机(Support Vector Machine, SVM)^[14]等。其中,SVM利用支持向量样本集决定预测结果,具有一定的鲁棒性,但存在计算复杂、耗时长、惩罚系数难确定、核函数受Mercer条件限制等问题。RVM^[15]是一种将Bayesian理论与SVM相结合的机器学习算法,相较于SVM,RVM具有良好的稀疏性和泛化能力,已在趋势预测领域得到一定的应用。

基于此,本文提出了基于LSTM-ES-RVM的滚动轴承RUL预测方法。首先,为了减少对先验知识和人工经验的依赖,提出了LSTM模型,直接从频域上提取轴承的初步性能退化HI;其次,提出了ES模型,以消除退化HI曲线振荡并增强其整体单调性;最后,将提取的性能退化健康指标按时间维度拆分为训练集和测试集,利用RVM模型对HI曲线进行趋势预测,实现滚动轴承的RUL预测。

2 基于LSTM-ES-RVM的滚动轴承RUL预测框架

本文提出的基于LSTM-ES-RVM的滚动轴承RUL预测框架主要分为五个部分:数据采集、数据预处理、性能退化HI曲线构建、健康阶段划分、RUL预测,如图2所示。

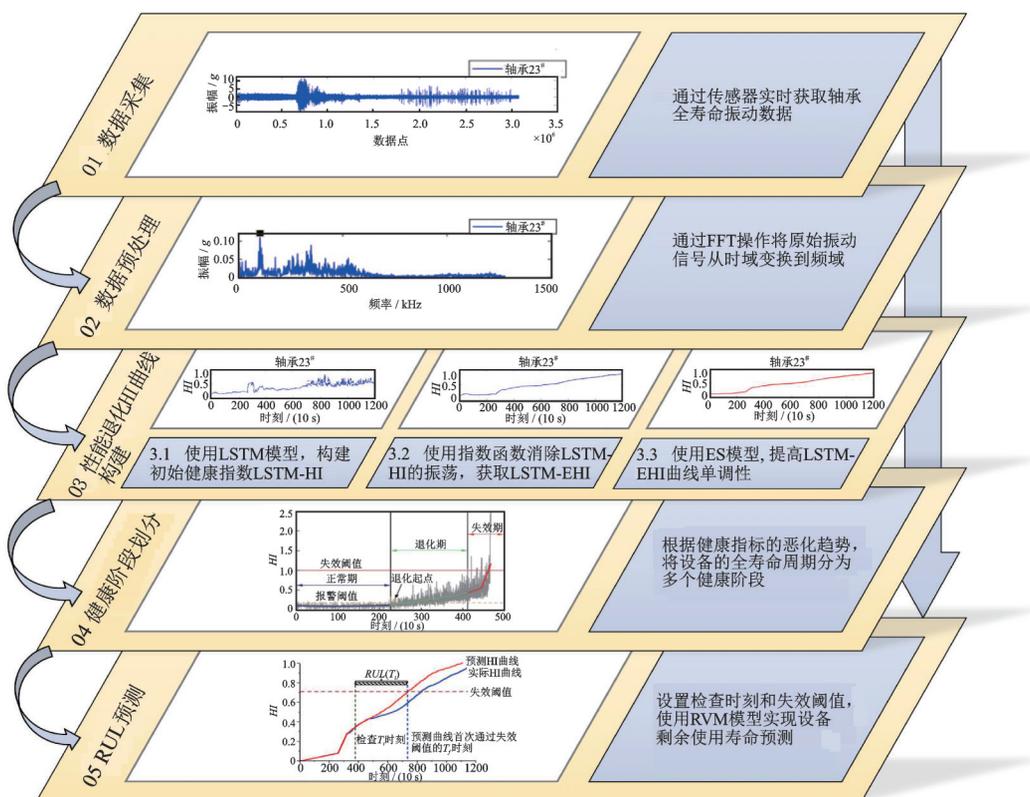


图2 基于LSTM-ES-RVM的RUL预测框架图

Fig. 2 RUL prediction framework based on LSTM-ES-RVM

(1)数据采集:通过振动传感器实时采集滚动轴承的全寿命振动数据;

(2)数据预处理:通过快速傅里叶变换(Fast Fourier Transform,FFT)操作将原始振动信号从时域变换到频域;

(3)性能退化HI曲线构建:先使用LSTM模型提取初始HI曲线,再通过ES模型利用曲线的极小值数据点来替代HI曲线局部剧烈振荡区域中的极值点,从而提高HI曲线的单调性;

(4)健康阶段划分:在进行剩余寿命预测之前,可以根据设备退化HI的变化趋势将设备的全生命周期划分为健康阶段、退化阶段和失效阶段;

(5)RUL预测:采用RVM模型对退化HI曲线进行趋势预测,通过外推预测退化HI曲线超过失效阈值的时间,间接获得设备剩余寿命的预测值。

2.1 基于LSTM-ES的健康指标构建

为了获得图1(g)中的曲线,本文提出了ES模型,以消除图1(f)的振荡区域并增强曲线的整体单调性。ES模型的主要思想是在健康指标曲线中检测所有局部最小值数据点,并根据这些数据点的斜率按降序进行排序,通过将这些局部最小值点连接起来消除振荡区域,从而增强HI曲线的整体单调性。算法流程如图3所示,其中LSTM-EHI作为ES模型的输入,经过ES方法处理后的HI曲线称为LSTM-EHI with ES曲线。

具体的计算步骤如下所示:

步骤1:查找所有局部最小点(极小值点),并将其保存到相应极小值点数据序列MP。

步骤2:计算MP中两个相邻极小值点的斜率,

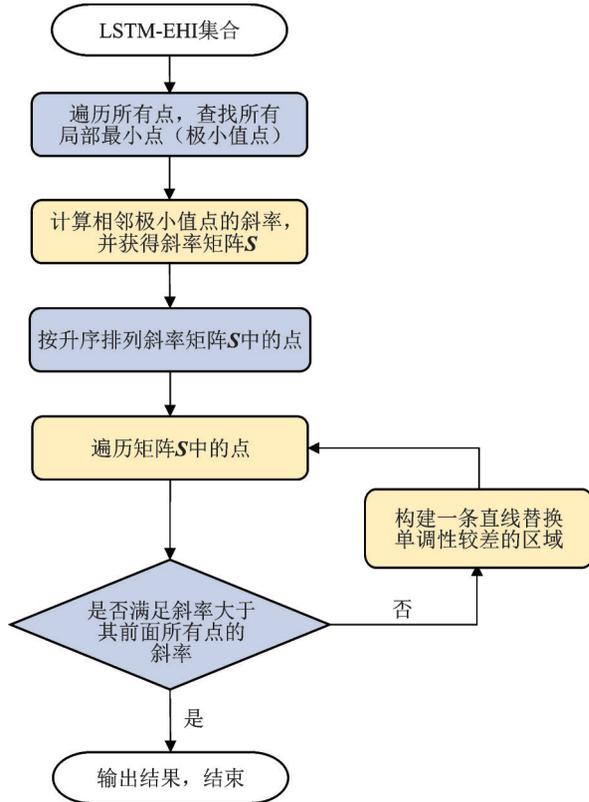


图3 ES算法流程图

Fig. 3 Flow chart of ES algorithm

并得到斜率矩阵 S , 若斜率矩阵中的 S_j 为 0 (斜率为 0 代表横坐标轴), 则设置 $S_j=1$ 。

步骤 3: 将 S_j 按照升序排列获得斜率数据集 $sorta$ 和相应的顺序下标集合 $sortb$ 。

步骤 4: 使用临时变量 Temp 保存当前选择的最小极值点 $P(x_j, y_j)$, 该数据点是根据 $sorta$ 和 $sortb$ 获取的; x_j 表示相应数据点的下角标, y_j 为 LSTM-EHI 数值, y_1 表示集合 $sorta$ 中最小斜率的点。Temp 变量的第一个数据点为 $P(x_1, y_1)$, 第二个数据点为 $P(x_2, y_2)$, 当数据点 $P(x_j, y_j)$ 具有第二最小斜率时, 选择局部极值点 $P(x_{j-1}, y_{j-1})$ 和 $P(x_j, y_j)$, 并使用这两点构建的直线替换 HI 曲线单调性较差的区域。

步骤 5: 重复步骤 4 直到满足当前数据点 $P(x_j, y_j)$ 的斜率大于前面所有点的斜率时为止。

2.2 基于 RVM 的剩余寿命预测模型构建

旋转机械的 RUL 预测一直备受关注, 剩余寿命预测是指确定当前检查时刻至失效阈值的时间间隔^[16], 通常用一个时间段来表示从当前时刻到机器失效的剩余寿命。定义如下:

$$RUL(T_i) = T_f - T_i \quad (2)$$

式中 T_i 为当前检查时刻, 由用户指定; T_f 为预测退化 HI 曲线首次穿过失效阈值所对应的时刻;

$RUL(T_i)$ 为当前检查时刻 T_i 所预测的轴承剩余寿命。如图 4 所示。

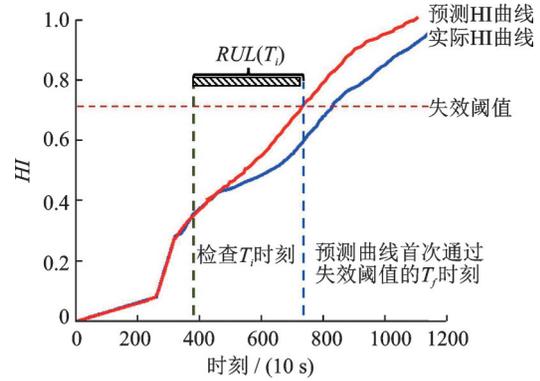


图4 剩余寿命预测示意图

Fig. 4 Schematic diagram of remaining useful life prediction

为了降低寿命预测模型的参数复杂度并提高训练速度, 使用 RVM 模型对 HI 曲线进行趋势预测, 从而实现了滚动轴承的 RUL 预测。RVM 基本原理如下:

假设存在训练样本集为 $\{x_i, t_i\}_{i=1}^S$, 其中, S 为样本总数, x_i 为输入向量, t_i 为对应输出值, 且输入与输出均为独立分布样本, 并认为 t_i 为关于输入 x_i 的非线性回归函数 $f(x_i)$ 的噪声输出, 即:

$$t_i = f(x_i) + \varepsilon_i \quad (3)$$

式中 ε_i 为零均值的高斯分布 $N(0, \sigma^2)$ 中各样本的高斯噪声误差, 其中 σ^2 为噪声误差的方差, 是未知量。因此, 容易得到 $p(t_i|x) = N(t_i|f(x_i), \sigma^2)$ 服从高斯正态分布, 其分布由 $t_i, f(x_i)$ 及方差 σ^2 所决定。

回归函数 $f(x_i)$ 的常见表达形式为:

$$f(x|\mathbf{w}) = \sum_{i=1}^S \mathbf{w}_i k(x, x_i) + \mathbf{w}_0 \quad (4)$$

式中 $k(x, x_i)$ 为核函数; $\mathbf{w} = [\mathbf{w}_1 \ \mathbf{w}_2 \ \mathbf{w}_3 \ \cdots \ \mathbf{w}_s]^T$ 为权重向量; \mathbf{w}_0 为偏差。

由于高斯核 (Gaussian kernel) 函数具有局部性好、参数数量少、计算量小等优势, 本文采用高斯核函数 $k(x, x_i) = \exp\{-\|x - x_i\|^2 / (2\tau^2)\}$, 其中 τ 为核宽度参数。

则最大似然概率可以定义为:

$$p(\mathbf{t}|\mathbf{w}, \sigma^2) = (2\pi\sigma^2)^{-\frac{S}{2}} \exp\{-\|\mathbf{t} - \boldsymbol{\phi}\mathbf{w}\|^2 / (2\sigma^2)\} \quad (5)$$

式中 $\mathbf{t} = [t_1 \ t_2 \ t_3 \ \cdots \ t_s]^T$ 为训练样本的输出向量; $\boldsymbol{\phi} = [\varphi(x_1) \ \cdots \ \varphi(x_s)]^T$ 为 $S \times (S+1)$ 维的核函数矩阵, 其中 $\varphi(x_i) = [1 \ k(x_i, x_1) \ \cdots \ k(x_i, x_s)]^T$ 。

如果直接对上述似然函数关于 \mathbf{w} 和 σ^2 求极大似然, 会导致“过拟合”的问题。Tipping^[15]从贝叶斯定理的角度出发, 认为权重向量 \mathbf{w} 服从一定的先验

分布,即:

$$p(\boldsymbol{w}|\boldsymbol{\alpha}) = \prod_{i=1}^S N(\boldsymbol{w}_i|0, \boldsymbol{\alpha}_i^{-1}) \quad (6)$$

式中 $\boldsymbol{\alpha}_i$ 为权重 \boldsymbol{w}_i 对应的超参数。

定义先验分布后,根据贝叶斯定理,未知参数的后验分布为:

$$p(\boldsymbol{w}, \boldsymbol{\alpha}, \sigma^2|t) = \frac{p(t|\boldsymbol{w}, \boldsymbol{\alpha}, \sigma^2) p(\boldsymbol{w}, \boldsymbol{\alpha}, \sigma^2)}{p(t)} = \frac{p(\boldsymbol{w}|t, \boldsymbol{\alpha}, \sigma^2) p(\boldsymbol{\alpha}, \sigma^2|t)}{p(t)} \quad (7)$$

根据贝叶斯定理,得到权重向量 \boldsymbol{w} 的后验分布为:

$$p(\boldsymbol{w}|t, \boldsymbol{\alpha}, \sigma^2) = \frac{p(t|\boldsymbol{w}, \sigma^2) p(\boldsymbol{w}|\boldsymbol{\alpha})}{p(t|\boldsymbol{\alpha}, \sigma^2)} = (2\pi\sigma^2)^{-\frac{s+1}{2}} |Q|^{-\frac{1}{2}} \exp\left[-\frac{1}{2}(\boldsymbol{w} - \boldsymbol{\mu})^T Q^{-1}(\boldsymbol{w} - \boldsymbol{\mu})\right] \quad (8)$$

其中,后验分布的协方差 Q 和均值 $\boldsymbol{\mu}$ 分别为:

$$Q = (\sigma^{-2} \boldsymbol{\Phi}^T \boldsymbol{\Phi} + A)^{-1} \quad (9)$$

$$\boldsymbol{\mu} = \sigma^{-2} Q \boldsymbol{\Phi}^T t \quad (10)$$

且 $A = \text{diag}(\boldsymbol{\alpha}) = \text{diag}(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \dots, \boldsymbol{\alpha}_s)$ 。

此时,RVM的训练学习过程转化为如何使超参数后验分布 $p(\boldsymbol{\alpha}, \sigma^2|t) \propto p(t|\boldsymbol{\alpha}, \sigma^2) p(\boldsymbol{\alpha}) p(\sigma^2)$ 最大化的问题。这一过程只需最大化边际分布 $p(t|\boldsymbol{\alpha}, \sigma^2)$,通过对参数 \boldsymbol{w} 进行边缘积分求得:

$$p(t|\boldsymbol{\alpha}, \sigma^2) = \int p(t|\boldsymbol{w}, \sigma^2) p(\boldsymbol{w}|\boldsymbol{\alpha}) d\boldsymbol{w} = (2\pi\sigma^2)^{-\frac{s}{2}} |\sigma^2 \mathbf{I} + \boldsymbol{\Phi} A^{-1} \boldsymbol{\Phi}^T|^{-\frac{1}{2}} \exp\left[-\frac{1}{2} t^T (\sigma^2 \mathbf{I} + \boldsymbol{\Phi} A^{-1} \boldsymbol{\Phi}^T)^{-1} t\right] \quad (11)$$

从而得到超参数 $\boldsymbol{\alpha}, \sigma^2$ 的边缘似然:

$$p(t|\boldsymbol{\alpha}, \sigma^2) = N(0, C) \quad (12)$$

式中 $C = \sigma^2 \mathbf{I} + \boldsymbol{\Phi} A^{-1} \boldsymbol{\Phi}^T$ 。

当超参数通过序贯稀疏贝叶斯学习算法优化获得了最优解超参数 $(\bar{\boldsymbol{\alpha}}, \bar{\sigma}^2)$ 之后,如果给定一个新的输入向量 $\boldsymbol{x}_{\text{new}}$,新的预测概率分布为:

$$p(t_{\text{new}}|\boldsymbol{x}_{\text{new}}, \bar{\boldsymbol{\alpha}}, \bar{\sigma}^2) = \int p(t_{\text{new}}|\boldsymbol{w}, \bar{\sigma}^2) p(\boldsymbol{w}|t_{\text{new}}, \bar{\boldsymbol{\alpha}}, \bar{\sigma}^2) d\boldsymbol{w} \sim N(\mu_{\text{new}}, \sigma_{\text{new}}^2) \quad (13)$$

由于式(13)服从正态分布,因此有:

$$p(t_{\text{new}}|\boldsymbol{x}_{\text{new}}, \bar{\boldsymbol{\alpha}}, \bar{\sigma}^2) = N(\mu_{\text{new}}, \sigma_{\text{new}}^2) \quad (14)$$

其中,预测值的平均值 μ_{new} 和方差 σ_{new}^2 分别为:

$$\mu_{\text{new}} = \boldsymbol{\mu}^T \boldsymbol{\Phi}(\boldsymbol{x}_{\text{new}}) \quad (15)$$

$$\sigma_{\text{new}}^2 = \bar{\sigma}^2 + \boldsymbol{\Phi}^T(\boldsymbol{x}_{\text{new}}) Q \boldsymbol{\Phi}(\boldsymbol{x}_{\text{new}}) \quad (16)$$

基于上述理论基础,本文提出了一种基于RVM的剩余寿命预测模型。首先,将提取的性能退化健康指标按时间维度拆分为训练集和测试集,

作为RVM模型的输入;然后,拟合相关向量中未知的参数,从而获得相关向量;接着,设置预测过程中所需的检查时刻和失效阈值等参数;最后,通过外推拟合的退化HI曲线超过失效阈值的时间,间接获取设备剩余寿命的预测值。

3 实验结果与对比分析

3.1 评估指标定义

3.1.1 单调性指标

使用单调性指标 (Monotonicity, Mon) 来评估构建的HI曲线的单调性^[17], Mon的计算公式如下:

$$\text{Mon} = \left| \frac{\text{Number}(dF > 0) - \text{Number}(dF < 0)}{T - 1} \right| \quad (17)$$

式中 $dF = \frac{(HI)_{t+1} - (HI)_t}{\Delta t}$ ($t = 1, 2, 3, \dots, T$) 为任意两个相邻点之间的差值; $\text{Number}(dF > 0)$ 表示差值大于零的健康因子数量; $\text{Number}(dF < 0)$ 表示差值小于零的健康因子数量。

如果 $\text{Mon} = 0$ 表示HI曲线平滑不具有单调性,即 $\text{Number}(dF > 0)$ 等于 $\text{Number}(dF < 0)$; $\text{Mon} = 1$ 表示HI曲线持续平稳的单调上升或下降,即 $\text{Number}(dF > 0)$ 或 $\text{Number}(dF < 0)$ 为0。因此, Mon值越接近1, HI曲线的单调性越好。

3.1.2 预测精度指标

为了定量评价预测模型的有效性和优越性,本文使用RE^[18] (Relative Error, 相对误差)、RMSE^[18] (Root Mean Square Error, 均方根误差) 作为评价指标对轴承RUL预测模型进行评估,其评估公式分别为:

$$\text{RE} = \frac{|y_i - \tilde{y}_i|}{y_i} \times 100\% \quad (18)$$

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

式中 n 为数据总数; y_i 表示第 i 时刻的真实值; \tilde{y}_i 表示第 i 时刻的预测值。

3.2 实验数据采集平台及数据预处理

本文采用IEEE PHM 2012挑战数据集对所提出的方法进行有效性验证。该数据集来自PRO-NOSTIA轴承退化试验平台(如图5所示),共采集了17个滚动轴承的数据,分为三个工况组,其中工况1组和工况2组各有7个轴承,工况3组有3个轴承。实验平台每10 s采集一个样本,每个样本长度

为2560个数据点,采样频率为25.6 kHz。表1提供了有关数据集的详细信息^[19]。

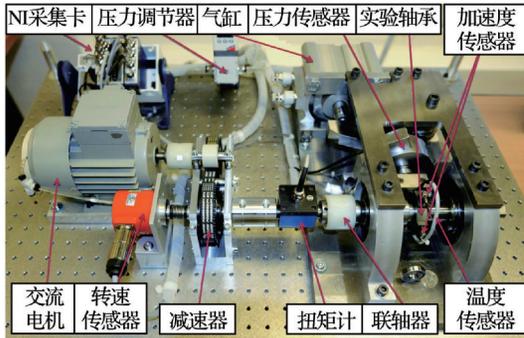


图5 实验数据采集平台

Fig. 5 The experimental data acquisition platform

表1 不同工况下滚动轴承原始振动信号信息

Tab. 1 The roller bearing's raw vibration signal information under different working conditions

工况组别	轴承编号	转速/ ($r \cdot \text{min}^{-1}$)	负载/ N	样本 总数	长度/ 个
工况1组	轴承11 [#]	1800	4000	2803	2560
	轴承12 [#]	1800	4000	871	2560
	轴承13 [#]	1800	4000	180	2560
	轴承14 [#]	1800	4000	1139	2560
	轴承15 [#]	1800	4000	2302	2560
	轴承16 [#]	1800	4000	2302	2560
	轴承17 [#]	1800	4000	1502	2560
工况2组	轴承21 [#]	1650	4200	911	2560
	轴承22 [#]	1650	4200	797	2560
	轴承23 [#]	1650	4200	1202	2560
	轴承24 [#]	1650	4200	612	2560
	轴承25 [#]	1650	4200	2002	2560
	轴承26 [#]	1650	4200	572	2560
	轴承27 [#]	1650	4200	172	2560
工况3组	轴承31 [#]	1500	5000	515	2560
	轴承32 [#]	1500	5000	1637	2560
	轴承33 [#]	1500	5000	352	2560

本文选取了轴承13[#]~15[#]、轴承23[#]~25[#]和轴承33[#]作为观察数据集。由于滚动轴承的振动信号具有非线性和随机的特性,从时域数据中很难提取到有用的信息,例如图6(c)中轴承15[#]的信号在曲线上就没有明显的变化。

本文采用FFT方法将原始振动信号从时域变换到频域,以提取有用的信息。从图7中观察到,大多数轴承的特征频率主要分布在103 kHz左右,特别是轴承13[#]的特征频率为103 kHz,正好是系统工作频率(25.6 kHz)的4倍,说明特征频率主要分布在工作频率的倍频。

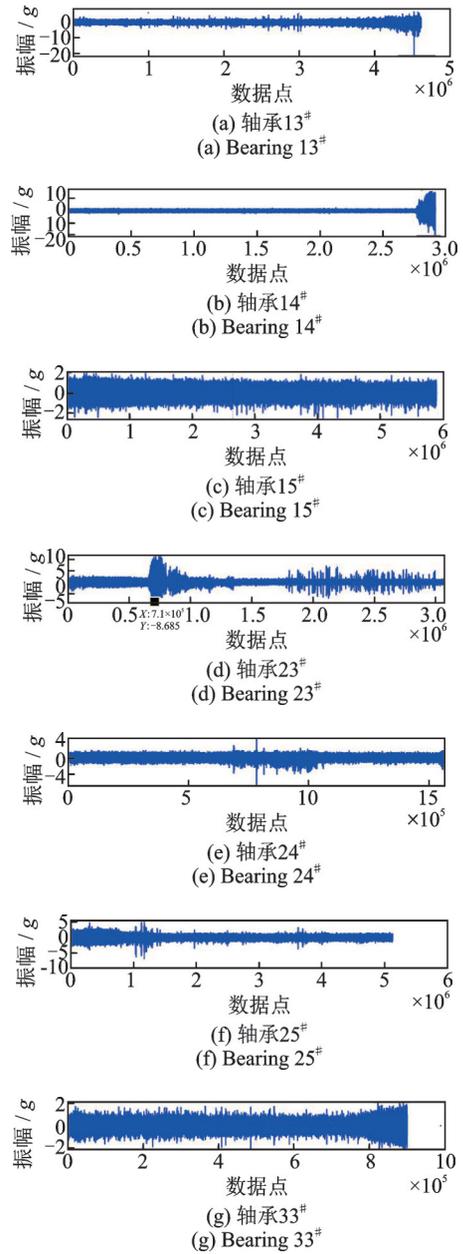


图6 不同轴承的原始时域信号

Fig. 6 The original time domain information of different bearings

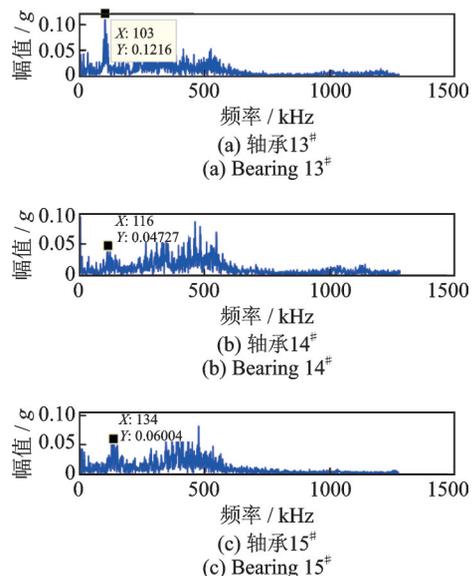


图7 不同轴承的原始频域信息

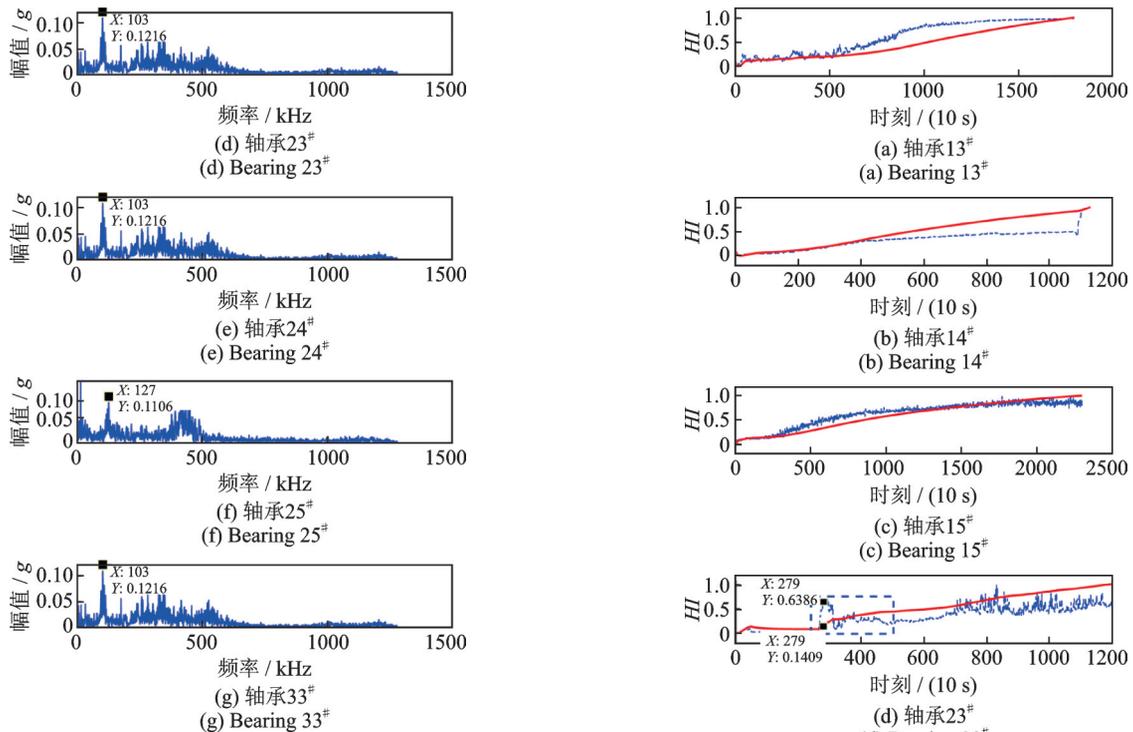


图 7 通过 FFT 方法转换的不同轴承的频域信号

Fig. 7 The frequency domain information of different bearings converted by FFT method

3.3 基于 LSTM-ES 的健康指标实验结果与分析

3.3.1 基于 LSTM-ES 的健康指标 (HI) 构建

在使用 LSTM 算法提取 LSTM-HI 曲线时, 设定初始学习率为 0.01, 最大迭代次数为 300。在训练过程中, 采用均方根误差 (RMSE) 作为损失函数, 自适应矩估计 (Adam) 作为优化器进行网络模型的参数优化更新。模型的输出值是第 i 个样本在 t 时刻的退化百分比, 例如, 若轴承 24# 的全生命周期样本总数为 5120000 个且输出大小被设置为 1, 则 3072000 个样本点的目标退化百分比为 0.6。

在利用 LSTM 算法提取轴承的 HI 曲线时, 本文采用同工况组中的其他轴承作为 LSTM 的训练数据集, 比如, 图 8 中轴承 13# 的 HI 曲线是以工况 1 组中的其他轴承作为训练数据集。

(a) 在图 8 (d) 中, 可以看到轴承 23# 的 LSTM-HI 曲线在第 279 点之后急剧增加; 与此同时, 在图 6 (d) 中, 轴承 23# 的原始振动信号曲线在 7.1×10^5 个点的振幅也明显增加。这是因为一个原始振动信号样本包含 2560 个数据点, 则图 8 (d) 中轴承 23# 的 279 点刚好对应图 6 (d) 中第 7.1424×10^5 (279×2560) 个数据点。这说明 LSTM 构建的 HI 曲线 (LSTM-HI) 能很好地反映轴承振动信号的变化趋势。

(b) 在图 8 中, 所有的 LSTM-EHI 曲线总体上表现出比 LSTM-HI 曲线更平滑的趋势, 并且振荡显著减弱。特别在图 8 (d) 的蓝色虚线矩形区域中,

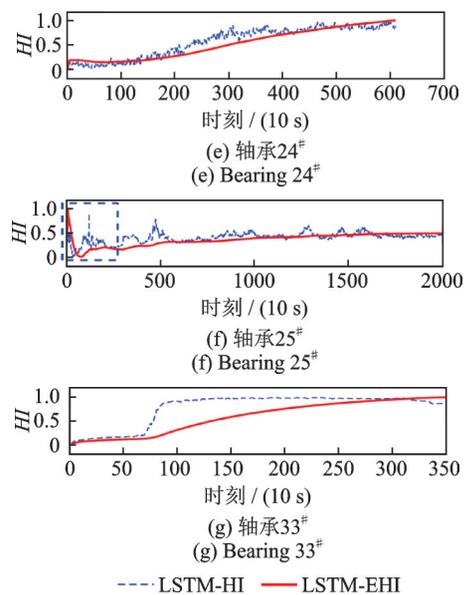


图 8 LSTM-HI 曲线和 LSTM-EHI 曲线

Fig. 8 The LSTM-HI curves and LSTM-EHI curves

可以明显看出 LSTM-HI 值 (0.6383) 在第 279 点有所增加。由于 LSTM-EHI 曲线在计算过程中利用指数函数 (式 (1)) 对从开始时间到当前时间的平均值进行处理, 因此其值 (0.1409) 明显小于 LSTM-HI 曲线的值 (0.6386), 振荡明显减弱。

(c) 以轴承 23# 为例, 使用式 (1) 中指数函数时, 从开始时间到当前时间的 HI 值的平均值作为指数函数的输入, 如图 9 (a) 所示。在图 9 (a) 中, 第 279 点附近的值正在逐渐增加; 相反, 第 263 点是图中的局部最小点, 这是因为在图 8 (d) 中, 第 263 点附近对应的 LSTM-HI 值较小且第 263 点附近点的值彼此接近。图 9 (a) 中第 263 点到第 279 点的值也略有增

长,这部分增长有效地缓冲了图8(d)中在第279点之后曲线的增长,这表明基于平均值的指数函数可以有效消除局部振荡。为了进一步加强对指数函数优越性的解释,表明其可以有效地消除曲线的整体振荡,LSTM-HI和LSTM-EHI曲线的两个相邻点的差值如图9(b)所示。在图9(b)中,LSTM-HI差值(蓝线)出现了明显的波动;相反,LSTM-EHI差值(红线)非常稳定,接近于0。此外,图9(c)为LSTM-EHI曲线的两个相邻点的差值放大扩展图。

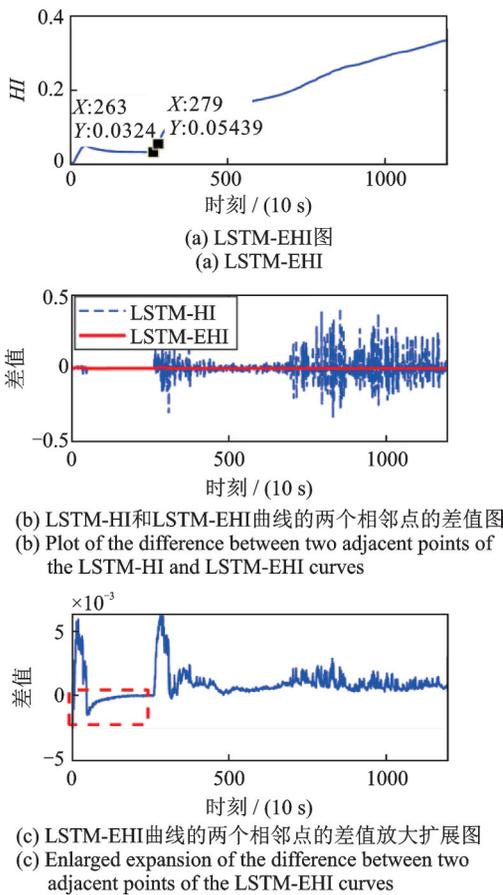


图9 使用公式(1)中指数函数时从开始时间到当前时间的平均值以及所有LSTM-EHI的差值
Fig.9 The mean value from starting time until the current time when the exponent function is used in formula (1) and the difference values for all LSTM-EHI

相较于其他轴承,轴承25#的LSTM-HI和LSTM-EHI曲线表现出较差的平滑度。在图8(f)中的蓝色虚线矩形区域(前300个数据点之前),LSTM-HI和LSTM-EHI曲线都呈现出明显的波动。这些振荡曲线掩盖了曲线在300个数据点之后逐渐增长的趋势,这种变化情况与图6(f)中轴承25#的原始振动信号一致。因此,需要删除并替换蓝色虚线区域中的LSTM-EHI曲线,以保持曲线整体平稳的增长单调性。

为了改善LSTM-EHI的单调性,并去除异常点和强烈振荡区域(如图8(f)中的蓝色虚线区域),本文采用ES模型进行数据处理。图10展示了不同轴承的LSTM-EHI with ES曲线。

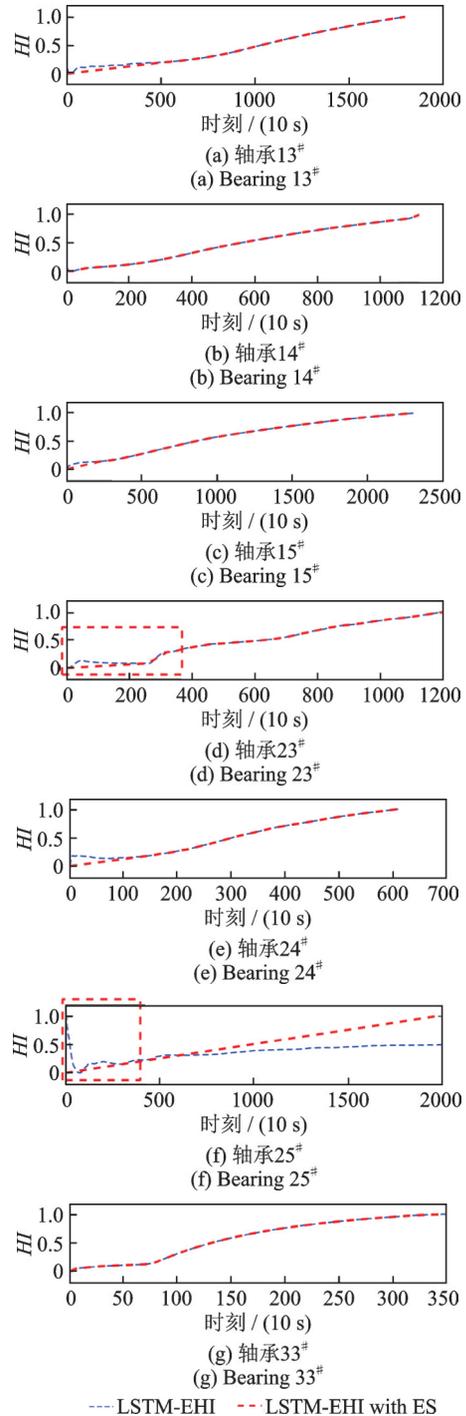


图10 LSTM-EHI曲线和LSTM-EHI with ES曲线
Fig.10 The LSTM-EHI curves and LSTM-EHI with ES curves

(a)在图10中,轴承23#和25#的LSTM-EHI曲线中的振荡区域被ES模型替换,且LSTM-EHI with ES的两条红色虚线比LSTM-EHI的两条蓝色虚线明显上升。特别是在轴承25#的红色虚线区域中,LSTM-EHI with ES的单调性显著增强,在前500个数据点之前,LSTM-EHI曲线呈下降趋势,之后则急

剧上升。这表明 ES 模型可以增强曲线的单调性。

(b)以轴承 24[#]为例,用 *Mon* 的结果和比较分析来证明 ES 模型可以增强曲线的单调性。将轴承 24[#] 的 LSTM-EHI 输入到 ES 模型中,第一个局部最小斜率点编号为 136(如图 11 中点 A),以坐标轴零点作为起始点和第 136 个数据点作为终点所对应的直线斜率为 $K_1=0.1691/136=0.00124$ (此处 0.1691 为 HI 值);第二个局部最小斜率点编号为 128 号所对应的直线的斜率为 $K_2=0.1627/128=0.00127$, $K_2 > K_1$,因此编号为 136 点的斜率为局部最小值。从第 136 个数据点开始,所有后续点的斜率值都大于第 136 点的斜率值,迭代结束。最后,图 11 中蓝色矩形虚线区域中的红色虚线 LSTM-EHI 曲线被黄色虚线曲线替换,从图中可以看出,ES 模型具有良好地增强 LSTM-EHI 曲线单调性的能力。

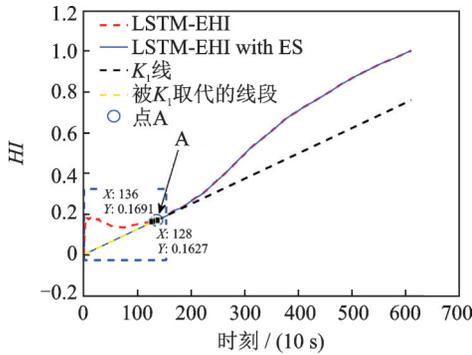


图 11 轴承 24[#] 的 LSTM-EHI with ES 曲线生成过程图
Fig. 11 The generation process figure of LSTM-EHI with ES curves of bearing 24[#]

3.3.2 与其他 HI 构建模型的对比分析

为了证明 ES 模型具有良好的增强 HI 曲线单调性的能力,本文与文献中的其他模型进行比较,如 RMS^[2],kurtosis^[2],EMD-SVD-k-means/k-medoids^[11],频率指标(Time-frequency)^[11],带指数函数的 SAE (SAE-EHI)^[20],如图 12 所示。

(a)所有轴承的 HI 曲线通过 RMS 和 kurtosis 处理之后趋势都逐渐增加,但曲线中也存在一些噪声,这些噪声很容易导致整体单调性变差,如图 12 (a),(b)所示。经 EMD-SVD-k-means/k-medoids 和 Time-frequency 处理过的 HI 曲线单调减小但趋势是平滑的,不会急剧下降或上升,如图 12(c)~(e)所示。LSTM-EHI with ES 所有曲线明显倾向于单调增加,如图 12(f)所示。

(b)图 12(g)中轴承 25[#] 的 SAE-EHI 曲线在开始阶段具有明显的振荡,但在图 12(f)中,25[#] 轴承的 LSTM-EHI with ES 曲线的变化趋势更加稳定。

(c)本文对不同 HI 构造模型的 *Mon* 值(式(17))进行了对比分析,如表 2 所示,LSTM-EHI with ES

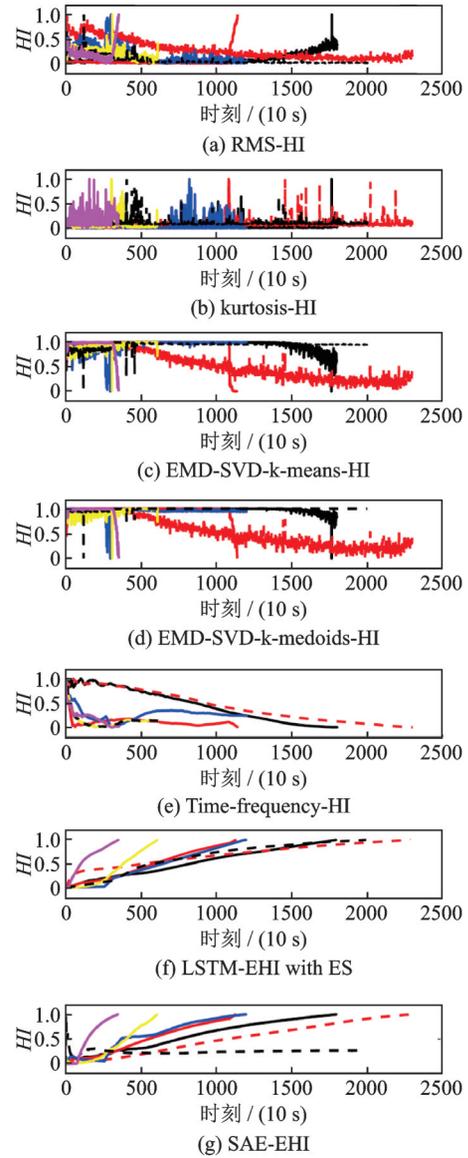


图 12 不同轴承以不同模型构造 HI 的对比图

Fig. 12 Comparison diagram of the constructed HI obtained from different models of different bearings

表 2 使用各种模型时的 *Mon* 对比结果

Tab. 2 The comparison results of *Mon* with various models

HI 模型	<i>Mon</i>						
	轴承编号						
	13 [#]	14 [#]	15 [#]	23 [#]	24 [#]	25 [#]	33 [#]
RMS-HI	0.01	0.04	0.01	0.04	0.01	0.02	0.01
kurtosis-HI	0.00	0.01	0.00	0.00	0.01	0.01	0.03
EMD-SVD-k-means-HI	0.02	0.03	0.01	0.00	0.00	0.01	0.03
EMD-SVD-k-medoids-HI	0.01	0.03	0.00	0.01	0.00	0.04	0.04
Time-frequency-HI	0.49	0.01	0.77	0.09	0.05	0.04	0.11
SAE-EHI	0.84	0.92	0.96	0.51	0.61	0.07	0.85
LSTM-EHI with ES	1.00	0.98	1.00	1.00	1.00	1.00	1.00

的所有轴承HI的Mon值均明显高于其他模型。

3.4 基于RVM的剩余寿命预测模型结果分析

为了验证所提出的RVM模型对寿命预测的有效性,本文选择IEEE PHM 2012挑战数据集中工况2组内的轴承23[#]的全生命周期数据作为实验对象,并选定检查时刻 $T(362, 483, 655) \times 10$ s对轴承进行剩余寿命预测。

3.4.1 模型参数设置

本文在进行剩余寿命预测时,采用RVM模型,选择高斯核函数,最大迭代次数为2500,目标函数最小误差为 10^{-3} ,失效阈值为819(10 s)时刻(轴承23[#]的初始故障点为8190 s)。

若高斯核宽度参数过小易导致过学习,过大易导致过平滑,因此,过大或过小会引起回归能力的下降。图13显示了不同高斯核宽度的预测结果平均相对误差。从图13中可以看出,高斯核宽度参数为0.1时,对应的平均相对误差最小,故本文的高斯核宽度参数设置为0.1。

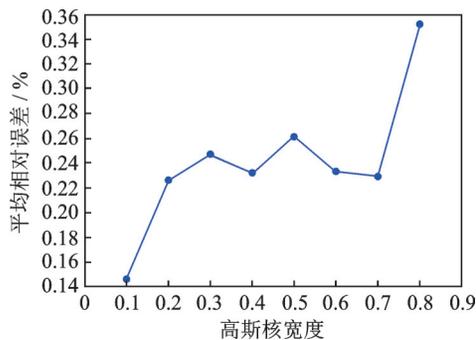


图13 不同高斯核宽度的预测结果平均相对误差

Fig. 13 Average relative error of prediction results with different Gaussian kernel widths

3.4.2 寿命预测结果

图14~16显示了预测过程和相应的结果。表3给出了各检查时刻利用RVM模型所得到的详细预测结果。从图14~16和表3中可以看出,随着更多的训练样本数据被提供,预测模型能够获得更好的预测结果。

3.4.3 RVM模型的优越性

为了证明本文所提出的RVM模型的优越性,以轴承23[#]的寿命预测为例,与SVM^[14],BP^[12]和PSO^[13]等模型进行了对比。

在SVM神经网络中,采用高斯核函数,将惩罚因子、核函数参数和不敏感系数分别设置为9,0.002和1;在BP神经网络中,神经网络层数设置为3,神经元激励函数采用Sigmoid函数,输入层、隐含层、输出层的各层神经元个数分别设置为10,14,1;在

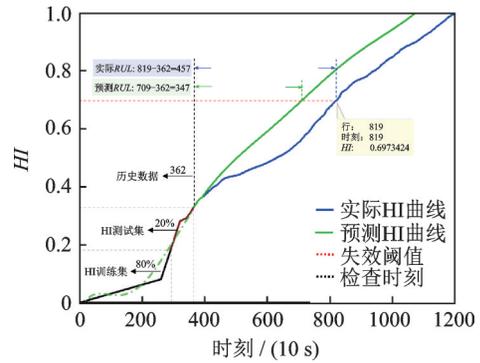


图14 轴承23[#]在检查时刻362使用RVM模型获得的预测结果
Fig. 14 Prediction results of bearing 23[#] at inspection time 362 with RVM model

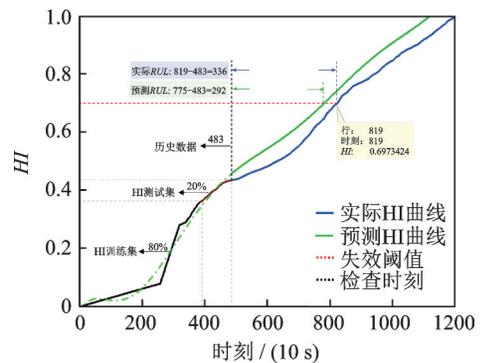


图15 轴承23[#]在检查时刻483使用RVM模型获得的预测结果
Fig. 15 Prediction results of bearing 23[#] at inspection time 483 with RVM model

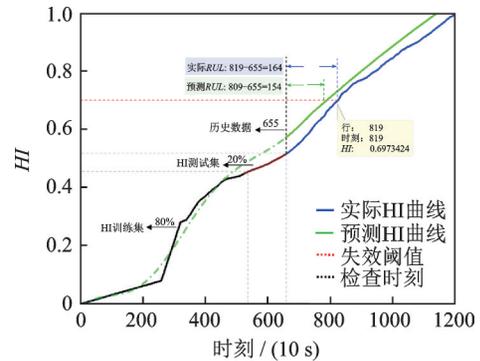


图16 轴承23[#]在检查时刻655使用RVM模型获得的预测结果
Fig. 16 Prediction results of bearing 23[#] at inspection time 655 with RVM model

表3 轴承23[#]在不同检查时刻T使用RVM模型获得的详细结果

Tab. 3 Detailed results obtained by RVM model at different inspection time T of bearing 23[#]

检查时刻/ (10 s)	预测RUL/ (10 s)	实际RUL/ (10 s)	误差/(10 s)
362	347	457	110
483	292	336	44
655	154	164	10

PSO 中,粒子数、惯性权重、学习因子等基本参数分别设置为 50,0.5,1;对比模型的最大迭代次数、目标函数最小误差、失效阈值等参数均与 RVM 模型一致。

分别对检查时刻 $T(326, 483, 655)$ (10 s) 的预测结果进行计算,各模型的预测结果对比如图 17~19 和表 4 所示。

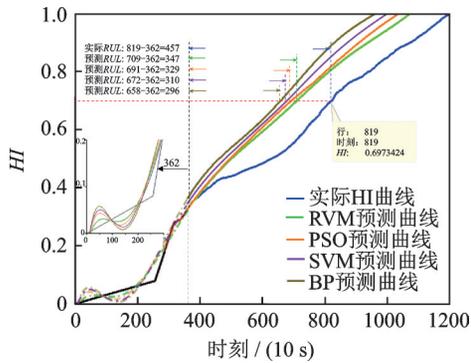


图 17 轴承 23[#]在检查时刻 362 使用不同模型的预测对比图
Fig. 17 Comparison diagram of life prediction of bearing 23[#] using different models at inspection time 362

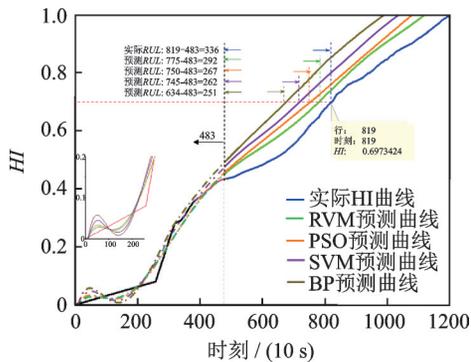


图 18 轴承 23[#]在检查时刻 483 使用不同模型的预测对比图
Fig. 18 Comparison diagram of life prediction of bearing 23[#] using different models at inspection time 483

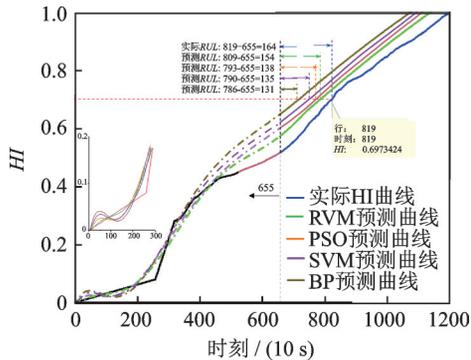


图 19 轴承 23[#]在检查时刻 655 使用不同模型的预测对比图
Fig. 19 Comparison diagram of life prediction of bearing 23[#] using different models at inspection time 655

从图 17~19 和表 4 中可以看出,在所有对比模型中,RVM 取得了最好的性能。另外,在检查时刻

表 4 轴承 23[#]在不同检查时刻 $T(10\text{ s})$ 使用不同模型的结果分析

Tab. 4 Result analysis of bearing 23[#] using different models at different inspection time $T(10\text{ s})$

模型	362		483		655	
	预测/ (10 s)	相对 误差/%	预测/ (10 s)	相对 误差/%	预测/ (10 s)	相对 误差/%
RVM	347	24.07	292	13.09	154	6.09
PSO	329	28.01	267	20.53	138	15.85
SVM	310	32.16	262	22.02	135	17.68
BP	296	35.23	251	25.30	131	20.12

$T=6550\text{ s}$ 的情况下,PSO 方法也取得了不错的预测效果,这表明提高 HI 的单调性有利于开展设备剩余寿命预测工作。

3.5 结果分析

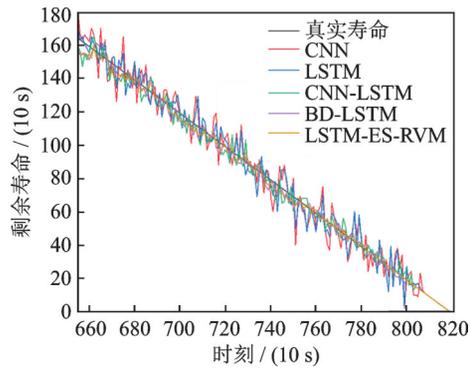
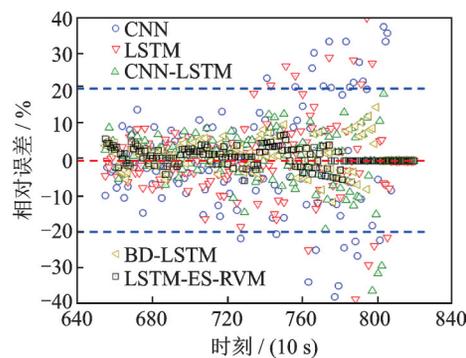
为了证明本文所提出的 LSTM-ES-RVM 方法的优越性,以轴承 23[#] 的全生命周期振动数据为例,与现有文献的方法进行了对比试验。其中,对比方法包括 CNN^[21], LSTM^[22], CNN-LSTM^[23] 和 BD-LSTM^[21]。本文分别计算了不同方法在运行时刻 [655, 819] 之间的 RUL 预测值,采用均方根误差 (RMSE) 指标进行评估,结果如表 5 所示。从表 5 的结果中可以看出,所提出的 LSTM-ES-RVM 优于其他对比方法。

表 5 轴承 23[#] 使用不同方法的均方根误差 RMSE 比较结果
Tab. 5 The comparative result of root mean square error RMSE of bearing 23[#] using different methods

模型	RMSE/(10 s)
CNN	8.63
LSTM	6.99
CNN-LSTM	4.69
BD-LSTM	3.39
LSTM-ES-RVM	2.75

为了直观地分析各种方法的预测性能,本文绘制了各对比方法在运行时刻 [655, 819] 之间的 RUL 预测值的对比曲线和相对误差图,分别如图 20 和 21 所示。

所有实验在 Win10 操作系统环境下进行,硬件环境为 Intel Core (TM) I7-7700 CPU-2.80 GHz 和 16 GB 内存,表 6 列出了轴承 23[#] 不同模型的计算时间。从表 6 中可以看出,在单模型的运行时间中,CNN 的计算时间最短;而在多模型的运行时间中,本研究提出的方法计算时间最短。

图20 轴承23[#]使用不同方法的预测结果Fig. 20 Prediction results of using different methods of bearing 23[#]图21 轴承23[#]使用不同方法的预测结果相对误差Fig. 21 Relative error of using different methods of bearing 23[#]表6 轴承23[#]不同模型的计算效率结果Tab. 6 The calculation efficiency results of different models of bearing 23[#]

模型	时间/s
CNN	489.73
LSTM	582.07
CNN-LSTM	829.45
BD-LSTM	728.56
LSTM-ES-RVM	678.34

4 结论

本文提出了一种基于LSTM-ES-RVM的滚动轴承剩余寿命预测方法。该方法减少了对人工经验和先验知识的依赖,改善了性能退化HI曲线的整体单调性,提高了滚动轴承的RUL预测精度。为验证本文所提方法,使用公开数据集(PHM2012)的实验数据进行验证,结果表明相较于文献中的其他方法,所提方法取得了最优的预测性能。

在未来的研究中,可以考虑找到一种有限数据集下自动识别失效阈值的方法,结合无监督深度学习的寿命预测模型,提高设备RUL预测的自适应性。

参考文献:

- [1] 沈长青, 汤盛浩, 江星星, 等. 独立自适应学习率优化深度信念网络在轴承故障诊断中的应用研究[J]. 机械工程学报, 2019, 55(7): 81-88.
SHEN Changqing, TANG Shenghao, JIANG Xingxing, et al. Bearings fault diagnosis based on improved deep belief network by self-individual adaptive learning rate [J]. Journal of Mechanical Engineering, 2019, 55(7): 81-88.
- [2] Lei Y G, Li N P, Lin J. A new method based on stochastic process models for machine remaining useful life prediction[J]. IEEE Transactions on Instrumentation and Measurement, 2016, 65(12): 2671-2684.
- [3] Antoni J. The spectral kurtosis: a useful tool for characterising nonstationary signals[J]. Mechanical Systems and Signal Processing, 2006, 20(2): 282-307.
- [4] Rai A, Upadhyay S H. Bearing performance degradation assessment based on a combination of empirical mode decomposition and k-medoids clustering[J]. Mechanical Systems and Signal Processing, 2017, 93: 16-29.
- [5] Zhang Y, Tang B P, Han Y, et al. Bearing performance degradation assessment based on time-frequency code features and SOM network[J]. Measurement Science and Technology, 2017, 28(4): 045601.
- [6] Cheng Y W, Zhu H P, Wu J, et al. Machine health monitoring using adaptive kernel spectral clustering and deep long short-term memory recurrent neural networks [J]. IEEE Transactions on Industrial Informatics, 2019, 15(2): 987-997.
- [7] 申彦斌, 张小丽, 夏勇, 等. Bi-LSTM神经网络用于轴承剩余使用寿命预测研究[J]. 振动工程学报, 2021, 34(2): 411-420.
SHEN Yanbin, ZHANG Xiaoli, XIA Yong, et al. Bi-LSTM neural network for remaining useful life prediction of bearings[J]. Journal of Vibration Engineering, 2021, 34(2): 411-420.
- [8] 黄宇, 冯坤, 高俊峰, 等. 结合LSTM和Self-Attention的滚动轴承剩余使用寿命预测方法[J/OL]. 振动工程学报: 1-11 [2023-05-19]. <http://kns.cnki.net/kcms/detail/32.1349.TB.20220922.0942.002.html>.
HUANG Yu, FENG Kun, GAO Junfeng, et al. Combining LSTM and Self-Attention for remaining life prediction of rolling bearings[J/OL]. Journal of Vibration Engineering: 1-11 [2023-05-19]. <http://kns.cnki.net/kcms/detail/32.1349.TB.20220922.0942.002.html>.
- [9] Tse Y L, Cholette M E, Tse P W. A multi-sensor approach to remaining useful life estimation for a slurry pump[J]. Measurement, 2019, 139: 140-151.
- [10] 谢雨洁, 肖友刚, 王田天, 等. 基于异常检测的轴承退化阶段识别方法[J]. 中南大学学报(自然科学版), 2022, 53(5): 1740-1749.

- XIE Yujie, XIAO Yougang, WANG Tiantian, et al. Degradation stage recognition of bearings based on outlier detection [J]. *Journal of Central South University (Science and Technology)*, 2022, 53(5): 1740-1749.
- [11] Tse P W, Wang D. Enhancing the abilities in assessing slurry pumps' performance degradation and estimating their remaining useful lives by using captured vibration signals[J]. *Journal of Vibration and Control*, 2017, 23(12): 1925-1937.
- [12] Ma Y, Yao M H, Liu H C, et al. State of Health estimation and Remaining Useful Life prediction for lithium-ion batteries by Improved Particle Swarm Optimization-Back Propagation Neural Network[J]. *Journal of Energy Storage*, 2022, 52: 104750.
- [13] Yao F, He W X, Wu Y X, et al. Remaining useful life prediction of lithium-ion batteries using a hybrid model [J]. *Energy*, 2022, 248: 123622.
- [14] Nieto P J G, Garcia-Gonzalo E, Lasheras F S, et al. Hybrid PSO-SVM: based method for forecasting of the remaining useful life for aircraft engines and evaluation of its reliability [J]. *Reliability Engineering & System Safety*, 2015, 138: 219-231.
- [15] Tipping M E. Sparse Bayesian learning and the relevance vector machine [J]. *Journal of Machine Learning Research*, 2001, 1: 211-244.
- [16] Xu Z F, Bashir M, Liu Q, et al. A novel health indicator for intelligent prediction of rolling bearing remaining useful life based on unsupervised learning model[J]. *Computers & Industrial Engineering*, 2023, 176: 108999.
- [17] Ren L, Cui J, Sun Y Q, et al. Multi-bearing remaining useful life collaborative prediction: a deep learning approach[J]. *Journal of Manufacturing Systems*, 2017, 43(2): 248-256.
- [18] 王豪,董广明,陈进.遗传规划提取优化特征在轴承寿命预测中的应用[J].*振动工程学报*, 2021, 34(3): 626-632.
- WANG Hao, DONG Guangming, CHEN Jin. Application of optimization feature extraction from bearing based on genetic programming for life prediction[J]. *Journal of Vibration Engineering*, 2021, 34(3): 626-632.
- [19] Nectoux P, Gouriveau R, Medjaher K, et al. PRO-NOSTIA: an experimental platform for bearings accelerated degradation tests[C]. *IEEE International Conference on Prognostics and Health Management*. Denver, USA, 2012: 1-8.
- [20] Xu F, Huang Z L, Yang F F, et al. Constructing a health indicator for roller bearings by using a stacked auto-encoder with an exponential function to eliminate concussion[J]. *Applied Soft Computing*, 2020, 89: 106119.
- [21] Zhang J J, Wang P, Yan R Q, et al. Long short-term memory for machine remaining life prediction[J]. *Journal of Manufacturing Systems*, 2018, 48: 78-86.
- [22] Shen Y Z, Tang B P, Li B, et al. Remaining useful life prediction of rolling bearing based on multi-head attention embedded Bi-LSTM network [J]. *Measurement*, 2022, 202: 111803.
- [23] Zhou J Z, Shan Y H, Liu J, et al. Degradation tendency prediction for pumped storage unit based on integrated degradation index construction and hybrid CNN-LSTM model[J]. *Sensors*, 2020, 20(15): 4277.

Remaining useful life prediction method of rolling bearing based on LSTM-ES-RVM networks

ZHOU Sheng-wen¹, GUO Shun-sheng^{1,2}, DU Bai-gang^{1,2}

(1. School of Mechanical and Electronic Engineering, Wuhan University of Technology, Wuhan 430070, China;

2. Hubei Digital Manufacturing Key Laboratory, Wuhan 430070, China)

Abstract: In order to enhance life prediction accuracy of rolling bearing, as well as its degradation performance trend and fluctuation range, a remaining useful life prediction method of rolling bearing based on LSTM-ES-RVM networks is proposed. The LSTM model is used to learn frequency domain and establish the preliminary Health Indicator (HI), so as to reduce dependence on prior knowledge and artificial experience. An improved local extremum model is proposed to eliminate violent oscillation and improve the global monotony of Health Indicator. The RVM model is used to predict remaining useful life of rolling bearing. The experimental results show that the proposed method achieves the optimal prediction performance compared to other prediction methods in the field.

Key words: remaining useful life prediction; rolling bearing; long short-term memory neural networks; health indicators; slope extreme inflection point model

作者简介:周圣文(1985—),男,博士研究生。电话:(027)87651793; E-mail: swzhou1007@163.com。

通讯作者:杜百岗(1987—),男,博士,副教授。电话:(027)87651793; E-mail: dbg767@163.com。