

基于轴承退化状态评估和改进图注意力双向门控循环单元网络的轴承剩余寿命预测

宋李俊 刘松林 辛 玉* 马婧华 谢正邱

重庆理工大学机械工程学院,重庆,400054

摘要:受工作条件和运行工况影响,滚动轴承全寿命周期的运行状态监测数据存在强噪声干扰,且轴承运行寿命退化呈非线性,严重影响剩余寿命预测的准确性。提出了一种结合高精度故障始发时刻退化状态评估和改进图注意力双向门控循环单元网络的轴承剩余寿命预测方法,并利用 XJTU-SY 全寿命周期轴承数据集验证了所提方法的有效性。研究表明,所提预测方法能有效捕获表征轴承退化状态的深度时空特征,与 CNN-LSTM 等方法相比,剩余寿命预测精度显著提高。

关键词:滚动轴承;故障始发时刻;剩余寿命预测;改进图注意力双向门控循环单元

中图分类号: TP182

DOI:10.3969/j.issn.1004-132X.2025.07.019

开放科学(资源服务)标识码(OSID):



Residual Life Prediction for Bearings Based on Bearing Degradation State Assessment and IGAT-BiGRU Network

SONG Lijun LIU Songlin XIN Yu* MA Jinghua XIE Zhengqiu

School of Mechanical Engineering, Chongqing University of Technology, Chongqing, 400054

Abstract: Due to the influences of working conditions and operating conditions, the collected status monitoring data was interfered with strong noise in full life cycle of rolling bearings, and the bearing operating life degradation was nonlinear, which seriously affected the accuracy of residual life prediction. So, a bearing residual life prediction method was proposed based on a joint high-precision FPT degradation state evaluation and an IGAT-BiGRU network, and the XJTU-SY full life cycle bearing dataset was used to verify the effectiveness of the proposed method. The results show that the proposed prediction method may effectively capture the deep spatiotemporal features that characterize the bearing degradation states, and significantly improve the residual life prediction accuracy, compared with methods such as CNN-LSTM.

Key words: rolling bearing; first predicting time(FPT); residual life prediction; improved graph attention bidirectional gate recurrent unit(IGAT-BiGRU)

0 引言

滚动轴承是机械设备内部支承旋转部件的关键部件,其工作环境恶劣,在运行过程中会发生性能退化,产生故障^[1]。精确捕捉轴承性能退化点,实时评估其健康状态和剩余寿命,能最大限度地减少维修成本,并为建立精准的维修方案提供理论支撑。

服役过程中,轴承全寿命退化过程可分为健康阶段和退化阶段^[2]。健康阶段中,轴承退化指标变化微小,达到退化点后,退化指标随服役时间呈指数增大,因此准确判断轴承的寿命退化点对提高轴承寿命预测精度具有重要意义。KONG

等^[3]使用均方根阈值法确定轴承寿命退化点。ZHU 等^[4]采用隐马尔可夫模型自动检测轴承退化状态,确定故障始发时刻(first predicting time, FPT)点。杨小东等^[5]根据初始时刻的轴承故障频率能量划分退化阶段。单一轴承退化特征难以有效表征轴承的复杂退化过程,易出现退化点误判,严重影响剩余寿命预测模型的预测精度,而复杂的 FPT 点识别模型会消耗大量的计算资源。因此,本文融合能表征机械设备退化状态的多个特征,并结合 PELT(pruned exact linear time)算法对突变点进行识别,精准识别 FPT 点。

基于深度学习的剩余寿命预测(remaining useful life, RUL)方法得到广泛使用^[6]。王玉静等^[7]将卷积神经网络(convolutional neural networks, CNN)提取的空间特征输入长短时记忆网络(long short term memory, LSTM)来预测

收稿日期:2024-06-29

基金项目:国家自然科学基金(52205144);重庆市教育委员会科学技术研究项目(KJQN202101119);重庆市博士“直通车”科研项目(CSTB2022BSXM-JCX0163)

轴承剩余寿命。WANG 等^[8]结合深度可分离卷积与注意力机制,有效提取信号的深度空间特征,使得 RUL 结果更精确。ZHU 等^[9]构建多尺度 CNN 来预测轴承寿命,并通过挖掘轴承的全局特征和局部特征来更全面和准确地提取特征。沈天浩等^[10]利用图卷积神经网络的时序信息表征能力、长短时记忆网络的长时序特征刻画能力,建立了轴承信号和剩余寿命的映射关系。传统 CNN 无法捕捉轴承振动信号中的时间依赖关系。循环神经网络(recurrent neural networks, RNN)提取的特征会随时间步的增加迅速衰退而无法有效捕捉长时间依赖关系。文献[7-9]所提方法没有很好利用信号之间的显式时间依赖关系,导致轴承的退化特征无法充分提取。因此,本文利用图数据和时空特征融合的方法进行剩余寿命预测。首先,通过黑翅鸢算法(black-winged kite algorithm, BKA)优化变分模态分解(variational mode decomposition, VMD)参数,精确提取轴承退化特征。然后,使用改进的图注意力模型、双向门控循环单元(bidirectional gate recurrent unit, BiGRU)模型分别挖掘路图中的空间特征和时间特征。最后,通过交叉注意力机制自适应地获取轴承的深度时空特征,实现轴承的剩余寿命预测。

1 图注意力网络 v2

图注意力网络(graph attention networks, GAT)通过自注意力机制学习不同邻居节点对目标节点的权重来加权聚合邻居节点,从而提高模型捕获空间特征的能力^[11]。通过自注意力机制,GAT 网络可以为每个节点获取对应的注意力系数:

$$e(\mathbf{h}_i, \mathbf{h}_j) = \text{LeakyReLU}(\alpha^\top [\mathbf{W}\mathbf{h}_i \parallel \mathbf{W}\mathbf{h}_j]) \quad (1)$$

式中: $\mathbf{h}_i, \mathbf{h}_j$ 分别为图中的节点 i 和节点 j 的坐标, $\mathbf{h}_i, \mathbf{h}_j \in \mathbf{R}^s$; $\text{LeakyReLU}(\ast)$ 为泄露线性整流函数; α^\top, \mathbf{W} 为可训练权重矩阵, $\alpha^\top \in \mathbf{R}^{2s}, \mathbf{W} \in \mathbf{R}^{s' \times s}$; s 为节点的特征维度; s' 为节点的新特征维度; \parallel 表示拼接操作。

为便于比较不同节点的注意力系数,通过 softmax 函数对注意力系数进行正则化:

$$\hat{e}_{ij} = \text{softmax}(e(\mathbf{h}_i, \mathbf{h}_j)) = \frac{\exp(e(\mathbf{h}_i, \mathbf{h}_j))}{\sum_{k \in \Omega_i} \exp(e(\mathbf{h}_i, \mathbf{h}_k))} \quad (2)$$

式中: Ω_i 为节点 V_i 的邻居节点集。

通过式(2)得到 \hat{e}_{ij} 后,可获取第 i 个节点经过图注意力层的输出特征:

$$F'_i = \sigma\left(\sum_{j \in \Omega_i} \hat{e}_{ij} \mathbf{W}\mathbf{h}_j\right) \in \mathbf{R}^{s'} \quad (3)$$

式中: $\sigma(\ast)$ 为非线性激活函数。

GAT 具有较好的噪声鲁棒性,其注意力系数

的输出进一步强化了模型决策的可追溯性。传统 GAT 无条件查询邻居节点,影响图的整体表征性能。为此,BRODY 等^[12]提出一种图注意力网络 v2(graph attention networks v2, GATv2),将 GAT 中注意力系数的 $\text{LeakyReLU}(\ast)$ 移至 \mathbf{a}^\top 和两个线性操作之间,使 GATv2 有效应用神经网络计算每个查询键对的分值:

$$e_{ij} = \mathbf{a}^\top \text{LeakyReLU}(\mathbf{W}[\mathbf{h}_i \parallel \mathbf{h}_j]) \quad (4)$$

2 基于 IGAT-BiGRU 网络的轴承剩余寿命预测方法

使用 GATv2 网络和 BiGRU 网络以及交叉注意力机制构建改进图注意力双向门控循环单元(improved graph attention bidirectional gate recurrent unit, IGAT-BiGRU)网络,以提取 VMD 去噪后的轴承信号的深度时空特征。为精确表征振动信号的退化状态,使用融合退化特征确定 FPT 点。

2.1 基于 BKA 优化 VMD 信号去噪算法

滚动轴承运行条件恶劣、工况复杂,采集的监测数据存在大量噪声干扰,严重影响轴承退化状态估计和剩余寿命预测的精度,因此,使用 VMD 算法^[13]进行信号去噪。VMD 算法的分解效果主要取决于分解层数 K 和惩罚系数 α 的选取,其中, K 直接决定了分解后的模态分量数目, α 影响模态分量的带宽。在先前的研究中, (K, α) 主要通过人工经验选择,缺乏客观的评判标准。本文使用 BKA^[14]选择超参数,其适应度函数选择最小包络熵:

$$f(k, \alpha) = \min(p_{\text{IMF}}(1), p_{\text{IMF}}(2), \dots, p_{\text{IMF}}(k)) \quad (5)$$

$$p_{\text{IMF}}(k) = - \sum_{j=1}^N P_{k,j} \log_2 P_{k,j} \quad (6)$$

$$P_{k,j} = a_{k,j} / \sum_{j=1}^N a_{k,j} \quad (7)$$

式中: $p_{\text{IMF}}(k)$ 为变分模态分解后信号第 k 个分量的包络熵, $k = 1, 2, \dots, K$; $a_{k,j}$ 为希尔伯特解调第 k 个分量后所得包络信号序列的第 j 个值, $j = 1, 2, \dots, J$ 。

综上所述,基于 BKA 优化 VMD 算法参数的具体流程如图 1 所示,算法的具体流程如下:

1) 参数初始化。输入原始信号,初始化 BKA 参数,即参数组合 (K, α) 的范围、最大迭代次数 T_{max} 和种群个数 P 等。

2) 首先,根据 (K, α) 对信号进行变分模态分解,根据式(5)计算适应度 $f(K, \alpha)$,保存当前的最小包络熵及对应的局部最优解。

3) BKA 参数更新。

① 攻击阶段中,第 i 个黑翅鸢在第 j 维 $T+1$

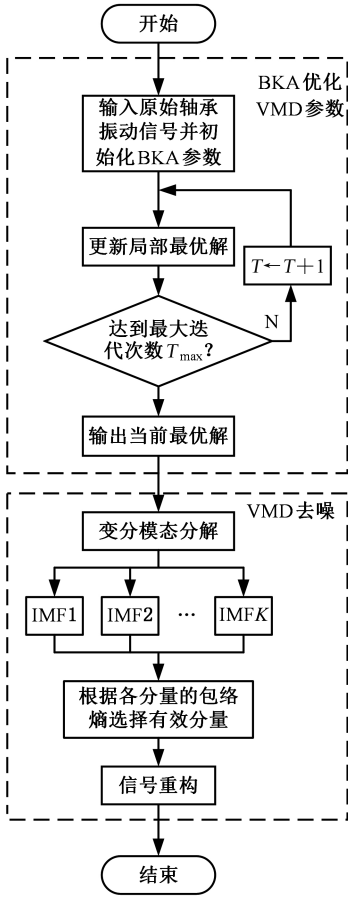


图 1 BKA 算法优化的 VMD 降噪流程图
Fig.1 Block diagram of VMD noise reduction process optimized by BKA algorithm

次迭代中的位置为

$$x_{i,j}(T+1) = \begin{cases} (n+1+n_r \sin r)x_{i,j}(T) & 0.9 < r \\ (2n_r r - n_r + 1)x_{i,j}(T) & \text{其他} \end{cases} \quad (8)$$

$$n_r = \frac{1}{20} \exp(-2(\frac{T}{T_{\max}})^2) \quad (9)$$

式中： r 为 0 ~ 1 的随机数。

② 迁徙阶段中，如果当前种群的适应度小于随机种群的适应度，则领导者不适合领导迁徙种群前进，应放弃领导并加入迁徙种群，反之，则继续引导种群到达目的地，对应计算公式为

$$x_{i,j}(T+1) = \begin{cases} x_{i,j}(T) + C_{i,t}(x_{i,j}(T) - L_j(T)) & F_{i,t} < F_{r,i,t} \\ x_{i,j}(T) + C_{i,t}(L_j(T) - m_r x_{i,j}(T)) & \text{其他} \end{cases} \quad (10)$$

$$m_r = 2 \sin(r + \pi/2) \quad (11)$$

$$C_{i,t} = \frac{1}{\pi(z_{i,t}^2 + 1)} \quad (12)$$

式中： $L_j(T)$ 为到目前为止第 T 次迭代的第 j 维中黑翅鹰的领先得分者位置； $F_{r,i,t}$ 为第 T 次迭代中第 i 只黑翅鹰在第 j 维随机位置的适应度； $F_{i,t}$ 为第 T 次迭代中第 i 只黑翅鹰在第 j 维当前位置的适应度； $z_{i,t}$ 为第 i 只黑翅鹰在第 j 个维度的均匀分布概率。

4) 依据适应度迭代优化更新后的种群，并记录当前最小包络熵及其对应的局部最优解。

5) 达到最大迭代次数后停止循环，输出最小包络熵对应的 (K, α) 。

6) 根据所求的最佳参数组合对原始信号进行变分模态分解，重构包络熵较小的 IMF 分量实现信号去噪。

2.2 基于融合退化特征确定 FPT 方法

相比于线性退化模型，分段线性退化模型能更好地反映滚动轴承退化状况，提高寿命预测的精度。为精确捕捉 FPT 点，使用均值、标准差、峰峰值、均方根值和峰度来识别 FPT 点。均值、标准差反映信号平均能量的大小和波动情况；峰峰值反映故障信号幅值的变化，对冲击型故障更敏感；均方根值反映信号的整体能量或幅度的变化，对冲击信号不够敏感；峰度是点蚀型损伤故障的常见特征指标，适用于轴承故障的早期检测^[15]。上述 FPT 识别流程中的信号特征如图 2 所示，具体过程如下：

1) 对去噪后的轴承信号确定健康监测区间，并计算均值 F_{mean} 、标准差 F_{std} 、峰峰值 F_{pp} 、均方根值 F_{rms} 和峰度 F_{kurtosis} ：

$$F_{\text{mean}} = \frac{1}{N} \sum_{i=1}^N x_i \quad (13)$$

$$F_{\text{std}} = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_i - F_{\text{mean}})^2} \quad (14)$$

$$F_{\text{pp}} = \max\{x_i\} - \min\{x_i\} \quad (15)$$

$$F_{\text{rms}} = \sqrt{\frac{1}{N} \sum_{i=1}^N x_i^2} \quad (16)$$

$$F_{\text{kurtosis}} = \frac{1}{N} \sum_{i=1}^N \frac{(x_i - F_{\text{std}})^4}{F_{\text{std}}^4} \quad (17)$$

2) 定义融合特征：

$$F_{\text{max}} = \sqrt{F_{\text{mean}}^2 + F_{\text{std}}^2 + F_{\text{pp}}^2 + F_{\text{rms}}^2 + F_{\text{kurtosis}}^2} \quad (18)$$

3) 为确定退化特征的首次大幅突变点，求解退化特征初始点 $(p_x(0), p_y(0))$ 到结束点 $(p_x(1), p_y(1))$ 的二维欧氏距离：

$$d_2 = \sqrt{(p_x(0) - p_x(1))^2 + (p_y(0) - p_y(1))^2} \quad (19)$$

4) PELT 算法^[16] 通过最小化给定数据序列中的损失函数：

$$C[S] = -\max_{\theta} \sum_{i=1}^n \log_2 f(S_i | \theta) \quad (20)$$

式中： θ 为序列 $S \in \mathbf{R}^n$ 的对数似然最大化期望值； S_i 为序列 S 的第 i 个值。

来识别数据序列中的欧氏距离突变点。本文设置惩罚项 β 来选择突变点，每增加一个突变点都会增加固定的惩罚值。PELT 算法伪代码如下：

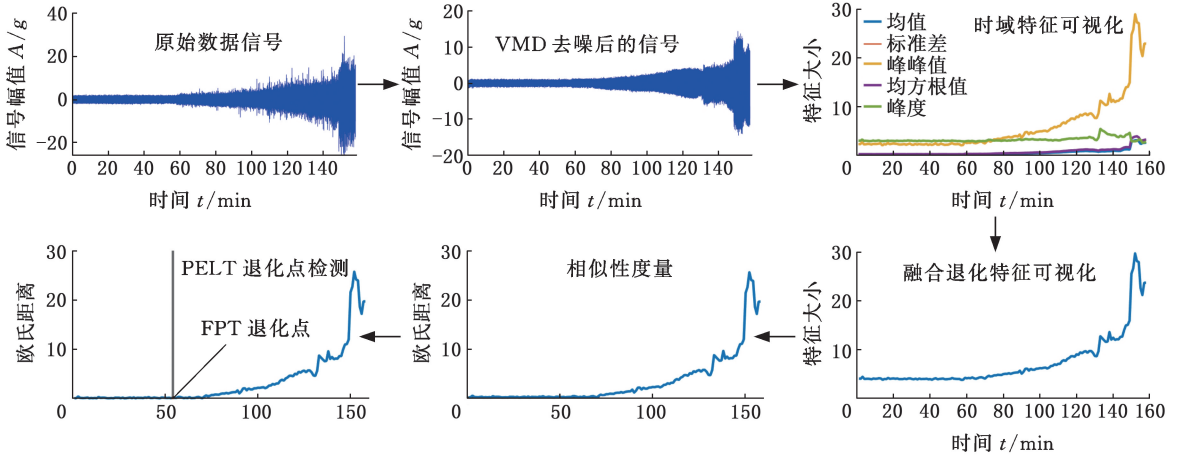


图2 FPT点识别流程不同阶段的结果

Fig.2 The results of different stages of the FPT point recognition process

输入: $y \in \mathbf{R}^n$, 损失函数 $C(*)$, 惩罚常数 β , 满足 $C(y_{(t+1),t'}) + C(y_{(t+1),t''}) + H \leq C(y_{(t+1),t''})$ 的剪枝常数 H , 最小代价序列 $F \in \mathbf{R}^n$, 待识别的突变点序列 $cp \in \mathbf{R}^n$, 其中, n 为输入序列的长度, $C(y_{(t+1),t'})$, $C(y_{(t+1),t''})$, $C(y_{(t+1),t''})$ 分别为从时间 $t+1$ 到时间 t' 、时间 $t'+1$ 到时间 t'' 和时间 $t+1$ 到时间 t'' 的代价, $t < t' < t''$ 。

初始化: $F(0) = -\beta \cdot cp(0) = \text{NULL}$, 候选突变点集合 $pts = \{0\}$ 。

迭代: 对点编号 $\tau^* = 1, 2, \dots, n$

1. 计算 $F(\tau^*) = \min_{\tau \in pts} [F(\tau) + C(y_{(t+1),\tau^*}) + \beta]$, 其中, τ 为不可能成为最优解的点编号
2. 记录最优分割点 $\hat{\tau} = \arg\{\min_{0 \leq \tau < \tau^*} [F(\tau) + C(y_{(\hat{\tau}+1),\tau^*}) + \beta]\}$
3. 更新突变点序列 $cp(\tau^*) = [cp(\hat{\tau}), \hat{\tau}]$
4. 剪枝, 更新候选集 $pts_{\tau^*+1} = \arg\{F(\tau) + C(y_{\hat{\tau}+1,\tau^*}) + \beta + H > F(\tau^*) + \beta\}$

输出: 识别出的突变点序列 $cp(n)$

2.3 IGAT-BiGRU 网络寿命预测模型

图数据^[17] 通常表示为 $G(\mathbf{X}_{n_a \times n_f}, \mathbf{A}, \mathbf{E})$, 其中, $\mathbf{X}_{n_a \times n_f}$ 为节点特征矩阵, \mathbf{A} 为邻接矩阵, \mathbf{E} 为边的集合, n_a 为节点的数量, n_f 为特征维度^[18]。

路图适用于单传感器时间序列数据的建模^[19], 边连接相邻时间段的轴承振动信号, 边的权重表示相邻时间段之间的时间依赖关系, 路图结构如图3所示。

交叉注意力机制可自适应地融合提取到的2种信号特征, 获取轴承信号的深度有效特征, 减少模型对轴承无效特征的计算。交叉注意力的计算公式为

$$C_A(\mathbf{X}_1, \mathbf{X}_2) = \text{Softmax}\left(\frac{\mathbf{X}_1 \mathbf{W}_Q (\mathbf{X}_1 \mathbf{W}_K)^T}{\sqrt{d_2}}\right) \mathbf{X}_2 \mathbf{W}_V \quad (21)$$

式中: $\mathbf{X}_1 \in \mathbf{R}^{n \times d_{qk}}$, $\mathbf{X}_2 \in \mathbf{R}^{n \times d_{qv}}$ 分别为查询矩阵和键值矩阵; d_{qk} , d_{qv} 分别为查询矩阵和键值矩阵的特征维度;

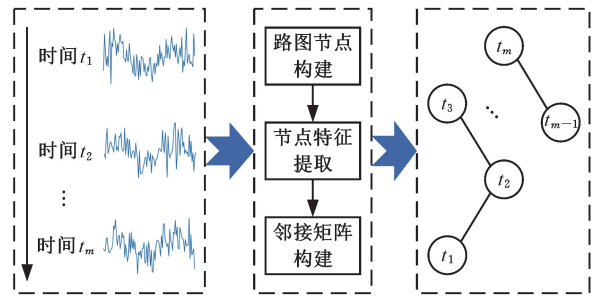


图3 路图结构示意图

Fig.3 Schematic diagram of the path graph structure

$\mathbf{W}_Q \in \mathbf{R}^{d_{qk} \times d_k}$, $\mathbf{W}_K \in \mathbf{R}^{d_{qv} \times d_k}$, $\mathbf{W}_V \in \mathbf{R}^{d_{qv} \times d_k}$ 分别为学习到的查询参数矩阵、键参数矩阵和值参数矩阵; d_k 为查询矩阵和键值矩阵的维度。

为提取轴承信号的深度时空退化特征, 寿命预测模型不仅需要提取信号的深度特征, 而且需要提取振动信号之间的时间依赖关系, 而传统的神经网络难以协同优化时空特征。为此, 提出图4所示的IGAT-BiGRU寿命预测模型。该模型使用并行的GATv2、BiGRU网络分别提取滚动轴承的空间特征和时间特征。GATv2通过动态注意力机制传递和聚合轴承信号有效退化空间特征, 提高模型的表达能力和泛化能力。BiGRU网络通过双向门控机制全面捕捉轴承信号的有效退化时间特征, 全面提取轴承信号的上下文退化信

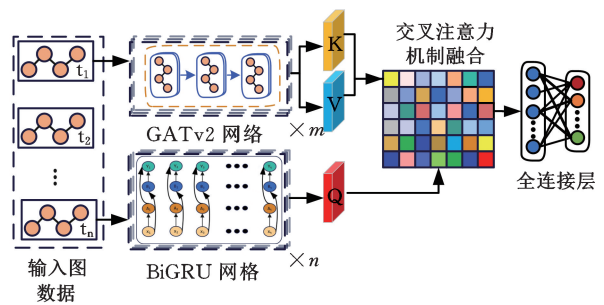


图4 IGAT-BiGRU 模型

Fig.4 IGAT-BiGRU model

息。最后,使用交叉注意力机制自适应地获取滚动轴承的深度融合时空特征帮助模型精确捕捉振动信号中的复杂时空关系,增强模型对噪声和异常值的鲁棒性,提高滚动轴承剩余寿命预测的精度。

2.4 轴承 RUL 流程

基于 IGAT-BiGRU 模型,提出图 5 所示的轴承 RUL 预测框架。

1) 数据预处理。首先,使用 BKA 算法优化 VMD,得到最佳参数集合 (K, α) ,依照最佳参数组合 (K, α) 对信号进行变分模态分解,计算各个分量对应的包络熵,重构包络熵较小的分量,实现信号去噪。将去噪后的轴承振动信号通过无重叠滑动窗口划分为 n_s 个样本,并对每个样本提取特征。

2) FPT 退化点识别。首先,利用提取的样本特征计算退化特征 F_{mex} 。然后,求解退化特征的欧氏距离。最后,将 PELT 算法识别的欧氏距离曲线的突变点作为 FPT 点,从而将全寿命信号分为健康阶段和退化阶段。

3) 剩余寿命预测。首先,利用提取的数据特征构建路图,并对路图特征进行归一化处理。然后,使用 IGAT-BiGRU 模型挖掘路图的时间特征和空间特征,并通过交叉注意力机制自适应地获取轴承深度时空特征,利用全连接层获取每个时间段的剩余寿命,同时保存最佳模型。最后,将测试集输入到训练完成的 IGAT-BiGRU 模型,并评价模型的有效性和先进性。

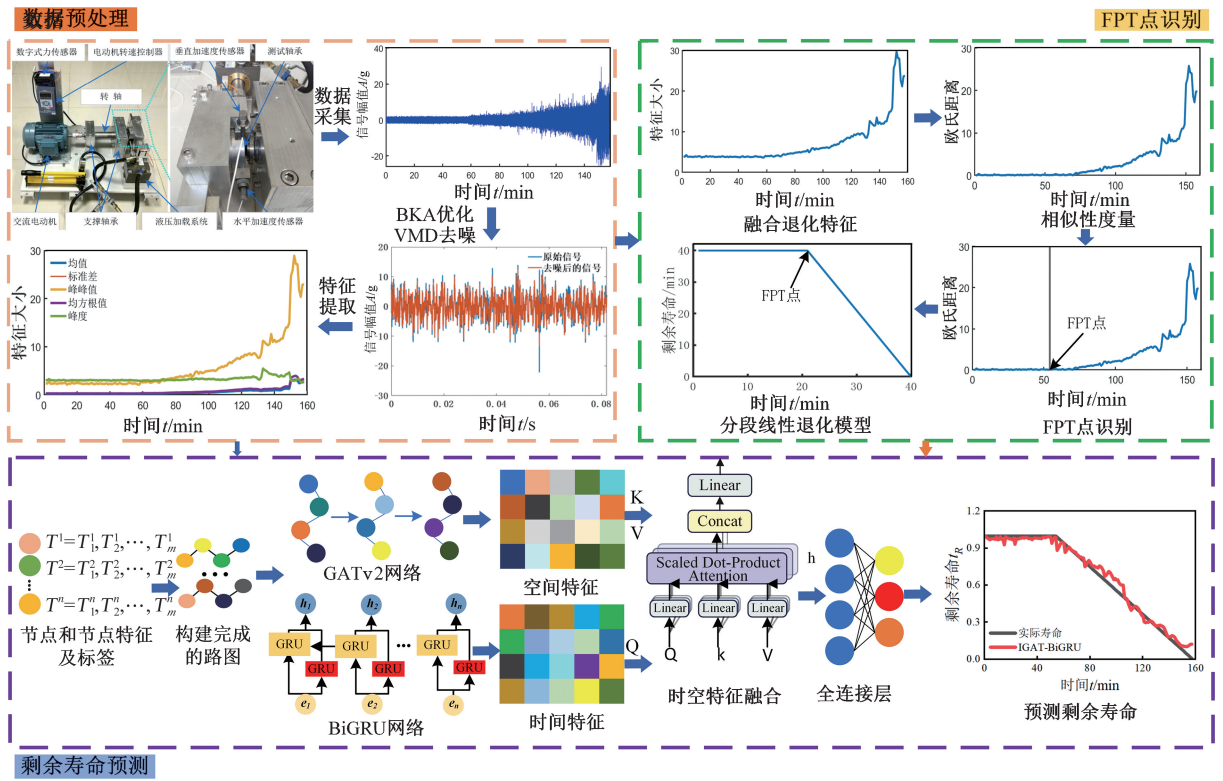


图 5 基于 IGAT-BiGRU 的滚动轴承 RUL 预测框架

Fig.5 RUL prediction framework of rolling bearing based on IGAT-BiGRU

2.5 模型评价指标

为验证模型的有效性,采用平均绝对误差 L_{MAE} 、 L_{RMSE} 和决定系数 L_{R^2} :

$$L_{MAE} = \frac{1}{m} \sum_{i=1}^m |y_i - \hat{y}_i| \quad (22)$$

$$L_{RMSE} = \sqrt{\frac{1}{m} \sum_{i=1}^m (y_i - \hat{y}_i)^2} \quad (23)$$

$$L_{R^2} = 1 - \frac{\sum_{i=1}^m (\hat{y}_i - y_i)^2}{\sum_{i=1}^m (\bar{y} - y_i)^2} \quad (24)$$

式中: m 为样本点个数; y_i 为第 i 个样本的实际值; \hat{y}_i 为第 i 个样本的预测值; \bar{y} 为实际值的均值。

来评价模型预测效果,其中, L_{MAE} 、 L_{RMSE} 越小,拟

合精度越高; L_{R^2} 值越接近 1,拟合精度越高。

3 试验验证与分析

为验证所提方法的有效性,采用 XJTU-SY 数据集验证基于 IGAT-BiGRU 的剩余寿命预测性能。算法的运行环境是 Python3.11 和 PyTorch Geometric 2.5.2,实验所用 CPU 为 i5-12400F,内存 16GB,显卡 NVIDIA RTX3050。

3.1 数据来源

XJTU-SY 数据集^[20-21]包括 3 组工况 15 组轴承的全寿命周期振动信号,如表 1 所示。为方便计算,将一次采集的 32 768 个信号设为一个样本。

表 1 XJTU 轴承数据集描述

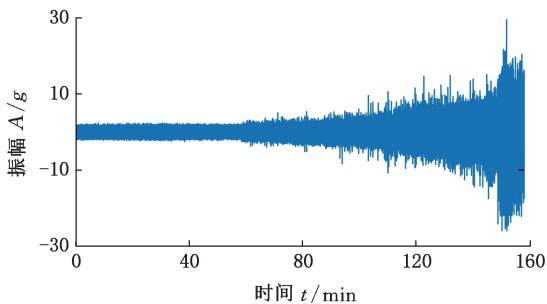
Tab.1 Description of XJTU bearing data set

工况	数据集	故障部位
工况 1 转速 2100 r/min 负载 12kN	轴承 1-1	外圈
	轴承 1-2	外圈
	轴承 1-3	外圈
	轴承 1-4	保持架
	轴承 1-5	内外圈
工况 2 转速 2250 r/min 负载 11 kN	轴承 2-1	内圈
	轴承 2-2	外圈
	轴承 2-3	保持架
	轴承 2-4	外圈
	轴承 2-5	外圈
工况 3 转速 2400 r/min 负载 10 kN	轴承 3-1	外圈
	轴承 3-2	内外圈、滚动体和保持架
	轴承 3-3	内圈
	轴承 3-4	内圈
	轴承 3-5	外圈

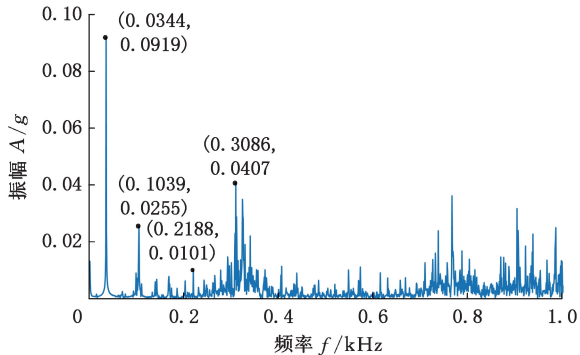
3.2 基于 BKA 优化 VMD 的信号去噪

通过 BKA 算法确定 VMD 算法的最优参数组合 (K, α) , 以最小包络熵为适应度函数。算法相关参数设置如下: 优化变量数目为 2, K 的迭代范围为 3~10, α 的迭代范围为 100~2500, 最大迭代次数 T 为 20, 种群大小 P 为 30, 对每个轴承进行参数寻优。

为评估 BKA 优化 VMD 算法参数的有效性, 以轴承 1-3 为例进行验证分析, 其信号时域波形如图 6 所示。使用 BKA 寻优 VMD 算法得到的



(a) XJTU-SY 轴承数据集轴承 1-3 时域信号图



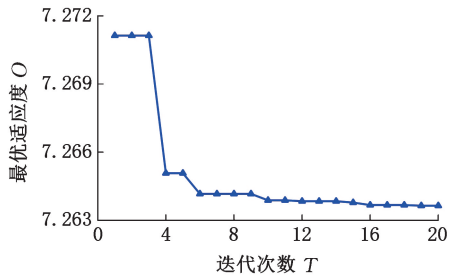
(b) XJTU-SY 轴承数据集轴承 1-3 时域信号图

图 6 XJTU-SY 轴承数据集轴承 1-3 时域和频域信号对比图

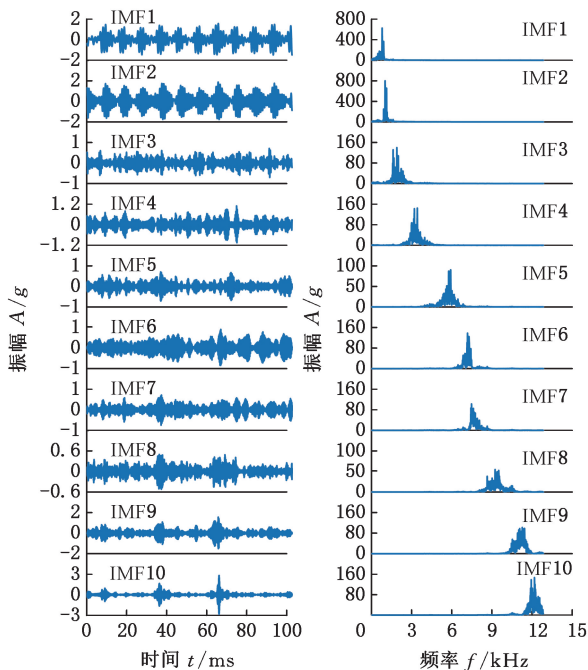
Fig.6 Comparison chart of time-domain and frequency-domain signals of bearings 1-3 in the XJTU-SY bearing dataset

最佳参数组合 $(K, \alpha) = (10, 1200)$ 。

如图 7 所示, 原始信号 10 个分量的中心频率由低到高均匀分布, 未发生模态混叠, 且与原始信号频谱图对应。为有效提取强噪声信号包含的故障特征, 计算各 IMF 分量与原故障信号的包络熵。包络熵越小, 代表信号相关程度越高, 所包含的故障特征越丰富。计算轴承 1-3 各个分量的包络熵, 如表 2 所示。



(a) BKA 优化 VMD 迭代过程图



(b) BKA 优化 VMD 后的时域信号 (c) BKA 优化 VMD 后的频域信号

图 7 BKA 优化 VMD 后的分解效果图

Fig.7 The decomposition effect after optimization of VMD by BKA

由表 2 可得, IMF2-IMF10 的包络熵较小, 因此对它们的信号进行信号重构, 得到最终的去噪信号, 计算得到去噪信号的信噪比为 11.81dB, 去噪后的信号与原始信号如图 8 所示。对部分轴承进行超参数寻优, 得到的最佳 (K, α) 见表 3。

3.3 基于路图的图数据集构建

FPT 点识别是剩余寿命预测的重要环节, 可帮助模型准确理解轴承退化过程。采用 2.2 节的方法识别 XJTU-SY 轴承数据集的 FPT 点, 结果如表 4 所示。

表 2 各分量的最小包络熵

Tab.2 Minimum envelope entropy for each component

分量	包络熵	分量	包络熵
IMF1	7.525	IMF6	7.463
IMF2	7.468	IMF7	7.457
IMF3	7.488	IMF8	7.472
IMF4	7.460	IMF9	7.439
IMF5	7.468	IMF10	7.264

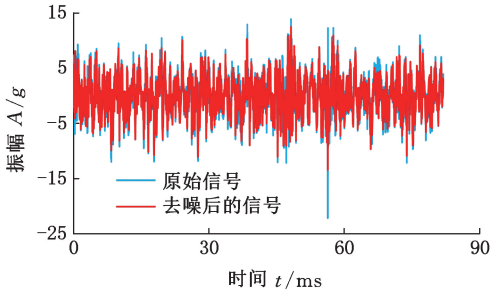


图 8 去噪信号与原始信号的对比图

Fig.8 Comparison between denoised signal and original signal

表 3 轴承去噪的 VMD 最佳参数组合

Tab.3 The optimal parameter combination of VMD for bearing denoising

编号	(K, α)	编号	(K, α)
轴承 1-1	(4, 1899)	轴承 2-5	(7, 2403)
轴承 1-2	(10, 1220)	轴承 3-3	(8, 2500)
轴承 1-3	(10, 1200)	轴承 3-4	(9, 2500)
轴承 2-2	(5, 125)	轴承 3-5	(9, 168)
轴承 2-4	(10, 395)		

表 4 XJTU-SY 数据集轴承初始故障点的识别结果

Tab.4 Initial fault point identification results of XJTU-SY dataset bearing min

工况	轴承	试验时长	初始故障点	参照故障点 ^[10]
1	1-1	123	64	77
	1-2	161	49	35
	1-3	158	54	59
2	2-2	161	52	46
	2-4	42	26	30
	2-5	339	144	122
3	3-3	371	330	340
	3-4	1515	1416	1418
	3-5	114	6	7

为构建路图数据集,将提取的数据特征作为路图的节点特征。表示振动信号的时间依赖关系的边的权重系数由余弦相似度度量。

数据特征量纲不一致对模型的预测效果造成较大影响,因此将数据特征归一化以提高模型收敛速度:

$$\gamma_{i, \text{norm}} = \frac{\gamma_i - \gamma_{\min}}{\gamma_{\max} - \gamma_{\min}} \quad (25)$$

式中: γ_i 为第 i 个时间点的特征值; γ_{\max} 、 γ_{\min} 分别为特征的最大值和最小值。

本文采用水平方向的振动信号进行实验分析^[22]。实验数据集的划分如表 5 所示。

表 5 XJTU 数据集实验数据集划分

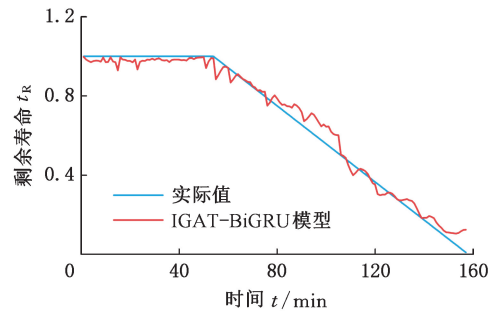
Tab.5 Experimental dataset division of XJTU dataset

任务	训练集	测试集
1	轴承 1-1、轴承 1-2	轴承 1-3
2	轴承 2-2、轴承 2-5	轴承 2-4
3	轴承 3-3、轴承 3-5	轴承 3-4

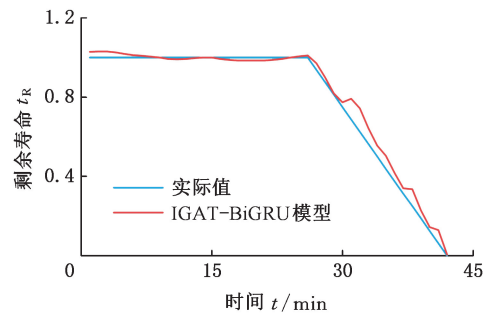
3.4 模型验证

3.4.1 实验结果分析

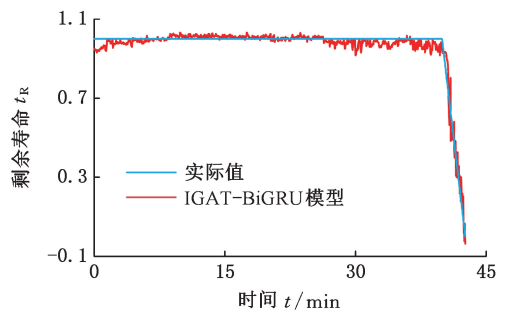
首先,使用训练集训练出最佳的 IGAT-BiGRU 模型。然后,将测试集输入训练完的模型,得到剩余寿命预测值,如图 9 所示,其中,模型的优化器选择 Adam,学习率选择 0.0008。



(a) 轴承 1-3



(b) 轴承 2-4



(c) 轴承 3-4

图 9 3 个轴承的 RUL 预测图

Fig.9 RUL prediction plot for 3 bearings

为分析 IGAT-BiGRU 模型的有效性,采用 t-SNE 算法将各阶段提取到的轴承退化特征进行降维和可视化分析。以任务 1 为例,将数据按照轴承的退化时间分为 5 组,然后可视化原始信号输入阶段、GATv2 模块输出阶段、BiGRU 模块输出

阶段和 CrossAttention 模块输出阶段提取得到的轴承退化特征。由图 10 可知,在原始信号输入阶段,轴承退化特征随机混淆分布,无法捕捉轴承振动信号中的时序关系。在 GATv2 模块和 BiGRU

模块输出阶段,模型基本能捕捉振动信号中的时序关系,但数据特征还存在随机混淆现象。在 CrossAttention 模块输出阶段,捕捉的时空特征已按照轴承运行时间顺序首尾相接,分类聚集较为清晰。这表明该模型可精确捕捉振动信号之间的时空依赖关系,准确识别轴承性能的退化过程。

3.4.2 与相关构图方式的结果比较

为评价路图显式建模信号时间依赖关系的有效性,选取完全图、K 近邻图、半径图^[23]进行对比实验。实验选择任务 1,模型选择 IGAT-BiGRU 模型,不同构图方法的 RUL 预测评估指标计算结果如表 6 所示,路图得到的评估指标效果都最优,且该方法在 FPT 点之后的剩余寿命预测结果更优。相较于路图,另外 3 种构图方式均无法显式建模轴承振动信号间的时间依赖关系。因此,显式建模监测轴承振动信号之间隐含的时间依赖关系可大幅提高轴承剩余寿命的预测精度。

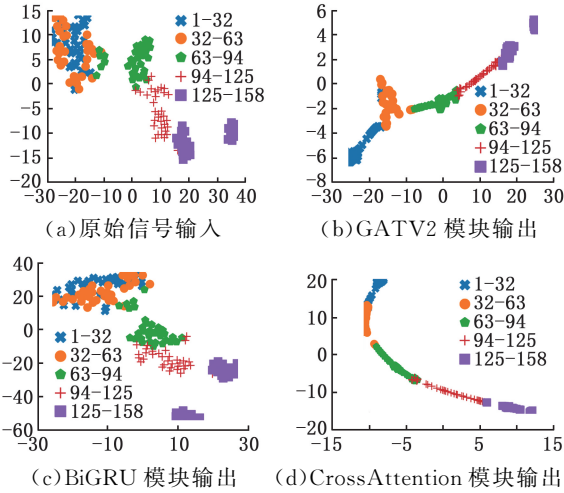


图 10 轴承 1-3 时序特征可视化结果

Fig.10 Bearing 1-3 timing feature visualization results

表 6 基于 XJTU-SY 数据集的不同构图方法的实验结果

Tab.6 Experimental results of different composition methods based on XJTU-SY dataset

构图方式	MAE			RMSE			R ²			时间/s
	均值		标准差	均值		标准差	均值		标准差	
	全局退化	FPT 点退化		全局退化	FPT 点退化		全局退化	FPT 点退化		
完全图	0.093	0.123	0.009	0.113	0.139	0.008	0.905	0.841	0.016	17.71
K 近邻图	0.103	0.134	0.011	0.122	0.144	0.014	0.844	0.796	0.024	18.45
半径图	0.104	0.136	0.011	0.124	0.142	0.013	0.870	0.801	0.018	17.52
路图	0.032	0.038	0.006	0.042	0.048	0.005	0.984	0.971	0.006	15.79

3.4.3 模型的消融实验

为证明所提出方法的有效性和优越性,对 GATv2-BiGRU-Conconrate 模型、GATv2-BiLSTM-CrossAttention 模型、数据未去噪的 IGAT-BiGRU 模型、数据未阶段划分的 IGAT-BiGRU 模型与 IGAT-BiGRU 模型进行消融实验分析,并比较 MAE、RMSE 和 R²。图 11 为不同模型的 RUL 预测效果图,表 7 所示为 RUL 预测的消融实验评价指标。

对比 IGAT-BiGRU 模型与 GATv2-BiGRU-Conconrate 模型可知,交叉注意力更自适应融合

时空特征。对比 IGAT-BiGRU 模型与 GATv2-BiLSTM-CrossAttention 可知,IGAT-BiGRU 模型的预测曲线与实际寿命曲线的拟合精度更高且训练更快,证明了 GRU 网络能充分利用轴承退化信息,且训练时间也有所缩短。对比 IGAT-BiGRU 模型与数据未去噪的 IGAT-BiGRU 模型可知,去噪后的数据特征更为显著,有助于减少特征提取的复杂性。对比 IGAT-BiGRU 模型与数据未阶段划分的 IGAT-BiGRU 模型可知,分段线性退化模型可更精确地表征轴承退化状态,提高模型捕捉退化特征的能力。

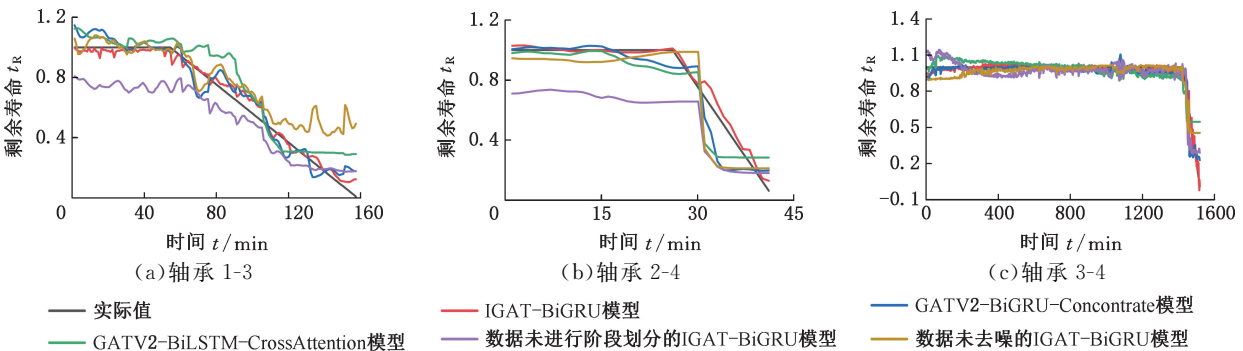


图 11 所提模型与消融分析效果对比

Fig.11 Comparison of the proposed model with the effect of ablation analysis

表 7 各轴承消融实验结果对比

Tab.7 Comparison of the results of ablation experiments for each bearing

模型	轴承	MAE			RMSE			R ²			时间/s
		均值		标准差	均值		标准差	均值		标准差	
		全局退化	FPT点退化		全局退化	FPT点退化		全局退化	FPT点退化		
GATv2-BiGRU-Conconrate	1-3	0.065	0.073	0.027	0.078	0.085	0.025	0.859	0.811	0.052	17.14
	2-4	0.081	0.132	0.016	0.107	0.159	0.013	0.883	0.730	0.029	17.86
	3-4	0.042	0.123	0.015	0.061	0.154	0.014	0.810	0.730	0.089	33.42
GATv2-BiLSTM-CrossAttention	1-3	0.106	0.127	0.020	0.127	0.148	0.019	0.852	0.733	0.041	17.91
	2-4	0.091	0.143	0.012	0.126	0.172	0.007	0.837	0.674	0.018	20.21
	3-4	0.048	0.096	0.009	0.068	0.143	0.007	0.767	0.730	0.051	34.80
数据未阶段划分的IGAT-BiGRU	1-3	0.110	0.167	0.013	0.130	0.209	0.015	0.796	0.646	0.047	19.57
	2-4	0.129	0.178	0.015	0.155	0.208	0.014	0.727	0.647	0.055	20.72
	3-4	0.117	0.139	0.082	0.141	0.165	0.327	0.761	0.711	0.127	33.26
数据未去噪的IGAT-BiGRU	1-3	0.112	0.114	0.089	0.127	0.127	0.117	0.847	0.803	0.131	17.84
	2-4	0.142	0.194	0.062	0.126	0.223	0.107	0.837	0.604	0.078	18.45
	3-4	0.106	0.125	0.049	0.087	0.155	0.079	0.867	0.728	0.079	34.37
IGAT-BiGRU	1-3	0.032	0.038	0.006	0.042	0.048	0.005	0.984	0.971	0.006	15.79
	2-4	0.026	0.050	0.004	0.039	0.060	0.007	0.985	0.957	0.018	18.24
	3-4	0.020	0.056	0.012	0.028	0.069	0.003	0.960	0.948	0.009	30.31

3.4.4 模型的对比实验

为直观验证模型的性能,将 IGAT-BiGRU 模型与 CNN-LSTM^[24]、GAT-LSTM^[25]、GAT-TCN^[26]进行对比,使用相同数据集预测轴承的剩余寿命。图 12 为不同模型的 RUL 预测对比图,

表 8 所示为 RUL 预测的评价指标。从表 8 中可以看出,IGAT-BiGRU 模型具有更优的 MAE、RMSE 和 R²,且本方法具有较好的稳健性,证明了所提模型的有效性和先进性。

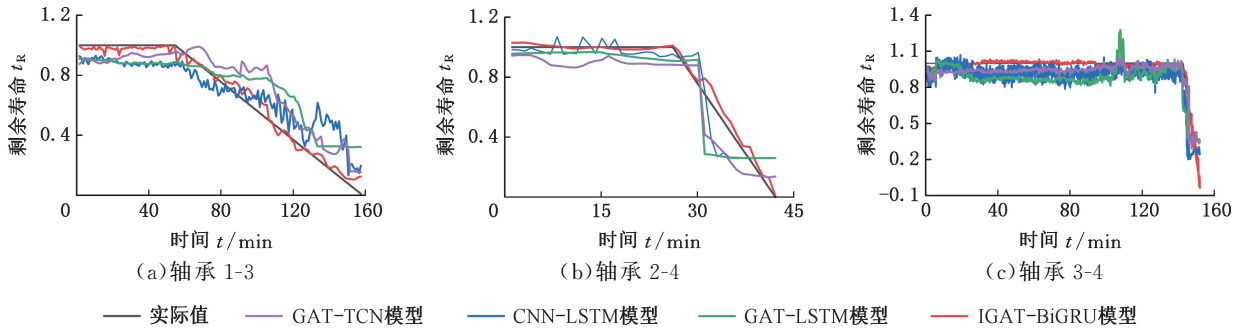


图 12 所提模型与 3 种主流模型寿命预测曲线

Fig.12 Lifetime prediction curves of the proposed model and three mainstream models

表 8 各轴承对比实验结果对比

Tab.8 Comparison of test results for each bearing

模型	轴承	MAE			RMSE			R ²			时间/s
		均值		标准差	均值		标准差	均值		标准差	
		全局退化	FPT点退化		全局退化	FPT点退化		全局退化	FPT点退化		
CNN-LSTM	1-3	0.125	0.131	0.022	0.145	0.162	0.023	0.809	0.680	0.060	26.46
	2-4	0.076	0.136	0.007	0.106	0.163	0.008	0.885	0.565	0.020	7.63
	3-4	0.085	0.217	0.012	0.106	0.253	0.012	0.765	0.588	0.109	17.62
GAT-LSTM	1-3	0.143	0.159	0.027	0.162	0.184	0.029	0.785	0.584	0.075	8.21
	2-4	0.093	0.163	0.012	0.129	0.198	0.016	0.799	0.527	0.050	6.46
	3-4	0.114	0.208	0.020	0.127	0.232	0.015	0.801	0.599	0.034	34.07
GAT-TCN	1-3	0.128	0.153	0.023	0.147	0.171	0.019	0.830	0.642	0.030	11.25
	2-4	0.112	0.132	0.011	0.125	0.155	0.007	0.823	0.711	0.021	9.71
	3-4	0.059	0.132	0.009	0.070	0.160	0.014	0.784	0.714	0.038	41.17
IGAT-BiGRU	1-3	0.032	0.038	0.006	0.042	0.048	0.005	0.984	0.971	0.006	15.79
	2-4	0.026	0.050	0.004	0.039	0.060	0.007	0.985	0.957	0.018	18.24
	3-4	0.020	0.056	0.012	0.028	0.069	0.003	0.960	0.948	0.009	30.31

4 结论

1) 路图可显式建模振动信号之间的时间依赖关系, 显著提高了模型捕捉时间特征的能力和模型预测剩余寿命的精度。

2) 对轴承振动信号的融合特征进行 FPT 退化点识别可使模型精确捕捉轴承的退化状态, 提高了模型剩余寿命的预测精度。

3) 设计的 IGAT-BiGRU 网络挖掘出振动信号之间隐含的时空依赖关系, 帮助模型捕捉更加复杂和非线性的轴承退化过程, 增强了模型对噪声和异常数据的鲁棒性, 进一步提高了轴承性能退化过程信息挖掘的能力。

参考文献:

- [1] 曾大懿, 杨基宏, 邹益胜, 等. 基于并行多通道卷积长短时记忆网络的轴承寿命预测方法[J]. 中国机械工程, 2020, 31(20):2454-2462.
ZENG Dayi, YANG Jihong, ZOU Yisheng, et al. Bearing Life Prediction Method Based on PMCCNN-LSTM[J]. China Mechanical Engineering, 2020, 31(20):2454-2462.
- [2] YAN Bingxin, MA Xiaobing, HUANG Guifa, et al. Two-stage Physics-based Wiener Process Models for Online RUL Prediction in Field Vibration Data [J]. Mechanical Systems and Signal Processing, 2021, 152:107378.
- [3] KONG Xuefeng, YANG Jun. Remaining Useful Life Prediction of Rolling Bearings Based on RMS-MAVE and Dynamic Exponential Regression Model [J]. IEEE Access, 2019, 7:169705-169714.
- [4] ZHU Jun, CHEN Nan, SHEN Changqing. A New Data-driven Transferable Remaining Useful Life Prediction Approach for Bearing under Different Working Conditions [J]. Mechanical Systems and Signal Processing, 2020, 139:106602.
- [5] 杨小东, 纪国宜. 基于融合 CNN 的滚动轴承剩余寿命预测[J]. 国外电子测量技术, 2021, 40(3):62-67.
YANG Xiaodong, JI Guoyi. Remaining Useful Lifetime Prediction Method of Rolling Bearing Based on Fusion-CNN [J]. Foreign Electronic Measurement Technology, 2021, 40(3):62-67.
- [6] 付国忠, 杜华, 张志强, 等. 基于注意力机制和 CNN-BiLSTM 模型的滚动轴承剩余寿命预测[J]. 核动力工程, 2023, 44(增刊 2):33-38.
FU Guozhong, DU Hua, ZHANG Zhiqiang, et al. Remaining Useful Life Prediction of Rolling Bearings Based on Attention Mechanism and CNN-BiL-

- STM [J]. Nuclear Power Engineering, 2023, 44(S2):33-38.
- [7] 王玉静, 李少鹏, 康守强, 等. 结合 CNN 和 LSTM 的滚动轴承剩余使用寿命预测方法[J]. 振动测试与诊断, 2021, 41(3):439-446.
WANG Yujing, LI Shaopeng, KANG Shouqiang, et al. Method of Predicting Remaining Useful Life of Rolling Bearing Combining CNN and LSTM [J]. Journal of Vibration, Measurement & Diagnosis, 2021, 41(3):439-446.
- [8] WANG Biao, LEI Yaguo, LI Naipeng, et al. Deep Separable Convolutional Network for Remaining Useful Life Prediction of Machinery [J]. Mechanical Systems and Signal Processing, 2019, 134:106330.
- [9] ZHU Jun, CHEN Nan, PENG Weiwen. Estimation of Bearing Remaining Useful Life Based on Multi-scale Convolutional Neural Network [J]. IEEE Transactions on Industrial Electronics, 2019, 66(4):3208-3216.
- [10] 沈天浩, 丁康, 黎杰, 等. 图结构联合时序数据驱动的装备剩余使用寿命预测方法[J]. 机械工程学报, 2023, 59(12):183-194.
SHEN Tianhao, DING Kang, LI Jie, et al. Graph Structure and Temporal Data Driven Remaining Useful Life Prediction Method for Machinery [J]. Journal of Mechanical Engineering, 2023, 59(12):183-194.
- [11] VELIKOVI P, CUCURULL G, CASANOVA A, et al. Graph Attention Networks [J]. arXiv, 2018:1710.10903.
- [12] BRODY S, ALON U, YAHAV E. How Attentive Are Graph Attention Networks? [J]. arXiv, 2021:2105.14491.
- [13] DRAGOMIRETSKIY K, ZOSSO D. Variational Mode Decomposition [J]. IEEE Transactions on Signal Processing, 2014, 62(3):531-544.
- [14] WANG Jun, WANG Wenchuan, HU Xiaoxue, et al. Black-winged Kite Algorithm: a Nature-inspired Meta-heuristic for Solving Benchmark Functions and Engineering Problems [J]. Artificial Intelligence Review, 2024, 57(4):98.
- [15] GUO Xin, TU Jiesong, ZHAN Shengpeng, et al. A Novel Method for Online Prediction of the Remaining Useful Life of Rolling Bearings Based on Wavelet Power Spectrogram and Transformer Structure [J]. Engineering Research Express, 2023, 5(4):045074.
- [16] KILLICK R, FEARNHEAD P, ECKLEY I A. Optimal Detection of Change-points with a Linear Computational Cost [J]. Journal of the American Statistical Association, 2012, 107(500):1590-

- 1598.
- [17] 左栋, 黄钊, 陈明, 等. 基于格拉姆角场与改进图卷积网络的滚动轴承故障诊断[J]. 装备制造技术, 2023(3):48-50.
ZUO Dong, HUANG Zhao, CHEN Ming, et al. Fault Diagnosis of Rolling Bearing Based on Gram Angular Field and Improved Graph Convolution Network[J]. Equipment Manufacturing Technology, 2023(3):48-50.
- [18] 段礼祥, 李涛, 唐瑜, 等. 基于多源异构信息融合的机械故障诊断方法[J]. 石油机械, 2021, 49(2):60-67.
DUAN Lixiang, LI Tao, TANG Yu, et al. Mechanical Fault Diagnosis Method Based on Multi-source Heterogeneous Information Fusion[J]. China Petroleum Machinery, 2021, 49(2):60-67.
- [19] 高艺源, 于德介, 王好将, 等. 基于图谱指标的滚动轴承故障特征提取方法[J]. 航空动力学报, 2018, 33(8):2033-2040.
GAO Yiyuan, YU Dejie, WANG Haojiang, et al. Fault Feature Extraction Method of Rolling Bearing Based on Spectral Graph Indices[J]. Journal of Aerospace Power, 2018, 33(8):2033-2040.
- [20] WANG Biao, LEI Yaguo, LI Naipeng, et al. A Hybrid Prognostics Approach for Estimating Remaining Useful Life of Rolling Element Bearings [J]. IEEE Transactions on Reliability, 2020, 69(1):401-412.
- [21] 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[J]. 机械工程学报, 2019, 55(16):1-6.
LEI Yaguo, HAN Tianyu, WANG Biao, et al. XJTU-SY Rolling Element Bearing Accelerated Life Test Datasets: a Tutorial[J]. Journal of Mechanical Engineering, 2019, 55(16):1-6.
- [22] 宋浏阳, 金焯, 郭旭东, 等. 基于自适应权重时间卷积网络的剩余使用寿命预测方法[J]. 北京化工大学学报(自然科学版), 2024, 51(3):76-87.
SONG Liuyang, JIN Ye, GUO Xudong, et al. Remaining Useful Life Prediction Based on an Adaptive Weight Temporal Convolutional Network[J]. Journal of Beijing University of Chemical Technol-
ogy (Natural Science Edition), 2024, 51(3):76-87.
- [23] LI Tianfu, ZHOU Zheng, LI Sinan, et al. The Emerging Graph Neural Networks for Intelligent Fault Diagnostics and Prognostics: a Guideline and a Benchmark Study[J]. Mechanical Systems and Signal Processing, 2022, 168:108653.
- [24] ZHOU Jianzhong, SHAN Yahui, LIU Jie, 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.
- [25] 杨超颖, 刘颀, 周凯波. 基于路图注意力网络的轴承剩余寿命预测方法[J]. 机械工程学报, 2023, 59(12):195-201.
YANG Chaoying, LIU Jie, ZHOU Kaibo. Path Graph Attention Network-based Bearing Remaining Useful Life Prediction Method[J]. Journal of Mechanical Engineering, 2023, 59(12):195-201.
- [26] 王朋飞, 刘长良, 徐健, 等. 基于图注意力和时间卷积网络的风电齿轮箱故障预警方法[J]. 电子测量与仪器学报, 2023, 37(8):204-213.
WANG Pengfei, LIU Changliang, XU Jian, et al. Wind Turbine Gearbox Fault Warning Method Based on Graph Attention and Temporal Convolutional Network[J]. Journal of Electronic Measurement and Instrumentation, 2023, 37(8):204-213.

(编辑 张 洋)

作者简介: 宋李俊, 男, 1974 年生, 副教授, 硕士研究生导师。研究方向生产调度管理、故障诊断。发表论文 50 余篇。E-mail: sgljn@163.com。辛 玉* (通信作者), 男, 1987 年生, 讲师, 硕士研究生导师。研究方向为机械装备状态监测与智能运维。发表论文 13 篇。E-mail: yu_xin@cqut.edu.cn。

本文引用格式:

宋李俊, 刘松林, 辛玉, 等. 基于轴承退化状态评估和改进图注意力双向门控循环单元网络的[J]. 中国机械工程, 2025, 36(7): 1562-1572.

SONG Lijun, LIU Songlin, XIN Yu, et al. Residual Life Prediction for Bearings Based on Bearing Degradation State Assessment and IGAT-BiGRU Network[J]. China Mechanical Engineering, 2025, 36(7):1562-1572.