双向建模增强 TKAN 和全局注意力机制融合的滚动轴承剩余寿命预测

姜 蕾 [1] 郑建飞 [1] 胡昌华 [1] 赵瑞星 [2] 韩其辉 [1] 杨立浩 [1]

摘 要 滚动轴承剩余使用寿命 (RUL) 的精准预测是确保设备或系统安全可靠运行的关键. 针对滚动轴承 RUL 预测中多维退化特征的长期依赖关系难以有效建模的问题, 提出一种双向时间序列建模与注意力机制融合的预测模型——双向时序科尔莫戈洛夫–阿诺尔德注意力网络 (Bi-TKAN-Att). 该模型兼具了时序科尔莫戈洛夫–阿诺尔德网络的强时序建模能力和全局注意力机制的关键特征提取能力, 采用双向建模的方式捕捉前后向信息, 最终实现了具有长期依赖多维退化特征的滚动轴承 RUL 预测. 所提方法在滚动轴承数据集上进行实验验证, 结果表明 Bi-TKAN-Att 模型在捕获滚动轴承退化特性和提升 RUL 预测精度方面具有显著优势, 并通过消融实验证明了模型各组件的合理性和有效性, 为滚动轴承的寿命预测提供了全新可行的解决方案.

关键词 剩余使用寿命预测, 滚动轴承, 时序科尔莫戈洛夫–阿诺尔德网络, 双向建模, 全局注意力机制

引用格式 姜蕾, 郑建飞, 胡昌华, 赵瑞星, 韩其辉, 杨立浩. 双向建模增强 TKAN 和全局注意力机制融合的滚动轴承剩余寿命预测. 自动化学报, 2025, **51**(8): 1857−1868

DOI 10.16383/j.aas.c250064 **CSTR** 32138.14.j.aas.c250064

Bidirectional Modeling-enhanced TKAN and Global Attention Mechanism Fusion for Rolling Bearing Remaining Useful Life Prediction

JIANG Lei[1] ZHENG Jian-Fei[1] HU Chang-Hua[1] ZHAO Rui-Xing[2] HAN Qi-Hui[1] YANG Li-Hao[1]

Abstract The accurate prediction of the remaining useful life (RUL) of rolling bearings is crucial for ensuring the safe and reliable operation of equipment or system. To address the challenge of effectively modeling the long-term dependencies of multi-dimensional degradation features in RUL prediction, this paper proposes a prediction model that integrates bidirectional time-series modeling and attention mechanisms——Bidirectional temporal Kolmogorov-Arnold attention networks (Bi-TKAN-Att). This model combines the powerful time-series modeling capability of temporal Kolmogorov-Arnold networks with the key feature extraction capability of global attention mechanisms. It uses a bidirectional modeling approach to capture both forward and backward information, ultimately achieving RUL prediction for rolling bearings with long-term dependent multi-dimensional degradation features. The proposed method is experimentally validated on the rolling bearing dataset. Experimental results demonstrate that the Bi-TKAN-Att model has significant advantages in capturing rolling bearing degradation features and improving RUL prediction accuracy. Ablation experiments further confirm the rationality and effectiveness of each component of the model, providing a novel and feasible solution for the lifetime prediction of rolling bearing.

Key words Remaining useful life prediction, rolling bearing, temporal Kolmogorov-Arnold networks, bidirectional modeling, global attention mechanism

Citation Jiang Lei, Zheng Jian-Fei, Hu Chang-Hua, Zhao Rui-Xing, Han Qi-Hui, Yang Li-Hao. Bidirectional modeling-enhanced TKAN and global attention mechanism fusion for rolling bearing remaining useful life prediction. *Acta Automatica Sinica*, 2025, **51**(8): 1857−1868

作为旋转机械的核心支撑部件, 滚动轴承被广泛应用于风力发电机、航空发动机和内燃机等各类旋转机械设备中[1]. 滚动轴承在复杂多变的工作环境中, 容易因磨损、腐蚀和变形等问题发生故障, 影响机械设备的精度、寿命和可靠性. 通过精确的状态监测, 可以有效减少故障、提升生产效率. 因此, 对滚动轴承的剩余使用寿命 (Remaining useful life,

收稿日期 2025-02-24 录用日期 2025-06-09
Manuscript received February 24, 2025; accepted June 9, 2025
国家重大科研仪器研制项目 (62227814), 国家自然科学基金 (62233017, 62373369) 资助
Supported by National Major Scientific Research Instrument Development Project (62227814) and National Natural Science Foundation of China (62233017, 62373369)
本文责任编委 唐漾
Recommended by Associate Editor TANG Yang
1. 火箭军工程大学导弹工程学院 西安 710025 2. 火箭军装备部驻成都地区某军事代表室 成都 610000
1. College of Missile Engineering, Rocket Force University of Engineering, Xi′an 710025 2. Military Representative Office of the Rocket Force Equipment Department Stationed in Chengdu, Chengdu 610000

RUL) 进行精准预测是确保其所在设备或系统安全可靠运行的关键, 也是对滚动轴承进行科学健康管理的前提.

滚动轴承 RUL 预测方法主要分为两类: 基于模型的方法[2] 和数据驱动的方法[3]. 基于模型的预测方法通过利用已有知识与系统工作状态建立物理模型, 对机械设备的 RUL 进行预测[4], 例如 Paris-Erdogan 模型和指数退化模型[5]. 然而, 随着机械结构的日益复杂, 在不同运行条件下为特定滚动轴承构建精确的机理模型变得极为困难, 导致预测结果因适应性较差而不够准确[6]. 数据驱动的预测方法无需构建特定模型, 从而在预测任务中逐渐展现出显著优势. 尤其是基于深度学习的数据驱动方法, 由于可以避免对物理失效机制的研究或机械退化过程模型的构建, 能够自适应地建立监测数据与 RUL 之间的非线性映射关系, 已广泛应用于 RUL 预测. 因此, 基于深度学习的 RUL 预测方法逐渐成为研究的热点[7–8].

常用于 RUL 任务中的深度学习方法包括卷积神经网络 (Convolutional neural network, CNN)[9]、循环神经网络 (Recurrent neural network, RNN)[10]、长短期记忆网络 (Long short term memory, LSTM)[11]、门控循环单元 (Gated recurrent unit, GRU)[12]、时间卷积网络 (Temporal convolutional network, TCN)[13] 等. 在 RUL 预测的多数场景中, 时间序列是最常见的监测数据形式, 而 CNN 虽然在提取局部特征方面具有显著优势, 但在处理时间序列数据时, 其对长距离依赖关系的捕捉能力较为有限. 这对需要解析长期趋势和模式的 RUL 预测任务造成了一定限制, 进而影响预测精度的进一步提升. 为克服这些缺陷, RNN 及其变种通过引入循环机制与记忆单元, 使之更适合处理各种时间序列任务. Zhou 等[14] 提出一种改进型的增强记忆 GRU 网络, 该网络针对 RNN 预测模型易遗忘历史趋势信息的问题, 通过重复利用前一时刻的状态信息来提高预测性能. Zhao 等[15] 提出一种融合 CNN 和 LSTM 的 RUL 预测模型. 该模型利用 CNN 自动提取矩阵参数中的局部特征和空间信息, 避免了手动特征提取的复杂性, 并通过 LSTM 学习与剩余使用寿命相关的全局依赖关系, 有效提高了预测精度. 然而, 在滚动轴承的 RUL 预测任务中, 轴承的退化过程通常涉及复杂的长期依赖关系, 仅使用单向网络容易导致信息丢失, 不能充分考虑过去和未来的信息.

为解决上述问题, 康守强等[16] 引入双向长短时记忆网络 (Bi-directional long short-term memory, Bi-LSTM) 实现滚动轴承的 RUL 预测, 获得了更准确的预测结果. 夏然等[17] 提出一种融合健康因子筛选与混合神经网络的锂离子电池剩余寿命预测方法. 该方法通过随机森林算法评估并筛选多种健康因子, 利用遗传算法优化的广义回归神经网络模型进行容量估计, 并构建 Bi-LSTM 与非线性自回归混合模型进行残差修正和寿命预测, 显著提高了预测的精度与稳定性. 双向建模的引入在一定程度上推动了 RUL 预测精度的提升. 然而, 上述研究均基于传统的多层感知机 (Multilayer perceptron, MLP) 架构, 存在模型灵活性不足的问题. 当面对复杂的非线性关系和时间序列中各元素之间长期依赖性的全局学习问题时, 传统的时间序列预测模型采用固定激活函数, 尽管部分模型使用了可变激活函数, 但这些激活函数的选择依然是有限制的, 无法完全自适应学习, 限制了模型的拟合能力, 从而影响滚动轴承 RUL 预测的精度.

近期出现的科尔莫戈洛夫–阿诺尔德网络 (Kolmogorov-Arnold networks, KAN) 是一种受科尔莫戈洛夫–阿诺尔德表示定理启发的神经网络架构. KAN 将传统 MLP 中的权重矩阵替换为可学习的单变量函数, 从而将固定激活函数替换为可变激活函数. KAN 在多项科学建模任务中, 达到了更优的精度[18]. 在 KAN 基础上, Genet 等[19] 增加一个外部记忆模块, 得到循环科尔莫戈洛夫–阿诺尔德网络 (Recurrent Kolmogorov-Arnold network, RKAN), 之后进一步增加了门控机制, 提出一种更适合时间序列预测的网络, 即时序科尔莫戈洛夫–阿诺尔德网络 (Temporal Kolmogorov-Arnold networks, TKAN). 但 TKAN 为单向网络, 双向结构网络具有更全面的上下文感知能力, 在一定程度上可以缓解单向模型对长期依赖关系捕捉不良的问题. 为进一步增强长期依赖特征的全局捕捉能力, 注意力机制逐渐受到关注. He 等[20] 提出一种基于改进深度脉冲残差神经网络的预测方法, 用于轨道 RUL 的估计. 该方法通过把多通道振动数据直接作为输入, 利用分离式脉冲卷积编码模块将时间序列数据自适应地转换为脉冲信号, 同时结合注意力机制对脉冲特征图进行重新标定, 并通过改进的脉冲残差连接减轻梯度消失和爆炸问题, 从而提高了复杂时空关系的建模能力与预测性能.

综上所述, 本文提出一种融合双向 TKAN 和注意力机制的滚动轴承 RUL 预测方法——双向时序科尔莫戈洛夫–阿诺尔德注意力网络 (Bidirectional temporal Kolmogorov-Arnold attention

networks, Bi-TKAN-Att). 该方法使用具有时间依赖性更新机制的 TKAN 作为单向网络, 利用其具有可学习激活函数的特殊架构, 显著提升了预测模型的建模能力, 双向建模的引入明显增强了网络对长时间序列的处理能力, 通过进一步融入全局注意力机制增强模型对于局部关键特征的捕获能力, 长短期特征得到了综合表达, 实现了滚动轴承 RUL 的精准预测.

1　Bi-TKAN-Att 网络结构

1.1　原理介绍

KAN 本质上为静态网络, 难以直接对时序数据中的动态依赖关系建模. 为此, RKAN 通过引入递归机制, 将 KAN 的层间连接扩展为时间依赖的隐藏状态传递, 使其能够捕捉序列的短期记忆特征. 随着序列长度的增加, RKAN 仍面临长期依赖建模的挑战. TKAN 通过在 RKAN 中嵌入门控机制实现了信息流的长期记忆与更新. 在此基础上, 本文进一步提出 Bi-TKAN-Att 网络, 通过对 TKAN 结构进行双向化以同时整合前向与后向上下文, 通过叠加注意力层关注关键时间片段, 实现了长期依赖关系的捕捉. 下面将依次介绍 RKAN、TKAN 的具体架构.

1.1.1　RKAN

在 RNN 处理时间序列数据时, 递归权重矩阵用于对隐藏状态进行转换. 该操作通常会经过加法步骤, 转换结果再通过一个非线性激活函数进行处理. 更新过程可表示为:

$$h_t = f(W_{hh}h_{t-1} + W_{hx}x_t + b_h) \tag{1}$$

其中, $f(\cdot)$ 是非线性激活函数; h_t 是时间步 t 的隐藏状态; W_{hh} 和 W_{hx} 分别是前一时刻隐藏状态 h_{t-1} 和当前输入的权重矩阵; x_t 是 t 时刻的输入; b_h 是偏置项. 这种结构使得 RNN 能够在处理时间序列时捕捉到时序信息, 并能够通过不断迭代来学习长期依赖关系. 为更好地处理时间依赖性, RKAN 提出一种针对 KAN 的全新更新机制. 该机制通过将前一时间步的隐藏状态整合到当前状态中, 保留了历史输入的记忆, 从而使得网络能够展现动态的时间特性. 具体来说, 在 RKAN 的每一层中, 变换函数被设计为具有时间依赖性的函数 $\phi_{l,\,j,\,i,\,t}$, 从而使得每一层的输出不仅受当前输入的影响, 还会受到历史状态的影响. 具体地, 第 $l+1$ 层的第 j 个节点在 t 时刻的输出 $x_{l+1,\,j}(t)$ 可以表示为:

$$x_{l+1,\,j}(t) = \sum_{i=1}^{n_l} \tilde{x}_{l,\,j,\,i}(t) = \sum_{i=1}^{n_l} \phi_{l,\,j,\,i,\,t}\left(x_{l,\,i}(t),\ h_{l,\,i}(t)\right) \tag{2}$$

其中, $\tilde{x}_{l,\,j,\,i}(t)$ 是第 l 层的第 i 个节点在时间步 t 的计算输出. 在此框架下, “记忆”步骤 $h_{l,\,i}(t)$ 被定义为过去隐藏状态的加权组合, 具体形式为:

$$h_{l,\,i}(t) = W_{hh}h_{l,\,i}(t-1) + W_{hz}x_{l,\,i}(t) \tag{3}$$

其中, W_{hz} 为输入权重矩阵. 在 RKAN 中, 网络在每一层都嵌入了记忆管理机制, 通过下式进行计算:

$$\mathrm{KAN}(x,\ t) = (\Phi_{L-1,\,t} \circ \Phi_{L-2,\,t} \circ \cdots \circ \Phi_{0,\,t})\,(x,\ t) \tag{4}$$

其中, $\mathrm{KAN}(x,\ t)$ 为 KAN 模型的结构公式; $\Phi_{L-1,\,t}$ 为 KAN 结构的第 $L-1$ 层; $\circ$ 表示函数的复合运算. 这种机制使得每一层的计算不仅依赖于当前的输入数据, 还结合了历史的状态信息, 从而有效捕捉时间序列数据的动态性和时间依赖性.

1.1.2　TKAN

LSTM 是一种特殊的循环神经网络结构, 旨在解决 RNN 在处理长序列时的梯度消失问题. LSTM 通过门控机制, 能够有选择性地保留或遗忘信息. 其核心包括三个主要的门控机制: 遗忘门、输入门和输出门. TKAN 使用了与 LSTM 相类似的门控机制来控制信息的流动, 其结构图如图 1 所示.

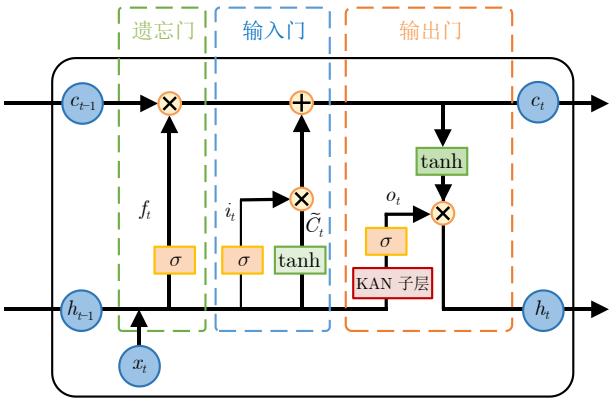


图 1　TKAN 结构图

Fig. 1　TKAN structure diagram

首先遗忘门控制从先前状态 h_{t-1} 中需要遗忘的信息量, 其激活向量 f_t 的计算公式为:

$$f_t = \sigma(w_f\,[h_{t-1},\ x_t] + b_f) \tag{5}$$

输入门决定了当前时刻需要引入的新信息, 其激活向量 i_t 的计算公式为:

$$i_t = \sigma(w_i\,[h_{t-1},\ x_t] + b_i) \tag{6}$$

式中, w_f, w_i 分别表示遗忘门、输入门的权重矩阵; b_f, b_i 分别表示遗忘门、输入门的偏置; σ 是 sigmoid 激活函数.

输出门决定了当前时刻记忆状态中需要输出的信息, 其激活向量 o_t 的计算公式为:

$$o_t = \sigma(\mathrm{KAN}(x_t,\ t)) \tag{7}$$

当前的记忆状态 c_t 综合了遗忘门和输入门的作用, 遗忘门控制上一时刻的单元状态 c_{t-1} 传递到当前时刻的单元状态, 输入门决定当前时刻候选状态 $\tilde{C}_t$ 中保留多少信息.

$$c_t = f_t \otimes c_{t-1} + i_t \otimes \tilde{C}_t \tag{8}$$

最终输出向量 h_t 定义如下:

$$h_t = o_t \odot \tanh(c_t) \tag{9}$$

其中, $\otimes$ 表示矩阵乘法; $\odot$ 表示矩阵点乘.

1.2 Bi-TKAN 结构

TKAN 是一个单向模型, 这意味着在处理当前时间步的信息时, TKAN 只能利用当前时间步之前的信息, 而不考虑后续的输入信息. 当上下文依赖关系很长或很复杂时, 尤其是在当前时间步需要依赖未来信息的情况下, 这种单向处理可能会导致获取信息的不完整, 因此单向 TKAN 需要新的结构来获得更全面的输入信息. 为解决该问题, 本文提出一种新的时间序列预测网络: 双向时序科尔莫戈洛夫–阿诺尔德网络 (Bidirectional TKAN, Bi-TKAN). Bi-TKAN 包含两个独立的 TKAN 隐层, 其结构如图 2 所示.

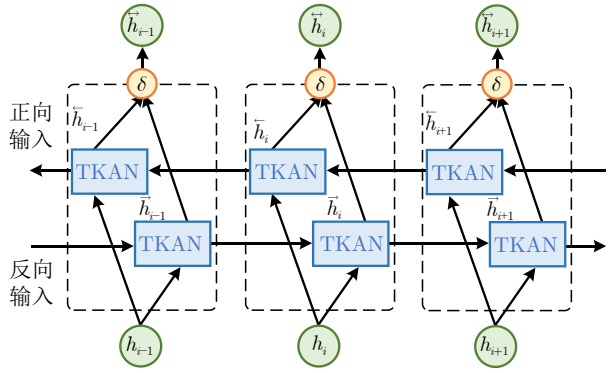


图 2 Bi-TKAN 结构图

Fig. 2 Bi-TKAN structure diagram

TKAN 层的所有隐藏状态分别按时间顺序和逆序排列, 从而得到顺序序列 $[h_1,\ h_2,\ \cdots,\ h_n]$ 及逆序序列 $[h_n,\ h_{n-1},\ \cdots,\ h_1]$, 将这两个序列输入两个独立 TKAN 中. 通过这种双向建模方式, 可以有效地捕捉特征数据的前向和后向依赖信息. 图 2 中 δ 是连接前向输入和逆向输入的函数, Bi-TKAN 在 i 时刻的输出 $\overleftrightarrow{h}_i$ 可以表示为:

$$\overleftrightarrow{h}_i = \delta(\overrightarrow{h}_i,\ \overleftarrow{h}_i) \tag{10}$$

其中, $\overrightarrow{h}_i$ 表示 i 时刻正向 TKAN 隐藏状态, $\overleftarrow{h}_i$ 表示反向 TKAN 隐藏状态.

1.3 Bi-TKAN-Att 结构

注意力机制是一种能够自适应调整网络中不同信息重要性权重的算法, 在时间序列建模任务中具有重要应用. 在时间序列预测任务中, 来自历史输入中不同时刻的数据通常具有不同的显著性和重要性. 然而所提出的 Bi-TKAN 模型并未考虑这一点, 即历史输入在不同时间步长上的权重可能存在差异, 而这些差异可能对提升模型性能具有关键作用. 因此, 在 Bi-TKAN 后端引入注意力层, 该层用于识别模型中输入特征的重要性, 并为重要信息分配更高的权重因子, 使模型能够专注于关键的输入数据, 从而提高准确性, 如图 3 所示.

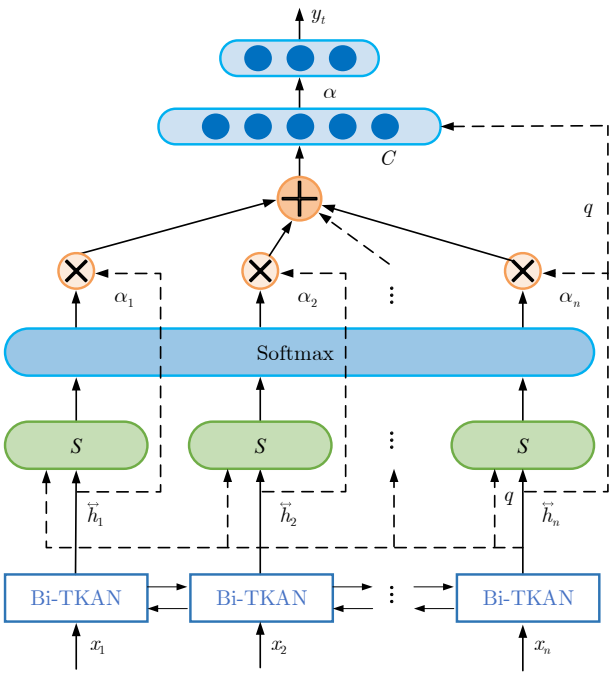


图 3 注意力机制结构图

Fig. 3 Structure diagram of attention mechanism

图 3 中, $\overleftrightarrow{h}_i$ 为 Bi-TKAN 的输出, $i=1,\ 2,\ \cdots,\ n$; S 是注意力评分函数; q 是查询向量, $q = \overleftrightarrow{h}_n$; α_i 是 Bi-TKAN 输出的注意力权重; C 是 Bi-TKAN 输出的加权平均. 为充分利用 Bi-TKAN 的隐藏状态

信息, C 和 $\overleftrightarrow{h}_n$ 被组合后输入一个激活函数为 tanh 的全连接神经网络, 最后, α 被输入全连接网络中, 得到 RUL 预测值 y_t, 具体公式如下:

$$y_t = W_t\alpha \tag{11}$$

$$\alpha = f(C,\ \overleftrightarrow{h}_i) = \tanh(W_c[C;\ \overleftrightarrow{h}_i]) \tag{12}$$

$$C = \sum_{i=1}^{n} \alpha_i \overleftrightarrow{h}_i \tag{13}$$

$$\alpha_i = \frac{\exp(S(\overleftrightarrow{h}_i,\ q))}{\sum\limits_{j=1}^{n} \exp(S(\overleftrightarrow{h}_j,\ q))} \tag{14}$$

$$S(\overleftrightarrow{h}_i,\ q) = \overleftrightarrow{h}_i^{\mathrm{T}} W q \tag{15}$$

其中, W、W_c 和 W_t 是该网络的权重.

本文提出一种全新的时间序列预测模型 Bi-TKAN-Att.该模型在 TKAN 的基础上融合了双向时间序列建模以及注意力机制, 通过整合正向和逆向的时间序列特征, 全面捕捉数据的前后依赖关系以及全局信息. Bi-TKAN-Att 的结构如图 4 所示, 其核心结构包括两个独立的 TKAN 层, 用于捕捉输入序列的前向和后向信息; 一个注意力层, 用于分配输入特征的重要性权重; 一个全连接层, 用于生成最终预测结果. 模型的设计充分利用了 TKAN 的非线性拟合能力以及其在时间序列预测上的优势, 结合了双向时序网络的建模能力和注意力机制, 使得 Bi-TKAN-Att 能够在复杂的时间序列数据中高效提取关键退化信息, 并显著提高预测性能.

2　基于 Bi-TKAN-Att 的时间序列预测框架

本节主要介绍所提 RUL 预测方法的框架, 该框架主要包含两个部分: 第一部分是对退化特征进行提取和筛选; 第二部分是基于 Bi-TKAN-Att 对滚动轴承的 RUL 进行预测. 所提方法的预测流程如图 5 所示.

2.1　退化特征提取

滚动轴承原始监测数据量较大且包含冗余信息, 这增加了精确退化建模的复杂性. 因此, 需要从原始振动信号中提取退化特征, 以实现数据降维. 首先提取原始振动信号常见的时域特征, 如峰度、峰值等, 然后提取振动信号的频域特征. 通过将振动信号从时域转换到频域, 频谱数据可以沿时间维度排列, 同时呈现出退化趋势. 对于非周期信号 $x(t)$, 采用傅里叶变换表示为:

$$X(f) = \int_{-\infty}^{+\infty} x(t)\mathrm{e}^{-\mathrm{j}2\pi ft}\mathrm{d}t \tag{16}$$

其中, $X(f)$ 表示频域信号, 通过傅里叶变换从时域信号 $x(t)$ 得到; j 是虚数单位; f 表示频率.

从转换后的信号中提取多种频域特征参数, 如中心频率等. 最后, 为了分析非平稳信号的频谱随时间的变化规律, 使用变分模态分解 (Variational mode decomposition, VMD) 对时频域特征进行提取. 传统的经验模态分解 (Empirical mode decomposition, EMD) 存在端点效应和模态分量混叠问题. 作为一种自适应的完全非递归方法, VMD 可用

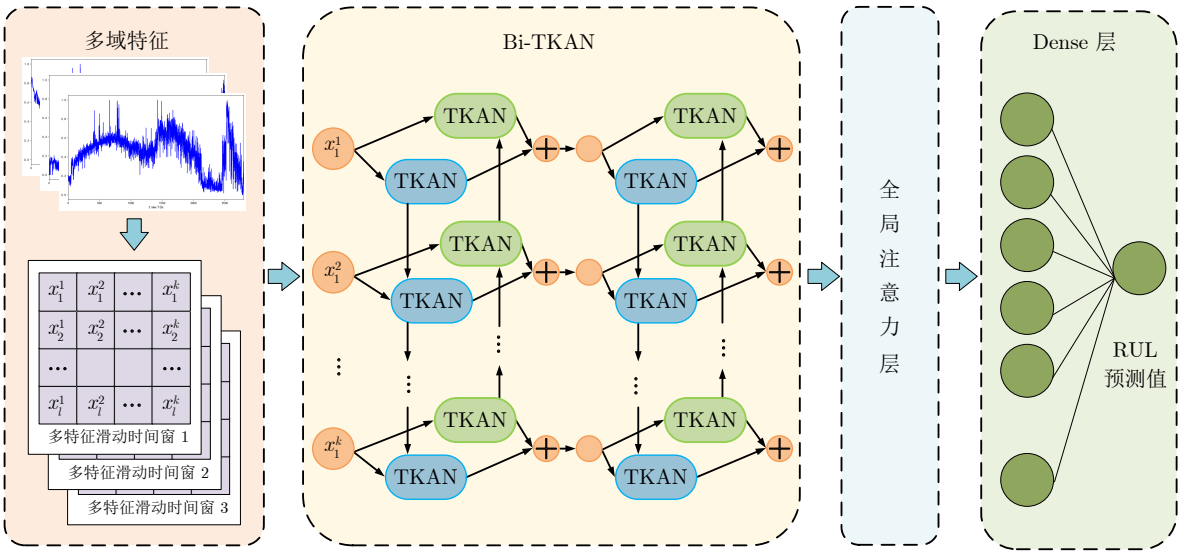


图 4　Bi-TKAN-Att 结构图

Fig. 4　Bi-TKAN-Att structure diagram

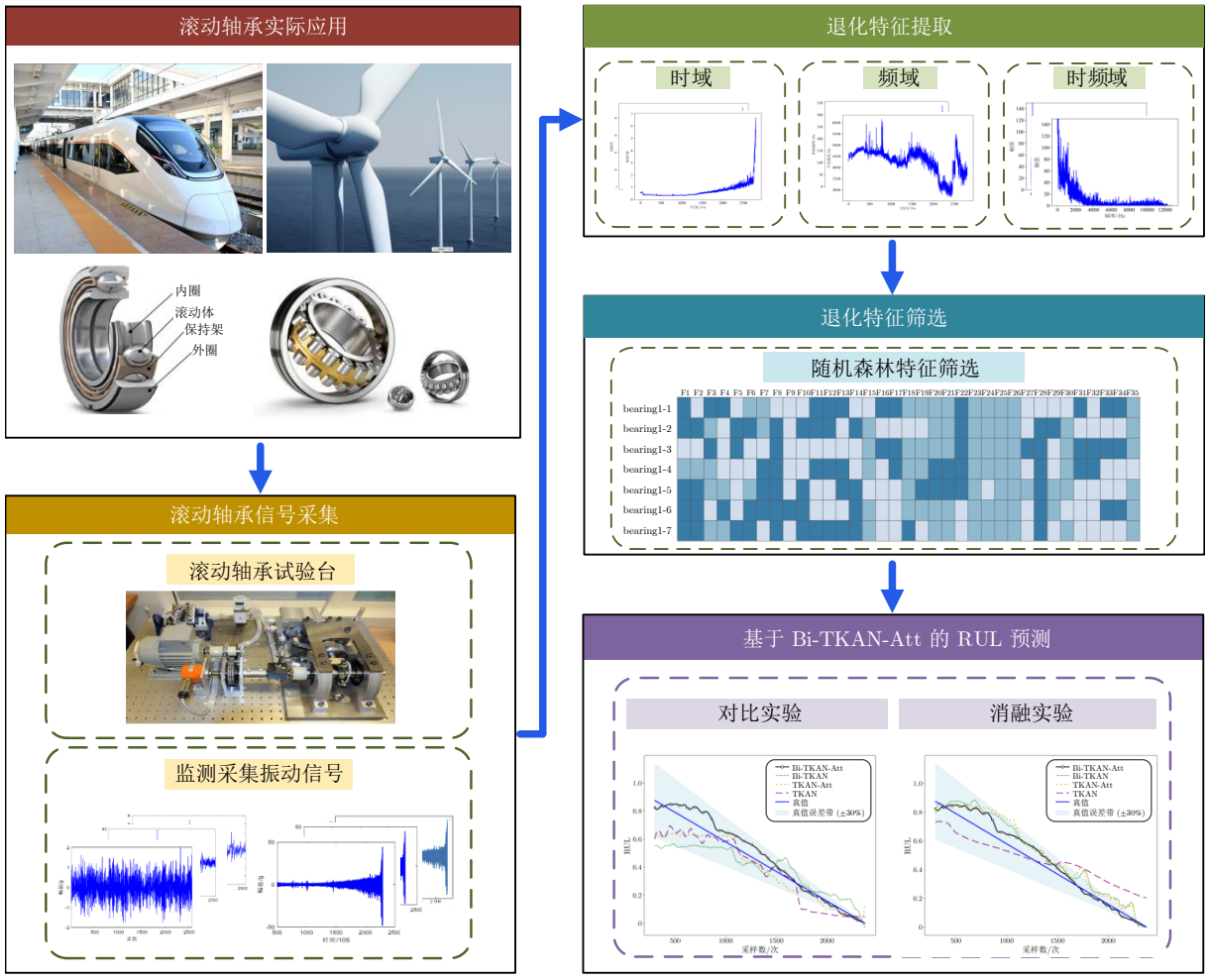


图 5 基于 Bi-TKAN-Att 的时间序列预测框架

Fig. 5 Time-series prediction framework based on Bi-TKAN-Att

于模态变化和信号处理. VMD 减少了具有高复杂度和强非线性时间序列的非平稳性, 并将其分解为包含多个不同频率尺度的相对平稳的子序列. 在 VMD 分解中, 每个模态表示为 $v_k(t)$. VMD 的目标是最小化各模态的估计带宽之和, 并满足各模态之和等于原始数据的约束条件, 数学形式如下:

$$\min\left\{\sum_k\left\|\frac{\partial\left(\left(\delta(t)+\frac{\mathrm{j}}{\pi t}\right)*v_k(t)\right)}{\partial t}\times \mathrm{e}^{-\mathrm{j}\omega_k t}\right\|_2^2\right\}$$
$$\text{s.t.}\quad \sum_k v_k=f(t) \tag{17}$$

其中, v_k 表示分解得到的各模态分量; ω_k 表示各模态分量对应的中心频率; $*$ 表示卷积操作; $\delta(t)$ 为狄拉克函数; k 表示分解得到的模态总数; $f(t)$ 是原始信号. 通过 VMD 将振动信号分解为一系列模态分量, 有效提取不同频率尺度下的信号特征. 能量熵作为一种反映信号能量分布复杂程度的指标, 能够量化各模态分量的能量特征. 当滚动轴承发生故障时, 振动信号的频率分布会发生变化, 各模态分量的能量值也会随之变化. 结合能量熵和样本熵对各模态信号分量进行量化, 能够在保留信号重要退化信息的同时实现数据降维和关键特征提取.

2.2 退化特征筛选

在滚动轴承剩余使用寿命预测中, 特征筛选是提高模型性能和减少计算复杂度的关键步骤. 它能有效去除冗余和无关特征, 减小噪声干扰. 同时, 保留重要特征有助于增强模型的泛化能力, 提升预测精度, 从而为滚动轴承寿命预测提供更可靠的支持. 本文使用随机森林对特征重要性进行评估, 随机森林的特征重要性指标基于决策树节点的分裂效果进行计算. 在训练过程中, 每次节点分裂都会带来目标变量预测误差的减少. 特征重要性即为某一特征在所有树中分裂时所带来的误差减少的总和. 通过对所有特征的重要性进行归一化处理, 可以得到每

个特征的相对重要性.

2.3 RUL 预测

本文采用 Bi-TKAN 来捕获时间序列中的复杂依赖关系, 结合注意力机制进一步提升模型的预测能力. 具体步骤如下:

步骤 1. 对输入的多特征数据进行标准化, 并采用滑动时间窗技术进行预处理. 预处理后的数据输入 Bi-TKAN 模块.

步骤 2. Bi-TKAN 模块由两个独立的 TKAN 网络组成, 分别处理时间序列的正向和反向信息. 正向 TKAN 网络提取从起始时间步到当前时间步的累积时序特征, 能够捕获序列中基于过去信息的依赖关系; 反向 TKAN 网络提取从当前时间步到起始时间步的时序特征, 能够结合未来信息提高序列上下文的信息完整性. 通过这种双向结构, Bi-TKAN 可以综合利用正反两个方向的时间序列特征, 从而更有效地建模复杂的依赖关系, 包括短期和长期的动态变化.

步骤 3. 本文在 Bi-TKAN 模块后引入了全局注意力机制. 注意力模块通过为不同时间步分配权重, 动态关注对当前预测任务最重要的时序特征. 注意力机制使得模型能够突出重要的时序特征, 同时抑制对预测无关的冗余信息, 从而显著提高了预测的准确性.

步骤 4. 综合处理后的时序特征被输入全连接层. 在全连接层中, 通过加权和偏置的作用对特征进行进一步处理, 最终输出 RUL 预测值.

3 实验验证

3.1 数据集描述

本文使用 IEEE PHM 2012 挑战赛数据集来验证所提方法在滚动轴承 RUL 预测中的性能. 该数据集由 PRONOSTIA 平台采集, 是滚动轴承退化预测领域中广泛应用的数据集. 该数据集通过模拟滚动轴承在不同工作条件下的运行状态, 提供了全面的退化数据, 为开发和验证预测模型提供了重要支持. 数据集包含三个不同的工作条件: 工作条件 1 设定为电机转速 1 800 r/min、负载 4 000 N; 工作条件 2 设定为电机转速 1 650 r/min、负载 4 200 N; 工作条件 3 设定为电机转速 1 500 r/min, 负载 5 000 N. 在实验中, 振动信号以 25.6 kHz 的采样频率采集, 每次采样时长为 0.1 s, 共记录 2 560 个数据点, 并分别从水平和垂直方向采集信号[21]. 本文选取工作条件 1 下的 3 组数据进行 RUL 预测.

3.2 性能评价指标

为全面验证所提预测方法的优越性, 引入 3 种常用的回归评价指标对 RUL 预测结果进行评估, 分别为平均绝对误差 (Mean absolute error, MAE)、均方根误差 (Root mean square error, RMSE)、决定系数 (R-square, R^2), 具体定义如下:

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \tag{18}$$

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

$$\mathrm{R}^2 = 1 - \frac{\sum\limits_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum\limits_{i=1}^{n}(y_i - \bar{y}_i)^2} \tag{20}$$

式中, y_i, $\hat{y}_i$ 分别为第 i 时刻的实际 RUL 值和预测 RUL 值; $\bar{y}_i$ 为实际 RUL 的均值; n 为测试样本的数量.

3.3 实验设置

Bi-TKAN 模块由两层组成, 其隐藏神经元数分别设置为 32 和 40. B 样条函数在 TKAN 中作为一种激活函数, 用于实现更灵活的非线性变换, 它的超参数主要控制样条的平滑程度、分段细化程度等. 在本实验中使用的样条阶数为 3, 样条的网格分段数设置为 10, 多头注意力的头数设置为 4. 实验过程中, 4 个注意力头的设置在性能和计算效率之间达到了较好的平衡, 满足了实验任务对建模能力的需求. 为防止模型过拟合, 批处理大小设置为 128, Dropout 正则化参数设定为 0.4. 学习率设为 0.0008, 以确保模型训练过程的稳定性和效率. 滑动时间窗大小设置为 300. 具体配置详见表 1.

本实验在一台高性能工作站上进行, 工作站配置包括 64 GB 内存的 Intel(R) Xeon(R) Gold 6226R 处理器和一块显存容量为 16 GB 的 NVIDIA RTX A4000 显卡. 实验代码使用 Python 语言, 开发环境为 PyCharm. 模型的训练以及优化基于 TensorFlow 框架完成.

3.4 特征提取和筛选

为全面分析和描述滚动轴承退化过程, 本文从滚动轴承振动信号中提取 12 个时域特征、9 个频域特征和 14 个时频域特征, 构建了 35 维的特征集. 这些特征从不同角度表征了滚动轴承的健康状态, 有效弥补了单一特征可能带来的局限性. 然而, 并

表 1 Bi-TKAN-Att 的结构参数

Table 1 Structural parameters of Bi-TKAN-Att

网络结构超参数	最优值
Bi-TKAN 层	2
Bi-TKAN 层 1	32
Bi-TKAN 层 2	40
全连接层	(40, 1)
批次大小	128
训练周期	100
学习率	0.0008
滑动窗口大小	300

非所有特征对 RUL 预测都具有很高的重要性, 部分特征可能存在冗余或噪声. 为此, 本文使用前述随机森林算法对 35 维特征进行筛选, 通过特征筛选不仅能够降低计算复杂度, 还可以剔除不相关特征, 减少模型过拟合的风险. 将 bearing1-1 到 bearing1-7 的特征重要性进行排序, 特征的颜色深浅反映了它们在 RUL 预测中的重要性. 颜色越深的特征表示其对 RUL 预测的影响越大, 反之则表示其影响越小.

由图 6 可见, 不同滚动轴承的相同特征在 RUL 预测中的重要性存在差异. 本文特征的选择是基于这些颜色深浅的分布, 选取在图 6 中占据深色色块最多的特征, 这些特征在多个轴承数据中具有较高的重要性. 本实验将 bearing1-1、bearing1-2 的数据用作训练集, bearing1-3 的数据用作测试集. 最终筛选出工作条件 1 下的敏感特征为: 模态 2 能谱、频谱带宽、模态 1 能谱、平均绝对偏差、模态 4 样本熵、模态 1 样本熵、基本频率、总功率. 为进一步验证所提预测模型的有效性, 本文按照所提方法提取了工作条件 2 下的八个特征, 并将其应用于对比实验中, 以评估模型在不同工作条件下的表现. 工作条件 2 下的八个特征分别为频谱平坦度、均方根频率、模态 1 能谱、模态 3 样本熵、谱质心、模态 2 样本熵、中心频率、峰–峰值. 图 6 中 F1 ~ F35 分别为模态 1 能谱、模态 1 样本熵、模态 2 能谱、模态 2 样本熵、模态 3 能谱、模态 3 样本熵、模态 4 能谱、模态 4 样本熵、模态 5 能谱、模态 5 样本熵、模态 6 能谱、模态 6 样本熵、模态 7 能谱、模态 7 样本熵、均值、均方根、方差、偏度、峰度、峰值、峰–峰值、平均绝对偏差、波形因子、峰值因子、脉冲因子、边际因子、频谱质心、频谱带宽、频谱平坦度、频谱峰值、总功率、信号能量、中心频率、基本频率、均方根频率.

3.5 模型消融实验

为验证所提方法的合理性和有效性, 针对所提出的 Bi-TKAN-Att, 本文基于工作条件 1 设计以下消融实验:

Bi-TKAN-Att: 完整的预测模型, 包括双向建模和注意力机制;

Bi-TKAN: 移除注意力机制, 仅保留双向建模;

TKAN-Att: 移除双向建模, 仅保留注意力机制;

TKAN: 移除双向建模和注意力机制, 仅保留单向 TKAN.

实验结果如图 7 和表 2 所示. 由图 7 可知各个模型在 RUL 预测中的表现具有明显差异. Bi-TKAN-Att 的预测结果最接近真实值, 在精度和稳定性上表现优异, 能够有效捕捉长时间序列中的 RUL 变化. Bi-TKAN 和 TKAN-Att 虽然精度略低, 但仍能较好地跟随真实值的变化趋势. 相比之下, TKAN 的预测精度较差, 未能准确反映 RUL 的变化趋势. 由表 2 可知在基础模型 TKAN 的基础上引入双向建模机制后, MAE 从 10.506 9% 降低至 8.129 8%, RMSE 从 11.846 0% 降低至 9.438 1%, R^2 从 0.779 6 提升至 0.860 1. 这表明双向建模机制显著增强了模型捕捉时间序列全局特征的能力, 能够同时建模序列的正向和反向依赖关系, 从而更好地捕获滚动轴承的退化趋势. 相比之下, 单向建模的 TKAN 无法利用未来信息对当前退化模式进行补充, 因而在长时间序列预测中表现欠佳. 在基础模型 TKAN 的基础上引入注意力机制, MAE 从 10.506 9% 降低至 8.188 8%, RMSE 从 11.846 0% 降低至 9.889 3%, R^2 从 0.779 6 提升至 0.846 4. 注意力机制通过动态

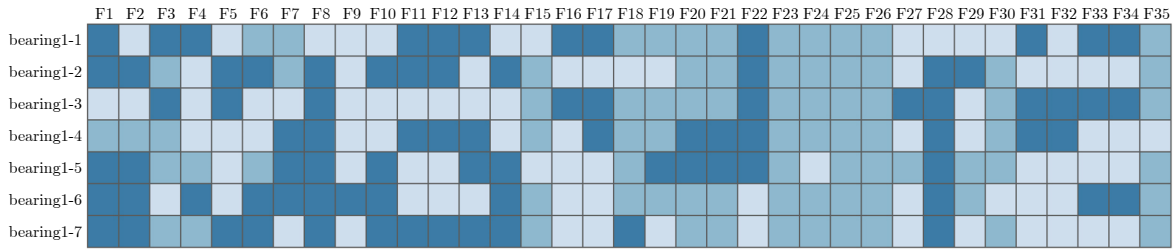


图 6 特征重要性分布图

Fig. 6 Feature importance distribution map

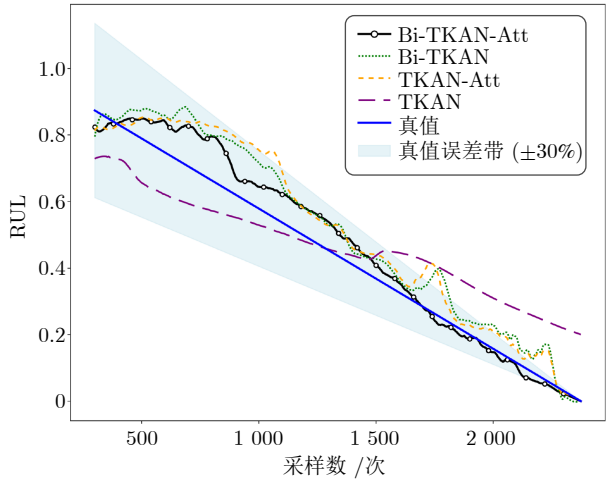


图 7　消融实验预测结果对比

Fig. 7　Comparison of ablation experiment prediction results

表 2　消融实验预测结果评价指标对比

Table 2　Comparison of evaluation metrics for ablation experiment prediction results

评价指标	Bi-TKAN-Att	Bi-TKAN	TKAN-Att	TKAN
R^2	**0.9373**	0.8601	0.8464	0.7796
RMSE (%)	**6.3199**	9.4381	9.8893	11.8460
MAE (%)	**5.0744**	8.1298	8.1888	10.5069

分配权重, 突出关键时间步的特征信息, 从而有效改善了模型在长时间序列预测中应对冗余信息的能力. 这一改进对于单向建模的 TKAN 尤其重要, 因为单向模型难以同时建模全局特征, 而注意力机制的引入能够在一定程度上弥补该不足. 在 Bi-TKAN 的基础上引入注意力机制, 形成完整模型 Bi-TKAN-Att, 其 MAE 进一步降低至 5.0744%, RMSE 降低至 6.3199%, R^2 提升至 0.9373. 相比仅包含双向建模的 Bi-TKAN, Bi-TKAN-Att 的 MAE 降低了 3.0554%, RMSE 降低了 3.1182%, R^2 提升了 0.0772. 这表明, 注意力机制与双向建模机制的结合能在捕捉时间序列全局依赖的同时, 自适应地关注关键退化特征. 这种结合使模型既能全面捕获长时间序列的动态变化, 又能聚焦于对 RUL 预测最重要的时间步, 大幅提高了 RUL 预测精度.

由消融实验结果可知, 双向建模与注意力机制在滚动轴承 RUL 预测任务中均发挥了关键作用, 双向建模机制通过捕捉时间序列的前向和后向依赖关系, 显著增强了模型的全局特征表达能力; 注意力机制通过动态聚焦关键特征, 进一步提升了模型对重要时间步的关注能力. 二者的有机结合使得 Bi-TKAN-Att 能够全面适应长时间序列预测的复杂性, 为复杂机械设备的寿命预测提供了更精准可行的解决方案.

3.6　对比实验

为进一步评估本文所提的 Bi-TKAN-Att 模型在滚动轴承 RUL 预测中的性能, 基于工作条件 1 和工作条件 2 设计对比实验, 与其他常用的深度学习 RUL 预测方法进行比较, 分别为 Transformer、TCN-BiLSTM[22]、Bi-LSTM, 对比结果如图 8 和表 3 所示.

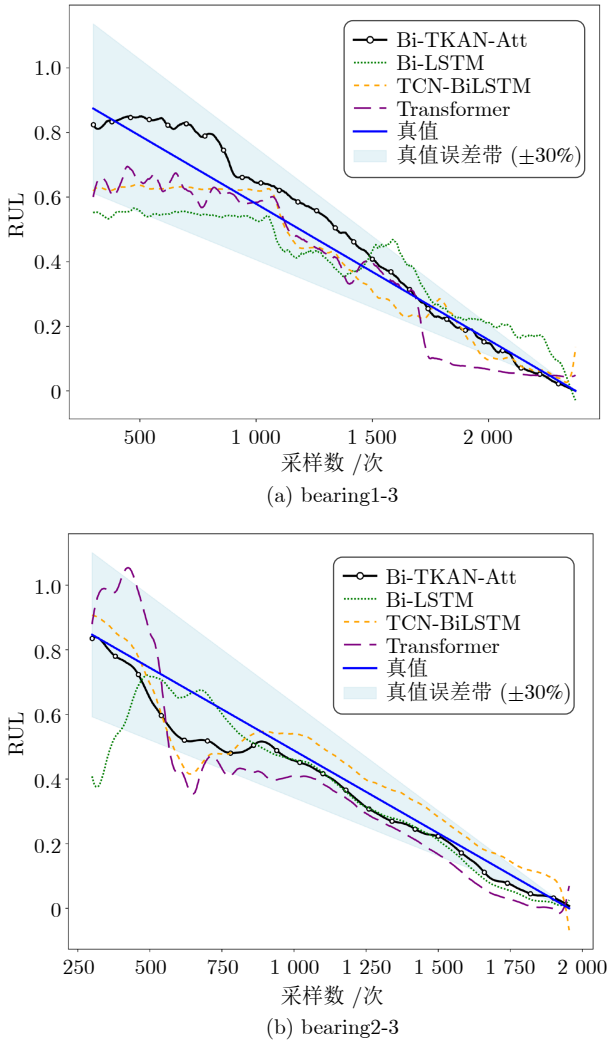


图 8　Bi-TKAN-Att 与其他方法的预测结果对比

Fig. 8　Comparison of the prediction results of Bi-TKAN-Att with those of other methods

由图 8(a) 可知, 各模型的 RUL 预测结果在 bearing1-3 上存在明显差异, Bi-TKAN-Att 的预测结果最接近真实值, 其变化趋势平稳且准确, 能够较好地捕捉到 RUL 的长期变化, 显示出其在长时间序

列预测中的优越性能. 相比之下, Bi-LSTM 虽然能够反映 RUL 的逐渐下降趋势, 但与真实值相比存在一定的误差, 预测精度相对较低. TCN-BiLSTM 的表现略优于 Bi-LSTM, 但仍未能准确捕捉到 RUL 的实际变化, 其预测曲线与真实值的偏差较为明显, 表明该模型在处理局部特征时存在一定的局限性. 而 Transformer 模型的预测结果波动较大, 未能有效地拟合真实值的变化. 在工作条件 2 下, Bi-TKAN-Att 模型依旧表现优异, 预测结果与真实 RUL 值高度一致, 能够平稳且准确地捕捉 RUL 的变化趋势. 相比之下, Bi-LSTM 虽然能够反映 RUL 的整体下降趋势, 但预测精度较低, 与真实值存在较大偏差. TCN-BiLSTM 的表现略优于 Bi-LSTM, 但仍未能准确拟合 RUL 的变化. Transformer 模型波动较大, 未能有效跟踪真实 RUL 的变化.

表 3 为不同预测方法在 bearing1-3 和 bearing2-3 上的 R^2、RMSE、MAE 三个评价指标的比较. 结果显示, Bi-TKAN-Att 在所有指标上均优于其他方法. 在两个工作条件下, Bi-TKAN-Att 的 R^2 值分别为 0.9373 和 0.9270, 显著高于其他模型, 表明其在拟合度上具有明显优势. Bi-TKAN-Att 的 RMSE 和 MAE 值分别为 6.3199% 和 5.0744% 以及 6.6083% 和 4.9011%, 在误差控制方面也表现最佳. 相比之下, Bi-LSTM、Transformer 和 TCN-BiLSTM 的预测结果稍显逊色, Transformer 的 RMSE 和 MAE 值较高, 分别为 9.2119% 和 6.9968% 以及 13.2034% 和 10.9906%, 表明其预测精度较低. 总体而言, 所提模型 Bi-TKAN-Att 在各项指标上均表现出色, 验证了其在长时间序列 RUL 预测中的优越性能.

此外, 本文还与其他相关文献中的方法进行了对比分析. 表 4 和图 9 展示了所提方法与其他文献预测方法在 bearing1-3 上的对比结果. 结果表明, 所提方法在两个指标上均取得了最佳性能, 显著优于其他方法.

3.7 交叉验证

为评估模型在同一工作条件下对不同轴承样本的泛化能力, 本文选取工作条件 1 中前五组轴承数据进行交叉验证实验. 在每轮验证中, 选定其中一个轴承作为测试对象, 其余四组用于训练, 轮流以五个轴承依次作为测试样本, 共进行五轮实验. 通过这种方式, 可以观察模型在未见轴承上的性能表现, 从而衡量其对设备间差异的适应能力. 结果如表 5 所示.

整体结果表明, 模型在大多数轴承上的预测性能良好. 然而 bearing1-2 预测结果较差, 可能是由于其在早期故障后迅速失效, 退化数据较少, 未能完整展现轴承的退化过程, 导致模型难以有效捕捉退化特征, 进而影响预测精度.

4 结束语

针对滚动轴承 RUL 预测中多维特征长期依赖关系难以有效建模的问题, 提出一种基于 Bi-TKAN-Att 的多特征时序滚动轴承 RUL 预测方法. 主要结论如下:

1) 基于 TKAN 提出一种新的双向时间建模框架, 可有效捕获时间序列的前向和后向依赖关系, 同时引入全局注意力机制, 对输入数据中关键时间步的特征信息进行自适应聚焦, 动态分配权重, 有效抑制了冗余信息对预测结果的干扰, 进一步提高了预测的精度和稳定性.

2) 在公开的滚动轴承数据集上, 通过退化特征

表 3 Bi-TKAN-Att 与其他方法预测结果评价指标对比

Table 3 Comparison of evaluation metrics of prediction results between Bi-TKAN-Att and other methods

轴承编号	评价指标	Bi-TKAN-Att	Bi-LSTM	TCN-BiLSTM	Transformer
bearing1-3	R^2	**0.9373**	0.7390	0.8898	0.8667
	RMSE (%)	**6.3199**	12.8902	8.3768	9.2119
	MAE (%)	**5.0744**	10.0602	6.0953	6.9968
bearing2-3	R^2	**0.9270**	0.8223	0.8708	0.7084
	RMSE (%)	**6.6083**	10.3079	8.7897	13.2034
	MAE (%)	**4.9011**	5.7579	6.8132	10.9906

表 4 Bi-TKAN-Att 与其他预测方法的结果对比

Table 4 Comparison of the results of Bi-TKAN-Att with other prediction methods

	Bi-TKAN-Att	TCN-SABG[23]	DT-DA[24]	Shrinkage-Att-TL[25]	Adaptive-GPR[26]	WTE-Trans[27]
RMSE (%)	**6.32**	7.7	28.4	15	77.9	17.3
MAE (%)	**5.07**	6.4	33.5	13	38.3	14.6

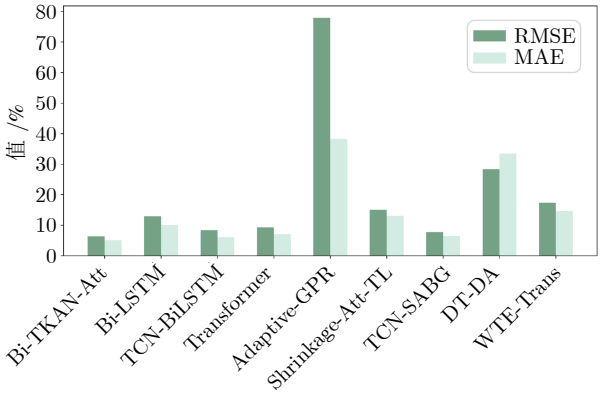


图 9　不同网络结构的 RMSE 和 MAE 比较

Fig. 9　Comparison of RMSE and MAE for different network architectures

表 5　工作条件 1 下预测结果评价指标

Table 5　Evaluation metrics of prediction results under working condition 1

轴承编号	R^2	RMSE (%)	MAE (%)
bearing1-1	0.9231	7.7832	4.6578
bearing1-2	0.8976	9.8970	8.9352
bearing1-3	0.9378	6.3178	5.0723
bearing1-4	0.9047	8.4254	7.6311
bearing1-5	0.9207	7.7601	4.5507

提取和随机森林筛选方法, 建立了多特征输入的模型训练框架. 消融实验验证了双向时间建模和注意力机制对模型性能提升的关键作用, 显示了二者结合对多阶段退化特性建模的有效性.

3) 对比实验进一步将所提 RUL 预测方法与其他文献中的方法进行比较, 结果表明 Bi-TKAN-Att 模型在 R^2、RMSE 和 MAE 性能指标上显著优于其他深度学习方法.

尽管本文提出的 Bi-TKAN-Att 方法在公开数据集上表现出色, 但在工程实际中, 进一步考虑滚动轴承所经历的复杂多变条件对 RUL 预测结果的影响, 提升其在工业场景中的泛化能力, 是未来研究需要重点关注的方向.

References

1 Dong Qing, Zheng Jian-Fei, Hu Chang-Hua, Li Bing, Mu Han-Xiao. Remaining useful life prognostic method based on two-stage adaptive Wiener process. *Acta Automatica Sinica*, 2022, **48**(2): 539−553
(董青, 郑建飞, 胡昌华, 李冰, 牟含笑. 基于两阶段自适应 Wiener 过程的剩余寿命预测方法. 自动化学报, 2022, **48**(2): 539−553)

2 Ye Z S, Revie M, Walls L. A load sharing system reliability model with managed component degradation. *IEEE Transactions on Reliability*, 2014, **63**(3): 721−730

3 Kong Y N, Ye Z S. Goodness-of-fit tests in the multi-state Markov model. *Reliability Engineering & System Safety*, 2017, **166**: 16−24

4 Cubillo A, Perinpanayagam S, Esperon-Miguez M. A review of physics-based models in prognostics: Application to gears and bearings of rotating machinery. *Advances in Mechanical Engineering*, 2016, **8**(8): 1−21

5 Wang T, Liu Z, Mrad N. A probabilistic framework for remaining useful life prediction of bearings. *IEEE Transactions on Instrumentation and Measurement*, 2021, **70**: Article No. 3503412

6 Liu L, Song X, Chen K, Hou B C, Chai X D, Ning H S. An enhanced encoder-decoder framework for bearing remaining useful life prediction. *Measurement*, 2021, **170**: Article No. 108753

7 Zhao R, Yan R Q, Chen Z H, Mao K Z, Wang P, Gao R X. Deep learning and its applications to machine health monitoring: A survey. *Mechanical Systems and Signal Processing*, 2019, **115**: 213−237

8 Mo Ren-Peng, Li Tian-Mei, Si Xiao-Sheng, Zhu Xu. Remaining useful life prediction for equipment using residual network and convolutional attention mechanism. *Journal of Xi′an Jiaotong University*, 2022, **56**(4): 194−202
(莫仁鹏, 李天梅, 司小胜, 朱旭. 采用残差网络与卷积注意力机制的设备剩余使用寿命预测方法. 西安交通大学学报, 2022, **56**(4): 194−202)

9 Guo L, Lei Y G, Li N P, Yan T, Li N B. Machinery health indicator construction based on convolutional neural networks considering trend burr. *Neurocomputing*, 2018, **292**: 142−150

10 Guo L, Li N P, Jia F, Lei Y G, Lin J. A recurrent neural network based health indicator for remaining useful life prediction of bearings. *Neurocomputing*, 2017, **240**: 98−109

11 Zheng Jian-Fei, Mu Han-Xiao, Hu Chang-Hua, Zhao Rui-Xing, Zhang Bo-Wei. Remaining useful life prediction of degradation equipment considering multiple performance indexes correlation. *Journal of Harbin Engineering University*, 2022, **43**(5): 620−629
(郑建飞, 牟含笑, 胡昌华, 赵瑞星, 张博玮. 考虑多性能指标相关性的退化设备剩余寿命预测. 哈尔滨工程大学学报, 2022, **43**(5): 620−629)

12 Lin R G, Wang H W, Xiong M L, Hou Z G, Che C C. Attention-based gate recurrent unit for remaining useful life prediction in prognostics. *Applied Soft Computing*, 2023, **143**: Article No. 110419

13 Xu Z Q, Zhang Y J, Miao Q. An attention-based multi-scale temporal convolutional network for remaining useful life prediction. *Reliability Engineering & System Safety*, 2024, **250**: Article No. 110288

14 Zhou J H, Qin Y, Chen D L, Liu F Q, Qian Q. Remaining useful life prediction of bearings by a new reinforced memory GRU network. *Advanced Engineering Informatics*, 2022, **53**: Article No. 101682

15 Zhao B, Zhang W G, Zhang Y R, Zhang C P, Zhang C, Zhang J W. Lithium-ion battery remaining useful life prediction based on interpretable deep learning and network parameter optimization. *Applied Energy*, 2025, **379**: Article No. 124713

16 Kang Shou-Qiang, Zhou Yue, Wang Yu-Jing, Xie Jin-Bao. RUL prediction method of a rolling bearing based on improved SAE and Bi-LSTM. *Acta Automatica Sinica*, 2022, **48**(9): 2327−2336
(康守强, 周月, 王玉静, 谢金宝. 基于改进 SAE 和双向 LSTM 的滚动轴承 RUL 预测方法. 自动化学报, 2022, **48**(9): 2327−2336)

17 Xia Ran, Su Chun. Remaining useful life prediction for lithium-ion batteries based on health indicators and hybrid Bi-LSTM-NAR model. *China Mechanica Engineering*, 2024, **35**(5): 851−859
(夏然, 苏春. 基于健康因子和混合 Bi-LSTM-NAR 模型的锂离子电池剩余寿命预测. 中国机械工程, 2024, **35**(5): 851−859)

18 Liu Z M, Wang Y X, Vaidya S, Ruehle F, Halverson J, Soljacic

M. KAN: Kolmogorov-Arnold networks. arXiv preprint arXiv: 2404.19756, 2024.

19 Genet R, Inzirillo H. TKAN: Temporal Kolmogorov-Arnold networks. arXiv preprint arXiv: 2405.07344v2, 2024.

20 He J, Xiao Z G, Zhang C F. Predicting the remaining useful life of rails based on improved deep spiking residual neural network. *Process Safety and Environmental Protection*, 2024, **188**: 1106−1117

21 Lei Y G, Li N P, Guo L, Li N B, Yan T, Lin J. Machinery health prognostics: A systematic review from data acquisition to RUL prediction. *Mechanical Systems and Signal Processing*, 2018, **104**: 799−834

22 Liang Hao-Peng, Cao Jie, Zhao Xiao-Qiang. Remaining useful life prediction method for bearing based on parallel bidirectional temporal convolutional network and bidirectional long and short-term memory network. *Control and Decision*, 2024, **39**(4): 1288−1296
(梁浩鹏, 曹洁, 赵小强. 基于并行双向时间卷积网络和双向长短期记忆网络的轴承剩余使用寿命预测方法. 控制与决策, 2024, **39**(4): 1288−1296)

23 Liu Ye-Feng, Wang Shuai, Liu Jing-Jing, Ma Yi-Hang. Bearing remaining life prediction based on TCN-SA and Bi-GRU. *Computer Integrated Manufacturing Systems*, DOI: 10.13196/j.cims.2024.0260
(刘业峰, 王帅, 刘晶晶, 马祎航. 基于 TCN-SA 和 Bi-GRU 的轴承剩余寿命预测. 计算机集成制造系统, DOI: 10.13196/j.cims.2024.0260)

24 Li X, Zhang W, Ma H, Luo Z, Li X. Data alignments in machinery remaining useful life prediction using deep adversarial neural networks. *Knowledge-Based Systems*, 2020, **197**: Article No. 105843

25 Li W X, Shang Z W, Gao M S, Qian S Q, Feng Z H. Remaining useful life prediction based on transfer multi-stage shrinkage attention temporal convolutional network under variable working conditions. *Reliability Engineering and System Safety*, 2022, **226**: Article No. 108722

26 Hou W J, Peng Y Z. Adaptive ensemble Gaussian process regression-driven degradation prognosis with applications to bearing degradation. *Reliability Engineering and System Safety*, 2023, **239**: Article No. 109479

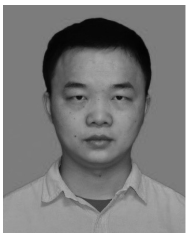
27 Zhang M Y, He C, Huang C X, Yang J H. A weighted time embedding Transformer network for remaining useful life prediction of rolling bearing. *Reliability Engineering and System Safety*, 2024, **251**: Article No. 110399

姜 蕾 火箭军工程大学导弹工程学院硕士研究生. 主要研究方向为预测与健康管理, 剩余寿命估计.
E-mail: jl00092025@163.com
(**JIANG Lei** Master student at the College of Missile Engineering, Rocket Force University of Engineering. Her research interest covers prognostics and health management, and remaining useful life estimation.)

郑建飞 火箭军工程大学教授. 主要研究方向为预测与健康管理, 剩余寿命估计. 本文通信作者.
E-mail: zjf302@126.com
(**ZHENG Jian-Fei** Professor at Rocket Force University of Engineering. His research interest covers prognostics and health management, and remaining useful life estimation. Corresponding author of this paper.)

胡昌华 火箭军工程大学教授. 主要研究方向为故障诊断, 可靠性工程和预测与健康管理.
E-mail: hch66603@163.com
(**HU Chang-Hua** Professor at Rocket Force University of Engineering. His research interest covers fault diagnosis, reliability engineering, and prognostics and health management.)

赵瑞星 火箭军装备部驻成都地区某军事代表室工程师. 主要研究方向为质量管理.
E-mail: zhruixing@163.com
(**ZHAO Rui-Xing** Engineer at Military Representative Office of the Rocket Force Equipment Department Stationed in Chengdu. His main research interest is quality management.)



韩其辉 火箭军工程大学讲师. 主要研究方向为图像处理, 计算摄影, 故障诊断与预测和模式识别.
E-mail: hanqihui2013@163.com
(**HAN Qi-Hui** Lecturer at Rocket Force University of Engineering. His research interest covers image processing, computational photography, fault diagnosis and prediction, and pattern recognition.)

杨立浩 火箭军工程大学讲师. 主要研究方向为可靠性工程, 故障诊断与预测.
E-mail: yanglihao003@163.com
(**YANG Li-Hao** Lecturer at Rocket Force University of Engineering. His research interest covers reliability engineering and fault diagnosis and prediction.)