

DOI: 10.20189/j.cnki.CN/61-1527/E.202505012

基于空间重构与并行化建模的滚动轴承 剩余寿命预测方法

韩其辉, 杨立浩, 郑建飞, 杜党波, 裴 洪
(火箭军工程大学, 西安 710025)

摘 要: 针对滚动轴承剩余寿命 (remaining useful life, RUL) 预测领域存在的网络训练并行化困难、人工特征提取导致信号高频细节与非线性退化信息丢失等问题, 提出基于空间重构与并行化建模的剩余寿命预测方法, 构建了空间编码与并行化融合的深度预测模型。首先, 利用莫顿空间填充曲线将一维振动信号重构为二维矩阵, 有效解决了原始信号完整性保持与空间相关性表征的关键问题; 其次, 通过分组卷积网络构建多尺度时频特征提取模块, 结合特征张量融合技术解决了多维度退化信息联合与时间属性嵌入的难题; 最后, 在 XJTU-SY 滚动轴承标准数据集上进行了寿命预测实验。结果表明: 所提方法较 Bi-LSTM-Att 与 CDBN-D 模型预测精度更高, 所预测出的寿命曲线更贴近真实值; 此外, 所提方法在保持原始信号完整性的同时, 能够提升计算效率, 成功验证了空间编码与并行化建模的协同增效作用。

关键词: 滚动轴承; 健康管理; RUL 预测; 空间重构; 莫顿编码

中图分类号: TH133

文献标志码: A

开放科学标识码(OSID):

文章编号: 2097-2741(2025)05-0127-10



听语音
与作者互动
聊科研

Spatial Reconstruction and Parallelizing Modeling-Based Remaining Useful Life Prediction Method of Rolling Bearings

HAN Qihui, YANG Lihao, ZHENG Jianfei,
DU Dangbo, PEI Hong

(Rocket Force University of Engineering, Xi'an 710025, China)

ABSTRACT: To address the challenges in predicting the remaining useful life (RUL) rolling bearings, such as the parallelized difficulties in network training, information loss of high-frequency signal details, and nonlinear degradation information arising from manual feature extraction—anovel RUL prediction method based on spatial reconstruction and parallelized modeling was proposed. A deep predictive model integrating spatial encoding and parallel fusion mechanisms was constructed. First, one-dimensional vibration signals were reconstructed into two-dimensional matrices using the Morton space-filling curve, effectively preserving the integrity of the original

收稿日期: 2025-04-20

修回日期: 2025-05-22

录用日期: 2025-06-05

第一作者: 韩其辉 (1991—), 男, 讲师, 主要从事设备寿命预测与健康管理研究。E-mail: hanqihui2013@163.com

引用格式: 韩其辉, 杨立浩, 郑建飞, 等. 基于空间重构与并行化建模的滚动轴承剩余寿命预测方法[J]. 火箭军工程大学学报, 2025, 39 (5): 127-136. HAN Qihui, YANG Lihao, ZHENG Jianfei, et al. Spatial reconstruction and parallelizing modeling-based remaining useful life prediction method of rolling bearings [J]. Journal of Rocket Force University of Engineering, 2025, 39 (5): 127-136.

signal and characterizing spatial correlations. Second, a multi-scale time-frequency feature extraction module was constructed using grouped convolutional network, and the problems of joint multidimensional degradation information and temporal attribute embedding were addressed through feature tensor fusion technology. Finally, life prediction experiments were conducted on the XJTU-SY rolling bearing standard dataset. The results show that the proposed method achieves higher prediction accuracy than both Bi-LSTM-Att and CDBN-D, and the predicted lifespan curves exhibit closer alignment with the true values. It maintains original signal integrity while optimizing computational efficiency, thereby successfully verifying the synergistic effects of spatial encoding and parallelized modeling.

KEYWORDS: rolling bearing; health management; RUL prediction; spatial reconstruction; Morton code

随着工业技术的快速发展,我国装备制造业的智能化转型加速推进,基于状态监测的预测性维护技术已成为设备健康管理领域的研究热点。作为旋转机械系统的关键承载部件,滚动轴承的剩余使用寿命(remaining useful life, RUL)预测在保障设备运行安全性和经济性方面具有重要工程价值^[1]。然而,传统RUL预测方法主要基于物理退化模型和统计回归分析,在复杂工况下,面临非线性退化过程难以建立精确数学模型和动态工况下的模型泛化能力不足等众多挑战^[2]。

近年来,深度学习技术在时序信号处理领域展现出显著优势^[3]。基于循环神经网络(recurrent neural network, RNN)架构的预测方法,特别是其改进型长短期记忆(long short-term memory, LSTM)网络和门控循环单元(gated recurrent unit, GRU),通过门控机制有效捕捉时序依赖关系。黄宇等^[4]利用包络谱分段计算皮尔逊相关系数以提取退化特征,再通过LSTM提取时序相关性,结合自注意力模块筛选关键信息,构建健康度指标并预测滚动轴承剩余使用寿命;陈东楠等^[5]利用Mann-Kendall检验和奇异值相关系数划分滚动轴承退化阶段,随后融合双向LSTM和注意力机制,构建剩余使用寿命预测模型;张菀等^[6]融合振动、电流、温度等多传感器数据,通过自编码器提取振动特征,结合多头注意力机制,优化多源特征权重分配,最终实现RUL预测;孟琳书等^[7]提出基于贝叶斯优化的GRU网络,结合编解码器的网络结构和注意力机制,利用小波包降噪提取时域特征而实现预测。

此外,卷积神经网络(convolutional neural network, CNN)凭借局部感受野和权值共享特性,也可以被用于振动信号的时频特征提取^[8],

而Transformer在时序数据处理中也展现出了优越性能^[9]。因此,有部分研究者尝试融合RNN与CNN、Transformer的优势,构建融合各方长处的RUL预测网络。例如,唐贵基等^[10]提出Transformer和GRU并行的网络模型,结合Transformer的全局依赖建模与GRU的时序处理能力,动态优化模型的超参数;付国忠等^[11]引入注意力机制筛选高贡献特征,结合小波包降噪和健康指标构建方法,优化输入数据质量,最后构建CNN-BiLSTM-AM混合模型实现RUL预测;惠憬明等^[12]通过CNN提取多时域指标的空间特征,而利用BiLSTM捕捉时序退化信息,最后结合自注意力模块强化关键特征权重。

然而,现有基于RNN架构的预测方法仍存在两大本质性局限。一是序列化计算机制导致训练过程难以并行化,制约了工业级大数据场景下的计算效率^[13];二是人工特征提取策略(如时频域统计、小波变换等)易造成高频细节和非线性退化信息损失,削弱模型对早期故障的敏感性。随着传感技术飞速发展,滚动轴承监测数据体量大幅增加,运行状态数据开始呈现出大规模、高维化等新特点。现有的RUL预测对原始采样信号的处理多采用构建健康指标的方法,运用统计或信号处理方法(如时频域统计、小波变换等),从状态信号中人为地提取与故障物理机制密切相关的特征,例如振动信号的均方根、峰峰值等。有时也会将多个健康指标或者传感器信号融合为新的健康指标,这些指标虽没有现实的物理意义,但能在一定程度上描述机械设备的退化趋势。然而,人工提取此类特征更容易受到经验的影响,无法适应复杂多变的工况环境。

针对RNN序列化计算问题,Tan等^[14]提出时

空注意力单元 (temporal attention unit, TAU), 将时空注意力机制解构为 2 个并行模块; Zhang 等^[15]据此提出多尺度时空注意力网络, 以提升局部特征捕获能力; Lai 等^[16]将 TAU 与频域特征提取结合, 弥补高频动态细节捕捉的不足; Teng 等^[17]将 TAU 思想迁移至人体活动识别领域, 通过双解耦方法增强了对动态特征的捕捉; Mo 等^[18]将 TAU 与多模态融合技术结合, 应用于卫星图像序列预测。研究表明, TAU 的思想可以突破 RNN 的计算瓶颈, 提高长期依赖关系的建模效率。

基于此, 本文提出一种基于空间重构与并行化建模的滚动轴承 RUL 预测方法。首先, 采用莫顿空间编码, 将一维振动信号映射为保留空间邻近性的二维栅格矩阵, 建立几何关联基础; 其次设计多分支分组卷积网络, 分别捕获信号局部冲击成分与全局趋势分量, 提升高频细节保留能力; 然后, 构建包含时间-频率-空间三维属性的高阶特征张量, 建模动态传播特性; 最后, 通过深度卷积神经网络, 实现端到端 RUL 预测。

1 理论基础

所提的基于空间重构与并行化建模的方法主要依托于莫顿空间填充曲线和分组卷积来实现, 莫顿空间填充曲线将一维振动信号映射为二维矩阵, 分组卷积代替人工提取输入二维矩阵中的特征信息。

1.1 莫顿空间填充曲线

莫顿空间填充曲线也被称为莫顿编码, 最早由 G.M. Morton 于 1966 年提出, 用于地理数据库的文件序列化。莫顿编码是一种建立多维数据到一维空间映射的编码方法, 其核心思想是通过特定的位组合方式, 将多维坐标转换为一个一维的编码值, 从而实现多维空间到一维空间的映射, 且能在一定程度上保持多维空间中的局部性特征。

如图 1 (a) 所示, 莫顿编码通过定义一条 Z 形的空间填充曲线, 使得 N 维空间中相对接近的坐标在映射到一维数据时, 仍然保持有相近的坐标, 即保持空间位置的相对性。此外, 根据莫顿编码的特性, 对于本身比较接近的块, 也可以作为一个整体而产生一个单一索引。

莫顿编码映射关系如图 1 (b) 所示, 以二维空间和一维空间映射为例, 假设有一个序列为 [1, 2, ..., 64], 想要将这 64 个数字根据莫顿编码映

射到二维空间。莫顿编码的工作过程为: 首先, 将相邻的 4 个数, 例如 1-4、5-8, 作为一个 Z 字放进第一层; 然后, 把每个 Z 字当作一个单独的位置, 每 4 个相邻的位置再以 Z 字的顺序放在第二层, 此时一个位置中包含了相邻的 16 个数; 最后, 把相邻的 16 个数当作一个单独的位置, 每 4 个相邻的位置再以 Z 字的顺序放在第三层。这种编码方式可以看作是按照特定的 Z 字形空间填充曲线, 对二维空间进行遍历和映射, 使得在编码值上相近的点在二维空间中具有一定的相邻性, 从而在一定程度上保留空间的局部性特征^[19]。

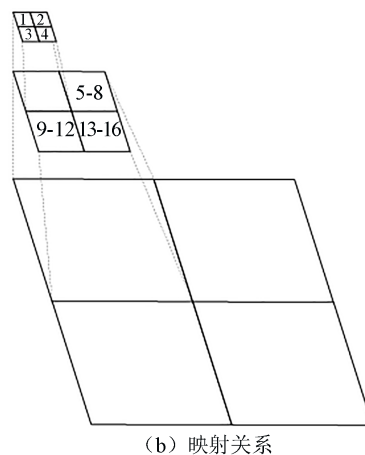
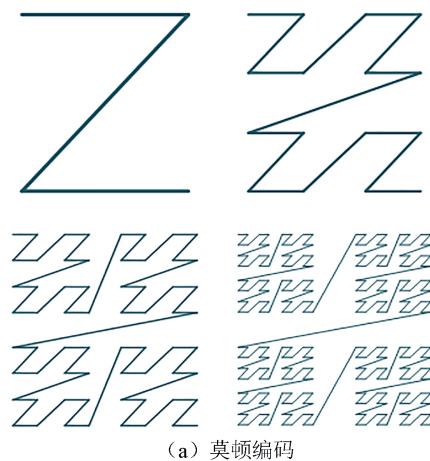


图 1 莫顿编码及映射关系

Fig. 1 Morton code and mapping relationship

由于莫顿编码按照特定的空间填充曲线对空间进行遍历, 不仅完整保留了整个采样点的数据信息, 还保存了空间信息, 可以为后续卷积神经网络的特征提取提供几何关联基础, 更利于多尺度时频信息的提取。

1.2 分组卷积

卷积神经网络在计算机视觉、自然语言处理等领域取得了一定进展。而卷积操作作为 CNN 的

核心组成部分,其设计和优化对模型性能有着深远的影响。常规卷积和分组卷积是2种常见的卷积方式,其在结构、计算效率、参数数量以及模型性能等方面存在显著差异。

常规卷积是最基本的卷积操作形式,通过一个卷积核在输入特征图上滑动,对每个位置的局部区域进行加权求和,从而生成输出特征图。如图2(a)所示,假设输入特征图的大小为 $H_{in} \times W_{in} \times C_{in}$ (输入特征高度×输入特征宽度×输入通道数),卷积核的大小为 $H_k \times W_k \times C_{in}$ (卷积核高度×卷积核宽度×输入通道数),输出特征图的大小为 $H_{out} \times W_{out} \times C_{out}$ (输出特征高度×输出特征宽度×输出通道数)。每个卷积核会生成一个输出通道,所有卷积核的输出通道叠加在一起,形成最终的输出特征图。常规卷积在每个输出通道上,对所有输入通道的信息进行融合,能够捕捉到输入特征图中不同通道之间的复杂关系。

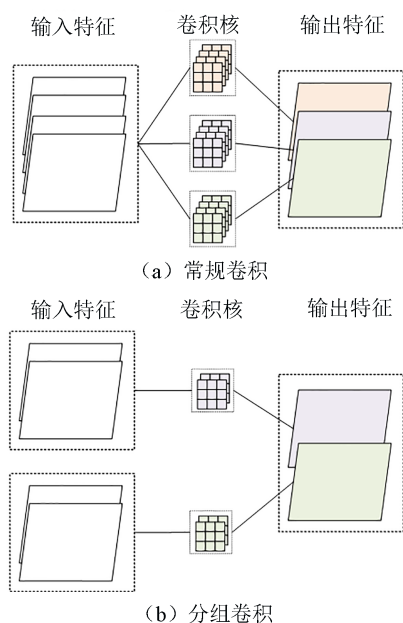


图2 卷积原理

Fig. 2 Principles of convolutions

分组卷积是常规卷积的一种变体,将输入特征图的通道分成多个组,每个组分别进行卷积操作。如图2(b)所示,假设输入特征图的通道数为 C_{in} ,将其分成 G 个组(图示为2),每个组包含 C_{in}/G 个通道。同样地,卷积核也按照组进行划分,每个组的卷积核数量为 C_{out}/G 。在每个组内,卷积操作与常规卷积类似,但不同组之间的通道是相互独立的,即一个组内的卷积核只与该组内的输入通道进行卷积,不会跨组操作。分组卷积

通过减少卷积核与输入通道之间的连接,降低了模型的计算复杂度和参数数量,同时引入了一种通道间的分层结构,有助于模型学习到更具区分性的特征。

需要注意的是,当分组卷积的分组数量 G 和输入特征图通道数量 C_{in} 一致时,每个组包含 $C_{in}/G=1$ 个通道,即每张特征图分别计算卷积,不会受到相邻通道特征的影响^[20]。同时,由于每组特征图输入通道数为1,卷积核的通道数相应地减少到1,此时每组所有的卷积操作都在二维平面上进行,可以极大地减轻算力消耗。因此,分组卷积经常出现在轻量级特征提取网络中^[21]。

2 基于空间重构与并行化建模的滚动轴承剩余寿命预测

所提的基于空间重构与并行化建模的滚动轴承剩余寿命预测方法不采用循环神经网络及其变体结构网络,而是使用卷积神经网络来并行化地处理时间依赖,将特征提取分为采样点内特征提取和相邻点间动态特征提取2个部分。采样点内特征提取使用小核心分组卷积网络和池化操作来实现大感受野,从而捕捉点内的长距离依赖关系。相邻点间动态特征提取使用深度卷积网络的方式来并行地学习相邻点之间的通道权重,从而捕捉相邻点的变化趋势。其流程如图3所示,主要包含数据空间重构、多尺度特征提取、特征并行处理和全连接RUL预测4个关键步骤。

2.1 数据空间重构

很多统计或信号处理的方法具有显著的参数敏感性,人工提取此类特征也更容易受到经验的影响。例如,时频域统计中的窗口长度、小波基类型等参数。因此,人工提取特征的策略易造成高频细节和非线性退化信息的损失,削弱模型对早期故障的敏感性。具体来说,传统时频域方法依赖固定窗函数,难以捕捉瞬态冲击信号和非平稳特征,导致早期故障的微弱高频成分被平滑或忽略^[22]。而傅里叶变换的全局频域分析会模糊信号的时间局部特性,小波变换虽具有多尺度特性,但基函数选择和尺度参数固定仍可能遗漏非平稳信号中的快速瞬变特征。此外,基于线性假设的统计特征(如均值、方差)无法表征信号中的非线性动力学行为,而早期故障往往表现为微弱的

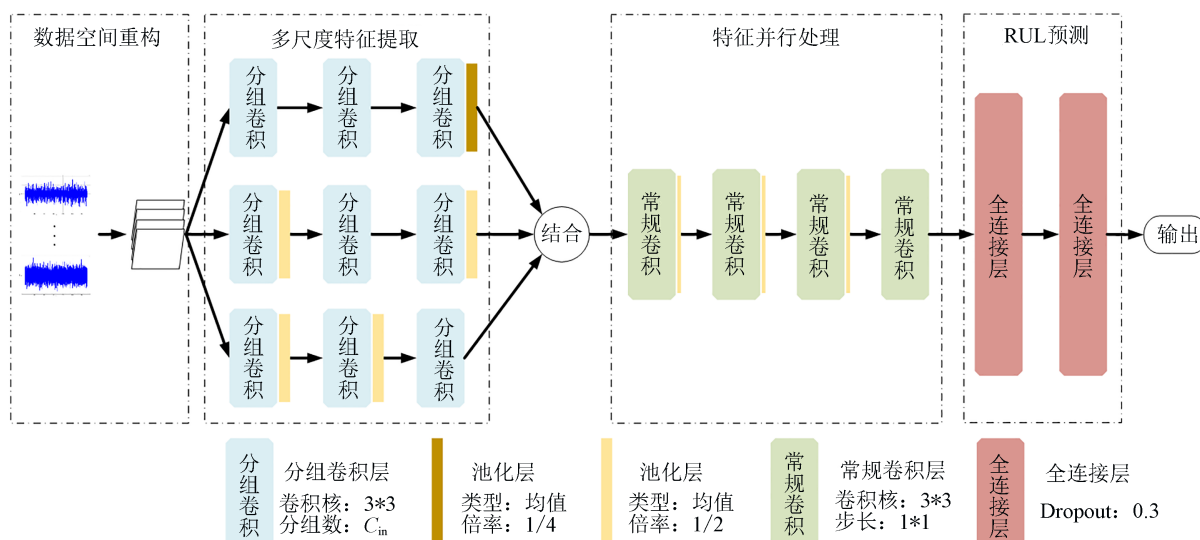


图 3 基于空间重构与并行化建模的滚动轴承剩余使用寿命预测方法流程

Fig. 3 Flowchart of the spatial reconstruction and parallelized modeling-based RUL prediction method for rolling bearings

非线性模式。

针对这些问题, 本文采取数据空间重构的方式保留完整的采样数据, 具体操作过程如图 4 所示。单个采样点传感器每分钟以 25.6 kHz 的频率收集 1.28 s 的原始振动数据, 因此, 单采样点样本数据长度为 32 768 个。单个采样点的数据经过莫顿编码后, 一维离散数据转化成大小为 128×256 的二维矩阵。莫顿编码的特性使得输入数据经过池化后, 仍能保持相对位置的邻近特性, 为提取不同频率的特征奠定了基础。最后, 将转化好的数据按照采样点的原始时间序列顺序, 采用滑动时间窗算法进行数据组合, 窗口大小设置为 16, 最终组合成带有时间维度的三维数据, 完成数据重构操作。

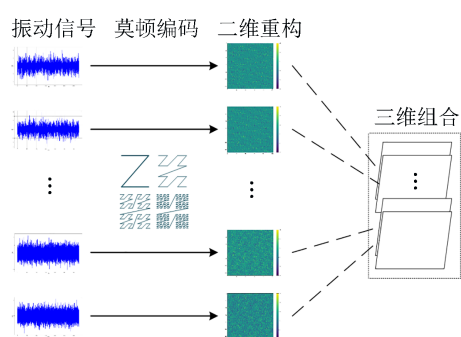


图 4 数据空间重构流程

Fig. 4 Flowchart of data space reconstruction

2.2 基于分组卷积的多尺度特征提取

数据重构操作利用莫顿编码的多尺度覆盖特性, 将不同频率的原始数据分布到栅格矩阵的特

定区域, 既保留了数据的完整性, 也保留了数据的频率特性。在特征提取中, 可以设计多尺度卷积, 使得单次前向传播中可以同时捕获低频段的全局调制特征和高频段的局部共振特征^[23], 从而避免人工特征提取造成的高频细节和非线性退化信息损失问题。

为了减少相邻采样点之间的相互影响, 利用分组卷积的方法实现采样点数据特征的单独提取。如图 2 (b) 所示, 分组卷积将输入通道分成多个组, 每个组分别进行卷积操作, 各组之间信息隔绝, 避免相互影响。将分组数设置为输入通道的通道数, 不仅减少了计算复杂度和参数数量, 还能够引入通道间的分层结构, 使得网络能够学习到更具区分性的特征。

为了捕获这些局部冲击成分和长期变化趋势, 设计了具有不同池化策略的多分支结构, 以实现不同尺度的特征提取。局部冲击成分通常指信号中的短时高频变化, 可能预示着某种重要的事件或故障。而长期变化趋势则指信号的基线漂移, 反映了滚动轴承的整体状态。为了捕获局部冲击成分, 第一分支采用 3 组分组卷积串联+池化的形式, 以更准确地识别和分类高频特征。为了有效地捕获长期变化趋势分量, 第三分支采用分组卷积+池化+分组卷积+池化+分组卷积的特征提取策略, 使得该分支具有较大感受野, 更充分地提取输入数据的低频分量。第二分支采用介于第一、第三分支感受野之间的池化策略, 采用分组卷积+池化+分组卷积+池化+分组卷积的特征提取方式,

增加提取特征的尺度特性。如图 3 所示, 3 个分支结构的最后组合为一个总体特征张量, 图中黄色条块为池化操作, 为了保持维度的一致性, 第一分支后增加了 1/4 的池化操作, 第二分支后增加了 1/2 池化操作, 其他池化倍率均为 1/2。

在滚动轴承健康管理中, 振动信号通常包含丰富的高频细节, 提取这些高频细节对于识别故障模式至关重要。通过多分支分组卷积网络, 小感受野卷积核组能够捕获振动信号中的局部冲击成分, 这些成分往往对应着滚动轴承的故障情况; 大感受野卷积核组则能够捕获全局趋势分量, 反映滚动轴承的整体状态。多尺度特征提取网络中的每个分支和卷积核组, 无需人工干预, 即可自动提取最适合输入数据的特征表示, 避免了人工提取特征带来的经验陷阱和信息损失。这不仅提高了特征提取的效率, 还能够更好地保留数据中的有效信息。

2.3 时间序列特征并行处理网络

循环神经网络及其变体在时序信号处理领域占据重要地位, 能通过循环结构能够对序列数据中的时间依赖关系进行建模, 使得 RNN 及其变体在语音识别、自然语言处理、时间序列预测等任务中取得显著成果。然而, 由于 RNN 类模型循环结构的特点, 其在训练过程中存在难以并行化的固有缺陷。因为每个时间步的计算, 都需要依赖于前一个时间步的隐藏状态, 在处理序列数据时只能按照顺序依次计算, 无法像其他模型那样进行并行计算。随着数据规模的不断增长, 这种计算方式导致其计算效率受到严重限制。例如, 在大规模的智能物联网场景中, 设备产生的时序信号数据可能达到海量级别, RNN 类模型的训练时间会显著增加, 难以满足实时或近实时处理的需求。

因此, 通过引入卷积神经网络来实现时间维度特征非线性融合的并行化处理。多尺度卷积网络提取的特征, 经过组合后组成了带有时间维度的特征图。卷积核在时间维度上进行操作, 整合时间窗内各采样点的特征信息, 对局部时间窗口内的特征进行卷积运算。此外, 多个卷积核可同时工作以分别提取不同的特征, 最终实现并行化的特征融合提取。如图 3 所示, 本方法采用常规卷积+池化+常规卷积+池化+常规卷积+池化+常规卷积的网络结构, 不断提升特征提取质量。同时, 通过池化操作减少总体数据量, 为后续预测提供有效特征。

由于所提方法能够融合多维度信息, 且可以动态捕捉滚动轴承运行状态的变化, 对复杂退化模式具有更强的适应性, 因此, 无论退化状态呈现缓慢的线性趋势, 还是出现突发的非线性变化, 所提方法均能较好地提取其有效特征, 为后续 RUL 预测提供有力支持。与循环神经网络及其变体相比, 这种并行化处理方式在大规模数据场景下仍能够保持特征提取效率, 可以更高效地完成复杂的时序信号处理任务。

2.4 剩余使用寿命预测

通过多尺度特征提取网络提取各采样点的特征后, 再使用神经网络并行处理之前提取的各采样点特征, 最后, 通过全连接网络处理融合后的特征, 以预测滚动轴承的剩余使用寿命。全连接网络由 2 个全连接层组成, 每个全连接层都使用多个神经元来学习输入与输出数据之间的复杂非线性关系, 能够实现从数据到 RUL 完整的端到端预测功能。

模型的训练就是通过反向传播来优化模型的权重和偏差, 以最小化损失函数。由于所提方法为单值预测, 经过几种常见损失函数的对比, 最终采用 Huber Loss 作为训练时的损失函数。Huber Loss 是一种结合均方误差和平均绝对误差优点的损失函数。在单值预测中, 当预测值与真实值之间的误差较小时, Huber Loss 采用均方误差的形式, 充分利用均方误差在误差较小时对误差敏感的优势, 从而精确地优化模型参数。而当误差较大时, Huber Loss 切换为平均绝对误差的形式, 避免了均方误差在处理大误差时可能出现的过优化问题。Huber Loss 函数可表示为

$$L_{\delta} = \begin{cases} \frac{1}{n} \sum_i \frac{1}{2} (y_i - \hat{y}_i)^2, & |y_i - \hat{y}_i| \leq \delta \\ \frac{1}{n} \sum_i \delta |y_i - \hat{y}_i| - \frac{1}{2} \delta^2, & |y_i - \hat{y}_i| > \delta \end{cases} \quad (1)$$

式中: y 和 $\hat{y}$ 分别为真实值和预测值; n 为训练样本的批次大小; δ 为判断阈值。

3 实验和结果分析

3.1 数据集介绍

采用西安交通大学开源的滚动轴承退化数据集来评估所提方法的有效性^[24]。数据集由滚动轴承加速寿命测试平台采集, 测试数据由 2 个安装在受测滚动轴承水平和竖直方向上的加速度计提供,

主要包含水平和竖直2个方向上的振动信号。由于该实验平台的载荷是在水平方向施加的, 而竖直方向上仅受到滚动轴承的轴向力, 故水平方向的振动信号包含更多退化信息。因此, 仅使用水平方向的振动信号作为输入数据进行RUL预测。实验中采样频率为25.6 kHz, 采样时间为1.28 s, 每隔1 min记录一次, 即每次采样记录数据包含32 768个采样点。该数据集包括3个工况数据, 工况一的负载为1 200 N, 转速为2 100 r/min; 工况二的负载为1 100 N, 转速为2 250 r/min; 工况三的负载为1 000 N, 转速为2 400 r/min。分别选取工况一和工况二下的5个全寿命周期的数据集进行实验。

3.2 实验设置

所提模型中分组卷积和常规卷积的卷积核大小均为 3×3 , 步长均为 1×1 , 卷积层后都跟有批量归一化层和Leaky ReLU激活函数。分支结构的3个卷积层卷积核的个数分别为32、64、128。第一分支的池化核的大小和步长均为 4×4 , 其他所有池化核的大小和步长均为 2×2 。4个常规卷积层的卷积核个数分别为64、128、128、256。全连接层的2个输出分别为4 096和1, 第一个全连接层后设置dropout为0.3。

3.3 训练过程

此次实验在搭载 NVIDIA 4090 显卡及 i9-14900 CPU 的 Windows11 平台上开展, 基于 PyTorch 2.6.0 框架与 Python 编程语言, 全程启用 GPU 加速。训练时选用 Adam 优化器更新网络参数, 初始学习率为 0.001, 随着迭代逐步降至 0.000 05, 最大迭代次数为 100 次, 批次大小为 32。

鉴于不同工况下数据特征各异, 分别针对工况一和工况二进行2次实验。工况一以滚动轴承1_2数据为测试集, 其余4组数据为训练集; 工况二则选取滚动轴承2_3数据作为测试集, 其他4组数据作为训练集。训练开始前, 首先对原始数据进行预处理。利用莫顿编码对水平方向的原始振动信号进行数据重构, 并采用滑动时间窗算法进行数据组合, 最后对数据进行归一化处理。工况一的模型训练完成后, 使用工况二的轴承2_3数据作为交叉验证集; 工况二的模型训练完成后, 使用工况一的轴承1_2数据作为交叉验证集。在上述实验条件下, 工况一的训练过程数据如图5所示。

由图5可以看出, 模型在迭代30次左右开始收敛, 训练集损失由1.5下降至0.08。随着学习率的下降, 训练集损失虽然还在缓慢下降, 但测试

集的损失在第50次左右达到最小, 之后开始波动向上, 表明模型开始有过拟合的风险。

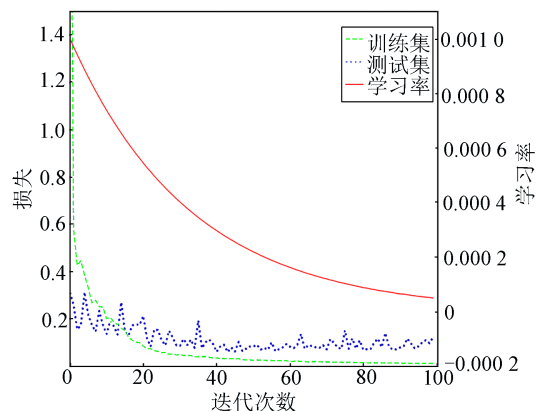


图5 工况一的训练过程数据

Fig. 5 Training process data of condition 1

3.4 结果分析

模型训练完成后, 为全面评估模型性能, 在同工况与跨工况条件下分别进行测试, 并运用配对 t 检验对同一模型在不同测试集上的表现进行深入分析。图6展示了测试集滚动轴承1_2和滚动轴承2_3在同工况和跨工况模型上的预测结果对比。

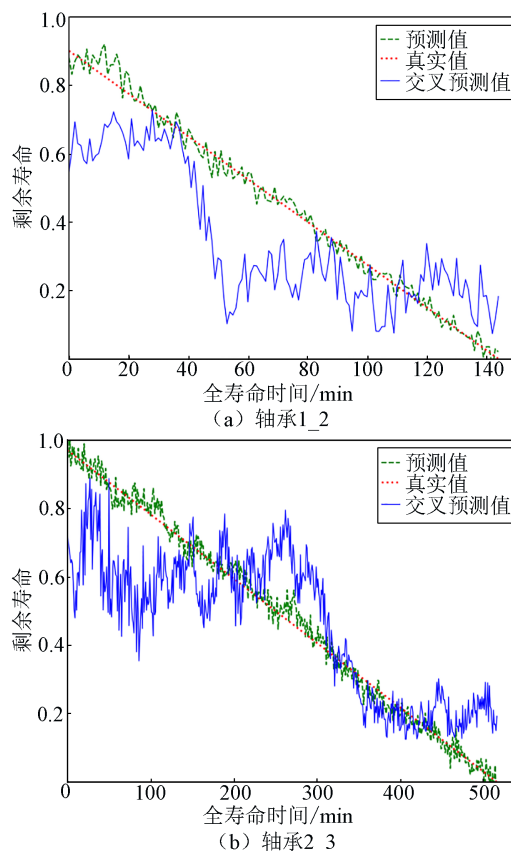


图6 预测结果与真实值对比

Fig. 6 Comparison between predicted and actual values

值得注意的是, 由于滑动时间窗设为 16, 2 组数据均未对前 15 点的剩余使用寿命进行计算。由图 6 可以看出, 模型在同工况测试中表现出色, 预测曲线紧密贴合真实值曲线, 凸显了所提方法在预测滚动轴承剩余使用寿命方面的准确性与有效性。

然而, 进一步分析发现, 同一模型在不同测试集上的表现存在显著差异。对滚动轴承 1_2 与 2_3 的同工况和跨工况测试结果进行配对 t 检验 ($p < 0.05$) 发现, 在滚动轴承 1_2 的跨工况测试中, 预测结果多集中于 0.6 和 0.2 附近, 而滚动轴承 2_3 也呈现类似趋势, 表明模型在不同工况下的预测准确性存在明显差距。分析原因, 应该是源于 2 个数据集在特征上的固有差异, 模型在训练过程中高度拟合特定工况数据, 导致在其他工况下的泛化能力受限。总体而言, 实验结果有力证明了所提方法在同工况预测中的优秀性能, 同时也揭示了跨工况预测面临的挑战。

3.5 对比实验

为了更好地验证所提方法的性能, 选取近年的 2 个滚动轴承寿命预测领域方法 Bi-LSTM-Att^[5] 和 CDBN-D^[25] 作为对比, 2 个方法均为针对滚动轴承剩余使用寿命预测的较新成果。图 7 展示了 3 种方法对测试集滚动轴承 1_2 和滚动轴承 2_3 的预测结果。

由图 7 可以看出, 3 种方法均能有效预测滚动轴承的剩余使用寿命, 但方法 Bi-LSTM-Att 和 CDBN-D 在滚动轴承运行初期预测的结果和真实值有较大误差, 且 CDBN-D 在全周期都有较大的波动, 预测稳定性不高。而所提方法全过程的预测结果均贴近真实值, 虽然初期有一些波动, 但后期波动逐渐变小, 可以稳定地预测滚动轴承的剩余使用寿命。

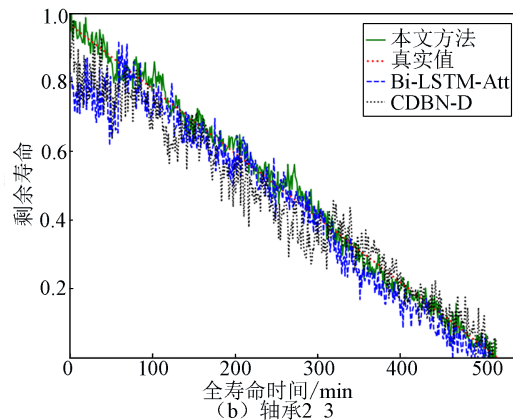
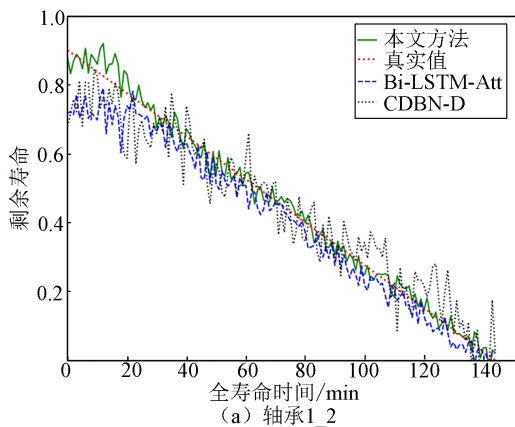


图 7 不同方法的预测结果对比

Fig. 7 Comparison of predicted results obtained using different methods

4 结论

针对现有基于 RNN 架构的预测方法在并行化计算和特征提取方面的局限性, 提出一种基于空间重构与并行化建模的滚动轴承剩余寿命预测方法, 主要结论如下。

1) 与传统时间序列分析模型相比, 所提方法创新性地引入了并行化计算方法, 不仅显著提升了计算效率, 而且在一定程度上有效缓解时间序列前后依赖所带来的问题;

2) 相较于传统通过手动提取特征再输入网络进行预测的方法, 所提方法巧妙地运用莫顿编码技术, 将时间序列重构为二维矩阵, 极大保留了原始数据的完整性和丰富性, 为后续分析和预测工作奠定了坚实的数据基础;

3) 通过在 XJTU-SY 滚动轴承运行数据集上的实验表明, 本文方法具有较高的预测准确性和显著的优越性。

本研究仍存在一些不足之处, 例如在特征提取过程中, 虽然多尺度分组卷积特征提取方法能够有效保留高频细节信息, 但对于不同工况下的数据, 最优的卷积核尺寸和分组策略仍需进一步研究和优化。此外, 时空融合张量表征方法虽然能够捕捉设备退化的动态传播特性, 但在处理大规模、高维度的时序数据时, 计算复杂度和存储需求可能会增加。后续工作将重点关注以下几个方面: 一是进一步优化模型架构和参数设置, 以提高模型在不同工况数据的适应性和预测性能; 二是结合实际工业应用场景, 开展更广泛的数据采集和实验验证, 以验证模型的实用性和可靠性;

三是探索与其他先进技术和方法的融合, 如深度强化学习、迁移学习等, 以进一步提升模型的性能和泛化能力; 四是研究如何降低模型的计算复杂度和存储需求, 以适应大规模、高维度时序数据的处理需求, 推动装备健康管理技术在工业领域的更广泛应用。

参考文献

- [1] 吕延军, 赵晓伟, 陈锐搏, 等. 关节轴承摩擦学性能与剩余寿命预测综述[J]. 交通运输工程学报, 2025, 25(1): 29-47.
LYU Y J, ZHAO X W, CHEN R B, et al. Review on frictional properties and residual life prediction of spherical plain bearing[J]. Journal of Traffic and Transportation Engineering, 2025, 25(1): 29-47.
- [2] 熊子龙, 曹哲. 考虑测试履历影响的惯导系统健康状态评估方法[J]. 火箭军工程大学学报, 2025, 39(2): 22-28.
XIONG Z L, CAO Z. Health assessment method for inertial navigation systems with consideration of test impact[J]. Journal of Rocket Force University of Engineering, 2025, 39(2): 22-28.
- [3] 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 C H, SI X S, et al. Review of machine learning based remaining useful life prediction methods for equipment[J]. Journal of Mechanical Engineering, 2019, 55(8): 1-13.
- [4] 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]. 振动工程学报, 2023, 36(6): 1744-1753.
HUANG Y, FENG K, GAO J F, et al. Combination of LSTM and Self-Attention for remaining life prediction of rolling bearings[J]. Journal of Vibration Engineering, 2023, 36(6): 1744-1753.
- [5] 陈东楠, 胡昌华, 郑建飞, 等. 状态划分下基于 Bi-LSTM-Att 的轴承剩余寿命预测[J]. 空间控制技术与应用, 2023, 49(4): 29-39.
CHEN D N, HU C H, ZHENG J F, et al. Prediction of bearing residual life based on Bi-LSTM-Att under state partition[J]. Aerospace Control and Application, 2023, 49(4): 29-39.
- [6] 张苑, 张泰瑀, 贾民平, 等. 多传感器融合和 MHA-LSTM 的电机轴承剩余寿命预测[J]. 仪器仪表学报, 2024, 45(3): 84-93.
ZHANG W, ZHANG T Y, JIA M P, et al. Prediction of remaining life of motor bearings using multi-sensor fusion and MHA-LSTM[J]. Chinese Journal of Scientific Instrument, 2024, 45(3): 84-93.
- [7] 孟琳书, 张音旋, 张起, 等. 基于贝叶斯优化的 GRU 网络轴承剩余使用寿命预测方法[J]. 机电工程, 2024, 41(1): 130-136.
MENG L S, ZHANG Y X, ZHANG Q, et al. Prediction method of RUL of bearings in GRU network based on Bayesian optimization[J]. Journal of Mechanical & Electrical Engineering, 2024, 41(1): 130-136.
- [8] 余江鸿, 彭雄露, 刘涛, 等. 融合 Inception V1-CBAM-CNN 的轴承剩余寿命预测模型[J]. 机电工程, 2024, 41(1): 107-114.
YU J H, PENG X L, LIU T, et al. Residual life prediction model of bearings based on Inception V1-CBAM-CNN[J]. Journal of Mechanical & Electrical Engineering, 2024, 41(1): 107-114.
- [9] 邓飞跃, 陈哲, 郝如江, 等. 基于 MsTCN-Transformer 模型的轴承剩余使用寿命预测研究[J]. 振动与冲击, 2024, 43(4): 279-287.
DENG F Y, CHEN Z, HAO R J, et al. Research on bearing remaining useful life prediction based on an MsTCN-Transformer model[J]. Journal of Vibration and Shock, 2024, 43(4): 279-287.
- [10] 唐贵基, 刘叔杭, 陈锦鹏, 等. 基于 Transformer-GRU 并行网络的滚动轴承剩余寿命预测[J]. 机床与液压, 2024, 52(19): 188-195.
TANG G J, LIU S H, CHEN J P, et al. Remaining useful life prediction of rolling bearings based on Transformer-GRU parallel network[J]. Machine Tool & Hydraulics, 2024, 52(19): 188-195.
- [11] 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(S2): 33-38.
FU G Z, DU H, ZHANG Z Q, et al. Remaining useful life prediction of rolling bearings based on attention mechanism and CNN-BiLSTM[J]. Nuclear Power Engineering, 2023, 44(S2): 33-38.
- [12] 惠憬明, 王健, 吴双, 等. 基于自注意力 CNN-BiLSTM 的滚动轴承剩余使用寿命预测[J]. 轴承, 2024(3): 92-98.
HUI J M, WANG J, WU S, et al. Remaining useful life prediction for rolling bearings based on self-attention CNN-BiLSTM[J]. Bearing, 2024(3): 92-98.
- [13] 潘雪娇, 董绍江, 朱朋, 等. 基于 TCN 和残差自注意力的变工况下滚动轴承剩余寿命迁移预测[J]. 振动与冲击, 2024, 43(1): 145-152.
PAN X J, DONG S J, ZHU P, et al. Transfer prediction of RUL of rolling bearing under variable operating conditions based on TCN and residual self-attention[J]. Journal of Vibration and Shock,

- 2024, 43(1): 145-152.
- [14] TAN C, GAO Z Y, WU L R, et al. Temporal attention unit: towards efficient spatiotemporal predictive learning[C]//2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ, USA: IEEE, 2023: 18770-18782.
- [15] ZHANG F, CHENG Y Y, HUA Q, et al. A multiscale spatiotemporal attention network for ground-based remote sensing cloud image sequence prediction[J]. IEEE Transactions on Geoscience and Remote Sensing, 2024, 62: 5646113.
- [16] LAI J Y, GAN L Q, ZHU J H, et al. Exploring spatial frequency information for enhanced video prediction quality[J]. IEEE Transactions on Multimedia, 2024, 26: 8955-8968.
- [17] TENG Q, LI W, HU G W, et al. Innovative dual-decoupling CNN with layer-wise temporal-spatial attention for sensor-based human activity recognition[J]. IEEE Journal of Biomedical and Health Informatics, 2025, 29(2): 1035-1048.
- [18] MO F B, HUANG Y X, WU M, et al. MMS-ISP: a satellite image sequence prediction network with multi-factor decoupling and multi-modal fusion [C]//International Conference on Pattern Recognition. Cham, Switzerland: Springer Nature, 2025: 221-236.
- [19] 袁瑶, 徐骏, 顾剑锋. 一种基于莫顿码及镜像编码的平衡八叉树模型[J]. 计算力学学报, 2024, 41(3): 467-473.
- YUAN Y, XU J, GU J F. A balanced octree based on Morton code and mirror code[J]. Chinese Journal of Computational Mechanics, 2024, 41(3): 467-473.
- [20] CHOLLET F. Xception: deep learning with depthwise separable convolutions[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ, USA: IEEE, 2017: 1800-1807.
- [21] NASIR T, MALIK M K. Efficient CRNN: towards end-to-end low resource Urdu text recognition using depthwise separable convolutions and gated recurrent units[J]. Information Processing & Management, 2024, 61(1): 103544.
- [22] 任守纲, 张景旭, 顾兴健, 等. 时间序列特征提取方法研究综述[J]. 小型微型计算机系统, 2021, 42(2): 271-278.
- REN S G, ZHANG J X, GU X J, et al. Overview of feature extraction algorithms for time series [J]. Journal of Chinese Computer Systems, 2021, 42(2): 271-278.
- [23] 崔智高, 兰云伟, 王念, 等. 基于多教师引导的知识蒸馏图像去雾算法[J]. 火箭军工程大学学报, 2024, 38(5): 35-43.
- CUI Z G, LAN Y W, WANG N, et al. An image dehazing algorithm based on multi-teacher guided knowledge distillation[J]. Journal of Rocket Force University of Engineering, 2024, 38(5): 35-43.
- [24] WANG B, LEI Y G, LI N P, 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.
- [25] PEI H, SI X S, HU C H, et al. An adaptive prognostics method for fusing CDBN and diffusion process: application to bearing data[J]. Neurocomputing, 2021, 421: 303-315.

(编辑: 于福兴)