

DOI: 10.20189/j.cnki.CN/61-1527/E.202503008

改进 Transformer 架构的混合时频域增强分解模型的重大装备剩余寿命预测

蔡志强¹, 胡昌华^{2*}, 王兆强², 李榕泽¹

(1. 西北工业大学 机电学院, 陕西 西安 710072;

2. 火箭军工程大学, 陕西 西安 710025)

摘 要: 针对航空航天领域重大装备的剩余寿命 (remaining useful life, RUL) 预测问题, 提出一种基于改进 Transformer 架构的混合时频域增强分解模型。首先, 通过融合时域与频域特征, 全面提取多层次的设备退化特征; 然后, 设计了频域增强模块, 结合傅里叶变换和小波变换, 从频域角度提取信号的周期性和局部特征; 接着, 利用分解趋势学习模块对信号进行多尺度趋势分解, 提取周期性与趋势性成分; 最后, 通过基于分层预测的残差学习模块, 进一步提高 RUL 预测的精度。为验证模型的性能, 使用 NASA C-MAPSS 数据集进行了实验。结果表明: 模型在 FD001、FD002、FD003、FD004 四个数据集上的均方根误差分别为 10.83、17.53、11.42 和 20.56, 相较于现有的深度学习方法, 具有更优的预测性能和更快的收敛速度, 在多工况和多传感器环境下具有更强的鲁棒性与泛化能力。

关键词: 重大装备; 健康管理; 剩余寿命预测; 改进 Transformer 架构; 混合模型; 深度学习

中图分类号: TP391.4

文献标志码: A

文章编号: 2097-2741(2025)03-076-14

Improved Transformer Architecture-Based Hybrid Time-Frequency Domain Enhanced Decomposition Model of Useful Life Prediction of Major Equipment

CAI Zhiqiang¹, HU Changhua^{2*}, WANG Zhaoqiang², LI Rongze¹

(1. School of Machelor Engineering, Northwestern Polytechnical University, Xi'an 710072, China;

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

ABSTRACT: A hybrid time-frequency domain enhanced decomposition model based on an improved transformer architecture was proposed to address the problems in remaining useful life (RUL) prediction of major equipment in the aerospace field. First, multi-level degradation features were extracted by integrating both time and frequency-domain features com-

收稿日期: 2024-11-30

修回日期: 2025-04-09

录用日期: 2025-04-15

基金项目: 国家自然科学基金 (72271200); 军事科研重点实验室开放基金 (ICL-2023-0304);

陕西省杰出青年科学基金 (2023-JQ-JC-10)

第一作者: 蔡志强 (1981—), 男, 教授, 主要从事重大装备健康管理及寿命预测研究。E-mail: caizhiqiang@nwpu.edu.cn

***通信作者:** 胡昌华 (1966—), 男, 教授, 主要从事自动测试、故障诊断与寿命预测研究。E-mail: hch66603@163.com

引用格式: 蔡志强, 胡昌华, 王兆强, 等. 改进 Transformer 架构的混合时频域增强分解模型的重大装备剩余寿命预测 [J]. 火箭军工程大学学报, 2025, 39 (3): 76-89. CAI Zhiqiang, HU Changhua, WANG Zhaoqiang, et al. Improved transformer architecture-based hybrid time-frequency domain enhanced decomposition model of useful life prediction of major equipment [J]. Journal of Rocket Force University of Engineering, 2025, 39 (3): 76-89.

hensively. Next, a frequency-domain enhancement module was designed by combining Fourier transform and wavelet transform to extract periodic and local features of signals. Subsequently, a decomposition-based fusion trend learning module was applied to separate and extract periodic and trend components using multi-scale analysis. Finally, a residual learning module based on hierarchical prediction was introduced to further improve the accuracy of RUL prediction. To validate the performance of the model, experiments were conducted using the NASA C-MAPSS dataset. The results show that the root mean square error values of the proposed model on the four datasets FD001, FD002, FD003, and FD004 are 10.83, 17.53, 11.42, and 20.56, respectively. Compared with existing deep learning methods, the proposed model demonstrates superior predictive performance, faster convergence speed, and higher robustness and generalization ability in multicondition and multi-sensor environments.

KEYWORDS: major equipment; health management; remaining useful life prediction; improved Transformer architecture; hybrid model; deep learning

发动机作为现代航空器的核心部件,其运行可靠性和健康状态对飞行安全和经济效益至关重要,发动机健康状态的评估和剩余寿命的准确预测日益成为确保飞行安全和优化维修策略的重要任务^[1]。剩余寿命(remaining useful life, RUL)预测技术作为健康管理的重要组成部分,其预测效率和准确度对航空发动机的维修决策和相关设备生命周期管理的重要性日益凸显,已成为航空领域预测与健康管理(prognostics and health management, PHM)的关键问题,具有显著的理论价值和实践意义。近年来,剩余寿命预测领域的研究方法主要分为3类,包括基于物理模型的方法、基于数据驱动的方法和混合方法。

基于物理模型的方法通过对设备退化机理的深入理解,建立描述退化行为的数学模型来预测系统的性能。Balli等^[2]提出一种基于热力学与环境效应的涡扇发动机性能预测方法,利用物理模型考虑不同工作条件下的复杂动力学过程,评估了涡扇发动机在能量消耗和热力学性能方面的表现。Huang等^[3]提出一种基于定位精度可靠性的公差优化模型,平衡了故障维护成本和制造成本间的矛盾。段磊光等^[4]通过交流阻抗方法,基于阻抗参数采集模块,拟合退化行为的函数曲线,实现了对石墨烯丁羟推进剂的老化检测。基于物理模型的方法在较为稳定的环境下能准确预测RUL,但在面对非线性、非平稳的退化过程时,仍面临一定局限。且其通常过度依赖于系统的先验知识和领域专家的经验,缺乏对复杂退化行为的自适应建模能力。因此,尽管物理建模方法在一些充分了解物理特性的特定设备中能够有效应对RUL

预测任务,但是在更加常见的非线性复杂问题中适用性较为有限。

随着数据获取、存储和分析技术的不断发展,基于数据驱动的方法逐渐成为PHM领域主流的解决方案。尤其是深度学习模型,不依赖于设备的物理退化机理,而是通过学习大量传感器数据中的潜在特征来进行预测。这类方法不但能够处理复杂的非线性关系,而且不过度依赖特定领域中的专业知识。Aygun等^[5]针对高功率模式下涡扇发动机的排放参数,设计了一种CNN-LSTM模型用于估计发动机重要排放参数,取得了较高准确度。Zhang等^[6]提出了基于Transformer的双重自注意力机制模型,显著提高了RUL预测精度,该方法能够自动从时序数据中提取时空特征,并在多维数据的处理上表现出色。陆文灏等^[7]针对发动机健康管理领域的装配优化问题,提出了面向多目标优化的发动机关键特征选择方法,提升了发动机的装配通过率。Kraus等^[8]针对设备的预防性维护决策,提出一种基于贝叶斯理论的可解释深度学习方法,提升了剩余寿命预测模型的可解释性,从而更好地指导维修决策。朱晓峰等^[9]以关键工装为研究对象,提出一种集成主成分分析和随机森林相结合的模型,取得了良好的预测性能。Gu等^[10]针对常规算法处理非线性振动信号的不足,提出一种通过Angular Domain重新采样方法和深层卷积神经网络模型,提升了故障诊断的分类精度。杨静宗等^[11]为提高现有诊断模型的准确度,提出了基于固有时间尺度分解、奇异值分解和多点最优最小熵反褶积相结合的故障诊断方法。此外,随着大模型理论及应用研究的发展,

Transformer 模型及其应用日益成为研究热点^[12]。在健康管理和寿命预测领域,基于 Transformer 模型处理复杂的寿命预测任务受到了学者们的关注,并在面对时间序列数据处理方面表现出了良好性能^[13]。尽管数据驱动方法具有较强的建模能力,但在多工况场景下的适应性和对噪声的鲁棒性方面仍有较大的研究空间,处理复杂时序数据时仍存在模型解释性不足、计算复杂度较大等缺陷。

近年来,混合方法逐渐兴起,得到 PHM 研究领域的广泛关注。这类方法通过结合物理模型和数据驱动方法的优势,在提高预测精度的同时增强了模型的鲁棒性。Zhuang 等^[14]结合功率和热管理系统以及发动机系统的研究成果,通过组合 PT-MS 和发动机系统开发了一种新型的集成模型,并评估了系统间的耦合效应。尽管该方法有效解决了时序数据特征提取不完善和模型可解释性较差的问题,但在处理复杂系统中多层次、非线性退化行为时,仍然面临较大挑战。此外,混合方法在计算复杂度和模型融合过程中对技术要求较高,其应用也面临着更高的挑战。

尽管现有的 RUL 预测方法在一定程度上解决了设备退化建模和预测的问题,但仍然面临以下挑战。1) 现有模型通常依赖于单一的时域或频域特征,难以全面表征复杂的退化过程;2) 深度学习方法在多传感器数据融合中往往无法有效地提取关键特征,且特征学习过程可解释性较差;3) 现有方法在面对复杂时序数据的任务时仍面临较高的计算复杂度,亟需提升模型效率。

针对现有模型和方法存在的局限性,提出了一种基于混合时频域增强分解的改进 Transformer 模型 (TFD-Hformer)。该模型融合了时域与频域特征,结合深度学习方法的优点,提升了 RUL 预测的精度和鲁棒性。

1 RUL 预测方法

RUL 是指预测设备从当前时刻开始,计算至出现功能失效或故障的剩余运行时间或周期。在 PHM 领域,准确的 RUL 预测能够显著提升设备维护决策的科学性和合理性。本文以航空发动机等重大装备为研究对象,通过分析发动机从运行至失效过程中传感器采集的时序数据,预测发动机的剩余使用寿命。具体而言,将从 n 个发动机样本传感器获得的数据定义为 $X = \{X_i | i = 1, 2, \dots,$

$n\}$, 表示其从开始运行至运行故障的全运行周期数据。每个发动机配备 m 个传感器,持续记录运行过程中各种参数的变化情况。将第 i 个发动机的监测数据定义为 $X_i = \{x_{i,1}, x_{i,2}, \dots, x_{i,m}\}$ 。式中: $x_{i,m}$ 表示第 i 个发动机的第 m 个传感器获取到的时间序列数据。将某个具体的传感器获取到的监测数据定义为 $x_{i,j} = \{x_{i,j}^1, x_{i,j}^2, \dots, x_{i,j}^{T_i}\}$, 其中 T_i 表示从运行到故障的总时间步长。 $X_i^t = \{x_{i,1}^t, x_{i,2}^t, \dots, x_{i,m}^t\}$ 表示第 i 个发动机的第 m 个传感器在第 t 个时间步所在窗口下的时间序列数据。设所有发动机的剩余寿命值为 $Y = \{Y_i | i = 1, 2, \dots, n\}$, 对于某一台发动机,其剩余寿命值的变化表示为 $Y_i = \{y_i^1, y_i^2, \dots, y_i^{T_i}\}$, 其中 $y_i^{T_i}$ 表示第 i 个发动机在第 T_i 个时间步的 RUL 值。因此, RUL 预测的目标即为学习一个映射函数 f , 来准确捕捉到第 t 个时间步下监测的一段时序数据同当前时刻下的剩余寿命之间的关系,即 $f: X_i^t \rightarrow y_i^t$ 。

1.1 RUL 预测流程

所提的 TFD-Hformer 模型在应对 RUL 预测问题的实现流程如图 1 所示。总体包括数据预处理、模型构建、模型训练及预测评估 3 个阶段。数据预处理阶段主要是为模型训练构建提供充足的数据样本,模型构建阶段主要是设计并搭建 TFD-Hformer 模型,模型训练及预测评估阶段主要是通过合适的模型训练策略,获得 RUL 预测模型,并实现准确的 RUL 预测。具体可划分为如下步骤。

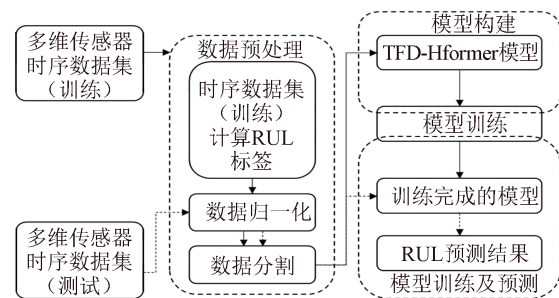


图 1 RUL 预测方法流程

Fig. 1 Flowchart of RUL prediction

- 1) 数据预处理: 包括训练集数据的 RUL 计算、归一化处理和数据分割;
- 2) 模型构建: 根据参数配置搭建模型的主体结构;
- 3) 模型训练及预测评估: 采用监督学习方法训练模型, 并对模型性能进行评估。

1.2 TFD-Hformer 预测模型

所提的 TFD-Hformer 模型是一种基于改进

Transformer 架构的时序预测模型, 在装备剩余寿命预测任务应用中取得了良好的效果。本模型以 Transformer 模型结构为基础, 结合门控循环单元 (GRU)、时间序列分解模块、残差学习模块和频域注意力机制, 以及数据集中时序数据的特点进

行参数设置, 核心架构及组件如图 2 所示。
基于模型的主体结构和输入的时序数据特性, TFD-Hformer 模型的部分关键参数设置细节如表 1 所示, 基于最优参数搭建并训练模型, 并对测试集进行 RUL 预测及模型性能评价。

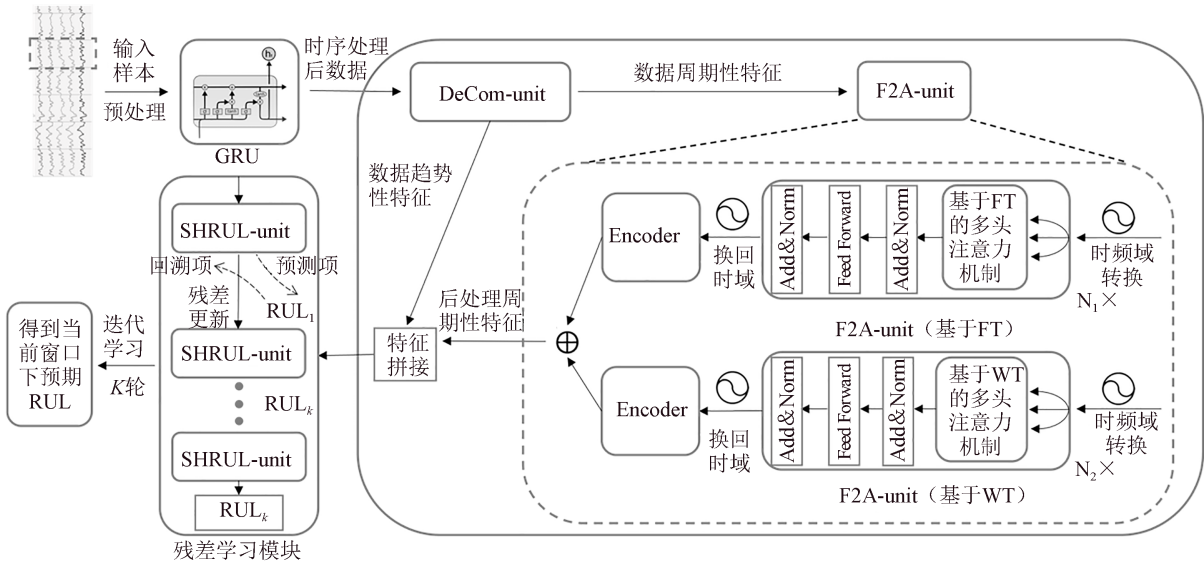


图 2 TFD-Hformer 模型架构
Fig. 2 Architecture of TFD-Hformer model

表 1 TFD-Hformer 模型关键参数
Table 1 Key parameters of TFD-Hformer model

参数		FD001	FD002	FD003	FD004
最大周期数(未触发早停)		50			
批大小		64	32	64	32
滑动窗口长度 L 、步长 Step		30、1	25、1	30、1	25、1
GRU	input_size	输入数据每个时间窗口下的特征数量(21)			
	hidden_size	隐藏层状态维度(128)			
	num_layers	GRU 层数(2)			
	dropout	避免过拟合(0.2)			
DeCom-unit	kernel_size	平滑窗口大小(25), 用于划分周期性、趋势性特征			
F2A-unit	in_channels	输入通道数(64)			
	out_channels	频域增强后, 输出通道数(256)			
	modes	频域模式数量(4)			
Encoder	num_layers	编码器层数(2)			
	in_channels	输入通道数(256)			
	out_channels	输出通道数(64)			
SHRUL-unit	in_channels	输入通道数(128), 周期性、趋势性特征拼接			
	n_layers	SHRUL_unit 层数(4)			
	n_hidden	隐藏层单元数(256)			
	dropout	避免过拟合(0.2)			
损失函数		使用均方根误差(RMSELoss)作为损失函数			
优化器		使用 Adam 优化器(学习率 0.001 7)			

2 基于混合时频域增强分解的重大装备剩余寿命预测方法

2.1 RUL 计算方法

对于某一台发动机而言, 将其剩余寿命值随运行时间周期的变化表示为 $Y_i = \{y_i^1, y_i^2, \dots, y_i^{T_i}\}$, 其中 y_i^t 表示从当前时间 t 到发动机故障时刻 T_i 的时间间隔, 可表示为

$$y_i^t = T_i - t \quad (1)$$

然而, 在实际运行过程中, 设备在运行初期往往处于性能稳定阶段, 其退化行为并不明显, 因此, 此阶段的 RUL 变化对模型训练的贡献十分有限, 甚至可能造成干扰。为了更准确地反映设备的退化特征, 常采用如下分段函数形式对 RUL 进行处理。

$$y_i^t = \begin{cases} y_{\max}, & t \leq t_{\lambda} \\ T_i - t, & t > t_{\lambda} \end{cases} \quad (2)$$

式中: $y_{\max}$ 表示分段函数在初期运行的最大阈值; t_{λ} 表示达到阈值时的最大时间间隔。

2.2 多传感器数据的归一化方法

在设备运行过程中, 不同传感器采集的数据常常在数值尺度和物理单位上存在显著差异。若直接输入模型, 易出现数值主导、梯度不平衡等问题, 严重影响模型训练效果。因此, 需要对模型输入的多维传感器序列数据采取归一化处理。同时, 为了确保模型在训练过程中获得更好的收敛表现, 本文对样本标签 RUL 值也进行了归一化操作。

对多维传感器时间序列数据采取最大-最小归一化方法, 可表示为

$$\bar{x}_{ij}^t = \frac{x_{ij}^t - x_j^{\min}}{x_j^{\max} - x_j^{\min}} \quad (3)$$

式中: $x_j^{\max}$ 和 $x_j^{\min}$ 分别表示第 j 个传感器在训练集中的最大值和最小值; $\bar{x}_{ij}^t \in [0, 1]$ 为样本多维传感器时间序列数据的归一化结果。

在实验中对样本标签 RUL 值也采取最大-最小归一化方法, 以在模型的训练中获得更好的收敛表现, 可表示为

$$\bar{y}_i^t = \frac{y_i^t - y_{\min}}{y_{\max} - y_{\min}} \quad (4)$$

式中: $y_{\max}$ 和 $y_{\min}$ 分别表示 RUL 值分段函数在初期运行的最大阈值和最小阈值, $y_{\min}$ 取值为 0。

2.3 数据分割方法

划分时间窗口是处理时间序列类型数据的常用方法, 采用不同窗口大小 L 和步长 Step 的滑动时间窗口来划分每台发动机预处理后的多变量时间序列数据, 如图 3 所示。具体而言, 当 $\text{Step}=1$ 时, 通过切分时间窗口, 将最大运行时间周期为 T_i 的发动机样本个数增加至 $T_i - L + 1$ 。对于经过滑动窗口数据分割后的样本, 取当前窗口的最后时刻计算样本的剩余寿命。那么该发动机运行过程中剩余寿命值的变化过程可表示为 $Y_i = \{y_i^1, y_i^2, \dots, y_i^{T_i-L+1}\}$, 可以看出, 可供模型训练的样本个数有所增加。

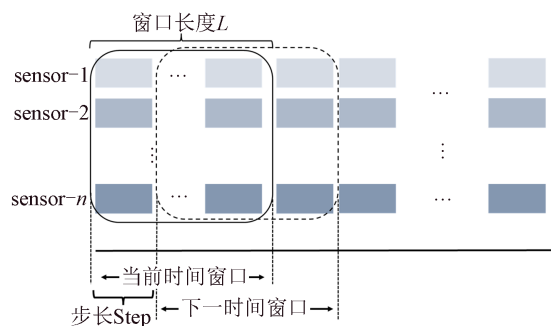


图 3 时间窗口划分步骤

Fig. 3 Steps of time window division

2.4 预测框架

剩余寿命预测任务中的时间序列数据有着较强的时频域特征, 为了更加有针对性地捕捉数据样本在退化过程的数据特征, 提出了一种改进 Transformer 架构下的基于混合时频域增强分解的重大装备 RUL 预测方法, 命名为 TFD-Hformer 模型。所提模型的核心思想是在模型学习运行数据总体趋势之外, 关注部分特征属性在时频域融合视角下展示出的周期性和趋势性, 以缓解传统 Transformer 模型在处理 RUL 预测任务时表现出的无法准确捕捉时间序列总体趋势的缺点, 同时提升模型在面对样本存在不同属性间时频域特征差异时的学习能力。

TFD-Hformer 模型由 3 部分组成, 分别为基于频域特征的信号处理模块 F2A-unit、基于分解的融合趋势学习模块 DeCom-unit 以及基于分层预测的残差学习模块 SHRUL-unit, 如图 2 所示。

2.4.1 基于频域特征的信号处理模块 F2A-unit

本部分模型基于 Transformer 架构进行改进, 由于传统的 Transformer 模型采取逐步关注、独立预测的策略, 无法准确地捕捉到设备退化过程中

产生的时间序列数据的总体趋势。因此, 针对模型在处理 RUL 预测任务时的局限性, 对 Transformer 模型的 encoder 模块及注意力模块进行系统性改进, 命名为 F2A-unit 子模块, 以实现模型在 RUL 预测任务的有效性。

在传统的信号处理领域, 傅里叶变换 (Fourier transform, FT) 和小波变换 (wavelet transform, WT) 是常用的信号处理方法。FT 是一种分析信号频率成分的常用工具, 对于分析数据所表现出的全局周期性有着良好的表现。WT 是一种通过引入可伸缩的小波函数来描述信号的局部结构, 对于获取非平稳信号的局部特征监测有着广泛的应用。鉴于 2 种方法分别关注频域上的全局和局部特征, 本文设计了带有基于 FT 和 WT 频域增强结构的信号处理模块 F2A-unit 用于系统性地提取数据的频域特征, 其核心结构如图 4 所示。

针对不同的数据集, 在选取的适当时间窗口内应用 FT 或 WT, 以提取数据中的全局和局部时频特征。通过时频域融合分解方法进一步增强特征的表达能力, 最终将提取到的频域特征与时域特征一起输入到后续模型中。其中, FT 能够有效地揭示设备在较长时间尺度上的周期性规律, 而 WT 能够准确刻画数据的局部时频信息, 为捕捉设备退化过程中潜在的突变和非平稳特征提供独特的视角。本模型通过混合时频域增强分解, 以提高在复杂工况下对设备退化过程的建模能力, 并增强模型在不同工况变化下的鲁棒性。此外, 为了将频域特征高效融合, 还采用动态调整傅里叶特征和小波特征在最终融合时权重的方法, 提高模型对不同特征的学习能力, 进一步提升 RUL 预测的准确性。

1) 基于 FT 的信号处理模块

假设 F 为傅里叶变换, F^{-1} 即为逆傅里叶变换。将一段长为 T 的时域的特征序列 $x = \{x_1, x_2, \dots, x_n, \dots, x_N\}$ 进行傅里叶变换, 其过程定义为

$$X_l = \sum_{n=0}^N x_n^* e^{-i\omega n l} \quad (5)$$

式中: l 表示频域上的复数序列。

同理, 逆傅里叶变换可以定义为

$$x_n = \sum_{l=0}^{L-1} X_l^* e^{i\omega n l} \quad (6)$$

基于傅里叶变换的频域增强结构应用于 Encoder 模块, 并结合基于分解的融合趋势学习模块 DeCom-unit 模块, 可以增强模型在频域方面的全局特征学习能力, 其结构如图 4 上部所示。设输入的信号为 $x \in \mathbb{R}^{N \times D}$, 其中 D 为模型隐藏层的单元序列, 首先对使用权重 $\omega \in \mathbb{R}^{D \times D}$ 进行线性投影, 定义 $q = x \cdot \omega$, $q \in Q$, 转换到频域后进行傅里叶变换后得到 $Q = F(q)$, $Q \in \mathbb{R}^{N \times D}$, 筛选其中 m 个特征进行频域增强, 那么傅里叶变换的频域增强模块对应函数 f_{FT1} 定义为

$$f_{FT1}(q_{(m)}) = F^{-1}(\text{padding}(Q \odot R)), Q \in \mathbb{R}^{M \times D} \quad (7)$$

式中: R 为参数核; $\odot$ 运算定义为

$$S(m, f) = Q \odot R = \sum_{d_i=0}^D Q(m, d_i) \cdot R(d_i, f, m),$$

$$d_i = 1, 2, \dots, D, f = 1, 2, \dots, D \quad (8)$$

式中: f 表示输出维度; d_i 表示输入维度。

完成运算后将 $Q \odot R \in \mathbb{R}^{M \times D}$ 使用零填充到空间 $\mathbb{R}^{N \times D}$, 并使用逆傅里叶变换回到时域分布。

本文同时将频域增强结构应用于注意力模块。在基础的 Transformer 架构中, 查询向量 q 、键向

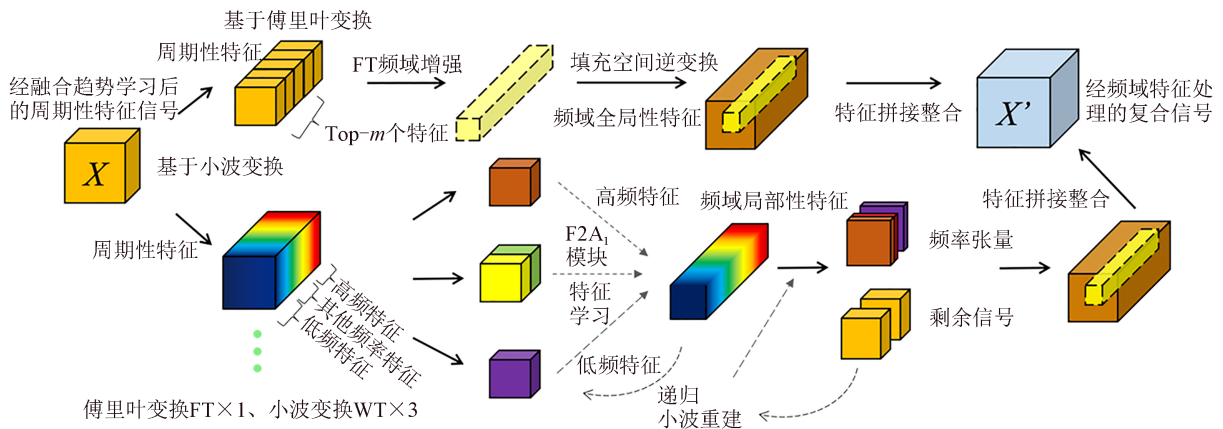


图 4 F2A-unit 模块核心结构

Fig. 4 Core structure of F2A-unit module

量 $\mathbf{k}$ 和值向量 $\mathbf{v}$ 是核心概念。定义

$$\mathbf{q} \in \mathbb{R}^{L_q \times D_k}, \mathbf{k} \in \mathbb{R}^{L_k \times D_k}, \mathbf{v} \in \mathbb{R}^{L_v \times D_v} \quad (9)$$

式中: L_q 为查询的长度; D_k 为特征维度; L_k 为键长度; D_v 为值的维度。交叉注意力机制公式定义为

$$\text{Attention}(\mathbf{Q}, \mathbf{K}, \mathbf{V}) = \text{Soft max} \left(\frac{\mathbf{Q}\mathbf{K}^T}{\sqrt{d_k}} \right) \mathbf{V} \quad (10)$$

定义 X_{en} 为 Decoder 的输入, X_{de} 为 Encoder 的输出。式 (10) 可以变为

$$\text{Attention}(X_{\text{en}}, X_{\text{de}}) = \text{Soft max} \left(\frac{(X_{\text{en}} \cdot \mathbf{W}_Q) \cdot (X_{\text{de}} \cdot \mathbf{W}_K)^T}{\sqrt{d_k}} \right) \cdot (X_{\text{de}} \cdot \mathbf{W}_V) \quad (11)$$

式中: $\mathbf{W}_Q \in \mathbb{R}^{D \times D}$ 表示查询的权重矩阵; $\mathbf{W}_K \in \mathbb{R}^{D \times D}$ 表示键的权重矩阵; $\mathbf{W}_V \in \mathbb{R}^{D \times D}$ 表示值的权重矩阵。

筛选其中 m 个特征进行注意力增强, 那么原定义可以替换为 $\mathbf{Q} \in \mathbb{R}^{M \times D}, \mathbf{K} \in \mathbb{R}^{M \times D}, \mathbf{V} \in \mathbb{R}^{M \times D}$, 则注意力增强模块对应函数可定义为

$$f_{\text{FT2}}(\mathbf{q}, \mathbf{k}, \mathbf{v}) = F^{-1}(\text{padding}(\tanh(\mathbf{Q} \cdot \mathbf{K}^T) \cdot \mathbf{V})) \quad (12)$$

完成运算后将 $\tanh(\mathbf{Q} \cdot \mathbf{K}^T) \cdot \mathbf{V} \in \mathbb{C}^{M \times D}$ 使用零填充到空间 $\mathbb{R}^{N \times D}$, 并使用逆傅里叶变换回到时域分布。

2) 基于 WT 的信号处理模块

为解决复杂信号中频域出现局部特征耦合导致的建模困难, 提出基于 WT 的频域增强结构, 通过递归分解和频率增强机制, 有效捕捉了复杂信号中局部多尺度数据特征, 其结构如图 4 下部所示。与基于 FT 的信号处理模块类似, 本模块通过将频域增强结构应用于 Encoder 模块实现局部特征提取。

基于小波变换的频域增强结构应用于 Encoder 模块, 主要通过固定次数的递归轮次将信号数据区分为高频、低频和其他频率数据, 并使用 3 个 F2A₁ 模块学习局部特征。在小波重建部分, 对于

每个递归周期, 从分解部分产生经过变换的低频张量、高频张量和剩余信号部分, 并为下一个递归周期输出待处理的信号数据。

3) 全局-局部特征融合模块

在频域特征提取阶段, 使用 FT 和 WT 分别提取全局频域特征和局部时频特征。傅里叶和小波特征的融合主要通过注意力加权融合方法。即为每种特征计算一个注意力分数, 该分数反映了每类特征对最终预测任务的贡献。然后, 通过加权求和将 2 类特征融合, 最终得到一个加权的融合特征表示。分别计算两组特征的注意力得分, 可表示为

$$\alpha_f = \exp(\mu_f^T f) \quad (13)$$

$$\alpha_w = \exp(\mu_w^T w) \quad (14)$$

式中: μ_f^T 和 μ_w^T 为可学习的参数向量; f 和 w 分别为全局频域特征和局部时频特征。

归一化权重后, 形成的融合特征 y 可表示为

$$y = \alpha_f f + \alpha_w w \quad (15)$$

2.4.2 基于分解的融合趋势学习模块 DeCom-unit

与基于小波分解的模块不同, 所提的 DeCom-unit 模块用于多次区分数据的周期性特征和趋势特征, 并最终融合形成数据的最终趋势, 其核心结构如图 5 所示。

由于原始数据中观察到的周期性和趋势性存在着明显的特征耦合, 因此, 使用固定长度划分的时间窗口往往难以同时进行学习, DeCom-unit 模块是一组含有不同池化内核的平均池化层, 结合多尺度分析和自适应加权机制, 从输入信号中提取不同的趋势分量, 并使用数据相关的权重对这些趋势分量进行组合, 表示为

$$\text{DataTrend} = \text{soft max}(L(x)) * \text{MF}(x) \quad (16)$$

式中: $\text{MF}(x)$ 表示一个滤波器; $\text{soft max}(L(x))$ 表示该滤波器对应的权重。权重 $L(x)$ 采用 Average Pooling 计算方法。

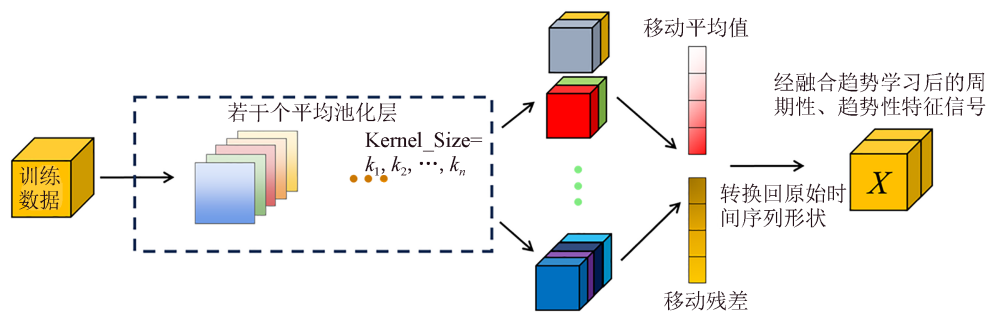


Fig. 5 Core structure of DeCom-unit module

2.4.3 基于分层预测的残差学习模块 SHRUL-unit

基于频域特征的信号处理模块 F2A-unit 和基于分解的融合趋势学习模块 DeCom-unit 分别从频域和时域的角度对输入信号进行了分解和增强。然而, 航空航天领域重大装备的退化过程往往具有复杂的非线性和多尺度特征, 单纯依靠上述模块直接预测 RUL, 可能出现特征学习不完全、误差较大等缺陷。因此, 本文提出基于分层预测的残差学习模块 SHRUL-unit, 旨在通过迭代建模, 进一步提高对剩余寿命的预测精度。

SHRUL-unit 的核心思想是利用分层残差学习的策略, 对时间序列数据中的未解释部分进行逐层建模。具体来说, 模型通过堆叠多个神经网络学习模块, 对残差进行多次迭代处理, 每个模块都递归地捕捉前一层未能捕捉到的特征, 从而逐步减少预测误差, 该模块的核心结构如图 6 所示。此结构能够充分利用深层网络的优势, 充分捕捉经过时频域融合增强后数据中的非线性和复杂模式, 提高模型的泛化能力。

每个 SHRUL-unit 由若干个子模块组成, 每个子模块都包括输入层、隐藏层和输出层。输入层接收来自上一层的残差信息, 隐藏层由多层全连接网络组成, 输出层则生成回溯项和预测项, SHRUL_unit 分层预测的具体过程在表 2 中以算法形式进行详细的描述。假设输入残差向量为 $\mathbf{r} \in \mathbb{R}^n$, 经过由 L 层全连接网络组成的隐藏层, 每一层的输出通过 ReLU 激活函数, 并使用 Dropout 进行正则化, 防止过拟合。第 l 层的输出为

$\mathbf{h}^{(l)} = \text{ReLU}(\mathbf{W}^{(l)}\mathbf{h}^{(l-1)} + \mathbf{b}^{(l)}), l = 1, 2, \dots, L$ (17)

式中: $\mathbf{h}^{(0)} = \mathbf{r}$; $\mathbf{W}^{(l)}$ 和 $\mathbf{b}^{(l)}$ 分别为第 l 层的权重矩阵和偏置向量。

最后一层全连接网络将隐藏层的输出映射到参数向量 $\boldsymbol{\theta} \in \mathbb{R}^{2n}$, 即

$$\boldsymbol{\theta} = \mathbf{W}^{(L+1)}\mathbf{h}^{(L)} + \mathbf{b}^{(L+1)} \quad (18)$$

式中: $\boldsymbol{\theta}$ 包含回溯项信息 $\mathbf{b} = \boldsymbol{\theta}_{[1:n]}$ 和预测项信息

$$\mathbf{f} = \boldsymbol{\theta}_{[(n+1):2n]}$$

表 2 SHRUL_unit 分层预测算法

Table 2 SHRUL_unit hierarchical prediction algorithm

算法 1 SHRUL_unit 分层预测算法及迭代训练过程

输入及初始化: $\mathbf{x} \in \mathbb{R}^n$, 残差向量 $\mathbf{r}_k, \mathbf{r}_0 = \mathbf{x}$, 最大迭代周期

K , 预测值 $F^{(0)} = 0$ 。

输出: 当前样本的预测 RUL 值 $F^{(K)}$ 。

步骤 具体迭代训练过程

- 1 对于每一个 k 从 0 至 K , 执行如下操作。
- 2 计算第 k 个残差学习模块的前向传播值:
 $\boldsymbol{\theta}^{(k)} = \mathcal{NN}^{(k)}(\mathbf{r}_{k-1})$;
- 3 计算当前的回溯项: $\mathbf{b}^{(k)} = \boldsymbol{\theta}_{[1:n]}^{(k)}$;
- 4 计算当前的预测项: $\mathbf{f}^{(k)} = \boldsymbol{\theta}_{[(n+1):2n]}^{(k)}$;
- 5 依据回溯项更新剩余残差: $\mathbf{r}_k = \mathbf{r}_{k-1} - \mathbf{b}^{(k)}$;
- 6 依据预测项信息变换至一维, 叠加更新预测值:
 $F^{(k+1)} = F^{(k)} + \mathbf{f}^{(k)}$;
- 7 达到最大迭代周期 K 时, 当前样本的预测 RUL 值即为 $F^{(K)}$;
- 8 迭代结束, 输出当前预测 RUL 值。

在分层预测的过程中, 每个模块可能捕捉到不同的数据模式或特征, 将这些信息进行有效融合, 可以增强模型的整体预测性能。

2.4.4 协作机制

基于频域特征的信号处理模块 F2A-unit、基于分解的融合趋势学习模块 DeCom-unit 以及基于分层预测的残差学习模块 SHRUL-unit 分别聚焦于特征学习的不同方面, 同时也有着密切的协作机制。

1) F2A-unit 和 DeCom-unit 模块间协作机制

DeCom-unit 模块接收来自 GRU 层时序特征提取后的数据, 并进一步对其进行时域分解。该模块通过多尺度池化和加权机制, 将信号的周期性成分和趋势性成分分离开来, 为 F2A-unit 模块提供更加精细的信息。F2A-unit 模块通过时频域转换方法, 采取基于 FT 和 WT 的方法, 从频域角

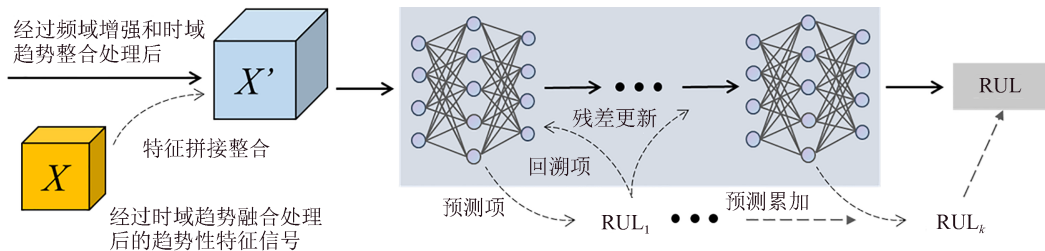


图 6 SHRUL-unit 模块的核心结构

Fig. 6 Core structure of SHRUL-unit module

度提取输入信号的全局周期性和局部时频特征。F2A-unit 提供了关于信号在频域上的关键信息,为后续的时域建模提供丰富的特征输入。DeCom-unit 模块对于将信号的初步特征分解为 F2A-unit 模块,并进行有针对性的特征学习提供了可能,确保了模型能够充分地学习信号数据在不同频域下的综合特征。

2) DeCom-unit 和 SHRUL-unit 的协作机制

DeCom-unit 模块对信号进行特征分解,将周期性和趋势性特征分离并输出。这些分解后的特征经过不同的处理,最终被送入 SHRUL-unit 模块,用于剩余寿命的预测建模。SHRUL-unit 运用分层残差学习策略,通过递归的方式逐层捕捉前一层未能捕捉到的特征,并生成最终的预测结果。DeCom-unit 提供的分解信号能够区分地表示数据的周期性和趋势性特征,SHRUL-unit 模块也会在残差学习的过程中,同时学习到数据的多维特征,多层次中对残差进行建模,从而使得模型能够逐步减少误差,提升预测精度。

3) F2A-unit、DeCom-unit 和 SHRUL-unit 的协作机制

F2A-unit 在多频域数据环境下的特征提取能力、DeCom-unit 提供的特征分解功能以及 SHRUL-unit 基于分层残差学习的 RUL 预测,在 TFD-Hformer 模型中形成了一个信息流动的循环。DeCom-unit 对特征的初步划分为特征的准确学习提供了良好的数据基础,F2A-unit 模块为模型提供了强有力的频域特征提取能力,并汇集学习到的多维特征输入 SHRUL-unit 模块,采用分层残差学习机制,从而形成完整的特征提取和预测过程。

模块间的协作机制在特征提取上互为补充,确保了 TFD-Hformer 模型能够在面对复杂的退化信号时,充分利用信号的全局周期性、局部变化和时域趋势性,最终提高剩余寿命预测的精度,提高了整体模型的鲁棒性和预测准确性。

3 实验设计与分析

3.1 实验设计

采用 C-MAPSS 公开数据集验证所提的 TFD-Hformer 模型在 RUL 预测任务中的性能。在使用所提出的模型进行预测之外,本实验同时使用了不同类型的模型架构进行对比实验,包括基于

CNN、RNN、基础 Transformer 模型以及近年来较为流行的其他 RUL 预测模型。为保证实验的公平性和结果的客观性,所有对比模型均采用相同的数据集、一致的预处理方法、软硬件环境、训练策略和训练轮次。

3.1.1 C-MAPSS 航空发动机数据集介绍

C-MAPSS 数据集是由美国国家航空航天局提供的一组仿真数据,旨在模拟发动机在不同运行条件下的性能变化及其随时间的退化过程^[15]。如表 3 所示,C-MAPSS 数据集由 4 个子数据集组成,每个子集都模拟了不同的故障模式和运行条件下发动机的退化过程。FD001 和 FD002 中的发动机展示了高压压缩机退化的故障模式,FD003 和 FD004 则同时包括了高压压缩机和风扇退化 2 种故障模式。此外,FD001 和 FD003 的发动机仅在单一运行条件下运行,而 FD002 和 FD004 则在 6 种不同的运行条件下进行测试并记录传感器信息。

表 3 C-MAPSS 数据集概况汇总

Table 3 Summary of the C-MAPSS dataset overview

数据集概况	FD001	FD002	FD003	FD004
训练集样本	100	260	100	249
测试集样本	100	259	100	248
工况数量	1	6	1	6
失效模式	1	1	2	2

C-MAPSS 数据集通过发动机的高保真仿真平台运行生成,涵盖多种运行工况和故障模式,包含传感器监测的多种运行参数及时序数据,为 RUL 预测研究提供了广泛的数据支持,全部的监测传感器基本信息如表 4 所示。

从发动机传感器获得的时间序列数据由多个变量构成,图 7 展示了数据集中编号为 FD001_#002 号样本的部分传感器数据信息。

3.1.2 数据预处理

对收集到的传感器数据进行预处理,处理后的数据将用于模型的训练过程,并依据训练后的模型对测试集样本展开 RUL 预测。预处理阶段的具体流程包括 RUL 计算、归一化和时序数据分割操作。本文以 C-MAPSS 数据集作为实验数据集,将所有的发动机样本 RUL 最大阈值 $y_{\max}$ 设定为 125,则式 (2) 中 RUL 值分段函数的曲线如图 8 所示。

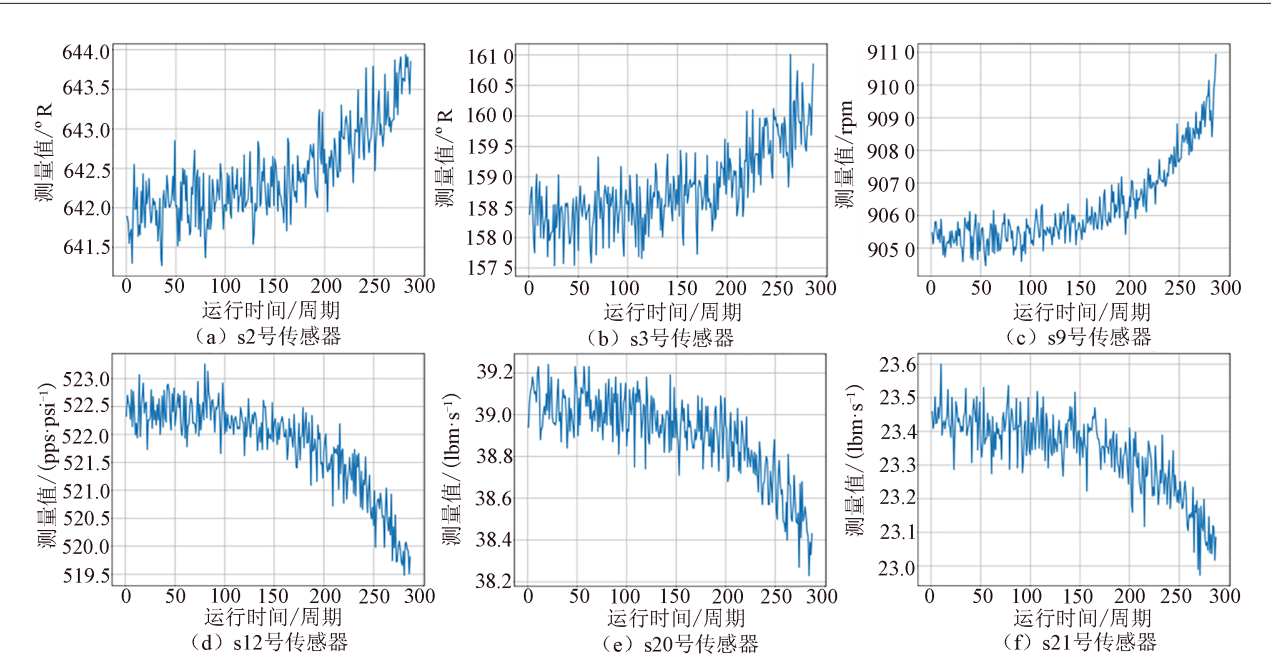


图 7 C-MAPSS数据集部分时间序列数据

Fig. 7 Partial time series data in the C-MAPSS dataset

表 4 C-MAPSS 数据集中的传感器信息		
Table 4 Sensor information in the C-MAPSS dataset		
序号	符号	描述
s1	T2	风扇入口总温/°R
s2	T24	低压压缩机入口总温/°R
s3	T30	高压压缩机入口总温/°R
s4	T50	低压涡轮机出口总温/°R
s5	P2	风扇入口压力/psia
s6	P15	外涵总压/psia
s7	P30	高压压气机出口总压/psia
s8	Nf	风扇物理转速/rpm
s9	Nc	核心机物理速度/rpm
s10	epr	发动机压力比(P50/P2)
s11	Ps30	高压压缩机出口静压/psia
s12	phi	燃油流量与 Ps30 之比/(pps·psi ⁻¹)
s13	NRf	校正风扇转速/rpm
s14	NRc	校正核心转速/rpm
s15	BPR	涵道比
s16	farB	燃烧室油气比
s17	htBleed	排气焓值
s18	Nf-dmd	风扇转速目标值/rpm
s19	PCNfR-dmd	风扇转速的修正目标值/rpm
s20	W31	高压涡轮机冷却液流速/(lbm·s ⁻¹)
s21	W32	低压涡轮机冷却剂流量/(lbm·s ⁻¹)

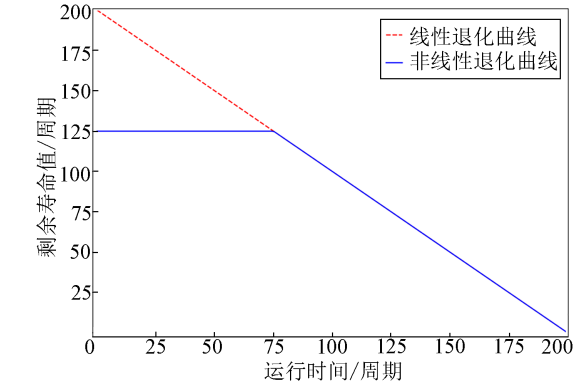


图 8 RUL 值分段函数

Fig. 8 Segmentation function of RUL values

在样本标签 RUL 值的归一化操作中，考虑到 $y_{\max}$ 取值为 125、 $y_{\min}$ 取值为 0，因此，将具体数值代入式 (4) 后得到

$$\bar{y}_i^t = \frac{y_i^t}{125} \tag{19}$$

在计算训练损失及模型评价阶段，由于模型的原始输出为归一化后的 RUL 预测值。因此，本文借助反归一化方法重新将 RUL 预测值转换到原始区间，以在模型的评价中直观地反映模型性能。与归一化方法类似，具体的反归一化方法可表示为

$$y_i^t = 125 \times \bar{y}_i^t \tag{20}$$

经过归一化处理后的数据，仍需要选择恰当的时间窗口进行数据分割，才能变换为可供模型学习和训练的样本。对于数据集 FD001 和 FD003，设置滑动窗口大小 $L=30$ ；对于数据集 FD002 和

FD004, 设置窗口大小 $L=25$ 。在 FD001~FD004 共 4 个数据集上, 滑动窗口的步长 Step 均为 1。在每一个时间窗口下, 以当前窗口的最后一个时刻为基准, 计算 RUL 值, 作为该采样样本的 RUL 标签, 以此构建模型的训练集样本。在预测环节中, 选取发动机当前运行的最后一个时间窗口下的数据作为采样样本, 经过式 (3) 的归一化处理并输入模型, 预测当前发动机的剩余寿命。

3.1.3 软硬件环境设置

实验硬件使用的处理器为 12 vCPU Intel (R) Xeon (R) Silver 4214R, 显卡为 NVIDIA GeForce RTX 3080Ti, 显存为 12 GB, 内存为 90 GB。软件使用 Anaconda 环境、PyTorch 2.3.0、Python 3.12 和 Cuda 12.1。

3.1.4 模型训练策略

在模型训练过程中, 采取标准的监督学习方法。在数据预处理中, 首先对所有数据进行标准化处理, 将输入数据归一化到 $[0, 1]$ 范围内, 以消除不同传感器之间的量纲差异。对原始数据的训练集进行划分, 其中 80% 用于训练模型, 20% 用于在每轮训练后对模型的当前性能进行评估。若干轮次的训练结束后, 使用原始数据的测试集进行模型性能评价。训练过程中, 采用 Adam 优化器进行参数更新。设置最大训练轮次为 50, 采用早停策略, 当验证集的损失连续 10 次未下降时, 自动停止训练。批次大小依据不同的数据集分别设置为 32 或 64, 以平衡计算效率与内存消耗。训练结束后, 冻结模型中的各项参数并使用测试集进行 RUL 预测。

3.2 评价指标

本研究采用均方根误差 (RMSE) 作为评估 RUL 预测模型性能的主要指标。在模型训练阶段, 由于对数据进行了统一的归一化处理。因此评价时, 本模型将输出反归一化至原始 RUL 取值区间, 并以此计算均方根误差, 从而可以量化模型的预测值和实际值之间的差异, 计算表达式为

$$E_{\text{RMSE}} = \sqrt{\frac{1}{N} \sum_{i=1}^N (\hat{y}_i^t - y_i^t)^2} \quad (21)$$

式中: y_i^t 为发动机在第 i 个长度为 t 的时间窗口下的实际 RUL 值; $\hat{y}_i^t$ 为对应模型输出的 RUL 预测值。 E_{RMSE} 值越低, 表明模型的预测性能越优。

3.3 TFD-Hformer 模型的 RUL 预测结果

从以下 2 个方面来评价所提出的模型在 C-MAPSS 数据集上的性能。

1) 模型在所有数据集上的 E_{RMSE} 表现

使用 TFD-Hformer 模型在 FD001~FD004 共 4 个数据集上进行 RUL 预测。具体分析 TFD-Hformer 模型在 RUL 预测任务中的性能, 在 4 个数据集上的 E_{RMSE} 分别为 10.83、17.53、11.42 和 20.56。在模型训练阶段, 所提模型在 RUL 预测结果上能够实现较快且较为稳定的收敛, 基本在 15~20 轮训练后即可展现出良好的预测性能, 相比于同预测精度的深度学习模型有着较快的收敛速度, 图 9 展示了模型在反归一化后计算的损失函数变化情况。

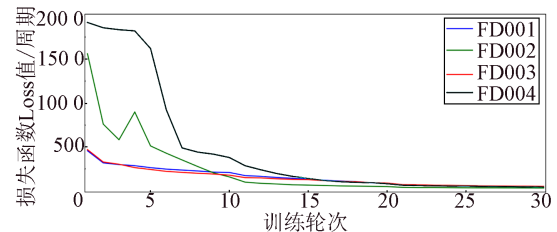


图 9 TFD-Hformer 模型训练的收敛过程

Fig. 9 Convergence process of TFD-Hformer model training

可以看出, 在训练过程中, 模型在相对较少的轮次内即达到良好的预测精度, 体现了其在高效学习特征和收敛方面的能力。这种较快的收敛速度能够适应快速响应的工业环境需求, 在处理设备健康管理任务中具有良好的应用潜力。此外, 相较于一些传统深度学习模型, TFD-Hformer 模型的计算需求较低, 从而在部署时减少了硬件要求和运行成本。

2) 模型在具体数据样本上的 RUL 拟合表现

为更加直观地描述模型预测结果的准确性, 4 个数据集上所有样本的真实 RUL 值和模型预测值汇总如图 10 所示。可以看出, 所提模型的准确度较高, 预测值与真实 RUL 值的偏差较小, 说明模型能够准确反映当前运行时间窗口下样本的运行健康状态。

同时, 为体现模型在预测 RUL 任务中的可靠性, 绘制 4 个数据集上部分样本的全周期 RUL 退化曲线, 如图 11 所示。可以看出, 在发动机运行早期, 其剩余寿命值恒定时, 所提模型预测 RUL 值也较为平稳; 当发动机真实的剩余寿命出现下降趋势时, 模型预测的 RUL 值也基本实现了同步衰减; 验证了模型在反映发动机寿命退化过程的有效性。随着剩余寿命的减少, 模型的拟合也更加精确, 进一步说明模型在学习退化特征方面

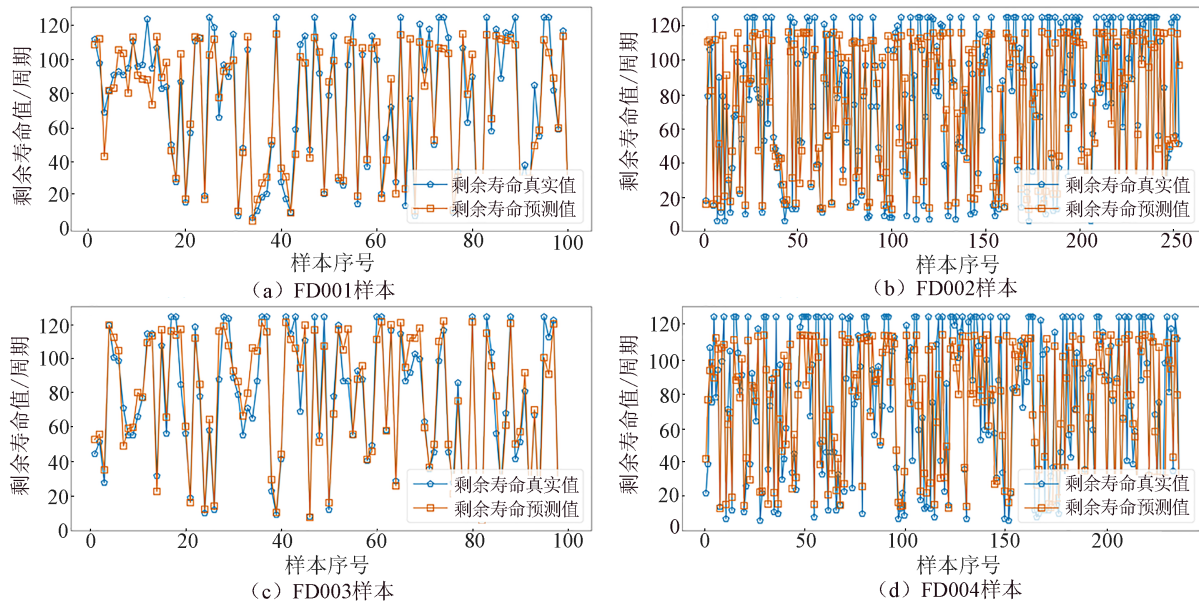


图 10 样本真实RUL值和预测值对比

Fig. 10 Comparison of true and predicted RUL values

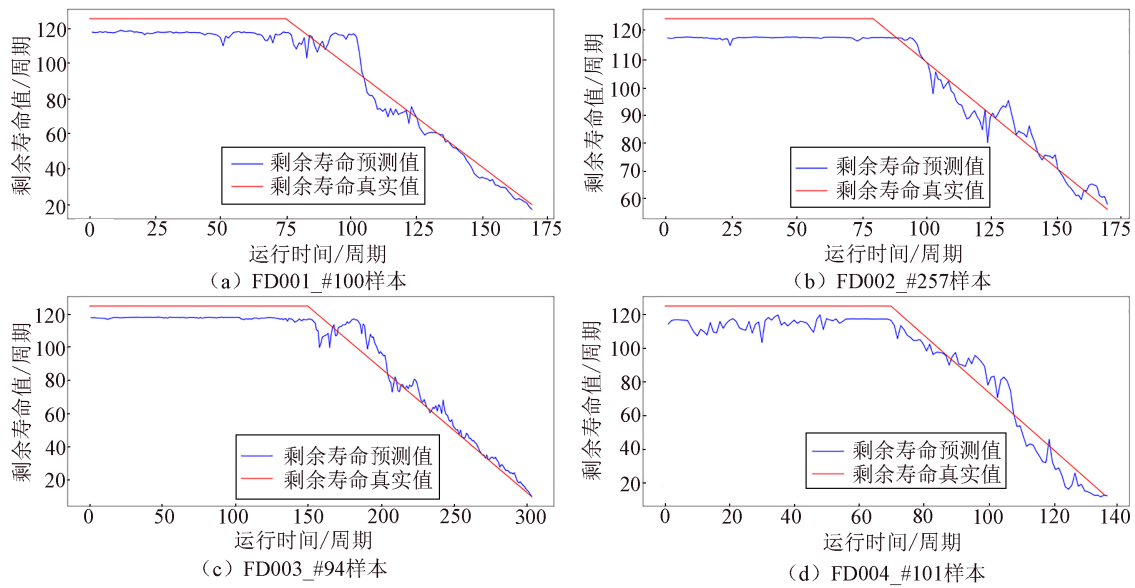


图 11 样本全周期RUL退化曲线

Fig. 11 Degradation curves of full-life RUL

有着良好的表现。

3.4 不同模型的RUL预测结果

为进一步评价所提模型在预测RUL任务中的有效性, 本文开展了其与近年来流行的模型及经典模型的对比实验, 并评价不同模型的RMSE表现, 结果如表5所示。

由表5可以看出, TFD-Hformer模型在FD001和FD002数据集上的RMSE分别为10.83和17.53, 其他模型在FD001、FD002数据集上的RMSE最优表现分别落后本模型0.44、1.41; 在

FD003数据集上GCU-Transformer模型表现与本模型基本持平; 在FD004数据集上, 所提模型在RMSE表现上仅略微落后于BiGRU-TSAM模型。由此可见, 本文所提模型相较于现有的模型解决了深度学习中多尺度特征提取不充分以及处理时序数据的复杂退化过程效率较低的问题, 具备更全面的特征提取能力, 有效避免了信息损失, 使得模型能够更加准确地捕捉退化过程且鲁棒性较强。通过结合时频域增强分解方法, 模型能够在较少的训练轮次内实现良好的预测精度。

表 5 不同模型 RMSE 表现
Table 5 RMSE performances of different models

模型	RMSE			
	FD001	FD002	FD003	FD004
SVR ^[16]	18.28	30.50	21.37	34.11
CNN ^[17]	15.84	28.73	13.53	30.66
LSTM ^[18]	14.91	28.68	13.37	30.32
BiLSTM-ED ^[19]	14.74	22.07	17.48	23.49
AGCNN ^[20]	12.42	19.43	13.39	21.50
BiGRU ^[21]	—	26.54	—	29.13
BiGRU-TSAM ^[22]	12.56	18.94	12.45	20.47
CNN-LSTM-DA ^[23]	14.40	27.23	14.32	26.69
RBM-LSTM-FNN ^[24]	12.56	22.73	12.10	22.66
Transformer	13.32	19.83	13.92	21.88
GCU-Transformer ^[25]	11.27	22.81	11.42	24.86
AdaBN-DCNN	13.17	20.87	14.97	24.57
TFD-Hformer	10.83	17.53	11.42	20.56

注：黑体为各模型 RMSE 的最优数据。

4 结论

针对航空航天领域中航空发动机等重大装备的 RUL 预测问题，提出了一种基于改进 Transformer 架构的 TFD-Hformer 模型。该模型通过结合时域和频域特征，设计了频域增强模块、分解趋势学习模块和分层预测残差学习模块，显著提升了重大装备 RUL 预测的精度和鲁棒性，主要结论如下。

1) TFD-Hformer 模型在 C-MAPSS 数据集上的表现优于传统的深度学习方法。结果表明，模型能够更准确地捕捉复杂的退化过程，尤其是在多工况、多传感器环境下表现出较强的鲁棒性。

2) TFD-Hformer 模型在运行效率方面也有良好的表现。模型训练过程中，其收敛速度较快，训练过程在 15~20 个轮次内便可实现较好的收敛效果。结果表明，所提模型的计算效率较高，为有效适应工业环境中对计算时效性的要求提供了可能。

3) 与现有的 RUL 预测方法相比，TFD-Hformer 模型在处理复杂退化模式和多工况数据时具备良好的表现。通过同时融合时域与频域特征，在多个维度上捕捉设备退化过程中的复杂特征，

改善了现有深度学习方法在特征提取方面的不足。为解决传统方法在多尺度、非线性退化过程建模方面的局限性，提供了可行的解决方案。

未来的研究可以进一步优化模型结构，拓展处理时序数据的思路和方法，以应对更加复杂的退化模式。

参考文献

[1] 袁晓静, 罗伟蓬, 张宏强, 等. 基于区块链的装备维修支持系统设计分析[J]. 火箭军工程大学学报, 2024, 38(1): 54-63.
YUAN X J, LUO W P, ZHANG H Q, et al. Design and development of equipment maintenance support system based on blockchain[J]. Journal of Rocket Force University of Engineering, 2024, 38 (1): 54-63.

[2] BALLI O, CALISKAN H. Turbofan engine performances from aviation, thermodynamic and environmental perspectives[J]. Energy, 2021, 232: 121031.

[3] HUANG P, GU Y K, LI H, et al. An optimal tolerance design approach of robot manipulators for positioning accuracy reliability[J]. Reliability Engineering & System Safety, 2023, 237: 109347.

[4] 段磊光, 霍怡伟, 郭剑锋, 等. 基于交流阻抗法的复合多层石墨烯丁羟推进剂老化监测研究[J]. 火箭军工程大学学报, 2024, 38(3): 66-73.
DUAN L G, HUO Y W, GUO J F, et al. Study on aging monitoring of composite multi-layer graphene HTPB propellant based on AC impedance method[J]. Journal of Rocket Force University of Engineering, 2024, 38 (3): 66-73.

[5] AYGUN H, DURSUN O O, TORAMAN S. Machine learning based approach for forecasting emission parameters of mixed flow turbofan engine at high power modes[J]. Energy, 2023, 271: 127026.

[6] ZHANG Z Z, SONG W, LI Q Q. Dual-aspect self-attention based on transformer for remaining useful life prediction[J]. IEEE Transactions on Instrumentation and Measurement, 2022, 71: 1-11.

[7] 陆文灏, 柯勇伟, 郭永强, 等. 面向多目标优化的航空发动机装配特征选择[J]. 工业工程, 2024, 27(4): 1-8.
LU W H, KE Y W, GUO Y Q, et al. Feature selection for aero-engine assembly using multi-objective optimization[J]. Industrial Engineering Journal, 2024, 27(4): 1-8.

- [8] KRAUS M, FEUERRIEGEL S. Forecasting remaining useful life: interpretable deep learning approach via variational Bayesian inferences[J]. Decision Support Systems, 2019, 125: 113100.
- [9] 朱晓峰, 徐曼菲, 刘治红, 等. 集成主成分分析与随机森林模型的关键工装寿命预测方法[J]. 工业工程, 2023, 26(5): 149-158.
- ZHU X F, XU M F, LIU Z H, et al. Life span prediction of key fixtures integrating principal component analysis and random forest model[J]. Industrial Engineering Journal, 2023, 26(5): 149-158.
- [10] GU Y K, ZENG L, QIU G Q. Bearing fault diagnosis with varying conditions using angular domain resampling technology, SDP and DCNN[J]. Measurement, 2020, 156: 107616.
- [11] 杨静宗, 杨天晴, 吴丽玫. 基于 ITD-SVD 和 MOMEDA 的故障特征提取方法[J]. 工业工程, 2021, 24(6): 48-56.
- YANG J Z, YANG T Q, WU L M. A fault feature extraction method based on ITD-SVD and MOMEDA[J]. Industrial Engineering Journal, 2021, 24(6): 48-56.
- [12] WAHID A, YAHYA M, BRESLIN J G, et al. Self-attention transformer-based architecture for remaining useful life estimation of complex machines[J]. Procedia Computer Science, 2023, 217: 456-464.
- [13] SU X Y, LIU H M, TAO L F, et al. An end-to-end framework for remaining useful life prediction of rolling bearing based on feature pre-extraction mechanism and deep adaptive transformer model[J]. Computers & Industrial Engineering, 2021, 161: 107531.
- [14] ZHUANG L H, XU G Q, DONG B S, et al. Study on performance and mechanisms of a novel integrated model with power & thermal management system and turbofan engine[J]. Applied Thermal Engineering, 2023, 219: 119481.
- [15] SAXENA A, GOEBEL K, SIMON D, et al. Damage propagation modeling for aircraft engine run-to-failure simulation[C]//2008 International Conference on Prognostics and Health Management. Piscataway, NJ, USA: IEEE, 2008: 1-9.
- [16] KHELIF R, CHEBEL-MORELLO B, MALINOWSKI S, et al. Direct remaining useful life estimation based on support vector regression[J]. IEEE Transactions on Industrial Electronics, 2017, 64(3): 2276-2285.
- [17] SATEESH BABU G, ZHAO P L, LI X L. Deep convolutional neural network based regression approach for estimation of remaining useful life[C]//Database Systems for Advanced Applications. Cham: Springer International Publishing, 2016: 214-228.
- [18] XIA T B, SONG Y, ZHENG Y, et al. An ensemble framework based on convolutional bi-directional LSTM with multiple time windows for remaining useful life estimation[J]. Computers in Industry, 2020, 115: 03182.
- [19] YU W N, KIM I Y, MECHEFSKE C. Remaining useful life estimation using a bidirectional recurrent neural network based autoencoder scheme[J]. Mechanical Systems and Signal Processing, 2019, 129: 764-780.
- [20] LIU H, LIU Z Y, JIA W Q, et al. Remaining useful life prediction using a novel feature-attention-based end-to-end approach[J]. IEEE Transactions on Industrial Informatics, 2021, 17(2): 1197-1207.
- [21] HUANG C G, HUANG H Z, LI Y F. A bidirectional LSTM prognostics method under multiple operational conditions[J]. IEEE Transactions on Industrial Electronics, 2019, 66(11): 8792-8802.
- [22] ZHANG J S, JIANG Y C, WU S M, et al. Prediction of remaining useful life based on bidirectional gated recurrent unit with temporal self-attention mechanism[J]. Reliability Engineering & System Safety, 2022, 221: 108297.
- [23] WU Z Y, YU S Y, ZHU X N, et al. A weighted deep domain adaptation method for industrial fault prognostics according to prior distribution of complex working conditions[J]. IEEE Access, 2019, 7: 139802-139814.
- [24] LISTOU ELLEFSEN A, BJØRLYKHAUG E, ILMARÆ, et al. Remaining useful life predictions for turbofan engine degradation using semi-supervised deep architecture[J]. Reliability Engineering & System Safety, 2019, 183: 240-251.
- [25] MO Y, WU Q H, LI X, et al. Remaining useful life estimation via transformer encoder enhanced by a gated convolutional unit[J]. Journal of Intelligent Manufacturing, 2021, 32(7): 1997-2006.

(编辑: 于福兴)