

文章编号: 1006-3080(2026)03-0409-09

DOI: 10.14135/j.cnki.1006-3080.20251023001

融合多尺度时空图和注意力的航空发动机 RUL 预测

吴丽娟, 郑 红, 廖振宏, 徐 贤

(华东理工大学信息科学与工程学院, 上海 200237)

摘要:针对现有航空发动机剩余使用寿命(RUL)预测方法忽略时空融合以及局部与全局特征整合的不足,提出了一种基于多尺度时空图和融合注意力的航空发动机剩余寿命预测方法。在时间维度上,设计短期与长期因果卷积分支,结合注意力机制实现多尺度时间特征融合;在空间维度上,分别建模子系统内局部图与子系统间的超图结构,通过门控机制自适应融合局部与全局特征。最后,引入跨尺度时空注意力模块以强化关键时空依赖,提升预测精度。实验表明,该模型在 C-MAPSS 的多个子数据集上的性能均优于现有主流方法,展现出良好的预测准确性。

关键词:剩余使用寿命预测;时空图神经网络;多尺度融合;注意力机制;故障预测

中图分类号:TP18;V240.2

文献标志码:A

随着航空装备智能化水平的不断提升,剩余使用寿命(RUL)预测已成为航空发动机健康管理领域的研究热点。精准预测发动机的退化趋势与寿命状态,对于提升设备安全性、降低维修成本具有重要意义。

近年来,基于深度学习的模型常用于航空发动机 RUL 预测。如卷积神经网络(CNN)^[1],长短期记忆^[2]等基于深度学习的模型已被广泛应用于航空发动机 RUL 的预测。相较于传统模型,这些模型在捕捉时间序列数据的时间信息方面表现出色。然而,它们无法很好地捕捉空间依赖性^[3-4],即航空发动机传感器之间的相关性,这限制了它们在表示学习中的表现^[5]。

图神经网络因其具有挖掘数据空间关系的能力,被用于航空发动机剩余使用寿命的预测。Zhang 等^[6]提出了一种自适应时空图信息融合框架,学习传感器信号的空间特性和时间特性,使用扩张卷积解决长信号过度平滑问题,提高预测准确性。Jia 等^[7]提出基于领域泛化的多视图时空图卷积网络,利用图卷积和时间卷积分别提取空间拓扑特征与睡眠阶段转换规律,并引入注意力机制聚焦关键时空信息,

提取不受个体影响的睡眠特征。Xiao 等^[8]搭建 CNN 和双向门控循环单元(BiGRU)的混合模型,CNN 负责从多传感器数据中提取局部特征,BiGRU 能同时捕捉数据的正向和反向时间关联。Cao 等^[9]使用双通道架构提取退化特征,图神经网络用于建模发动机传感器数据之间的复杂关联性,设计自适应通道分别处理短期动态退化特征和长期静态退化趋势。

现有图模型仍面临两个关键挑战:一是无法有效建模不同时间尺度上的退化过程,如短期振动与长期性能衰减的交织影响;二是空间结构建模粒度单一,既缺乏对子系统局部耦合的刻画,也难以捕捉系统级跨子系统协同作用。

针对上述挑战,本文提出基于多尺度时空图和融合注意力的 RUL 预测方法,提出时间和空间的多尺度特征提取模块,引入多尺度的时空注意力机制,有效捕捉关键时空交互关系,从而实现对航空发动机 RUL 的更精确预测,本文模型在商用模块化航空推进系统仿真数据集^[10]C-MAPSS 上进行了实验分析和验证。

收稿日期: 2025-10-23

基金项目: 上海市 2024 年度“科技创新行动计划”资助(24BC3200500, 24BC3200300)

作者简介: 吴丽娟(2000—),女,浙江人,硕士生,主要研究方向为人工智能、航空发动机剩余寿命预测。E-mail: Y80230110@mail.ecust.edu.cn

通信联系人: 郑 红, E-mail: zhenghong@ecust.edu.cn

引用本文: 吴丽娟,郑 红,廖振宏,等.融合多尺度时空图和注意力的航空发动机 RUL 预测[J].华东理工大学学报(自然科学版),2026,52(3): 409-417.

Citation: WU Lijuan, ZHENG Hong, LIAO Zhenhong, *et al.* Remaining Useful Life Prediction for Aircraft Engines Based on Multi-Scale Spatio-Temporal GNN and Fusion Attention[J]. Journal of East China University of Science and Technology, 2026, 52(3): 409-417.

1 模型相关工作

模型的总体架构如图 1 所示。该模型接收传感器序列数据 X_t ，它包含多个通道的传感器数据，将原始数据输入滑动窗进行降噪处理，并设计两个多尺度特征提取模块分别进行特征提取，将提取后的时空特征经过多尺度的时空注意力进行融合，随后用于 RUL 预测。

1.1 时间多尺度特征提取模块

为有效建模设备运行过程中存在的时间多尺度动态行为，提出一个双分支的时间多尺度模块，分别面向不同时间粒度建模信号特征，并在此基础上引入注意力机制来实现尺度间信息的自适应融合，以增强模型对复杂时间依赖的表达能力。

模型接收来自多个传感器的观测数据，设输入窗口记作 $X_t \in \mathbb{R}^{S \times T}$ ，其中 S 表示传感器数量， T 为时间窗口长度，该时间片段被分别输入至短期与长期两个分支。短期分支采用膨胀率为 1 的因果卷积操作，用于捕捉近邻时间步间的局部变化模式，即高频动态。该过程形式化表达为：

$$T_s = \text{Conv}(X_t; K, d=1, C) \quad (1)$$

卷积核大小设为 K ，输出通道为 C 。长期分支则

通过膨胀率为 8 的卷积网络扩大感受野，从而捕捉更长时间尺度上的演化趋势，如缓慢的性能衰退或周期性变化。其操作可表示为：

$$T_l = \text{Conv}(X_t; K, d=8, C) \quad (2)$$

通过该设计，长期分支能够在保持卷积核大小不变的情况下扩展时间感知范围，有效建模长距离依赖结构。在获得两个尺度下的特征表示后，需对其进行融合，融合表示为：

$$T_{\text{concat}} = [T_s, T_l] \in \mathbb{R}^{S \times T \times 2C} \quad (3)$$

$$\alpha_{\text{scale}} = \text{softmax}(W_a \cdot \tanh(T_{\text{concat}}) + b_a) \quad (4)$$

其中， W_a 与 b_a 为可学习参数， α_{scale} 表示每个时间步与每个传感器通道在两个尺度上的注意力得分。最终融合后的时间特征通过加权求和获得：

$$Z_{\text{temporal}} = \alpha_1 \cdot T_s + \alpha_2 \cdot T_l \quad (5)$$

其中， α_1 、 α_2 分别对应短期与长期分支的注意力权重。融合后输出的时间特征不仅保留了短期的敏感响应，也兼顾了长期的趋势演化。

1.2 空间多尺度特征提取模块

该模块的输入为时间多尺度模块处理前的原始时序特征的某一时间片段 $X_t \in \mathbb{R}^{S \times d}$ ，其中 d 表示每个传感器的特征维度。在这一时间截面下，空间多尺度模块分为两个并行分支：局部子图建模与全局

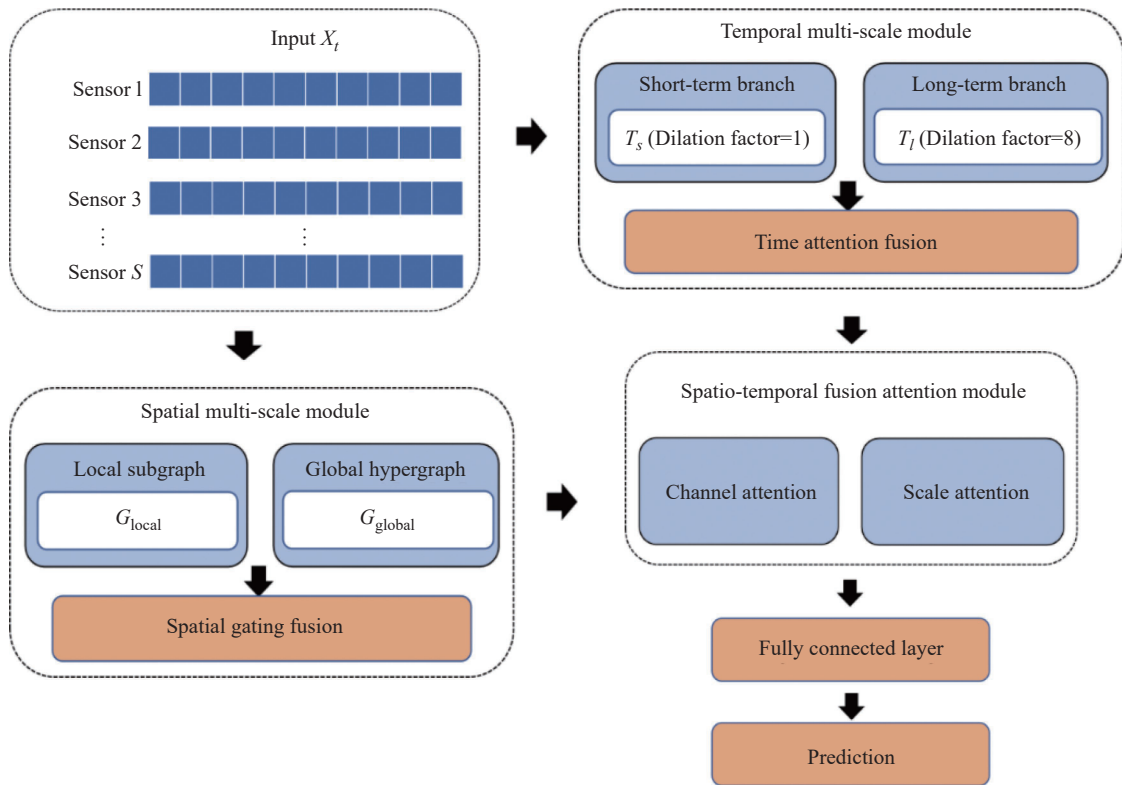


图 1 模型架构

Fig. 1 Architecture of model

超图建模。在局部建模分支中,传感器节点按物理子系统(如压气机、燃烧室、涡轮等)进行分组,每组构建局部邻接图 $A_{\text{local}}^{(g)}$,图中的边表示组内传感器之间的物理连接或统计相关性。对于每一组 g ,使用图卷积网络对其特征进行编码:

$$Z_{\text{local}}^{(g)} = \sigma(\tilde{D}^{-\frac{1}{2}} \tilde{A}_{\text{local}}^{(g)} \tilde{D}^{-\frac{1}{2}} X_t^{(g)} W_{\text{local}}) \quad (6)$$

其中, $\tilde{A}_{\text{local}}^{(g)} = A_{\text{local}}^{(g)} + I$,为加自环邻接矩阵; $\tilde{D}$ 为度矩阵, W_{local} 为可训练参数。所有局部子图的输出再进行拼接,形成整体局部特征:

$$Z_{\text{local}} = \text{Concat}(Z_{\text{local}}^{(1)}, \dots, Z_{\text{local}}^{(G)}) \quad (7)$$

其中, Z_{local} 为局部子图特征, $Z_{\text{local}}^{(G)}$ 为第 G 个子系统的特征。

另一方面,全局建模分支通过构建超图结构来建模跨组别的高阶空间依赖。将每一个子系统抽象为一个超边,连接其下所有传感器节点,构建超图关联矩阵 $H_{\text{hyper}} \in \mathbb{R}^{S \times E}$,其中 E 为超边数量。利用超图卷积传播公式提取跨组全局特征,其中 D_v 、 D_e 分别为节点与超边的度矩阵, W_{global} 为超图路径的可学习参数。

$$Z_{\text{global}} = \sigma(D_v^{-\frac{1}{2}} H_{\text{hyper}} D_e^{-1} H_{\text{hyper}}^T D_v^{-\frac{1}{2}} X_t W_{\text{global}}) \quad (8)$$

其中, Z_{global} 为全局超图特征。

并行提取局部与全局特征之后,为了实现多尺度空间信息的有效融合,引入空间门控机制进行自适应加权。首先将两个分支输出进行拼接,并通过门控网络计算融合权重:

$$\beta = \sigma(W_g [Z_{\text{local}}; Z_{\text{global}}] + b_g) \quad (9)$$

其中, $\beta \in \mathbb{R}^{S \times 1}$,为每个节点的门控因子, W_g 和 b_g 为训练参数。最终融合后的空间特征表达为:

$$Z_{\text{spatial}} = \beta \odot Z_{\text{local}} + (1 - \beta) \odot Z_{\text{global}} \quad (10)$$

其中, Z_{spatial} 为最终融合后的空间特征。

该结构实现了从局部耦合到系统协同的空间信息整合,具备良好的可解释性与泛化能力,特别是在航空发动机等结构复杂、传感器布局异构的系统中,能够精准捕捉影响 RUL 预测的关键空间特征。

1.3 时空融合注意力模块

该模块以时间多尺度模块输出的时间特征 $Z_{\text{temporal}} \in \mathbb{R}^{T \times d_t}$ 与空间多尺度模块输出的空间特征 $Z_{\text{spatial}} \in \mathbb{R}^{S \times d_s}$ 为输入,其中 T 表示时间窗口长度, d_t 、 d_s 分别为两个分支的特征维度。为了对接后续统一建模过程,首先对两路特征进行线性变换,使其投影到统一的维度 d 上:

$$\tilde{Z}_{\text{temporal}} = Z_{\text{temporal}} W_t, \tilde{Z}_{\text{spatial}} = Z_{\text{spatial}} W_s \quad (11)$$

其中, $W_t \in \mathbb{R}^{d_t \times d}$ 、 $W_s \in \mathbb{R}^{d_s \times d}$,均为可学习参数。为了有效捕捉时空之间的交互模式,模块引入一种跨尺度交叉注意力机制。通过构建时间特征对空间特征的注意力映射与空间特征对时间特征的注意力映射,形成双向融合策略。令 $\tilde{Z}_{\text{temporal}}$ 为查询项 Q ,并将空间特征 $\tilde{Z}_{\text{spatial}}$ 通过线性变换映射为键 K 与值 V ,可计算注意力权重和融合表示:

$$A_{\text{time} \rightarrow \text{space}} = \text{Softmax}\left(\frac{QK^T}{\sqrt{d}}\right) \in \mathbb{R}^{T \times S} \quad (12)$$

$$F_{ts} = A_{\text{time} \rightarrow \text{space}} V \quad (13)$$

其中, F_{ts} 为时间到空间方向的融合特征。

空间引导时间中,令空间特征为查询项 Q' ,时间特征通过线性变换映射为键 K' 和值 V' ,得到最终融合特征表示:

$$A_{\text{space} \rightarrow \text{time}} = \text{Softmax}\left(\frac{K'Q'^T}{\sqrt{d}}\right), F_{st} = A_{\text{space} \rightarrow \text{time}} V' \quad (14)$$

$$Z_{\text{fused}} = \text{Concat}(F_{ts}, F_{st}) W_f \quad (15)$$

为了提升模型对于关键时间区段和关键节点区域的聚焦能力,模块进一步引入了通道注意力与尺度注意力机制,分别用于强化具有判别性的特征通道与时空尺度维度。通道注意力通过全局平均池化与 Sigmoid 激活函数生成通道权重 $\alpha_{\text{channel}} \in \mathbb{R}^d$,对融合特征逐通道调节:

$$Z'_{\text{fused}} = \alpha_{\text{channel}} \odot Z_{\text{fused}} \quad (16)$$

而尺度注意力则为时间路径和空间路径分别引入权重 $\lambda_t, \lambda_s \in [0, 1]$,其值通过小型前馈神经网络自适应学习,控制两个分支在最终融合中的占比,满足 $\lambda_t + \lambda_s = 1$:

$$Z_{\text{final}} = \lambda_t \cdot F_{ts} + \lambda_s \cdot F_{st} \quad (17)$$

该模块的引入,显著增强了预测结果的准确性与鲁棒性,为复杂工业设备的健康管理提供了更具实用价值的技术支持。

2 结果与分析

2.1 数据集描述

本文采用 C-MAPSS 进行仿真实验。该数据集对分布于发动机不同部位的 21 个传感器状态参数进行采集,包括速度、压力、温度等参数,包含 4 个子集 (FD001, FD002, FD003, FD004),数据规模和具体信息详见表 1。

2.2 数据预处理

由于不同传感器退化数据与剩余使用寿命标签

表 1 数据集详细信息
Table 1 Details of datasets

Dataset	Training unit	Testing unit	Fault mode	Operating condition	Training instance	Test instance
FD001	100	100	1	1	20 631	13 096
FD002	260	259	1	6	53 759	33 991
FD003	100	100	2	1	24 720	16 596
FD004	249	248	2	6	61 249	41 214

的相关性存在显著差异,采用可视化技术进行相关性分析,筛选并剔除与 RUL 关联性较弱的传感器,具体相关性如图 2 所示,可观察到传感器 Sensor 1、Sensor 5、Sensor 6、Sensor 10、Sensor 16、Sensor 18 和 Sensor 19 在整个发动机生命周期中保持恒定数值。因此,选定其余传感器作为模型输入。

在实际应用中,当新型航空发动机投入使用时,其性能在初期阶段稳定,性能下降轻微,随着设备持续运转,老化进程会逐渐加速,导致设备剩余寿命呈现递减趋势。因此,本文采用分段线性模型,将健康状态阶段的 RUL 标签阈值设定为 125。

2.3 评价指标和实验环境

本文采用均方根误差 (RMSE) 和评分函数 (Score) 作为评估模型性能的指标。

模型基于 Python 3.8 和 Pytorch 1.9 构建,并在 NVIDIA GeForce RTX 4060 GPU 上进行训练,在模型训练过程中,批量大小设置为 100,选择自适应矩估计优化器(ADAM),初始学习率设置为 0.001,训练模型设置 40 个周期。

2.4 对比实验

为了证明所提出模型的优越性,将本文方法与以下几种先进的方法进行了比较,包括 HierCorrPool^[11]、GCN^[12]、InFormer^[13]、AConvLSTM^[14]、MAGNN^[15]、AGCNN^[16]、DCNN^[17]、FA-GRU-GAN^[18]、BACE-RUL^[19]、CfC-F2RN^[20]、AD-TCN-MSCDIM^[21]、HDA-GNN^[22]、CNN-BiGRU-Attention^[8]、DCDM^[9]。表 2 对比了不同方法在 C-MAPSS 数据集上的实验结果,结果显示,本文方法在 C-MAPSS 数据集的 FD001、FD002、FD003、FD004 4 个子集上分别取得 11.23、12.64、11.15、13.41 的均方根误差值;对应评分值分别为 152、673、187 和 789。相较于每个子数据集上对应的 RMSE 次优的方法,本模型在 FD001、FD002 和 FD004 子集上分别实现 1.40%、0.71% 和 3.25% 的 RMSE 提升;在 FD001、FD002 和 FD004 子集上评分分别提升 10.59%、4.54% 和 7.61%。

为直观展示所提方法在 C-MAPSS 数据集上的

预测能力,本文通过可视化对比了模型在 FD001 和 FD003 两个数据集上的实际 RUL 与预测使用寿命,如图 3 所示。

本文方法通过引入多尺度的时间卷积和空间图卷积,分别捕捉短期突变与长期退化趋势、局部物理耦合与全局系统协同效应,显著增强了对复杂退化模式的表征能力。在时间多尺度特征提取方面,传统 LSTM 或 TCN 仅依赖单一时间尺度,而本文模型的双分支设计能更全面覆盖退化过程的复杂动态。在空间多尺度特征提取方面,普通 GCN 或 CNN 仅用单一邻接矩阵,无法区分局部和全局空间依赖,而本文模型的多层级结构更贴合发动机的真实物理拓扑。尽管本文模型整体优势明显,但在 FD003 上略逊于 AD-TCN-MSCDIM,原因可能在于 FD003 包含两种故障,其退化路径可能相反,模型空间图结构未显示区分故障类型。其次,FD003 训练样本仅 100 台,少于 FD002 和 FD004,模型可能对少数故障模式欠拟合。

2.5 消融实验结果及分析

为验证所提出的两个多尺度特征提取模块和融合注意力的有效性,对本文模型进行消融实验,如表 3 所示。

可以发现,将时间多尺度特征提取模块加入到基线 GCN 模型,在 FD001、FD002、FD003 和 FD004 4 个子集中 RMSE 分别降低了 2.78%、3.63%、1.51% 和 3.19%, Score 分别降低了 7.59%、6.01%、7.80% 和 5.58%。这些结果表明,时间多尺度模块可以提取不同膨胀因子下的多尺度特征,有效捕捉长短期依赖关系。类似的,将空间多尺度特征提取模块加入到基线 GCN 模型中,在 FD001、FD002、FD003 和 FD004 4 个子集中 RMSE 分别降低了 1.03%、1.81%、0.25% 和 1.45%, Score 分别降低了 3.80%、3.53%、4.13% 和 2.48%,说明空间多尺度模块可以提取局部子图和全局超图之间的依赖关系,更好地捕捉空间关系。将融合注意力模块加入到多尺度时空图中,使得模型在 FD001、FD002、FD003 和 FD004 4 个子集中的表现显著提高, RMSE 分别降低了 10.73%、8.27%、6.46% 和 7.14%, Score 分别降低了 35.86%、20.73%、14.22% 和 18.41%,表明各个模块对预测模型效果的有效性。

2.6 超参数实验结果及分析

2.6.1 片段大小实验分析 在实际训练过程中,将每个多元时间序列样本分割为多个片段,用于数据特征提取,为评估片段大小对预测精确度的影响,进行了片段大小实验分析,由于 C-MAPSS 数据集样本时

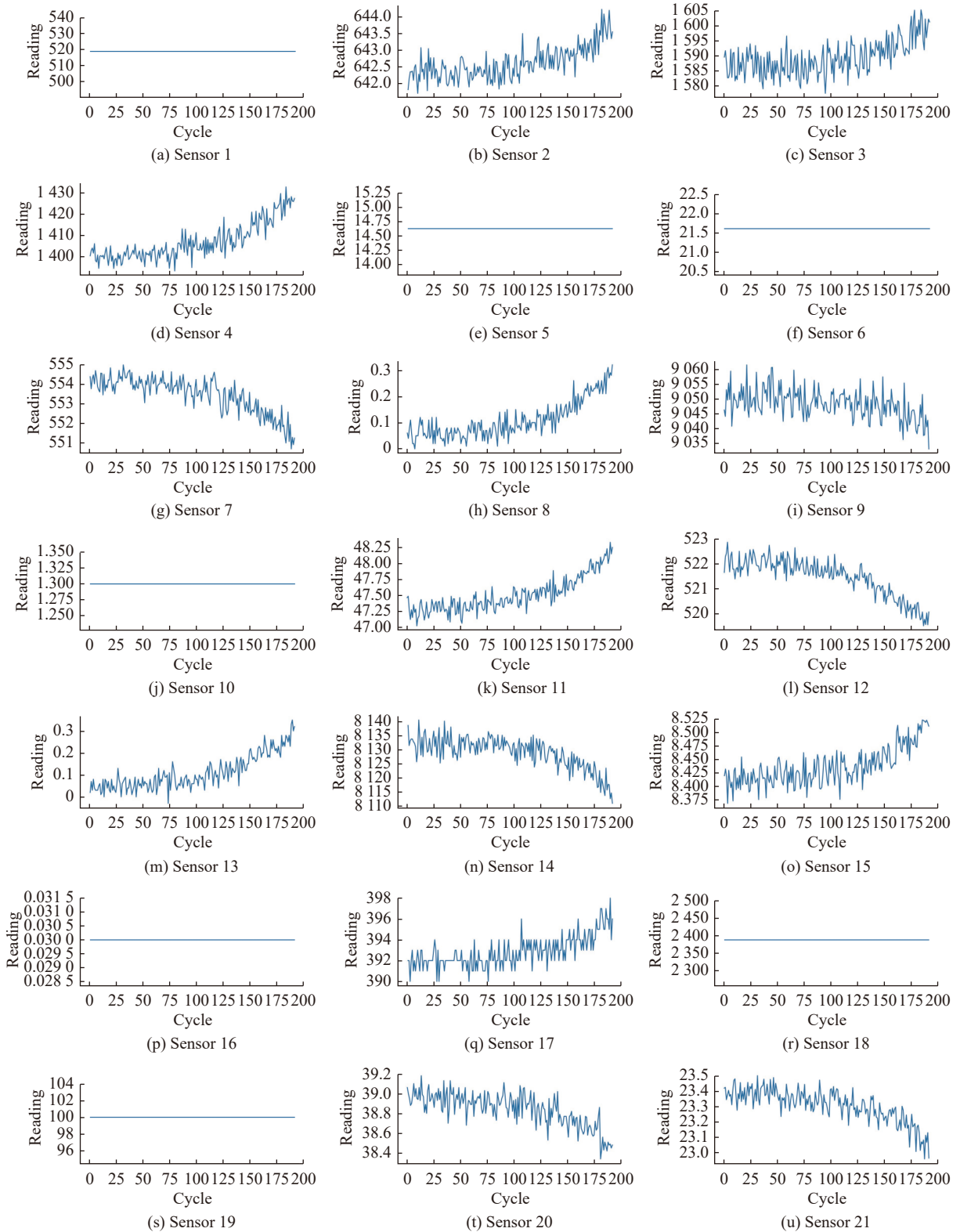


图2 传感器数值在航空发动机全生命周期中的变化趋势

Fig. 2 Variation trend of sensor values during the full life cycle of aero-engines

间长度相对较短,选择在 $[2, 4, 6, 8, 10]$ 范围内选取较小片段进行样本分割,结果如图4所示。由图可知,相对较小的片段尺寸更有利于获得更优性能,这一趋势在各子数据集中普遍存在,当片段尺寸设置为4或6时,通常能获得最优表现。

2.6.2 滑动窗口大小实验分析 滑动窗口定义了模型在每次前向传播中接收的输入序列长度,即每条输入样本中包含的连续时间步数量。每个时间步对应一次发动机的传感器采样,该参数直接影响模型对历史退化趋势的建模能力。

表 2 本文模型与基线模型性能比较
Table 2 Performance comparison of baseline models and ours

Dataset	FD001		FD002		FD003		FD004	
	RMSE	Score	RMSE	Score	RMSE	Score	RMSE	Score
HierCorrPool ^[11]	12.64	227	13.23	709	12.30	220	13.86	854
GCN ^[12]	12.58	237	13.78	849	11.92	218	14.44	967
InFormer ^[13]	13.13	263	13.20	715	12.58	228	14.16	1023
AConvLSTM ^[14]	13.10	286	13.11	737	12.13	276	14.64	1011
MAGNN ^[15]	12.63	246	13.09	714	12.15	253	14.30	978
AGCNN ^[16]	12.42	225	19.43	1492	13.39	227	21.05	3392
DCNN ^[17]	12.61	273	22.36	4020	12.64	284	23.31	5027
FA-GRU-GAN ^[18]	13.12	259	15.19	1047	13.24	278	14.32	1002
BACE-RUL ^[19]	21.81	72099	23.19	340131	29.36	94867	26.41	710811
CfC-F2RN ^[20]	11.74	198	12.75	705	11.10	195	14.25	863
AD-TCN-MSCDIM ^[21]	11.39	170	13.41	763	10.17	134	15.28	1483
HDA-GNN ^[22]	12.25	244	19.06	1343	12.56	274	18.27	2215
CNN-BiGRU-Attention ^[8]	11.69	189	12.73	689	14.89	325	18.20	563
DCDM ^[9]	11.75	195	14.98	1539	11.23	198	16.15	1750
Ours	11.23	152	12.64	673	11.15	187	13.41	789

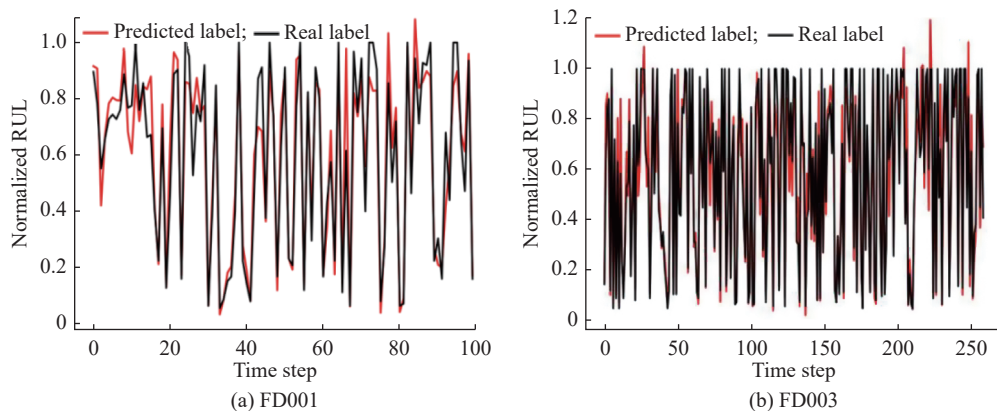


图 3 FD001 和 FD003 数据集上的剩余使用寿命预测

Fig. 3 Remaining useful life prediction on the FD001 and FD003 datasets

表 3 不同模块的消融实验

Table 3 Ablation study of different modules

Dataset	FD001		FD002		FD003		FD004	
	RMSE	Score	RMSE	Score	RMSE	Score	RMSE	Score
GCN	12.58	237	13.78	849	11.92	218	14.44	967
GCN+MS-TG	12.23	219	13.28	798	11.74	201	13.98	913
GCN+MS-SG	12.45	228	13.53	819	11.89	209	14.23	943
Ours	11.23	152	12.64	673	11.15	187	13.41	789

MS-TG—Multi-scale temporal graph; MS-SG—Multiscale spatial graph

为了评估滑动窗口大小对模型性能的影响, 本文在保持其他超参数不变的前提下, 设置滑动窗口

长度为 [10, 20, 30, 40, 50, 60], 并在 4 个 C-MAPSS 子数据集上分别进行实验, 比较不同设置下的预测精

度。模型在每次输入中以固定窗口长度截取传感器序列,通过多尺度时空图融合模块进行特征提取,最终输出对应时间步的RUL预测值,如图5所示。

实验结果表明,较小窗口可能导致信息不足,难以捕捉全局的退化特征,而过大窗口可能引入噪声或冗余,影响模型泛化能力。综合各项指标,本文最终选择窗口大小40作为默认配置,在各子数据集间表现稳定。

2.6.3 膨胀因子实验分析 为了进一步验证膨胀因子的合理性与敏感性,本文在多个数据集上开展了超参数实验。分别设置长时间分支的膨胀率[2, 4,

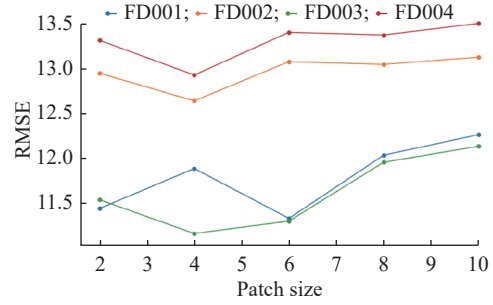


图4 划分片段大小实验分析

Fig. 4 Analysis of segment size experiments

6, 8, 10],短时间分支固定为1,实验结果如图6所示。

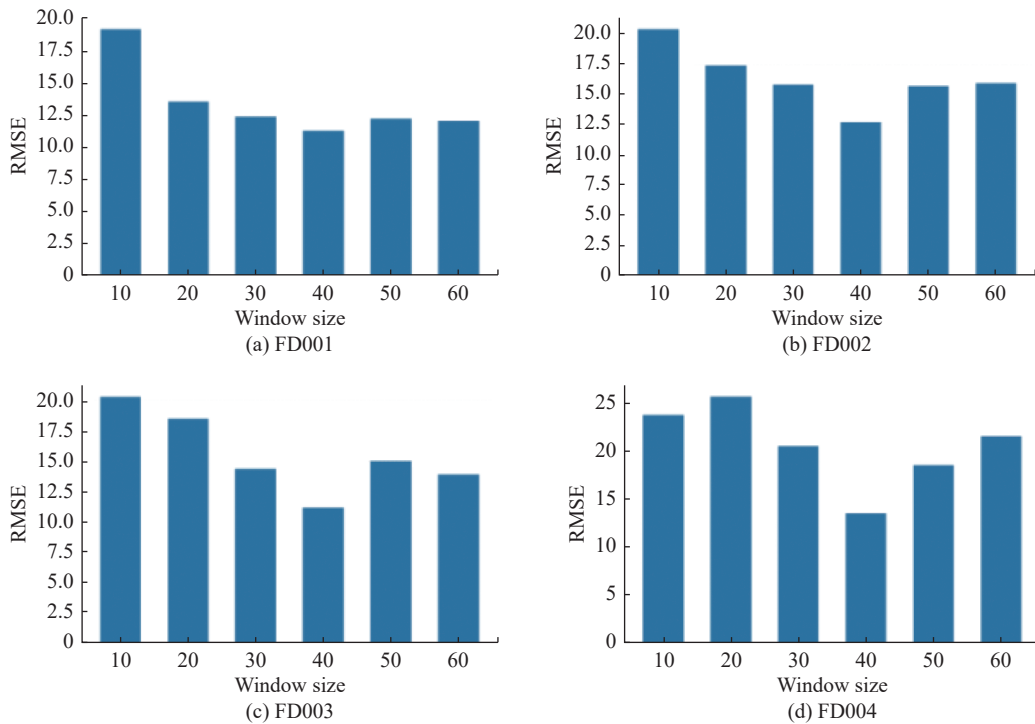


图5 滑动窗口尺寸实验分析

Fig. 5 Analysis of sliding window size experiments

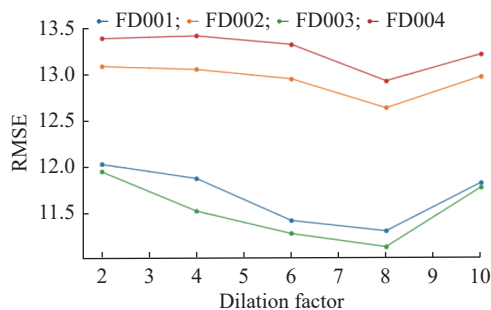


图6 膨胀因子大小实验分析

Fig. 6 Analysis of expansion factor size experiments

由图6可知,膨胀率越大,单层卷积感受野越广,模型对长期退化趋势的建模能力越强。然而,过大的膨胀率可能造成特征提取稀疏化,导致信息遗

漏。考虑到性能与稳定性,本文最终选择长时间分支的膨胀率为8,在主流数据集上获得最优表现。

3 结 论

本文提出的基于多尺度时空图和融合注意力模型进一步提升了航空发动机RUL预测的准确率。该模型将多尺度因果卷积引入时间建模通道,利用不同膨胀因子分别建模高频瞬变特征与长期退化趋势;同时,构建局部子图与全局超图的空间多尺度图结构,捕捉发动机不同子系统之间的物理关联与协同演化机制。通过融合注意力机制对多尺度时空特征进行交互建模,有效提升模型对关键退化模式的

感知能力与预测精度,为航空发动机健康管理提供了新的解决方案。本研究未来工作将重点关注多工况和复杂故障场景下的模型预测能力,并结合物理方法分析发动机退化原理,提高模型的可解释性和泛化能力。

参考文献:

- [1] GAURAV D, SEETARAM M. Fault diagnosis of machines using deep convolutional Beta-variational autoencoder[J]. IEEE Transactions on Artificial Intelligence, 2021, 3(2): 287-296.
- [2] 郭津延, 郑红, 杜佳宇, 等. 基于双尺度自适应令牌注意力的交通流量预测[J]. 华东理工大学学报(自然科学版), 2025, 51(6): 817-826.
- [3] RAMI N K, ANGKOON P, ALI H A, *et al.* Recursive multi-signal temporal fusions with attention mechanism improves EMG feature extraction[J]. IEEE Transactions Artificial Intelligence, 2020, 1(2): 139-150.
- [4] 孙俊静, 顾幸生. 基于注意力机制多尺度卷积神经网络的轴承故障诊断[J]. 华东理工大学学报(自然科学版), 2024, 50(2): 247-256.
- [5] ZHENG H C, MIN W, RUI Z, *et al.* Machine remaining useful life prediction via an attention-based deep learning approach[J]. IEEE Transactions on Industrial Electronics, 2021, 68(3): 2521-2531.
- [6] ZHANG Y X, LI Y X, WANG Y L, *et al.* Adaptive spatio-temporal graph information fusion for remaining useful life prediction[J]. IEEE Sensors Journal, 2022, 22(4): 3334-3347.
- [7] JIA Z Y, LIN Y F, WANG J, *et al.* Multi-view spatial-temporal graph convolutional networks with domain generalization for sleep stage classification[J]. Computing Research Repository, 2021, 29: 1977-1986.
- [8] XIAO H, WANG J H, CENG M J, *et al.* A CNN-BiGRU predictive maintenance method for industrial equipment based on attention mechanism[C]//2025 7th International Conference on Artificial Intelligence Technologies and Applications (ICAITA). Wenzhou, Zhejiang, China: IEEE, 2025: 88-91.
- [9] CAO L A, LI Y F, CHEN J W, *et al.* Dual-channel degradation monitoring based on graph neural network for aero-engine remaining useful life prediction[J]. IEEE Transactions on Instrumentation and Measurement, 2025, 74: 1-19.
- [10] ABHINAV S, KAI G, DON S, *et al.* Damage propagation modeling for aircraft engine run-to-failure simulation[C]//IEEE Conference on Prognostics and Health Management. Denver, CO, USA: IEEE, 2008: 1-9.
- [11] YU C W, MIN W, XIAO L L, *et al.* Multivariate time-series representation learning via hierarchical correlation pooling boosted graph neural network[J]. IEEE Transactions on Artificial Intelligence, 2024, 5(1): 321-333.
- [12] THOMAS N, MAX W. Semi-supervised classification with graph convolutional networks [EB/OL]. (2017-02-22). <https://doi.org/10.48550/arxiv.1609.02907>.
- [13] ZHOU H Y, ZHANG S H, PENG J Q, *et al.* Informer: Beyond efficient transformer for long sequence time-series forecasting [C]//AAAI Conference on Artificial Intelligence. California, USA: AAAI Press, 2021: 11106-11115.
- [14] YU T X, HONG S Y, YU D Z, *et al.* A dual-stage attention-based ConvLSTM network for spatio-temporal correlation and multivariate time series prediction[J]. International Journal of Intelligent Systems, 2021, 36(5): 2036-2057.
- [15] LING C, DONG H C, ZONG J S, *et al.* Multi-scale adaptive graph neural network for multivariate time series forecasting[J]. IEEE Transactions on Knowledge and Data Engineering, 2023, 35(10): 10748-10761.
- [16] HUI L, ZHEN Y L, WEI Q J, *et al.* Remaining useful life prediction using a novel feature-attention-based end-to-end approach[J]. IEEE Transactions on Industrial Informatics, 2020, 17(2): 1197-1207.
- [17] XIANG L, QIAN D, JIAN Q S. Remaining useful life estimation in prognostics using deep convolution neural networks[J]. Reliability Engineering & System Safety, 2017, 172: 1-11.
- [18] 袁烨, 黄虹, 程骋, 等. 基于特征注意力机制的 GRU-GAN 航空发动机剩余寿命预测[J]. 中国科学: 技术科学, 2022, 52(1): 198-212.
- [19] ZHANG Z K, LI D, WU S Y, *et al.* BACE-RUL: A bi-directional adversarial network with covariate encoding for machine remaining useful life prediction[C]//Collaborative Computing: Networking, Applications and Worksharing, Cham: Springer, 2025: 473-490.
- [20] XUAN L L, YA W H, HANG W, *et al.* A closed-form continuous-depth neural-based hybrid difference features re-representation network for RUL prediction[J]. Reliability Engineering & System Safety, 2025, 253: 110540.
- [21] ZHE L, BING L, CHANG Y F, *et al.* A Multi-scale cross-dimension interaction approach with adaptive dilated TCN for RUL prediction[J]. Scientific Reports, 2025, 15(1): 19229.
- [22] JIN Y, MENG Z G, HU Q H, *et al.* Heterogeneous dynamic-aware GNN for RUL prediction of aero-engine[C]//IEEE Transactions on Instrumentation and Measurement. [s.l.]: IEEE, 2025: 1-14.

Remaining Useful Life Prediction for Aircraft Engines Based on Multi-Scale Spatio-Temporal GNN and Fusion Attention

WU Lijuan, ZHENG Hong, LIAO Zhenhong, XU Xian

(School of Information Science and Engineering, East China University of Science and Technology, Shanghai 200237, China)

Abstract: Remaining Useful Life (RUL) prediction is a critical task in the health management of aircraft engines. To effectively model the complex temporal dynamics and spatial structural information, this paper proposes a novel RUL prediction method based on Multi-Scale Spatio-temporal Graphs with Fusion Attention (MSFA-STG). In the temporal dimension, we design parallel branches of short-term and long-term causal convolutions, which are fused using an attention mechanism to capture multi-scale temporal features. In the spatial dimension, both local graphs within subsystems and hypergraph structures across subsystems are constructed, and a gating mechanism is adopted to adaptively fuse local and global spatial features. Furthermore, a cross-scale spatio-temporal attention module is introduced to enhance the modeling of critical spatio-temporal dependencies, thereby improving prediction accuracy. Experimental results on the C-MAPSS dataset demonstrate that the proposed model outperforms mainstream methods, showing strong generalization and high predictive accuracy.

Key words: remaining useful life prediction; spatio-temporal graph neural network; multi-scale fusion; attention mechanism; fault prognostics

(责任编辑: 李娟)