

DOI:10.19651/j.cnki.emt.2518881

基于多尺度的双轴注意力 GCN 剩余寿命预测模型^{*}郑森潇¹ 郭志涛^{1,2} 李义博¹ 负智¹

(1.河北工业大学电子信息工程学院 天津 300130;2.河北工业大学创新研究院(石家庄) 石家庄 050299)

摘要: 准确预测设备的剩余使用寿命(RUL)能够优化维护策略、降低成本并提高效率。然而现有的方法大多依赖于分离式地提取时间特征与空间特征,无法充分融合时间信息与空间信息。为此本文提出了一种基于多尺度特征提取的双轴注意力图卷积剩余寿命预测模型。模型首先通过级联的尺度可塑卷积模块对原始特征进行多尺度时空特征提取,获得不同维度的时空特征;随后利用这些不同维度的时空特征构造时空图,通过图卷积操作挖掘数据深层次的依赖关系;最后设计双轴注意力机制,对时间维度与空间维度的特征进行动态加权,实现关键特征的增强。在 C-MAPSS 数据集的 FD001 和 FD004 子集的实验验证中, RMSE 和 Score 分别为 11.87、236 和 13.44、816。结果表明,该方法相较于其他方法具有更高的精度。

关键词: 剩余寿命预测;GCN;自注意力机制;多尺度特征提取

中图分类号: TN807 文献标识码: A 国家标准学科分类代码: 520.20

Multi-scale dual-axis attention GCN model for remaining useful life prediction

Zheng Senxiao¹ Guo Zhitao^{1,2} Li Yibo¹ Yun Zhi¹

(1. College of Electronic Information Engineering, Hebei University of Technology, Tianjin 300130, China;

2. Innovation Research Institute, Hebei University of Technology (Shijiazhuang), Shijiazhuang 050299, China)

Abstract: Accurate prediction of equipment remaining useful life (RUL) can optimize maintenance strategies, reduce costs, and improve overall efficiency. However, most existing methods rely on separately extracting temporal and spatial features, which hinders the effective fusion of temporal and spatial information. To address this issue, this paper proposes a dual-axis attention graph convolutional network based on multi-scale feature extraction for RUL prediction. The model first utilizes a cascaded scale-adaptive convolution module to perform multi-scale spatiotemporal feature extraction from raw sensor data, capturing spatiotemporal features across different dimensions. These features are then used to construct a spatiotemporal graph, where graph convolution operations are applied to uncover deep dependencies within the data. Finally, a dual-axis attention mechanism is designed to dynamically weight features along both the temporal and spatial dimensions, thereby enhancing the representation of critical features. In the experimental validation on the FD001 and FD004 subsets of the C-MAPSS dataset, the RMSE and Score were 11.87 and 236 for FD001, and 13.44 and 816 for FD004, respectively. The results show that this method has higher accuracy compared with other methods.

Keywords: remaining useful life prediction; graph convolutional network; self-attention mechanism; multi-scale feature extraction

0 引言

健康管理预测技术在现代工业中被广泛应用,以确保机械设备的安全性和可靠性,其中状态维修是健康管理预测系统的重要组成部分^[1]。在设备故障发生前,准确预测其剩余使用寿命(remaining useful life, RUL)是至关重要

的,不仅有助于保障设备的安全性,还能够优化维护策略、降低运营成本、提高系统的可靠性^[2]。在传统工业维护中,定期检查和被动维修策略通常效率较低且成本高昂。因此,数据驱动的预测性维护方法受到广泛关注,它利用传感器数据和机器学习技术预测设备退化趋势并估计 RUL^[3]。

近年来,深度学习技术的快速发展极大地推动了 RUL

收稿日期:2025-05-22

* 基金项目:河北工业大学创新研究院(石家庄)石家庄市科技合作专项基金(SJZZXB23005、SJZZXC24011)项目资助

预测研究,使得模型能够自动从高维传感器数据中提取有意义的特征。在这些技术中,卷积神经网络(convolutional neural networks, CNN)和循环神经网络(recurrent neural networks, RNN)在捕捉时间序列数据的时空依赖性方面取得了显著成效。例如,Mo 等^[4]利用进化算法自动搜索 CNN-LSTM 模型的最优架构,有效提升了 RUL 预测精度。Zhang 等^[5]提出了结合一维卷积神经网络和双向 GRU 的混合神经网络,综合了时间和空间特征,建立了特征与 RUL 之间的映射关系。此外,基于注意力机制的模型 Transformer 在 RUL 预测领域同样获得了显著成果。Zhou 等^[6]通过稀疏化的 ProbSparse 自注意力机制以及生成式解码器,提高了 Transformer 用于长序列预测的推理速度。Liu 等^[7]通过结合通道注意力和时间注意力,提出了一种结合 CNN 与 Transformer 的端到端模型,取得了显著的预测效果。图神经网络(graph neural networks, GNN)作为一种新兴方法,能够有效建模不同传感器信号之间的复杂关系。Wang 等^[8]基于多分支自注意力机制和图神经网络提出了端到端框架 SEensor Alignment,用于多变量时间序列的无监督领域自适应,取得了良好的性能。Wang 等^[9]通过构建层次化的时空依赖性学习框架,在 GNN 的基础上提出了多变量时间序列预测模型 HSDGNN,有效解决了现有模型对属性内依赖和动态空间依赖建模不足的问题。

尽管剩余使用寿命预测技术已取得广泛关注,但其在实际工程应用中仍面临诸多挑战。一方面,发动机运行环境复杂多变,载荷、温度、湿度和气压等工况条件具有高度不确定性,导致传感器数据表现出非线性与非平稳性^[10],模型难以准确捕捉退化规律。另一方面,多变量传感器信号之间存在潜在的非线性相关性和动态变化关系,如何融合时间信息与空间信息、避免信息冗余并突出关键特征,是提升模型泛化性能和预测精度的核心问题。

当前,GNN 通常与时间编码器组合以捕获时空依赖性,尽管在一定程度上提升了时空依赖性的建模能力,但仍存在一些局限性。首先,方法依赖独立的时间编码器和 GNN 模块进行时空特征提取,为每个时间戳分别构造图来表示传感器间的时空关系,如此会导致时间信息与空间信息融合不充分,限制模型的表达能力。其次,时间编码器通常基于固定的时间窗口或预定义的时间模式,难以适应不同设备的动态变化,导致模型泛化能力受限。此外,现有方法未充分考虑不同传感器通道的重要性,通常将所有通道等权重处理,忽略了不同传感器对 RUL 预测的重要性差异。

为了解决上述问题,本文提出了一种融合多尺度时空特征提取与双轴注意力机制的图卷积网络模型(multiscale dual-axis attention graph convolutional network, MDAGCN)。MDAGCN 模型由两个关键模块组成:多尺度时空特征提取模块和双轴注意力图卷积模块。多尺度时

空特征提取模块将输入的时间序列转换为多个空间尺度,通过一维卷积和非线性变换在空间维度进行信息交互,提取不同维度的时空相关性特征。双轴注意力图卷积模块则同时关注空间拓扑信息和时间拓扑信息,提取节点的全局特征并强化关键特征,使模型具备更强的时空特征提取和自适应优化能力。最后,通过拼接与全连接处理所获取的特征,得到 RUL 的预测值。主要贡献如下:

1)设计了一种基于一维卷积神经网络的多尺度时空特征提取器(multi-scale spatio-temporal feature extractor, MSTFE)。通过利用级联的尺度可塑卷积模块(scale-plastic convolutional module, Sp-Conv),在划分的时间窗口中捕捉传感器之间的空间关系,从而在强化时间特征的同时提取空间特征信息,初步获得不同维度的时空信息。

2)提出了一种双轴注意力图卷积网络(dual-axis attention graph convolutional network, DAGCN),用于捕获多变量时间序列的深层次时空特征。DAGCN 通过移动窗口将相邻时间点的所有传感器特征建立连接,构造时空图并进行图卷积,随即对时间特征轴与空间特征轴进行特征的自适应特征分配,增强模型对关键特征的感知能力。

3)在 NASA 提供的 C-MAPSS 数据集上验证了所提出的 RUL 预测方法。实验结果表明,该方法较其他主流 RUL 预测方法具有更高的预测准确性。

1 剩余使用寿命预测框架

1.1 框架概述

本文所提出的 RUL 预测框架如图 1 所示。在离线训练阶段,首先对原始传感器数据进行筛选,剔除具有恒定数值的传感器数据并计算其 RUL 值,然后对其进行 min-max 归一化处理。其次,通过滑动窗口方法构造时间序列输入,并进行初步的特征提取。为提高模型对长期依赖信息的感知能力,提出尺度可塑卷积模块 Sp-Conv,通过级联的 Sp-Conv 生成不同尺度的时空特征表示,实现对原始序列的多尺度时空特征提取。继而构造时空图以捕捉不同时间戳和传感器之间的时空相关性。经过位置编码增强特征后,设计双轴注意力图卷积网络 DAGCN 来充分提取序列的空间特征,并对不同通道的特征进行动态加权,学习高级传感器特征。最后,将不同尺度生成的特征进行拼接,通过输出层获得 RUL 预测结果。在在线预测阶段,使用在验证集上表现最优的模型,输入实时传感器数据进行 RUL 预测。

1.2 数据预处理

在使用模型进行训练和预测之前,有必要对数据集中的原始数据进行预处理。预处理会将数据转换成模型可以识别和利用的格式,其流程包括数据筛选、标签设计、数据归一化、时间窗口划分以及图的构建。

数据筛选:C-MAPSS 数据集包含来自 21 个传感器的测量数据,这些传感器覆盖了发动机多个关键部件和功能区域,如压气机、燃烧室、高压涡轮、低压涡轮等。每个传感

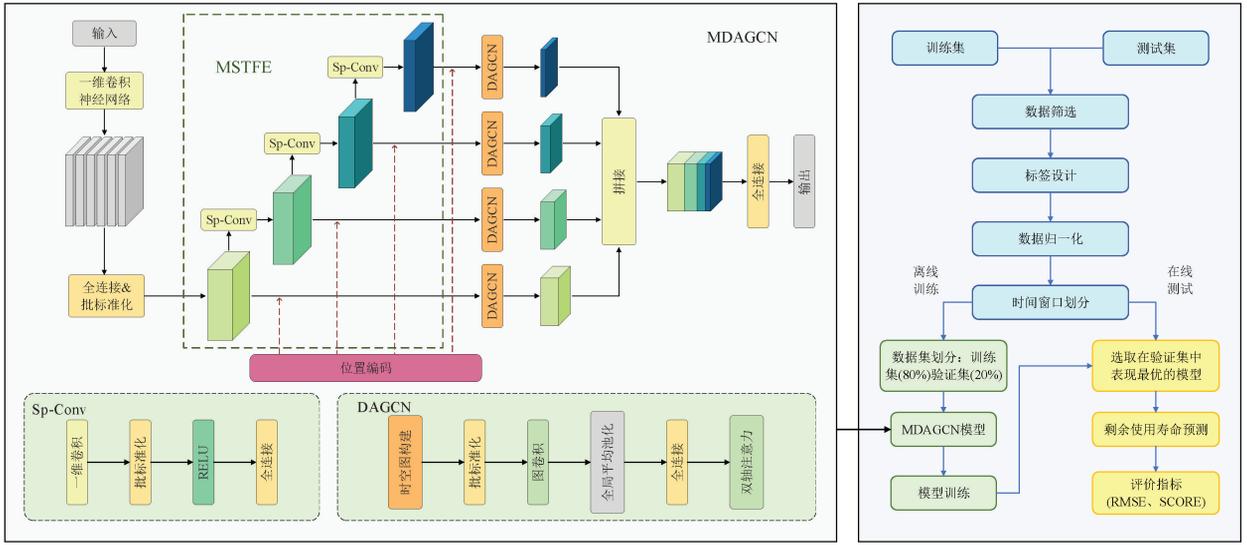


图 1 MDAGCN 模型用于 RUL 预测整体框架

Fig. 1 The overall framework of the MDAGCN model for RUL prediction

器以高频率采样的方式记录与发动机运行状态密切相关的物理参数,包括温度、压力、流速等。其中部分传感器在整个生命周期内数值几乎保持恒定(s_1, s_5, s_6, s_10, s_16, s_18, s_19),直到故障发生。这说明它们不能代表发动机的传感器退化趋势。为了防止这些数据影响模型的预测性能,实验时将它们去除,使其不参与模型的训练。

标签设计:有效地对 RUL 标签进行建模的能力将影响模型的预测准确性。在模型进行训练时,每台发动机的 RUL 可以通过 $RUL = \max(\text{cycle}) - \text{current cycle}$ 进行计算,其中 $\max(\text{cycle})$ 代表该发动机的最后一个运行周期, current cycle 代表当前的时间步。此外,由于故障模拟遵循稳定工况下的逐步性能退化过程,与线性退化趋势具有高度一致性^[11],因此本文假设发动机的 RUL 在开始运行时保持设定的最大值 \max_RUL , 并将其限制为 125, 然后线性退化直到故障。当计算出的 RUL 超过该值时将其截断: $RUL = \min(RUL, \max_RUL)$ 。其中, RUL 是实际剩余使用寿命, RUL_{target} 是训练模型的目标剩余使用寿命。

$$RUL_{target} = \begin{cases} 125, RUL \geq 125 \\ RUL, RUL < 125 \end{cases} \quad (1)$$

数据归一化:由于 C-MAPSS 数据集中不同传感器的测量值具有不同的尺度和单位且数值相差较大,直接分析原始数据会导致数值较大的特征在计算损失时主导梯度更新,数值较小的特征可能被忽略,影响模型的准确性。因此,本文采用 min-max 归一化方法,将来自不同传感器的数据缩放为 $[0, 1]$, 使不同特征对模型贡献相对均衡。其中, x 是传感器测量值, $\min(x)$ 和 $\max(x)$ 分别是该传感器测量值的最小值和最大值, x' 是归一化的传感器测量值。

$$x' = \frac{x - \min(x)}{\max(x) - \min(x)} \quad (2)$$

此外,为了避免 RUL 数值过大带来的梯度不稳定问题,本文使用 \max_RUL 对 RUL_{target} 进行归一化处理。其中, RUL_{norm} 为归一化后的 RUL_{target} 。

$$RUL_{norm} = \frac{RUL_{target}}{\max_RUL} \quad (3)$$

时间窗口的划分:由于模型需要输入时间窗口格式的数据,即每个时间点的预测应依赖于过去一段时间的传感器状态,因此需要将数据按照滑动窗口的方式构造样本。对于给定的多变量时间序列 $X \in \mathbb{R}^{N \times L}$, 其中 N 表示传感器数量, L 表示时间长度,假设窗口大小为 f , 则对于时间步 w , 输入时间片序列为:

$$X_w = [x_{w-f+1}, x_{w-f+2}, \dots, x_w] \quad (4)$$

其中,每个时间片为 $X_w \in \mathbb{R}^{N \times f}$, 第 w 个时间片中第 i 个传感器的局部序列为 $x_{w,i} \in \mathbb{R}^f$, 第 w 个时间片中第 m 个时间戳的第 i 个传感器为 $x_{w,m,i}$ 。

1.3 多尺度时空特征提取器 (MSTFE)

在本文提出的 MDAGCN 模型中, MSTFE 旨在对滑动窗口所得到的原始数据进行逐层抽象与聚合,从而在不同传感器尺度上捕捉潜在的模式与结构,并提取出多层次、多尺度的有效特征。具体而言,原始数据首先经过两次一维卷积处理生成初步的特征表示,随后通过全连接层与 BatchNorm 操作得到初步特征的线性重构与归一化结果。此时,初步特征已经具备了一定的局部信息表达能力,但仅凭单一尺度的信息仍不足以充分描述复杂的时空序列。因此,本文引入了基于尺度可塑卷积模块 (Sp-Conv) 的多尺度时空特征提取策略,以进一步捕捉短期瞬态变化与长期趋势信息,从而在初始的原始数据中提取多尺度的时空特征,实现对时空序列局部与全局信息的有效

融合。

Sp-Conv 是 MSTFE 的核心组件之一,其作用在于通过一维卷积和非线性变换,在传感器维度进行信息交互,从而有效捕捉传感器节点间的局部关系,并保留数据的空间拓扑结构,进而提取空间相关性特征,为后续模块提供更具代表性的特征表示。在此基础上,结合 BatchNorm 归一化和 ReLU 激活,增强模型的泛化能力,并降低梯度消失问题。最后,通过全连接层进行特征映射,优化特征表示。

Sp-Conv 的输入为 $X \in \mathbb{R}^{B \times T \times N \times F}$,其中 B 代表批次的大小, T 代表时间步长, N 代表传感器的数量, F 表示经降维映射后的节点嵌入特征,输出记为 Z 。其具体实现表示为:

$$\begin{cases} H_{b,t,n,f'} = \sum_{f=1}^F W_{f',f,k} \cdot X_{b,t,n+k,f} + b_{f'} \\ \hat{H}_{b,t,n,f'} = \frac{H_{b,t,n,f'} - \mu_{f'}}{\sigma_{f'}} \gamma_{f'} + \beta_{f'} \\ Z_{b,t,n,d} = W_{d,f'} \cdot \text{ReLU}(\hat{H}_{b,t,n,f'}) + b_d \end{cases} \quad (5)$$

其中, k 为卷积核大小, $W_{f',f,k}$ 与 $b_{f'}$ 分别为待学习的卷积核权重与偏置项, $n+k$ 为节点在 k 邻域内进行信息聚合, f' 是卷积输出的特征维度。 $\mu_{f'}$ 、 $\sigma_{f'}$ 分别是批次维度上的均值和标准差, $\gamma_{f'}$ 、 $\beta_{f'}$ 是可学习的缩放和平移参数。 $W_{d,f'}$ 是全连接层的权重, b_d 是偏置项, d 是输出的特征维度。

MSTFE 通过级联多层的 Sp-Conv 模块,可以捕获从局部到全局、从低阶到高阶的传感器关系。这种结构使得在特征空间中邻接关系稀疏、直接联系较弱的传感器之间的信息可以通过多层特征传播,从而提升特征的全局性。不同传感器的数据存在动态相关性,但这种关系并不总是线性的,还可能受到环境因素的影响。多层 Sp-Conv 允许模型逐步学习不同层次的跨传感器交互模式,不仅能够平滑噪声,还能提高模型的泛化能力。MSTFE 为图卷积操作提供了丰富的时空特征,为模型后续在全局范围内捕捉更精确的空间和时间依赖关系奠定了基础。

1.4 双轴注意力图卷积网络(DAGCN)

图卷积网络(graph convolutional network,GCN)通过图结构对多维时序数据的空间相关性进行深度挖掘,能够高效建模多维数据间的空间依赖关系,并结合时间建模以提升动态特征表达能力^[12]。然而,GCN 在特征聚合过程中对不同维度的特征采用等权重处理,未能区分不同特征的重要性,导致对潜在关键信息的表达能力受限,进而影响模型的预测性能。

为解决上述问题,本文提出了双轴注意力图卷积网络 DAGCN,旨在充分挖掘数据中潜在的空间关联结构并提升模型对关键特征的感知能力,从而提高对发动机剩余使用寿命的预测精度。具体而言,考虑到传感器节点和时间

节点的嵌入特征分别承载空间和时间维度的抽象信息,若仅关注某一维度的全局特征,可能会忽略另一维度的重要信息,甚至放大冗余。为此本文设计了一种双轴注意力机制(dual-axis squeeze-and-excitation, Da-SE),在空间和时间两个维度上分别引入注意力机制:首先通过全局平均池化对每个维度的特征进行压缩,提取全局上下文信息;随后由两个全连接层组成的激励模块生成权重系数,用于对各维度的嵌入特征进行重标定,增强关键特征响应,抑制无关信息干扰,从而更准确地捕捉复杂特征。DAGCN 融合了 GCN 与 Da-SE 机制,其整体结构如图 2 所示。

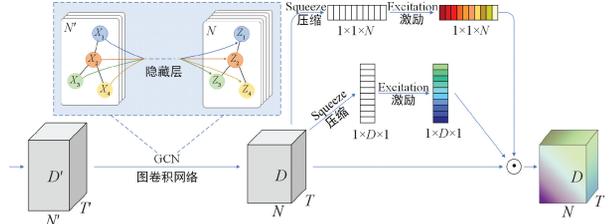


图 2 DAGCN 整体框架

Fig. 2 The overall framework of the DAGCN model

在 DAGCN 中,图结构的构建是不可或缺的一部分。本文将所有时间块中的所有传感器节点当作独立的节点,每个节点代表一个传感器在特定时间点的特征,节点之间的相似度则表示传感器之间的相关性。随后,利用移动窗口的方法,将每个窗口内的传感器节点相互连接来构造时空图,进而充分捕捉不同时间戳之间以及不同传感器之间的时空相关性。具体而言,首先对传感器特征进行位置编码,以增强其空间和时间信息。然后,通过点积计算两个节点之间的相似度,并利用 softmax 函数进行归一化,得到节点之间的权重矩阵。相关计算如下:

$$\begin{cases} z_{w,m,i} = f_c(Z_{b,t,n,d}) + p_m \\ p_m^{(q)} = \begin{cases} \sin(\omega_k \cdot m), q = 2k \\ \cos(\omega_k \cdot m), q = 2k + 1 \end{cases} \\ e_{(m,i),(n,j)}^{(w)} = \text{softmax}(g_s(z_{w,m,i}) \cdot g_s(z_{w,n,j})^T) \end{cases} \quad (6)$$

其中, $f_c(Z_{b,t,n,d}) \in \mathbb{R}^d$ 表示提取的 $Z_{b,t,n,d}$ 的表示特征, $p_m \in \mathbb{R}^d$ 表示时间位置编码,定义如 $p_m^{(j)}$ 所示, $z_{w,m,i}$ 表示第 w 个时间块中第 m 个时间戳的第 i 个传感器特征, $z_{w,r,j}$ 表示第 w 个时间块中第 r 个时间戳的第 j 个传感器特征, $e_{(m,i),(n,j)}^{(w)}$ 表示节点 (w,m,i) 与节点 (w,n,j) 间边的相似度权重,函数 $g_s(\cdot)$ 通过可学习的权重进行映射,softmax 函数对相似度进行归一化处理。

由于单纯基于特征相似度构造图未能考虑不同时间戳间的时序距离,本文设计了时间衰减因子 c_{mn} ,其每个元素由时间间隔决定。时间间隔越近,衰减值越大(即相关性越高),时间间隔较远,则相关性被按比例降低。最终通过特征相似度和时间衰减因子相乘得到图中边的权重,即:

$$\begin{cases} c_{mn} = \delta^{|m-n|}, \delta \in (0, 1) \\ a_{(m,i),(n,j)}^{(w)} = e_{(m,i),(n,j)}^{(w)} \cdot c_{mn} \end{cases} \quad (7)$$

$a_{(m,i),(n,j)}^{(w)}$ 代表节点 (w, m, i) 与节点 (w, n, j) 间边的最终相似度权重。通过不同节点之间的相似度构建邻接矩阵 $A^{(w)} \in \mathbb{R}^{(Nf) \times (Nf)}$ ，滑动窗口内部节点集合 $V^{(w)} = \{z_{w,m,i}\}$ ，构造时空图 $G^{(w)} = (V^{(w)}, A^{(w)})$ ，如此不仅反映了传感器在同一时间点内的空间相关性，还捕捉了不同时间点的关联信息。DAGCN 在时空图的基础上采用邻居信息聚合与更新进行图卷积操作，并在窗口内的时间维度上引入了全局平均池化进行特征融合，形成第 w 个时间块的特征表示：

$$\begin{cases} z_{w,m,i}^{out} = ReLU\left(\sum_{n=1}^M \sum_{j=1}^N a_{(m,i),(n,j)}^{(w)} \cdot z_{w,m,i} \cdot W_g + b_g\right) \\ z_{w,i}^{pool} = \frac{1}{M} \sum_{m=1}^M z_{w,m,i}^{out} \\ \mathbf{Z}^{(w)} = [z_{w,1}^{pool}, z_{w,2}^{pool}, \dots, z_{w,N}^{pool}] \end{cases} \quad (8)$$

其中， $z_{w,m,i} \in \mathbb{R}^d$ 为位置编码后的节点特征， $W_g \in \mathbb{R}^{d \times d}$ 、 $b_g \in \mathbb{R}^d$ 为图卷积的可训练参数， $z_{w,m,i}^{out}$ 为经消息聚合与节点更新后的节点特征， $z_{w,i}^{pool}$ 为每个传感器 i 在滑动窗口 w 中不同时间戳的表示平均， $\mathbf{Z}^{(w)} \in \mathbb{R}^{N \times d}$ 表示窗口 w 的特征矩阵。

由于不同传感器节点以及不同时间步对模型预测的贡献并不均衡，本文基于图卷积后的特征张量 $\mathbf{Z}^{(w)} \in \mathbb{R}^{N \times T \times d}$ ，设计了双轴注意力机制，从传感器维度和时间维度两个方向动态建模重要性。

$$\begin{cases} y_n^{(w)} = \frac{1}{T \times D} \sum_{t=1}^T \sum_{d=1}^D \mathbf{Z}_{n,t,d}^{(w)} \quad (n = 1, \dots, N) \\ \alpha_N = \sigma(\mathbf{W}_2^{(N)} \cdot ReLU(\mathbf{W}_1^{(N)} \cdot y_n^{(w)})) \\ y_d^{(w)} = \frac{1}{N \times T} \sum_{n=1}^N \sum_{t=1}^T \mathbf{Z}_{n,t,d}^{(w)} \quad (d = 1, \dots, D) \\ \alpha_D = \sigma(\mathbf{W}_2^{(D)} \cdot ReLU(\mathbf{W}_1^{(D)} \cdot y_d^{(w)})) \\ Z_{out}^{(w)} = Z^{(w)} \odot \alpha_N \odot \alpha_D \end{cases} \quad (9)$$

其中， $y_n^{(w)} \in \mathbb{R}^N$ 、 $y_d^{(w)} \in \mathbb{R}^D$ 分别代表传感器特征节点与时间特征维度上进行全局平均池化， $\sigma(\cdot)$ 代表 Sigmoid 激活函数， α_N, α_D 代表经过两层全连接网络进行非线性映射与压缩，生成的传感器维度与时间特征维度的注意力权重， $\mathbf{W}_1^{(N)} \in \mathbb{R}^{r \times N}$ 、 $\mathbf{W}_2^{(N)} \in \mathbb{R}^{N \times r}$ 、 $\mathbf{W}_1^{(D)} \in \mathbb{R}^{r' \times D}$ 、 $\mathbf{W}_2^{(D)} \in \mathbb{R}^{D \times r'}$ 分别代表全连接层的权重矩阵，最终通过 Hadamard 乘积将权重信息与图卷积后的特征张量相融合。

DAGCN 兼顾空间拓扑信息和关键特征选择，不仅能够精准捕捉不同时间戳、不同传感器之间的复杂时空结构关系，还加强了对关键节点的关注。通过动态调节不同维度的注意力权重，模型增强了特征选择能力以及自适应优化能力，提高了模型的预测精度和鲁棒性。

2 实验与结果

2.1 数据集介绍

实验数据集是来自美国国家航空航天局提供的商用模块化航空推进系统仿真(C-MAPSS)数据集。C-MAPSS 数据集包含了多个飞机发动机的模拟退化数据，每个飞机发动机配备 21 个传感器，从不同的组件和位置收集数据。

C-MAPSS 数据集具体情况如表 1 所示，数据集包含基于各种操作条件和故障模式的 4 个子数据集(FD001~FD004)。FD002 和 FD004 包含 6 种用于预测的故障模式，更具有挑战性。此外，每个子数据集都包括训练集和测试集，训练集涵盖了每台发动机从初始运行状态到故障的整个生命周期。测试集则是对每台发动机随机选择从发动机的初始运行状态到发动机故障之间的时间点，涵盖了从初始运行状态到随机时间点之间的部分生命周期。

表 1 飞机发动机数据集 C-MAPSS

Table 1 Aircraft engine dataset C-MAPSS

数据集	训练集引擎数	测试集引擎数	工况条件	故障状态
FD001	100	100	1	1
FD002	260	259	6	1
FD003	100	100	1	2
FD004	249	248	6	2

2.2 评估指标

本文采用回归问题和 RUL 预测中常用的度量，即 RMSE (均方根误差) 和 Score，来评估模型的性能。RMSE 是回归任务中广泛使用的性能指标，其定义如下：

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (\hat{Y}_i - Y_i)^2} \quad (10)$$

其中， \hat{Y}_i 和 Y_i 分别表示第 i 个样本的估计和实际 RUL。N 是测试样本的总数。对于设备的 RUL 预测任务，与早期预测相比，延迟预测会带来更严重的后果^[13]。因此，使用 Score 函数来弥补 RMSE 在这方面的不足，定义如下：

$$Score = \begin{cases} \sum_{i=1}^N [e^{-\frac{(\hat{Y}_i - Y_i)}{13}} - 1], & \hat{Y}_i < Y_i \\ \sum_{i=1}^N [e^{\frac{(\hat{Y}_i - Y_i)}{10}} - 1], & \hat{Y}_i \geq Y_i \end{cases} \quad (11)$$

2.3 实验参数分析

表 2 为实验所使用的最优超参数配置。由于 C-MAPSS 中 4 种子数据集设备的运行条件和故障模式不同，同一模型的最佳滑动窗口大小也可能在不同的子数据集中有所不同，因此有必要探索不同滑动窗口大小对每个子数据集中模型性能的影响。在其他参数设置相同的情况下，对滑动窗口大小在范围 [20, 30, 40, 50, 60] 内进行测

试,得到 MDAGCN 模型在 FD001~FD004 子数据集上的最佳滑动窗口尺寸分别为 30、50、40 和 50,实验结果如图 3 所示,其中实线代表 RMSE 值,虚线代表 Score 值。

表 2 实验超参数配置

Table 2 Experimental hyperparameter configuration

类别	超参数	值
FD001~FD004 的窗口尺寸 30/50/40/50		
通用超参数	时间块大小	5
	批量大小	256
	学习率	0.001
	decay 权重衰减	0.7
	有效传感器数量	14
多尺度时空特征提取器	特征维度	16
	Sp-Conv 层数	4
双轴注意力图卷积网络	Sp-Conv 中的激活函数	ReLU
	时空图的大小	28×28
	池化操作	平均池化
全连接层(FC)中的激活函数		ReLU

在 MSTFE 中,Sp-Conv 层数和特征维度的大小是极为关键的超参数,为评估其对模型 RUL 预测结果的影响,

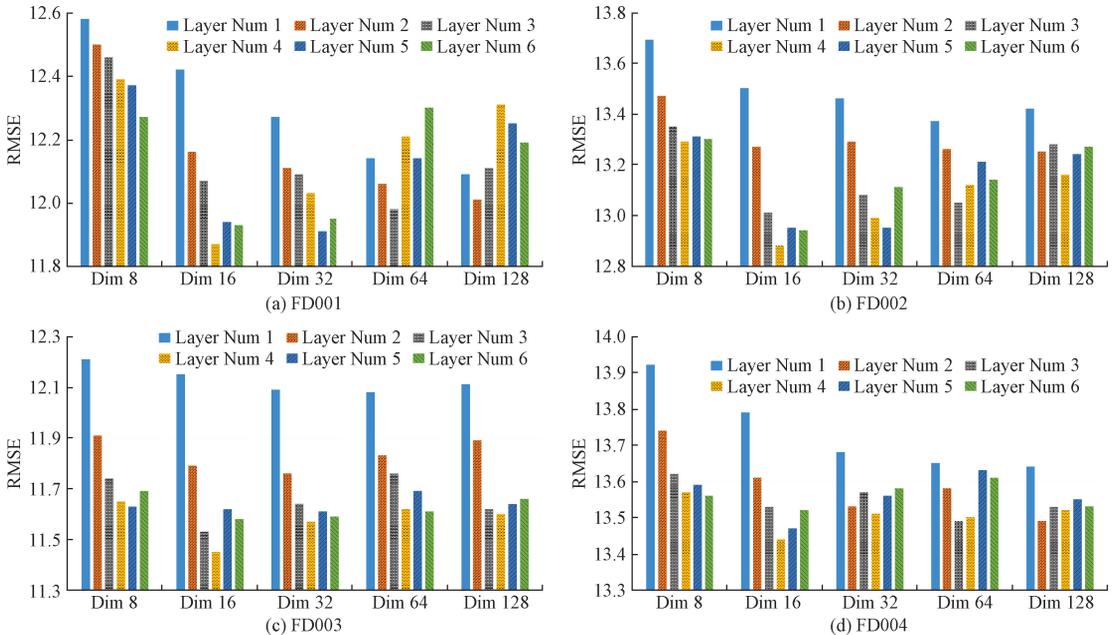


图 4 FD001-FD004 数据集中 Sp-Conv 层数和特征维度的大小对 RMSE 的影响

Fig. 4 The influence of the number of Sp-Conv layers and the size of feature dimensions in the FD001-FD004 dataset on RMSE

2.4 实验对比

为进一步验证 MDAGCN 的优越性能,本文选择了当前主流的 RUL 预测模型进行对比。其中对比模型包括 BLSTM^[14], 基于 RNN-CNN 的 CNN-LSTM^[4], 基于 Transformer 的 InFormer^[6] 和 DA-Transformer^[7], 以及基

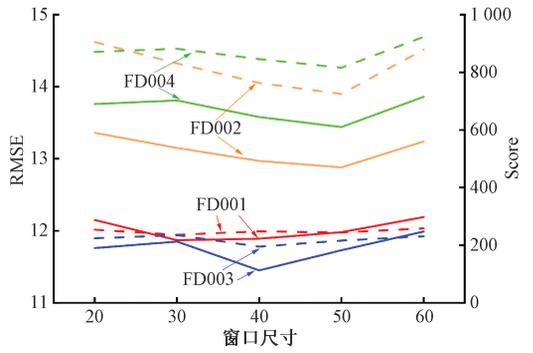


图 3 不同窗口尺寸对 RMSE 与 Score 的影响

Fig. 3 The influence of different window size on RMSE and Score

本文在 4 个数据集上设计了交叉验证实验。其中,Sp-Conv 级联层数在范围[1,2,3,4,5,6]内调整,特征维度在范围[8,16,32,64,128]内调整。实验结果如图 4 所示,每个子图的 X 轴和 Y 轴分别表示模型的特征维度和 RMSE 值,不同的颜色表示 Sp-Conv 的级联层数。实验证明,特征维度为 8 或 Sp-Conv 层数为 1 时的性能通常较差,说明此时非线性映射能力不足,出现了欠拟合现象;当 Sp-Conv 层数和特征维度过高时,网络消耗的计算量与过拟合的风险也随之上升;当 Sp-Conv 的层数和特征维度大小分别为 4 和 16 时,模型具有最佳的 RUL 预测能力。

于 GNN 的 GCN^[15]、HAGCN^[16] 和 MAGNN^[17]。与当前主流的 RUL 预测模型相比,MDAGCN 在多数情况下都有很大的提升。MDAGCN 在 FD002、FD003 和 FD004 数据集上的效果均为最优, RMSE 分别比次优的方法提高了 1.6%、3.9% 和 5.1%, 并且在 FD001 数据集上也达到了次

优的效果,接近最优模型。实验结果如表 3 所示,其中粗体和下划线分别表示实验中最优和次优的结果。实验证明,本文提出的 MDAGCN 比其他主流 RUL 预测模型具

有更好的性能,且模型在 C-MAPSS 数据集的 4 个子数据集上均表现出色,体现了模型的有效性和优越性。

表 3 不同方法在 C-MAPSS 数据集上的预测性能比较

Table 3 Comparison of the prediction performance of different methods on the C-MAPSS dataset

模型	FD001		FD002		FD003		FD004	
	RMSE	Score	RMSE	Score	RMSE	Score	RMSE	Score
BLSTM ^[14]	—	—	25.11	4 793	—	—	26.61	4 971
CNN-LSTM ^[4]	11.56	247	17.67	1 743	12.98	808	20.19	3 051
InFormer ^[6]	13.13	263	13.20	<u>715</u>	12.58	228	<u>14.16</u>	1 023
DA-Transformer ^[7]	12.25	198	17.08	1 575	13.39	290	19.86	1 741
GCN ^[15]	12.58	237	13.78	849	<u>11.92</u>	<u>218</u>	14.44	967
HAGCN ^[16]	13.10	263	14.92	1 086	13.46	327	14.66	<u>880</u>
MAGNN ^[17]	12.63	246	<u>13.09</u>	714	12.15	253	14.30	978
MDAGCN	<u>11.87</u>	<u>236</u>	12.88	725	11.45	195	13.44	816

为了进一步证明 MDAGCN 的有效性,从 4 个子数据集中随机选择一组发动机单元进行 RUL 预测,结果如图 5 所示。可以观察到,对于 4 个子数据集,MDAGCN 的预测曲线十分贴合真实的 RUL 曲线,平均 RUL 预测误差控制在 6 个周期以内,表明了 MDAGCN 的良好性能。此外,中期预测误差往往大于早期和晚期预测误差,这是由于发动机刚刚进入退化阶段,时间序列中仍包含发动机早期健康状态的历史信息,但包含的退化信息较少。随着包含在输入时间序列中的退化信息增加,预测误差也随之降低。

2.5 消融实验

为了更好的评估提出模块的有效性,本文提出了 3 种模型变体来评估模型中各模块的有效性,并使用 RMSE 与 Score 评估 3 种变体的性能,结果如表 4 所示,其中粗体和下划线分别表示实验中最优和次优的结果。3 种模型变体分别为模型 1 无 MSTFE 并将 DAGCN 替换为 GCN 的 MDAGCN、模型 2 将 DAGCN 替换为 GCN 的 MDAGCN、模型 3 无 MSTFE 的 MDAGCN。其中,与模型 1 相比,模型 2 与模型 3 在各个数据集上均显著降低了预测误差,证明了 DAGCN 与 MSTFE 在模型中的重要性。此外,完整的 MDAGCN 模型在 RMSE 与 Score 指标上均优于其他 3 种模型变体,证明了 MSTFE 与 DAGCN 在联合使用时能

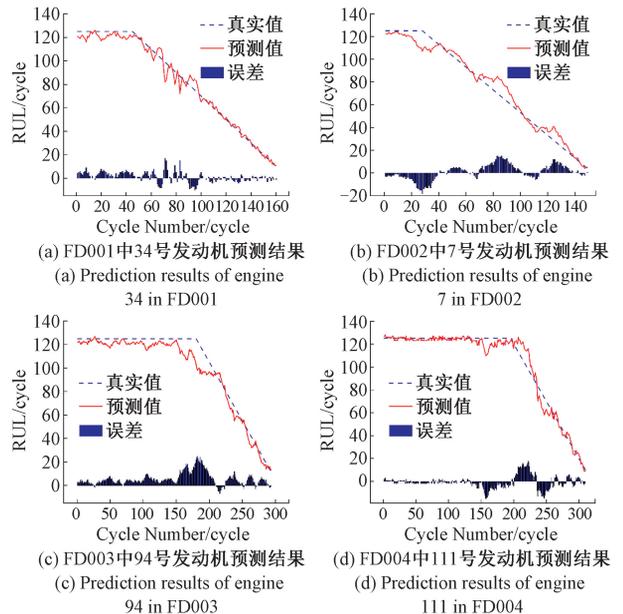


图 5 四个随机发动机引擎上的 RUL 预测结果

Fig. 5 The RUL prediction results on four random engine engines

够协同增强特征表示能力,验证了本文提出的各模块在发动机剩余使用寿命预测任务中的有效性和必要性。

表 4 MDAGCN 模型与其三种变体在 C-MAPSS 数据集上的 RMSE 与 Score 值

Table 4 The RMSE and Score values of the MDAGCN model and its three variants on the C-MAPSS dataset

模型	FD001		FD002		FD003		FD004		均值	
	RMSE	Score								
模型 1	12.67	271	13.89	964	12.44	271	14.31	940	13.33	612
模型 2	12.52	264	<u>13.34</u>	869	<u>11.56</u>	<u>203</u>	14.03	942	<u>12.86</u>	570
模型 3	<u>12.42</u>	<u>256</u>	13.5	<u>761</u>	12.15	238	<u>13.79</u>	<u>880</u>	12.97	<u>534</u>
MDAGCN	11.87	236	12.88	725	11.45	195	13.44	816	12.41	493

3 结 论

RUL 预测是 PHM 确保机器可靠性和安全性的重要手段,有助于帮助复杂维修决策的开展。本文提出了一种融合多尺度时空特征提取与双轴注意力机制的 GCN 网络模型。首先引入了多尺度时空特征提取模块对卷积处理后的初步特征进行多尺度的时空特征提取,帮助模型建立不同尺度的感受野;通过将双轴注意力机制融合到 GCN 结构中,弥补了 GCN 网络在不同特征通道关注度方面的不足,增强了对关键特征的提取能力。本文通过设计一系列消融实验验证了模型各部分的有效性,并在 C-MAPSS 数据集上与其他先进的 RUL 预测模型进行对比,结果表明 MDAGCN 模型的预测精度优于其他先进的模型。在未来,如何在保证模型预测精度的前提下降低其计算量,使其能够在有限资源的设备上部署,仍是一个值得深入研究的方向。

参考文献

- [1] JIANG L, ZHANG T AO, LEI W, et al. A new convolutional dual-channel Transformer network with time window concatenation for remaining useful life prediction of rolling bearings [J]. *Advanced Engineering Informatics*, 2023, 56: 101966.
- [2] 张鑫阳,王可庆,贾新旺,等. 基于并行 TCN-SE-BiLSTM 模型的涡扇发动机剩余寿命预测[J]. *电子测量技术*, 2024, 47(24): 179-187.
ZHANG X Y, WANG K Q, JIA X W, et al. Prediction of remaining useful life for turbofan engines based on parallel TCN-SE-BiLSTM model [J]. *Electronic Measurement Technology*, 2024, 47(24): 179-187.
- [3] ZHAO K, JIA ZH, JIA F, et al. Multi-scale integrated deep self-attention network for predicting remaining useful life of aero-engine[J]. *Engineering Applications of Artificial Intelligence*, 2023, 120:105860.
- [4] MO H, CUSTODE L L, IACCA G. Evolutionary neural architecture search for remaining useful life prediction [J]. *Applied Soft Computing*, 2021, 108: 107474.
- [5] ZHANG J, TIAN J L, LI M L, et al. A parallel hybrid neural network with integration of spatial and temporal features for remaining useful life prediction in prognostics [J]. *IEEE Transactions on Instrumentation and Measurement*, 2023, 72: 1-12.
- [6] ZHOU H Y, ZHANG SH H, PENG J Q, et al. Informer: Beyond efficient transformer for long sequence time-series forecasting [C]. *AAAI Conference on Artificial Intelligence*, 2021, 35(12): 11106-11115.
- [7] LIU L, SONG X, ZHOU ZH T. Aircraft engine remaining useful life estimation via a double attention-based data-driven architecture [J]. *Reliability Engineering & System Safety*, 2022, 221: 108330.
- [8] WANG Y CH, XU Y C, YANG J F, et al. SENSOR alignment for multivariate timeseries unsupervised domain adaptation[C]. *AAAI Conference on Artificial Intelligence*, 2023, 37(8), 10253-10261.
- [9] WANG Y CH, WU M, LI X 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.
- [10] 白晏年,李小波,杨志豪,等. 基于变点处数据处理的剩余寿命预测[J]. *电子测量技术*, 2025, 48(2): 108-114.
BAI Y N, LI X B, YANG ZH H, et al. Remaining useful life prediction based on data processing at change points [J]. *Electronic Measurement Technology*, 2025, 48(2): 108-114.
- [11] GUO J, LEI SH CH, DU B G. MHT: A multiscale hourglass-transformer for remaining useful life prediction of aircraft engine [J]. *Engineering Applications of Artificial Intelligence*, 2024, 128: 107519.
- [12] 徐震震,薛林,马凯,等. 基于时空特征的航空发动机剩余使用寿命预测[J]. *电子测量技术*, 2023, 46(23): 63-67.
XU ZH ZH, XUE L, MA K, et al. Remaining useful life prediction of based on aero-engine spatio-temporal feature [J]. *Electronic Measurement Technology*, 2023, 46(23): 63-67.
- [13] ZHANG Y, XIN Y Q, LIU ZH W, et al. Health status assessment and remaining useful life prediction of aero-engine based on BiGRU and MMoE [J]. *Reliability Engineering & System Safety*, 2022, 220: 108263.
- [14] HUANG CH G, HUANG H ZH, LI Y F. A

- bidirectional LSTM prognostics method under multiple operational conditions [J]. IEEE Transactions on Industrial Electronics, 2019, 66(11): 8792-802.
- [15] KIPF T N, WELLING M. Semi-supervised classification with graph convolutional networks [J]. ArXiv preprint arXiv: 1609.02907, 2016.
- [16] LI T F, ZHAO ZH B, SUN CH, et al. Hierarchical attention graph convolutional network to fuse multisensor signals for remaining useful life prediction[J]. Reliability Engineering & System Safety, 2021, 215: 107878.
- [17] CHEN L, CHEN D H, SHANG Z J, 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.

作者简介

郑森潇, 硕士研究生, 主要研究方向为人工智能与机器学习。

E-mail: z_senxiao9589@163.com

郭志涛(通信作者), 教授, 主要研究方向为物联网、智能传感器、智能优化算法。

E-mail: 2002089@hebut.edu.cn

李义博, 硕士研究生, 主要研究方向为深度强化学习。

E-mail: yibo_li2001@163.com

贲智, 硕士研究生, 主要研究方向为嵌入式与人工智能。

E-mail: 942775630@qq.com