

# 基于多尺度工况增强网络及 Informer 的设备剩余寿命预测

刘付渝杰<sup>1,2</sup>

(1. 广东省茂名市质量计量监督检测所, 广东 茂名 525000;

2. 广东工业大学 计算机学院, 广州 510006)

**摘要:** 设备 RUL 预测在提高设备可靠性、安全性、降低维护成本等方面具有重要意义; 通过提前发现设备的健康状态和潜在故障, RUL 预测有助于降低突发故障风险、延长设备寿命, 提高工作效率, 确保任务正常运行; 然而在面对设备越来越复杂, 采集到的传感器数据维度越来越高, 传统方法和某些深度学习方法在处理特征关系、长时间序列数据和挖掘重要传感器数据方面存在限制; 为了提高预测准确性, 提出一种基于 MWCEN 结合 Informer 的混合模型——MWCEN-Informer, MWCEN 通过动态工况编码算法对设备时序数据进行工况编码, 对设备传感器进行一维多尺度混合卷积充分提取特征信息, 使用多分支通道注意力机制增强有效特征, 增强后的传感器数据输入 Informer 用于分析设备传感器时序数据的关联性, 以实现更准确的设备 RUL 预测; 以基于 C-MAPSS 的通用涡扇发动机数据集进行验证, 结果表明, 该模型在 4 个子集上的 RMSE 平均减少了 5.5%, S-Score 平均减少了 4.7%, 能有效提高设备在复杂工况和复杂故障下的 RUL 预测精度。

**关键词:** 剩余寿命预测; 多尺度卷积; 工况编码; 注意力机制; Informer

## RUL Prediction of Device Based on Multi-scale Working Condition Enhancement Network and Informer

LIUFU Yujie<sup>1,2</sup>

(1. Quality & Metrology Supervision Testing Institute, Maoming 525000, China;

2. School of Computer Science and Technology, Guangdong University and Technology, Guangzhou 510006, China)

**Abstract:** It is of great significance for the remaining useful life (RUL) prediction of devices to improve their reliability safety, and reducing maintenance costs. By discovering the health status and potential faults of devices in advance, RUL prediction helps to reduce the risk of sudden failure, extend device life, improve work efficiency, and ensure the normal operation of tasks. However, with the increasing complexity of devices, and the collected sensor data have increasingly higher dimensions, traditional methods and some deep learning methods have limitations in processing the feature relationships, long time series data and mining important sensor data. Based on the multi-scale work condition enhancement network (MWCEN) and Informer model, this paper proposes a hybrid model of MWCEN-Informer to improve the prediction accuracy. The MWCEN encodes the device time series data by using the dynamic work condition coding algorithm, fully extracts the feature information by performing one-dimensional multi-scale hybrid convolution on the device sensor information, enhances the effective features by using the multi-branch channel attention mechanism, inputs the enhanced sensor data into the Informer model to analyze the correlation of the device sensor timing data, and achieves more accurate RUL prediction of the device. Validation is carried out on a generic turbofan engine data set based on C-MAPSS, the results show that the model reduces the RMSE by an average of 5.5% and the S-Score by an average of 4.7% on the four subsets, which effectively improves the RUL prediction accuracy of the device under complex operating conditions and complex faults.

**Keywords:** RUL prediction; multi-scale convolution; work condition coding; attention mechanism; Informer

## 0 引言

设备剩余寿命 (RUL, remaining useful life) 预测<sup>[1]</sup>是指在用设备在已经使用一段时间后, 预计还能够继续正常运行或提供所需功能的时间段。RUL 预测在制造业、航空航天、能源、检验检测等领域具有重要意义, 它对设备和系统的可靠性、安全性、维护成本和效率等方面具有积极

影响<sup>[2-3]</sup>。通过 RUL 预测, 可以提前识别设备的健康状况和潜在故障, 从而采取预防性维护措施, 降低设备发生突发故障的风险, 延长设备的使用寿命, 减少设备的突发故障和停机时间对工作效率产生的负面影响。随着科技进步, 设备系统越来越精密, 装备的传感器种类及数量、运行工况复杂性、运行荷载量也日益增加, 导致设备的退

收稿日期: 2024-02-19; 修回日期: 2024-03-20。

基金项目: 广东省市场监督管理局科技项目(2024CZ11); 茂名市科技计划项目(230506164551410)。

作者简介: 刘付渝杰(1991-), 男, 硕士, 高级工程师。

引用格式: 刘付渝杰. 基于多尺度工况增强网络及 Informer 的设备剩余寿命预测[J]. 计算机测量与控制, 2024, 32(8): 115-122.

化过程通常具有突变性、阶段性、模糊性等多种复杂特征,给设备 RUL 的准确预测带来新的挑战,如何准确预测复杂退化过程中的设备 RUL 仍是目前研究的难点与热点。

RUL 预测的目标是根据设备的历史状态监测数据,准确预测当前时刻的剩余寿命,从当前的相关应用研究来看,RUL 预测技术主要分为以下几类:1) 统计方法:传统的统计方法通常依赖于历史数据的分析,如均值、方差、回归分析等,如马尔科夫模型(Markov Model)<sup>[4]</sup>、自回归模型(ARMA<sup>[5]</sup>、GARCH<sup>[6]</sup>)等,这些方法简单易用,但对于复杂的设备系统来说,它们的性能有限;2) 物理建模方法<sup>[7]</sup>:物理建模方法基于设备的物理结构和工作原理,通过建立设备数学模型来预测 RUL。这些方法需要深刻的设备制造技术领域知识和模型参数的准确估计,因此通常用于高度工程化的系统;3) 机器学习方法:在剩余使用寿命(RUL)预测领域,得益于传感器技术的发展,结合采集到的大量高维监测数据,支持向量机<sup>[8]</sup>、随机森林<sup>[9]</sup>和神经网络<sup>[10]</sup>等方法广泛应用于特征提取和模型构建,它们能够自动从数据中提取关键特征并建模非线性关系,因此适用于各种设备类型。此外,在深度学习领域,一些成熟技术如卷积神经网络(CNN)<sup>[11]</sup>、循环神经网络(RNN)<sup>[12]</sup>、长短时记忆网络(LSTM)以及注意力机制(Attention)<sup>[13]</sup>等,也开始在 RUL 预测中崭露头角。这些深度学习方法具备处理大规模高维度数据的能力,有助于有效捕捉复杂的时序特征。此外,一些研究工作还结合了不同的方法,如融合物理模型和机器学习模型,以充分利用领域知识和数据,从而提高了 RUL 预测的准确性和鲁棒性。

随着设备的复杂性的日益增加,物理建模方法成本也随之提高,而目前越来越多的各类传感器广泛嵌入设备内部用于采集设备状态,基于数据驱动的机器学习方法是未来 RUL 预测的主流。文献[14]提出了一种基于 CNN-LSTM 和自我关注机制(SAM)的深度学习融合算法,通过 SAM 能建立时序数据的直接依赖关系,从而解决 LSTM 信息丢失问题,提高对飞机发动机故障的提前预测精度。文献[15]在改进卷积神经网络中引入了软阈值残差注意力机制,解决人工特征提取的局限性,从而提高设备的 RUL 预测精度。文献[16]结合了深度特征表示和长短时记忆网络,提出一种名为支持向量数据归一相关系数的新标准,以改进维护策略。文献[17]设计了一种改进的 Transformer 模型来进行设备 RUL 预测,使用一种没有循环结构的前馈网络,训练效率得到大幅提高。

尽管设备 RUL 预测方法已有大量的研究,随着机械设备复杂化、精密化程度不断提高,以上方法难以在海量的设备状态时序监测数据中有效地提取特征之间的关系和识别重要特征;其次,设备不同传感器数据在模型学习过程中不分离,无法学习到独立传感器的趋势和模式;另外,设备监测数据的历史周期尺度一般较大,提取长时间周期中的退化特征花费的时间代价较大,进一步制约了准确性和运行速度。

为此,针对上述问题,本文提出一种多尺度工况增强

网络(MWCEN, Multi-scale work condition enhancement network),并结合 Informer 对设备 RUL 进行预测,采用 MWCEN 提取设备时序监测状态数据的深层局部趋势特征,采用 Informer 分析长时间序列数据的关联性,实现设备 RUL 预测。首先,对设备工况信息进行编码。其次,对设备每个传感器进行独立一维多尺度卷积操作提取不同传感器的时序信息。然后,结合多分支通道注意力机制关注全局和局部范围最有价值的信息,最后对设备运行周期的故障事件进行事件编码输入到 Informer 中。相比其他方法,本文方法充分考虑了工况变化引入的影响、传感器的独立变化趋势和整体变化趋势以及关键时序特征,从而能够实现设备 RUL 的准确预测。

本文的主要工作与贡献如下:

1) 针对不同工况对设备 RUL 预测的影响差异,如极端工况会导致 RUL 骤减,普通工况则影响轻微,设计了基于设备动态工况的编码算法(CTE, condition-based time encoding),对设备运行周期的所处的工况进行编码,以帮助模型能够更好地理解不同工况对设备 RUL 的影响,该算法考虑工况类型及单一工况持续时间,结合嵌入编码(Embedding Code)和反映相对持续时间的标量进行多类别时间编码,能有效帮助模型学习到历史工况状态信息。

2) 为了提取不同传感器时空可分性以及传感器间的关联,提出了一维多尺度混合卷积模块(ODMHC, one-dimensional multi-scale hybrid convolution),对传感器时序数据分别进行独立和全局多尺度卷积,通过不同尺度的感受野能够有效提取到充分的时序特征,确保模型输入信息的丰富性。

3) 针对卷积操作后得到的大量特征,模型需要聚焦关键信息,本文对通道注意力机制进行了改进,提出多分支通道注意力算法(MBCA, multi-branch channel attention)关注不同特征,并使用各自的学习参数增强重要特征,消除无用信号带来的噪声效应,提高预测效果和模型泛化能力。

## 1 卷积神经网络与 Informer

### 1.1 卷积神经网络

卷积神经网络(CNN, convolutional neural network)<sup>[11]</sup>是一种深度学习模型,最初被设计用于图像处理任务。CNN 是一种具有多层结构的神经网络,主要由卷积层、池化层和全连接层组成。它们的核心特点是卷积操作,该操作允许网络有效地捕捉输入数据中的局部模式和特征。传统 CNN 模型结构如图 1 所示。

卷积层应用一系列滤波器(也称为卷积核)以滑动窗口的方式在输入数据上执行卷积操作。根据卷积核的维度不同,卷积核的滑动方向也有不同,二维卷积核在两个不同维度方向进行滑动,区别于图像数据,对于时序数据,二维卷积核提取深层特征时会同时综合考虑来自不同传感器数据和时序数据的关系,这样有时会对结果造成混淆,降低识别能力。相比之下,一维卷积核只发生在传感器数据的时间维度上,如图 1 所示,在滑动窗口的作用下,他

们提取特征时主要考虑传感器的局部时序特征，例如趋势、季节性、重复模式等，这有助于从原始时序数据中提取有用的信息，并可以有效压缩数据量，加快运算速度。

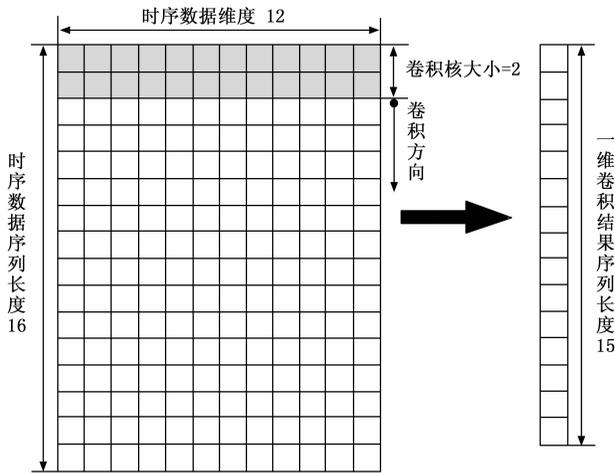


图 1 一维卷积操作

### 1.2 Informer 模型

Informer<sup>[18]</sup>是基于注意力机制的监督学习模型，整体上由编码器（Encoder）和解码器（Decoder）两部分构成，由 Zhou H 等人于 2021 年提出，在 Tranformer<sup>[19]</sup>的基础上主要对编码器和解码器进行优化改进，结合注意力特征图的长尾分布提出多头概率稀疏自注意力机制（Prob Attention），重点考虑那些对注意力分数贡献重大的主导点积对，忽略那些次要点积对。与传统的自注意力机制对比，概率稀疏自注意力机制的关键在于它的查询矩阵的稀疏特性，不仅显著地降低了计算负担，同时也提升了在处理长期时间序列数据时的性能。另一方面，它通过运用多个独立的投影空间来捕获时序数据的多样化特点，相较于传统的单头注意力机制具有更优的特征分辨力。

除此之外，还引入了一种蒸馏操作（Self-attention Distilling），其主要目的是凸显关键注意力节点。同时，通过最大池化，进一步强化了主要特征，并缩减了时间序列的长度，有效降低计算成本。然而，由于 Informer 模型蒸馏部分并未集成自注意力机制，其序列简化操作仅限于蒸馏本身，这会导致部分关键信息的丢失。

## 2 MWCEN-Informer

本文提出的 MWCEN-Informer 框架如图 2 所示，数据经过处理输入 MWCEN 进行时序特征提取，然后输入 Informer 进行 RUL 预测，其中 MWCEN 主要包括 3 个部分：用于考虑工况影响的动态工况编码模块，提取丰富时序特征的一维多尺度混合卷积模块和基于通道注意力机制的多分支通道特征增强模块。

### 2.1 动态工况编码

设备所运行的工况与设备退化过程关系密切，在设备全生命周期中，不同时刻设备往往处于多种不同工况中，高负

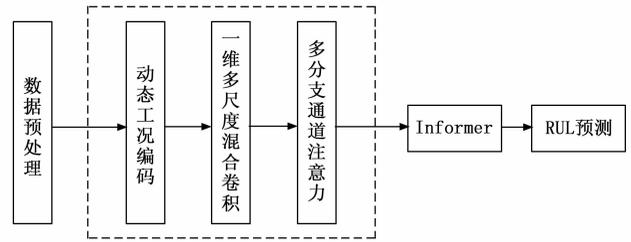


图 2 MWCEN-Informer 框架

载、高温、高湿等极端工况会导致 RUL 变化幅度增加，普通工况则影响较轻。若不考虑运行工况直接使用传感器时序数据进行预测，一方面违背了机器学习的独立同分布假设，会导致模型的泛化能力弱，另一方面忽略了历史工况状态对设备 RUL 的影响，无法更准确地预测设备 RUL。

本文设计了基于设备工况的动态工况编码模块（DCTE, dynamic condition-based time encoding）的工况编码模块。该算法考虑工况类型及单一工况持续时间，结合嵌入（Embedding）编码和反映相对持续时间的标量进行多类别时间编码，能有效帮助模型学习到历史工况状态信息。首先，初始化嵌入维度，取工况数量  $n$  作为嵌入维度。然后，将工况从 1 开始排序，将工况序号通过嵌入层映射为嵌入向量。其次，通过全连接层将嵌入编码映射为最终编码，最终编码维度与时序数据维度保持一致，即设备传感器的数量  $m$ 。接着，按式 1 计算工况持续系数：

$$c_j^i = \frac{t_j^i}{T_j} \quad (1)$$

其中： $c_j^i$  为第  $j$  台设备的第  $i$  种工况的工况持续系数， $t_j^i$  为第  $j$  台设备的第  $i$  种工况的工况持续时间， $T_j$  为第  $j$  台设备的全生命周期时间。

最后，将得到的工况持续系数矩阵  $C$  与嵌入编码矩阵  $S$  相乘，与原始设备时序数据  $D$  叠加，作为输出，其计算表达式如下：

$$\begin{cases} S = X \times E \times FC \\ D^E = C \times S + D \end{cases} \quad (2)$$

其中： $X$  为工况独热编码矩阵， $E$  为 Embedding 参数矩阵， $FC$  为全连接层参数矩阵， $C$  为工况持续系数矩阵， $D$  为原始设备时序数据。

动态工况编码模块如图 3 所示。

过程如算法 1 所示。

算法 1：动态工况编码

1) 输入：工况数量  $n$ ，传感器数量  $m$ ，设备数量  $s$ ，设备时序数据  $D$

2) 初始化：

- (a) 初始化嵌入层  $E$ ；
- (b) 初始化全连接层  $FC$ ；

3) 计算工况持续系数：

- (a)  $j=1, i=1$ ；
- (b) 按设备序号  $j$  对设备时序数据  $D$  进行分离得到第  $j$  台设备的全生命周期时序数据  $D_j$ ；

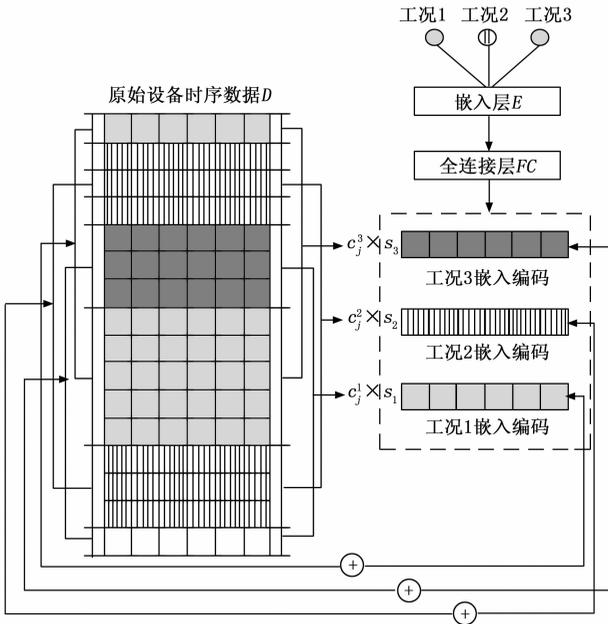


图 3 动态工况编码

- (c) 计算每台设备的全生命周期时间  $T_j$ ;
- (d) 遍历  $D_j$  得到第  $i$  中工况的全局持续时间  $t_j^i$ , 计算  $c_j^i = \frac{t_j^i}{T_j}$ , 直到  $i = n$ ;
- (e)  $j = j + 1$ ;
- 4) 直至  $j = s$ , 得到每台设备的各个工况持续系数;
- 5) 将工况序号输入嵌入层  $E$  和输入全连接层  $FC$  映射到嵌入编码;
- 6) 将工况持续系数与嵌入编码  $S$  相乘;
- 7) 输出: 最终动态工况编码与  $D$  叠加后结果  $D^E$

最后得到可嵌入的 DCTE 编码, 和原始时序数据进行叠加, 作为该模块的输出。

### 2.2 一维多尺度混合卷积

传统一维卷积对所有传感器数据进行卷积操作, 会将某些传感器的噪声和无关信息引入卷积结果, 且难以适应每个传感器的特殊特征。为了学习到每个传感器的时序特征和全局信息, 本文设计了一维多尺度混合卷积模块。

首先, 输入经过动态工况编码后的设备时序数据, 并对其进行传感器层面的数据分离, 为每个传感器应用不同尺度的卷积核, 可以学习到不同感受野下的时序特征, 包括短期记忆和长期记忆。其次, 由于同一传感器时序数据进行不同尺度卷积后序列长度不一致, 使用前置零填充方法, 在卷积后序列数据前端进行零填充, 以确保卷积后序列的时间关系。然后, 对全部传感器数据进行传统一维卷积, 提取传感器之间的特征关系。最后, 将多尺度独立一维卷积后的各个传感器序列数据和传统一维卷积后的全局数据进行随机错位拼接作为输出, 其计算表达式如下:

$$\begin{cases} D^{IC} = \sum_{j=1}^s W_1^{IC} \otimes D_j^E + \sum_{j=1}^s W_3^{IC} \otimes D_j^E + \sum_{j=1}^s W_5^{IC} \otimes D_j^E \\ D^{CC} = W_1^{CC} \otimes D^E + W_3^{CC} \otimes D^E + W_5^{CC} \otimes D^E \\ D^C = randIndex(D^{IC} + D^{CC}) \end{cases}$$

(3)

其中:  $\otimes$  表示卷积操作,  $W_1^{IC}$ 、 $W_3^{IC}$ 、 $W_5^{IC}$ 、 $W_1^{CC}$ 、 $W_3^{CC}$ 、 $W_5^{CC}$  分别代表尺度 1 独立卷积核、尺度 3 独立卷积核、尺度 5 独立卷积核、尺度 1 全局卷积核、尺度 3 全局卷积核、尺度 5 全局卷积核,  $s$  为设备传感器数量,  $D_j^E$  为第  $j$  台经过工况编码后的设备时序数据,  $D^E$  为全部经过工况编码后的全部设备时序数据,  $randIndex$  表示随机排列,  $D^{IC}$  表示独立卷积结果,  $D^{CC}$  表示混合卷积结果。

一维独立多尺度卷积操作如图 4 所示, 其中卷积操作分别使用尺度为 1, 3, 5 大小的卷积核进行一维卷积操作。

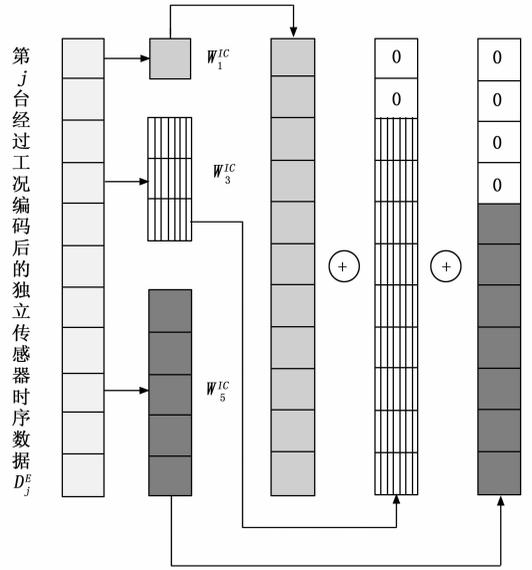


图 4 一维独立多尺度卷积

具体流程如算法 2 所示。

#### 算法 2: 一维多尺度混合卷积

- 1) 输入: 传感器数量  $m$ , 设备数量  $s$ , 动态工况编码叠加后的设备时序数据  $D^E$
- 2) 独立卷积:
  - (a)  $j = 1$ ;
  - (b) 按设备序号  $j$  对  $D^E$  进行分离得到第  $j$  台设备编码后的全生命周期时序数据  $D_j$ ;
  - (c) 对  $D_j$  分别使用尺度为 1, 3, 5 的一维卷积核进行卷积, 得到卷积结果  $D_j^{IC}$ ;
  - (d) 对卷积后结果  $D_j^{IC}$  进行前置零填充;
  - (e)  $j = j + 1$
  - (f) 重复以上步骤, 直至  $j = s$ ;
  - (g) 对以上卷积结果进行拼接, 得到  $D^{CC}$
- 3) 全局卷积:
  - (a) 对  $D^E$  分别使用尺度为 1, 3, 5 的一维卷积核进行卷积并拼接, 得到卷积结果  $D^{CC}$ 。
  - (b) 对卷积后结果  $D^{CC}$  进行前置零填充;
- 4) 对  $D^{IC}$  和  $D^{CC}$  进行随机错位, 再进行拼接;
- 5) 输出:  $D^C$

### 2.3 多分支通道增强

设备的监测信号往往是由多个传感器数据共同组成的,

经过 2.2 处理后，会产生大量的特征数据，然而并不是所有序列数据都对设备 RUL 预测提供有效帮助，为了降低数据冗余，突出关注那些对预测任务至关重要的特征通道，同时抑制那些对任务无贡献的通道。本文提出一种改进的多分支通道增强模块，首先，输入经过独立多尺度混合卷积的序列数据，输入序列数据维度为  $D^C$ ，将其分为  $b$  个分支，每个分支的维度为  $d_k$ 。然后，对于每个分支，通过通道注意力机制学习每个分支下各通道的权重。其次，将每个分支下的通道按照权重进行加权拼接，得到每个分支的输出。最后，将所有分支的结果进行拼接作为输出，其计算表达式如下：

$$D^F = \sum_{j=1}^b W_j^f \times D_j^C \quad (4)$$

其中： $W_j^f$  为第  $j$  分支的注意力权重参数矩阵， $D_j^C$  为经过卷积操作后第  $j$  分支的时序数据。

多分支通道增强操作如图 5 所示。

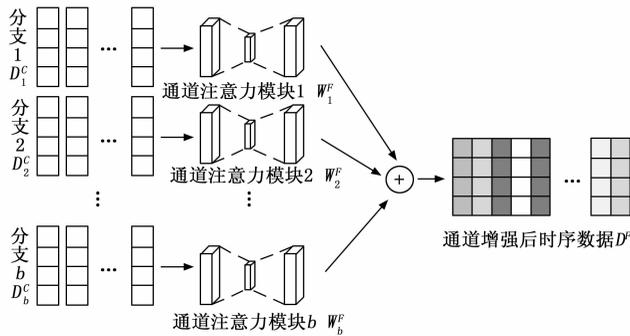


图 5 多分支通道增强

具体流程如算法 3 所示。

算法 3：多分支通道增强

- 1) 输入：分支数量  $b$ ，独立多尺度混合卷积后设备时序数据  $D^C$
- 2) 初始化：
  - (a) 对  $D^C$  进行分支划分，每个分支维度为  $d_k = d_{model} / b$ ；
  - (b) 对每个分支进行通道注意力学习；
  - (c) 将每个分支下的通道按照学习到的权重进行加权拼接，得到每个分支的输出；
  - (d) 将所有分支结果进行拼接
- 3) 输出： $D^F$

### 3 实验结果及讨论

本文采用 NASA Turbofan Jet Engine Data Set<sup>[20]</sup> 作为研究数据集，该数据集包含由 NASA 卓越故障预测中心使用 C-MAPSS 在不同操作条件和故障模式组合上模拟得到的 4 组不同的发动机传感器数据子集，包含多种工况，已经被广泛地应用于设备故障预测研究。本文实验软硬件环境如下：GPU 为 NVIDIA GeForce GTX 960，CPU 为 Intel Core i5-3470，内存为 16 G，操作系统为 Windows 10，编程语言为 Python 3.10，深度学习框架为 Pytorch。

#### 3.1 实验步骤

首先对传感器采集到的时序数据进行处理，然后对数

据集按比例进行划分，再输入 MWCEN-Informer 混合模型进行训练，最后利用测试数据集验证模型 RUL 预测性能，具体步骤如下：

1) 数据预处理：采用均值滤波 (Mean Filtering) 来平滑数据集；使用 Canopy-Kmeans 算法<sup>[21]</sup> 对设备时序传感器数据进行工况聚类并进行 Z-score 归一化处理；采用分段线性函数为发动机设置 RUL 标签，将 RUL 阈值设置为 125 作为初期值；

2) 模型训练：按照表 1 超参数训练 MWCEN-Informer 模型；

表 1 模型超参数设置

超参数名称	参数值	超参数名称	参数值
训练轮数	100	窗口大小	30
优化器	Adam	标签数据长度	10
批大小	32	预测数据长度	1
学习率	0.000 1	DropoutRate	

3) RUL 预测及评估：本文的训练集与测试集的划分比例为 4 : 1，将 20% 的时序数据进行上述同样的处理，通过滑动窗口构建测试数据集，进行模型验证，计算设备的剩余寿命值；为了更好地评估 MWCEN-Informer 混合模型在设备 RUL 预测的性能，采用均方根误差 (RMSE) 和分段误差评分函数 (S-Score) 作为评价指标，计算公式如下：

$$Error = y'_i - y_i \quad (5)$$

$$RMSE = \sqrt{\frac{1}{M} \sum_1^M Error^2} \quad (6)$$

$$S_m = \begin{cases} e^{-Error/13} - 1, & Error < 0 \\ e^{Error/10} - 1, & Error \geq 0 \end{cases} \quad (7)$$

$$S-score = \sum_1^M S_m \quad (8)$$

式中， $y_i$  和  $y'_i$  分别为当前时刻设备 RUL 的真实值和预测值， $M$  为测试集中样本的数量。

RMSE 对于提前 RUL 和滞后 RUL 的惩罚程度都是一致的，而 S-score 对提前 RUL 预测结果的惩罚远小于滞后 RUL 预测结果，更符合实际应用场景，提前预测失效时间能规避严重风险。

#### 3.2 预测结果分析

为了验证本文所提出模型与主流技术 LSTM<sup>[22]</sup> 的性能差异，本文从 4 个独立测试集中随机抽取了一台发动机进行测试。如图 6 所示的预测曲线所示，无论是在发动机正常工作还是在其退化期间，与 LSTM 相比，本文模型的 RUL 预测值与真实值之间具有更好的一致性。对于 FD001、FD003，由于只存在一种工况，本文模型 (Proposed) 与 LSTM 相比差异不大，动态工况编码无法通过工况变化来学习到更多信息，反而引入了部分噪声，导致发动机退化早期预测偏差较大，但随着时间推移，通过一维多尺度卷积和多分支通道增强，RUL 预测迅速靠近真实 RUL。对于 FD002 和 FD004 而言，受工况复杂性的影响，相较于

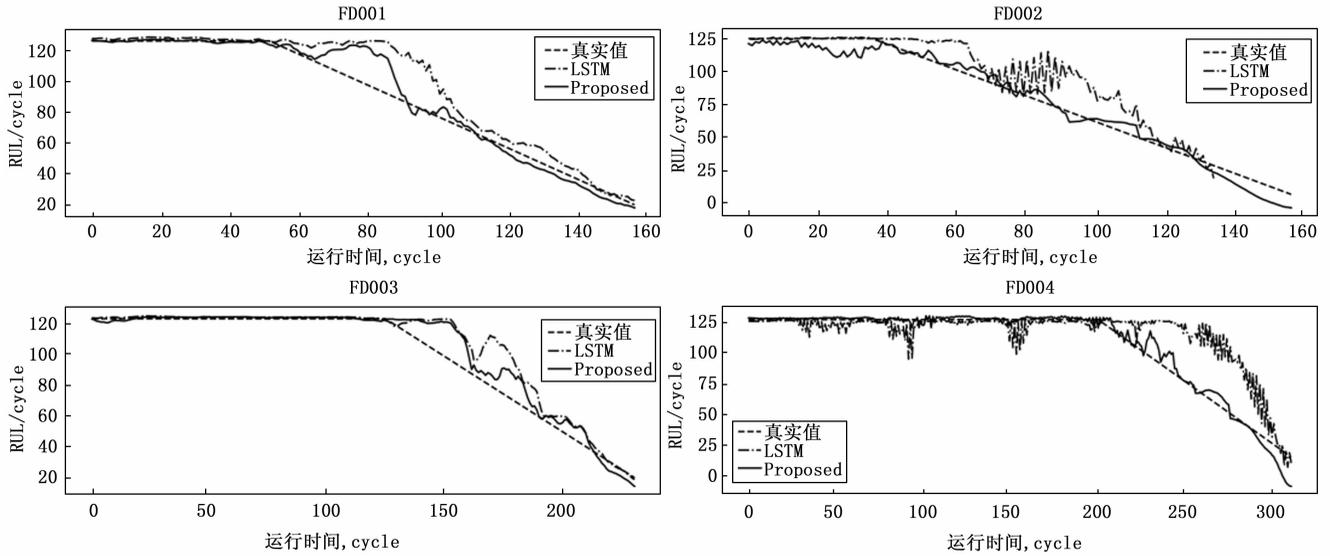


图 6 随机单个发动机 RUL 预测结果

LSTM 的预测结果, 本文模型表现出显著的差异性。尽管在早期的预测阶段出现了波动, 但随着 RUL 的缩短, 预测精度逐步提高。此外, 在退化末期, 预测值普遍低于真实值, 这一保守的预测倾向对于真实的工程应用更具有价值, 有助于更有效地预防发动机的潜在故障。

图 7 展示了本研究的模型在多个数据集上, 对所有发动机 RUL 预测结果的偏差分布情况。通过该图明显可以观察到, 在涵盖不同运行工况以及故障类型的发动机数据集中, 本方法所预测的 RUL 偏差主要集中在零点附近, 展现了预测的稳定性和可靠性。

对于具有单一运行工况的 FD001 和 FD003 数据集, 整体预测偏差偏向负偏差, 集中在  $-10 \sim 0$  区间, 对于含有多种运行工况与故障类型的 FD002 和 FD004 数据集, 集中分布在  $-20 \sim 0$  区间, 在 4 个不同的数据集中进行的测试表明, 本文模型展示出良好的泛化能力, 更加贴合真实世界的工程应用场景。它能够及时预测出设备的退化情况, 由于动态工况编码模块的工况信息抽取能力, 使得模型能根据设备的实际运行情况动态地进行工况特征抽取, 从而构建更为准确的设备退化趋势模型。综上, 本文方法在不同的工况条件下具有较高的预测准确度, 并且本文模型的适用性不局限于特定领域, 能够被应用于其他类似的设备系

统中, 具有较高实际工程应用价值。

### 3.3 消融实验

为了确认所提模型各部分的有效性, 使用了各个模块进行组合测试, 并分析这些组合模块对整体预测能力的具体影响。同样地, 性能评价指标为 RMSE 与 S-score, 用于比较本文模型 (Proposed) 和其他 4 个组合模型在预测性能上的差异。4 种对照组合模型具体包括:

- 1) 基础模型 (Informer);
- 2) 动态工况编码结合基础模型 (CTE-Informer);
- 3) 一维多尺度混合卷积结合基础模型 (ODMHC-Informer);
- 4) 多分支通道注意力结合基础模型 (MCBA-Informer)。

图 8、图 9 展示了本文模型与这 4 种变体之间的比较结果。首先, 可以看出动态工况编码、一维多尺度混合卷积以及多分支通道注意力均能有效提高基础模型 RUL 预测精度, 且动态工况编码对预测效果的影响大于其他部分。一方面, 由于 FD002、FD004 数据集有 6 种不同运行工况, 工况变化复杂, 引入工况编码能让模型学习到历史工况知识; 另一方面, 一维多尺度混合卷积虽然能够为模型提供更多的时序特征, 一定程度上能提高 RUL 预测准确度, 但同时也引入了更多噪声, 且缺乏引入充分的先验工况知识, 提高效果有限。另外, 多分支注意力虽然能够关注不同的

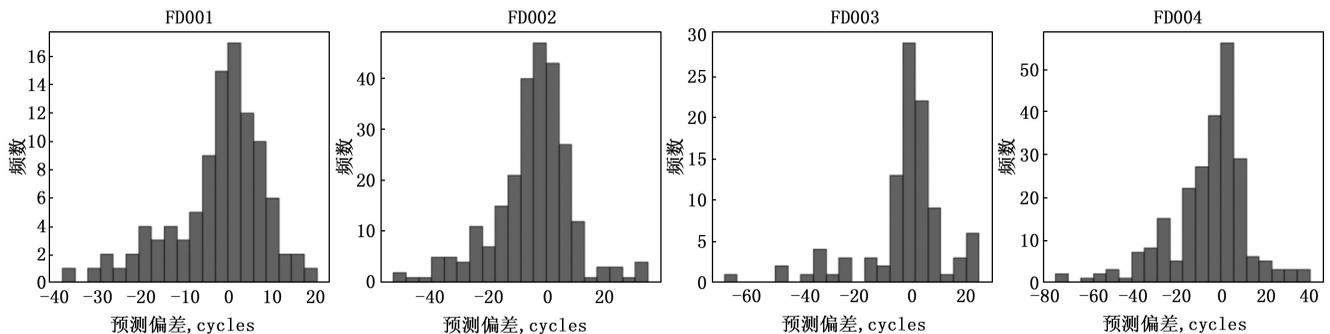


图 7 RUL 预测偏差直方图

特征，但由于缺乏充分的特征提取，效果一般。最后，在所有子数据集的测试中，本文模型均取得了最优结果，证明了 3 个独立模块能够有机融合协同提高整体性能，动态工况编码模块能够有效抽取历史工况信息，一维多尺度混合卷积模块能够提取不同尺度范围的特征，增加原始数据的信息表征能力，多分支通道注意力模块能够关注重要特征，抑制无用信号，提高对设备 RUL 的准确预测能力。

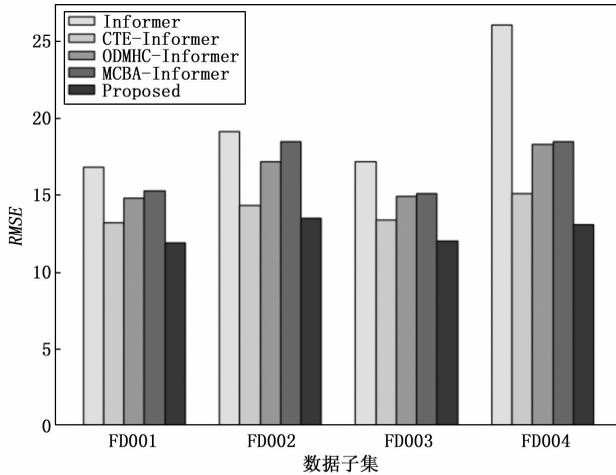


图 8 消融实验 RMSE 对比

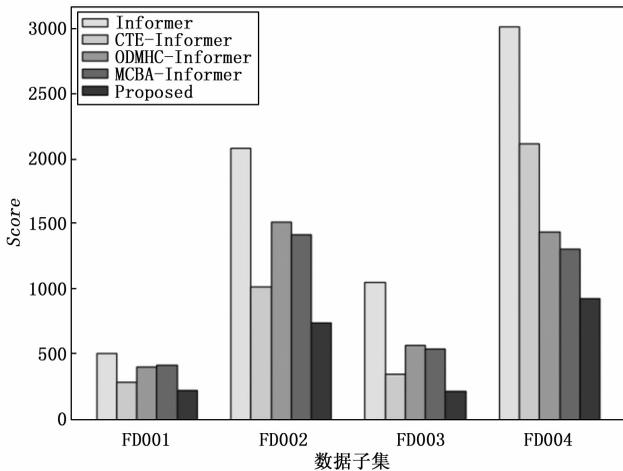


图 9 消融实验 S-Score 对比

### 3.4 对比实验

为了对比分析网络模型的预测性能，使用过去 3 年之中不同机器学习的时序预测模型（LSTM<sup>[22]</sup>、AGCNN<sup>[23]</sup>、PE-Net<sup>[24]</sup>、ASTS2S<sup>[25]</sup>、Transformer<sup>[19]</sup>、Informer<sup>[18]</sup>）与本文模型（Proposed）进行对比。结果如表 2 和表 3 所示，其中粗体突出那些在具体子数据集中表现最优的模型，而下划线用于标识子数据集中表现次优的模型。

从表 2 可知，提出的模型在 3 个子集上的 RMSE 均最低，比其他次优模型的下降了 4.3%、7.8% 以及 14.7%，在 FD003 子集中，ASTS2S 模型的 RMSE 略低于本文模型，综合计算可得 RMSE 平均下降 5.5%。然而 ASTS2S 相对

于本文模型 S-Score 更高，表明了本文模型在确保预测精度的同时，更贴切地符合实际的工程应用场景需求。

从表 3 可以看出，在工况较为复杂的 FD002 和 FD004 子集中，本文所提模型的 S-Score 表现明显优于其他对比模型。具体来说，与次优模型相比，本文模型在 FD002 和 FD004 子集中分别实现了 14.2% 和 12.4% 的降幅，整体平均下降了 4.7%。尽管在单一工况下 AGCNN 模型的 S-Score 略低于本文模型，但本文模型在复杂的工况下的 S-Score 却比 AGCNN 降低了较大幅度，分别是 49.6% 和 72.3%，同时在 RMSE 上亦表现出优势，充分展现了本文模型在复杂工况下的优越性。相较于传统的 LSTM、Transformer 等主流技术，本文模型在预测复杂工况的设备剩余使用寿命时，表现了更好的退化特征抽取和故障趋势预测能力，显著提升了预测的精度。此外，该方法无需依赖太多物理背景或专家经验，可以广泛应用于不同领域和场景的设备剩余使用寿命预测问题，在实际应用中具有较高的实用价值和广阔的应用前景。

表 2 不同模型 RMSE 对比

Models	RMSE			
	FD001	FD002	FD003	FD004
LSTM <sup>[22]</sup>	16.14	24.49	16.18	28.17
AGCNN <sup>[23]</sup>	12.42	19.43	13.39	21.50
PE-Net <sup>[24]</sup>	13.98	14.69	12.33	15.40
ASTS2S <sup>[25]</sup>	12.63	14.65	11.44	16.66
Transformer <sup>[19]</sup>	17.35	21.38	18.89	27.33
Informer <sup>[18]</sup>	16.80	19.13	17.23	26.11
Proposed	11.88	13.51	12.01	13.13

表 3 不同模型 S-Score 对比

Models	S-Score			
	FD001	FD002	FD003	FD004
LSTM <sup>[22]</sup>	523.04	3016	1067	3351
AGCNN <sup>[23]</sup>	225.51	1492	227.09	3392
PE-Net <sup>[24]</sup>	280.87	881.73	272.85	1103.18
ASTS2S <sup>[25]</sup>	243	876	263	1074
Transformer <sup>[19]</sup>	671.36	3678	2136	3843
Informer <sup>[18]</sup>	523.81	2093	1067	3031
Proposed	239.31	752.13	231.88	940.35

### 4 结束语

为了充分挖掘设备传感器时序原始数据中蕴含的信息，并且让模型深入洞察并把握设备时间序列数据的时间相关性，增强设备 RUL 预测精度，本文提出了一种基于多尺度工况增强网络及 Informer 混合模型的设备 RUL 预测方法，多尺度工况增强网络主要分为 3 个部分，分别为动态工况编码模块、一维多尺度混合卷积模块和多分支通道注意力模块。通过动态工况编码模块额外提取出工况特征，多尺度混合卷积模块能有效提取充分的传感器时序特征，最后通过多分支通道注意力模块聚焦重要特征，通过在航空涡

扇发动机数据集上进行大量实验验证,可以有效提高设备故障诊断的准确性,对及时发现潜在故障,保障设备的正常运行具有重要的价值。

未来的研究方向将聚焦于考虑不同操作条件、不同故障频次等更多原因对设备 RUL 预测的影响,并在真实工业场景的数据集上进行测试,进一步提高模型的泛化能力,使得该模型能够更好地适应不同情况下的设备 RUL 预测,为设备使用和维保提供更可靠的质量控制工具。

#### 参考文献:

- [1] LI Y, CHEN Z, HU C, et al. Bearing remaining useful life prediction with an improved CNN-LSTM network using an artificial gorilla troop optimization algorithm [J]. Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability, 2024: 1748006X231222397.
- [2] MITICI M, PATER ID, BARROS A, et al. Dynamic predictive maintenance for multiple components using data-driven probabilistic RUL prognostics; the case of turbofan engines [J]. Reliability Engineering and System Safety, 2023, 234: 109199.
- [3] COSTA N, SANCHEZ L. Variational encoding approach for interpretable assessment of remaining useful life estimation [J]. Reliability Engineering and System Safety, 2022, 222: 108353.
- [4] CHUNGKHAM H S. Estimation of working life expectancy: application of a multistate Markov model [J]. Innovation in Aging, 2023, 7 (s1): 314.
- [5] ARDIAN H, SULETRA I W, PRIYANDARI Y, et al. Remaining useful life prediction of induction motor using ARMA model and thermal imaging sensor [C] //AIP Conference Proceedings, New York: AIP Publishing, 2023, 2824 (1): 1-10.
- [6] CAESARENDRA W, WIDODO A, PHAM H T, et al. Machine degradation prognostic based on RVM and ARMA/GARCH model for bearing fault simulated data [C] //2010 Prognostics and System Health Management Conference, Macau: IEEE, 2010: 1-6.
- [7] VICHARE N M, PECHT M G. Prognostics and health management of electronics [J]. IEEE Transactions on Components and Packaging Technologies, 2006, 29 (1): 222-229.
- [8] YAN M, WANG X, WANG B, et al. Bearing remaining useful life prediction using support vector machine and hybrid degradation tracking model [J]. ISA Transactions, 2020, 98: 471-482.
- [9] ALFARIZI M G, TAJIANI B, VATN J, et al. Optimized random forest model for remaining useful life prediction of experimental bearings [J]. IEEE Transactions on Industrial Informatics, 2022: 1-10.
- [10] 袁志响, 卢文壮, 刘杰, 等. 基于改进一维卷积神经网络和通道注意机制的刀具磨损预测研究 [J]. 工具技术, 2023, 57 (6): 42-46.
- [11] DEUTSCH J, HE D. Using deep learning-based approach to predict remaining useful life of rotating components [J]. IEEE Transactions on Systems, Man, and Cybernetics: Systems, 2017, 48 (1): 11-20.
- [12] 曾友渝, 谢强. 基于改进 RNN 和 VAR 的船舶设备故障预测方法 [J]. 计算机科学, 2021, 48 (6): 184-189.
- [13] DA COSTA P R D O, AKCAY A, ZHANG Y, et al. Attention and long short-term memory network for remaining useful lifetime predictions of turbofan engine degradation [J]. International Journal of Prognostics and Health Management, 2019, 10 (4): 1-12.
- [14] LI J, JIA Y, NIU M, et al. Remaining useful life prediction of turbofan engines using CNN-LSTM-SAM approach [J]. IEEE Sensors Journal, 2023: 10241-10251.
- [15] ZHANG L, WANG B, YUAN X, et al. Remaining useful life prediction via improved CNN, GRU and residual attention mechanism with soft thresholding [J]. IEEE Sensors Journal, 2022, 22 (15): 15178-15190.
- [16] MAO W, HE J, TANG J, et al. Predicting remaining useful life of rolling bearings based on deep feature representation and long short-term memory neural network [J]. Advances in Mechanical Engineering, 2018, 10 (12): 1687814018817184.
- [17] ZHANG Z, SONG W, LI Q. Dual-aspect self-attention based on transformer for remaining useful life prediction [J]. IEEE Transactions on Instrumentation and Measurement, 2022, 71: 1-11.
- [18] ZHOU H, ZHANG S, PENG J, et al. Informer: Beyond efficient transformer for long sequence time-series forecasting [C] //Proceedings of the AAAI Conference on Artificial Intelligence, New York, 2021, 35 (12): 11106-11115.
- [19] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need [J]. Advances in Neural Information Processing Systems, 2017, 30: 1-11.
- [20] 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: IEEE Press, 2008: 1-9.
- [21] 郑小东. 设备全寿命周期管理与故障预测研究 [D]. 绵阳: 西南科技大学, 2021.
- [22] LAN G, LI Q, CHENG N. Remaining useful life estimation of turbofan engine using LSTM neural networks [C] //2018 IEEE CSAA Guidance, Navigation and Control Conference (CGNCC), Xiamen: IEEE, 2018: 1-5.
- [23] 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 Trans on Industrial Informatics, 2020, 17 (2): 1197-1207.
- [24] JIN R B, WU M, WU K Y, et al. Position encoding based convolutional neural networks for machine remaining useful life prediction [J]. IEEE/CAA Journal of Automatica Sinica, 2022, 9 (8): 1427-1439.
- [25] RAGAB M, CHEN Z H, WU M, et al. Attention-based sequence to sequence model for machine remaining useful life prediction [J]. Neuro Computing, 2021, 466: 58-68.