

基于多注意力和变分编码时序网络的发动机 剩余使用寿命预测方法

杨凯旋¹, 赵书健², 魏佳隆¹, 李良², 苏本淦², 刘扬¹, 赵振¹

(1. 青岛科技大学 信息科学技术学院, 山东 青岛 266061;

2. 青岛淄柴博洋柴油机股份有限公司, 山东 青岛 266701)

摘要: 准确可靠的剩余使用寿命 (RUL) 预测结果可为决策者提供有价值的信息, 以采取适当的维护策略, 最大限度地利用设备, 避免昂贵的故障维修费用; 为了从高噪声的真实工况中对发动机故障进行有效诊断, 提出了一种融合多注意力机制和变分编码的时序预测模型 (MA-VBLSTM), 通过嵌入多注意力机制获得所有特征在空间维度和通道维度的不同权重, 以提高退化特征的提取能力; 采用变分编码器进行退化信息编码并学习数据间深度隐藏的信息; 利用双向长短期记忆网络的长短期时序数据双向处理能力实现发动机 RUL 的预测; 实验结果表明, 该模型在发动机 CMAPSS 数据集的 FD001、FD002、FD003、FD004 子数据集上, RMSE 和 Score 值相比现有方法分别平均降低 5.27% 和 10.70%、1.37% 和 1.68%、6.37% 和 26.94%、3.02% 和 2.06%。

关键词: 剩余使用寿命; 多注意力; 变分编码; 时序预测; 深度学习

Prediction Method of Engine Remaining Useful Life Based on Multi-attention and Variational Coding Sequential Network

YANG kaixuan¹, ZHAO Shujiang², WEI Jialong¹, LI Liang², SU Bengan², LIU Yang¹, ZHAO Zhen¹

(1. Qingdao University of science and technology, Qingdao 266061, China;

2. Qingdao ZiChaiBoYang Diesel Engine Co., Ltd., Qingdao 266701, China)

Abstract: The prediction results of accurate and reliable residual service life (RUL) can provide valuable information for decision makers, appropriate maintenance strategies are adopted to maximize the use of equipment, and avoid expensive maintenance costs. In order to effectively diagnose engine faults in real working condition with high noise, a sequential prediction model (MA-VBLSTM) integrating multiple attention mechanisms and variational coding is proposed. The different weights of all features in the spatial and channel dimensions are obtained by embedding the multiple attention mechanisms to improve the extraction ability of degraded features; The variational autoencoders are used to encode the degraded information and learn the deeply hidden information between the data; The prediction of the engine RUL is realized by using the bidirectional processing capability with the long and short time series data in the bidirectional long short-term memory network. Compared with existing methods, the RMSE and Score values of the proposed model on FD001, FD002, FD003 and FD004 sub-data sets are reduced by 5.27% and 10.70%, 1.37% and 1.68%, 6.37% and 26.94%, 3.02% and 2.06%, respectively.

Keywords: remaining useful life; multi-attention; variational coding; sequential prediction; deep learning

0 引言

预测与健康管理 (PHM, prognostics and health management) 可以根据各种传感器采集的监测数据提供机械系统的健康信息, 并有助于制定合理的维修计划^[1], 确保机械系统可靠性和安全性。剩余使用寿命 (RUL, remaining

useful life) 估计是 PHM 的主要任务之一, 准确地预测剩余使用寿命可为预知维护决策提供依据, 减少冗余维护操作和成本, 对于评估健康状态和避免灾难性故障^[2]至关重要。

一般而言, RUL 估计方法大致可以分为基于模型方法和数据驱动方法^[3]。基于模型的方法利用机械系统的先验

收稿日期: 2023-02-25; 修回日期: 2023-03-02。

项目基金: 国家自然科学基金(62201314); 山东省自然科学基金(ZR2020QF007); 强链计划(23-1-2-qdjh-18-gx)。

作者简介: 杨凯旋(1998-), 男, 山东滨州人, 硕士研究生, 主要从事设备寿命预测方向的研究。

通讯作者: 赵振(1982-), 男, 山东德州人, 博士研究生, 副教授, 主要从事人工智能和大数据分析方向的研究。

引用格式: 杨凯旋, 赵书健, 魏佳隆, 等. 基于多注意力和变分编码时序网络的发动机剩余使用寿命预测方法[J]. 计算机测量与控制, 2023, 31(7): 42-48, 56.

知识, 建立数学模型来描述机械设备的退化过程, 然后利用实时监测数据更新模型参数实现 RUL 估计。常见的基于模型的方法包括 Wiener 过程模型^[4], Gamma 模型^[5], Weibull 分布^[6]等。然而, 该类方法需要充分的先验知识来描述退化过程, 但在复杂的系统中很难获得准确的失效信息。数据驱动的预测方法基于历史监测数据, 与对应系统的健康状态之间建立一个非线性的映射关系。经典的数据驱动模型, 如支持向量回归^[7]、人工神经网络^[8]、马尔科夫模型^[9]等, 主要包含人工特征设计和退化行为学习两个步骤。然而, 这些方法需要过多的人工干预, 且忽略了人工特征设计与退化行为学习之间的联系, 从而限制了估计精度。

随着人工智能的快速发展, 深度学习具有良好的非线性映射效果和自主的特征提取能力, 可以通过多层连接从大量数据中提取高效的特征, 已成功应用于许多领域。在 RUL 估计中, 可以通过建立一种多层网络连接组合的模型, 在已获取的原始数据序列上独立地学习层次化的特性, 而无需人工设计特征。如深度神经网络 (DNN, deep neural networks)^[10]、深度置信网络 (DBN, deep belief networks)^[11]、卷积神经网络 (CNN, convolutional neural networks)^[12]、循环神经网络 (RNN, recurrent neural network)^[13]、长短时记忆网络 (LSTM, long short term memory networks)^[14]等, 已被应用于 RUL 预测, 并取得了较好的预测结果。但是现有深度学习方法通常假设不同传感器采集的监测数据对 RUL 估计的贡献相等。但实际上, 不同时间步长的监测数据往往包含不同程度的退化信息, 且它们对 RUL 预测的贡献是不等的。

注意力机制广泛应用于各种任务中, 例如机器翻译和视觉识别等, 注意力机制通过计算不同特征的注意力概率, 对模型中的不同特征赋予不同权重, 有效地提高网络的表达能力。王欣等^[15]使用长短时记忆神经网络提取序列特征, 然后使用注意力网络学习不同时间域蕴含的退化信息, 以提高估计效果。赵志宏等^[16]提出直接将原始振动信号输入双

向长短期记忆网络 (BiLSTM, bidirectional long short-term memory) 提取特征, 并使用注意力机制对特征分配不同的权重以提升预测精度。张加劲^[17]采用卷积神经网络提取特征和双向长短期记忆网络获取特征中的长短期依赖关系, 并使用注意力机制来突出特征中的重要部分, 提高模型预测的准确率。但上述方法只是考虑运行数据随时间变化趋势蕴含着的退化信息, 未考虑同一时间窗内多维时域指标数据蕴含的反映指标间关联性的空间信息以及同一时间窗内不同的时域指标数据对 RUL 预测的不同重要性, 导致深度学习网络的预测结果不能更加精确。

为解决发动机退化数据中信息无法被充分发掘的问题, 提出了一种融合多注意力 (MA, multi-attention) 和变分编码时序网络 (VBLSTM, variational coding bidirectional long short-term memory networks) 的剩余寿命预测方法。首先, 空间注意力通过自注意力 (SA, self-Attention) 和卷积神经网络提取多维时域指标间的空间信息并分配注意力权重。通道注意力通过池化层和多层感知机赋予不同时域指标以不同的注意力权重, 使模型高效的关注与剩余寿命关系密切的指标; 然后, 通过变分自编码器 (VAE, variational auto-encoder) 的潜在分布映射能力进行低维深度隐藏特征提取, 再利用双向长时记忆网络双向学习复杂的时间前后关联信息, 捕获时间序列中长短期的时序特征; 最终, 通过全连接层融合通道维度时序特征和空间维度时序特征进行剩余使用寿命的预测。实验结果表明, 提出的 MA-VBLSTM 模型在发动机 CMAPSS 数据集的 FD001、FD002、FD003、FD004 子数据集上 RMSE 和 Score 值相比现有方法分别平均降低 5.27% 和 10.70%、1.37% 和 1.68%、6.37% 和 26.94%、3.02% 和 2.06%。

1 MA-VBLSTM 算法

如图 1 所示, 本文提出的 MA-VBLSTM 模型由多注意力网络、变分编码时序网络和全连接输出网络组成。首先, 通过标准化及数据分割等操作将原始发动机运行数据转化为网络的标准输入数据; 然后, 通过多注意力网络有效地

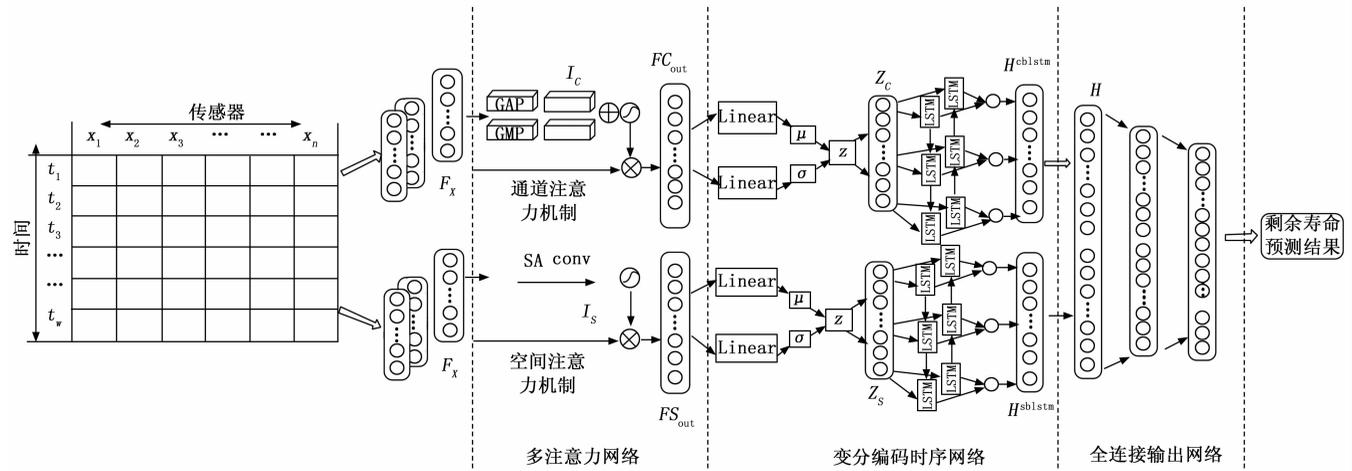


图 1 MA-VBLSTM 模型结构图

进行输入数据中的空间特征和时间特征的权重配比, 过滤或弱化冗余特征和抽取更加重要和关键的退化特征信息; 然后, 通过 VAE 建模两个复杂的条件概率密度函数输出服从一定分布的隐藏变量并使用 BiLSTM 综合考虑序列信息的历史数据和未来数据进行有效地捕获时间序列中远距离相关特征。最后, 通过多层全连接网络将加权后的空间时序特征和通道时序特征输出结果进行转换和整合, 借助其非线性拟合能力将维度逐层降低到与输出标签一致后, 输出整个模型的剩余寿命预测结果。

1.1 多注意力网络

发动机退化数据是高维度且复杂、大数据量、含有信息冗余的时序信息, 其中的噪声和无关信号会影响剩余使用寿命预测的精准度^[18], 因此构建注意力模块突出输入数据中与设备退化关联性更大的特征, 充分提取出与发动机故障的相关信息, 进而输入到模型中进行剩余使用寿命预测。

为了挖掘出具有显著的退化特征数据和区分故障时刻的关键信息, 本文所提的通道注意力机制对通道间信息关系进行建模, 区分不同通道特征的重要性; 空间注意力机制利用自注意力进行权重配比, 结合卷积结构进行达到空间上的关键退化信息捕捉。

通道注意力: 如图 2 所示, 通道注意力通过压缩和聚合操作, 对各通道的的时间维度上下关系进行独立编码进而捕获关键的退化特征。

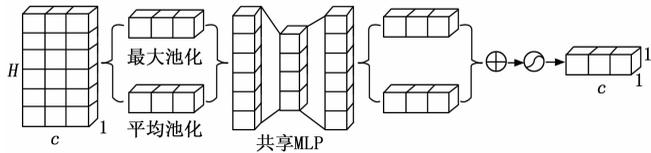


图 2 通道注意力网络结构图

首先, 输入数据 F_x 通过全局平均池化 (GAP, global average pooling) 和全局最大池化 (GMP, global max pooling) 生成压缩通道平均描述符 $F_{avg}^C \in R^{C \times 1 \times 1}$ 和通道最大描述符 $F_{max}^C \in R^{C \times 1 \times 1}$ 。然后, 通过 MLP 对 F_{avg}^C 和 F_{max}^C 进行聚合计算。最后, 将 F_{avg}^C 和 F_{max}^C 映射到原始数据上, 通过 2 个全连接层获得通道权重输出 $I_c \in R^{C \times 1 \times 1}$ 。通道注意力计算方式如下:

$$I_c = fsm(W_1(W_0 \cdot GAP(F_x)) + W_1(W_0 \cdot GMP(F_x))) = fsm(W_1(W_0(F_{avg}^C) + W_1(W_0(F_{max}^C))) \quad (1)$$

式中, fsm 为 sigmoid 函数, W_0 和 $W_1 \in R^{C \times c}$ 为权重参数。最终, 通道注意力的输出表示如下:

$$FC_{out} = I_c * F_x \quad (2)$$

空间注意力: 如图 3 所示, 空间注意力使用自注意力替代双池化的聚焦操作。相对比池化操作, 自注意力机制具备全局的感受野, 卷积结构经自注意力加权后的输出考虑了所有特征的信息, 能够实现不同位置特征间的直接点乘融合、不受特征间距离的影响, 从而得到更全局的特征提

取效果。

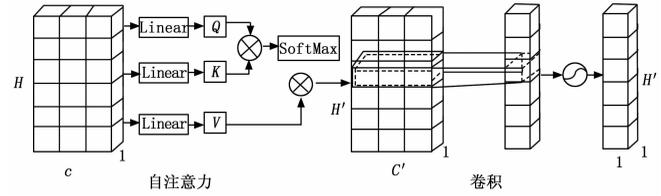


图 3 空间注意力网络结构图

首先, 将输入数据 F_x 分别通过 3 个线性层 (Linear, linear layer), 依次得到查询矩阵 Q 、键矩阵 K 和值矩阵 V ; 随后, 将 Q 与 K 的转置相乘并除以缩放因子后经 Softmax 函数得到自注意力权重矩阵 A ; 然后将 V 和自注意力权重矩阵 A 相乘, 得到自注意力加权后的结果 $attn(F_x)$; 然后, 对 $attn(F_x)$ 采用卷积核大小为 7×1 进行卷积计算生成空间特征描述符。最后, 通过 sigmoid 函数处理获得空间特征权重 $I_s \in R^{1 \times H \times 1}$ 。空间注意力的计算方法如下:

$$I_s = fsm(conv^{7 \times 1}(attn * (F_x))) = fsm(conv^{7 \times 1}(Softmax * (\frac{QK^T}{\sqrt{d_k}})V)) \quad (3)$$

式中, fsm 为 sigmoid 函数, d_k 是缩放因子, $conv^{7 \times 1}$ 表示与滤波器大小为 7×1 的卷积操作。最终, 空间注意力的输出表示如下:

$$FS_{out} = I_s * F_x \quad (4)$$

1.2 变分编码时序网络

如图 4 所示, 变分编码时序网络一方面结合 VAE, 通过变分推理学习机械设备健康状态的分布并映射到隐空间中, 借助隐空间学习到的连续概率分布重建去干扰的深层低维高质量数据, 另一方面通过引入 BiLSTM 对发动机运行数据建立时间序列的依赖关系, 充分挖掘运行数据中的长短期时序特征, 对运行工况、设备差异等因素引起的噪声具有一定的容忍性和鲁棒性^[19]。

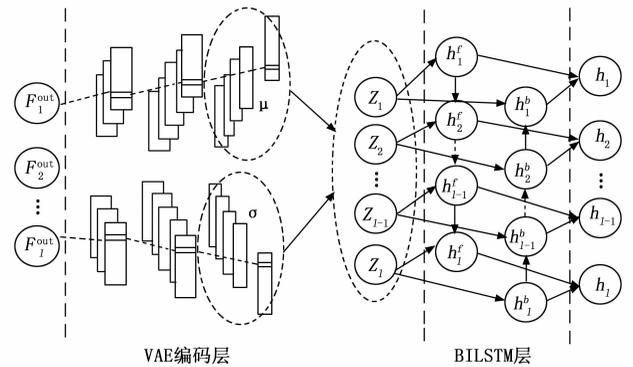


图 4 变分编码时序预测网络结构

变分自编码器 (VAE): VAE 是一种无监督深度学习生成模型的一种形式, 可以通过对训练数据特征的学习进行建模, 进而对测试数据进行纠正。VAE 的结构图如图 5 所示, $x \rightarrow z$ 是 VAE 的编码过程, 记 $q_z(z|x)$, 这个过程是

通过神经网络训练而学习来的。而 $z \rightarrow \hat{x}$ 则是 VAE 的解码过程, 记 $p_\theta(x|z)$ 。

VAE 基于输入数据不断训练编码器生成隐变量 z 的均值 μ 和标准差 σ , 结合标准正态分布随机采样参数 ϵ , 经过重采样生成 z , 见式 (5) 所示, 避免了神经网络训练中反向梯度断裂的问题。

$$z = \mu + \epsilon \cdot \sigma \quad (5)$$

VAE 的损失函数有两个部分组成, 公式如下:

$$\mathcal{L}(\theta, \varphi; x) = -D_{KL}[q_\varphi(z|x) \parallel p(z)] + E_q[\log \cdot 10p\theta(x|z)] \quad (6)$$

该式由两部分组成, 一个是重构损失, 即生成数据与输入数据的重构误差 $E_q[\log \cdot 10p\theta(x|z)]$, 另一个训练生成的模型分布与由数据本身隐变量所服从高斯分布 $p(z)$ 的比较, 这里使用 Kullback-Leibler (KL) 散度 $D_{KL}[q_\varphi(z|x) \parallel p(z)]$ 来衡量两个分布的差距。

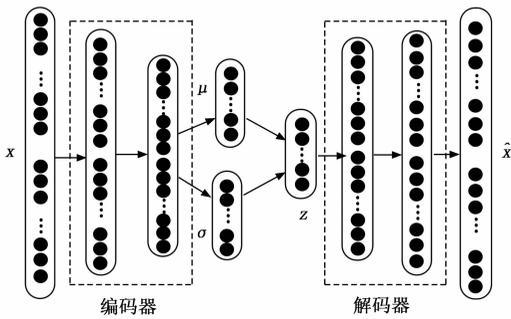


图 5 VAE 网络结构图

双向长短时记忆网络 (BiLSTM): 长短时记忆网络是一种循环神经网络, 更好的捕捉到较长距离的依赖关系, 适用于处理时序数据。它的基本思路是引入了门控装置来处理记忆单元的内存、遗忘、输入程度、输出程度的问题。通过训练, 可以自主适应调整各个门的开启程度, 并通过门控中神经网络参数控制信息的记忆。LSTM 的计算公式为:

$$\begin{cases} i_t = f_{sm}(W_i \cdot [h_{t-1}, z_t] + b_i) \\ \tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, z_t] + b_c) \end{cases} \quad (7)$$

$$f_t = f_{sm}(W_z \cdot [h_{t-1}, z_t] + b_f) \quad (8)$$

$$C_t = f_t \otimes C_{t-1} + i_t \otimes \tilde{C}_t \quad (9)$$

$$\begin{cases} O_t = f_{sm}(W_o \cdot [h_{t-1}, z_t] + b_o) \\ h_t = o_t \otimes \tanh(C_t) \end{cases} \quad (10)$$

式 (7) ~ (10) 分别为输入门、遗忘门、记忆门, 以及输出门。 C_t 为 LSTM 细胞单元状态; z_t 和 h_t 为细胞单元输入和隐藏层输出; f_{sm} 为 sigmoid 激活层, W_i, W_c, W_z, W_o 表示属于不同门或状态的权重矩阵; b_i, b_c, b_z, b_o 表示对应权重的偏置向量; \otimes 为 2 个矩阵对应元素逐个相乘方法。

如图 6 所示, BiLSTM 是将结构相同但信息流方向相反的两个 LSTM 叠加而成。因此可以在前向和后向两个方向上处理序列, 每个隐藏层在一个特定的时间步长可以同时获得过去 (前向) 和未来 (后向) 的信息, 可以提取更全

面的发动机退化特征, 提高网络的预测性能。

将 z_t 输入前向层, 从 0 时刻到 t 时刻正向计算出向前隐含层的输出向量 h_f ; 输入至反向层, 从 t 时刻到 0 时刻反向计算出向后隐含层的输出 h_b 。最后, 将前向层和反向层的输出结果输入至全连接层, 得到最终输出 h 。

$$h = f(h_f + h_b) \quad (11)$$

式中, $f(\cdot)$ 是全连接层的映射函数。

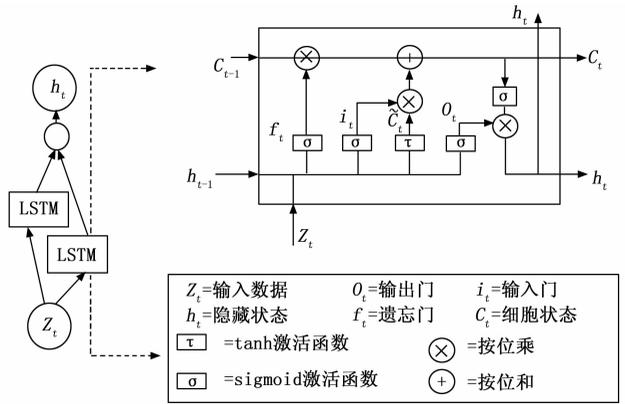


图 6 BiLSTM 网络结构图

1.3 全连接输出网络

全连接输出网络包含两个全连接层, 全连接层的作用是把并行的变分编码时序网络得到的时间维度特征向量和空间维度进行重新展平组装拟合, 并将其作为全连接层的输入, 并通过 ReLU 激活函数实现设备剩余寿命的预测。

$$H = \text{Concat}(H^{\text{tblstm}}, H^{\text{slstm}}) \quad (12)$$

$$y = f_n(W_r H + b_r) \quad (13)$$

式中, H^{tblstm} 为通道时序输出特征向量; H^{slstm} 为空间时序输出特征向量 f_n 为 ReLU 激活函数; W_r 为权重矩阵; b_r 为偏差参数。

2 实验设置

2.1 实验数据介绍

本研究使用的 CMAPSS 数据集是美国宇航局通过商用模块化航空航天推进系统仿真生成的涡扇发动机退化数据。它由 4 个子数据集组成, 具有不同的故障模式和运行条件。每个子数据集包含若干个涡扇发动机退化数据, 其中包括从涡扇发动机不同部件采集到的 21 个包含退化信息的传感器信号, 如温度、压力、速度等, 如表 1 所示。在本实验中, 使用多台涡扇发动机从启动到故障的退化传感器数据的训练数据集对 MA-VBLSTM 模型进行离线训练。然后, 采用退化数据的测试数据集实现在线 RUL 预测。

表 1 航空发动机数据集

数据状况	FD001	FD002	FD003	FD004
训练集发动机台数	100	260	100	249
测试集发动机台数	100	259	100	248
操作模式	1	6	1	6
故障模式	1	2	2	1

2.2 剩余使用寿命标注

发动机在工作初期运行状态良好，性能退化可忽略不计，但到工作末期，发动机的性能会随着时间大幅度降低。若发动机性能快速劣化，将运行数据的标签设定为总运行周期数和现运行周期数之差，会加大 RUL 预测结果的滞后性。

因此通常采用分段线性退化模型设置 RUL 标签，该方法在 C-MAPSS 数据集上已被验证是可行的^[20]。在初始阶段，发动机稳定运行，RUL 值保持不变，即为训练集的 RUL 标签设置阈值，而在退化阶段，RUL 值线性下降。研究表明，将训练集 RUL 标签阈值设置为 130 个运行周期的预测效果较好，RUL 标签设置结果如图 7 所示。

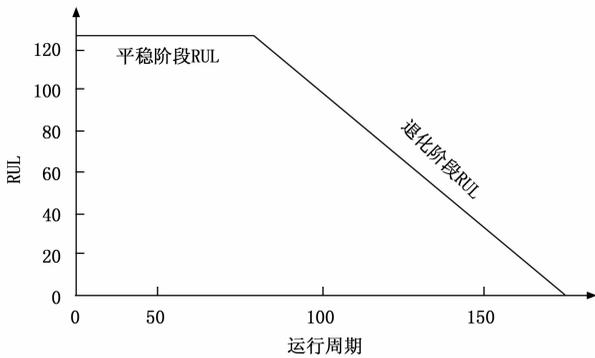


图 7 RUL 标签设置

2.3 数据增强

Mixup 是基于邻域风险最小化（VRM, vicinal risk minimization）原则的数据增强方法，使用线性插值得到新样本数据，对离散样本空间进行连续化，提高邻域内的平滑性，提升深度学习模型的泛化性。

在邻域风险最小化原则下，根据特征向量线性插值将导致相关目标线性插值的先验知识，可得出简单且与数据无关的 Mixup 公式：

$$x_n = \lambda x_i + (1 - \lambda)x_j \quad (14)$$

$$y_n = \lambda y_i + (1 - \lambda)y_j \quad (15)$$

其中： x_i, y_i 是从训练集中随机挑选的时序数据， x_j, y_j 是对应的独热编码标签，通过先验知识确定特征向量的线性插值和对应目标的线性插值，插值生成新数据样本 x_n, y_n 。其中 λ 通过贝塔分布获得，取值范围介于 0 到 1。该算法利用线性插值法，在一定程度上扩展了训练数据的分布空间，提高了模型的泛化性。

2.4 评价指标

本文采用 RMSE（root mean square error）和非对称评分函数（score）两个指标客观地评价了所提出的 MA-VBLSTM 算法的估计性能。RMSE 在回归任务中应用广泛，其定义如下：

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (\hat{r}_i - r_i)^2} \quad (16)$$

式中， N 为样本总个数， \hat{r}_i 为预测值， r_i 为真实值。RMSE

值越小，估计误差越小，估计精度越高。score 函数被 2008 年预后和健康管理数据挑战所采用。公式如下：

$$score = \begin{cases} \sum_{i=1}^N (e^{\hat{r}_i - r/13} - 1), & \text{if } (\hat{r}_i - r < 0) \\ \sum_{i=1}^N (e^{\hat{r}_i - r/10} - 1), & \text{if } (\hat{r}_i - r > 0) \end{cases} \quad (17)$$

与 RMSE 一样，性能越好的 RUL 预测方法所产生的 score 分值越低。但与 RMSE 不同的是，score 函数的早期预测的罚分函数和后期预测是不对称的，由于较晚的预测更有可能造成严重的灾难和重大的经济损失，需要比早期预测更严厉的惩罚。因此，评分函数会对较大的误差进行严厉的惩罚。

3 实验结果与讨论

在本节中，利用 C-MAPSS 数据集验证 MA-VBLSTM 方法的 RUL 预测准确性。首先，采用 CMPASS 数据集验证 MA-VBLSTM 方法有效性并分析预测结果。然后，设计对比实验分析多注意力机制和变分编码时序网络带来的预测效果提升。最后，将所提的 MA-VBLSTM 方法与现有 RUL 预测方法进行了比较。

3.1 模型预测与结果分析

为验证 MA-VBLSTM 方法可以更加充分地挖掘退化数据以提高 RUL 预测准确性的有效性，通过 CMPASS 数据集进行 MA-VBLSTM 模型的训练并分析预测结果。CMPASS 数据集的训练集和测试集记录了发动机在若干运行周期下 3 个操作设置值和 21 个传感器监测数据。为减少噪声提高计算的稳定性，剔除在发动机退化过程中从未变化的操作设置和传感器数据，使用有效的原始数据进行训练。在训练过程中，对原始训练集数据按 8:2 的比例划分训练集和验证集，并采用提前终止训练（early stopping）的正则化方法避免模型出现过拟合现象。将 MA-VBLSTM 方法训练得到的模型应用于测试集，测试集的预测结果如图 8 所示。

从 FD001、FD002、FD003、FD004 的测试集中分别选取 64、11、71 和 37 号测试样本，得到的样本预测结果如图 9 所示。可以观察到，在 4 个测试样本的早期阶段，因引擎开始运转初期时缺少历史数据支持，导致 RUL 估计偏差相对大，同时变化较为剧烈且具有显著的滞后性。但运行一段时间后，预测 RUL 逐渐收敛到了实际 RUL 左右，预测性能比在运行初期明显提升。由于在运行阶段的后期阶段是健康管理的重要时期，这个阶段对故障状况的正确判断能大大提高运行的可靠性和安全性程度，减少维修成本^[22]。所以本文所提模型可以更充分地挖掘深度退化特征并具有较长时序特征的记忆能力，有效提升 RUL 预测精度。

3.2 变分编码时序网络影响分析

本文利用 VAE 的编码器对数据进行编码并学习数据间深度隐藏的信息，输出编码优化处理的高效数据作为输入数据训练 BiLSTM 网络。为了验证 VAE 编码器降低训练阶段输入数据的过拟合性和解决数据特征丢失问题的有效性，在使用相同的 CMPASS 数据进行训练后，利用相同测试数据对 MA-VBLSTM、MA-BLSTM 网络（去除变分编码器）、

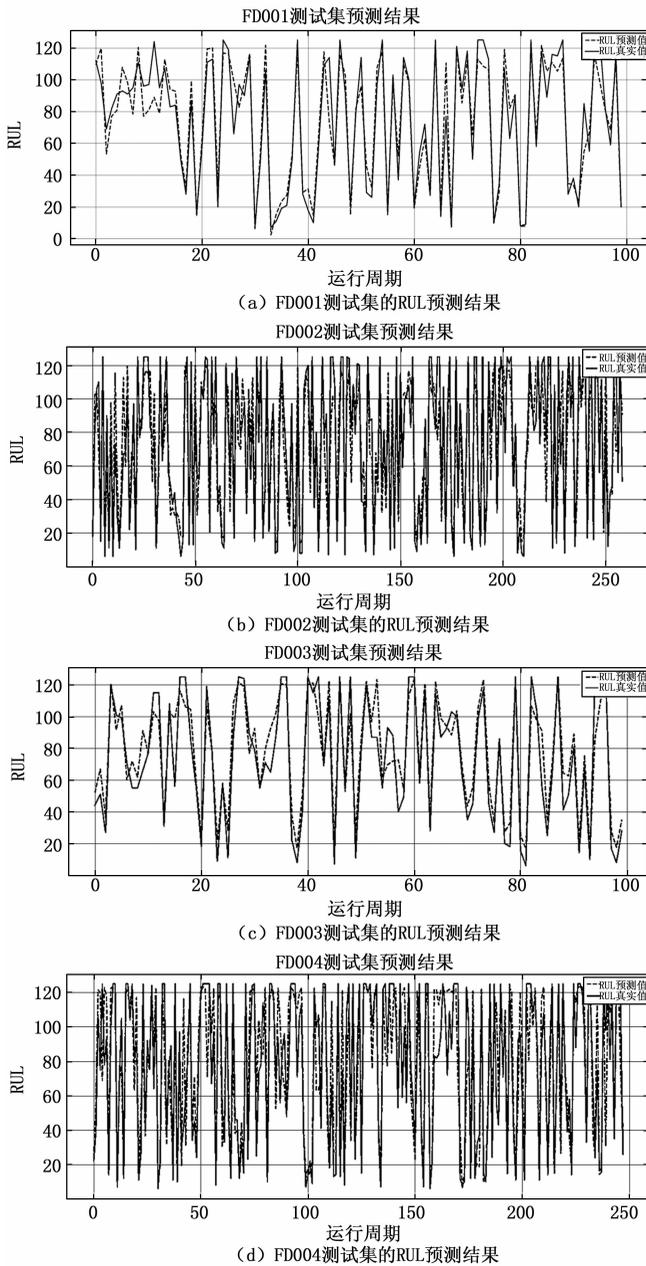


图 8 4 个测试集的 RUL 预测结果

MA-EBILSTM 网络 (使用自编码器代替变分编码器) 进行对比测试。对比实验结果如表 2 所示, MA-VBILSTM 方法与 MA-BLSTM 方法相比, 分别在 FD001、FD002、FD003、FD004 数据集上 RMSE 和 Score 评估值降低 5.43% 和 9.61%、8.02% 和 33.5%、8.69% 和 28.4%、9.41% 和 16.08%。与 MA-EBLSTM 方法相比, MA-VBILSTM 方法分别在 FD001、FD002、FD003、FD004 数据集上 RMSE 和 Score 评估值降低 6.73% 和 12.60%、0.53% 和 9.65%、2.84% 和 14.44%、5.48% 和 9.47%。因此本文所提方法 MA-VBLSTM 方法结合 VAE 将输入数据及各数据维度之间关联性映射到隐空间中, 可以有效降

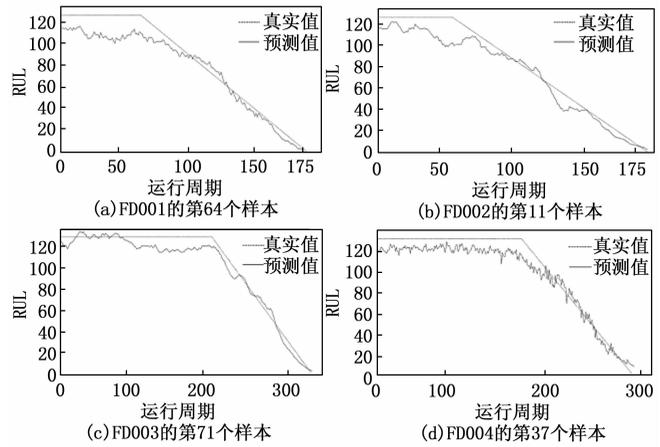


图 9 单个引擎 RUL 预测结果

低数据的过拟合性; 同时引入 BiLSTM 网络双向提取数据的长短期时间依赖, 有效地提升剩余寿命预测的准确率。

3.3 多注意力网络影响分析

为验证本文提出的基于多注意力机制从通道维度和空间维度中自适应选取性能退化关键特征的重要性。本节构建对比实验, 以清楚地说明多注意力特征提取的优势: (1) 仅通道特征作为输入的单流通道注意神经网络, 记作 CA-VBLSTM; (2) 仅空间特征作为输入的单流空间注意神经网络, 记作 SA-VBLSTM。对比实验结果如表 3 所示, 可以观察到本文所提出的 MA-VBLSTM 方法相比 SA-VBLSTM 方法在 RMSE 和 Score 评估值分别在 FD001、FD002、FD003、FD004 数据集上降低 4.74% 和 11.9%、0.88% 和 21.9%、3.45% 和 10.79%、1.29% 和 14.06%。相比 CA-VBLSTM 方法, MA-VBLSTM 方法在 RMSE 和 Score 评估值分别在 FD001、FD002、FD003、FD004 数据集上降低 3.63% 和 5.57%、1.40% 和 18.8%、0.48% 和 6.06%、0.38% 和 1.94%。因此多注意力网络可在通道维度和空间维度分别进行重要退化特征提取, 更好地在每个时间步长挖掘多维变量之间的空间相关性以及在所有时间步长上自适应挖掘与退化信号关联性更大的关键的特征, 进而更加精准地预测发动机剩余使用寿命。

3.4 现有方法的对比实验

为了说明 MA-VBLSTM 网络的有效性和优越性, 将本文方法与 CNN^[12]、LSTM^[14]、基于注意力的 LSTM 模型 (LSTM-A)^[15]、基于注意力的 BiLSTM 模型 (BiLSTM-A)^[16]、基于注意力的 CNN-GRU 模型 (AGCNN)^[18]、基于注意力和残差卷积的 BiLSTM 模型 (RCNN-ABiLSTM)^[20]、时间卷积网络 (TCN)^[21] 进行了对比。对比实验结果如表 4 所示, 本文提出的 MA-VBLSTM 在 FD001、FD002、FD003、FD004 子数据集上 RMSE 和 Score 值相比现有方法分别平均降低 5.27% 和 10.70%、1.37% 和 1.68%、6.37% 和 26.94%、3.02% 和 2.06%。因此本文提出的 MA-VBLSTM 方法的整体性能优于其他方法, 利用多注意力机制自适应学习发动机运行数据存在的通道特征和空

表 2 变分编码对比结果

方法	FD001		FD002		FD003		FD004	
	RMSE	Score	RMSE	Score	RMSE	Score	RMSE	Score
MA-BLSTM	12.3	220.0	18.3	1 876.76	13.4	330.0	20.1	2 933.32
MA-EBLSTM	12.4	227.5	16.9	1 380.94	12.6	275.9	19.3	2 719.30
MA-VBLSTM	11.6	198.9	16.8	1 247.61	12.2	236.0	18.2	2 461.52

表 3 双注意力并行对比实验

方法	FD001		FD002		FD003		FD004	
	RMSE	Score	RMSE	Score	RMSE	Score	RMSE	Score
SA-VBLSTM	12.23	225.97	17.00	1 598.87	12.72	264.63	18.51	2 864.36
CA-VBLSTM	12.09	210.64	17.09	1 536.67	12.34	251.30	18.34	2 510.22
MA-VBLSTM	11.65	198.90	16.85	1 247.61	12.28	236.07	18.27	2 461.52

表 4 现有方法对比实验结果

方法	FD001		FD002		FD003		FD004	
	RMSE	Score	RMSE	Score	RMSE	Score	RMSE	Score
CNN ^[12]	12.27	237.58	16.90	1 405.52	12.96	326.19	18.89	2 492.15
LSTM ^[14]	12.29	221.62	17.02	1 509.64	12.98	332.36	19.48	2 592.34
LSTM-A ^[15]	12.31	215.84	17.51	1 668.61	13.33	348.82	18.60	2 511.07
TCN ^[21]	12.45	229.54	17.10	1 547.17	13.72	294.46	18.73	2 525.44
BiLSTM-A ^[16]	12.64	245.11	16.93	1 465.58	13.41	307.71	18.83	2 488.28
AGCNN ^[18]	11.95	207.52	17.09	1 540.95	12.87	334.27	18.63	2 484.17
RCNN-ABiLSTM ^[20]	12.20	207.48	17.05	1 399.27	12.60	324.11	18.74	2 502.62
MA-VBLSTM	11.65	198.90	16.85	1 247.61	12.28	236.07	18.27	2 461.52

间特征，并对表现明显的退化特征数据和可区分故障时刻的关键信息进行动态加权，然后利用变分编码时序网络进行退化信息编码学习并强化全局上下时序特征的相关性，实现精准地反映健康趋势。

4 结束语

本文提出了一种融合多注意力和变分编码的时序网络进行设备剩余使用寿命预测。首先，使用通道注意力和空间注意力分别选取通道维度和空间维度的关键退化特征；然后，融合空间和通道特征学习的优势，通过自动编码器提取高维运行数据的深层低维特征，并利用 BiLSTM 网络双向捕捉长短期的时序数据特征；最后，通过全连接输出网络进行双维度特征的融合并实现 RUL 估计。将所提出的 MA-VBLSTM 模型应用于 C-MAPSS 数据集，并将实验结果与多种主流方法进行对比，实验结果验证了所提出的 MA-VBLSTM 算法能够获得更准确的 RUL 估计。

在未来的研究中，RUL 方法还可以考虑多种失效模式和复杂工况，以增强对更复杂问题的预测能力。此外，未来将进一步剖析机械系统的健康状态和退化状态的特征处理方式，这有助于获得复杂多变运行条件下运行数据的高效内部差异表示。

参考文献:

[1] 邱立军, 吴明辉. PHM 技术框架及其关键技术综述 [J]. 国

外电子测量技术, 2018, 37 (2): 10 - 15.

[2] 金晓航, 李建华, 孙毅. 基于二元维纳过程的轴承剩余寿命预测 [J]. 仪器仪表学报, 2018, 39 (6): 89 - 95.

[3] ZHU J, CHEN N, PENG W W. Estimation of bearing remaining useful life based on multiscale convolutional neural network [J]. IEEE Transactions on Industrial Electronics, 2018, 66 (4): 3208 - 3216.

[4] 任子强, 司小胜, 胡昌华, 等. 融合多传感器数据的发动机剩余寿命预测方法 [J]. 航空学报, 2019, 40 (12): 134 - 145.

[5] ZHANG Y B, JIA Y X, et al. Remaining useful life prediction model of planetary carrier in helicopter main gear-box based on Gamma degradation process [J]. Zhendong yu Chongji (Journal of Vibration and Shock), 2012, 31 (14): 47 - 51.

[6] KUNDU P, DARPEA K, KULKARNI M S. Weibull accelerated failure time regression model for remaining useful life prediction of bearing working under multiple operating conditions [J]. Mechanical systems and signal processing, 2019, 134: 106302.1 - 106302.19.

[7] 徐国平. 基于支持向量机的动调陀螺仪寿命预测方法研究 [D]. 上海: 上海交通大学, 2008.

[8] GEBRAEEL, NAGI, et al. Residual life predictions from vibration-based degradation signals: a neural network approach [J]. IEEE Transactions on industrial electronics, 2004, 51 (3): 694 - 700.

(下转第 56 页)