

# 基于卷积 GRU 注意力的设备剩余寿命预测\*

赵志宏<sup>1,2</sup>, 李 晴<sup>1</sup>, 李春秀<sup>1</sup>

(1. 石家庄铁道大学信息科学与技术学院 石家庄, 050043)

(2. 石家庄铁道大学省部共建交通工程结构力学行为与系统安全国家重点实验室 石家庄, 050043)

**摘要** 为了直接利用神经网络从采集的全寿命振动信号中自动提取特征信息,避免对人工提取特征的依赖,提出了一种基于卷积门控循环单元(convolutional gated recurrent unit,简称 ConvGRU)注意力的剩余寿命预测方法。首先,对于采集的设备振动信号预处理,输入 ConvGRU 注意力模型,ConvGRU 通过卷积神经网络(convolutional neural networks,简称 CNN)提取设备状态的空间局部特征,门控循环神经网络(gate recurrent unit,简称 GRU)提取时序特征信息,从而有效提取设备状态特征;其次,利用注意力机制对特征信息分配不同的权重;然后,进行中间网络层特征输出的可视化实验,验证了本研究方法特征提取的有效性;最后,进行了 2 个机械设备数据集 PHM2012 轴承数据集和 NASA 发动机数据集的实验,并与已有方法进行对比。实验结果表明,笔者提出的基于 ConvGRU 注意力的剩余寿命预测方法预测准确性更好,并具有较好的泛化性。

**关键词** 机械设备;剩余寿命预测;深度学习;卷积门控循环单元;注意力

**中图分类号** TH133; TH17

## 引 言

当今工业行业亟待解决的问题是面临繁重的机械设备健康维护管理工作时如何提高机械设备的安全性和稳定性,提高工作效率,降低维护成本<sup>[1]</sup>。故障预测与健康管理系统包括设备故障检测、实时监测和剩余寿命(remaining useful life,简称 RUL)预测等。RUL 是预测机械设备从当前时刻到无法正常工作的运行时间<sup>[2]</sup>。现有的剩余寿命预测方法主要分为基于失效机理的方法、基于数据驱动的方法和两者融合的方法<sup>[3]</sup>。基于失效机理的方法需要针对研究对象建立失效退化模型,一般不具有泛化性。由于工况复杂多变,机械设备愈加复杂,退化机理不同,获得失效模型过程复杂,预测效果难以保证<sup>[4]</sup>。基于数据驱动的方法是通过机器学习和统计学的方法从传感器采集的数据中探究与剩余寿命之间的关系<sup>[5]</sup>。传统的基于数据驱动的方法(例如支持向量机<sup>[6]</sup>、贝叶斯<sup>[7]</sup>、神经网络<sup>[8]</sup>等)在剩余寿命预测方面获得一定的成果,但随着机械设备的复杂化和集成化,采集的传感器数据愈加庞大,其中蕴含的特征关系难以获取,剩余寿命预测结果准确性存在一定的误差。

深度学习具有强大的非线性映射能力和提取特

征能力,被越来越多地应用在剩余寿命预测和健康监测领域<sup>[9]</sup>。Ren 等<sup>[10]</sup>应用 CNN 的轴承剩余寿命预测方法,通过 CNN 提取特征向量,利用深度神经网络预测剩余寿命。王奉涛等<sup>[11]</sup>应用基于长短期记忆网络(long short-term memory,简称 LSTM)预测模型,将提取到的退化特征输入到 LSTM 中预测滚动轴承的剩余寿命,并取得了较好的预测结果。肖丽<sup>[12]</sup>将 GRU 应用到滚动轴承的剩余寿命预测研究中,将 GRU 与粒子滤波器相结合预测轴承的 RUL,有效提高了预测的精度。王玉静等<sup>[13]</sup>将 CNN 和 LSTM 相结合,提出了一种新的网络结构预测滚动轴承的剩余寿命,提高了预测的准确度和稳定性。以上方法从振动信号中提取空间特征或时序特征,只是利用了一方面的特征信息,会造成特征信息的丢失。部分研究者开始从振动信号中同时提取空间特征和时序特征。王久健等<sup>[14]</sup>将卷积长短期记忆网络(convolutional long short-term memory,简称 ConvLSTM)应用于轴承 RUL 预测方法,通过 CNN 提取局部特征信息,LSTM 提取时间信息,获取退化程度的特征。ConvLSTM 存在的问题是参数量太大,为了进一步减少 ConvLSTM 的参数量并提高泛化能力,Shi 等<sup>[15]</sup>提出一种 CNN 与 GRU 结合的时空

\* 国家自然科学基金资助项目(11972236,11790282);石家庄铁道大学研究生创新基金资助项目(YC2021077)

收稿日期:2021-07-25;修回日期:2021-09-15

序列网络结构 ConvGRU。Chen 等<sup>[16]</sup>提出一种基于注意力机制的 RUL 预测模型,利用 LSTM 提取与时间有关的特征,通过注意力为不同传感器特征和时间步加权,进一步提高剩余寿命预测的准确性。ConvGRU 门控单元数目和参数量少于 ConvLSTM,提高了计算性能;注意力机制能对重要的特征信息分配更大权重。

基于以上考虑,笔者提出一种基于 ConvGRU 注意力的剩余寿命预测模型。ConvGRU 作为 CNN 和 GRU 结合体,既通过 CNN 提取设备状态的空间局部特征,又利用 GRU 提取时序特征,可更有效地提取设备状态特征;通过注意力机制对于特征信息分配不同的权重,有效利用特征信息。实验表明,本研究方法在剩余寿命预测方面取得了较好结果。

## 1 理论分析

### 1.1 GRU

LSTM 解决了 RNN 存在的长期依赖的问题,LSTM 通过 3 个门控结构(输入门、输出门和遗忘门)控制网络的输入值、输出值和记忆值。文献[17]提出了一种不同于 LSTM 的结构 GRU,对 LSTM 进行简化,它只包含 2 个门:更新门和重置门,其参数少于 LSTM,提高了网络计算效率。GRU 广泛应用于时间序列分类和预测问题<sup>[18]</sup>。GRU 网络单元结构如图 1 所示。

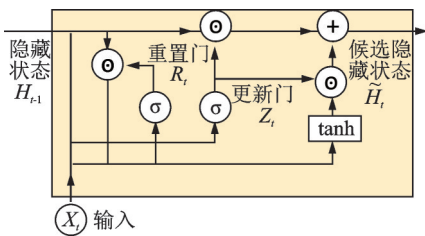


图1 GRU 网络单元结构

Fig.1 structure of GRU network

GRU 的计算公式如式(1)~(4)所示

$$R_t = \sigma(X_t W_{xr} + H_{t-1} W_{hr} + b_r) \quad (1)$$

$$Z_t = \sigma(X_t W_{xz} + H_{t-1} W_{hz} + b_z) \quad (2)$$

$$\tilde{H}_t = \tanh(X_t W_{xh} + (R_t \odot H_{t-1}) W_{hh} + b_h) \quad (3)$$

$$H_t = Z_t \odot H_{t-1} + (1 - Z_t) \odot \tilde{H}_t \quad (4)$$

其中: $\sigma$ 为激活函数; $X_t$ 为  $t$ 时刻网络层的输入; $H_{t-1}$ 为  $t-1$ 时刻隐藏状态; $R_t$ 为重置门,用于控制  $t-1$ 时刻的状态信息被传入  $t$ 时刻候选集  $\tilde{H}_t$  的数量, $t-1$ 时刻的状态信息传入  $t$ 时刻候选集  $\tilde{H}_t$  的数量越少,

则重置门越小; $Z_t$ 为更新门,用于控制  $t-1$ 时刻的状态信息更新至  $t$ 时刻的情况, $t-1$ 时刻的状态信息更新至  $t$ 时刻越少,更新门越小。

### 1.2 ConvGRU

Shi 等<sup>[15]</sup>提出的 ConvGRU 的门控循环机制与 GRU 相同,不同之处在于特征计算时 ConvGRU 采用卷积网络结构,而 GRU 采用全连接结构。ConvGRU 与 ConvLSTM 相比,结构更加简单,参数量更少,计算速度更快,网络收敛速度更快<sup>[15]</sup>。ConvGRU 同时具有 CNN 与 GRU 的优点,通过 CNN 提取空间局部特征,利用 GRU 提取时序特征,从而可以同时提取时间序列中的时空特征信息。ConvGRU 内部结构如图 2 所示。

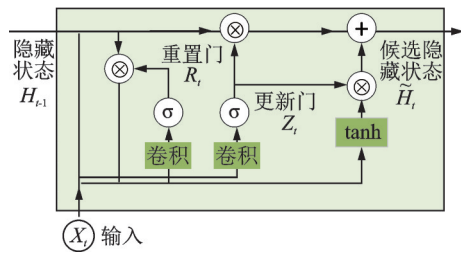


图2 ConvGRU 内部结构

Fig.2 Internal structure of ConvGRU

ConvGRU 的计算公式如式(5)~(8)所示

$$R_t = \sigma(X_t * W_{xr} + H_{t-1} * W_{hr} + b_r) \quad (5)$$

$$Z_t = \sigma(X_t * W_{xz} + H_{t-1} * W_{hz} + b_z) \quad (6)$$

$$\tilde{H}_t = \tanh(X_t * W_{xh} + (R_t \odot H_{t-1}) * W_{hh} + b_h) \quad (7)$$

$$H_t = Z_t \odot H_{t-1} + (1 - Z_t) \odot \tilde{H}_t \quad (8)$$

其中: $\sigma$ 为激活函数; $*$ 表示卷积操作; $\odot$ 表示元素相乘; $R_t$ 为重置门; $Z_t$ 为更新门; $X_t$ 为  $t$ 时刻网络层的输入; $H_{t-1}$ 为  $t-1$ 时刻隐藏状态; $\tilde{H}_t$ 为候选集。

### 1.3 注意力机制

注意力机制<sup>[19]</sup>最初应用于机器翻译领域,随着注意力机制的快速发展,已广泛应用于时序信息处理的各个方面。在深度学习模型中,通过注意力对提取到的特征信息赋予不同的权重,突出重要信息。笔者所使用的注意力分配权重过程如图 3 所示,计算过程分为 3 个阶段。

1) 通过点积运算计算输入特征向量  $Q, K_i$  的相

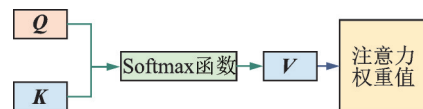


图3 注意力分配权重过程

Fig.3 Attention weight allocation process

关性,得到权重系数,计算公式为

$$P_{\text{similarity}}(\mathbf{Q}, \mathbf{K}_i) = \mathbf{Q} \cdot \mathbf{K}_i \quad (9)$$

2) 在Dense层利用Softmax函数对上一步得到的权重系数进行归一化处理,数值转换得到 $\mathbf{w}_i$

$$\mathbf{w}_i = \text{Softmax}(P_{\text{similarity}}(\mathbf{Q}, \mathbf{K}_i)) \quad (10)$$

3) 通过 $\mathbf{w}_i, \mathbf{V}_i$ 加权求和,得到对应的注意力权重值,计算公式为

$$A(\mathbf{Q}, \mathbf{K}_i) = \sum_{i=1}^n \mathbf{w}_i \cdot \mathbf{V}_i \quad (11)$$

其中: $\mathbf{Q}$ 为注意力层输入; $(\mathbf{K}_i, \mathbf{V}_i)$ 为一组数据对,通过计算与 $\mathbf{Q}$ 的相似性得到 $\mathbf{K}_i$ 对应的 $\mathbf{V}_i$ 。

## 2 寿命预测模型

笔者提出的剩余寿命预测模型ConvGRU注意力模型网络结构如图4所示,主要由ConvGRU层、批处理化层(BatchNormalization,简称BN)层、注意力层、平铺层和全连接层组成。寿命预测的具体步

骤如下:

1) 对于输入的原始数据序列首先经过数据转换,得到 $n$ 个特征向量 $\{x_1, x_2, \dots, x_n\}$ ,将其作为ConvGRU注意力网络结构的输入特征序列;

2) 通过2层ConvGRU隐藏层特征矩阵 $H_1, H_1$ 经过BN层进行正则化;

3) 将经过正则化后的特征矩阵 $H_2$ 输入到注意力机制进行加权,在注意力机制中,首先通过输入矩阵的点积运算得到对应的权重,将得到的权重经过Softmax函数对生成的权重归一化,最后将得到的权重和对应的值加权求和生成新的特征矩阵序列;

4) 由于预测目标与输入维数不同,将预测网络中的所有状态经过平铺层连接起来,并将其输入到2个全连接层以多对一的形式输出预测结果,即为当前剩余寿命与全寿命结果的比值。

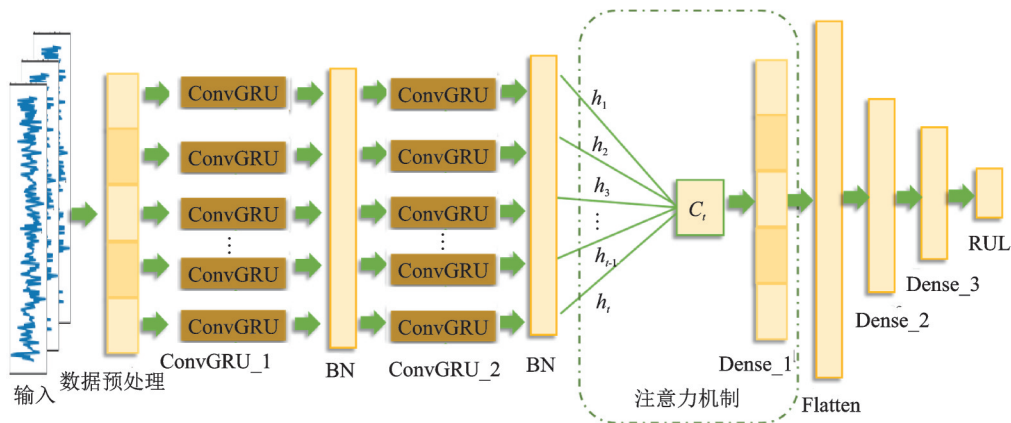


图4 ConvGRU注意力模型网络结构

Fig.4 Network structure of ConvGRU-attention model

## 3 实验与分析

### 3.1 数据集介绍

在公开的数据集<sup>[20]</sup>上进行实验验证笔者所提出的基于ConvGRU注意力的剩余寿命预测模型的性能。该数据集包含3个工况,17个轴承的全寿命数据。前2种工况各包括7个轴承,命名分别为轴承1-1~轴承1-7,轴承2-1~轴承2-7。第3种工况包括3个轴承,命名为轴承3-1~轴承3-3。该数据集每间隔10s采集一次水平和垂直2个方向的传感器数据,每次采集记录时间为0.1s,采样频率为25.6kHz。笔者采用水平方向采集的数据。实验选用11个轴承的数据作为训练集,剩余6个轴承数据

作为测试集。以轴承1~3为例,水平方向采集全寿命周期原始信号如图5所示。

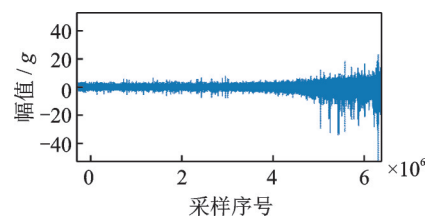


图5 轴承全寿命周期原始信号

Fig.5 Full life cycle original signal of bearing

### 3.2 数据预处理

将轴承方向的原始数据整合处理,对于整合的轴承数据在输入模型之前采用最大最小标准化进行

数据预处理,缩小数据范围,将数据集最小值和最大值转化至 $[0, 1]$ 之间,以提高模型计算速度,计算公式为

$$\overline{X}_t = \frac{X_t - X_{\min}}{X_{\max} - X_{\min}} \quad (12)$$

其中: $X_t$ 为数据集 $t$ 时刻的输入; $X_{\min}$ 为数据集的最小值; $X_{\max}$ 为数据集的最大值; $\overline{X}_t$ 为 $t$ 时刻通过数据标准化后的值。

对于每行数据进行标签化处理,每行数据的标签为轴承对应的剩余寿命和全寿命的比值 $Z_i$ , $Z_i$ 计算公式为

$$Z_i = \frac{L - i}{L - 1} \quad (13)$$

其中: $L$ 为每个轴承整合后的行数,即为轴承的实际全寿命长度; $i$ 为当前行。

对剩余寿命标签进行归一化处理,可以降低轴承不同工况和剩余寿命不同等因素的影响,提高剩余寿命预测的准确性。

### 3.3 剩余寿命计算步骤

利用一次线性回归预测,得到 $t$ 时刻对应的剩余寿命 $RUL_t$ ,其计算步骤如下:

1) 通过线性回归建立预测,得到当前剩余寿命与全寿命结果之间的线性方程为

$$t = ay_t + b \quad (14)$$

2) 计算轴承的全寿命周期,当 $y_t = 0$ 时,轴承剩余寿命为0,其全寿命周期为

$$m = ay_t + b = b \quad (15)$$

3) 计算轴承任意时刻 $t$ 的剩余寿命为

$$RUL_t = L - t \quad (16)$$

其中: $t$ 为轴承已运行时间; $y_t$ 为预测的当前剩余寿命与全寿命结果的比值; $L$ 为轴承的全寿命周期; $RUL_t$ 为 $t$ 时刻轴承的剩余寿命。

### 3.4 实验分析

#### 3.4.1 轴承RUL预测结果对比

为验证笔者提出的基于ConvGRU注意力的剩余寿命预测模型的有效性,利用PHM2012数据集的误差百分比进行评价。误差百分比的计算公式为

$$Er_i = \frac{RUL_{acti} - RUL_i}{RUL_{acti}} \times 100\% \quad (17)$$

其中: $RUL_{acti}$ 为第 $i$ 组测试数据的实际RUL; $RUL_i$ 为第 $i$ 组测试数据的预测RUL。

不同预测方法对比如表1所示。可以看出,以轴承1~3为例,本研究方法与ConvGRU相比误差降低了79.87%;与ConvLSTM相比误差降低了45.61%;与基于自组织映射(self-organizing map,简称SOM)的健康指数(health index,简称HI)<sup>[21]</sup>相比误差降低了43.32%;与CNN-HI相比误差降低了62.21%;与RNN-HI<sup>[22]</sup>相比误差降低了57.67%。因此,笔者基于ConvGRU注意力的剩余寿命预测模型具有更低的误差率,表明了该模型的有效性。

表 1 不同预测方法对比

Tab.1 Comparison of different prediction methods

测试集	当前时刻	实际寿命	预测寿命	ConvGRU注意力	ConvGRU	ConvLSTM <sup>[14]</sup>	CNN-HI <sup>[14]</sup>	SOM-HI <sup>[21]</sup>	RNN-HI <sup>[22]</sup>
轴承 1-3	1 801	573	678	-18.32	-89.53	33.68	48.52	-31.76	43.28
轴承 1-4	1 138	290	216	25.52	-44.21	47.24	53.57	62.76	67.55
轴承 1-6	2 301	146	175	-19.86	-29.17	23.28	19.39	-32.88	21.23
轴承 1-7	1 501	757	649	14.27	34.87	-3.30	16.27	-11.09	17.83
轴承 2-5	2 001	309	421	-36.25	-60.84	-39.80	56.13	68.61	54.37
轴承 2-6	571	129	139	-8.33	34.88	8.52	-18.65	-51.94	-13.95

#### 3.4.2 与ConvLSTM注意力比较

ConvGRU注意力和ConvLSTM注意力模型的参数量和运行时间对比如表2所示。其中,2个模型的层数相同且训练次数均为100。运行的硬件设备参数:内存为64G;处理器为Intel Xeon(R) Silver 4110 CPU @ 2.10GHz \*16;操作系统为64位。从表2可以看出,ConvGRU注意力模型与ConvL-

表 2 参数量和运行时间对比

Tab.2 Comparison of parameters and running time

模型	参数量	运行时间/s
ConvLSTM注意力	25 539	22 500
ConvGRU注意力	19 789	19 000

STM注意力模型相比,参数量降低了22.51%,运行时间减少了15.56%。



以轴承 1~3 为例, ConvLSTM 注意力和 ConvGRU 注意力的预测结果如图 6 所示。可以看出, ConvGRU 注意力方法的实际 RUL 和预测 RUL 相比,波动小于 ConvLSTM 注意力方法。

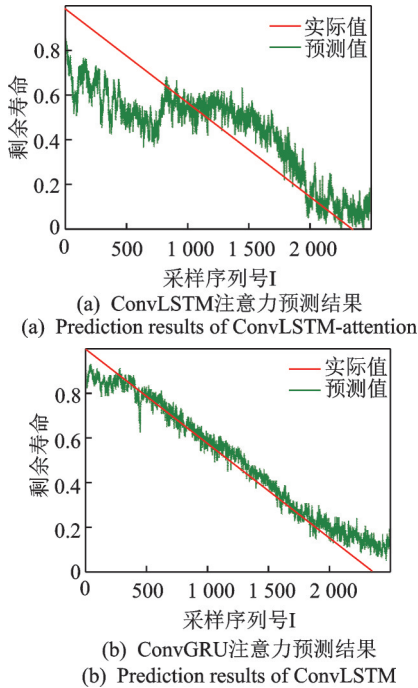


图 6 轴承实际 RUL 和预测 RUL 对比

Fig.6 Comparison of actual RUL and predicted RUL of bearing

### 3.4.3 可视化分析

为进一步验证笔者所提的基于 ConvGRU 注意力剩余寿命预测模型的有效性,对该模型不同层所提取的特征可视化。选取可视化的层分别为输入层、ConvGRU\_2 层输出,注意力层输出和 Dense\_2 层输出。隐藏层特征可视化如图 7 所示,图中的每个点表示不同时刻  $t$  的特征,颜色则对应点的剩余寿命和全寿命的比值  $Z_i$ 。根据图 7

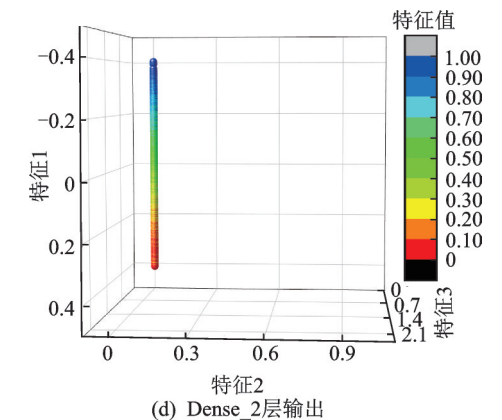
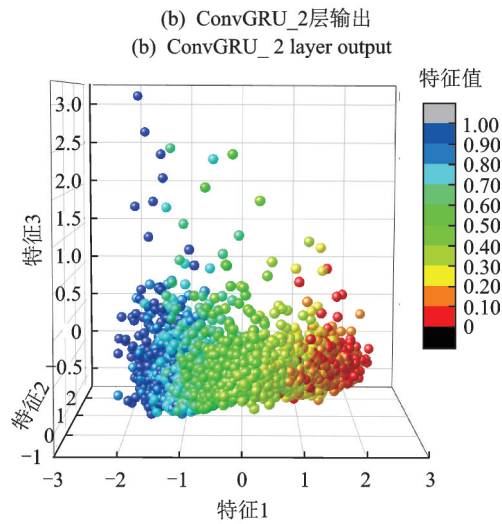
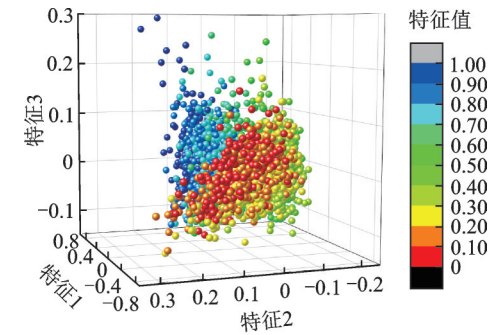
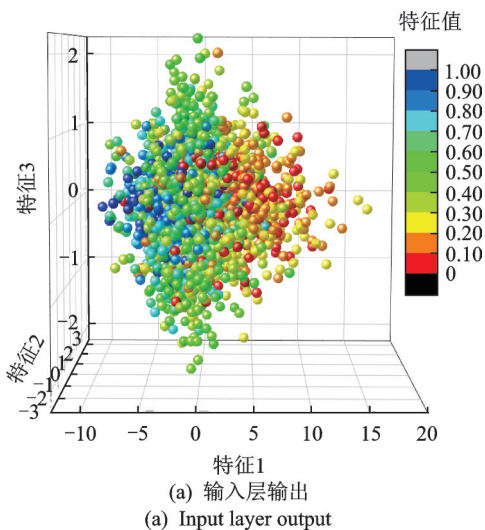


图 7 隐藏层特征可视化

Fig.7 Visualization of hidden layer features

可知,随着模型的深入,所提取的特征由输入层的杂乱无序,经过 ConvGRU\_2 层后的逐渐分离,至注意力层输出的逐渐有序,到 Dense\_2 层的有序排列,反映了该模型较好的提取特征能力。

### 3.5 C-MAPSS 数据集实验

为验证笔者提出模型的泛化能力,使用 NASA 仿真 C-MAPSS 数据集。该数据集包括 4 个子集,具体情况如表 3 所示。每个子集包括对应的训练集、测试集以及测试集对应的实际剩

余寿命。该数据集使用的评价指标为:评分函数(PHM 2008数据挑战评价标准中评分函数)和均方根误差(root mean square error,简称RMSE)。

表3 C-MAPSS数据集

Tab.3 C-MAPSS data set

参数设置	数据集			
	FD001	FD002	FD003	FD004
训练发动机单元个数	100	260	100	249
测试发动机单元个数	100	259	100	248
操作条件	1	6	1	6
故障类型	1	1	2	2

评分函数计算公式为

$$S = \begin{cases} \sum_{i=1}^n e^{-\left(\frac{d}{13}\right)} - 1 & (d < 0) \\ \sum_{i=1}^n e^{\left(\frac{d}{10}\right)} - 1 & (d \geq 0) \end{cases} \quad (18)$$

其中: $n$ 为每个测试集中的发动机个数; $d = RUL_{estimated} - RUL_{real}$ 。

RMSE具有无偏性,用于衡量实际值和预测值之间的偏差,计算公式为

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (19)$$

其中: $y_i$ 为实际值; $\hat{y}_i$ 为预测值。

FD001~FD004子集测试集中剩余寿命实际值与预测值对比如图8所示,按照从小到大对发动机索引排序。由图8可知,FD001与FD003对应的剩余寿命预测结果与实际寿命相比,波动幅度小于FD002与FD004,说明FD001与FD003剩余寿命预测结果更准确。

如表4所示,为验证笔者所提的基于ConvGRU注意力的剩余寿命预测模型的有效性,与已有的使用C-MAPSS数据集RUL预测方法进行对比。方法包括双向长短期记忆网络(bi-directional long-short term memory,简称Bi-LSTM)<sup>[23]</sup>、长短期记忆网络(long-short term memory,简称LSTM)<sup>[24]</sup>、梯度提升(gradient boosting,简称GB)<sup>[25]</sup>、随机森林(random forest,简称RF)<sup>[25]</sup>、支持向量机(support vector machine,简称SVM)<sup>[25]</sup>、基于时间窗神经网络(time window based neural network,简称TWBNN)<sup>[26]</sup>、深度卷积神经网络(deep convolutional neural network,简称DCNN)<sup>[27]</sup>。从表4可以看出,笔者基于ConvGRU

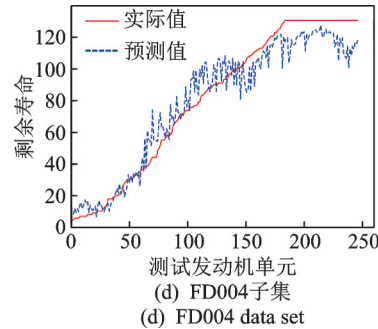
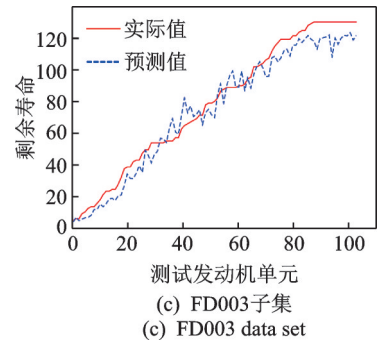
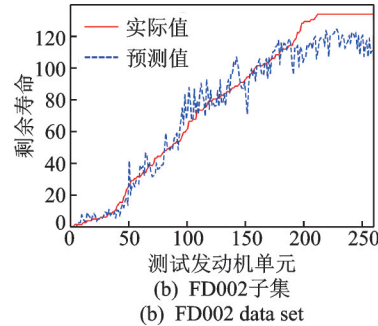
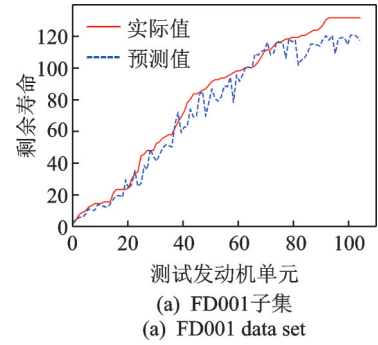


图8 FD001~FD004子集测试集中剩余寿命实际值与预测值对比

Fig.8 FD001~FD004 subset test set remaining useful life actual value and predicted value comparison

注意力的剩余寿命预测模型与已有的方法相比,评分函数和RMSE都取得了较好的结果。其中:和Bi-LSTM比较,子集FD001的评分函数降低了3.25%,RMSE减少了6.26%;子集FD002的评分函数降低了50.65%,RMSE减少了23.48%;子集FD003的评分函数降低了47.54%,RMSE减少了1.84%;子集FD004的评分函数降低了15.26%,RMSE减少了25.10%。

表 4 基于 C-MAPSS 数据集的不同方法与本研究方法对比

Tab.4 Comparison of different methods based on C-MAPSS dataset and the method in this paper

数据集	评价指标	ConvGRU 注意力	Bi-LSTM <sup>[23]</sup>	LSTM <sup>[24]</sup>	GB <sup>[25]</sup>	RF <sup>[25]</sup>	SVM <sup>[25]</sup>	TWBNN <sup>[26]</sup>	DCNN <sup>[27]</sup>
FD001	评分函数	327	338	—	474.01	479.75	7 703.33	—	1 286.7
	RMSE	15.13	16.14	12.81	15.67	17.91	40.72	—	18.45
FD002	评分函数	2 196	4 450	—	87 280.06	70 456.86	316 483.31	—	13 570
	RMSE	18.74	24.49	—	29.09	29.59	52.99	—	30.29
FD003	评分函数	447	852	—	576.72	711.13	22 541.58	—	1 596.2
	RMSE	15.88	16.18	—	16.84	20.27	46.32	—	19.82
FD004	评分函数	4 703	5 550	—	17 817.92	46 567.63	141 122.19	—	7 886.4
	RMSE	21.10	28.17	—	29.01	31.12	59.96	—	29.16

## 4 结 论

1) 针对 ConvLSTM 参数量大的问题,笔者采用 ConvGRU 网络结构,既利用 CNN 提取空间局部特征,又通过 GRU 提取时序特征,直接从采集的全寿命振动信号中提取时空特征信息,从而全面有效地获取具有一定趋势的退化特征信息。

2) 针对不同特征信息对于剩余寿命的影响不同,笔者引入注意力机制,对提取的特征信息分配不同的权重,突出重要信息。通过特征可视化实验,说明了注意力机制的有效性,提高了 RUL 预测的准确性。

3) 笔者所提模型在 2 个不同数据集上进行实验验证,并与其他已有的方法进行对比。结果表明,基于 ConvGRU 注意力的剩余寿命预测模型在设备剩余寿命预测方面更加有效,且具有较好的泛化性。

## 参 考 文 献

[1] MOGHADDASS R, ZUO M J. An integrated framework for online diagnostic and prognostic health monitoring using a multistate deterioration process[J]. Reliability Engineering and System Safety, 2014, 124: 92-104.

[2] LIU G. A study on remaining useful life prediction for prognostic applications [D]. New Orleans: University of New Orleans, 2011.

[3] LIAO L, KÖTTIG F. Review of hybrid prognostics approaches for remaining useful life prediction of engineered systems, and an application to battery life prediction[J]. IEEE Transactions on Reliability, 2014, 63(1): 191-207.

[4] 陈杰,邓二平,赵子轩,等.不同老化试验方法下 SiC MOSFET 失效机理分析[J].电工技术学报,2020,35(24): 5105-5114.

CHEN Jie, DENG Erping, ZHAO Zixuan, et al. Failure

mechanism analysis of SiC MOSFET under different aging test methods[J]. Transactions of China Electrotechnical Society, 2020, 35(24):5105-5114. (in Chinese)

[5] 裴洪,胡昌华,司小胜,等.基于机器学习的设备剩余寿命预测方法综述[J].机械工程学报,2019,10(5):1-13.

PEI Hong, HU Changhua, SI Xiaosheng, et al. Review of machine learning based remaining useful life prediction methods for equipment[J]. Journal of Mechanical Engineering, 2019, 10(5): 1-13. (in Chinese)

[6] 胡昌华,胡锦涛,张伟,等.支持向量机用于性能退化的可靠性评估[J].系统工程与电子技术,2009,31(5): 1246-1249.

HU Changhua, HU Jintao, ZHANG Wei, et al. Reliability assessment of performance degradation using support vector machines[J]. Systems Engineering and Electronics, 2009, 31(5): 1246-1249. (in Chinese)

[7] 杜坤.基于贝叶斯算法的齿轮箱故障诊断和剩余寿命预测[D].长沙:湖南工业大学,2018.

[8] 王新春,赵金金.基于小波神经网络的 IGBT 寿命研究[J].电工技术,2020(10): 114-116.

WANG Xinchun, ZHAO Jinjin. Research on IGBT life based on wavelet neural network[J]. Electric Engineering, 2020(10): 114-116. (in Chinese)

[9] ZHANG W, PENG G L, LI C H, et al. A new deep learning model for fault diagnosis with good anti-noise and domain adaptation ability on raw vibration signals [J]. Sensors, 2017, 17(2): 425-425.

[10] REN L, SUN Y, WANG H, et al. Prediction of bearing remaining useful life with deep convolution neural network [J]. IEEE Access, 2018, 6: 13041-13049.

[11] 王奉涛,刘晓飞,邓刚,等.基于长短期记忆网络的滚动轴承寿命预测方法[J].振动、测试与诊断,2020,40(2): 303-309.

WANG Fengtao, LIU Xiaofei, DENG Gang, et al. Remaining useful life prediction method for rolling bear-

- ing based on the long short-term memory network[J]. Journal of Vibration, Measurement & Diagnosis, 2020, 40(2): 303-309. (in Chinese)
- [12] 肖丽. 基于GRU网络的滚动轴承剩余寿命预测研究[D]. 武汉:武汉科技大学, 2020.
- [13] 王玉静, 李少鹏, 康守强, 等. 结合CNN和LSTM的滚动轴承剩余使用寿命预测方法[J]. 振动、测试与诊断, 2021, 41(3): 439-446.  
WANG Yujing, LI Shaopeng, KANG Shouqiang, et al. Method of predicting remaining useful life of rolling bearing combining CNN and LSTM[J]. Journal of Vibration, Measurement & Diagnosis, 2021, 41(3): 439-446. (in Chinese)
- [14] 王久健, 杨绍普, 刘永强, 等. 一种基于空间卷积长短期记忆神经网络的轴承剩余寿命预测方法[J]. 机械工程学报, 2021, 57(21): 88-95.  
WANG Jiujian, YANG Shaopu, LIU Yongqian, et al. A method of bearing remaining useful life estimation based on convolutional long short-term memory neural network[J]. Journal of Mechanical Engineering, 2021, 57(21): 88-95. (in Chinese)
- [15] SHI X, GAO Z, LAU S L, et al. Deep learning for precipitation nowcasting: a benchmark and a new model [C]//Advances in Neural Information Processing Systems. Long Beach, CA: [s. n.], 2017: 5617-5627.
- [16] CHEN Z, WU M, ZHAO R, et al. Machine remaining useful life prediction via an attention-based deep learning approach[J]. IEEE Transactions on Industrial Electronics, 2020, 68(3): 2521-2531.
- [17] LITJENS G, KOOI T, BEJNORDI B E, et al. A survey on deep learning in medical image analysis[J]. Medical Image Analysis, 2017, 42: 60-88.
- [18] 陈海文, 王守相, 王绍敏, 等. 基于门控循环单元网络与模型融合的负荷聚合体预测方法[J]. 电力系统自动化, 2019, 43(1): 65-74.  
CHEN Haiwen, WANG Shouxiang, WANG Shaomin, et al. Aggregated load forecasting method based on gated recurrent unit networks and model fusion[J]. Automation of Electric Power Systems, 2019, 43(1): 65-74. (in Chinese)
- [19] 汪琪, 段湘煜. 基于注意力卷积的神经机器翻译[J]. 计算机科学, 2018, 45(11): 226-230.  
WANG Qi, DUAN Xiangyu. Neural machine translation based on attention convolution[J]. Computer Science, 2018, 45(11): 226-230. (in Chinese)
- [20] NECTOUX P, GOURIVEAU R, MEDJAHHER K, et al. Pronostia: an experimental platform for bearings accelerated degradation tests [C]//IEEE International Conference on Prognostics and Health Management. Denver, Colorado, USA: [s. n.], 2012: 1-8.
- [21] SHENG H, ZHENG Z, ENRICO Z, et al. Condition assessment for the performance degradation of bearing based on a combinatorial feature extraction method[J]. Digital Signal Processing, 2014, 27: 159-166.
- [22] GUO L, LI N P, JIA F, et al. A recurrent neural network based health indicator for remaining useful life prediction of bearings [J]. Neurocomputing, 2017, 240: 98-109.
- [23] SHUAI Z, RISTOVSKI K, FARAHAT A, et al. Long short-term memory network for remaining useful life estimation [C]//IEEE International Conference on Prognostics and Health Management. Dallas, TX, USA: [s. n.], 2017: 88-95.
- [24] HSU C S, JIANG J R. Remaining useful life estimation using long short-term memory deep learning [C]//IEEE International Conference on Applied System Innovation. Chiba, Japan: [s. n.], 2018: 58-61.
- [25] ZHANG C, LIM P, QIN A K, et al. Multiobjective deep belief networks ensemble for remaining useful life estimation in prognostics [J]. IEEE Transactions on Neural Networks and Learning Systems, 2016, 28(10): 2306-2318.
- [26] LIM P, CHI K G, TAN K C. A time window neural network based framework for remaining useful life estimation [C]//IEEE International Joint Conference on Neural Networks. Vancouver, BC, Canada: [s. n.], 2016: 1746-1753.
- [27] BABU G S, ZHAO P, LI X L, et al. Deep convolutional neural network based regression approach for estimation of remaining useful life [C]//International Conference on Database Systems for Advanced Applications. [S. l.]: Springer, 2016: 214-228.



**第一作者简介:**赵志宏,男,1972年8月生,博士、教授、博士生导师。主要研究方向为机械故障诊断,机器学习,信号处理及动力学分析。曾发表《多小波系数特征提取方法在故障诊断中的应用》(《振动、测试与诊断》2015年第35卷第2期)等论文。

E-mail:hb\_zhaozhong@126.com