

电子科技大学  
UNIVERSITY OF ELECTRONIC SCIENCE AND TECHNOLOGY OF CHINA

# 博士学位论文

DOCTORAL DISSERTATION



论文题目 基于监测大数据的产品剩余寿命预测

方法研究

学科专业 机械工程

学 号 201511080101

作者姓名 黄承赓

指导教师 黄洪钟 教 授

分类号 \_\_\_\_\_ 密级 \_\_\_\_\_

UDC <sup>注1</sup> \_\_\_\_\_

# 学 位 论 文

## 基于监测大数据的产品剩余寿命预测方法研究

(题名和副题名)

黄承赓

(作者姓名)

|      |        |     |
|------|--------|-----|
| 指导教师 | 黄洪钟    | 教 授 |
|      | 电子科技大学 | 成 都 |
|      |        |     |
|      |        |     |

(姓名、职称、单位名称)

申请学位级别 博士 学科专业 机械工程

提交论文日期 2019 年 10 月 论文答辩日期 2019 年 12 月

学位授予单位和日期 电子科技大学 2019 年 12 月

答辩委员会主席 \_\_\_\_\_

评阅人 \_\_\_\_\_

注 1: 注明《国际十进分类法 UDC》的类号。

**Research on methods for remaining useful life  
prediction of products driven by  
condition monitoring big data**

**A Doctoral Dissertation Submitted to  
University of Electronic Science and Technology of China**

Discipline: **Mechanical Engineering**

Author: **Cheng-Geng Huang**

Supervisor: **Prof. Hong-Zhong Huang**

School: **School of Mechanical and Electrical Engineering**

## 独创性声明

本人声明所呈交的学位论文是本人在导师指导下进行的研究工作及取得的研究成果。据我所知，除了文中特别加以标注和致谢的地方外，论文中不包含其他人已经发表或撰写过的研究成果，也不包含为获得电子科技大学或其它教育机构的学位或证书而使用过的材料。与我一同工作的同志对本研究所做的任何贡献均已在论文中作了明确的说明并表示谢意。

作者签名：黄承康

日期：2020年1月6日

## 论文使用授权

本学位论文作者完全了解电子科技大学有关保留、使用学位论文的规定，有权保留并向国家有关部门或机构送交论文的复印件和磁盘，允许论文被查阅和借阅。本人授权电子科技大学可以将学位论文的全部或部分内容编入有关数据库进行检索，可以采用影印、缩印或扫描等复制手段保存、汇编学位论文。

（保密的学位论文在解密后应遵守此规定）

作者签名：黄承康

导师签名：黄胜军

日期：2020年1月6日

## 摘 要

随着传感器技术、信息与通信技术及物联网等技术的快速发展,状态监测技术已被广泛应用于各类产品。不同类型的传感器被广泛地部署于产品的各个部位并以极高的采样频率对其退化状态进行实时在线监测,所获得监测数据呈现井喷式的增长,这不可避免地将我们带入了工业大数据时代。在工业大数据背景下,有效地利用海量的监测数据能实现对产品剩余寿命更为准确的估计,有助于产品实现提前报警,从而预防灾难性事故的发生,进一步优化产品的维护策略,最终降低昂贵且非必需的维护费用,同时还能提升产品的可靠性、可用性和安全性。因此,本文将重点研究基于监测大数据的产品剩余寿命预测方法。针对产品群规模大、传感器多、数据采样频率高、动态工况和多失效模式耦合、复杂产品失效物理模型不可获知等诸多挑战,本文研究了基于退化轨迹相似性分析的剩余寿命预测方法,构建了产品动态工况和多失效模式耦合作用下基于深度学习模型的剩余寿命预测方法,再融合粒子滤波技术和深度学习模型提出了混合剩余寿命预测方法,最后在Bootstrap 实施框架下提出了更泛化的基于深度卷积神经网络的剩余寿命预测方法。

本文主要研究内容及创新成果如下:

(1) 提出一种改进的基于退化轨迹相似性的产品剩余寿命预测方法。当存在大量相似产品历史监测数据时,传统的基于退化轨迹相似性的方法能精确地预测该类工程背景下产品的剩余寿命,但仅能给出剩余寿命点估计的结果。因此,需对传统的基于退化轨迹相似性的方法进行更加深入的研究以突破仅能对产品剩余寿命进行点估计的应用限制。基于轨迹相似性分析方法和核密度估计技术,同时结合 $\beta$ 准则,提出了改进的基于退化轨迹相似性的产品剩余寿命预测方法。

(2) 提出一种基于深度学习模型的产品剩余寿命预测方法。现代工程系统常常工作在动态工况下且存在多失效模式耦合的情形,传统的基于数据驱动的方法尚不能处理该类情况下复杂产品的剩余寿命预测问题。本文利用深度学习模型中双向长短期记忆神经网络对时间序列数据的建模优势与极强的非线性建模分析能力,构建一种基于深度学习模型的方法,从而能够为复杂产品剩余寿命预测提供一种端对端的解决方案,且能给出比其他方法更为精确的估计结果。

(3) 提出融合粒子滤波技术和深度学习模型的混合剩余寿命预测方法。当某些特定产品的关键失效模式和失效模型已知时,传统的混合方法能够依据产品的失效物理模型并利用产品的状态监测数据对产品剩余寿命分布进行预测。但传统的混合方法较为繁琐,包含特征提取、选择、降维和退化指标回归分析等步骤。上

述步骤的实施不仅需要充分利用领域内的专业知识，并且最终的预测性能较大地依赖于各个步骤的精确性。因此，传统的混合剩余寿命预测方法效率低且稳健性差。本文在粒子滤波的技术框架下，有效地利用双向长短期记忆神经网络和前馈神经网络的复杂结构，提出融合粒子滤波技术和深度学习模型的混合剩余寿命预测方法。

（4）提出基于深度卷积神经网络和 **Bootstrap** 的剩余寿命预测方法。在工程实践中，当某些复杂产品关键失效模式未知且失效模型难以构建时，基于物理模型和深度学习模型的混合方法不能有效地对其剩余寿命预测区间进行量化。因此，本文在 **Bootstrap** 实施框架内提出一种更加泛化的基于深度卷积神经网络的剩余寿命预测方法，该方法能突破混合方法需要建立产品退化模型的限制，进一步拓宽基于深度学习模型的方法在复杂产品剩余寿命预测中的应用范围。

**关键词：** 剩余寿命预测，状态监测，深度学习模型，粒子滤波，**Bootstrap** 方法

## ABSTRACT

The rapid advancements of sensor, information and communication technology, and Internet-of-Thing (IoT) technique, etc, the condition monitoring (CM) technique has been adopted on a large scale in various products. In order to implement real-time online monitoring toward the health condition of the products, different types of sensors have been deployed on some specific parts of the products to acquire data by high sampling frequency. Consequently, the growing amount of sensors result in the continuous collection of high volume data, which inevitably creates an industrial big data environment. Under such a circumstance, the available massive CM data can be leveraged to continuously track the product's health status and make an accurate and reliable estimation about its remaining useful life (RUL). With these actionable information in hand, catastrophic failures can be avoided by the means of advance warning, maintenance schedule can be further optimized to reduce the cost of unnecessary and expensive maintenance actions, and meanwhile, the reliability, availability, and safety of the products can also be improved. Thus, the research of this dissertation will focus on prognostic approaches for RUL prediction of products. In the era of industrial big data, the amount of products group monitored and the number of the sensors for each product are large, massive CM data are acquired by the high sampling frequency, products are working under dynamic operating conditions and multiple failure modes coexist, and the failure physical models for some complex products are unavailable. Aiming at aforementioned practical challenges, this dissertation firstly conducts the research on RUL prediction method based on degradation trajectory similarity based prognostic method, then develops the deep learning (DL)-based prognostic framework to tackle the issue of RUL prediction of products under dynamic operating conditions and multiple failure modes coexisting. In addition, particle filter (PF) technique and DL-based model are strategically integrated to develop the hybrid prognostic approach. Finally, in the bootstrap implementation framework, the deep convolutional neural network (DCNN)-based prognostic approach is proposed to quantify prediction interval of the RUL without utilization of the physical model of the products.

The main contents and innovations of this dissertation are summarized as follows:

- (1) Development of an improved degradation trajectory similarity-based prognostic

(TSBP) method toward RUL prediction of the products. Considering the scenario where there are plenty of similar products with massive historical CM data, traditional TSBP method can accurately provide the point estimation of the RUL about the operating products. However, this might impose restrictions on the application of TSBP method to some safety-critical products, such as aircraft engine, etc. Thus, it is necessary to conduct in-depth study about the TSBP method to break through the limitation of point estimation of the RUL. This dissertation proposes an improved TSBP method, which integrates the kernel density estimation (KDE) technique and  $\beta$ -criterion.

(2) Development of a deep learning (DL)-based prognostic method. Modern engineering systems generally work under dynamic operating conditions and multiple failure modes coexist. Most of the traditional data-driven prognostic methods still lack an effective model to handle this prognostic issue. Due to the fact that bi-directional long short-term memory (BLSTM, one kind of DL-based model) networks can effectively extract long-term temporal dependences hidden in the time series data and has powerful nonlinear modelling capacity. Thus, this dissertation proposes a BLSTM-based approach that can provide an end-to-end prognostic solution. And the proposed DL-based method also can obtain more accurate RUL prediction results compared with other state-of-the-art data-driven prognostic methods.

(3) Development of a hybrid prognostic method that integrates particle filter (PF) technique with DL-based model. For some specific products that the critical failure mode is known and the failure physical model has been established, traditional hybrid prognostic method is based on the physical model of the products and leverage the CM data to estimate the probability density function (PDF) of the RUL. However, a limitation associated with traditional hybrid approaches is that numerous cumbersome steps including but not limited to feature extration, selection, reduction and regression analysis about the degradation indicator of the products. And the aforementioned processing steps take advantage of in-depth domain knowledge. In addition, the prognosis performance greatly relies on these specific steps, which might make the traditional hybrid approaches inefficient and not that robust. Thus, this dissertation proposes an enhanced DL-based hybrid prognostic method in the PF framework, which fully utilizes BLSTM networks and feed-forward neural networks (FFNN) to automatically extract and select features and implement regression analysis to obtain the degradation indicator of the products.

(4) Development of a novel DCNN-Bootstrap prognostic method. For some

complex products or engineering systems, the corresponding failure modes and failure physical models are unavailable, thus the aforementioned PF-DL fusion prognostic method cannot effectively quantify the RUL prediction interval. Accordingly, this dissertation proposes a generalized prognostic approach based on DCNN and is implemented in the Bootstrap framework. The proposed DCNN-Bootstrap prognostic method do not require any prior physical information about the products, such as its failure physical model and the corresponding parameters statistical distribution, and this novel characteristic will enhance the applicability of the DL-based prognostic method toward the complex products and engineering systems.

**Keywords:** remaining useful life prediction, condition monitoring, deep learning-based model, particle filter, bootstrap

# 目 录

|                                      |    |
|--------------------------------------|----|
| 第一章 绪 论 .....                        | 1  |
| 1.1 研究背景及意义 .....                    | 1  |
| 1.2 产品剩余寿命预测方法研究综述 .....             | 2  |
| 1.2.1 基于物理模型的剩余寿命预测方法研究 .....        | 3  |
| 1.2.2 基于数据驱动的剩余寿命预测方法研究 .....        | 5  |
| 1.2.3 基于物理模型和数据驱动的混合剩余寿命预测方法研究 ..... | 11 |
| 1.3 产品剩余寿命预测方法研究的发展趋势分析 .....        | 13 |
| 1.4 论文结构框架和主要研究内容 .....              | 14 |
| 第二章 基于退化轨迹相似性的剩余寿命预测改进方法 .....       | 17 |
| 2.1 引言 .....                         | 17 |
| 2.2 传统的基于相似性的剩余寿命预测方法 .....          | 17 |
| 2.2.1 方法假设与实施流程 .....                | 17 |
| 2.2.2 数据前处理 .....                    | 18 |
| 2.2.3 退化轨迹相似性分析与剩余寿命预测 .....         | 21 |
| 2.3 基于核密度估计的相似性剩余寿命预测改进方法 .....      | 24 |
| 2.3.1 改进 TSBP 方法的实施框架与数据预处理 .....    | 25 |
| 2.3.2 基于核密度估计算法的剩余寿命概率密度函数近似 .....   | 28 |
| 2.3.3 基于改进 TSBP 方法的剩余寿命预测 .....      | 30 |
| 2.4 算例分析与方法验证 .....                  | 32 |
| 2.4.1 剩余寿命预测方法评价性能指标 .....           | 32 |
| 2.4.2 航空发动机退化仿真数据集 .....             | 33 |
| 2.4.3 结果分析与讨论 .....                  | 36 |
| 2.5 本章小结 .....                       | 39 |
| 第三章 基于双向长短期记忆神经网络的产品剩余寿命预测方法 .....   | 40 |
| 3.1 引言 .....                         | 40 |
| 3.2 问题描述与数据前处理 .....                 | 41 |
| 3.2.1 问题描述 .....                     | 41 |
| 3.2.2 数据前处理 .....                    | 42 |
| 3.3 基于双向长短期记忆神经网络的剩余寿命预测方法 .....     | 43 |
| 3.3.1 循环神经网络简介 .....                 | 44 |

|                                                 |           |
|-------------------------------------------------|-----------|
| 3.3.2 基于 BLSTM 网络的产品多工况下剩余寿命预测模型.....           | 50        |
| 3.3.3 模型训练与正则化方法.....                           | 54        |
| 3.4 算例分析与方法验证.....                              | 56        |
| 3.4.1 多工况下航空发动机退化仿真数据集简介及数据前处理.....             | 56        |
| 3.4.2 结果分析与讨论.....                              | 59        |
| 3.5 本章小结.....                                   | 66        |
| <b>第四章 基于双向长短期记忆神经网络的混合剩余寿命预测方法.....</b>        | <b>67</b> |
| 4.1 引言.....                                     | 67        |
| 4.2 数据采集与前处理.....                               | 68        |
| 4.2.1 数据采集.....                                 | 68        |
| 4.2.2 数据前处理.....                                | 68        |
| 4.3 基于 BLSTM 模型的产品混合剩余寿命预测方法.....               | 69        |
| 4.3.1 基于 BLSTM 模型的产品退化指标的构建.....                | 70        |
| 4.3.2 融合粒子滤波算法的混合剩余寿命预测方法.....                  | 72        |
| 4.4 算例分析与方法验证.....                              | 75        |
| 4.4.1 数据集简介.....                                | 75        |
| 4.4.2 数据前处理.....                                | 76        |
| 4.4.3 方法验证与结果讨论.....                            | 78        |
| 4.5 本章小结.....                                   | 83        |
| <b>第五章 基于深度卷积神经网络和 BOOTSTRAP 的剩余寿命预测方法.....</b> | <b>84</b> |
| 5.1 引言.....                                     | 84        |
| 5.2 数据前处理.....                                  | 85        |
| 5.2.1 首次预测时间识别.....                             | 85        |
| 5.2.2 基于连续小波变换的时频表征提取.....                      | 87        |
| 5.3 基于深度卷积神经网络的 RUL 预测模型.....                   | 88        |
| 5.3.1 卷积神经网络简介.....                             | 88        |
| 5.3.2 基于深度卷积神经网络的 RUL 预测模型.....                 | 90        |
| 5.4 基于 DCNN 模型和 BOOTSTRAP 框架的 RUL 预测方法.....     | 92        |
| 5.4.1 预测区间的构建.....                              | 93        |
| 5.4.2 基于 DCNN-Bootstrap 方法的 RUL 预测区间量化.....     | 95        |
| 5.5 算例分析与方法验证.....                              | 98        |
| 5.5.1 轴承退化实验数据简介与前处理.....                       | 98        |
| 5.5.2 结果分析与讨论.....                              | 102       |

|                     |     |
|---------------------|-----|
| 5.6 本章小节 .....      | 107 |
| 第六章 总结与展望 .....     | 108 |
| 6.1 本文主要研究结论 .....  | 108 |
| 6.2 后续工作展望 .....    | 109 |
| 致 谢 .....           | 111 |
| 参考文献 .....          | 112 |
| 攻读博士学位期间取得的成果 ..... | 123 |

## 图目录

|                                            |    |
|--------------------------------------------|----|
| 图 1-1 论文结构框架图 .....                        | 15 |
| 图 2-1 基于 TSBP 方法的实施框架图 .....               | 18 |
| 图 2-2 基于 TSBP 方法的数据前处理流程图 .....            | 21 |
| 图 2-3 退化轨迹相似性分析与剩余寿命预测 .....               | 22 |
| 图 2-4 基于 TSBP 改进方法的剩余寿命预测框架 .....          | 26 |
| 图 2-5 按权重值排序后的 $(w^j, PRL^j)$ 三维示例图 .....  | 26 |
| 图 2-6 按权重值排序后的 $(w^j, PRL^j)$ 二维示例图 .....  | 27 |
| 图 2-7 PRLs 的异常值剔除示例图 .....                 | 27 |
| 图 2-8 PRLs 等权重样本扩充方法 .....                 | 29 |
| 图 2-9 等权重样本扩充后所引入不确定性示例图 .....             | 30 |
| 图 2-10 基于 RUL 的点估计值的容许误差界示例图 .....         | 31 |
| 图 2-11 NASA 涡扇发动机结构示意图 .....               | 34 |
| 图 2-12 训练发动机运行至失效时 14 个传感器的监测信号 .....      | 35 |
| 图 2-13 训练发动机退化模型库示例图 .....                 | 35 |
| 图 2-14 传统方法与改进方法就发动机#43 的 RUL 预测结果对比 ..... | 37 |
| 图 2-15 传统方法与改进方法就发动机#74 的 RUL 预测结果对比 ..... | 38 |
| 图 3-1 滑动时间窗处理技术示例图 .....                   | 43 |
| 图 3-2 基于 BLSTM 的 RUL 预测方法实施流程 .....        | 44 |
| 图 3-3 FFNN 的网络结构示意图 .....                  | 44 |
| 图 3-4 RNN 的网络结构示意图 .....                   | 46 |
| 图 3-5 LSTM 单元的结构示意图 .....                  | 47 |
| 图 3-6 GRU 单元的结构示意图 .....                   | 49 |
| 图 3-7 本章所提的基于 BLSTM 的网络结构 .....            | 51 |
| 图 3-8 深度栈式 BLSTM 的网络结构 .....               | 54 |
| 图 3-9 FD002 训练发动机 21 个传感器监测信号 .....        | 58 |
| 图 3-10 分段线性退化模型 .....                      | 58 |
| 图 3-11 某训练发动机多传感器原始监测数据 .....              | 59 |
| 图 3-12 某发动机归一化多传感器监测数据 .....               | 59 |
| 图 3-13 不同网络结构 20 次运行的预测性能箱线图 .....         | 61 |
| 图 3-14 不同长度序列数据的 RUL 预测结果对比示意图 .....       | 62 |

|                                                  |     |
|--------------------------------------------------|-----|
| 图 3-15 考虑或忽略辅助输入数据对 RUL 预测结果的影响 .....            | 63  |
| 图 4-1 一种基于 BLSTM 模型的混合 RUL 预测方法实施流程 .....        | 70  |
| 图 4-2 所提出的基于 BLSTM 模型的网络结构 .....                 | 71  |
| 图 4-3 高速铣削铣床刀具磨损状态监测实验平台 .....                   | 75  |
| 图 4-4 基于刀具磨损实验 C1 的六个信号通道监测数据简化图 .....           | 76  |
| 图 4-5 所提取的归一化处理的特征示意图 .....                      | 77  |
| 图 4-6 基于 BLSTM-PF 混合方法的 RUL 预测实施流程 .....         | 78  |
| 图 4-7 BLSTM-PF 混合方法在不同情况下 RUL 预测的结果 .....        | 81  |
| 图 4-8 等间距观测时间下的刀具 RUL 概率密度分布估计结果 .....           | 82  |
| 图 5-1 滚动轴承振动测量信号示意图 .....                        | 85  |
| 图 5-2 轴承首次预测时间识别示意图 .....                        | 86  |
| 图 5-3 提取特征与 RUL 标记值示意图 .....                     | 86  |
| 图 5-4 基于 CWT 算法所提取的轴承时频域特征 .....                 | 88  |
| 图 5-5 CNN 网络的基本结构 .....                          | 89  |
| 图 5-6 卷积与池化运算简化示意图 .....                         | 90  |
| 图 5-7 所提出的基于 DCNN 模型的网络结构 .....                  | 91  |
| 图 5-8 模型预测方差与预测误差示意图 .....                       | 94  |
| 图 5-9 Bootstrap 实施流程示意图 .....                    | 95  |
| 图 5-10 基于 DCNN-Bootstrap 方法的 RUL 预测区间量化流程图 ..... | 97  |
| 图 5-11 滚动轴承加速退化测试平台 .....                        | 99  |
| 图 5-12 振动信号采样设置 .....                            | 99  |
| 图 5-13 实验中滚动轴承四种不同的失效模式 .....                    | 100 |
| 图 5-14 针对两组轴承所提取的两个通道的时域特征 .....                 | 100 |
| 图 5-15 轴承原始时频特征和其对应的压缩时频特征 .....                 | 102 |
| 图 5-16 针对 Bearing1_3 的 RUL 预测结果对比图 .....         | 104 |
| 图 5-17 针对 Bearing2_5 的 RUL 预测结果对比图 .....         | 104 |
| 图 5-18 针对 Bearing3_4 的 RUL 预测结果对比图 .....         | 105 |
| 图 5-19 针对 Bearing1_5 的 RUL 预测区间量化结果 .....        | 106 |
| 图 5-20 针对 Bearing2_2 的 RUL 预测区间量化结果 .....        | 106 |
| 图 5-21 针对 Bearing3_3 的 RUL 预测区间量化结果 .....        | 106 |

## 表目录

|                                                       |     |
|-------------------------------------------------------|-----|
| 表 2-1 常用核函数.....                                      | 28  |
| 表 2-2 航空发动机退化仿真数据集-有效传感器描述.....                       | 34  |
| 表 2-3 不同参数组( $\Delta t, M$ )下改进方法的 $M_{PS}$ 预测结果..... | 36  |
| 表 2-4 改进方法与公开发表方法的性能对比结果.....                         | 37  |
| 表 2-5 传统方法与改进方法的 RUL 预测性能对比.....                      | 38  |
| 表 3-1 NASA 航空发动机退化仿真数据集简介.....                        | 57  |
| 表 3-2 基于网格搜索法的超参数选取列表.....                            | 60  |
| 表 3-3 不同的网络结构的平均预测性能.....                             | 60  |
| 表 3-4 不同长度序列数据的平均预测性能.....                            | 63  |
| 表 3-5 考虑或忽略辅助输入数据的平均预测性能.....                         | 64  |
| 表 3-6 与基于 RNN 及其变体结构方法的预测性能对比.....                    | 64  |
| 表 3-7 与最新的基于 AI 技术的相关方法的预测性能对比.....                   | 65  |
| 表 4-1 特征选择步骤对预测结果的影响.....                             | 79  |
| 表 4-2 特征融合步骤对预测结果的影响.....                             | 79  |
| 表 4-3 SVR 模型参数设置.....                                 | 80  |
| 表 4-4 所提方法与相关方法的预测性能对比.....                           | 80  |
| 表 4-5 不同观测时刻点的 RUL 预测结果.....                          | 82  |
| 表 5-1 XJTU-SY 轴承数据集详细信息.....                          | 101 |
| 表 5-2 Hybrid DCNN 方法与各对比方法的预测性能结果.....                | 103 |

## 第一章 绪 论

### 1.1 研究背景及意义

随着传感器、信息与通信技术及物联网技术等快速发展，状态监测技术已经广泛地应用于各类复杂的产品，例如：航空发动机、高档数控机床、风力发电机等。各类传感器部署于产品的各个部位并以极高的采样频率对产品的健康状态进行实时在线监测，所获得的有关产品的状态监测数据呈现井喷式的增长，这不可避免地将我们带入了工业大数据时代。例如，空客公司在 A350 飞机上部署了超过 6000 个传感器以实时捕捉各类数据，包括燃料流量、压力、温度、飞行高度、飞行速度和外部环境温度等，每架飞机每天将会产生高达 2.5 TB 的监测数据；昆仑数据和新疆金风科技有限公司合作开发了风力装备 KMX 大数据平台来监测风机的实际运行状况，每台风机的数据测点约为五百个，数据回传频率为每秒 50 组，在峰值状态下，20000 台风机每秒会产生超过千万条数据，每天会新增近 1TB 的监测数据和运行日志数据；通用电气公司对某风力发电厂的风力涡轮机（共计 200 台）部署了约 50 个传感器和相应的数据采集的控制回路，并结合超过上千台风力涡轮机组的历史数据实时分析涡轮机组的健康状态；中联重科公司所开发的大数据平台收集了来自该公司的工程机械、农业机械等产品的物联网数据。当前中联重科物联网平台已累积了近 10 年数据，监控产品约为 12 万台，存量数据约 40 TB，每月新增数据约 300 GB。

通过上述众多公司对各自旗下产品的实时监测案例可知，产品的监测大数据具备以下特点：（1）相似设备群规模大、历史监测数据充足，需要构建利用海量历史监测数据与在线监测数据进行分析的新方法；（2）监测数据规模大，且每日新增数据量多，不能仅依赖于专家经验或专业人员人工设计故障特征提取算法，需要利用最新的技术以实现特征自动提取并进一步设计端对端的智能算法；（3）监测数据类型多，所采集的监测数据包含各类传感器数据和时变的工况数据等，需要构建能融合复杂装备系统多源监测数据的新方法。

《国家中长期科学和技术发展规划纲要（2006-2020 年）》（以下简称《纲要》）明确将重大产品和重大设施寿命预测技术列为先进制造领域中需要攻克的前沿技术。《纲要》还进一步指出：重大产品和重大设施寿命预测技术是提高产品运行可靠性、安全性、可维护性的关键技术。在《机械工程学科发展战略报告（2011-2020 年）》（以下简称《战略报告》）中将寿命预测方法列为未来 5~10 年的研究前沿与重大科学问题。《战略报告》还指出：在极端环境下服役的重大机械产品具有较高的

失效风险，基于信号采集、特征提取、分析等系列方法的寿命监测技术可以为重大机械产品的服役安全保驾护航。因此，在工业大数据背景下，如何基于产品监测大数据所呈现的特点，开发解析海量监测数据的新方法并进一步探明复杂产品的健康状态，对其剩余寿命进行更加精确的预测对保障产品的安全运行至关重要。

产品剩余寿命预测是预测与健康管理（Prognostic and Health Management, PHM）技术的核心概念，更是产品进行健康管理的基础决策信息。对产品剩余寿命进行高效、准确且稳健的预测有助于产品提前报警，从而预防灾难性事故的发生，进一步优化其维护策略以降低昂贵且非必需的维护费用，同时还能进一步提升产品的可靠性、可用性和安全性。此外，产品剩余寿命预测技术研究是智能制造的关键组成部分，同时也是国家战略“中国制造 2025”的关键核心技术。在当今工业大数据背景下，现有的产品剩余寿命预测理论和技术遇到了全新的挑战：（1）如何有机地结合海量的相似产品的历史监测数据与在线监测数据以实施精确且稳健的剩余寿命预测；（2）如何利用最新的深度学习技术减少对专家经验和领域内相关专业知识的依赖，进一步实施产品监测大数据的自动解析，最终实现产品端对端的剩余寿命预测；（3）如何有效地融合产品在时变工况和多失效模式耦合作用下的多源监测信息，从而提升模型预测精度；（4）针对失效机理清晰且失效物理模型已知的产品，如何将产品失效物理模型与传感器监测大数据进行有效的融合，最终构建基于失效物理模型和数据驱动的混合剩余寿命预测方法；（5）针对失效机理不可获知且失效模型难以构建的复杂产品，如何在深度学习技术自动解析产品监测大数据的框架内构建更加泛化的数据驱动方法，以对产品的剩余寿命预测区间进行进一步量化。因此，研究新的基于监测大数据驱动的产品剩余寿命预测理论和技术具有重大的战略意义，同时也是保障复杂产品安全运行的关键。

## 1.2 产品剩余寿命预测方法研究综述

产品剩余寿命主要是指产品运行至当前时刻与产品失效时刻之间的时间间隔。具体来说，当某一产品  $X$  运行至当前时刻  $t$ ，根据该产品运行至  $t$  时刻的状态监测信息和/或同类型产品的海量历史监测信息，预测产品  $X$  由当前运行状态至失效状态的剩余使用寿命（Remaining Useful Life, RUL）。Jardine 等<sup>[1]</sup>将产品剩余寿命定义为：

$$T - t | T > t, Z(x), [Z_i(y), E_i] \quad (1-1)$$

式中， $T$  表示产品的失效时刻，为一随机变量， $t$  为当前时刻， $T - t$  为产品在当前时刻的剩余寿命， $Z(x)$  为产品  $X$  运行至当前时刻的监测信息且  $x \in [0, t]$ ， $Z_i(y)$

表示同类型产品  $i$  当其运行至  $y$  时刻的状态监测数据,  $E_i$  为同类型产品  $i$  对应的事件数据 (Event Data: 这里可以是产品的失效时间数据也可以是产品维护或更换时间数据等),  $y \in [0, T_i]$ ,  $i = 1, 2, \dots, M$ ,  $M$  为同类型产品的个数,  $T_i$  为同类型产品  $i$  的事件数据 (大多数时候指同类型产品的失效时间数据)。变量  $T-t$  即为产品的剩余寿命, 根据 Jardine<sup>[1]</sup> 的综述论文可知, 产品剩余寿命预测主要有两种内涵, 其一为对变量  $T-t$  的概率密度分布进行估计, 其二为直接对变量  $T-t$  的期望进行估计, 即为,

$$E[T-t | T > t, Z(x), [Z_i(y), E_i]] \quad (1-2)$$

目前, 关于产品剩余寿命预测方法的分类尚存在争议, 不同学者或科研组织都提出了不尽相同的分类方法。例如: 多伦多大学 Jardine 等<sup>[1]</sup>将其分为“基于统计模型、基于物理模型、基于人工智能”的方法, 辛辛那提大学 Jay Lee 等<sup>[2]</sup>将其分为“基于物理模型、基于数据驱动、基于混合模型”的方法, 国际标准化组织 ISO<sup>[3]</sup>于 2015 年发表了有关机器状态监测、诊断和预测的通用指南, 在指南中将机器剩余寿命预测方法分类为“基于物理模型、基于统计模型、基于启发式模型、基于数据驱动、基于混合模型”的方法, 西安交通大学雷亚国等<sup>[4]</sup>将其分为“基于物理模型、基于统计模型、基于人工智能技术、基于混合模型”的方法。本文根据各类方法的具体技术细节和整体的方法论, 以及学术界在该研究领域所发表文献中常用命名方式, 将产品剩余寿命预测分为基于物理模型的方法, 基于数据驱动的方法和基于混合模型的方法三大类。下文将对这三类产品剩余寿命预测方法的研究现状分别进行综述。

### 1.2.1 基于物理模型的剩余寿命预测方法研究

基于物理模型的剩余寿命预测方法依据产品的失效机理或损伤法则构建相应的能具体阐释产品退化规律的物理模型, 并有效地结合实时在线监测数据实现对设备的剩余寿命评估。一般来说, 物理模型中的各个参数往往与产品组成材料的特性和外部应力水平相关, 这些参数通常可以通过特定的实验、有限元分析和其他相关技术进行确定。

Paris 等<sup>[5]</sup>于 1963 年针对机械材料疲劳裂纹扩展问题建立了 Paris-Erdogan 模型 (简称 Paris 模型)。由于大多数机械产品的退化过程都直接或间接的与疲劳裂纹扩展相关, 因此 Paris 模型被广泛应用于机械产品的剩余寿命预测。严如强等<sup>[6]</sup>提出基于改进的 Paris 裂纹增长模型来刻画轴承缺陷传播的过程, 并结合高维自回归模型和相空间卷曲技术实现对轴承寿命的预测。Zio 等<sup>[7-8]</sup>针对测量不确定性下

部件的裂纹扩展问题,采用粒子滤波技术对模型参数进行估计并更新,最终实现部件剩余寿命的评估。Billington 等<sup>[9]</sup>基于 Paris 裂纹增长模型,利用非线性递归最小二乘法对滚珠轴承的剩余寿命进行了预测。Orchard 等<sup>[10]</sup>针对非线性、非高斯系统提出了粒子滤波技术的在线故障监测和剩余寿命预测框架,并对行星齿轮箱盘轴向裂纹扩展的剩余寿命进行预测。Zhao 等<sup>[11]</sup>利用贝叶斯更新方法针对齿轮构建了基于物理模型的剩余寿命预测框架,通过引入贝叶斯方法,融合所采集到的齿轮的状态监测信息并更新齿轮剩余寿命的后验分布。所提出的物理模型主要涵盖三个部分:用于齿轮应力分析的有限元模型、用于动力学载荷计算的齿轮动力学模型、用于刻画齿轮损伤传播的 Paris 模型。雷亚国等<sup>[12]</sup>针对滚动轴承的剩余寿命预测问题提出了两阶段的基于模型的预测方法。在退化指标构建阶段,一种新的基于权重最小化量化误差的健康指标被提出以融合多特征信息。在剩余寿命估计阶段,使用了极大似然估计方法对基于 Paris 的状态空间模型中的各个参数进行初始化,并采用粒子滤波技术更新模型中的各参数估计,最终实现对滚动轴承剩余寿命的估计。Pecht 等<sup>[13]</sup>同样基于 Paris 模型构建了裂纹增长的状态空间退化模型,并在贝叶斯更新框架内实现了退化状态的估计和预测,且所提出的方法在经典的 Virkler 裂纹增长数据集中得到了成功的应用。Liao<sup>[14]</sup>和 Wang<sup>[15]</sup>等均将 Paris 模型转变为状态空间模型,并分别使用贝叶斯更新方法和粒子滤波方法实现轴承剩余寿命的预测。除了被广泛应用的 Paris 模型以外,还存在刻画机械产品失效演进过程的其他模型。例如,Zio 等<sup>[16-17]</sup>采用 Norton 法则描述核能发电机厂涡轮叶片的蠕变退化演进过程,并分别利用卡尔曼滤波技术和粒子滤波技术实现了涡轮叶片的在线剩余寿命预测。Goebel 等<sup>[18]</sup>建立了基于经典的 shear-lag 模型来刻画聚合物材料的疲劳退化过程,并结合粒子滤波技术对其时变可靠性模型中的各个参数进行估计进而得到其剩余寿命的分布。Chan 等<sup>[19]</sup>构建了与时间相关的裂纹增长模型,并成功应用在推进系统热端部件的寿命预测中。Khaled 等<sup>[20]</sup>基于非线性累积损伤法则提出了一种解析式的剩余寿命预测框架,并在所提出的框架内考虑所施加载荷和初始裂纹长度的随机性,最后将其应用到石油化工产业输油管道的剩余寿命预测。

综上所述,基于模型的剩余寿命预测方法主要应用于一些失效机理已知且失效物理模型能准确刻画退化演进过程的特定研究对象,例如,齿轮、轴承、涡轮叶片、聚合物材料等。所构建的物理模型参数一旦被准确识别,那么基于预先设定的失效阈值即可通过模型随着时间的演进从而估计研究对象的剩余寿命。但是,针对某些产品其失效过程复杂且存在多种失效模式耦合的现象,很难构建其准确的失效物理模型以刻画对应的退化过程。此外,基于模型的方法往往需要结合已知有关产品监测到的退化数据(例如裂纹扩展的长度等)实施剩余寿命预测,但大多数复

杂产品的退化状态难以通过直接的手段进行实时在线观测，这无疑又大大增加了基于物理模型的剩余寿命预测方法在实际工程应用中的难度。

### 1.2.2 基于数据驱动的剩余寿命预测方法研究

基于数据驱动的剩余寿命预测方法主要适用于有关产品的失效物理模型不可知，但同时存在大量的历史监测数据可供使用的情况。基于数据驱动的方法通过统计模型、人工智能技术或相似性分析技术对所获得的监测数据直接进行建模以探明产品的退化规律进而获得对应的剩余寿命估计。根据对产品历史监测数据的建模方式，本文进一步将基于数据驱动的剩余寿命预测方法划分为“基于统计模型的方法、基于人工智能的方法、基于相似性的方法”。本部分将依次综述这三种不同数据驱动的剩余寿命预测方法。

#### (1) 基于统计模型的方法。

基于统计模型的预测方法通过构建依赖于经验知识的统计模型估计产品的剩余寿命概率密度分布。在这类方法中，剩余寿命预测模型是通过拟合在线监测数据为一随机系数模型或统计过程模型建立的。其中，一些随机变量被引入为模型参数以刻画多源不确定性信息，例如，样本之间的差异性、测量误差等。因此，基于统计模型的预测方法可以有效地描述退化过程的不确定性及其对剩余寿命估计的影响。

随机系数模型通过引入随机系数到产品的退化模型中以刻画产品退化过程的随机性。Gebrael 等<sup>[21]</sup>提出了带误差项的双参数指数退化模型，其中误差项被假定为独立同分布的标准正态分布，并引入贝叶斯更新方法估计所提出指数退化模型的随机参数分布，从而获得了剩余寿命分布的解析解。为了验证所提出的退化模型的有效性，所提出的带误差项的指数退化模型被应用在轴承的剩余寿命估计中。Coble 等<sup>[22]</sup>基于贝叶斯统计的方法将先验信息引入到随机系数模型中，并进一步使用贝叶斯更新方法得到模型参数的后验估计。所提出方法被有效地应用到了航空发动机退化仿真数据集与数控机床刀具耗损的剩余寿命预测任务中。Park 等<sup>[23]</sup>提出了带随机误差项的双指数模型以刻画退化过程，并利用极大似然方法对模型中的各个参数进行估计，最终将所提出的模型应用于有机发光二极管的失效时间分布的估计。Jin 等<sup>[24]</sup>基于对轴承磨损期的监测数据的回归分析提出了双指数和函数的随机系数模型以描述轴承的退化过程，并利用扩展卡尔曼滤波算法实现了对所提出模型的参数更新和轴承的剩余寿命估计。Hu 等<sup>[25]</sup>同样提出了双指数和函数的随机系数模型以刻画油砂泵叶轮的退化过程，并结合相关向量机方法估计了叶轮的剩余寿命分布。随机系数模型能有效地给出产品剩余寿命的概率密度分布的估

计,但是该模型假定随机系数服从正态分布且不能有效地刻画产品剩余寿命预测的时间变化,这也就限制了该方法在实际工程中的应用。

基于统计过程模型的剩余寿命预测方法主要假定产品的退化过程为一随机过程模型。Huang 等<sup>[26]</sup>构建了线性维纳过程对工业设备的退化过程进行建模,并提出在线滤波算法以充分地利用先验信息和历史数据对设备的状态进行估计,并基于极大似然估计方法提出了估计模型参数的两阶段算法,最终采用旋转电机滚珠轴承的剩余寿命预测问题对所提方法进行验证。Liao 等<sup>[27]</sup>考虑设备工作在时变工况环境下同样提出了线性维纳过程模型,并使用贝叶斯更新方法实现了对模型参数的后验分布估计和剩余寿命预测,最终所提出的方法在仿真实验和有关滚珠轴承的实际案例分析中得到了验证。针对基于线性维纳过程模型的机械产品剩余寿命预测方法研究还可参考文献[28-32]。Si 等<sup>[33]</sup>针对退化过程中的动态性和非线性问题提出了基于维纳过程的非线性模型,并在适当的假设条件下得到了剩余寿命分布的近似解析解,其中退化模型的未知参数通过极大似然估计方法进行估计。由于有效地考虑了退化过程非线性,因此所提出模型在有关激光器数据和裂纹增长数据中获得了更高的剩余寿命预测精度。一些学者针对所提出的基于维纳过程的非线性模型进行了拓展研究,相关研究内容可参考文献[34-35]。除了基于维纳过程的模型,常见的用于产品剩余寿命预测的随机过程模型还有伽马过程<sup>[36-40]</sup>和逆高斯过程<sup>[41-45]</sup>,在这里不一一赘述。值得注意的是,上述三种随机过程模型均具有马尔可夫特性,即假定产品未来的退化状态仅与当前的退化状态相关而与之前的历史退化状态无关,但是该假设在一些实际的工程应用中并不全都适用。

除了上述基于随机系数模型和基于统计过程模型的方法,基于马尔可夫模型的方法也常用来估计产品剩余寿命分布。马尔可夫模型假定产品退化过程是在服从马尔可夫特性的有限个状态空间内的转移过程,该模型最初由 Jeffrey<sup>[46]</sup>提出以解决产品随时间的累积损伤是一连续耗损过程情况下的失效时间分布的估计问题。Jeffrey 等<sup>[47-48]</sup>在随后的研究中进一步完善了在马尔可夫工况下和半马尔可夫工况下针对产品的剩余寿命估计的相关理论和方法。刘宇等<sup>[49]</sup>针对退化过程分别服从齐次连续时间马尔可夫过程和非齐次连续时间马尔可夫过程的多状态系统,提出了利用系统监测数据的动态可靠性评估方法,使用了递归贝叶斯方法动态地更新模型的参数并获得了多状态系统剩余寿命分布。基于马尔可夫模型的方法通常假定产品退化过程的退化状态为可观测的情形,但是在实际工程应用中大多数产品的退化状态均不能直接进行在线观测,例如,数控机床加工过程中刀具的耗损状况不能直接在线进行监测,为了获取刀具的退化状况必须停机进行离线测量。因此,提出隐马尔可夫模型<sup>[50]</sup>以解决上述产品退化状态不能直接获取的问题。Ramasso 等

[51]利用隐马尔可夫模型表征产品不可观测的健康状态,在信任函数框架内表示有关产品不可观测的健康状态的部分信息,引入了极大证据期望算法对模型的参数进行估计,并采用基于 Viterbi 的变体算法解决有关状态推理的问题。所提出的方法在航空发动机数据集上获得了较好的剩余寿命估计结果。Syed 等[52]利用隐马尔可夫模型对汽车直流电机齿轮进行了故障预测分析,将发动机电流的时频特征提取为机器的健康指标并作为模型的输入数据以训练模型参数和预测发动机故障。此外,比例风险模型也是一类被用来估计产品剩余寿命的模型。比例风险模型假定系统的失效率函数为两个相乘因子组成,其一为与时间相关的基底失效率函数,再者为一协变量函数。Liao 等[53]提出了用于轴承剩余寿命估计的比例风险模型,其中轴承振动信号的均方根和峭度值被用作比例风险模型的协变量,试验表明,所提出的比例风险模型所获得的剩余寿命预测精度显著高于基于 logistic 回归模型的方法。Jardine 等[54]针对大型运输机的转动系统建立了比例风险模型,在所提出的模型中失效过程被表示为一随机离散马尔可夫过程,并进一步给出了相应的条件可靠度函数和剩余寿命计算的方法。此外,火箭军工程大学司小胜等[55]也对基于统计模型的剩余寿命预测方法进行了综述。

## (2) 基于人工智能的方法。

基于人工智能的剩余寿命预测方法在近年来随着人工智能理论和应用的进一步发展备受关[56]。基于人工智能的方法主要通过利用先进的人工智能技术直接对所监测的传感器数据和产品的寿命数据进行建模。由于该方法不依赖于任何与产品相关的物理模型和与退化过程相关的统计模型,所以基于人工智能的方法对复杂产品的剩余寿命预测研究具有巨大的应用前景。由于深度学习理论与技术在近些年蓬勃发展,基于人工神经网络的方法受到了广泛且深入的关注。因此,本部分将着重对基于人工神经网络的剩余寿命预测方法进行综述。

人工神经网络由于其具有高度非线性、自适应性和任意函数拟合的功能,从而被广泛地应用于复杂产品的剩余寿命预测研究中。此外,随着近年来深度学习研究的进一步推进,基于人工神经网络的方法更是在学术界与工业界得到了前所未有的关注。多层前向神经网络(FFNN)是一类被广泛用于产品剩余寿命预测的传统神经网络[57-60],其中有关产品的健康指标(或退化指标)被用于 FFNN 的输入,产品的剩余寿命被直接用作 FFNN 的输出,并基于特定的优化算法 FFNN 迭代调整其模型内部的参数值以对输入输出的高度非线性关系进行建模分析。Wu 等[57]将滚珠轴承的振动监测信号的均方根值作为其退化指标,并将轴承当前时刻和前一时刻的退化指标值及其当前监测时间点信息作为 FFNN 模型的输入,轴承的寿命百分数作为模型的输出。此外,基于 Levenberg-Marquardt (LM) 优化算法对 FFNN

模型进行训练,最后将完成训练的模型用于轴承的实时在线剩余寿命预测。Tian<sup>[58]</sup>进一步改进了 Wu 所提出的基于 FFNN 的模型。在 Tian 所提出的模型中,可以输入有关产品的多个退化指标的连续观测值和其对应的监测时刻点信息。此外, Tian 等<sup>[59]</sup>还进一步拓展他之前基于 FFNN 的研究工作。在所提出的基于 FFNN 模型的拓展框架内,同时利用轴承的失效和截尾监测数据以进一步提升所提出模型的预测性能。由于 FFNN 模型在本质上属于浅层神经网络,因此仅能对产品简单的退化信息进行特征提取并利用其回归分析能力对产品的剩余寿命进行直接预测。

近年来深度学习的兴起为提取产品复杂退化信息的特征提供了另外一种建模的思路。其中循环神经网络<sup>[61-65]</sup>、卷积神经网络<sup>[66-68]</sup>和深度信念网络<sup>[69-72]</sup>是学术界最为常用的实现产品剩余寿命预测的三种深度学习神经网络。

袁梅等<sup>[61]</sup>利用循环神经网络中的长短期记忆神经网络 (Long Short-Term Memory, LSTM) 对时间序列数据建模的优势,实现了对航空发动机多传感器数据的特征提取和剩余寿命评估。Ellefsen 等<sup>[62]</sup>同样提出利用长短期记忆神经网络和前馈神经网络的复合网络结构,实现了对航空发动机传感器监测数据的时间相关性的建模和剩余寿命的回归分析。Zhao 等<sup>[63]</sup>将循环神经网络中 GRU (Gated Recurrent Unit, GRU) 网络应用于机器健康监测中,首先提取机器监测数据的时域、频域和时频域特征,再将所提取的局部特征序列输入到双向 GRU 网络实现对机器的深层次的故障特征挖掘,最后通过网络的不同功能输出层的定义分别实现了对机器的故障诊断 (使用 softmax 层) 和健康状态监测 (使用回归层)。Ren 等<sup>[64]</sup>针对轴承的剩余寿命预测问题提出了多尺度全连接 GRU 网络,其中网络特征提取层的参数通过预训练受限玻尔兹曼机 (Restricted Boltzman Machine, RBM) 进行初始化。针对工作在多工况和复杂失效模式下的航空发动机剩余寿命预测问题, Huang 等<sup>[65]</sup>提出了一种基于双向长短期记忆神经网络 (Bi-directional Long Short-Term Memory, BLSTM) 和 FFNN 的复合深度神经网络结构模型。其中经归一化和序列化处理的多传感器信息被用作所提出模型的主输入,动态工况信息被用作模型的辅助输入。BLSTM 网络被用来实现对融合输入数据的深度特征提取,FFNN 网络被用来进一步完成对所提取特征和发动机剩余寿命的回归分析任务。

Li 等<sup>[66]</sup>提出了一维深度卷积神经网络 (Deep Convolution Neural Network, DCNN) 模型以实现航空发动机的剩余寿命预测。在所提出的框架内,应用了时间窗处理技术以处理多传感器原始数据为序列数据,并将所获得序列传感器数据输入到一维 DCNN 模型中进行特征提取和剩余寿命的回归分析。值得注意的是 DCNN 模型具备很强的提取二维数据 (图片数据) 层级特征的能力。因此一些研究者尝试通过基于原始传感器数据并利用信号处理技术以构建产品的二维退化特征

信息，并利用 DCNN 模型进一步提取产品深层次故障特征并预测其剩余寿命。陈楠等<sup>[67]</sup>提出利用连续小波变换对轴承的振动信号实施时频分析并进一步获得轴承的时频特征（二维数据）。其次，利用多尺度卷积神经网络对轴承的时频信息里的全局特征和局部特征进行深度特征提取，并同样利用最后的回归层实现了对所提取特征和轴承的剩余寿命的回归分析。Li 等<sup>[68]</sup>提出了采用短时傅里叶变化对轴承的振动信号进行时频特征提取，并同样基于多尺度卷积神经网络对轴承的时频序列信息的全局特征和局部特征进行深度特征提取，最后利用回归层得到了轴承的剩余寿命百分比。

He 等<sup>[69]</sup>构建了基于深度信念网络（Deep Belief Networks, DBN）和 FFNN 的复合神经网络结构以实现轴承的剩余寿命预测。其中利用 DBN 的自我特征学习能力实现对轴承振动信号的特征挖掘，此后 FFNN 被用来完成对所提取特征和轴承剩余寿命的回归分析。Zhang 等<sup>[70]</sup>提出了一种多目标深度信念网络集成的方法以预测航空发动机的剩余寿命。在所提出的方法中，同时利用多目标遗传算法和传统的 DBN 训练方法来完成对所提出的 DBN 网络的训练。此后，多个训练完成的 DBN 网络按权重进行有效地结合以形成一个集成 DBN 的剩余寿命预测模型。其中，权重是通过单目标差分遗传算法优化而获得。Ma 等<sup>[71]</sup>针对机器健康状态评估问题提出了基于 DBN 和蚁群算法的预测框架，其中 DBN 用于提取机器监测数据的故障特征，蚁群算法用于自动优化 DBN 模型参数。最后将所提出的方法分别用于轴承和航空发动机的健康状态评估以验证其有效性。Zhao 等<sup>[72]</sup>提出了基于 DBN 和相关向量机的锂离子电池剩余寿命预测框架，其中采用 DBN 提取锂离子电池的退化指标（电池容量）的特征，并使用相关向量机对所提取特征和锂离子电池的剩余寿命进行建模分析。经试验验证，所提出的方法比传统的基于相关向量机的方法能给出更高精度且更稳定的剩余寿命预测结果。

除了上述基于人工神经网络的剩余寿命预测方法，当前基于人工智能方法所采取的人工智能技术还包括支持向量机、相关向量机、神经模糊系统等。支持向量机（Support Vector Machine, SVM）理论是由 Vapnik 等<sup>[73]</sup>于 1995 年提出且主要分为支持向量分类（Support Vector Classification, SVC）和支持向量回归（Support Vector Regression, SVR）两种方法，其中 SVR 被广泛用于产品的剩余寿命预测研究<sup>[74-75]</sup>。但是，基于 SVR 的方法只能给出产品剩余寿命的点估计值，因此为解决产品剩余寿命分布的估计问题，提出了基于相关向量机（Relevance Vector Machine, RVM）的方法<sup>[76-78]</sup>。与基于人工神经网络的方法相比，基于 SVR 或 RVM 的方法不需要大量的训练样本数据，但是该类方法的预测性能与核函数的选取密切相关，并且用于产品剩余寿命预测的核函数选取标准也尚未确立<sup>[4]</sup>。此外，该类方法需要

首先人工提取监测数据（例如，振动数据，声发射数据等）中的特征，并进一步对所提取的特征进行选择 and 降维处理，此后将所获得的有效特征集输入到 SVR 或 RVM 的模型中，最后通过构建产品退化指标（或健康指标）与有效特征集的回归模型（与核函数相关），进而外推其退化指标至产品预设的失效阈值处从而获得产品的剩余寿命估计。因此，该类方法往往涵盖了复杂的处理流程且极大地受相关专业知识的限制。神经模糊系统（Neural Fuzzy System, NF system）是基于神经网络的模糊推理系统，其中预测部分由模糊逻辑系统完成。模糊逻辑系统的结构由专家经验决定，隶属函数是由神经网络训练而获得。基于神经模糊系统的方法常常被用作一种时间序列预测技术以实施产品退化指标的多步提前预测，并结合其他算法（如粒子滤波）以估计产品的剩余寿命概率密度分布<sup>[79-80]</sup>。基于神经模糊系统的方法充分利用了专家经验信息，但是该方法和基于人工神经网络方法一样，都需要大量的高质量训练数据以获得模型高精度的预测性能。

### （3）基于相似性的方法。

基于相似性的方法是另一种常用的数据驱动的产品剩余寿命预测方法，该方法最初由辛辛那提大学 Wang 等<sup>[81]</sup>于 2008 年 PHM 数据挑战竞赛中提出，此后该方法在工业界和学术界得到了广泛的推广应用。该方法的核心思想是基于目前正处于实时监测下的服役的产品与已经失效的产品如果来自一批相同或相似的产品且工作在类似的工况环境下完成大致一样的工作任务，那么它们的退化趋势将呈现相似性。因此，可以认为正在服役的产品的剩余寿命可通过对已经失效的相似产品在某一时刻剩余寿命的加权平均进行计算，其中权重值可通过服役产品与失效产品之间的相似性进行计算，相似性可进一步通过服役产品与失效产品之间退化趋势或状态监测数据进行相似性匹配而确定。Zio<sup>[82]</sup>等采用基于模糊相似性的分析技术对具有大量失效历史信息的某核能系统进行了剩余寿命分析。孟光等<sup>[83]</sup>认为越接近设备失效的时刻在相似性分析的时候应被分配更高的权重以反映设备最新的性能退化状态，提出了泛化相似性分析技术，并成功应用于球栅阵列焊点处于振动环境下的剩余寿命预测。施建明等<sup>[84]</sup>提出了采用聚类分析技术获取航空发动机的健康指标，并基于所获得的一维健康指标实施相似性分析，对航空发动机的剩余寿命进行评估。胡小锋等<sup>[85]</sup>提出了考虑距离相似性和空间方向相似性的剩余寿命预测方法，并对铣槽机刀具的剩余寿命进行了预测。Simon 等<sup>[86]</sup>提出了一种改进型的基于退化轨迹相似性的剩余寿命预测方法，在所提出的模型中仅利用相似性匹配过程中得分最高的前几组进行计算运行产品的剩余寿命，并用所提方法预测了航空发动机的剩余寿命。Liao 等<sup>[87]</sup>提出了带预测正则项的受限玻尔兹曼机自动提取轴承的特征，并将所提取的特征输入到基于相似性分析框架完成对轴承的剩余

寿命预测。候晓东等<sup>[88]</sup>基于 T/R 组件的故障特点,有效识别出能反映组件退化过程的健康指标,并进一步提出了基于多指标相似性的 T/R 组件剩余寿命融合预测方法。梁泽明等<sup>[89]</sup>针对传统相似性剩余寿命预测方法中退化指标构建相对困难的问题,提出了一种多参数相似性信息融合的剩余寿命预测方法,所提方法被应用于预测航空发动机的剩余寿命。尤明懿<sup>[90-91]</sup>针对当前基于相似性的剩余寿命预测方法仅能利用运行至失效的历史样本的问题,提出了一种扩展的基于退化轨迹相似的剩余寿命预测方法,同时利用失效与未失效的历史样本以进一步提高设备的剩余寿命预测精度。

### 1.2.3 基于物理模型和数据驱动的混合剩余寿命预测方法研究

基于物理模型和数据驱动的混合剩余寿命预测方法尝试将基于模型的方法和基于数据驱动的方法进行有效的结合,利用这两种方法优势的同时抑制各自方法的劣势,从而实现产品剩余寿命的预测<sup>[92]</sup>。由于大多数复杂产品不能通过传感器测量技术直接在线获取其退化状态的数据,因此该类混合方法往往首先使用数据驱动方法处理所获得有关产品在线状态监测信息,进而对产品的退化状态做出间接的推断。然后,利用产品的失效物理模型对间接获取的有关产品的退化状态进行建模并估计其剩余寿命分布。

Robert 等<sup>[93]</sup>针对刀具的磨损预测问题提出了混合预测的方法。所提出的方法首先提取刀具振动信号和切削力信号的时域、频域和时频域的相关特征,然后采用核主成分分析技术对所提取的众多原始特征进行选择 and 融合,再使用支持向量机回归的方法对融合特征与刀具磨损状况的相关性进行建模并对刀具退化状态进行实时推断。最后,基于刀具磨损的失效物理模型采用粒子滤波技术实现刀具磨损的在线预测。雷亚国等<sup>[94]</sup>提出了针对滚动轴承剩余寿命的混合预测方法。首先提取轴承的振动信号的特征,其次引入相关向量机回归技术实现对所提取特征的相关向量的分析,并且利用指数退化模型来刻画轴承的退化过程并实现对相关向量的拟合分析。最后,通过外推轴承退化曲线至所设定的失效阈值,从而实现滚动轴承的剩余寿命分布的评估。雷亚国等<sup>[95]</sup>基于循环神经网络构建了融合多个经筛选后的轴承的时域、频域和时频域的特征为一合成的健康指标,并借助于粒子滤波技术实现了对风力涡轮发电机轴承的剩余寿命预测。Chang 等<sup>[96]</sup>针对锂离子电池剩余寿命预测提出了一种混合相关向量机和粒子滤波技术的方法。其中,相关向量机用于分析所获得的电池容量数据的测量不确定性,并采用双指数模型对锂离子电池的退化过程进行描述,最后基于粒子滤波技术实现了对锂离子电池退化模型参数的更新与剩余寿命的在线预测。此外,还存在一些尝试将两种不同的数据驱动方法

进行有机结合的混合剩余寿命预测方法。其中,第一个数据驱动方法用于对多传感器数据进行建模分析以得到产品的退化状态(此时假定产品的退化状态不能直接获取),另一个数据驱动方法被用于外推产品的退化状态以预测其剩余寿命。吴军等<sup>[97]</sup>提出了针对刀具剩余寿命的混合预测方法。首先利用集成经验模式分解技术对所获得的原始监测信号进行降噪处理。其次,进一步提取经过前处理的传感器信号的时域和频域特征。再后,通过对所提取特征的归一化、耦合分析、单调性分析等实现最优特征选择,进一步利用自适应模糊神经网络模型实现最优特征信息融合分析并获得刀具的退化状态。最后,基于多项式曲线拟合技术实现对刀具剩余寿命的预测。孙惠斌等<sup>[98]</sup>提出了针对刀具的剩余寿命的混合预测方法。基于经验模式分解技术,首先提取刀具的切削力信号、振动信号和声发射信号的有效特征。其次,利用传统的反向传播神经网络来实现对所提取的特征和刀具耗损状况的映射关系建模。最后,引入维纳过程模型以描述刀具的耗损过程并基于该模型对刀具剩余寿命分布进行估计。Liao 等<sup>[87]</sup>提出了带正则项的受限玻尔兹曼机自动提取轴承的振动信号的有效特征,并使用自组织映射方法融合所提取的特征并对轴承的退化状态进行评估,最后使用相似性分析的方法对轴承的剩余寿命进行预测。

上述有关混合剩余寿命预测方法研究中均假定已知产品退化模型的解析式,但针对某些产品失效物理模型解析式无法获取的情况,现有研究指出可以采用数据驱动方法所得结果代替产品的物理模型。Chen 等<sup>[99]</sup>提出了一种混合自适应神经模糊推论系统(ANFIS)和高阶粒子滤波的混合方法以预测直升机行星齿轮托板和轴承的剩余寿命。所提出的方法用产品的历史监测数据训练数据驱动算法 ANFIS 以对产品的退化过程进行建模,并采用高阶粒子滤波算法实施多步提前预测以外推产品的退化指标的演进过程,最终对其剩余寿命分布进行估计。此外也有研究指出,利用数据驱动算法以预测产品未来退化指标的值,并将其输入到基于产品物理模型的方法中,将有助于进一步提升混合方法的剩余寿命预测性能。Liu 等<sup>[80, 100]</sup>基于当前所采集到的锂离子电池的容量数据,提出采用基于数据驱动方法的预测算子,例如,神经网络、神经模糊网络和循环神经网络以预测未来的电池容量数据。并进一步将新预测的值输入到粒子滤波框架内以更新电池退化模型参数的分布并对其进行剩余寿命做出长期预测。所提出的混合剩余寿命方法的预测性能相比单独使用粒子滤波方法和单独使用数据驱动方法的性能更好。

针对混合剩余寿命预测方法研究中除了将数据驱动方法与基于物理模型方法进行串行连接的思路(即将数据驱动方法的输出当做基于物理模型方法的输入),还有并行组合的思路。该思路的核心思想在于同时采用数据驱动方法与模型方法对产品的剩余寿命分别进行预测,并设计一种适当的融合机制将两种方法的预测

结果进行有机结合。Goebel 等<sup>[101]</sup>提出了一种结合数据驱动算法预测结果和基于模型方法预测结果的混合方法,以提升轴承剩余寿命的预测精度且降低预测不确定性,并使用 Dempster-Shafer (DS) 回归方法融合两种预测结果。Goebel 等<sup>[102]</sup>进一步改进了他们之前的研究工作,用自适应神经模糊推论系统对轴承的在线监测数据进行退化评估,引入一种基于核函数回归的方法以融合两种方法的预测结果。此外,也存在一些尝试将两种不同的数据驱动方法所得的剩余寿命预测结果进行有机融合的研究。Youn 等<sup>[103]</sup>提出了一种混合数据驱动的剩余寿命预测方法以加权求和的方式对多个不同数据驱动算法输出的剩余寿命估计结果进行最终的融合。并进一步探究了三种不同权重分配机制,即基于准确度的权重分配、基于分散度的权重分配和基于优化的权重分配。最终将所提方法应用于航空发动机的剩余寿命预测,结果表明所提出的三种基于不同权重分配机制的混合方法在预测性能上均显著优于基于单个数据驱动的方法。Xu 等<sup>[104]</sup>同样针对航空发动机剩余寿命预测提出了一种混合方法。所提方法分别使用了三种不同的数据驱动算法,即 DS 回归方法、支持向量机方法和循环神经网络方法。最终基于信息熵理论实现了对上述三种不同数据驱动方法预测结果的有效融合,且融合结果显著好于单个方法的预测结果。

### 1.3 产品剩余寿命预测方法研究的发展趋势分析

通过以上综述可见,众多学者针对产品剩余寿命的不同类型预测方法,即基于模型的方法、基于数据驱动的方法和基于物理模型与数据驱动的混合方法均开展了广泛且深入的研究,并取得了大量的研究成果。通过对研究现状的深入分析,并结合工业大数据背景下产品状态监测数据所呈现出的显著特点,具体梳理出以下四点发展趋势。

#### (1) 基于产品退化轨迹相似性的剩余寿命预测方法研究。

针对相似设备群规模大且历史监测数据充足的情形,例如,通用电气公司大数据中心拥有超过千台风力涡轮机组的历史状态监测数据,基于产品退化轨迹相似性分析的方法可以对该类情况的产品剩余寿命给出精确的预测结果。但是,传统的基于退化轨迹相似性的方法仅能提供有关产品剩余寿命的点估计结果,而不能给出其剩余寿命的区间估计和对应的置信度水平,这将限制该方法对某些产品的工程应用。因此需要进一步结合其他技术以扩展基于产品退化轨迹相似性的方法,进而对产品的剩余寿命概率分布进行估计。

#### (2) 动态工况和多失效模式耦合下的产品剩余寿命预测方法研究。

在工业大数据背景下,产品的监测数据往往呈现出数据量大且数据类型丰富

的特点，因此不能仅依赖于专家经验或具备专业知识的领域内从业人员人工设计故障特征提取、选择、降维、多传感器融合等算法，需要利用最新的技术实现对产品监测大数据的自动实时解析。利用深度学习的理论与方法是解决这一关键问题的有效途径。然而现有的绝大多数基于深度学习的方法均不能有效地解决产品在动态工况和多失效模式耦合作用下的剩余寿命预测问题。因此对于动态工况和多失效模式耦合下的产品，基于深度学习模型的剩余寿命预测方法研究是一个重要的趋势。

### （3）基于深度学习模型和产品物理模型的混合剩余寿命预测方法研究。

基于数据驱动和物理模型的混合剩余寿命预测方法往往首先使用数据驱动方法处理所获得有关产品的在线状态监测数据（例如，振动数据和声发射数据等），并进一步对产品的退化状态做出间接的推断。然后，利用产品的失效物理模型对间接获取的有关产品的退化状态进行建模并估计其剩余寿命分布。然而，现有的大部分混合剩余寿命预测方法都涵盖了繁琐的处理步骤以获取监测数据与产品退化指标的非线性关系。包括但不限于以下步骤：特征提取、特征选择、特征降维及回归分析等，这导致了现有混合方法在工程应用中效率和智能化水平不高。为了进一步简化上述监测数据的分析流程，需要进一步研究结合深度学习方法与产品失效物理模型，以构建更加智能且更加精确的混合剩余寿命预测方法。

### （4）基于深度学习模型和 Bootstrap 的剩余寿命预测方法研究。

针对基于深度学习模型的剩余寿命预测方法，量化产品剩余寿命的预测区间为一难点问题。针对该问题，可以从融合深度学习模型和产品失效物理模型的角度切入以解决。但是，当针对某些复杂产品其失效过程往往存在多失效模式耦合的情况，难以准确地构建其对应的失效物理模型以有效地实施产品混合剩余寿命预测方法，所以这将大大限制该方法在复杂产品中的剩余寿命预测的工程应用。此外，Bootstrap 方法为一量化预测区间的经典方法且不需要有关产品的失效物理模型。因此，需要进一步研究整合深度学习模型到 Bootstrap 实施框架内的新方法，以在不需产品失效模型的前提下量化其剩余寿命预测区间。

## 1.4 论文结构框架和主要研究内容

本文共分为六章，其结构框架如图 1-1 所示。总体分为两大部分：第一部分为传统的数据驱动剩余寿命预测方法，具体为第二章；第二部分为基于深度学习模型的剩余寿命预测方法，包括第三章、第四章和第五章。本文主要针对产品在监测大数据背景下所呈现的不同特点展开剩余寿命预测方法研究。

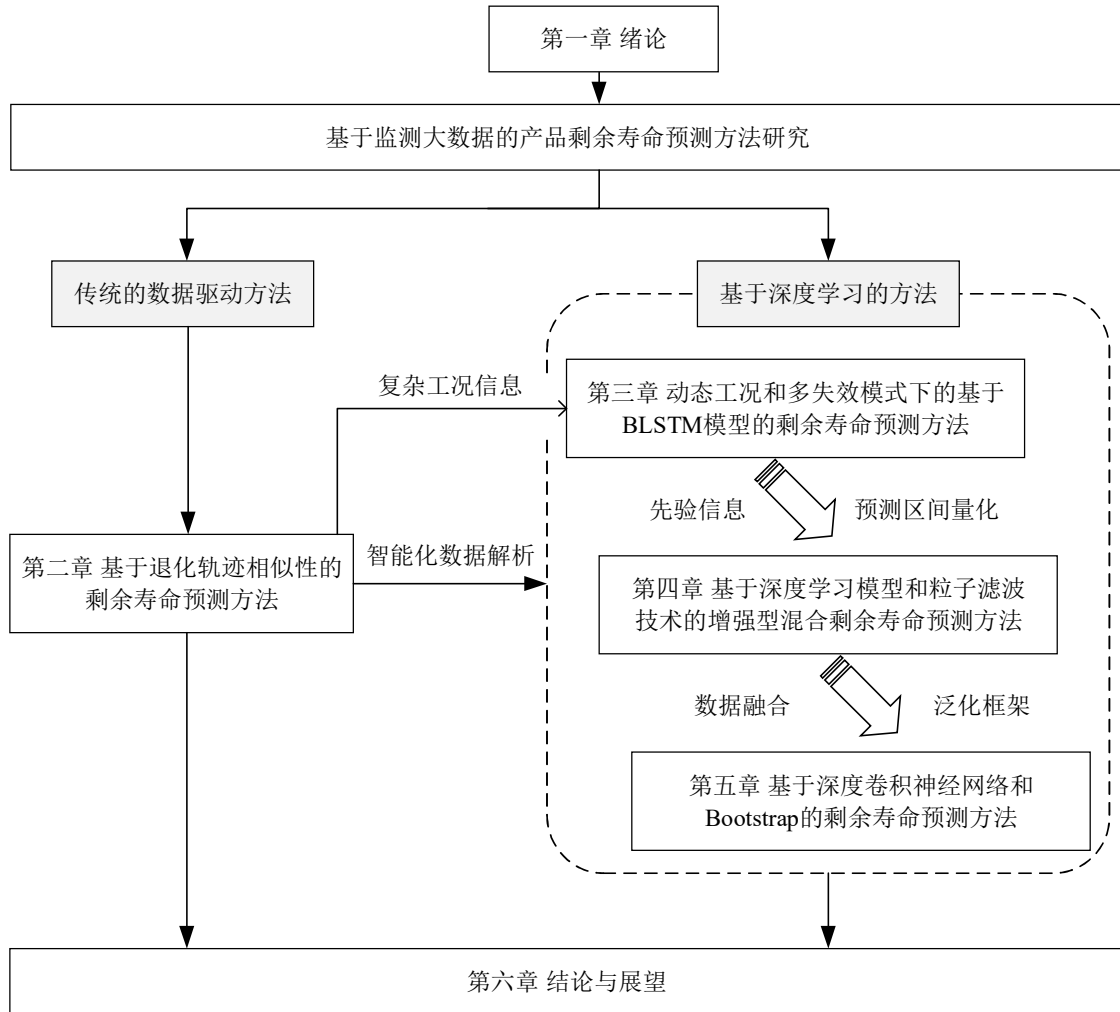


图 1-1 论文结构框架图

各章主要研究内容包括：

第一章首先对论文的研究背景、意义进行介绍，其次对产品剩余寿命预测方法进行广泛的文献调研，进一步总结各类方法的特点并指出产品在监测大数据背景下的剩余寿命预测方法的发展趋势，最后简要概述本文的主要研究内容。

第二章研究基于产品退化轨迹相似性的剩余寿命预测方法。对于存在海量历史监测数据的相似性产品提出一种改进的基于退化轨迹相似性的剩余寿命预测方法，突破该类传统方法仅能对产品剩余寿命进行点估计的应用限制。基于核密度估计技术和相似性分析方法，完成对产品剩余寿命概率密度的近似估计，同时结合  $\beta$  准则给出产品在某一置信度下剩余寿命的区间估计。此外，依据产品在实际工程中的预测需求，进一步提出更为精确的基于退化轨迹相似性的剩余寿命点估计方法。

第三章研究基于深度学习模型的剩余寿命预测方法。针对工作在动态工况环境和多失效模式耦合作用下的产品提出一种基于深度学习模型的剩余寿命预测方

法。所提出的剩余寿命预测方法将有效地融合多传感器监测数据与动态工况数据。通过使用深度学习模型中的双向长短期记忆神经网络，完成对融合数据的时间相关性建模。基于所提出的深度学习模型极强的非线性建模能力和泛化能力，进而实现对多失效模式耦合作用下的融合数据与产品剩余寿命的非线性回归分析。为了验证所提方法的适用性和准确性，拟对某航空发动机分别工作在多工况单失效模式和多工况多失效模式下的剩余寿命进行预测，并将预测结果与其他最新的基于数据驱动方法的预测结果进行对比。

第四章研究基于深度学习模型和基于粒子滤波技术的混合剩余寿命预测方法。考虑某些产品关键失效模式已知、失效机理已被阐明、且失效模型已经建立，本文拟结合传感器在高采样频率下所采集的监测数据与产品的失效物理模型，构建基于深度学习模型和粒子滤波技术的混合剩余寿命预测方法。基于双向长短期记忆神经网络对时间序列数据建模的优势，实现对监测数据前处理信息的深度特征提取、选择和融合分析，并进一步结合前馈神经网络回归分析的能力识别出产品的退化指标。此外，构建基于粒子滤波技术的剩余寿命预测框架，有效地结合产品失效物理模型与识别的退化指标数据，进一步给出产品剩余寿命的概率密度估计。为了验证所提方法的有效性，拟对某数控机床刀具在加工过程中的剩余寿命进行预测。

第五章研究基于深度卷积神经网络和 Bootstrap 的剩余寿命预测方法。针对某些复杂产品，其失效模式未知、失效机理难以阐明且失效模型构建困难，本文拟结合深度卷积神经网络和 Bootstrap 实施框架，构建不依赖于产品失效物理模型的更加泛化的剩余寿命预测方法。为了验证所提方法的有效性，拟对某滚动轴承的剩余寿命进行预测。

第六章对全文进行总结，并对未来的研究工作进行展望。

## 第二章 基于退化轨迹相似性的剩余寿命预测改进方法

### 2.1 引言

故障预测与健康管理技术是有效实现产品预防性维修(Predictive Maintenance)的基础。其中产品的故障预测(RUL 预测)是 PHM 技术中最具挑战性的核心组成部分之一。RUL 预测结果的准确、可靠且稳健是 PHM 技术得以实施的先决条件。第一章已对现有的产品故障预测方法进行了综述,本章主要聚焦在基于退化轨迹相似性分析的 RUL 预测方法(Trajectory Similarity-Based Prognostic, TSBP)。在基于 TSBP 方法的框架中,在役产品的 RUL 可通过对已失效相似产品在某一时刻 RUL 的加权平均进行计算,其中权重值可通过在役产品与失效产品之间的相似性进行计算,而相似性可进一步通过在役产品与失效产品之间退化趋势或状态监测数据进行相似性匹配而确定。该方法对存在海量失效产品状态监测数据的情形具有很强的泛化能力,尤其适用于工业大数据背景下的产品 RUL 预测。

在当前基于 TSBP 方法研究中,针对产品健康指标的合成<sup>[84, 87, 89-90]</sup>和相似性测度函数的构建<sup>[83,85-86]</sup>两方面开展了大量的研究工作,但目前绝大多数基于 TSBP 方法<sup>[83-90, 105-106]</sup>仅能给出 RUL 的点估计结果。另外,在产品 RUL 预测的实际过程中往往存在各种不确定性,例如,测量不确定性、工况不确定性、建模不确定性和输入数据的不确定性等。因此,一个能有效适用于工程实际的 RUL 预测方法,尤其是针对一些安全性要求较高的产品<sup>[107]</sup>,不仅要能准确地给出 RUL 的点估计,而且还必须能给出某一置信度下 RUL 的区间估计。本章基于核密度估计(Kernel Density Estimation, KDE)技术和 $\beta$ 准则提出一种改进的 TSBP 方法,以进一步给出更为准确的产品 RUL 点估计和某一置信度下的区间估计。

### 2.2 传统的基于相似性的剩余寿命预测方法

#### 2.2.1 方法假设与实施流程

为便于讨论,做以下假设:

(1) 测试部件指在役产品,且其健康状态正被  $N$  个传感器实时监测。因此,其实际的 RUL 在任何一个监测时刻点都是未知的,需要通过实时处理其相应监测信号以对其 RUL 做出实时推断。相反,训练部件指已经失效的产品,它的健康状态在寿命周期内被同样的  $N$  个传感器所记录且保存。因此训练部件的状态监测信息和对应的 RUL 在其完整的监测周期内都是已知的。

(2) 测试部件和训练部件来自于同一批次产品，且工况环境大致相同，可将测试部件与训练部件看做相似产品。因此，测试部件和训练部件在失效发生之前都具有相似的退化行为，且这一退化行为可被多传感器信号所刻画。

(3) 测试部件和训练部件的退化过程被认为是连续的，且在这一过程中不考虑部件维修、替换等行为。

基于部件退化轨迹相似性的 RUL 预测方法实施流程如图 2-1 所示。该实施流程主要包括两阶段：离线建模阶段（2.2.2 节中展开说明）和在线 RUL 预测阶段（2.2.3 节中展开说明），其中离线阶段主要涵盖两个部分。首先，对训练部件多传感器运行至失效（Run to Failure, RtF）的历史监测数据进行数据前处理，包含数据采集，数据归一化，训练部件健康指标提取，和训练部件退化轨迹建模等具体实施步骤；其次，经过数据前处理流程，多传感器数据可转化为健康指标时间序列数据（通常为一维合成健康指标）。通过合适的退化轨迹建模方法进一步提取对应的退化轨迹，记录所有训练部件的退化轨迹并保存于训练部件退化模型库，方便在线 RUL 预测阶段的调用。此外，在线阶段则主要涵盖三个部分。首先，类似于离线阶段，测试部件的实时多传感器在线监测数据同样经过数据前处理；其次，将所获得的测试部件的退化轨迹输入到退化轨迹相似性分析模块中，并结合训练部件的退化模型库展开相似性分析；最后，将上述相似性分析结果进一步输入到 RUL 预测模块中，对测试部件的实时 RUL 进行在线预测。

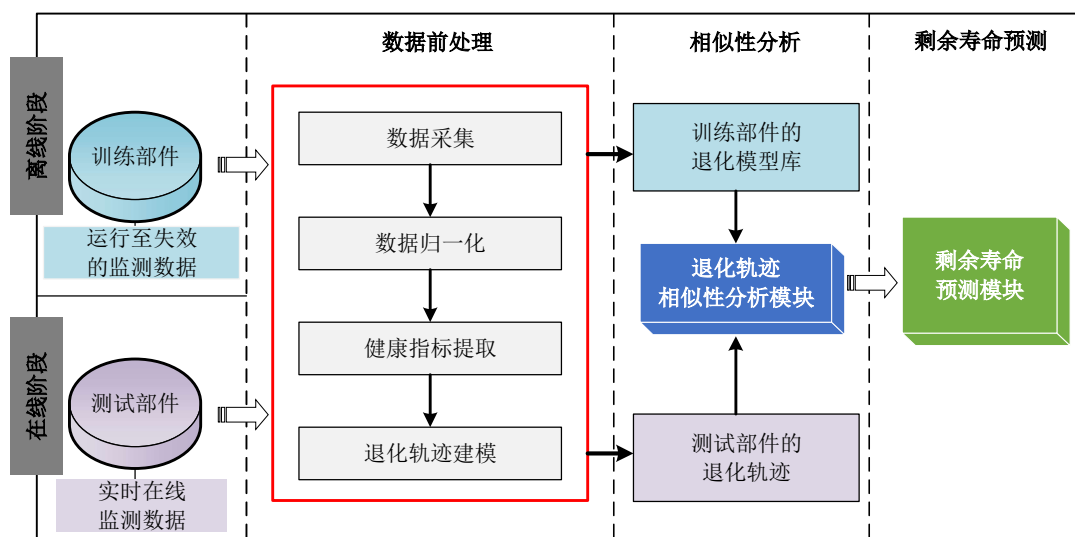


图 2-1 基于 TSBP 方法的实施框架图

## 2.2.2 数据前处理

为了有效地监测产品的健康状态，往往需要在产品的不同部位部署多个传感器以收集相关数据。例如，美国国家航空航天局（National Aeronautics and Space

Administration, NASA) 所发布的涡扇航空发动机退化仿真数据<sup>[108]</sup>中涵盖了 21 个传感器收集到的相关信号, 包括风扇入口总温度、低压压气机出口总温度、低压涡轮出口总温度、风扇物理转速、高压涡轮冷却引气流量等。不同传感器所监测的信号往往具备完全不同的尺度范围, 如图 2-2 步骤一所示。因此, 为了便于后续健康指标提取和退化轨迹建模, 首先需将原始的多传感器数据进行归一化处理。常用的数据归一化方法主要有两种: z-score 归一化和 max-min 归一化, 分别用式(2-1)和(2-2)表示。

$$\mathbf{x}_i' = \frac{\mathbf{x}_i - \mu_i}{\sigma_i} \quad (2-1)$$

式中,  $\mathbf{x}_i$  表示第  $i$  个传感器的原始数据,  $\mathbf{X} = [\mathbf{x}_i]_{i=1}^N$  为一矩阵表示全部  $N$  个传感器的监测数据,  $N$  为所部署的传感器数量,  $\mathbf{x}_i'$  为第  $i$  个传感器的归一化数据,  $\mu_i$  和  $\sigma_i$  分别代表第  $i$  个传感器在监测周期内的均值和方差。值得注意的是  $\mu_i$  和  $\sigma_i$  是将所有训练部件的第  $i$  路传感器数据当做总体所计算出的统计均值和方差。

$$\mathbf{x}_i' = \frac{\mathbf{x}_i - x_i^{\min}}{x_i^{\max} - x_i^{\min}} \quad (2-2)$$

其中,  $x_i^{\max}$  和  $x_i^{\min}$  分别代表第  $i$  个传感器在监测周期内的最大值和最小值。同样的  $x_i^{\max}$  和  $x_i^{\min}$  是通过对所有训练部件的第  $i$  路传感器数据为总体所计算出的最大值和最小值。

经过数据归一化处理后, 多传感器数据被转化到归一化尺度范围内。随后, 归一化的多传感器数据再次经过步骤三(健康指标提取)的处理, 将  $N$  维传感器信号转化为一维合成健康指标, 该合成健康指标即可明确地刻画产品健康状态的演变情况。值得注意的是: 在产品剩余寿命预测研究领域, 退化指标的获取一共存在三种途径: (1) 通过直接测量的方式进行获取, 例如轮盘裂纹扩展的长度<sup>[109]</sup>, 叶片的蠕变长度<sup>[16]</sup>和锂电池的容量退化<sup>[109]</sup>; (2) 针对原始监测数据在时域观测范围内没有显著退化的产品, 需提取其时域、频域和时频域内的多个有效特征<sup>[95]</sup>, 并进一步融合这些特征从而得到一合成的健康指标, 例如轴承、齿轮和数控机床的刀具等; (3) 针对原始多传感器数据已经存在明显退化特征的产品, 直接在此基础上融合多传感器数据并获得一合成的健康指标<sup>[81, 110]</sup>。本章主要处理的是第三种情况。

目前关于多传感器数据融合以获取其健康指标的研究较多, 主要涵盖了线性回归<sup>[81, 86]</sup>、人工神经网络<sup>[87, 89]</sup>和聚类分析<sup>[84]</sup>等方法。由于多元线性回归(Multiple Linear Regression, MLR)能有效地融合多传感器数据且最大限度地保留其原有的退化模式<sup>[81]</sup>, 所以本章采用 MLR 方法融合归一化的多传感器数据并提取其对应的

退化指标。MLR 模型表示为：

$$y = \beta_0 + \sum_{i=1}^N \beta_i x'_i + \varepsilon \quad (2-3)$$

式中， $\beta_0, \beta_1, \dots, \beta_N$  表示 MLR 的模型参数， $\varepsilon$  为模型残差项， $y$  为合成健康指标。为获取 MLR 模型中的  $N+1$  个模型参数值，需要  $K$  ( $K > N$ ) 个观测值，基于最小二乘回归法，可拟合得到一  $(N+1)$  维平面最小化  $K$  个观测值的残差平方和。

$$\sum_{j=1}^K e_j^2 = \sum_{j=1}^K \left( y_j - \beta_0 - \sum_{i=1}^N \beta_i x'_{ji} \right)^2 \quad (2-4)$$

式(2-4)中， $x'_{ji}$  表示第  $j$  个观测值中的第  $i$  个传感器归一化数据， $y_j$  表示  $x'_{ji}$  所对应的健康指标值。在产品实际 RUL 预测中，可依据工程经验得到这  $K$  组观测值  $(x'_{ji}, y_j) (j=1, 2, \dots, K)$ ，例如可假定所有训练部件的最开始三个和最后两个监测时刻点的传感器数据分别对应了产品接近健康状态和接近失效状态（即  $y_j \approx 1$  和  $y_j \approx 0$ ）<sup>[81]</sup>。最后，求解式(2-4)的最小化问题即可得到 MLR 的模型参数，并保存所求得的 MLR 模型参数组以便在线阶段的调用。经步骤三处理后，所有训练部件的  $N$  维传感器信号可转化为一合成的一维健康指标，且该健康指标可以有效地刻画部件的健康状态。

由于所提取的合成健康指标时间序列在监测周期内存在较大的波动，因此，需要进一步对其进行处理以获得平滑的退化轨迹，方便后续的相似性分析。将上述过程定义为部件的退化轨迹建模，具体的建模方法主要分为参数化建模和非参数化建模两大类。首先，针对某些已知其退化模式的产品，参数化建模方法可以有效地对其合成健康指标进行建模并得到其合理的退化轨迹。例如，针对 NASA 所发布的涡扇航空发动机退化仿真数据<sup>[108]</sup>，指数退化模型<sup>[111]</sup>被用来刻画发动机故障演化行为且其具体模型为：

$$y = 1 - d - \lambda_1 \exp(\lambda_2 t + \lambda_3) \quad (2-5)$$

式中， $y$  表示训练部件平滑处理后的退化轨迹； $d > 0$  为初始损伤且与制造偏差和初始耗损有关； $\lambda_1, \lambda_2, \lambda_3$  为模型参数，可以通过求解一非线性最小二乘曲线拟合得到。如图 2-2 步骤四所示，黑色曲线表示训练部件平滑处理后的退化轨迹。最后，记录所有训练部件的退化轨迹并保存于训练部件退化模型库，方便在线 RUL 预测阶段的调用。此外，针对某些退化模式未知的产品，非参数化方法是更为有效的对产品合成退化指标时间序列进行平滑处理以获取其退化轨迹的手段。常用的非参数建模方法包括核回归<sup>[112]</sup>、局部加权回归散点平滑法<sup>[113]</sup>、相关向量机等方法。由于上述方法均具备成熟的理论框架且在大多数商用软件上已嵌入了对应的算法实

现，此处不再赘述。

本小节内容涵盖了离线建模阶段的主要四个步骤，通过对训练部件逐步实施包括数据采集、数据归一化、健康指标提取、退化轨迹建模等步骤，实现了对训练部件多传感器数据的处理且构建了训练部件退化模型库。接下来的 2.2.3 小节主要涵盖了在线 RUL 预测阶段的主要技术内容。

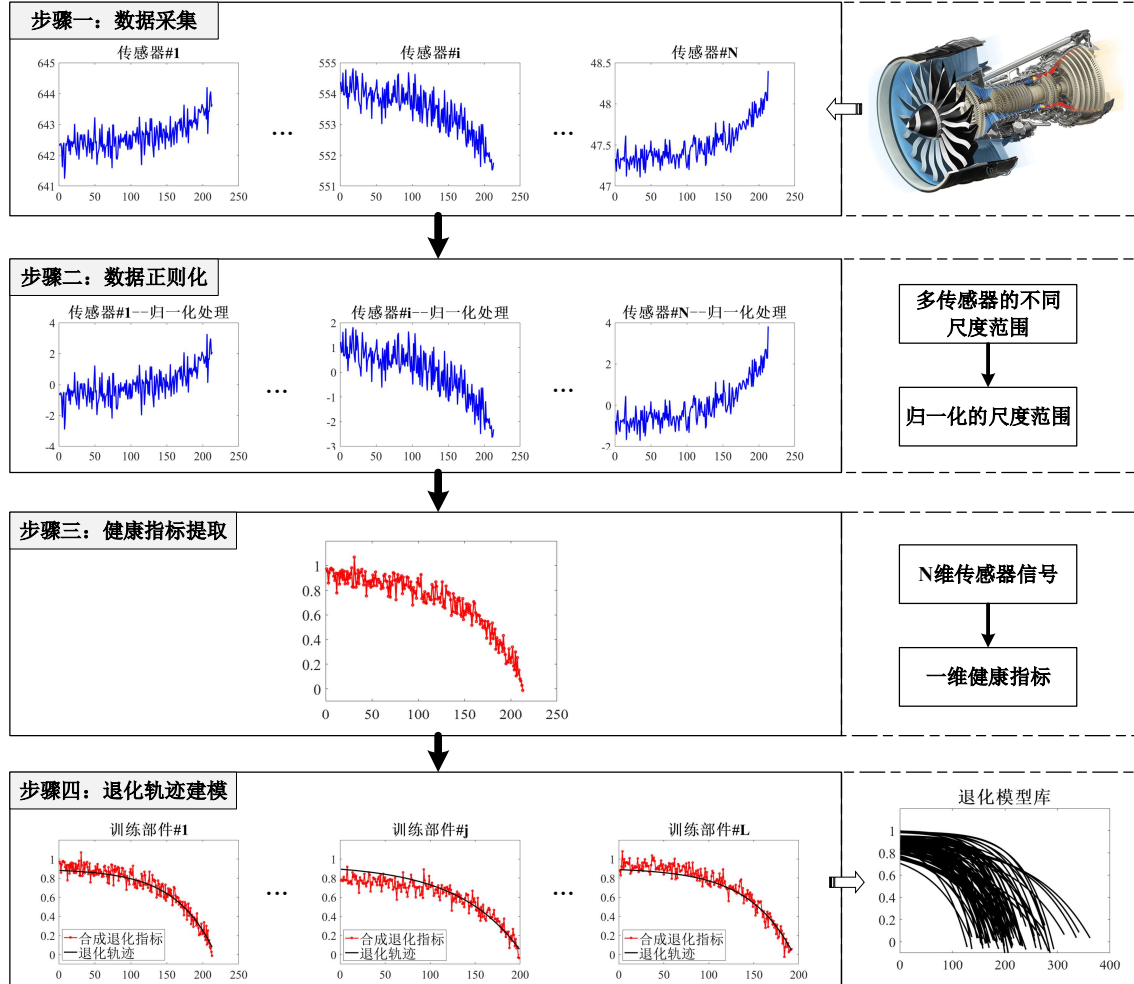


图 2-2 基于 TSBP 方法的数据前处理流程图

### 2.2.3 退化轨迹相似性分析与剩余寿命预测

针对在线 RUL 预测阶段，测试部件的多传感器数据同样需要经过上述数据前处理中的四个步骤从而获得测试部件的实时在线退化轨迹。基于离线阶段所获取并保存的 MLR 模型的参数  $\beta_0, \beta_1, \dots, \beta_N$ ，测试部件的健康指标时间序列计算如下：

$$s = \beta_0 + \sum_{i=1}^N \beta_i \sigma_i + \varepsilon \quad (2-6)$$

式中， $\sigma_i$  为测试部件归一化后的第  $i$  路传感器数据，值得注意的是在针对测试部件

归一化处理的过程中所需用到的第  $i$  路传感器的均值和方差数值（z-score 归一化方法）或最大值和最小值（max-min 归一化方法）都需要以所有训练部件的第  $i$  路传感器数据为总体进行计算。 $s$  为所计算得到的测试部件的健康指标时间序列，且由于所提取的  $s$  在局部范围存在较大波动，因此采用 LOWESS 方法对其进行平滑处理以获取其退化轨迹，为了表示的简洁性同样采用  $s$  表示经过平滑处理后的测试部件的退化轨迹，如图 2-3 所示。

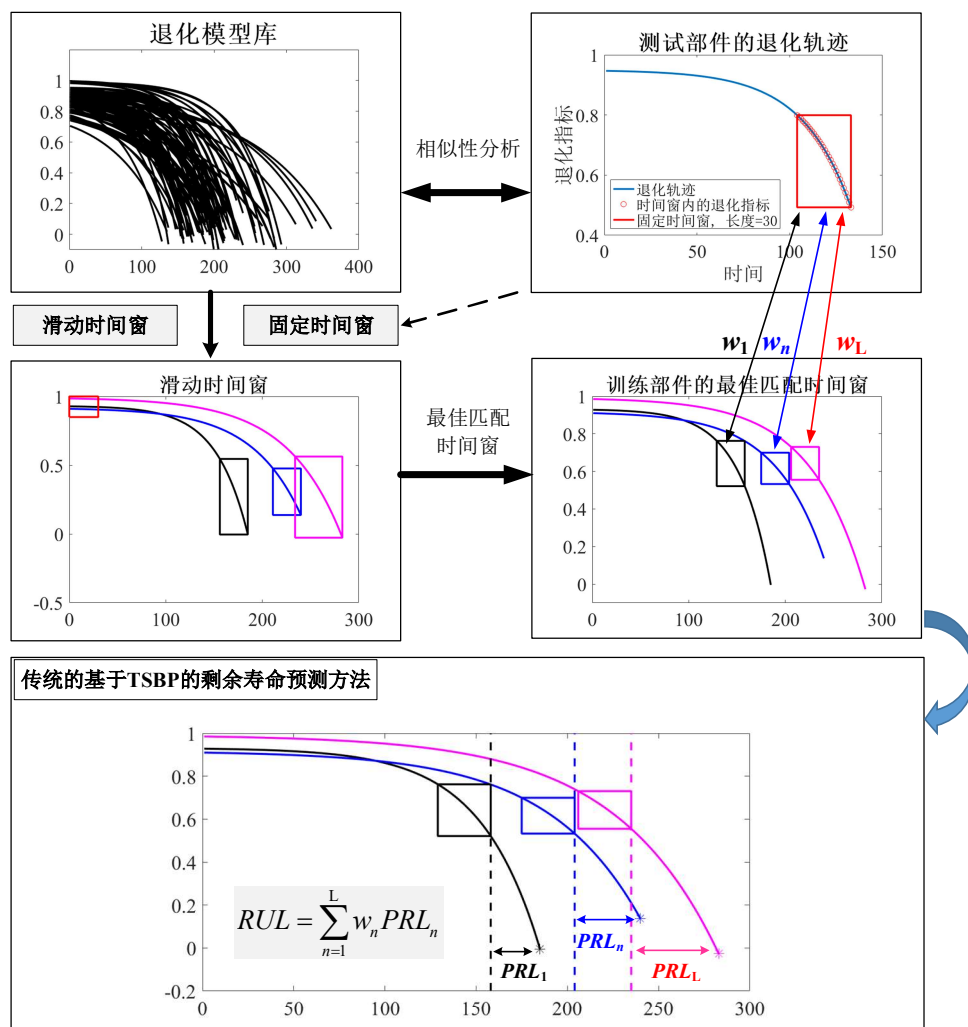


图 2-3 退化轨迹相似性分析与 RUL 预测

为了进行相似性分析和测试部件的 RUL 预测，共计需要实施以下四个步骤。第一步需要确定测试部件的固定时间窗（Fixed Time Window, FTW）。FTW 指最近的连续退化指标时间序列，图 2-3 中右上角红色方框即表示测试部件的一固定时间窗，其长度假定为 30。测试部件的 FTW 可表示为：

$$s^m = [s_1^m, \dots, s_i^m, \dots, s_{\Delta t}^m] \quad (2-7)$$

式中,  $s^m$  表示第  $m$  个测试部件的固定时间窗,  $s_i^m$  表示其在第  $i$  个监测时刻点的退化指标,  $\Delta t$  为时间窗的长度。值得注意的是  $\Delta t$  是一个影响 TSBP 方法性能的关键参数, 该参数的取值将在 2.4 节的算例分析中进行详细的讨论。

在第二个步骤中, 需要进一步定义相似性测度函数以量化测试部件与各个训练部件之间退化轨迹的相似度。欧几里得距离函数是基于 TSBP 方法所常用的相似性测度函数, 本部分以欧氏距离函数为例, 相似度具体计算为:

$$d_n^m(j) = \sqrt{\sum_{i=1}^{\Delta t} (s_i^m - y_{i+j-1}^n)^2} \quad (2-8)$$

式中,  $s_i^m$  表示第  $m$  个测试部件在第  $i$  个监测时刻点的退化指标, 且取值为  $s^m = [s_1^m, \dots, s_i^m, \dots, s_{\Delta t}^m]$  即式(2-7)中所定义的 FTW。  $y_{i+j-1}^n$  表示第  $n$  个训练部件连续变化的退化指标且随着指标  $i$  的变化取值为  $[y_j^n, y_{j+1}^n, \dots, y_{j+\Delta t-1}^n]$ , 在这里参数  $j$  的取值范围为  $[1, l(y^n) - \Delta t + 1]$  且表示控制训练部件滑动时间窗的参数, 滑动时间窗的滑动步长为 1,  $l(y^n)$  表示第  $n$  个训练部件退化轨迹的长度。因此,  $[y_j^n, y_{j+1}^n, \dots, y_{j+\Delta t-1}^n]$  可以进一步表示为:

$$[y_1^n, y_2^n, \dots, y_{\Delta t}^n] \rightarrow [y_{l(y^n) - \Delta t + 1}^n, \dots, y_{l(y^n)}^n] \quad (2-9)$$

式中,  $[y_1^n, y_2^n, \dots, y_{\Delta t}^n]$  表示第  $n$  个训练部件的第一个时间窗, 而  $[y_{l(y^n) - \Delta t + 1}^n, \dots, y_{l(y^n)}^n]$  表示第  $n$  个训练部件最后一个时间窗。如图 2-3 滑动时间窗 (Sling Time Window, STW) 的示例图所示, 三种不同颜色的曲线表示三个不同的训练部件的退化轨迹, 左上角的红色方框即统一代表第一个时间窗, 右下角的不同颜色的方框即表示对应不同训练部件的最后一个时间窗。由于时间窗在各自训练部件对应的退化轨迹上以步长为 1 进行滑动, 因此在这里将其定义为滑动时间窗的概念。

值得注意的是测试部件的固定时间窗 (FTW) 与训练部件的滑动时间窗 (STW) 之间的相似度以步长为 1 每次均进行计算, 该过程通过参数  $j$  进行控制。因此, 第  $m$  个测试部件的 FTW 与第  $n$  个训练部件在整个监测周期的 STW 之间的相似度向量  $d_n^m(j)$  可表示为:

$$d_n^m(j) = [d_n^m(1), d_n^m(2), \dots, d_n^m(l(y^n) - \Delta t + 1)] \quad (2-10)$$

通过式(2-10)所计算出的相似度向量, 可以进一步得到其中的最小相似度的值, 由于欧氏距离越小即表明两者之间越相似。因此, 该最小值即可最终量化训练部件与测试部件之间退化轨迹的相似性。此外, 该最小值所对应的参数  $j$  的取值 (即为训练部件的第  $j$  个 TW) 也就对应了最佳匹配时间窗的位置。最小相似度的值  $d_n^m$

可计算如下：

$$d_n^m = \min_{1 \leq j \leq l(y^n) - \Delta t + 1} d_n^m(j) \quad (2-11)$$

在第三个步骤中，为了计算测试部件的 RUL，需要进一步基于所获得  $d_n^m$  定义权重函数  $w_n^m$ 。定义  $z_n^m$  为：

$$z_n^m = \frac{\sum_{n=1}^L d_n^m}{d_n^m} \quad (2-12)$$

式中， $L$  表示训练部件的总数量，之后权重函数  $w_n^m$  可计算为：

$$w_n^m = \frac{z_n^m}{\sum_{n=1}^L z_n^m} \quad (2-13)$$

基于表达式(2-11)中所获得的最佳匹配时间窗的位置，测试部件的 RUL 预测值（Predicted Remaining Life, PRL）为：

$$PRL_n^m = EoL_n - \left[ \left( \arg \min_{1 \leq j \leq l(y^n) - \Delta t + 1} d_n^m(j) \right) + \Delta t - 1 \right] \quad (2-14)$$

式中， $EoL_n$  表示第  $n$  个训练部件使用寿命终止时刻（End of Life, EoL）， $\arg \min_{1 \leq j \leq l(y^n) - \Delta t + 1} d_n^m(j) + \Delta t - 1$  表示第  $n$  个训练部件退化轨迹上用于计算 PRL 所对应的

时刻点。如图 2-3 下方图例（传统的基于 TSBP 的 RUL 预测方法）所示，不同颜色的虚线即代表针对不同训练部件用于计算 PRL 所对应的时刻点。最后考虑全体  $L$  个训练部件的退化轨迹并基于其对应的权重函数  $w_n^m$ ，第  $m$  个测试部件的 RUL 为：

$$RUL^m = \sum_{n=1}^L w_n^m PRL_n^m \quad (2-15)$$

### 2.3 基于核密度估计的相似性剩余寿命预测改进方法

传统的基于 TSBP 方法仅能以加权和的形式给出测试部件 RUL 的点估计值。但是，在产品 RUL 预测的实际过程中往往存在各种不确定性，所以一个能有效适用于工程实际的 RUL 预测算法不仅能准确的给出 RUL 的点估计结果，而且还必须能给出其在某一置信度下的 RUL 的区间估计。本部分基于核密度估计技术提出了一种 TSBP 的改进方法以进一步给出更为准确的 RUL 点估计结果和对应的在某一置信度下的区间估计结果。

### 2.3.1 改进 TSBP 方法的实施框架与数据预处理

在相似性分析模块中，对测试部件的退化轨迹与退化模型库中的各个训练部件的退化轨迹进行相似性分析，并基于各个训练部件已知的  $EoL_i$  可进一步推算出测试部件的 RUL 预测值。记  $PRL_n^j$  为第  $j$  个测试部件根据与第  $n$  个训练部件的退化轨迹相似性分析所得出的 RUL 预测值结果，且其对应的权重值记为  $w_n^j$ 。在这里统一用  $(w_n^j, PRL_n^j)$  表示第  $j$  个测试部件根据与第  $n$  个训练部件的退化轨迹在相似性分析模块中所得到的输出结果。考虑整个退化模型库，进一步用  $(w^j, PRL^j) = \{(w_1^j, PRL_1^j), \dots, (w_n^j, PRL_n^j), \dots, (w_L^j, PRL_L^j)\}$  表示第  $j$  个测试部件与全体训练部件经过相似性分析模块所得到的结果，其中  $L$  表示所有训练部件的数量。

传统的 TSBP 方法和改进的 TSBP 方法针对产品 RUL 预测部分的实施框架如图 2-4 所示。测试部件经过相似性分析模块所得的结果  $(w^j, PRL^j)$  需先进行数据预处理，并记  $(w_*^j, PRL_*^j)$  为经过数据预处理模块所处理后的结果。针对传统的 TSBP 方法，其测试部件的 RUL 可通过加权平均的方法（式 2-15）进行求解，但是传统的方法仅能给出测试部件 RUL 的点估计结果。针对传统的 TSBP 方法不能给出在某一置信度下的 RUL 的区间估计的问题，本节提出一种改进的 TSBP 方法以解决上述实际工程问题。具体地，基于 KDE 技术，用  $(w_*^j, PRL_*^j)$  估计测试部件剩余寿命预测值（PRLs）的概率分布。最后基于所得的 PRLs 的概率分布，可以得到测试部件更为精确的 RUL 点估计结果，并结合  $\beta$ -准则<sup>[107]</sup>进一步给出测试部件在某一置信度水平下的 RUL 区间估计。

测试部件经过相似性分析模块所得的结果为  $(w^j, PRL^j) = \{(w_n^j, PRL_n^j)\}_{n=1}^L$ ，但是并不是所有的训练部件的 PRLs 都被用来计算测试部件最终的 RUL，这主要是因为某些训练部件与测试部件的退化轨迹相差较大，并不具备参考的价值。值得注意的是，某训练部件对应的权重值越高即表明该训练部件与测试部件之间的退化轨迹越相似，所以用该训练部件根据式(2-14)所推导得到的剩余寿命估计值（PRL）也就越精确。因此，仅选取根据权重值排序结果的前  $M$  个训练部件的结果来对测试部件 RUL 进行点估计和区间估计，其中  $M$  为所提出的基于 TSBP 的改进方法的另一个关键参数，它的具体取值将在 2.4 节算例分析部分进行讨论。

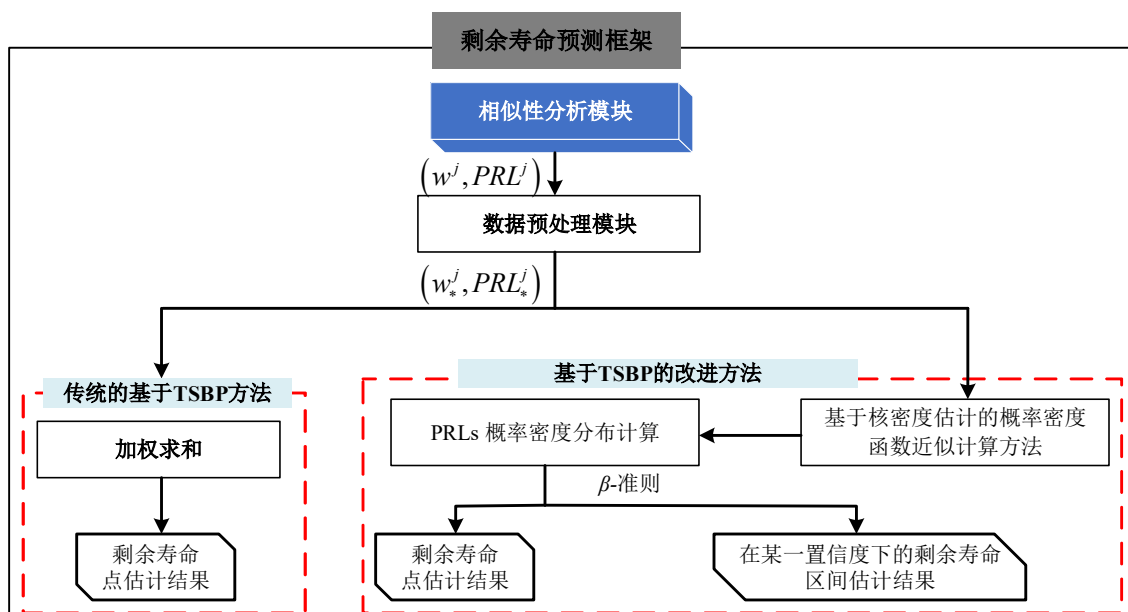


图 2-4 基于 TSBP 改进方法的 RUL 预测框架

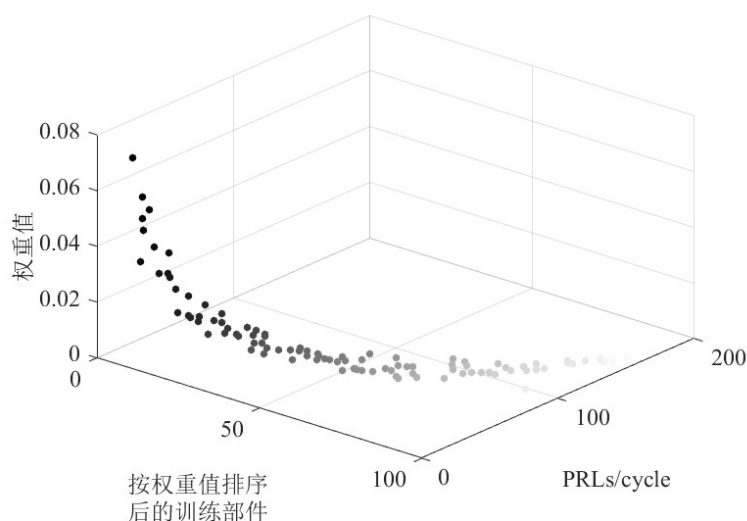


图 2-5 按权重值排序后的  $(w^j, PRL^j)$  三维示例图

图 2-5 给出了某航空发动机 RUL 预测的案列，其中共计有 100 个训练发动机，并对某一测试发动机的 RUL 进行预测，该测试发动机的实际 RUL 为 28 个 cycle。图中直接给出了基于测试发动机与各个训练发动机退化轨迹相似性分析结果  $(w^j, PRL^j)$  的三维示例图，并已重新根据其权重值进行了排序。在图 2-5 中，实心圆点代表排序后的训练部件在某一权重值下对应的 PRL 值，且颜色越深代表其所占的权重越大。图 2-6 具体地给出了上述图 2-5 在去掉权重值后的示例图，从图中可以看出权重值越大所对应的 PRLs 的值越接近测试部件的实际 RUL，相反权重值越小所对应的那些 PRLs 逐渐偏离测试部件实际的 RUL。黑色方框内为权重值

排序前 30 个训练部件的 PRLs, 从图中可知黑色方框内的 PRLs 基本接近测试部件实际 RUL。因此, 训练部件所对应的权重值一方面可以表明该训练部件与测试部件退化轨迹之间的相似性, 同时另一方面也可以表明基于该训练部件所得的 PRL 值的重要程度。

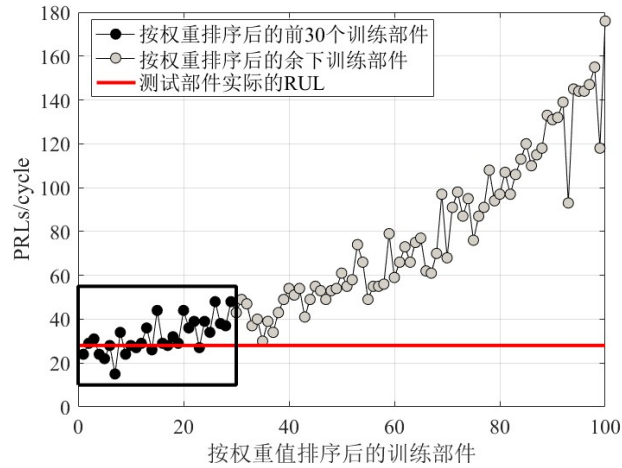


图 2-6 按权重值排序后的  $(w^j, PRL^j)$  二维示例图

虽然根据权重值排序结果的前  $M$  个训练部件所对应 PRLs 与测试部件的实际 RUL 比较接近且大致分布在实际 RUL 的两侧, 但是在某些特定情况下例如测试部件的监测周期较短且没有展示出明显的退化趋势, 所得的前  $M$  个训练部件所对应 PRLs 存在很强的波动性。因此, 需要进一步对  $\{(w_n^j, PRL_n^j)\}_{n=1}^M$  进行异常值剔除, 此处假设  $\{PRL_n^j\}_{n=1}^M$  为一组来自独立同分布的样本且服从正态分布, 则可以将那些超过  $[\mu - 2\delta, \mu + 2\delta]$  的范围的样本视为异常值, 由于这些异常值与 PRLs 的均值存在明显的偏差, 因此对其进行剔除, 图 2-7 为 PRLs 异常值剔除的示例图。

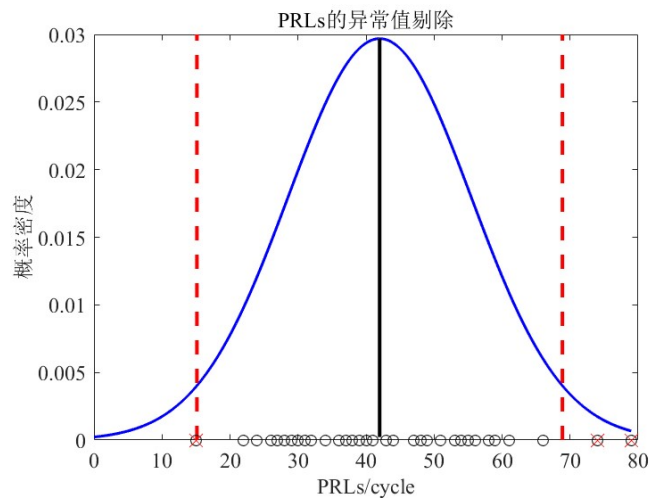


图 2-7 PRLs 的异常值剔除示例图

通过实施上述两个步骤，可以得到预处理之后的结果并记为 $(w_*^j, PRL_*^j)$ 。基于 KDE 技术，用 $(w_*^j, PRL_*^j)$ 估计测试部件剩余寿命预测值 (PRLs) 的概率密度分布。

### 2.3.2 基于核密度估计算法的剩余寿命概率密度函数近似

本章选择 KDE 技术主要是针对两方面情况。首先，某些产品的 RUL 难以通过合适的概率密度函数的解析解表示；其次，当存在多失效模式耦合的情况，其 RUL 的概率密度分布可能由多个不同的概率密度分布组合而成。基于非参数方法的 KDE 技术可以很好地解决上述两种情况，因此 KDE 技术被用来近似测试部件的 RUL 概率密度函数。

假设  $x = (x_1, x_2, \dots, x_n)$  是独立同分布的随机变量，且来源于某一未知的概率密度函数  $f(x)$ 。为了有效地估计  $f(x)$ ，定义 KDE 算子  $\hat{f}_h(x)$ ：

$$\hat{f}_h(x) = \frac{1}{n} \sum_{i=1}^n K_h(x - x_i) = \frac{1}{nh} \sum_{i=1}^n K\left(\frac{x - x_i}{h}\right) \quad (2-16)$$

式中， $K(\bullet)$  为核函数， $h$  为平滑参数，又称为带宽 (Bandwidth) 且为一正数， $K_h(\bullet)$  为有标尺的核函数且  $K_h(x) = [K(x/h)]/h$ 。从上式可以看出， $\hat{f}_h(x)$  不仅与核函数的选择和带宽有关，还与样本点的集合密切相关。核函数  $K(\bullet)$  通常符合以下规则：

- (1)  $K(-u) = K(u)$
- (2)  $\text{Sup}|K(u)| < \infty$  且  $\int_{-\infty}^{+\infty} K(u)du = 1$

在核密度估计中，常用的核密度函数如表 2-1 所示。

表 2-1 常用核函数

| 核函数              | 定义                                                                                                        |
|------------------|-----------------------------------------------------------------------------------------------------------|
| 均匀核函数            | $K(u) = \begin{cases} \frac{1}{2} & \text{如果 }  u  \leq 1 \\ 0 & \text{如果 }  u  > 1 \end{cases}$          |
| 三角核函数            | $K(u) = \begin{cases} 1 -  u  & \text{如果 }  u  \leq 1 \\ 0 & \text{如果 }  u  > 1 \end{cases}$              |
| Epanechnikov 核函数 | $K(u) = \begin{cases} \frac{3}{4}(1 - u^2) & \text{如果 }  u  \leq 1 \\ 0 & \text{如果 }  u  > 1 \end{cases}$ |
| 高斯核函数            | $K(u) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}u^2\right)$                                           |

由于高斯核函数具备良好的数学性质且为最常用的核函数<sup>[14]</sup>，因此本章采用高斯核函数对测试部件 RUL 的概率密度函数进行近似求解。高斯核函数可以通过以下式进行描述：

$$K\left(\frac{x-x_i}{h}\right) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(x-x_i)^2}{2h}\right) \quad (2-17)$$

为了简洁, 将  $(w_k^l, PRL_k^l)$  改写为  $(w_k^l, r_k^l(i))$ 。其中  $r_k^l(i)$  表示第  $l$  个测试部件在第  $i$  个监测周期基于第  $k$  个训练部件经过退化轨迹相似性分析所得的 PRL 值, 且  $w_k^l$  表示其对应的权重值。 $(w^l, r^l(i)) = \{(w_k^l, r_k^l(i))\}_{k=1}^M$  表示经过数据预处理模块后的结果, 将该结果作为 KDE 模型的输入数据, 测试部件 RUL 的概率密度函数  $\hat{f}_h(r_u)$  可近似表示为:

$$\hat{f}_h(r_u) = \sum_{k=1}^M \frac{w_k^l}{\sqrt{2\pi}h} \exp\left(-\frac{(r_u - r_k^l(i))^2}{2h^2}\right) \quad (2-18)$$

式中,  $\{(w_k^l, r_k^l(i))\}_{k=1}^M$  为已知的输入数据,  $M$  为所提出 TSBP 改进方法的一个参数。最终确定式(2-18), 还需对其带宽  $h$  进行估计。本章采用 Botev 等<sup>[114]</sup>提出的基于线性扩散过程的自适应 KDE 算法对带宽  $h$  进行自动选择。此外, Botev 等也开源了针对高斯核函数的一维数据带宽  $h$  自动选择的快速 KDE 的 MATLAB 实现代码<sup>[115]</sup>, 该算法在保证运行效率的同时其估计结果的精度也可以得到保证<sup>[115]</sup>。然而, 用于测试部件 RUL 的概率密度函数近似的输入数据  $(w^l, r^l(i))$  为一组带权重的 PRLs 的数据, 并不适用于快速一维数据的 KDE 算法, 因此本章还进一步提出了 PRLs 等权重样本扩充的方法, 如图 2-8 所示。

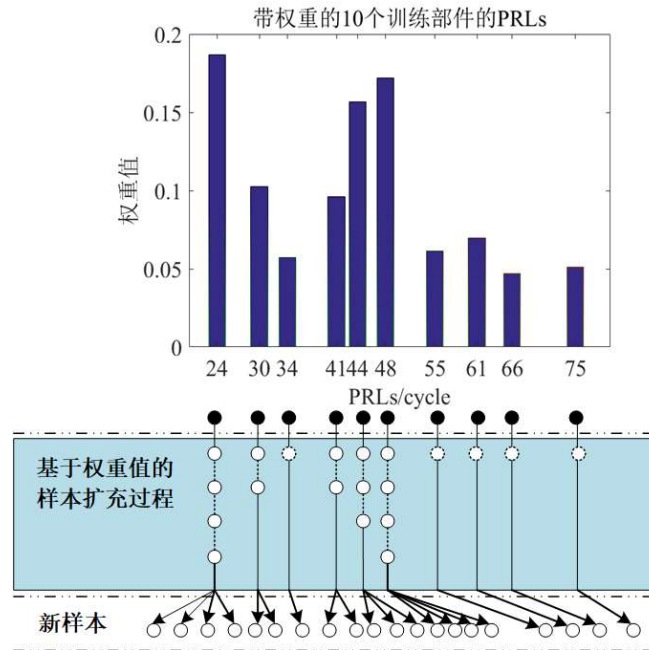


图 2-8 PRLs 等权重样本扩充方法

图 2-8 中,原始的带权重的 PRLs 的输入数据  $(w^l, r^l(i))$  按照其所具有的权重值进行扩充。具体地,以 PRL 为 24 个监测周期的训练部件为例,其对应的权重值接近 0.2,然而所有训练部件所对应的最小的权重值约为 0.05,因此该训练部件的 PRL 样本即可扩增四倍。按照该方法,带权重的 PRLs 的输入数据  $(w^l, r^l(i))$  即可进一步转化为等权重的一维输入数据  $r_k^l(i)$ ,然后即可输入到一维数据的快速 KDE 算法中以对带宽  $h$  进行自动选择并估计测试部件 RUL 的概率密度函数。值得注意的是采用本章所提出的 PRLs 等权重样本扩充方法所引入的不确定性很小并可以忽略。如图 2-9 所示,基于等权重样本扩充的 KDE 结果与原始的带权重的 KDE 结果<sup>[116]</sup>大致重合。但是,如果不采用等权重样本扩充的方法而直接将所得的 PRLs 数据(忽略其对应的权重值)输入到一维数据的快速 KDE 算法中,其所得 PDF 估计结果与实际的估计结果相去甚远。因此,本部分所提出的 PRLs 等权重样本扩充方法在保证计算速度的同时,并不会引入额外的不确定性。

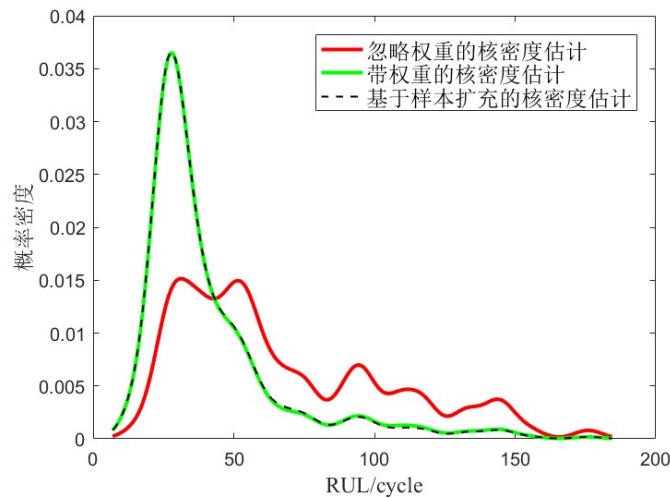


图 2-9 等权重样本扩充后所引入不确定性示例图

### 2.3.3 基于改进 TSBP 方法的剩余寿命预测

传统的基于 TSBP 方法通过对  $(w^l, r^l(i)) = \{(w_k^l, r_k^l(i))\}_{k=1}^M$  进行加权求和的方式得到测试部件的 RUL, 具体可用式(2-19)表示:

$$\hat{r}_l(i) = \sum_{k=1}^M w_k^l r_k^l(i) \quad (2-19)$$

式中,  $\hat{r}_l(i)$  表示第  $l$  个测试部件在第  $i$  个在线监测周期的 RUL 的点估计值。实际产品的 RUL 预测往往存在多种类型的不确定性, RUL 的估计值常常与实际值相差甚远。因此,工程实际中常基于 RUL 的点估计值给出一对应的容许误差界来刻画产品的 RUL。如图 2-10 所示,蓝色曲线表示由 KDE 算法估计所得的测试部件 RUL

的概率密度函数  $\hat{f}_h(r_u)$ ，在  $[\alpha_1^-, \alpha_1^+]$  和  $[\alpha_2^-, \alpha_2^+]$  的容许误差界内其对应的概率分别为 90% 和 50%，传统基于 TSBP 方法的 RUL 点估计值大致处于其均值附近。因此，相比于传统方法的单一 RUL 点估计，给出一对应的容许误差界将会使得产品的预防性维护决策更加可靠且稳健。

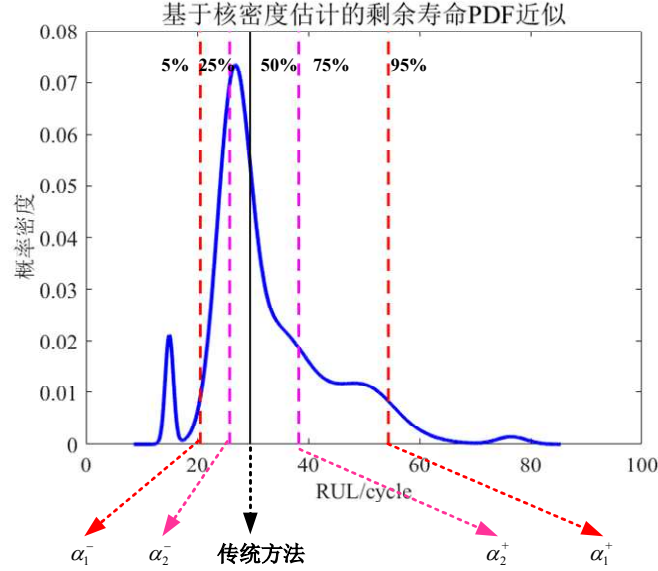


图 2-10 基于 RUL 的点估计值的容许误差界示例图

本章中，KDE 算法估计测试部件 RUL 的概率密度函数为一非参数方法，难以通过显式解析的方法求取其容许误差界的概率。本章采用基于直方图求和的方法<sup>[107]</sup>来计算容许误差界的概率，具体表达式为：

$$\pi[\hat{f}(r_u)]_{\alpha^-}^{\alpha^+} = \sum_{\alpha^-}^{\alpha^+} \hat{f}(r_u) \quad (2-20)$$

式中， $\pi[\hat{f}(r_u)]_{\alpha^-}^{\alpha^+}$  表示非参数概率密度分布  $\hat{f}(r_u)$  在容许误差界  $[\alpha^-, \alpha^+]$  内对应的概率。本章采用 Saxena 等<sup>[107]</sup>提出的  $\beta$ -准则来获得测试部件在某一置信度水平下的 RUL 区间估计， $\beta$ -准则定义为：

$$[r_-^l(i), r_+^l(i)] = [\alpha^-, \alpha^+] = \arg \min_{\min(r^l(i)) < \alpha^-, \alpha^+ < \max(r^l(i))} \left| \pi[\phi(x)]_{\alpha^-}^{\alpha^+} - \beta \right| \quad (2-21)$$

其中， $[r_-^l(i), r_+^l(i)]$  表示在  $\beta$  置信度水平下的 RUL 区间估计， $\min(r^l(i))$  和  $\max(r^l(i))$  分别表示 PRLs 对应的最小值与最大值。 $\beta \in (0, 1)$  其具体取值与测试部件的风险吸收能力有关<sup>[107]</sup>，针对某些安全关键产品，如航空发动机，参数  $\beta$  建议取较大的值（如 90%）以保证维修决策的有效性，进而避免灾难性事故的发生。

在工程实际中，RUL 点估计对产品的后续维修决策同样重要，本章根据式(2-

21), 提出改进 TSBP 方法, 该方法中 RUL 点估计为:

$$r_{\sigma}^l(i) = \arg \min_{\min(r^l(i)) < r^l(i) < \max(r^l(i))} \left| \pi[\hat{f}(r_u)]_{\min(r^l(i))}^{r^l(i)} - \sigma \right| \quad (2-22)$$

式中,  $r_{\sigma}^l(i)$  为第  $l$  个测试部件在第  $i$  个在线监测周期的 RUL 点估计值, 参数  $\sigma$  可有效控制 RUL 点估计值。工程实际中, 产品剩余寿命早预测 (RUL 的预测值小于其对应的实际值) 优于迟预测 (RUL 的预测值大于其对应的实际值)。因此, 当测试部件的传感器数据或合成的健康指标指示该产品大致处于退化初期, 且无明显退化趋势时, 较难准确预测 RUL, 此时建议参数  $\sigma$  取较小的值 (例如 10% 或 15%) 以保守估计 RUL。相反, 当测试部件的传感器数据或合成的健康指标指示该产品大致处于退化中期或晚期, 且具有明显退化趋势时, 建议参数  $\sigma$  取较大的值 (例如 25% 或 50%) 以准确估计 RUL。此外, 参数  $\sigma$  的取值也可结合具体产品进一步细化分析。

## 2.4 算例分析与方法验证

本节将用 NASA 阿姆斯研究中心所开源的关于航空发动机退化仿真数据集<sup>[108]</sup>对所提出的基于退化轨迹相似性分析的方法进行验证。

### 2.4.1 剩余寿命预测方法评价性能指标

为有效地验证 RUL 预测算法的预测性能, 需要用到不同的算法性能评价指标, 常用的性能指标分别为平均预测分数 (Mean Prediction Score, M\_PS)<sup>[117]</sup>、均方根误差 (Root Mean Square Error, RMSE)<sup>[117]</sup> 及漏报率 (False Negative Rate, FNR)<sup>[117]</sup>。

#### (1) 平均预测分数 (M\_PS)

针对 NASA 阿姆斯研究中心所开源的关于航空发动机退化仿真数据集, 预测分数是官方所定义的评价不同研究者所提的预测方法性能的关键指标, 定义为:

$$s = \begin{cases} \sum_{i=1}^n \left( e^{-\left(\frac{d_i}{\alpha_1}\right)} - 1 \right), & \text{如果 } d_i < 0 \\ \sum_{i=1}^n \left( e^{\left(\frac{d_i}{\alpha_2}\right)} - 1 \right), & \text{如果 } d_i \geq 0 \end{cases} \quad (2-23)$$

式中,  $s$  为总预测分数,  $n$  为测试部件数量,  $d_i = \hat{r}_{\sigma}^i - r_{*}^i$  为第  $i$  个测试部件 RUL 的预测误差, 其中  $\hat{r}_{\sigma}^i$  为第  $i$  个测试部件 RUL 的点估计值,  $r_{*}^i$  为第  $i$  个测试部件 RUL 的真实值。注意,  $d_i < 0$  表示早预测,  $d_i > 0$  表示迟预测, 根据文献<sup>[111]</sup>可知,

$\alpha_1 = 13$  和  $\alpha_2 = 10$  分别对应早预测和迟预测的惩罚系数。平均预测分数为总预测分数  $s$  在所有测试部件上的平均表现，其定义为：

$$M\_PS = \frac{s}{n} \quad (2-24)$$

### (2) 均方根误差 (RMSE)

均方根误差是评价 RUL 预测方法的常用性能指标。与平均预测分数相比，均方根误差分配给早预测和迟预测两种情况的惩罚系数相同。均方根误差具体定义为：

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n d_i^2} \quad (2-25)$$

### (3) 漏报率 (FNR)

工程实际中，早预测往往可以比迟预测更为有效地避免产品发生灾难性的失效，甚至可以进一步减少经济和人员的重大损失。漏报率是衡量预测算法是否出现迟预测情况的性能评价指标，其定义为：

$$FN(i) = \begin{cases} 1, & \text{如果 } r_o^i(t) - r_*^i(t) > 0 \\ 0, & \text{其他} \end{cases} \quad (2-26)$$

$$FNR = \frac{\sum FN(i)}{n} \quad (2-27)$$

注意，上述三个性能评价指标，即  $M\_PS$ 、 $RMSE$ 、 $FNR$ ，其值越小表明 RUL 预测算法的性能越好。

## 2.4.2 航空发动机退化仿真数据集

NASA 阿姆斯研究中心的航空发动机退化仿真数据集由 MATLAB 和 Simulink 编程环境下所开发的商用模块化航空推进系统仿真软件 (Commercial Modular Aero-Propulsion System Simulation, C-MAPSS) 所生成。C-MAPSS 仿真所用涡扇发动机的结构如图 2-11 所示，其中主要部件包括发动机风扇、燃烧室、低压涡轮、低压压气机、高压涡轮、高压压气机等。该航空发动机退化仿真数据集一共包含了 4 个数据子集 (Dataset #1-Dataset #4)，本章主要采用 Dataset #1 来验证基于 TSBP 的改进方法的有效性。Dataset #1 是在涡扇发动机单一故障模式 (高压压气机性能退化) 且单一工况下 (海平面) 所生成的。在 Dataset #1 中，共计有 100 个训练发动机和 100 个测试发动机，且所有发动机可视为来自于同一批次型号的产品。但是，每个发动机由于初始耗损和制造偏差均存在不同程度的初始损伤。

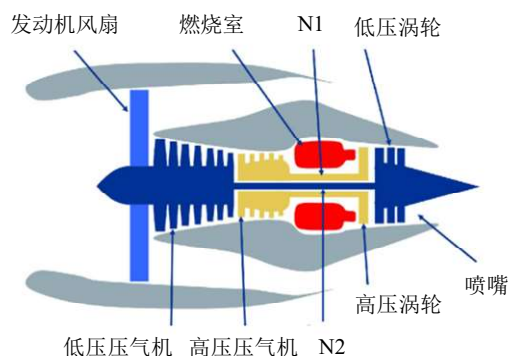


图 2-11 NASA 涡扇发动机结构示意图

为了有效地监测发动机的健康状态，共部署了 21 个不同的传感器。在这 21 个传感器的监测数据中，部分传感器的测量数据在整个监测周期内均保持恒定，因此，在发动机 RUL 预测的过程中不考虑该部分传感器的测量数据。表 2-2 具体给出了剩余的 14 个传感器的详细信息。其中， $^{\circ}\text{R}$  表示兰氏温度单位， $\text{psia}$  表示压力单位（磅力/平方英寸）， $\text{rpm}$  表示转速单位（转/分钟）， $\text{lbm/s}$  表示流量单位（磅/秒）。

表 2-2 航空发动机退化仿真数据集-有效传感器描述

| 传感器序号 | 表示符号    | 传感器描述        | 单位                 |
|-------|---------|--------------|--------------------|
| 2     | T24     | 低压压气机出口温度    | $^{\circ}\text{R}$ |
| 3     | T30     | 高压压气机出口温度    | $^{\circ}\text{R}$ |
| 4     | T50     | 低压涡轮出口温度     | $^{\circ}\text{R}$ |
| 7     | P30     | 高压压气机出口总压    | $\text{psia}$      |
| 8     | Nf      | 风扇物理转速       | $\text{rpm}$       |
| 9     | Nc      | 核心机物理转速      | $\text{rpm}$       |
| 11    | Ps30    | 高压压气机出口静压    | $\text{psia}$      |
| 12    | phi     | 燃油流量与 P30 比值 | $\text{pps/psi}$   |
| 13    | NRf     | 风扇换算转速       | $\text{rpm}$       |
| 14    | NRc     | 核心机换算转速      | $\text{rpm}$       |
| 15    | BPR     | 涵道比          | -                  |
| 17    | htBleed | 引气焓值         | -                  |
| 20    | W31     | 高压涡轮冷却引气流量   | $\text{lbm/s}$     |
| 21    | W32     | 低压涡轮冷却引气流量   | $\text{lbm/s}$     |

图 2-12 为对应于表 2-2 中所有训练发动机运行至失效的 14 个传感器的历史监测信号。其中，0 时刻表示发动机的失效时刻，因此之前的所有监测时刻可对应发动机的运行状态。由图 2-12 可知，这 100 个训练发动机的传感器监测信号中，T24、T30、T50、P30 等已经存在明显的退化趋势，而 Nc 和 NRc 没有退化趋势。因此，

这 14 个传感器中选择与发动机退化趋势关联性更强的传感器组合是后续发动机 RUL 预测的关键步骤。目前已存在大量关于该数据集传感器选择方面的研究工作和成果<sup>[104, 118]</sup>，本章通过综合分析不同传感器选择组合对所提出方法性能的影响，发现 Xu 等<sup>[104]</sup>所提出的传感器选择组合能产生较好的预测性能，因此，本章将采用文献[104]中的传感器选择策略（传感器#2, #4, #7, #8, #11, #12 和#15）。

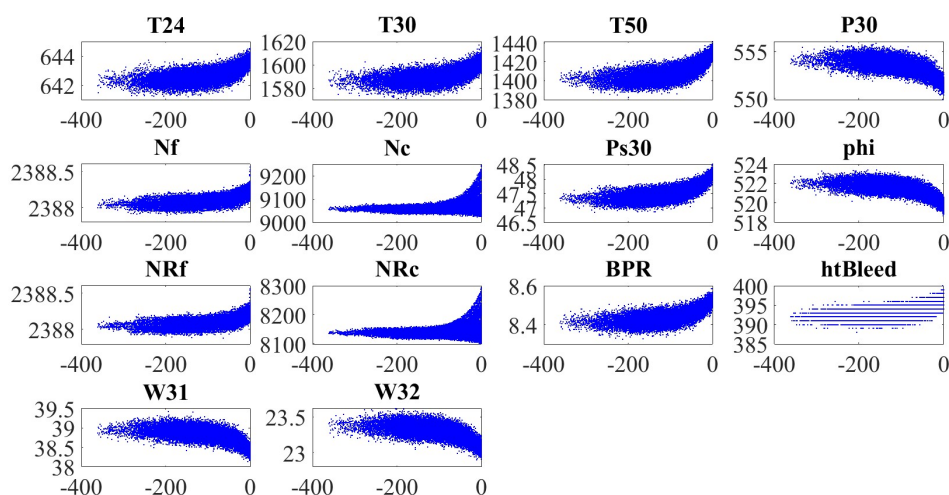


图 2-12 训练发动机运行至失效时 14 个传感器的监测信号

由于所选择的 7 个传感器监测信号已经存在明显的退化趋势，因此不用再提取额外的特征而将这 7 组传感器信号直接用于数据前处理。图 2-13 为训练发动机传感器原始监测信号经数据前处理后所得的退化模型库。

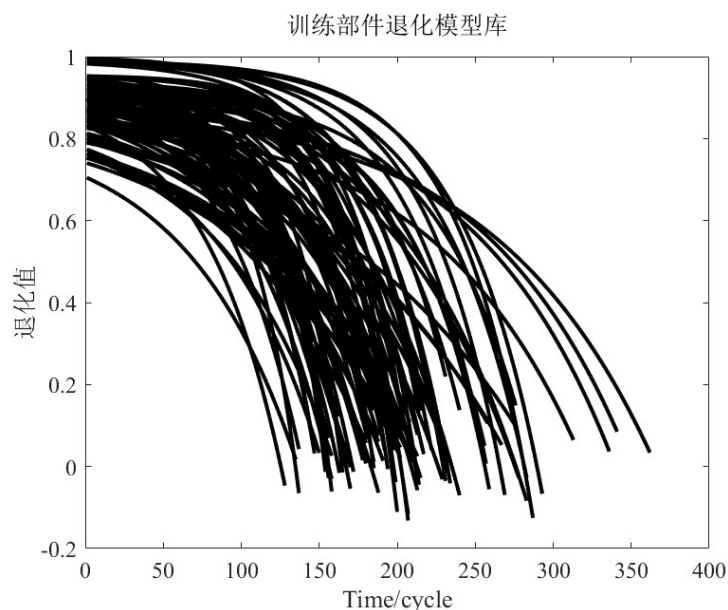


图 2-13 训练发动机退化模型库示例图

### 2.4.3 结果分析与讨论

本章所提出的基于 TSBP 的改进方法存在两个影响 RUL 预测性能的参数，分别为时间窗的长度  $\Delta t$  和  $(w^j, PRL^j)$  根据权重值排序后用于测试部件 RUL 预测的数据  $M$ 。为综合测试参数组  $(\Delta t, M)$  对所提方法 RUL 预测性能的影响，本小节对 100 个测试发动机在不同参数组  $(\Delta t, M)$  组合下的平均预测分数进行分析，其结果如表 2-3 所示。

表 2-3 不同参数组  $(\Delta t, M)$  下改进方法的  $M_{PS}$  预测结果

| $M$ | $\Delta t$ |      |            |            |            |            |     |
|-----|------------|------|------------|------------|------------|------------|-----|
|     | 20         | 25   | 30         | 35         | 40         | 45         | 50  |
| 25  | 15.1       | 10.4 | 8.5        | 7.9        | 6.5        | 5.8        | 6.3 |
| 30  | 12.9       | 8.6  | 7.0        | 6.2        | 5.7        | 5.4        | 6.1 |
| 35  | 11.5       | 7.0  | 6.2        | 5.7        | 5.2        | 5.2        | 6.0 |
| 40  | 10.1       | 6.2  | <b>5.7</b> | <b>5.4</b> | <b>5.2</b> | <b>5.4</b> | 6.3 |
| 45  | 9.7        | 5.7  | <b>5.6</b> | <b>5.2</b> | <b>5.4</b> | <b>5.2</b> | 6.3 |
| 50  | 9.2        | 5.4  | <b>5.4</b> | <b>5.4</b> | <b>5.1</b> | <b>5.2</b> | 6.4 |
| 60  | 8.9        | 6.1  | 5.5        | 5.4        | 5.5        | 5.4        | 6.6 |

由表 2-3 可知，随着时间窗长度的增加，基于 TSBP 改进方法的预测性能也逐渐变好，这主要是由于较长的时间窗可以涵盖更多的有关测试发动机局部退化趋势的信息。然而，当  $40 \text{ cycle} \leq \Delta t \leq 45 \text{ cycle}$  时， $M_{PS}$  逐渐稳定；当  $\Delta t = 50 \text{ cycle}$  时，预测性能变差。此外，随着  $M$  逐渐增大，所提方法的预测性能也在逐渐变好，但是当  $M \geq 40$  时，改进方法的预测性能同样趋于稳定。值得注意的是，当参数组  $(\Delta t, M)$  取值较大时，如表 2-3 中黑色方框内的数据，所提方法的预测性能相差不大。当基于 TSBP 的改进方法用于其他产品的 RUL 预测时，参数  $M$  的值建议取为总体训练部件数量的一半。另外，参数  $\Delta t$  的取值与测试部件的实际监测周期有关，当测试部件的监测周期小于 40 时，建议直接取测试部件的整个监测周期内的数据，当测试部件的监测周期大于 40 时，参数  $\Delta t$  可以直接取值 40。

为验证本章所提方法的有效性，下文将以相同的数据集 (Dataset #1) 为基础，分别采用基于 TSBP 的原方法及其改进方法<sup>[81, 86]</sup>、SKF 集成方法<sup>[119]</sup>、基于 CNN 的方法<sup>[120]</sup>以及本章所提的基于 TSBP 的改进方法等计算预测算法评价指标  $M_{PS}$  和 RMSE，计算结果如表 2-4 所示。TSBP 原方法由 Wang 等<sup>[81]</sup>于 2008 年 PHM 数据挑战竞赛中所提出。此后，Malinowski 等<sup>[86]</sup>对该方法进行改进，并提出可利用相似性匹配过程中相似度最高的前几组预测测试部件的 RUL，且根据不同的相似

性匹配的载体（原始传感器数据和合成退化指标）给出对应改进方法的预测性能。在表 2-4 中后面两种对比的方法均为基于人工智能的方法，其中 SKF 集成方法为应用开关卡尔曼滤波（Switching Kalman Filtering, SKF）集成的方法，基于 CNN 的方法为利用卷积神经网络的方法。从表中可以看出，本章所提出的方法就性能指标 M\_PS 而言取得了最好的预测性能，就 RMSE 性能指标而言也取得了较好的预测性能。

表 2-4 改进方法与公开发表方法的性能对比结果

| 预测方法                         | NASA C-MAPSS Dataset #1 |       |
|------------------------------|-------------------------|-------|
|                              | M_PS                    | RMSE  |
| 基于 TSBP 的原方法 <sup>[81]</sup> | 7.91                    | -     |
| TSBP 的改进方法 1 <sup>[86]</sup> | 8.07                    | -     |
| TSBP 的改进方法 2 <sup>[86]</sup> | 6.52                    | -     |
| SKF 集成方法 <sup>[119]</sup>    | 7.15                    | 19.08 |
| 基于 CNN 的方法 <sup>[120]</sup>  | 12.87                   | 18.45 |
| 本章所提的基于 TSBP 的改进方法           | 5.09                    | 18.82 |

最后为了验证本章所提方法对连续监测下部件 RUL 预测的有效性，本章随机选择了两个训练发动机（发动机#43 和发动机#74）作为测试集，余下的 98 个发动机作为训练集。图 2-14 和图 2-15 分别给出各自的 RUL 预测结果，需要指出的是预测算法在第 51 个监测周期被第一次触发，最终有效预测的时间设置为 10 个周期（即在发动机 EoL 之前 10 个周期停止 RUL 预测）。

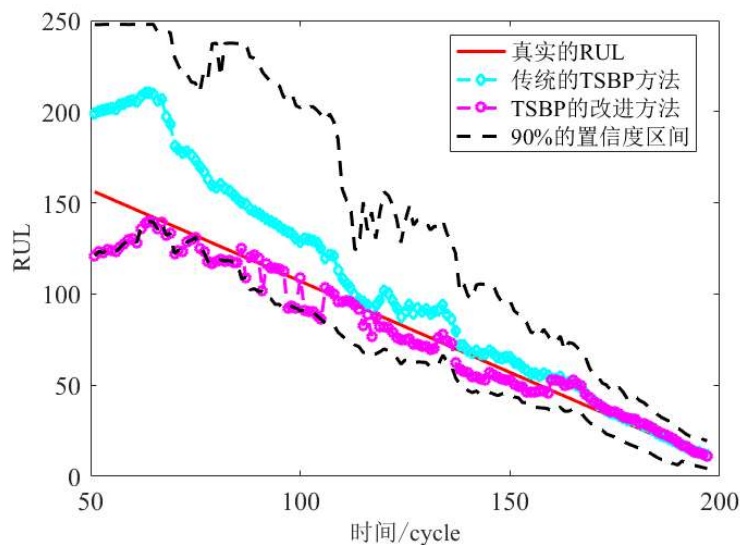


图 2-14 传统方法与改进方法就发动机#43 的 RUL 预测结果对比

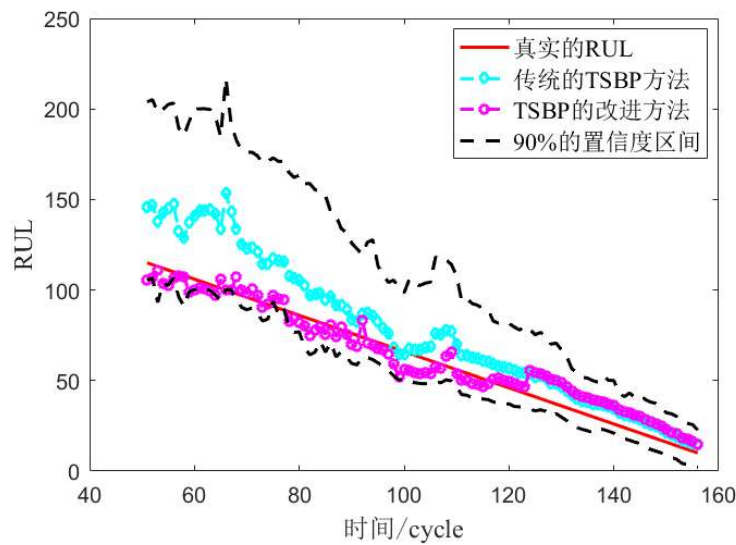


图 2-15 传统方法与改进方法就发动机#74 的 RUL 预测结果对比

表 2-5 传统方法与改进方法的 RUL 预测性能对比

| 预测性能指标 | 测试发动机编号 |      |         |      |
|--------|---------|------|---------|------|
|        | 发动机 #43 |      | 发动机 #74 |      |
|        | 原方法     | 改进方法 | 原方法     | 改进方法 |
| M_PS   | 57.35   | 1.15 | 9.98    | 0.88 |
| RMSE   | 26.64   | 9.82 | 18.99   | 7.17 |
| FNR    | 1       | 0.39 | 0.98    | 0.54 |

图 2-14 和图 2-15 中传统的 TSBP 方法和 TSBP 的改进方法分别就两个发动机 RUL 预测的性能如表 2-5 所示。对上述图和表中的内容分析如下：

(1) 就预测性能指标 M\_PS 和 RMSE，本章提出的 TSBP 改进方法在两个发动机的 RUL 预测性能上都要显著好于传统 TSBP 方法。具体来说，针对发动机早期运行阶段，例如发动机#43 在 51-100 个监测周期内，本章所提出的 TSBP 改进方法的 RUL 点估计值均给出了较为保守的估计，这主要是由参数  $\sigma$  进行控制，从图中可以看出该阶段的 RUL 点估计值均在 5%-15%分位数附近。其次，针对发动机中期运行阶段即为 101-150 个监测周期内，RUL 的点估计值为 25%分位数附近的估计值。最后，针对发动机末期运行阶段即 151-200 个监测周期内，RUL 点估计值为均值且与传统 TSBP 方法的估计相近。此外，从图中可以看出，传统的 TSBP 方法的点估计主要集中在 RUL 概率密度分布的均值处。

(2) 正是由于本章所提出的 TSBP 改进方法的参数  $\sigma$  有效地控制了发动机运行的各个阶段 RUL 点估计值，因此改进的方法可以取得相比于原方法更低的 FNR 预测性能。且由于性能指标 M\_PS 对 RUL 迟预测情况的惩罚系数更大，因此这也

进一步提升了改进方法就  $M_{PS}$  指标的预测表现。本章所提方法可以有效地控制 RUL 预测中迟预测的这一特征，也可以进一步增加该方法在安全关键性产品中的应用前景。

(3) 尽管在发动机预测初期，无论是 TSBP 原方法还是所提出的 TSBP 改进方法均存在一定程度上的预测误差，但是针对发动机预测末期（即最后 50 个监测周期内），本章所提的改进方法和原方法均可以取得较为良好的预测结果。实际上在发动机运行末期阶段，具备高精度的预测算法可以更加有效地支持发动机的维护决策，同时还可以显著地降低总的维修费用。

(4) 针对两个发动机在任何监测周期内，本章所提出的 TSBP 改进方法所给出的在 90%置信度水平下的 RUL 区间估计均能有效地将 RUL 实际值涵盖其中。因此，相比于 TSBP 原方法仅能给出 RUL 的点估计，在改进方法的 RUL 估计结果的基础上所做出的发动机维护决策将更加稳健且可靠。

## 2.5 本章小结

基于产品退化轨迹相似分析的原方法仅能给出产品剩余寿命的点估计结果，从而大大限制了该方法在某些产品中 RUL 预测的应用。本章采用核密度估计技术并结合  $\beta$ -准则，提出了基于产品退化轨迹相似性分析的改进方法。所提方法不仅能给出产品更为精确的 RUL 点估计结果，还能给出在某一置信度水平下的产品 RUL 的区间估计结果。通过算例分析，阐释了各个参数对所提出的改进方法性能的影响，且与多种当前最新的方法进行了性能对比，进一步验证了本章改进方法的有效性。最后经过与原方法性能的分析与比较，证明了改进方法的预测精度更高，且预测结果对产品的维护决策更为有效。本章所提方法将进一步拓展基于产品退化轨迹相似性分析方法的应用范畴，且为实际产品 RUL 预测的工程实践提供了坚实的理论与技术支撑。

## 第三章 基于双向长短期记忆神经网络的产品剩余寿命预测方法

### 3.1 引言

基于产品退化轨迹相似性分析 (TSBP) 的方法可以预测产品的 RUL, 但该方法存在两方面的弊端: 其一是该方法实施流程较复杂, 从产品原始监测信号的采集到 RUL 预测过程中包含了数据归一化、退化指标提取、退化轨迹建模、测试部件与训练部件之间的退化轨迹相似性分析、RUL 预测等步骤。值得注意的是: 产品 RUL 预测的最终结果很大程度上依赖于上述各个步骤实施的准确性, 这将使得基于 TSBP 方法所得的 RUL 预测结果的稳健性降低。其次就是该方法主要聚焦在解决工作在单一工况下产品的 RUL 预测问题, 这也将极大限制该方法针对工作在复杂多工况下产品的 RUL 预测的实际工程应用。

基于人工智能的方法对复杂产品的 RUL 预测问题具备强大的泛化能力, 因此该方法在产品 RUL 预测研究领域得到了极大的关注。与第二章研究思路不同的是: 基于人工智能的 RUL 预测方法直接对产品的监测原始信号和其对应的 RUL 之间的映射关系进行建模分析。在所有基于人工智能的方法中, 基于人工神经网络 (Artificial Neural Network, ANN) 的方法具有高度非线性、自适应性以及任意函数拟合的特性, 因而被认为是解决产品 RUL 预测问题的强有力工具。近年来兴起的深度学习 (Deep Learning, DL) 技术更是为传统的产品 RUL 预测问题提供了更加有效的解决途径。

当前, 基于 DL 方法的研究主要聚焦在利用不同的深度学习模型对产品的状态监测信号或基于状态监测信号所提取的时域、频域及时频域的初步特征进行深度特征提取, 并通过回归分析直接获取产品的 RUL。大多数基于 DL 方法的研究均仅利用了产品的多传感器状态监测数据, 而对产品工作在复杂多工况环境下的 RUL 预测问题鲜有研究, 尤其是融合多传感器状态监测数据和多工况监测数据的相关研究更是少之又少。在实际工程应用中, 为了完成特定的任务, 产品往往需要工作在多工况环境下。例如针对 NASA 阿姆斯研究中心所开源的有关航空发动机退化仿真数据子集 2 和数据子集 4<sup>[108]</sup>, 发动机在服役阶段经历了不同的工况环境, 例如飞行的海拔高度范围历经海平面至 40,000 英尺、温度范围为 -60°F-103°F, 飞行马赫数范围为 0-0.9。并且上述工况监测数据同发动机的状态监测数据一样均被实时采集并保存于数据中心。然而工况环境的变化会显著地加速或减速产品的退化过程, 进而影响产品的 RUL。例如, 相较于气候温和的工况环境, 飞机在热带沙漠性气候中航行将会显著地加速发动机风扇叶片的退化。因此, 本章基于双向长

短期记忆神经网络模型提出一种 RUL 预测方法，以融合多传感器监测数据和多工况监测数据，进而对复杂多工况下产品的 RUL 预测问题给出一种端对端的解决方案。

## 3.2 问题描述与数据前处理

### 3.2.1 问题描述

考虑同一批次产品，监测其健康状态且保存于多传感器数据中，这里用  $\mathbf{X}$  表示产品的多传感器状态监测数据。与此同时，也采集并保存该产品在运行过程中的环境工况数据，这里用  $\mathbf{U}$  表示产品的工况数据。此外，产品的总监测数据用  $(\mathbf{X}, \mathbf{U}, \mathbf{t})$  表示，其中， $\mathbf{t}$  表示多传感器状态监测数据  $\mathbf{X}$  和环境工况数据  $\mathbf{U}$  所对应的采集时刻点。训练部件指一批已失效的产品且其运行至失效的监测数据被保存于训练数据  $(\mathbf{X}_i^{\text{training}}, \mathbf{U}_i^{\text{training}}, \mathbf{t}_i^{\text{training}})$  中。此外，测试部件指一批正在服役的产品，其多传感器状态监测数据和环境工况数据也被实时采集并保存于测试数据  $(\mathbf{X}_j^{\text{testing}}, \mathbf{U}_j^{\text{testing}}, \mathbf{t}_j^{\text{testing}})$  中，测试部件的 RUL 需要根据测试集数据进行预测。

假定  $\{(\mathbf{X}_i^{\text{training}}, \mathbf{U}_i^{\text{training}}, \mathbf{t}_i^{\text{training}}), y_i\}$  表示一训练集，其中  $y_i$  表示其对应的产品真实 RUL。基于 ANN 的 RUL 预测方法的核心思想即确定一映射函数  $f: \hat{y}_i = f(\mathbf{X}_i^{\text{training}}, \mathbf{U}_i^{\text{training}}, \mathbf{t}_i^{\text{training}}; \boldsymbol{\theta})$  以最小化损失函数：

$$\min_{\boldsymbol{\theta}} \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - f(\mathbf{X}_i^{\text{training}}, \mathbf{U}_i^{\text{training}}, \mathbf{t}_i^{\text{training}}; \boldsymbol{\theta}))^2} \quad (3-1)$$

式中， $N$  表示训练样本量， $\boldsymbol{\theta}$  为 ANN 模型的参数集。通过合适的优化算法可以得到上述损失函数的最小值，该过程即为模型训练。假定  $\boldsymbol{\theta}^*$  表示模型训练完成后所获得的最小损失函数对应的模型参数集。最终测试部件的 RUL 可以通过将测试数据输入到已完成训练的 ANN 模型中进行计算，测试部件 RUL 预测表示为：

$$\hat{y}_j^{\text{testing}} = f(\mathbf{X}_j^{\text{testing}}, \mathbf{U}_j^{\text{testing}}, \mathbf{t}_j^{\text{testing}}; \boldsymbol{\theta}^*) \quad (3-2)$$

基于 ANN 的 RUL 预测框架主要涵盖三个关键组成部分。首先是数据前处理部分，即对原始监测数据进行合适的前处理，为 ANN 模型输入做准备；其次就是构建 ANN 模型以实现对输入数据（经前处理后的产品监测数据）和输出数据（产品的 RUL 信息）之间的映射关系建模；最后是模型训练，即对所定义的 ANN 模型进行优化以获取模型的最优参数解。本章也将主要围绕上述三个关键组成部分进行展开。

### 3.2.2 数据前处理

在产品 RUL 预测的工程实践中，首先获取多传感器数据和工况数据，多数情况下这些数据的尺度不一样。如果直接将这些数据输入到 ANN 模型将会导致模型的权重参数分布不均，使得某些具备大尺度的输入数据将被赋予更大的权重，相反某些仅有小尺度的输入数据对应的权重有可能被忽略。此外，这也可能降低 ANN 模型在训练过程中的收敛速度。因此，在前处理步骤中，通过数据归一化处理将不同尺度下的输入数据归一化到相同的尺度下是必要的。本部分采用 z-score 归一化方法，由于单一工况下的 z-score 归一化方法已在本文 2.2.2 部分进行了介绍，此处将介绍多工况下的 z-score 归一化方法<sup>[121]</sup>，具体如下：

$$x'_{(m,d)} = \frac{x_{(m,d)} - \mu_{(m,d)}}{\sigma_{(m,d)}}, \forall m, d \quad (3-3)$$

式中， $x_{(m,d)}$  表示产品监测原始数据，且假设产品工作在  $M$  个环境工况下， $m$  为这  $M$  个环境工况之一。 $d$  表示输入数据的第  $d$  个特征，这里的第  $d$  个特征表示产品的第  $d$  路传感器信号或工况信号。 $x'_{(m,d)}$  表示经过 z-score 处理后的归一化监测数据。 $\mu_{(m,d)}$  和  $\sigma_{(m,d)}$  分别表示输入数据的第  $d$  个特征在第  $m$  个工况下的均值和标准差。

BLSTM 模型能对时间序列数据的时间相关性进行建模并提取深层特征，因此，本章采用滑动时间窗处理技术对产品的整个监测周期内的数据进行处理并获取适当长度的序列数据。图 3-1 给出了滑动时间窗处理技术的实施流程。如图所示，多传感器状态监测数据和环境工况数据统称为多元监测数据  $(X, U) = [x_i, u_i]_{i=1}^{EoL}$ ，其中  $x_i = (x_i^1, x_i^2, \dots, x_i^p)$  表示在第  $i$  个数据采样周期的多传感器监测数据， $p$  表示传感器的数量； $u_i = (u_i^1, u_i^2, \dots, u_i^q)$  表示在第  $i$  个数据采样周期的工况监测数据， $q$  表示环境工况的数量，图中的  $N$  个特征即表示  $p + q = N$  且不同特征在图中用不同颜色进行表示。用一固定长度时间窗来表示一固定长度序列的多元监测数据  $[x_i, u_i]_{i=1}^L$ ，这里  $L$  表示时间窗的长度。然后时间窗向前滑动一个数据采集周期得到下一个多元监测数据序列  $[x_i, u_i]_{i=2}^{L+1}$ ，时间窗以步长为 1 滑动至 EoL 处并得到最后一个多元监测数据序列  $[x_i, u_i]_{i=EoL-L+1}^{EoL}$ 。这里应用滑动时间窗处理技术将产品整个监测周期内多元监测数据转化为一系列长度为  $L$  的序列数据，上述过程表示为：

$$(X, U) = [x_i, u_i]_{i=1}^{EoL} \rightarrow \left\{ [x_i, u_i]_{i=1}^L, [x_i, u_i]_{i=2}^{L+1}, \dots, [x_i, u_i]_{i=EoL-L+1}^{EoL} \right\} \quad (3-4)$$

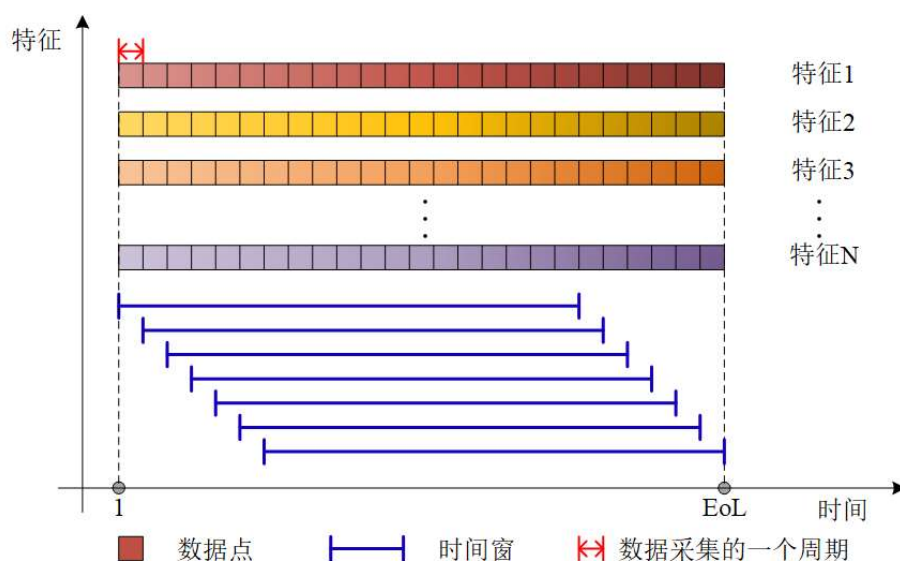


图 3-1 滑动时间窗处理技术示例图

将经时间窗处理后的序列数据输入到 BLSTM 模型，便能进行时间相关性建模并提取深层特征。需要指出的是更长的序列数据可以涵盖产品更多的局部退化信息，这将有助于 BLSTM 模型提取隐藏在序列数据中更加有效的深层特征，进而提升算法的预测精度。与此同时，这也将增加模型的复杂度从而模型的训练速度将变慢且有可能出现过拟合的现象。因此，序列数据长度这一关键参数会显著影响所提方法的性能，在本章 3.4 节将结合实际算例对此展开讨论。

### 3.3 基于双向长短期记忆神经网络的剩余寿命预测方法

本章所提出的基于 BLSTM 的 RUL 预测方法的实施流程如图 3-2 所示。其中离线阶段的实施流程包括：（1）提取训练部件运行至失效的监测数据并进行数据前处理，对前处理的监测数据进行标记以获取训练部件的 RUL 标记值也即真实值；（2）将前处理后的多传感器监测数据当作 BLSTM 模型的主输入数据，并且将前处理后的环境工况监测数据和对应的监测时刻点数据当作 BLSTM 模型的辅助输入数据；（3）将所获取的主输入数据和辅助输入数据均输入到所提出的基于 BLSTM 的模型中，进行模型的前向传播且计算得到最终的模型输出，即 RUL 预测值  $\hat{y}$ ；（4）结合训练部件 RUL 的真实值  $y$  和通过前向传播所得到的 RUL 预测值  $\hat{y}$  构建模型的损失函数，并基于合适的优化算法对模型进行后向传播以训练所提出的 BLSTM 模型。在线阶段的实施流程主要包括：（1）提取测试部件的实时在线监测数据并同样进行数据前处理（该部分实施步骤与离线阶段类似），并进一步获取测试部件的主输入数据和辅助输入数据；（2）直接将所获得测试部件的主输入和辅

助输入数据输入到在离线阶段已完成训练的 BLSTM 模型中以进行前向传播,并最终获得测试部件的实时 RUL 预测结果。

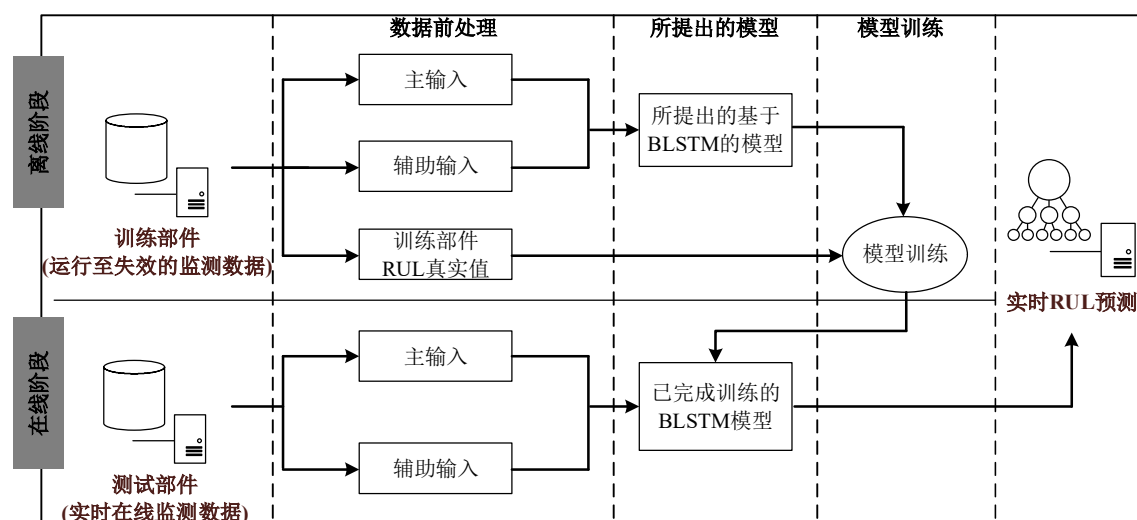


图 3-2 基于 BLSTM 的 RUL 预测方法实施流程

### 3.3.1 循环神经网络简介

前馈神经网络 (FFNN) 也常被称为多层感知机 (Multi-Layer Perceptron, MLP), 其简化的网络结构如图 3-3 所示。从图中可以看出 FFNN 在  $t$  时刻的输出  $y^t$  仅与当前时刻的输入  $(x_1^t, x_2^t, x_3^t)$  形成一一映射的关系, 因此 FFNN 网络结构不能有效地利用时间序列数据在  $t$  时刻之前的历史数据。

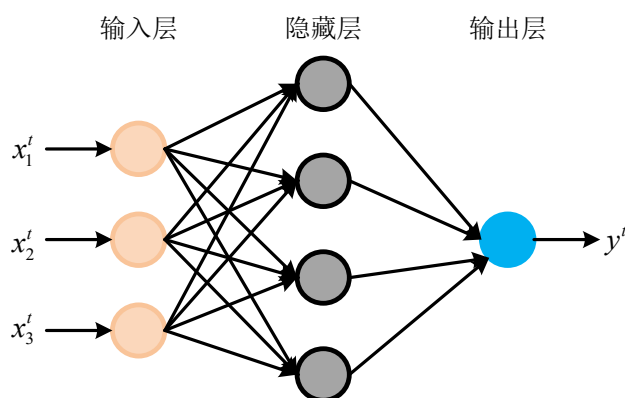


图 3-3 FFNN 的网络结构示意图

针对产品 RUL 预测的问题, 其监测数据尤其是多传感器状态监测数据在产品的整个监测周期内有时会呈现退化趋势, 因此监测数据在时间维度上是密切相关的, 有效地利用当前  $t$  时刻之前的监测数据将会有助于辅助当前决策且进一步提升

模型的预测效果<sup>[122-123]</sup>。不同于 FFNN 的网络结构, 循环神经网络 (Recurrent Neural Network, RNN) 具有带自反馈的神经元结构, 因此能有效地记忆当前时刻之前的监测信息并应用于计算当前的输出。图 3-4 为 RNN 的网络结构示意图, 其中  $\mathbf{X}$  表示多元时间序列数据为 RNN 模型的输入层,  $\mathbf{h}$  为 RNN 模型的隐藏层,  $\mathbf{y}$  表示模型的输出层。此外  $\mathbf{U}$ ,  $\mathbf{W}$ ,  $\mathbf{V}$  表示三种不同连接的权重矩阵 (分别为输入层-隐藏层、隐藏层-隐藏层、隐藏层-输出层), 且图中的环状结构即为 RNN 的自反馈结构, 黑色方块表示隐藏层激励函数。将该带自反馈的 RNN 结构在时间维度上进行展开即可得到右侧的示意图, 从图中可知权重矩阵  $\mathbf{U}$ ,  $\mathbf{W}$ ,  $\mathbf{V}$  在各个时刻点都是共享的, 该权重共享的策略将有助于 RNN 进一步提取序列数据的关键信息而不受其所在序列中的位置的影响。以  $t$  时刻为例, RNN 前向传播的过程可用以下方程表示:

$$\begin{aligned} \mathbf{h}^t &= \sigma(\mathbf{U}\mathbf{x}^t + \mathbf{W}\mathbf{h}^{t-1} + \mathbf{b}_h) \\ \hat{\mathbf{y}}^t &= \mathbf{V}\mathbf{h}^t + \mathbf{b}_y \end{aligned} \quad (3-5)$$

式中,  $\mathbf{x}^t$  表示当前时刻的输入数据,  $\mathbf{h}^t$  和  $\mathbf{h}^{t-1}$  分别表示当前时刻和上一时刻的隐藏层,  $\mathbf{b}_h$  表示隐藏层的偏置向量,  $\sigma(\cdot)$  表示一激励函数,  $\hat{\mathbf{y}}^t$  为 RNN 网络当前时刻的预测输出,  $\mathbf{b}_y$  为输出层的偏置向量。值得注意的是本章所聚焦解决的产品 RUL 预测问题本质属于一回归分析问题, 因此输出层没有施加额外的激励函数 (例如针对分类问题, softmax 函数或 sigmoid 函数常被用作输出层激励函数) 而直接使用线性函数。进一步具体展开当前  $t$  时刻的网络结构如下虚线框内的示意图所示, 由图可知, 输入层、隐藏层和输出层之间的连接均为全连接。假定  $\mathbf{x}^t = (x_1^t, x_2^t, \dots, x_N^t)$  表示输入数据具有  $N$  维特征,  $\mathbf{h}^t = (h_1^t, h_2^t, \dots, h_M^t)$  表示隐藏层具有  $M$  个隐藏节点, 输出层仅有一个节点  $\hat{\mathbf{y}}^t$ , 表示模型所预测的产品 RUL。RNN 模型向前传播的过程从  $t=1$  时刻开始, 随着时间  $t$  的增加采用式(3-5)依次完成迭代计算, 需要指出的是,  $t=1$  时刻对应模型的初始化步骤, 为了提升模型训练稳定性和预测性能, 通常假定隐藏层初始值  $\mathbf{h}^0$  为比较小但不为 0 的数<sup>[124]</sup>。

前面部分详细阐释了 RNN 模型的前向传播的过程, 即从输入数据到 RNN 模型输出的全过程。此外, RNN 模型的反向传播则主要涉及到模型的训练过程, 在定义好模型的损失函数的基础上, 一般采用沿时反向传播的算法 (Backpropagation Through Time, BPTT)<sup>[125]</sup>对 RNN 模型进行梯度计算, 进而优化 RNN 模型的参数, 该部分内容将在 3.3.3 节进行详细阐述。

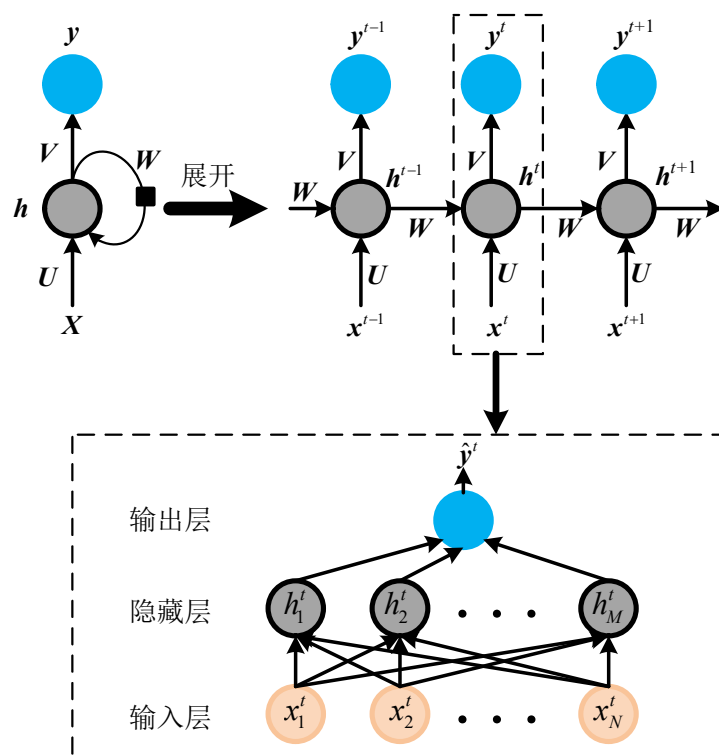


图 3-4 RNN 的网络结构示意图

RNN 模型所具备参数共享和自反馈的两大特点理论上使其能对任意长度的时间序列数据进行建模分析，但是在实际应用中 RNN 模型的训练极为困难尤其是针对较长的时间序列数据。从图 3-4 中可以看出，RNN 模型本质上就是参数共享的多层前馈神经网络，且层数对应了时间节点的数量。应用 RNN 模型对较长的时间序列数据进行分析时，由于其庞大且极深的网络结构将会显著地降低模型的训练速度，并且在应用 BPTT 算法的过程中常会出现所谓的“梯度爆炸”或“梯度消失”等问题<sup>[126]</sup>。这些问题将会最终导致 RNN 模型对序列数据长时间依赖关系（Long-Term Dependencies）建模能力的下降，从而进一步影响 RNN 模型对序列数据的时间相关性的建模能力。

针对 RNN 模型在训练中常出现的“梯度爆炸”或“梯度消失”等问题，研究人员提出了长短期记忆神经元（Long Short-Term Memory, LSTM）<sup>[126]</sup>和门限循环单元（Gated Recurrent Unit, GRU）<sup>[127]</sup>等 RNN 模型的改进版本。通过新引入的记忆单元和各种调节门的机制，使得 LSTM 模型和 GRU 模型能在较长的时间跨度内解决梯度值“消失”或“爆炸”的问题，以进一步稳定模型的整个训练过程。相比于传统的 RNN 模型，LSTM 模型和 GRU 模型能够真正地利用历史数据以实现对序列数据长时间依赖关系的建模与分析。下面依次对 LSTM 模型和 GRU 模型进行简单的介绍。

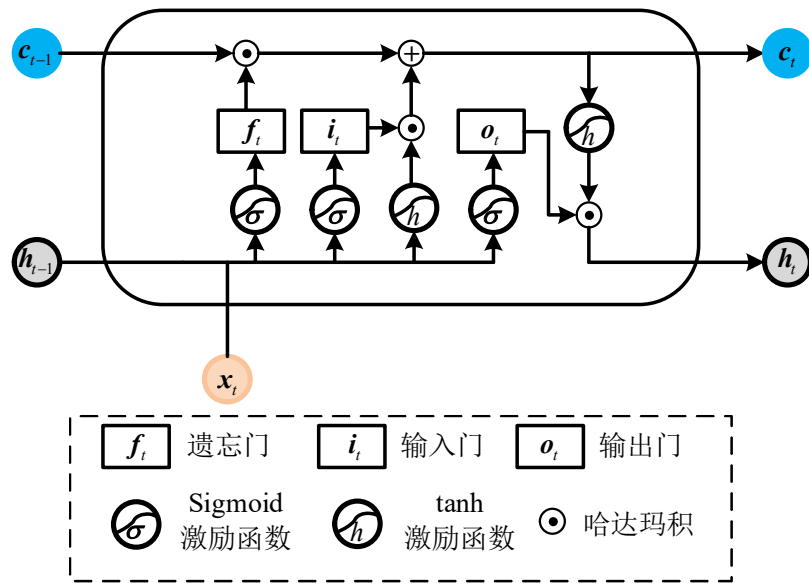


图 3-5 LSTM 单元的结构示意图

相比于简单的传统 RNN 模型，LSTM 模型内部增加了记忆单元和各种调节门的机制，用于实现对信息在序列数据中的有效传递。具体来说，LSTM 模型引入了 3 个调节门，分别为输入门、遗忘门和输出门。该 3 种调节门的引入主要用于对历史数据实现选择性的记忆与遗忘和对当前信息选择性的输出。此外，LSTM 模型中也引入了类似 RNN 模型中的隐藏状态功能的记忆单元，从而用于记录额外的信息。图 3-5 为一个 LSTM 单元的结构示意图，假定  $x_t$  表示当前时刻的输入数据， $h_t$  和  $h_{t-1}$  分别表示当前时刻和上一时刻的隐藏层输出， $c_t$  和  $c_{t-1}$  分别表示当前时刻和上一时刻记忆单元的状态。LSTM 模型的内部运算过程描述如下：

**Step1:** 确定遗忘门的值  $f_t$ 。遗忘门主要用于决定让哪些前一时刻的记忆单元  $c_{t-1}$  的信息通过当前时刻的记忆单元  $c_t$  中，可用以下方程计算：

$$f_t = \sigma(W_f x_t + R_f h_{t-1} + b_f) \quad (3-6)$$

式中， $W_f$  为输入数据与遗忘门之间的连接权重值矩阵， $R_f$  为隐藏层与遗忘门之间的连接权重值矩阵， $b_f$  为对应的偏置向量。需要指出的是， $\sigma(\cdot)$  激励函数的取值区间为  $[0, 1]$ ，当取值为 0 时表示遗忘掉所有前一时刻的记忆单元信息，当取值为 1 时则表示让所有前一时刻的记忆单元信息都进入到当前的记忆单元中。

**Step2:** 确定输入门的值  $i_t$ 。输入门主要用于决定让多少新信息添加到当前时刻的记忆单元  $c_t$  中，可用以下方程确定：

$$i_t = \sigma(W_i x_t + R_i h_{t-1} + b_i) \quad (3-7)$$

式中， $W_i$  为输入数据与输入门之间的连接权重值矩阵， $R_i$  为隐藏层与输入门之间

的连接权重值矩阵， $\mathbf{b}_i$  为对应的偏置向量。

**Step3:** 确定输出门的值  $\mathbf{o}_t$ 。输出门主要用于决定输出当前时刻的记忆单元  $\mathbf{c}_t$  的哪部分信息，可用以下方程计算：

$$\mathbf{o}_t = \sigma(\mathbf{W}_o \mathbf{x}_t + \mathbf{R}_o \mathbf{h}_{t-1} + \mathbf{b}_o) \quad (3-8)$$

式中， $\mathbf{W}_o$  为输入数据与输出门之间的连接权重值矩阵， $\mathbf{R}_o$  为隐藏层与输出门之间的连接权重值矩阵， $\mathbf{b}_o$  为对应的偏置向量。

**Step4:** 计算当前时刻候选记忆单元的值  $\mathbf{z}_t$ 。它的计算与上面三种门的计算过程类似，但是使用了双曲正切函数（ $\tanh$ ）作为其激励函数，可用以下方程表示：

$$\mathbf{z}_t = \tanh(\mathbf{W}_z \mathbf{x}_t + \mathbf{R}_z \mathbf{h}_{t-1} + \mathbf{b}_z) \quad (3-9)$$

式中， $\mathbf{W}_z$  为输入数据与记忆单元之间的连接权重值矩阵， $\mathbf{R}_z$  为隐藏层与记忆单元之间的连接权重值矩阵， $\mathbf{b}_z$  为对应的偏置向量。

**Step5:** 计算当前时刻记忆单元的值  $\mathbf{c}_t$ 。上述 **Step1-Step2** 步骤中使用元素值域范围在[0, 1]的遗忘门和输入门主要是用来控制 LSTM 单元内的隐藏状态中信息的流动。当前时刻记忆单元的值  $\mathbf{c}_t$  的计算需结合上一时刻记忆单元的值  $\mathbf{c}_{t-1}$  和当前时刻候选记忆单元的值  $\mathbf{z}_t$  的信息，具体可用以下方程表示：

$$\mathbf{c}_t = \mathbf{z}_t \odot \mathbf{i}_t + \mathbf{c}_{t-1} \odot \mathbf{f}_t \quad (3-10)$$

式中， $\odot$  表示哈达玛乘积，即两个矩阵对应元素的乘积。输入门  $\mathbf{i}_t$  控制当前时刻的输入数据  $\mathbf{x}_t$  通过候选记忆单元  $\mathbf{z}_t$  如何流入当前时刻的记忆单元状态  $\mathbf{c}_t$  内，遗忘门  $\mathbf{f}_t$  则控制上一时刻记忆单元的状态  $\mathbf{c}_{t-1}$  流入当前时刻记忆单元状态  $\mathbf{c}_t$  的程度。

**Step6:** 计算当前时刻隐藏层的值  $\mathbf{h}_t$ 。**Step5** 中得到了记忆单元的值  $\mathbf{c}_t$ ，之后可以进一步结合 **Step3** 中所得到的输出门用以控制从记忆单元到隐藏层之间的信息流动，上述过程可用以下方程表示：

$$\mathbf{h}_t = \tanh(\mathbf{c}_t) \odot \mathbf{o}_t \quad (3-11)$$

式中，双曲正切激励函数保证了隐藏层元素的值在-1 至 1 之间。

值得注意的是式(3-6)-式(3-11)中所使用的激励函数  $\sigma(\cdot)$  和  $\tanh(\cdot)$  分别表示 S 型生长曲线（Sigmoid Function）和双曲正切函数，其对应公式如下：

$$\sigma(x) = \frac{1}{1 + e^{-x}} \quad (3-12)$$

$$\tanh(x) = \frac{e^{2x} - 1}{e^{2x} + 1} \quad (3-13)$$

除了上文所介绍的 LSTM 基本模型，还存在一些改进的 LSTM 模型，其中最为主流的有以下两种：其一是在 LSTM 单元内增加窥视孔连接<sup>[128]</sup>以将记忆单元的状态值融合到 3 个调节门的计算中，其二为耦合输入门和遗忘门的 LSTM 改进模型。有关 LSTM 模型的各种变体结构请参考文献[129]，此处不再赘述。

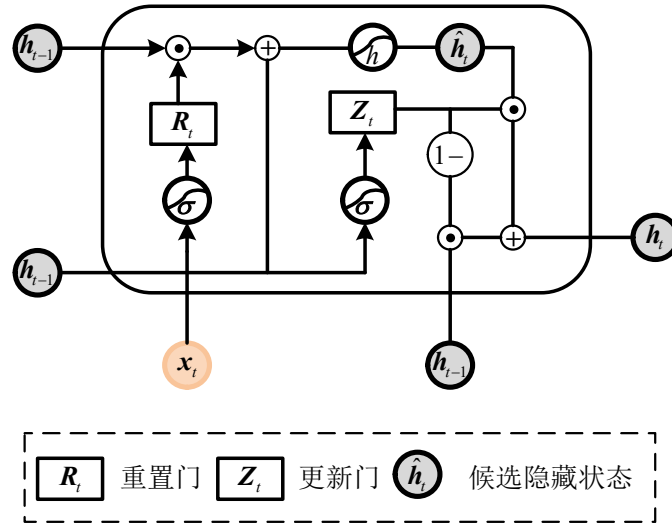


图 3-6 GRU 单元的结构示意图

由于 LSTM 模型有效地解决了 RNN 模型在训练过程中的梯度“爆炸”或梯度“消失”问题，因此 LSTM 被广泛应用到各类与序列数据相关的建模分析中，例如机器翻译<sup>[130]</sup>、语音识别<sup>[131]</sup>、移动物体识别与探测<sup>[132]</sup>等。但是，LSTM 模型内部结构较为复杂，相比之下 GRU 的结构在 LSTM 单元结构的基础上做了一定的简化，且可以取得与 LSTM 模型相当的性能。GRU 单元中仅含重置门和更新门，其结构如图 3-6 所示，其内部的运算过程描述如下：

**Step1:** 确定重置门的值  $R_t$ 。重置门主要用于控制对 GRU 隐藏状态信息不同程度的丢弃或保留，可用以下方程计算：

$$R_t = \sigma(W_{xr}x_t + W_{hr}h_{t-1} + b_r) \quad (3-14)$$

式中， $W_{xr}$  为输入数据与重置门之间的连接权重值矩阵， $W_{hr}$  为隐藏状态与重置门之间的连接权重值矩阵， $b_r$  为对应的偏置向量。 $\sigma(\cdot)$  激励函数取值为 0 时表示完全丢弃前一时刻的隐藏状态信息  $h_{t-1}$ ，取值为 1 时表示完全保留前一时刻的隐藏状态信息  $h_{t-1}$ 。

**Step2:** 确定更新门的值  $Z_t$ 。更新门主要用于控制隐藏状态应该如何被包含当前时刻的候选隐藏状态  $\tilde{z}_t$  所更新，具体可用以下方程确定：

$$Z_t = \sigma(W_{xz}x_t + W_{hz}h_{t-1} + b_z) \quad (3-15)$$

式中,  $W_{xz}$  为输入数据与更新门之间的连接权重值矩阵,  $W_{hz}$  为隐藏状态与更新门之间的连接权重值矩阵,  $b_z$  为对应的偏置向量。

**Step3:** 确定 GRU 候选隐藏状态的值  $\tilde{z}_t$ , 具体可用以下方程计算:

$$\tilde{z}_t = \tanh(W_{xh}x_t + W_{hh}(R_t \odot h_{t-1}) + b_h) \quad (3-16)$$

式中,  $W_{xh}$  为输入数据与隐藏层之间的连接权重值矩阵,  $W_{hh}$  为隐藏层权重值矩阵,  $b_h$  为对应的偏置向量。值得注意的是, 重置门控制了前一时刻的隐藏状态  $h_{t-1}$  如何流入当前时刻的候选隐藏状态  $\tilde{z}_t$ , 并且隐藏状态  $h_{t-1}$  可能包含了截止  $t-1$  时刻的所有历史信息。因此, 基于重置门激活函数  $\sigma(\cdot)$  的不同取值可以灵活地控制历史信息的保留或丢弃的程度。

**Step4:** 计算当前时刻隐藏状态输出值  $h_t$ , 它的计算主要是基于更新门  $Z_t$  对前一时刻隐藏状态  $h_{t-1}$  和对当前时刻的候选隐藏状态  $\tilde{z}_t$  的组合, 可用以下方程表示:

$$h_t = Z_t \odot h_{t-1} + (1 - Z_t) \odot \tilde{z}_t \quad (3-17)$$

式中, 值得注意的是更新门  $Z_t$  被用于控制隐藏状态应当如何被包含当前时刻候选隐藏状态  $\tilde{z}_t$  所更新, 假设更新门  $Z_t$  在  $t^*$  时刻到当前  $t$  时刻 ( $t^* < t$ ) 之间一直接近 1, 那么由  $t^*$  时刻到当前  $t$  时刻之间输入信息基本没有流入当前时刻的隐藏状态  $h_t$  中。当然这也可以视为隐藏状态  $h_{t-1}$  一直被保存至  $h_t$  中, 这个设计也确保了 GRU 模型具备对长时间序列数据的时间相关性的建模能力。

本部分主要对 RNN 模型及两种最为主流的变体结构, 即 LSTM 模型和 GRU 模型进行了简单的介绍。后续部分将主要基于 LSTM 模型提出用于产品在多工况下的 RUL 预测的方法。

### 3.3.2 基于 BLSTM 网络的产品多工况下剩余寿命预测模型

本章所提出的基于双向长短期记忆 (BLSTM) 神经网络的产品多工况下 RUL 预测模型如图 3-7 所示。从图中可知, 所提出的基于 BLSTM 的网络结构主要包括以下四大部分: (1) 提取隐藏在主输入序列数据中的时间相关性特征, 该提取过程由一 BLSTM 网络进行实现, 具体地 BLSTM 网络将逐层并分别采用前向和后向的计算过程对主输入的长时间依赖关系进行特征提取; (2) 结合辅助输入与前面的 BLSTM 网络所提取的特征向量为一新的融合输入序列数据; (3) 提取新的融合输入序列数据的时间相关性特征, 该提取过程由另一 BLSTM 网络完成; (4) 将第三部分所提取的特征进一步输入到全连接层中, 并最终通过一线性回归输出层对产品 RUL 进行预测。本章所提出的基于 BLSTM 网络的 RUL 预测模型将主输入和辅助输入视为独立的输入, 且仅用 BLSTM 网络对主输入序列数据的长时间依赖关系

进行特征提取。这主要是因为相比于产品多传感器状态监测数据（即文中主输入的概念）而言，环境工况数据（即文中辅助输入的概念）相对简单且易于理解，因此没用额外的 BLSTM 网络对辅助输入的序列数据单独进行特征提取，该种设计也进一步精简了网络的结构，有利于增强网络的训练效率。

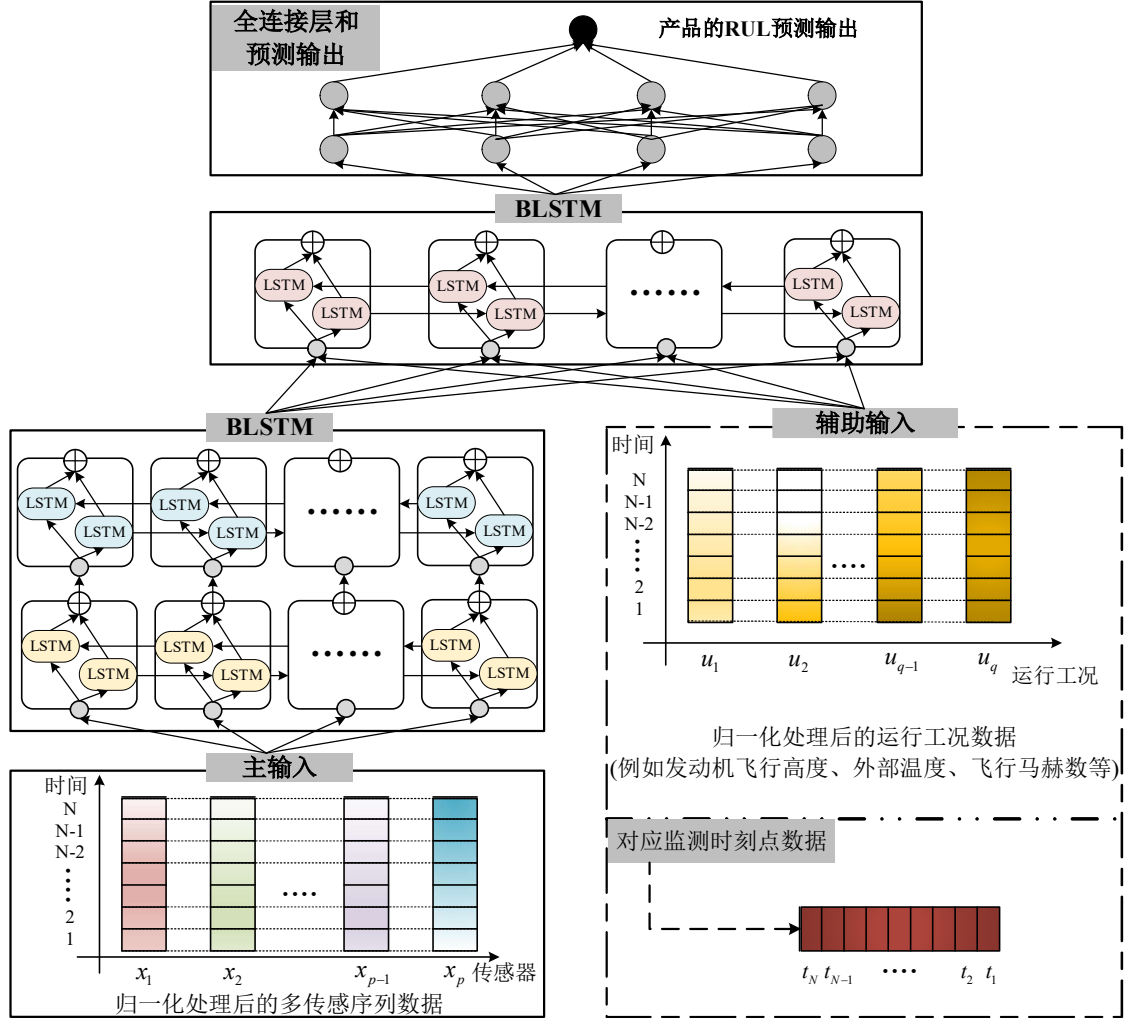


图 3-7 本章所提的基于 BLSTM 的网络结构

训练部件在多工况下所采集的监测原始信号将首先进行数据前处理，具体处理流程请参考文中 3.2.2 节内容。此后，经过前处理的数据将进一步形成训练集  $\{X_i, U_i, t_i, y_i\}_{i=1}^L$ ，其中  $X_i = [x_1, \dots, x_j, \dots, x_N]$  和  $U_i = [u_1, \dots, u_j, \dots, u_N]$  分别表示第  $i$  个多传感器数据序列和工况数据序列， $X_i$  和  $U_i$  数据序列均来源于连续  $N$  个数据采样周期  $t_i = \{t_j\}_{j=1}^N$ ， $x_j \in \mathbb{R}^p$  和  $u_j \in \mathbb{R}^q$  分别表示在第  $t_j$  个采样周期的归一化多传感器数据和归一化工况数据，且传感器数据和工况数据的维度分别为  $p$  和  $q$ 。此外，假定  $Y = \{y_i\}_{i=1}^L$  为训练数据序列  $\{X_i, U_i\}_{i=1}^L$  所对应的 RUL 标记值的集合，其中  $y_i$  为

序列  $\{X_i, U_i\}$  最后时刻监测数据  $(x_N, u_N)$  所对应的产品 RUL 真实值。图 3-7 中,  $\{X_i\}_{i=1}^L$  表示主输入的序列数据,  $\{U_i, t_i\}_{i=1}^L$  表示辅助输入的序列数据,  $\{y_i\}_{i=1}^L$  表示产品 RUL 的标记值将主要用于模型的训练阶段。

本章所采用的 BLSTM 网络是建立在基本的 LSTM 单元的基础上, 3.3.1 节已对 LSTM 单元内部的运算机制进行了简要的介绍, 本质上 3.3.1 节中式(3-6)-(3-11)刻画的是正向 LSTM 单元的计算流程, 即结合历史隐藏状态信息  $h_{t-1}$  和当前时刻的输入信息  $x_t$  实现对当前时刻隐藏状态信息更新输出, 上述正向 LSTM 单元计算过程可总结如下:

$$\vec{h}_t = f(\vec{x}_t, \vec{h}_{t-1}, \vec{\Theta}_{LSTM}) = \begin{cases} \vec{f}_t = \sigma(\vec{W}_f \vec{x}_t + \vec{R}_f \vec{h}_{t-1} + \vec{b}_f) \\ \vec{i}_t = \sigma(\vec{W}_i \vec{x}_t + \vec{R}_i \vec{h}_{t-1} + \vec{b}_i) \\ \vec{o}_t = \sigma(\vec{W}_o \vec{x}_t + \vec{R}_o \vec{h}_{t-1} + \vec{b}_o) \\ \vec{z}_t = \tanh(\vec{W}_z \vec{x}_t + \vec{R}_z \vec{h}_{t-1} + \vec{b}_z) \\ \vec{c}_t = \vec{z}_t \odot \vec{i}_t + \vec{c}_{t-1} \odot \vec{f}_t \\ \vec{h}_t = \tanh(\vec{c}_t) \odot \vec{o}_t \end{cases} \quad (3-18)$$

式中,  $\vec{\Theta}_{LSTM}$  表示正向 LSTM 单元的参数集合, 包括在正向计算过程中与当前输入信息  $\vec{x}_t$  相关的权重矩阵  $\vec{W}_f, \vec{W}_i, \vec{W}_o, \vec{W}_z$  和与隐藏状态  $\vec{h}_{t-1}$  相关的权重矩阵  $\vec{R}_f, \vec{R}_i, \vec{R}_o, \vec{R}_z$  及其对应偏置向量  $\vec{b}_f, \vec{b}_i, \vec{b}_o, \vec{b}_z$ 。上述正向 LSTM 单元仅考虑了历史信息对当前隐藏状态的影响, 但是实际上时间序列数据未来的信息同样对当前隐藏状态具有重要的影响, 因此同时考虑历史信息和未来信息将会进一步增强 LSTM 单元对时间序列数据长时间依赖关系的建模能力。在产品 RUL 预测领域, 有研究表明<sup>[133]</sup>, 同时考虑产品监测信号的历史信息和未来信息将会进一步提升基于 LSTM 模型的预测性能。因此, 本章将采用双向 LSTM 网络对产品在多工况下的监测数据进行特征提取进而预测产品的剩余寿命。双向 LSTM 单元的核心思想在于采用两个分开的隐藏层从正向和逆向处理序列数据, 以分别对历史信息和未来信息对当前隐藏状态的影响进行建模。式(3-18)刻画了正向 LSTM 单元内部计算过程, 逆向 LSTM 单元主要利用未来信息实现对当前隐藏状态信息的更新, 具体可用如下公式描述:

$$\bar{\mathbf{h}}_t = f(\bar{\mathbf{x}}_t, \bar{\mathbf{h}}_{t+1}, \bar{\boldsymbol{\Theta}}_{LSTM}) = \begin{cases} \bar{\mathbf{f}}_t = \sigma(\bar{\mathbf{W}}_f \bar{\mathbf{x}}_t + \bar{\mathbf{R}}_f \bar{\mathbf{h}}_{t+1} + \bar{\mathbf{b}}_f) \\ \bar{\mathbf{i}}_t = \sigma(\bar{\mathbf{W}}_i \bar{\mathbf{x}}_t + \bar{\mathbf{R}}_i \bar{\mathbf{h}}_{t+1} + \bar{\mathbf{b}}_i) \\ \bar{\mathbf{o}}_t = \sigma(\bar{\mathbf{W}}_o \bar{\mathbf{x}}_t + \bar{\mathbf{R}}_o \bar{\mathbf{h}}_{t+1} + \bar{\mathbf{b}}_o) \\ \bar{\mathbf{z}}_t = \tanh(\bar{\mathbf{W}}_z \bar{\mathbf{x}}_t + \bar{\mathbf{R}}_z \bar{\mathbf{h}}_{t+1} + \bar{\mathbf{b}}_z) \\ \bar{\mathbf{c}}_t = \bar{\mathbf{z}}_t \odot \bar{\mathbf{i}}_t + \bar{\mathbf{c}}_{t+1} \odot \bar{\mathbf{f}}_t \\ \bar{\mathbf{h}}_t = \tanh(\bar{\mathbf{c}}_t) \odot \bar{\mathbf{o}}_t \end{cases} \quad (3-19)$$

式中,  $\bar{\boldsymbol{\Theta}}_{LSTM}$  表示逆向 LSTM 单元的参数集合, 包括在逆向计算过程中与当前输入信息  $\bar{\mathbf{x}}_t$  相关的权重矩阵  $\bar{\mathbf{W}}_f$ ,  $\bar{\mathbf{W}}_i$ ,  $\bar{\mathbf{W}}_o$ ,  $\bar{\mathbf{W}}_z$ , 和与隐藏状态  $\bar{\mathbf{h}}_{t+1}$  相关的权重矩阵  $\bar{\mathbf{R}}_f$ ,  $\bar{\mathbf{R}}_i$ ,  $\bar{\mathbf{R}}_o$ ,  $\bar{\mathbf{R}}_z$ , 及其对应偏置向量  $\bar{\mathbf{b}}_f$ ,  $\bar{\mathbf{b}}_i$ ,  $\bar{\mathbf{b}}_o$ ,  $\bar{\mathbf{b}}_z$ 。

由图 3-7 可知, 本章所提出的基于 BLSTM 网络的多工况下产品剩余寿命预测模型主要包含四个部分, 现对模型的技术细节描述如下:

**Part1:** 将主输入序列数据即  $\mathbf{X}_i = [\mathbf{x}_1, \dots, \mathbf{x}_j, \dots, \mathbf{x}_N]$  输入到第一个 BLSTM 网络中以提取序列数据的时序特征, 输出特征可计算如下:

$$\mathbf{H}_i^{(1)} = [\mathbf{h}_1^{(1)}, \mathbf{h}_2^{(1)}, \dots, \mathbf{h}_t^{(1)}, \dots, \mathbf{h}_N^{(1)}] = f_1(\mathbf{X}_i; \bar{\boldsymbol{\Theta}}_{BLSTM}^{(1)}, \bar{\boldsymbol{\Theta}}_{BLSTM}^{(1)}) \quad (3-20)$$

式中,  $f_1(\cdot)$  表示 BLSTM 网络隐藏层函数且具体由式(3-18)和式(3-19)所刻画,  $\mathbf{H}_i^{(1)}$  表示对应于主输入序列数据  $\mathbf{X}_i$  经过第一个 BLSTM 网络处理后的输出特征序列数据, 且第一个 BLSTM 网络由参数集  $(\bar{\boldsymbol{\Theta}}_{BLSTM}^{(1)}, \bar{\boldsymbol{\Theta}}_{BLSTM}^{(1)})$  所刻画。需要指出的是每一个时刻点完整的输出结果均是由正向和逆向输出结果按元素逐点求和所得到, 具体以  $t$  时刻为例, 其完整的输出  $\mathbf{h}_t^{(1)}$  可由下式计算求得:

$$\mathbf{h}_t^{(1)} = \bar{\mathbf{h}}_t^{(1)} \oplus \mathbf{h}_t^{(1)} \quad (3-21)$$

**Part2:** 将辅助输入序列数据  $\{\mathbf{U}_i, \mathbf{t}_i\}$  同 **Part1** 中经过第一个 BLSTM 网络所提取的输出特征序列进行结合, 并形成新的融合输入特征序列数据  $\{\mathbf{H}_i^{(1)}, \mathbf{U}_i, \mathbf{t}_i\}$ 。此后, 采用第二个 BLSTM 对序列数据  $\{\mathbf{H}_i^{(1)}, \mathbf{U}_i, \mathbf{t}_i\}$  进一步提取其时序特征, 且  $N$  个时刻点中分别对应正向和逆向 LSTM 网络的最后时刻点的输出特征将进一步融合为第二个 BLSTM 网络的输出特征  $\mathbf{h}_N^{(2)} = \bar{\mathbf{h}}_N^{(2)} \oplus \mathbf{h}_N^{(2)}$ , 该过程可由图 3-8 和以下公式描述:

$$\mathbf{H}_i^{(2)} = \mathbf{h}_N^{(2)} = f_2(\left(\mathbf{H}_i^{(1)}, \mathbf{U}_i, \mathbf{t}_i\right); \bar{\boldsymbol{\Theta}}_{BLSTM}^{(2)}, \bar{\boldsymbol{\Theta}}_{BLSTM}^{(2)}) \quad (3-22)$$

式中,  $\mathbf{H}_i^{(2)} = \mathbf{h}_N^{(2)}$  表示第二个 BLSTM 网络的输出特征, 且第二个 BLSTM 由参数集

$(\vec{\theta}_{\text{BLSTM}}^{(2)}, \overleftarrow{\theta}_{\text{BLSTM}}^{(2)})$  所刻画。需要指出的是，经过第二个 BLSTM 网络处理所得到的输出特征并不是一特征序列，而是正向和逆向 LSTM 网络的最后时刻点的输出特征按元素求和后的一融合特征。

**Part3:** 将第二个 BLSTM 网络的输出特征  $H_i^{(2)}$  进一步输入到全连接层即一前馈神经网络，该过程描述如下：

$$o_i = f_3(H_i^{(2)}; \Theta_{\text{FC}}) = g(W_F h_N^{(2)} + b_F) \quad (3-23)$$

式中， $o_i$  为全连接层的输出特征， $\Theta_{\text{FC}}$  为全连接层对应的参数集，包括权重矩阵  $W_F$  和偏置向量  $b_F$ 。 $g(\cdot)$  表示全连接层的激励函数，在这里设置为线性整流函数 (ReLU)。值得注意的是式(3-23)定义的一单层结构，实际上可以堆叠更多的全连接层以形成深度结构。在深度结构中，前一层的输出特征即为下一层的输入数据。这里为了简化说明，假定  $o_i$  为全连接层的最终输出。

**Part4:** 将全连接层的输出特征  $o_i$  最后输入到一线性回归层以计算产品的 RUL，该过程可用下列公式描述：

$$\hat{y}_i = f_4(o_i; \Theta_R) = W_R o_i \quad (3-24)$$

式中， $\hat{y}_i$  表示产品的预测剩余寿命， $W_R$  表示最后线性回归层的权重向量。

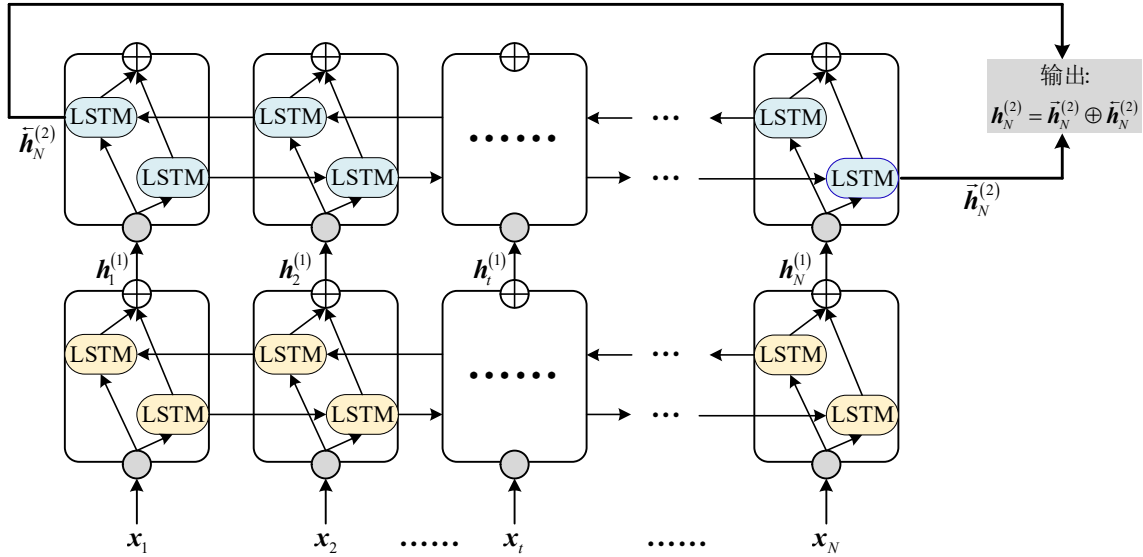


图 3-8 深度栈式 BLSTM 的网络结构

### 3.3.3 模型训练与正则化方法

模型的前向传播刻画了从产品监测数据获取到产品 RUL 预测整个计算过程。

相反，模型的反向传播则主要聚焦在利用所提供的对应于监测数据的产品 RUL 标记值并定义恰当的模型损失函数，且在适当的模型反向传播规则和优化算法的基础上对模型的参数集进行更新。但是，在实际的工程应用中，由于训练样本有限和所提出的深度学习模型过于复杂等原因将会产生模型过拟合的现象，因此采用模型正则化方法以降低模型的过拟合现象并增强模型的泛化能力。本部分内容将简要介绍相关模型训练的技术细节和后续采用的模型正则化方法。

通过模型前向传播可以得到产品的预测剩余寿命  $\hat{y}_i$ ，并且结合产品 RUL 的标记值  $y_i$ ，可以定义基于均方差（Mean Squared Error, MSE）的模型损失函数如下：

$$J(\boldsymbol{\theta}) = \frac{1}{L} \sum_{i=1}^L (y_i - f_{\text{BLSTM}}(X_i, U_i, t_i; \boldsymbol{\theta}))^2 \quad (3-25)$$

式中， $\hat{y}_i = f_{\text{BLSTM}}(X_i, U_i, t_i; \boldsymbol{\theta})$  表示当输入序列数据为  $(X_i, U_i, t_i)$  时，利用所提出的 BLSTM 网络模型计算所得的产品预测 RUL 输出  $\hat{y}_i$ 。所提出的 BLSTM 网络模型可用  $f_{\text{BLSTM}}(\boldsymbol{\theta})$  表示，且具体由式(3-18)-(3-24)所刻画。模型的参数集  $\boldsymbol{\theta}$  表示所提出 BLSTM 网络模型中所有参数的集合，具体包括第一个 BLSTM 网络模型的参数集  $(\bar{\boldsymbol{\theta}}_{\text{BLSTM}}^{(1)}, \bar{\boldsymbol{\theta}}_{\text{BLSTM}}^{(1)})$ ，第二个 BLSTM 网络模型的参数集  $(\bar{\boldsymbol{\theta}}_{\text{BLSTM}}^{(2)}, \bar{\boldsymbol{\theta}}_{\text{BLSTM}}^{(2)})$ ，全连接层的参数集  $\boldsymbol{\theta}_{\text{FC}}$  和最后线性回归层的参数集  $\boldsymbol{\theta}_{\text{R}}$ 。因此，模型的整体参数集  $\boldsymbol{\theta}$  可表示为  $\boldsymbol{\theta} = (\bar{\boldsymbol{\theta}}_{\text{BLSTM}}^{(1)}, \bar{\boldsymbol{\theta}}_{\text{BLSTM}}^{(1)}; \bar{\boldsymbol{\theta}}_{\text{BLSTM}}^{(2)}, \bar{\boldsymbol{\theta}}_{\text{BLSTM}}^{(2)}; \boldsymbol{\theta}_{\text{FC}}; \boldsymbol{\theta}_{\text{R}})$ 。针对所提出模型参数集  $\boldsymbol{\theta}$  的初始化问题，根据 Bengio<sup>[134-135]</sup>和 Hinton<sup>[136]</sup>等的研究成果，两个 BLSTM 网络中与隐藏状态  $h_{t-1}$  相关的权重矩阵均采用正交初始化方法<sup>[137]</sup>，与输入序列数据  $x_t$  相关的权重矩阵则采取 Xavier 均匀初始化方法<sup>[138]</sup>，遗忘门的偏置向量初始化为 1<sup>[139]</sup>，其余的偏置向量均进行零初始化。此外，全连接层  $\boldsymbol{\theta}_{\text{FC}}$  和线性回归层  $\boldsymbol{\theta}_{\text{R}}$  则均使用 Keras 模块默认的初始方法，即权重矩阵进行 gloriot 正态初始化<sup>[138]</sup>，偏置向量进行零初始化。

本章所提出的基于 BLSTM 网络的模型是属于 RNN 网络的一种改进版本，因此同样应用沿时反向传播 BPTT 算法通过将输出层的梯度值进行反向传播，并基于偏导数链式法则进一步计算各层每一个参数的对应梯度。在模型参数集  $\boldsymbol{\theta}$  的更新过程中，选用适应性矩估计（Adaptive Moment Estimation, Adam）方法<sup>[140]</sup>为模型的优化算法，Adam 方法是一种替代传统随机梯度下降过程的一阶优化算法，它能基于训练数据迭代更新神经网络权重且自适应地调节学习率。相比其他优化算法，Adam 方法收敛精度高且速度快，并且 Adam 方法有效地解决了学习率消失等问题。Adam 算法现已内嵌至绝大多数深度学习的编程平台中，实现较为简单，其有关理论及实现细节不属于本章研究内容，因此不再赘述。

由于本章所提出的基于 BLSTM 网络的模型结构复杂且在产品 RUL 预测的工

程实践中往往会面临训练数据有限等困境，因此引入模型正则化技术以有效地缓解模型在训练过程中出现的过拟合问题。本章主要采用两种正则化方法，分别为早停法<sup>[141]</sup>和丢弃法（Dropout Method）。其中早停法主要是在模型完成每次迭代训练后计算模型在验证集上的预测性能，当模型在验证集的预测性能开始下降时，则停止整个训练过程并输出模型参数更新之后的结果。早停法的应用可以有效防止模型在训练过程中由于过多的训练迭代次数（Epoch）发生模型过拟合的不利现象。此外本章还采用了由 Hinton 等<sup>[142]</sup>所提出的 Dropout 正则化方法以提升模型的泛化能力。Dropout 正则化方法的核心思想是在模型训练过程中，从 BLSTM 网络内以一定概率随机丢弃一些神经元以及其对应的连接，该策略可以有效地防止神经元之间由于过度的协同适应所产生的过拟合现象。需要指出的是，由于所提模型中两个 BLSTM 网络结构庞大且复杂，因此仅在这两个 BLSTM 网络中使用了 Dropout 正则化方法。

### 3.4 算例分析与方法验证

本章同样采用 NASA 阿姆斯研究中心所开源的有关航空发动机退化仿真数据集<sup>[108]</sup>对所提出的方法进行全方面的验证与讨论。本部分将先介绍多工况下航空发动机退化仿真数据和对应的数据前处理结果。其次，在结果分析与讨论部分，将对所提出的基于 BLSTM 网络的模型进行全面的分析与讨论，之后将本章所提方法与最新的基于人工智能的多种方法进行综合对比，以进一步验证本章所提方法的有效性。

#### 3.4.1 多工况下航空发动机退化仿真数据集简介及数据前处理

有关 NASA 阿姆斯研究中心所开源的航空发动机退化仿真数据集的基本信息已在本文 2.4.2 节进行了简要的介绍，值得注意的是用户可以通过 C-MAPSS 仿真模块设置一组输入参数以仿真发动机在运行过程中不同的环境工况，主要包括外界温度、飞行马赫数和飞行高度。该航空发动机退化仿真数据集一共涵盖了 4 个数据子集，具体信息如表 3-1 所示。从表 3-1 可知，数据集 FD001 和 FD003 均是针对运行在单工况下的发动机退化仿真数据，相反数据集 FD002 和 FD004 则是针对运行在多工况（共计 6 种不同的环境工况）下的发动机退化仿真数据。第二章 2.4 部分主要应用的是 FD001 对所提出的基于 TSBP 改进方法进行的验证，本部分则主要应用数据集 FD002 和 FD004 验证本章所提出的基于 BLSTM 网络的产品多工况下 RUL 的预测方法。

表 3-1 NASA 航空发动机退化仿真数据集简介

| 数据集描述   | NASA 航空发动机退化仿真数据集 |       |       |       |
|---------|-------------------|-------|-------|-------|
|         | FD001             | FD002 | FD003 | FD004 |
| 训练发动机数量 | 100               | 260   | 100   | 248   |
| 测试发动机数量 | 100               | 259   | 100   | 248   |
| 运行工况    | 1                 | 6     | 1     | 6     |
| 失效模式    | 1                 | 1     | 2     | 2     |

航空发动机退化仿真数据集共计有 21 个传感器数据、3 个运行工况数据和对应的监测时刻点数据,且 3 个运行工况数据共组成了 6 种不同的发动机运行工况。不同的工况将会对发动机性能造成显著的影响,并进一步影响其对应的传感器监测数据,造成监测数据在不同工况下的明显波动。图 3-9 给出了基于 FD002 数据子集的所有训练发动机的 21 个传感器监测信号,需要指出的是 0 时刻对应了发动机处于失效的时刻,所有之前时刻则对应了发动机的运行阶段。通过上述方式将所有 260 个训练发动机的在其对应的整个监测周期内的 21 个传感器监测信号均绘制于图 3-9 中。从图中可以清晰地看出由于发动机工作在 6 个不同的工况下,因此各个传感器监测信号均存在不同的聚类分布。所有 21 个传感器监测信号中共有 7 个传感器监测信号在整个监测时间内其聚类分布保持不变,分别是传感器 T2、P2、P15、epr、farB、Nf-dmd 和 PCNfR-dmd,因此部分传感器监测信号不会用于后续的建模分析中。剩余的 14 个传感器监测信号中还存在一些传感器监测信号在整个监测时间内其聚类分布变化很小,例如传感器 T24、P30、Nf、phi 和 NRf,同样地该部分传感器监测信号也不会用于后续的建模分析中,这将进一步提升后续基于 BLSTM 网络的模型训练效率。结果,本部分共选取了剩下的 9 个传感器监测信号作为所提出模型的主输入数据。此外,研究表明<sup>[143]</sup>,分段线性退化模型能较好地处理 NASA 航空发动机退化仿真数据集的 RUL 预测问题,因此本章将采用该模型对发动机的真实 RUL 值进行标记。图 3-10 给出了本章具体所采用的分段线性模型,需要指出的是 RUL 的最大值被限定在 130 以内,当 RUL 值小于 130 时,采用线性递减函数对发动机真实 RUL 值进行标记。

图 3-11 给出了针对某训练发动机的整个监测周期内多传感器原始监测数据,从图中可以清晰地看出 9 个所选取的传感器原始监测数据之间其变化范围各不相同且差异巨大。此外由于发动机运行在多工况下,因此 9 个传感器原始监测数据在整个监测周期内不存在明显的退化趋势。

图 3-12 给出了该训练发动机多传感器原始监测数据经过多工况 z-score 归一化(详见本章 3.2.2 节内容)处理之后的结果。从图中可以看出,9 个传感器监测

数据经过归一化处理之后其变化范围基本一致，并且在整个监测周期内存在明显的退化趋势，这将有利用于后续应用基于 BLSTM 网络的模型进行序列数据建模和 RUL 回归分析。

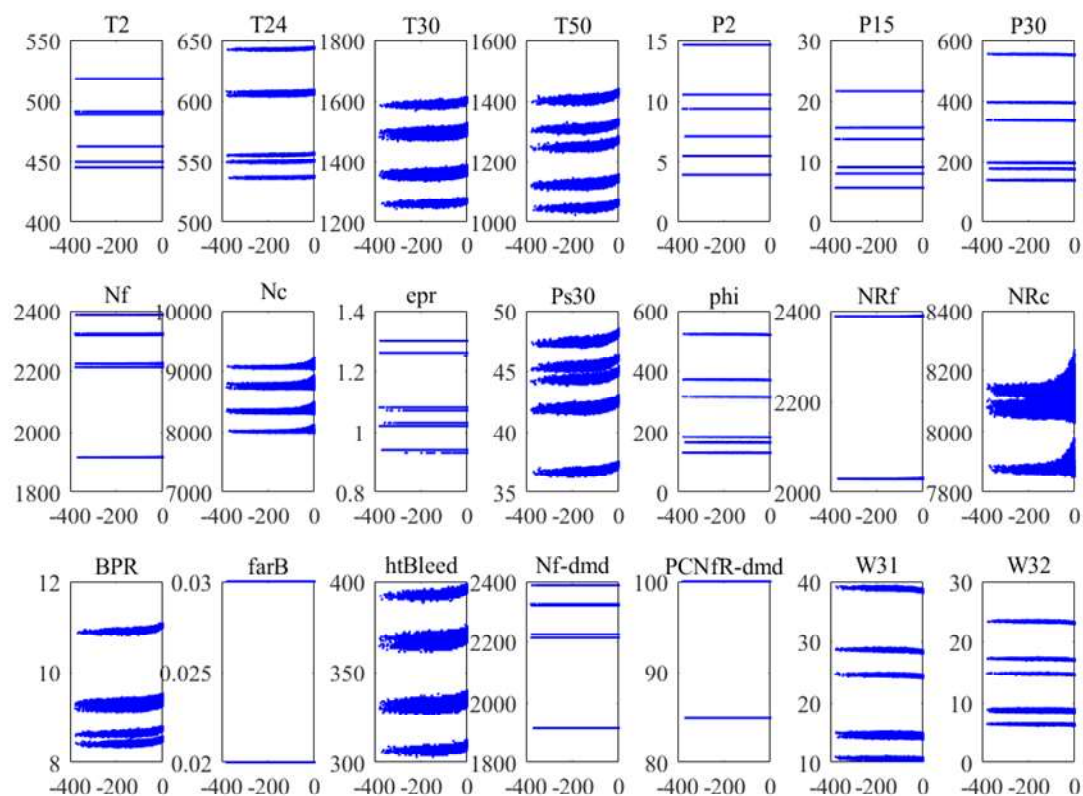


图 3-9 FD002 训练发动机 21 个传感器监测信号

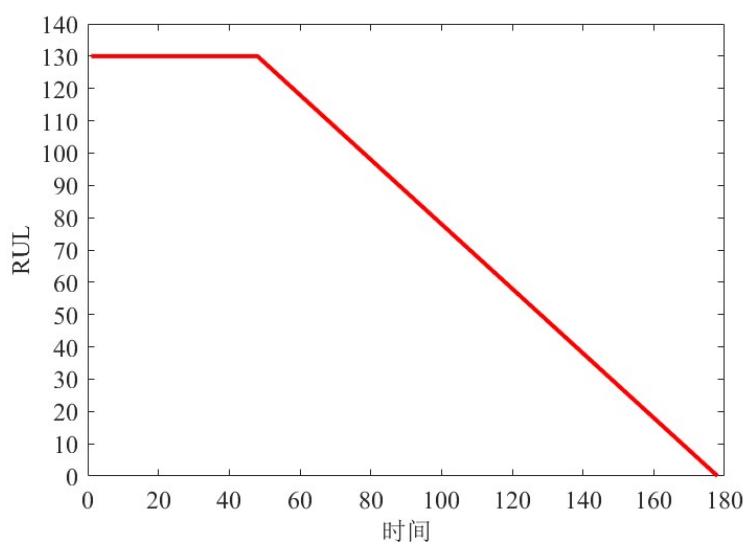


图 3-10 分段线性退化模型

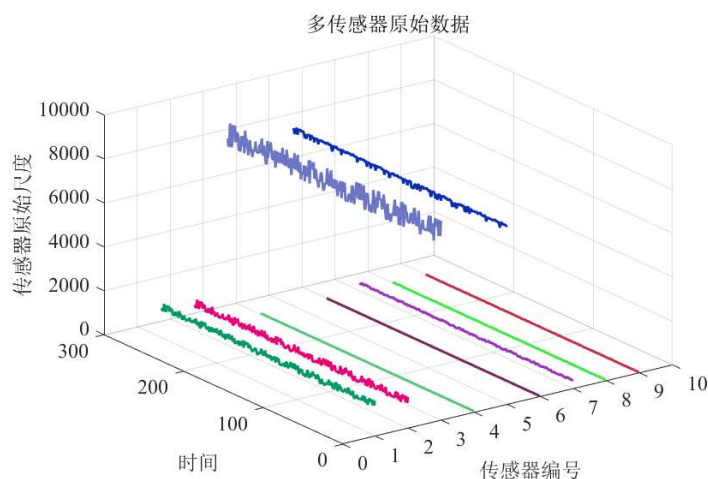


图 3-11 某训练发动机多传感器原始监测数据

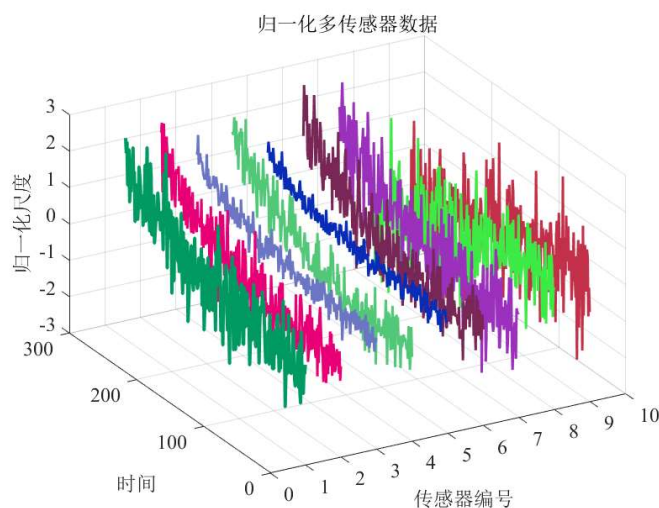


图 3-12 某发动机归一化多传感器监测数据

### 3.4.2 结果分析与讨论

本部分将首先针对显著影响所提出模型 RUL 预测性能的两个关键因素进行研究，这两个因素分别为所提出模型的网络结构和监测数据的序列长度。其次，为了验证本章所提出方法对利用辅助输入数据（即考虑产品多工况监测数据）的有效性，进一步开展了相应的对比实验，需要指出的是该部分的对比实验均基于 FD002 数据集展开。最后，同时使用 FD002 和 FD004 两个数据集对本章所提出的基于 BLSTM 网络的方法和其他的最新的基于 AI 的方法进行预测性能的综合对比。另外，本部分所选用的 RUL 预测方法性能评价指标主要是预测分数和均方根误差，本文 2.4.1 节已对这两个性能评价指标进行了介绍。

由于本章所提模型的 RUL 预测的最终结果将会受到很多方面的影响，例如模

型参数初始化、丢弃法（Dropout）的使用和迭代优化过程等，因此本部分的实验将重复运行 20 次，且给出其最终的平均预测结果以量化模型的预测性能。对于超参数的选取问题，诸如批量大小、丢弃率、训练迭代次数和早停基准，采用了网格搜索法且最终所获得超参数见表 3-2。此外，本部分仿真实验的编程平台为 sever Intel Xeon CPU E5-2698 v4 2.2-GHz, 64-GB RAM，操作系统为 Windows server 2016 datacenter，编程语言选用 Python 3.5 版本并搭载了以 tensorflow 为后端的 Keras 深度学习库。

表 3-2 基于网格搜索法的超参数选取列表

| 超参数    | 超参数描述                   | 建议取值 |
|--------|-------------------------|------|
| 批量大小   | 模型一次前向传播或后向传播所使用的训练样本数量 | 10   |
| 丢弃率    | BLSTM 网络单元的丢弃概率         | 0.20 |
| 训练迭代次数 | 针对所有训练数据模型前向传播或后向传播的次数  | 10   |
| 早停基准   | 模型在验证集的预测性能开始下降的次数      | 2    |

该部分首先研究不同的网络结构对所提出模型预测性能的影响。模型的网络结构主要由三个参数决定，分别是第一个 BLSTM 网络的隐藏层数目和对应各层的 LSTM 隐藏单元数目，第二个 BLSTM 网络的隐藏层数目和对应各层的 LSTM 隐藏单元数目，和全连接层的数目和隐藏单元的数目。为了清晰阐述， $L_1(18,18)L_2(18)N_3(18)$ 表示所提出模型共计有 4 个隐藏层，其中第一个 BLSTM 网络  $L_1$  有两个隐藏层，且各具有 18 个 LSTM 隐藏单元，第二个 BLSTM 网络  $L_2$  仅有一个隐藏层，且同样具有 18 个 LSTM 隐藏单元，最后全连接层  $N_3$  同样具有一个隐藏层，且同样包含 18 个隐藏单元。有研究<sup>[70]</sup>建议各个隐藏层的隐藏单元数目尽量一致，因此本部分不同的网络结构中的隐藏单元数目均保持一致。

表 3-3 不同的网络结构的平均预测性能

| 网络结构                                  | 预测分数 | RMSE  |
|---------------------------------------|------|-------|
| I: $L_1(12)L_2(12)N_3(12)$            | 6149 | 25.76 |
| II: $L_1(18)L_2(18)N_3(18)$           | 5875 | 25.77 |
| III: $L_1(36)L_2(36)N_3(36)$          | 6249 | 25.88 |
| IV: $L_1(64)L_2(64)N_3(64)$           | 6184 | 26.04 |
| V: $L_1(18,18)L_2(18)N_3(18)$         | 5880 | 25.78 |
| VI: $L_1(18,18)L_2(18)N_3(18,18)$     | 5714 | 25.73 |
| VII: $L_1(18,18)L_2(18,18)N_3(18,18)$ | 7261 | 26.05 |

表 3-3 给出了七种不同的网络结构重复运行 20 次的平均预测性能。从表中结

果可以看出,网络结构 I 至网络结构 IV 均采用了三个隐藏层,但是其对应的隐藏单元数目逐步递增,其中具有较少隐藏单元数目的结构 I 和较多隐藏单元数目的结构 III 和结构 IV 均不能获得较好的预测性能。相反,具有适中隐藏单元数目的网络结构 II 在四组不同的网络结构中取得了较好的预测性能,因此网络结构 II 中对应的隐藏单元数目将作为后续实验中默认隐藏单元数目。

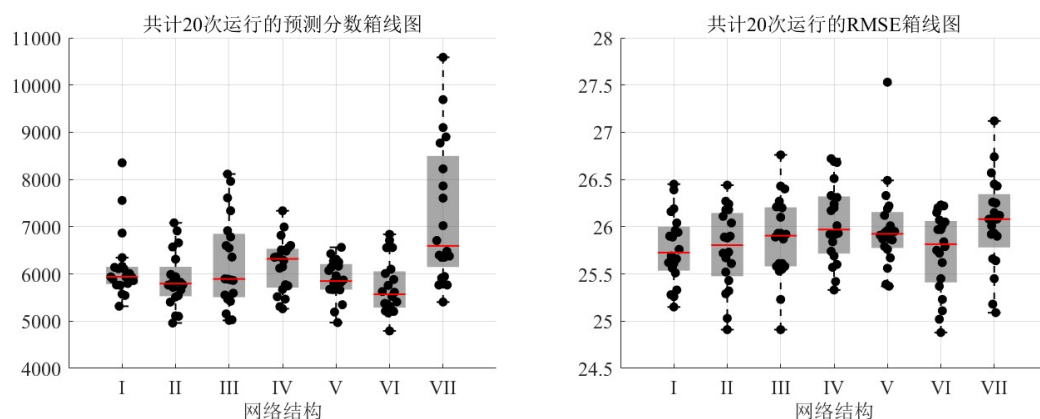


图 3-13 不同网络结构 20 次运行的预测性能箱线图

我们也研究了更加复杂的网络结构的预测性能,但是考虑到 FD002 数据集有限的训练样本量会导致模型过拟合的风险,因此仅将网络结构隐藏层数目设置在 3~6 之间。从表 3-3 和图 3-13 中均可以看出,具有 6 个隐藏层的网络结构 VII 在所有网络结构中的预测性能是最差的。从表 3-3 中可以看出,网络结构 VI 取得了最佳的平均预测性能,并且从图 3-13 中可以进一步看出,在重复 20 次的运行中其预测性能的稳健性也较高。对比网络结构 VI 和网络结构 II 可以发现,适当深度的网络结构有助于所提出的基于 BLSTM 网络的模型提取隐藏在序列数据中更加有效的时序特征,并产生更加精确的 RUL 预测结果。因此,网络结构 VI 将在后续的实验中被设置为默认的结构。

此后,还研究了监测序列数据的长度对所提出的 RUL 预测方法性能的影响。本部分的仿真实验中,一随机训练发动机(发动机#216)的监测数据被选作测试集,余下的 259 训练发动监测数据将用作训练集。本章所提出的基于 BLSTM 网络的方法仅需对该测试发动机的最后 150 个监测时刻内的 RUL 进行预测,图 3-14 给出了本章所提方法采用两种不同长度的监测序列数据(序列长度分别为 2 和 15)的 RUL 预测对比结果。

图 3-14 中灰色区域表示 RUL 标记值的 10%误差界,采用不同长度的监测序列数据的 RUL 预测结果用不同符号进行表示,且其对应颜色代表了预测结果是否在误差界内。从图中可以看出,当采用长度为 15 的监测序列数据时,尽管其 RUL

预测结果在标记值附近存在一定的波动，但是所有预测值均落于误差界内。然而，当采用长度为 2 的监测序列数据时，共计有 38 个预测结果位于误差界之外。值得注意的是采用两种不同长度的监测序列数据的预测结果在最后 30 个监测时刻内都具有很高的预测精度。

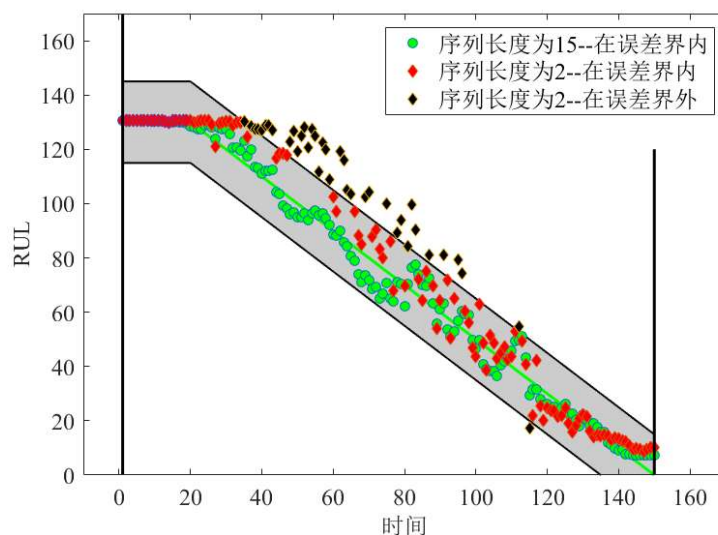


图 3-14 不同长度序列数据的 RUL 预测结果对比示意图

表 3-4 给出了更多有关采用不同长度的监测序列数据且重复运行 20 次的平均 RUL 预测性能结果。从表中可以看出，所提出的方法无论是采用较短的长度（比如序列数据长度为 2）或是采用较长的长度（例如序列数据长度为 30 和 40）均不能得到很好的预测性能，这主要可以从两方面进行解释。首先，当使用较短长度的监测序列数据作为模型的输入时，所提出的基于 BLSTM 网络的模型不能提取有效的时序特征，并且最终产生了较差的 RUL 预测结果。其次，当使用较长长度的监测序列数据作为模型的输入时，尽管输入序列数据已经包含了有关发动机更多的局部退化信息，但是由于 FD002 数据集有限的训练样本量和较长的时间步数均会给模型训练增加过拟合的风险，且最终导致不理想的 RUL 预测结果。因此，所提出方法使用适中长度（例如长度为 15 和 20）的监测序列数据将会产生最佳的 RUL 预测性能。此外，在模型训练阶段，随着序列长度的增加模型的训练时间也将会显著的变长。由于采用序列长度为 15 的监测序列数据作为模型输入时，本章所提方法获得最好的预测性能且其对应的训练时间适中，因此本章将监测数据的序列长度设置为 15 并应用于后续的仿真实验中。

为了验证本章所提出方法对利用辅助输入数据（考虑产品多工况监测数据）的有效性，也进一步开展了相应的对比实验。本部分主要对比了两种不同的 BLSTM 网络结构的 RUL 预测性能，其中第一个所采用的 BLSTM 网络结构即为文中 3.3.2

节中所详细阐释的本章所提出的网络结构，该结构能有效地对多传感器数据和工况数据以主输入和辅助输入的形式进行时序关系的建模。其次，第二个所采用的 BLSTM 网络结构为 Zheng 等<sup>[144]</sup>基于大量的仿真对比实验所提出的结构，该结构共采用了四个隐藏层，其中前两个 BLSTM 隐藏层均使用了 32 个隐藏单元，后两个全连接层则使用了 8 个隐藏单元，最后一层同样为线性回归层。此外，与前面对监测序列数据的长度研究类似，同样随机选取一训练发动机（发动机#223）的监测数据作为测试集，余下的 259 训练发动监测数据作为训练集。本部分两种不同的 BLSTM 网络仅需对该测试发动机的最后 150 个监测时刻内的 RUL 进行预测，图 3-15 给出了两种不同的 BLSTM 网络分别在考虑和忽略辅助输入数据时对 RUL 预测结果的影响。

表 3-4 不同长度序列数据的平均预测性能

| 序列长度 | 预测分数 | RMSE  | 训练时间（秒） |
|------|------|-------|---------|
| 2    | 290  | 10.32 | 60      |
| 10   | 178  | 8.01  | 196     |
| 15   | 141  | 7.50  | 295     |
| 20   | 154  | 8.33  | 348     |
| 30   | 168  | 8.86  | 548     |
| 40   | 187  | 9.84  | 723     |

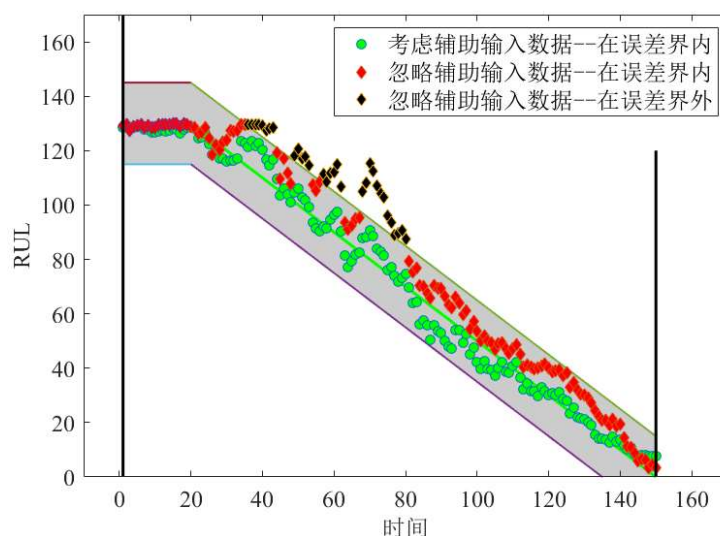


图 3-15 考虑或忽略辅助输入数据对 RUL 预测结果的影响

从图中可以看出，采用本章所提出的 BLSTM 网络结构在有效地考虑了辅助输入数据的情况下，能将最后 150 个监测时刻内的所有 RUL 预测值控制在误差界内。

其中在前 100 个监测时刻内, RUL 预测值大致分布在其标记值的两侧, 并且在最后 50 个监测时刻内, RUL 预测值与其对应的标记值基本一致。然而, 当采用传统的 BLSTM 网络结构在忽略辅助输入的情况下, 共计有 32 个预测结果位于误差界之外。此外, 表 3-5 给出了对应图 3-15 在 20 次运行中的平均预测性能结果, 从表中可以清晰地看出本章所提方法在考虑了辅助输入数据的情况下能有效地提升 RUL 预测的性能, 且其对应所增加的训练时间也在可接受的范围内。

表 3-5 考虑或忽略辅助输入数据的平均预测性能

| 方法       | 预测分数 | RMSE | 训练时间 (秒) |
|----------|------|------|----------|
| 考虑辅助输入数据 | 107  | 6.37 | 440      |
| 忽略辅助输入数据 | 177  | 8.50 | 417      |

我们还同时使用了 FD002 和 FD004 两个数据集对本章所提出的基于 BLSTM 网络的方法和其他的最新的基于 AI 的方法进行了性能对比。首先, 对比了其他基于 RNN 单元和其变体结构的方法, 主要包括 Vanilla RNN、LSTM、GRU、Bi-direction GRU (BiGRU), 并且这些对比方法均采用与本章所提的基于 BLSTM 网络的方法一致的网络结构和超参数。表 3-6 给出了本章所提方法与其他基于 RNN 及其变体结构方法的 RUL 预测性能对比结果。

表 3-6 与基于 RNN 及其变体结构方法的预测性能对比

| 方法    | FD002 |       |          | FD004 |       |          |
|-------|-------|-------|----------|-------|-------|----------|
|       | 预测分数  | RMSE  | 训练时间 (秒) | 预测分数  | RMSE  | 训练时间 (秒) |
| RNN   | 7787  | 27.83 | 175      | 12736 | 31.93 | 172      |
| GRU   | 6618  | 26.76 | 360      | 8080  | 29.68 | 329      |
| BiGRU | 6352  | 26.54 | 394      | 6339  | 29.13 | 361      |
| LSTM  | 6614  | 26.77 | 444      | 7341  | 29.44 | 420      |
| BLSTM | 4793  | 25.11 | 562      | 4971  | 26.61 | 550      |

从表中可以看出本章所提出的基于 BLSTM 网络的方法在 FD002 和 FD004 两个数据集上均取得了最好的预测性能。与基于 LSTM 网络的方法对比, 可以发现同时利用历史信息 and 未来信息 (即双向长短期记忆神经网络, BLSTM) 将会显著地提升 RUL 预测的性能但却要更多的训练时间。该现象同样可以在基于 GRU 网络和 BiGRU 网络的方法中得到验证。此外, 尽管基于 Vanilla RNN 的方法在所有对比方法中性能最差, 但由于其数学模型简单, 它的训练效率是最高的。在产品 RUL 预测的工程实践中, 可以考虑使用图形处理单元 (Graphic Processing Unit, GPU) 以进一步缩短本章基于 BLSTM 网络的方法的训练时间。

最后, 将本章所提的方法与其他已公开发表的最新的基于 AI 技术的相关方法 [66, 70, 119-120] 进行了性能对比, 以进一步验证所提方法的有效性和优先性。具体地, 本章所用到的对比方法主要包括基于 DCNN 的方法<sup>[66, 120]</sup>、基于多目标深度信念网络集成 (Multi-Objective Deep Belief Networks Ensemble, MODBNE) 的方法<sup>[70]</sup>、基于随机森林 (Random Forest, RF) 的方法<sup>[70]</sup>、基于梯度提升 (Gradient Boosting, GB) 的方法<sup>[70]</sup>和基于开关卡尔曼滤波 (Switching Kalman Filtering, SKF) 的方法<sup>[119]</sup>, 表 3-7 给出了上述方法 RUL 预测性能对比结果。

表 3-7 与最新的基于 AI 技术的相关方法的预测性能对比

| 方法                            | FD002 |       | FD004 |       |
|-------------------------------|-------|-------|-------|-------|
|                               | RMSE  | 预测分数  | RMSE  | 预测分数  |
| DCNN <sup>[66]</sup>          | 22.36 | 10412 | 23.31 | 12466 |
| MODBNE <sup>[70]</sup>        | 25.05 | 5585  | 28.66 | 6558  |
| RF <sup>[70]</sup>            | 29.59 | 70457 | 31.12 | 46568 |
| GB <sup>[70]</sup>            | 29.09 | 87280 | 29.01 | 17818 |
| Earlier DCNN <sup>[120]</sup> | 30.29 | 13570 | 29.16 | 7886  |
| SKF <sup>[119]</sup>          | 27.45 | 23834 | 26.76 | 7544  |
| Simple BLSTM*                 | 26.41 | 7150  | 28.12 | 6389  |
| Proposed BLSTM                | 25.11 | 4793  | 26.61 | 4971  |

\*Simple BLSTM 方法指所提出的基于 BLSTM 网络的模型但是忽略采用工况数据作为模型的辅助输入数据。

从表中可以看出, 与基于 RF<sup>[70]</sup>、GB<sup>[70]</sup>、SKF<sup>[119]</sup>、Earlier DCNN<sup>[120]</sup>等方法相比, 本章所提方法在两个数据集上均要显著地优于这些方法。需要指出的是 Earlier DCNN 表示 Babu 等<sup>[120]</sup>早期所提出的基于 DCNN 的航空发动机 RUL 预测方法。此外, 相比于最新所构建的基于深度学习的相关方法, 例如 DCNN<sup>[66]</sup>和 MODBNE<sup>[70]</sup>, 本章所提方法的预测性能同样具有优势, 讨论如下:

(1) 尽管基于 DCNN 的方法<sup>[66]</sup>在两个数据集上就预测性能指标 RMSE 而言均取得了最好的结果, 但是该方法所取得的对应预测分数结果均偏高 (这里预测分数结果越高表示该方法的预测精度越低), 这可能是由于该方法大多数 RUL 预测的结果晚于 RUL 预测的真实值。相比基于 DCNN 的方法, 本章所提的基于 BLSTM 网络的方法在 FD002 和 FD004 两个数据集上针对预测性能指标 RMSE 和预测分数而言均取得了较好的预测性能。

(2) 本章所提出的方法与其他基于 DL 的对比方法 (例如 Earlier DCNN<sup>[120]</sup>、DCNN<sup>[66]</sup>和 MODBNE<sup>[70]</sup>) 之间最大差别主要在于网络结构和输入模型的数据。需要指出的是, 其他基于 DL 的对比方法均只利用了多传感器监测数据作为模型的输

入而没有考虑使用发动机的工况监测数据。因此，在这里为了进行有效的对比，也给出所提出的基于 BLSTM 网络的方法但是忽略采用工况监测数据作为模型的辅助输入，并在表 3-7 中表示为基于 Simple BLSTM 的方法。从 FD002 和 FD004 两个数据集上的对比结果可以发现，有效地结合发动机多传感器监测数据和工况监测数据能进一步提升所提出方法的 RUL 预测性能。

(3) 由于 FD004 数据集是在两个失效模式（即为高压压气机性能退化和风扇退化）和六种不同工况下所采集的发动机退化数据，由此可见本章所提出的方法针对工作在多失效模式耦合与多工况环境下的产品也具有很强的泛化能力，并最终能对产品的 RUL 进行精确的预测。

### 3.5 本章小结

本章利用双向长短期记忆神经网络，提出了一种基于深度学习模型的产品 RUL 预测方法，所提方法能有效地对工作在多工况环境下和存在多失效模式耦合情况下的产品 RUL 进行精确预测。与其他基于深度学习方法不同的是本章所提方法能同时考虑多传感器监测数据和多工况监测数据，且在所提方法的框架内将归一化的多传感器监测数据当作模型的主输入数据，其次将归一化的多工况监测数据和对应监测时刻点数据当作模型的辅助输入数据。由于本章所提模型较为复杂且训练样本量有限，故采用丢弃法和早停法被用以缓解模型训练过程中的过拟合问题。通过算例分析进一步阐释了具体网络结构和输入数据序列长度对所提方法预测性能的影响，并通过对比实验进一步验证了本章所提方法采用工况信息作为模型辅助输入的有效性。最后将本章所提方法与众多最新的基于人工智能的方法进行了预测性能的对比，结果同样验证了本章方法的有效性。此外，由于本章方法能简洁且有效地实现从产品监测原始数据采集到 RUL 预测等一系列任务，因此本章研究能为实现产品端对端的 RUL 预测提供新的解决途径和技术支撑。

## 第四章 基于双向长短期记忆神经网络的混合剩余寿命预测方法

### 4.1 引言

第三章所提出的基于双向长短期记忆神经网络（BLSTM）的多工况下产品 RUL 预测方法，通过利用 BLSTM 网络对时间序列数据长时间依赖关系的建模优势，并有效地结合产品状态监测数据和多工况数据，从而实现了对工作在复杂多工况下产品 RUL 的端对端预测。尽管基于深度学习（DL）模型的 RUL 预测方法能在很多实际产品 RUL 预测任务中取得不错的预测精度，但是该方法仍然面临诸多挑战，例如基于 DL 的方法本质上解决的是产品监测数据和其对应 RUL 的回归分析问题，目前大多数基于 DL 的方法仅能得到产品 RUL 确定性的估计值而不能给出其在某一置信度下的区间估计值，这将大大限制基于 DL 方法在某些产品中的应用。

基于物理模型的 RUL 预测方法依据产品的失效机理或损伤法则构建相应能具体阐释产品退化规律的物理模型，并有效地结合实时在线监测数据与所构建的退化模型实现对产品的 RUL 的评估。在众多基于物理模型的方法框架中，粒子滤波（PF）算法常被用来解决一些非线性/非高斯系统的 RUL 预测问题。这里本文依据产品退化状态监测数据的获取方式，大致将该类方法分为两类，即基于直接测量的方法和间接测量的方法。基于直接测量的方法针对的是一些研究对象的退化状态可以直接获取的情况，例如锂离子电池的集总参数  $R_E + R_{CT}$  与电池的容量成反比例关系，可以直接用于表征锂离子电池的退化状态<sup>[100]</sup>，裂纹扩展长度测量数据可直接用以刻画与裂纹增长导致失效的产品的退化状态<sup>[109]</sup>，集电极和发射极之间的开路电压（ $V_{CE}$ ）的电气参数测量值，也可被用来描述绝缘栅双极晶体管的退化状态<sup>[145]</sup>。基于物理模型的 PF 方法就可以直接结合产品的退化模型和所获得产品退化监测数据，以更新模型的参数并最终实现产品的 RUL 预测。

相比之下，间接测量的方法则主要应用于一些研究对象的退化状态难以直接测量或测量成本极高的情况。例如针对数控机床由于加工过程中切削液的使用和照明等问题，刀具磨损量很难通过显微镜或 CCD 相机等手段进行直接测量以实时在线监测刀具的磨损状况。在这种情况下，间接测量技术以测量加工过程中一些辅助变量（例如切削力信号和刀具的振动信号）的变化来进一步推断刀具的磨损状况则是一种更加合适的在线监测实施方案。一些传统的混合 RUL 预测方法针对间接测量的情况，首先将提取监测原始信号（例如振动信号、切削力信号和声发射信号）的时域、频域和时频域特征。其次，应用降维方法（例如主成分分析法）进行特征

选择和融合。然后采用回归分析方法，例如支持向量机回归（SVR），进一步完成对融合特征和产品的退化状态的相关性建模分析。最后，用粒子滤波算法来处理上述通过回归分析所得到的产品退化状态数据并最终实现其 RUL 估计。但是，传统的混合 RUL 预测方法涵盖了繁琐的步骤，具体包括特征提取、选择、融合与回归分析，且产品最终的 RUL 估计结果的精度将会极大地依赖于上述各个步骤的准确实施。

针对基于 DL 的 RUL 预测方法和传统的混合 RUL 预测方法中所存在的不足，本章将提出基于双向长短期记忆神经网络的混合 RUL 预测方法。

## 4.2 数据采集与前处理

### 4.2.1 数据采集

考虑一批相似产品在运行过程中的退化状态被多个传感器（例如加速度传感器、测力传感器和声发射传感器等）连续监测的情形，这里将产品在运行过程中所采集的多传感器数据定义为在线测量数据。针对训练产品，记录其运行至失效的监测数据且保存于训练数据集  $\{s^i, y^i\}_{i=1}^L$ ，其中  $s^i \in \mathbb{R}^{M \times p}$  表示第  $i$  个训练样本的在线监测数据，其被  $p$  个传感器所监测，且每个传感器各采样  $M$  个数据点（ $M$  具体与传感器的采样频率和采集周期相关）， $L$  为总的训练样本数量。另外， $y^i$  表示产品对应第  $i$  个训练样本的退化状态，本章考虑间接测量的情况，即产品的退化状态  $y^i$  不能直接在线获取，需停机进行离线监测，例如针对数控机床刀具的磨损状态，需通过显微镜和 CCD 相机等测量技术在其停机状态下对刀具的退化状态  $y^i$  进行离线测量。针对测试产品，其多传感器数据同样被实时记录并保存于测试数据集  $\{s^j\}_{j=1}^C$ ，其中  $s^j \in \mathbb{R}^{M \times p}$  表示第  $j$  个测试样本的在线实时监测数据， $C$  为总的测试样本数量。针对测试产品其退化状态和剩余寿命等信息未知，需要依据其在线监测数据  $\{s^j\}_{j=1}^C$  进一步估计。

### 4.2.2 数据前处理

本章所涉及的数据前处理流程主要涵盖三个步骤，分别为基本特征提取、数据正则化处理和滑动时间窗处理，具体介绍如下。

对监测原始数据进行初步特征提取可以有效地压缩数据规模，并进一步降低后续深度学习模型的复杂度，最终可以显著地提升基于深度学习模型方法的实施效率。此外，带噪音的监测原始数据和数据中的一些异常点都将会影响后续基于深度学习模型方法的预测精度，合适的特征提取可以有效地抑制上述不利的影响<sup>[146]</sup>。

具体地,多传感器监测原始数据的时域特征可以首先被提取,例如均值、标准差、均方根、偏斜度、峭度。其次,基于快速傅里叶变换技术,可以进一步提取多传感器监测原始数据的频域特征,例如平均频率、谱偏斜度、谱峭度、谱功率等。最后,基于经验模式分解和小波变化等技术,还可以提取其时频域特征。由于本章的贡献点主要聚焦在产品的 RUL 预测方法上,因此有关 FFT, EMD 和 WT 等技术在此不再赘述。通过上述特征提取步骤,训练数据集和测试数据集中的多传感器监测原始数据  $\{\mathbf{s}^i\}_{i=1}^L$  和  $\{\mathbf{s}^j\}_{j=1}^C$  可以分别转换为一组训练初步特征集  $\{\mathbf{x}^i\}_{i=1}^L$  和测试初步特征集  $\{\mathbf{x}^j\}_{j=1}^C$ , 其中  $\mathbf{x}^i \in \mathbb{R}^{1 \times N}$  和  $\mathbf{x}^j \in \mathbb{R}^{1 \times N}$  分别表示第  $i$  个训练样本  $\mathbf{s}^i$  和第  $j$  个测试样本  $\mathbf{s}^j$  所对应初步特征,  $N$  表示所提取的初步特征的数量。需要指出的是,本文的第二章和第三章均处理的是低采样频率下的监测数据,由于总的监测数据规模不大且已经呈现出一定的退化趋势,因此没有进行额外的特征提取而直接将监测原始数据输入到所提的方法中进而进行产品的 RUL 预测,故本章就产品的监测数据而言是有别于第二章和第三章内容的。

由于所提取的初步特征具备完全不同的数值尺度,因此同样对上述所获得的特征集进行数据归一化处理。本章采用的是 z-score 正则化方法:

$$\hat{x}_d^k = \frac{x_d^k - \mu_d}{\sigma_d}, \quad \forall k, d \quad (4-1)$$

式中,  $x_d^k$  表示针对第  $k$  个样本的第  $d$  个特征量,  $\hat{x}_d^k$  表示经过归一化处理后的结果,  $\mu_d$  和  $\sigma_d$  分别表示特征集中针对第  $d$  个特征对应的均值和标准差。

最后,同样使用滑动时间窗技术对上述经过归一化的特征进行序列化处理,滑动时间窗处理技术在本文第三章 3.2.2 节进行了详细的介绍,此处不再说明。

### 4.3 基于 BLSTM 模型的产品混合剩余寿命预测方法

本章所提出基于 BLSTM 模型的产品混合 RUL 预测方法的实施流程如图 4-1 所示,主要包含数据采集与前处理、基于 BLSTM 模型的产品退化指标构建和基于 PF 算法的产品 RUL 预测三个不同的阶段,其中第一阶段数据采集与前处理已在 4.2 节中进行了详细的阐释。在第二阶段基于 BLSTM 模型的产品退化指标构建中,前处理后的序列特征数据被输入到 BLSTM 的网络结构中进行前向传播,并结合产品的离线监测数据和优化算法对模型进行训练,值得注意的是在模型训练过程中采用正则化方法以增强模型的泛化能力。当完成 BLSTM 模型训练后,测试产品经前处理的在线序列特征数据也被输入到模型中以实时预测产品的退化指标。在第三阶段基于 PF 算法的产品 RUL 预测中,阶段二所获取的产品退化指标预测输出

被进一步输入到测量方程中，此外基于领域内的物理知识所构建的产品经验退化模型被用作系统方程，粒子滤波算法被用来处理产品最新的退化指标预测输出数据，以实时更新模型参数并最终外推得到产品的 RUL 估计结果。

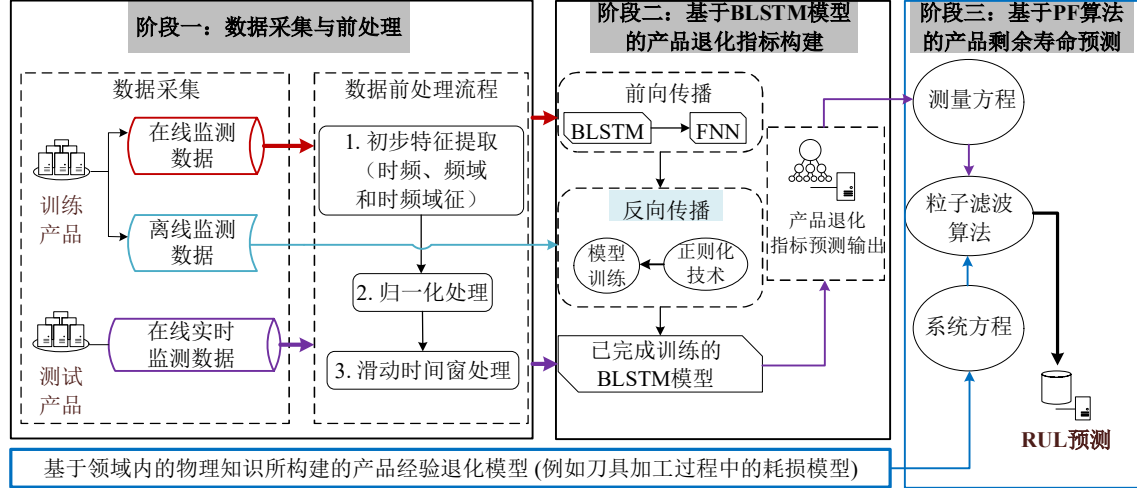


图 4-1 一种基于 BLSTM 模型的混合 RUL 预测方法实施流程

#### 4.3.1 基于 BLSTM 模型的产品退化指标的构建

本章所提出的基于 BLSTM 模型的网络结构如图 4-2 所示，从图中可以看出该网络结构主要包含两个部分：（1）提取隐藏在归一化处理后的输入特征序列数据中的时间相关性，该提取过程由一简单的 BLSTM 网络实现，具体地该 BLSTM 网络将分别从正向和逆向对输入特征序列数据的长时间依赖关系进行特征提取；（2）将 BLSTM 所提取的特征进一步输入到一全连接层中，并最终通过一线性回归层实现对产品退化指标的预测输出。具体描述如下：

**Part1:** 采集训练产品的多传感器监测原始数据且实施前处理后可得到训练集  $\{X^k, y^k\}_{k=1}^K$ ，其中  $X^k = [x_1^k, x_2^k, \dots, x_t^k, \dots, x_{t_{TW}}^k]$  表示第  $k$  个训练特征序列数据， $t_{TW}$  表示序列数据的长度， $x_t^k \in \mathbb{R}^{1 \times N}$  表示  $t$  时刻的归一化特征数据。此外， $y^k$  表示对应第  $k$  个训练特征序列数据  $X^k$  的产品退化状态，需要说明的是  $y^k$  的取值为序列数据终点  $t_{TW}$  时产品的离线退化状态测量数据。将归一化处理后的特征序列数据  $X^k$  输入到 BLSTM 网络中以提取序列数据的时序特征，所得的输出特征计算如下：

$$[h_1, h_2, \dots, h_t, \dots, h_{t_{TW}}] = f_1(X^k; \bar{\theta}_{BLSTM}, \bar{\theta}_{BLSTM}) \quad (4-2)$$

式中， $f_1(\cdot)$  表示 BLSTM 网络隐藏层函数，其在本文 3.3.2 节中已定义，BLSTM 网络的参数集由  $(\bar{\theta}_{BLSTM}, \bar{\theta}_{BLSTM})$  所刻画。另外，需要说明的是式(4-2)中给出了 BLSTM 网络对应  $X^k$  的所有输出特征序列数据，但是仅 BLSTM 网络第二层最后

时刻点的输出特征将用于输入到全连接层中，其计算过程如下：

$$\mathbf{h}_{t_{TW}} = \bar{\mathbf{h}}_{t_{TW}} \oplus \tilde{\mathbf{h}}_{t_{TW}} \quad (4-3)$$

式中， $\bar{\mathbf{h}}_{t_{TW}}$  和  $\tilde{\mathbf{h}}_{t_{TW}}$  分别表示正向和逆向 LSTM 网络在最后时刻点的输出特征， $\oplus$  表示按元素逐点求和运算。

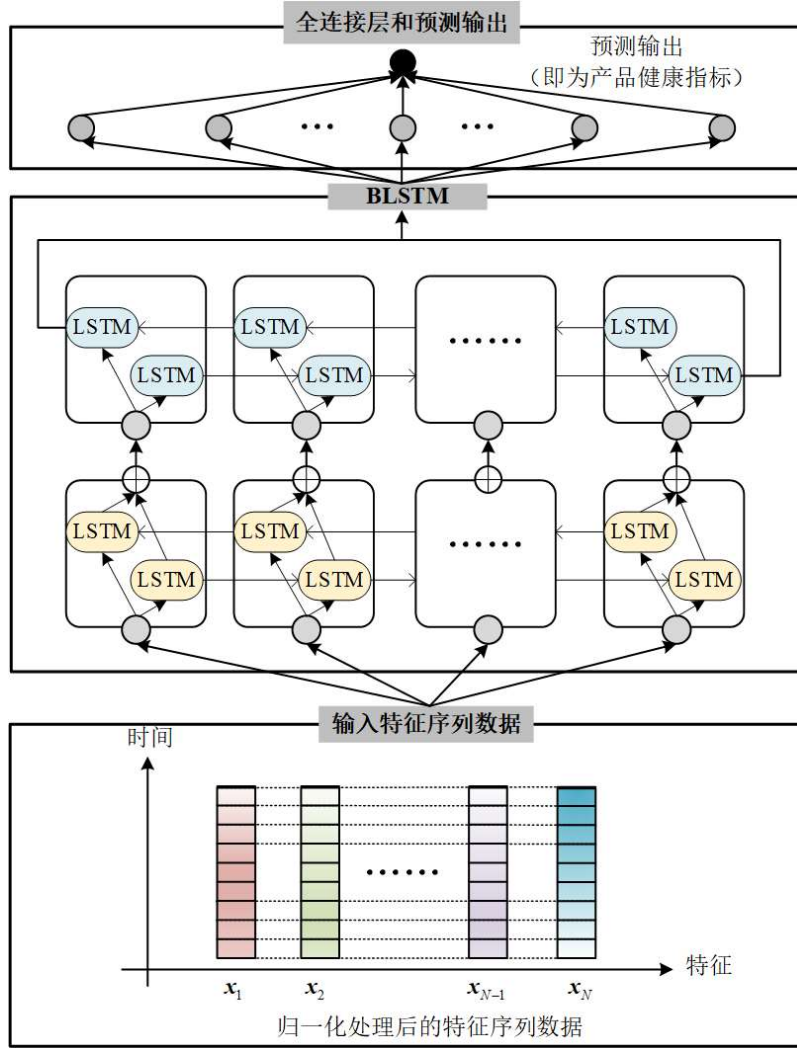


图 4-2 所提出的基于 BLSTM 模型的网络结构

**Part2:** 将 BLSTM 网络的输出特征  $\mathbf{h}_{t_{TW}}$  进一步输入到全连接层中以进一步提取更高层的特征，其对应计算过程如下：

$$\mathbf{o}^k = f_2(\mathbf{h}_{t_{TW}}; \boldsymbol{\Theta}_{FC}) = \mathbf{W}_F \mathbf{h}_{t_{TW}} + \mathbf{b}_F \quad (4-4)$$

式中， $\mathbf{o}^k$  为全连接层的输出特征， $\boldsymbol{\Theta}_{FC}$  为全连接层对应的参数集，包括权重矩阵  $\mathbf{W}_F$  和偏置向量  $\mathbf{b}_F$ 。最后，将全连接层的输出特征  $\mathbf{o}^k$  输入到一线性回归层以预测产品的退化状态，该过程可用如下公式描述：

$$\hat{y}_k = f_3(\mathbf{o}^k; \boldsymbol{\Theta}_R) = \mathbf{W}_R \mathbf{o}^k \quad (4-5)$$

式中,  $\hat{y}_k$  为产品的退化指标预测输出,  $\mathbf{W}_R$  为最后线性回归层的权重向量。

需要指出的是 **Part1** 和 **Part2** 部分均采用的是单层双向长短期记忆神经网络和全连接神经网络的表达式, 在实际应用中, 可以进行多层堆叠从而形成深度结构。但是深度神经网络结构则意味着更加复杂的模型, 因此需要更多的训练数据和资源完成对模型的训练。

上述部分刻画了模型前向传播的计算过程, 即从产品的在线监测数据至归一化处理后的特征序列数据到最终产品的退化指标预测输出。此外, 可以定义基于均方差的模型损失函数如下:

$$L(\boldsymbol{\Theta}) = \frac{1}{K} \sum_{k=1}^K \left( y^k - f_{\boldsymbol{\Theta}}(\mathbf{X}^k) \right)^2 \quad (4-6)$$

式中,  $y^k$  表示产品处于停机状态下离线测量的退化指标归一化处理后的值,  $\hat{y}_k = f_{\boldsymbol{\Theta}}(\mathbf{X}^k)$  表示当输入数据序列为  $\mathbf{X}^k$  时, 利用所提出的 BLSTM 网络模型计算所得的产品退化指标预测输出,  $f_{\boldsymbol{\Theta}}(\cdot)$  表示所提出的 BLSTM 网络模型具体由式(4-2)-(4-5)所定义。此外, 本部分也采用沿时反向传播 BPTT 算法, 并同样基于 Adam 方法对上述模型  $f_{\boldsymbol{\Theta}}(\cdot)$  进行优化。在模型训练过程中, 采用 Dropout 正则化方法防止模型过拟合以增强模型的泛化预测能力。

#### 4.3.2 融合粒子滤波算法的混合剩余寿命预测方法

对于某些复杂产品而言, 其表征内部退化状态的量往往由于一些实际的工程限制<sup>[93]</sup>很难直接在线获取。因此, 常采取一些在线监测技术作为辅助手段, 以实时获取表征产品退化状态的间接测量数据, 并基于所获取的间接测量数据和特定的模型完成对产品内部退化状态的精确估计。例如, 可以通过在线测量数控机床在加工过程中刀具的切削力信号、振动信号和声发射信号, 以间接对刀具的实时磨损状态进行推断。针对上述情况贝叶斯学习框架给出了一有效的解决方案。为了定义产品退化状态估计的问题, 首先基于产品的状态转移模型给出其退化状态随时间的变化如下<sup>[147]</sup>:

$$y_k = f(y_{k-1}, v_{k-1}) \Leftrightarrow p(y_k | y_{k-1}) \quad (4-7)$$

式中,  $k$  表示时刻点,  $f(\cdot)$  表示基于领域内的物理知识所构建的产品状态转移模型,  $y_{k-1}$  表示产品在  $k-1$  时刻内部退化状态,  $v_{k-1}$  表示一独立且同分布的过程噪音。在任一时刻点  $k$  可以获取有关产品的在线测量数据  $\mathbf{s}_k \in \mathbb{R}^{M \times p}$ , 通过观测模型进一步刻画在线测量数据  $\mathbf{s}_k$  与不可直接观测的产品内部退化状态  $y_k$  之间的非线性关系,

具体表述为：

$$\mathbf{s}_k = h(y_k, w_k) \Leftrightarrow p(\mathbf{s}_k | y_k) \quad (4-8)$$

式中， $h(\cdot)$ 为观测模型， $w_k$ 为一独立且同分布的测量噪音。

由于在线测量数据往往是来源于多个传感器在极高采样频率下采集的信号且其信号噪声比常常较低<sup>[93]</sup>，所以很难得到有关产品在线监测数据与其内部退化状态之间准确的显示解析模型。因此，本章先对产品的在线测量数据进行前处理，具体包括初步特征提取、归一化处理和序列化处理等步骤，该部分内容已在本章 4.2 节中进行了详细说明。此后在线监测数据可转化为一组归一化的特征序列数据  $\mathbf{x}_k = [\mathbf{x}_1^k, \mathbf{x}_2^k, \dots, \mathbf{x}_t^k, \dots, \mathbf{x}_{t_{TW}}^k] = f_{\text{Pre-processing}}(\mathbf{s}_k)$ 。最后，将所得归一化特征序列数据  $\mathbf{x}_k$  进一步输入到基于 BLSTM 网络的模型中，以完成对所提取特征  $\mathbf{x}_k$  与产品离线测量的内部退化状态  $y_k$  之间的非线性建模，该部分内容在本章 4.3.1 节中进行了介绍。上述基于 BLSTM 网络的产品观测模型可简要表述为：

$$z_k = h_{\text{BLSTM}}(y_k, \mathbf{x}_k, w_k) \Leftrightarrow p(z_k | y_k) \quad (4-9)$$

式中， $z_k$  表示经过 BLSTM 网络处理后的产品退化指标预测输出值（即前文中的  $\hat{y}_k$ ）， $y_k$  表示产品处于停机状态下基于离线测量技术所获取的产品内部退化状态， $w_k$  为测量噪音，具体可以通过 Bagged Ensemble 的方法<sup>[148-149]</sup>进行确定，这里为了模型的简洁和计算的便利假定其服从正态分布<sup>[109]</sup>。

在贝叶斯学习的框架内，主要任务为在给定实时在线监测数据  $\mathbf{s}_{k+1}$  的情况下计算产品退化状态的后验概率分布  $p(y_{k+1} | \mathbf{s}_{k+1})$ 。需要指出的是在线监测数据  $\mathbf{s}_{k+1}$  经过前处理后输入到 BLSTM 网络中最终得到产品退化指标预测输出值  $z_{k+1}$ ，因此实质上求得的是后验概率分布  $p(y_{k+1} | z_{k+1})$ 。总体来说，后验概率分布  $p(y_{k+1} | z_{k+1})$  可以通过迭代执行预测步骤和更新步骤来获得。假定先验概率分布  $p(y_k | z_k)$  在  $k$  时刻已知，预测步骤主要利用产品的状态转移模型（式(4-7)）并基于 Chapman-Kolmogorov 方程去估计概率密度函数  $p(y_{k+1} | z_k)$ ，具体表示如下：

$$p(y_{k+1} | z_k) = \int p(y_{k+1} | y_k) p(y_k | z_k) dy_k \quad (4-10)$$

在  $k+1$  时刻，新的在线测量数据  $\mathbf{s}_{k+1}$  经过前处理后输入到 BLSTM 网络得到产品的退化指标预测输出  $z_{k+1}$ ，其次在更新步骤中，后验概率分布  $p(y_{k+1} | z_{k+1})$  即可基于贝叶斯准则更新如下：

$$p(y_{k+1} | z_{k+1}) = \frac{p(z_{k+1} | y_{k+1}) p(y_{k+1})}{p(z_{k+1})} = \frac{p(z_{k+1} | y_{k+1}) p(y_{k+1} | z_k)}{p(z_{k+1} | z_k)} \quad (4-11)$$

式中,  $p(z_{k+1}|y_{k+1})$  表示  $k+1$  时刻的似然函数且可由式(4-9)进行定义, 概率密度分布  $p(y_{k+1}|z_k)$  则可通过预测步骤计算得到,  $p(z_{k+1}|z_k)$  表示归一化因子, 具体计算如下:

$$p(z_{k+1}|z_k) = \int p(z_{k+1}|y_{k+1})(y_{k+1}|z_k) dy_{k+1} \quad (4-12)$$

由于大多数产品实际的状态转移模型均为非线性模型, 所以很难得到以上相关表达式的解析解。因此引入粒子滤波以解决上述问题, 并通过实施基于蒙特卡洛仿真的递推贝叶斯滤波来近似后验概率密度分布  $p(y_{k+1}|z_{k+1})$ 。假定  $k$  时刻点  $p(y_k|z_k)$  已知且可通过一系列离散值  $y_k^1, \dots, y_k^i, \dots, y_k^N$  和对应的权重值  $\omega_k^1, \dots, \omega_k^i, \dots, \omega_k^N$  所刻画, 具体表示如下:

$$p(y_k|z_k) \approx \sum_{i=1}^N \omega_k^i \delta(y_k - y_k^i) \quad (4-13)$$

式中,  $\sum_{i=1}^N \omega_k^i = 1$ ,  $\delta(\cdot)$  表示狄拉克  $\delta$  函数。

在粒子滤波框架内, 式(4-10)中所定义的预测步骤可改写如下:

$$p(y_{k+1}|z_k) \approx \sum_{i=1}^N \omega_k^i \delta(y_k - y_k^i) p(y_{k+1}|y_k) = \sum_{i=1}^N \omega_k^i p(y_{k+1}|y_k) \quad (4-14)$$

式中,  $i$  表示第  $i$  个粒子,  $N$  表示粒子的总数,  $p(y_{k+1}|y_k)$  可直接基于产品的状态转移模型计算得到。

采集  $k+1$  时刻的在线测量数据  $s_{k+1}$  且经过数据前处理后输入到 BLSTM 模型, 最终可得到预测输出  $z_{k+1}$ 。在更新步骤中, 基于式(4-9)权重值  $\omega_k^i$  即可采用重要性采样的方法实现递归更新, 具体在更新步骤中粒子权重的更新过程可表示为:

$$\omega_{k+1}^i \propto \omega_k^i p(z_{k+1}|y_k^i) \quad (4-15)$$

式中,  $p(z_{k+1}|y_k^i)$  表示在给定第  $i$  个粒子的状态  $y_k^i$  下得到预测输出  $z_{k+1}$  的似然函数。

为了避免粒子滤波算法实施过程中粒子的退化问题 (即随着循环次数的增加, 粒子的权重集中在少数粒子上), 采用基于逆累积函数密度的重采样方法<sup>[8]</sup>以获得等权重的粒子。具体实施细节如下: (1) **Step1**: 基于似然函数  $p(z_{k+1}|y_k^i)$  构建一累积密度函数; (2) **Step2**: 从均匀分布  $U(0,1)$  抽样一随机值并作为 **Step1** 中的累积密度函数的取值; (3) **Step3**: 将具备与 **Step2** 中 CDF 值最接近的粒子取出以作为重采样之后的结果; (4) **Step4**: 重复实施  $N$  次 **Step1-Step3** 的计算流程, 所获得的等权重  $N$  个重采样粒子即可用于近似后验概率密度分布  $p(y_{k+1}|z_{k+1})$ 。

## 4.4 算例分析与方法验证

### 4.4.1 数据集简介

为了验证本章所提出的 BLSTM-PF 混合 RUL 预测方法的有效性,本部分将采用 PHM 协会于 2010 年公开的数据集<sup>[150-151]</sup>,该数据集主要针对的是解决有关高速铣削铣床刀具磨损状态预测的问题。所采用的实验平台如图 4-3 所示,其中实验所采用铣床型号为 Rödgers Tech RFM760,刀具和工件分别采用的是球头硬质合金铣刀和不锈钢 HRC52。实验所采用的加工参数如下:(1)主轴转速为 10400 rpm;(2)进给速率为 1555 mm/min;(3)铣削径向深度 Y 和轴向深度 Z 分别为 0.125mm 和 0.2mm;(4)工件长度为 108mm。

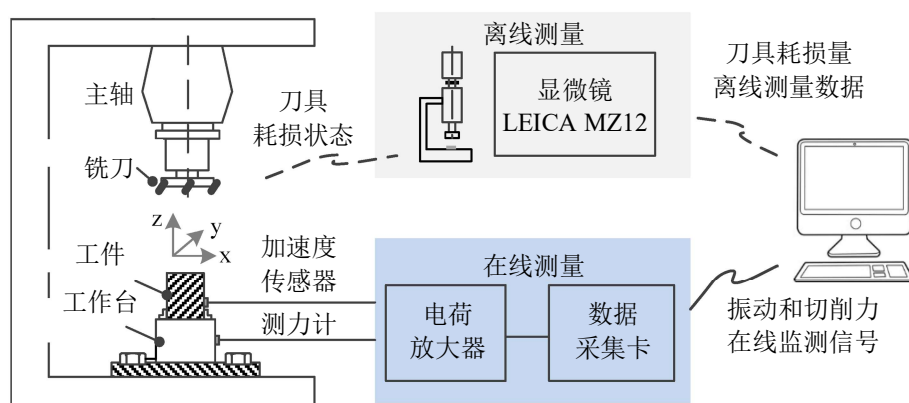


图 4-3 高速铣削铣床刀具磨损状态监测实验平台

实验采集了三把球头铣刀(记为 C1, C4, C6)的端面铣削监测数据作为训练数据集,具体针对每把铣刀,共计含有 315 次端面铣削的传感器在线监测数据,且每次铣削后铣刀的三个切削刃上的磨损量  $l_1$ ,  $l_2$ ,  $l_3$  均进行离线测量以作为模型训练的标记值。如图 4-3 所示,工件与工作台之间布置了一三分量测力传感器(型号为 Kistler 9265B 三向测力仪)以采集电压的方法分别对三个方向上的切削力进行实时采集,并且在工件上还布置了振动传感器(型号为 Kistler 8635C 加速度传感器)以实时监测刀具在加工过程中三个方向上的振动信号。本章所采用的刀具在线监测数据涵盖三个方向的切削力信号和三个方向上的振动信号,共计六个通道的信号,并且各个通道的信号均以 50KHZ 的采样频率(采集卡型号为 NI DAQ PCI 1200)被采集到 PC 端中。此外,采用 LEICA MZ12 显微镜以离线测量一次完整走刀后切削刃上的磨损量,离线测量数据同样地被保存到 PC 端中。

## 4.4.2 数据前处理

本章所采用的刀具磨损实验记为 *C1*, *C4* 和 *C6*, 三个实验中均涵盖了来自于六个信号通道的在线监测数据 (记为 *forece\_x*, *forece\_y*, *forece\_z*, *vibration\_x*, *vibration\_y*, *vibration\_z*), 且每个实验共计采集了 315 组刀具在加工过程在线监测数据阵列和其对应的刀具磨损量离线测量数据。由于传感器较高的采样频率, 因此每组在线监测数据阵列均包括了大量的数据点, 所有 315 组在线监测数据阵列涵盖了超过 65,000,000 个数据点。为清晰阐述在线监测数据, 对每组在线监测数据阵列仅提取 25 个等距离的数据监测点, 并将所有的 315 组所提取的数据监测点有序地拼接在一起, 其结果如图 4-4 所示。

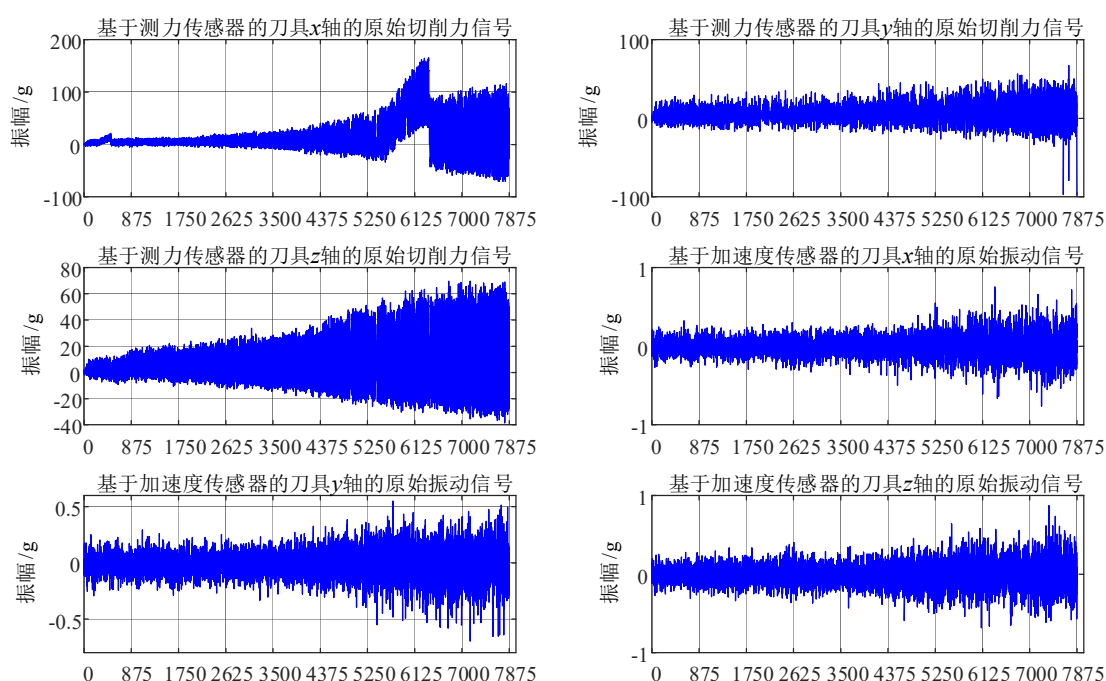


图 4-4 基于刀具磨损实验 *C1* 的六个信号通道监测数据简化图

由于六个信号通道的在线监测数据本质上属于时间序列数据, 因此首先对其时域特征进行提取, 主要包括均值 (Mean)、方差 (Variance)、均方根 (RMS)、最大值 (Max)、峰峰值 (P2P)、偏斜度 (Skewness)、峭度 (Kurtosis)。其次, 基于快速傅里叶变换技术将原始的六个通道的时域信号转换为对应的频域信号, 并进一步提取其频域特征, 主要包括谱偏斜度 (Spectral Skewness)、谱峭度 (Spectral Kurtosis)、谱功率 (Spectral Power)。对于六个信号通道, 其中每个信号通道均提取了十个时域及频域内的相应特征, 图 4-5 给出了经归一化处理后的 60 个特征。

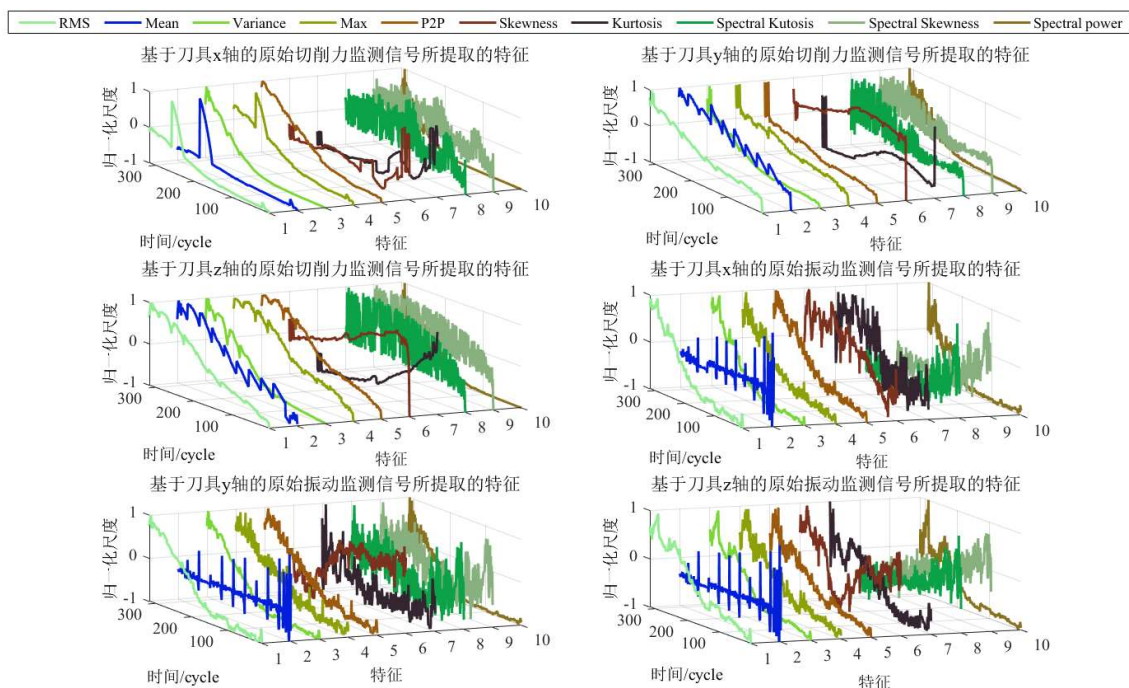


图 4-5 归一化处理后的特征示意图

所提取的六个信号通道的时域频域特征降低了在线监测原始信号的数据规模, 且将有效地提升后续 BLSTM 模型训练和预测的效率。从图 4-5 也可以看出, 所提取的某些特征彼此是相关的, 例如针对刀具  $x$  轴、 $y$  轴和  $z$  轴的原始切削力所提取的均方根特征和均值特征。此外, 所提取的某些特征在整个监测周期内并不具备很好的单调一致性, 例如针对刀具  $x$  轴、 $y$  轴和  $z$  轴原始振动监测信号所提取的均值特征与谱峭度特征。因此, 传统的方法<sup>[93, 97]</sup>将会依次实施特征选择和降维的处理步骤以选取最具表征性的特征组合, 并基于所挑选的特征组合进一步实施特征融合与回归分析。相比之下, 本章所提出的基于 BLSTM 模型的方法将特征选择、降维、融合与回归分析等步骤结合到统一的框架内, 通过模型训练调整网络权重参数的方式自动且自适应地实现了上述冗杂的处理流程。

本章所提出的基于 BLSTM-PF 混合方法针对刀具 RUL 预测的实施流程如图 4-6 所示。首先, 提取来自六个信号通道的传感器监测原始数据的 60 个时域和频域特征, 并进行数据前处理。其次, 将前处理后的归一化训练特征序列和其对应的归一化离线测量数据输入到 BLSTM 模型用于训练模型, 当模型训练完成后即可将前处理后的归一化测试特征序列输入到 BLSTM 模型, 用于刀具磨损状态的实时在线预测。此后, 将刀具磨损状态的预测输出值用于粒子滤波框架内的测量方程, 并结合刀具磨损的状态转移模型以实现刀具磨损状态的估计与模型的参数更新。最后, 通过外推刀具磨损状态至失效阈值处即可得到刀具的 RUL 估计值。需要指

出的是，由于刀具磨损状态已进行了归一化处理，所以此处的失效阈值即为 1。另外，Robert 等<sup>[93]</sup>针对刀具在加工过程中的磨损问题提出了一经验指数模型，具体形式如下：

$$dy/dt = Cy^m \quad (4-16)$$

式中， $y$  表示刀具磨损量， $dy/dt$  表示刀具在加工过程中的磨损速率， $C$  和  $m$  表示与工件和刀具等相关的参数，由于刀具的磨损量  $y$  是在离散时刻点进行测量，因此式(4-16)可以进一步改写为以下形式：

$$y_k = C_{k-1}y_{k-1}^{m_{k-1}} + y_{k-1} \quad (4-17)$$

上式即为本章所采用的的刀具磨损的状态转移模型，其中  $C_{k-1}$  和  $m_{k-1}$  的概率密度分布可以刻画刀具在加工过程中的随机特性，且可以根据刀具磨损状态的预测输出值进行更新。

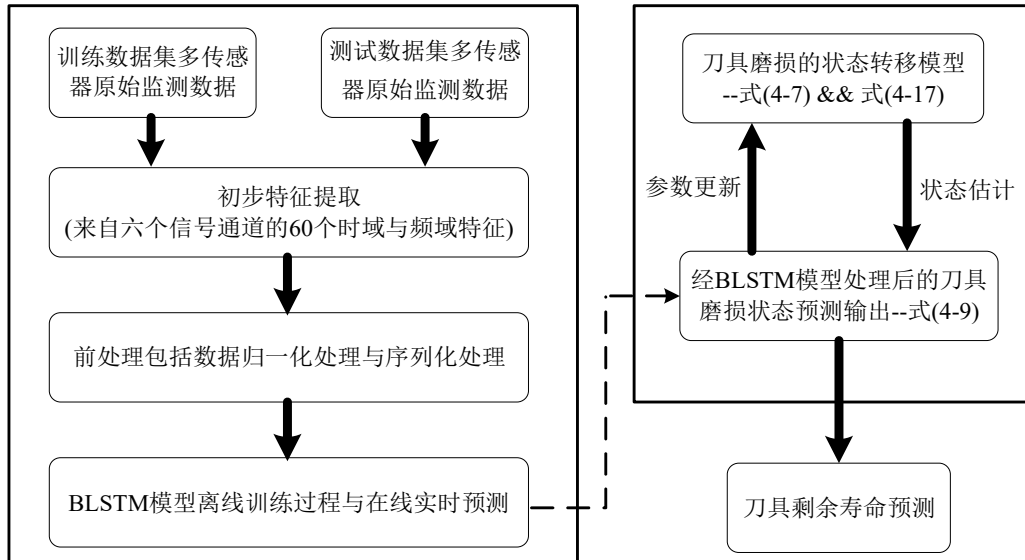


图 4-6 基于 BLSTM-PF 混合方法的 RUL 预测实施流程

#### 4.4.3 方法验证与结果讨论

本部分将首先验证所提出的基于 BLSTM 模型的方法对减少传统方法复杂的实施流程的有效性，其中所对比的传统方法实施流程主要包括以下四个关键步骤：

**Step1:** 特征提取。提取相同的经归一化处理的 60 个特征（如图 4-5 所示）。

**Step2:** 特征选择。由于所提取的 60 个特征存在一定的冗余特性且某些特征在整个监测周期内不具备很好的单调一致性，因此实施特征选择步骤以选取最优特征集。本部分基于两个常用特征选择的指标（即相关性和单调一致性）进行最优特征集选取，其中相关性指标是通过计算每个所提取特征与刀具耗损量之间的

Pearson 相关系数而得到，且上述两个指标的平均值被用作特征选择的评判分数。

**Step3:** 特征融合。选用主成分分析法 (Principle Component Analysis, PCA) 以对所选择的特征进行融合分析。

**Step4:** 回归分析。采用 SVR 以完成对融合特征与刀具归一化磨损量之间映射关系的回归分析。

本部分实验选取 C1 作为测试样本，C4 和 C6 作为训练样本。首先分析了有关特征选择步骤对最终刀具耗损量预测结果的影响，其对应结果如表 4-1 所示。表中最优特征集表示经过特征选择步骤后所得的不同的最优特征组合（依据特征选择评判分数对所有 60 个特征进行排序后的结果），此后采用 PCA 以处理表中不同的最优特征集为一融合特征，且该融合特征最后被输入到 SVR 模型中以预测刀具的耗损量。从表中可以看出，基于评判分数前五个特征所组成的最优特征集可以取得最好的预测性能。

表 4-1 特征选择步骤对预测结果的影响

| 最优特征集 | 15   | 12   | 10   | 8    | 5    |
|-------|------|------|------|------|------|
| RMSE  | 28.8 | 24.3 | 26.8 | 14.6 | 13.9 |
| MAE   | 21.0 | 22.2 | 23.5 | 13.2 | 12.8 |

其次，有关特征融合步骤对最终刀具耗损量预测结果的影响也开展了相应的研究，其结果如表 4-2 所示。需要说明的是表 4-2 中所采用的最优特征集为基于评判分数前八个特征所形成的特征组合，且经过 PCA 处理后所得的不同主成分分别经过 SVR 模型进行回归分析。从表中可以看出基于最高主成分（保留了 95% 最优特征集的累积方差）所得的融合特征能得到最好的预测性能，且基于前两个或前三个主成分所得的融合特征的预测性能均显著下降。

表 4-2 特征融合步骤对预测结果的影响

| 所保留的主成分数量 | 1—95% | 2—99% | 3—99.5% |
|-----------|-------|-------|---------|
| RMSE      | 14.6  | 33.8  | 35.4    |
| MAE       | 13.2  | 18.6  | 25.3    |

再次，还需进一步确定有关 SVR 模型相关参数的取值。本部分主要分析 SVR 模型的四个参数：（1）SVR 模型所用核的类型 (Kernel)；（2）回归误差项的惩罚系数 ( $C$ )；（3）SVR 模型回归误差的上限 (Epsilon)；（4）所选取核的系数 (Gamma)。应用网格搜索法确定上述四个与 SVR 模型相关的参数，且各参数具体建议取值如表 4-3 所示。

表 4-3 SVR 模型参数设置

| 参数   | Kernel   | C   | Epsilon | Gamma |
|------|----------|-----|---------|-------|
| 建议取值 | Gaussian | 500 | 0.1     | 0.1   |

最后,有关传统方法、基于 BLSTM 模型的方法和一些公开发表方法就刀具耗损量预测任务的性能对比结果如表 4-4 所示。表中所采取的对比方法分别包括基于线性回归 (LR) 模型<sup>[133]</sup>、多层感知机 (MLP) 模型<sup>[133]</sup>、循环神经网络 (RNN) 模型<sup>[133]</sup>的方法和上述传统方法。需要指出的是针对传统方法,基于评判分数前五个特征所组成的最优特征集被选取且进一步被 PCA 处理以形成一融合特征,并采用基于表 4-3 的 SVR 模型实施该融合特征与刀具耗损量之间的回归分析。

表 4-4 所提方法与相关方法的预测性能对比

| 方法                   | MAE  | RMSE |
|----------------------|------|------|
| LR <sup>[133]</sup>  | 24.4 | 31.1 |
| MLP <sup>[133]</sup> | 24.5 | 31.2 |
| RNN <sup>[133]</sup> | 13.1 | 15.6 |
| 传统方法                 | 12.8 | 13.9 |
| BLSTM                | 7.9  | 10.2 |

从表 4-4 中可见,基于 BLSTM 模型的方法均好于传统的方法和其余三种对比方法。相比于传统方法,本章所提方法将特征选择、融合与回归分析均整合到统一的 BLSTM 模型中,因此可有效减少传统方法繁琐的实施流程。另外,从表 4-1 至表 4-3 可见,传统方法中各个具体实施步骤均对最终预测结果存在显著的影响。相反,由于本章所提出基于 BLSTM 模型的方法端对端的处理思路,因此其预测性能相比于传统方法更加稳健。

前面部分验证了本章所提的 BLSTM 模型基于刀具在线监测数据预测刀具归一化磨损量(这里可视为刀具的退化指标)的有效性。其后,结合所获得刀具退化指标、刀具磨损的状态转移模型及其基于 BLSTM 的观测模型,将粒子滤波算法用于预测刀具退化指标随时间的演变规律,并且当其超过失效阈值,可以进一步得到刀具的 RUL 估计结果。本部分主要分析所提出的 BLSTM-PF 混合 RUL 预测方法的短期和长期的 RUL 预测能力,且其对应的预测结果如图 4-7 所示。图中提前  $N$  步预测表示在前  $(300-N)$  个刀具退化指标可以利用 BLSTM 模型获取的情况下,预测其退化指标后续随时间的演进规律。

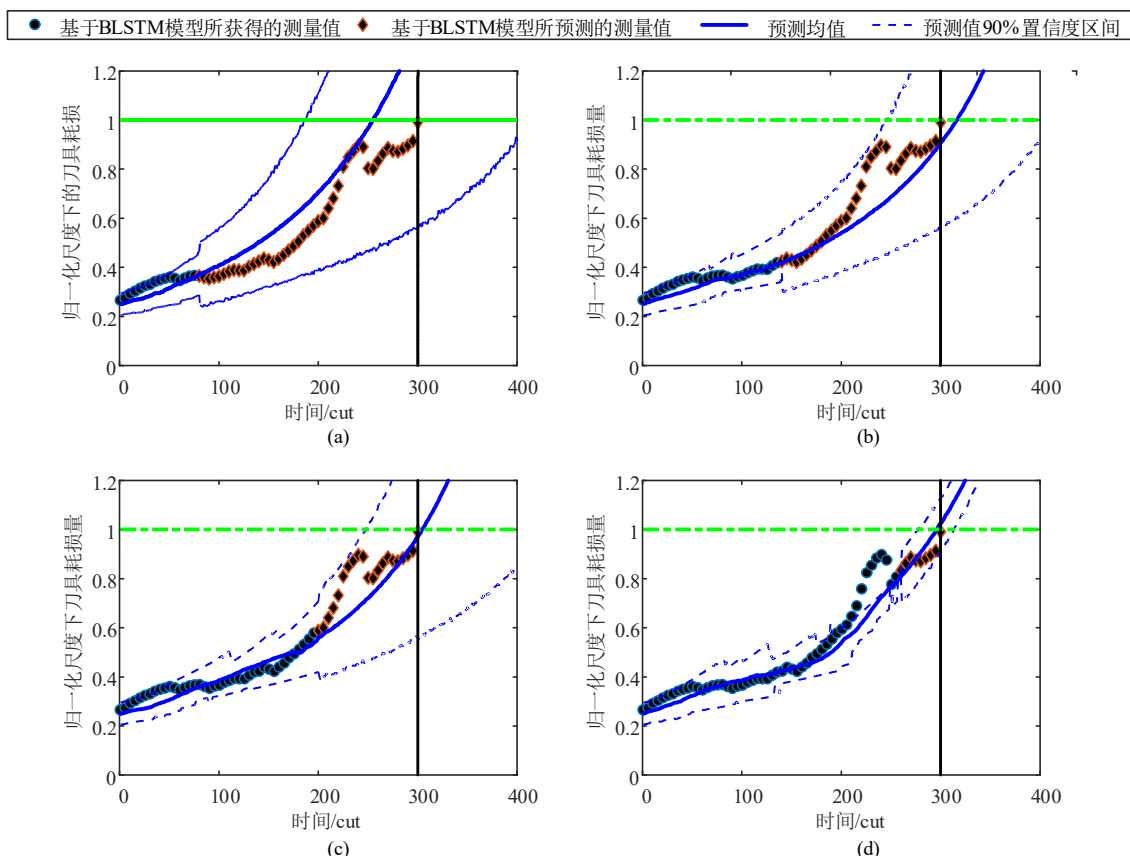


图 4-7 BLSTM-PF 混合方法在不同情况下 RUL 预测的结果。(a)提前 220 步预测结果；(b)提前 160 步预测结果；(c)提前 100 步预测结果；(d)提前 40 步预测结果

从图中可以看出，对于 BLSTM-PF 混合方法长期预测的情况，例如针对提前 220 步和 160 步的刀具耗损预测的案例，由于有限的刀具磨损量观测值被 PF 算法所利用，因此其对应的预测均值偏离了 BLSTM 模型所预测的刀具退化指标，其最终的 RUL 估计结果与真值之间存在一定偏差，并且其相应 90%置信度下的预测区间较宽。另外，对于 BLSTM-PF 混合方法短期预测的情况，例如针对提前 100 步和 40 步的刀具耗损预测的案例，由于更多的刀具磨损量观测值被 PF 算法所利用，因此其对应的预测均值更加靠近 BLSTM 模型所预测的刀具退化指标，最终的 RUL 估计结果也与真实值比较接近，并且其相应 90%置信度下的预测区间较窄。为了全面地验证 BLSTM-PF 方法 RUL 预测的有效性，基于核密度估计技术近似刀具在等间隔观测时间下的 RUL 概率密度分布，其结果如图 4-8 所示。

为了清晰展示，图 4-8 仅给出了在 5 个等间隔观测时刻点上的刀具 RUL 概率密度分布的估计结果，表 4-5 进一步给出了在更多等间隔观测时刻点上刀具的 RUL 估计结果，具体包括其估计均值和 95%置信区间 (Confidence Interval, CI)。基于图 4-8 和表 4-5 中结果，具体讨论如下：

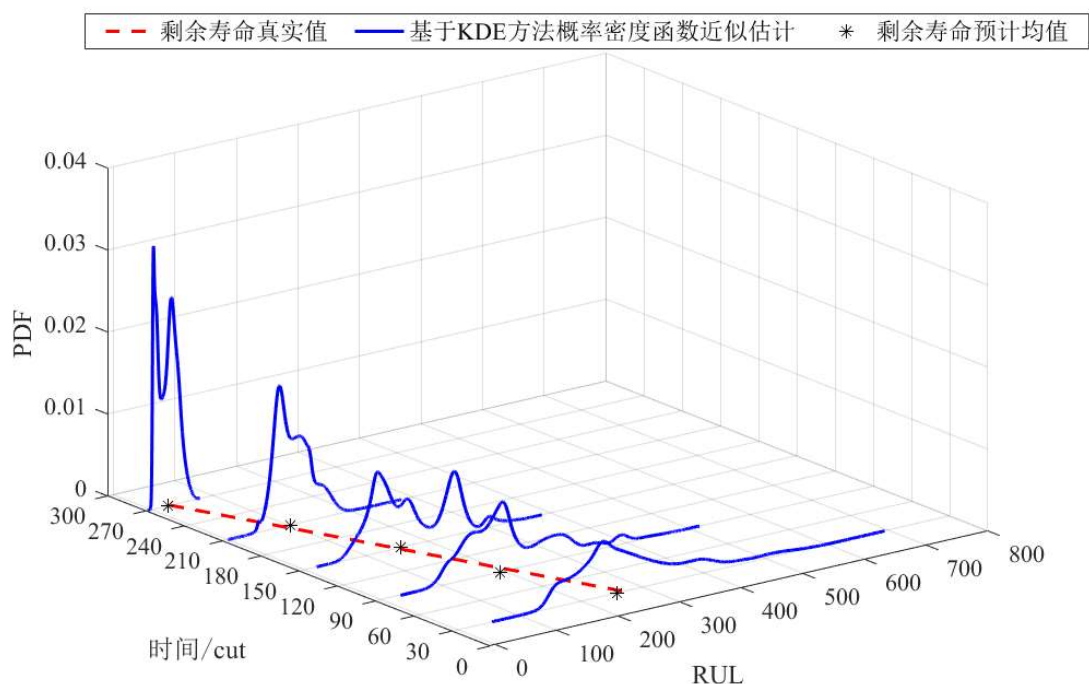


图 4-8 等间距观测时间下的刀具 RUL 概率密度分布估计结果

表 4-5 不同观测时刻点的 RUL 预测结果

| 观测时刻点 | 真实 RUL | RUL 估计均值 | 95% CI    |
|-------|--------|----------|-----------|
| 30    | 270    | 260      | [163,444] |
| 60    | 240    | 228      | [125,389] |
| 90    | 210    | 194      | [112,358] |
| 120   | 180    | 172      | [95,264]  |
| 150   | 150    | 157      | [90,267]  |
| 180   | 120    | 135      | [76,223]  |
| 210   | 90     | 101      | [67,160]  |
| 240   | 60     | 48       | [25,76]   |
| 270   | 30     | 26       | [3,51]    |

(1) 在观测区间 30~180 内，由于在线监测数据有限，因此本章所提出的 BLSTM-PF 混合方法所预测的 RUL 均值一定程度上偏离了对应真实值，并且其对应的 95%置信区间范围较大。需要说明的是该部分预测结果可以反映 BLSTM-PF 混合预测方法的长期预测性能。

(2) 在观测区间 210~270 内，由于刀具在加工过程中大部分的在线监测数据已知，所以混合方法能更好地预测刀具的 RUL，且对应的 95%置信区间范围更窄。该部分预测结果反映了 BLSTM-PF 混合预测方法的短期预测性能。

(3) 从表 4-5 还可以进一步看出针对任一观测时刻点, 刀具的真实 RUL 均落在了 95%置信区间内。因此, 相比于当前绝大多数仅能给出产品 RUL 点估计结果的基于深度学习的相关方法<sup>[61, 65, 67, 70]</sup>, 基于本章所提出的 BLSTM-PF 混合 RUL 预测方法所制定的刀具预防性维护决策将会更可靠且稳健。

## 4.5 本章小结

本章基于双向长短期记忆神经网络和粒子滤波技术, 提出了一种产品混合剩余寿命预测方法, 所提出的方法充分利用了双向长短期记忆神经网络对时间序列数据的处理优势, 同时将粒子滤波技术结合到所提方法的框架内以突破当前基于深度学习的方法仅能给出产品 RUL 点估计结果的限制。与传统的混合 RUL 预测方法相比, 本章所提方法由于采用了深度学习技术, 因此极大地简化了特征选择、降维、融合与回归分析等冗杂且耗时的处理流程。通过算例分析, 首先阐释了本章所提出的基于 BLSTM 模型的方法在预测刀具退化指标任务上的有效性, 其次通过对刀具的 RUL 预测进一步验证了本章所提出的 BLSTM-PF 混合 RUL 预测方法的有效性。最后, 由于本章所提方法部分解决了当前基于深度学习的 RUL 预测方法的局限性, 因此本章研究将进一步推进深度学习技术在某些对安全性要求较高产品上的 RUL 预测的工程应用。

## 第五章 基于深度卷积神经网络和 Bootstrap 的剩余寿命预测方法

### 5.1 引言

第四章所提出的基于双向长短期记忆神经网络的混合 RUL 预测方法，通过结合双向长短期记忆神经网络和粒子滤波算法能对产品 RUL 进行有效的区间估计。但是，在粒子滤波框架下需建立产品的状态转移模型。因此，当一些复杂产品的状态转移模型不可知时，第四章所提方法将会面临应用限制。

针对当前绝大多数基于深度学习模型（DL）的 RUL 预测方法，其所利用的模型输入数据主要分为以下两类，分别是时间序列数据和图像数据。其中，时间序列数据指原始的传感器监测数据，例如航空发动机退化仿真数据（具体可参考第三章内容及其研究工作<sup>[61-62, 65-66, 70]</sup>），或基于原始传感器监测数据所提取的有关时域、频域和时频域内的对应特征（具体可参考第四章内容及其研究工作<sup>[63-64, 95, 133, 146]</sup>）。此外，这里图像数据是指基于原始传感器监测数据和相关算法所提取的有关图像特征，例如 Zhu<sup>[67]</sup>和 Li<sup>[68]</sup>等分别基于连续小波变换（Continuous Wavelet Transform, CWT）和短时傅里叶变换对轴承的原始振动信号进行处理以提取其对应的时频表征（Time-Frequency Representations, TFRs），由于所提取的 TFRs 本质上即一矩阵数据，因此可以将其视为一图片数据。总结相关研究工作发现，当前大多数基于 DL 模型的方法仅利用时间序列数据或图片数据对产品的 RUL 进行预测，很少有结合上述两类数据的基于 DL 模型的 RUL 预测方法的相关研究。其次，绝大多数基于 DL 模型的 RUL 预测方法均只能给出产品 RUL 点估计结果，尽管第四章给出了结合 DL 模型与粒子滤波算法的混合方法，但是该方法在实际应用中依然存在诸多应用限制。

针对上述基于 DL 模型的方法和混合方法在实际应用中所面临的挑战，本章提出基于深度卷积神经网络（DCNN）和 Bootstrap 的 DCNN-Bootstrap 剩余寿命预测方法。一方面，本章所提出的基于 DCNN 的网络结构能同时利用从原始传感器监测数据中所提取的时间序列特征和对应的图片特征，以进一步增强基于 DL 方法的 RUL 预测精度。另一方面，由于采用了 Bootstrap 的实施框架，因此本章所提出的 DCNN-Bootstrap 方法能突破混合方法（即第四章主要内容）需要建立产品退化模型的限制，进一步拓宽基于 DL 模型的方法在复杂产品 RUL 预测中的应用范围。

## 5.2 数据前处理

### 5.2.1 首次预测时间识别

针对某些产品例如滚动轴承，其振动测量信号在整个监测周期内既包括了轴承处于健康阶段的数据，同时也包含了轴承处于退化阶段的数据。图 5-1 为某滚动轴承在整个监测周期内的振动测量原始信号，从图中可以看出在红线以前其对应振动测量信号的振幅变化不大，相反在红线以后其对应振动测量信号的振幅呈现逐步增大的趋势。因此，可以进一步根据红线，即轴承首次预测时间（First Prediction Time, FPT）所处位置将轴承整个监测周期依次划分为健康状态和退化状态。

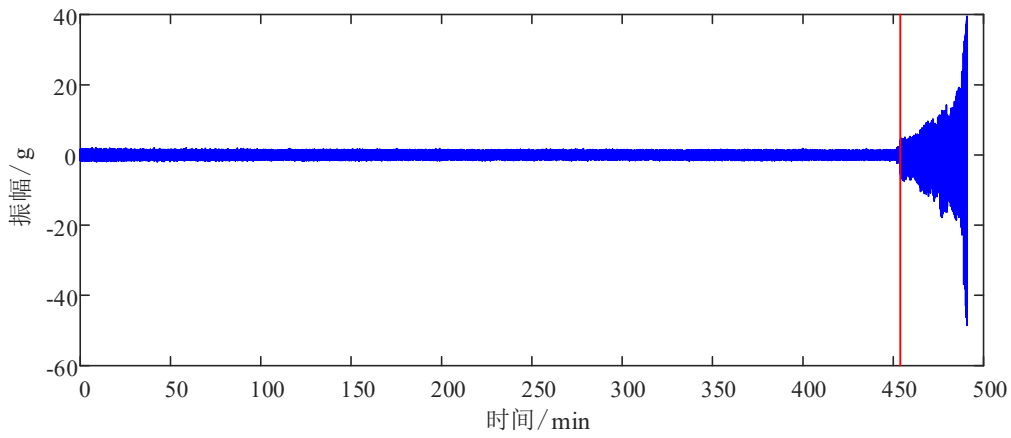


图 5-1 滚动轴承振动测量信号示意图

由于轴承峭度（Kurtosis）特征对其早期故障极为敏感<sup>[152]</sup>，因此采用雷亚国<sup>[152]</sup>等构建的基于振动信号峭度特征的方法以识别产品的首次预测时间，该方法实施流程如下<sup>[152]</sup>：

**Step1:** 根据对产品整个监测周期内振动测量信号的分析以确定产品处于早期健康状态的监测区间（如图 5-1 中监测区间 0~450），并进一步计算产品处于该健康状态阶段的 Kurtosis 的均值  $\mu$  与方差  $\sigma$ 。

**Step2:** 计算 Kurtosis 的  $3\sigma$  区间，即  $[\mu - 3\sigma, \mu + 3\sigma]$ ，并定义当所获取的 Kurtosis 值超过  $3\sigma$  区间即为异常状态点。

**Step3:** 定义一阈值  $i$  以最终确定 FPT 的值，例如  $i=2$  表示当连续出现两个异常状态点时，即可认为产品进入了退化阶段，具体可表示如下：

$$\{|k_{t-i} - \mu| > 3\sigma\}_{i=0,1,2} \quad (5-1)$$

式中， $k_{t-i}$  表示根据第  $(t-i)$  个振动数据序列计算得到的对应 Kurtosis 值，满足式 (5-1) 所确定的  $t$  即为所需的 FPT 值。由于设置了一阈值  $i$  的缘故，因此该方法可以

有效地消除随机噪音对 FPT 的影响。图 5-2 给出了上述 FPT 识别流程的示意图。

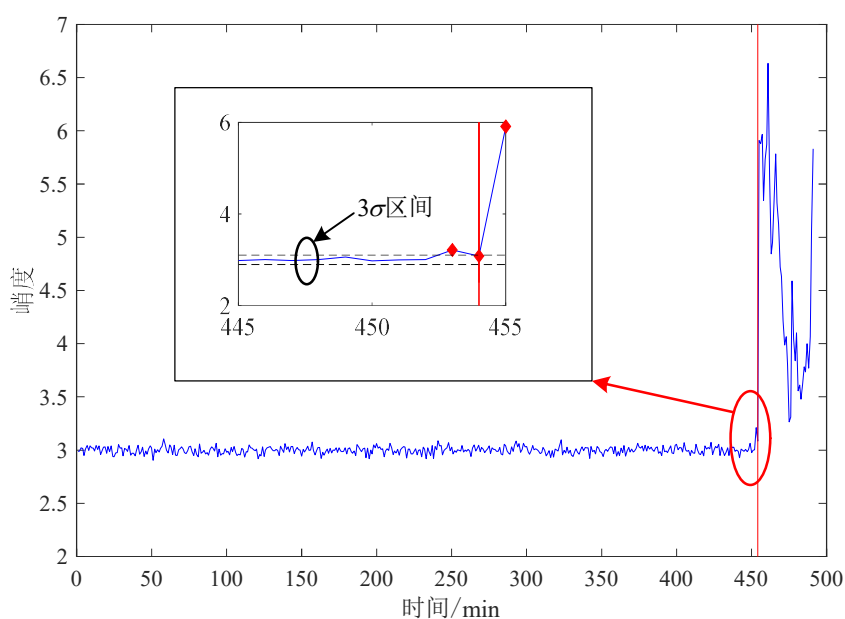


图 5-2 轴承首次预测时间识别示意图

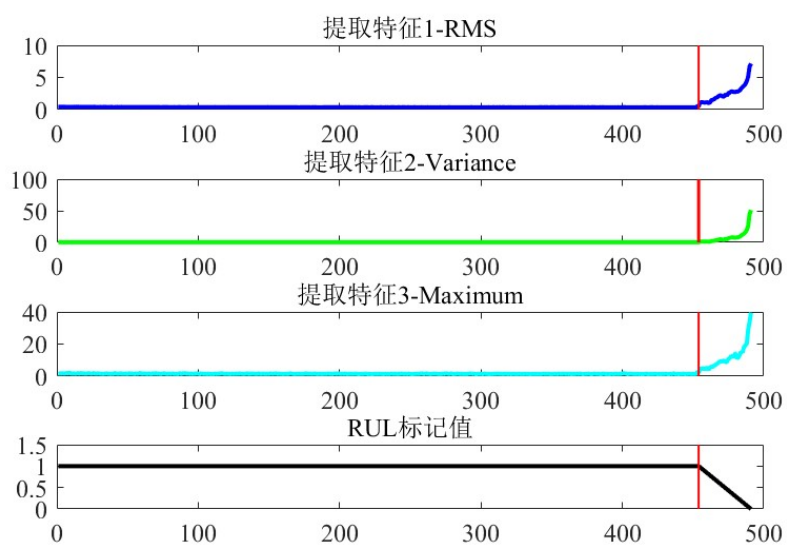


图 5-3 提取特征与 RUL 标记值示意图

当 FPT 值确定后,即可进一步提取有关轴承的时域、频域和时频域特征,并对其真实的 RUL 值进行标记。图 5-3 给出了所提取的三个特征与 RUL 标记值的示意图,从图中可以看出,所提取的特征在 FPT 之后呈现一定的退化趋势,因此该部分数据将会被用作模型输入数据。此外,需要进一步说明的是本章采取常用的线

性递减标记的方法<sup>[68]</sup>对 RUL 进行标记。

### 5.2.2 基于连续小波变换的时频表征提取

基于数据驱动的 RUL 预测方法其中一关键步骤就是提取能有效刻画产品退化状态的特征，例如振动信号时域内的均方根特征、方差特征、峭度和偏斜度等特征，及其频域内的谱峭度和谱偏斜度等特征。上述特征均聚焦在时域或频域内，除此之外信号的时频域内的特征同样包含了有关产品退化状态的丰富信息<sup>[67]</sup>。时频域分析技术是一研究非稳态信号的有效手段，目前常用的时频域分析技术主要包括小波变换、经验模式分解和短时傅里叶变换等方法。本章采用 CWT 技术<sup>[154]</sup>对振动信号进行时频域特征提取。

CWT 是一种用来分解一连续时间函数为数个小波的方法，相比于傅里叶变换方法，CWT 可以有效地构建一具备良好时域和频域局部特征的时频信号，且常被用于旋转机械的状态监测中<sup>[67]</sup>。从数学角度切入，一具有连续时间性质且可积分的函数  $x(t)$  可表示为以下积分形式：

$$\begin{aligned} X(a, b) &= \langle x(t), \psi_{a,b} \rangle \\ &= \int_{-\infty}^{+\infty} x(t) \bar{\psi}_{a,b}(t) dt \\ &= \frac{1}{\sqrt{|b|}} \int_{-\infty}^{+\infty} x(t) \bar{\psi}\left(\frac{t-a}{b}\right) dt \end{aligned} \quad (5-2)$$

式中， $x(t)$  表示一维振动信号， $\psi(t) \in L^2(R)$  为小波母函数且为一平方可积函数， $\bar{\psi}(t)$  为其对应的共轭复数形式， $a$  和  $b$  分别为平移因子和缩放因子。小波母函数  $\psi(t)$  的用途为提供一可产生子波的根源函数，而子波则是通过小波母函数平移或缩放或平移和缩放产生。由于不存在选取小波母函数的标准，因此本章选择与振动信号较为类似的 Morlet Wavelet 作为小波母函数<sup>[67]</sup>。通过小波变换，一维振动信号  $x(t)$  即可映射为二维小波变换系数矩阵  $X(a, b)$ ，在这里称  $X(a, b)$  为  $x(t)$  的时频特征。

经过 CWT 算法所提取的时频域特征如图 5-4 所示，图中主要给出了 FPT 之后的轴承振动信号及其在等间隔采样区间所提取的对应时频特征（这里采用小波量值图 Scalogram 表示）。其中横轴和纵轴分别表示时间和频率，在二维图片中则采用颜色深浅及其对应的色度条表示小波量值图中的强度（也可称之为振幅）。从图中可以看出，随着轴承的退化，其时频特征（小波量值图）在某一对应频带范围内的强度对应呈现逐渐变大的趋势。因此，除了采用时域和频域内的时间序列特征以刻画轴承的退化趋势外（如图 5-3），时频域内的对应二维图片特征也可以从另一

个角度切入以有效地表征轴承的退化状态。

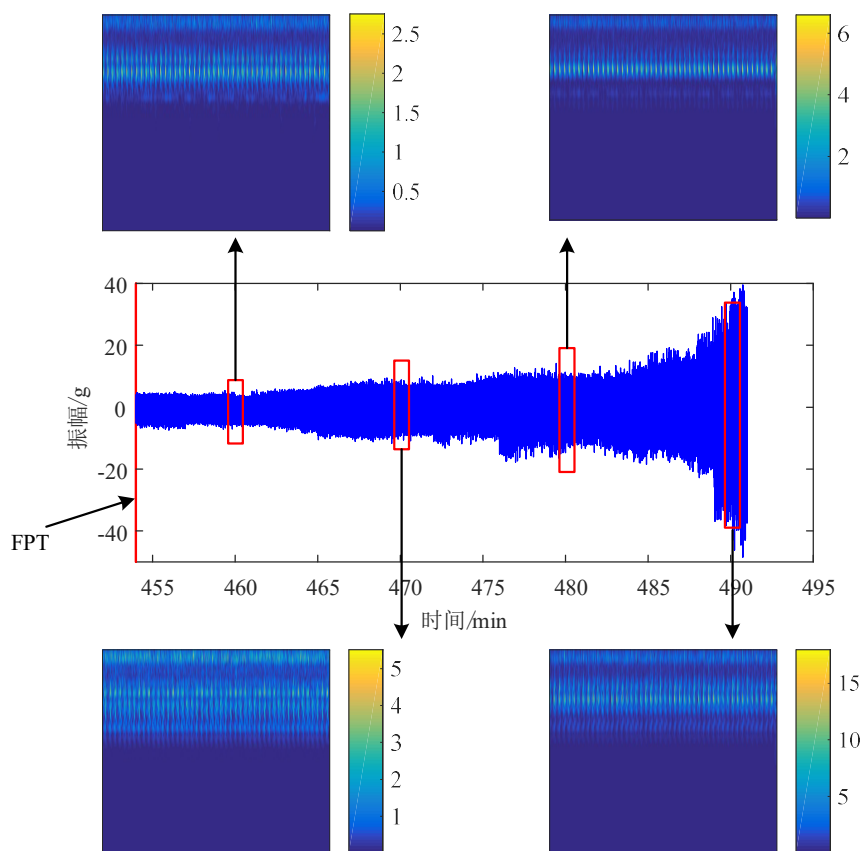


图 5-4 基于 CWT 算法所提取的轴承时频域特征

## 5.3 基于深度卷积神经网络的 RUL 预测模型

### 5.3.1 卷积神经网络简介

相比于全连接前馈神经网络、循环神经网络等，卷积神经网络（CNN）因其局部连接和权重共享等特性更适合处理图像信息<sup>[154]</sup>，例如本章所研究的基于 CWT 算法所提取的时频域特征。卷积神经网络主要是受生物学上感受野机制的启发而提出的一种具备特殊结构的神经网络。生物学上感受野机制主要是指一些神经元只接受其所支配的刺激区域内的信号，例如针对视觉神经元的感受野其实就是指视网膜上的某一特定区域，仅有该区域内的刺激才能激活该神经元，进而将对应的神经冲动信号传导至视觉皮层<sup>[155]</sup>。目前常用的 CNN 网络主要由卷积层、池化层和全连接层交叉堆叠而组成，上述各层分别对应了卷积运算、池化运算和全连接运算等操作，具体如图 5-5 所示。

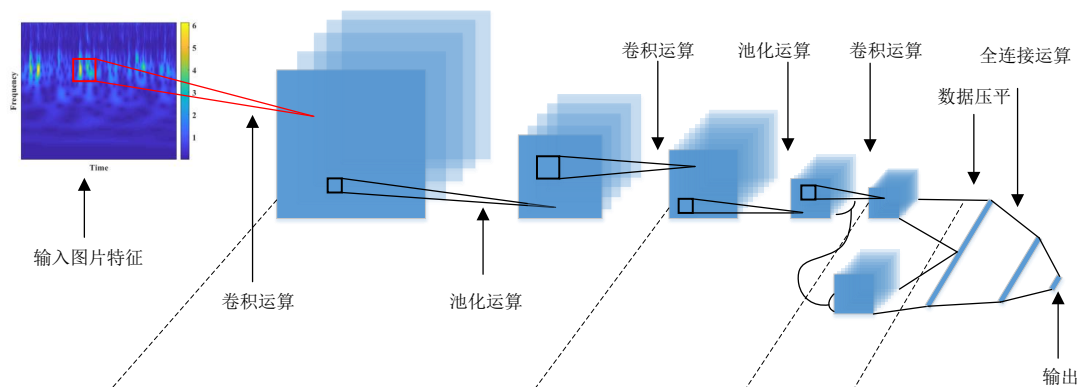


图 5-5 CNN 网络的基本结构

卷积层是经过卷积运算而获得，该计算过程使得卷积层具备了局部连接的性质，且卷积操作为 CNN 网络提取特征的关键方法。卷积层的计算过程主要是将原始的输入图片特征或前一层输出特征图与多个可训练的卷积核进行卷积操作从而得到的一组输出特征图，并进一步通过对应的激活函数而得到最终的输出特征图，该计算过程描述如下：

$$\mathbf{X}_j^r = g \left( \sum_i \mathbf{X}_i^{r-1} * l_{ij}^r + b_j^r \right) \quad (5-3)$$

式中， $\mathbf{X}_j^r$  表示经过卷积运算后的第  $r$  层的对应第  $j$  个输出特征图， $\mathbf{X}_i^{r-1}$  表示第  $(r-1)$  层的第  $i$  个输入特征图， $*$  表示卷积运算， $l_{ij}^r$  和  $b_j^r$  分别表示卷积核与偏置向量， $g(\cdot)$  表示激活函数。需要指出的是多个卷积核将产生组成该层神经元的多个输出特征图（如图 5-5 所示），且该层神经元之间将共享权值。

在卷积操作的实际使用过程中，需要定义如下参数以确定卷积操作具体执行的技术细节，分别是（1）卷积核的大小：确定卷积运算的感受野的大小；（2）步长：确定卷积核遍历输入图像特征的步幅，当步长大于 1 时则表示卷积运算的同时对输入图像特征进行下采样操作；（3）输出通道数：确定卷积核的个数；（4）边界扩充：确定处理输入图像特征边界的方式。一简化输入图片特征为  $6 \times 6$  数据，当采用 3 个卷积核且其大小为  $3 \times 3$ ，遍历输入图片特征的步长为 2，那么其对应的卷积运算则可表示为图 5-6。如图所示，输入的简化图像数据首先采用全零扩充的方式形成新的输入数据，其次采用三个不同的卷积核对其进行步长为 2 的卷积运算，最终得到三个对应的输出特征图，也称为输出特征映射。由于采用的步长为 2，因此实际上通过下采样得到了比原始输入图片数据维度更小但是个数更多的输出特征映射数据。

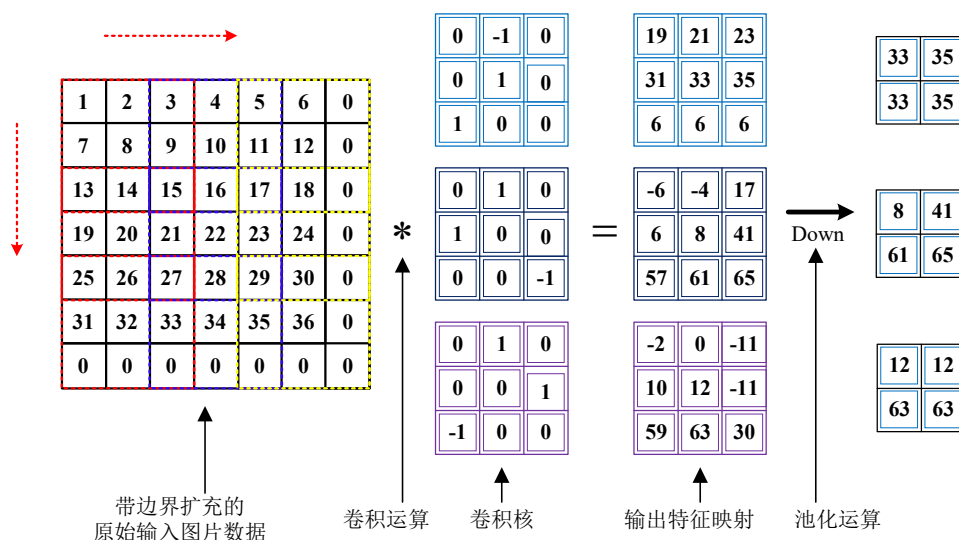


图 5-6 卷积与池化运算简化示意图

池化运算用于卷积运算之后并形成池化层，由于输出特征映射具备局部不变形，所以池化运算可以进一步减小所需参数并扩大感受野。池化运算通过将输出特征映射划分为更小块区域，并对各个区域进行过滤，留下一特定值作为最终输出值以组成新的输出特征映射。常见的池化运算主要有最大池化和平均池化两种，其中最大池化表示选取池化核内的最大值作为新的输出值，平均池化则表示求取池化核内的平均值作为新的输出值。池化运算同样包括：（1）池化核的大小；（2）遍历图像步长；（3）边界扩充等，图 5-6 最右侧即表示池化核为  $2 \times 2$ ，步长为 1 的最大池化运算过程。此外，当原始输入图片数据依次经过几个卷积层和池化层后（如图 5-5 所示），常常还需连接一全连接层以对所提取的高层特征进行输出。

### 5.3.2 基于深度卷积神经网络的 RUL 预测模型

本章所提出的基于深度卷积神经网络（DCNN）的 RUL 预测模型的结构如图 5-7 所示，从图中可以看出，该网络结构主要包含三个部分：（1）采用三个带步长的卷积运算依次提取第一路图片输入数据（基于 CWT 算法所提取的有关振动信号的时频特征）的深层表征，需要指出的是本章没有采用池化层，而是直接使用带步长的卷积运算以实现数据的下采样操作；（2）采用三层全连接神经网络的结构以提取隐藏在第二路时间序列输入特征的深层表征；（3）结合深度卷积神经网络所提取的输入图片数据的深层表征（且经过数据压平处理）和全连接神经网络所提取的时间序列数据的深层表征为一输出特征融合层，并最终通过一线性回归层以预测产品的 RUL。

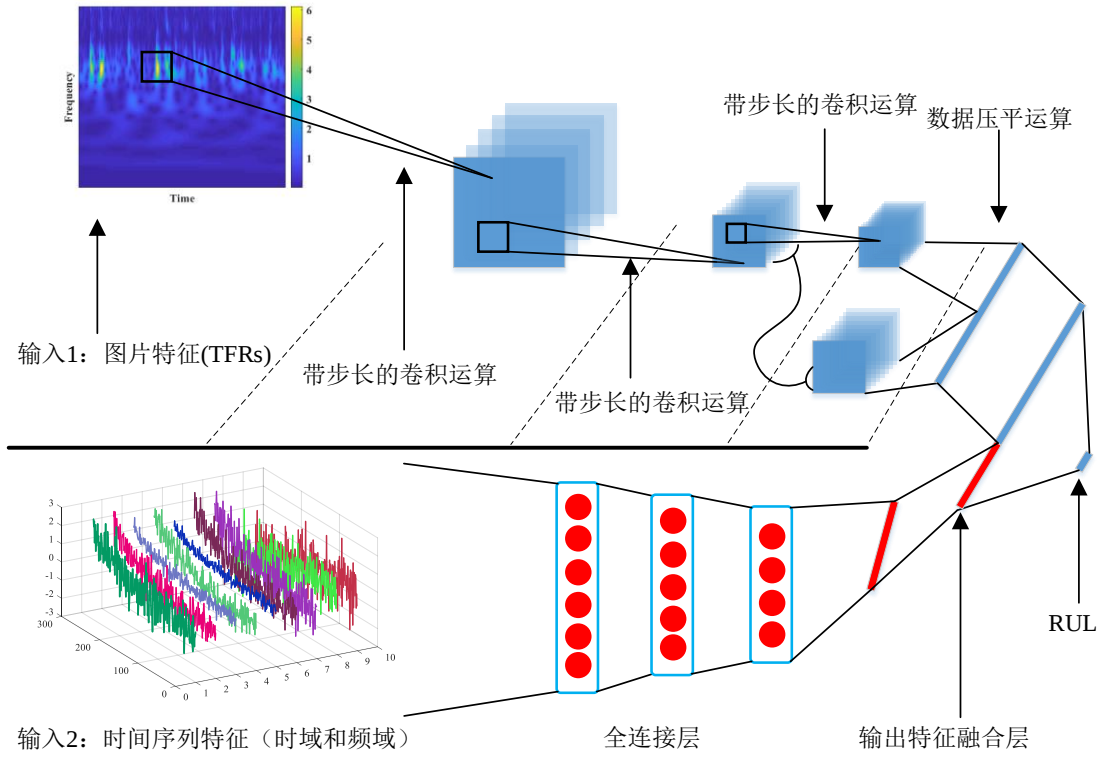


图 5-7 所提出的基于 DCNN 模型的网络结构

采集训练产品的振动监测原始信号，其次基于对应的峭度特征对其 FPT 进行有效识别，此后再对 FPT 之后的时域和频域的时间序列特征和时频域的图像特征进行提取并形成训练集  $\{\mathbf{v}^k, \mathbf{x}^k, y^k\}_{k=1}^N$ ，其中  $\mathbf{v}^k$  表示经过 z-score 归一化处理后的第  $k$  个时域和频域提取的数据特征， $\mathbf{x}^k$  表示基于 CWT 算法所提取的第  $k$  个时频域图像特征， $y^k$  表示对应  $\{\mathbf{v}^k, \mathbf{x}^k\}$  的产品 RUL 标记值。本章所提出的基于 DCNN 的网络结构的三个组成部分依次描述如下：

**Part1:** 将采用 CWT 算法所提取的时频域图像特征  $\mathbf{x}^k$  作为第一路输入数据，并依次进行带步长的卷积运算，该过程计算如下：

$$\mathbf{X}_j^r = g\left(\sum_i \mathbf{X}_i^{r-1} * l_{ij}^r + b_j^r\right) \quad (5-4)$$

该式同式(5-3)一致， $g(\cdot)$  表示激活函数且设置为 ReLU 函数， $*$  表示带步长的卷积运算，并且需要指出的是当  $r=2$  时， $\mathbf{X}^1 = \mathbf{x}^k$  表示第一路的输入 TFRs 数据， $\mathbf{X}_j^r$  表示经过带步长的卷积运算后的第  $r$  层的对应第  $j$  个输出特征图，文中采用三个带步长的卷积操作后即可得到输出特征图集合为  $[\mathbf{X}_j^4]_{j=1}^{d_3}$ ，其中  $d_3$  表示在第三次卷积运算过程中所使用的卷积核的个数。此后，对所获得的输出特征图集合进行数据压平处理即可得到展开的一维特征向量。该数据压平处理过程表示如下：

$$\mathbf{h}_{\text{DCNN}} = \varphi_{\text{Flatten}}\left(\left[\mathbf{X}_j^4\right]_{j=1}^{d_3}\right) \quad (5-5)$$

式中,  $\varphi_{\text{Flatten}}(\cdot)$  表示数据压平运算,  $\mathbf{h}_{\text{DCNN}}$  表示所获得的展开一维特征向量。

**Part2:** 将时域和频域内所提取的时间序列数据特征  $\mathbf{v}^k$  作为第二路输入数据, 并依次经过多层全连接神经网络的处理, 该过程计算如下:

$$\mathbf{V}^r = g\left(\mathbf{W}_{\text{FC}}\mathbf{V}^{r-1} + \mathbf{b}_{\text{FC}}\right) \quad (5-6)$$

式中,  $\mathbf{V}^r$  和  $\mathbf{V}^{r-1}$  分别表示第  $r$  层输出特征向量与第  $(r-1)$  层输入特征向量,  $\mathbf{W}_{\text{FC}}$  和  $\mathbf{b}_{\text{FC}}$  表示由第  $(r-1)$  层到第  $r$  层的对应神经连接的权重向量和偏置向量,  $g(\cdot)$  为激活函数, 在这里同样设置为 ReLU 函数, 需要进一步指出的是  $\mathbf{V}^1 = \mathbf{v}^k$  表示第二路输入数据, 同样地当采用三层全连接结构, 可以得到对应第二路输入数据的输出特征向量  $\mathbf{h}_{\text{FC}} = \mathbf{V}^4$ 。

**Part3:** 结合 DCNN 所提取的第一路输入数据且经过数据压平处理后的一维特征向量  $\mathbf{h}_{\text{DCNN}}$  和全连接神经网络所提取的第二路输入数据的深层表征  $\mathbf{h}_{\text{FC}}$  为一输出特征融合层  $\mathbf{h}_{\text{fuse}} = [\mathbf{h}_{\text{DCNN}}, \mathbf{h}_{\text{FC}}]$ , 并最终将其输入到一线性回归层以预测产品的 RUL, 该过程描述如下:

$$\hat{y} = \mathbf{W}_{\text{R}}\mathbf{h}_{\text{fuse}} \quad (5-7)$$

式中,  $\hat{y}$  表示产品 RUL 的预测值,  $\mathbf{W}_{\text{R}}$  表示线性回归层的权重参数向量。

进一步可以根据产品 RUL 的标记值  $y^k$  定义基于 MSE 的模型损失函数为:

$$L(\boldsymbol{\theta}) = \frac{1}{N} \sum_{k=1}^N \left( f(\mathbf{v}^k, \mathbf{x}^k; \boldsymbol{\theta}) - y^k \right)^2 \quad (5-8)$$

式中,  $f(\cdot)$  表示本章所提的基于 DCNN 的模型,  $\boldsymbol{\theta}$  表示对应的模型参数集, 具体可以简化为  $\boldsymbol{\theta} = [\boldsymbol{\theta}_{\text{DCNN}}, \boldsymbol{\theta}_{\text{FC}}, \mathbf{W}_{\text{R}}]$ , 其中  $\boldsymbol{\theta}_{\text{DCNN}}$ ,  $\boldsymbol{\theta}_{\text{FC}}$  和  $\mathbf{W}_{\text{R}}$  分别在 **Part1~Part3** 中进行了定义。模型的参数集  $\boldsymbol{\theta}$  可以直接通过反向传播算法, 并基于 Adam 优化方法对模型损失函数  $L(\boldsymbol{\theta})$  进行优化, 最终求得最优参数集  $\boldsymbol{\theta}^*$ 。此外, 在模型训练的过程中为了进一步提升模型的泛化能力, 本部分也采用了 Dropout 正则化方法。最后, 本部分针对卷积运算采取了常用的批量归一化 [156], 以提升本章基于 DCNN 模型的训练效率和预测精度。

## 5.4 基于 DCNN 模型和 Bootstrap 框架的 RUL 预测方法

本章及其第四章的主要贡献点之一都在于对产品 RUL 预测区间的量化, 其中第四章主要聚焦在结合 BLSTM 模型与 PF 算法的混合 RUL 预测方法上, 产品 RUL 预测区间的量化主要是基于 PF 算法所完成。但是, 由于在 PF 框架内系统转移模

型为一关键组成部分，因此必须建立有关产品的退化模型。相比于第四章的融合 DL 模型与 PF 算法的方法，本章所提的 DCNN-Bootstrap 方法为一数据驱动方法，在方法的实施过程中不需要产品的退化模型、相关参数的分布及其失效阈值等先验信息。本部分内容将首先简要介绍预测区间构建的基本理论，其次将详细介绍有关基于 DCNN-Bootstrap 方法的 RUL 预测区间量化的相关内容。

#### 5.4.1 预测区间的构建

构建预测区间的理论方法<sup>[157-158]</sup>首先假定预测模型为一针对真实值的无偏估计，且预测误差服从一均值为 0 的确定分布。现对其基本理论简介如下：首先假定一数据集  $\{\mathbf{x}_i, y_i, \bar{y}_i\}_{i=1}^K$ ，其中  $\mathbf{x}_i$  表示模型的输入数据， $y_i$  为模型的真实标记值， $\bar{y}_i$  表示通过回归模型所得的回归均值。本章所提出的基于 DCNN 模型的 RUL 预测本质上同样属于一回归分析问题，即确定一复杂的数学模型  $f(\cdot)$  以准确地拟合输入数据  $\mathbf{x}_i$  与真实标记值  $y_i$  之间的映射关系。针对训练产品由于  $\{\mathbf{x}_i, y_i\}$  均已知，因此可以基于特定的优化算法获取模型的最优参数集  $\Theta^*$ （该部分内容已在 5.3.2 节中进行了详细的阐述）。相反针对测试产品其真实标记值  $y_i$  未知，因此通过将测试产品的输入数据  $\mathbf{x}_i$  导入至模型中以对其真实标记值  $y_i$  进行估计，该过程描述如下：

$$\bar{y}_i = f_{\text{ensemble}}(\mathbf{x}_i; \Theta^*) \quad (5-9)$$

式中， $f_{\text{ensemble}}(\cdot)$  为基于  $f(\cdot)$  的一集成模型（具体技术细节将在后文进行详细介绍）， $\bar{y}_i$  为通过该集成模型计算所得的回归均值，且该回归均值  $\bar{y}_i$  与真实的标记值  $y_i$  之间往往会存在一预测误差，具体表示如下：

$$y_i = \bar{y}_i + \varepsilon_i \quad (5-10)$$

式中， $\varepsilon_i$  表示预测误差项。此外，集成模型  $f_{\text{ensemble}}(\cdot)$  中某一具体的模型  $f_i(\cdot)$  的预测输出值  $\hat{y}_i$  常常不等于回归预测均值  $\bar{y}_i$ ，因此可以进一步得到以下表达式：

$$y_i - \hat{y}_i = (\bar{y}_i - \hat{y}_i) + \varepsilon_i \quad (5-11)$$

式中，假定  $(\bar{y}_i - \hat{y}_i)$  与  $\varepsilon_i$  两项之间相互独立，那么可以得到真实标记值  $y_i$  与模型的预测输出值  $\hat{y}_i$  之间的方差如下：

$$\sigma_i^2 = \sigma_{\hat{y}_i}^2 + \sigma_{\varepsilon_i}^2 \quad (5-12)$$

式中， $\sigma_i^2$  表示真实标记值与模型预测值之间的方差， $\sigma_{\hat{y}_i}^2$  表示模型的预测方差（Prediction Variance）， $\sigma_{\varepsilon_i}^2$  表示模型预测误差项（Prediction Error Term）的方差。由于  $\sigma_{\hat{y}_i}^2$  与  $\sigma_{\varepsilon_i}^2$  的概念比较混淆，采用图 5-8 中的示例以对上述两项概念进行描述。从图中可以看出，模型预测方差主要是由于模型在优化过程中每次所获得最优参

数集不同, 进而导致每次预测时的结果存在差异, 这在基于深度学习的模型中尤为常见。相反, 模型预测误差则更加关注的是预测均值与真实值之间的差异 (这里直接称之为预测误差)。

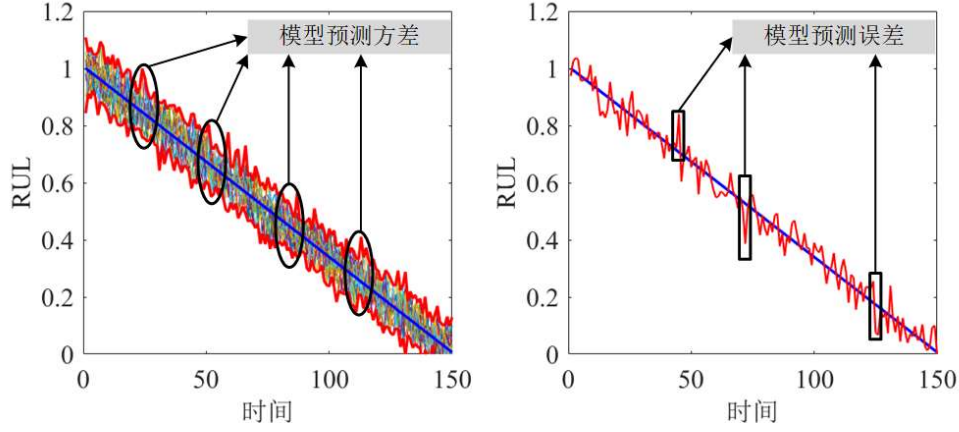


图 5-8 模型预测方差与预测误差示意图

假定预测误差项  $\varepsilon_i$  服从一零均值的正态分布, 那么模型预测值  $\hat{y}_i$  即可认为处在真实标记值  $y_i$  附近, 其  $(1-\alpha) \times 100\%$  置信度下的 PI 可表示如下<sup>[157]</sup>:

$$\left( \bar{y}_i \pm z^{1-\frac{\alpha}{2}} \sigma_i \right) \quad (5-13)$$

Bootstap 方法是非参数统计中一种进行统计量的区间估计的重要方法, 其核心思想在于采用重复抽样技术从原始样本中抽取一定数量 (一般与原始样本量相同) 的样本, 且该抽样过程可为有放回抽样。那么当完整地重复上述抽样过程  $B$  次, Bootstrap 方法即可采用  $B$  个集成回归模型处理上述  $B$  次抽样的整体样本以生成对真实回归值偏差较小的估计, 图 5-9 给出了上述 Bootstrap 过程的示意图。最后, 采用学生  $t$  检验 (Student's  $t$ -test) 的 Bootstrap 分布所得到的分位数  $t_{d_f}^{1-\frac{\alpha}{2}}$  替换式(5-13)中的近似分位数  $z^{1-\frac{\alpha}{2}}$ , 并结合式(5-12)可得到下式<sup>[157]</sup>:

$$\left( \bar{y}_i \pm t_{d_f}^{1-\frac{\alpha}{2}} \sqrt{\sigma_{\hat{y}_i}^2 + \sigma_{\varepsilon_i}^2} \right) \quad (5-14)$$

式中,  $t_{d_f}^{1-\frac{\alpha}{2}}$  表示具有  $d_f$  自由度的  $t$ -分布函数 ( $t$ -Distribution Function) 的  $(1-\alpha/2)$  分位数, 一般情况下可以直接将自由度  $d_f$  设置为整体抽样重复次数  $B$  的大小, 并且  $t_{d_f}^{1-\frac{\alpha}{2}}$  的具体数据可以直接查表获得。

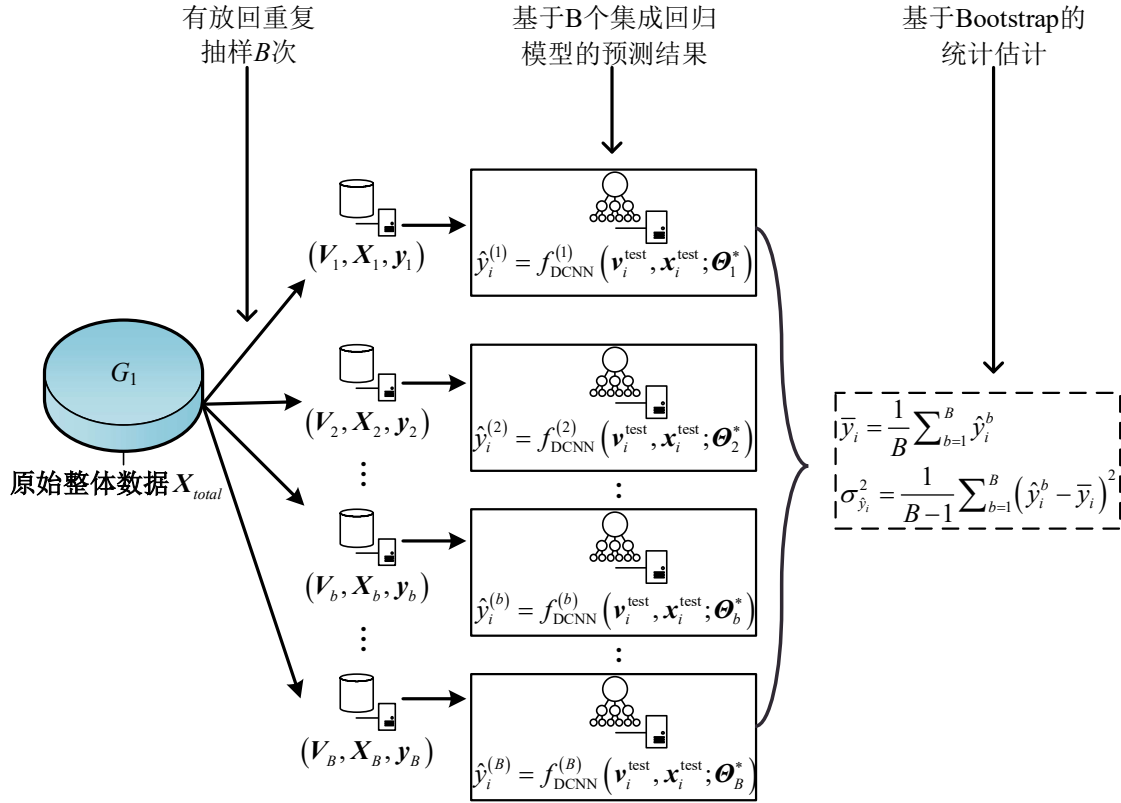


图 5-9 Bootstrap 实施流程示意图

#### 5.4.2 基于 DCNN-Bootstrap 方法的 RUL 预测区间量化

从式(5-14)可知, 求取 RUL 预测区间主要取决于三个关键变量, 分别是基于集成回归模型的回归均值  $\bar{y}_i$ 、预测方差  $\sigma_{\hat{y}_i}^2$  和模型预测误差项的方差  $\sigma_{\varepsilon_i}^2$ 。因此, 本章所提出的基于 DCNN-Bootstrap 方法主要聚焦在求取上述三个关键变量, 其对应的 RUL 预测区间量化流程如图 5-10 所示。总体来说本章所提出的 DCNN-Bootstrap 方法主要涵盖三个步骤, 分别是: (1) 基于 Bootstrap 方法和第一个数据子集构建  $B$  个基于 DCNN 的集成回归模型, 并应用该集成模型实现对前两个关键变量 (回归均值  $\bar{y}_i$  和预测方差  $\sigma_{\hat{y}_i}^2$ ) 的估计; (2) 应用上述集成模型实现对另一数据子集的回归分析, 并进一步创建一新的数据集用于构建另一新的基于 DCNN 的预测误差模型, 该模型即可实现对第三个关键参数 (模型预测误差项的方差  $\sigma_{\varepsilon_i}^2$ ) 的估计; (3) 结合上述  $B$  个基于 DCNN 的集成回归模型和另一基于 DCNN 的预测误差模型, 以实现对测试产品在线监测数据的处理, 并基于式(5-14)获得其对应的 RUL 预测区间的估计。针对上述基于 DCNN-Bootstrap 方法实施流程中所包含的三个关键步骤具体描述如下:

**Step1:** 构建  $B$  个基于 DCNN 的集成回归模型以实现对  $\bar{y}_i$  和  $\sigma_{\hat{y}_i}^2$  的估计。

首先将全体训练数据集  $G$  分为两个训练子集  $G_1$  和  $G_2$ , 其中  $G$  可以表示为

$(V_{\text{total}}, X_{\text{total}}, y)$ ,  $G_1$  和  $G_2$  分别用  $\{v_i^{\text{train}}, x_i^{\text{train}}, y_i^{\text{train}}\}_{i=1}^{N_1}$  和  $\{v_j^{\text{train}}, x_j^{\text{train}}, y_j^{\text{train}}\}_{j=1}^{N_2}$  表示。其次, 针对训练子集  $G_1$  采用有放回抽样的方式生成  $B$  个 Bootstrap 样本, 并简化表示为  $\{V_b, X_b, y_b\}_{b=1}^B$ , 需要说明的是  $B$  个 Bootstrap 样本量均与训练子集  $G_1$  的样本量大小一致。然后, 应用上述  $B$  个 Bootstrap 样本分别训练  $B$  个基于 DCNN 的集成回归模型 (该过程可参考 5.3.2 节中的对应内容), 现简要表示如下:

$$\begin{cases} f_{\text{DCNN}}^{(1)}(V_1, X_1, y_1; \Theta_1^*) \\ \vdots \\ f_{\text{DCNN}}^{(b)}(V_b, X_b, y_b; \Theta_b^*) \\ \vdots \\ f_{\text{DCNN}}^{(B)}(V_B, X_B, y_B; \Theta_B^*) \end{cases} \quad (5-15)$$

式中,  $f_{\text{DCNN}}^{(b)}(V_b, X_b, y_b; \Theta_b^*)$  表示应用第  $b$  个 Bootstrap 样本完成基于 DCNN 模型的前向传播和反向优化训练后所获得的模型, 其中  $\Theta_b^*$  表示模型训练完成后的最优参数集。

最后, 即可将上述完成训练的  $B$  个基于 DCNN 的集成回归模型用于处理测试产品在线监测数据前处理后所获得的测试特征集  $(V_{\text{test}}, X_{\text{test}}) = \{v_i^{\text{test}}, x_i^{\text{test}}\}_{i=1}^K$ , 该过程可表示如下:

$$\hat{y}_i^{(b)} = f_{\text{DCNN}}^{(b)}(v_i^{\text{test}}, x_i^{\text{test}}; \Theta_b^*) \quad (5-16)$$

式中,  $\hat{y}_i^{(b)}$  为测试特征数据  $(v_i^{\text{test}}, x_i^{\text{test}})$  经过第  $b$  个已完成训练的回归模型  $f_{\text{DCNN}}^{(b)}(\Theta_b^*)$  处理之后所得的 RUL 预测输出值。那么当测试特征数据  $(v_i^{\text{test}}, x_i^{\text{test}})$  经过  $B$  个基于 DCNN 的集成回归模型后所得的 RUL 预测输出均值可估计如下:

$$\bar{y}_i = \frac{1}{B} \sum_{b=1}^B \hat{y}_i^{(b)} \quad (5-17)$$

此外, 由  $B$  个基于 DCNN 的集成回归模型在处理测试特征数据  $(v_i^{\text{test}}, x_i^{\text{test}})$  时所产生的预测方差可以结合式(5-17)进一步估计如下:

$$\sigma_{\hat{y}_i}^2 = \frac{1}{B-1} \sum_{b=1}^B (\hat{y}_i^{(b)} - \bar{y}_i)^2 \quad (5-18)$$

**Step2:** 构建一新的基于 DCNN 的预测误差模型以实现  $\sigma_{\hat{y}_i}^2$  的估计。

应用上述  $B$  个基于 DCNN 的集成回归模型以对 **Step1** 中所分割的第二个数据集  $\{v_j^{\text{train}}, x_j^{\text{train}}, y_j^{\text{train}}\}_{j=1}^{N_2}$  进行回归分析, 该过程可表示如下:

$$\hat{y}_j^{(b)} = f_{\text{DCNN}}^{(b)}(v_j^{\text{train}}, x_j^{\text{train}}; \Theta_b^*) \quad (5-19)$$

同式(5-17)与式(5-18)，可以进一步计算针对  $G_2$  数据子集的输入特征  $\{\mathbf{v}_j^{\text{train}}, \mathbf{x}_j^{\text{train}}\}_{j=1}^{N_2}$  经过  $B$  个集成回归模型后的 RUL 预测输出均值  $\bar{y}_j$  与预测方差  $\sigma_{\hat{y}_j}^2$ 。由于  $G_2$  数据子集的 RUL 标记值  $y_j$  已知，因此可以直接估计真实标记值与模型预测值之间的方差  $\sigma_j^2$ ，那么进一步结合式(5-12)即可得到模型预测误差项的方差的近似估计值，该计算过程描述如下：

$$\sigma_{\varepsilon_j}^2 = \sigma_j^2 - \sigma_{\hat{y}_j}^2 \simeq E\left\{\left(y_j - \bar{y}_j\right)^2\right\} - \sigma_{\hat{y}_j}^2 \quad (5-20)$$

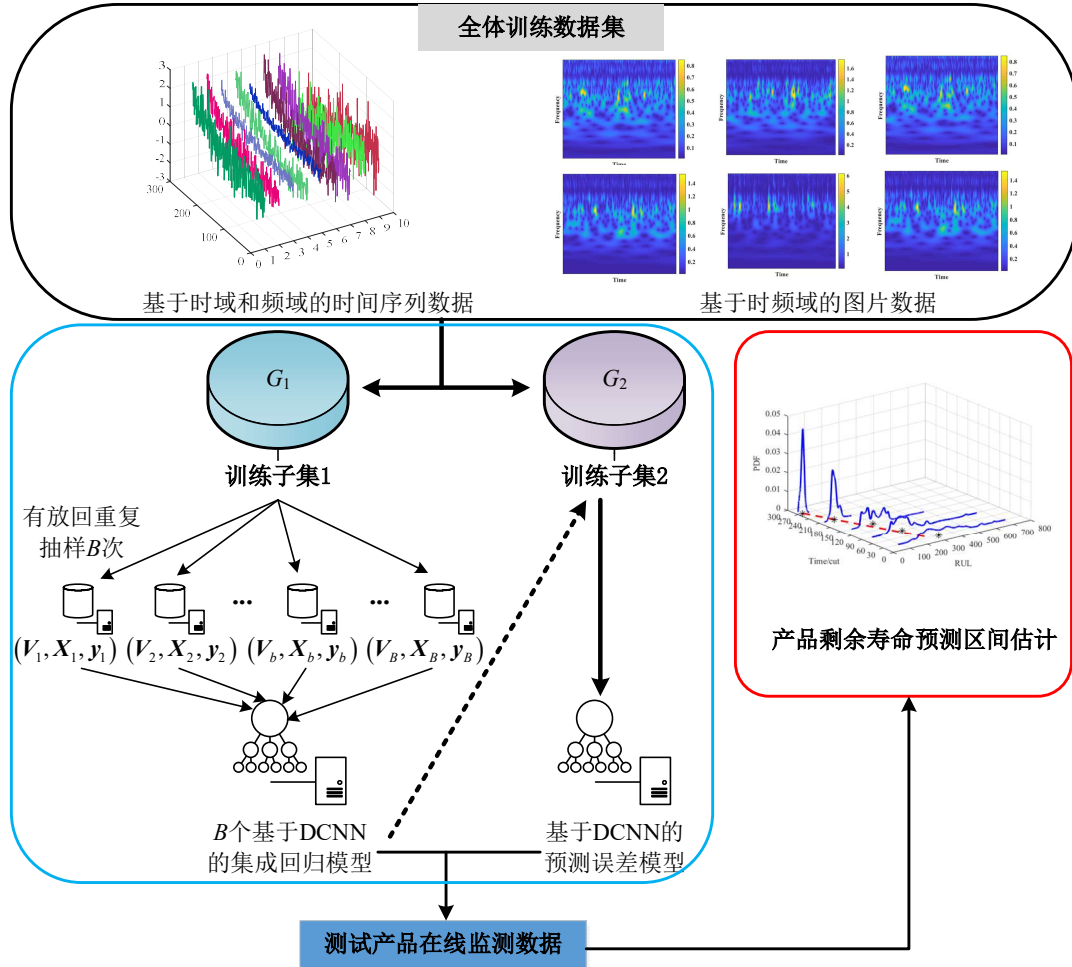


图 5-10 基于 DCNN-Bootstrap 方法的 RUL 预测区间量化流程图

上式在实际计算的过程中有可能会出现  $E\left\{\left(y_j - \bar{y}_j\right)^2\right\} < \sigma_{\hat{y}_j}^2$  的情况，因此采用以下近似表达式求取模型预测误差项的方差，具体描述如下：

$$r_j^2 = \max\left\{\left(y_j - \bar{y}_j\right)^2 - \sigma_{\hat{y}_j}^2, 0\right\} \quad (5-21)$$

式中， $r_j^2$  表示模型预测误差项的方差。因此，可以基于  $G_2$  数据子集与所获得的对应模型预测误差项的方差构建一新的数据集  $G_{r^2}$  并简要表示如下：

$$G_{r^2} = \left\{ \mathbf{v}_j^{\text{train}}, \mathbf{x}_j^{\text{train}}, r_j^2 \right\}_{j=1}^{N_2} \quad (5-22)$$

此后，基于新的数据集  $G_{r^2}$  可进一步构建一基于 DCNN 的预测误差模型以对输入训练特征数据  $\left\{ \mathbf{v}_j^{\text{train}}, \mathbf{x}_j^{\text{train}} \right\}_{j=1}^{N_2}$  与模型预测误差项的方差  $\left\{ r_j^2 \right\}_{j=1}^{N_2}$  之间的映射关系进行建模，对应建模过程描述如下：

$$f_{\text{DCNN}}^{(B+1)} \left( \mathbf{V}_{G_2}, \mathbf{X}_{G_2}, \mathbf{r}_{G_2}^2; \boldsymbol{\Theta}_{B+1}^* \right) \quad (5-23)$$

式中， $\left( \mathbf{V}_{G_2}, \mathbf{X}_{G_2}, \mathbf{r}_{G_2}^2 \right)$  表示新的数据集  $G_{r^2} = \left\{ \mathbf{v}_j^{\text{train}}, \mathbf{x}_j^{\text{train}}, r_j^2 \right\}_{j=1}^{N_2}$ ， $f_{\text{DCNN}}^{(B+1)} \left( \boldsymbol{\Theta}_{B+1}^* \right)$  表示应用  $G_{r^2}$  完成基于 DCNN 模型的前向传播和反向优化训练后所获得预测误差模型，其中  $\boldsymbol{\Theta}_{B+1}^*$  表示对应模型训练完成后最优参数集。

最后，即可将上述完成训练的基于 DCNN 的预测误差模型，用于处理测试产品在线监测数据前处理后所获得的测试特征集  $\left( \mathbf{V}_{\text{test}}, \mathbf{X}_{\text{test}} \right) = \left\{ \mathbf{v}_i^{\text{test}}, \mathbf{x}_i^{\text{test}} \right\}_{i=1}^K$ ，以估计模型预测误差项的方差，该过程可表示如下：

$$\sigma_{\varepsilon_i}^2 = f_{\text{DCNN}}^{(B+1)} \left( \mathbf{v}_i^{\text{test}}, \mathbf{x}_i^{\text{test}}; \boldsymbol{\Theta}_{B+1}^* \right) \quad (5-24)$$

式中， $\sigma_{\varepsilon_i}^2$  为测试特征数据  $\left\{ \mathbf{v}_i^{\text{test}}, \mathbf{x}_i^{\text{test}} \right\}$  输入到预测误差模型  $f_{\text{DCNN}}^{(B+1)} \left( \boldsymbol{\Theta}_{B+1}^* \right)$  后估计得到的模型预测误差项的方差。

**Step3:** 结合上述  $(B+1)$  个模型完成对测试产品的 RUL 预测区间的量化。

通过 **Step1-Step2** 中所构建的  $(B+1)$  个模型即可有效地估计与产品 RUL 预测区间相关的三个关键参数。具体地，RUL 预测均值  $\bar{y}_i$  和预测方差  $\sigma_{\hat{y}_i}^2$  可通过 **Step1** 中所构建的  $B$  个基于 DCNN 的集成回归模型进行估计，模型预测误差项的方差  $\sigma_{\varepsilon_i}^2$  可通过 **Step2** 中所建立的第  $(B+1)$  个基于 DCNN 的预测误差模型进行估计。最后，针对测试产品在线监测数据前处理后所获得的测试特征集  $\left\{ \mathbf{v}_i^{\text{test}}, \mathbf{x}_i^{\text{test}} \right\}$ ，即可量化其对应的 RUL 预测区间，具体可参考式(5-14)。

## 5.5 算例分析与方法验证

### 5.5.1 轴承退化实验数据简介与前处理

为了验证本章所提出的基于深度卷积神经网络和 Bootstrap 的 RUL 预测方法的有效性，本部分采用西安交通大学设计科学与基础部件研究所和长兴昇阳科技有限公司联合发布的 XJTU-SY 轴承数据集。该数据集包含了 15 个滚动轴承在加速退化实验中所获取的运行至失效的振动监测信号<sup>[94]</sup>。所采用的滚动轴承加速退化测试平台如图 5-11 所示，主要包括交流感应电动机、电机速度控制器、支撑轴、两个支承轴承（重型滚动轴承）和液压加压系统等，有关本章所选用的测试轴承的关键参数可参考文献[94]，在此不再赘述。该测试平台可针对滚动轴承在不同工况

(即不同的径向力和转速)下开展相应的加速退化实验,其中所提供的数据集共包括了三种不同的工况,在每种工况下各对五组不同的轴承开展了加速退化实验,且采集了对应的运行至失效的振动监测信号,三种工况分别为:(1)转速:2100 rpm (35 HZ),径向力:12KN;(2)转速:2250 rpm (37.5 HZ),径向力:11KN;(3)转速:2400 rpm (40 HZ),径向力:10KN。

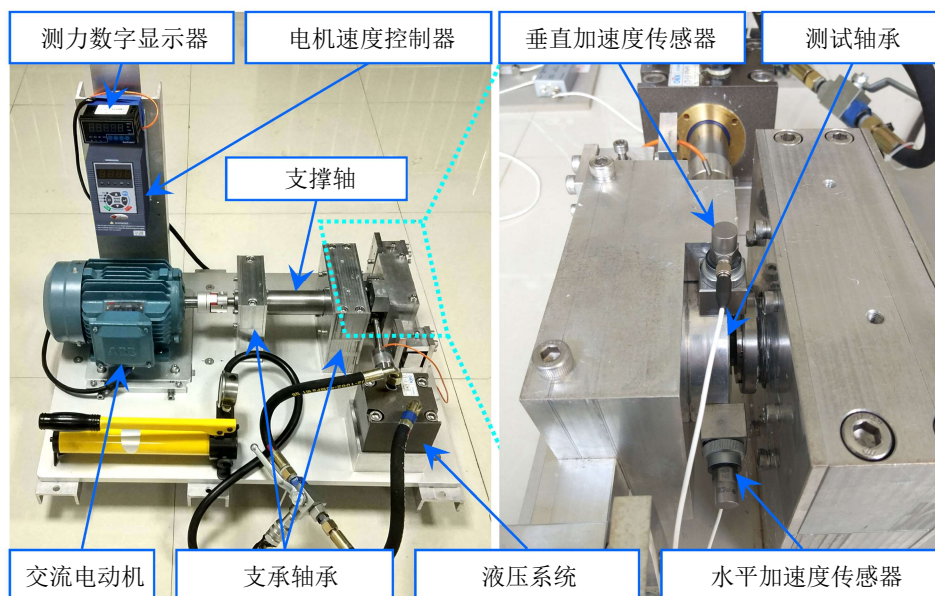


图 5-11 滚动轴承加速退化测试平台

两个加速度传感器被安置在测试轴承外壳的水平和垂直两个方向上以采集测试轴承的水平和垂直两个方向上的振动信号,传感器的采样频率为 25.6 kHz 且每次均采样 1.28 秒,此外采样间隔时间为 1 分钟,具体如图 5-12 所示。

振动信号采集频率为25.6 kHz

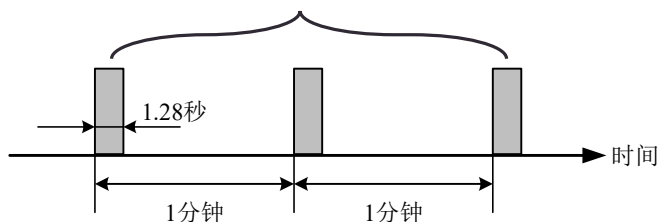


图 5-12 振动信号采样设置

为了有效地观测轴承从健康状态至失效状态完整的退化过程,在每次加速退化实验中当水平或垂直振动信号最大幅值超过 $10 \times A_h$ 时停止测试,其中 $A_h$ 表示当轴承处于健康运行状态下的水平或垂直振动信号最大幅值。本实验中测试滚动轴承的失效部件涵盖了外圈、内圈、保持架和滚动体,图 5-13 给出了不同失效模式对应的示意图。

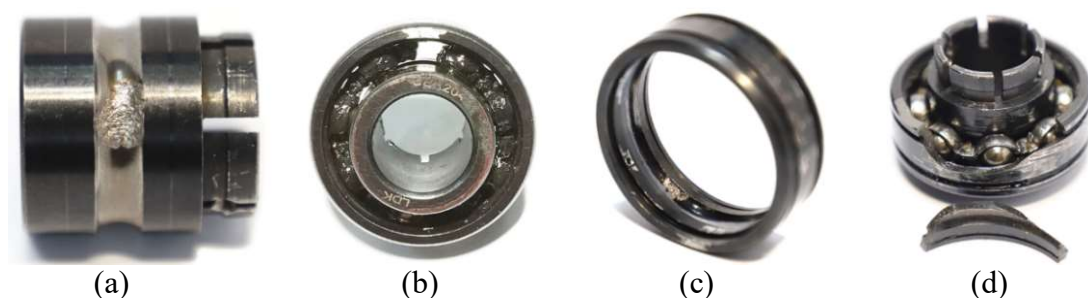


图 5-13 实验滚动轴承四种不同的失效模式。(a)内圈磨损；(b)保持架断裂；(c)外圈磨损；(d)外圈裂损

实验中采用的 XJTU-SY 轴承数据集所涵盖的 15 个滚动轴承其对应的运行工况、所采集的振动信号文件数量、轴承寿命及其对应的失效模式等相关信息如表 5-1 所示，表 5-1 中第四栏也相应给出了基于轴承峭度特征  $3\sigma$  区间所识别出的 FPT 值。需要指出的是 Bearing1\_4 由于其具有突然失效的特征，因此没有给出其对应的 FPT 信息且不会将该组轴承用于后续的分析。此外，由于 Bearing3\_2 的失效部件涵盖了内外圈、滚动体及其保持架，因此其振动监测信号与余下的轴承振动信号存在极大的差异性，因此也没有将该组轴承用于后续的分析。

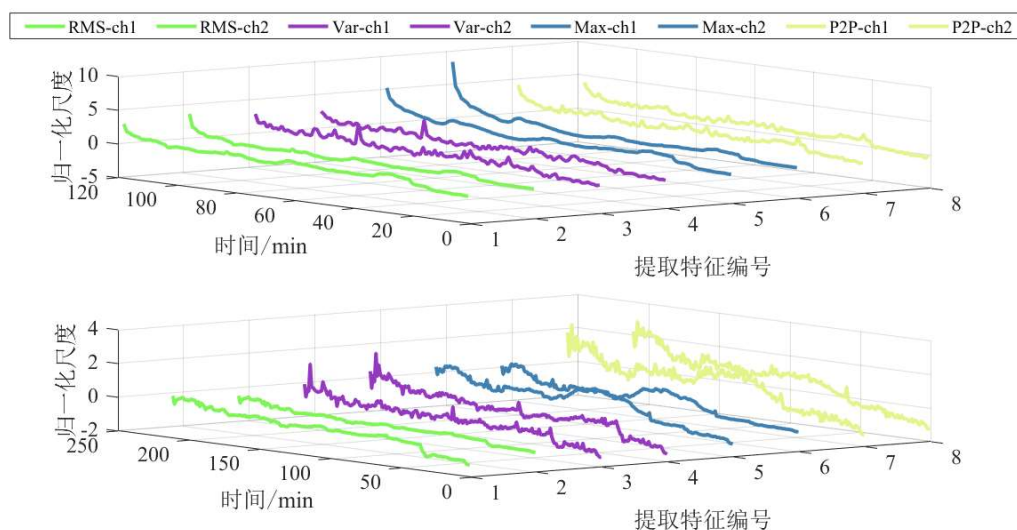


图 5-14 针对两组轴承所提取的两个通道的时域特征

当各个轴承的 FPT 值确定后，即可提取其 FPT 之后的时域和频域内的时间序列特征。针对轴承水平和垂直两个通道的振动监测信号，分别提取时域和频域内的特征包括：均值、方差、均方根、最大值、峰峰值、偏斜度和峭度。此外，还基于快速傅里叶变换技术进一步将两通道振动监测信号转换为对应的频域信号，并提取了其频域特征，具体包括谱偏斜度、谱峭度、谱功率三个特征。经过对上述所提

取的时域和频域内十个时间序列特征进行分析,发现时域内的四个特征(均方根、方差,最大值和峰峰值)具备能有效刻画轴承退化状态的良好性质,其余特征均不能显著地表征轴承完整的退化过程。因此,本章仅选用了时域内的四个特征作为所提出的基于 DCNN 模型的第二路输入时间序列特征(图 5-7 中的输入 2)。图 5-14 给出了针对 Bearing1\_2 和 Bearing2\_3 所提取的两个通道的时域特征示意图,其中 ch1 和 ch2 分别表示通道一和通道二(轴承水平和垂直振动监测信号)所提取的特征。

表 5-1 XJTU-SY 轴承数据集详细信息

| 运行工况                    | 轴承编号       | 文件数  | 轴承寿命 & FPT                  | 失效部件              |
|-------------------------|------------|------|-----------------------------|-------------------|
| 运行工况 1<br>(35Hz/12kN)   | Bearing1_1 | 123  | 2 h 3 min & 1h 15 min       | 外圈                |
|                         | Bearing1_2 | 161  | 2 h 41 min & 43 min         | 外圈                |
|                         | Bearing1_3 | 158  | 2 h 38 min & 59 min         | 外圈                |
|                         | Bearing1_4 | 122  | <b>2 h 2 min &amp; --</b>   | 保持架               |
|                         | Bearing1_5 | 52   | 52 min & 38 min             | 内圈与外圈             |
| 运行工况 2<br>(37.5Hz/11kN) | Bearing2_1 | 491  | 8 h 11 min & 7 h 34 min     | 内圈                |
|                         | Bearing2_2 | 161  | 2 h 41 min & 47 min         | 外圈                |
|                         | Bearing2_3 | 533  | 8 h 53 min & 5 h 22 min     | 保持架               |
|                         | Bearing2_4 | 42   | 42 min & 30 min             | 外圈                |
|                         | Bearing2_5 | 339  | 5 h 39 min & 2 h 21 min     | 外圈                |
| 运行工况 3<br>(40Hz/10kN)   | Bearing3_1 | 2538 | 42 h 18 min & 39 h 4 min    | 外圈                |
|                         | Bearing3_2 | 2496 | <b>41 h 36 min &amp; --</b> | 内圈、外圈、<br>滚动体和保持架 |
|                         | Bearing3_3 | 371  | 6 h 11 min & 5h 39 min      | 内圈                |
|                         | Bearing3_4 | 1515 | 25 h 15 min & 23 h 38 min   | 内圈                |
|                         | Bearing3_5 | 114  | 1 h 54 min & 9 min          | 外圈                |

此外,基于 CWT 算法进一步提取了轴承 FPT 之后的水平振动监测信号的时频特征(文中的图片特征),由于水平和垂直两个通道的振动监测信号所提取的时频特征基本一致,因此仅选用由轴承水平振动监测信号所提取的时频特征。需要指出的经 CWT 算法所提取的轴承原始时频特征具备极高的维度,卷积神经网络需要花费较高的计算资源以处理上述高维图片特征并完成模型的训练。因此,为了减轻计算负担并提升模型训练效率,本章还采用了图像处理领域内常用的双线性插值方法(Bilinear Interpolation)<sup>[67]</sup>以对原始的时频特征进行降维处理。图 5-15 给出了 Bearing2\_1 分别在 8 h, 8 h 5 min, 和 8 h 10 min 三个采样间隔内的原始时频特征(上三图,原始维度为  $135 \times 32768$ ),及其采用双线性插值方法所得到的对应的压

缩时频特征（下三图，压缩维度为  $80 \times 80$ ），从图中可以看出压缩时频特征基本可以准确地捕捉到原始时频特征所具备的典型特征，且其色度条所表示的小波量值图中的强度也进行了对应的压缩。最后，经双线性插值法压缩处理后的时频特征将作为所提出的基于 DCNN 模型的第一路图片特征输入（图 5-7 中的输入 1）。

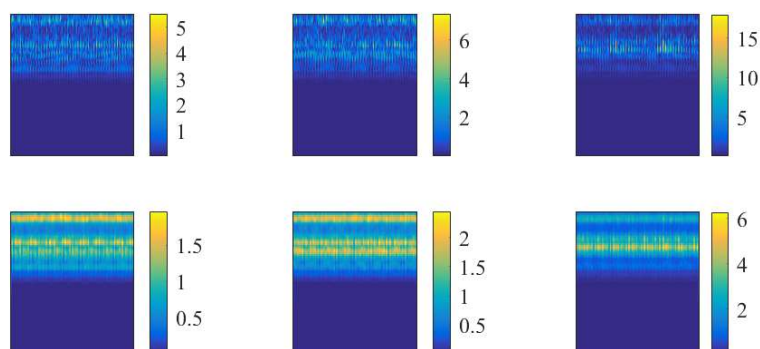


图 5-15 轴承原始时频特征和其对应的压缩时频特征

### 5.5.2 结果分析与讨论

本部分将首先验证本章所提出的基于深度卷积神经网络的 RUL 点预计的有效性。对比基于多层感知机 (MLP) 的方法、基于 LSTM 的方法和普通的基于 DCNN 的方法。上述各对比方法的网络结构简介如下：（1）基于 MLP 的方法采用的是三层全连接结构且各层均采用 Dropout 正则化方法提升其泛化性能；（2）基于 LSTM 的方法采用的四层结构，其中前两层为 LSTM 网络且同样采用了 Dropout 正则化方法，第三层为全连接网络且最后为回归输出层；（3）基于普通 DCNN 的方法采用的六层网络结构，其中前两层为第一个卷积层和池化层，中间两层为第二个卷积层和池化层，最后两层分别为数据压平层和回归输出层，采用 Dropout 方法和批归一化方法以提升模型的泛化能力。需要进一步指出的是，基于 MLP 的方法和基于 LSTM 的方均仅采用轴承时间序列特征作为模型输入，基于普通 DCNN 的方法则仅采用轴承时频特征（图像特征）作为输入。由于本章所提出的基于 DCNN 的方法可同时利用轴承的时间序列特征和图像特征作为模型的输入数据，因此简称本章所提方法为基于混合深度卷积神经网络的方法 (Hybrid DCNN)。此外，本章所采用的三种对比方法及其所提出的 Hybrid DCNN 方法的超参数均采用的是网格搜索后所得的最优超参数组合。本部分仿真实验的编程平台为 sever Intel Xeon CPU E5-2698 v4 2.2-GHz, 64-GB RAM，操作系统为 Windows server 2016 datacenter，编程语言选用的是 tensorflow。表 5-2 给出了各种方法基于性能指标 RMSE 和 MAE 的预测性能对比结果，并且表中的结果均为 10 次运行之后所得的平均预测性能。

表 5-2 Hybrid DCNN 方法与各对比方法的预测性能结果

| 工况   | 轴承编号       | MLP    |       | LSTM  |       | DCNN         |              | Hybrid DCNN  |              |
|------|------------|--------|-------|-------|-------|--------------|--------------|--------------|--------------|
|      |            | RMSE   | MAE   | RMSE  | MAE   | RMSE         | MAE          | RMSE         | MAE          |
| 工况 1 | Bearing1_1 | 0.206  | 0.186 | 0.205 | 0.184 | 0.256        | 0.211        | <b>0.191</b> | <b>0.153</b> |
|      | Bearing1_2 | 0.235  | 0.208 | 0.232 | 0.193 | 0.192        | 0.171        | <b>0.187</b> | <b>0.160</b> |
|      | Bearing1_3 | 0.299  | 0.253 | 0.280 | 0.239 | 0.141        | 0.116        | <b>0.093</b> | <b>0.077</b> |
|      | Bearing1_5 | 0.215  | 0.176 | 0.109 | 0.090 | 0.218        | 0.203        | <b>0.082</b> | <b>0.065</b> |
| 工况 2 | Bearing2_1 | 0.190  | 0.170 | 0.176 | 0.164 | 0.116        | 0.104        | <b>0.096</b> | <b>0.073</b> |
|      | Bearing2_2 | 0.290  | 0.235 | 0.253 | 0.203 | <b>0.080</b> | <b>0.065</b> | 0.086        | 0.067        |
|      | Bearing2_3 | 0.178  | 0.149 | 0.169 | 0.146 | 0.125        | 0.108        | <b>0.101</b> | <b>0.082</b> |
|      | Bearing2_4 | 0.183  | 0.163 | 0.089 | 0.081 | 0.131        | 0.100        | <b>0.089</b> | <b>0.081</b> |
|      | Bearing2_5 | 0.134  | 0.102 | 0.143 | 0.111 | 0.153        | 0.118        | <b>0.092</b> | <b>0.072</b> |
| 工况 3 | Bearing3_1 | 0.277  | 0.241 | 0.263 | 0.228 | 0.184        | 0.148        | <b>0.175</b> | <b>0.147</b> |
|      | Bearing3_3 | 0.1855 | 0.172 | 0.166 | 0.160 | 0.157        | 0.141        | <b>0.132</b> | <b>0.122</b> |
|      | Bearing3_4 | 0.159  | 0.128 | 0.148 | 0.127 | 0.130        | 0.108        | <b>0.112</b> | <b>0.087</b> |
|      | Bearing3_5 | 0.233  | 0.203 | 0.246 | 0.214 | 0.252        | 0.204        | <b>0.230</b> | <b>0.184</b> |

表 5-2 中针对各工况下具体每组轴承的 RUL 预测均采用的是留一法，例如 Bearing1\_1 为测试轴承时，余下三组轴承即 Bearing1\_2, Bearing1\_3 和 Bearing1\_5 均为训练轴承。从表 5-2 中各组轴承的预测性能对比结果可以看出，本章所提出的 Hybrid DCNN 的方法针对所有 13 组轴承的 RUL 预测问题，其中 12 组轴承的预测结果均要显著优于其他对比方法，并且针对 Bearing2\_2 本章所提的 Hybrid DCNN 方法与基于普通 DCNN 方法最优的预测性能也较为接近。此外，从表中还可以看出，基于 LSTM 的方法由于有效地考虑时间序列特征的长时间依赖关系，因此其针对大多数轴承的预测性能均要优于基于 MLP 的方法。但是，将基于 LSTM 的方法与基于 DCNN 的方法进行对比会发现，两种方法之间不存在明显的预测性能的优势。在共计 13 组中的 7 组轴承 RUL 预测任务中，基于 DCNN 的方法表现优于基于 LSTM 的方法，这主要是由于基于 LSTM 的方法与基于 DCNN 的方法分别采用了不同的信息源作为模型的输入数据。因此，正是由于本章所提出的基于 Hybrid DCNN 的方法能同时融合上述两种不同的信息源（即时间序列特征和基于时频表征的图片特征）的数据作为模型的输入，所以本章提出的 Hybrid DCNN 的方法在绝大多数情况下的预测性能均要显著优于其他对比方法。最后，为了清晰阐释本章所提方法及其对比方法针对不同轴承 RUL 预测的结果，图 5-16, 5-17 和图 5-18 分别给出了 Bearing1\_3, Bearing2\_5 和 Bearing3\_4 的 RUL 预测对比结果，图中底部的预测误差条表示 RUL 预测值与其对应真实标记值之间的差值，从图中也可以看

出，本章所提出的基于 Hybrid DCNN 方法在三组不同的轴承 RUL 预测中均要优于其余三种对比方法的预测结果。

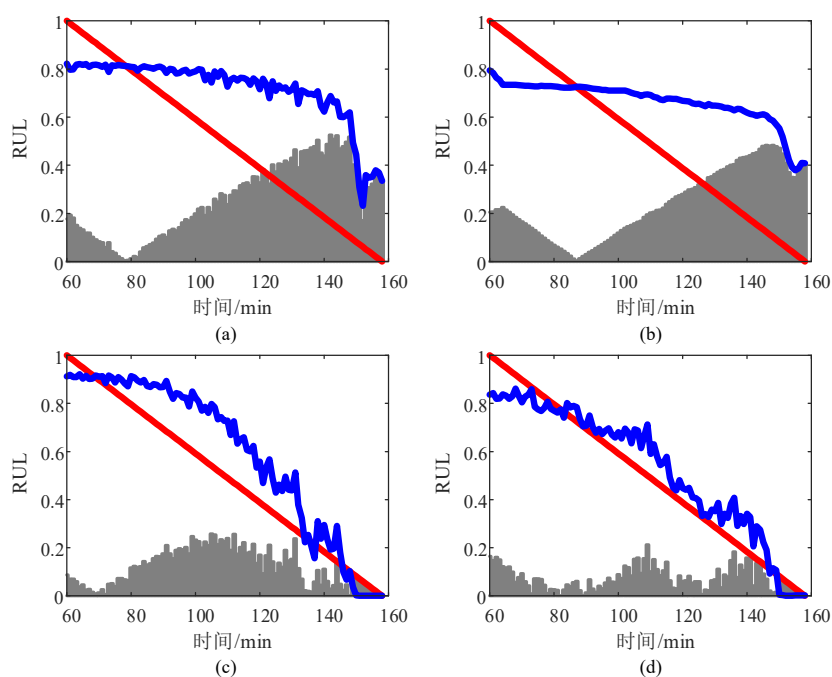


图 5-16 针对 Bearing1\_3 的 RUL 预测结果对比图。(a)基于 MLP 方法的结果；(b)基于 LSTM 方法的结果；(c)基于 DCNN 方法的结果；(d)基于 Hybrid DCNN 方法的结果

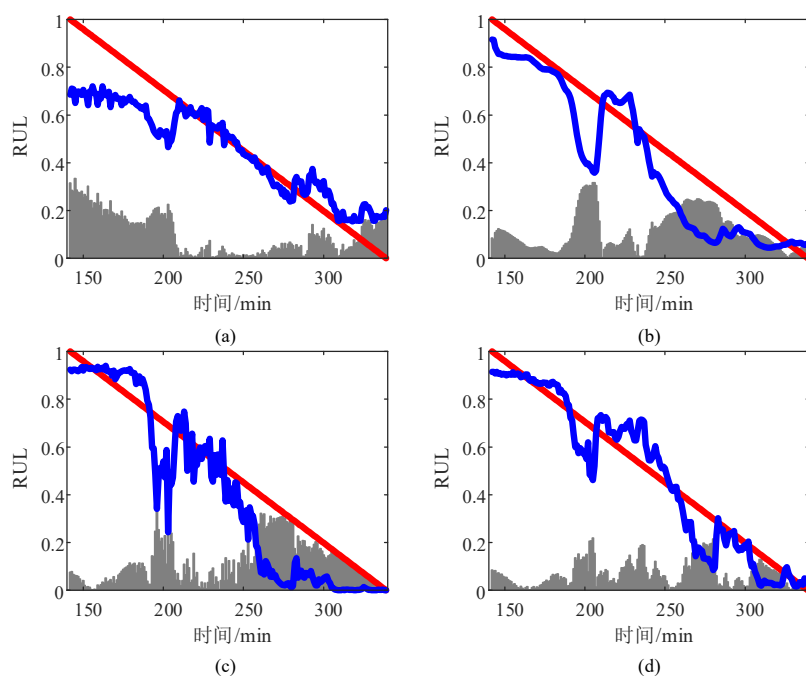


图 5-17 针对 Bearing2\_5 的 RUL 预测结果对比图。(a)基于 MLP 方法的结果；(b)基于 LSTM 方法的结果；(c)基于 DCNN 方法的结果；(d)基于 Hybrid DCNN 方法的结果

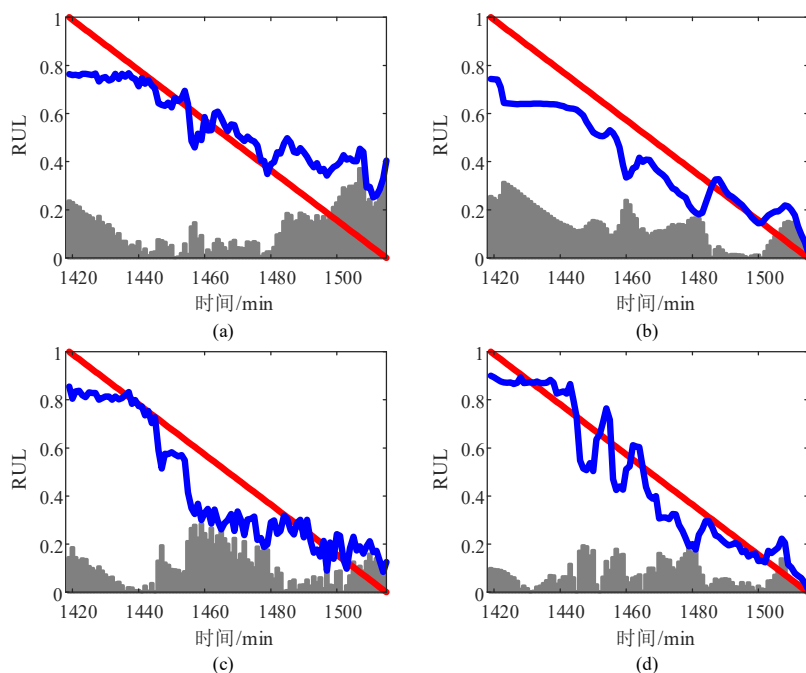


图 5-18 针对 Bearing3\_4 的 RUL 预测结果对比图。(a)基于 MLP 方法的结果；(b)基于 LSTM 方法的结果；(c)基于 DCNN 方法的结果；(d)基于 Hybrid DCNN 方法的结果

最后，为了验证本章所提出的 DCNN-Bootstrap 方法对量化轴承 RUL 预测区间的有效性，本章分别对工作在三种不同工况下的三组轴承进行了 RUL 区间预测。具体地针对运行工况 1，Bearing1\_5 被选作为测试轴承，Bearing1\_1 和 Bearing1\_2 被用作第一个训练子集  $G_1$  且采用有放回抽样的方式生成了 10 个 Bootstrap 样本，并采用上述 10 个 Bootstrap 样本训练了 10 个基于 Hybrid DCNN 的集成回归模型（请参考 5.4.2 节 **Step1** 的具体内容）。此外，Bearing1\_3 被用作为训练子集  $G_2$ ，此后将上述已完成训练的 10 个基于 Hybrid DCNN 的集成回归模型用于处理训练子集  $G_2$  以构建一新的数据集  $G_{r_2}$ ，并利用新生成的数据集  $G_{r_2}$  进一步训练一基于 Hybrid DCNN 的预测误差模型（请参考 5.4.2 节 **Step2** 的具体内容）。最终所获得的测试轴承 Bearing1\_5 的 RUL 预测区间量化结果如图 5-19 所示。同样地针对运行工况 2，Bearing2\_2 选为测试轴承，Bearing2\_3, Bearing2\_4 和 Bearing2\_5 作为第一个训练子集  $G_1$ ，Bearing2\_1 被作为训练子集  $G_2$ 。最终所获得的测试轴承 Bearing2\_2 的 RUL 预测区间量化结果如图 5-20 所示。最后针对运行工况 3，Bearing3\_3 选为测试轴承，Bearing3\_4 和 Bearing3\_5 作为第一个训练子集  $G_1$ ，Bearing3\_1 作为训练子集  $G_2$ 。所获得的测试轴承 Bearing3\_3 的 RUL 预测区间量化结果如图 5-21 所示。

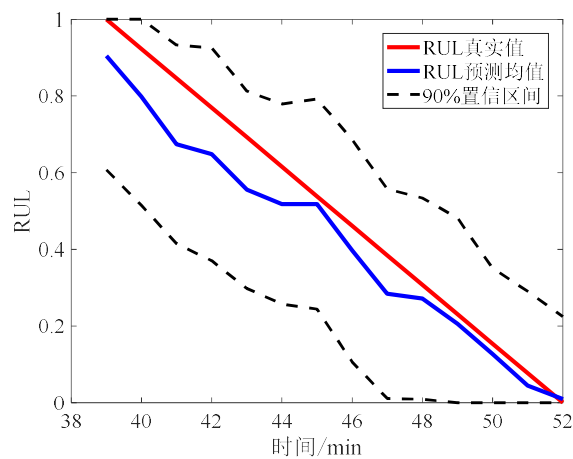


图 5-19 针对 Bearing1\_5 的 RUL 预测区间量化结果

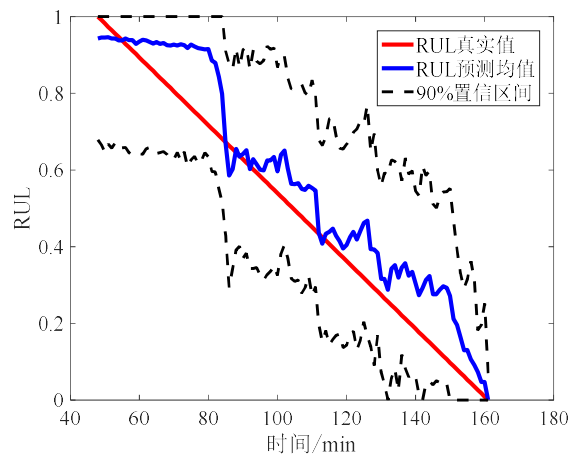


图 5-20 针对 Bearing2\_2 的 RUL 预测区间量化结果

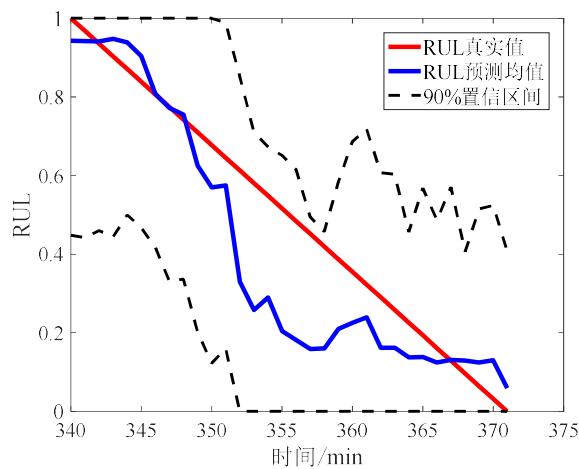


图 5-21 针对 Bearing3\_3 的 RUL 预测区间量化结果

基于图 5-19, 5-20 和 5-21, 具体讨论如下:

(1) 针对上述三组轴承任一观测时刻点, 轴承真实的 RUL 值均落在本章提出的基于 DCNN-Bootstrap 方法的 90%置信区间内。因此基于本章所提的 DCNN-Bootstrap 方法所制定的有关轴承的预防性维护决策将会更加稳健。

(2) 本章内容同第四章内容的其中一个贡献点均在于扩展基于深度学习模型的方法, 以实现产品 RUL 的区间估计的问题, 但是本章内容不同于第四章的地方在于, 本章所提的基于 DCNN-Bootstrap 方法不需要产品退化模型及其模型中各个参数的分布等先验信息, 该特点使本章所提方法更加适用于一些复杂的安全关键性产品。

(3) 尽管本章所提的 DCNN-Bootstrap 方法能有效地对轴承 RUL 预测区间进行量化, 但是针对各组轴承其对应的 90%置信区间较大。若在训练样本更加充足的情况下, 增加训练子集 G2 的数量将会有助于量化 RUL 预测区间结果。

## 5.6 本章小节

基于深度卷积神经网络和 Bootstrap 技术, 本章提出了一种 DCNN-Bootstrap 的 RUL 预测方法。一方面所提出的 Hybrid DCNN 模型能同时利用基于产品时频和频域所提取的时间序列特征和基于时频域的图片特征, 另一方面由于本章所提的 DCNN-Bootstrap 方法有效地将 Hybrid DCNN 模型纳入了 Bootstrap 的实施框架内, 因此能进一步实现产品 RUL 预测区间的量化。相比于其他基于深度学习模型的产品 RUL 预测方法, 本章所提方法具备两个新颖的特点, 其一为本章所提方法可以同时利用能表征产品退化的时间序列信息和图片信息, 以进一步增强产品 RUL 预测的精度, 其次为本章所提方法在不需要有关产品具体的退化模型和其对应各个参数的分布等先验信息的基础上, 实现产品 RUL 预测区间的量化。通过算例分析验证了本章所提出的 Hybrid DCNN 模型在对滚动轴承 RUL 点估计的有效性, 通过对工作在不同工况下的轴承的 RUL 预测区间的量化进一步阐释了本章所提出的 DCNN-Bootstrap 方法的有效性。

## 第六章 总结与展望

### 6.1 本文主要研究结论

传感器技术、物联网技术和 5G 等技术的快速发展推动产品状态监测研究进入了工业大数据时代，基于可获得的海量监测数据对重大产品和重大设施进行 RUL 预测为先进制造领域内一前沿技术，同时该技术也是提高产品运行可靠性、安全性和可维护性的关键所在。基于数据驱动的产品 RUL 预测方法，尤其是近年来兴起的基于深度学习模型的方法因其强大的特征提取和回归分析等能力且不依赖于有关产品失效机理和失效模式等先验信息的优点，更是为解析产品海量监测数据并进一步探明产品健康状态且最终对其 RUL 进行更加精确的预测等问题提供了一行之有效的途径。然而，现有的产品 RUL 预测理论和技术在海量监测数据背景下遇到了全新的挑战，亟需研究新的基于数据驱动的 RUL 预测方法，并整合最新的深度学习等相关技术以构建更加符合当下工业大数据背景下的产品 RUL 预测的全新理论和方法。

本文在对现有 RUL 预测模型与方法进行全面且深入调研的基础上，分别研究了基于产品退化轨迹相似性的 RUL 预测方法、在复杂运行工况下基于双向长短期记忆神经网络的产品 RUL 预测方法、结合双向长短期记忆神经网络与粒子滤波技术的混合方法及其基于深度卷积神经网络和 Bootstrap 的产品 RUL 预测方法，主要研究成果如下：

(1) 提出了基于核密度估计技术的退化轨迹相似性 RUL 预测改进方法。本文结合核密度估计技术和  $\beta$ -准则提出了基于产品退化轨迹相似性的改进方法，进一步突破原方法仅能给出 RUL 点估计的限制。通过将改进方法应用于航空发动机的 RUL 预测发现，改进方法相比于原方法具有更高的预测精度且还能给出一合理的 RUL 区间估计结果，更适合解决一些对安全性要求较高产品的 RUL 预测问题。

(2) 提出了基于双向长短期记忆神经网络的产品多工况下 RUL 预测方法。本文基于双向长短期记忆神经网络提出了一 RUL 预测方法以有效地融合产品多传感器监测数据与复杂工况数据，且进一步对工作在复杂多工况下产品的 RUL 预测问题给出了一端对端的解决方案。所提方法简化了传统数据驱动 RUL 预测方法（第二章）的实施流程，且对工作在多工况下的产品具有很高的预测精度。

(3) 提出了基于双向长短期记忆神经网络和粒子滤波技术的混合 RUL 预测方法。所提方法通过有效整合粒子滤波技术能突破基于深度学习模型的方法（第三章）在 RUL 区间估计上的限制，并且由于采用了双向长短期记忆神经网络，因此

所提方法简化了传统的混合 RUL 预测方法的处理流程，并进一步提升了混合方法的实施效率和稳健性。

(4) 提出了基于深度卷积神经网络和 Bootstrap 的 RUL 预测方法。所提出的基于 Hybrid DCNN 的模型能同时利用产品的时间序列特征数据和对应的图片特征数据，可进一步提高基于深度学习方法的 RUL 预测精度。此外，由于采用了 bootstrap 的实施框架，因此本文所提出的 Hybrid DCNN-bootstrap 方法能突破基于深度学习模型的混合方法（第四章）需要产品退化模型等先验信息的限制。

## 6.2 后续工作展望

在工业大数据背景下有关产品 RUL 预测的研究正处于蓬勃发展的阶段，基于目前已完成的相关工作，后续拟开展的工作如下：

(1) 基于深度学习模型的方法与基于深度学习模型的混合方法均能在工业大数据背景下有效地实现产品监测大数据的自动解析，并进一步精确地给出其 RUL 估计。但是，上述方法的缺点是需要产品大量运行至失效的完整监测数据以训练深度神经网络。然而在实际工程中，针对一些对风险性要求较高的产品，例如，高铁轴承、直升飞机轴承等，很难获得足够的完整监测数据。相反由于定时维修、更换、或停机检查等原因，往往存在产品大量的运行至非失效的截尾监测数据。因此，对于只存在少量完整监测数据与大量截尾监测数据的特殊产品的 RUL 预测问题，拟在今后的研究中考虑使用半监督类的深度学习框架，例如对抗生成网络、变分自编码器，从而建立有效利用产品海量截尾监测数据的基于深度学习模型的新方法。

(2) 在产品 RUL 预测的工程实践中，针对同类型产品大致具备相似的失效模式与失效机理，但是其工程应用的背景却各不相同，服役的工况环境差异较大，例如实验室轴承、高铁轴承、风力涡轮机轴承与车用轴承等。其中实验室轴承运行至失效的监测数据大而全，可以基于对应的海量监测数据构建泛化能力强且预测精度高的基于深度学习模型的相关方法。然而，在其他相关实际工程背景下由于产品刚投入使用或其对安全性要求较高等原因，往往难以获取产品足够的运行至失效的监测数据。因此，拟在下一步研究中考虑使用迁移学习的方法以实现将实验室所构建的深度学习模型迁移至各种不同工程背景中，并完成对应的产品 RUL 预测任务。

(3) 尽管基于深度学习模型的方法在众多产品的 RUL 预测中展示出了极高的预测精度，但是基于深度学习的模型仍然被视为黑盒子模型，其超参数选择、模型优化等内部运算机理常常难以解释。因此，为增强基于深度学习模型的方法的可解释性，拟在下一步研究中采用各种特征可视化技术加深对各种深度学习模型所

学到的特征的认识，进而构建可解释性更强的基于深度学习模型的方法。

## 致 谢

行文至此，我在电子科技大学的十年求学生涯也即将结束。回首成电十年，往昔记忆依旧让我感觉似梦似幻，如痴如醉，各种情绪顿时袭上心头。博士四年艰辛且不易，有过论文被拒后的低落和遗憾，也有过论文接收后的兴奋与喜悦，但此时此刻内心中更多的还是感恩。

我要感恩我的导师黄洪钟教授。感谢您将我引入可靠性工程的学术殿堂，并加入您所创立的电子科技大学系统可靠性与安全性研究中心这一大家庭。“古人学问无遗力，少壮工夫老始成”，恩师亲身践行着这一治学信仰，且深深的影响着我，让我明白了厚积薄发的道理。“勤奋固然重要，但方法更为珍贵”，每当我遇到研究瓶颈的时候，恩师这句话总会成为我攻克研究难题的指明灯。感谢恩师鼓励并帮助我出国进行联合培养，让我拥有了更加开阔的学术视野；在生活上，恩师更似慈父，教会了我感恩和与人为善的道理。我在博士期间所取得的成果无不凝聚着您的心血，今后踏上工作岗位，学生定将铭记师恩，砥砺前行，再次向恩师说声感谢，感谢有您！

特别感谢我在新加坡国立大学两年联合培养期间给予我学术和生活上极大帮助的董润桢老师和叶志盛老师，感谢你们在异国他乡给我提供了宝贵的研究平台和学术支持；感谢翟庆庆博士、胡丁骊博士、陈飘博士、孙秋壮博士、洪兰青博士、等，感谢你们给予我在国外学习和生活期间的无私帮助。

感谢研究中心的刘宇老师、李彦锋老师、汪忠来老师、朱顺鹏老师、许焕卫老师等在学术上对我的帮助；感谢我的师兄彭卫文博士、杨圆鉴博士、吕志强博士、王海琨博士、彭兆春博士、付国忠博士、殷毅超博士、张小强博士、李翔宇博士、郑波博士，师姐米金华博士、高会英博士、刘征博士，感谢你们在学术上对我的指点，让我获益匪浅；感谢我的同窗同学郭骏宇和周杰，谢谢你们在艰苦岁月里的陪伴；感谢李贺、李享、黄鹏、钱华明、李懿凡、余奥迪、柏松、曾颖、黄土地等同学，与你们相遇于系统可靠性与安全性研究中心这一大家庭，我倍感快乐。

特别感谢我的父母，是你们的支持、理解、极具启发式的教导和无条件的付出，让我有绝对的勇气和自信完成我的博士生涯。

特别感谢我的妻子谭玲，感恩在最美的年华于成电与你相识、相知、相爱，并且在我一事无成的时候伴我左右，理解我，全心全意地支持我，让我能在求学生涯披荆斩棘。

最后，衷心感谢评阅本论文的教授和专家！

## 参考文献

- [1] A. K. S. Jardine, D. M. Lin, D. Banjevic. A review on machinery diagnostics and prognostics implementing condition-based maintenance[J]. *Mechanical Systems and Signal Processing*, 2006, 20: 1483-1510
- [2] J. Lee, F. J. Wu, W. Y. Zhao, et al. Prognostics and health management design for rotary machinery systems-reviews, methodology and applications[J]. *Mechanical Systems and Signal Processing*, 2014, 42(1-2): 314-334
- [3] International Standards Organisation. ISO 13381-1. Condition monitoring and diagnostics of machines-prognostics-part 1: general guidelines[S], UK: British Standards Institution, 2015
- [4] Y. Lei, N. Li, L. Guo, et al. Machinery health prognostics: a systematic review from data acquisition to RUL prediction[J]. *Mechanical Systems and Signal Processing*, 2018, 104: 799-834
- [5] P. C. Paris, F. Erdogan. A critical analysis of crack propagation laws[J]. *Journal of Basic Engineering*, 1963, 85(4): 528-533
- [6] Y. Qian, R. Yan, R X. Gao. A multi-time scale approach to remaining useful life prediction in rolling bearing[J]. *Mechanical Systems and Signal Processing*, 2017, 83: 549-567
- [7] F. Cadini, E. Zio, D. Avram. Monte Carlo-based filtering for fatigue crack growth estimation[J]. *Probabilistic Engineering Mechanics*, 2009, 24: 367-373
- [8] E. Zio, G. Peloni. Particle filtering prognostic estimation of the remaining useful life of nonlinear components[J]. *Reliability Engineering & System Safety*, 2011, 96(3): 403-409
- [9] Y. Li, S. Billington, C. Zhang, et al. Adaptive prognostics for rolling element bearing condition[J]. *Mechanical Systems and Signal Processing*, 1999, 13(1): 103-113
- [10] M. E. Orchard, G. J. Vachtsevanos. A particle-filtering approach for on-line fault diagnosis and failure prognosis[J]. *Transactions of the Institute of Measurement and Control*, 2009, 31(3-4): 221-246
- [11] F. Zhao, Z. Tian, Y. Zeng. Uncertainty quantification in gear remaining useful life prediction through an integrated prognostics method[J]. *IEEE Transactions on Reliability*, 2013, 62(1): 146-159
- [12] Y. Lei, N. Li, S. Gontarz, et al. A model-based method for remaining useful life prediction of machinery[J]. *IEEE Transactions on Reliability*, 2016, 65(3): 1314-1326
- [13] J. Sun, H. Zuo, W. Wang, et al. Prognostics uncertainty reduction by fusing on-line monitoring data based on a state-space-based degradation model[J]. *Mechanical Systems and Signal Processing*, 2014, 45(2): 396-407

- 
- [14] L. Liao. Discovering prognostic features using genetic programming in remaining useful life prediction[J]. IEEE Transactions on Industrial Electronics, 2014, 61(5): 2464-2472
- [15] J. Wang, R. X. Gao, Z. Yuan, et al. A joint particle filter and expectation maximization approach to machine condition prognosis[J]. Journal of Intelligent Manufacturing, 2016, 30(2): 605-621
- [16] P. Baraldi, F. Mangili, E. Zio. A Kalman filter-based ensemble approach with application to turbine creep prognostics[J]. IEEE Transactions on Reliability, 2012, 61(4): 966-977
- [17] P. Baraldi, F. Mangili, E. Zio. Investigation of uncertainty treatment capability of model-based and data-driven prognostic methods using simulated data[J]. Reliability Engineering & System Safety, 2013, 112: 94-108
- [18] J. Chiachío, M. Chiachío, S. Sankararaman, et al. Condition-based prediction of time-dependent reliability in composites[J]. Reliability Engineering & System Safety, 2015, 142: 134-147
- [19] K. S. Chan, M. P. Enright, J. P. Moody, et al. Life prediction for turbopropulsion systems under dwell fatigue conditions[J]. Journal of Engineering for Gas Turbines and Power, 2012, 134(12): 1-8
- [20] K. El-Tawil, A. A. Jaoude. Stochastic and nonlinear-based prognostic model[J]. Systems Science & Control Engineering: An Open Access Journal, 2013, 1(1): 66-81
- [21] N. Z. Gebraeel, M. A. Lawley, R. Li, et al. Residual-life distributions from component degradation signals: a Bayesian approach[J]. IIE Transactions, 2005, 37(6): 543-557
- [22] J. Coble, J. W. Hines. Incorporating prior belief in the general path model: a comparison of information sources[J]. Nuclear Engineering and Technology, 2014, 46(6): 773-782
- [23] J. I. Park, S. J. Bae. Direct prediction methods on lifetime distribution of organic light-emitting diodes from accelerated degradation tests[J]. IEEE Transactions on Reliability, 2010, 59(1): 74-90
- [24] X. Jin, Y. Sun, Z. Que, et al. Anomaly detection and fault prognosis for bearings[J]. IEEE Transactions on Instrumentation and Measurement, 2016, 65(9): 2046-2054
- [25] J. Hu, P. W. Tse. A relevance vector machine-based approach with application to oil sand pump prognostics[J]. Sensors, 2013, 13(9): 12663-12686
- [26] Z. Huang, Z. Xu, X. Ke, et al. Remaining useful life prediction for an adaptive skew-Wiener process model[J]. Mechanical Systems and Signal Processing, 2017, 87: 294-306
- [27] H. Liao, Z. Tian. A framework for predicting the remaining useful life of a single unit under time-varying operating conditions[J]. IIE Transactions, 2013, 45(9): 964-980
- [28] W. B. Wang, M. Carr, W. J. Xu, et al. A model for residual life prediction based on Brownian motion with an adaptive drift [J]. Microelectronics Reliability, 2011, 51(2): 285-293

- [29] S. Liu, Y. Hu, C. Liu, et al. Real-time reliability self-assessment in milling tools operation[J]. *Quality and Reliability Engineering International*, 2015, 32(7): 2245-2252
- [30] L. Hao, L. Bian, N. Gebraeel, et al. Residual life prediction of multistage manufacturing processes with interaction between tool wear and product quality degradation[J]. *IEEE Transactions on Automation Science and Engineering*, 2017, 14(2): 1211-1224
- [31] C. Y. Peng, S. T. Tseng. Statistical lifetime inference with skew-wiener linear degradation models[J]. *IEEE Transactions on Reliability*, 2013, 62(2): 338-350
- [32] C. Y. Peng, S. T. Tseng. Mis-specification analysis of linear degradation models[J]. *IEEE Transactions on Reliability*, 2009, 58(3): 444-455
- [33] X. S. Si, W. B. Wang, C. H. Hu, et al. Remaining useful life estimation based on a nonlinear diffusion degradation process[J]. *IEEE Transactions on Reliability*, 2012, 61(1): 50-67
- [34] K. Le Son, M. Fouladirad, A. Barros, et al. Remaining useful life estimation based on stochastic deterioration models: a comparative study[J]. *Reliability Engineering & System Safety*, 2013, 112: 165-175
- [35] X. S. Si, W. B. Wang, C. H. Hu, et al. Estimating remaining useful life with three-source variability in degradation modeling[J]. *IEEE Transactions on Reliability*, 2014, 63(1): 167-190
- [36] S. P. Kuniewski, J. A. van der Weide, J. M. van Noortwijk. Sampling inspection for the evaluation of time-dependent reliability of deteriorating systems under imperfect defect detection[J]. *Reliability Engineering & System Safety*, 2009, 94(9): 1480-1490
- [37] V. Bagdonavicius, M. S. Nikulin. Estimation in degradation models with explanatory variables[J]. *Lifetime Data Analysis*, 2001, 7(1): 85-103
- [38] J. Lawless, M. Crowder. Covariates and random effects in a gamma process model with application to degradation and failure[J]. *Lifetime Data Analysis*, 2004, 10(3): 213-227
- [39] W. Peng, Y. F. Li, Y. J. Yang, et al. Leveraging degradation testing and condition monitoring for field reliability analysis with time-varying operating missions[J]. *IEEE Transactions on Reliability*, 2015, 64(4): 1367-1382
- [40] H. K. Wang, H. Z. Huang, Y. F. Li, et al. Condition-based maintenance with scheduling threshold and maintenance threshold[J]. *IEEE Transactions on Reliability*, 2016, 65(2): 513-524
- [41] 潘东辉. 基于退化数据的产品可靠性建模与剩余寿命预测方法研究[D]. 武汉: 华中科技大学, 2014
- [42] Z. S. Ye, L. P. Chen, L. C. Tang, et al. Accelerated degradation test planning using the inverse Gaussian process[J]. *IEEE Transactions on Reliability*, 2014, 63(3): 750-763
- [43] 彭卫文. 重型数控机床可靠性建模与评估技术研究[D]. 成都: 电子科技大学, 2015

- [44] W. Peng, Y. F. Li, Y. J. Yang, et al. Bivariate analysis of incomplete degradation observations based on inverse Gaussian processes and copulas[J]. IEEE Transactions on Reliability, 2016, 65(2): 624-639
- [45] 胡姚刚, 李辉, 廖兴林, 等. 风电轴承性能退化建模及其实时剩余寿命预测[J]. 中国电机工程学报, 2016, 36(6): 1643-1649
- [46] J. P. Kharoufeh. Explicit results for wear processes in a Markovian environment[J]. Operations Research Letters, 2003, 31(3): 237-244
- [47] J. P. Kharoufeh, S. M. Cox. Stochastic models for degradation-based reliability[J]. IIE Transactions, 2005, 37(6): 533-542
- [48] J. P. Kharoufeh, C. J. Solo, M. Y. Ulukus. Semi-Markov models for degradation-based reliability[J]. IIE Transactions, 2010, 42(8): 599-612
- [49] Y. Liu, M. J. Zuo, Y. F. Li, et al. Dynamic reliability assessment for multi-state systems utilizing system-level inspection data[J]. IEEE Transactions on Reliability, 2015, 64(4): 1287-1299
- [50] R. B. Chinnam, P. Baruah. Autonomous diagnostics and prognostics in machining processes through competitive learning-driven HMM-based clustering[J]. International Journal of Production Research, 2009, 47(23): 6739-6758
- [51] E. Ramasso, T. Denoeux. Making use of partial knowledge about hidden states in HMMs: an approach based on belief functions[J]. IEEE Transactions on Fuzzy Systems, 2014, 22(2): 395-405
- [52] S. S. H. Zaidi, S. Aviyente, M. Salman, et al. Prognosis of gear failures in DC starter motors using hidden Markov models[J]. IEEE Transactions on Industrial Electronics, 2011, 58(5): 1695-1706
- [53] H. Liao, W. Zhao, H. Guo. Predicting remaining useful life of an individual unit using proportional hazards model and logistic regression model[C]. Annual Reliability and Maintainability Symposium, California, 2006, 127-132
- [54] D. Banjevic, A. Jardine. Calculation of reliability function and remaining useful life for a Markov failure time process[J]. IMA Journal of Management Mathematics, 2006, 17(2): 115-130
- [55] X. S. Si, W. Wang, C. H. Hu, et al. Remaining useful life estimation-a review on the statistical data driven approaches[J]. European Journal of Operational Research, 2011, 213(1): 1-14
- [56] 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, 55(8): 1-13
- [57] S. J. Wu, N. Gebraeel, M. A. Lawley, et al. A neural network integrated decision support system for condition-based optimal predictive maintenance policy[J]. IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans, 2007, 37(2): 226-236

- [58] Z. Tian. An artificial neural network method for remaining useful life prediction of equipment subject to condition monitoring[J]. *Journal of Intelligent Manufacturing*, 2012, 23(2): 227-237
- [59] Z. Tian, L. Wong, N. Safaei. A neural network approach for remaining useful life prediction utilizing both failure and suspension histories[J]. *Mechanical Systems and Signal Processing*, 2010, 24(5): 1542-1555
- [60] N. Gebræel, M. Lawley, R. Liu, 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
- [61] Y. Wu, M. Yuan, S. Dong, et al. Remaining useful life estimation of engineered systems using vanilla lstm neural networks[J]. *Neurocomputing*, 2018, 275: 167-179
- [62] A. L. Ellefsen, E. Bjørlykhaug, V. Æsøy, et al. Remaining useful life predictions for turbofan engine degradation using semi-supervised deep architecture[J]. *Reliability Engineering & System Safety*, 2019, 183: 240-251
- [63] R. Zhao, D. Wang, R. Yan, et al. Machine health monitoring using local feature-based gated recurrent unit networks[J]. *IEEE Transactions on Industrial Electronics*, 2018, 65(2): 1539-1548
- [64] L. Ren, X. Cheng, X. Wang, et al. Multi-scale dense gate recurrent unit networks for bearing remaining useful life prediction[J]. *Future Generation Computer Systems*, 2019, 94: 601-609
- [65] C. G. Huang, H. Z. Huang, Y. F. Li. A bi-directional lstm prognostics method under multiple operational conditions[J]. *IEEE Transactions on Industrial Electronics*, 2019, 66(11): 8792-8802
- [66] X. Li, Q. Ding, J. Q. Sun. Remaining useful life estimation in prognostics using deep convolution neural networks[J]. *Reliability Engineering & System Safety*, 2018, 172: 1-11
- [67] J. Zhu, N. Chen, W. Peng. Estimation of bearing remaining useful life based on multiscale convolutional neural network[J]. *IEEE Transactions on Industrial Electronics*, 2019, 66(4): 3208-3216
- [68] X. Li, W. Zhang, Q. Ding. Deep learning-based remaining useful life estimation of bearings using multi-scale feature extraction[J]. *Reliability Engineering & System Safety*, 2019, 182: 208-218
- [69] J. Deutsch, D. He. Using deep learning-based approach to predict remaining useful life of rotating components[J]. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 2018, 48(1): 11-20
- [70] C. Zhang, P. Lim, A. K. Qin, et al. Multiobjective deep belief networks ensemble for remaining useful life estimation in prognostics[J]. *IEEE Transactions on Neural Networks and Learning Systems*, 2017, 28(10): 2306-2318
- [71] M. Ma, C. Sun, X. Chen. Discriminative deep belief networks with ant colony optimization for health status assessment of machine[J]. *IEEE Transactions on Instrumentation and Measurement*, 2017, 66(12): 3115-3125

- [72] G. Zhao, G. Zhang, Y. Liu, et al. Lithium-ion battery remaining useful life prediction with deep belief network and relevance vector machine[C]. International Conference on Prognostics and Health Management, Dallas, 2017, 7-13
- [73] V. N. Vapnik. An overview of statistical learning theory[J]. IEEE Transactions on Neural Networks, 1999, 10(5): 988-999
- [74] J. Liu, V. Vitelli, E. Zio, et al. A novel dynamic-weighted probabilistic support vector regression-based ensemble for prognostics of time series data[J]. IEEE Transactions on Reliability, 2015, 64(4): 1203-1213
- [75] 申中杰, 陈雪峰, 何正嘉, 等. 基于相对特征和多变量支持向量机的滚动轴承剩余寿命预测[J]. 机械工程学报, 2013, 49(2): 183-189
- [76] M. E. Tipping. Sparse bayesian learning and the relevance vector machine[J]. Journal of Machine Learning Research, 2001, 1: 211-244
- [77] A. Widodo, B. S. Yang. Application of relevance vector machine and survival probability to machine degradation assessment[J]. Expert Systems with Applications, 2011, 38(3): 2592-2599
- [78] D. Liu, J. Zhou, D. Pan, et al. Lithium-ion battery remaining useful life estimation with an optimized relevance vector machine algorithm with incremental learning[J]. Measurement, 2015, 63: 143-151.
- [79] C. Chen, B. Zhang, G. Vachtsevanos, et al. Machine condition prediction based on adaptive neuro-fuzzy and high-order particle filtering[J], IEEE Transactions on Industrial Electronics, 2011, 58(9): 4353-4364
- [80] J. Liu, W. Wang, F. Golnaraghi. A multi-step predictor with a variable input pattern for system state forecasting[J]. Mechanical Systems and Signal Processing, 2009, 23(5): 1586-1599
- [81] T. Wang, J. Yu, D. Siegel, et al. A similarity-based prognostics approach for remaining useful life estimation of engineered system[C]. International Conference on Prognostics and Health Management, Denver, 2008, 1-6
- [82] E. Zio, F. Di Maio. A data-driven fuzzy approach for predicting the remaining useful life in dynamic failure scenarios of a nuclear system[J]. Reliability Engineering & System Safety, 2010, 95(1): 49-57
- [83] M. Y. You, G. Meng. A generalized similarity measure for similarity-based residual life prediction[J]. Proceedings of the Institution of Mechanical Engineers, Part E: Journal of Process Mechanical Engineering, 2011, 225(3): 151-160
- [84] J. Shi, Y. Li, G. Wang, et al. Health index synthetization and remaining useful life estimation for turbofan engines based on run-to-failure datasets[J]. Eksploatacja i Niezawodnosc-Maintenance and Reliability, 2016, 18(4): 621-631

- [85] Y. Liu, X. Hu, W. Zhang. Remaining useful life prediction based on health index similarity[J]. Reliability Engineering & System Safety, 2019, 185: 502-510
- [86] S. Malinowski, B. Chebel-Morello, N. Zerhouni. Remaining useful life estimation based on discriminating shapelet extraction[J]. Reliability Engineering & System Safety, 2015, 142: 279-288
- [87] L. Liao, W. Jin, R. Pavel. Enhanced restricted Boltzmann machine with prognosability regularization for prognostics and health assessment[J]. IEEE Transactions on Industrial Electronics, 2016, 63(11): 7076-7083
- [88] 候晓东, 杨江平, 邓斌, 等. 基于多指标相似性的 T/R 组件剩余寿命融合预测方法[J]. 国防科技大学学报, 2019, 41(4): 190-198
- [89] 梁泽明, 姜洪权, 周秉直, 等. 多参数相似性信息融合的剩余寿命预测[J]. 计算机集成制造系统, 2018, 24(4): 813-819
- [90] 尤明懿. 一个拓展的基于相似性的剩余寿命预测框架[J]. 电子产品可靠性与环境试验, 2012, 30(30): 40-48
- [91] 尤明懿. 基于状态监测数据的产品寿命预测与预测维护规划方法研究[D]. 上海: 上海交通大学, 2012
- [92] L. Liao, F. Köttig. 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
- [93] J. Wang, Y. Zheng, P. Wang, et al. A virtual sensing based augmented particle filter for tool condition prognosis[J]. Journal of Manufacturing Processes, 2017, 28: 472-478
- [94] B. Wang, Y. Lei, N. Li, et al. A hybrid prognostics approach for estimating remaining useful life of rolling element bearings[J]. IEEE Transactions on Reliability, 2018, DOI: 10.1109/TR.2018.2882682
- [95] L. Guo, N. Li, F. Jia, et al. A recurrent neural network based health indicator for remaining useful life prediction of bearings[J]. Neurocomputing, 2017, 240: 98-109
- [96] Y. Chang, H. Fang. A hybrid prognostic method for system degradation based on particle filter and relevance vector machine[J]. Reliability Engineering & System Safety, 2019, 186: 51-63
- [97] J. Wu, Y. Su, Y. Cheng, et al. Multi-sensor information fusion for remaining useful life prediction of machining tools by adaptive network based fuzzy inference system[J]. Applied Soft Computing, 2018, 68: 13-23
- [98] H. Sun, D. Cao, Z. Zhao, et al. A hybrid approach to cutting tool remaining useful life prediction based on the wiener process[J]. IEEE Transactions on Reliability, 2018, 67(3): 1294-1303

- [99] C. Chen, B. Zhang, G. Vachtsevanos, et al. Machine condition prediction based on adaptive neuro-fuzzy and high-order particle filtering[J]. IEEE Transactions on Industrial Electronics, 2011, 58(9): 4353-4364
- [100] J. Liu, W. Wang, F. Ma, et al. A data-model-fusion prognostic framework for dynamic system state forecasting[J]. Engineering Applications of Artificial Intelligence, 2012, 25(4): 814-823
- [101] K. Goebel, N. Eklund, P. Bonanni. Fusing competing prediction algorithms for prognostics[C]. IEEE Aerospace Conference, Montana, 2006, 1-10
- [102] K. Goebel, N. Eklund. Prognostic fusion for uncertainty reduction[C]. 2007 AIAA InfoTech at Aerospace Conference, California, 2007, 1373-1382
- [103] C. Hu, B. D. Youn, P. Wang, et al. Ensemble of data-driven prognostic algorithms for robust prediction of remaining useful life[J]. Reliability Engineering & System Safety, 2012, 103: 120-135
- [104] J. Xu, Y. Wang, L. Xu. PHM-oriented integrated fusion prognostics for aircraft engines based on sensor data[J]. IEEE Sensors Journal, 2014, 14(4): 1124-1132
- [105] 谷梦瑶. 不同退化变量下机电设备相似性寿命预测的关键技术研究[D]. 重庆: 重庆大学, 2018
- [106] Ö. F. Eker, F. Camci, I. K. Jennions. A similarity-based prognostics approach for remaining useful life prediction[C]. Second European Conference of the Prognostics and Health Management (PHM) Society, Nantes, 2014, 1-5
- [107] A. Saxena, J. Celaya, B. Saha, et al. Metrics for offline evaluation of prognostic performance[J]. International Journal of Prognostics and Health Management, 2010, 1(1): 4-23.
- [108] NASA Ames Research Center. Turbofan engine degradation simulation data set[EB/OL]. <https://ti.arc.nasa.gov/tech/dash/groups/pcoe/prognostic-data-repository/>, July 7, 2018
- [109] D. An, J. H. Choi, N. H. Kim. Prognostics 101: a tutorial for particle filter-based prognostics algorithm using Matlab[J]. Reliability Engineering & System Safety, 2013, 115: 161-169
- [110] T. Wang. Trajectory similarity based prediction for remaining useful life estimation[D]. Ohio: University of Cincinnati, 2010
- [111] A. Saxena, K. Goebel, D. Simon, et al. Damage propagation modeling for aircraft engine run-to-failure simulation[C]. International Conference on Prognostics and Health Management, Denver, 2008, 1-9
- [112] S. J. Russell, P. Norvig. Artificial intelligence: a modern approach[M]. Malaysia: Pearson Education Limited, 2016
- [113] W. S. Cleveland. Robust locally weighted regression and smoothing scatterplots[J]. Journal of the American Statistical Association, 1979, 74(368): 829-836

- [114] Z. I. Botev, J. F. Grotowski, D. P. Kroese. Kernel density estimation via diffusion[J]. The Annals of Statistics, 2010, 38 (5): 2916-2957.
- [115] Z. I. Botev. Adaptive kernel density estimation in one-dimension[EB/OL]. [https://www.mathworks.com/matlabcentral/fileexchange/58309-adaptive-kernel-density-estimation-in-one-dimension?s\\_tid=prof\\_contriblnk](https://www.mathworks.com/matlabcentral/fileexchange/58309-adaptive-kernel-density-estimation-in-one-dimension?s_tid=prof_contriblnk), September 27, 2019
- [116] MathWorks. Kernel density estimation toolbox for matlab (r13) [EB/OL]. <https://www.ics.uci.edu/~ihler/code/kde.html>, May 18, 2011
- [117] A. Saxena, J. Celaya, E. Balaban, et al. Metrics for evaluating performance of prognostic techniques[C]. International Conference on Prognostics and Health Management, Denver, 2008, 1-17
- [118] L. Liu, S. Wang, D. Liu, et al. Entropy-based sensor selection for condition monitoring and prognostics of aircraft engine[J]. Microelectronics Reliability, 2015, 55(9-10): 2092-2096
- [119] P. Lim, C. K. Goh, K. C. Tan, et al. Multimodal degradation prognostics based on switching kalman filter ensemble[J]. IEEE Transactions on Neural Networks and Learning Systems, 2017, 28 (1): 136-148.
- [120] G. S. Babu, P. Zhao, X. L. Li. Deep convolutional neural network based regression approach for estimation of remaining useful life[C]. International Conference on Database Systems for Advanced Applications, Dallas, 2016, 214-228
- [121] L. Peel. Data driven prognostics using a Kalman filter ensemble of neural network models[C]. International Conference on Prognostics and Health Management, Denver, 2008, 1-6
- [122] 林蕾. 基于循环神经网络模型的遥感影像时间序列分类及变化检测方法研究[D]. 北京: 中国科学院大学, 2018
- [123] 仇悦. 基于循环神经网络的胎儿心电提取方法研究[D]. 杭州: 浙江大学, 2018
- [124] I. Sutskever, J. Martens, G. Dahl, et al. On the importance of initialization and momentum in deep learning[C]. International Conference on Machine Learning, Atlanta, 2013, 1139-1147
- [125] P. J. Werbos, Backpropagation through time: what it does and how to do it[J]. Proceedings of the IEEE, 1990, 78(10): 1550-1560
- [126] S. Hochreiter, J. Schmidhuber. Long short-term memory[J]. Neural Computation, 1997, 9(8): 1735-1780
- [127] J. Chung, C. Gulcehre, K. Cho, et al. Empirical evaluation of gated recurrent neural networks on sequence modeling[J]. arXiv preprint:1412.3555, 2014
- [128] F. A. Gers, J. Schmidhuber. Recurrent nets that time and count[C]. International Joint Conference on Neural Networks, Como, 2000, 189-194

- [129] K. Greff, R. K. Srivastava, J. Koutník, et al. LSTM: a search space odyssey[J]. IEEE Transactions on Neural Networks and Learning Systems, 2017, 28(10): 2222-2232
- [130] T. Mikolov, M. Karafiát, L. Burget, et al. Recurrent neural network based language model[C]. Eleventh Annual Conference of the International Speech Communication Association, Makuhari, 2010, 1045-1048
- [131] D. Bahdanau, K. Cho, Y. Bengio. Neural machine translation by jointly learning to align and translate[J]. arXiv preprint:1409.0473, 2014
- [132] S. Bell, C. Lawrence Zitnick, K. Bala, et al. Inside-outside net: detecting objects in context with skip pooling and recurrent neural networks[C]. IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, 2016, 2874-2883
- [133] R. Zhao, R. Yan, J. Wang, et al. Learning to monitor machine health with convolutional bi-directional lstm networks[J]. Sensors, 2017, 17(2): 273-290, 2017
- [134] Y. Wu, S. Zhang, Y. Zhang, et al. On multiplicative integration with recurrent neural networks[C]. Advances in Neural Information Processing Systems, Barcelona, 2016, 2856-2864
- [135] K. Cho, B. van Merriënboer, C. Gulcehre, et al. Learning phrase representations using RNN encoder-decoder for statistical machine translation[J]. arXiv preprint:1406.1078, 2014
- [136] Q. V. Le, N. Jaitly, G. E. Hinton. A simple way to initialize recurrent networks of rectified linear units[J]. arXiv preprint:1504.00941, 2015
- [137] A. M. Saxe, J. L. McClelland, S. Ganguli. Exact solutions to the nonlinear dynamics of learning in deep linear neural networks[J]. arXiv preprint:1312.6120, 2013
- [138] X. Glorot, Y. Bengio. Understanding the difficulty of training deep feedforward neural networks[C]. International Conference on Artificial Intelligence and Statistics, Sardinia, 2010, 249-256
- [139] R. Jozefowicz, W. Zaremba, I. Sutskever. An empirical exploration of recurrent network architectures[C]. International Conference on Machine Learning, Lille, 2015, 2342-2350
- [140] D. P. Kingma, J. Ba. Adam: A method for stochastic optimization[J]. arXiv preprint:1412.6980, 2014
- [141] G. B. Orr, K. R. Müller. Neural networks: tricks of the trade[M]. Berlin: Springer, 1998
- [142] N. Srivastava, G. Hinton, A. Krizhevsky, et al. Dropout: a simple way to prevent neural networks from overfitting[J]. The Journal of Machine Learning Research, 2014, 15(1): 1929-1958
- [143] F. O. Heimes. Recurrent neural networks for remaining useful life estimation[C]. IEEE Conference on Prognostics and Health Management, Denver, 2008, 1-6.
- [144] S. Zheng, K. Ristovski, A. Farahat, et al. Long short-term memory network for remaining useful life estimation[C]. IEEE Conference on Prognostics and Health Management, Dallas, 2017, 88-95

- [145] M. S. Haque, S. Choi, J. Baek. Auxiliary particle filtering-based estimation of remaining useful life of IGBT[J]. IEEE Transactions on Industrial Electronics, 2017, 65(3): 2693-2703
- [146] R. Zhao, R. Yan, Z. Chen, et al. Deep learning and its applications to machine health monitoring[J]. Mechanical Systems and Signal Processing, 2019, 115: 213-237
- [147] M. S. Arulampalam, S. Maskell, N. Gordon, et al. A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking[J]. IEEE Transactions on Signal Processing, 2002, 50(2): 174-188
- [148] P. Baraldi, M. Compare, S. Saucio, et al. Ensemble neural network-based particle filtering for prognostics[J]. Mechanical Systems and Signal Processing, 2013, 41(1-2): 288-300
- [149] Y. Liao, L. Zhang, C. Liu. Uncertainty prediction of remaining useful life using long short-term memory network based on bootstrap method[C]. IEEE International Conference on Prognostics and Health Management, Seattle, 2018, 1-8
- [150] PHM Society. 2010 PHM Society Conference Data Challenge[EB/OL]. <https://www.phmsociety.org/competition/phm/10>, June 20, 2010
- [151] X. Li, B. S. Lim, J. H. Zhou, et al. Fuzzy neural network modelling for tool wear estimation in dry milling operation[C]. Annual Conference of the Prognostics and Health Management Society, San Diego, 2009, 1-11
- [152] N. Li, Y. Lei, J. Lin, et al. An improved exponential model for predicting remaining useful life of rolling element bearings[J]. IEEE Transactions on Industrial Electronics, 2015, 62(12): 7762-7773
- [153] A. Grossmann, J. Morlet. Decomposition of Hardy functions into square integrable wavelets of constant shape[J]. SIAM Journal on Mathematical Analysis, 1984, 15(4): 723-736
- [154] Y. LeCun, Y. Bengio. Convolutional networks for images, speech, and time series[J]. The Handbook of Brain Theory and Neural Networks, 1995, 3361(10): 1-14
- [155] I. Goodfellow, Y. Bengio, A. Courville. Deep learning[M]. London: MIT Press, 2016
- [156] S. Ioffe, C. Szegedy. Batch normalization: Accelerating deep network training by reducing internal covariate shift[J]. arXiv preprint:1502.03167, 2015
- [157] K. Li, R. Wang, H. Lei, et al. Interval prediction of solar power using an improved bootstrap method[J]. Solar Energy, 2018, 159: 97-112
- [158] Y. S. Lee, S. Scholtes. Empirical prediction intervals revisited[J]. International Journal of Forecasting, 2014, 30(2): 217-234

## 攻读博士学位期间取得的成果

### 一、在学期间参与的项目研究

- [1] 国家自然科学基金项目：小子样多源异种不确定信息下基于故障物理的主轴轴承寿命和可靠性评估方法，项目编号：51875089
- [2] 航空发动机及燃气轮机重大专项基础研究项目：多场载荷/环境下 XX 疲劳可靠性设计理论、方法与验证，项目编号：2017-IV-0009
- [3] 总装备部武器装备预研基金项目：航空发动机 XX 诊断与预测方法研究，项目编号：9140A27020212DZ02310

### 二、在学期间发表的学术论文与专利申请

- [1] **Cheng-Geng Huang**, Hong-Zhong Huang, Yan-Feng Li. A Bidirectional LSTM prognostics method under multiple operational conditions[J]. IEEE Transactions on Industrial Electronics, 2019, 66(11): 8792-8802 (SCI 检索)
- [2] **Cheng-Geng Huang**, Xian-Hui Yin, Hong-Zhong Huang, Yan-Feng Li. An enhanced deep learning-based fusion prognostic method for RUL prediction [J]. IEEE Transactions on Reliability, 2019, DOI: 10.1109/TR.2019.2948705 (已录用, SCI 检索源期刊)
- [3] **Cheng-Geng Huang**, Hong-Zhong Huang, Weiwen Peng, Tudi Huang. Improved trajectory similarity-based approach for turbofan engine prognostics[J]. Journal of Mechanical Science and Technology, 2019, 33(10): 1-14 (SCI 检索)
- [4] Hong-Zhong Huang, **Cheng-Geng Huang**, Zhaochun Peng, Yan-Feng Li, Hengsu Yin. Fatigue life prediction of fan blade using nominal stress method and cumulative fatigue damage theory[J]. International Journal of Turbo & Jet-Engines, 2017, DOI: 10.1515/tjj-2017-0015 (SCI 检索源期刊)
- [5] **Cheng-Geng Huang**, Yan-Feng Li, Shu-Ying Li, Hai-Qing Li. Dynamic fault tree analysis for explosive logic network with two-input-one-output[J]. Journal of Donghua University (English Edition), 2017, 32(1): 140-143 (EI 检索)
- [6] **Cheng-Geng Huang**, Xiang-Yu Li, Hong-Zhong Huang, Yan-Feng Li. Fault prognosis of engineered systems: a deep learning perspective[C]. Proceedings of 2019 Annual Reliability and Maintainability Symposium (RAMS 2019), Orlando, 2019, 1-7 (EI 检索)

- [7] Xiang Li, Hong-Zhong Huang, Xiang-Yu Li, **Cheng-Geng Huang**. Reliability evaluation for the C4ISR communication system via propagation model[C]. Proceedings of 2019 Annual Reliability and Maintainability Symposium (RAMS 2019), Orlando, 2019, 1-5 (EI 检索)
- [8] **Cheng-Geng Huang**, Hong-Zhong Huang, Weiwen Peng, Jie Zhou, Chaoyang Xie. PoF assessment of thermomechanical fatigue in solder joints in multichip module[C]. Proceedings of 2016 International Conference on Quality, Reliability, Risk, Maintenance, and Safety Engineering and 2016 World Congress on Engineering Asset Management (QR2MSE 2016 & WCEAM 2016), 2016, Jiuzhaigou, 317-324
- [9] 黄洪钟, **黄承庚**, 郭来小, 李彦锋, 殷毅超, 郭骏宇, 米金华. 电子封装焊点疲劳寿命分析方法 [P], 中国, 发明专利, 申请号: CN201710073143.9, 2017
- [10] 黄洪钟, 郭来小, 黄思思, 黄土地, 曾颖, **黄承庚**, 郭骏宇. 基于故障物理的功率电子器件失效分析方法 [P], 中国, 发明专利, 公开号: CN108256276A, 2018.