

分类号:TP273

10710-2020132056



长安大学

硕士学位论文

基于数据驱动与融合技术的锂电池剩余使用寿命预测研究

马福荣

导师姓名职称	邱彦章 副教授		
申请学位类别	工学硕士	学科专业名称	控制科学与工程
论文提交日期	2023 年 4 月 3 日	论文答辩日期	2023 年 6 月 5 日
学位授予单位	长安大学		

Research on Prediction of Remaining Useful Life of Lithium Batteries based on Data-driven and Fusion Technology

A Thesis Submitted for the Degree of Master

Candidate: Ma FuRong

Supervisor: Prof. Qiu Yanzhang

Chang'an University, Xi'an, China

答辩委员会名单

答辩委员会	姓名	职称	工作单位
主席	巨永锋	教授	长安大学
委员	文常保	教授	长安大学
委员	李演明	副教授	长安大学
委员	朱玮	副教授	长安大学
委员	解俊虎	高级工程师	中国兵器工业第二〇五研究所
秘书	杜凯	副教授	长安大学

摘 要

锂电池作为新能源产业中的重要组成部分, 凭借诸多优点得到广泛的应用与发展。对于锂电池老化, 可靠的锂电池剩余使用寿命(Remaining Useful Life, RUL)预测可以保障其安全高效运行, 及时做出预警等, 因此锂电池 RUL 的有效预测成为锂电池管理系统的核心研究内容之一。数据驱动方式只需电池历史数据而无需电池内部机理知识, 因此本文使用数据驱动型的深度学习融合算法对锂电池 RUL 预测进行研究, 主要内容为:

建立了基于卷积神经网络(Convolutional Neural Network, CNN)和基于长短期记忆(Long Short Term Memory, LSTM)网络的两种锂电池 RUL 预测模型。使用滑动窗口将历史容量数据作为输入, 分别通过 CNN 和 LSTM 网络两种模型来提取锂电池容量退化特性, 对未来容量的变化做出预测从而确定 RUL 终止周期, 根据终止周期获得剩余寿命。通过两模型各自的实验结果, 验证了深度学习模型在锂电池 RUL 预测中的有效性, 并且表明擅长处理时序数据的 LSTM 网络在锂电池 RUL 预测中具有更好的表现。

为进一步提高 RUL 预测精度, 结合 CNN 与 LSTM 的优势, 提出了一种基于混合网络与注意力机制(Attention)的融合型模型 CNN-LSTM-Attention。使用滑动窗口输入设定步长的锂电池连续历史容量值, 利用 CNN 对历史容量数据进行空间特征提取, 然后用 LSTM 网络对 CNN 输出的向量进行时间特征提取, 最后通过注意力机制来捕获 LSTM 输出向量的关键时空特性, 给出单步前向容量预测结果, 进行连续迭代运算获得未来时序容量, 进而获得 RUL 预测值, 并在网络中加入 Dropout 以避免过拟合现象。实验结果表明, CNN-LSTM-Attention 网络能精准地对锂电池 RUL 做出预测, 并且与单一的模型及未引入注意力机制的混合模型进行对比, 表明该网络具有更高的精度与泛化能力。

RUL 结果的不确定性分析可为锂电池的维护等提供决策意见, 为获得锂电池 RUL 预测结果的不确定性, 提出蒙特卡洛(Monte Carlo, MC) Dropout 方法结合 CNN-LSTM-Attention 网络来获得 RUL 预测结果不确定性。基于变分推断后的贝叶斯网络可以给出结果的概率分布, 提出加入 Dropout 技术可使得 CNN-LSTM-Attention 网络近似于贝叶斯网络, 再结合 MC 采样来获得 RUL 预测结果的不确定性。经过实验结果对比与分析, 验证了 MC Dropout 对于预测结果不确定性量化的有效性, 并且从不确定性量化结果验证了本课题提出的 CNN-LSTM-Attention 比其他网络具有更强的鲁棒性与确定性。

关键词: 锂电池, 剩余使用寿命, CNN, LSTM, 注意力机制

Abstract

As an important part of the new energy industry, Li-ion battery has been widely used and developed with many advantages. For the aging of Li-ion batteries, reliable Remaining Useful Life (RUL) prediction of Li-ion batteries can guarantee their safe and efficient operation and make timely warnings, etc. Therefore, the effective prediction of Li-ion battery RUL has become one of the core research contents of Li-ion battery management system. The Data-driven approach only requires battery historical data without battery internal mechanism knowledge, so this thesis uses Data-driven deep learning fusion algorithm for Li-ion battery RUL prediction, and the main contents are:

Two RUL prediction models for lithium batteries based on Convolutional Neural Network (CNN) and Long Short Term Memory (LSTM) networks were developed. Historical capacity data were used as input using sliding windows, and the capacity degradation characteristics of Li-ion battery were extracted by the CNN and LSTM networks respectively, and the future capacity change was predicted to determine the RUL termination period, and the RUL was obtained according to the termination period. The experimental results of each of the two models validated the effectiveness of the deep learning model in Li-ion battery RUL prediction, and showed that the LSTM network, which is good at handling temporal data, has better performance in Li-ion battery RUL prediction.

To further improve the RUL prediction accuracy, a fusion type model CNN-LSTM-Attention based on hybrid network and Attention mechanism (Attention) was proposed. Using a sliding window to input the continuous historical capacity value of lithium battery with set step, CNN was used to extract spatial features from the historical capacity data, and then the LSTM network was used to extract temporal features from the CNN output vector, and finally the attention mechanism was used to capture the key spatio-temporal characteristics of the LSTM output vector to give the single-step forward capacity prediction results, and successive iterative operations were performed to obtain the future temporal capacity, and then the RUL prediction values were obtained, and Dropout was added to the network to avoid the overfitting phenomenon. The experimental results showed that the CNN-LSTM-Attention network can

make accurate predictions of Li-ion battery RUL, and the comparison with two single models and a hybrid model without the Attention mechanism showed that the CNN-LSTM-Attention network has higher accuracy and generalization ability.

The uncertainty analysis of RUL results can provide decision advice for the maintenance of lithium batteries, etc. In order to obtain the uncertainty of the RUL prediction results for lithium batteries, the Monte Carlo (MC) Dropout method combined with CNN-LSTM-Attention network was proposed to obtain the uncertainty of RUL prediction results. Based on the fact that the Bayesian network after variational inference can give the probability distribution of the results, it is proposed that adding the Dropout technique can make the CNN-LSTM-Attention network approximate the Bayesian network, and then combined with MC sampling to obtain the uncertainty of the RUL prediction results. After comparing and analyzing the experimental results, the effectiveness of MC Dropout for quantifying the uncertainty of prediction results was verified, and the results from the uncertainty quantification verified that the CNN-LSTM-Attention proposed in this project has stronger robustness and certainty than other networks.

Keywords: Lithium battery, Remaining Useful Life, CNN, LSTM, Attention mechanism

目 录

第一章 绪论.....	1
1.1 课题背景及研究的目的和意义	1
1.2 国内外研究现状.....	2
1.2.1 基于机理模型的锂电池剩余寿命预测	3
1.2.2 基于数据驱动的锂电池剩余寿命预测	5
1.2.3 基于融合技术的锂电池剩余寿命预测	8
1.2.4 研究现状分析.....	10
1.3 论文主要研究内容与结构安排	10
第二章 锂离子电池概述与剩余使用寿命理论	13
2.1 锂离子电池原理.....	13
2.1.1 锂离子电池结构与工作过程.....	13
2.1.2 锂电池退化机制.....	14
2.2 锂电池管理系统 BMS	15
2.3 基于历史容量的锂电池剩余寿命预测流程	17
2.4 锂电池剩余寿命的不确定性	20
2.5 本章小结.....	22
第三章 基于 CNN 与 LSTM 的锂电池剩余使用寿命预测	23
3.1 神经网络与深度学习.....	23
3.2 数据集与相关实验设定.....	27
3.2.1 数据集介绍.....	28
3.2.2 实验设定与评价指标.....	29
3.3 基于 CNN 的锂电池剩余寿命预测	30
3.3.1 卷积神经网络 CNN 原理	31
3.3.2 CNN 模型的运行学习过程	33
3.3.3 CNN 网络模型构建	35
3.3.4 实验结果与分析.....	36
3.4 基于 LSTM 的锂电池剩余寿命预测.....	39
3.4.1 循环神经网络 RNN	39
3.4.2 长短时记忆网络 LSTM.....	41
3.4.3 LSTM 网络运行训练算法.....	43
3.4.4 LSTM 网络模型搭建	44
3.4.5 实验结果与分析.....	45
3.5 本章小结.....	49
第四章 基于融合型网络 CNN-LSTM-Attention 的锂电池剩余寿命	

预测	51
4.1 深度学习 Dropout 技术	51
4.2 改进型 CNN-LSTM 混合网络	53
4.3 注意力机制 Attention	56
4.4 基于融合技术的 CNN-LSTM-Attention 网络	57
4.4.1 CNN-LSTM-Attention 网络结构与原理	58
4.4.2 基于 CNN-LSTM-Attention 网络的锂电池 RUL 预测建模	59
4.5 实验结果与分析	61
4.5.1 CNN-LSTM-Attention 网络实验结果	61
4.5.2 实验对比分析	64
4.6 本章小结	66
第五章 基于 MC Dropout 与 CNN-LSTM-Attention 网络的锂电池 RUL 不确定性分析	69
5.1 基于时序预测的锂电池剩余使用寿命不确定性	69
5.2 基于蒙特卡洛模拟的电池 RUL 概率分布	70
5.3 基于贝叶斯方法的预测结果不确定性	71
5.3.1 贝叶斯神经网络的概率模型	71
5.3.2 基于变分推断的贝叶斯神经网络	72
5.4 近似贝叶斯推理的 MC Dropout 算法	74
5.4.1 Dropout 的随机性	74
5.4.2 MC Dropout 算法的近似贝叶斯过程	75
5.5 基于 MC Dropout 与 CNN-LSTM-Att 模型的 RUL 不确定性建模	76
5.6 实验结果与分析	77
5.7 本章小结	81
总结与展望	83
参考文献	85

第一章 绪论

1.1 课题背景及研究的目的和意义

近年来,随着全球经济与科技的飞速进步发展,发展以新能源汽车等为代表的新能源运载装备,是我国在应对气候环境污染、推动生态发展、实现“碳中和”与“碳达峰”战略目标的重要举措^[1]。2020 年 11 月国务院发布了《中共中央关于制定国民经济和社会发展第十四个五年规划和二〇三五年远景目标的建议》,其中明确提出了“发展战略性新兴产业,加快壮大新能源汽车等产业,推动互联网、大数据、人工智能等同各产业深度融合”^[2]。以新能源产业中的电动汽车为典型代表,《新能源汽车产业发展规划(2021-2035)》中明确提到,到 2025 年我国的新能源汽车市场竞争力要明显增强,锂离子动力电池等关键系能源技术需要取得关键性突破,安全水平得到显著提升^[3]。在新能源产业的不断发展中,锂电池依据其独特的优势,发挥了举足轻重的作用。如图 1.1 所示为从 2017 年至 2022 年的中国锂电池出货量统计情况。

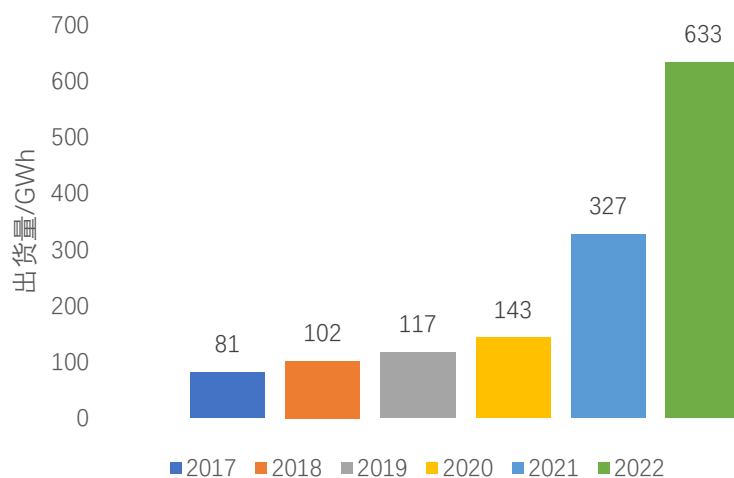


图 1.1 2017-2022 年中国锂电池出货量

对于锂离子电池在新能源产业中的主要应用,其优势主要体现在以下几个方面:

(1) 能量密度高。相较于镍铬电池,锂离子动力电池的能量密度是其能量密度的四倍,并且锂电池能量密度可达镍氢电池的两倍以上至 $350\text{W}\cdot\text{h/kg}$ 。

(2) 工作电压高。锂离子动力电池的常用工作电压是镍铬电池的三倍,可达到 3.6-3.7V。

(3) 循环寿命长。目前一些锂离子电池的循环使用寿命次数可高达 2000 次以上,

在一些低放电深度状况下甚至可达几次,其水平远超镍铬电池等。

(4)自放电率低。有些锂离子电池的自放电率低至 6%每月,远低于镍镉电池(25%)与镍氢电池(15%)。

(5)无记忆性。锂离子电池可以根据应用场景的不同需求而进行充放电,却不会轻易导致电池性能降低。

(6)环境友好。锂离子电池不含有害物质,属于环境友好型电池。

(7)体积灵活。锂离子电池在各种不同应用场景下均能满足体积布置要求。

锂电池在新能源电动汽车等众多领域被广泛使用,然而锂离子动力电池在一些应用场景下被成组使用时,由于每个锂电池单体制造条件与使用场景具有差异性,电池单体老化机制与衰减轨迹和速率都会产生不同的结果,因此锂电池单体电池间的不一致性会随着时间与使用场景的发展而显得尤为明显,从而使得锂电池组的系统状态估计精度下降。当部分电池的健康状况发生恶化时,整个电池系统的整体性能都将受到消极影响,加剧寿命衰减速率,甚至将引发安全隐患。因此,对锂电池的电池保护、精准预测使用寿命、提高电池适应能力、延缓电池性能衰减、保障锂电池预期使用寿命的完成等,成为了锂电池系统管理的重要内容,是保障新能源产业中锂电池的可持续发展的关键所在,同时也是树立锂电池的新能源产业市场信心的基本所在。

锂电池在长期使用过程中,由于高频次的充放电循环,会引起电池活性锂物质与活性材料的损失,造成电池容量退化,从而使得电池剩余使用寿命变短^[4,5]。锂电池剩余使用寿命(Remaining Useful Life,RUL)通常是指从当前观测循环周期到电池寿命终止点周期之间的周期数,可靠的锂电池 RUL 可以保障电池组安全高效地运行,很大程度上可以确保锂电池在使用过程中的安全性以及可靠性,降低电池使用成本和电池故障率,提高锂电池的使用体验,由此可靠的 RUL 预测技术成为锂电池管理系统的核心研究内容之一。

1.2 国内外研究现状

锂离子电池 RUL 预测技术是一个不断发展的热门课题,高精度的 RUL 预测成为众多科研机构不断深入探究的追求目标,目前国内众多高校与研究所都相继展开高水平的锂电池 RUL 研究,同时国外研究机构比如美国国家航空航天局 NASA、新加坡国立大

学及美国马里兰大学等也在 RUL 预测领域具有显著性技术成果。锂电池 RUL 的预测方法有很多，但是由于锂电池作为一个典型的非线性系统，其 RUL 的预测研究具有很大的研究空间。

现阶段的 RUL 预测方式从模型构建角度可以主要分为三类，第一种是根据锂电池的实体结构与特性等系统原理来建立预测模型，第二种是无需依赖系统机理而根据数据来获取电池退化规律与趋势的数据驱动型预测方法，第三种是融合不同种类的预测方法或者优化算法来获得融合性预测模型。但是各种预测方法之间是没有严格的界限的，即许多方法之间都是相互交融的。从模型构建主体角度出发，对于锂电池 RUL 预测方法的分类方法如图 1.2 所示：

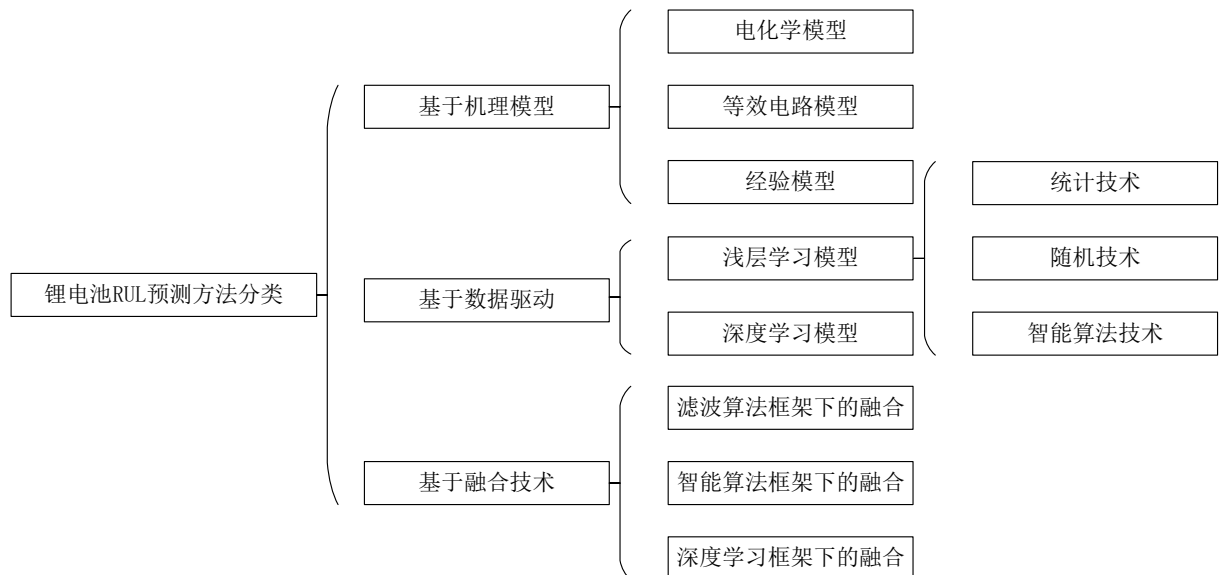


图 1.2 锂电池 RUL 预测方法分类方式

1.2.1 基于机理模型的锂电池剩余寿命预测

锂电池 RUL 预测的机理模型，主要是从影响电池寿命的物理结构或化学原理角度出发，建立锂电池等效模型来获取电池退化机制，来进行 RUL 的预测研究，这种方法的关键点在于对电池构造合适的机理模型。机理模型可以分为电化学模型、等效电路模型与经验模型^[6, 7]。

（1）电化学模型

主要是根据锂电池内部材料及化学机理的动态变化特征来建立退化模型，主要依赖于材料因素如电解液、扩散系数等。Ran 等通过引入电化学阻抗谱，结合电化学原理与最小二乘来拟合锂电池的荷电状态 SOC 等，通过锂电池 SOC 的角度来探究锂电池的退

化机制^[8]。电化学建模方式虽然对电池内部原理的探究较为清晰，但由于其内部结构材料复杂，难以实现对锂电池退化过程趋势的准确把握，尤其是电池退化耦合较为严重时，对于 RUL 预测的建模变得更为困难^[9]。

（2）等效电路模型

等效电路模型是将电池内部物理结构状态通过常见的物理变量如电容、电阻等来构造与电池相同机理的电路模型，这个电路模型与电池具有相同的功能特征。Remmlinger 等人对于电池终端电流与电压进行收集，建立等效电路模型后更新其电路内阻自适应阻抗，以此方式进行退化因子的构造，并对于退化影响因素温度等进行一定的补偿^[10]。Andre 等人对锂电池建立了两个具有差异性的等效电路，对等效电路进行参数化处理，使用 SOC 测量值的阻抗谱数据与环境温度来进行电路参数估计，从而展开对老化机制的研究^[11]。Santhanagopalan 等人构造了一种高速率多孔电极模型，在其中加入了无味滤波器来进行电池 SOC 与老化机制的研究^[12]。等效电路模型相较于电化学模型，具有便于建模与计算的优势，但同时对于电池内部阻抗等参数依赖性较强，实际应用中采集数据等过程具有不便性^[13]。

（3）经验模型

经验模型是利用锂电池经验知识与相关数学理论知识如数理统计等来建立锂电池的容量退化模型。经验模型的建模方式中，较为常用的是基于电池容量退化的指数模型，同时在其中加入一些滤波算法来进行优化，最常用的为卡尔曼滤波与粒子滤波。An 等人从锂电池内阻变化角度建立经验指数增长型模型，对于不同时间状态下的系统退化趋势，结合了标准例子滤波 PF 和重采样技术来进行拟合^[14]。Su 等人构建了基于两类经验指数的模型来进行电池 RUL 的预测，使用 PF 算法来进行参数寻优，获得较为良好的预测效果^[15]。王等人提出一种基于球形立方粒子滤波 SCPF 的状态空间模型，使用该模型分别对 26 个不同的锂电池进行 RUL 预测，其模型的预测精度高于以单一 PF 为构造的模型^[16]。

以上基于机理模型的 RUL 预测方法比较倾向于应用在稳定的场景中，模型相对容易受到环境变化的负面影响，例如电流、温度等的突变，并且在不同的电池环境中产生不同的预测效果甚至产生无法适用该模型的极端情况。

1.2.2 基于数据驱动的锂电池剩余寿命预测

基于数据驱动的锂电池 RUL 预测方法类似于黑匣子模型，主要是通过一定量的数据来近似化地拟合变化趋势，研究模型输入与输出关系，而不必理解电池内部机理，并且具有自适应性，将锂电池 RUL 预测模型的重点放在对数据特征的提取之上。基于数据驱动的锂电池 RUL 预测方法可以分为基于数学理论的统计技术与随机技术、智能算法技术和深度学习模型，其中前两者为浅层学习模型。

(1) 统计技术与随机技术

根据数学统计技术与随机过程原理建立的电池 RUL 预测模型是数据驱动型预测模型重要的一部分，具有十分广泛的应用^[17]。基于自回归 AR 统计理论及其变体的模型主要是利用电池数据来建立一个线性化的关系，将未来的状态值通过统计原理，设定为过去的状态值的线性函数，以此实现拟合目的。Liu 等人在自回归理论的基础上，建立了非线性退化的自回归(ND-AR)的预测模型，用于对时间序列问题的处理，并在其中加入正则化 PF 技术来进行不确定性问题的研究^[18]。Zhou 等人提出将自回归统计模型的变体 ARMA 模型用于锂电池的 RUL 预测中，将该模型与经验模式分解(EMD)相结合，将锂电池的容量全局退化趋势与锂电池健康状态相分离，从而得到 SOH 与 RUL 两个主要电池状态指标^[19]。

在统计技术中，基于卡尔曼滤波 KF 与粒子滤波 PF 的统计原理有很大占比，两种滤波技术不仅常用于机理模型优化，在数据驱动型模型中也具有广泛应用。卡尔曼滤波 KF 主要是使用历史数据来进行未来数据的预测，并且伴随的噪声必须为高斯噪声。例如，Yan 等人提出将具有 Lebesgue 采样技术的扩展卡尔曼滤波模型来对锂电池进行寿命预测^[20]。基于 PF 模型具有更好的可扩展性，粒子滤波相较于卡尔曼滤波具有更广泛的应用。粒子滤波的计算原理主要是基于蒙特卡洛模拟方法，对于样本，使用其发生频率来近似化地模拟该样本的概率分布。PF 相较于 KF，其先进性主要体现在数据分布更为广泛，不局限于单一的高斯分布。Zhang 等人为了了解决 UPF 模型中的样本贫化问题，使用基于马尔可夫链的蒙特卡洛技术于 UPF 相结合，建立改进型的无味粒子滤波器 IUPF，实现对锂电池的 RUL 的高效预测^[21]。基于之前的研究，Yu 等人建立了基于量子粒子群优化算法的粒子滤波算法，通过实验验证，该模型比传统 PF 算法具有更加出色的表现^[22]。Ma 等人提出基于高斯-厄米粒子滤波的锂电池 RUL 预测模型，使得基于 PF 的预

测模型有更进一步的发展^[23]。还有基于其他的数学统计理论算法的 RUL 预测模型,例如 Ng 等人建立了朴素贝叶斯算法框架下的锂电池容量退化模型,对不同工况、温度等实验条件下对比验证 RUL 预测精度,取得了良好的效果^[24]。

基于随机技术的预测模型也有广泛的应用,例如 Li 等人建立了一种基于高斯过程混合(Gaussian Process Mixture,GPM)的新型的 RUL 预测模型,为了更好处理其中的多模态问题,该模型创新地将不同容量退化轨迹加入到不同轨迹相适应的多类型 GPR 模型中,使用 GPM 模型得到预测结果与预测结果的置信区间,都充分验证该模型的优越性^[25]。维纳过程也是较为常用的随机过程模型,其核心是马尔科夫过程,例如 Tang 等人基于维纳过程原理,向中加入了实验测量误差,建立了一种创新性的 RUL 预测模型,并且使用多个实例验证了该方式的有效性与先进性^[26]。

(2) 智能算法技术

基于数据驱动型的智能算法技术已经在锂电池的 RUL 预测方面有个许多应用与研究。其中一些典型的方法包括支持向量机、相关向量机与 Box-Cox 等。

支持向量机(Support Vector Machine,SVM)是一类广义线性分类器,在避免局部最优之类的问题之上取得了显著的效果,具有较高的鲁棒性与稀疏性。SVM 结合核方法可以在非线性高维识别问题上取得良好的效果,并且支持向量回归可以实现回归预测。Li 等人提出基于 SVM 的锂电池 RUL 预测模型,通过多组实验对比验证,SVM 比传统的神经网络具有更高的预测精度和更少的计算量^[27]。Doo 等人建立了基于支持向量回归(Support Vector Regression.SVR)的电池寿命预测模型,该模型在电池循环到达电池寿命终止点处时,可以实现较为精准的 RUL 估计^[28]。在支持向量回归的理论成果基础之上,Wang 等人^[29]建立了一种基于灵活支持向量回归(Flexible Support Vector Regression,F-SVR)的预测模型,该模型相较于其他的模型,其特点在于具有非迭代性,并将该模型与具有多步迭代的 SVR 模型相结合,从而对低维度的输入数据进行很好的处理。Zhao 等人创新性地提出两个电池健康指标,即相等充电电压的时间差以及相等放电电压的时间差,同时结合了特征向量的选择和支持向量回归 SVR 来建立锂电池的 RUL 预测模型^[30]。SVR 在电池的 RUL 预测中取得了出色的成绩,但是对于大数据下的电池 RUL 预测需求,表现出一定的非适应性。锂电池的 RUL 预测不仅对于预测精度有要求,并且对于模型要求可以给出结果的概率分布,由此,RVM 得到了相比 SVM 更为广泛的应用。

相关向量机(Relevance Vector Machine, RVM)是一种贝叶斯稀疏核方法,该方法比 SVM 具有一定的泛化能力,避免了 SVM 要求核函数为正定核之类的局限性。Hu 等人首次使用稀疏贝叶斯方法来处理 RUL 问题,对于所建立的 RVM 模型,将电流电压作为模型输入,该方法被应用于医疗设备中的电池设备中,取得了良好的预测效果^[31]。

Box-Cox 变换算法是由 George Box 与 David Cox 提出的一种广义幂变换算法^[32],是在数学统计建模领域常用的数据变换方式之一,常用于处理误差以及提高变量关联性,对于不满足正态分布的连续性数据具有强大的处理能力,在锂电池的 RUL 预测领域也具有广泛的应用。例如 Zhang 等人建立了一种基于蒙特卡洛算法与 Box-Cox 算法的预测方法,其中使用 Box-Cox 方法来构建电池容量与电池循环之间的线性模型,再使用蒙特卡洛模拟采样来产生预测结果的不确定性^[33]。Liu 等人将 Box-Cox 变换方法加入到所建立的 SOH 于 RUL 预测模型中的健康因子构建环节,主要被用来提升健康因子与电池退化状态的关联性^[34]。

使用机器学习智能算法对于锂电池的 RUL 具有出色的预测结果,得到了较为广泛的应用,比统计方式与随机过程具有更高的适用范围,但同时大部分机器学习智能算法对于电池 RUL 预测结果的不确定性分析缺少支撑。

(3) 深度学习模型

除了上述的基于统计技术与随机技术和机器学习智能算法的锂电池 RUL 预测方法之外,基于深层次神经网络模型的深度学习预测方法得到了越来越广泛的应用,深度学习再图像处理、语义分析等各方面都有非常出色的表现,并且也逐渐被应用于锂电池的 RUL 预测研究之中。深度学习模型中,常见的基本模型有卷积神经网络 CNN 与循环神经网络 RNN。

卷积神经网络是包含了卷积操作、池化操作等技术的前馈型运算网络,是深度学习模型典型代表之一,对于大型图像处理等空间特征明显的问题具有强大的处理能力。在 RUL 预测中,Cheng 等人首先构建了基于希尔伯特-黄变换的退化指标提取模型,用于对轴承振动信号提取指标,然后利用 CNN 拟合了锂电池容量退化曲线,最后使用向量回归方法来实现最终的 RUL 预测^[35]。Ince 等人建立了基于一维卷积神经网络模型的电机故障检测系统,该 CNN 模型可以通过原始信号的特征提取来对电机的运行状态做出实时监测,从而提高整个系统的检测效率,通过实验验证了 CNN 网络对于电机状态监

测的可行性^[36]。

循环神经网络在深度学习中主要被用于处理时序问题，该网络擅长计算时间依赖型的数据。Liu 等人建立了一种基于自适应的 RNN 预测模型来对系统状态进行状态监测，在其中的自适应机制中使用递归文伯格-马夸特方法来对 RNN 网络权重参数进行优化更新，经过实验对比分析，验证了该方法比常规的机器学习算法具有更强的学习能力，例如 SVM 等^[37]。Zhang 等人在常规 RNN 的理论基础上，使用长短期记忆网络 LSTM 来对 RNN 进行提升，克服了 RNN 对于长时间依赖关系的局限性，并在网络中加入 Dropout 技术来避免过拟合现象，最后还对 RUL 预测结果进行了不确定性分析，经过实验验证表明，LSTM 比常规 RNN 具有更强的时间序列问题处理能力^[38]。

除了常见的深度学习模型 CNN 和 RNN 外，还有一些其他的深度学习模型，例如 Khumprom 等人使用基于深度神经网络 DNN 的锂电池 RUL 模型来对锂电池进行状态 SOH 与 RUL 预测，并与其他机器学习算法进行比较，验证了该模型的先进性与有效性^[39]。相较于其他的数据驱动类模型，深度学习方法具有更加强大的数据运算能力，并具有可以自主训练学习的独特优势，受到越来越广泛的应用。

1.2.3 基于融合技术的锂电池剩余寿命预测

模型的融合技术是为了补充单一模型具有的某些不足而结合不同种类的模型，来达到更高标准的组合方法，从而使模型更好地从全局捕捉模型特征，获得更好的预测效果^[40]。在融合模型中，有一部分主要是基于滤波算法的融合机制，滤波技术包含粒子滤波与卡尔曼滤波以及两种滤波的变体。还有一部分是将一些机器学习等智能算法与相关模型相结合的融合机制。

（1）滤波算法框架下的融合模型

基于滤波算法框架的融合模型也是主要包含了卡尔曼滤波与粒子滤波。对于卡尔曼滤波算法框架下的融合模型，与常规 EKF 相对比，郑等人提出将无味卡尔曼滤波 UKF 与一种特殊的基于非线性时间序列预测方法相结合，构造的融合模型可以根据 UKF 与电池容量回归数值共同更新电池等效模型的实时状态，该模型的预测精度相较于扩展卡尔曼滤波具备更高的精度与鲁棒性^[41]。Song 等人将卡尔曼滤波与机器学习算法相关向量机 RVM 相结合构造融合模型来进行电池剩余寿命的预测，同时根据 RVM 长期预测效果弱于短期效果的特性，提出一种对 SVM 进行更新迭代的优化算法，以此达到长期

预测的目的^[42]。Chang 等人在相关理论基础上,将 UKF、RVM 及经验模式分解 EMD 三者结合起来构造融合模型,该种结合三种技术的 RUL 预测模型在预测精度上取得了明显的效果^[43]。

相较于卡尔曼滤波框架下的融合模型,粒子滤波 PF 框架下的融合机制受到更加广泛的应用。Dong 等人将粒子滤波与基于布朗运动过程容量退化模型相结合起来,构造 RUL 的融合预测模型,该模型相较于高斯回归模型,具有更高的精度与稳定性^[44]。Zhang 等人将粒子滤波与指数模型相结合起来组建 RUL 预测模型,通过与自回归积分滑动平均算法进行比较,验证该融合模型具有一定的先进性^[45]。Song 等人提出将粒子滤波与一种优化自回归模型 IND-AR 结合,形成一种结合二者优点的混合型模型,该模型适用于广泛的非线性容量退化预测,适合用于改善锂电池 RUL 模型的长期退化预测失效现象^[46]。Cadini 等人使用 PF 算法与多层感知器 MLP 的算法相结合,主要是将 PF 算法中的观测方程表示为 MLP 模型,构建的预测模型取得了高精度的结果^[47]。

(2) 智能算法与深度学习框架下的融合模型

机器学习智能算法框架下的融合模型也具有广泛的应用。Hu 等人数据驱动型的模型与机理模型相融合,来建立锂电池的 RUL 预测模型,该融合型预测模型由两大主要部分组成,第一部分首先利用稀疏贝叶斯方法拟合锂电池的电荷特征,从而计算电池容量,第二部分则是使用递归贝叶斯来更新从第一部分得到的容量值,通过两部分的组合来实现未来容量预测目的^[48]。He 等人建立了基于贝叶斯蒙特卡洛算法与 DS 算法的融合型预测模型,其中的 DST 模块主要训练数据的参数初始化,贝叶斯蒙特卡罗模块主要用于更新模型参数^[49]。Peng 等人将混合高斯过程函数 HGPFR 与小波去噪 WD 算法相结合,其中的 WD 模块主要用于减少数据噪声,通过实验对比,融合了 WD 模块的模型比单一的 HGPFR 模型据有更高的精度^[50]。

随着机器学习与深度学习的发展,越来越多的学者将智能算法与深度学习融合,构建优化型的深度学习预测模型。例如 Chen 等人构建了灰色神经网络 GNNM 与粒子群优化算法的融合模型,并且使用了移动窗口的方式对构建的网络进行迭代训练,通过实验对比验证,表明 PSO-GNNM 具有良好的 RUL 预测效果^[51]。有些学者将不同的深度学习网络模型相融合,例如 Li 等人构建了一种基于 LSTM 网络与 Elman 网络的融合模型用于电池的 RUL 预测,该模型使用经验模式分解方法将锂电池容量进行分层处理,然

后再用 LSTM 与 Elman 实现不同层的协作处理,实验结果表明该模型可以准确预测电池 RUL^[52]。

1.2.4 研究现状分析

对于锂电池的 RUL 预测,基于机理模型的预测方式具有无需依靠大量数据便可完成电池退化过程建模的优势,但同时也存在预测不够精准、建模方式困难等诸多劣势,该方法需要对锂电池的结构及退化机制具备充分的理论知识与深刻的理解,涉及物理结构、化学组成等诸多研究领域,同时也较容易受到实验环境的影响,因此,基于机理模型的锂电池 RUL 预测方式不是主流方式。

与机理模型相对应,基于数据驱动的锂电池 RUL 建模方式则是主要依靠数据来实现,无需了解电池内部机理,通过对电池数据的深度挖掘与特征分析,来建立锂电池的容量退化模型。传统的数据驱动方式主要是依赖于数学建模理论与智能算法,对于电池数据是浅层的学习过程,并且一般需要手动提取数据特征,而深度学习的方式则可以直接从数据中自动提取潜在的特征与趋势,极大的提高了数据驱动型建模的效率,因而数据驱动型的深度学习方式成为锂电池 RUL 预测与电池类非线性系统研究的主流方式之一。

随着锂电池 RUL 预测技术的不断发展,目前仍存在对数据利用率不够高以及各类模型参数优化策略不够完善等诸多问题,但主要问题始终还是聚焦于电池寿命预测的精准性。由于各种模型利弊共存,因此基于融合技术的预测模型顺理成章成为了近些年对于电池 RUL 预测研究的热门课题。深度学习的方式作为锂电池研究领域的后起之秀,在电池系统领域的研究尚未完全成熟,因此探究基于深度学习的融合模型具有更多的研究价值与现实意义。

1.3 论文主要研究内容与结构安排

本文在已有锂电池相关研究的基础上,对锂电池的剩余使用寿命预测方法展开研究。首先从锂离子电池结构组成与工作原理的角度入手,分析了锂电池的退化机制与老化因素,并且介绍锂电池管理系统及锂电池寿命预测对于电池保护和电池管理的重要性。以此为基础,提出数据驱动型的电池 RUL 预测流程和分析锂电池 RUL 预测结果不确定性的研究意义。在相关神经网络与深度学习的理论基础之上,建立了基于 CNN 网络与

LSTM 网络的锂电池 RUL 预测模型，对实验结果进行对比与分析，为进一步提高锂电池 RUL 预测精度，结合 CNN 与 LSTM 优点构建混合网络，向其中加入注意力机制，建立了融合型的 CNN-LSTM-Attention 网络，并在其中加入 Dropout 技术，进一步提高了锂电池 RUL 预测模型的精度与鲁棒性。最后，基于贝叶斯网络相关理论与神经网络 Dropout 技术的特点，建立了基于蒙特卡洛 Dropout 方法与 CNN-LSTM-Attention 网络的锂电池 RUL 预测模型并给出其置信区间与概率分布，对锂电池 RUL 预测结果的不确定性进行了分析，验证所提出的模型的有效型与先进性。

本文主要研究内容将通过以下组织结构来展开：

第二章主要介绍锂电池的原理，其中包括锂电池结构与工作过程，分析锂电池的老化因素与退化机制，为锂电池寿命预测提供理论基础。介绍锂电池管理系统的结构与功能，展示其对锂电池系统的重要意义，并以此揭示锂电池剩余寿命预测对于电池与电池管理系统的现实价值。介绍基于历史容量数据驱动的锂电池 RUL 预测流程，其中重点介绍电池容量时序数据的操作过程。介绍锂电池剩余使用寿命不确定性的研究意义，并对其进行成因、分类、度量等相关理论的介绍，为后文的研究提供理论基础。

第三章首先从神经网络和深度学习的相关原理与运算过程出发，分析数据驱动型的深度学习方法在锂电池 RUL 预测中的可行性。介绍本文实验所使用的 NASA 数据集与 CALCE 数据集的电池容量退化数据来源及数据预处理，对本课题的相关实验做出统一设定，并建立相关预测结果的评价指标。介绍深度学习典型代表 CNN 网络与 LSTM 网络的原理与学习过程，建立基于 CNN 网络与 LSTM 网络的锂电池剩余使用寿命预测模型，经过实验验证与对比分析，分析两种模型对于锂电池 RUL 预测的现实价值与意义。

第四章针对前文中 CNN 网络与 LSTM 网络对于锂电池寿命预测的改进必要性，结合 CNN 网络与 LSTM 网络的优势，建立了混合型的 CNN-LSTM 网络。介绍深度学习中的 Dropout 技术，将其加入至混合网络中来避免学习过程中的过拟合，同时为后文奠定基础。为进一步提高锂电池 RUL 预测精度，引入注意力机制，建立融合型的 CNN-LSTM-Attention 网络，充分利用 CNN 网络与 LSTM 网络各自优势，通过注意力机制加强相关信息的关联，经过实验验证该模型的有效性，并且通过不同模型的对比分析，展示所提出融合型网络的先进性与泛化能力。

第五章针对所提出的 CNN-LSTM-Attention 网络无法给出 RUL 预测结果不确定性

的问题，首先对电池 RUL 预测结果不确定性进行介绍与分析，介绍在锂电池 RUL 预测领域常用的蒙特卡洛采样方法计算概率分布的过程。根据贝叶斯网络相关理论与变分推断，验证贝叶斯方法在 RUL 预测结果不确定性的可行性，提出神经网络中加入 Dropout 来逼近贝叶斯的变分推断，从而使得含有 Dropout 的网络也可给出预测结果概率分布与置信区间，通过蒙特卡洛采样方式得到锂电池 RUL 预测结果的概率分布，从而对预测模型进行不确定性度量，通过实验验证该 MC Dropout 方法的可行性及 CNN-LSTM-Attention 网络相较于其他网络的确定性与稳定性。

第二章 锂离子电池概述与剩余使用寿命理论

锂电池是现代各类电子产品与各类动力设备中最具代表性的能量储备电池类型，基于其寿命长、能量高等诸多优点受到广泛的应用^[53]。锂电池作为本课题的研究对象，本章基于锂电池的应用，介绍锂电池原理中的结构与工作过程，以及锂电池容量退化机制；介绍电池管理系统的主要内容及优点，以及锂电池的剩余寿命预测对于电池管理系统的重要意义；介绍使用数据驱动的方式来实现锂电池 RUL 预测的流程，着重研究基于历史容量类的时序数据作为模型输入的预测流程；最后介绍了在锂电池剩余寿命预测中的不确定性及其在预测过程中的来源与度量。

2.1 锂离子电池原理

2.1.1 锂离子电池结构与工作过程

锂离子电池结构的主要组成部分有正负两电极、正负极隔膜、温度系数元件、锂盐电解液以及安全阀等。

(1) 锂电池正极主要是用来接收锂离子，含有接受锂离子的空间以及用于扩散的路径，对锂电池的整体性能有着很大的影响。锰酸锂、磷酸铁锂等多种金属氧化物三元材料都是常见的锂电池正极材料。

(2) 锂电池负极材料一般是由碳物质与专用化学粘合物质构成，并在其中一般含有一些具有专用特性的溶剂，这些材料构成粘稠型混合物覆盖于铜基之上，然后按照薄层结构分布，一些新型的快充电池使用钛酸锂作为电池负极材料。

(3) 锂电池的正负级间的隔膜，主要是用隔绝两电极间电子的通行，是专用于锂离子的转移通道，其构成材料通常为聚乙烯或聚丙烯等构成的多孔薄膜材料。

(4) 锂电池电解液存在的意义主要是转移离子，是由某些含有专用化学有机溶剂的锂盐溶液组成，该锂盐溶液具有高导电性。锂电池的电解液须具有一定的化学稳定性，对于反应中的氧化还原反应要有一定的适应性。

(5) 从锂电池使用安全角度出发，其内部通常含有异常电流关断装备，这个装备为温度系数原件，面对电池中常发生的内压骤升异况，通常装配安全阀来释放高压气体，以免强大压差破坏电池结构。

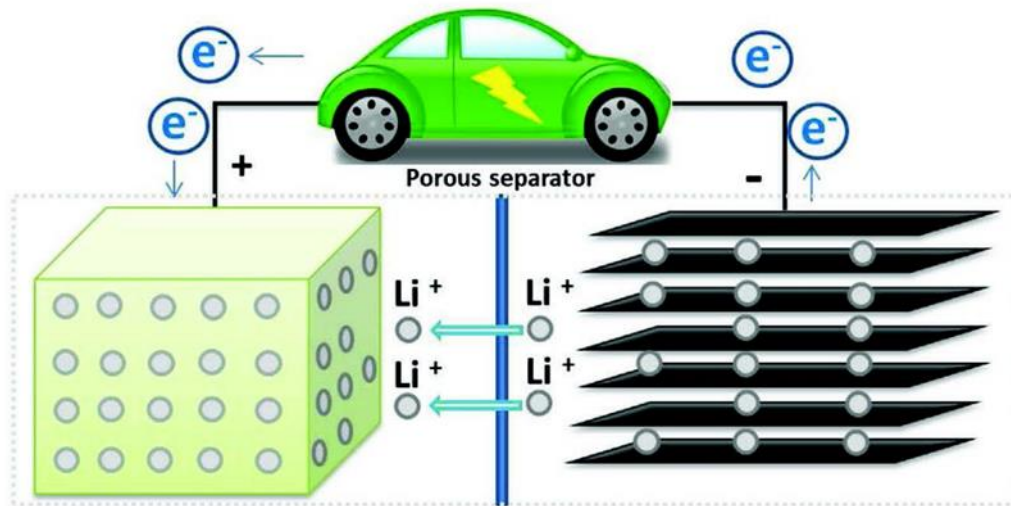


图 2.1 锂电池工作原理

如图 2.1 所示为锂离子动力电池在电动汽车的应用场景下的工作原理，该电池的正极与负极都在电解液之中。对于锂离子电池而言，其充放电过程，主要是基于锂离子不停地在正极与负极材料上进行的嵌入动作与脱离动作。在电池充电过程中，锂电池正极材料中的锂离子脱离而出，进入电解液中，基于电解液中不同电极之间产生锂离子浓度差，从而产生驱动力使得锂离子穿过专用转移通道，该通道即为锂电池正负极隔膜，穿过的锂离子朝着负极移动，靠近负极后靠电荷吸引力从而嵌入电极中。由于电子从正极转移至负极，则在外电路中由于外部驱动力，电子从正极出发，转移至负极，从而形成电流。在整个过程中，正极失去电子发生的是氧化反应，负极则与之相反，得到电子发生还原反应，完成一个完整的锂电池充电过程。放电过程与充电过程相反，即锂离子从负极脱离而出转移至正极，外电路中电子从负极转移至正极，从而驱动电动车等用电器工作，完成整个放电过程，正极上发生还原反应，负极发生氧化反应。在锂离子电池的第一次充电过程中，其负极上会自然形成一种固态膜，被称为固体电解质面(SEI)膜，其形成主要目的是保护负极，随着时间推移与充放电过程的循环，SEI 膜将会被逐渐溶解，从而导致充放电过程中的锂离子与其他化学物质不断减少，由此 SEI 膜对于锂电池的容量退化具有重要影响。

2.1.2 锂电池退化机制

对于锂离子电池的老化现象，主要包含两种老化模式，即日历老化与循环老化。其中，日历老化是指锂电池在未使用情况下，静置状态发生自放电现象，其电池容量衰减

速度较为缓慢,适宜的电池环境或者应用场景可以延缓这种日历老化。而循环老化是指锂电池在使用过程中的充放电循环下,其内部环节各部分性能都出现衰减以及物理老化等^[54]。诱发电池老化的主要因素包括有温度、电流倍率、充放电深度以及充放电截至电压等,还有长期静置或者各类突发环境状态等。

从微观角度,锂离子电池发生老化的本质原因是电池内部的活性物质与锂离子的减少。其中的电池负极老化主要包括 SEI 膜增长、粘结剂分解、集流体的腐蚀等,电池正极的老化则包括活性锂的脱落、结构坍塌、集流体腐蚀以及电解液与正极材料间的化学反应等。

从宏观角度,锂电池的老化主要体现在电池容量的退化、内阻的变大、故障率增加等各方面^[55]。当单体锂电池进行不同组合后,其工作状况与单体状况发生差异,导致内部反应也发生差异。由此电池组内的单体 SOC 与电池剩余可用容量存在一定程度的不一致性,这种不一致性随着老化而逐渐增大,这种机理将进一步促进整个电池组总容量的衰减。

锂电池的容量与剩余寿命估计是监控整个电池的生命周期的主要方面之一,各种电池异常情况与电池老化息息相关。基于耦合关系,电池的容量衰减将导致电池内部反应异常,电极机构老化等,由此更容易引发某些早期微小故障,然而一些微小故障发展成为热失控等较为显著的故障,则将进一步使得电池容量大幅度衰减^[1]。从时间角度,如果电池容量衰减变化尺度拉长,则电池异常故障变化尺度也将变长,从而使得电池安全性能大幅降低。

2.2 锂电池管理系统 BMS

电池管理系统(Battery Management System,BMS)是指以一些方式对电池进行相应控制与管理的产品或者技术^[56]。BMS 主要是由固化算法的各类控制器、执行装置以及其他功能性传感器所组成,其主要任务是提供对用电系统的控制与能量管理所需状态的各类信息,确保用电系统与锂电池的安全可靠性,同时在使用场景下的各类突发状况做出相应干预措施,BMS 结构图如图 2.2 所示。

BMS 可以通过采样电路来实时采集获取各个单体电池或者电池组的端电压、工作电流以及温度等相关数据,然后使用一些既定算法和控制策略来估算电池的 SOC、SOH、

SOP 及 RUL 等，进而将相关参数输出至用电整体控制器，为用电环境的能量控制管理与用电分配情况提供参考依据。BMS 的基本功能主要有以下几点：

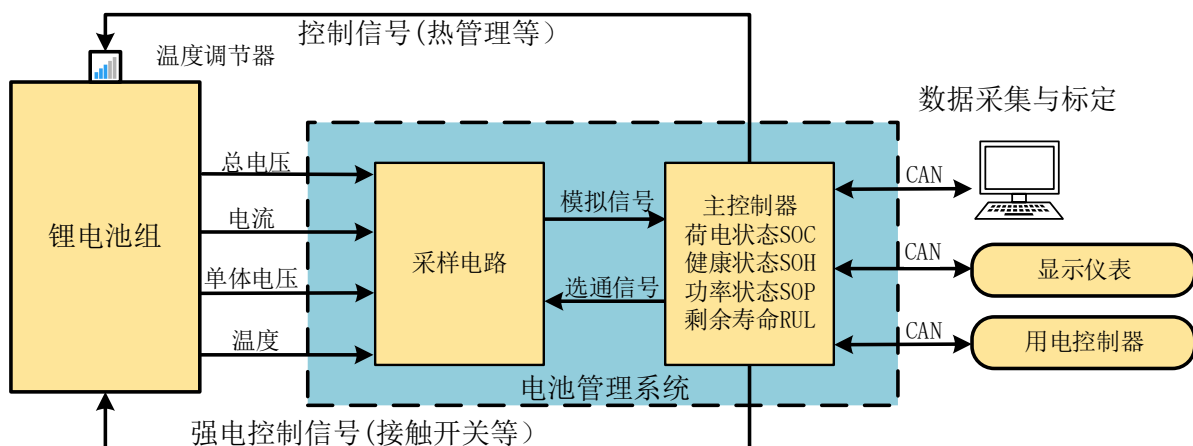


图 2.2 电池管理系统 BMS 结构图

(1) 数据采集。为了准确获取锂电池的运行状况与工作效率，更好地进行对整个系统的管理，BMS 需要通过信息采集系统来准确采集电压电流等关键参数信息^[57]。

(2) 状态监测。锂电池是一个非常复杂的非线性时变系统，它具有诸多实时变化的需要监测的状态量，BMS 需要将检测到的实时电池数据，通过相应控制算法与策略来进行电池各项状态的实时估计，从而获得整个系统的变化状态，其中包括 SOC、SOH、SOP、RUL 及 SOE 等。

(3) 安全保护。锂电池安全保护主要是指电池及其组成系统的意外故障诊断以及安全控制，通过一系列控制系统的传感器获取到的实时信息，BMS 需要及时做出故障反应判断，并且采取相应应对策略来进行及时干预处理，保证整个用电系统的安全。

(4) 充电控制。锂电池的充电过程对于电池寿命与安全具有巨大的直接影响，基于此，BMS 需要集成构建一个充电管理模块，依据锂电池的实时状态与特性、温度变化、功率大小等因素，对电池充电模块进行实时控制，确保安全充电。

(5) 能量控制管理。为保证电池的安全性以及经济性，BMS 需要通过状态信息，合理控制电池的能量输出以及再生制动的能量回收。若用电装置搭配有复合型电源，则 BMS 需要根据复合电源各组成部分的状态信息来进行电池能量的优化分配，保证复合电源整体最佳性能。

(6) 均衡管理。根据生产工艺、存储方式以及电子元器件之间的误差累积效应，各

单体电池之间必然产生不一致性,为保证用电安全与电池保护,充分发挥单体电池的性能,BMS 需要根据单体电池实时信息,采取相对应的均衡措施,尽可能来降低各单体电池间的不一致性给整体系统带来的负面效应。

(7) 热管理。锂电池在工作中不仅受到环境温度的影响,同时也会受到自身充放电产生热量的影响,基于此 BMS 需要集成电池热管理模块,可以根据温度分布情况以及充放电的需求情况来决定加热与散热操作,使得锂电池工作于一个合适的温度中,充分发挥电池的性能,减少温度负面影响。

(8) 信息管理。BMS 根据用电环境的需求,需要集成多个功能性模块,合理协调各功能模块之间的协作运行。由于电池使用过程中的庞大信息量,BMS 需要根据信息功能与重要程度,对信息进行相应处理与筛选,存储关键信息数据,保证整个控制网络节点进行通信,提升管理水平。

2.3 基于历史容量的锂电池剩余寿命预测流程

在现在这个大数据占主导地位信息化时代,数据驱动是锂电池实现 RUL 预测的主要技术手段,数据驱动的核心主要是通过历史容量数据的衰减轨迹进行挖掘提炼与分析推广^[58]。在使用锂电池历史容量数据的 RUL 预测过程中,获取锂电池历史容量数据是进行锂电池 RUL 预测的第一步,建立模型从中发掘内在规律,进而对未来容量数据进行延伸及推广,最终实现锂电池的剩余寿命预测。与其他一些预测机理相较,基于数据驱动型的 RUL 预测方式具有诸多优点,例如计算量少、过程简单且无需来考虑锂电池复杂的应用环境,对于 BMS 的运行有效减轻负担,适用于广泛的应用场景。

对于锂电池的 RUL 预测,依据预测建模思路的不同,使用历史容量数据的 RUL 预测方式可以分为三个主要类型:经验预测法、滤波预测法及时序预测法。

经验预测法是先设定锂电池的容量衰减轨迹遵循于某一种确定性的数学模型,然后构造一个数学表达式,以循环周期数作为这个表达式的输入,输出则为锂电池的最大可用容量,以这个表达式来作为锂电池的老化规律描述方式。一般从数据拟合的角度作为建模切口,以不断试探构造不同数学表达式来作为建模手段,经过多次拟合之后,选用最合适的数学表达式来作为锂电池容量衰减轨迹的表达式,这个确定性的数学式即为锂电池寿命经验模型。常用的基于寿命经验预测的表达关系式如表 2.1 所示:

表 2.1 常用于锂电池寿命经验模型的数学表达式

名称	数学表达式	模型参数
双指数模型	$C_{\max} = b_1 e^{b_2 n} + b_3 e^{b_4 n}$	b_1, b_2, b_3, b_4
单指数模型	$C_{\max} = a_1 e^{a_2 n} + a_3$	a_1, a_2, a_3
多项式模型	$C_{\max} = d_1 n^2 + d_2 n + d_3$	d_1, d_2, d_3
线性模型	$C_{\max} = c_1 n + c_2$	c_1, c_2
Verhulst 模型	$C_{\max} = \frac{e_1 / e_2}{1 + [e_1 / (e_2 C_0) - 1] e^{-e_1 n}}$	e_1, e_2

基于不同的工况与使用阶段，对于锂电池而言，需要匹配不同的经验模型，因此需要提前在 BMS 中录入多种经验模型，设置其适用场景。BMS 需要在不同状况下调用合适的经验寿命模型来逼近求参，以此来得到较为精准的容量衰减拟合轨迹，然后带入相应电池的失效阈值，完成最终的 RUL 预测。基于经验预测法的模型具有建模原理简单、计算过程清晰容易等诸多优点，但由于考虑的因素简单，其预测精度与模型泛化能力相对受限。

滤波预测法的核心是状态估计，主要是通过对数据的实时观测和更新来校正寿命经验模型的参数，是目前较为常见的电池 RUL 预测技术。相对于经验预测法的数据拟合，滤波预测法的参数获取方式具有更强大的适应能力，基于此，滤波预测法也是经验预测法的改进算法，其基础模型也是寿命经验模型。

在使用滤波预测法进行锂电池 RUL 预测时，BMS 将提取锂电池历史容量数据，调用适用的经验寿命模型后，根据相应寿命经验模型建立对应状态空间方程，将数据拟合过程转化为状态估计过程，以双指数寿命经验模型为例，其对应的状态空间如式(2.1)所示：

$$\begin{aligned}
 x_n &= [b_{1,n} \ b_{2,n} \ b_{3,n} \ b_{4,n}]^T \\
 \begin{cases} b_{1,n+1} &= b_{1,n} + w_{b1,n}, \ w_{b1,n} \sim N(0, \sigma_{b1}) \\ b_{2,n+1} &= b_{2,n} + w_{b2,n}, \ w_{b2,n} \sim N(0, \sigma_{b2}) \\ b_{3,n+1} &= b_{3,n} + w_{b3,n}, \ w_{b3,n} \sim N(0, \sigma_{b3}) \\ b_{4,n+1} &= b_{4,n} + w_{b4,n}, \ w_{b4,n} \sim N(0, \sigma_{b4}) \end{cases} \quad (2.1)
 \end{aligned}$$

式(2.1)中, x_n 为状态向量, $w_{b1} \sim w_{b4}$ 代表均值为 0 的噪声, $\sigma_{b1} \sim \sigma_{b2}$ 代表参数噪声的方差。相对应的观测方程为:

$$C_{\max,n} = b_{1,n} e^{b_{2,n} n} + b_{3,n} e^{b_{4,n} n} + v_n, v_n \sim N(0, \sigma_v) \quad (2.2)$$

式(2.2)中, v_n 代表均值为 0 的观测噪声, σ_v 代表观测噪声的方差。

在建立模型的状态空间之后, BMS 将以历史容量数据作为模型观测数据, 通过滤波算法对状态向量进行更新与校正, 从而获取预先设立的寿命经验模型的参数。完成参数更新校正后, 在模型中带入对应电池容量失效阈值, 从而可以估算出该容量失效阈值所对应的循环周期数, 从而完成对锂电池 RUL 的有效预测。滤波预测法对于寿命经验模型具有有效的改善效果, 提高寿命经验模型的 RUL 预测精度, 但同时预测性能也受到因素简单的限制, 因此该方式在 RUL 预测领域还有很大的发展空间。在滤波预测法中常用的滤波方法有扩展卡尔曼滤波、粒子滤波及球形容积粒子滤波等。

同样是数据驱动型的 RUL 预测方式, 时序预测法是从时间序列角度出发, 不需要建立寿命经验模型, 将锂电池 RUL 的预测问题转变成为时间序列预测问题, 这是与前两种方式最本质的区别。时间序列预测方法的数学表示为:

$$C_{\max,n+1} = f(C_{\max,n}, C_{\max,n-1}, \dots, C_{\max,n-m}), m < n \quad (2.3)$$

式(2.3)中, $f(\cdot)$ 为时间序列预测模型。

时间序列预测模型的常用功能是实现对未来序列数据的预测, 首先将研究对象设定为基于时间依赖性的连续变量, 使用历史变化序列来捕捉变化特征, 从而推测未来时序变化趋势。对于我们的研究对象锂电池, 可以将其循环中的充放电过程导致的容量衰减轨迹当作是一个相对稳定且长期的时间序列, 从而可以根据处理时间序列的方式来对容量衰减轨迹做出拟合, 进而对锂电池 RUL 进行预测研究。

基于式(2.3)及时间序列预测模型 $f(\cdot)$, 可以使用 $n-m$ 循环周期到 n 次循环周期之间的历史容量数据作为模型输入来得到第 $n+1$ 次循环周期时对应的容量值, 如此类推, 反复使用该模型与方式来依次得到未来时间序列的容量数据, 该算法过程最核心的技术在于构建合理的时间序列预测模型。基于此, 本课题主要是使用现在最为流行的神经网络的延伸即深度学习模型来建立时间序列预测模型, 以此来实现锂电池 RUL 预测。使用时间序列预测模型的锂电池 RUL 预测模型如图 2.3 所示。

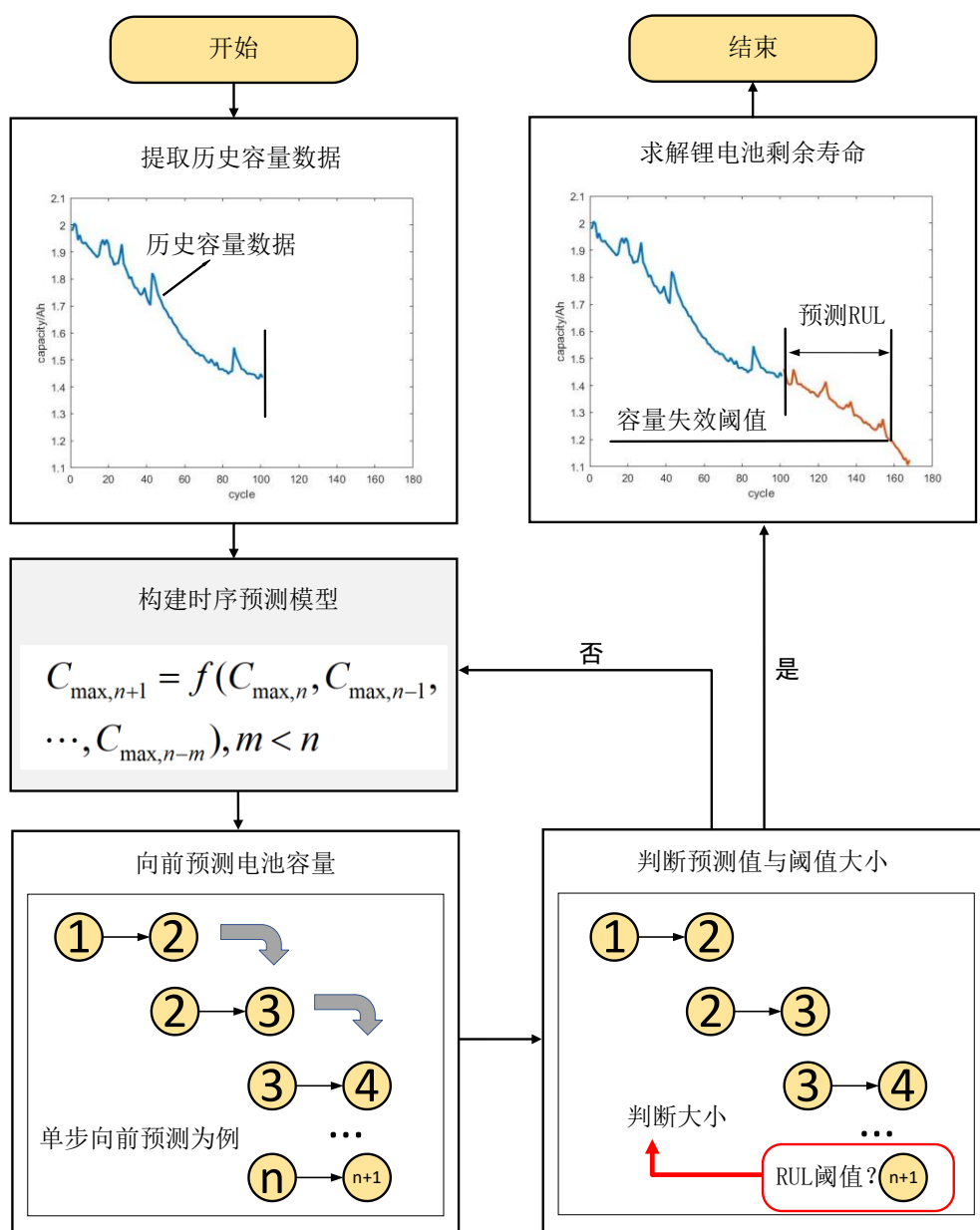


图 2.3 历史容量值为输入的时间序列预测流程

图 2.3 详尽展示了基于时序预测法进行锂电池 RUL 预测的流程，由图中所示步骤可知，本课题采用的时序预测法不需要依赖耗时漫长且复杂多变的电池老化离线测试实验，也无需建立寿命经验模型，只需要根据 BMS 中存储下来的锂电池历史容量数据进行未来时序预测便可获得电池 RUL 的预测结果。

2.4 锂电池剩余寿命的不确定性

锂电池 RUL 的预测机制，是基于历史数据的运用一定的数学算法与策略来对锂电池的残值寿命进行计算的技术。关于预测问题的研究，通常不可避免对于不确定性的分

析^[59]。一般来说，锂电池的 RUL 的不确定性研究比 RUL 预测结果更具有现实意义，而不确定性的表征方式一般为 RUL 的概率分布，工程应用中可以根据 RUL 概率分布来获得 RUL 的分布规律，来对锂电池的管理与检修提供理论支撑与决策建议

锂离子电池剩余使用寿命预测算法随着神经网络与机器学习算法的日益发展而获得强大进步，各种基于数据的算法使得锂电池 RUL 的预测准确度不断提高。同时也不难发现，现有的 RUL 计算方式往往只能给出剩余使用寿命的点估计值，仅仅只是一个输出结果，而无法给出 RUL 的可靠度以及置信区间。在锂电池 RUL 预测的实际工程应用中，由于各类设备的偏差导致的数据误差以及计算模型的能力强弱和噪声影响，导致得到的 RUL 预测结果值距离实际的结果仍然存在一定程度的偏差，说明模型给出的点估计值对于 RUL 的研究是不够深入的，对于 RUL 结果处理的决策建议是不够全面的，甚至可能会导致较为复杂的不良后果。因此对于锂电池的 RUL 预测，不仅需要得到点估计值，同时也需要得到 RUL 预测结果值的可靠性，由此可以避免预测过程中的一些偶然结果，可以为工程应用提供一些实质性的决策指导意见。

对于基于深度学习方法锂电池 RUL 算法，对预测结果不确定性的分类主要分为偶然不确定性和认知不确定性。其中偶然不确定性可分为同方差不确定性和异方差不确定性；认知不确定性可分为参数不确定性和结构不确定性，分类方式如下图 2.4 所示：

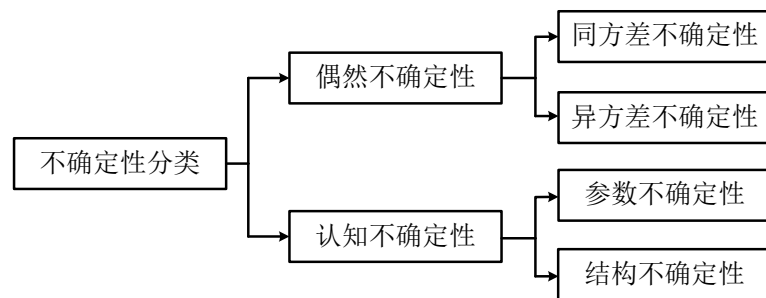


图 2.4 不确定性分类方式

对于基于深度学习的 RUL 预测不确定性分类中，偶然不确定性指的是整个观测数据中存在的固有的干扰噪声，其来源主要是整个观测过程，其中的同方差不确定性不依赖于输入数据，也不是模型输出，而是对于所有的输入数据都相同存在的一些常量，对不同的任务才有不同的变量，基于此同方差不确定也被叫做任务依赖性不确定性；而异方差不确定性主要是表征输入的数据带来的不确定性，这种不确定性也是较为普遍的。认知不确定性主要指的是由 RUL 预测模型带来的不确定性，这种不确定性在深度学习是可以增加训练数据的量来进行一定程度的改善，认知不确定性中的参数不确定

性主要表征深度学习网络中的神经元和各类参数带来的不确定性，而结构不确定性主要表征的是在深度学习方式中的选择上带来的不确定性，依赖于模型整体架构的创建。

对于锂电池不确定性量化环节，对于数据驱动型的三种主流预测方式的不确定性产生环节如表 2.2 所示：

表 2.2 数据驱动型预测方式主要不确定性产生环节

预测方式	主要不确定性产生环节
经验预测法	基于寿命经验模型对于历史容量数据的拟合过程
滤波预测法	基于状态空间方程构造的状态向量初始化过程
时序预测法	锂电池 RUL 预测过程的时序预测模型输入环节

Smith 等人^[60]对于深度学习算法中的不确定度进行了详尽的介绍，解释了互信息是有效检测对抗性样本的。对于不确定性的度量方式主要包含有标准差、预测方差、方差比、预测熵和互信息等，而在我们一般性的研究中，为了使得结果度量变得简洁，通常采用预测方差来对不确定性进行定量分析。在本课题中，对于锂电池剩余寿命的不确定性分析，采用标准差与方差等作为度量指标，通过对模型预测输出结果的后验分布进行统计量求取来实现不确定性评估。

2.5 本章小结

本章从锂电池结构与原理出发，介绍了锂电池的工作原理，从不同角度分析了锂电池的老化原理与退化机制；介绍了锂电池管理系统以及其现实价值，从电池管理系统的组成部分与工作原理角度出发，分析了锂电池剩余使用寿命作为 BMS 主要内容之一，对于电池管系统的重要意义；从数据驱动型 RUL 预测方式出发，介绍了数据驱动型预测方式的实现流程，重点介绍了基于时序数据的预测流程；介绍了锂电池剩余寿命不确定性的概念与分类，分析其不确定性的研究意义与度量方式，为下一步的工作奠定了基础。

第三章 基于 CNN 与 LSTM 的锂电池剩余使用寿命预测

本课题使用锂电池的容量变化来表征其老化过程，推导预测锂电池的容量变化来实现对 RUL 的预测。基于深度学习在锂电池研究领域的广泛应用，本课题使用数据驱动型的深度学习方法来建立锂电池的 RUL 预测模型。本章将首先介绍神经网络及其延伸深度学习的原理，介绍其计算过程与训练方式，为之后的预测模型奠定理论基础；对本课题使用的锂电池容量退化数据集进行介绍，便于统一，对本课题的实验环境与实验流程进行设定介绍，并且对实验结果的评价指标进行介绍；使用卷积神经网络 CNN 与长短期记忆网络 LSTM 来分别构建锂电池容量序列预测模型，通过介绍两种模型各自的原理与训练过程，建立预测模型，使用预测挖掘锂电池容量退化的内在特征，得到锂电池的容量预测终止周期，从而实现 RUL 的预测，通过实验结果来探究两种网络对锂电池 RUL 预测的研究价值。

3.1 神经网络与深度学习

神经网络是一种模仿动物大脑神经来处理数据系统的特殊行为特征，主要是以人造神经元为基础，其内部各神经元之间进行相互之间的连接，以此进行信号的传输过程。随着日益增长的大数据处理需求，神经网络的发展也越来越壮大。由早期的机器学习发展壮大为如今的神经网络，神经网络与机器学习息息相关，神经网络代替了数据处理中的模型，使用学习规则代替了机器学习。

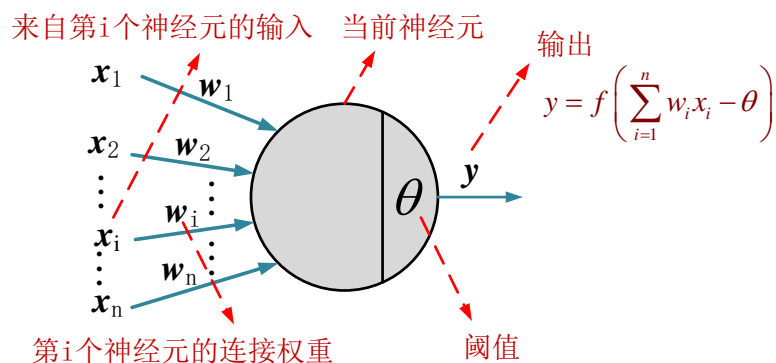


图 3.1 M-P 神经元

神经网络中最基本的单元是神经元，也称为节点或单元。在生物神经网络之中，每个神经元之间进行相互的连接，通过“兴奋—超过阈值—激活”的方式进行信息传输，

在人工神经网络中，此结构的神经元模型被定义为“M-P 神经元”神经元，“M-P”神经元的工作机制如图 3.1 所示。

将诸多上述的“M-P”神经元模型进行一定规则的连接后，得到的结构模型便为神经网络模型。神经网络常用的激活函数常为阶跃函数，阶跃函数可将输入信号转换为 0 和 1 的输出值，分别对应为神经元的抑制与兴奋，由于阶跃函数的不连续性，sigmoid 函数的使用场景更为广泛，sigmoid 可以为结果限定 0 至 1 的范围，便于进行更广泛的研究应用。阶跃函数与 sigmoid 函数如下图 3.2 所示：

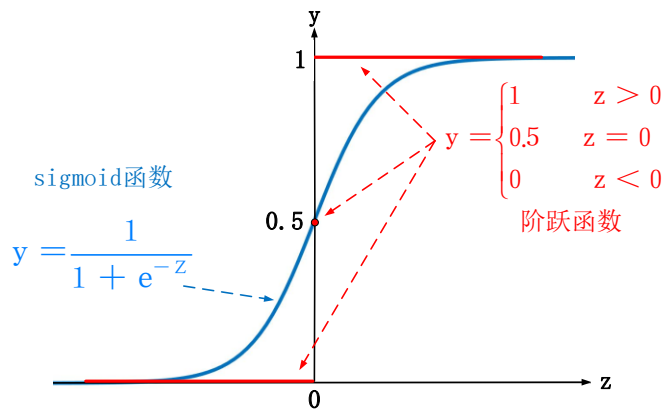


图 3.2 阶跃函数与 sigmoid 函数

神经网络是由众多神经元的节点所构成的网络，而节点之间不同的连接方式则构成了神经网络的巨大差异性，由此可根据节点连接方式来建立多种神经网络。其中最常用的则是分层节点结构，如下图 3.3 所示。

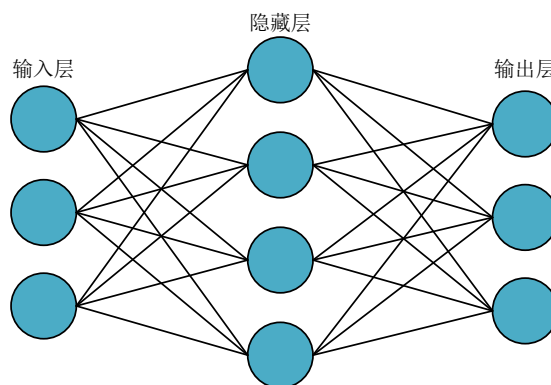


图 3.3 神经网络分层节点结构

神经网络的分层模式主要将神经网络的结构划分为三个层次，分别为输入层、隐含层和输出层。其中输入层仅被用作传递输入信号的通道，不计算信号的权重与激活函数；输出层主要用左神经网络计算结果的输出通道，该层的节点输出便是整个神经网络模型

的最终结果；在输入层与输出层之间的层，被叫做隐含层，通常无法通过外部对隐含层进行直接性的访问。根据其中隐含层的数量和复杂程度将神经网络区分为浅度神经网络和深度神经网络，包含两个及以上的隐含层就是深度神经网络。

神经网络中的感知机是由两层结构的神经元构成，可以实现逻辑运算的结构体，其输入层用于接受外部输入，输出层即为 M-P 神经元。假设给定数据集，神经元权重 w_i 与阈值 θ 是可以通过神经网络进行学习来修正的，对于给定的样例 (x, y) ，设定感知机输出为 \hat{y} ，则有感知机的调整过程如下所示：

$$w_i \leftarrow w_i + \Delta w_i \quad (3.1)$$

$$\Delta w_i = \eta (y - \hat{y}) x_i \quad (3.2)$$

其中 η 为学习率，其范围为 $(0, 1)$ 。从式(3.1)和(3.2)中可以看出，假若感知机的训练是正确的，得到的结果与真实值是相同的，即 $y = \hat{y}$ ，那么感知机不会再进行学习，倘若结果有偏差，将依据错误程度来对权重进行修正，感知机的修正过程是神经网络进行学习的基本所在。

对于多层网络的训练，基于感知器修正原理，发展出了更强大的误差逆传播（BackPropagation, BP）算法，由 BP 算法构成的 BP 网络具有良好的自学习性、自适应性、鲁棒性和泛化能力^[61]。其核心 BP 算法被广泛应用于各类神经网络之中，BP 算法过程如算法 3.1 所示，BP 算法原理与符号如下图 3.4 所示：

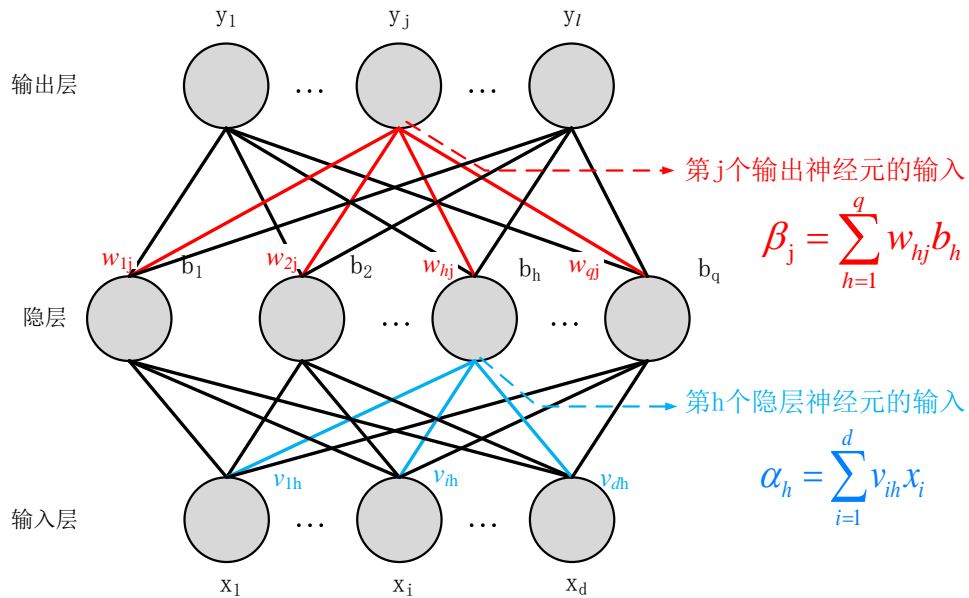


图 3.4 BP 算法过程与符号示意

BP 算法的主要目标是使得训练集 D 上的累计误差最小化，累计误差如下所示：

$$E = \frac{1}{m} \sum_{k=1}^m E_k \quad (3.3)$$

$$E_k = \frac{1}{2} \sum_{j=1}^l (\hat{y}_j^k - y_j^k)^2 \quad (3.4)$$

神经网络的训练过程也是一种参数的寻优过程，即在所拥有的参数空间中，寻找出一组最优的参数组合，该参数组合将会使得误差函数 E 为最小。基于梯度的搜索是神经网络中最常用的参数寻优方式，例如随机梯度下降法。

算法 3.1 BP 算法

输入：训练集 $D = \{(x_k, y_k)\}_{k=1}^m$ ；

学习率 η 。

过程：

1: 在 $(0,1)$ 范围内随机初始化网络中所有的连接权与阈值

2: repeat

3: for all $(x_k, y_k) \in D$ do

4: 根据当前参数与公式 $\hat{y}_j^k = f(\beta_j - \theta_j)$ 计算当前样本的输出 \hat{y}_j^k ；根据

$$g_j = -\frac{\partial E_k}{\partial \hat{y}_j^k} \cdot \frac{\partial \hat{y}_j^k}{\partial \beta_j} = \hat{y}_j^k (1 - \hat{y}_j^k) (y_j^k - \hat{y}_j^k) \text{ 计算输出层神经元的梯度项 } g_j；$$

5: 根据 $g_j = -\frac{\partial E_k}{\partial \hat{y}_j^k} \cdot \frac{\partial \hat{y}_j^k}{\partial \beta_j} = \hat{y}_j^k (1 - \hat{y}_j^k) (y_j^k - \hat{y}_j^k)$ 计算输出层神经元的梯度项 g_j ；

6: 根据 $e_h = -\frac{\partial E_k}{\partial b_h} \cdot \frac{\partial b_h}{\partial \alpha_h} = b_h (1 - b_h) \sum_{j=1}^l w_{hj} g_j$ 计算隐层神经元的梯度项 e_h ；

7: 根据 $\Delta w_{hj} = \eta g_j b_h, \Delta \theta_j = -\eta g_j, \Delta v_{ih} = \eta e_h x_i, \Delta \gamma_h = -\eta e_h$ 更新连接权 w_{hj}, v_{ih} 以及阈值 θ_j, γ_h

8: end for

9: until 达到停止条件

输出：连接权与阈值确定的多层前馈神经网络

从理论研究层面，一个模型的参数越多，随之带来的是模型的复杂度与“容量”的增长，同时意味着模型处理问题的能力也越强，但同时也会带来低效率和过拟合之类的问题。随着技术的发展与需求的日益增长，为了提高模型计算能力以及达到高效率的建模，以深度学习为典型代表的复杂模型越来越受到关注。

深度学习即是更深层次的神经网络，是对神经网络的延伸，是一种主要运用深度神经网络的机器学习技术。一般而言，深度学习技术主要是对神经网络模型中增加隐含层的数目，相应增加的是神经元与参数的数目，模型的复杂度也随着这些神经网络主要元

素的增加而得到大幅度提高。并且从复杂度而言，增加隐藏层的数目要比在某一层单纯增加神经元的数目要高效率，这是因为增加隐藏层数不仅可以增加了激活函数的神经元数量，并且还使得激活函数嵌套层数增加，深度学习基于神经网络增加隐藏层结构的示意如图 3.5 所示。

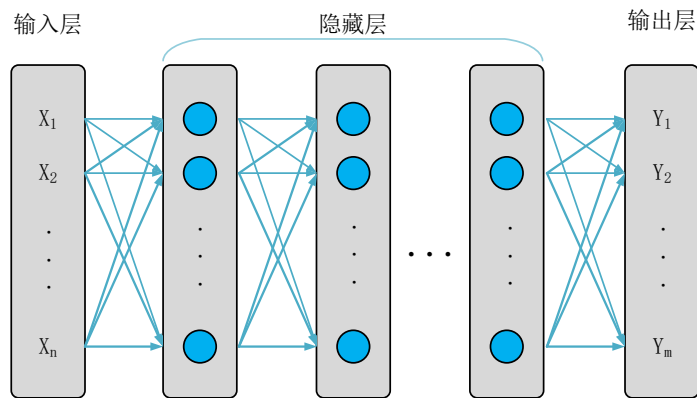


图 3.5 深度学习示意图

在深度学习模型中，往往容易出现误差逆传播时“发散”而无法收敛至稳定态的问题。无监督逐层训练往往是处理多隐层网络结构的有效方式，其主要思想是每次训练一层隐节点，训练过程中，将上一隐层的输出作为本隐层的输入，本隐层的输出作为下一隐层的输入，这样的过程被称为预训练，待每一层的预训练结束，再对全局网络进行微调，微调可使用全局网络常用的误差逆传播算法来实现。这种“预训练+微调”的训练方式，可看作是对大量参数进行了分组处理，先对每组进行局部的优化，然后再基于每组的优化结果进行全局的参数寻优，这样可以达到有效节省训练开销的目的，提高优化效率。另一种有效节省训练资源的主流方式是“权共享”，即让同一层的神经元共享使用相同的权值参数，使用这种方式的典型网络为卷积网络，其具体原理将在后文介绍。

总体而言，深度学习可以视作是对输入信号进行堆叠加工后，将初始的、与目标结果联系不太紧密的表示转化为了输入信息与输出表示联系更加密切的强化表示，使得神经网络基于简单的输出映射而无法完成的任务变成了可实现的有效过程，通过多层处理，使得简单的低维特征表示转化为了高维的多特征表示。深度学习的杰出成就不在于某项决定性的技术，而是在于对深度神经网络诸多细微之处的改进。

3.2 数据集与相关实验设定

本课题进行锂离子电池 RUL 预测研究所采用的开源数据集为美国国家航天航空局

(National Aeronautics and Space Administration, NASA)Ames 研究中心公布的锂离子电池性能退化数据集与马里兰大学先进寿命周期工程中心(Center for Advanced Life Cycle Engineering, CALCE)公开的锂离子电池数据集。在本课题的研究中,将电池额定容量的70%设定为电池寿命终止点(End Of Life, EOL)。

3.2.1 数据集介绍

NASA 电池退化数据集以 18650 型号的锂离子电池为研究对象。本数据集引用的电池有三块,分别为 B05、B06 和 B07。室温条件为 24° , 分别对电池进行充电、放电与阻抗测量的操作,以恒流 1.5A 的电流模式,将电池充电至 4.2V,再以恒压电压状态将充电电流调至 20mA;使用 2A 的恒定电流进行放电过程,恒流状态下使电压降至 2.5V。最后使用 0.1Hz 到 5kHz 的电化学阻抗谱来对电池进行扫描,得到电池阻抗。三组电池的额定容量均为 2Ah,数据集包含有 168 个循环周期数据。EOL 设定为额定容量 2Ah 的 70%即 1.4Ah,电池容量退化曲线如图 3.6 所示。

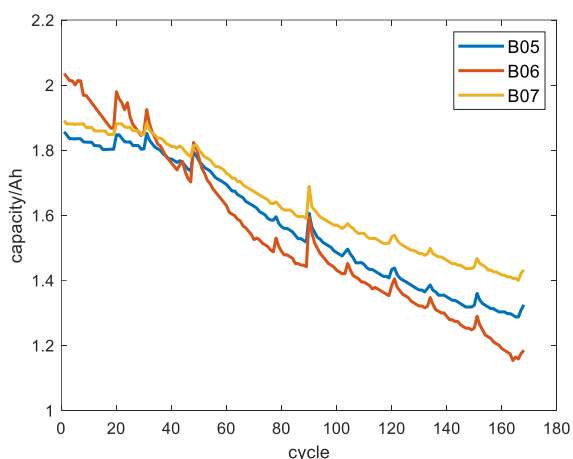


图 3.6 NASA 数据集电池容量退化曲线

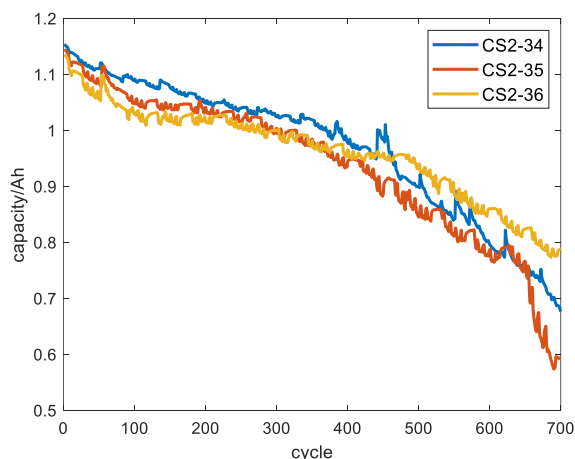


图 3.7 CALCE 数据集电池容量退化曲线

CALCE 数据集集中的三组电池为 CS 型号的 CS2-34、CS2-36、CS2-37 锂离子电池,数据集中包含电池充放电循环次数和每次循环结束时对应的容量记录,包含 700 个循环的记录数据。该数据集是使用 ArbinBT2000 的锂电池测试系统得到,电池额定容量为 1.1Ah,在室温条件下进行充放电。实验具体过程为:首先以 0.45A 的恒定电流对电池进行充电,使电压增至 4.2V,然后保持恒压 4.2V 状态,将充电电流在充电过程逐渐降低至 0.05A,之后以 0.45A 的恒定电流放电至电压降至 2.7V,实验过程中的放电速率保持为恒定 1C。EOL 设定为额定容量 1.1Ah 的 70%即 0.77Ah,电池容量退化曲线如图 3.7 所示。

为了保证本课题使用的深度学习模型在训练过程中提高效率，提高模型基于损失下降的收敛速度，使得参数寻优过程更加高效，对本课题实验数据进行线性归一化处理，过程如下所示：

$$C = \frac{C - C_{\min}}{C_{\max} - C_{\min}} \quad (3.5)$$

3.2.2 实验设定与评价指标

本课题实验数据采用 NASA 实验数据中的 B05 号电池和 B07 号电池退化数据集及马里兰 CALCE 电池容量数据中的 CS-34 号电池与 CS-36 号电池容量退化数据集。仿真实验在 Windows10 系统 8GB 的个人计算机上进行，实验环境为 Matlab2021b 版本及 Matlab 深度学习工具箱。

为本课题各模型间进行有效实验对比，本课题所涉及实验均实现前向单步的 RUL 预测，输入步长为 5，即 5 个历史连续容量数据为模型输入，前向一步的 RUL 预测值为模型输出。本课题对于锂电池的 RUL 预测原理如图 3.8 所示。

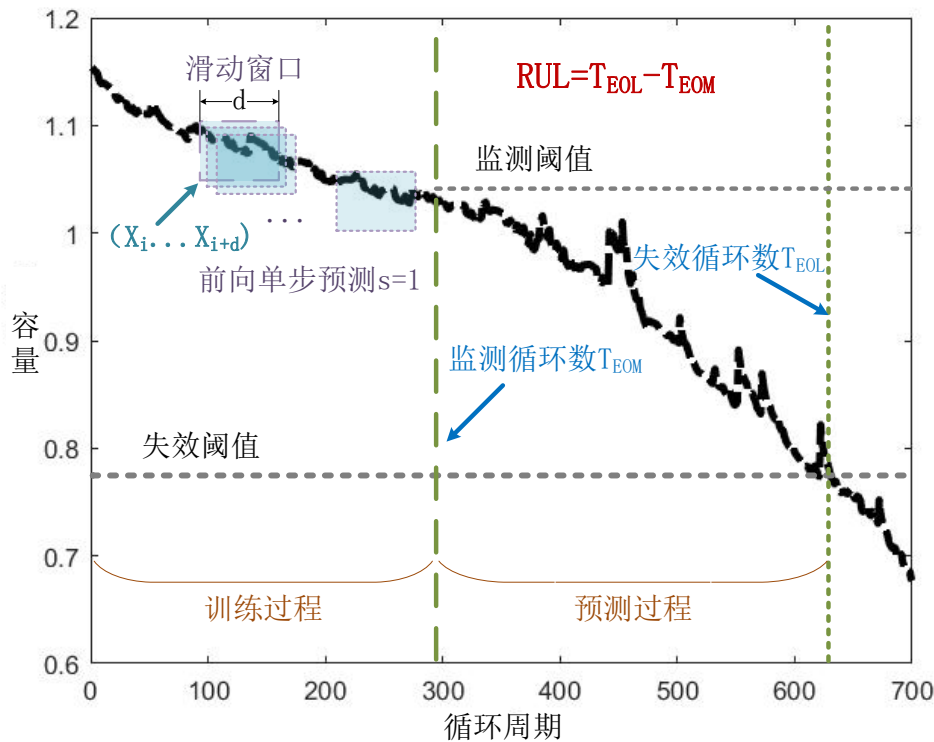


图 3.8 锂电池 RUL 预测中滑动窗口与单步前向预测原理

当电池容量退化到一定程度时开始预测 RUL 的电池循环周期数，即 RUL 预测起点

为 T_{EOM} ，对应容量为监测阈值，到达失效阈值时的循环周期数设为 T_{EOL} 。此时不难看出，RUL 即为失效阈值对应周期与检测周期对应周期，即：

$$RUL = T_{EOL} - T_{EOM} \quad (3.6)$$

本课题的 RUL 预测中基于容量的评估标准采用均方根误差(Root Mean Squared Error,RMSE)、平均绝对百分比误差(Mean Absolute Percent Error, MAPE)及剩余寿命误差 E_{RUL} 三种评价指标，三种评价指标均体现了模型的精度，可以进行客观的模型评价。

均方根误差（Root Mean Squared Error,RMSE）如下式(3.7)所示，该误差在数量级上比较直观，比如 $RMSE=5$ ，可以认为回归效果相比真实值平均相差 5。

$$RMSE = \sqrt{\frac{\sum_{i=1}^N (C_i - \hat{C}_i)^2}{N}} \quad (3.7)$$

平均绝对百分比误差（Mean Absolute Percent Error, MAPE）如下式(3.8)所示，该误差是用于评估预测性能的最受欢迎的指标之一，其值越小，说明模型精度越高。

$$MAPE = \frac{\sum_{i=1}^N \left| \frac{C_i - \hat{C}_i}{C_i} \right|}{N} \quad (3.8)$$

剩余寿命误差为预测值到达寿命终止循环次数与电池真实值到达终止寿命时循环次数之差的绝对值，如式(3.9)所示：

$$E_{RUL} = |RUL_{real} - RUL_{predict}| \quad (3.9)$$

基于 CALCE 数据集中各电池之间的寿命终止周期相差较大，在实验中分别选用前 40%、50%及 60%周期数作为模型训练数据，来进行后续未来容量的预测。而对于 NASA 数据集，各电池间容量退化趋势较为相近，选取前 70、80 及 90 个周期作为训练数据，同时即为未来容量预测起点。

3.3 基于 CNN 的锂电池剩余寿命预测

卷积神经网络(Convolutional Neural Network,CNN)是深度学习领域最常用的神经网络模型之一，在处理计算机视觉相关任务、自然语言分析处理、医学图像处理研究以及时间序列问题方面均有广泛的应用^[62]。CNN 通常被用于处理二维图像数据，可以有效提取数据空间特征，取得良好的处理效果，因此本课题使用 CNN 网络来对锂电池 RUL

进行预测。在锂电池 RUL 预测研究中,输入的容量历史数据为一维数据,由此本课题采用一维 CNN 网络来进行锂电池的 RUL 预测。

3.3.1 卷积神经网络 CNN 原理

卷积神经网络主要使用前向传播的方式来计算神经网络的输出值,而对神经网络节点之中的权重和偏置则主要使用反向传播和链式求导的法则来进行训练时期的更新,对于层与层之间的连接并不是同 BP 网络(Back Propagation,BP)一样全部连接,而是只进行部分连接。

CNN 的核心思想主要在于以下几方面:

(1) 权重的共享性。对于 CNN 中的每一个的神经元而言,其与上一层神经元进行连接时,同层的每个神经元的权重参数都具有有一致性,即卷积核是一样的,而每一个卷积核只能用于提取一种特征,因此对于多特征的滤波提取,需要拥有多个卷积核。相对 BP 网络而言, CNN 通过减少参数的方式,大大降低网络的复杂程度。

(2) 局部区域感知。在 CNN 网络之中,卷积层中的每个神经元不需要与前面的每个输入都进行连接,而是与输入数据的局部进行连接,随着卷积网络层数的逐递增加,最后 CNN 再将每一部分汇聚而成总体特征,得到输入数据的总体信息特征。

(3) 时间或者空间中的降采样。由于 CNN 在学习过程中会得到巨高维的特征,容易出现过拟合的现象从而导致输出结果不理想。根据每次卷积得到的特征位置的相对不变性,对相对位置进行细微变换并不会很影响特征,于是可以对卷积后的特征区域进行降采样,这样便可以使计算量得到大幅度的减少,这种降采样的方式即池化(pooling)。

CNN 的结构主要由卷积层、池化层和全连接层构成, CNN 的结构如图 3.9 所示:

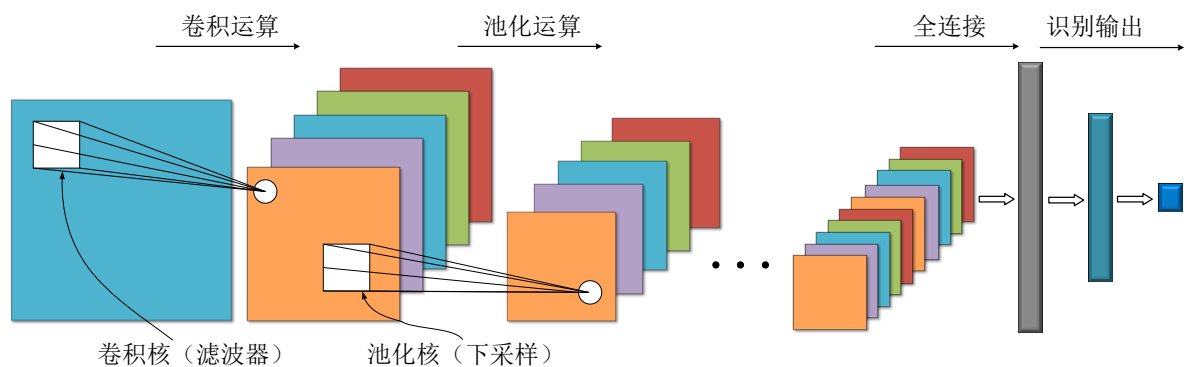


图 3.9 卷积神经网络结构原理

卷积层是 CNN 算法中的核心之一,其主要包含一系列的卷积核,也就是滤波器。

其主要工作机制可以概括为通过对输入数据进行卷积运算来得到数据的特征向量，再通过非线性激活函数对上一步卷积的局部区域添加权重与偏置的参数来生成特征映射。卷积运算原理可描述为：

$$C_j^n = \sigma(w_j^n * C_i^{n-1} + b_j^n) \quad (3.10)$$

式(3.10)中， C_j^n 为第 n 卷积层的第 j 个特征映射； w_j^n 为第 n 卷积层中的第 j 权重矩阵； b_j^n 为第 n 卷积层中的第 j 个偏置向量； $\sigma(\cdot)$ 为当前非线性激活函数； $*$ 为卷积运算； C_i^{n-1} 为第 $n-1$ 个卷积层的第 i 个特征输出。卷积操作原理如图 3.10 所示。

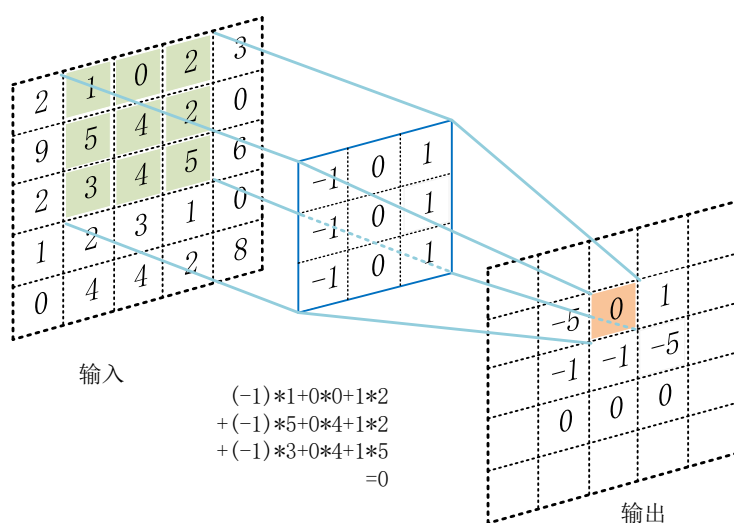


图 3.10 卷积操作原理

池化层的主要功能是对卷积层输出的特征映射先进行采样处理，通过这种方式来对上一层得到的特征进行降维，以此降低整体运算量，同时可避免过拟合现象。常见的池化层的池化操作过程有最大值池化和平均值池化。最大池化操作原理如图 3.11 所示。

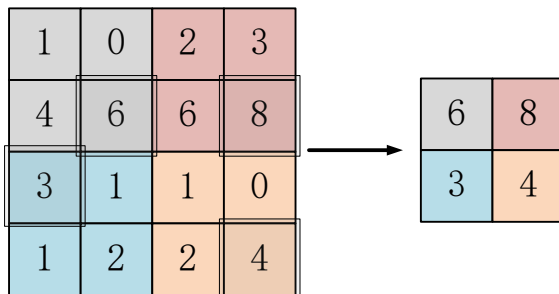


图 3.11 最大池化操作

最大值池化过程如式(3.11)所示：

$$P_j^n = \max(C_j^{n-1}, s) \quad (3.11)$$

式(3.11)中, C_j^{n-1} 代表从卷积层得到的特征向量, P_j^n 代表池化层进行池化处理后的输出, $\max(\cdot)$ 代表最大值的采样操作函数。

全连接层是将池化层所得到的所有结果特征向量进行平展, 然后作为全连接层的输入, 进行一些非线性化的变换, 主要使用激活函数作用, 进而将结果输出, 作为跟在其后面各个层的输入。其原理与传统的多层感知器原理类似, 如图 3.12 所示:

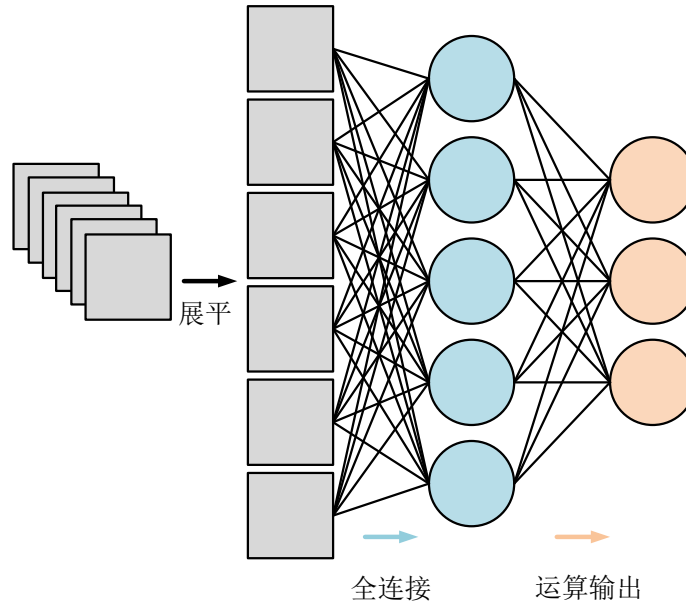


图 3.12 全连接层原理

全连接运算得到的每一个输出神经元与每一个输入神经元的运算都有密切联系, 因此可以充分提取每一个输入数据之间的特征关系, 全连接层的运算过程如式(3.12)所示:

$$V^n = \sigma(w^n P^n + b^n) \quad (3.12)$$

式(3.12)中, w^n 代表全连接层的权重参数矩阵, P^n 代表池化层的输出, b^n 代表全连接层的偏置向量, V^n 代表第 n 个全连接层的输出。

3.3.2 CNN 模型的运行学习过程

通过上述对 CNN 网络结构的介绍可知, CNN 的运行过程主要是基于前向运算的, 即网络的输出是基于传入的特征矩阵进行卷积和池化等操作而得。对于 CNN 网络中各个层的权值和偏置等参数, 则需要在网络训练过程中进行相应规则的学习, CNN 网络参数的学习与 BP 算法是类似的, 主要不同之处在于 CNN 具有权值共享和局部连接的特性, 基于此, CNN 网络的参数学习需要进行分层分析。

对于卷积层的参数学习主要针对于权重矩阵 w 和偏置 b , 卷积层获得的特征结果是

基于输入特征与卷积核的相应运算得到的，其节点结果如(3.13)所示：

$$a_j^k = f(\sum_{i \in M_j} a_i^{k-1} * cv_{i,j}^k + b_j^k) \quad (3.13)$$

式(3.13)中， k 代表网络的层数， j 代表个数， M_j 代表输入其中的 j 个特征的集合， $cv_{i,j}^k$ 代表卷积层中的卷积核， a_j^k 代表第 k 层的第 j 个输出值， b_j^k 代表对应的偏置量。与前文提到的感知器的正向传播类似，可得卷积层内部特征值与输出值之间的层间递推运算过程如下所示：

$$\begin{cases} E = \frac{1}{2} \times \sum_i (a_i - y_i)^2 \\ z^k = w^k \times a^{k-1} + b^k \\ a^k = f(z^k) \end{cases} \quad (3.14)$$

式(3.14)中， E 代表传播误差， y_i 代表类别向量的第 i 维的分量， a_i 代表输出层的第 i 维的输出值。在此同样类比感知器的反向传播，可得权值 w 和偏置 b 的梯度，其计算如下所示，其中的中间变量 δ 如式(3.15)和(3.16)所示，从中可知，第 k 层的变量 δ 是根据下一层的 δ 值、下一层权值以及激活函数进行相乘而得的，基于此，求出 δ^k 与 δ^{k+1} 是优化权值 w^k 的前提。

$$\begin{cases} \frac{\partial E}{\partial b} = \frac{\partial E}{\partial z} \times \frac{\partial z}{\partial b} = \delta, \text{ 令 } \delta = \frac{\partial E}{\partial z} \\ \frac{\partial E}{\partial w^k} = \frac{\partial E}{\partial z} \times \frac{\partial z}{\partial w^k} = a^{k-1} \times (\delta^k)^T \end{cases} \quad (3.15)$$

$$\delta^k = (w^{k+1}) \times \delta^{k+1} \times f' \quad (3.16)$$

CNN 结构中是在一定的区域之中进行采样，这个区域叫做采样区域，该采样区域对于下一采样层的特征单元值有决定作用，基于此， δ 的递推公式可进一步表示为式(3.17)，其中的操作符 \uparrow 代表上采样。

$$\delta_j^k = w_j^{k+1} \times (f' \times \uparrow(\delta_j^{k+1})) \quad (3.17)$$

基于以上推导过程，卷积层的权值矩阵 w 与偏置矩阵 b 的梯度表示如下式(3.18)所示，其中的 $(r_i^{k-l})_{u,v}$ 表示在 a_i^{k-1} 中的特征图的局部区域， u 和 v 代表其中的特征图的水平方向和竖直方向。

$$\begin{cases} \frac{\partial E}{\partial b_j} = \sum_{u,v} (\delta_j^k)_{uv} \\ \frac{\partial E}{\partial w_{i,j}^k} = \sum_{u,v} (\delta_j^k)_{uv} (r_i^{k-1})_{uv} \end{cases} \quad (3.18)$$

对于采样层，其中需要学习的参数值是线性偏置 b 与乘性偏置 β ，其后面的网络层为全连接层时，其参数学习过程与 BP 算法相同；其后面连接的是卷积层时，参数 β 与 b 的梯度计算过程如下式(3.19)所示，其中的 \downarrow 操作符代表下采样， (u, v) 代表下采样层内的特征图组中第 j 个特征图中的数据。

$$\begin{cases} \frac{\partial E}{\partial b_j} = \sum_{u,v} (\delta_j^k) \\ \frac{\partial E}{\partial \beta_j} = \sum_{u,v} (\delta_j^k \times \downarrow (a_j^{k-1}))_{uv} \end{cases} \quad (3.19)$$

3.3.3 CNN 网络模型构建

基于所处理数据之间的差异，CNN 可以分为一维卷积、二维卷积和三维卷积，在处理时间序列问题及自然语言的问题领域中一般使用一维卷积。在一维的卷积网络中，可以理解是提取数据在一个方面上平移的特征向量，因此在本课题的基于时间序列的容量数据中，是主要提取序列的平移空间特征。一维卷积平移工作过程如下图 3.13 所示，本节选取一维卷积网络来进行电池容量的预测，达到剩余使用寿命的预测目的。

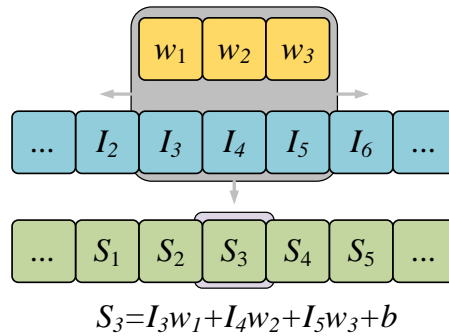


图 3.13 一维卷积平移操作

不论基于何种卷积网络，其模型原理与结构都是相同的。根据 CNN 算法原理以及本课题输入数据特点，提出本课题基于 CNN 模型的锂离子电池寿命预测方法，建立模型，模型输入为 5 个步长的连续历史容量值，输出为单步前向容量预测值，建立两个卷积层，一个全连接层，使用 Relu 激活函数，在模型的训练过程中，使用均方根误差损失函数来获得 LSTM 预测容量值与真实容量值之间的误差，并且使用自适应矩估计优化

器 Adam 来实现反向传播计算参数梯度并实时更新模型的参数值，初始学习率设定为 0.01，学习率下降因子设定为 0.5。具体模型参结构与参数如表 3.2 所示。

表 3.2 CNN 网络结构与参数

名称	类型	结构	激活数	可学习参数
imageinput	输入	归一化	$5 \times 1 \times 1$	-
conv_1	卷积	10, $3 \times 1 \times 1$, 步幅[1 1], 填充[0 0 0 0]	$3 \times 1 \times 10$	Weight: $3 \times 1 \times 1 \times 10$ Bias: $1 \times 1 \times 10$ Offset: $1 \times 1 \times 10$ Scale: $1 \times 1 \times 10$
batchnorm_1	批量归一化	10 个通道	$3 \times 1 \times 10$	
relu_1	ReLU	ReLU	$3 \times 1 \times 10$	-
conv_2	卷积	20, $3 \times 1 \times 10$, 步幅[1 1], 填充[0 0 0 0]	$1 \times 1 \times 20$	Weight: $3 \times 1 \times 10 \times 20$ Bias: $1 \times 1 \times 10$ Offset: $1 \times 1 \times 20$ Scale: $1 \times 1 \times 20$
batchnorm_2	批量归一化	20 个通道	$1 \times 1 \times 20$	
relu_2	ReLU	ReLU	$1 \times 1 \times 20$	-
fc	全连接	全连接层	$1 \times 1 \times 1$	Weight: 1×32 Bias: 1×1
regressionoutput	回归输出	mean-squared-error	-	-

3.3.4 实验结果与分析

对于 NASA 数据集的 B05 号电池与 B07 号电池使用 70、80、90 周期作为输入，实验结果如图 3.14 所示，对于 CALCE 数据集的 CS2-34 号电池与 CS2-36 号电池分别使用 40%、50%、60%周期作为输入，实验结果如图 3.15 所示，实验结果误差如表 3.2 所示。

表 3.2 容量预测实验结果误差

数据集	电池型号	周期数	RMSE	MAPE(%)
NASA	B05	70	0.180	10.28
		80	0.065	4.149
		90	0.065	4.308
	B07	70	0.054	3.221
		80	0.041	2.417
		90	0.065	4.308
CALCE	CS2-34	40%	0.076	6.834
		50%	0.074	8.375
		60%	0.054	4.827
	CS2-36	40%	0.084	7.675
		50%	0.047	5.382
		60%	0.040	4.338

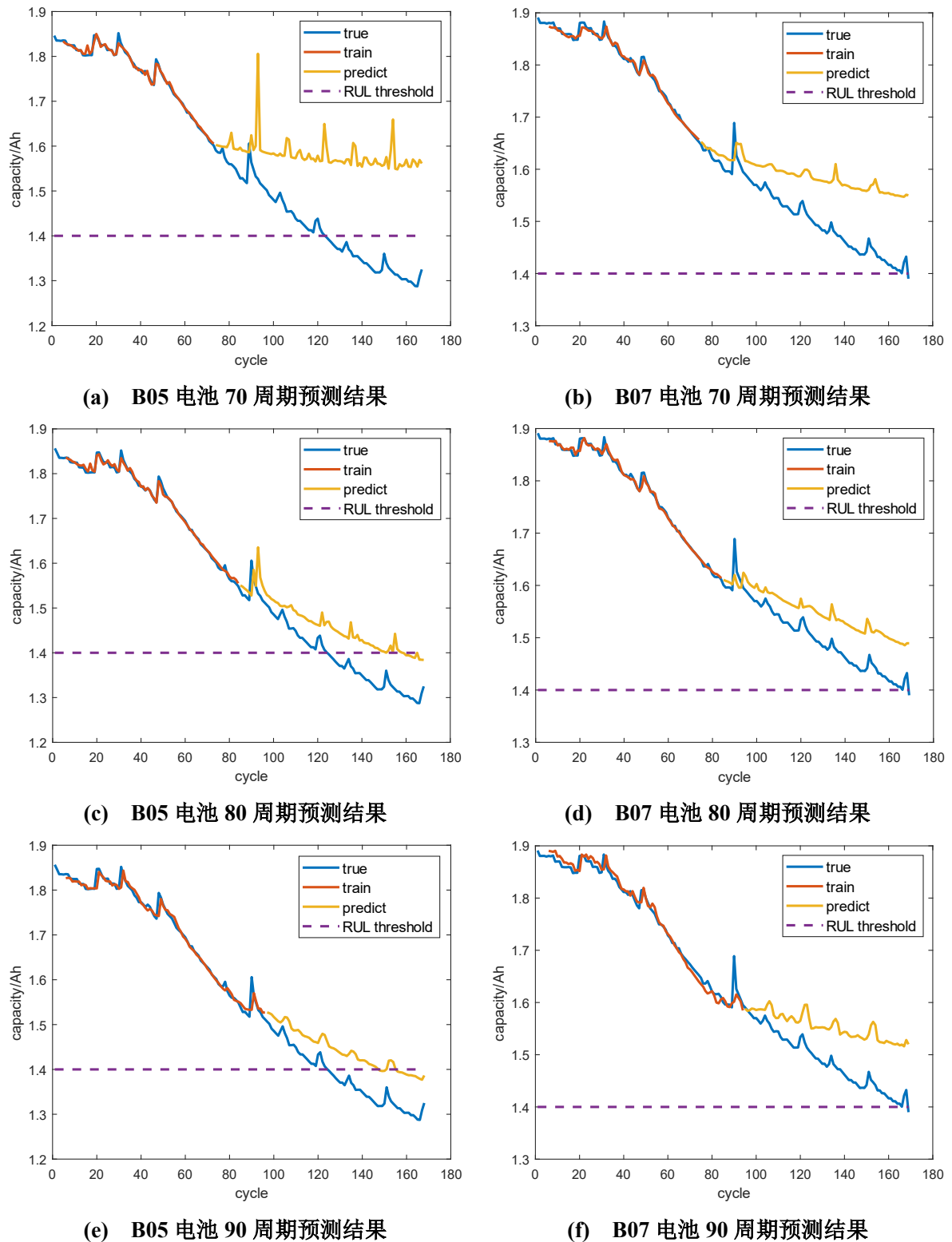


图 3.14 基于 CNN 模型的 NASA 数据集预测结果

表 3.2 展示了四个电池的预测误差,结果误差表明,对于大部分电池容量预测过程,预测结果的精度与锂电池预测周期呈正比关系,这种变化关系对于 CNN 模型而言是较为明显的,表明在锂电池的 RUL 预测中,越多的历史数据可以得到更好的预测结果。对于个别电池不同周期预测过程,例如 B07 号电池从 80 周期至 90 周期的变化中,预测

精度呈下降趋势，这是主要受到 90 周期左右自充电导致的容量波动变化较大的影响。

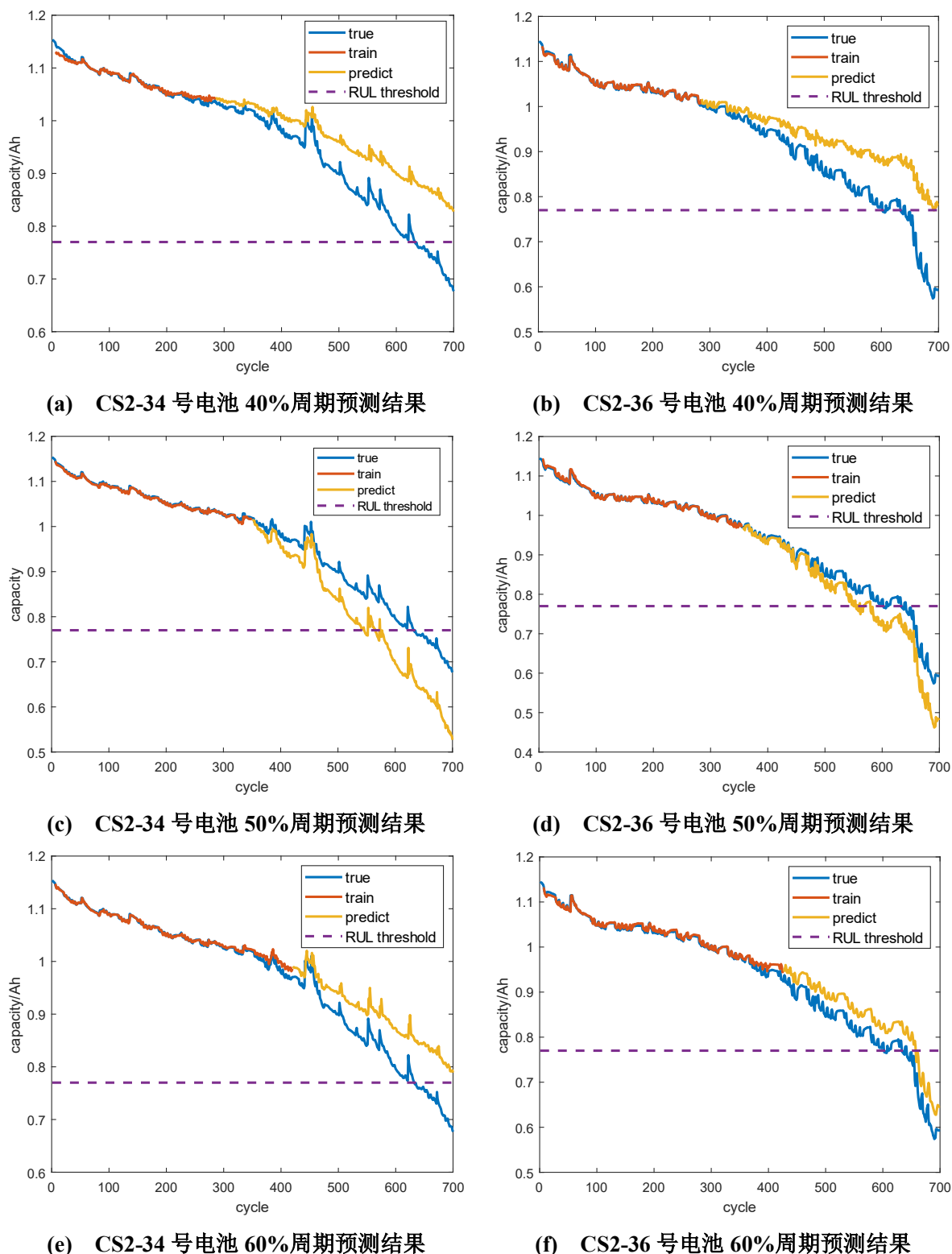


图 3.15 基于 CNN 模型的 CALCE 数据集预测结果

图 3.14 展示了基于 NASA 数据集的 B05 号电池与 B07 号电池不同预测周期的容量预测结果，可以看出 CNN 对两电池的未来序列容量可以做出一定的预测，在 B05 号电池 70 周期的预测效果最不理想，随着预测周期的增加，预测精度逐渐提高。在预测起

始周期较大情况下，受容量波动影响，B05 号电池的预测效果优于 B07 号电池。

图 3.15 展示了基于 CALCE 数据集的 CS2-34 号电池与 CS2-36 号电池的不同预测周期下的容量预测结果，从结果可以看出，CS2-36 号电池的预测结果优于 CS2-34 号电池，这是由于在历史输入数据中，容量波动变化 36 号电池比 34 号电池更明显，这也显示 CNN 网络更容易捕捉这种波动变化，并且两电池都表现出预测精度随着预测周期的增加而提高的现象，符合深度学习框架的普遍现象。

根据图 3.14 和图 3.15 所示的实验结果，可见通过 CNN 网络对锂电池未来容量的变化趋势做出一定程度的预测，可以较好地拟合容量波动点的变化，即捕捉到了历史容量数据的空间特征，这种容量再生点处的预测变化在 CALCE 数据集的 CS2-34 号电池上表现地最为明显，并且所有的实验结果都表明随着输入数据量的增长，预测精度也相应得到提高。锂电池容量阈值所对应的循环周期即为 RUL 终止点，RUL 终止周期减去当前周期，即为锂电池的剩余使用寿命。

基于 CNN 模型的锂电池 RUL 预测结果表明，CNN 模型可以对历史容量进行一定程度的拟合，这种趋势在空间跳跃变化与波动上较为明显，但是在时间关系上不够理想，CNN 模型在锂电池 RUL 预测中有较大的提升空间。

3.4 基于 LSTM 的锂电池剩余寿命预测

在普通的全连接神经网络与 CNN 网络中，由于信号只是在相邻层的节点中进行传播，而不具备处理时间依赖关系问题的能力，因此解决实现时序问题的循环神经网络应运而生。

3.4.1 循环神经网络 RNN

循环神经网络（Recurrent Neural Network，RNN）是深度学习中常用的网络之一，是专门用于处理时间序列数据问题的神经网络。

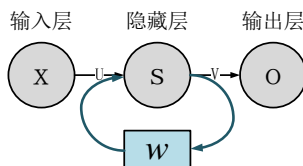


图 3.16 基本 RNN 结构

循环神经网络在之前介绍的原有全连接网络的基础之上，引入了时间轴的概念，产

生了神经网络的记忆模块，因此一个序列的结果输出不仅与当前输入有关，与前面的序列输入与输出也有密切联系。一个最基本的 RNN 结构如图 3.16 所示，RNN 结构中， x 为输入信息， o 为输出数据， s 为隐藏层的值， U 为连接输入层与隐藏层的权重值， v 为连接隐藏层与输出层的权重值。根据其结构可以看出，RNN 的隐藏层的值与之前的值密切联系，并且权重矩阵 W 即隐藏层前一个值作为当前输入的权重。现实中使用的循环神经网络拥有许多改进变体，按照时间序列展开，最经典的 RNN 模型如图 3.17 所示：

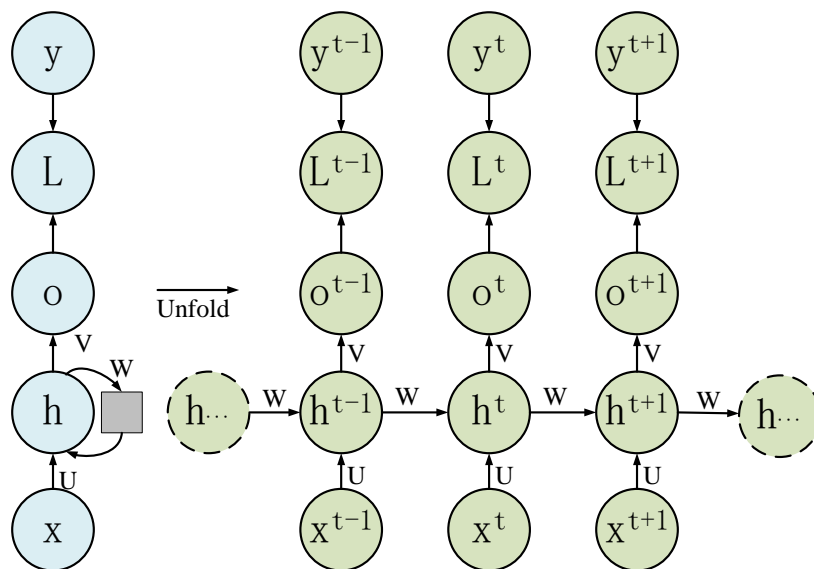


图 3.17 RNN 模型原理

其中， U 、 W 、 V 是贯穿于整个 RNN 模型之中的线下参数矩阵量，具有使整体模型前后紧密相连的重要意义； L^t 代表在 t 时刻的网络损失函数； o^t 代表在 t 时刻的网络输出结果； h^t 代表在 t 时刻的网络结构中的隐藏状态值； x^t 代表在时刻 t 的模型输入值； $y^{(t)}$ 代表在 t 时刻的真实结果值。

RNN 网络的反向传播算法使用的主要方式是梯度下降法，为了提高模型的计算精度，需要不断的进行各项参数的更新迭代，其中计算精度主要通过计算损失函数来衡量。对于普通的循环神经网络而言， $t-2$ 与 $t-1$ 时刻的网络特征是可以在 t 时刻的运算体现出来的，因此 RNN 相较于传统神经网络在处理时序数据方面是具有显著优势的，然而，在现实情况中，RNN 在学习过程中反向传播时，会带来连续的叠加，由此会导致相距较远的两数据之间会逐渐失去联系，因此 RNN 在处理长时间依赖性的数据时效果并不理想^[63]。在处理长时间依赖关系的时间序列数据时，长短时记忆网络具有更强的处理能力。

3.4.2 长短时记忆网络 LSTM

长短时记忆网络的一般结构为具有记忆功能单元的子网结构之间递归连接，各个记忆单元中含有专门处理记忆数据的“门”结构，以此来实现 LSTM 的记忆信息的各种处理操作。每一个记忆单元模块类似于一个状态机，其中的门结构主要在网络学习过程中学习获得网络权重。门结构之中主要包含 sigmoid 网络层与点乘运算操作，门结构可以通过 sigmoid 网络结构来产生一个取值于 0 至 1 的范围的输出值，此输出值可以用来描述此记忆单元的信息通过量，0 代表不允许任何值通过，而 1 代表信息可以全部通过。一般的门结构如图 3.18 所示：

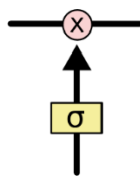


图 3.18 门结构

循环神经网络主要通过反向传播的方式来进行网络的训练，长短时记忆网络也是通过此方式来进行训练，反向传播过程主要使用梯度下降法，并且具有克服梯度消失的能力，但从结构角度出发，LSTM 比 RNN 多增加了三个门，能够克服梯度消失的主要原因也是因为增加了这三个门，这三个门分别为输入门、输出门和遗忘门，三个门负责各自不同的功能实现，但执行操作是相同的，在门结构中的选择工具为 sigmoid 函数，信息变换工具则为 tanh 函数。一个基本记忆单元的内部结构如图 3.19 所示：

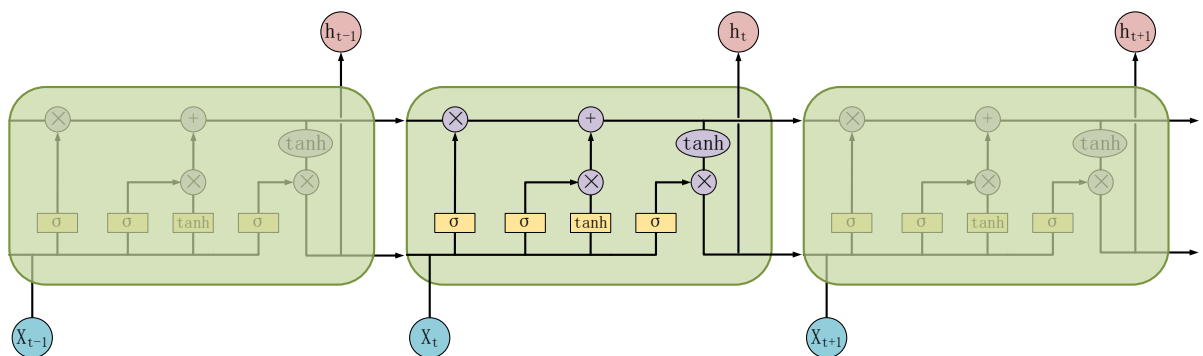


图 3.19 基本记忆单元内部结构

(1) 遗忘门：

遗忘门的结构如图 3.20 所示，遗忘门可以读取 h_{t-1} 和 x_t 的值，然后通过 sigmoid 函数给出一个 0 至 1 范围内的数值，将该数值传递给上一时刻的细胞状态 C_{t-1} ，因此遗忘

门的主要作用即一些记忆信息的丢弃处理。

遗忘门的输出为:

$$f_t = \sigma(W_f[h_{t-1}, x_t] + b_f) \quad (3.20)$$

(2)输入门:

输入门的结构如图 3.21 所示,此主要作用为确定更新细胞状态的信息。输入门中的计算量如(3.21)所示:

$$i_t = \sigma(W_i[h_{t-1}, x_t] + b_i), \tilde{C}_t = \tanh(W_C[h_{t-1}, x_t] + b_C) \quad (3.21)$$

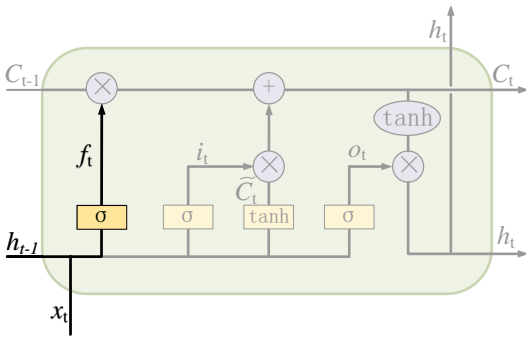


图 3.20 LSTM 遗忘门

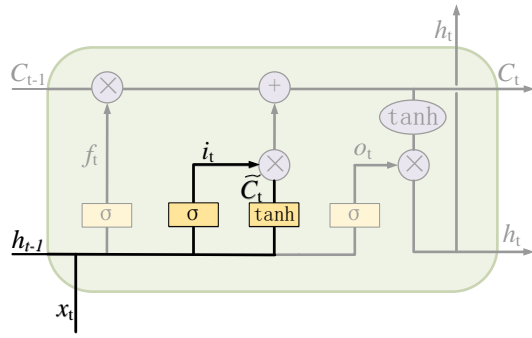


图 3.21 LSTM 输入门

细胞状态贯穿于整个 LSTM 网络层,是 LSTM 网络中最为关键的部分之一,细胞状态的线性作用很微弱,如图 3.22 所示,类似于传送带上的运作方式,是一个链式结构,因此信息可以通过这种方式轻松传播却不会轻易改变状态。在输入门的数据运算完成之后,可以进行细胞状态的更新,如图 3.23 所示。

新的细胞状态的运算为:

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t \quad (3.22)$$

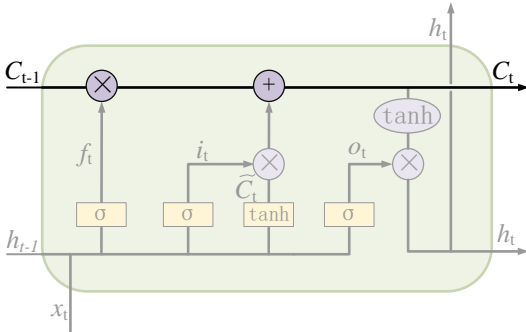


图 3.22 LSTM 模型中细胞状态传输

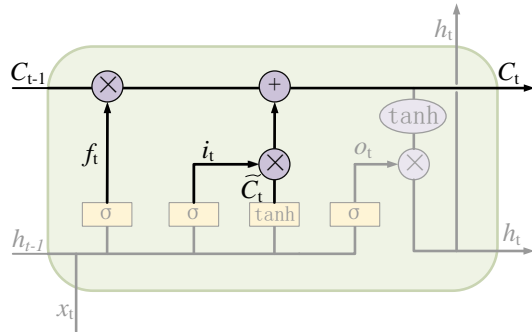


图 3.23 LSTM 模型中更新细胞状态

式(3.21)中,通过上一时刻状态 C_t 与遗忘门输出结果 f_t 相乘来丢弃一些无关数据信息,再加上 $i_t * \tilde{C}_t$ 得到新的细胞状态候选值。

(3)输出门:

输出门的主要功能为确定每个记忆单元的输出值,当细胞状态完成更新的操作之后,输出门就可以基于新的细胞状态来进行输出信息的确定。输出门的结构如图 3.24 所示:

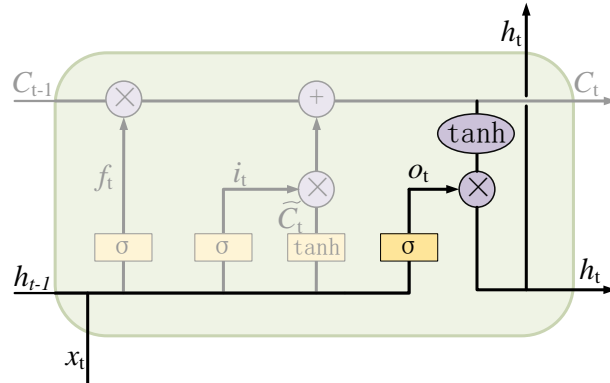


图 3.24 LSTM 输出门

根据输出门的结构有:

$$o_t = \sigma(W_o[h_{t-1}, x_t] + b_o) \quad (3.23)$$

$$h_t = o_t \cdot \tanh(C_t) \quad (3.24)$$

LSTM 基于以上介绍的三个门,实现对含有时间依赖关系的信息进行选择性的删除、遗忘及更新,基于此 LSTM 具备了处理长时间序列问题的能力,同时也可在一定程度上避免反向传播过程中梯度消失问题。

3.4.3 LSTM 网络运行训练算法

上文基于 LSTM 网络的结构层面,介绍了网络根据三个门的原理而进行信息的处理,即网络的前向传播算法过程。对于一个网络,学习过程是最主要的部分之一,与传统网络类似,LSTM 网络的学习过程也主要使用反向传播算法。LSTM 的反向传播也是使用梯度下降法,相较于 RNN 网络,LSTM 含有更多的内部参数,对于网络内参数的更新,也主要是基于参数对于损失函数的导数,实现过程如下所示。

基于 LSTM 中含有两个中间隐藏状态,可设定两个量 δ 和 φ ,如式(3.25)所示:

$$\delta_t = \frac{\partial L}{\partial h_t}, \varphi_t = \frac{\partial L}{\partial C_t} \quad (3.25)$$

在时刻 t 时, δ_t 的梯度计算过程中包含有当前 t 时刻的梯度误差,还包含有 $t+1$ 时刻的梯度误差,如式(3.26)所示:

$$\delta_t = \frac{\partial L}{\partial h_t} = \frac{\partial l}{\partial h_t} + \left(\frac{\partial h_{t+1}}{\partial h_t}\right)^T \frac{\partial L_{t+1}}{\partial h_{t+1}} = V^T (\hat{y}_t - y_t) + \left(\frac{\partial h_{t+1}}{\partial h_t}\right)^T \delta_{t+1} \quad (3.26)$$

与 δ_t 的推导过程类似，在时刻 t ， φ_t 的反向梯度误差有：

$$\varphi_t = \left(\frac{\partial C_{t+1}}{\partial C_t}\right)^T \frac{\partial L}{\partial C_{t+1}} + \left(\frac{\partial h_{t+1}}{\partial C_t}\right)^T \frac{\partial L}{\partial h_{t+1}} = \varphi_{t+1} \cdot f_{t+1} + \delta_t \cdot o_t \cdot (1 - \tanh^2(C_t)) \quad (3.27)$$

基于 δ_t 与 φ_t 的计算，可实现对 LSTM 细胞结构中的参数梯度推导，LSTM 中含有的参数较多，此处以其中的遗忘门的权重参数为例，其梯度计算过程如式(3.28)所示：

$$\frac{\partial L}{\partial W_f} = \sum_{t=1}^T [\varphi_t \cdot C_{t-1} \cdot f_t \cdot (1 - f_t)] (h_{t-1})^T \quad (3.28)$$

LSTM 网络在学习过程中几乎不会发生梯度消失或者梯度爆炸的问题，这种情况在 RNN 中较为容易发生的原因是因为激活函数的导函数的取值在(0,1)范围内，连续的小数进行叠加相乘会导致结果逐渐变小，直至接近 0。但是对于 LSTM 网络而言，在其中加入了遗忘门，遗忘门可以使得在神经元之间的数据信息并不是单纯的叠加相乘，而是在一定筛选结果的基础之上进行叠加。如式(3.29)所示连乘项：

$$\frac{\partial c_t}{\partial c_{t-1}} = \sigma(W_f x_t + b) \quad (3.29)$$

在上式所示连乘项中，所得结果均为小于 1 的非负数。在参数的更新过程中，可以通过调整 b 的取值大小，来使得梯度逐渐接近 1，以便梯度可以在 LSTM 网络中更好地进行传播，不会出现叠加相乘而逐渐消失的情况。基于此，可见 LSTM 相较于 RNN，在解决梯度消失问题上更具有优良的性能，使得 LSTM 在解决具有长时间依赖关系的时序问题上具有更为广泛的使用与研究。

3.4.4 LSTM 网络模型搭建

本课题研究使用 LSTM 神经网络模型来对锂离子电池的未来容量进行预测，通过对未来容量的预测来进行对剩余寿命进行预测，将锂电池的容量退化至寿命终止阈值时的周期数作为锂电池 RUL 寿命终止点。对于 LSTM 网络，输入值为 5 个时间步长的容量序列，即滑动窗口为 5，通过网络迭代训练与预测，将模型输出得到的单步的寿命预测值作为历史数据，带入到新的输入序列中，继续进行之后的容量预测，以此使得除最后一步的预测值以外的所有结果值都将被作为模型输入，其运行方式如图 3.25 所示。

本课题中，LSTM 网络的除输入、输出层外的隐藏层设定为三层，前两层为 LSTM 层来提取历史容量数据的特征，第三层作为全连接层来特征向量进行低维处理从而得到预测结果值。在模型的训练过程中，初始学习率设定为 0.01，学习率下降因子设定为 0.5 使用均方根误差损失函数来获得 LSTM 预测容量值与真实容量值之间的误差，并且使用 Adam 优化器来实现反向传播计算参数梯度并实时更新模型的参数值。

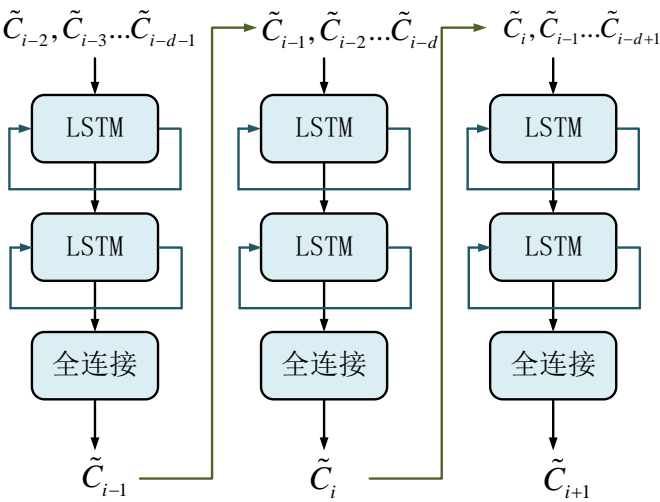


图 3.25 基于 LSTM 的锂电池容量预测过程

LSTM 网络的具体网络结构与参数设置如表 3.3 所示：

表 3.3 LSTM 网络结构与参数

名称	类型	结构	激活数	可学习参数
sequenceinput	序列输入	5 个维度	5	-
lstm_1	LSTM	10 个隐含单元	10	InputWeights:40×5 RecurrentWeights: 40×10 Bias: 40×1
lstm2	LSTM	20 个隐含单元	20	InputWeights:80×10 RecurrentWeights: 80×20 Bias: 40×1
relu	ReLU	ReLU	20	-
fc	全连接	全连接层	1	Weights:1×20 Bias: 1×1
Regressionoutput	回归输出	mean-squared-error	-	-

3.4.5 实验结果与分析

构建基于 LSTM 网络的容量预测模型，对于 NASA 数据集中的 B05 号电池与 B07 号电池，分别使用其前 70、80、90 周期作为训练集，同时第 70、80、90 周期为未来容量序列的预测起点，预测结果如图 3.26 所示。

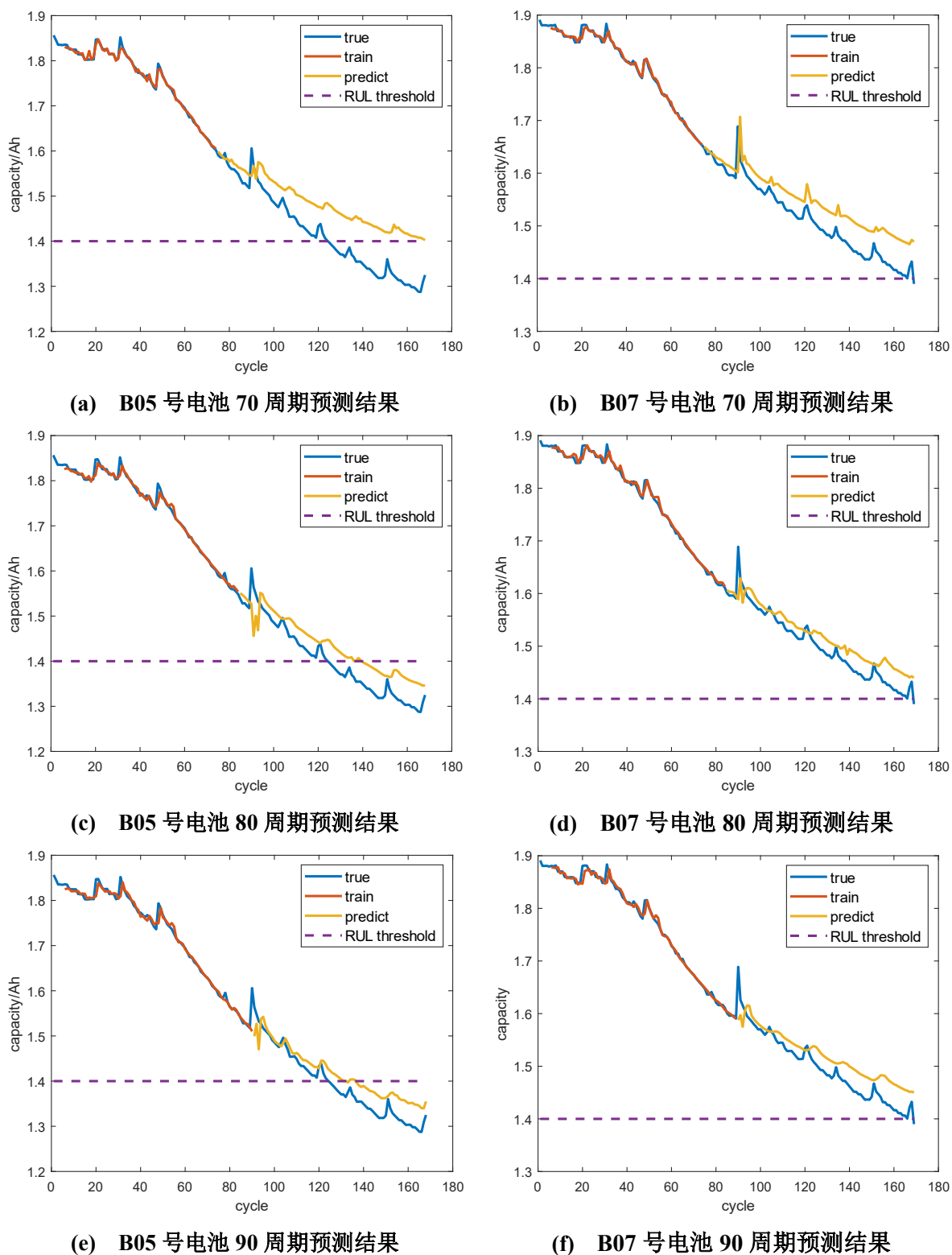


图 3.26 基于 LSTM 模型 NASA 数据集预测结果

对于 CALCE 数据集中的 CS2-34 号电池与 CS2-36 号电池,分别使用其前 40%、50% 和 60%周期作为训练集,同时也是预测起点,预测结果如图 3.27 所示。根据预测结果中的容量阈值获取 RUL 终止周期,根据终止周期可得锂电池的 RUL 预测结果。基于 LSTM 模型的容量预测结果误差如表 3.4 所示。

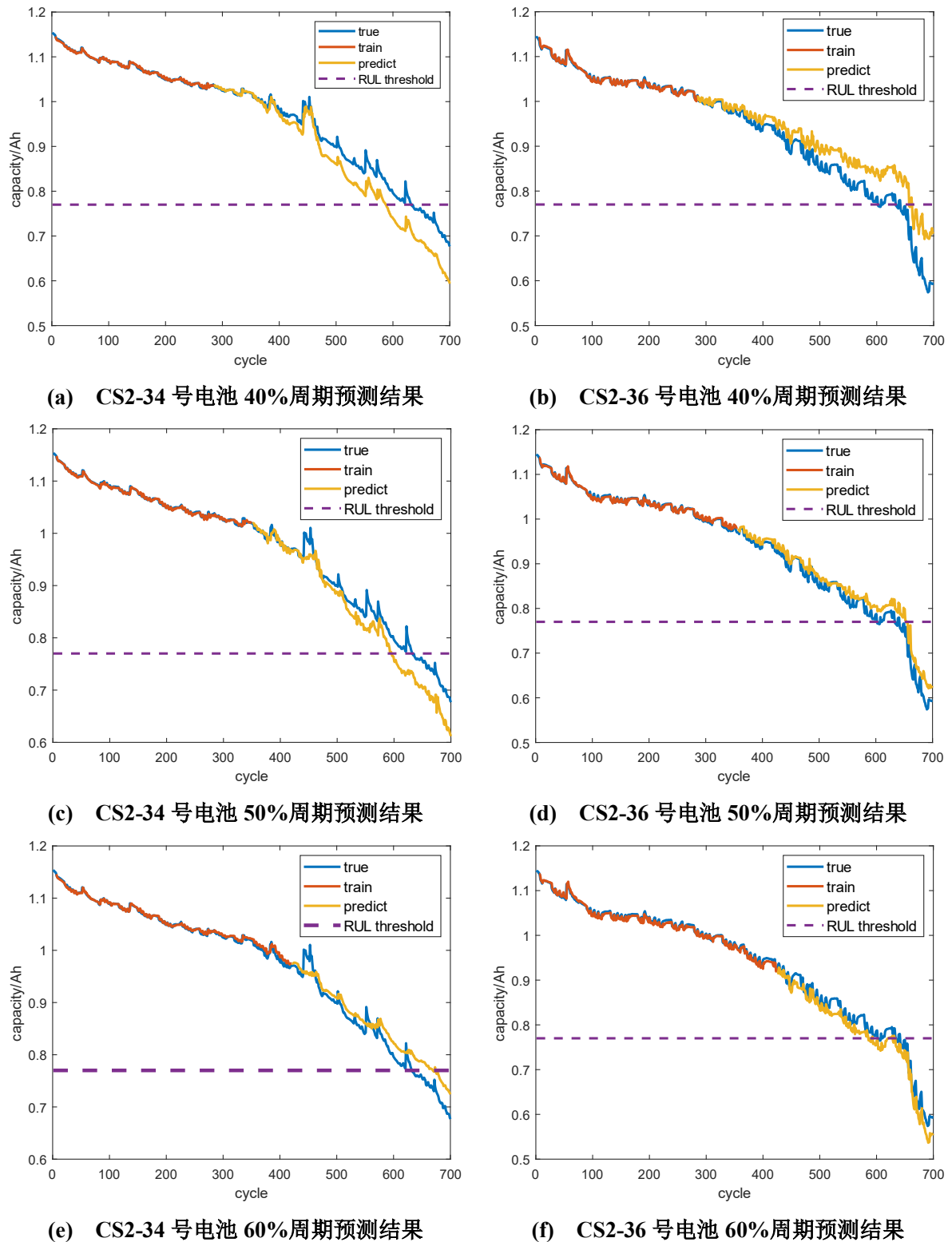


图 3.27 基于 LSTM 模型 CALCE 数据集预测结果

表 3.4 展示了 NASA 数据集与 CALCE 数据集的预测结果误差，包含了基于 RMSE 与 MAPE 两个评价指标的实验结果。根据实验结果误差可见，训练周期数与预测结果精度成正比关系，尤其这种精度的提升变化在 CALCE 数据集的两个电池的容量预测结果上表现得更为明显，这是主要因为 CALCE 数据集的周期数要比 NASA 数据集的周期数

在数量上更占优势，使得模型在训练阶段得到更多的更迭训练。

表 3.4 基于 LSTM 的容量预测实验结果误差

数据集	电池型号	周期数	RMSE	MAPE(%)
NASA	B05	70	0.048	4.683
		80	0.044	2.846
		90	0.035	2.160
	B07	70	0.045	2.505
		80	0.028	1.578
		90	0.033	1.889
CALCE	CS-34	40%	0.043	4.549
		50%	0.029	2.776
		60%	0.020	1.915
	CS-36	40%	0.050	4.847
		50%	0.027	2.866
		60%	0.021	2.451

图 3.26 展示了 NASA 数据集中 B05 与 B07 号锂电池的容量预测结果拟合曲线。从容量序列拟合曲线看，容量下降趋势与真实下降趋势较为相近，对于两个电池而言，预测精度都随着预测起点 70、80 及 90 而逐级递增，并且从周期 70 到周期 80 的精度提高比从 80 到 90 的提高要显著，这是由于 80 至 90 周期之间的容量再生情况影响了后续的迭代预测过程，受到影响的数据量占比较多，导致精度提高受到一定影响，并且两电池在容量抖动处的预测结果偏差稍大。

图 3.27 展示了 CALCE 数据集中 CS2-34 号电池与 CS2-36 号电池的容量预测结果拟合曲线，从拟合曲线可以看出，两电池在容量预测上表现得良好，随着周期数的增加而精度提高，34 号电池从 40%周期至 50%周期的精度提升更为明显，同样也是 34 号电池受到容量再生抖动点的影响。对于 36 号电池，其容量在到达寿命终止点时，容量急剧下降，与实际情况相符合，在现实中此时电池已不再使用。

根据 NASA 数据集与 CALCE 数据集的容量预测结果，LSTM 网络相较于 CNN 网络在预测精度上提高了很多，容量下降趋势与真实下降趋势较为相近，且容量预测值与真实值的偏差随着历史数据的增加而递减，即预测起点周期越靠后，预测精度越高，此时将预测容量终止周期减去当前循环周期，即为剩余寿命使用循环次数。在同一周期内的预测过程中，预测过程的后期精度明显低于预测前期的精度，这是由于对于时序数据的预测是基于前期结果的迭代计算，预测后期将受到越来越大的误差迭代累积的影响。

相较于 CNN 网络的未来容量序列预测结果,以 B05 号电池为例,LSTM 网络的预测结果精度在 70、80、90 周期的 MAPE 上分别下降了 5.597%、1.303%和 3.148%,可见 LSTM 在预测精度上得到了一定程度的提高,在时间退化趋势上展现出了较为良好的预测效果,但同时也可看出 LSTM 在容量波动处的处理效果并不是很理想,对于容量再生等波动情况的敏感程度较低于 CNN 网络。虽然 LSTM 网络随着周期数增加而预测精度得到提升,总体预测水平高于 CNN 网络,但是同时也存在对于波动数据处理不够精准,迭代误差累积等,模型在基于锂电池时序容量的预测与剩余使用寿命预测仍存在很大的进步空间。

3.5 本章小结

本章通过对深度学习相关技术与所使用数据集的介绍,对锂电池 RUL 预测过程进行介绍与设定,分别建立了基于 CNN 神经网络和 LSTM 神经网络的锂电池容量预测模型,并展示了其实验结果,通过两种网络的实验结果展示了深度学习模型进行时间序列预测与分析的可行性。CNN 可以捕捉锂电池容量退化数据的空间特征,LSTM 网络可以解决时间问题,二者都可以根据历史容量数据对未来的容量衰减趋势做出预测。通过结合容量曲线对 CNN 与 LSTM 网络预测的结果进行分析与比较,表明在锂电池的容量退化趋势预测问题上,LSTM 网络比 CNN 网络取得了更好的效果,通过研究与分析,得到 LSTM 网络更适用于研究时间序列问题的结论,同时也为后续的研究提供理论基础与出发点。

第四章 基于融合型网络 CNN-LSTM-Attention 的锂电池剩余寿命预测

为了更好地捕捉锂电池容量数据集的容量退化特征,进一步提高锂电池 RUL 的预测精度,基于电池历史容量数据所具备的时序相关性,本章根据前文建立的 CNN 预测模型与 LSTM 预测模型,结合二者优点,建立了一种含注意力机制的融合型卷积神经网络-长短期记忆-注意力(convolutional neural networks-long short term memory-Attention,CNN-LSTM-Attention)网络预测模型。首先使用滑动窗口按照设定步长输入锂电池历史容量数据,通过该模型进行连续迭代计算预测容量序列,从而获得锂电池 RUL 预测值,并且在网络中加入了防止过拟合的 Dropout 技术,提高模型训练效率,同时也是为后文进行锂电池 RUL 不确定性研究奠定基础。

4.1 深度学习 Dropout 技术

在各类机器学习的训练过程中,容易出现过拟合的现象,导致模型在测试阶段预测精度无法提高,其主要原因是模型的训练参数众多,而供以训练的数据在数量上却不足以支撑如此多的参数量,模型在训练阶段的损失函数较小,而在测试阶段的损失函数较大。过拟合是深度学习模型中很常见的现象,同时也是模型精度受阻的主要成因之一。在模型的学习过程中,是指学习器把训练样本学的“太好了”,很可能把个别的训练样本的一些特点当作了所有样本数据的一般特征^[64]。如果模型参数数量和迭代次数庞大,将导致训练时间过久,造成效率低下。基于以上两个问题,Dropout 技术应运而生,Dropout 技术可以较高水平地缓和模型的过拟合问题,一定程度上起到对模型正则化的作用,即提高模型的泛化能力,同时为第五章的剩余寿命不确定性分析提供理论基础。

Dropout 作为神经网络学习过程中的一种可供选择的技巧,在每一次的训练过程中,忽略掉神经网络层中的某一部分特征检测,从而减少各特征检测器之间的相互影响,以此来达到减少过拟合现象的发生。从技术角度出发,Dropout 就是在网络训练的前向传播过程中,以一定概率 P 来使得某个神经元失去工作能力,从而使得模型提高其泛化能力,从而不依赖网络层中某些局部的特征。其结构原理如图 4.1 所示:

Dropout 技术对深度学习的影响主要体现在以下几个方面:

(1) 取平均作用。不同网络的训练过程可能会导致相对相反的过拟合结果，那么 Dropout 可以通过取平均作用的方式来使得相反结果相互抵消，从而防止整体层面的过拟合。在训练过程中删除掉不同的隐藏层神经元就相当于每次训练的是不同的网络，因为神经元的存在与否导致了网络结构的改变，那么整个 Dropout 过程就相当于是对不同的神经网络取了平均，不同网络相反结果的过拟合互相抵消就可达到宏观结构的减少过拟合。

(2) 减轻不同神经元之间的共适应作用。Dropout 技术会导致不同的神经元在同一个 dropout 层不会共存，这样的话部分权值的更新不会再依赖于某些含有固定联系的神经元之间的共同作用，从而减弱了某些局部特征对于整体效果的影响，使得网络的训练往鲁棒性更强的方向去发展，将网络的重心聚焦到整体的共同特征，提高泛化能力。

(3) 提高网络的适应力。Dropout 技术类似于在生物发展过程中为了适应环境变化而做出的反应力，犹如有性生殖的存在可以使得生物繁衍出适应新环境的下一代变种，Dropout 技术最核心的防止过拟合作用相当于有性繁殖避免了物种的灭绝。

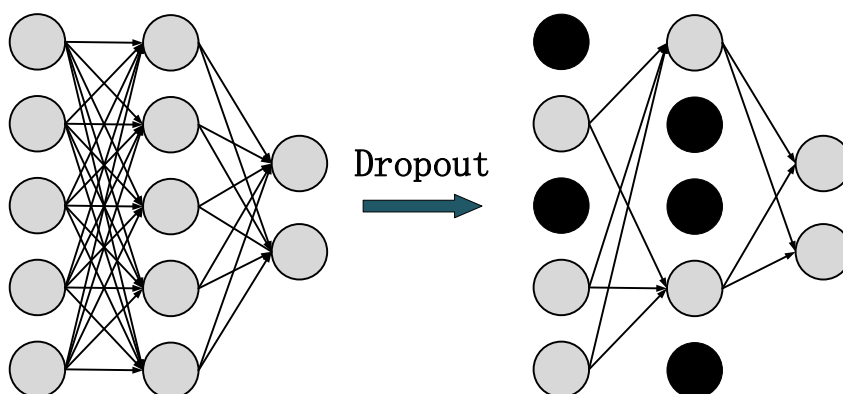


图 4.1 Dropout 原理

Dropout 技术实现流程如下所示：

(1) 随机地删除隐藏网络层中的部分神经元，而不改变输入层和输出层。

(2) 将模型输入 x 根据上一步修改后的网络进行前向计算，将得到的结果根据损失函数计算误差，将误差使用网络再进行反向传播来计算梯度，一小批次的训练数据进行前向计算与误差反向传播过程后，在为被删除的的神经元上进行参数的更新，从而完成基于当前训练样本的学习过程，整个过程中被随机删除的神经元则保留了原始参数值，未被更新。

(3) 恢复第一步中被停止工作的神经元，从隐藏层中再次随机选择神经元进行删

除, 重复第二步的操作过程。不断地重复以上三个过程。Dropout 计算过程如 (4.1)所示:

$$\begin{cases} P_l^i \sim \text{Bernoulli}(p) \\ Z_{l+1}^i = w_{l+1}^i * p_l^i * y_l^{\text{input}} + b_{l+1}^i \\ y_{l+1}^i = f(Z_{l+1}^i) \end{cases} \quad (4.1)$$

其中, P_l^i 代表在网络中第 l 层中第 i 个神经元的随机丢弃系数, 其取值一般为 1 或者 0, 分别代表在此次学习迭代过程中是被保留还是被丢弃, w_{l+1}^i 、 b_{l+1}^i 分别代表了在 $l+1$ 层中神经元的权值与偏置量, y_l^{input} 代表经过 Dropout 层之后的第一层神经元, y_{l+1}^i 代表第 $l+1$ 层中经过前向处理即激活函数处理后的网络层输出。

4.2 改进型 CNN-LSTM 混合网络

CNN 与 LSTM 是深度学习模型中最典型的网络之一, 在深度学习中各自含有独特的优势, 可针对不同的问题^[65]。CNN 擅长处理数据的空间问题, 而 LSTM 擅长处理数据的时间问题。基于此, 本课题提出结合二者优势的混合网络 CNN-LSTM。

CNN-LSTM 从模型功能角度而言是指先利用 CNN 模型提取电池历史容量数据的空间特征, 在提取出的空间特征数据的基础之上, 再利用 LSTM 网络提取出电池历史容量数据的时间特征, 最后再进行电池未来容量的估计输出。模型结构如图 4.2 所示:

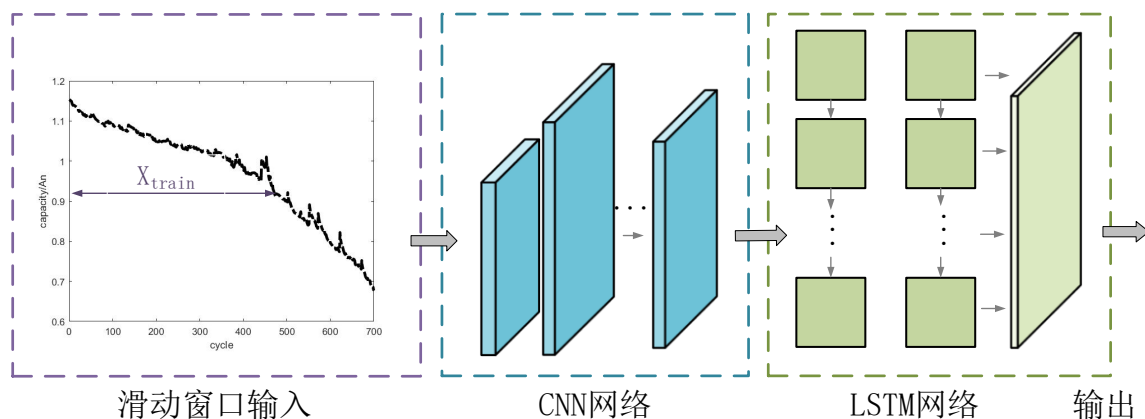


图 4.2 CNN-LSTM 混合模型

CNN-LSTM 混合模型的上层由 CNN 网络构成, 可以接受电池历史容量数据变量, CNN 主要由输入层、卷积层、ReLU 以及池化层构成。其中输入层用来接收数据, 卷积层用来对电池容量类的时序数据进行一定的卷积运算, 将得到的特征数据传输给下一层。如果设定历史容量数据为 $x = (x_1, x_2, \dots, x_n)$, 其中 n 代表变量的个数, 经过标准化处理后的数据为 x_i , 那么第一个卷积层的输出有:

$$y_{ij}^1 = \sigma \left(\sum_{m=1}^M W_{m,j}^1 x_{i+m-1,j}^0 + b_j^1 \right) \quad (4.2)$$

式(4.2)中 W 代表核的权重, M 代表滤波器数目, σ 代表 ReLU 为激活函数。那么第 l 个卷积层的输出为:

$$y_{ij}^l = \sigma \left(\sum_{m=1}^M W_{m,j}^l x_{i+m-1,j}^0 + b_j^l \right) \quad (4.3)$$

数据经过卷积处理后, 对卷积输出进行池化处理, 则最大的池化操作为:

$$p_{ij}^l = \max_{r \in R} y_{i \times T + r, j}^{l-1} \quad (4.4)$$

式(4.4)中 R 代表池化大小, 其需小于输入 y 大小; T 代表了合并区域移动步长大小。

CNN-LSTM 混合模型的下层由 LSTM 网络构成, 主要用于处理通过 CNN 网络提取出的容量序列特征的时间信息。LSTM 网络主要处理大范围时间问题, 此功能主要源于其可以对先前隐藏状态的内存单元进行合并后更新, 以此来保留长期的内存信息。LSTM 擅长于解决时序数据模型在训练阶段容易发生的梯度消失和爆炸问题, 其特有的三个门结构可以通过运算来确定独立单元的状态存储机制, 这是相较于 RNN 的先进性。LSTM 通过对各个门的激活态来对各个存储单元的当前时刻状态进行更新, 这样使得门单元的激活量被控制为是 0 至 1 之间的连续值, 各个单元的隐藏状态 h_t 每 t 步更新一次。

$$i_t = \sigma(W_{pi} p_t + W_{hi} h_{t-1} + W_{ci} * c_{t-1} + b_i) \quad (4.5)$$

$$f_t = \sigma(W_{pf} p_t + W_{hf} h_{t-1} + W_{cf} * c_{t-1} + b_f) \quad (4.6)$$

$$o_t = \sigma(W_{po} p_t + W_{ho} h_{t-1} + W_{co} * c_t + b_o) \quad (4.7)$$

式(4.5)、(4.6)、(4.7)分别代表 LSTM 记忆单元的输入门、遗忘门以及输出门的操作过程, 其中 i_t 、 f_t 、 o_t 分别代表输入门、遗忘门和输出门的输出, 由三个门确定的单元状态与隐藏状态分别由 c 和 h 表示, 如下所示:

$$c_t = f_t * c_{t-1} + i_t * \sigma(W_{pc} p_t + W_{hc} h_{t-1} + b_c) \quad (4.8)$$

$$h_t = o_t * \sigma(c_t) \quad (4.9)$$

式(4.8)和(4.9)中, W 代表门单元权重矩阵, b 代表偏置量, σ 代表具有非线性的激活函数, 其同时具有压缩输入至 -1 至 1 范围的功能; p_t 作为时间 $t-1$ 时刻 CNN 池化层的输

出, 包含有容量数据的关键特征, 并且也是 LSTM 网络的存储单元输入。CNN-LSTM 混合模型基于 LSTM 部分对信号的时间信息进行建模, 使得整个模型在未来容量预测有了更进一步的性能提升。

CNN-LSTM 网络的最后一部分是全连接层, 全连接层的运算如式(4.10)所示:

$$d_i^l = \sum_j w_{ji}^{l-1} \left(\sigma(h_i^{l-1}) + b_i^{l-1} \right) \quad (4.10)$$

式(4.10)中, l 代表 LSTM 网络最后一层的单元数目, LSTM 的输出被展平为特征向量 $h^l = (h^1, h^2, \dots, h^l)$, LSTM 的输出也即是全连接层的输入, σ 代表激活函数, w 代表 $l-1$ 层中第 i 节点与 l 层第 j 节点之间连接的权重, b_i^{l-1} 代表偏差, d 代表全连接层的输出, 即整个 CNN-LSTM 混合网络的输出。CNN-LSTM 混合网络的训练框架如算法 4.1 所示:

算法 4.1 CNN-LSTM 模型训练框架

输入: 训练集 D

输出: 训练后的 CNN-LSTM 模型参数

初始化模型参数;

1: For $i=1$: epoch do:

2: 模型前馈

3: 电池容量退化数据与 CNN 网络进行卷积操作, 使用 ReLUc 进行非线性变换;

4: LSTM 分析并拟合处理 CNN 输出的特征, 输出新的时空特征;

5: 全连接层接收 LSTM 输出, 实现最终的预测结果输出

6: 利用均方误差损失函数计算预测值与真实值之间误差;

7: 反向传播;

8: Adam 优化器反向传播计算梯度并更新模型参数

9: End

通过以上理论可知, 对于整个 CNN-LSTM 混合网络而言, 其中分别包括了 CNN 网络跟 LSTM 网络的设计。CNN 作为前一部分, 包含了卷积网络 CNN 层数、卷积核数目与大小、池化层函数和大小以及全连接层的层数和神经元个数。LSTM 结构的设计包括确定 LSTM 的层数、每层的单元数。

4.3 注意力机制 Attention

2015 年, Bahdanau 等人首次提出了自然语言处理领域里的注意力机制, 展示了最经典的注意力结构并将其应用于机器学习, 直观地反映了注意力机制对于深度学习模型带来的积极影响^[66]。

深度学习中的注意力机制与人体大脑的反应机制类似, 人类在观察某些事物是, 一般而言先大概浏览全局, 然后根据大脑判断定位至比较关注的部位^[67]。这种对信息进行进行一定机制的取舍, 是包括人类在内的生物在大脑经过长期的进化而形成的生存机理, 被称作注意力机制, 可以使个体在一定资源受限的状况下, 可以根据自己的经验, 在纷繁复杂的环境区域中迅速捕捉到对自己有用的信息或者资源, 从而使得相应活动变得更为高效和快速。

Attention 机制的最早应用是在 E-D 任务之中, 其核心思想便是模仿人体大脑的注意力集中模式。一般而言, 对于一个待解决的问题, 人的大脑往往会将注意力集中在一些少数的且比较关键的角度, 而忽略那些无挂紧要的信息。对于一个常规性的传统 E-D 模型, 定义输入数据为集合 $X = (x_1, x_2, x_3 \cdots x_t)$, 输出为 $Y = (y_1, y_2, y_3 \cdots y_t)$, 设定中间变量为 C , 则编码过程 Encode 为: $C = E(x_1, x_2, x_3 \cdots x_t)$, 那么解码过程则可设定为 $y_i = D(C, y_1, y_2, y_3 \cdots y_{i-1})$, 在解码过程中需要利用中间变量 C 与历史结果数据。根据 Encoder-Decoder 模型, 对于解码过程的解码源是相同的, 即具有相同的中间语义编码, 这种机理可能会导致输入序列的元素对于模型的输出序列的每一个元素具有相同的影响力。同时, 中间语义的编码结果跟每一个输入序列都有相关性, 同时这种相关性将随着输入序列的增多而变强, 也基于此, 传统 Encoder-Decoder 受序列的先后顺序的影响较大。

在 LSTM 网络后面加入注意力层, 经过 LSTM 网络处理过的向量, 作为 Attention 层的输入, 在 Attention 中根据输入的向量计算权重向量, 进而使用当前输入向量与计算好的权重向量进行合并处理, 将计算出的新特征向量输入至全连接层, 从而计算最终结果值。Attention 模型的结构运行过程如图 4.3 所示:

图 3.4 所示过程中, x 表示输入的序列, h 对应输入序列隐藏层的状态值, 模型中的计算过程如下所示:

$$h_t = f(x_t, h_{t-1}) \quad (4.11)$$

$$P(y'_t | y_1 \dots y'_{t-1}) = P(y'_{t-1}, S'_t, c'_t) \quad (4.12)$$

$$S'_t = g(y'_{t-1}, c'_t, S'_{t-1}) \quad (4.13)$$

$$a_{ki} = \frac{\exp(e_{ki})}{\sum_{j=i}^{T_x} \exp(e_{kj})} \quad (4.14)$$

$$e_{ki} = V^T \tanh(WS_{t-1} + Uh_i + b) \quad (4.15)$$

$$c = \sum_{i=1}^{T_x} a_{ki} h_i \quad (4.16)$$

其中, a_{ki} 代表输出节点 k 对于输入节点 i 的注意力 Attention 权重, 其计算模式对于注意力机制的功能与效率起决定性作用, V 、 W 、 U 表示矩阵。在编码阶段, h_t 代表在 LSTM 网络中的隐藏层变量, C 为个隐藏层关于注意力机制概率分布的乘积项累加而得的中间值。

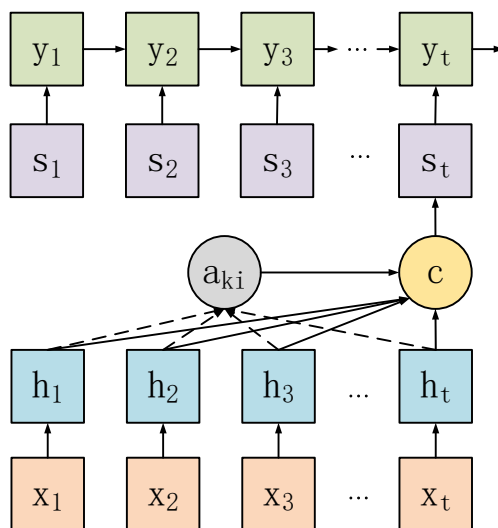


图 4.3 Attention 运行过程

注意力机制对 LSTM 特征向量注意力分布的概率进行计算, 可以使得模型将那些对目标结果有贡献的部分进行有选择性的学习, 从而增强了特征表达的稀疏性能力。Attention 计算 LSTM 层的每一个隐藏状态与最末状态的关联性, 得到每个状态的 Attention 权重, 以此来对强相关时刻状态的信息特征表达进行强化, 对弱相关的信息特征进行了一定的弱化, 从而提高模型的整体运算水平。

4.4 基于融合技术的 CNN-LSTM-Attention 网络

为了进一步提高本课题提出的 CNN-LSTM 网络对于锂电池剩余使用寿命预测结果

的预测精度，将注意力机制与 CNN-LSTM 网络进行融合，构建一种改进的融合型混合网络 CNN-LSTM-Attention。

4.4.1 CNN-LSTM-Attention 网络结构与原理

为了使得锂电池 RUL 的预测更为精准，本课题融合了混合网络 CNN-LSTM 与注意力机制，构建了含注意力机制的混合网络 CNN-LSTM-Attention 网络，用于锂电池的剩余使用寿命预测，其结构如图 4.4 所示：

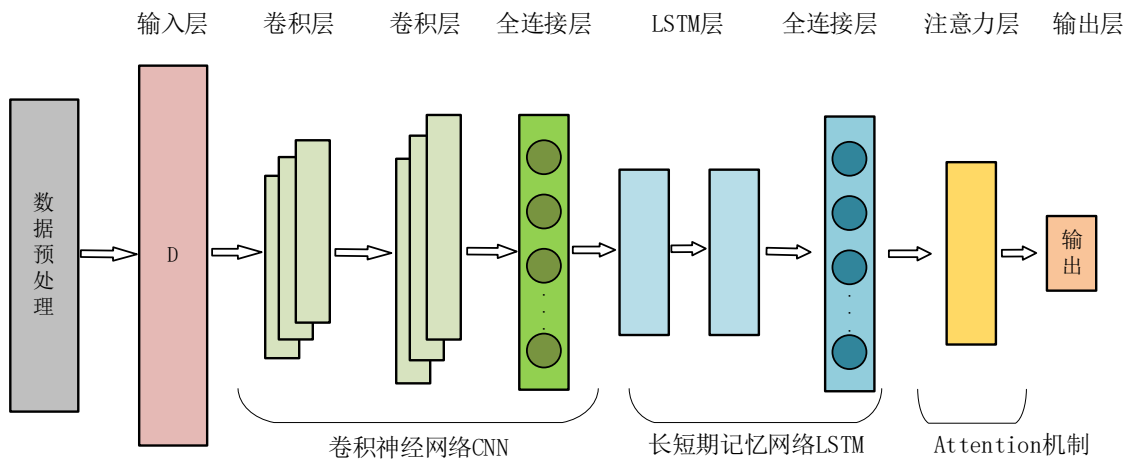


图 4.4 CNN-LSTM-Attention 网络结构

CNN-LSTM 结构融合了 CNN 与 LSTM 各自的优点，利用 CNN 擅长提取空间特征的能力与 LSTM 擅长提取数据时间特征的能力相融合，再使用注意力机制对 LSTM 输出向量通过分配概率权重的方式来实现优化过程，捕捉 CNN-LSTM 混合网络提取到的时空特征的关键信息，进而提高模型的预测精度。CNN-LSTM-Attention 网络主要包含了三大主体部分，即卷积网络 CNN、长短期记忆网络 LSTM 与注意力层 Attention，其中在 LSTM 网络中的每一个 LSTM 层后都添加了 Dropout 层，添加 Dropout 层主要是为了提高模型训练效率，防止过拟合，进而提升深度学习模型的泛化能力与鲁棒性。

在 CNN-LSTM 模型中引入注意力机制，主要是为了提升模型结构第二部分 LSTM 的 Encoder-Decoder 模型的效果，在 LSTM 网络后面引入 Attention 机制，它可以使得历史节点对于当前节点的影响得到更为合理的权重分配，从而可以防止一些关键的信息在历史数据的更迭中被丢弃，这样的融合方式可以为模型预测算法提供优化后的特征向量，从而提高模型预测的精度与鲁棒性，与此同时，不会使模型的计算量与存储开销随着模型的训练而变大。从 CNN 网络输出的特征向量，输入至 LSTM 网络中，经 LSTM 网络处理后的特征向量，作为 Attention 的输入，计算出注意力机制的权重向量，将权重向量

与当前输入向量合并后再输送至全连接层,得到最终的预测结果。CNN-LSTM-Attention 网络与 CNN 和 LSTM 网络的训练方式相同,使用正向前馈进行预测计算,使用反向误差传播来进行参数更新学习。

加入了注意力机制的融合型网络 CNN-LSTM-Attention 不仅可以对输入的锂电池历史容量数据的时间与空间特征进行有效学习,更可以进一步归纳表达特征,使得模型对于特征的表达更加明确,减弱次要特征的干扰力,从而使模型的预测精度进一步提高。

4.4.2 基于 CNN-LSTM-Attention 网络的锂电池 RUL 预测建模

本课题中使用 CNN-LSTM-Attention 融合型网络对锂电池剩余使用寿命进行预测。使用滑动窗口,模型输入为五个步长的历史容量数据,得到单步前向的容量预测值,通过重复的单步前向迭代运算,将锂电池剩余寿命终止阈值对应循环数减去当前循环数,即为最终的 RUL 预测结果值,滑动窗口与单步前向预测原理如图 4.5 所示。

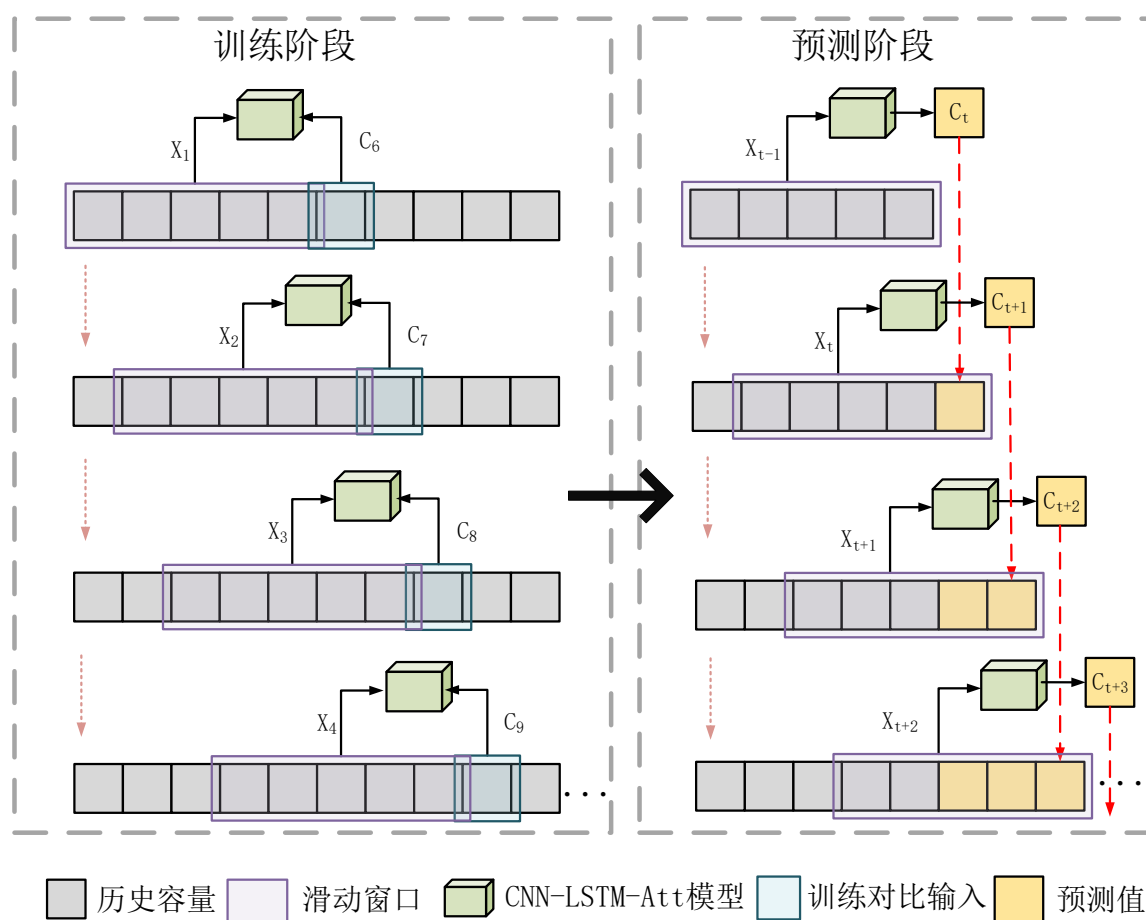


图 4.5 滑动窗口与单步前向预测原理图

CNN-LSTM-Attention 网络的输入层,对长度为 5 的历史容量数据进行预处理,然后将该序列数据作为模型输入。

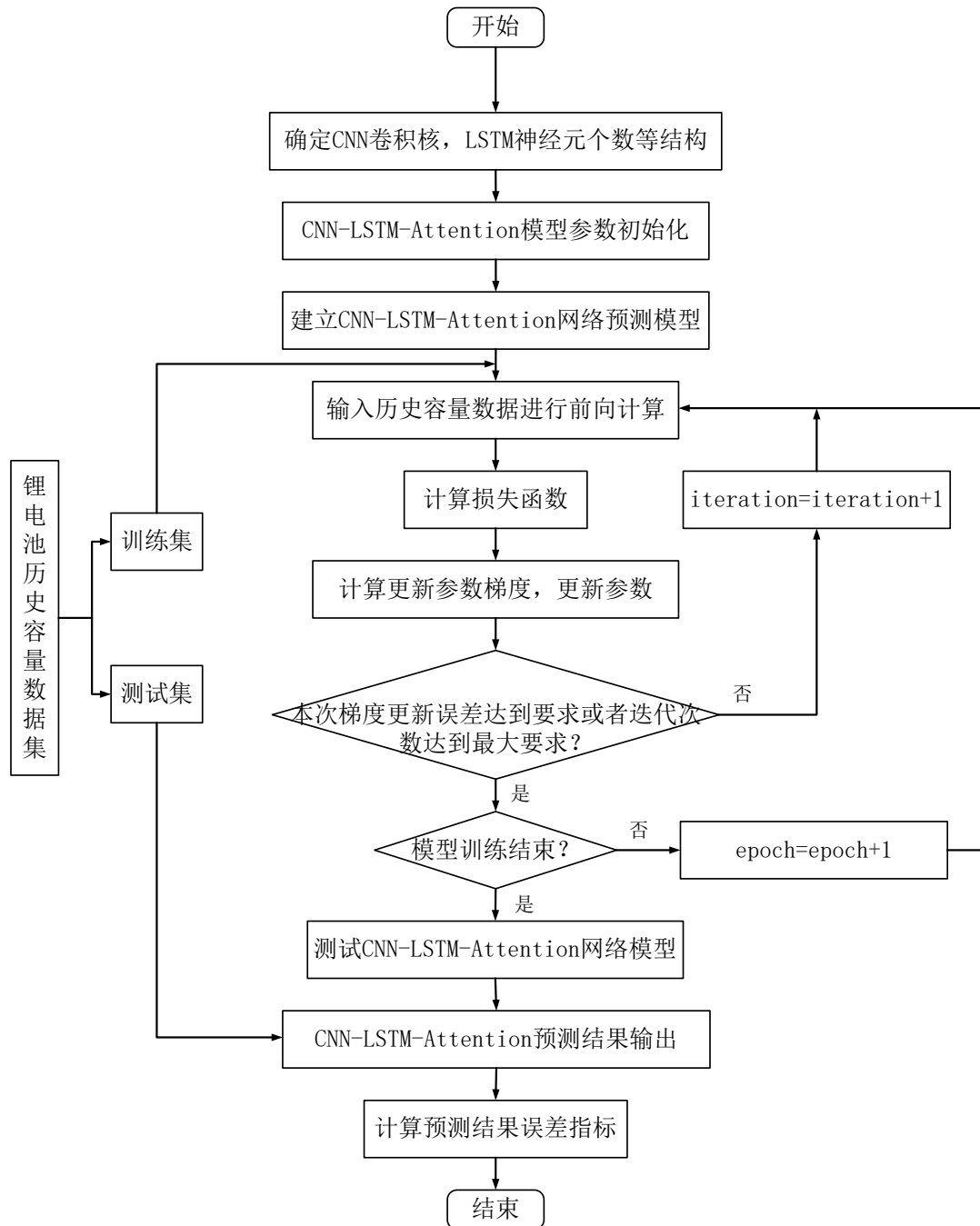


图 4.6 CNN-LSTM-Attention 模型训练流程

在第一主体部分 CNN 网络中, CNN 层设计为 2 层卷积层与 1 层全连接层, 两个卷积层均为一维卷积, 卷积核大小为 3, 第一个卷积层含 16 个卷积核, 第二个卷积层含 32 个卷积核, 全连接层将卷积处理后的特征向量输出, 卷积层使用的激活函数为 Relu 函数, 将卷积网络输出的特征向量进行扁平化处理后, 输入至 LSTM 网络中。LSTM 网络中构建两层 LSTM 层和一个全连接层, LSTM 的单个数设定为 10, LSTM 层之后的 Dropout 层中 dropout 率设置为 0.2, LSTM 输出的特征向量经过全连接层后, 输入至注意力层, 在注意力层用概率分配权重的方式来计算权重矩阵, 最后将计算得到的结果经

输出层输出。CNN-LSTM-Attention 网络使用前向传播进行预测运算，使用 Adam 优化器通过反向传播的方式来更新模型参数。整个 CNN-LSTM-Attention 预测模型的训练流程如图 4.6 所示。

4.5 实验结果与分析

本章与第三章实验环境保持一致，实验数据仍然采用 NASA 数据集 B05 号电池与 B07 电池和 CALCE 数据集 CS2-34 号电池与 CS2-36 号电池的容量退化数据集。为体现模型对比度以及泛化，滑动窗口大小与第三章 CNN 网络和 LSTM 网络保持一致，也是单步前向预测，利用五个连续的历史容量数据作为模型的输入，得到前向一步的容量预测结果。从实验结果图中可以看出，CNN-LSTM-Attention 模型可以很出色地拟合数据集训练周期内的容量退化趋势，具有较强的学习能力，对未来序列容量做出了很好的预测，具有较高的精度。

4.5.1 CNN-LSTM-Attention 网络实验结果

如表 4.1 所示为使用 CNN-LSTM-Attention 网络所得到的实验结果误差。

表 4.1 基于 CNN-LSTM-Attention 容量预测实验结果

数据集	电池型号	周期数	RMSE	MAPE(%)
NASA	B05	70	0.018	1.063
		80	0.013	0.893
		90	0.011	0.551
	B07	70	0.022	1.160
		80	0.020	0.940
		90	0.017	0.708
CALCE	CS2-34	40%	0.015	1.439
		50%	0.010	0.996
		60%	0.008	0.591
	CS2-36	40%	0.017	1.635
		50%	0.012	1.051
		60%	0.011	1.149

从两类数据集的预测结果误差来看，该模型展现出优秀的预测效果，预测精度很高，RMSE 与 MAPE 也是总体上呈现出随训练周期数的增加而减小的现象，说明该模型也同样会受训练数据量的影响，对于其中的 CS2-36 号电池，50%训练周期情况下的误差最低。NASA 数据集的 B05 号电池与 B07 号电池的预测结果如图 4.7 所示，CALCE 数

数据集的 CS2-34 号电池与 CS2-36 号电池的预测结果如图 4.8 所示。

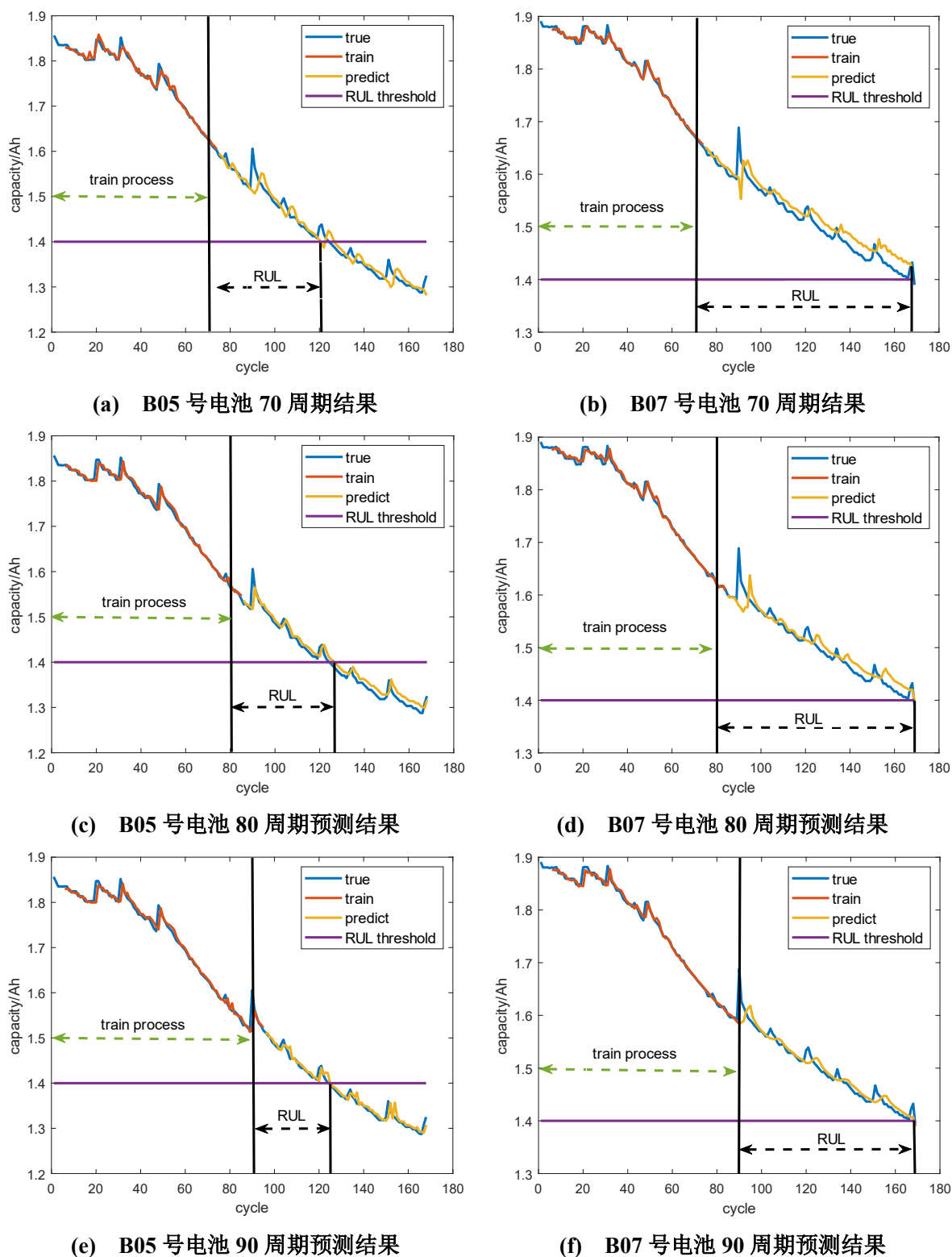


图 4.7 NASA 数据集预测结果

图 4.7 展示了 NASA 数据集的预测结果，从两个电池容量预测结果可以看出，同样随着周期数的增加，预测精度也是逐渐提高，从最终表现而言，B5 号电池 90 周期的预测结果为最佳，而 B7 电池 70 的预测结果不够理想，并且主要体现在对容量再生点的预

测，增加了后续迭代预测的误差。从同一个预测起始周期角度可发现，B05 号电池的预测拟合曲线比 B07 号电池更为接近真实曲线，主要是因为 B7 号电池的容量再生情况比较显著，这类抖动较大的点，对于模型预测的影响还是较为明显的。

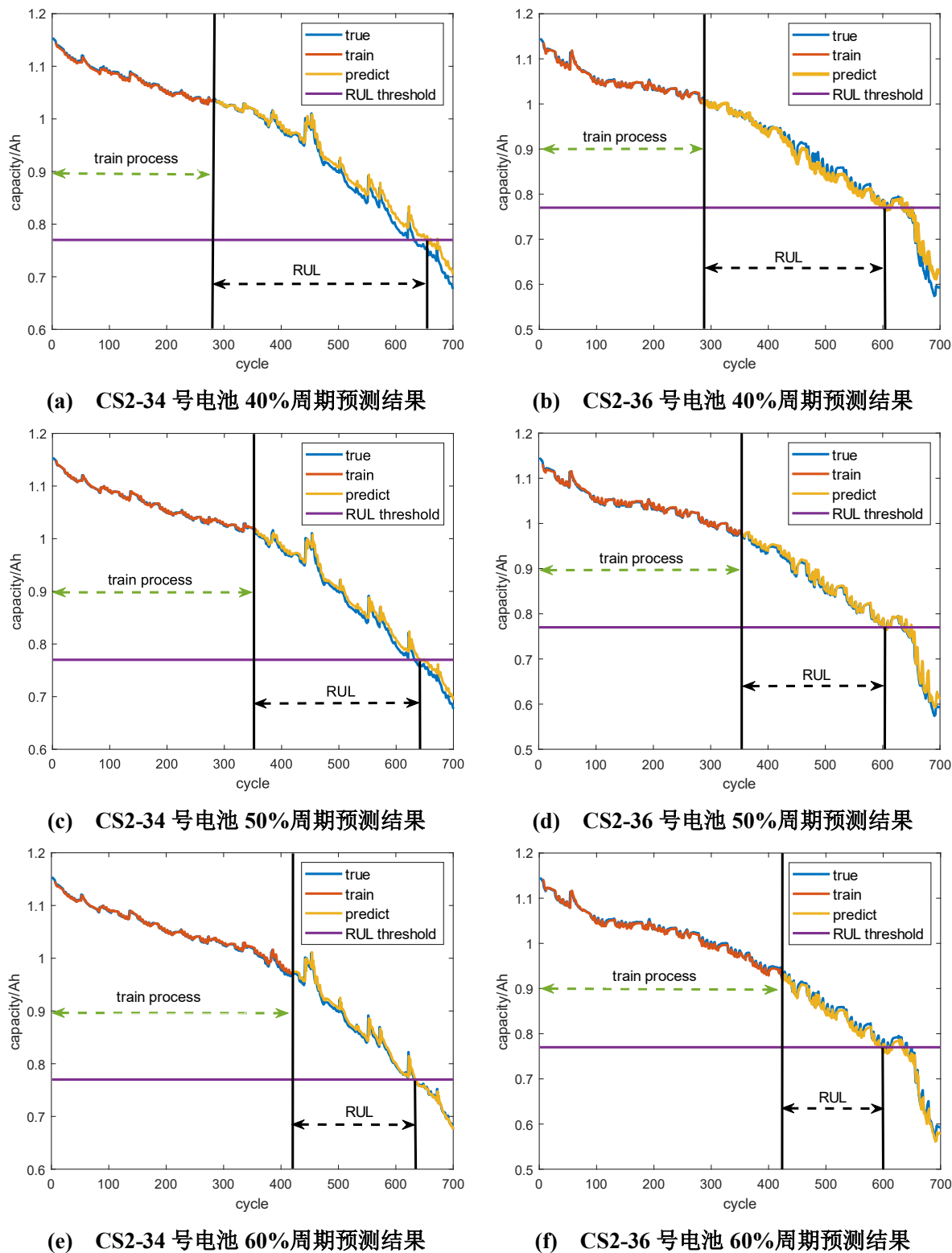


图 4.8 CALCE 数据集预测结果

图 4.8 展示了 CALCE 数据集的 CS2-34 号电池与 CS2-36 号电池的容量预测结果，

从容量预测拟合曲线来看,两电池的未来序列曲线与真实曲线均有良好的贴合效果,并且二者精度也是随着训练周期从 40%至 50%的增加而逐渐提高。60%周期时,34 号电池表现出最好的拟合效果,但在 50 周期时,36 号电池比 34 号电池有更好的表现,这是由于 34 号电池总体容量上比 36 号电池受到容量再生的影响较大。36 号电池在 50%周期情况下就已经展现出最好的拟合效果,这主要是由于 36 号电池的容量变化波动比 34 号电池更为规律。总体预测效果上,34 号电池优于 36 号电池,在 RUL 预测终止周期之前,34 号电池的容量曲线比 36 号电池更为平滑,因此误差迭代影响也相应较小。

从 NASA 数据集与 CALCE 数据集二者的预测拟合曲线都可以看出,从预测起点开始,模型可以很好地进行预测,预测结果曲线与真实曲线十分接近,很好地延续了训练时期的退化趋势特征。在每一个预测过程中,随着充放电循环次数的增加,后续的预测结果与真实值的误差逐渐变大,预测效果逐渐产生不良变化,这主要是受到了预测误差逐步累积的影响,在预测过程中每一步的预测结果都将受到上一步的误差影响,使得整个前向预测误差逐渐累积增大。这样的情况对于 RUL 终止周期较早电池的影响是较小的,例如 NASA 的 B05 号电池与 CALCE 的 CS2-34 号电池。容量再生、预测累计误差以及退化噪声等众多不可控因素使得锂电池剩余使用寿命的精准预测相较于其他类型的时间序列数据,其研究与学习过程更为困难和复杂。

4.5.2 实验对比分析

以 CALCE 数据集的 CS2-34 号电池的 60%训练周期为例,分别取出各模型预测结果中锂电池 RUL 预测终止点对应的周期,求出其剩余寿命误差,如表 4.2 所示。

表 4.2 CALCE 数据集 CS2-34 号电池不同模型 60%周期 RUL 终止点预测误差

模型	真实 RUL 终止点周期	预测 RUL 终止点周期	剩余寿命误差 E_{RUL}
CNN	632	678	46
LSTM	632	651	19
CNN-LSTM	632	646	11
CNN-LSTM-Attention	632	636	4

该 CS2-34 号电池真实 RUL 终止周期为 632,相较于单一的 CNN 网络、LSTM 网络和未加入注意力机制的 CNN-LSTM 网络,CNN-LSTM-Attention 网络基于 60%历史周期的剩余寿命误差 E_{RUL} 分别降低了 42、15 及 7,本课题提出的 CNN-LSTM-Attention 模型的 RUL 终止周期预测结果误差最低,说明该模型对于锂电池 RUL 预测具有更高的精

度，基于容量预测的锂电池 RUL 预测是有效可行的方法。在进行 RUL 计算时，只需将 RUL 终止周期预测值减去当前周期，即为当前锂电池剩余使用寿命。

以两种不同电池为例，图 4.9 和图 4.10 展示了 CNN、LSTM、CNN-LSTM 与本章提出的 CNN-LSTM-Attention 网络模型对于锂电池未来容量预测结果的误差对比图，图 4.9 为 NASA 数据集 B05 号电池关于不同周期的 RMSE 结果对比，图 4.10 是 CALCE 数据集 CS2-36 号电池关于不同周期的 MAPE 结果对比图。

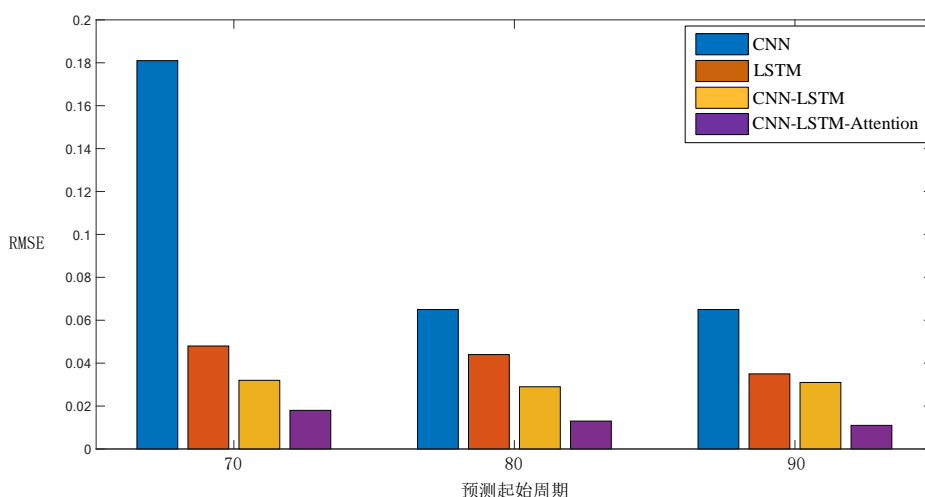


图 4.9 NASA B05 号电池不同预测模型不同周期 RMSE 对比统计图

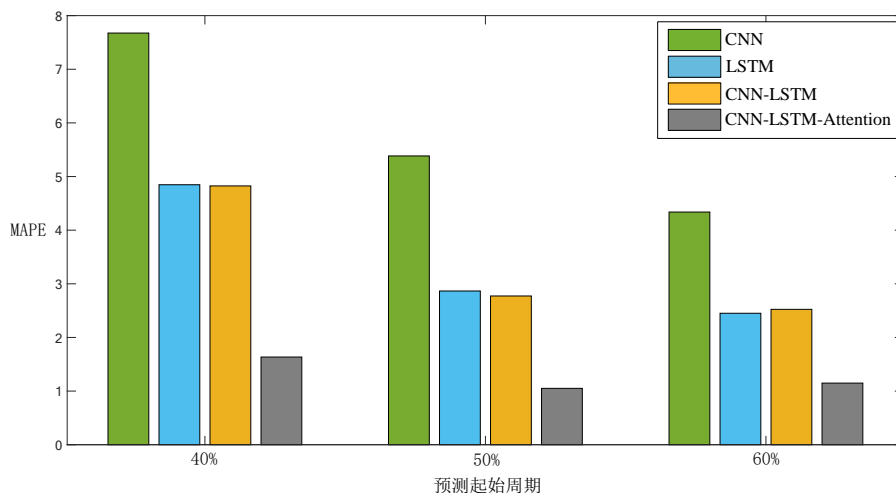


图 4.10 CALCE CS2-36 号电池不同预测模型不同周期 MAPE 对比统计图

从图 4.9 和图 4.10 中可以看出，与前三个模型对比，本课题提出的 CNN-LSTM-Attention 网络模型预测精度高，预测性能较好，未引入注意力机制的 CNN-LSTM 网络虽使得预测精度提高，但效果不显著，与其对比，加入了注意力机制后，模型预测精度提升较为明显，这说明注意力机制在预测模型中提取到了时间序列的关键时空特征，以

此使得预测结果更加准确。同时,可以观察到,在个别误差结果中呈现出来,CNN-LSTM 网络虽然使得预测精度提高,但误差降低幅度不够明显,说明在个别训练过程中,混合模型中一些非关键特征在一定程度上限制了预测精度的提高,这同时也说明了注意力机制 Attention 对于 CNN-LSTM 混合模型具有重要的优化作用。

从未来容量序列曲线拟合角度,可见 CNN 拟合结果在容量空间波动处较为敏感,但是在时间退化趋势的拟合中不够理想,而 LSTM 网络在时间退化趋势上拟合效果相较于 CNN 表现更好,预测精度得到大幅提高,但容量波动处的拟合结果较为平缓。CNN-LSTM-Attention 的未来容量序列的拟合结果不仅在时间退化趋势上进一步提升,在容量再生情况处的容量波动上也表现出良好的拟合效果,可见 CNN-LSTM-Attention 网络结合了 CNN 和 LSTM 的优势,使得总体预测精度相较于单一 CNN 和单一 LSTM 得到进一步的提高。

从各模型的预测结果来看,预测精度不光受到模型的影响,同时也受到训练数据数量的影响,预测精度与训练数据数量从总体角度成正比关系,并且在同比例数据数量增加过程中,后期精度提升空间低于前期提升空间,这主要是后期过程中迭代误差累积的缘故。个别情况下精度甚至随周期数的增加而降低,这说明对深度学习模型而言,训练数据过多有时不一定会使精度提高,因此要选取合适的数据量,例如对于 CALCE 数据集的 CS2-36 号电池,选取 50%训练周期时,预测效果相对较好。

4.6 本章小结

本章结合了 CNN 网络与 LSTM 网络的优点,提出了基于 CNN 与 LSTM 的混合网络,并在其中加入了注意力机制 Attention,构建了融合型的 CNN-LSTM-Attention 网络模型来实现锂电池 RUL 预测。首先使用滑动窗口按照设定步长将历史容量数据作为模型输入,利用 CNN 网络提取历史数据的空间特征,CNN 输出的特征向量作为 LSTM 网络的输入,利用 LSTM 来提取输入向量的时间特征,LSTM 保留了长序列数据间的依赖关系,最后引入注意力机制,以概率分配权重的方式赋予 LSTM 输出向量不同的权重,突出了重要信息的作用,并在网络中加入 Dropout 技术来防止深度学习模型的过拟合现象,提高了模型训练效率。经过实验验证,高精度的实验结果充分展示了 CNN-LSTM-Attention 网络对于锂电池 RUL 预测的有效性,分别从不同数据集、不同电池、训练集

大小等多个角度充分展示了该模型可以很准确地实现锂电池的剩余寿命预测，并且具有较强的泛化能力。通过对比 CNN-LSTM-Attention 模型与单独的 CNN-LSTM、CNN、LSTM 的实验结果误差，展示了融合型 CNN-LSTM-Attention 的先进性与鲁棒性。

第五章 基于 MC Dropout 与 CNN-LSTM-Attention 网络的锂电池 RUL 不确定性分析

融合型模型 CNN-LSTM-Attention 可以对锂电池的 RUL 实现精准预测,却无法得到预测结果的置信区间与不确定性表达,而贝叶斯统计方法可以解决不确定性表述问题。本章对电池容量类的时序数据进行不确定性的介绍,分析在电池 RUL 预测过程中产生不确定性的环节与原因;研究使用蒙特卡洛采样方法计算锂电池 RUL 概率分布的过程;对贝叶斯神经网络进行介绍,并且对其进行变分推导,研究其对于不确定性表达的理论支撑;利用贝叶斯网络的变分推断可由 Dropout 进行逼近这一技术优势,提出蒙特卡洛 Dropout(MC Dropout)的不确定性表达方式,结合本课题提出的 CNN-LSTM-Attention 模型,得到锂电池 RUL 预测结果的概率分布,经过实验结果以及不同模型对比分析,验证 MC Dropout 方法的有效性以及 CNN-LSTM-Attention 网络的鲁棒性与确定性。

5.1 基于时序预测的锂电池剩余使用寿命不确定性

锂电池的 RUL 预测过程产生的不确定性主要来源于从数学理论到现实建模逼近的环节、对于数学预测模型中的参数更新求解环节以及在样本数据的采集环节,那么对于锂电池剩余使用寿命预测的不确定性分析中,定性量化这些不确定性的影响,将会对锂电池 RUL 预测提供更加有力的现实应用价值。

对于锂电池的充放电循环过程,锂电池的不确定性的主要可能来源有以下几点:

(1)操作过程差异性。电池在不同的应用场景下按照不同方式运行,将会造成在退化过程中产生退化机制的差异。

(2)电池内部差异性。电池在制造过程中,由于其生产环节或者制作材料的不同导致的内部结构与参数产生差异性,因此在实际运行充放电过程中展现出不同的电池特性。

(3)监测手段的差异性。不同的监测方式或者检测传感器的不同及准确度差异将会对监测数据产生直接影响,例如电子干扰、数据噪声等,进而促进预测结果的不确定度。

(4)模型误差累积效应。

对于本课题所采用的时序预测模型,其主要不确定性来源于锂电池 RUL 预测起点的模型输入环节。从时间序列的研究角度出发,锂电池的老化过程通常需要经历一个比

较长的时间，整个过程将覆盖一个比较大的循环量，锂电池的容量衰减与循环次数是紧密相关的，时序数据经过迭代预测后的误差累积效应是较为显著的，因此电池性能退化模型在预测过程中的误差累积也会导致锂电池 RUL 预测结果的不确定性。

基于前文分析，由锂电池历史容量数据中固有的噪声导致不确定性是偶然不确定性，而由于预测模型的结构或者参数等问题导致的对于数据的认知缺乏所引入的不确定性是认知不确定性，在深度学习领域，贝叶斯深度学习框架可以对这两种不确定进行建模。

在本课题基于时序预测方法的锂电池 RUL 建模过程中，由于缺乏对电池内部退化机制的详细了解，以及无法获取历史容量数据对于电池内部差异及操作环境的差异性，因此主要通过预测模型角度来对锂电池的 RUL 不确定性进行研究和分析，具体过程是将前文中建立的混合模型与深度学习中常见的 Dropout 技术结合起来，使用蒙特卡洛采样方式来获取锂电池 RUL 结果的概率分布，通过概率分布来进行 RUL 不确定性的度量分析，以此来验证模型的精确性与鲁棒性。

5.2 基于蒙特卡洛模拟的电池 RUL 概率分布

蒙特卡洛(Monte Carlo,MC)方法在不确定性研究中，通常与不同的预测方式相结合来计算预测结果的概率分布，其核心思想是以概率为算法基础，对于样本进行重复性的随机试验来进行复杂过程的数值计算^[56]。基于蒙特卡洛模拟的锂电池 RUL 概率分布的计算过程如下所示：

(1)定位锂电池 RUL 不确定性产生的环节并进行量化。对于时间序列数据驱动的预测模型，需要量化模型建立环节中产生的不确定性因素对于锂电池 RUL 结果的影响，在一定情况下也需要结合多种不同因素的共同影响，对于本课题提出的模型，可以从网络层中存在的随机性 Dropout 环节进行考虑。

(2)对于不确定性来源因素的初始分布进行量化确定。对于前文中提到的经验预测方法，在寿命经验模型中数据拟合环节中的不确定性分布规律可以由模型参数的均值、方差进行量化确定；与经验预测法类似，对于滤波预测法，在模型状态向量初始化的环节中，可以根据寿命经验模型拟合历史数据的参数均值与方差来确定不确定性分布规律；对于本课题选用的时序预测法，可以根据预测起点附近的历史容量数据分布规律来对起点样本采集环节不确定性分布规律进行量化确定。

(3)使用 MC 来模拟初始不确定性分布的传播。根据第(2)步所获得的不确定性因素的分布规律来随机生成模拟化样本数据,然后使用生产的样本数据作为输入,根据相对应的预测模型来产生模型预测结果进行输出,获得数量一致的若干 RUL 预测结果数据。

(4)计算第(3)步得到的 RUL 预测结果的概率分布。对于锂电池 RUL 的概率密度函数计算过程如式(5.1)所示:

$$\hat{f}_n(C_{\max}) = \sum_{i=1}^N \left[K_p \left(\frac{C_{\max} - C_{\max,i}^-}{h_p} \right) + K_p \left(\frac{C_{\max} - C_{\max,i}}{h_p} \right) + K_p \left(\frac{C_{\max} - C_{\max,i}^+}{h_p} \right) \right] \quad (5.1)$$

式(5.1)中, $\hat{f}_n(C_{\max})$ 代表锂电池 RUL 预测结果概率密度函数, $K_p(\cdot)$ 代表高斯核函数, h_p 代表带宽, $C_{\max,i}^-$ 与 $C_{\max,i}^+$ 计算过程式(5.2)所示:

$$\begin{cases} C_{\max,i}^- = 2L_c - c_{\max,i} \\ C_{\max,i}^+ = 2U_c - c_{\max,i} \end{cases} \quad (5.2)$$

式(5.2)中, U_c 与 L_c 分别为蒙特卡洛模拟结果的上下界, $C_{\max,i}$ 为 i 个 RUL 预测结果。

5.3 基于贝叶斯方法的预测结果不确定性

对于深度学习领域中不确定性的研究,经常使用贝叶斯神经网络来进行解决。鉴于传统神经网络的权值固定这一特点,贝叶斯神经网络则是通过将网络权重由确定的数值变成某种分布,通过变分推断的方式来得到模型权重的后验分布,从而可以提供对神经网络模型的概率解释^[68]。

5.3.1 贝叶斯神经网络的概率模型

设定一个神经网络模型的输入为 x ,输出结果为 y , w 为神经网络中的权值,如果将这个神经网络的模型设为服从条件概率分布的模型 $P(y|x, w)$,如果待解决的问题是分类问题,则输出结果是分布对应的结果类型的概率,倘若待解决的问题是回归结果值输出,则此结果值为这个分布对应所有结果的均值。将训练数据集设为 D ,对于这个分布对应的神经网络的学习过程可以视为一个最大似然估计(MLE)的过程:

$$w^{\text{MLE}} = \operatorname{argmax}_w \log P(D|w) = \operatorname{argmax}_w \sum_i \log P(y_i | x_i, w) \quad (5.3)$$

在分类问题之中,需要通过交叉熵来得到结果的输出,那么在模型之中需要代入相应的逻辑运算函数,而在回归问题中,需要通过均方误差 MSE 来逐步逼近得到输出值,

在模型之中需要带入正态分布,对于权值等参数的更新主要使用的还是主流的反向传播下的梯度下降法。对于最大似然估计,则是先将 w 取不同值的事件视作相同概率事件,因此对于权值 w 的先验估计是没有必要的。但如果为权值 w 提前设定好先验估计,则可得最大后验估计(MAP):

$$w^{MAP} = \operatorname{argmax}_w \log P(w|D) = \operatorname{argmax}_w \log P(D|w) + \log P(w) \quad (5.4)$$

在贝叶斯估计方法之中,首先对网络权重参数 w 设定正态分布 $N(\mu, \sigma)$,那么贝叶斯神经网络的训练过程之中,更新调整的参数即变为 μ 和 σ ,同时参数的数量就变为 $2n$ 。对贝叶斯神经网络的参数引入先验假设,为了完成对贝叶斯网络的结果进行不确定性度量,将 w 的后验分布进行求解,后验概率为:

$$P(w|D) = \frac{P(D|w)P(w)}{P(D)} \quad (5.5)$$

式(5.5)中, $P(D|w)$ 代表似然概率,是数据集 D 基于观测变量 w 的概率; $P(w)$ 为先验概率,表征矩阵变量 w 在模型训练前的概率分布; $P(D)$ 是对所有潜在变量进行求和之后的积分,即:

$$P(D) = \int P(D|w)P(w)dw \quad (5.6)$$

但是上述的 $P(D)$ 过程是个非常复杂的积分过程,难以对其进行求解,则需要更有效的算法来对上述后验概率实现近似化推导。

5.3.2 基于变分推断的贝叶斯神经网络

贝叶斯网络的变分推理首先设定一种中间过程的概率分布 $q(w; \theta)$,用变分参数的方式来对后验分布 $P(w|D)$ 进行近似,为了使得计算简单,首先假设贝叶斯变分后验分布服从的是正态分布,那么设定的变分向量中参数为:

$$\theta = (\mu, \rho) \quad (5.7)$$

式(5.7)中, μ 为均值, ρ 是一个向量,其为正向量,其功能是主要用于产生后续的标准差,有 $\sigma = \ln(1 + e^\rho)$ 。这样的话就使得求解后验分布 $P(w|D)$ 转变为了求解变分参数 θ ,然后利用求解两个分布最小KL散度的方式来求解变分参数 θ ,并且在此定义损失函数 F ,过程如下所示:

$$\begin{aligned}\theta^* &= \operatorname{argmin}_{\theta} D_{KL} [q(w|\theta) \| P(w|D)] \\ &= \operatorname{argmin}_{\theta} \int q(w|\theta) \log \frac{q(w|\theta)}{P(w)P(D|w)} dw\end{aligned}\quad (5.8)$$

$$= \operatorname{argmin}_{\theta} D_{KL} [q(w|\theta) \| P(w)] - E_{q(w|\theta)} [\log P(D|w)]$$

$$F(D|\theta) = D_{KL} [q(w|\theta) \| p(w)] - E_{q(w|\theta)} [\log P(D|w)] \quad (5.9)$$

式(5.8)和(5.9)可见,对于 θ^* 的计算中,第一项主要是两个分布的 KL 散度的表达式,而第二项表征了取值是依赖于训练数据 D 的,在此处,将 $P(w)$ 设定为两个正态分布的叠加计算,这两个正态分布的均值和标准差均不相同,即:

$$P(w) = \prod_j \pi N(w_j | 0, \sigma_1^2) + (1 - \pi) N(w_j | 0, \sigma_2^2) \quad (5.10)$$

在目标函数 F 中,因为难以求解其中包含的期望,则可对其进行蒙特卡洛近似化,即:

$$F(D, \theta) \approx \sum_{i=1}^n \log q(w^{(i)} | \theta) - \log P(w^{(i)}) - \log P(D | w^{(i)}) \quad (5.11)$$

式(5.11)中, $w^{(i)}$ 代表蒙特卡洛对第 i 个数据点进行采样的权值结果,进行高斯变换,可对单位化的高斯变化量进行相对应的缩放与平移操作,可得 w 样本:

$$w = \mu + \epsilon * \ln(1 + e^{\rho}) \quad (5.12)$$

式(5.12)中, $*$ 代表按元素设置的乘法操作, ϵ 代表单位化的高斯变量,其服从 $\epsilon \sim N(0, I)$ 。

通过高斯变分推理,就可以利用权值的后验分布 $P(w|D)$ 对贝叶斯神经网络进行推理。通过对后验分布输出结果进行平均化处理来达到对后验分布的充分表达,以此产生整个网络的结果输出,即输出 y 为:

$$y = E_{p(w|D)} [g(x_0, w)] \quad (5.13)$$

式(5.13)中, $g(x_0, w)$ 代表网络函数, x_0 代表网络输入。整个推导过程是基于变分推理的,则可以使用之前设定的参数 (μ, ρ) 控制的概率分布 $q(w; \theta)$ 来进行对后验分布 $P(w|D)$ 的近似,那么神经网络的输出则表示为:

$$y \approx E_{q(w; \theta)} [g(x_0, w)] \quad (5.14)$$

对于上面的输出 y ,在进行蒙特卡洛采样处理,可将输出进一步近似化:

$$y \approx \frac{1}{N} \sum_{i=1}^N g(x_0, w^i) \quad (5.15)$$

式(5.15)中, w^i 代表从变分后验分布之中蒙特卡洛采样得到的第 i 个样本数据, N 代表样本总数, 网络估计输出可进一步表示为:

$$y \approx \frac{1}{N} \sum_{i=1}^N g(x_0, \mu + \sigma^* \epsilon^i) \quad (5.16)$$

式(5.16)中, ϵ^i 代表通过蒙特卡洛采样得到的第 i 个单位高斯变量。在贝叶斯神经网络之中, 需要对每一个权重进行一个单位高斯变量的相应采样, 基于以上分析和推理, 贝叶斯神经网络可进行与传统神经网络相同的推理运算和学习过程。在寿命预测研究中, 贝叶斯变分推断常被引入来获得剩余寿命的不确定性。

5.4 近似贝叶斯推理的 MC Dropout 算法

Damianou 等人提出在神经网络层中引入 Dropout, 就相当于是一个高斯过程^[69]。而根据前文的理论推导研究, 贝叶斯变分推断求解过程复杂, 并且需要对常规神经网络进行结构上的大改进才能获得 RUL 预测的不确定性。Y.Gal 等人提出含有 Dropout 的神经网络等价于对神经网络参数进行变分推断, 结合蒙特卡洛采样可以得到预测结果的不确定性, 对于推导过程做了详尽的说明^[70]。

5.4.1 Dropout 的随机性

在上一章已经对 Dropout 技术进行了详细的介绍, 在混合模型中引入 Dropout 不仅是为了防止过拟合现象的发生和提高训练效率, 同时也是为本章的锂电池剩余使用寿命不确定性分析提供理论基础。

Dropout 指的是在神经网络中随机性的丢弃一些神经元, 即暂时性删除该神经元, 如图 5.1 所示, 这些神经元的遗弃是随机的, 每一个神经元在一次迭代过程中都有被丢弃的可能, 被丢弃的神经元在前向传播过程不会起作用, 在预测过程中, 模型会根据 Dropout 概率设置调整模型的权重参数。Dropout 的加入, 在一定程度上避免了深度学习过程中常见的过拟合现象, 同时也是减弱了部分权重对于整个网络的过度影响, 提高模型的泛化能力。

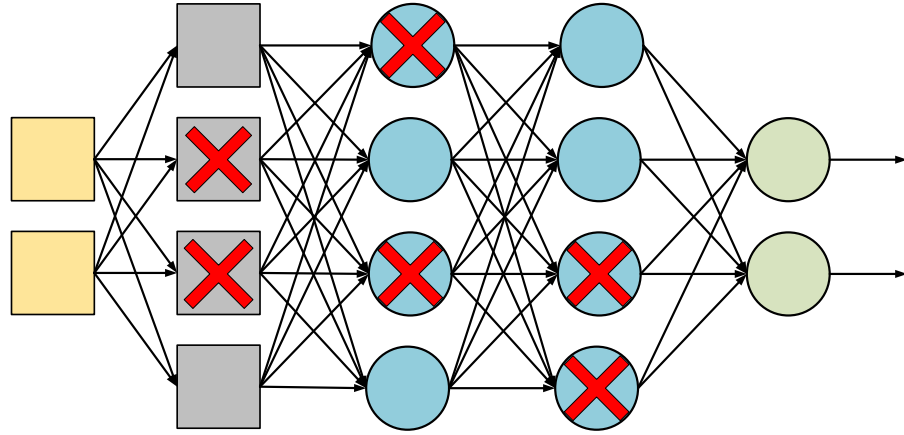


图 5.1 神经元的随机删除

在神经网络结构加入 Dropout，相当于是对原有的神经网络进行了一定程度的简化操作。假设现有 n 个神经元，则网络结构相当于包含有 2^n 个简化神经网络的集合，这些神经网络之间是共享参数的，由此减少了整个平均化的大型复杂网络的结构参数量。在测试阶段，应用了 Dropout 机制的神经网络将通过较为简易的近似化方式来得到预测结果，每一个神经元的权重参数将根据 Dropout 的概率进行比例放缩，从而使得神经元的期望输出值是相对于实际输出值的无偏估计。

5.4.2 MC Dropout 算法的近似贝叶斯过程

首先使用非线性化的映射函数来设定一个协方差函数，其中的非线性函数可以使用 ReLU 函数或者 tanh，此协方差函数为：

$$K(x, y) = \int P(w) P(b) \sigma(w^T x + b) \sigma(w^T y + b) dw db \quad (5.17)$$

式(5.17)中， $P(w)$ 代表多维正态分布， $P(b)$ 代表一维标准正态分布。通过 K 此采样的蒙特卡洛积分，可得到一个有限秩协方差函数：

$$\hat{K}(x, y) = \frac{1}{K} \sum_{k=1}^K \sigma(w_k^T x + b_k) \sigma(w_k^T y + b_k) \quad (5.18)$$

式(5.18)中， w_k 满足于 $P(w)$ 分布， b_k 满足于 $P(b)$ 分布，参数 K 可视作神经网络中隐藏层之中的节点数目。对上面的协方差函数进行整合，可得如下的预测分布：

$$P(Y|X) = \int p(Y|W_1, W_2, X, b) p(W_1) p(W_2) p(b) \quad (5.19)$$

基于以上过程，可获得基于参数可调的高斯过程模型，进而可对其进行后验分布的近似变分推断。对参数 W_1 、 W_2 及 b 进行基于高斯过程混合分布的建模，进而根据如下所

示过程的 ELBO 过程进行推断：

$$ELBO = \int q(W_1, W_2, b) \log(p(Y|W_1, W_2, X, b)) - KL(q(W_1, W_2, b) \| p(W_1, W_2, b)) \quad (5.20)$$

对于 ELBO 的推断过程此处不再进行详细赘述，使用蒙特卡洛方式积分过程，可获得如下的 ELBO：

$$ELBO_{MC} = \sum_{n=1}^N \log(p(y_n | x_n, \hat{W}_1^n, \hat{W}_2^n, \hat{b}^n)) - KL(q(W_1, W_2, b) \| p(W_1, W_2, b)) \quad (5.21)$$

式(5.21)中， \hat{W}_1^n 、 \hat{W}_2^n 及 \hat{b}^n 各自代表对伯努利分布和多维正态分布的线性组合通过采样得到的多维向量。综合上述推导，ELBO 关于样本数据和相关分布之间的关系如式(5.22)所示：

$$ELBO_{MC} \propto -\frac{1}{N} \sum_{n=1}^N \|y_n - \hat{y}_n\|_2^2 - \frac{p_1}{\tau N} \|M_1\|_2^2 - \frac{p_2}{\tau N} \|M_2\|_2^2 - \frac{p}{\tau N} \|M\|_2^2 \quad (5.22)$$

式(5.22)表明，在神经网络之中传递的含有服从伯努利分布的变量，基于此，对于后验分布的变分推断就相当于在神经网络中加入 Dropout。

通过近似变分推断，使得贝叶斯网络的变分推断可由普通神经网络加入 Dropout 来近似，这样可以在不改变原本神经网络结构的前提下近似构建贝叶斯网络，从而进一步获得 RUL 预测值的不确定性，同时 Dropout 也可以避免神经网络模型的过拟合现象。

5.5 基于 MC Dropout 与 CNN-LSTM-Att 模型的 RUL 不确定性建模

根据之前章节对于神经网络的相关介绍，鉴于目前已有的相关技术与设备基础，考虑到随机过程验证算法以及贝叶斯方法理论关于不确定性分析的复杂性和技术难度，在综合考量之下，本课题使用神经网络中的 Dropout 方法来近似贝叶斯神经网络的建模过程，并且使用蒙特卡洛采样方式来近似锂离子电池剩余使用寿命的概率分布，以此完成对锂电池剩余使用寿命的不确定性的量化研究，实现过程如图 5.2 所示。

图 5.2 展示了通过神经网络中加入 Dropout 技术实现对 RUL 不确定性分析的概率分布建模过程。在预测模型的训练阶段，可使用历史样本数据进行拟合得到一个拥有足够表达性的预测模型。在模型的训练过程中，神经网络隐藏层之中的 Dropout 可以在一定程度上避免神经网络常见的过拟合问题，模型训练完成之后，在模型预测阶段，打开 Dropout 结构设置，从而得到锂电池未来容量序列结果，进而进行多次反复迭代，得到

最终的容量序列预测结果。

根据前文分析,通过 MC Dropout 方式获得的 RUL 预测值服从高斯分布^[71]。使用样本均值和样本标准差分别代表该预测过程获得的 RUL 预测值分布的 μ 与 σ ,则可得 RUL 预测值的 95%置信区间如式(5.23)所示:

$$P(\mu - 1.96 \frac{\sigma}{\sqrt{n}} \leq \bar{y} \leq \mu + 1.96 \frac{\sigma}{\sqrt{n}}) \approx 0.95 \quad (5.23)$$

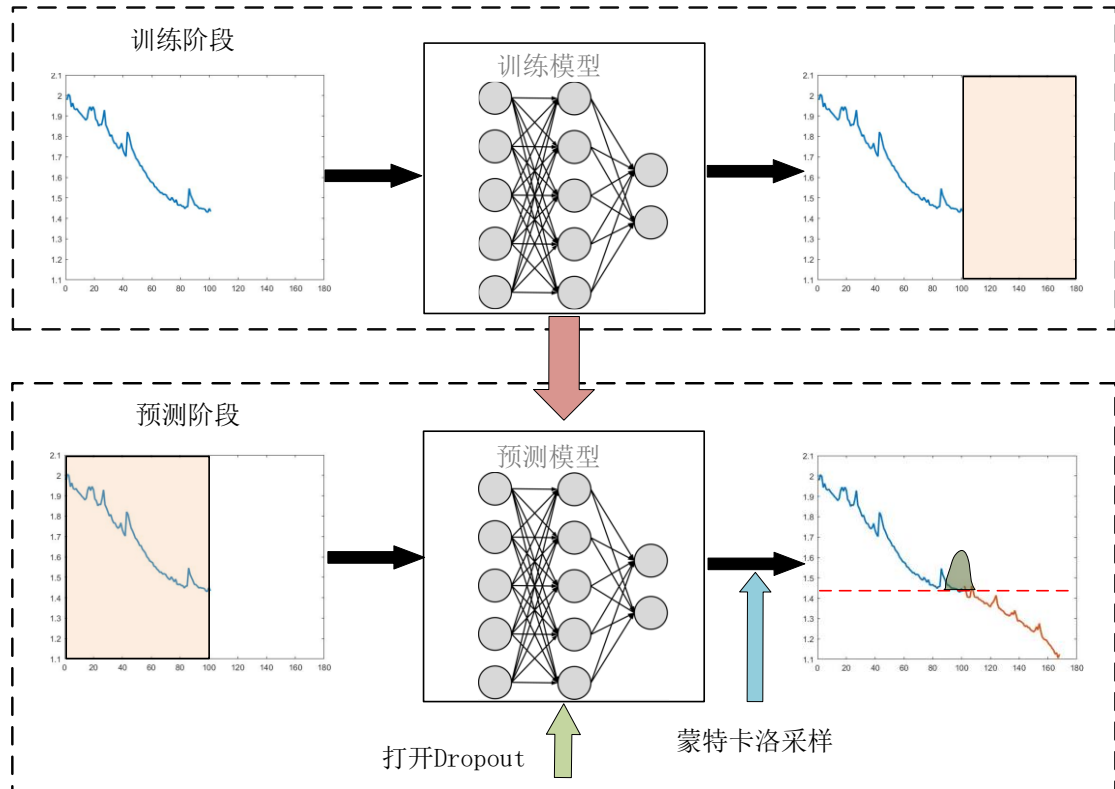
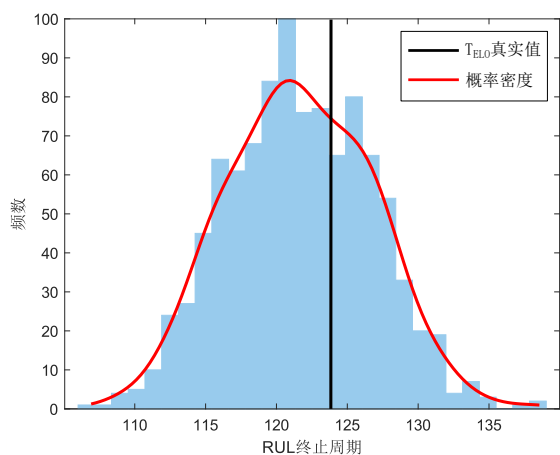


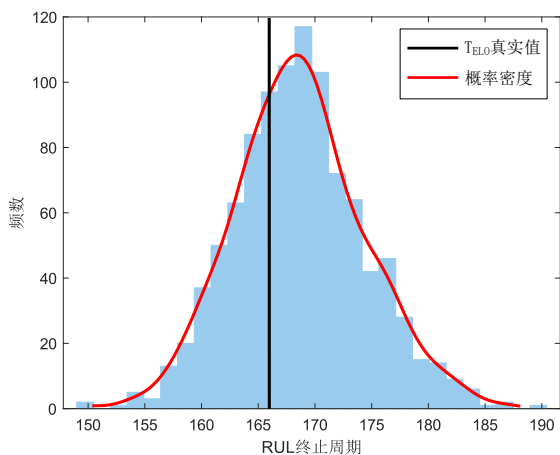
图 5.2 MC Dropout 方法工作过程

5.6 实验结果与分析

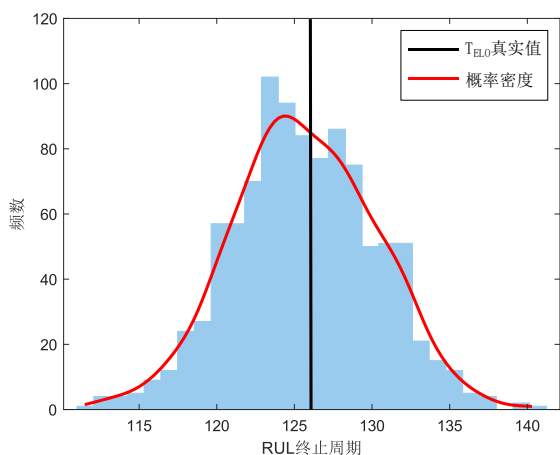
基于上一章提出的 CNN-LSTM-Attention 网络锂电池 RUL 预测模型,对本章实验在预测阶段打开网络层中的 Dropout 设置,使用在训练阶段相同的 Dropout 概率设置,以此来对其中的神经元进行随机性的保留与丢弃,来实现未来时间序列的容量预测。对预测结果根据蒙特卡洛采样方式进行 1000 次的采样,从而可得 1000 份的基于 CNN-LSTM-Attention 模型的容量预测结果样本,基于本课题设定的 RUL 终止点(额定容量的 70%)时的周期数,从 1000 份容量预测结果中提取出 RUL 终止周期预测值,共 1000 个 RUL 终止周期预测样本,建立 RUL 终止周期的概率分布模型。



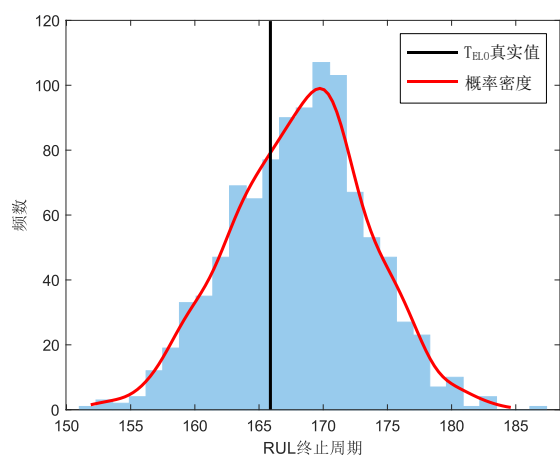
(a) B05 电池 70 周期 RUL 终止周期预测分布



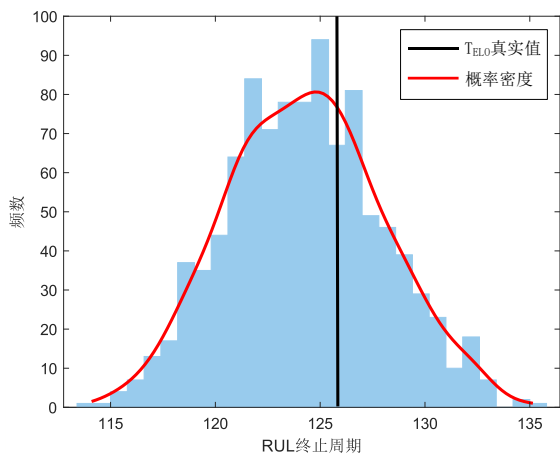
(b) B07 电池 70 周期 RUL 终止周期预测分布



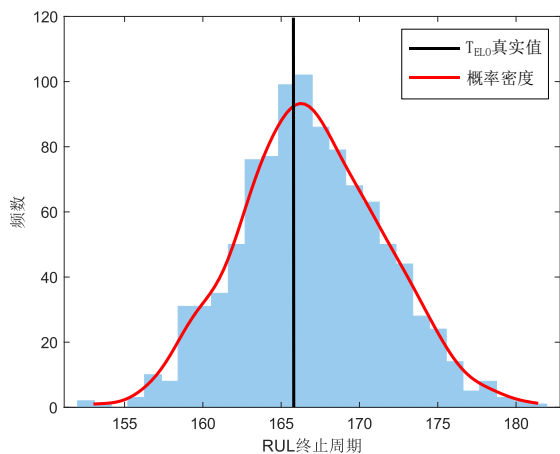
(c) B05 号电池 80 周期 RUL 终止周期预测分布



(d) B07 电池 80 周期 RUL 终止周期预测分布



(e) B05 号电池 90 周期 RUL 终止周期预测分布



(f) B07 电池 90 周期 RUL 终止周期预测分布

图 5.3 NASA 数据集 RUL 终止周期预测结果概率分布

图 5.3 为使用 NASA 数据集的 B5 号电池与 B7 号电池，分别以前 70、80 及 90 周期为训练集，也同样为预测起点而得到的基于 1000 个样本的 RUL 终止周期概率分布模型。红色曲线代表预测结果概率密度函数，黑色直线代表 RUL 终止周期真实值。根据得到的两个电池的 RUL 终止周期概率分布可知，通过蒙特卡洛采样得到的 RUL 终止周

期近似分布于正态分布，充分验证了在上节中证明的原理：在含有 Dropout 机制的神经网络中，通过 Dropout 服从伯努利分布的原理建立近似贝叶斯网络，得到的神经网络等价于高斯过程的模型。图 5.4 为基于 NASA 数据集 B5 号电池 90 周期的预测结果置信区间。

表 5.1 NASA 数据集剩余寿命概率分布统计量

电池	真实值	周期数	中位数	均值	标准差	方差
B5	124	70	121	122	5.13	26.28
		80	125	126	4.61	21.27
		90	124	124	3.68	13.54
B7	166	70	168	169	5.76	33.14
		80	167	168	5.32	28.29
		90	166	167	4.60	21.19

通过两电池的 RUL 终止周期概率分布情况，可见相较于 70 周期，80 周期与 90 周期更加集中分布于真实 RUL 终止周期周围。并且可以看出，对于多数分布结果，大部分的 RUL 终止周期预测值都分布于真实 RUL 终止周期之前。根据 RUL 终止预测周期分布统计表 5.1 所示，不同电池不同预测起点的电池，其 RUL 终止周期分布的中位数与均值都十分接近真实值，展现出本课题提出的 CNN-LSTM-Attention 预测模型的优越性，同时标准差与方差随着训练周期的增加而减小，从这个角度也能展现出训练样本的数据量对于模型预测效果的影响，同时也表明，数据量较少的情况下模型将具有更大的不确定性。

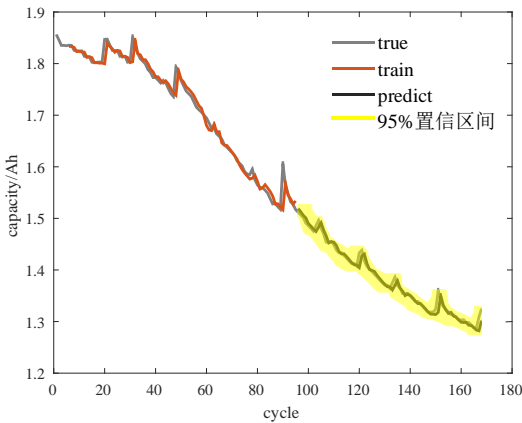


图 5.4 CNN-LSTM-Att 模型 B5 号电池置信区间

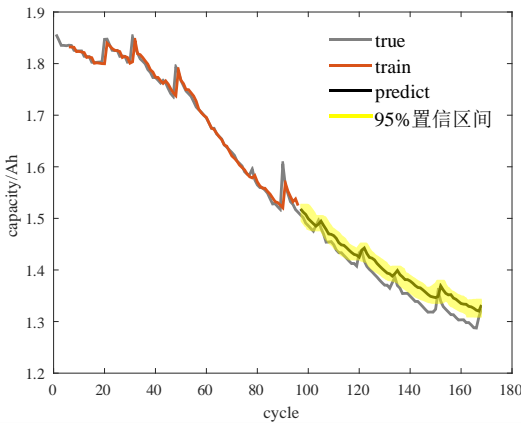


图 5.5 CNN-LSTM 模型 B5 号电池置信区间

为验证不同预测模型之间的不确定性差异，基于上一章提出的 CNN-LSTM 网络同样含有 Dropout，对其进行跟 CNN-LSTM-Attention 网络相同的 MC-Dropout 不确定性

建模过程。

对于不确定性建模的 CNN-LSTM 网络，使用 CALCE 数据集 CS2-34 号电池的容量数据，即在预测阶段打开同训练阶段相同的 Dropout 设置进行蒙特卡洛采样，得到 1000 份 RUL 终止周期预测样本，得到如图 5.6 所示的 RUL 终止周期预测结果的概率分布模型。对 NASA 数据集 B05 号电池进行相同操作，预测周期为 90 周期，得到图 5.5 所示置信区间。

根据 CNN-LSTM 网络和 CNN-LSTM-Attention 网络基于 B05 号电池 90 周期的置信区间，可以看出，真实容量预测值基本都落于置信区间内，相较于点估计值，置信区间则可以给出预测值的可信程度。

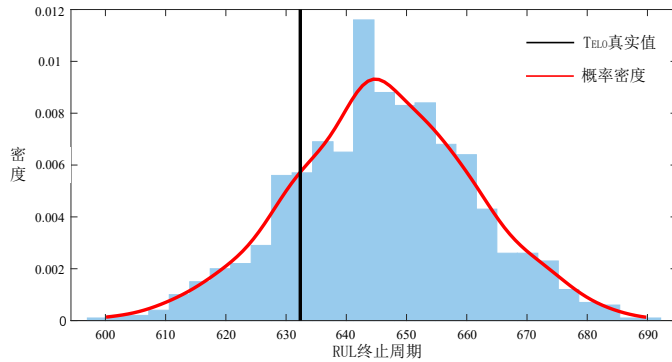


图 5.6 CS2-34 号电池 CNN-LSTM 模型 60%周期 RUL 终止周期预测分布

根据 CS2-34 号电池的 RUL 终止周期预测结果概率分布图 5.6 所示，其分布也近似于正态分布，同样验证了前文的推理过程。CNN-LSTM-Attention 模型与 CNN-LSTM 模型对于 CS2-34 号电池 60%预测周期的 RUL 终止周期的概率分布统计如表 5.2 所示。

表 5.2 CALCE 数据集 CS2-34 号电池基于不同模型 60%周期 RUL 终止周期概率分布统计量

模型	真实值	中位数	均值	标准差	方差
CNN-LSTM	632	645	646	15.07	227.24
CNN-LSTM-Attention	632	634	636.01	5.02	25.25

根据前文分析，锂电池 RUL 预测结果的不确定性主要来源于偶然不确定性与认知不确定性，偶然不确定性主要指数数据噪声的影响，这个影响对于不同的模型其不确定性是相差不大的，而认知不确定性主要是模型结构与参数带来的不确定性，则本实验环节不同模型的预测结果不确定性对比表征了模型之间的不确定性和稳定性差异。

根据分布图 5.6 与统计表 5.2，相较于 CNN-LSTM-Attention 网络的结果分布，CNN-LSTM 网络 60%周期的 RUL 概率分布的中位数和均值都与真实终止周期具有较大偏差，

其方差远大于前者，这也表明本课题提出的 CNN-LSTM-Attention 预测模型具有更强的稳定性和更高的准确度，未含有注意力机制的混合网络具有更高的不确定度。

5.7 本章小结

本章首先对基于锂电池 RUL 预测结果的不确定性相关理论进行介绍，分析了在锂电池 RUL 预测过程中产生不确定性的因素和环节，并提出使用蒙特卡洛方式对锂电池 RUL 预测结果进行概率分布的计算过程；通过对贝叶斯神经网络的原理与统计方法介绍，对其参数后验概率进行建模，研究其推导过程，得到贝叶斯网络通过变分推断后的输出为概率分布的结论，能表达锂电池 RUL 预测结果的不确定性；但基于贝叶斯网络需要对本课题提出的融合型网络进行权重的较大改进，并且将引入大量计算过程，本章提出了 MC Dropout 近似贝叶斯推理的方法来实现 RUL 预测结果的不确定性建模，通过变分推断，使得贝叶斯网络中的变分推断可由神经网络结构中加入的 Dropout 来进行近似，从而避免了巨大计算开销，另一方面 Dropout 也可避免深度学习网络中的 Dropout 现象。实验结果表明，本章提出的 MC Dropout 方法只需在 CNN-LSTM-Attention 网络建模过程中引入 Dropout 便可获得类似于贝叶斯网络的概率分布结果，并且可以给出 RUL 预测结果的置信区间，得到不确定性的表达，同时通过 CNN-LSTM-Attention 模型与其他模型进行不确定性概率分布的对比，展示了 MC Dropout 方法的有效性，验证了本课题提出的 CNN-LSTM-Attention 模型在泛化能力与稳定性方面优于其他模型，具有工程实用价值。

总结与展望

锂电池的剩余使用寿命预测对于电池管理系统有着至关重要的作用,本文使用基于驱动型的深度学习融合模型进行锂电池 RUL 预测研究。首先介绍了本课题的研究背景与意义,从不同的 RUL 预测方法入手,介绍国内外各专家学者的研究成果,对现有的 RUL 预测方法研究现状进行分析,从而明确了基于数据驱动与融合技术的研究方式,确定了深度学习为模型主体。本文所做的主要工作和研究内容为:

(1)从锂电池结构与工作过程出发,分析了锂电池老化机制与退化因素,介绍电池管理系统对于锂电池的重要意义,同时表明有效锂电池 RUL 预测作为 BMS 重要组成部分的重要研究意义。从历史容量数据驱动型的锂电池 RUL 预测流程出发,重点研究了基于时序数据的预测算法,分析锂电池 RUL 预测结果不确定性研究的重要性。

(2)基于深度学习模型原理,分别提出了基于卷积网络 CNN 与长短期记忆 LSTM 网络的两种锂电池未来容量序列预测模型,通过容量预测实现锂电池 RUL 的预测。通过滑动窗口按照设定输入历史容量数据,分别使用 CNN 与 LSTM 两种网络根据锂电池的历史容量来对未来容量序列做出预测,根据寿命终止周期来获得锂电池的 RUL 预测值。两种模型的实验结果表明,基于数据驱动的深度学习方法在锂电池的 RUL 预测中具有可行性,并且 LSTM 网络模型比 CNN 模型更适合处理锂电池容量类的时序数据。

(3)为进一步提高锂电池 RUL 预测精度,提出了一种基于混合模型与注意力机制的融合型网络 CNN-LSTM-Attention。该模型融合了 CNN 与 LSTM 的优势,通过滑动窗口输入数据后,首先使用 CNN 网络结构对锂电池历史容量数据进行空间特征提取,将 CNN 输出的特征向量输入至 LSTM 结构中来提取其中的时间特征,最后将 LSTM 的输出传递注意力层中,捕捉其中的关键时空特征,从而得到最终的预测结果,通过失效阈值对应的终止周期来获得 RUL 值。并且在网络中加入 Dropout,避免深度学习模型中的过拟合现象,同时为后续的不确定性分析奠定基础。经过实验结果的对比与分析,本文提出的融合型网络 CNN-LSTM-Attention 在锂电池 RUL 预测中拥有出色的表现,并且与单一的 CNN 模型、LSTM 模型及未引入注意力机制的 CNN-LSTM 进行对比,本文提出的融合型模型就有较高的精度与泛化能力。

(4)为获得锂电池 RUL 预测结果的不确定性,提出了一种通过 MC Dropout 方法结

合 CNN-LSTM-Attention 网络的不确定性描述方法。基于经过变分推断的贝叶斯网络可以给出预测结果的概率分布，在网络模型中引入 Dropout 技术来逼近贝叶斯网络，使用蒙特卡洛 MC 采样方式，从而获得锂电池 RUL 预测结果的置信区间与概率分布，实现不确定性的描述与分析，通过实验结果的对比与分析，验证了 MC Dropout 与融合模型结合的方式在锂电池 RUL 模型不确定性分析的有效性，同时表明本文提出的 CNN-LSTM-Attention 网络相较于其他网络具有更好的稳定性与确定性。

由于个人能力有限，在之后的研究中，还将致力研究以下问题：

(1)基于数据集规模限制，今后将致力于研究借助更大规模数据集的网络与算法，对预测模型进行更多的调整与优化以面对 BMS 的 SOC 等其他指标的监测以及更大更多种类的数据集。

(2)对于深度学习模型在锂电池 RUL 预测结果不确定性表达，今后将继续研究，实现更多种类的不确定性建模方式，使不确定性建模方式更加简便。

(3)目前的研究主要是基于给定数据进行 RUL 研究，今后将致力于更加贴近现实中应用与实验环境的预测方式，解决更多元化的问题。

参考文献

- [1] 徐瑞琳, 曾涛, 刘欢, 等. 磷酸铁锂电池循环初期衰减快原因分析及性能改善 [J]. 无机盐工业, 2023, 55(03): 92-97.
- [2] 中共中央关于制定国民经济和社会发展第十四个五年规划和二〇三五年远景目标的建议 [J]. 中国民政, 2020, No.690(21): 8-21.
- [3] 国务院办公厅关于印发新能源汽车产业发展规划(2021—2035 年)的通知 [J]. 中华人民共和国国务院公报, 2020, No.1714(31): 16-23.
- [4] 来鑫, 孟正, 韩雪冰, 等. 基于特征电压模型的锂离子电池容量估计与 RUL 预测 [J]. 汽车安全与节能学报, 2022, 13(1): 194-201.
- [5] 刘伯峥, 王静波, 曾涛, 等. 磷酸铁锂电池寿命初期与末期安全性差异 [J]. 化工学报, 2022, 73(12): 5555-5563.
- [6] Wei J, Dong G, Chen Z. Remaining Useful Life Prediction and State of Health Diagnosis for Lithium-Ion Batteries Using Particle Filter and Support Vector Regression [J]. IEEE Transactions on Industrial Electronics, 2018, 65(7): 5634-5643.
- [7] Yong Z, Zidong W, Alsaadi F E. Detection of intermittent faults for nonuniformly sampled multi-rate systems with dynamic quantisation and missing measurements [J]. International Journal of Control, 2018: 1-23.
- [8] Ran L, Wu J, Wang H, et al. Prediction of state of charge of Lithium-ion rechargeable battery with electrochemical impedance spectroscopy theory; proceedings of the Industrial Electronics & Applications, F, 2010 [C].
- [9] Kemper P, Li S E, Kum D. Simplification of pseudo two dimensional battery model using dynamic profile of lithium concentration [J]. Journal of Power Sources, 2015, 286(jul.15): 510-525.
- [10] Remmlinger J, Buchholz M, Meiler M, et al. State-of-health monitoring of lithium-ion batteries in electric vehicles by on-board internal resistance estimation - ScienceDirect [J]. Journal of Power Sources, 2011, 196(12): 5357-5363.
- [11] Andre D, Meiler M, Steiner K, et al. Characterization of high-power lithium-ion batteries by electrochemical impedance spectroscopy. I. Experimental investigation [J]. Journal of Power Sources, 2011, 196(12): 5334-5341.
- [12] Santhanagopalan S, White R E. State of charge estimation using an unscented filter for high power lithium ion cells [J]. International Journal of Energy Research, 2010, 34(2): 152-163.
- [13] Eddahech A, Briat O, Woirgard E, et al. Remaining Useful Life prediction of Lithium batteries in calendar ageing for automotive applications [J]. Microelectronics Reliability, 2012, 52(9-10): 2438-2442.
- [14] An D, Choi J H, Kim N H. Prognostics 101: A tutorial for particle filter-based prognostics algorithm using Matlab [J]. Reliability Engineering System Safety, 2013, 115: 161-169.
- [15] Su X, Wang S, Pecht M, et al. Prognostics of lithium-ion batteries based on different dimensional state equations in the particle filtering method [J]. Transactions of the Institute of Measurement Control, 2016, 39(10): 1537-1546.

- [16] Wang D, Yang F, Kwok-Leung T, et al. Remaining Useful Life Prediction of Lithium-Ion Batteries Based on Spherical Cubature Particle Filter [J]. IEEE Transactions on Instrumentation Measurement, 2016.
- [17] 彭宇, 刘大同. 数据驱动故障预测和健康管理综述 [J]. 仪器仪表学报, 2014, 35(03): 481-495.
- [18] Liu D, Luo Y, Liu J, et al. Lithium-ion battery remaining useful life estimation based on fusion nonlinear degradation AR model and RPF algorithm [J]. Neural Computing Applications, 2014, 25(3): 557-572.
- [19] Zhou Y, Huang M. Lithium-ion batteries remaining useful life prediction based on a mixture of empirical mode decomposition and ARIMA model [J]. Microelectronics Reliability, 2016, 65: 265-273.
- [20] Yan W, Zhang B, Wang X, et al. Lebesgue-Sampling-Based Diagnosis and Prognosis for Lithium-Ion Batteries [J]. IEEE Transactions on Industrial Electronics, 2016, 63(3): 1804-1812.
- [21] Zhang X, Miao Q, Liu Z. Remaining useful life prediction of lithium-ion battery using an improved UPF method based on MCMC [J]. Microelectronics Reliability, 2017, 75: 288-295.
- [22] Yu J, Mo B, Tang D, et al. Remaining useful life prediction for lithium-ion batteries using a quantum particle swarm optimization-based particle filter [J]. Quality Engineering, 2017, 29(3): 536-546.
- [23] Ma Y, Chen Y, Zhou X, et al. Remaining Useful Life Prediction of Lithium-Ion Battery Based on Gauss-Hermite Particle Filter [J]. IEEE Transactions on Control Systems Technology, 2019, 27(4): 1788-1795.
- [24] Ng S, Xing Y, Tsui K L. A naive Bayes model for robust remaining useful life prediction of lithium-ion battery [J]. Applied Energy, 2014, 118(apr.1): 114-123.
- [25] Li L, Wang P, Chao K-H, et al. Remaining Useful Life Prediction for Lithium-Ion Batteries Based on Gaussian Processes Mixture [J]. Plos One, 2016, 11(9): 13.
- [26] Tang S, Yu C, Wang X, et al. Remaining Useful Life Prediction of Lithium-Ion Batteries Based on the Wiener Process with Measurement Error [J]. Energies, 2014, 7(2): 520-547.
- [27] Li X, Shu X, Shen J, et al. An On-Board Remaining Useful Life Estimation Algorithm for Lithium-Ion Batteries of Electric Vehicles [J]. Energies, 2017, 10(5).
- [28] Doo, Seokgwang, Tagade, et al. A novel multistage Support Vector Machine based approach for Li ion battery remaining useful life estimation [J]. Applied energy, 2015, 159: 285-297.
- [29] Wang S, Zhao L, Su X, et al. Prognostics of Lithium-Ion Batteries Based on Battery Performance Analysis and Flexible Support Vector Regression [J]. Energies, 2014, 7(10): 6492-6508.
- [30] Zhao Q, Qin X, Zhao H, et al. A novel prediction method based on the support vector regression for the remaining useful life of lithium-ion batteries [J]. Microelectronics Reliability, 2018, 85(jun.): 99-108.
- [31] Chao H, Jain G, Schmidt C, et al. Online Estimation of Lithium-Ion Battery Capacity Using Sparse Bayesian Learning; proceedings of the ASME 2015 International Design

- Engineering Technical Conferences and Computers and Information in Engineering Conference, F, 2015 [C].
- [32] Box G, Cox D R. An analysis of transformations [J]. Journal of the Royal Statistical Society, 1964, 26(2): 211-243.
- [33] Zhang Y, Xiong R, He H, et al. Lithium-ion battery remaining useful life prediction with Box-Cox transformation and Monte Carlo simulation [J]. IEEE Transactions on Industrial Electronics, 2018, 66(2): 1585-1597.
- [34] Datong L, Jianbao Z, Haitao L, et al. A Health Indicator Extraction and Optimization Framework for Lithium-Ion Battery Degradation Modeling and Prognostics [J]. IEEE Transactions on Systems, Man, and Cybernetics: Systems, 2015, 45(6): 915-928.
- [35] Cheng C, Ma G, Zhang Y, et al. Online Bearing Remaining Useful Life Prediction Based on a Novel Degradation Indicator and Convolutional Neural Networks [J]. arXiv, 2018.
- [36] Ince T, Kiranyaz S, Eren L, et al. Real-Time Motor Fault Detection by 1D Convolutional Neural Networks [J]. IEEE Transactions on Industrial Electronics, 2016, 63(11): 7067-7075.
- [37] Saxena A, Saha B, Liu J, et al. An Adaptive Recurrent Neural Network for Remaining Useful Life Prediction of Lithium-ion Batteries; proceedings of the Annual Conference of the Prognostics and Health Management Society, F, 2010 [C].
- [38] Zhang Y, Xiong R, He H, et al. Long Short-Term Memory Recurrent Neural Network for Remaining Useful Life Prediction of Lithium-Ion Batteries [J]. IEEE Transactions on Vehicular Technology, 2018, 67(7): 5695-5705.
- [39] Khumprom P, Yodo N. A Data-Driven Predictive Prognostic Model for Lithium-ion Batteries based on a Deep Learning Algorithm [J]. Energies, 2019, 12(4).
- [40] 刘月峰, 赵光权, 彭喜元. 多核相关向量机优化模型的锂电池剩余寿命预测方法 [J]. 电子学报, 2019, 47(06): 1285-1292.
- [41] Zheng X, Fang H. An integrated unscented kalman filter and relevance vector regression approach for lithium-ion battery remaining useful life and short-term capacity prediction [J]. Reliability Engineering System Safety, 2015, 144: 74-82.
- [42] SONG Y, LIU D, HOU Y, et al. Satellite lithium-ion battery remaining useful life estimation with an iterative updated RVM fused with the KF algorithm %J Chinese Journal of Aeronautics [J]. 2018, 31(01): 31-40.
- [43] Chang Y, Fang H, Zhang Y. A new hybrid method for the prediction of the remaining useful life of a lithium-ion battery [J]. Applied Energy, 2017, 206(nov.15): 1564-1578.
- [44] Dong G, Chen Z, Wei J, et al. Battery Health Prognosis Using Brownian Motion Modeling and Particle Filtering [J]. IEEE Transactions on Industrial Electronics, 2018, 65(11): 8646-8655.
- [45] Zhang L, Mu Z, Sun C. Remaining Useful Life Prediction for Lithium-ion Batteries Based on Exponential Model and Particle Filter [J]. IEEE Access, 2018, 6: 17729-17740.
- [46] Song Y, Liu D, Yang C, et al. Data-driven hybrid remaining useful life estimation approach for spacecraft lithium-ion battery [J]. Microelectronics Reliability, 2017, 75: 142-153.
- [47] Cadini F, Sbarufatti C, Cancelliere F, et al. State-of-life prognosis and diagnosis of lithium-ion batteries by data-driven particle filters [J]. Applied Energy, 2019,

- 235(FEB.1): 661-672.
- [48] Hu C, Ye H, Jain G, et al. Remaining useful life assessment of lithium-ion batteries in implantable medical devices [J]. Journal of Power Sources, 2018, 375(JAN.31): 118-130.
- [49] He W, Williard N, Osterman M, et al. Prognostics of lithium-ion batteries based on Dempster-Shafer theory and the Bayesian Monte Carlo method [J]. Journal of Power Sources, 2011, 196(23): 10314-10321.
- [50] Peng Y, Hou Y, Song Y, et al. Lithium-Ion Battery Prognostics with Hybrid Gaussian Process Function Regression [J]. Energies, 2018, 11(6).
- [51] Chen, Laghrouche, Salah, et al. Degradation prediction of proton exchange membrane fuel cell based on grey neural network model and particle swarm optimization [J]. Energy Conversion and Management, 2019, 195: 810-818.
- [52] Li X, Zhang L, Wang Z, et al. Remaining useful life prediction for lithium-ion batteries based on a hybrid model combining the long short-term memory and Elman neural networks [J]. Journal of Energy Storage, 2019, 21: 510-518.
- [53] Yao F, He W, Wu Y, et al. Remaining useful life prediction of lithium-ion batteries using a hybrid model [J]. Energy, 2022, 248(Jun.1): 1-13.
- [54] Berecibar M, Gandiaga I, Villarreal I, et al. Critical review of state of health estimation methods of Li-ion batteries for real applications [J]. Renewable and Sustainable Energy Reviews, 2016, 56: 572-587.
- [55] Lin C P, Cabrera J, Yu D, et al. SOH Estimation and SOC Recalibration of Lithium-Ion Battery with Incremental Capacity Analysis & Cubic Smoothing Spline [J]. Journal of The Electrochemical Society, 2020, 167(9).
- [56] 熊瑞. 动力电池管理系统核心算法 [M]. 北京: 机械工业出版社, 2021.
- [57] 宋绍剑, 王志浩, 林小峰. 基于SOC的锂动力电池组双向主动均衡控制 [J]. 系统仿真学报, 2017, 29(03): 609-617.
- [58] 王萍, 范凌峰, 程泽. 基于健康特征参数的锂离子电池 SOH 和 RUL 联合估计方法 [J]. 中国电机工程学报, 2022, 42(04): 1523-1534.
- [59] 赵申坤, 姜潮, 龙湘云. 一种基于数据驱动和贝叶斯理论的机械系统剩余寿命预测方法 [J]. 机械工程学报, 2018, 54(12): 115-124.
- [60] Smith L, Gal Y. Understanding Measures of Uncertainty for Adversarial Example Detection [J]. 2018.
- [61] 吴铁洲, 刘思哲, 张晓星, 等. 基于 FA-BP 神经网络的锂离子电池 SOH 估算 [J]. 电池, 2021, 51(1): 21-25.
- [62] 杨礼吉, 王家祺, 景丽萍, 等. 基于张量计算的卷积神经网络语义表示学习 [J]. 计算机学报, 2023, 46(03): 568-578.
- [63] 李鹏程. 基于循环神经网络氢燃料电池寿命预测技术 [D]; 电子科技大学, 2021.
- [64] 冯芳. 基于蒙特卡洛神经网络算法的不平衡数据分类研究 [D]; 兰州大学, 2019.
- [65] 黄立威, 江碧涛, 吕守业, 等. 基于深度学习的推荐系统研究综述 [J]. 计算机学报, 2018, 41(07): 1619-1647.
- [66] Bahdanau D, Cho K, Bengio Y. Neural Machine Translation by Jointly Learning to Align and Translate; proceedings of the International Conference on Learning Representations, F, 2015 [C].

-
- [67] 张仲伟, 曹雷, 陈希亮, 等. 基于神经网络的知识推理研究综述 [J]. 计算机工程与应用, 2019, 55(12): 8-19+36.
- [68] 郑芮. 贝叶斯深度标签分布学习不确定度的研究与应用 [D]; 中国科学技术大学, 2021.
- [69] Damianou A C, Titsias M K, Lawrence N D. Variational inference for latent variables and uncertain inputs in Gaussian processes [J]. JMLRorg, 2016, 17(1): 1425-1486.
- [70] Gal Y, Ghahramani Z. Dropout as a Bayesian Approximation: Representing Model Uncertainty in Deep Learning; proceedings of the International Conference on Machine Learning, F, 2016 [C].
- [71] Peng W, Ye Z S, Chen N. Bayesian Deep Learning based Health Prognostics Towards Prognostics Uncertainty [J]. IEEE Transactions on Industrial Electronics, 2020, 67(3): 2283-2293.