

基于CNN-LSTM混合深度学习模型的轴承故障检测方法

郑亚坤

上海理工大学机械工程学院, 上海

收稿日期: 2024年12月15日; 录用日期: 2025年1月5日; 发布日期: 2025年1月15日

摘要

随着近年来工业的快速发展, 对机械设备稳定运行的要求也不断提高。因此, 如何确保机械设备高效且平稳地运行成为了一个重要课题。对于传统的机械设备故障检测方法, 其过于依赖人的经验, 效率低下和检测结果出入过大的情况, 本文提出一种基于混合深度学习模型的轴承故障检测方法。该方法能够准确捕捉到轴承信号中的空间特征和时间特征, 通过美国凯斯西储大学的轴承数据集实验验证, 显著提升了模型的效率和准确性, 相比于传统的检测方法提高了检测效率和检测准确率。

关键词

卷积神经网络, 长短时期记忆网络, 故障诊断, 分类识别

Detection Method of Bearing Faults Based on CNN-LSTM Hybrid Deep Learning Model

Yakun Zheng

School of Mechanical Engineering, University of Shanghai for Science and Technology, Shanghai

Received: Dec. 15th, 2024; accepted: Jan. 5th, 2025; published: Jan. 15th, 2025

Abstract

With the rapid development of industry in recent years, the requirements for the stable operation of mechanical equipment have continuously increased. As a result, ensuring the efficient and smooth operation of machinery has become an important issue. Traditional fault detection methods for mechanical equipment heavily rely on human experience, leading to inefficiency and significant discrepancies in detection results. This paper proposes a bearing fault detection method based on a hybrid deep learning model. The method effectively captures both spatial and temporal features of

the bearing signals. Experimental validation using the bearing dataset from Case Western Reserve University in the United States demonstrates a significant improvement in model efficiency and accuracy, enhancing both detection efficiency and accuracy compared to traditional methods.

Keywords

Convolutional Neural Networks, Long Short-Term Memory Networks, Fault Diagnosis, Classification Recognition

Copyright © 2025 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

随着科学技术的快速进发和工业规模的日益增长, 轴承作为机械设施和工厂高效率的运作过程中发挥着举足轻重的作用。但是, 由于长时间高负荷运转、环境因素和保养不充分等原因, 轴承可能出现多种多样的故障现象, 如磨损、裂纹和剥落, 严重的时候还可能会造成整个设备的损坏和安全事故[1]。因此, 能够及时和准确地检测出轴承的故障对于整个工业的生产和运作极其重要。以往的过于依赖人工技术的检测方法存在效率较低和检测准确性低等情况。

将机械振动信号用在轴承的故障检测之上, 是一种现在智能检测的一种重大突破, 在工程实践中被应用广泛。Li 等[2]介绍了一种基于 GRU-DeepAR 模型的滚动轴承剩余使用寿命预测方法, 利用自适应故障阈值提高预测精度。通过提取振动信号特征, 划分操作阶段, 建立降级数据集, 并进行模型训练和误差评估, 实验验证了方法的有效性和稳定性。王克定等[3]提出了一种基于深度迁移学习的矿井通风机轴承故障诊断方法, 结合卷积神经网络(CNN)、双向门控循环单元(BiGRU)和随机森林(RF)分类器, 解决负样本少导致的诊断率低问题。柳雅倩等[4]提出了一种基于深度卷积生成对抗网络(DCGAN)的方法, 用于在小样本条件下扩充滚动轴承故障库。通过设计上采样卷积(USCONV)层, 解决了传统方法的棋盘格效应问题, 从而提高故障诊断的准确率。谭启瑜等[5]提出了一种基于图卷积神经网络(GCN)的滚动轴承故障诊断方法。通过快速傅里叶变换将振动信号转换为频域信号, 并构建全连接图, 利用 GCN 提取节点特征, 最终使用全连接层和 Softmax 分类器进行故障识别。实验结果表明, 该方法在强噪声环境下相比传统机器学习和其他深度学习模型, 具有更高的准确率和抗噪性。Lin 等[6]提出了一种基于深度学习的自动诊断方法, 用于检测通用航空活塞发动机轴承故障。通过改进 YOLOv5 模型, 结合多尺度检测层、坐标注意机制和 Alpha-CIoU, 显著提升了小目标检测性能, 提高了检测准确率, 为发动机维护提供了有力支持。韩争杰等[7]提出了一种基于注意力机制的改进残差神经网络, 用于滚动轴承故障诊断。通过 SE-ResNet 和 CBAM-ResNet 模型, 提升了模型在不同工况下的准确率和泛化性能, 在实验中优于传统 ResNet, 验证了方法的有效性。Kaya 等[8]提出了一种结合连续小波变换(CWT)和深度迁移学习方法的自动轴承故障尺寸诊断技术。通过将振动信号转化为时频图像, 利用多种深度学习模型进行故障尺寸预测, 成功率达到 96.67%至 100%。

近年来, 循环神经网络(RNN)在时间序列数据特征挖掘领域有着很高的地位, 门控循环单元(GRU)和长短期记忆网络(LSTM)及其变体在时间序列预测领域取得了显著的成果。本文将轴承故障振动信号这一典型的时间序列问题作为研究对象, 重点采用 LSTM 方法进行分析与建模。同时引入 CNN 来提取空间特征并减小对 LSTM 网络模型的数据输入量, 减小运算时间和快速拟合。

2. 算法原理

2.1. RNN 模型

RNN 是一种专门针对序列数据处理的人工神经网络，其主要原理是引入一个循环隐藏层，使得网络能够保存历史信息，并利用这些过去的的数据来影响当前的输出结果。如图 1，RNN 在每个时间步都接受输入和隐藏状态，并输出一个新的隐藏状态和预测值。通过这种方式，RNN 能够有效地建模数据中的时间依赖性。

在处理轴承数据时，RNN 尤其能够发挥其时序建模的优势。轴承的振动信号通常具有明显的时间变化规律，而 RNN 可以捕获这些规律，识别出潜在的故障特征。这种能力使得 RNN 在故障预测和监测领域变得非常有用，为后续的维护决策提供了坚实的基础。

RNN 能够通过分析历史振动数据，了解轴承在正常和异常状态下的表现，从而在早期阶段预测可能的故障。这种预警机制可以有效减少设备停机时间，降低维护成本。此外，RNN 还可以与其他传感器数据结合，综合评估设备健康状况，提高预测的准确性。

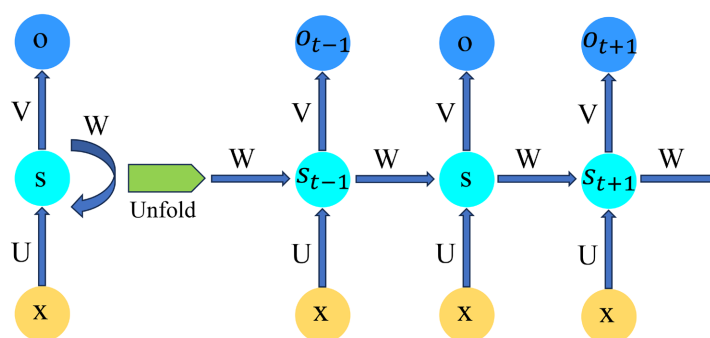


Figure 1. RNN structure diagram
图 1. RNN 结构图

通过不断学习新的数据，RNN 不仅可以适应不同的工作条件，还能不断优化预测模型，从而提升整个系统的可靠性和安全性。

2.2. CNN-LSTM 模型

CNN 是一种具有卷积结构的深度神经网络，是深度学习中的经典算法，具备强大的特征提取能力。如图 2，其核心构成包括卷积层和池化层。

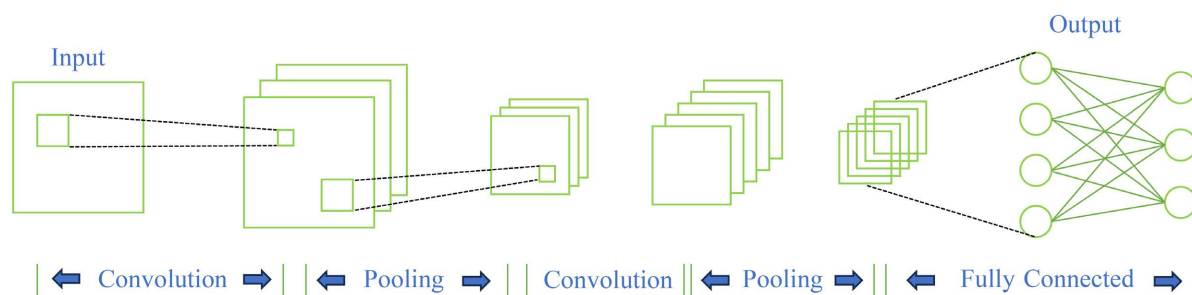


Figure 2. CNN structure diagram
图 2. CNN 结构图

卷积层通过卷积操作提取输入特征图的局部特征，而池化层则通过下采样来降低特征图的维度，进而减少网络的参数和计算量，能够达到节省资源的效果。卷积神经网络的优势在于局部感知和参数共享机制，这使得模型具有更强的泛化能力和对平移的不变性。然而，对于时间序列数据的处理，卷积神经网络的表现并不十分理想。

不同于 CNN 良好的局部特征提取能力。如图 3，LSTM 作为改进的循环神经网络，其可以提取长短时间序列的复杂的特征关系，所以在处理时间序列上的问题有着更好的效果。故可以将输入的特征数据经过 CNN 模型特征提取和扁平化处理之后传入到 LSTM 模型中进一步操作。

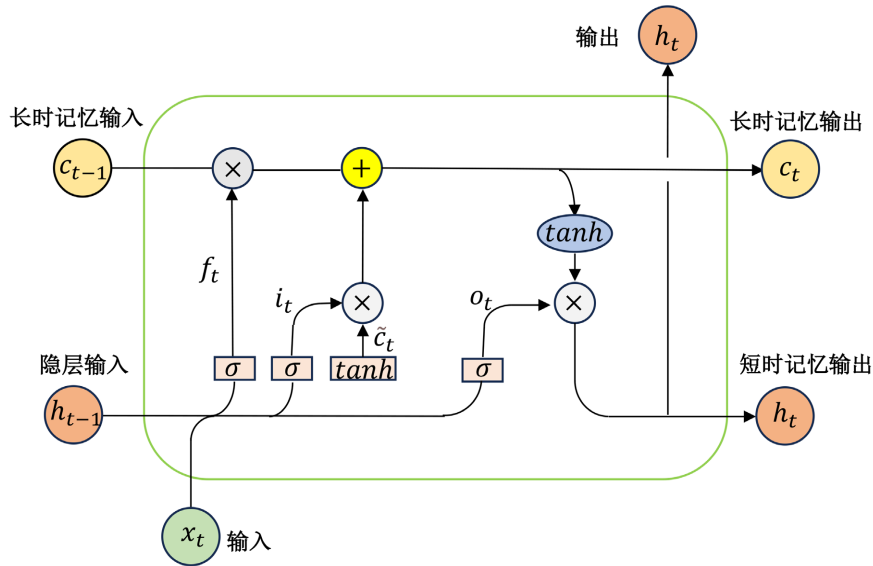


Figure 3. Schematic diagram of LSTM structure

图 3. LSTM 结构原理图

在 LSTM 中，首先需要决定哪些信息从细胞状态中丢弃，这个决定是通过忘记门来实现的。

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

第二步是确定哪些新信息将被存储在细胞状态中，这包括两个部分。输入门决定哪些值需要更新，而 tanh 层则生成一个新的候选值向量，为下一步的计算做准备。

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \quad (2)$$

$$C_t = \sigma(W_c \cdot [h_{t-1}, x_t] + b_c) \quad (3)$$

第三步更新旧细胞状态，把旧状态与 f_t 相乘，丢弃不重要的信息，接着加上 $i_t * \tilde{C}_t$ ，这就是新的候选值，决定每个状态的程度进行变化。

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t \quad (4)$$

最后，需要决定输出哪些值，这个输出是基于细胞状态。通过一个 sigmoid 层来确定细胞状态需要输出的部分，然后将细胞状态放入 tanh 层进行处理，并且与 sigmoid 层的输出相乘达到效果。

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \quad (5)$$

$$h_t = o_t * \tanh(C_t) \quad (6)$$

在以上公示中 W_f 、 W_i 、 W_c 、 W_o 、 b_f 、 b_i 、 b_c 、 b_o 分别为不同输入状态的权重系数矩阵与偏置； σ 和 $\tanh$ 为 sigmoid 函数的双曲与正切激活函数； h_t 为 LSTM 网络的输出值。

3. 实验过程与分析

3.1. 实验数据

本文利用美国凯斯西储大学的轴承数据集进行模型有效性验证实验。实验测试包括内圈故障、滚动体故障和外圈故障(6 点钟位置)三种类型，每种类型有四种不同的损伤直径：0.007、0.014、0.021 和 0.028 英寸。我们分析正常状态、内圈故障、滚动体故障和外圈故障(6 点钟位置)四种类别的数据，每类故障有三种程度，总共 10 种故障状态。

数据的采样频率为 48 kHz，电机转速为 1750 r/min，轴承每转一圈采样 1646 个点。为了确保每个样本包含完整周期，样本长度设置为 1680。增加神经网络训练数据时，使用 50%重叠进行采样，而测试数据则采用无重叠方式截取。

3.2. 混合模型设计

滚动轴承是机械设备中的关键部件，其运行状态直接影响设备的稳定性。混合 CNN-LSTM 模型是一种常见的轴承故障诊断方法，能够有效地捕捉时空特征，从而提高故障诊断的准确性，并实现对轴承健康状态的有效监控。

数据预处理：数据预处理首先需要收集振动数据，然后进行去噪、降采样、滤波等操作，得到处理后的数据集。接着按照 7:2:1 的臂力划分训练集、验证集、测试集进行随机抽样划分。这些数据集将用于轴承故障检测的模型训练和测试。

基础模型选择：在引入 CNN-LSTM 模型之前，首先需要选定一个基础模型进行对比。通常，CNN 或 RNN 等模型被用于从数据中提取特征并进行处理。

构建混合 CNN-LSTM 模型：CNN-LSTM 混合模型将 CNN 与 LSTM 相结合，旨在有效处理二维时间序列数据。在混合模型中，各种训练参数的选取 Epochs 为 50、Batch_Size 为 32、学习率为 0.003 在模型训练过程中将训练数据输入网络模型中，在网络中完成一次前向和反向传播通过多层的 CNN 网络模型提取数据的空间特征，多层的 LSTM 模型提取数据的时间特征，从而实现对轴承运行状态的检测。滚动轴承故障检测流程见图 4。

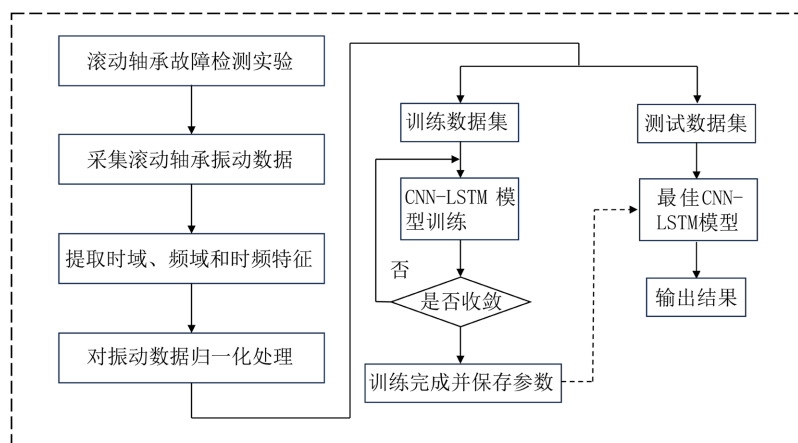


Figure 4. Design of the CNN-LSTM model framework
图 4. CNN-LSTM 模型框架设计

CNN 层模块：CNN 层数设置为 3 层，每层包含卷积、激活函数和最大池化。本文设置的输入通道数为 1，输出通道数为 128，卷积核的大小设置为 3×3 ，步长大小设置为 1，填充大小设置为 1。**LSTM 层模块：**LSTM 单元的层数为 2，输入特征维数为 32，隐藏层大小设置为 128，不使用偏置项，并设置第一个维度的大小为 Batch_size。**全连接层：**经过 CNN 层和 LSTM 层处理过之后的数据输入全连接层来达到所需的分类类别数为 10。模型构造详见表 1。

Table 1. Specific parameters of the network model

表 1. 网络模型具体参数

层类型	输入通道数	输出通道数	卷积核大小	LSTM 层数	参数数量
Conv1d(1)	1	128	3		512
Conv1d(2)	128	128	3		147,584
Conv1d(3)	128	128	3		147,584
LSTM				2	65,536
Linear		10			1290

3.3. 实验结果分析

利用特征融合的方法，把 CNN 的浅层数据特征和深层数据特征输入到 LSTM 单元中增强网络模型的计算能力，高效的获取到了故障信号中的时间特征和空间特征，进而对轴承发生的故障检测效率进一步提高。

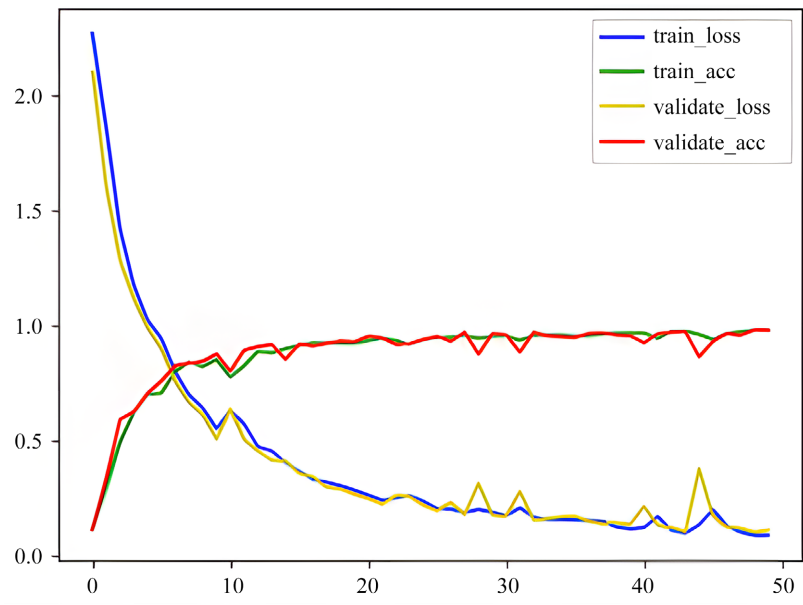


Figure 5. Model training accuracy

图 5. 模型训练精度

实验结果表现出，本文所使用的研究方法能准确对轴承的故障进行分类计算，该模型实现了轴承运转过程的实时监控，为轴承故障诊断提供了新的方法和思路。根据图 5 所示，通过 50 轮训练，结果表明该模型在轴承故障分类任务中的准确率达到了 98.06%。从结果中足以看出应用混合 CNN-LSTM 模型进行实际的轴承故障检测的可行性。

3.4. 不同网络模型实验结果对比

将本文所使用的轴承数据集输入到 BP 神经网络和 CNN 中进行分析和训练。见表 2，通过这两种网络模型和本文所使用的网络模型的检测结果进行比较，我们可以更深入地理解它们在轴承故障识别方面的表现。

CNN-LSTM 模型相较于 BP 神经网络和 CNN 神经网络，在轴承故障检测中的优势在于其结合了 CNN 强大的特征提取能力和 LSTM 处理时序数据的优势。CNN 能够自动提取故障信号的空间特征，而 LSTM 通过其记忆机制能够有效捕捉信号的时间依赖性，适应轴承故障信号中的动态变化。这种结合提高了模型对复杂故障模式的识别准确性和鲁棒性，进一步反映出本文所采用的网络模型在实际应用中有着更高的有效性和可靠性。

Table 2. Comparison of different network models

表 2. 不同网络模型对比

诊断方法	平均准确率
BP 神经网络	87.3%
CNN	94.6%
本文方法	98.06%

4. 结论与展望

本文研究的内容将深度学习和人工智能进行深度结合，给轴承故障检测提出了一种新的方法和理念。本文所研究的内容为将 CNN 和 LSTM 两种网络进行融合，目的是为了提升轴承故障检测的准确性和稳定性。

本文所提出来的检测方法不仅在轴承故障检测领域起到了显著性的效果，还使得在智能故障检测的领域中提供了一种新的方法。结果表明，将 CNN 和 LSTM 两种网络模型结合起来，弥补了各自的缺点，将两种模型的优点进行融合，使轴承故障检测迈入了一个新的台阶。

当然，本文的研究还是有一些不足之处。在往后的研究之中，可以使用深度卷积神经网络对振动信号的特征进一步的提取，还可以融科迁移学习方法，进一步提高轴承故障检测的精确性和稳定性。

参考文献

- [1] 吴定海, 任国全, 王怀光, 张云强. 基于卷积神经网络的机械故障诊断方法综述[J]. 机械强度, 2020, 42(5): 1024-32.
- [2] Li, J., Wang, Z., Liu, X. and Feng, Z. (2023) Remaining Useful Life Prediction of Rolling Bearings Using GRU-DeepAR with Adaptive Failure Threshold. *Sensors*, **23**, Article No. 1144. <https://doi.org/10.3390/s23031144>
- [3] 王克定, 李敬兆, 石晴, 胡迪. 基于深度迁移学习的矿井通风机轴承故障诊断[J]. 机床与液压, 2023, 51(22): 209-14.
- [4] 柳雅倩, 蔡浩原, 李文宽, 等. 小样本条件下轴承故障的 DCGAN 诊断方法[J]. 振动测试与诊断, 2023, 43(4): 817-823+836.
- [5] 谭启瑜, 马萍, 张宏立. 基于图卷积神经网络的滚动轴承故障诊断[J]. 噪声与振动控制, 2023, 43(6): 101-108+116.
- [6] Lin, M., Liu, Q., Zeng, R., Bai, Y. and Zhang, G. (2023) An Automatic Diagnosis Method for Bearing Failure of General Aviation Piston Engine with Deep Learning Networks. *3rd International Conference on Artificial Intelligence, Virtual Reality, and Visualization (AIVRV 2023)*, Chongqing, 7-9 July 2023. <https://doi.org/10.1117/12.3011402>
- [7] 韩争杰, 牛荣军, 马子魁, 等. 基于注意力机制改进残差神经网络的轴承故障诊断方法[J]. 振动与冲击, 2023, 42(16): 82-91.

- [8] Kaya, Y., Kuncan, F. and Ertunç, H.M. (2022) A New Automatic Bearing Fault Size Diagnosis Using Time-Frequency Images of CWT and Deep Transfer Learning Methods. *Turkish Journal of Electrical Engineering and Computer Sciences*, **30**, 1851-1867. <https://doi.org/10.55730/1300-0632.3909>