

# 深度卷积长短期记忆网络的轴承故障诊断

孙洁娣<sup>1,2</sup>,毛新茹<sup>1,2</sup>,温江涛<sup>3</sup>,张鹏程<sup>3</sup>,张光宇<sup>3</sup>

(1. 燕山大学 信息科学与工程学院,河北秦皇岛 066004;

2. 燕山大学 河北省信息传输与信号处理重点实验室,河北秦皇岛 066004;

3. 燕山大学 河北省测试计量技术及仪器重点实验室,河北秦皇岛 066004)

**摘要:** 针对传统数据驱动故障诊断方法难以从轴承信号中自适应提取有效特征、没有充分利用故障数据的时序特点以及缺乏自适应处理动态信息能力的问题,提出了一种深度卷积神经网络与长短期记忆网络相结合的智能故障诊断方法。本文方法构建的深度模型能够从轴承原始信号中自适应地提取鲁棒性特征,然后利用长短期记忆网络学习特征中的时间依赖关系实现了高准确度的轴承故障诊断。该方法克服了传统特征提取方法依赖专家经验和信息利用不完全等问题,实现了故障的智能、准确诊断。实验结果表明,该方法可以提取更准确的特征而且由于利用了故障演变过程中的时序信息,使得故障诊断更加智能、可靠。

**关键词:** 轴承故障诊断; 卷积神经网络; 循环神经网络; 时序序列

**中图分类号:** TH865; TP29

**文献标志码:** A

**DOI:** 10.13433/j.cnki.1003-8728.20200170 **文章编号:** 1003-8728(2021)07-1091-09



OSID 码

## Bearing Fault Diagnosis Using Deep CNN and LSTM

SUN Jiedi<sup>1,2</sup>, MAO Xinru<sup>1,2</sup>, WEN Jiangtao<sup>3</sup>, ZHANG Pengcheng<sup>3</sup>, ZHANG Guangyu<sup>3</sup>

(1. School of Information Science and Engineering, Yanshan University, Qinhuangdao 066004, Hebei, China;

2. Hebei Key Laboratory of Information Transmission and Signal Processing, Yanshan University, Qinhuangdao 066004, Hebei, China;

3. Key Laboratory of Measurement Technology and Instrumentation of Hebei Province, Yanshan University, Qinhuangdao 066004, Hebei, China)

**Abstract:** There exists some problems in traditional data-driven fault diagnosis methods, for example, it is difficult to adaptively extract effective features from bearing signals, cannot make full use of the timing characteristics of fault data, and lacks of the ability to adaptively process dynamic information. An intelligent bearing fault diagnosis method combined deep convolutional neural network and long-term and short-term memory network is proposed. This method constructed a kind of deep networks and can adaptively extract the robust features from the original bearing signals, and then utilized the long short-term memory network to learn the time-dependent relationship in these features, and achieved high-accuracy bearing fault diagnosis. The proposed method overcame the problems existed in the traditional feature extraction methods, such as heavy dependence on expert experience and incomplete utilization for time series information, and realized intelligent and accurate diagnosis of faults. The experimental results show that the proposed method can extract more accurate features and make the fault diagnosis more intelligent and reliable by utilizing the timing information in the process of fault degradation

**Keywords:** bearing fault diagnosis; convolutional neural network; long short-term memory network; temporal sequence

收稿日期: 2019-12-18

基金项目: 国家自然科学基金项目(61973262, 62073282)、河北省自然科学基金项目(E2020203061)、河北省引进留学人员资助项目(C201827)、河北省高等学校科学技术研究项目(QN2019133)及京津冀联合攻关重点研发计划(19YFSLQY00080)

作者简介: 孙洁娣(1975-) 教授,博士,研究方向为机械智能故障诊断, wjtsjd@163.com

滚动轴承是旋转机械的重要组成部分,对机械的健康状态产生直接的影响。滚动轴承故障可能导致意外停机或经济损失<sup>[1]</sup>,因此准确的轴承故障诊断在机械健康管理中起着关键作用。

目前,基于振动信号分析的轴承故障检测得到广泛应用。该类方法主要是通过传感器采集大量振动信号后,采用多种信号分析方法实现故障类别分

析,主要包括两个步骤:数据特征提取和筛选、故障诊断,其中时域、频域和时频域方法<sup>[2-4]</sup>是特征提取常用的处理方法。在故障诊断方面,通常用支持向量机<sup>[5]</sup>、模糊推断及神经网络等人工智能方法对特征进行学习训练,进而实现故障诊断<sup>[6]</sup>。学者们已提出了许多传统智能诊断方法<sup>[7-9]</sup>,这些方法虽然实现了故障诊断,特征提取却受限于先验知识和专家经验,而且这些特征是根据特定诊断问题选取的,具有一定局限性,缺乏泛化能力;此外传统故障分类方法多是基于浅层网络的学习方法,简单的结构限制了神经网络在故障诊断中学习复杂非线性关系的能力。针对上述问题,研究自适应特征提取和诊断方法在轴承智能故障诊断中更具吸引力。

近年来,深度学习在许多领域中取得了令人瞩目的成绩。它通过多层非线性变换自适应地学习高维特征表示,摆脱传统智能方法的限制<sup>[10-12]</sup>,特别是卷积神经网络(Convolutional neural networks, CNN)和循环神经网络(Recurrent neural network, RNN)在很多领域获得了应用。CNN具有强大的特征提取能力,并逐渐被用于机械故障诊断中。Ince等<sup>[13]</sup>利用原始数据作为输入并提出利用一维CNN将特征提取和分类集成在一起,完成实时故障诊断。Lin等<sup>[14]</sup>利用CNN从振动信号的频谱中学习特征,实现变速箱的健康状态监测。曲建岭等<sup>[15]</sup>提出了一维卷积神经网络故障诊断算法,实现了基于深度网络的自适应特征提取和故障诊断。Lu等<sup>[16]</sup>通过分段组合的方式将原始监测信号映射到二维矩阵中,经深度CNN网络自动学习高维特征表示,实现轴承的故障诊断。

分析已有方法可见,CNN用于故障诊断主要有两种方式:一类是源于CNN在图像处理中的应用,将一维振动信号通过时频变换或者分段组合的方式形成二维矩阵,通过CNN强大的学习能力获取故障信息;另一类则多是在一维CNN基础上,通过滑动窗口将传感数据分割为固定长度序列段,利用一维CNN学习故障信息。轴承振动信号是一种自然时间序列信号,故障信息随着时间的推移不断增多,数据间的动态变化包含大量轴承衰退信息,可用的历史数据越多,诊断将越准确。由于数据分割时的滑动窗口是独立存在的,轴承原始数据中各时刻数据点的关联性会受到破坏,与故障相关的信息可能会丢失,因此单纯CNN网络难以捕捉轴承信号随时间变化的故障信息。长短期记忆网络(Long short-term memory, LSTM)是一种特殊的RNN结构,具有很强的动态建模学习能力。它通过状态的时间序列全

面地学习故障信息,弥补了单独使用CNN处理时序数据的不足,许多涉及顺序数据的机器学习问题都证明了LSTM在分析时间信息上的有效性<sup>[17-19]</sup>。

轴承故障诊断中,不仅要利用空间联系,还需考虑数据在时间维度上的变化,本文将卷积神经网络与循环神经网络结合,构建了深度卷积循环神经网络(Deep convolution long short-term memory, DCLSTM),该模型具有时间和空间的特征表达能力,可以充分利用传感数据的序列信息,并能挖掘更多隐藏的时间信息。本文方法基于这一网络模型,以原始信号作为输入,可以自适应获取特征表示,避免了主观、片面选择特征及分类产生的问题;利用LSTM网络,获得更多潜在信息充分利用信号间的序列关系准确识别故障。

## 1 基础理论

### 1.1 基于CNN的特征提取

CNN是一种受视觉系统结构启发的深度学习模型<sup>[20-21]</sup>,实质是通过建立多个滤波器提取输入数据的特征。随着网络的加深,提取出更高级别的特征,最终从原始数据中得到具有移不变性的鲁棒特征。轴承振动时序数据存在局部相关性,因此利用CNN可以较好的处理局部特征。

CNN特征提取中,卷积层与池化层交替排列组成卷积组,逐层提取特征。卷积层通过设计适宜的卷积核来卷积上一层输出的特征矢量。设给定输入轴承振动信号 $X$ ,一维卷积特征映射可表示为

$$y_j^k = (X * W) = g\left(\sum W_{ij}^k \cdot X_i^{k-1} + b_j^k\right) \quad (1)$$

式中: $*$ 表示卷积操作; $k$ 为第 $k$ 层网络; $W$ 为卷积核; $y_j^k$ 为第 $k$ 层输出; $g$ 为非线性激活函数ReLU,它的一部分神经元输出为0。ReLU的使用使得梯度下降以及反向传播更有效,同时还可以简化计算过程,加速训练。

池化层是对输入特征进行缩放,提取主要特征。本文采用最大池化,它对输入特征执行局部最大操作,以减少参数并获得位置不变特征。这使网络对于振动信号中的小偏移和失真更具鲁棒性。此外,本文在卷积过程中利用边缘零填充方法确保在卷积输出后数据大小不变,保证数据的结构顺序不变,以便后续的LSTM能够学习分析数据的时间动态。通过CNN发掘数据之间的关联并从中剔除噪声和不稳定成分,将处理后相对稳定的信息作为整体传入LSTM网络进行长序列故障诊断。

## 1.2 基于 LSTM 的故障分类

传统的神经网络是建立输入层  $x_i$  与输出层  $y_i$  之间的关系,而循环神经网络则是建立前  $t$  个输入  $x_{i-t}$  与输出层的关系。1997 年 Hochreiter 和 Schmidhuber 提出了 LSTM<sup>[22]</sup>。LSTM 通过门函数,控制历史数据的传输,具有时间序列处理能力。LSTM 的另一个创新是存储单元  $C_t$ ,它充当状态信息的累加器,用来保持长期记忆,能从隐藏层中提取动态特征,这是其在时间序列分析中应用的关键。图 1 显示了在  $t$  时刻 LSTM 的典型结构。

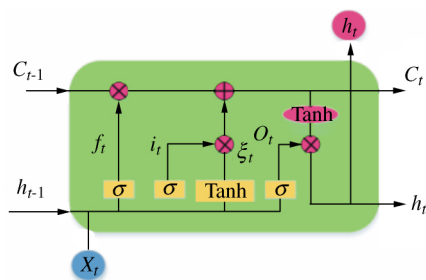


图 1 LSTM 内部结构图

首先,遗忘门决定是否保留先前细胞状态信息  $C_{t-1}$ 。该门读取上一时刻隐藏状态  $h_{t-1}$  和新输入的数据  $X_t$ ,即

$$f_t = \sigma(W_f[h_{t-1}, X_t] + b_f) \quad (2)$$

接下来,输入门计算可以从输入中学习什么信息,并决定将要使用和保存的值:

$$i_t = \sigma(W_i[h_{t-1}, X_t] + b_i) \quad (3)$$

$$\xi_t = \tanh(W_\xi[h_{t-1}, X_t] + b_\xi) \quad (4)$$

下一步操作是更新细胞状态所呈现的长期记忆,即

$$C_t = f_t \cdot C_{t-1} + i_t \cdot \xi_t \quad (5)$$

最后,使用输出门从输入、单元状态和先前隐藏状态更新隐藏状态:

$$O_t = \sigma(W_o[h_{t-1}, X_t] + b_o) \quad (6)$$

$$h_t = O_t \cdot \tanh(C_t) \quad (7)$$

式中:  $f_t, i_t, O_t$  分别表示遗忘门、输入门和输出门,这些门的值在 0 到 1 之间;  $h_t, \xi_t$  分别表示隐藏状态、候选细胞状态;  $W_f, W_\xi, W_i, W_o, b_f, b_\xi, b_i, b_o$  对应于各个门的权重矩阵和偏差;  $\sigma$  是 sigmoid 函数。对于轴承故障诊断问题,构建的 LSTM 模型为多对一形式(即输入为序列,输出为类别)。将待处理的数据输入到 LSTM 网络,通过前向计算得到 LSTM 网络的最后一个输出状态  $h_t$  作为轴承故障诊断的最终结果。

## 2 基于 DCLSTM 的轴承故障诊断

针对浅层网络故障识别的局限性和时序数据时间动态特性利用不充分的问题,提出 DCLSTM 网络模型用于轴承故障诊断,该网络包含 5 个卷积层,5 个池化层,两个全连接层和 3 个 LSTM 循环层。网络内部结构框架及信号处理过程如图 2 所示。

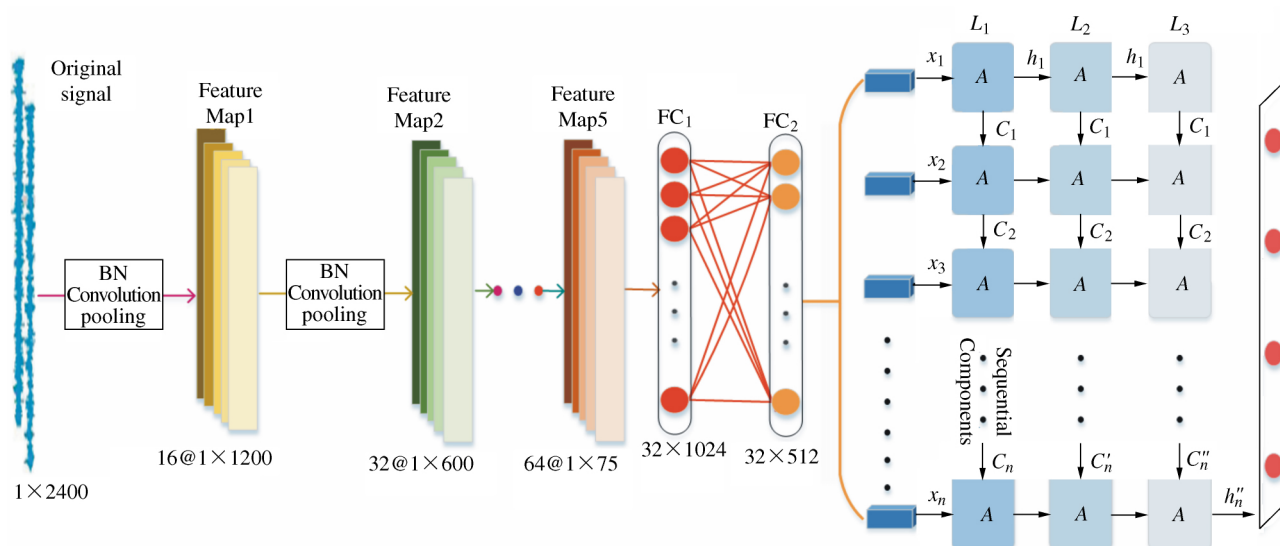


图 2 本文方法的结构框图

该模型由深度一维 CNN 及 LSTM 网络组成,能够学习原始信号中的深层特征信息和时间动态信息。首先,一维 CNN 利用其强大特征自学习能力从原始振动数据中自适应提取故障特征,减少对大量

信号处理方法与诊断经验的依赖。其次,考虑到轴承振动信号的时间关联特性,LSTM 被用来学习 CNN 网络输出的故障特征,建立动态时间序列分类模型。最后,通过 softmax 完成特征分类,以识别轴

承故障的类型。

## 2.1 基于 LSTM 的故障分类

本文的 CNN 网络由 5 个交替堆叠的卷积池化层和 2 个全连接层构成,如图 2 所示。首先,卷积层沿时间轴处理输入样本以获得特征表示,其中第一个卷积层使用大小为  $1 \times 3$  的 16 个内核对输入数据进行特征提取,除内核数量之外,其他卷积层采用与第一个卷积层类似的结构;第二和第三层有 32 个内核,第四和第五层有 64 个内核。每个卷积层处理前一层产生的特征映射分为三步:卷积运算、批标准化运算 (Batch normalization, BN) 和整流线性单元 (ReLU)。BN 可以加速网络训练并提高深度神经网络的性能,减少内部协方差的转移。在等式 (1) 的基础上,使用 BN 变换来规范化  $WX+b$ ,即

$$y = (X^* W) = \sum BN(WX; \gamma) + b \quad (8)$$

基于批标准化的 CNN 结构,每层神经网络的输入保持相同分布,使得通过训练数据获得的模型在测试集上也有好的效果。随后池化层使用最大池化函数计算某个位置及其相邻位置参数的整体特征,即执行大小为  $1 \times 2$  且步长为 1 的最大池化操作,实现采样聚合,从输入数据中获得最重要的局部特征。最后经过多层卷积池化提取的特征被传送到全连接层以进行特征融合。

CNN 特征提取的结果作为 LSTM 网络输入。LSTM 引入时间对网络的影响,增加时间维度,将数据转化为三维矩阵,通过 1.2 节介绍的 LSTM 网络结构,仅将 LSTM 的最后时刻的隐藏层输出馈送到完全连接的层中以进行轴承故障分类。本文采用了 3 层 LSTM,以便能捕获由卷积运算提取的特征的时间依赖性,全面刻画时序数据。

## 2.2 故障的分类识别

图 2 显示了用于轴承故障诊断的 DCLSTM 架构。输入数据通过所构建的网络结构学习故障特表示,多层网络有助于获得输入数据的良好表示并改善网络性能,最终由 softmax 全连接进行特征分类。

在训练期间,根据梯度函数逐渐更新权重和偏差,反向传播算法用于监视学习过程并调整每层参数,计算每次网络迭代的误差。所提出模型的损失函数是输出概率分布与目标概率分布之间的交叉熵。将目标分布表示为  $y'$ ,输出分布表示为  $y$ 。交叉熵损失函数定义为

$$H(y) = - \sum_b y' \log_2(y) \quad (9)$$

每次训练的损失函数由前向计算得到,利用误差反向传播算法进行网络参数更新,直到最大训练步数。一旦导出损失函数,就可以用很多优化算法

训练网络,例如随机梯度下降<sup>[14]</sup>,Adam 随机优化。本文中,应用了易于实现且高效的随机梯度下降法训练网络,将训练好的网络参数保存用于网络测试。

本文方法的思想是利用 CNN 的特征抽象能力提取高维特征,利用 LSTM 引入时序诊断,保留长期有效信息,实现故障特征的自动提取和故障类别的智能诊断。详细过程如图 3 所示。

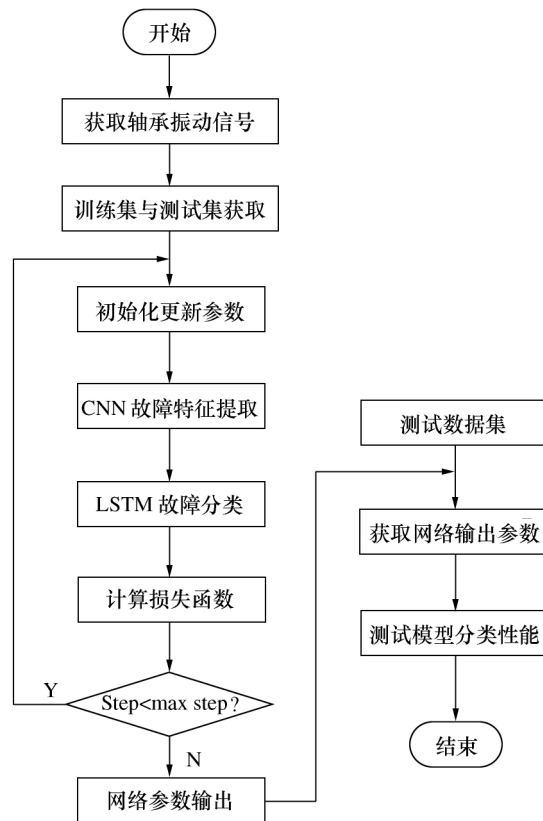


图3 本文方法处理过程

## 3 实验结果与分析

### 3.1 数据描述与处理

本文利用西储大学轴承数据验证所提出的方法<sup>[23]</sup>。西储大学的测试设备包括两个电机,一个耦合器和其他设备。加速度计用于收集不同故障类型的信号。驱动端轴承 (DE) 加速计沿垂直方向安装在 DE 壳体上。风扇端轴承 (FE) 加速计在垂直方向安装在 FE 壳体上。

研究采用驱动端轴承数据,其故障采样频率为 48 kHz 和 12 kHz。待识别的故障类型有 4 种: 外圈故障 (OF)、内圈故障 (IF)、滚动故障 (RF) 和正常情况 (N)。对于正常的轴承,振动模式是恒定的,而当一些故障发生时,振动模式发生变化。图 4 为训练数据集的某些振动信号原始波形。

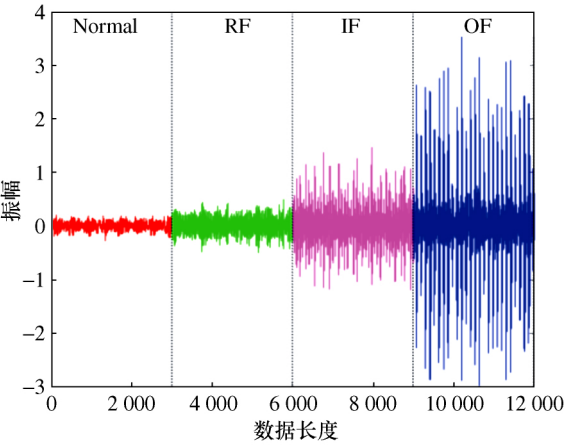


图 4 振动信号图

将收集的信号整理成数据集  $\{X_i, l_i\}_{i=1}^m$ ,  $m$  表示样本数量,  $X_i$  表示第  $i$  个样本,  $l_i$  表示  $X_i$  对应的数据标签。由于循环网络进行的是时间序列分析, 因此训练集和测试集需按时间顺序排列, 不能随机抽取。实验台在 4 种条件下运行: 0, 1 hp, 2 hp, 3 hp, 构建 4 种数据集  $A, B, C, D$ , 如表 1 所示, 每个数据集对应一种操作条件, 并包含 4 种故障类型。

表 1 振动信号数据集

名称	负载	采样点数
A	0	2 400
B	1	2 400
C	2	2 400
D	3	2 400

3.2 诊断结果分析

根据本文方法, 首先构建深度 DCLSTM 模型, 将归一化后的振动信号由网络中的 DCNN 进行自适应特征提取, 而 LSTM 则进行数据的时序分析, 最后进行故障类型的识别。研究中轴承故障信号样本长度取 2400, 采用分类正确样本数与测试总样本数的比值来衡量本文方法的总分类准确率, 统计多次实验结果的平均分类准确率为 99.96%, 损失收敛于 0.08。

下面以一组 48 kHz 采样频率下驱动端采集的信号为例进行结果展示及分析。模型主要参数设定如表 2 所示。

表 2 模型的主要参数选择

网络参数	设置值	训练参数
CNN 层数	5	学习率为 $8 \times 10^{-4}$
池化移动步长	2	训练集比率为 80%
全连接层	2	最大轮数为 100
LSTM 层数	3	Batch_size 为 32
时间步长	20	SGD

本文方法某次分类结果的混淆矩阵如图 5 所示。图 5 中对角线表示正确分类的样本数量, 其余每一块显示出错分为其他类别的样本数量。本文方法中每一类都有较高的分类精度, 只将少部分的第三类故障错分为第四类故障。

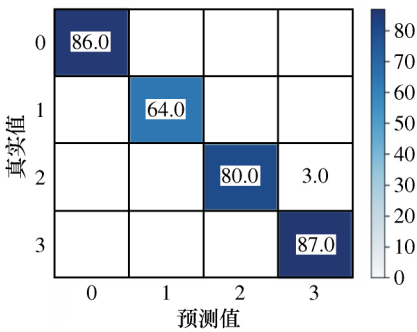


图 5 某次分类结果统计的混淆矩阵

3.3 主要参数的选择与分析

研究中发现模型中的主要参数会对故障诊断的准确度和耗费时间产生影响, 因此后文将对这些参数进行深入研究并分析其对诊断结果的影响。为了减小随机因素的影响, 每种方法的最终统计数据为 10 次实验结果的平均值。

3.3.1 CNN 层数的影响

理论上, 网络结构越深误差越小, 准确率越高。但与此同时网络复杂度也会随之增加, 训练更困难。综合目前轴承诊断已有成果通常采用的网络层数, 并且考虑计算复杂度和实时故障检测的要求, 本文实验设计了 5 种不同的 CNN 结构, 将一个交替堆叠的卷积层和池化层简称为 Conv-Maxpool (记为 L), 全连接层记为 FC。为了能够实现轴承原始信号的逐级特征抽取, 通常 CNN 网络中会设计三层以上的卷积池化层, 本文中主要比较了不同卷积池化层的特征抽取效果, 比较的网络结构记为  $CNN_1 \sim CNN_5$ , 结构如表 2 参数所示。  $CNN_1$  有 3 个 Conv-Maxpool 层,  $CNN_2 \sim CNN_4$  有 4 个 Conv-Maxpool 层且有不同的全连接层,  $CNN_5$  有 5 个 Conv-Maxpool 层。最终的分类结果也显示在表 3 中。



表 3 不同 CNN 的结构及其诊断结果

	CNN <sub>1</sub>	CNN <sub>2</sub>	CNN <sub>3</sub>	CNN <sub>4</sub>	CNN <sub>5</sub>
$L_1$	Conv-Maxpool	Conv-Maxpool	Conv-Maxpool	Conv-Maxpool	Conv-Maxpool
$L_2$	Conv-Maxpool	Conv-Maxpool	Conv-Maxpool	Conv-Maxpool	Conv-Maxpool
$L_3$	Conv-Maxpool	Conv-Maxpool	Conv-Maxpool	Conv-Maxpool	Conv-Maxpool
$L_4$	-	Conv-Maxpool	Conv-Maxpool	Conv-Maxpool	Conv-Maxpool
$L_5$	-	-	-	Conv-Maxpool	Conv-Maxpool
FC <sub>1</sub>	1 024	1 024	1 024	1 024	1 024
FC <sub>2</sub>	1 024	1 024	1 024	1 024	512
FC <sub>3</sub>	-	-	512	1024	-
Acc/(%)	91.79	99.33	98.04	95.49	99.96

由表 3 中可以看出随着 Conv-Maxpool 层数的增加故障识别准确率逐渐上升,说明更深的卷积网络可以获得更好的特征表示。此外,不同的全连接层数及隐藏神经元数也决定着诊断的准确度。

图 6 和图 7 分别展示出了不同 CNN 结构的损失值变化及在测试数据集上的诊断结果。

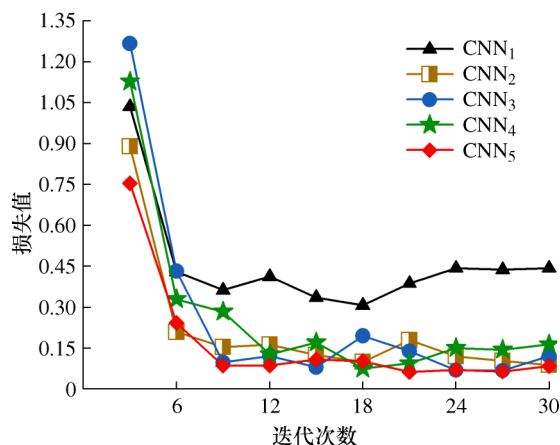


图 6 不同 CNN 结构的损失对比图

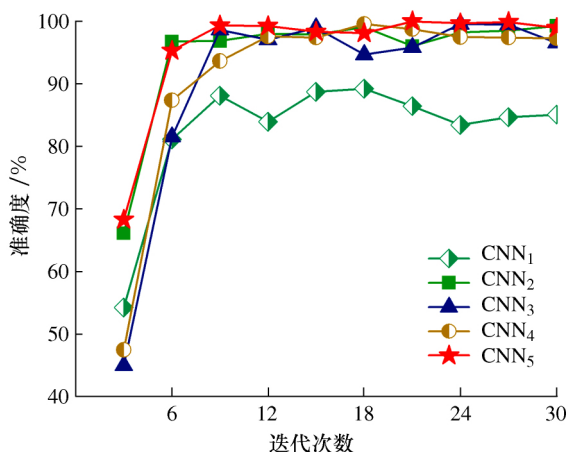


图 7 不同 CNN 结构识别准确度对比

由图 6 可以看出, CNN<sub>5</sub> 的收敛速度最快且在训练后期波形稳定达到最优; CNN<sub>1</sub> 统计结果波动较大且难以收敛; CNN<sub>2</sub>~CNN<sub>4</sub> 相对来说较为稳定,但整体效果略次于 CNN<sub>5</sub>。图 7 结果表明, CNN<sub>1</sub> 的识别结果最差而 CNN<sub>5</sub> 识别较为准确且稳定性好。由 CNN<sub>1</sub>、CNN<sub>2</sub> 和 CNN<sub>5</sub> 可以看出随着卷积网络层数增加诊断效果变好, CNN<sub>2</sub> 与 CNN<sub>4</sub> 的比较展示出全连接对诊断结果的影响。综合考虑本文选用了 CNN<sub>5</sub> 的网络结构。

### 3.3.2 时间步长的影响

时间步长  $N$  是与分类准确度直接相关的重要参数之一,同时也决定了循环神经网络的复杂度。通常时间步长越长信息的利用率越高,即故障诊断越准确。但是  $N$  过大会消耗大量的训练时间,分类更复杂。为了研究  $N$  的变化对模型的影响,进行了多组实验。采用单变量控制方法,对  $N$  分别为 5, 15, 20, 30, 50 的情况进行多次重复测试,结果如图 8 所示。

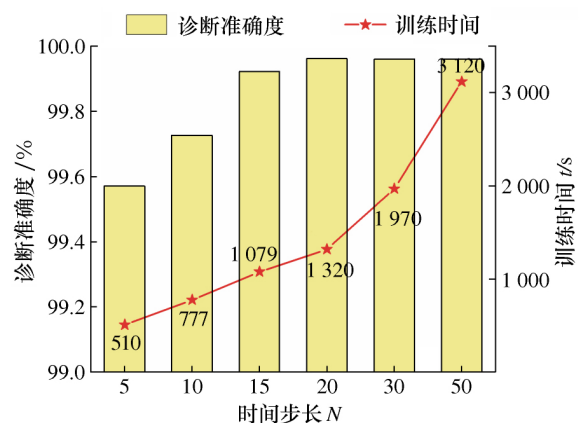


图 8 不同时间步长对识别准确率的影响

由统计结果可知诊断准确度与时间步长正相关,呈逐渐上升趋势。理论上如果想得到更高的模型诊断精度,增大  $N$  即可。但是实际上  $N$  受样本的约束,  $N$  越大,数据处理的时间越长。由图 8 中曲线可看出,训练时间随着  $N$  的增大近似指数增长,时间从 510 s 增加到 3 120 s,当  $N$  大于 20 时,训练时间急剧增长,而故障诊断准确度在  $N$  大于 20 后趋于稳定。综合考虑多种因素,选取时间步长为 20。

### 3.4 不同方法诊断效果的对比

#### 3.4.1 与传统方法的比较

为了验证本文方法的有效性,选择其他传统方法与其进行比较。传统智能故障诊断中多用 BP 和 SVM,因此选用这两种方法的处理结果进行对比。BP 以原始信号为输入,其他网络参数设置为:网络层数设为 3 层,隐藏节点设置为 512,全局参数与本

文方法一致。故障信号常用时频统计参数作为 SVM 输入,惩罚参数为 2,核函数参数为 1。所提出方法的参数设置与上文实验一致,实验结果比较如图 9 所示。

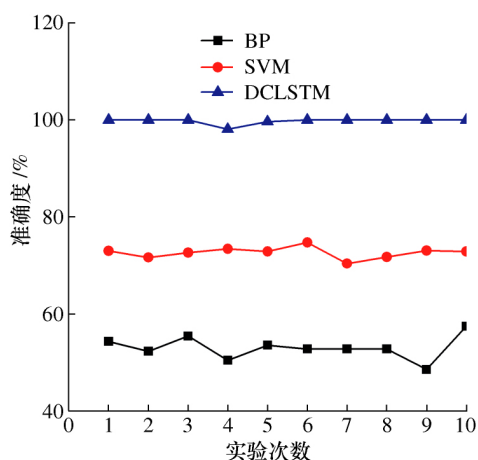


图 9 与传统方法的诊断准确度对比

从图 9 可以看出,与传统方法相比,本文所提出的基于 DCLSTM 模型的方法 10 次测试结果的准确率都超过 99%,且波动较小、稳定性强。BP 神经网络的效果最差,波动较大,说明其泛化能力较差。SVM 方法虽然结果相对稳定,但是诊断准确率远低于本文提出方法。这一结果表明所提出的方法有更好的学习分析能力,可靠性高,有效的提高了网络的鲁棒性和泛化性能。随机 10 次实验故障诊断准确率及标准差统计结果如表 4 所示。

表 4 不同方法 10 次实验平均诊断结果

方法	平均准确率 / %	准确率的标准差 / %
本文提出方法	99.961	0.59
常用统计特征+SVM 方法	72.692	1.17
BP 方法	53.085	2.47

3 种方法相比,从准确率角度,本文提出的方法故障诊断准确率高达到了 99.961%,BP 方法最低,平均准确率只有 53.085%。从诊断稳定性角度,本文提出方法稳定性最高,其标准差为 0.59%。实验表明所提出方法在故障诊断方面的表现优于传统智能方法。

### 3.4.2 神经网络的重要作用

为了验证文中 LSTM 的重要作用,将 DCLSTM 模型与普通的深度 CNN 模型识别方法进行比较,其结构与 CNN5 相同。对于两种情况,卷积层中的卷

积核数量和全连接层中的神经单元是相同的。DCLSTM 与普通 CNN 之间的主要区别在于全连接层的拓扑不同。DCLSTM 中卷积神经网络连接的是 LSTM 层,而在一般 CNN 模型的情况下,卷积网络连接的是全连接层,网络结构不同导致模型性能的差异。两种模型的诊断结果与方差对比分别展示在如图 10 和表 5 中。

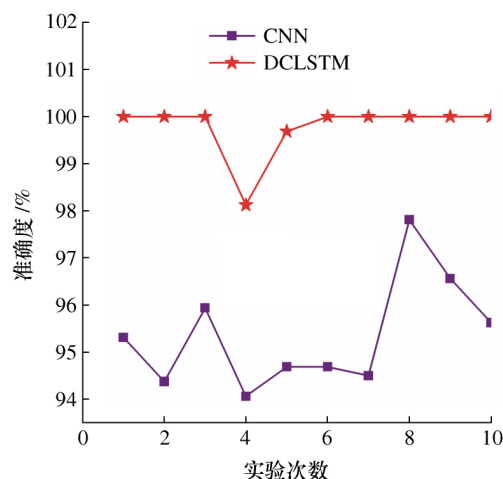


图 10 与典型 CNN 的诊断准确度对比图

表 5 深度学习方法平均诊断结果

网络结构	准确率 / %	标准方差 / %
DCLSTM	99.96	0.5902
单一 CNN	95.35	1.1626

由图 10 可看出,单独使用 CNN 方法诊断准确率波动较大,本文方法的诊断准确度整体略高于普通 CNN 模型且相对稳定。由此可知循环神经网络对于时序数据分类是有效的,可以充分利用数据间的时间动态,提高了模型稳定度。

另一方面,模型的性能评估不仅仅取决于最终的准确率,每类的识别精度也是重要指标。 $F_1$  分数被广泛应用于分类任务中作为性能评估指标,它说明了每个类别的分类性能,定义如下:

$$F_1 = \frac{2 \text{precision} * \text{recall}}{\text{precision} + \text{recall}} \quad (10)$$

$$\text{precision} = \frac{TP}{TP + FP} \quad (11)$$

$$\text{recall} = \frac{TP}{TP + FN} \quad (12)$$

式中:  $TP$  为正确分类样本数;  $FP$  为将其他样本分类为所选样本数;  $FN$  所选样本分类为其他样本数; 普通 CNN 模型和本文所提出方法的  $F_1$  分数列于表 6 中。

表6 不同方法的  $F_1$  分数

类别	本文提出方法			CNN 方法		
	准确率/%	召回率/%	$F_1$ 分数	准确率/%	召回率/%	$F_1$ 分数
1	100	100	100	98	85	91
2	100	100	100	89	99	93
3	100	96	98	99	100	98
4	97	100	98	89	100	99

表6中统计了每个故障类别的准确率、召回率和  $F_1$  分数。对于分类任务应在保证准确率的情况下尽可能的提高召回率。从表中可以看出,所提出的方法对于每种故障类别都有较高的  $F_1$  分数,不同故障类别之间波动非常小,并且准确率和召回率都在95%以上。从以上分析可知,在循环神经网络作用下,每类故障都有准确的判断,使得整个网络的泛化能力和稳定性都有所提升。

DCLSTM 结合两种网络模型,一定程度上综合了全局特征和局部特征,在故障诊断方面表现优于单一的 CNN。本文所提出的模型各项指标的结果均优于常用的方法,比单一 CNN 方法准确率提升了 5.2%,  $F_1$  分数增加了 9%。本文模型对各类故障分类基本准确,说明从深度 CNN 提取的有用特征中, LSTM 可以获得更准确的时序规律,避免了重要信息的丢失。以上实验结果验证了本文通过 CNN 与 LSTM 的结合,较好地处理实际轴承故障诊断规律。

#### 3.4.3 不同负载条件下的性能评估

实际应用中,机械轴承会在不同条件下工作。当工作条件改变时,测得的振动信号也会改变,使得故障诊断更加困难,本节中将进行实验以评估所提出方法在不同负载条件下的性能。如 3.1 所述,对 12 k 不同工况建立 4 个数据集: A, B, C, D, 选择 DE 数据进行训练比较。每个电机负载下样本数量为 2 364,同样采用平均准确度作为评估标准。结果显示在表 7 和图 11 中。

表7 不同负载下不同方法的比较结果

方法	0	1 hp	2 hp	3 hp
SVM	70.792	72.870	73.592	72.201
CNN	94.238	93.828	91.602	97.598
DCLSTM	99.883	99.805	99.219	99.688

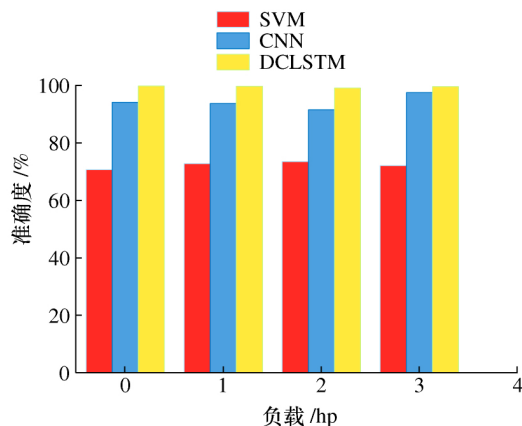


图 11 不同负载下的诊断精度

从诊断结果看,所提出的 DCLSTM 诊断精度在不同的电机负载下仍保持最高,且准确波动仅有 0.66%。而 SVM 和 CNN 的波动分别为 2.8%、5.99%。该结果证明所提出方法在不同工况下故障诊断的高效性和稳定性。本文提出的 DCLSTM 模型在变负载下,较其他方法有更高的准确度。这说明引入 LSTM 在处理时序序列有较好的改进效果。

## 4 结论

本文提出深度 CNN 结合 LSTM 构建深度网络进行轴承故障诊断的方法。该方法一方面充分利用了 CNN 实现轴承故障的自适应特性提取,另一方面充分挖掘故障演变过程中的时间序列信息,实现了更为准确的轴承故障诊断,并采用西储大学轴承数据集数据验证了方法的有效性。相较于传统的 BP 神经网络和 SVM 等智能诊断方法,所提出方法的主要优势在于无需手动特征提取即可完成高精度诊断,摆脱对先验知识的依赖。除此以外,利用循环神经网络的时序分析特性减少信息丢失,深层次模型能更好地学习有效特征提高诊断准确率。在不同负载情况下,训练模型的鲁棒性和稳定性也比较好。本文提出的 DCLSTM 方法可用于滚动轴承故障诊断。

## [参考文献]

- [1] LI Y F, LIANG X H, ZUO M J. Diagonal slice spectrum assisted optimal scale morphological filter for rolling element bearing fault diagnosis [J]. Mechanical Systems and Signal Processing, 2017, 85: 146-161



- [2] ZHANG Z Y, WANG Y, WANG K S. Fault diagnosis and prognosis using wavelet packet decomposition, Fourier transform and artificial neural network [J]. Journal of Intelligent Manufacturing, 2013, 24(6): 1213-1227
- [3] 马伦, 康建设, 孟妍, 等. 基于 Morlet 小波变换的滚动轴承早期故障特征提取研究 [J]. 仪器仪表学报, 2013, 34(4): 920-926  
MA L, KANG J S, MENG Y, et al. Research on feature extraction of rolling bearing incipient fault based on Morlet wavelet transform [J]. Chinese Journal of Scientific Instrument, 2013, 34(4): 920-926 (in Chinese)
- [4] 陈彦龙, 张培林, 徐超, 等. 基于 DCT 和 EMD 的滚动轴承故障诊断 [J]. 电子测量技术, 2012, 35(2): 121-125  
CHEN Y L, ZHANG P L, XU C, et al. Fault diagnosis of rolling bearing based on DCT and EMD [J]. Electronic Measurement Technology, 2012, 35(2): 121-125 (in Chinese)
- [5] 熊景鸣, 潘林, 朱昇, 等. DBN 与 PSO-SVM 的滚动轴承故障诊断 [J]. 机械科学与技术, 2019, 38(11): 1726-1731  
XIONG J M, PAN L, ZHU S, et al. Bearing fault diagnosis based on deep belief networks and particle swarm optimization support vector machine [J]. Mechanical Science and Technology for Aerospace Engineering, 2019, 38(11): 1726-1731 (in Chinese)
- [6] LEI Y G, JIA F, LIN J, et al. An intelligent fault diagnosis method using unsupervised feature learning towards mechanical big data [J]. IEEE Transactions on Industrial Electronics, 2016, 63(5): 3137-3147
- [7] AHARON M, ELAD M, BRUCKSTEIN A. K-SVD: an algorithm for designing overcomplete dictionaries for sparse representation [J]. IEEE Transactions on Signal Processing, 2006, 54(11): 4311-4322
- [8] 蒋永华, 程光明, 阚君武, 等. 基于 NGA 优化 SVM 的滚动轴承故障诊断 [J]. 仪器仪表学报, 2013, 34(12): 2684-2689  
JIANG Y H, CHENG G M, KAN J W, et al. Rolling bearing fault diagnosis based on NGA optimized SVM [J]. Chinese Journal of Scientific Instrument, 2013, 34(12): 2684-2689 (in Chinese)
- [9] DENG S, JING B, SHENG S, et al. Impulse feature extraction method for machinery fault detection using fusion sparse coding and online dictionary learning [J]. Chinese Journal of Aeronautics, 2015, 28(2): 488-498
- [10] 雷亚国, 贾峰, 周昕, 等. 基于深度学习理论的机械装备大数据健康监测方法 [J]. 机械工程学报, 2015, 51(21): 49-56  
LEI Y G, JIA F, ZHOU X, et al. A deep learning-based method for machinery health monitoring with big data [J]. Journal of Mechanical Engineering, 2015, 51(21): 49-56 (in Chinese)
- [11] SHAO H D, JIANG H K, LIN Y, et al. A novel method for intelligent fault diagnosis of rolling bearings using ensemble deep auto-encoders [J]. Mechanical Systems and Signal Processing, 2018, 102: 278-297
- [12] 温江涛, 闫常弘, 孙洁娣, 等. 基于压缩采集与深度学习的轴承故障诊断方法 [J]. 仪器仪表学报, 2018, 39(1): 171-179  
WEN J T, YAN C H, SUN J D, et al. Bearing fault diagnosis method based on compressed acquisition and deep learning [J]. Chinese Journal of Scientific Instrument, 2018, 39(1): 171-179 (in Chinese)
- [13] INCE T, KIRANYAZ S, EREN L, et al. Real-time motor fault detection by 1-D convolutional neural networks [J]. IEEE Transactions on Industrial Electronics, 2016, 63(11): 7067-7075
- [14] JING L Y, ZHAO M, LI P, et al. A convolutional neural network based feature learning and fault diagnosis method for the condition monitoring of gearbox [J]. Measurement, 2017, 111: 1-10
- [15] 曲建岭, 余路, 袁涛, 等. 基于一维卷积神经网络的滚动轴承自适应故障诊断算法 [J]. 仪器仪表学报, 2018, 39(7): 134-143  
QU J L, YU L, YUAN T, et al. Adaptive fault diagnosis algorithm for rolling bearings based on one-dimensional convolutional neural network [J]. Chinese Journal of Scientific Instrument, 2018, 39(7): 134-143 (in Chinese)
- [16] LU C, WANG Z Y, ZHOU B. Intelligent fault diagnosis of rolling bearing using hierarchical convolutional network based health state classification [J]. Advanced Engineering Informatics, 2017, 32: 139-151
- [17] YANG B L, SUN S L, LI J Y, et al. Traffic flow prediction using LSTM with feature enhancement [J]. Neurocomputing, 2019, 332: 320-327
- [18] LU W N, LI Y P, CHENG Y, et al. Early fault detection approach with deep architectures [J]. IEEE Transactions on Instrumentation and Measurement, 2018, 67(7): 1679-1689
- [19] LEI J H, LIU C, JIANG D X. Fault diagnosis of wind turbine based on Long Short-term memory networks [J]. Renewable Energy, 2019, 133: 422-432
- [20] LECUN Y, BENGIO Y, HINTON G. Deep learning [J]. Nature, 2015, 521(7553): 436-444
- [21] BENGIO Y. Learning deep architectures for AI [M]. Hanover: The Essence of Knowledge, 2009: 1-127
- [22] HOCHREITER S, SCHMIDHUBER J. Long short-term memory [J]. Neural Computation, 1997, 9(8): 1735-1780
- [23] SMITH W A, RANDALL R B. Rolling element bearing diagnostics using the case western reserve university data: a benchmark study [J]. Mechanical Systems and Signal Processing, 2015, 64-65: 100-131