

# 改进 CNN-LSTM 模型在滚动轴承故障诊断中的应用<sup>①</sup>



曹正志, 叶春明

(上海理工大学 管理学院, 上海 200093)

通讯作者: 曹正志, E-mail: 1760882099@qq.com

**摘 要:** 滚动轴承的运行状态对整机工作状态影响重大, 但目前其故障诊断方法存在依赖手工特征提取、鲁棒性不高等问题. 因此, 本文提出了一种基于改进的一维卷积神经网络 (1D-CNN) 和长短期记忆网络 (LSTM) 集成的滚动轴承故障诊断方法 (1D-CNN-LSTM). 首先, 利用改进的 1D-CNN-LSTM 模型对滚动轴承 6 种不同的工作状态进行了分类识别实验, 实验结果表明提出的分类模型能够以较快的速度识别出滚动轴承的不同状态, 平均识别准确率达 99.83%; 其次, 将提出的模型与部分传统算法模型进行对比实验, 结果表明所提方法在测试精度方面有较大优势; 最后, 引入迁移学习测试模型的鲁棒性和泛化能力, 实验结果表明提出的改进模型在不同工况下有较好的适应性和高效性, 模型有较强的泛化能力, 具备工程应用的可行性.

**关键词:** 故障诊断; 卷积神经网络; 长短期记忆网络; 深度学习; 迁移学习

引用格式: 曹正志, 叶春明. 改进 CNN-LSTM 模型在滚动轴承故障诊断中的应用. 计算机系统应用, 2021, 30(3): 126–133. <http://www.c-s-a.org.cn/1003-3254/7830.html>

## Application of Improved CNN-LSTM Model in Fault Diagnosis of Rolling Bearings

CAO Zheng-Zhi, YE Chun-Ming

(Business School, University of Shanghai for Science and Technology, Shanghai 200093, China)

**Abstract:** The state of rolling bearings has a great influence on the working state of the whole machine, but the fault diagnosis method of the rolling bearings at present has some problems, such as dependency on manual feature extraction and low robustness. Therefore, we propose a fault diagnosis method of rolling bearings (1D-CNN-LSTM) based on the improved integration of 1D Convolutional Neural Network (1D-CNN) and Long Short-Term Memory (LSTM) network. Firstly, the 1D-CNN-LSTM model is used to classify and identify six different working states of rolling bearings. The experimental results indicate that the proposed classification model can identify different states of rolling bearings at a high speed, with an average identification accuracy of 99.83%. Secondly, the proposed model is compared with some traditional algorithm models and shows great advantages in measuring accuracy. Finally, transfer learning is introduced to test the robustness and generalization ability of the proposed model. The experimental results demonstrate that the model proposed in this study has good adaptability and high efficiency under different working conditions, featuring strong generalization ability and engineering application feasibility.

**Key words:** fault diagnosis; Convolutional Neural Network (CNN); Long-Short Term Memory (LSTM) network; deep learning; transfer learning

① 基金项目: 国家自然科学基金 (71840003); 上海理工大学科技发展基金 (2018KJFZ043)

Foundation item: National Natural Science Foundation of China (71840003); Technology Development Fund of University of Shanghai for Science and Technology (2018KJFZ043)

收稿时间: 2020-07-16; 修改时间: 2020-08-13, 2020-08-21; 采用时间: 2020-08-25; csa 在线出版时间: 2021-03-03

滚动轴承作为许多机械的基础零部件,其运行状态往往会影响整机机械的工作状态,对生产和安全造成直接影响<sup>[1]</sup>.有关资料记载,大型企业因滚动轴承故障而引发的一次生产线非计划停产造成的经济损失可达数千万元<sup>[2]</sup>.因此对滚动轴承进行科学有效的故障诊断具有十分重要的意义.

故障诊断方法的研究主要分为基于解析模型和基于数据驱动两方面.Hsiao等<sup>[3]</sup>提出了一种分层多模型方案来检测和隔离机器人机械手的执行器故障.代祥<sup>[4]</sup>提出了一种电网信息物理模型故障诊断优化模型,将故障问题表示成求解目标函数是极值的0-1整数规划问题,从而通过严密的数学方法来确定故障元件.基于解析模型的方法需要对故障诊断问题进行解析化表达,对于复杂度较高的系统建模难度大,且建立的模型在其他系统上的普适性较低<sup>[4]</sup>,实际推广使用具有一定局限性.近年来,随着机器学习研究的兴起,基于数据驱动的故障诊断方法已成为重点研究领域之一<sup>[5]</sup>.姚德臣等<sup>[6]</sup>将改进后的支持向量机(SVM)应用于轴承的故障诊断研究当中.Peng等<sup>[7]</sup>将主成分分析(PCA)算法应用到电动潜水泵轴的损坏原因检测中.Yang等<sup>[8]</sup>将BP人工神经网络(BPNN)应用于滚动轴承的故障诊断.这些传统的基于数据驱动的方法都取得了相对不错的效果,但随着人工智能技术的发展,现已证实这些浅层网络结构,因其特征提取能力不足,难以挖掘提取故障数据中更深层次的微小特征<sup>[9]</sup>,从而限制了诊断准确率的提升.

随着互联网、物联网等快速兴起与普及,当前社会数据的增长速度比以往任何时期都要迅猛<sup>[10]</sup>.大数据给深度神经网络提供了充足的训练“原料”,给基于数据驱动的机械智能故障诊断的深入研究和应用提供了新的机遇<sup>[11]</sup>.

深度学习理论由Hinton等<sup>[12]</sup>于2006年提出,近年来,深度学习技术因其强大的特征提取和学习能力在语音识别<sup>[13]</sup>和计算机视觉<sup>[14]</sup>领域迅速发展,并产生了许多新的突破.卷积神经网络(CNN)由LeCun等<sup>[15]</sup>于1989年提出.2012年,Krizhevsky等<sup>[14]</sup>将卷积神经网络与深度学习理论结合提出“AlexNet”网络结构.深度卷积网络能够由浅到深逐步抽象特征,自动特征提取,其独特的网络结构能有效的在保留数据特征的同时减少参数数量降低数据复杂度,并通过多层次的非线性映射关系学习深层次的故障特征<sup>[16]</sup>.将深度学习

技术应用于故障诊断领域产生了不错的效果,宫文峰等<sup>[17]</sup>通过引入全局均值池化技术代替传统CNN的全连接层部分,使用改进后的CNN算法识别轴承故障种类.杜小磊等<sup>[18]</sup>提出一种基于SSST和DCCNN的滚动轴承故障诊断方法提高了信号的时频分辨率.

上述研究都只是单独使用深度CNN进行训练,忽略了滚动轴承在发生故障时的时序特征.滚动轴承性能退化是依存于服役时间的连续演化过程,相较于常规“事后诊断”,变工况下滚动轴轻微损伤甚至早期退化状态的准确识别对于指导预测性维护工作等有更大价值<sup>[19]</sup>.

CNN有着强大的图像特征提取能力,但在处理带时序问题时准确率和效率都没有循环神经网络(RNN)高,RNN可以学习到历史信息,因而RNN更适合处理时间序列.作为RNN的变体,长短期记忆网络(LSTM)通过遗忘门、输入门和输出门等机制解决了普通RNN不易处理的远距离信息上下文依赖、梯度消失或梯度爆炸等问题.LSTM在语音识别、文本识别等方面有成功的应用,同时也被用于故障诊断领域提取故障信号时间序列的特征.Qu等<sup>[20]</sup>使用基于LSTM的深度学习对研磨系统进行故障诊断,诊断错误率小于3%.于洋等<sup>[21]</sup>使用LSTM结合迁移学习实现了多种类型工况下轴承故障声发射信号特征的自适应提取与智能识别.

针对以上分析,本文拟将CNN与LSTM的优势结合,提出一种首先使用CNN提取数据特征,再结合LSTM处理时序特征的滚动轴承故障诊断方法.为了最大程度的保留振动信号的时序特征,采用一维卷积神经网络(1D-CNN)进行特征提取,用全局池化层代替传统CNN网络架构中的Flatten层以及全连接层,从而避免Flatten操作和全连接层带来的参数特征的割裂.以达到减少人工特征提取时间、适应时序问题、提高故障诊断精度的目的.

## 1 1D-CNN-LSTM 诊断模型

### 1.1 CNN 模型

CNN模型通常包含3个主要组成部分:卷积层、池化层、全连接层.卷积层的作是通过对输入数据的局部区域与卷积核进行卷积运算,通过滑动卷积核窗口使局部感受野遍历整个输入数据.卷积计算公式如下:

$$x_i^l = f(W_i^l * X^{(l-1)} + b_i^l) \quad (1)$$

式中,  $x_i^l$  表示第  $l$  层的输出值的第  $i$  个特征;  $W_i^l$  表示第  $l$  层的第  $i$  个卷积核的权重矩阵;  $*$  运算符表示卷积运算;  $X^{(l-1)}$  为第  $l-1$  层的输出;  $b_i^l$  表示偏置项; 函数  $f$  表示输出的激活函数, CNN 通过非线性的激活函数来解决现实世界中的非线性问题, 选择整流线性单元 (ReLU) 作为卷积神经网络的激活函数。

池化层的作用是空间合并也叫做子采样或者下采样, 可以在保持最重要信息的同时降低特征图的维度。它有多种类型, 一般采用平均池化或者最大池化, 采用最大池化表达式为:

$$y_i^{(l+1)}(j) = \max x_i^j(k) \quad k \in D_j \quad (2)$$

式中,  $y_i^{(l+1)}(j)$  为经过池化后的第  $l+1$  层的第  $i$  个特征图中的元素;  $D_j$  表示第  $j$  个池化区域;  $x_i^j(k)$  表示第  $l$  层第  $i$  个特征图在池化核范围内的元素。

全连接层是一个传统的多层感知器, 在输出层使用一个 Softmax 激活函数, 主要作用就是将前面提取到的特征结合在一起进行非线性激活输出各分类的概率分布然后进行分类, 表达式为:

$$p(y_j) = \frac{\exp(y_j)}{\sum_{k=1}^m \exp(y_k)} \quad (3)$$

式中,  $p(y_j)$  为神经元经过 Softmax 的概率输出;  $p(y_j)$  为输出层第  $j$  个神经元的输出值;  $m$  为所目标分类的数量, 即轴承状态的种类数量。

## 1.2 LSTM 模型

长短期记忆网络 (LSTM)<sup>[22]</sup>, 是一种带有记忆功能的神经网络, 是 RNN 的一种变种, LSTM 对时序型数据处理具有极为优秀的表现, 被广泛应用于自然语言处理等领域。LSTM 使用输入门、输出门与遗忘门实现对信息的控制。单个 LSTM 神经元如图 1 所示, 图中  $\sigma$  表示激活函数 Sigmoid, tanh 函数用于调节数值大小, 输出范围为 -1 到 1 之间。

遗忘门用于控制先前时刻的状态是否保留到当前神经元状态, 实现对记忆的筛选。输入门将前一时刻的状态值与当前输入值输入激活函数 Sigmoid, 得到一个重要度值来决定信息的更新情况, 再通过 tanh 函数来处理前一时刻的状态值和输入信息得到候选单元状态。输出门控制单元状态的最终输出, 单元状态通过输出门的过滤, 经由 tanh 函数压缩得到单元最终输出。

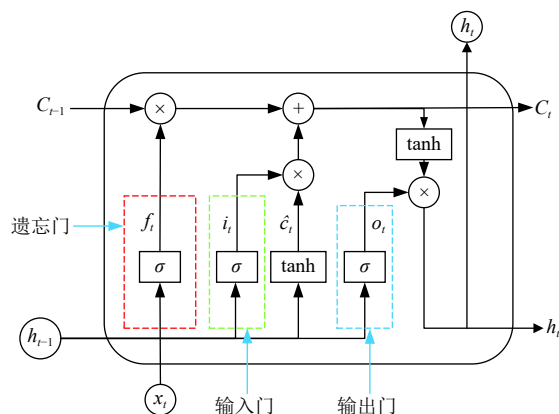


图 1 LSTM 神经元结构图

## 1.3 改进的 1D-CNN-LSTM 故障诊断模型

在使用 CNN 处理一般二维图像信号时通常会选用二维卷积核 (2D-CNN), 而滚动轴承性能退化一般是依存于服役时间的连续演化过程, 因此原始的滚动轴承故障振动信号一般为基于时间序列的一维数据。宫文峰等<sup>[18]</sup>通过人工裁剪和堆叠将一维振动信号处理成了二维图像进行诊断, 这种处理方法割裂了数据的时间序列连续性, 导致模型难以捕捉振动信号的时间序列特性。本文模型为了保留输入振动信号的时间序列信息, 以及尽量减少人工处理信息操作, 直接使用一维卷积核对一维的时间序列振动信号进行卷积处理 (1D-CNN), 避免了时间序列的割裂。

传统的 CNN 在卷积层之后通常会使用 Flatten 层降维再使用全连接层得到目标形状的特征向量进行分类或预测。Flatten 操作通过将二维矩阵按行或列展开来实现数据降维, 其在按行或列拆分图形矩阵时改变了各数据的空间位置, 从而丢失了部分有用特征。本文采用最大池化层代替 Flatten 层和全连接层作为 1D-CNN 层与 LSTM 层之间的连接, 来避免这部分特征的丢失。这种类似全卷积网络的结构支持网络采用反卷积层对最后一个卷积层的特征图进行上采样, 使它恢复到与输入图像相同的尺寸, 因此通过这种方法输入到下一步即 LSTM 层中的特征图保留了原始输入的空间信息。

本文提出的基于 1D-CNN-LSTM 的故障诊断方法网络结构如图 2 所示, 模型主要分为 1D-CNN 部分、LSTM 部分以及分类输出部分, 损失函数采用交叉熵损失函数, 梯度下降采用 Adam 优化器。输入信号为同一工况下滚动轴承不同状态的振动信号。1D-CNN 部分通过一系列的一维卷积层来提取振动信号图像特征,



并通过 MaxPooling 操作逐渐降低特征图维度. 这一操作降低了输入 LSTM 部分数据的复杂度, 既可以加快 LSTM 网络处理信号的速度, 同时又避免了 Flatten 操作, 尽可能的保留了输入数据的时序特征, 确保了模型的精度. LSTM 部分由两层 LSTM 网络构成, 借由 LSTM

独特的网络结构, 通过遗忘门、输入门、输出门的选择过滤操作可以进一步提取出 1D-CNN 部分所忽略的时间序列特征, 从而提高故障诊断模型的精度. 最后通过 Softmax 层分类输出该振动信号所表示的滚动轴承的工作状态, 模型各层具体参数如表 1 所示.

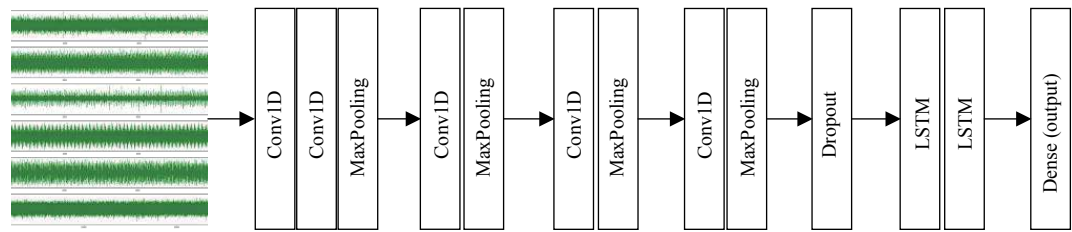


图2 模型结构图

表1 1D-CNN-LSTM 结构参数

结构部分	网络层	输出维度	参数量
1	InputLayer	(None, 400, 1)	0
2	Conv1D	(None, 400, 16)	272
	Conv1D	(None, 400, 32)	4128
	MaxPooling	(None, 50, 32)	0
3	Conv1D	(None, 50, 64)	8256
	MaxPooling	(None, 12, 64)	0
4	Conv1D	(None, 12, 256)	33 024
	MaxPooling	(None, 6, 256)	0
5	Conv1D	(None, 6, 512)	131 584
	MaxPooling	(None, 6, 512)	0
6	Dense	(None, 6, 256)	131 328
	Dropout	(None, 6, 256)	0
7	LSTM	(None, 6, 32)	36 992
	LSTM	(None, 16)	3 136
8	Dense	(None, 6)	102

由于提出的模型具有较深的网络结构, 为了增强模型鲁棒性, 防止发生过拟合现象, 模型在 1D-CNN 部分与 LSTM 部分连接处引入了随机丢弃机制 (dropout), 随机丢弃神经元之间的权重, 从而降低网络对某一单一神经元的依赖, 该操作同样可以降低输入振动信号中带来的噪声影响, dropout 层按一定的比例随机将神经元权重置为 0, 其表达式为:

$$r_i^l \sim \text{Bernoulli}(p) \tag{4}$$

$$\tilde{X}^l = r^l X^l \tag{5}$$

式中,  $r_i^l$  表示服从伯努利分布的概率向量;  $\tilde{X}^l$  表示经过随机丢弃机制后的输出.

2 实验验证

2.1 实验数据集及预处理

本实验数据来自美国凯斯西储大学 (CWRU) 的轴

承实验平台. 如图 3 所示, 实验平台包括一个 2 马力的电机, 一个转矩传感器, 一个功率计以及电子控制设备 (没有显示), 被测试轴承支承电机轴. 模拟现实中的点蚀等故障, 实验使用电火花加工技术在轴承上布置了单点故障. 实验中使用加速度采集振动信号, 传感器安放在电机壳体上. 振动数字信号的采样频率为 12 kHz, 驱动端轴承故障数据同时以 48 kHz 的采样频率采集.



图3 轴承振动数据采集试验台

外圈故障是固定不变的, 为了对该故障相对于轴承承载区域的位置对电机/轴承系统的振动响应直接影响进行定量研究, 实验中分别对驱动和风扇端的轴承外圈布置 3、6 以及 12 点钟方向的故障.

本实验选择了在同一工况下驱动端滚动轴承的 6 种不同状态的 12 kHz 采样振动信号数据作为实验数据集, 滚动轴承数据集详细信息如表 2 所示.

实验选取样本为载荷为 1 马力, 转速约为 1772 r/min 的驱动端滚动轴承的 6 种工作状态数据作为训练数据. 12 kHz 采样频率下每秒采集 12 000 个点, 转轴每转一圈传感器采集的点数为 406 个点 (12 000×60/1772≈406), 在保障数据可信度的情况下考虑到数据集的长度, 每

种工作状态的每个样本长度设置为 400 个采样点. 由于各故障数据集采样点数量不一致, 最少为 121 410 最多为 122 426, 因此全部取前 120 000 采样点, 每个样本长度为 400 个采样点, 每种工作状态包含 300 个样本. 按 8:2 的比例将 300 个样本分成训练集与测试集进行训练. 训练样本共计 1440 个, 测试样本共计 360 个.

表 2 滚动轴承故障数据集

标签	电机载荷 (马力)	电机近似 转速(r/min)	轴承状态
0	1	1772	正常
1	1	1772	滚动体故障
2	1	1772	内圈故障
3	1	1772	外圈相对位置符合区(中心位置在 6点方向)正交方向@3:00
4	1	1772	外圈相对位置符合区(中心位置在 6点方向)中心方向@6:00
5	1	1772	外圈相对位置符合区(中心位置在 6点方向)相对方向@12: 00

2.2 实验结果及分析

不同的 dropout 比率对模型的表现存在一定的影响, 该值取值一般在 0.2 到 0.5 之间, 为了选择最佳的 dropout 比率, 本文对 0.2、0.3、0.4、0.5 这 4 个常用比率分别进行了 5 组实验, 实验结果如图 4 所示.

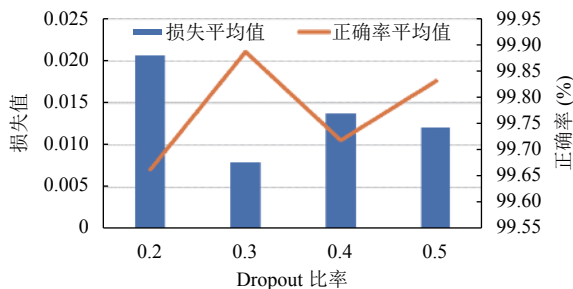


图 4 不同 dropout 比率结果对比图

如图 4 所示, 曲线表示不同 dropout 比率下的模型 5 组实验预测平均准确率, 柱形表示模型的平均损失值. 实验结果表明 dropout 比率为 0.3 时模型平均损失值最低, 且正确率最高, 因此本文模型的 dropout 比率定为 0.3.

本文采用上述模型进行了 10 次实验, 迭代次数为 50 次. 10 次实验结果表明 1D-CNN-LSTM 模型在滚动轴承故障诊断问题中最高准确率可达 100%, 平均准确率达到了 99.833%. 结果如表 3 所示.

表 3 1D-CNN-LSTM 模型实验结果

实验序号	训练集损失	训练集准确率(%)	测试集损失	测试集准确率(%)
1	0.0013	100	0.0013	100
2	0.0019	100	0.0158	99.72
3	0.0014	100	0.0014	100
4	0.0017	100	0.0035	100
5	0.0013	100	0.0012	100
6	0.0059	99.93	0.0143	99.44
7	0.0017	100	0.0019	100
8	0.0022	100	0.0025	100
9	0.0013	100	0.0013	100
10	0.0016	100	0.0404	99.17
平均值	0.002 03	99.993	0.008 36	99.833

第 10 次实验的训练损失率下降曲线以及正确率曲线如图 5 所示. 随着训练迭代次数增加, 损失率下降, 准确率逐步上升, 模型表现良好.

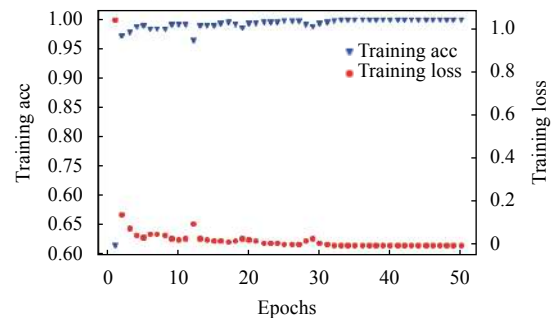


图 5 训练损失和训练精度

为验证本方法在故障诊断精度上的优势, 本文利用相同数据集使用不同的算法模型另外进行了 5 组对比实验, 每个模型运行 5 次, 迭代次数均为 50, 结果如表 4 所示. 实验 1 采用本文所提出的改进 1D-CNN-LSTM 模型; 实验 2 采用未改进的 1D-CNN-LSTM 模型, 该模型 CNN 与 LSTM 的连接部分采用了传统的 Flatten 层和全连接层; 实验 3 单独使用 1D-CNN 模型; 实验 4 单独使用 LSTM 模型; 实验 5 单独使用 2D-CNN 模型; 实验 6 使用 2D-CNN 与 LSTM 组合的模型. 实验结果表名本文所提出的改进 1D-CNN-LSTM 模型在故障诊断准确率上有最好的表现, 准确率达到了 99.83%.

由实验 1 和实验 2 对比可以发现, 改进后的 1D-CNN-LSTM 网络在精度和训练速度都有更好的表现, 通过卷积池化层连接 CNN 和 LSTM 两部分网络相对于 Flatten 层和全连接层来说输入信号的有效特征保留的更加全面, 降维效果也更加优秀; 实验 1 和实验 3 结果对比可以看出, 在引入了 LSTM 后, 模型精度确实有相应提高;

实验 1 和实验 4 对比可以发现, 通过卷积操作降低特征图的维度可以大大加快 LSTM 模型的训练速度; 实验 3 和实验 5 对比可以发现, 一维卷积网络在处理一维的滚动轴承振动信号数据方面可以保留更多有效的特征, 在故障诊断精度方面比二维的卷积网络更具优势; 实验 1 和实验 6 对比可以看出, 相对与二维卷积网络改良后的一维卷积网络结构可以保留下更多可以被 LSTM 所提取的时间序列特征, 从而提高模型的诊断精度。

表 4 对比实验结果

实验序号	模型名称	训练集平均准确率(%)	测试集平均准确率(%)	平均训练时间(s)
1	改进 1D-CNN-LSTM	99.99	99.83	48.89
2	传统 1D-CNN-LSTM	99.68	99.72	52.78
3	1D-CNN	99.93	99.72	40.28
4	LSTM	98.33	98.06	510.24
5	2D-CNN	99.92	98.83	25.00
6	2D-CNN-LSTM	99.94	99.00	35.54

首先通过 1D-CNN 提取特征并简化特征图维度再输入 LSTM 进行时序特征提取的方法比直接使用 LSTM 进行故障诊断训练速度减少了 461.35 s. 在引入 LSTM 后改进的 1D-CNN-LSTM 模型训练时间仅增加了 8.61 秒. 实际使用环境中一般采用已训练好的模型对现有故障进行诊断分类, 且对模型精度的要求远高于训练速度, 本文所提出的模型在对包含 360 个样本的测试集进行诊断分类操作时所需时间不足 1 s, 可以满足绝对对数

的使用场景要求, 因此相对于精度的提高训练时长的增加是可以接受的.

以上实验分析表明, 采用 1D-CNN 与 LSTM 组合的结构, 利用全局最大池化层规避使用 Flatten 层的操作, 可以有效的保留并利用输入信号的时序特征, 从而提高模型在故障诊断时的精度; 通过 1D-CNN 提取并简化信号特征, 减少输入 LSTM 的参数量, 可以有效降低 LSTM 的训练时间, 增强模型时序特征提取能力. 因此本文所提出的模型改进方案是有意义的.

2.3 不同负载迁移实验

为验证改进的 1D-CNN-LSTM 模型的鲁棒性和泛化能力, 采用迁移学习的方法评估算法模型在不同负载下的迁移适应性, 同时可以解决一部分对训练时间有较高要求的问题.

迁移学习能够学习到以往任务中的知识和经验, 并用于新任务中. 其目的是从一个或多个源任务中抽取知识、经验, 应用于一个新的目标领域中. 本文采用基于参数的迁移学习 (parameter-transfer learning): 目标领域和源领域的任务之间共享相同的模型参数. 本次迁移实验通过冻结上文所述在 1 马力载荷下的 1D-CNN-LSTM 模型的主要参数从而保留已训练好的模型的特征提取能力, 再添加一层全连接层使其适应目标领域, 并将模型运用到 3 马力载荷下的轴承故障识别诊断中, 迁移学习模型结构如图 6 所示.

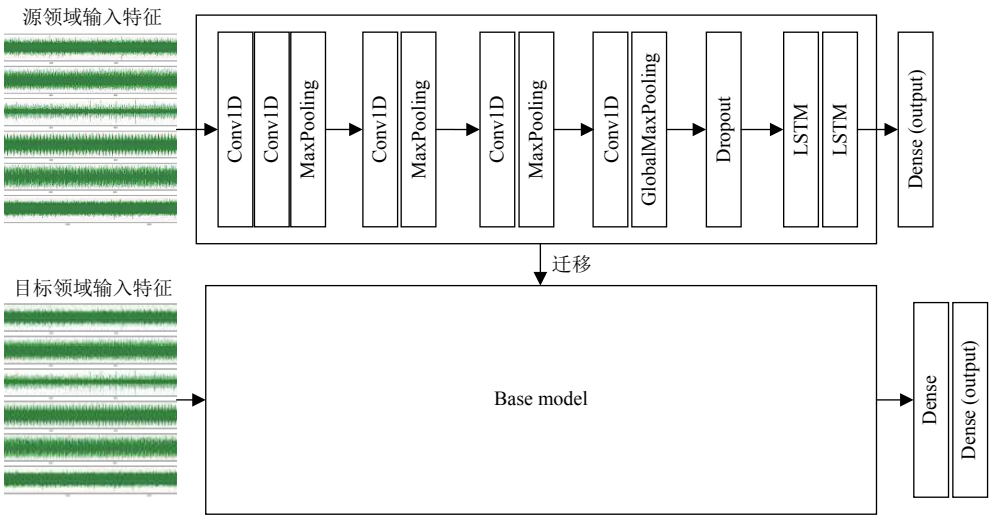


图 6 迁移学习模型结构

使用迁移学习模型对 3 马力载荷工况下得滚动轴承信号数据进行故障诊断, 模型所使用的超参数与

1 马力载荷工况下的相同, 5 次实验结果如表 5 所示, 测试集准确率达 99.72%, 表明本文所使用的算法模型



在不同工况下仍具有较高准确率, 有较强的泛化能力, 且平均训练用时仅有 18.024 s, 相比于源领域训练用时下降了 63.13%。

表5 迁移学习实验结果

实验序号	训练集损失率	训练集准确率(%)	测试集损失率	测试集准确率(%)	训练时间(s)
1	0.0856	99.72	0.0833	99.72	18.12
2	0.0790	99.79	0.0780	99.72	17.51
3	0.1059	99.58	0.1055	99.72	18.26
4	0.0624	99.86	0.0658	99.72	18.08
5	0.0911	99.79	0.0833	99.72	18.15
平均值	0.0848	99.748	0.08318	99.72	18.024

第5次实验的训练损失率下降曲线以及正确率曲线如图7所示。随着训练迭代次数增加, 损失率平滑下降, 准确率逐步上升, 在15次迭代左右, 故障诊断正确率到达相对稳定状态, 因此考虑通过减少迭代次数到20次迭代, 进一步压缩模型训练时间, 从而适应对模型训练时间有极端要求的场景。通过实验表明, 在仅20次迭代训练情况下迁移学习模型仍能达到99.72%的故障诊断准确率, 且训练用时仅为8.43 s, 与源领域相比下降了82.76%, 这对紧急情况下的快速故障诊断有着指导性的意义。

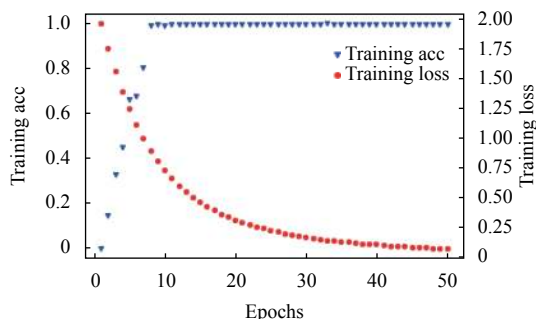


图7 迁移学习训练损失和训练精度

### 3 结束语

针对传统 CNN 以及现在故障诊断算法的不足, 本文提出了基于改进的 ID-CNN-LSTM 的深度学习算法用于电机滚动轴承的智能化故障诊断。所提方法改进了传统 CNN 模型的结构, 引入最大池化层来替代 Flatten 层和全连接层避免了特征时序特征割裂, 并引入 LSTM 来提取时序特征。该方法无需手工特征提取, 端到端的算法结构有较好的可操作性和通用性。通过对比实验, 验证了该方法故障诊断精度的优越性, 将所提的方法与单一结构的深度学习算法以及基于传统的

二维 CNN 的算法相关算法进行实验对比, 实验结果表明所提方法模型的测试精度方面具有一定优势。通过迁移学习实验, 缩短了模型训练时间并验证了该算法模型在其他工况下仍有较好的表现, 模型具有较好的泛化能力。然而, 由于提出的模型网络结构较深, 模型在训练速度上并不具有明显优势。在以后的研究中将对其进行深入研究, 提高模型的训练速度。

### 参考文献

- 唐立力, 陈国彬. 基于 MEA 优化 BP 神经网络的农机滚动轴承故障诊断. 农机化研究, 2019, 41(3): 214–218. [doi: 10.3969/j.issn.1003-188X.2019.03.038]
- 阳建宏, 黎敏, 丁福焰, 等. 滚动轴承诊断现场实用技术. 北京: 机械工业出版社, 2015. 1–5.
- Hsiao T, Weng MC. A hierarchical multiple-model approach for detection and isolation of robotic actuator faults. Robotics and Autonomous Systems, 2012, 60(2): 154–166. [doi: 10.1016/j.robot.2011.10.003]
- 代祥. 基于电网信息物理模型的故障诊断优化模型. 电力学报, 2019, 34(2): 158–166.
- 文成林, 吕菲亚, 包哲静, 等. 基于数据驱动的微小故障诊断方法综述. 自动化学报, 2016, 42(9): 1285–1299.
- 姚德臣, 杨建伟, 程晓卿, 等. 基于多尺度本征模态排列熵和 SA-SVM 的轴承故障诊断研究. 机械工程学报, 2018, 54(9): 168–176.
- Peng L, Han GQ, Pagou AL, et al. Electric submersible pump broken shaft fault diagnosis based on principal component analysis. Journal of Petroleum Science and Engineering, 2020, 191: 107154. [doi: 10.1016/j.petrol.2020.107154]
- Yang Y, Yu DJ, Cheng JS. A roller bearing fault diagnosis method based on EMD energy entropy and ANN. Journal of Sound and Vibration, 2006, 294(1–2): 269–277.
- Zhou Q, Li YB, Tian Y, et al. A novel method based on nonlinear auto-regression neural network and convolutional neural network for imbalanced fault diagnosis of rotating machinery. Measurement, 2020, 161: 107880. [doi: 10.1016/j.measurement.2020.107880]
- 雷亚国, 贾峰, 孔德同, 等. 大数据下机械智能故障诊断的机遇与挑战. 机械工程学报, 2018, 54(5): 94–104.
- 李国杰, 程学旗. 大数据研究: 未来科技及经济社会发展的重大战略领域——大数据的研究现状与科学思考. 中国科学院院刊, 2012, 27(6): 647–657. [doi: 10.3969/j.issn.1000-3045.2012.06.001]
- Hinton GE, Salakhutdinov RR. Reducing the dimensionality

- of data with neural networks. *Science*, 2006, 313(5786): 504–507. [doi: [10.1126/science.1127647](https://doi.org/10.1126/science.1127647)]
- 13 Hinton G, Deng L, Yu D, *et al.* Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups. *IEEE Signal Processing Magazine*, 2012, 29(6): 82–97. [doi: [10.1109/MSP.2012.2205597](https://doi.org/10.1109/MSP.2012.2205597)]
- 14 Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. *Proceedings of the 25th International Conference on Neural Information Processing Systems*. Lake Tahoe, NV, USA. 1097–1105.
- 15 LeCun Y, Boser B, Denker JS, *et al.* Backpropagation applied to handwritten zip code recognition. *Neural Computation*, 1989, 1(4): 541–551. [doi: [10.1162/neco.1989.1.4.541](https://doi.org/10.1162/neco.1989.1.4.541)]
- 16 LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature*, 2015, 521(7553): 436–444. [doi: [10.1038/nature14539](https://doi.org/10.1038/nature14539)]
- 17 宫文峰, 陈辉, 张泽辉, 等. 基于改进卷积神经网络的滚动轴承智能故障诊断研究. *振动工程学报*, 2020, 33(2): 400–413.
- 18 杜小磊, 陈志刚, 张楠, 等. 基于同步挤压 S 变换和深度学习的轴承故障诊断. *组合机床与自动化加工技术*, 2019, (5): 90–93, 97.
- 19 Zhang Y, Tang BP, Han Y, *et al.* Bearing performance degradation assessment based on time-frequency code features and SOM network. *Measurement Science and Technology*, 2017, 28(4): 045601. [doi: [10.1088/1361-6501/aa56c9](https://doi.org/10.1088/1361-6501/aa56c9)]
- 20 Qu XY, Zeng P, Xu CC, *et al.* RNN-based method for fault diagnosis of grinding system. *Proceedings of the IEEE 7th Annual International Conference on CYBER Technology in Automation, Control, and Intelligent Systems (CYBER)*. Honolulu, HI, USA. 2017. 673–678.
- 21 于洋, 何明, 刘博, 等. 基于 TL-LSTM 的轴承故障声发射信号识别研究. *仪器仪表学报*, 2019, 40(5): 51–59.
- 22 Hochreiter S, Schmidhuber J. Long short-term memory. *Neural Computation*, 1997, 9(8): 1735–1780. [doi: [10.1162/neco.1997.9.8.1735](https://doi.org/10.1162/neco.1997.9.8.1735)]