

doi: 10.7690/bgzdh.2019.03.010

## 基于卷积神经网络的滚动轴承故障诊断

贺思艳<sup>1</sup>, 任利娟<sup>2</sup>, 田新诚<sup>2</sup>

(1. 山东电子职业技术学院智能制造工程系, 济南 250200; 2. 山东大学控制科学与工程学院, 济南 250061)

**摘要:** 针对传统故障诊断方法诊断过程复杂、效果不佳的问题, 提出一种基于卷积神经网络的滚动轴承故障诊断方法。首先选取不同故障的振动信号进行归一化处理, 然后把 1 维的振动信号转化成 2 维的灰度图像, 利用每个元素与其相邻元素之间的关系, 并且采用重叠采样的方法加强数据集。在卷积神经网络方面利用 tensorflow 搭建网络框架, 采用 4 种不同的卷积神经网络结构对样本进行训练。为避免实验的随机性, 对每种方案进行多次训练, 采其结果的均值。根据测试集的准确率选取最好的适合轴承故障诊断的模型, 同时对网络的结构参数进行优化改进, 提高模型的识别率和运行效率。实验结果表明, 该方法可以准确地将滚动轴承的故障进行识别和分类。

**关键词:** 滚动轴承; 卷积神经网络; 故障诊断

**中图分类号:** TP202 **文献标志码:** A

## Rolling Bearing Fault Diagnosis Based on Convolution Neural Network

He Siyan<sup>1</sup>, Ren Lijuan<sup>2</sup>, Tian Xincheng<sup>2</sup>

(1. Department of Intelligent Manufacturing Engineering, Shandong College of Electronic Technology, Ji'nan 250200, China;  
2. School of Control Science & Engineering, Shandong University, Ji'nan 250061, China)

**Abstract:** This paper presents a rolling bearing fault diagnosis method based on convolution neural network for the complicated diagnosis progress, and bad effect of traditional fault diagnosis method. First of all, different fault vibration signals which were normalized. Next, converted the one-dimensional vibration signals into two-dimensional grey image, to take advantage of the relationship between each element and its neighbors, use the method of the overlap sampling to strengthen data sets. In the convolution neural network, the tensorflow was used to build a network framework to establish network structures. 4 different convolution neural network structures were used to train the samples. In order to avoid the randomness of the experiments, train many times for each scheme, the average of the results were selected to as the optimal model. According to the accuracy of the test sets, the best model for bearing fault diagnosis was selected. At the same time, the structural parameters of the network were optimized to improve the recognition rate and operation efficiency of the model. The experimental results showed that the method can identify and classify the faults of rolling bearings well.

**Keywords:** rolling bearing; convolution neural network; fault diagnosis

### 0 引言

滚动轴承作为机电设备中应用最为广泛的机械零部件, 也是最易损坏的部位之一。很多专家学者围绕传统的故障诊断开展了大量工作。Wade A.Smith 对西储大学的数据做出了一个基准研究<sup>[1]</sup>, 采用传统的频谱分析和包络谱分析对滚动轴承进行故障诊断。小波分析在故障诊断中也起着非常重要的作用<sup>[2]</sup>, 林京等提出了连续小波变换在故障诊断中的应用, 利用连续小波变换的方法来提取隐含的周期性脉冲成分。通过判断故障轴承中的特殊成分从而实现滚动轴承的故障诊断。

目前, 出现了许多基于机器学习和神经网络的故障诊断方法, 包括 K 近邻、SVM 等。大多数的方法操作包括 2 步: 特征提取<sup>[3]</sup>和故障分类。把提取出来的特征用于被训练和测试的特征向量。这样

造成诊断过程较为复杂, 需要较多信号处理先验知识, 而且诊断的效果也不佳。

笔者将 1 维的振动信号转化成 2 维的灰度图像, 用卷积神经网络判断故障的类别, 省去了传统方法的特征提取阶段, 提高了工作效率, 可以准确地判断滚动轴承的故障类型<sup>[4]</sup>。

### 1 数据样本的创建

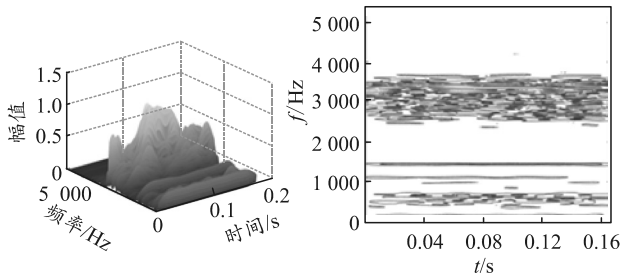
#### 1.1 1 维数据转化为 2 维图像

图 1 显示了几种振动信号的短时傅里叶变化。由图可看出: 在轴承故障诊断中, 随着时间的变化, 频率也随之发生变化, 但是很难从时域上分析不相邻样本之间的关系。如果把这些振动信号转化成 2 维图像, 就可以更好地看出给定元素和相邻元素之间存在的关系。

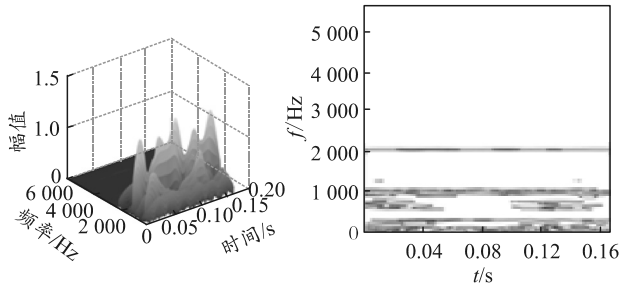
收稿日期: 2018-10-22; 修回日期: 2018-12-08

基金项目: 山东省重点研发计划项目(2016ZDJS02B03); 山东省重大科技创新工程项目(2017CXGC0601)

作者简介: 贺思艳(1967—), 女, 山东人, 硕士, 副教授, 从事电力系统及其自动化研究。



(a) 内圈故障 STFT 变换



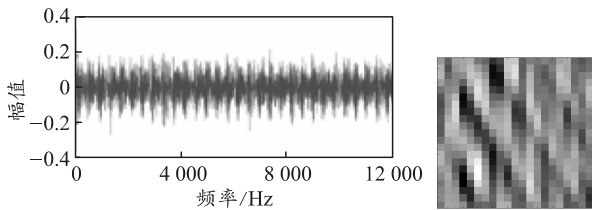
(b) 滚动体故障 STFT 变换

图 1 短时傅里叶变换

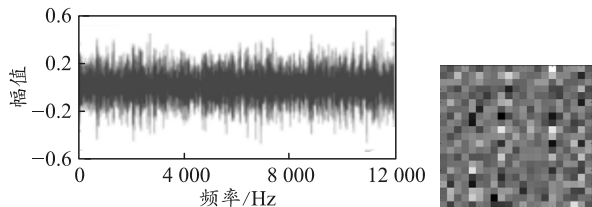
首先将振动信号中每个样本的振幅归一化成 0~1, 把每个样本的幅度归一化成对应像素的强度, 转化成  $M \times N$  的像素图片。转化的公式如下:

$$P[i, j] = A[(i-1) \times M + j] \quad (1)$$

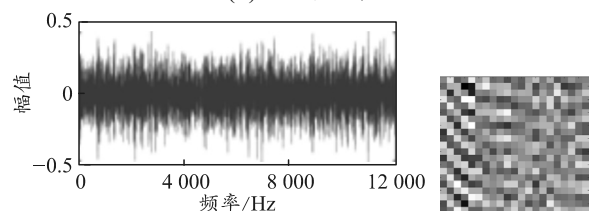
其中:  $i=1:N$ ;  $j=1:M$ ;  $P[i, j]$ 是在  $M \times N$  的振动灰度图像中有关于像素  $(i, j)$  强度的矩阵;  $A[\cdot]$ 是振动信号中对样本幅值归一化的矩阵。利用 Matlab 软件将振动幅值信号通过式(1)转化得到 2 维图像如图 2 所示。左边是 1 维的数据, 右边是转化得到的灰度图。



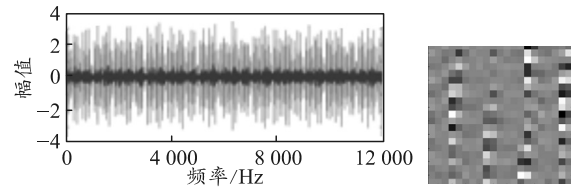
(a) 正常的信号



(b) 内圈故障



(c) 滚动体故障



(d) 外圈故障

图 2 振动信号与相关的 2 维图像

### 1.2 创建数据集

如图 3, 为了充分挖掘振动信号中存在的信息, 采用重叠采样的方法, 即将原始信号的每一段信号与后一段信号之间进行重叠采样。

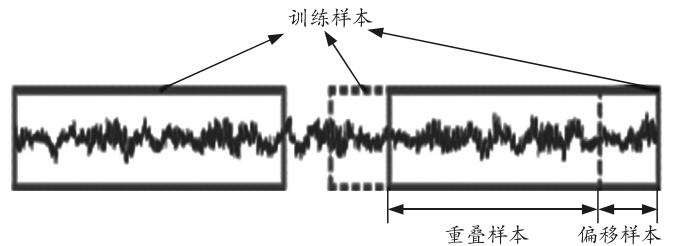


图 3 数据采集方式

如表 1 所示, 实验数据集共 4 个。每个数据集包括 8 000 个训练样本和 2 000 个测试样本。为了增强数据集, 对每个样本进行了随机翻转等操作。batch1, batch2, batch3, batch4 分别是在负载为无负载, 负载为 1hp, 2hp, 3hp 条件下的数据集。样本标签采用 One-hot 形式, 分为 4 种类型, 分别是: 没有故障、滚动体故障、内圈故障和外圈故障。

表 1 实验数据集描述

数据	数据类型	正常	滚动体	内圈	外圈
batch1	数据集	2 000	2 000	2 000	2 000
	测试集	500	500	500	500
batch2	数据集	2 000	2 000	2 000	2 000
	测试集	500	500	500	500
batch3	数据集	2 000	2 000	2 000	2 000
	测试集	500	500	500	500
batch4	数据集	2 000	2 000	2 000	2 000
	测试集	500	500	500	500
样本标签		[1000] <sup>T</sup>	[0100] <sup>T</sup>	[0010] <sup>T</sup>	[0001] <sup>T</sup>

## 2 卷积神经网络故障诊断

### 2.1 卷积神经网络及其结构

卷积神经网络的网络结构由输入层、隐藏层、全链接层和输出层组成。卷积层和池化层组成了隐藏层。卷积层采用了局部链接和权值共享。它的数学表达式可以这样表述, 如果卷积层的输入为

$$z^{(l+1)} = w^{(l)} + x^{(l)} + b^{(l)} \quad (2)$$

其中  $x^{(l)}$  是上一层的输出, 输出为

$$x^{(l+1)} = f(z^{(l)}) \quad (3)$$

其中： $w^{(l)}$ 为第  $l$  层的卷积核； $b^{(l)}$ 为偏置； $f$ 是激活函数。如果卷积层后面直接就是分类器，那么就会因为输入的维度过高造成过拟合现象，因此需要加上池化层进行降维处理。

如果定义池化函数是 lower，则表达式可为

$$X_k^{l+1} = f(w^{l+1} \text{lower}(R_k) + b^{l+1}) \quad (4)$$

其中： $w^{l+1}$ 是权重； $b^{l+1}$ 是函数的偏置项。卷积层的作用是取神经网络中每一个小的部分进行更加深入的分析，将其抽象化得到更高的特征，也就是将节点矩阵加深。池化层的作用是缩小最后全连接层的节点数。卷积层和池化层相当于图像特征的特征提取过程。全链接层和输出层可以将样本进行分类处理。输出层例如有 softmax, svm 等。通过 softmax 层可以得到当前样本属于不同种类的概率分布。总之，每个神经网络都只影响邻层的一部分神经元<sup>[5]</sup>，具有局部性，因此具有极强捕捉局部特征的能力。另一方面，通过权值共享和池化，显著降低了网络计算的复杂度。

如图 4，笔者采用 Tensorflow 建立的神经网络框架。网络结构有 5 个卷积层，5 个池化层和 2 个全连接层。将 Matlab 每生成的 100 个图像组成一个 batch，送入卷积神经网络中进行训练。在评判模型方面，采用 softmax 函数与交叉熵组合，来判定实际输出和期望输出的接近程度。

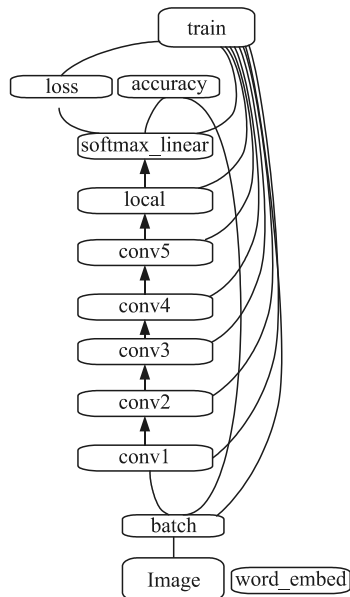


图 4 网络结构框架

### 2.2 结构参数选择与优化

为了能够更好地识别故障类型，笔者采用了 4 种方案对样本进行训练。为了消除随机因素对实验

结果造成的干扰，保证模型的稳定性，对每次实验重复 10 次，然后采其均值。如表 2， $C_1, C_2, C_3, C_4, C_5$  分别表示第 1 个卷积层到第 5 个卷积层中卷积核的个数。每个池化层过滤器的边长都是 2，移动步长为 1，且使用全 0 填充。

其中优化器选择的 RMSPropOptimizer，RMSProp 由 Geoff Hinton 提出，是一种自适应学习率方法。这种优化器的特点是收敛速度快，且稳定性很好<sup>[6]</sup>，比使用广泛的 Gradient-DescentOptimizer 优化器收敛速度要快很多，因而提高了模型的效率。学习参数中，学习率设置为 0.001，激活函数是 Relu。

表 2 不同的网络结构模型的卷积核

方案	卷积层				
	$C_1$	$C_2$	$C_3$	$C_4$	$C_5$
1	8	8	15	15	32
2	8	15	15	32	32
3	15	15	32	32	32
4	15	32	32	32	32

## 3 滚动轴承故障诊断实验

### 3.1 实验装置

实验平台的配置如下：Windows10 的 64 位操作系统，CPU 是 i5-7500@3.40 GHz，程序运行的环境是 python3.6。如图 5，实验装置是由一个电动机，一个扭转传感器/译码器，一个功率测试计组成。

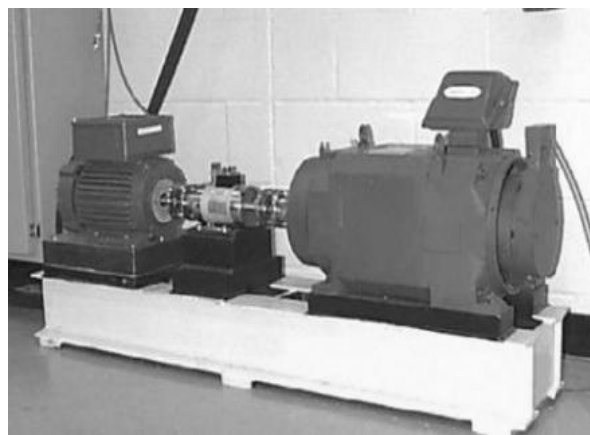


图 5 实验装置组成

滚动轴承的几何尺寸如表 3。

表 3 滚动轴承的几何尺寸 mm

型号	内圈直径	外圈直径	厚度	滚动体径
SKF6203	17.000	39.999	11.999	6.746
SKF6205	25.001	51.998	0.590	15.001

其中驱动端的轴承是 SKF6205，风扇端的轴承是 SKF6203，振动信号是数据记录仪采集得到<sup>[7]</sup>，采集通道是 16 位，采样频率为 12 kHz。

45-60.

[6] MOUSAVI S, KHORASANI K. Fault detection of reaction wheels in attitude control subsystem of formation flying satellites: A dynamic neural network-based approach[J]. International Journal of Intelligent Unmanned Systems, 2014, 2(1): 2-26.

[7] WEN X, ZHANG X W, ZHU Y P. Design of Fault Detection Observer Based on Hyper Basis Function[J]. Tsinghua Science and Technology, 2015, 20(2): 200-2005.

[8] HAN S J, CHO S B. Evolutionary neural networks for anomaly detection based on the behavior of a program[J]. IEEE Transactions on Systems Man & Cybernetics Part B Cybernetics A Publication of the IEEE Systems Man & Cybernetics Society, 2006, 36(3): 559-70.

[9] 张威, 魏炳翌, 闻新. 国内航天器故障诊断技术应用状况分析与展望[J]. 航空兵器, 2017(4): 66-74.

[10] ISLAM M M, SATTAR M A, AMIN M F, et al. A New Adaptive Merging and Growing Algorithm for Designing Artificial Neural Networks[J]. IEEE Transactions on Systems Man & Cybernetics Part B Cybernetics, 2009, 39(3): 705-722.

[11] HORNIK K, STINCHCOMBE M, WHITE H. Multilayer

feedforward networks are universal approximators[J]. Neural Networks, 1989, 2(5):359-366.

[12] SATTAR M A, ISLAM M M, MURASE K. A New Constructive Algorithm for Designing and Training Artificial Neural Networks[C]//Neural Information Processing. Springer-Verlag, 2007: 317-327.

[13] LAROCHELLE H, ERHAN D, Courville A, et al. An empirical evaluation of deep architectures on problems with many factors of variation[C]//International Conference on Machine Learning. ACM, 2007: 473-480.

[14] BOONSERM C, SURESH C S, DHADBANJAN T. A neural network based power system stabilizer suitable for on-line trainin-A practical case study for EGAT system[J]. IEEE Transactions on energy conversion, 2000, 15(1): 103-109.

[15] 张威, 王嘉轶. 一种微纳航天器编队的故障构型重构方法[J]. 航空兵器, 2017(5): 75-82.

[16] 李新, 闻新, 罗立生, 等. 基于小波神经网络的一类非线性系统的故障检测[J]. 盐城工学院学报(自然科学版), 2016, 29(1): 10-16.

[17] 王嘉轶, 闻新. 航天器故障诊断技术的研究现状与进展[J]. 航空兵器, 2016(5): 71-76.

\*\*\*\*\*  
(上接第 41 页)

### 3.2 实验结果

图 6 是 tensorboard 中的标量图表示的训练集的准确率图像。随着训练次数的增加, 4 种方案的准确率不断提高, 到达 10 000 次趋于平稳状态, 因此训练次数定为 10 000 次。方案 1 最终准确率均值是 90.62%, 方案 2 的最终准确率均值是 92.58%, 方案 3 的准确率均值达到了 93.58%, 方案 4 为 87.99%。笔者再将测试样本输入到模型中, 得到对测试样本的识别率, 方案 1 是 90.62%, 方案 2 是 92.16%, 方案 3 是 93.04%, 方案 4 是 85.41%。从测试集和训练集的精确率来看, 方案 3 取得的效果最好, 对故障类别的分类更加精确。

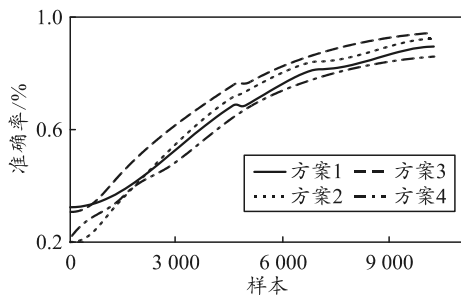


图 6 训练集准确率趋势

### 4 结束语

笔者提出了一种基于卷积神经网络的滚动轴承

故障诊断方法, 通过将 1 维的振动信号转化成 2 维的灰度图像, 采用 5 层的卷积神经网络来对其有的特征进行充分的学习。实验结果表明: 相比人为的故障识别, 该方法提高了特征提取效率与识别准确性。由于采用 RMSPropOptimizer 优化器, 增强了模型的稳定性, 提高了处理样本数据的效率, 实现了轴承故障的智能诊断。

### 参考文献:

[1] 包红燕. 基于 MEMD 和条件熵相空间重构的滚动轴承故障诊断[D]. 秦皇岛: 燕山大学, 2014: 43-47.

[2] 林京, 屈梁生. 基于连续小波变换的信号检测技术与故障诊断[J]. 机械工程学报, 2000, 36(12): 95-100.

[3] 陈果. 滚动轴承早期故障的特征提取与智能诊断[J]. 航空学报, 2009, 30(2): 362-367.

[4] 王皓, 衣同胜. 基于神经网络 CA/OS-CFAR 检测方法[J]. 兵工自动化, 2018, 37(2): 15-18.

[5] 柳小桐. BP 神经网络输入层数据归一化研究[J]. 机械工程与自动化, 2010(3): 122-123, 126.

[6] 李红利, 张晓彤, 兰立柱, 等. 基于遗传算法的 RBF 神经网络的优化设计方法[J]. 计算机仿真, 2003, 20(11): 67-69.

[7] 贾小权, 曲林伟, 王强, 等. 滚动轴承内圈局部故障碰撞的 NeimarkSacker 分岔研究[J]. 兵器装备工程学报, 2017(2): 122-126.