

DOI:10.16825/j.cnki.cn13-1400/tb.2020.03.002

基于余弦下降学习率的卷积网络轴承故障诊断

王 萌¹, 曾 艳¹, 彭 飞¹, 杨成刚¹, 王 康²

(1. 唐山工业职业技术学院, 河北 唐山 063299; 2. 北京首钢自动化信息技术有限公司, 北京 100043)

摘 要: 滚动轴承在旋转机械中起到重要作用, 其工作环境多变并存在多种失效模式。传统的轴承故障诊断方法需要依赖技术人员的经验提取时频域故障特征, 并且诊断准确率不高。提出了一种基于余弦下降学习率的卷积神经网络模型, 该模型以数据驱动方式自主提取轴承的故障特征。在训练的过程中采用余弦下降学习率, 使损失函数可以快速收敛完成分类模型的训练。应用凯斯西储大学(CWRU)滚动轴承数据集验证该方法的有效性。结果显示, 本模型能够有效识别轴承不同损伤程度以及不同故障, 整体识别率达到99.91%。

关键词: 深度学习; 变学习率; 卷积神经网络; 滚动轴承; 故障诊断

中图分类号: TH17

文献标志码: B

文章编号: 1674-943X(2020)03-0004-05

Bearing Fault Diagnosis Based on Cosine Descent Learning Rate Convolution Network

WANG Meng¹, ZENG Yan¹, PENG Fei¹, YANG Chenggang¹, WANG Kang²

(1. Tangshan Polytechnic College, Tangshan 063299, China;

2. Beijing Steel Group Automation Information Technology Co., Ltd, Beijing 100043, China)

Abstract: Rolling bearing plays a significant role in rotating machinery equipment. Its working environment is complex and there are many failure modes. Conventional bearing fault diagnosis methods depend on the experience of technicians to extract time-frequency fault features, and the diagnosis accuracy is not high. In this paper, a convolutional neural network model based on cosine descent learning rate strategy is proposed. The model extracts the fault features of bearings by data-driven method. And in the process of training, the cosine descent learning rate is used to make the loss function converge quickly and complete the training of classification model. The effectiveness of this method is verified by using the rolling bearing data set of Case Western Reserve University (CWRU). The results show that the proposed method can effectively identify different degrees of damage and different faults of bearings, and the overall recognition rate reaches 99.91%.

Key words: deep learning; variable learning rate; CNN; rolling bearing; fault diagnosis

0 引言

滚动轴承作为旋转机械的关键部件, 具有工况复杂、故障类型多的特点^[1]。旋转机械中滚动轴承是一种关键零部件并且属于易损件。其轻度损伤可能导致整机停车故障, 甚至产生难以估计的经济损失和安全事故。因此, 国内外众多研究学者积极开展滚动轴承的预测性故障维护方面的研究。及时对滚动轴承进行故障分类有利于改进旋转机械的故障模式识别并提高维修效率。但是由于滚动轴承的工作环境复杂并且伴随着大量环境噪音, 故障特征信号提取非常困难^[2]。传统的故障诊断时频域分析方法倾向于自原始信号提取时域和频域特征, 例如快速傅立叶变换^[3]、奇异值分解^[4]、小波变换^[5]和集合经验模态分解^[6]等。这些方法依靠研究者根据已有经验甄选合理的工作状态特征并根据数据特点设置最终分类器核心参数, 并获得比较理想的运行状态识别效果。然而, 这些诊断模型的泛化能力差, 工程技术人员难以选择合适的模型提取最全面的特征从而胜任不同的诊断任务。此外, 振动信号容易受到不同环境噪声的干扰, 研究人员从原始振动信

号中提取的统计特征通常无法对故障保持良好的灵敏度和稳定性^[7]。因此, 如何实现滚动轴承故障特征的自动提取和模式识别成为国内外研究学者的研究热点。

近年来, 智能滚动轴承故障诊断模型应运而生, 受到越来越多的研究关注^[8]。人工智能主要算法中卷积神经网络(Convolutional Neural Network, CNN)由于其参数量小、模型回归速度快等特点, 被广泛应用于预测性维护等工业大数据领域。胡晓依等基于深度学习理论, 有效组合卷积神经网络和SVM分类器构造适于滚动轴承故障诊断的改进型卷积神经网络诊断模型。该模型以原始轴承的振动信号为基准, 逐层学习优化, 实现特征提取与状态识别^[9]。Jiedi Sun等提出将卷积神经网络与压缩感知提取轴承振动信号结合, 构建不需要故障提取先验知识的高效轴承故障检测模型, 并取得良好效果^[10]。H.O.A. Ahmed提出根据过完备稀疏表示的稀疏自编码器对轴承振动信号进行压缩测量分类, 利用压缩感知方法, 结合深度置信网络对轴承进行工作状态识别^[11]。温江涛等提出基于压缩感知和深度学习理论, 研究用随机高斯矩阵实现轴

收稿日期: 2020-07-31

基金项目: 唐山工业职业技术学院院级课题“基于机器学习的轴承故障诊断研究”(课题编号: YJKT201949), 主持人王萌。

作者简介: 王 萌(1985-), 男, 河北唐山人, 硕士, 讲师, 主研方向为深度学习工业故障诊断。

承信号的变换域压缩采集,并将此信号输入深度神经网络,实现故障的智能诊断。该方法直接利用含有大量故障信息的压缩感知域采集信号训练深度神经网络,充分利用深度学习挖掘少量压缩采集数据中隐藏的故障信息^[12]。乔林涵基于 CV+SGDBPNN 方法的单隐层结构泛化能力和表达能力不够高效,利用特征学习对经典网络 Le Net 做出修改,提出了一种 2 卷积层串联型卷积神经网络,应对大数据集下的多故障诊断任务^[13]。赵春华等利用深度学习方法、鲸鱼优化算法 (Whale Optimization Algorithm, WOA) 和支持向量机 (Support Vector Machine, SVM) 等技术,提出了一种基于深度学习特征提取和 WOA-SVM 状态识别相结合的故障诊断模型。该模型由深度学习自适应提取轴承频域特征,并将其与鲸鱼优化算法提取的时域特征相融合,利用 WOA-SVM 对融合后的综合特征进行状态识别并取得良好实验效果^[14]。陈志刚等提出基于改进 EWT-CS 联合降噪和导联卷积网络的轴承故障诊断方法,具有较好的泛化能力和模式识别能力^[15]。近年来涌现大量基于深度学习理论建立的故障诊断模型研究成果,但均存在模型收敛速度慢、训练参数多、模型损失函数降低困难等突出问题。由于深度学习的固有模型反向传播回归计算特点,造成模型在最优解附近反复震荡的情况出现。

本文使用余弦下降学习率来训练卷积神经网络模型,构建适合轴承故障诊断的 CNN-SVM 深度学习模型;可以把加速度传感器采集到的轴承振动信号直接作为模型的输入,自适应地从振动加速度信号中挖掘分类特征。模型末端的支持向量机最终完成轴承工作状态的分类任务。本文模型的主要工作流程如图 1 所示。

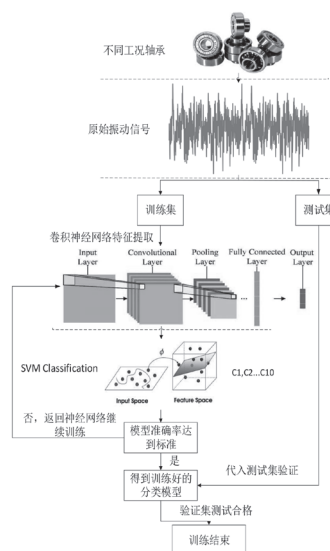


图 1 模型训练过程

1 卷积神经网络算法原理

1.1 一维卷积神经网络

CNN 是一种经典的 BP(back propagation) 神经网络,不同于全连接层神经网络。卷积层的神经元只与前一层的神经元节点部分相连,并且同一个卷积层中神经元之间的权重系数 w 和偏置系数 b 被设置为共享,从而极大减少需要训练参数的数量。CNN 网络作为近几年的研究热点,一大批优秀模型如: AlexNet、GoogleNet 和 ResNet 等被研发并应用于工程领域。图 2 是 AlexNet 的主体网络架构,主要由二维卷积层、最大值池化层和作为分类器的 Softmax 函数连接的全连接层组成。

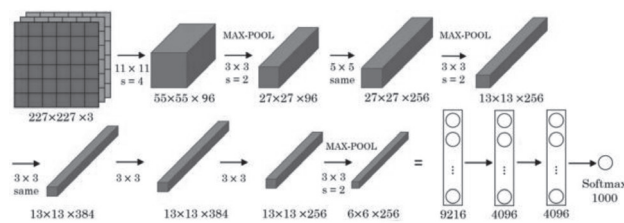


图 2 AlexNet 网络结构

经典的卷积神经网络 (CNN) 主要包括以下几个部分: 1) 输入层,接受输入数据; 2) 卷积层,利用卷积核实现自适应特征提取和特征映射; 3) 激励层,卷积本质上是一种线性运算,用 ReLU 函数等增加非线性映射; 4) 池化层,常用最大值池化来压缩数据和网络参数的量,减小过拟合; 5) 全连接层,通常在 CNN 的尾部进行重新拟合,进行特征信息的整合计算; 6) 输出层,输出网络模型的计算结果。

在卷积层中,每一层特征图的计算公式为

$$X_j^l = f\left(\sum_{i \in M_j} X_i^{l-1} * w_{ij}^l + b_j^l\right) \quad (1)$$

式 (1) 中: M_j 为输入特征图 (feature map); l 为网络的当前层数; X_i^{l-1} 为当前层第 i 个输入特征图; w 为当前层卷积核权重参数并且可以实现共享; b 为当前层偏置参数也可以实现共享; X_j^l 为第 l 层第 j 个特征图; $f()$ 代表激活函数,常用 ReLU 等; $*$ 表示离散卷积计算法则。

在误差反向传播 (Back Propagation) 的计算过程,一般应用随机梯度下降法对网络参数进行更新。由损失函数 (如交叉熵或者 MSE 等) 的导数,自动更新权重参数 w 和偏置参数 b ,降低模型产生的总误差。更新计算方法为:

$$\begin{cases} w' = w - \eta \frac{\partial E}{\partial b} \\ b' = b - \eta \frac{\partial E}{\partial b} \end{cases} \quad (2)$$

式(2)中: w' 和 b' 为梯度下降法运算更新后的权重系数和偏置系数; w 和 b 为原有的权重和偏置系数; η 为学习率决定权值更新的速度。 η 过大, 会使模型无法最终得到最优解; η 过小, 则会损害模型收敛的速度。1D-CNN 模型是将卷积神经网络中的卷积核定为一维结构, 将原始振动信号一维数据直接输入到网络模型中, 非常适合轴承等旋转零件产生的时序信号。

本文提出的卷积神经网络模型结构如下:

```
AutoEncoder(
(encoder): Sequential(
(0): Conv1d(1, 32, kernel_size=(32,),
stride=(2,), padding=(15,))
(1): ReLU()
(2): MaxPool1d(kernel_size=2, stride=2,
padding=0, dilation=1, ceil_mode=False)
(3): Conv1d(32, 16, kernel_size=(16,),
stride=(2,), padding=(7,))
(4): ReLU()
(5): MaxPool1d(kernel_size=2, stride=2,
padding=0, dilation=1, ceil_mode=False)
(6): Conv1d(16, 16, kernel_size=(8,),
stride=(2,), padding=(3,))
(7): ReLU()
(8): MaxPool1d(kernel_size=2, stride=2,
padding=0, dilation=1, ceil_mode=False)
(9): Flatten()
(10): Linear(in_features=256, out_
features=200, bias=True)
(11): Dropout(p=0.75, inplace=False)
(12): Linear(in_features=200, out_features=10,
bias=True)
)
)
```

1.2 变学习率优化方法

学习率 (Learning rate) 是设计神经网络时, 通过损失函数的梯度调整网络权重的超参数。学习率越小, 则损失函数的变化速度就越慢, 损失函数收敛时间也会延长。使用低学习率可以保证模型不会错过任何局部极小值, 但也意味着模型需要花费更

长的时间达到收敛。学习率对模型参数的优化如式(2)所示。

学习率是深度学习中关键的超参数, 合理选择的学习率能够使损失函数快速收敛到局部最小值。学习率太低会导致训练速度会非常慢, 需要消耗大量的训练时间和计算资源, 训练过程如图3所示。

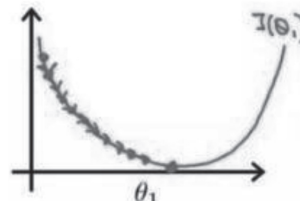


图3 学习率过小收敛速度慢

模型训练过程选择过高的学习率, 可能在损失函数最小值附近反复震荡, 从而不能到达全局最小值。因为模型总是在全局最小值附近, 但是损失函数未收敛到最小值。在训练的过程产生振荡波动, 以至于产生损失函数在最小值附近反复震荡的现象。训练过程如图4所示。

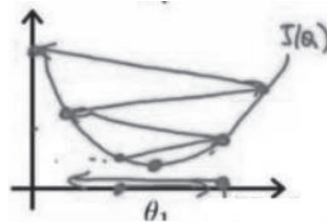


图4 学习率过大造成损失函数反复震荡

为了能够让学习率可以适应模型的训练过程, 随着训练批次的进行学习率逐渐衰减。本文采用余弦下降策略来调整学习率。实现在模型训练初期, 应用较大步长的学习率让模型可以快速迭代至最小值附近, 到了模型训练的后半期, 应用较小步长的学习率让模型可以平滑地收敛至最小值。学习率变换公式为

$$lr = LR * \cos(\text{epoch}/40) \quad (3)$$

其中初始学习率 $LR=0.01$, 最终训练批次为40, epoch 为当前训练批次, lr 为当前学习率。

2 实验结果

2.1 实验数据描述

本文采用业内公认的凯斯西储大学 (CWRU) 振动实验室的开源数据来进行验证实验。电机驱动端和风扇端的轴承作为轴承振动数据采集端, 分别在被测试轴承的内圈、外圈和滚动体上采用电火花点蚀加入单点损伤从而实现轴承的三类故障状

态, 其点蚀尺寸由小到大分别为 0.18、0.36、0.54 mm。在相近的工作负载和转速条件下, 由安装在电机驱动端上侧采样频率为 12kHz 的加速度传感器采集得到轴承振动信号。实验数据采集装置如图 5 所示。

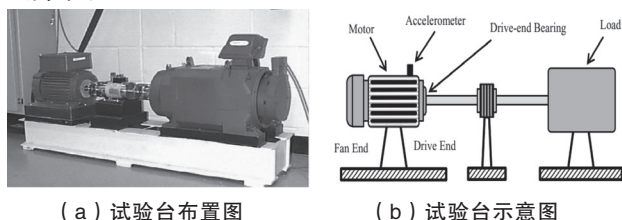


图 5 轴承实验平台

参照大部分轴承实验的惯例, 根据轴承故障的位置和尺寸将轴承的工作状态分为 10 种, 以 2 048 个振动采样数据为一个样本。每种工作状态选择 800 个样本, 共 8 000 个采样样本。全部实验数据如表 1 所示。

采集的原始振动信号如图 6 所示。

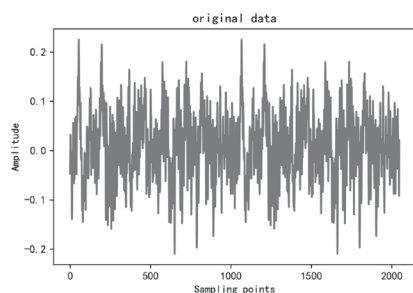


图 6 原始振动信号

2.2 实验参数设置

本文基于 Python-Pytorch 搭建网络结构, 在 Inter-i7-CPU 的计算条件下进行试验。实验中设定数据的 60% 为训练集, 剩余的 40% 为测试集, 采用分类交叉熵损失函数。模型按照以下方法设置

超参数: 训练批次=40, batchsize=64, 初始学习率=0.01。随着试验批次的增加学习率以余弦规律下降。本模型设置 SVM 惩罚因子 $C=10.33$, 核函数 $\text{gamma}=3.26$ 。

该模型的分类型准确率和 Loss 函数值如图 7、8 所示。模型训练到 10 次时, 已经几乎完全收敛, 训练集和测试集的分类型准确率都超过 99%。试验表明, 采用余弦下降学习率训练策略的卷积神经网络模型提取出的故障特征在 SVM 分类器中可以准确快速判断轴承的故障类型。该方法训练得到的模型具有较快的收敛速度和较高的分类型准确率, 并且具备较强的模型泛化能力。

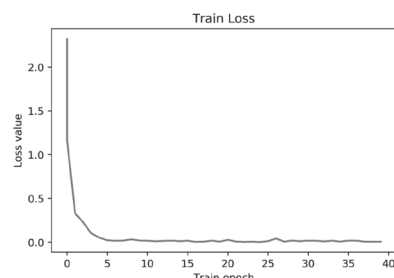


图 7 模型训练损失函数变化

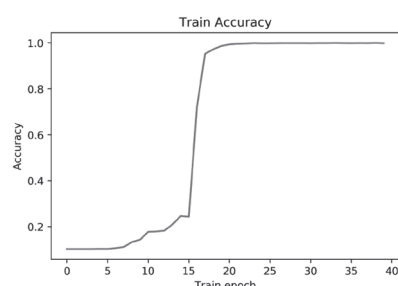


图 8 模型分类准确率变化

2.3 模型的训练效果对比

为了进一步验证该轴承故障诊断模型的泛化能力, 将该模型与几种基于传统深度学习和时频

表 1 轴承采样数据

类型	轴承工作状态	故障直径 (mm)	数据长度	样本数量	负荷 (hp)	转速(rpm)
1)	正常轴承	0	2048	800	3	1730~1797
2)	内圈轻微磨损	0.18	2048	800	3	1730~1797
3)	滚动体轻微磨损	0.18	2048	800	3	1730~1797
4)	外圈轻微磨损	0.18	2048	800	3	1730~1797
5)	内圈中度磨损	0.36	2048	800	3	1730~1797
6)	滚动体中度磨损	0.36	2048	800	3	1730~1797
7)	外圈中度磨损	0.36	2048	800	3	1730~1797
8)	内圈严重磨损	0.54	2048	800	3	1730~1797
9)	滚动体严重磨损	0.54	2048	800	3	1730~1797
10)	外圈严重磨损	0.54	2048	800	3	1730~1797

域分析的轴承分类方法做对比试验^[9]。本文根据引言各文献提出的方法,对信号通过经验模态分解(empirical mode decomposition)进行分析,提取信号的各阶模态特征构造 SVM 的输入向量;将小波分解得到的相对小波能量作为输入分类特征,通过 SVM 判定轴承工作状态;提取波形因子、裕度因子、峭度、均方根频率、频率标准差等 8 个时频域特征,输入 SVM 分类器进行工作状态分类。搭建以上几种算法的分类模型,以相同试验数据进行验算,在测试数据集上得到的准确率如表 2 所示。相较于其他几种方法, CNN-SVM+ 余弦下降学习率训练模型的计算准确率更高。

表 2 不同模型下的轴承故障诊断准确率

诊断方法	测试集 (%)
EMD+SVM	91.94
时频特征+SVM	90.52
小波+SVM	92.21
CNN-SVM	99.82
CNN-SVM+变学习率	99.91%

3 结论

本文提出的余弦下降学习率的卷积神经网络模型,可以在提取特征后配合 SVM 分类器准确判断轴承的故障类型。该方法训练得到的模型具有较快的收敛速度和较高的分类准确率,并且具备较强的模型泛化能力。在后续的研究中可以考虑进一步优化学习率下降策略,采用更加复杂并且适合轴承故障检测的学习率下降方法,比如指数型学习率衰减等方法使得学习率下降更为平滑,加快模型的收敛速度和诊断准确率。

【参考文献】

- [1] Xinpeng Zhang, Neil hu, Lei Hu, LingChen. A bearing fault diagnosis method based on the low-dimensional compressed vibration signal[J]. Advances in Mechanical Engineering, 2015, 7(07):1-12.
- [2] 吕跃刚,何洋洋. EWT 和 ICA 联合降噪在轴承故障诊断中的应用[J]. 振动与冲击, 2019, 38(16): 42-48.
- [3] 解晓婷,李少波,杨观赐,刘国凯,姚雪梅. 基于 FFT 与 CS-SVM 的滚动轴承故障诊断[J]. 组合机床与自动化加工技术, 2019(04): 90-94.
- [4] Hou F, Chen J, Dong G. Weak fault feature extraction of rolling bearings based on globally optimized sparse coding and approximate SVD[J]. Mechanical Systems and Signal Processing, 2018(111): 234-250.
- [5] Jiang Y, Zhu H, Li Z. A new compound faults detection method for rolling bearings based on empirical wavelet transform and chaotic oscillator[J]. Chaos Solitons & Fractals, 2016(89):8-19.
- [6] Yan D, Yingpin C, Guoying M, et al. Fault Severity Monitoring of Rolling Bearings Based on Texture Feature Extraction of Sparse Time-Frequency Images[J]. Applied Sciences, 2018, 8(09):1538.
- [7] Tianchen Liang, Shuaipeng Wu, Wenjing Duan, Ruishan Zhang. Bearing fault diagnosis based on improved ensemble learning and deep belief network[J]. Journal of Physics: Conference Series, 2018, 1074(1).
- [8] 刘慧斌,李少波,张安思,杨万里,黄金苑,姚勇. 基于深度卷积神经网络的轴承多故障诊断研究[J]. 组合机床与自动化加工技术, 2020(05):12-16.
- [9] 荆云建,胡晓依,宋志坤,侯银庆. 基于 CNN-SVM 的深度卷积神经网络轴承故障识别研究[J]. 振动与冲击, 2019, 38(18):173-178.
- [10] Sun J, Yan C, Wen J. Intelligent Bearing Fault Diagnosis Method Combining Compressed Data Acquisition and Deep Learning[J]. IEEE Transactions on Instrumentation and Measurement, 2017, 67(99):185-195.
- [11] Ahmed H.O.A, Wong M.L.D, Nandi A.K. Intelligent condition monitoring method for bearing faults from highly compressed measurements using sparse over-complete features[J]. Mechanical Systems & Signal Processing, 2018, 15(99):459-477.
- [12] 温江涛,闫常弘,孙洁娣,乔艳雷. 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(01):171-179.
- [13] 乔林翰. 基于卷积神经网络的电机故障诊断方法研究[D]. 兰州:兰州理工大学, 2018:7.
- [14] 赵春华,胡恒星,陈保家,张毅娜,肖嘉伟. 基于深度学习特征提取和 WOA-SVM 状态识别的轴承故障诊断[J]. 振动与冲击, 2019, 38(10):31-37.
- [15] 陈志刚,杜小磊,张楠,张俊玲. 基于改进 EWT-CS 联合降噪和导联卷积网络的轴承故障诊断[J/OL]. 哈尔滨工程大学学报:1-10[2020-08-30]. <http://kns.cnki.net/kcms/detail/23.1390.U.20200320.1408.012.html>.