

基于一维卷积神经网络的齿轮箱故障诊断

吴春志, 江鹏程, 冯辅周, 陈 汤, 陈祥龙

(陆军装甲兵学院 车辆工程系 北京 100072)

摘 要: 传统故障诊断方法通常需要先人工提取特征再用模式识别方法进行分类,难以解决端到端故障诊断的问题。为此,提出了一种利用一维卷积神经网络的齿轮箱故障诊断模型。其特点是可以直接从原始振动信号中学习特征并完成故障诊断。采用 PHM 2009 Challenge Data 和某型坦克变速箱的复合故障数据对三种传统模型和一维卷积神经网络模型进行测试。结果表明,1-DCNN 模型对单一和复合故障诊断准确率均高于传统诊断方法。

关键词: 卷积神经网络; 故障诊断; 齿轮箱; 特征学习

中图分类号: TP206

文献标志码: A

DOI: 10.13465/j.cnki.jvs.2018.22.008

Faults diagnosis method for gearboxes based on a 1-D convolutional neural network

WU Chunzhi, JIANG Pengcheng, FENG fuzhou, CHEN Tang, CHEN Xianglong

(Department of Vehicle Engineering, Academy of Army Armored Forces, Beijing 100072, China)

Abstract: Traditional diagnosis methods need to extract features manually and classify faults by pattern recognition methods. It's difficult to solve the problem of end-to-end fault diagnosis. Therefore, a one-dimensional convolution neural network (1-DCNN) model suitable for analyzing vibration data was established. The model can learn features directly from raw vibration data and complete fault diagnosis in succession. Three traditional models and the 1-DCNN model were tested with the compound fault data collected from PHM 2009 Challenge Data and a tank gearbox. The results show that the precision of the 1-DCNN model for single and complex fault diagnosis is higher than that of traditional diagnostic methods.

Key words: convolutional neural network; faults diagnosis; gearbox; feature learning

齿轮箱是机械传动系统的关键部件,其广泛应用于风力发电机组、煤矿开采设备、装甲车辆传动机构等很多重要的机械设备。齿轮箱故障可能会造成生产停滞,甚至导致人员伤亡^[1]。因此对齿轮箱的实时状态监测与故障诊断就显得至关重要^[2]。

现有的实时状态监测大都基于振动信号的频谱分析,在多故障并发、频率耦合、重要频率模糊的情况下,很难进行诊断。而传统的信号处理方法,例如小波包变换 (Wavelet Packet Transform, WPT)^[3], EMD 分解^[4], 提升小波变换 (Lifting Wavelet Transform, LWT)^[5]等需要人工提取故障特征,再通过模式识别方法,例如人工神经网络 (Artificial Neural Network, ANN)^[6]、支持向量机 (Support Vector Machine, SVM)^[7]等进行故障诊断,虽然取得了一定的效果但是难以保证诊断的实时性,而且人工提取的某些特征往

往只对特定数据有效,诊断方法难以推广。

作为机器学习的一个重要分支,深度学习近年来借助计算设备的革新已经在各个领域有了广泛的应用^[8]。卷积神经网络 (Convolutional Neural Network, CNN) 是典型的深度学习方法,其最大的优势在于无需人工选取,可以从原始图片中自动学习特征,通过训练调整优化卷积核的参数,实现图片的分类识别^[9-10]。

一些专家学者已经将 CNN 应用到齿轮箱的故障诊断中,并且取得了一定的效果。Chen 等^[11]通过绘制振动信号的特征统计图作为 CNN 的输入,实现故障的分类诊断。韩涛等^[12-14]将信号做多小波变换和时频变换,以多小波系数图和时频图作为 CNN 的输入来诊断故障。

然而上述文献利用二维卷积神经网络对振动信号的特征进行学习,没有最大化地发挥深度学习自动学习原始信号特征的优势。

本文提出了一种利用一维卷积神经网络的齿轮箱故障诊断模型,直接以振动信号作为输入,实现端到端的齿轮箱及某型坦克变速箱的故障诊断。与三种传统特征选择加模式识别的方法进行了对比,同时对模型

基金项目: 装备预研基金重点项目 (9140A27020115JB35071)

收稿日期: 2017-12-13 修改稿收到日期: 2018-02-01

第一作者 吴春志 男, 博士生, 1991 年生

通信作者 冯辅周 男, 博士后, 教授, 1971 年生

的卷积核大小、CNN 层组数等参数的选择做了较深入的分析。

1 典型卷积神经网络结构

CNN 是由哺乳动物视觉皮层细胞感受野启发而建立的一种前馈神经网络,现在已经是图片和视频识别领域应用最为广泛的方法。一个传统的 CNN 模型如图 1 所示,由一个输入层,两组交替出现的卷积层和池化层,以及全连接层组成。卷积层中每一个特征图都对应一个卷积核,这些卷积核通过一组权重来卷积前一层的输入并组成一组特征输出,成为下一层的输入。与传统滤波器需要人工设定参数不同,卷积神经网络的权值和偏置通过 BP 算法进行训练。

池化层主要是为了实现两个目的,一是大幅减少输入卷积层的空间维度,使权重参数减少 75%,从而降低计算成本。二是控制过拟合,使测试集准确度更接近训练集准确度。主流的池化方式有最大值池化、平均池化、L2-norm 池化,本模型采用的是最大值池化。

经过两组卷积层和池化层后,会接一个全连接层。与传统的神经网络类似,全连接层可以应用到不同的分类模型。全连接层后接一个隐藏层,最后由 Softmax 回归层完成分类。假设有 k 类的分类问题,Softmax 回归的输出可以计算如下^[15]:

$$O = \begin{bmatrix} P(y = 1 | x; W_1, b_1) \\ P(y = 2 | x; W_2, b_2) \\ \dots \\ P(y = K | x; W_k, b_k) \end{bmatrix} = \frac{1}{\sum_{j=1}^k \exp(W_j x + b_j)} \begin{bmatrix} \exp(W_1 x + b_1) \\ \exp(W_2 x + b_2) \\ \dots \\ \exp(W_k x + b_k) \end{bmatrix} \quad (1)$$

式中: W 和 b 分别是权重矩阵和偏置值, O 是 CNN 的最终输出。

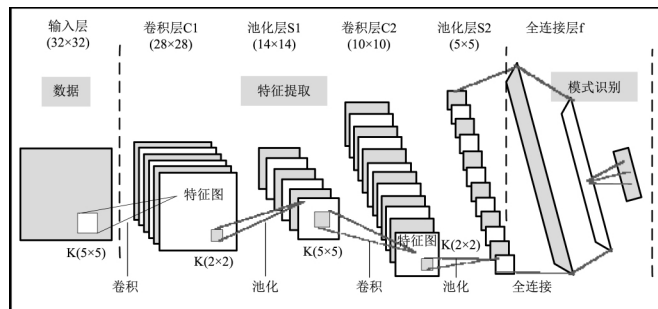


图1 典型卷积神经网络结构图

Fig. 1 Typical convolution neural network structure

区别于传统故障诊断模式,卷积神经网络将特征提取和模式识别过程整合起来,所有的参数都通过 BP 算法根据输入的训练集数据进行训练,因此模型有较强的鲁棒性和可迁移性。

2 一维卷积神经网络模型

由于振动信号和图片的维度不同,一维的信号最好用一维的模型来解决。因此,本文将图 1 中所示模型优化为图 2 所示的一维卷积神经网络模型。优化后的模型由输入层、CNN 层组(由卷积层和池化层组成)、全连接层、输出层组成。其中 CNN 层组可以根据所需分类数据的复杂度增加或减少,本模型使用 2 个 CNN 层组。第一个 CNN 层组中卷积核数量为 N_1 ,第二个 CNN 层组中卷积核数量为 N_2 ,相应卷积核的权值向量和偏置值分别为 W 和 b 。

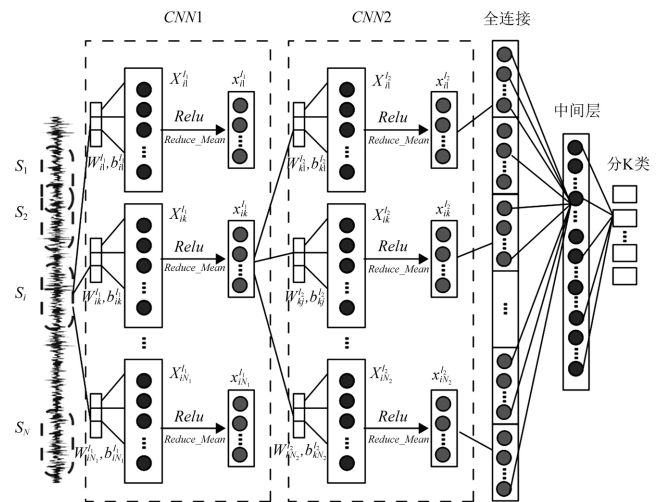


图2 优化后 1-DCNN 网络结构图

Fig. 2 Optimized 1-DCNN network structure

对截取的信号 S_i 经过第一层卷积后得到 $[X_{i1}^{l1}, X_{i2}^{l1}, \dots, X_{iN_1}^{l1}]$ 其中:

$$X_{ik}^{l1} = b_{ik}^{l1} + \text{conv } 1D(S_i, W_{ik}^{l1}) \quad (2)$$

在对 X_{ik}^{l1} 进行池化操作($\downarrow \text{Subconv}$)之前,需要激活函数增加模型的非线性特性。激活函数采用应用最广的 $\text{Relu}(\max(0, x))$ 函数。同时为了增加模型的稀疏性,提高模型训练速度,本模型在激活函数之前对 X_{ik}^{l1} 增加了去均值操作,因此第一个 CNN 层组的输出为:

$$x_{ik}^{l1} = \text{Relu}(X_{ik}^{l1} - \overline{X_{ik}^{l1}}) \downarrow \text{Subconv} \quad (3)$$

第二个 CNN 层组中的 X_{ik}^{l2} 是第一层的输出再经过卷积后得到 N_1 个输出的平均值,即:

$$X_{ik}^{l2} = \frac{1}{N_1} \sum_j f(x_{ij}^{l1}) \quad (4)$$

第二个 CNN 层的输出为:

$$x_{ik}^{l2} = \text{Relu}(X_{ik}^{l2} - \overline{X_{ik}^{l2}}) \downarrow \text{Subconv} \quad (5)$$

将 $x_{i1}^{l2}, x_{i2}^{l2}, \dots, x_{iN_2}^{l2}$ 首尾相连得到全连接层。全连接层、中间层和输出层构成一个传统神经网络。

卷积神经网络训练的目的是为了使最终的损失达到最小,模型采用的损失函数为交叉熵,优化器选用

AdamOptimizer。

3 基于 1-DCNN 的齿轮箱复合故障诊断

为了证实模型的有效性,本节采用 PHM 2009 Challenge Data^[16]中小型实验台斜齿轮的复合故障数据和结构更复杂、背景噪声更大的某型坦克的变速器单一及复合故障数据对模型进行对比分析。对一维卷积神经网络模型中卷积核长度、CNN 层组数等做了分析。四种诊断模型如图 3 所示,为了保证对比的公平性,在模式识别部分都采用中间层神经元数量相同的神经网络算法。

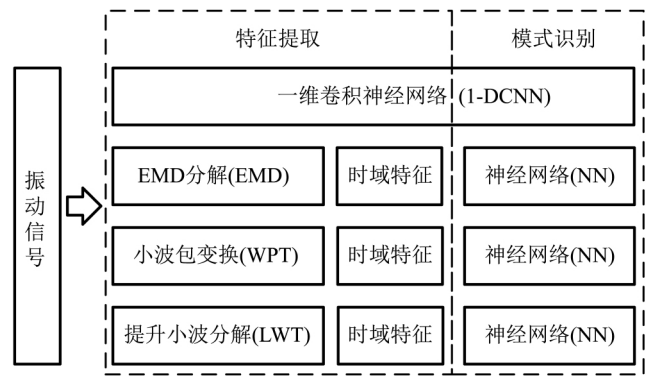


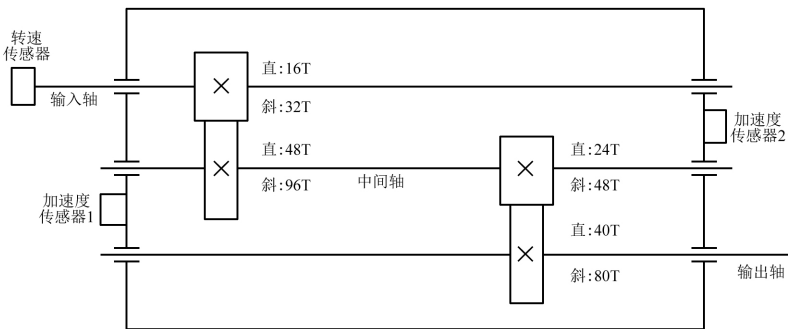
图 3 4 种故障诊断模型
Fig. 3 Four kinds of fault diagnosis model

3.1 数据准备

PHM 2009 Challenge Data 是 PHM 协会在 2009 年国际竞赛的全套齿轮箱数据,包括齿轮、轴承及轴的故障。其实验平台结构如图 4 所示,内部有三根轴(输入轴 IS,中间轴 ID,输出轴 OS)。实验采用的两组啮合齿轮有直齿轮模式和斜齿轮模式两种啮合模式。箱体两侧分别安装了振动传感器用以采集数据,采样频率为 66.67 kHz,采样时间为 4 s,采样点数为 266 656。

本文以斜齿轮故障为例,斜齿轮啮合模式下共有 8 种复合故障模式(见表 1),5 个工作频率和两种负载状态(低负载和高负载)。为了在充分利用实验数据的同时尽可能提高数据样本组数,每种故障模式包含三种相近转速的数据,采用 30 Hz、35 Hz、40 Hz 三种转速下低负载通道 1 的斜齿轮数据对几种模型进行对比。为了优化模型的运算速度,在有限数据样本内包含更多周期的信号,对原信号进行降采样 1/3 处理。取样本数据段长度为 1 600,可以包含 2 到 3 周期的振动数据。

降采样后,每隔 170 点为起点取一组样本,这样每种故障模式就可以得到 $266\ 656/3170 \times 3 \approx 1\ 500$ 组数据样本,其中选 900 组为训练集,600 组为测试集。



(a) 齿轮箱结构简图



(b) PHM 2009 Challenge Data实验台结构

图 4 齿轮箱结构简图和 PHM 2009 Challenge Data 实验台结构

Fig. 4 Schematic of the gearbox and Overview of the gearbox of PHM 2009 Challenge Data

表 1 斜齿轮复合故障模式表

Tab.1 Pattern label description of the spur gearbox

齿轮					轴承						轴	
工况	32T	96T	48T	80T	输入轴 输入端	中间轴 输入端	输出轴 输入端	输入轴 输出端	中间轴 输出端	输出轴 输出端	输入	输出
Spur 1	正常	正常	正常	正常	正常	正常	正常	正常	正常	正常	正常	正常
Spur 2	裂纹	正常	偏心	正常	正常	正常	正常	正常	正常	正常	正常	正常
Spur 3	正常	正常	偏心	正常	正常	正常	正常	正常	正常	正常	正常	正常
Spur 4	正常	正常	偏心	断齿	滚动体	正常	正常	正常	正常	正常	正常	正常
Spur 5	裂纹	正常	偏心	断齿	内圈	滚动体	外圈	正常	正常	正常	正常	正常
Spur 6	正常	正常	正常	断齿	内圈	滚动体	外圈	正常	正常	正常	不平衡	正常
Spur 7	正常	正常	正常	正常	内圈	正常	正常	正常	正常	正常	正常	键槽磨损
Spur 8	正常	正常	正常	正常	正常	滚动体	外圈	正常	正常	正常	不平衡	正常

3.2 1-DCNN 参数选择分析

3.2.1 卷积核长度

卷积神经网络每一层卷积核的数量一般以 8 的倍数递增,因此第一个 CNN 层组中卷积核的数量设为 8,第二层为 16。为了保证训练的稳定性,CNN 层组中卷积核长度依次减半。在训练 500 次后,其长度 L 与识别准确率 P 的关系如表 2 所示。卷积核长度过小识别率较低,而过大又会出现过拟合现象,在考虑不浪费计算资源的情况下,模型选择卷积核长度为 23 和 12。

表 2 卷积核长度与诊断准确率对照表

Tab.2 The relationship between the length of convolution kernel and the precision of model diagnosis				
L	5-3	11-6	17-9	23-12
p	0.94	0.95	0.97	0.993
L	29-15	35-18	41-21	47-24
p	0.993	0.98	0.975	0.963

3.2.2 CNN 层组数量

在对模型进行训练前,需要初始化卷积核的权重向量和偏置,初始化后权重向量呈标准差为 0.1 的正态分布。训练过程中权值不断更新变化,因此可以通过查看训练后权值向量分布变化来判断所需的 CNN 层组数量。图 5 是采用 4 个 CNN 层组训练 500 次后各层组权重出现频次的分布图。第 3 和第 4 层组的权重仍近似于初始的正态分布,而且模型准确率没有提升,该层组对诊断模型的贡献度不高,因此只需 2 个 CNN 层组。

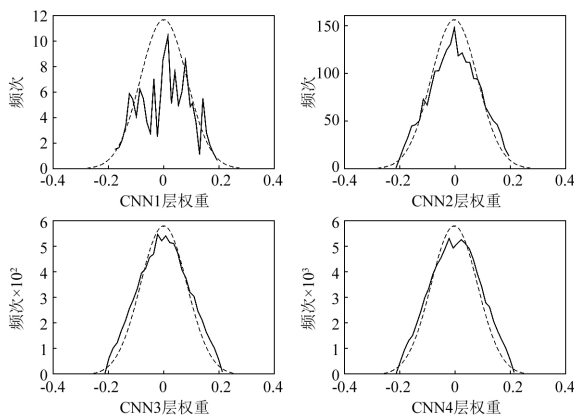


图 5 500 次训练后各层权重频次分布图

Fig.5 Layers of the weight of the frequency distribution after 500 training

3.3 模型对比

为了尽可能多的提供神经网络输入的特征,首先采用三种不同的信号分解方法对每个样本信号进行多层分解。其中 EMD 算法分解后取其前 8 个 IMF 分量信号;小波包变换采用“db3”对信号进行三层分解,对 8 个小波包系数重构得到 8 个分量信号;提升小波分解采用“db3”的提升小波方案进行四层分解,对低频和高

频部分分别重建四层近似系数得到 8 个分量信号。然后对每个样本得到的 8 个分量信号分别求取峰峰值、标准差、方差、均值、均方根值、峭度、整流平均值、波性因子、峰值因子、峭度因子、脉冲因子、裕度因子等 12 个时域特征组成长度为 96 的特征向量,作为神经网络的输入。

每个模型进行 500 次训练,每次训练更新权重后从测试集中随机选取 100 组进行测试,得到准确率变化曲线如图 6 所示(曲线经过 smooth 函数拟合)。在训练完成后,用全部 4 800 组测试集数据对模型进行测试。全部测试完成四种模型用时均在 3 s 左右,平均每组测试时间约为 0.6 s。传统模型由于输入数据少,需要训练的参数远小于 1-DCNN,因此在模式识别阶段,模型训练时间相对较少,但是数据准备时间及复杂度远大于 1-DCNN。多次训练测试后取平均,得到模型的准确率及召回率,如表 3 所示。准确率表示的是预测为正的样本中有多少是真正的样本,而召回率表示样本中的正例有多少被预测正确了。由图 6 及表 3 可以看出,一维卷积神经网络模型的准确率及召回率均高于另外三种模型,而且到达最高准确率所需训练次数更少,到达最高点后准确率波动平稳,模型比较稳健。

表 3 模型测试结果统计及共有参数表

Tab.3 Model test results statistics and common parameter				
	准确率%	召回率%	共有参数	
1-DCNN	99.34	99.33	学习率:0.000 4 训练次数:500	
EMD-NN	96.32	96.29	训练集数量:900×8 测试集数量:600×8	
WPT-NN	94.47	94.06	批次大小:100 损失计算:交叉熵	
LWT-NN	96.09	96.04	优化器:AdamOptimizer 中间层神经元数量:1 024	

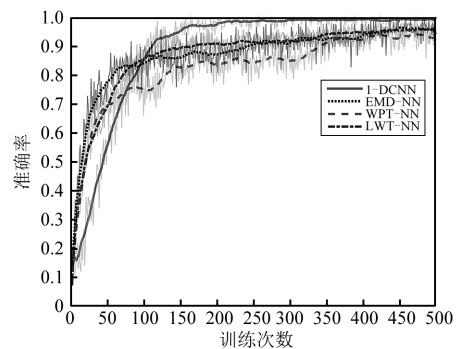


图 6 准确率变化图

Fig.6 Changes in precision

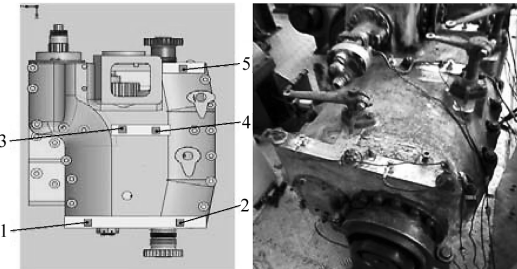
3.4 变速箱故障诊断

图 7(a) 为某型坦克变速箱复合故障模拟实验装置,由转速控制台、三相异步电动机、测试平台和电磁测功机组成,变速箱齿轮为渐开线直齿圆柱齿轮。图 7(b)和图 7(c)为箱体测点位置,选用测点 3 的三档振

动数据为测试数据,测定工况为转速 1 000 r/min,负载 150 N·m,采样频率 20 kHz。



(a) 实验台整体结构



(b) 箱体测点示意位置 (c) 箱体测点实际位置

图 7 实验台整体结构,箱体测点示意位置
和箱体测点实际位置

Fig. 7 Overview of the experimental platform, Schematic diagram of test points and actual location of test points

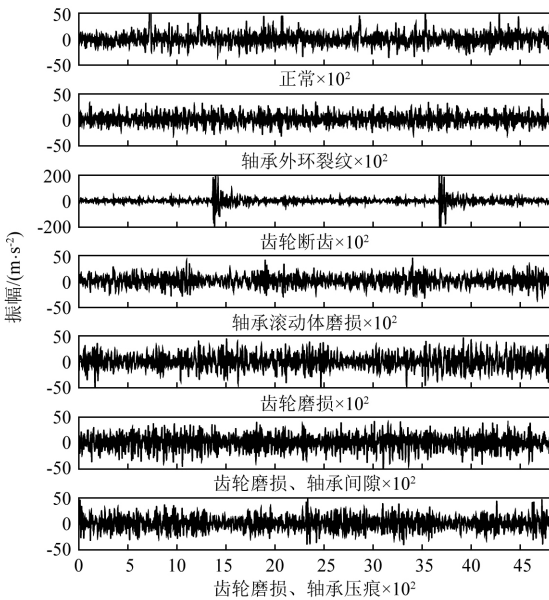


图 8 故障设置及时域信号图

Fig. 8 Fault settings and time domain signal diagram

样本数据长度取略多于两倍整周期的 2 400 点,故障设置及时域信号如图 8 所示。可以看出,除了断齿故障可以明显区分外,其余故障很难在时域图中甄别,尤其是后三个齿轮磨损及齿轮磨损和轴承的复合故障,更是难以区分。因此,在诊断复合故障时,1-DCNN 将 CNN 层组增加为 3 个,而诊断单一故障时仍为 2 个。

实验进行两组测试,测试一:考虑前 5 个单一故障的诊断;测试二:考虑 5 个单一故障及 2 个复合故障的诊断。最终诊断结果如表 4 所示,1-DCNN 在诊断单一

和复合故障时都取得了很高的准确率,而另外三种模型在诊断复合故障时准确率有所下降。准确率变化如图 9 和图 10 所示,两次测试 WPT-NN 和 LWT-NN 的准确率波动变化较大,而 1-DCNN 训练过程表现都很稳定,在 250 次左右达到最高值。说明通过调整 CNN 层组数量,1-DCNN 模型对复杂信号分类诊断具有很好的泛化性。

表 4 两次测试结果统计表

Tab. 4 Results of two test

	测试一: 单一故障检测		测试二: 单一及复合故障检测	
	准确率%	召回率%	准确率%	召回率%
1-DCNN	99.38	99.45	99.32	99.32
EMD-NN	95.97	95.95	93.82	93.68
WPT-NN	98.47	98.45	94.56	94.53
LWT-NN	98.38	98.35	93.69	93.64

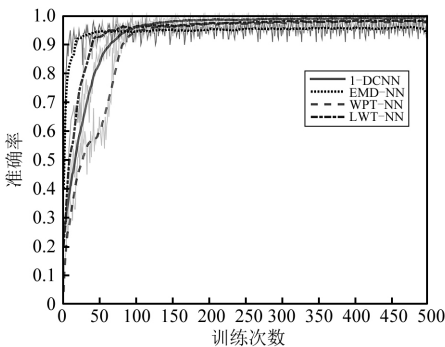


图 9 测试一准确率变化图

Fig. 9 The precision of the first test change

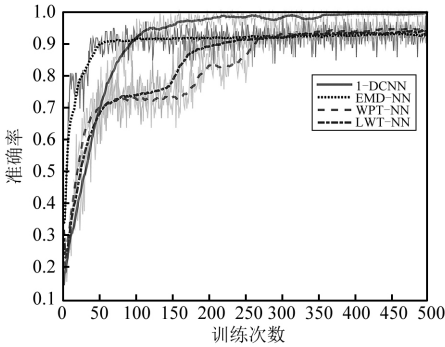


图 10 测试二准确率变化图

Fig. 10 The precision of the second test changes

4 结 论

本文提出了一种利用一维卷积神经网络的齿轮箱复合诊断模型,可以直接从原始振动信号中自动学习特征完成分类诊断。采用 PHM 2009 Challenge Data 斜齿轮复合故障和某型坦克变速箱单一及复合故障对 1-DCNN 以及传统的 EMD-NN、WPT-NN、LWT-NN 诊断模型进行了三次测试,实验结果表明:

(1) 训练好的 1-DCNN 模型可以使用采集到的振动信号直接作为输入,诊断的时间短,真正实现了端到

端的快速故障诊断。

(2) 三次测试 1-DCNN 均取得了 99.3% 左右的准确率,高于另外三种模型。由于测试数据来源既有小型齿轮箱,又有结构更为复杂,噪声干扰更为严重的实装变速箱,说明 1-DCNN 有很高的泛化能力。对于诊断难度更大的变速箱复合故障,通过增加 CNN 层组数量,依然能得到较高的准确率,说明模型的可塑性较强。

参考文献

- [1] LI C, SANCHEZ R V, ZURITA G, et al. Gearbox fault diagnosis based on deep random forest fusion of acoustic and vibratory signals [J]. Mech. Syst. Signal Process. 2016, 76: 283–293.
- [2] POLAK S, Gearbox and gear system problems [C] // in: Gears Conf. London, GEAR TECH., Imech seminar, 1999: 65–72.
- [3] 罗毅,甄立敬. 基于小波包与倒频谱分析的风电机组齿轮箱齿轮裂纹诊断方法[J]. 振动与冲击, 2015, 34(3): 210–214.
LUO Yi, ZHEN Lijing. Diagnosis method of turbine gearbox gearcrack based on wavelet packet and cepstrum analysis [J]. Journal of Vibration and Shock, 2015, 34(3): 210–214.
- [4] 赵德尊,李建勇,程卫东,等. 变转速及齿轮噪声干扰下基于 IDMM 与 EMD 的滚动轴承故障诊断方法[J]. 振动与冲击, 2016, 35(10): 101–107.
ZHAO Dezun, LI Jianyong, CHENG Weidong, et al. Method for rolling element bearing fault diagnosis based on IDMM and EMD under time-varying rotational speed and gear noise [J]. Journal of Vibration and Shock, 2016, 35(10): 101–107.
- [5] 季忠,黄捷,秦树人,等. 提升小波在齿轮箱故障诊断中的应用[J]. 振动、测试与诊断, 2010, 30(3): 291–294.
JI Zhong, HUANG Jie, QIN Shuren, et al. Gearbox fault diagnosis using lifting wavelet [J]. Journal of Vibration Measurement & Diagnosis, 2010, 30(3): 291–294.
- [6] 褚青青,肖涵,吕勇,等. 基于多重分形理论与神经网络的齿轮故障诊断[J]. 振动与冲击, 2015, 34(21): 15–18.
CHU Qingqing, XIAO Han, LÜ Yong, et al. Gear fault diagnosis based on multifractal theory and neural network [J]. Journal of Vibration and Shock, 2015, 34(21): 15–18.
- [7] 张超,陈建军,郭迅,等. 基于 EEMD 能量熵和支持向量机的齿轮故障诊断方法[J]. 中南大学学报(自然科学版), 2012, 43(3): 932–939.
ZHANG Chao, CHEN Jianjun, GUO Xun, et al. Gear fault diagnosis method based on ensemble empirical mode decomposition energy entropy and support vector machine [J]. Journal of Central South University, 2012, 43(3): 932–939.
- [8] HINTON G E, SALAKHUTDINOV R R. Reducing the dimensionality of data with neural networks [J]. Science, 2006, 313(5786): 504–507.
- [9] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation [C] // in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2014.
- [10] TOMPSON J, GOROSHIN R, JAIN A, et al. Efficient object localization using convolutional networks [J]. IEEE Computer Society, 2015: 648–656.
- [11] CHEN Z Q, LI C, SANCHEZ R V. Gearbox fault identification and classification with convolutional neural networks [J]. Shock & Vibration, 2015, 2015(2): 1–10.
- [12] 韩涛,袁建虎,唐建,等. 基于 MWT 和 CNN 的滚动轴承智能复合故障诊断方法[J]. 机械传动, 2016, 40(12): 139–143.
HAN Tao, YUAN Jianhu, TANG Jian, et al. An approach of intelligent compound fault diagnosis of rolling bearing based on MWT and CNN [J]. Journal of Mechanical Transmission, 2016, 40(12): 139–143.
- [13] 袁建虎,韩涛,唐建,等. 基于小波时频图和 CNN 的滚动轴承智能故障诊断方法[J]. 机械设计与研究, 2017, 33(2): 93–97.
YUAN Jianhu, HAN Tao, TANG Jian, et al. An approach to intelligent fault diagnosis of rolling bearing [J]. Machine Design & Research, 2017, 33(2): 93–97.
- [14] SUN W F, BIN Y, ZENG N Y, et al. An intelligent gear fault diagnosis methodology using a complex wavelet enhanced convolutional neural network [J]. Materials, 2017, 10(7): 790.
- [15] JING L, 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.
- [16] PHM data challenge 2009. Available from: <https://www.phmsociety.org/competition/PHM/09>.