

基于应力波卷积神经网络的齿轮故障诊断方法

花海波, 于洪亮, 闫 锦, 廖建彬

(集美大学轮机工程学院, 福建 厦门 361021)

[摘要] 针对齿轮箱的故障诊断问题, 提出一种应力波特征提取算法, 并结合卷积神经网络进行齿轮箱故障诊断。对加速度传感器采集的振动信号进行高通滤波、等间隔重采样、峰值提取, 以获取应力波特征信号; 使用小波变换处理应力波信号, 获得二维时频谱; 将该谱图输入到卷积神经网络中, 训练获得符合要求的故障诊断模型。基于凯斯西储大学的轴承数据集对该故障诊断算法进行初步有效性验证, 以齿轮箱多故障状态下的振动加速度实验数据为基础进行故障特征分析。结果表明, 在使用应力波特征提取算法后, 卷积神经网络的故障诊断率以及诊断效率均有明显提升, 证明了该算法的有效性。

[关键词] 齿轮箱; 应力波; 故障诊断; 卷积神经网络

[中图分类号] TM 311

Fault Diagnosis of Gear Based on Stress Wave Feature and Convolutional Neural Network

HUA Haiibo, YU Hongliang, YAN Jin, LIAO Jianbin

(School of Marine Engineering, Jimei University, Xiamen 361021, China)

Abstract: Aiming at the problem of gearbox fault diagnosis, a stress wave feature extraction algorithm is proposed and combined with convolutional neural network for gearbox fault diagnosis. Firstly, the vibration signal collected by the accelerometer was subjected to high-pass filtering, equal interval resampling, peak extraction, etc. to obtain the stress wave feature signal, and then the wavelet transform was used to process the stress wave signal to obtain a two-dimensional time frequency spectrum. Finally, the spectrum was used as an input parameter entering the convolutional neural network and trained to obtain a fault diagnostic model that meets the requirements. The preliminary effectiveness verification of the fault diagnosis algorithm was conducted based on the bearing dataset of Case Western Reserve University, fault feature analysis was conducted based on experimental data of vibration acceleration signal under multiple fault states of the gearbox. It indicates that after using the stress wave feature extraction algorithm, the fault diagnosis rate and diagnosis efficiency of the convolutional neural network can be significantly improved, which proves the effectiveness of the algorithm.

Keywords: gearbox; stress wave; fault diagnosis; convolutional neural network

0 引言

齿轮箱作为机械设备中的一种必不可少的连接和传递动力的通用零部件, 在农业机械、运输机

[收稿日期] 2022-11-07

[基金项目] 福建省自然科学基金项目(2022J01808); 福建省科技计划项目(2021H0020, 2020H0018)

[作者简介] 花海波(1996—), 男, 硕士生, 从事机械振动与噪声故障诊断研究。通信作者: 于洪亮(1963—), 男, 教授, 博士生导师, 从事模块造船及可能性、振动检测及故障诊断、现代轮机管理工程研究。

械、冶金机械等现代化工业设备中得到了广泛的应用^[1]。在船舶制造领域,船用齿轮箱是船舶的核心部件,目前研究的重点主要集中在船用齿轮箱的非线性振动特性、动力学分析等^[2],关于船用齿轮箱故障机理方面的研究较少。齿轮的不同类型及不同程度的齿轮故障远超轴承故障^[3],研究齿轮故障诊断技术,分析齿轮故障机理,可以更加及时准确地判断齿轮的运转状态,并以此为依据对设备进行保养和维修,这对保障机械设备平稳运行,避免因齿轮故障而造成重大事故等具有重要的意义。

传统的齿轮故障诊断方法包括倒谱法和 Hilbert 解调法^[4],这两种方法分别针对齿轮运行过程中由故障冲击引起的频谱边带及共振调制现象。从本质上来说,由于齿轮的故障而导致运行过程中产生额外力载荷,加速度传感器在采集振动信号时会影响旋转机械的正常振动信号,这种影响在使用传统诊断方式时,对正常信号的畸变只占其比例的 5%~10%^[5]。当齿轮处于初始故障时,其影响甚至更低,因此,对于齿轮的故障诊断方式也一直在不断改进。近年来,深度学习成为机械智能故障诊断的一种主流趋势^[6]。为了克服 BP 神经网络学习效率低、收敛速度慢的缺陷,陈立爱等^[7]将 SOM 神经网络与 BP 结合,构成复合神经网络,其实时性与自主学习性均得到较大提升。薛萍等^[8]针对 BRF 神经网络在非平稳工况下齿轮故障诊断检测困难的问题,引入了 CEEMDAN 算法,提高了故障诊断的有效性和准确性。上述几种方式均是针对某一类信号的故障诊断算法,在诊断算法效率一致的状况下,其诊断准确性非常依赖于原始信号。目前有两种方式可以弥补单一信号在故障诊断上容错能力不佳,故障信息有限的缺点^[9]。1) 提高信号的信噪比,使信号的有效分量提升,凸显故障特征,包括诊断分析^[10]、声发射分析^[11]、热成像技术^[12]等。胡葛庆等^[13]通过经验模态分解对采集到的信号进行提取,结合深度卷积神经网络进行特征融合,实现对齿轮箱的故障诊断。Zhang 等^[14]提出改进的复合多尺度模糊熵,并将其应用于故障特征提取。2) 使用信息融合的方式可以较为全面、准确地反映设备运行状态,从而提高故障诊断精度^[15]。王伟源^[16]为了对齿轮箱故障做出准确的判断,测试数据不仅包含振动信号,还包括油液中磨粒含量数据,这使得诊断算法具备更多维的判断依据。以上算法均是在基于信号本身,而不考虑故障的应力传播特征情况下进行的,或者是采用特殊传感器或多传感器进行协同诊断,在实际故障诊断环境中实施难度大。本文采用最常见的加速度信号传感器,使用应力波信号提取算法对故障特征进行提取,并结合小波变换及卷积神经网络以达到对齿轮故障进行诊断的目的。

1 基本理论

1.1 故障齿轮运转过程中应力波现象

当动力载荷(冲击载荷)作用于弹性物体时,在物体内部会产生应力波^[17]。齿轮早期的断齿、磨损、点蚀等故障或齿轮缺陷都会产生应力波。在旋转机械中,应力波的主要成分为弯曲波,在金属表面,这种波从激励点向外的传递速度 c 与波的频率 f 满足以下关系: $c \propto \sqrt{f}$ 。

因此在激励点之外应力波会随着距离增加而产生衰减,而持续时间会增加。当故障发生时,机器表面产生应力波,由此对运动传感器波产生一个输入信号,传感器收到应力波信号后随之产生一个对应的响应。应力波现象是一种短时瞬态事件,其分布频带宽,频率范围 1~15 kHz,甚至可以扩展到 50 kHz。一般来说,在采集的信号中,正常的振动信号其倍频分量分布位于比应力波分量更低的频带中。为了将应力波信号分离,需要将信号通过高通高阶滤波器进行滤波。经过滤波后,应力波信号比重增大,为了确定故障的特征与类型,需要量化该应力波信号的一些特征,主要包括应力波的幅值,故障激励特征信号的持续时间以及应力波冲击信号的周期性(与部件转速及机器转速有相关性)。

抽取高频应力波分量后还需要对信号进行峰值抽取,而峰值则是滤波信号中一段固定时间间隔中的值。一般情况下,时间间隔的选取是根据分析频率带宽而定的,选取原则需保证在最高的故障频率周围也能够得到足够的采样点数。

峰值数据块的长度取决于轴转速和原始数据时间间隔。图 1 中的原始加速度信号在经过高通滤波(截止频率为 1 ~ 2 kHz) 包络检波后可初步将低频干扰信号从振动信号中剔除, 之后再对信号进行等间隔重采样, 以匹配应力波冲击特征 (一般重采样频率为 $2.56 f_{\max}$, f_{\max} 为分析频率), 为了保持应力波信号的强度, 在重采样区间内对每个区块进行峰值提取, 以提升应力波信号的频谱峰值占比。经过上述处理后即可得图 2 所示的应力波信号。可以看出, 经过以上的预处理, 应力波的幅值和频率得到了较好的保持。

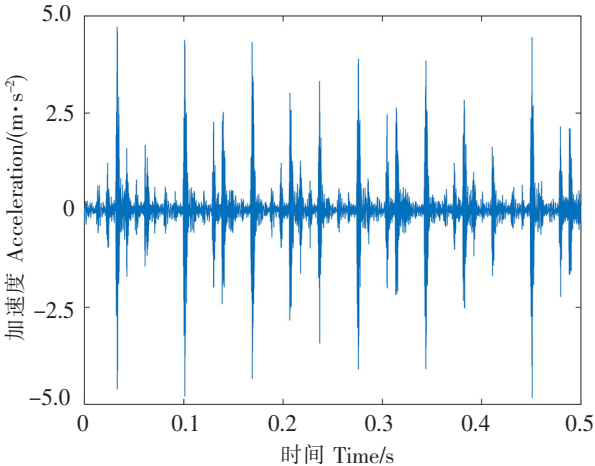


图 1 加速度信号
Fig.1 Acceleration signal

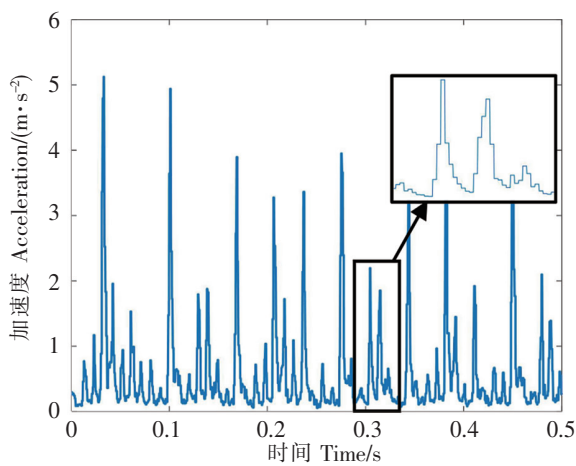


图 2 应力波信号
Fig.2 Stress wave signal

1.2 卷积神经网络

卷积神经网络 (CNN) 作为一种前馈神经网络, 具有强大的特征提取能力, 可以将输入层原始数据逐层转化为更易于识别的特征^[18]。CNN 通过设定卷积核, 对图像特征进行提取, 有效地从大量样本中学习特征表达, 模型泛化能力更强。

在使用 CNN 算法对应力波信号进行处理前, 为了使数据更加符合卷积神经网络的处理需要, 使用连续小波变换 (continuous wavelet transform, CWT) 对应力波信息进行前处理。经过小波变换可以将一维时域信号转换为在时域和频域两个维度上能量强度的时频图谱。另外为了充分利用二维卷积神经网络在图像特征提取方面的优势, 采用 morlet 小波作为小波基函数对应力波信号进行连续小波变换, 改变 morlet 小波基函数的尺度与平移, 应力波在时域上与信号的相关性以积分运算的形式体现, 通过积分运算可以获得数据片段的频率特征。

应力波信号经过 CWT 处理后, 一维应力波信号转换为二维时频信号, 在反映应力波信号频率的同时, 还可以反映变化趋势, 为卷积神经网络识别信号特征提供更多判断依据。

经过上述步骤处理所得的二维时频应力波信号即构成了 CNN 算法的输入层, 除输入层外, CNN 主要组成部分还包括卷积层、池化层和全连接层, CNN 算法的整体结构如图 3 所示。

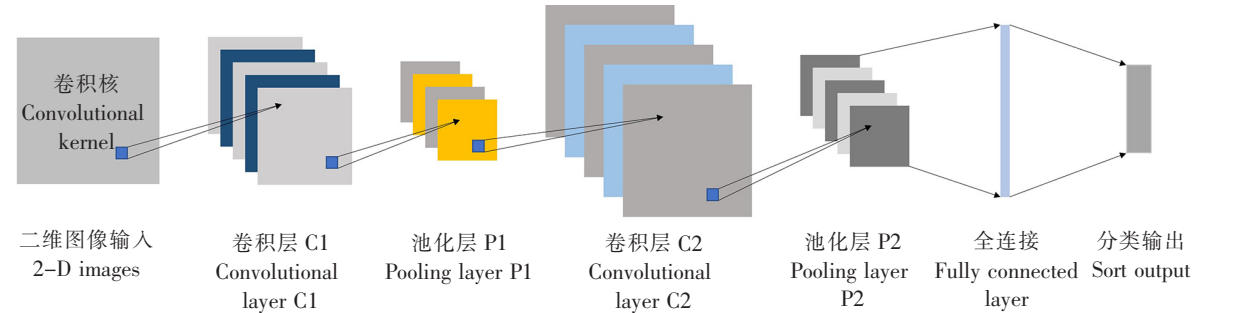


图 3 卷积神经网络结构示意图
Fig.3 Structure of CNN

一个卷积层一般包含一个或多个不同的卷积核，卷积核以指定步长进行滑动卷积，提取输入数据的图形特征。卷积层的数学模型为

$$X_j^{(l)} = \varnothing(\sum_{k=1}^K W_{kj}^{(l)} \otimes X_j^{(l-1)} + B_k^{(l)}). \tag{1}$$

式中： l 表示第 l 层网络； K 表示对输入图片特征的索引所得集合， $k = 1, 2, \cdots, K$ ； J 表示对输出图片特征的索引所得集合， $j = 1, 2, \cdots, J$ ； $X_j^{(l)}$ 为第 l 层第 j 个输出特征； $W_j^{(l)}$ 表示第 j 个卷积核； $B_k^{(l)}$ 为偏置矩阵； \varnothing 为激活函数，作用是为网络空间引入非线性变换的能力，以使所得特征处于映射后的空间内有更强的线性可分性。

池化层用于压缩数据和参数的量，增强数据特征的不变性。池化层的数学模型为

$$P_k^{(i,j)} = \max_{(j-1)\omega+1 < t < j\omega} \{q_k^{(i,t)}\}, j = 1, 2, \cdots, m. \tag{2}$$

式中： $P_k^{(i,j)}$ 表示第 k 层第 i 个输入映射的第 j 个池化值； $q_k^{(i,t)}$ 表示第 k 层第 j 次池化操作的局部输入特征； ω 表示池化区宽度。

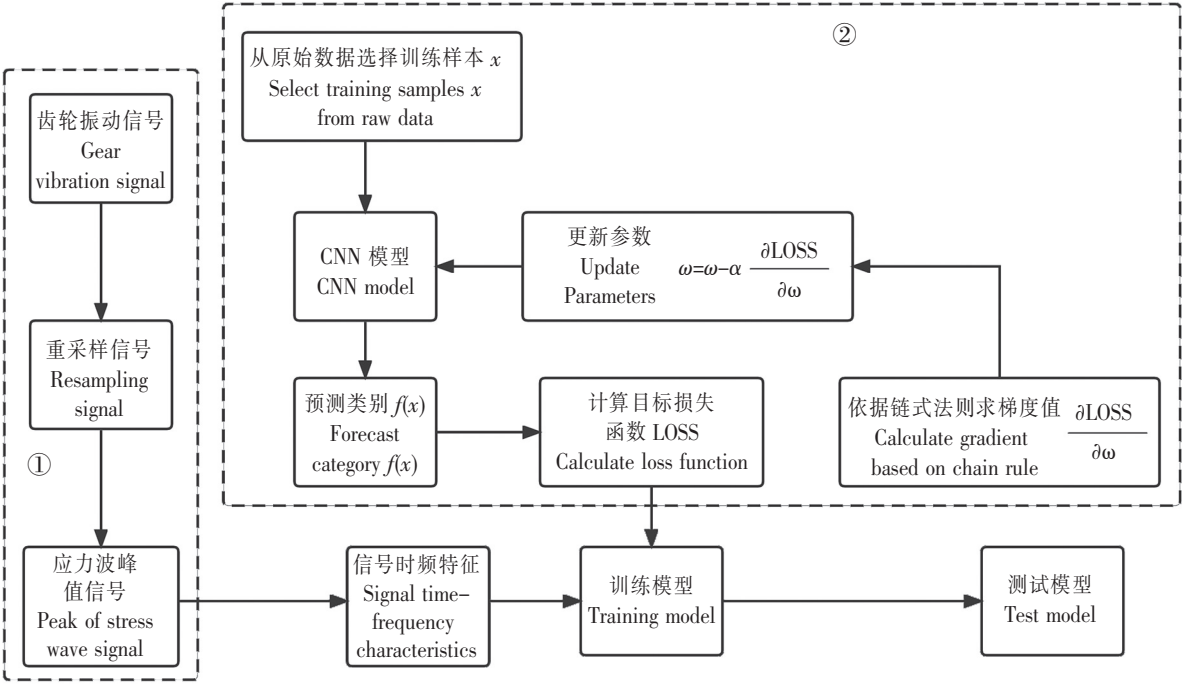
卷积神经网络的目标损失函数为

$$\text{Loss} = - \sum_{i=1}^n y_i \lg f(x_i). \tag{3}$$

式中： x_i 为样本数据； y_i 为样本数据的标签值； n 为样本集容量， $f(x_i)$ 为经过卷积神经网络处理的预测结果。

1.3 应力波特征提取及故障诊断算法流程

相比于原振动加速度信号，应力波信号与故障特征的相关性更强，在使用卷积神经网络对故障进行诊断前，采用应力波特征提取算法对振动加速度信号进行预处理，可以提升神经网络算法的收敛速率及诊断准确率。基于应力波特征提取的卷积神经网络故障诊断算法的具体步骤如图 4 所示。



① — 应力波特征提取 Strees wave feature extraction algorithm; ② — 卷积神经网络 CNN

图 4 故障诊断算法流程
Fig.4 Process of fault diagnosis

为了使卷积神经网络模型更加完善，预处理信号先按比例划分训练集、验证集、测试集。在对划

分好的数据集进行的应力波特征信号提取之后, 为了获取卷积神经网络所需的输入集, 将应力波特征信号进行小波变换, 其变换结果作为输入量参与模型的建立, 并以验证集对该模型进行完善, 最终以测试集作为检测数据获得模型精度及诊断误差等结果。以美国凯斯西储大学轴承数据集为基础, 使用采样率为 12 kHz 下的原始驱动端轴承振动数据, 在模型学习率设置为 0.5% 以及迭代 20 轮的条件下, 验证无负载时内圈故障、外圈故障、滚珠故障、正常轴承分别在 1797, 1772, 1750 r/min 工况下的诊断结果, 其输出特征降维图如图 5 所示。

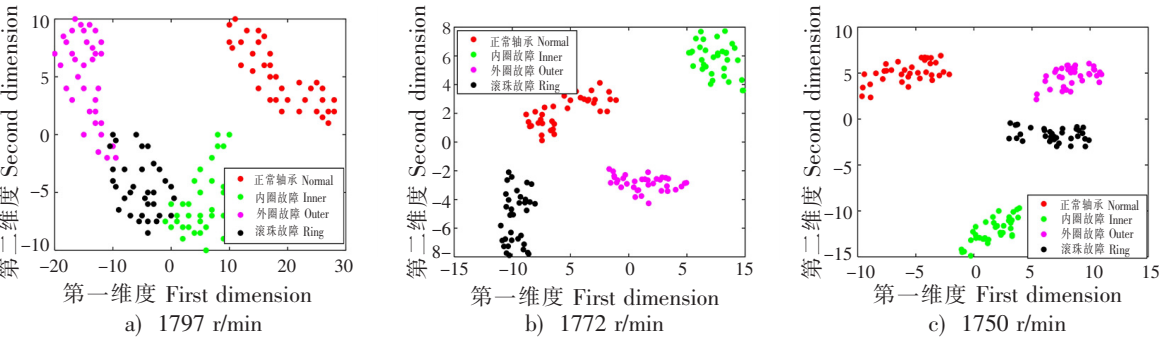


图 5 输出特征降维图
Fig.5 Reduced-dimension map

图 5 表明四种轴承在三种工况下的故障特征聚集性均较好, 对诊断结果进行统计可发现, 若使用原始信号二维时频谱作为输入集, 在 1797, 1772, 1750 r/min 转速下对故障轴承进行诊断的故障误判率分别为 1.4%、1.1%、1.3%, 而在对信号进行应力波特征提取后, 其诊断误判率分别降低至 0.9%、0.5%、0.8%, 初步验证了该算法的有效性。

2 试验方法与算法验证

针对基于应力波卷积神经网络的齿轮故障诊断算法, 设计齿轮故障检测及诊断实验, 实验采用旋转机械故障模拟试验台, 如图 6 所示。

伺服电机与齿轮箱以及齿轮箱与负载轴均采用联轴器连接, 齿轮箱包括两个齿轮, 小齿轮为主动轮, 齿数为 27, 大齿轮为副齿轮, 齿数为 73。

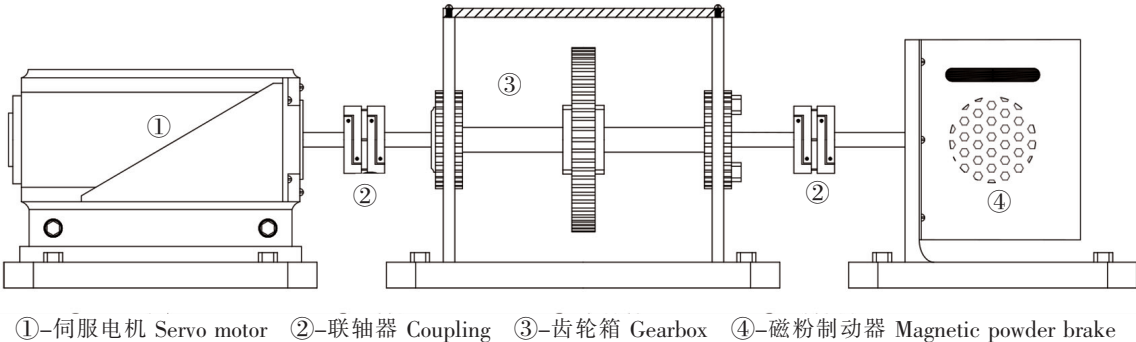


图 6 旋转机械故障模拟试验台
Fig.6 Rotating machinery fault simulation testing

实验中设置采样频率为 51.2 kHz, 采样时间为 21 s, 转速为 1500 r/min。两种模拟故障分别设置为大齿轮断齿和大齿轮点蚀。

实验采集得点蚀、断齿和正常齿轮的振动加速度信号, 使用 Morlet 小波为小波基函数对所采集的齿轮振动信号进行分析。将原始振动信号进行应力波信号提取后的波形以及对应的小波变换时频图如图 7 ~ 图 12 所示。

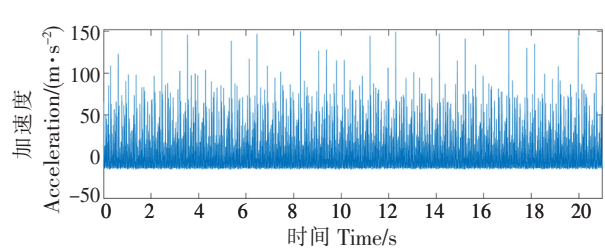


图 7 点蚀齿轮应力波信号
Fig.7 Stress wave signal of pitting fault

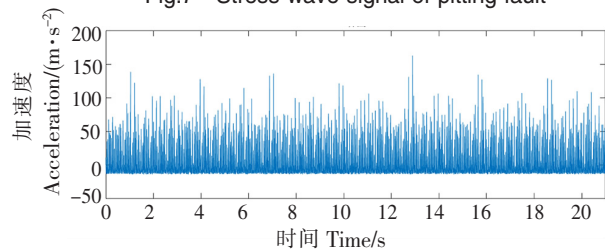


图 9 断齿齿轮应力波信号
Fig.9 Stress wave signal of tooth-breaking fault

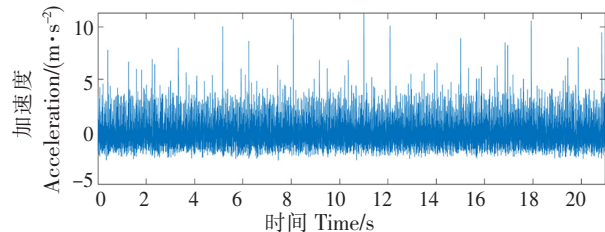


图 11 正常齿轮应力波信号
Fig.11 Stress wave signal of no fault

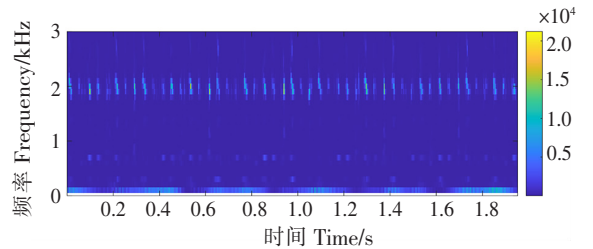


图 8 点蚀齿轮小波时频能量图
Fig.8 Time-frequency spectrum of pitting fault

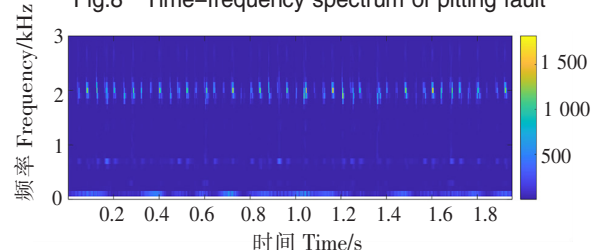


图 10 断齿齿轮小波时频能量图
Fig.10 Time-frequency spectrum of tooth-breaking fault

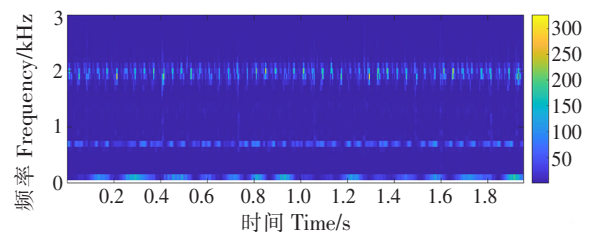


图 12 正常齿轮小波时频能量图
Fig.12 Time-frequency spectrum of no fault

合适的卷积神经网络对模型收敛速度及诊断正确率有明显的影响。第一层卷积主要对二维时频图像进行局部特征的提取，池化层对所提取的特征进行降重采样，去除冗余特征提取，减少模型多余的特征参数；第二层和第三层卷积层以及池化层进行重复的卷积运算，进一步提取与故障类型相关程度较高的特征参量，降低输入特征的模糊性；然后在全连接层的作用下将二维特征展开成一维向量，并将一维向量映射至样本空间；最后经过式（3）交叉熵损失函数实现齿轮故障类型的划分。

卷积神经网络的结构参数如表 1 所示。

表 1 卷积神经网络结构参数表

Tab.1 The Parameter of CNN

| 网络层名称 Name of network layer | 卷积核个数 Number of convolutional kernel | 卷积核大小 Size of convolutional kernel | 池化层大小 Size of pooling layer | 步长 Step width |
|--------------------------------|--|--|-----------------------------------|------------------|
| 卷积层 C1 Pooling layer C1 | 6 | 5 | — | 5 |
| 池化层 P1 Pooling layer C1 | — | — | 2 × 2 | 2 |
| 卷积层 C2 Convolutional layer C2 | 12 | 5 | — | 5 |
| 池化层 P2 Pooling layer P2 | — | — | 2 × 2 | 2 |

模型采取前 3/4 的数据作为训练集，后 1/4 的数据作为模型准确度验证集，将三种状态的小波时频能量图以二维向量的形式作为 CNN 网络的输入集进行模型训练，其最终可得模型误差矩阵，如图 13 所示。

从图 13 所示的模型误差矩阵可以发现，经过训练产生的故障诊断模型进行检验后，正常齿轮信号与点蚀、断齿的故障特征没有误诊断，但点蚀故障和断齿故障之间会产生一定的混淆。其中点蚀故

障有 2.17% 的概率可能会被误诊断为断齿故障, 断齿故障有 0.95% 的概率被误诊断为点蚀故障。从图 8, 图 10 和图 12 来看, 无论是点蚀还是断齿, 其应力波能量相较于正常齿轮要更为集中, 但在时域上并不十分稳定, 这可能是导致点蚀和断齿故障存在相互误判的原因之一。另外, 受限于模型初始化参数以及模型误差与迭代后的模型精度等因素, 也可能导致一定的诊断误差。从总体来说模型在不同故障样本上的分类具有较好的效果, 说明该方法是有效性的。

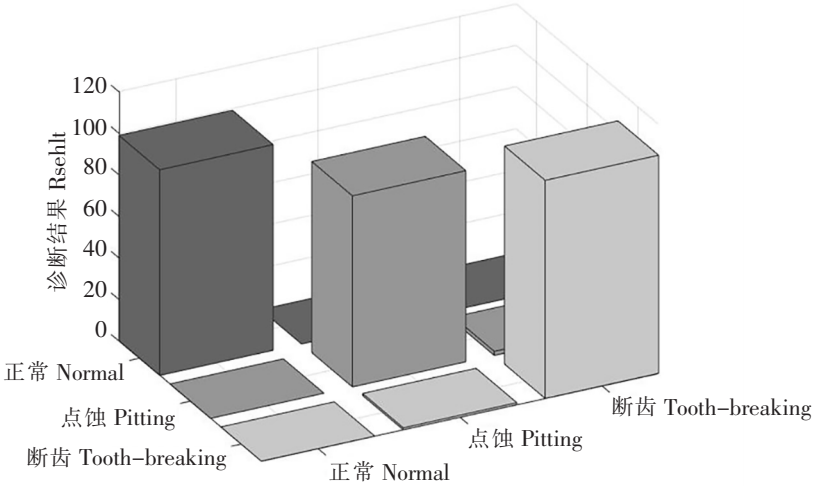


图 13 模型误差矩阵
Fig.13 Error matrix of the model

在应力波信号与原始信号两种输入集下, 模型迭代的准确率变化曲线如图 14 所示, 模型损失值变化趋势如图 15 所示。

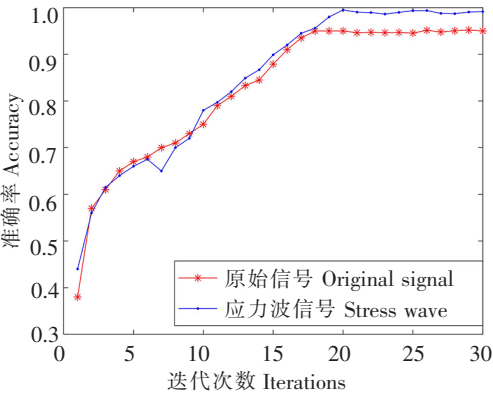


图 14 准确率变化趋势
Fig.14 Variation trend of accuracy from model

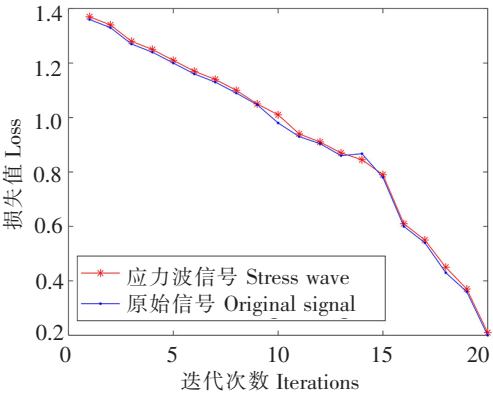


图 15 损失值变化趋势
Fig.15 Variation trend of Loss

由图 14 可见, 在多次迭代中, 随着横轴迭代次数增加, 纵轴原始数据集及应力波信号数据集准确率前期上升速度明显, 平稳收敛至最大值, 但两者仍有差异。依据原始振动加速度信号的诊断模型最终收敛接近于 95%, 而依据应力波信号的模型最终准确率近 99.5%, 模型准确度提升 4.74%, 效果明显。然而损失值变化趋势差异不明显, 以相似速率收敛, 未产生过大偏差, 说明模型在训练中未发生过拟合现象, 所使用的训练集训练所得模型在模型测试集上泛化性较优。

3 结论

本文提出了一种齿轮故障特征提取及故障诊断算法, 给出了基于振动加速度信号的应力波特征提取方法及应用卷积神经网络对齿轮箱故障进行诊断的方法。基于西储大学轴承数据集进行算法验证,

无负载,在转速分别为 1 797, 1 772, 1 750 r/min 时对四种轴承状态(内圈故障、外圈故障、滚珠故障、无故障)的诊断正确率分别达到 35.7%, 54.5%, 38.5%, 证明了该算法在轴承故障诊断中的有效性。

从故障诊断算法对实验原始数据及应力波特征信号的诊断结果来看,在迭代步数相近的情况下,模型诊断精度由 95% 提升至 99.5%, 诊断性能提升明显,说明该算法对于齿轮故障诊断也同样有效。后续可将该算法应用于其他传动结构或旋转部件的故障诊断。

[参 考 文 献]

- [1] 万畅. 齿轮箱轴承故障诊断的时频分析方法初步探讨[D]. 兰州:兰州理工大学,2007.
- [2] 张贵泉. 船用齿轮箱动力学特性及故障诊断研究[D]. 哈尔滨:哈尔滨工业大学,2021.
- [3] 刘士铭. 齿轮故障信号分析及诊断方法研究[D]. 成都:西南交通大学,2021.
- [4] 秦树人. 齿轮传动系统检测与诊断技术[M]. 重庆:重庆大学出版社,1999:173-179.
- [5] ROBINSON J C, CANADA R G, PIETY K P. Peakvue analysis-new methodology for bearing fault detection[J]. Sound & Vibration, 1996, 30(11):22-25.
- [6] 陈保家, 刘浩涛, 徐超, 等. 深度置信网络在齿轮故障诊断中的应用[J]. 中国机械工程, 2019, 30(2):205-211.
- [7] 陈立爱, 陈松, 张舒, 等. 基于 SOM-BP 复合神经网络的齿轮故障诊断研究[J]. 煤矿机械, 2021, 42(7):187-190.
- [8] 薛萍, 郝鹏, 王宏民. 基于径向基神经网络的新型齿轮故障诊断方法[J]. 控制与决策, 2022, 37(2):409-416.
- [9] 赵晓平, 魏旭全, 孙中波, 等. 基于信息融合与 CNN 的齿轮箱故障诊断方法[J]. 测控技术, 2022, 41(11):11-19.
- [10] GELMAN L, MURRAY B, PATEL T H, et al. Vibration diagnostics of rolling bearings by novel nonlinear non-stationary wavelet bicoherence technology[J]. Engineering Structures, 2014, 80(1):514-520.
- [11] JENA D P, PANIGRAHI S N. Automatic gear and bearing fault localization using vibration and acoustic signals[J]. Applied Acoustics, 2015, 98:20-33.
- [12] JANSSENS O, SCHULZ R, SLAVKOVIKJ V, et al. Thermal image based fault diagnosis for rotating machinery[J]. Infrared Physics & Technology, 2015, 73:78-87.
- [13] 胡芑庆, 陈徽鹏, 程哲, 等. 基于经验模态分解和深度卷积神经网络的行星齿轮箱故障诊断方法[J]. 机械工程学报, 2019, 55(7):9-18.
- [14] ZHANG X, ZHAO J M, LI H P, et al. Rolling bearings fault diagnosis under variable conditions using RCMFE and improved support vector machine[J]. International Journal of Acoustics and Vibration, 2020, 25(3):304-317.
- [15] 李益兵, 黄定洪, 马建波, 等. 基于深度置信网络与信息融合的齿轮故障诊断方法[J]. 振动与冲击, 2021, 40(8):62-69.
- [16] 王伟源. 基于信息融合技术的机车齿轮箱故障诊断[D]. 兰州:兰州交通大学,2018.
- [17] 陈长征. 设备振动分析与故障诊断技术[M]. 北京:科学出版社,2007:312-315.
- [18] 李益兵, 马建波, 江丽. 基于 SFLA 改进卷积神经网络的滚动轴承故障诊断[J]. 振动与冲击, 2020, 39(24):187-193.

(责任编辑 陈 敏 英文审校 郑青榕)