

DOI:10.19651/j.cnki.emt.2107869

基于广义 S 变换和迁移学习的轴承故障信号的识别算法*

徐文校¹ 张银胜^{1,2} 杨山山¹ 于心远¹ 徐永杰¹

(1.南京信息工程大学 南京 210044; 2.无锡学院 无锡 214105)

摘要: 滚动轴承是高科技机械设备的重要零部件,也是重要故障源之一。目前,轴承故障样本稀少,数据分布不均匀,传统轴承故障识别方法效果不稳定等给故障识别技术带来了巨大困难。将深度学习相关技术与轴承故障诊断技术相融合,利用深度学习模型识别图像的优势,提出一种广义 S 变换方法。广义 S 变换是小波变换和短时傅里叶变换的继承和发展,通过将其将一维轴承故障信号数据转换成二维时频图,对 Xception 网络进行模型的微调 and 超参数的优化,再将处理后的二维时频图输入改进后的 Xception 网络开展迁移学习。基于凯斯西储大学公开的滚动轴承数据进行了上述实验,针对不同工况的故障信号识别率达到 99.95%,实验结果证明基于广义 S 变换与迁移学习的识别方法真实、有效。

关键词: 轴承故障;广义 S 变换;迁移学习

中图分类号: TP306+.3 **文献标识码:** A **国家标准学科分类代码:** 520.2

Bearing fault signal recognition algorithm based on generalized S transform and transfer learning

Xu Wenxiao¹ Zhang Yinsheng^{1,2} Yang Shanshan¹ Yu Xinyuan¹ Xu Yongjie¹

(1. Nanjing University of Information Science & Technology, Nanjing 210044, China;

2. Wuxi University, Wuxi 214105, China)

Abstract: Rolling bearing is an important part of high-tech mechanical equipment and also one of the important fault sources. At present, few bearing fault samples, uneven data distribution and unstable effect of traditional bearing fault identification methods bring great difficulties to fault identification technology. A generalized S-transform method was proposed by combining deep learning correlation technology with bearing fault diagnosis technology, and taking advantage of deep learning model to recognize two-dimensional images. Generalized S-transform is the inheritance and development of wavelet transform and short-time Fourier transform. By transforming one-dimensional bearing fault signal data into two-dimensional time-frequency diagram, the model of Xception network is fine-tuned and the hyperparameters are optimized, and then the two-dimensional time-frequency diagram is input into the improved Xception network to carry out transfer learning. The above experiments were carried out based on rolling bearing data published by Case Western Reserve University, and the recognition rate of fault signals under different working conditions reached 99.95%. The experimental results prove that the recognition method based on generalized S-transform and transfer learning is real and effective.

Keywords: bearing fault; generalized S transform; transfer learning

0 引言

滚动轴承是现代机器中广泛应用的零部件之一,滚动轴

承能够对转动过程中轴与轴之间零件的正常运作起到很好的支撑作用,并能使轴处于正常位置,提高运作过程中旋转方面的精度,在工业技术发展中占有举足轻重的地位^[1]。但

收稿日期:2021-09-14

* 基金项目:国家自然科学基金(61372128)、教育部协同育人项目(202002179030)、南京信息工程大学滨江学院科研与教研项目(2020yng001, JGZDI201902)资助

是滚动轴承的金属退让性、抗冲击性、抗疲劳性一般,使得轴承在高强度的运转过程中极易发生故障,滚动轴承的故障必然导致机械设备的运行终止,甚至造成连锁反应,引发灾难性后果^[2]。因此,对滚动轴承的研究显得至关重要^[3]。

传统的故障诊断技术思路是将故障信号采用一定的技术手段进行滤波,去除噪声后得到特征较明显的故障信号,以此进行识别,在故障发生初期,故障信号的能量对于噪声信号而言,较为微弱。去除噪声信号的同时,容易将故障信号的部分特征连带去除,造成识别率的下降^[4]。利用向量机^[5]、随机森林^[6]、决策树^[7]等方法识别故障信号^[8],如 Chen 等^[9]对轴承信号进行解耦,将能量熵作为特征值输入向量机中分类, Li 等^[10]通过结合变分模态分解与深度学习方法,实现轴承的智能故障诊断。类似虽然取得了一定成就,但因为其内部结构的复杂性,使得识别结果具有很大偶然性。

近年来,把计算机相关视觉技术融入故障信号的诊断中^[11]成为众多学者的研究方向。尹文哲等^[12]使用 FFT 和传统的 CNN 网络对轴承故障信号进行识别,史光宇等^[13]提出一种二维数据灰度图的概念,将生成的灰度图输入到类似 lenet-5 的网络中。齐爱玲等^[14]提出融合特征的双通道的 CNN 网络识别方式。如 Li 等^[15]将源机器学习到的故障知识快速迁移到目标机器,成功对两个滚动轴承数据集进行实验,以上识别方式都取得了一定的效果,但在实际的操作中,轴承故障样本稀少,数据分布不均匀,常规识别系统的泛化能力不够强,识别精度不高,识别方法并不具备通用性,提出一种基于广义 S 变换与迁移学习的识别方法,广义 S 变换将轴承故障一维信号变成二维时频图,采用迁移学习的方法能够在轴承故障样本稀少的情况提高识别率并避免出现过拟合等现象,最后通过模型微调和算法优化,进一步提升识别率,扩大适用范围。

1 数据预处理

轴承故障数据是一维信号,不能用于卷积神经网络中二维图像信号的识别,需进行一维信号到二维信号的转换。常用的转换方法是短时傅里叶变换^[16]、小波变换^[17]。

短时傅里叶变换克服了傅里叶变换中不能反映局部区域上特征的缺陷,短时傅里叶变换通过加窗截取的思想,对截取后的信号做傅里叶变换,通过滑动窗口,完成对整段信号的频谱分析。短时傅里叶变换中,窗口大小一旦确定就不能改变,不能灵活满足现实具体信号分析中时频分辨率高低要求。小波变换引入有限长会衰减的小波基^[18],通过引入平移因子和尺度因子控制小波窗口的平移与伸缩,解决了短时傅里叶变换中窗口固定的缺陷,提高了分辨率的灵活性,需注意的是小波变换得到的结果并没有与时频谱建立直接联系,而是在时间-尺度平面上进行表示。

S 变换^[19]在短时傅里叶变换的基础上做出改进,引入一个可随频率变化而平移伸缩的高斯窗,克服了短时傅里

叶变换中窗口固定的缺陷,用 S 变换分析非平稳信号时既能保持与频率分辨率相联系又能与傅里叶变换有关。S 变换公式如式(1)所示。

$$S(f, \tau) = \int_{-\infty}^{+\infty} h(t)w(t - \tau)\exp(-i2\pi ft)dt \quad (1)$$

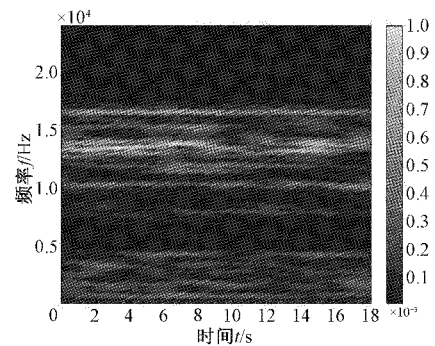
式中: $h(t)$ 为待分析信号, f 为信号频率, τ 为平移量, $w(t) = \frac{1}{\delta(f)\sqrt{2\pi}}\exp(\frac{-t^2}{2\sigma(f)^2})$ 为高斯窗函数,且 $\sigma(f) = 1/|f|$ 。

在 S 变换中,窗函数的形态只能随频率的变化而变化,无法灵活调节窗宽限定了其应用。广义 S 变换^[20]在 S 变换的基础上进行改进,通过引入参数 m 调节高斯窗的宽度,提高了 S 变换的时频分辨率。短时傅里叶变换、小波变换以及 S 变换在精度、分辨率以及灵活性方面都要弱于广义 S 变换。广义 S 变换公式如式(2)所示。

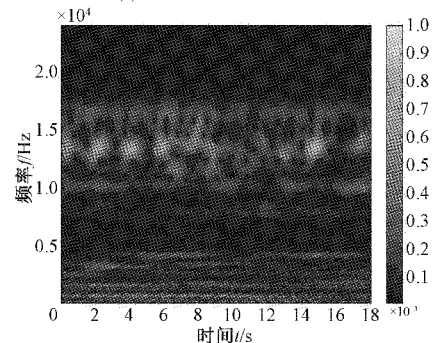
$$S(f, \tau) = \int_{-\infty}^{+\infty} w(t - \tau)h(t)\exp(-i2\pi ft)dt \quad (2)$$

式中: $w(t) = \frac{1}{\sigma_m(f)\sqrt{2\pi}}\exp(\frac{-t^2}{2\sigma_m(f)^2})$, 且 $\sigma_m(f) = 1/|f|^m$ 。

对 3 种变换进行实验比较,结果如图 1 所示。通过比较上述各类变换所得时频图,其中广义 S 变换得到的时频图较其他时频图而言,提取的细节特征更加明显,分辨率更高,选择将广义 S 变换所得时频图作为后续卷积神经网络的输入。本文将分别采用小波变换、短时傅里叶变换和广义 S 变换处理一维轴承故障数据得到二维时频图,将其作为改进后卷积神经网络的输入,进行对比实验,证明所述观点的准确性。



(a) 短时傅里叶变换时频图



(b) 小波变换时频图

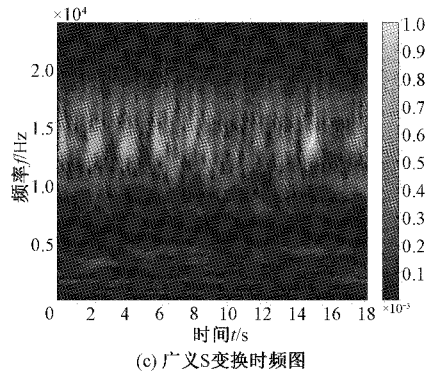


图1 各类变换时频图

2 迁移学习

迁移学习是一种机器学习的方法^[21],大多数的机器学习算法都有一个基本的假设前提,那就是训练样本和测试样本需服从同等的概率分布且同时标记训练的样本数也要足够多。在实际工程中,鉴于大批量样本采集的难度,所获数据的数量并不能很好地保证机器学习算法的成功运用,考虑到大部分数据或工程任务都存在很大的相关性,通过迁移学习可以把源领域学到的知识迁移到所要研究的目标领域,很好地解决了目标领域数据标签难获取,并克服了从头开始建立复杂模型耗时长的障碍,使得目标领域能够取得更好的学习效果。从学习方法的角度去看待迁移学习,有基于样本的迁移学习方法、基于特征的迁移学习方法、基于模型的迁移学习方法以及基于关系的迁移学习方法。其中基于模型的迁移学习方法是当下研究的热门,原理如图2所示。

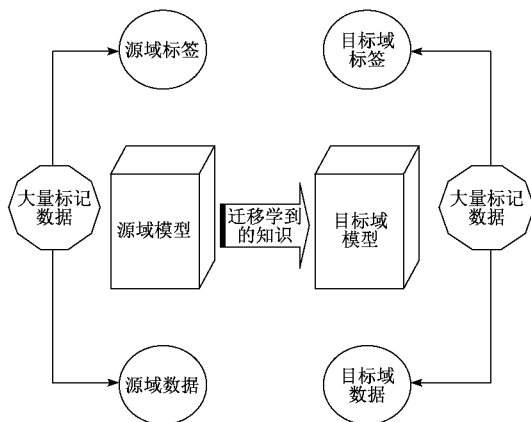


图2 迁移学习原理

考虑到滚动轴承的工况复杂,现实中故障信号采样难度大,本文提出了基于迁移学习的方法,采用在 Imagenet 数据集 (Imagenet 数据集是由斯坦福大学李飞飞教授带领创建的一个计算机视觉数据集) 上预训练好的 Xception 网络进行迁移学习,Imagenet 数据集有 1 400 多万幅图片,涵盖两万多个类别,在深度学习图像领域应用非常广,数据特

征的复杂度远高于轴承故障信号。因此,基于 Imagenet 数据集的预训练模型能够完全满足轴承故障信号的识别。随后通过模型的微调和超参数的优化,进一步提高对轴承故障信号的识别率。

3 模型的微调和优化

3.1 卷积神经网络

卷积神经网络 (convolutional neural networks, CNN) 是一种前馈神经网络,相比于其他全连接网络,卷积神经网络的核心思想是局部感受野和权值共享,降低了模型的复杂度,在图像分类、分割等领域具有出色表现。

卷积神经网络一般由输入层、卷积层、池化层、全连接层以及输出层构成。其结构如图3所示,卷积层通过权值共享提取图像特征,池化层通过合并附近单元,减小下层输入的尺寸,起到降维的作用,将卷积层与池化层提取的特征放入全连接层中,经输出层输出结果。

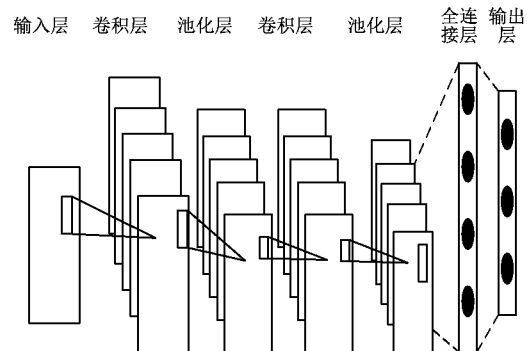


图3 卷积神经网络结构

3.2 Xception 网络

Xception 是谷歌继 Inception 后提出的对 Inception-v3 的另一种改进。作者首次提出将通道间的相关性与空间的相关性分离开单独处理的概念,在参数相同的情况下,增加了神经网络的深度,对神经网络的性能能够进一步的优化。Xception 的目标是成为易迁移、计算量小,能适应不同任务、且精度较高的模型。Xception 网络能用更少的参数和更少的计算开销学到更好的特征表示^[22]。对于本文所研究的轴承故障识别模型,Xception 网络天然的满足了不同工况下不同位置的轴承故障诊断所需求的不同网络模型,避免了重复设计网络,很大程度上提升了识别的效率。此外,Xception 加入的类似残差网络 (Resnet) 的残差机制^[23]也显著加快了模型的收敛速度并获得了更高的准确率。Xception 网络结构如图4所示。

3.3 模型的微调和参数的优化

神经网络的浅层结构往往提取的是图像的角点、边缘、轮廓、纹理等通用信息。这些信息一般不局限于在本网络中使用,也适用于其他网络,而靠近输出层的深层结构一般提取的是图像的细节特征,这种细节特征才是需要仔细把握的地方。

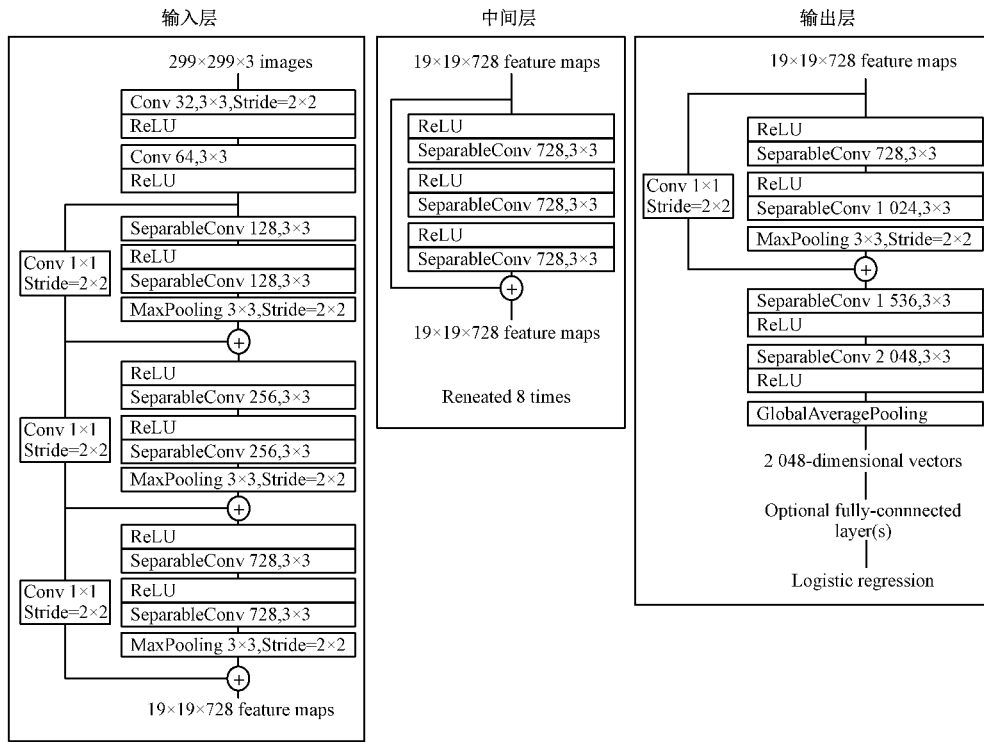


图 4 Xception 网络结构

由于本文所研究的故障信号种类数远不及 Imagenet 数据集的最终分类数(1 000),因此本文选择删除 Xception 网络的全连接层结构,添加符合本文所研究实际情况的全连接层(最终分类数 10),并冻结预训练模型的部分卷积层,通过大量实验,在训练时间和最终的识别率之间权衡,选择激活预训练模型的后 33 层(让最后 33 层重新训练,不使用预训练模型权重)。同时改进全连接层的激活函数^[24],用带泄露修正线性单元(LeakyReLU)函数代替传统的线性修正单元(ReLU)函数,函数图像如图 5、6 所示,ReLU 函数收敛速度、计算速度快,解决了正区间内存在的梯度消失问题,但是在负区间内,ReLU 函数将所有的负值置为 0,造成神经元死亡现象,LeakyReLU 函数在 ReLU 的基础上做出了改进,将所有的负值赋予一个很小的非零斜率,很好解决了 ReLU 函数存在的神经元坏死问题。针对本文所研究的故障信号识别,为证明本文所述观点的严谨性,在后续的实验部分,采用对比论证,证明了 LeakyReLU 函数相对于 ReLU 函数,有更好的识别效果。同时我们选用 SGD 梯度优化算法^[25],配合使用学习率调整函数,防止模型出现过拟合的情况,使模型具有更大的灵活性。改进后网络模型结构如图 7 所示。

4 实验验证

为验证本文所述观点的有效性,本文在美国凯斯西储大学公开的滚动轴承故障数据集上进行验证分析,本次实验系统配置处理器为 Intel(R)Core(TM)i5-10300H CPU@

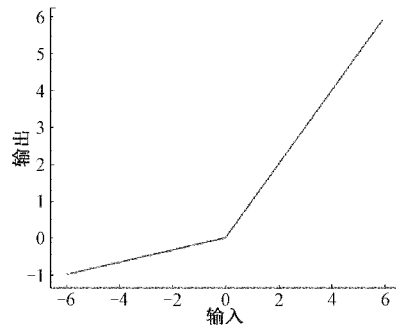


图 5 LeakyReLU 激活函数图像

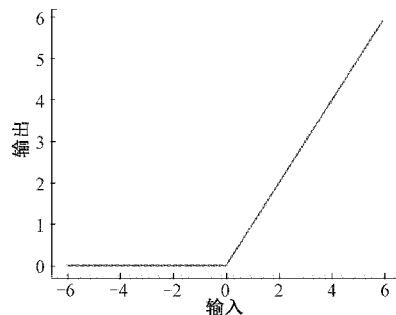


图 6 ReLU 激活函数图像

2.50 GHz, GPU 为 NVIDIA GeForce GTX1650Ti。

为了验证本文所提出模型的识别通用性,本文选取采样频率为 12 kHz 的风扇端轴承故障数据与正常轴承数据和采样频率为 48 kHz 的驱动端轴承故障数据与正常数据

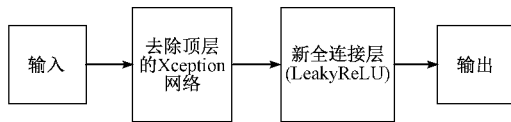


图7 改进网络结构简图

进行分析。设置3种故障直径,分别为0.007 inch、0.014 inch、0.021 inch,代表着不同的故障程度,在负载大小为0 hp(马力)下进行测试和数据收集。

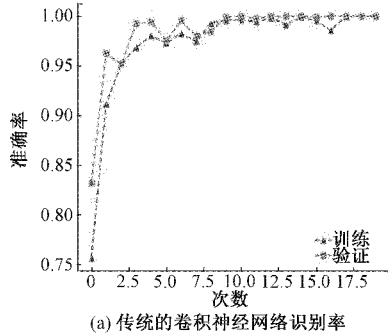
分别对风扇端和驱动端各10组故障数据,进行滑动窗口重叠采样,考虑到现实设备操作过程中,故障数据样本稀少,为此本文将每种故障状态下样本数采样为200,经数据增强技术处理后,样本数变为1000,每种状态下,训练样本集、验证样本集、测试样本集按照60%、20%、20%划分。表1为轴承故障数据集详细分布情况。

表1 轴承故障数据集

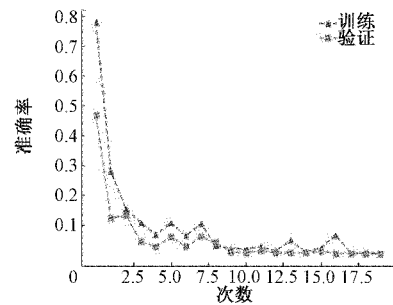
负载	故障位置	故障尺寸/inch	训练样本数	验证样本数	测试样本数	总计
0	外圈	0.007	600	200	200	200
		0.014	600	200	200	200
		0.021	600	200	200	200
	内圈	0.007	600	200	200	200
		0.014	600	200	200	200
		0.021	600	200	200	200
	滚动体	0.007	600	200	200	200
		0.014	600	200	200	200
	无故障	—	600	200	200	200

对处理好的数据集进行广义S变换得到二维时频图,将时频图输送到已经建立好的模型进行训练。本次实验神经网络全连接层的激活函数分别采用ReLU函数,和改进后的LeakyReLU函数。同时与传统的卷积神经网络(2个卷积-池化对和1个全连接层)的训练结果做对比,在12 kHz的采样频率下,风扇端故障训练效果如图8所示。

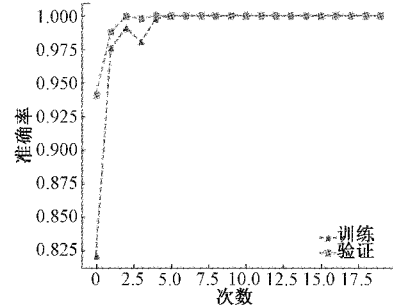
将处理好的数据集分别通过短时傅里叶变换与小波变换得到二维时频图,输入改进后的卷积神经网络中,实验结果如图9所示。



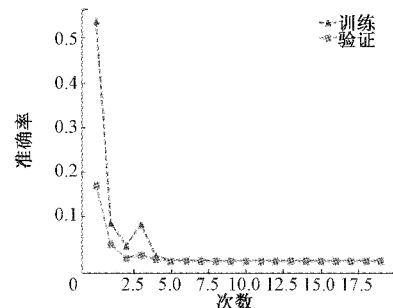
(a) 传统的卷积神经网络识别率



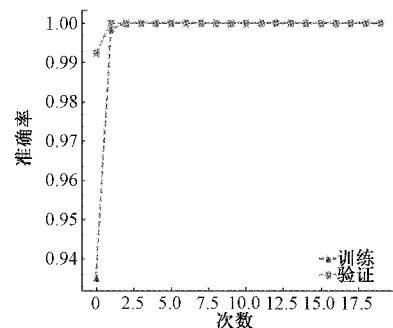
(b) 传统的卷积神经网络损失值



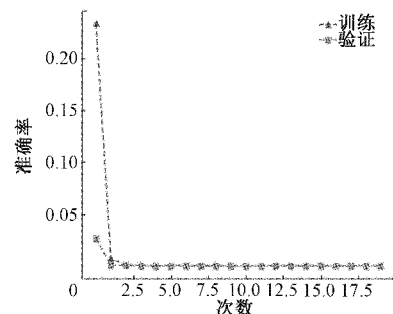
(c) ReLU识别率



(d) ReLU损失值

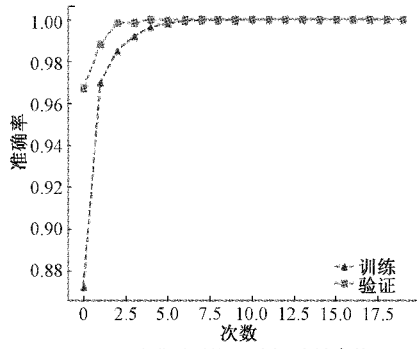


(e) 改进后网络识别率

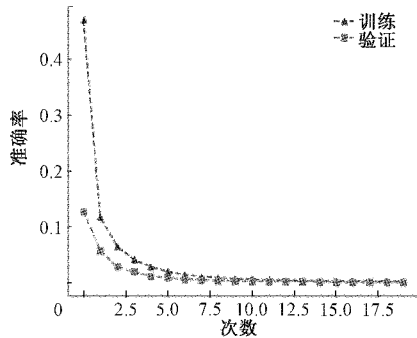


(f) 改进后网络损失值

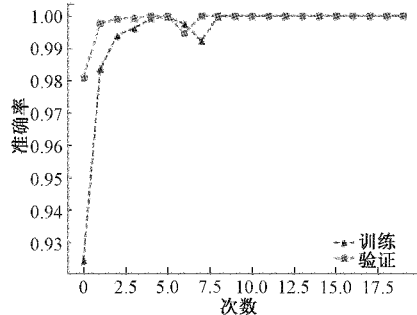
图8 相关网络识别率与损失值图



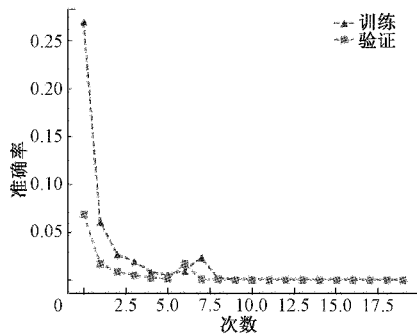
(a) 改进后网络识别率_小波变换



(b) 改进后网络损失值_小波变换



(c) 改进后网络识别率_短时傅里叶变换



(d) 改进后网络损失值_短时傅里叶变换

图 9 两种变换识别率与损失值图

把划分好的测试集传递到几种不同的卷积神经网络中,进一步验证模型性能好坏。测试结果如表 2 所示。

从图 8、9 与表 2 看出,传统的神经网络模型识别准确率与改进后的神经网络相比较低。通过广义 S 变换输入到改进后神经网络相对于其他两种变换后的结果作为输入识别率更高,收敛更快。

表 2 测试结果

诊断方法	传统的卷积神经网络	改进后卷积神经网络	改进后卷积神经网络_小波变换	改进后卷积神经网络_短时傅里叶变换
识别率/%	97.23	98.41	99.85	99.13

为了进一步了解模型对每个故障类别的识别率,排除对某一类故障存在识别率极端的情况,通过混淆矩阵,对测试集结果进行总结,结果如图 10 所示。

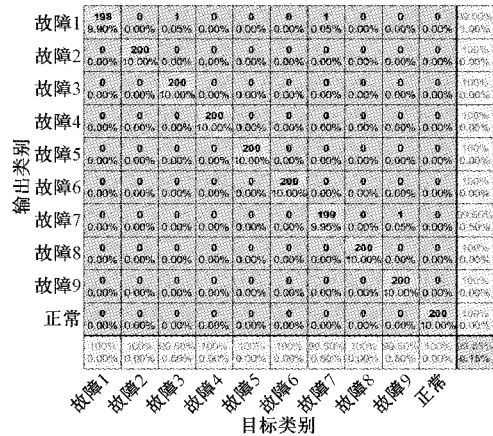


图 10 混淆矩阵结果 a

将改进后的神经网络模型(选用 LeakyReLU 作为激活函数)识别 48 kHz 采样频率下驱动端轴承故障。识别效果如图 11~12 所示,通过混淆矩阵对驱动端轴承故障验证集识别情况进行总结,如图 13 所示。

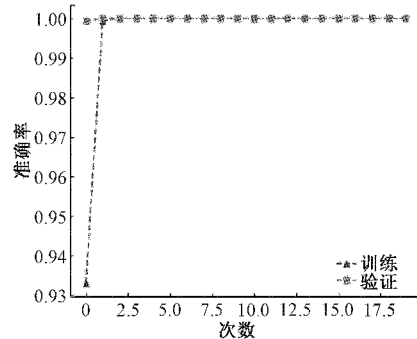


图 11 48 kHz_驱动端_识别率

对于采样频率为 48 kHz 的驱动端轴承故障数据,改进后的神经网络也有很好的训练效果,在测试集上的识别率达到 99.95%,混淆矩阵用来观察风扇端和驱动端每一类故障识别率,从图 10 与 13 看出,改进后的卷积神经网络相对而言对风扇端故障 1 识别效果最差,但识别率仍能达到 99%;驱动端同样如此,对故障 3 识别效果最差,识别率 99.5%,反映出对驱动端和风扇端的每一类故障均具有优秀的识别效果,再次证明此模型对于不同故障位置、不同轴承的诊断通用性和高精度性。

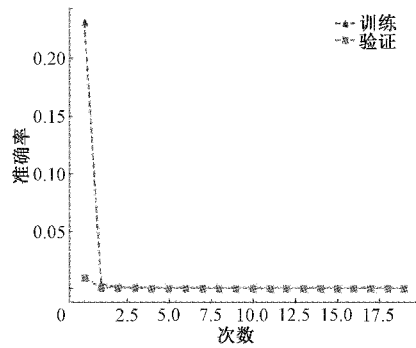


图12 48 kHz_驱动端_损失值

输出类别	故障1	故障2	故障3	故障4	故障5	故障6	故障7	故障8	故障9	正常
故障1	200	0	0	0	0	0	0	0	0	0
故障2	0	200	0	0	0	0	0	0	0	0
故障3	0	0	200	0	0	0	0	0	0	0
故障4	0	0	0	200	0	0	0	0	0	0
故障5	0	0	0	0	200	0	0	0	0	0
故障6	0	0	0	0	0	200	0	0	0	0
故障7	0	0	0	0	0	0	200	0	0	0
故障8	0	0	0	0	0	0	0	200	0	0
故障9	0	0	0	0	0	0	0	0	200	0
正常	0	0	0	0	0	0	0	0	0	200

图13 混淆矩阵结果 b

5 结 论

现实工业生产中机械设备的轴承故障数据样本稀少,数据分布不均匀,轴承设备多样性等特点造成传统的故障识别方法诊断效果不稳定。本文提出一种基于广义S变换与迁移学习的轴承故障识别方法,将一维故障数据经广义S变换得到二维时频图,将时频图作为改进后卷积神经网络的输入。

本文改进后的网络模型相对于传统的卷积神经网络而言对于故障信号的识别率更好,相对于小波变换和短时傅里叶变换,广义S变换得到的时频图特征更加明显,经广义S变换后所得的时频图输入改进后的网络模型相比于其它两种变换,对故障信号的识别精度更高。所述方法对故障信号的识别率有一定提升,但参数量会少许增加,计算时间略长,下一步的工作主要针对这一点进行后续研究。

参考文献

[1] 吴伟力,陈大力. 航空发动机振动故障诊断及发展趋势[J]. 航空动力,2021(3):69-72.
 [2] WANG Z, ZHOU J, WANG J, et al. A novel fault diagnosis method of gearbox based on maximum kurtosis spectral entropy deconvolution [J]. IEEE Access,2019: 29520-29532.

[3] 杨晓蔚. 滚动轴承产品技术发展的现状与方向[J]. 轴承,2020(8):65-70.
 [4] 唐贵基,闫晓丽,王晓龙. 基于Chen系统的滚动轴承早期微弱故障检测方法研究[J]. 振动与冲击,2020,39(20):128-133.
 [5] 王一鹏,陈学振,李连玉. 基于小波包混合特征和支持向量机的机床主轴轴承故障诊断研究[J]. 电子测量与仪器学报,2021,35(2):59-64.
 [6] 彭成,王松松,贺婧,等. 基于离散小波变换和随机森林的轴承故障诊断研究[J]. 计算机应用研究,2021,38(1):101-105.
 [7] 梁晓雯,蒋爱平,王国涛,等. 参数优化决策树算法的密封继电器多余物信号识别技术[J]. 电子测量与仪器学报,2020,34(1):178-185.
 [8] ZHANG W, PENG G, LI C, et al. A new deep learning model for fault diagnosis with good anti-noise and domain adaptation ability on raw vibration signals[J]. Sensors, 2017, 17(3):425.
 [9] CHEN X J, YANG Y M, CUI Z X, et al. Vibration fault diagnosis of wind turbine based on variational mode decomposition and energy entropy[J]. Energy, 2019, 174:1100-1109.
 [10] LI J M, YAO X F, WANG X D, et al. Multiscale local feature learning based on BP network for rolling bearing intelligent fault diagnosis[J]. Measurement, 2020, 153:107419.
 [11] 易静姝. 神经网络在滚动轴承故障诊断中的应用与发展[J]. 价值工程,2019,38(24):274-276.
 [12] 尹文哲,夏虹,彭彬森,等. 基于FFT和CNN的滚动轴承故障诊断方法[J]. 应用科技,2021:1-5.
 [13] 史光宇,徐健,杨强. 基于卷积神经网络的风电机组轴承机械故障智能诊断方法[J]. 华北电力大学学报(自然科学版),2020,47(4):71-79.
 [14] 齐爱玲,李琳,朱亦轩,等. 基于融合特征的双通道CNN滚动轴承故障识别[J]. 机械与电子,2021,39(5):15-19.
 [15] LI X, JIA X D, ZHANG W, et al. Intelligent cross-machine fault diagnosis approach with deep auto-encoder and domain adaptation[J]. Neurocomputing, 2020, 383(28): 235-247.
 [16] 谢丽娟,路锋,王旭,等. 一种用于信号去噪的小波阈值去噪算法[J]. 国外电子测量技术,2020,39(4):32-36.
 [17] 曹松. 小波变换在信号去噪中应用[J]. 电子测量技术, 2002(3):36-37.
 [18] 潘超,陈祥,蔡国伟,等. 基于小波包尺度-能量占比的变压器三相不平衡绕组振动特征辨识[J]. 仪器仪表学报,2020,41(4):129-137.
 [19] 陈明,马宏忠,潘信诚,等. 基于S变换和遗传算法优化SVM的GIS机械故障诊断[J]. 电力信息与通信技

- 术,2020,18(5):1-6.
- [20] 李卫国,许文文,王旭光,等. 基于广义 S 变换的有源配电网故障定位方法[J]. 电测与仪表,2021,58(6):105-112.
- [21] 庄福振,罗平,何清,等. 迁移学习研究进展[J]. 软件学报,2015,26(1):26-39.
- [22] 陈巧红,陈翊,李文书,等. 多尺度 SE-Xception 服装图像分类[J]. 浙江大学学报(工学版),2020,54(9):1727-1735.
- [23] 赵志成,罗泽,王鹏彦,等. 基于深度残差网络图像分类算法研究综述[J]. 计算机系统应用,2020,29(1):14-21.
- [24] 张焕,张庆,于纪言. 激活函数的发展综述及其性质分析[J]. 西华大学学报(自然科学版),2021,40(4):1-10.
- [25] 常永虎,李虎阳. 基于梯度的优化算法研究[J]. 现代计算机,2019(17):3-8.

作者简介

徐文校,硕士在读,主要研究方向为信号处理、深度学习等。

E-mail:971149614@qq.com

张银胜,副教授,硕士生导师,主要研究方向为深度学习、信号处理,无线通信技术等。

E-mail:yorkzhang@nuist.edu.cn