

DOI:10.19651/j.cnki.emt.2105644

时频分析与 VGG19 迁移学习的轴承故障检测*

李传鑫 刘增力

(昆明理工大学 信息工程与自动化学院 昆明 650500)

摘要:为了解决轴承故障诊断依赖专家经验的人工特征提取问题,提出时频分析与 VGG19 网络迁移学习的故障诊断方法。首先利用时频分析的方法将轴承的正常状态、内圈故障、外圈故障和滚动体故障的一维数据转换为时频样本图,同时也将上述数据生成谱峭度图;其次对 VGG19 网络模型中的全连接层进行网络替换并 Fine-tuning;最后通过网络调参实现卷积神经迁移学习网络对轴承故障的识别分类诊断。结果表明,在实验中滚动轴承故障诊断的时频样本图分类准确率高于谱峭度图样本分类的准确率高达 5.42%,验证了时频分析与 VGG19 迁移学习在信号处理方面应用的有效性;另外,迁移学习可以解决小样本的故障诊断问题。

关键词: VGG19;时频分析;故障诊断;迁移学习

中图分类号: TH133.3 **文献标识码:** A **国家标准学科分类代码:** 510.99

Time frequency analysis and bearing fault detection of
VGG19 transfer learning

Li Chuanxin Liu Zengli

(Department of Information Engineering and Automation, Kunming University of Science and Technology, Kunming 650500, China)

Abstract: In order to solve the problem of artificial characteristics of bearing fault diagnosis, this paper puts forward the fault diagnosis method of the time frequency analysis and VGG19 network migration learning. First, the normal state, the internal ring fault, the outer ring fault and the sliding fault of the rolling body are converted to the frequency sample diagram, and then the spectrum kurtosis graph is generated from the above data. Secondly, the full connection layer in the VGG19 network model is replaced and Fine-tuning. Finally, the convolutional neural transfer learning network is used to recognize and classify bearing faults through network parameter tuning. The results show that the classification accuracy of time-frequency sample graph for rolling bearing fault diagnosis in the experiment is 5.42% higher than that of spectral kurtosis graph, which verifies the validity of the application of time-frequency analysis and VGG19 transfer learning in signal processing. In addition, transfer learning can solve the problem of fault diagnosis of small samples.

Keywords: VGG19; time frequency analysis; fault diagnosis; transfer learning

0 引言

轴承是大型机械设备中关键的核心部件,在机床、航空、交通、精密仪器等方面得到了应用,为设备的正常运转起着重要作用。周济院长在“新一代智能制造-新一轮工业革命的核心驱动力”的报告中指出:“制造业数字化网络化智能化是新一轮工业革命的核心技术,应该作为《中国制造2025》的制高点、突破口和主攻方向”^[1]。因此,轴承故障智能诊断成为未来重要的发展趋势。

随着人工智能的发展,传统的机器学习方法早已运用

到了轴承故障诊断中去了,但是传统的机器学习方法需要大量的数据和人工经验做标注,耗时且重复性工作制约着智能化检测的发展。唐波等^[2]提出了短时傅里叶变换的二维时频谱和卷积神经网络的滚动轴承故障诊断。Sheng等^[3]提出了一种基于卷积神经网络的振动故障诊断新方法,该方法利用小波变换作为原始信号的时频域变换,利用小波变换对机械振动信号生成时频图,然后以连续小波变换时频图作为输入训练 CNN 模型进行故障诊断。黄驰城^[4]提出采用深度学习的图像分类技术对振动信号的时频图、频谱图进行分析相对于传统的图像分析方法更加快速

收稿日期:2021-01-01

* 基金项目:国家自然科学基金(61271007)项目资助

高效,具有很好的应用性。唐藁^[5]提出了基于时频图和能量特征的双输入 CNN 故障诊断模型,利用同步挤压小波变换处理所得的 RGB 时频图像作为分类模型的输入,将其转换为图像识别问题。杜小磊等^[6]提出基于同步挤压 S 变换和深度学习的轴承故障诊断。因此,时频图可以作为时频样本来分类,但是重新训练新网络要花费计算力和较大的数据样本。

刘鑫鹏等^[7]提出迁移学习主要解决的是以下两个问题:1)解决小数据问题。当训练数据过少时,经典监督学习会出现严重过拟合问题,而迁移学习可从源域的小数据中抽取并迁移知识,用来完成新的学习任务。2)解决个性化问题。王凯等^[8]提出利用迁移学习方法可以有效解决小样本条件下引起的分类不佳问题,泛化了模型的适应性。陈淑英等^[9]提出了基于迁移学习的多变量预测模型(TVPMCD)进行识别,有效解决了小样本的故障诊断问题。因此,迁移学习可以解决小样本的网络训练分类的问题。宫文峰等^[10]提出改进卷积神经网络的全连接层极大减少了训练参数和计算时间,实现轴承微小故障诊断。

因此,本文提出时频分析和 VGG19 模型迁移学习的轴承故障诊断方法,有效地解决了振动信号小样本训练的问题,也避免了传统的依赖专家经验的故障诊断方法,将数据分类问题转换为图像分类问题实现了信号处理与深度学习的结合,即使不是信号处理专家也可以实现故障诊断问题。同时,迁移学习利用经典的网络进行迁移用较少的数据样本来训练修改层并微调网络就可以实现分类问题和智能化的诊断,也避免花费较多的硬件资源。

1 典型的网络结构

卷积神经网络(convolutional neural network,CNN)是一种深度前馈神经网络,经典的结构包括:输入层、卷积层、池化层(采样层)、全连接层、输出层,卷积层和池化层交替搭建网络结构实现对图像特征的提取和降维操作。本文采用时频分析方法将一维振动数据转化为时频图,输入到卷积神经网络中输出目标,如图 1 所示。

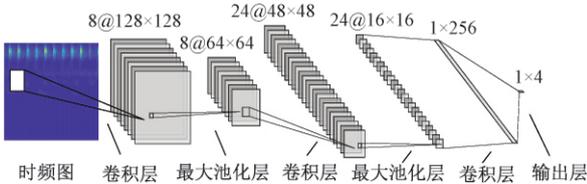


图 1 CNN 基本网络结构

卷积神经网络模型 VGG19,VGGNet 是牛津大学的 VGG(visual geometry group)小组在 2014 年的 ImageNet 大赛中提出的一种泛化能力强、结构清晰的网络模型,VGG19 模型在 ILSVRC-2014 大赛中以图像分类任务 Top-5 的错误率为 7.3%,在分类项目中获得第二名和定位项目中获得第一名的成绩,作为基础网络结构在迁移学习任务

中特征提取能力非常强大。VGGNet 结构中的所有卷积核为 3×3,池化核为 2×2。VGGNet 中的 3×3 卷积核可以更好地提取图片中的局部信息,在 VGG 网络中使用连续的 2 组 3×3 的卷积核(stride 分别为 2、1)的网络结构可以和 7×7 的卷积核产生相同的效果。VGG 团队为了测试网络的性能设计了 6 种不同的结构如图 2 所示。

ConvNet Configuration					
A	A-LPN	B	C	D	E
11 weight layers	11 weight layers	13 weight layers	16 weight layers	16 weight layers	19 weight layers
input (224×224 RGB image)					
conv3-64	conv3-64 LRN	conv3-64	conv3-64	conv3-64	conv3-64
maxpool					
conv3-128	conv3-128	conv3-128	conv3-128	conv3-128	conv3-128
maxpool					
conv3-256	conv3-256	conv3-256	conv3-256	conv3-256	conv3-256
conv3-256	conv3-256	conv3-256	conv3-256	conv3-256	conv3-256
maxpool					
conv3-512	conv3-512	conv3-512	conv3-512	conv3-512	conv3-512
conv3-512	conv3-512	conv3-512	conv3-512	conv3-512	conv3-512
maxpool					
conv3-512	conv3-512	conv3-512	conv3-512	conv3-512	conv3-512
conv3-512	conv3-512	conv3-512	conv3-512	conv3-512	conv3-512
maxpool					
FC-4096					
FC-4096					
FC-1000					
soft-max					

图 2 VGGNet 结构^[11]

本文将网络结构中的全连接中的第 1 层 FC-4096 改成 3×3 的卷积核(stride 为 2),第 2 层 FC-4096 改成 3×3 的小卷积核(stride 为 1),第 3 层 FC-1000 改成 FC-4;全连接层的网络替换使得参数比以前的 VGG19 参数少很多,提高了运算效率。具体卷积过程如图 3 所示。

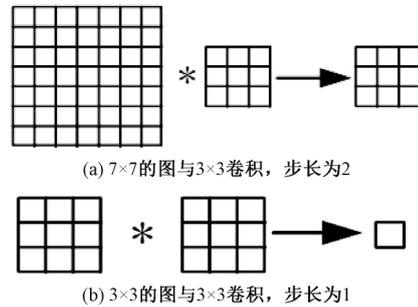


图 3 卷积过程

2 时频图与卷积神经网络的轴承故障识别

2.1 特征提取与数据处理

时频分析技术通常有短时傅里叶变换(STFT)、魏格纳-维尔(Wigner-Ville)分布、小波变换(WT)、小波包分解(WPD)等方法。凯斯西储大学轴承数据库(Case Western Reserve University, CWRU)的滚动轴承振动加速度数据,采样频率为 12 000 Hz,在这里采用的数据类型状态分别为:正常状态、外圈故障(0.177 8 mm)、内圈故障(0.177 8 mm)、滚动故障(0.177 8 mm)四种状态。在这里使用数据每种

状态为 48 000 个数据点,每 1 200 个数据点生成一个样本,然后依次进行连续小波变换(CWT)、小波包(WPD)时频转换、Wigner-Ville 分布(WVD)生成时频图,谱峭度(SK)生成图。以外圈故障为例,生成的图如图 4 所示。

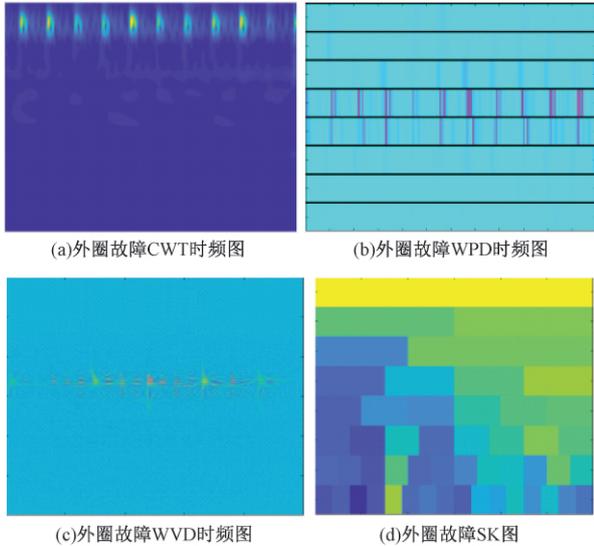


图 4 外圈故障图

2.2 实验步骤流程

对于数据处理技术有:批量归一化(batch normalizaion, BN)^[12]、Dropout^[13-14]、数据增强包括旋转、翻转、缩放、平移等操作。

1)采用连续小波分析、小波包时频分析、魏格纳-维尔(Wigner-Ville)分布、谱峭度分别生成 4 种不同状态的 RGB 特征图。

2)时频分析预处理通过 MATLAB 函数 Image Datastore 将特征图尺寸调整为 224×224×3。

3)预处理后的时频图作为 VGGNet 网络的提取器作为迁移层,全连接层改为 Conv3-512(stride 2)接 Conv3-512(stride 1)接 FC-4,接 soft-max 作为分类器,训练网络模型。具体步骤如图 5 所示。

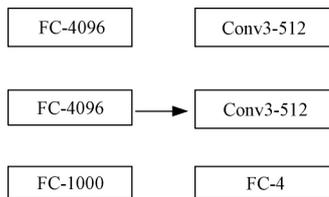


图 5 网络替换结构

4)根据模型的分类型准确率对网络进行参数优化,选择一个分类型准确率较好的网络。具体的算法流程如图 6 所示。

3 仿真结果及分析

平台为 GTX1050Ti 4G 运行环境和 MATLAB 软件,

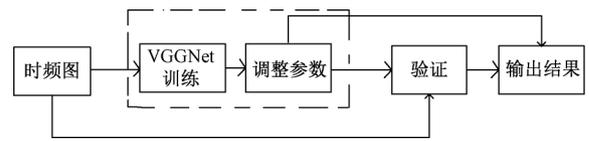


图 6 算法流程

利用经典的深度学习框架进行卷积神经网络的搭建。在图像数据存储中随机地将数据分为 70% 的训练集,30% 的测试集,而网络要求输入 224×224×3 尺寸的图像,但是图像存储的尺寸大小有所不同,使用图像数据存储函数可以自动调整为训练图像的大小。为了防止过拟合和训练图像中的细节特征提取,在这里将训练图像沿垂直轴方向随机翻转以及在水平和垂直方向上随机平移 30 pixel 的操作。

由于在模型的训练过程中,超参数的选择对模型的训练速度、分类准确率有直接的影响^[15]。从本文所构建的 CNN 模型及其任务来讲,将 VGG19 的特征提取器作为迁移网络层,从优化器^[16]、学习率、批处理样本数目 3 种参数配置因素重新训练网络模型来验证模型分类准确率的影响,来确定最优的模型匹配参数。

3.1 实验参数配置

从优化器、初始学习率、MinBatchSize 三个因素来对网络模型分类结果进行统计。其中 CWT_VGG 代表连续小波变化时频图的迁移网络分类,WPD_VGG 代表小波包时频图的迁移网络分类,WVD_VGG 代表 Wigner-Ville 分布时频图的网络分类,SK_VGG 代表谱峭度图的网络分类。

1)优化器类型

首先考虑所构建的 CNN 模型和目标分类任务,利用 MATLAB 深度学习框架,选择 Sgdm(SGD+M)、Adam 及 Rmsprop 三种优化算法,每组优化器分别随机训练 10 次,对 4 种样本平均识别准确率作为评价指标,结果如表 1 所示。优化器选择 Sgdm 的平均识别率为 99.43% 最高,作为本文的优化器。

优化器	不同优化器的模型分类结果				平均识别率
	CWT_VGG	WPD_VGG	WVD_VGG	SK_VGG	
Sgdm	100	99.79	100	97.92	99.43
Adam	99.79	97.29	99.58	93.54	97.55
Rmsprop	99.79	80.63	95.00	91.46	91.72

2)初始学习率

设置初始学习率为 0.001、0.000 1、0.000 01 分组,每组随机训练 10 次,对 4 种样本平均识别准确率作为评价指标,结果为如表 2 所示。初始学习率为 0.000 1 的平均识别率为 99.43% 最高,作为本文的初始学习率。

表 2 不同学习率的模型分类结果 %

初始学习率	CWT_VGG	WPD_VGG	WVD_VGG	SK_VGG	平均识别率
0.001	50.63	25.00	61.25	88.33	56.30
0.000 1	100	99.79	100	97.92	99.43
0.000 01	100	100	99.38	92.71	98.02

3)MinBatchSize

设置最小批处理为 8、16、32,每组随机运行 10 次,对 4 种样本平均识别准确率作为评价指标,如表 3 所示。MinBatchSize 为 16 时的平均识别率为 99.43%最高,作为本文的批量处理样本数量。

表 3 不同批处理的模型分类结果 %

最小批处理	CWT_VGG	WPD_VGG	WVD_VGG	SK_VGG	平均识别率
8	100	99.79	100	97.50	99.32
16	100	99.79	100	97.92	99.43
32	100	100	100	94.58	98.65

3.2 实验结果分析

由表 1~3 可以看出在相同的网络配置中,连续小波时频图、小波包时频图、Wigner-Ville 分布时频图的网络分类率均高于谱峭度图的网络识别率。如表 3 所示,当 MinBatchSize 为 32 时可以看出时频图的分分类率高谱峭图分类率达 5.42%。通过 3 个表中所有平均识别结果中的时频图对于迁移学习的网络分类高于谱峭度图的,说明滚

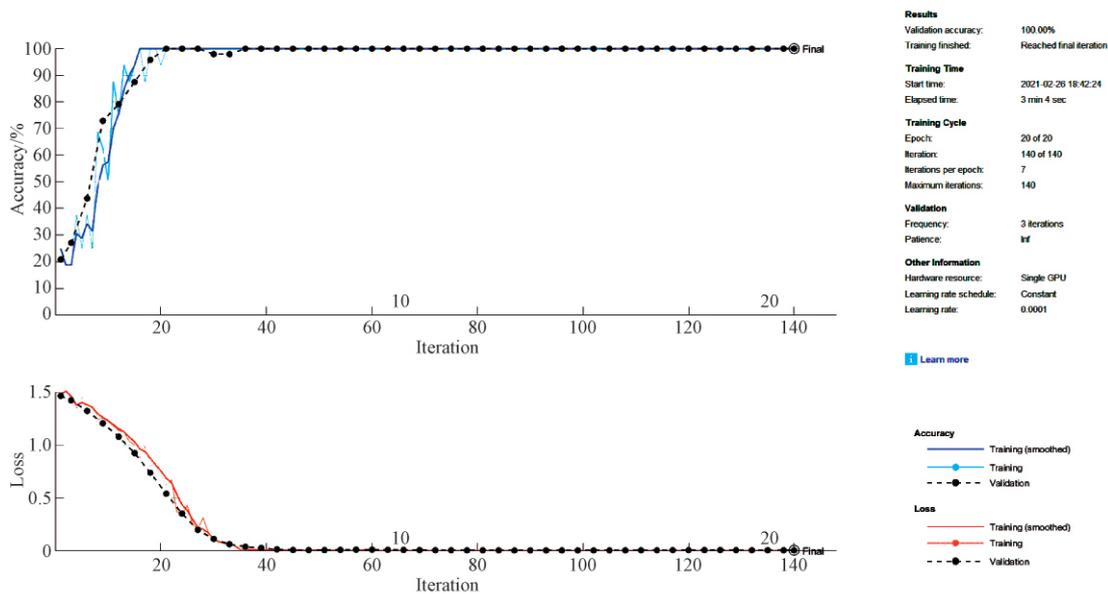
动轴承故障数据时频样本处理更适合网络识别。通过从优化器、初始学习率、MinBatchSize 三个因素的配置网络模型分类结果优化器选择 Sgdm、初始学习率为 0.000 1、MinBatchSize 为 16 时训练的模型分类结果比较好。将选择好的网络模型配置实验结果参数如表 4 所示。

表 4 网络模型部分参数设置

参数	参数值
优化器	Sgdm
初始学习率	0.000 1
最小批处理	16
最大轮	20
验证频次	3
执行环境	GPU

将上述配置参数微调网络随机抽取样本做实验,这里以连续小波变换时频图样本和去峭度图样本为例的训练过程如图 7 所示,其中 CWT_VGG 代表连续小波变化时频图的迁移网络分类训练过程,SK_VGG 代表谱峭度图的网络分类训练过程。

从图 7 中可以看出用连续小波生成的时频和迁移学习网络的分类效果为 100%,轴承类型的 4 种状态类型可以区分开来;而用峭度谱图结合迁移学习的网络分类为 91.67%,不能完全将故障类型准确分类。因此,用连续小波变换生成时频图样本的方法比谱峭度图的迁移学习网络分类效果好。



(a) CWT_VGG 训练过程

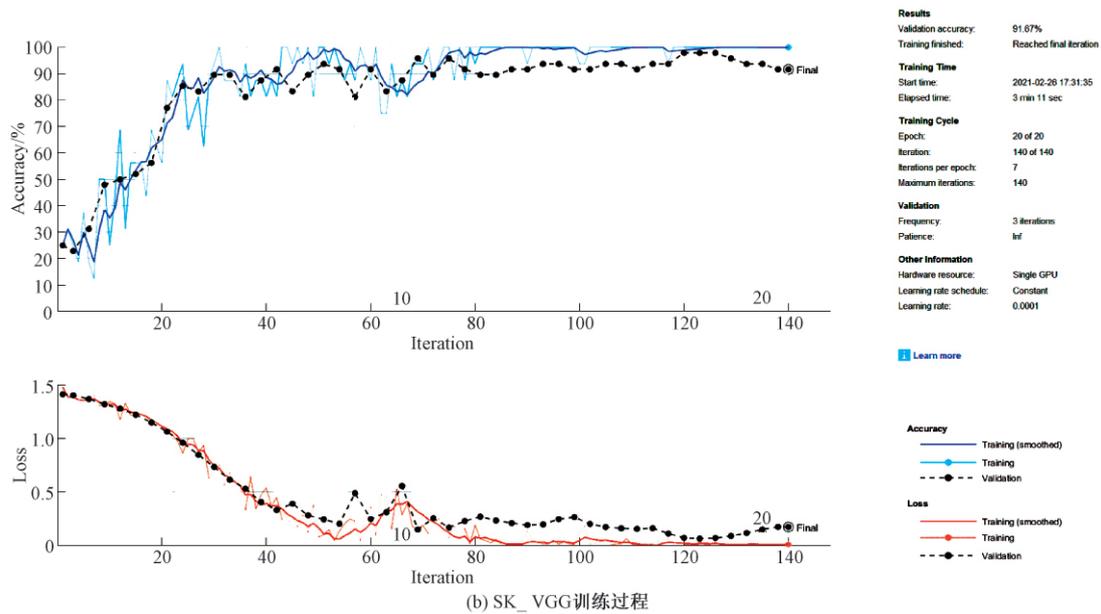


图 7 训练过程

4 结 论

本文研究的时频分析与 VGG19 迁移网络的滚动轴承故障诊断模型是一种卷积神经网络的故障诊断模型。通过采用轴承故障不同类型的数据做实验,结果表明时频图样本的卷积神经网络迁移学习的诊断识别率高于谱峭度图的卷积神经网络的迁移学习的诊断识别率,连续小波时频图的卷积神经网络迁移学习模型的识别率最高。利用现有的经典网络模型可以实现用较小的样本训练修改网络层并微调模型实现网络分类,在工程领域具有一定较高的应用价值,避免过多的时间成本和硬件开销。在迁移学习网络中振动信号小样本存在过拟合现象,参数的网络优化选取没有统一的标准需要根据分类任务要求凭借经验微调,参数的选择直接影响网络分类效果这是研究的难点和方向。

参考文献

- [1] 中国工程院周济院长关于“新一代智能制造——新一轮工业革命的核心驱动力”的主题报告[J]. 起重运输机械, 2018(1): 44.
- [2] 唐波,陈慎慎. 基于深度卷积神经网络的轴承故障诊断方法[J]. 电子测量与仪器学报, 2020, 34(3): 88-93.
- [3] SHENG G, TAO Y, WEI G, et al. A novel fault diagnosis method for rotating machinery based on a convolutional neural network [J]. Sensors (Basel, Switzerland), 2018, 18(5): 1429.
- [4] 黄驰城. 结合时频分析和卷积神经网络的滚动轴承故障诊断优化方法研究[D]. 杭州: 浙江大学, 2019.
- [5] 唐藁. 基于时频分析和 CNN 的滚动轴承故障诊断方法研究[D]. 杭州: 浙江大学, 2019.
- [6] 杜小磊,陈志刚,张楠,等. 基于同步挤压 S 变换和深度学习的轴承故障诊断[J]. 组合机床与自动化加工技术, 2019(5): 90-93, 97.
- [7] 刘鑫鹏,栾悉道,谢毓湘,等. 迁移学习研究和算法综

述[J]. 长沙大学学报, 2018, 32(5): 28-31, 36.

- [8] 王凯,李元辉. 迁移学习在机械设备预测性维护领域的应用综述[J]. 中国仪器仪表, 2019(12): 64-68.
- [9] 陈淑英,王利群. 基于迁移 VPMCD 的滚动轴承故障诊断方法[J]. 电子测量与仪器学报, 2019, 33(3): 93-98.
- [10] 宫文峰,陈辉,张美玲,等. 基于深度学习的电机轴承微小故障智能诊断方法[J]. 仪器仪表学报, 2020, 41(1): 195-205.
- [11] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition [J]. ArXiv Preprint, 2014, ArXiv: 1409.1556.
- [12] IOFFE S, SZEGEDY C. Batch normalization: Accelerating deep network training by reducing internal covariate shift[C]. Proceedings of the 32nd International Conference on Machine Learning, 2015: 448-456.
- [13] WAN L, ZEILER M, ZHANG S, et al. Regularization of neural networks using DropConnect [C]. International Conference on Machine Learning, 2013: 1058-1066.
- [14] SRIVASTAVA N, HINTON G, KRIZHEVSKY A, et al. Dropout: A simple way to prevent neural networks from overfitting [J]. Journal of Machine Learning Research, 2014, 15: 1929-1958.
- [15] 张荣,李伟平,莫同. 深度学习研究综述[J]. 信息与控制, 2018, 47(4): 385-397, 410.
- [16] 李俊,刘永葆,余又红. 卷积神经网络和峭度在轴承故障诊断中的应用[J]. 航空动力学报, 2019, 34(11): 2423-2431.

作者简介

李传鑫,硕士研究生,主要研究方向为信号处理。

E-mail: 2690451468@qq.com

刘增力,教授,博士生导师,主要研究方向为现代信号处理及其应用、水声信号处理、DSP 技术应用。

E-mail: liuzengli@hotmail.com