

基于卷积神经网络的汽车异响识别方法研究

施宇豪 许焕敏

(河海大学, 常州 213022)

【摘要】针对目前采用主观评价方法进行汽车异响诊断效率较低的问题,提出了一种基于卷积神经网络(CNN)的汽车异响识别方法。借助四立柱振动台在半消声室环境下采集信噪比较高的原始异响信号作为研究对象,提取信号的梅尔频谱(Mel-Spectrogram)作为神经网络的输入,然后搭建卷积神经网络对数据进行深层次的特征提取和压缩并进行分类识别,平均识别率可达90.5%,最后运用迁移学习的方法对模型进行优化,试验结果表明,VGG和ResNet模型均可以提高识别精度,且ResNet网络在测试集上的识别效果更好。

关键词:汽车异响 卷积神经网络 特征提取 分类识别 迁移学习

中图分类号:TP206 文献标志码:A DOI: 10.20104/j.cnki.1674-6546.20230178

Research on Vehicle Abnormal Sounds Recognition Based on Convolutional Neural Network

Shi Yuhao, Xu Huanmin

(Hohai University, Changzhou 213022)

【Abstract】To improve the low efficiency of the current diagnosis method using subjective evaluation method, this paper proposed a method for identifying vehicle abnormal noise based on Convolutional Neural Network (CNN). A four post vibration test stand was used to collect raw abnormal noise signals with high signal-to-noise ratio in a semi anechoic chamber environment as the research object to extract the Mel-spectrogram of the signal as input to the neural network, then, a convolutional neural network was constructed to perform deep level feature extraction, compression, and classification recognition on the data. It was found that the average recognition rate can reach 90.5%. Finally, the transfer learning method was used to optimize the model. The test results indicate that VGG and ResNet models can improve recognition accuracy, and the ResNet network has better recognition performance in the test set.

Key words: Vehicle abnormal sounds, Convolutional Neural Network (CNN), Feature extraction, Classification recognition, Transfer learning

【引用格式】施宇豪,许焕敏.基于卷积神经网络的汽车异响识别方法研究[J].汽车工程师,2023(12):37-42.

SHI Y H, XU H M. Research on Vehicle Abnormal Sounds Recognition Based on Convolutional Neural Network [J]. Automotive Engineer, 2023(12): 37-42.

1 前言

汽车异响会影响驾乘体验,甚至导致严重的故障。在整车研发阶段,主要通过振动台架试验与主观评价相结合的方法检查汽车异响问题,对试验人员的经验要求较高,且经过长时间的主观判断后,识别效率会大幅降低。利用人工智能技术开展声

音识别可以实现异响的自动诊断,并快速准确地判断异响种类,从而极大地提高识别效率,对改善汽车NVH性能具有重要意义。

声音识别技术广泛应用于语音识别、环境声音识别和机器声音识别等领域。构建声音识别模型是实现声音识别的重要基础,常用的声学模型包括隐马尔可夫模型(Hidden Markov Model, HMM)、高

斯混合模型(Gaussian Mixture Model, GMM)、支持向量机(Support Vector Machine, SVM)和人工神经网络(Artificial Neural Network, ANN)^[1]。杨军磊基于GMM识别模型和梅尔频率倒谱系数(Mel Frequency Cepstral Coefficient, MFCC)特征实现了对空调机组噪声的识别,从而进一步判断设备故障类型^[2]。杨兴国提出一种基于SVM分类器的发动机异响故障识别方法,提取异响声音的小波包能量谱和双谱峰值作为特征参数,可以较好地识别异响类型^[3]。

HMM、GMM、SVM与浅层神经网络等均属于浅层机器学习的模型,在处理一些特征较为复杂的数据时,存在数据表征能力不强的问题^[4],导致识别效果不佳。深度学习方法一经提出就在图像、声音识别等领域得到广泛应用。朱梦帆提取了电力设备故障声音声谱图的特征,并自建卷积神经网络(Convolutional Neural Network, CNN)模型完成了对电力设备不同放电故障声音和正常放电声音的分类^[5]。Li提出了一种具有多个卷积层的新型堆叠CNN模型搭配对数梅尔谱(Log-Mel)特征输入,使得对环境事件声音的识别效果优于采用其他特征输入的模型^[6]。章杰设计了一种基于残差网络(Residual Network, ResNet)和门控循环单元(Gated Recurrent Unit, GRU)的模型结构,在车辆噪声识别方面实现了96%的准确率^[7]。

本文借助四立柱振动试验台架采集了多辆试验车的异响数据,将同类异响问题进行标注,针对异响数据内部结构复杂的特点,提取其梅尔频谱(Mel-Spectrogram),采用卷积神经网络对特征参数进行深层次的学习以实现其分类。

2 汽车异响数据集建立

2.1 汽车异响试验环境

本文试验在整车异响试验室进行,试验室内部为半消声室结构,大幅减少了背景噪声对原始信号的干扰,如图1所示。四立柱试验台通过4台独立加振机控制,通过将加速度信号、位移信号等进行多次迭代来复现垂直方向的振动,从而进行整车四通道道路模拟试验。4台液压伺服作动器分别对4个车轮施加振动,可按设定的规律或采集的路谱特征对待测车辆施加激振力,从而激起待测车辆的异响。



图1 汽车异响试验室

2.2 异响信号采集试验

本文的试验对象为来自多个整车制造商的车型,并采用多条路谱进行激励,以保证异响样本来源的多样性,使数据库更具代表性。试验采用的设备如表1所示。

表1 测试仪器

序号	仪器名称	型号
1	整车四立柱试验台	
2	精密声传感器	G.R.A.S.46AE
3	标定器	B&K 4228
4	LMS数据采集前端	SCM09 SERIAL

人耳听觉的频率范围一般为20~20 000 Hz,结合奈奎斯特采样定理,试验时信号采样频率设置为44 100 Hz,采样值可包含原始信号的所有信息。测试前调整标定器输出频率和输出声压级分别为1 000 Hz、114 dB,然后插入声传感器进行标定,从而找到最适应当前试验环境的灵敏度,最后将声传感器放置在距离采样点5~20 cm处固定。

在控制端软件中输入迭代好的路谱来模拟试验车在不同路面上的行驶状态,在异响工程师判断出异响问题后即可开始信号采集。

2.3 异响数据集整理

本文采集了减振器异响、安全气囊异响、发动机罩异响、空调出风口异响4种测试阶段常见的异响数据。另外,考虑到实车道路异响识别会被发动机声和周围车辆的汽笛声干扰,因此加入了Urbansound8K数据集中的发动机空转声和汽笛声来避免异响识别系统的误触发,并提高训练模型的性能。将上述几种数据以6:2:2的比例随机分割为训练集、验证集、测试集用于模型训练及测试。

本文的汽车异响识别总体流程如图2所示。

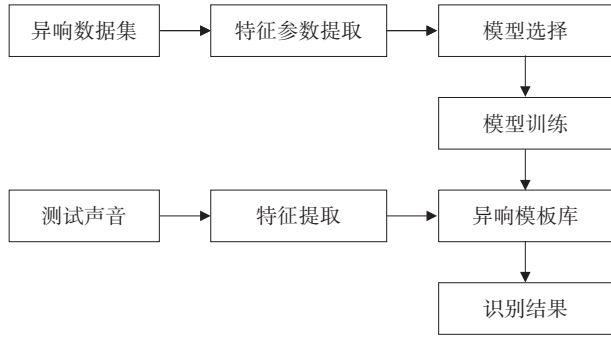


图2 异响识别流程

3 异响信号的特征提取

CNN的输入是图像,因此需要提取异响信号的特征图谱。以往,由于机器学习算法无法处理复杂的图像信息,HMM、GMM等模型多将梅尔频率倒谱系数用于特征提取。深度学习算法不需过多的预处理来提炼信息,故可以直接提取梅尔频谱图作为CNN的输入^[8],提取流程如图3所示。

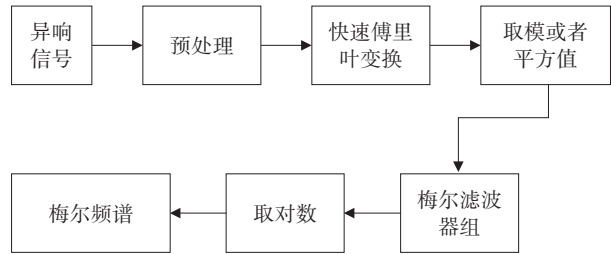


图3 特征提取流程

梅尔频谱参数的计算过程如下:

a. 预处理。预处理包括预加重、分帧、加窗函数。预加重的目的是补偿高频信息的损失,通常采用一阶滤波器 $H(z)$ 实现:

$$H(z)=1-\alpha z^{-1} \quad (1)$$

式中, α 为常数,一般取0.9~1.0。

分帧的目的是使相邻帧之间的参数平稳过度,因此两帧之间存在重叠部分,加窗函数可以减少信号频域的谱泄漏^[9]。语音信号 $x(n)$ 经预处理后的第 i 帧信号为 $x_i(m)$ 。

b. 快速傅里叶变换(Fast Fourier Transform, FFT)。对每一帧信号进行快速傅里叶变换,将信号从时域转化到频域:

$$X(i,k)=f_{\text{FFT}}[x_i(m)] \quad (2)$$

式中, $X(i,k)$ 为第 i 帧信号中第 k 条频域谱线的频域数据; f_{FFT} 为快速傅里叶变换。

c. 计算傅里叶变换后第 i 帧信号中第 k 条频域谱线的能量谱:

$$E(i,k)=|X(i,k)|^2 \quad (3)$$

d. 计算通过梅尔(Mel)滤波器组的第 i 帧信号第 m 个三角带通滤波器的对数能量谱:

$$S(i,m)=\ln\left(\sum_{k=0}^{N-1} E(i,k)H_m(k)\right) \quad (4)$$

式中, $H_m(k)$ 为通过第 m 个三角带通滤波器的传递函数。

这里得到的对数能量为每一帧信号的对数梅尔谱特征,即梅尔频谱^[10],后续可用作CNN的输入参数。

4 基于卷积神经网络的汽车异响识别

4.1 卷积神经网络

CNN在图像分类上的应用非常成功,本文对声音信号提取特征图谱后,使用CNN进行声音识别、自然语言处理等研究。

CNN模型中使用卷积层和池化层代替传统神经网络中的隐藏层,通过局部感知、权值共享、池化采样实现对数据样本深层特征的提取与参数的简化^[11]。典型的卷积神经网络由输入层、卷积层、池化层、全连接层及输出层构成,其基本结构如图4所示。

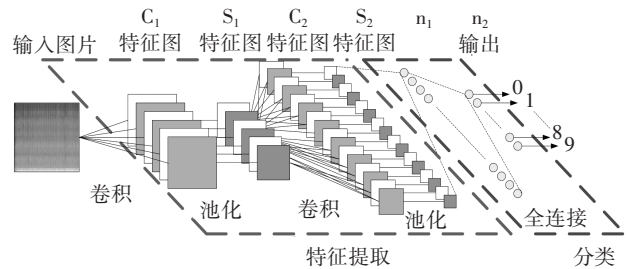


图4 典型卷积神经网络

输入图像经过卷积层进行深层次的特征提取,由池化层对特征参数进行压缩凝练,在多次卷积池化操作后使用全连接层将所有参数展开成一维,通过层层神经元之间的输入、输出调整权重,最终得到各输出类别的概率。

4.2 异响识别的CNN模型

本文搭建CNN异响识别模型,记为Model1。输入图片大小为 224×224 ,然后进行3次卷积池化,卷积核尺寸为 3×3 ,步长为1,池化层选择最大池化,池化窗口大小为 2×2 ,激活函数选择线性整流函数(Rectified Linear Unit, ReLU),最后用Softmax函数实现六分类。模型结构如表2所示。

表2 卷积神经网络结构

层	输出结构(Output Shape)
卷积层1(Conv2d_1)	(224,224,16)
池化层1(MaxPooling2D_1)	(112,112,16)
卷积层2(Conv2d_2)	(112,112,32)
池化层2(MaxPooling2D_2)	(56,56,32)
卷积层3(Conv2d_3)	(56,56,16)
池化层3(MaxPooling2D_3)	(28,28,16)
展平层(Flatten)	(12 544)
全连接层1(Dense_1)	(512)
全连接层2(Dense_2)	(6)

模型学习率设置为0.001,损失函数设为多分类交叉熵,反向传播的优化器选择Adam,并使用准确率(Accuracy)函数绘制训练及验证精度曲线。训练轮次(Epoch)设置为100。

4.3 模型训练及结果分析

将前文提取的梅尔频谱通过图像数据生成器(ImageDataGenerator)模块的旋转、平移等操作完成预处理,从而变相地增加了训练样本,通过Reshape函数修改图像分辨率为224×224,然后开始训练模型。完成100个轮次训练后的验证精度和损失如图5、图6所示。

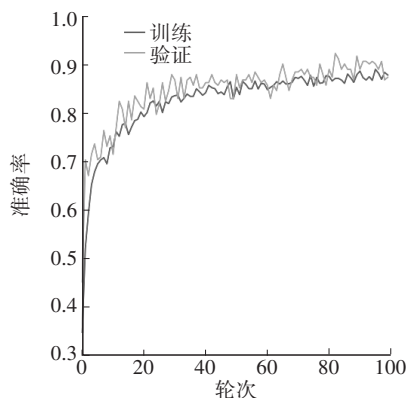


图5 训练及验证精度

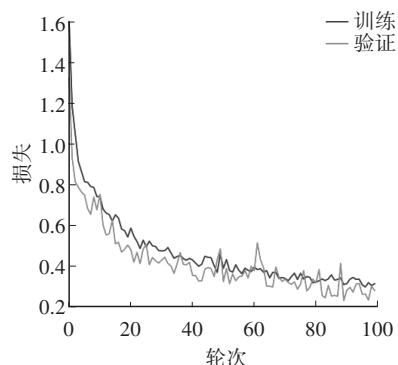


图6 训练及验证损失

由图5可以看出,训练及验证精度在前20个轮次左右上升较快,此后趋向平缓,处于缓慢上升过

程,损失精度与之相反。训练过程中的最高精度停留在89.06%,验证的最高精度约为90.07%,略高于训练精度,精度最高的模型在第95轮。

4.4 异响识别试验

模型训练完成后,使用精度最高的模型识别测试集数据,以考查模型的泛化能力。

在模型训练的基础上开发识别程序,在程序中导入训练好的.h5模型文件,只需输入测试集数据的相对路径即可展示识别结果,其中单个测试样本的识别结果如图7所示。

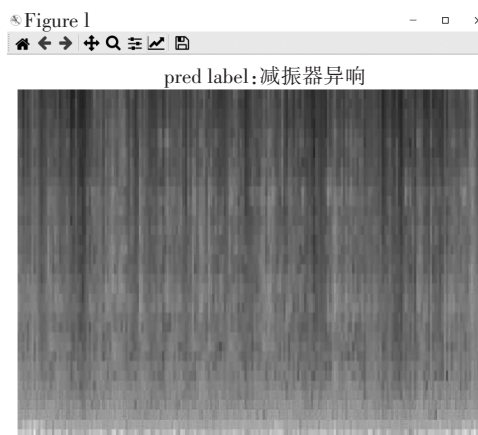


图7 单个测试样本识别结果

混淆矩阵是一个误差矩阵,可以用来进行可视化评估,从而监督学习算法的性能,如图8所示。对角线上的数据可以理解为各类别的识别率,其余部分则为误差数据。例如,减振器异响中有2%被错误识别为发动机空转声、2%被识别为安全气囊异响、1%被识别为空调异响。

实际 \ 预测	减振器异响	发动机空转	安全气囊异响	发动机罩异响	汽笛声	空调异响
减振器异响	0.86	0.02	0.02	0.00	0.00	0.1
发动机空转	0.1	0.88	0.00	0.00	0.00	0.01
安全气囊异响	0.00	0.05	0.95	0.00	0.00	0.00
发动机罩异响	0.00	0.05	0.00	0.93	0.00	0.02
汽笛声	0.00	0.00	0.00	0.00	0.97	0.03
空调异响	0.00	0.1	0.00	0.05	0.00	0.84

图8 混淆矩阵

本文搭建的异响识别CNN模型的平均识别精度达到90.5%,其中汽笛声和气囊异响识别率较高,

达95%以上,而其余几类大多低于90%,因此,该模型性能仍有提高空间,可通过调整参数进一步优化。

4.5 基于迁移学习的优化方法

迁移学习作为机器学习的一种常用方法,多用于计算机视觉和自然语言处理等任务。迁移学习通过将预训练模型置于目标任务中来帮助训练可靠的决策函数,从而解决目标域中已标记样本较少的问题^[12]。异响信号的采集较为繁琐,很难获得大量的已标注样本,因此将迁移学习的方法引入汽车异响识别具有一定可行性。本文选择基于模型的迁移学习方法,以VGG16、ResNet50作为预训练模型,使用2种经典模型的卷积池化层来提取特征参数,修改Softmax函数的分类数量来满足试验的具体需求。

VGG模型^[13]广泛用于图像分类问题,其模型结构如图9所示。

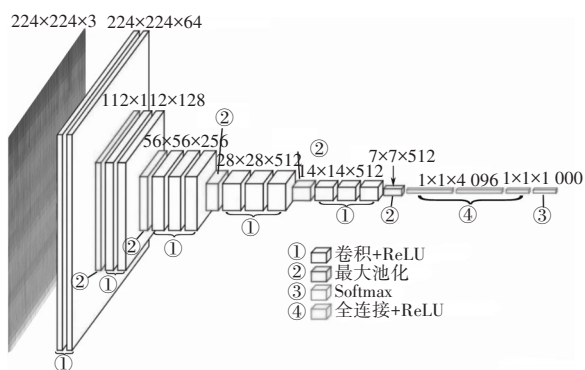


图9 经典VGG网络模型

ResNet网络使用层间残差跳连的方式传递信息,可以解决图像经多层卷积提取特征后,特征图像丢失原图本质信息的问题,有效抑制了网络退化^[14]。ResNet结构块如图10所示。

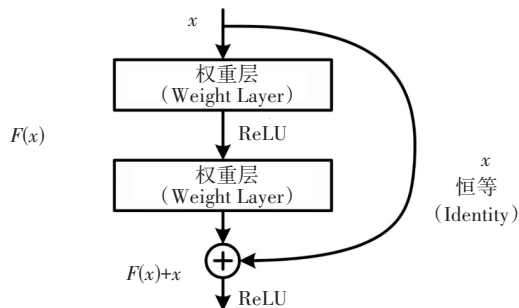


图10 残差块

ResNet块在层与层的传递中给出2条路线,如果卷积的效果不佳,就将当次卷积重置,直接将上一层参数输入给下一层,以避免某次卷积效果较差而影响后续训练。

在使用上述2种模型训练时,学习率、优化器、激活函数、训练轮次等参数均不变。汇总各模型的训练精度、验证精度、平均测试精度如表3所示。

表3 不同模型各精度汇总 %

模型结构	训练精度	验证精度	测试精度
Model1	89.06	90.07	90.50
VGG16	93.41	94.26	93.33
ResNet50	93.55	94.70	94.33

由表3可知,迁移学习引入的模型在训练、验证、测试精度上均获得了提升。相较于自建模型,VGG16模型的识别率提升了2.83个百分点,ResNet50的识别率提升了3.83个百分点。

5 结束语

本文提出了基于卷积神经网络的汽车异响识别方法,经数据采集、特征提取、模型训练等流程实现异响识别。首先,在整车异响试验室中采集高信噪比的异响数据,然后,在传统梅尔频率倒谱系数的基础上去除离散余弦变化,提取了更适合作为神经网络输入的梅尔频谱,再通过搭建模型,调整模型参数来观察模型训练的精度和泛化能力,最后,基于迁移学习的方法解决了训练样本数据不足的问题。试验结果表明:本文提出的汽车异响识别方法可以较好地识别异响问题,在原始数据集上的识别率可达90.5%;迁移学习可以在小样本异响识别中取得更好的效果,VGG16模型的识别率达到93.33%,ResNet50的识别率达到94.33%。

参考文献

- 卢林,王东. 浅谈声音识别模型发展趋势[J]. 汽车实用技术, 2021, 46(12): 186-188.
LU L, WANG D. On the Development Trends of Sound Recognition Models[J]. Automotive Practical Technology, 2021, 46(12): 186-188.
- 杨军磊. 基于噪声的设备故障识别[D]. 武汉: 华中科技大学, 2016.
YANG J L. Equipment Fault Identification Based on Noise [D] Wuhan: Huazhong University of Science and Technology, 2016.
- 杨兴国. 基于小波包和双谱的发动机异响模式识别[J]. 噪声与振动控制, 2018, 38(6): 146-149+177.
YANG X G. Engine Abnormal Noise Pattern Recognition Based on Wavelet Packet and Bispectrum[J]. Noise and Vibration Control, 2018, 38(6): 146-149+177.
- TRAN H D, NG W Z T. Generalized Gaussian Distribution

- Kullback-Leibler Kernel for Robust Sound Event Recognition[C]// 2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). Florence, Italy: IEEE, 2014(1): 5949-5953.
- [5] 朱梦帆. 基于卷积神经网络的电力设备放电故障声音识别研究[D]. 抚州: 东华理工大学, 2022.
- ZHU M F. Research on Sound Recognition of Power Equipment Discharge Fault Based on Convolutional Neural Network[D]. Fuzhou: East China University of Technology, 2022.
- [6] LI S B, YAO Y, HU J, et al. An Ensemble Stacked Convolutional Neural Network Model for Environmental Event Sound Recognition[J]. Applied Sciences, 2018, 8(7).
- [7] 章杰. 基于 ResNet 和 GRU 的噪声车辆识别[D]. 阜阳: 阜阳师范大学, 2021.
- ZHANG J. Noise Vehicle Recognition Based on ResNet and GRU[D]. Fuyang: Fuyang Normal University, 2021.
- [8] 邱志斌, 卢祖文, 王海祥, 等. 基于 Mel 频谱图和 CNN 的电网涉鸟故障鸟声识别[J]. 华南理工大学学报(自然科学版), 2022, 50(2): 129-136.
- QIU Z B, LU Z W, WANG H X, et al. Bird Sound Recognition for Power Grid Bird Related Faults Based on Mel Spectrum and CNN[J]. Journal of South China University of Technology (Natural Science Edition), 2022, 50(2): 129-136.
- [9] 宋立新, 孙东梓, 安佳星, 等. 离散傅里叶变换泄露及其加窗抑制仿真实验设计[J]. 实验室研究与探索, 2018, 37(7): 106-109.
- SONG L X, SUN D Z, AN J X, et al. Simulation Experiment Design of Discrete Fourier Transform Leakage and Windowing Suppression[J]. Laboratory Research and Exploration, 2018, 37(7): 106-109.
- [10] 张留军, 王玫, 罗丽燕. 基于改进对数梅尔谱特征的街道环境声事件检测方法[J]. 桂林电子科技大学学报, 2020, 40(5): 411-417.
- ZHANG L J, WANG M, LUO L Y. A Street Environmental Sound Event Detection Method Based on Improved Logarithmic Mel Spectral Features[J]. Journal of Guilin University of Electronic Technology, 2020, 40(5): 411-417.
- [11] 周飞燕, 金林鹏, 董军. 卷积神经网络研究综述[J]. 计算机学报, 2017, 40(6): 1229-1251.
- ZHOU F Y, JIN L P, DONG J. Review of Convolutional Neural Network[J]. Journal of Computer Science, 2017, 40(6): 1229-1251.
- [12] 周旷, 姜名. 基于迁移学习的小样本目标识别研究进展与展望[J]. 航空科学技术, 2023, 34(2): 1-9.
- ZHOU K, JIANG M. Research Progress and Prospect of Small Sample Target Recognition Based on Transfer Learning[J]. Aerospace Science and Technology, 2023, 34(2): 1-9.
- [13] BAE J, KIM M, LIM J S. Combining VGG16 Convolutional Layer and Neuro-Fuzzy Network Model for Classifying Normal and Abnormal Heart Sounds[J]. Basic & Clinical Pharmacology & Toxicology, 2020, 126.
- [14] HE K M, ZHANG X Y, REN S Q, et al. Deep Residual Learning for Image Recognition[C]// 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas, NV, USA: IEEE, 2016.

(责任编辑 斛 畔)

修改稿收到日期为 2023 年 5 月 23 日。