

DOI:10.19344/j. cnki. issn1671-5276.2024.01.054

基于声音信号的汽车安全带卷收器质量检测方法研究

刘洪达, 左敦稳, 王勇, 靳萌萌

(南京航空航天大学 机电学院, 江苏 南京 210016)

摘要:为提高汽车安全带生产现场质量检测效率,根据《QC/T987—2014 汽车安全带卷收器性能要求和试验方法》搭建实验平台,采集卷收器合格品与次品工作过程中的声音信号,将卷积注意力模块(CBAM)嵌入残差网络(ResNet-18)残差块之前,设计CBAM-ResNet-18“Before Blocks”模型,对采集到的卷收器声音信号进行分类。与不加注意力机制的ResNet-18模型、在残差块后加注意力机制的CBAM-ResNet-18“Within Blocks”模型、传统分类模型支持向量机和随机森林相比,模型在卷收器声音信号分类任务中的混淆矩阵、准确率、精确率、召回率和 F_1 值等方面均表现良好。实验结果表明:所设计的模型对于基于声音信号的汽车安全带卷收器质量检测十分有效。

关键词:汽车安全带;声音信号;卷收器;质量检测;CBAM-ResNet

中图分类号:TP391.9 **文献标志码:**B **文章编号:**1671-5276(2024)01-0271-05

Reserch on Quality Detection Method for Automobile Safety Belt Retractor Based on Sound Signal

LIU Hongda, ZUO Dunwen, WANG Yong, JIN Mengmeng

(College of Mechanical and Electrical Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China)

Abstract:In order to improve the quality inspection efficiency of the automobile seat belt production site, an experimental platform is built according to the “QC/T987—2014 Automotive Seat Belt Retractor Performance Requirements and Test Methods” to collect the sound signals during the working process of the retractor qualified and defective products. Before the convolutional attention module (CBAM) is embedded into the residual network (ResNet-18) residual block, a CBAM-ResNet-18 “Before Blocks” model is designed to classify the collected retractor sound signals. Compared with the ResNet-18 model without the attention mechanism, the CBAM-ResNet-18 “Within Blocks” model with the attention mechanism after the residual block, the traditional classification model support vector machine and random forest, the designed model performs well in the aspects of confusion matrix, accuracy, precision, recall rate and F_1 value in the task of retractor sound signal classification, which is very effective for the quality detection of automobile safety belt retractors based on the sound signal.

Keywords: car seat belts; sound signal; retractor; quality inspection; CBAM-ResNet

0 引言

安全带卷收器在以往的质量检测过程中,通过总成检验工的手感来判断卷收器功能是否合格,检测效率低、成本高^[1]。

由于检测过程中卷收器会产生各种声音信息,这些声音信息可以通过非接触测量安全、快速并方便地采集与获取。近年来国内外学者在声音信号的分类研究方面展开了一些研究^[2]。

随着深度学习技术的发展,卷积神经网络(convolutional neural network, CNN)受到了广泛的关注,它在声音信号分类方面得到许多应用。TRAN等^[3]利用卷积神经网络结构从声音信号的

图像中提取丰富特征对钻机故障声音进行分类。郑思宇^[4]将采集故障噪声信号特征优化后输入到卷积神经网络模型中进行故障类型判断。

为了解决CNN网络过深导致的梯度弥散以及网络退化问题,HE等^[5]将残差神经网络(residual neural network, ResNet)引入到CNN中。SHETHWALA^[6]利用ResNet和肺音信号的Mel谱图对肺音进行分类。LOEY^[7]用ResNet对不同咳嗽声进行分类用于COVID-19领域。

对于声音信号特征差异不明显的分类,引入注意力机制能够取得更好的分类效果。目前结合残差网络与注意力机制对声音信号分类的研究还不够深入,但是在图像和视频领域已经取得了一

些成果。ZHAO 等^[8]将改进的注意力机制嵌入到 ResNet 中,用于区分植物病害之间的微小变化。ULLAH 等^[9]利用注意力机制 CBAM 和 ResNet 在视频识别中识别马来西亚手语(MSL)。

1 分类模型

1.1 Within Blocks

残差网络 ResNet-18 中的残差块有 4 个卷积层的深度。“Within Blocks”是指在网络架构的每个残差块第 2、第 4 卷积层的后面嵌入 CBAM 的方法。这种结构推断“Within Blocks”方法使得 CBAM 模块在这些残差块之间被应用 8 次。这种 CBAM 残差网络可以将中间特征映射细化到更能代表输入的重要信息。图 1 为 CBAM-ResNet-18 “Within Blocks”(以下简称 CBAM-ResNet-WB)中的单个残差块结构。

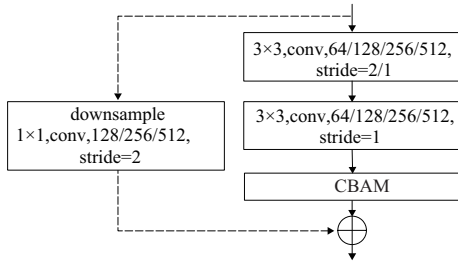


图 1 “Within Blocks”方法单个残差结构

卷积层的第 j 个单元的输出 a_j^l 计算如下:

$$a_j^l = f\left(\sum_{i \in M_j^l} a_i^{l-1} \times k_{i,j}^l + b_j^l\right) \quad (1)$$

式中: M_j^l 表示输入特征映射集; k 表示卷积核; $f(\cdot)$ 表示激活函数。

CBAM 是结合了通道注意力(channel attention)模块和空间注意力(spatial attention)模块的卷积注意力机制,每个子模块有自己的功能。具体结构如图 2 所示。

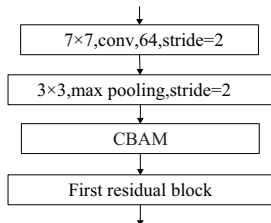


图 2 “Before Blocks”方法部分网络结构

通道注意力模块将输入的特征矩阵分别经过平均池化以及最大池化来筛选其空间信息,然后将结果输入到共享全连接层(shared MLP)中,再

将共享全连接层输出的特征进行相加合并操作,经过 sigmoid 函数激活,最终即可获得通道注意力模块的特征权重矩阵 $M_c(F)$ 。计算公式如下:

$$M_c(F) = \sigma[\text{MLP}(P_{\text{AvgPool}}(F)) + \text{MLP}(P_{\text{MaxPool}}(F))] \quad (2)$$

式中: F 为输入的特征矩阵; σ 为应用的 Sigmoid 激活函数; MLP 为多层感知机函数; P_{AvgPool} 和 P_{MaxPool} 分别表示平均池化和最大池化。

空间注意力模块首先将输入的特征矩阵做平均池化和最大池化,得到两个通道数为 1 个特征向量,然后将两个向量沿通道方向进行拼接操作,之后通过卷积操作映射为通道数为 1 的特征向量,最后经 Sigmoid 激活函数获得空间注意力特征权重矩阵 $M_s(F)$ 。计算公式如下:

$$M_s(F) = \sigma[f(P_{\text{AvgPool}}(F), P_{\text{MaxPool}}(F))] \quad (3)$$

式中 f 为卷积计算。

本文使用的 CBAM 注意力机制首先使用通道注意力对输入特征 F 进行加权得到 F_1 ,然后再针对加权后的特征 F_1 使用空间注意力进行再次加权得到最终的输出 F_2 ,其计算过程如下:

$$\begin{cases} F_1 = M_c(F) \odot F \\ F_2 = M_s(F_1) \odot F_1 \end{cases} \quad (4)$$

式中 $\odot$ 表示逐元素相乘。

残差块中包含下采样(downsample)模块,其输出被添加到块的输入中,计算如下:

$$H(x) = F' + x \quad (5)$$

式中: x 为结构的输入; F' 为 CBAM 的输出。

1.2 Before Blocks

与之前的方法不同,本文采用的“Before Blocks”方法将 CBAM 模块嵌入到第一个残差块之前。与“Within Blocks”方法相比,它具有更低的网络复杂性和更少的计算成本。图 2 展示了 CBAM 模块的具体位置。

2 验证实验和结果分析

2.1 数据集和数据预处理

本文基于实际工况环境下对信号进行采集,背景噪声复杂,噪声信号、安静环境下和实际工况中的卷收器信号图像如图 3 所示。

根据《QC/T987—2014 汽车安全带卷收器性能要求和试验方法》,本文搭建实验平台如图 4 所示,实验平台主要包括 1 个型材架、1 个滑轮、1 条导向带和 1 个重物。

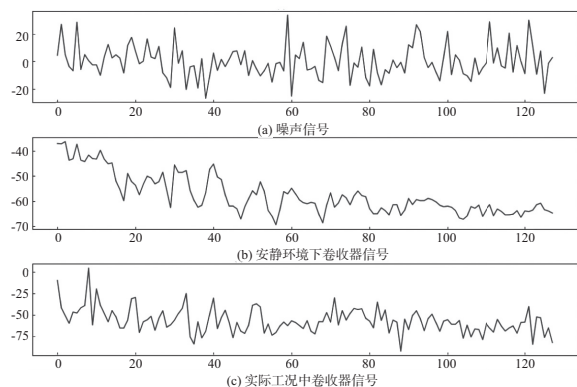


图3 噪声信号、安静环境下和实际工况中的卷收器信号

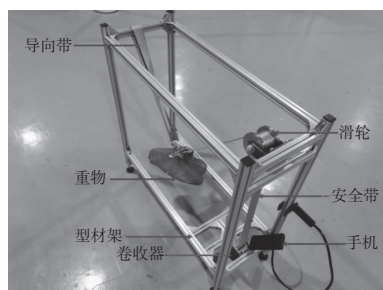


图4 采集卷收器声音数据实验平台

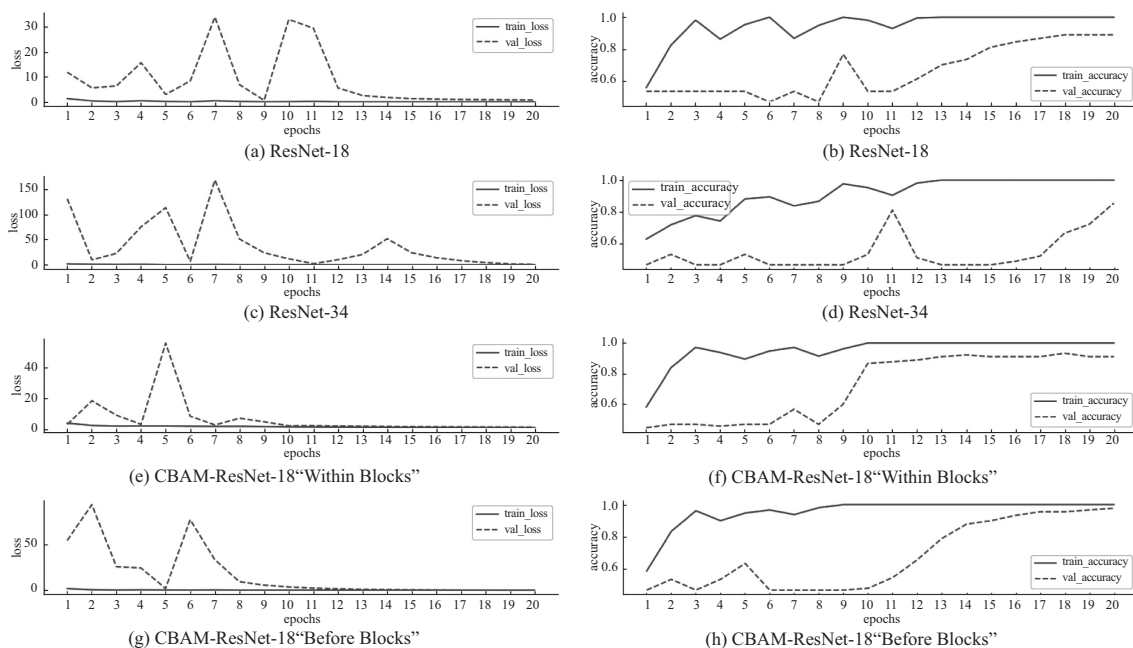


图5 ResNet-18、ResNet-34、CBAM-ResNet-WB、CBAM-ResNet-BB 模型的各轮次训练集与测试集的交叉熵和准确率对比图

ResNet-18 模型训练集的交叉熵从开始时的 1.34, 最后逐渐减小到 5×10^{-5} 。测试集交叉熵前 12 轮在 2.98~33.91 之间波动, 第 13 轮到 20 轮持续减小, 在 20 轮达到最小值 0.79。训练集准确率在前 3 轮从 55.71% 提升到 98.1% 后始终保持较高水平, 在第 13 轮达到 100% 后一直稳定。测试

该实验平台将卷收器固定于型材架上, 安全带绕过滑轮连接重物。重物为 5 kg, 在其重力的带动和导向带的牵引下, 安全带卷收器进行往复工作。录制过程中人工手持手机在卷收器旁, 采集其工作声音, 每次录制时间 3~5 s, 每一个单独保存为 WAV 文件, 其中包括 150 个合格品和 150 个次品, 共计 300 个音频文件。

对所有音频文件提取 Log-Mel Spectrogram 特征, 经过裁剪调整为 128×256 , 加载到分类模型中, 随机选择 70% 用于训练, 30% 用于测试。分类模型初始学习率为 0.000 5, 批量大小为 128。在训练过程中采用 RMSprop 优化器, 进行 20 轮训练。

2.2 验证结果分析

1) 训练与测试过程比较

本文采用 ResNet-18、ResNet-34、CBAM-ResNet-WB、CBAM-ResNet-BB 模型进行比较, 比较结果如图 5 所示。

集准确率前 10 轮波动较大, 第 9 轮提升到 76.67%, 第 10 轮降到 53.33%, 第 11 轮到 20 轮持续升高, 最终在 18 轮达到最大值 88.89%。模型训练准确率和测试准确率最终相差 11.11%, 存在过拟合问题。

ResNet-34 模型训练集的交叉熵从开始时的

1.93,最后逐渐减小到 6×10^{-6} 。测试集交叉熵前 12 轮在 2.13~168.35 之间波动,第 13 轮到 16 轮从 10.17 持续升高到 51.73,之后逐渐减小,在 20 轮达到最小值 0.67。训练集准确率呈持续升高趋势在第 13 轮达到 100%后一直稳定。测试集准确率前 13 轮波动较大,第 11 轮提升到 81.11%,第 13 轮降到 46.67%,第 14 轮到 20 轮持续升高,最终在 20 轮达到最大值 85.56%。模型训练准确率和测试准确率最终相差 14.44%,模型深度增加后导致过拟合更加严重。

CBAM-ResNet-WB 模型训练集的交叉熵从开始时的 3.92 最后逐渐减小到 0.93。测试集交叉熵前 8 轮在 2.77~55.89 之间波动,第 9 轮到 20 轮持续减小,在 20 轮达到最小值 1.32。训练集准确率在前 3 轮从 58.1%提升到 97.14%后始终保持较高水平,在第 10 轮达到 100%后一直稳定。测试集准确率前 8 轮处于较低水平,第 9 轮开始从 60%提升到第 14 轮的 92.22%,第 18 轮提升到 93.33%。模型训练准确率和测试准确率最终相差 6.67%,存在一定的过拟合问题。加入注意力机制的模型使特征更加符合分类要求,过拟合问题减小。

CBAM-ResNet-BB 模型训练集的交叉熵从开始时的 1.59,最后逐渐减小到 0.06。测试集交叉熵前 8 轮在 1.78~93.29 之间波动,第 9 轮到 20 轮持续减小,在 20 轮达到最小值 0.12。训练集准确率在前 3 轮从 58.57%提升到 96.19%后始终保持较高水平,在第 9 轮达到 100%后一直稳定。测试集准确率前 11 轮处于较低水平,第 12 轮开始从 65.56%提升到第 20 轮的 97.78%。模型训练准确率和测试准确率最终相差 2.22%,说明还是存在一些过拟合问题。本文模型过拟合问题更小,准确率更高,收敛速度仅次于参数更多的 CBAM-ResNet-WB 模型,但计算量相对较小。说明该模型能够加快收敛速度且获得更好的分类效果。

2) 验证分类结果和混淆矩阵

图 6 为 ResNet-18 模型、ResNet-34 模型、

CBAM-ResNet-WB 模型和 CBAM-ResNet-BB 模型 的分类结果混淆矩阵,其中“0”代表合格品,“1”代表次品。左上角和右下角单元格显示每个类别分类正确的数量,其他单元格显示分类错误的数量。

图 6 可知,ResNet-18 模型分类结果较差,其中有 10 个合格品被检测成为次品;ResNet-34 模型分类结果差,其中有 13 个次品被检测成为合格品;CBAM-ResNet-WB 模型分类效果稍好一些,其中有 6 个合格品被检测成为次品,2 个次品被检测成为合格品;CBAM-ResNet-BB 模型分类效果最好,只有 2 个正品被检测为次品。

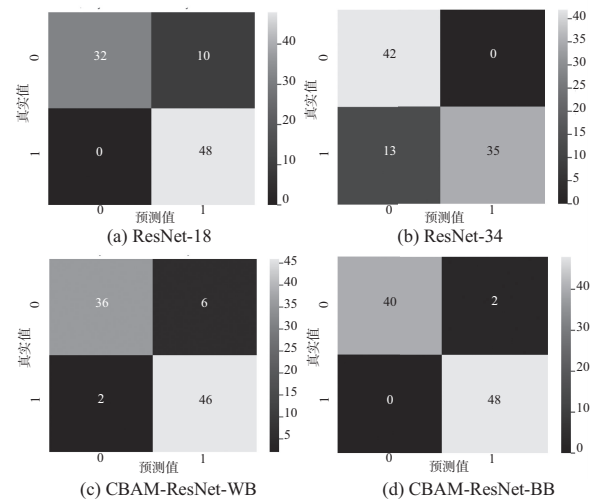


图 6 模型分类混淆矩阵

表 1 列出了 ResNet-18 模型、ResNet-34 模型、CBAM-ResNet-WB 模型和 CBAM-ResNet-BB 模型分类结果的精确率、召回率和 F_1 值。ResNet-18 和 CBAM-ResNet-BB 模型合格品的精确率都为 1,所有被检测为合格品的卷收器均为合格品;同时,两个模型次品召回率都为 1,能够检测出所有次品,确保出厂的卷收器质量,保障安全性;CBAM-ResNet-WB 模型精确度和召回率相对较为平均,合格品和次品的分类均有错误。CBAM-ResNet-BB 模型的 F_1 值最高,合格品和次品 F_1 值均为 0.98,说明该模型分类效果最好。

表 1 ResNet-18、ResNet-34、CBAM-ResNet-WB 和 CBAM-ResNet-BB 模型的性能对比

分类	精确率				召回率				F_1 值			
	ResNet-18	ResNet-34	Within Blocks	Before Blocks	ResNet-18	ResNet-34	Within Blocks	Before Blocks	ResNet-18	ResNet-34	Within Blocks	Before Blocks
合格品	1.00	0.76	0.95	1.00	0.76	1.00	0.86	0.95	0.86	0.86	0.90	0.98
次品	0.83	1.00	0.88	0.96	1.00	0.73	0.96	1.00	0.71	0.84	0.92	0.98

3.3 与其他方法的对比

表2列出了CBAM-ResNet-BB模型与传统分类模型支持向量机、随机森林分类结果的准确率、精确率、召回率和 F_1 值。实际工况下采集到的信号噪声较大,传统分类效果较好的支持向量

机和随机森林模型在本次分类任务中结果较差,CBAM-ResNet-BB模型分类结果的准确率、精确率、召回率和 F_1 值均优于支持向量机和随机森林,模型具有实际应用价值。

表2 支持向量机、随机森林和CBAM-ResNet-BB模型的分类性能对比

分类	准确率			精确率			召回率			F_1 值		
	支持 向量机	随机 森林	Before Blocks	支持 向量机	随机 森林	Before Blocks	支持 向量机	随机 森林	Before Blocks	支持 向量机	随机 森林	Before Blocks
合格品	0.79	0.77	0.98	0.74	0.79	1.00	0.86	0.80	0.95	0.80	0.80	0.98
次品				0.85	0.74	0.96	0.72	0.72	1.00	0.78	0.73	0.98

注:表中随机森林最大特征数为30,最大深度为35;支持向量机核函数选择为“poly”,惩罚项大小为0.6。

3 结语

本文提出一种CBAM-ResNet-BB模型,将注意力机制融入残差网络,能够基于声音信号对汽车安全带卷收器的质量进行分类检测,分类准确率达98%。

ResNet-18模型的检测准确率为88.89%,ResNet-34模型的检测准确率为85.56%。对于本文来说,ResNet-34模型网络深度相对较深,但是过拟合问题更加严重,导致准确率更低,因此选用ResNet-18模型进行改进。

CBAM-ResNet-BB模型的训练集和测试集最终准确率相差2.22%,对比ResNet-18模型的11.11%和ResNet-18 Within Blocks模型的6.67%,过拟合问题明显改善,准确率提高8.89%。

CBAM-ResNet-BB模型分类准确率为98%,传统分类模型支持向量机的分类准确率为79%,随机森林的分类准确率为77%,进一步验证了该模型的优势,能够完成噪声较大的分类任务。

参考文献:

- [1] 徐晓敏. 安全带卷收器总成统计过程控制[J]. 上海电机学院学报, 2015, 18(5): 288-292.
- [2] 李常有,徐敏强,郭算. 利用声信号对滚动轴承进行故障诊断的研究[J]. 应用声学, 2008, 27(4): 315-320.
- [3] TRAN T, LUNDGREN J. Drift fault diagnosis based on

the scalogram and mel spectrogram of sound signals using artificial intelligence[J]. IEEE Access, 2020, 8: 203655-203666.

- [4] 郑思宇. 基于声音信号的机械设备故障诊断[J]. 内燃机与配件, 2021(13): 129-132.
- [5] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition [EB/OL]. (2015-12-10) [2022-04-27]. <https://arxiv.org/abs/1512.03385>.
- [6] SHETHWALA R, PATHAR S, PATEL T, et al. Transfer learning aided classification of lung sounds—wheezes and crackles [C] // 2021 5th International Conference on Computing Methodologies and Communication (ICCMC). Erode, India: IEEE, 2021: 1260-1266.
- [7] LOEY M, MIRJALILI S. COVID-19 cough sound symptoms classification from scalogram image representation using deep learning models [J]. Computers in Biology and Medicine, 2021, 139: 105020.
- [8] ZHAO Y, CHEN JG, XU X, et al. SEV-Net: residual network embedded with attention mechanism for plant disease severity detection [J]. Concurrency and Computation: Practice and Experience, 2021, 33(10): e6161.
- [9] ULLAH KHAN R, SHENG WONG W, ULLAH I, et al. Evaluating the efficiency of CBAM-resnet using malaysian sign language [J]. Computers, Materials & Continua, 2022, 71(2): 2755-2772.

收稿日期: 2022-08-01