

基于深度学习的车辆振动信息识别研究

罗智元^{1,2} 刘小勇²
LUO Zhiyuan LIU Xiaoyong

摘要

车辆振动信息识别对于维护交通安全和道路状况具有重要作用,可以节省大量人力物力。机器视觉、红外线成像等基于图像的方法虽然可以达到较高的准确率,但是容易受到环境比如照明条件、天气等因素干扰,相比之下,车辆的振动信号可以提供相对较稳定的车辆信息,传感器装配成本更低、维护更容易。近年来,从振动信号识别车辆信息的方法集中在传感器改进、信号处理及识别等方面。在使用神经网络的方法中,一类方法将计算出的特征输入全连接网络或一维卷积神经网络进行特征识别,这类方法需要专家知识对特征进行设计且计算方法复杂,另一类使用原始数据作为输入,这类方法则面临计算资源消耗大、训练慢等问题。针对上述问题,提出了基于深度学习的车辆振动信息识别方法,分别构建了1dCNN和2dCNN轴型信息识别网络模型,并对模型进行了鲁棒性测试、模型大小评估及推理速度测试。结果表明,1dCNN和2dCNN两个模型在测试集上分别达到了0.986和0.973的识别准确率,1dCNN比2dCNN有更好的鲁棒性、更快的推理速度。对于采样率500 Hz、长度为632的输入信号,1dCNN在CPU上的平均推理速度可达0.006 9 s/样本,2dCNN为0.18 s/样本。

关键词

车辆信息识别; 振动信号; 深度学习; 小波变换; 卷积神经网络

doi: 10.3969/j.issn.1672-9528.2024.02.021

0 引言

车辆信息识别是构建智能交通系统(ITS)的重要环节,在维护交通规则、路面行车秩序中发挥重要作用,可为交通管理部门节省大量人力物力。轴型是车辆的一个重要的信息,由轴型可以推测车辆的种类,比如货车中的平板式货车、罐式货车、箱式货车的轴型分别是六轴、三轴和二轴。不同车辆有不同的限速、负载限值和高度限值等,自动识别车辆的轴型,进而推断车辆种类可以辅助判断通过的车辆是否超过规定的限值。

为了识别车辆信息,常对车辆的图像、声音、振动等信号进行处理和识别,众多研究采用声音、图像、红外传感器等进行信息识别。然而,声音传感器可能会因为车辆的移动而受到多普勒效应的不利影响,图像和红外传感器容易受到照明条件、温度、湿度变化的影响^[1]。振动信号是生活中一种常见的信号,表现为某种测量物理量随着时间的波动,车辆行驶过的路面会产生形变,进而产生振动,可以在路面或者桥梁下面安装加速度计、光纤传感器、应变计等进行检测,振动信号可以反映车辆的丰富信息,相比于声音、图像和红外信号,振动信号不容易受到环境影响。

用来衡量振动幅度所测量的物理量有加速度、应力等。

Ma等人^[2]使用安装在桥梁下的应力传感器获取车辆振动信号,可以不打断车流的同时进行车辆信息识别^[3]。使用加速度计将路面振动转换为车轴的位置,磁力计监测并估计车辆到达和离开的速度从而推断车辆振动存在的时间窗口,再通过滤波分离出车辆经过时产生的峰值,从而推出车辆的轴型,在通过的226辆三轴以上的车辆中,准确分类了223辆。Khong等人^[4]开发了一种新型的低成本的压力传感器,将聚氨酯材料夹在铝电极中间,压力通过引起的两个电极间的距离变换导致电容的变化,从而将压力转换为电信号,车辆负荷可以引起该传感器的明显读数变化。Al-Tarawneh等人^[5]开发了一种三维玻璃纤维增强聚合物封装的纤维布拉格光栅传感器(3D GFRP-FBG),通过该传感器的中心波长的变化反映车辆经过产生的应力变换。Li等人^[6]使用一种先进的圆形电感线圈传感器获取车辆的振动数据,相比传统的电感线圈捕捉到了更高分辨率的振动波形。

深度学习算法使用计算出的特征或者原始信号作为输入,网络结构采用DNN、CNN等。Yu等人^[7]使用连续小波变换提取桥梁所受应力的信号的特征,使用这些特征辨别轴数和车速。Ying等人提出了一种基于决策树的算法,通过原始振动信号计算特征,利用这些特征构建决策树模型,在真实数据集中最高达到100%的准确率。Al-Tarawneh等人首先通过传感器信号计算车速、轴距等特征,再构建支持向量机(SVM)模型对提取到的特征进行分类,实现了98.5%的分

1. 新加坡南洋理工大学电子工程学院 新加坡 637616
2. 西安交通大学电信学部自动化学院 陕西西安 710049

类准确率。Yan 等人^[8]使用主元分析(PCA)从原始信号中提取出主要特征,将特征输入一个两层的神经网络,训练神经网络使之可以识别五类车辆,准确率可达 90%。Li 等人构建了一个六层的神经网络,每层有 256 个神经元,并用集成学习的方法进行训练,解决数据集不平衡的问题,达到了 87% 的平均准确率。Ma 等人构建了一个残差卷积神经网络,准确率可达 97.6%。

提升传感器性能是早期的研究中的重点,传感器采集数据和数据预处理承担了大部分的功能,这种方法的缺陷是成本高,需要采购、改装传感器,对改进的传感器进行多次测试,同时还需投入大量人力物力进行传感器维护。近年来,研究方向转向基于大数据和神经网络的车辆信息识别方法。该方法可以从采集到的大量数据中学习特征,减少对专家知识的依赖。深度学习在图像分类、目标检测、自然语言处理等领域展现强大的潜力。其强大的数据学习能力、鲁棒性、高效性在很多领域已经超过专家建模的方法^[9-10]。

1 振动信号的数据采集与标注

为了构建车辆振动信息识别模型,需要采集原始的振动信号。传感器的安装位置应该不影响车辆行驶或者不打断交通,振动信号采集需要噪声小同时能够捕捉足够信息的传感器,本文采用安装在路面下的称重传感器,信号采集现场示意图如图 1 所示,每个称重块的四角安装四个称重传感器。称重传感器的主要结构如图 2 所示,其主要功能结构是应变计。应变计是一种测量力的传感器,通过压敏材料,将地面的形变转换为电信号的变化。车辆行驶方向一共布置了六个测量模块,可以测得 24 路车辆振动信号,这些不同传感器测量的振动信号之间有一定的相关性,仅用部分测得的信号对系统的识别效果没有明显影响。

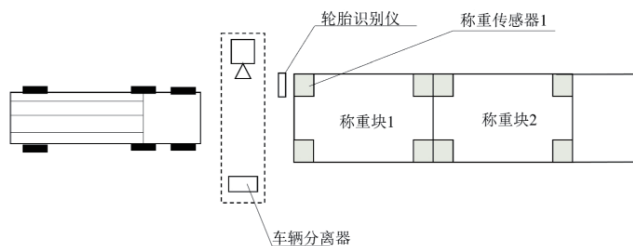


图1 振动信号采集现场示意图

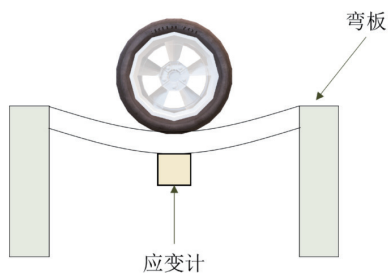


图2 称重传感器

振动信号信息识别是一种多分类问题,将输入的振动信号分类为对应车辆的轴型。多分类任务可以用监督学习实现,监督学习需要大量的标签数据。标注是记录数据类别的过程。将数据的地址和对应的类别记录在文档里,训练神经网络时,每读取一次数据,程序将根据文档里的地址找到数据和标签组成数据对。

首先对采集到的振动信号用 Excel 软件整理,将畸变严重、传感器故障和存在多个车辆通过的文件剔除,接着对信号文件进行标注。其次辅助信号记录了车辆通过的大致时间,通过辅助信号,配合工作日志的记录可以确定该时间内通过的车辆类型。然后,如图 3 所示,辅助信号处于低电平的时候表示有车辆经过,将辅助信号和传感器信号绘制在一起的时候可以确定该车辆的轴型。最后,通过工作日志检验人工的判断。

确定好该文件数据对应的车辆轴型后,就可以将文件的地址、标签等信息记录在文件中。为了从信号中提取特征,识别车辆轴型信息,以残差块和卷积层为结构基础构建了两种网络,1dCNN 和 2dCNN,基本结构分别是残差一维卷积和残差二维卷积,分别用来提取一维振动信号输入和二维频谱图输入的特征。

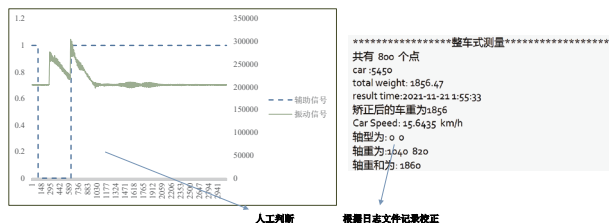


图3 数据标注过程

2 1dCNN 车辆信息识别网络模型

一维车辆信息识别网络模型需要实现计算输入的一维序列进行分析,输出这个序列的标签。模型的输入是以 500 Hz 采样好的振动序列 $x=[x_1, x_2, \dots, x_n]$, 输出标签 y , $y \in \{0, 1, 2, 3, 4\}$, 标签数 0、1、2、3、4 分别对应二轴、三轴、四轴、五轴、六轴。这个任务可以通过一个多分类网络解决。卷积神经网络是一种强大的网络结构,因此考虑用卷积神经网络设计这个多分类任务。

深度卷积网络导致了一系列在图像分类领域的突破,但是更深的网络也会导致梯度消失或者梯度爆炸等问题^[11]。本文采用残差块方法,残差块是一种分支结构,将输入层和输出层相加,效果是避免梯度消失和加快网络训练速度,残差块的基本结构如图 4 所示,设卷积层的映射为 $F(x)$, 残差学习的原理是基于学习 $F(x)-x$ 。

残差块的卷积通道上有权重层、BN 层、ReLU 层,其中权重层对输入进行卷积计算,BN 层用于对输入进行批标准

化 (batch normalization), ReLU 层对输入进行激活, 将输出映射到非线性空间。这三个层的顺序会影响模型效果。He 等人^[12]在基础残差块的结构基础上调整激活层、BN 层和权重层的顺序, 比较模型在 CIFAR-10/100 测试集上的分类损失, 结果显示, 加入预激活层可以将损失降得更低, 原残差块和使用预激活层残差块的结构如图 5 所示。

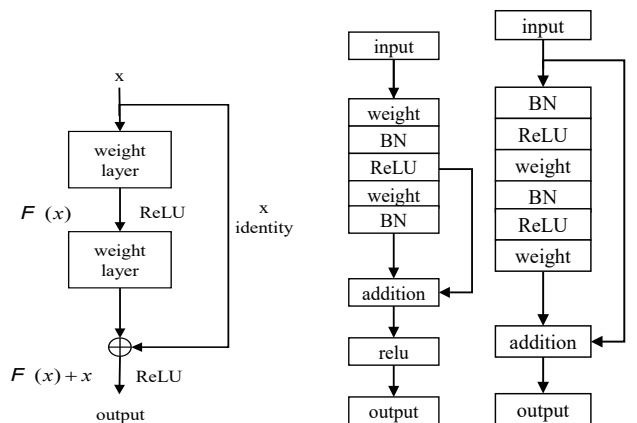


图 4 残差块结构

(a) 原残差块 (b) 预激活残差块
图 5 原残差块和预激活残差块

本文构建的 1dCNN 车辆信息识别网络先将信号依次输入衔接卷积层, 级联的五个残差块, 挺好输入全连接层, 输出经过 Softmax 函数转为各类置信度, 输入通道由 1 扩张到 512, 考虑到在网络的浅层进行池化可能会丢失重要信息, 因此删去池化层, 如图 6 所示。

残差块的输出需要有相同的维度才能够相加, 由于主通道提高了输入的通道数, 减小了输入的长度, 所以需要采取措施保持维度一致。常见的措施有下采样和在侧通道上使用同样的卷积层。车辆振动信号输入时长不是固定的, 需要通过补零统一输入长度。并且, 由于振动信号的幅值很大, 而且和车辆重量成正比, 实验中对所有输入进行了归一化处理, 振动的波谷可能是负值。如果在分支使用最大池化或者平均池化下采样可能会导致一些负值的信号点被过滤掉, 而这些负值的点在车辆振动信号中是有重要含义的。当选择不同的参数组合时产生的信号长度也是不同的, 如果在分支使用池化进行下采样, 每改动一次信号预处理参数都要调整相应的池化层参数, 在编程中带来很大的不方便。因此, 选择侧通道带卷积的残差块, 由于侧通道的卷积和主通道的卷积形状一致, 所以不用特别调整参数来保持输出特征大小一致。侧通道卷积残差块如图 7 所示, 其中, weight3 和 weight4 是独立学习的参数、激活层、BN 层省略。

综上所述, 本文构建的 1dCNN 车辆信息网络结构如图 8 所示。首先一个独立的卷积层处理输入信号, 该卷积层含有

8 个 1×3 的卷积核, 将输入通道由 1 扩张到 8。然后通过 5 个残差块, 图中残差块下面的 $8 \rightarrow 32$ 表示通道的变化, 每个残差块含有两层卷积, 逐渐将通道扩增到 512 个。最后, 将张量展开, 输入一个全连接层, 通过 SoftMax 函数全连接网络的输出映射到 0-1, 进行分类, 1dCNN 车辆信息识别网络的参数细节如表 1 所示。

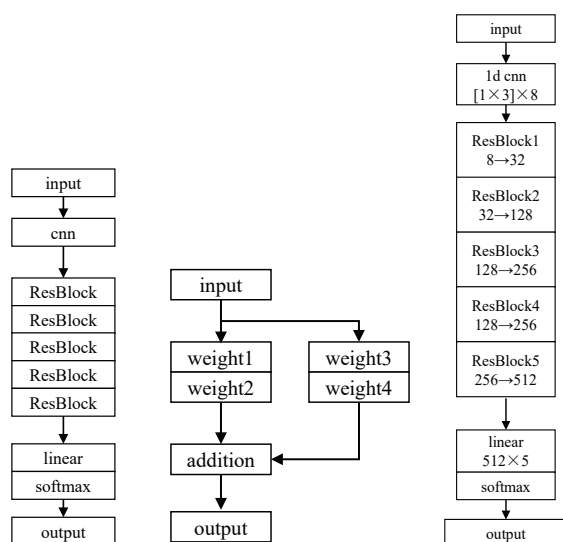


图 6 1dCNN 车辆信息识别网络框架 图 7 侧通道有卷积层的残差块 图 8 1dCNN 车辆信息识别网络

表 1 1dCNN 车辆信息识别网络参数

网络层	输入大小	卷积核大小	步长	填充	输出大小
conv1d	$16 \times 1 \times 484$	$[1 \times 3] * 8$	1	0	$16 \times 8 \times 482$
Residual Block1	$16 \times 8 \times 482$	$[1 \times 3] * 16$ $[1 \times 3] * 32$	1	0	$16 \times 32 \times 478$
Residual Block2	$16 \times 32 \times 478$	$[1 \times 3] * 64$ $[1 \times 3] * 128$	1	0	$16 \times 128 \times 237$
Residual Block3	$16 \times 128 \times 237$	$[1 \times 3] * 256$ $[1 \times 3] * 256$	2	0	$16 \times 256 \times 58$
Residual Block4	$16 \times 256 \times 58$	$[1 \times 3] * 256$ $[1 \times 3] * 256$	2	0	$16 \times 256 \times 13$
Residual Block5	$16 \times 256 \times 13$	$[1 \times 3] * 512$ $[1 \times 3] * 512$	2	0	$16 \times 512 \times 2$
Average Pool	$16 \times 512 \times 2$	—	—	—	$16 \times 512 \times 1$
Flatten	$16 \times 512 \times 1$	—	—	—	16×512
Linear	16×512	—	—	—	16×5

3 2dCNN 车辆信息识别网络模型

近年来, 对时间信号的频谱提取特征实现信号分类的研究逐渐增多, 频谱图不仅含有信号的频域信息还含有这些频率随着时间变化的信息, 非常适合振动信号这一种非平稳信号的信息提取。获取信号的频谱图一般有两种手段, 小波变换 (WT) 和短时傅里叶变换 (STFT)^[13], 在异常 ECG 信号识别的研究领域, 将信号转换到频谱图再用 CNN 处理的研究很多。Wang 等人^[14]使用连续小波变换 (CWT) 将 ECG

心跳信号分割后转换到时频域,接着输入一个六层的卷积神经网络训练分类,在测试集上实现了 98.7% 的准确率。

Huang 等人^[15]使用短时傅里叶变换得到 ECG 信号的频谱图,再输入一个三层的二维卷积神经网络中,达到了平均 99.00% 的准确率。Roberts 等人^[16]计算了 PCG 信号的梅尔频谱系数和 MFCC 系数,输入一个 5 层的前馈神经网络进行分类,在测试集上实现了 97.10% 的准确率。梅尔频谱图是利用短时傅里叶变换计算后再将频率转换为梅尔坐标轴上得到的。短时傅里叶变换是将信号分割为多个有重叠的窗,在每个时间窗内进行 FFT。梅尔频谱图适合处理声音类的信号,因为人类对频率的感知不是线性的,人类更加擅长感知低频的差异,而不是高频的差异,梅尔音阶将频率进行转换,使得同等距离的频率在人耳听来有同样的差别^[17]。

考虑到车的振动信号也可以认为是一种情景分类任务,人耳可以感知车辆的振动信号,并进行粗略分类,车辆的振动信号是一种低频和低频信号的结合,本文采用梅尔频谱图与 STFT 频谱图作为变换的方法对车辆振动信号的识别进行了研究。将车辆振动信号转换到梅尔频谱图,五类轴型车辆振动信号产生的梅尔频谱如图 9 所示。

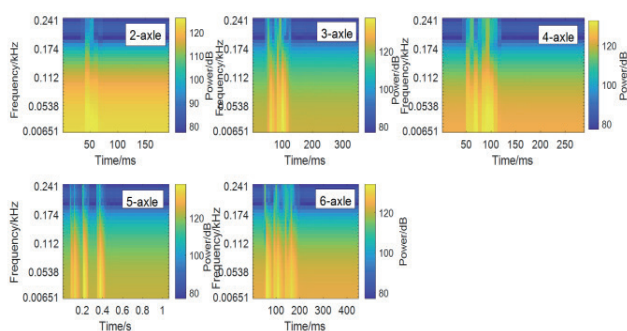


图 9 五类轴型车辆振动信号的梅尔频谱图

从图 9 可见,每个轴型车辆的振动信号的梅尔频谱图有较为明显差别,因此使用振动信号的梅尔频谱图进行车辆分类是可行的。设计 2dCNN 网络时采用了和 1dCNN 轴数识别网络基本一致的框架,但是删除了衔接卷积层,由于实验中发现加入该层的时候,验证集上的损失震荡幅度大。该网络的结构是五个残差块级联,最后通过全连接层输出。

由于二维的频谱图的数据量较大,并且二维卷积的参数比一维卷积的参数更多,训练需要耗费很长时间,且得到的模型参数文件大。为了提高模型训练的速度和减少模型大小,采用深度可分离卷积代替所有卷积层。深度可分离卷积是 Mobile Net 的核心结构^[18],这种结构将普通的卷积拆分为两步操作深度卷积和点卷积。其结构如图 10 所示,图 10 (a) 是深度卷积的过程,在这个过程中,通道数不变,输入通道数等于输出通道数,图 10 (b) 是点卷积的过程,将输入和 m 个 $n \times 1 \times 1$ 的卷积核做卷积,得到的输出通道数由 n 变为 m ,但是高和宽不变。

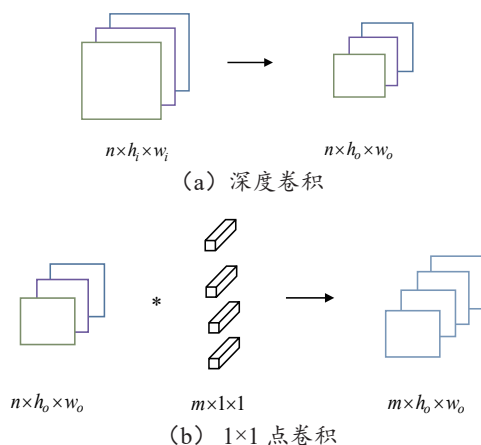


图 10 深度可分离卷积

假设一个普通的卷积过程,输入为 $C_i \times H \times W$,卷积核大小为 $C_i \times K_1 \times K_1$,输出通道大小为 C_o ,则需要 C_o 个 $C_i \times K_1 \times K_1$ 大小的卷积核,假设步长为 1,则总共的计算次数是 $C_i \times H \times W \times K_1 \times K_1 \times C_o$ 。如果采用深度可分离卷积,则第一步进行深度卷积,仅使用一个 $C_i \times K_1 \times K_1$ 的卷积核计算,对输入的每个通道分别卷积,总的计算操作数是 $C_i \times H \times W \times K_1 \times K_1$;接着,使用点卷积扩张通道,用 C_o 个大小为 $C_i \times 1 \times 1$ 的卷积核和输入进行计算,则计算操作数为 $C_i \times H \times W \times C_o$ 。深度可分离卷积总的操作数是 $C_i \times H \times W \times K_1 \times K_1 + C_i \times H \times W \times C_o = (C_i \times H \times W) \times (K_1 \times K_1 + C_o)$,相较于普通卷积的 $C_i \times H \times W \times K_1 \times K_1 \times C_o$,有明显的数量级上的减少。

2dCNN 车辆信息识别网络用到的基于深度可分离卷积的残差块如图 11 所示。2dCNN 车辆信息识别网络由 5 个图 11 所示的残差块组成,其结构示意图如图 12 所示。输入通过 5 个残差块将输入通道由 1 扩张到 512,再通过一个全连接层和 Softmax 输出各类置信度。该网络的详细参数如表 2 所示,假设输入是 batch size=1,维度为 128×127 的频谱图。

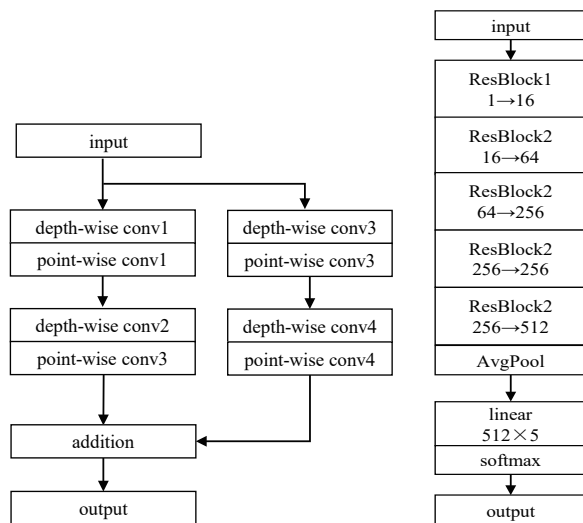


图 11 深度可分离卷积残差块

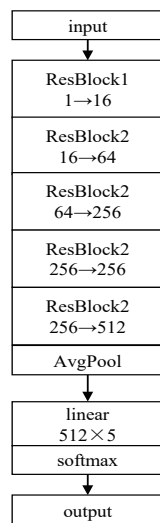


图 12 2dCNN 车辆信息识别网络

表 2 2dCNN 车辆信息识别网络参数

网络层	输入大小	卷积核大小	步长	填充	输出大小
Residual Block1	1×128×147	[3×3]*8 [3×3]*16	(1,1)	0	16×124×143
Residual Block2	16×124×143	[3×3]*32 [3×3]*64	(1,1) (1,1)	0	64×120×139
Residual Block3	64×120×139	[3×3]*128 [3×3]*256	(1,2) (1,2)	0	256×116×34
Residual Block4	256×116×34	[3×3]*256 [3×3]*256	(2,2) (2,2)	0	256×28×7
Residual Block5	256×28×7	[3×3]*256 [3×3]*512	(2,2) (2,2)	0	512×6×1
Average Pool	512×6×1	-	-	-	512×1×1
Flatten	512×1×1	-	-	-	1×512
Linear	1×512	-	-	-	1×5

4 训练过程、结果和分析

根据前述车辆真实行走实验，共采集了 1229 个数据。其中 1079 个数据被用于训练集和验证集，150 个数据为测试集。五类数据的分布如表 3 所示。以 8:2 的比例分割训练集和验证集，测试集是独立随机选择的 150 个数据。

表 3 数据集分布

标签	车辆类别	训练集	验证集	测试集	总计
0	2 轴	336	86	27	449
1	3 轴	146	24	36	206
2	4 轴	163	45	47	255
3	5 轴	74	12	11	97
4	6 轴	144	49	29	222

表 3 中，二轴车占总数据集比例 36.5%，三轴车占比 16.8%，四轴车占比 20.7%，五轴车占比 7.89%，六轴车占比 18.1%。实验全部流程使用 PyTorch 平台进行实现，整体实验框架如图 13 所示。

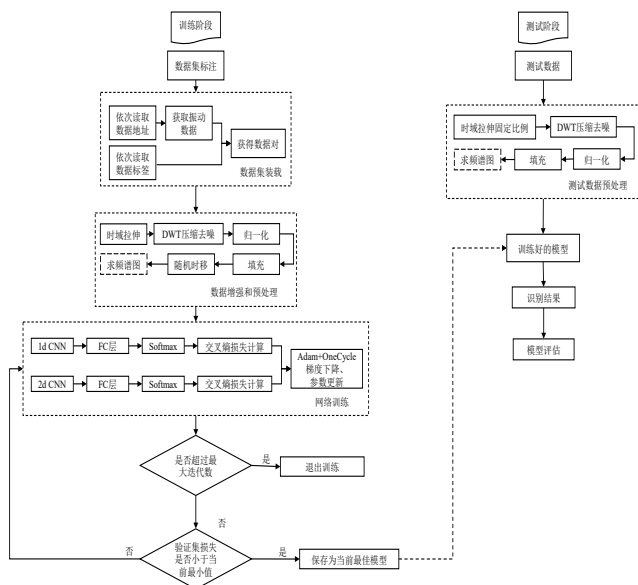


图 13 网络训练总体框架

实验使用了交叉熵损失函数，其公式为：

$$L_{CE} = -\sum_{i=1}^C t_i \log(p_i) \quad (1)$$

式中： C 为类别的总数，实验中 $C=5$ ； t_i 为二进制真实值，值为 1 代表是某类别，0 表示非某类别； p_i 为 Softmax 函数计算出的该类别的置信度。

为了减轻过拟合，加入 L2 正则化。L2 正则化是将权重的 L2 范数求和再乘以一个正则化系数作为一个惩罚量加在损失函数里面。网络的优化目标从减少网络预测的交叉熵损失变为不仅要减小网络预测的交叉熵损失还要使得网络的权重尽量小。加入 L2 正则化可以有效减轻过拟合，最终，网络的优化目标公式为：

$$J = \min_{\theta} (L_{CE} + \lambda \|w\|_2) \quad (2)$$

式中： θ 所有的网络参数； λ 正则化系数，取 0 相当于没有使用 L2 正则化，使用 L2 正则化可以有效避免过拟合。

神经网络的训练过程本质上是优化问题，是找到使得目标函数值最小或最大的一组参数的过程。而深度学习网络中最常用的优化算法是梯度下降，adam 是一种目标函数的优化算法，描述了梯度下降过程中参数更新的方式^[19]。在实验中，对于 1dCNN 模型，设置学习率为 0.001，L2 正则化系数为 0.0001；对于 2dCNN 模型，设置训练的学习率为 0.001，L2 正则化系数为 0.001，这是因为在训练过程中发现 2dCNN 模型更容易过拟合，所以设置了更大的 L2 正则化系数。实验中使用了 OneCycleLR 调度器调度学习率的变化，该调度器可以在训练过程中将学习率以某种规律慢慢增大到最大学习率，该算法最先在一篇关于超级收敛的论文^[20]中提出，这种方法从极小的学习率开始，逐渐增加学习率直到达到最大值。1dCNN 模型在训练集和验证集上的损失和准确率变化如图 14 和图 15 所示。2dCNN 模型在训练集和验证集上的损失曲线和准确率变化如图 16 和图 17 所示。

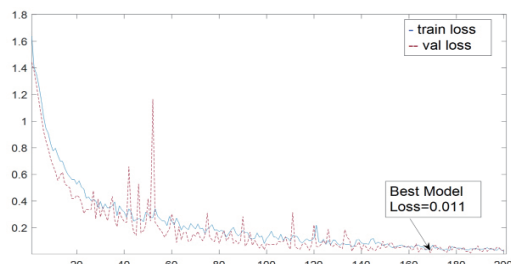


图 14 1dCNN 损失曲线，185 epoch 左右在验证集上有最小损失

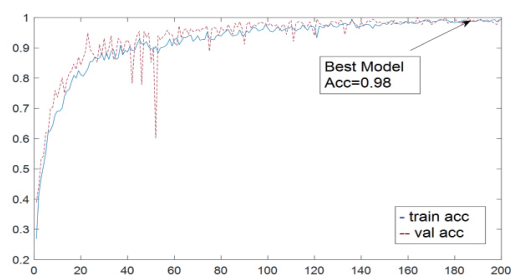


图 15 1dCNN 准确率变化曲线

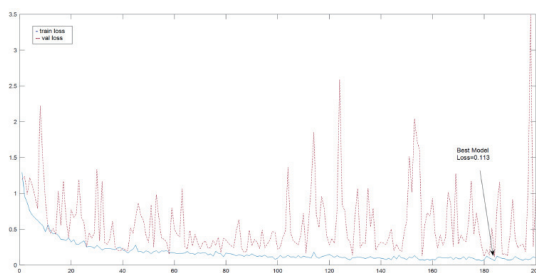


图 16 2dCNN 损失曲线, 185 epoch 左右在验证集上有最小损失

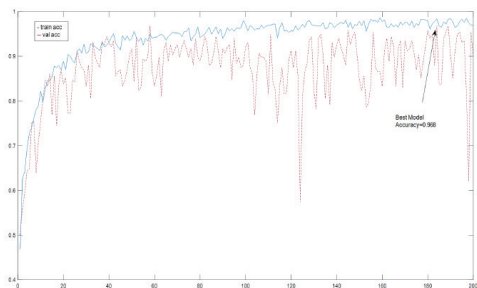


图 17 2dCNN 准确率变化曲线

从图 14 可以看出, 1dCNN 模型在 100 个 epoch 内很快收敛, 损失由初始的 1.65 左右降落到 0.2 左右, 而验证集上的损失在训练集附近震荡, 在 55 epoch 左右存在一个明显的尖端, 这可能是由于数据集中存在一个异常数据导致的。从图 16 可以看出, 2dCNN 模型在 100 个 epoch 内很快收敛。比较图 14 和图 16, 2dCNN 有过拟合倾向, 虽然训练集上的损失很平滑地随着迭代次数增加而减少, 但是验证集上的损失震荡幅度很大, 而 1dCNN 的验证集损失震荡幅度小, 基本可以跟随训练集的损失。由于上述的过拟合问题, 训练过程采取提前终止 (early stopping) 措施以避免过拟合。如图 14 所示, 1dCNN 模型在第 185 epoch 左右达到了验证集的最小损失, 损失为 0.011。如图 16 所示, 2dCNN 模型也是在第 185 epoch 左右达到了验证集上的最小损失, 损失为 0.113, 保存这两个时刻的模型权重文件作为最佳模型权重, 用作后续测试。

采用每个类别识别的精度、召回率和 F_1 分数以及全部类别的准确率四个指标来衡量模型的效果。准确率的计算公式为:

$$accuracy = \frac{T_P + T_N}{T_P + T_N + F_P + F_N} \quad (3)$$

式中: T_P 为真实的阳性, 即预测出属于某类的样本确实属于该类; T_N 为真实的阴性, 即预测出不属于某类的样本确实不属于该类; F_P 为虚假的阳性, 即预测出属于某类的样本实际上不属于该类; F_N 为虚假的阴性, 即预测出不属于某类的样本实际上属于该类。

精度指在预测出属于某类的样本中, 有多大比例是确实属于该类样本的, 其计算公式为:

$$precision = \frac{T_P}{T_P + F_P} \quad (4)$$

召回率指的是在所有真实属于某类的样本中, 有多大比例的样本可以被识别出来, 其计算公式为:

$$recall = \frac{T_P}{T_P + F_N} \quad (5)$$

F_1 分数同时考虑了精度和召回率, 考虑到假阳性和假阴性对模型性能的影响比真阴性更重要, F_1 分数给予假阳性和假阴性更多的权重^[21], 其计算公式如式 (6) 所示, F_1 分数越高说明模型综合效果越好。

$$F_1 = 2 \times \frac{precision \times recall}{precision + recall} \quad (6)$$

模型测试和模型训练有类似的数据预处理过程, 振动信号被读取、压缩去噪、数据增强, 但是为了降低识别结果的随机性, 数据增强部分只保留时域拉伸, 并且不再取随机的拉伸因子, 而是设置固定的拉伸因子。经过多次实验, 发现不同的时间拉伸因子对测试结果有影响。测试 1dCNN 的时候, 拉伸因子设置为 1.3, 测试 2dCNN 的时候拉伸因子测试 1.2, 即所有信号被拉伸 1.3 和 1.2 倍。当这样设置的时候, 模型取得最佳表现, 如果在测试的时候, 信号的时域拉伸因子随机取值, 将导致识别准确率向下浮动 0.0 ~ 0.2 左右。

为了研究数据压缩和去噪过程参数选择对模型效果的影响, 对多种数据预处理的方法进行了实验, 发现在使用 db10 小波函数, $k=4$ 重建信号的时候, 1dCNN 轴型识别网络可以达到最高的识别准确率 0.986; 在使用 coif5 小波函数, $k=4$ 重建的时候, 2dCNN 轴型识别网络可达到最高的识别准确率 0.973。

表 4 是 1dCNN 最佳权重文件在测试集上的混淆矩阵, 其中纵轴表示样本的真实值, 横轴是模型分类值。除了一个四轴车被识别为二轴车, 一个六轴车被识别为四轴车, 其余的车辆都可以被正确分类。

表 4 1dCNN 测试集混淆矩阵

		识别车辆类型				
真实车辆类型	车辆类型 / 轴	2	3	4	5	6
	2	27	0	0	0	0
	3	1	35	0	0	0
	4	0	0	47	0	0
	5	0	0	0	11	0
	6	0	0	1	0	28

表 5 是 2dCNN 最佳权重文件在测试集上的混淆矩阵, 除了三个三轴车被错误识别为二轴车、四轴车, 一个六轴车被识别为四轴车和一个五轴车被识别为六轴车, 其余的车辆可以被正确分类。

表 5 2dCNN 测试集混淆矩阵

真实车辆类型	识别车辆类型					
	车辆类型 / 轴	2	3	4	5	6
	2	27	0	0	0	0
	3	1	33	2	0	0
	4	0	0	47	0	0
	5	0	0	0	10	1
	6	0	0	1	0	29

两个模型在测试集上对每个轴型的详细分类精度、召回率和 F_1 分数以及全类准确率分别如表 6、表 7 所示。

表 6 1dCNN 测试集参数

	精度	召回率	F_1 分数	准确率
2 轴	0.964	1.000	0.981	0.986
3 轴	1.000	0.972	0.985	
4 轴	0.979	1.000	0.989	
5 轴	1.000	1.000	1.000	
6 轴	1.000	0.965	0.982	
平均	0.988	0.987	0.987	

表 7 2dCNN 测试集参数

	精度	召回率	F_1 分数	准确率
2 轴	0.964	1.000	0.981	0.973
3 轴	1.000	0.916	0.956	
4 轴	0.959	1.000	0.979	
5 轴	1.000	0.909	0.952	
6 轴	0.966	1.000	0.983	
平均	0.978	0.965	0.970	

从表 6 可以看出, 1dCNN 在测试集分类平均分类准确度为 0.986, 对三轴车、五轴车、六轴车的分类精确度可达到 1, 二轴车、三轴车、五轴车的召回率也达到了 1, 从 F_1 分数来看, 该模型对五轴车识别效果最佳, 二轴车最弱。这个结果可能有点违反常识, 因为二轴车在训练集中具有最多的样本, 占据总数据集的 36.5%。而且, 二轴车仅有两个波峰, 相比其他类型车辆的多个波峰, 二轴车应该是相对容易学习的样本。分析表 6 可知, 二轴车的召回率是 1, 即所有二轴车都可以被识别出来, 而识别的精度仅为 0.964, 是所有车轴类型中最低的, 这说明其他类型的车辆比较容易被该模型错误分类为二轴车。

表 7 可见, 2dCNN 在测试集分类平均准确度是 0.973, 对三轴车、五轴车的分类精确度最高达到了 1, 对二轴车、四轴车、六轴车的召回率也达到了 1; 从 F_1 分数来看, 该模型对六轴车的分类效果最好, 五轴车最弱。五轴车的精度为 1, 而召回率仅为 0.909, 说明很多五轴车没有被模型识别出来, 同样的情况也发生在三轴车上, 具有高识别精度和低的召回率。

对比表 6 和表 7 可以发现, 两个模型对二轴车和四轴车

的识别召回率都为 1, 两个模型对三轴车和五轴车的识别精度都为 1。而由表 4 和表 5 混淆矩阵可知, 几乎所有错误的识别都发生在其它车辆被识别成了二轴车或四轴车, 除了一个被识别为六轴车。所以, 可以认为二轴车和四轴车具有相比其它车轴类型更加一般的特征, 这导致其它车辆类型容易被错误识别为这两个类别。三轴车和五轴车具有非常显著的特征, 这导致了它们的识别精度高。精度和召回率是两个制衡的指标, 两者关系可以从表 6 和表 7 中看出, 大部分类别召回率和精度只有一个可以达到 1, 唯一存在的一个例外就是 6 中的 5 轴车, 精度和召回率都达到了 1, 这可能是由于 5 轴车样本少导致的。

总体来说, 1dCNN 的总体识别准确率, 每个类别的精度、召回率、 F_1 分数都不低于 2dCNN。同时, 从图 14 中 1dCNN 损失曲线来看, 1dCNN 的学习过程收敛更快, 更稳定; 2dCNN 在验证集上的损失振荡幅度比较大。

5 结论

本文主要构建了两种卷积神经网络, 结合小波变换等数据预处理方法, 实现了以较高的准确率从车辆振动信号中识别出对应的轴型信息, 主要结论如下。

(1) 针对车辆振动信号的特点构建了 1dCNN 和 2dCNN 两个网络模型, 并对原型网络进行了优化和修改, 如减少卷积扩张通道数目, 增加残差块中侧通道卷积层等。特别是车辆信息识别的研究很少涉及二维卷积神经网络, 本文在一维网络结构基础上构建了从信号频谱图提取特征的 2dCNN 模型。针对二维频谱图输入导致的计算量大的问题, 将普通卷积层替换为深度可分离卷积层, 获得了理想的基于振动参数的车辆信息识别结果。

(2) 根据车辆真实行走实验采集数据进行了 1dCNN 模型和 2dCNN 模型的训练, 1dCNN 在测试集上可以达到 0.986 的识别准确率, 各类平均精度为 0.988, 平均召回率为 0.987, 平均 F_1 分数为 0.987; 2dCNN 在测试集上可达 0.973 的准确率, 各类平均精度 0.978, 平均召回率 0.965, 平均 F_1 分数 0.970。两个模型在测试集上都可以达到较高的识别准确率、精度、召回率。

(3) 为了测试数据预处理方法对模型的影响, 训练了不同预处理参数下的模型。结果表明小波函数, 重建分量个数 k , 频谱图参数会对模型预测产生影响。当不对数据进行任何处理训练的 1dCNN 模型在测试集上可达 0.913 的准确率, 当采用小波变换和数据增强处理后准确率最高可达 0.986。在选择的几种小波函数和 k 参数中, 1dCNN 最低识别准确率为 0.933, 此时小波函数为 rbio2.2, $k=5$; 1dCNN 最高识别率为 0.986, 此时小波函数为 db10, $k=4$ 。2dCNN 的最低识别准确率为 0.793, 此时小波函数为 db10, $k=4$, 频谱图为梅尔频谱图; 2dCNN 最高识别准确率为 0.973, 小波函数为 coif5, $k=4$, 频谱图为梅尔频谱图。

参考文献:

- [1] YING K, AMERI A, TRIVEDI A, et al. Decision tree-based machine learning algorithm for in-node vehicle classification[C]//2015 IEEE Green Energy and Systems Conference (IGESC).Piscataway:IEEE, 2015: 71-76.
- [2] MA R, ZHANG Z, DONG Y, et al. Deep learning based vehicle detection and classification methodology using strain sensors under bridge deck[J]. Sensors, 2020, 20(18): 50-51.
- [3] MA W, XING D, MCKEE A, et al. A wireless accelerometer-based automatic vehicle classification prototype system[J]. IEEE transactions on intelligent transportation systems, 2014, 15(1): 104-111.
- [4] KHONG DUC C, HOANG V-P, TIEN NGUYEN D, et al. A low-cost, flexible pressure capacitor sensor using polyurethane for wireless vehicle detection[J]. Polymers, 2019, 11(8): 1247.
- [5] AL-TARAWNEH M, HUANG Y, LU P, et al. Vehicle classification system using in-pavement fiber bragg grating sensors[J]. IEEE sensors journal, 2018, 18(7): 2807-2815.
- [6] LI Y, TOK A, RITCHIE S G. Deep ensemble neural network approach for federal highway administration axle-based vehicle classification using advanced single inductive loops[J]. Transportation research record: journal of the transportation research board, 2022, 2676(3): 1-16.
- [7] YU Y, CAI C, DENG L. Vehicle axle identification using wavelet analysis of bridge global responses[J]. Journal of vibration and control, 2017, 23(17): 2830-2840.
- [8] YAN L, FRASER M, ELGAMAL A, et al. Neural networks and principal components analysis for strain-based vehicle classification[J]. Journal of computing in civil engineering, 2008, 22(2): 123-132.
- [9] 李健, 王晓明, 张英海, 等. 基于深度卷积神经网络的地震震相拾取方法研究[J]. 地球物理学报, 2020, 63(4): 1591-1606.
- [10] 陈伟, 孙强, 齐月月, 等. 深度卷积神经网络在心音分类方法中的应用[J]. 计算机工程与应用, 2021, 57(16): 182-189.
- [11] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR).Piscataway:IEEE, 2015:27-30.
- [12] HE K, ZHANG X, REN S, et al. Identity mappings in deep residual networks[EB/OL].(2016-05-16)[2023-09-26].https://arxiv.org/abs/1603.05027.
- [13] WANG T, LU C, SUN Y, et al. Automatic ECG classification using continuous wavelet transform and convolutional neural network[J]. Entropy, 2021, 23(1): 109-119.
- [14] Depthwise separable convolutions in pytorch : päpper's machine learning blog——this blog features state of the art applications in machine learning with a lot of pytorch samples and deep learning code. you will learn about neural network optimization and potential insights for artificial intelligence for example in the medical domain[EB/OL].(2022-05-05)[2023-09-28]. https://www.paeppler.com/blog/posts/depth-wise-separable-convolutions-in-pytorch/.
- [15] HUANG J, CHEN B, YAO B, et al. ECG arrhythmia classification using stft-based spectrogram and convolutional neural network[J]. IEEE access, 2019, 7: 92871-92880.
- [16] CHOWDHURY T H, POUDEL K N, HU Y. Time-frequency analysis, denoising, compression, segmentation, and classification of pcg signals[J]. IEEE access, 2020, 8: 160882-160890.
- [17] LELAND R, ANALYTICS V. Understanding the Mel Spectrogram[EB/OL].(2022-05-04)[2023-09-22]. https://medium.com/analytics-vidhya/understanding-the-mel-spectrogram-fca2afa2ce53.
- [18] HOWARD A G, ZHU M, CHEN B, et al. Mobilenets: efficient convolutional neural networks for mobile vision applications[EB/OL]. (2017-04-17)[2023-09-29]. https://arxiv.org/abs/1704.04861.
- [19] KINGMA D P, BA J. Adam: a method for stochastic optimization[EB/OL]. (2014-11-22)[2023-09-30]. https://arxiv.org/abs/1412.6980.
- [20] SMITH L N, TOPIN N. Super-convergence: very fast training of neural networks using large learning rates[EB/OL]. (2017-08-23)[2023-10-06].https://arxiv.org/abs/1708.07120.
- [21] RIGGIO C. What's the deal with Accuracy, Precision, Recall and F1?[EB/OL]. (2019-11-03)[2023-10-02]. https://towardsdatascience.com/whats-the-deal-with-accuracy-precision-recall-and-f1-f5d8b4db1021.

【作者简介】

罗智元(1999—), 女, 江西南昌人, 本科, 研究方向: 人工智能、机器人。

刘小勇(1972—), 男, 陕西西安人, 博士, 副教授, 研究方向: 智能机器人控制与边缘计算、物联网及实时信息处理等。

(收稿日期: 2023-11-28)