

· 计算机工程与应用 ·

基于卷积神经网络的驾驶员不安全行为识别

田文洪*, 曾柯铭, 莫中勤, 吝博强

(电子科技大学信息与软件工程学院 成都 610054)

【摘要】提出了一种基于卷积神经网络的驾驶员违规行为识别方法。首先,利用特定卷积神经网络对驾驶员的实时图像提取特征,然后并行对多种行为分别进行二分类。建立了一个真实场景下的驾驶员违规数据集,在此数据集上的测试说明了该方法的高效和良好的泛化能力。实验结果表明,该方法在约10万张图像的数据集中对打电话、吸烟、不系安全带3种行为分别达到了99.85%、99.62%、98.68%的识别率,同时使用当前较先进的Inception-v3和Xception模型测试,也获得了类似的识别效果。

关键词 卷积神经网络; 深度学习; 模式识别; 驾驶员不安全行为识别
中图分类号 TP391.4 **文献标志码** A **doi**:10.3969/j.issn.1001-0548.2019.03.012

Recognition of Unsafe Driving Behaviors Based on Convolutional Neural Network

TIAN Wen-hong*, ZENG Ke-ming, MO Zhong-qin, and LIN Bo-qiang

(School of Information and Software Engineering, University of Electronic Science and Technology of China Chengdu 610054)

Abstract The unsafe behavior of the driver is one of the important causes of many incidents. This paper presents a method to recognize unsafe driving behaviors based on the convolutional neural network. Firstly, the characteristics of the real-time image are extracted by the specific convolutional neural network, and then three kinds of behaviors are classified into two categories in parallel. The data set of unsafe driving behaviors in a real scene is established. The test on this dataset illustrates the efficiency and good generalization of the method. The experimental results show that the method achieves 99.85%, 99.62% and 98.68% accuracy for calling, smoking and unbelted in the data set of about 100 000 images, which is comparable to the results obtained by recent advanced Inception-v3 and Xception models.

Key words convolutional neural network (CNN); deep learning; pattern recognition; recognition of unsafe driving behaviors

截至2016年底,我国机动车保有量达2.9亿辆,其中汽车1.94亿辆,机动车驾驶人3.6亿人,而汽车驾驶人超过3.1亿人。海量的汽车导致了大量的交通事故,而这些交通事故中,很多是由不文明和不安全的交通行为引起,如一只手抽烟或打电话,单手操控汽车等;同时,由于司机对安全行为的认识不够,还存在大量不系安全带的情况。

对于重要货物运输或者载人客车,事故造成的危害是巨大的。因此,社会对于能够自动识别驾驶员违规行为并发出警报的系统有着强烈需求。

国内外对于驾驶员违规行为的研究较少且比较集中,数据集难以获得,大体可分为以下4个方面:

1) 对打电话的识别,文献[1]使用了基于支持向量机的模型识别开车打电话的情况,其数据集是模拟环境下的标准正脸。文献[2]也是基于支持向量机,但数据来自公共交通摄像。文献[3]使用了隐含马尔科夫模型,并使用了Adaboost方法,数据也来自室内的模拟生成。文献[4]在仪表盘上方安装摄像头采集数据,通过脸部、嘴、手部动作来进行识别,采用了隐条件随机场(hidden conditional random fields, HCRF)模型。文献[5]通过分析手的放置和面部图像进行识别,采用了快速区域卷积神经网络(faster regions with convolutional neural network, faster RCNN)进行目标检测。文献[6]利用了监督下降法(supervised

收稿日期: 2018-01-19; 修回日期: 2018-10-16

基金项目: 国家自然科学基金(61672136, 61828202); 四川省科技支撑计划(2016GZ0322)

作者简介: 田文洪(1972-),男,教授,主要从事大数据分析及深度学习方面的研究。Email: tian_wenhong@uestc.edu.cn

descent method, SDM)定位脸部特征,使用左右两个边界框来判断是否打电话,在分类过程中使用了梯度直方图(histogram of gradients, HOG)和迭代提升法Adaboost,在数据集上取得了93.9%的准确率。

2) 对面部和手部的分析:主要是美国加州大学圣地亚哥分校(University of California, San Diego, UCSD)智能驾车安全驾驶实验室的工作。文献[7]在车内使用了两个Kinect摄像头,一个观测面部,一个观测手部,这些信息可用于检测4种类型的行为:调镜子、调节音响、调节档位和控制空调。文献[8]的视角则更加专注于手的位置,车内顶部的摄像头监控着如下的图像:通过将一幅图像的3个区域(方向盘、档位、功能面板)分别进行手的识别,以此来识别驾驶员的行为。文献[9]又在以上研究中加入了对眼睛目光的追踪,但依旧是对方向盘、档位、功能面板3个部分的识别。

3) 对侧面半身图像的分析:这部分是来自东南大学的成果。文献[10]的工作中设计了更加完整的识别内容,图像拍摄于侧面,包括了抓方向盘、调控档位、吃东西以及打电话4个部分。在提取特征后,使用了随机森林、k近邻、多层感知机算法,其中随机森林取得了90.5%的识别率。文献[11]改进了上述工作,使多层感知机的结果达到了90.61%的识别率。文献[12]采用径向基函数(radial basis function, RBF)核的支持向量机方法,达到了94.25%的结果,而交叉核的支持向量机达到了每秒67图速度。在此之后,文献[13]改进了特征提取方法,使多层感知机的效果达到了94.75%的识别率。在此基础上,文献[14]用随机森林达到了96.56%的识别率,文献[15]使用卷积神经网络达到了99.78%的识别率。

4) 侧面的图像,数据集来自kaggle^[16]。不同于以上工作,这个数据集是唯一被公开的,虽然其只能被授权用于kaggle竞赛,但仍具有一定价值。该数据集的图像来自侧面,信息较为完整。标签分为10类:安全驾驶、发信息(左)、发信息(右)、打电话(左)、打电话(右)、喝饮料、操作音响、整理妆容、与乘客聊天、看向后方。

综合上述,本文提出一种基于卷积神经网络的轻量级驾驶员不安全行为识别方法。

1 整体系统结构

对驾驶员违规行为的识别可以看做多标签分类问题,即一个样本(行为)具有多个特性(是否系安全带、是否抽烟、是否打电话)。处理这样的问题通常采用两种做法:单个模型处理和多个模型并发处

理。本文采用多个模型并发处理的方法,整体流程如图1所示,待模型训练完成,即可重复用于识别实时行为。

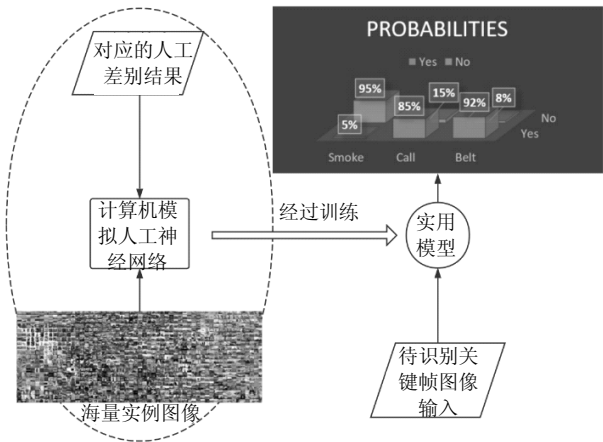


图1 整体流程图

1.1 一个模型解决所有问题

优点是实现简单,等同于多分类问题(一个正常类,多个交叉的错误类);缺点是不同行为之间的耦合较强,从软件工程的角度来讲,不利于软件开发。

1.2 3个模型解决3个独立问题

3种违规行为独立,判定过程可作为3个独立的二分类任务。其优点如下:

1) 对每种违规行为的判定可作为一个模块,易于功能扩展,加入新模块无需对已训练好的模块重新训练。

2) 由于不同模块的复杂度不同,处理同样的原图,相对于判定安全带,抽烟的判定需要更多像素的输入(因为烟的区域小),程序各模块独立性高。当需要加入新的模块(如对于其他违规行为的识别)时,这种方式有更大的优势。

2 神经网络模型

2.1 模型建立

基于卷积神经网络结构,本文将各层的激活函数设为线性整流函数(rectified linear unit, ReLU),最终分类为softmax,将模型理想输出与实际输出的交叉熵确定为loss,即优化目标应使两者的交叉熵足够小。理想输出的分布 P 与实际输出的分布 Q 的交叉熵 $H(P, Q)$ 的定义为:

$$H(P, Q) = \sum_i P(i) \times \log \frac{1}{Q(i)} \quad (1)$$

模型中的参数 w 使用梯度下降的方法进行更新,每一轮迭代都沿着梯度方向更加接近理想参数:

$$w \leftarrow w - \alpha \times \frac{\partial \text{loss}}{\partial w} \quad (2)$$

式中, α 是参数更新的步长(即学习率), 迭代终止的条件可以是loss低于某个阈值, 也可以是达到了本文设定的迭代次数。参数的初始化采用随机均匀的方式, 产生一系列均匀分布的随机参数, 经过尝试, 发现均匀分布更适合本文模型:

$$f(x) = \frac{1}{b-a} \quad x \in (a,b) \quad (3)$$

关于梯度下降的优化求解器较多, 在一阶梯度中, 随机梯度下降(stochastic gradient descent, SGD)是最常用的随机梯度下降方式, Adagrad是在SGD的基础上, 对更新步长进行优化的产物, 有:

$$n_t = n_{t-1} + \left(\frac{\partial \text{loss}}{\partial w}\right)^2 \quad (4)$$

$$\Delta w = -\frac{\partial \text{loss}}{\partial w} \frac{a}{\sqrt{n_t + \xi}} \quad (5)$$

式中, Δw 是每次更新的参数量; ξ 是一个常量; n

在每次迭代中对梯度与步长的乘积进行约束, 可看作是一个自适应的步长。Adam结合了Adagrad和RMSprop的优点, 引入梯度的二阶矩估计进行参数优化:

$$m_t = \alpha m_{t-1} + (1-\alpha) \frac{\partial \text{loss}}{\partial w} \quad (6)$$

$$n_t = \beta n_{t-1} + (1-\beta) \left(\frac{\partial \text{loss}}{\partial w}\right)^2 \quad (7)$$

$$\tilde{m}_t = \frac{m_t}{1-\alpha^t} \quad (8)$$

$$\tilde{n}_t = \frac{n_t}{1-\beta^t} \quad (9)$$

$$\Delta w = -\frac{\partial \text{loss}}{\partial w} \frac{\tilde{m}_t}{\sqrt{\tilde{n}_t + \xi}} \quad (10)$$

式中, m_t 表示对梯度的一阶矩估计; n_t 表示对梯度的二阶矩估计。在对其进行修正后, 最终它们共同对步长形成一个动态约束。

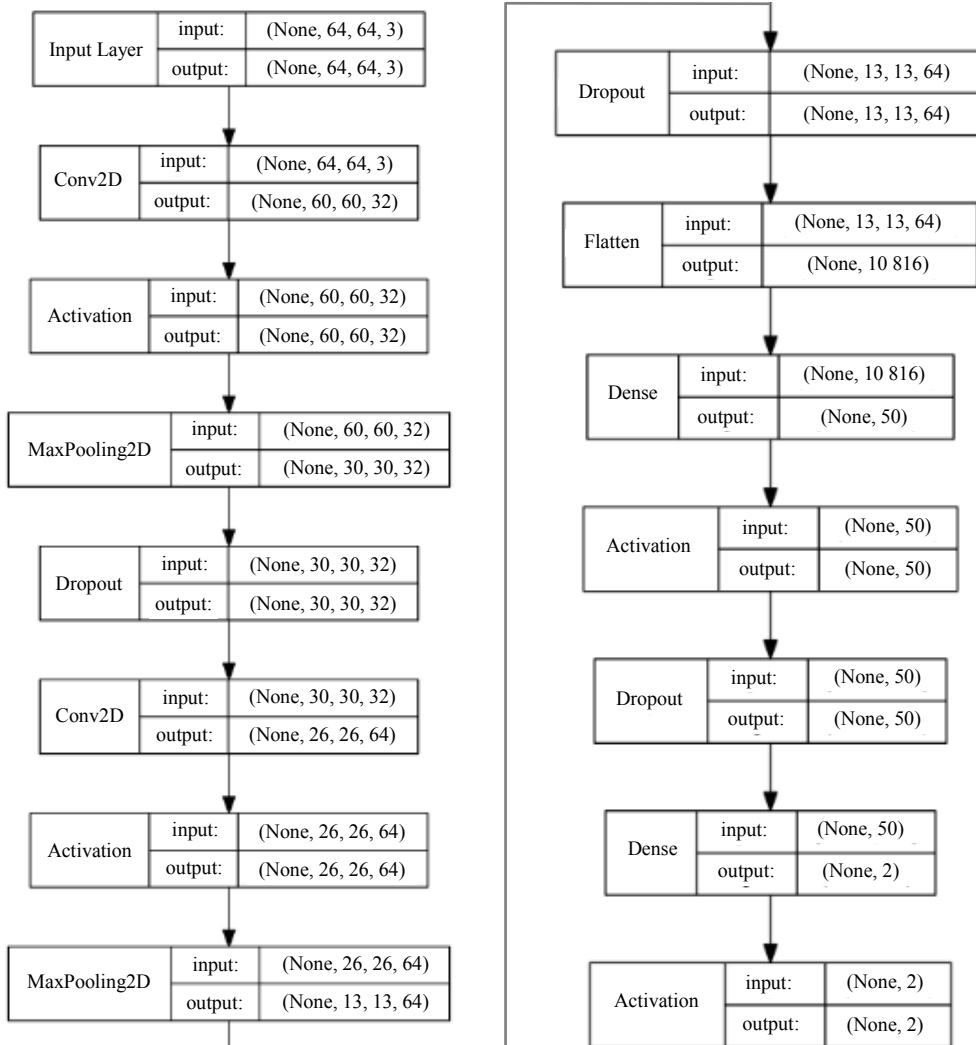


图2 模型结构

本文经过尝试,对比了SGD、Adagrad和Adam 3种优化方法,最终选用了相对效果最佳的Adam。

2.2 构造网络结构

根据奥勒姆剃刀原理,适当复杂的模型有利于防止过拟合。本文的模型构造基于以下3种考虑:模型参数量能满足在一个较小的内存空间内使用、模型计算量的减少不会过多影响识别率、模型本身具有较好的平移不变性。

在大量尝试复杂网络结构后,本文最终选用了类似于Lenet的卷积神经网络,它在不低于其他结构准确率的情况下达到了最小的参数量与计算量。输入为图像的红绿蓝(RGB)色度值,输出为分类情况。

卷积层与pool层可以考虑1~3层,最后再将特征全连接为一个向量,添加softmax分为两类。为了精简模型,本文将普通的卷积层更改为Separable Conv^[17],在不损失精度的情况下减少了参数量;同样的原因,本文没有添加BatchNormalization层(本文场景中使用BatchNormalization层并不会提高模型准确率)。

网络结构如图2所示,其中dropout参数设置为0.25,模型总参数的数量为544 067个。

3 建立真实数据集

虽然kaggle上的数据集公开,但其仅能应用于竞赛,暂时还没有得到其关于研究的授权。为了解决数据的问题,本文提出了两种方案:模拟拍摄和实地采集。

3.1 两种数据来源

3.1.1 模拟拍摄

模拟车内环境进行拍摄,图像采集条件如表1所示。

第一次采集得到图像总数为近1 200张,各正例反例接近1:1,用于初步验证方案。

表1 图像采集条件

情景	模拟道具
E1: 不系安全带	长背带(外观类似黑色安全带)
E2: 吸烟	多种外形香烟、多种姿势
E3: 打电话	多种不同外形的手机

对于这些所有采集的图像,选出90%作为训练集,10%作为测试集。90%的数据中随机选出80%进行训练,20%作为验证集。

3.1.2 实地采集

在多次真实图像采集未果的情况下,获得了某

集团提供的一辆运输车辆内部记录,包括9天内24小时对2个驾驶员的监控视频。

本文从中选取了一些能用于模型训练的内容,根据数据的特点,讨论了对打电话、系安全带、吸烟这3种违规行为的识别,最后处理获得了93 222张打电话行为图像(call),97 502张吸烟行为图像(smoke),101 917张系安全带行为图像(belt),每种行为均包含两类,如图3所示。

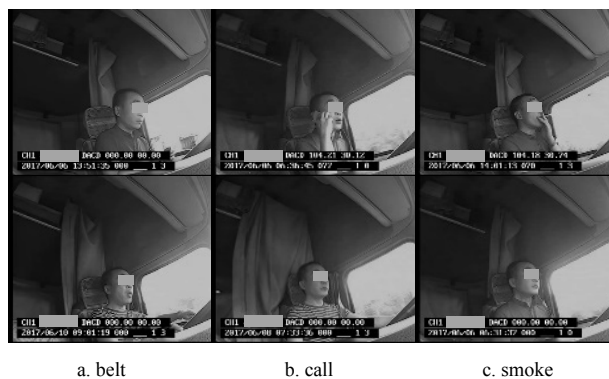


图3 3种违规行为(上图违规,下图为正常)

3.2 对数据的处理

3.2.1 视频

针对选中的视频,使用opencv对其进行帧采样。在尝试不同采样频率后,充分考虑到模型训练的过拟合和数据量本身的问题,最终决定1 s采样一次。

对于采样后的图像,按照以下原则进行筛选,最终得到大量图像。

1) 剔除无关图像:司机不在车内;图像受损;停车时司机做出的非常规动作(明显不像是在开车的动作),如吃泡面、翘脚、开门。停车时司机的常规动作可以保留,如吸烟、正常坐姿,只要像是司机在开车即可。

2) 按照吸烟与否对所有图像分类:吸烟动作是指香烟离司机嘴较近的动作,不包括只有烟雾的动作,也不包括烟离嘴较远的情况;在每个视频文件夹中建立一个smoke-yes文件夹与smoke-no文件夹;将所有图像分类到两个文件夹中。

3) 按照打电话与否对所有图像分类:打电话动作是指手机离耳朵较近的情况,不包括看手机;在每个视频文件夹中建立一个call-yes文件夹与call-no文件夹;将所有图像分类到两个文件夹中。

4) 按照系安全带与否对所有图像分类:在每个视频文件夹中建立一个belt-yes文件夹与belt-no文件夹;将所有图像分类到两个文件夹中;看不清楚安全带视作没系。

3.2.2 图像与标签

这个过程是对拍照所得图像的规范化, 使得系统模型的输入是同一规格的图像。具体过程根据图像采集情况如下:

- 1) 将图像左右裁边至合理尺寸。
- 2) 将图像上下裁边至合理尺寸。
- 3) 缩放图像(至64*64或299*299 pixels)。

不同摄像设备不同摄像角度输出图像的处理方式不同, 应具体情况具体应对。

对于图像数量较少的类(违规类), 适当采用crop的方式对图像进行扩充, 即对同一张图像随机做比例不超过0.3的裁剪, 随后缩放为网络的输入尺寸, 使得此类的动作有一定变形, 区别于原动作。这个过程在内存中进行, 是对图像进行的一种数据增强。

在拥有标准一致的图像之后, 对这些图像打上标签。本文采用0/1标记二分类, 1代表违规类即正类, 0代表正常类即负类。

4 模型的训练、测试与调优

4.1 基于模拟的少量数据

为了检验模型对识别过程的表达能力, 本文使用了前期在模拟环境采集的3种模式图像, 着力于识别不系安全带这种行为, 分别在不同的数据量和不同的网络深度进行了训练, 迭代50次(迭代50次之后可以看到算法已完全收敛), 结果如表2所示。

表2 前期的不同模型版本

模型版本	数据量	网络深度	正反类比例	训练数据占比/%	测试集准确率/%
E1_ver1	200	6	1:1	80	86.05
E1_ver2	200	4	1:1	80	77.52
E1_ver3	1 400	6	1:1	80	97.67
E1_ver4	1 400	4	1:1	80	94.57

可以看出, E1_ver3的模型达到了97.67%的泛化能力, 更多的数据带来了测试集准确率的显著提升。50次迭代可能并未达到极限, 所以对E1_ver3迭代训练100次, 测试得出96.9%的准确率。这个结果虽然略低于50次迭代的结果, 但由于测试集数量有限, 波动也在正常范围内。模型是否达到了全局最优, 还需要更多数据的测试。

4.2 基于采集的大量数据

本文利用实际车内视频转化得到的图像, 进行了更大规模的数据训练与测试, 迭代次数同样选择为50次, 训练策略与少量数据时相同。然而, 在训

练过程中, 不系安全带和打电话这两种行为识别都取得了较高的准确率, 吸烟行为的识别却并不好。

经过分析, 这是由于样本的正类与反类的数量差距过大造成的。为了解决这个问题, 提出了以下两种方案:

- 1) 对正类进行过采样或对反例进行欠采样;
- 2) 调整模型训练中正类与反类对loss贡献的权重。

经过分析与小规模测试, 发现对权重参数进行调整效果更佳。

4.3 权重调整

二分类中, 一般将要识别的那一类记作正类, 另一类记作反类。本文定义了一个权系数, 用于平衡训练时两类的权重:

$$\text{正类权重} = \frac{\text{反类数量}}{\text{正类数量}} \times \text{权系数} \quad (11)$$

将反类权重置为1, 则在不同权系数下, 对吸烟行为的识别模型结果如表3所示。

表3 不同权系数的结果

权系数	训练集准确率/%	测试集准确率/%	所有正类识别率/%	所有反类识别率/%
1/1	98.74	98.75	0	100.0
1/10	98.74	98.75	0	100.0
1/15	99.82	99.52	93.2	99.8
1/20	99.90	99.59	95.3	99.9
1/22	99.95	99.62	95.8	99.9
1/25	99.95	99.66	94.4	100.0
1/30	99.95	99.63	95.1	99.9

可以看出, 在权系数为1/22时, 模型取得了训练集与测试集准确率的一个高峰, 且对于正类的识别率较高。

4.4 结果

将打电话与不系安全带这两种行为的识别模型按照吸烟行为模型的调整方法, 进行调整, 在多次试错之后, 取得了如表4a、表4b所示的结果。

表4a 本文模型数据量

模型	总数据量	正类数据量	反类数据量	训练数据量占比/%
call	93 222	2 764	90 458	80
smoke	97 502	1 224	96 278	80
belt	101 917	34 699	67 218	80

表4b 本文模型识别率

模型	训练集准确率/%	测试集准确率/%	所有正类识别率/%	所有反类识别率/%
call	99.99	99.85	99.50	100.00
smoke	99.95	99.62	95.80	99.90
belt	99.47	98.68	98.54	99.71

4.5 识别结果分析

抽取3类分类失败的图像, 共计1 447张。可以看出这些图像的特点:

- 1) 约20%的图像质量不高, 有标注上的误判。
- 2) 约30%为逻辑判断上的“模棱两可”, 即单凭图像无法判断是否违规, 典型图像如图4所示。
- 3) 剩余错判来自模型本身的分类能力。



图4 第二种错误的典型图像

4.6 与Xception和Inception-v3的对比

为了与Xception^[17]和Google Inception^[18]对比, 将本文模型中64*64 pixels的输入扩展为299*299 pixels的输入, 引入了更多的信息量。分别在打电话、吸烟、安全带行为识别的场景下讨论了迁移模式(底层特征提取网络的参数不变)的结果, 结果对比如图5所示(这里使用的是Inception v3)。

其中, 两种模型均冻结(初始化为Imagenet比赛训练好的模型参数, 迭代中不再更新冻结部分)底层约1/2的参数(发现这个比例较其他比例更适合本文的数据量), 顶层使用2个神经元的softmax层, 梯度下降时尝试了各种方法的组合, 最终在先使用Adam优化, 再使用低学习率SGD+momentom调优的组合中取得了最佳的结果。

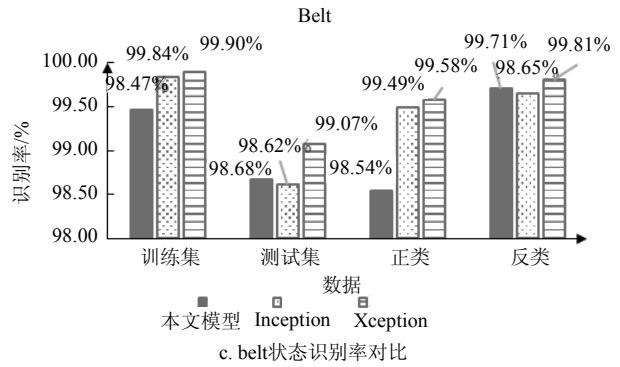


图5 3种行为与3种方法在不同数据上的识别率对比

可以看出, 本文模型采用的参数量为55万左右的网络结构在与参数量为2 000万左右的Inception与Xception的对比中, 并没有明显的劣势。其中, 对于打电话与吸烟行为, 本文模型的测试集识别率略低于另外两个结构; 对于不系安全带行为, 本文模型的测试集识别率高于Inception v3的结果。另一方面, 本文模型对正类的识别较好, 可以避免危险, 这在实际工业环境中是可用的。

5 讨论

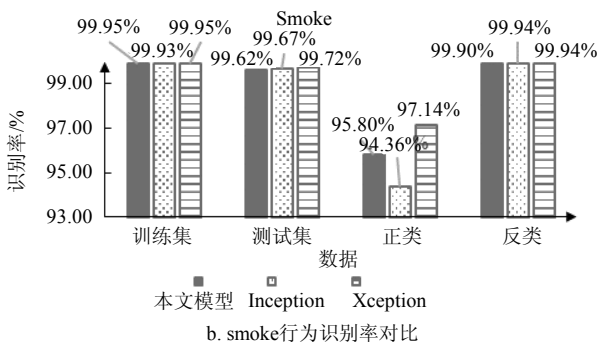
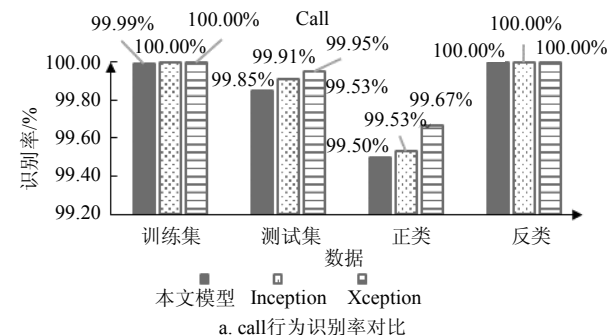
1) 对于网络结构, 本文尝试的是一种轻量级的卷积神经网络。通过与复杂模型的对比, 说明对于某些特定场景, 简单模型有更高的可用性。

2) 对于数据预处理, 本文没有使用翻转、模糊等增强方案, 是因为在应用场景中识别的对象较为固定(摄像头的位置已经固定), 在测试集中就没有太多其他风格的图像。但是考虑到模型的泛化性能, 这些预处理方案应该能够对模型有一些提升。

3) 在对行为的识别上, 本文只考虑了单帧图像反映的行为模式。但是后期可以考虑使用循环神经网络(recurrent neural network, RNN)或是长短期记忆网络(long short-term memory, LSTM)来处理一定长度的图像序列, 发掘出识别行为在时间维度上的连贯关系, 这样能够提升一些识别准确率。

4) 有鉴于Capsule网络^[19]在手写数字识别数据集(mixed national institute of standards and technology database, MNIST)上取得的成果, 本文将Flatten层之后的网络更换为PrimaryCap和两个Capsule用作分类, 但其在增加复杂性的同时并未提升准确率, 这一结果值得后续研究。

5) 本文采用的模型虽然在3种行为的识别上都取得了极好的结果, 受限于没有开放的数据集做更大规模对比测试, 希望学术界很快能有此类应用的



更多开放数据集。

6 结束语

本文提出了一种基于卷积神经网络的轻量级驾驶员不安全行为识别方法, 通过多层卷积神经网络提取特征并进行分类, 能够对打电话、吸烟、不系安全带这3种行为进行有效的识别。取得了如下成果:

- 1) 建立了一个真实场景下的驾驶员不安全行为数据集;
- 2) 采用多个模型并发处理多种不安全行为;
- 3) 提出一种网络层次少、参数较少、计算量较小且易于实用的卷积神经网络模型, 并采用权重参数调整方法解决正反类数据量差距过大的问题;
- 4) 对于不安全行为的识别率超出已有成果, 利于实用。

参 考 文 献

- [1] BERRI R A, SILVA A G, PARPINELLI R S, et al. A pattern recognition system for detecting use of mobile phones while driving[C]//International Conference on Computer Vision Theory and Applications. Berlin, Germany: IEEE, 2015: 411-418.
- [2] ARTAN Y, BULAN O, LOCE R P, et al. Driver cell phone usage detection from HOV/HOT NIR images[C]//Computer Vision and Pattern Recognition Workshops. Columbus, USA: IEEE, 2014: 225-230.
- [3] CRAYE C, KARRAY F. Driver distraction detection and recognition using RGB-D sensor[EB/OL]. [2015-02-01]. <https://arxiv.org/abs/1502.00250>.
- [4] ZHANG Xue-tao, ZHENG Nan-ning, WANG Fei, et al. Visual recognition of driver hand-held cell phone use based on hidden CRF[C]//IEEE International Conference on Vehicular Electronics and Safety. Beijing, China: IEEE, 2011:248-251.
- [5] LE T H N, ZHENG Yu-tong, ZHU Chen-chen, et al. Multiple scale faster-rcnn approach to driver's cell-phone usage and hands on steering wheel detection[C]//Computer Vision and Pattern Recognition Workshops. Las Vegas, USA: IEEE, 2016: 46-53.
- [6] SESHADRI K, JUEFEIXU F, PAL D K, et al. Driver cell phone usage detection on strategic highway research program (SHRP2) face view videos[C]//IEEE Conference on Computer Vision and Pattern Recognition Workshops. Boston, USA: IEEE Computer Society, 2015: 35-43.
- [7] MARTIN S, OHN-BAR E, TAWARI A, et al. Understanding head and hand activities and coordination in naturalistic driving videos[C]//Intelligent Vehicles Symposium Proceedings. Michigan, USA: IEEE, 2014: 884-889.
- [8] TRIVEDI M. Driver hand activity analysis in naturalistic driving studies: challenges, algorithms, and experimental studies[J]. Journal of Electronic Imaging, 2013, 22(4): 1119.
- [9] OHN-BAR E, MARTIN S, TAWARI A, et al. Head, eye, and hand patterns for driver activity recognition[C]//International Conference on proceedings of the Pattern Recognition. Stockholm, Sweden: IEEE, 2014.
- [10] ZHAO C, ZHANG B, HE J, et al. Recognition of driving postures by contourlet transform and random forests[J]. IET Intelligent Transport Systems, 2012, 6(2): 161-8.
- [11] ZHAO Chi-hang, GAO Yong-sheng, HE Jie, et al. Recognition of driving postures by multiwavelet transform and multilayer perceptron classifier[J]. Engineering Applications of Artificial Intelligence, 2012, 25(8): 1677-1686.
- [12] ZHAO C, ZHANG B, LIAN J, et al. Classification of driving postures by support vector machines[C]//Sixth International Conference on Image and Graphics. Anhui, China: IEEE Computer Society, 2011: 926-930.
- [13] ZHAO Cai-hang, ZHANG Bai-ling, ZHANG Xiao-zheng, et al. Recognition of driving postures by combined features and random subspace ensemble of multilayer perceptron classifiers[J]. Neural Computing & Applications, 2013, 22(1): 175-184.
- [14] YAN Chao, COENEN F, ZHANG Bai-ling. Driving posture recognition by joint application of motion history image and pyramid histogram of oriented gradients[J]. Advanced Materials Research, 2014, 846-847(12): 1102-1105.
- [15] YAN Chao, COENEN F, ZHANG Bai-ling. Driving posture recognition by convolutional neural networks[J]. IET Computer Vision, 2016, 10(2): 103-14.
- [16] Kaggle. State farm distracted driver detection[EB/OL]. [2017-11-13]. <https://www.kaggle.com/c/state-farm-distracted-driver-detection>.
- [17] CHOLLET F. Xception: Deep learning with depthwise separable convolutions[C]//IEEE Conference on Computer Vision and Pattern Recognition. Hawaii, USA: IEEE, 2017.
- [18] SZEGEDY C, VANHOUCKE V, IOFFE S, et al. Rethinking the inception architecture for computer vision[C]//Computer Vision and Pattern Recognition. Las Vegas, USA: IEEE, 2016: 2818-2826.
- [19] SABOUR S, FROSST N, HINTON G E. Dynamic routing between capsules[EB/OL]. [2017-10-26]. <https://arxiv.org/abs/1710.09829v2>.

编辑 漆 蓉