

• 第二十五届汽车安全技术国际学术会议专题 •

# 基于视觉的车内人员安全监测功能开发

曹立波<sup>1,2</sup> 龚溢鹏<sup>1</sup> 杨名海<sup>2</sup> 戴丽华<sup>2</sup> 朱李平<sup>2</sup> 陶强<sup>2</sup>

(1. 湖南大学, 汽车车身先进设计制造国家重点实验室, 长沙 410082; 2. 长沙立中汽车设计开发股份有限公司, 长沙 410205)

**【摘要】**为实现对分心驾驶和车内遗留情况的安全监测,分别采集了分心驾驶工况和车内乘员遗留工况数据,结合神经网络模型,采用端到端分心驾驶网络中的轻量化网络 MobileNet V2 实现分心驾驶识别,并通过串联 YOLO V5 目标检测器、DBFace 人脸检测器和 MobileNet V2 人脸分类器,设计了车内人员安全监测方案。针对 2 种车内人员安全监测任务分别设计实车试验验证方案的可行性,结果表明,所提出的方案在保证实时性的前提下具有较高的检测准确率。

**关键词:** 车辆安全 分心驾驶 车内遗留 神经网络

中图分类号: U463.6; TP391.41 文献标志码: A DOI: 10.20104/j.cnki.1674-6546.20230130

## Development of Health and Safety Monitoring Function in Vehicle Based on Vision

Cao Libo<sup>1,2</sup>, Gong Yipeng<sup>1</sup>, Yang Minghai<sup>2</sup>, Dai Lihua<sup>2</sup>, Zhu Liping<sup>2</sup>, Tao Qiang<sup>2</sup>

(1. State Key Laboratory of Advanced Design and Manufacture of Automobile Body, Hunan University, Changsha 410082;  
2. Changsha Lizhong Automobile Design and Development Co., Ltd., Changsha 410205)

**【Abstract】**Data on distracted driving conditions and passenger retention conditions in the car were collected, and combined with a neural network model, the lightweight network MobileNet V2 in the end-to-end distracted driving network was used to achieve distracted driving recognition. By connecting YOLO V5 target detector, DBFace face detector, and MobileNet V2 face classifier in series, an accurate health and safety monitoring plan was designed. Real vehicle test verification schemes were designed respectively, to verify the feasibility of two vehicle occupant health and safety tasks. The results show that the proposed scheme has high detection accuracy while ensuring real-time performance.

**Key words:** Vehicle safety, Distracted driving, Left in the vehicle, Neural network

**【引用格式】**曹立波, 龚溢鹏, 杨名海, 等. 基于视觉的车内人员安全监测功能开发[J]. 汽车工程师, 2023(6): 1-6.

CAO L B, GONG Y P, YANG M H, et al. Development of Health and Safety Monitoring Function in Vehicle Based on Vision[J]. Automotive Engineer, 2023(6): 1-6.

## 1 前言

数据显示,驾驶员注意力分散造成的交通事故占比极大,且逐年增加<sup>[1]</sup>。同时,儿童因被遗忘在车内造成中暑死亡的事件时有发生<sup>[2]</sup>。2017年9月,欧洲新车安全评鉴协会(Euro-New Car Assessment Program, E-NCAP)发布了2025路线图(Road Map 2025)<sup>[3]</sup>,将驾驶员监测(Driver Monitoring)和儿童存在检测(Child Presence Detection)分别列为初级安

全系统和第三级安全系统。E-NCAP鼓励车辆提供驾驶员监控功能和车内乘员遗留检测功能,并且将对配有这些功能的车辆予以奖励。

研究人员在分心驾驶方面开展了大量研究,其中针对驾驶员驾驶动作的研究较多,建立了很多分心驾驶动作识别数据集,如东南大学驾驶动作数据集(Southeast University Driving Posture Database)<sup>[4]</sup>、保险公司分心驾驶监测(State Farm Distracted Driver Detection)数据集<sup>[5]</sup>、开罗美国大学分心驾驶数据集

(American University in Cairo Distracted Driver's Dataset)<sup>[6]</sup>、Drive&Act 数据集<sup>[7]</sup>及多视角、多模式和多光谱驾驶员动作数据集(Multiview, Multimodal and Multispectral Driver Action Dataset, 3MDAD)<sup>[8]</sup>等,极大促进了分心驾驶研究的发展。但是,目前公开的且能够在真实应用场景中使用的数据集较少。

近年来,针对车内乘员遗留检测的研究较少,特别是对于车内遗留儿童检测的研究。Khamil等<sup>[9]</sup>使用负载传感器检测儿童座椅内是否存在遗留的儿童。Norhuzaini等<sup>[10]</sup>在后排座椅上方安装检测范围为37 cm的热传感器进行车内儿童的检测。现有研究大多使用射频(Radio Frequency, RF)信号检测儿童的生命体征,采用视觉方式进行车内乘员遗留检测的研究相对较少。

本文采集全天候的分心驾驶数据并进行相应处理,利用该数据集对分心驾驶动作进行分析,并使用卷积神经网络开展测试,同时,针对采集的车内遗留数据,提出一种用于视觉检测车内成人、儿童及宠物等生命体遗留的检测方案。针对分心驾驶识别和车内乘员遗留检测任务,分别对比选择最佳解决方案在实车上进行测试,并开发用于系统测试的用户界面。

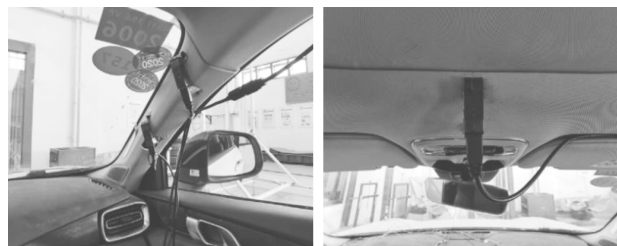
## 2 数据集制作

### 2.1 数据采集

针对国内现有的用于分心驾驶和车内遗留检测的数据集较少的情况,本文分别采集了用于分心驾驶和车内遗留检测的数据形成数据集。数据采集在实车内进行,将试验车辆布置在不同光照条件下,在驾驶员侧车窗范围内放置绿色幕布,以便后期替换车窗外背景,增强数据的多样性。使用Stellar 200 3D相机采集数据,该相机具有RGB和飞行时间(Time of Flight, ToF)摄像头模块,能够同时采集RGB、红外(Infrared Radiation, IR)和深度(Depth)数据,并且能够同时输出3种图像。该相机检测精度高、体积小、便于安装、价格低,且能够实时输出检测数据。为了能够较好地覆盖整车范围,本文同时使用2台Stellar 200 3D相机,分别布置在副驾驶员座椅一侧A柱上部和前排座椅上方车顶中心,不影响驾驶与乘坐,如图1所示。

为了更好地模拟车辆驾驶工况,本文分别采集了白天和夜晚的车内人员数据,数据采集频率为15 Hz,

每个相机输出RGB图像(分辨率为640×480)、深度图像(分辨率为240×180)和红外图像(分辨率为240×180),RGB图像和深度图像以8 bit数据的形式保存,红外图像以16 bit数据的形式保存,从而更好地保存原始的数据特征。8 bit图像采用视频录制的方式,同时保存对应帧的16 bit图像数据,使其能够与8 bit图像匹配。2台相机并非同步采集数据,本文对所有视频进行了时间戳上的对齐。



(a)安装位置1

(b)安装位置2

图1 摄像头安装位置示意

### 2.2 数据集相关指标

参与数据采集的志愿者包括37名成人和10名儿童,其中成人志愿者包括34名男性和3名女性。采集的数据包括37名驾驶员的分心驾驶数据、37名成人和10名儿童的车内乘员数据。为了保证试验数据的可处理性且便于在更多任务上使用,本文在数据采集前采取在墙上张贴标尺并拍照测量保存的方式对每个志愿者的体型进行测量。

为了保证驾驶动作的多样性,试验要求驾驶员按照指定要求分别做出本文所规定的10种驾驶动作,10种动作的选择参照美国汽车协会交通安全基金会提供的分散驾驶员注意力的潜在活动,且被美国交通部的致命事故报告系统数据库研究所证明。由于驾驶员的驾驶习惯不同,没有对驾驶员的动作标准进行规定,完全模拟真实的驾驶场景,后排乘员的数据采集同时进行,本文没有要求乘员做出指定的动作,乘员可以根据各自的乘车习惯进行试验。

为了更贴近真实驾乘状况,本文限制了后排乘员的人数。同成人一样,在车辆后排随机安排多名儿童,做出任意动作,为保证儿童安全,每次儿童数据采集都安排家长进行监护。

### 2.3 数据处理

#### 2.3.1 分心驾驶数据

本文对采集到的数据进行逐帧标注,针对每一个类别按照指定的帧数间隔提取图片,并且将规定之外的动作额外划分为一个类别。获得的各类别

图片数量如表1所示。

类别	描述	数量
C0	安全驾驶	3 025
C1	左手操作手机	2 590
C2	左手接打电话	2 291
C3	右手操作手机	2 399
C4	右手接打电话	2 334
C5	调节收音机和仪表盘	2 085
C6	喝水或饮料	2 625
C7	整理头发或化妆	1 840
C8	伸手到后排取东西	1 879
C9	与其他乘员交谈	2 013

本文按照9:1的比例划分训练集和验证集,以便后续开展分心驾驶识别试验,数据集图片示例如图2所示。



图2 分心驾驶数据集图片示例

### 2.3.2 车内乘员数据

车内乘员数据包括成人和儿童的二维和三维数据。数据采集的摄像头位置固定,因此图像的背景不会发生变化。车内乘员数据集的可变量主要为乘员的数量和体型,当前先进的人员识别网络能够很好地处理这一变化,同时可以结合深度数据对是否存在乘员进行判断,从而进行占据物体检测。此外,采集的车内乘员数据能够很好地记录车内乘员的特征,可用于乘员的人脸检测和人体姿态检测等。

### 2.4 背景干扰

RGB图像的采集时间和场景有限,不能很好地反映真实环境下车辆的工作状态,因此本文采用背景去除的方法手动实现驾驶场景的多样化。摄像头固定后,车辆行驶时,拍摄背景中只有车窗外的环境是变化的,故以车窗外的图像作为变量,利用背景叠加法,通过变换车窗外的环境来增加数据样

本多样性的。车窗部分背景通过绿幕去除获得,利用白色部分生成掩码(Mask),如图3a所示;在原始图像上叠加背景图像以模拟车辆在不同场景中的数据收集状态,如图3b所示。



(a)背景替换前 (b)背景替换后

图3 背景替换前、后效果

### 2.5 RGBD图像和点云数据的生成

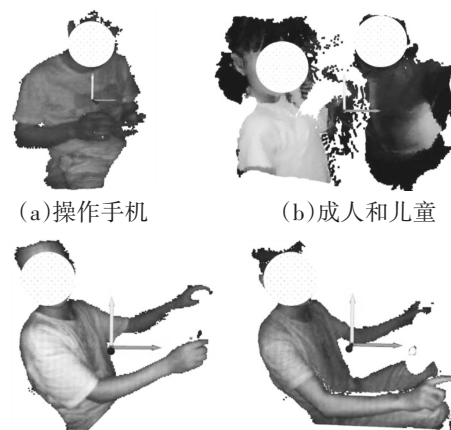
本文所采集的深度数据能够反映摄像头与物体的实际距离,利用摄像头的内部参数可以计算出物体的三维空间坐标,进而生成采集图像对应的RGBD图像数据和点云数据。RGB图像和深度图像的分辨率不同,故本文首先将RGB图像和深度图像进行对齐,并调整RGB图像的分辨率为 $240 \times 180$ ,最终得到RGBD图像和点云图。

相机坐标系到像素坐标系的转换关系为:

$$s \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \quad (1)$$

式中, $\begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix}$ 为相机内参矩阵; $(x,y,1)^T$ 为像素坐标系下的坐标; $(X,Y,Z)^T$ 为相机坐标系下的坐标; $s$ 为尺度因子; $f_x, f_y$ 分别为 $x, y$ 方向的焦距; $c_x, c_y$ 为相机成像的主点在像素坐标系上的 $x, y$ 坐标。

利用式(1)和摄像头模组的内部参数,使用RGB图像和深度图像可以生成如图4所示的点云信息。



(a)操作手机 (b)成人和儿童  
(c)与其他乘员交谈 (d)调节仪表盘

图4 点云数据示例

## 2.6 数据集总览

针对分心驾驶和车内遗留检测,本文建立了一个全天候、多工况的车内人员数据集,包含驾驶员和后排乘员等数据,可供人体姿态估计、分心驾驶和车内人员检测等多个任务使用,如图5所示。

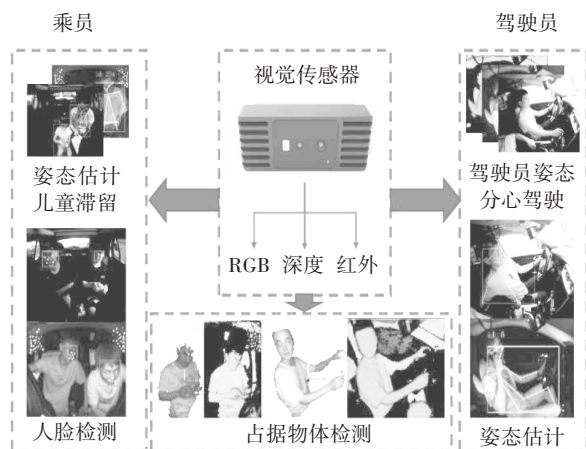


图5 车内人员安全监测系统

## 3 试验验证

### 3.1 分心驾驶识别试验

卷积神经网络能够提取图像的深层特征并得到特征的线性组合,实现对整幅图像的理解。对采集的数据集进行分析发现,驾驶员的特定分心驾驶动作存在一定规律,且分心动作出现的频率相对固定。为了验证数据集的有效性,本文利用经典的深度学习模型进行测试。为满足不同工况的应用需求,对数据进行多种方式的组合,利用不同模型进行交叉验证和测试。

#### 3.1.1 试验设置

本文选用经典的深度学习模型(AlexNet<sup>[11]</sup>、VGG<sup>[12]</sup>、ResNet<sup>[13]</sup>、MobileNet V2<sup>[14]</sup>)进行测试,选用不同的数据图像输入以适应不同场景的光照条件变化。试验使用处理后的数据集,且为满足不同工况,将不同的图像源进行分组,作为神经网络的输入。

#### 3.1.2 训练设置

训练利用Pytorch深度学习框架在2块RTX Titan X显卡上进行,训练基本参数设置为:输入图片分辨率为224×224,训练周期为200个周期(Epoch),批次大小(Batch\_Size)为32张,初始学习率 $l_r=0.01$ ,选用Adam作为优化器,选用交叉熵作为损失函数。不同模式下模型在测试集上的分类准确率如表2所示。

表2 测试集分类准确率

网络模型	RGB	红外	深度	深度+红外
AlexNet	94.2	93.7	92.6	93.8
VGG16	96.4	95.8	95.2	95.8
ResNet34	98.2	97.5	96.9	97.2
MobileNet V2	95.7	94.6	94.2	94.5

### 3.1.3 结果分析

数据集包含白天和夜晚的分心驾驶数据,能够满足不同光照条件下的实际应用。针对不同光照条件,选取不同的图像作为输入检验驾驶员分心检测的效果。由表2可知,ResNet34的准确率最高,但模型的数量和计算量较大。由于最终要在边缘计算设备上运行,因此选用参数数量和计算量较少的MobileNet V2作为算法骨干网络,融合深度图像和红外图像信息作为输入图像源。

### 3.1.4 检测结果

本文对MobileNet V2作为算法骨干网络的RGB图像检测结果进行了可视化,其结果如图6所示。



(a)安全驾驶

(b)左手操作手机

图6 分心驾驶检测结果

### 3.2 车内乘员遗留检测方法

为了验证车内乘员数据的实用性,本文利用人体检测、人脸检测、人体姿态检测方法对数据集进行测试,并对多种车内乘员遗留检测方案进行试验。

#### 3.2.1 点云占据物体检测

目前,对于车内乘员遗留的研究多基于非视觉传感器,本文通过视觉传感器采集的车内乘员数据可以不同的角度和方案实现车内乘员检测。停车后车内通常没有乘员,因此可以利用点云或深度图提取车辆的座椅背景,在车内有乘员的情况下,也可以通过点云或深度图像的差异得到车内遗留信息,如图7所示。利用点云进行聚类,区分不同个体并将不同个体的点云数据投影到二维图像上,得到对应的掩码,利用掩码裁剪出个体的图像范围,再针对不同的个体串联分类网络即可实现对成人、儿童及宠物的检测。



(a)原始点云 (b)去除背景后点云

图7 去除座椅等背景前、后点云数据

### 3.2.2 视觉检测方案

考虑到运用点云数据进行特征提取的复杂性,本文提出利用视觉方式检测成人、儿童以及宠物遗留的方案。摄像头模组获取图像后输入目标检测器,首先对视野范围内进行检测,如果检测到乘员,再利用串联的人脸检测器检测人脸图像,进而采用年龄分类算法进行分类,区分成人和儿童,检测方案流程如图8所示。

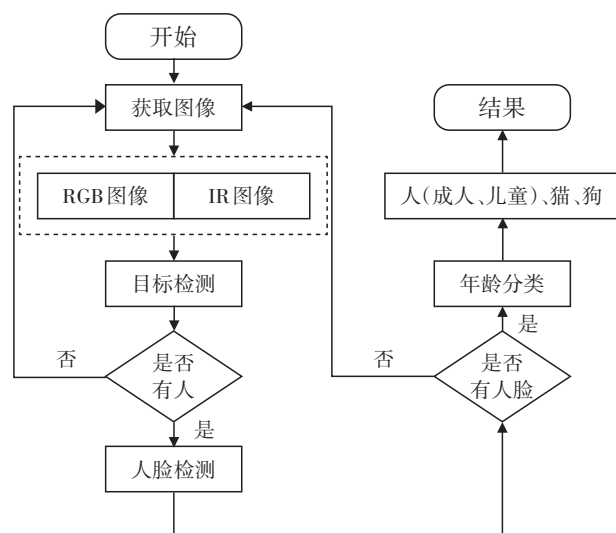


图8 检测方案流程

### 3.2.3 视觉检测方案试验结果

本文采用检测成功率对车内遗留检测任务进行评价:

$$P=R/A \quad (2)$$

式中, $P$ 为检测成功率; $R$ 为检测成功次数,本文将检测对象的位置和类别均正确视为检测成功; $A$ 为总检测次数。

本文进行了多次测试,测试在视频流上进行,试验的平均结果如表3所示。

表3 车内遗留检测结果

参数	人	成人	儿童	成人+儿童
正确次数/次	45	45	42	40
总次数/次	50	50	50	50
检测成功率/%	90	90	84	80

本文在实车场景下进行了多次试验,试验中摄像头布置在前排座椅上方车顶中心位置,试验结果与表3的结果相近,略有波动。

### 3.2.4 视觉检测方案检测结果

由于第2种视觉方式检测方案的实现过程较为简单,最终选作车内乘员遗留的检测方案,目标检测器选用YOLO<sup>[15]</sup>系列中的YOLO V5,人脸检测算法选用DBFace模型,年龄分类网络选用MobileNet V2网络,最终的检测结果如图9所示。



(a)检测结果1

(b)检测结果2

图9 车内遗留检测结果

## 4 测试界面可视化

针对车内人员安全的分心驾驶和车内遗留检测功能开发了测试界面,如图10和图11所示。



图10 分心驾驶测试界面



图11 车内遗留测试界面

## 5 结束语

本文建立了多模式、多工况的分心驾驶数据集和车内遗留人员数据集,选取MobileNet V2作为分心驾驶检测算法,在满足实时性的条件下达到了95.7%的检测准确率,实现了真实场景下的分心驾

驶识别,同时,设计了一种基于视觉的车内遗留检测方案,实现对车内成人、儿童及宠物的识别,检测成功率高达90%。结合本文所开发的测试软件,对车内人员安全监测2个任务进行了测试,结果表明,本文所提出的方案能够满足实际使用需求。

## 参 考 文 献

- [1] Center for Disease Control and Prevention. Distracted Driving [EB/OL]. (2022-01-15)[2023-05-05]. [https://www.cdc.gov/transportationsafety/distracted\\_driving/](https://www.cdc.gov/transportationsafety/distracted_driving/).
- [2] 刘丽萍,鞠伟男,王诗萌,等. 车内儿童存在探测技术及测评方法研究[J]. 汽车文摘, 2022(11): 39-44.  
LIU L P, JU W N, WANG S M, et al. Research on the Technology and Testing Method of Child Presence Detection [J]. Automotive Digest, 2022(11): 39-44.
- [3] Euro NCAP. Euro NCAP Launches Road Map 2025 - In Pursuit of Vision Zero[EB/OL]. (2017-09-12)[2023-04-03]. <https://www.euroncap.com/en/press-media/press-releases/euro-ncap-launches-road-map-2025-in-pursuit-of-vision-zero/>.
- [4] ZHAO C H, ZHANG B L, ZHANG X Z, et al. Erratum to Recognition of Driving Postures by Combined Features and Random Subspace Ensemble of Multilayer Perceptron Classifiers[J]. Neural Computing and Applications, 2013 (22): 175-184.
- [5] State Farm. State Farm Distracted Driver Detection[DB/OL]. <https://www.kaggle.com/datasets/rightway11/state-farm-distracted-driver-detection>.
- [6] ABOUENAGA Y, ERAQI H M, MOUSTAFA M N. Real-time Distracted Driver Posture Classification[C]// 32nd Conference on Neural Information Processing Systems (NIPS 2018). Montreal, Canada: Neural Information Processing Systems Foundation, 2018.
- [7] MARTIN M, ROITBERG A, HAURILET M, et al. Drive&Act: A Multi-Modal Dataset for Fine-Grained Driver Behavior Recognition in Autonomous Vehicles[C]// 2019 IEEE/CVF International Conference on Computer Vision (ICCV). Seoul, Korea: IEEE, 2019.
- [8] JEGHAM I, KHALIFA A B, ALOUANI I, et al. A Novel Public Dataset for Multimodal Multiview and Multispectral Driver Distraction Analysis: 3MDAD[J]. Signal Processing: Image Communication, 2020, 88.
- [9] KHAMIL K N, RAHMAN S I A, GAMBILOK M. Babycare Alert System for Prevention of Child Left in a Parked Vehicle [J]. ARPN Journal of Engineering and Applied Sciences, 2015, 10(22): 17313-17319.
- [10] NORHUZAINI A A, YAACOB M, NGAJIKIN N H. A System of Child Detection in Car Using Thermal Sensor[J]. Journal of Sustainable Manufacturing in Transportation, 2022, 2(1): 8-13.
- [11] KRIZHEVSKY A, SUTSKEVER I, HINTON G E, et al. ImageNet Classification with Deep Convolutional Neural Networks[J]. Communications of the ACM, 2017, 60(6): 84-90.
- [12] SIMONYAN K, ZISSERMAN A. Very Deep Convolutional Networks for Large-Scale Image Recognition[EB/OL]. (2015-04-10) [2023-04-03]. <https://arxiv.org/abs/1409.1556>.
- [13] 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.
- [14] SANDLER M, HOWARD A, ZHU M L, et al. Mobilenetv2: Inverted Residuals and Linear Bottlenecks[C]// 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, UT, USA: IEEE, 2018.
- [15] BOCHKOVSKIY A, WANG C Y, LIAO H. YOLOv4: Optimal Speed and Accuracy of Object Detection[EB/OL]. (2020-04-23) [2023-04-03]. <https://arxiv.org/abs/2004.10934>.

(责任编辑 弦 歌)

修改稿收到日期为2023年5月5日。