

# 基于深度学习与云计算的驾驶行为识别预警控制系统

作者：王子旭 许嘉晨 胡青娇 王婷 王磊

**摘要：**道路交通事故每年都给人民的生命财产带来巨大损失，根据事故调查显示，86%的交通事故是由驾驶人员的驾驶行为不当造成。驾驶行为识别预警控制系统基于深度学习、车联网技术，通过传感设备收集驾驶者面部表情、驾驶姿态、语音信息，并借助5G技术支撑的车辆网系统将这些样本上传至云端，利用深度学习技术评估驾驶者当前是否存在潜在的危险驾驶行为，并给以适当提示或采取强制措施，能有效降低因驾驶者自身原因带来的交通事故，同时结合云计算技术进一步降低人工智能对于硬件设备的限制。当一种风险驾驶行为被识别后，该系统将会立即发出警报，提醒驾驶员纠正驾驶行为，避免交通事故的发生。为了进一步降低硬件要求并达到识别系统越来越智能，搭建云端服务系统，实时地传输驾驶数据进入云端进行强化学习，以解决危险驾驶行为的个性化与普遍性的矛盾。

**关键字：**驾驶行为分析；深度学习；行为分析；车联网；人工智能；云计算；

## 1. 研究背景

2016年3月26日，我国正式出台《中共中央关于制定国民经济和社会发展第十三个五年规划的建议》，简称“十三五”规划，其中提到的完善现代综合交通运输体系，对着重提高交通安全和应急保障能力也有了进一步的强调。但是根据国家安全监管总局、交通运输部于2017年12月19日发布的最新研究报告显示，虽然近年来我国道路交通事故降幅明显，但依然高发。目前，我国道路交通事故年死亡人数仍高居世界第二位。报告中还指出2016年中国共接报道路交通事故864.3万起，同比增加65.9万起，上升16.5%。其中，涉及人员伤亡的道路交通事故212846起，造成63093人死亡、226430人受伤，直接财产损失12.1亿元。道路交通事故万车死亡率为2.14，同比上升2.9%。因此遏制道路交通事故高发、降低交通事故伤害仍然任重道远。并且据数据统计，80%的交通事故都是由于汽车驾驶员的不安全行为导致。因此遏制道路交通事故高发、降低交通事故伤害的重点在于矫正驾驶者的不安全驾驶行为。

## 2. 驾驶行为的分类与研究

根据我国道路交通事故的统计分析,占90%左右的事故主要由于驾驶人造成。驾驶人在开车过程中产生的情绪对交通安全具有重大的影响。在开车过程中,不稳定的情绪会导致驾驶人不能集中思想、妨碍驾驶人技术的正常发挥、驱使驾驶人不讲职业道德、让驾驶人感到疲劳、使驾驶人感情用事,很容易引发交通事故。驾驶人在悲哀、愤怒、恐惧、忧愁时,理智性和感受性将会降低,观察和思考能力差,在这种情感支配下驾车很容易发生

交通事故。心理学中把人们情绪的外部表现称为表情,按照表达情绪的身体部位可以将表情划分为面部表情、姿态表情、言语表情三种[1]。

驾驶员分心是驾驶员自愿或非自愿的转移驾驶人任务,而且转移任务是因为驾驶员执行另外的任务,而短暂的将注意力集中在与驾驶无关的时间、物体和人上,这种转变降低了驾驶员的环境意识、决策水平和操作技能。美国国家高速公路安全管理局(NHTSA)在实际的上路驾驶实验中发现近 80%的碰撞和 65%的临街碰撞的原因与驾驶非专注有关[2]。“路怒症”概念最早来自国外心理学。随着中国汽车数量增多,这一心理问题也受到国内越来越多专家的关注。“路怒”(road rage)是形容在交通阻塞情况下,开车压力与挫折所导致的愤怒情绪,发作者会袭击他人的汽车,有时无辜的同车乘客也会遭殃。“路怒症”现已成为当今世界的通病,2001 年到 2003 年年间,美国研究人员对 9282 位美国成年人进行了面对面调查,发现患“路怒症”的美国司机达 5%到 7%。其中公交车、出租车和长途车司机患这种心里疾病的比例更高,到 30%以上。每年 10%到 20%的车祸都与“路怒症”有关[3]。驾驶人疲劳时,会出现视线模糊、腰酸背疼、动作呆板、手脚发胀或有精力不集中、反应迟钝、思考不周全、精神涣散、焦虑、急躁等现象。如果仍勉强驾驶车辆,则可能导致交通事故的发生。据统计估计每年因疲劳驾驶引发的事故致 9 万多人死亡或重伤,而美国汽车协会(AAA)交通安全基金会的一项调查页表明了疲劳驾驶在美国的交通事故死亡事件中占据 21%的比例。因此本文将交通事故的主要成因分为非专注驾驶、愤怒驾驶、疲劳驾驶三个方面。

综合以上研究,从面部表情、姿态表情、言语表情三方面来检测驾驶员是否存在非专注驾驶、愤怒驾驶、疲劳驾驶的驾驶状态。当驾驶员处于非专注驾驶状态时,其表现有接听电话、吃喝东西、化妆等;当驾驶员处于愤怒驾驶状态时,常表现为双目圆瞪、咬牙切齿等;当驾驶员处于疲劳驾驶状态时,常常会频频打哈欠、闭眼、频繁打盹。因此可将上述三种表情具体化为一些可检测的细节行为,其分类如图 2.1

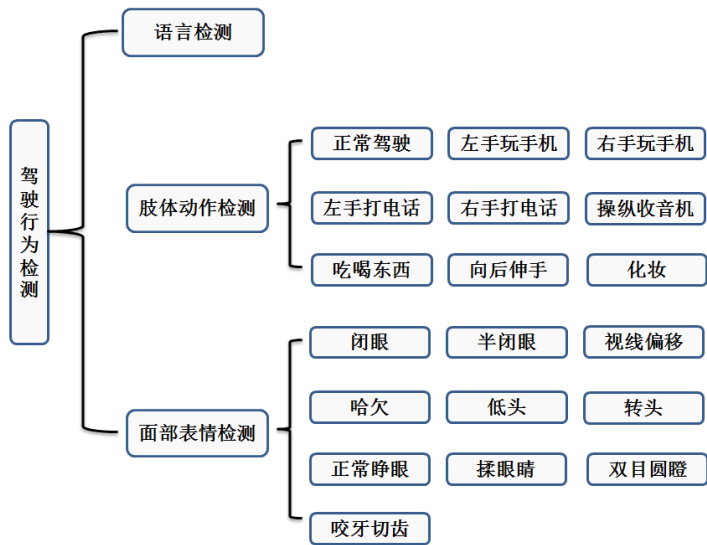


图 2.1 三种表情的具体分类

---

### 3.相关研究发展

目前国内外众多学者分别在不同的方面对驾驶状态进行了研究,并提出了一些相关的监测方法,主要有基于驾驶员生理参数的监测方法、基于车辆运行信息的监测方法以及基于驾驶员外部生理变化的监测方法三种。

#### 3.1 基于驾驶员生理参数的监测方法

测量驾驶员生理参数包括脑电图(Electro-Encephalogram, EEG)、心电图 (Electro-Cardiograph, ECG)、眼电图(Electro-Oculography, EOG)、脉搏跳动次数、呼吸频率、血压等等。在上述生理参数指标中,脑电图 EEG 和心电图 ECG 由于指标性好而成为该领域研究的热点。

Mark. A[4]研究分析了驾驶员在驾驶过程中注意力与脑电信号的关系后指出,注意力的变化趋势与脑电信号的变化趋势是同步的,从而得到了判断驾驶员状态的依据。国内学者王炳浩等[5]也做了相关方面的研究,将瞌睡时与清醒时的脑电波功率谱密度进行了对比,从而得到了判断驾驶员是否瞌睡的依据。Riemersma 等[6]指出经过长时间夜间驾驶的驾驶员心率有所下降,认为心率的变化对判断驾驶员是否处于疲劳状态具有一定的指导意义。郭士杰等[7]根据人体疲劳时呼吸频率降低的思想,设计了专门用于监测驾驶员呼吸频率的系统,以达到监测驾驶员是否疲劳的目的。

以上基于驾驶员参数的监测方法比较简单且容易实现,由于参数的客观性很强所以可以得到一个比较准确的检测结果。但是由于测量生理参数的仪器非常昂贵且必须与人体接触,属于侵入性的测量,会使驾驶员觉得不舒适而影响到驾驶行为,导致不安全驾驶,所以很少应用在实际中。

#### 3.2 基于驾驶员操作行为的监测方法

此类方法可以通过监测车辆硬件系统来监测驾驶员是否疲劳驾驶,如把握方向盘的压力或者位置变动、踩刹车或者操作档位的频率等等。Jarek Krajewski[8]指出通过监测驾驶员对方向盘的操作具有方法简单、性价比高的优势,文中通过提取方向盘的转角特征来判断驾驶员是否疲劳驾驶,并指出疲劳驾驶时的转角具有以下特征:转动频率较平时低甚至出现长时间维持在某个角度的现象,有时会发生快速的大角度修正行为。Eskandarian[9]也指出驾驶员疲劳驾驶时注意力分散,方向盘大幅度转动的频率明显提高。Federico Baronti[3]等人研究出一种由 16 个可以独立监测受力情况的小单元组成的传感器,用来实时测量驾驶员在驾驶过程中把握方向盘的力度大小,根据力度的变化程度来判断驾驶员的驾驶状态。

以上这种基于驾驶员操作行为的监测方法的优点是:仪器安装在汽车上,不与驾驶员直接接触所以不会使驾驶员产生不适感,得到的驾驶员操作行为的数据比较准确;缺点是这种得到操作行为数据很难度量驾驶员的疲劳程度,而且会受到车辆类型、驾驶员的驾驶习惯、路况条件、天气原因以及其他因素的影响而使得检测结果不够准确,因此

---

很少应用于实际中。

### 3.3 基于驾驶员外部生理变化的监测方法

此类方法使用摄像头实时获取驾驶员头部动作以及面部特征等外部生理变化，然后通过计算机视觉技术进行处理以达到监测驾驶员状态的目的。这种方法提取的驾驶员特征非常直观，而且属于非侵入性方法，不会对驾驶员产生影响，是当前该领域研究的热点方向。

文献通过获取驾驶员单位时间内眼睛闭合时间所占的百分率(Percentage of Eyelid Closure over the Pupil over Time, PERCLOS)来判断驾驶员是否疲劳驾驶，算法分为三个步骤，首先对人脸进行定位后进一步定位眼睛的位置并对眼睛进行跟踪，接下来继续准确定位瞳孔的位置，最后根据统计出 PERCLOS 分数来判断驾驶员的驾驶状态。文献提出了一种对人脸中的眼睛和嘴巴进行跟踪识别的疲劳检测系统，经过试验后证实此系统的准确性较高，有一定的实际应用价值。

## 4.项目实现

随着智能时代的来临，人工智能已经逐渐被全世界的各个行业所关注，计算机中的许多先进的人工智能技术已经被应用进了生活中的各个领域，例如智能家居，智能物流，智慧医疗等，先进计算机技术的引入能够使原本的领域被发掘出更多的可能性和在一定程度上提高效率。因此，将传统的交通行业与新兴的计算机技术结合，已经成为当代智慧交通的主流趋势。

### 4.1 研究目标与内容

本作品从当下驾驶员在驾驶过程中可能出现的不良行为作为出发点，综合考虑面部状态信息、肢体动作信息、语言表述信息等多个方向，利用机器学习算法以及深度学习模型对驾驶员的驾驶状态做出判断，以此对可能潜在的驾驶风险做出预警，避免危险的发生同时也矫正驾驶者的不良行为。

本次作品设计实现的主要功能和特色主要有以下几项：基于集成学习的驾驶者面部状态信息分析、基于 VGGNet 的驾驶者肢体动作分析、基于循环神经网络 LSTM 的驾驶者语言表述分析、智能预警防范系统。

#### 4.1.1 基于集成学习的驾驶者面部状态信息分析

所谓面由心生，就是在人心情愤怒时，面部表情会有很明显的改变，如：眉头紧皱、双目圆睁、咬牙露齿等；在人感到疲惫时，面部也会有所表现，如：眼睛半眯，频繁点头（打盹）等；同样，人在着急或者慌张时也会有不同的面部表现。



图 4.1 愤怒、疲惫时面部变化

当驾驶者驾驶车辆正常行驶时，通过高精度摄像头不断采集驾驶者面部图像，综合考虑驾驶者所表现出的状态细节，给予精确的状态分类，判断驾驶者是否处于疲劳驾驶、愤怒驾驶、分心驾驶等情况，为便于从图像中提取必要的细节并判断这些状态，选用卷积神经网络（CNN）模型做图像分析处理。卷积神经网络可以自动地从面部图像中提取高层抽象细节并进行精确分类，现已广泛应用与各种人脸检测与面部识别应用中。而且，一旦卷积神经网络提前训练完成，在使用过程中前向传播速度非常快，可以瞬间判断驾驶者当前所属状态，因此可以做到实时检测。

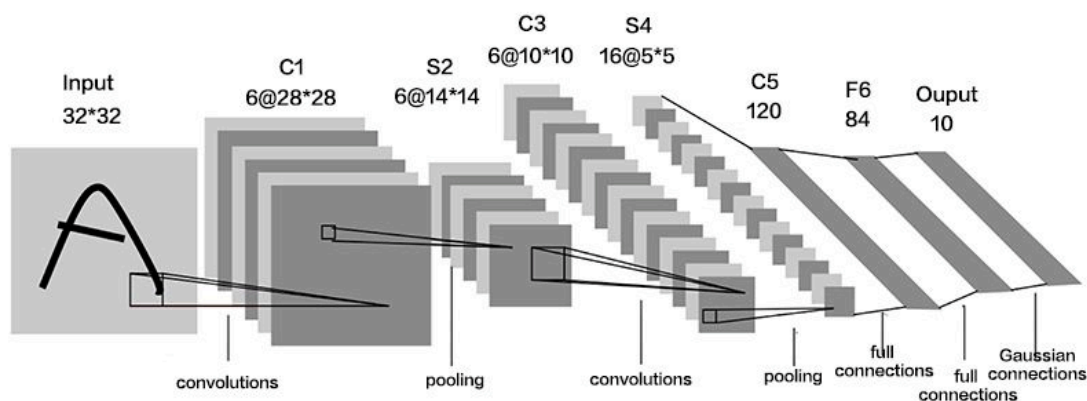


图 4.2 卷积神经网络（识别图片中手写字体）

考虑到 CNN 训练时速度的局限性，采用小型 CNN 模型，为弥补精度不足的缺陷，采用集成学习方法，同时训练多个小型 CNN，最终决策阶段所有神经网络共同参与，并采用多数表决法的到驾驶者状态分类，整个流程如下图：

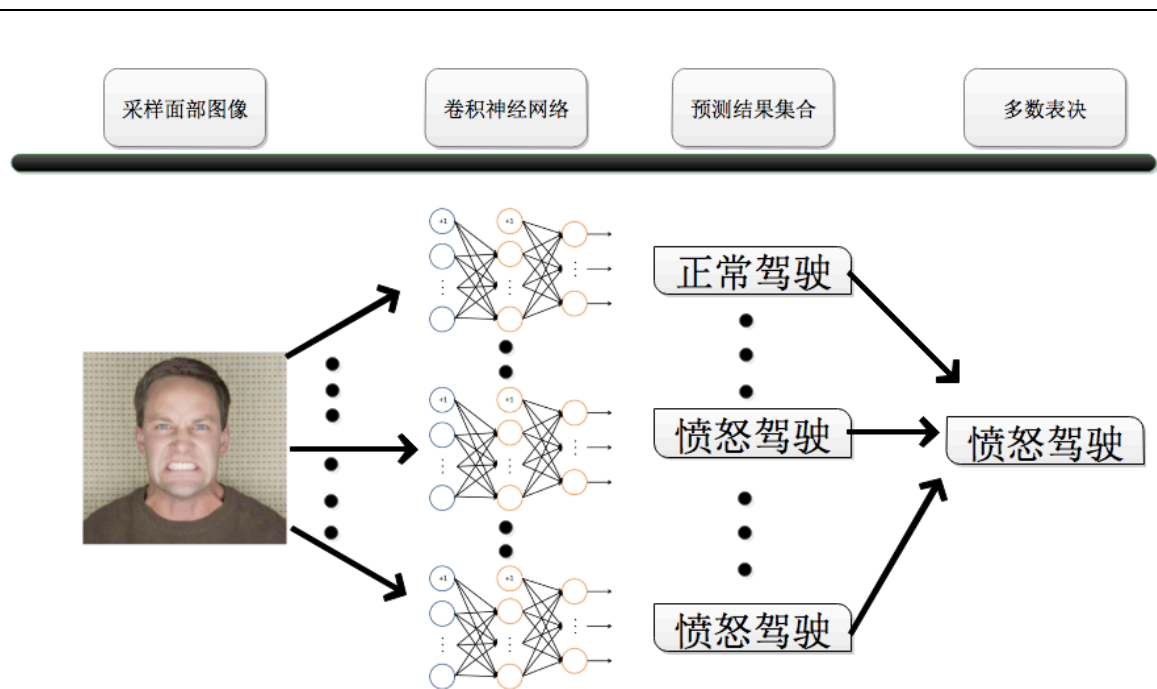


图 4.3 基于卷积神经网络的驾驶者面部状态信息分析

#### 4.1.2 基于 VGGNet 的驾驶者肢体动作分析

人处于愤怒、疲惫、焦虑时，除了面部表情会发生变化外，一些不经意的肢体动作也会随之表现出来。例如，人在紧张的时候可能会有频繁挠头、咬手、拍打方向盘等动作；在疲惫的时候可能频繁挥手打哈欠、揉眼...通过对这些肢体动作的捕捉，可以进一步精确驾驶者所处的状态。

同时，有一些驾驶不良行为非常危险，如驾驶中接打电话，向副驾驶、后排长时间转头等，通过安装在汽车上的高帧率运动摄像头，实时地采集驾驶者的行为，通过神经网络分析这些行为，当识别出存在安全隐患的驾驶行为时，提醒驾驶者修正驾驶动作，保持安全行驶的状态。







图 4.4 驾驶中的不良行为（上肢行为）

把驾驶行为分为十个类别，然后逐次将十种驾驶行为的图片进行分类学习，十种驾驶行为分类如下：

- 1) 正常驾驶
- 2) 右手玩手机
- 3) 右手打电话
- 4) 左手玩手机
- 5) 左手打电话
- 6) 调节收音机
- 7) 喝饮料、吃东西
- 8) 身体大幅度扭转
- 9) 整理头发、化妆
- 10) 转头说话谈论

从开源数据库中找到不同肤色、不同种族、不同体型、不同性别和不同年龄的数千名驾驶者，并对它们在驾驶过程中的各类行为进行标注后作为训练数据集。在训练时采用 VGGNet 卷积神经网络，并利用在计算机视觉系统识别比赛 ImageNet 中遇先训练好的网络作初始化，合理利用 CUDA 和 Fine-tuning 技术加速训练，最终在验证集上精度可高达 97.33%。

为了进一步降低错误率，在实际应用中每间隔一段时间，就连续在 4 秒内采样得到 12 张图像，送至训练好的 VGGNet 预测出驾驶者在这 4 秒内的 12 个驾驶状态，经过多数表决方案投票得出驾驶者当前最可能处于的驾驶状态，并传送到中央预警系统处理。

#### 4.1.3 基于循环神经网络 LSTM 的驾驶者语言表述分析

综合采用卷积神经网络模型和摄像头分别对驾驶者面部细节以及肢体动作做了分析预警，但仅仅从面部和动作这些“视觉”的方向是不足够判断一个人的驾驶状态的，言语也是很重要的一个部分。

据调查，有很高比例的驾驶者有“路怒症”，而且在驾驶过程中会出现自言自语的情况，可以从他们所说的话中提取有关信息，作为判断驾驶者是否出现异常情况的依据。人在愤怒的时候容易说脏话，在疲劳的时候可能会打哈欠发出“啊——”的声音，可以利用循环神经网络分析处理这些零散的语音信息。

首先，使用 CNN 实现的基于字的情感分析，并将其量化为一个确定的当前不正常状态值（可能有多个维度）。

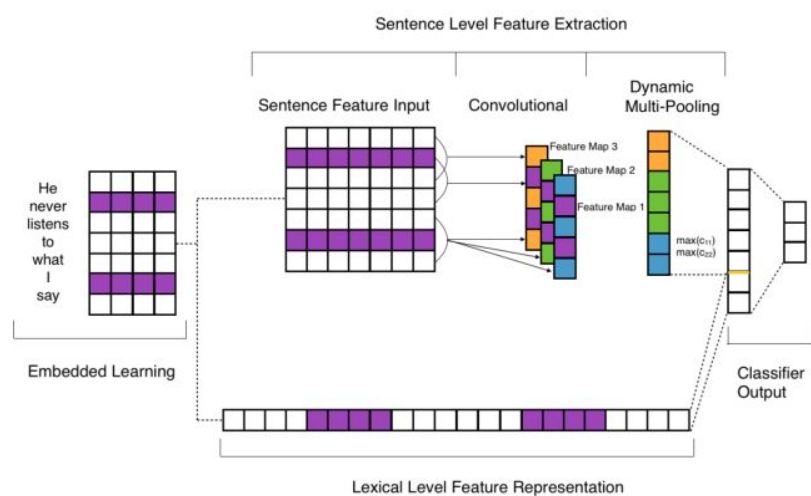


图 4.5 CNN 情感分析

得到每个时间节点的不正常状态值，可将其视为一个时间序列，并利用 LSTM 预测未来潜在的危险，并作出相应的举措。LSTM（LongShort-TermMemory）是长短期记忆网络，是一种时间递归神经网络，适合于处理和预测时间序列中间隔和延迟相对较长的重要事件。

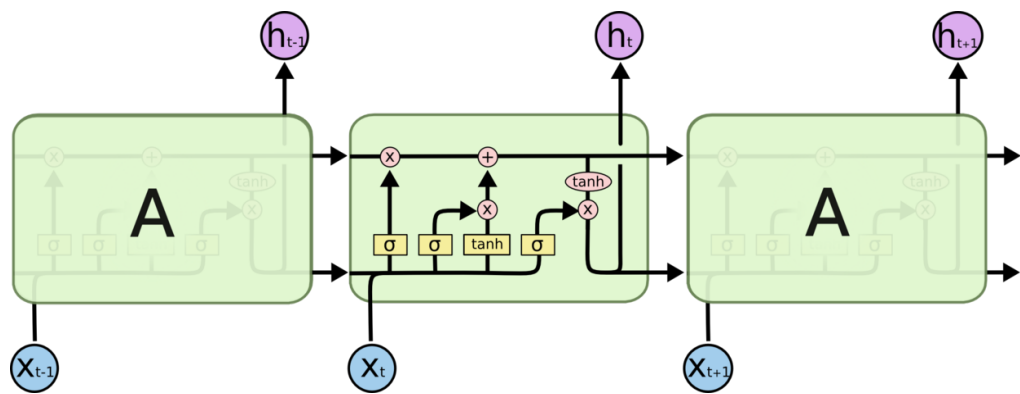


图 4.6 LSTM 网络结构

4.1.4 基于云计算的智能预警防范系统

直接在车内安装中央决策系统不仅价格昂贵，而且对硬件设备也有严格的限制，因此为保证系统的实时性与有效性， 只在车内安置传感器模块与通信组件，并借助 5G 技术带来的车联网系统，将传感器获取的图像与音频样本上传至云服务器，在云端利用预先训练好的多个神经网络完成上述三个模块的预测，并由云端搭建的中央决策系统分析驾驶者当前的驾驶状态，给出合理的决策并将结果加密后返回车内通信组件，通信组件根据云端的决策结果做出适当的预警提示。这样不仅能合理利用云计算的高效性，也能



借助 5G 车联网技术低延迟的实时性，及时发现潜在的驾驶风险并给以适当提示，协助其纠正驾驶行为。

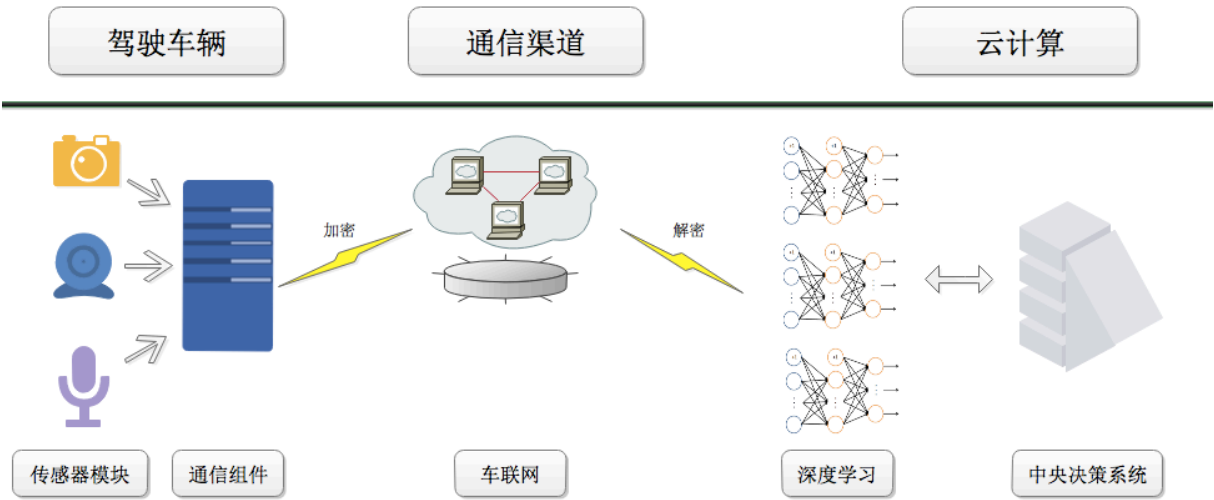


图 4.7 云计算架构

在云端采用上述三个模块的功能得到驾驶者当前的驾驶状态，根据驾驶状态的不同将进行合理范围内的智能预警与防范操作，目的是有效避免这些潜在的危险行而引起的风险，暂时将驾驶者驾驶状态分为以下几种：正常驾驶、疲劳驾驶、极度疲劳驾驶、愤怒驾驶、紧张驾驶、非专注驾驶、极度非专注驾驶。

文献[10]里的研究表明，驾驶人观察后视镜或仪表表等区域的时长如果小于 1.0s，则可认为驾驶者在获取与驾驶相关的必要信息，而当驾驶人视线离开前方区域的时间高于 1.0s 则认为此过程与正常驾驶过程无关，即分心驾驶，会对安全行车存在一定隐患。因此我们将 1.0s 作为系统进入警戒状态的时间节点。驾驶者驾驶分心时，通常会导致跑偏事故，引起交通事故的发生。长安大学汽车运输安全保障技术交通行业重点实验室[11]做了相关的研究，实验室选取了 12 名驾驶者在实际道路上进行行车实验，记录了 398 次分心过程的车辆运动数据。选取车道偏离的均值、方差和偏离程度来表征车道偏离的程度，分别对该 3 项统计指标与分心时长进行相关性检验和拟合分析，并按时长分组对比分心驾驶和正常驾驶时的三项该指标，得到了车道偏离随驾驶分心时长的变化关系。结果表明：驾驶分心过程中的车道偏离高于正常驾驶；车道偏离随着驾驶分心时长的增加而增大，并且增大速度越来越快；当分心时长超过 3.0s 时，车辆的车道偏离量在短时间内迅速增大。虽然该研究室研究的是驾驶分心时长对车道偏离的影响，但美国公路交通安全管理局[12]将分心行为分为视觉分心、听觉分心、操作分心和认知分心 4 类，因而疲劳驾驶和非专注驾驶也属于分心驾驶。所以，本文也选取 3.0s 作为一个分隔点，用于区别一般危险驾驶行为和极度危险驾驶行为。据生物学家研究，由于人的生理心理因素的限制，人对刺激的反应速度是有限的，一般条件下，反应时间约为 0.1~

---

0.5s。并且有资料显示，疲劳时调节时间可在 1.5s~2.0s。因而，驾驶者对语音提示的反应时间及调节时间最长为 2.5s。

当驾驶者车速高于某个速度阈值时（一般为正常路面行驶速度，为了排除车内休息、倒车入库等特殊情况）预警系统启动，持续启用三个模块监测驾驶者驾驶状态，并每隔一段时间对在这段时间内采取的样本进行分析，一旦其中一个模块出现异常情况，即当出现驾驶者出现危险驾驶行为，若持续 1.0s，系统进入警戒状态，在警戒状态中，采样频率升高并持续分析驾驶者状态，如果驾驶者立即恢复正常驾驶状态则退出警戒状态，若危险行为持续到 3.0s，系统开始判定驾驶者处于〔疲劳驾驶〕状态或〔非专注驾驶〕状态，系统开始做出相应提示，若提示后 2.5s 驾驶者仍未恢复到正常驾驶状态，系统开始判定驾驶者处于〔极度疲劳驾驶〕或〔极度非专注驾驶〕状态。

1) 当驾驶者处于〔疲劳驾驶〕时，可以激活驾驶座背后的震动器，震动座椅刺激驾驶者恢复正常状态，同时由语音系统给出“高亮男音”的“请集中注意力，谨防疲劳驾驶，如有需要，请休息片刻后继续驾驶”等语音提示，提示完成后可以播放一些摇滚音乐帮助驾驶者恢复状态。当给出提示后驾驶者仍然没有改观，便进入〔极度疲劳驾驶〕状态，此时将会采取强制降速、强制鸣笛的措施，避免出现危险。

2) 当驾驶者处于〔愤怒驾驶〕时，系统将会采取“女性化”声音安抚驾驶者情绪，例如“请不要这么生气啦，好好开车吧”，并播放一些柔和的轻音乐调节紧张的气氛。

3) 当驾驶者处于〔紧张驾驶〕时，系统会播放一些流行歌曲帮助驾驶者缓解气氛，并适当检测车速，控制车速不会突然大幅度升高。

4) 当驾驶者处于〔非专注驾驶〕时，系统会针对不同的行为作出不同的语音预警，这些行为包含：接打电话、长时间向后回头，长时间向两边转头，视线集中在下部等，会采用“高亮男音”语音包给出相应的提示。当给出提示后驾驶者仍然没有改观，便进入〔极度非专注驾驶〕状态，此时将会采取强制降速、强制鸣笛的措施，避免出现危险。



图 4.8 智能预警防范系统流程图

## 4.2 相关工作

### 4.2.1 驾驶者潜在危险行为造成的恶劣影响调研

驾驶者在行车过程中可能会出现具有潜在危险性的行为有多种，对各种类别的危险行为所可能造成的后果进行了相应的调研。

首先是驾驶者的疲劳驾驶，相比于其他一些行为，疲劳驾驶往往体现在其面部特征上，其表现有频频打哈欠，闭着眼睛，频繁打盹等，它往往会导致驾驶人疲劳时判断能力下降、反应迟钝和操作失误增加。而疲劳驾驶也分为不同的疲劳程度，当驾驶人处于轻微疲劳时，会出现换挡不及时、不准确；而驾驶人处于中度疲劳时，操作动作呆滞，有时甚至会忘记操作；驾驶人处于重度疲劳时，往往会下意识操作或出现短时间睡眠现象，严重时会失去对车辆的控制能力。

驾驶人疲劳时，会出现视线模糊、腰酸背痛、动作呆板、手脚发胀或有精力不集中、反应迟钝、思考不周全、精神涣散、焦虑、急躁等现象。如果仍勉强驾驶车辆，则可能导致交通事故的发生。据统计估计每年因疲劳驾驶引发的事故致 9 万多人死亡或重伤，而美国汽车协会(AAA)交通安全基金会的一项调查页表明了疲劳驾驶在美国的交通事故死亡事件中占据 21%的比例。同时深夜、凌晨或午后则被认为是“疲劳驾驶”事故高发期，因为驾驶者在该段时间下往往最容易犯困，出现疲劳的状况，但是许多驾驶者却总是过分的高估自己对紧急情况的处理能力以及自己的驾驶技术，置疲劳驾驶于不顾。而正是这样看似“不值一提”的疲劳驾驶却成为了高速公路上发生事故的最大起因，主要是由于高速公路的路面宽阔信号灯较少也不存在行人和非机动车的打扰，这样长期持续不

---

变的驾驶便会给驾驶者造成疲惫的状态，从而造成疲劳驾驶事故的发生。

其次是单手操控方向盘，其常常表现在其肢体上。行车中有些驾驶者习惯开车抽烟或把手放在窗外，仅用右手控制，这样很容易出意外。因为车辆在快速行驶或遇到紧急情况时，单手对汽车的操控能力不及双手的一半。除此之外，由于现在智能手机的普及，以及其应用的多样化，许多驾驶者都深陷“手机综合症”，尤其在开车过程中常常会觉得枯燥又无聊，很多人都忍不住打开手机玩起了微信，语音群聊已是家常便饭，甚至不少人还会选择刷刷朋友圈。

据验证，瞟一眼短信并大致看清内容，最快也需要 2 秒。当驾驶者以 60km/h 车速行驶时，2 秒钟的时间，车辆已经开出了 33.3 米。如果期间发生点意外，后果将不堪设想，这还仅仅是看短信，如果是在发短信、聊 QQ、查地图等情况下，需要投入更长时间注意力，从而可能会造成更大的安全隐患。还有就是在行车过程中接打电话，虽然上述行为被交通法规明文禁止了，但是在行车过程中接打电话的场景绝对是屡见不鲜，接打电话的影响不仅仅是使驾驶者只能使用一只手开车，更会转移驾驶者的注意力使其不能够专注于处理行车过程中遇到的事情。

据统计估计，开车时打电话，驾驶者会精力分散，妨碍其对路面情况和周边环境的观察，一旦遇到紧急或突发情况，将会大大消弱驾驶者的应变能力和反应时间，极易造成交通事故。英国一家交通研究所最近公布的测试结果表明，开车时使用手机，大脑的反应速度比酒后驾车时慢 30%，年轻人驾驶车辆接打手机相当于 70 岁老人驾驶车辆，开车打手机发生车祸的风险比正常驾驶时高 4 倍以上，有 70% 的致命事故是司机注意力不集中造成的。还有研究表明，开车打手机导致司机注意力下降 20%，如果通话内容重要，注意力甚至下降 37%。同时，拨打手机的驾驶员行车速度比正常状态慢 9%，刹车的反应速度也要慢 19%，不单单会影响到自己的驾驶安全，还会影响到整个交通状况，造成交通的阻塞。

最后则是“路怒症”。“路怒症”概念最早来自国外心理学。随着中国汽车数量增多，这一心理问题也受到国内越来越多专家的关注。“路怒”(roadrage)是形容在交通阻塞情况下，开车压力与挫折所导致的愤怒情绪，发作者会袭击他人的汽车，有时无辜的同车乘客也会遭殃。医学界把“路怒症”归类为阵发型暴怒障碍，指多重的怒火爆发出来，猛烈程度叫人大感意外，因此其也被认作是一种心理障碍。路怒症发作的人经常会口出威胁、动粗甚至毁损他人财物，也就是攻击性驾驶。根据我国的一项调查显示，长途汽车司机心理障碍发生率高达 80%，私家车主为 44.4%。而“路怒症”对应的表现的最多的就是在面对一些交通状况下爆粗口。爆粗口这样的行为代表着驾驶者当前的心理状态并非出于一个平静的状态下，而相应引起的攻击性驾驶行为，例如强行变道，是严重的交通违法行为，扰乱交通秩序、危害交通安全。而根据调查表明，在当下这样快节奏的生活中这一交通陋习常见多发，严重的妨碍汽车社会健康发展，造成了很多出人意料的安全事故。

上述三种就是当前驾驶者当前在行车过程中常常会出现的具有潜在危险性的行为，

而利用现代计算机技术所设计的智能预警防范系统也正是被设计来用作检测和处理驾驶者所显现出得上述三种危险驾驶行为。

4.2.2 卷积神经网络的分析与选取

1).神经网络模型

神经网络是一种机器学习模型。从感知机发展到现在的深度学习，都是使用的神经网络这个基本模型，神经网络由很多节点和边组成，边连接节点并具有一个权值，每个节点具有一个激活函数，模拟神经元的兴奋与抑制状态，整个神经网络可以看作对人体神经系统的一个模拟，从得到的刺激（输入）可以不断调整自身权重从而达到期望的输出，神经网络模型训练的过程其实就是权重调整的过程，最终训练好的神经网络模型可以直接对输出做出判断，如同人体神经元的机制。神经网络由输入层、输出层、隐藏层组成，隐藏成很多的神经网络一般称为深度神经网络，在此之上扩展近似于基本的深度学习模型。

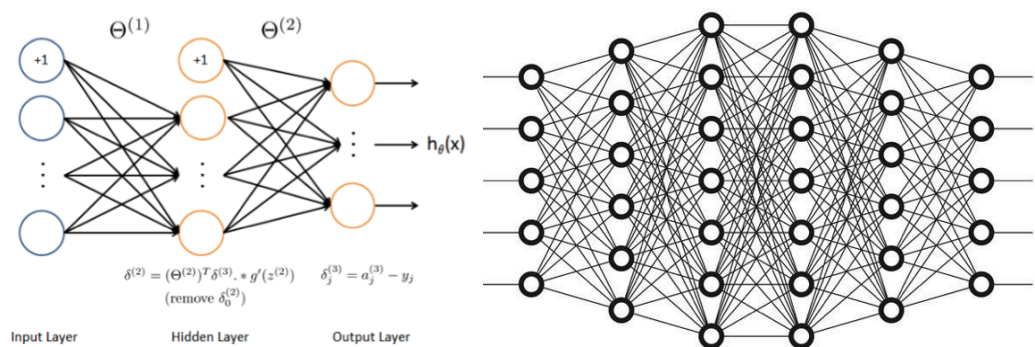


图 4.9 单隐层神经网络模型与深度神经网络模型

2).卷积神经网络

在 2012 年计算机视觉大赛上，AlexNet 卷积神经网络获得惊人的准确率，斩获比赛冠军，卷积神经网络和深度学习瞬间涌入学术、工业界，并持续称霸图像识别与人脸检测领域。卷积神经网络有效避免了因层次太深而带来的计算复杂问题，同时利用仿生的卷积层对图像进行处理，可以逐渐提取图像中的细节特质，因此在图像处理上有着天然的优势，有关卷积神经网络的发展也是如火如荼。

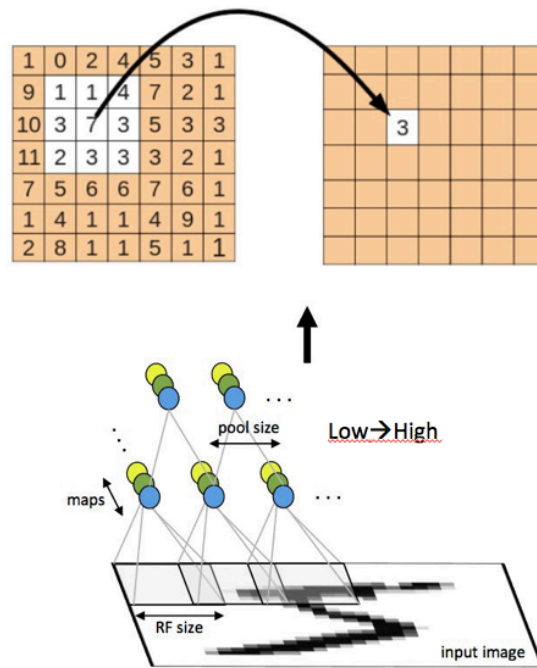


图 4.10 卷积神经网络剖析

### 3).Bagging 集成学习方式

由于计算成本和手中设备的限制，无法使用 AlexNet 或 VGGNet 这类层次很深的神经网络，因此采用集成学习的方法，集成多个小型卷积神经网络进行学习，并在决策阶段采用多数表决制得到预测的分类。

在机器学习的监督学习算法中，目标是学习出一个稳定的且在各个方面表现都较好的模型，但实际情况往往不这么理想，有时只能得到多个有偏好的模型（弱分类器，比如在此选取的小型卷积神经网络，它们可能在某些方面表现的比较好）。集成学习就是组合这里的多个弱学习器以得到一个更好更全面的强监督模型，集成学习潜在的思想是即便某一个弱分类器得到了错误的预测，其他的弱分类器也可以将错误纠正回来。

Bagging 是 bootstrapaggregating 的简写。bootstrap 也称为自助法，它是一种有放回的抽样方法，目的是为了得到统计量的分布以及置信区间。具体步骤如下：

- （1）采用重抽样方法（有放回抽样）从原始样本中抽取一定数量的样本
- （2）根据抽出的样本计算想要得到的统计量  $T$
- （3）重复上述  $N$  次（一般大于 1000），得到  $N$  个统计量  $T$
- （4）根据这  $N$  个统计量，即可计算出统计量的置信区间

在 Bagging 方法中，利用 bootstrap 方法从整体数据集中采取有放回抽样得到  $N$  个数据集，在每个数据集上学习出一个模型，最后的预测结果利用  $N$  个模型的输出得到，具体地：分类问题采用  $N$  个模型预测投票的方式，回归问题采用  $N$  个模型预测平均的方式。本项目中用到的是多个 CNN 的 Bagging 集成。



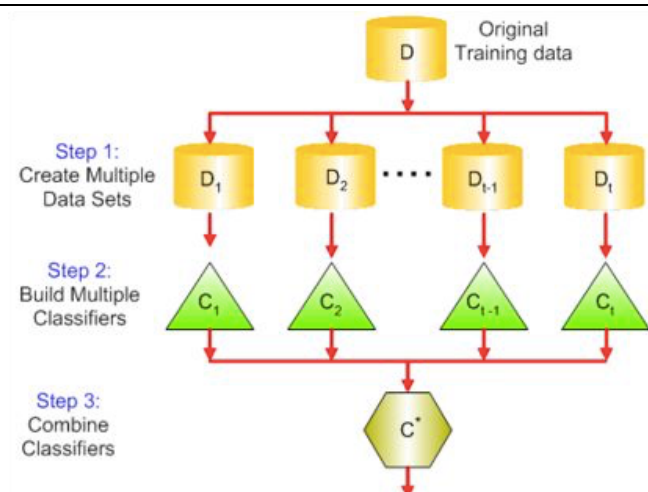


图 4.11 Bagging 集成学习方法

### 4.2.3 VGGNet 的分析和选取

#### 1) .Fine-tuning 与迁移学习

由于设备性能一般，无法训练高复杂度的神经网络，所以在 ImageNet 竞赛中预训练好的 VGGNet-16 模型中三个全连接层去除，只留下卷积层和池化层，并将它们全部冻结，在训练过程中不发生改变。因为据研究得知卷积层的功能主要是提取图像中的高阶特征，而不会对分类产生影响。因此仅仅加入自己定义的全连接层做图像分类。最终在预训练网络后添加一个包含隐藏层的全连接网络，得到十种驾驶行为的输出。这样只需要训练一部分神经网络而不是从头开始训练，合理的弥补了硬件设备的限制。

#### 2) .训练网络

将训练轮数设置为 30，学习速率的数值设置为 0.001，每次都把准备好的图片数据分批输入。在通过 NVIDIA GeForce 显存核心计算运行 30 小时之后，最终 Loss 函数值稳定在 0.023 后，可以认为训练已经完成，此时结束训练。

#### 3) .检验结果

模型训练完毕之后，在验证集上精度可以达到 99.73%，证明模型训练成功，参数选取合理；在 50,000 张全新的样本上做了测试，并得到精度约等于 90%，可认为模型在特定情况下适用。

#### 4) .适用场景微调

之前得到的模型适用于大多情况，但对于不同的车辆型号，摄像头的放置位置不同可能导致拍摄到的图像角度略有不同，因此在不同型号的车上重新获取图像进行新一批的训练，使得不同模型可以适配不同车辆型号，以降低车辆不同而带来的精度损失。

#### 5) .简单预测

现在就可以使用初步训练好的神经网络对校验集中的图片进行识别，先简单输入 10 张图片进行检验。这 10 张图片正确的驾驶行为编号分别是[0 7 8 2 5 3 9 4 1 6]，而通过神经网络识别的结果是[0 7 8 2 2 3 9 4 1 6]，其中只有 1 张识别错误，训练效果还是比较

理想。

4.2.4 LSTM 分析与应用价值

人类并不是每时每刻都从一片空白的大脑开始他们的思考。在你阅读这篇文章时候，你都是基于自己已经拥有的对先前所见词的理解来推断当前词的真实含义。不会将所有的东西都全部丢弃，然后用空白的大脑进行思考，人的思想拥有持久性。

传统的神经网络并不能做到这点，看起来也像是一种巨大的弊端。例如，假设你希望对电影中的每个时间点的时间类型进行分类。传统的神经网络应该很难来处理这个问题——使用电影中先前的事件推断后续的事件。

RNN 解决了这个问题。RNN 是包含循环的网络，允许信息的持久化。LSTM 是一种 RNN 特殊的类型，可以学习长期依赖信息。在很多问题上，LSTM 都取得相当巨大的成功，并得到了广泛的使用。

在对行驶过程中驾驶者的状态的研究中发现，这是一个随着时间变化的连续过程，即一个时间序列，并且过去所产生的情绪会对之后的状态产生影响，因此可以合理高效地 LSTM 来完成预测。

5. 项目的技术路线、解决的问题及成果

5.1 系统技术路线

本系统分为三个部分：样本采集部分、样本处理部分、智能预警系统部分。

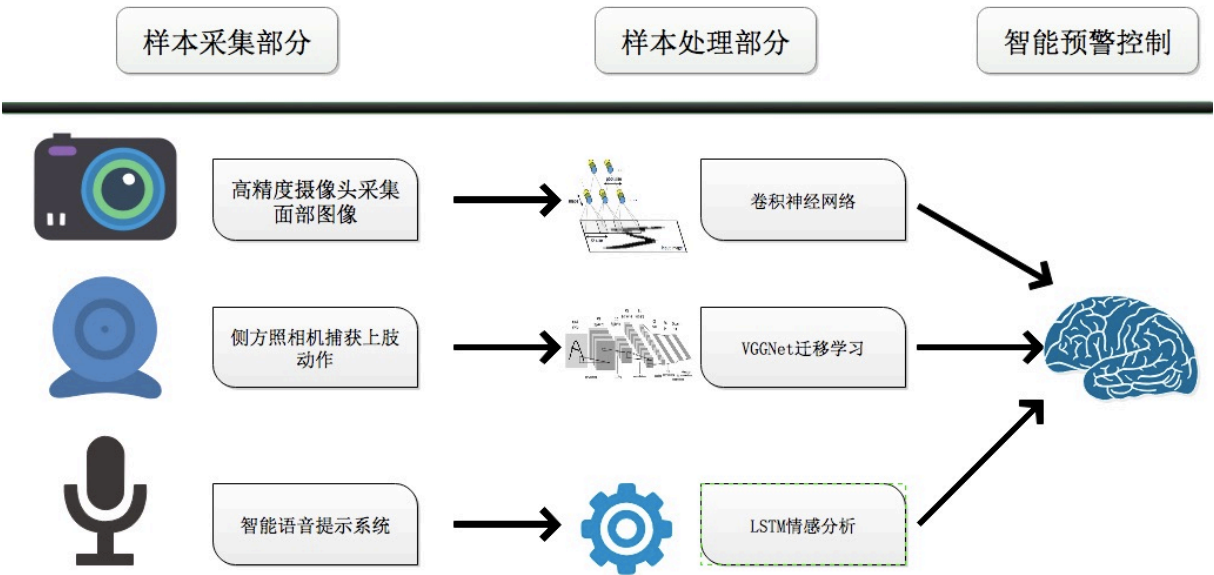


图 5.1 系统架构

5.1.1 样本采集部分

样本采集部分的主要功能是对驾驶者驾驶状态做采样，具体包括利用高精度摄像头

对驾驶者面部图像和肢体动作的采样、利用麦克风实时获取驾驶者的语音信息两个模块。

1) 为了得到驾驶者驾驶过程中肢体上是否有不良行为以及驾驶者面部表情, 选取高清高速摄像头(KS4A986), KS4A986摄像头分辨率可达到 $1920 \times 1080$ , 帧率为60fps, 可以实现运动抓拍和高速录像。因为面部识别和肢体识别所需采集的图片精度不一样, 因而选用两个摄像头, 一个用于采集面部信息, 一个用于采集肢体信息。

2) 对于语音样本的采集, 选用普通麦克风即可。

### 5.1.2 样本处理部分

1) 对于高精度摄像头获取到的面部图像, 首先对其预处理。考虑到不同车内背景不同、每个人衣着颜色款式也不相同, 不希望卷积神经网络在分析面部图像时把这些无关因素考虑在内, 因此首先用 OpenCV 工具对图片进行预处理, 扣除除了面部以外的背景, 将背景(包含衣着)全部作为黑色, 这样可以避免冗余特征的影响。

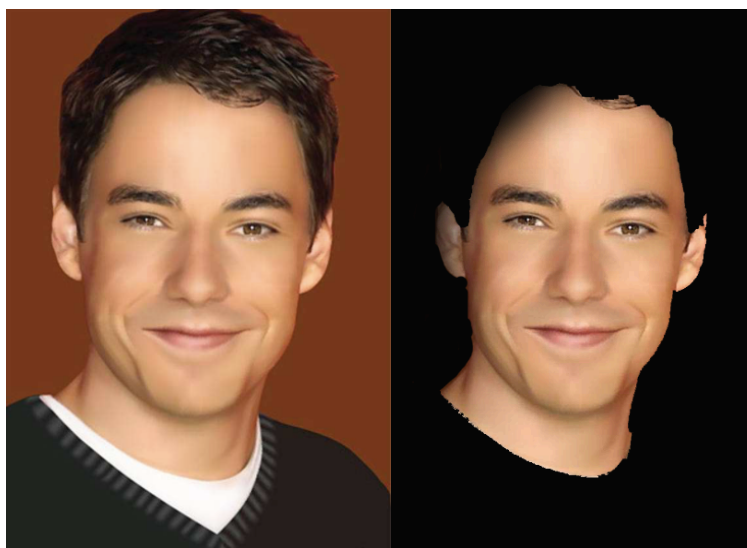


图 5.2 图像预处理示范(左图为原始图像, 右图为处理后的面部)

2) 当预处理完样本后, 将它们放入卷积神经网络中进行训练, 并分出验证集用来对模型参数进行调整, 分出测试集用来检测模型准确率, 当准确率不再上升时停止训练, 保存模型以便于在系统中使用。

3) 对于高帧率运动摄像头捕获到的肢体行为图像, 会每隔一段时间就在 4s 内连续采样 12 张, 将他们传入 VGGNet 得到判断的 12 个状态类别, 采用多数表决法得到当前最可能的驾驶行为, 如果系统判别为危险行为, 程序会把它们记录下来并告知预警系统处理。

4) 对于语音信息, 首先利用开源库将语音转变为对应的文字, 然后利用 CNN 完成情感分析并量化, 在此基础上用 LSTM 完成对可能情况的预测, 并作出相应预警。

5.1.3 智能预警部分

将上述三个模块的功能得到样本处理结果作为智能预警系统的输入，即驾驶者当前的驾驶状态，根据驾驶状态的不同将进行合理范围内的智能预警与防范操作，目的是有效避免这些潜在的危险行而引起的风险，暂时将驾驶者驾驶状态分为以下几种：正常驾驶、疲劳驾驶、极度疲劳驾驶、愤怒驾驶、紧张驾驶、非专注驾驶、极度非专注驾驶。

智能预警部分的中央控制系统会综合考虑驾驶者面部细节变化、上肢动作细节以及言语中的异常信息，合理做出相应的处理，当驾驶者车速高于某个速度阈值时（一般为正常路面行驶速度，为了排除车内休息、倒车入库等特殊情况）预警系统启动，持续启用三个模块监测驾驶者驾驶状态，并每隔一段时间对在这段时间内采取的样本进行分析，一旦其中一个模块出现异常情况，立刻进入警戒状态，在警戒状态中，采样频率升高并持续分析驾驶者状态，如果驾驶者立即恢复正常驾驶状态则退出警戒状态，如果驾驶者仍存在危险行为则对危险行为进行分类，并做出一些相应的处理

5.2 系统运行环境

5.2.1 硬件环境

系统设计和开发过程中考虑到可实现性和经济性，采用高清高速摄像头(KS4A986)。

5.2.2 软件环境

本系统的软件部分采用 Python 作为开发语言，并使用了 TensorFlow 框架。详细的运行环境信息见表 5.1。

表 5.1 系统运行环境

	名称	版本
开发语言	Python3	3.6
	JupyterNotebook	4.3
开发环境	Pycharm	2018.1
	UbuntuServer	18.04
运行环境	Windows/Linux/Ma	-
	cOS/Android/IOS	
数据库信息	MySQL	5.7.18

5.3 系统解决的问题与成果

将整个系统分为了三个模块来进行：调研模块、软件开发模块与软件测试模块。

1) 调研模块

在项目进展初期需要大量的调研工作，包括摄像头应装在车辆具体的哪个位置，车

---

内光线变化对摄像头采集到图像的影响，以及摄像头的种类和型号等等。在调研阶段队员相互合作，从互联网新闻和期刊中搜寻相关资料，并与指导老师积极沟通。

### 2) 软件开发模块

软件开发模块与调研模块同时进行，在调研模块进行的过程中，首先完成神经网络模型的搭建和初步测试，在确定数据格式之后，首先利用开源数据库做迁移学习，预先训练神经网络初始化权重，可以省下很大一部分的时间。在训练过程中，将完成对总控程序的设计与实现。

### 3) 软件测试模块

软件测试模块随着软件开发阶段一同执行，当任何部分的软件开发之后，小组成员对这部分软件进行综合测试，就出现的异常商量改进方向。

## 6.系统的功能特点及创新优势

随着人们生活水平的极大提升，汽车已经进入家家户户，而近年来交通事故的发生频率越来越高，驾驶者的安全问题已经成为了一个受到多方关注的焦点问题。针对此问题，一些汽车公司已经在想方设法让汽车碰撞时最大限度地减少人员损失，比如改变车的材质让车在碰撞时吸收更多的能量，制造独特的安全气囊减少车在翻滚时人脑部受到的伤害。但这些都是事故发生时的抢救，如果能在事故发生前就进行预警防范，就能把被动的抢救变为主动的预防，将会大大降低事故率。因此，对驾驶者的驾驶行为检测和预警成为目前研究的一个重点。

国外已有公司发布了自己的专利及产品。现代研发的一套用于防止疲劳驾驶的系统是通过监控驾驶员心率和车辆行驶状态。例如，当驾驶者心率低于正常值或者车辆忽左忽右地前行时，系统就会判定为疲劳驾驶并发出警告。目前现代尚未透露这项专利将合适进行商业化生产。丰田汽车选择通过监测驾驶员眼睛及面部表情来判定是否是疲劳驾驶。丰田旗下的高端车品牌雷克萨斯的一些车型，例如 LS600hL 可选配驾驶员注意力监视器，通过安装在转向杆周围的红外 LED 显示器来记录驾驶员的面部表情，如果出现不断眨眼、闭眼、视线偏离的情况，警报就会响起。大众则是通过驾驶员转向动作来确定是否处于疲劳状态。这个疲劳程度是取决于来自转向角传感器或电子助力转向系统的信息。根据这些信息，驾驶员疲劳探测系统会不断分析驾驶员的转向行为来确认其状态，计算出疲劳指数。如果指数超过一个特定值，且车速超过 65km/h，系统会判定驾驶员精力集中不够，有疲劳驾驶的迹象。可以看到，国外研发的这几款产品价格十分昂贵，目前只能配备在高端车型上，并且每款产品检测的角度都比较单一，单一的角度可能会对系统的准确度造成一定影响。但以上系统均只用于高端汽车，价格昂贵，不支持独立安装，不满足大众需求。

而本系统采用的硬件设备是两个普通的高精度摄像头，价格仅为 500 元左右，系统灵活小巧，摄像头仅为 500mm\*500mm\*200mm，方便安装，灵活可调，可使用多款车型，普通的车载电源 USB 接口就能驱动，不需要再独立添加驱动电源；系统工作环境为

---

-20~70℃，适应温度跨度大。通过智能驾驶监测预警系统能够较为精确地监测出驾驶者潜在的危险行为，并及时地提供一个友好的预警，在较大程度上给驾驶者提供一个安全舒适的驾驶环境，提高驾驶者的驾驶安全，而该系统的推广也势必能够满足各大厂商的需求。虽然本系统现在试用的环境是小汽车，但应用环境绝不止于小汽车，只要稍加调整，系统可以应用到货车、公交车、甚至是高铁火车。

该系统综合分析了有可能给驾驶者造成安全隐患的三大类驾驶行为，分别对其提供了检测预防的方法。首先是利用了 CNN 卷积神经网络提取驾驶者的人脸特征，并判断其是否正在处于一些疲劳驾驶等可能会造成极大安全隐患的状态下，然后利用了高精度摄像头，完成对驾驶者肢体行为的检测，判断其是否正在处于一些不安全的行为中，例如正在打电话或是四处张望，除此之外，还利用了 LSMT 循环神经网络用来进行语音识别分析，判断当前驾驶者是否处于一个不平稳的心态，防止其因此而出现一些过激的驾驶行为，影响到各方的行车安全。如果系统监测到了上述情况，将会自动发出对应的人性化且友好的预警，用来保证驾驶者的安全。

随着驾驶行为识别系统的不断扩张，势必将会对车载硬件的计算能力和学习能力提出更高的要求，而 5G 通信技术和普及将会很好的解决这个问题。5G 通信技术能够为未来生活带来超低延迟和更大数据量的实时传输，可以在移动端以按需、已扩展方式连接到远端的服务器，获得所需的资源。对于系统开发者来说，需要在云端搭建深度学习的网络。不断对用户产生的数据进行学习并返回学习结果，以实现对于复杂的、多元的驾驶行为进行不断的更新。同时在云端引入增强学习算法，满足系统越来越智能，及时发现危险驾驶行为打的前端行为，扼杀风险驾驶与萌芽，同时又能够有效控制风险驾驶行为的后果，杜绝交通事故的后患。

## 7.系统改进

系统目前使用的硬件是高速高清 1080P-60 帧、无畸变 100 度摄像头，适合运动抓拍和高清录像，对于光线充足、驾驶环境复杂度低的驾驶条件下可以实现精确的识别。如要普及这项系统，该系统需要适应不同天候、不同人群、多元的行为习惯、精确识别以及安全性和可靠性更高的目标，因此需要在以下几个方面进行更加深入的研究。

### 1) 适应多元环境

在未来的工作中，系统对环境的试用率会更高，不只是一般驾驶环境，系统还可以在强光、弱光、驾驶员佩戴眼镜墨镜口罩等复杂环境进行高精度的识别。在强光条件下，可以将目前的系统再辅助以抑制强光摄像头；在暗光环境下，将目前系统再辅红外线摄像头就可以进行准确的识别；在驾驶员佩戴眼镜墨镜口罩时，不断扩充深度学习的数据集，加强行为的识别精度。外部环境条件恶劣以及车体内环境复杂的情况下，需要对驾驶人的生理特点进行准确抓取，同时要应对随时可能的驾驶影响做出及时的反应，这都需要在系统算法上不断改进。

### 2) 适应多元人群



---

驾驶员的生理特点、行为习惯既有普遍性又拥有个体化差异，这将对系统识别驾驶人的情绪、行为的识别带来困难，使得系统在应对不同人群时可靠度不一致。在系统的总控制平台，要不断扩张驾驶行为的实验数据库，进行不断地测试和实验，更要开发一套全体驾驶员共享的学习库，已实现总体的智能化，不断识别和学习新的驾驶行为状态。在车辆内部系统平台，建立更加完善的个人驾驶行为学习和储存机制，记录和识别专属于该车辆的驾驶人行为属性，能够专一地服务于该车辆高频率的驾驶人群。

### 3) 提高系统可靠性

目前该系统所能够识别的行为数量有限，驾驶行为多属于普遍存在同时对引发交通事故概率更高的行为和情绪，而且对于系统实车测试、实验次数、测试范围还不能满足未来商用范围。因此在开发过程中需要对人机功能进行合理分配，增大试验次数和实验范围。对于设备的更新与测试要经过更加严格的流程。同时要对驾驶环境进行更加真实的模拟或者实验，在这个过程中增加识别判定的机制，在硬件和软件层面都增加一些冗余度设计，增加一套以上完成相同功能的功能通道、工作部件，以保证当该部分系统出现识别误差时，该系统仍然能够实现高精度的识别。而当常规系统出现故障时，仍有其他系统能够正常工作。提高系统可靠性。

### 4) 搭建与汽车之间的接口

目前系统还未与汽车控制系统建立联系，因而还不能实现强制制动、开关双闪，如果能搭建与汽车之间的接口，那么该系统能在驾驶者处于极度危险状态时发送信息给汽车控制系统，实行强制制动等主动防护措施。

### 5) 降低音频噪音

目前系统语音识别部分采用的是麦克风采集驾驶者语音信息，但驾驶者行车环境并不是一个很安静的环境，所以麦克风采集的并不只有驾驶者语音，还混杂了车内车外各种声音，因而，如果对采集到的音频先做一个降噪处理，再进行驾驶者语言表述分析，这样分析结果将会更加准确。

## 8. 总结

世界卫生组织提供的数据显示，全世界每年因道路交通事故死亡人数约有 125 万，相当于全球每天有 3500 人因交通事故死亡。数据显示，每年还有几千万人因此而受伤或致残。80%的交通事故都是由于汽车驾驶员的不安全行为导致。

本文从面部，肢体，语言三方面对驾驶者的驾驶行为进行检测，以将传统的交通行业与新兴的计算机技术结合，在驾驶舱内安装高精度摄像头以及麦克风来获取驾驶者的面部表情信息，肢体动作信息，以及语言表达信息，使用卷积神经网络以及循环神经网络进行数据处理，并且建立车联网云端系统对信息进行综合分析，得到驾驶者的状态后采取相应的预警防范措施。基于深度学习和云计算的驾驶行为识别预警系统结合了人工智能技术和云计算技术，能够识别不同类型的危险驾驶行为和风险驾驶行为，并给出相

---

应的预警，能够让驾驶人员由于驾驶行为不正确所带来的交通安全隐患降到最低，给人的生活出行带来更加可靠的保障。

---

## 参考文献

- [1] 雷虎. 愤怒情绪下的汽车驾驶行为特征及其对交通安全的影响研究[D]. 武汉理工大学, 2011.
- [2] 吴志周, 贾俊飞. 驾驶分心行为及应对策略研究综述[J]. 交通信息与安全, 2011, 29(5):5
- [3] Federico Baronti, Francesco Lenzi, Roberto Roncella, et al, Distributed Sensor for Steering Wheel Grip Force Measurement in Driver Fatigue Detection[A]. Proceedings of the Conference on Design, Automation and Test[C], 2009:894-897
- [4] Mark A. Sehler. Changes in EEG Alpha Power during Simulated Driving: a Demonstration. International Journal of Psychophysiology. 2000,(37):155-162
- [5] 王炳浩, 魏建勤, 吴永红. 汽车驾驶员瞌睡状态脑电波特征的初步探索[J]. 汽车工程, 2004, 26(1):70-73
- [6] J.B.J.Riemersma, A.Randers, C. Hildervanek. Performance decrement during Prolonged night driving. In: Mackie.R. (Ed.), Vigilance: Theory, Operational Performance and Physiological Correlates. Plenum Press, New York, 1977:41~58
- [7] 郭士杰, 杨雪翠等. 汽车驾驶员专用呼吸监测系统. 中国, 201020535549.8[P], 2011-4-3
- [8] Krajewski J, Sommer D, Trutschel U, et al. Steering wheel behavior based estimation of fatigue[C]Proceedings of the Fifth International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design. 2009: 118-124
- [9] Eskandarian A, Sayed R, Delaigue P, et al. Advanced driver fatigue research[R]. 2007
- [10] Zhang H, Smith M R H, Witt G J. Identification of real-time diagnostic measures of visual distraction with an automatic eye-tracking system [J]. Human Factors: The Journal of the Human Factors and Ergonomics Society, 2006, 48(4):805-821.
- [11] 马勇, 石涌泉, 付锐, 等. 驾驶人分心时长对车道偏离影响的实车试验[J]. 吉林大学学报(工学版), 2015, 45(4):1095-1101.
- [12] Ranney T A, Mazzae E, Garrott W R, et al. NHTSA driver distraction research: past, present, and future [EB/OL]. (2000-07-05). [2005-12-20]. <http://www-nrd.nhtsa.dot.gov/departments/nrd-13/driver-distraction/Wel-come.Htm>.