

分类号	<u>TP391.4</u>	密级	<u>公开</u>
UDC	<u>621.3</u>	学位论文编号	<u>D-10617-308-(2022)-01174</u>

## 重庆邮电大学硕士学位论文

中文题目	<u>基于面部多特征融合的</u> <u>疲劳驾驶检测算法研究及实现</u>
英文题目	<u>Research and Implementation of</u> <u>Fatigue Driving Detection Algorithm</u> <u>Based on Facial Multi-feature Fusion</u>
学 号	<u>S190101177</u>
姓 名	<u>熊 昆</u>
学位类别	<u>工学硕士</u>
学科专业	<u>信息与通信工程</u>
指导教师	<u>代少升 教授</u>
完成日期	<u>2022 年 5 月 11 日</u>

## 独创性声明

本人声明所呈交的学位论文是本人在导师指导下进行的研究工作及取得的研究成果。尽我所知，除了文中特别加以标注和致谢的地方外，论文中不包含他人已经发表或撰写过的研究成果，也不包含为获得重庆邮电大学或其他单位的学位或证书而使用过的材料。与我一同工作的人员对本文研究做出的贡献均已在论文中作了明确的说明并致以谢意。

作者签名：熊昆

日期：2022年6月14日

## 学位论文授权使用授权书

本人完全了解重庆邮电大学有权保留、使用学位论文纸质版和电子版的规定，即学校有权向国家有关部门或机构送交论文，允许论文被查阅和借阅等。本人授权重庆邮电大学可以公布本学位论文的全部或部分内容，可编入有关数据库或信息系统进行检索、分析或评价，可以采用影印、缩印、扫描或拷贝等复制手段保存、汇编本学位论文。

（注：保密的学位论文在解密后适用本授权书。）

作者签名：熊昆

日期：2022年6月14日

导师签名：代书

日期：2022年6月14日

## 摘要

疲劳驾驶是影响交通安全的主要因素，对生命财产有很大的损害。为保障驾驶人员及行人安全，减少交通事故，研究及实现一款车载疲劳驾驶检测系统有着极大的应用价值。基于驾驶员面部视觉信息的判定方法具有易采集、准确、不接触等特点，在各类疲劳驾驶检测中极具发展潜力和应用前景。受实际驾驶过程中头部姿态变化、面部部分遮挡、光照条件的随机性、疲劳表现多样性等因素的影响，全天候、高鲁棒的疲劳驾驶检测仍然存在许多技术瓶颈。本文围绕驾驶员眼睛和嘴部特征进行疲劳状态识别，开发疲劳驾驶检测算法，并进行软硬件设计和实现。

在定位眼睛和嘴部区域时，为提高大姿态、部分遮挡情况人脸特征点定位的准确性，本文提出一种多视角约束级联回归的视频人脸特征点定位和跟踪算法。该方法利用 3 维和 2 维稀疏点集建立的变换关系估计初始形状，并使用仿射变换对人脸图像做一个姿态矫正。在构造形状回归模型时，分别采用左脸、正脸、右脸构造级联回归模型。最后采用了重新初始化机制，在特征点正确定位时，使用归一化互相关模板匹配跟踪建立连续帧之间的形状关系。

基于人脸眼睛和嘴部的特征点定位结果，本文将闭眼和打哈欠特征参数相结合，提出利用加权和实现疲劳状态等级的划分。针对眼睛状态的判定，本文提出一种基于改进 MobileNet-SSD 网络的人眼状态判别方法，同时为提升瞳孔部分遮挡时眼睛状态判定的准确性，引入眼睛张角辅助判定。而针对嘴部状态的判定，使用常用的嘴部纵横比来衡量嘴巴的张开程度，并将嘴部张开较大程度的持续帧数作为哈欠动作的判定特征。

针对疲劳驾驶检测算法的具体实现，本文结合软硬件设计疲劳驾驶检测系统。为克服夜晚工况对驾驶员进行疲劳检测，系统采用近红外摄像头获取驾驶员面部图像。为提高系统实时性，系统采用具有深度学习加速功能的 NT96580 芯片作为核心处理器，并在此基础上设计软件系统。最后经测试实验表明：该系统白天和夜晚都具有较高的准确率和实时性，其中疲劳状态等级划分准确率均在 90% 以上，具有一定使用价值。

**关键词：**疲劳驾驶检测，人脸特征点定位，级联回归，MobileNet-SSD

## Abstract

Fatigue driving is the main factor affecting traffic safety, and it has great damage to life and property. In order to ensure the safety of drivers and pedestrians and reduce traffic accidents, it is of great application value to research and implement a vehicle-mounted fatigue driving detection system. The determination method based on the driver's facial visual information has the characteristics of easy collection, accuracy, and no contact, and has great development potential and application prospects in various types of fatigue driving detection. Affected by factors such as head posture changes, partial facial occlusion, randomness of lighting conditions, and diversity of fatigue performance during actual driving, there are still many technical bottlenecks in all-weather and highly robust fatigue driving detection. This thesis focuses on the driver's eyes and mouth features to identify the fatigue state, develop a fatigue driving detection algorithm, and design and implement software and hardware.

In order to improve the accuracy of facial feature point location in large pose and partial occlusion when locating eyes and mouth regions, this thesis proposes a video facial feature point location and tracking algorithm based on multi-view constrained cascade regression. The method estimates the initial shape using the transformation relationship established by 3D and 2D sparse point sets, and uses affine transformation to perform a pose correction on the face image. When constructing the shape regression model, the left face, front face and right face are respectively used to construct the cascade regression model. Finally, a re-initialization mechanism is adopted to establish the shape relationship between consecutive frames using normalized cross-correlation template matching tracking when the feature points are located correctly.

Based on the feature point localization results of eyes and mouth, this thesis combines the feature parameters of eyes closed and yawning, and proposes the use of weighted sum to achieve the classification of fatigue state levels. For the determination of eye state, this thesis proposes a method for human eye state discrimination based on the improved MobileNet-SSD network. At the same time, in order to improve the accuracy of eye state determination when the pupil is partially occluded, an auxiliary determination of eye opening angle is introduced. For the judgment of the mouth state, the commonly used mouth aspect ratio is used to measure the opening degree of the mouth, and the continuous

frame number of the mouth opening to a greater extent is used as the judgment feature of the yawning action.

Aiming at the specific realization of the fatigue driving detection algorithm, this thesis designs a fatigue driving detection system combining software and hardware. In order to detect the driver's fatigue under nighttime conditions, the system uses a near-infrared camera to obtain the driver's face image. In order to improve the real-time performance of the system, the system adopts the NT96580 chip with deep learning acceleration function as the core processor, and designs the software system on this basis. Finally, the test experiments show that the system has high accuracy and real-time performance during the day and night, and the accuracy of the classification of fatigue status is above 90%, which has a certain use value.

**Keywords:** fatigue driving detection, facial feature point location, cascade regression, MobileNet-SSD

## 目录

图录 .....	IX
表录 .....	XIII
第 1 章 引言 .....	1
1.1 研究背景及意义 .....	1
1.2 疲劳驾驶检测研究现状 .....	2
1.2.1 基于生理信号的检测方法 .....	2
1.2.2 基于车辆行为的检测方法 .....	3
1.2.3 基于驾驶员行为的检测方法 .....	4
1.3 研究内容及结构安排 .....	5
1.3.1 研究内容 .....	5
1.3.2 结构安排 .....	6
第 2 章 疲劳驾驶检测相关技术 .....	9
2.1 深度学习技术 .....	9
2.1.1 深度学习开发框架 .....	9
2.1.2 卷积神经网络 .....	10
2.1.3 目标检测网络 .....	11
2.2 面部分析相关技术 .....	12
2.2.1 人脸检测 .....	12
2.2.2 人脸特征点检测 .....	14
2.2.3 头部姿态估计 .....	15
2.3 疲劳判定指标 .....	16
2.4 本章小结 .....	17
第 3 章 人脸特征点定位和跟踪算法研究 .....	19
3.1 基于级联回归特征点定位的不足 .....	19
3.2 MCCR 视频人脸特征点定位和跟踪 .....	20
3.2.1 算法概述 .....	20

3.2.2 初始化.....	21
3.2.3 多视角约束.....	22
3.2.4 多视角级联形状回归.....	25
3.2.5 模板匹配跟踪.....	27
3.2.6 重新初始化.....	28
3.3 实验结果分析.....	29
3.4 本章小结.....	33
第 4 章 疲劳驾驶状态检测算法研究 .....	35
4.1 算法概述.....	35
4.2 眼部状态识别 .....	36
4.2.1 常用眼部状态识别方法的不足.....	36
4.2.2 基于改进 MobileNet-SSD 的眼部状态识别 .....	38
4.2.3 眼睛张角辅助判定眼部状态 .....	41
4.2.4 实验结果分析.....	43
4.3 嘴部状态识别 .....	44
4.3.1 纵横比嘴部状态识别.....	44
4.3.2 实验结果分析.....	46
4.4 疲劳状态等级划分 .....	47
4.4.1 多特征融合的划分规则.....	47
4.4.2 实验结果分析.....	48
4.5 本章小结 .....	49
第 5 章 疲劳驾驶检测系统的设计与实现 .....	51
5.1 系统硬件设计.....	51
5.1.1 设计方案.....	51
5.1.2 图像采集模块.....	52
5.1.3 NT96580 处理器.....	52
5.1.4 数据存储模块.....	54
5.1.5 显示模块.....	54
5.2 系统软件设计.....	55

---

5.2.1 软件系统架构 .....	55
5.2.2 软件系统工作流程 .....	56
5.3 疲劳驾驶检测系统测试与分析 .....	57
5.3.1 测试环境 .....	57
5.3.2 实验结果与分析 .....	58
5.4 本章小结 .....	61
第 6 章 总结与展望.....	63
6.1 论文工作总结 .....	63
6.2 工作展望 .....	64
参考文献 .....	65
致谢 .....	71
攻读硕士学位期间从事的科研工作及取得的成果 .....	73



## 图录

图 2.1 卷积网络的完整结构 .....	10
图 2.2 卷积细节 .....	10
图 2.3 池化层示意图 .....	11
图 2.4 MTCNN 工作流程图 .....	13
图 2.5 P-Net 网络结构 .....	14
图 2.6 R-Net 网络结构 .....	14
图 2.7 O-Net 网络结构 .....	14
图 2.8 头部姿态示意图 .....	16
图 2.9 PERCLOS 的 P80 标准 .....	16
图 3.1 级联回归方法示意图 .....	19
图 3.2 MCCR 算法流程 .....	20
图 3.3 不同投影映射关系变换结果 .....	21
图 3.4 初始化结果 .....	22
图 3.5 特征点部分重叠 .....	22
图 3.6 不同增量倍数叠加后的人脸形状 .....	24
图 3.7 24 种不同增量的形状 .....	25
图 3.8 调整前后的形状 .....	25
图 3.9 仿射变换前后 .....	26
图 3.10 特征点回归过程 .....	26
图 3.11 NCC 跟踪 .....	27
图 3.12 NCC 模板匹配跟踪结果 .....	28
图 3.13 重新初始化步骤 .....	29
图 3.14 MCCR 的示例图像 .....	30
图 3.15 重新初始化验证结果 .....	31
图 3.16 累积误差曲线和前 400 帧归一化误差 .....	32
图 3.17 300VW 结果 .....	32

图 4.1 疲劳驾驶状态检测算法流程 .....	35
图 4.2 眼睛纵横比 .....	36
图 4.3 头部姿态与 EAR 的变化 .....	36
图 4.4 眼睛的灰度投影 .....	37
图 4.5 复杂环境下眼睛灰度投影 .....	37
图 4.6 基于 Hough 方法的检测结果 .....	37
图 4.7 MobileNet-SSD 网络结构 .....	38
图 4.8 数据集中眼睛长宽分布图 .....	39
图 4.9 改进网络的架构图 .....	40
图 4.10 增大人脸检测方法流程图 .....	41
图 4.11 增大人脸检测区域效果图 .....	41
图 4.12 眼睛张角示意图 .....	41
图 4.13 眼睛状态识别流程图 .....	42
图 4.14 睁闭眼张角统计 .....	43
图 4.15 数据集标注示例 .....	43
图 4.16 嘴巴纵横比示意图 .....	45
图 4.17 打哈欠 MAR 统计 .....	45
图 4.18 不同场景下的 MAR .....	46
图 4.19 数据集标准示例 .....	47
图 4.20 $w_1$ 和 $w_2$ 寻优 .....	48
图 5.1 NT96580 整体功能结构 .....	51
图 5.2 高清近红外摄像头 .....	52
图 5.3 深度学习加速模块开发流程 .....	53
图 5.4 系统硬件电路板 .....	54
图 5.5 LCD 显示屏 .....	55
图 5.6 系统软件架构 .....	55
图 5.7 软件工作总流程图 .....	56
图 5.8 硬件设备 .....	57
图 5.9 摄像头安装位置 .....	58

---

图 5.10 特征点定位效果 .....	58
图 5.11 眼睛和嘴部状态识别结果 .....	59
图 5.12 设备显示结果 .....	61

表录

表 3.1 平均误差与现有方法的比较 ..... 30

表 4.1 网络检测层输出统计结果 ..... 39

表 4.2 网络硬件平台测试 ..... 44

表 4.3 眼睛状态识别方法比较 ..... 44

表 4.4 嘴部状态识别方法比较 ..... 47

表 4.5 疲劳状态等级划分规则 ..... 48

表 4.6 疲劳状态检测准确率 ..... 48

表 4.7 使用不同特征的检测准确率 ..... 49

表 5.1 特征点定位算法运行时间 ..... 58

表 5.2 不同环境下疲劳驾驶检测性能 ..... 60

## 第 1 章 引言

### 1.1 研究背景及意义

随着经济的发展，国民消费能力的不断提高，我国民用汽车的数量大幅度增长。根据国家权威数据显示，2020 年我国已拥有民用汽车 27340.92 万辆，而相关的机动车驾驶员有 45702.49 万人。庞大数字的面前，我国道路安全问题也面临着巨大的挑战。《中国统计年鉴》统计数据中，我国 2018 至 2020 年交通事故年均死亡人数为 6.3 万人，受伤人数为 24.14 万人，其中交通事故发生次数为 23.19 万次<sup>[1]</sup>。面对交通事故引发的人员伤亡和经济损失，如何保证道路交通安全是需要面临的重大问题。

关于道路交通事故的调查显示，引发道路交通安全问题的主要因素是人、车和环境，其中人为因素的占比达到了 90%<sup>[2]</sup>。可见道路交通安全与驾驶员有着极大的关联，驾驶员的驾驶行为关系到事故的发生。影响驾驶员安全驾驶的因素有很多，其中主要有酒驾、超速和疲劳驾驶<sup>[3]</sup>。其中疲劳驾驶在生活中广泛存在，20%的交通事故与其相关。驾驶员在过多的重复操作中，会因心理和生理的变化引起疲劳，影响驾驶员的注意力和反应能力，致使错误驾驶或者驾驶失能而导致交通事故<sup>[4]</sup>。因此，行车过程中出现的疲劳驾驶是一个不容忽视的问题。

道路交通安全中疲劳驾驶已经是一个常见的问题，公安交管部门也有相关的法律法规，例如驾驶机动车不能连续超过 4 小时，对客运驾驶人员还有更加严格的规定，这些都是为了避免疲劳驾驶而出台的规定。与超速、超重和酒驾等不安全驾驶行为相比，疲劳驾驶是一个很难实际去管控的行为。日常交通管制中，交管人员针对疲劳驾驶检测没有严格的设备支持，简单从驾驶员连续驾驶的时间去判断并不是可靠的方法，它有着一些局限性。

研究发现，驾驶员的反应速度与交通事故息息相关，在事故发生前仅快 0.5 秒就能减少一半以上的交通事故发生<sup>[5]</sup>。当司机出现疲劳时，若能及时进行检测并引导驾驶人员，这将减少部分交通事故的发生。因此为保障驾驶人员和行人安全，减少交通事故发生，研究及实现一款准确率高、实时性强的小型车载疲劳驾驶检测系统具有重大意义。

## 1.2 疲劳驾驶检测研究现状

随着国内外学者对疲劳驾驶的深入研究，越来越多的检测方法被提出，涉及的标准也不尽相同，主要可分为三类：基于生理信号的检测方法、基于车辆行为的检测方法、基于驾驶员行为的检测方法。

### 1.2.1 基于生理信号的检测方法

通过驾驶员的生理信号判断疲劳程度是最为准确和客观的检测方法，该方法是接触式的，检测时需要使用传感器采集驾驶员的相关生理参数，最后分析参数的变化实现对疲劳的检测。

脑电图(Electroencephalography, EEG)直接反映了驾驶员的大脑活动，被称作判断疲劳的“黄金标准”。文献[6]对脑电信号大量研究后发现，清醒状态与疲劳状态相比脑电信号有着明显不同，当驾驶员逐渐疲劳时，Theta 波也会随之发生变化，因此可通过此生理参数检测驾驶员的疲劳状态。文献[7]对大量模拟驾驶的脑电信号进行采集，发现疲劳判断准则 P80 与脑电信号之间存在很大的关联性。文献[8]在模拟实验中使用样本熵描述脑电信号复杂度的变化，发现疲劳状态下大脑的复杂性会降低，主要表现为  $\beta$  节律下大脑额叶和中枢区域样本熵会明显降低。文献[9]在研究脑电信号对驾驶疲劳的分类问题时，将提取到的功率特征和连通性特征进行多种方式的融合，最终发现特征级和决策级融合与单个特征分类相比较，准确率均有明显提高。文献[10]在疲劳识别时使用了深度迁移学习的方法，将脑电信号作为数据，创新了疲劳的识别方法。

心电图(Electrocardiogram, ECG)是心电仪根据心脏跳动产生的生物电位变化，该生理参数反映了人体心脏跳动的规律。使用心电图检测疲劳主要是通过心率的变化判断。文献[11]在研究中发现当驾驶员处于疲劳状态时，ECG 信号有着规律性下降的趋势。文献[12]提出一种使用短周期心电信号进行疲劳识别的方法，并解决了一些心电信号特征提取的问题，有效地提高了检测的实时性。文献[13]使用深度学习体系结构进行心电疲劳检测，并将时域、频率和非线性心率变异性特征结合，达到了高准确率的疲劳检测。文献[14]为了提高噪声环境下驾驶疲劳的检测性能，研究如何融合脑电信号和心电信号，提出一种新的深度学习网络模型。文献[15]采用

自适应滤波器的方法捕获心电信号，利用心率变异性和心率对驾驶员进行疲劳分析，设计了一套无干扰的疲劳评估系统，有效地减少了复杂监控设备和辅助设备给驾驶员带来的行车干扰。

肌电图(Electromyography, EMG)分为肌肉电图和表面肌电图，反应了人体肌肉活动的情况。为防止驾驶员衣服的影响，疲劳检测的相关研究大多使用表面肌电信号，涉及的特征参数有肌电平均频率、肌电图幅值等。文献[16]在研究驾驶员行车过程中的 EMG 信号时发现，当司机进入疲劳状态时采集到的表面肌电信号幅值会显著增长，而平均频率下降，因此这两个特征参数可作为判断疲劳驾驶的依据。文献[17]提出一种新的疲劳驾驶检测方法，将心电与肌电信号特征融合作为疲劳特征，并引入迁移学习检测疲劳。文献[18]通过在方向盘上嵌入表面肌电信号传感器，以非侵入的方式采集驾驶员生理信号，这种方法能较早对驾驶员的疲劳状态进行检测。

使用生理信号检测疲劳的方法准确率相对较高，但是会佩戴一些复杂的生理信号采集装置，这将会影响到驾驶员的安全操作，造成不必要的安全事故。因此，这类方法不易在行车环境中普及。

### 1.2.2 基于车辆行为的检测方法

基于车辆行为的检测方法主要是通过车辆信息来判断疲劳状态，如车辆方向盘、行驶轨迹、加速踏板和制动踏板等。同正常情况下驾驶员操作车辆相比，当驾驶员处于疲劳时，会因身体机能下降而使行驶过程中方向盘角度、车辆轨迹、踏板压力等长时间未变动。目前此类疲劳驾驶检测方法分为基于车辆参数和基于车辆行驶轨迹两种。

基于车辆信息的疲劳驾驶检测方法主要是对方向盘、制动和加速踏板等的运行情况进行分析，信息的收集相对容易。文献[19]研究了司机驾驶时方向盘的操作情况，并以转角和速率设置坐标系，分析发现正常状态时数据在原点为中心的椭圆内分布，而疲劳状态时数据分布较远，因此可根据此特点判断驾驶员的疲劳状态。文献[20]在行驶模拟器上使用数据采集系统采集数据，得到大量车辆加速、制动、换挡和转向的变化信号，经过分析车辆信息实现对驾驶员疲劳的检测。文献[21]将方向盘握力信号和驾驶员脑电信号作为 BP 神经网络的训练数据，建立了两者的关系，实现了对驾驶员的疲劳检测。

基于车辆行驶轨迹的疲劳驾驶检测方法主要通过分析车辆的一些运行轨迹，进而判断驾驶员的状态。文献[22]对车辆横向位移标准差研究后发现，当司机疲劳程度不断加重时，车辆横向位移标准差会不断变大，然而该方法受实际的道路环境影响。文献[23]使用摄像头获取车道线的数据，利用对车道线的建模和追踪分析驾驶员的疲劳状态，最终验证该方法是实际有效的。文献[24]在研究车辆行驶过程中的偏航角后发现，正常行驶时车辆的偏航角会在一定范围内波动，当驾驶员疲劳程度加深时，偏航角的波动会随之增大。文献[25]也在研究中发现，当驾驶员疲劳时车辆的偏航角变化加剧，因此提出通过偏航角的标准差判断驾驶员的疲劳状态。

通过车辆行为判断驾驶员疲劳状态，数据获取相对生理参数比较容易，并且数据处理简单。在驾驶员安全操作过程中不会对驾驶员产生影响，能达到实时检测。但是车辆行为易受外界环境影响，导致该类方法存在很大的误差。

### 1.2.3 基于驾驶员行为的检测方法

基于驾驶员行为的检测方法主要分析行为特征检测疲劳状态，如面部特征、头部运动轨迹等，这些信息与疲劳有一定的关系，能真实反应驾驶员的生理状态。

驾驶员处于疲劳状态时往往会打瞌睡，并且伴随着闭眼现象，因此闭眼频率和闭眼幅值等可直接反映驾驶员的疲劳程度。文献[26-28]构建专用的卷积神经网络检测人眼状态，并最终使用闭眼频率和眼睑闭合率判断司机的疲劳状态。文献[29]提出一种基于层次插值的方法，从低质量的眼动数据中提取眼动疲劳指标，并给出适用于人体疲劳检测的注视和扫视定义，通过实验验证该方法获得的眼球跟踪疲劳指标相对稳定，在人体疲劳检测中准确率高。文献[30]研究了一种利用双向投影积分进行眼睛定位的方法，利用瞳孔半径的变化曲线检测驾驶员的睁眼和闭眼频率，从而判断驾驶员的疲劳状态。文献[31]为提高行车光照环境下的精度，首先采用同态滤波预处理图像，消除各种光照环境的影响，然后定位眼睛的特征点，最后根据特征点计算多种眼睛特征参数，并使用支持向量机构建分类器检测驾驶员的疲劳状态。文献[32]提出一种基于迁移学习的眼睛状态识别网络，首先使用多任务卷积神经网络得到人脸和双眼，然后提取眼部图像并用眼睛识别网络识别，这种方法在光照和头部姿态变化条件下有良好的准确率。文献[33]针对复杂背景、光照等环境，基于LeNet-5提出一种变体模型，这种模型有效地增加了眼部状态的识别率。



打哈欠也是疲劳表现的一种特征，但是光靠嘴部动作检测率很低，通常是将嘴部和眼睛特征进行结合判断，增加识别的准确性。文献[34, 35]使用 AdaBoots 算法得到驾驶员的人脸区域，然后使用人脸特征点定位算法得到面部的五官位置，最后根据眼睛和嘴巴特征实现疲劳检测。文献[36]提出一种基于多指标融合和状态识别的疲劳驾驶检测算法，使用一个多任务级联卷积神经网络检测人脸和面部特征点，然后通过构造双眼图像的眼睛状态识别网络和嘴状态识别网络，最后结合眼睛和嘴巴状态建立疲劳判断模型。文献[37]构建了一种能够同时识别视频图像中人脸和眼睛的分类器，并利用跟踪算法实现人眼的稳定跟踪，最后结合多种面部特征确定人体疲劳状态。文献[38]采用级联神经网络检测疲劳，将改进后的人脸检测网络与设计的嘴巴和眼睛检测网络进行级联，提升检测准确率。文献[39]根据人脸检测器得到的人脸区域，使用面部器官的几何分布计算眼睛和嘴部区域，然后根据似圆度判断嘴部的状态，而眼部状态的识别则使用灰度直方图统计。文献[40]在获得人脸区域的情况下，分别使用 EyeMap 算法和肤色差异获取眼部和嘴部区域，并利用卷积神经网络进行疲劳检测。

头部的运动状态也是疲劳的一种判定特征，加入头部运动信息的疲劳识别方法准确率有一定提高。文献[41]通过人脸检测来得到头部的运动信息，将打哈欠和头部运动相结合达到了较好的疲劳识别率。文献[42]使用运动捕捉系统获得驾驶员的头部姿态运动信息，设计了一种基于修正线性单元层的双向长短时记忆深度神经网络，并对三维时间序列的头部角速度数据进行训练和测试，用于序列间的分类，该方法在识别驾驶员主动状态、疲劳状态和过渡状态方面的性能优于现有的方法。

基于驾驶员行为的疲劳驾驶检测方法不需要设备与驾驶员直接接触，并且信息容易采集，对疲劳的判定也较准确。目前该方法已得到广泛应用，并有许多相关研究论文。因此，本文将基于此类方法研究一种准确、实时的疲劳驾驶检测算法。

## 1.3 研究内容及结构安排

### 1.3.1 研究内容

本文主要针对疲劳驾驶检测进行深入研究和嵌入式端实现，主要的工作内容包含以下几个方面：

### 1. 疲劳驾驶检测方法研究现状

对现有的疲劳驾驶检测方法进行详细的研究，整理了各类方法的国内外研究现状，并分析总结了方法的优点。

### 2. 人脸特征点定位和跟踪算法研究

在人脸图像中预测面部特征点位置是疲劳驾驶检测中不可缺少的部分，可以使得眼睛和嘴部的定位更加准确。本文将人脸特征点的检测与跟踪相结合，提出一种多视角约束级联回归的视频人脸特征点定位和跟踪算法。为应对人脸部分遮挡，本文采用稀疏特征点估计密集特征点的方法，在连续的视频帧之间还使用了归一化互相关匹配建立帧与帧之间的联系。该算法适合在摄像头视频流中检测和跟踪人脸特征点，针对大姿态和部分遮挡具有较强的鲁棒性。

### 3. 疲劳驾驶状态检测算法研究

为提高疲劳驾驶检测的准确率，本文将闭眼和打哈欠特征参数相结合，提出利用加权和实现疲劳状态等级的划分。针对眼睛状态的识别，本文提出一种基于改进 MobileNet-SSD 网络的人眼状态判别方法，同时为提升瞳孔部分遮挡时眼睛状态判定的准确性，引入眼睛张角辅助判定。而针对嘴部状态的判定，使用常用的嘴部纵横比衡量嘴巴的张开程度，并将嘴部张开较大程度持续帧数作为打哈欠动作的判定特征。

### 4. 设计和实现疲劳驾驶检测系统

针对本文提出和改进的疲劳驾驶检测算法，将在嵌入式端进行软硬件的设计和实现，并对疲劳驾驶检测系统进行整体测试分析。

## 1.3.2 结构安排

本文各章节的内容安排如下：

第 1 章，对疲劳驾驶的背景进行阐述，通过对疲劳驾驶危害的描述引出本课题的研究意义。同时对各类疲劳驾驶检测方法的研究进展进行介绍，最后阐述了本文具体研究内容和结构安排。

第 2 章，介绍疲劳驾驶检测算法涉及的一些理论知识，主要是深度学习、面部分析相关技术、疲劳判定指标。

第 3 章，针对驾驶员面部特征点定位，本章首先阐述传统级联回归方法的不足，提出一种多视角约束级联回归的视频人脸特征点定位和跟踪算法，然后对算法涉及的初始化、多视角约束、多视角级联形状回归、归一化互相关(Normalized Cross Correlation, NCC)模板匹配跟踪、重新初始化部分进行详细的描述，最后通过测试验证该算法的性能。

第 4 章，针对驾驶员疲劳状态的划分，首先对复杂环境下部分眼睛状态识别方法的性能进行测试，然后提出一种基于改进 MobileNet-SSD 网络的人眼状态判别方法，同时为提升瞳孔部分遮挡时眼睛状态判定的准确性，引入眼睛张角辅助判定。接着针对嘴部状态的识别，使用常用的嘴部纵横比衡量嘴巴的张开程度，并将嘴部张开较大程度持续帧数作为打哈欠动作的判定特征。最后将闭眼和打哈欠特征参数加权进行疲劳状态等级的划分，并进行实验和分析。

第 5 章，主要对本文实现的疲劳驾驶检测系统进行整体介绍，首先对硬件系统的各部分进行介绍，接着对软件系统的工作流程进行介绍，最后对系统进行整体的测试和分析。

第 6 章，总结本文的研究工作，同时分析本文提出的疲劳驾驶检测算法的不足，并对今后的研究和改进方向进行阐述。



## 第2章 疲劳驾驶检测相关技术

本章主要对疲劳驾驶检测涉及的一些技术理论进行介绍，首先对深度学习涉及的概念和技术进行叙述，然后对人脸检测和人脸特征点检测的一些方法进行介绍，最后介绍关于疲劳判定的指标。

### 2.1 深度学习技术

#### 2.1.1 深度学习开发框架

深度学习是机器学习中一种基于对数据进行表征学习的方法，作为热门领域，吸引了 Google、Facebook、Microsoft 等科技公司的重点投资，这包括一些深度学习框架，以下将对目前流行的深度学习框架进行介绍。

Tensorflow 是一款使用数据流图形式进行计算的开源数学计算软件，它是由深度学习基础架构 DistBelief 二次开发获得。Tensorflow 架构灵活，可应用在许多场景中，如移动设备、服务器等。此外，由于 Google 公司频繁的更新和维护，再加上有着多种编程语言的接口，使其成为社区最庞大的一个框架。

Caffe 是由伯克利视觉中心维护的一个高效的开源深度学习框架，主要运用 C++ 实现。它可以在多种处理器上运行，并且有多种语言的接口，如 Python、Matlab 等。Caffe 使用起来简单快捷，但是灵活性较低。

Theano 是有着较大影响力的 Python 学习框架，它是专门设计来处理大型神经网络算法中的计算问题，被认为是行业的标准，后续一些深度学习框架也借鉴了它的设计思想。Theano 能获取用户的代码结构，核心是一个数学表达式的编译器。

MXNet 被亚马逊所使用，有着良好的性能、占用显存低、超强的分布式支持等优点。它支持命令和符号编程，可以运行在 CPU、GPU、集群、服务器以及移动设备上，同时还提供了多种编程语言的接口。但是 MXNet 版本更新快，其文档更新速度却很慢，新手用户难以快速掌握，因此在一定程度上限制了它的发展。

PyTorch 的前身是 Torch，它是难得的简单、高效、快捷的深度学习框架。PyTorch 与 Torch 相比，有着更加灵活的开发方式，并且还有 Python 接口，可看作

numpy 增加了对 GPU 的支持。除 Facebook 以外,采用的还有 Twitter、Salesforce 和 CMU 等公司。

### 2.1.2 卷积神经网络

卷积神经网络是一种带卷积结构的前馈型神经网络,具有强大的特征提取能力。它的使用大大减少了计算时的内存占用,与普通神经网络相比参数量更少。目前卷积神经网络已经在图像分类、目标检测等领域广泛使用。

卷积神经网络的结构较为复杂,涉及很多的层数,每一层有自己特定的功能,主要使用了局部感知和权值共享机制。其具体结构如图 2.1 所示,下面重点对卷积操作和池化层进行介绍。

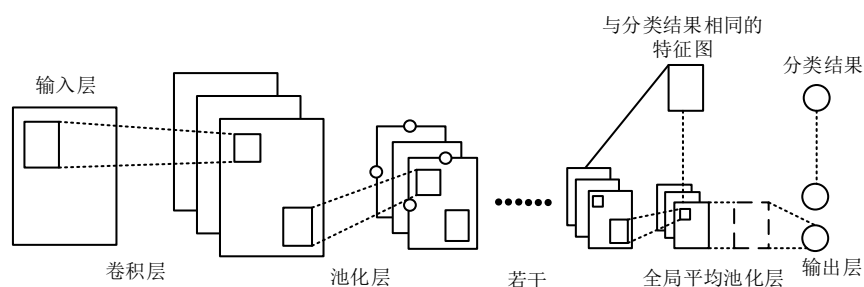


图 2.1 卷积网络的完整结构

#### 1. 卷积操作

卷积操作的目的是为了提取图像的特征,它可分为窄卷积、同卷积和全卷积。卷积操作的核心就是步长,不同的步长可以得到不同类型的卷积操作,在图 2.2 中本文将以窄卷积为例进行相关操作的介绍。

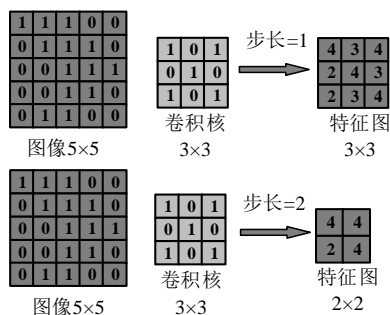


图 2.2 卷积细节

如图 2.2 所示,图像用  $5 \times 5$  的矩阵表示,卷积核为  $3 \times 3$  矩阵。卷积操作时需要找到卷积核与图像数据的对应关系,然后将数据对应相乘后相加。如结果特征图中

第一个元素，它是卷积核与图像中前3行3列的数据卷积的结果，即  $4=1\times1+1\times0+1\times1+0\times0+1\times1+1\times0+0\times1+0\times0+1\times1$ 。卷积核与图像的对应关系是可以移动的，移动的格数表示为步长。当步长为1时，最右侧特征图中的第二个元素为3，它是卷积核在计算完第一个元素以后，向右移动一个再计算卷积的结果，即  $3=1\times1+1\times0+0\times1+1\times0+1\times1+1\times0+0\times1+1\times0+1\times1$ 。当步长为2时，则每次移动两格，最终可得到最右侧2×2的特征图。

## 2. 池化层

在卷积操作后可对得到的特征进行分类，但是此时得到的特征数据量大，计算资源消耗多而导致处理速度较慢，因此需要对特征进行降维。池化的作用就是为了降低特征图的维度，它与卷积操作不同，池化只注重尺寸问题，并且和卷积一样也具有步长。图2.3(a)为最大池化方法，其原理是将滤波器映射区域元素的最大值作为池化特征，图2.3(b)为均值池化方法，其主要思想是将滤波器映射区域元素的均值作为池化特征。

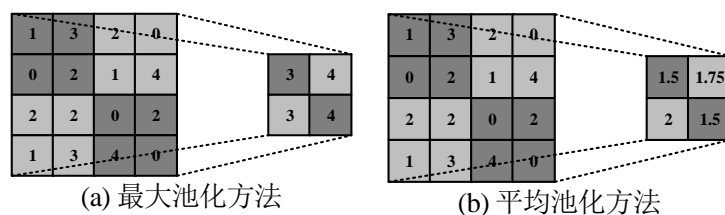


图 2.3 池化层示意图

### 2.1.3 目标检测网络

卷积神经网络在图像分类中取得巨大成就后，许多学者又将其引入目标检测、图像分割等领域。目标检测领域发展至今，出现了许多具有巨大影响力的网络结构和网络设计思想，以下分别对几种经典的目标检测网络进行介绍。

#### 1. R-CNN

R-CNN<sup>[43]</sup>是将区域生成(Region Proposal)与卷积神经网络相结合的算法，它不仅用区域生成实现图像分类与目标检测问题的转化，而且运用了卷积神经网络优秀的特征提取和分类性能，它是目标检测领域的一个具有重大意义的创新。R-CNN先是提取候选框，再经过卷积神经网络提取候选框图像的特征，接着再进行分类。但它并不穷举所有可能的区域来分类，而是基于 selective search 方法创建了约 2000 个候选区域。尽管 R-CNN 相较于传统目标检测方法有很大的改进，但是候选框的

数量依然很多,而且整个训练会耗费大量的时间和存储空间,因此后续很多目标检测网络在此基础上进行了改进。

## 2. Fast R-CNN

Fast R-CNN<sup>[44]</sup>是 R-CNN 的改进,它减少了网络训练的耗时,并降低了训练占用的空间。网络首先提取整个图像的特征,并在结果特征图上提取候选框,它不像 R-CNN 需要在图像上多次提取特征。然后使用感兴趣区域(Region of Interest, RoI)池化层对特征图大小进行统一处理,最后再输入全连接层分类和定位。因此,这种改进后的方法在处理速度上得到明显优化。为进一步优化提升 Fast R-CNN,还提出了 Faster R-CNN<sup>[45]</sup>,它使用了另外的方法得到候选框。

## 3. YOLO

前面所介绍的目标检测网络都是两个阶段(two-stage),运算速度还有待提升。而 YOLO<sup>[46]</sup>则是一个阶段(one-stage)的网络,它将目标分类和定位在一个步骤里完成。YOLO 将目标检测问题转化为回归问题,直接在输出层回归目标的位置和类别。具体的做法是先对图片划分网格,对每一个网格都会检测内部是否有目标的中心点落在该区域,如果有则会启动该任务来检测 n 个边框对象。通过这种方式的创新,网络达到了 45 帧每秒的运算速度,这完全满足实时性的要求。

## 4. SSD

SSD<sup>[47]</sup>是继 YOLO 之后的又一个 one-stage 网络,相比 YOLO 其检测准确率高。YOLO 有两个缺点:一是每个网格内目标个数的预测是固定的,容易漏检。二是对目标尺寸较为敏感。而 SSD 主要设计思想是多尺度、多长宽比的密集锚点设计和特征金字塔,它引入了区域生成网络的思想,使用了多种尺度的特征图,能对大小目标进行检测,网络具有较好的性能。

## 2.2 面部分析相关技术

### 2.2.1 人脸检测

人脸检测是当前深度学习领域重要研究热点,现在已广泛应用于许多领域,如人脸识别、表情识别、身份识别、智能安防等。目前人脸检测的方法很多,大致可分为基于知识、基于模板匹配和基于统计模型的方法。



基于知识的方法首先需要编写一些规则，然后再进行人脸检测。这种检测方法是自上而下的，先得到人脸的各种特征，再根据规则找到实际的联系。所谓的特征包括人脸的结构、灰度、纹理、肤色等，这类方法就是利用这些特征将人脸检测问题进行转化。

基于模板匹配的方法主要利用的是人脸图像与模板的自相关性，这类方法可分为固定模板和可变模板两种。前者将面部器官的关系做成固定模板，使用时将模板与图像匹配计算自相关性，后者模板的参数是可调节的，使用时还需要构建相应的能量函数。

基于统计模型的方法种类有很多，使用较广泛的是基于 AdaBoost 和神经网络的方法。基于 AdaBoost 的人脸检测主要是使用人脸数据训练得到许多弱分类器，通过融合使用得到的弱分类器构建一个强分类器。神经网络的方法是当前研究的热门，其涉及的检测网络有很多，移动端较流行的有多任务卷积神经网络(Multi-task Convolutional Neural Network, MTCNN)<sup>[48]</sup>，下面将对其进行简单介绍。

MTCNN 是多任务级联的卷积网络，它将人脸检测和关键点的定位相结合以提升检测指标。该网络主要采用了 P-Net(Proposal Network)、R-Net(Refine Network)、O-Net(Output Network)三个级联的网络，整体检测流程如图 2.4 所示。首先输入图像会进行多种尺度的缩放，这样可提高对不同大小人脸的检测能力。接着这些图像将被输入到 P-Net 网络，该子网络结构如图 2.5 所示，它能满足不同大小图像的输入。P-Net 网络将在不同尺寸的图像上以滑动窗口的方式筛选可能的人脸候选框，然后使用非极大值抑制算法对这些候选框进行过滤，边框回归后将得到 P-Net 网络输出结果。然后 P-Net 网络的输出结果会被缩放至  $24 \times 24$  并输入 R-Net 网络，该子网络结构如图 2.6 所示，R-Net 网络拥有更复杂的结构。R-Net 网络会对候选框进行选择 and 调整，同样也会利用非极大值抑制算法过滤候选框。最后 R-Net 网络的结果会被缩放至  $48 \times 48$  并送入 O-Net 网络，该子网络结构如图 2.7 所示，主要任务是对人脸候选框进行再一次筛选并得到五个特征点（鼻子、左右眼睛、左右嘴角）坐标。



图 2.4 MTCNN 工作流程图

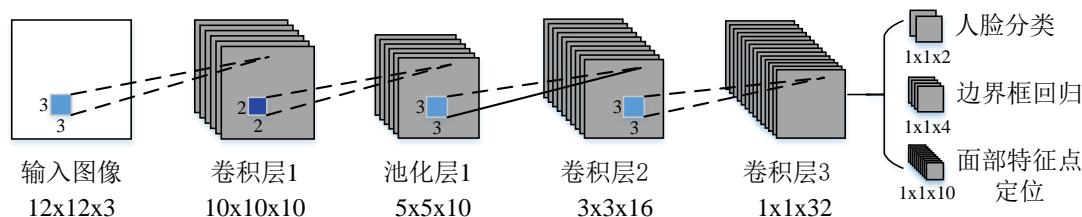


图 2.5 P-Net 网络结构

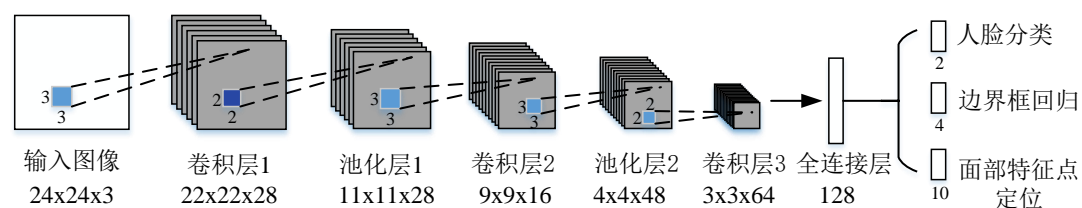


图 2.6 R-Net 网络结构

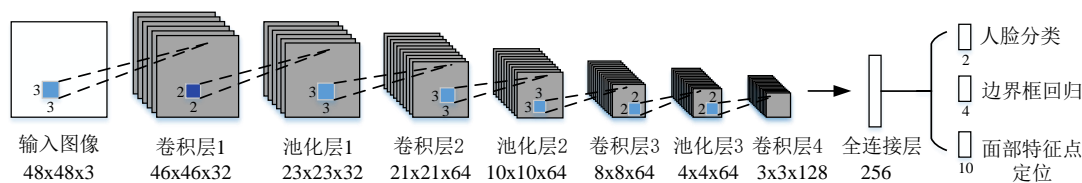


图 2.7 O-Net 网络结构

## 2.2.2 人脸特征点检测

在许多的面部应用中，人脸特征点检测是至关重要的，这包括面部行为分析、人机界面、人机交互、情感计算、唇读和监视，目前的人脸特征点检测方法主要有以下几种：

### 1. 基于传统方法

主动形状模型(Active Shape Model, ASM)是人脸特征点检测中一种典型的传统方法，该方法通过许多的点来描述形状，它是一种统计学模型。只要是在训练过程中加入过的变化模式，ASM 都能对检测的图像定位。具体定位的实现结合了形状模型和点的局部灰度分布，而相似度的表示使用当前局部灰度分布和经验灰度分布匹配的加权和。该方法的优点是模型架构简单易实现、应用方便，并且约束轮廓形状较强，但是其运算效率低。主动表观模型(Active Apperance Model, AAM)是另一种较为典型的传统方法，它是 ASM 方法的改进，这两种方法都分为构建模型和模型匹配两个阶段，不同之处是 AAM 加入了纹理特征。

### 2. 基于级联回归方法

基于级联回归方法的人脸特征点检测在近些年有大量的研究，许多的方法被提

出, 包括 SDM<sup>[49]</sup>、ESR<sup>[50]</sup>、ERT<sup>[51]</sup>、LBF<sup>[52]</sup>等方法。级联回归方法的主要思想是利用回归器对初始的人脸形状进行细化, 这种方法使用多个回归器级联进行预测, 即每个回归器的输入都是上一个回归器的预测结果。一般级联回归方法在检测时包含多个阶段, 各个阶段都会在预测的特征点位置提取形状索引特征, 然后回归器会利用形状索引特征估计特征点位置增量, 上个阶段的预测结果将根据增量被调整, 最后经过多次迭代得到最终的特征点位置。这种方法运算速度相对较快, 模型非常适合在嵌入式端运行, 因此本文采用级联回归的方法检测人脸特征点。

### 3. 基于深度学习方法

基于深度学习方法的人脸特征点检测是近些年研究的热门领域, 利用深度学习可直接对人脸图像进行处理, 而不必像级联回归那样提取特征。近些年有许多相关算法被提出, 其中最早是文献[53]提出的 DCNN 网络, 该方法首次使用 CNN 检测人脸特征点。DCNN 网络共有三层, 每层都使用了多个独立的 CNN 模型预测特征点, 最后经过平均得到本层的结果。2014 年又提出了 TCDCN<sup>[54]</sup>方法, 该方法运行速度和特征点定位的精度较好, 但是在姿态情况下特征点容易产生漂移。之后基于深度学习又提出了许多改进方法, 包括引入关键点热图的 DAN<sup>[55]</sup>方法, 借鉴人体姿态估计将边界信息引入特征点回归的 LAB<sup>[56]</sup>方法, 以及通过优化损失函数达到对特征点未知性预测的 LUVLi<sup>[57]</sup>方法等。

#### 2.2.3 头部姿态估计

头部姿态估计是利用人脸图像估计三维空间下的头部姿态角, 可表示为俯仰角(Pitch)、偏航角(Yaw)、滚转角(Roll)三个欧拉角, 如图 2.8 所示。头部姿态估计应用领域很多, 如注意力检测、行为分析、人机互动、视线追踪等, 具有很高的研究价值。目前头部姿态估计的方法很多, 有利用人脸特征点估计的方法, 也有使用支持向量机(Support Vector Machine, SVM)训练和预测的方法, 还有深度学习的方法。虽然使用深度学习的头部姿态估计有较多的研究论文, 并且有着很高的精度, 但是考虑本文方法需在嵌入式端部署, 且对精度要求不高。最终本文选择了使用人脸特征点估计头部姿态的方法, 这类方法有基于回归和基于几何的方法。基于回归的方法需要预先标定大量三维人脸特征点, 然后随机旋转平移得到更多的头部姿态, 最后通过投影可得到所有的二维特征点, 这样只需要建立二维特征点与头部姿态的映

射关系，就可通过二维特征点估计头部姿态。基于几何的方法是将二维人脸特征点与三维标准特征点相联系，利用之间的刚体变换求解旋转矩阵，进而估计头部姿态。

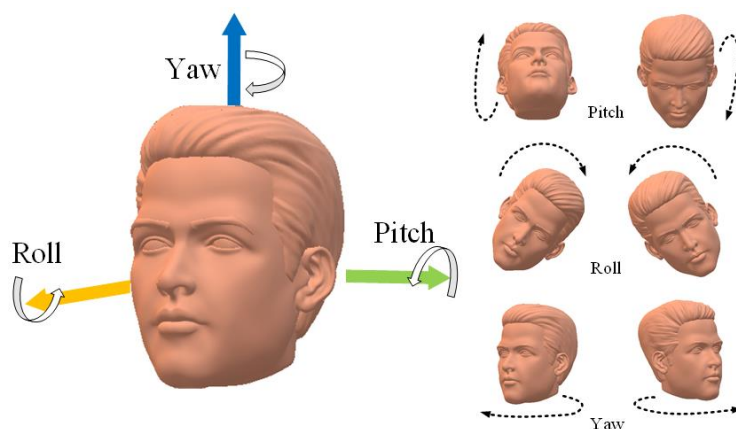


图 2.8 头部姿态示意图

## 2.3 疲劳判定指标

疲劳判定的指标有很多，本文使用驾驶员行为来判别疲劳，这主要考虑驾驶员面部的特征信息，这主要是眼睛和嘴部的状态。卡内基梅隆研究所提出了度量瞌睡的物理量 PERCLOS<sup>[58]</sup>，它指出人眼闭合时间可以作为疲劳判定指标。PERCLOS 有 EM、P70 和 P80 三种标准，他们分别表示眼睑遮挡瞳孔面积超过 50%、70% 或者 80% 视为闭眼，然后再进行参数计算。实验表明 P80 更能反映驾驶员的疲劳状态，因此本文采用了 PERCLOS 中的 P80 标准，其原理如图 2.9 所示。

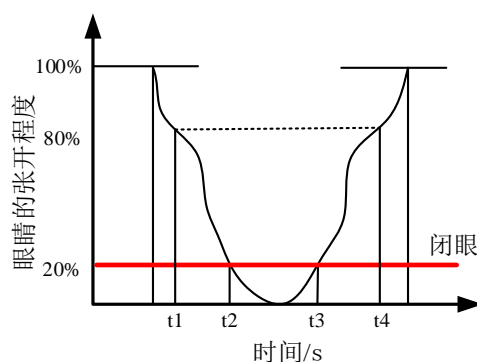


图 2.9 PERCLOS 的 P80 标准

图 2.9 是 P80 标准中一次闭眼过程中眼睛的张开程度变化，其中  $t_1$  和  $t_4$  为眼睛张开程度在 80% 的时刻， $t_2$  和  $t_3$  为为眼睛张开程度在 20% 的时刻。在计算一次闭眼行为中闭眼时间的占比时，只需计算  $t_2$  到  $t_3$  持续的时间与  $t_1$  到  $t_4$  持续时间的

比值。实际运用中每秒检测的视频帧是有限的，并且闭眼判断也存在一定的误差。为了方便计算一定时间内闭眼时间的占比，本文使用的是在一定时间范围内闭眼帧数占总检测帧数的比值。

驾驶员在处于疲劳状态时会出现打哈欠动作，因此可根据打哈欠判断驾驶员的程度，其中的判定指标包括动作持续时间、哈欠频率等。打哈欠时嘴部动作较大，并且会保持一小段时间，这相比于普通说话和唱歌有明显的区别。打哈欠频率是指在一定时间内打哈欠的次数，它可以在一定程度上反映驾驶员的疲劳程度。本文借鉴了 PERCLOS 值计算的思想，检测时将嘴部张开程度超过一定阈值视为打哈欠，计算频率时用一定时间内打哈欠帧数与总帧数的比值代替。

## 2.4 本章小结

本章对疲劳驾驶检测相关技术进行了研究和探讨，主要介绍研究过程中涉及的一些技术。首先对深度学习的一些理论知识进行介绍，主要是开发框架、卷积神经网络和目标检测网络，然后介绍面部分析中涉及的人脸检测、人脸特征点检测和头部姿态估计技术，最后介绍了本文使用的一些疲劳判定指标。



## 第3章 人脸特征点定位和跟踪算法研究

人脸特征点的定位是实现疲劳驾驶检测中一个重要的组成部分，并且采用适宜的特征点跟踪替代每一帧进行人脸检测，这有利于算法在嵌入式端的速度提升。目前人脸特征点定位方法在大姿态、部分遮挡情况下定位误差较大，因此本章基于级联回归方法的不足，将人脸特征点的定位和跟踪相结合，提出一种多视角约束级联回归的视频人脸特征点定位和跟踪算法，本章命名为 MCCR(Multi-view Constrained Cascade Regression)。该算法合理采用“分而治之”的原理，有效降低了大姿态、部分遮挡下的定位误差。

### 3.1 基于级联回归特征点定位的不足

#### 1. 级联回归方法的原理

使用级联回归的特征点定位方法是 will 一组形状回归器  $r^{(1 \dots T)}$  进行级联，回归过程如图 3.1 所示。形状  $S$  指的是人脸图像中各特征点坐标的集合，整个回归过程从初始的形状  $S^0$ （由训练集得到的平均形状）开始，经过  $T$  个回归器得到特征点形状  $S^T$ 。其中每一步是训练得到的不同回归器，每个回归器的输入都依赖前一级的输出。级联回归框架下整体的计算复杂度较低，并且能达到有效的特征点定位，因此近些年得到广泛的应用，并产生了许多改进算法。

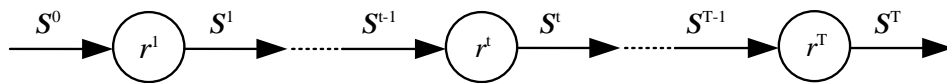


图 3.1 级联回归方法示意图

#### 2. 级联回归方法存在的不足

传统级联回归方法虽然能够实现较精确的人脸特征点定位，但是该方法在一些情况下，仍然存在以下不足：

(1) 特征点初始位置的定位通常是一个平均形状，这种定位方式较为粗糙，使得迭代的次数增加，完成特征点定位的时间变长。并且每一帧图像都需要先前的人脸检测得到人脸区域，这又使得整个特征点定位的时间变长。

(2) 级联回归方法在正脸或接近正脸的定位精度相对较高，但是在在大姿态和部

分遮挡情况下定位效果较差，这使得在复杂情况下人脸特征点出现漂移，影响后续疲劳检测的准确性。

## 3.2 MCCR 视频人脸特征点定位和跟踪

### 3.2.1 算法概述

在应对较大姿态和遮挡条件时，现有可应用于嵌入式端的人脸特征点定位算法达不到理想效果，为此本章以级联回归方法为基础，提出一种多视角约束级联回归的视频人脸特征点定位和跟踪算法，如图 3.2 为算法总体流程。

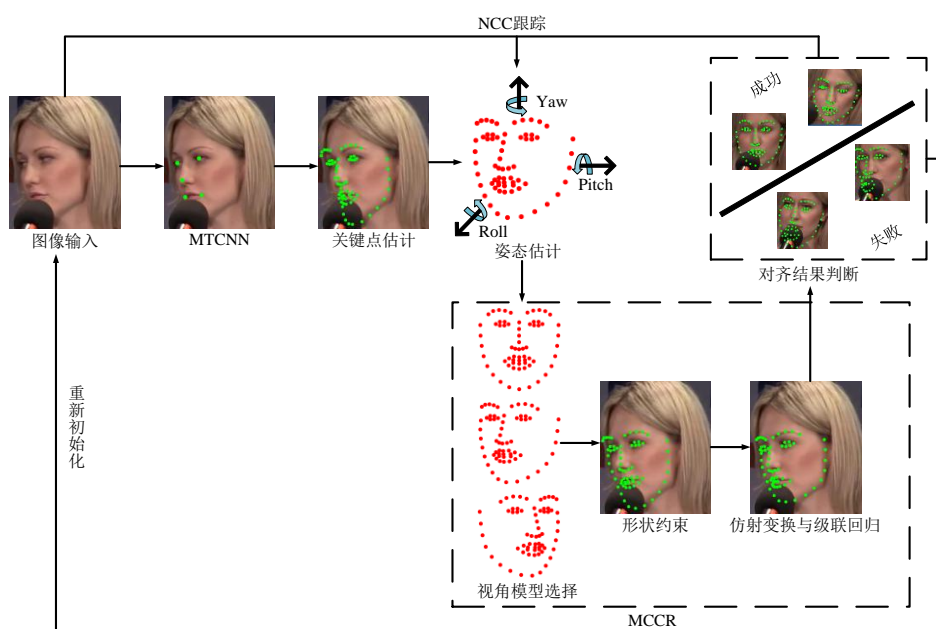


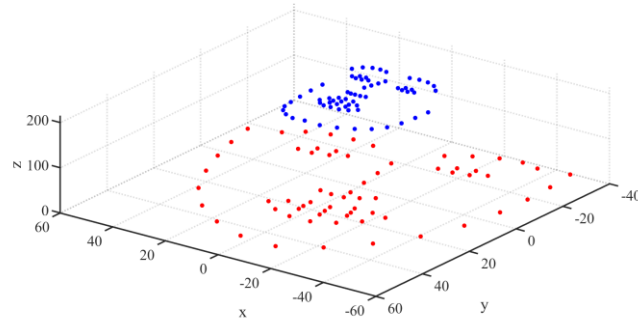
图 3.2 MCCR 算法流程

MCCR 算法开始执行时会进行一系列的初始化，首先使用预先训练的 MTCNN 人脸检测器进行人脸检测，得到初始人脸位置和 5 个初始人脸特征点。利用这 5 个稀疏特征点与三维标准特征点估计初始的二维密集特征点，详见 3.2.2 节。其次利用特征点姿态估计得到面部视角（左脸、正脸、右脸），然后在相应视角下进行特征点的约束和级联回归，将得到当前帧的人脸特征点，详见 3.2.3 和 3.2.4 节。最后对特征点对齐结果进行评分估计，当对齐分数大于系统阈值时，从当前帧到下一帧将使用 NCC 模板匹配进行特征点的跟踪，详见 3.2.5 节。当对齐分数小于阈值时，为避免后续人脸特征点跟踪过程的定位错误，本章将对算法进行重新初始化，详见在 3.2.6 节。

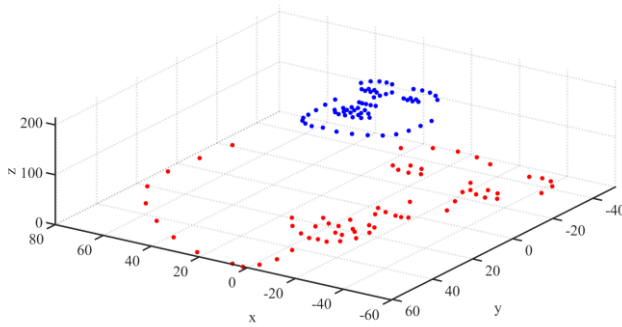


### 3.2.2 初始化

传统的级联回归算法大多数使用固定的平均形状开始回归，这种策略在应对较大的遮挡和头部姿态变化时，人脸特征点的回归定位容易错位导致初始化失败。而从 MTCNN 人脸检测器得到的 5 个特征点在人脸中最突出，并且三维特征点中保留了大量的面部形状和角度信息。综合这两个特点，本章提出一种新的初始化方法。利用一组二维稀疏点集和三维稀疏点集建立投影映射关系，并根据投影映射关系将三维密集特征点转换为二维密集特征点，从而得到对应的初始形状。这种方法将形状和角度信息引入初始形状中，有效地提高了级联回归算法的性能。如图 3.3 所示，蓝色为标准三维特征点，红色是根据投影映射关系转换的二维特征点。



(a) 正脸



(b) 侧脸

图 3.3 不同投影映射关系变换结果

在初始化时，MTCNN 人脸检测器得到了 5 个人脸特征点，联系对应三维空间下的 5 个标准特征点可建立如下关系：

$$\mathbf{S}_{3D5} \mathbf{C} = \mathbf{S}_{2D5} \quad (3.1)$$

其中， $\mathbf{S}_{3D5}$  是三维空间下 5 个标准特征点的  $x$ 、 $y$ 、 $z$  坐标集合，是一个  $5 \times 3$  的矩阵。

$\mathbf{C}$  表示投影映射矩阵， $\mathbf{S}_{2D5}$  是 MTCNN 人脸检测器得到的 5 个特征点的  $x$ 、 $y$  坐标

集合，是一个  $5 \times 2$  的矩阵。这些数据在计算前需要进行简单的归一化处理，即减去均值。式 3.1 在求解投影映射矩阵  $C$  时不一定有解，为此本章利用最小二乘的矩阵形式求解，近似成立的最小二乘解为：

$$C = (S_{3D5}^T S_{3D5})^{-1} S_{3D5}^T S_{2D5} \quad (3.2)$$

在经式 3.2 计算得到三维投影变换到二维人脸特征点的矩阵  $C$  后，即可将三维 66 特征点  $S_{3D66}$  投影变换为二维 66 特征点  $S_{2D66}$ ，如式：

$$S_{2D66} = S_{3D66} C \quad (3.3)$$

此时再与 MTCNN 人脸检测器得到的 5 个特征点计算比例、平移关系，最终经过简单微调即可得到初始形状。如图 3.4 所示，白色为 MTCNN 人脸检测器得到的 5 个特征点，红色是经本文所提初始化方法估计的 66 点。

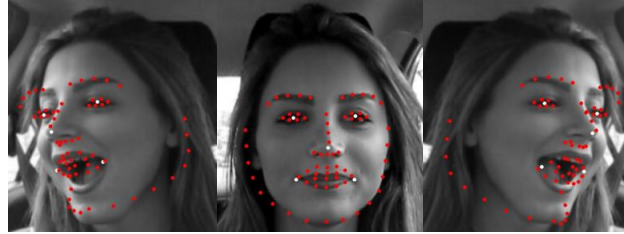


图 3.4 初始化结果

### 3.2.3 多视角约束

初始化时，过大的人脸姿态偏转会使估计的初始特征点发生部分重叠，这将对后续的特征点级联回归产生影响，如图 3.5 绿色点。为使后续级联回归更加稳定，本章对初始特征点或前一帧预测特征点进行头部姿态估计，根据估计的偏航角  $Yaw$  对人脸视角进行分类，即正脸 ( $-15^\circ < Yaw < 15^\circ$ )、左脸 ( $Yaw < -15^\circ$ ) 和右脸 ( $Yaw > 15^\circ$ )。然后组建不同视角下不同形状增量组成的约束模型，并对相应视角下的特征点进行形状调整。最后在相应视角下使用各自训练好的级联形状回归模型，从而完成人脸特征点的定位。

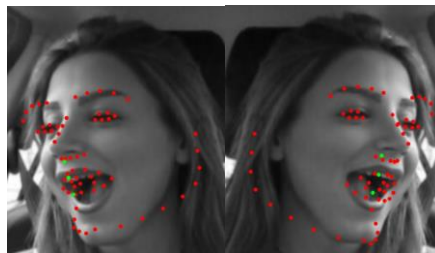


图 3.5 特征点部分重叠

## 1. 头部姿态估计

人的头部姿态可由偏航角(Yaw)、俯仰角(Pitch)和滚转角(Roll)三个欧拉角表示, 它们与旋转矩阵  $\mathbf{R}$  之间有如下关系:

$$\mathbf{R} = \mathbf{R}_x(\alpha) \mathbf{R}_y(\beta) \mathbf{R}_z(\gamma) \quad (3.4)$$

其中,

$$\mathbf{R}_x(\alpha) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \alpha & -\sin \alpha \\ 0 & \sin \alpha & \cos \alpha \end{bmatrix} \quad (3.5)$$

$$\mathbf{R}_y(\beta) = \begin{bmatrix} \cos \beta & 0 & -\sin \beta \\ 0 & 1 & 0 \\ \sin \beta & 0 & \cos \beta \end{bmatrix} \quad (3.6)$$

$$\mathbf{R}_z(\gamma) = \begin{bmatrix} \cos \gamma & \sin \gamma & 0 \\ -\sin \gamma & \cos \gamma & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (3.7)$$

为此可以根据旋转矩阵求得欧拉角, 已知二维空间下的人脸点形状和三维空间下的标准人脸点模型, 可建立如下数学形式:

$$(\mathbf{R}, \mathbf{T}) = \arg \min_{\mathbf{R}, \mathbf{T}} \sum_{i=1}^n w_i \|\mathbf{R} \mathbf{q}_i + \mathbf{T} - \mathbf{s}_i\|^2 \quad (3.8)$$

其中,  $\mathbf{R}$  为旋转矩阵,  $\mathbf{T}$  为平移矩阵,  $w_i$  表示每个点对之间的权重,  $\mathbf{S} = \{\mathbf{s}_1, \mathbf{s}_2, \dots, \mathbf{s}_n\}$  为二维人脸特征点集合,  $\mathbf{Q} = \{\mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_n\}$  为三维空间下标准人脸模型的点集合。

根据式 3.4~式 3.8, 可通过奇异值分解方法求得旋转矩阵  $\mathbf{R}$  和其对应的欧拉角  $\alpha$ 、 $\beta$ 、 $\gamma$ , 在得到欧拉角之后还需要进行相应的弧度转换, 如下式:

$$\text{Yaw} = \frac{180 \times \alpha}{\pi}, \text{Pitch} = \frac{180 \times \beta}{\pi}, \text{Roll} = \frac{180 \times \gamma}{\pi} \quad (3.9)$$

## 2. 形状约束

在平均形状的基础上变化的一个形状可以构造如下关系:

$$\mathbf{e}_l(\mathbf{S}_l - \mathbf{S}_{mean}) = \mathbf{P}_l \quad (3.10)$$

其中,  $\mathbf{S}_l$  均值为 0, 它是根据平均形状变化而来的人脸形状。  $\mathbf{S}_{mean}$  为平均形状, 两者皆为特征点  $x$ 、 $y$  坐标组成的  $2N \times 1$  的列向量, 两者的差为形状增量

$\Delta S = [x_1, y_1, \dots, x_i, y_i]^T$ ,  $i \in (1, N)$ , 为特征点数。 $e_i = [\frac{x_1}{\sqrt{2N}\sigma}, \frac{y_1}{\sqrt{2N}\sigma}, \dots, \frac{x_i}{\sqrt{2N}\sigma}, \frac{y_i}{\sqrt{2N}\sigma}]$ , 表示由形状增量构建的单位增量,  $\sigma$  为形状增量  $\Delta S$  的标准差, 此时可计算出  $P_i = \sqrt{x_1^2 + y_1^2 + \dots + x_i^2 + y_i^2} = \|\Delta S\|_2$ , 表示为增量倍数。

由上述可知, 一个在平均形状附近变化且均值为 0 的形状, 可以求得其单位增量  $e_i$ , 结合式 3.10 可建立如下关系:

$$e_i^T P'_i + S_{mean} = S'_i \quad (3.11)$$

根据上述关系在单位增量不变的情况下, 给定不同的增量倍数  $P'_i$  后, 叠加平均形状可以得到不同的人脸形状  $S'_i$ 。如图 3.6 分别为  $P'_i = P_i$ 、 $2P_i$ 、 $3P_i$ 、 $-P_i$ 、 $-2P_i$ 、 $-3P_i$ , 并且单位增量为嘴部变化时的人脸形状变化。

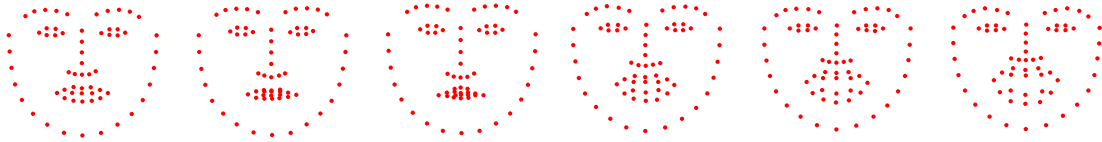


图 3.6 不同增量倍数叠加后的人脸形状

由此使用多个人脸形状时, 可构建如下矩阵公式:

$$V^T P' + S_{mean} = S' \quad (3.12)$$

其中,  $V = [e_1^T, e_2^T, \dots, e_k^T]^T$ ,  $k$  为脸型形状总数,  $P' = [P'_1, P'_2, \dots, P'_k]^T$ , 为不同增量倍数组成的向量,  $S'$  为叠加后的形状, 可见改变  $P'$  可得到不同的叠加形状。

根据式 3.12, 对于一个需要调整的形状  $S$ , 可以计算得到一个基于  $S$  的增量矩阵  $P$ , 即式:

$$(VV^T)^{-1}V(S - S_{mean}) = P \quad (3.13)$$

由此任意形状  $S$  可视为单位增量基础上不同增量倍数的叠加, 若对不同单位增量的增量倍数加以一定的限制范围, 即可对特征点形状进行调整。本章规定了每个形状上增量倍数的最大范围, 增量倍数超过这个范围就将其纠正为这个形状上的最大增量倍数, 最后根据式 3.12 即可得到调整后的形状。

本章在正脸视角下建立了 24 种不同的形状, 如图 3.7 所示。约束只在构建的这些形状下约束, 但不仅局限于这 24 种。其中第一行前两个形状是在平均形状基础上进行缩放, 这样就得到了两个不同方向的缩放单位增量, 即可在不同的增量倍

数叠加后得到不同大小和不同旋转角度的形状。最后一行作为形状调整使用，变化相对平均形状较小。

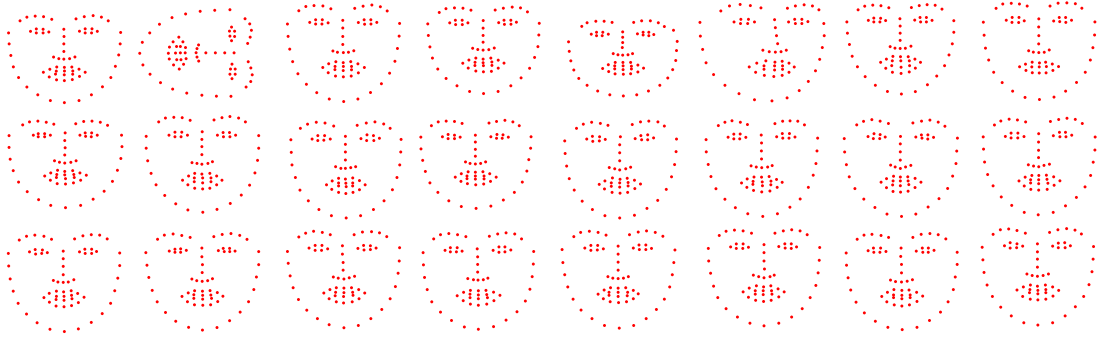


图 3.7 24 种不同增量的形状

利用上述方法可对三个视角下的估计形状进行调整，部分结果如图 3.8 所示。可见第一行重叠部分的点得到了调整，人脸整体形状也得到了部分调整，且在平均形状附近变化，有利于下一步级联回归的进行。



图 3.8 调整前后的形状

### 3.2.4 多视角级联形状回归

由多视角约束调整后可得到级联回归的初始形状，此时初始形状的大小，头部滚转角 Roll 是多样的，为此本章引入仿射变换对级联回归前的形状进行简单统一，使得级联回归效果更佳。对于只进行平移、旋转、缩放的仿射变换，可表示为：

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} s\cos\theta & -s\sin\theta & T_x \\ s\sin\theta & s\cos\theta & T_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \quad (3.14)$$

其中， $s$  为等比例缩放因子， $\theta$  为旋转角度， $T_x$  为水平平移量， $T_y$  为垂直平移量。

$(x, y, 1)^T$  是源坐标系下的坐标， $(x', y', 1)^T$  是目标坐标系下的坐标。

在相应人脸视角下, 结合本视角的标准形状和级联回归前的初始形状计算仿射变换矩阵  $M$ , 即 3.14 式中的  $3 \times 3$  矩阵, 然后利用仿射变换可得到简单统一后的特征点坐标, 如图 3.9 为三个视角下的仿射变换结果。



图 3.9 仿射变换前后

此时, 统一后的特征点即可作为级联回归的初始形状。本章在人脸图像中计算 HOG 特征, 并根据偏航角 Yaw 选择最佳视角模型进行级联回归, 算法如下式:

$$\hat{S}^{(t+1)} = \hat{S}^{(t)} + r_t(I, \hat{S}^{(t)}) \quad (3.15)$$

其中,  $\hat{S}^{(t)}$  是由特征点坐标组成的向量, 它是上层回归器的输出结果,  $t$  表示回归的次数,  $I$  为输入图像,  $r_t$  表示当前层数的回归器。

回归器的计算过程为:

$$r_t(I, \hat{S}^{(t)}) = W^t \Phi^t(I, \hat{S}^{(t)}) \quad (3.16)$$

其中,  $W^t$  是一个线性回归矩阵,  $\Phi^t$  是根据图像和上一层特征点位置提取的图像特征信息。图 3.10 是该过程的一个实例, 展示了每次迭代过程特征点的变化。

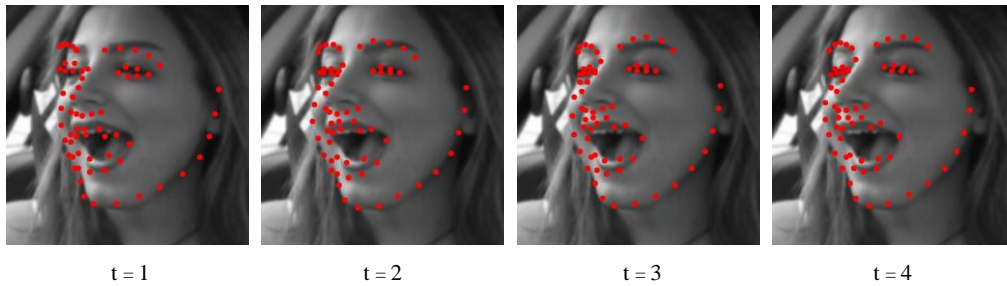


图 3.10 特征点回归过程

最终输出的特征点坐标是仿射变换后图像中的位置坐标, 对坐标进行一个逆变

换即可得到原图像中的特征点坐标, 如式 3.17 所示,  $x'_i$  是仿射变换后的坐标系下的位置坐标,  $x_i$  是源坐标系下特征点位置的坐标。

$$x_i = \mathbf{M}^{-1} x'_i \quad (3.17)$$

### 3.2.5 模板匹配跟踪

在视频特征点定位中, 通常需要在每一帧图像上检测人脸, 而本章算法中则还需要进行初始化, 这些需要耗费大量时间。此外, 连续帧的图像特征点之间坐标信息相差不大, 建立一种关联是非常重要的。因此本章引入 NCC 模板匹配跟踪算法, 使用它计算上一帧特征点在本帧图像中的位置偏移量, 进行初步的跟踪。前一帧特征点叠加偏移量后即可得到当前帧的初始形状, 并可在下一步继续使用 MCCR 进行特征点定位。使用这种方法不仅省略了每一帧都进行人脸检测和初始化, 而且还建立了连续帧之间的关联。整个人脸特征点跟踪过程如图 3.11 所示, IS 为初始形状, PS 为预测形状。

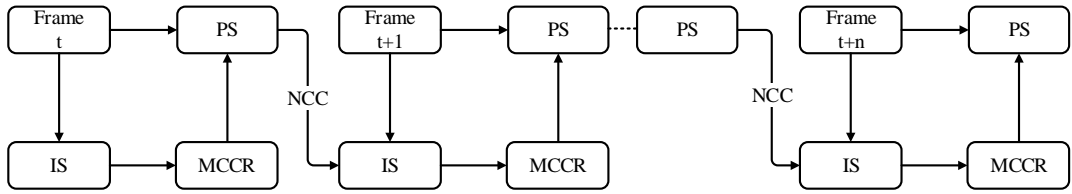


图 3.11 NCC 跟踪

NCC 模板匹配算法也称为归一化互相关匹配算法, 它适用于没有几何误差的情况。本章通过计算得出互相关值来确定目标的坐标位置, 假设搜索图像  $I$  的尺寸为  $M \times M$ , 模板  $T$  的尺寸为  $N \times N$ , 其中  $M > N$ , 单位为像素。模板  $T$  在搜索图像  $I$  上平移,  $I^{i,j}$  为模板在搜索图像所覆盖的子图, 子图在搜索图像  $I$  中左上角顶点的坐标为  $(i, j)$ 。在实际匹配中, 模板和搜索图像的相似性通过度量函数来衡量, 则归一化互相关匹配度量定义为:

$$R(i, j) = \frac{\sum_{m=1}^M \sum_{n=1}^N I^{i,j}(m, n) T(m, n)}{\sqrt{\sum_{m=1}^M \sum_{n=1}^N (I^{i,j}(m, n) - \bar{I}^{i,j})^2} \sqrt{\sum_{m=1}^M \sum_{n=1}^N (T(m, n) - \bar{T})^2}} \quad (3.18)$$

在本章算法中, 首先从前一帧图像中截取包含所有特征点的最小人脸, 并缩放为  $48 \times 48$  的图片作为匹配模板, 如图 3.12(a)。然后在当前帧中, 将截取前一帧模板



位置为中心向外扩大一倍的区域，并缩放为  $96 \times 96$  作为待匹配区域，如图 3.12(b)。最后使用 NCC 模板匹配得到模板在匹配区域的偏移量，即为特征点的偏移量，以前一帧特征点加上偏移量可得到当前帧的初始形状，如图 3.12(c)，绿色为 NCC 模板匹配跟踪之前、红色 NCC 模板匹配跟踪之后。



图 3.12 NCC 模板匹配跟踪结果

### 3.2.6 重新初始化

多视角约束级联回归被用来预测每一帧图像的特征点位置，而 NCC 模板匹配跟踪被用来建立连续帧之间的联系。两个步骤同时工作时，前一帧可靠的对齐结果可以准确的预测当前帧的特征点位置。当前一帧的对齐结果有漂移时，会导致本帧特征点预测出现偏移或丢失。因此，采用一种重新初始化机制有效、准确地应对特征点定位失败问题是必不可少的。在这项工作中，本章引入了对齐分数，它对应对齐的优度。当对齐分数低于设定的阈值  $m$  时，算法将重新初始化。为此，本章使用 SVM 训练了一个分类器，根据最后的形状索引特征来区分对齐和未对齐的图像。本章从标注的样本中生成正样本，然后在实际特征点周围随机生成负样本，以训练后的 SVM 的得分作为判断对齐效果的标准。在本章实验中，所有特征点的平均对齐分数高于  $m$  被认为是一个成功的特征点定位。

如图 3.13 所示，给定一个人脸视频，如果前一帧人脸特征点的平均对齐分数低于  $m$ ，MTCNN 人脸检测器将在下一帧进行人脸检测，并重新根据 5 个特征点估算得到初始形状。如果前一帧定位成功，将不会再次进行初始化，而是结合下一帧进行 NCC 跟踪。这种重新初始化机制不仅减少了重复初始化带来的计算消耗，而且保障了 NCC 跟踪时特征点的定位优度。



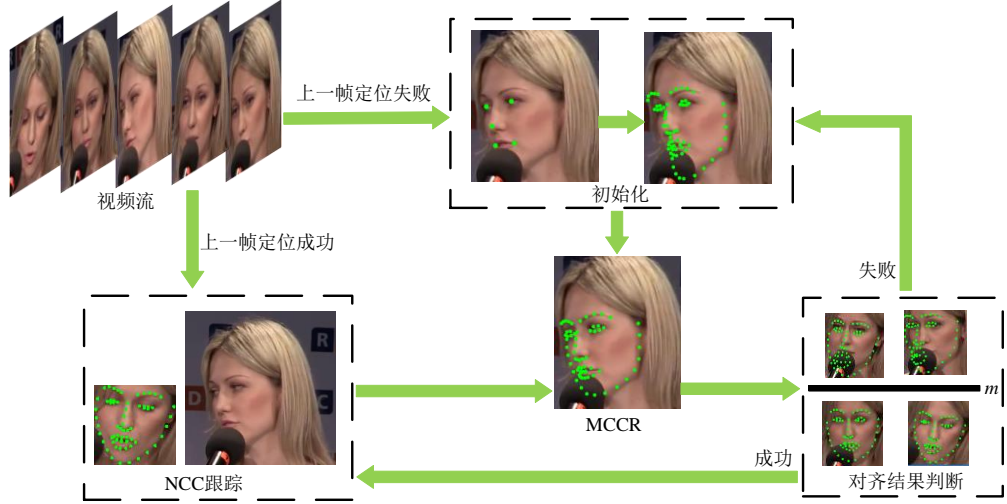


图 3.13 重新初始化步骤

### 3.3 实验结果分析

MCCR 视频人脸特征点定位和跟踪算法构建完成后，需要制作大量的训练数据集。但是目前已经公开的数据集中，没有关于夜晚近红外摄像头的数据样本。因此，本章在公开数据集 300W 人脸图片数据集和 300VW 人脸视频数据集基础上，增加了真实环境下拍摄的近红外人脸图像和视频并手工标注。除此之外，算法还有几个自由参数：回归器级联阶数  $T$ 、对齐分数阈值  $m$ 、形状的增量倍数限制范围  $P_\lambda$ ，经过大量测试和验证，最后设置  $T=4$ ， $m=0.5$ ， $-2P < P_\lambda < 2P$ （ $P$  为构建多视角约束模型时各形状的增量倍数）。

#### 1. 静态图像结果

人脸特征点的定位通常用归一化平均误差(Normalized Mean Error, NME)来评价。特别地，利用平均欧几里得距离归一化误差，如式 3.19：

$$E = \frac{\frac{1}{N} \sum_{j=1}^N \|x_j - g_j\|_2}{\|l - r\|_2} \quad (3.19)$$

其中， $E$  是平均欧几里得距离归一化误差， $N$  是特征点总数， $x$  是预测特征点的位置， $g$  是真实特征点的位置。在静态图像的评估中， $l$  和  $r$  分别是左眼中心和右眼中心的位置。

本章首先在静态图像中评估了 MCCR，以显示此方法在单张图像上人脸特征点定位的性能。注意，在测试时是将约束后的估计形状作为级联回归的初始形状。算法将与一些常用方法进行比较，包括 DRMF<sup>[59]</sup>，GN-DPM<sup>[60]</sup>，ERT，ESR，

RCPR<sup>[61]</sup>, SDM, LBF。实验结果如表 3.1 所示, 可以观察到本章所提出的方法在这些数据集上优于大多数以前的方法, 特别是在挑战集上有明显的性能优势。

表 3.1 平均误差与现有方法的比较

方法	Lfpw	Helen	300W		
			common	challenge	full
DRMF	6.57	6.70	6.65	19.79	9.22
RCPR	6.56	5.93	6.18	17.26	8.35
ESR	—	—	5.28	17.00	7.58
SDM	5.67	5.50	5.57	15.40	7.50
ERT	—	—	—	—	6.40
LBF	—	—	<b>4.95</b>	11.98	6.32
LBF fast	—	—	5.38	15.50	7.37
GN-DMP	5.92	5.69	5.78	—	—
MCCR	<b>5.51</b>	<b>5.31</b>	5.39	<b>9.72</b>	<b>6.24</b>

图 3.14 是本章方法的人脸特征点定位结果示例, 其中第一行图像来自 Lfpw 数据集, 第二行图像来自 Helen 数据集, 第三行来自 300W 的 Ibug 挑战集。可以注意到, 本章方法在这些数据集中遇到表情、姿态、光照和部分遮挡的巨大变化时, 具有较强的鲁棒性。

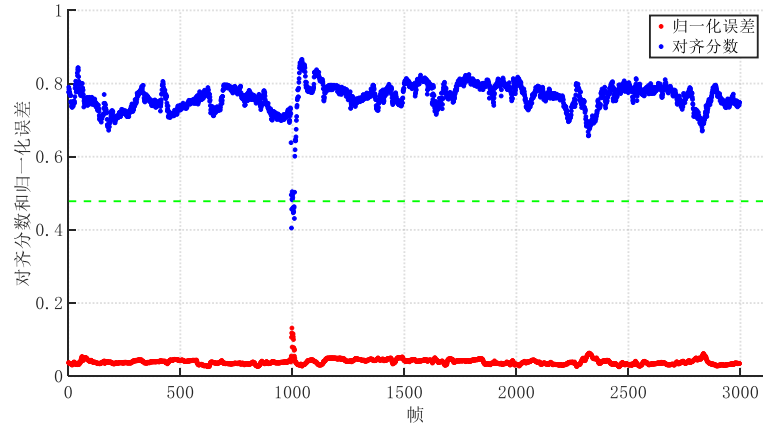


图 3.14 MCCR 的示例图像

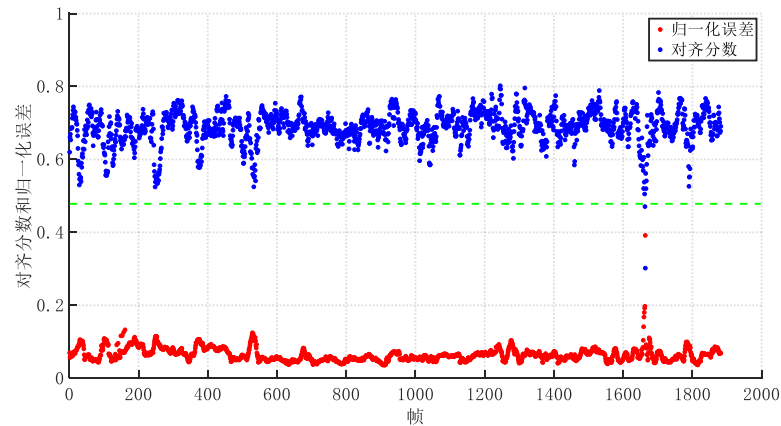
## 2. 评估重新初始化

如果先前的对齐结果出现特征点漂移, 将把错误的信息传递给下一帧, 造成特征点预测失败。本实验中当对齐分数低于设定的阈值  $m=0.5$  时, 进行人脸检测和重新初始化。对齐分数在 0.5 以上的特征点定位被认为是一个成功定位, 将进行 NCC 模板匹配跟踪。

图 3.15 中的实验结果是 300VW 测试集中随机选择的 2 个视频，本章测试并绘制了每一帧的 NME（以瞳孔距离归一化）及对应的对齐分数，其中对齐分数为蓝色，NME 为红色，绿线为对齐分数阈值线。结果表明，本章使用的重新初始化机制是合理有效的，当 NME 增大时，预测形状的对齐分数变低，这种情况是本文想要达到的效果。



(a) 视频序号 125



(b) 视频序号 407

图 3.15 重新初始化验证结果

### 3. 评估 NCC 模板匹配跟踪

为评估 NCC 模板匹配跟踪在视频特征点定位中的效果，本章从 300VW 测试集中挑选了一个人脸区域偏移较大的视频（视频序号为 557），绘制了基于该视频的累计误差曲线和每一帧的 NME（以眼角距离归一化），并分别对使用 NCC 模板匹配跟踪和未使用 NCC 模板匹配跟踪（即直接将上一帧级联回归结果作为当前帧的初始形状）进行测试。测试结果如图 3.16 所示，实验表明引入了 NCC 模板匹配跟踪后系统的性能得到了提高，并且在人脸区域出现较大偏移时更加的稳定。

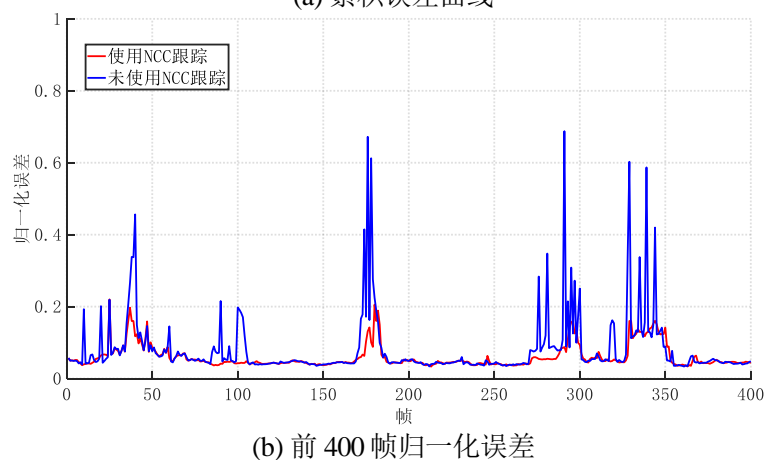
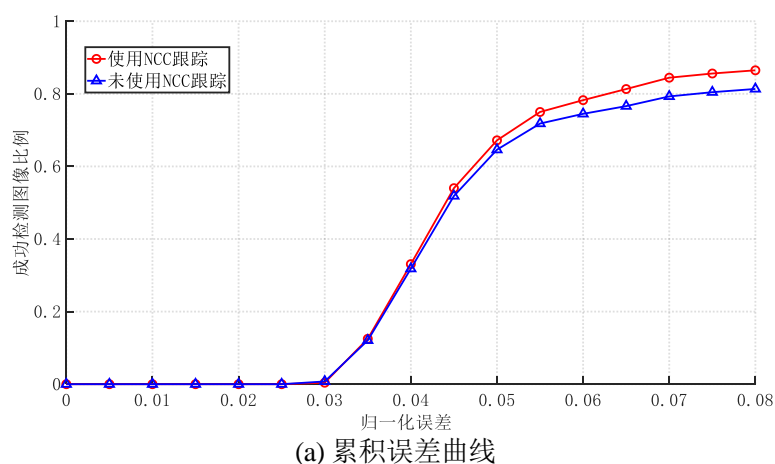


图 3.16 累积误差曲线和前 400 帧归一化误差

#### 4. 评估整体性能

为测试算法的整体性能，本章在 300VW 和自采集的驾驶视频中进行实验。如图 3.17 展示了部分视频中的对齐结果，其中最后一行为自采集近红外视频。可见在面临光照、姿态变化和遮挡的挑战时，本章算法可以很好地处理这些困难。

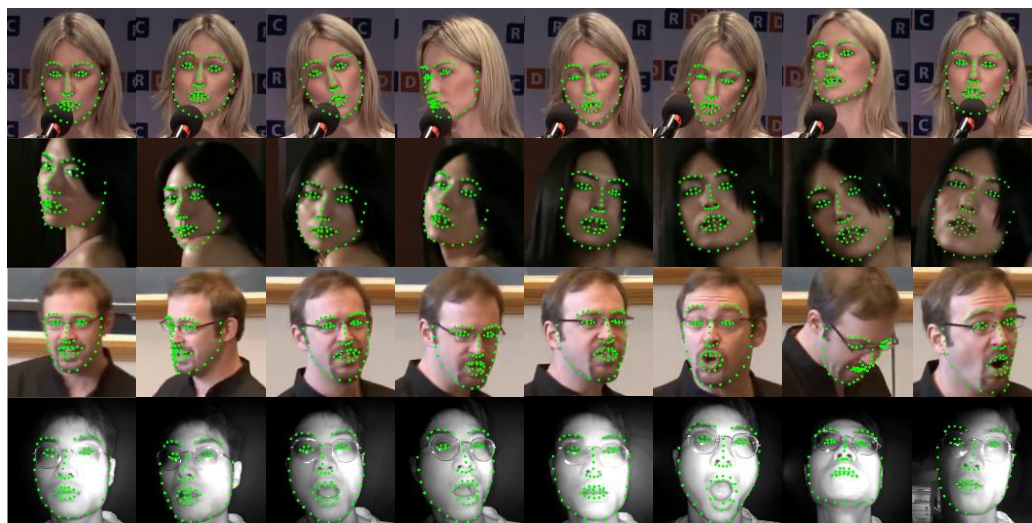


图 3.17 300VW 结果

### 3.4 本章小结

本章在级联回归方法的基础上，将特征点的定位和跟踪相结合，研究和设计了一种多视角约束级联回归的视频人脸特征点定位和跟踪算法。该算法涉及初始化、多视角约束、多视角级联形状回归、NCC 模板匹配跟踪、重新初始化等步骤。实验证明本章提出的方法在应对视频人脸图像的大姿态、部分遮挡和光照变化时，具有很强的鲁棒性。





## 第4章 疲劳驾驶状态检测算法研究

根据 MCCR 算法预测人脸特征点之后，即可定位眼睛和嘴巴的位置。本章主要是对眼睛和嘴部状态进行分析，为闭眼和打哈欠特征参数的计算提供技术支撑，最后使用这些疲劳信息完成对驾驶员疲劳状态的等级划分。

### 4.1 算法概述

疲劳驾驶状态检测算法主要包括眼睛状态识别、嘴部状态识别和疲劳状态等级划分，其总体算法流程如图 4.1 所示。

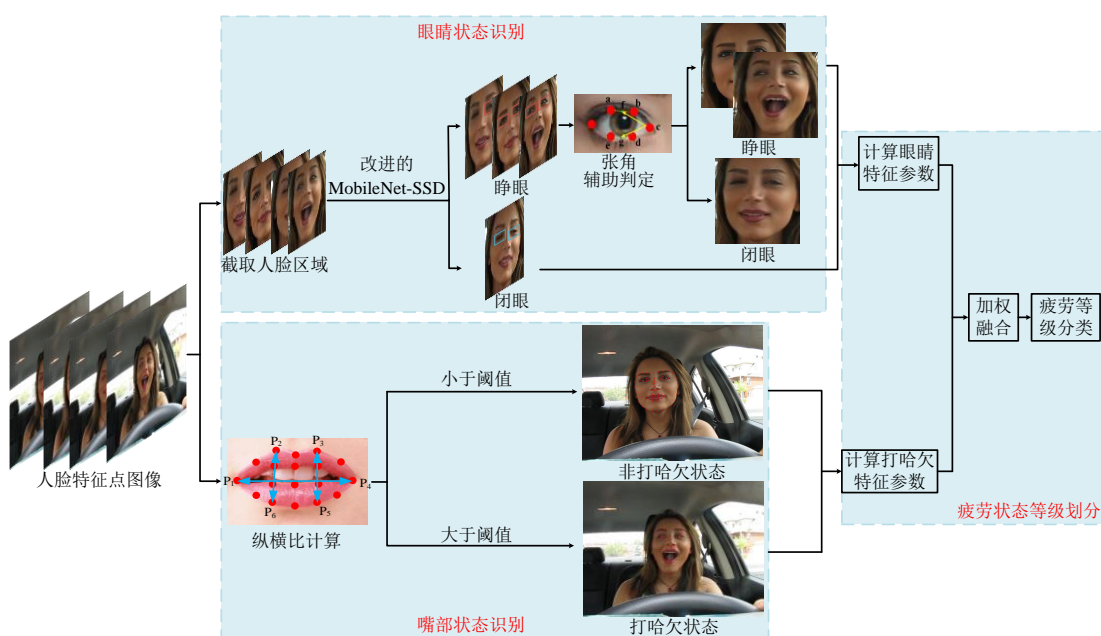


图 4.1 疲劳驾驶状态检测算法流程

疲劳驾驶状态检测算法执行时，首先根据人脸特征点计算人脸区域，并进行眼部状态识别。本章为适应更多复杂光照环境和眼镜干扰，将眼部状态识别转化为目标检测问题，并针对目标检测网络 MobileNet-SSD 进行优化改进。同时为提升瞳孔部分遮挡时眼睛状态识别的准确性，本章结合人脸眼睛特征点引入眼睛张角辅助判定。最后在计算眼睛特征参数时，采用常用的闭眼 PERCLOS 参数。

其次进行嘴部状态识别时，本章结合人脸嘴部的特征点，使用嘴巴纵横比衡量嘴巴的张开程度，当嘴巴张开程度超过设定阈值时，嘴巴状态为打哈欠状态。为了更准确的区分打哈欠与说话、唱歌等动作的区别，本章基于纵横比和持续帧

数对打哈欠动作进行判定，即计算嘴巴打哈欠特征参数时，使用嘴巴张开较大程度持续帧数占总检测帧数的百分比。

最后进行疲劳状态等级划分时，本章将闭眼 PERCLOS 参数和打哈欠特征参数进行加权融合，以得到加权疲劳参数，最终根据加权疲劳参数所在范围进行疲劳分类。

## 4.2 眼部状态识别

### 4.2.1 常用眼部状态识别方法的不足

目前通常使用的眼睛状态识别方法包括眼睛纵横比(Eye Aspect Ratio, EAR)、灰度投影、Hough 找圆和模板匹配。眼睛纵横比主要是根据已经得到的眼部人脸特征点计算，如图 4.2 所示，每个特征点有对应的编号，则可由公式 4.1 计算眼睛的纵横比。

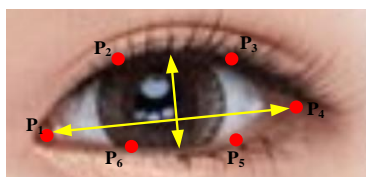


图 4.2 眼睛纵横比

$$EAR = \frac{\|p_2 - p_6\| + \|p_3 - p_5\|}{2\|p_1 - p_4\|} \quad (4.1)$$

驾驶员人脸图像采集时，摄像头的位置固定不变，并且眼睛目标较小，此时驾驶员头部姿态将对眼睛纵横比产生较大影响。如图 4.3 所示，头部姿态偏航角 Yaw 变化时，眼睛的纵横比也随之改变，闭眼与睁眼之间准确的纵横比阈值界限是难以确定的。因此使用 EAR 判断睁闭眼的方法，判断条件单一易产生误判。

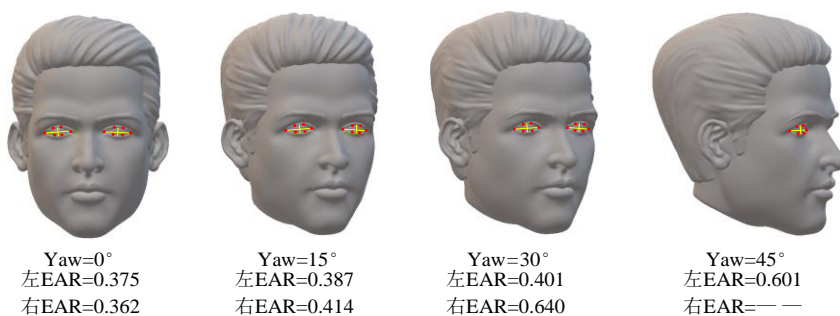


图 4.3 头部姿态与 EAR 的变化



通过灰度投影判断眼睛的状态，它的原理是将眼睛部位的灰度图像进行水平投影。如图 4.4 所示，正常情况下，睁眼状态的灰度投影会有波峰且较宽，而闭眼状态时没有波峰或波峰较窄，因此利用波峰的宽度大小可判断眼睛状态。

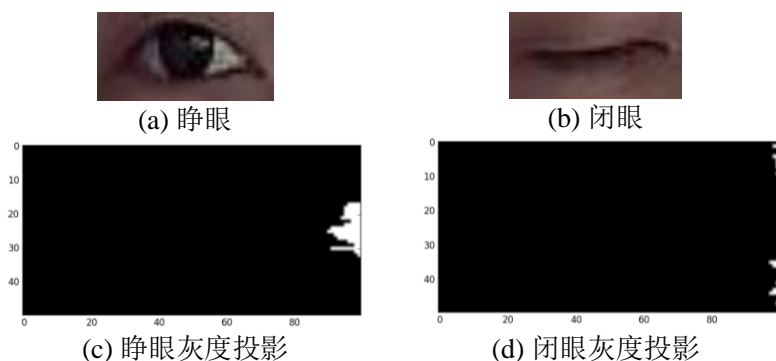


图 4.4 眼睛的灰度投影

通过实际测试，灰度投影也存在着一些不足。如对光照比较敏感，在复杂环境下对眼睛状态的判断准确率较低。如图 4.5 所示，当面部光照不均匀出现弱光、侧光、夜间近红外人脸图像眼睛亮瞳等现象时，可能导致灰度水平投影形成的波峰界限不明显。

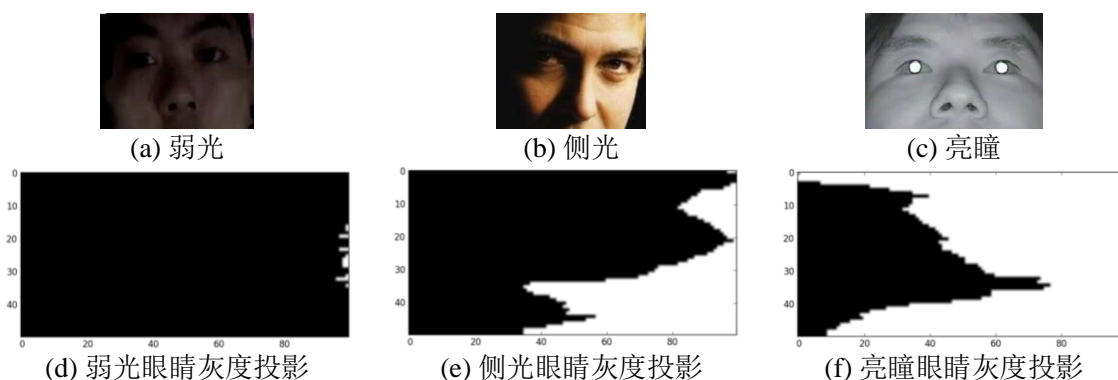


图 4.5 复杂环境下眼睛灰度投影

Hough 找圆的方法是通过 Hough 变换寻找眼睛图像中的圆形，眼睛睁开时瞳孔可视为圆形，闭眼则没有圆形。由图 4.6 所示，当正常睁眼时，利用 Hough 变换可在图像中找到圆形瞳孔，但是在一些瞳孔部分遮挡的情况下，利用 Hough 变换检测不到圆形，这种状态应该是处于睁眼状态，因此利用 Hough 找圆的方法判断眼睛状态准确率较低。

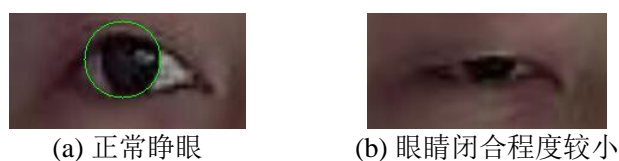


图 4.6 基于 Hough 方法的检测结果

模板匹配的方法是使用闭眼和睁眼的标准图片模板与目标图片比较计算相似度，若待测图像与闭眼模板相似度高，则是闭眼。由于眼睛部位光照、角度和遮挡等的不同，需要匹配大量的模板才能得到准确的结果，这使得计算时间变长，不利于嵌入式设备的集成。

## 4.2.2 基于改进 MobileNet-SSD 的眼部状态识别

眼睛状态识别需要应对不同角度、不同光照、眼镜遮挡和不同张开程度等情况，而传统的方法无法完全应对这些干扰因素，为此本章提出一种基于改进 MobileNet-SSD 的眼部状态识别方法，这种方法训练得到的网络模型小、检测速度快，为本文在实现阶段移植到嵌入式设备提供了基础。

### 1. 网络改进

SSD 网络在 2016 年首次提出，是目前 one-stage 方法中比较常用的目标定位与识别网络。它在不同层的特征图上设置不同尺寸和长宽比的默认框，预测时会每个默认框中的每个存在对象类别进行打分，并对框进行调整以更好的匹配对象形状。网络对不同分辨率的多个特征图的预测结果进行组合，越在前面的检测层对小目标的分类与定位效果越好，适用于多尺寸目标的检测。SSD 原论文中使用 VGG-16 作为整个网络的主干网络，而使用针对移动端提出的轻量级深度网络 MobileNet\_v1 作为主干网络时，参数量相对原网络大大减少，并且网络的准确率相差不大。结合本文在嵌入式端的实际应用，最后选择了 MobileNet\_v1 为主干网络的 MobileNet-SSD 网络，网络结构如图 4.7 所示。

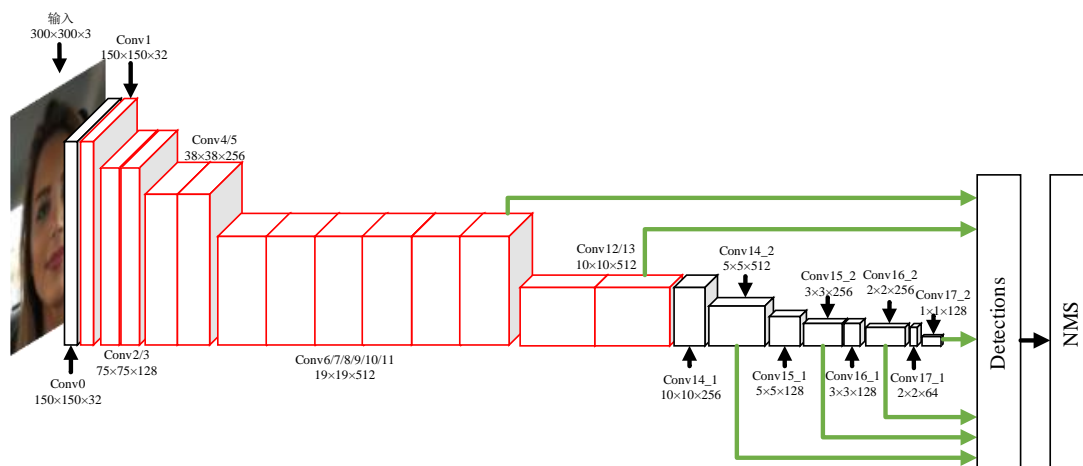


图 4.7 MobileNet-SSD 网络结构

在使用 MobileNet-SSD 网络进行训练时，本章使用近红外摄像头采集白天和夜晚的人脸图像，并根据人脸所在位置进行裁剪，得到包含人脸的多个尺寸图像作为数据集。数据集总共 13282 张，其中白天睁眼 3625 张，晚上睁眼 3107 张，白天闭眼 3461 张，晚上闭眼 3089 张。将图片大小统一缩放至 300×300 像素时眼睛的长宽分布如图 4.8 所示。

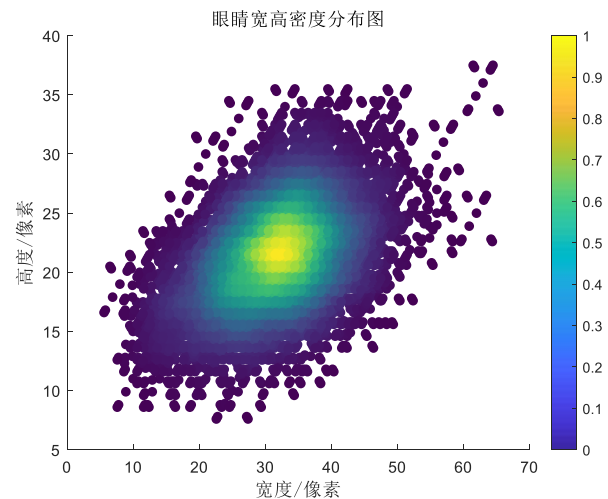


图 4.8 数据集中眼睛长宽分布图

小目标有两种定义方式，一种是相对尺寸定义，即目标尺寸小于原图像尺寸的 0.1，另外一种为绝对尺寸定义，即目标尺寸小于 32×32 像素。从图 4.8 可知本章数据集中的眼睛目标大部分是小目标，而 MobileNet-SSD 是一个多尺度检测的网络，前面的检测层用于检测小目标，后面的检测层在检测大目标时很可能没有使用。为了验证这个猜想，本章使用采集得到的数据集训练网络，训练到模型收敛后统计了 6 个检测层的输出占比结果，如表 4.1 所示。

表 4.1 网络检测层输出统计结果	
输出层	输出统计(%)
Conv11	91.3
Conv13	8.7
Conv14_2	0
Conv15_2	0
Conv16_2	0
Conv17_2	0

由表 4.1 的统计结果显示，MobileNet-SSD 网络从 Conv14\_2 开始，后面的检测层输出占比为 0，据此可知在对眼睛目标进行检测时，该网络结构存在冗余。

由于 MobileNet-SSD 网络检测层是按照线性串联的, 将后面用于检测大目标的检测层结构去掉不会对前面小目标检测层产生影响, 所以本章对 MobileNet-SSD 网络进行冗余架构裁剪, 改进后的 MobileNet-SSD 网络整体架构如图 4.9 所示, 把原有网络中 Conv13 以后的网络结构全部去掉, 并保留了生成分类和定位检测框的 Detections 层和非极大值抑制(Non-Maximum Suppression, NMS)部分。

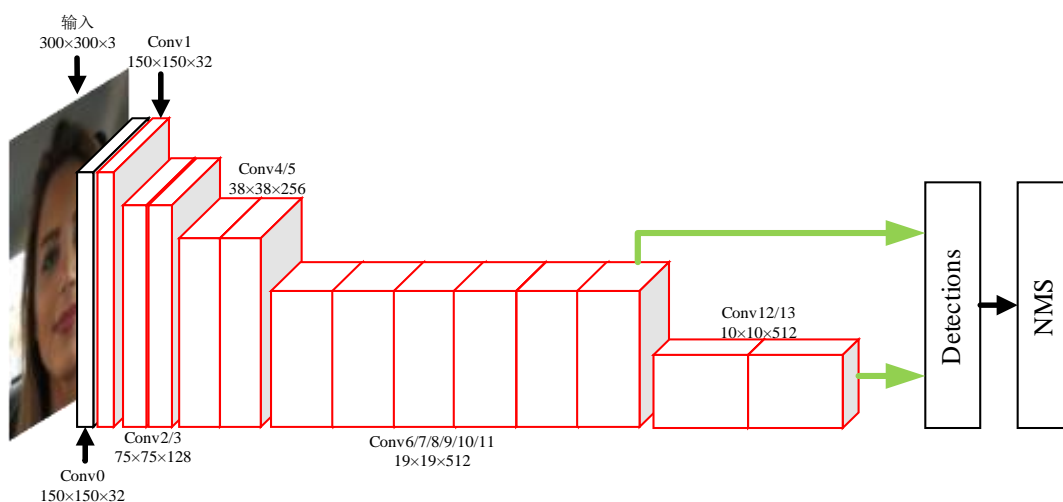


图 4.9 改进网络的架构图

## 2. 网络检测优化

针对眼睛数据集进行网络改进后, 本章继续对网络的实际检测进行优化研究。由于眼睛区域的特征存在睁开和闭合两种状态, 因此在对数据集进行标注时分为睁眼和闭眼两类。检测时在同一区域检测两种类别, 提高了模型对睁闭眼检测的能力。

在对近红外摄像头采集的视频图像帧进行检测时, 本章为提高检测的准确性, 在不改变检测图像尺寸的同时, 增大人脸区域的尺寸, 这样在不降低模型检测速度的情况下提高了检测精度。增大人脸区域的优化流程如图 4.10 所示, 首先利用特征点定位算法得到人脸关键点, 再通过关键点计算并截取出人脸区域图像。为了对应网络的输入, 需把截取的图像统一缩放为  $300 \times 300$  的尺寸, 然后使用改进的 MobileNet-SSD 网络模型进行眼睛状态的判断。只对人脸区域的图像进行检测不仅增大了检测尺寸, 而且在不改变检测速度的同时提高了抗干扰能力。并且这种方式在检测前可对是否为驾驶员人脸进行判断, 如始终对摄像头最近的人脸区域进行检测, 这样极大地减少了背景干扰。部分检测结果如图 4.11 所示。

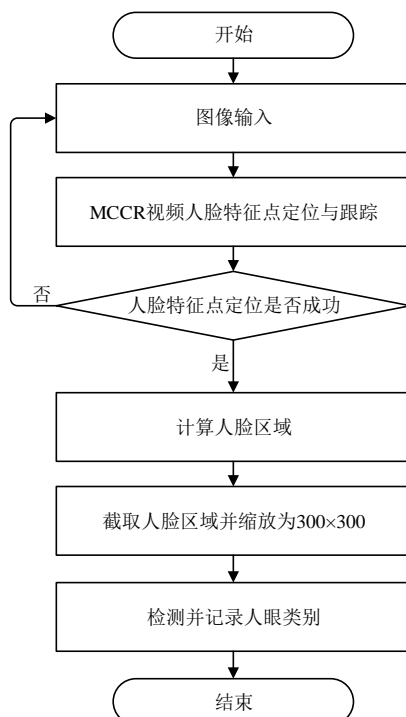


图 4.10 增大人脸检测方法流程图

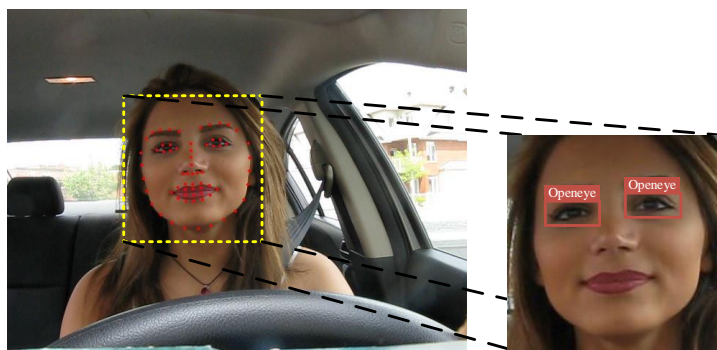


图 4.11 增大人脸检测区域效果图

### 4.2.3 眼睛张角辅助判定眼部状态

本章通过实验对改进 MobileNet-SSD 网络模型的检测率进行了实际测试，发现模型对瞳孔部分遮挡的闭眼识别率较低。为确保最终识别的准确性，本章使用 MCCR 人脸特征点跟踪得到的眼部特征点作进一步的判定。在对常用眼部状态识别方法的不足进行探讨时，发现利用眼睛宽高比 EAR 判断很容易受头部姿态的影响，为此本章引入眼睛张角对识别结果进行二次判定，如图 4.12 所示。

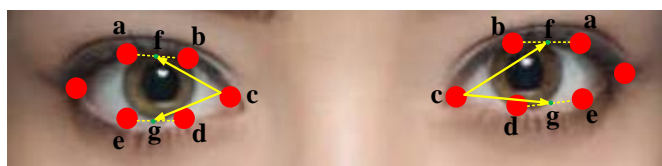


图 4.12 眼睛张角示意图

图 4.12 眼睛中红色特征点由上一章人脸特征点定位与跟踪算法获得，绿色点 f、g 分别为点 a 和 b、e 和 d 的中点，眼睛的张角为 cf 与 cg 的夹角，计算公式如 4.2 和 4.3 所示。

$$d_{ij} = \sqrt{(y_j - y_i)^2 + (x_j - x_i)^2} \quad (4.2)$$

$$A = \frac{(\arccos \frac{d_{cf}^2 + d_{cg}^2 - d_{fg}^2}{2d_{cf}d_{cg}})}{\pi \times 180^\circ} \quad (4.3)$$

其中， $x_i$  和  $y_i$  为特征点的平面坐标， $d_{ij}$  为两点间的距离，而  $A$  为根据特征点计算得到的张角大小。

针对眼睛半开半闭状态时的判定，本章将眼皮遮挡眼睛超过 80% 分为闭眼，这符合 PERCLOS 中的 P80 规范。对于改进的 MobileNet-SSD 网络模型，这种瞳孔遮挡在 80% 的眼睛很容易被识别为睁眼，为提高对这种类型眼睛的判定，本章将网络检测和张角判定相结合。整个判定流程如图 4.13 所示，判断睁眼状态时，将由网络检测和张角两个条件共同决定。

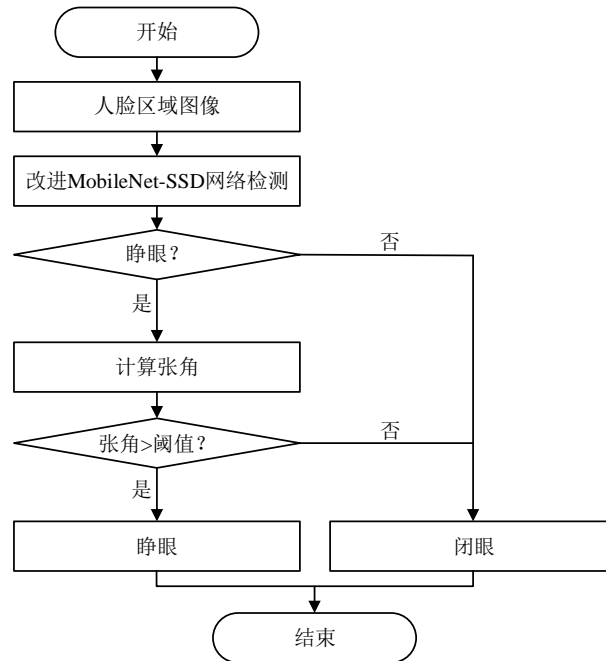


图 4.13 眼睛状态识别流程图

为确保张角判定的准确性，本章对大量睁眼和闭眼状态下的眼睛进行了张角计算，统计结果如图 4.14 所示。通过对数据的分析，通常在闭眼情况下，眼睛的张角都是在  $22^\circ$  范围内，因此本章将闭眼判定的张角阈值设定为  $22^\circ$ 。

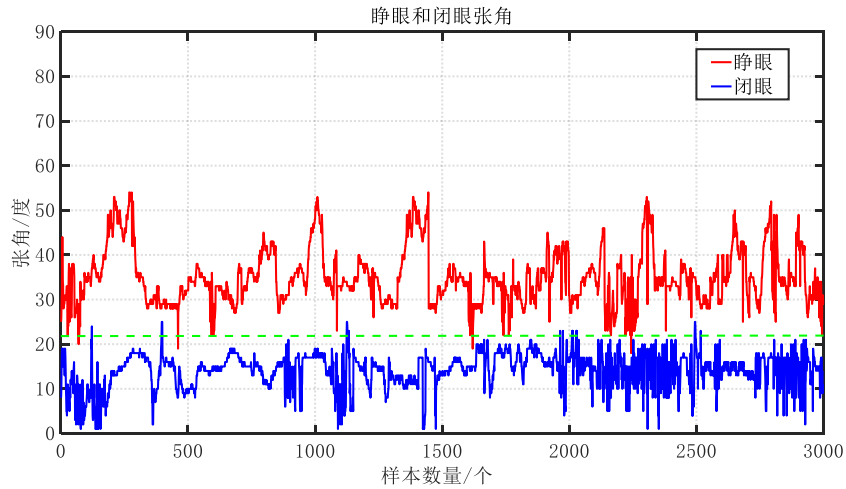


图 4.14 睁闭眼张角统计

针对疲劳驾驶检测，卡内基梅隆研究所经过反复实验和论证，提出了度量疲劳瞌睡的物理量 PERCLOS，其定义为单位时间内眼睛闭合一定比例所占的时间。本章基于 P80 规范计算闭眼的 PERCLOS 参数，公式为：

$$f_e = \frac{n_e}{n} \times 100\% \quad (4.4)$$

其中， $n_e$  为 30 秒内闭眼的累计帧数， $n$  为 30 秒内检测的总帧数， $f_e$  为眼睛 PERCLOS 参数，对  $f_e$  的使用将在疲劳状态等级划分中介绍。

#### 4.2.4 实验结果分析

为训练得到基于改进的 MobileNet-SSD 网络的眼睛状态识别模型，本章使用近红外摄像头采集了白天和夜晚的人脸图像作为数据集。通过人工筛选获得数据集共 13282 张，其中白天睁眼 3625 张，晚上睁眼 3107 张，白天闭眼 3461 张，晚上闭眼 3089 张。最后对数据集进行了手工标注，部分标注示例如图 4.15 所示。

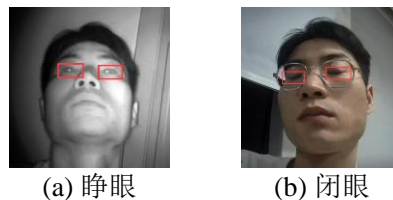


图 4.15 数据集标注示例

实验测试阶段，本章针对睁眼和闭眼各自挑选 1000 张图像作为测试集。为验证改进 MobileNet-SSD 网络的性能，本章与原 MobileNet-SSD 网络分别进行了硬件平台的加载测试，其中网络输入为人脸特征点定位后裁剪出的人脸区域，测试



结果如表 4.2 所示。

表 4.2 网络硬件平台测试

检测方法	检测准确率/%	平均检测时间/ms
MobileNet-SSD	92.3	261
改进 MobileNet-SSD	92.1	249

根据表 4.2 测试数据可知，本章针对眼睛数据集裁剪改进的 MobileNet-SSD 网络与原网络相比，硬件平台端的准确率只相差 0.2%，但是平均检测时间减少了 12ms，这说明本章对网络的裁剪改进是实际有效的。

为了解增加张角辅助判定后眼睛状态识别的实际效果，本章使用和上述同样的测试集，在测试集上与其他眼睛状态识别方法进行对比，测试结果如表 4.3 所示。

表 4.3 眼睛状态识别方法比较

检测方法	睁眼正确数	闭眼正确数	睁眼准确率/%	闭眼准确率/%
Hough 找圆	738	704	73.8	70.4
灰度水平投影	751	717	75.1	71.7
EAR	845	811	84.5	81.1
眼睛张角	864	836	86.4	83.6
改进 MobileNet-SSD	936	901	93.6	90.1
改进 MobileNet-SSD +眼睛张角	934	925	93.4	92.5

根据表 4.3 中的测试结果可知，眼睛张角判断比宽高比判断的准确率高，这得益于头部姿态对张角影响较小。本章只使用改进 MobileNet-SSD 网络的眼睛状态识别方法对睁闭眼的准确率分别为 93.6%、90.1%，与传统方法相比准确率较高。但是该方法对闭眼的准确率较低，而引入张角辅助判断后闭眼准确率得到提高，此时睁闭眼的准确率分别为 93.4%、92.5%，表现较好。

## 4.3 嘴部状态识别

### 4.3.1 纵横比嘴部状态识别

在疲劳驾驶检测中，嘴部状态的识别也是一个判定疲劳的重要条件。正常情况下，驾驶者的嘴部状态是闭合或者微张状态，但是在疲劳状态下会有打哈欠现



象, 这种情况下嘴部张开程度由闭合到较大程度, 并且会在较大程度时保持一定时间。对于打哈欠的识别使用深度学习的方法不易对张开程度进行衡量, 普通说话、唱歌等易误判为打哈欠。通常嘴部目标较大, 受姿态等影响较小。因此本章结合人脸嘴部的特征点, 使用嘴巴纵横比来衡量嘴巴的张开程度, 如图 4.16 所示。

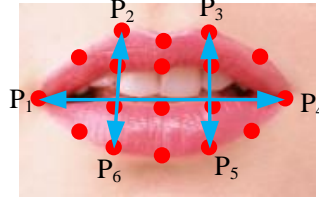


图 4.16 嘴巴纵横比示意图

图中的嘴部特征点是由 MCCR 人脸特征点定位和跟踪算法获得, 利用图 4.16 中 6 个红色特征点的  $x$  和  $y$  坐标可计算得到嘴部纵横比(Mouth Aspect Ratio, MAR), 计算公式为:

$$MAR = \frac{\|p_2 - p_6\| + \|p_3 - p_5\|}{2\|p_1 - p_4\|} \quad (4.5)$$

在得到嘴部张开程度的度量后, 还需要确定嘴部纵横比的一个阈值, 以确定驾驶员嘴部处于打哈欠范围。为保证阈值设置的合理性, 本章采集了不同人在不同视角下的打哈欠数据, 并统计嘴部纵横比, 如图 4.17 所示。根据最终的统计结果, 打哈欠时嘴部的纵横比基本在 0.6 以上, 因此本章将纵横比阈值设置为 0.6, 即当嘴部纵横比大于此值时为打哈欠状态。

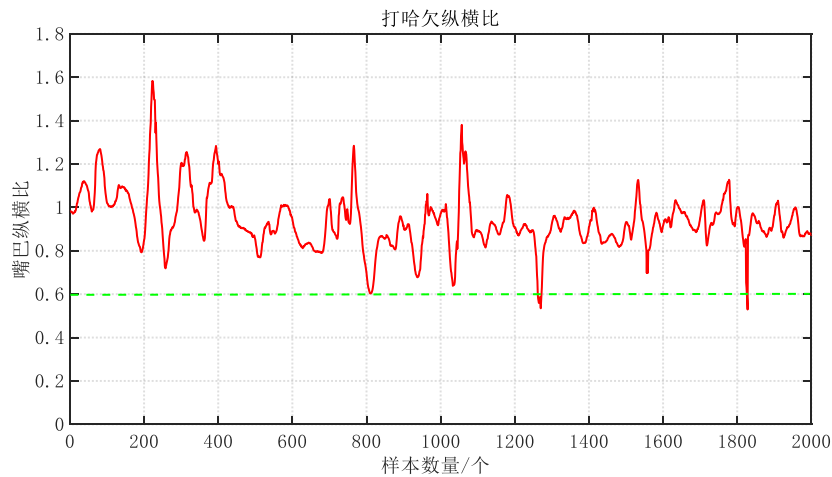


图 4.17 打哈欠 MAR 统计

使用嘴部纵横比进行打哈欠的识别是不准确的, 这只能确定当前检测帧中驾驶员嘴部处于打哈欠范围, 并不能确定驾驶者有打哈欠的动作。当驾驶者在行驶

过程中存在说话、唱歌等行为时，在某一帧中可能检测到嘴部纵横比高于设定的阈值，这时并不一定是打哈欠，因此打哈欠动作的判断不能只依赖某一帧的结果。为了更准确的区分打哈欠与说话、唱歌等动作的区别，本章对多种嘴部动作的视频进行逐帧的纵横比统计，如图 4.18 所示。

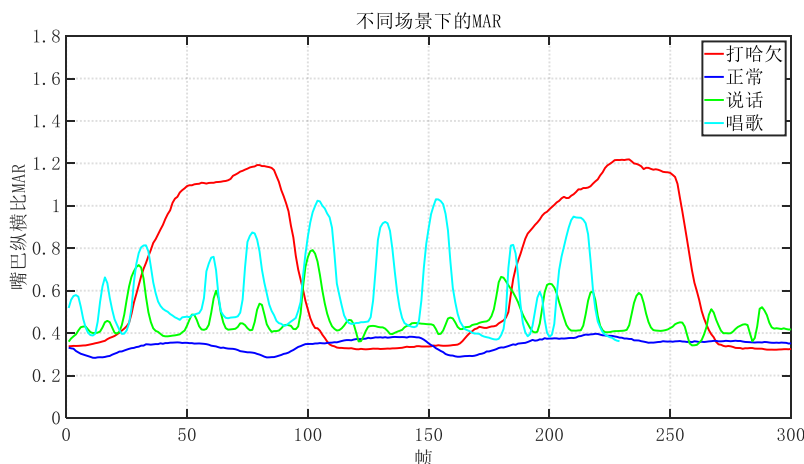


图 4.18 不同场景下的 MAR

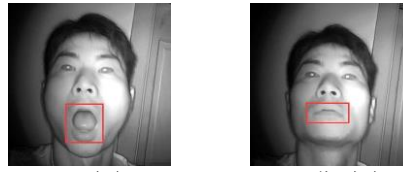
从图 4.18 可知，打哈欠时嘴部连续张开较大程度持续的帧数较多，即持续时间较长，而说话和唱歌时持续的帧数较短。因此本章基于纵横比和持续帧数对打哈欠动作进行判定。为方便计算和衡量，本章将计算一段时间内嘴部纵横比连续超过阈值的帧数与检测总帧数的百分比，如公式 4.6 所示。

$$f_m = \frac{n_m}{n} \times 100\% \quad (4.6)$$

其中， $n_m$  为统计时间段内嘴部纵横比连续高于阈值的最大帧数， $n$  为统计时间段内的总帧数， $f_m$  为高于阈值的帧数所占总帧数的百分比，对  $f_m$  的使用将在疲劳状态等级划分中介绍。

### 4.3.2 实验结果分析

为训练得到基于 MobileNet-SSD 网络的嘴部状态识别模型，本章使用近红外摄像头采集了白天和夜晚的人脸图像作为数据集。通过人工筛选得到不同姿态的人脸图像数据集共 8306 张，其中白天非哈欠 2156 张，白天哈欠 2207 张，夜晚非哈欠 1956 张，夜晚哈欠 1987 张。最后对数据集进行了手工标注，部分标注示例如图 4.19 所示。



(a) 哈欠 (b) 非哈欠

图 4.19 数据集标准示例

实验测试阶段，本章针对哈欠和非哈欠挑选了 1000 张图像作为测试集，在测试集上分别对嘴部纵横比和 MobileNet-SSD 网络判断打哈欠的方法进行测试，测试结果如表 4.4 所示。

表 4.4 嘴部状态识别方法比较

检测方法	哈欠正确数	非哈欠正确数	哈欠准确率/%	非哈欠准确率/%
MobileNet-SSD	905	826	90.5	82.6
MAR	912	874	91.2	87.4

根据表 4.4 中的测试结果可知，使用 MobileNet-SSD 网络的哈欠检测方法与 MAR 方法相比，打哈欠的识别准确率相差不大，但是对非哈欠识别时准确率低，这是因为网络对介于哈欠与非哈欠之间的嘴部状态识别存在误差，会将张开程度较小的嘴部识别为打哈欠，造成误识别。

## 4.4 疲劳状态等级划分

### 4.4.1 多特征融合的划分规则

驾驶者在行驶过程中疲劳是逐渐产生的，疲劳程度由轻到重，因此本章计算 30 秒内眼睛的 PERCLOS 参数  $f_e$  和哈欠参数  $f_m$ ，并使用加权平均方法进行疲劳状态的度量，计算公式如下：

$$F = w_1 f_e + w_2 f_m \quad (4.7)$$

其中， $F$  为加权疲劳参数，眼睛 PERCLOS 参数  $f_e$  的权重为  $w_1$ ，哈欠参数  $f_m$  的权重为  $w_2$ ，并且  $w_1 + w_2 = 1$ 。

为寻求权重参数  $w_1$  和  $w_2$  的最优值，本章采集了大量驾驶者在行驶过程中的视频数据，并在不同的驾驶环境下对不同的疲劳等级进行模拟。得到关于疲劳等级的视频共 400 个，每个视频 60 秒，每秒 30 帧。其中 200 个视频用于分析疲劳等

级的划分, 剩余 200 个视频用于实验测试。通过对数据的分析和实际测试, 得到权重设置对检测准确率影响的关系。如图 4.20 所示, 最终将  $w_1$  设置为 0.4,  $w_2$  设置为 0.6。

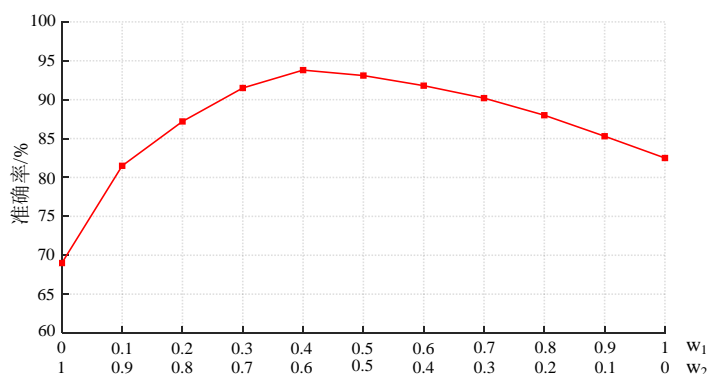


图 4.20  $w_1$  和  $w_2$  寻优

根据疲劳参数值的不同, 本章将疲劳等级分为正常、轻度疲劳、和严重疲劳。结合实际的疲劳数据分析和模拟测试, 设计了划分三种疲劳等级的规则, 如表 4.5 所示。

表 4.5 疲劳状态等级划分规则

疲劳等级	正常	轻度疲劳	严重疲劳
加权疲劳参数	$F < 0.09$	$0.09 < F < 0.18$	$F > 0.18$

#### 4.4.2 实验结果分析

为验证多特征疲劳状态等级划分的性能, 本章使用采集的 200 个实验测试视频进行实验, 视频中正常状态 60 个, 轻微疲劳 90, 严重疲劳 50 个。首先对疲劳状态等级划分的性能进行测试, 实验结果如表 4.6 所示, 本章方法对各疲劳等级的准确率均在 90% 以上, 其中严重疲劳的准确率达到 96%。

表 4.6 疲劳状态检测准确率

疲劳状态	正确检出个数	准确率/%
正常	55	91.7
轻微疲劳	83	92.2
严重疲劳	48	96.0

然后本章将多特征融合与单一特征疲劳检测的性能进行测试, 如表 4.7 所示。可见, 使用多特征融合的疲劳检测比单一特征的疲劳检测的准确率高。因此, 本

文将在后续嵌入式端疲劳驾驶检测系统实现中使用这种方法。

表 4.7 使用不同特征的检测准确率

特征类型	准确率/%
$f_m$	68.7
$f_e$	82.1
加权 $f_e + f_m$	93.6

4.5 本章小结

在人脸特征点定位和跟踪基础上，本章为识别眼睛状态，提出一种改进 MobileNet-SSD 网络的人眼状态识别方法，同时为提升瞳孔部分遮挡时眼睛状态识别的准确性，引入眼睛张角辅助判定。而对嘴部状态的识别，本章利用嘴部 ERA 进行判断，且将嘴部张开较大程度持续帧数占总检测帧数的百分比作为打哈欠的特征参数。最后将闭眼 PERCLOS 参数与打哈欠特征参数加权进行疲劳状态等级的划分，实验表明多特征融合的疲劳驾驶检测准确率高且实用性好。



## 第5章 疲劳驾驶检测系统的设计与实现

本章主要是将本文研究和设计的 MCCR 人脸特征点定位和跟踪方法、基于改进 MobileNet-SSD 和张角辅助判定的眼部状态识别、基于纵横比的嘴部状态识别和疲劳状态等级划分进行集成，结合 NT96580 处理器进行软硬件的设计和实现，并对疲劳驾驶检测系统进行整体测试分析。

### 5.1 系统硬件设计

#### 5.1.1 设计方案

根据疲劳驾驶检测系统实时性和所需实现的功能，可将其分为图像采集、数据存储、深度学习加速、核心处理器、图像显示和语音提醒等 6 个模块。针对算法中大量图像数据和神经网络相关的运算，本系统采用智能多媒体芯片 NT96580 作为核心处理器，芯片整体功能结构如图 5.1 所示。

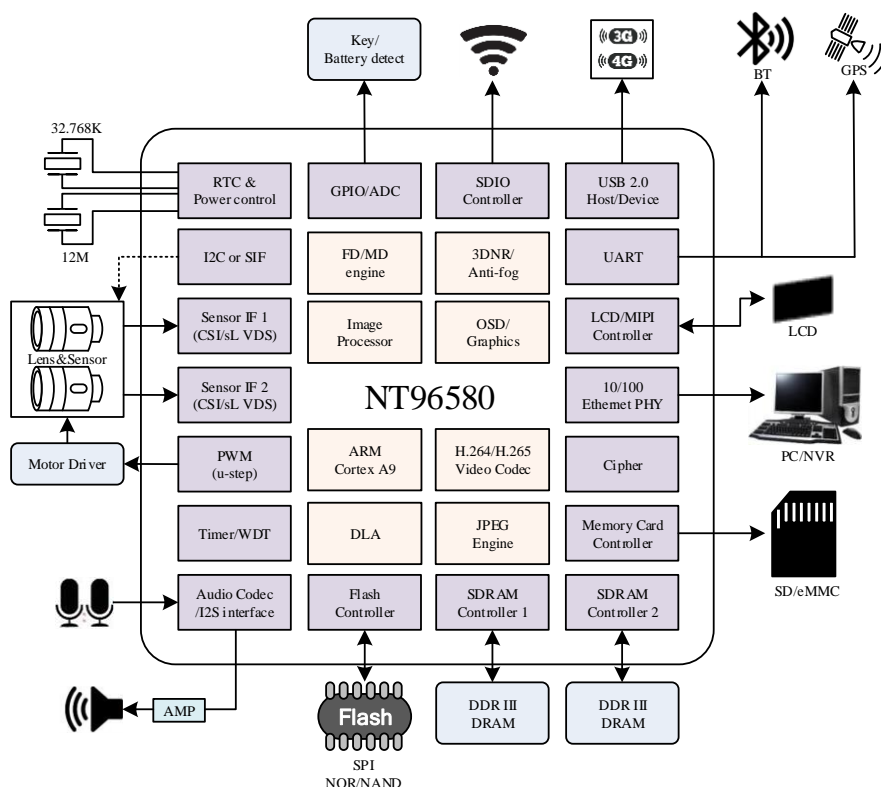


图 5.1 NT96580 整体功能结构

本章系统中, ARM 端主要负责运行 Linux 系统, 对整个软件系统的内存、进程、从设备和算法加速模块等进行管理, 硬件 CNN 引擎用于神经网络模型加载和检测。整个系统运行时, 先是加载存储模块中的 Linux 系统, Linux 系统运行后会加载神经网络模型为后续算法处理做准备。ARM 端会调用近红外摄像头采集图像, 并利用芯片的 ISP 功能进行图像质量的调整, 最终输出的是 YUV420 格式的图像。选取一定的感兴趣区域的图像后, ARM 端会将图像传入硬件 CNN 引擎进行检测和识别。AMR 端相应的应用程序会对检测结果进行分析, 并执行后续的算法, 最终给出显示图像和报警信号。

### 5.1.2 图像采集模块

驾驶员在行驶过程中, 驾驶环境是多变的, 并且获取图像的质量对检测结果会有极大的影响。为使设备能在白天和黑夜都能正常工作, 本章使用高清近红外摄像头采集车内驾驶图像, 如图 5.2 是本文所使用的摄像头, 该摄像头使用索尼的图像处理芯片 IMX291。它使用红外波段为 850nm, 最大支持每秒 120 帧的读出速率和 1920×1080 的分辨率, 并且内置定时调整电路和串行通信电路, 方便数据传输。



图 5.2 高清近红外摄像头

### 5.1.3 NT96580 处理器

疲劳驾驶检测算法运算复杂, 涉及许多浮点运算和矩阵运算, 尤其是神经网络相关的算法特别耗时。为了兼顾实时性和成本, 本系统采用智能多媒体芯片 NT96580 作为核心处理器, 它是一款高度集成的 Soc, 具有高图像质量、低波特率、低功耗, 适用于边缘智能相机应用。芯片集成了 ARM CortexA9 处理器内核, 新一代 ISP、H.265/H.264 视频压缩编解码器、高性能硬件 CNN 引擎、图形引擎、显示



控制器、音频编解码器和 SD/SDIO3.0 等。尤其是其强大的硬件 CNN 引擎，CNN 卷积神经网络加速与计算能力高达 0.75TOPS。其主要的特点归纳如下：

- (1) 高性能 32 位 CPU，单核心 ARM CortexA9，拥有 32kb 指令缓存、32kb 数据缓存和 128kb L2 缓存。
- (2) 传感器接口引擎，支持 MIPI、sub-LVDS、HiSPi 等多种高速串行接口。
- (3) 图像处理引擎，支持 YUV 输出压缩，动态缺陷像素隐藏，以及强大的时空降噪技术。
- (4) 深度学习加速模块，其高达 0.75TOPS 的计算能力，可支持卷积神经网络加速。NUE 网络效用引擎，可加速预处理和分类任务。
- (5) 拥有 H.265/H.264 视频压缩编解码和 JPEG 编解码器。

该处理器的计算能力可支持多路摄像头，是一种专用于相机图像的高性能视觉处理器。官方提供了方便开发的 SDK 包，可供嵌入式软件开发参考。本系统运行的是 Linux 系统，在此之上开发疲劳驾驶算法，主要使用芯片支持的图像处理引擎和深度学习加速引擎。图像处理引擎可使用硬件加速常见的图像滤波算法和边缘检测算法。芯片最强大的功能就是它的深度学习加速引擎，其开发流程如图 5.3 所示。

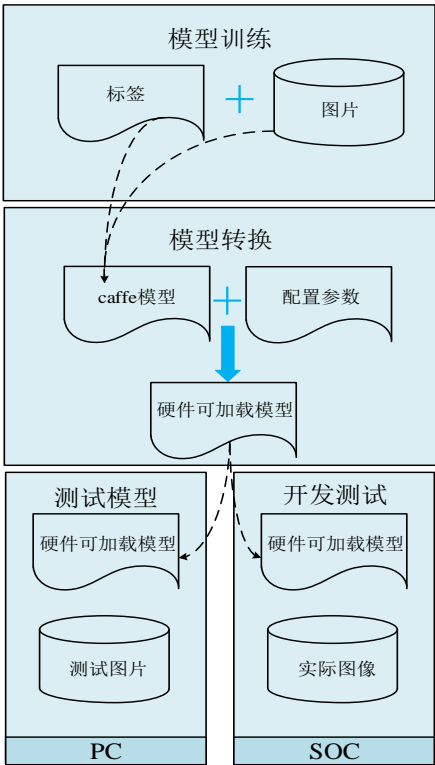


图 5.3 深度学习加速模块开发流程

### 5.1.4 数据存储模块

NT96580 芯片所相关的存储介质有内部存储器和外部存储器。图像和神经网络相关的算法会占用相当大的内存，为满足疲劳驾驶检测算法的要求，本系统使用 512MB 的 DDR RAM 作为内部存储器，这种存储器能在一个时钟周期读写两次数据，数据的存取更快，并且它的成本也具有巨大优势。

在外部存储器方面使用了 Flash，目前主要有 NOR Flash 和 NAND Flash。NOR Flash 的数据读写和常见的 SDRAM 是相同的，CPU 直接与它总线连接，上电即可直接读写。NAND Flash 不能随机读取，它的每次读写是以块为单位进行的，一般每次读写 512 个字节。使用这种技术的 Flash 比 NOR Flash 便宜很多，但是上电使用前需要进行相应的初始化。综合考虑，本章系统只是存储少量的数据，因此使用的是 128MB 的 NAND Flash，其读写接口为 SPI。此外为了在系统调试过程中存储一些图片、视频等原始及中间数据，方便后期的参考和测试。如图 5.4 为本系统的硬件电路板，其中增加了 SD 卡接口，可以使用其保存一些检测数据及图片到 SD 卡中。



图 5.4 系统硬件电路板

### 5.1.5 显示模块

本系统在检测出疲劳驾驶结果后，还需要对结果进行显示和相应的语音报警，因此使用了控制简单的小分辨率 LCD 彩色显示屏和小功率扬声器。如图 5.5 为所使用的显示屏，其功耗低、体积小，使得其安装更加容易。在程序中系统会将采集到的图像缩放为 320×240 的分辨率，并显示出来。其将和扬声器结合对驾驶员进行实时疲劳预警，以达到提醒司机的目的。



图 5.5 LCD 显示屏

5.2 系统软件设计

5.2.1 软件系统架构

疲劳驾驶检测系统主要是采集驾驶者行驶过程中的图像信息，通过图像处理和深度学习等方法，对驾驶者的面部特征进行分析，从而判断驾驶者的疲劳状态，最终以语音方式警告司机。整个软件系统的架构如图 5.6，主要包括驱动层、数据加载层、任务调度层和应用层，以下将对这四个层详细说明。

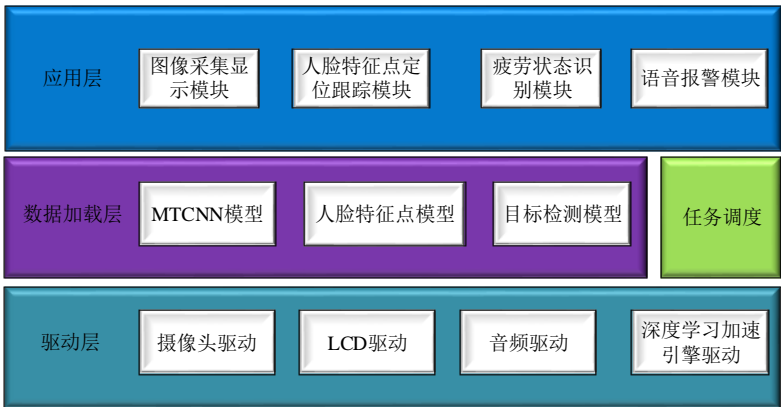


图 5.6 系统软件架构

1. 驱动层

为了方便应用对硬件的控制，驱动是必不可少的。本系统的驱动层主要是深度学习加速引擎驱动、LCD 驱动、音频驱动、摄像头驱动等。其中深度学习加速引擎驱动主要是将神经网络模型在专用硬件端进行加速，达到对输入图像进行快速检测的目的。

2. 数据加载层

本文所研究的疲劳驾驶检测算法主要涉及图像处理和深度学习，因此本层的作用是对相关的神经网络模型进行加载部署，其中包括人脸特征点级联回归模型、

MTCNN 模型和改进 MobileNet-SSD 目标检测模型。

### 3. 任务调度层

为方便各个任务之间的数据传输，本章主要是基于多线程进行各应用的实现，因此本层主要是对应用线程任务进行管理和调度。除此之外还涉及一些线程之间的通信，并且需要预防线程之间出现死锁的情况。

### 4. 应用层

应用层主要实现了图像采集显示模块、人脸特征点定位和跟踪模块、疲劳状态识别模块和语音报警模块，这些都依赖其它层并对下层 API 进行调用。

## 5.2.2 软件系统工作流程

在对软件系统的架构进行设计后，为在疲劳驾驶检测中使得应用层各个模块之间协调工作，本章对整个系统的软件工作流程进行整体设计，如图 5.7 所示。

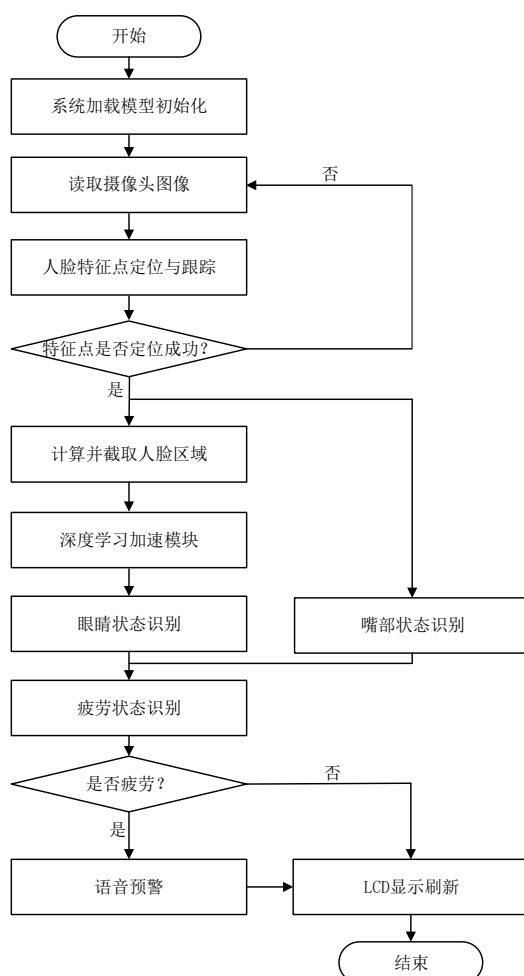


图 5.7 软件工作总流程图

由图 5.7 可知，本章所设计的疲劳驾驶检测软件系统主要包括以下执行流程：

(1) 系统开始运行时，需要对硬件进行初始化，接着加载各种模型用于检测，包括 MTCNN 人脸检测模型、人脸特征点回归模型、基于改进 MobileNet-SSD 网络模型，最后初始化系统参数，如眼睛张角阈值、嘴部纵横比阈值、疲劳状态划分参数等。

(2) 系统初始化后，需要开启摄像头并读取摄像头的数据，采集到的数据还需要经过处理器 ISP 模块进行图像优化，接着就可以交付给应用程序使用。

(3) 获得图像数据后，利用人脸特征点定位和跟踪模块检测特征点，如果检测失败则需接收下一帧图像。

(4) 利用得到人脸特征点检测嘴部状态，并计算人脸区域送入深度学习加速模块得到眼睛状态。

(5) 基于眼睛和嘴部状态计算疲劳参数并进行疲劳状态的划分，接着显示图像数据到 LCD 显示屏，此时如果出现疲劳则进行语音预警。

## 5.3 疲劳驾驶检测系统测试与分析

### 5.3.1 测试环境

为对疲劳驾驶检测系统进行性能分析，本章使用完整的硬件设备在多种环境下进行测试，硬件设备如图 5.8 所示，集成了显示屏和近红外摄像头。



图 5.8 硬件设备

设备的安装位置和角度也是有一定要求的。首先不能影响驾驶员的正常行驶，其次对驾驶员的面部图像采集要清晰，目标大小要合适。综合考虑设备的大小和摄像头焦距等问题，本章设备安装位置如图 5.9 所示，位于车辆正中间。



图 5.9 摄像头安装位置

为便于记录测试过程中图像和参数数据，本章不仅将一些测试效果显示至设备 LCD 显示屏并语音报警，还使用了 SD 卡记录了测试过程中的图像、视频和参数数据，便于后续的结果分析。

### 5.3.2 实验结果与分析

#### 1. 人脸特征点定位实验分析

为验证本文提出的人脸特征点定位和跟踪算法在实际设备中的运行效果，本章在不同光照、不同头部姿态下对移植入设备的算法进行测试，并将结果灰度图保存至 SD 卡，部分结果如图 5.10 所示。同时也验证了算法运行时间，如表 5.1 所示。通过结果表明，算法在各条件下定位效果较好，并可满足实时性。



图 5.10 特征点定位效果

表 5.1 特征点定位算法运行时间

环境	平均运行时间/ms
白天	203
夜晚	211



2. 眼睛和嘴部状态检测实验分析

针对眼睛状态和嘴部状态识别方法在实际设备中运行的效果，本章分别在正常、轻度疲劳和严重疲劳三种状态下对驾驶员测试，并将测试得到的眼睛状态和嘴巴 MAR 数据记录到 SD 卡中，最后绘制了图 5.11。

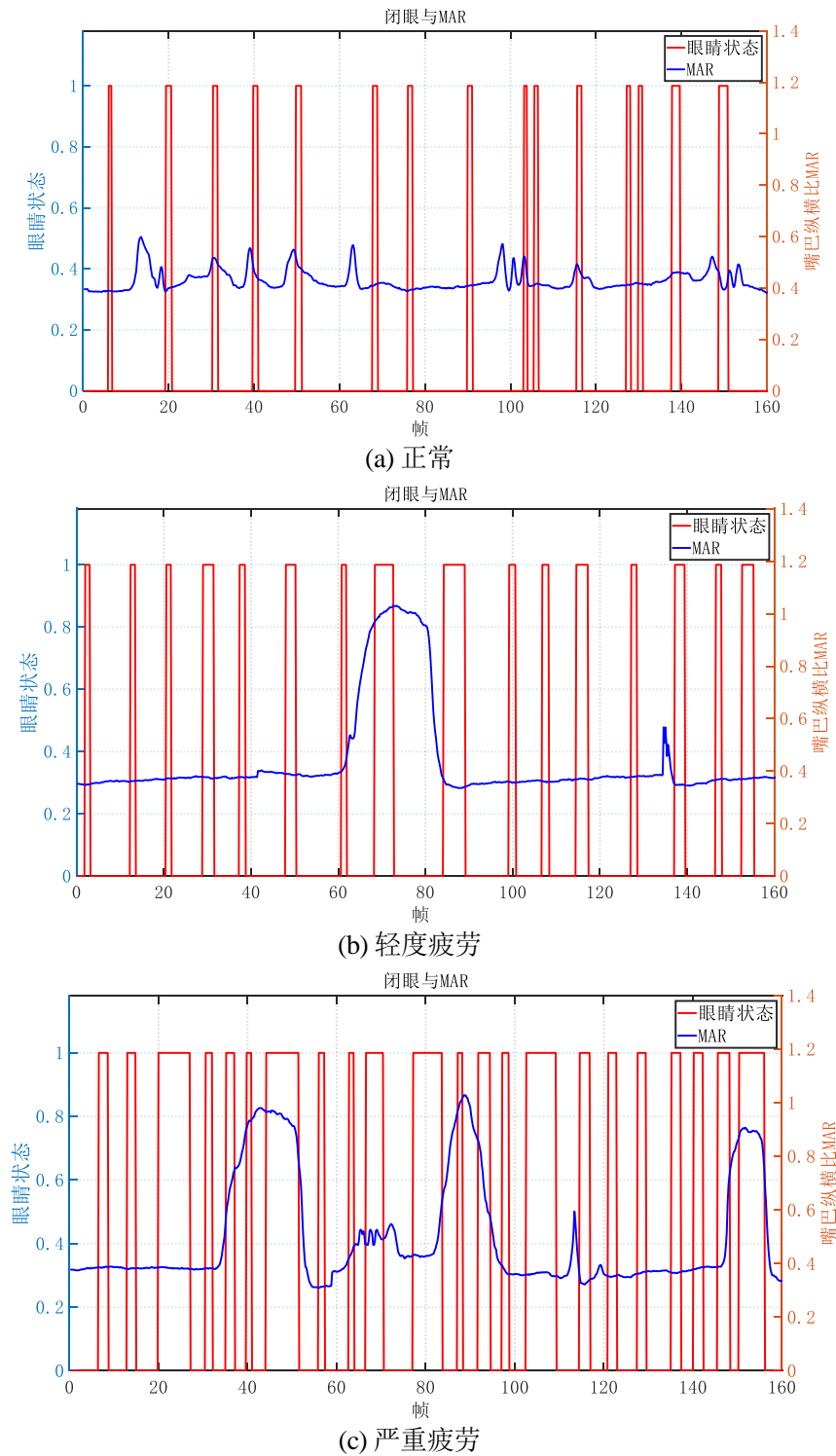


图 5.11 眼睛和嘴部状态识别结果

图 5.11(a)、5.11(b)、5.11(c)中分别是本疲劳驾驶检测系统在正常、轻度疲劳和严重疲劳状态下的测试结果，其中左边纵坐标表示眼睛状态识别情况，0 表示睁眼，1 表示闭眼，右边纵坐标是嘴巴 MAR，横坐标表示检测的帧数。实验结果表明，本系统在三种疲劳状态下能有效检测驾驶员的眼睛和嘴部状态。正常状态时，检测到的闭眼帧数较少、时间间隔较长，并且嘴巴 MAR 均在阈值 0.6 以下。轻度疲劳时，检测到的闭眼帧数较多，并且嘴巴 MAR 出现连续多帧超过阈值 0.6，这表明检测到打哈欠。严重疲劳时，检测到闭眼帧数最多，并有短时间内连续多帧闭眼，并且嘴巴 MAR 有多次出现连续多帧超过阈值 0.6 的情况。综合分析，本系统能及时检测驾驶员的眼睛和嘴部状态，并能通过这两个状态特征明显区分驾驶员的疲劳状态。

### 3. 疲劳驾驶检测系统实验分析

为进一步验证疲劳驾驶检测系统的性能，本章在不同环境下多次模拟了正常、轻度疲劳和严重疲劳，并对系统的检测结果进行记录，如表 5.2 所示。

表 5.2 不同环境下疲劳驾驶检测性能

检测数据	正常		轻度疲劳		严重疲劳	
	白天	夜晚	白天	夜晚	白天	夜晚
测试总数	120	120	150	150	150	150
正确检测数	113	111	140	137	145	144
误检数	7	9	10	13	5	6
准确率/%	94.2	92.5	93.3	91.3	96.7	96.0

由表 5.2 的测试结果可知，本章设计实现的疲劳驾驶检测系统能在全天候进行检测，并且准确率都在 90% 以上。白天环境与夜晚环境相比，三种疲劳状态的划分都比夜晚准确率高，但是准确率相差不大，这表明本文系统能够适应夜晚工况的检测。而在三种疲劳状态中，严重疲劳的检测准确率最高，这是因为严重疲劳特征明显，同时这也满足本系统对检测严重疲劳准确率的要求。

图 5.12 是本系统的部分实验测试结果图，图中是设备显示到 LCD 显示屏的摄像头数据及检测结果。其中显示 E0、E1 分别表示睁眼和闭眼，M0、M1 分别表示未打哈欠和打哈欠、F0、F1、F2 分别表示正常、轻度疲劳和严重疲劳。当检测到疲劳时，设备会语音报警，并将疲劳信息显示到 LCD 显示屏。





(a) 正常



(b) 轻度疲劳



(c) 严重疲劳

图 5.12 设备显示结果

## 5.4 本章小结

本章针对提出的疲劳驾驶检测算法，在嵌入式端对疲劳驾驶检测进行硬件和软件的设计和实现，同时对系统硬件的各个模块和软件工作流程进行了介绍，最后对疲劳驾驶检测系统进行了整体测试分析。



## 第6章 总结与展望

### 6.1 论文工作总结

本文以校企合作项目“车载疲劳驾驶预警系统研究”为基础，完成了疲劳驾驶检测算法的研究及实现。在硬件设计中，为满足实时性、可靠性等要求，本文选用了联咏公司的智能多媒体芯片 NT96580 作为核心处理器，该芯片自带深度学习加速模块，可对本文采用的深度学习算法进行加速。本系统经大量真实场景测试，具有良好的实时性和可靠性。本课题主要完成的工作内容如下：

(1) 对国内外关于疲劳驾驶检测的研究进行了比较分析，并对各类疲劳驾驶检测方法的研究现状进行了梳理，结果表明基于驾驶员行为的疲劳检测方法是当前使用最为广泛的方法。

(2) 基于级联回归的人脸特征点定位，研究和设计了一种多视角约束级联回归的视频人脸特征点定位与跟踪算法。首先，利用 3 维和 2 维稀疏点集建立变换关系，并估计初始形状。其次，由于人脸图像存在较大的姿态差异，使用仿射变换对人脸图像做一个姿态矫正。在构造形状回归模型时，采用多视角约束级联回归模型来减小形状方差，使学习到的回归模型对形状方差具有更强的鲁棒性。最后，采用了重新初始化机制，在特征点正确定位时使用归一化互相关匹配跟踪建立连续帧之间的形状关系。该算法面对人脸大姿态、部分遮挡具有较好的鲁棒性，适合用于视频流中的人脸特征点定位和跟踪。

(3) 针对利用单一特征进行疲劳驾驶检测的局限性，本文将闭眼和打哈欠特征参数相结合，提出利用加权和实现疲劳状态等级的划分。同时针对眼睛状态的识别，本文设计了一种基于改进 MobileNet-SSD 网络的人眼状态判别方法，同时为了提升检测结果的可靠性，引入了眼睛张角辅助判定。而针对嘴部状态的识别，本文利用嘴部纵横比衡量嘴巴的张开程度，并将嘴部张开较大程度持续帧数作为哈欠动作的判定特征。

(4) 针对本文改进的算法，对疲劳驾驶检测系统的软硬件进行设计，并对算法进行优化和移植，最后对系统进行整体的测试和分析。

## 6.2 工作展望

针对本文提出和改进的疲劳驾驶检测算法，本文结合 NT96580 处理器进行了系统软硬件设计，最终实现了疲劳驾驶检测系统，但在真实场景下的测试还不足，不能完全适应复杂多变的真实情况。下面还需研究的工作：

(1) 对真实驾驶环境还需要进行更多的测试，并进一步验证系统的实时性和准确性，针对复杂环境对检测的影响，需研究相应的算法并优化。

(2) 驾驶员在佩戴墨镜时进行疲劳检测，将无法对眼睛状态进行识别，这将影响疲劳检测的准确性。在后续的工作中，可考虑实现墨镜状态下的眼睛状态识别，或者引入其他驾驶特征进行辅助判定，如驾驶员头部姿态、头部运动轨迹等。

(3) 本文在疲劳检测时使用了眼睛和嘴部的特征，但是所使用的人脸特征点定位包含了许多冗余的特征点，实际本文只使用了眼部和嘴部的特征点。因此，后续工作可考虑在预测精度允许情况下减少特征点数量，这有利于减少计算复杂度，提高实时性。

(4) 在 NT96580 平台上的算法设计还可以进行改进，后续可以针对处理器功能进行特定优化，可进一步提升系统的检测性能。

## 参考文献

- [1] 李响. 基于深度学习的面部疲劳信息检测方法研究与实现[D]. 长春: 东北师范大学, 2021.
- [2] 秦恩国. 我国道路交通事故现状与对策研究[J]. 内燃机与配件, 2018(16): 184-185.
- [3] Li Xia, Zhuge Chengxiang, Yu Bingying. Analysis on the impact of illegal driver behaviors on road traffic accidents case study on china[C]//International Conference on Intelligent Human-Machine Systems and Cybernetics (IHMSC). Hangzhou, China: IEEE Press, 2019: 250-253.
- [4] 李都厚, 刘群, 袁伟, 等. 疲劳驾驶与交通事故关系[J]. 交通运输工程学报, 2010, 10(02): 104-109.
- [5] Carroll R. J. Ocular-based measures of driver alertness[C]//International Truck and Bus Safety Symposium. Tennessee, USA: University of Tennessee Transportation Center, 1999: 47-64.
- [6] Göran Kecklund, Torbjörn Åkerstedt. Sleepiness in long distance truck driving: an ambulatory EEG study of night driving[J]. Ergonomics, 1993, 36(9): 1007-1017.
- [7] Fang Ruixue, Zhao Xiaohua, Rong Jian, et al. Study on driving fatigue based on EEG signals[J]. Journal of Highway and Transportation Research and Development, 2009, 26: 124-126.
- [8] Han Chunxiao, Yang Yaru, Sun Xiaozhou, et al. Complexity analysis of EEG signals for fatigue driving based on sample entropy[C]//2018 11th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI). Beijing, China: IEEE Press, 2018: 1-9.
- [9] Harvy J., Sigalas E., Thakor N., et al. Performance improvement of driving fatigue identification based on power spectra and connectivity using feature level and decision level fusions[C]//2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). Honolulu, HI, USA: IEEE Press, 2018: 102-105.
- [10] 王斐, 吴仕超, 刘少林, 等. 基于脑电信号深度迁移学习的驾驶疲劳检测[J]. 电子与信息学报, 2019, 41(09): 2264-2272.

- [11] Lal S., Craig A. Driver fatigue: electroencephalography and psychological assessment[J]. *Psychophysiology*, 2002, 39(3): 313-321.
- [12] 金礼. 基于心电信号的疲劳驾驶研究[D]. 重庆: 重庆大学, 2017.
- [13] Bhardwaj R., Natrajan P., Balasubramanian V. Study to determine the effectiveness of deep learning classifiers for ECG based driver fatigue classification[C]//2018 IEEE 13th International Conference on Industrial and Information Systems (ICIIS). Rupnagar India: IEEE Press, 2018: 98-102.
- [14] Du Guangzhou, Long Shuaiying, Li Chunquan, et al. A product fuzzy convolutional network for detecting driving fatigue[J]. *IEEE Transactions on Cybernetics*, 2022, 2: 2168-2175.
- [15] Ma Yonghao, Tian Fuze, Zhao Qinglin, et al. Design and application of mental fatigue detection system using non-contact ECG and BCG measurement[C]//2018 IEEE International Conference on Bioinformatics and Biomedicine (BIBM). Madrid, Spain: IEEE Press, 2018: 1508-1513.
- [16] Katsis C. D., Goletsis Y., Likas A., et al. A novel method for automated EMG decomposition and MUAP classification[J]. *Artificial Intelligence in Medicine*, 2006, 37(1): 55-64.
- [17] 谢平, 齐孟松, 张园园, 等. 基于多生理信息及迁移学习的驾驶疲劳评估[J]. *仪器仪表学报*, 2018, 39(10): 223-231.
- [18] Lu Jianchao, Zheng Xi, Tang Lihong, et al. Can steering wheel detect your driving fatigue?[J]. *IEEE Transactions on Vehicular Technology*, 2021, 70(6): 5537-5550.
- [19] King D. J., Mumford D. K., Siegmund G. P. An algorithm for detecting heavy-truck driver fatigue from steering wheel motion[C]//16th International Technical Conference on the Enhanced Safety of Vehicles. Windsor, Ontario, Canada: National Highway Traffic Safety Administration, 1998: 873-882.
- [20] Wang Hailin, Liu Hanhui, Song Zhumei. Fatigue driving detection system design based on driving behavior[C]//International Conference on Optoelectronics and Image Processing. Haikou, China: IEEE Press, 2010: 549-552.
- [21] 沙春发, 李瑞, 张明明. 基于方向盘握力的疲劳驾驶检测研究[J]. *科学技术与工程*, 2016, 16(30): 299-304.
- [22] Wylie C. D., Shultz T., Miller J. C., et al. Commercial motor vehicle driver fatigue and alertness study: Technical summary[J]. *commercial drivers*, 1996, 20(6): 431-439.

- [23] Zhao Kun, Meuter M., Nunn C., et al. A novel multi-lane detection and tracking system[C]//2012 IEEE Intelligent Vehicles Symposium. Madrid, Spain: IEEE Press, 2012: 1084-1089.
- [24] Akita S., Kago Y. Zigzag running warning system for automotive vehicles. US, US4673913A[P]. 1987.
- [25] Shimoura H., Tenmoku K. Apparatus for assisting driver in carefully driving. US, US6046671A[P]. 1998.
- [26] Huang Rui, Wang Yan, Guo Lei. P-FDCN based eye state analysis for fatigue detection[C]//2018 IEEE 18th International Conference on Communication Technology (ICCT). Chongqing, China: IEEE Press, 2018: 1174-1178.
- [27] Sharan S. S., Viji R., Pradeep R., et al. Driver fatigue detection based on eye state recognition using convolutional neural network[C]//2019 International Conference on Communication and Electronics Systems (ICCES). Coimbatore, India: IEEE Press, 2019: 2057-2063.
- [28] Liu Shiwang, Yu Long, Hou Mingbin. An efficient method for driver fatigue state detection based on deep learning[C]//2019 2nd International Conference on Safety Produce Informatization (IICSPI). Chongqing, China: IEEE Press, 2019: 172-176.
- [29] Li Fan, Chen Chun, Xu Gangyan, et al. Hierarchical eye-tracking data analytics for human fatigue detection at a traffic control center[J]. IEEE Transactions on Human-Machine Systems, 2020, 50(5): 465-474.
- [30] Liu Guoliang, Yan Dongwen, Chen Ziyu. Research on early warning of driver fatigue status based on image processing[C]//2021 33rd Chinese Control and Decision Conference (CCDC). Kunming, China: IEEE Press, 2021: 2678-2681.
- [31] Liu Xing, Cai Lecai, Wu Zhiming, et al. Evaluation of motor vehicle driver fatigue based on eye movement signals[C]//2021 International Conference on Computer Engineering and Artificial Intelligence (ICCEAI). Shanghai, China: IEEE Press, 2021: 94-98.
- [32] 徐莲, 任小洪, 陈闰雪. 基于眼睛状态识别的疲劳驾驶检测[J]. 科学技术与工程, 2020, 20(20): 8292-8299.
- [33] 沈英超. 基于眼部特征的疲劳驾驶检测系统的研究与实现[D]. 桂林: 桂林电子科技大学, 2019.

- [34] Huang Juan, Lin Zihui. Multi-feature fatigue driving detection based on computer vision[C]//Journal of Physics Conference Series. Guilin, China: IOP Publishing, 2020: 178-188.
- [35] Yu Qichao, Ke Xianxin, Yang Dezhi, et al. Mental fatigue testing based on deep learning[C]//2020 IEEE 9th Joint International Information Technology and Artificial Intelligence Conference (ITAIC). Chongqing, China: IEEE Press, 2020: 32-35.
- [36] Ji Yingyu, Wang Shigang, Zhao Yan, et al. Fatigue state detection based on multi-index fusion and state recognition network[J]. IEEE Access, 2019, 7: 64136-64147.
- [37] Liu Deqi, Zhang Chao, Zhang Qinbei, et al. Design and implementation of multimodal fatigue detection system combining eye and yawn information[C]//2020 IEEE 5th International Conference on Signal and Image Processing (ICSIP). Nanjing, China: IEEE Press, 2020: 65-69.
- [38] 敖邦乾, 杨莎, 令狐金卿, 等. 基于级联神经网络疲劳驾驶检测系统设计[J]. 系统仿真学报, 2022, 34(02): 323-333.
- [39] 邹昕彤, 王世刚, 赵文婷, 等. 基于眼睛与嘴部状态识别的疲劳驾驶检测[J]. 吉林大学学报(信息科学版), 2017, 35(02): 204-211.
- [40] 季映羽. 基于面部特征分析与多指标融合的疲劳状态检测算法研究[D]. 长春: 吉林大学, 2019.
- [41] Raman K. J., Azman A, Arumugam V, et al. Fatigue monitoring based on yawning and head movement[C]//2018 6th International Conference on Information and Communication Technology (ICoICT). Bandung, Indonesia: IEEE Press, 2018: 343-347.
- [42] Ansari S, Naghdy F, Du H, et al. Driver mental fatigue detection based on head posture using new modified reLU-BiLSTM deep neural network[J]. IEEE Transactions on Intelligent Transportation Systems, 2021, 8: 1313-1325.
- [43] Girshick R., Donahue J., Darrell T., et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. Columbus, OH, USA: IEEE Press, 2014: 580-587.
- [44] Girshick R. Fast r-cnn[C]//Proceedings of the IEEE international conference on computer vision. Santiago, Chile: IEEE Press, 2015: 1440-1448.



- 
- [45] Ren Shaoqing, He Kaiming, Girshick R., et al. Faster r-cnn: Towards real-time object detection with region proposal networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(6): 1137-1149.
- [46] Redmon J., Divvala S., Girshick R., et al. You only look once: unified, real-time object detection[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. Las Vegas, NV, USA: IEEE Press, 2016: 779-788.
- [47] Liu Wei, Anguelov D., Erhan D., et al. Ssd: Single shot multibox detector[C]//European conference on computer vision. Amsterdam, Netherlands: Springer Cham, 2016: 21-37.
- [48] Zhang Kaipeng, Zhang Zhanpeng, Li Zhifeng, et al. Joint face detection and alignment using multitask cascaded convolutional networks[J]. IEEE Signal Processing Letters, 2016, 23(10): 1499-1503.
- [49] Xiong Xuehan, Fernando D. Supervised descent method and its applications to face alignment[C]//Proceedings of the 2013 IEEE Conference on Computer Vision and Pattern Recognition. Portland, OR, USA: IEEE Press, 2013: 532-539.
- [50] Cao Xudong, Wei Yichen, Fang Wen, et al. Face alignment by explicit shape regression[J]. International Journal of Computer Vision, 2014, 107(2): 177-190.
- [51] Kazemi V., Sullivan J. One millisecond face alignment with an ensemble of regression trees[C]//Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition. Columbus, OH, USA: IEEE Press, 2014: 1867-1874.
- [52] Ren Shaoqing, Cao Xudong, Wei Yichen, et al. Face alignment at 3000 FPS via regressing local binary features[C]//Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition. Columbus, OH, USA: IEEE Press, 2014: 1685-1692.
- [53] Sun Yi, Wang Xiaogang, Tang Xiaoou. Deep convolutional network cascade for facial point detection[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. Portland, OR, USA: IEEE Press, 2013: 3476-3483.
- [54] Zhang Zhanpeng, Luo Ping, Loy C. C., et al. Facial landmark detection by deep multi-task learning[C]//European conference on computer vision. Zurich, Switzerland: Springer Cham, 2014: 94-108.
- [55] Kowalski M., Naruniec J., Trzcinski T. Deep alignment network: A convolutional neural network for robust face alignment[C]//Proceedings of the IEEE conference on

- computer vision and pattern recognition workshops. Honolulu, HI, USA: IEEE Press, 2017: 88-97.
- [56] Wu Wenyan, Qian Chen, Yang Shuo, et al. Look at boundary: A boundary-aware face alignment algorithm[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. Salt Lake City, UT, USA: IEEE Press, 2018: 2129-2138.
- [57] Kumar A., Marks T. K., Mou W., et al. LUVLi face alignment: estimating landmarks' location, uncertainty, and visibility likelihood[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle, WA, USA: IEEE Press, 2020: 8236-8246.
- [58] Hanowski R. J., Bowman D., Alden A., et al. PERCLOS+: Development of a robust field measure of driver drowsiness[C]//15th World Congress on Intelligent Transport Systems and ITS America's 2008 Annual Meeting. New York, USA: ITS America, 2008:1-13
- [59] Asthana A., Zafeiriou S., Cheng S., et al. Robust discriminative response map fitting with constrained local models[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. Portland, OR, USA: IEEE Press, 2013: 3444-3451.
- [60] Tzimiropoulos G., Pantic M. Gauss-newton deformable part models for face alignment in-the-wild[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2014: 1851-1858.
- [61] Burgos-Artiz X. P., Perona P., Dollár P. Robust face landmark estimation under occlusion[C]//Proceedings of the IEEE international conference on computer vision. Piscataway: IEEE Press, 2013: 1513-1520.

## 致谢

匆匆三年，转瞬即逝，至此二十多年的求学之路即将圆满结束。如今回忆这段时光，可谓“南山南，南山难，上山容易，下山难”，其中既有艰辛也有收获。三年间遇到过很多人很多事，有失望有开心，有迷茫有坚定，我学会了很多也成长了很多。这一路走来最少不了的就是你们的支持和帮助，在此致以衷心的感谢。

首先，感谢我的祖国提供如此安全舒适的生活环境，让我有机会接受教育，学习先进的科学文化知识。感谢学校提供干净整洁的校园，舒适宽敞的宿舍，以及那间井然有序的实验室，让我能够安心学习，专心科研，能在逸夫楼里学习知识技能。祝愿重邮为祖国培养更多优秀人才，蒸蒸日上，创造新辉煌。

其次，感谢我的导师代少升教授。回忆起刚来实验室的时候，代老师面带微笑地向我介绍实验室的情况，这瞬间消除了我的压迫感。其间代老师为人谦虚、做事严谨、尽职尽责的品质深深地影响了我，他在这三年里给了我许多锻炼的机会，让我学到很多专业的知识，丰富了我的专业领域。同时也感谢实验室的师兄师姐们，不管是学习还是项目上都给了我许多的经验和帮助，让我迅速融入这个大家庭。感谢师弟师妹们，主动分担、尽职尽责完成项目。感谢“快乐砖厂”的同门们，在这三年里大家互帮互助，相互学习。感谢我的室友杨锡政、董坤明、岳渤涵，让我的学习生活更加有趣。

然后，感谢给我支持帮助的朋友和重要的人，希望你们永远快乐。最重要的要感谢我的父母，谢谢你们无条件的支持，并在背后默默的为我付出。二十余载对我无微不至的照顾与支持，你们永远是我最强坚实的后盾，也是我昂首前进的不竭动力。养育之恩，无以为报，希望有朝一日能成为你们的骄傲。

最后，真诚感谢百忙之中评阅本论文的老师，以及参与答辩的各位老师，感谢你们提出的宝贵意见。



## 攻读硕士学位期间从事的科研工作及取得的成果

### 参与科研项目：

- [1] “车载疲劳检测与辅助驾驶系统”，校企合作项目，2020 年 10 月至 2021 年 12 月.

### 发表及完成论文：

- [1] 代少升, 肖佳伟, **熊昆**, 等. 基于机器学习的车道线精确检测算法[J]. 半导体光电, 2021.
- [2] 代少升, **熊昆**, 吴云铎, 等. 多视角约束级联回归的视频人脸特征点跟踪[J]. 计算机应用, 2022.
- [3] 代少升, 吴云铎, **熊昆**, 等. 一种基于网络的实时限速牌识别算法[J]. 电讯技术, 2022.

### 获奖：

- [1] **熊昆**. 第十一届“蓝桥杯”，重庆市三等奖, 2020.
- [2] **熊昆**. 第十二届“蓝桥杯”，重庆市一等奖, 2021.
- [3] **熊昆**, 肖佳伟, 吴云铎. “兆易创新杯”第十六届中国研究生电子设计竞赛, 西南赛区三等奖, 2021.