

分类号: TU311

UDC: \_\_\_\_\_

密 级: 公开

单位代码: 10076

## 工学硕士学位论文

# 基于面部多特征融合的驾驶员疲劳 检测方法研究

作 者 姓 名 : 张闯  
指 导 教 师 : 朱天军  
企 业 导 师 : 杨建森  
申 请 学 位 级 别 : 工学硕士  
学 科 专 业 : 机械  
培 养 单 位 : 机械与装备工程学院  
授 予 学 位 单 位 : 河北工程大学

**A Dissertation Submitted to  
Hebei University of Engineering  
For the Academic Degree of Master of Engineering**

**Research on Driver Fatigue Detection Method  
Based on Facial Multi-Feature Fusion**

Candidate : Zhang Chuang  
Supervisor : Prof. Zhu Tianjun  
Pluralistic Supervisor : Dr. Yang jiansen  
Academic Degree Applied for : Master of Mechanical  
Specialty : Mechanical  
College/Department : College of Mechanical and  
Engineering

Hebei University of Engineering

June, 2023

## 摘要

由于全球社会经济的快速发展，车辆总量与日俱增，同时因驾驶员疲劳驾驶导致的车祸大幅度增加，造成大量的人员伤亡。如果能够实时准确检测到驾驶员的疲劳状态并给予警示提醒，就能有效减少疲劳行驶造成的车祸。本文主要围绕基于驾驶员面部多特征融合的检测方法进行研究，主要研究内容包含以下几个方面：

（1）基于软硬件系统搭建驾驶员疲劳驾驶实验平台，同时设计驾驶员疲劳驾驶检测系统，邀请不同年龄段的志愿者进行疲劳驾驶实验，采集志愿者在相同的测试环境下清醒、疲劳和重度疲劳状态的面部特征。

（2）驾驶员疲劳检测过程中，人脸检测是最重要的一步。本文将采用深度级联多任务学习的卷积神经网络框架，获得面部图像和五个关键点，然后再对面部区域进行划分，提取眼睛、嘴巴和鼻子各自的特征。对获得的面部特征进行多任务学习，使每个特征点都能够共享，通过这种方式不仅能够快速准确提取面部的全部特征，而且能够提高人脸特征点检测的鲁棒性。

（3）对采集的驾驶员睁眼、闭眼、嘴巴闭合、打哈欠和是否戴眼镜或太阳镜等特征先进行离线训练，再应用任务约束深度卷积网络对所采集的图像进行面部特征识别，判断眼睛和嘴巴的开合状态，使用 AdaBoost 分类器对眼睛和嘴巴的状态进行分类。在离线训练的基础上，采用一个实时模块：在线测试。在实时视频序列中检测驾驶员的疲劳特征，实时获得眼睛和嘴巴的特征地标，通过这些特征地标可以得出眼睛纵横比（EAR），眼睛闭合百分比（Percentage of Eyelid Closure over the pupil over time, PECLORS）和嘴巴高宽比（MAR）的值。

（4）利用权重比例的方法对疲劳状态下的面部特征进行融合，基于获得的 EAR 值、PERCLOS 值和 MAR 值按不同权值相加得到参数 M，在一定的时间内累加  $M > 0.605$  的图像帧数记为参数 N，通过 N 值的大小来判断驾驶员的疲劳程度。通过在封闭且安全的道路中对不同年龄段的志愿者进行疲劳测试，试验结果表明：本文基于驾驶员面部多特征融合的方法能够有效地利用视频图像实时检测驾驶员疲劳状态，其准确率为 93.1%，对于保护驾驶员及车辆行驶安全具有重要意义。

**关键词：**疲劳驾驶检测；眼睛纵横比；眼睛闭合百分比；嘴巴高宽比；深度学习

## Abstract

Due to the rapid development of the global social economy, the total number of vehicles is increasing day by day, while the number of car accidents caused by driver fatigue has increased significantly, resulting in a large number of casualties. If the driver's fatigue status can be accurately detected in real time and given a warning, it can effectively reduce accidents caused by fatigue driving. This article mainly focuses on the research of detection methods based on driver facial multi feature fusion, and the main research content includes the following aspects:

(1) Establish a driver fatigue driving experiment platform based on software and hardware systems, and design a driver fatigue driving detection system. Invite volunteers of different ages to conduct fatigue driving experiments, and collect facial features of volunteers in awake, fatigue, and severe fatigue states under the same test environment.

(2) Face detection is the most important step in driver fatigue detection. In this thesis, a deep cascaded multi task learning convolutional neural network framework is used to obtain facial images and five key points, and then the facial regions are divided to extract the respective features of the eyes, mouth, and nose. Multitasking learning of acquired facial features enables each feature point to be shared. This approach not only enables rapid and accurate extraction of all facial features, but also improves the robustness of facial feature point detection.

(3) Conduct offline training on the collected features such as driver's eyes open, eyes closed, mouth closed, yawning, and whether to wear glasses or sunglasses. Then, apply a task constrained depth convolution network to recognize facial features of the collected images, judge the opening and closing status of the eyes and mouth, and use an AdaBoost classifier to classify the status of the eyes and mouth. Based on offline training, a real-time module is adopted: online testing. Detect the fatigue characteristics of drivers in real-time video sequences, and obtain real-time eye and mouth feature landmarks. Through these feature landmarks, the values of eye aspect ratio, eye closure percentage, and mouth aspect ratio can be obtained.

(4) The method of weight proportion is used to fuse facial features under fatigue status. Based on the obtained EAR value, PERCLOS value, and MAR value, the

parameter  $M$  is obtained by adding them with different weights. The number of image frames with  $M > 0.605$  is recorded as the parameter  $N$  within a certain period of time, and the degree of fatigue of the driver is judged by the size of the  $N$  value. Through fatigue testing on volunteers of different ages on closed and safe roads, the experimental results show that the method based on driver facial multi feature fusion in this thesis can effectively use video images to detect driver fatigue status in real time, with an accuracy of 93.1%, which is of great significance for protecting driver and vehicle driving safety.

**Keywords:** Fatigue detection; Eye aspect ratio; Percentage of eye closure; Mouth aspect ratio; Deep learning

# 目 录

第 1 章 绪论.....	1
1.1 研究背景和意义.....	1
1.2 疲劳驾驶的因素.....	1
1.3 国内外研究现状分析.....	1
1.3.1 主观评价法.....	2
1.3.2 客观检测法.....	3
1.4 研究现状存在的问题.....	5
1.5 研究的内容和方法.....	5
第 2 章 人脸关键点的检测 .....	7
2.1 多任务卷积神经 MTCNN 网络模型 .....	7
2.1.1 P-Net .....	8
2.1.2 R-Net .....	9
2.1.3 O-Net.....	10
2.2 深度学习的深度卷积网络 DCNN 模型 .....	10
2.2.1 DCNN 网络模型 .....	10
2.2.2 level 1 网络结构 .....	11
2.2.3 level 2 网络结构 .....	12
2.2.4 level 3 网络结构 .....	13
2.3 多任务约束的深度卷积网络 TCDCN 模型.....	13
2.3.1 多任务学习.....	14
2.3.2 TCDCN 网络结构 .....	15
2.3.3 损失函数.....	16
2.4 MTCNN、DCNN 和 TCDCN 网络模型对比分析 .....	17
2.5 本章小结.....	18
第 3 章 基于深度卷积网络模型的疲劳特征识别 .....	19
3.1 神经网络与深度学习.....	19
3.2 离线训练和在线测试模块 .....	23
3.3 任务约束的深度卷积网络 .....	25
3.4 驾驶员面部特征点.....	26
3.5 驾驶员疲劳识别特征.....	27
3.6 本章小结.....	27
第 4 章 面部特征的标定与计算 .....	29
4.1 基于 dlib 68 个地标及面部特征标定 .....	29
4.2 EAR 计算.....	29

4.3 PERCLOS 值计算 .....	32
4.4 MAR 值计算 .....	33
4.5 疲劳状态在线检测技术路线 .....	35
4.5.1 KNN 分类器 .....	37
4.5.2 面部多特征融合 .....	38
4.6 本章小结.....	38
第 5 章 实验数据和结果 .....	39
5.1 环境和数据集.....	39
5.2 驾驶员疲劳驾驶检测系统 .....	40
5.2.1 软件系统.....	41
5.2.2 硬件系统.....	42
5.3 短期实验.....	43
5.4 长期实验.....	45
5.5 不同性别对疲劳检测结果的影响 .....	47
5.6 驾驶员疲劳验证.....	50
5.6.1 疲劳主观问卷调查验证.....	50
5.6.2 疲劳检测系统验证.....	51
5.7 本章小结.....	51
结论.....	53
参考文献.....	55

## 第 1 章 绪论

### 1.1 研究背景和意义

由于全球社会经济和智能科技的快速发展，使得人民的生活水平逐步提升。同时车辆生产成本大幅下降，机动车数量与日俱增，造成的交通事故频繁出现，进而导致大量的人员伤亡。研究发现造成交通事故的因素中，驾驶员疲劳驾驶是极为重要的一部分<sup>[1]</sup>。据国家公安部统计 2021 年国内汽车保有量达到了 3.95 亿辆，同时国家统计局在 2021 年统计年鉴中统计的交通事故共计 244674 起，造成 61703 人死亡，250723 人受伤，直接造成的经济损失累计 131360.6 万元；其中因疲劳驾驶导致的交通事故占总交通事故的 10%~20%左右<sup>[2-3]</sup>，占重大交通事故的比例高达 40%<sup>[4]</sup>。不仅我国存在因疲劳驾驶导致的交通事故，而且世界上其他国家也是如此。例如美国国家交通管理局（National Highway Traffic Safety Administration, NHTSA）依赖于警方和医院的报告，得出驾驶员因疲劳驾驶导致撞车事故的发生率，NHTSA 估计在 2017 年涉及疲劳驾驶的司机多达 9.1 万人，导致 5 万多人受伤，近 800 人死亡<sup>[5]</sup>。

通过上述分析可知，驾驶员疲劳驾驶是造成道路交通事故的重要因素之一。若能够实时检测出驾驶员疲劳驾驶状态并及时提醒，这样能够很大程度上避免交通事故的发生。因此，疲劳驾驶检测系统的研究十分重要，一种能够准确且迅速判断出驾驶员处于疲劳驾驶状态，并能及时提醒驾驶员注意安全驾驶，对人民生命财产安全的保护和降低交通事故的发生率具有重大意义。

### 1.2 疲劳驾驶的因素

导致驾驶员疲劳驾驶的原因包括驾驶环境因素和自身因素。驾驶环境包括：驾驶室内长时间不通风、长时间在高速公路上行驶、驾驶室内温度过高等都会导致驾驶员产生睡意；驾驶员自身因素包括：缺乏充足的睡眠、连续 4 小时以上驾驶车辆、饮酒、吃药等导致驾驶员产生睡意。因此，充足的睡眠或短暂的休息是减轻驾驶员嗜睡和疲劳驾驶的最好方式。

### 1.3 国内外研究现状分析



疲劳驾驶检测方法可以分为主观评价和客观检测。主观评价是通过问卷调查判断驾驶员疲劳的方法<sup>[6]</sup>，例如斯坦福嗜睡量表和派珀疲劳量表；对驾驶员的面部表情特征进行评分，例如打哈欠、眨眼睛、视线转移、头部姿势等。客观检测分为接触式和非接触式两类。接触式测试主要通过对驾驶员生理状态的特征参数进行检测分析，例如脑电、心电、眼电、呼吸等信号，其好处是检测精度高，不足是检测仪器成本太高，要求司机穿戴复杂的检测设备，而且这些复杂的检测设备对驾驶员驾驶车辆造成干扰。相对于接触式检测，非接触式检测更加便宜、方便；例如基于车辆行为特征检测，检测车辆状态和参数，包括车道位置、方向盘转角、行驶速度、加速度、制动等<sup>[7-11]</sup>；基于驾驶员面部特征检测，主要通过眼睛纵横比（EAR），眼睛闭合百分比（PECLORS）、嘴巴高宽比（MAR）的值对驾驶员的状态进行判断。本文采用客观检测的方法，基于驾驶员面部多特征融合对驾驶员疲劳驾驶进行检测。

### 1.3.1 主观评价法

主观评价法主要利用驾驶员面部表情变化、主动应急反应时间、问卷调查等方法对驾驶员驾驶的状态进行判定。主观评价法具有操作简单、花费较低等优点被广泛利用，但是存在误差、实时性较差等劣势。主观评价法包括问卷调查、口头问答、面部表情特征分析方法

#### （1）问卷调查分析

研究者通过对驾驶员填写问卷调查表的形式，对驾驶员的疲劳状态进行分析。Guo 等人为了了解军飞行人员的疲劳状况和特点，采用疲劳评定量表（Fatigue Assessment Instrument, FAI）和疲劳量表（Fatigue Scale-14, FS-14）对 924 位军事飞行人员进行问卷调查<sup>[12]</sup>。李晷等人通过随机抽取 510 名空军飞行人员填写自行设计的飞行人员飞行疲劳状况调查表，疲劳人员出现全身卷懒、腰酸痛、头沉、打哈欠、眨眼等特征<sup>[13]</sup>。万斌等人对现代矿工心理疲劳状况进行了研究，采用问卷调查的方式对山西高平某煤矿一线矿工进行了调查<sup>[14]</sup>。田春雨等人通过对海军某舰队 373 名远航任务舰员填写眼表疾病指数量表（Ocular Surface Disease Index, OSDI）和视疲劳调查表，分析远航任务舰员眼部疲劳发生率和严重程度<sup>[15]</sup>。

#### （2）口头问答分析

口头问答分析主要对测试者进行简单的询问，其受主观意识的影响存在一定的局限性。Lee 等人通过对 929 名飞行员进行口头问答分析，飞机的飞行方向、天气状况、工作类型等都会影响飞行员的疲劳状况<sup>[16]</sup>。Filtness 等人通过小组讨论的方法对铁路驾驶员的疲劳相关的数据进行归纳，得出疲劳驾驶的原因、后果和

应对措施<sup>[17]</sup>。

### 1.3.2 客观检测法

#### (1) 基于驾驶员生理特征的疲劳检测

对于驾驶员生物电信号的早期研究主要依赖于脑电信号。Mervyn 等人建立支持向量机 (Support Vector Machines, SVM), 根据驾驶员的脑电信号 (Electroencephalogram, EEG) 进行建模, 识别了 95% 的睡意驾驶状态<sup>[18]</sup>。Lin 等人基于无线脑电图的实时脑计算机接口系统和基于脑电图的自组织模糊神经网络来预测疲劳驾驶<sup>[19-20]</sup>。Dong 等人通过嵌入人脑的电极获得脑电信号, 并对这些信号进行分析, 以确定驾驶员是否处于疲劳状态<sup>[21]</sup>。Sang 等人利用嵌入在方向盘上的心电传感器来采集驾驶者的心电信号, 并利用心率变异性 (Heart Rate Variability, HRV) 时频域指数来提高信号的准确性<sup>[22]</sup>。Lei 等人提出一种面向短周期心电信号的方法, 提高心电信号的可识别性和频率分辨率的预处理方案, 能够在仅 5 秒间隔的心电信号下有效地识别疲劳特征<sup>[23]</sup>。Zhang 等人提出将不同电极放置在前额, 并提取眼球转动的各种特征用于疲劳检测<sup>[24]</sup>。Bulling 等人发现由眼球运动引起的 EOG 信号是心理活动的良好指标<sup>[25]</sup>。Hu 等人根据眼电图 (Electro-oculogram, EOG) 中提取与眼睑相关的参数作为数据输入, 并利用 SVM 将驾驶员的状态分为清醒、疲劳和重度疲劳<sup>[26]</sup>。同时还结合 EEG、心电图 (Electrocardiogram, ECG)、EOG 来提取相应的特征参数来检测疲劳驾驶。Fu 等人基于驾驶员司机 HRV 的睡意检测算法能力分析与基于 EEG 的睡眠比较<sup>[27]</sup>。Ahn 等人收集 EEG、ECG、EOG 等多个模态和功能近红外光谱仪数据, 对驾驶员的神经生理相关性的状态进行识别<sup>[28]</sup>。Wang 等人结合 EOG 和 EEG 信号分析融合熵, 检测驾驶员的疲劳状态<sup>[29]</sup>。

王磊等人针对 EEG 信号的非平稳性, 在基于通过利用 Relief F 通道选择算法和稀疏表示特征融合算法的基础上, 能够有效地分析疲劳脑电的共性通道<sup>[30]</sup>。杨硕等人通过去趋势的模糊熵值实现对昏昏欲睡的状态进行辨别, 研究表明昏昏欲睡状态的熵值比正常状态的值要小<sup>[31]</sup>。张春翠等人利用非线性复杂度参数, 对疲劳状态下的顶叶、额叶、颞叶和认知运动等有关脑区的脑电信号进行特征描述<sup>[32]</sup>。李增勇等人通过分析司机疲劳行驶过程中 HRV 功率频谱的变化, 结合主观评价分析驾驶员的心脏自主神经功能状态<sup>[33]</sup>。王琳等人以传感器为载体, 收集驾驶员在驾驶车辆的过程中其股二头肌做出相关反应所发射出的生理信号, 再将这种生理信号独立分解成肌电和心电两种信号, 综合这 3 个特征参数区分驾驶员清醒或疲劳两种状态<sup>[34]</sup>。王福旺等人通过对脑电信号进行小波包分析, 获得信号中的 $\alpha$

波,通过统计相对功率谱和眨眼次数,依据脑电和眼电信号的融合特征确定司机是否处于疲劳状态<sup>[35]</sup>。

### (2) 基于车辆行为特征的疲劳检测

当驾驶员疲劳驾驶汽车时,汽车行驶状态会有所改变,进而运用车道偏置、方向盘角度、车速等参数进行疲劳检测。汽车车道偏移量是指在汽车行驶过程中偏移当前轨迹的位移量,一般是运用计算机视觉技术辨认以及检测汽车行驶轨迹。Riera 等人设计一个能够检测和跟踪车道偏移量的方法来识别疲劳驾驶,因此车道偏移量也可用作疲劳检测的参数<sup>[36]</sup>。Takei 等人利用方向盘转向角信息对驾驶员的疲劳状态进行识别,通过使用快速傅立叶变换和小波变换对方向盘转角信号进行预处理,进而获得相应的特征<sup>[37]</sup>。汽车的行驶车速可以通过汽车传感器测量得出,当司机处于疲劳状态时,车速就会发生大幅度变化。D Sandberg 等人采集汽车的速度、汽车的横向位置、方向盘转角和气动角度等数据,通过前馈神经网络对疲劳信息进行分析,辨别司机是否处于疲劳状态<sup>[38]</sup>。

蔡素贤等人利用控制局域网络(Controllor Area Network, CAN)收集汽车行驶情况信息,获得车辆状态相关的参数,通过随机森林算法对获取的相关参数进行处理,再对驾驶员的疲劳状态进行辨别<sup>[39]</sup>。张弛等人通过选择与疲劳状况存在重要关联的数据因素,给出数据预处理和因素标准化的方法,利用层次分析法理论实现正确的数据因素权重划分,进而建立疲劳状况辨别模型<sup>[40]</sup>。屈肖蕾等人研究司机在不同疲劳情况下的转向行为和驾驶状态变量的波动幅度、车速、频度的变化特征,通过统计学方法对疲劳参数显著性变化进行验证<sup>[41]</sup>。黄皓等人通过驾驶模拟器提供仿真驾车环境,并按照需求设定收集的信息,包括方向盘转动角度、汽车横摆角度、汽车横向位移等,通过自主评分技术,确定实验过程中驾驶者的状况,形成疲劳样本数据库<sup>[42]</sup>。

### (3) 基于驾驶员面部特征的疲劳检测

随着机器视觉和人工智能科技的迅速发展,计算机计算能力有了长足的提升。基于驾驶员面部特征融合的疲劳检测方法日益受到人们的关注。Zhao 等人提出一种卷积神经网络(Convolutional Neural Networks, CNN)进行疲劳检测的算法,该算法的第一网络对人眼训练,将人眼和非人眼进行特征划分,第二网络标定人眼特征点位置,按照特征点位置计算眼睛的睁开度,进而得出 PERCLOS 值来确定驾驶员的疲劳<sup>[43]</sup>。Zhang 等人利用面部行为分析的检测方法,结合 Adaboost 分类器和 KFC 相关滤波器(Kernel Correlation Filter, KCF)将红外收集装置收集的人脸图像实施人脸检测和跟踪,通过级联回归定位特征点提取眼睛和嘴巴区域,利用 CNN 识别眼睛和嘴巴的外观表现<sup>[44]</sup>。为进一步提高司机疲劳检测的精确度和时效性,Liu 等人采用深度学习的方法分析驾驶员的面部表情,以识别驾驶员

是否处于疲劳状态<sup>[45]</sup>。其先利用采集的眼睛和嘴巴的大量数据进行模型训练，再使用所建模型对眼睛和嘴巴的状态进行识别，最终通过模糊推理系统推理出司机行驶过程中的疲劳状况。

廖明明等人对采集的驾驶员图像进行预处理，然后采用多任务级联神经网络进行人脸检测并提取眼部和嘴部特征；同时设计眼部状态识别网络和嘴部状态识别网络分别对眼睛和嘴巴区域的状态进行识别<sup>[46]</sup>。苏锦瑾等人经过计算眼睛纵横比与阈值 EAR 比较以统计眨眼次数，通过嘴巴高宽比与阈值 MAR 比较统计打哈欠次数，根据驾驶员眨眼的次数和打哈欠的次数判断驾驶员的疲劳状态<sup>[47]</sup>。刘炜煌等人通过多任务级联卷积网络定位驾驶员的嘴部和眼睛，从驾驶员嘴部和眼睛的图像中获取静态特征信息，结合 CNN 从嘴部、眼睛光流图中获取动态特征进行分类训练，并通过静态图像实现驾驶员的疲劳检测<sup>[48]</sup>。周云鹏等人通过眼睛、嘴部的状况、头部的位置描述驾驶者状况，使用模糊系统推理出驾驶者最终的疲劳状况<sup>[49]</sup>。

## 1.4 研究现状存在的问题

经过检索国内外关于疲劳检测的文献获悉，检测驾驶员疲劳的方式主要包括：基于驾驶员生理特征、车辆行为特征、驾驶员面部特征等。驾驶员生理特征包括 EEG、ECG、EOG、呼吸等信号用作判别疲劳的参数，其优点精度较高，缺点成本较高、要求司机穿戴复杂的检测设备，干扰司机的驾车状况。汽车的行为特征检测主要通过车辆的行驶路径、方向盘转向角、车速等参数确定驾驶员的疲劳状况，该方法不能直接反馈出司机的驾驶状况，受外界影响因素较大。驾驶员面部特征检测主要根据眨眼、打哈欠、点头的频率等来判断驾驶员的驾驶状态，相对于以上两种检测方法，其能够较为直观的检测出驾驶员的疲劳状态，不需要复杂的检测设备，只需要一颗高清摄像头就能够判断驾驶员的驾驶状态。目前，国内外对疲劳驾驶检测研究方法所面临的主要问题：要求驾驶者佩戴复杂的检测装置影响驾驶员驾驶，或仅采用单个特征或参数确定驾驶者是否疲劳驾驶等，而关于上述问题本文将通过基于驾驶员面部多特征融合的方式进行解决。

## 1.5 研究的内容和方法

本文主要针对基于驾驶员面部多特征融合的疲劳检测系统进行研究，应用深度学习理论、可靠性理论等研究方法，进行以下工作：

第1章，绪论。阐述了疲劳驾驶检测方法的研究背景和意义，对国内外驾驶

员疲劳检测方法的研究现状进行归纳，总结出不同检测方法的优缺点。

第2章，人脸关键点检测。阐述了多任务卷积神经网络模型、深度学习的深度卷积网络模型、任务约束的深度卷积网络模型在人脸关键点检测方面的优缺点，对不同的网络模型结构进行详细阐述。通过三种模型的对比，本文采用任务约束的深度卷积网络模型对人脸关键点进行检测，其模型能够解决人脸检测耗时过长的的问题，并且能使各个任务的特征点共享。在能够保证快速且精确的检测到人脸及标定，为提取眼睛和嘴巴区域打下了坚实的基础。

第3章，疲劳检测。阐述了疲劳检测中神经网络和深度学习之间的关联，对神经网络的卷积层、池化层和全连接层的作用进行了说明，详细阐述了具有代表性的卷积神经网络模型。同时对提出的离线训练模块和在线测试模块的作用进行详细阐述，采取多任务训练、多任务学习的方式能够准确快速获得驾驶员面部关键点以及面部特征，并且提供稳定的人脸图像。

第4章，面部特征的标定与计算。采用Dlib的68个地标标记面部轮廓、左眼、右眼、鼻子和嘴巴区域，通过地标点计算眼睛的纵横比和嘴巴的高宽比。同时对眨眼过程、打哈欠过程进行检测，对各自相应的曲线图进行分析。对在线监控工作流程图进行了详细阐述、确定最近邻K值以及描述了面部多特征参数的融合。

第5章，实验数据与结果。对实验环境、数据收集、设备软硬件系统进行了详细阐述，为了验证疲劳检测系统的准确性与实时性，展开了两种实验：一种是短期实验，验证疲劳状态下的眼睛和嘴巴特征对应的核心阈值是否有效；另一种是长期实验，在封闭且安全的道路上对测试者进行测试，实时监测测试者的疲劳状态。通过对实验数据的分析及验证，得出基于面部多特征融合的疲劳检测系统的准确率达到93.1%；最后对测试者进行问卷调查，对测试者的疲劳程度进行评估，对疲劳检测系统进行验证。

## 第2章 人脸关键点的检测

人脸关键点检测是指在一张包含人脸的图片中，检测出人脸的关键点位置，有的方法能够定位人脸的5个关键点，有的方法能够定位出68个关键点，甚至有的方法能够定位出108个关键点。人脸关键点检测方式主要包括以下三类：利用级联回归、基于活动形状模型（Active Shape Model, ASM）和主动外观模型（Active Appearance Model, AAM）以及深度学习。本章将重点阐述利用级联回归的多任务卷积神经网络模型（Multi-task Cascaded Convolutional Networks, MTCNN）、基于深度学习的深度卷积网络模型（Deep Convolutional Network, DCNN）和任务约束的深度卷积网络模型（Tasks-Constrained Deep Convolutional Network, TCDCN）。

### 2.1 多任务卷积神经 MTCNN 网络模型

关于深度学习和面部多特征融合的方法检测驾驶员疲劳的过程中，人脸检测是最重要的一步。由于驾驶员头部姿势的变化以及黑暗的环境因素影响，在实际驾驶过程中很难快速准确检测出人脸的全部特征。因此，疲劳驾驶检测系统既要检测速度快又要检测精度高，为满足这一要求，在检测系统中采用卷积神经网络模型，其可以提高对图像特征识别的精确度，同时可以提高算法的速度。

经典的人脸关键点检测模型是 MTCNN 网络<sup>[50]</sup>，由于使用级联结构，它能迅速进行人脸检测和对齐，也能准确提取人脸的全局特征，进而得到人脸、右眼、左眼和嘴巴的左右两角位置。原始图像通过 MTCNN 模型分析后，可以获得人脸框架和人脸五个关键点，并利用先验知识“三庭五眼”的概念，然后再通过人脸框架和关键点坐标获得眼睛、嘴巴和鼻子的区域。MTCNN 网络的损失函数主要包括：面部5个关键点、人脸区域分类、边框回归（Bounding Box Regression）三部分，以下为各个部分的解释：

（1）人脸五个关键点：其与边框回归相似，都是利用模型生成的回归值和真值（Ground Truth）两者间的偏移量。因此，可以根据偏移量得出网络输出值和真实值的差值。

（2）人脸区域分类：将其定义为一个二分类问题，检测图片中是否含有人脸。

（3）边框回归：针对模型生成的各种候选窗口，都可以获得其偏移量的数

据，这些数据分别是回归的高度和宽度以及回归框的左侧和顶部，可以获得回归与给定窗口两者间的差值。

MTCNN 的网络结构如图 2-1 所示，它是由三个级联子网络构成，分别为建议网络（Proposal Network, P-Net）、精细网络（Refine Network, R-Net）和输出网络（Output Network, O-Net），从粗略到精细依次检测人脸和特征点的位置。

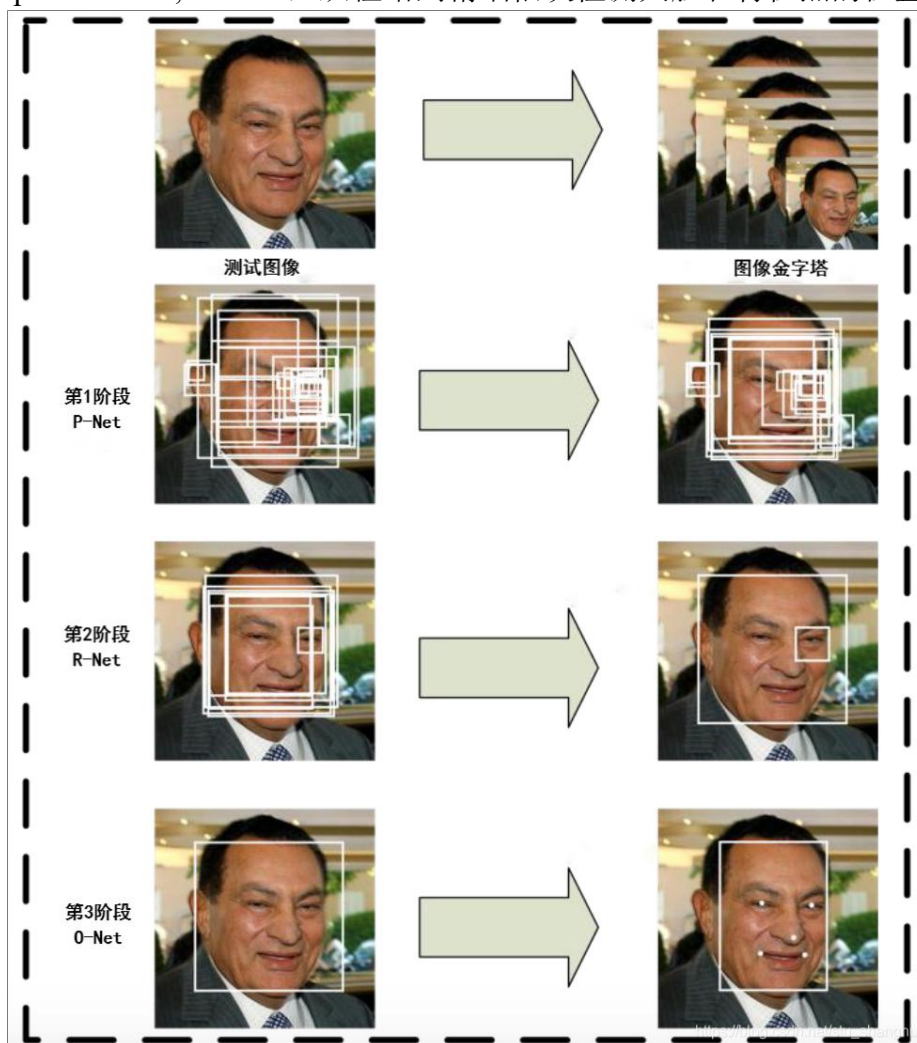


图 2-1 MTCNN 网络结构

Fig.2- 1 Structure specification for MTCNN

### 2.1.1 P-Net

P-Net 基本构造如图 2-2 所示，通过图片金字塔获得图像，其图像尺寸各不相同，使用全卷积网络（Fully Convolutional Networks, FCN）确定各个 12x12 的图像中是否包含一张人脸，提取候选面部范围和回归向量的边框；通过边界框回归（Bounding-Box Regression, BBR）和非极大值抑制（Non Maximum Suppression, NMS）完成这些窗口的筛选。P-Net 是区分人脸区域的建议性网络，然后利用人

脸分类器判定这些区域中是否都含有人脸，在这个部分会获得多个存在人脸的范围，然后再把这些范围输入 R-Net 实现进一步筛选。

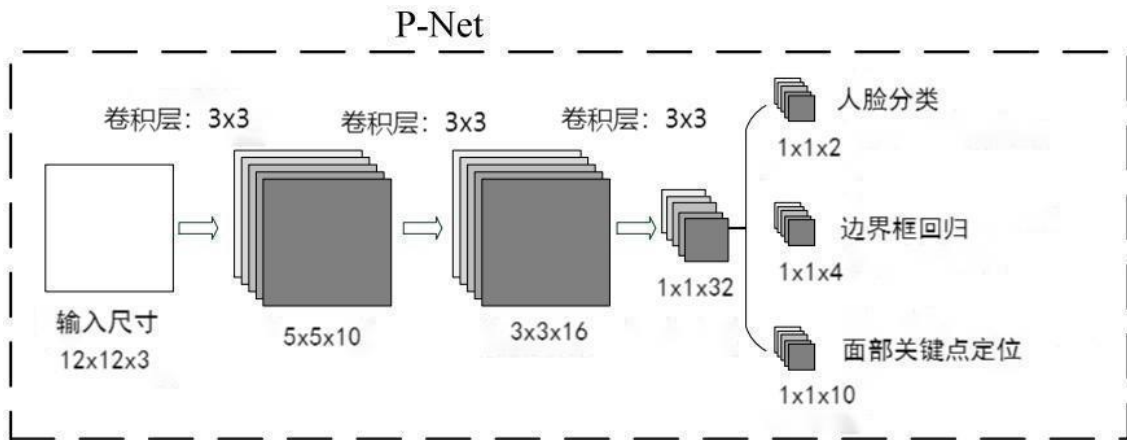


图 2-2 P-Net 网络

Fig.2-2 P-Net

### 2.1.2 R-Net

R-Net 的基础结构是一种卷积式神经网络如图 2-3 所示，相比 P-Net 网络而言，添加一个全新的全连接层，这样对于输入内容的过滤就会比较严格。当图像通过 P-Net 时，将生成多个预测窗口。假设将所有预测的图像都被输入到 R-Net 网络中，网络会将一些比较差的候选框筛选出来。最后，通过 BBR 和 NMS 过滤判断结果。由于 R-net 输出是一个足够精确的面部范围，所以在这种网络技术中，输入是经过精炼和过滤的。因此所有不正确的输入都会被抛弃，然后通过回归和面部关键点定位器完成面部的关键点定位。最后，获得可靠的面部范围，并传输到 O-Net 网络。与使用全卷积输出 1x1x32 特点的 P-Net 网络相比较而言，只是把 128 的全连接层放在 R-Net 网络的卷积层之后，让图像特征得到更多的保存。

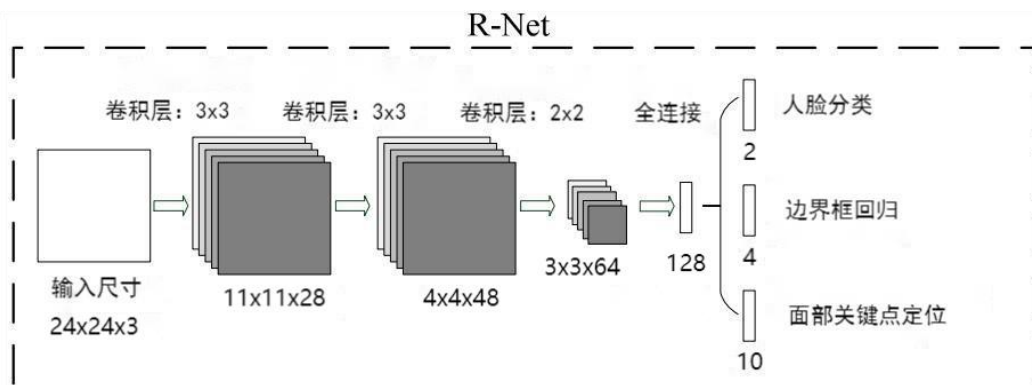


图 2-3 R-Net 网络

Fig.2-3 R-Net



### 2.1.3 O-Net

O-Net 相对于 R-Net 多了一个  $3 \times 3$  的卷积层如图 2-4 所示。O-Net 和 R-net 两者的差别：O-Net 结构通过更多的监督来缩小面部的范围，减少回归的面部特征点，最终对面部的 5 个特征关键点进行输出。在网络结构的末尾，是一个更大的全连接层，保存更多的图像功能，实现人脸识别、校准人脸区域边界回归和人脸特征。

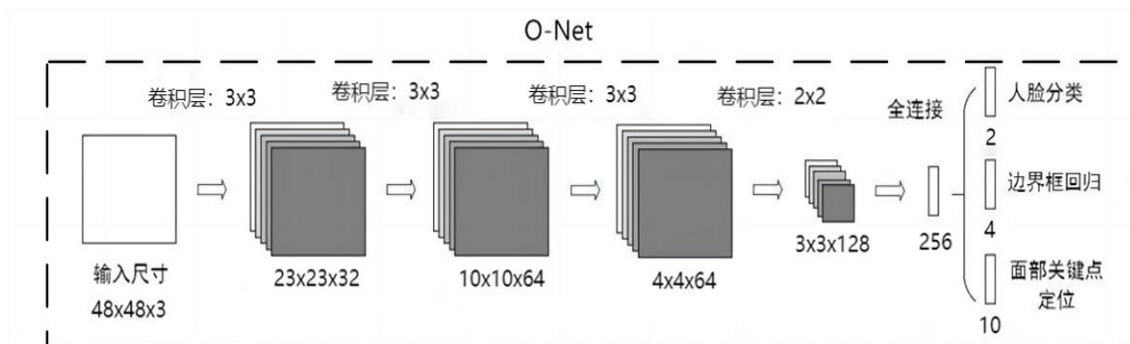


图 2-4 O-Net 网络

Fig.2-4 O-Net

## 2.2 深度学习的深度卷积网络 DCNN 模型

DCNN 网络模型是将 CNN 运用于人脸关键点识别的初步展示方法，由 Sun 等人在 2013 年设计的<sup>[51]</sup>。Sun 等人运用级联网络的设计思路，设置三种不同的卷积神经网络，以便能更好的获取特征点，并且能更加精确的定位特征。如果使用卷积神经网络训练回归特征点位置，仅仅通过一个网络完成回归训练，可能会出现特征点位置不够准确的情况，而通过级联回归 CNN 的方式，可以实现分段式特征点的位置标定，从而能够更加迅速、精确地标定人脸关键点。如果使用比较大的网络系统，对特征点的检测会更加精确，但花费时间会有所延长；为了在效率上和性能上寻找一个平衡点，可以采用较小的网络系统以及级联的思想，先通过粗过滤，之后再微调关键点，而 DCNN 就是采用这种思路。

### 2.2.1 DCNN 网络模型

DCNN 网络模型分成三个部分，依次是：level 1、level 2、level 3。level 1 代表粗糙级别，level 2 代表精细级别，level 3 代表更加精细级别，level 里面包括不同个数的 CNN 模型。DCNN 模型检测分析，首先经过人脸检测器，裁剪出人脸图片，提取哈尔特征（Haar-like Features）；然后从输入 level 1 的图片中裁剪人脸

图像。以下将分阶段地阐述各个 level 的具体细节。

### 2.2.2 level 1 网络结构

输入图像进入人脸检测器后，level 1 网络将面部的边界框裁剪出来，接着将裁剪的图像转换成灰度图像，随后将灰度图像尺寸大小转换成  $39 \times 39$ ，最后将其输入到 level 1 网络中，level 1 的网络结构如图 2-5 所示。第一层的网络结构是由三种神经网络构成，这三种卷积神经网络依次是：F1（输入包含人脸的图像）、EN1（输入的图像包括眼睛和鼻子）、NM1（输入的图像包括鼻子和嘴巴）。

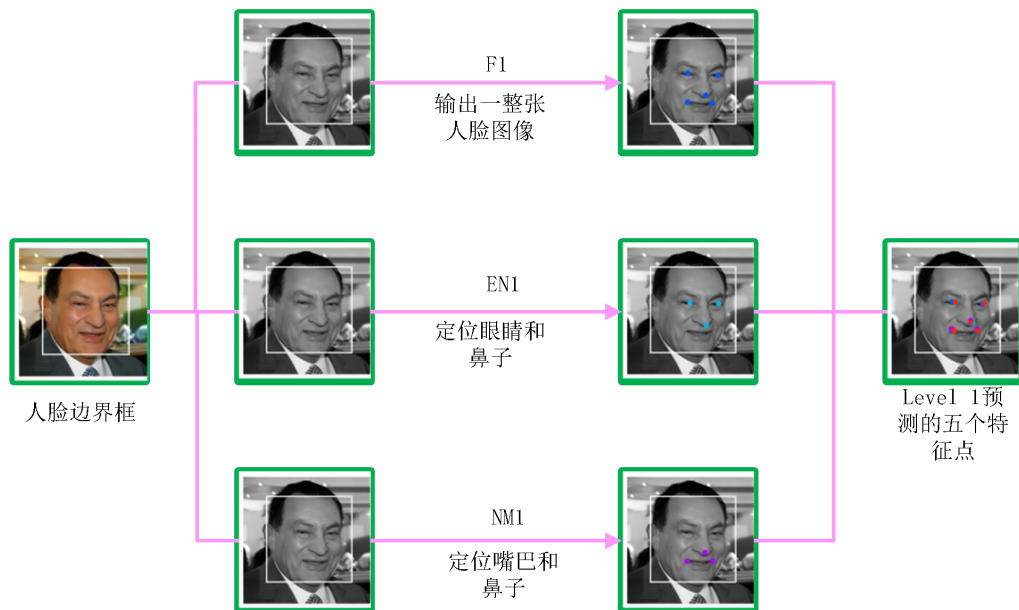


图 2-5 level 1 网络结构

Fig.2-5 Structure specification for level 1

F1 输入是  $39 \times 39$  的人脸图像，输出是检测的五个关键点。F1 的网络结构如图 2-6 所示，将包含人脸的图像输出，再通过卷积神经网络，再将 10 维的特征向量输出。

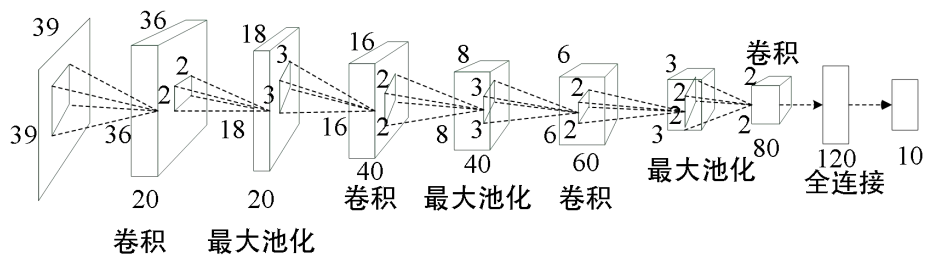


图 2-6 F1 网络结构

Fig.2-6 Structure specification for F1

F1 的基本结构为：第一个卷积特征图选取 20，第二个卷积特征图选取 40，第三个卷积特征图选取 60 和第四个卷积特征图选取 80。EN1 可以用来标定左右眼和鼻子的特征点，所以在模型设计时，输出层的神经元数量是 6。输入的图像是根据比例剪裁的，把  $39 \times 39$  图像的上半部分提取出来，再裁剪出  $39 \times 39$  图像，并确定裁剪出的图像包括眼睛和鼻子的区域。NM1 被用于标定左右嘴角和鼻子的区域。类似的，输出 6 个神经元以及输入的图像，也是根据  $39 \times 39$  的大小裁剪出来，裁剪的部分包括嘴巴和鼻子的区域。

### 2.2.3 level 2 网络结构

通过 level 1 就可以估计出图像上各个关键点的坐标，level 2 的目的是将搜索区域缩小，以 level 1 检测的 5 个关键点为中心，使用一个较小的边界框 (bounding box)，从图片中把 5 个关键点的区域提取出来。level 2 网络结构如图 2-7 所示，其网络包含 10 个 CNN，每个 CNN 有 6 层。这 10 个 CNN，依次用来检测 5 个关键点，每个关键点占用两个 CNN，最后对检测关键点的位置进行平均分配。

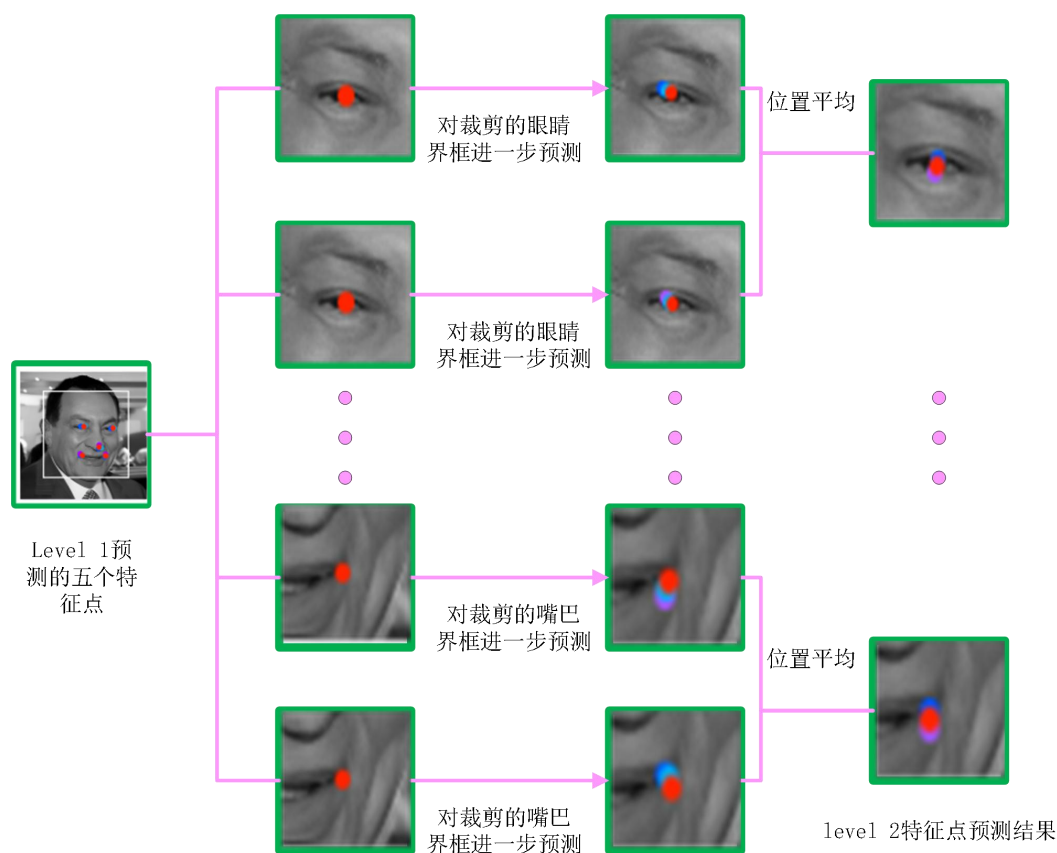


图 2-7 level 2 网络结构

Fig.2- 7 Structure specification for level 2

### 2.2.4 level 3 网络结构

level 3 利用在 level 2 获得的位置基础上，再次进行裁剪。根据 level 2 网络获悉，5 个关键点的区域更加精确化，距离获得关键点的具体位置就更加贴近，level 3 对 level 2 裁剪后的图片再进行一次更加精确的裁剪，对新裁剪的区域进行特征点检测。level 3 的结构如图 2-8 所示，其包括 10 个 CNN，一个特征点检测需要两个 CNN。level 3 与 level 2 的不同是对裁剪范围进行改变，使 level 3 裁剪图片中包含关键点的区域更加具体，关键点识别精度更高。

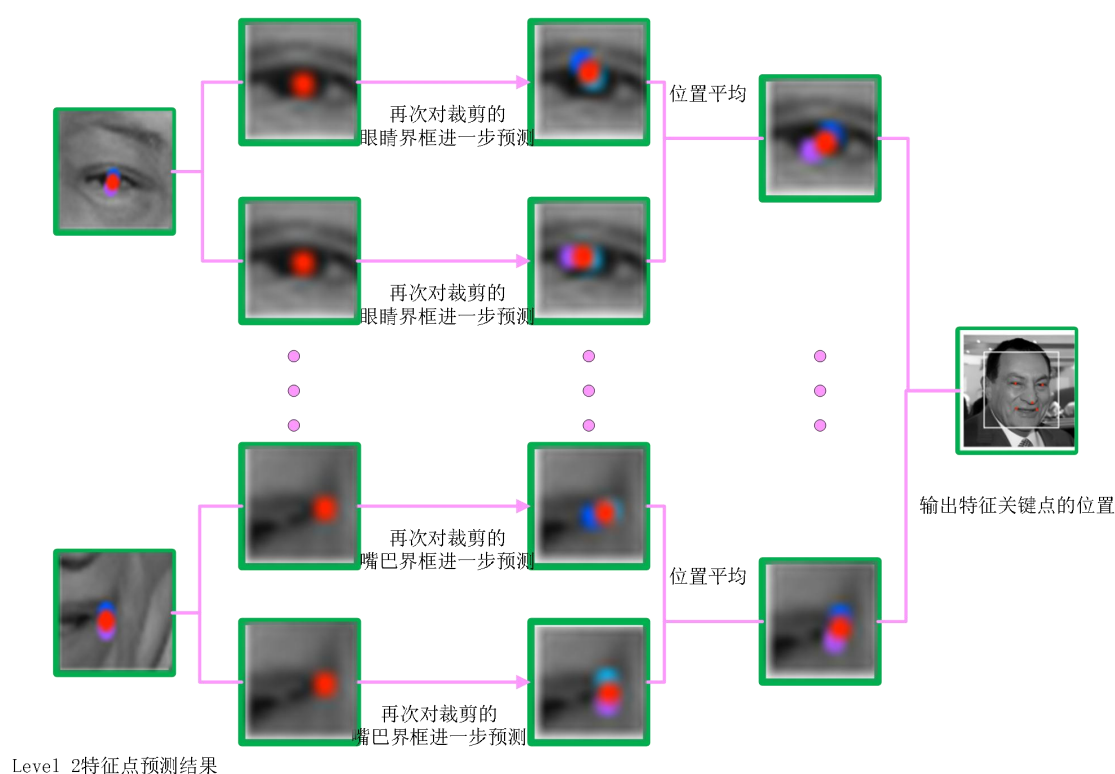


图 2-8 level 3 网络结构

Fig.2- 8 Structure specification for Level 3

## 2.3 多任务约束的深度卷积网络 TCDCN 模型

2014 年 Zhang 等人建立了 TCDCN 网络，该网络最大的贡献就是把多任务学习的思想运用到检测人脸关键点上<sup>[52]</sup>。在对人脸关键点模型的训练过程中发现某些辅助数据（例如性别、面部表情、是否戴眼镜、环境是否黑暗等），这些辅助数据都有可能干扰模型的训练结果，对检测产生一定的影响。通过 TCDCN 训练人脸的关键点位置，正是利用多任务学习的方式，来提高人脸特征点检测的鲁棒性。

### 2.3.1 多任务学习

多任务学习是 TCDCN 通过一些辅助数据实现多任务学习，其好处是检测速度快和多种任务共同学习，而且在人脸关键点检测上使用端到端的技术。并通过辅助特征可以更好的标定人脸关键点，如眼睛和嘴巴的实时位置。依据 Zhang 等人的研究成果表明，对人脸关键点检测技术有一定的帮助。多任务学习（Multi-task Learning, MTL）的问题就是：不同的任务对应着各自的特征，以及不同的收敛方式，而关于这两种情况，提出的处理方法分别是采用任务约束的深度模型（Task-constrained Deep Model）和提前停止任务（Task-wise Early Stopping）的方法。Zhang 等人的模型对处理有覆盖和姿态变化时的效果很好，并且模型也比较简单<sup>[52]</sup>。

为了验证 TCDCN 对人脸关键点检测的准确性及优越性，将 TCDCN、CNN 和级联 CNN（Cascaded Convolutional Neural Networks）进行对比。准备 7 张不同的人脸图像进行测试，比如性别、面部的朝向、是否戴眼镜、是否微笑等特征。CNN、级联 CNN 和 TCDCN 在相同的实验条件下对 7 张图像进行测试，对眼睛、鼻子和嘴巴的检测结果如图 2-9 所示。

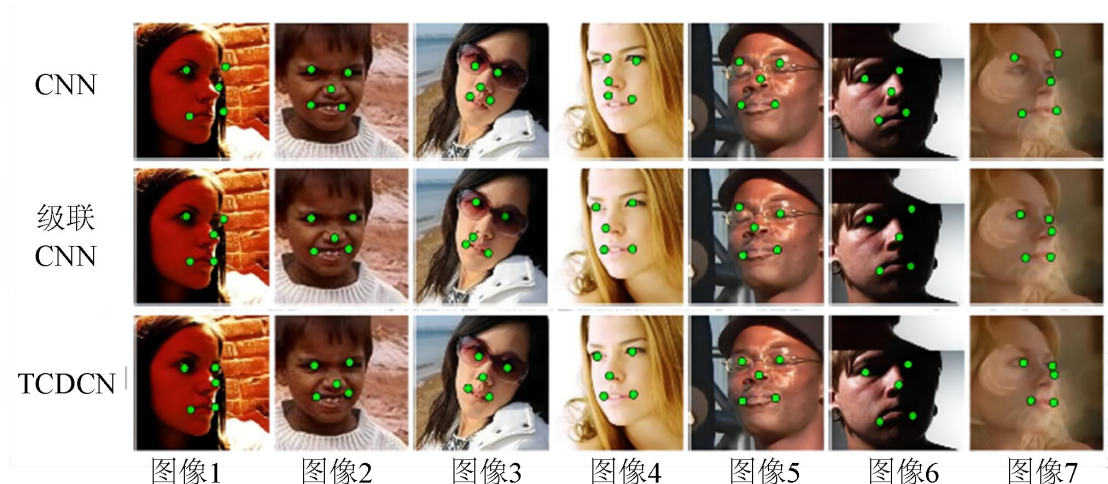


图 2-9 人脸特征点对比

Fig.2-9 Face feature point comparison

CNN 在面部朝向右侧或者左侧时对眼睛和嘴巴的定位点存在偏差；例如图像 1、图像 2、图像 3、图像 4 和图像 6 的嘴角关键点定位存在偏差，以及图像 1、图像 3、图像 5、图像 6 和图像 7 对眼睛关键点定位同样存在偏差，戴眼镜情况下受到一定的影响；而在面部朝向为正面时都不能够准确标定嘴巴、眼睛和鼻子的关键点位置，关键点检测表现力较差。级联 CNN 在面部不同朝向时对嘴巴的关键点检测相对准确，但是在面部朝向不同时对鼻子的关键点定位存在误差，例如

图 2-9 中图像 3 的鼻子定位点在鼻孔下方没有准确定位到鼻子的位置；戴眼睛的情况下级联 CNN 同样受到了一定的干扰导致关键点定位出现偏差，例如图 2-9 中戴眼镜的图像 5 定位点出现在眼睛的上方；相对于 CNN 而言，级联 CNN 对关键点的检测表现较好，综合情况显示级联 CNN 的效果一般。TCDCN 网络无论面部朝向不同、是否戴眼镜都能够对嘴巴、眼睛和鼻子的关键点进行准确定位，正确区分嘴巴、眼睛和鼻子的区域；因此 TCDCN 对嘴巴、眼睛和鼻子的关键点检测效果最佳。

在 TCDCN 能够准确的对嘴巴、眼睛和鼻子关键点定位的情况下，对 TCDCN 进行进一步的测试。图像的面部在不同朝向下对是否戴眼镜、是否微笑和性别进行测试，测试结果如表 2-1 所示，通过这 7 张图像的测试结果来看，TCDCN 对图像是否戴眼镜、是否微笑和性别均能检测正确，具有较强的表现力。

表 2-1 TCDCN 测试结果  
Table 2-1 TCDCN test results

类型	图像 1	图像 2	图像 3	图像 4	图像 5	图像 6	图像 7
戴眼镜	否	否	是	否	是	否	否
微笑	否	是	否	否	否	否	是
性别	女性	男性	女性	女性	男性	男性	女性
姿势	右侧	正面	正面	左侧	正面	正面	右侧

### 2.3.2 TCDCN 网络结构

TCDCN 的网络比较简单，如图 2-5 所示。其网络的输入图像是  $40 \times 40$  的灰度图，图片经过  $5 \times 5$  卷积层和  $2 \times 2$  最大池化层之后得到  $18 \times 18 \times 16$  的图片； $18 \times 18 \times 16$  的图片分别经过两次  $3 \times 3$  卷积层和  $2 \times 2$  最大池化层后依次得到  $8 \times 8 \times 48$  和  $3 \times 3 \times 64$  的图像，再经过一次  $2 \times 2$  卷积层后得到  $2 \times 2 \times 64$  的图像，激活函数选择双曲正切函数 ( $|\tanh|$ )。而  $2 \times 2 \times 64$  的多层经过全连接层后变成 100 个共享特征，因此可以把整个过程看作对图像共享特征的提取，末尾对这些共享特征实现了不同的提取。



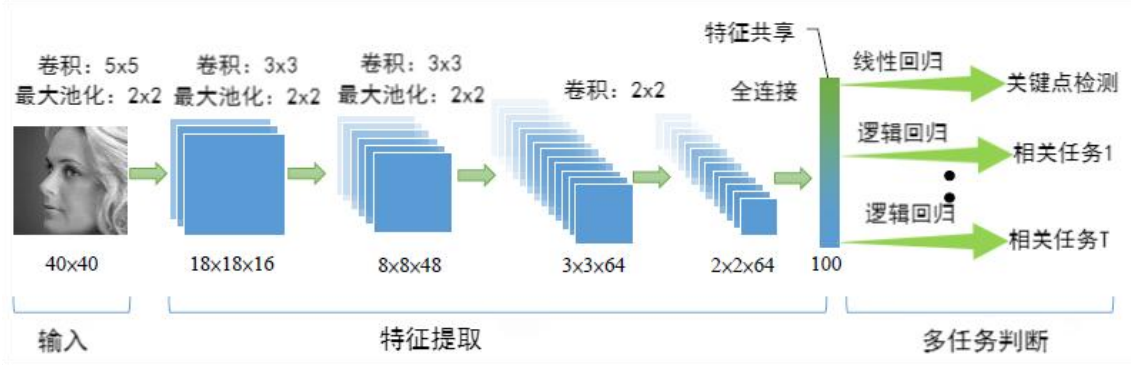


图 2-10 TDCN 网络结构

Fig.2-10 Structure specification for TDCN

该模型将关键点输入所有的子程序之中，从而使各个子程序都能共用到该图像的特征，实现特征共享。通过线性回归实现关键点检测，其余选择逻辑回归的方式，最后将人脸五个关键点的结果输出。

### 2.3.3 损失函数

将多个任务的特征共同享用就是 TDCN 网络的主要优点，线性回归模型用来处理驾驶员面部特征关键点的检测问题，而逻辑回归用来处理面部特征分类问题。在传统的 MTL 中，每个任务都具有同等重要的程度，其目标方程如 2-1 所示：

$$\arg \min_{\{w^t\}_{t=1}^T} \sum_{t=1}^T \sum_{i=1}^N \ell(y_i^t, f(x_i^t; w^t)) + \phi(w^t) \quad (2-1)$$

相对而言，每个任务  $t$  的重要性都是相同的，但是针对更多任务学习时，由于各个任务学习难易有所不同，所以使用同样的损失权重，将使得下一个任务无法收敛。针对多任务学习时，不同的任务处理方式进行了调整，其权值的目标函数如 2-2 所示：

$$\arg \min_{w^r, \{w^a\}_{a \in A}} \sum_{i=1}^N \ell^r(y_i^r, f(x_i; w^r)) + \sum_{i=1}^N \sum_{a \in A} \lambda^a \ell^a(y_i^a, f(x_i; w^a)) \quad (2-2)$$

主任务的损失函数是公式 2-2 的左侧部分，即人脸关键点检测的损失函数，其余各子任务的损失函数是公式 2-2 的右侧部分。针对人脸关键点检测任务，结合三个子任务，分别是：性别、脸部位置、戴眼镜，最终目标函数如 2-3 所示：

$$\sum_{i=1}^N \|y_i^r, f(x_i; w^r)\|^2 + \sum_{i=1}^N \sum_{a \in A} \lambda^a y_i^a \log(p(y_i^a | x_i; w^a)) + \sum_{t=1}^T \|w\|_2^2 \quad (2-3)$$

平方和误差是公式2-3 中最左侧的部分，是人脸关键点的损失函数，分类任务是公式2-3 中间的部分，使用的是交叉熵误差，正则项是该公式 2-3 中最右侧的部分。

多任务学习中各个任务收敛的速度不同，在模型训练过程中是一个极为重要的问题，使用提前停止任务的方式。如果在某个子任务中取得良好的成绩后，对于主任务来说这个子任务就没有任何意义，那么就可以终止这个任务。关于自动终止子任务的计算公式如 2-4 所示：

$$\frac{k \cdot \text{med}_{j=t-k}^t E_{tr}^a(j)}{\sum_{i=t-k}^t E_{tr}^a(j) - k \cdot \text{med}_{j=t-k}^t E_{tr}^a(j)} \cdot \frac{E_{val}^a(t) - \min_{j=1..t} E_{tr}^a(j)}{\lambda^a \cdot \min_{j=1..t} E_{tr}^a(j)} > \theta \quad (2-4)$$

公式 2-4 中的  $\theta$  表示阈值，公式乘积左侧为训练误差的趋势，右侧为泛化误差与训练误差之比，当阈值  $\theta$  大于两项乘积时，则该任务终止。

TCDCN 通过多任务学习来识别人脸关键点。在人脸关键点识别任务中，多任务学习的主要问题是：根据不同任务间的难度和收敛速度的差异，提出一种新的目标函数和提前停止方法，并进行相应的改进。

## 2.4 MTCNN、DCNN 和 TCDCN 网络模型对比分析

通过对 MTCNN、DCNN 和 TCDCN 进行对比结果如表 2-2 所示。在表 2-2 中，可以发现 MTCNN 的主要特征是：能够同时完成人脸检测和人脸关键点标定，且能够挑选难以学习的样本（Hard Sample）；缺点为：耗时太高且定位一般。DCNN 的网络的优点为：首次将深度学习引入关键点检测领域；缺点为：卷积神经网络的数量太多，耗时太长，关键点定位一般。TCDCN 的好处在于：完成了多任务的训练学习后，可以迅速标定关键点。因此，检测系统通常选用 TCDCN 网络为脸关键点的检测与标定。

表 2-2 各网络人脸特征点检测以及定位的优缺点

Table 2-2 The advantages and disadvantages of facial feature point detection and localization in each network

方法	优点	比例缺点
MTCNN	可以同时进行人脸检测和人脸关键点定位两个任务，且能够进行挑选难以学习的样本。	耗时太高且定位一般
DCNN	首次将深度学习引入关键点检测领域	卷积神经网络的数量太多，耗时太长，关键点定位一般
TCDCN	实现了多任务的训练学习，且能够快速定位关键点。	在本文场景中无问题



## 2.5 本章小结

本章节主要阐述了人脸关键点检测技术在神经网络和深度学习领域中的应用，并阐述了关键点检测技术在计算机等不同应用领域强大的表现能力。并阐述 MTCNN、DCNN 和 TCDCN 在人脸关键点检测领域的应用，对各自网络结构以及优缺点进行阐述；接着对 MTCNN、DCNN 和 TCDCN 在人脸关键点测试中进行比较，经过多次实验验证之后，最终确定 TCDCN 多任务学习作为检测驾驶员人脸关键点的网络模型。

## 第3章 基于深度卷积网络模型的疲劳特征识别

对驾驶员疲劳检测系统中应用的神经网络和深度学习进行了详细的阐述，对比各个网络模型的优缺点，本文采用任务约束的深度卷积网络模型对疲劳特征进行识别。提出离线训练和在线测试模块，增加疲劳检测的实时性以及准确性；通过多任务训练可以准确获得驾驶员的面部特征，提供稳定的人脸图像。

### 3.1 神经网络与深度学习

机器模型中存在一种学习类型被称作神经网络，其模仿动物神经网络的动作特性，运用数学模型以实现分布式网络并行信息处理的目的。典型的神经网络包括输入、输出和中间（也叫隐藏）三层，其结构如图 3-1 所示，从左至右依次是输入层、隐藏层和输出层。

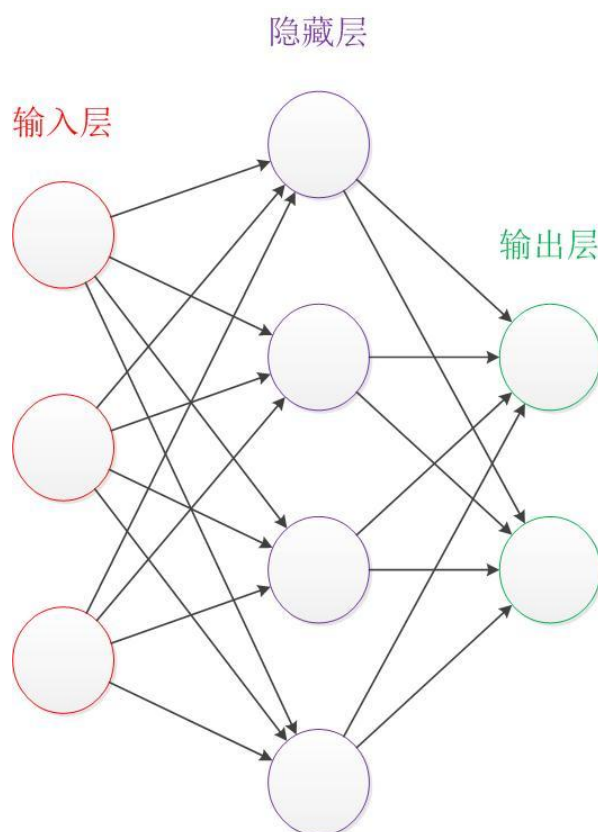


图 3-1 神经网络结构图

Fig.3-1 Neural network structure diagram

一般来说，神经网络的架构可以分为四类：前馈神经网络、循环网络、对称连接网络、卷积神经网络。

卷积神经网络包括卷积层、池化层和全连接层三个网络层结构，以下将分别说明各个网络层结构。

### (1) 卷积层

卷积作为最有效地提取图像信息特征的方式，通常采用逐点式，其卷积核为正方形，对图像上所有像素点进行遍历。各个像素值表示为图像与卷积核相重合的位置，然后对卷积核中对应的权值进行相乘，最后对每个权值乘积相加同时再加上偏置就是图像输出的每一个像素值。其中灰度图和彩色图在包含卷积核的数量方面可能是单个或者多个，比如单通道（这里单通道是指灰度图）的卷积核数为1个，如图3-2所示。图3-2输入的灰度图是 $5 \times 5 \times 1$ 图片，其中1表示单通道， $5 \times 5$ 表示分辨率，而灰度值是5行5列，卷积核是 $3 \times 3 \times 1$ 的。用 $3 \times 3 \times 1$ 的卷积核对 $5 \times 5 \times 1$ 的灰度图片进行卷积计算，偏置项 $b=1$ ，则求卷积结果是： $(-1) \times 1 + 0 \times 0 + 1 \times 2 + (-1) \times 5 + 0 \times 4 + 1 \times 2 + (-1) \times 3 + 0 \times 4 + 1 \times 5 + 1 = 1$ 。

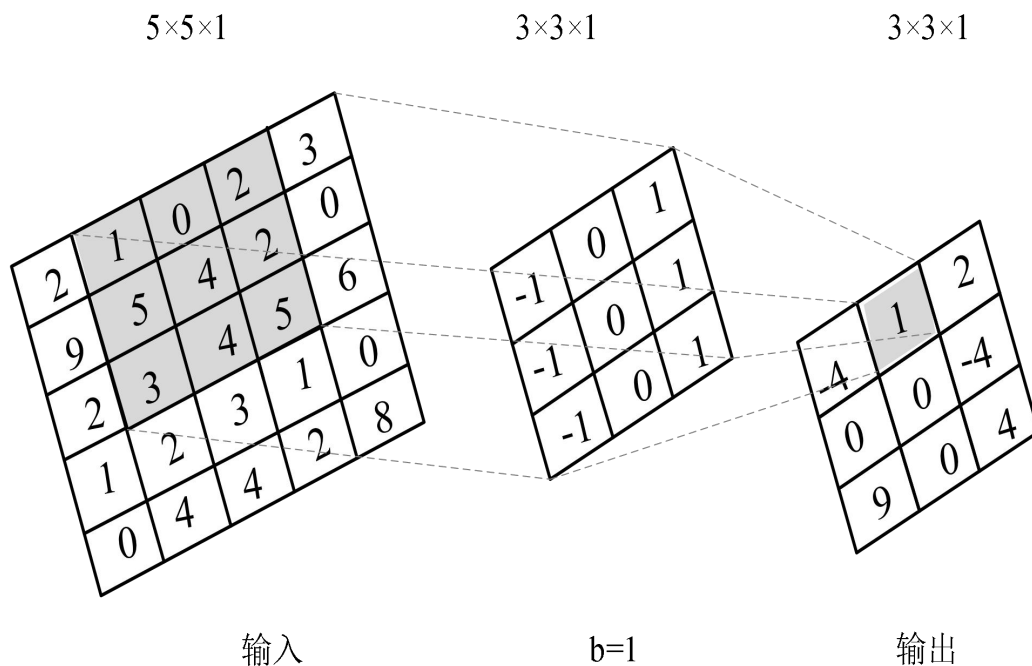


图 3-2 单通道卷积层

Fig.3-2 Single channel convolution

### (2) 池化层

池化层是某种类型的降采样，而非线性池化函数有不同的种类，但其“最大池化”是最普遍的。输入的图像被池化层裁剪出许多个矩形区域，并且输出每个区域的最大值。池化层会进一步降低信息的空间尺度，随之图像参数的数量和运

算量也会相应的降低很多，进而防止出现过拟合。使用池化层的效果与其输入相比大大减少了运算量，因为池化层的输入主要是模拟人的视觉系统。对于卷积神经网络普遍认为池化层三种作用：一是特征不变性。采用池化方式让模型比较注重一些特征而不是特征精确的坐标，其中不变性包含平移不变性、旋转不变性和尺度不变性。例如图 3-3 所示，输入  $(1,0,4,6)$ 、 $(2,3,6,8)$ 、 $(3,1,1,2)$ 、 $(1,0,2,4)$  经过最大池化后会获取 6、8、3、4，不改变其伸缩的方向，原来的神经元在通过最大池化后输出的数值会变大，但在通过伸缩（尺度变换）后，最大池化输出最大值的概率可能在该神经元上。二是下采样（特征降维），相当于池化在空间区域内进行维度约减，模型就可以得到更加宽泛的区域特征，但会使输入下一层的维度减小，进而减少图像的运算量和参数数量。一定程度上避免过拟合现象的发生。

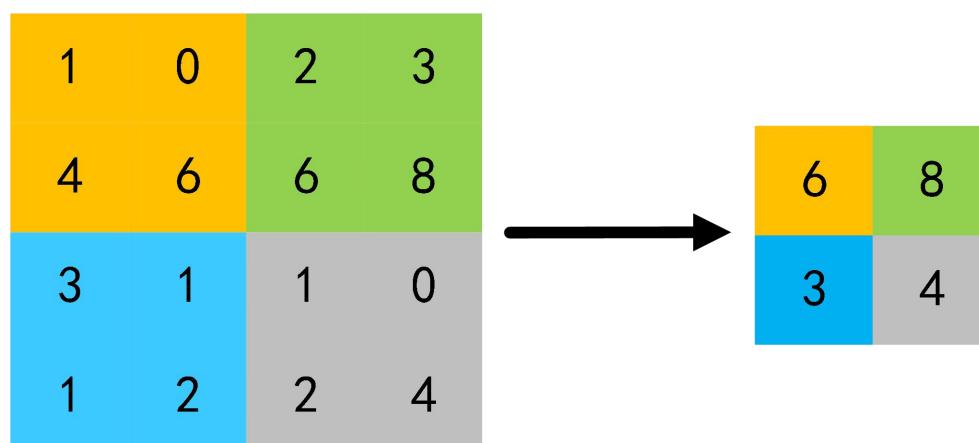


图 3-3 池化层

Fig.3-3 pooling layer

### （3）全连接层

全连接层在卷积神经网络中具有分类器的功能，假如卷积层、池化层和激活函数层等只是将原始数据投射到隐藏层特征空间，而全连接层则是将学到的分布式特征表示投射到样本标记空间。其主要功能就是将前层（卷积、池化层等层）运算得出的特征空间投射样本标记空间。简洁的讲便是把所有特征合并为一个值，其优点在于降低特征位置对分类结果的干扰，进而提高整个网络的鲁棒性。在整个卷积神经网络的最后通常会设置全连接层，负责把卷积输出的二维特征图转化成一维的某个向量，进而完成端到端的学习流程，即输入一张图片或者一段语音，输出一个向量或信息。

为奠定本文的研究基础，将在本小节阐述现有研究者应用较多且具备一定代表性的网络模型。

### (1) LeNet-5

LeNet-5, 此名称取自于研究者 LeCun 的姓名, 5 是作者第五次改进后的模型结构, 在 LeNet-5 之前还有 LeNet-4 和 LeNet-1 鲜为人知。LeNet-5 神经网络主要应用于识别手写数字和机器印刷字符。LeNet-5 描述了图像中像素特征之间的关联可以通过参数共享的卷积方法所获得, 并且采用卷积、下取样(池化)以及非线性映射等的组成架构, 是当前大多数深度图像识别网络的基石。LeNet-5 虽然是早期提出的一种小网络系统, 却涵盖多数深度学习卷积神经网络的基础: 卷积运算层、池化层和全连接层。而 LeNet-5 中总共包括 7 层(输入层不作为网络结构), 它们由 2 个卷积运算层、2 个池化层和 3 个全连接层组合成, 其中层和全部连接层的核尺寸分别代表采样范围和连接矩阵尺寸。LeNet-5 模型的特点: 实现卷积、池化、非线性映射三层序列的组合, 为目前深层卷积系统提供支持。

### (2) AlexNet

AlexNet 是由 Geoffrey 和他的研究生 Alex 首次提出的, 并于 2012 年在 ImageNet 大规模视觉识别竞赛(ImageNet Large Scale Visual Recognition Challenge, ILSVRC)比赛中夺得冠军。Alexnet 共有八层结构, 前面五层为卷积层, 后三层为全连接层, 并在激活方法中选择非直线非饱和度的 relu 函数, 训练过程中梯度衰减速度变化时, relu 函数比常规神经网络中选择的非直线饱和度参数要快很多。同时 AlexNet 还在双 GPU 上工作, 大大提高整个网络系统的计算效率; 其引入的局部响应归一化方法(Local Response Normalization, LRN), 对非饱和函数 relu 而言, 并不要求对其输入加以标准化, 而 Alex 等人则认为在 relu 层添加了 LRN, 就会产生一定形式的横向抑制, 改善网络的泛化性能; 而池化方法则基于重叠池化(Overlapping Pooling), 即池化窗口的大小大于步长, 因此池化都有重合部分, 这种重叠的池化方法与一般无重叠的池化方法相比更具有有效性, 并能够减少过拟合。

### (3) VGGNet

VGGNet 是由牛津大学的计算机视觉中心与谷歌 DeepMind 研究院联合开发的深层卷积计算网络, 其网络深入研究了卷积神经网络的计算与其特性之间的关系, 并使用可以反复堆叠的小型卷积核和最大池化层, 从而成功建立 16~19 层深的卷积计算神经网络。在 2014 年的 ILSVRC 赛事上获得亚军以及定位项目的冠军。VGGNet 在图像特征提取方面依旧被广泛应用, 该网络在 top5 上的错误率仅为 7.5%。而 VGGNet 通过  $3 \times 3$  的卷积核和  $2 \times 2$  的池化核来改善其网络的结构和特性, 卷积核层数上的增加并不会引起参数在计算量上的巨增, 主要问题在于末尾的三个全连接层汇集大部分的卷积核层数。因此, 1 个  $5 \times 5$  的卷积层等同于 2 个串联的  $3 \times 3$  卷积层, 1 个  $7 \times 7$  的卷积层等同于 3 个串联的  $3 \times 3$  卷积层, 而 3

个串联的 $3 \times 3$ 卷积层与1个 $7 \times 7$ 卷积层的感受野是相同的，但是在参数数量方面3个串联的 $3 \times 3$ 卷积层只是1个 $7 \times 7$ 卷积层的一小部分。3个串联的 $3 \times 3$ 卷积层具有3个非线性操作，而1个 $7 \times 7$ 卷积层只有1个非线性操作，这也造成了3个串联的 $3 \times 3$ 卷积层在特征学习方面具有更好的表现能力。

#### (4) GoogLeNet

AlexNet在2012年ImageNet的图像分析比赛拿到了冠军，这也导致深度学习和卷积型神经网络研究获得高速进展。在2014年的ImageNet图像分析比赛上，GoogLeNet拿到了冠军的好成绩，其模型参数约为AlexNet的十二分之一。GoogLeNet的发展首先是因为Inception功能，而整个GoogLeNet的核心结构是由几个Inception功能所堆砌而成。在GoogLeNet以前的卷积计算神经网络，一般都是由多个卷积层和池化层所堆砌而成，然后接入一个或多个全连接层来预测输出。在卷积神经网络全连接层之前的卷积层和池化层的主要目的是为了获取各种图像特征，而这种图像特征对于在全连接层中的整个系统来说就被拉成一堆向量，通常这也造成了整个系统数据主要集中在全连接层，从而避免了过拟合，而在全连接层中一般都是采用丢弃（Dropout）的方法来减少过拟合的可能性。另外，池化方法主要包括平均池化层和最大池化层二种。平均池化层主要保存图像的背景数据，而最大池化方法的目的是保存图像数据，池化层的主要目的则是减少特征与网络参数，在GooLeNet之前通常使用最大池化层。但是最大池化层可能会造成空间信息的损失，大大降低模型的表达能力，为了解决此问题，Lin等人于2013年发明了Inception，Inception模块，主要在CNN中增加一个额外的卷积层，利用Relu作为激活函数，其主要作用是在不牺牲网络模型性能的前提下，实现网络特征的降维，减少计算量，有利于训练更深更广的网络<sup>[53]</sup>。

## 3.2 离线训练和在线测试模块

本文研究的结构如图3-4所示，该方法包括以下两部分：

离线训练：为了能够准确且快速实现驾驶员的疲劳识别，进行离线训练。驾驶员疲劳状态通过面部特征来识别，比如睁开眼睛的大小和打哈欠时张嘴的程度。因此，本文首先收集了驾驶员戴眼镜或太阳镜的驾驶员面部特征图像，例如睁眼、闭眼、张嘴的程度等。然后，应用任务约束的深度卷积网络（TDCDN）算法对所采集的每幅图像进行面部特征识别。该算法的优点是可以同时进行多任务学习（包括性别、是否戴眼镜或太阳镜以及面部姿势），这些辅助属性可以更好的帮助定位面部特征点。此外，本文还使用两种类型的闭眼时间百分比（PERCLOS）方法，如图3-4所示。PERCLOS1是戴眼镜或不戴眼镜睁眼时计算的值，

PERCLOS2 是驾驶员戴眼镜或睁眼状态时计算的值，也可以实时计算出嘴巴高度和宽度的比值（MAR）。最后，本文采取上述三种类型的数据作为样本进行训练，三种类型的数据分别为：睁眼时的 PERCLOS1，闭眼时的 PERCLOS2 以及嘴巴高度和宽度的比值。使用 AdaBoost 分类器来确定眼睛是睁开还是闭合，以及嘴巴是张开还是闭合。AdaBoost 是自适应增强功能的简写，这是由 Yoav Freund 和 Robert Schapire 制作的一个机器学习算法。其核心思想是为相同的训练集训练不同的分类器（弱分类器），然后将这些弱分类器结合在一起，形成更强的分类器 [54]。

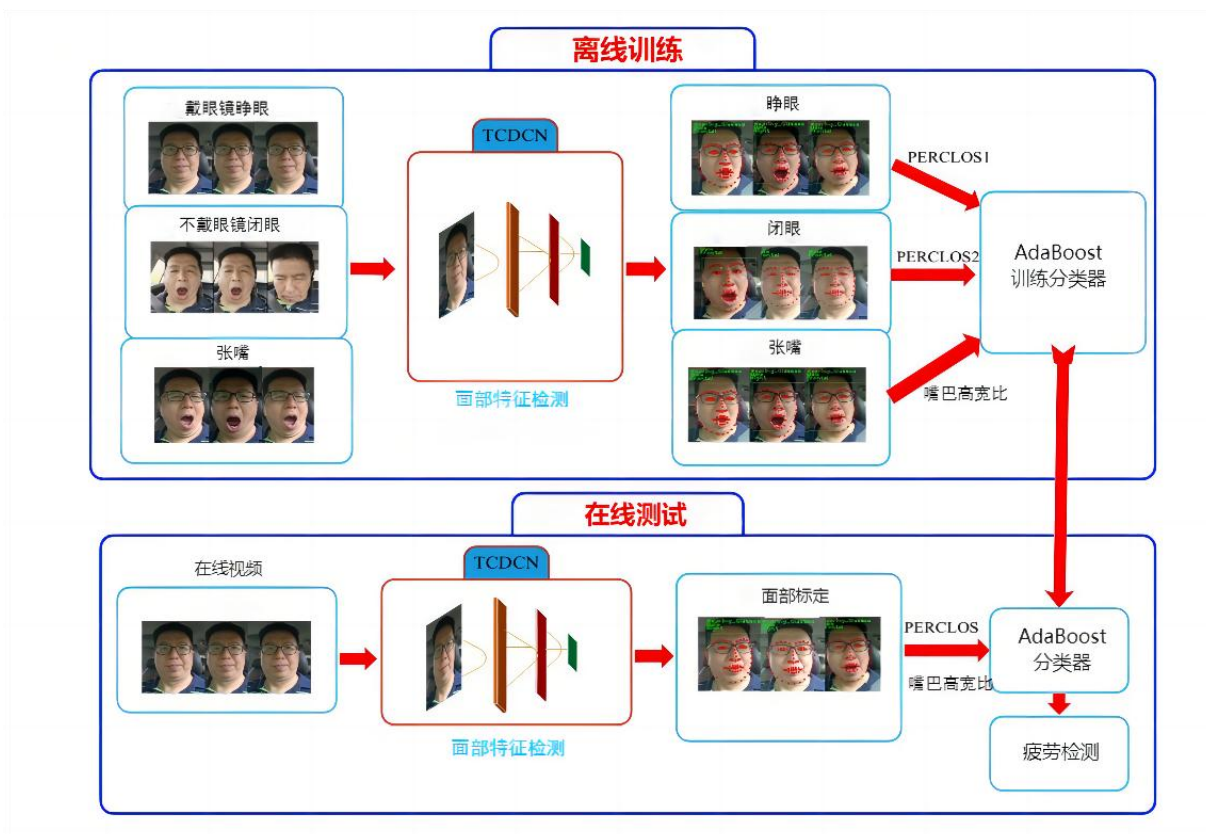


图 3-4 离线训练和在线测试模块

Fig.3-4 Offline training and online testing module

在线测试：是一个实时模块，从实时的视频序列中检测驾驶员的疲劳状态。将所有视频图像导入 TDCN 中识别驾驶员的面部，实时获得眼睛和嘴巴的特征地标<sup>[52]</sup>。然后，使用离线训练的 AdaBoost 分类器将驾驶员的眼睛和嘴巴的状态进行判断<sup>[55]</sup>。最后，根据一段时间内的平均闭合时间（或张嘴时间）超过选定阈值的情况，对驾驶员是否困倦进行分类。此外，为了更准确地识别驾驶员疲劳，TDCN 可以消除戴眼镜、头部姿势等因素的影响。



### 3.3 任务约束的深度卷积网络

TCDCN 算法的基本结构如图 3-5 所示。对于输入一个  $40 \times 40$  的灰度图像的特征获取将分为四个层次： $40 \times 40$  的灰度图像分别经过  $5 \times 5$ 、两个  $3 \times 3$  和  $2 \times 2$  不同尺寸的卷积层，其使用的三个池化层都是  $2 \times 2$ ，最后一次特征提取不再使用池化层；全连接层的目的是将相关的特征层进行分类，之后将分类好的特征层输出；把分类好的特征层分给不同任务处理，任务一代表性别，任务二代表面部位置，任务三代表戴眼镜。激活函数选择绝对值的双曲正切，并进行四次卷积后在全连接层可以获得特征向量。这些特征向量可以在计算中与多个任务共用，包括通过线性回归得到地标点的正确位置，以及通过多元逻辑回归处理多个其他任务，例如性别、面部位置、戴眼镜等。

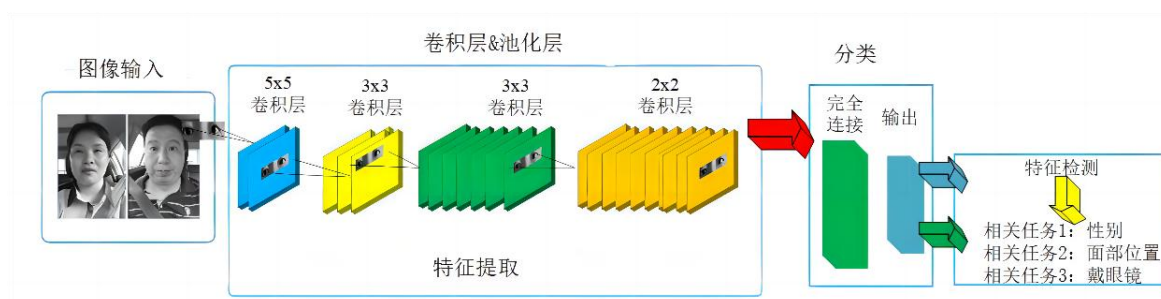


图 3-5 TCDCN 算法框架

Fig.3-5 The architecture of TCDCN

通常来说，驾驶员的面部特征在图像中时，首先会检测到面部在图像中的位置，然后确定性别、面部特征的位置、驾驶员是否戴眼镜等。当系统执行这些任务时，它们通常需要设计独立的算法来解决不同的任务；但是，TCDCN 可以设计一个复杂的网络来同时完成所有的任务，并利用这些任务之间的关系。

多任务学习的难度在于不同的任务具有不同的特征和收敛速度。本文提出的方法优于现有的方法，特别是在处理严重遮挡和位置变化的情况，并降低了模型的复杂性。

为能够让多任务学习上的特征得到应用，例如驾驶员的面部、面部的位置、性别、是否戴眼镜等信息，用多个网络训练不同任务的数据集，达到特征共享的目的。通过多任务学习的方式，TCDCN 网络整个域的特征参数得到共享，而不是适应特定的任务域。

在 TCDCN 的训练阶段，使用多任务面部地标（Multi-Task Facial Landmark, MTFL）数据集<sup>[56]</sup>作为训练数据，该数据集共有 12995 张人脸图像。图像上标定了面部位置、性别、是否戴眼镜和头部姿态的属性。

TCDCN 网络使用动态参数调整来达到模型的最优化，并在训练过程中引入



一个损失函数，目的是为了避开参数的不同对 TCDCN 网络模型训练的干扰。本文主要任务的损失函数是最小二乘法，交叉熵损失函数作为次要任务使用。主要的损失函数如 3-1 所示：

$$L(y, f(x)) = \sum_{i=1}^N (y_i - f(x_i))^2 \quad (3-1)$$

公式 3-1 中  $L$  为损失函数， $i$  表示第  $i$  个样本， $N$  表示样本总数， $y$  为预测值， $f(x)$  为真值。

辅助任务（性别、面部位置、是否戴眼镜）的交叉熵损失函数如 3-2 所示：

$$L(\gamma, p(\gamma | X)) = -\log p(\gamma | X) \quad (3-2)$$

公式 3-2 中损失函数  $L$  为分类  $\gamma$  条件下样本  $X$  的概率  $p(\gamma | X)$  达到的最大值。

通过对 TCDCN 的训练，可以准确地获得驾驶员的面部，为以下的算法提供了稳定的人脸图像。

### 3.4 驾驶员面部特征点

在本文中，疲劳是由驾驶员的眼睛和嘴巴状态来实时检测，获取眼睛和嘴巴的地标和形状是关键。Dlib 是一个用于机器学习算法和面部地标检测的开源工具包。在这项研究中使用 Dlib 的训练模型，用 OpenCV 校准和成像 68 个地标，并在面部绘制 68 个地标。如上所述，TCDCN 网络采用多任务学习模式学习面部相关的属性，能够更好检测到面部特征点的位置区域。具有相关属性的面孔姿态由人脸 68 个地标表示，如图 3-6 所示。整个面部由 68 个点组成，其中关键特征区域（比如眼睛、嘴巴和鼻子）也被标注出来。左眼和右眼各占用 6 个关键点，通过这 6 个关键点可以区分驾驶员是否眨眼睛；嘴巴的区域共由 20 个关键点组成，也可以通过这 20 个关键点之间的关系判断驾驶员是否打哈欠；鼻子部分总共占用 9 个关键点，本文不对鼻子的关键点以及除眼睛和嘴巴区域以外的关键点做讨论。下一章节会着重揭示了眼睛和嘴巴关键点的作用。

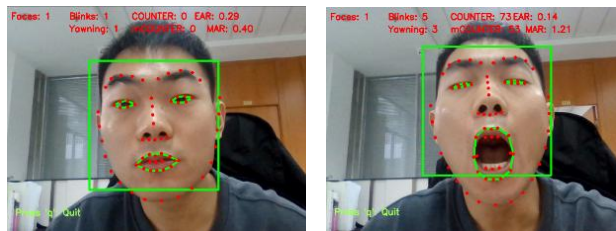


图 3-6 人脸特征点标定

Fig.3-6 Face feature point calibration

### 3.5 驾驶员疲劳识别特征

在回顾相关文献后，驾驶员面部表情有两种普遍的疲劳状态：打哈欠（嘴巴张开，在较长时间内保持这种状态）和眨眼（或者稍微闭上眼睛，眨眼增加或变慢）。本文研究主要关注眼睛睁闭状态（眼睛纵横比：EAR）、眼睛闭合百分比（眼睑覆盖眼球的比例：PERCLOS）和嘴巴高宽比（MAR）、权重比例之和（M）和视频帧数  $N$ 。基于这些数据，可以实时计算出驾驶员的注意力集中程度，并可以及时分析驾驶员是否处于昏昏欲睡的状态。

Ru 等人提出一个利用面部地标进行实时眨眼检测的模型概念<sup>[57]</sup>。然而，检测性能可能受以下实时的限制：（1）当驾驶员戴眼镜时，EAR 值波动很大；（2）处理图像像素具有挑战性。图像的像素值越低，内存占用越小，处理的速度就越快，但是图像需要一定的分辨率才能提高识别能力。

综上所述，本文提出的方法不同于传统的眨眼图像处理方法。使用 TCDCN 的 EAR、PERCLOS 和 MAR 是一个更简洁的解决方案。无论驾驶员是否戴眼镜，该方案都能准确识别眼睛的特征点。

### 3.6 本章小结

本章阐述了神经网络和深度学习之间的相互作用，对卷积神经网络的卷积层、池化层和全连接层进行了详细的阐述以及对一些常见且具有代表性的卷积神经网络模型进行阐述，例如 LeNet-5、AlexNet、VGGNet 和 GoogLeNet 网络。同时本文所采用的 TCDCN 中的离线训练和在线测试模块进行了详细的阐述，离线训练模块是为了能够准确且快速的实现驾驶员的疲劳识别进行的训练；在线测试模块是一个实时模块，从实时的视频序列中检测驾驶员的疲劳状态。通过任务约束的深度卷积网络，对 TCDCN 算法框架进行多任务训练，多任务训练可以准确快速地获得驾驶员面部的关键点以及面部特征，为后续的检测提供稳定的人脸图像。本章节最后对驾驶员面部标志及辅助任务、驾驶员疲劳识别特征和眼睛、嘴巴状态识别的突出特征进行了简单的阐述。



## 第 4 章 面部特征的标定与计算

### 4.1 基于 dlib 68 个地标及面部特征标定

Dlib<sup>[58]</sup>工具包中含机器学习算法，它在很多机器学习领域都有应用，并协助人们处理实际问题。在 Dlib 工具包中，由面部 68 个地标对检测的面部标志进行对齐如图 4-1 所示。在人脸的基准上将左眼、右眼、嘴巴以及面部轮廓等区域标定出来，因此称之为人脸对齐，例如点 37-40、43-46、49-68 分别为左眼、右眼、嘴巴的基准关键点，获取当前检测图像对应位置的关键点。总之，Dlib 的 68 个地标标记是根据左眼、右眼、嘴巴的动态变化，保证在面部五官不发生变形的基础上，使人脸与基准人脸重合。

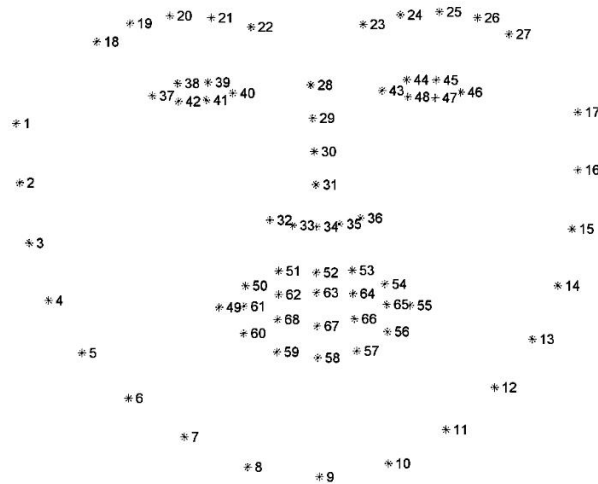


图 4-1 68 个人脸地标

Fig.4- 1 68 human face landmarks

### 4.2 EAR 计算

据统计，普通人每分钟眨眼 15 次左右，两次眨眼中间大约有 4 秒的间隔，嗜睡时眨眼的频率会加快。单位时间内累计的眨眼次数称之为眨眼频率，其计算如公式 4-1 所示：

$$B_{link} = \frac{N_B}{N_P} \quad (4-1)$$

公式 4-1 中,  $N_b$  为时间段内累计的眨眼次数, 由  $EAR$  来确定, 视频帧数用  $N_p$  表示。基于 Dlib 工具包获取的面部标志, 采用六个定位点对每只眼睛进行标定如图 4-2 所示。可以根据眼睛的坐标位置计算出  $EAR$  的值, 计算公式如 4-2 所示:

$$EAR = \frac{\|P_2 - P_6\| + \|P_3 - P_5\|}{2\|P_1 - P_4\|} \quad (4-2)$$

其中,  $P_2$ 、 $P_3$ 、 $P_5$ 、 $P_6$  关键点为纵坐标的值,  $P_1$ 、 $P_4$  关键点为横坐标的值。 $EAR$  值取左眼  $EAR$  值和右眼  $EAR$  值得的平均值。公式 4-2 中  $\|P_2 - P_6\| + \|P_3 - P_5\|$  表示上眼皮与下眼皮之间的垂直距离, 即为眼睛的高度;  $2\|P_1 - P_4\|$  是计算左眼角到右眼角的水平距离, 即为眼睛的宽度。因为垂直距离有两组数据, 而水平距离只有一组数据, 为了确保两组数据的权重一致性, 所以水平距离要乘以 2, 依靠眼睛纵横比确定驾驶员是否眨眼。

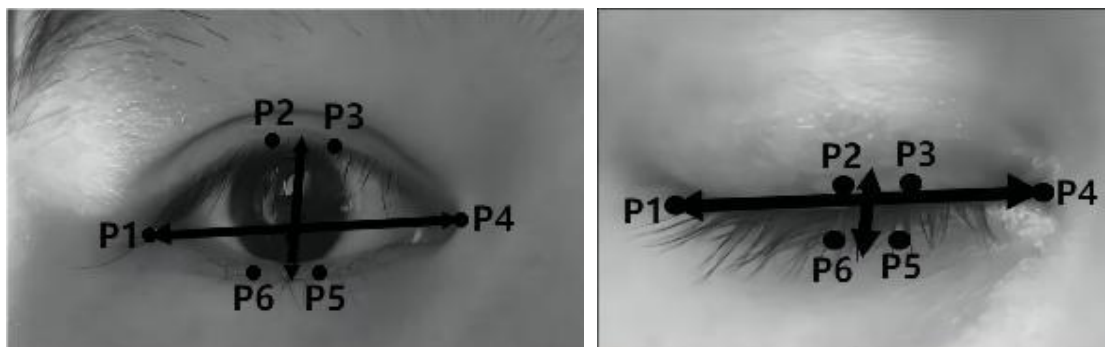


图 4-2 眼睛关键点标定

Fig.4-2 Eye key calibration

一次完整闭眼过程的  $EAR$  变化曲线如图 4-3 所示, 图中横纵坐标分别是视频的帧数和  $EAR$  值, 设置  $EAR$  阈值为 0.2, 通过 K 最近邻 (K Near Neighbor, KNN) 分类器将  $EAR < 0.2$  的值进行权值计算, 权值比重为 0.2。  $EAR$  值取左眼的  $EAR$  值 ( $leftEAR$ ) 和右眼的  $EAR$  值 ( $rightEAR$ ) 的平均值, 模拟一组驾驶员眨眼或闭眼计算  $EAR$  变化曲线, 如图 4-4 所示。

从图 4-4 曲线图中可以看出有两次眨眼和三次闭眼过程, 帧数在 34 到 38 帧和 118 到 122 帧为眨眼过程, 171 到 184、300 到 315 和 449 到 462 帧为闭眼过程与模拟过程中眨眼和闭眼的次数相同, 分析结果如表 4-1 所示。从模拟驾驶员眨眼或闭眼以及实际疲劳检测过程得到的  $EAR$  曲线图, 可以得出: 系统在检测驾驶员眨眼或闭眼过程中, 几乎不存在漏检或者误检, 检测的实时性和计算的准确度达到了较高的水平。

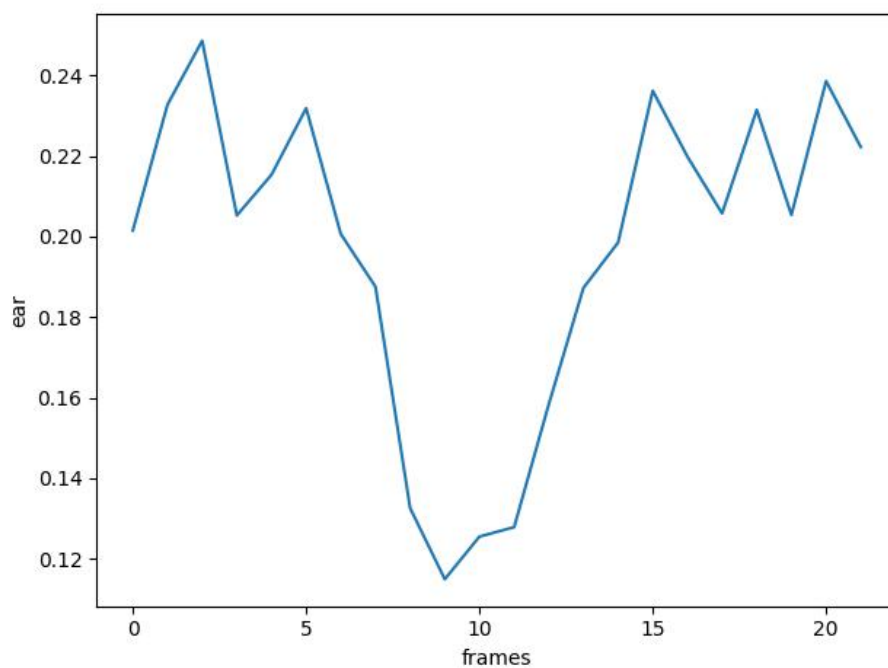


图 4-3 EAR 曲线图

Fig.4- 3 EAR curve

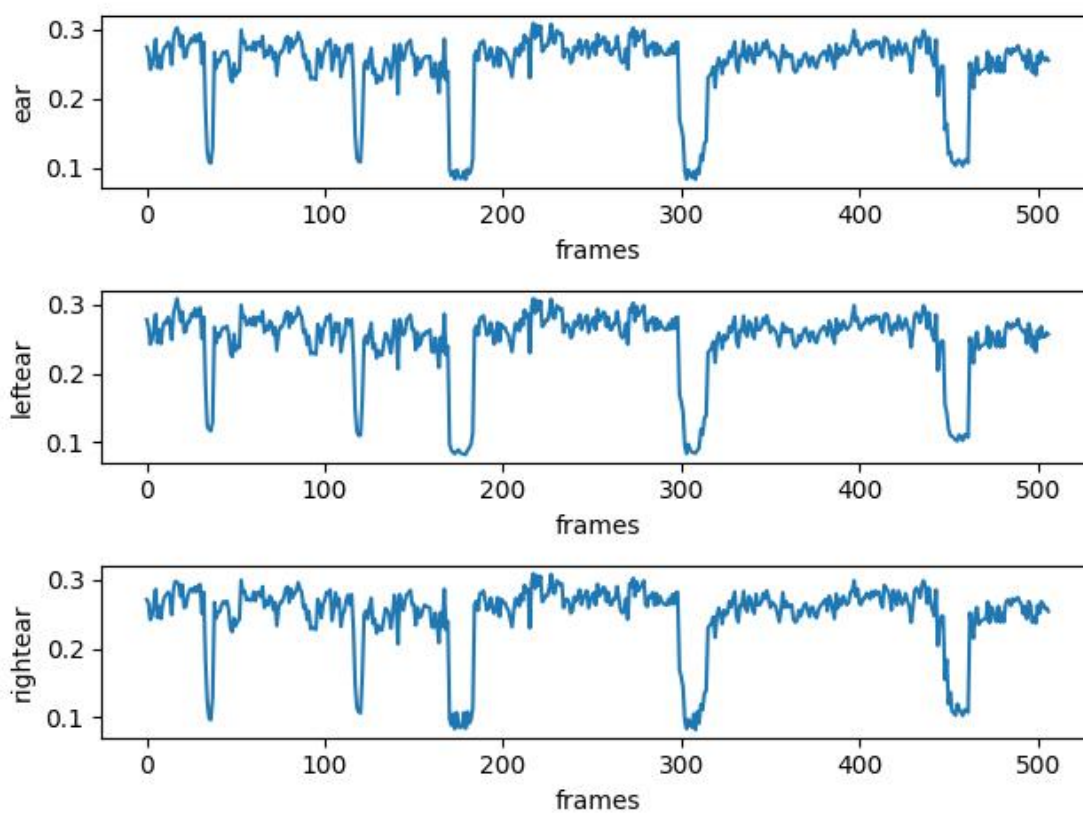


图 4-4 EAR 综合曲线图

Fig.4- 4 EAR comprehensive curve

表 4-1 EAR 曲线分析结果  
Table 4-1 Curve analysis results

眼睛状态	区段	帧数
眨眼	34~38	5
眨眼	118~122	5
闭眼	171~184	14
闭眼	300~315	16
闭眼	449~462	14

### 4.3 PERCLOS 值计算

卡内基梅隆研究院的研究员经过长时间的测试与验证，最终给出测量疲劳的物理量值 PERCLOS，并得到了 NHTSA 的认可。PERCLOS 定义为一定时间内眼睛的闭合度，PERCLOS 是疲劳检测最有效的参数。由于眼睛的大小因人而异，眼睛的范围因受所在环境和头部运动的限制，眼睛的睁开程度是相比于自身的最大睁开状况而言的，因此可将时间转换为视频帧数进行计算<sup>[59]</sup>。

随着持续时间的延长，当眼睑覆盖瞳孔面积超过一定的百分比时被定义为在一定时间内闭眼的时长，它是一个度量疲劳状态的有效手段。由于眼睛的大小因人而异，眼睛的大小受到环境因素和头部活动的限制而动态变化。时间可以转换为视频的帧数进行计算，驾驶员的嗜睡程度可以根据阈值进行判断<sup>[60]</sup>。

目前，PERCLOS 方法有三种判断眼睛闭合的不同准则，分别为 EM 准则、P70 准则、P80 准则。在实际研究中，研究者普遍认为最有效的是 P80 准则，P80 准则表示一定时间内眼睑覆盖眼球的百分比超过 80% 的比例。

PERCLOS 原理曲线的横坐标表示时间  $t$ ，单位为秒；纵坐标为瞳孔被眼睑覆盖的百分比。详细的测量方法是：用 1080 的摄像头获取驾驶员的面部图像，定位出每一帧图像的眼睛区域，对眼睛特征区域进行提取，通过计算在一定时间内眼睑覆盖眼球的百分比超过 80% 的帧数，确定闭眼所占总时间的比例。

一周期的闭眼百分比如式 4-3 所示：

$$PERCLOSE(t_0 \sim t_5) = \frac{t_3 - t_2}{t_5 - t_0} * 100\% \quad (4-3)$$

闭眼时长如式 4-4 所示：

$$T(t_0 \sim t_5) = t_3 - t_2 \quad (4-4)$$

同样，在一定时间内可以计算得到 *PERCLOS* 值和眼睛的平均闭眼时长，假设在一定时间  $T$  内包含有  $n$  个周期， $T_k$  为  $k$  个闭眼周期内眼睑覆盖眼球的比例超过 80% 所用的时间，当  $k < n$  时可以获得公式 4-5：

$$PERCLOSE = \frac{\sum_{k=1}^n T_k}{T} \quad (4-5)$$

平均闭眼时长如式 4-6 所示：

$$t = \frac{\sum_{k=1}^n T_k}{n} \quad (4-6)$$

通过连续 5 帧图像的 EAR 值，得出眼睑覆盖眼球的平均百分比为 *PERCLOS* 值，*PERCLOS* 值的曲线图如图 4-5 所示，横坐标为视频的帧数，纵坐标为眼睛闭合的百分比，虚线代表的值为 0.8。*PERCLOS* 的计算过程是依次取 *EAR* 值的每一帧，每连续 5 帧循环一次计算，步长为 1。

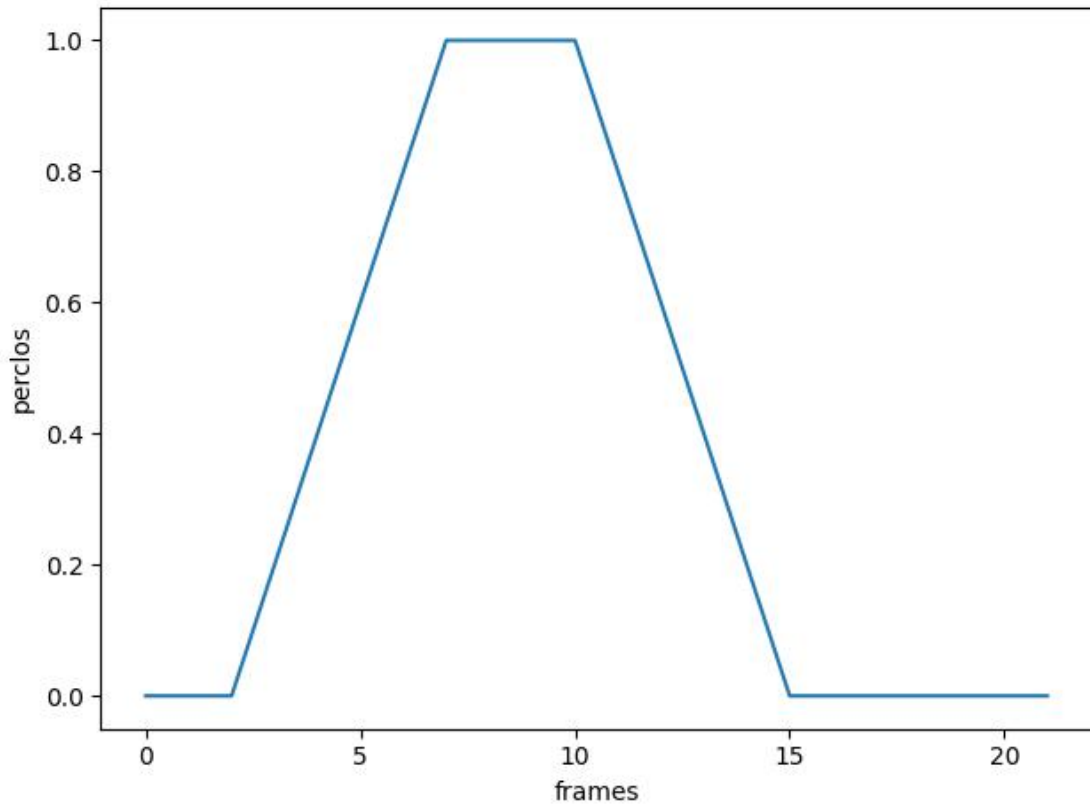


图 4-5 PERCLOS 曲线图

Fig.4-5 PERCLOS curve

#### 4.4 MAR 值计算

驾驶员嘴巴状态通常有三种情况，分别为闭合、说话和打哈欠，在疲劳状态



下，人会频繁的打哈欠。提取嘴巴区域的关键点类似于眼睛区域疲劳特征的提取，嘴巴区域的关键点分布如图 4-6 所示。

嘴巴的高度  $H$  和宽度  $W$  计算公式如 4-7 所示：

$$\begin{cases} W = |x_{\max} - x_{\min}| \\ H = |y_{\max} - y_{\min}| \end{cases} \quad (4-7)$$

$H$  与  $W$  的比值  $MAR$  如式 4-8 所示：

$$MAR = \frac{H}{W} \quad (4-8)$$

公式 4-8 中的  $x_{\max}$ 、 $x_{\min}$  分别为 68 个面部关键点中的第 54 和第 48 关键点的位置，取两者之间差值的绝对值为嘴巴的宽度，记作  $W$ ；68 个面部关键点中第 56 个和第 58 关键点表示  $y_{\max}$  以及第 50 个和第 52 个关键点表示  $y_{\min}$ ，同样两者之间的差值的绝对值为嘴巴的高度。

打哈欠的频率  $f_{yawn}$  如式 4-9 所示：

$$f_{yawn} = \frac{N_y}{N_T} \quad (4-9)$$

$N_y$  表示在  $T$  时间段内累计打哈欠的次数， $N_T$  表示视频的帧数。

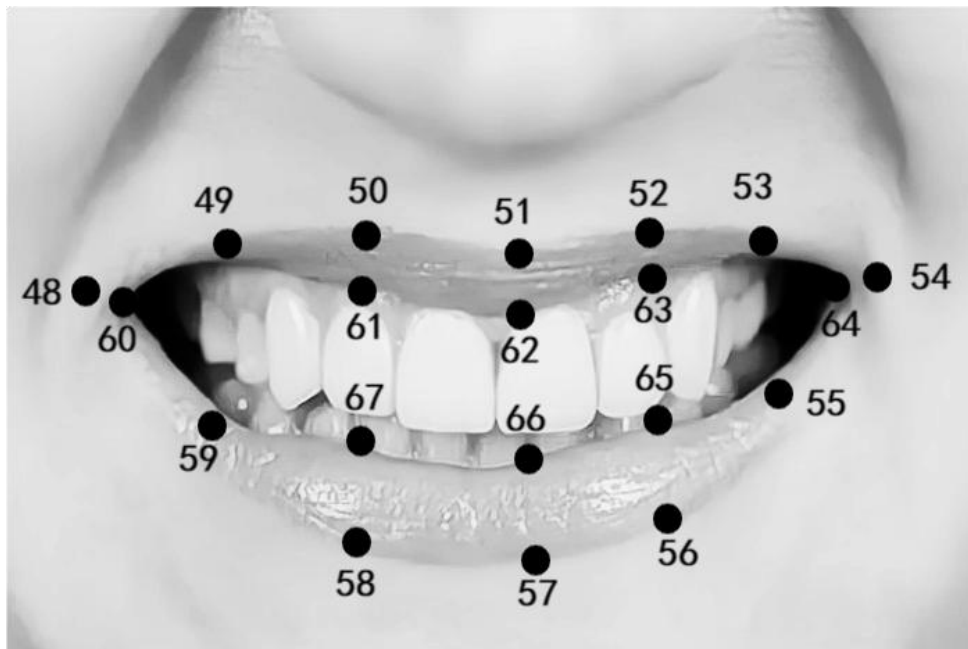


图 4-6 嘴巴关键点标定

Fig.4-6 Mouth key calibration

驾驶员嘴部状态的检测如图 4-7 所示，横坐标为视频帧数，纵坐标为  $MAR$  值。图 4-7 (a) 为驾驶员嘴巴闭合状态， $MAR$  值在 0.350 到 0.375 之间；图 4-7

(b) 为驾驶员嘴巴说话状态, MAR 值在 0.375 到 0.60 之间; 图 4-7 (c) 为驾驶员嘴巴打哈欠状态, MAR 值在 0.60 以上。通过 KNN 分类器将 MAR 大于 0.60 的值进行权值计算, 权值比重为 0.1。

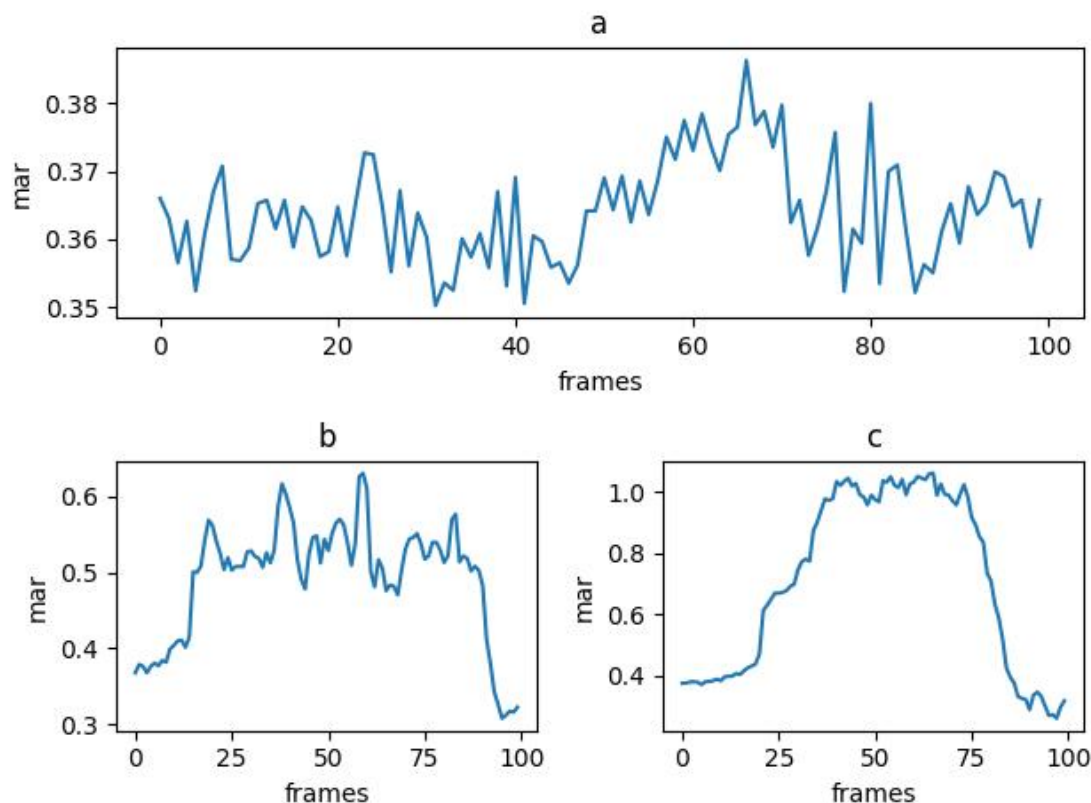


图 4-7 嘴巴状态曲线图

Fig.4-7 Mouth state curve

## 4.5 疲劳状态在线检测技术路线

如前所述, 驾驶员疲劳状态是一个动态变化的过程, 需要实时进行检测。本文提出一种判断驾驶员疲劳状态的方法, 与传统的单一疲劳指标相比, 该方法具有识别精度高和可靠性好的特点。

驾驶员疲劳状态在线监测系统的流程图如图 4-8 所示, 系统工作过程如下:

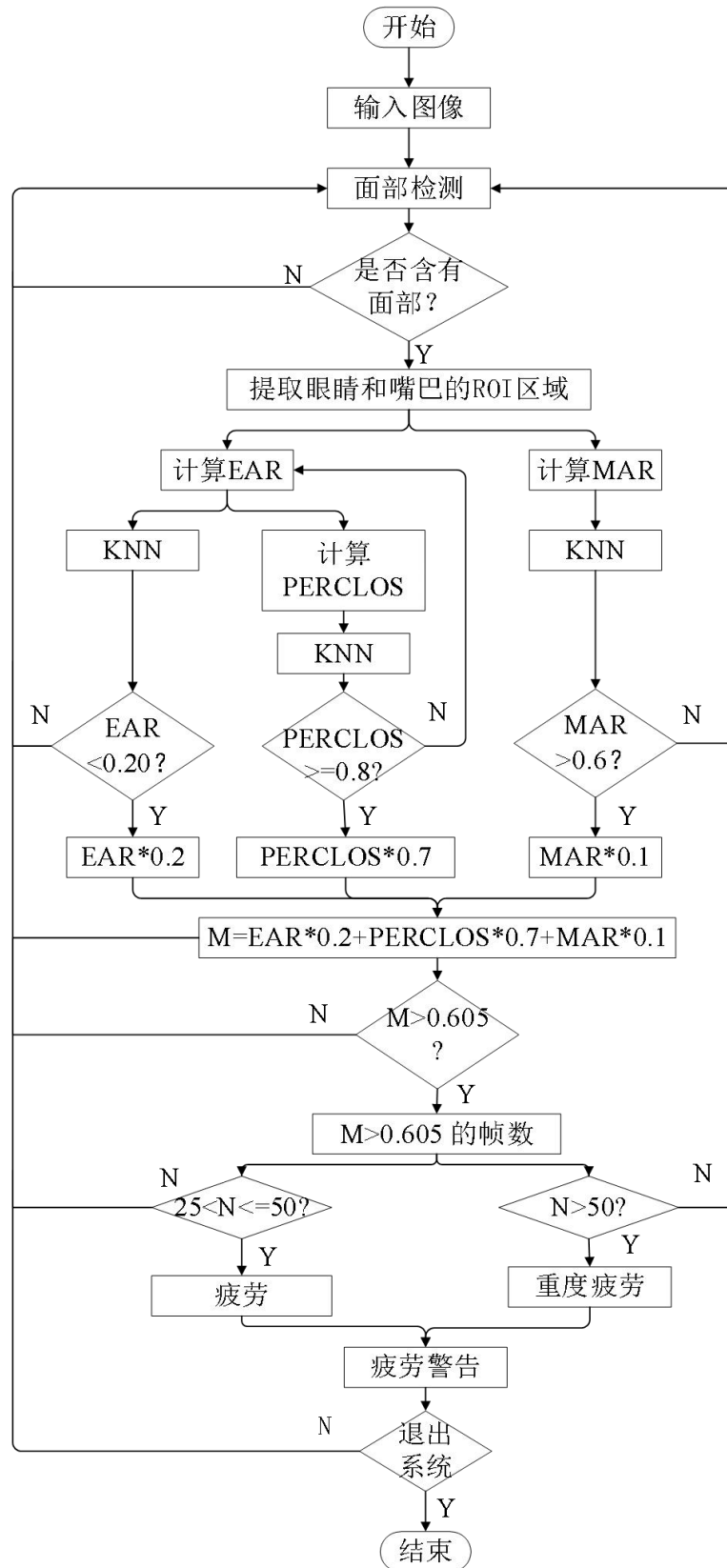


图 4-8 在线监控工作流程图

Fig.4- 8 Online monitoring work flow chart

(1) 摄像头采集驾驶状态下的视频输入系统中, 系统对视频中的每一帧图像进行检测, 筛选并过滤包含驾驶员面部的图像。

(2) 筛选出含有驾驶员面部的图像并进行 ROI 特征提取。

(3) 针对提取到的 ROI 面部区域特征, 获取眼睛纵横比和嘴巴高宽比, 计算 EAR 和 MAR 值。根据得到的 EAR 值, 计算 PERCLOS 值, 并将得到的 EAR、PERCLOS 和 MAR 值分别应用于 KNN。

(4) 通过实验测试分析得出, 将 EAR 的阈值设置为 0.2, 根据 P80 准则, 将 PERCLOS 阈值为 0.8。

(5) 当 MAR 值大于 0.6 时, 表示驾驶员打哈欠, 进而将 MAR 阈值设置为 0.6。

(6) 将 EAR 和 PERCLOS 的权值分别设置为 0.2 和 0.7, MAR 值为辅助值, 其权重设置为 0.1。

(7) 通过实验测试分析可知, 将 M 阈值设置为 0.605 较为合理。在这一部分中, N 被记录为  $M > 0.605$  的帧数, 即驾驶员在一定时间内闭上眼睛累积的时长。

(8) 根据实验结果, 将驾驶员状态划分为三种情况, 当  $0 < N \leq 25$  时为清醒状态; 当  $20 < N \leq 50$  时为疲劳状态;  $N > 50$  为重度疲劳状态。

#### 4.5.1 KNN 分类器

利用输入数据训练一个 KNN 分类器。KNN 算法<sup>[61]</sup>广泛应用于数据的分类和处理领域中, 同时是一种实用的分类器模型, 其功能是能够划分和识别任何指定的样本数据点。完成这一目标需经历两个过程, 一是收集完成目标所使用的数据集, 二是设置 K 个分类阈值来求得样本数据参考点与其他任意数据点之间的距离。

首先, 在 KNN 分类器模型中, 经常以数据集  $D(x, y)$  与设置阈值 K 的样本参考点之间的计算距离当作被测数据的参照指标, 从而可以更高效地处理样本点  $(x, y)$  之间的匹配优化问题。然后, 再选择欧式距离  $D(x, y)$  用作被测样本数据的衡量标准, 欧式距离  $D(x, y)$  的计算公式如 4-10 所示。

$$D(x, y) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \dots + (x_n - y_n)^2} = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (4-10)$$

再选择 K 的数值。在 KNN 算法中, 计算的精确度较大程度上取决于 K 值的选择。所以, 一旦数据集分类器中的 K 值设置过大或过小, 都有可能使分类结果出现错误, 不利于分类模型的最优拟合结果。针对不同的数据集, 并没有一定的具体应用数值 (通常默认设置为 5)。通常情形下, 可以根据样本数据集参数的分

布情况，选取较小的参数。另外，经过反复交叉迭代验证后，可得出一个适合的  $K$  值。

$K$  值的选择对算法结果有显著的影响。如果  $K$  值很小，则相当于在一个小的领域中使用训练示例进行预测。在极端情况下， $K = 1$ ，测试示例只与最近的样本相关，并且训练的误差非常小。若样本遇到噪声，那么预测是错误的，而且测试的错误非常大。当  $K$  的数值比较小时，会出现过拟合；当  $K$  数值比较大时，则相当于在一个更大的区域中使用训练示例加以检测。在特殊情况下  $K = n$ ，如果测试示例是训练数据集中实例数量最多的类，也会出现过拟合。本文一般选择较小的  $K$  值， $K$  值为 7。

#### 4.5.2 面部多特征融合

对识别出驾驶员面部的五个特征点，由 ROI 以特征点为中心对眼睛、嘴巴和鼻子的特征区域进行提取，这些特征区域经过 TCDCN 获取眼睛和嘴巴的特征地标，再由离线训练的 AdaBoost 分类器对睁眼、闭眼、张嘴和闭嘴状态进行分类，筛选出闭眼和张嘴的图像。闭眼和张嘴的图像根据特征地标计算出 EAR 和 MAR 值，在 EAR 值的基础上计算出 PERCLOS 的值；利用 KNN 对 EAR、PERCLOS 和 MAR 进行分类，筛选出 EAR 小于 0.2、PERCLOS 大于等于 0.8 和 MAR 大于 0.6 的值。对 EAR、PERCLOS 和 MAR 特征参数进行融合，融合得到的参数定义为  $M$ 。 $M$  为 EAR、PERCLOS 和 MAR 按照各自不同的权值比重进行相加所得，其中 EAR 占 20%、PERCLOS 占 70% 和 MAR 占 10%。将  $M > 0.605$  的帧数  $N$  来判断驾驶员的疲劳程度， $N \leq 25$  时标记为清醒状态； $25 < N < 50$  时标记为疲劳； $N \geq 50$  时为重度疲劳。

#### 4.6 本章小结

本章对 dlib 68 个地标做简要的说明，这是面部特征点标定的主要依据。随后针对标定的眼睛和嘴巴区域获得眼睛纵横比和嘴巴高宽比，从而计算得到 EAR 值和 MAR 值。根据 EAR 和 MAR 值的大小判断驾驶员是否眨眼和打哈欠，从而检测驾驶员眼睛和嘴巴的状态。在 EAR 值的基础上计算 PERCLOS 值，通过 P80 准则来判断眼睑覆盖眼球的百分比，此值作为判断疲劳驾驶的重要参数。最后详细阐述疲劳检测系统的在线监控工作流程图、KNN 分类器以及面部多特征融合。

## 第 5 章 实验数据和结果

在本章中，首先评估所提出的 TCDCN 在人脸检测 Fddb 数据集中的有效性，然后讨论 EAR、MAR 和 PERCLOSE 之间的相关性。通过疲劳主观问卷调查对采集的视频进行清醒、疲劳和重度疲劳状态进行标记，疲劳检测系统对标记的视频进行验证，确定检测系统的准确率。

### 5.1 环境和数据集

该实验平台包括了一个具有 X86 架构的英特尔酷睿 i7-10750（主频率：2.60GHz）、NVIDIA GeForce GTX 1650、16GB DDR4、OpenCV 3.4.11 图像库、Dlib 19.17.0 工具包和深度学习框架 Tensorflow 2.5。

在本文中，主要采用了以下两种类型的数据集进行训练。第一种为脸部检测及标准数据库（Face Detection Data Set and Benchmark, Fddb），它是一种专门用来进行无约束人脸检测的数据集，该数据集包括 5171 个面孔如图 5-1 所示。



图 5-1 Fddb 数据集

Fig.5- 1 Fddb dataset



第二种是志愿者在不同场景下采集的视频数据集，收集驾驶车辆的视频数据集，如图 5-2 所示。本文收集的视频数据集来自 8 名司机（20 至 45 岁，4 名女性司机和 4 名男性司机）和 30 次驾驶测试。所有的参与者都要求有驾驶证，每年开车里程超过 2000 公里，并且没有相关的疾病、吸毒、和饮酒记录。测试要求在测试前避免饮酒或喝含咖啡因的饮料，参与者在测试前需要熟悉车辆驾驶状态和驾驶路线，驾驶路线均在封闭安全路段。每次驾驶测试的持续时间分别为 30 分钟、60 分钟和 120 分钟，驾驶路线是一个模拟双车道的高速公路，有一些平稳的转弯，参与者在测试期间要求连续行驶约 40 公里、80 公里和 160 公里。



图 5-2 不同场景下的数据集收集

Fig.5-2 Data collection in different scenarios

## 5.2 驾驶员疲劳驾驶检测系统

为了能够实现在实际环境中验证本文第四章在线监控系统的真实性，本文所采用的疲劳驾驶流程如图 5-3 所示。疲劳检测终端和后台分析系统两者构成一个完整的疲劳检测系统。疲劳检测终端主要分为摄像头（采集驾驶员的面部状态，传输给图像处理单元）、图像处理单元（对摄像头传输过来的图像进行分析处理，并将驾驶员的状态传输给后台分析系统）、语音警告设备（对驾驶员处于疲劳状态驾驶时，进行语音警告提醒：您已疲劳驾驶，请注意休息！）三个部分。其次，后台分析系统主要放在驾驶室内，主要对终端传输过来的驾驶员状态信息进行分析，并让语音警告设备对驾驶员做出警告提醒。

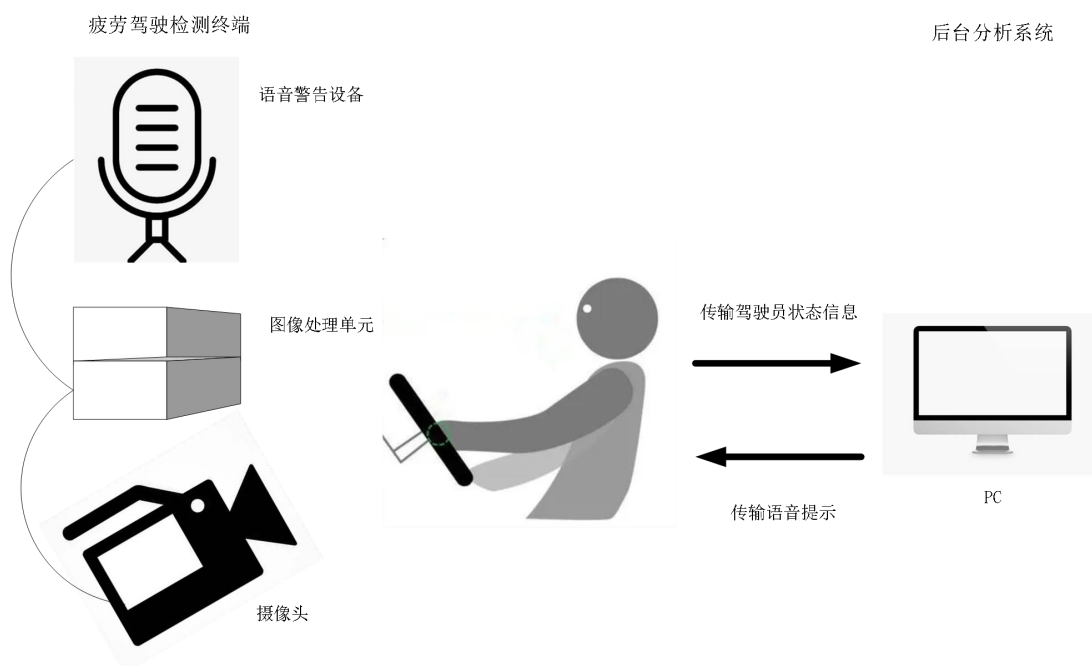


图 5-3 疲劳驾驶检测系统结构

Fig.5-3 Structure of fatigue driving detection system

### 5.2.1 软件系统

疲劳驾驶检测系统的软件部分也包括两部分：一是疲劳驾驶检测终端的软件部分；二是后台分析系统的硬件部分。下面将详细地对这两个部分基本构成加以说明。

#### 1、疲劳驾驶检测终端

本文基于 python 脚本语言完成疲劳驾驶检测终端的软件部分，一方面，部分和终端以及后台分析系统的交互部分由 python 语言来实施算法进而得以实现，另一方面，整个 python 程序由脚本进行激活。

对于整个算法的实现，疲劳驾驶检测系统采用 tensorflow 作为深度学习的框



架，来编写程序从而实现卷积神经网络，采用 Dlib 68 个地标对人脸特征点进行标定，采用 opencv 和 opencv-contrib-python 对图像进行处理。疲劳检测系统的摄像头负责采集图像，TCDCN 网络负责人脸检测、眼睛和嘴巴定位，采用 KNN 分类器对眼睛状态和嘴巴状态进行分类。对图像的帧数进行统计，图像的帧数达到一定的帧数，则计算这段时间内的 EAR、PERCLOS、MAR 和 M 的值，其各自的阈值分别为 0.2、0.8、0.6 和 0.605，各自的值与其各自的阈值进行比较，计算出 M 值。在一定帧数内连续累加 M 大于 0.605 的帧数得到 N 值，通过 N 的大小来判断驾驶员的状态。在脚本语言部分，为了实现能够自动启动运行算法部分，使疲劳检测系统能够完全实现自动实时监督驾驶员的驾驶车辆状态。本文采用 bat 批量处理脚本程序。

## 2、后台分析系统

本文的检测系统在 python 程序语言的基础上，设计一种专门用于疲劳驾驶检测系统的管理程序，使用图片服务器完成终端与后台之间的图像互传，并通过云服务完成终端与后台之间的文字、语音通讯。当发现驾驶员正处在疲劳驾车状况时，后台语音系统会向疲劳检测终端发出语音提醒驾驶员已经处于疲劳驾驶，请注意安全驾驶。

为了验证上述驾驶员疲劳驾驶检测算法的准确性，本文分别进行了短期测试和长期测试进行验证。短期测试验证的目的主要是识别驾驶员昏昏欲睡驾驶的行为。长期测试主要验证是所提算法能否准确判断驾驶员在长期驾驶时是否打瞌睡，进而根据程度（清醒、疲劳和严重疲劳）进行分类。

## 5.2.2 硬件系统

疲劳驾驶检测系统的硬件部分包括两个部分：一是疲劳驾驶检测终端的设备部分；二是后台分析系统的硬件部分。以下将详细地对这两个部分的构成进行说明。

### 1、疲劳驾驶检测终端

疲劳驾驶检测终端分为两种：驾驶模拟器采集终端和实车道路测试终端，实验中所用的实验设备如图 5-4 所示。驾驶模拟器采集终端主要由三块 32 英寸的显示屏组成，可以形成 130 度的驾驶视野，经过多次不同调整，将摄像头安装在显示屏上方。通过驾驶模拟器模拟不同的驾驶场景，采集驾驶员清醒、疲劳和重度疲劳三种状态下的样本。实车道路测试终端主要由摄像头、笔记本电脑、汽车三部分组成，摄像头安装在驾驶员面部正前方的中控台上，实时监控驾驶员的面部状态，将采集到的视频传输到笔记本终端。图 5-3 的语音设备由电脑笔记本上的

麦克风和扬声器所代替。



图 5-4 实验设备

Fig.5-4 Experimental equipment

## 2、后台分析系统

后台分析系统由一台联想 Y7000 笔记本电脑所代替，电脑安装 Windows10 的操作系统并插入无线网卡，笔记本电脑放在驾驶室内的副驾驶座椅上。笔记本电脑负责分析驾驶员的面部特征信息，实时的将驾驶员的驾驶状态显示在显示屏上。

## 5.3 短期实验

为了验证所提出具有个体差异的疲劳驾驶检测算法的有效性，本文进行了一系列实验。首先，使用疲劳阈值 0.605、EAR 阈值 0.2、PERCLOS 阈值 0.8 和

MAR 阈值 0.6 分析疲劳检测的准确性。然后，通过未打哈欠和连续打哈欠的驾驶员驾驶车辆来确定疲劳识别算法中的核心阈值是否有效。

短期实验将根据眼睛和嘴巴的不同状态进行实验，图 5-5 为未打哈欠状态下驾驶员的疲劳检测结果，图 5-6 为打哈欠状态下的驾驶员疲劳检测结果。图 5-5 和图 5-6 中的纵坐标分别为 EAR、MAR、PERCLOS 和 M 的值，横坐标为视频帧数（frames），灰色点划线为 EAR、MAR、PERCLOS 和 M 各自的阈值。图 5-5 中的 EAR 图为视频帧数在 60 帧附近左眼和右眼一次眨眼的过程，EAR 值取 leftEAR 和 rightEAR 值的平均值。从图中可以看出眼睛在正常的情况下 EAR 值均大于 0.2，发生闭眼的时 EAR 值迅速减小且小于 0.2，因此设置 EAR 的阈值为 0.2，其权重比例为 0.2。通过对图 5-5 和图 5-6 中的 MAR 曲线图的对比，人在打哈欠时 MAR 的值会迅速增加，未打哈欠的状态下 MAR 的值保持在 0.33 附近，通过多次实验验证设定 MAR 的阈值为 0.6，考虑打哈欠的频率较低将其权重比例设置为 0.1。根据 P80 准则设定 PERCLOS 的阈值为 0.8，此值为判断是否疲劳驾驶的重要参数，其权重比例为 0.7。图 5-5 和图 5-6 中 M 的曲线图分别是在闭眼过程中未打哈欠状态下和打哈欠状态下的曲线图，图中红色曲线部分为超过 M 的阈值 0.605 的部分。

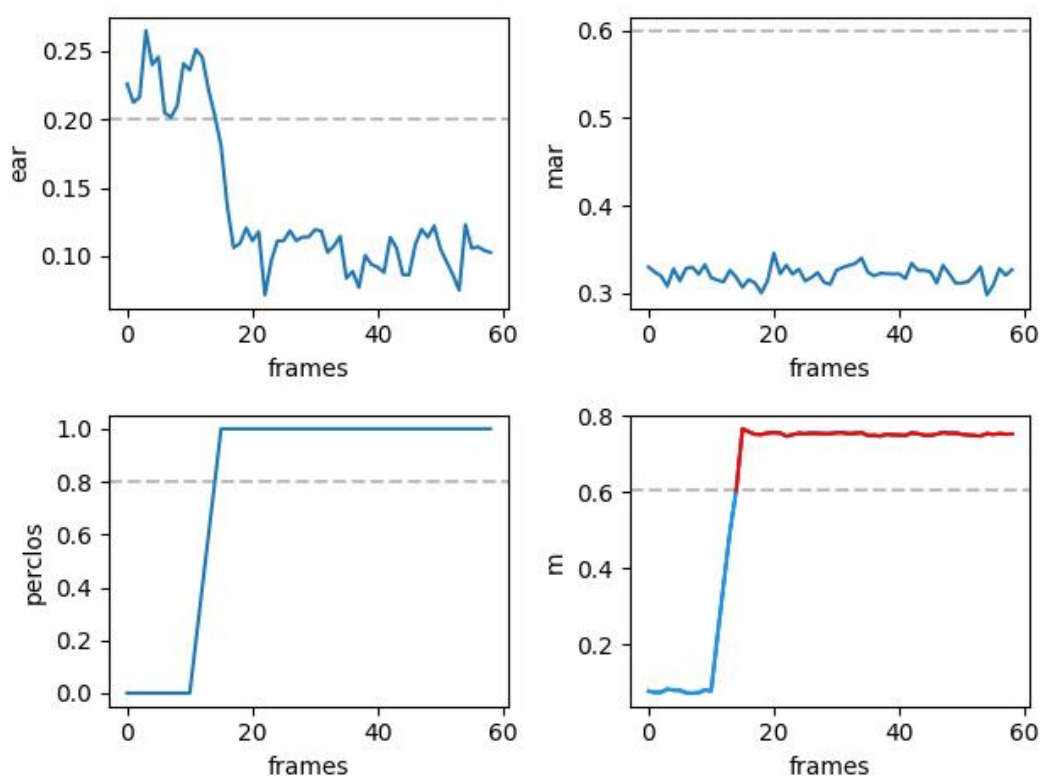


图 5-5 未打哈欠驾驶员的疲劳检测结果

Fig.5-5 Fatigue detection results for drivers who do not yawn

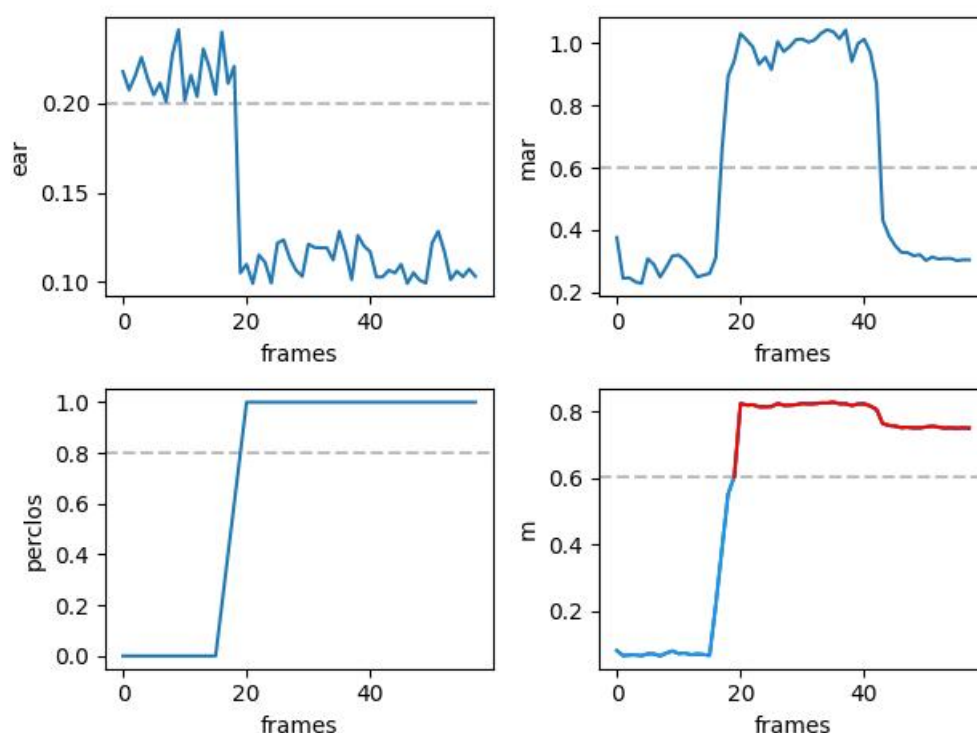


图 5-6 打哈欠驾驶员的疲劳检测结果

Fig.5-6 Fatigue detection results for drivers who do not yawn

对于图 5-5 和图 5-6 的分析，当驾驶员未打哈欠时  $M$  大于 0.605 的帧数相对较少；驾驶员在打哈欠的情况下往往伴随着长时间闭眼，使  $M$  大于 0.605 的帧数相对增加。驾驶员开始打哈欠时身体会出现疲劳犯困的情况，从图 5-6 可以看出驾驶员长时间打哈欠时  $M$  大于 0.605 值的帧数在 25 到 50 之间，可以给予驾驶员警告提醒。图 5-5 和图 5-6 显示了在睡意时间内对眼睛、嘴巴和疲劳阈值  $M$  的识别结果。

## 5.4 长期实验

在短期测试中，发现 EAR、PERCLOS、MAR 和疲劳警告阈值  $M$  与驾驶员疲劳状态之间的关系。为了直观的评估在不同驾驶条件下眼睛和嘴巴状态的变化，依据提前设计好的路线对驾驶员进行实车实验。长期实验检测的结果如图 5-7 和图 5-8 所示，表示男生驾驶员的驾驶状态。图 5-7 在整个驾驶期间 EAR、MAR、PERCLOS 和  $M$  的值都远远偏离它们的阈值， $M$  值连续红线部分的帧数增加，超过一定范围表明驾驶员处于疲劳状态下；而图 5-8 出现一些红色部分的区域，是由于驾驶员偶尔眨眼或者打哈欠的情况下导致 EAR、PERCLOS、MAR 和  $M$  的值偏离了它们的阈值，表明驾驶员在长期驾驶期间是清醒的。

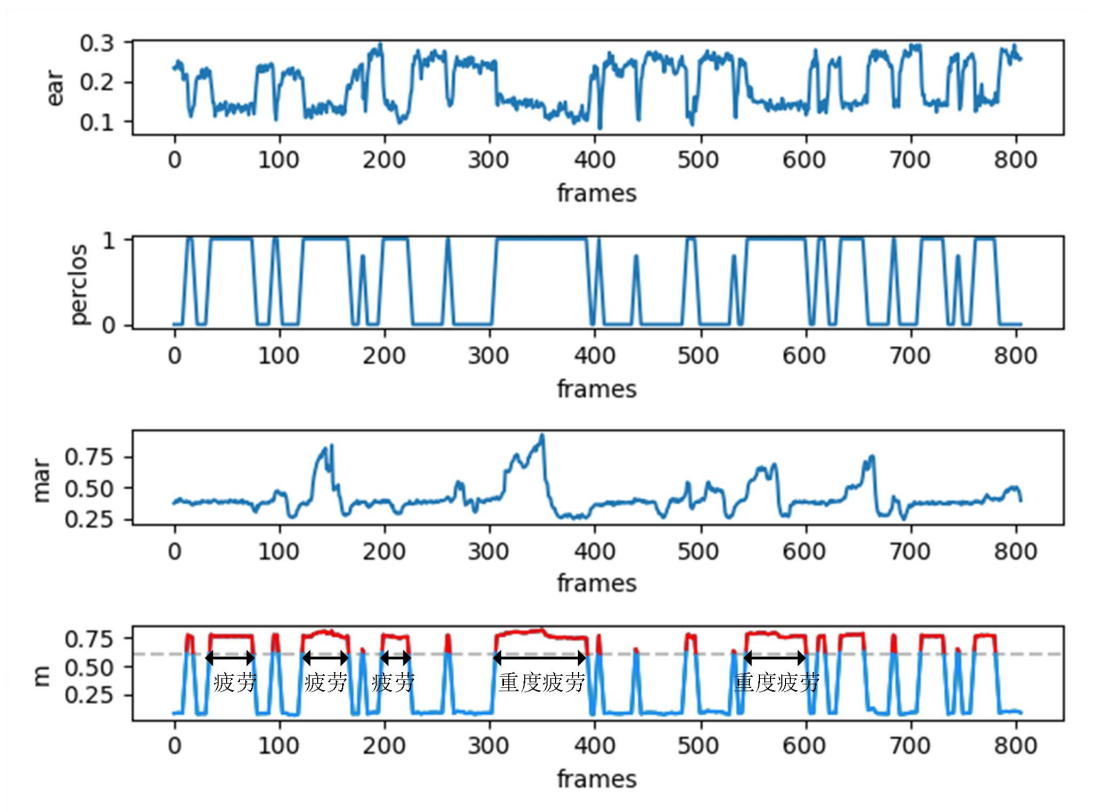


图 5-7 男生疲劳驾驶状态

Fig.5- 7 Fatigue driving state of men

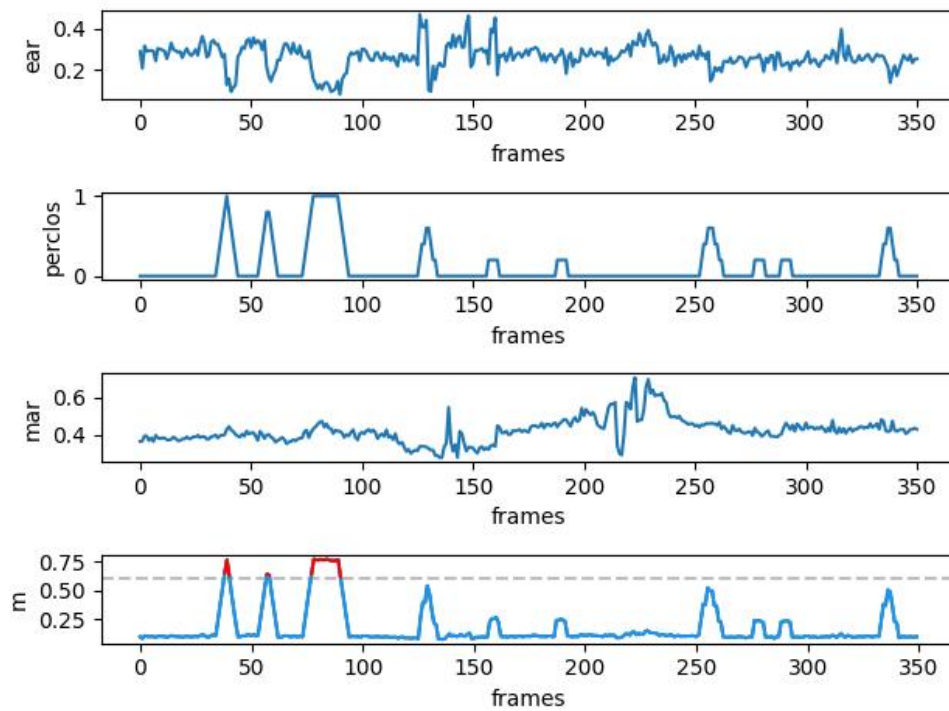


图 5-8 男生未疲劳驾驶状态

Fig.5- 8 Driving without fatigue of men

图 5-7 和图 5-8 中,  $M$  是判断驾驶员面部多特征融合的判断指数, 用于判断驾驶员的疲劳程度, 每个参数根据不同的权重进行融合和计算, 既保证了检测的准确性, 也避免了缺失检测或检测错误, 此外还可以根据  $N$  的大小来判断驾驶员的疲劳状态。

在图 5-8 的  $M$  曲线中疲劳检测系统识别出 3 段疲劳状态曲线和 2 段重度疲劳曲线, 其余部分为清醒状态下的曲线, 分析结果如表 5-1 所示。图 5-7 中  $M$  曲线的部分没有出现  $N$  大于 25 的部分, 则表明没有出现疲劳状态。根据图 5-7 和图 5-8 中  $M$  曲线得出的结果显示, 检测系统没有出现错误检测或缺失检测的情况, 满足实际的检测要求。

通过对图 5-7 和图 5-8 的视频帧数提取分析, 设置驾驶员清醒、疲劳和重度疲劳三种状态,  $N$  为视频的帧数。 $N$  小于等于 25 帧时为清醒状态,  $N$  大于 25 小于等于 50 为疲劳状态,  $N$  大于 50 为重度疲劳状态。如表 5-1 所示中可以看出在视频帧数在  $N$  大于 25 小于等于 50 的有 35~76、123~167 和 199~224, 视频帧数分别为 42、45 和 26 为疲劳状态;  $N$  大于 50 的帧数有 307~394 和 545~602, 视频帧数为 88 和 58 为重度疲劳。

依据短期实验和长期测试对本文提出的疲劳驾驶检测系统进行验证, 为了进一步确认本文提出的检测系统在性别上是否存在差异性, 本文对女生进行同样的测试, 验证疲劳检测系统的稳定性及精确性。

表 5-1 男生驾驶员状态分析  
Table 5-1 Driver status analysis of men

状态	条件	区段	帧数
清醒	$N \leq 25$		
疲劳	$25 < N \leq 50$	35~76	42
疲劳	$25 < N \leq 50$	123~167	45
疲劳	$25 < N \leq 50$	199~224	26
重度疲劳	$N > 50$	307~394	88
重度疲劳	$N > 50$	545~602	58

## 5.5 不同性别对疲劳检测结果的影响

在长期测试的基础上对相同年龄段男生和女生进行同样的测试, 以验证本文提出基于面部特征融合的驾驶员疲劳检测系统的准确性以及实时性。同样在封闭且安全的道路情况下对年龄相同的男生和女生进行测试, 为了确保测试的准确性,



保持测试环境一致性，比如同一辆汽车和同样的驾驶路线。

图 5-7 和图 5-8 是一位年龄在 26 岁左右男生测试的数据；同样找一位年龄在 26 岁左右的女生进行测试，驾驶路线、环境、时长与男生保持一致。通过疲劳检测系统获得女生的驾驶数据如图 5-9 和图 5-10 所示，两张曲线图中的四条线从上往下依次表示为 EAR、PERCLOS、MAR 和 M 的数据曲线图，横坐标为视频的帧数，纵坐标为 EAR、PERCLOS、MAR 和 M 的值。

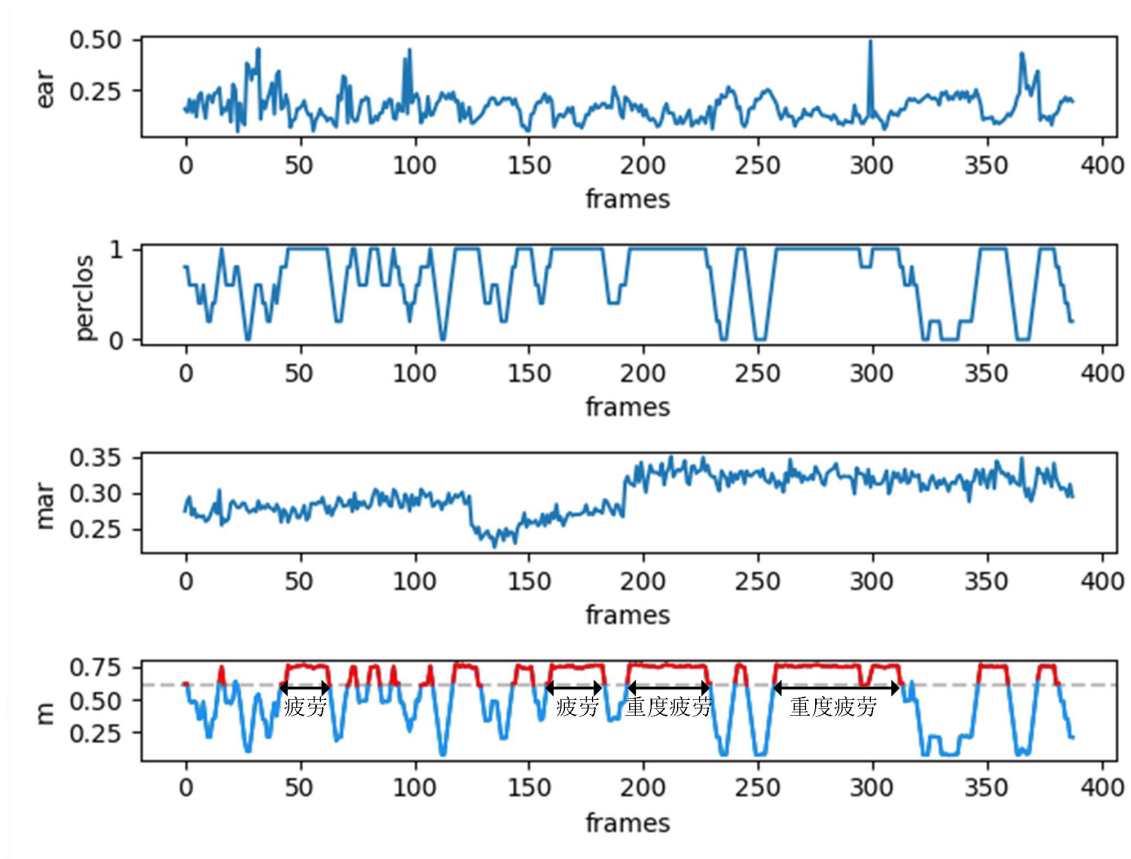


图 5-9 女生疲劳驾驶状态

Fig.5-9 Fatigue driving state of girls

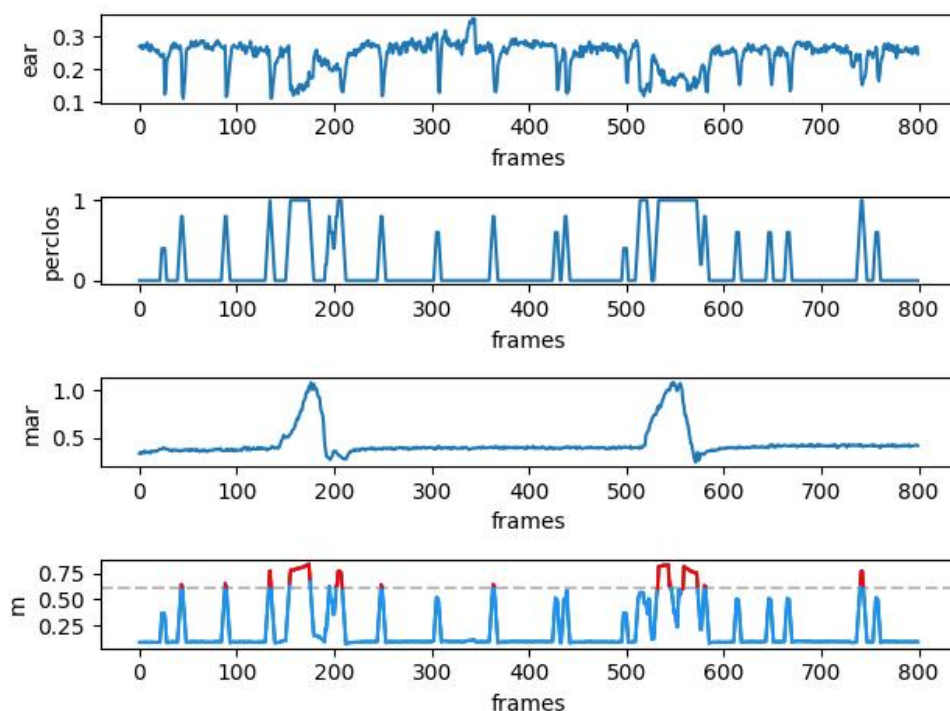


图 5-10 女生未疲劳驾驶状态

Fig.5-10 Driving without fatigue of girls

图 5-9 中的 M 曲线中包含 2 段疲劳状态曲线和 2 段重度疲劳曲线，其余部分为清醒状态下的曲线。在图 5-10 中 M 曲线的部分也没有出现 N 大于 25 的部分，证明未曾出现疲劳状态。图 5-9 中的 M 曲线分析结果如表 5-2 所示，在表 5-2 中可以看出女生驾驶车辆视频帧数在 N 大于 25 小于等于 50 的有 43~68 和 151~179，视频帧数分别为 26 和 29 为疲劳状态；N 大于 50 的帧数有 189~241 和 262~343，视频帧数为 53 和 82 为重度疲劳。

表 5-2 女生驾驶员状态分析  
Table 5-2 Driver status analysis of girls

状态	条件	区段	帧数
清醒	$N \leq 25$		
疲劳	$25 < N \leq 50$	43~68	26
疲劳	$25 < N \leq 50$	151~179	29
重度疲劳	$N > 50$	189~241	53
重度疲劳	$N > 50$	262~343	82

通过图 5-7、图 5-9 和图 5-8、图 5-10 男生和女生分别在疲劳状态和清醒状态下的各个曲线对比结果显示，在同样的驾驶条件下，本系统可以精确识别不同性别的驾驶员疲劳状态。



5.6 驾驶员疲劳验证

为了验证本文提出的基于面部多特征融合的疲劳检测系统的准确性，对 8 名参加测试的测试者分别进行同样的测试，测试条件：48 小时内不能喝酒、吃药、喝咖啡和吃刺激性食物等；测试工况：（1）驾驶模拟器，进行高速公路、绕城高速和城市道路测试；（2）封闭道路测试，进行长直线行驶、连续弯道行驶和绕圈行驶。实时记录 8 名测试者的驾驶状态数据和面部特征，做成相应的样本数据库。

在 8 名测试者的样本数据库随机抽取 480 个样本数据，每名测试者抽取 60 个样本数据。通过主观问卷调查的方式，将样本数据进行标记，分别是清醒状态、疲劳状态和重度疲劳状态。将标记的样本数据输入到疲劳检测系统进行验证，确定本文疲劳检测系统的准确率。

5.6.1 疲劳主观问卷调查验证

在这项研究中，当司机完成测试驾驶时，他们被要求根据回顾视频完成疲劳主观问卷调查，自我报告评估问卷调查的具体内容如表 5-3 所示。8 名测试者根据自己的驾驶视频回顾自己的驾驶状态，在当时的驾驶状态回答问卷调查中的问题，回答问题时是否存在困难：（1）根本不存在困难；（2）有轻微困难；（3）有困难。通过主观问卷调查的得分将驾驶员的状态分为清醒状态、疲劳状态和重度疲劳状态。驾驶员主观问卷调查的得分值划分为三个区域：5-8 分为清醒状态，9-12 分为疲劳状态，13-15 分为重度疲劳状态；8 名测试者问卷调查得分的结果如表 5-4 所示。按照主观问卷调查的得分对每名测试者的样本数据进行标记，标记内容是：清醒、疲劳和重度疲劳；将标记好的样本数据传输到疲劳检测系统进行逐一验证。

表 5-3 对疲劳程度的自我报告进行评估  
Table 5-3 Assess self-report of fatigue level

项目	描述	分值
1	你感到昏昏欲睡吗？	1~3
2	你对清晰的思考有困难吗？	1~3
3	你马上回答一个问题有困难吗？	1~3
4	你对控制自己的肌肉有困难吗？	1~3
5	你对注意力集中有困难吗？	1~3

注：1 表示根本不存在困难；2 表示有轻微困难；3 表示有困难。

表 5-4 问卷调查得分表

Table 5-4 Questionnaire Score Table

测试驾驶员	1	2	3	4	5	6	7	8
清醒	5	5	5	5	6	5	5	5
疲劳	11	12	11	12	12	9	12	10
重度疲劳	13	15	13	15	14	13	15	15

### 5.6.2 疲劳检测系统验证

对疲劳检测系统进行验证，将识别系统的准确率作为评价标准<sup>[62]</sup>，识别准确率的计算公式如 5-1 所示：

$$P_n = \frac{q_n}{Q_n}, n=1,2,3 \quad (5-1)$$

公式中， $n$  表示驾驶员的疲劳程度：1 代表清醒，2 代表疲劳，3 代表重度疲劳； $p_n$  表示模型对第  $n$  种疲劳程度识别的正确率； $q_n$  表示第  $n$  种的疲劳驾驶测试样本数据中被准确识别出的样本个数； $Q_n$  表示第  $n$  种的疲劳驾驶测试样本总数。

疲劳检测系统将对标记的 480 个样本数据进行验证，识别每名测试者清醒、疲劳、重度疲劳状态的准确率，准确率如表 5-5 所示。

表 5-5 疲劳驾驶检测系统对不同驾驶员的识别结果

Table 5-5 Identification Results of Fatigue Driving Detection System for Different Drivers

测试驾驶员		1	2	3	4	5	6	7	8	平均值
清醒	P1	97.2	96.6	95.7	98.3	97.1	93.6	92.8	98.2	96.2
疲劳	P2	92.8	93.4	91.7	94.8	92.6	90.8	91.2	95.8	92.9
重度疲劳	P3	92.2	94.9	92.6	95.2	93.7	90.1	91.8	96.7	93.4

注：单位为百分比。

由表 5-5 中的数据可以得出，本文基于驾驶员面部多特征融合的疲劳检测系统对不同驾驶员的检测的准确率有所不同，8 名测试者清醒状态的平均准确率为 96.2%，疲劳状态的平均准确率为 92.9%，重度疲劳状态的平均准确率为 93.4%。疲劳和重度疲劳的准确率的平均值为疲劳检测系统的准确率。因此基于驾驶员面部多特征融合的疲劳检测系统的准确率为 93.1%，满足实际情况下的疲劳检测系统的要求。

## 5.7 本章小结

本章主要对实验的数据和结果进行详细阐述，验证本文所提疲劳驾驶疲劳检测系统的合理性。本章首先对实验环境和数据集的收集以及对疲劳检测设备的软件系统和硬件系统进行了阐述。然后进行了两次实验，一次是短期实验，另一次

是长期实验。通过主观问卷调查的方式将测试者的驾驶视频标记为清醒、疲劳和重度疲劳，疲劳检测系统对标记的视频进行逐一识别验证。实验结果表明本文采用基于面部多特征融合的疲劳检测系统的准确率达到了 93.1%，满足工业应用驾驶员疲劳实时检测系统的要求。

## 结论

驾驶员疲劳驾驶检测技术的研究减少因疲劳引起交通事故的首要任务，对交通安全以及公共安全具有重要意义。因此，研究驾驶员疲劳驾驶具有重要意义。针对驾驶员长时间驾驶引起的疲劳驾驶问题，本文在现有研究成果的基础上，提出了一种基于驾驶员面部多特征融合的疲劳检测系统。该方法不仅结合 TCDCN 网络对驾驶员面部进行准确识别，而且通过驾驶员面部多个特征的融合对疲劳状态进行检测，实现了对驾驶员疲劳驾驶的预警。本文工作总结如下：

（1）针对国内外对驾驶员疲劳检测技术的相关研究，总结了现今疲劳检测系统的优缺点以及存在的关键技术问题。通过对国内外的研究现状的分析，本文提出了基于驾驶员面部多特征融合的疲劳检测系统。

（2）深入研究了疲劳驾驶系统中驾驶员面部检测的关键技术，对面部检测网络 MTCNN、DCNN 和 TCDCN 进行了深入分析。

（3）为了能够准确且快速的实现对驾驶员疲劳状态的识别和检测的实时性，本文提出了离线训练模块和在线测试模块。离线训练模块通过利用多任务学习的优点，对驾驶员的面部疲劳特征进行训练；在线测试模块是为了能够从摄像头中实时的检测驾驶员的驾驶状态。

（4）对眼睛、嘴巴的特征进行了单独检测分析。通过对眼睛睁眼、闭眼以及眨眼的过程进行检测，得到的 EAR 数据使用 Python 中的 Matplotlib 画出眨眼过程的曲线图，确定 EAR 的阈值为 0.2，EAR 小于 0.2 即为闭眼状态；在 EAR 数据的基础上通过公式计算得到 PERCLOS 的值，画出 PERCLOS 的曲线图，确定了 PERCLOS 的阈值为 0.8，PERCLOS 大于 0.8 连续帧数为闭眼时长。对嘴巴张开、闭合、打哈欠和说话的状态逐一进行了检测分析，通过对比嘴巴在不同状态下的 MAR 曲线图，确定 MAR 的阈值为 0.6，MAR 大于 0.6 为打哈欠状态。为了使疲劳检测的精度更加准确，本文将眼睛、嘴巴的特征进行融合，既是将 EAR、PERCLOS 和 MAR 参数融合成一个新参数 M，利用  $M > 0.605$  的帧数 N 来判断驾驶员的疲劳程度，驾驶员驾驶状态分为清醒状态、疲劳状态、重度疲劳状态三种情况。

（5）为了保证实验数据的可靠性以及检测系统的准确性，搭建了疲劳检测平台。在该平台上分别进行了两种实验验证：短期实验和长期实验。短期测试验证的目的主要是识别驾驶员昏昏欲睡驾驶的典型迹象；长期测试主要验证是所提算法能否准确判断驾驶员在长期驾驶时是否打瞌睡，并且根据疲劳程度（清醒、

疲劳和严重疲劳)进行分类。通过实验表明疲劳检测系统的准确率为 93.1%，达到了目前工业应用检测的准确率，可用于驾驶员疲劳驾驶检测。

本文的创新点如下：

(1) 本文利用了驾驶员面部特征(眼睛、嘴巴等)来融合视觉信息，有效的提高了驾驶员疲劳检测的准确性。

(2) 通过 TCDCN 网络对人脸的关键点进行定位，提取出眼睛和嘴巴的区域，从而计算的 EAR、PERCLOS 和 MAR 是一个更简洁的解决方案。

(3) 本文提出新的参数 M 和 N。M 值为 EAR、PERCLOS 和 MAR 的加权值，即眼睛和嘴巴综合状态；N 为判断驾驶员疲劳状态的参数值，依据 N 值的累加次数判断驾驶员是否疲劳。

本文结合了深度学习、机器视觉、多特征融合等技术，提出了基于驾驶员面部多特征融合的疲劳检测方法，同时将其与现有的方案进行了对比，实验结果显示本文检测系统表现出良好的技术优势。但本文提出的方案还是有许多欠缺的地方，需要进一步改进，在未来的研究工作中，将关注以下几个方面：

(1) 上述研究将安装在车辆上，进一步验证夜间驾驶条件的疲劳检测的识别效果。

(2) 添加其他视觉特征来融合信息，如驾驶员头部姿势。

## 参考文献

- [1] Amodio A, Ermidoro M, Maggi D, et al. Automatic detection of driver Impairment based on pupillary light reflex[J]. IEEE Transactions on Intelligent Transportation Systems, 2019, 20(8): 3038-3048.
- [2] Li Y Y, Zhang G N. The effect of fatigue driving on injury severity considering the endogeneity [J]. Journal of Safety Research, 2018, 64: 11-19.
- [3] Zhang G N, Yan K K W, Zhang X, et al. Traffic accidents involving fatigue driving and their extent of casualties[J]. Accident Analysis and Prevention, 2016, 87(2): 34-42.
- [4] Ganderp H, Marshall L N S, James I, et al. Investigating driver fatigue based on vehicle operating data[J]. Journal of Transportation Systems Engineering and Information Technology, 2020, 20(4): 77-82.
- [5] Higgins J S, Michael J, Austin R, et al. Asleep at the wheel—The road to addressing drowsy driving[J]. Journal of Sleep, 2017, 40(2): 1-9.
- [6] Pei Z. Study Situation and Developing Trend on Detecting and Evaluating Techniques of Motor Driver Fatigue[J]. Journal of China Agricultural University, 2001, 6(6): 01–105.
- [7] Liang Y, Reyes M L, Lee J D. Real-Time Detection of Driver Cognitive Distraction Using Support Vector Machines[J]. IEEE Transactions on Intelligent Transportation Systems, 2007, 8(2): 340–350.
- [8] Liu C C, Hosking S G, Lenne M G. Predicting driver drowsiness using vehicle measures: Recent insights and future challenges[J]. Journal of Safety Research, 2009, 40(4): 239-245.
- [9] Forsman P M, Vila B J, Short R A, et al. Efficient driver drowsiness detection at moderate levels of drowsiness[J]. Accident Analysis and Prevention, 2013, 50: 341–350.
- [10] Wang J, Zhu S, Gong Y. Driving Safety Monitoring Using Semi-supervised Learning on Time Series Data[J]. IEEE Transactions on Intelligent Transportation Systems, 2010, 11(3): 728-737.
- [11] Wu, B F, Chen, et al. Reasoning-Based Framework for Driving Safety Monitoring Using Driving Event Recognition[J]. IEEE Transactions on Intelligent Transportation Systems, 2013, 14(3): 1231-1241.
- [12] 郭建生,陈士伟,白菁等.军事飞行人员疲劳状况调查与分析(英文)[J].解放军医学杂志, 2012, 37(03):249-252.
- [13] 李晷,詹皓,郭华等.空军飞行人员飞行疲劳状况调查与分析[J].人民军医, 2014, 57(11): 1167-1169.
- [14] 万斌,贾爱芳.矿工心理疲劳状况调查分析[J].能源技术与管理, 2022, 47(05): 52-55.
- [15] 田春雨,邵维阳,崔蓓等.海军舰员远航条件下视疲劳状况调查分析[J].国际眼科杂志, 2021, 21(09): 1657-1660.

- [16] Lee S, Kim J K. Factors contributing to the risk of airline pilot fatigue[J]. Journal of air transport management, 2018, 67: 197-207.
- [17] Filtness A J, Naweed A. Causes, consequences and countermeasures to driver fatigue in the rail industry: The train driver perspective[J]. Applied ergonomics, 2017, 60: 12-21.
- [18] Mvmy A, Xla B, Ks B, et al. Can SVM be used for automatic EEG detection of drowsiness during car driving[J]. Safety Science, 2009, 47(1): 115-124.
- [19] Lin C T, Chang C J, Lin B S, et al. A Real-Time Wireless Brain-Computer Interface System for Drowsiness Detection[J]. IEEE Transactions on Biomedical Circuits & Systems, 2010, 4(4): 214-222.
- [20] Lin F C, Ko L W, Chuang C H, et al. Generalized EEG-Based Drowsiness Prediction System by Using a Self-Organizing Neural Fuzzy System[J]. IEEE Transactions on Circuits & Systems I Regular Papers, 2012, 59(9): 2044-2055.
- [21] Dong N, Li Y, Gao Z, et al. A WPCA-Based Method for Detecting Fatigue Driving From EEG-Based Internet of Vehicles System[J]. IEEE Transactions on Intelligent Transportation Systems, 2019, 7:124702–124711.
- [22] Jung S J, Shin H S, Chung W Y. Driver fatigue and drowsiness monitoring system with embedded electrocardiogram sensor on steering wheel[J]. Iet Intelligent Transport Systems, 2014, 8(1): 43-50.
- [23] Lei J, Liu F, Han Q, et al. Study on Driving Fatigue Evaluation System Based on Short Time Period ECG Signal[C]// 2018 IEEE International Conference on Intelligent Transportation Systems (ITSC). IEEE, 2018: 2466-2470.
- [24] Zhang Y F, Gao X Y, Zhu J Y, et al. A novel approach to driving fatigue detection using forehead EOG[J]. International IEEE/EMBS Conference on Neural Engineering, 2015, 7:707-710.
- [25] Bulling A, Ward J A, Gellersen H, et al. Eye Movement Analysis for Activity Recognition Using Electrooculography[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2011, 33(4): 741-753.
- [26] Hu S, Zheng G. Driver drowsiness detection with eyelid related parameters by Support Vector Machine[J]. Expert Systems with Applications, 2009, 36(4): 7651-7658.
- [27] Koichi, Fujiwara, Erika, et al. Heart Rate Variability-based Driver Drowsiness Detection and its Validation with EEG[J]. IEEE transactions on bio-medical engineering, 2018, 66(6): 1769-1778.
- [28] Sangtae A, Thien N, Hyojung J, et al. Exploring Neuro-Physiological Correlates of Drivers' Mental Fatigue Caused by Sleep Deprivation Using Simultaneous EEG, ECG, and fNIRS Data[J]. Frontiers in Human Neuroscience, 2016, 10(848): 1-14.
- [29] Wang H, Wu C, Li T, et al. Driving Fatigue Classification Based on Fusion Entropy Analysis Combining EOG and EEG[J]. IEEE Access, 2019, 7: 61975–61986.
- [30] 王磊.面向脑疲劳检测的 EEG 信号通道选择与多特征融合研究[D]. 武汉: 武汉理工大学, 2019.

020.

- [31] 杨硕,李润泽,丁建清,徐桂芝.基于 EMD 去趋势波动的脑疲劳模糊熵分析[J].中国生物医学工程学报, 2020, 39(01): 33-39.
- [32] 张春翠.体疲劳对脑疲劳影响的脑电信息分析与处理[D]. 天津: 天津大学, 2014.
- [33] 李增勇,焦昆,陈铭,王成焘.汽车驾驶员驾驶过程中的心率变异性功率谱分析[J].中国生物医学工程学报, 2003(06): 574-576.
- [34] 王琳,张陈,尹晓伟,付荣荣,王宏.一种基于驾驶员生理信号的非接触式驾驶疲劳检测技术[J].汽车工程, 2018, 40(03): 333-341.
- [35] 王福旺,王宏,罗旭.基于 EEG 与 EOG 信号的疲劳驾驶状态综合分析[J].东北大学学报(自然科学版), 2014, 35(02): 175-178.
- [36] Riera L, Ozcan K, Merickel J. Detecting and tracking unsafe lane departure events for predicting driver safety in challenging naturalistic driving data[C]//2020 IEEE Intelligent Vehicles Symposium (IV). IEEE, 2020: 238-245.
- [37] Takei Y, Furukawa Y. Estimate of driver's fatigue through steering motion[C]// Systems, Man and Cybernetics, 2005 IEEE International Conference on. IEEE, 2005(2):1765-1770.
- [38] Sandberg D, Wahde M. Particle swarm optimization of feedforward neural networks for the detection of drowsy driving[C]// IEEE International Joint Conference on Neural Networks. IEEE, 2008:788-793.
- [39] 蔡素贤,杜超坎,周思毅,王雅斐.基于车辆运行数据的疲劳驾驶状态检测[J]. 交通运输系统工程与信息, 2020, 20(04): 77-82.
- [40] 张驰.高危车辆驾驶员疲劳驾驶监测及预警系统研究[D]. 锦州: 辽宁工业大学, 2018.
- [41] 屈肖蕾.基于转向操作和车辆状态的疲劳驾驶检测方法研究[D]. 北京: 清华大学, 2012.
- [42] 黄皓.基于驾驶操作及车辆状态的疲劳驾驶行为检测研究[D]. 南京: 东南大学, 2016.
- [43] Xuepeng Z, Chunning M, Mingkui F, et al. Eye Feature Point Detection Based on Single Convolutional Neural Network[J]. Iet Computer Vision, 2017, 12(4): 453-457.
- [44] Fang Z, Su J, Lei G, et al. Driver Fatigue Detection Based on Eye State Recognition[C]//2017 International Conference on Machine Vision and Information Technology (CMVIT). IEEE, 2017: 105-110.
- [45] Liu, Z, Peng, Y, Hu W. Driver Fatigue Detection based on Deeply-Learned Facial Expression Representation[J]. Journal of Visual Communication and Image Representation, 2020, 71(6): 12.
- [46] 廖明明,赵波.基于面部特征融合的驾驶员疲劳检测[J]. 智能计算机与应用, 2021, 11 (10): 77-81.
- [47] 苏锦瑾,霍春宝,王特特.基于面部特征的驾驶员疲劳检测系统[J]. 信息技术与信息化, 2021 (09): 114-116.
- [48] 刘炜煌,钱锦浩,姚增伟,焦新涛,潘家辉.基于多面部特征融合的驾驶员疲劳检测算法[J]. 计算机系统应用, 2018, 27(10): 177-182.



- [49] 周云鹏,朱青,王耀南,卢笑,凌志刚.面部多特征融合的驾驶员疲劳检测方法[J]. 电子测量与仪器学报, 2014, 28(10): 1140-1148.
- [50] Freund Y. A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting[C]. J. Comput. Syst. Sci. 1997, 55, 119–139.
- [51] Yi S, Wang X, Tang X. Deep Convolutional Network Cascade for Facial Point Detection [C]// Computer Vision and Pattern Recognition (CVPR), 2013 IEEE Conference on. IEEE, 2013: 3476-3483.
- [52] Zhang Z, Luo P, Loy. Facial Landmark Detection by Deep Multi-task Learning[J]. Lecture Notes in Computer Science, 2014, 8694: 94-108.
- [53] Lin M, Chen Q, Yan S. Network in Network[J]. arXiv e-prints, 2013.
- [54] Freund Y, Schapire R E. A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting[C]. Conference on Learning Theory. Academic Press, 1997, 55:119–139.
- [55] Cheng R, Zhao Y, Dai Y, et al. An On-board Embedded Driver Fatigue Warning System Based on Adaboost Method[J]. Beijing Daxue Xuebao Ziran Kexue Ban Acta Scientiarum Naturalium Universitatis Pekinensis, 2012, 48(5): 719-726.
- [56] Min J, Xiong C, Zhang Y, et al. Driver fatigue detection based on prefrontal EEG using multi-entropy measures and hybrid model[J]. Biomedical Signal Processing and Control, 2021, 69(3):102857–102865.
- [57] Shi-Ru Q U, Peng J C. Design of multi-feature fusion driver fatigue state detection system based on FPGA[J]. Transducer and Microsystem Technologies, 2013, 32(5): 86-105.
- [58] King D E. Dlib-ml: A Machine Learning Toolkit[J]. Journal of Machine Learning Research, 2009, 10(3): 1755-1758.
- [59] B Range, K West, N Air, et al. U.S. Department of Transportation, Federal Highway Administration[J]. Seismic Retrofitting Manual for Highway Bridges, 1995, 15(5): 766-770.
- [60] Heng Q, Wang W, Jiang X, et al. Assessment of Driver Mental Fatigue Using Facial Landmarks [J]. IEEE Access, 2019, 7: 150423–150434.
- [61] Nonis F, Dagnes N, Marcolin F, et al. 3D Approaches and Challenges in Facial Expression Recognition Algorithms—A Literature Review[J]. Journal of Applied Sciences, 2019, 9:3904.
- [62] 郭孜政,谭永刚,马国忠,等.基于 BP 神经网络的驾驶精神疲劳识别方法[J]. 哈尔滨工业大学学报, 2014, 46(08): 118-121.