

# 基于神经网络的疲劳驾驶检测方法研究

张育榕, 谷 昆, 张轩雄

上海理工大学, 光电信息与计算机工程学院, 上海

收稿日期: 2023年4月14日; 录用日期: 2023年5月18日; 发布日期: 2023年5月25日

## 摘 要

由于疲劳驾驶是导致交通事故的重要原因。因此, 针对疲劳驾驶检测, 本文提出一种基于神经网络的疲劳检测方法。对于传统疲劳检测技术检测精度不高、综合检测能力不足等问题提出了相应的改进策略, 这对预防疲劳驾驶、提高驾驶安全具有重要的意义。本文的主要工作如下: 1) 采用基于Retina Face算法的人脸检测网络, 同时本文收集了大量局部遮挡和大姿态偏转的人脸图像作为训练样本, 使得人脸检测模型更加灵活, 解决了传统人脸检测算法对于人脸姿态要求过严的缺点。2) 本文从驾驶员眼部、嘴部以及头部姿态出发, 基于神经网络的方法, 综合分析驾驶员的疲劳状态。当系统存在某一指标失效时, 依然可以正常检测出驾驶员的疲劳状态, 具有一定的容错性。此外, 借鉴PERCLOS经验准则, 提出了基于一定时间周期的疲劳判定法, 使系统对于驾驶员的疲劳预估更加准确。实验结果表明, 使用本文提出的方法, 能够实现多特征点的定位与识别, 且基于眼部状态以及嘴部状态预测疲劳的精度分别达到了97.1%和97.5%, 根据头部状态预测疲劳的精度达到了88.1%, 处理速度也达到了每张图片62 ms。相较于传统只基于眼部状态检测疲劳的方法, 本文提出的算法在保证准确率和效率的前提下, 提高了综合检测能力。

## 关键词

交通事故, 疲劳驾驶检测, 人脸检测, PERCLOS, 眼部状态, 嘴部状态, 头部姿态, 神经网络

## Research on Fatigue Driving Detection Method Based on Neural Network

Yurong Zhang, Kun Gu, Xuanxiong Zhang

School of Optical-Electrical and Computer Engineering, University of Shanghai for Science and Technology, Shanghai

Received: Apr. 14<sup>th</sup>, 2023; accepted: May 18<sup>th</sup>, 2023; published: May 25<sup>th</sup>, 2023

## Abstract

Fatigue driving is an important cause of traffic accidents. Therefore, for fatigue driving detection,

this paper puts forward a fatigue detection technology based on deep learning, and puts forward corresponding improvement strategies for the problems of low detection accuracy and poor real-time performance of traditional fatigue detection technology, which is of great significance to prevent fatigue driving and improve driving safety. The main work of this paper is as follows: 1) The face detection network based on Retina Face algorithm is adopted, and a large number of face images with partial occlusion and large posture deflection are collected as training samples, which makes the face detection model more flexible and solves the shortcoming that the traditional face detection algorithm requires too strict face posture. 2) Based on the attitude of the driver's eyes, mouth and head, this paper comprehensively analyzes the driver's fatigue state based on the method of neural network. When a certain index of the system fails, the fatigue state of the driver can still be detected normally, which has a high enough fault tolerance rate. In addition, a fatigue judgment method based on a certain time period is proposed by referring to PERCLOS, which makes the system more accurate for driver fatigue prediction. The experimental results show that the method proposed in this paper can realize the location and recognition of multi-feature points, and the accuracy of fatigue prediction base on the eye state and mouth state is 97.1% and 97.5% respectively. The accuracy of fatigue prediction base on the head state is 88.1%, and the processing speed of each image is 62 ms. Compared with the traditional fatigue detection methods of only detecting eye state, the algorithm proposed in this paper improves the comprehensive detection capability, based on sufficient accuracy and efficiency.

## Keywords

Traffic Accidents, Drowsy Driving Detection, Face Detection, PERCLOS, Eye State, Mouth State, Head Posture, Neural Networks

Copyright © 2023 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 引言

疲劳驾驶检测一直都是国内外学者研究的热点,近年来,随着深度学习、计算机视觉的迅猛发展,一些国内外的研究学者将计算机视觉以及深度学习引入到疲劳驾驶领域,取得了巨大的发展。从疲劳驾驶的检测方式来看,主要分为以下三大类:基于车辆运动特征的检测方法、基于驾驶员生理特征的检测方法以及基于驾驶员面部表情的检测方法。

基于车辆运动特征的方法主要通过安装在车上安装各种传感器来采集车辆在行驶过程中的各项数据,进而判断驾驶员是否疲劳驾驶。Sayed [1]等人根据采集的方向盘角度信息建立了基于人工神经网络的疲劳驾驶检测模型。柴萌[2]根据车辆的方向盘参数、转向速度以及驾驶员的反应参数建立了高斯混合隐马尔夫模型。陈志勇[3]等人提取出了车辆的速度、加速度以及方向盘的角速度等信息,最后送入BP神经网络训练出了疲劳检测模型。

基于生理特征的方法是通过采集驾驶人员的生理信息,如脑电波、心电波、脉搏等,然后将这些数据进行降噪处理,从中提取出疲劳特征信息,用以研究生理特征和疲劳度的关系。王斐[4]等人通过采集驾驶员的脑电波以及方向盘信息来分析疲劳程度。Chuang [5]等人提出了一种感知功能整合系统,徐礼胜[6]等人通过对驾驶员的心电信号进行特征提取,通过随机森林算法对特征进行分类进而判断驾驶员的疲劳状态。

基于面部特征的方法需要在驾驶车内部安装相机,采集车辆在行驶过程中车舱内的图像,结合计算

机视觉相关算法对图像进行实时分析。Kim [7]等人提出了基线模型检测技术。李文学[8]等人根据 Harr [9]特征人脸检测算法定位人脸区域,并结合积分投影、分线性中来预测驾驶员的疲劳状态。潘志庚[10]等提出使用 Adaboost 算法[11]来定位人脸,运算方法取出眉毛部分,最后通过眼睛纵横比、瞳孔面积占比等信息综合判断出驾驶员的疲劳程度。

综合国内外现状,基于人脸特征的疲劳驾驶检测技术方法检测精度较高,相对于其他两种,安装价格也较为便宜,更容易被大众接受。目前,许多已有的疲劳检测系统仅仅依靠眼睛状态或嘴巴状态来判断驾驶员是否疲劳,当系统因为某些原因检测不到眼睛或者嘴巴等信息时,系统就无法正常工作。因此,本文对基于面部特征的疲劳检测不够全面的不足之处展开研究,提出了不局限于单一状态的检测方法,而是基于眼睛状态、嘴巴状态和头部姿态相结合的疲劳检测方法,当系统实际运行时,某些指标比如眼睛状态检测不到时,还可以根据其他特征例如头部以及嘴部状态来检测驾驶员的当前状态,具有综合性。此外,我们在疲劳度分析模块中引入 PERCLOS 准则,即疲劳状态评判标准,使得系统在检测时更加可靠。

本文结构安排如下:

第一章介绍了疲劳驾驶检测方法的整体流程框架;第二章介绍了人脸检测和脸部特征点提取的具体方法;第三章说明了疲劳分析方法的设计思路与具体实现;第四节给出了实验结果与分析;

#### 1) 整体框架

本文所讨论的疲劳检测系统对于驾驶员疲劳分析主要是从三个方面入手,分别是:眼睛状态、嘴巴状态、头部姿态,从多方面检测来判断驾驶员是否疲劳,使得系统能够全面和高效地服务于驾驶人员。系统主要由人脸采集、预处理、人脸定位、特征提取、眼睛状态检测、嘴巴状态检测、头部姿态检测、疲劳估计和报警这几个模块组成,系统的整体框架图如下图 1 所示。

从下面的框架图可以看出本文研究的疲劳驾驶检测方法的具体工作流程。下面按照系统的工作流程简单介绍一下各个模块的大致工作。

1) 相机采集人脸:由于本文研究的方法是通过分析人脸的特征来判断驾驶员是否疲劳,因此采集人脸图是第一步;

2) 图像预处理:由相机采集到的图片难免会出现伴有噪声、细节模糊不清等情况,因此在对人脸图做具体分析之前,我们需要将图像进行预处理,使其变成高质量图;

3) 人脸定位:人脸检测作为人脸属性分析的第一步,精准定位人脸的位置可以使整个系统对驾驶员疲劳度分析更加准确和高效;

4) 特征提取:通过特征提取,我们可以快速定位到人脸中关键部位,如眼睛,嘴巴。这些特征主要用于后面各个部位的状态识别网络中;

5) 眼睛定位:本文对于眼睛状态的检测是通过神经网络模型来进行的,因此我们需要裁减出人脸中的眼部区域,送入眼部状态检测模型中进行预测。

6) 眼部状态检测:由于眼睛是判断人是否疲劳的最直观因素,因此我们对于眼睛状态的判断单独设计了分类检测模型,使得系统对于眼睛的检测更加准确;

7) 嘴部状态检测:对于嘴部状态的检测我们选择基于宽高比的计算方法,由于嘴部关键点相对于眼部关键点较分散,因此使用关键点坐标计算的方式高效且精确;

8) 头部姿态检测:同嘴部姿态一样,我们从特征提取模块获得的关键点中选取一定数量的关键点来计算头部的下倾角,从而判断当前驾驶员的视线是否正常;

9) 疲劳分析:在获得了眼睛、嘴巴、头部状态之后,我们需要对这些参数进行融合分析,并通过经验准则给出驾驶员当前是否疲劳的准确估计;

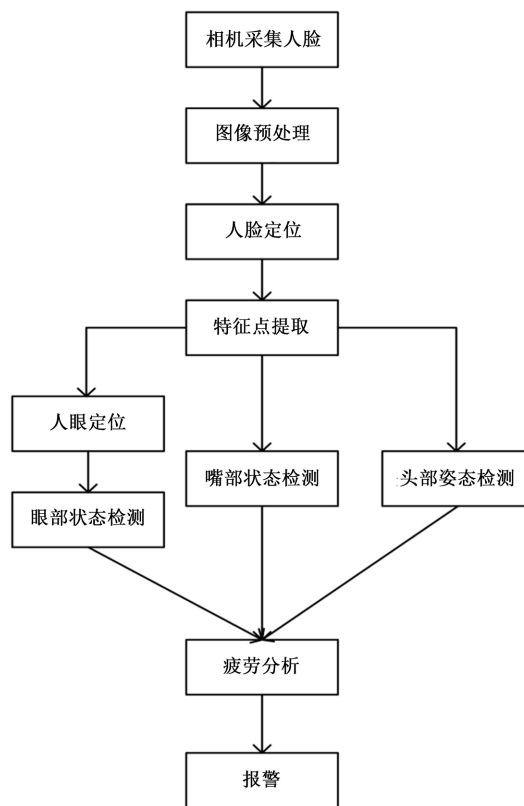


Figure 1. System framework  
图 1. 整体框架

10) 报警：如果疲劳检测系统检测出驾驶员当前处于疲劳状态，需要通过报警模块发出报警信息及干预驾驶员当前的驾驶行为，使得驾驶员快速纠正到正常驾驶状态，从而避免意外事故发生。

## 2. 人脸检测与特征点提取

### 2.1. 人脸检测算法

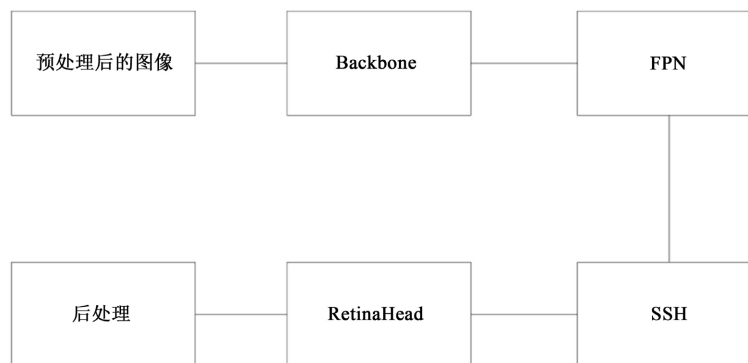
#### 2.1.1. 人脸数据收集

本文所采用的人脸数据集主要包含 WilderFace 数据集和自标注自拍数据集两部分。首先 RetinaFace 原论文作者已经在 WilderFace 数据集上标注了自己的 landmark 关键点信息因此我们可以直接将 WilderFace 拿过来训练。其次，我们仅靠 WilderFace 训练人脸检测网络还不够，因为本文的人脸检测网络要使用于机动车驾驶场景下，为了不对驾驶人员的驾驶姿势做太大的限制，必须在人脸检测训练集中加入一部分特殊的图像数据。驾驶员在驾驶的过程中，经常会有转头观望或者脸部局部遮挡等现象，如果因此这些原因检测不到人脸，那么训练出的人脸检测模型在实际应用中的体验会很差。因此，需要在训练集中加入随机遮挡的人脸图像和头部转动的图像。对于随机遮挡的数据图，主要包含手部遮挡以及其他的一些局部遮挡情况；对于头部大角度转动，主要包含头部左右转动以及低头的人脸图。其中低头的人脸图是我们判断驾驶员是否睡着的一个关键因素之一，如果我们因此低头检测不到人脸，就会导致疲劳检测系统误判而不会发出报警，从而导致意想不到的事情发生。

#### 2.1.2. 基于 Retina Face 的人脸检测

本文使用 Insight Face 在 2019 年提出的人脸检测算法框架 Retina Face [12], 该算法使用基于 one-stage

人脸检测网络同时实现人脸定位和人脸对齐, 相比于过去基于 tow-stage 人脸检测方法, Retina Face 最大的优点就是其检测的速度不会受到图像中人脸数量的影响。此外, 作者提出了一种自编码的人脸编码器用于人脸检测, 如果我们在主干特征提网络中选用 resnet, 我们就可以使整个模型拥有相当高的检测精度; 如果把 resnet 换成具有深度分离卷积结构的 mobileNet, 就可以使系统的检测速度达到实际驾驶场景的需求。



**Figure 2.** The specific process of face detection network  
**图 2.** 人脸检测网络具体流程

本文所使用的 Retina Face 网络主要包括四个部分, 分别是特征提取网络(Backbone)、特征金字塔(FPN)、上下文建模(SSH)、多任务监督(Retina Head), 整个算法的流程如图 2 所示。

## 2.2. 基于回归树的人脸特征点定位

回归分析法是一种比较经典的分析预测方法, 能够很快地建立出输入与输出之间的映射关系。将回归技术应用人脸特征定位任务中, 可以使得模型结构更加直观, 输入输出关系也更为明确, 不需要像统计类方法那样进行大量推断演化。在基于回归的特征点定位算法当中比较经典的算法有: 局部二值化(LBF [13])、回归树联合(ERT [14])、显示形状回归(ESR)等。其中 ESR [15]算法作为一种改进的级联形状回归方法, 它在形状索引特征的基础上, 采用两个级联 boost 回归层实现人脸特征点定位。其算法基本流程为: 首先给出初始的人脸形状, 然后计算出预估形状与真实形状之间的差值, 最后通过回归方法一步步调整预估的人脸形状, 使其越来越接近于真实人脸形状。ESR 算法在特征点附近选取像素点, 并将像素点的差值作为形状索引特征 ESR 在两个特征点附近选取两个点, 然后将这两个像素点的差值作为形状索引特征进行回归分析。由于只需要进行差值运算, LBF 算法是对 ESR 算法的一种改进, 它通过使用全局线性回归和随机森林相结合的方式实现人脸特征点定位。ESR 算法是在两个特征点附近选取两个像素点并将其差值作为形状索引特征, 同时采用随机森林的结构作为预测单元, 在其叶子节点中存储对应的形变作为预测输出。与 ESR 不同的是, LBF 只在当前特征点附近区域随机选择两个点并将其差值作为形状索引特征, 同时采用随机树作为预测单元, 并将随机森林叶子节点的输出组合成一种特征即局部二值特征。ERT 算法同样采用回归器的级联进行特征点定位, 不同的是, ERT 方法将回归树修改为梯度提升决策树(GBDT), 使得每级回归器学习的都是当前人脸形状与真实人脸形状的差值。每个回归器在级联中预测一个人脸形状的更新向量, 人脸图像的初始形状经过一级一级回归器更新之后形成最终模型预测的人脸形状。

本文经过分析以上特征点定位方法的特点后, 决定使用 ERT 的方法来进行疲劳驾驶检测系统的人脸特征点定位。该算法不仅对缺少标注信息的人脸图具有很好的处理能力, 还具有实习性好、检测精度高

的特点，可以在毫米级的时间内精准定位出人脸的 68 个关键点，有利于疲劳检测系统后期任务的进行。

### 3. 疲劳分析模块设计

#### 3.1. 基于卷积神经网络的眼部状态分析

##### 3.1.1. 模型搭建

Table 1. Constitutional neural network with eyes closed

表 1. 睁闭眼卷积神经网络

层数	层级名称	卷积核大小	步长	输出大小
第一层	深度卷积	(3, 3)	(1, 1)	(64, 64, 3)
	卷积	(1, 1)*32	(1, 1)	(64, 64, 32)
	池化	(2, 2)	(2, 2)	(32, 32, 32)
	深度卷积	(3, 3)	(1, 1)	(32, 32, 32)
	卷积	(1, 1)*64	(1, 1)	(32, 32, 64)
	第二层	深度卷积	(3, 3)	(1, 1)
卷积		(1, 1)*64	(1, 1)	(32, 32, 64)
池化		(2, 2)	(2, 2)	(16, 16, 64)
深度卷积		(3, 3)	(1, 1)	(16, 16, 64)
卷积		(1, 1)*128	(1, 1)	(16, 16, 128)
深度卷积		(3, 3)	(1, 1)	(16, 16, 128)
卷积		(1, 1)*128	(1, 1)	(16, 16, 128)
第三层		深度卷积	(3, 3)	(1, 1)
	卷积	(1, 1)*128	(1, 1)	(16, 16, 128)
	池化	(2, 2)	(2, 2)	(8, 8, 128)
第四层	全连接			2048
	全连接			2

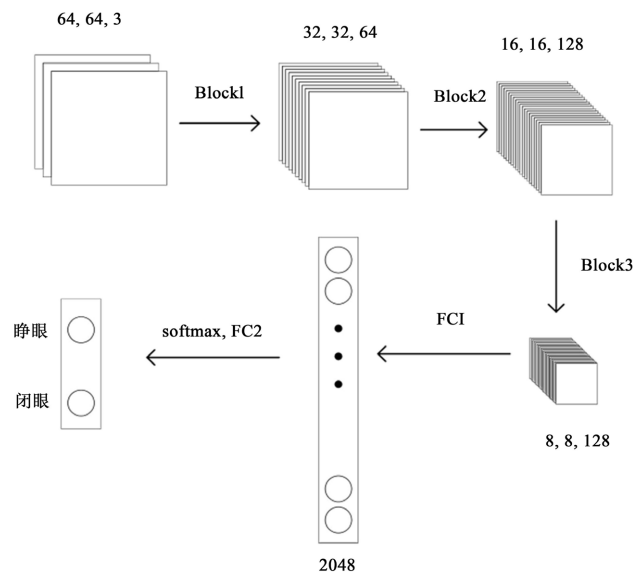


Figure 3. Illustration of the closed-eye network model

图 3. 睁闭眼网络模型图解

基于卷积神经网络的方法主要是利用眼睛数据作为训练集，输入到搭建好的神经网络中进行训练。卷积神经网络通过大量样本自我学习数据的特征，经过多次训练后，就可以得到一个检测精度高、检测速度快且鲁棒性好的模型。本文针对该分类任务提出了一种基于深度可分离[16]卷积神经网络的二分类模型，并且网络的整体搭建采用简单卷积下采样堆叠，最后通过全连接层将图像特征抽象成两个概率数值用以表示人眼是否闭合。整个网络结构如表 1 以及如图 3 所示。

### 3.1.2. 疲劳分析

眼部疲劳分析可以根据特征提取网络输出的关键点信息定位出眼睛位置，我们通过一定的方法来检测当前帧的人脸图像眼部是否闭合，最后我们可以通过 PERCLOS [17]方法来判断驾驶员是处于眨眼还是睡着状态。下面我们分别介绍两种判断眼睛闭合的方法，以及 PERCLOS 算法的基本原理。

#### 1) 眼睛纵横比

基于人眼纵横比的判断方法主要是根据眼睛的 6 个关键点来判断眼睛是否闭合。如下图 4 所示，我们首先计算 EAR 值，然后根据设定的阈值来判断眼睛是否处于闭合状态。对于闭合阈值的选择我们一般通过实验取得，我们可以采集 10 到 20 个人的眼睛从睁开到闭合的过程，记录人眼刚闭合时候的 EAR 值，最后将这些 EAR 值求平均作为判断人眼闭合的阈值。该方法虽然可以快速计算出人眼睁开状态，但其容易受到人眼和摄像头距离的影响，当人眼离摄像头较远时，会由于眼部关键点密集而导致检测精度下降。

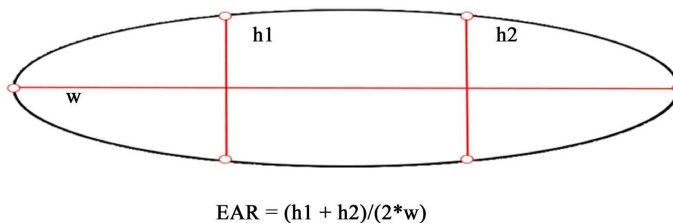


Figure 4. Definition of eye aspect ratio  
图 4. 眼睛宽高比定义图

#### 2) 二分类卷积网络模型

为了避免人眼状态检测受到眼睛和相机距离的影响，我们使用二分类卷积网络模型来判断眼睛是否闭合。该方法只要在训练数据充分的情况下可以很好预测人眼状态并且具有较高鲁棒性。对于训练样本的选取，为了让人眼图像的背景面积最小化，我们可以根据人眼周围关键点取出左右眼图作为训练数据。这样不仅在模型前向推理的过程中可以节省大量计算还可以更容易训练高精度模型。因此，使用这种二分类卷积网络来检测眼睛的状态更加适合本文所设计的疲劳驾驶检测系统。

#### 3) PERCLOS 算法

PERCLOS 算法在 1999 年提出，PERCLOS 是指在一定的时间周期内，眼睛处于闭合状态的时间所占的比例。PERCLOS 作为公认的疲劳状态评判标准，被广泛应用于疲劳驾驶检测技术中。PERCLOS 通过分析人眼开闭状况和疲劳程度之间的关系，得出了根据闭眼时间比例就可以快速准确地判断出驾驶员疲劳度的结论。PERCLOS 有 P70、P80 以及 EM 三个标准，三种标准分别表示眼睑遮住瞳孔面积的 70%、80% 和 50% 时为人眼处于闭眼状态，然后据此求出单位时间内的人眼闭合率。数据显示 PERCLOS 算法在 P80 标准下对疲劳驾驶的检测效果最有效。本论文选择基于 P80 标准的 PERCLOS 算法来判断驾驶员的疲劳度，P80 标准定义：如果眼睑遮住瞳孔的面积在 80% 以上我们认为人眼处于闭合状态，并据此统计人眼闭合时间占规定时间的百分比。如下图 5 为眨眼过程示意图：

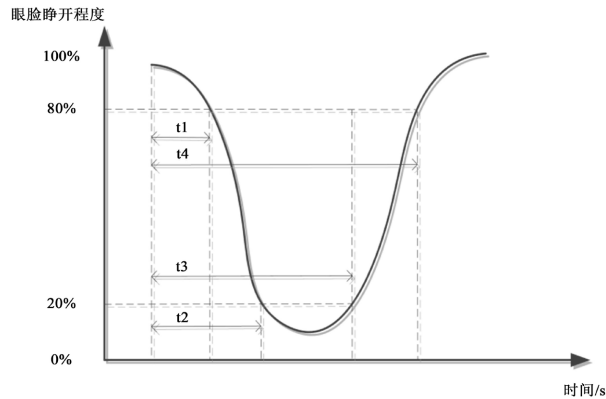


Figure 5. Schematic diagram of P80 blinking process  
图 5. P80 眨眼过程示意图

图 5 中横轴表示时间，纵轴表示眼睑睁开程度。图中  $t_1$  到  $t_4$  表示眼睛闭合过程中不同阶段所用的时间，其中  $t_3 - t_2$  表示眼睑睁开度小于 20% 的时间， $t_4 - t_1$  表示眼睑睁开度小于 80% 的时间。因此 P80 的计算公式为：

$$P80 = \frac{t_3 - t_2}{t_4 - t_1} \times 100\%$$

由上式可以看出，当前 P80 参数越大的时候，疲劳程度越高，反之清醒程度越高。因此我们使用 P80 指标来实时监测驾驶员的疲劳状况。

通过上文对于睁闭眼检测网络介绍，我们可以精确、快速地完成睁闭眼的判断，但是单凭某帧图像的眼部状态还不足以判定驾驶员是否疲劳。我们在概要设计阶段介绍了 PERCLOS 算法，通过计算单位时间内闭眼时间占比来确定驾驶员是否处于疲劳状态。然而在实际的编码过程中，我们将时间概念替换成帧，我们通过统计指定数量帧[18]内的闭眼帧所占比例来近似 PERCLOS 指标。

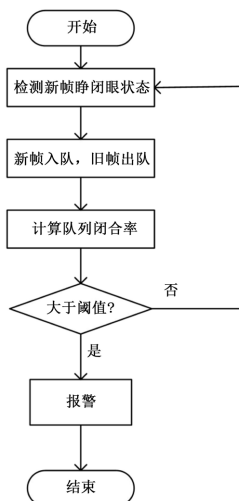
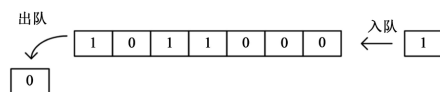


Figure 6. Eye fatigue detection process  
图 6. 眼部疲劳检测流程



在实际的编码过程中，我们通过一个先进先出队列来存储指定数量图片的检测结果。当相机采集到新的一帧图片时，我们将队头元素弹出，并将该帧图片的睁闭眼检测结果存入队尾。接着我们重新计算队列中闭眼帧占的比例，通过将计算出来的闭眼率与设定的阈值比较，当闭眼率大于设定的阈值时，疲劳检测系统应及时发出报警来规范驾驶人员的驾驶行为。对于该阈值的选择，我们通过统计 10 到 20 人模拟闭眼睡着状态下的闭眼率得到，实验表明，通过该方法选取的阈值在实际的疲劳检测中具有较高的准确性。PERCLOS 算法在本文中的实际应用流程如下图 6 所示。

### 3.2. 基于纵横比的嘴部状态分析

打哈欠次数增多是疲劳驾驶的明显特征之一。一般来说，打哈欠的次数越多、哈欠持续时间越长，表明驾驶员疲劳程度越深。本文对于嘴巴的状态检测方法使用基于宽高比计算的方法，如下图 7 所示。

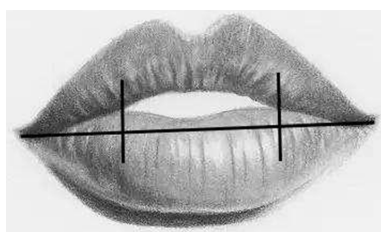


Figure 7. Schematic diagram of mouth aspect ratio

图 7. 嘴巴纵横比示意图

我们假设图中的两条高分别为  $h_1$ ,  $h_2$ ，宽为  $w$ ，嘴部纵横比为  $MAR$ ，则有如下计算公式：

$$MAR = \frac{h_1 + h_2}{2 \times w}$$

计算出  $MAR$  值之后，我们将其与张嘴阈值对比，当  $MAR$  大于阈值的时候，系统将该帧人脸图检测为嘴部张大状态。我们对嘴部是否打哈欠的估计仍然使用 PERCLOS 准则，取前 80 帧连续人脸图的嘴巴检测结果作为嘴部疲劳判定队列，当队列中嘴部张大率大于 40% 时，系统将驾驶员的当前状态判定为疲劳。

### 3.3. 基于坐标变换的头部姿态估计

在驾驶员驾驶过程中处于疲劳状态时，驾驶员会出现注意力分散、颈部肌肉控制力下降等现象，导致驾驶员的头部姿态处于下倾状态。因此我们在疲劳驾驶检测任务中需要将低头睡着的情况考虑进去，以防在低头睡着时，系统由于检测不到人脸信息而不能及时发出警报提醒驾驶员。

本文采用基于坐标变换的头部姿态估计算法，首先根据关键点检测模块提取到的人脸特征点，取出眉毛左右角点、眼睛左右角点、鼻子左右角点、嘴巴左右角点、嘴巴下角点以及下巴下角点共 14 个人脸关键点作为计算头部姿态的基准点。求出人脸图像中的这 14 个点与三维世界坐标系中对应的 14 个点之间的映射关系，最后我们从变换函数中取出头部下倾角的欧拉角来判断驾驶员是否处于低头睡着状态。下面我们从坐标变化、相机标定以及欧拉角计算三个方面展开具体的介绍。

#### 3.3.1. 坐标变换

从三维空间的坐标点映射到图像中的坐标点需要经过两次变换：1) 将世界坐标系中的点变换到相机坐标系中；2) 将相机坐标系的点变换到图像平面中。

其中三维坐标系是基于我们在真实世界中定义的坐标系，其中蕴含了三维空间物体的关系；相机坐

标系在相机内部定义的一种三维坐标系，是世界坐标点到图像坐标点之间转换的桥梁；图像坐标系作为一种二维坐标系，可以提供一种直观的二维感官。在下图 8 中展示了从世界坐标系到图像坐标系的转换过程。

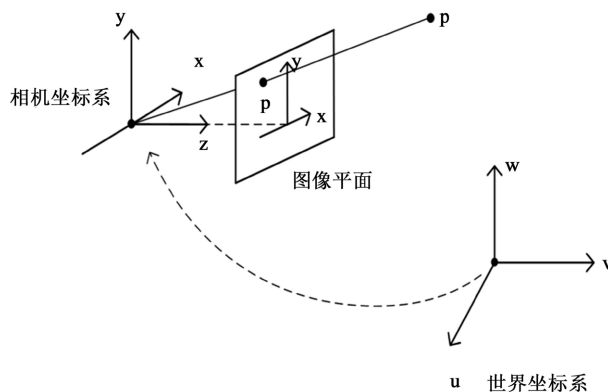


Figure 8. Schematic diagram of coordinate transformation  
图 8. 坐标变换示意图

其中世界坐标系到相机坐标系是一种三维点到三维点之间的转换，假设世界坐标系中某点坐标为 $(U, V, W)$ ，对应到相机坐标系中的点坐标为 $(X, Y, Z)$ ，我们有如下变换公式：

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = R \begin{bmatrix} U \\ V \\ W \end{bmatrix} + t$$

将其展开：

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} r_{00} & r_{01} & r_{02} & t_x \\ r_{10} & r_{11} & r_{12} & t_y \\ r_{20} & r_{21} & r_{22} & t_z \end{bmatrix} \begin{bmatrix} U \\ V \\ W \\ 1 \end{bmatrix}$$

其中  $R$  为旋转矩阵， $t$  为平移向量，这两个参数属于相机的外参。如果我们有成对的世界坐标系和相机坐标系点，我们就可以利用线性代数的知识求解出相机外参。相机坐标系到图像坐标系的坐标转换公式如下：

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \lambda \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}$$

式中  $x, y$  是相机坐标点在图像坐标系中的投影点坐标， $\lambda$  是比例因子， $f_x, f_y, c_x, c_y$  是相机内部参数，通过相机标定可以得到。最后我们将图像坐标系中的点通过放缩和平移变换就可以得到像素坐标中的点，也就是通过相机采集到的图像中的点坐标。通过上面的公式我们可以看出，如果有了相机内参、像素坐标系中的点坐标以及世界坐标系中的点坐标，就可以求出从相机坐标系到世界坐标系变换参数：旋转矩阵和平移向量，有了这两个参数我们就可以就出头部旋转欧拉角了。

### 3.3.2. 相机标定

由于图片从相机坐标映射到图像坐标时会产生畸变，为了减少相机采集到的图像与人眼观察图像之间的差异，我们就需要选择合适的相机内参和相机外参，相机标定就是为了解决这一任务而诞生的。我

们选择一个黑白棋盘作为模板图片，将其内外角点信息作为标定输入，在标定过程中我们根据结果对相机进行校正，最后输出相机的内参矩阵和畸变系数(如下表 2)。其中相机的焦距为(6.5308, 6.5308)，相机焦点为(3.1950, 2.3950)。

**Table 2.** Camera internal parameter matrix  $K$  and distortion coefficient  $D$   
**表 2.** 相机内参矩阵  $K$  和畸变系数  $D$

内参矩阵 $K$				
6.5308		0		3.1950
0		6.5308		2.3950
0		0		1
畸变系数 $D$				
7.0835	6.9140	0	0	-1.3073

### 3.3.3. 欧拉角计算

从上文介绍中，我们已经知道根据像素坐标点以及世界坐标点可以求出旋转矩阵  $R$ 。根据罗德里格斯旋转[19]公式，我们可以将旋转矩阵表示成欧拉角组成的正余弦矩阵，如下式：

$$R = \begin{bmatrix} \cos \alpha \cos \beta & \sin \gamma \cos \alpha - \cos \gamma \sin \alpha & \cos \gamma \sin \beta \cos \alpha + \sin \gamma \sin \alpha \\ \cos \alpha \sin \beta & \sin \gamma \sin \beta \sin \alpha + \sin \gamma \sin \alpha & \cos \gamma \sin \beta \sin \alpha - \sin \gamma \cos \alpha \\ -\sin \beta & \sin \alpha \cos \beta & \cos \alpha \cos \beta \end{bmatrix}$$

我们可以得到欧拉角计算公式：

$$\begin{cases} \alpha = \arctan\left(\frac{r_{21}}{r_{22}}\right) \\ \beta = \arctan\left(\frac{-r_{20}}{\sqrt{r_{21}^2 + r_{22}^2}}\right) \\ \gamma = \arctan\left(\frac{r_{21}}{r_{11}}\right) \end{cases}$$

根据欧拉角计算公式，我们就可以计算出头部下倾角。在实际的编码过程中，我们同样借鉴 PERCLOS 思想，选取指定数量的图像，计算这些图像中处于低头状态的图像所占的比例。当低头率大于指定阈值的时候，系统应当发出报警，本文选取的低头率阈值为 40%，指定的图像数量为 80 帧。

## 4. 实验测试与结果分析

本文的实验环境为一般的实验室环境，使用 Visual Studio 2015 作为开发工具，opencv3.0 和 C++ 开发语言进行。我们将 windows 平台训练的模型部署到 Ubuntu 需要经过模型转换过程，本论文使用的是 MNN 模型。本论文基于 Ubuntu 系统对模型的功能、性能等指标做了简单的测试。我们通过下面三个小节详细介绍测试过程及结果。

### 4.1. 功能性测试

#### 4.1.1. 通过眼睛状态检测疲劳

我们在通过眼睛部位特征检测疲劳度的时候，首先需要对每一帧照片的睁闭眼做出准确估计，如下

图 9 所示为某一帧照片当前眼睛是处于睁开还是处于闭合状态。图中左上角 cur eye 就是当前帧的眼睛状态，当眼睛睁开时显示 normal (右图)，当眼睛闭合时显示为 close (左图)。图中眼睛部位通过黄色的矩形框框出，对于眼部的检测我们是通过二分类卷积神经网络模型预测的。

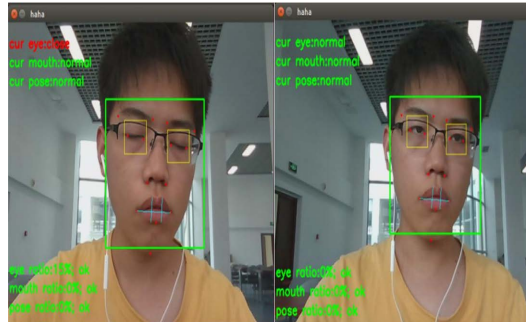


Figure 9. The state of the eye in the current frame  
图 9. 当前帧的眼睛状态

然而，我们并不能仅仅通过某一帧的睁闭眼状态来判断驾驶员是否疲劳，因为还有眨眼的情况存在。因此我们通过设置一个队列来判断驾驶员是否疲劳，在这个队列中会记录前 80 帧图像的睁闭眼结果，我们通过计算这个队里中闭眼帧的占比来判断当前驾驶员是否疲劳。经过频繁测试，我们选择 40% 作为疲劳判定阈值，当队列中图片的闭眼率大于 40%，则发出报警提示驾驶员当前已经处于疲劳状态。测试结果如下图 10 所示。

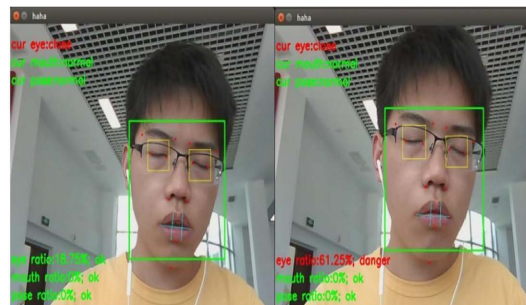


Figure 10. Detecting fatigue based on eye condition  
图 10. 根据眼部状态检测疲劳

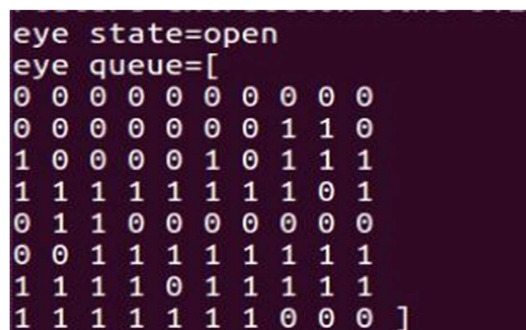


Figure 11. Eye closure rate for all images in the current queue  
图 11. 当前队列所有图像的闭眼率

图 11 eye ratio 表示当前队列所有图像的闭眼率, 当闭眼率小于 40% 时, 系统判为 ok (如左上图所示); 当闭眼率大于 40% 时, 系统会判为 danger (如右上图所示) 表明此时处于疲劳状态。除此之外我们在系统打印日志的地方打印了前 80 帧图像的眼睛状态, 其中 0 表示眼睛睁开, 1 表示眼睛闭合。队列日志上面的 eye state 就是当前这帧照片的眼睛状态, 其与眼睛状态队列的最后一帧照片的 0 相对应。

#### 4.1.2. 通过嘴巴状态检测疲劳

我们在对嘴巴做系统测试的时候同眼睛一样, 分为每一帧的嘴巴状态识别以及当前根据嘴巴队列判断驾驶员是否疲劳。如下图 12 所示为某一帧的嘴巴检测结果, 图中左边的 cur mouth 显示的就是当前帧图像的检测结果, 从图中我们可以看出当嘴巴处于正常状态时, cur mouth 显示为 normal; 当嘴巴处于张大状态时, cur mouth 显示为 yaw。图中嘴巴区域的三条线分别表示 1 条宽 2 条高, 我们通过嘴巴的宽高比来计算当前状态下嘴巴是否处于张大状态。

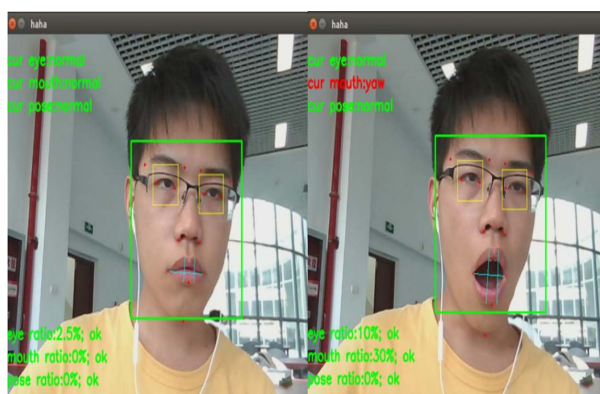


Figure 12. The state of the mouth in the current frame  
图 12. 当前帧的嘴部状态

同样, 仅凭某一帧图像中的嘴巴状态我们还不足以判定驾驶员就在打哈欠。因此, 我们对嘴巴也设定了一个队列, 记录前 80 帧照片的嘴巴状态。当嘴巴张大比率超过 40% 时, 系统发出报警提示。如下图 13 所示, 左下角的 mouth ratio 即为嘴巴队列的张大比率, 当小于 40% 时, 显示为 OK (左图); 当大于 40% 时, 显示为 yawing (右图)。图 14 中左下图显示为当前帧的嘴巴宽高比和当前的嘴巴队列记录的每一帧嘴巴识别结果, 0 表示嘴巴为正常状态, 1 表示嘴巴为张大状态(图 14)。

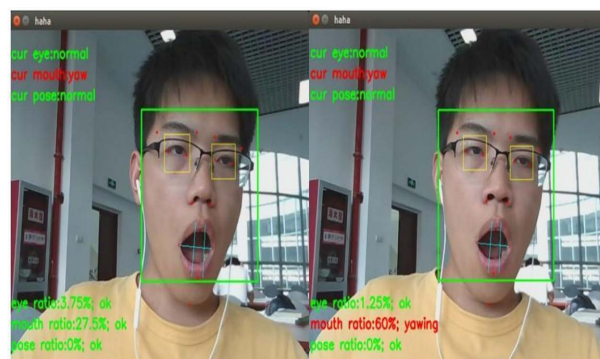


Figure 13. Detects fatigue based on mouth status  
图 13. 根据嘴巴状态检测疲劳

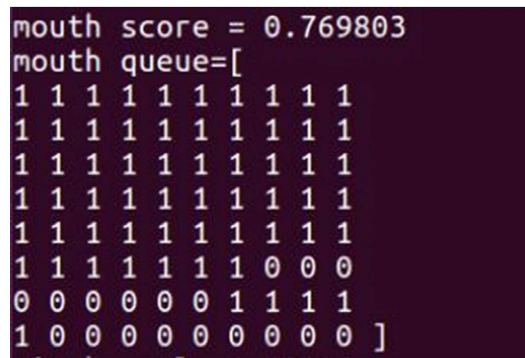


Figure 14. Mouth opening rate of all images in the current queue

图 14. 当前队列所有图像的张嘴率

#### 4.1.3. 通过头部姿态检测疲劳

对于某帧图像的头部姿态判定我们是通过人脸 14 个关键点来计算的，这 14 个关键点位置如下图 15 中的红色点表示。当头部处于正常驾驶状态时，图 15 中的左上角 cur pose 显示为 normal (左图)；当人头部低下时，cur pose 显示为 down (右图)。

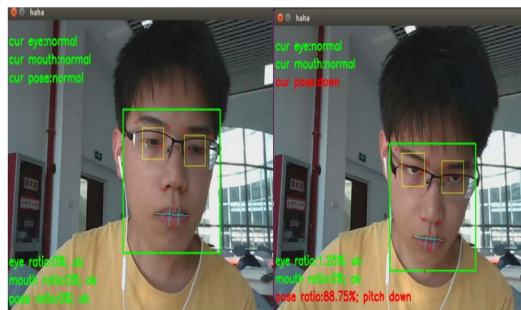


Figure 15. Head state for the current frame

图 15. 当前帧的头部状态

我们同样使用队列来判断驾驶员是否睡着，当队列中的图像低头率大于 40%时，系统将会判断为低头睡着，并发出报警提示；如下图 16 所示，图中左下角的 pose ratio 即为队列中图像的低头率，当其小于 40%时，pose ratio 处显示 ok；当其大于 40%时，pose ratio 出显示为 danger。图 17 为当前时刻前 80 帧图像的头部姿态检测结果，0 表示头部姿态正常，1 表示头部低下。pitch angle 表示当前帧的头部下倾角。

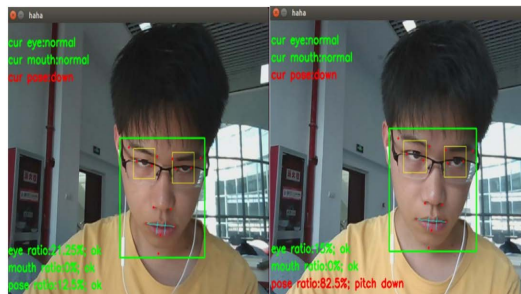


Figure 16. Detects fatigue based on the state of the head

图 16. 根据头部状态检测疲劳



Figure 17. Head pose detection results for the current frame

图 17. 当前帧头部姿态检测结果

## 4.2. 精度测试

本文对于疲劳驾驶检测基于眼睛、嘴巴、状态的精度进行了详细测试，测试结果如下表 3 所示。从表中我们可以看出对于眼睛和嘴巴状态的预测精度分别是 97.1% 和 97.5%，对于头部姿态的预测准确率为 88.1%，相对偏低。

Table 3. Accuracy of eye, mouth, and head pose prediction  
表 3. 眼睛、嘴巴、头部姿态预测精度

任务名称	测试图片总数	测试结果	准确率
预测眼睛状态	934	正确: 907 错误: 27	97.1%
预测嘴巴状态	881	正确: 859 错误: 22	97.5%
预测头部状态	789	正确: 695 错误: 94	88.1%

接下来我们分别模拟驾驶员疲劳场景来检测由眼睛检测出疲劳、由嘴巴检测出打哈欠、由头部姿态检测出疲劳的精度，如下表 4 所示。从表中可以看出对于由眼睛和嘴巴检测出疲劳的准确率相对较高，对于头部姿态的预测精度相对较低。

Table 4. Accuracy of predicting fatigue scenarios by eye, mouth, and head posture  
表 4. 通过眼睛、嘴巴、头部姿态预测疲劳场景的精度

任务名称	测试场景数	测试结果	准确率
通过眼睛测试疲劳	100	正确: 95 错误: 5	95%
通过嘴巴测试疲劳	100	正确: 97 错误: 3	97%
通过头部疲测试劳	100	正确: 85 错误: 15	85%

### 4.3. 速度测试

对疲劳检测系统的速度测试我们主要通过对图像预处理时间、人脸检测时间、特征提取时间、眼睛状态预测时间进行测试,由于嘴部状态和头部姿态我们都是通过数学计算得出,因此对其速度测试略过。

下面我们一次性测试 1000 帧图片,对各个阶段的时间求平均,可以得到如下表 5 所示结果。从表中我们可以看出疲劳检测系统对于一张照片的总的检测时间在 62 ms 左右,因此平均 1 秒可以检测 16 张左右的人脸图,基本满足实时性需求。

**Table 5.** Average time for each stage of a fatigue detection system  
**表 5.** 疲劳检测系统各个阶段的平均时间

阶段名称	平均时间
人脸检测	56.49 ms
特征点提取	4.309 ms
眼部状态检测	0.935 ms

## 5. 结论与展望

本文基于神经网络对检测疲劳驾驶的方法进行了一定的研究,提出了基于神经网络的疲劳检测方法:疲劳估计模块包含三个子模块:通过眼睛检测疲劳度、通过嘴巴检测疲劳度、通过头部姿态检测疲劳度。此外,借鉴 PERCLOS 经验准则,提出了基于一定时间周期的疲劳判定法,使系统对于驾驶员的疲劳预估更加准确。实验结果表明,使用本文提出的方法,能够实现多特征点的定位与识别,且根据眼部状态以及嘴部状态预测疲劳的精度分别达到了 97.1%和 97.5%,根据头部状态预测疲劳的精度达到了 88.1%,处理速度也达到了每张图片 62 ms。相较于传统只基于单一特征检测疲劳的方法,本文提出的算法在保证准确率和效率的前提下,提高了综合检测能力。

本文对所设计的疲劳驾驶检测方法与目前已有的解决方案进行了对比,并在测试过程中表现出一定的优越性。但是,本文所提出的方案仍然有一些不足之处,有如下展望:1)我们在测试阶段发现,系统对于头部姿态的预测精度仅仅只有 85%左右,这还不能满足实际场景的需求,因此对于头部姿态估计的方案有待优化;2)系统对于鲁棒性的考虑还不够充分,在实际的驾驶场景中可能会存在曝光过度、机动车出现震动、驾驶员各种不良的驾驶习惯等。因此,在实际投入运行前,还需要对这些场景做出优化方案。3)在保证精度的前提下,处理效率为每张图片 62 ms,有待优化与提高。

## 基金项目

国家自然科学基金(62276167)。

## 参考文献

- [1] Sayed, R. and Eskandarian, A. (2001) Unobtrusive Drowsiness Detection by Neural Network Learning of Driver Steering. *Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering*, **215**, 969-975. <https://doi.org/10.1243/0954407011528536>
- [2] 柴萌. 长途客车驾驶员疲劳状态辨识与预警[D]: [博士学位论文]. 长春: 吉林大学, 2019.
- [3] 陈志勇, 杨佩, 彭力, 莫子兴, 蔡岗. 基于 BP 神经网络的驾驶员疲劳监测研究[J]. *计算机科学*, 2015, 42(S1): 67-69+93.
- [4] 王斐, 王少楠, 王惜慧, 彭莹, 杨乙丁. 基于脑电图识别结合操纵特征的驾驶疲劳检测[J]. *仪器仪表学报*, 2014, 35(2): 398-404.



- [5] Chuang, C.-H., Huang, C.-S., Ko, L.W. and Lin, C.-T. (2015) An EEG-Based Perceptual Function Integration Network for Application to Drowsy Driving. *Knowledge-Based Systems*, **80**, 143-152. <https://doi.org/10.1016/j.knosys.2015.01.007>
- [6] 徐礼胜, 张闻勛, 庞宇轩, 吴承暘. 基于短时心电信号的疲劳驾驶检测算法[J]. 东北大学学报(自然科学版), 2019, 40(7): 937-941.
- [7] Reddy, B., Kim, Y.H., Yun, S., et al. (2017) Real-Time Driver Drowsiness Detection for Embedded System Using Model Compression of Deep Neural Networks. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, Honolulu, 21-26 July 2017, 121-128. <https://doi.org/10.1109/CVPRW.2017.59>
- [8] 李文学, 谢凯. 基于深度学习的疲劳驾驶检测方法研究[J]. 电子世界, 2019(17): 51-52.
- [9] Papageo, R., Giou, C.P., et al. (1998) A General Framework for Object Detection. *Proceedings of the 6th International Conference on Computer Vision*, Bombay, 4-7 January 1998, 555-562.
- [10] 潘志庚, 刘荣飞, 张明敏. 基于模糊综合评价的疲劳驾驶检测算法研究[J]. 软件学报, 2019, 30(10): 2954-2963.
- [11] Ed-Doughmi, Y., Idrissi, N. and Hbali, Y. (2020) Real-Time System for Driver Fatigue Detection Based on a Recurrent Neuronal Network. *Journal of Imaging*, **6**, 8.
- [12] Deng, J.K., Guo, J., Zhou, Y.X., et al. (2019) RetinaFace: Single-Stage Dense Face Localisation in the Wild. 2020 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Seattle, 13-19 June 2020, 5202-5211. <https://doi.org/10.1109/CVPR42600.2020.00525>
- [13] Ren, S.Q., Cao, X.D., Wei, Y.C., et al. (2016) Face Alignment via Regressing Local Binary Features. *IEEE Transactions on Image Processing*, **25**, 1233-1245. <https://doi.org/10.1109/TIP.2016.2518867>
- [14] Valle, R., et al. (2019) Face Alignment Using a 3D Deeply-Initialized Ensemble of Regression Trees. *Computer Vision and Image Understanding*, **189**, Article ID: 102846. <https://doi.org/10.1016/j.cviu.2019.102846>
- [15] 万杨. 显性形状回归的人脸对齐方法实验研究[D]: [硕士学位论文]. 大连: 大连理工大学, 2015.
- [16] 程泽, 林富生, 靳朝, 周鼎贺. 基于轻量化卷积神经网络的疲劳驾驶检测[J]. 重庆理工大学学报(自然科学), 2022, 36(2): 142-150.
- [17] 周壮壮, 陈艳, 唐苏, 龚桂玉, 周丹. 基于 PERCLOS 的眼部疲劳检测系统研究[J]. 集成电路应用, 2021, 38(2): 48-49.
- [18] Li, W. and Liu, K. (2021) Confidence-Aware Object Detection Based on MobileNetv2 for Autonomous Driving. *Sensors (Basel, Switzerland)*, **21**, Article No. 2380. <https://doi.org/10.3390/s21072380>
- [19] 李庆臣. 基于面部特征的疲劳驾驶检测系统设计[D]: [硕士学位论文]. 郑州: 郑州大学, 2019.