

# 基于深度学习的疲劳驾驶检测算法<sup>①</sup>

戴诗琪, 曾智勇

(福建师范大学 数学与信息学院, 福州 350108)

通讯作者: 曾智勇, E-mail: [zzyong@fjnu.edu.cn](mailto:zzyong@fjnu.edu.cn)

**摘要:** 针对现有疲劳驾驶检测算法实用性差或准确率低的问题, 本文提出了一种基于深度学习的疲劳驾驶检测算法. 首先, 使用 HOG (Histogram of Oriented Gradient) 特征算子检测人脸的存在; 其次, 利用特征点模型实现人脸的对齐, 同时实现眼睛、嘴巴区域的分割; 最后通过深度卷积神经网络提取驾驶员的眼部疲劳特征, 并融合驾驶员嘴部的疲劳特征进行疲劳预警. 大量的实验表明, 该方法在疲劳驾驶检测的准确率、实时性等方面都取得明显的性能提升.

**关键词:** HOG 算子; 特征点模型; 深度学习; 卷积神经网络; 疲劳驾驶检测

引用格式: 戴诗琪, 曾智勇. 基于深度学习的疲劳驾驶检测算法. 计算机系统应用, 2018, 27(7): 113-120. <http://www.c-s-a.org.cn/1003-3254/6415.html>

## Fatigue Driving Detection Algorithm Based on Deep Learning

DAI Shi-Qi, ZENG Zhi-Yong

(College of Mathematics and Informatics, Fujian Normal University, Fuzhou 350108, China)

**Abstract:** A fatigue driving detection algorithm based on deep learning was proposed to solve the problem of poor practicability or low accuracy of existing fatigue driving detection algorithms. First, the Histogram of Oriented Gradient (HOG) feature operator was used to detect the presence of human faces. Secondly, the feature points model was used to realize the face alignment and segment the eye and mouth regions. Finally, the fatigue features of the driver's eyes were extracted by the deep convolutional neural network, and the fatigue features of the driver's mouth were fused to carry out fatigue warning. Experiments show that the proposed method can achieve significant performance improvement in terms of accuracy and real-time performance of fatigue driving detection.

**Key words:** Histogram of Oriented Gradient (HOG) operator; feature point model; deep learning; convolutional neural network; fatigue driving detection

随着社会经济飞速发展, 城市道路的大规模建设, 机动车保有量呈指数增加, 公路通车里程逐年增长, 交通事故的发生率日益增高. 同时随着现代人生活节奏加快, 熬夜现象严重、缺乏锻炼、饮食不规律等等因素造成了现代人容易疲劳精力差, 随之出现的疲劳驾驶成为造成交通事故的主要原因之一. 如何有效地预防疲劳驾驶已经成为世界性的难题, 国内外有不少学

者采用不同技术进行了研究, 现有的疲劳驾驶检测算法按疲劳特征可以分为两大类: 基于生理特征的算法和基于行为特征的算法. 其中生理特征包括生理信号特征和生理反应特征, 文献[1-5]提出了基于脑电信号 (EElectroEncephaloGram, EEG) 的睡眠检测模型, 结果表明, EEG 能够正确地区分清醒状态和睡眠状态, 但脑电信号容易受到噪音的干扰并且难以收集. 眼电信号

<sup>①</sup> 基金项目: 福建省科技重点项目 (2013H0020)

Foundation item: Key Science and Technology Project of Fujian Province (2013H0020)

收稿时间: 2017-10-30; 修改时间: 2017-11-21; 采用时间: 2017-12-01; csa 在线出版时间: 2018-06-27

(ElectroOculoGram, EOG)<sup>[6]</sup>相对于 EEG 而言更易于收集并且可以免疫轻微的噪音,但仍需安装头部设备进行采集.文献[7]运用远程视频 CCD 相机进行实时的图像采集,监视眼睑运动、面部朝向和视线方向三种生理反应特征.此外,行为特征包括驾驶人行为特征和车辆行为特征,Baronti 等人<sup>[8]</sup>将分布式压力传感器嵌入方向盘里面测量手握方向盘的压力以及位置,集合每个元件的测量数据统计分析进行疲劳预警,这种传感器结构简单且成本较低,在实际应用中取得了初步的成果,缺点是会受到驾驶人的个人习惯的干扰.文献[9-11]利用一个摄像头捕捉道路前方路况,参照路上的白线判断车辆行驶是否规范,从而判断车辆是否处于失控状态.这类方法不需要驾驶人直接接触检测装置,并且在汽车现有装置的基础上对设备需求较低,实用性很强,易于推广,但会受到路面条件和车辆型号的限制.

随着近年来深度学习<sup>[12-19]</sup>在目标检测等领域的巨大成功,深度学习技术也推动了疲劳驾驶检测的研究.Zhu X 等人<sup>[20]</sup>提出了一个基于 EOG 的疲劳检测回归模型,该模型使用卷积神经网络 (Convolutional Neural Network, CNN) 进行无监督的学习特征表达,代替人工设计特征提取的过程,并在后处理中使用 LDS (Linear Dynamic System) 算法减少不利的干扰.文献[21]利用 CNN 提取手部位置的判别信息,从而无监督的学习和预测安全和不安全的四种驾驶姿势.文献[22]提出了一种基于 CNN 的计算眨眼频率的眼睛状态识别方法,采用红外视频检测提升了驾驶员佩戴太阳镜情况的准确率,但疲劳参数过于单一.研究表明,使用深度神经网络直接从图像中学习视觉特征表示,相较于手工设计的特征对光照、姿势等条件变化具有更好的鲁棒性,预测精度显著提高.

以上的疲劳驾驶检测算法都取得了一定的效果与进展,但仍然存在很大的局限性:基于生理信号特征的检测算法虽然具有较高的精度,但需要驾驶人与检测设备有直接的身体接触以采集信号,可能干扰到驾驶员的驾驶操作而且设备成本太高,更适用于实验室环境,不适用于实际应用;基于行为特征的检测算法不需要驾驶人直接接触检测装置,并且在汽车现有装置的基础上对设备需求较低,实用性很强,易于推广,但会受到驾驶人个人习惯、路面条件和车辆型号的限制,在雨雪天气或路面情况不理想时检测的准确率较低.在保证一定的准确率和良好的实时性的情况下,基于计

算机视觉处理生理反应特征的算法更易于推广.但是人工设计的特征分辨能力较差,同时,单一依靠某一个视觉特征可能遇到困难,如驾驶人戴眼镜时眼镜反光可能造成摄像头无法捕捉眼球运动,此外,眼睛的开闭程度可能因人而异,不规则的头部运动可能产生误报等等.本文利用深度学习提取眼部部位的视觉特征,建立视觉模型,以提升视觉特征的辨别能力,然后采用 PERCLOS<sup>[23]</sup>算法结合打哈欠的频率判断驾驶员是否疲劳,减少单一特征的误检率.实验结果表明,本文提出的基于深度学习的疲劳驾驶检测方法是有效的.

## 1 基于深度学习的疲劳驾驶检测

本文设计的疲劳驾驶检测算法分为 4 步:首先利用 HOG 特征模型检测出摄像头实时获取的画面中的驾驶员脸部图像,然后根据面部标识算法定位出人眼和嘴部区域,接着将定位出的眼睛区域图像输入深度卷积神经网络提取眼睛的视觉特征并分别判断眼睛和嘴巴的开闭程度,最后依据 PERCLOS 算法求出单位时间内眼睛闭合状态所占的百分比,同时结合打哈欠的频率完成疲劳驾驶预警机制,整体框架如图 1 所示.



图 1 整体框架

### 1.1 基于 HOG 的人脸检测

HOG 最早由 Dalal 和 Triggs 等人<sup>[24]</sup>在 2005 年提出,利用梯度及梯度方向的统计信息来描述图像.因为梯度往往产生在边缘处,可以描述局部目标的形状,且梯度受到光照条件变化的影响较小,同一个目标的梯度特征在或亮或暗的图像中具有一致性.与其它的图像描述特征相比,HOG 特征具有较好的光学不变性与几何不变性,尤其适用于目标检测.

在进行人脸检测时,生成 HOG 特征的主要步骤如下:

(1) 首先采用 Gamma 压缩法将由摄像头捕获的视频帧图像进行颜色空间归一化处理,减少颜色数据和光照因素的影响,公式如下:

$$I(x,y) = I(x,y)^{\text{Gamma}} \quad (1)$$

(2) 遍历图像,分别计算每个像素点 $(x,y)$ 在水平方向上 $G_x(x,y)$ 和垂直方向上的梯度 $G_y(x,y)$ , $H(x,y)$ 表示点 $(x,y)$ 的像素值,公式如下:

$$G_x(x,y) = H(x+1,y) - H(x-1,y) \quad (2)$$

$$G_y(x,y) = H(x,y+1) - H(x,y-1) \quad (3)$$

从而计算出该像素位置的梯度值 $G(x,y)$ 和梯度方向 $\alpha(x,y)$ :

$$G(x,y) = \sqrt{G_x(x,y)^2 + G_y(x,y)^2} \quad (4)$$

$$\alpha(x,y) = \tan^{-1} \left( \frac{G_y(x,y)}{G_x(x,y)} \right) \quad (5)$$

(3) 将输入图像划分为若干个单元区域, 每个单元区域包含  $n \times n$  个像素, 并将  $0^\circ$  至  $360^\circ$  划分为  $N$  个方向区间, 统计每个单元区域像素的梯度方向并根据梯度值加权投影, 得到该单元区域的梯度方向直方图。

(4) 将相邻的  $K$  个单元区域组成单元块, 把单元块中所有单元区域的 HOG 合并得到特征向量, 并对其进行归一化, 对光照等条件变化具有更好的鲁棒性。

(5) 按一定的步长  $L$  滑动单元块遍历整个图像, 将所有单元块的 HOG 特征合并形成最终的 HOG 特征向量描述整个图像。

本文选用的单元区域大小为  $8 \times 8$  像素 (即  $n$  值为 8), 方向区间  $N$  值为 9, 单元块大小为  $16 \times 16$  像素 (即  $K$  值为 4), 滑动步长  $L$  为 8, 当输入图像大小为  $128 \times 128$  像素时, 则生成的 HOG 特征向量为  $((128-16)/8+1)^2 \times 4 \times 9 = 8100$  维。最后通过提取大量的已标定人脸样本的 HOG 特征, 并使用 SVM 对其进行训练生成人脸检测器, 具体训练算法如算法 1。

算法 1. 基于 HOG 特征的 SVM 分类训练算法

1. 输入人脸训练样本及样本中人脸的位置坐标  $T$ ,  $pos_{i,j}$  代表第  $i$  副图像的第  $j$  个人脸包围框的 4 维位置坐标 ( $top, left, width, height$ );  
 $T = \left\{ \left( (I_i, pos_{i,j}), (I_i, pos_{i,j})^+, (I_i, pos_{i,j})^- \right) \right\}_{i=1, j=1}^N$
2. 对样本图像进行预处理后逐个扫描, 提取每幅图片的 HOG 特征;
3. 将已标定人脸的位置设为正样本, 在其余位置随机选取非人脸区域作为负样本, 进行 SVM 训练, 当  $loss$  小于设定值 0.01 时训练结束, 表明 SVM 优化完成, 输出训练好的人脸检测器模型  $detector$ 。

图 2 为使用  $detector$  进行人脸检测的示例, 将由  $detector$  返回的最大人脸包围框作为需采集驾驶员的人脸图像。

### 1.2 人脸关键点定位

在驾驶过程中, 驾驶员的面部会产生不同角度的旋转, 为了提高识别的准确率, 本文使用 2014 年 Kazemi 和 Sullivan<sup>[25]</sup> 提出的面部标志估算算法检测人脸特征点并对齐人脸。首先选取人脸的 68 个特殊标志点模型,

从眉毛外沿至下颌底部, 包括眼睛轮廓和嘴部轮廓等, 如图 3 所示。然后利用该算法提出的基于梯度增强 (radient boosting) 的框架, 通过优化损失函数和误差的总和来学习回归树集合 (Ensemble of Regression Trees, ERT), 检测出脸部图像中的 68 个关键点, 最后对齐人脸, 并根据关键点序号定位出所需要的眼部、嘴部图像, 如图 4 所示。



图 2 人脸检测示例

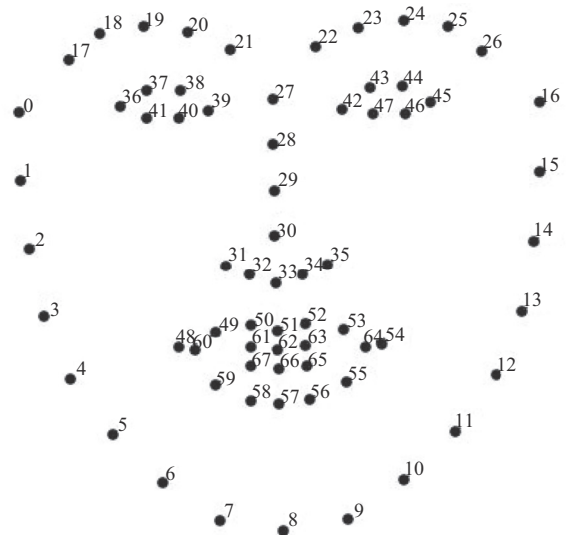


图 3 68 个面部标识点

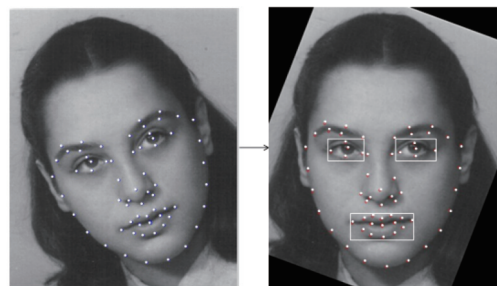


图 4 眼部、嘴部定位

### 1.3 卷积神经网络识别眼部状态

在定位眼睛和嘴部区域后, 分割上述局部区域, 然后建立眼睛的视觉特征模型。传统的基于深度学习的

目标识别方法主要针对单张的静态图片,使用 CNN 网络提取特征,将某一中间层作为特征描述然后利用 softmax 等分类器对其分类,中间层的特征描述通常维数较高.而在疲劳驾驶检测过程中需要对视频监控的连续多副图像进行处理,包含了更多的光照、遮挡等条件变化信息.本文没有采用传统的损失函数而是采用基于三元组损失函数的深度卷积神经网络学习眼睛特征.传统的损失函数主要计算最小化预测标记与真实标记之间的误差,它虽然可以保障类间分离,但不能实现类内聚集.三元组损失函数计算最大化不同类对

象之间的欧式距离与同类对象之间的欧式距离的差值,这可以保障类间距离最大化,类内距离最小化,实现深度学习网络类内聚集、类间远离,提高了深度网络的分类能力.区分睁眼和闭眼两个状态集合,将独立的样本分类任务转换为样本间的距离度量学习任务,保留了更多数据内部的关联性信息,直接通过两幅图像特征的欧式距离来判断两幅图像是否为同一类别,去除使用中间层作为特征描述的间接性,实现端到端的学习,降低计算量,提高系统的实时性,网络结构如图 5 所示.

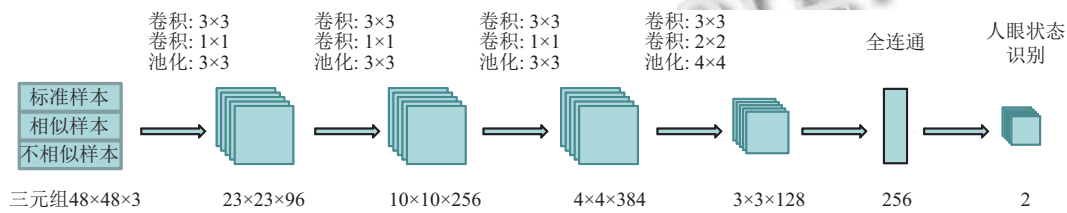


图 5 网络结构

从图 5 可以看出 CNN 结构的输入尺寸大小为 \$48 \times 48\$, 输出层为 2 维向量, 表示 2 个类别 (眼睛闭合和眼睛睁开), 中间包含四个阶段的卷积层、归一化层、池化层的操作. 第一阶段的卷积核大小分别为 \$3 \times 3\$ 和 \$1 \times 1\$, 池化层以 \$3 \times 3\$ 的大小平均池化对比归一化的特征图, 后续阶段的相应操作与第一阶段类似.

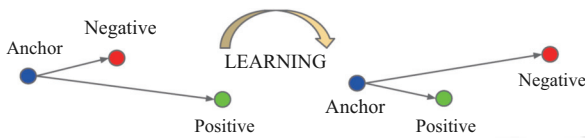


图 6 三元组示例

如图 6 所示, 一个三元组包含三个样本: 一个随机选择的眼睛样本 Anchor, 一个与随机样本相似的样本 Positive, 一个与随机样本不相似的样本 Negative, 核心思想为通过学习使得随机样本与相似样本间的距离相差最小, 与不相似样本间的距离相差最大, 公式为:

$$\|f(x_i^a) - f(x_i^p)\|_2^2 + \alpha < \|f(x_i^a) - f(x_i^n)\|_2^2, \quad (6)$$

$$\forall (f(x_i^a), f(x_i^p), f(x_i^n)) \in D$$

式中,  $x_i^a$  代表随机样本,  $x_i^p$  代表相似样本,  $x_i^n$  代表不相似样本,  $f(x)$  代表通过网络提取的特征, 下标  $i$  代表每个三元组的序号,  $D$  代表于训练的三元组集. 通过迭代训练, 使  $N$  个三元组都满足上述条件, 其损失函数构造为:

$$L = \sum_i^N \left[ \|f(x_i^a) - f(x_i^p)\|_2^2 - \|f(x_i^a) - f(x_i^n)\|_2^2 + \alpha \right]_+ \quad (7)$$

其中,  $L$  表示三元组分类损失,  $f(x_i^a)$ 、 $f(x_i^p)$ 、 $f(x_i^n)$  分别表示随机样本、正例样本和负例样本的卷积池化后的激活特征,  $\alpha$  表示  $x_i^a$  与  $x_i^n$  之间的距离与  $x_i^a$  与  $x_i^p$  之间的距离之间的最小间隔,  $\alpha$  设置为一个常数.

在式 (7) 中, 三元组分类损失对三个样本特征的梯度分别为:

$$\frac{\partial L}{\partial f(x_i^a)} = 2 * (f(x_i^a) - f(x_i^p)) - 2 * (f(x_i^a) - f(x_i^n))$$

$$= 2 * (f(x_i^n) - f(x_i^p)) \quad (8)$$

$$\frac{\partial L}{\partial f(x_i^p)} = 2 * (f(x_i^a) - f(x_i^p)) * (-1)$$

$$= 2 * (f(x_i^p) - f(x_i^a)) \quad (9)$$

$$\frac{\partial L}{\partial f(x_i^n)} = -2 * (f(x_i^a) - f(x_i^n)) * (-1)$$

$$= 2 * (f(x_i^a) - f(x_i^n)) \quad (10)$$

公式 (7) 的偏微分可表示为:

$$\frac{\partial L}{\partial w} = 2 * (f(x_i^a) - f(x_i^p)) * \left( \frac{\partial L}{\partial f(x_i^a)} - \frac{\partial L}{\partial f(x_i^p)} \right)$$

$$- 2 * (f(x_i^a) - f(x_i^n)) * \left( \frac{\partial L}{\partial f(x_i^a)} - \frac{\partial L}{\partial f(x_i^n)} \right) \quad (11)$$

我们使用随机梯度下降算法训练上面描述的基于三元组损失深度卷积神经网络模型,具体训练算法如算法2.

算法2. 基于三元组随机梯度下降训练算法

```

输入: 眼睛训练样本  $D = \{(I_i^-, I_i^+, I_i^c)\}_{i=1}^N$ 
输出: 网络参数  $\{w\}$ 
while  $t < T$ , do
     $t = t + 1$ 
     $\frac{\partial L}{\partial w} = 0$ 
    for 所有训练三元组样本  $D$ , 循环
        利用前向传播算法计算  $f(x_i^c)$ ,  $f(x_i^+)$  和  $f(x_i^-)$ 
        利用后向传播算法计算  $\frac{\partial L}{\partial f(x_i^c)}$ ,  $\frac{\partial L}{\partial f(x_i^+)}$ ,  $\frac{\partial L}{\partial f(x_i^-)}$ 
        利用公式(11)计算  $\frac{\partial L}{\partial w}$ 
    for 循环结束
        更新网络权重参数  $w = w^{t-1} - \lambda_t \frac{\partial L}{\partial w}$ 
while 循环结束

```

利用深度卷积神经网络判别单帧图像的眼部开闭状态后,根据 PERCLOS 定义计算单位时间内眼睛闭合状态所占时间的百分比.计算公式如下:

$$f_{eye} = \frac{t_{clo}}{t} \times 100\% \quad (12)$$

$t$  为单位时间内的视频帧总数,  $t_{clo}$  为单位时间内眼睛闭合的帧总数.

#### 1.4 嘴部状态识别

在驾驶过程中驾驶员的嘴部通常有闭合、说话或唱歌、打哈欠三种状态,打哈欠的状态下嘴部的张开程度将明显并持续大于前两种状态,因此在嘴部定位精确的情况下,可直接根据嘴部图像的宽高比判断嘴巴的张开程度,快速区分这三种状态,减小计算量.具体判断步骤如下:

(1) 设嘴部图像的宽为  $w$ , 高为  $h$ , 则其宽高比  $m = \frac{w}{h}$ , 当  $m$  小于某一阈值  $T$ , 可判断当前图像帧为打哈欠;

(2) 当嘴巴的宽高比  $m$  值大于阈值  $T$  的视频帧连续达到  $K$  帧时,记作打哈欠一次,即打哈欠从嘴巴张开到闭合结束为打哈欠一次,统计单位时间内驾驶员的打哈欠次数记作  $f_{mouth}$ .

分别判断眼部和嘴部的疲劳状态后,融合眼睛的疲劳参数与嘴部的疲劳参数,分别赋予一定的权重  $w_1$ 、 $w_2$ , 计算最终的疲劳值  $F$ , 公式如下:

$$F = w_1 f_{eye} + w_2 f_{mouth} \quad (13)$$

## 2 实验结果与分析

### 2.1 数据集与网络训练

本文采用了 ZJU<sup>[26]</sup> 眨眼视频数据集和 CEW<sup>[27]</sup> 眼睛数据集, ZJU 包含 20 个人的共 80 个 avi 剪辑视频,每个人包含四类片段:正面视图未戴眼镜、正面视图戴薄边框眼镜、正面视图戴黑色框架眼镜、向上的视角未戴眼镜.在眨眼的过程中手动选择眼睛睁开、半开半闭、闭合的眼睛图像,并将左右眼分开存储,共产生 7334 幅图像,包括 1574 张闭眼样本和 5770 张睁眼样本. CEW 包含 2423 幅人脸图像,其中 1192 幅为从网上收集的受试者双眼闭合的图像,1231 幅从 LFW<sup>[28]</sup> 选取的已标记人脸的睁眼图像,分别收集左右眼的图像产生 2384 幅闭眼样本和 2462 幅睁眼样本.将所有样本图像归一化为  $48 \times 48$ , 如图 7 所示,分别合并睁眼、闭眼样本集(不区分左右眼),然后分别从中划分 20% 的样本作为测试集,其余为训练集.使用随机选择策略在训练集中生成 1000 组三元组用于网络迭代训练,本实验 Ubuntu 环境下使用 torch 框架实现.

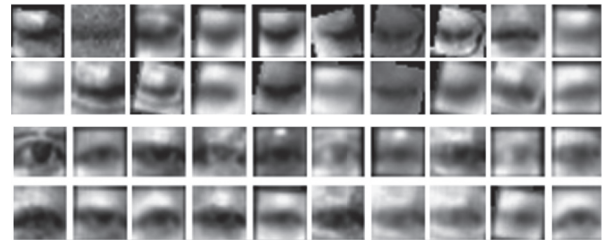


图7 ZJU 和 CEW 数据集示例

### 2.2 实验结果

为了验证本文提出的基于深度学习的防疲劳驾驶检测方案能否达到预期效果,分别从以下几个方面进行了测试实验.

#### 2.2.1 眼睛状态识别测试

表1为利用本文方法训练的模型对测试集进行测试的结果,可以看出对眼部状态的识别准确率较为理想.

表1 本文方法眼睛状态识别结果

眼睛状态	正确数	识别率 (%)
睁眼	1530	92.8
闭眼	730	92.2

表2为利用传统的 CNN 方法在相同测试集上的测试结果.

表2 传统 CNN 方法眼睛状态识别结果

眼睛状态	正确数	识别率 (%)
睁眼	1432	87
闭眼	673	84.9

表3 为本文方法与传统的 CNN 方法的对比, 可见本文方法在准确率与实时性上均有提升.

表3 眼睛状态识别方法对比

	本文方法	CNN
准确率 (%)	92.5	86
人脸检测时间 (s)	0.28	0.45
眼、嘴定位时间 (s)	0.007	0.01
眼部识别时间 (s)	0.15	0.40

### 2.2.2 疲劳状态识别测试

本文采用 YawDD<sup>[29]</sup>数据集作为疲劳检测的标准数据集, 该数据集包含不同性别、不同种族的驾驶员在不同光照环境下拍摄的驾驶视频, 如图8所示.



图8 YawDD 数据集示例

图9 为嘴部宽高比  $m$  值与连续视频帧数  $K$  值的设定对准确率的影响.

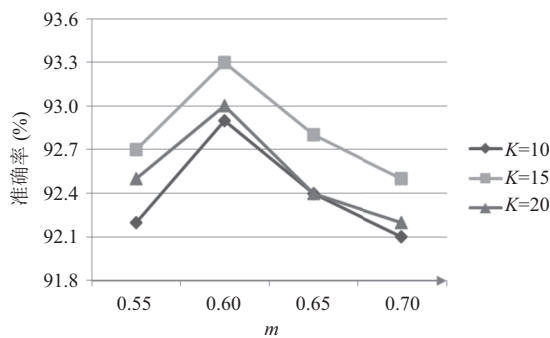


图9  $m$ 、 $K$  值寻优

图10 为计算疲劳值的权重  $w_1$ 、 $w_2$  的设定对准确率的影响,  $w_1 + w_2 = 1$ . 可以看出当只依靠眼部疲劳特征或只依靠嘴部疲劳特征时准确率均不如将两个疲劳特征融合的设置.

本文选取了三个视频测试者 (女性、男性、戴眼镜的男性) 的视频进行疲劳驾驶检测, 每名测试者分别

截取了其在正常状态和模拟疲劳状态的视频进行测试, 如图11所示. 在正常状态下测试者保持正常的说话或唱歌状态以及眨眼频率, 在模拟疲劳状态下测试者模拟打哈欠、降低眨眼频率并加深眨眼深度. 程序实时统计 30 s 单位时间测试者的 PERCLOS 值和打哈欠次数, 每间隔 3 s 计算一次疲劳值并发出预警, 如图4所示.

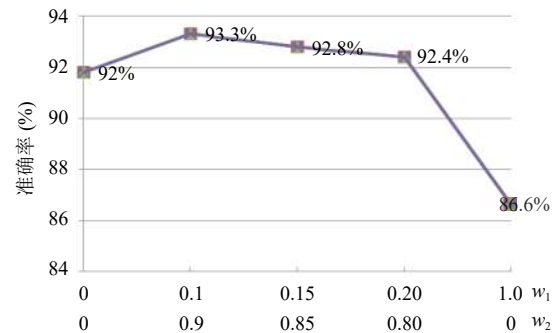
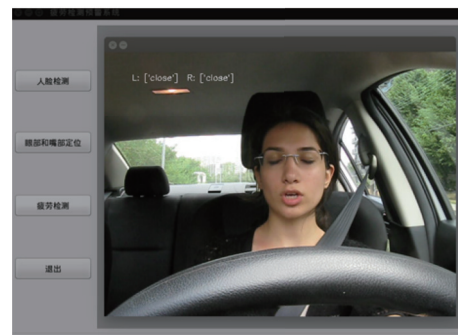


图10  $w_1$ 、 $w_2$  值寻优



(a) 正常状态示例



(b) 模拟疲劳状态示例

图11 疲劳检测示例

从表4 可以看出在测试者的正常状态下检测出的 PERCLOS 值均低于 0.4, 而疲劳状态下相反, 在 PPERCLOS 值及打哈欠次数增高时, 预警次数也随之

增多,两种状态下均未出现误判。

### 3 总结

疲劳特征的选择是疲劳驾驶检测的关键,本文提出了一种基于深度学习的疲劳驾驶检测方法。该方法首先利用 HOG 特征检测视频中的人脸图像,然后利用特征点模型实现人脸器官的定位和分割,接着使用深度卷积神经网络学习眼睛区域的特征并实现眼睛闭合

状态的识别,最后融合眼睛闭合状态与嘴部闭合状态实现疲劳驾驶检测预警。与其它的疲劳驾驶检测方法相比,该方法无需接触驾驶员,能够自主提取更有判别力的眼部特征,并且具有良好的实时性,特征的融合也进一步提高了准确性。在 ZJU、CEW 和 YawDD 数据集上与 CNN 的视觉特征的检测方法的对比试验和分析结果表明,该方法具有更好的检测性能。

表4 疲劳驾驶检测结果

视频	实际闭合帧数	检测闭合帧数	准确率(%)	PERCLOS 均值	实际哈欠次数	检测哈欠次数	预警数
femaleA	171	158	92.4	0.25	1	1	1
femaleB	379	351	92.6	0.55	3	3	5
maleA	150	141	94.0	0.22	0	0	0
maleB	347	325	93.6	0.51	2	2	3
glassesA	142	128	90.0	0.20	0	0	0
glassesB	351	313	89.2	0.49	2	2	2

### 参考文献

- Shi LC, Lu BL. EEG-based vigilance estimation using extreme learning machines. *Neurocomputing*, 2013, (102): 135–143. [doi: [10.1016/j.neucom.2012.02.041](https://doi.org/10.1016/j.neucom.2012.02.041)]
- Acharya UR, Faust O, Kannathal N, *et al.* Non-linear analysis of EEG signals at various sleep stages. *Computer Methods and Programs in Biomedicine*, 2005, 80(1): 37–45. [doi: [10.1016/j.cmpb.2005.06.011](https://doi.org/10.1016/j.cmpb.2005.06.011)]
- Liang NY, Saratchandran P, Huang GB, *et al.* Classification of mental tasks from EEG signals using extreme learning machine. *International Journal of Neural Systems*, 2006, 16(1): 29–38. [doi: [10.1142/S0129065706000482](https://doi.org/10.1142/S0129065706000482)]
- Jenke R, Peer A, Buss M. Feature extraction and selection for emotion recognition from EEG. *IEEE Transactions on Affective Computing*, 2014, 5(3): 327–339. [doi: [10.1109/TAFFC.2014.2339834](https://doi.org/10.1109/TAFFC.2014.2339834)]
- Barry RJ, De Blasio FM. EEG differences between eyes-closed and eyes-open resting remain in healthy ageing. *Biological Psychology*, 2017, (129): 293–304. [doi: [10.1016/j.biopsycho.2017.09.010](https://doi.org/10.1016/j.biopsycho.2017.09.010)]
- Suman D, Malini M, Anchuri S. EOG based vigilance monitoring system. *Proceedings of 2015 Annual IEEE India Conference*. New Delhi, India. 2016. 1–6.
- Ji Q, Yang XJ. Real-time eye, gaze, and face pose tracking for monitoring driver vigilance. *Real-Time Imaging*, 2002, 8(5): 357–377. [doi: [10.1006/rtim.2002.0279](https://doi.org/10.1006/rtim.2002.0279)]
- Baronti F, Lenzi F, Roncella R, *et al.* Distributed sensor for steering wheel rip force measurement in driver fatigue detection. *Proceedings of 2000 Design, Automation & Test in Europe Conference & Exhibition*. Nice, France. 2009. 894–897.
- Rao MK, Prakah-Asante KO. Driver behavior based vehicle application recommendation: US, 9272714 B2. [2016-03-01].
- Shankwitz CR, Baker III AR. Apparatus and method for autonomous control and balance of a vehicle and for imparting roll and yaw moments on a vehicle for test purposes: US, 9310808 B2. [2016-04-12].
- Silver A, Lewis L. Automatic identification of a vehicle driver based on driving behavior: US, 8938227 B2. [2015-01-20].
- He KM, Zhang XY, Ren SQ, *et al.* Spatial pyramid pooling in deep convolutional networks for visual recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2015, 37(9): 1904–1916. [doi: [10.1109/TPAMI.2015.2389824](https://doi.org/10.1109/TPAMI.2015.2389824)]
- Girshick R. Fast R-CNN. *Computer Science*, 2015, arXiv: 1504.08083.
- Ren SQ, He KM, Girshick R, *et al.* Faster R-CNN: Towards real-time object detection with region proposal networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, 39(6): 1137–1149. [doi: [10.1109/TPAMI.2016.2577031](https://doi.org/10.1109/TPAMI.2016.2577031)]
- Sun Y, Wang XG, Tang XO. Deeply learned face representations are sparse, selective, and robust. *Proceedings of 2015 IEEE Conference on Computer Vision and Pattern Recognition*. Boston, MA, USA. 2015. 2892–2900.
- Taigman Y, Yang M, Ranzato M, *et al.* DeepFace: Closing the gap to human-level performance in face verification.

- Proceedings of 2014 IEEE Conference on Computer Vision and Pattern Recognition. Columbus, OH, USA. 2014. 1701–1708.
- 17 Zhao Q, Ge SS, Ye M, *et al.* Learning saliency features for face detection and recognition using multi-task network. *International Journal of Social Robotics*, 2016, 8(5): 709–720. [doi: [10.1007/s12369-016-0347-x](https://doi.org/10.1007/s12369-016-0347-x)]
- 18 Naz S, Ahmed A, Mubarak QUA, *et al.* Intelligent driver safety system using fatigue detection. Proceedings of the 19th International Conference on Advanced Communication Technology. Bongpyeong, South Korea. 2017. 89–93.
- 19 Mandal B, Li LY, Wang GS, *et al.* Towards detection of bus driver fatigue based on robust visual analysis of eye state. *IEEE Transactions on Intelligent Transportation Systems*, 2017, 18(3): 545–557. [doi: [10.1109/TITS.2016.2582900](https://doi.org/10.1109/TITS.2016.2582900)]
- 20 Zhu XM, Zheng WL, Lu BL, *et al.* EOG-based drowsiness detection using convolutional neural networks. *International Joint Conference on Neural Networks*. Beijing, China. 2014. 128–134.
- 21 Yan C, Zhang BL, Coenen F. Driving posture recognition by convolutional neural networks. Proceedings of the 11th International Conference on Natural Computation. Zhangjiajie, China. 2015. 680–685.
- 22 Zhang F, Su JJ, Geng L, *et al.* Driver fatigue detection based on eye state recognition. *International Conference on Machine Vision and Information Technology*. Singapore. 2017. 105–110.
- 23 Dinges DF, Grace R. PERCLOS: A valid psychophysiological measure of alertness as assessed by psychomotor vigilance. FHWA-MCRT-98-006, 1998.
- 24 Dalal N, Triggs B. Histograms of oriented gradients for human detection. Proceedings of 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. San Diego, CA, USA. 2005. 886–893.
- 25 Kazemi V, Sullivan J. One millisecond face alignment with an ensemble of regression trees. *IEEE Conference on Computer Vision and Pattern Recognition*. Columbus, OH, USA. 2014. 1867–1874.
- 26 Pan G, Sun L, Wu ZH, *et al.* Eyeblink-based anti-spoofing in face recognition from a generic webcam. Proceedings of the 11th International Conference on Computer Vision. Rio de Janeiro, Brazil. 2007. 1–8.
- 27 Song FY, Tan XY, Chen SC, *et al.* A literature survey on robust and efficient eye localization in real-life scenarios. *Pattern Recognition*, 2013, 46(12): 3157–3173. [doi: [10.1016/j.patcog.2013.05.009](https://doi.org/10.1016/j.patcog.2013.05.009)]
- 28 Huang GB, Mattar M, Berg T, *et al.* Labeled faces in the wild: A database for studying face recognition in unconstrained environments. Technical Report 07-49, Amherst: University of Massachusetts, 2007.
- 29 Abtahi S, Omidyeganeh M, Shirmohammadi S, *et al.* YawDD: A yawning detection dataset. *ACM Multimedia Systems*. Singapore. 2014. 24–28.