

文章编号: 2095-2163(2021)10-0077-06

中图分类号: TP391.41

文献标志码: A

基于面部特征融合的驾驶员疲劳检测

廖明明, 赵波

(上海工程技术大学 机械与汽车工程学院, 上海 201620)

摘要: 针对复杂驾驶环境中光照强度、驾驶员头部姿态变化对疲劳驾驶检测精度的影响, 提出了一种基于面部特征融合和卷积神经网络的驾驶员疲劳检测方法。首先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(multitask cascade convolutional neural network, MTCNN)进行人脸检测并提取眼部和嘴部特征; 设计了眼部状态识别网络 SOE-Net 和嘴部状态识别网络 SOM-Net 分别对眼部和嘴部的状态进行识别。最后结合 PERCLOS 准则以及持续张嘴时间 (CMOT) 进行疲劳状态判定。实验结果表明, 该方法可快速准确地定位出人脸区域, 疲劳检测能够达到较高准确率, 同时能够满足实时检测的需求。

关键词: 卷积神经网络; 特征提取; 疲劳检测; 人脸检测

Driver fatigue detection based on facial feature fusion

LIAO Mingming, ZHAO Bo

(School of Mechanical and Automotive Engineering, Shanghai University of Engineering Science, Shanghai 201620, China)

[Abstract] Aiming at the influence of the light intensity and the change of the driver's head posture in the complex driving environment on the accuracy of fatigue driving detection, a driver fatigue detection method based on facial feature fusion and convolutional neural network is proposed. Firstly, the collected driver images are preprocessed, then the multitask cascade convolutional neural network (MTCNN) is used for face detection and eye and mouth features are extracted; the eye state recognition network SOE-Net and mouth state recognition network SOM-Net respectively recognize the state of the eyes and the mouth. Finally, the PERCLOS criterion and the continuous mouth opening time (CMOT) are used to determine the fatigue state. The experimental results show that the method can quickly and accurately locate the face area, the fatigue detection can reach a high accuracy rate, and it can meet the needs of real-time detection.

[Key words] convolutional neural network; feature extraction; fatigue detection; face detection

0 引言

疲劳驾驶引起的道路交通事故是造成人员伤亡的重要原因之一^[1]。因此, 研究疲劳检测方法并开发自动检测预警系统具有重要意义。疲劳检测的方法可分为: 基于生理、基于行为、基于视觉三种。相对前两种方法, 基于视觉的方法有成本低、检测精度高和非侵入的优点, 是现在疲劳检测的主流方法。视觉的疲劳检测方法包括 2 个步骤: 人脸检测、特征提取以及疲劳分析。

对于人脸检测, Viola 等人提出利用 Haar 特征和 Adaboost 训练级联分类器进行人脸检测, 但是该方法在头部姿态偏转情况下的检测精度有待提高^[2]。Zhang 等人^[3]提出了一种多任务的卷积神经网络模型 MTCNN, 将人脸检测分为 3 层网络进行训

练, 从而提高了检测的准确率和鲁棒性。对于面部特征提取以及疲劳分析, 眼部和嘴部包含丰富的信息, 且不易受外界干扰和人为因素影响。Ursulescu 等人^[4]使用眨眼检测来检测眼睛所表征的睡意, 并测量每一次连续眨眼之间的持续时间。在 Bhone 的工作中^[5], 通过计算每帧图像眼睛的纵横比 (EAR) 判断眼睛状态, 如果低于阈值, 则表明眼睛处于闭合状态。文献[6]提出一种两级神经网络的检测方法。该算法通过训练第一级网络对人眼和非人眼进行分类, 用第二级网络检测眼睛特征点的位置, 根据特征点来计算眼睛张开程度, 用 PERCLOS 进行疲劳状态评估。打哈欠被认是疲劳的突出表现之一, 大量研究致力于使用打哈欠来判断疲劳状态。Knapik 等人^[7]利用热图像和平均温度分析的方法对驾驶员打哈欠进行检测, 但是在驾驶员脸部快速

作者简介: 廖明明(1996-), 男, 硕士研究生, 主要研究方向: 图像处理、深度学习; 赵波(1963-), 女, 硕士, 副教授, 硕士生导师, 主要研究方向: 智能汽车。

通讯作者: 赵波 Email: 1131689745@qq.com

收稿日期: 2021-04-19

运动的情况下可能会检测失败。Yang 等人^[8]提出使用三维卷积网络以及双向长短期记忆网络对嘴部进行时空特征提取,采用 SoftMax 对嘴部状态进行分类,然而低分辨率的图形会降低该方法的有效性。Anund 等人^[9]首先通过视频捕捉设备获取嘴部的图像,然后输入嘴部的开合特征,神经网络系统根据嘴部开合的持续时间来判断疲劳。

为减小真实驾驶场景下光照变化、头部姿态变化等不利因素的影响。本文采用 MTCNN 进行实时人脸检测,并获取眼部和嘴部区域。在此基础上,考虑到眼部和嘴部形状以及动态特征的差异,分别设计了眼部和嘴部状态识别网络,有效地提高了检测的准确性。

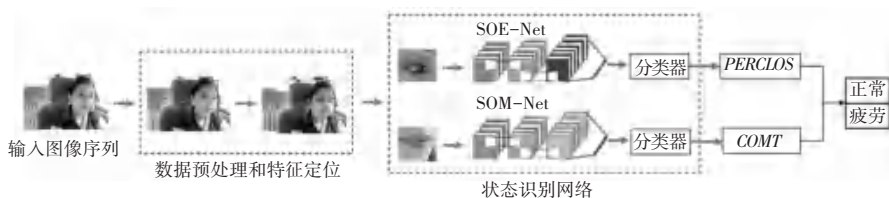


图1 算法流程

Fig. 1 The flow chart of the algorithm

2 人脸检测及特征提取

驾驶过程中驾驶室内的光照强度发生变化会影响图像的质量,导致人脸的特征变得模糊。这通常发生在某些特殊场景下(阴天、下雨天、晚上)。为了提高检测的准确性,在对驾驶员面部进行检测之前,采用光照增强的方法对图像进行预处理,使用直方图均衡化来提高图像帧的亮度以及对对比度。

相对于传统的人脸检测的算法,深度学习的算法具有精度高、鲁棒性好的优势。因此,本文采用目前比较先进的 MTCNN 来开展人脸检测的工作。MTCNN 主要包括 3 个子网络:P-Net、R-Net 和 O-Net,具体结构如图 2 所示。对此拟展开分述如下。

(1)P-Net。利用边界框的回归向量校正候选窗口,然后利用非极大抑制(non-maximum suppression, NMS)方法合并高度重叠的候选帧。此外,为了解决多尺度目标的问题,构造了图像金字塔,然后将不同大小的图像依次送入级联网络中。

(2)R-Net。主要通过边界框回归和非最大抑制消除假阳性样本。将 P-Net 生成的候选人脸框大小调整为 24×24 ,输出仍然是人脸分类和边界框回归的结果。

1 总体框架

提出的疲劳检测方法包括数据采集及预处理、人脸检测、口眼特征提取、疲劳状态识别。具体工作流程为:首先通过相机实时捕获驾驶员图像;然后将原始图像灰度化处理并进行直方图均衡化。接着使用 MTCNN 进行人脸检测,并定位眼睛和嘴部区域;然后,使用设计的 SOE-Net 和 SOM-Net 分别对眼睛和嘴部的状态进行识别;最后,根据 PERCLOS 和持续张嘴时间(COMT)对眼部和嘴部的状态进行量化分析,再结合以上 2 个疲劳判别准则对驾驶员的疲劳状况进行综合性评估。算法流程如图 1 所示。

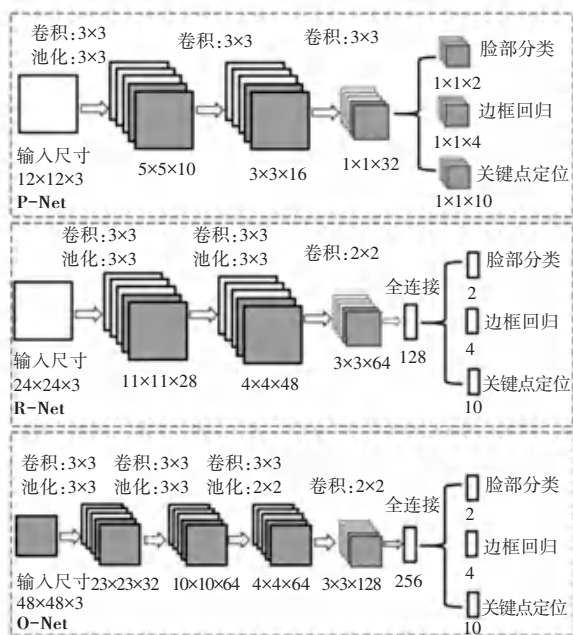


图2 MTCNN 网络

Fig. 2 MTCNN network

(3)O-Net。对经过 R-Net 滤波后的候选框区域再次滤波,将图像的大小调整为 48×48 ,并计算特征点在每一帧人脸上的位置。

通过图像的预处理极大地消除了较差光照条件

对面部特征提取的影响。MTCNN 具有很强的鲁棒性,能够适应头部姿态变化特殊的驾驶场景,其在驾驶员头部左右、上下旋转一定角度时,也能实现精准人脸定位。通过 MTCNN 网络可以得到人脸 5 个关键点,可根据关键点获得眼睛和嘴部区域。检测效果如图 3 所示。



图 3 MTCNN 检测结果

Fig. 3 Detection results of MTCNN

3 面部特征状态识别网络

3.1 眼部状态识别

通过 MTCNN 定位得到眼部区域后,需要进行眼部状态的识别。本文构建了一个眼睛状态识别网络 (SOE-Net) 对眼睛状态进行分类,网络结构如图 4(a) 所示。首先需要将输入图像的大小调整为 42×30 ;所有卷积层的卷积核大小均为 3×3 ,无填充且步长为 1;所有池化层的卷积核大小均为 2×2 ,步长为 2。池化层可以在保留主要特征的同时减小图像尺寸,并将图像尺寸减小到原始图像的一半,这极大减少了模型的参数量,可以有效防止模型过拟合。全

连接层汇集了先前层提取的特征,然后通过 Softmax 分类器实现二分类任务,即眼睛的睁开或闭合。Softmax 函数表达式为:

$$p_i = \frac{\exp(z_i)}{\sum_{k=0}^1 \exp(z_k)}, m = 0, 1 \quad (1)$$

其中, p_i 为第 i 类的概率, z_i 为全连接层的输出, z_i 公式为:

$$z_i = \sum_n x_n w_{n,i} + b_i \quad (2)$$

激活函数采用收敛速度较快 ReLU 函数,以减小反向传播过程中梯度消失问题带来的不利影响。网络使用带动量的梯度下降法进行训练。损失函数为交叉熵损失函数,具体表达式为:

$$L = - \sum [y_i \log(p_i) + (1 - y_i) \log(1 - p_i)] \quad (3)$$

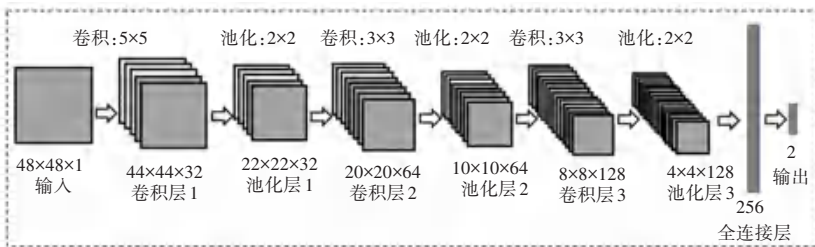
其中, y_i 为样本 i 的标签,正例为 1,负类为 0; p_i 表示样本预测为正的的概率。

3.2 嘴部状态识别

构建嘴部状态识别网络 (SOM-Net) 对提取的嘴部特征进行分类,网络结构如图 4(b) 所示。与眼部状态识别网络不同, SOM-Net 的输入图像大小是 48×48 ,并且加深了神经网络的层数,采用 3 层卷积层和 3 层池化层的网络。相同的是, SOM-Net 采用 ReLU 作为激活函数,并且仍使用交叉熵损失函数来衡量模型的质量。



(a) SOE-Net



(b) SOM-Net

图 4 面部特征识别网络

Fig. 4 Network of facial feature recognition

3.3 多指标融合的疲劳判别模型

驾驶员在疲劳的时候经常会表现出眨眼频率增

加、闭眼持续时间延长和持续张嘴等特征。在前人研究中,根据眼睛和嘴部状态来识别疲劳的常用指标

有: $PERCLOS$ ^[10]、持续闭眼时间^[11]、嘴巴张合频率^[12]、持续张嘴时间^[11]。本文使用 $PERCLOS$ 值进行眼部状态评估。 $PERCLOS$ 为 1, 表示在单位时间内持续闭眼的情况。虽然嘴巴张合频率能够一定程度反映出驾驶员的疲劳状况, 但是在说话、唱歌、大笑的情况下可能会出现误判, 而用持续张嘴时间来评估能避免这类误判。所以结合 $PERCLOS$ 和持续张嘴时间两个指标来进行疲劳判定。具体研究论述如下。

(1) $PERCLOS$ 。正常情况下, 一个司机每 60 s 平均会眨眼 10 次, 每 2 ~ 6 s 眨眼一次, 每次耗时 0.2 ~ 0.4 s。 $PERCLOS$ 是眼睛在特定时间段内闭合帧数占总帧数的比率, 能够反映驾驶员疲劳程度, $PERCLOS$ 的计算公式可表示为:

$$PERCLOS = N_c / N_t \quad (4)$$

其中, N_c 是指在单位时间内眼睛闭合状态的总帧数, N_t 是指单位时间内的总帧数。

(2) 持续张嘴时间。在疲劳的驾驶条件下, 司机经常表现出打哈欠的特征。一般情况下, 一个人打呵欠的时候, 嘴巴会张大且会保持几秒钟。提出的 SOM-Net 能判断嘴部是张开、还是闭合的状态。在打哈欠、说话和大笑时, 人的嘴部均会变为张开的状态, 但研究表明, 在正常情况下, 司机打哈欠约 5 s, 然而在说话和大笑的场景下, 人嘴部张开的的时间很短且相对于打哈欠嘴张开的的时间会很少。因此, 本文通过持续张嘴时间 ($COMT$) 来判断打哈欠, $COMT$ 表达式如下:

$$COMT = (F_{end} - F_{start}) / f \quad (5)$$

其中, F_{start} 表示开始闭合嘴部时图像帧的序号; F_{end} 表示结束闭合嘴部时图像帧的序号; f 表示采集图像的帧率。

(3) 疲劳状态检测。对于 $PERCLOS$ 的研究, Wierwille 等人^[13] 提出当进入疲劳时, $PERCLOS > 0.15$; Chu 等人^[14] 通将 $PERCLOS$ 阈值设为 0.25。本文通过实验, 为避免因强光导致的驾驶员眯眼造成对疲劳误判, 将 $PERCLOS$ 的阈值设定为 0.4, 这表示当 $PERCLOS$ 超过 0.4 的时候, 即视为驾驶员表现出疲劳的状态。对于嘴部张开持续时间, 实验中视频帧率为 20 帧/s, 每帧的时长为 50 ms, 当张嘴帧数连续超过 80 帧 (4 s) 时, 驾驶员出现打哈欠的情况。为提高疲劳检测准确性, 需要同时考虑到 $PERCLOS$ 和持续张嘴时间 ($COMT$) 两个指标, 所以提出一种疲劳判定的方案: 当 $PERCLOS \geq 0.4$ 或者 $COMT \geq 4$ s, 判定为疲劳状态; 其它情况皆为正常非疲劳状态。

4 实验结果与分析

4.1 实验平台与数据集

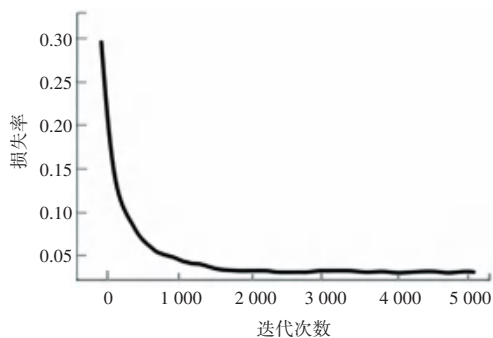
本节将介绍实验配置及结果分析, 主要包括疲劳特征识别网络性能评估和整体疲劳检测性能分析。实验的硬件环境为: Intel (R) Core (TM) i7-9800X CPU (3.80 GHz), 32 GB 内存, RTX2080 显卡, 实验在 Ubuntu 18.04 下进行。

网络训练的数据来源于自采集数据和 YawDD^[15] 数据集截取的有效部分。数据集内容包括: 清醒状态、大笑、说话、打哈欠、瞌睡的驾驶场景。

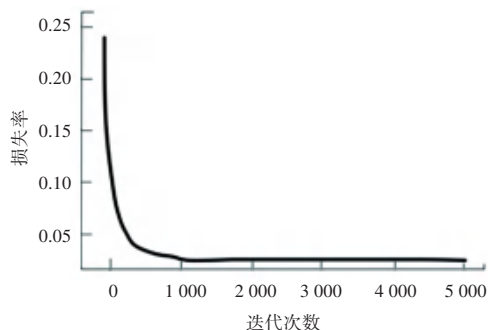
4.2 眼睛和嘴部状态识别

从数据集中获取眼睛和嘴部的图片共 16 500 张作为实验样本。训练前需要将正负样本进行灰度处理, 并将眼睛图片归一化尺寸为 42×30, 将嘴部图片归一化尺寸为 48×48。

在训练过程中, 学习率设为 0.01。为了提高评估模型的性能, 防止过拟合, 采用交叉验证法来划分数据集, 训练集和测试集的比例为 9 : 1。如图 5 (a) 所示, SOE-Net 在训练的最初损失率约为 0.29, 在经过 1 500 次迭代以后, 损失值基本上稳定在 0.02。SOM-Net 训练的过程与 SOE-Net 类似, 如图 5 (b) 所示, 在训练的最初损失率约为 0.24, 在经过 1 000 次迭代以后, 损失值基本上稳定在 0.01。



(a) SOE-Net loss



(b) SOM-Net loss

图 5 损失函数曲线

Fig. 5 Curves of loss functions

由实验结果可知, SOE-Net 在测试集的分类平均准确率高达 90.4%, 结果见表 1。SOM-Net 在平均准确率高达 90.7%, 结果见表 2。

表 1 眼睛状态识别结果

Tab. 1 The result of eye state recognition

眼睛状态	训练样本数	测试样本数	共计	准确率/%
睁眼	3 600	400	4 000	90.5
闭眼	3 600	400	4 000	90.3

表 2 嘴部状态识别结果

Tab. 2 The result of mouth state recognition

嘴部状态	训练样本数	测试样本数	共计	准确率/%
张开	4 050	450	4 500	91.1
闭合	3 600	400	4 000	90.2

4.3 疲劳检测

为进一步验证面部疲劳识别网络的性能, 使用自采数据集的 3 段视频样本进行测试。视频为实验人员在极度疲劳的状态下采集的, 用于进一步验证 SOE-Net 和 SOM-Net 的泛化能力以及实时疲劳检测的鲁棒性, 检测结果见表 3。

表 3 疲劳检测结果

Tab. 3 The result of fatigue detection

测试视频	疲劳次数	检测次数	准确率/%
视频 1	51	46	90.2
视频 2	84	77	91.6
视频 3	95	86	90.5

视频 1 中一部分疲劳检测分析如图 6 所示。由图 6 可知, 在第 80~220 帧, 由于持续张嘴的时间超过 4 s, 所以被判定为打哈欠; 而在第 395~415 帧, 持续张嘴的时间不超过 4 s, 为驾驶员说话的场景。在第 600~700 帧, PERCLOS 值达到了 0.78, 为驾驶员瞌睡的场景。

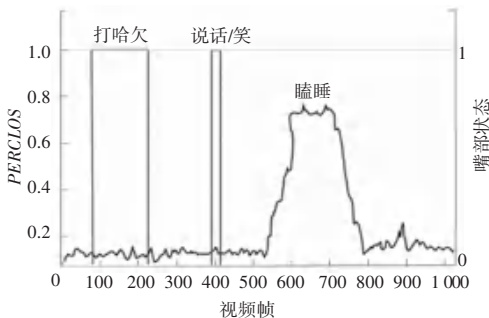


图 6 疲劳检测分析

Fig. 6 Analysis of fatigue detection

5 结束语

本文提出了一种基于卷积神经网络的疲劳驾驶

检测方法, 该方法通过融合了驾驶员眼部和嘴部的特征来判断驾驶员疲劳状况。采用 MTCNN 定位驾驶员脸部、眼部和嘴部区域; 将眼部和嘴部的特征分别输入 SOE-Net 和 SOM-Net 进行状态识别。最后, 结合 PERCLOS 和持续张嘴时间进行疲劳判定。该方法能够适应光照变化和驾驶员头部姿态变化的复杂驾驶环境。在满足实时检测的需求的前提下, 疲劳检测的平均准确率可达 90.8%。提出的神经网络模型结构简单, 便于未来在移动端部署。

参考文献

- [1] MANDAL B, LI L, WANG G S, et al. Towards detection of bus driver fatigue based on robust visual analysis of eye state[J]. IEEE Transactions on Intelligent Transportation Systems, 2017, 18(3): 545-557.
- [2] LI X, HONG L, WANG N, et al. Fatigue driving detection model based on semi-supervised learning[J]. Tongji Daxue Xuebao/Journal of Tongji University, 2018, 46:81-86,189.
- [3] ZHANG Kaipeng, ZHANG Zhanpeng, LI Zhifeng, et al. Joint face detection and alignment using multitask cascaded Convolutional Networks[J]. IEEE Signal Processing Letters, 2016, 23(10):1499-1503.
- [4] URSULESCU O, ILIE B, SIMION G. Driver drowsiness detection based on eye analysis[C]// 2018 International Symposium on Electronics and Telecommunications (ISETC). Timisoara, Romania :IEEE,2018;1-4.
- [5] BHONE R A. Computer Vision based drowsiness detection for motorized vehicles with Web Push Notifications[C]//2019 4th International Conference on Internet of Things: Smart Innovation and Usages (IoT-SIU).Ghaziabad, India :IEEE, 2019;1-4.
- [6] ZHAO Xuepeng, MENG Chunling, FENG Mingkui, et al. Eye feature point detection based on single Convolutional Neural Network[J]. Iet Computer Vision, 2017, 12(4):453-457.
- [7] KNAPIK M, CYGANEK B. Driver's fatigue recognition based on yawn detection in thermal images[J]. Neurocomputing, 2019, 338: 274-292.
- [8] YANG Hao, LIU Li, MIN Weidong, et al. Driver yawning detection based on subtle facial action recognition[J]. IEEE Transactions on Multimedia, 2020, 23:572-583.
- [9] ANUND A, FORS C, AHLSTROM C. The severity of driver fatigue in terms of line crossing: A pilot study comparing day- and night time driving in simulator[J]. European Transport Research Review, 2017, 9(2):31.
- [10] DASGUPTA A, RAHMAN D, ROUSTRAY A. A smartphone-based drowsiness detection and warning system for automotive drivers[J]. IEEE Transactions on Intelligent Transportation Systems, 2019,20(11):4045-4054.
- [11] JI Yingyu, WANG Shigang, ZHAO Yan, et al. Fatigue state detection based on multi-index fusion and state recognition network[J]. IEEE Access, 2019, 7:64136-64147.
- [12] 顾王欢, 朱煜, 陈旭东, 等. 基于多尺度池化卷积神经网络的疲劳检测方法研究[J]. 计算机应用研究, 2019, 36(11): 3471-3475.

(下转第 88 页)