

基于深度学习的驾驶员疲劳检测算法研究

郑 瀚¹

ZHENG Han

摘要

疲劳驾驶是造成交通事故频繁发生的主要原因之一，对疲劳驾驶的有效检测对于提高行车安全具有重要意义。针对驾驶环境的多样性、驾驶员的差异性以及疲劳驾驶状态特征参数的不确定性，提出一种基于深度学习的多特征融合疲劳检测算法。首先采用 ShuffleNet 网络获取复杂情况下的面部区域，解决疲劳检测存在的环境适应性差，光线不足等问题。其次获取驾驶员眼睛最大值和嘴巴最小值进行计算。最后通过信息融合，建立驾驶员疲劳状态评估模型对于驾驶员疲劳状态进行检测。实验的准确率为 98.8%，召回率为 90.2%， F -score 为 94.3%。本研究提出的深度学习疲劳检测算法，显著提升了驾驶安全性。

关键词

疲劳驾驶；ShuffleNet 网络；信息融合

doi: 10.3969/j.issn.1672-9528.2024.12.010

0 引言

随着铁路运输里程持续延展，汽车保有量亦不断攀升。汽车在为人们生活缔造快捷便利的同时，交通事故也如影随形，频繁发生。而在众多诱发交通事故的因素中，疲劳驾驶占据显著位置，是主要致因之一。据《中国统计年鉴 2021》数据显示，2021 年，中国交通事故死亡人数为 6.17 万人，中国交通事故受伤人数约为 25 万人。在交通事故中，由疲劳驾驶引起的交通事故占交通事故的 20%~30%，特别是高速上因疲劳驾驶发生的事故超过 30%。因此对疲劳驾驶状态进行检测并做出相应的预警至关重要^[1]。

目前，在驾驶员状态检测领域中，基于面部特征展开检测是最为常用的一种手段。其主要通过对驾驶员眼睛的运动状况以及头部的位置变动予以监测，进而判定疲劳程度^[2]。然而，此类算法大多忽视了驾驶员的个体特性差异，运用固定不变的阈值去判别眼睛与嘴巴的状态，显然会导致较高的误判率。

基于改进的 ShuffNet V2-16 卷积神经网络，本文设计了一种驾驶员面部特征点监测体系结构。并利用开源数据集 YawDD^[3] 对该网络进行训练。与 PFLD、Pafpif 等其他算法进行比较，改进的 ShuffNet V2-16 的算法提高了人脸识别的准确性，简化了网络机构，减少了计算时间且更便于移植到移动端。

现有的检测算法大都基于 MAR 和 EAR 的数值，利用驾驶员的眼睛和嘴巴特征点阈值作为判断特征来判断驾驶员的

疲劳与否。事实上，不同驾驶员的眼睛尺寸有别，嘴巴的开合幅度也不尽相同，对应的参数阈值也不一样。根据驾驶员实际的眼睛尺寸和嘴巴张合大小，选定前 20 帧算出眼睛比值的最大值和嘴巴张合比值的最小值并不断进行带入替换，选择适用于每一个司机的特定的疲劳阈值来判断疲劳程度，具有较高的精度。

1 系统框架

系统主要由 ShuffleNet V2k16 进行人脸关键点定位。算法流程图如图 1 所示。

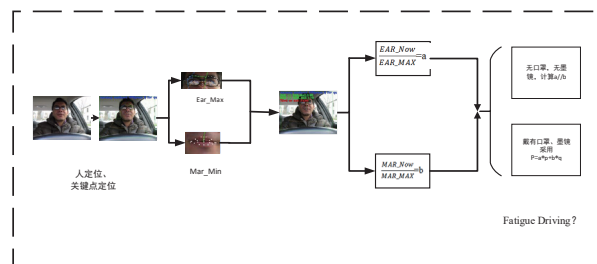


图 1 算法流程图

ShuffleNet 是一个轻量化的 CNN 模型，主要是将各部分的特征有序打乱，构成新的特征图，解决由于分组卷积带来的“信息流通不畅”问题^[5]。ShuffleNet 利用分组卷积和通道重组。在 ShuffleNet 下将各组通道进行平均分，然后依次序重新构成特征图，见图 2。图 2 表示两个通道进行重组，GConv 代表组卷积。（a）代表两个堆叠的卷积层，组数相同，每个输出通道只与组内的输入通道有关，为交叉；（b）代表将两个通道进行重新组合，输出为 68 个坐标点及背景信息，共 69 个类输出。

1. 河池学院人工智能与制造学院 广西河池 546300

[基金项目] 2024 度广西高校中青年教师基础能力提升项目“驾驶员情绪识别的关键技术研究”（2024KY0629）

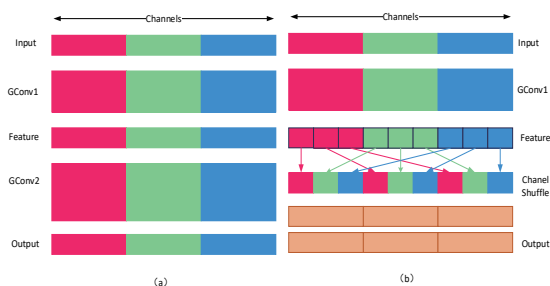


图2 结构特征图

改进的 ShuffleNet V2 相比其他模型速度较快而且其准确性有所提高。引入了新的运算 Channel Split。ShuffleNet V2 输出通道与输入到通道相等时, MAC 消耗最小; 当组卷积的组增大时, MAC 也会变大; 网络的碎片化程度越高, 速度越慢; 元素之间的操作减少。网络结构图如图3所示。

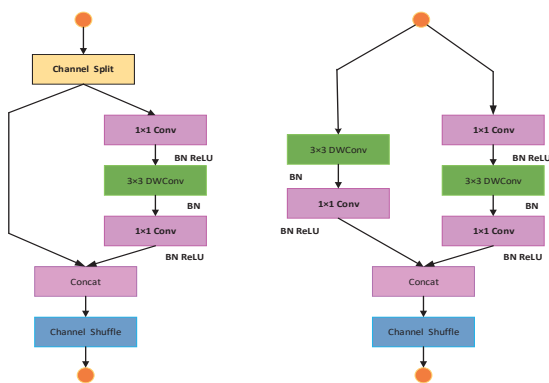


图3 ShuffleNet V2 结构图

2 疲劳状态检测

在疲劳检测中, 通过驾驶员眼睛和嘴巴的状态来判断驾驶员的疲劳程度^[6]。通过定位人脸结构来获取面部特征, 具体图像如图4所示。

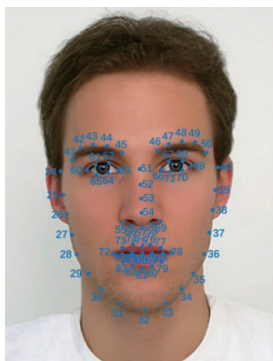


图4 面部特征点定位

2.1 眼睛状态评价指标

在眼睛睁开和闭合时, 眼睛上下特征点之间的距离会发生一定的变化。EAR 是通过使用眼睛特征点之间的相对距离来计算获得。从图4中可以看出左右眼睛点的坐标分别为60~65和66~71。左眼 EAR 计算公式为:

$$EAR = \frac{\|P_{61} - P_{65}\| + \|P_{62} - P_{64}\|}{2\|P_{60} - P_{63}\|} \quad (1)$$

右眼睛的计算公式类似。

2.2 嘴巴状态评价指标

当驾驶员打哈欠时, 嘴巴的张合类似于眼睛的张合, 有学者使用外唇的关键点进行检测^[7]。通过对嘴巴特征点之间的相对距离来计算获得 MAR 值。从图4可以看出嘴巴的关键点为72~88。MAR 计算公式为:

$$MAR = \frac{\|P_{73} - P_{83}\| + \|P_{77} - P_{79}\|}{2\|P_{72} - P_{78}\|} \quad (2)$$

式中: 分子部分表示为垂直特征点之间的欧式距离, 分母分标表示为眼睛和嘴巴水平特征点之间的欧式距离。

2.3 多重指标融合

实际驾驶环境中, 疲劳驾驶是一种复杂的心理和生理状况。检测结果容易受到各种环境的干扰。因此在本文中提出一种 MAX-MIN 算法。该算法是基于 EAR 和 MAR 公式的基础上提出的新的计算方式。在实际情况中, 驾驶员的眼睛和嘴巴因为个人的差异而有所区别, 在此前的研究中, 通常是将 EAR 和 MAR 确定一个固定的值来判断驾驶员是否疲劳而忽视了驾驶员个体的差异。每一位驾驶员的眼睛和嘴巴的大小并不相同^[8]。因此 EAR 和 MAR 的值也会存在一定的差异。在本文中采用数据集每段视频的前 100 帧获取 EAR_Max 和 MAR_Min 的值。并将后续每帧图像的 EAR 和 MAR 的值与 EAR_Max 和 MAR_Min 进行比值并赋予新值 a 和 b 。在驾驶员行驶过程中, a 和 b 是动态调整的。若 $a < a_{Th}$ (a_{Th} 为驾驶员疲劳眼睛状态判断阈值) 或 $a > a_{Th}$ (b_{Th} 为驾驶员嘴巴疲劳状态判断阈值) 则会判断该驾驶员为疲劳的。

$$a = \frac{EAR_Now}{EAR_Max} \quad (3)$$

$$b = \frac{MAR_Min}{MAR_Now} \quad (4)$$

3 实验

为了验证算法的有效性。本文利用公开数据集对改进后的 MMPose 网络的性能进行了评价。在此基础上, 进行设计对比实验, 验证基于人脸多特征融合的疲劳驾驶检测算法的准确性。

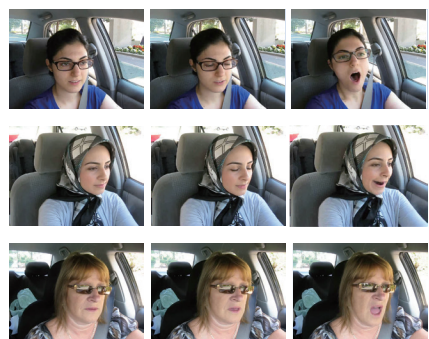
3.1 数据集

为了验证所提出的 MAX-MIN 算法在驾驶员疲劳检测任务中的有效性。本文主要使用两类数据集进行实验。第一类选择了公开的 YawDD 视频数据集, 用于眼睛和嘴巴检测的数据集。用来测试驾驶员闭眼和打哈欠的算法和模型。其中包含了多名驾驶员在驾驶位上的视频。视频分为两部分, 一部分是摄像头位于车内后视镜下方, 测对着驾驶员录制得到的; 另一部分是摄像头位于仪表盘上方, 正对着驾驶员,

采集到驾驶员的正面图像，每秒采集 30 帧。每个驾驶员录制 3~4 段视频。此数据集用于验证模型的泛化性。本文取正面部分数据集。在这个数据集中，有 322 个不同种族的男女司机戴和不戴眼镜以及墨镜的视频。YawDD 中的一些数据集中的样本见图 5。在视频中，驾驶员有打哈欠、闭眼睛、说话等实际情况。随机选取了 30 秒钟的模拟疲劳视频进行了检测。



(a) 男性驾驶员图像



(b) 女性驾驶员图像

图 5 YawDD 数据集。

3.2 目标检测的训练及评价指标

以 YawDD 数据集为训练数据集。训练过程迭代 50 次，前 20 次的学习率设置为 0.01，后 20 次的学习率设置为 0.001。实验中选择的评价指标包括：Precision、Recall、F-score。

(1) Precision

精度用来衡量检测驾驶员疲劳的准确性，公式为：

$$\text{Precision} = \frac{TP}{TP + FP} \quad (5)$$

式中：TP 表示驾驶员疲劳时被正确检测出来的次数；FP 表示驾驶员未处于疲劳状态但被误判为疲劳的次数。

(2) Recall

召回率是测量疲劳检测的查全率。也就是显示系统驾驶员疲劳检测的漏检率。

$$\text{Recall} = \frac{TP}{TP + FN} \quad (6)$$

式中：FN 表示驾驶员处于疲劳状态但系统错误的判断为非疲劳状态的次数。

(3) F-score

F-score 是综合指标，为了平衡准确率和召回率的影响，

较为全面的评价系统对于驾驶员疲劳检测的效果，是精确率和召回率的调和平均，F-score 越高，性能越好。

$$F\text{-score} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (7)$$

3.3 疲劳检测实验

疲劳检测中使用 YawDD 视频作为实验数据。每个视频包括闭眼打哈欠等疲劳状态。检测结果如图 6 所示。分别采用 EAR_Max 和 MAR_Min 作为评价指标来判断疲劳状态。为了验证提出的基于个体差异的疲劳检测算法有效性，进行了一系列的实验。通过不同驾驶人进行驾驶过程中眨眼睛和打哈欠时眼睛和嘴巴的特征值来判断是否疲劳^[9]。

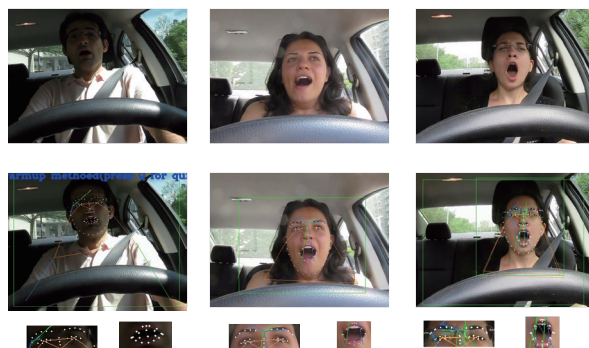


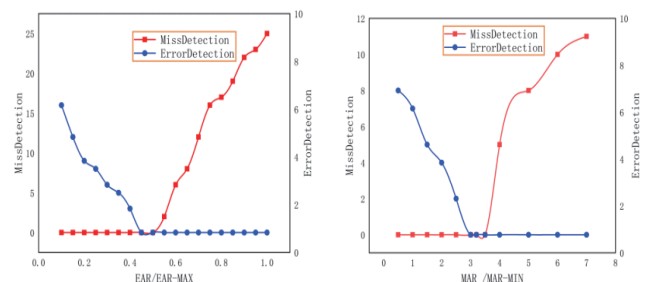
图 6 人脸定位和关键点检测及眼睛和嘴巴区域检测

3.3.1 MAX-MIN 阈值设置

本实验选取了 12 段 YawDD 中驾驶员视频，其中包括了男性驾驶员和女性驾驶员，也包括了佩戴眼镜和未佩戴眼镜的驾驶员状况。以 12 段驾驶员在驾驶环境中获取的视频作为设置 MAX-MIN 阈值的依据。这些视频包括了清醒和疲劳的状态。

另外，为了使 MAX-MIN 的阈值达到最佳状态，设置了两个参数进行优化，分别为：未检测和误检测。其中，未检测表示驾驶员处于疲劳状态但是系统未检出；误检测表示驾驶员未处于疲劳状态但系统误认为驾驶员为疲劳^[10]。

从图 7 可以看出，当 EAR/EAR-MAX 的阈值为 0.5 时，驾驶员疲劳检测的误检率和未检率最低。当 MAR/MAR-MIN 为 3 或 4 时驾驶员疲劳检测的误检率和未检率最低。



(a) 不同 EAR/EAR-MAX 下 错检和漏检图像 (b) 不同 MAR/MAR-MIN 下错检和漏检图像

图 7 不同比值的错检和漏检实验

驾驶员实际驾驶时,相比于眼睛频繁张开闭合而言嘴巴张开和闭合频率会较低。因此在实际实验中眼睛的检测会比嘴巴有更多的干扰。因此眼睛检测的误检率也会明显高于嘴巴。为了避免眼睛检测中出现较多的误检和漏检,将EAR/EAR-MAX的阈值设置的较低,MAR/MAR-MIN设置的阈值较高。综合考虑,设置阈值方案为:

(1) 当EAR/EAR-MAX的值小于0.5时,判定驾驶员处于疲劳驾驶状态。

(2) 当MAR/MAR-MIN的值大于3.2时,判定驾驶员处于疲劳驾驶状态。

为了验证方案的准确性,用评价指标精确度、召回率和F-score来衡量。结果表明,准确率为98.8%、召回率90.2%、F-score为94.3%。

3.3.2 疲劳检测对比实验

为了充分证明本文提出的MAX-MIN的有效性,在相同的实验环境下,进行了两组对比实验:(1)第一组是将采用EAR和MAR进行驾驶员疲劳检测;(2)第二组对比实验是采用MAX-MIN算法,分别获取EAR/EAR-MAX以及MAR/MAR-MIN的值,通过MAX-MIN的值取得驾驶员疲劳检测的阈值从而来判断驾驶员是否疲劳。其中,所涉及的方法都在同一个数据集上进行训练,并从Precision、Recall、F-score三个方面进行评估。

4 结论

本文旨在研究驾驶员疲劳检测的方法,并采用三种不同网络进行了对比。采用ResNet50、Vgg以及改进的ShuffleNet V2k16对于数据集进行人脸进行检测并对于特征进行定位。通过对比最终采用ShuffleNet V2k16对于人脸检测及特殊点定位。此外,本文设计了MAX-MIN算法用于驾驶员脸部嘴巴和眼睛特征点的计算。在YawDD中进行了测试,并在自建数据集(SBD)上进行了测试。实验的准确率为98.8%、召回率90.2%、F-score为94.3%。因为驾驶员眼睛和嘴巴的大小因为个体差异所以存在阈值无法统一的问题。为了解决以上问题。本文提出了一种MAX-MIN算法,利用眼睛和嘴巴的特征进行计算。有效提高驾驶疲劳检测的准确性。实验表明所提出的算法能够在不同的驾驶工况下大大提高驾驶疲劳的检测精度。在未来的研究工作中,将重点关注以下几个方面:(1)将以上研究安装在车上,进一步研究类似夜间行驶条件下的识别效果;(2)研究关于驾驶员头部姿势偏转情况下识别效果的提高;(3)疲劳检测是一个实时的过程,需要较好的处理速度。因此,优化网络结构也是未来研究的重点。

参考文献:

- [1] 中华人民共和国国家统计局.中国统计年鉴2021[M].北京:中国统计出版社,2021.
- [2] HE J Y, CHEN J Q, LIU J, et al. A lightweight architecture for driver status monitoring via convolutional neural networks[C]//2019 IEEE International Conference on Robotics and Biomimetics (ROBIO). Piscataway:IEEE, 2019: 388-392.
- [3] ABTAHI S, OMIDYEGANEH M, SHIRMOHAMMADI S, et al. YawDD: a yawning detection dataset[C]//Proceedings of the 5th ACM Multimedia Systems Conference. New York: ACM, 2014: 24-28.
- [4] 杨巨成,魏峰,林亮,等.驾驶员疲劳驾驶检测研究综述[J].山东大学学报(工学版),2024,54(2):1-12.
- [5] MA N N, ZHANG X Y, ZHENG H T, et al. ShuffleNet V2: practical guidelines for efficient CNN architecture design[C]//European Conference on Computer Vision. Berlin:Springer, 2018:122-138.
- [6] 林梅,徐佳富,王宇辰,等.基于人工智能的汽车驾驶员精神状态监测系统研究[J].科技创新与应用,2023,13(25):27-32.
- [7] 蒋启超,余成波,宣以国,等.基于轻量级主干的YOLOv5驾驶员疲劳检测算法[J].科学技术与工程,2024,24(16): 6766-6774.
- [8] 徐文奇,胡耀聪.基于面部多特征跨层融合网络的驾驶员疲劳检测方法[J].安徽工程大学学报,2023,38(6):64-71.
- [9] SUGANIYA M, JERRITTA S, ARUN S. Detection and analysis: driver state with electrocardiogram (ECG)[J]. Physical and engineering sciences in medicine, 2020(43):525-537.
- [10] 马明佳.车辆驾驶员的疲劳状态检测与缓解[D].吉林:东北电力大学,2024.

【作者简介】

郑瀚(1987—),女,江苏兴化人,硕士研究生,副教授,研究方向:人工智能、智慧交通。

(收稿日期:2024-09-06)