

基于并行短时面部特征的驾驶人疲劳检测方法研究*

刘强¹ 谢谦¹ 方玺² 李波³ 解孝民⁴

(1.中山大学,智能工程学院,深圳 518107;2.国家邮政局发展研究中心,北京 100868;3.广州汽车集团股份有限公司汽车工程研究院,广州 511434;4.广东玛西尔电动科技有限公司,肇庆 523268)

【摘要】为实现更快速、准确的疲劳预警,提出了一种基于并行短时面部特征的驾驶人疲劳检测方法。基于加入了MicroNet模块、CA注意力机制、Wise-IoU损失函数的YOLOv7-MCW目标检测网络提取驾驶人面部的短时面部特征,再使用并行Informer时序预测网络整合YOLOv7-MCW目标检测网络得到的面部时空信息,对驾驶人疲劳状态进行检测与预警。结果表明:在领域内公开数据集UTA-RLDD和NTHU-DDD上,YOLOv7-MCW-Informer模型的准确率分别为97.50%和94.48%,单帧检测时间降低至28 ms,证明该模型具有良好的实时疲劳检测性能。

关键词:智能交通 疲劳检测 目标检测 注意力机制 时序预测

中图分类号:U492.8*4 **文献标志码:**A **DOI:**10.19620/j.cnki.1000-3703.20230617

Research on Driver Fatigue Detection Method Based on Parallel Short-Term Facial Features

Liu Qiang¹, Xie Qian¹, Fang Xi², Li Bo³, Xie Xiaomin⁴

(1. School of Intelligent Systems Engineering, Sun Yat-sen University, Shenzhen 518107; 2. Development & Research Center of State Post Bureau, Beijing 100868; 3. Automobile Engineering Research Institute of Guangzhou Automobile Group Co., Ltd., Guangzhou 511434; 4. Guangdong Marshall Electric Technology Co., Ltd., Zhaoqing 523268)

【Abstract】A driver fatigue detection method based on parallel short-term facial features is proposed to achieve faster and more accurate fatigue warning. The method utilizes the YOLOv7-MCW object detection network, which incorporates the MicroNet module, CA attention mechanism, and Wise-IoU loss function, to extract short-term facial features of the driver's face. The parallel Informer temporal prediction network is then used to integrate the spatiotemporal information obtained from the YOLOv7-MCW object detection network, enabling the detection and warning of driver fatigue. The results demonstrate that the YOLOv7-MCW-Informer model achieves accuracy rates of 97.50% and 94.48% on the publicly available datasets UTA-RLDD and NTHU-DDD, respectively, with a single-frame detection time reduced to 28 ms, proving the excellent real-time fatigue detection performance of the model.

Key words: Intelligent transportation, Fatigue detection, Object detection, Attention mechanism, Time series prediction

【引用格式】刘强,谢谦,方玺,等.基于并行短时面部特征的驾驶人疲劳检测方法研究[J].汽车技术,2024(5):15-21.

LIU Q, XIE Q, FANG X, et al. Research on Driver Fatigue Detection Method Based on Parallel Short-Term Facial Features [J]. Automobile Technology, 2024(5): 15-21.

1 前言

基于驾驶人面部特征的疲劳检测方法因具有快速、准确的优点被广泛用于交通安全研究。驾驶人的面部特征主要包括单位时间内闭眼百分比(Percentage of

Eyelid Closure Over Time, PERCLOS)^[1]、眨眼频率、视线方向、单位时间内张口百分比(Percentage of Mouth Open Over the Pupil over Time, POM)^[2]、哈欠频率、点头次数和头部偏转角等。在较短的单位时间(一般为1 min)内具有较为明显的变化规律的面部特征,本文称为短时

*基金项目:重庆市科技创新重大研发项目(CSTB2023TIAD-STX0030),广东省重点领域研发计划项目(2022B0701180001),广东省基础与应用基础研究基金项目(2022A1515010692),肇庆市第四批西江创新创业团队与领军人才项目。

通信作者:刘强(1981—),男,江西赣州人,教授,博士,研究方向为交通安全,liuq32@mail.sysu.edu.cn。

面部特征,如PERCLOS、POM等。

国内外围绕基于短时面部特征的驾驶人疲劳检测展开了相关研究。Bai等^[3]提出使用双流时空图卷积网络检测驾驶人疲劳,采用面部标志检测法从实时视频中提取驾驶人面部标志,然后通过双流时空图卷积网络得到驾驶人疲劳检测结果,试验表明,该方法显著提高了疲劳检测性能,准确率高达92.70%,但该模型结构较为复杂,检测时间较长,实时性不足。姜平等^[4]通过改进的多任务卷积神经网络模型定位人脸区域并截取眼部、嘴部图像,再通过AlexNet图像分类网络对眼、嘴状态进行分类,并基于PERCLOS和POM判定疲劳状态。该模型的准确率达93.50%,但该方法仅捕获驾驶员面部空间特征,在处理驾驶员在不同驾驶阶段、驾驶环境下的短时面部特征规律差异上存在不足。Tamanani等^[5]使用基于Haar特征的Cascade分类器,从输入的视频流中截取人脸并捕获眼睛、嘴等面部特征,并使用LeNet-5模型进行二分类来确定驾驶人的疲劳状态,该模型的准确率达91.80%,但该模型较少考虑并行检测中眼部特征与嘴部特征存在的规律差异且检测速度较慢,模型的鲁棒性与实时性存在提升空间。

综上所述,当前基于短时面部特征的驾驶人疲劳驾驶检测方法的模型实时性能有待改进,融合时空特征的疲劳检测研究较少,关于面部多特征的并行检测有待深入研究。因此,本文将基于YOLO(You Only Look Once) v7-MCW (Micro-Net Coordinate Attention Wise-IoU)-Informer模型,深入探究基于并行短时面部特征的驾驶人疲劳检测问题。

2 短时面部特征提取

2.1 YOLOv7模型

YOLOv7^[6]在速度和准确性方面具有极佳表现,并提供了对边缘计算设备的良好支持。YOLOv7网络主要包含主干网络(Backbone)、颈部(Neck)、头部(Head)3个部分。

2.2 改进后的YOLOv7检测器

本文针对YOLOv7进行适应性改进以在保证良好准确性的基础上提升单帧检测速度。主要改进内容包括轻量化主干网络、注意力机制以及损失函数。

2.2.1 轻量化主干网络

为实现面部特征快速捕获,需使用轻量化网络结构重构Backbone部分。在处理极低计算成本问题时,主流轻量化网络结构MobileNet和ShuffleNet等存在较为严重的性能下降,因此引入微网络(MicroNet)^[7]提升极

低计算成本下的网络性能。MicroNet建立在微分解卷积(Micro-Factorized Convolution, MFC)模块和动态最大化(Dynamic Shift-Max, DSM)激活函数的基础上。MFC模块通过在点卷积和深度可分离卷积上的低秩近似值来实现通道数和输入输出连接之间的平衡,DSM激活函数则动态地融合了连续的通道组,增强了节点的连接性和非线性,以弥补主干网络深度的减少。

2.2.2 注意力机制

目前,轻量级网络的注意力机制大多采用通道注意力,仅考虑了通道间的信息,忽略了位置信息。尽管后来的瓶颈注意模块和卷积块注意模块尝试在降低通道数后通过卷积提取位置注意力信息,但卷积只能提取局部关系,缺乏长距离关系提取的能力。因此,引入高效坐标注意力机制(Coordinate Attention, CA)^[8]模块,其结构如图1所示,其中C、W和H分别为通道的数量、宽度和高度。

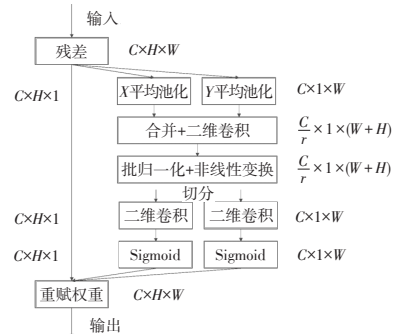


图1 坐标注意力模块结构

2.2.3 损失函数

边界框回归(Bounding Boxes Regression, BBR)损失函数对于目标检测至关重要。

引入明智交并比(Wise Intersection over Union, Wise-IoU)v3损失函数^[9],该BBR损失函数具有动态非单调静态聚焦机制(Focusing Mechanism, FM):

$$L_{wv3} = rR_{wIoU} L_{IoU}, r = \frac{\beta}{\delta\alpha^{\beta-\delta}} \quad (1)$$

$$R_{IoU} = \exp\left(\frac{(x-x_{gt})^2 + (y-y_{gt})^2}{(W_g^2 + H_g^2)^*}\right) \quad (2)$$

式中: L_{wv3} 为 Wise-IoU v3 损失函数, L_{IoU} 为交并比损失函数, R_{wIoU} 为正则惩罚项, β 为锚框的离群度, r 为梯度增益, α 、 δ 为超参数, R_{IoU} 为交并比正则惩罚项, x 、 y 为边界框的坐标值, x_{gt} 、 y_{gt} 为目标框的坐标, W_g 、 H_g 为最封闭几何框的宽高。

当 $\beta = \delta$ 时, δ 使 $r=1$ 。当 $\beta=C$ (C 为常数) 时, 锚框将获得最高的梯度增益。

2.2.4 整体结构

图像经过数据增强等预处理后,进入基于MicroNet汽车技术

(包含 MicroBlock-A 与 MicroBlock-B)、扩展高效层聚合网络(Extended Efficient Layer Aggregation Networks, E-ELAN)模块以及空间金字塔池化和全连接空间金字塔卷积(Spatial Pyramid Pooling and Fully Connected Spatial Pyramid Convolution, SPPCSPC)模块组合而成的主干网络。E-ELAN 模块基于原始 ELAN 结构,改变原始计算模块的同时保持过渡层结构,并利用扩张、混洗、合并基数的思想来增强网络学习的能力,而不破坏原有

的梯度路径。SPPCSPC 模块在一串卷积中加入并行的多次最大池化操作,可避免由于图像处理操作所造成的图像失真等问题,同时也解决了卷积神经网络提取到重复特征的难题;主干提取到的特征经过 CA 模块再次提取,再经过 Neck 模块特征融合处理得到大、中、小 3 种尺寸的特征;最终,融合后的特征被送入头部网络,经过检测之后输出结果。YOLOv7-MCW 模型的网络结构如图 2 所示。

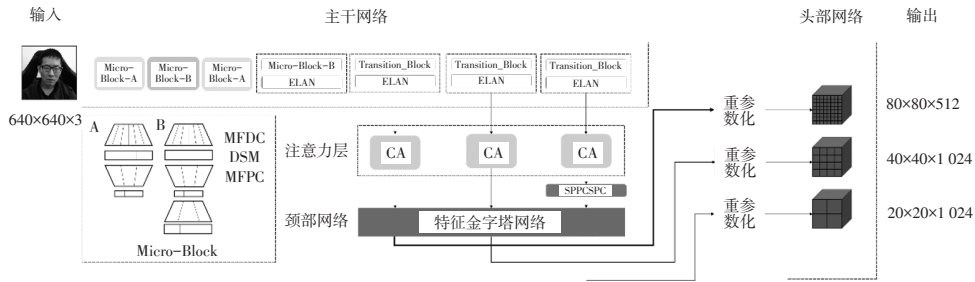


图2 YOLOv7-MCW 网络结构

2.3 特征选取

模拟驾驶系统被认为是研究驾驶行为的重要工具^[10]。本文通过模拟驾驶试验得到了具有明显规律的驾驶员短时间面部特征。为更好地模拟真实驾驶工况,本试验搭建了具备环绕屏幕、转向盘、油门踏板、制动踏板、离合器踏板、换挡操纵杆、可调节座椅、透明封闭舱、采集摄像头的模拟驾驶室,并按照真实车型设定了模拟驾驶系统的加速度、可视范围、反馈力度等相关参数,以保证接近真实的驾驶体验。试验招募的被试人员均为驾龄超过 3 年、具备丰富驾驶经验的人员,并有过疲劳驾驶的经历。另外,试验选用了更易产生疲劳的平直高速公路,驾驶员在身体健康、睡眠充足、精神状态良好的情况下开始驾驶,中途开始产生驾驶疲劳,直至主观感到过于疲劳不能驾驶时试验停止。如图 3 所示,试验共 20 组,20 位驾驶员的年龄构成包括:8 位 20~30 岁的学生

和 12 位 30~50 岁的公司职员、社会人士;性别构成包括 10 位男性和 10 位女性。试验选用简单场景,驾驶环境参考了北京某试验场的环形道路,其中直道长度为 100 km,模拟驾驶车速限制为最高 100 km/h,样本可以保持长达 1 h 的直线行驶。所选短时间面部特征为闭眼百分比、张口百分比、最长单次闭眼时长、打哈欠频率,在试验过程中部分样本的参数记录如表 1 所示。

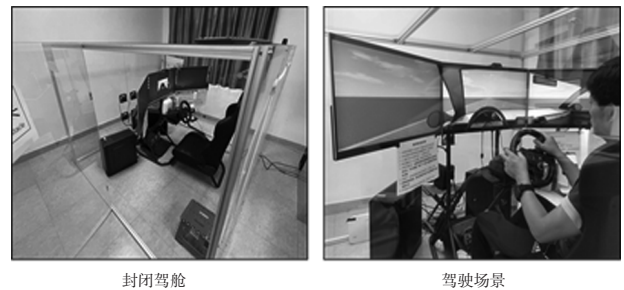


图3 驾驶试验模拟场景

表1 部分样本参数

试验编号	闭眼百分比最大值/%	闭眼百分比最大值时段/min	张口百分比最大值/%	张口百分比最大值时段/min	单次持续闭眼最长时间/s	单次最长持续闭眼时段/min	打哈欠频率最高时段/min	哈欠最高频次/次·min ⁻¹
1	32	38.00	10	37.00	4.20	38.00	37.00	4.00
2	61	41.00	9	22.00	5.34	25.00	27.00	4.00
3	73	41.00	15	52.00	6.00	37.00	42.00	4.00
4	63	37.00	2	43.00	5.40	32.00	43.00	4.00
5	55	32.00	12	18.00	5.12	34.00	18.00	5.00
6	74	36.00	7	35.00	6.04	40.00	36.00	7.00
7	72	24.00	30	13.00	7.94	22.00	13.00	7.00

2.4 眼部判断

卡内基-梅隆研究中心 Wierwille 提出了“闭眼百分比”(PERCLOS)的概念,用于衡量人类疲劳状态(嗜

睡),被定义为单位时间内眼睛闭上的时间。美国联邦公路管理局和国家公路交通安全管理局在实验室中开展模拟驾驶,验证了 PERCLOS 在描述驾驶人疲劳方面

的有效性,PERCLOS是基于面部特征的检测方法中最准确的指标之一^[11],共有P70、P80、EM3种测量标准。本文采用P80标准,该标准是指在一定时间内,当80%以上的瞳孔被眼睑覆盖时,眼睛闭合的时间比例,时间尺度为分钟级。PERCLOS计算公式为:

$$P_{PER} = \frac{\sum_i^N f_i}{N} \times 100\% \quad (3)$$

式中: i 为视频帧顺序, f_i 为第 i 帧眼睛闭合的帧数, $\sum_i^N f_i$ 为单位时间内眼睛闭合的总帧数, N 为单位时间内视频总帧数。

最长单次闭眼时长 E_M 也是衡量眼部状态的重要参数,时间尺度为秒级。通过对眼睛连续闭合的视频帧数进行计数,根据视频帧率,可以得到最长单次闭眼时长:

$$E_M = g_i \times F_v \quad (4)$$

式中: g_i 为眼睛闭合的帧数, F_v 为视频帧率。

2.5 嘴部判断

单位时间内张口百分比(Percentage of Mouth Open Over the Pupil over Time, POM)计算公式为:

$$P_{POM} = \frac{\sum_i^N h_i}{N} \times 100\% \quad (5)$$

式中: h_i 为第 i 帧嘴张开的帧数, $\sum_i^N h_i$ 为单位时间内张口的总帧数。

打哈欠频率 N_Y 是衡量嘴部状态的重要参数。POM统计一定时间内,驾驶员张口时间所占比例,但除打哈欠外,驾驶员还可能存在说话等其他张口行为; N_Y 统计一定时间内,驾驶员打哈欠的次数,强调打哈欠这一具体行为。区分这两个指标,有助于细化嘴部行为,提升准确性。通过对嘴巴连续张开的视频帧数进行计数,可以得到打哈欠的频率:

$$N_Y = \frac{F_Y}{T} \quad (6)$$

式中: F_Y 为单位时间内打哈欠的总次数, T 为单位时间。

2.6 模型训练

为增加数据准确性,进行了15组单次时长为1~2h的驾驶模拟试验,并通过摄像头采集驾驶人面部图像。从这些图像中,截取了15 340个不同状态,并进行标注,制作为短时疲劳检测面部图片数据集(Short-time Fatigue Driving Detection Image Dataset, SFDDID),用来训练YOLOv7-MCW模型,如图4所示。

2.7 初始化

本文通过驾驶模拟试验来检验模型检测效果并

得到相应参数的初始范围,为模型提供较好先验。结果显示,驾驶员在疲劳时,闭眼百分比、张口百分比、最长闭眼时间和打哈欠频率存在明显变化,验证了本文选取规律的科学性,并作为模型训练初始值,如表2所示。



图4 模拟驾驶试验场景

表2 不同状态下的参数范围初始值

参数	清醒	轻度疲劳	重度疲劳
PERCLOS/%	0~7.5	>7.5~15	>15
POM/%	0~5	>5~10	>10
E_M/s	0~2	>2~4	>4
$N_Y/次 \cdot min^{-1}$	0~1	>1~3	>3

3 疲劳检测

本文疲劳检测主要基于Informer^[12]框架实现,通过对目标检测网络得到的面部数据进行分析,对驾驶员的疲劳状态进行判别。时序预测算法运用到疲劳驾驶检测上,可使分类预测结果结合时间特性,更符合疲劳发生的过程^[13]。

Informer模型整体上由编码器(Encoder)和解码器(Decoder)两个部分组成,如图5所示。其中,编码器用于捕获长序列输入的内部依赖关系,解码器进一步实现序列检测。本文先利用检测器检测到驾驶人的面部空间信息,再输送进Informer中分析时序数据,模型学习到面部时空特征之间的依赖关系,并基于这些关系来检测驾驶人是否处于疲劳状态。

眼部特征与嘴部特征的时空变化存在差异,使用两组Informer模型进行组合,分别为Informer-A和Informer-B,二者具有不同的权重,分别处理眼部特征与嘴部特征,形成并行结构。

将眼部特征指标、嘴部特征指标分别输入,并行计算。 P_{PER} 、 E_M 、 P_{POM} 以及 N_Y 4个指标分别作为并行Informer网络的4个时间序列输入,每个时间序列输入包括多个时间步的指标值。将Informer组合网络的输出作为疲劳状态的预测概率值。

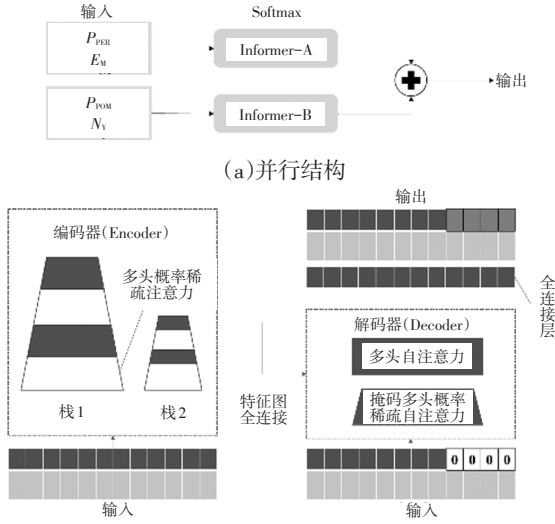
3.1 数据标准化

通过目标检测网络定位驾驶员的面部特征后,计算

P_{PER} 、 E_M 、 P_{POM} 以及 N_V 并对这4个参数进行归一化,本文采用最大最小归一化:

$$s' = \frac{s_i - s_{\min}}{s_{\max} - s_{\min}} \quad (7)$$

式中: s' 为指标的最终得分, s_i 为实时检测到的指标的得分, s_{\min} 、 s_{\max} 分别为检测过程中疲劳指标的最小值和最大值。



(b)Informer
图5 模型框架

3.2 分类器

在原有编码器至解码器的线路外增加去时序空间特征辅助分类器,由1个输入层、40个残差卷积层、1个平滑层、2个全连接层构成,用于跳过时序辅助进行空间状态分类,以增加准确性,其结构如图6所示。

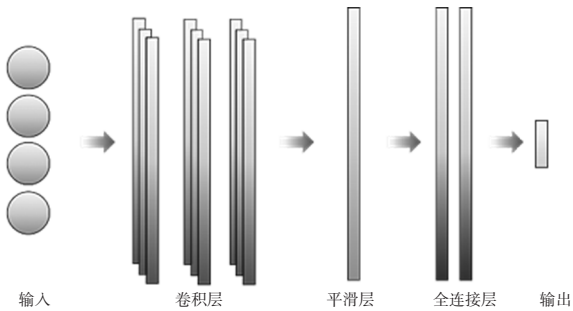


图6 辅助分类器网络结构

辅助分类器中每个隐含层的激活函数都使用线性修正单元(Rectified Linear Unit, ReLU),输出层的激活函数使用Sigmoid函数,将输出值映射到0~1的范围内。使用随机梯度下降(Stochastic Gradient Descent, SGD)优化器,迭代50次,批尺寸设置为32,初始学习率设置为0.001。模型整体检测流程如图7所示。

4 试验验证

4.1 数据集

本文选用2个公开数据集,德州大学阿灵顿分校真
2024年 第5期

实生活瞌睡数据集(University of Texas at Arlington Real-Life Drowsiness Dataset,UTA-RLDD)^[14]和台湾清华大学计算机视觉实验室的驾驶员疲劳检测数据集(National Tsing Hua University Drowsy Driver Dataset, NTHU-DDD)^[15]。

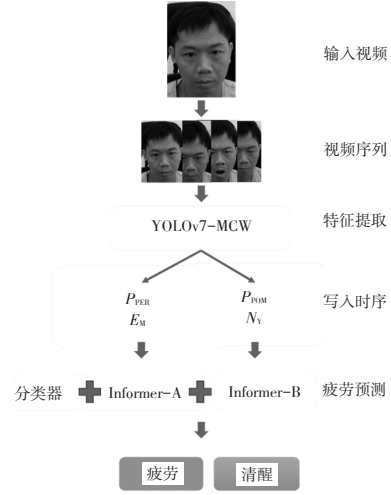


图7 整体检测框架

为了测试模型性能,进行了30组单次时长为1~2 h的模拟驾驶试验,并通过摄像头采集驾驶人面部图像,制作为短时疲劳检测面部视频数据集(Short-time Fatigue Driving Detection Video Dataset, SFDDVD),采集了30位年龄在20~30岁之间,驾龄在1年以上的受试者在高速道路上持续驾驶的正面面部视频。通过人工切分与标定,得到了600个时长为1 min的标签为“清醒”状态与“疲劳”状态的视频片段。

4.2 试验平台

本试验的平台为一台配备Intel® Core™ i7-10700K CPU 3.80 GHz处理器,内存为32G,配置NVIDIA GeForce RTX3060 12 GB显卡,Ubuntu 22.04.1操作系统的计算机。

4.3 在UTA-RLDD数据集上的表现

本文将UTA-RLDD进行了重新划分,将数据标注为“清醒”和“疲劳”2类,分别获得了60个“疲劳”类视频片段和60个“清醒”类视频片段。准确率 P_{Acc} 、精确率 P_{Pre} 和召回率 P_{Rec} 分别为:

$$P_{Acc} = \frac{T_p + T_n}{T_p + F_n + F_p + T_n} \quad (8)$$

$$P_{Pre} = \frac{T_p}{T_p + F_p} \quad (9)$$

$$P_{Rec} = \frac{T_p}{T_p + F_n} \quad (10)$$

式中: T_p 为被正确分类为“疲劳”的样本, F_p 为被错误分类为“疲劳”的样本, T_n 代表被正确分类为“清醒”的样本, F_n 为被错误分类为“清醒”的样本。

准确率、精确率和召回率计算结果如表3所示。

表3 在UTA-RLDD上的准确率、精确率和召回率 %

评价指标	准确率	精确率	召回率
结果	97.50	95.08	98.33

4.4 在NTHU-DDD数据集上的表现

对于NTHU-DDD数据集,根据给定的每一帧标签,从中裁剪出多个标记为“清醒”或“疲劳”的视频片段,如表4所示,试验结果如表5所示。

表4 测试集划分

类别	人数/人	清醒片段数量/个	疲劳片段数量/个
戴眼镜	14	29	24
夜晚戴眼镜	14	40	33
夜晚无眼镜	14	34	27
无眼镜	14	27	17
戴太阳镜	14	48	41
合计	70	178	142

表5 在NTHU-DDD上的准确率、精确率和召回率 %

类别	准确率	精确率	召回率
戴眼镜	94.34	95.65	91.67
夜晚戴眼镜	93.15	91.18	93.94
夜晚无眼镜	95.08	92.86	96.30
无眼镜	95.45	94.12	94.12
戴太阳镜	94.38	92.86	95.12
均值	94.48	93.33	94.23

4.5 消融试验

YOLOv7-MCW-Informer模型对主干网络、注意力机制、损失函数进行了改进。为评估不同结构改进与组合对算法性能的提升,设计了消融试验,共计8组。所有模型均在相同数据集上进行训练,得到最佳权重后在同一数据集SFDDVD上进行测试。结果如表6所示,M0~M7分别代表YOLOv7-Informer、YOLOv7-M-Informer、YOLOv7-C-Informer、YOLOv7-W-Informer、YOLOv7-MC-Informer、YOLOv7-MW-Informer、YOLOv7-CW-Informer、YOLOv7-MCW-Informer 8种模型。可以看出,本文提出的3个改进方法相比于原始的YOLOv7算法,性能方面均具有小幅提升。相比于最终算法YOLOv7-MCW-Informer(M7)算法,去除任何一个改进方法都会使得模型性能下降。消融试验结果证明了本文所提出的改进方法的有效性。

4.6 与现有模型的对比

表7给出了本文提出的模型与其他算法模型在性能上的比较。在UTA-RLDD上准确率可达97.50%,单

帧检测时间为33 ms;在NTHU-DDD上准确率可达94.48%,单帧检测时间为28 ms。相较于文献[3]、文献[5]、文献[16]、文献[17]中给出的方法,本文方法在检测准确率上有一定提升;与文献[18]、文献[19]对比,在同一数据集上,单帧检测速度接近的前提下,本文提出的模型具有较高的准确率,分别高出21.88%和7.74%;本模型在保持检测准确性有所提升的前提下,大幅缩减了检测时间;与文献[5]对比,在同一数据集上,准确率提升的前提下,单帧检测时间大幅下降,本文提出的模型检测速度接近其4倍。由此可以看出,YOLOv7-MCW-Informer模型通过并行网络分析短时面部特征,融合多维面部信息,提升了检测准确性与实时性。

表6 消融试验结果 %

模型	准确率	精确率	召回率
M0	89.50	87.74	91.83
M1	90.83	89.01	93.17
M2	91.83	90.35	93.67
M3	92.42	91.93	93.00
M4	93.33	92.62	94.17
M5	94.75	93.94	95.67
M6	94.58	93.35	96.00
M7	97.75	97.51	98.00

表7 多种算法模型性能对比

研究	数据集	检测器	分类器	整体准确率/%	单帧检测时间/ms
文献 ^[5]	UTA-RLDD	Haar Cascades	LeNet CNN	91.80	119
文献 ^[3]	NTHU-DDD	2s-STGCN	Softmax	92.70	
文献 ^[16]	ZJU	SVM	BLS	97.05	91
文献 ^[17]	NFDVD	RetinaFace	ShuffleNet V2	95.70	83
文献 ^[18]	NTHU-DDD	3DCNN	3DCNN	72.60	36
文献 ^[19]	NTHU-DDD	DNN	DNN	86.74	39
本文方法	UTA-RLDD	YOLOv7-MCW	Informer	97.50	33
本文方法	NTHU-DDD	YOLOv7-MCW	Informer	94.48	28

5 结束语

本文提出了一种基于并行短时面部特征的驾驶人疲劳驾驶检测方法。YOLOv7-MCW-Informer模型相较于现有方法能够提升疲劳检测的准确性,YOLOv7-MCW-Informer模型在领域内公开数据集汽车技术

UTA-RLDD、NTHU-DDD上分别达到97.50%和94.48%的准确率;相较于现有工作,保持准确性的基础上具备更好的实时性,单帧检测时间最低达到28 ms,时序预测模型能够更好地提取驾驶人疲劳状态的变化趋势,有利于及早预警和干预。

参 考 文 献

- [1] DU G, ZHANG L, SU K, et al. A Multimodal Fusion Fatigue Driving Detection Method Based on Heart Rate and PERCLOS[J]. IEEE Transactions on Intelligent Transportation Systems, 2022, 23(11): 21810-21820.
- [2] YE M, ZHANG W, CAO P, et al. Driver Fatigue Detection Based on Residual Channel Attention Network and Head Pose Estimation[J]. Applied Sciences, 2021, 11(19).
- [3] BAI J, YU W, XIAO Z, et al. Two-Stream Spatial-Temporal Graph Convolutional Networks for Driver Drowsiness Detection[J]. IEEE Transactions on Cybernetics, 2021, 52(12): 13821-13833.
- [4] 姜平, 杨欣, 胡辑伟, 等. 基于边缘计算的疲劳驾驶检测方法[J]. 计算机工程, 2021, 47(7): 13-20+29.
- LOU P, YANG X, HU J W, et al. Fatigue Driving Detection Method Based on Edge Computing[J]. Computer Engineering, 2021, 47(7): 13-20+29.
- [5] TAMANANI R, MURESAN R, AL-DWEIK A. Estimation of Driver Vigilance Status Using Real-Time Facial Expression and Deep Learning[J]. IEEE Sensors Letters, 2021, 5(5): 1-4.
- [6] WANG C Y, BOCHKOVSKIY A, LIAO H Y M. YOLOv7: Trainable Bag-of-Freebies Sets New State-of-the-Art for Real-Time Object Detectors[C]// Oxford: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023: 7464-7475.
- [7] LI Y, CHEN Y, DAI X, et al. Micronet: Improving Image Recognition with Extremely Low Flops[C]// Proceedings of the IEEE/CVF International Conference on Computer Vision. Montreal: IEEE, 2021: 468-477.
- [8] HOU Q, ZHOU D, FENG J. Coordinate Attention for Efficient Mobile Network Design[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Nashville: IEEE, 2021: 13713-13722.
- [9] WANG W, LIU W. Small Object Detection with YOLOv8 Algorithm Enhanced by MobileViTv3 and Wise-IoU[C]// Proceedings of the 2023 12th International Conference on Computing and Pattern Recognition. Qingdao: ACM, 2023: 174-180.
- [10] SOARES S, FERREIRA S, COUTO A. Driving Simulator Experiments to Study Drowsiness: A Systematic Review[J]. Traffic Injury Prevention, 2020, 21(1): 29-37.
- [11] KHUNPISUTH O, CHOTCHINASRI T, KOSCHAKOSAI V, et al. Driver Drowsiness Detection Using Eye-Closeness Detection[C]// 2016 12th International Conference on Signal-Image Technology & Internet-Based Systems (SITIS). Naples: IEEE, 2016: 661-668.
- [12] ZHOU H, ZHANG S, PENG J, et al. Informer: Beyond Efficient Transformer for Long Sequence Time-Series Forecasting[C]// Proceedings of the AAAI Conference on Artificial Intelligence. Vancouver: AAAI 2021, 35(12): 11106-11115.
- [13] 廖冬杰. 基于Dlib和变种Transformer的哈欠检测方法[J]. 汽车技术, 2023(3): 42-48.
- LIAO D J. Yawn Detection Method Based on Dlib and Variant Transformer[J]. Automobile Technology, 2023(3): 42-48.
- [14] GHODDOOSIAN R, GALIB M, ATHITSOS V. A Realistic Dataset and Baseline Temporal Model for Early Drowsiness Detection[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops. Long Beach: IEEE, 2019.
- [15] WENG C H, LAI Y H, LAI S H. Driver Drowsiness Detection via A Hierarchical Temporal Deep Belief Network[C]// ACCV Workshops. Taipei, China: Springer International Publishing, 2017.
- [16] 任俊, 魏霞, 黄德启, 等. 基于眼睛状态多特征融合的疲劳驾驶检测[J]. 计算机工程与设计, 2022, 43(11): 3187-3194.
- REN J, WEI X, HUANG D Q, et al. Fatigue Driving Detection Based on Multi-Feature Fusion of Eye Status[J]. Computer Engineering and Design, 2022, 43(11): 3187-3194.
- [17] 王鹏, 神和龙, 尹勇, 等. 基于深度学习的船舶驾驶员疲劳检测算法[J]. 交通信息与安全, 2022, 40(1): 63-71.
- WANG P, SHEN H L, YIN Y, et al. Ship Driver Fatigue Detection Algorithm Based on Deep Learning[J]. Traffic Information and Safety, 2022, 40(1): 63-71.
- [18] YU J, PARK S, LEE S, et al. Representation Learning, Scene Understanding, and Feature Fusion for Drowsiness Detection[C]// ACCV 2016 Workshops. Taipei, China: Springer International Publishing, 2017.
- [19] BEKHOUCHE S E, RUCHEK Y, DORNAIKA F. Driver Drowsiness Detection in Video Sequences Using Hybrid Selection of Deep Features[J]. Knowledge-Based Systems, 2022, 252.

(责任编辑 王 一)

修改稿收到日期为2023年8月7日。