

DOI:10.19651/j.cnki.emt.2518506

基于 MTCNN 和 PFLD 的疲劳驾驶检测方法^{*}

宋志强^{1,2} 李明阳² 周鹏²

(1. 无锡学院自动化学院 无锡 214105; 2. 南京信息工程大学自动化学院 南京 210044)

摘要: 针对驾驶员疲劳检测方法中存在面对光照变化或复杂背景时人脸检测精度下降的局限性,提出一种改进的 MTCNN 网络。通过对 MTCNN 网络进行优化,在 3 个子网络中均引入坐标注意力机制和批量归一化算法,提高模型对驾驶员面部的定位精度,提升网络的收敛速度和稳定性,并增强对过拟合的抑制。实验结果表明:改进 MTCNN 模型在疲劳驾驶数据集上的准确率达到 98.78%,比原模型提高了 2.43%,且模型参数量仅为 0.5 M,具有良好的 人脸检测精度和可部署性。此外,将改进 MTCNN 模型与 PFLD 模型结合,根据实验设定了合理的疲劳参数阈值,并实现了较为准确的疲劳驾驶检测。

关键词: 疲劳驾驶检测;人脸检测;MTCNN;PFLD;关键点检测

中图分类号: TP391.4;TN911.73 **文献标识码:** A **国家标准学科分类代码:** 510.40

Fatigue driving detection method based on MTCNN and PFLD

Song Zhiqiang^{1,2} Li Mingyang² Zhou Peng²

(1. School of Automation, Wuxi University, Wuxi 214105, China;

2. School of Automation, Nanjing University of Information Science and Technology, Nanjing 210044, China)

Abstract: Aiming at the limitation of face detection accuracy degradation when facing light changes or complex background in driver fatigue detection methods, an improved MTCNN network is proposed. By optimising the MTCNN network, the coordinate attention mechanism and batch normalisation algorithm are introduced in all three sub-networks to improve the model's localisation accuracy of the driver's face, enhance the convergence speed and stability of the network, and enhance the suppression of overfitting. The experimental results show that the accuracy of the improved MTCNN model on the fatigue driving dataset reaches 98.78%, which is 2.43% higher than that of the original model, and the number of parameters of the model is only 0.5 M, which has good face detection accuracy and deployability. In addition, combining the improved MTCNN model with the PFLD model, a reasonable fatigue parameter threshold is set based on the experiments, and a more accurate fatigue driving detection is achieved.

Keywords: fatigue driving detection; face detection; MTCNN; PFLD; key point detection

0 引言

疲劳驾驶是导致交通事故的主要原因之一。相关研究数据表明,约有 20% 的交通事故是由疲劳驾驶引起的,而在涉及特大伤亡交通事故时,疲劳驾驶导致的交通事故约占其中的 40%^[1]。因此,为了有效预防疲劳驾驶的发生,开发一种高效精确的驾驶员疲劳检测方法,以降低疲劳驾驶引发的交通事故风险,具有重要意义。

基于当前的交通状况,大多数国家对疲劳驾驶展开了深入研究,在实际应用中,相关领域的车企也相继研发了疲劳驾驶检测系统^[2]。目前主流的疲劳驾驶检测方法根据信息采集方式可分为侵入式和非侵入式两大类。根据提取特

征类型的不同,这些方法又可进一步划分为 3 种类型:驾驶员生理特征、车辆驾驶信息和驾驶员面部特征^[3]。

1 疲劳检测方法

1.1 基于驾驶员生理特征的疲劳检测方法

当驾驶员陷入疲劳状态时,其生理指标会出现明显的变化。通过相关生理特征采集设备,可以在行车过程中实时收集数据,并对其进行处理与深入分析,从而判断驾驶员是否存在疲劳状态^[4]。生理特征^[5]包括脑电图(electroencephalogram, EEG)信号、眼电图(electro-oculogram, EOG)信号以及心电图(electro-cardiogram, ECG)信号等。

收稿日期:2025-04-03

* 基金项目:江苏省产学研合作项目(BY20230688)资助

Dogan等^[6]提出了一种基于EEG的驾驶疲劳检测方法,其信号分类框架结合了小波包分解(wavelet packet decomposition, WPD)、16个滤波器、3个特征提取器、3个特征选择器、分类器和迭代多数投票法(iterative majority voting, IMV)算法。实验结果表明,该框架在分类性能上表现卓越,验证其在准确检测疲劳方面的有效性。Wang等^[7]通过整合脑电图和眼动追踪模式,引入了一种驱动疲劳检测的新架构,并提出一种便携式半干式电极用于长时间有效采集ECG信号。最后,研究使用小波散射网络对驾驶疲劳进行分类,结果表明,与传统的分类方法相比,该方法在分类精度和运行速度上均表现更优。为了准确识别驾驶员的疲劳状态,王琳等^[8]采用生物力学分析方法,提取了驾驶过程中颈部和腰部的肌电信号以及EEG信号,并研究了这些信号特征的变化趋势。研究人员通过主成分分析实现了各特征参数的有效整合,进而基于多元回归理论构建了一个能够有效预测疲劳驾驶的数学模型。验证结果表明,该模型在疲劳状态判别方面表现出良好的性能。

通过检测驾驶员的生理特征,可以对其疲劳状态进行有效评估。该方法具有实时性和准确率高等优势。然而,由于其属于侵入式检测,需要驾驶员佩戴相关检测设备,因此存在便捷性不足、成本较高以及检测精度易受环境干扰等问题,同时也可能分散驾驶员的注意力,从而对驾驶安全产生不利影响。

1.2 基于车辆驾驶信息的疲劳检测方法

疲劳状态会导致驾驶员反应迟钝,进而引发车辆操控异常,如车速波动、转向不稳或车道偏离。基于车辆驾驶信息的疲劳检测方法,通过分析方向盘转角、加速度等动态数据,可间接评估驾驶员的疲劳状态^[9]。

Mcdonald等^[10]考虑了驾驶员在困倦和清醒状态之间转换的时间依赖性,创新性地引入时间序列关联分析,运用动态贝叶斯网络构建检测模型,综合转向操作、踏板位移、加速度等多源信号进行状态判别。研究表明,该方法有效提升了疲劳检测的准确性。蔡素贤等^[11]基于控制器局域网(controller area network, CAN)总线采集车辆运行状态数据,并从中提取了18项与驾驶行为相关的特征,用于分析和检测驾驶员的疲劳状态,并采用随机森林算法识别疲劳状态。实验结果表明,该方法具有较好的检测效果。国内哈弗的汽车疲劳检测系统通过监测车辆状态参数评估驾驶员状态,并利用声音警报及仪表盘显示疲劳信息等方式向驾驶员发出警告。

通过分析车辆的转向角、航向角、加速度等参数,可以评估驾驶员的疲劳程度,并通过多参数融合可以有效提高模型的检测精度。然而,该方法难以直接反映驾驶员生理和心理状态的变化,同时存在一定的滞后性,无法准确、实时地反馈驾驶员的疲劳状态。

1.3 基于驾驶员面部特征的疲劳检测方法

在驾驶疲劳状态下,驾驶员常呈现特定的面部行为特

征,如哈欠频发、眼部快速眨动或头部持续性低垂等。基于面部特征的疲劳检测方法^[12]通过图像采集设备实时获取驾驶员面部数据,并利用人工智能算法精准提取眼部、嘴部及头部的动态变化特征。通过分析不同状态下的面部特征,特别是眼部闭合持续时间、嘴部开合幅度等关键特征,同时结合周期性眨眼行为与哈欠动作的频次特征,从而判断驾驶员是否处于疲劳状态。

为了提高疲劳驾驶检测方法的实时性和准确性,史瑞鹏等^[13]针对多任务卷积神经网络算法实施了三重加速优化,在保持检测准确率的前提下显著提升了检测速度。在人脸特征点精确定位的基础上,根据三庭五眼理论,提出了基于稀有特征点提取图像(extracting images based on rare feature points, ERFP)方法快速准确地提取目标区域图像,并利用构建的眼和嘴数据集 EMSD(eye and mouth state datasets),通过卷积神经网络 EM-Net(eye and mouth network)训练得到眼、嘴部状态分类模型。最后,通过设计的判定算法,实现了疲劳驾驶的识别。实验结果表明,该方法不仅具有较高的检测准确率,而且具备良好的实时性。针对疲劳驾驶检测中存在的特征识别精度不足、鲁棒性差及传统方法侵入性强等问题,薛丽^[14]提出了一种基于改进多任务卷积神经网络的多特征融合检测方法。该方法首先对三个子网进行针对性优化,并结合目标跟踪算法,引入方向梯度直方图(histogram of oriented gradients, HOG)特征且改进了跟踪失败检测机制,以实现驾驶员人脸部位的稳定追踪。此外,构建了基于支持向量机的多特征融合决策模型,综合利用眼部纵横比、嘴部纵横比等多维生理特征,建立了具有时空关联性的疲劳判定体系。实验结果表明,该方法具有良好的检测准确率。

Zhao等^[15]提出了一种基于驾驶图像的全自动驾驶员疲劳状态检测算法,该算法采用多任务级联卷积神经网络架构,用于人脸检测和特征点定位,并通过提取关键特征点来确定感兴趣区域。此外,研究设计了眼部-嘴部卷积神经网络(eye-mouth convolutional neural network, EM-CNN)网络,用于监测眼部和嘴部的状态。通过分析眼睑闭合的时间比例、瞳孔变化率以及嘴部开合度等关键参数实现疲劳检测。实验结果表明,该算法能够高效地基于驾驶图像检测驾驶员的疲劳状态。王诗皓^[16]开发了一种集成卷积神经网络与长短期记忆网络的疲劳检测系统,有效捕获驾驶员行为与面部动态特征。通过引入数据增强策略与迁移学习机制,显著提升了模型的泛化能力与适应性能。实验结果表明,该深度学习模型在疲劳状态识别准确度上较传统方法具有显著优势。

基于驾驶员面部特征的疲劳检测方法采用计算机视觉技术捕捉面部关键区域的动态变化,以实现非侵入式识别。该方法不仅部署灵活、干扰性低,且检测精度较高,已成为当前广泛应用的疲劳驾驶检测技术之一。

通过对驾驶员疲劳检测方法的研究回顾表明:尽管已经进行了大量的研究,但仍存在一些缺陷。例如,人脸检测算法在面对实际驾驶环境可能存在光线剧烈变化情况时检测精度下降,以及目前大部分疲劳检测模型由于参数量较大,在实际部署时会占用过多计算资源,难以满足实时性要求。

针对这些挑战,本文提出了相应的改进措施。本文的主要贡献可以概括为以下几点:

1) 在轻量级网络多任务卷积神经网络 (multi-task convolutional neural network, MTCNN) 的 3 个子网络中均嵌入坐标注意力机制,提升模型对人脸信息的敏感度并增强其在复杂背景下的鲁棒性,确保模型能够精确定位人脸及其他空间位置。

2) 在 MTCNN 的 3 个子网络中引入批量归一化算法,提升网络的收敛速度和稳定性,增强对过拟合的抑制。

3) 本文设计改进的 MTCNN 检测模型在数据集上进行实验,结果表明该模型具有较高的人脸检测精度,与轻量级 PFLD 模型结合使用,能够满足疲劳驾驶检测需求。

2 MTCNN 模型

MTCNN 作为当前主流的轻量级人脸检测框架之一^[17]。该模型对自然环境下光线轻微变化、角度偏差和人脸表情的变化具有良好的鲁棒性,并在人脸检测任务中的表现明显优于传统目标检测算法。此外,MTCNN 采用分

阶段网络设计策略,通过三级联结构的协同优化,实现面部区域的快速定位与特征提取。其轻量化特性体现在低参数规模与低内存占用的双重优势,能够有效满足嵌入式设备对实时处理性能的需求。MTCNN 网络的主要结构包括图像金字塔以及 P-Net (proposal network)、R-Net (refine network)、O-Net(output network) 3 个子网络。该网络通过多尺度图像变换生成不同尺寸的候选矩形框,并采用滑动窗口方法检测人脸区域,同时提取矩形框的特征。为了提高人脸检测的准确性,MTCNN 引入了非极大值抑制 (non-maximum suppression, NMS) 算法^[18],用于去除同一目标中高度重叠的候选框,从而保留最佳边界框。经过 3 个子网络的逐步处理,最终模型输出人脸预测框及面部对应的关键点位置。

3 改进的 MTCNN 模型

本文在基础网络 MTCNN 的基础上进行了改进。为了提升模型对人脸信息的敏感度并增强其在复杂背景下的鲁棒性,确保模型能够精确定位人脸及其他空间位置,在 MTCNN 的 3 个子网络中均引入了坐标注意力机制。为了保证 MTCNN 较高检测精度的同时,减少推理时间,并增强模型面对未经充分预处理的数据时的鲁棒性,引入了批量归一化算法。改进后的 MTCNN 网络结构如图 1 所示,其中,圆柱体部分代表坐标注意力机制,长方体部分代表批量归一化算法。

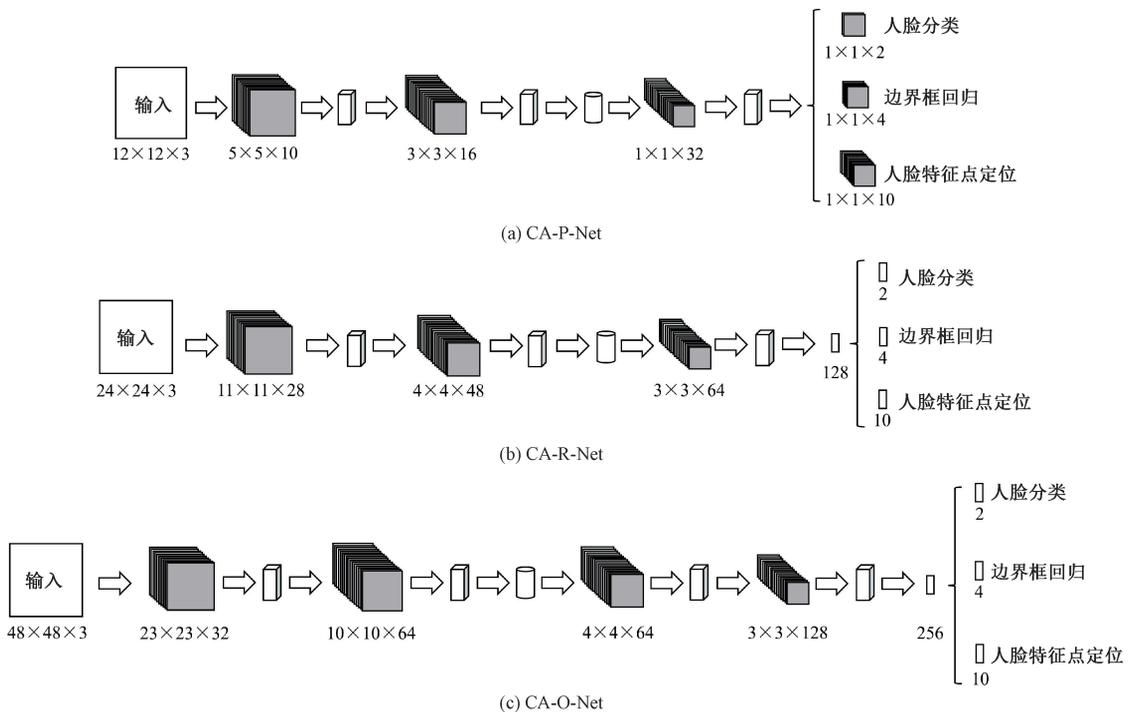


图 1 CA-MTCNN 网络结构

Fig. 1 Network architecture of CA-MTCNN

3.1 结合坐标注意力机制的MTCNN

坐标注意力(coordinate attention, CA)机制旨在增强图像的空间关系^[19]。其核心思想是在目标检测任务中引入坐标信息,通过调整特征图的注意力分布,进一步聚焦于目标位置以提升检测性能。CA将二维全局池化解为两个一维池化操作,即分别沿水平方向和垂直方向进行全局平均池化。这种分解方法不仅能够沿单一维度捕获跨区域的远程依赖关系,同时保持正交维度的精确定位能力,实现了大范围上下文感知与局部定位精度的协同优化。在特征编码阶段,CA模块通过方向敏感的注意力掩码生成算法,构建具有空间感知能力的增强特征图。最终通过特征加权融合机制,显著提升目标区域的特征表达能力,使模型在复杂背景干扰下仍能保持优异的判别性能与定位鲁棒性。

CA作为一种轻量级注意力机制,参数量较小,能够嵌入位置信息以提高模型在定位人脸五官等结构特征时的精度。CA模块通过为特征图的各个通道及空间位置分配差异化权重,有效增强了神经网络的特征学习能力。因此,CA模块适用于本研究的MTCNN模型,其结构如图2所示。

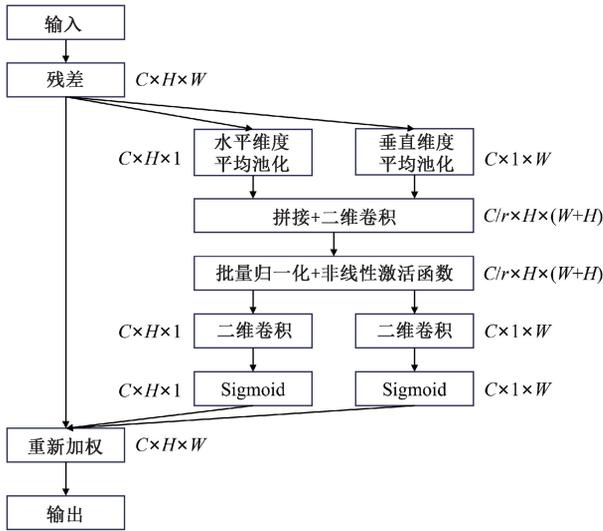


图2 CA模块
Fig.2 CA module

如图2所示,首先,针对给定大小为 $C \times H \times W$ 的特征图 X 采用双轴向特征编码机制实施空间信息提取。在特征处理阶段,分别在水平维度和垂直维度应用池化核进行编码,水平方向采用尺寸为 $(1, W)$ 的池化窗口,垂直方向则配置 $(H, 1)$ 维度的池化滤波器,从而得到大小为 $C \times H \times 1$ 和 $C \times 1 \times W$ 的特征图。该方法在水平维度维持序列元素的全局关联特性,同时确保了垂直方向的精确位置信息。特别的,当对特征图 X 执行水平方向池化运算时,第 c 个特征通道在垂直坐标为 h 处的输出可通过式(1)进行形式化描述:

$$z_c^h(h) = \frac{1}{W} \sum_{0 \leq i \leq W} x_c(h, i) \tag{1}$$

其中, $z_c^h(h)$ 表示第 c 个特征通道在垂直坐标为 h 处的水平方向池化输出值, $x_c(h, i)$ 输入特征图 X 的第 c 个通道在空间位置 (h, i) 处的特征值。

同理,对于特征图沿垂直方向进行池化后,第 c 个通道在横坐标为 w 处的输出如式(2)所示。

$$z_c^w(w) = \frac{1}{H} \sum_{0 \leq j \leq H} x_c(j, w) \tag{2}$$

其中, $z_c^w(w)$ 表示第 c 个特征通道在水平坐标为 w 处的垂直方向池化输出值, $x_c(j, w)$ 输入特征图 X 的第 c 个通道在空间位置 (j, w) 处的特征值。

将经过水平和垂直方向池化操作得到的特征图进行拼接,并通过 1×1 卷积进行降维,得到大小为 $(C/r) \times 1 \times (W+H)$ 的特征图,其计算公式如式(3)所示。

$$f = \delta(F_1([z^h, z^w])) \tag{3}$$

其中, δ 表示非线性激活函数, F_1 表示 1×1 卷积。

接着,将得到的特征图 f 沿空间维度分解为两个独立张量 f^h 和 f^w ,其维度参数分别为 $(C/r) \times H$ 和 $(C/r) \times W$ 。为实现通道维度的特征适配,利用 1×1 卷积操作 F_h 和 F_w 对张量 f^h 和 f^w 的通道数进行调整,通过该过程使其与原始输入特征图 X 通道维度 C 精准匹配。经此通道重校准操作后,最终生成具有空间感知特性的注意力权重分布,其数学表征如式(4)与(5)所示。

$$g^h = \sigma(F_h(f^h)) \tag{4}$$

$$g^w = \sigma(F_w(f^w)) \tag{5}$$

最后,坐标注意力模块的输出如式(6)所示。

$$y_c(i, j) = x_c(i, j) \times g_c^h(i) \times g_c^w(j) \tag{6}$$

在模型迭代训练过程中,CA模块通过参数动态调整与渐进式优化机制,实现权值空间参数向全局最优解的渐进式逼近。由于在面对光照剧烈变化或复杂背景时,MTCNN的检测精度会下降。因此,本研究在MTCNN的3个子网络P-Net、R-Net、O-Net中都引入了CA模块,提升模型对人脸信息的敏感度并增强其在复杂背景下的鲁棒性,确保模型能够精确定位人脸及其他空间位置。

3.2 结合批量归一化的MTCNN

深度神经网络的训练过程较为复杂。在训练过程中,前一层网络参数的更新会影响下一层的输入分布。此外,较低的学习率和过于精细的参数初始化可能会减缓模型训练速度,尤其会导致具有饱和和非线性激活函数的模型收敛困难,并容易陷入梯度消失问题。

批量归一化(batch normalization, BN)算法^[20]通过对中间层的输出进行归一化处理,将输入值的差异减小到相同的范围,在提高梯度收敛程度、加快模型训练速度的同时,使得各层可以尽可能在面对相同特征分布的输入值时,降低变化带来的不确定性,也减少对后一层网络的冲击,使各层网络变得相对独立,缓解训练中梯度消失的问题。

题,从而提高网络性能。在训练过程中,BN 通常被插入到每一层的线性变换(如全连接层或卷积层)和非线性激活函数(如 ReLU)之间。其基本步骤如下:

1)假设同一批次的所有输入特征为 $X = [x_1, x_2, \dots, x_m]$,对于每个特征计算其均值 μ 和方差 σ^2 ,结果如式(7)所示。

$$\mu = \frac{1}{m} \sum_{i=1}^m x_i, \sigma^2 = \frac{1}{m} \sum_{i=1}^m (x_i - \mu)^2 \quad (7)$$

其中, m 表示批次中的样本数量, x_i 表示第 i 个样本在某一维度的特征值。

2)在特征维度上,利用计算得到的均值和方差对每个特征进行归一化处理,其计算公式如式(8)所示。

$$\hat{x}_i = \frac{x_i - \mu}{\sqrt{\sigma^2 + \epsilon}} \quad (8)$$

其中, ϵ 是一个很小的常数,避免分母为零。

3)为了保证模型具有足够的表达能力,在归一化处理引入两个可学习参数,缩放参数 γ 和偏移参数 β ,对归一化后的特征进行线性变换后的结果如式(9)所示。

$$y_i = \gamma \hat{x}_i + \beta \quad (9)$$

其中, γ 控制特征尺度, β 控制特征偏移。这些参数通过反向传播训练,最终使网络能够学习到适应目标任务的特征分布。

MTCNN 在人脸检测与特征点定位任务中具有显著优势,然而,其在实际应用中仍存在一些局限性。由于 MTCNN 采用了级联架构,其多阶段处理方式虽然有助于逐步提高检测精度,但也显著增加了推理时间。这种架构在高分辨率图像或实时处理场景中表现出一定的效率瓶颈,难以满足对速度要求较高的应用场景。此外,MTCNN 对输入图像的尺度、对比度以及噪声较为敏感。这种对输入数据预处理的高度依赖性使得模型在面对未经充分预处理的数据时表现出鲁棒性不足,进而影响其在复杂实际场景中的检测效果。本研究在 MTCNN 的 3 个子级网络中引入 BN 算法,使得网络的训练效率和性能得到显著提升,尤其是在收敛速度、稳定性以及对过拟合的抑制方面。

4 实验结果与分析

4.1 数据集与评价指标

本文实验采用疲劳驾驶数据集(yawning detection dataset, YawDD)^[21],该数据集包含驾驶过程中拍摄的驾驶员面部视频数据,分为驾驶员正面和侧面两种情况。通过对上述视频进行抽帧和筛选,分别获得打哈欠、正常状态数据各 700 余帧。这些人脸数据在光照变化、姿态调整等方面呈现出多样性,为本文的研究提供了重要且可靠的实验数据。图 3 为部分数据集示例图。将数据集按 7:2:1 的比例分类为训练集、验证集、测试集进行模型训练。

为了更加精确的评估疲劳检测模型的表现,在模型训练前,通常需要找到各类模型的通用评价标准。因此,在



图 3 数据集示例图

Fig. 3 Example diagram of the dataset

疲劳检测中,本研究选用准确率、精确率、召回率、F1-得分作为模型性能评判指标。

4.2 实验环境

实验环境配置如表 1 所示。

表 1 实验环境配置

Table 1 Experimental environment configuration

设备名称	设备信息
操作系统	Windows 11
CPU	i5-13490F
GPU	RTX3080(10.0 GB)
Python 版本	3.8
CUDA 版本	11.3
Pytorch 版本	1.12.0

4.3 实验结果与分析

为了进一步深入评估改进后的 MTCNN 模型的贡献和优势,本文进行了消融实验,逐步评估其改进部分对模型的影响,结果如表 2 所示。

表 2 消融实验

Table 2 Ablation experiments

模型	BN	CA	准确率	精确率	召回率	F1-得分	%
MTCNN	—	—	96.35	96.4	96.32	96.35	
MTCNN+BN	✓	—	96.57	96.59	96.55	96.58	
MTCNN+AC	—	✓	97.71	97.73	97.68	97.69	
改进 MTCNN	✓	✓	98.78	98.8	98.77	98.78	

如表 2 所示,本文对提出的 CA 模块和 BN 算法进行了组合实验,通过准确率、精确率、召回率、F1-得分等评价指标全面验证了本文中所提方法的有效性能。消融实验结果表明:本文的改进显著提升了原模型的整体性能,验证了其在疲劳驾驶检测中的有效性。此外,改进 MTCNN 模型的参数量仅为 0.5 M,符合轻量级检测模型的标准,未来将其部署在嵌入式设备上的难度会大幅降低。

为了验证模型改进的有效性,本文将改进后的模型与 RetinaFace^[22]、YOLOv8-SWF^[23]、YOLOv5^[24]、4AC-YOLOv5^[25]和 SCRFD-10-GF^[26]等主流人脸检测模型在同数据集上进行了性能对比,评价指标包括检测准确率和参数量,具体结果如表 3 所示。通过对比不同模型的实验结

果,本文提出的改进MTCNN模型相较于RetinaFace、YOLOv5、4AC-YOLOv5和SCRFD-10 GF模型在参数量和准确率方面均表现出领先水平。尽管与YOLOv8-SWF对比,本文改进的模型的准确率降低了0.48%,但是参数量却大幅度减少。实验结果证明:本文模型在与各种主流模型的对比中展现出显著优势,验证了本文方法的有效性和性能优势。

表3 对比实验

Table 3 Comparative experiments

模型	准确率/%	参数量/M
RetinaFace	95.60	25.10
YOLOv8-SWF	99.26	11.13
YOLOv5	91.20	8.16
4AC-YOLOv5	93.08	6.79
SCRFD-10 GF	95.20	3.86
本文模型	98.78	0.50

由于数据集中驾驶环境比较的单一,而实际疲劳驾驶检测场景较为复杂,需要应对行车途中光照强度剧烈变化等情况。为此,本研究采用数据增强技术对数据集进行扩充,以模拟上述驾驶场景并提高实验数据的多样性。数据增强技术主要采用光度调整和伽马校正^[27]的方式对原数据图像进行图像转化。图4为经过数据增强后的数据集的部分示例图。



图4 数据集示例图

Fig. 4 Example diagram of the dataset

如图5所示为训练模型在经过数据增强后的数据集上所得的人脸检测结果,默认带有人脸的5个关键点定位。实验结果表明,改进MTCNN的模型能够在精准地定位人脸,满足疲劳驾驶人脸检测需求。

5 面部关键点检测与疲劳检测

5.1 眼部疲劳特征提取方法

人脸关键点检测能够精确定位人脸特征的坐标,从而

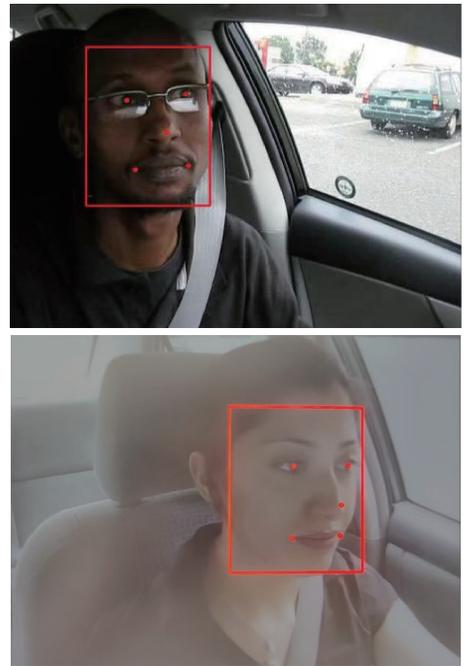


图5 人脸检测结果示例

Fig. 5 Example of face detection results

表征出人脸的各个部分,包含眉毛、眼部、鼻子、嘴部以及脸部轮廓等。随着人脸检测技术的发展,其应用场景变得愈加复杂,且对检测精度的要求也日益提高。由于MTCNN只能检测出5个关键点,在实际人脸检测场景中无法满足需求,本文选择实用人脸关键点检测器(practical facial landmark detector, PFLD)模型^[28]用于检测驾驶员的人脸关键点。PFLD在大多数正常环境下展现出较高的准确性,且由于其具有模型小、参数量少、检测速度快等优势,特别适合未来集成到移动嵌入式设备中进行应用。

PFLD模型基于YawDD图像和视频数据集训练。图6展示了基于PFLD模型的人脸关键点位置分布示意图,其中人脸被检测出68个关键点的坐标。根据图中标定的关键点坐标,能够精准识别输入图像中的眼部与嘴部位置,从而提取相应的局部区域图像。

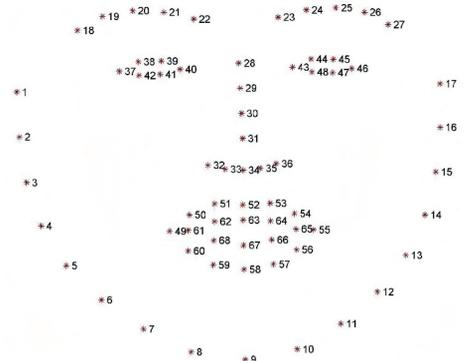


图6 人脸68关键点分布图

Fig. 6 Distribution of key points of face 68

5.2 眼部疲劳特征提取方法

驾驶员在出现疲劳状态时,眼部的表现最为突出。因此,本文将眼部疲劳特征列为判断驾驶员疲劳的重要标准。

为精准判断驾驶员的眼部状态,并有效区分睁眼与闭眼情况,Soukupova 等^[29]创新性地提出了一种基于眼部纵横比(eye aspect ratio, EAR)的疲劳状态判定方法。该方法通过计算眼部关键点垂直与水平方向的距离比例,得到 EAR 值,并基于该参数值的变化趋势实现对驾驶员疲劳状态的量化评估,其中,左眼 6 个关键点对应的坐标编号为(37~42),右眼 6 个关键点对应的坐标号为(43~48)。左右眼的眼部纵横比计算公式分别如式(10)及(11)所示。

$$EAR_{\text{left}} = \frac{\|P_{38} - P_{42}\| + \|P_{39} - P_{41}\|}{2 \|P_{37} - P_{40}\|} \quad (10)$$

$$EAR_{\text{right}} = \frac{\|P_{44} - P_{48}\| + \|P_{45} - P_{47}\|}{2 \|P_{43} - P_{46}\|} \quad (11)$$

其中, $P_{37} - P_{42}$ 表示左眼轮廓的 6 个关键点坐标, $\|P_{38} - P_{42}\|$ 为坐标 P_{38} 到 P_{42} 计算的垂直距离,其余部分同理。

为了减少驾驶姿势对眼部疲劳检测的影响,眼部疲劳特征参数通过计算左右眼部纵横比值的平均值来确定。其计算公式如式(12)所示。

$$EAR = \frac{EAR_{\text{left}} + EAR_{\text{right}}}{2} \quad (12)$$

考虑到不同驾驶员的眼睛形状及大小不同,为确保 EAR 阈值的设置更加合理准确,本文选取公开数据集 YawDD 中放置在后视镜上采集的 20 段视频进行 EAR 阈值设定实验。

图 7 记录了某驾驶员样本在 20 s 实验周期内 EAR 值的时序变化曲线。横轴代表该时间段内的视频总帧数,纵坐标对应 EAR 值。可以看出,当驾驶员睁眼时,EAR 值在一定范围内小幅波动;当发生眨眼动作时,该参数会在连续帧内急剧降低,随后恢复至正常波动范围,此时 EAR 曲线呈现出明显的波谷形态。

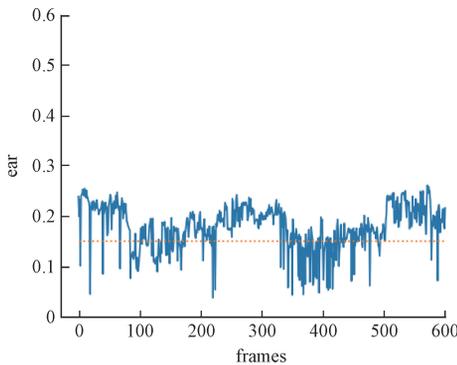


图 7 EAR 值变化曲线图

Fig. 7 EAR value change curve graph

在数据分析阶段,将采集的 20 组 EAR 值划分为 5 组,

并统计了 3 次正常睁眼与闭眼状态下对应的 EAR 值,具体结果如表 4 所示。实验结果表明,当驾驶员处于睁眼状态时,EAR 值在 0.20~0.26 之间波动;在闭眼状态下,该值大多分布在 0.05~0.11 之间。综合分析后,最终将 EAR 阈值设置为 0.15,以此作为区分眼部是否处于闭合状态的判定标准。

表 4 EAR 值实验结果

Table 4 Experimental results of EAR values

组别	第 1 次	第 2 次	第 3 次
第 1 组	0.24/0.10	0.22/0.05	0.21/0.09
第 2 组	0.26/0.08	0.22/0.09	0.23/0.07
第 3 组	0.24/0.11	0.20/0.08	0.21/0.07
第 4 组	0.21/0.08	0.23/0.06	0.22/0.09
第 5 组	0.25/0.07	0.26/0.10	0.25/0.09

5.3 嘴部疲劳特征提取方法

在轻度或中度疲劳状态下,驾驶员常常会出现打哈欠的现象。因此,嘴部特征也成为评估疲劳驾驶的重要指标之一。

借鉴 EAR 值的概念,本文引入嘴部纵横比(Mouth Aspect Ratio, MAR)进行嘴部状态检测^[30],通过 PFLD 模型提取驾驶员嘴部区域的关键点,从中提取 6 个关键点,并利用这些关键点计算 MAR 值,所需 6 个关键点对应的坐标编号为(61, 62, 64, 65, 66, 68)。MAR 值计算与 EAR 值类似,如式(13)所示。

$$MAR = \frac{\|P_{62} - P_{68}\| + \|P_{64} - P_{66}\|}{2 \|P_{61} - P_{65}\|} \quad (13)$$

针对驾驶疲劳引发的面部特征变化展开分析,重点关注哈欠动作这一典型生理反应。本文采用 MAR 值作为哈欠检测指标。当嘴部处于闭合状态时,MAR 值处于最低水平;随着嘴部逐渐张开时,该参数逐渐递增,并在最大张口阶段达到峰值。通过实验验证确定的最优 MAR 临界值,可实现哈欠状态的有效识别。

为了确定一个合理的 MAR 阈值,本文使用 YawDD 数据集中从仪表盘上方镜头采集的 29 段视频进行阈值优化分析,其中男性 16 例,女性 13 例。

图 8 展示了哈欠周期内 MAR 参数的动态变化过程,其中横轴为视频帧序列,纵轴对应 MAR 值,曲线对应连续变化轨迹。由图 8 可知,当嘴部处于闭合状态时,MAR 值会在一个稳定范围内小幅波动;当驾驶员打哈欠时,MAR 值会迅速上升,导致折线图中出现明显的波峰。在嘴部轻微张开的情况下,MAR 值会有略微增加,但不会达到波峰位置。

将得到的 29 组 MAR 值划分为 5 组,并统计 3 次从嘴部闭合到张嘴打哈欠时的 MAR 值,如表 5 所示。

实验结果发现嘴部处于闭合或微张状态时,MAR 值

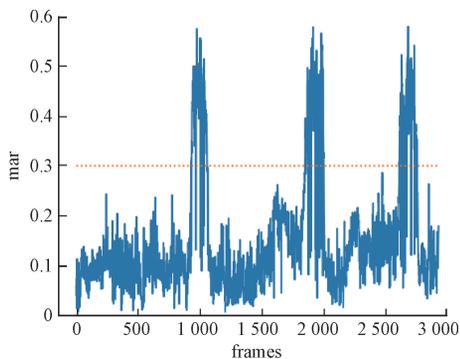


图8 MAR值变化曲线图

Fig. 8 MAR value change curve graph

表5 MAR值实验结果

Table 5 Experimental results of MAR values

组别	第1次	第2次	第3次
第1组	0.18/0.48	0.13/0.60	0.16/0.43
第2组	0.16/0.54	0.19/0.56	0.20/0.63
第3组	0.13/0.42	0.15/0.46	0.11/0.53
第4组	0.13/0.51	0.19/0.49	0.14/0.46
第5组	0.21/0.53	0.19/0.62	0.17/0.58

在0.11~0.21范围内。当嘴部处于张嘴打哈欠时, MAR值在0.42~0.63范围内。为了排除说话张嘴带来的误差,嘴部疲劳阈值设为0.3。

5.4 疲劳判定实验与分析

利用单一疲劳特征判定驾驶员是否疲劳存在一定的局限性,且准确率低。因此,为了提高疲劳检测和判定的准确性,本文采用眼部与嘴部疲劳特征相结合的方法。

为验证本文驾驶疲劳检测模型的性能优势,实验采用YawDD数据集中摄像头固定在后视镜上采集的视频片段,共有320段视频,其中涵盖了驾驶员正常、交谈和疲劳3种典型状态。为了确保实验的准确性,剔除了用于EAR阈值设定的部分视频,最终从中挑选了224段视频。这些视频中涵盖了不同年龄和性别的志愿者,且所有视频均在模拟驾驶环境中录制,以确保实验数据的多样性和实际应用的广泛性。

利用本文提出的疲劳驾驶检测方法对上述224段视频进行了测试,测试结果如表6所示。根据表6可知,本文提出的疲劳驾驶检测方法在整体准确率上表现较好。

表6 YawDD视频数据集检测结果

Table 6 YawDD video dataset detection results

视频类别	视频数量	疲劳数量	检测数量	准确率/%
正常	64	0	2	96.88
交谈	60	0	4	93.33
疲劳	100	100	3	97.00

为了进一步验证本文提出疲劳检测模型的先进性及有效性,本文进行了对比实验,本次对比实验选择了SSD^[3]、YOLOv7^[31]、YOLOv5^[32]、YOLOv7-TMC^[33]、CBAM-YOLOv5s^[34]等主流的疲劳检测模型。由表7可以看出,通过对比不同模型的实验结果,本文提出的疲劳检测模型相较于其他模型在准确率方面表现出领先水平。实验结果证明:本文模型在与各种主流模型的对比中展现出显著优势,验证了本文方法的有效性和性能优势。

表7 疲劳检测模型对比实验

Table 7 Comparative experiments of fatigue detection models

模型	准确率/%
SSD	94.60
YOLOv7	95.35
YOLOv5	95.60
CBAM-YOLOv5s	95.60
YOLOv7-TMC	96.10
本文模型	97.00

图9展示了驾驶员模拟疲劳状态时的检测结果,当驾驶员处于疲劳状态时,本文所提方法能够精确检测并及时发出预警。



图9 疲劳判定结果示例图

Fig. 9 Example of fatigue determination results

6 结论

针对MTCNN模型在面对强烈光照变化、复杂背景时检测精度下降的局限性,本文提出了一种改进的MTCNN人脸检测模型,通过引入CA模块和BN算法对其进行改进,提高模型对驾驶员面部的定位精度,提升网络的收敛

速度和稳定性,增强对过拟合的抑制。实验结果表明:改进后的 MTCNN 模型检测准确率达到 98.78%,相较于原始网络提升了 2.43%。改进后的模型不仅提升了检测精度,还符合轻量级模型的标准。未来的研究可以将改进模型部署至移动嵌入式设备,以提高模型的实用性。

通过计算 EAR 值和 MAR 值获取眼部及嘴部状态,并进行了相关部位的疲劳阈值设定实验,以确定合适的疲劳阈值。综合分析眼部和嘴部状态,判断驾驶员是否处于疲劳状态。根据 YawDD 视频数据集上的实验结果,在面部无遮挡的条件下,该方法能够实现较高的疲劳检测准确率。

由于驾驶者存在面部特征多样性及行为模式差异性,导致基于通用型阈值参数的评估标准存在适用性局限。因此,未来的研究应考虑引入个性化特征学习机制,建立动态参数自适应调整模型,以提升状态判别的个体适配性。此外,探索在复杂环境中能够有效实现疲劳检测也是本文未来的重要研究方向之一。

参考文献

- [1] ZHOU H N, DAI H. Research on fatigue driving detection based on deep learning [J]. *Engineering Letters*, 2025, 33(2): 348-356.
- [2] 罗通强,李仰光,刘坚坚,等. 驾驶员疲劳监测技术研究现状及发展趋势[J]. *中国汽车*, 2024(5): 25-31.
LUO T Q, LI Y G, LIU J J, et al. Research status and development trend of driver fatigue monitoring technology[J]. *China Automotive*, 2024(5): 25-31.
- [3] 吕秀丽,刘希凤,白永强. 基于 SSD 的多因素融合的驾驶疲劳检测研究[J]. *电子测量技术*, 2022, 45(15): 138-143.
LYU X L, LIU X F, BAI Y Q. Research on driving fatigue detection based on multi-factor fusion with SSD[J]. *Electronic Measurement Technology*, 2022, 45(15): 138-143.
- [4] SALEEM A A, SIDDIQUI H U R, RAZA M A, et al. A systematic review of physiological signals based driver drowsiness detection systems [J]. *Cognitive Neurodynamics*, 2023, 17(5): 1229-1259.
- [5] 张瑞,朱天军,邹志亮,等. 驾驶员疲劳驾驶检测方法研究综述[J]. *计算机工程与应用*, 2022, 58(21): 53-66.
ZHANG R, ZHU T J, ZOU ZH L, et al. Research review on driver fatigue detection methods [J]. *Computer Engineering and Applications*, 2022, 58(21): 53-66.
- [6] DOGAN S, TUNCER I, BAYGIN M, et al. A new hand-modeled learning framework for driving fatigue detection using EEG signals[J]. *Neural Computing and Applications*, 2023, 35(20): 14837-14854.
- [7] WANG F W, CHEN D P, ZHANG X L. Real-time driving fatigue detection of ECG signals acquired based on novel electrodes using wavelet scattering networks [J]. *Measurement*, 2025, 243(4): 116438-116452.
- [8] 王琳,化成城,姜鑫,等. 基于颈腰部肌电及脑电信号的疲劳驾驶检测[J]. *东北大学学报(自然科学版)*, 2018, 39(1): 102-107.
WANG L, HUA CH CH, JIANG X, et al. Fatigue driving detection based on cervical and lumbar EMG and EEG signals [J]. *Journal of Northeastern University(Natural Science Edition)*, 2018, 39(1): 102-107.
- [9] 于立娇,吴振昕,王文彬,等. 驾驶员疲劳状态监测系统综述[J]. *汽车文摘*, 2019(3): 24-30.
YU L J, WU ZH X, WANG W B, et al. A review of driver fatigue state monitoring systems [J]. *Automotive Digest*, 2019(3): 24-30.
- [10] MCDONALD A D, LEE J D, SCHWARZ C, et al. A contextual and temporal algorithm for driver drowsiness detection [J]. *Accident Analysis & Prevention*, 2018, 113(4): 25-37.
- [11] 蔡素贤,杜超坎,周思毅,等. 基于车辆运行数据的疲劳驾驶状态检测[J]. *交通运输系统工程与信息*, 2020, 20(4): 77-82.
CAI S X, DU CH K, ZHOU S Y, et al. Fatigue driving state detection based on vehicle operation data [J]. *Transport Systems Engineering and Information*, 2020, 20(4): 77-82.
- [12] 杨艳艳,李雷孝,林浩. 提取驾驶员面部特征的疲劳驾驶检测研究综述[J]. *计算机科学与探索*, 2023, 17(6): 1249-1267.
YANG Y Y, LI L X, LIN H. A review on fatigue driving detection by extracting drivers' facial features [J]. *Computer Science and Exploration*, 2023, 17(6): 1249-1267.
- [13] 史瑞鹏,钱屹,蒋丹妮. 一种基于卷积神经网络的疲劳驾驶检测方法[J]. *计算机应用研究*, 2020, 37(11): 3481-3486.
SHI R P, QIAN Y, JIANG D N. A fatigue driving detection method based on convolutional neural network[J]. *Computer Application Research*, 2020, 37(11): 3481-3486.
- [14] 薛丽. 基于改进 MTCNN 网络的多特征疲劳驾驶检测研究[D]. 西安:长安大学, 2021.
XUE L. Research on multi-feature fatigue driving detection based on improved MTCNN network [D]. Xi'an: Chang'an University, 2021.
- [15] ZHAO Z P, ZHOU N N, ZHANG L, et al. Driver fatigue detection based on convolutional neural networks using EM-CNN [J]. *Computational*

- Intelligence and Neuroscience, 2020, 2020(1): 7251280-7251290.
- [16] 王诗皓. 基于深度学习的疲劳驾驶检测技术研究[J]. 时代汽车, 2024(17): 196-198.
WANG SH H. Research on fatigue driving detection technology based on deep learning [J]. Times Automotive, 2024(17): 196-198.
- [17] ZHANG N, LUO J M, GAO W Q. Research on face detection technology based on MTCNN [C]. International Conference on Computer Network, Electronic and Automation, 2020: 154-158.
- [18] HOSANG J, BENENSON R, SCHIELE B. Learning non-maximum suppression[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2017: 4507-4515.
- [19] HOU Q B, ZHOU D Q, FENG J SH. Coordinate attention for efficient mobile network design [C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition, Nashville, 2021: 13713-13722.
- [20] BJORCK N, GOMES C P, SELMAN B, et al. Understanding batch normalization[C]. International Conference on Neural Information Processing Systems, 2018: 7705-7716.
- [21] 陈之坤, 迟万达, 高尚, 等. 基于深度学习的疲劳检测算法[J]. 计算机与网络, 2020, 46(8): 68-71.
CHEN ZH K, CHI W D, GAO SH, et al. A fatigue detection algorithm based on deep learning [J]. Computers and Networks, 2020, 46(8): 68-71.
- [22] DENG J K, GUO J, ZHOU Y X, et al. Retinaface: single-stage dense face localisation in the wild [J]. ArXiv preprint arXiv: 1905.00641, 2019.
- [23] 孙涵, 田野, 孙春风. 基于改进YOLOv8s的密集多人脸检测[J]. 软件, 2024, 45(4): 142-146.
SUN H, TIAN Y, SUN CH F. Dense multiple face detection based on improved YOLOv8s[J]. Software, 2024, 45(4): 142-146.
- [24] 宋传旗. YOLOv5算法的人脸识别检测方法研究[J]. 计算机时代, 2023(7): 15-19.
SONG CH Q. Research on face recognition detection method of YOLOv5 algorithm [J]. Computer Age, 2023(7): 15-19.
- [25] JIANG B, JIANG H B, ZHANG H L, et al. 4AC-YOLOv5: An improved algorithm for small target face detection [J]. EURASIP Journal on Image and Video Processing, 2024, 2024(1): 10-23.
- [26] GUO J, DENG J K, LATTAS A, et al. Sample and computation redistribution for efficient face detection [C]. International Conference on Learning Representations, 2022: 1-17.
- [27] RAHMAN S, RAHMAN M M, ABDULLAH-AL-WADUD M, et al. An adaptive gamma correction for image enhancement [J]. EURASIP Journal on Image and Video Processing, 2016, 2016(1): 35-47.
- [28] GUO X J, LI S Y, YU J K, et al. PFLD: A practical facial landmark detector [J]. ArXiv preprint arXiv: 1902.10859, 2019.
- [29] SOUKUPOVA T, CECH J. Eye blink detection using facial landmarks [C]. Computer Vision Winter Workshop, 2016.
- [30] SAVAŞ B K, BECERIKLI Y. Real time driver fatigue detection based on SVM algorithm [C]. International Conference on Control Engineering & Information Technology, 2018: 1-4.
- [31] 李威, 张婧. 基于改进YOLOv7的疲劳驾驶检测算法研究[J]. 信息与电脑(理论版), 2023, 35(24): 64-66.
LI W, ZHANG J. Research on fatigue driving detection algorithm based on improved YOLOv7 [J]. Information and Computer (Theoretical Edition), 2023, 35(24): 64-66.
- [32] 鲁佳儒, 胡文勋, 肖运虹. 基于改进YOLOv5的人脸疲劳检测[J]. 信息与电脑(理论版), 2023, 35(7): 111-114.
LU J R, HU W X, XIAO Y H. Face fatigue detection based on improved YOLOv5 [J]. Information and Computer (Theoretical Edition), 2023, 35(7): 111-114.
- [33] 马明. 基于轻量型YOLOv7-TMC网络的疲劳驾驶状态检测方法[J]. 电脑与信息技术, 2024, 32(6): 6-11.
MA M. Fatigue driving state detection method based on lightweight YOLOv7-TMC network [J]. Computer and Information Technology, 2024, 32(6): 6-11.
- [34] 金云峰, 路志展, 王瑞利, 等. 基于改进YOLOv5s的疲劳驾驶检测[J]. 北华大学学报(自然科学版), 2024, 25(2): 255-261.
JIN Y F, LU ZH ZH, WANG R L, et al. Fatigue driving detection based on improved YOLOv5s [J]. Journal of Beihua University (Natural Science Edition), 2024, 25(2): 255-261.

作者简介

宋志强(通信作者), 博士, 教授, 主要研究方向为多智能系统协同与控制、计算机应用。

E-mail: zqsong@cw Xu. edu. cn

李明阳, 硕士研究生, 主要研究方向为视觉识别、多智能体系统协作与控制等。

E-mail: 202212490507@nuist. edu. cn

周鹏, 硕士研究生, 主要研究方向为无人驾驶目标检测, 无人驾驶环境感知等。

E-mail: 202312490378@nuist. edu. cn