



激光与光电子学进展  
*Laser & Optoelectronics Progress*  
ISSN 1006-4125, CN 31-1690/TN

## 《激光与光电子学进展》网络首发论文

题目：低照度下基于图像增强和人脸状态识别的疲劳驾驶检测  
作者：赵洋，苗佳龙，刘雪枫，赵锦程，徐森  
收稿日期：2024-02-20  
网络首发日期：2024-04-24  
引用格式：赵洋，苗佳龙，刘雪枫，赵锦程，徐森. 低照度下基于图像增强和人脸状态识别的疲劳驾驶检测[J/OL]. 激光与光电子学进展.  
<https://link.cnki.net/urlid/31.1690.TN.20240422.1635.098>



**网络首发：**在编辑部工作流程中，稿件从录用到出版要经历录用定稿、排版定稿、整期汇编定稿等阶段。录用定稿指内容已经确定，且通过同行评议、主编终审同意刊用的稿件。排版定稿指录用定稿按照期刊特定版式（包括网络呈现版式）排版后的稿件，可暂不确定出版年、卷、期和页码。整期汇编定稿指出版年、卷、期、页码均已确定的印刷或数字出版的整期汇编稿件。录用定稿网络首发稿件内容必须符合《出版管理条例》和《期刊出版管理规定》的有关规定；学术研究成果具有创新性、科学性和先进性，符合编辑部对刊文的录用要求，不存在学术不端行为及其他侵权行为；稿件内容应基本符合国家有关书刊编辑、出版的技术标准，正确使用和统一规范语言文字、符号、数字、外文字母、法定计量单位及地图标注等。为确保录用定稿网络首发的严肃性，录用定稿一经发布，不得修改论文题目、作者、机构名称和学术内容，只可基于编辑规范进行少量文字的修改。

**出版确认：**纸质期刊编辑部通过与《中国学术期刊（光盘版）》电子杂志社有限公司签约，在《中国学术期刊（网络版）》出版传播平台上创办与纸质期刊内容一致的网络版，以单篇或整期出版形式，在印刷出版之前刊发论文的录用定稿、排版定稿、整期汇编定稿。因为《中国学术期刊（网络版）》是国家新闻出版广电总局批准的网络连续型出版物（ISSN 2096-4188，CN 11-6037/Z），所以签约期刊的网络版上网络首发论文视为正式出版。

# 低照度下基于图像增强和人脸状态识别的疲劳驾驶检测

赵洋<sup>1,2</sup>, 苗佳龙<sup>1\*</sup>, 刘雪枫<sup>1</sup>, 赵锦程<sup>3</sup>, 徐森<sup>1,2</sup>

<sup>1</sup> 沈阳化工大学计算机科学与技术学院, 辽宁 沈阳 110142;

<sup>2</sup> 辽宁省化工过程工业智能化技术重点实验室, 辽宁 沈阳 110142;

<sup>3</sup> 沈阳化工大学信息工程学院, 辽宁 沈阳 110142

**摘要** 针对现有的疲劳驾驶检测模型存在精度低、模型体积较大且在低光环境下性能降低等问题, 提出一种基于图像增强和人脸状态识别的疲劳驾驶检测模型。首先通过改进 YOLOv5s 模型用于低照度下人脸检测和关键点定位, 在 YOLOv5s 中引入 SCI(Self-Calibrated Illumination) 模块对低照度图像进行图像增强; 下采样层替换为 StemBlock 模块, 提高特征表达能力; 主干网络替换为 ShuffleNetv2 网络, 大幅降低模型的参数量和计算量; 提出 CIBC3 (cbam inverted bottleneck) 模块代替 C3 模块, 降低噪声对检测的干扰, 增强模型的全局感知能力; 总损失函数中添加 Wingloss 损失函数用于人脸关键点回归。之后使用疲劳状态识别网络判断人脸检测模型定位到的眼嘴部位的开闭状态, 最后使用评价指标判断疲劳状态。改进后的 YOLOv5s 模型在 DARK FACE 数据集上的实验表明, 在参数量和计算量上分别下降了 62.12%、63.41% 的同时精度提高了 2.38%。提出的疲劳驾驶检测模型在 YawDD 正常光照数据集与自建低光照数据集上分别达到了 96.07%, 94.5% 的准确率, 优于其他模型且处理单张图像的时间为 27ms。验证了提出的疲劳驾驶检测模型在保证正常环境与低光环境下的检测准确率的同时满足实时性要求, 且具备能部署在算力有限的边缘计算设备上的能力。

**关键词** 图像处理; 疲劳驾驶检测; 人脸状态识别; YOLOv5s; 轻量化算法

**中图分类号** TP391.4 **文献标志码** A

## Fatigue driving detection under low illumination based on image enhancement and face state recognition

Zhao Yang<sup>1,2</sup>, Miao Jialong<sup>1\*</sup>, Liu Xuefeng<sup>1</sup>, Zhao Jincheng<sup>3</sup>, Xu Sen<sup>1,2</sup>

<sup>1</sup> The College of Computer Science and Technology, Shenyang University of Chemical Technology, Shenyang 110142, Liaoning China;

<sup>2</sup> Key Laboratory of Intelligent Technology for Chemical Process Industry of Liaoning Province, Shenyang 110142, Liaoning China;

<sup>3</sup> The College of Information Engineering, Shenyang University of Chemical Technology, Shenyang 110142, Liaoning China

**Abstract** To tackle the shortcomings of current fatigue driving detection models, such as their suboptimal accuracy, considerable model size, and diminished performance in low-light conditions, a new fatigue driving detection model based on image enhancement and face state recognition is proposed. Firstly, the YOLOv5s model is improved for face detection and key

**基金项目** 辽宁省教育厅基本科研项目面上项目 (LJKMZ20220782)

**通信作者**: \*E-mail: [jgxiaodingdang@163.com](mailto:jgxiaodingdang@163.com);

point positioning in low light conditions. The self-calibrated illumination (SCI) module is incorporated into YOLOv5s, aiming to enhance images under low illumination. The subsampling layer is replaced with StemBlock module to improve the feature expression ability. The backbone network is replaced by ShuffleNetv2 network, which greatly reduces the number of parameters and calculation amount of the model. The cbam inverted bottleneck(CIBC3) module is proposed to replace the C3 module to reduce the interference of noise to the detection and enhance the global perception ability of the model. Wingloss loss function is added to the total loss function for face key point regression. After that, the fatigue state recognition network is used to judge the opening and closing state of the eye and mouth parts located by the face detection model. Finally the evaluation metrics is used to judge the fatigue state. The experiment of the improved YOLOv5s model on the DARK FACE dataset shows that the number of parameters and the amount of computation decreased by 62.12% and 63.41% respectively, while the accuracy increased by 2.38%. The proposed fatigue driving detection model achieved an accuracy of 96.07% and 94.5% respectively on YawDD normal light dataset and self-built low light dataset, which is better than other models and the processing time of a single image is 27ms. It is verified that the proposed fatigue driving detection model meets the real-time requirements while ensuring the detection accuracy in normal and low light environments, and can be deployed on edge computing devices with limited computing power.

**Key words** image processing; fatigue driving detection; face state recognition; YOLOv5s; lightweight algorithm

## 1 引 言

疲劳驾驶已经成为导致交通事故的一个重要因素<sup>[1]</sup>, 准确检测并阻止疲劳驾驶对减少交通事故与保护人民生命安全具有重要的现实意义。

目前, 针对疲劳驾驶的检测方法主要基于三种方式: 基于车辆行驶信息<sup>[2]</sup>、基于驾驶人生理特征<sup>[3]</sup>和基于驾驶人面部特征<sup>[4]</sup>。相较于前两种方式, 基于驾驶人面部特征的检测方法优势在于准确率高且不影响正常驾驶。因此该方法得到了广泛的关注。

近年来, 由于深度学习在图像处理<sup>[5]</sup>与目标检测<sup>[6]</sup>等领域中取得了巨大的成就。许多研究者将其应用于疲劳驾驶检测中。张志威等<sup>[7]</sup>使用 VGG19 网络来检测不规范驾驶行为。但该模型稳定性较差, 面对特殊情况时无法有效定位到主驾驶人。敖邦乾等<sup>[8]</sup>将 MTCNN 网络与眼、嘴检测网络进行级联用于检测疲劳驾驶。由于该模型检测精度较差难以实际应用。Wang 等<sup>[9]</sup>设计了一款基于深度学习的个性化框架用来判断驾驶人的疲劳状态。但该模型综合性较差, 在检测脸部特征时存在特征不易区分等问题。吕秀丽等<sup>[10]</sup>使用改进后的 SSD 网络, 并结合人脸特征点、眼睛纵横比和嘴巴纵横比用来判断疲劳驾驶。由于该模型体积较大而无法部署在资源有限的边缘计算设备上。

此外,这些模型并未考虑到在夜晚低光环境下会导致成像设备在拍摄图像时存在对比度下降和噪声等情况,这些情况会导致检测性能大幅下降。因此,亟需设计一款准确率高、实时性好且在复杂低光环境下仍能保持良好性能的轻量级疲劳驾驶检测模型。

在低照度环境下,能够快速准确地检测到人脸是提高疲劳驾驶检测性能的核心。不少研究者通过使用图像增强方法提高低照度图像的质量来增加人脸检测精度。Chen 等<sup>[11]</sup>提出图像增强网络并与 SSD 模型进行结合提升了人脸检测性能。SaSagawa 等<sup>[12]</sup>提出 YOLO-in-the-Dark 模型应用于低光照下的人脸检测,该模型是将低光照模型与 YOLO 模型进行融合。Ma 等<sup>[13]</sup>提出自校准照明 SCI (Self-Calibrated Illumination) 模型并与检测模型相结合用于低光照人脸检测。因此本文采用将图像增强方法与目标检测算法相结合的方式用于低光照人脸检测。

综上,本文提出一种基于图像增强和人脸状态识别的疲劳驾驶检测模型。该疲劳驾驶检测模型主要分为三部分:人脸检测及人脸关键点定位、人脸状态识别和疲劳评估。在人脸检测及关键点定位中本文将 YOLOv5s<sup>[14]</sup>模型作为基准模型进行改进,旨在提高人脸检测的准确率、降低模型的参数量与计算量并对人脸关键点定位,同时融合图像增强方法对低照度图像进行处理;在人脸状态识别中本文通过关键点对驾驶人眼、嘴部位进行定位分割,然后将各部位图像分别送入疲劳状态识别网络进行检测;在疲劳评估中本文设计了两个评价指标用于判断疲劳状态。

## 2 人脸检测及关键点定位模型

### 2.1 YOLOv5s 模型

YOLOv5s 模型以泛化性能好、准确率高等优点,被广泛应用于人脸检测中。本文选用 YOLOv5s 模型作为基础网络进行改进。

模型共由输入(input)、主干网络(Backbone)、颈部网络(Neck)、检测层(Head)四部分组成。输入端主要采用 Mosaic 数据增强、自适应锚框计算、自适应图片缩放等方法对输入图像进行预处理;主干网络负责特征提取,主要包括 CBS (ConvBNSILU) 模块、C3 模块和 SPPF 模块;颈部网络采用特征金字塔与路径聚合结构,在缩短信息传输路径的同时为整个特征融合网络增加丰富特征信息。检测层用于识别图像中的物体,该层采用非极大值抑制对检测的目标框进行筛选,增强对目标的检测能力。模型结构见图 1。

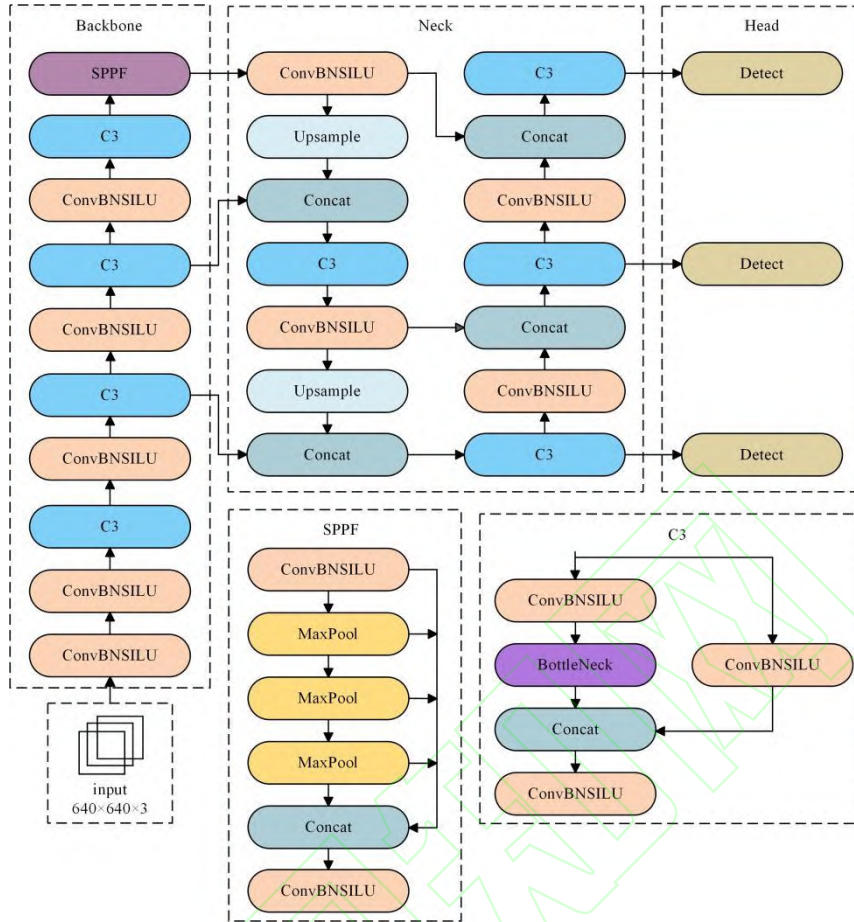


图1 YOLOv5s模型结构  
Fig. 1 YOLOv5s model structure

## 2.2 图像增强模块

在低光照环境下，成像设备拍摄出的图像往往会存在对比度下降、亮度低和噪声干扰等情况，这导致模型难以提取到有效特征，检测精度会大幅下降。图像增强算法在最大程度上保留特征信息的同时，可以提高图像的对比度和亮度、降低噪声产生的影响。因此本文将图像增强方法与 YOLOv5s 模型进行融合，达到提高检测精度的目的。本文选取三种图像增强方法对图像进行增强，其可视化结果见图 2。





图2 各图像增强方法后的可视化结果。(a) 原图；(b) LIME；(c) EnlightenGAN；(d) SCI  
Fig. 2 The visualization results of each image enhancement method. (a) Image; (b) LIME; (c) EnlightenGAN; (d) SCI

如图 2 所示，LIME<sup>[15]</sup> (Low-Light Image Enhancement) 方法在提高对比度的同时放大了噪声；EnlightenGAN<sup>[16]</sup> (Enlighten Generative Adversarial Network) 方法在图像上增强的效果很不均匀，仍存在一些阴影区域；SCI 方法处理后的图像整体效果较好，细节较为清晰。本文采用 SCI 图像增强方法作为增强模块对低照度图像进行图像增强，同时本文将在 5.1.3 小节进行具体实验。

### 2.3 StemBlock 模块

YOLOv5s 模型中的原下采样操作是通过两个卷积层实现的，该操作特征提取效率较低。本文采用 StemBlock<sup>[17]</sup> 模块替换原下采样操作，在减少参数量与计算量的同时，可以有效地提高特征表达能力。该模块见图 3。

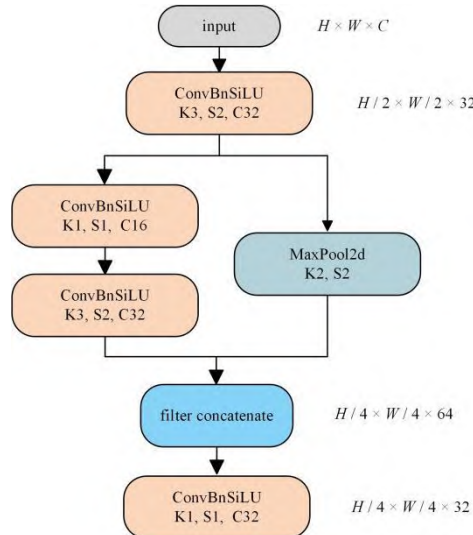


图3 StemBlock模块

Fig. 3 StemBlock module

StemBlock 模块首先将输入特征图经过  $3 \times 3$  的卷积用于下采样操作并改变通道数。之后分为两个支路，一个支路分别经过  $1 \times 1$  和  $3 \times 3$  的卷积用于降低通道数和下采样操作，该分支引入瓶颈层目的在于减少模型参数量；另一支路进行最大池化操作，目的是防止特征图的特征信息丢失。最后将这两个支路的输出结果进行合并后再经过一个  $1 \times 1$  的卷积还原通道数。因此该模块在保证减少参数量和计算量的同时能够保留足够的特征信息，不会造成信息的过度损失。

## 2.4 主干网络改进

YOLOv5s 模型虽然检测性能良好，但由于其参数量计算量较大导致无法满足实时性。近年来，研究人员提出了许多轻量化模型，其中 ShuffleNetV2<sup>[18]</sup> 模型在降低模型的参数量和计算量的同时，具备高效提取特征的能力。本文采用 ShuffleNetV2 模型替换原主干网络，实现模型的轻量化。

ShuffleNetV2 模型有两个模块，分别是基本模块和下采样模块，具体结构见图 4。

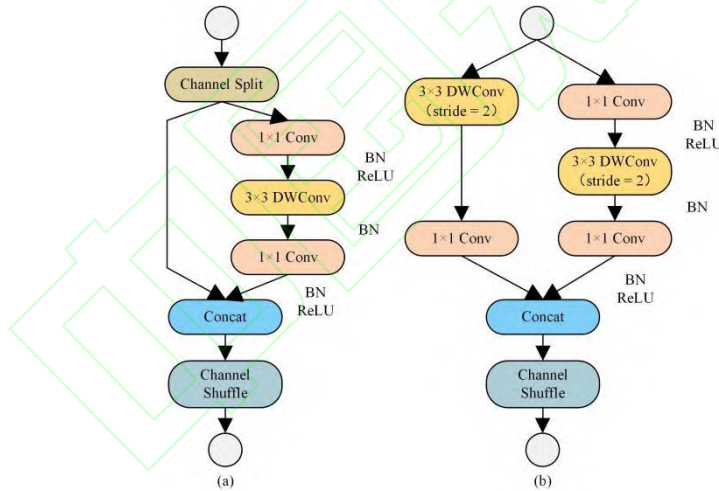


图4 ShuffleNetV2模块。(a) 基本模块S\_block1；(b) 下采样模块S\_block2

Fig. 4 ShuffleNetV2 module. (a) Basic module S\_block1; (b) subsampling module S\_block2

基本模块通过引入通道切分操作，将输入特征通道数分成两部分。其中一个分支不做任何操作，另一个分支进行卷积操作和批量归一化等处理。之后将两个分支进行特征融合，最后通过通道重排进行不同组之间的特征融合。

下采样模块由于没有通道切分操作，两个分支在经过卷积操作和批量归一化等处理并进行融合之后，特征图的大小将减半，输出通道数为输入通道数的两倍。

## 2.5 改进 C3 模块

YOLOv5s 模型中的 C3 模块是一种特征提取模块。然而随着模型的网络层数加深，会出现特征信息丢失的情况。同时在低光环境下，很容易出现漏检误检。针对上述问题，本文提出了 CIBC3 (cbam inverted bottleneck C3) 模块。

### 1) 反向瓶颈模块

瓶颈模块是通过  $1 \times 1$  与  $3 \times 3$  的卷积核进行特征提取。这种方式可能会存在信息丢失问题，使得模型精度降低。因此本文引入反向瓶颈模块，该模块的思想是将特征提取阶段提前并采取大卷积核增大感受野，这种方式可以减少特征信息的丢失并可以有效地提取到丰富的特征；同时利用深度卷积 DWConv (depthwise convolution) 与点卷积 PWConv (pointwise convolution) 减少计算成本。瓶颈模块和反向瓶颈模块见图 5。

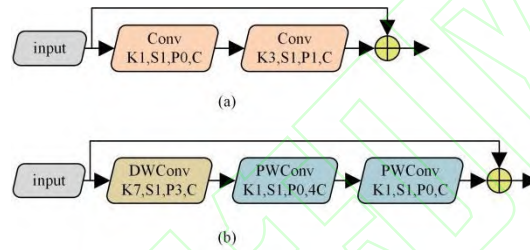


图5 瓶颈模块与反向瓶颈模块。(a) 瓶颈模块；(b) 反向瓶颈模块

Fig. 5 Bottleneck module and inverted bottleneck module. (a) Bottleneck module; (b) inverted bottleneck module

### 2) CBAM 注意力机制

由于实际的检测环境更加复杂，很容易出现误检和漏检的情况。计算机视觉中的注意力机制的功能是让模型学会关注图像中的重点信息，忽略无关信息。因此本文引入 CBAM<sup>[19]</sup> 注意力机制，CBAM 注意力机制将空间和通道结合在一起，强调空间和通道两个维度上的关键特征用来增强模型的全局感知能力，同时调整特征图的权重，使模型能够提取到有用的特征信息从而提高模型精度。其结构见图 6。

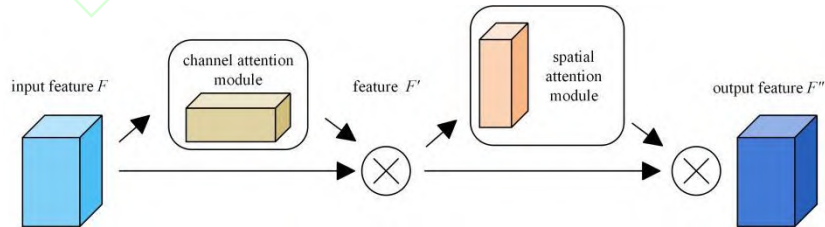


图6 CBAM结构

Fig. 6 The structure of CBAM

### 3) 构建 CIBC3 模块

本文首先将 CBAM 注意力机制嵌入反瓶颈模块中构建 CIB (cbam inverted bottleneck) 模块，然后将 C3 中的瓶颈模块替换为 CIB 模块形成 CIBC3 模块。CIB 模块见图 7，CIBC3 模



块见图 8。

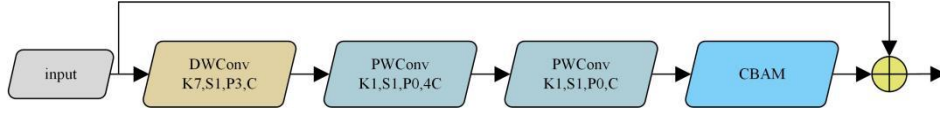


图7 CIB模块

Fig. 7 CIB module

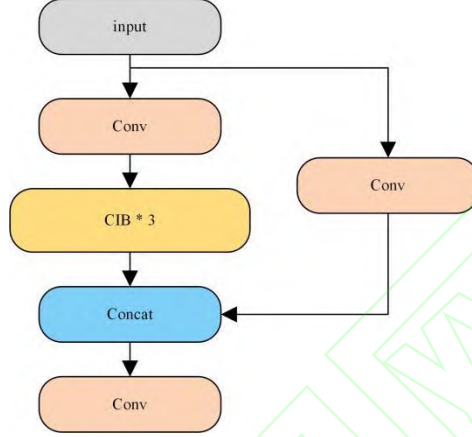


图8 CIBC3模块

Fig. 8 CIBC3 module

CIBC3 模块可以保留更多的特征信息并能增强模型的全局感知能力，使其可以有效地提取到丰富的特征。

## 2.6 人脸关键点检测

如何准确地检测到人脸状态很大程度上依赖于精确的人脸关键点所提供的人脸面部组件特殊位置。传统的人脸检测关键点由 68 个点组成<sup>[20]</sup>，但在实际检测过程中发现仅获取驾驶人眼嘴部位就可以判断是否疲劳。因此本文只选取其中 4 个关键点，分别为右、左眼中心点，右、左嘴角点。由于 YOLOv5s 模型没有人脸关键点预测，无法将其直接应用，本文设计了一种用于关键点回归的翼型分段损失函数 Wingloss，如式 (1) 所示。

$$wing(x) = \begin{cases} w \cdot \log_e(1 + |x|/e), & \text{if } x < w \\ |x| - C, & \text{othersize} \end{cases}, \quad (1)$$

式中： $wing(\cdot)$  为损失函数； $(-w, w)$  ( $w$  为非负数) 是 Wingloss 损失函数的非线性区； $e$  为约束非线性区域的曲率； $C = w - w \log_e(1 + w/e)$  是一个常数，用于平滑的连接分段的线性部分和非线性部分。

Wingloss 函数在不同参数下的图像见图 9。

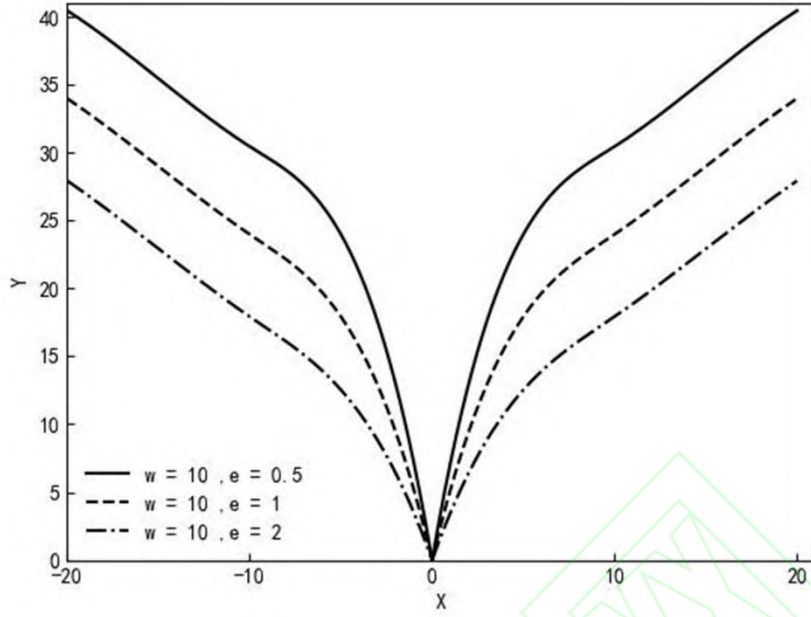


图9 Wingloss在不同参数下的函数图像

Fig. 9 Wingloss function image with different parameters

关键点的坐标值  $s = \{s_i\}$  与其预测值  $s' = \{s'_i\}$  的损失函数见式 (2)。

$$loss_L = \sum_i wing(s_i - s'_i), \quad (2)$$

式中： $loss_L$  为关键点损失函数； $s_i (i=1,2,\dots,n)$  为第  $i$  个坐标点的真实值；

$s'_i (i=1,2,\dots,n)$  为第  $i$  个坐标点的预测值。

将该损失函数加入 YOLOv5s 模型总损失函数  $loss_O$  中，YOLOv5s 模型新的总损失函数见式 (3)。

$$loss = loss_O + \lambda \cdot loss_L, \quad (3)$$

式中： $loss$  为 YOLOv5s 模型新的总损失函数； $loss_O$  为 YOLOv5s 模型原损失函数； $\lambda$  为关键点损失函数的权重因子。

## 2.7 改进后的 YOLOv5s 模型

本文以 YOLOv5s 模型为基础网络进行改进，首先将输入的低照度图像经过图像增强模块 SCI 进行增强处理，其次将下采样层替换为 StemBlock 模块，在减少参数量与计算复杂度的同时，可以有效地提高特征表达能力；然后将主干网络替换为 ShuffleNetV2 网络，实现模型的轻量化；之后提出 CIBC3 模块代替 C3 模块，可以保留更多的特征信息并能增强模型的

全局感知能力；最后将 Wingloss 损失函数加 YOLOv5s 总损失函数中用于人脸关键点回归。

综上，改进后的 YOLOv5s 模型结构见图 10，其检测人脸的结果示例见图 11。

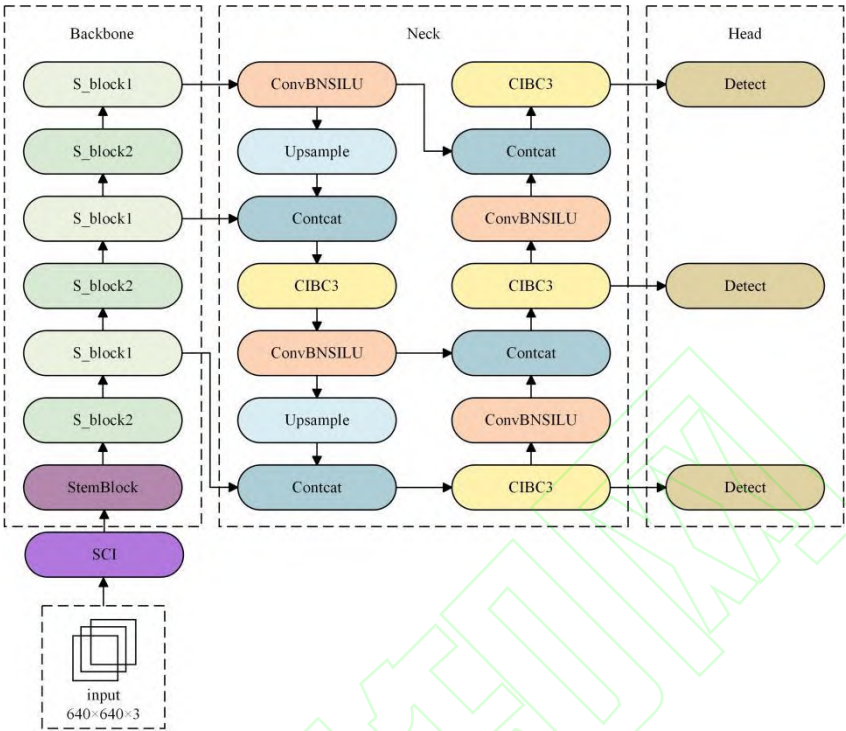


图10 改进后的YOLOv5s模型  
Fig. 10 Improved YOLOv5s model

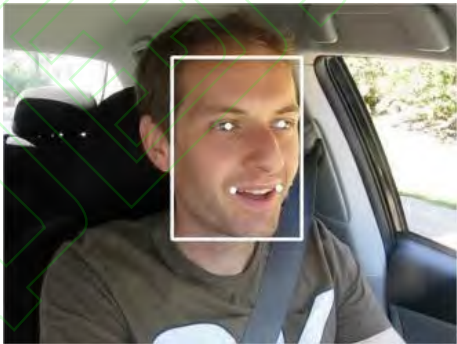


图11 结果实例  
Fig. 11 Result instance

### 3 人脸状态识别

#### 3.1 截取眼部和嘴部区域

在实际的图像采集过程中，当驾驶人头部发生偏转时，所获取到的眼部区域也有一定倾斜，这势必增大了检测难度。在将眼部图像送入分类网络之前，需要将眼部图像旋转至水平。

头部倾斜后的眼部示意图见图 12，双眼距离与截取图像宽高之间的关系见式（4）。本

文将通过图 12 与式（4）来截取眼部区域图像。

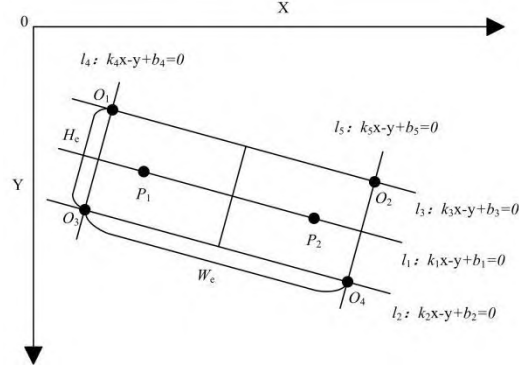


图12 眼部区域示意图

Fig. 12 Schematic diagram of eye area

$$\begin{cases} d_e = \sqrt{(x_1^{(e)} - x_r^{(e)})^2 + (y_1^{(e)} - y_r^{(e)})^2} \\ W_e = \frac{5}{3} \times d_e \\ H_e = \frac{10}{9} \times d_e \end{cases}, \quad (4)$$

式中：  $P_1(x_1^{(e)}, y_1^{(e)})$ 、 $P_2(x_r^{(e)}, y_r^{(e)})$  为左、右眼中心点坐标；  $d_e$  为双眼中心相距距离；  $W_e$ 、 $H_e$  分别为截取的眼部图像宽度和高度。利用文献[21]中的研究结果，本文设置截取图像的宽度为双眼中心距离的 5/3，高度为双眼中心距离的 10/9。

可以根据图 12 中  $l_1$ 、 $l_2$ 、 $l_3$ 、 $l_4$ 、 $l_5$  计算出眼部图像的四个边界点  $O_1$ 、 $O_2$ 、 $O_3$ 、 $O_4$ 。

$l_1$ 、 $l_2$ 、 $l_3$ 、 $l_4$ 、 $l_5$  公式见式（5）。

$$l_n : k_n x - y + b_n = 0, \quad (5)$$

式中：  $k_n (n=1,2,\dots,5)$  为  $l_n$  的斜率；  $b_n$  为  $l_n$  的截距。

$l_1$  计算见式（6）。

$$\begin{cases} k_1 = \frac{y_1^{(e)} - y_r^{(e)}}{x_1^{(e)} - x_r^{(e)}} \\ b_1 = y_1^{(e)} - k_1 \cdot x_1^{(e)} \end{cases}, \quad (6)$$

由于  $l_2$ 、 $l_3$  与  $l_1$  平行且距离为  $5/9 d$ ，因此  $l_2$ 、 $l_3$  的斜率与  $l_1$  相等，即为  $k_1$ ， $b_1$ 、 $b_2$  计算分别见式（7）、（8）。

$$b_2 = b_1 + \frac{5}{9} d_e \cdot \sqrt{k_1^2 + 1}, \quad (7)$$

$$b_3 = b_1 - \frac{5}{9}d_e \cdot \sqrt{k_1^2 + 1}, \quad (8)$$

由于  $l_4$ 、 $l_5$  与  $l_1$  垂直，因此斜率为  $-1/k_1$ ， $b_4$ 、 $b_5$  计算分别见式 (9)、(10)。

$$b_4 = \frac{y_1^{(e)} + y_r^{(e)}}{2} + \frac{x_1^{(e)} + x_r^{(e)}}{2k_1} - \frac{5}{6}d_e \cdot \sqrt{\frac{1}{k_1^2} + 1}, \quad (9)$$

$$b_5 = \frac{y_1^{(e)} + y_r^{(e)}}{2} + \frac{x_1^{(e)} + x_r^{(e)}}{2k_1} + \frac{5}{6}d_e \cdot \sqrt{\frac{1}{k_1^2} + 1}, \quad (10)$$

由此可以得到  $l_1$ 、 $l_2$ 、 $l_3$ 、 $l_4$ 、 $l_5$  五条线段，其中截取图像左上角点  $O_1$  是  $l_3$  与  $l_4$  交点， $O_1$  计算见式 (11)，其余三个点同理可得。

$$\begin{cases} k_1 x_{O1} - y_{O1} + b_3 = 0 \\ k_2 x_{O1} - y_{O1} + b_4 = 0 \end{cases}, \quad (11)$$

式中： $x_{O1}$ 、 $y_{O1}$  分别为要截取图像左上角点  $O_1$  的横、纵坐标值。

根据这四个点可以将眼部区域确定，并旋转原图像至水平后进行截取。

嘴角相距距离与截取图像宽高相应关系见式 (12)。

$$\begin{cases} d_m = \sqrt{(x_1^{(m)} - x_r^{(m)})^2 + (y_1^{(m)} - y_r^{(m)})^2} \\ W_m = \frac{6}{5} \times d_m \\ H_m = \frac{6}{5} \times d_m \end{cases}, \quad (12)$$

式中： $x_1^{(m)}$ 、 $y_1^{(m)}$  为左嘴角中心点横、纵坐标值； $x_r^{(m)}$ 、 $y_r^{(m)}$  为右嘴角中心点横、纵坐标值； $d_m$  为左右嘴角相距距离； $W_m$ 、 $H_m$  分别为截取的嘴部图像宽度和高度。

采用上述相同原理完成嘴部区域的截取。最终，眼部和嘴部的截取图像示意图见图 13。



图13 眼部嘴部截取图像示意图。(a) 眼部截取图像；(b) 嘴部截取图像



Fig. 13 Schematic diagram of an image capture of the eye and mouth. (a) Cropped image of the eye region; (b) cropped image of the mouth region

### 3.2 疲劳状态识别网络

在获取眼部与嘴部图像后，本文构建了用于判断眼嘴开闭状态的疲劳状态识别网络 FSR-Net (fatigue state recognition network)。

FSR-Net 采用了残差网络结构，输入图像首先通过 32 个  $3 \times 3$  卷积核生成 32 个特征图，这些特征图经过  $2 \times 2$  的最大池化层进行降维处理。上述特征图通过 64 个  $3 \times 3$  卷积核做进一步的特征提取，特征提取完后经过一个残差块，生成 64 个特征图。经过  $2 \times 2$  的平均池化层进行降维处理，上述特征图在经过由 256 个神经元组成的全连接层，用于接收前层提取的特征。最后通过 softmax 层对输入图像进行二分类，即眼睛和嘴巴张开或闭合的状态。网络结构见图 14。

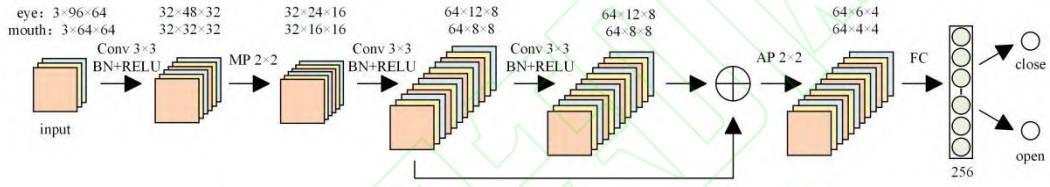


图14 FSR-Net结构

Fig. 14 The structure of FSR-Net

## 4 疲劳评估与疲劳驾驶检测模型

本文提出两个评价指标用于对疲劳状态进行评估，分别为最长闭眼时间 MCT (maximum closing time)、最长打哈欠时间 MYD (maximum yawn duration)。

### 4.1 最长闭眼时间指标

在实际驾驶过程中，驾驶人会因为疲劳而长时间闭眼，这种情况应该被预警。本文提出最长闭眼时间指标，当驾驶人开始闭眼时，计算连续闭眼的时间见式 (13)。

$$T_e = (E_{\text{end}} - E_{\text{start}}) \times T, \quad (13)$$

式中： $T_e$  为总共闭眼的时间； $E_{\text{end}}$  表示从连续闭眼到睁眼的帧序列号； $E_{\text{start}}$  表示开始闭眼的帧序列号； $T$  表示采集每帧图像的时间间隔。

本文设定当最长闭眼时间超过 2s 时，驾驶人处于疲劳状态。在数据集中每段视频以平均每秒 20 帧的速率播放，即每帧图像的时间间隔为 50ms，因此当连续闭眼帧数超过 40 帧

后，判定驾驶人当前处于疲劳状态。

#### 4.2 最长打哈欠时间指标

除闭眼外，当驾驶人处于疲劳状态时还会出现打哈欠行为。本文提出最长打哈欠时间指标，计算持续张嘴的时间见式（14）。

$$T_m = (M_{\text{end}} - M_{\text{start}}) \times T, \quad (14)$$

式中：  $T_m$  表示持续张嘴的时间；  $M_{\text{end}}$  表示从持续张嘴到闭嘴的帧序列号；  $M_{\text{start}}$  表示开始张嘴的帧序列号。

本文设定当最长打哈欠时间指标超过 3s 时，驾驶人处于疲劳状态。即当持续张嘴超过 60 帧后，判定驾驶人处于疲劳状态。

#### 4.3 疲劳驾驶检测模型

综上，本文最终提出了疲劳驾驶检测模型，模型结构见图 15。

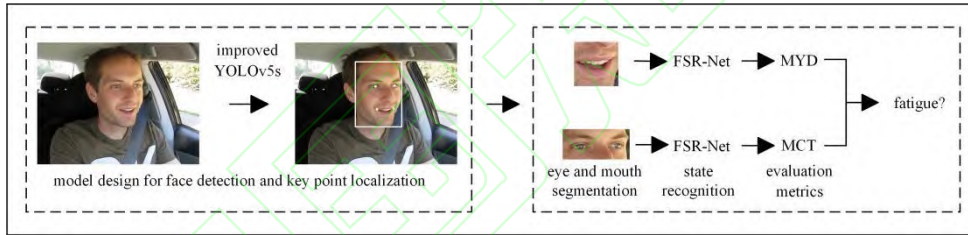


图15 疲劳驾驶检测模型

Fig. 15 Fatigue driving detection model

本文提出的疲劳驾驶检测算法流程如下：

首先将视频进行分帧，转化为帧图像，将帧图像依次输入到疲劳驾驶检测模型中。该模型通过改进后的YOLOv5s模型对输入图像进行人脸检测及关键点定位；随后通过关键点对驾驶人眼、嘴部位进行定位分割；然后将各部位图片分别送入FSR-Net对驾驶人状态进行检测；最后通过MYD、MCT指标判断驾驶人当前是否处于疲劳状态。

## 5 实验与结果分析

本文实验环境及配置如下：Intel(R) Core(TM) i7-9300H 处理器，32GB 运行内存，NVIDIA GeForce RTX 1050Ti 显卡，Win10 操作系统，Pytorch1.13.0。

#### 5.1 改进后的YOLOv5s模型的实验与结果分析

### 5.1.1 实验数据集与参数设置

该实验选用的数据集为 DARK FACE 公共数据集。DARK FACE 数据集中包含了 6000 张真实环境下拍摄的低光照图像，本文将该数据集划分为训练集、验证集、测试集，比例为 5: 1: 4。

参数设置如下：epochs（总迭代次数）为 300；输入图像尺寸为  $640 \times 640$ ；学习率为 0.01；batch（每次迭代的训练样本）为 16；动量为 0.937；权重衰减系数为 0.005；关键点损失函数的权重因子本文通过手动调参的方式最终确定为 0.005。

### 5.1.2 评估指标

评估网络性能的指标有精确率  $P(\text{precision})$ ，召回率  $R(\text{recall})$ ，平均精度  $AP(\text{average precesion})$ ，平均精度均值  $mAP(\text{mean average precesion})$ 。计算公式见式（15）。

$$\begin{aligned} P &= \frac{TP}{TP + FP} \\ R &= \frac{TP}{TP + FN} \\ AP &= \int_0^1 P(R) \\ mAP &= \frac{1}{c} \sum_{j=1}^c AP_j \end{aligned}, \quad (15)$$

式中： $TP$  为预测正确的正样本数量； $FN$  为预测错误的负样本数； $FP$  为预测错误的正样本数； $c$  为总样本数。

### 5.1.3 不同图像增强方法对比

在 2.2 节中本文展示了不同图像增强方法下的可视化结果。为继续探究不同图像增强方法对模型精度的影响，本文设置了在不同图像增强方法下的实验，并将未使用图像增强方法的 YOLOv5s 模型作为基准进行对比。实验结果如表 1 所示，表中“—”处代表未使用图像增强方法。

表 1 各图像增强方法对比实验

Table 1 Comparative test of each image enhancement method

enhancement method	mAP@0.5/%
—	71.36
LIME	70.17
EnlightenGAN	70.33
SCI	72.19

LIME 在增强过程中放大了噪声，EnlightenGAN 增强的效果不是很均匀，仍存在一些阴影区域，这些问题导致了人脸检测性能的下降，而 SCI 处理后的图像因为整体效果较好细节较为清晰使人脸检测性能提升，因此本文选用 SCI 图像增强方法对低照度图像进行处理。

#### 5.1.4 对比实验

为验证改进后的 YOLOv5s 模型的先进性，将其与主流的目标检测模型 SSD<sup>[22]</sup>、Faster RCNN<sup>[23]</sup>、YOLOv7<sup>[24]</sup>、YOLOv8<sup>[25]</sup>、YOLOv5s 在 DARK FACE 数据集上进行对比分析，实验结果如表 2 所示。

表 2DARK FACE 数据集上的实验结果  
Table 2 Experimental results on the DARK FACE dataset

Model	mAP@0.5/%	Params/MB	FLOPs/G
SSD	57.34	26.29	242.71
Faster RCNN	62.52	137.18	398.12
YOLOv5s	69.81	7.26	16.78
YOLOv7	71.67	37.21	103.58
YOLOv8	69.28	11.19	27.96
Ours	72.19	2.75	6.14

从表 2 中可以看出本文提出的改进模型在参数量和计算量大幅下降的同时检测精度最高，对比基准模型 YOLOv5s 不仅在参数量和计算量上分别下降了 62.12%、63.41%，在检测精度上还提高了 2.38%。对比最新的检测模型 YOLOv7 和 YOLOv8，在参数量上分别下降了 92.61%、75.45%，计算量上分别下降了 94.07%、78.04%，检测精度分别提升了 0.52%、2.91%。本文提出的改进模型在参数量和计算量大幅下降的前提下，对比其它模型仍具有较高的检测精度。

#### 5.1.5 消融实验

为了验证不同改进方法对算法的影响程度，本文在 DARK FACE 公共数据集上进行了消融实验，实验结果如表 3 所示。

表 3 消融实验  
Table 3 Ablation experiment

Parts	1	2	3	4	5	6
ShuffleNetV2	×	√	√	√	√	√
StemBlock	×	×	√	√	√	√
IBC3	×	×	×	√	√	√
CIBC3	×	×	×	×	√	√

SCI	×	×	×	×	×	√
mAP@0.5/%	69.81	68.93	69.75	70.49	71.36	72.19
Params/MB	7.26	2.92	2.91	2.73	2.75	2.75
FLOPs/G	16.78	7.47	6.31	6.0	6.14	6.14

在 DARK FACE 公共数据集上, 将 ShuffleNetV2 网络代替原 YOLOv5s 网络中的主干网络后, mAP 值下降了 0.88%, 参数量和计算量分别下降了 59.78%、55.48%, 在第三组实验中, 将下采样模块替换为 StemBlock 模块, 模型的 mAP 值增加了 0.82%, 同时参数量下降 0.01MB, 计算量下降 1.16G。在第四组实验中, 将 C3 模块替换为 IBC3 模块, mAP 值上升了 0.74%, 参数量下降了 0.18MB, 计算量减少了 0.31G。在第五组实验中, 在 IBC3 模块中添加了 CBAM 注意力机制, mAP 值增加了 0.87%, 参数量增加了 0.02MB, 计算量增加了 0.14G。在第六组实验中添加了图像增强模块 SCI, mAP 值提高了 0.83%。本文提出的改进模型更适合部署在边缘计算设备中并能满足复杂低光环境下的需求。

## 5.2 FSR-Net 实验结果与分析

### 5.2.1 实验数据集

该实验选用的数据集为 CEW 公开数据集、YawDD 公开数据集和自建数据集。CEW 数据集是由 Song 等<sup>[26]</sup>收集的数据, 用于检测眼睛状态。YawDD 数据集是由 Abtahi 等<sup>[27]</sup>构建的视频数据集, 本文先将视频格式转换为帧图像, 并对图像进行预处理和分类用于构建数据集。自建数据集是在模拟驾驶环境下进行拍摄而成的数据集。将各数据集分别划分为训练集、验证集、测试集, 比例为 8: 1: 1。

最终划分的眼部、嘴部数据集信息如表 4、5 所示。

表 4 眼部数据集信息

Table 4 Information of eye dataset

Dataset	Training set	Validation set	Test set	Total
CEW	1939	242	242	2423
YawDD	2880	360	360	3600
Selt-built dataset	2880	360	360	3600

表 5 嘴部数据集信息

Table 5 Information of mouth dataset

Dataset	Training set	Validation set	Test set	Total
YawDD	2880	360	360	3600
Selt-built dataset	2880	360	360	3600



数据集实例见图 16。



图16 数据集实例

Fig. 16 Example of dataset

### 5.2.2 对比实验

FSR-Net 在各个数据集下的测试集平均准确率为 97.89%，具体准确率如表 6 所示。

表 6 FSR-Net 检测准确率

Table 6 FSR-Net detection accuracy

Dataset	Accuracy of Eye/%	Accuracy of Mouth/%
CEW	97.22	—
YawDD	97.50	98.06
Selt-built dataset	98.06	98.61
Average Accuracy/%	97.89	

为验证 FSR-Net 的有效性，在同数据集上将该网络与目前先进的分类网络 MobileNetV2<sup>[28]</sup>、ShuffleNetV2、GhostNet<sup>[29]</sup>、ResNet50<sup>[30]</sup> 和同类型网络 EMs-Net<sup>[8]</sup>、EMSD-Net<sup>[31]</sup> 进行对比，检测准确率和参数量如表 7 所示。

表 7 FSR-Net 与不同网络的实验结果对比

Table 7 Comparison of detection results between FSR-Net and different networks

Algorithm	Average Accuracy/%	Params/MB
MobileNetV2	95.53	3.50
ShuffleNetV2	95.71	1.36
GhostNet	96.49	7.30
ResNet50	90.64	25.64
EMs-Net	97.41	4.84
EMSD-Net	97.52	0.84
Ours	97.89	0.45

由表 7 可知，本文提出的 FSR-Net 检测平均准确率为 97.89%，在参照网络中准确率最高，同时参数量最小为 0.45MB。综上可知，FSR-Net 在眼嘴状态分类任务中具有可行性和有效性。

### 5.3 疲劳驾驶模型实验结果与分析

#### 5.3.1 实验数据集

该实验采用了 YawDD 公开数据集和自建数据集。YawDD 公共数据集是在正常光照环境下进行拍摄的视频数据集，共有 178 段视频。由于目前很少有公开的在夜晚环境下疲劳驾驶数据集，因此本文自制了相关数据集。该数据集在夜间真实驾驶场景下进行拍摄，参与者共有 20 名（14 名男性，6 名女性），每位参与者被要求分别执行正常驾驶和疲劳驾驶的行为活动。最终拍摄视频 200 段，每段视频大约 10s 左右。其中标签为疲劳驾驶的视频数据为 93 段，标签为正常驾驶的视频数据为 107 段。数据集实例见图 17。图像增强后的自建数据集实例见图 18。

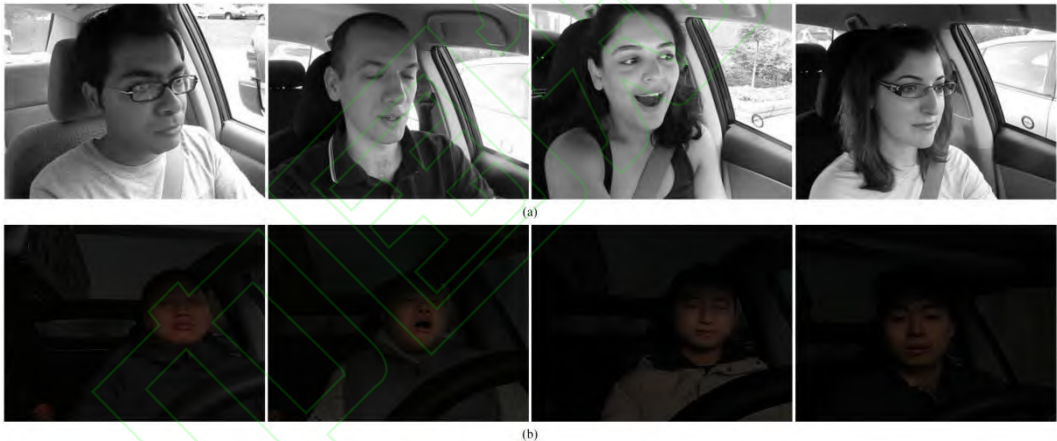


图17 疲劳驾驶数据集实例 (a) YawDD公共数据集实例； (b) 自建数据集实例

Fig. 17 Example of fatigue driving dataset (a) Example of the YawDD public dataset; (b) example of the Self-built dataset



图18 图像增强后的自建数据集实例

Fig. 18 Example of image enhanced Self-built dataset

#### 5.3.2 对比实验

为验证本文所提模型的检测性能，将其与目前先进的几种疲劳驾驶检测模型 MTCNN+EMs-Net<sup>[8]</sup>，DWC-based one-stage algorithm<sup>[9]</sup>，RetinaFace+EMSD-Net<sup>[31]</sup> 分别在 YawDD 数据集与自建数据集上进行对比。结果如表 8 所示。

表 8 不同模型实验结果

Table 8 Experimental results of different models

Model	YawDD/%	Selt-built dataset/%	Average time expenditure/ms
MTCNN+EMs-Net	93.26	81.50	19
DWC-based one-stage algorithm	94.38	79.00	24
RetinaFace+EMSD-Net	95.50	83.50	40
Ours	96.07	94.50	27

注：平均时间开销为检测一帧图像所耗费时间。

由表 8 可知，本文所提模型在 YawDD 数据集与自建数据集上的准确率分别为 96.07%、94.5%，相较于对比模型准确率最高。YawDD 数据集是在正常光照下拍摄的，对比模型在该数据集上的准确率相对稳定。由于自建数据集是在低光环境下拍摄的，这导致了对比模型的准确率大幅下降。综上，本文所提模型无论在正常光照还是在低光照环境下都能保持良好的准确性，并且该模型处理单张图片的时间为 27ms，每秒可以处理大约 37 帧图像，能够满足实时性要求。

疲劳检测结果实例见图 19。

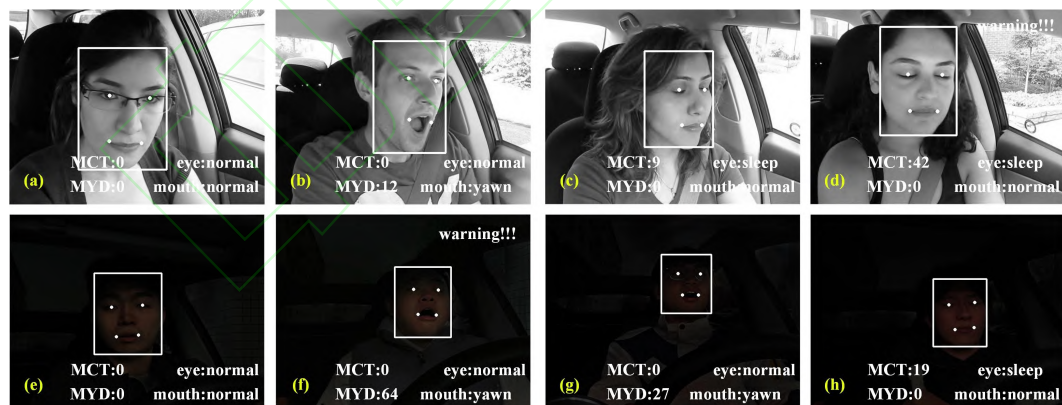


图19 疲劳检测结果实例。(a) (b) (c) (d) YawDD公共数据集检测结果；(e) (f) (g) (h) 自建数据集检测结果

Fig. 19 Examples of fatigue detection results. (a)(b)(c)(d) YawDD public dataset detection results; (e)(f)(g)(h) Self-built dataset detection results

图 19 (a) (e) 为正常驾驶状态；图 19 (b) (f) (g) 为检测到驾驶人打哈欠，由于 (b) (g) 的 MYD 指标分别为 12、27，没有达到设定阈值故不会发出警报，(f) 的 MYD 指标为 64，超过设定的阈值故发出警报；图 19 (c) (d) (h) 为检测到驾驶人闭眼，由于 (c)

(h) MCT 指标分别为 9、19, 没有达到设定阈值故不会发出警报, 但是 (d) 的 MCT 指标为 42, 超过设定的阈值故发出警报。

## 6 结 论

本文提出一种基于图像增强和人脸状态识别的疲劳驾驶检测模型。首先通过改进 YOLOv5s 模型用于低光环境下人脸检测及关键点定位; 之后提出了 FSR-Net 用于判断眼嘴图像的开闭状态; 最后设计了 MCT 指标与 MYD 指标判断驾驶人当前是否处于疲劳驾驶。所提模型在 YawDD 正常光照数据集与自建低光照数据集上分别达到了 96.07%, 94.5% 的准确率, 并能每秒处理 37 帧的图像, 验证了所提模型在正常环境与低光环境下都具有良好的实时性和准确性。

## 参考文献

- [1] 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 T L, et al. Review of research on driver fatigue driving detection methods[J]. Computer Engineering and Applications, 2022, 58(21): 53-66.
- [2] Arefnezhad S, Samiee S, Eichberger A, et al. Applying deep neural networks for multi-level classification of driver drowsiness using Vehicle-based measures[J]. Expert Systems with Applications, 2020, 162(1): 113778.
- [3] Zheng W, Gao K, Li G, et al. Vigilance estimation using a wearable EOG device in real driving environment[J]. IEEE Transactions on Intelligent Transportation Systems, 2020, 21(1): 170-184.
- [4] Jia H J, Xiao Z J, Ji P. Real-time fatigue driving detection system based on multi-module fusion[J]. Computers Graphics, 2022, 108: 22-33.
- [5] 李迟件, 姚靖, 高玉峰, 等. 利用深度学习扩展双光子成像视场[J]. 中国激光, 2023, 50(9): 0907107.  
Li C J, Yao Y, Gao Y F, et al. Extending Field-of-View of Two-Photon Microscopy Using Deep Learning[J]. Chinese Journal of Lasers, 2023, 50(9): 0907107.
- [6] 赵菲, 邓英捷. 融合多异构滤波器的轻型弱小目标检测网络[J]. 光学学报, 2023, 43(9): 0915001.  
Zhao F, Deng Y J. Light Dim Small Target Detection Network with Multi-Heterogeneous Filters[J]. Acta Optica Sinica, 2023, 43(9): 0915001.
- [7] 张志威. 基于机器视觉的异常驾驶行为检测方法研究[D]. 长沙: 湖南大学, 2020.  
Zhang Z W. Research on the detection method of abnormal driving behavior based on machine vision[D]. Changsha: Hunan University, 2020.
- [8] 敖邦乾, 杨莎, 令狐金卿, 等. 基于级联神经网络疲劳驾驶检测系统设计[J]. 系统仿真学报, 2022, 34(02): 323-333.

- Ao B Q, Yang S, Linghu J Q, et al. Design of fatigue driving detection system based on cascaded neural network[J]. Journal of System Simulation, 2022, 34(02): 323-333.
- [9] Wang J, Huang S, Lin J, et al. Driver fatigue detection using improved deep learning and personalized framework[J]. International Journal on Artificial Intelligence Tools, 2022, 31(02): 2250024.
- [10] 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.
- Lv X L, Liu X F, Bai Y Q, et al. Research on driving fatigue detection based on SSD multi-factor fusion[J]. Electronic Measurement Technology, 2022, 31(02): 2250024.
- [11] Chen W, Shah T. Exploring low-light object detection techniques[EB/OL]. (2021-07-30) [2024-02-09]. <https://arxiv.org/abs/2107.14382>.
- [12] Sasagawa Y, Nagahara H. Yolo in the dark-domain adaptation method for merging multiple models[C]//Computer Vision-ECCV 2020: 16th European Conference, August 23-28, 2020, Glasgow, UK. Berlin: Springer, 2020: 345-359.
- [13] Ma L, Ma T, Liu R, et al. Toward fast, flexible, and robust low-light image enhancement[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, June 18-24, 2022, LA, USA. Piscataway: IEEE, 2022: 5637-5646.
- [14] 刘丰, 韩民, 万军, 等. 基于深度学习的牙齿病变自动检测算法[J]. 中国激光, 2022, 49(20): 2007207.
- Liu F, Han M, Wang J, et al. Automatic Detection of Dental Lesions Based on Deep Learning[J]. Chinese Journal of Lasers, 2022, 49(20): 2007207.
- [15] Guo X J, Li Y, Ling H B. LIME: Low-light image enhancement via illumination map estimation[J]. IEEE Transactions on image processing, 2016, 26(2): 982-993.
- [16] Jiang Y, Gong X, Liu D, et al. Enlighten: Deep light enhancement without paired supervision[J]. IEEE transactions on image processing, 2021, 30: 2340-2349.
- [17] Wang R J, Li X, Ling C X. Pelee: a real-time object detection system on mobile devices[EB/OL]. (2018-04-18) [2024-03-17]. <https://arxiv.org/abs/1804.06882>.
- [18] Yang H, Liu J, Mei G, et al. Research on real-time detection method of rail corrugation based on improved ShuffleNetV2[J]. Engineering Applications of Artificial Intelligence, 2023, 126: 106825.
- [19] Wang Q, Sheng J, Tong C, et al. A fast facet-based SAR imaging model and target detection based on YOLOv5 with CBAM and another detection head[J]. Electronics, 2023, 12(19): 4039.
- [20] 崔馨方. 关于人脸关键点检测的若干问题研究[D]. 南京: 东南大学, 2020.
- Cui Q F. Research on several issues about face keypoints detection[D]. Nanjing: Southeast University, 2020.
- [21] Ji Y, Wang S, Zhao Y, et al. Fatigue state detection based on multi-index fusion and state recognition network[J]. IEEE Access, 2019, 7: 64136-64147.
- [22] Yin Z B, Liu F Y, Geng H, et al. A high-precision jujube disease spot detection based on SSD during the sorting process[J]. Plos one, 2024, 19(1): e0296314.
- [23] Kong X, Li X, Zhu X, et al. Detection model based on improved faster-RCNN in apple orchard environment[J]. Intelligent Systems with Applications, 2024, 21: 200325.
- [24] Li Z, Zhu Y, Sui S, et al. Real-time detection and counting of wheat ears based on improved YOLOv7[J]. Computers and Electronics in Agriculture, 2024, 218: 108670.



- [25] Guan H, Deng H, Ma X, et al. A corn canopy organs detection method based on improved DBi-YOLOv8 network[J]. *European Journal of Agronomy*, 2024, 154: 127076.
- [26] Song F, Tan X, Liu X, et al. Eyes closeness detection from still images with multi-scale histograms of principal oriented gradients[J]. *Pattern Recognition*, 2014, 47(9): 2825-2838.
- [27] Abtahi S, Omidyeganeh M, Shirmohammadi S, et al. YawDD: a yawning detection dataset[C]//*Proceedings of the 5th ACM multimedia systems conference*, March 19, 2014, Tainan, Taiwan. New York: Association for Computing Machinery, 2014.
- [28] Anil B K, Mohan B. Face mask detection on photo and Real-Time video images using Caffe-MobileNetV2 transfer learning[J]. *Applied Sciences*, 2023, 13(2): 935-935.
- [29] Du Z, Xu X, Bai Z, et al. Feature fusion strategy and improved GhostNet for accurate recognition of fish feeding behavior[J]. *Computers and Electronics in Agriculture*, 2023, 214:108310.
- [30] Hasanah S A, Pravitasari A A, Abdullah A S, et al. A deep learning review of ResNet architecture for lung disease identification in CXR Image[J]. *Applied Sciences*, 2023, 13(24): 13111.
- [31] 张博熠, 者甜甜, 赵新旭, 等. 基于眼嘴状态识别网络的疲劳驾驶检测[J]. *计算机工程*, 2023, 49(05): 310-320.
- Zhang B O, Zhe T T, Zhao X X, et al. Fatigue driving detection based on eye and mouth state recognition network[J]. *Computer Engineering*, 2023, 49(05): 310-320.

网络首发：

标题：低照度下基于图像增强和人脸状态识别的疲劳驾驶检测

作者：赵洋, 苗佳龙, 刘雪枫, 赵锦程, 徐森

收稿日期：2024-02-20

录用日期：2024-03-25

DOI：10.3788/LOP240711

引用格式：

赵洋, 苗佳龙, 刘雪枫, 赵锦程, 徐森. 低照度下基于图像增强和人脸状态识别的疲劳驾驶检测[J]. 激光与光电子学进展, 2024, 61(22): 2215005.

网络首发文章内容与正式出版的有细微差别，请以正式出版文件为准！

您感兴趣的其他相关论文：

车底危险物图像快速拼接算法

庄建军 金鑫

南京信息工程大学电子与信息工程学院，江苏 南京 210044

激光与光电子学进展, 2024, 61(8): 0837011

基于YOLOv5s的自动扶梯乘客异常行为实时检测算法

王源鹏 万海斌 黄凯 迟兆展 张金旗 黄智星

广西大学计算机与电子信息学院，广西 南宁 530004

激光与光电子学进展, 2024, 61(8): 0812004

基于载物台运动信息的显微图像拼接算法研究

黄家广 玉振明 彭国晋 甘辉 吕美妮

桂林电子科技大学信息与通信学院，广西 桂林 541004

激光与光电子学进展, 2024, 61(8): 0837012

基于DeepLabv3+的轻量化路面裂缝检测模型

夏晓华 苏建功 王耀耀 刘洋 李明臻

长安大学工程机械学院，陕西 西安 710000

激光与光电子学进展, 2024, 61(8): 0812001

深度学习在甲状腺结节良恶性分类中的应用进展

张文凯 王晓燕 刘静 周启香 贺鑫

山东中医药大学智能与信息工程学院，山东 济南 250355

激光与光电子学进展, 2024, 61(8): 0800002