

基于改进 YOLOv11 的驾驶员面部疲劳检测研究

赵治淳¹ 朱立忠¹
ZHAO Zhichun ZHU Lizhong

摘 要

由疲劳驾驶导致的交通事故频发，已成为交通安全隐患之一。为减少道路安全隐患，提高交通安全性，文章提出了一种基于改进 YOLOv11n 的驾驶员疲劳检测方法。首先，使用 EfficientNetV2 网络替换 Backbone 网络，在基础网络中嵌入注意力机制，提升了检测精确度与速度；其次，引入 NWD 损失函数（normalized wasserstein distance），提升小目标检测精度；最后，在检测头模块引入了 Detect-ASFF 模块，提升模型在复杂场景下的检测精度。测试结果表明，与原模型相比，改进后的算法平均精度提高了 1.9%，准确度提高了 2.2%，参数量下降了 50%。在 CEW 数据集（closed eyes in the wild）上表现出较高的准确率与实时性、鲁棒性增强。

关键词

疲劳检测；YOLOv11；疲劳驾驶

doi: 10.3969/j.issn.1672-9528.2025.10.019

0 引言

近年来，疲劳驾驶导致的交通事故逐渐增多，保证驾驶安全成为一个问题。若驾驶员在长时间连续驾驶、睡眠不足或身体疲劳的状态下驾驶车辆，易出现反应迟钝、注意力不集中的情况，从而显著增加交通事故发生的概率。然而，准确地判断不同情况下驾驶员的疲劳程度成为一个问题。年龄、性别及种族不同的驾驶员的面部差异，环境中不同的光照，阴影等因素均可能干扰检测效果。因此，实时地检测驾驶员疲劳程度，开发相应的系统成为疲劳驾驶检测的重要研究方向。

驾驶员疲劳具体表现为眼睛闭合或半闭合，嘴部微张、

头部倾倒及生理信号波动等。通过眼部特征分析、头部姿态分析、表情识别、3D 建模以及生理信息检测等方法判断疲劳程度。生理信号检测虽然精度较高，但依赖接触式传感器，用户体验较差，难以大规模应用。基于行为特征的方法属于间接检测，易受驾驶习惯影响，误报率较高。基于计算机视觉的非接触式疲劳检测成为主流研究方向，其核心是通过人脸特征分析疲劳状态。不同方法中，基于驾驶员面部特征分析检测的速度较快、准确性高、鲁棒性好、适用范围广，是近年来的主流研究方向之一。其中，相比于双阶段方法，单阶段方法中的 YOLO 系列模型算法具有较快的处理速度。因此本文采用单阶段检测方法^[1]。

蔡忠祺等人^[2]在 YOLOv8n 网络中引入 Ghost 卷积、添

1. 沈阳理工大学自动化与电气工程学院 辽宁沈阳 110159

[6] 邵炜世,皮德常,邵仲世.学习驱动的分分布式异构混合流水车间批量流能效调度优化 [J]. 控制理论与应用, 2024, 41(6): 1018-1027.

[7]HARENI M, KRISHNA S, AYYAPPAN S, et al. Sharma dynamic resource allocation with entropy-Based CPU weighting in mobile edge computing[C]//2024 15th International Conference on Computing Communication and Networking Technologies (ICCCNT).Piscataway:IEEE,2022: 1892-1905.

[8] XIE Z G, SONG X, CAO J,et al.Energy-aware task scheduling via sigmoid-based battery modeling in distributed systems[J]. ETRI journal, 2022, 44(5): 746-758.

[9] LIU F Z,YU H. HUANG J W,et al.Joint service migration and

resource allocation in edge-IoT system based on deep reinforcement learning[J].IEEE internet of things journal, 2024, 11(7): 11341-11352.

[10] 陈运,陈新,陈伟建.信息论与编码(第4版)[M].北京:电子工业出版社,2023.

【作者简介】

邱秀荣（1983—），女，河南商丘人，硕士，副教授，研究方向：物联网负载均衡。

杨瑞丽（1990—），女，河南商丘人，硕士，讲师，研究方向：物联网通信技术。

（收稿日期：2025-05-20 修回日期：2025-10-13）

通道注意力模块 M_c 计算公式为:

$$\begin{aligned} M_c &= \sigma \left(\text{MLP}(\text{AvgPool}(F)) + \text{MLP}(\text{MaxPool}(F)) \right) \\ &= \sigma \left(W_1(W_0(F_{\max}^c)) + W_1(W_0(F_{\text{avg}}^c)) \right) \end{aligned} \quad (2)$$

式中: σ 表示 Sigmoid 激活函数; W_0 和 W_1 表示 MLP 的权重参数。

空间注意力模块 M_s 计算公式为:

$$\begin{aligned} M_s &= \sigma(f^{7 \times 7}([\text{AvgPool}(F); \text{Maxpool}(F)])) \\ &= \sigma(f^{7 \times 7}([F_{\text{avg}}^c; F_{\max}^c])) \end{aligned} \quad (3)$$

式中: $f^{7 \times 7}$ 表示 7×7 卷积。

CBMA 混合注意力机制模块如图 3 所示。

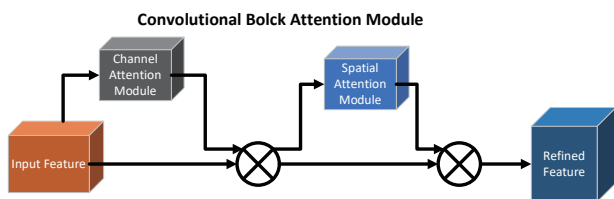


图 3 CBAM 混合注意力机制

该模块有效地增强了特征表达能力和适应性,一定程度上提高了模型性能。

2.4 检测头改进

在提高复杂场景下的检测精度时,常用特征金字塔或多级特征塔等方法,但在单阶段检测器中,不同尺度之间的一致性影响了其性能,本文引入 ASFF 自适应空间特征融合模块。ASFF 结构如图 4 所示。

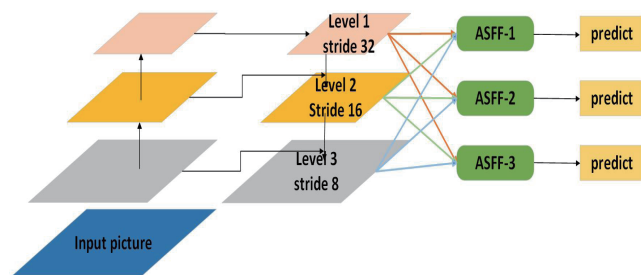


图 4 ASFF 模块

设调整后的特征向量为 $x_{ij}^{n \rightarrow l}$, 融合公式为:

$$y_{ij}^l = \alpha_{ij}^l \times x_{ij}^{1 \rightarrow l} + \beta_{ij}^l \times x_{ij}^{2 \rightarrow l} + \gamma_{ij}^l \times x_{ij}^{3 \rightarrow l} \quad (4)$$

ASFF 通过自适应学习不同尺度特征图的融合权重,使网络能动态选择对当前目标最有效的特征层级,在不增加计算开销的情况下,提升复杂情况下的检测精度。与传统方法如 FPN、PANet 等相比表现更好^[10]。

3 实验与分析

3.1 实验环境

本文模型训练的实验平台硬件环境为 Intel Core i5-9300H

CPU@2.40 GHz、NVIDIA GeForce GTX 1650。软件环境为 Ubuntu 系统、Python3.9、CUDA11.7、PyTorch0.15.2 及其他依赖包。

3.2 实验数据集

本文数据集由当前疲劳检测领域具有权威性的 CEW (closed eyes in the wild) 数据集、FDD-dataset 数据集以及自制数据整合而成,包含了 4 507 张被测者驾驶状态的图片,被测者的差异化以及环境、光照等因素的不同体现出图片的差异化。对部分图片进行处理后按照 3:1:1 的比例划分训练集、测试集和验证集。

3.3 实验内容

本文通过平均精度均值 (mAP)、权重大小 (ModelSize)、精确率 (Precision) 和召回率 (Recall) 衡量性能。

精确率 P 与召回率 R 的公式分别为:

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

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

式中: TP 表示图片被检测为正样本; FN 表示图片被检测为负样本; FP 表示负样本被检测为正样本。

平均精度是根据精确率与召回率计算得出,其公式为:

$$AP = \int_0^1 P(R) dR \quad (7)$$

mAP 是所有类别的 AP 值的平均,其公式为:

$$mAP = \frac{\sum_{n=0}^N AP_n}{N} \quad (8)$$

参数设置为 device = '0', epochs = 500, batch = 4, mosaic = 0, patience = 500, workers = 0, imgsz = 640。

3.4 结果分析

3.4.1 对比试验

本文将部分主流模型与本文模型做了对比实验,通过准确率、召回率、模型大小以及平均精度作为指标,使用同一数据集进行训练,实验结果如表 1 所示。

表 1 对比试验

Models	mAP/%	P/%	R/%	Params/ 10^6
YOLOv8n	92.3	92.2	90.5	6.2
YOLOv11n	93.1	93.0	91.5	2.6
本文模型	95.8	94.9	92.9	1.5

通过对比可知, YOLOv11 比 YOLOv8n 在参数量上大幅减少,召回率高出 1.0%,准确率高出 0.8%,平均精度高出 0.8%。本文所用模型与 YOLOv11n 相比,参数量减少 42%,召回率提高 1.4%,准确率提高 1.9%,平均精度提高 2.7%,

证明本实验有一定效果。

3.4.2 消融实验

本文在同一个数据集和实验环境下，对 YOLOv11n 模型进行了消融实验，以此来验证每个模块的有效性以及其性能，结果如表 2 所示。

表 2 消融实验

Models	MAP/%	P/%	R/%	Params/10 ⁶	FLOPs/10 ⁹
YOLOv11n	93.1	93.0	91.5	2.6	6.5
YOLOv11n + EfficientNetV2	92.9	92.9	91.4	1.5	6.5
YOLOv11n + EfficientNetV2 + NWD	94.8	94.8	93.2	1.5	6.5
YOLOv11n + EfficientNetV2 + NWD + CBAM	94.8	94.8	92.6	1.5	4.7
YOLOv11n + EfficientNetV2 + NWD + CBAM+Detect-ASFF	95.8	94.9	92.9	1.5	4.7

从消融实验结果看出，模型引入 EfficientNetV2 后，参数量减少 42%，但平均精度降低 0.2%，准确率降低 0.1%，召回率降低 0.1%。引入 NWD 损失函数后，平均精度提高 1.9%，准确率提高 1.9%，召回率提高 1.8%。再引入 CBAM 模块后，召回率降低 0.6%，运算次数减少 27%，其他参数不变，最后引入 Detect-ASFF 模块后，平均精度提高 1%，准确率提高 0.1%，召回率提高 0.3%，参数量和运算次数不变。总体来说，本文模型比原模型在平均精度上提高 2.7%，准确率提高 1.9%，召回率提高 1.4%，参数量减少 42%，运算次数降低 27%，证明了本文算法有一定提高。

3.4.3 驾驶员面部检测

在测试数据集上使用改进后的模型进行测试，检测结果如图 5 所示。

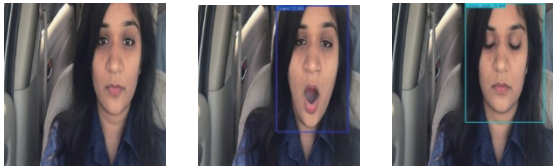


图 5 检测结果图

图 5 中分别为正常驾驶面部状态、打哈欠面部状态和闭眼面部状态。本文可以顺利识别出不同状态下驾驶员的面部状态，对疲劳状态识别的精准度提高。

4 结论

本文基于 YOLOv11n 模型，设计了一个驾驶员面部疲劳

检测系统。首先，使用 EfficientNetV2 网络完成了模型的轻量化。其次，引入 NWD 损失函数提升检测精度。后在模型中嵌入 CBAM 注意力机制，提升检测的准确度与速度。最后，在检测头部分引入了 Detect-ASFF 模块，增强了模型的表达能力。通过实验对比发现，改进后的算法与标准 YOLOv11n 模型对比平均精度提高了 1.9%，准确度提高了 2.2%，参数量下降了 50%。在 CEW 数据集上表现出较高的准确率与实时性，鲁棒性增强。

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.

[2] 蔡忠祺, 林珊玲, 林坚普, 等. 基于改进 YOLOv8n-Pose 的疲劳驾驶检测 [J]. 液晶与显示, 2025, 40(4): 617-629.

[3] 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.

[4] 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.

[5] 杜威, 宁武, 孟丽园, 等. 基于改进 YOLO 的矿卡驾驶员疲劳检测算法 [J]. 现代电子技术, 2025, 48(7): 126-131.

[6] TAN M X, LE Q V. EfficientNetV2: smaller models and faster training[EB/OL]. (2021-06-23)[2025-06-12]. <https://doi.org/10.48550/arXiv.2104.00298>.

[7] WANG J W, XU C, YANG W, et al. A normalized gaussian wasserstein distance for tiny object detection[EB/OL]. (2022-06-14)[2025-05-23]. <https://doi.org/10.48550/arXiv.2110.13389>.

[8] HU J, SHEN L, SUN G. Squeeze-and-excitation networks[EB/OL]. (2019-05-16)[2025-06-22]. <https://doi.org/10.48550/arXiv.1709.01507>.

[9] WOO S, PARK J, LEE J-Y, et al. CBAM: convolutional block attention module[C]//Proceedings of the European conference on computer vision (ECCV). Berlin: Springer, 2018: 3-19.

[10] LIU S T, HUANG D, WANG Y H. Learning spatial fusion for single-shot object detection[EB/OL]. (2019-11-25)[2025-06-12]. <https://doi.org/10.48550/arXiv.1911.09516>.

【作者简介】

赵治淳 (2001—)，男，辽宁锦州人，硕士研究生，研究方向：目标检测。

朱立忠 (1967—)，男，辽宁沈阳人，硕士，教授，研究方向：图像处理、智能控制。

(收稿日期：2025-05-19 修回日期：2025-10-14)