

基于 YOLOv8n 的列车疲劳驾驶检测算法

杨松¹ 彭素真¹

YANG Song PENG Suzhen

摘要

为解决铁路驾驶员疲劳驾驶检测的困难和提高检测精度,文章提出一种基于 YOLOv8n 的轻量化疲劳驾驶检测技术。首先,在 YOLOv8n 的基础框架上,将 Conv 替换为动态通道注意力卷积 DY-CNNs,以增强检测精度。其次,引入损失函数 SIOU,以提高检测精度和鲁棒性。接着,颈部网络引入了跨尺度特征融合模块 CCFM,增强了模型对疲劳特征的细节和上下文信息的捕捉能力。最后,在检测头模块引入了 Detect-improve,以提取更为丰富和精确的特征。测试结果表明,改进后的模型在平均精度上比原模型提高了 3%,准确度提升了 2.8%,并且在计算量、参数量和模型大小上分别下降了 46.3%、50% 和 45.2%。这些改进使得轻量化模型在保持高准确率的同时,为铁路驾驶员疲劳驾驶的预防和检测提供了一个有效的技术手段。

关键词

疲劳检测; YOLOv8n; DY-CNNs; CCFM; 损失函数

doi: 10.3969/j.issn.1672-9528.2025.04.015

0 引言

随着城市化进程加快,城市交通压力日益增大,驾驶员在道路上能否安全行驶,成为一个极为重要的社会问题。疲劳驾驶作为导致交通事故的主要原因之一,其危害性不容忽视^[1]。据统计,疲劳驾驶引发的交通事故占交通事故总量的 30% 以上。因此,开发有效的疲劳驾驶监测系统,对于减少铁路交通事故、保障铁路安全具有重要意义。

驾驶人员在持续驾驶过程中会身心疲劳,从而导致驾驶能力显著降低,这种现象被称为疲劳驾驶^[2]。为解决这一安全隐患,本文采用深度学习方法,通过分析列车司机的眼部和嘴部特征来进行疲劳状态识别。在目标检测技术领域,深度学习已取得显著进展,主要分为单阶段和双阶段检测方法。单阶段方法如 YOLO 系列算法^[3]和 SSD^[4],具有较快的处理速度;双阶段方法如 Faster R-CNN 算法,虽然检测准确率较高,但由于处理速度较慢,难以适应某些实际应用场景的要求,因此本研究选择采用单阶段检测方法。

以往的疲劳监测系统主要基于对驾驶人员生理参数和车辆运行数据的采集,这类系统往往依赖复杂的硬件设备,不仅响应时间长,准确性也较差,同时难以做到对驾驶员状态的持续监控,实际应用价值有限。针对这些问题,本文设计了一套采用轻量级 YOLOv8n 模型的疲劳驾驶识别系统。本系统基于高效的 YOLOv8n 单阶段检测框架,经过优化升级,

成功实现了准确且快速的疲劳状态检测功能。

1 YOLOv8n 模型

YOLOv8 包括了 YOLOv8s、YOLOv8n、YOLOv8m、YOLOv8l 和 YOLOv8x 等不同规模的模型。这些模型通过调整深度和宽度因子来适应不同的应用需求,其中较大的模型通常能提供更精确的检测结果,但相应的推理速度会有所下降。YOLOv8n 虽然在 YOLOv8 系列中是一个中等规模模型,但在保持较高检测精度的同时,也能保证较快的处理速度,所以本文在 YOLOv8n 的基础上进行实验。

2 改进 YOLOv8n 模型

尽管 YOLOv8n 在目标检测领域已经取得了不错的效果,但在实际的驾驶场景中,原始 YOLOv8n 模型仍存在检测精度不高、运行效率不足以及环境适应性较差等局限性。为解决上述问题,本文选择在 YOLOv8n 的基础架构上进行优化和调整。如图 1 所示,为提高模型对疲劳驾驶算法特征的识别能力,引入 SIOU 损失函数;为增强模型的表示能力,引入 DY-CNNs;为增加对小尺度目标的识别,加入 CCFM。最后,为实现更高效的目标检测进行了检测头改进。

2.1 DY-CNNs

为提高模型检测的准确性,本文对 YOLOv8n 主干网络进行了优化,将常规卷积(Conv)替换为动态卷积。动态卷积神经网络(dynamic convolutional neural networks, DY-

1. 沈阳工业大学化工过程自动化学院 辽宁辽阳 111003

CNNs), 该方法的主要优点是可以基于输入特征自适应地更新卷积参数, 在保持网络规模不变的同时提升模型的性能表现^[5]。

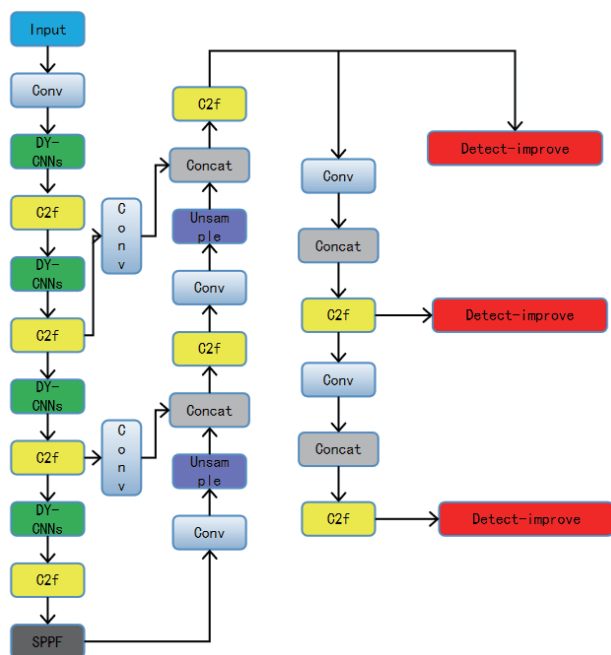


图1 模型结构图

DY-CNNs的计算流程如图2所示。首先, 对输入数据进行全局平均池化, 通过两个全连接层(中间有ReLU激活函数)处理后, 利用softmax函数计算得到归一化的注意力权重。然后, 根据注意力权重对预先设定的多个卷积核进行加权聚合, 得到最终用于卷积操作的卷积核。最后, 用此聚合得到的卷积核对输入数据进行卷积操作, 并经过批归一化和激活函数处理后得到输出。采用这一动态卷积机制, 可根据输入信息灵活地修正卷积参数, 从而在维持原有网络结构规模的前提下增强模型的性能表现。

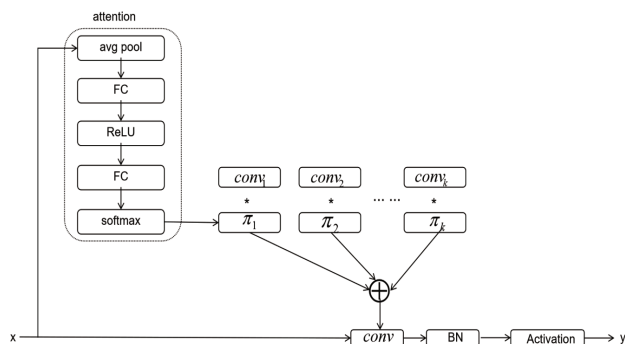


图2 动态卷积层

2.2 损失函数 SIoU

为进一步提升疲劳驾驶检测模型的准确性, 本文引入SIoU用来代替YOLOv8n的CIoU。SIoU损失函数更加关注边界框的位置和尺寸, 还考虑到边界框之间的向量角度, 从

而能够更加全面地评估预测框与真实框之间的一致性。SIoU由角度损失、距离损失、形状损失、IoU损失组成。

通过引入SIoU损失函数, 本文的模型在列车疲劳驾驶检测任务中能够更全面地评估预测框的准确性, 从而提高了模型对疲劳特征的识别能力。

2.3 CCFM

为提高模型对疲劳特征的检测能力, 尤其是对小尺度目标的识别, YOLOv8n颈部网络部分引入了(cross-scale feature fusion module, CCFM)^[6]。CCFM的核心思想是将不同尺度的特征图进行有效融合, 以增强模型对细节和上下文信息的捕捉能力。

为增强特征的表达能力, 在特征提取阶段引入深度可分离卷积来降低计算量, 同时通过在每个尺度的特征提取后加入注意力机制, 使模型更加专注于疲劳特征相关的区域。在模型训练和正则化方面, 采用知识蒸馏技术, 将预训练模型的知识迁移到当前模型中, 增强对小尺度目标的检测能力。

2.4 检测头改进

在目标检测任务中, 检测头的作用是将特征图中的信息转换为最终的边界框和类别置信度。为进一步提升模型的检测性能, 本文对检测头进行了创新性的改进, 引入了Detect-improve模块, 结合了新提出(partial convolution, PConv)和传统的卷积操作, 以实现更高效的目标检测^[7]。

在检测头中, 将PConv与传统的卷积操作相结合, 以提取更为丰富和精确的特征, 使得检测头有更高准确度。图3展示了2种卷积操作图对比: 一种是常规卷积, 另一种是部分卷积。其中, 图3(a)表示常规卷积操作, 有一个输入特征图、一个卷积核和一个输出特征图, 卷积核在输入特征图上滑动进行运算得到输出。图3(b)是部分卷积操作, 利用特征图中的冗余, 通过仅对部分输入通道应用常规卷积, 而对其余通道不进行操作, 从而减少了计算量和内存访问。

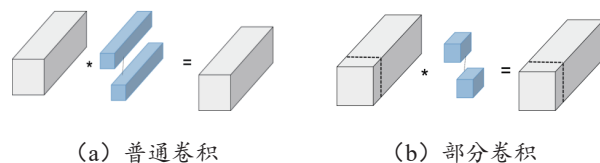


图3 三种卷积图对比

3 试验与分析

3.1 数据集及试验环境

本文采用FDD-dataset数据集作为实验数据集, 该数据集利用车内摄像头记录了司机在驾驶过程中的各种状态, 能

够真实反映实际驾驶中的疲劳情况。整个数据集包含 2 914 张图像, 其中疲劳状态(闭眼和打哈欠)的图像有 545 张。对扩增后的数据按照 3:1:1 的比例进行了训练集、验证集和测试集的随机划分。具体的数据集划分情况和实验环境配置分别在表 1 和表 2 中列出。

表 1 数据集划分

训练集		验证集		测试集	
正常驾驶	疲劳驾驶	正常驾驶	疲劳驾驶	正常驾驶	疲劳驾驶
1 422 张	326 张	474 张	109 张	474 张	109 张

表 2 试验平台

类别	版本
操作系统	Ubuntu20.04
内存	80 GB
CPU	Intel Xeon Platinum 8474C
GPU	RTX 4090D
PyTorch	PyTorch 1.11.0
Python	Python 3.8
CUDA	CUDA 11.3

3.2 评价指标

在本项研究中, 利用以下性能指标来对疲劳检测模型进行性能评估: 精确度 P 、召回率 R 和平均精度均值 mAP 等指标^[8]。准确率是指在所有的疲劳驾驶的样本中, 真正属于疲劳驾驶的比例。而召回率则是指在所有实际疲劳驾驶的样本中, 模型成功被识别出的比例^[9]。计算公式为:

$$P = \frac{TP}{TP + FP} \quad (1)$$

$$R = \frac{TP}{TP + FN} \quad (2)$$

式中: 真正例 TP 指的是正确识别为疲劳驾驶的样本数量; 假正例 FP 指的是模型错误地将非疲劳驾驶识别为疲劳驾驶的样本数量; 假负例 FN 则是指模型未能识别出实际疲劳驾驶的样本数量。

平均精度是通过计算精确率与召回率之间的积分来得到, 反映了模型在整个召回率范围内的精度表现。对于多类别问题, mAP 是所有类别 AP 值的平均。计算公式为:

$$AP = \int_0^1 P(R) dR \quad (3)$$

$$mAP = \frac{\sum_{n=0}^N AP_n}{N} \quad (4)$$

式中: N 是类别的总数; AP_n 是第 n 个类别的 AP 值。

3.3 试验结果与分析

3.3.1 对比实验

为进一步分析改进算法的性能, 本文对模型做了对比实验^[10]。通过准确度、召回率、平均精度等作为评价指标, 并使用同一训练集、验证集和测试集, 对 YOLOv5n、YOLOv8n 及改进后的 YOLOv8n 模型分别做了实验, 实验结果见表 3。

表 3 对比实验

Models	MAP/%	P/%	R/%	MS/MB
YOLOv5n	92	91.5	90	5.3
YOLOv8n	92.4	92.2	90.5	6.2
Ours	96	94.9	93.3	3.4

通过 YOLOv5n 和 YOLOv8n 对比可知, YOLOv8n 虽然模型体积增加 14.5%, 但在准确度上高出 0.8%, 召回率提升 0.5%; 本文所用模型和 YOLOv8n 相比, 参数量减少 50%, 准确率提高 2.8%, 召回率提高 3%, 证明了本实验的有效性。

3.3.2 消融实验

为验证引入各个模块后的有效性, 本文利用相同的数据集和实验环境, 对 YOLOv8n 模型以及改进后的模型设计了消融实验, 以用来验证每个模块改进后实现的算法性能, 实验结果如表 4 所示^[11]。

表 4 模型试验结果

Models	MAP /%	P/%	R/%	Params / 10^6	MS /MB	FLOPs
YOLOv8n	92.4	92.2	90.5	3	6.2	8.2
YOLOv8n+SiIoU	93.7	93.7	91.8	3	6.2	8.2
YOLOv8n+DY-CNNs	93.6	93.6	91.5	3	6.2	8.2
YOLOv8n+CCFM	94.9	92	90.1	1.9	4.2	6.7
YOLOv8n+Detect-improve	93.7	94.4	93.9	2.5	5.3	5.7
YOLOv8n+SiIoU+DY-CNNs	95.8	94.3	93.5	3	6.2	8.2
YOLOv8n+DY-CNNs+CCFM+Detect-improve	95.8	94.5	92.6	1.5	3.4	4.4
YOLOv8n+SiIoU+DY-CNNs+CCFM+Detect-improve	96	94.9	93.3	1.5	3.4	4.4

由表 4 可以看出, 在 YOLOv8n 的基础上引入 SiIoU 损失后, 准确率提升 1.6%, 召回率提升 1.4%。当单独引入 DY-CNNs 时, 准确率、召回率和平均精度分别提升 1.5%、1.1%、1.3%。YOLOv8n 加入 CCFM 后, 准确率和召回率分别降低

0.2%、0.4%，参数量、模型体积和计算量分别减少 37%、32%、18%。YOLOv8 引入 Detect-improve 后，准确率、召回率和平均精度分别提升 2.3%、3.6%、1.4%，参数量减少 17%。改进后的疲劳驾驶检测算法模型 YOLOv8n+Siou+DY-CNNs+CCFM+Detect-improve 相比于原 YOLOv8n，准确率提高了 2.8%，检测精度上比原模型高 3%，并在计算量、参数量和模型大小上分别下降了 46.3%、50% 和 45.2%，证明了本文算法的有效性。

3.3.3 疲劳驾驶检测

将优化后的疲劳检测系统应用于测试数据集，检测效果如图 4 所示，图中展示了三种典型状态的检测结果：正常驾驶、眼睛闭合，以及打哈欠状态。实验结果表明，本文提出的改进模型通过重点关注驾驶员面部的眼部和嘴部区域，成功实现了对疲劳状态的精准识别与判断。

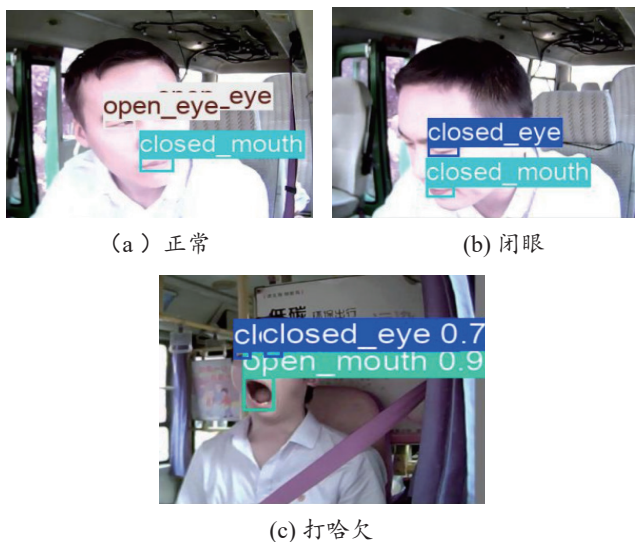


图 4 疲劳状态识别效果

4 结语

本文以 YOLOv8n 为基础框架，设计并实现了一个优化的列车司机疲劳状态检测系统。首先，将 Conv 替换为 DY-CNNs，提高了模型的检测性能。接着，引入 Siou 损失函数，提升了模型的准确度。进一步引入 CCFM，增强了模型对复杂特征的捕获能力。最后，对检测头部分进行了创新性的改造，提出了 PConv，增强了模型的代表能力。

通过实验对比发现，优化后的检测系统较基准 YOLOv8n 模型的准确率提升了 2.8%，检测精度上比原模型提高了 3%，计算量、参数量和模型大小分别下降了 46.3%、50% 和 45.2%。通过不断对模型的改进，使得改进后的模型更加符合实时列车疲劳驾驶检测的要求。未来，通过不断的技术创新和改进，本文的模型将能够为减少由疲劳驾驶引起的铁路交通事故做出更大的贡献。

参考文献：

- [1] 龚浩为, 文志强. 车辆驾驶人疲劳状态的识别与监测系统[J]. 福建电脑, 2025, 41(4): 75-79.
- [2] 李文全, 彭新东, 郭润秋. 基于计算机视觉的疲劳驾驶预警系统的实现[J]. 现代计算机, 2024, 30(22): 207-211.
- [3] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: unified, real-time object detection[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE, 2016: 779-788.
- [4] LIU W, ANGUELOV D, ERHAN D, et al. SSD: single shot multibox detector[C]//Computer Vision ECCV 2016. Berlin: Springer, 2016: 21-37.
- [5] GUO Z C, ZHANG X Y, MU H Y, et al. Single path one-shot neural architecture search with uniform sampling[C]//Computer Vision ECCV 2020. Berlin: Springer, 2020: 544-560.
- [6] ZHAO Y A, LÜ W Y, XU S L, et al. DETRs beat YOLOs on real-time object detection[C]//Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. Piscataway: IEEE, 2024: 16965-16974.
- [7] CHEN J, KAO S H, HE H, et al. Run, don't walk: chasing higher FLOPS for faster neural networks[C]//Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. Piscataway: IEEE, 2023: 12021-12031.
- [8] 宋巍. 基于深度学习的违规驾驶行为视觉识别方法研究[D]. 武汉: 武汉科技大学, 2023.
- [9] 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.
- [10] 郭占广. 基于神经网络的安全驾驶辅助系统设计与研发[D]. 青岛: 青岛科技大学, 2021.
- [11] 史瑞秀. 自适应融合图卷积双流特征的驾驶员疲劳检测[D]. 西安: 长安大学, 2024.

【作者简介】

杨松(1968—), 男, 辽宁庄河人, 硕士, 教授、硕士生导师, 研究方向: 化工过程控制算法和人工神经网络。

彭素真(1998—), 通信作者(email: 2278568484@qq.com), 女, 河南周口人, 硕士研究生, 研究方向: 化工过程控制算法和人工神经网络。

(收稿日期: 2025-04-15)