

DOI:10.19651/j.cnki.emt.2416897

基于改进 YOLOv8n 的轻量化分心驾驶行为检测方法^{*}

沈 塞 张 磊 张 宇 翔 李 易 刘 世 豪

(淮阴工学院自动化学院 淮安 223003)

摘要: 分心驾驶行为是导致出现道路交通安全问题的主要原因之一。针对现有基于深度学习的检测算法计算复杂度高、泛化能力有限、检测精度不理想等问题,本文构建了一种基于改进 YOLOv8n 的轻量化分心驾驶行为检测算法。首先,将 CAA 注意力机制融合进 StarNet,形成 StarNet-CAA,并且将 StarNet-CAA 集成到 YOLOv8n 的主干网络中,提高了模型的全局特征提取能力,显著降低了计算复杂度。随后,将 FasterBlock 与 CGLU 相结合加入到颈部网络中,形成 C2f-Faster-CGLU 模块,降低计算成本。此外,在检测头中引入共享卷积,进一步减少了计算量和参数量。实验结果表明,改进的 YOLOv8n 算法显著提高了分心驾驶行为检测的效率,在 StateFarm 数据集上达到了 99.4% 的准确率。模型的参数量减少 46.7%,计算量减少 41.5%。此外,在 100-Driver 数据集上进行了泛化实验,结果表明,与 YOLOv8n 相比,该方案的泛化效果有所提高。因此,该算法在保持高可靠性和泛化能力的同时,显著降低了计算量。

关键词: 分心驾驶; YOLOv8n; 目标检测; 深度学习

中图分类号: TN0; TP391 文献标识码: A 国家标准学科分类代码: 520.6

Lightweight distracted driving behavior detection method based on improved YOLOv8n

Shen Qian Zhang Lei Zhang Yuxiang Li Yi Liu Shihao

(School of Automation, Huaiyin Institute of Technology, Huai'an 223003, China)

Abstract: Distracted driving is one of the main causes of road traffic safety problems. Aiming at the problems of high computational complexity, limited generalization ability and unsatisfactory detection accuracy of existing detection algorithms based on deep learning, this paper constructs a lightweight distracted driving behavior detection algorithm based on improved YOLOv8n. Firstly, the Context Anchor Attention mechanism was fused into StarNet to form StarNet-CAA, and StarNet-CAA was integrated into the backbone network of YOLOv8n to improve the global feature extraction ability of the model and significantly reduce the computational complexity. Subsequently, FasterBlock combined with CGLU is added to the neck network to form the C2f-Faster-CGLU module, which reduces the computational cost. In addition, the shared convolution is introduced into the detection head to further reduce the computational burden and parameter size. Experimental results show that the improved YOLOv8n algorithm significantly improves the efficiency of distracted driving behavior detection, reaching an accuracy of 99.3% on the StateFarm dataset. The number of parameters of the model is reduced by 46.7%, and the amount of calculation is reduced by 41.5%. In addition, the generalization experiment is carried out on the 100-Driver dataset, and the results show that the generalization effect of the proposed scheme is improved compared with YOLOv8n. Therefore, the proposed algorithm significantly reduces the computational burden while maintaining high reliability and generalization ability.

Keywords: distracted driving; YOLOv8n; object detection; deep learning

0 引言

随着近年来汽车数量的急剧增加,分心驾驶已成为一

个普遍且日益严重的全球性问题,它不仅对驾驶员和乘客的安全构成重大威胁,并且也有可能对道路旁的行人带来危险。随着科技的进步,这一现象已经引起了广泛关注,越

来越多的车辆配备了 L2 级甚至 L3 级的自动驾驶能力,可以在紧急情况下快速控制车辆,减少事故的发生。然而,有效识别分心驾驶行为是辅助驾驶的先决条件。因此,一种精确、高效的分心驾驶行为检测算法对于提高道路安全,减少交通事故具有至关重要的意义。

目前的研究分析包括采集车辆的自然驾驶数据。该方法主要利用传感器测量车辆速度、加速度和方向盘转动角度来评估驾驶状态。但实时性较差,无法实时准确判断驾驶状态。另一种方法是使用可穿戴设备检测驾驶员生理参数,虽然该方法精度较高,但会干扰驾驶员操作。目前,检测分心驾驶的主要方法是计算机视觉配合摄像头识别分心驾驶行为。

近年来,深度学习在包括分心驾驶行为检测在内的各个领域都取得了巨大的成功。卷积神经网络类的模型之前被大量应用在分心驾驶行为识别领域。例如:Koesdwiady 等^[1]利用 VGGNet 进行分心驾驶行为检测。然而,该文献中所识别的分心驾驶类别太少,无法包括目前大多数分心驾驶行为。Hesham 等^[2]捕获人脸和手,利用 AlexNet 提取特征,在 AUC 数据集上取得了较高的检测精度。该文献构建了一个新的分心驾驶数据集,为后续的分心驾驶行为检测做出了很大的贡献。张志威^[3]改进了姿态估计算法 OpenPose,将算法输出的骨架置信度图、骨架亲和场和原始图像输入到 VGG19 深度学习网络中,对分心和异常驾驶行为进行分类。白中浩等^[4]利用图卷积网络提取驾驶员姿态特征,并结合目标检测算法对分心驾驶行为进行分类,在 StateFarm 数据集上准确率达到 90%。Tran 等^[5]提出了一种基于双目摄像头的驾驶员行为检测系统,使用双目摄像头,并且进行数据融合之后,识别准确率明显大于单目摄像头。这些在传统 CNN 模型上做出优化之后的模型,参数量与计算量都过于庞大,这些模型更加适合用于解决更加复杂的问题。Li 等^[6]提出了一种轻量级的卷积神经网络 OLCMNet,该网络采用通道模块减少模型参数,同时提高分类精度。Bekhouche 等^[7]利用 YOLO 和 ResNet-50 开发了一种驾驶员困倦检测方法,通过混合特征选择在疲劳驾驶检测数据集上实现了较高的准确率。尹智帅等^[8]提出了一种 2D 姿态估计算法的分心驾驶行为检测模型。虽然该模型在 StateFarm 数据集上的准确率达到 95.77%,但在自建的实际驾驶集上,模型的准确率下降了 10.35%。张志远^[9]提出了一种改进的 YOLOv5s 网络模型,主要采用 MobileNet-v3 结构替代主干网,增加 SPPF 结构,将注意机制与瓶颈结构相结合,检测危险驾驶行为。但是,该文献没有测试光照、天气等因素对算法的影响。Peng 等^[10]提出了一种基于时空双流深度学习网络和因果或图的分心驾驶行为识别方法。该方法处理连续帧来识别分心驾驶行为,准确率大大超过了其他方法。李少凡等^[11]将 YOLOv5 与 HRNet 相结合,在开罗美国大学分心驾驶数据集上的准确率为 96.17%,在自建数据集上的准确率为 96.97%,显著

优于传统方法。Lou 等^[12]提出了一种基于 YOLOv5s 的改进算法,将其与 GhostNet 相结合,减少了计算量和参数计数。此外,在 C3 模块中加入了 CBAM 注意机制,提高了模型的性能。在一个自建的分心驾驶行为数据集上,该模型的准确率相比 YOLOv5s 提高了 1.5%,计算负载减少了 7.6 GFLOPs。Du 等^[13]提出了一种基于 YOLOv8n 的改进算法:YOLO-LBS。该研究将 GhostBottleneck 与 C2f 模块集成形成 GhostC2f,并将 YOLOv8n 中的 PAN 结构替换为双向特征金字塔网络。由于 StateFarm 数据集的不完善,该工作扩充了 StateFarm 数据集,在其中加入了多种分心驾驶行为,该方法使模型精度提高了 5.1%,同时显著降低了计算负荷。Guo 等^[14]提出了基于 S-ASPP 和轻量级特征重组的 YOLOX-B,用于增强实时驾驶员行为检测。Fu 等^[15]提出了 GD-YOLO,通过 DLAN 和 GLAN 模块增强 YOLOv7,实时检测吸烟和手机使用行为。该方法达到了较好的检测效果,但是并没有使用公开的数据集。Du 等^[16]提出了一种 YOLOv5-GBC 算法,该方法结合了 GhostConv、BiFPN 与通道注意力机制,最终验证得出该算法具有较好的性能。Li 等^[17]通过添加空间和通道重构卷积和全局注意力机制,以及改进的空间金字塔池化快速方,增强了模型对多尺度空间-通道特征的捕获能力。通过大量实验,得出结论。YOLO-SGC 在平均准确率上比其他主流模型有显著提高,同时保持了高检测速度。以上这些方法都是 YOLO 系列的变式,虽然相比于原始网络,优化之后的模型计算量大幅缩减,但是想要在嵌入式设备上流畅运行,模型的复杂度还是较高。贺宜等^[18]提出了一种用于检测营运车辆驾驶员分心驾驶的 MobileViT-CA 模型。该研究基于 MobileViT 网络,通过大量实验,确定 MobileViT-CA 为最优识别模型。在自建的数据集上实现了 96.57% 的准确率,并且在模型的大小和检测精度上都有优势。

目前,深度学习在检测分心驾驶行为方面取得了重大进展。然而,现有的大多数方法仍然存在计算复杂度高、泛化能力差和精度不足的问题,这对它们的广泛采用构成了很大的障碍,特别是在资源受限的嵌入式设备上。为了解决这些问题,本文提出了一种基于改进 YOLOv8n 的轻量化分心驾驶行为检测算法。本文的主要贡献和方法总结如下:本文将 StarNet^[19]与上下文锚点注意力机制(context anchor attention, CAA)^[20]融合形成 StarNet-CAA,并且将 StarNet-CAA 与 YOLOv8n 主干网络集成在一起。StarNet 利用星型运算在低维计算空间内高效处理高维特征,增强了模型对分心驾驶行为的特征提取能力。对 YOLOv8n 的 C2f 模块进行改进,用 FasterBlock^[21]与卷积门控线性单元(convolutional gated linear unit, CGLU)模块结合取代原始的 Bottleneck 结构,形成 C2f-Faster-CGLU 模块。这种改进不仅提高了检测精度,而且优化了模型的运行效率。在检测头中加入共享卷积构造轻量化共享卷积检测头(lightweight shared convolutional detection

head, LSCD), 显著降低了计算量。

1 基于改进 YOLOv8n 的分心驾驶行为检测模型架构

YOLOv8^[22-23]是 Ultralytics 公司 YOLO 系列目标检测算法的最新成果,在其上一代的模型基础上做出了大量

有效的改进,大幅提高了模型的性能和适应性。YOLOv8 提供 5 种不同大小的模型版本: YOLOv8n^[24]、YOLOv8s^[25]、YOLOv8m、YOLOv8l 和 YOLOv8x。这些模型具有相同的基本架构,但深度和宽度不同。其中, YOLOv8n^[26-28]的检测速度最快,参数量最少。本文提出的方法也是基于对 YOLOv8n 的改进,如图 1 所示。

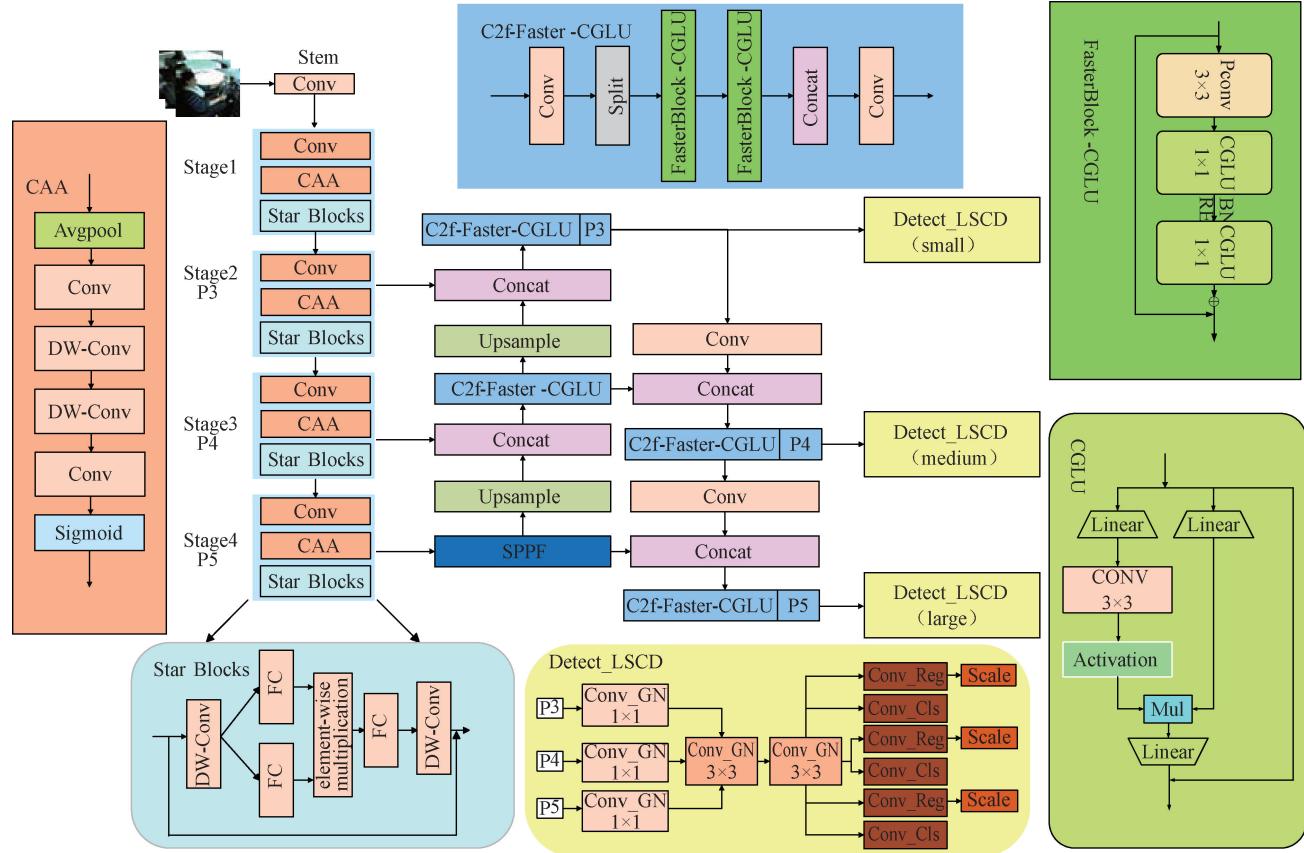


图 1 整体模型结构图

Fig. 1 Overall model structure diagram

1.1 StarNet 网络

为了进一步提高模型在目标检测任务中的性能,增强特征提取能力,减少计算量和参数量,本文使用 StarNet 代替原有的主干。StarNet 利用星型运算,在低维输入空间中有效捕获复杂、高维和非线性特征空间。在单层神经网络中,StarNet 将权重矩阵和偏差合并成一个单一实体,表示为 $w = \begin{bmatrix} W \\ B \end{bmatrix}$, 其中 W 表示权重部分, B 表示偏置项。相应地,将输入向量 \mathbf{x} 扩展为包含一个常数项(通常为 1)的矩阵, $\mathbf{x} = \begin{bmatrix} \mathbf{X} \\ 1 \end{bmatrix}$ 。通过这种方法,StarNet 实现了星形运算, $(w^T \mathbf{x}) \times (w^T \mathbf{x})$ 。为了简化分析,首先关注单输入单输出的场景。具体来说,定义 w_1 和 w_2 , $\mathbf{x} \in \mathbf{R}^{(d+1) \times 1}$, 其中 d 表示输入通道数。这可以很容易地扩展到多个输出

通道,其中 $w_1, w_2 \in \mathbf{R}^{(d+1) \times n}$ 。星型运算可以表示为:

$$\begin{aligned} w_1^T \mathbf{x} \times w_2^T \mathbf{x} &= \left(\sum_{i=1}^{d+1} w_1^i \mathbf{x}^i \right) \times \left(\sum_{j=1}^{d+1} w_2^j \mathbf{x}^j \right) = \\ &\sum_{i=1}^{d+1} \sum_{j=1}^{d+1} w_1^i w_2^j \mathbf{x}^i \mathbf{x}^j = \alpha_{(1,1)} \mathbf{x}^1 \mathbf{x}^1 + \dots + \alpha_{(d+1,d+1)} \mathbf{x}^{d+1} \mathbf{x}^{d+1} \end{aligned} \quad (1)$$

$$\alpha_{(i,j)} = \begin{cases} w_1^i w_2^j, & i = j \\ w_1^i w_2^j + w_1^j w_2^i, & i \neq j \end{cases} \quad (2)$$

其中, i 和 j 对通道进行索引, α 表示每个项目的系数。

星型运算最终可以展开为 $\frac{(d+2)(d+1)}{2}$ 种不同的项目组合,如式(1)所示。接下来,通过多层叠加,隐藏维度可以递归增加到接近无穷大。设 S_n 为第 n 次迭代的星型运算输出:

$$\begin{aligned}
 \mathbf{S}_1 &= \sum_{i=1}^{d+1} \sum_{j=1}^{d+1} \mathbf{w}_{(1,1)}^i \mathbf{w}_{(1,2)}^j \mathbf{x}^i \mathbf{x}^j \\
 \mathbf{S}_2 &= \mathbf{W}_{2,1}^T \mathbf{S}_1 \times \mathbf{W}_{2,2}^T \mathbf{S}_1 \\
 \mathbf{S}_3 &= \mathbf{W}_{3,1}^T \mathbf{S}_2 \times \mathbf{W}_{3,2}^T \mathbf{S}_2 \\
 &\dots \\
 \mathbf{S}_n &= \mathbf{W}_{n,1}^T \mathbf{S}_{n-1} \times \mathbf{W}_{n,2}^T \mathbf{S}_{n-1}
 \end{aligned} \tag{3}$$

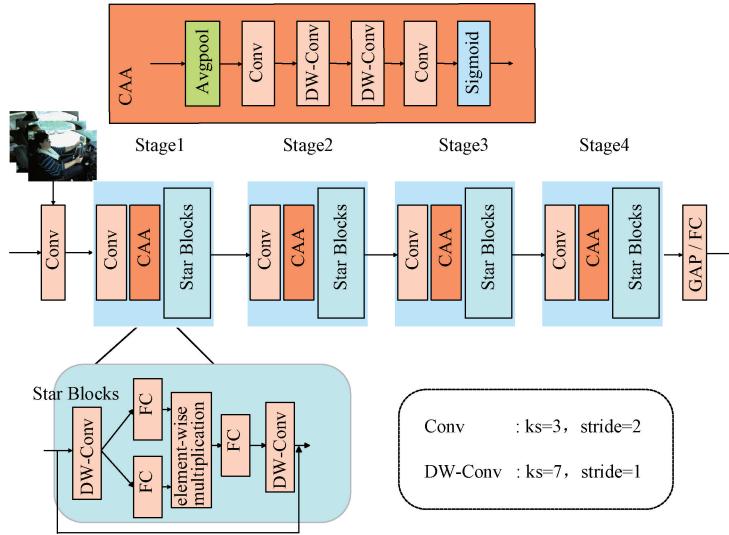


图 2 StarNet-CAA 网络结构图

Fig. 2 Diagram of the StarNet-CAA network structure

并且,为了更好的提取特征,本文将 CAA 注意力机制融合进 StarNet 网络中,形成 StarNet-CAA,其模型结构图如图 2 所示。首先,CAA 模块通过全局平均池化操作来获取整个特征图的统计信息。然后通过 1×1 卷积有助于提取细长形状物体的特征。并且使用深度可分离卷积增加感受野,并且不会显著增加计算复杂度。

1.2 C2f-Faster-CGLU 模块

本文使用 FasterBlock 与 CGLU 模块相结合替换 C2f 模块中的 Bottleneck 模块。

FasterBlock 在保持了较高的浮点运算次数,并提出了一种新的局部卷积(partial convolution, PConv)。原理框图如图 3 所示。它使用常规卷积从一部分输入通道中提取特征,在保持通道数量不变的情况下,将第一个通道作为整个特征映射进行计算。部分通道数为 c_p ,输入特征图和输出特征图具有相同的通道数。PConv 的 FLOPs 可表示为:

$$FLOPs = h \times w \times k^2 \times c_p^2 \tag{4}$$

其中, c_p 和 c 共同构成了分离比: $r = \frac{c_p}{c}$, 当 $r = \frac{1}{4}$ 时, PConv 的 FLOPs 只有普通卷积的 $1/16$,并且 PConv 的内存访问也更小:

$$h \times w \times 2c_p + k^2 \times c_p^2 \approx h \times w \times 2c_p \tag{5}$$

本文将 TransNeXt 中的 CGLU 模块与 FasterBlock 模

块相结合,得到 FasterBlock-CGLU 模块。CGLU 模块的具体结构如图 4 所示。在 GLU 门控分支的激活函数前加入 3×3 深度可变形卷积,使其结构符合门控通道注意的设计理念,并将其转化为基于最近邻特征的门控通道注意机制。并且将 C2f 中的 Bottleneck 替换为 FasterBlock-CGLU 模块,得到 C2f-Faster-CGLU 模块。PConv 使得 FasterBlock 具有速度更快和参数更少的优点,同时保持有限的精度损失。FasterBlock 中的 BN 模块允许它与相邻的 Conv 模块相结合,以加快推理速度。

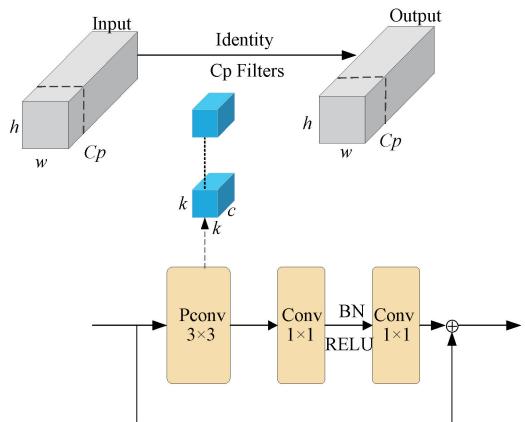


图 3 FasterBlock 结构图

Fig. 3 FasterBlock structure diagram

块相结合,得到 FasterBlock-CGLU 模块。CGLU 模块的具体结构如图 4 所示。在 GLU 门控分支的激活函数前加入 3×3 深度可变形卷积,使其结构符合门控通道注意的设计理念,并将其转化为基于最近邻特征的门控通道注意机制。并且将 C2f 中的 Bottleneck 替换为 FasterBlock-CGLU 模块,得到 C2f-Faster-CGLU 模块。PConv 使得 FasterBlock 具有速度更快和参数更少的优点,同时保持有限的精度损失。FasterBlock 中的 BN 模块允许它与相邻的 Conv 模块相结合,以加快推理速度。

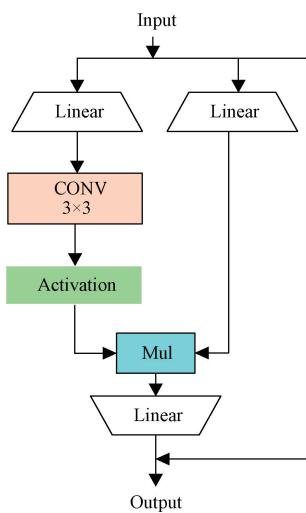


图 4 CGLU 模块结构图

Fig. 4 Diagram of the CGLU module structure

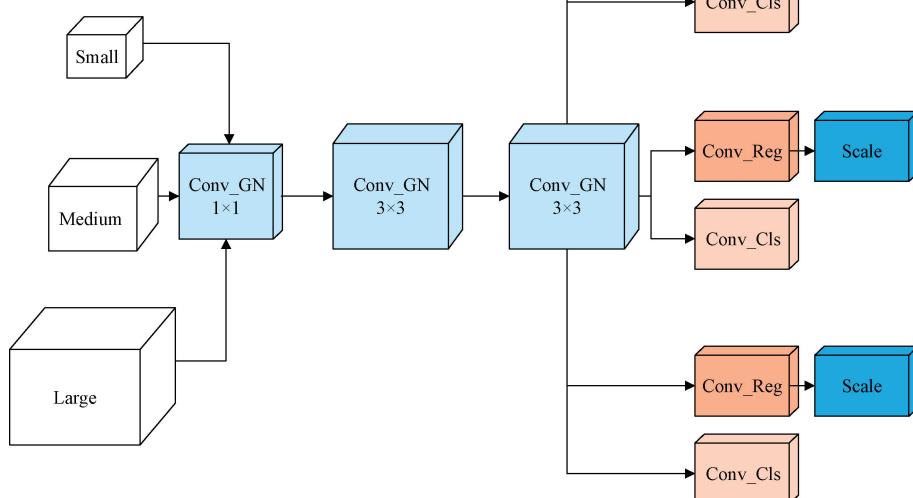


图 5 LSVD 模型结构图

Fig. 5 LSVD model structure diagram

保在不同尺度上可以进行一致与有效的检测分心驾驶行为。

2 实验与结果分析

在本节中,本文将首先介绍实验的软件和硬件设置。接下来,本文将介绍实验中使用的数据集,并提出分心驾驶检测的评估指标。随后,本文将替换 YOLOv8n 主干网络与其他主流模型进行比较,之后本文进行消融实验来验证各个模块的有效性,最后进行泛化实验来验证本文提出的方法的泛化能力。

2.1 实验环境及参数设置

实验平台信息: Windows 11, NVIDIA GeForce RTX

1.3 LSVD 检测头

在 YOLO 系列中,检测头通常包括 3 个分支,每个分支处理关于同一对象的不同尺度的信息。传统上,这些分支独立运行,可能导致模型参数的使用效率低下以及具有过度拟合的风险。为了解决这些问题,本文在检测头中引入了共享卷积的概念。模块的结构如图 5 所示。如图所示,小型、中型和大型目标通常通过 3 个独立的卷积层进行处理,每个层都有自己的一组参数,这大大增加了计算负荷。在分心驾驶行为检测中,LSVD 检测头有效减少了参数量,降低了模型复杂度,提高了计算性能。此外,参数共享有助于寻找全局最优解,并有助于避免过拟合的风险。

为了解决单个检测头中目标尺度不一致的问题,本文提出在每个检测头中合并一个 scale 层。这些模块配备了专门的卷积层,其任务是预测各自特征映射的边界框坐标和类概率。在此之后,整合到一个完整的张量中,并且确

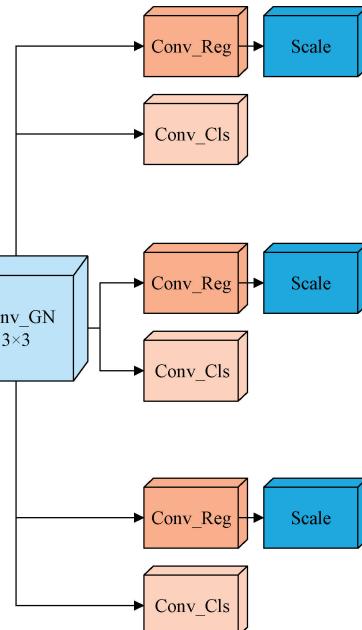


图 5 LSVD 模型结构图

Fig. 5 LSVD model structure diagram

4080 Super 16GB, Python 3.11, PyTorch 2.1.0, CUDA 12.1。参数设置:训练 150 轮, batch size 为 64, 优化器为 SGD。

2.2 数据集

1) StateFarm 数据集。本文使用 StateFarm 分心驾驶员数据集对模型进行了训练与验证,该数据集包含 26 名不同国籍驾驶员的 10 类分心驾驶行为。数据集中捕获的这些行为包括安全驾驶、在驾驶时发短信(用左手或右手)、在驾驶时打电话(用左手或右手)、与乘客交谈、调整中控、拿东西、整理妆容和喝水,如图 6 所示。该数据集的每一类包含大约 2 000 张图像,每张图像的高清分辨率为



图 6 StateFarm 数据集样张

Fig. 6 Sample of the StateFarm dataset

640×480 像素, 总共有 22 424 张高质量图像, 确保了数据集的丰富性和多样性。具体分布如表 1 所示。

表 1 数据集分类汇总

Table 1 Summary of the dataset

类别	名称	图片数量
a	安全驾驶	2 489
b	右手打字	2 267
c	右手打电话	2 317
d	左手打字	2 346
e	左手打电话	2 326
f	收音机	2 312
g	喝水	2 325
h	拿东西	2 002
i	整理妆容	1 911
j	聊天	2 129
Total		22 424

2) 100-Driver 数据集^[29]。为了证明模型的泛化能力, 本文进一步利用了 100-Driver 分心驾驶数据集。本数据集白天采用 RGB 格式, 晚上切换为近红外(NIR)格式。这个庞大的数据集包括从 4 个不同角度捕获的 100 名驾驶员的数据, 21 种不同的分心驾驶行为以及 1 种正常驾驶行为。本文选定的照片是从第 4 个相机角度拍摄的 10 类分心驾驶行为, 每个类随机选取 100 张白天和 100 张夜间照片, 测试算法的泛化能力。总共有 1 000 张 RGB 格式的图像和 1 000 张 NIR 格式的图像。

2.3 评价指标

本研究的目的是在保持较高检测精度的同时, 通过减小模型尺寸来优化 YOLOv8n。因此, 选择 mAP@0.5 和 mAP@0.5:0.95 作为评估模型精度的指标。mAP 表示平均精度。0.5% 和 0.5:0.95% 表示“预测边界框”与“真实边界框”的重叠度分别达到 50% 和 50~95%。此外, 使用

计算量与参数量对模型复杂度进行评估。精度和召回率作为次要指标, 定义如下, 其中 n 表示分类的类别数量:

$$Recall = \frac{TP}{TP + FN} \times 100\% \quad (6)$$

$$Precision = \frac{TP}{TP + FP} \times 100\% \quad (7)$$

$$AP = \int_0^1 P(r) dr \quad (8)$$

$$mAP = \frac{1}{n} \sum_{i=0}^n AP(i) \quad (9)$$

$$F1 = \frac{Precision \times Recall}{Precision + Recall} \quad (10)$$

2.4 替换 YOLOv8n 模型主干网络对比实验结果与分析

为了评估 StarNet-CAA 在替换 YOLOv8n 主干网络时的性能优势, 本文进行了一系列对比实验。选取了 YOLOv8n 原始模型和过去几年流行的五款轻量化模型作为参考。所有模型均在统一的实验环境下进行测试, 以确保结果的公平性。具体实验结果如表 2 所示。本文模型与 YOLOv8n+FasterNet 的 F1-Confidence 曲线如图 7 所示。

在本实验中, 与其他模型相比, HGNet^[30]略微降低了计算复杂度和参数复杂度。EfficientViT^[31] 和 MobileNetV4^[32]需要更多的参数量和计算资源, 但模型的性能并没有提高。FasterNet 在精度上比原始模型略有提高, 但计算和参数成本增加。与 YOLOv8n 相比, StarNet-CAA 具有更好的性能, 同时减少 19.5% 的计算量和 23.3% 的参数量。综上所述, 采用 StarNet-CAA 作为主干网络不仅增强了模型检测分心驾驶行为的性能, 同时缩小了模型的大小。另外观察 F1 得分, 可以看出 FasterNet 与本文模型差距不大, 但是从图 7 的置信度中可以看出, 本文所使用模型可以在置信度较高的水平中仍然保持准确率。

2.5 消融实验与分析

本研究以 YOLOv8n 为基准模型, 在保持软硬件环境

表 2 替换不同主干网络对比试验

Table 2 Comparison test of replacing different backbone networks

算法	mAP@0.5	mAP@0.5:0.95	GFLOPs	Parameters/M	P	R	F1
YOLOv8n	98.6	68.8	8.2	3.0	98.4	98.4	98.4
YOLOv8n+HGNetV2	98.8	68.8	6.9	2.3	98.6	98.8	98.7
YOLOv8n+EfficientViT	98.8	68.7	9.4	4.0	98.6	98.8	98.7
YOLOv8n+FasterNet	99.1	69.0	10.7	4.2	99.1	99.1	99.1
YOLOv8n+MobileNetV4	98.8	69.0	23.2	5.7	98.8	99.2	98.9
本文模型	99.1	69.1	6.6	2.2	99.3	99.1	99.2

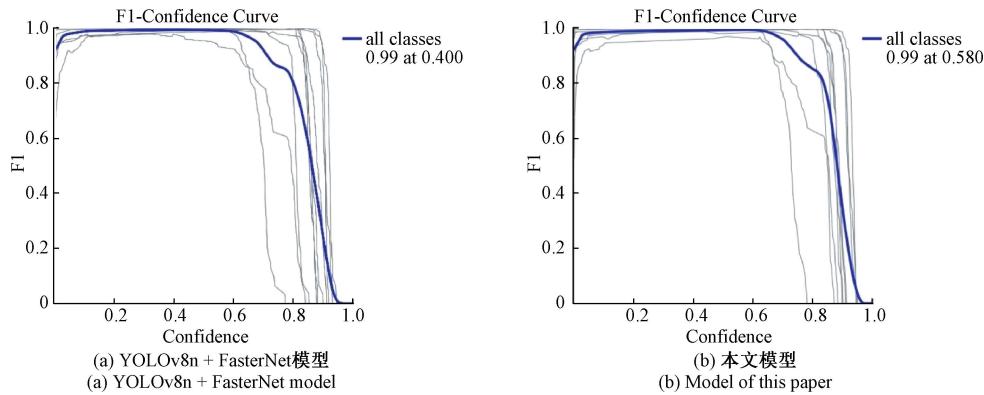


图 7 F1-Confidence 曲线对比

Fig. 7 Comparison of F1-Confidence curves

一致的前提下,进行消融实验,评估模型各组成部分对其整体性能的影响。表 3 给出了 8 组实验数据。值得注意的是,单独替换每个模块都降低了计算量和参数量。具体而言,将 StarNet-CAA、C2f-Faster-CGLU 和 LSCD 集成后,计算量分别降低了 19.5%、15.8% 和 26.8%,参数量分别减少了 23.3%、16.7% 和 23.3%。此外,与基准模型相比,

这些模块也略微提高了精度。实验结果说明了本文提出改进策略的有效性和实用性。遵循简化模型的目标,在保持检测精度的同时降低计算与参数负担。图 8 为 YOLOv8n 与本文所使用模型的混淆矩阵,从图中可以看出,改进之后的模型可以更好的识别分心驾驶行为,识别的准确率更高。图 9 为本文改进模型的整体结果曲线。

表 3 消融实验

Table 3 Ablation experiment

StarNet-CAA	C2f-Faster-CGLU	LSCD	mAP@0.5	mAP@0.5:0.95	GFLOPs/G	Parameters/M
	YOLOv8n		98.6	68.8	8.2	3.0
✓			99.1	69.1	6.6	2.3
	✓		99.0	69.4	6.9	2.5
		✓	99.1	69.3	6.0	2.3
✓	✓		99.2	69.1	6.2	2.0
✓		✓	99.2	69.2	4.9	1.6
✓	✓	✓	99.2	69.1	5.3	1.9
✓	✓	✓	99.4	69.5	4.8	1.6

2.6 泛化实验

泛化实验使用 100-Driver 数据集。由于本文用于训练的照片是从副驾驶的角度拍摄的,并且包括十种分心驾驶行为,因此本文从 100-Driver 数据集中随机选择具有相应角度和分类的照片。每个类别随机选择 100 张照片,包括不

同车辆、司机以及不同照明条件的图片。此外,100-Driver 数据集包括两种类型的图像:白天拍摄的 RGB 照片和夜间拍摄的 NIR 照片,每种模式有 1 000 张图像,所有这些图像都用于验证。在实验环境不变的情况下,分别使用 YOLOv8n 和本文所提出的模型进行实验。实验结果如

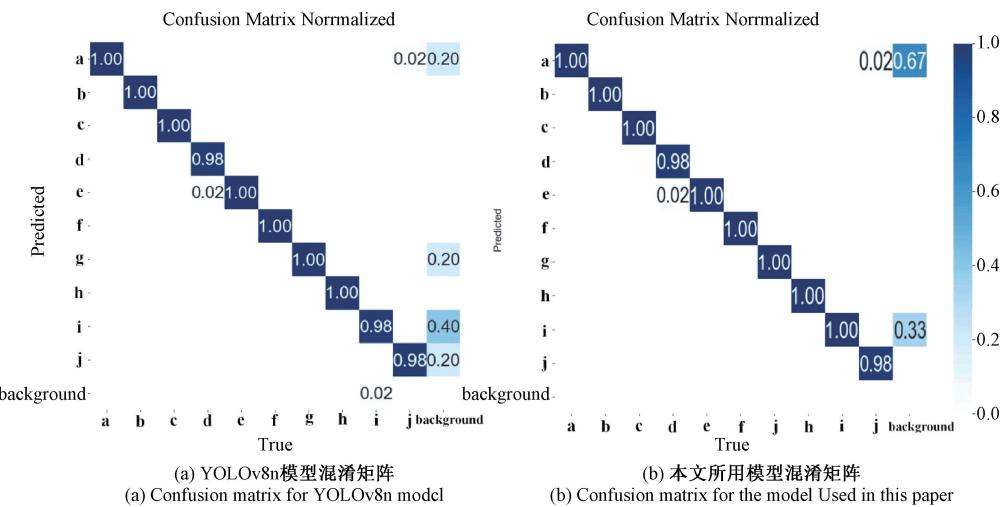


图 8 混淆矩阵的对比

Fig. 8 Comparison of confusion matrices

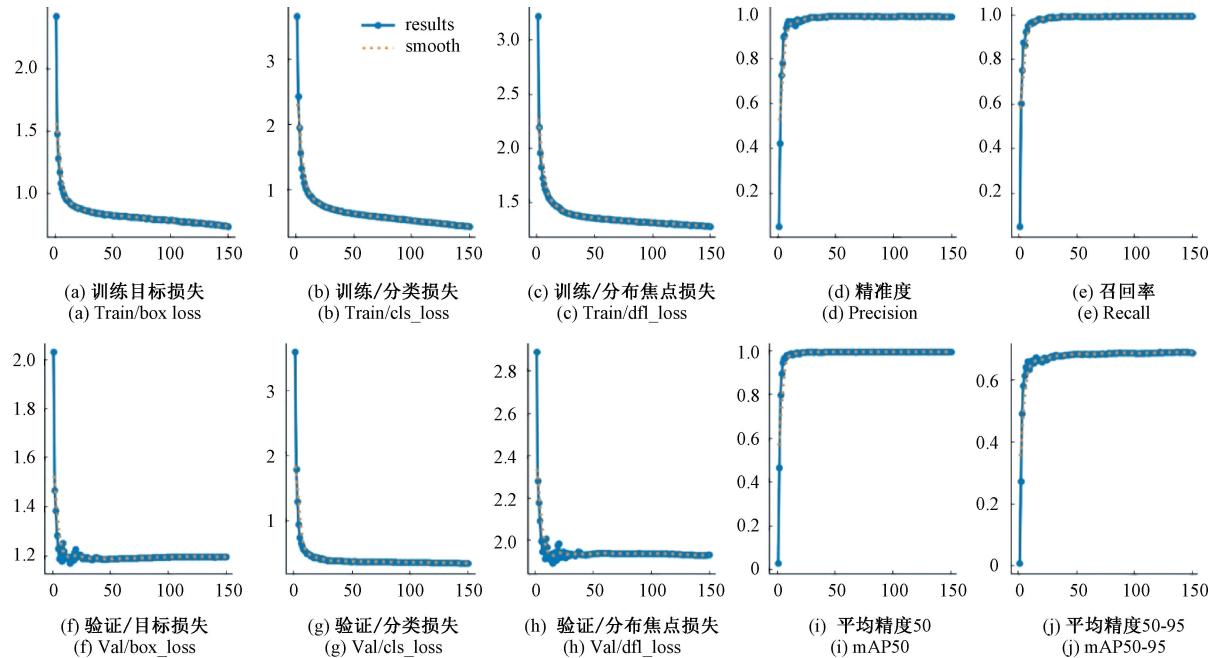


图 9 本文改进模型的整体结果曲线

Fig. 9 The overall result curve of the improved model in this paper

表 4 所示。

表 4 白天与夜间的泛化实验

Table 4 Day and night generalization experiments

算法	mAP@0.5(白天)	mAP@0.5(夜晚)
YOLOv8n	24.3	18.2
本文模型	32.1	27.1

在本文的泛化实验中,本文模型与 YOLOv8n 模型相比,白天和夜间的识别准确率分别提高了 7.8% 和 8.9%。

结果表明该模型具有较强的泛化能力。值得注意的是,白天过于强烈或微弱的照明可能导致识别精度下降。同样,在夜间,该模型往往难以识别分心驾驶行为,经常将车载终端误认为是智能手机。图 10 为实验结果样张。

2.7 Jetson Nano 嵌入式硬件平台测试

将本文改进之后的模型部署到 Jetson Nano 嵌入式平台中进行实际测试,如图 11 所示。并且测试了本文模型分别在计算机与 Nano 设备上的运行帧率,如表 5 所示。实验结果表明,在资源非常有限的 Jetson Nano 设备上也可以使用。



图 10 泛化实验样张

Fig. 10 Sample of the generalization experiment



图 11 实际嵌入式平台测试样例图

Fig. 11 Actual embedded platform test sample diagram

表 5 帧率表

Table 5 Frame rate table

算法	Nano 帧率/fps	计算机帧率/fps
本文模型	9.96	110.29

3 结 论

本文提出的算法提高了对车辆驾驶员分心驾驶行为的识别精度,降低了模型的计算复杂度和参数复杂度。具体改进包括:首先,将CAA注意力机制融合进StarNet,形成StarNet-CAA,再用StarNet-CAA取代YOLOv8n的主干网络,采用星型运算从低维输入中捕获高维和非线性特征空间。其次,将颈部C2f模块中的残差块替换为FasterBlock-CGLU,减少了计算量。最后,在检测头中引入共享卷积的思想,大大降低了计算成本。实验结果表明,在标准光照条件下,本文所提出的算法在大量减少计算负担的情况下,可以较为准确的识别出分心驾驶行为,

并且在部署在资源有限的实际平台上进行了测试。

然而,所提出的方法在强光以及弱光条件下仍需要进一步改进。为了克服这些限制,未来的研究将集中在融合不同的分心驾驶数据集、提高复杂条件下模型稳定性以及探索更加先进网络结构。

参考文献

- [1] KOESDWIADY A, BEDAWI S M, OU C, et al. End-to-end deep learning for driver distraction recognition[C]. In Image Analysis and Recognition: 14th International Conference, ICIAR 2017, Montreal, QC, Canada, 2017: 11-18.
- [2] HESHAM E H, ABOURLNAGA Y, SAAD M H, et al. Driver distraction identification with an ensemble of convolutional neural networks[J]. Journal of Advanced Transportation, 2019(1): 4125865.
- [3] 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, 2021.
- [4] ZHANG ZH W. Research on abnormal driving behavior detection method based on machine vision[D]. Changsha: Hunan University, 2021.
- [5] 白中浩, 王韫宇, 张林伟. 基于图卷积网络的多信息融合驾驶员分心行为检测[J]. 汽车工程, 2020, 42(8): 1027-1033.
- [6] BAI ZH H, WANG W Y, ZHANG L W. Driver distraction behavior detection with multi-information fusion based on graph convolution networks [J]. Automotive Engineering, 2020, 42(8): 1027-1033.
- [7] TRAN D, DO H M, LU J, et al. Real-time detection of distracted driving using dual cameras [C]. 2020 IEEE/RSJ International Conference on Intelligent

- Robots and Systems, IROS 2020, Las Vegas, NV, USA, 2020: 2014-2019.
- [6] LI P H, YANG Y F, GROSU R, et al. Driver distraction detection using octave-like convolutional neural network[J]. IEEE Transactions on Intelligent Transportation Systems, 2021, 23(7): 8823-8833.
- [7] 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: 109436.
- [8] 尹智帅, 钟恕, 聂琳真, 等. 基于人体姿态估计的分心驾驶行为检测[J]. 中国公路学报, 2022, 35(6): 312-323.
- YIN ZH SH, ZHONG SH, NIE L ZH, et al. Distracted driving behavior detection based on Human Pose Estimation[J]. China Journal of Highway and Transport, 2022, 35(6): 312-323.
- [9] 张志远. 基于深度学习的机动车驾驶员危险驾驶行为检测方法研究[D]. 杭州: 杭州电子科技大学, 2024.
- ZHANG ZH Y. Research on detection method of dangerous driving behavior of motor vehicle drivers based on deep learning [D]. Hangzhou: Hangzhou Electronic Science and Technology University, 2024.
- [10] PENG P, HUANG C, DING, W, et al. Distracted driving detection based on the fusion of deep learning and causal reasoning[J]. Information Fusion, 2023, 89: 121-142.
- [11] 李少凡, 高尚兵, 张莹莹. 用于驾驶员分心行为识别的姿态引导实例感知学习[J]. 中国图象图形学报, 2023, 28(11): 3550-3561.
- LI SH F, GAO SH B, ZHANG Y Y. Pose-guided instance perception learning for driver distracted behavior recognition [J]. Journal of Image and Graphics, 2023, 28(11): 3550-3561.
- [12] LOU CH CH, NIE X. Research on lightweight-based algorithm for detecting distracted driving behaviour [J]. Electronics, 2023, 12(22): 4640.
- [13] DU Y J, LIU X F, YI Y W, et al. Optimizing road safety: Advancements in lightweight YOLOv8 models and GhostC2f design for real-time distracted driving detection[J]. Sensors, 2023, 23(21): 8844.
- [14] GUO X, MA M, ZHANG J Q, et al. YOLOX-B: A better YOLOx model for real-time driver behavior detection[C]. ICASSP 2023-2023 IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP 2023, Rhodes Island, Greece, 2023: 1-5.
- [15] FU Y Y, RAN T, XIAO W D, et al. GD-YOLO: An improved convolutional neural network architecture for real-time detection of smoking and phone use behaviors [J]. Digital Signal Processing, 2024, 151: 104554.
- [16] DU Y J, LIU X F, YI Y W, et al. Incorporating bidirectional feature pyramid network and lightweight network: A YOLOv5-GBC distracted driving behavior detection model [J]. Neural Computing and Applications, 2024, 36(17): 9903-9917.
- [17] LI R J, YU CH D, QIN X R, et al. YOLO-SGC: A dangerous driving behavior detection method with multiscale spatial-channel feature aggregation [J]. IEEE Sensors Journal, 2024, 24(21): 36044-36056.
- [18] 贺宜, 鲁曼可, 高嵩, 等. 基于 MobileViT-CA 模型的营运车辆驾驶人分心行为检测[J]. 中国公路学报, 2024, 37(1): 194-204.
- HE Y, LU M K, GAO S, et al. Distraction behavior detection of commercial vehicle driver based on MobileViT-CA model[J]. Chinese Journal of Highway and Transport, 2024, 37(1): 194-204.
- [19] MA X, DAI X Y, BAI Y, et al. Rewrite the stars[C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition, CVPR 2024, Seattle, Washington, USA, 2024: 5694-5703.
- [20] CAI X H, LAI Q X, WANG Y W, et al. Poly kernel inception network for remote sensing detection[C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition, CVPR 2024, Seattle, Washington, USA, 2024: 27706-27716.
- [21] CHEN J R, KAO S H, HE H, et al. Run, don't walk: Chasing higher FLOPS for faster neural networks[C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition, CVPR 2023, New Orleans, Louisiana, USA, 2023: 12021-12031.
- [22] 李志星, 杨啸龙, 李天昊, 等. 基于 YOLOv8 的煤矿用钢丝绳损伤检测算法[J]. 电子测量技术, 2024, 47(9): 1-9.
- LI ZH X, YANG X L, LI T H, et al. Damage detection algorithm of steel wire rope in coal mine based on YOLOv8 [J]. Electronic Measurement Technology, 2024, 47(9): 1-9.
- [23] 曾勇杰, 范必双, 杨涯文, 等. 改进 YOLOV8 算法在风机叶片缺陷检测上的应用[J/OL]. 电子测量与仪器学报, 1-12 [2024-10-28]. <http://kns.cnki.net/kcms/detail/11.2488.tn.20240829.1721.004.html>.
- ZENG Y J, FAN B SH, YANG Y W, et al. Application of improved YOLOV8 algorithm in defect detection of wind turbine blades[J/OL]. Journal of

- Electronic Measurement and Instrumentation, 1-12 [2024-10-28]. <http://kns.cnki.net/kcms/detail/11.2488.tn.20240829.1721.004.html>.
- [24] 梁天添, 杨淞淇, 钱振明. 基于改进 YOLOv8s 的恶劣天气车辆行人检测方法[J]. 电子测量技术, 2024, 47(9): 112-119.
- LIANG T T, YANG S Q, QIAN ZH M. Vehicle and pedestrian detection method in bad weather based on improved YOLOv8s [J]. Electronic Measurement Technology, 2024, 47(9): 112-119.
- [25] 赵栓峰, 李乐平, 王茂权. 基于深度学习的驾驶员分心行为检测方法[J]. 电子测量技术, 2024, 47(9): 145-153.
- ZHAO SH F, LI L P, WANG M Q. Driver distraction behavior detection method based on deep learning [J]. Electronic Measurement Technique, 2024, 47(9): 145-153.
- [26] 王德伟, 刘小芳. 基于改进 YOLOv8n 的带钢表面缺陷检测方法[J]. 国外电子测量技术, 2024, 43(7): 158-169.
- WANG D W, LIU X F. Steel strip surface defect detection method based on improved YOLOv8n [J]. Foreign Electronic Measurement Technology, 2024, 43(7): 158-169.
- [27] 姜源, 付波, 权轶, 等. 基于改进 YOLOv8n 的 PCB 缺陷检测算法[J]. 国外电子测量技术, 2024, 43(6): 22-32.
- JIANG Y, FU B, QUAN Y, et al. PCB defect detection algorithm based on improved YOLOv8n [J]. Foreign Electronic Measurement Technology, 2024, 43(6): 22-32.
- [28] 叶彦斐, 胡龙癸, 张成龙. 基于改进 YOLOv8n-Pose 的轨道作业人员跨轨安全动作识别[J]. 国外电子测量技术, 2024, 43(8): 181-188.
- YE Y F, HU L G, ZHENG CH L. Cross-track safety action recognition of track operators based on improved YOLOv8n-Pose [J]. Foreign Electronic Measurement Technology, 2024, 43(8): 181-188.
- [29] WANG J, LI W J, LI F, et al. 100-driver: A large-scale, diverse dataset for distracted driver classification [J]. IEEE Transactions on Intelligent Transportation Systems, 2023, 24(7): 7061-7072.
- [30] YAO T, LI Y H, PAN Y W, et al. Hgnet: Learning hierarchical geometry from points, edges, and surfaces [C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition, CVPR 2023, New Orleans, Louisiana, USA, 2023: 21846-21855.
- [31] LIU X Y, PENG H W, ZHENG N X, et al. Efficientvit: Memory efficient vision transformer with cascaded group attention [C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition, CVPR 2023, New Orleans, Louisiana, USA, 2023: 14420-14430.
- [32] QIN D F, LEICHNER C, DELAKIS M, et al. MobileNetV4-universal models for the mobile ecosystem [J]. ArXiv preprint arXiv, 2404.10518, 2024.

作者简介

沈 雯(通信作者),讲师,博士,主要研究方向为深度学习,数字信号处理、图像加密、图像识别等。

E-mail:qianshen@hyit.edu.cn

张 磊,硕士研究生,主要研究方向为深度学习,数字信号处理、图像识别等。

E-mail:leizhang1201@foxmail.com