

# 基于 Faster-YOLO11 的危险驾驶行为检测算法

马 骁<sup>1</sup>  
MA Xiao

## 摘要

全球范围内, 交通事故发生率不断增长, 成为日益突出、不容忽视的严重社会问题。基于此, 文章针对危险驾驶行为检测提出了一种目标检测算法——Faster-YOLO11, 旨在提高危险驾驶行为检测的效率和准确性。算法以 YOLO11 作为基础, 使用 FasterNet 作为主干网络, 在 SPPF (spatial pyramid pooling-fast) 中 LSK 注意力机制, 并将颈部网络的 C3K2 替换为 RCSOSA (reparameterized convolution based on channel shuffle and one-Shot aggregation), 以及后处理阶段的 Soft-NMS, 实现了在保持高检测精度的同时, 显著提升了模型的推理速度, 为危险驾驶检测领域提供了一种新的解决方案。

## 关键词

目标检测; 危险驾驶行为检测; YOLO11; 深度膨胀卷积; 大型可分离核注意; 减少通道空间对象注意力

doi: 10.3969/j.issn.1672-9528.2025.03.042

## 0 引言

随着现代社会的发展和生活节奏的加快, 交通事故已成为全球范围内一个日益严重的社会问题。其中, 由于驾驶员危险驾驶引发的交通事故占较大比例, 给人们的生命安全带来了巨大威胁。因此, 开发有效的危险驾驶检测系统对于减少交通事故、提高道路安全具有重要意义。近年来, 随着深度学习技术的快速发展, 基于计算机视觉的危险驾驶检测技术取得了显著进展。研究者们不断探索新的算法, 结合深度学习的优势, 以提高危险驾驶检测系统的准确性和实时性。

危险驾驶检测主要分为疲劳检测与危险驾驶行为检测两个方面, 其中疲劳检测是指对驾驶员面部表情进行分析, 观察驾驶员有无闭眼、打哈欠等疲劳行为; 危险驾驶行为主要是指驾驶过程中出现的抽烟、打电话等行为。对于疲劳检测任务, 主流的处理方法主要分为基于关键点分析的方法与基于面部整体纹理的方法, 李孟成等人<sup>[1]</sup>基于 YOLOv5 目标检测算法和 dlib 人脸识别库, 通过分析人脸朝向、位置、眼睛开合度等数据, 实时计算驾驶员是否存在疲劳驾驶和危险驾驶行为, 并及时给出安全提示; 许旻等人<sup>[2]</sup>设计了一种基于深度学习的危险驾驶状态检测预警系统, 通过深度学习中的人脸检测等相关算法模型以及多技术配合研究, 实现了驾驶员危险驾驶状态的快速、准确检测和预警; 张樱己<sup>[3]</sup>针对驾驶分心问题, 提出了基于眼动与手部行为识别的驾驶分心检测算法。研究提出了多通道卷积神经网络用于识别驾驶员注视区域, 并基于手部行为特征提出危险驾驶行为识别模型。

危险驾驶行为检测则是典型的目标检测任务, 宋巍<sup>[4]</sup>聚

焦于基于深度学习的违规驾驶行为视觉识别方法, 从系统工程和人因工程的角度分析了违规驾驶行为, 并构建了违规驾驶行为视觉识别方法体系结构。研究提出了基于 ResNet 和视觉注意力模块的危险驾驶行为识别方法, 以及基于改进 ShuffleNet 轻量化模型的高效识别方法, 旨在提高识别精度和效率。郑天赐<sup>[5]</sup>针对驾驶员手持通话检测问题, 设计了基于剪枝算法、注意力机制和特征融合的目标检测网络模型, 并优化算法移植到嵌入式设备。研究创建了 hand\_phone 驾驶员打电话数据集, 并设计了新型目标检测网络 YOLO-PAI 和 YOLO-RESDAI, 提高了小目标物体的检测精度和速度。

在实际应用中, 驾驶员检测会受到外部场景与不均匀光照的影响, 暗光环境下检测难度大幅提升。唐天俊等人<sup>[6]</sup>提出了一种基于深度学习的夜间危险驾驶行为检测算法, 该算法包括弱光增强模块和检测模块, 通过提高图像曝光度和基于 NanoDet-Plus 模型的检测, 实现了夜间弱光环境下的高准确率检测, 且模型参数量小, 检测速度快, 可部署在移动设备上实时检测。

基于上述分析, 危险驾驶行为检测中的两个任务可以合并为统一的目标检测任务, 同时算法应当兼顾实时性、准确性, 并且有着较强的鲁棒性, 同时对遮挡、小目标等较难识别的场景也有较好的性能。本文提出了一种基于 YOLO11n 的改进型目标检测算法——Faster-YOLO11, 旨在提高危险驾驶行为检测的效率和准确性。算法以 YOLO11 作为基础, 使用 FasterNet 作为主干网络, 在 SPPF 中 LSK 注意力机制, 并将颈部网络的 C3K2 替换为 RCSOSA, 以及后处理阶段的 Soft-NMS, 实现了在保持高检测精度的同时, 显著提升了模型的推理速度, 为危险驾驶检测领域提供了一种新的解决方案。

1. 辽宁公安司法管理干部学院 辽宁沈阳 110000

### 1 Faster-YOLO11 算法设计

针对危险驾驶行为检测任务，本文提出了一种高效的目标检测算法 Faster-YOLO11，算法以当前的 SOTA 目标检测模型 YOLO11n 为基础，主干网络使用 FasterNet，在不降低检测精度的前提下大幅降低网络的参数量与计算量，以达到更好的推理效率；在 SPPF 模块中引入 LSK 注意力机制，有效提升主干网络的特征提取与特征融合性能；使用 RCSOSA 模块替换颈部网络中的 C3K2 模块，丰富梯度流的同时加强多尺度的特征重组性能，同时提升了对香烟这种纵横比差异过大的小目标的识别能力；在后处理阶段，使用 Soft-NMS 提高候选框的精度。Faster-YOLO11 的网络结构如图 1 所示。

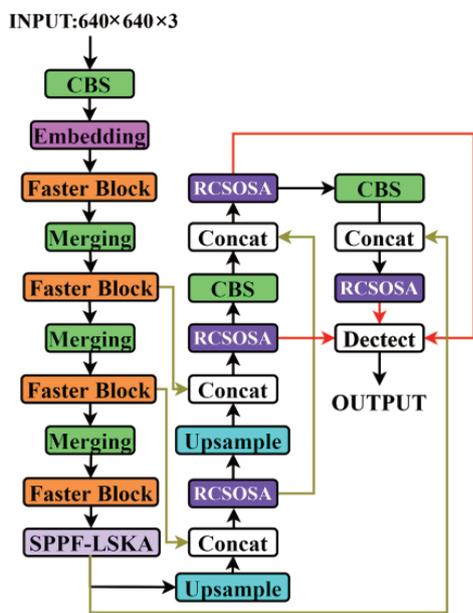


图 1 Faster-YOLO11 网络结构图

对比 YOLO11n，Faster-YOLO11 通过更精简的结构设计实现了轻量化与精度的有效提升，更加适用于危险驾驶行为检测任务中对驾驶员表情、吸烟与打电话等危险驾驶行为的检测。

#### 1.1 FasterNet 主干网络

FasterNet<sup>[7]</sup> 是一种轻量化的主干网络，能够在不牺牲检测精度的前提下有效提高算法的推理速度。FasterNet 对传统卷积进行频繁内存访问导致的性能瓶颈进行优化，其核心是部分卷积 (PConv)，这种新型卷积方法通过仅处理输入通道的一部分来减少计算量和内存访问，从而降低整体的计算复杂度。与传统方法不同，FasterNet 不仅关注减少浮点运算 (FLOPs)，还强调提升每秒浮点运算数 (FLOPS)，以实现更快的处理速度。FasterNet 的架构包含 4 个层次化的阶段，每个阶段由一系列 FasterNet 块组成，其结构如图 2 所示。

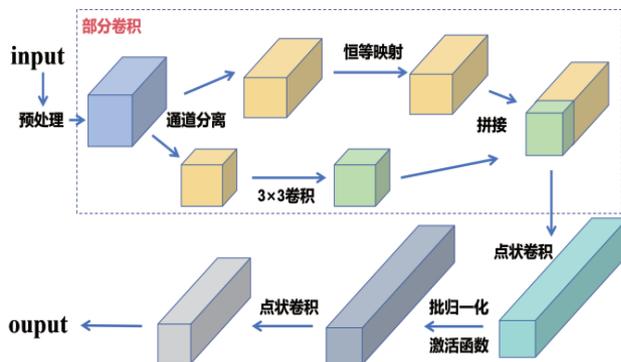


图 2 FasterNet 块结构图

FasterNet 块以嵌入或合并层开始，每个块中包含一个部分卷积层后跟两个点状卷积层。这种设计类似于倒置残差块，其中间层具有扩展的通道数量，并放置了 Shortcut 以重用输入特征。

#### 1.2 SPPF-LSKA 模块

LSKA (large separable kernel attention)<sup>[8]</sup> 是一种视觉注意力网络模块，旨在解决大内核卷积在计算效率方面的挑战。与传统的 2D 深度卷积不同，LSKA 通过将卷积核分解为级联的水平和垂直 1-D 卷积核，有效降低了参数数量和内存占用，同时保持了计算效率。这种设计使得 LSKA 更倾向于捕捉对象的形状特征而非纹理，从而提高了模型的鲁棒性和泛化能力。

LSKA 能够有效捕捉图像中的长距离依赖，这一点通过其有效感受野的生成方法得到了验证。随着卷积核尺寸的增加，LSKA 能够覆盖更广泛的区域，表明其在捕捉长距离依赖方面的有效性。LSKA 继承了 LKA 的设计，包括空间注意力和通道注意力，这两种策略能够根据上下文依赖适应性地重新校准特征权重，与自注意力相比，具有更低的参数和计算复杂度。

在 SPPF (spatial pyramid pooling-fast) 结构中，LSKA 的引入进一步增强了特征聚合能力，使得模型能够更有效地捕捉和融合不同层次的特征信息。LSKA 通过生成注意力图来加权原始特征图，集中注意力于图像中的重要部分，从而提升特征的表达能力。LSKA 的优势在于其计算效率，通过将深度卷积层的二维卷积核分解为级联的水平和垂直一维卷积核，减少了计算和内存占用，同时能够实现与标准 LKA 模块相当的性能。此外，LSKA 设计使模型在卷积核大小增加时更倾向于关注物体的形状而非纹理，这在许多视觉任务中至关重要，因为形状信息通常对目标识别更为关键。LSKA 在 SPPF 中最大池化层之后进一步加工特征，改善了不同尺度特征图之间的联系，从而提升特征融合效果。SPPF-LSKA 模块结构如图 3 所示。

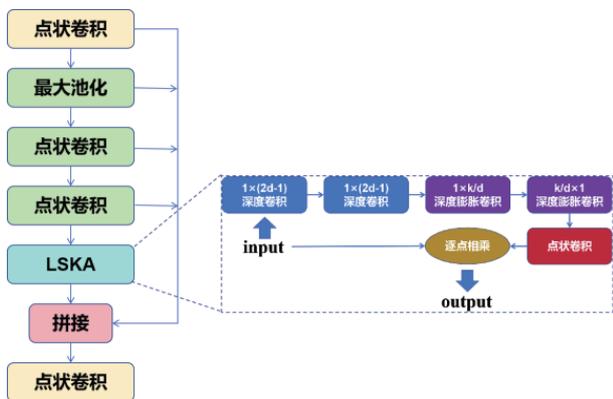


图3 SPPF-LSKA 结构图

### 1.3 RCSOSA 模块

RCSOSA (reparameterized convolution based on channel shuffle and one-Shot aggregation)<sup>[9]</sup> 是一种创新的神经网络模块，专门设计来提升目标检测任务的速度和准确性。该模块融合了通道混洗和重参数化卷积技术 (RCS) 以及一次性聚合 (one-shot aggregation, OSA) 策略。在训练阶段，RCSOSA 利用多分支结构学习丰富的特征表示，而在推理阶段，通过结构化重参数化简化为单一分支，以减少内存消耗。通过堆叠 RCS 模块，RCSOSA 实现了特征级联，加强了不同层之间的信息流动，并通过上采样和下采样实现了多尺度特征融合，促进了不同预测特征层之间的信息交换，从而提高了检测精度。RCSOSA 的设计注重计算效率，通过维持输入通道和最小输出通道的数量来降低内存访问成本。

在工作机制上，RCSOSA 在训练时利用多分支拓扑架构学习特征信息，在推理时通过重参数化简化架构，加快推理速度。RCSOSA 结构如图4所示。

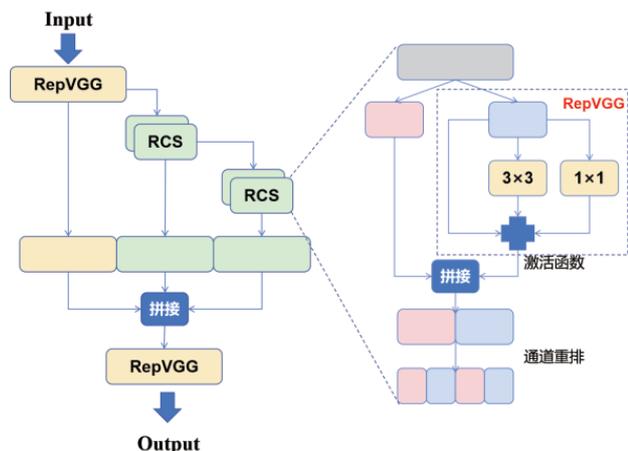


图4 RCSOSA 结构图

与 C3k2 模块相比，RCSOSA 具有更优的计算效率与更强的特征提取能力。在计算效率方面，RCSOSA 通过通道混洗和重参数化卷积的设计显著降低了计算复杂度，特别是在推理阶段，经通道分割和混洗操作，实现了计算复杂度减半

的同时保持了通道间的信息交换，使得处理高维特征更加高效。其次，RCSOSA 模块通过堆叠 RCS，不仅确保了特征的重用，还增强了相邻层特征之间不同通道的信息流动，这有助于提取更丰富的特征信息，并降低内存访问成本。此外，RCSOSA 采用一次性聚合策略，减少了重复的特征计算和存储需求，提高了网络的计算和能源效率，同时有效聚合不同层次的特征，增强了模型的语义信息提取能力。

### 1.4 Soft-NMS

Soft-NMS (soft non-maximum suppression)<sup>[10]</sup> 是一种非极大抑制 (NMS) 的改进算法，旨在解决标准 NMS 在目标检测中因硬阈值导致的候选框丢失问题。与传统 NMS 直接移除得分低且与高分框重叠的检测框不同，Soft-NMS 通过应用一个得分衰减函数来降低重叠检测框的得分，而不是直接将其移除。这种方法首先选取得分最高的检测框作为基准，然后根据剩余检测框与基准框的 IoU (交并比) 来更新它们的得分。采用高斯惩罚函数作为得分衰减函数。

Soft-NMS 的核心思想是在保留尽可能多的候选框的同时，降低与高分检测框重叠的低得分框的优先级。这种逐步衰减得分的方法减少了因硬阈值而误删目标的情况，提高了检测的准确性，尤其是在处理小目标或高密度目标时更为有效。

## 2 实验结果分析

### 2.1 数据集与训练平台

在 Windows10 系统上使用 GeForce GTX 3060 (8 GB) 显卡与 PyTorch 平台进行模型训练与验证，训练参数如表1所示。

表1 网络训练参数

参数	数值
Epoch/ 训练轮数	300
Batchsize/ 批大小	32
lr0/ 初始学习率	0.001
Momentum/ 学习率动量	0.937
box/iou 损失系数	7.5
cls/ 分类损失系数	0.5

在视频网站上收集 1 000 小时的车内驾驶员视角视频，为保证数据多样性，行车场景包括白天、黑夜、黄昏等，车型包括轿车、公交车、卡车等，同时保证驾驶员性别与年龄的多样性，最终获得 12 000 张图片，并将其中的驾驶员表情、抽烟、打电话进行标注，标注类别包含正常驾驶、闭眼、打哈欠、手机、香烟五类，使用数据增强手段扩充数据到 15 000 张，根据 6:3:1 的比例划分训练集、测试集与验证集。

## 2.2 评价标准

使用召回率、平均精度均值 (mAP) 以及推理速度作为模型的评价指标。召回率衡量的是模型预测正确的正类别样本占有所有实际正类别样本的比例。AP 是在不同召回率水平上精确率的平均值, mAP 是所有类别 AP 的平均值, 特别适用于多类别目标检测任务。

## 2.3 对比试验

将 Faster-YOLO11 与主流目标检测模型 YOLOv5n、YOLOv8n、YOLO11n 进行对比试验, 实验结果如表 2 所示。

表 2 基于危险驾驶数据集的对比实验结果

模型	召回率 /%	mAP50 /%	mAP50-95 /%	推理速度 /ms
YOLOv5n	71.2	81.1	48.6	27.9
YOLOv8n	71.8	81.3	48.9	27.6
YOLO11n	72.4	81.9	49.1	26.6
Faster-YOLO11	73.5	82.5	49.6	24.3

对比实验结果表明, YOLO11n 全面优于其他 YOLO 系列模型, 同时本文提出的 Faster-YOLO11 相较于基准模型 YOLO11n 也有明显提升, 召回率提升了 1.52%, mAP50 与 mAP50-95 分别提升了 0.73% 与 1.02%, 推理速度减少了 8.64%。

## 2.4 消融实验

为验证各改进模块的有效性, 对各模块进行消融实验, 实验结果如表 3 所示。

表 3 基于危险驾驶数据集的消融实验结果

Faster Net	LSKA	RCS OSA	Soft-NMS	mAP50 /%	推理速度 /ms
√				81.3	22.6
√	√			81.8	23.4
√	√	√		82.2	24.3
√	√	√	√	82.5	24.3

消融实验结果表明, 各个模块的改进有效优化了推理速度或推理精度, FasterNet 作为主干减少了 15% 的推理速度, LSKA 的引入使 mAP50 提升了 0.5%, RCSOSA 的替换也使 mAP50 提升了 0.4%, Soft-NMS 的使用使 mAP50 提升了 0.3%。需要说明的是, Soft-NMS 作为一种后处理技术, 其计算时间不参与模型推理速度的统计。

## 3 结论

本文针对危险驾驶行为检测领域提出了一种高效的目标检测算法 Faster-YOLO11, 模型基于 YOLO11n, 主干网络替换为 FasterNet, 并在 SPPF 中引入 LSKA 注意力机制; 使用 RCSOSA 模块替换颈部网络中的 C3K2 模块; 在后处理阶

段, 使用 Soft-NMS 提高候选框的精度。算法在精度与速度方面对比 YOLO11n 均匀明显提升。根据对比实验, Faster-YOLO11 对比基准模型召回率提升了 1.52%, mAP50 与 mAP50-95 分别提升了 0.73% 与 1.02%, 推理速度减少了 8.64%, Faster-YOLO11 有着明显的性能提升。在实际应用中, 危险驾驶检测算法通常部署于端侧设备, 因此还需要对模型进行剪枝、量化等优化手段, 并且针对不同目标硬件 (CPU、GPU、NPU) 进行针对性移植, 因此还要考虑改进模块在上述优化手段中的适用性。

## 参考文献:

- [1] 李孟成, 罗甜, 张琰, 等. 基于 YOLOv5 的疲劳与危险驾驶行为检测系统研究 [J]. 内蒙古科技与经济, 2023(6):107-110.
- [2] 许昊, 马晨东, 罗紫琳. 基于深度学习的危险驾驶状态检测预警系统设计 [J]. 无线互联科技, 2022, 19(12): 50-52.
- [3] 张樱己. 基于眼动与手部行为识别的驾驶分心检测算法研究 [D]. 济南: 济南大学, 2022.
- [4] 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.
- [5] 郑天赐. 基于卷积神经网络驾驶员手持通话检测算法的研究与应用 [D]. 北京: 中国矿业大学, 2023.
- [6] 唐天俊, 宋平. 基于深度学习的夜间危险驾驶行为检测算法 [J]. 电脑与信息技术, 2024, 32(4):9-13.
- [7] CHEN J R, KAO S H, HE H, et al. Run, don't walk: chasing higher FLOPS for faster neural networks[C]//2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE, 2023:12021-12031.
- [8] LAU K W, PO L M, REHMAN Y A U. Large separable kernel attention: rethinking the large kernel attention design in CNN[J]. Expert systems with applications, 2024, 236: 121352.
- [9] KANG M, TING C M, TING F F, et al. RCS-YOLO: a fast and high-accuracy object detector for brain tumor detection[C]//Medical Image Computing and Computer Assisted Intervention—MICCAI 2023:26th International Conference. NewYork: ACM, 2023:600-610.
- [10] BODLA N, SINGH B, CHELLAPPA R et al. Soft-NMS—improving object detection with one line of code[C]//2017 IEEE International Conference on Computer Vision (ICCV). Piscataway: IEEE, 2017:5562-5570.

## 【作者简介】

马骁 (1981—), 男, 辽宁沈阳人, 硕士, 副教授, 研究方向: 信息技术研究。

(收稿日期: 2024-11-22)