

DOI:10.19651/j.cnki.emt.2519268

## 基于深度学习的高速公路交通事故检测方法研究\*

凌锐<sup>1,2</sup> 闫坤<sup>1,2</sup> 梁宏宇<sup>1,2</sup> 韦焯淇<sup>1,2</sup> 郝航勃<sup>1,2</sup>

(1.桂林电子科技大学信息与通信学院 桂林 541004; 2.桂林电子科技大学卫星导航定位与位置服务国家地方联合工程研究中心 桂林 541004)

**摘要:** 现有单阶段深度模型的交通事故检测在高速公路场景中误报率高、计算冗余大,严重制约了实际部署。为此,本文提出一种基于双阶段架构的高速公路交通事故检测方法,采用“静止车辆筛选+外观特征识别”的处理流程。第一阶段结合 YOLO11 与 Bot-SORT,实现对车辆的检测与跟踪,并通过帧间速度分析筛选出疑似事故的静止车辆。第二阶段引入改进模型 YOLO-EA,仅对静止车辆执行外观检测,并采用多帧投票机制提高判断稳定性与鲁棒性。YOLO-EA 基于 YOLO11 架构,引入 EAS-Stem 模块与 AWD-Conv 模块,前者增强了输入阶段边缘与轮廓的特征提取,后者提升了下采样效率并在保留关键特征同时降低计算负担。实验结果表明,YOLO-EA 在 Precision、mAP@0.5 和 mAP@0.5:0.95 上分别提升 10.9%、3.4% 和 2.8%,参数量减少 21%;在构建的事故视频数据集上,本方法的事事故识别率达 81.25%,相对于单阶段检测策略误报率降低 24.46%。该方法在准确性与推理效率之间实现良好平衡,具备较强的实际部署潜力。

**关键词:** 交通安全;事故检测;多目标跟踪;深度学习;高速公路

**中图分类号:** TN911.73;TP391.41;U491.3 **文献标识码:** A **国家标准学科分类代码:** 520.2040;580.80

## Research of deep learning-based methods for highway traffic accident detection

Ling Rui<sup>1,2</sup> Yan Kun<sup>1,2</sup> Liang Hongyu<sup>1,2</sup> Wei Zhuoqi<sup>1,2</sup> Hao Hangbo<sup>1,2</sup>

(1. School of Information and Communication, Guilin University of Electronic Technology, Guilin 541004, China;

2. National and Local Joint Engineering Research Center for Satellite Navigation, Positioning and Location Services,

Guilin University of Electronic Technology, Guilin 541004, China)

**Abstract:** Existing single-stage deep models for traffic accident detection often suffer from high false alarm rates and computational redundancy in highway scenarios, severely limiting their practical deployment. To address these issues, this paper proposes a two-stage traffic accident detection method tailored for highways, following a "stationary vehicle filtering+appearance-based recognition" strategy. In the first stage, YOLO11 and Bot-SORT are integrated to detect and track vehicles, and inter-frame speed analysis is used to identify stationary vehicles as potential accident candidates. In the second stage, an improved model named YOLO-EA is introduced to perform appearance-based detection exclusively on the stationary vehicles, combined with a multi-frame voting mechanism to enhance stability and robustness. Built upon the YOLO11 architecture, YOLO-EA incorporates an EAS-Stem module and an AWD-Conv module. The former enhances edge and contour extraction in the input stage, while the latter improves downsampling efficiency by retaining critical features and reducing computational cost. Experimental results show that YOLO-EA improves Precision, mAP@0.5 and mAP@0.5:0.95 by 10.9%, 3.4% and 2.8% respectively, while reducing parameter count by 21%. On the constructed accident video dataset, the proposed method achieves an accident recognition rate of 81.25%, with a 24.46% reduction in false alarm rate compared to single-stage detection strategies. This method achieves a favorable balance between accuracy and inference efficiency, demonstrating strong potential for real-world deployment.

**Keywords:** traffic safety; accident detection; multi-object tracking; deep learning; highway

## 0 引言

随着机动车保有量持续增长,高速公路上的交通事故

频发,且事故多发生在高速行驶状态下,极易引发重大人员伤亡和财产损失。近年来,基于深度学习的视频分析技术已被广泛应用于交通事故检测场景,显著提高了自动化水

收稿日期:2025-07-03

\* 基金项目:国家自然科学基金(62101147)、广西自然科学基金(桂科 2020GXNSFAA159146)、广西创新驱动发展专项(桂科 AA21077008)、教育部重点实验室基金(CRKL190108)项目资助

平。但现有主流方法多采用单阶段检测策略,直接对每帧图像中的所有车辆执行外观分析,导致在高速公路复杂背景下误报率高、计算冗余严重,难以兼顾检测精度与推理效率,限制了实际部署效果。因此,亟需设计一种高效、低误报的检测架构,以提升交通事故检测系统在高速公路场景下的实用性与稳定性。

目前事故检测的研究主要包含以下 3 类方法:

1) 基于外观特征,通过目标检测或图像分类模型识别车辆在事故中产生的显著视觉变化,如碰撞损伤、车体变形、侧翻等。Rezaei 等<sup>[1]</sup>基于 Faster R-CNN,并在时序上加入动态空间注意力来捕捉事故先兆。柳圣等<sup>[2]</sup>先用 YOLOv5 检测候选目标,基于目标关系构建图卷积递归网络(graph convolution recurrent network, GCRN),再以贝叶斯神经网络(Bayesian neural network, BNN)预测事故概率。Arifeen 等<sup>[3]</sup>采用 GoogLeNet 提取帧级深度特征,用一类支持向量机(support vector machine, SVM)进行事故检测,对检测出的事故帧再用多类 SVM 细分碰撞类型。Chen 等<sup>[4]</sup>提出 YOLOv8 改进模型 YOLO-SED,通过 SPD-Conv 模块和 SE 注意力机制提升远距离小目标检测能力。该类方法适用于可视异常明显的事故场景,结构成熟,但对轻微碰撞等低显著性情形识别效果有限,易受遮挡与光照干扰。

2) 基于运动特征,通过分析车辆速度、加速度、轨迹偏移、边界框 IoU 变化等运动信息判断事故发生。Essam 等<sup>[5]</sup>提出融合碰撞估计与暴力流描述符(violent flow descriptor, ViF)的双层验证框架,采用 YOLOv3 与 MOSSE 实现轨迹预测,并通过 SVM 分类器进行碰撞决策。Ghahremanzhad 等<sup>[6]</sup>利用 YOLOv4 进行车辆检测,采用卡尔曼滤波与匈牙利算法跟踪目标,并通过轨迹冲突分析完成事故识别。Dharmadasa 等<sup>[7]</sup>利用 YOLOv5 + StrongSORT,通过分析车辆速度、加速度、轨迹异常及位置区域异常等特征,实现实时事故识别与预警。Nallasivan<sup>[8]</sup>通过 Haar 级联分类器识别车辆目标,车辆影响位置生成器构建车辆轨迹,最后利用 CNN 分类器分析轨迹特征实现事故检测。该类方法适用于动态行为异常识别,实时性强,但在车辆静止或低速碰撞场景中效果有限,且易受遮挡、ID 切换等多目标跟踪问题影响。

3) 融合视觉外观与轨迹运动信息的多模态检测方法,通过跨模态特征协同提升检测精度。Tee 等<sup>[9]</sup>构建混合交通场景事故识别框架,分别提取帧级外观、轨迹运动与时序特征并融合判断事故。Zhang 等<sup>[10]</sup>先用 YOLOv5 + DeepSORT 生成轨迹,利 CNN 与 Transformer 进行多模态特征融合。王晨等<sup>[11]</sup>则采用改进双流网络结构,通过辅助网络与三元损失优化外观特征提取,结合光流网络增强运动特征捕捉能力。此类方法在复杂环境与多种事故类型下表现优越,但模型结构复杂、计算开销大,对数据标注与同步要求严格,部署成本相对较大。

综上,针对现有方法存在的局限性,提出一种基于双阶

段架构的交通事故检测方法。第 1 阶段使用 YOLO11<sup>[12]</sup>与 Bot-SORT<sup>[13]</sup>对车辆检测与追踪,结合帧间速度估算识别出静止车辆作为潜在事故目标;第 2 阶段使用改进模型 YOLO-EA 对静止车辆进行外观分析,并通过多帧投票机制提升判定稳定性与准确性。该双阶段策略具有如下优势:1) 降低计算资源消耗:仅对静止车辆进行外观检测,避免了对所有目标逐帧推理;2) 提升误报抑制能力:通过多帧时序一致性判断,降低 YOLO-EA 模型的偶发误判影响;3) 规避轨迹断裂风险:以静止状态作为事故触发信号替代对完整轨迹的依赖,有效避免了因 ID 切换或遮挡导致的轨迹中断问题;4) 相比于融合轨迹与图像多模态特征的方法,本文结构更加简洁,计算资源占用更低,更适于部署于实时事故检测场景。

需要指出的是,本文所提出的双阶段检测策略适用于高速公路中车辆因发生事故而最终而停下的典型场景。在极端恶劣天气(如暴雨、大雾)导致能见度下降,或夜间及隧道出入口等光照剧烈变化的场景下,第 1 阶段的车辆检测与跟踪精度,以及第 2 阶段 YOLO-EA 模型的外观特征提取能力均存在精度下降的风险,从而对整体性能带来挑战。此外,当交通极度拥堵、存在大量车辆长时间静止时,第 1 阶段筛选出的候选目标数量会显著增加,虽仍可依靠第 2 阶段进行判别,但计算负荷也会相应提升。上述边界条件限定了本方法的当前适用范围,其在不同环境下的鲁棒性与泛化能力将是未来研究中有待进一步深入验证和优化的重要方向。

## 1 双阶段交通事故检测方法

### 1.1 系统框架

本文基于“静止车辆筛选+外观特征识别”的双阶段交通事故检测系统,整体结构如图 1 所示,第 1 阶段:静止车辆筛选。系统基于 YOLO11 与 Bot-SORT 算法对高速公路监控视频中的车辆进行目标检测与多目标跟踪,并通过帧间位置变化估算车辆速度。当车辆的速度  $Speed$  低于设定的静止速度阈值  $S$ (设为 1.5 像素/帧),且累积静止时间  $Stationary\ time$  超过设定阈值  $T$ (设为 3 s)时,系统将该车标记为“静止状态”。该策略可有效过滤因临时减速、堵车等非事故行为导致的短时静止情况,确保进入下一阶段的车辆具备较高的事故可能性,从而提升系统整体的准确率与效率。第 2 阶段,外观特征识别。对已标记为静止状态的车辆,启用事故外观检测模型 YOLO-EA,判断其是否存在事故特征(如翻车、碰撞、损坏等)。仅检测车辆静止后的连续  $Window-N$  帧(设为 10 帧),并统计其中被判定为“事故”的帧数  $Accident-F$ 。若事故帧数超过设定的阈值  $Pos-K$ (设为 7 帧),则该车被最终判定为发生事故,并关闭该目标的后续外观检测以节省计算资源;否则,判定为非事故状态。该多帧投票机制可有效抵抗单帧误报,提高检测稳定性与判定鲁棒性,尤其适用于模型存在一定误差或

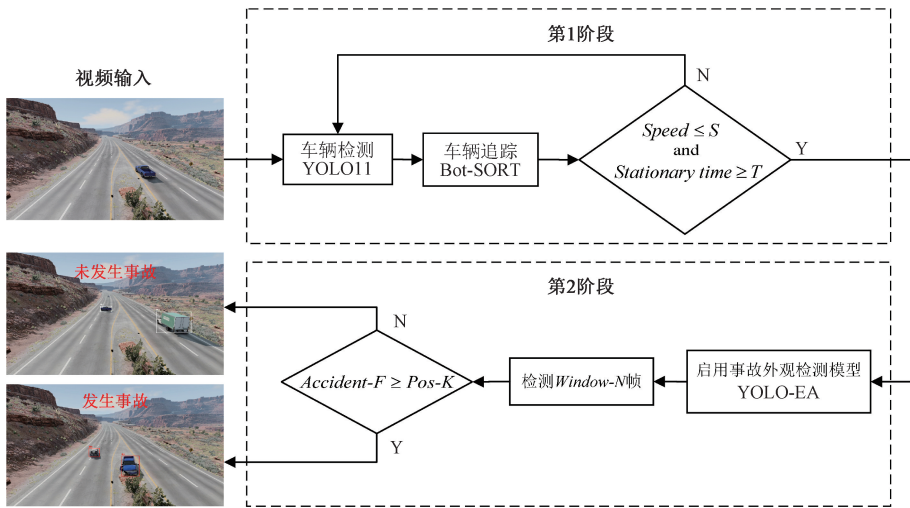


图 1 系统框图

Fig.1 System block diagram

视频质量较低的实际应用场景。

为保证系统判定的可靠性,本文对关键触发参数进行了合理设定与验证:

1)静止速度阈值  $S$ :其设定主要考虑到目标检测框在静态场景下存在轻微抖动的问题。当  $S < 0.5$  像素/帧时,系统难以识别真实静止车辆;当  $S > 2$  像素/帧时,又会导致误报数量明显增加。通过对多段高速公路监控视频的对比实验,发现当  $S = 1.5$  像素/帧时,能够较好地区分真实静止车辆与因检测抖动引起的误差。

2)静止时间阈值  $T$ :其设定主要用于区分临时停车/缓慢行驶与真实事故导致的停滞。过低(如 1 s)易将临时停车、堵车车辆误判为静止;过长则在短时事故视频中难以捕捉到真正的静止车辆。综合考虑实验视频的时长分布及车辆停滞特征,选取  $T = 3$  s 作为折中值,在检测准确性与时效性之间取得合理折中。

3)多帧投票窗口  $Window-N$ :固定 10 帧,以抑制单帧检测可能出现的偶然误判,在保证系统实时性的同时增强判定稳定性。

4)投票阈值  $Pos-K$ :通过实验对比优化为 7(详见第 2.4 节),在  $Window-N = 10$  的条件下兼顾了较高的识别率与较低的误报率。

上述参数设置的有效性与合理性均在后续系统评估中得到了验证。

## 1.2 第 1 阶段:静止车辆筛选

### 1)YOLO11 车辆检测

YOLO11 网络结构可划分为 3 大部分:Backbone、Neck 与 Head,如图 2 所示。Backbone 部分包括多个卷积模块(Conv)、残差结构 C3k2、空间金字塔池化模块(spatial pyramid pooling-fast,SPPF)<sup>[14]</sup>,及增强通道注意力感知的 C2PSA 模块,旨在自底向上逐层提取金字塔多尺度特征图,逐步压缩图像空间维度,捕捉丰富的高层语义信息。

C3k2 是一种轻量化多分支残差结构,具备良好的特征表达能力,在兼顾检测精度同时有效控制计算开销;SPPF 通过多尺度池化操作串联不同感受野特征,有效增强模型对大目标的感知能力;C2PSA 则引入通道注意力机制,强化特征图中关键通道的响应权重,提升模型对显著区域的关注度。NecK 这一部分采用路径聚合网络(path aggregation network,PAN-FPN)结构,并结合动态加权融合机制,实现高层语义与底层细节特征的高效整合。该结构通过上采样与下采样操作完成多尺度特征的交互,显著提升对小目标和边缘目标的检测性能。Head 部分在每个尺度分支上均输出 3 个分支:类别概率、目标置信度和边界框位置回归参数。YOLO11 采用 anchor-free 检测机制与动态标签分配策略,减少先验框设计依赖,简化训练流程的同时提升了定位精度与收敛效率。综上,YOLO11 通过多尺度融合、注意力增强与轻量化设计,在处理密集交通场景时表现出色<sup>[15]</sup>,具备较高的检测精度与实时性,能够胜任高速公路复杂交通场景下多类车辆的快速识别任务,为后续静止状态判定与事故分析提供了高质量候选区域。

### 2)Bot-SORT 车辆跟踪

Bot-SORT 是一种基于检测的多目标跟踪算法,本文结合 YOLO11 获取每帧图像中的车辆目标框及其置信度分数,进一步完成目标的时序跟踪与 ID 关联。该算法在 ByteTrack<sup>[16]</sup> 框架基础上,融合了 SORT 的高效结构,并对运动建模与匹配机制进行了增强,适合本研究中第 1 阶段静止车辆的持续状态监测。Bot-SORT 的主要增强点包括:(1)基于 Kalman 滤波器的轨迹预测与更新:通过状态空间建模,使用 Kalman 滤波器预测目标在下一帧的位置,实现对短期遮挡的稳健恢复;(2)引入相机运动补偿机制(camera motion compensation,CMC):考虑到高速公路监控画面中可能存在轻微抖动或摄像头平移,算法引入稀疏

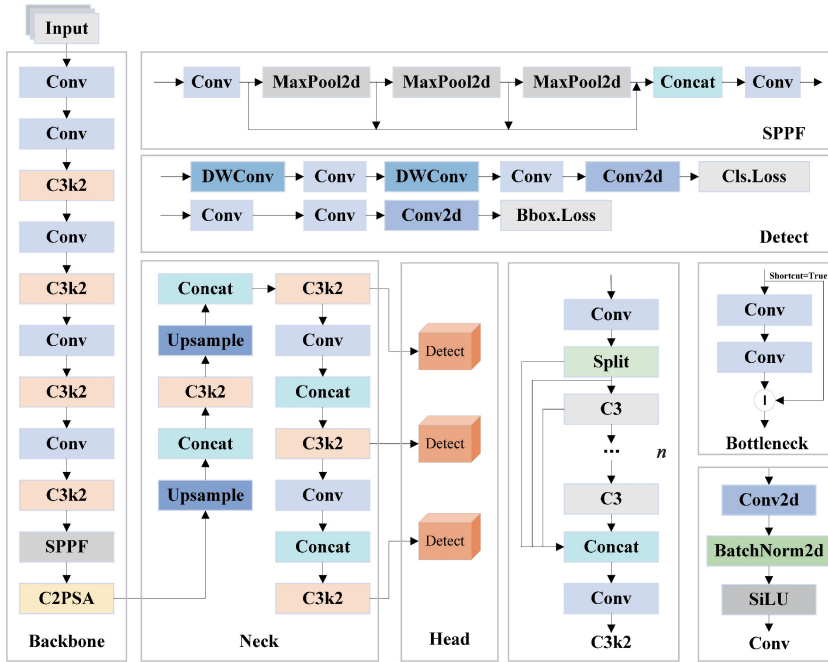


图 2 YOLO11 结构图

Fig. 2 YOLO11 architecture diagram

光流 (sparseOptFlow) 进行全局运动估计, 减少因背景运动造成的匹配错误; (3) 融合 IoU 距离与 ReID 余弦相似度的双重匹配策略: 检测框之间的关联不仅依赖几何重叠度 (IoU), 还计算目标外观特征的余弦距离。通过加权融合两者进行级联匹配, 使得在目标间距接近、尺度变化或遮挡时仍能维持 ID 的持续一致性。

其核心流程如图 3 所示。置信度低于 0.1 的检测框将直接丢弃<sup>[17]</sup>; 高于 0.5 的高置信度检测框将参与第 1 轮匹配, 与 Kalman 滤波器预测的轨迹进行 IoU 与外观特征的联合匹配。若匹配成功则视为同一目标并

更新其轨迹状态; 若高置信度检测框未能匹配任一已有轨迹, 则将其视为新目标并初始化新轨迹。未完成第 1 次匹配的历史轨迹将进入第 2 轮匹配, 该轮仅使用 IoU 距离与低置信度检测框进行匹配, 若匹配成功则继续跟踪; 若匹配失败, 则该轨迹将被保留最多 30 帧, 若在此期间仍无法完成有效匹配, 则该目标轨迹将被判定为丢失并删除。Bot-SORT 兼顾外观特征与运动状态的融合策略, 保障实时性的同时提升目标 ID 的持续一致性与遮挡恢复能力, 为后续的静止状态判定提供了高质量的轨迹数据支持。

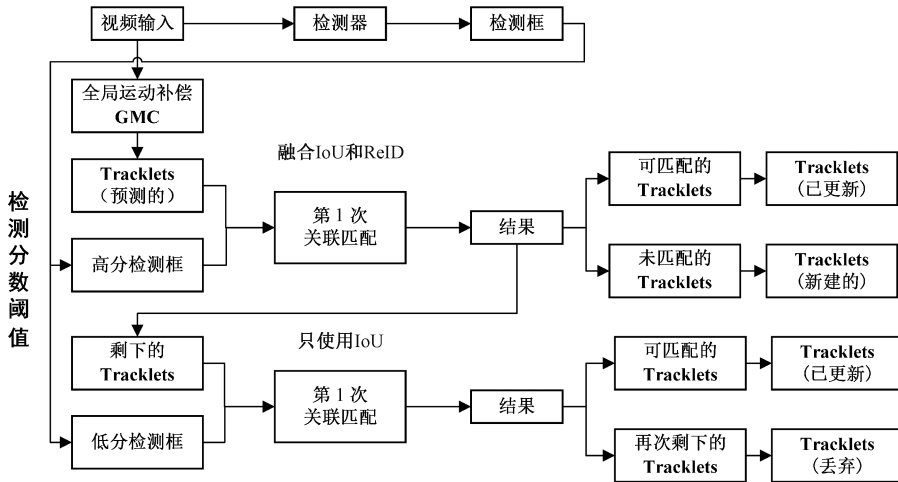


图 3 Bot-SORT 流程图

Fig. 3 Bot-SORT flowchart

为了确保方法透明性, Bot-SORT 的关键参数配置如 表 1 所示。

表 1 Bot-SORT 关键参数配置

Table 1 Critical parameter settings of Bot-SORT

参数	设定值	说明
track_high_thresh	0.5	高置信度阈值,用于第 1 轮匹配。
track_low_thresh	0.1	低置信度阈值,用于第 2 轮匹配。
new_track_thresh	0.6	新轨迹初始化的最低置信度要求。
match_thresh	0.8	IoU 匹配阈值。
proximity_thresh	0.5	空间接近度阈值。
appearance_thresh	0.25	外观相似度阈值。
track_buffer	30	轨迹最大保留帧数。
fuse_score	True	是否融合检测置信度与 IoU 距离

1.3 第 2 阶段:外观特征识别

在完成第 1 阶段的静止车辆筛选后,为进一步判断是否发生交通事故,本文设计了一种改进的外观检测模型 YOLO-EA。结构图如图 4 所示,该模型在 YOLO11 的基础上引入两个结构性改进模块:1)边缘与空间感知模块(edge and spatial-aware stem, EAS-Stem),用于增强输入阶段对边缘与轮廓特征的感知能力,提升模型在复杂背景下对细节变化的检测灵敏度;2)自适应加权下采样卷积模块(adaptive weighted downsampled convolution, AWD-Conv),用于构建轻量化的中高层特征提取路径,在有效减少参数量的同时保持对关键结构的表达能力。相关研究表明,在复杂交通场景中对 YOLO 系列进行轻量化与结构改进可同时提升鲁棒性与实时性<sup>[18]</sup>。

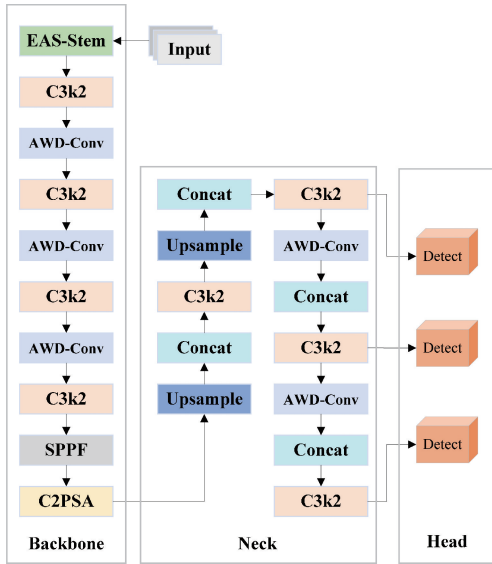


图 4 YOLO-EA 结构图

Fig. 4 YOLO-EA architecture diagram

1) EAS-Stem 模块

传统卷积神经网络在处理细粒度目标时,常因浅层特征提取能力有限而导致对关键细节的漏检,尤其在识别车辆碰撞、裂痕、损坏等微小异常时效果不佳。为提升模型

在输入阶段的特征感知能力,本文提出一种边缘与空间感知模块 EAS-Stem,用于替换 YOLO11 中的前两层标准卷积结构。EAS-Stem 模块设计了两个并行子路径,分别感知图像的局部边缘变化与空间结构信息,并通过特征融合机制进行高效整合,在兼顾计算效率的同时增强了输入层对显著视觉变化的响应能力。模块结构如图 5 所示。

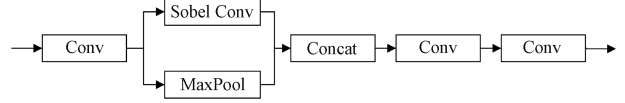


图 5 EAS-Stem 模块

Fig. 5 EAS-Stem module

(1) Sobel Conv 分支:提取图像边缘特征。结构如图 6 所示。

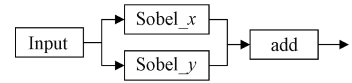


图 6 Sobel Conv 分支

Fig. 6 Sobel Conv branch

该分支通过 Sobel Conv 模块进行边缘提取,使用固定的  $3 \times 3$  Sobel 卷积核输入特征图进行处理,分别计算图像在水平方向与垂直方向的梯度图:

$$\begin{cases} G_x = I * K_x \\ G_y = I * K_y \end{cases} \quad (1)$$

式中:  $I$  为输入图像;本文中,“ $*$ ”表示卷积运算;“ $\times$ ”仅用于尺寸与维度;  $K_x$  与  $K_y$  分别为 Sobel 卷积核,如式(2)所示,用于提取图像水平与垂直方向的边缘信息。

$$K_x = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}, K_y = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix} \quad (2)$$

最终边缘特征图通过通道融合表示为:

$$G = \sqrt{G_x^2 + G_y^2} \quad (3)$$

在实现上,该模块采用 3D 深度可分离卷积(Conv3D with groups=channel)对每通道独立计算梯度,确保边缘信息在通道级别保持一致性,且卷积核参数 requires\_grad=False,不参与反向传播,从而实现结构性边缘建模。

(2) MaxPool 分支:提取空间结构信息。该分支利用最大化池化操作提取图像的宏观结构信息。为了确保池化后特征尺寸与 Sobel 分支对齐,使用 ZeroPad2d 对输入的边界进行对称填充,随后施加  $2 \times 2$  核、stride=1 的最大池化操作:

$$F_{pool} = MaxPool_{2 \times 2} (Pad(I)) \quad (4)$$

式中:  $F_{pool}$  表示最大化池化分支输出特征图,  $Pad(\cdot)$  表示使用 ZeroPad2d 进行的对称填充操作。

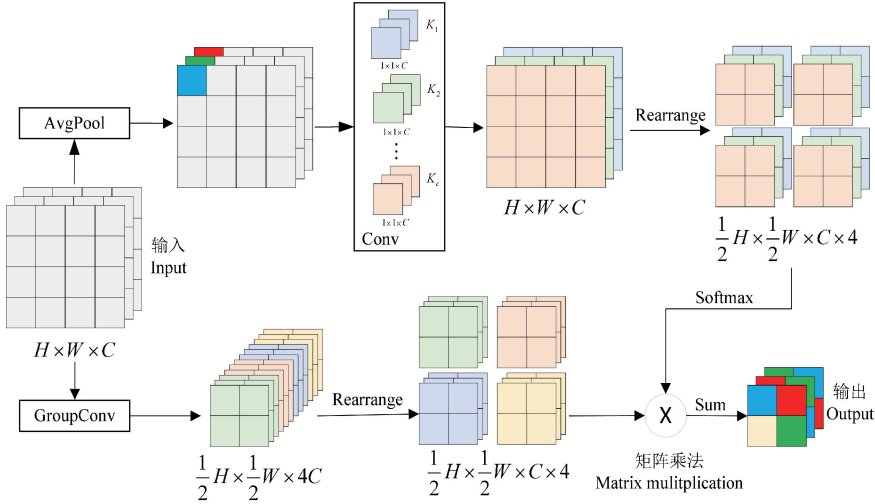
此分支可有效压缩冗余背景噪声,同时保留车辆轮廓、姿态等空间结构,为模型提供全局感知辅助信息。

(3) 特征融合模块:上述两个分支的输出在通道维度拼接后,进入两个连续卷积模块进行融合与压缩:

$$\mathbf{x}_{out} = Conv3(Conv2(Concat[\mathbf{G}, \mathbf{F}_{pool}])) \quad (5)$$

式中:  $\mathbf{x}_{out}$  为模块最终输出特征图;  $Concat[\mathbf{G}, \mathbf{F}_{pool}]$  表示将边缘与空间特征图在通道维度上拼接;  $Conv2(\cdot)$  表示  $3 \times 3$  卷积, 步幅为 2, 进行空间降采样;  $Conv3(\cdot)$  表示  $1 \times 1$  卷积, 用于通道压缩与非线性变换。该设计可高效整合边缘与空间特征, 同时避免参数膨胀。

综上, EAS-Stem 模块在保证计算开销可控的前提下, 增强了模型对车辆外观异常细节的检测敏感性, 有效缓解了浅层表征不足导致的误检与漏检问题, 为整体检测性能提升奠定坚实基础。



Avg Pool: 平均池化层,  $K_1 \sim K_n$ :  $1 \sim C$  的卷积核, Rearrange: 张量维度调整, Sum: 元素求和

图 7 AWD-Conv 结构图

Fig. 7 AWD-Conv architecture diagram

(1) 权重生成路径(上分支): 空间注意力加权。首先对输入特征图  $\mathbf{x}$  进行平均池化处理以提取局部上下文信息:

$$\mathbf{x}_{avg} = AvgPool(\mathbf{x}) \quad (6)$$

随后通过  $1 \times 1$  卷积提取通道间交互信息并映射为注意力特征图:

$$\phi(\mathbf{x}) = Conv_{1 \times 1}(\mathbf{x}_{avg}) \quad (7)$$

该特征图在空间上被重新排列为  $\mathbf{R}^{C \times H \times W \times 4}$ , 即将每个  $2 \times 2$  区域展开为 4 个子区域通道维度。随后, 对该特征沿最后一维进行了 Softmax 归一化, 得到自适应注意力权重:

$$\tilde{\mathbf{A}}_{ijkq} = \frac{\exp(\phi(\mathbf{x})_{ijkq})}{\sum_{s=1}^4 \exp(\phi(\mathbf{x})_{ijks})} \quad (8)$$

式中:  $i, j, k$  分别表示通道、高度位置与宽度索引,  $q \in \{1, 2, 3, 4\}$  表示每个  $2 \times 2$  区域中第  $q$  个子区域的位置索引。该注意力矩阵将用于对下采样特征进行空间加权融合, 从而增强模型的空间感知能力。

(2) 特征提取路径(下分支): 分组卷积 + 重组融合。输入特征图  $\mathbf{x}$  同时输入下采样路径, 通过  $3 \times 3$  分组卷积生成四倍通道的特征图:

## 2) AWD-Conv 模块

标准卷积操作在图像空间中的每个感受野上共享相同的卷积核参数, 忽略了不同位置可能具有的信息差异, 从而在空间建模上存在一定的表达能力瓶颈。为提升特征表达的灵活性并保持模型的轻量化, 本文在 RFACConv<sup>[19]</sup> 思路的基础上提出了一种轻量化的自适应加权下采样卷积模块——AWD-Conv, 用于替代 YOLO11 部分标准卷积, 增强模型对空间位置变化与局部差异的建模能力。AWD-Conv 模块整体结构如图 7 所示, 主要由权重生成路径和特征提取路径两个分支构成:

$$\varphi(\mathbf{x}) = GroupConv_{3 \times 3}^{groups}(\mathbf{x}) \in \mathbf{R}^{4C \times H \times W} \quad (9)$$

随后将其重排为与注意力权重相匹配的维度  $\mathbf{R}^{C \times H \times W \times 4}$ , 与上分支输出逐元素相乘后, 在最后一维进行加权求和, 最终得到下采样后的输出特征图:

$$AWDConv(\mathbf{x}) = \sum_{s=1}^4 Softmax(\phi(\mathbf{x}))_s \odot \varphi(\mathbf{x})_s \quad (10)$$

式中:  $Softmax(\phi(\mathbf{x}))_s$  表示经过 Softmax 归一化后的注意力权重向量(针对子区域  $s$ ),  $\varphi(\mathbf{x})_s$  表示分组卷积后第  $s$  个区域的特征, “ $\odot$ ” 表示逐元素乘法。

该结构通过对每个  $2 \times 2$  空间区域的子区域进行动态加权融合, 能够自适应地选择保留信息重点, 在提升空间建模能力的同时实现高效的降采样操作。

## 2 实 验

### 2.1 实验环境与配置

本研究的所有模型训练与测试实验均在本地计算机上完成, 操作系统为 Windows 11, 处理器为 Intel(R) Core (TM) i5-12400F, 搭载 NVIDIA GeForce RTX 4060 Ti

16 GB 显卡。实验采用 Pytorch 2.2.2 深度学习框架, CUDA 版本为 12.1 进行模型构建与训练, 编程语言为 Python 3.10.14, 开发环境为 PyCharm。

## 2.2 数据集介绍

### 1) 事故图像数据集

本研究构建的事故图像数据集“LTAD”主要用于训练第 2 阶段的事故外观检测模型 YOLO-EA。数据集包含国内外图片网站、交通事故视频截图, 以及开源数据集中的部分图片, 涵盖了白天、夜间、黄昏、晴天、雨天等多种光照与天气条件, 以及城市快速路、高速公路、隧道、匝道等不同场景, 确保了数据在视觉外观和环境背景上的多样性与复杂性。所涵盖的事故类型包括车辆正面/侧面碰撞、追尾、翻车、撞护栏以及严重车体损坏等典型场景。

LTAD 原始图像共计 2 080 张, 其中事故图像(正样本)1 582 张, 非事故图像(负样本)498 张, 正负样本比例约为 3:1。在数据划分上, 严格遵循样本独立性原则, 保证来自同一来源、同一场景或同一视频序列的图像仅出现在一个子集中, 正负样本分别按 8:1:1 的比例划分为训练集、验证集和测试集, 并在各集合中保持比例一致性。此后, 对训练集图像施加亮度调整、水平翻转、添加高斯噪声等数据增强操作, 以提升模型的鲁棒性与泛化能力。验证集与测试集则保持原始图像数据, 此举旨在模拟模型在真实、未见过的数据上的性能, 确保评估结果的可靠性与无偏性。最终, 扩充后的数据集总量为 4 574 张图像, 并统一采用 YOLO 格式进行标注, 类别分为两类: “accident” 与 “non-accident”。数据示例如图 8 所示。

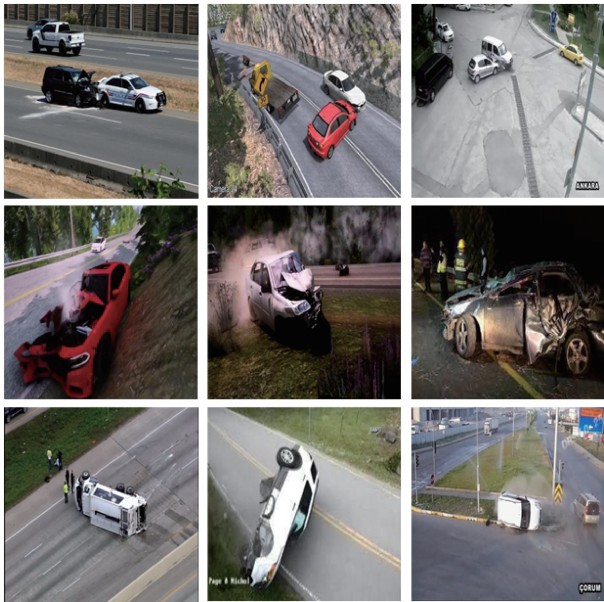


图 8 数据示例  
Fig. 8 Data sample

### 2) 事故视频数据集

为评估本文方法在实际应用场景中的有效性, 构建了

一个涵盖真实与仿真数据的事故视频数据集, 用于测试整体系统的检测性能。由于公开可用的交通事故视频资源有限, 本文除了使用 TAD<sup>[20]</sup> 数据集中的部分视频, 还采用了物理仿真平台 BeamNG.drive 生成高保真度的仿真交通事故视频。BeamNG.drive 是一款具备高度真实性的物理驾驶模拟器, 具备优异的车辆建模、碰撞还原与物理引擎支持能力。最终构建的事故视频数据集包含 90 段长度为 10~30 s 的视频片段, 其事故视频(正样本)45 段, 正常行驶视频(负样本)45 段, 正负样本比例为 1:1。所有视频片段在划分时均确保独立性, 即任意两段视频均取自不同的时间、地点或仿真场景, 从源头上避免了内容重叠, 保证评估结果的可靠性。该数据集被用于第 1 阶段静止车辆检测与第 2 阶段外观判别的联合实验, 全面验证所提方法在不同交通状态、光照环境与事故类型条件下的稳定性与检测精度。

## 2.3 评估指标

为全面评估所提出的交通事故检测方法的性能, 本文从两个层面进行评价: 一是对整个系统的综合检测效果; 二是对第 2 阶段外观检测模型(YOLO-EA)的性能进行评估。

### 1) 系统级评估指标:

识别率(Recognition rate): 表示被成功检测的静止车辆或事故车辆在所有实际静止/事故车辆中的比例。

$$Recognition\ rate = \frac{TP}{TP + FN} \quad (11)$$

漏检率(Miss detection rate): 表示实际存在的静止车辆或事故车辆中, 未被检测出的比例。

$$Miss\ detection\ rate = \frac{FN}{TP + FN} \quad (12)$$

误报率(False alarm rate): 表示系统错误地将非静止或非事故车辆判断为静止/事故的比例。

$$False\ alarm\ rate = \frac{FP}{TP + FP} \quad (13)$$

式中:  $TP$  表示检测为正且实际为正的样本数,  $FP$  表示检测为正但实际为负的样本数,  $FN$  表示检测为负但实际为正的样本数。

2) 针对第 2 阶段的事故外观检测模型 YOLO-EA, 采用目标检测领域常用的指标进行评价:

精准率(Precision,  $P$ ): 表示所有被预测为“事故”类别的样本中, 实际为事故的比例。

$$P = \frac{TP}{TP + FP} \quad (14)$$

召回率(Recall,  $R$ ): 表示所有实际为“事故”类别的样本中, 被正确预测为事故的比例。

$$R = \frac{TP}{TP + FN} \quad (15)$$

mAP@0.5: 在 IoU 阈值为 0.5 时, 分别对每个类别计算平均精度(AP), 并对所有类别的 AP 取平均, 得到的平均精度均值(mean average precision, mAP)。

mAP@0.5:0.95:在 IoU 阈值从 0.5 到 0.95(步长为 0.05)下分别计算 mAP 的平均值。评价模型在不同定位精度要求下的综合检测能力,是更严格的性能指标。

#### 2.4 双阶段事故检测系统

基于构建的事故视频数据集,并结合上述评估指标,对本文的双阶段事故检测系统进行了系统性实验。第 1 阶段实验结果如表 2 所示。

表 2 第 1 阶段实验结果

Table 2 First-stage experimental results

	TP	FN	FP	识别率/%	漏检率/%	误报率/%
结果	104	13	20	88.89	11.11	16.13

从表 2 可知,第 1 阶段的静止车辆识别率为 88.89%,系统能较为准确地筛选出具有潜在事故风险的静止目标,为第 2 阶段外观检测提供了可靠的候选对象,误报率为 16.13%,主要误报来自高速路监控视角远处缓行车辆,被系统误判为静止。

在第 2 阶段,系统对第 1 阶段筛选出的静止车辆启用 YOLO-EA 模型进行外观分析,并引入多帧投票机制以提升判定稳定性。为选取最优投票阈值参数  $Pos-K$ (即在  $N=10$  帧窗口中被判为事故的最少帧数),本文测试了  $Pos-K \in \{1, 3, 5, 7, 9\}$  的配置,最终的实验结果如表 3 所示。

从表 3 可知, $Pos-K$  设为 1 时虽能取得最高识别率(82.81%),但误报率也升至 29.33%,表明偶发帧的误判容易被放大;而当  $Pos-K$  设为 9 时,误报率虽降至最低

(23.88%),但漏检率明显上升(20.31%),影响整体判定的可靠性。综合考虑检测准确性与误判容忍度,最终选择  $Pos-K=7$  作为默认配置,在保持较高识别率(81.25%)的同时有效控制了误报率(24.63%),实现了精度与鲁棒性的平衡。

表 3 第 2 阶段实验结果

Table 3 Second-stage experimental results

	TP	FN	FP	识别率/ %	漏检率/ %	误报率/ %
$Pos-K=1$	53	11	22	82.81	17.19	29.33
$Pos-K=3$	52	12	19	81.25	18.75	26.76
$Pos-K=5$	52	12	18	81.25	18.75	25.71
$Pos-K=7$	52	12	17	81.25	18.75	24.63
$Pos-K=9$	51	13	16	79.69	20.31	23.88

#### 2.5 YOLO-EA 改进模块消融实验

为验证 EAS-Stem 和 AWD-Conv 模块在提升 YOLO-EA 性能方面的具体作用,本文在 LTAD 数据集上设计了系统的模块消融实验,用以评估各模块对整体性能的提升效果,实验结果如表 4 所示。实验共设置 4 组模型结构,具体说明包括:

M1:YOLO11s 原始模型,作为 Baseline;

M2:在 M1 基础上引入 EAS-Stem 模块;

M3:在 M1 基础上引入 AWD-Conv 模块;

M4:同时引入 EAS-Stem 与 AWD-Conv,即完整的 YOLO-EA 模型。

表 4 YOLO-EA 模块消融实验结果

Table 4 YOLO-EA module ablation study results

模块	Precision	Recall	mAP0.5	mAP0.5:0.95	GFLOPs	Parameters
M1	0.667	<b>0.737</b>	0.737	0.515	21.3	9 413 574
M2	0.761	0.696	0.756	0.538	21.8	9 416 230
M3	0.746	0.701	0.721	0.495	<b>19.8</b>	<b>7 451 334</b>
M4	<b>0.776</b>	0.691	<b>0.771</b>	<b>0.543</b>	20.2	7 453 990

从表 4 可以看出,两种改进模块在不同维度上均带来了性能的提升:

1)EAS-Stem 模块(M2):在不显著增加计算复杂度的前提下,实现明显的检测精度提升。与基准模型 M1 相比,mAP@0.5 提升 1.9%,Precision 提升 9.4%。该模块结构上引入基于 Sobel 算子的边缘感知分支与空间结构保持路径,能够在输入阶段突出图像中的细粒度变化与显著区域,有助于模型更有效地区分碰撞、裂痕等局部异常,从而提升整体检测准确率。

2)AWD-Conv 模块(M3):与 M1 相比,在参数量下降约 21%、GFLOPs 降低约 7%的同时,仍能保持相对稳定的

检测性能(mAP@0.5 仅下降 1.6%,Precision 提升 7.9%),表明其通过空间加权机制在下采样过程中保留了较强的判别信息,在结构压缩与表达能力之间实现了较好平衡。

3)YOLO-EA 模型(M4):结合两种改进模块后,精度与效率指标均获得进一步提升。相较于 Baseline 模型 M1,Precision 提升 10.9%、mAP@0.5 提升 3.4%、mAP@0.5:0.95 提升 2.8%;同时 GFLOPs 降低 5.2%,参数量下降约 21%。该结果充分验证了所提模块在兼顾检测精度与推理效率方面的有效性与协同性。

综合上述,EAS-Stem 通过增强输入阶段的边缘与空

间结构响应,改善了模型对异常区域的局部感知能力;AWD-Conv 则以较低计算代价实现空间注意力下采样,两者结合可实现对高速公路事故目标的精确检测与高效处理。

## 2.6 不同检测策略对比实验

为进一步验证所提出双阶段交通事故检测方法在降低误报率和提升实用性方面的优势,本文基于构建的事故视频数据集,并设计了两组实验与其对比:

1)单阶段方法:对每帧中所有检测到的车辆执行 YOLO-EA 模型进行事故判定,不进行静止筛选及多帧投票;

2)单帧判定方法:仅在车辆被判定为静止后,对其中一帧执行 YOLO-EA 模型进行事故判断,结果不经投票机制。

实验结果如表 5 所示。

表 5 对比试验结果

Table 5 Results of comparative experiment

方法	静止筛选	多帧投票	识别率/%	误报率/%
单阶段	×	×	87.50	49.09
单帧判定	✓	×	82.81	29.33
本文	✓	✓	81.25	24.63

从表 5 可知,单阶段方法虽具备最高识别率(87.50%),但误报率高达 49.09%,这表明在未经过静止筛选直接对所有车辆进行外观判断时,模型更容易受到背景干扰、遮挡等因素影响,误将非事故车辆识别为事故;单帧判定方法虽能一定程度缓解误报问题,但稳定性不足,仍存在较高误判风险。本文的两阶段方法通过静止筛选有效缩小检测范围,通过多帧投票机制抑制偶发误判,与单阶段相比,误报率降低了 24.46%,系统推理量大幅减少,提升推理效率;与单帧判定相比,误报率下降了 4.7%,检测结果更为稳定。综上,本文的方法在保证准确率的前提下,有效抑制误报并提升推理效率,具备更高的实用性与部署价值。

## 3 结 论

本文提出了一种基于深度学习的双阶段交通事故检测方法,分别从系统结构与模型精度优化两个层面进行了深入探索与改进。

在系统设计方面:构建了“静止车辆筛选+外观特征识别”的双阶段检测架构。第 1 阶段基于 YOLO11 与 Bot-SORT 对高速公路上的车辆检测与追踪,并根据车辆速度与静止时间判断其是否静止。该策略有效缩小检测范围,降低计算冗余,尤其适用于事故发生后车辆停滞的典型高速路场景。第 2 阶段则仅对静止车辆使用改进模型 YOLO-EA 进行事故识别,并引入多帧投票机制,在降低

误报率的同时提升了判定的稳定性与鲁棒性,显著减少了对 YOLO 模型的频繁调用,优化了整体推理效率。

在模型结构方面:对 YOLO11 网络进行结构优化,构建了 YOLO-EA 外观检测模型。该模型引入了两种模块:一是 EAS-STEM 模块,通过边缘增强和空间建模提升模型对翻车、碰撞等外观细节的感知能力;二是 AWD-Conv 模块,在显著减少参数量与计算复杂度的同时保持了特征提取能力。实验表明,这两个模块对提升检测性能具有明显贡献,最终构建的 YOLO-EA 模型在 mAP@0.5 与 mAP@0.5:0.95 等关键指标上均优于基线 YOLO11,验证了改进策略的有效性。

需要指出的是,本文的方法仍存在一定局限性:其有效性主要依赖于事故导致车辆停滞的典型场景,在光照剧烈变化或恶劣天气条件下,第 1 阶段的检测与跟踪性能以及第 2 阶段的外观特征提取精度均可能受到影响;当前系统为纯视觉方案,对无明显损伤的事故判别能力有限,需结合运动时序特征(如加速度突变、轨迹异常)或引入多源信息加以补偿;此外,训练数据的规模与采集环境的多样性仍不足,限制了模型在复杂场景下的泛化能力。

综合来看,本文提出的双阶段检测策略不仅有效降低了单阶段检测中普遍存在的误报率高、计算开销大等问题,还提升了系统的稳定性与实时性,具备较高的部署价值。未来研究可从以下方向拓展:一是引入时序建模机制,融合车辆运动轨迹信息,进一步提升事故过程识别能力;二是结合多源异构数据(如雷达、地图或语义分割信息)与行为语义线索(如分心驾驶识别等)<sup>[21]</sup>,增强系统对复杂环境下的泛化与鲁棒性,推动交通事故检测向更智能、稳定、实时的方向发展。

## 参考文献

- [1] REZAEI Z, EBRAHIMPOUR-KOMLEH H. Prediction and detection of car accidents in video by deep learning[C]. 2021 5th International Conference on Pattern Recognition and Image Analysis (IPRIA), 2021: 1-9.
- [2] 柳圣,丁奕,王心莹,等.基于 BNN 的交通事故预警系统[J].信息与电脑(理论版),2023,35(20):116-119.  
LIU SH, DING Y, WANG X Y, et al. Traffic accident early warning system based on BNN[J]. Information and Computer (Theoretical Edition), 2023, 35(20): 116-119.
- [3] ARIFEEN Z U, HONG J E, SEO B S, et al. Traffic accident detection and classification in videos based on deep network features [C]. 2023 Fourteenth International Conference on Ubiquitous and Future Networks(ICUFN), 2023: 491-493.
- [4] CHEN J, JIN Y, CHEN X, et al. Surveillance-based road accident detection method using improved

- YOLOv8 for IIoT applications [C]. 2024 IEEE/CIC International Conference on Communications in China (ICCC), 2024: 1413-1418.
- [5] ESSAM M, GHANEM N M, ISMAIL M A. Detection of road traffic crashes based on collision estimation[J]. ArXiv preprint arXiv:2207.12886, 2022.
- [6] GHAREMANNEZHAD H, SHI H, LIU C. Real-time accident detection in traffic surveillance using deep learning[C]. 2022 IEEE International Conference on Imaging Systems and Techniques (IST), 2022: 1-6.
- [7] DHARMADASA K A D D, SAHOO G K, DAS S K, et al. Video-based road accident detection on highways: A less complex YOLOv5 approach [C]. 2023 International Conference on Computer, Electronics & Electrical Engineering & their Applications(IC2E3), 2023: 1-6.
- [8] NALLASIVAN G. A vision-based traffic accident analysis and tracking system from traffic surveillance video [C]. 2024 Third International Conference on Intelligent Techniques in Control, Optimization and Signal Processing(INCOS), 2024: 1-6.
- [9] TEE F S, THENG L B, SHIONG B L C, et al. A multi-stream approach to mixed-traffic accident recognition using deep learning [J]. IEEE Access, 2024, 12: 185232-185249.
- [10] ZHANG Y, SUNG Y. Hybrid traffic accident classification models[J]. Mathematics, 2023, 11(4): 1050.
- [11] 王晨, 周威, 严隽逸, 等. 一种用于道路交通事故自动检测的改进双流网络[J]. 中国公路学报, 2023, 36(5): 185-196.
- WANG CH, ZHOU W, YAN J Y, et al. An improved two-stream network for automatic detection of road traffic accidents[J]. China Journal of Highway and Transport, 2023, 36(5): 185-196.
- [12] KHANAM R, HUSSAIN M. YOLOv11: An overview of the key architectural enhancements [J]. ArXiv preprint arXiv:2410.17725, 2024.
- [13] AHARON N, ORFAIG R, BOBROVSKY B Z. BoT-SORT: Robust associations multi-pedestrian tracking[J]. ArXiv preprint arXiv:2206.14651, 2022.
- [14] 刘奥奇, 杨成佳. 改进 YOLOv8n 的车辆检测算法研究[J]. 吉林建筑大学学报, 2024, 41(5): 83-88.
- LIU AO Q, YANG CH J. Research on an improved YOLOv8n vehicle-detection algorithm [J]. Journal of Jilin Jianzhu University, 2024, 41(5): 83-88.
- [15] 霍爱清, 郭岚洁, 冯若水. 面向密集场景的多目标车辆检测算法[J]. 电子测量技术, 2024, 47(9): 129-136.
- HUO A Q, GUO L J, FENG R S. Multi-object vehicle detection algorithm for dense scenes [J]. Electronic Measurement Technology, 2024, 47(9): 129-136.
- [16] ZHANG Y, SUN P, JIANG Y, et al. Bytetrack: Multi-object tracking by associating every detection box [C]. European Conference on Computer Vision (ECCV). Cham: Springer Nature Switzerland, 2022: 1-21.
- [17] 占丽琴. 基于 YOLOv8 的车流量检测平台设计与实现[J]. 自动化应用, 2025, 66(10): 46-49.
- ZHAN L Q. Design and implementation of traffic flow detection platform based on YOLOv8 [J]. Automation Application, 2025, 66(10): 46-49.
- [18] 卫策, 吕进, 曲晨阳. 改进 YOLOv5s 的复杂交通场景下目标检测算法 [J]. 电子测量技术, 2024, 47(2): 121-130.
- WEI C, LYU J, QU CH Y. An improved YOLOv5s object detection algorithm for complex traffic scenes [J]. Electronic Measurement Technology, 2024, 47(2): 121-130.
- [19] ZHANG X, LIU C, YANG D, et al. RFACConv: Innovating spatial attention and standard convolutional operation [J]. ArXiv preprint arXiv:2304.03198, 2023.
- [20] XU Y, HU H, HUANG C, et al. TAD: A large-scale benchmark for traffic accidents detection from video surveillance [J]. IEEE Access, 2024, 13: 2018-2033.
- [21] 沈骞, 张磊, 张宇翔, 等. 基于改进 YOLOv8n 的轻量化分心驾驶行为检测方法 [J]. 电子测量技术, 2024, 47(24): 65-75.
- SHEN Q, ZHANG L, ZHANG Y X, et al. Lightweight distracted-driving detection method based on improved YOLOv8n [J]. Electronic Measurement Technology, 2024, 47(24): 65-75.

## 作者简介

**凌锐**, 硕士研究生, 主要研究方向为图像处理、信号处理。  
E-mail: 1194984426@qq.com

**闫坤**(通信作者), 博士, 教授, 主要研究方向为自适应信号处理、机器学习、图像处理等。  
E-mail: kyan5702@gmail.com

**梁宏宇**, 硕士研究生, 主要研究方向为图像处理、信号处理。  
E-mail: 1098246384@qq.com

**韦焯淇**, 硕士研究生, 主要研究方向为图像处理、信号处理。  
E-mail: 2521711281@qq.com

**郝航勃**, 硕士研究生, 主要研究方向为图像处理、信号处理。  
E-mail: 1628759348@qq.com