

基于ARV-YOLO的交通标志检测算法

闫家瑞, 杨宗才, 李星星*

五邑大学电子与信息工程学院, 广东 江门

收稿日期: 2026年3月15日; 录用日期: 2026年5月6日; 发布日期: 2026年5月15日

摘要

本文针对复杂道路场景中交通标志目标尺度小、尺度变化大以及背景干扰强等问题, 提出一种基于改进YOLOv8n的交通标志检测方法。首先, 在主干网络中引入RFCACnv模块, 通过结合感受野建模与通道注意力机制增强网络对多尺度特征的表达能力; 其次, 借鉴YOLOv7的下采样, 以多分支结构在降低特征分辨率的同时尽可能保留空间细节信息减小背景干扰; 最后, 构建ASF-P2多尺度特征融合模块, 充分利用浅层高分辨率特征, 提高模型对小尺度交通标志的检测能力。实验结果表明, ARV-YOLO模型在TT100K数据集上的mAP@0.5达到80.8%, 较原YOLOv8n提升9.9%。同时, 模型检测精度与召回率分别提升9.3%和7.1%, 有效提升了交通标志检测中多尺度标志和小目标检测性能, 解决了背景干扰强问题, 具有较好的鲁棒性和应用价值。

关键词

交通标志检测, YOLOv8n, 小目标, 多尺度

Traffic Sign Detection Algorithm Based on ARV-YOLO

Jiarui Yan, Zongcai Yang, Xingxing Li*

School of Electronic and Information Engineering, Wuyi University, Jiangmen Guangdong

Received: March 15, 2026; accepted: May 6, 2026; published: May 15, 2026

Abstract

This paper addresses the challenges of small-scale traffic signs, large scale variations, and strong background interference in complex road scenarios, proposing a traffic sign detection method based on an improved YOLOv8n. First, an RFCACnv module is introduced into the backbone network

*通讯作者。

文章引用: 闫家瑞, 杨宗才, 李星星. 基于 ARV-YOLO 的交通标志检测算法[J]. 人工智能与机器人研究, 2026, 15(3): 776-787. DOI: 10.12677/airr.2026.153073

to enhance the network's ability to represent multi-scale features by combining receptive field modeling and channel attention mechanisms. Second, borrowing downsampling from YOLOv7, a multi-branch structure is used to reduce feature resolution while preserving spatial details and minimizing background interference. Finally, an ASF-P2 multi-scale feature fusion module is constructed to fully utilize shallow high-resolution features, improving the model's ability to detect small-scale traffic signs. Experimental results show that the ARV-YOLO model achieves an mAP@0.5 of 80.8% on the TT100K dataset, a 9.9% improvement over the original YOLOv8n. Simultaneously, the model's detection accuracy and recall are improved by 9.3% and 7.1%, respectively, effectively enhancing the detection performance of multi-scale signs and small targets in traffic sign detection, solving the problem of strong background interference, and demonstrating good robustness and application value.

Keywords

Traffic Sign Detection, YOLOv8n, Small Targets, Multi-Scale

Copyright © 2026 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

随着智能交通系统(Intelligent Transportation Systems, ITS)和自动驾驶技术不断发展成熟,车辆对道路环境信息获取的精准度要求越来越高。交通标志作为道路上传递关键信息的基础设施——比如限速提示、危险预警、方向指引等——它能否被准确识别出来,直接影响着车辆行驶过程中的安全保障和驾驶决策的可靠性。当前,打造一套能精准识别、快速响应且不易受外界干扰的交通标志检测算法,正成为计算机视觉与智能交通领域里大家重点研究的方向之一。

在实际道路场景中开展交通标志检测工作时,还存在不少需要解决的难题。交通标志本身的特点就增加了检测难度——比如尺寸普遍偏小、大小差异大、不同类型的标志数量不均衡,尤其是车辆离得远或者标志密集排列时,小尺寸标志很容易被周围背景覆盖;同时道路上还会遇到光照忽明忽暗、标志被其他物体遮挡、车辆行驶带来的图像模糊以及背景环境复杂等情况,这些因素都让检测工作变得更复杂。这些问题都需要检测模型在提取特征信息、识别不同大小目标以及实时处理数据等方面具备更强能力。

近年来,基于深度学习的目标检测方法在交通标志检测任务中取得了显著进展[1]。两阶段检测算法如 Faster R-CNN 在检测精度方面表现较好,但其模型结构复杂、计算量大,难以满足车载系统对实时性的需求。相比之下,以 YOLO (You Only Look Once)系列为代表的单阶段检测算法在速度与精度之间取得了较好的平衡。其中, YOLOv8n 具有参数量少、计算复杂度低、推理速度快等优势,尤其适合部署在算力受限的车载终端或边缘设备上。然而,轻量化模型在复杂交通标志场景中仍存在特征表达能力不足、对小目标和多尺度目标检测精度有限等问题,直接将 YOLOv8n 应用于交通标志检测任务,往往难以同时兼顾检测精度与鲁棒性[2]。

针对上述问题,国内外学者围绕 YOLOv8 架构在交通标志检测领域开展了大量研究工作,例如, Zhang 等人提出的 TS-YOLO 模型在 YOLOv8n 框架中嵌入轻量级通道注意力模块,并结合改进的损失函数,在雨雾等恶劣天气条件下有效抑制了背景干扰,提升了小尺度交通标志的召回率[3]。类似地, Wang 等人设计了一种融合注意力尺度序列与 P2 小目标特征层的 Neck 结构,通过增强浅层高分辨率特征的语义信息,显著改善了对远距离微小标志的定位能力[4]。为应对多尺度变化与背景杂乱的双重挑战, Zhao

等人构建了 CPB-YOLOv8 框架,其核心创新在于引入跨阶段部分划分的 Transformer 块(CSP-PTB),在保留局部细节的同时捕获长距离依赖关系,从而在 TT100K 和 GTSDB 数据集上实现了优于基线模型的 mAP 表现[5]。上述研究在特征提取和特征融合方面取得一定进展,但仍存在复杂场景交通标志检测中对多尺度特征和小目标特征捕捉不充分,适应性差的问题。

基于此,本文提出交通标志检测模型 ARV-YOLO,包括如下三个改进策略:

(1) 在主干网络中引入 RFCACnv 模块,通过感受野建模与通道注意力机制增强特征表达能力,提高多尺度检测能力。

(2) 构建 ASF-P2 多尺度特征融合模块并引入 P2 检测头,强化浅层特征利用,提升小目标交通标志检测性能。

(3) 引入 V7DS 下采样模块,通过多分支结构减少下采样过程中的复杂背景中交通标志信息损失。

2. ARV-YOLO 交通标志检测模型

ARV-YOLO 选取 YOLOv8n 作为基础网络框架。YOLOv8n 本身具备多尺度特征融合、无锚点检测模式以及高效的特征提取能力——这些特性让它在维持检测精度的同时,还能兼顾实时响应速度与轻量化优势,因此很适合交通标志检测这类对实时性要求较高的应用场景。不过在实际处理交通标志识别任务时,原始 YOLOv8n 仍存在一些不足:比如很多交通标志在图像中占比很小、标志尺度变化范围大、背景环境复杂等情况,都可能影响它的检测效果。

针对这些问题,本文设计了一种改进后的 ARV-YOLO 交通标志检测算法。在主干网络部分引入 RFCACnv 模块,该模块通过结合感受野特征建模与通道注意力机制,强化网络对关键区域特征的捕捉能力,让模型能更好地感知不同尺度的交通标志,从而更充分地提取标志的细粒度特征信息。在特征融合阶段,构建 ASF-P2 多尺度特征融合结构,并新增 P2 检测层。这样一来,浅层的高分辨率特征就能更深入地参与到目标检测过程中,有助于提升模型对小尺度交通标志的检测表现。另外,为了缓解传统下采样过程中空间信息丢失的问题,在网络结构里加入 v7DS 下采样模块——该模块采用最大池化分支与卷积分支并行处理的方式,在降低特征图分辨率的同时保留更多空间细节,进一步增强网络的特征表达能力。ARV-YOLO 在保持模型轻量化和实时性优势的前提下,有效提升了对小尺寸、多尺度交通标志的检测能力,以及在复杂交通场景下的识别精度。改进后的 ARV-YOLO 网络结构可参考图 1。

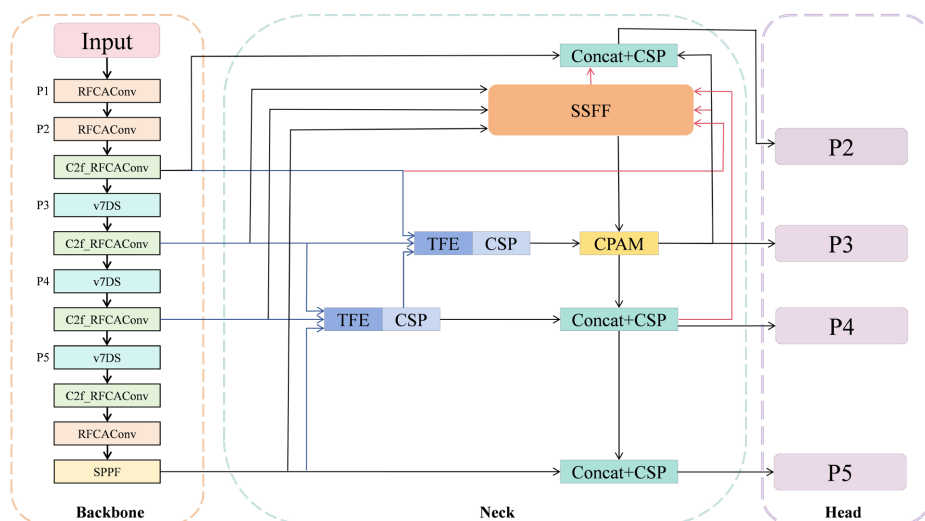


Figure 1. ARV-YOLO network architecture
图 1. ARV-YOLO 网络结构

2.1. ASF-P2 模块

YOLOv8 在处理骨干网络输出的 P3、P4、P5 三层特征图时，只是将这三层直接送入 FPN + PANet 结构，仅通过简单的上采样和拼接操作完成多尺度融合。这种方式没有深入挖掘不同尺度特征之间的关联，当交通标志因拍摄距离不同出现明显的大小变化时，模型很难做出有效应对。同时，它缺乏针对小目标设计的注意力机制，在树木遮挡、建筑物干扰等复杂交通背景下，难以聚焦到关键的标志区域。相比之下，ASF-YOLO 首先提出的 Triple Feature Encoder (TFE)模块[6]能同时处理大、中、小三种尺度的特征图，并将它们在空间维度上拼接起来，这样可以更好地保留浅层特征图里包含的细节信息，这对检测远处的小交通标志来说十分关键。其次是 Scale Sequence Feature Fusion (SSFF)模块，该模块采用 3D 卷积对多尺度特征序列进行融合，能够更精准地捕捉不同尺度特征之间的语义联系，让模型在面对交通标志大小变化时表现得更稳定。第三是 Channel and Position Attention Mechanism (CPAM)模块，它通过通道注意力筛选出关键的特征通道，再借助位置注意力精确定位空间坐标，帮助模型在复杂交通场景中同时关注标志的类别特征和具体位置。最后，ASF-YOLO 还引入了 P2 检测头，接收来自主干网络浅层的高分辨率特征，专门负责检测远距离的交通标志。

2.1.1. SSFF 模块

SSFF 模块先接收主干网络输出的 P3、P4 和 P5 三级特征图，用 1×1 卷积把 P4 和 P5 的通道数都调成 256 维，和 P3 保持一致。之后给每一层特征加上标准差 σ 越来越大的 2D 高斯滤波，做出尺度空间表示——高斯核的平滑作用能生成不同模糊程度的特征序列，这样就能捕捉不同大小的目标特征了。接着用最近邻插值法把深层的 P4 和 P5 上采样到 P3 的高分辨率尺寸(80×80)，让它们在空间上对齐。再通过 `unsqueeze` 操作把各层的 3D 特征张量[H, W, C]变成 4D 张量[depth, H, W, C]，然后沿着深度维度拼起来，形成包含多尺度信息的 4D 特征堆叠体。最后把这个张量放进 3D 卷积层，让卷积核在深度维度上滑动，这样就能直接建模跨尺度序列之间的联系。经过 3D 批归一化和 SiLU 激活函数处理后，输出融合了深层语义信息和浅层细节特征的统一表示。这种方式能有效应对交通标志因拍摄距离不同而产生的大幅尺度差异问题。SSFF 模块结构如图 2 所示。

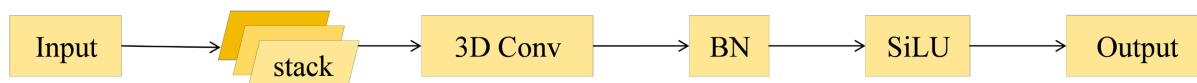


Figure 2. SSFF module structure

图 2. SSFF 模块结构

2.1.2. TFE 模块

TFE (Triple Feature Encoder)模块的输入来源是主干网络输出的三种不同尺度特征图——大、中、小尺度特征图。其中大尺度特征图包含较多空间细节，但语义信息不够深入；小尺度特征图能提供深层语义内容，但空间分辨率较低；中尺度特征图则处于两者之间，同时具备一定的空间细节和语义层次。针对大尺度分支，模块先借助 ConvBNSiLU 卷积模块将其通道数调整到目标维度 C。之后采用最大池化与平均池化结合的混合下采样结构来降低空间维度，这种结构可以增强网络的平移不变性，让网络对输入图像的空间变化有更强的鲁棒性编码能力。小尺度分支同样先通过 ConvBNSiLU 将通道数调整至 C，不过采用最近邻插值法进行上采样以恢复空间分辨率。最近邻插值法通过填充邻域像素来补充特征图，能有效保留小目标的局部细节，避免上采样过程中出现信息损失。中尺度分支直接利用 ConvBNSiLU 完成通道调整与特征提取，保持原有的空间尺度不变。当三个分支都达到统一的空间尺寸(1S)和通道维度(1C)后，模块会对每个分支的输出分别进行一次卷积操作。最后在通道维度上将大、中、小三种特征图拼接

(Concat)融合，形成同时包含丰富细节信息、多尺度上下文以及深层语义的综合特征表示(3C × 1S)，为后续的小目标检测提供全面的特征支持。TFE 模块的具体结构可参考图 3。

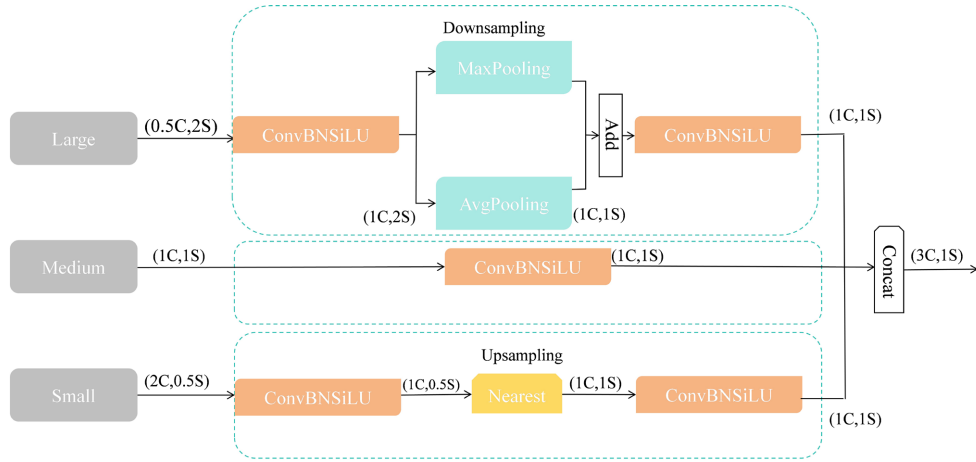


Figure 3. TFE module structure
图 3. TFE 模块结构

2.2. RFCACConv 模块

交通标志通常具有多尺度的特征，RFCACConv 通过将坐标注意力机制(Coordinate Attention)与感受野空间特征相结合更精确地定位多尺度目标。RFCACConv 的完整计算流程可表示为：

$$F_{rf} = \text{ReLU}(\text{Norm}(g^{k \times k}(X))) \quad (1)$$

$$A_{rf} = \sigma(f^h) \times \sigma(f^w) \quad (2)$$

$$F = A_{rf} \odot F_{rf} \quad (3)$$

式中 X 表示输入特征图， $g^{k \times k}$ 代表核大小为 $k \times k$ 的分组卷积(Group Convolution)，Norm 表示批量归一化(Batch Normalization)操作，ReLU 为修正线性单元激活函数，而 F_{rf} 则表示经过变换后得到的感受野空间特征(Receptive-Field Spatial Feature)，即由不重叠滑窗组成的特征图。第二个公式用于计算感受野注意力权重，其中 f^h 和 f^w 分别表示通过对感受野空间特征沿高度和宽度方向进行平均池化(或全局平均池化)得到的高度编码和宽度编码， σ 代表 Sigmoid 激活函数，用于将注意力值映射至(0, 1)区间，符号 \times 表示逐元素相乘，最终得到的 A_{rf} 即为感受野注意力图，用于表征感受野内各特征的重要性权重。第三个公式表示最终的加权操作，其中符号 \odot 代表 Hadamard 积，通过将注意力权重 A_{rf} 与感受野空间特征 F_{rf} 进行逐元素相乘，得到最终的输出特征图 F 。

输入特征图通过分组卷积(Group Conv)进行感受野空间特征提取，将原始空间特征转换为感受野空间特征(Receptive-Field Spatial Feature) [7]，此时特征图的尺寸变为 $C \times KH \times KW$ ，其中 K 为卷积核大小；随后，该特征图被送入坐标注意力分支，分别沿高度和宽度方向进行平均池化(H AvgPool 和 W AvgPool)，生成两个方向的特征描述符，这两个描述符经过拼接、卷积、归一化和非线性激活后，再分别通过卷积和 Sigmoid 函数生成高度方向和宽度方向的注意力权重；与此同时，输入特征图还通过全局平均池化(Global AvgPool)提取全局信息，经过线性层和 ReLU 激活后与感受野特征进行融合；最终，将坐标注意力生成的权重与感受野空间特征进行重加权(Re-weight)操作，再通过一个步长为 K 的 $K \times K$ 卷积调整形状并输出结果。RFCACConv 模块结构如图 4 所示。

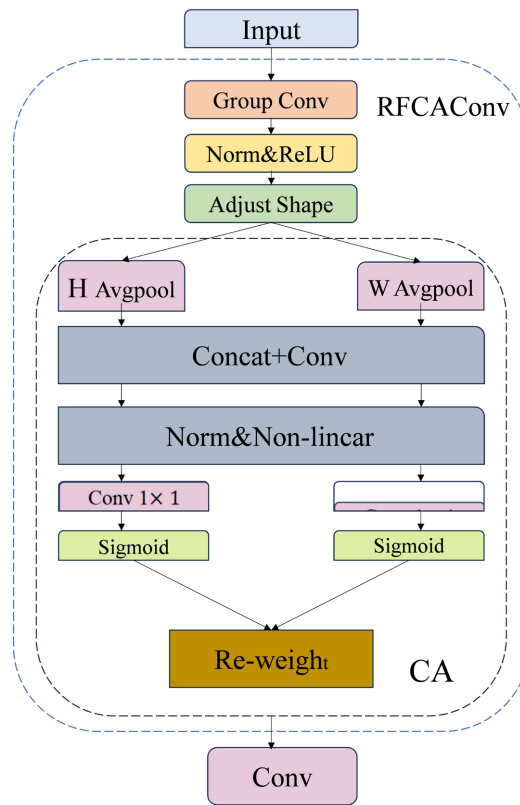


Figure 4. RFCACConv module structure
图 4. RFCACConv 模块结构

2.3. v7DS 模块

实际道路场景中常遇到光照变化、物体遮挡、图像模糊等复杂情况。YOLOv8 模型里的下采样模块主要依靠卷积操作来完成，这种方式虽然能较好地学习图像特征，但经过多次下采样后，很容易把目标的细节信息弄丢，进而影响交通标志检测结果。v7DS 模块能在一定程度上解决这个问题。它主要通过同时使用两种不同方式——最大池化(Max Pooling)和步长为 2 的卷积(Conv with stride = 2)，来实现空间维度压缩和特征通道调整[8]。一个分支采用 3×3 的最大池化操作，步长设为 2，用于保留显著的特征响应并降低特征图尺寸；另一个分支则使用 3×3 卷积，同样以步长为 2 进行下采样，用于学习可训练的特征变换。两个分支的输出在通道维度上进行拼接(Concatenate)，从而在不丢失过多信息的前提下实现高效的下采样。v7DS 模块结构如图 5 所示。

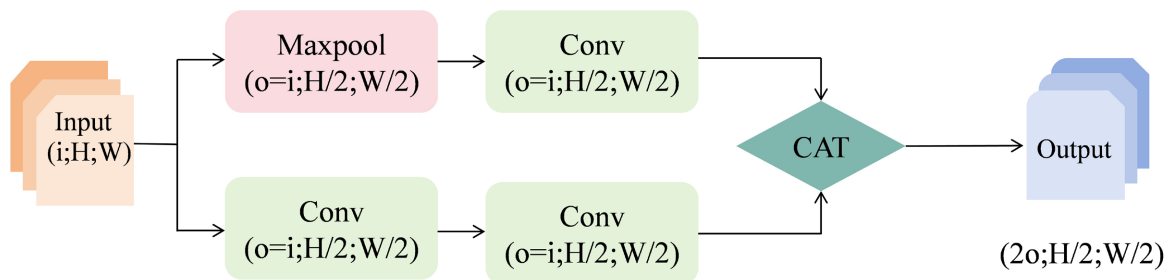


Figure 5. v7DS module structure
图 5. v7DS 模块结构

3. 实验结果及分析

3.1. 实验环境

实验平台使用 NVIDIA GeForce RTX 4060Ti 的 GPU，操作系统为 Windows。深度学习框架为 Pytorch 2.2.2，Python 3.9，Cuda 12.1。实验环境参数设置 Epoch = 300，Batch = 8，optimizer = “SGD”。

3.2. 数据集介绍

TT100K (Tsinghua-Tencent 100K)数据集由清华大学和腾讯公司共同推出，是一个规模较大的交通标志检测数据集，主要用于支持智能交通和自动驾驶方向的研究工作。这个数据集里的街景图像来自国内多个城市的道路场景，涵盖了不同天气状况、光照条件、交通环境和道路类型，复杂程度和真实感都比较高，在目前公开的交通标志检测相关数据集中算是很有代表性的一个。

TT100K 数据集大概有 10 万张交通场景图像，但因为里面包含的标志类别很多，而且各类别的数量分布差别很大，所以文章根据实验的实际需求挑选了部分常见类别来做训练和验证——比如只保留出现次数较多的 45 类，训练用了 6793 张、验证用 1949 张、测试用 996 张。这样的划分方式一方面能让模型有足够多的训练样本，另一方面也可以比较客观地衡量模型在没见过的样本上能不能准确检测。交通标志的具体例子可以看文中的图 6。



Figure 6. Examples of traffic signs

图 6. 交通标志示例

3.3. 评价指标

交通标志检测任务中，通常从计算量(GFLOPS)、参数量(Parameters)、精确率(Precision)、召回率(Recall)、平均精度均值(mAP)来评价。这些指标能够从不同角度反映检测模型在复杂交通环境下识别交通标志的能力。各指标计算公式如下：

$$P = \frac{TP}{TP + FP} \quad (4)$$

$$R = \frac{TP}{TP + FN} \tag{5}$$

$$AP = \int_0^1 P(R) dR \tag{6}$$

$$mAP@0.5 = \frac{1}{N} \sum_{i=1}^N AP_i|_{IoU=0.5} \tag{7}$$

式中 TP 表示真正例(正确检测到的交通标志数量), FP 表示假正例(误检为背景或错误类别的数量), FN 表示假反例(漏检的交通标志数量), $P(R)$ 为召回率 - 精确率曲线(PR 曲线), N 为交通标志类别总数(本文中 $N = 45$), AP_i 为第 i 类的平均精度。在交并比(IoU)阈值为 0.5 时计算得到的 mAP。

3.4. 消融实验

为评估每个模块对模型的贡献, 在 TT100K 数据集上进行了消融实验, 对比基线模型 YOLOv8n。在基线模型的基础进行消融实验。

Table 1. Ablation experiment results data

表 1. 消融实验结果数据

Model	ASF-P2	RFCACConv	V7DS	Parameters	GFLOPS	P	R	MAP@50	MAP@50-95
YOLOV8N	×	×	×	3.0	8.1	0.722	0.632	0.709	0.54
A	√	×	×	2.6	13.0	0.714	0.719	0.773	0.589
R	×	√	×	3.1	8.9	0.767	0.66	0.751	0.576
V	×	×	√	2.7	7.9	0.715	0.661	0.727	0.543
AV	√	×	√	2.4	12.9	0.741	0.728	0.782	0.597
ARV	√	√	√	2.5	13.7	0.815	0.703	0.808	0.623

为确认 ARV-YOLO 模型里各个改进模块的实际作用, 研究团队在 TT100K 数据集上开展了一系列消融实验, 相关结果整理在表 1 中。作为基础参照的 YOLOv8n 模型, 其 mAP@0.5 指标为 70.9%, 模型参数数量是 3.0 M, 计算量达到 8.1 GFLOPS。

单独引入 ASF-P2 模块的模型(记为模型 A)参数量减少了 0.4 M, 精确率下降了 0.8 个百分点, 但召回率提高了 8.7 个百分点, mAP@0.5 由 70.9%提升到 77.3%, 提高了 6.4 个百分点, mAP@0.5:0.95 也由 54.0%提升到 58.9%。这说明 ASF-P2 模块能够有效加强浅层高分辨率特征在检测过程中的利用效果, 使模型对远距离、小尺寸交通标志更加敏感, 从而提升目标检出能力和整体检测精度。虽然该模块在精确率上略有下降, 说明随着更多目标被检出, 误检情况也有所增加, 但从召回率和 mAP 指标的提升可以看出, ASF-P2 在改善小目标漏检问题方面表现更为突出, 对交通标志检测任务具有较明显的促进作用。单独引入 RFCACConv 模块的模型(记为模型 R)参数只增加了 0.1 M, 计算量多了 0.8 GFLOPS。这个模块让精确率提升 4.5 个百分点——从 72.2%到 76.7%; mAP@0.5 提高 4.2 个百分点, 达到 75.1%。这 RFCACConv 通过引入更合理的感受野建模方式, 并结合坐标注意力机制, 对关键区域和有效通道进行强化, 使网络能够更加关注交通标志本身具有辨识性的边缘、轮廓和局部结构信息, 削弱复杂背景中的冗余响应。主要效果体现在抑制误检、提高分类与定位结果的可靠性上, 这会直接反映为精确率的提升更加明显。换言之, RFCACConv 并不是单纯“检测出更多目标”, 而是让模型“更准确地区分什么是真正的交通标志”, 因此其对 Precision 的贡献通常强于对 Recall 的贡献。但召回率的提升幅度不大——只有 2.8 个百分点, 说明单靠这一个特征增强模块来提升小目标召回能力是有局限的。只加入 v7DS 下采样模块的模型(记为

模型 V)参数降到 2.7 M, 计算量降到 7.9 GFLOPS, 是所有测试模型里最低的。这个模块让 mAP@0.5 提升了 1.8 个百分点, 达到 72.7%。虽然提升幅度不算大, 但 v7DS 模块在降低计算复杂度的时候还保留了更多空间细节信息, 这证明多分支下采样的设计思路在轻量化模型开发中是有用的。

接下来看模块组合的效果。模型 AV (同时用 ASF-P2 和 v7DS 模块)的 mAP@0.5 达到 78.2%, 比基础模型高 7.3 个百分点。这说明把浅层特征融合和多分支下采样结合起来, 可以在检测精度和计算效率之间找到较好的平衡。最终整合了三个模块(ASF-P2 + RFCACConv + v7DS)的完整模型 ARV 表现最好: mAP@0.5 达到 80.8%, 比 YOLOv8n 高 9.9 个百分点; 精确率和召回率分别是 81.5%和 70.3%——分别提升 9.3 和 7.1 个百分点; mAP@0.5:0.95 达到 62.3%, 提升 8.3 个百分点。值得注意的是 ARV 模型的参数数量(2.5 M)和计算量(13.7 GFLOPS)还是保持在较低水平, 做到了精度和效率的平衡。

各个改进模块从不同方向提升了模型性能——ASF-P2 模块强化了小目标召回能力、RFCACConv 模块增强了多尺度特征判别能力、v7DS 模块优化了空间信息保留能力。把这三个模块结合起来后, ARV-YOLO 模型能在复杂交通场景里更准确地识别小尺度和多尺度的交通标志。研究团队还绘制了基础模型和改进后模型的可视化对比图, 直观展示 ARV-YOLO 在交通标志检测上的提升效果。像图 7 里展示的那样, 基础模型识别的准确性和有效性较低, 在密集的交通背景里很难准确定位小尺寸标志; 而改进后的模型借助 ASF-P2 模块, 能识别更小尺寸以及准确定位不同尺度交通标志。

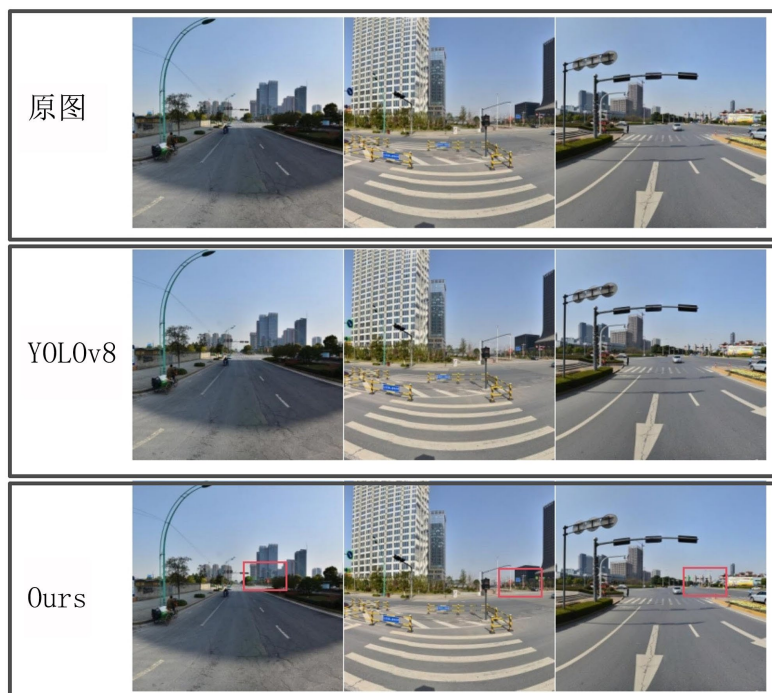


Figure 7. Visual comparison: before and after improvements
图 7. 改进前后可视化对比图

3.5. 对比实验

为进一步验证本文所提 ARV-YOLO 模型的先进性与泛化能力, 将其与当前主流的目标检测算法进行横向对比实验, 包括两阶段检测器 Cascade R-CNN、Faster R-CNN [9]以及单阶段轻量化检测器 RTMDet [10]、YOLOX [11]。实验在 tt100k 数据集上开展, 评价指标涵盖参数量、计算量、精确率、召回率及平均精度均值。

Table 2. Comparison of experimental results data
表 2. 对比实验结果数据

Model	Parameters	GFLOPS	<i>P</i>	<i>R</i>	MAP@50	MAP@50-95
Cascade-rcnn	69.828	163	0.750	0.616	0.721	0.559
Faster-rcnn	41.574	134	0.730	0.523	0.630	0.498
Rtmdet	4.866	8.061	0.841	0.665	0.786	0.606
Yolox	5.046	7.606	0.670	0.736	0.746	0.549
Yolov10n	2.3	6.6	0.671	0.63	0.682	0.515
Ours	2.5	13.7	0.815	0.703	0.808	0.623

从表 2 可以看出，相较于 Faster R-CNN、Cascade R-CNN 这类传统两阶段检测模型，ARV-YOLO 在模型轻量化上有很大优势。它的参数量只有 2.5 M，计算复杂度是 13.7 GFLOPS，和 Cascade R-CNN 比起来，这两个数值分别降低了 96.4%和 91.6%。在检测精度方面，ARV-YOLO 的 mAP@0.5 达到了 80.8%，比 Faster R-CNN 高 17.8 个百分点，比 Cascade R-CNN 高 8.7 个百分点。这些数据说明这个模型能大幅减少计算资源消耗，同时还能在复杂环境里准确检测交通标志。和 RTMDet、YOLOX 这些单阶段轻量化检测器比起来，ARV-YOLO 的参数量分别降低了 48.6%和 50.5%，压缩效果更好。RTMDet 的精确率虽然稍微高一点(84.1%对比 81.5%)，但召回率比 ARV-YOLO 低 3.8 个百分点，导致 mAP@0.5 比 ARV-YOLO 低 2.2 个百分点；YOLOX 召回率比较高(73.6%)，但精确率只有 67.0%，整体 mAP@0.5 比 ARV-YOLO 低 6.2 个百分点。这说明 ARV-YOLO 把多尺度特征融合和注意力机制结合起来后，能在复杂交通场景里更好地平衡检测精度和召回率。

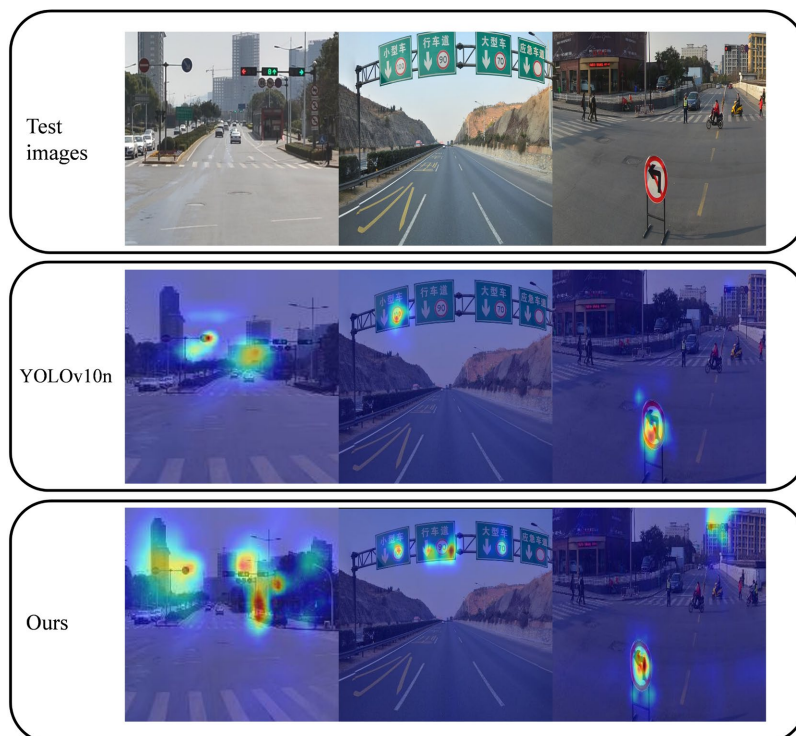


Figure 8. Heatmap comparison
图 8. 热力图对比

ARV-YOLO 在 $mAP@0.5:0.95$ 这个指标上达到了 62.3%，不仅比两阶段检测器提升明显，和 RTMDet、YOLOX 比也分别高 1.7 和 7.4 个百分点。这说明它对不同大小的交通标志都能准确识别和定位，鲁棒性很强。ARV-YOLO 参数量很低，但检测性能和主流算法相当甚至更好，尤其是处理小目标、多尺度交通标志时优势明显，更适合装在车载终端、边缘计算设备这些算力有限的地方，有很高的工程应用价值。为进一步验证 ARV-YOLO 模型对交通标志区域的关注机制及关键特征定位能力，对 YOLOv10n 与改进后的模型在 TT100K 测试集上的热力图进行可视化对比分析。热力图对比如图 8 所示。

可视化结果表明，YOLOv10n 在识别远处的小交通标志时，图像上显示的激活区域不仅亮度低，而且边缘也不清楚。这种情况容易让系统漏掉一些标志，或者把不是标志的东西错认成标志。ARV-YOLO 模型对应的热力图不一样。它的激活区域更集中，边缘也更清晰。就算是远处的小目标，激活区域的亮度也很高，背景里那些没用的干扰信息被压下去了。在路口这种标志大小变化很大的场景里，ARV-YOLO 有个多尺度特征融合的设计。这个设计让热力图既能覆盖近处大尺寸的标志，也能覆盖远处小尺寸的标志。从图 8 上能看出来，它对不同大小的目标都能关注到。

4. 结论

针对复杂道路场景中交通标志尺度变化大、目标小及背景干扰强等难点问题，本文提出了一种基于改进 YOLOv8n 的交通标志检测算法 ARV-YOLO。通过在主干网络引入 RFCACConv 模块，有效融合感受野空间特征与坐标注意力机制，增强了网络对多尺度关键区域特征的表达能力；构建 ASF-P2 多尺度特征融合结构并增加 P2 检测层，充分利用浅层高分辨率特征，显著提升了小尺度交通标志的检测能力；同时引入 V7DS 下采样模块，通过最大池化与卷积分支并行策略，在降低特征分辨率的同时有效保留了空间细节信息。上述改进在保持模型轻量化的前提下，实现了特征提取、多尺度融合与空间信息保留的协同优化。

在 TT100K 数据集上的实验结果表明，ARV-YOLO 模型的 $mAP@0.5$ 达到 80.8%，较基线 YOLOv8n 提升 9.9 个百分点，精确率和召回率分别提升 9.3%和 7.1%，且参数量仅为 2.5 M。与 Faster R-CNN、Cascade R-CNN、RTMDet 及 YOLOX 等主流算法相比，ARV-YOLO 在参数量降低 48.6%~96.4%的情况下，仍取得了更优的检测精度和鲁棒性。实验结果与可视化分析进一步验证了各模块在增强小目标与多尺度目标检测、抑制背景干扰及提升定位精度方面的有效性。

后续研究将聚焦于进一步优化模型的实时推理速度，探索模型在边缘计算设备上的高效部署策略，并结合时序信息提升视频序列中交通标志检测的稳定性，以更好地满足自动驾驶系统在复杂动态环境下的实际应用需求。

参考文献

- [1] 山显英, 张琳, 李泽慧. 深度学习驱动下的目标检测研究进展综述[J]. 计算机工程与应用, 2025, 61(1): 24-41.
- [2] 李建东, 徐米雪. 基于改进 YOLOv8 的交通标志检测算法研究[J/OL]. 软件导刊: 1-9. <https://link.cnki.net/urlid/42.1671.TP.20251030.1706.012>, 2026-03-15.
- [3] Zhang, Z., Feng, B. and Liu, H. (2026) TS-YOLO: A Small Traffic Sign Detection Algorithm for Various Harsh Driving Conditions in Bad Weather. *Scientific Reports*, **16**, Article No. 4444. <https://doi.org/10.1038/s41598-025-34544-9>
- [4] Wang, G., Jin, P., Qi, Z. and Li, X. (2025) Traffic Sign Detection Method Based on Improved Yolov8. *Scientific Reports*, **15**, Article No. 19385. <https://doi.org/10.1038/s41598-025-03792-0>
- [5] Zhao, W., Li, L. and Gong, X. (2025) CPB-YOLOv8: An Enhanced Multi-Scale Traffic Sign Detector for Complex Road Environment. *Information*, **16**, Article 798. <https://doi.org/10.3390/info16090798>
- [6] Kang, M., Ting, C., Ting, F.F. and Phan, R.C. (2024) ASF-YOLO: A Novel YOLO Model with Attentional Scale Sequence Fusion for Cell Instance Segmentation. *Image and Vision Computing*, **147**, Article ID: 105057. <https://doi.org/10.1016/j.imavis.2024.105057>
- [7] Zhang, X., Liu, C., Yang, D., et al. (2023) RFACConv: Innovating Spatial Attention and Standard Convolutional Operation. arXiv: 2304.03198.

-
- [8] Wang, C., Bochkovskiy, A. and Liao, H.M. (2023) YOLOv7: Trainable Bag-Of-Freebies Sets New State-of-the-Art for Real-Time Object Detectors. *2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Vancouver, 17-24 June 2023, 7464-7475. <https://doi.org/10.1109/cvpr52729.2023.00721>
- [9] Ren, S., He, K., Girshick, R. and Sun, J. (2017) Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **39**, 1137-1149. <https://doi.org/10.1109/tpami.2016.2577031>
- [10] Lyu, C., Zhang, W., Huang, H., *et al.* (2022) RTMDet: An Empirical Study of Designing Real-Time Object Detectors. arXiv: 2212.07784.
- [11] Ge, Z., Liu, S., Wang, F., *et al.* (2021) YOLOX: Exceeding YOLO Series in 2021. arXiv: 2107.08430.