

ATS-YOLO: 基于改进YOLO11的轻量化无人机小目标检测方法

刘 爽, 杨继芸

华北理工大学电气工程学院, 河北 唐山

收稿日期: 2026年6月8日; 录用日期: 2026年6月30日; 发布日期: 2026年7月13日

摘 要

无人机目标检测面临小目标尺度小且密集、视角变化大与目标姿态多样, 以及黑夜和强光等复杂光照环境带来的特征退化问题, 同时受限于平台计算资源, 使得检测精度、鲁棒性与实时性之间难以兼顾。为解决以上问题, 本文提出了基于YOLO11的改进检测模型ATS-YOLO。首先, 在主干网络中引入C2f-CA模块以增强特征表达能力, 在检测头中引入TADDH结构以提升多尺度特征融合效果, 并在训练阶段对边界框回归损失进行改进以强化对小目标的关注。通过上述设计, 从特征提取、特征融合与损失约束三个层面实现了协同优化。实验结果表明, 所提出组合ATS-YOLO在VisDrone2019数据集上相较于基线YOLO11模型在小目标检测性能上取得了明显提升, ATS-YOLO相比YOLO11n在mAP@0.5和mAP@0.5:0.95上分别取得了0.025、0.018的提升, 在Params和Model size上分别减少了0.292 M和0.61 MB。此外, 在对比实验上, 改进模块的指标也显著优于其他算法。实验结果表明, ATS-YOLO在保持模型轻量化的同时实现了更优的检测效果, 验证了该组合策略的有效性, 为无人机场景下高效、鲁棒的目标检测提供了有效解决方案。

关键词

无人机目标检测, 小目标, 多尺度特征, 轻量化

ATS-YOLO: An Improved YOLO11-Based Lightweight Detector for Small Targets in UAV Imagery

Shuang Liu, Jiyun Yang

College of Electrical Engineering, North China University of Science and Technology, Tangshan Hebei

Received: June 8, 2026; accepted: June 30, 2026; published: July 13, 2026

Abstract

UAV target detection is faced with the problems of small and dense small targets, large changes in perspective, diverse target poses, and feature degradation caused by complex illumination environments such as dark nights and strong light. At the same time, it is limited by platform computing resources, making it difficult to balance detection accuracy, robustness and real-time performance. In order to solve the above problems, this paper proposes an improved detection model ATS-YOLO based on YOLO11. Firstly, the C2f-CA module is introduced into the backbone network to enhance the feature expression ability, the TADDH structure is introduced into the detection head to improve the multi-scale feature fusion effect, and the bounding box regression loss is improved in the training stage to strengthen the focus on small targets. Through the above design, collaborative optimization is realized from three aspects: feature extraction, feature fusion and loss constraint. The experimental results show that the proposed combined ATS-YOLO has achieved a significant improvement in small target detection performance compared with the baseline YOLO11 model on the VisDrone2019 dataset. Compared with YOLO11n, ATS-YOLO has achieved 0.025 and 0.018 improvements on mAP@0.5 and mAP@0.5:0.95, respectively, and reduced 0.292 M and 0.61 MB on Params and Model size, respectively. In addition, in the comparative experiment, the index of the improved module is also significantly better than other algorithms. The experimental results show that ATS-YOLO achieves better detection results while maintaining the lightweight of the model, which verifies the effectiveness of the combined strategy and provides an effective solution for efficient and robust target detection in UAV scenarios.

Keywords

UAV Target Detection, Small Target, Multi-Scale Feature, Lightweight

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

随着无人机技术的快速发展以及低空经济的持续推进, 无人机在城市巡检[1]、农作物分析[2]、地质灾害检测[3]、交通监测[4]、应急救援[5]和科学研究[6]等领域得到了广泛应用。无人机平台能够获取大范围、高分辨率的航拍图像, 为智能感知和自动化分析提供了丰富的数据来源。然而, 在实际应用场景中, 无人机拍摄的目标通常具有尺度小、密集度高、背景复杂以及姿态变化明显等特点, 使得目标检测任务面临较大的挑战。尤其是在低空航拍数据集中, 大量目标在图像中仅占据极少像素区域, 传统目标检测算法往往难以准确提取其有效特征, 从而影响检测精度。因此, 研究面向小目标场景的高性能目标检测算法成为计算机视觉领域的重要研究方向之一。

近年来, 针对上述问题, 研究者们基于 YOLO 系列算法提出了多种改进方案。刘涛[7]等提出改进 YOLOv5s 方法, 通过重构锚框、引入 SEC2f 注意力模块及解耦检测头, 提升小目标检测精度, 但对复杂环境鲁棒性仍有提升空间。奉志强[8]等提出改进 YOLOv5 方法, 通过构建 SCAM 注意力模块及 SC-AFF 特征融合结构, 并引入 Transformer, 提高密集小目标检测精度, 但模型复杂度增加。聂源[9]等提出 MFF-YOLOv7 方法, 通过多尺度特征融合、注意力机制及解耦检测头提升小目标检测精度并减少漏检, 但模型结构较复杂, 计算开销较大。张光华[10]等提出改进 YOLOv7-tiny 方法, 通过引入 ConvMixer、小目标

检测层及解耦检测头,提升无人机航拍小目标检测性能,但模型复杂度有所增加。陈海秀[11]等提出改进 YOLOv10n 方法,通过引入 C2f-MC 模块与小目标特征增强结构,提高密集行人小目标检测精度,但主要面向固定摄像场景,对无人机复杂环境适应性不足。高立鹏[12]等提出 REIW-YOLOv10n 算法,通过优化骨干网络、特征融合及损失函数,有效提升井下安全帽小目标检测精度,但方法依赖特定场景,泛化能力仍有待提升。

随着计算机视觉与深度学习技术的快速发展,小目标检测逐渐成为研究热点,并展现出广泛的应用潜力。在此背景下,学术界与工业界围绕 YOLO11 等先进检测模型展开了大量研究与改进工作。彭哲轩[13]等提出 ADE-YOLO 方法,通过轻量化结构设计、自适应特征融合检测头及图像增强策略,提高退化环境下小目标检测精度与鲁棒性,但在多场景泛化能力方面仍存在一定局限。张志豪[14]等提出 ACFI-YOLO11 方法,通过跨层特征交互模块、Tiny Head 及 SPD 卷积增强小目标特征表达,在提升检测精度的同时实现轻量化,但模型结构仍较复杂。黄毅[15]等提出 LED-YOLOv10 方法,通过引入 MCIR、C2f-CM 及 Focal-EIoU 损失函数,实现多尺度特征增强与模型轻量化,在复杂背景下显著提升检测精度,但结构较复杂,推理效率仍有优化空间。

综上,现有方法主要通过改进特征提取结构、多尺度融合机制及检测头设计来提升无人机小目标检测性能,但仍存在模型结构复杂、计算开销增加以及在极端复杂场景下鲁棒性不足等问题。基于上述分析,本文在 YOLO11n 模型的框架基础上提出一种创新的无人机目标检测轻量级的组合方法。首先,在检测头部分引入 TADDH (Task-Adaptive Dynamic Detection Head)结构,通过任务自适应的特征调整机制增强多尺度特征表达能力,提高检测头对复杂场景目标的感知能力。其次,在网络主干与特征融合阶段引入 C2f-CA (C2f with Coordinate Attention) [16]模块,利用坐标注意力机制在保持轻量化结构的同时强化空间位置信息建模,从而提升网络对小尺度目标和细粒度结构的特征捕获能力。此外,为进一步提高小目标检测性能,本文的 SADR-Loss 在原有 IoU 类损失的基础上,引入尺度相关的权重因子,对小目标赋予更高权重,对大目标适当降低权重,从而实现对不同尺度目标的动态平衡。

通过上述改进,所提出的方法在保持模型轻量化和实时性的同时,显著提升了模型对小目标的检测能力。实验结果表明,在 VisDrone2019 无人机目标检测数据集上,改进模型在检测精度与特征表达能力方面均取得了更好的表现,验证了所提出方法的有效性。

2. YOLO11 模型

YOLO11[17]是 YOLO 系列算法的最新版本之一,在网络结构设计、特征表达能力以及计算效率方面进行了进一步优化,使模型在检测精度与推理速度之间取得更好的平衡。

YOLO11 模型通常提供多种不同规模的网络版本,包括 YOLO11n、YOLO11s、YOLO11m、YOLO11l 和 YOLO11x 等。YOLO11 系列模型的参数量、计算量如表 1 所示。其中,YOLO11n 版本是最轻量化的模型,由输入预处理、轻量化骨干(Backbone)、精简颈部(Neck)、无锚解耦检测头(Head)四部分串联构成,其整体架构如图 1 所示。YOLO11n 的参数量和计算量相对较小,能够在保证一定检测精度的同时显著降低模型复杂度,因此更加适用于资源受限设备以及实时检测任务。

而 YOLO11n 模型的默认输入尺寸为 640×640 。然而,在无人机航拍场景中,图像中的目标通常尺寸较小且分布密集。VisDrone2019 数据集中的大量目标在原始图像中仅占据较少像素,如果输入分辨率较低,小目标在特征图中容易被下采样削弱,从而影响检测效果。因此,本文在实验中将输入图像尺寸设置为 960×960 ,以保留更多图像细节信息,提高模型对小目标的特征表达能力。由于 YOLO 系列模型采用全卷积结构,能够适应不同尺寸的输入图像,因此在不改变网络结构的情况下即可使用更高分辨率进行训练与推理。

Table 1. Parameter count and computational cost of the YOLO11 series models
表 1. YOLO11 系列模型的参数量与计算量

Model	Input size/pixels	Params/M	FLOPs/B
YOLO11n	640	2.6	6.5
YOLO11s	640	9.4	21.5
YOLO11m	640	20.1	68.0
YOLO11l	640	25.3	86.9
YOLO11x	640	56.9	194.9

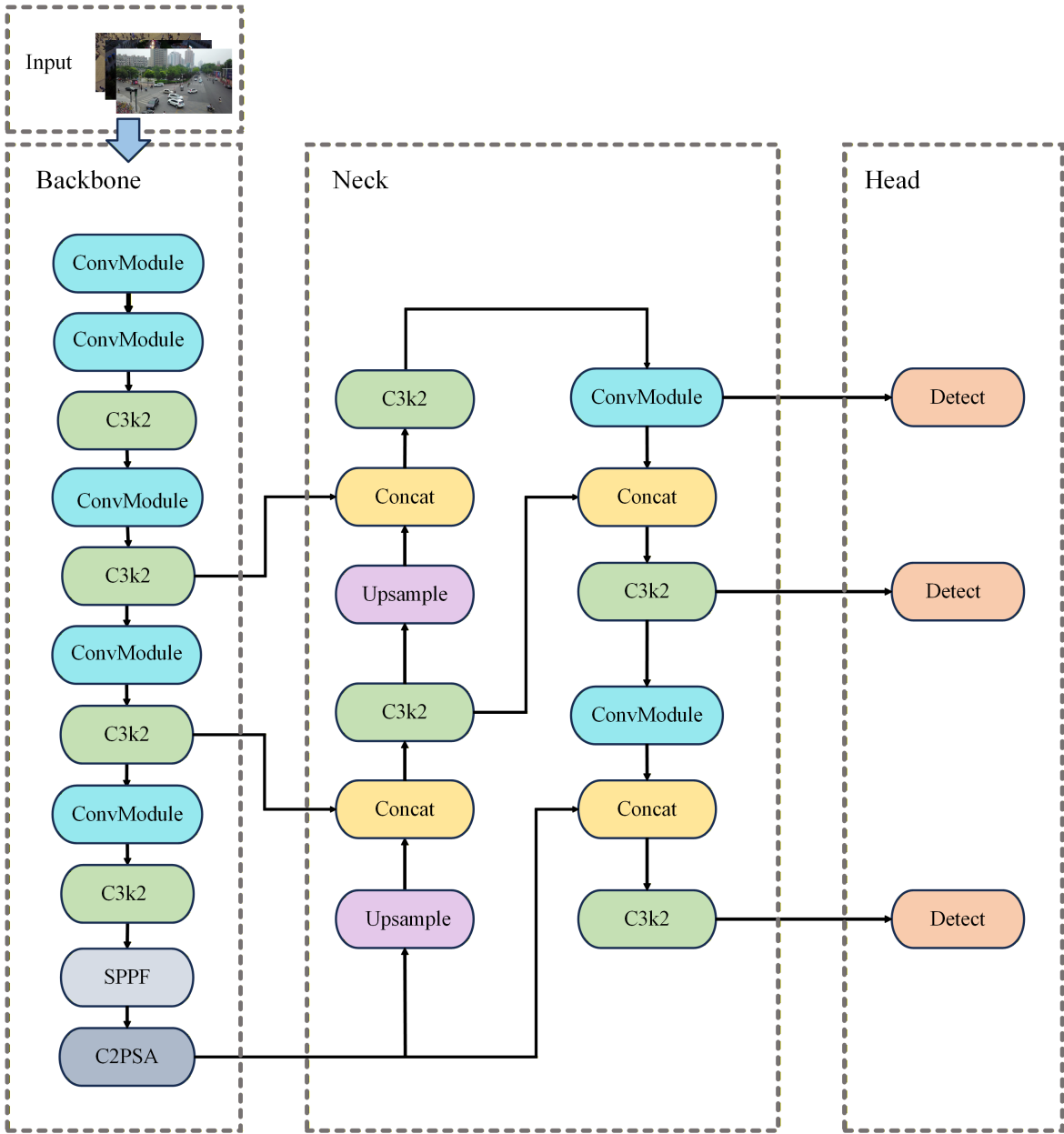


Figure 1. YOLO11 network structure
图 1. YOLO11 网络结构

3. ATS-YOLO 模型

3.1. ATS-YOLO 的总体结构

ATS-YOLO 是在 YOLO11n 模型基础上提出的一种改进目标检测网络，其整体结构仍由主干网络 (Backbone)、特征融合网络(Neck)和检测头(Head)三部分组成。在保持原有网络整体框架不变的前提下，本文针对小目标检测任务对 YOLO11n 的关键模块进行了模块的轻量级组合优化。

具体而言，本文在主干网络中引入 C2f-CA 模块，将融合了 Coordinate Attention 机制的 C2f 结构替换原有的 C3k2 模块，以增强网络对空间位置信息的建模能力，从而提升小目标的特征表达能力。在检测头部分，引入 TADDH (Task-Adaptive Dynamic Detection Head) [18]结构，增强多尺度特征的动态建模能力，提升模型在复杂场景下的检测性能。此外，在训练阶段对边界框回归损失函数进行优化，在标准 IoU 损失的基础上，引入尺度感知加权机制，针对小目标赋予更高的损失权重，从而提升模型对小目标的关注程度。

需要指出的是，本文的核心创新并不在于提出全新的基础模块，而是在于对这些已有先进模块进行有效组合与适配。通过将 C2f-CA、TADDH 以及损失加权策略有机融合到 YOLO11n 框架中，构建了一个轻量化且针对小目标检测任务优化的检测网络。实验结果表明，该组合策略在保持模型轻量化的同时，显著提升了小目标检测的精度与鲁棒性。改进后的 YOLO11 模型整体结构如图 2 所示。

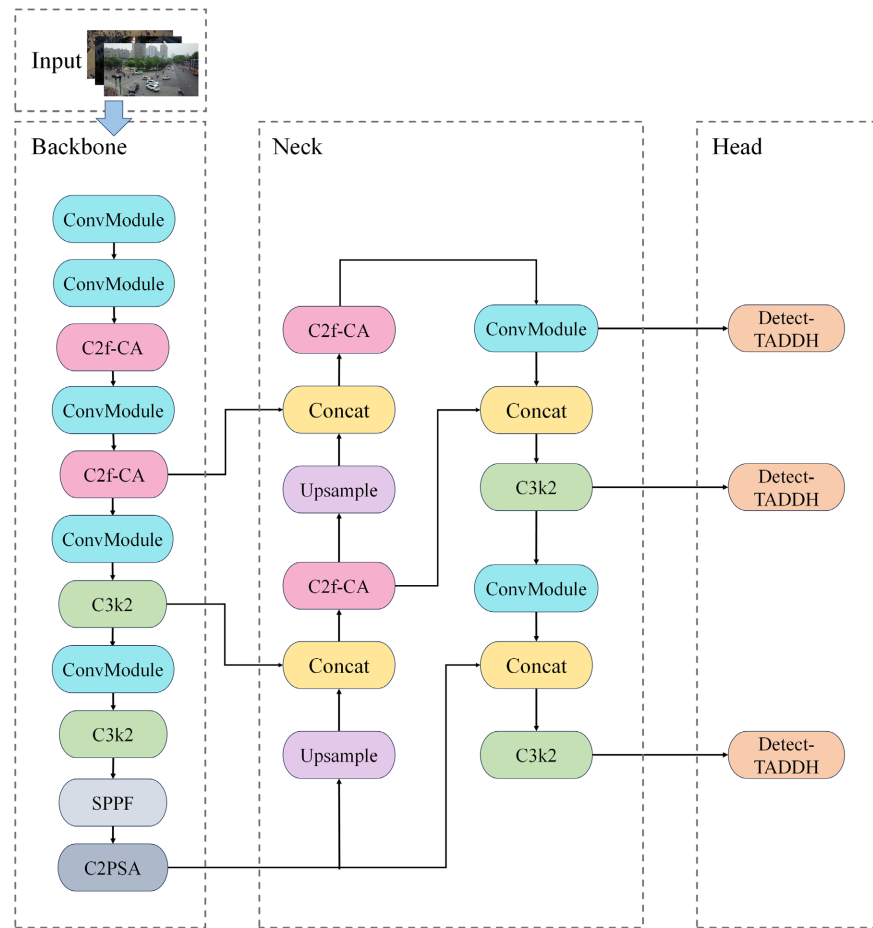


Figure 2. ATS-YOLO network structure
图 2. ATS-YOLO 网络结构

3.2. C2f-CA 模块

在目标检测网络中, 主干网络主要负责从输入图像中提取多层次特征。然而, 在无人机航拍场景中, 由于目标尺寸较小且背景复杂, 传统卷积结构在特征提取过程中容易忽略关键区域的空间位置信息, 从而影响小目标的特征表达能力。为增强网络对空间信息的感知能力, 本文在 YOLO11n 主干网络中引入 C2f-CA 模块, 其结构如图 3 所示。Coordinate Attention (CA) 模块是一种轻量级注意力机制, 结构如图 4 所示, 通过分别在水平和垂直方向进行全局池化, 将位置信息融入通道注意力中, 从而在增强特征表达的同时保留空间位置信息。其结构简单、计算开销小, 能够有效提升模型在小目标和复杂场景下的特征感知能力。

C2f-CA 模块是在 C2f 结构基础上引入 Coordinate Attention (CA) 机制形成的特征增强模块。首先, 设输入特征图 $X \in R^{B \times C \times H \times W}$, 其中 B 表示批大小, C 表示通道数, H 和 W 分别表示特征图的高度和宽度。为了降低计算复杂度并提高特征表达效率, 首先通过 1×1 卷积按照公式(1)对输入特征进行通道压缩, 其中 X_c 为压缩后的特征表示。

$$X_c = \text{Conv}_{1 \times 1}(X) \quad (1)$$

随后, 将压缩后的特征输入到 C2f 结构中进行特征提取。C2f 模块通过多个 Bottleneck 单元对特征进行逐级提取, 并通过特征拼接实现多分支信息融合, 其过程可表示为

$$F = \text{Concat}(B_1(X_c), B_2(X_c), \dots, B_n(X_c)) \quad (2)$$

其中 $B_i(\cdot)$ 表示第 i 个 Bottleneck 单元, $\text{Conv}(\cdot)$ 表示特征拼接操作。通过该结构能够在保持较低计算量的同时增强网络的特征表达能力。

随后, 将 F 输入 Coordinate Attention 模块, 沿高度和宽度方向分别进行全局平均池化:

$$z_h(x) = \frac{1}{W} \sum_{i=1}^W F(x, i), z_w(y) = \frac{1}{H} \sum_{j=1}^H F(j, y) \quad (3)$$

其中 z_h 和 z_w 分别表示沿高度方向和宽度方向的空间编码特征。

再通过卷积变换生成对应的注意力权重, 并对原始特征进行加权, 其表达式为:

$$Y = F \odot \sigma(A_h + A_w) \quad (4)$$

其中 A_h 和 A_w 分别表示两个方向生成的注意力权重, $\sigma(\cdot)$ 表示激活函数, $\odot$ 表示逐元素乘法操作。

最终得到增强后的特征 Y , 作为 C2f-CA 模块的输出。该模块通过融合空间坐标信息, 增强了特征图对关键区域的表达能力, 适用于目标检测等任务。

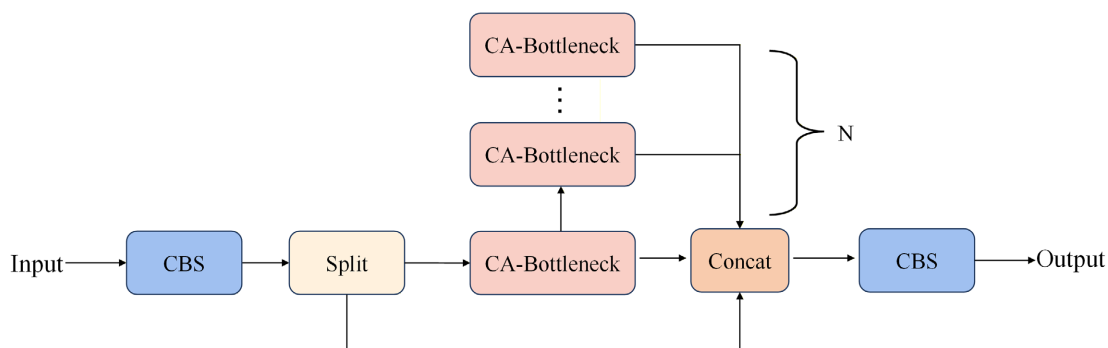


Figure 3. C2f-CA module structure

图 3. C2f-CA 模块结构

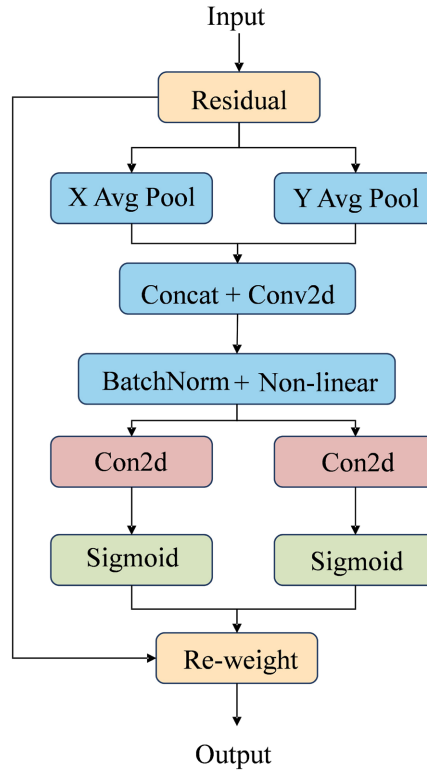


Figure 4. CA module structure

图 4. CA 模块结构

3.3. TADDH 模块

TADDH (Task-Adaptive Dynamic Detection Head)是一种任务自适应动态检测头结构, 其整体框架如图 5 所示。

首先, 设来自特征金字塔网络(FPN)的输入特征图为 $\{F_1, F_2, \dots, F_L\}$, 其中 $F_i \in R^{B \times C \times H_i \times W_i}$ 表示第 i 层特征。首先对各层特征分别进行共享卷积变换, 得到统一维度的共享特征表示:

$$F_s^i = \phi(F_i), i=1, \dots, L \quad (5)$$

其中 $\phi(\cdot)$ 表示共享权重的卷积层, 用于将不同层特征映射到同一特征空间。

随后, 将共享特征输入任务自适应动态模块。考虑到目标检测任务同时包含分类与边界框回归两个子任务, 两者对特征表达的需求存在差异, 因此通过动态权重生成机制对特征进行任务自适应调整。动态权重 W_t 从共享特征中自适应生成:

$$W_t = \text{Softmax}(\psi_t(F_s)), t \in \{cls, reg\} \quad (6)$$

任务调整后的特征表示为:

$$F_t = W_t \odot F_s \quad (7)$$

其中 $\odot$ 表示逐元素乘法。

对于分类任务, 任务调整后的特征为 $F_{cls} = W_{cls} \odot F_s$, 分类预测为:

$$P_{cls} = \sigma(\text{Conv}_{cls}(F_{cls})) \quad (8)$$

其中 $\sigma(\cdot)$ 为激活函数, $Conv_{cls}$ 为分类卷积层。

对于回归任务, 任务调整后的特征为 $F_{reg} = W_{reg} \odot F_s$, 回归预测为:

$$B_{reg} = Conv_{reg}(F_{reg}) \quad (9)$$

其中 B_{reg} 为预测的边界框参数。

通过上述过程, TADDH 模块能够根据不同检测任务的需求动态调整特征表达, 使网络在保持共享特征信息的同时适应不同任务需求。

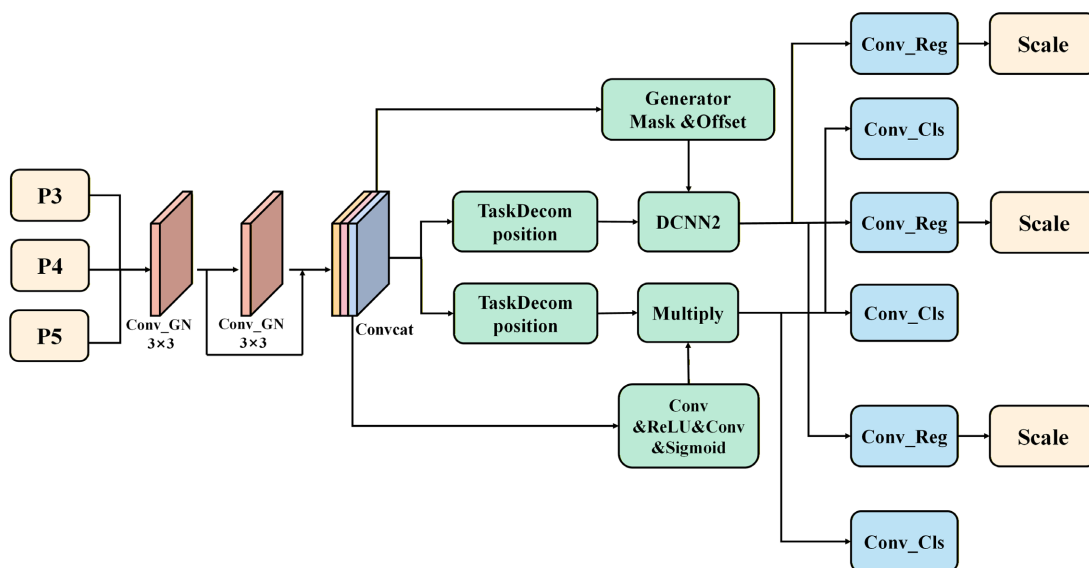


Figure 5. TADDH module structure

图 5. TADDH 模块结构

3.4. SADR-Loss

SADR-Loss (Scale-Aware Dynamic Reweighting Loss) 相比传统 IoU 损失, 该方法无需额外结构修改, 仅通过简单的尺度加权策略即可有效提升小目标检测性能, 具有良好的通用性与轻量化优势。

首先, 为提升模型对小目标的定位能力, 本文在传统边界框回归损失的基础上, 引入尺度感知加权机制, 并结合分布式焦点损失(DFL), 构建如下总损失函数:

$$L = L_{IoU} + L_{DFL} \quad (10)$$

其中, IoU 回归损失定义为:

$$L_{IoU} = \frac{1}{\sum_{i \in \mathcal{F}} w_i} \sum_{i \in \mathcal{F}} w_i \cdot s_i (1 - CIoU_i) \quad (11)$$

式中, $CIoU_i$ 表示预测框与真实框之间的重叠度, w_i 为分类置信度权重, 用于衡量不同样本的重要性。

为增强小目标的学习能力, 引入尺度权重因子 s_i , 其定义为:

$$s_i = \begin{cases} \alpha, & A_i < T \\ 1, & A_i \geq T \end{cases} \quad (12)$$

其中, A_i 表示目标面积:

$$A_i = (x_i^2 - x_i^l)(y_i^2 - y_i^l) \quad (13)$$

T 为小目标阈值(设为 40×40), α 为加权系数(设为 1.5)。当目标面积小于阈值时, 其回归损失被放大, 从而在训练过程中获得更大的梯度贡献。

同时, 为提升边界框回归的精细度, 引入分布式焦点损失:

$$L_{DFL} = \frac{1}{\sum_i w_i} \sum_i w_i \cdot DFL(p_i, t_i) \quad (14)$$

最终损失通过 IoU 项与 DFL 项共同优化, 实现对目标位置与边界的联合约束。

4. 实验与分析

4.1. 数据集

VisDrone2019-DET 数据集[19]由天津大学 AISKEYEYE 团队构建, 是一个面向无人机视角的大规模公共基准数据集涵盖中国 14 个城市、城乡环境、10 类目标(行人、人群、自行车、小轿车、厢式货车、卡车、三轮车、带篷三轮车、公交车、摩托车)以及稀疏/拥挤场景。该数据集目标具有尺度小、分布密集、背景复杂等特点, 对检测算法的空间位置感知能力、多尺度动态建模能力以及小目标定位精度提出了较高要求。

上述挑战与本文所提 C2f-CA 模块(增强空间位置建模)、TADDH 模块(提升任务自适应多尺度表达)以及损失加权策略(强化小目标定位)的组合改进方向高度契合。因此, VisDrone2019-DET 能够有效验证本文组合策略在无人机小目标检测任务中的整体性能与有效性。

4.2. 评价指标

本文采用平均精度均值(mean Average Precision, mAP)中的 $mAP@0.5$ 和 $mAP@0.5:0.95$ 、参数量(Param)和模型大小(Model-size)作为评估模型性能的核心指标。

4.2.1. 平均精度均值

平均精度均值(mean Average Precision, mAP)为数据集中所有类别 AP 的平均值, 是评估多类别检测模型整体性能的重要指标。

$mAP@0.5$ 表示在 IoU (Intersection over Union)阈值为 0.5 时计算的 mAP:

$$mAP@0.5 = mAP|_{IoU=0.5} \quad (15)$$

$mAP@0.5:0.95$ 表示在多个 IoU 阈值下的平均 mAP, 阈值从 0.5 到 0.95, 步长为 0.05:

$$mAP@0.5:0.95 = \frac{1}{10} \sum_{t \in \{0.5, 0.55, \dots, 0.95\}} mAP|_{IoU=t} \quad (16)$$

4.2.2. 参数量(Params)

模型参数量是指网络中所有可训练参数的总数, 反映模型的复杂度和表达能力。计算公式为:

$$Params = \sum_{l=1}^L Params_l \quad (17)$$

其中 L 为网络总层数, $Params$ 为第 l 层的参数量。

4.2.3. 模型大小(Model Size)

模型大小是指存储模型参数所需的磁盘空间, 通常以兆字节(MB)为单位, 直接影响模型的部署可行性。计算公式为:

$$\text{Model-size(MB)} = \frac{\text{Params} \times b}{1024^2} \quad (18)$$

其中:

- Params: 模型总参数量
- b : 每个参数存储所需的字节数(bytes per parameter)
- $1024^2 = 1,048,576$: 字节(Bytes)转换为兆字节(MB)的换算系数

4.3. 消融实验

为验证本文所提改进模块组合的有效性, 本文在 VisDrone2019 数据集上设计消融实验进行对比分析。消融实验通过控制变量法, 在基准模型(YOLO11n)基础上逐步引入各改进模块, 观察模型性能的变化, 从而定量评估每个模块以及该组合对整体检测精度的贡献。实验结果如表 2 所示。

Table 2. Ablation experiment results of the proposed algorithm module on the VisDrone2019 validation set

表 2. 提出的算法模块在 VisDrone2019 验证集上的消融实验结果

方法	C2f-CA	TADDH	SADR-Loss	mAP@0.5	mAP@0.5:0.95	Params/M	Model size/MB
YOLO11n				0.438	0.268	2.592	5.25
方法(1)	√			0.452	0.274	2.742	5.57
方法(2)		√		0.445	0.273	2.207	4.49
方法(3)			√	0.441	0.271	2.584	5.25
方法(4)		√	√	0.444	0.272	2.207	4.49
方法(5)	√		√	0.442	0.273	2.660	5.40
方法(6)	√	√		0.452	0.278	2.304	4.69
方法(7)	√	√	√	0.463	0.286	2.300	4.64

实验中, 依次将不同改进模块引入 YOLO11n 模型, 以量化各模块在检测精度与模型复杂度方面的实际贡献。首先, 在基线模型中引入 C2f-CA 模块后, mAP@0.5 由 0.438 提升至 0.452, mAP@0.5:0.95 提升至 0.274, 说明该模块通过引入位置信息增强机制, 有效提升了模型对小目标及细粒度特征的表达能力, 但参数量略有增加。接着, 仅引入 TADDH (方法 2) 时, mAP@0.5 提升至 0.445, 同时参数量由 2.592 M 降至 2.207 M, 表明其在提升多尺度特征融合能力的同时有效降低了模型复杂度。随后, 引入 SADR-Loss (方法 3) 后, mAP@0.5 提升至 0.441, 说明改进损失函数能够在一定程度上优化定位精度。

进一步地, 模块组合实验表明不同模块之间具有良好的协同作用。当引入 TADDH 与 SADR-Loss (方法 4) 时, 模型精度较单一模块进一步提升; 结合 C2f-CA 与 SADR-Loss (方法 5) 时, 模型在保证精度提升的同时保持了较低的复杂度; 当 C2f-CA 与 TADDH 共同作用(方法 6) 时, mAP@0.5 达到 0.452, mAP@0.5:0.95 提升至 0.278, 在精度与参数量之间取得了较优平衡。

最终, 当三种改进模块全部引入(方法 7) 时, 模型性能达到最优, mAP@0.5 和 mAP@0.5:0.95 分别提升至 0.463 和 0.286, 较基线分别提升 2.5% 和 1.8%。同时, 模型参数量为 2.300 M, 模型体积为 4.64 MB, 较基线未出现明显增加。综合来看, 所提出的多模块协同优化策略在保持模型轻量化的同时, 有效提升了检测精度, 验证了各改进模块的有效性与互补性。

针对表 2 的消融实验结果, TADDH 单独使用时(方法 2) 精度提升微弱(mAP@0.5 仅提升 0.007), 原因是其动态重加权机制需要高质量的空间特征作为依据, 而基线主干网络缺乏显式的空间注意力, 导致小目标位置信息模糊, 动态卷积核难以找到最优采样策略。只有当 C2f-CA 先行引入、增强空间编码后,

TADDH 的潜力才得以释放(如方法 7 提升 0.025)。

4.4. 对比实验

为检验 ATS-YOLO 算法在下的综合性能, 将其与为验证所提出模型在轻量化与性能方面的优势, 将改进模型与 YOLOv5s、YOLOv8n、YOLOv10n 和 YOLOv11n 等主流轻量化目标检测模型进行了对比实验。与其他模型相比, 改进模型在 $mAP@0.5$ 上达到 0.463, $mAP@0.5:0.95$ 为 0.286, 优于所有对比模型。同时, 参数量为 2.3, 模型大小为 4.64, 在常见的轻量化模型中也是优异的存在。其中, 与 YOLOv5s 相比, $mAP@0.5$ 提升 8.6 个百分点, 参数量仅增加了 0.523; 与 YOLOv8n 相比, $mAP@0.5$ 提升 3.3 个百分点, 参数量减少了 0.713。与 YOLOv10n 相比, $mAP@0.5$ 提升 2.9 个百分点, 参数量减少了 0.411。与上述模型相比, 改进模型在召回率上有显著提升, 精度率仅次于 YOLO11n, 同时模型大小仅高于 YOLOv5s。本文模型在保持较高检测精度的同时, 保持了较低的参数量和模型大小, 表现出更优的轻量化与实时检测能力, 适用于小目标检测的应用场景。结果如表 3 所示。

Table 3. Comparison of common model experiments
表 3. 常见模型实验对比

方法	P	R	$mAP@0.5$	$mAP@0.5:0.95$	Params/M	Model size/MB
YOLOv5s	0.467	0.372	0.377	0.215	1.777	3.81
YOLOv8n	0.530	0.416	0.430	0.262	3.013	5.97
YOLOv10n	0.545	0.412	0.434	0.267	2.711	5.51
YOLO11n	0.553	0.412	0.438	0.268	2.592	5.25
ATS-YOLO	0.548	0.444	0.463	0.286	2.300	4.64

4.5. 实验结果可视化对比

为分析模型在各类别上的检测性能, 采用归一化混淆矩阵进行可视化, 如图 6 所示。图(a)为 Baseline, 图(b)为改进模型。对角线表示正确分类, 左下三角反映漏检, 右上三角反映误检。

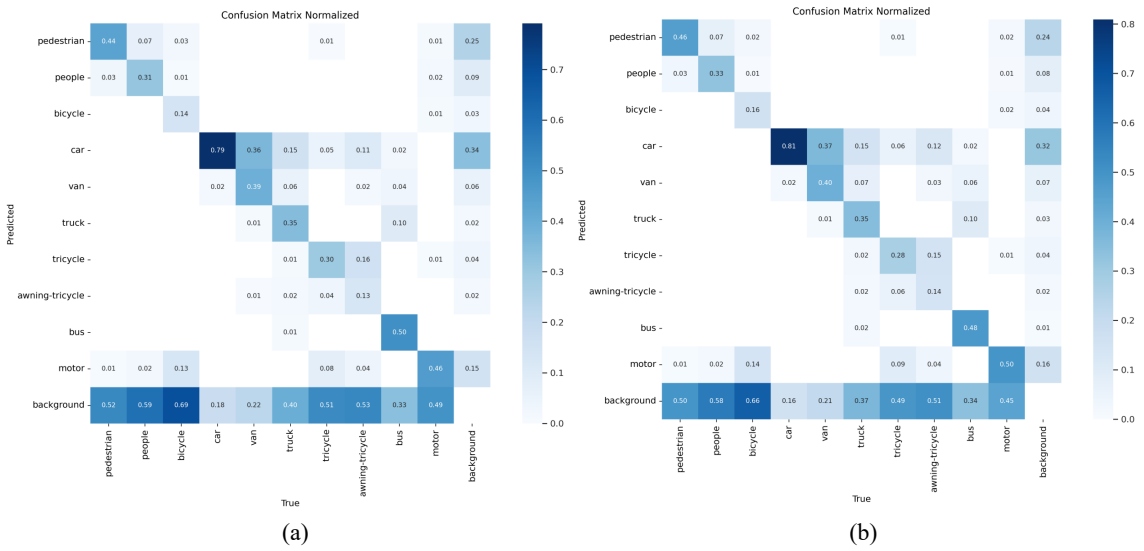


Figure 6. Confusion Matrix (a) YOLO11n Confusion Matrix; (b) ATS-YOLO Confusion Matrix
图 6. 混淆矩阵(a) YOLO11n 混淆矩阵; (b) ATS-YOLO 混淆矩阵

由图(a)可见, Baseline 在小目标类别及相似类别间存在一定混淆, 误检与漏检现象较为明显。相比之下, 图(b)中改进模型的对角线响应增强, 非对角区域明显减弱, 说明整体识别准确率得到提升。

此外, 在小目标类别上误分类现象显著减少, 同时相似类别间的混淆得到缓解, 表明改进组合方法有效提升了模型的判别能力与鲁棒性。

4.6. 检测结果可视化

为进一步分析模型的特征关注区域, 采用 Grad-CAM 对检测结果进行可视化, 如图 7 所示。其中深色区域代表模型的高关注区。改进模型的热力响应更加集中于车辆等目标区域, 激活区域更加清晰且连续, 对背景区域的响应明显减弱。同时, 在远距离及小尺度目标上, 改进模型仍能保持较强的响应强度, 说明其对小目标的特征感知能力得到提升。



Figure 7. Heatmap visualization comparison
图 7. 热力图可视化对比

同时为了更直观地展示改进模型的检测性能, 本文从 VisDrone2019 验证集中选取三组代表性图像, 分别对应黑夜场景、小目标密集和强光环境, 并以此为基础将改进模型与 YOLO11 基线模型进行对比分析。面对无人机视角下的黑夜场景(如图 8 所示)、小目标密集(如图 9 所示)和强光环境(如图 10 所示)等复杂场景, ATS-YOLO 检测精度更高以及目标漏检效果更好。检测结果对比如下图所示。

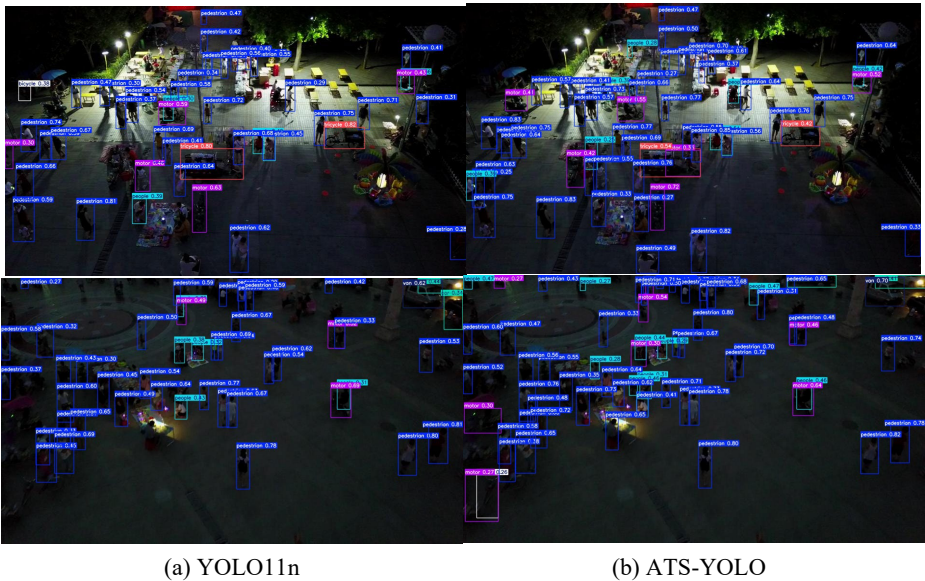


Figure 8. Missed detections of small targets in dim scenes
图 8. 昏暗场景下小目标的漏检

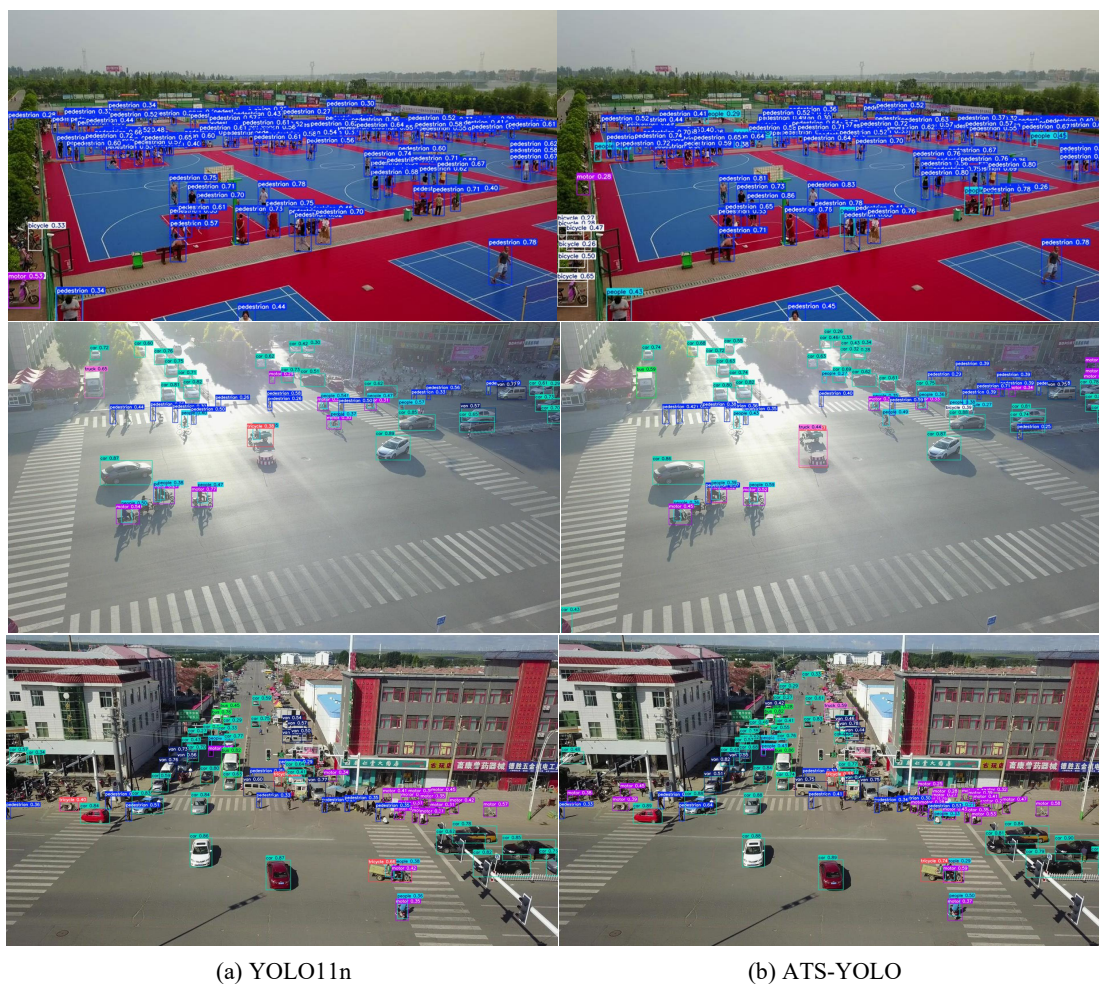


Figure 9. Missed detections of dense small targets

图 9. 密集小目标的漏检



Figure 10. Missed detections in strong light conditions

图 10. 强光环境下的漏检

改进后的新的模型组合在黑夜场景、小目标密集分布以及强光干扰等复杂环境下, 所提出的模型均表现出优于基准模型的检测性能。得益于更强的特征表达与多尺度信息融合能力, 模型能够在低照度条件下保持稳定识别, 在密集目标场景中有效减少漏检, 并在强光干扰下抑制背景噪声影响, 从而实现更准确、可靠的目标检测, 整体鲁棒性与适应性显著提升。

5. 结论

无人机航拍图像中的目标检测具有目标尺度小、背景复杂等特点, 尤其在小目标与极小目标检测方面面临较大挑战。针对传统 YOLO11 小目标识别能力不足与多尺度特征融合不佳的问题, 本文基于 YOLO11n 提出了一种轻量化小目标检测模型 ATS-YOLO, 通过融合 C2f-CA 坐标注意力、TADDH 动态检测头和强化回归损失完成模型改进。实验结果表明, 与原始 YOLO11n 的 mAP@0.5 为 0.438 相比, 改进后的 ATS-YOLO 在 mAP@0.5、参数量、召回率等核心指标上分别达到 0.463、2.3M 和 0.444, 模型体积仅为 4.64MB, 在参数量与模型规模得到有效控制的前提下实现了精度显著提升, 相较于同量级轻量化模型展现出更优的精度—效率权衡。通过黑暗场景、小目标密集场景及强光场景下的可视化对比可以看出, ATS-YOLO 能够更准确地定位并识别目标本体, 在多种复杂光照与密集目标条件下表现出较好的鲁棒性。但该模型仍存在一定局限性, 在部分极端低照度及超远极小目标条件下仍存在少量漏检, 且 TADDH 模块需配合注意力机制才能充分发挥性能。总体而言, 本文提出的 ATS-YOLO 在有效提升小目标检测精度的同时保持了网络轻量化特性, 可为无人机在复杂场景下的智能作业提供参考, 为无人机智能感知、智慧安防及低空监测等领域提供高效可靠的技术支撑, 具有良好的工程应用价值与推广前景。

参考文献

- [1] 蔡伯根, 李智宇, 王剑, 等. 面向铁路巡检的低空无人机智能感知与精密定位研究[J]. 铁路通信信号工程技术, 2025, 22(1): 1-12.
- [2] Duan, S.M., Zhang, X.F., Jiang, H. and Sun, K.G. (2025) CEDL YOLO: Multi-Scale Lightweight Cotton Field Weed Detection Method. *Computer Engineering and Applications*, **61**, 261-272.
- [3] 黄海峰, 林海玉, 吕奕铭, 等. 基于小型无人机遥感的单体地质灾害应急调查方法与实践[J]. 工程地质学报, 2017, 25(2): 447-454.
- [4] 徐慧智, 古旭楠. 无人机视角下交通小目标图像检测算法优化研究[J]. 计算机工程与应用, 2024, 60(21): 194-204.
- [5] 李宁, 刘青, 熊俊, 等. 基于无人机的火灾检测系统设计与实现[J]. 火灾科学(中英文), 2022, 31(1): 46-51.
- [6] 李腾, 张宝钢, 程晓, 等. 无人机在南极科学研究的应用: 进展与展望[J]. 武汉大学学报(信息科学版), 2022, 47(5): 651-664.
- [7] 刘涛, 高一萌, 柴蕊, 等. 改进 YOLOv5s 的无人机视角下小目标检测算法[J]. 计算机工程与应用, 2024, 60(1): 110-121.
- [8] 奉志强, 谢志军, 包正伟, 等. 基于改进 YOLOv5 的无人机实时密集小目标检测算法[J]. 航空学报, 2023, 44(7): 251-265.
- [9] 聂源, 赖惠成, 高古学. 改进 YOLOv7+Bytetrack 的小目标检测与追踪[J]. 计算机工程与应用, 2024, 60(12): 189-202.
- [10] 张光华, 李聪发, 李钢硬, 等. 基于改进 YOLOv7-tiny 的无人机航拍图像小目标检测算法[J]. 工程科学与技术, 2025, 57(3): 235-246.
- [11] 陈海秀, 陈子昂, 宁馨, 等. 改进 YOLOv10n 的密集行人小目标检测方法[J]. 小型微型计算机系统, 2026, 47(4): 999-1007.
- [12] 高立鹏, 周孟然, 胡锋, 等. 基于 REIW-YOLOv10n 的井下安全帽小目标检测算法[J]. 煤炭科学技术, 2025, 53(S2): 417-429.
- [13] 彭哲轩, 应展烽, 葛昊, 等. ADE-YOLO: 基于改进 YOLO11 的退化环境下小目标检测算法[J]. 计算机工程与应用, 2026, 62(1): 140-150.
- [14] 张志豪, 厉小润, 陈淑涵. 基于改进 YOLO11 的无人机航拍图像小目标检测算法[J]. 液晶与显示, 2025, 40(6): 915-930.
- [15] 黄毅, 周纯, 刘欣军, 等. 基于 YOLOv10 的无人机复杂背景下多尺度检测模型[J]. 光通信研究, 2025(5): 45-52.
- [16] Jia, Q., Yang, J., Han, S., et al. (2024) CAMLLA-YOLOv8n: Cow Behavior Recognition Based on Improved YOLOv8n. *Animals*, **14**, 3033. <https://doi.org/10.3390/ani14203033>
- [17] Khanam, R. and Hussain, M. (2024) Yolov11: An Overview of the Key Architectural Enhancements.

- [18] Chen, H., Wu, P., Wen, W. and Zeng, N. (2025) DLA-Net: A Dynamically Learnable Attention Network for Intelligent Surface Visual Inspection of Aero-Engine Blades. *IEEE Transactions on Instrumentation and Measurement*, **74**, 1-14. <https://doi.org/10.1109/tim.2025.3561440>
- [19] Du, D., Zhu, P., Wen, L., *et al.* (2019) VisDrone-DET2019: The Vision Meets Drone Object Detection in Image Challenge Results. *Proceedings of the IEEE/CVF International Conference on Computer Vision Workshops*, Seoul, 27-28 October 2019, 213-226.