

DOI:10.19651/j.cnki.emt.2519255

改进 YOLOv5s 的轻量化无人机航拍小目标识别算法*

陈文轩 杨风暴 李 蕊 王肖霞 吉琳娜

(中北大学信息与通信工程学院 太原 030051)

摘要: 无人机航拍是目前主流的物体检测技术之一,该任务中面临目标物体小、尺度变化大、复杂背景干扰等问题。如何在有限的计算资源下提升检测精度是一个重要的挑战。针对以上问题,提出了一种轻量级无人机航拍目标检测方法。首先设计了一种层次依赖感知剪枝算法减少模型的冗余计算。此外,将检测头分辨率提升至 160×160 以增强小目标检测能力,利用 GhostConv 替换网络的标准卷积块以减少计算冗余,并引入紧凑型架构 StarNet 重新设计 Neck 网络中的 C3 模块以减少特征融合过程的复杂度,增强特征表达能力。最后在 Backbone 层引入注意力机制来提高模型的特征提取能力。实验结果表明:在 VisDrone2019 数据集中,模型的 mAP_{0.5} 提升了 1.8%。同时相比于原模型,参数量下降了 50.4%,计算量降低了 35.44%。综上所述,模型满足无人机平台在小目标检测任务中对精度和轻量化的需求。

关键词: 无人机航拍;YOLOv5;小目标检测;GhostConv;StarNet;GAM 注意力机制;层次依赖剪枝

中图分类号: TN911.73 **文献标识码:** A **国家标准学科分类代码:** 520.2060

Improve the small target recognition algorithm of YOLOv5s for lightweight UAV aerial photography

Chen Wenxuan Yang Fengbao Li Bo Wang Xiaoxia Ji Linna

(College of Information and Communications Engineering, North University of China, Taiyuan 030051, China)

Abstract: UAV aerial photography is one of the mainstream object detection technologies, and this task faces problems such as small target objects, large scale changes, and complex background interference. How to improve the detection accuracy with limited computing resources is an important challenge. In order to solve the above problems, a lightweight UAV aerial target detection method was proposed. Firstly, a hierarchical dependence-aware pruning algorithm was designed to reduce the redundancy of the model. In addition, the resolution of the detection head is increased to 160×160 to enhance the detection ability of small targets, the standard convolution blocks of the network are replaced by GhostConv to reduce the computational redundancy, and the C3 module in the Neck network is redesigned by introducing the compact architecture StarNet to reduce the complexity of the feature fusion process and enhance the feature expression ability. Finally, the attention mechanism is introduced in the backbone layer to improve the feature extraction ability of the model. The experimental results show that in the VisDrone2019 dataset, the mAP_{0.5} of the model is increased by 1.8%. At the same time, compared with the original model, the number of parameters is reduced by 50.4%, and the amount of computation is reduced by 35.44%. In summary, the model satisfies the requirements of the UAV platform for accuracy and lightweight in small target detection tasks.

Keywords: UAV aerial photography; YOLOv5; small target detection; GhostConv; StarNet; GAM attention mechanism; hierarchical dependence-perceptual pruning

0 引 言

航拍目标识别算法作为计算机与无人机技术结合的前

沿研究方向,相较于传统检测手段,航拍识别算法具备高空广域视角、高分辨率及灵活机动性等优势^[1],其消除了地面监测任务中的视觉盲区,加强了大区域的信息监测分析能

收稿日期:2025-07-02

* 基金项目:山西省基础研究计划青年科学研究项目(202303021212195)、中国博士后科学基金第 76 批面上项目(2024M76303)、四川省机器人与智能系统国际联合研究中心开放研究课题重点项目(JQZN2023-002)资助

力。这些技术在灾害救援^[2]任务中可以迅速构建受灾区域全局态势,精确定位幸存人员与可利用资源;在交通监控领域^[3]能够实时掌握路网流量分布与异常事件,为交通调度提供数据支撑;在城市管理方面^[4],能够有效监测环境变化并评估城市扩张趋势。该技术为现代智能任务决策系统提供了重要支持。

计算机算力提升与模型的不断优化,航拍目标识别算法在检测精度与泛化能力方面取得了显著进步,常见的目标检测方法有SSD^[5]和YOLO系列^[6-8]的单阶段算法和Fast R-CNN^[9]的双阶段算法。为强化模型对小目标的捕捉能力与检测精度往往需要以模型复杂度与增加计算需求为代价。无人机平台受限于算力与载重等因素,无法搭载高性能的大模型设备,而许多航拍任务要求低延迟响应,例如在应急救援任务中,将数据传输地面并处理会造成信号延迟;在动态目标追踪任务中,处理延迟会直接导致目标损失;在安防监控应用中,实时预警能力是系统核心价值所在。因此将模型轻量化部署于无人机端,实现“端侧智能”,不仅大大缩短了检测响应,还能显著降低通信负担,提升任务的实时性,为各类实时性任务提供关键技术保障^[10]。

随着对无人机端实时目标识别需求的提升,轻量化算法成为研究热点,许多学者提出了专为资源受限设备设计的轻量级神经网络结构:Howard等^[11]提出了一种基于深度可分离卷积的轻量级网络MobileNets,通过引入两个全局超参数实现了延迟与精度的灵活权衡,该模型在ImageNet分类及目标检测、细粒度分类等多个视觉任务中都展现出了良好性能。Tan等^[12]系统研究了模型的深度、宽度和分辨率的协同缩放,提出了一种统一缩放三者的新方法,并通过复合系数实现高效模型扩展。实验表明,该方法在扩展MobileNet和ResNet时能显著提升性能。星操作可在不增加网络宽度的情况下,将输入映射到高维非线性特征空间,类似于核技巧。基于此,Ma等^[13]设计了高效且结构紧凑的StarNet模型,在保持低延迟的同时取得了优异性能,展现了星操作在轻量化网络设计中的巨大应用前景。为了更好的剔除模型中的冗余参数,越来越多的研究人员开始研究模型剪枝方法。Zheng等^[14]针对Transformer进行特征提取时复杂度高的问题,通过引入无参数的轻量级剪枝方法,对输入图像进行自适应采样,过滤掉不重要的单元,有效的减少了不相关的输入。许多研究人员在进行模型剪枝时发现,由于残差连接使得Hessian矩阵得特征值分布更加均匀,很难区分出哪些参数对模型性能影响小得冗余参数。针对于残差连接应该如何修剪的问题,Luo等^[15]提出通过基于KL发散的标准剪枝残差连接内部和外部的通道从而优化网络结构。结构化剪枝方面,已有多种方法被提出以加速神经网络的推理速度。然而,大多数现有的方法依赖于架构特定的手工设计分组方案,无法很好地推广到不同的网络架构。针对此问题,Fang等^[16]提出了一种通用的方法,以显式建模层之间

的依赖关系,并对耦合参数进行全面分组以进行有效的剪枝。

虽然模型的轻量化方法对于其复杂度等有所改善,但由于参数量及计算量的大幅减少,精度也会随之降低,因此要在压缩后的模型基础上进行优化与微调。吴旭红等^[17]在YOLOv7的基础上增加了一个极小目标检测层,以适应不同尺度下的物体目标,降低了小目标的漏检率,也同时增加了模型的参数量。Li等^[18]修剪了Backbone和Neck部分后,设计了一款轻量级的C3模块并且融入了注意力机制从而提高模型的特征提取能力,但修剪和注意力机制的结合可能导致模型对训练数据的依赖性增强导致其泛化能力受限。Tang等^[19]对YOLOv5进行通道剪枝后采用蒸馏学习的操作来回调精度,在大幅减少计算量的同时使平均准确率提升了4.6%,但仍依赖BN层 γ 值可能忽略通道间的高阶相关性,使得模型忽略掉与其他通道协同作用时对特征至关重要的通道。程期浩等^[20]通过引入SConv、C2f-LSK、P2检测头对YOLOv8协同改进提升精度与鲁棒性,但受计算量、数据集覆盖及底层技术缺失影响,在实时性、泛化性等方面仍需优化。严嘉旭等^[21]通过MobileSAM骨干、LSCD检测头、NWD损失的协同设计,在轻量化、小目标检测精度、泛化性上远超YOLOv8s,但存在极端算力适配差、模块灵活性低等问题,需后续从动态参数适配、模块解耦方向进一步优化。张亚军等^[22]通过C2f-DWR、FHSFPN及ShapeIoU对YOLOv8s进行改进,实现无人机航拍路面损伤检测的轻量化与准确性平衡,有效改善漏检误检问题,但在算力受限场景的进一步压缩及极端损伤检测上仍有提升空间。贾玮迪等^[23]通过dwCSP_n、SPDC等模块替换兼顾轻量化与检测精度,但未结合剪枝、量化等技术,未针对边缘设备进行部署适配。

结合上述思想,以YOLOv5s为基准网络框架,提出了一种基于轻量化无人机航拍目标检测模型,主要贡献如下:

- 1)将轻量级StarNet网络引入到YOLOv5的Neck网络中,替换小目标检测头,在压缩模型的同时保证其检测精度与速度。
- 2)对于全局网络采用Ghost卷积替换标准卷积,并且引入GAM注意力机制,提高模型的特征表达能力。
- 3)采用层次依赖感知剪枝,自动剪除以模型影响较低的神经元,从而提升模型的推理速度。

1 方 法

1.1 YOLOv5 模型

YOLOv5延续经典的单阶段检测器设计范式。目前共有5种版本,分别是YOLOv5n、YOLOv5s、YOLOv5m、YOLOv5l、YOLOv5x。每个版本的深度和宽度各不相同,由n到x依次增加,检测精度也是同理。YOLOv5由输入端(Input)、骨干网络(Backbone)、特征金字塔网络(Neck)和检测头(Head)三部分组成,结构示意图如图1所示。

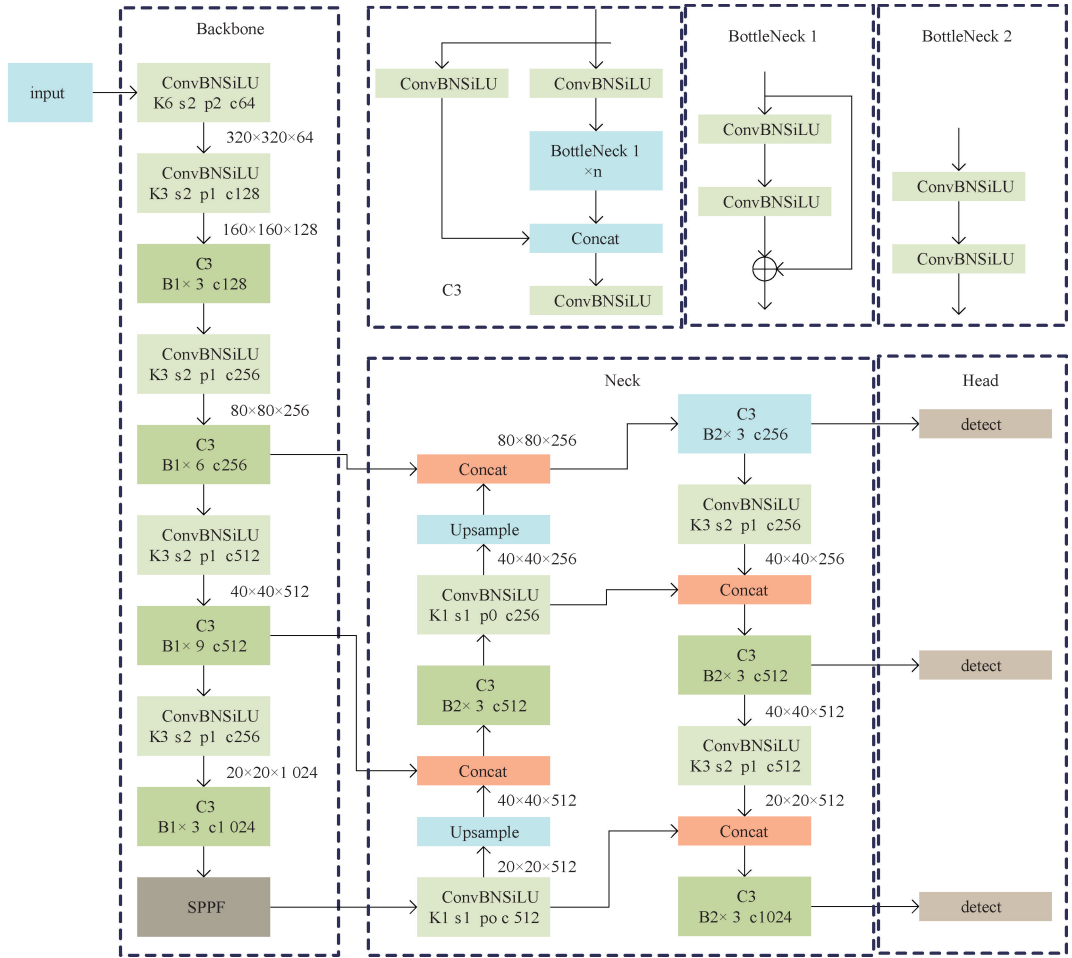


图 1 YOLOv5s 结构框架

Fig. 1 YOLOv5s structural framework

1.2 C3_Star 与多尺度检测改进

由于无人机等边缘设备的算力普遍较低,较为复杂的模型难以部署,所以首先考虑引入轻量级架构 StarNet 模型的特征融合部分来减轻模型的参数量及计算量。其次为了弥补轻量化网络造成的精度下降问题及针对于无人机航拍目标较小的特点,本文在原来模型的基础上移除了较大物体的检测头替换为更加适用于小目标的检测头,从而提升模型的精准度与鲁棒性。

高维空间可以抓捕数据之间的细微差距,提供了更加丰富的信息和表达能力。传统的深度学习通过增加网络的宽度和深度来实现高维映射,这显著增加了计算量和复杂度。Ma 等^[13]研究的 StarNet 采用了星操作(逐元素相乘)来融合不同子空间特征。星操作能够在不增加网络宽度的情况下实现特征的高维映射,能够更有效地建模特征之间地交互关系,而不仅仅是简单地将其相加。将 StarNet 引入到模型中与 C3 模块结合设计出一种新型的轻量化 C3(C3star)模块。StarNet 模块结构如图 2 所示,数据流向为:输入→归一化处理→深度可分离卷积→分叉

两路 FC→融合→FC→输出。

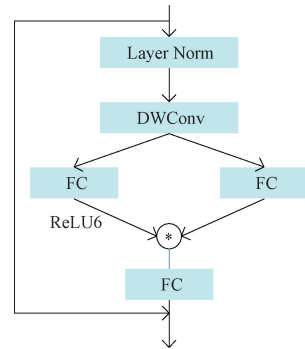


图 2 StarNet 模块

Fig. 2 StarNet module

本文将 YOLOv5 中 C3 模块的 Bottleneck 替换为 StarNet 模块,保留了原先的模型架构,只处理内部单元。并将其命名为 C3star,其结构图如图 3 所示。

1.3 Ghostconv

Ghostconv^[24]是 GhostNet 的核心模块,相对于传统卷

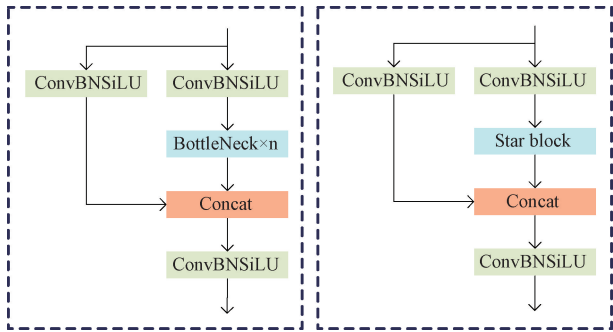


图3 C3与C3star对比图

Fig. 3 C3 vs. C3star

积直接通过密集的卷积核生成所有特征图,Ghostconv并不全部依赖普通卷积,而是将卷积操作分为两步进行。首先使用 1×1 的卷积聚合通道间的信息特征,然后采用“廉

价操作”对固有特征图处理生成新的特征图,最后将这两组特征图进行通道拼接得到最终结果。标准卷积与Ghostconv卷积的对比图如图4、5所示。

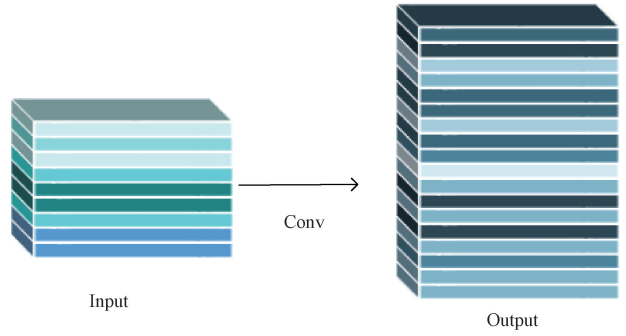


图4 标准卷积

Fig. 4 Standard convolution

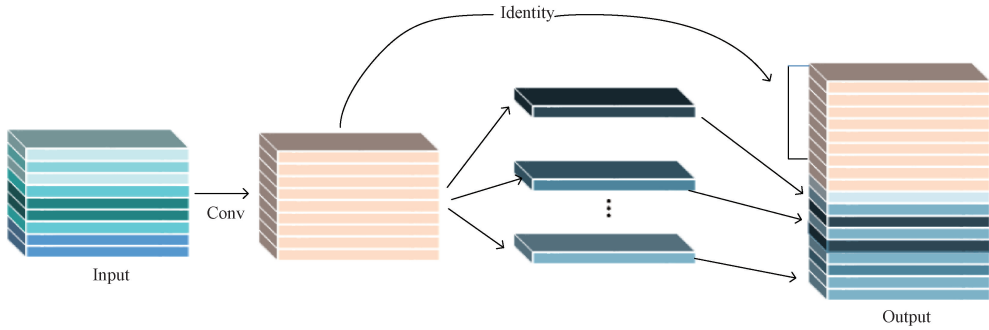


图5 Ghostconv

Fig. 5 Ghostconv

为了进一步减少模型的复杂度,本文采用Ghostconv替换原先的标准卷积块。假设该层的输入经过 n 个 $k \times k$ 大小的卷积得到输出 $Y = R^{h' \times w' \times n}$,经过Ghostconv生成 m 个固有特征图,每个特征图经过廉价操作可以生成 $t-1$ 个新的特征图,故可以得一下等式:

$$mt = n \quad (1)$$

其中, h' 为输出后的特征图高度, w' 为输出后的特征图宽度,廉价操作的卷积核大小为 $d \times d$ 。标准卷积与Ghostconv的加速比如式(2)所示。

$$Rate = \frac{cnk^2}{cnk^2 + m(t-1)d^2} = \frac{tk^2}{k^2 + \frac{t-1}{c}d^2} \approx t \quad (2)$$

当输入通道 c 足够大时, $\frac{t-1}{c}d^2$ 可忽略,标准卷积计算量是Ghostconv的 t 倍。标准卷积与Ghostconv的参数比如式(3)所示。

$$paramratio = \frac{cnk^2 + n}{cmk^2 + m + m(t-1)d^2 + m(t-1)} = \frac{t(ck^2 + 1)}{ck^2 + (t-1)d^2 + t} \quad (3)$$

由于模型Backbone开始部分的卷积块作用是提取图

像特征,Head检测头部分的卷积块作用是输出分类特征,所以不对其进行修改。

1.4 GAM注意力机制

GAM^[25]是一种结合了通道注意力和空间注意力的双重注意力机制,针对CNN特征图进行自适应特征增强,如图6所示。GAM在三个维度上并行捕捉重要特征,旨在保留信息以放大全局跨维度的交互作用,旨在解决单一维度注意力机制的局限性,通过全面考虑特征图的通道关系和空间分布,有效提升模型性能。

GAM采用顺序级联的方式对特征图处理,给定输入特征图 $X_{shape} \in R^{B \times C \times H \times W}$,GAM注意力机制计算过程如下:

通道注意力计算:

$$M_{channel} = \sigma(W_2 \cdot \delta(W_1 \cdot X_{shape})) \quad (4)$$

空间注意力计算:

$$M_{spatial} = \sigma(B_2(C_2(\delta(B_1(C_1(X')))))) \quad (5)$$

级联计算:

$$Y = (XM_{channel})M_{spatial} \quad (6)$$

其中, δ 表示ReLU激活函数, σ 表示sigmoid激活函数, B_n 表示BN层, C_n 表示卷积层。

通道注意力模块使用三维排列来在三个维度上保留

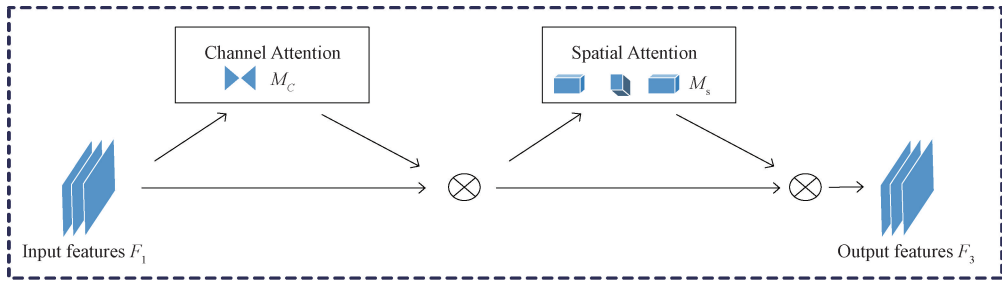


图 6 GAM 注意力机制

Fig. 6 GAM attention mechanisms

信息。然后,它用一个两层的 MLP(多层感知器)放大跨维通道,从而提高模型的特征表达能力。空间注意力子模块中,为了关注空间信息,使用两个卷积层进行空间信息融合,充分学习空间特征。

将 GAM 注意力模块加入到 Backbone 部分末端,SPFF 模块前。GAM 全局注意力模块的引入可以在不给模型添加更多参数的情况下提取到更多有用的特征,从而提升模型检测的准确率。

1.5 层次依赖感知剪枝

通道剪枝作为结构化剪枝的一种重要方法,通过移除贡献较小的卷积通道,能够直接减少模型计算量和存储需

求。与非结构化剪枝不同,通道剪枝不依赖特殊的硬件支持,移除整个通道后可立即获得计算加速效果。在目标检测任务中,通道剪枝面临一个关键挑战:如何准确评估不同通道对检测性能的贡献。传统方法通常采用单一指标(例权重 L1 范数)进行评估,这种方式难以全面反映通道的真实功能重要性。本节提出的通道重要性评估方法,结合了参数静态特性和动态激活行为,更加精准地识别模型中的关键结构。如图 7 所示,本方法的整体流程包括 4 个主要阶段:模型分析与依赖图构建、通道重要性评估、基于层级依赖的通道剪枝策略以及模型重构与微调。

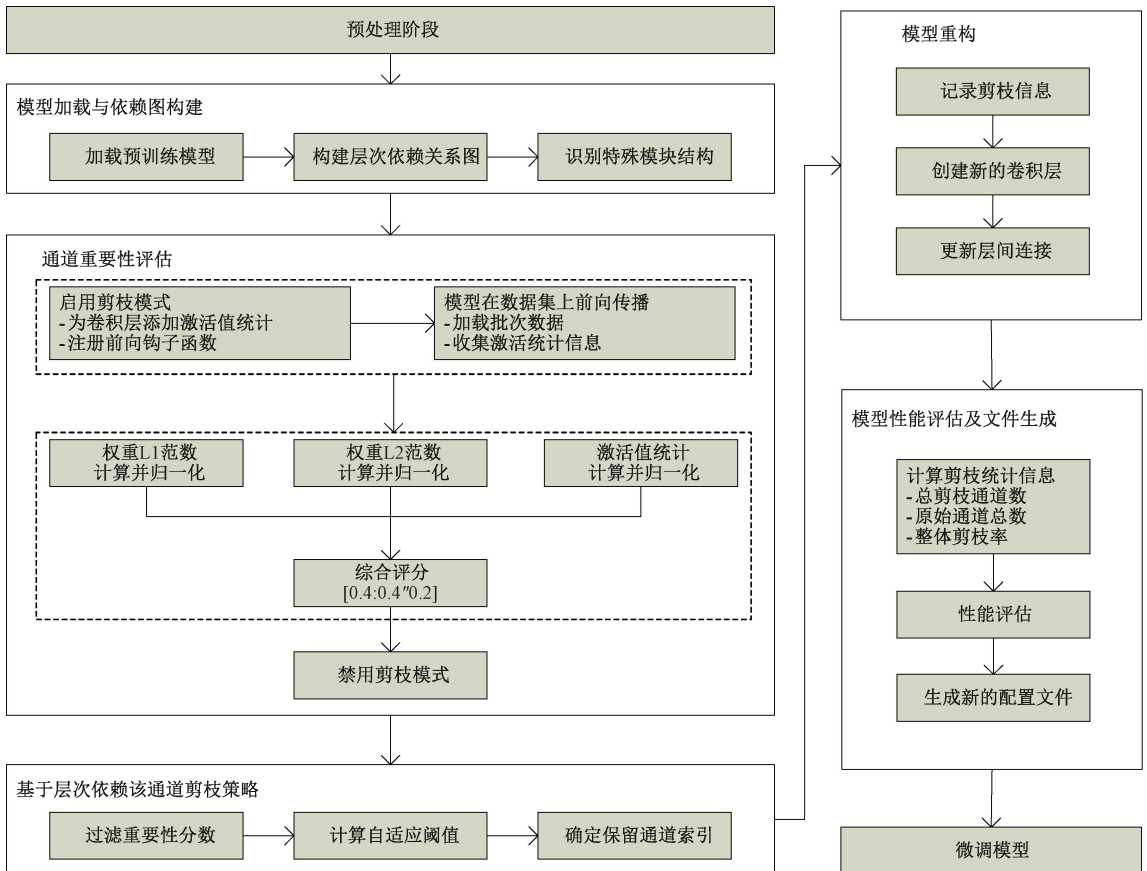


图 7 剪枝流程图

Fig. 7 Pruning flowchart

目标检测网络具有复杂的层级结构,盲目剪枝可能导致结构不匹配和信息流中断。为了精确捕捉这些复杂关系,首先对网络进行拓扑遍历,记录每一层的基本信息,包括层类型、输入输出通道数、卷积核大小等参数。然后分析模型的前向传播逻辑,建立层间的连接关系,构建网络依赖图 $G=(V,E)$,协调相互关联层的剪枝过程。其中 V 表示网络中的各层节点, E 表示层间的依赖边。

$$G=(V,E), \text{ where } V=\{v_1, v_2, \dots, v_n\} \quad (7)$$

其中, $E=\{v_i, v_j\} |$ 层 v_i 的输出层是 v_j 的输出。

在构建依赖图的过程中,还需要识别网络中的关键节点,这些节点在剪枝过程中需要特殊处理。小目标识别任务中对检测头大量剪枝会大幅影响模型的检测效果,需谨慎处理。在识别特殊模块结构方法中分离出了模型的特征提取、特征融合与检测头部分。在先前的处理中并未对骨干网络做较多的轻量化处理,造成了较多的冗余参数,故在此对模型进行分布剪枝,主要是针对于特征提取部分。在得到依赖图后收集模型的信息进行综合评估。其中静态特性由 $L1$ 、 $L2$ 范数共同反应,动态激活行为由激活值反应。

$L1$ 范数反映了卷积对输入特征的敏感程度,公式如下:

$$\|\omega\|_1 = \sum_i |\omega_i| \quad (8)$$

绝对值的大小指通道权重的绝对值总和,代表该通道对输入的总响应强度,较大的 $L1$ 范数意味着该通道在总体上对输入信号有更强的反应幅度。

$L2$ 范数反映了通道权重的能量分布,公式如下:

$$\|\omega\|_2 = \sqrt{\sum_i \omega_i^2} \quad (9)$$

能量分布高的通道往往对特定模式有较强的选择性效应,对于异常值更加敏感,能够有效区分关键特征和背景噪声。

为了捕捉模型在实际数据上的动态行为,本文设计了一种高效的激活值采集策略。该策略通过注册前向传播钩子函数,在模型推理过程中自动收集通道激活统计信息。前向钩子是神经网络训练和优化中的一种非侵入式监测技术,允许在不修改模型结构的情况下,获取并分析特定层的中间特征表示。对于卷积层 M_k 的输出特征 O_k ,通道 c 的激活强度 A_c 计算如下:

$$A_c = \frac{1}{N} \sum_{n=1}^N \left(\frac{1}{H \times W} \sum_{h=1}^H \sum_{w=1}^W |O_{k,n,c,h,w}| \right) \quad (10)$$

其中, N 为收集的数据批次总数, $|\cdot|$ 表示绝对值运算, $O_{k,n,c,h,w}$ 表示在第 n 个样本的特征图在通道 c 和空间位置 (h, w) 处的激活值。

经过多个样本的统计,计算每个通道的平均激活强度,激活值统计反应了通道在实际数据上的活跃程度。公式如下:

$$Activation = \frac{1}{N} \sum_{i=1}^N a_i \quad (11)$$

活跃程度指处理实际数据时的平均激活强度,反映了该通道在特定数据集上的实际利用率,高活跃度表明该通道频繁被激活且输出信号强度大。

将得到的评估结果归一化并利用加权融合策略整合为统一的重要性分数:

$$I_{\omega}(i) = \alpha \times I_{L1}(i) + \beta \times I_{L2}(i) + \gamma \times I_{act}(i) \quad (12)$$

其中, α 、 β 、 γ 是各维度的权重系数,基于权重范数提供了稳定的静态结构信息,应占据主导地位; $L1$ 与 $L2$ 范数具有互补性,应给予相同的权重;激活值统计提供了任务相关的动态信息,作为重要补充,所以设置为 $\alpha = \beta = 0.4$, $\gamma = 0.2$ 。

为提高激活值的效率,本研究探索了数据采样策略对评估结果的影响。实验表明,使用有限次的样本进行激活值统计已经能够获得稳定的重要性评估结果。图8展示了不同批次数量对通道重要性排序稳定性的影响。

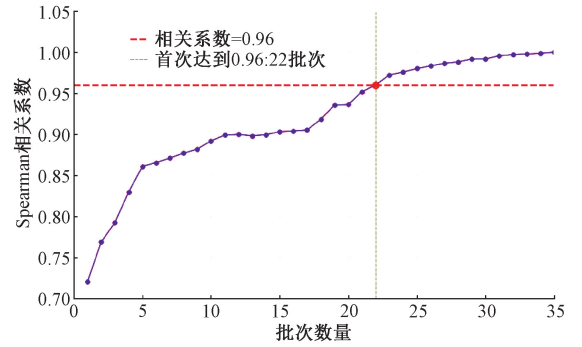


图8 各批次数量对通道重要性排序稳定性的影响
Fig.8 Effect of the number of batches on the stability of channel importance sequencing

可以看出,在22个批次后,通道重要性的排序已经趋于稳定,继续增加批次数量带来的变化很小。这种稳定性可以用 Spearman 相关系数量化:

$$\rho = 1 = \frac{6 \sum_{i=1}^n d_i^2}{n(n^2 - 1)} \quad (13)$$

其中, d_i 是第 i 个通道在两种评估方法下的排名的差异, n 是通道总数。

在层次依赖感知剪枝策略中,对于模型的剪枝率进行了自适应的阈值计算:

$$T(layer) = \begin{cases} P_{10}(I_{layer}), & \text{layer} \in \text{detect}_{layer} \\ P_{r \times 100}(I_{layer}), & \text{其他} \end{cases} \quad (14)$$

其中, $P_Q(I)$ 表示重要性分数 I 的第 Q 百分数, r 为全局剪枝率。

为了防止过度剪枝,该方法采用了安全保护机制,保证每层至少保留30%。

$$C_{keep} = \{i | I(i) > T(layer)\} \quad (15)$$

$$|C_{keep}| = \text{MAX}(|C_{keep}|, [0.3 \times C_{out}]) \quad (16)$$

其中, $I(i)$ 是索引为 i 的通道的重要性分数,

$T(layer)$ 表示当前层的剪枝阈值, C_{out} 是输出通道数。基于确定的保留通道索引,对模型的每一层进行剪枝。剪枝完成后对其进行一定程度的微调以恢复性能。

2 实 验

2.1 实验环境

本文使用的实验环境如表 1 所示。

表 1 实验环境

Table 1 Experimental environment

参数	配置
操作系统	Windows 11
CPU	Intel(R) i9-13900H
GPU	NVIDIA GeForce RTX 4060 Laptop
编程语言	Python
深度学习框架	Pytorch
CUDA 版本	CUDA Version 12.7

2.2 实验数据

本文选取了开源的 VisDrone2019 数据集^[26],该数据集是面向无人机视角下的视觉目标检测、跟踪和动作识别任务的多任务数据集。它由天津大学研究团队于 2019 年发布,旨在为无人机视频监控和其他与无人机视觉相关的研究提供数据支持。其中包括训练集 6 471 张、测试集 548 张、验证集 1 610 张。VisDrone2019 数据集主要用于支持目标检测、目标跟踪、目标行为识别等多种视觉任务,尤其是在户外复杂环境中的应用。推动了目标识别技术在该领域的应用和发展。

2.3 评估指标

本文选用的主要评估指标为 mAP、模型大小、参数量及计算量。其中 mAP 又称“均值平均精度”是结合精确率 (Precision) 和召回率 (Recall) 的综合评估指标, AP 是 PR 曲线下的面积,用于衡量模型在不同召回率下的平均精度,在得到该值后,取所有类别的 AP 平均值,即可以衡量模型在多类别任务中的整体性能。公式如下:

$$P = \frac{TP}{TP + FP} \quad (17)$$

$$R = \frac{TP}{TP + FN} \quad (18)$$

$$AP = \int_0^1 p_r dr \quad (19)$$

$$mAP = \frac{\sum_{i=1}^K AP_i}{K} \quad (20)$$

其中, TP 表示正确预测的正样本数; FP 表示错误预测的正样本数; FN 表示错误预测的负样本数; P_r 表示该类的 PR 曲线; K 表示类别。

2.4 实验分析

1) 消融实验

为了验证本文方法对模型的有效性,在 YOLOv5-7.0 的基础上进行修改,将各改进与原模型逐步融合。消融实验如表 2 所示。C-D 表示在特征融合 (Neck) 层融入轻量级模块 C3star 并添加检测头 (detect), B-G 表示在特征提取 (Backbone) 层加入全局注意力机制 GAM, A-G 表示修改整个模型除特征提取层的开始部分和检测头外的其他卷积层为 Ghost 卷积, HRP 表示层次依赖感知剪枝方法。改进 1~4 是各修改方法的代称。

表 2 消融实验

Table 2 Ablation experiments

名称	C-D	B-G	A-G	HRP	Volume/MB	Params	GFLOPS	mAP_0.5/%
baseline	—	—	—	—	13.7	7.04	15.8	32.9
改进 1	√	—	—	—	10.1	4.97	14.6	34.6
改进 2	√	√	—	—	10.23	5.03	14.6	35.1
改进 3	√	√	√	—	8.40	4.05	12.2	34.7
改进 4	√	√	√	√	7.34	3.49	10.2	34.7

从表 2 可以看出,相较于基准模型,改进 1 在经过轻量化设计与添加小目标检测头后,模型的参数量降低了 29%,计算量降低了 7.5%,同时 mAP_0.5 上升了 1.7%。改进 2 在改进 1 的基础上添加了全局注意力机制,参数量和计算量几乎不变,mAP_0.5 提高了 0.5%。为了进一步减少模型的复杂度的同时保证模型的精度,改进 3 在改进 2 的基础上添加了 Ghost 卷积,其参数量降低了 19%,计算量降低了 16.4%,mAP_0.5 少量减少。改进 4 在前者的基础上对模型进行剪枝,剔除其中的冗余参数,在 mAP_

0.5 不变的情况下,显著降低模型的复杂度。最终与 Base 模型对比,参数量降低了 50.4%,计算量降低了 35.4%,mAP_0.5 提升了 1.8%。

2) 对比试验

为了进一步验证模型的效果,本文选取了不同的目标识别模型在 VisDrone 数据集上进行训练验证。通过表 3 可知,本文提出的方法在保证大幅度减少模型参数量和计算量的同时,相对于原模型来讲精确度和召回率得到了提升。与 YOLOv5s、TPH-YOLOv5 和 YOLO-ERF 相比、

本文提出的方法在精度上分别提升了1.8%、1.3%和1.1%;参数量分别减少了3.55、1.92和2.41M;计算量分别减少了5.6、3.9和12.0。与YOLOv7-tiny、YOLOv8s和NDM-YOLOv8相比,虽然精度与其相比略低,但是参数量和计算量的减少量更多,更易在无人机平台部署。在FPS对比中本文的方法占据优势,由于在无人机航拍任务

中需要满足实时性,平台对于模型的大小和计算量有着较高要求,根据帕累托最优理论,本文提出的方法有效的平衡了无人机航拍任务中的准确性与实时性,最适用于无人机航拍任务。对于SSD和Faster R-CNN两大主流目标识别算法,本文改进后的模型不论是在轻量化角度还是在准确率角度对于无人机航拍小目标的任务都占有大幅优势。

表3 对比试验

Table 3 Comparative experiments

模型	P/%	R/%	mAP _{0.5} /%	Params/M	GFLOPS	FPS
YOLOv5s	44.7	33.1	32.9	7.04	15.8	152
YOLOv7-tiny	43.2	39	34.6	6.2	13.5	176
SSD	21.2	35.3	24.1	60.39	23.0	93
Faster R-CNN	33.5	36.4	30.8	38.6	39.7	34
YOLOv8s	49.4	38.6	39.5	11.1	28.5	160
NDM-YOLOv8 ^[20]	49.6	39.4	39.7	7.74	16.4	185
TPH-YOLOv5 ^[27]	44.9	34.7	33.4	5.41	14.1	160
YOLO-ERF ^[28]	44.7	35.1	33.6	5.9	12.5	105
Ours	44.9	35.4	34.7	3.49	10.2	190

3) 算法可视化分析

为了更加直观的评估本文提出的改进算法的检测结果,在VisDrone2019数据集中抽取了4张测试集数据进行

可视化分析,结果如图9所示。由4张数据图可清楚的看出在不同场景下,本文提出的方法对于无人机航拍小目标识别有着更加准确的识别效果及更低的漏检率与误检率。



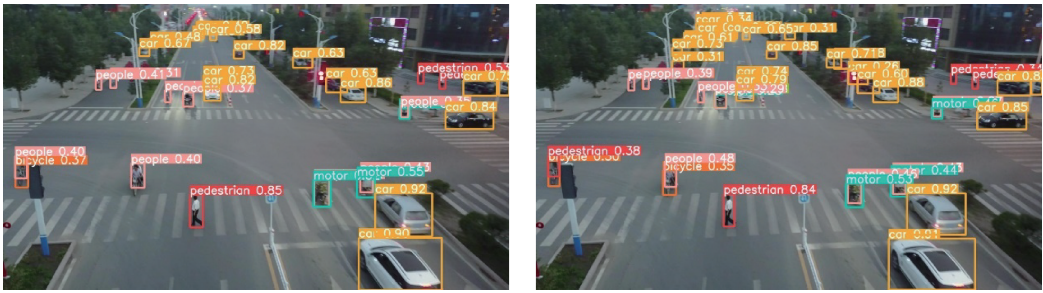


图 9 结果对比,左侧为 baseline,右侧为改进后检测结果

Fig. 9 The results are compared with the baseline on the left and the improved detection results on the right

3 结 论

针对航拍目标物体小,边缘设备部署困难等问题,本文提出了一种基于 YOLOv5s 轻量级的无人机航拍模型。利用 Ghostconv 及紧凑型网络 StarNet 大幅减少了模型的计算参数。基于层次依赖感知的剪枝算法有效抑制模型的冗余计算。修改模型的检测头以增强小目标的特征信息。针对当前目标识别算法对在小目标检测任务中因背景、密集目标导致的漏检和误检问题,引入 GAM 注意力机制,利用通道与空间的全局注意力机制增强模型对于小目标的识别能力,显著增强了模型表现。

通过在 VisDrone 数据集中的表现可知,本文提出的算法实现了参数量减少 50.4% 的同时 mAP 提升了 1.8%,对于资源有限的设备有着非常好的适应性。对于密集目标也有着较好的检测效果,具备一定的实用性。为无人机航拍场景提供了理论创新性及工程实用性的方案。

参考文献

- [1] 李琼, 考月英, 张莹, 等. 面向无人机航拍图像的目标检测研究综述[J]. 图学学报, 2024, 45(6): 1145-1164.
LI Q, KAO Y Y, ZHANG Y, et al. A review of target detection research for drone aerial images[J]. Journal of Graphic Science, 2024, 45(6): 1145-1164.
- [2] 芦艳春, 周开园, 张建杰. 无人机的发展现状及其在航空应急救援领域的应用综述[J]. 医疗卫生装备, 2023, 44(10): 108-113.
LU Y CH, ZHOU K Y, ZHANG J J. Overview of the current status of drone development and its application in aerial emergency rescue field [J]. Medical and Health Equipment, 2023, 44(10): 108-113.
- [3] VOHRA D S, GARG P K, GHOSH S. Real-time vehicle detection for traffic monitoring by applying a deep learning algorithm over images acquired from satellite and drone [J]. International Journal of Intelligent Unmanned Systems, 2022, 11(4): 441-452.

- [4] 刘德政. 未来无人机在城市规划、管理、运营中的应用探讨[J]. 未来城市设计与运营, 2023(7): 42-44.
LIU D ZH. Discussion on the application of drones in urban planning, management, and operations in the future[J]. Future City Design and Operation, 2023(7): 42-44.
- [5] LIU W, ANGUELOV D, ERHAN D, et al. SSD: Single shot MultiBox detector[C]. Computer Vision-ECCV 2016, 2016:21-37.
- [6] LI CH Y, LI L L, JIANG H L, et al. YOLOv6: A single-stage object detection framework for industrial applications[J]. ArXiv preprint arXiv: 2209.02976, 2022.
- [7] WANG C Y, BOCHKOVSKIY A, LIAO H Y M. YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors[C]. 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023: 7464-7475.
- [8] VARGHESE R, M. S. YOLOv8: A novel object detection algorithm with enhanced performance and robustness [C]. 2024 International Conference on Advances in Data Engineering and Intelligent Computing Systems, 2024: 1-6.
- [9] GIRSHICK R. Fast R-CNN [C]. 2015 IEEE International Computer Vision, 2015: 1440-1448.
- [10] 汤伟博, 方强, 李沛根, 等. 基于 RSD-YOLO 的无人机航拍图像的小目标检测[J/OL]. 计算机工程, 1-15 [2025-09-22]. <https://doi.org/10.19678/j.issn.1000-3428.0070151>.
TANG W B, FANG Q, LI P G, et al. Small target detection in drone aerial images based on RSD-YOLO[J/OL]. Computer Engineering, 1-15[2025-09-22]. <https://doi.org/10.19678/j.issn.1000-3428.0070151>.
- [11] HOWARD A G, ZHU M, CHEN B, et al. MobileNets: Efficient convolutional neural networks for mobile vision applications [J]. ArXiv preprint arXiv:1704.04861, 2017.

- [12] TAN M X, LE Q V. EfficientNet: Rethinking model scaling for convolutional neural networks[J]. ArXiv preprint arXiv:1905.11946, 2019.
- [13] MA X, DAI X Y, BAI Y, et al. Rewrite the Stars[C]. 2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2024:5694-5703.
- [14] ZHENG W F, LU S Y, YANG Y SH, et al. Lightweight transformer image feature extraction network[J]. PeerJ. Computer Science, 2024, 10:e1755.
- [15] LUO J H, WU J X. Neural network pruning with residual-connections and limited-data [C]. IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2020: 1455-1464.
- [16] FANG G F, MA X Y, SONG M L. DepGraph: Towards any structural pruning[C]. 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023: 16091-16101.
- [17] 吴旭红, 赵清华. 基于改进YOLOv7的无人机航拍图像目标检测[J]. 电光与控制, 2024, 31(2): 35-40, 111. WU X H, ZHAO Q H. Target detection in drone aerial images based on improved YOLOv7[J]. Electric Light and Control, 2024, 31(2): 35-40, 111.
- [18] LI W J, ZHANG G, CUI L. A novel lightweight traffic sign recognition model based on YOLOv5[J]. Journal of Transportation Engineering, Part A: Systems, 2023, 149 (5), DOI: 10.1061/jtepbs.teeng-7461.
- [19] TANG X, TANG K W, SUN K X, et al. Ebp-YOLOv5: Channel pruning-based lightweight YOLOv5 for QR code detection[J]. Signal, Image and Video Processing, 2024, 18: 9515-9526.
- [20] 程期浩, 陈东方, 王晓峰. 基于NDM-YOLOv8的无人机图像小目标检测[J]. 计算机技术与发展, 2024, 34(9):63-69. CHENG Q H, CHEN D F, WANG X F. Small target detection in drone images based on NDM-YOLOv8[J]. Computer Technology and Development, 2024, 34(9): 63-69.
- [21] 严嘉旭, 苏天康, 宋慧慧. 改进YOLOv8的轻量化无人机航拍目标检测. 计算机系统应用, 2025, 34(9): 151-161. YAN J X, SU T K, SONG H H. Improving the lightweight drone aerial target detection of YOLOv8[J]. Computer Systems Applications, 2025, 34(9):151-161.
- [22] 张亚军, 苗皓源, 马薇, 等. 基于YOLOv8改进的无人机航拍路面损伤检测算法[J/OL]. 电子测量技术, 2025-09-23, 1-12. ZHANG Y J, MIAO H Y, MA W, et al. An improved drone aerial imaging pavement damage detection algorithm based on YOLOv8 [J/OL]. Electronic Measurement Technology, 2025-09-23, 1-12.
- [23] 贾玮迪, 余鹏飞, 余国豪, 等. 基于改进YOLOv5m的轻量化车脸检测方法[J]. 电子测量技术, 2023, 46(12):125-133. JIA W D, YU P F, YU G H, et al. A lightweight car face detection method based on the improved YOLOv5m, [J]. Electronic Measurement Technology, 2023, 46(12):125-133.
- [24] KAI H, WANG Y H, TIAN Q, et al. GhostNet: More features from cheap operations[C]. 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019: 1577-1586.
- [25] LIU Y CH, SHAO Z R, HOFFMANN N. Global attention mechanism: Retain information to enhance channel-spatial interactions[J]. ArXiv preprint arXiv: 2112.0556, 2021.
- [26] DU D W, ZHU P F, WEN L Y, et al. VisDrone-DET2019: The vision meets drone object detection in image challenge results [C]. 2019 IEEE/CVF International Conference on Computer Vision Workshop, 2019: 213-226.
- [27] ZHOU X K, LYU SH CH, WANG X, et al. TPH-YOLOv5: Improved YOLOv5 based on transformer prediction head for object detection on drone-captured scenarios [J]. 2021 IEEE/CVF International Conference on Computer Vision Workshops (ICCVW), 2021:2788.
- [28] WANG X, HE N, HONG CH, et al. YOLO-ERF: Lightweight object detector for UAV aerial images[J]. Multimedia Systems, 2023: 29 (6): 3329-3339.

作者简介

陈文轩, 硕士研究生, 主要研究方向为目标检测、模型压缩。

E-mail: 1575979752@qq.com

杨风暴(通信作者), 教授, 博士生导师, 主要研究方向为信息融合、信息探测与处理、光电信息技术与系统。

E-mail: yfengb@163.com

李菠, 讲师, 博士, 主要研究方向目标识别。

E-mail: libo5y@nuc.edu.cn

王肖霞, 副教授, 硕士生导师, 主要研究方向为多模态目标智能识别与检测技术。

E-mail: wangxiaoxia@muc.edu.cn

吉琳娜, 副教授, 硕士生导师, 主要研究方向为多源信息智能融合与识别。

E-mail: jlnmuc@163.com