

基于YOLOv8的无人机目标检测改进算法

姜家瑞, 刘兆丰, 傅迎华

(上海理工大学 光电信息与计算机工程学院, 上海 200093)

摘要: 随着无人机技术的迅猛发展, 其在日常活动以及军事领域中所引发的安全隐患与日俱增。在此背景下, 研究以无人机为检测目标的高效检测算法迫在眉睫。针对现有目标检测算法在复杂场景中对无人机小目标的检测精度与计算量难以兼顾的问题, 提出一种基于YOLOv8的改进算法。首先, 在YOLOv8的主干网络和颈部网络引入C2fRepGhost模块, 用于降低模型的计算复杂度; 其次, 引入微小目标检测头, 并剔除大目标检测头与冗余网络, 提升模型对小型目标的识别精度; 最后, 在检测头结构中引入CBAM, 增强模型对关键特征的注意能力。在公共数据集DUT Anti-UAV上的实验结果表明, 改进算法在保持快速推理速度和低参数数量的同时, 达到了95.3%的平均精度, 相较于YOLOv8m提高了2.2%, 计算量降低了69.2 GFLOPs。

关键词: YOLOv8; Ghostnet; CBAM; 无人机检测; 注意力机制

DOI: 10.11907/rjtk.251022

中图分类号: TP391

文献标识码: A

文章编号: 1672-7800(2026)003-0179-09

扫描二维码阅读全文:



Improved Algorithm for UAV Target Detection Based on YOLOv8

JIANG Jiarui, LIU Zhaofeng, FU Yinghua

(School of Optical-Electrical and Computer Engineering, University of Shanghai for Science and Technology, Shanghai 200093, China)

Abstract: With the rapid development of drone technology, the security risks it poses in daily activities and military fields are increasing day by day. In this context, it is urgent to study efficient detection algorithms targeting drones. A improved algorithm based on YOLOv8 is proposed to address the problem of difficulty in balancing the detection accuracy and computational complexity of small unmanned aerial vehicle targets in complex scenes using existing object detection algorithms. Firstly, the C2fRepGhost module is introduced into the backbone and neck networks of YOLOv8 to reduce the computational complexity of the model; Secondly, introducing a small target detection head and eliminating large target detection heads and redundant networks to improve the recognition accuracy of the model for small targets; Finally, CBAM is introduced into the detection head structure to enhance the model's ability to pay attention to key features. The experimental results on the public dataset DUT Anti UAV show that the improved algorithm achieves an average accuracy of 95.3% while maintaining fast inference speed and low parameter count, which is 2.2% higher than YOLOv8m and reduces computational complexity by 69.2GFLOPs.

Key Words: YOLOv8; GhostNet; CBAM; UAV detection; attention mechanism

0 引言

无人机的普及应用为人们的日常生活与工作带来了诸多便利,但与此同时也引发了一系列干扰问题。凭借自身的高隐蔽性与远程操控特性,无人机沦为非法拍摄、危险物品喷洒、非法监听等违法活动的执行工具,给公共安全

和隐私保护造成了潜在威胁^[1]。在此形势下,如何借助现有传感技术实现高效的无人机目标检测,已然成为当下备受关注的热门研究课题。

根据数据来源,无人机检测技术可分为音频检测技术、图像或视频检测技术和混合检测技术3类。由于复杂的空域环境常受到云层、光照和物体遮挡等因素的影响,音频、电磁等信号在无人机检测中的应用往往面临较高的

收稿日期:2025-02-19

基金项目:上海航天科技创新基金项目(SAST2021-005)

作者简介:姜家瑞(1998-),男,上海理工大学光电信息与计算机工程学院硕士研究生,研究方向为计算机视觉;刘兆丰(1997-),男,上海理工大学光电信息与计算机工程学院硕士研究生,研究方向为计算机视觉;傅迎华(1976-),女,上海理工大学光电信息与计算机工程学院副教授、硕士生导师,研究方向为数据挖掘与计算机视觉。

误检和漏检率。基于图像的目标检测方法成本较低、不易受低空杂波干扰,在公共安全监控、智能交通系统、国防安全、人机交互和安全生产等多个领域得到广泛应用^[2]。

传统的基于图像的无人机目标检测算法利用特征点的对应关系将图像对配准,再对配准图像进行变化检测^[3]。由于该类方法依赖于特征点的提取与匹配,在不同场景或光照变化条件下的检测结果鲁棒性较差。基于深度学习的目标检测方法能够通过卷积神经网络(Convolutional Neural Network, CNN)自动学习并提取图像中的复杂特征,摆脱了传统手工特征提取的依赖,在复杂检测任务中表现出更强的适应性,主要分为两阶段目标检测方法和单阶段目标检测方法^[4]。代表性两阶段目标检测方法 Faster R-CNN 通过生成候选区域并对每个区域进行分类和边界框回归,具有较高的检测精度,但其推理速度较慢,难以满足实时检测需求。代表性单阶段目标检测方法 YOLO(You Only Look Once)将图像划分为多个网格,在每个网格上同时进行边界框回归和类别预测,从而实现端到端的目标检测,具有更快的推理速度,可满足实时检测需求^[5]。YOLO 框架主要包括骨干网络、颈部网络和检测头 3 个部分,主要通过调整主干网络的通道数和卷积层数来获得检测性能更高的版本^[6]。然而,过多的通道数和卷积层会显著增加模型的复杂度,导致检测精度与计算量难以平衡。如何在保证高精度的同时优化计算开销仍是一大挑战。例如,文献[7]向 YOLOv8 中添加小目标检测头并采用 AFPN(Asymptotic Feature Pyramid)替换颈部网络结构,取得了较高的检测精度,但计算复杂度与参数数量的增加不利于实际应用。文献[8]基于 YOLOv10 提出一种小目标检测算法,通过多尺度特征聚合扩散金字塔网络,在主干网络中对不同的特征图进行尺寸变化,再与颈部网络的对应特征图融合,以增强模型的多尺度表达能力。尽管通过调整特征融合策略可以在保持计算量的同时有效提高多尺度目标的检测性能,但由于主干网络在提取特征尺寸过程中有可能导致小目标信息丢失,对于小目标的检测效果提升有限。

针对无人机检测任务中小目标检测困难、检测精度与模型复杂度难以平衡的问题,本文通过向 YOLOv8n 中引入小目标检测头、C2fRepGhost 和注意力机制来提升网络性能。首先,在引入小目标检测头的同时删除大检测头和冗余网络;其次,C2fRepGhost 结合 GhostNet 与 RepVGG 结构重新构建了 YOLOv8 的核心模块 C2F(Faster Implementation of CSP Bottleneck with 2 Convolutions),利用 GhostNet 简化特征提取过程,并通过重参数化技术解耦 GhostNet 的训练推理过程,使模型在保持较高检测精度的同时减少计算量和推理延迟^[9-10];最后,在检测头前引入 CBAM(Convolutional Block Attention Module)注意力机制来优化特征权重分配,增强模型区分复杂背景下背景与目标的能力。所提方法在 Anti-UAV 数据集上的检测精度显著优于 YOLO 系

列最新版本 YOLOv11 及其他相关模型^[11]。

1 YOLOv8 算法

YOLOv8 算法通过骨干网络、颈部网络和检测头 3 个部分实现端到端的目标检测。首先,骨干网络通过卷积操作从输入图像中提取多层次的特征图,提炼目标的关键特征和语义信息;其次,颈部网络基于特征金字塔(Feature Pyramid Network, FPN)与路径聚合(Path Aggregation Network, PAN)实现不同尺度的特征融合,增强模型对多尺度目标的感知能力;最后,检测头负责为每个物体生成类别、位置和置信度,并通过非最大抑制(Non-Maximum Suppression, NMS)算法去除冗余框,输出最终检测结果。YOLOv8 结构如图 1 所示。

YOLOv8 的骨干网络采用 CSPDarknet53 的优化版本,相较 CSPDarknet53 框架新增了特征提取模块 C2F 与特征融合模块 SPPF(Spatial Pyramid Pooling Fast)^[12-13]。C2F 模块通过一个卷积块提取全局特征,将提取到的粗粒度特征经过多个 Bottleneck 模块处理,以细化特征的多层次表达,再利用恒等链接将各分支提取到的部分细化特征拼接,获取语义信息更为丰富的特征图。SPPF 模块取代了传统的空间金字塔池化(Spatial Pyramid Pooling, SPP),通过串行多个池化操作来快速获取不同大小的池化窗口,替代并行的不同尺度的池化操作,得到具有不同感受野的特征图,最后对这些特征图进行拼接后输出。以上操作同时兼顾了图像的局部特征和全局特征,提高了目标检测的准确性和鲁棒性。

颈部结构采用与 PANet 相同的结构,通过构建自顶向下和自底向上的路径,采用路径聚合来融合不同尺度的特征^[14]。颈部网络将经过 SPPF 处理后的特征图通过反卷积上采样至与主干网络相同的尺寸,随后对主干网络与颈部网络的特征图进行拼接,实现不同尺度特征的卷积信息共享,以增强特征的综合表达能力,提升特征多样性。

检测头采用分离式架构,针对目标分类和边界框回归分别设计了独立的处理流程,避免两个任务相互干扰,有助于提升不同任务的精度和速度。目标分类任务和边界框回归任务分别应用不同的损失函数:目标分类采用二元交叉熵损失(Binary Cross Entropy Loss);边界框回归采用分布焦点损失(Distribution Focal Loss)和 CIoU 损失。

2 YOLOv8 算法的改进策略

2.1 C2fRepGhost

C2fRepGhost 基于改进 YOLOv8 的 C2F 构建而成,通过将重参数化后的 GhostModel——RepGhost 集成至 Bottleneck,根据步长不同构建两种新型的 RepGhostBottleneck 模块,结构如图 2(a)(b)所示。将 RepGhostBottleneck 模块

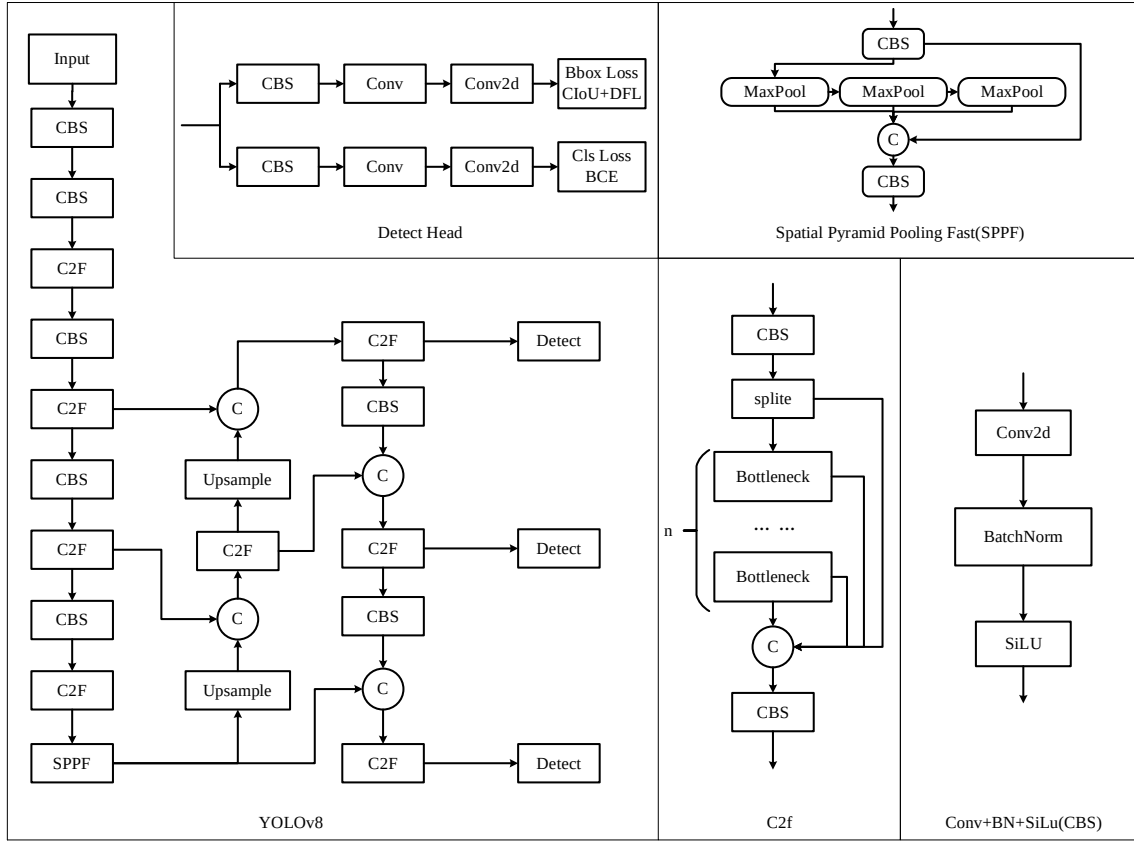


Fig. 1 YOLOv8 architecture

图 1 YOLOv8 结构

应用于 C2F 模块的瓶颈部分,构建全新的 C2fRepGhost 结构,具体如图 2(c)所示。C2fRepGhost 被用于替换 YOLOv8 的 C2f 模块,解耦了模型的训练与推理过程,降低了计算量与推理延迟,保持了与原始模型相当的性能。

GhostModel 是 GhostNet 采用的一种轻量化模块,与传统模块相比,其使用较少的卷积核提取输入特征,结构如图 3 所示。输入特征经过 1×1 卷积获得较少的特征图,再使用 5×5 深度卷积生成 Ghost 特征来补充通道数损失特征,通过拼接操作生成与输入相同的特征图。GhostModel 通过合并原始特征与 Ghost 特征来维持较大的通道数,从而降低模型复杂度。

公式(1)和(2)量化了相同条件下 Ghost 卷积与传统卷积的浮点运算次数(Floating Point Operations per Second, FLOPs)比率 r_s 和参数量比率 r_c :

$$r_s = \frac{h \times w \times c \times H \times W \times n}{\left(\frac{n}{s} \times H \times W \times k^2 \times c\right) + ((s-1) \times \frac{n}{s} \times H \times W \times d^2)} \quad (1)$$

$$= \frac{c \times k^2}{\frac{1}{s} \times c \times k^2 + \frac{s-1}{s} \times d^2} \approx s$$

$$r_c = \frac{n \times c \times k^2}{\frac{n}{s} \times c \times k^2 + (s-1) \times \frac{n}{s} \times d^2} = \frac{s \times c}{s \times c - 1} \approx s \quad (2)$$

式中: h 、 w 和 c 分别表示输入特征图的高度、宽度和通道数; n 表示卷积数量, k 表示卷积核的维度; d 表示线性变

换核的大小, s 表示参数压缩因子。可以看出,参数量与计算量的减少程度与压缩因子 s 成正比。这意味着通过深度卷积生成的 Ghost 特征越多,模型的轻量化效果越显著。

特征拼接需要将原始特征复制到内存,会增加在硬件设备上的 IO 开销,影响模型的推理速度。为避免原始特征图的内存复制,本文利用重参数技术,在推理时将恒等连接与 Ghost 卷积通过数学上的等价变换在权重空间中融合为一个卷积,省去内存复制的操作,缓解推理速度的下降程度。重参数技术利用卷积的可加性实现推理阶段的卷积融合。在训练阶段,假设有两个并行的卷积分支独立地学习图像特征。根据卷积的可加性,可将这两个分支学到的卷积核参数 $F^{(1)}$ 、 $F^{(2)}$ 以及偏置项 $b^{(1)}$ 、 $b^{(2)}$ 逐元素相加,创建一个新的卷积核 F' 与新的偏置项 b' ,使其能得到与原两个分支相加相同的输出结果^[15]。表示为:

$$F' = F^{(1)} + F^{(2)} \quad (3)$$

$$b' = b^{(1)} + b^{(2)} \quad (4)$$

在推理阶段,直接使用新卷积核计算,将多个卷积层学到的信息汇总到一个输出,不需要分别计算两个分支,减少了计算量,隐式地实现了特征重用。如图 4 所示,为了实现 GhostModel 的重参数化,首先需要将特征拼接操作替换为相加操作;其次,将恒等链接转换成与深度卷积相同大小的卷积核;最后在推理阶段将其学习到的参数相加,合并成一个卷积。

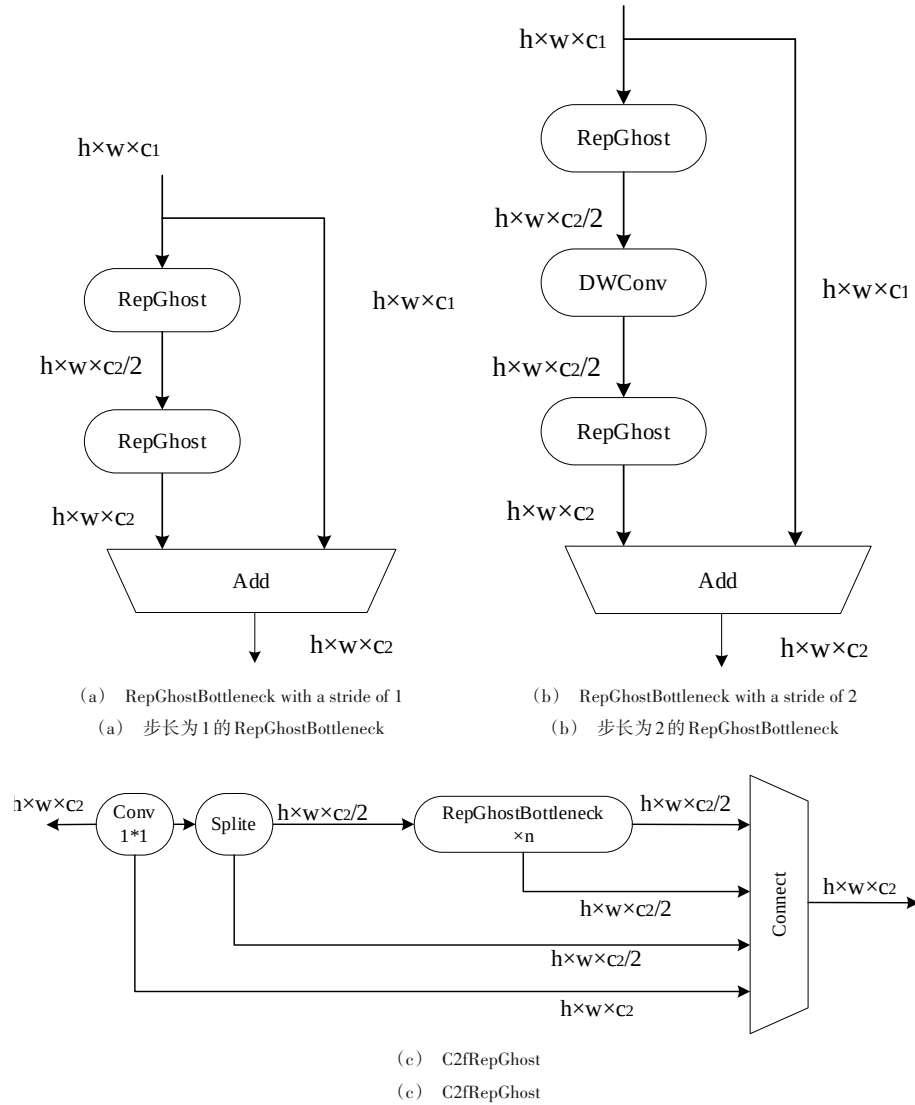


Fig. 2 C2fRepGhost construction process

图2 C2fRepGhost构建过程

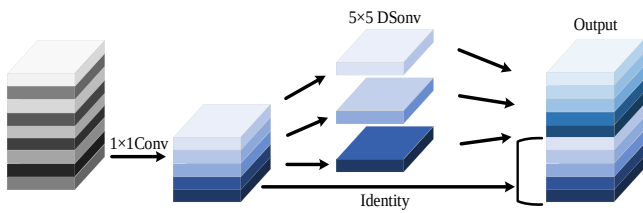


Fig. 3 GhostModel structure

图3 GhostModel结构

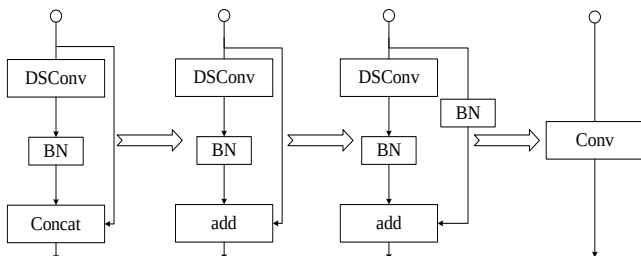


Fig. 4 GhostModule reparameterization schematic

图4 GhostModule重参数化示意图

具体来说,为了将恒等连接转换为与深度卷积相同大小的卷积核,首先采用一个BN(Batch Normalization)层替换恒等链接,BN层对于原始特征的操作等价于一个 1×1 的卷积核,而 1×1 的卷积核又可以通过零填充(Padding with Zeros)来构成一个与深度卷积相同大小的卷积核。具体如图5所示。

重参数化串联的 3×3 卷积与BN的过程可用公式表达为:

$$Output = (Input \otimes F - \mu_j) \frac{\gamma_j}{\sigma_j} + \beta_j \quad (5)$$

$$F' = \frac{\gamma_j}{\sigma_j} F, b' = -\frac{\mu_j \gamma_j}{\sigma_j} + \beta_j \quad (6)$$

式中: j 为对应通道编号; μ_j 和 σ_j 分别为累积信道均值和标准差; γ_j 和 β_j 分别为学习到的缩放因子和偏置项; F' 与 b' 分别为融合后的卷积核与偏置。

最后,将串联的传统卷积和BN融合后的卷积,与恒等

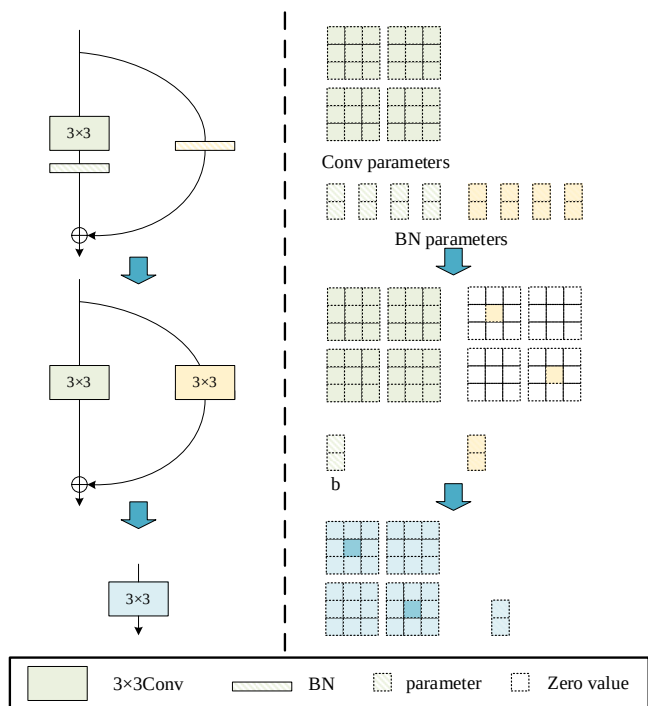


Fig. 5 Schematic diagram of identity shortcut reparameterization

图5 恒等链接重参数化示意图

连接转换的卷积相加,构建RepGhost。

2.2 微小目标检测头

原始YOLOv8的检测头结构如图6所示。原始YOLOv8有大、中、小3个不同的检测层,分别通过处理 20×20 、 40×40 、 80×80 的特征图来检测不同尺度的目标。在主干网络与颈部网络,YOLOv8通过FPN与PAN来整合小目标多层次的特征信息,获得融合浅层位置信息与深层语义信息的特征图,以增强检测头对小目标的感知能力。然而,主干网络在 160×160 特征提取至 80×80 特征图的过程中便会丢失微小目标的特征信息,导致颈部网络在聚合主干网络 80×80 特征图时也无法获取目标的位置信息,容易造成微小目标的漏检。

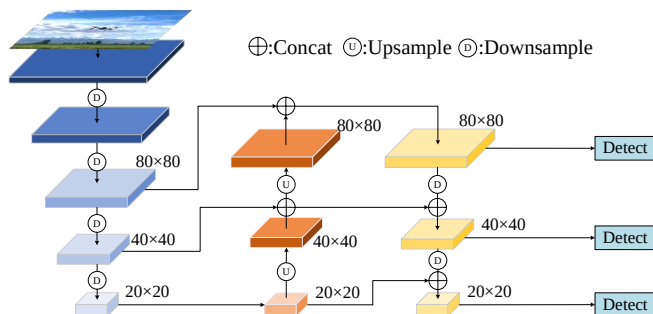


Fig. 6 YOLOv8 detection head structure

图6 YOLOv8检测头结构

本文在原始YOLOv8的检测头结构中加入新的微小目标检测头,并剔除了原始的大目标检测头与冗余网络,在提高微小目标检测精度的同时缓解了计算量的增加。改进的检测头结构如图7所示。具体来说,通过增加模型颈部网络的上采样操作生成 160×160 的高分辨率特征图,采

用原始路径聚合策略,与主干网络 160×160 的浅层位置信息特征图融合,在融合的特征图后添加一个小目标检测头来提高模型对微小目标的感知能力。删除原始YOLOv8的大检测头预期对应的特征图生成网络与路径聚合网络,提高模型的检测效率。

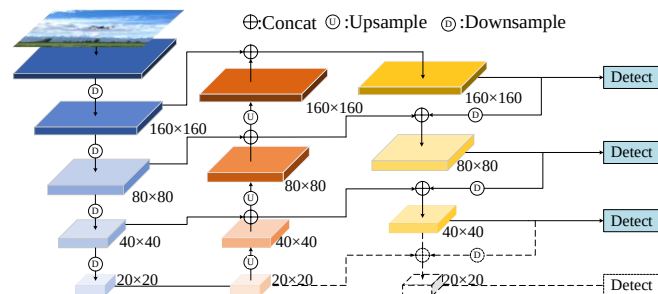


Fig. 7 Improved detection head structure

图7 改进检测头结构

2.3 注意力机制

注意力机制是计算机视觉领域提高网络模型对感兴趣目标关注程度的一种技术。该机制可以模拟人类视觉系统的注意力过程,通过引入权重来指导网络模型对输入图像的不同区域进行加权处理。在目标检测任务中,注意力机制主要分为空间注意力机制与通道注意力机制,分别通过对空间或通道进行调节,使模型更加准确地关注目标并抑制不相关信息。改进模型采用的CBAM是一种基于CNN设计的用于增强特征表达能力的注意力机制^[16]。结构如图8所示。CBAM将空间注意力与通道注意力串联起来,实现对输入特征图的精细化调整。通道注意力单元首先对特征通道执行全局最大和平均池化操作;其次通过一个共享的多层感知机进行特征转换;最后通过Sigmoid函数进行激活以获得通道注意力图。该步骤对通道间的特征重要性进行编码,在强化具有信息量的特征表示的同时削弱了不重要的特征响应。空间注意力则关注特征图在空间维度上的信息,通过全局最大和平均池化操作捕获空间特征的全局依赖关系,帮助模型集中关注图像中具有显著性的空间区域,进一步优化特征的空间分布。CBAM通过这两个注意力的协同作用实现对特征图的动态调整,增强模型对图像关键信息的捕捉能力,并抑制无关信息。

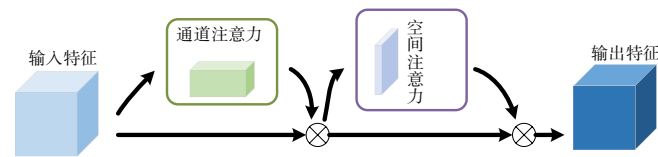


Fig. 8 Convolutional block attention module structure

图8 CBAM结构

3 实验方法与结果分析

3.1 实验环境与参数设置

所有实验均在处理器为Intel(R) Core(TM) i7-

9750H, 显存为 24 GB 的 NVIDIA GeForce RTX 4090 的 Windows 系统下进行。在训练阶段通过 Adam(随机梯度下降)和 YOLOv8 网络中 AdamW 的默认参数进行优化, 学习率为 0.01, 批次大小为 32, 输入图片尺寸设置为 640×640, 迭代次数为 200 次。实验采用有助于防止模型过拟合的早停机制, 当模型在验证集上的精度在 15 个 epoch 间没有提升时则停止训练^[17]。

3.2 数据集

采用由大连理工大学开发的开源无人机数据集 DUT Anti-UAV 进行模型训练与测试, 具体为 5 243 张图像用于模型训练, 2 621 张图像用于验证, 2 245 张图像用于测试^[18]。为评估模型的泛化能力, 使用由西湖大学构建的 Det-Fly 公共数据集进行实验^[19]。该数据集包含 13 000 张经过标注的无人机图像, 覆盖了不同角度, 为模型的泛化性测试提供了丰富样本。表 1 为两个数据集的无人机尺寸分布。

Table 1 Size distribution of drones in two datasets

表 1 两个数据集的无人机尺寸分布

数据集	小目标	中等目标	大目标
Anti-UAV	8 223	431	1 454
Det-Fly	12 731	417	122

3.3 评估指标

采用准确率(Precision, P)、召回率(Recall, R)、平均准确率(mean Average Precision, mAP)评价模型性能。计算公式为:

$$P = \frac{TP}{TP + FP} \quad (7)$$

$$R = \frac{TP}{TP + FN} \quad (8)$$

Table 2 Ablation study of lightweight strategies at different positions in the model

表 2 轻量化策略在模型不同位置的消融实验

基线模型	改进策略	改进位置	P/%	R/%	mAP/%	Params/Million	FLOPs/Giga	InfTime/ms	RepInfTime/ms
YOLOv8n	C2FRepGhost	BackBone	92.7	84.9	90.8	2.624	7.1	7.1	6.3
		Neck	93.8	81.5	89.4	2.651	7.1	7.3	5.9
		All	91.2	84.3	90.1	2.206	6.3	7.6	6.3
	C3Ghost	All	91.9	83.0	89.6	2.007	5.6	10.7	-
	ShuffleNetV2	BackBone	89.8	80.6	86.7	1.849	5.1	5.8	-

由表 2 数据可知, 采用 C3Ghost 和 ShuffleNetV2 改进策略的 YOLOv8n 在 FLOPs 上表现更优, 但 C3Ghost 的推理时间增加至 10.7 ms, ShuffleNetV2 的精度下降至 86.7%^[20-21]。C3Ghost 对 GhostNet 的 C3 模块进行了改造。由于 C3 模块没有采用重参数化技术, 且与 C2F 模块相比, C3 包含更多的冗余计算, 导致其推理时间大幅增加。使用 ShuffleNetV2 替换 YOLOv8n 主干的模型精度显著下降, 这主要是由于 ShuffleNetV2 通过通道重排来减少每层卷积的通道数, 降低了模型的特征表达能力。在复杂背景下的小目标检测任务中, 细节信息的丢失使得精度大幅降低。

为进一步验证各个改进模块对模型性能的影响, 通过组合不同改进策略在 Anti-UAV 数据集上进行 8 组消融实

$$mAP = \frac{\sum_{i=1}^{classes} AP_i}{classes} = \frac{\sum_{i=1}^{classes} \int_0^1 P(R) d(R)}{classes} \quad (9)$$

式中: TP、FP、FN 分别表示真阳性、假阳性、假阴性, 其中真阳性表示正确地检测到目标, 假阳性表示错误地将非目标标记为目标, 假阴性表示错过了真实的目标; classes 表示类别数量, 由于无人机只有一个类别, 因此 classes=1。

在计算量方面, 本文采用的指标主要包括参数量(Params)、FLOPs(Floating Point Operations)、inf 与 Repinf。其中, Params 表示模型中所有可训练参数的总数, 是衡量模型复杂度的一个重要指标; FLOPs 表示模型在前向传播过程中所需的浮点运算次数, 是衡量模型计算量的一个指标; Inf 与 RepInf 分别表示模型处理一张图像所需要的推理时间和模型重参数化后所需要的推理时间, 是衡量模型推理速度的指标。

3.4 消融实验

为验证所提改进模块对模型检测性能的影响, 进行以下消融实验。首先, 设计 C2fRepGhost 模块的消融实验, 分析其在不同位置的替换对模型精度、参数量、计算量和推理时间的影响, 并与几种常用的轻量化改进策略进行比较。实验结果如表 2 所示。可以看出, 全量替换 C2fRepGhost 模块后的模型计算量相较基线模型减少了 1.8 GFLOPs, 参数量下降了 0.8 M, 且引入重参数化技术有效降低了推理延迟。相较仅替换主干网络的模型, 全量替换模型的准确率与召回率分别下降了 1.5% 与 0.6%, 但模型计算量更低, 为不同的检测需求提供了不同的替换策略。由于本研究的目标为实现高效的无人机检测, 后续的 C2fRepGhost 改进均采用全量替换策略。

验。结果如表 3 所示, 其中“√”表示在基线模型中实现了该改进策略。可以看出, 模型在引入小目标检测头后, mAP 显著提高至 94.7%, FLOPs 增加至 11.7 GFLOPs。注意力机制的引入进一步提升了模型在精度上的表现, 提高了 0.2%, 表明小目标检测头协同注意力机制通过在更广泛的特征图中自适应地关注图像中的关键信息, 有效提升了特征的表达能力, 帮助模型更好地捕获目标。在此基础上引入 C2fRepGhost 的模型通过 Ghost 卷积与重参数化缓解计算量的增加, 将计算量与推理速度分别控制在 9.5 GFLOPs 与 7.7 ms, 在计算量、推理时间与精度之间达到了较好的平衡。

如图 9 所示, 改进模型的 Precision-Recall 曲线整体位

Table 3 Ablation study on Anti-UAV dataset
表 3 Anti-UAV 数据集上的消融实验

基线模型	C2FRepGhost	小目标检测头	注意力机制	mAP/%	FLOPs/ Giga	Params/ Million	InfTime/ ms
YOLOv8n				89.7	8.1	3.006	5.7
	✓			90.1	6.3	2.221	6.3
		✓		94.7	11.7	2.047	7.0
			✓	90.8	8.2	3.093	6.3
	✓	✓		94.6	10.3	2.228	7.8
	✓		✓	89.5	6.3	2.308	7.2
		✓	✓	94.8	11.5	2.032	6.4
	✓	✓	✓	95.3	9.5	1.487	7.7

于 YOLOv8n 之上。在低召回率区间,改进模型能够维持更高的精确率。进一步分析 mAP 曲线可知,在宽松的 mAP@0.5 和严格的 mAP@0.95 下,改进模型的 mAP 值均高于 YOLOv8n,表明改进模型对小目标的粗定位更为敏感,且边界框预测精度更高。此外,改进模型的训练收敛速度更快,展现出高效的特征提取与利用能力。

综上所述,改进模型采用的 C2FRepGhost、小目标检测头与注意力机制能够在不显著增加计算复杂度的情况下

提高检测精度。各模块通过各自独立的优化策略增强了模型在复杂场景下的表现。

3.5 Anti-UAV 数据集实验结果

在 Anti-UAV 数据集上将本文模型与 YOLOv8m、YOLOv9c^[22]、YOLOv10s^[23]、YOLOv11、RTDETR^[24]、YOLOv3-Tiny^[25]等多种模型进行比较。结果如表 4 所示。可以看出,本文模型相较其他目标检测算法展现出显著优势。相较于基线模型,本文模型的平均准确率和召回率分别提升了 5.6% 和 7.3%,在小目标检测中的平均准确率最高,达到 51.9%,而 YOLOv3-Tiny 最低,仅为 22.9%。这主要由于 YOLOv3-Tiny 仅配置了两个检测头,限制了其小目标的检测能力。在计算量上,本文模型比基线模型增加了 1.4 GFLOPs,但与更大的模型,如 YOLOv8m 和 YOLOv9c 等相比,仍表现出较强的竞争力。YOLOv8m 和 YOLOv9c 的计算量分别高达 78.7 GFLOPs 和 102.3 GFLOPs,平均准确率分别为 93.1% 和 93.4%。本文模型的计算量为 9.5 GFLOPs,平均准确率为 95.3%,推理时间与基线模型相当。本文模型在提升检测精度的同时保持了更低的计算开销,维持了精度与计算效率之间的平衡。

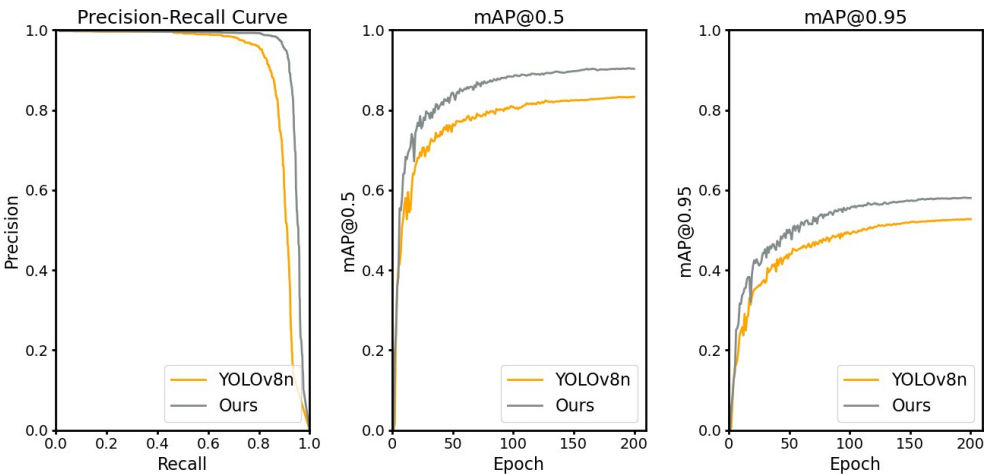


Fig. 9 Precision-Recall curves and mAP convergence graphs

图 9 Precision-Recall 曲线与 mAP 收敛图

Table 4 Detection results of different models on Anti-UAV dataset
表 4 不同模型在 Anti-UAV 数据集上的检测结果

模型	P/%	R/%	mAP@0.5/%	mAP@0.5- 0.95/%	mAPs/%	FLOPs/Giga	Params/Million	InfTime/ms
YOLOv3-Tiny	91.2	63.4	72.2	47.7	17.4	18.9	12.128	3.2
YOLOv5s	92.9	86.3	91.5	62.3	44.4	23.8	9.112	6.7
YOLOv6	92.4	83.2	89.7	59.5	40.5	11.8	4.234	5.8
YOLOv8n	94.0	81.8	89.7	60.0	42.7	8.1	3.006	5.6
YOLOv8m	94.8	87.0	93.1	64.9	46.0	78.7	25.840	8.7
YOLOv9s	93.6	86.0	91.3	63.1	44.5	26.7	7.167	14.1
YOLOv9c	94.9	87.9	93.4	66.5	47.9	102.3	25.320	12.9
YOLOv10n	91.7	81.4	89.4	59.5	42.4	8.2	2.695	7.4
YOLOv10s	94.2	84.8	92.3	63.1	44.9	24.4	8.036	8.2
YOLOv11	92.8	83.3	89.6	59.8	42.4	6.3	2.582	7.1
RTDETR	92.6	91.8	93.5	62.2	50.0	125.6	41.937	15.8
本文模型	96.6	89.1	95.3	64.0	51.9	9.5	1.487	7.7

3.6 Det-Fly 数据集实验结果

不同模型在 Det-Fly 数据集上的实验结果如表 5 所示。可以看出,本文模型在未见数据集上的各项性能指标均优于其他对照模型,具有出色的泛化性能。具体而言,本文模型的 mAP@0.5 达到 44.2%,较次优模型 YOLOv9c 提升了 2.6%,较基准模型 YOLOv8n 提升了 3.8%。本文模型的小

目标检测指标 mAPs 达到 6.6,为 YOLOv8m(4.1)的 1.6 倍,证实了其对微小目标的特征提取与泛化能力。值得注意的是,YOLOv10s 模型表现出明显的泛化能力缺陷,mAP@0.5 为 38.4,低于轻量化版本 YOLOv10n,这可能与模型复杂度过高导致的过拟合现象有关。

Table 5 Detection results of different models on Det-Fly dataset

表 5 不同模型在 Det-Fly 数据集上的检测结果

模型	P/%	R/%	mAP@0.5/%	mAP@0.5-0.95/%	mAPs/%	FLOPs/Giga	Params/Million	InfTime/ms
YOLOv3-Tiny	56.8	23.4	25.9	11.4	1.2	18.9	12.128	3.6
YOLOv5s	71.0	34.8	39.9	18.0	2.9	23.8	9.112	7.1
YOLOv6	65.2	37.3	40.7	17.9	3.2	11.8	4.234	6.0
YOLOv8n	70.3	35.5	40.4	18.3	2.9	8.1	3.006	6.1
YOLOv8m	65.1	38.0	41.5	18.7	4.1	78.7	25.840	9.0
YOLOv9c	71.6	35.5	41.6	19.3	3.4	102.3	25.320	13.6
YOLOv10n	70.0	37.5	41.4	17.5	3.3	8.2	2.695	7.5
YOLOv10s	62.3	37.2	38.4	16.5	3.2	24.4	8.036	11.0
YOLOv11	67.1	35.6	39.7	17.2	3.0	6.3	2.582	7.9
RTDETR	60.5	40.6	40.2	18.4	4.5	125.6	41.937	15.8
本文模型	79.4	40.5	44.2	19.3	6.6	9.5	1.487	7.8

3.7 可视化实验结果

图 10 和图 11 分别展示了 YOLOv8n 与本文模型在不同背景和不同尺度无人机图像上的检测效果。可以看出,在复杂背景图像中,YOLOv8n 能够检测到较大尺度的目标,但检测框的准确性较低,存在重复检测和错误检测的现象。相比之下,本文模型的检测置信度更高,有效减少了错误检测的情况。对于微小无人机目标图像,YOLOv8n 的检测性能较差,出现了漏检现象,而本文模型则能在复杂背景中准确检测出微小目标。



Fig. 10 Detection results of YOLOv8n

图 10 YOLOv8n 检测结果



Fig. 11 Detection results of the proposed model

图 11 本文模型检测结果

在 DUT Anti-UAV 和 Det-Fly 数据集上的实验结果表明,本文模型在保持较低参数量与计算量的同时实现了超越大多数模型的检测精度与泛化性。

然而,本文模型在处理新数据集时的检测性能仍面临较大挑战,且在处理遮挡情况时存在目标丢失的风险。未来研究将集中在跟踪算法与检测算法的集成,以及构建更加多样化的无人机数据集方面,以实现更加鲁棒的无人机实时检测效果。

参考文献:

- [1] ZHAO J, ZHANG J, LI D, et al. Vision-based anti-UAV detection and tracking [J]. IEEE Transactions on Intelligent Transportation Systems, 2022, 23(12): 25323-25334.
- [2] GARCIA A J, LEE J M, KIM D S. Anti-drone system: a visual-based drone detection using neural networks [C]//2020 International Conference on Information and Communication Technology Convergence, 2020: 559-561.
- [3] CHIPER F L, MARTIAN A, VLADEANU C, et al. Drone detection and defense systems: survey and a software-defined radio-based solution [J].

4 结语

本文提出一种基于 YOLOv8n 的无人机检测模型。该模型在调整检测头策略的同时优化颈部网络,提供更广泛的特征表达,提高了对小目标的检测精度;同时通过在检测头前引入注意力机制进一步提升模型对目标的感知能力,通过重参数化与轻量化平衡模型的计算量与参数量。

- Sensors, 2022, 22(4): 1453.
- [4] HAO Z X, WANG Q. Enhanced algorithm for small target detection in UAV aerial images based on YOLO-v7[J]. Software Guide, 2024, 23(1): 167-172.
郝紫霄,王琦. 基于YOLO-v7的无人机航拍图像小目标检测改进算法[J]. 软件导刊, 2024, 23(1): 167-172.
- [5] VIJAYAKUMAR A, VAIRAVASUNDARAM S. YOLO-based object detection models: a review and its applications[J]. Multimedia Tools and Applications, 2024, 83: 83535-83574.
- [6] ZHAI X, HUANG Z, LI T, et al. YOLO-Drone: an optimized YOLOv8 network for tiny UAV object detection[J]. Electronics, 2023, 12(17): 3664.
- [7] ZHANG L G, YUAN Y L, JIN M, et al. Research on UAV target detection algorithm based on improved YOLOv8n[J]. Journal of Metrology, 2024, 45(10): 1487-1493.
张立国,袁煜琳,金梅,等. 基于改进YOLOv8n的无人机目标检测算法研究[J]. 计量学报, 2024, 45(10): 1487-1493.
- [8] JIANG W Y, WANG L, YAO Y P, et al. Small target detection algorithm based on multi-scale feature aggregation diffusion and edge information enhancement[J]. Computer Engineering and Applications, 2025, 61(7): 105-116.
江旺玉,王乐,姚叶鹏,等. 多尺度特征聚合扩散和边缘信息增强的小目标检测算法[J]. 计算机工程与应用, 2025, 61(7): 105-116.
- [9] HAN K, WANG Y, TIAN Q, et al. Ghostnet: more features from cheap operations[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020: 1580-1589.
- [10] DING X, ZHANG X, MA N, et al. Repvgg: making vgg-style convnets great again[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021: 13733-13742.
- [11] BAKIRCI M, DMYTROVYCH P, BAYRAKTAR I, et al. Multi-class vehicle detection and classification with YOLO11 on UAV-captured aerial imagery[C]//Kyiv: 2024 IEEE 7th International Conference on Actual Problems of Unmanned Aerial Vehicles Development, 2024.
- [12] FARHADI A, REDMON J. Yolo3: an incremental improvement[DB/OL]. <https://arxiv.org/abs/1804.02767>.
- [13] ZHOU H, WANG Y, YE M. A method of CNN traffic classification based on spynet[C]//2018 14th International Conference on Computational Intelligence and Security, 2018: 390-394.
- [14] WANG K, LIEW J H, ZOU Y, et al. Panet: few-shot image semantic segmentation with prototype alignment[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision, 2019: 9197-9206.
- [15] DING X, ZHANG X, HAN J, et al. Diverse branch block: building a convolution as an inception-like unit[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021: 10886-10895.
- [16] WOO S, PARK J, LEE J Y, et al. CBAM: convolutional block attention module[C]//Proceedings of the European Conference on Computer Vision, 2018: 3-19.
- [17] RASKUTTI G, WAINWRIGHT M J, YU B. Early stopping and non-parametric regression: an optimal data-dependent stopping rule[J]. The Journal of Machine Learning Research, 2014, 15(1): 335-366.
- [18] ZHAO J, ZHANG J, LI D, et al. Vision-based anti-UAV detection and tracking[J]. IEEE Transactions on Intelligent Transportation Systems, 2022, 23(12): 25323-25334.
- [19] ZHENG Y, CHEN Z, LV D, et al. Air-to-air visual detection of micro-UAVs: an experimental evaluation of deep learning[J]. IEEE Robotics and Automation Letters, 2021, 6(2): 1020-1027.
- [20] WANG C, LI Z, GAO Q, et al. Lightweight and efficient air-to-air unmanned aerial vehicle detection neural networks[C]//2023 IEEE International Conference on Unmanned Systems, 2023: 1575-1580.
- [21] MA N, ZHANG X, ZHENG H T, et al. Shufflenet v2: practical guidelines for efficient CNN architecture design[C]//Proceedings of the European Conference on Computer Vision, 2018: 116-131.
- [22] WANG C Y, YEH I H, MARK LIAO H Y. Yolo9: learning what you want to learn using programmable gradient information[C]//European Conference on Computer Vision, 2025: 1-21.
- [23] WANG A, CHEN H, LIU L, et al. Yolo10: real-time end-to-end object detection[DB/OL]. <https://arxiv.org/html/2405.14458>.
- [24] ZHAO Y, LV W, XU S, et al. Detsr beat yolos on real-time object detection[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2024: 16965-16974.
- [25] NGO H T, PHAN M K, LE N L, et al. YOLOv3 tiny-based human detection and tracking using vitis AI and VVAS on FPGA[C]//2024 5th International Conference on Communications, Information, Electronic and Energy Systems, 2024: 1-6.

(责任编辑:尹晨茹)