DOI: 10. 13382/j. jemi. B2407893

# 基于 YOLOv8 的遥感小目标检测算法\*

冉 宁<sup>1,2</sup> 施高朗<sup>1</sup> 张少康<sup>3</sup> 郝晋渊<sup>4,5</sup>

(1.河北大学电子信息工程学院 保定 071002;2.河北大学节能技术研发中心 保定 071002;
 3.河北大学网络空间安全与计算机学院 保定 071002;4.河北大学中央兰开夏传媒与创意学院 保定 071002;5.河北大学物联网智能技术研究中心 保定 071002)

摘 要:遥感小目标图像存在检测目标过于密集、尺寸过小、特征信息难以提取等缺陷,导致现有的目标检测算法在遥感小目标 图像上的检测精度不高。为了解决上述问题,提出了一种基于 YOLOv8 的遥感小目标检测算法 SBC-YOLOv8 并采用了切片辅 助超推理(SAHI)方法。首先,采用了 SAHI 切片方法对遥感小目标图像进行了切片操作,有效地改善了检测目标过于密集、尺 寸过小的缺陷;其次,在 YOLOv8 的 Backbone 部分加入 Space-to-Depth 模块,增强了网络结构的特征提取能力,有效地改善了小 目标特征信息难以提取的缺陷;然后,采用了加权双向特征金字塔(BiFPN)的特征融合方式并且将原本的 P5 层替换为 P2 层, 增强了网络的多尺度特征融合能力,有效地提升了检测精度;最后,采用 CSP-OmniFusion 模块,进一步改善了遥感小目标特征 信息难以提取的缺陷。实验结果表明,相比于原 YOLOv8 算法,采用 SAHI 加上 SBC-YOLOv8 算法的改进在 VisDrone2019 数据 集的验证集和测试集平均精度均值(mAP)mAP@0.5 分别提升了 23.4%和 18.3%;mAP@0.5:0.95 分别提升了 17.4%和 12.4%,同时在 CARPK 数据集和 HRSID 数据集上 mAP@0.5 分别提升了 1.6%和 1%,mAP@0.5:0.95 分别提升了 6.1%和 2.7%。因此,所提算法有效地提升了遥感小目标图像的检测效果。

关键词:遥感小目标;目标检测;YOLOv8;多尺度特征融合

中图分类号: TP11; TN98 文献标识码: A 国家标准学科分类代码: 510.80

## Remote sensing small target detection algorithm based on YOLOv8

Ran Ning<sup>1,2</sup> Shi Gaolang<sup>1</sup> Zhang Shaokang<sup>3</sup> Hao Jinyuan<sup>4,5</sup>

(1. School of Electronic Informational Engineering, Hebei University, Baoding 071002, China; 2. Laboratory of Energy-Saving Technology, Hebei University, Baoding 071002, China; 3. School of Cyber Security and Computer, Hebei University, Baoding 071002, China; 4. HBU-UCLAN School of Media, Communication and Creative Industries, Hebei University, Baoding 071002, China; 5. Laboratory of IoT Technology, Hebei University, Baoding 071002, China)

Abstract: Remote sensing small target images often suffer from issues such as overly dense targets, small target sizes, and difficulty in feature extraction, leading to low detection accuracy for existing object detection algorithms. To address these problems, this paper proposes an SBC-YOLOv8 algorithm for remote sensing small target detection based on YOLOv8 and incorporates the SAHI slicing method. First, the SAHI slicing method is applied to slice the remote sensing small target images, effectively mitigating the problems of excessive target density and small sizes. Second, a Space-to-Depth module is added to the Backbone of YOLOv8 to enhance the network's feature extraction capability, effectively addressing the challenge of extracting small target features. Then, a BiFPN feature fusion method is employed, and the original P5 layer is replaced with the P2 layer, strengthening the network's multi-scale feature fusion ability and improving detection accuracy. Finally, the CSP-OmniFusion module is adopted to further address the difficulty of

收稿日期: 2024-10-16 Received Date: 2024-10-16

\*基金项目:国家自然科学基金(62373132)、中央引导地方科技发展资金项目(236Z1602G)、石家庄市驻冀高校基础研究项目(241791367A)、保 定市科技计划项目(2472P006)、河北大学优秀青年科研创新团队建设项目(QNTD202411)、河北大学多学科交叉研究计划(DXK202409)项目 资助 extracting remote sensing small target features. Experimental results show that, compared to the original YOLOv8 algorithm, the SBC-YOLOv8 algorithm with SAHI improvements yields a 23.4% and 18.3% increase in mAP@ 0.5 on the validation and test sets of the VisDrone2019 dataset, respectively; mAP@ 0.5 : 0.95 increases by 17.4% and 12.4%, respectively. Additionally, on the CARPK and HRSID datasets, mAP@ 0.5 increases by 1.6% and 1%, and mAP@ 0.5 : 0.95 increases by 6.1% and 2.7%, respectively. Therefore, the proposed algorithm effectively improves the detection performance of remote sensing small target images.

Keywords: remote sensing small targets; object detection; YOLOv8; multi-scale feature fusion

## 0 引 言

随着无人机和卫星遥感技术的不断进步,基于无人 机和卫星拍摄的遥感图像在各个领域的应用日益广泛。 遥感小目标检测在环境监测、军事侦察、城市管理和农业 管理等方面具有重要意义<sup>[1]</sup>。遥感图像往往具有目标密 集和目标尺寸较小、难以检测等缺陷,如何准确的检测出 小目标对象是一个急需解决的难题。

传统的遥感小目标检测方法如边缘检测、模板匹配、 形态学处理和统计学习方法(如支持向量机),虽然在一 些特定场景下有效,但普遍存在适应性差的缺点。这些 方法依赖于手工特征提取,容易受到噪声干扰,对复杂背 景和小目标的检测效果较差,且在目标尺度变化较大或 形态变化时,表现不理想。传统的目标检测算法的鲁棒 性和实时性有限,难以满足现代遥感图像中小目标检测 的需求<sup>[2]</sup>。

近年来,随着深度学习方法的兴起,基于深度学习的 检测算法在图像检测领域取得了显著的成功[34]。基于 深度学习的检测算法一般被分为两类,一类是两阶段算 法,两阶段目标检测算法通常通过首先生成候选区域,然 后对这些区域进行分类和回归,以实现目标的精确检测。 经典的两阶段检测算法包括区域卷积神经网络(region based convolutional neural networks, R-CNN) 系列(如 Fast R-CNN<sup>[5]</sup>、Faster R-CNN<sup>[6]</sup>和 Mask R-CNN<sup>[7]</sup>)。这种两 阶段算法通常具有较高的精度,能够精确度识别小目标 且具有良好的鲁棒性,但是当面对小目标密集的场景时, 可能会生成大量的候选框,导致冗余计算。而另一类算 法为单阶段算法,单阶段算法将目标检测任务直接转化 为一个端到端的回归问题,直接对输入图像进行目标定 位和分类,不在生成候选区域,节省了大量的算力。经典 的单阶段算法包括单次多框检测器(SSD)<sup>[8]</sup>、YOLO(you only look once)<sup>[9-10]</sup>系列等,其中 YOLO 系列算法因其极 低的计算成本和高效率得到了广泛的应用,成为了目标 检测领域的主流算法之一。但 YOLO 系列算法在面对遥 感小目标检测难题时仍然存在对于密集目标的处理较 差、检测精度不高、容易出现漏检和错检等问题。

为了解决 YOLO 系列算法面对遥感小目标时检测效 果较差的问题,大量学者进行了相关工作的研究。Qiu 等<sup>[11]</sup>提出了一种基于多尺度特征融合的无人机道路检

测方法 ASFF-YOLOv5,利用 ASFF 检测头和快速空间金 字塔池化(SPPF)结构提高特征尺度不变性,增强目标检 测效果。尽管 ASFF 模块可以改进多尺度特征的融合, 但对于特别小的目标,仍然可能面临检测精度不足的问 题。Zhu 等<sup>[12]</sup> 提出了 TPHYOLOv5 网络模型,将 Transformer 与 YOLO 相结合,并增加了一个小目标检测 头,有效地提升了遥感小目标的检测精度,但精度提升有 限,同时也带来了参数量增加、网络体积变大等问题。 Sahin 等<sup>[13]</sup> 提出了 YOLODrone, 这是一种基于 YOLOv3 的改进网络,增加了检测层以适应多尺度变化的目标,但 是极大的增加了网络的复杂度。Deng 等<sup>[14]</sup>提出了一种 轻量化的网络模型 LAI-YOLOv5s,该模型采用了一种轻 量化的检测头结构,并且改进了原模型特征融合的方式, 极大的提高了模型的实时性,可以做到实时检测,但是对 遥感小目标的精度提升有限,面对复杂场景的检测任务 表现不佳。

虽然上述研究在遥感小目标的检测上有所进展,在 一定程度上改进了检测效果,但是遥感小目标检测仍然 面临着巨大的挑战。当目标过于密集、尺寸过小时,上述 算法的检测效果较差、检测精度依然不高,同时在复杂场 景下容易出现漏检和错检情况。为了解决上述问题,提 出了一种基于 YOLOv8 的改进算法 SBC-YOLOv8,结合了 切片辅助超推理(slicing aided hyper inference, SAHI)数 据增强方法并将 Space-to-Depth 模块、加权双向特征金字 塔网络(BiFPN+P2)结构和 CSP-OmniFusion 模块集成到 YOLOv8 中。实验结果表明,SBC-YOLOv8 可以有效地提 升遥感小目标的检测精度,为小目标检测提供了更多的 思路。

## 1 YOLOv8 算法

#### 1.1 YOLOv8 概述

YOLOv8 是 YOLO 系列目标检测模型的第8个版本, 由 Ultralytic 公司于 2023 年提出,旨在提升目标检测的精 度和效率,YOLOv8 具有较高的检测精度和实时检测速 度。YOLOv8 主要由主干部分(backbone)、颈部(neck) 以及检测头(head)3 部分组成。其中 Backbone 部分由 Conv、C2f、SPPF 模块组成,C2f、SPPF 模块可以获取更加 丰富的梯度流信息,其主要作用为从输入图像中提取出 不同尺度的有效的特征信息,相比于之前的 YOLO 版本, YOLOv8 对主干部分进行了优化,加强了主干网络的特 征提取能力。Neck 采用经典的 PANet 架构,由 Conv、 C2f、Concat 模块组成,Neck 部分可以将主干所提取的高 级特征信息和低级特征信息进行融合,从而得到输入图 像更加丰富的特征信息,增强了网络模型的特征融合能 力。Head 部分由 3 个检测头组成,分别对小、中、大 3 种 不同尺寸的特征图进行检测,最后输出检测目标的类别、 坐标、置信度信息。YOLOv8 包含 5 种不同尺寸的模型, 分别是 YOLOv8n、YOLOv8s、YOLOv8m、YOLOv8l、 YOLOv8X,其中 YOLOv8s 很好的平衡了检测精度与检测 速度,因此,选取 YOLOv8s 作为基准模型。YOLOv8 的具 体结构如图 1 所示。



Fig. 1 YOLOv8 architecture diagram

#### 1.2 YOLOv8 的不足与局限

虽然 YOLOv8 以其极高的精度和推理速度得到了广 泛的认可和应用,但是当面对遥感小目标的检测难题时, YOLOv8依旧存在着一些不足与局限。相较于以往的 YOLO 系列模型,尽管 YOLOv8 在小目标检测上有所改 进,但仍然可能出现小目标识别能力较差的情况, YOLOv8 的主干部分提取小目标特征信息的能力较差, 无法很好的捕捉到小目标的特征信息。这是由于小目标 物体所占的像素较少,导致其特征经过多次卷积操作之 后可能会出现特征信息丢失的情况,极大的降低了 YOLOv8 的检测精度。此外, YOLOv8 网络结构设计主要 是针对大型目标和中型目标,3个检测头的网络结构可 能无法处理多尺度变化的检测目标,导致 YOLOv8 无法 适应多尺度变化的图像,面对极小尺寸的目标,YOLOv8 易出现漏检的情况。另外,当检测图像存在背景遮挡、小 目标物体重叠等情况出现时, YOLOv8 易出现漏检情况。 这些问题导致了面对遥感小目标图像时 YOLOv8 的性能 表现不佳。

## 1.3 YOLOv8 改进思路

为了改善 YOLOv8 在遥感小目标检测方面的性能, 提升遥感小目标图像检测的精准度,降低误检率和漏检 率,重新设计了 YOLOv8 的主干部分并采用了新的特征 融合方式替代了原本 YOLOv8 的特征融合方式,同时采 用了小目标检测头代替了大目标检测头,这些改进增强 了 YOLOv8 多尺度特征的提取和融合能力,提高小目标 特征的表达效果,使得网络能够适应多尺度变化的小目 标物体。同时,采用数据增强的方式提高了图像中的小 目标占比,有效地降低了漏检率。

## 2 SBC-YOLOv8 算法

#### 2.1 SBC-YOLOv8 概述

针对上述 YOLOv8 算法在面对遥感小目标检测时所存在的问题,对原 YOLOv8 算法进行了改进,提出了一种针对遥感小目标图像的检测算法 SBC-YOLOv8。首先采取了 SAHI 图像切片方法,提高了遥感图像中小目标的占比,有效地改善了遥感小目标图像检测目标过小、过于密集的问题;然后对 YOLOv8 模型进行了改进,在主干部分加入了 Space-to-Depth 模块,有效地改善了特征信息丢失的问题;同时采用 BiFPN 特征融合方式代替了原本的PANet,并且用小目标检测层 P2 替换了原本的 P5 层,得到了更加丰富的有关遥感小目标的特征信息,有效地提

升了遥感小目标的检测精度,最后在 YOLOv8 颈部加入 了 CSP-OmniFusion 模块,进一步改善了遥感小目标特征 信息难以提取的问题。改进后的 SBC-YOLOv8 模型的结 构如图 2 所示。



图 2 SBC-YOLOv8 结构 Fig. 2 Structure of SBC-YOLOv8

## 2.2 SAHI 图像切片技术

SAHI 是一种针对目标检测任务的优化技术,旨在提高大尺寸图像和复杂背景图像的检测精度和效率,通常被用于模型推理预测阶段。SAHI 首先将图像切分成多个较小的块(通常称为切片)。这些切片的大小是根据模型的输入要求和计算资源来确定的。切分过程可以是均匀的,也可以是基于图像内容的动态调整。然后每个切片被独立地送入目标检测模型进行处理。最后对每个切片的检测结果进行汇总。SAHI 使用精确的拼接和去重技术,将切片中检测到的目标信息整合成完整图像的检测结果。这样可以有效地避免重复检测和漏检问题。推理过程如图 3 所示。

为了改善 VisDrone2019 数据集背景复杂、目标遮挡、目标过小等缺点,对 VisDrone2019 数据集采用 SAHI 切片的方法<sup>[15]</sup>,得到了切片后的 VisDrone2019 数据集。以 原 VisDrone2019 数据集的一张图片为例,选取切片后子 图尺寸为 640×640,交叉重叠率为 0.2 对图片进行切片 处理,原图和切片后子图分别如图 4 和 5 所示。可以看出,SAHI 切片方法有效地改善了 VisDrone2019 数据集检测目标过于密集、尺寸过小的缺陷。



图 3 SAHI 切片辅助超推理方法 Fig. 3 SAHI slicing-assisted hyper-inference method



图 4 VisDrone2019 数据集原图 Fig. 4 Original image of the VisDrone2019 dataset



图 5 切片后 VisDrone2019 数据集子图 Fig. 5 Subplot of the VisDrone2019 dataset after slicing

## 2.3 空间到深度卷积(Space-to-Depth)

近些年来,卷积神经网络(CNN)在深度学习领域取 得了巨大的成功,它也被广泛的运用于目标检测领域,然 而,面对背景复杂度较高的遥感小目标图像,传统的卷积 神经网络表现不佳。这是由于 CNN 设计缺陷所导致的, 传统的卷积神经网络采用了跨行卷积或者池化层,会造 成大量特征信息的损失,当面对大目标检测任务时,由于 特征信息充足,CNN 的跨步卷积和池化操作会去掉大量 的冗余信息,保留主要特征,所以 CNN 的性能表现依旧 出色;但是当面对遥感小目标的检测任务时,跨步卷积和 池化操作会造成大量特征信息的丢失,导致卷积神经网 络无法很好的学习遥感小目标的特征,造成性能和精度 的大幅度下降。为了解决上述问题,在 YOLOv8 的卷积 操作之后加入了 Space-to-Depth 模块(空间到深度卷积模 块).有效地避免了遥感小目标的特征信息丢失的问题。

Space-to-Depth<sup>[16]</sup>模块是一种用于将特征图的空间 维度转换为深度维度的技术。Space-to-Depth 模块的主 要作用是将特征图的空间信息重组到深度(通道)维度 中。它通常用于增加特征图的深度(通道数),同时减少 其空间维度。通过将空间信息重组到深度维度中,可以 增加特征图的深度,从而提高特征图的空间分辨率。在 一些网络架构中,尤其是超分辨率网络中,使用 Space-to-Depth 模块可以提升模型对细节的恢复能力,使得生成的 图像更加清晰和细腻。该操作可以增强特征图的表达能 力,保留更多上下文的细节和特征信息。其具体实现办 法为将输入张量重排成多个小块,然后将这些小块的像 素值组合到新的深度通道中。例如:假设输入的张量形 状为(N, C, H, W),其中 N 是批量大小(batch size), C 是 通道数(depth), H和 W 分别为特征图的高度和宽度,下 采样因子为r,输入张量的每个r×r的空间块将合并到 一个新的通道中,输出的张量形状变为 $(N, C \times r^2, H/r)$ , W/r),其中 $C \times r^2$ 是新的通道数,因为每个 $r \times r$ 的块被合 并为一个通道,H/r和W/r为特征图新的高度和宽度。 Space-to-Depth 模块示意图如图 6 所示。



图 6 Space-to-Depth 模块示意图 Fig. 6 Schematic of the Space-to-Depth module

## 2.4 改进特征融合方式(BiFPN+P2)

YOLOv8 的 Neck 部分采用 PANet<sup>[17]</sup> 架构,它是在 FPN 的基础之上改进而来,具体改进为在特征金字塔结 构(FPN)的基础上再建立一条自下而上的通路。这种特 征融合的结构使得高层的特征图具有更强的语义信息, 底层的特征图具有更强的位置信息,能够更好的融合浅 层和深层特征,极大的提高了检测精度。虽然 PANet 支 持多尺度特征融合,但主要是单向特征融合,从高分辨率 到低分辨率,不支持双向流动。并且特征融合方式较为 固定,主要依靠传统的加权方法,还会造成部分特征信息 的损失。PANet 的结构如图 7(a) 所示。

因此,为了改进原本的特征融合方式,采用 BiFPN<sup>[18]</sup>特征融合替换了原本的PANet,BiFPN通过双 向特征融合,将特征图从高分辨率到低分辨率及反向融 合,增强了多尺度特征的交互信息流动。这种双向融合 帮助捕捉更加丰富和多样的特征信息。同时引入了可学 习的加权机制,通过自适应地调整不同尺度特征的权重, 提高了特征融合的质量。网络中的加权层通过训练自动 学习最优的融合方式。并且采用了高效的特征融合策 略,通过减少计算复杂度和内存消耗来提高性能。这种 高效的设计使得 BiFPN 在保持高精度的同时,计算开销 更小。BiFPN 的结构如图 7(b)所示。

同时,采用 P2 层替换掉 YOLOv8 原本的 P5 层,P2 层的特征图具有更高的空间分辨率,相比于 P5 层的较低 分辨率特征图,更适合捕捉小目标的细节。这对于遥感 图像中的小目标检测尤为重要,因为高分辨率特征图可 以保留更多的细节信息,有助于提高小目标的检测精度, 高分辨率的特征图能够更好地表达小目标的形状和位 置,有助于模型更准确地识别和定位这些小目标。通过 将 YOLOv8 原本的 P5 层替为 P2 层,模型在遥感小目标 检测中的精度得到了明显提升<sup>[19]</sup>。



图 7 PANet、BiFPN 结构 Fig. 7 Structure diagram of PANet and BiFPN

## 2.5 提升特征提取能力(CSP-OmniFusion)

Omni-Kernel 模块<sup>[20]</sup>由3个分支组成,即全局分支、 大分支和局部分支。特征图经过1×1的卷积之后被输入 Omni-Kernel 模块,之后,会被反馈经过3个分支。其中, 全局分支经过DCAM 模块和FSAM 模块的处理,模块结 构如图8(b)、(c)所示;大分支则分别经过3个不同大小 卷积核的深度可分离卷积(DConv);局部分支则经过1× 1的Dconv处理,之后将特征图进行加法操作拼接在一 起。Omni-Kernel 模块可以确保神经网络有效地学习全 局到局部的特征表征,增强了神经网络的特征提取和特 征融合能力。Omni-Kernel 模块结构如图8(a)所示。

CSP(cross stage partial)结构源于 CSPNet<sup>[21]</sup>,CSP 结构被有效地应用于 YOLOv5 中。CSP 结构是由跨阶段部 分连接组成的,主要思想为将特征图按通道分割为两个

部分,其中一部分经过多个卷积处理,另外一部分则直接 进行下一步的计算。两部分特征图处理完毕之后经过残 差连接合并在一起。CSP模块允许网络同时处理不同的 信息流,使得特征提取更加丰富和准确。通过在网络中 使用部分特征图,提高了计算效率,同时特征图在进行高 层特征提取时,可以保留更多的底层特征,提高了网络的 感受野和特征提取能力。YOLOv5 中 CSP模块结构如图 8(d) 所示。

CSP-OmniFusion 模块集成了 CSP 思想和 Omni-Kernel 模块,其结构如图 8(e)所示。CSP-OmniFusion 模 块的基本思想是将特征图分割为两个部分,一部分经过 Omni-Kernel 模块处理,另一部分不做处理,保留原始的 特征信息,最后将两部分特征图拼接到一起输出。P2 和 P3 层的特征图包含了丰富的小目标特征信息。将这些 特征图输入到 CSP-OmniFusion 模块中,该模块通过融合 来自不同卷积操作的特征图,并进一步通过频域处理进 行增强,从而更全面地捕捉图像中的各类特征信息,显著 提升了特征融合能力。此外,CSP-OmniFusion 模块通过 特征的分割与重组,改善了网络中的梯度流动。



图 8 CSP-OmniFusion 结构

Fig. 8 Structure diagram of CSP-OmniFusion

## 3 实验数据与结果分析

## 3.1 实验环境与超参数设置

实验环境如表 1 所示,实验所采用的硬件加速卡为 两张英伟达的 RTX 3090(24 G)显卡,中央处理器 CPU 为 Intel Xeon Gold 6226R;所采用的深度学习实验环境为 Ubuntu 22.04、Python3.10、Cuda12.1、pytorch2.1.0;实验 超参数设置如下:训练轮数 Epochs 设置为 200,批大小 Batch size 设置为 16, Image size 设置为 640,其余超参数 均采用默认值。

## 3.2 数据集与评价指标

为了对 SBC-YOLOv8 算法进行评估,采用了经典的 遥感 小 目 标 数 据 集—VisDrone2019 数 据 集<sup>[22]</sup>。 VisDrone2019 数据集是由天津大学发布,专门用于无人

表1 实验环境

Table 1	Experimental environment
名称	配置信息
操作系统	Ubuntu 22. 04
开发语言	Python3. 10
深度学习框架	Cuda12. 1+pytorch2. 1. 0
GPU	RTX 3090(24 G) *2
CPU	Intel Xeon Gold 6226R

机视觉任务的大型基准数据集,是经典的遥感小目标数据集之一,涵盖了从中国 14 个不同城市收集的广泛数据。这些图像和视频由各种无人机摄像头捕获,记录了多种天气和光照条件下的场景。VisDrone2019 数据集包含 8 629 张图像,其中 6 471 张用于训练,548 张用于验证,1 610 张用于测试,图中共包含 10 类标注物体,分别是行人、人、汽车、面包车、巴士、卡车、摩托车、自行车、遮

阳篷-三轮车和三轮车,共计超过 260×10<sup>4</sup> 个标注框。因为 VisDrone2019 数据集存在目标遮挡严重、部分目标过小以及数据分布不均等问题,导致 VisDrone2019 数据集仍然是当前比较具有挑战性的遥感小目标数据集之一。

为了评估模型对遥感小目标对象的检测精度,主要 采用精准度(precision,P)和召回率(recall,R)以及平均 精度均值(mAP)评价指标,参数量和计算量来评估模型 大小。

1) 精准度和召回率的计算公式<sup>[23]</sup> 如式(1) 和(2) 所示。

$$P = \frac{TP}{TP + FP} \tag{1}$$

$$R = \frac{TP}{TP + FN} \tag{2}$$

精准度衡量的是模型预测为正样本的实例中,实际 为正样本的比例。TP(true positives)表示模型正确预测 为正类的样本数;FP(false positives)表示模型错误预测 为正类的样本数。召回率衡量的是实际正样本中,被模 型正确预测为正样本的比例。FN(false negatives)表示 模型错误预测为负类的实际正类样本数。

2) 在目标检测任务中, 所有类别的 mAP 是一个常用 的性能评估指标。mAP 主要用来评估模型在不同类别 上的平均检测精度。mAP@0.5和mAP@0.5:0.95是 在不同交并比(intersection over union, IoU)阈值下计算的 mAP。mAP@0.5和mAP@0.5:0.95的计算公式如 式(3)和(4)所示。

$$mAP@ 0.5 = \frac{1}{C} \sum_{i=1}^{C} AP_{i,\text{IoU}=0.5}$$
(3)

$$mAP@ 0.5: 0.95 = \frac{1}{N} \sum_{j=1}^{N} AP_{i, \text{IoU}=j}$$
(4)

式中:*C* 是类别数; *AP*<sub>*i*,IOU=0.5</sub> 是第 *i* 类别在 IoU 阈值为 0.5 时的平均精度(AP); *N* 是 IoU 阈值的数量; *AP*<sub>*i*,IOU=*j*</sub> 是第 *i* 类别在 IoU 阈值为 *j* 时的 *AP*。

3)模型的参数量指的是模型中所有可训练参数的总数,主要用于评估模型的复杂度和存储需求。模型的计算量通常指的是在模型训练或推理过程中所需要消耗的计算资源。

## 3.3 消融实验

为了验证所采用的数据增强方法 SAHI 图像切片和 所采用的 3 个模块的有效性(Space-to-Depth(模块 A)、 BiFPN+P2(模块 B)、CSP-OmniFusion(模块 C)),在 VisDrone2019数据集上进行了消融实验,其中" $\sqrt{}$ "表示 添加了该模块或采用了该方法。消融实验结果分别如表 2 和 3 所示。

## 表 2 VisDrone2019 验证集消融实验 Table 2 Ablation study on the VisDrone2019 validation set

Yolov8s	模块 A	模块 B	模块 C	SAHI	mAP@ 0. 5/%	mAP@ 0. 5 : 0. 95/%	精确度/%	召回率/%	参数量/(×10 <sup>6</sup> )
$\checkmark$					39.0	23.2	49.9	38.3	11.13
$\checkmark$					41.7	25.1	53.8	39.7	12.42
$\checkmark$		$\checkmark$			48.6	29.9	57.8	46.1	7.45
$\checkmark$					40. 7	24.7	51.1	39.8	12.03
$\checkmark$					49.2	30.4	59.2	45.7	8.50
$\checkmark$			$\checkmark$		42.6	25.7	53.6	40.4	13.08
$\checkmark$		$\checkmark$	$\checkmark$		49.2	30.2	58.2	46.5	7.60
$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$		50.4	31.2	57.7	47.7	8.65
$\checkmark$				$\checkmark$	56.2	35.7	64.4	52.6	11.13
		$\checkmark$			62.4	40.6	67.2	59.0	8.65

Table 3	Ablation	study	on the	VisDrone2019	test	set
---------	----------	-------	--------	--------------	------	-----

Yolov8s	模块 A	模块 B	模块 C	SAHI	mAP@ 0. 5/%	mAP@ 0. 5 : 0. 95/%	精确度/%	召回率/%	参数量/(×10 <sup>6</sup> )
$\checkmark$					31.7	18.0	44.2	33.6	11.13
$\checkmark$	$\checkmark$				34.1	19.5	46.0	35.5	12.42
$\checkmark$		$\checkmark$			37.8	21,7	50.4	38.5	7.45
$\checkmark$			$\checkmark$		32.7	18.6	44.4	34.6	12.03
$\checkmark$	$\checkmark$	$\checkmark$			39.3	22.7	50.2	40.6	8.50
$\checkmark$	$\checkmark$		$\checkmark$		34.3	19.8	46.3	36.7	13.08
$\checkmark$		$\checkmark$	$\checkmark$		38.6	22. 2	50.9	39.0	7.60
$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$		40.1	23.0	50.6	41.0	8.65
				$\checkmark$	45.8	27.6	56.0	45.8	11.13
	$\checkmark$		$\checkmark$	$\checkmark$	50.0	30.4	58.7	50.0	8.65

由表2和3可以得出,在使用 SAHI 方法对 VisDrone2019 数据集进行切片操作之后, YOLOv8 算法在 验证集和测试集的平均精度 mAP@0.5 分别提升了 17.2%和14.1%,SBC-YOLOv8算法在验证集和测试集的 mAP@0.5分别提升了12%和9.9%,证明了SAHI切片 方法可以有效地改善遥感小目标图像目标密集、尺寸过 小的问题。Space-to-Depth 模块与 YOLOv8 相结合之后, 在验证集和测试集上的 mAP@ 0.5 分别提升了 2.7% 和 2.4%,有效地改善了小目标特征信息容易丢失的问题。 而 BiFPN+P2 结构改进了特征融合的方式并用小目标检 测头替换了大目标检测头,在验证集和测试集上的 mAP@0.5分别提升了9.6%和6.1%,大幅的提高了检 测精度。CSP-OmniFusion 增强了特征融合能力,在验证 集和测试集的 mAP@ 0.5 分别提升了 1.7% 和 1%, 有效 地改善了小目标特征信息难以提取的问题。而在 BiFPN +P2结构的基础之上融入 Space-to-Depth 模块,在验证集 和测试集上的 mAP@ 0.5 分别提升了 10.2% 和 7.6%, 进 一步的提升了检测精度。BiFPN+P2结构与CSP-OmniFusion 模块相结合之后,mAP 分别提升了 10.2%和 6.9%,进一步提升了模型的特征提取能力。Space-to-Depth 模块与 CSP-OmniFusion 结合之后, mAP@ 0.5 分别 提升了 3.6% 和 2.6%, Space-to-Depth 模块确保模型提取 出更多的关于小目标的特征信息,CSP-OmniFusion 提升 特征融合能力,保留了更加丰富的小目标的特征信息。 同时采用3项改进之后,mAP@0.5分别提升了11.4% 和 8.4%,同时参数量下降了 22.28%,极大的提高了遥感 小目标的检测精度并降低了模型的参数量。证明了 SAHI 切片方法和 SBC-YOLOv8 算法的有效性。

## 3.4 模型泛化能力评估

模型的泛化能力(generalization ability)是指模型在 没有见过的新数据集上的表现,衡量模型是否能够学习 到数据的本质规律,并将其应用于新样本。泛化实验可 以评估模型在训练数据和真实环境(或新数据)中的表 现,帮助判断其在实际应用中的有效性。为了验证所提 出的 SBC-YOLOv8 算法和所采用的 SAHI 切片方法的有 效性和泛化能力,选取了两个具有挑战性的遥感小目标 数据集,即 CARPK 数据集和 HRSID 数据集,进行了模型 的泛化实验。CARPK 数据集由 Hsieh 等<sup>[24]</sup>提出,是通过 无人机在距离地面约 40 m 的高度拍摄而成的。该数据 集包含 989 张用于训练的图像和 459 张用于测试的图 像,涵盖了来自4个不同停车场的大约9×10<sup>4</sup>辆汽车的 详细信息。HRSID 数据集则由 Wei 等<sup>[25]</sup>提出,该数据集 包括 5 604 张高分辨率 SAR 图像以及 16 951 个船舶实 例。为了评估 SBC-YOLOv8 算法的泛化性能,分别在这 两个数据集上对 SBC-YOLOv8 算法进行了训练和测试. 其中"√"表示采用了 SAHI 方法对数据集进行了切片操 作,其泛化实验的结果如表4和5所示。

表4 CARPK 数据集实验结果

Table 4 Experimental results of CARTR uatase	Fable 4	Experimental	results	of	CARPK	dataset
--	---------	--------------	---------	----	-------	---------

模型	SAHI	mAP@ 0. 5/% m	hAP@ 0.5 : 0.95/%	<i>P/%</i>	R/%
YOLOv8s		96.3	67.7	96.0	90.8
YOLOv8s	$\checkmark$	97.8	73.7	97.4	94.1
SBC-YOLOv8s	$\sim$	97.9	73.8	97.3	94.4

表 5 HRSID 数据集实验结果 Table 5 Experimental results of HRSID dataset

模型	SAHI	mAP@ 0. 5/% n	mAP@ 0. 5 : 0. 95/%	P/%	R/%
YOLOv8s		90.6	66.8	90.1	83.2
YOLOv8s	$\checkmark$	90.8	67.8	90.6	83.8
SBC-YOLOv8s	$\sim$	91.6	69.5	90.6	84.5

由表 4 和 5 可以得出,对 CARPK 数据集采用 SAHI 切片之后,平均精度 mAP @ 0.5 提升了 1.5%, mAP@0.5:0.95 提升了 6%,在 SAHI 切片方法结合 SBC-YOLOv8 算法之后,mAP@0.5 提升了 1.6%,mAP@ 0.5:0.95 提升了 6.1%。在 HRSID 数据集上,在使用 SAHI 方法对数据集进行切片之后,mAP@0.5 提升了 0.2%,mAP@0.5:0.95 提升了 1%,结合 SBC-YOLOv8 算法之后,mAP@0.5 提升了 1%,mAP@0.5:0.95 提升 了 2.7%。证明了 SAHI 切片方法对于遥感图像的适用 性和 SBC-YOLOv8 算法针对遥感小目标检测的有效性。

## 3.5 对比实验

为了验证所提出的 SBC-YOLOv8 算法的有效性,并 全面评估其在小目标检测任务中的性能表现,选取了近 年来在小目标检测领域具有代表性的经典模型和改进模 型进行对比实验。具体而言,选取的对比模型包括两阶 段检测器 Faster RCNN、单阶段检测器 RetinaNet<sup>[26]</sup>,以及 多种 YOLO 系列模型,如 YOLOv5s、TPH-YOLOv5s、 YOLOv8s 和最新的 YOLOv10s<sup>[27]</sup>。所有模型均在 VisDrone2019数据集上进行训练和测试,具体的实验结 果如表 6 所示。

쿳	₹6	VisDrone2019 测试集对比实验结果
Table 6	Ex	perimental results of VisDrone2019 dataset

	mAP@ 0. 5/	mAP@ 0. 5 : 0. 95/	计算量/	参数量/
候望	%	%	GFLOPs	$(\times 10^{6})$
Faster RCNN	32.9	19.4	208.0	41.39
RetinaNet	27.6	16.4	210.0	36.52
YOLOv5s	28.7	15.5	15.8	7.04
TPHYOLOv5s	29.7	15.7	36.7	7.46
YOLOv8s	31.7	18.0	28.5	11.13
YOLOv8m	34.6	20.0	78.7	25.85
YOLOv10s	31.1	17.6	24.5	8.04
SBC-YOLOv8s	40.1	23.0	78.8	8.65

由表6可以得出,经典的目标检测算法 Faster RCNN 具有较高的检测精度,但是其计算量 GFLOPs 和参数量 过高,存在着网络体积太过庞大的问题;RetinaNet 同样 也面临网络体积庞大的问题,同时其检测精度也不高;而 YOLOv5 虽然具有极低的参数量和计算量以及轻量的网 络结构,但是无法平衡检测精度,面对遥感小目标检测任 务时检测精度过低;TPHYOLOv5 在 YOLOv5 的基础之上 添加了小目标检测头,虽然提升了检测精度,但也带来了 网络体积变大、计算成本高昂等问题,且精度提升十分有 限:YOLOv8 和 YOLOv10 很好的平衡了检测精度和计算 成本的开销,但是其模型架构在面对密集场景下的遥感 小目标检测时易出现误检和漏检情况,导致检测精度不 高。而 SBC-YOLOv8 算法具有极高的精度和较低的参数 量,相比于原 YOLOv8 算法, mAP@ 0.5 提升了 8.4%, mAP@0.5:0.95 提升了 5%,参数量下降了 22.28%。 但是由于小目标检测层 P2 层的加入,不可避免的增加了 模型的计算量。对比表 6 中 YOLOv8m 算法与 SBC-YOLOv8s 算法发现,在模型计算量大致相等的情况下, SBC-YOLOv8 算法具有明显的精度优势,mAP@0.5 提升 5.5%, mAP@0.5:0.95 提升 3%, 并且大幅度的降低了 模型的参数量,验证了 SBC-YOLOv8s 算法的有效性, SBC-YOLOv8 有效地提升了遥感小目标图像的检测 效果。

## 3.6 可视化结果分析

为了验证 SBC-YOLOv8 算法在实际复杂场景中的有 效性,从 VisDrone2019 数据集的测试集中选取了具有代 表性的不同类型图像作为检测对象。这些图像涵盖了多 种复杂环境,包括交通流量密集的城市路口、夜间强光照 射下的开放广场以及停放大量车辆的居住区。这些场景 具有不同的挑战性,例如光照条件复杂、目标密集分布、 目标尺寸较小等,能够充分考验检测模型的鲁棒性和适 应性。

实验中,分别采用原 YOLOv8 模型和改进后的 SBC-YOLOv8 模型对上述图像进行目标检测,并对比分析两 者的检测效果,检测结果如图9所示。从检测结果对比 可以观察到,在目标尺寸较小、目标过于密集的情况下, 原 YOLOv8 模型容易出现漏检和误检,尤其是在背景复 杂的场景中。相比之下, SBC-YOLOv8 模型能够有效克 服这些问题,准确检测出部分原模型漏检的小目标,同时 降低了误检率,在目标密集区域的检测效果明显优于原 YOLOv8 模型。综合实验结果表明,SBC-YOLOv8 在应对 小目标、密集目标以及复杂光照条件方面具有更强的鲁 棒性,能够有效降低漏检率和误检率,提升检测性能。实 验结果证明了 SBC-YOLOv8 在实际应用场景中的可行性 和优越性。



(a) YOLOv8

(b) SBYOLOv8

图 9 YOLOv8 与 SBC-YOLOv8 检测结果对比 Fig. 9 Comparison of detection results between YOLOv8 and SBC-YOLOv8

## 4 结 论

为了解决遥感小目标图像目标密集、目标尺寸过小、 特征信息提取困难和检测精度较低等问题,提出了基于 YOLOv8 的改进遥感小目标算法 SBC-YOLOv8。首先采 用了 SAHI 方法对数据集进行切片处理,有效地改善了 遥感小目标图像存在的目标密集和目标尺寸过小等问 题;其次,对 YOLOv8 算法进行了改进,采用了 Space-to-Depth 模块,有效地改善了在小目标特征信息易丢失、难 以提取的问题:然后采用了 BiFPN+P2 的检测头架构, BiFPN 特征融合方式可以捕捉更加丰富和多样的特征信 息,P2 层替换 P5 层明显的提高了小目标的检测精度;最 后,所采用的 CSP-OmniFusion 模块可以确保模型有效地 学习从全局到局部的特征表征,从而提高小目标的检测 性能。与其他先进的模型相对比可以得出,SBC-YOLOv8 算法具有较优的小目标识别能力,有效地提升了遥感小 目标图像的检测精度。SBC-YOLOv8 算法虽然具有较高 的检测精度,但是同样也面临着网络体积过于庞大的问 题,未来将继续模型轻量化的研究,探讨如何在保持模型 的检测精度的同时,减小模型的网络体积和计算资源的 消耗。

## 参考文献

[1] 李嘉新,侯进,盛博莹,等. 基于改进 YOLOv5 的遥感 小目标检测网络[J]. 计算机工程, 2023, 49(9): 256-264.

> LI J X, HOU J, SHENG B Y, et al. Remote sensing small target detection network based on improved YOLOv5[J]. Computer Engineering, 2023, 49(9): 256-264.

[2] 邝先验,程福军,吴翠琴,等.基于改进 YOLOv7-tiny 的 高效轻量遥感图像目标检测方法[J].电子测量与仪 器学报,2024,38(7):22-33.

> KUANG X Y, CHENG F J, WU C Q, et al. Efficient and lightweight remote sensing image object detection method based on improved YOLOv7-tiny[J]. Journal of Electronic Measurement and Instrumentation, 2024, 38(7): 22-33.

 [3] 薛林雁,李轩昂,齐晁仪,等. 基于改进 YOLOv5s 的肠 镜息肉多分类实时检测方法[J].河北大学学报(自然 科学版),2024,44(4):424-432.

> XUE L Y, LI X A, QI CH Y, et al. Multi-classification real-time detection method of colonoscopic polyps based on improved YOLOv5s [J]. Journal of Hebei University (Natural Science Edition), 2024, 44(4):424-432.).

[4] 陈向阳,周扬,杨文柱.融合深度特征的改进 KCF 行人 追踪算法[J].河北大学学报(自然科学版),2023, 43(6):665-672.

CHEN X Y, ZHOU Y, YANG W ZH. Improved KCF pedestrian tracking algorithm based on fusion of depth features [J]. Journal of Hebei University (Natural Science Edition), 2023, 43(6):665-672.

- [5] GIRSHICK R. Fast R-CNN[J]. ArXiv preprint arXiv: 1504.08083, 2015.
- [6] REN S, HE K, GIRSHICK R, et al. Faster R-CNN: Towards real-time object detection with region proposal networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(6): 1137-1149.
- [7] HE K, GKIOXARI G, DOLLÁR P, et al. Mask R-CNN[C].
   Proceedings of the IEEE International Conference on Computer Vision, 2017: 2961-2969.
- [8] LIU W, ANGUELOV D, ERHAN D, et al. SSD: Single shotmultibox detector [C]. 14th European Conference on Computer Vision, 2016: 21-37.
- [9] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: Unified, real-time object detection [C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016: 779-788.
- [10] WANG C Y, BOCHKOVSKIY A, LIAO H Y M. YOLOv7: Trainable bag-of-freebies sets new state-of-theart for real-time object detectors [C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2023: 7464-7475.
- [11] QIU M, HUANG L, TANG B H. ASFF-YOLOv5: Multielement detection method for road traffic in UAV images based on multiscale feature fusion [J]. Remote Sensing, 2022, 14(14): 3498.
- ZHU X, LYU S, WANG X, et al. TPH-YOLOv5: Improved YOLOv5 based on transformer prediction head for object detection on drone-captured scenarios [C]. Proceedings of the IEEE/CVF International Conference on Computer Vision, 2021: 2778-2788.
- SAHIN O, OZER S. YOLODrone: Improved YOLO architecture for object detection in drone images[C]. 2021
   44th International Conference on Telecommunications and Signal Processing (TSP). IEEE, 2021: 361-365.
- [14] DENG L, BI L, LI H, et al. Lightweight aerial image object detection algorithm based on improved YOLOv5s [J].
   Scientific Reports, 2023, 13(1): 7817.
- [15] 白宇,周艳媛,安胜彪.改进 YOLOv5 的无人机小目标 检测方法研究[J].计算机工程与应用,2024,60(10): 276-284.

BAI Y, ZHOU Y Y, AN SH B. Research on drone small target detection method based on improved YOLOv5[J]. Computer Engineering and Applications, 2024, 60(10):

第5期

#### 276-284.

- [16] SUNKARA R, LUO T. No more strided convolutions or pooling: A new CNN building block for low-resolution images and small objects [C]. Joint European Conference on Machine Learning and Knowledge Discovery in Databases. Cham: Springer Nature Switzerland, 2022: 443-459.
- [17] LIU S, QI L, QIN H, et al. Path aggregation network for instance segmentation [C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018: 8759-8768.
- [18] TAN M, PANG R, LE Q V. Efficientdet: Scalable and efficient object detection [C]. Proceedings of the IEEE/ CVF Conference on Computer Vision and Pattern Recognition, 2020: 10781-10790.
- [19] 翁俊辉,成乐,黄曼莉,等. 基于 CS-YOLOv5s 的无人机 航拍图像小目标检测 [J]. 电子测量技术, 2024, 47 (7): 157-162.

WENG J H, CHENG L, HUANG M L, et al. Small target detection in UAV aerial images based on CS-YOLOv5s [J]. Electronic Measurement Technology, 2024, 47(7): 157-162.

- [20] CUI Y, REN W, KNOLL A. Omni-kernel network for image restoration [C]. Proceedings of the AAAI Conference on Artificial Intelligence, 2024, 38 (2): 1426-1434.
- [21] WANG C Y, LIAO H Y M, WU Y H, et al. CSPNet: A new backbone that can enhance learning capability of CNN[C]. Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops. 2020: 390-391.
- [22] DU D, ZHU P, WEN L, et al. VisDrone-DET2019: The vision meets drone object detection in image challenge results [C]. Proceedings of the IEEE/CVF International Conference on Computer Vision Workshops, 2019.
- [23] PADILLA R, NETTO S L, DA SILVA E A B. A survey on performance metrics for object-detection algorithms [C].
  2020 International Conference on Systems, Signals and Image Processing (IWSSIP). IEEE, 2020; 237-242.
- [24] HSIEH M R, LIN Y L, HSU W H. Drone-based object counting by spatially regularized regional proposal network [ C ]. Proceedings of the IEEE International Conference on Computer Vision, 2017; 4145-4153.

- [25] WEI S, ZENG X, QU Q, et al. HRSID: A highresolution SAR images dataset for ship detection and instance segmentation [J]. IEEE Access, 2020, 8: 120234-120254.
- [26] LIN T Y, GOYAL P, GIRSHICK R, et al. Focal loss for dense object detection [C]. Proceedings of the IEEE International Conference on Computer Vision, 2017: 2980-2988.
- [27] WANG A, CHEN H, LIU L, et al. YOLOv10: Realtime end-to-end object detection [J]. ArXiv preprint arXiv:2405.14458, 2024.

#### 作者简介



**冉宁**,2010年于河北大学获学士学位, 2013年于华北电力大学获硕士学位,2017 年于浙江大学获博士学位,现为河北大学副 教授,主要研究方向为无人机图像处理、路 径规划等。

E-mail: ranning87@ hotmail. com

**Ran Ning** received B. Sc. degree from Hebei University in 2010, M. Sc. degree from North China Electric Power University in 2013, and Ph. D. degree from Zhejiang University in 2017. He is currently an associate professor at Hebei University. His main research interests include UAV image processing and UAV path planning.



施高朗,2023年于长江大学获得学士 学位,现为河北大学硕士研究生,主要研究 方向为深度学习、目标检测。

E-mail: 2803718919@ qq. com

Shi Gaolang received his B. Sc. degree

from Yangtze University in 2023. He is

currently a M. Sc. candidate at Hebei University. His main research interests include deep learning and object detection.



**郝晋渊**(通信作者),2013 年于河北大 学获学士学位,2015 年于新南威尔士大学 获硕士学位,现为河北大学讲师,主要研究 方向为人工智能、信息技术、新媒体传播等。 E-mail: hjy@ hbu. edu. cn

 ${\bf Hao} \ \ {\bf Jinyuan} \ \ ( \ {\rm Corresponding} \ \ {\rm author} \ )$ 

received the B. Sc. degree from Hebei University in 2013, and M. Sc. degree from the University of New South Wales in 2015. She is currently a lecturer at Hebei University. Her main research interests include artificial intelligence, information technology and new media communication.