

DOI: 10.13382/j.jemi.B2508108

# 基于 YOLO11 的轻量级遥感图像检测算法\*

冉宁<sup>1,2</sup> 施高朗<sup>1</sup> 张昊侯<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:**针对现有的目标检测算法在面对遥感小目标检测难题时存在网络体积大、模型复杂度高、检测准确率低等问题,提出了一种基于 YOLO11 的轻量级遥感算法 EDB-YOLO11。该算法主要从模型结构和损失函数两个方面进行了改进。首先,采用 EMBC 模块改进了原本的 C3K2 模块,有效地提升了网络结构的特征表达能力;其次,采用 Downsample 下采样模块替换了原本的卷积下采样,在降低模型参数量和计算量的同时提升了主干部分的特征提取能力;然后,采用 BiFPN 特征融合方式替换了原本的 PANet,增强了网络的特征融合能力,实现了更高效的多尺度特征融合;最后,结合 Focal 机制与 WIoU 的优势,引入了一种新的损失函数 Focal-WIoU,使模型更关注高质量的训练样本,从而降低了低质量训练样本的影响,有效地提升了模型的检测精度。实验结果表明,EDB-YOLO11 算法的参数量下降了 27.18%,计算量下降了 16.43%,同时在 VisDrone2019 数据集上的 mAP@0.5 提升了 3.4%,mAP@0.5:0.95 提升了 1.8%,在泛化数据集 SIMD 和 MAR20 上,mAP@0.5 分别提升了 3.8% 和 0.3%,mAP@0.5:0.95 分别提升了 3% 和 0.2%。上述实验结果验证了 EDB-YOLO11 算法在遥感小目标检测任务中的有效性与优越性。

**关键词:** 遥感小目标; YOLO11; 轻量化; 深度学习

**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

## Lightweight remote sensing image detection algorithm based on YOLO11

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

(1. School of Electronic Information 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:** To address the limitations of existing object detection algorithms in remote sensing small object detection—such as large model size, high computational complexity, and low detection accuracy—this paper proposes a lightweight remote sensing detection algorithm based on YOLO11, named EDB-YOLO11. The algorithm introduces improvements from two aspects: network architecture and loss function. First, the EMBC module is designed to replace the original C3K2 module, which effectively enhances the feature representation capability of the network. Second, a novel Downsample module is employed instead of the traditional convolution-based downsampling, which reduces the number of parameters and computational cost while improving the feature extraction ability of the backbone. Third, BiFPN is adopted to replace the original PANet for feature fusion, significantly enhancing the network's multi-scale feature integration efficiency. Finally, a new loss function called Focal-WIoU is proposed by combining the advantages of the Focal mechanism and WIoU loss. This design enables the model to focus more on high-quality training samples and reduce the impact of low-quality samples, thereby

收稿日期: 2025-01-09 Received Date: 2025-01-09

\* 基金项目: 国家自然科学基金(62373132)、河北省自然科学基金(F2025201023)、石家庄市驻冀高校基础研究项目(241791367A)、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项目(DXK202409)资助

improving overall detection accuracy. Experimental results show that EDB-YOLO11 reduces the number of parameters by 27.18% and the computational cost by 16.43%. On the VisDrone2019 dataset, the mAP@0.5 increases by 3.4%, and the mAP@0.5:0.95 improves by 1.8%. On the generalized remote sensing datasets SIMD and MAR20, the mAP@0.5 improves by 3.8% and 0.3%, respectively, while the mAP@0.5:0.95 improves by 3.0% and 0.2%, respectively. These results demonstrate the effectiveness and superiority of the proposed EDB-YOLO11 algorithm in remote sensing small object detection tasks.

**Keywords:** remote sensing small objects; YOLO11; lightweight; deep learning

## 0 引言

随着科学技术的快速发展,无人机和卫星空间技术取得了重大的突破并且得到了广泛的应用,遥感图像的分辨率得到了显著地提高,这也推动了遥感图像在多个领域的应用。然而,随之而来的还有图像中的遥感小目标的检测难题。遥感图像中的小目标主要包括行人、建筑物、道路、船舶、车辆等,遥感小目标在灾害监测<sup>[1]</sup>、城市规划、军事侦察<sup>[2]</sup>、商业应用<sup>[3]</sup>等领域都具有重要的研究意义,如何准确地检测出遥感图像中的小目标对象是一个急需解决的难题。

遥感小目标具有尺寸较小、形态复杂、存在背景遮挡和检测场景复杂等特点,尤其是在高分辨率的遥感图像中,小目标同时还会受到噪声、阴影、天气变化等其他因素的影响,从而影响到遥感小目标的检测精度,这使得传统的图像处理方法难以解决遥感小目标的检测难题。随着近年来深度学习的兴起,深度学习在计算机视觉任务中得到了广泛的应用<sup>[4-5]</sup>。通过深度学习模型架构强大的特征提取能力,能够学习到遥感小目标更加复杂的特征,有效地提高了遥感小目标的检测效果。因此,一大批学者致力于通过深度学习算法来解决遥感小目标的检测难题<sup>[6]</sup>。

基于深度学习的目标检测方法主要是将卷积神经网络应用于遥感图像,提取图像中有关遥感小目标的有效特征之后,再对遥感小目标进行分类与识别。现阶段的目标检测算法大致分为两类,一类是两阶段算法,另外一类是单阶段算法。两阶段算法的核心思想是将目标检测任务分为两个阶段,第1阶段生成候选区域,第2阶段对这些候选区域进行分类和回归。其中最经典的是 Girshick 等<sup>[7]</sup>于2014年提出的基于区域的卷积神经网络(region-based convolutional neural network, R-CNN)系列的无人机目标检测算法。相比于传统的特征提取方法, R-CNN 采用了卷积神经网络提取图像特征,具有明显的精度提升,同时适用于复杂背景环境下的小目标检测问题。但是 R-CNN 的计算成本和训练成本都非常高昂,基于此, Girshick 等<sup>[8]</sup>在2015年又提出了 Fast R-CNN, Fast R-CNN 采用共享卷积特征,极大地减少了模型的计算量,但是无法满足实时性要求。为了解决这个问题,在 Fast R-CNN 的基础上, Ren 等<sup>[9]</sup>于2015年提出了 Faster

R-CNN。Faster R-CNN 是一个端到端训练的目标检测框架, RPN 和 Fast R-CNN 共享卷积特征,并且候选区域生成和目标检测可以通过统一的损失函数进行联合优化,极大地提高了检测精度。但是 Faster R-CNN 的实时性和响应速度仍然较低,难以满足一些场景下的实时检测要求。两阶段的目标检测算法难以满足实时性的要求,因此,学者们设计出了单阶段检测算法。其主要特点是简化了检测流程,通过直接回归目标边界框并进行分类,避免了候选区域生成的步骤,通常能实现更快的检测速度。其中的经典之作作为 YOLO(you only look once)系列算法。YOLO 由 Redmon 等<sup>[10]</sup>于2016年提出,至今已经发布了多个版本<sup>[11-14]</sup>。YOLO 算法以其极快的检测速度闻名,可以做到实时检测应用,但同时 YOLO 算法在精度上往往低于两阶段算法,尤其在小目标检测和密集场景中表现不佳。面对 YOLO 系列算法检测精度不高的难题, Zhu 等<sup>[15]</sup>提出了 TPHYOLOv5 算法,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,但同时也带来了网络体积变大等问题。Wang 等<sup>[16]</sup>提出了 UAV-YOLOv8 网络,用于无人机场景下的小目标检测,以 WIoU 作为边界框回归损失提高定位能力,利用 BiFormer 注意力机制优化主干网络。然而,仍然无法解决无人机数据分布不均造成的分类问题。

针对遥感小目标检测效果差、网络模型复杂等检测难题,提出了一种基于 YOLO11 的改进算法 EDB-YOLO11。主要贡献与创新如下:1)采用了 EMSE 模块改进了 YOLO11 的主干部分,提升了网络的特征提取能力;2)采用 YOLOv7 的下采样方式替换了 YOLO11 原本的下采样模块,有效地降低了模型的计算量与参数量;3)采用了新的 Neck 结构,采用 BiFPN 特征融合方式改进了原本的 PANet,增强了网络的特征融合能力;4)引入了一种新的损失函数 Focal-WIoU,使模型更加关注高质量的样本,降低了低质量样本对于训练过程的影响,有效地提高了 EDB-YOLO11 模型对遥感小目标图像的检测精度。

## 1 YOLO11 算法概述

### 1.1 YOLO11 算法简介

YOLO11 是 Ultralytics 公司推出的 YOLO 系列的第

11 个版本,它是在 YOLOv8 的基础之上改进而来。相较于之前的 YOLO 版本,YOLO11 主要在以下方面进行了改进:首先,YOLO11 通过重新设计主干部分和颈部,增强了模型的特征提取能力。具体而言,YOLO11 采用了新的 C3K2 模块代替了原本的 C2f 模块,同时新增了 C2PSA 模块,有效地提升了模型从图像中提取特征的能力和最终的检测精度,同时降低了模型的参数量和计算量;其次,YOLO11 优化了训练流程,有效地提升了对于图像的处理速度。使得模型的训练和收敛速度加快;最后,YOLO11 支持多平台部署,包括云平台、边缘设备和移动设备。YOLO11 的网络结构与 YOLOv8 保持一致,由主干、颈部和头部 3 个部分组成。其中主干部分由

Conv、C3K2、SPPF、C2PSA 等模块组成,主要作用为从输入图像中提取图像的主要特征;颈部与 YOLOv8 保持一致,采用了 PANet 架构,其主要作用为对主干部分提取的特征进行特征融合,帮助网络更加精细地融合不同层次的特征;头部包含 3 个不同尺寸的检测头,分别对应小、中、大 3 类不同大小的检测目标,最终输出检测目标的类别、坐标、置信度信息。与 YOLOv8 相似,YOLO11 同样包含了 5 种不同尺寸的网络模型,它们按照从小到大的排序依次为 YOLO11n、YOLO11s、YOLO11m、YOLO11l、YOLO11x,其中 YOLO11s 以较小的网络体积保持了较高的检测精度,平衡了检测的精度与速度。因此,选取了 YOLO11s 作为基准模型,YOLO11 的网络结构如图 1 所示。

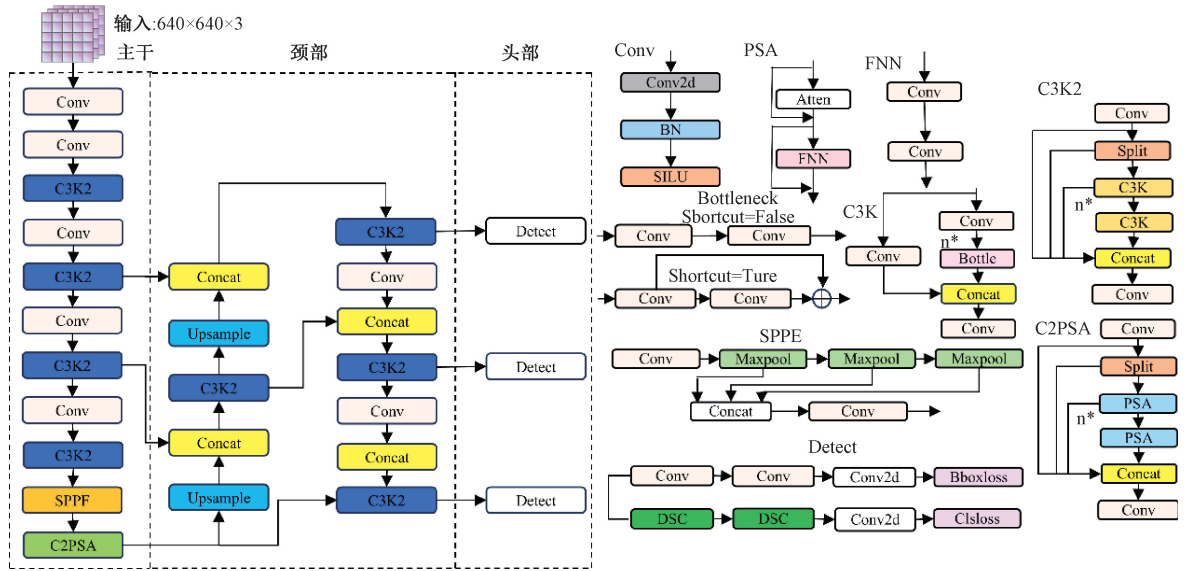


图 1 YOLO11 模型结构  
Fig. 1 YOLO11 model architecture

### 1.2 YOLO11 的不足之处

尽管 YOLO11 重新设计了网络结构并且优化了训练流程,有效地提升了检测精度和检测速度,但是当它面对遥感小目标图像的检测难题时,依然存在着不足与局限。虽然 YOLO11 的主干部分引入了新的 C3K2 模块和 C2PSA 模块,增强了主干部分的特征提取能力,但对于遥感小目标的检测精度提升不大。这主要是因为遥感小目标的像素占比较小,常规的下采样卷积操作在进行跨步卷积操作时会损失大量的有关遥感小目标的特征信息直接影响了最终的检测精度。同时,YOLO11 依旧采用了传统的 PANet 特征融合方式,这种单向的特征融合方式可能会造成部分特征信息的损失。此外,相比于 YOLOv8,YOLO11 虽然已经进行了模型的轻量化,减少了模型的参数量与计算量,但是网络体积依旧庞大,还有进一步提升的空间。

### 1.3 YOLO11 的改进方向

针对上述 YOLO11 在遥感小目标检测方向的不足,从以下方面对 YOLO11 进行了改进。首先,将 YOLO11 的传统下采样模块进行了替换,以改善主干部分在提取特征信息时出现的特征信息易丢失的问题;其次,对原模型中的 C3K2 模块进行了改进,提升了特征提取能力,然后,改进了原模型的特征融合方式,采用双向特征融合的方式替换了原本的单项特征融合方式,进一步提升了检测精度。最后,改进了原 YOLO11 的损失函数。

## 2 EDB-YOLO11 算法

### 2.1 EDB-YOLO11 算法简介

为了提升 YOLO11 算法在遥感小目标图像上的检测精度并且进一步轻量化网络模型,提出了一种基于

YOLO11 的轻量化遥感算法 EDB-YOLO11。首先,使用了 MobileNetv2<sup>[17]</sup> 中的 MBConv 与 EffectiveSE<sup>[18]</sup> 模块改进了 YOLO11 中的 C3k2 模块,有效地提升了网络的特征提取能力;其次,采用 YOLOv7<sup>[19]</sup> 中的 Downsample 下采样方式替换了原本网络中的普通卷积,有效地改善了遥感小目标易出现特征信息丢失的问题;然后,改进了 YOLO11 的特征融合方式,采用 BiFPN<sup>[20]</sup> 特征融合方式替换了原本的 PANet<sup>[21]</sup>,可以更好地融合有关小目标的特征信息,同时做到了模型的轻量化;最后,采用 Focal-WIoU 损失函数替换了原本的 CIoU 损失函数,有效地提升了检测精度。EDB-YOLO11 的网络结构如图 2 所示。

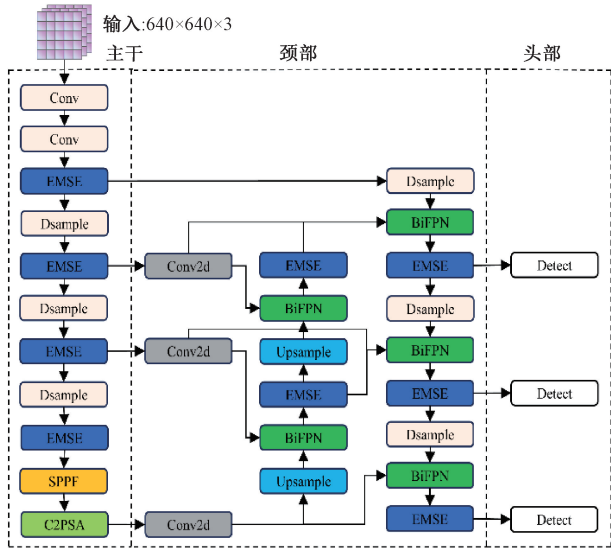


图2 EDB-YOLO11 模型结构  
Fig. 2 EDB-YOLO11 model structure

### 2.2 EMSE 模块

C3K2 模块是 YOLO11 的一项关键改进,它是基于 YOLOv8 的 C2f 模块和 YOLOv5 的 C3 模块改进而来,旨在增强模型的特征提取能力。C3K2 模块使用了不同尺寸的卷积核,传统的卷积操作所采取的卷积核一般是 3×3 大小的,而 C3K2 模块可以采用多种大小的卷积核,这一项改进使得 YOLO11 可以在不同的尺度上提取特征,有利于捕捉复杂的空间特征,有效地提升了模型的特征提取能力,同时做到了轻量化的设计。

但是当 YOLO11 面对遥感小目标的检测难题时,依旧会出现特征提取能力不足的情况。原因就在于小目标的特征一般比较微弱,其特征信息难以被有效地提取,需要更加精细的网络结构来进行小目标的特征提取,而 YOLO 算法的网络架构更适应较大目标的检测与识别。因此,采用 MobileNet2 网络结构中所提出的 MBConv 与 EffectiveSE 对 C3K2 模块进行了改进,得到了新的 EMSE 模块,有效地提升了网络的特征表达能力。EMSE 模块的结构如图 3 所示。SE 模块是由 SENet (squeeze-and-excitation networks) 提出的,通过引入通道注意力机制来增强模型的表现。SE 模块通过全局平均池化操作来对每个通道的特征进行压缩,然后通过全连接层来生成通道权重,这些权重再作用于每个通道的特征上,使得网络能够重点关注对任务最有帮助的特征。MBConv 是 MobileNet2 的重要创新之一,它基于倒置的残差结构,并且采用了深度可分离卷积和逐层线性瓶颈设置,同时引入了 SE 通道注意力机制,有助于保留更多的特征信息。因此采用 MBConv 模块替换 C3K2 中的 Bottleneck

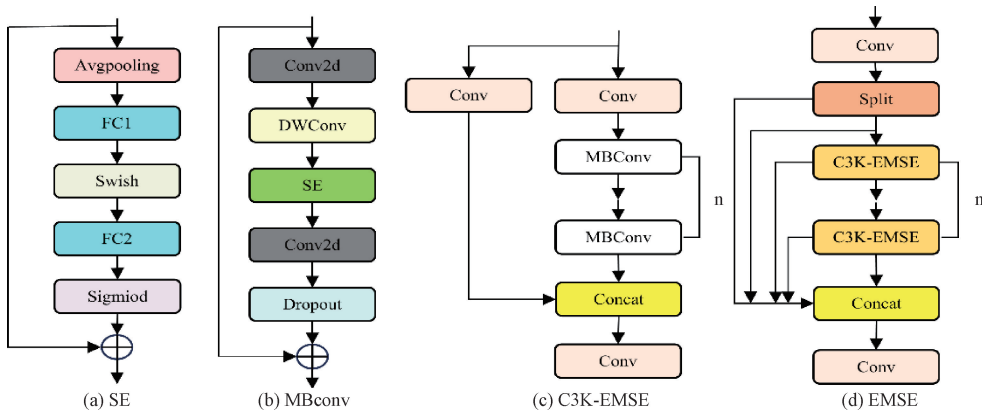


图3 EMSE 模块结构  
Fig. 3 Structure of the EMSE module

模块,形成了新的 EMSE 模块,引入了通道注意力机制和 MBConv 之后,提升了网络的特征表达能力和主干部分的特征提取能力。

### 2.3 Downsample 下采样方式

YOLO11 与 YOLOv8 采用了同样的下采样方式,主要是通过卷积操作进行下采样,并且结合了跨阶段局

部 (cross stage partial, CSP) 的思想, 这种结构更加注重网络的计算效率, 保证了模型的高精度和高推理速度, 尤其适用于较大对象的视觉检测任务。但是面对遥感小目标的检测任务时, 这种卷积下采样方式则存在一定的缺陷。原因就在于遥感小目标的特征较弱, 因为它们通常只占几百个像素点, 一些非常微小的小目标物体, 甚至只有几十个像素点, 这就导致遥感小目标可供学习的特征非常少。通过卷积进行下采样时可能无法提取到充足的有关小目标的特征信息, 这就导致网络针对遥感小目标对象的特征提取能力不足。因此, 采用了 YOLOv7 的 Downsample 下采样方式替换了原本的卷积下采样, 其结构如图 4 所示。Downsample 下采样具有两个分支, 其中一个分支由 1 个  $1 \times 1$  卷积和 1 个  $3 \times 3$  卷积组成, 而另外一个分支由  $2 \times 2$  最大池化操作和  $1 \times 1$  卷积组成。特征图同时经过两条分支后进行拼接, 然后输出特征信息。这种下采样方式的设计同时结合了卷积与池化的优势, 最大池化操作可以很好地保留图像的边缘信息特征, 而卷积则扩大了网络的感受野, 保留了更多的空间特征。此外, 模型可以同时获得来自卷积分路和池化分路的两种不同特征信息, 通过多层次的特征融合增强了网络的特征表达能力。

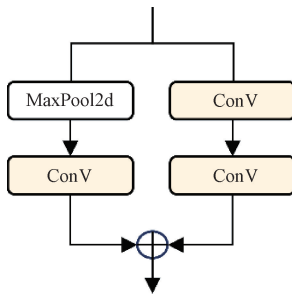


图 4 Downsample 下采样模块结构

Fig. 4 Structure of the downsampling module

## 2.4 双向特征金字塔网络 (BiFPN)

YOLO11 结合了 PAN 和 FPN 两种特征融合技术, 使得颈部可以从不同尺度的特征中抽取高质量的特征, 并进行多尺度的特征融合, 将高层语义信息与低层细节信息进行融合, 增强了模型对不同尺寸目标的检测能力, 有效地提升了模型的检测精度。虽然 YOLO11 的颈部支持多尺度特征融合, 但是只能支持单向的特征融合。具体来说, 这种颈部结构的设计只能支持从高分辨率特征图到低分辨率特征图的单向特征融合, 无法进行反向的特征融合。这种固定的特征融合方式可能会造成部分特征信息的损失。

为了解决这个问题, 需要改进 YOLO11 的颈部特征融合方式, 因此, 采取了 EfficientDet 中所提出的新的特征融合方式 BiFPN。BiFPN 是一种改进的特征金字塔网络

结构, 其核心思想就是支持双向的特征融合。传统的特征融合方式只支持自顶向下和自底向上两种特征融合方式, 每次融合只能进行单向的特征信息流动, 而 BiFPN 通过双向的特征融合, 允许特征信息在不同尺度之间进行传递, 使得不同尺度的特征信息可以进行双向的传递和融合, 保留了更加丰富的特征信息。BiFPN 双向特征融合方式的示意图如图 5 所示。

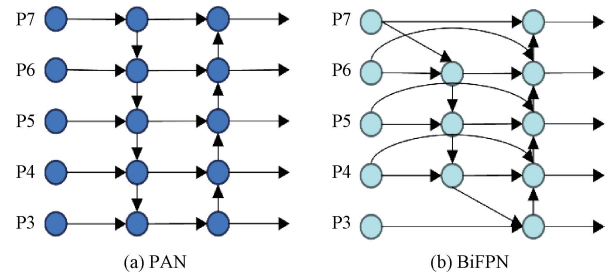


图 5 PAN 与 BiFPN 示意图

Fig. 5 Schematic diagram of PAN and BiFPN

## 2.5 Focal-WIoU 损失函数

YOLO11 的损失函数主要由边界框回归损失 (box loss)、分类损失 (classification loss) 和分布焦点损失 (distribution focal loss, DFL) 3 部分组成, 其中边界框回归损失是目标检测损失函数中的关键部分, 它的主要作用是优化预测边界框与真实边界框之间的差异。其主要目的是通过最小化中心点坐标损失和宽高损失, 使模型能够更准确地预测目标的位置和大小, 对于模型性能的提升有着至关重要的作用。IoU 和 IoU 损失函数的计算分别如式 (1) 和 (2) 所示。

$$IoU = \frac{|B \cap B^{gt}|}{|B \cup B^{gt}|} \quad (1)$$

$$L_{IoU} = 1 - IoU \quad (2)$$

式中:  $B^{gt}$  和  $B$  分别表示真实回归框和预测回归框。

YOLO11 采用的损失函数为 CIoU (complete intersection over union) 损失函数。其计算如式 (3) 所示。

$$L_{CIoU} = 1 - IoU + \frac{\rho^2(B, B^{gt})}{c^2} + av \quad (3)$$

式中:  $\rho^2(B, B^{gt})$  表示预测框中心点与真实框中心点之间的欧氏距离平方;  $c$  表示包含两个框最小外接框对角线长度;  $a$  为权重因子, 调节  $v$  的影响;  $v$  为宽高比的一致性度量。  $a$  和  $v$  的计算如式 (4) 和 (5) 所示。

$$a = \frac{v}{(1 - IoU) + v} \quad (4)$$

$$v = \frac{4}{\pi^2} \left( \arctan\left(\frac{w^{gt}}{h^{gt}}\right) - \arctan\left(\frac{w}{h}\right) \right)^2 \quad (5)$$

式中:  $w^{gt}$  和  $h^{gt}$  分别表示真实回归框的宽度和高度;  $w$  和  $h$  分别表示预测回归框的宽度和高度。CIoU 损失函数可

以确保模型在训练过程中更快收敛以及更稳定的回归。

遥感小目标图像具有目标尺寸小、密集分布、比例多样、目标与背景对比低、遮挡重叠严重等问题。因此,损失函数在训练阶段对于小目标定位精度和梯度分布的调节能力尤为关键。CIoU 损失函数忽略样本质量差异,容易影响高质量样本的收敛,且小目标样本的 IoU 本身波动大,CIoU 的中心距项和宽高比项影响有限。为了解决上述问题,对原本的损失函数 CIoU 进行了优化,引入了一种新的损失函数 Focal-WIoU 替换了原本的 CIoU 损失函数。

WIoU<sup>[22]</sup> 提出了一种基于质量评估动态加权的机制,更好地优化了边界框。WIoU 的计算方式如式(6)和(7)所示。

$$L_{\text{WIoU}} = R_{\text{WIoU}} L_{\text{IoU}} \quad (6)$$

$$R_{\text{WIoU}} = \exp\left(\frac{(x - x^{\text{gt}})^2 + (y - y^{\text{gt}})^2}{w_g^2 + H_g^2}\right) \quad (7)$$

式中:  $w_g$  和  $H_g$  分别表示最小封闭框的宽和高;  $x^{\text{gt}}$  以及  $y^{\text{gt}}$  分别表示真实回归框的中心点坐标。WIoU 损失函数更加关注中等质量的样本,忽略极差或极好样本的影响,在模型训练过程中可以更好地优化边界框。

而 Focal IoU<sup>[23]</sup> 的目的就是降低低质量预测框的梯度影响,提高训练鲁棒性通常是在其他 IoU 的基础之上加上聚焦因子(focal factor),Focal IoU 的计算方式如式(8)所示。

$$L_{\text{Focal IoU}} = (1 - \text{IoU})^\gamma \quad (8)$$

式中:  $\gamma$  为聚焦因子。

将 Wiou 和 Focal IoU 相结合得到了一种新的损失函数 Focal-WIoU 损失函数,其计算方式如式(9)所示。

$$L_{\text{Focal-WIoU}} = L_{\text{Focal IoU}} \cdot L_{\text{WIoU}} \quad (9)$$

Focal-WIoU 损失函数的引入有效地缓解了遥感小目标图像正负样本不均衡的问题,让模型更加关注高质量的样本,而降低了低质量样本的影响,有效地提升了模型对遥感小目标的检测精度。

## 3 实验结果及分析

### 3.1 实验环境及超参数设置

本文所采用的实验环境如表 1 所示,超参数设置如下:图像尺寸 Image size 设置为  $640 \times 640$ ,训练轮数 Epochs 设置为 300 轮,批大小 Batch size 设置为 16。此外,为保持实验环境一致,所有的实验均在同一硬件环境和相同的超参数设置下进行。

### 3.2 数据集及评价指标

为了验证所提出的模型在遥感小目标图像上的性能,采用了 VisDrone2019<sup>[24]</sup> 数据集来验证模型的有效性。

VisDrone2019 数据集由天津大学提出,是经典的遥感小目标数据集之一,共包含 8 629 张图像,其中训练集 6 471 张,验证集 548 张,测试集 1 610 张。VisDrone2019 数据集是由无人机在不同的光照条件下,在中国不同的城市拍摄所得,共包含 10 个类别的检测物体,它们分别是行人、人、自行车、汽车、面包车、卡车、三轮车、遮阳三轮车、巴士、摩托车。用于模型训练时,图像的分辨率统一为  $640 \times 640$ ,标注框超过 260 万个,其中分辨率小于  $32 \times 32$  像素的物体被称为小目标对象,而 VisDrone2019 数据集中的大多数对象为小目标物体。

表 1 实验环境设置

Table 1 Experiment environment settings

名称	配置信息
操作系统	Ubuntu 20.04
开发语言	Python3.10
深度学习框架	Cuda12.1+Pytorch2.1.0
GPU	RTX 3090(24G)×2
CPU	Intel Xeon Gold 6226R

实验采用了平均精度均值(mean average precision, mAP)和 mAP@0.5:0.95 以及精准度(precision, P)和召回率(recall, R)作为算法的精度评价指标,参数量和计算量作为模型的轻量化评价指标。各评价指标的计算如式(10)~(13)所示<sup>[25]</sup>:

$$P = \frac{TP}{TP + FP} \quad (10)$$

$$R = \frac{TP}{TP + FN} \quad (11)$$

$$\text{mAP@0.5} = \frac{1}{C} \sum_{i=1}^C \text{AP}_{i, \text{IoU}=0.5} \quad (12)$$

$$\text{mAP@0.5:0.95} = \frac{1}{N} \sum_{j=1}^N \text{AP}_{i, \text{IoU}=j} \quad (13)$$

模型的参数数量的定义是在训练过程中需要学习的权重和偏置的总数。计算量的定义是模型在训练或推理过程中所需的计算资源,通常用浮点运算次数(floating point operations, FLOPs)来衡量。

### 3.3 模型消融实验

为了分析所采用的 EMSE 模块、Downsample 下采样方式和 BiFPN 特征融合方式和 Focal-WIoU 损失函数对遥感小目标检测效果的影响,在 VisDrone2019 数据集上进行了消融实验,其中“√”表示添加了该模块,具体消融实验结果如表 2 所示。

从表 4 中的消融实验结果可以看出,相较于基准模型 YOLO11,提出的 EDB-YOLO11 在 VisDrone2019 数据集上的表现有显著提升。具体而言,EDB-YOLO11 在 mAP@0.5 指标上提升了 3.4%,在 mAP@0.5:0.95 指标

表 2 VisDrone2019 数据集消融实验结果

Table 2 Ablation results of the VisDrone2019 dataset

YOLO11s	EMSE	Downsample	BiFPN	Focal-WIoU	mAP@0.5	mAP@0.5:0.95	P	R	参数量/( $\times 10^6$ )	GFLOPs
✓					0.321	0.184	0.448	0.339	9.42	21.3
✓	✓				0.323	0.184	0.439	0.343	10.54	20.5
✓		✓			0.322	0.184	0.435	0.346	8.02	18.9
✓			✓		0.336	0.195	0.461	0.355	7.08	21.6
✓	✓	✓			0.329	0.187	0.455	0.346	9.14	18.0
✓	✓		✓		0.345	0.199	0.47	0.357	8.07	21.0
✓		✓	✓		0.336	0.195	0.462	0.351	<b>5.87</b>	18.4
✓	✓	✓	✓		0.348	0.2	0.475	0.359	6.86	17.8
✓				✓	0.328	0.186	0.453	0.343	9.42	21.3
✓	✓	✓	✓	✓	<b>0.355</b>	<b>0.202</b>	<b>0.485</b>	<b>0.363</b>	6.86	<b>17.8</b>

上提升了 1.8%,并且在精准度和召回率方面分别提升了 3.7%和 2.4%。这些提升表明,EDB-YOLO11 在检测精度上取得了明显的进步,同时还成功实现了模型的轻量化。实验结果显示,EDB-YOLO11 在减少模型计算量和参数数量的同时,依然能够有效提高性能,模型的参数量减少了 27.18%,计算量减少了 16.43%,这使得其在推理效率上得到了大幅优化。在所采用的各个模块中,BiFPN 对于精度提升的贡献最大,mAP@0.5 提升了 1.5%,同时参数量减少了 24.84%,并且计算量几乎保持不变,这表明 BiFPN 在提升检测精度的同时,不会显著增加计算负担。

此外,Downsample 下采样模块在降低参数量和计算量的同时,也带来了精度上的轻微提升,体现了该模块在模型压缩中的有效性。EMSE 模块则在保持精度稳定的基础上,成功地减少了计算量,尽管参数量略有增加,但其有效减轻了模型的计算压力,从而优化了推理效率。而所提出的 Focal-WIoU 损失函数在不增加模型计算开销的前提下,有效提升了 EDB-YOLO11 的检测精度。该损失函数结合了 Focal IoU 的梯度调控能力与 WIoU 的样本质量自适应加权机制,使模型在训练过程中能够更关注中等质量的目标框,同时抑制低质量预测对梯度更新的负面影响。实验结果表明,在引入 Focal-WIoU 损失函数后,EDB-YOLO11 在测试集上的 mAP@0.5 提升了 0.7%,而模型的参数量和计算量基本保持不变。这一结果充分说明了 Focal-WIoU 损失函数在提高检测性能的同时兼顾了模型的高效性,特别适用于对资源约束较高的遥感小目标检测场景。

综上所述,EDB-YOLO11 通过巧妙设计和模块优化,不仅在遥感图像中对小目标的检测效果上得到了显著提升,还实现了模型的轻量化,表现出了较高的实用性和应用潜力。

### 3.4 模型对比实验

为了验证所提出的 EDB-YOLO11 模型在目标检测

领域的先进性和优越性,本研究选取了近年来深度学习方向的几种经典且具有代表性的目标检测模型进行对比实验。这些模型包括:YOLOv7、YOLOv8、YOLOv9<sup>[26]</sup>、YOLOv10<sup>[27]</sup>、YOLOv12<sup>[28]</sup>以及 MHAF-YOLO<sup>[29]</sup>。为了全面评估模型的性能,本实验在广泛应用于无人机影像分析的 VisDrone2019 数据集上进行了对比。所有实验均在相同的硬件环境下进行,以保证结果的公平性和可比性。实验结果如表 3 所示,从而进一步体现 EDB-YOLO11 在检测精度、推理速度及鲁棒性等方面相较于其他经典模型的优势。

表 3 对比实验结果

Table 3 Comparison of experimental results

模型	mAP@0.5	mAP@0.5:0.95	参数量/ ( $\times 10^6$ )	GFLOPs
YOLOv7-tiny	0.303	0.154	6.04	13.3
YOLOv8s	0.32	0.184	11.13	28.5
YOLOv9s	0.346	0.204	9.75	39.6
YOLOv10s	0.318	0.181	8.04	24.5
YOLO11s	0.321	0.184	9.42	21.3
YOLOv12s	0.33	0.192	9.08	19.3
MHAF-YOLO	0.336	0.195	7.42	26.8
EDB-YOLO11s	<b>0.355</b>	<b>0.202</b>	<b>6.86</b>	<b>17.8</b>

从表 3 的对比实验结果可以得出,EDB-YOLO11 在精度指标 mAP@0.5 和 mAP@0.5:0.95 以及参数量和计算量指标上都取得了最优的结果。与其他的经典模型相比,EDB-YOLO11 在遥感小目标图像的检测效果更好,EDB-YOLO11 以更低的参数量和计算量保持了更高的精度,在提升检测精度的同时,做到了模型的轻量化,体现了 EDB-YOLO11 模型的先进性。

### 3.5 模型泛化实验

为了验证 EDB-YOLO11 模型在不同场景下的泛化能力,本研究选择了两个遥感图像数据集 SIMD<sup>[30]</sup>汽车数据集和 MAR20<sup>[31]</sup>军用飞机数据集,来评估 EDB-YOLO11 在遥感小目标检测中的表现。SIMD 汽车数据

集由巴基斯坦国立科学技术大学提出,专门用于车辆检测,包含了 5 000 幅遥感图像(每幅图像的尺寸为  $1\ 024 \times 768$  pixels)和 45 096 个汽车实例。该数据集中的汽车被细分为 15 个不同类别,涵盖了日常生活中常见的各类汽车类型。MAR20 数据集由西北工业大学提出,专门用于军用飞机目标检测。该数据集包括来自美国、俄罗斯等国家的 60 个军用机场采集的 3 842 张高分辨率遥感图像,图像中包含现役常见军用飞机的目标,共计 22 341 个飞机实例。MAR20 数据集的高分辨率和多样化的场景使其成为评估遥感图像中小目标检测算法,特别是在复杂背景和多样目标情况下表现的重要数据源。

为了全面评估 EDB-YOLO11 的性能,将原始的 YOLO11 模型与改进后的 EDB-YOLO11 模型分别在 SIMD 和 MAR20 数据集上进行了训练与测试,其中 SIMD 数据集训练了 200 轮, MAR20 数据集训练了 300 轮,实验数据如表 4 和 5 所示。

从表 4 和 5 的实验结果可以得出,相比于原模型,改进模型 EDB-YOLO11 在 SIMD 遥感汽车数据集上的  $mAP@0.5$  提升了 3.8%,  $mAP@0.5:0.95$  提升了 3%;在 MAR20 军用飞机数据集上的  $mAP@0.5$  提升了 0.3%,  $mAP@0.5:0.95$  提升了 0.2%,在保持轻量化模型的同时,精度均有小幅度的提升,验证了所提出的 EDB-YOLO11 算法的泛化性。

### 3.6 实验结果可视化与分析

为了验证所提出的 EDB-YOLO11 算法在实际场景中的检测效果,在 VisDrone2019 测试集上分别对原 YOLO11 和 EDB-YOLO11 算法进行了全面测试。为确保评估的多样性和实际意义,特意选择了具有挑战性的不同场景进行实验,具体包括车辆密集的高速公路、人员和车辆密集的十字路口、以及夜晚光照条件较差的停车场。这些场景均具备较高的复杂性,能够有效考察模型在各种实际条件下的表现。检测结果图如图 6 所示。



图 6 YOLO11 与 EDB-YOLO11 检测效果对比

Fig. 6 Comparison of the detection effect of YOLO11 and EDB-YOLO11

表 4 SIMD 数据集泛化实验结果

Table 4 Generalization experimental results on the SIMD dataset

模型	mAP@0.5	mAP@0.5:0.95	P	R
YOLO11	0.707	0.538	0.695	0.711
EDB-YOLO11	<b>0.745</b>	<b>0.568</b>	<b>0.740</b>	<b>0.685</b>

表 5 MAR20 泛化实验结果

Table 5 Generalization experimental results on the MAR20 dataset

模型	mAP@0.5	mAP@0.5:0.95	P	R
YOLO11	0.895	0.680	0.850	0.843
EDB-YOLO11	<b>0.898</b>	<b>0.682</b>	<b>0.851</b>	<b>0.837</b>

如图 6 所示,在复杂的遥感场景下,基准模型 YOLO11 存在一定程度的漏检与误检问题。具体而言,在车辆密集的高速公路场景中,YOLO11 经常漏检部分车辆,且存在将车辆类别错误分类的现象;在人员与车辆密集的十字路口场景中,模型将红绿灯误识别为行人,且对一辆自行车产生了多个重复检测框;而在夜间光照较差的停车场中,由于照明不足,模型检测结果进一步恶化,不仅将汽车误识别为厢式货车,还对一辆汽车产生了多个检测框。上述现象表明,YOLO11 在处理复杂背景、密集目标及低照度环境下的小目标检测任务时,鲁棒性仍有待提升。

与此相比,EDB-YOLO11 在这些复杂场景下展现出了显著的优势。通过引入针对性改进的模块,EDB-YOLO11 能够有效降低漏检率和错检率,显著提高了检测精度和召回率,尤其在目标密集和光照变化较大的场景中,其性能表现尤为突出。

这些实验结果证明,EDB-YOLO11 在面对具有挑战性的实际应用场景时,具备了出色的鲁棒性和适应性,特别是在复杂背景和环境,能够有效提升目标检测的精度和召回率,表现出了良好的实用性。因此,EDB-YOLO11 不仅提升了遥感图像中小目标的检测效果,也为实际应用中的复杂场景提供了可靠的解决方案。

## 4 结 论

针对现有的目标检测算法在面对遥感小目标的检测难题时存在的检测准确率低、模型参数量和计算量大、训练成本高昂等问题,提出了一种基于 YOLO11 的轻量级遥感算法 EDB-YOLO11。EDB-YOLO11 算法通过引入 EMSE 模块增强特征表达能力,优化下采样模块降低计算复杂度,采用 BiFPN 双向特征融合结构提升特征融合效果,并利用 Focal-WIoU 损失函数使模型聚焦于高质量样本。在 VisDrone2019、SIMD 和 MAR20 3 个遥感图像数

据集上的实验结果表明,EDB-YOLO11 算法在降低参数量和计算量、轻量化网络模型的同时,有效地提升了检测精度和遥感小目标的检测效果。本文的研究围绕遥感小目标检测精度不足、模型网络体积大这一核心思路展开,面临的主要技术难点是如何在提升小目标检测精度的同时控制模型的计算成本,所采用的核心技术包括网络结构重构(EMSE 模块与 BiFPN 融合结构)、高效特征采样设计(改进下采样模块)以及目标质量引导机制(Focal-WIoU 损失函数)。

## 参考文献

- [1] 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.
- [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 lightweight remote sensing image target detection method based on improved YOLOv7-tiny [J]. Journal of Electronic Measurement and Instrumentation, 2024, 38(7): 22-33.
- [3] 郝晋渊,张家明,张少康,等.基于改进 YOLO 的无人机入侵检测方法[J].电子测量与仪器学报,2024,38(7):143-151.  
HAO J Y, ZHANG J M, ZHANG SH K, et al. UAV intrusion detection method based on improved YOLO[J]. Journal of Electronic Measurement and Instrumentation, 2024, 38(7): 143-151.
- [4] 薛林雁,李轩昂,齐晔仪,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. A real-time multi-class detection method for colonoscopy polyps based on improved YOLOv5s [J]. Journal of Hebei University (Natural Science Edition), 2024, 44(4): 424-432.
- [5] 柯淇晟,花强,张博.基于改进 YOLOv8 的遥感图像目标检测[J].河北大学学报(自然科学版),2025,45(3):309-316.  
KE Q SH, HUA Q, ZHANG B. Remote sensing image object detection based on improved YOLOv8[J]. Journal of Hebei University (Natural Science Edition), 2025, 45(3): 309-316.

- [6] 李翔宇, 王伟, 王峰萍, 等. 面向密集场景结合 TC-YOLOX 的小目标检测方法[J]. 电子测量技术, 2023, 46(15): 133-142.
- LI X Y, WANG W, WANG F P, et al. Small target detection method for dense scenes based on TC-YOLOX [J]. Electronic Measurement Technology, 2023, 46(15): 133-142.
- [7] GIRSHICK R, DONAHUE J, DARRELL T, et al. Region-based convolutional networks for accurate object detection and segmentation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 38(1): 142-158.
- [8] GIRSHICK R. Fast R-CNN [C]. 2015 IEEE International Conference on Computer Vision, 2015: 1440-1448.
- [9] REN S, HE K, GIRSHICK R, et al. Faster R-CNN: Towards real-time object detection with region proposal networks[J]. Advances in Neural Information Processing Systems, 2015, 28: 91-99.
- [10] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: Unified, real-time object detection[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2016: 779-788.
- [11] REDMON J, FARHADI A. YOLO9000: Better, faster, stronger[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2017: 6517-6525.
- [12] REDMON J, FARHADI A. Yolov3: An incremental improvement [J]. arXiv preprint arXiv: 1804.02767, 2018.
- [13] BOCHKOVSKIY A, WANG C Y, LIAO H Y M. YOLOv4: Optimal speed and accuracy of object detection[J]. arXiv preprint arXiv:2004.10934, 2020.
- [14] LI C 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.
- [15] ZHU X K, LYU S C, WANG X, et al. TPH-YOLOv5: Improved YOLOv5 based on transformer prediction head for object detection on drone-captured scenarios [C]. IEEE/CVF International Conference on Computer Vision. 2021: 2778-2788.
- [16] WANG G, CHEN Y F, AN P, et al. UAV-YOLOv8: A small-object-detection model based on improved YOLOv8 for UAV aerial photography scenarios [J]. Sensors, 2023, 23(16): 7190.
- [17] SANDLER M, HOWARD A, ZHU M L, et al. Mobilenetv2: Inverted residuals and linear bottlenecks[C]. IEEE Conference on Computer Vision and Pattern Recognition. 2018: 4510-4520.
- [18] HU J, SHEN L, SUN G. Squeeze-and-excitation networks[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2018: 7132-7141.
- [19] 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]. IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023: 7464-7475.
- [20] TAN M X, PANG R M, LE Q V. Efficientdet: Scalable and efficient object detection [C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020: 10781-10790.
- [21] LIU S, QI L, QIN H F, et al. Path aggregation network for instance segmentation [C]. IEEE Conference on Computer Vision and Pattern Recognition, 2018: 8759-8768.
- [22] TONG Z J, CHEN Y H, XU Z W, et al. Wise-IoU: Bounding box regression loss with dynamic focusing mechanism [J]. arXiv preprint arXiv: 2301.10051, 2023.
- [23] ZHANG H, ZHANG S J. Focaler-iou: More focused intersection over union loss [J]. arXiv preprint arXiv: 2401.10525, 2024.
- [24] DU D W, ZHU P F, WEN L Y, et al. VisDrone-DET2019: The vision meets drone object detection in image challenge results [C]. IEEE/CVF International Conference on Computer Vision Workshops, 2019: 213-226.
- [25] PIDALLA 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), 2020: 237-242.
- [26] WANG C Y, YEH I H, MARK LIAO H Y. YOLOv9: Learning what you want to learn using programmable gradient information [C]. European Conference on Computer Vision. Springer, Cham, 2025: 1-21.
- [27] WANG A, CHEN H, LIU L H, et al. YOLOv10: Real-time end-to-end object detection [J]. arXiv preprint arXiv:2405.14458, 2024.
- [28] TIAN Y J, YE Q X, DOERMANN D. YOLOv12: Attention-centric real-time object detectors [J]. arXiv preprint arXiv:2502.12524, 2025.
- [29] YANG Z Q, GUAN Q, YU Z W, et al. MHAF-YOLO:

Multi-Branch heterogeneous auxiliary fusion YOLO for accurate object detection [J]. arXiv preprint arXiv: 2502.04656, 2025.

- [30] HAROON M, SHAHZAD M, FRAZ M M. Multisized object detection using spaceborne optical imagery [J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2020, 13: 3032-3046.

- [31] 禹文奇,程焱,王美君,等. MAR20:遥感图像军用飞机目标识别数据集 [J]. 遥感学报, 2023, 27(12): 2688-2696.

YU W Q, CHENG Z, WANG M J, et al. MAR20: A benchmark for military aircraft recognition in remote sensing images [J]. National Remote Sensing Bulletin, 2023, 27(12): 2688-2696.

## 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,现为河北大学电子信息工程学院教授,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 his B. Sc. degree from Hebei University, and M. Sc. degree from North China Electric Power University, and Ph. D. degree from Zhejiang University. He is now a professor at College of Electronic and Information Engineering, 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

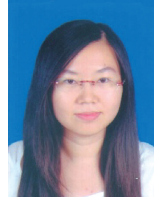
**Shi Gaolang** received his B. Sc. degree from Yangtze University in 2023. He is now a M. Sc. candidate at Hebei University. His main research interests include deep learning and object detection.



张昊侯,2022年于铜陵学院获学士学位,现为河北大学在读研究生,主要研究方向为深度学习、图像分割。

E-mail: 2892568146@qq.com

**Zhang Haoyu** received his B. Sc. degree from Tongling University in 2022. He is now a M. Sc. candidate at Hebei University. His main research interests include deep learning and image segmentation.



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

**Hao Jinyuan** (Corresponding author) received her B. Sc. degree from Hebei University in 2013, and M. Sc. degree from the University of New South Wales in 2015. She is now a lecturer with the HBU-UCLan School of Media, Communication and Creative Industries, Hebei University. Her main research interests include artificial intelligence, information technology and new media communication.