

光电工程

Opto-Electronic Engineering

中文核心期刊 中国科技核心期刊
Scopus CSCD

特征协同与细粒度感知的遥感图像小目标检测

肖振久, 张杰浩, 林瀚翰

引用本文:

肖振久, 张杰浩, 林瀚翰. 特征协同与细粒度感知的遥感图像小目标检测[J]. 光电工程, 2024, 51(6): 240066.

Xiao Z J, Zhang J H, Lin B H. Feature coordination and fine-grained perception of small targets in remote sensing images[J]. *Opto-Electron Eng*, 2024, 51(6): 240066.

<https://doi.org/10.12086/oe.2024.240066>

收稿日期: 2024-03-20; 修改日期: 2024-04-25; 录用日期: 2024-04-26

相关论文

基于改进YOLOv5s的无人机图像实时目标检测

陈旭, 彭冬亮, 谷雨

光电工程 2022, 49(3): 210372 doi: [10.12086/oe.2022.210372](https://doi.org/10.12086/oe.2022.210372)

基于多尺度特征融合的遥感图像小目标检测

马梁, 苟于涛, 雷涛, 靳雷, 宋怡萱

光电工程 2022, 49(4): 210363 doi: [10.12086/oe.2022.210363](https://doi.org/10.12086/oe.2022.210363)

面向遥感图像检索的级联池化自注意力研究

吴刚, 葛芸, 储珺, 叶发茂

光电工程 2022, 49(12): 220029 doi: [10.12086/oe.2022.220029](https://doi.org/10.12086/oe.2022.220029)

基于直接散射光谱的多环境要素激光遥感方法

郭圆新, 梁琨, 徐杨睿, 赵延鹏, 王元庆, 王蓓

光电工程 2024, 51(3): 240003 doi: [10.12086/oe.2024.240003](https://doi.org/10.12086/oe.2024.240003)

更多相关论文见光电期刊集群网站 



光电工程
Opto-Electronic Engineering

<http://cn.ojournal.org/oe>



OE_Journal



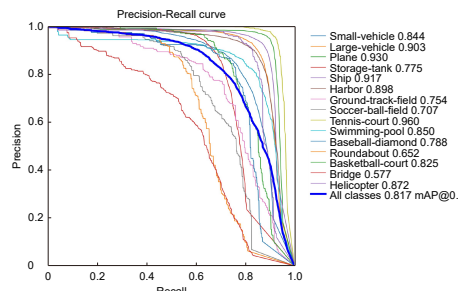
Website

DOI: 10.12086/oe.2024.240066

特征协同与细粒度感知的 遥感图像小目标检测

肖振久, 张杰浩*, 林瀚翰

辽宁工程技术大学软件学院, 辽宁 葫芦岛 125105



摘要: 针对遥感图像中小目标多、排列密集导致的漏检问题, 提出一种特征协同与细粒度感知的遥感图像小目标检测算法。首先, 构造精细特征协同策略, 通过智能调整卷积核参数, 优化了特征间的交互和整合过程; 通过精确控制信息流, 实现从粗糙到精细的渐进式特征精化。在此基础上, 本文设计一个细粒度感知模块, 将感知注意力与移动反向卷积结合形成一个增强型检测头, 显著增强网络对于极小尺寸物体的感知能力。最后, 为了提升模型训练的效率, 采用 MPDIoU 和 NWD 作为回归损失函数, 解决位置偏差, 加快模型收敛。在 DOTA1.0 数据集和 DOTA1.5 数据集上的实验结果表明, 改进后算法相比于基准方法, 平均精度分别提高 7.4% 和 6.1%, 相较于其他算法具有明显优势, 显著改善遥感图像中小目标的漏检情况。

关键词: 遥感图像; 小目标检测; 特征协同; 细粒度感知

中图分类号: TP391.4

文献标志码: A

肖振久, 张杰浩, 林瀚翰. 特征协同与细粒度感知的遥感图像小目标检测 [J]. 光电工程, 2024, 51(6): 240066

Xiao Z J, Zhang J H, Lin B H. Feature coordination and fine-grained perception of small targets in remote sensing images[J]. Opto-Electron Eng, 2024, 51(6): 240066

Feature coordination and fine-grained perception of small targets in remote sensing images

Xiao Zhenjiu, Zhang Jiehao*, Lin Bohan

School of Software, Liaoning University of Engineering and Technology, Huludao, Liaoning 125105, China

Abstract: Addressing the challenge of missed detection caused by many small targets and dense arrangement in remote sensing images, this study introduces a small target detection algorithm for remote sensing applications, leveraging a combination of feature synergy and micro-perception strategies. Initially, we propose a refined feature synergistic fusion strategy that optimizes the interaction and integration of features across different scales by intelligently adjusting the parameters of convolution kernels. This strategy facilitates progressive refinement of features from coarse to fine granularity. Building upon this foundation, a micro-perception unit is developed in this paper, incorporating perceptual attention mechanisms with moving inverse convolution to form an advanced detection head. This innovative approach substantially boosts the network's capability to detect very small objects. Furthermore, to augment the training efficiency of the model, we employ MPDIoU and NWD as regression loss functions, mitigating positional bias issues and expediting model convergence. Experimental evaluations on the DOTA1.0 dataset and DOTA1.5 dataset reveal that our algorithm achieves a substantial improvement in mean

收稿日期: 2024-03-20; 修回日期: 2024-04-25; 录用日期: 2024-04-26

基金项目: 辽宁省高等学校基本科研项目 (LJKMZ20220699); 辽宁工程技术大学学科创新团队项目 (LNTU20TD-23)

*通信作者: 张杰浩, zjhao0409@163.com

版权所有©2024 中国科学院光电技术研究所

Average Precision (mAP) by 7.4% and 6.1% over the baseline method, which has obvious advantages over other algorithms. The results underscore the algorithm's efficacy in significantly reducing the incidence of missed detections of small targets within remote sensing imagery.

Keywords: remote sensing images; small target detection; feature synergy; fine-grained aware

1 引言

随着遥感图像技术的快速发展, 遥感图像目标检测广泛应用于许多重要领域, 包括军事目标定位和识别、自然环境保护、灾害探测和城市规划建设。遥感图像目标检测的任务是精准辨识和定位图像中的特定目标, 并推测其类型和位置。遥感图像中的目标具有场景大、目标小、多尺度、背景复杂、重叠遮挡等特点, 同时光照、天气等自然环境因素影响, 会导致图像颜色失真、目标细节丢失或模糊, 因此准确检测特定物体是一项具有挑战性的任务。现阶段遥感图像目标检测算法已取得很大突破, 但对于小目标检测效果仍不理想, 小目标检测面临两大难点: ①目标特征信息少, 正样本稀缺、分类不均衡; ②目标定位困难, 背景复杂, 含有大量冗余信息, 对检测造成严重的干扰。这使得从航拍图像中提取小目标的边缘特征, 并将目标与背景区分开来成为一项具有挑战性的工作, 因而遥感图像中的小目标检测与应用研究具有重要理论与实际意义^[1-6]。

近年来, 对自然场景图像进行多分类的目标检测有很多优秀方法, 并逐渐发展出两个分支: 一是追求实时检测的单阶段目标检测方法, 如 SSD^[7] (single shot multibox detector)、RetinaNet^[8] (retina convolutional neural network)、YOLO^[9-13] (you only look once) 系列; 另一个是追求高精度的两阶段目标检测方法, 如 Fast R-CNN^[14] (fast region-based convolutional neural network)、R-FCN^[15] (region-based fully convolutional networks) 等算法。由于遥感图像尺度大、背景复杂多变、拍摄视角独特等原因, 自然场景下的目标检测算法不能很好地识别遥感图像中的小目标特征。

为了提高小目标的检测能力, Li^[16] 等在 YOLOv8-s 模型主干中使用 GhostblockV2 结构来代替部分 C2f (context to focus) 模块, 抑制在长距离特征传输过程中的信息丢失。Zhu 等^[17] 改进 BiFPN (双向特征金字塔网络) 代替 YOLOv4 模型中原有 PANet (path aggregation network) 作为特征融合模块, 通过共享权值方式实现多尺度特征融合。但上述方法都无法精准

定位小目标, 存在漏检问题。Zhai 等^[18] 在 YOLOv8 模型探测头组件中, 增加高分辨率探测头, 提高设备对小目标的探测能力, 同时砍掉大目标探测头和冗余网络层, 提高无人机探测速度。Zhou 等^[19] 在基于 ConvNext (convolutional neural network next) 和 Transformer 设计的 C3CNTR (ConvNext 和 Transformer 设计的区域建议网络) 模块中, 在预测头之前从局部和全局角度澄清目标在特征图中的位置。但他们忽略了周围背景干扰, 因此对小目标检测准确性不足。Zhu 等^[20] 在 SGA-GAN (spatial and global attention generative adversarial network) 网络上提出一个多维自注意特征选择模块, 通过串联跨注意力模块以及空间和通道注意力模块来引导生成器更好地利用全局信息。Xiao 等^[21] 在 ResNet50 网络上设计了一个双向多尺度特征融合网络, 通过融合语义特征和浅层特征, 以提高复杂背景下小物体检测效果。Wu 等^[22] 提出了一种具有旋转边界盒 (RBB) 的旋转自适应 YOLOv5 来检测任意方向的柔性对象。但他们的泛化能力有所局限, 收敛能力较差。Yang 等^[23] 通过添加 IoU 常数因子改进平滑 L1 损失, 来解决旋转目标边界框回归的边界问题。Zhang 等^[24] 通过预测物体的方向向量来确定物体的唯一方向, 实现物体检测的同时准确地估计物体的方向。Yu 等^[25] 提出空间变换解耦, 增强堆叠的 Transformer 块提取的特征, 有效地抑制背景信息, 同时突出显示前景对象。但增加了额外的计算量和参数量。

针对上述挑战, 本文在 YOLOv8 算法的基础上, 提出了一种特征协同与细粒度感知的遥感图像小目标检测算法。首先, 引入一种精细特征协同策略 (fine feature synergy, FFS), 通过 KernelWarehouse^[26] 智能化地调整卷积核参数, 优化了不同层级特征间的交互和整合, 实现了从粗糙到精细的特征逐步精化, 大幅提高了模型对复杂背景和微小目标的敏感性和响应速度。继而, 设计了一个细粒度感知模块 (fine-grained aware, FGA), 结合感知注意力机制和移动反向卷积技术, 构建了一个增强型检测头, 这一设计显著提升了对极小尺寸物体的感知能力。此外, 为了提高模型训

练效率, 本研究采用了 MPDIoU^[27] (modified progressive distance IoU) 和 NWD^[28] (normalized weighted distance) 结合作为回归损失函数, 有效解决了位置偏差问题, 加快了模型收敛速度。

2 基于特征协同与细粒度感知的遥感图像小目标检测

针对当前传统网络对遥感小目标图像存在漏检、识别精度低等问题, 本文提出一种特征协同与细粒度感知的遥感图像小目标检测算法。该算法的网络结构如图 1 所示, 主要有 3 个改进部分: 1) 精细特征协同模块, 通过 KernelWarehouse 动态卷积和 C2f_KW (C2f_KernelWarehouse) 模块的结合来聚焦小目标特征, 捕捉不同尺度特征信息; 2) 细节感知单元: 引入感知注意力与移动反向卷积 (mobile inverted residual bottleneck convolution, MBConv^[29]) 结构重构检测头, 改善对微小物体检测效果不佳的问题, 提升模型精度; 3) 为提升网络模型的边界框回归性能, 将 MPDIoU 和 NWD 结合起来作为回归损失函数, 以提高模型的收敛能力, 加快模型训练速度。

2.1 精细特征协同模块

为了强化模型对遥感图像中复杂背景和细小目标的识别能力, 采用一种从粗糙到精细的渐进式特征协同机制。首先, 使用 CSPDarkNet (cross stage partial DarkNet) 网络作为骨干网络提取输入图像的特征 X , 但是该网络所采用的传统卷积运算具有卷积核固定、感受野小等特点, 仅能提取局部特征, 在存在多尺度目标时易出现漏检现象。因此, 本文在主干网络部分引入特征提取能力强的动态卷积, 通过内核分区和仓库共享来增强同一层内和跨连续层的卷积参数依赖性, 使多个相邻卷积层可以共享同一个仓库, 根据当前输入的重要性自适应地调整卷积核权重, 在不同时间步骤上聚焦于关键区域, 增强主干网络特征提取能力, 提高对小目标的感知和检测性能, 减少漏检现象。具体结构如图 2 所示, 首先使 ConvNet 的任何规则卷积层的静态内核具有相同维度, 然后将每个内核单元计算为线性混合 w_i :

$$w_i = \alpha_{i1}e_1 + \cdots + \alpha_{in}e_n, \quad (1)$$

其中: $e_1, \dots, e_n$ 为 n 个相同尺寸的内核单元, $\alpha_{i1}, \dots, \alpha_{in}$ 为新的注意力函数计算所依赖的输入标量, 其被共享给所有相同阶段的卷积层, 并且最终通过按

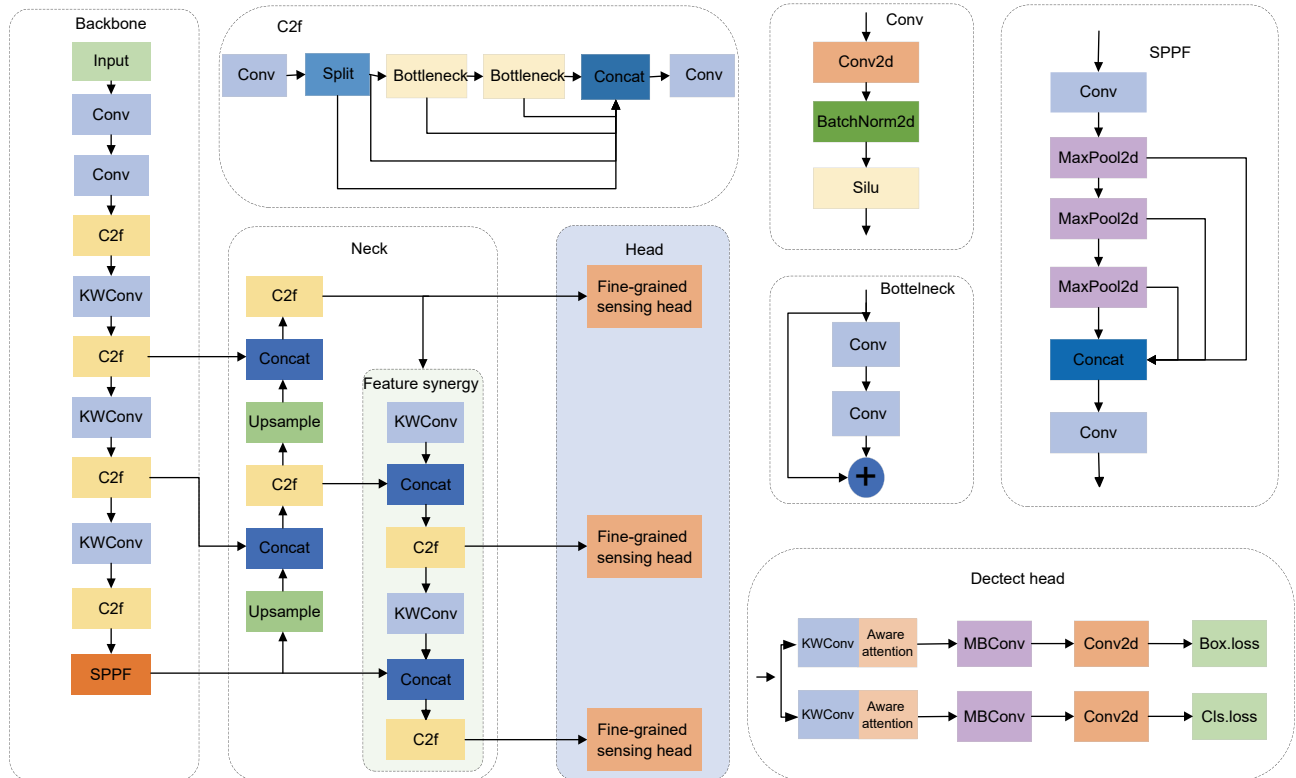


图 1 模型总体结构

Fig. 1 Overall model structure

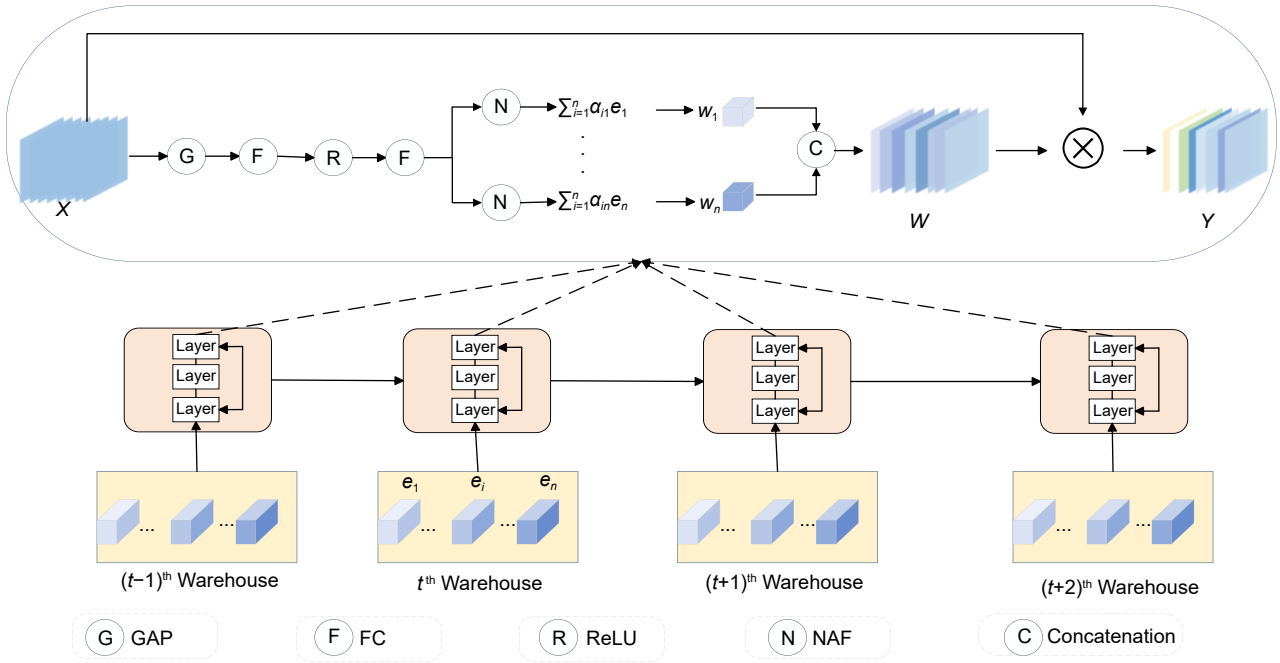


图 2 KernelWarehouse 的结构图

Fig. 2 Structure of KernelWarehouse

顺序组装其对应的 n 个混合来替换静态内核 W ，从而产生高自由度以拟合期望的卷积参数预算。

通过调整后的卷积核 W' ，对输入特征图 X 进行卷积操作，以实现特征的交互和整合。此外，为了实现从粗糙到精细的特征精化，引入一个信息流控制机制，通过权重系数 α 来调节不同层次特征图的贡献度：

$$Y_{\text{refined}} = \alpha \cdot Y + (1 - \alpha) \cdot X, \quad (2)$$

其中： Y_{refined} 是最终精化的特征图， α 是通过学习得到的权重系数，用于平衡原始特征图 X 和卷积后特征图 Y 的贡献。最后，为了进一步提升模型对于细小目标的检测能力，本文引入一个特征协同机制，其具体结构如图 3 所示，首先通过 KernelWarehouse 提高同一层内部及相邻层间卷积参数依赖关系，减少梯度冗余对网络性能的影响，以增加网络对复杂背景下小目标特征提取能力。然后使用一组动态卷积对输入特征进行滤波操作，再用 Split 操作进行特征切割，通过层叠 KW (KernelWarehouse) 卷积模块和跳跃连接来提高网络感受野，最后 Concat 操作将特征图进行连接，将 KW 卷积层输出特征映射到目标类别的预测结果，在保证模型计算量较小的前提下，获得更丰富的梯度流结构，提取更多的小目标特征，有效提升模型对复杂场景中小目标识别能力。

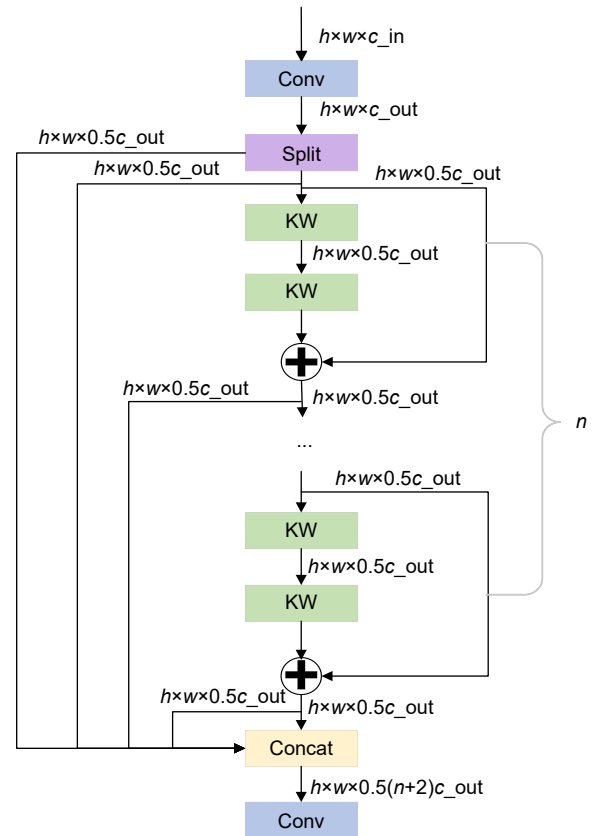


图 3 特征协同模块

Fig. 3 Feature synergy module

2.2 细粒度感知模块

为了进一步提升模型对小目标的检测性能，特别

是在处理遥感图像中密集小目标方面的能力, 提出一种细粒度感知增强的检测头, 其具体结构如图 4 所示, 该模块通过细节感知注意力机制聚焦于目标的细粒度信息, 同时通过移动反向卷积技术扩大网络感受野, 增强对小物体特征的捕捉能力。

2.2.1 感知注意力机制

注意力机制可以视为神经网络改进的关键途径, 本文通过将注意力函数 NAF^[30] (neural autoregressive flows) 与动态卷积结合, 提出了一种改进的感知注意力机制, 以感知每个内核单元的重要性权重, 加强网络对小目标密集图像特征的提取, 进一步提高内核单

元表达能力和适应性。其具体结构如图 5 所示, 对于特征向量 x 到叶子路径由红线描绘, 特征向量 x_j 以及对应输出值 y_j , Att 表示具有共享权重 θ (网络参数) 的 T 个网络。从图 5 中可以看出, 第 k 个神经网络实现了注意力操作, 并计算向量 $A_k(x)$ 和值 $B_k(x)$, 最终 GAtt 聚合所有键 $A_k(x)$ 和值 $B_k(x)$ 。该机制可以作用于具有静态内核 W 的任何卷积层, 特征向量 x 通过全局平均池化 (GAP) 将特征图像尺寸降低到固定大小, 通过全连接 (FC) 层减少特征向量长度, 提取图像中小目标的高级特征, ReLU 激活函数提高网络泛化能力和抗过拟合能力, 最终由 NAF 注意力函数逐组归

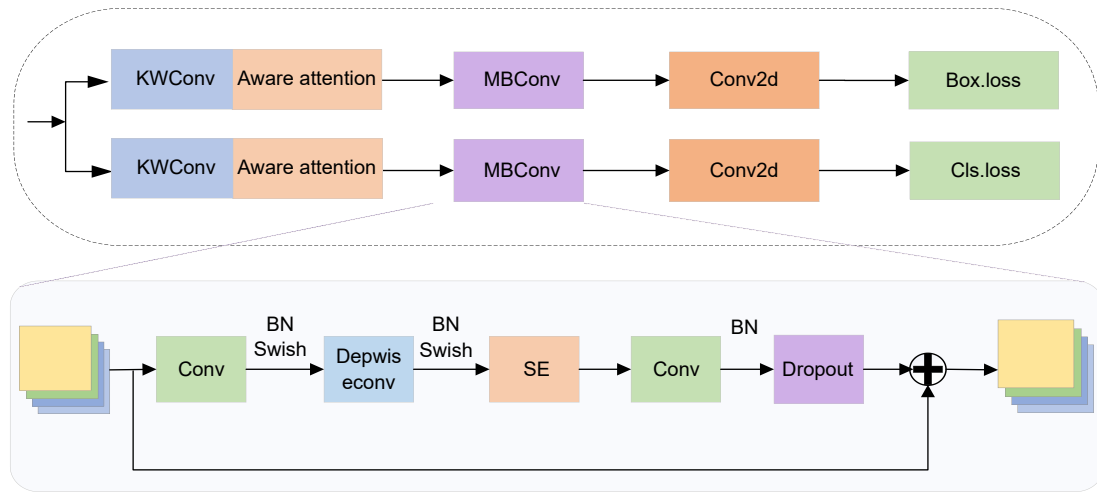


图 4 细粒度感知检测头

Fig. 4 Fine-grained aware detect head

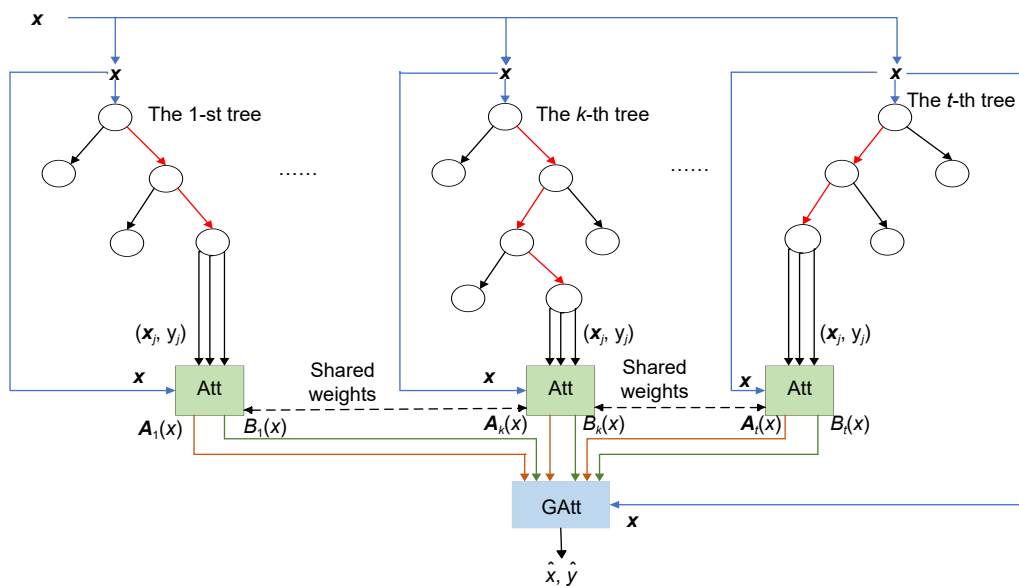


图 5 感知注意力机制

Fig. 5 Aware attention mechanism

一化, 强化遥感图像中的目标特征, 减少小目标信息损失, 提高模型性能。

2.2.2 移动反向卷积

YOLOv8 的检测头能够多尺度对目标进行检测, 但对微小目标的检测仍存在检测能力不佳的现象。为了提高 YOLOv8 检测性能并兼顾模型准确率, 受 EfficientNet^[31] (efficient neural network) 网络的启发, 将移动反向卷积模块重构到网络的检测头中。首先对每个通道进行单独的卷积计算, 再进行通道组合, 显著减少计算量, 提高模型效率。倒残差结构通过使用扩张卷积增加通道数, 逐点卷积进行降维, 实现了灵活的特征变换, 能处理不同尺度、不同分辨率的遥感图像, 尤其是对于背景复杂、目标密集的场景, 能够更加准确地进行定位, 提高模型性能。

MBConv 结构如图 4 所示, 首先输入特征图的通道数通过一个 1×1 卷积进行扩张, 以提取更丰富的特征表示, 捕捉更广阔的上下文信息。接着应用深度可分离卷积 (depth-wise convolution), 将扩张后的特征图分别进行独立卷积操作, 能够更好地处理和强化小目标特征表示, SE 模块通过动态学习通道间的相关性, 对特征图进行加权, 从而提高小目标的可分辨性和识别准确性。利用 MBConv 对多尺度信息的学习能力, 以及其对小尺寸物体的较好处理能力, 能够更好地适应不同尺度和大小的目标, 显著提升 YOLOv8 检测性能。

2.3 损失函数的优化

在 DOTA 遥感图像中, 针对传统 IoU 存在小目标定位偏差的问题, 本文将采用 NWD 和 MPDIoU 结合起来计算 IoU 部分, 作为回归损失函数, 以度量预测边界框与真实目标边界框之间的重叠程度, 提高模型收敛性。

2.3.1 NWD loss

IoU 对小目标非常敏感, 在小的位置偏移时, 会引起 Anchor 标签偏移, 从而造成两个样本之间特性相近, 难以收敛, 因此本文引入一种新的度量标准, 使用提出的 NWD (normalized wasserstein distance) 来度量导出的高斯分布的相似性代替标准 IoU。 (c_x, c_y) 表示中心坐标、 w 表示宽度和 h 表示高度的水平边界框 $R = (c_x, c_y, w, h)$ 可以被建模为 2D 高斯分布 $N(\mu, \Sigma)$, 其中:

$$\mu = \begin{bmatrix} c_x \\ c_y \end{bmatrix}, \Sigma = \begin{bmatrix} \frac{w^2}{4} & 0 \\ 0 & \frac{h^2}{4} \end{bmatrix}. \quad (3)$$

接着计算边界框 $A = (c_{xa}, c_{ya}, w_a, h_a)$ 和边界框 $B = (c_{xb}, c_{yb}, w_b, h_b)$ 建模的高斯分布 N_a 和 N_b 为

$$W_2^2(N_a, N_b) = \left\| \begin{bmatrix} c_{xa}, c_{ya}, \frac{w_a}{2}, \frac{h_a}{2} \\ c_{xb}, c_{yb}, \frac{w_b}{2}, \frac{h_b}{2} \end{bmatrix} \right\|_2^2. \quad (4)$$

由于无法直接量化边界框 N_a 和 N_b 之间的相似度, 将其归一化获得新的度量 NWD 为

$$NWD(N_a, N_b) = \exp\left(-\frac{\sqrt{W_2^2(N_a, N_b)}}{C}\right), \quad (5)$$

其中: C 是与数据集密切相关的常数。IoU 计算是基于像素级别的重叠程度, 对小目标检测不够敏感, NWD 对较大的重叠部分赋予更高权重, 对较小的重叠部分赋予较低权重, 可以减少边界框重叠度量的局限性, 更适合于处理复杂场景和小目标检测。

2.3.2 MPDIoU loss

MPDIoU 简化了两个边界框之间的相似性比较, 适应重叠或非重叠边界框回归。MPDIoU 的定义式:

$$MPDIoU = \frac{A \cap B}{A \cup B} - \frac{(x_1^B - x_1^A)^2 + (y_1^B - y_1^A)^2}{w^2 + h^2} - \frac{(x_2^B - x_2^A)^2 + (y_2^B - y_2^A)^2}{w^2 + h^2}, \quad (6)$$

其中: (x_1^A, y_1^A) , (x_2^A, y_2^A) 表示目标边界框 A 左上和右下点坐标, (x_1^B, y_1^B) , (x_2^B, y_2^B) 表示目标边界框 B 左上和右下点坐标, w 为输入图像宽度 h 为高度。

MPDIoU 作为损失函数, 表示式为

$$L_{MPDIoU} = 1 - MPDIoU. \quad (7)$$

结合上述两个损失后得到的最终回归损失为

$$L_{reg} = \alpha \cdot NWD + \beta \cdot L_{MPDIoU}. \quad (8)$$

通过实验分析, 最终取 $\alpha = 0.5$, $\beta = 0.5$, 结合后的损失函数使得模型训练效率得到明显提升。

3 实验结果与分析

3.1 数据集及实验环境

DOTA1.0 数据集是一个用于目标检测和分割的大规模遥感图像数据集。它包含 2806 个高分辨率遥感图像, 每个图像大小为 800×800 到 4000×4000 像素不等, 共有飞机 (PL)、轮船 (SH)、港口 (HA)、桥梁 (BR)、储罐 (ST)、直升机 (HC)、网球场 (TC)、篮球场 (BC)、足球场 (SBF)、棒球场 (BD)、田径场 (GTF)、游泳池 (SP)、大型车辆 (LV)、小型车辆 (SV)、环形

交叉路口 (RA) 十五个类别。与 DOTA1.0 数据集相比, DOTA1.5 数据集在此基础上增加了集装箱起重机 (CC) 类别。由于 DOTA 数据集图像尺寸较大, 很难直接用于网络训练, 因此对图像进行预处理, 将图像裁剪成大小为 640×640 像素的图像, 按照 7:1:2 的比例划分为训练集、验证集和测试集。

本实验使用的操作系统是 Windows 10, GPU 是 NVIDIA GeForce RTX 3090, 编程语言为 Python 3.8, 深度学习框架为 Pytorch 1.10.1, 设置 200 epoch。

3.2 评价指标

本实验用精度 Precision、召回率 Recall、平均精度 AP 和速度 FPS 来对模型进行评价, 表示检测效果。TP 为正确检测目标数目, FP 为误检目标数目, FN 为漏检目标数目。对应计算公式为

$$Precision = \frac{TP}{TP+FP}, \tag{9}$$

$$Recall = \frac{TP}{TP+FN}, \tag{10}$$

$$AP = \int_0^1 Precision(Recall)d(Recall), \tag{11}$$

$$mAP = \frac{1}{n} \sum_{i=1}^n \int_0^1 Precision(Recall)d(Recall), \tag{12}$$

$$FPS = \frac{Framenum}{ElapsedTime}. \tag{13}$$

3.3 消融实验

本文在 YOLOv8 原模型基础上进行改进, 在 DOTA1.0 数据集上进行检测, 对于改进部分进行消融实验, 实验所有数据参数和环境配置严格一致。实验结果如表 1 所示, A、B、C、D 分别表示应用动态

卷积 KW 模块、采用改进 Detect_MBConv 检测头、替换 NWD 和 MPDIoU 损失函数。

第一组实验为未改进的 YOLOv8 算法, 作为之后实验的对比基准, Precision 为 79%, Recall 为 68.7%, mAP@0.5 为 74.3%, mAP@0.5:0.95 为 50.9%。

第二组实验仅使用动态卷积 KW 提高模型的特征提取能力, 同时保证参数效率, 准确率提高 3.1%, 召回率提高 1.5%, mAP@0.5 提高 2.2%, mAP@0.5:0.95 提高 1.6%。

第三组实验仅使用 Detect_MBConv 检测头, 在进行特征提取时, 能够有效减少信息丢失, 在保持高效性能的同时, 具有更好的适应性, 准确率提升了 5.4%。

第四组实验仅将损失函数替换成 NWD 度量, 提高训练速度的同时准确率提升 1.2%。

第五组实验仅将损失函数替换成 MPDIoU, 训练速度有所提升, mAP@0.5 提高 1.5%。

第六组实验在引入 NWD 度量的基础上添加 MPDIoU, 能够提供更精确的目标边界框定位, 更准确地评估检测结果的位置偏移情况, 有助于定位精度的提升。mAP@0.5 提高 3.7%, 训练速度也得到显著提升。

第七组实验为同时引入 KW 动态卷积和 Detect_MBConv 检测头, 准确率提升至 84.9%, mAP@0.5 提升 2.1%, mAP@0.5:0.95 提升 1.9%。

第八组实验为同时引入 KW 动态卷积、NWD 度量和 MPDIoU, 检测速度提升明显, 并能更准确衡量微小物体边界框之间的相似性, 提高微小物体检测性能。mAP@0.5 提升了 5.5%, Precision 和 Recall 也均

表 1 所提算法在 DOTA1.0 数据集的消融实验

Table 1 Ablation experiments of the proposed algorithm in the DOTA1.0 dataset

Number	A	B	C	D	Precision/%	Recall/%	FPS	mAP@0.5 /%	mAP@0.5:0.95 /%
1	×	×	×	×	79.0	68.7	434	74.3	50.9
2	√	×	×	×	82.1	70.2	384	76.5	52.5
3	×	√	×	×	84.4	69.0	370	75.2	51.6
4	×	×	√	×	80.1	69.7	476	75.5	51.4
5	×	×	×	√	80.5	70.4	454	75.8	51.8
6	×	×	√	√	83.3	72.3	566	78.0	53.9
7	√	√	×	×	84.9	69.9	285	76.4	52.8
8	√	×	√	√	84.8	73.8	402	79.8	55.6
9	×	√	√	√	82.3	71.7	416	77.0	53.0
10	√	√	√	√	84.3	75.6	454	81.7	58.0

有明显提高。

第九组实验为同时引入 Detect_MBConv 检测头、NWD 度量和 MPDIoU, mAP@0.5 提升 2.7%, mAP@0.5:0.95 提升 2.1%。

第十组实验为本文所提算法, 通过引入 Kernel-W 动态卷积、Detect_MBConv 检测头、替换 NWD 度量和 MPDIoU, 将算法整体精度 mAP@0.5 提升至 81.7%, mAP@0.5:0.95 为 58.0%, 由数据分析可知, 本文所提算法能更好地改善漏检情况, 提升小目标检测精度, 加快训练速度, 验证了模型的有效性。

3.4 对比试验

为了评估本文算法, 分别在 DOTA1.0 数据集上

对原算法和本文算法进行训练, 训练结果分别如图 6 和图 7 所示。可以看出, 本文算法 mAP@0.5 值从 74.3% 提高到 81.7%, 整体精度提升 7.4%。其中, 小型车辆 (SV) AP 值从 73.5% 提升到 84.4%, 提升 10.9%。游泳池 (SP) AP 值从 64.4% 提升到 85.0%, 提升 20.6%。篮球场 (BC) AP 值从 66.0% 提升到 82.5%, 提升 16.5%。直升机 (HC) AP 值从 56.9% 提升到 87.2%, 提升 30.8%。同时, 飞机 (PL)、轮船 (SH)、桥梁 (BR)、储罐 (ST) 等也均有不同程度提升, 说明了本文算法的有效性。

此外, 本文算法分别和 YOLOv5、YOLOv7、CornerNet^[32]、R-FCN、YOLO-BiFPN^[17] (you only look once version with bidirectional feature pyramid network)、

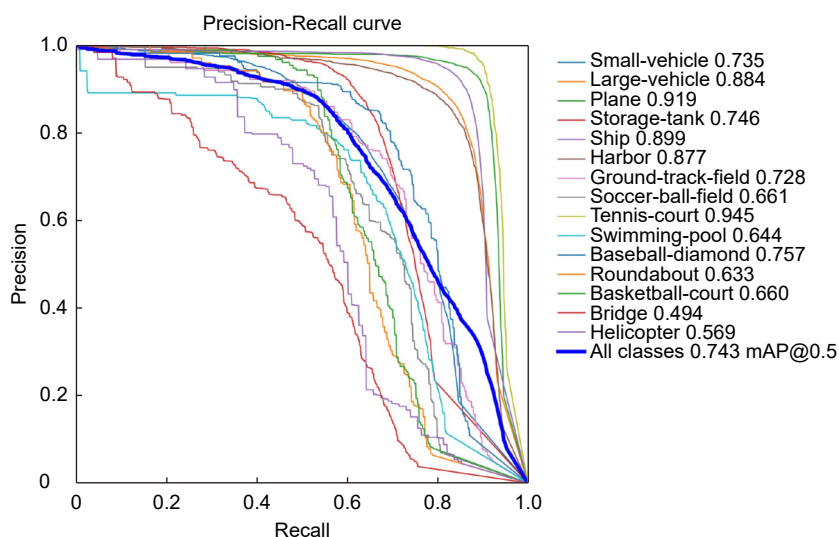


图 6 YOLOv8 算法训练结果图

Fig. 6 Training results of the YOLOv8 algorithm

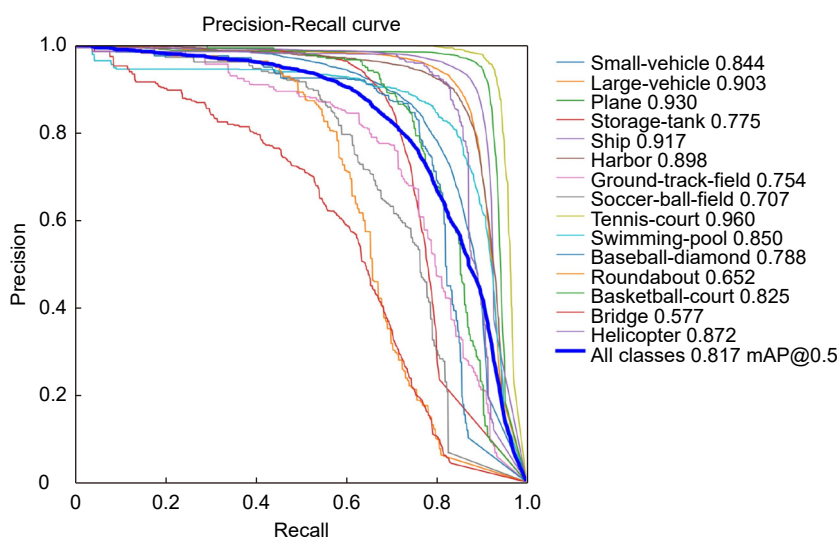


图 7 所提算法训练结果图

Fig. 7 Training results of the proposed algorithm

YOLO-PWCA^[33] 和 YOLO-DCTI^[34] 算法进行对比, 实验结果如表 2 所示, 本文算法在 DOTA1.0 数据集上检测性能表现优异, 相较之下, 其他算法检测精度明显不如本文算法。YOLOv5 在检测小目标方面, 由于小目标样本尺寸较小和下采样乘数较大, 学习小目标特征相对困难。YOLOv7 引入了小目标检测层, 可以有效提升模型对小物体识别精度, 但由于小目标像素少、覆盖面积小、信息少, 会出现目标定位不准确和漏检情况。CornerNet 模型无法捕捉到具有低对比度或复杂背景的小目标, 特征提取和预测过程导致一些小目标被误判为背景或被漏检, 在小目标检测方面表现较差。R-FCN 的设计由于通过 RoI 池化操作, 主要考虑大尺度目标的检测, 较小目标在特征图上表示不明显导致检测性能下降, 因此对于小目标检测相对较弱。YOLO-BiFPN 中改进后的特征融合模块无法充分捕捉到目标的全局上下文信息, 因此检测精度不高。YOLO-PWCA 在处理小目标时无法捕捉到足够的细节信息, 因此检测效果相对较差。YOLO-DCTI 为最近公开发表的小目标检测精度优秀的算法, 但对密集目标存在漏检问题。从实验结果来看, 本文算法在各对比的算法中精度值最佳, 并且对密集小目标如小型车辆 (SV)、直升机 (HC) 等目标的检测提升明显, 说明本文算法能有效提高检测精度, 实现高效目标检测。

为了进一步验证本文方法的有效性及其泛化性, 在 DOTA1.5 数据集上对比各算法的检测效果, 结果如表 3 所示。可以看出, 本文方法取得了最佳的 mAP (80.4%), 相较于其他算法提升了 1.7% ~ 9.6%。实验结果表明, 本文方法的精度优于其他方法。

在 DOTA 数据集上对本文算法进行测试, 并把测试结果进行可视化展示, 如图 8 所示。通过比较可以看出: YOLOv7 算法对边缘小目标存在漏检情况, YOLOv8 算法检测精度不高, YOLO-DCTI 对遥感图像中密集目标检测效果不好。本文通过引入 KW 动态卷积构建 C2f_KW 模块, 提取丰富的特征信息, 对局部区域进行更深入的感知来增强对小目标检测能力。利用 MBconv 结构重构检测头, 重组和整合特征信息, 使得模型能够更好地捕捉小目标细节和上下文信息。将 MPDIoU 和 NWD 结合起来作为回归损失函数, 归一化度量目标之间位置差异, 降低小目标在位置回归过程中的误差。综上所述, 本文算法检测精度明显提升, 能够改善边缘密集小目标漏检情况, 能够检测 YOLOv8 模型无法检测的物体, 从而解决传统方法中无法精确定位、无法准确识别小目标问题。

4 结 论

为解决遥感图像小目标检测中存在的漏检、精度低、速度慢等问题, 本文提出一种特征协同与细粒度

表 2 不同算法在 DOTA1.0 数据集上的实验结果
Table 2 Experimental results of different algorithms on DOTA1.0 dataset; unit: %

Category	YOLOv5	YOLOv7	CornerNet	R-FCN	YOLO-BiFPN	YOLO-PWCA	YOLO-DCTI	Ours
SV	75.1	76.5	10.1	49.8	81.7	77.6	86.8	88.4
LV	86.7	86.7	50.2	45.1	85.4	85.7	90.9	90.3
PL	93.6	92.3	64.7	81.1	91.8	92.6	91.9	93.0
ST	74.3	70.9	57.9	67.4	77.9	72.7	85.7	77.5
SH	89.6	89.1	31.3	49.3	89.1	87.5	81.1	91.7
HA	87.7	83.5	80.5	45.2	88.2	84.1	88.5	89.8
GTF	71.1	55.5	24.9	58.9	69.3	64.2	73.6	75.4
SBF	62.1	58.4	22.7	41.8	66.8	64.5	70.6	70.7
TC	94.0	94.9	85.5	68.9	93.4	94.0	93.8	96.0
SP	82.8	79.5	18.5	53.3	64.1	64.2	83.0	85.0
BD	76.1	71.1	38.2	58.9	74.0	78.7	77.3	78.8
RA	64.3	47.1	44.5	51.4	59.4	62.0	63.9	65.2
BC	78.3	72.2	62.5	52.1	76.2	75.5	82.1	82.5
BR	59.2	45.3	26.2	31.6	56.4	51.8	57.1	57.7
HC	84.4	81.7	12.1	33.9	83.0	76.3	85.2	87.2
mAP@0.5	78.6	73.6	42.0	52.6	80.2	78.8	80.6	81.7

表 3 不同算法在 DOTA1.5 数据集上的实验结果
Table 3 Experimental results of different algorithms on DOTA1.5 dataset; unit: %

Category	YOLOv5	YOLOv7	CornerNet	R-FCN	YOLO-BiFPN	YOLO-PWCA	YOLO-DCTI	Ours
SV	57.8	66.5	50.3	59.7	70.5	71.2	75.2	77.3
LV	71.4	82.1	59.6	58.9	77.8	86.3	88.5	89.6
PL	80.5	88.4	76.5	77.3	84.1	90.7	80.1	90.1
ST	77.8	80.9	68.1	70.5	76.2	73.1	75.7	77.6
SH	76.7	85.3	60.7	64.8	77.1	86.4	87.3	89.4
HA	82.6	81.7	77.8	75.1	86.9	80.6	89.1	90.5
GTF	73.7	80.6	64.1	60.3	77.5	75.3	74.9	76.1
SBF	63.2	68.4	58.3	61.6	73.4	66.8	75.1	75.2
TC	85.5	83.2	80.7	79.8	87.6	89.5	83.7	91.0
SP	76.1	78.6	72.9	73.4	60.5	70.1	80.4	81.3
BD	79.3	78.2	68.6	70.6	80.1	83.7	84.8	84.9
RA	73.4	75.4	70.2	66.5	74.6	68.7	70.1	76.9
BC	78.3	81.1	73.4	74.8	80.4	82.6	80.7	83.1
BR	60.3	62.5	63.2	66.3	67.1	69.8	62.6	70.2
HC	68.8	65.6	62.7	60.1	78.9	77.4	82.4	80.3
CC	62.7	67.8	64.5	66.9	73.1	70.3	74.2	75.6
mAP@0.5	76.3	75.6	72.0	70.8	78.7	77.9	78.1	80.4

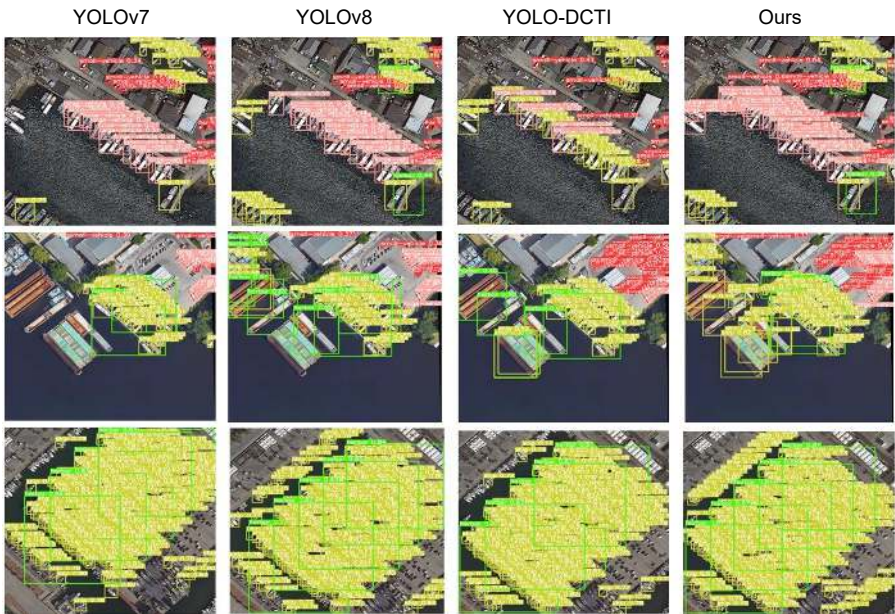


图 8 检测效果对比图
Fig. 8 Comparison of detection results

感知的遥感图像小目标检测算法，通过精细特征协同模块和细粒度感知模块的创新设计，显著提高了小目标的检测精度，减少了漏检现象。精细特征协同模块通过动态调整卷积核参数，优化特征间的交互和整合，实现对复杂背景下小目标的高效识别。细粒度感知模块进一步强化网络对极小尺寸物体的感知能力，特别

是在处理遥感图像中密集排列的小目标方面展现出了卓越的性能。此外，通过引入 MPDIoU 和 NWD 作为回归损失函数，本算法在模型训练效率和准确性上均得到了显著提升。实验结果表明，该算法在 DOTA1.0 数据集和 DOTA1.5 数据集上的表现相较于其他算法具有明显优势，证明了其在遥感图像小目标检测领域

的实际应用价值和研究意义。未来工作将集中在进一步优化算法性能, 并探索其在更广泛遥感应应用场景中的有效性。

利益冲突: 所有作者声明无利益冲突

参考文献

- [1] Ma L, Gou Y T, Lei T, et al. Small object detection based on multi-scale feature fusion using remote sensing images[J]. *Opto-Electron Eng*, 2022, **49**(4): 210363.
马梁, 苟于涛, 雷涛, 等. 基于多尺度特征融合的遥感图像小目标检测[J]. *光电工程*, 2022, **49**(4): 210363.
- [2] Chen X, Peng D L, Gu Y. Real-time object detection for UAV images based on improved YOLOv5s[J]. *Opto-Electron Eng*, 2022, **49**(3): 210372.
陈旭, 彭冬亮, 谷雨. 基于改进YOLOv5s的无人机图像实时目标检测[J]. *光电工程*, 2022, **49**(3): 210372.
- [3] Wang Y W, Guo Y, Shao X Y. Target detection in remote sensing images based on improved cascade algorithm[J]. *Acta Opt Sin*, 2022, **42**(24): 2428004.
王友伟, 郭颖, 邵香迎. 基于改进级联算法的遥感图像目标检测[J]. *光学学报*, 2022, **42**(24): 2428004.
- [4] Wang J B, Cheng G, Xie X X, et al. Multi-information supervision in optical remote sensing images[J]. *Natl Remote Sens Bull*, 2023, **27**(12): 2726–2735.
王家宝, 程臻, 谢星星, 等. 多元信息监督的遥感图像有向目标检测[J]. *遥感学报*, 2023, **27**(12): 2726–2735.
- [5] Zhang D Y, Zhao Z H, Xie Y G, et al. Research on aircraft target detection in remote sensing images based on improved YOLOv8[J]. *Autom Appl*, 2024, **65**(2): 193–195, 198.
张德银, 赵志恒, 谢逸戈, 等. 基于改进YOLOv8的遥感图像飞机目标检测研究[J]. *自动化应用*, 2024, **65**(2): 193–195, 198.
- [6] Zhang Y L, Jin H Y. Detector consistency research on remote sensing object detection[J]. *Remote Sens*, 2023, **15**(17): 4130.
- [7] Lyu Z, Jin H F, Zhen T, et al. Small object recognition algorithm of grain pests based on SSD feature fusion[J]. *IEEE Access*, 2021, **9**: 43202–43213.
- [8] Lin T Y, Goyal P, Girshick R, et al. Focal loss for dense object detection[J]. *IEEE Trans Pattern Anal Mach Intell*, 2020, **42**(2): 318–327.
- [9] Redmon J, Divvala S, Girshick R, et al. You only look once: unified, real-time object detection[C]//*Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition*, 2016: 779–788. <https://doi.org/10.1109/CVPR.2016.91>.
- [10] Redmon J, Farhadi A. YOLO9000: better, faster, stronger[C]//*Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition*, 2017: 6517–6525. <https://doi.org/10.1109/CVPR.2017.690>.
- [11] Redmon J, Farhadi A. YOLOv3: an incremental improvement[Z]. arXiv: 1804.02767, 2018. <https://doi.org/10.48550/arXiv.1804.02767>.
- [12] Bochkovskiy A, Wang C Y, Liao H Y M. YOLOv4: optimal speed and accuracy of object detection[Z]. arXiv: 2004.10934, 2020. <https://doi.org/10.48550/arXiv.2004.10934>.
- [13] Ge Z, Liu S T, Wang F, et al. YOLOX: exceeding YOLO series in 2021[Z]. arXiv: 2107.08430, 2021. <https://doi.org/10.48550/arXiv.2107.08430>.
- [14] Girshick R. Fast R-CNN[C]//*Proceedings of 2015 IEEE International Conference on Computer Vision*, 2015: 1440–1448. <https://doi.org/10.1109/ICCV.2015.169>.
- [15] Dai J F, Li Y, He K M, et al. R-FCN: object detection via region-based fully convolutional networks[C]//*Proceedings of the 30th International Conference on Neural Information Processing Systems*, 2016: 29.
- [16] Li Y T, Fan Q S, Huang H S, et al. A modified YOLOv8 detection network for UAV aerial image recognition[J]. *Drones*, 2023, **7**(5): 304.
- [17] Zhu F Z, Wang Y Y, Cui J Y, et al. Target detection for remote sensing based on the enhanced YOLOv4 with improved BiFPN[J]. *Egypt J Remote Sens Space Sci*, 2023, **26**(2): 351–360.
- [18] Zhai X X, Huang Z H, Li T, et al. YOLO-Drone: an optimized YOLOv8 network for tiny UAV object detection[J]. *Electronics*, 2023, **12**(17): 3664.
- [19] Zhou F Y, Deng H G, Xu Q G, et al. CNTR-YOLO: improved YOLOv5 based on ConvNext and transformer for aircraft detection in remote sensing images[J]. *Electronics*, 2023, **12**(12): 2671.
- [20] Zhu B Y, Lv Q B, Yang Y B, et al. Gradient structure information-guided attention generative adversarial networks for remote sensing image generation[J]. *Remote Sens*, 2023, **15**(11): 2827.
- [21] Xiao J S, Guo H W, Yao Y T, et al. Multi-scale object detection with the pixel attention mechanism in a complex background[J]. *Remote Sens*, 2022, **14**(16): 3969.
- [22] Wu J J, Su L M, Lin Z W, et al. Object detection of flexible objects with arbitrary orientation based on rotation-adaptive YOLOv5[J]. *Sensors*, 2023, **23**(10): 4925.
- [23] Yang X, Yan J C, Liao W L, et al. SCRDet++: detecting small, cluttered and rotated objects via instance-level feature denoising and rotation loss smoothing[J]. *IEEE Trans Pattern Anal Mach Intell*, 2023, **45**(2): 2384–2399.
- [24] Zhang H Y, Liu J. Direction estimation of aerial image object based on neural network[J]. *Remote Sens*, 2022, **14**(15): 3523.
- [25] Yu H T, Tian Y J, Ye Q X, et al. Spatial transform decoupling for oriented object detection[C]//*Proceedings of the 38th AAAI Conference on Artificial Intelligence*, 2024: 6782–6790. <https://doi.org/10.1609/aaai.v38i7.28502>.
- [26] Li C, Yao A B. KernelWarehouse: towards parameter-efficient dynamic convolution[Z]. arXiv: 2308.08361, 2023. <https://doi.org/10.48550/arXiv.2308.08361>.
- [27] Ma S L, Xu Y. MPDIoU: a loss for efficient and accurate bounding box regression[Z]. arXiv: 2307.07662, 2023. <https://doi.org/10.48550/arXiv.2307.07662>.
- [28] Wang J W, Xu C, Yang W, et al. A normalized Gaussian Wasserstein distance for tiny object detection[Z]. arXiv: 2110.13389, 2021. <https://doi.org/10.48550/arXiv.2110.13389>.
- [29] Sandler M, Howard A, Zhu M L, et al. MobileNetV2: inverted residuals and linear bottlenecks[C]//*Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2018: 4510–4520.
- [30] Konstantinov A V, Utkin L V, Lukashin A A, et al. Neural

attention forests: transformer-based forest improvement[C]//*Proceedings of the Seventh International Scientific Conference "Intelligent Information Technologies for Industry"*, 2023.
https://doi.org/10.1007/978-3-031-43789-2_14.

- [31] Tan M X, Le Q V. EfficientNet: rethinking model scaling for convolutional neural networks[C]//*Proceedings of the 36th International Conference on Machine Learning*, 2019: 6105–6114.
- [32] Law H, Deng J. CornerNet: detecting objects as paired keypoints[C]//*Proceedings of the 15th European Conference*

on Computer Vision (ECCV), 2018: 734–750.
https://doi.org/10.1007/978-3-030-01264-9_45.

- [33] Yu A Z, Wei W W, Wang P, et al. Small target detection algorithm for UAV based on patch-wise co-attention[J]. *Acta Aeronaut Astronaut Sin*, 2023: 1–12.
于傲泽, 魏维伟, 王平, 等. 基于分块复合注意力的无人机小目标检测算法[J]. *航空学报*, 2023: 1–12.
- [34] Min L T, Fan Z M, Lv Q Y, et al. YOLO-DCTI: small object detection in remote sensing base on contextual transformer enhancement[J]. *Remote Sens*, 2023, 15(16): 3970.

作者简介



肖振久 (1968-), 男, 副教授, 硕士研究生导师, 主要从事机器学习和图像与视觉信息计算方面的研究。

E-mail: xiaozhenjiu@lntu.edu.cn



【通信作者】张杰浩 (2000-), 女, 硕士研究生, 主要从事遥感图像目标检测方面的研究。

E-mail: zjhao0409@163.com



林渤翰 (1999-), 男, 硕士研究生, 主要从事目标检测方面的研究。

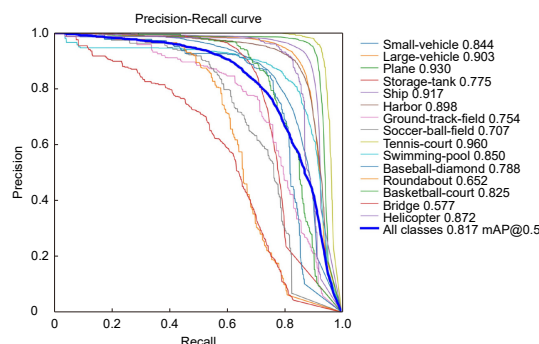
E-mail: 841608566@qq.com



扫描二维码, 获取PDF全文

Feature coordination and fine-grained perception of small targets in remote sensing images

Xiao Zhenjiu, Zhang Jiehao*, Lin Bohan



Overview: With the rapid development of remote sensing image technology, remote sensing image target detection is widely used in many important fields, including military target location and identification, natural environment protection, disaster detection, and urban planning and construction. The task of remote sensing image target detection is to accurately identify and locate the specific target in the image, and speculate its type and position. Different from targets in natural scenes, targets in remote sensing images have the characteristics of large scenes, small targets, multi-scale, complex backgrounds, overlapping occlusion, etc., so it is a challenging task to detect specific objects accurately. At present, great breakthroughs have been made in remote sensing image target detection algorithms, but the effect of small target detection is still not ideal. Small target detection faces two major difficulties: Little feature information of the target, scarce positive samples, and unbalanced classification; The target location is difficult, the background is complex, and contains a lot of redundant information, which causes serious interference to the detection. This makes it challenging to extract the edge features from aerial images and distinguish the object from the background. Therefore, the research on object detection and application in remote sensing images has important theoretical and practical significance. Addressing the challenge of missed detection caused by many small targets and dense arrangement in remote sensing images, this study introduces a small target detection algorithm for remote sensing applications, leveraging a combination of feature synergy and micro-perception strategies. Initially, we propose a refined feature synergistic fusion strategy that optimizes the interaction and integration of features across different scales by intelligently adjusting the parameters of convolution kernels. This strategy facilitates progressive refinement of features from coarse to fine granularity. Building upon this foundation, a micro-perception unit is developed in this paper, incorporating perceptual attention mechanisms with moving inverse convolution to form an advanced detection head. This innovative approach substantially boosts the network's capability to detect very small objects. Furthermore, to augment the training efficiency of the model, we employ MPDIoU and NWD as regression loss functions, mitigating positional bias issues and expediting model convergence. Experimental evaluations on the DOTA1.0 dataset and DOTA1.5 dataset reveal that our algorithm substantially improves mean Average Precision (mAP) by 7.4% and 6.1% over the baseline method, which has obvious advantages over other algorithms. The results underscore the algorithm's efficacy in significantly reducing the incidence of missed detections of small targets within remote sensing imagery.

Xiao Z J, Zhang J H, Lin B H. Feature coordination and fine-grained perception of small targets in remote sensing images[J]. *Opto-Electron Eng*, 2024, 51(6): 240066; DOI: [10.12086/oee.2024.240066](https://doi.org/10.12086/oee.2024.240066)

Foundation item: Project supported by Basic Scientific Research Project of Liaoning Provincial Universities (LJKMZ20220699), and Subject Innovation Team Project of Liaoning Technical University (LNTU20TD-23)

School of Software, Liaoning University of Engineering and Technology, Huludao, Liaoning 125105, China

* E-mail: zjhao0409@163.com