



融合分块注意力与小波特征聚合的遥感图像目标检测算法

牛为华, 郭迅

引用本文:

牛为华, 郭迅. 融合分块注意力与小波特征聚合的遥感图像目标检测算法[J]. *智能系统学报*, 2026, 21(3): 763-775.
NIU Weihua, GUO Xun. Remote sensing object detection algorithm integrating block attention and wavelet feature aggregation[J]. *CAAI Transactions on Intelligent Systems*, 2026, 21(3): 763-775.

在线阅读 View online: <https://dx.doi.org/10.11992/tis.202507006>

您可能感兴趣的其他文章

双向特征融合与注意力机制结合的目标检测

Target detection based on bidirectional feature fusion and an attention mechanism
智能系统学报. 2021, 16(6): 1098-1105 <https://dx.doi.org/10.11992/tis.202012029>

基于改进FCOS的拥挤行人检测算法

Crowded pedestrian detection algorithm based on improved FCOS
智能系统学报. 2021, 16(4): 811-818 <https://dx.doi.org/10.11992/tis.202010012>

融合视觉显著性再检测的孪生网络无人机目标跟踪算法

Siamese network combined with visual saliency re-detection for UAV object tracking
智能系统学报. 2021, 16(3): 584-594 <https://dx.doi.org/10.11992/tis.202101035>

结合模糊特征检测的鲁棒核相关滤波跟踪法

Robust KCF tracking algorithm combined with fuzzy feature detection
智能系统学报. 2021, 16(2): 323-329 <https://dx.doi.org/10.11992/tis.201912010>

基于特征融合及自适应模型更新的相关滤波目标跟踪算法

Correlation filter target tracking algorithm based on feature fusion and adaptive model updating
智能系统学报. 2020, 15(4): 714-721 <https://dx.doi.org/10.11992/tis.201803036>

基于反卷积和特征融合的SSD小目标检测算法

SSD small target detection algorithm based on deconvolution and feature fusion
智能系统学报. 2020, 15(2): 310-316 <https://dx.doi.org/10.11992/tis.201905035>

DOI: 10.11992/tis.202507006

网络出版地址: <https://link.cnki.net/urlid/23.1538.tp.20260306.1518.004>

融合分块注意力与小波特征聚合的 遥感图像目标检测算法

牛为华^{1,2}, 郭迅¹

(1. 华北电力大学 计算机系, 河北 保定 071003; 2. 复杂能源系统与智能计算教育部工程研究中心, 河北 保定 071003)

摘要: 针对遥感图像中目标背景复杂和小目标物体检测困难等问题, 提出融合分块注意力与小波特征聚合的遥感图像目标检测算法。在主干网络中引入层级分块注意力模块, 采用分块下采样操作进行下采样操作, 同时融合坐标注意力机制, 缓解了细粒度信息丢失的问题; 设计融合小波增强的对比驱动特征聚合模块, 采用 Haar 小波变换分离图像高低频信息, 并构建局部注意力机制, 增强模型对目标边缘与纹理细节的感知能力, 抑制复杂背景的干扰; 构建浅层注意力增强检测头, 融合 LSKA (large selective kernel attention) 注意力机制增强模型对小目标物体的检测精度。实验结果表明, 改进后算法在 ShipRSImageNet 与 VisDrone2019 数据集上的 mAP50 分别达到 87.3% 与 35.5%, 相比原始模型分别提升 5.9% 和 4.1%。改进后的模型在遥感图像目标检测任务上的性能得到了有效的提升。

关键词: 目标检测; 遥感影像; 小波变换; YOLO; 注意力机制; 小目标检测; 特征提取; 特征融合

中图分类号: TP391.41 **文献标志码:** A **文章编号:** 1673-4785(2026)03-0763-13

中文引用格式: 牛为华, 郭迅. 融合分块注意力与小波特征聚合的遥感图像目标检测算法 [J]. 智能系统学报, 2026, 21(3): 763-775.

英文引用格式: NIU Weihua, GUO Xun. Remote sensing object detection algorithm integrating block attention and wavelet feature aggregation[J]. CAAI transactions on intelligent systems, 2026, 21(3): 763-775.

Remote sensing object detection algorithm integrating block attention and wavelet feature aggregation

NIU Weihua^{1,2}, GUO Xun¹

(1. Department of Computer Science, North China Electric Power University, Baoding 071003, China; 2. Engineering Research Center of Intelligent Computing for Complex Energy System, Ministry of Education, Baoding 071003, China)

Abstract: A remote sensing image object detection algorithm is proposed to address the challenges of complex backgrounds and small object detection. A hierarchical split attention block attention module is introduced into the backbone network, utilizing block-based downsampling and coordinate attention to mitigate fine-grained information loss. Using the Haar wavelet transform, a wavelet-enhanced contrast-driven feature aggregation module separates high- and low-frequency information, while a local attention mechanism enhances edge and texture perception, suppressing background interference. A shallow enhancement attention detection head with the LSKA attention mechanism improves small object detection accuracy. Experimental results show that the algorithm achieves an mAP50 of 87.3% on the ShipRSImageNet dataset and 35.5% on the VisDrone2019 dataset, with improvements of 5.9% and 4.1%, respectively, over the original model. The improved model considerably enhances performance in remote sensing image object detection.

Keywords: object detection; remote sensing imagery; wavelet transform; YOLO; attention mechanism; small object detection; feature extraction; feature fusion

收稿日期: 2025-07-04. 网络出版日期: 2026-03-09.

基金项目: 国家自然科学基金项目 (62371188).

通信作者: 牛为华. E-mail: niuw@ncepu.edu.cn.

遥感图像是通过搭载在卫星和无人机上的遥感传感器所获取的对地表的观测数据, 其具备监

测范围广、成像时效性强以及空间和光谱信息丰富等优势。遥感图像目前被广泛应用于军事侦察、海洋监控和灾害预警等领域。因此,近年来遥感图像目标检测任务成为计算机视觉领域的研究热点之一。

与常规目标检测任务所检测的图像相比,遥感图像通常具有背景复杂、目标尺度差异大、目标数量稀疏且分布不均、图像分辨率较低和目标边界相对较模糊等一系列特征。这些特征使得遥感图像目标检测具有更高的挑战性和复杂性。

目前,针对遥感图像目标检测任务所存在的难题,主要有两类解决方法:第一类为基于手工特征提取的传统方法,第二类为基于深度学习的解决方法。传统的基于手工特征的方法主要有方向梯度直方图算法 (histograms of oriented gradients, HOG) 和可变形部件模型算法 (deformable part model, DPM)。HOG 系列算法^[1-2]首先在输入图像中创建一个网格,根据网格计算得到特征表,生成梯度直方图。通过提取感兴趣的区域生成特征,最后将特征输入到支持向量机 (support vector machine, SVM) 分类器中进行检测。DPM 系列算法^[3-4]是 HOG 算法的升级和扩展,其通过引入部件结构和可变形机制,有效地解决了目标在多种视角下的检测问题,但其在处理复杂背景和尺度变化较大的场景时,检测效果并不理想^[5]。在此背景下,姬晓飞等^[6]提出了一种基于多特征决策级融合的光学遥感图像多目标识别算法。该算法通过提取分层的 BoF-SIFT (bag of feature-scale invariant feature Transform) 特征、改进的 SC (shape context) 特征和 Hu 不变矩特征,并采用一对一支持向量机算法进行分类,提高了识别准确性。近年来,随着深度学习技术的快速发展,基于深度学习的目标检测算法在遥感图像目标检测任务上效果显著。这类算法通过采用端到端的训练框架,能够从大规模的数据中学习到更加鲁棒的目标特征,显著提升模型在复杂背景等场景下的检测性能。基于深度学习的目标检测算法目前可以分为双阶段检测算法和单阶段检测算法两大类。

双阶段检测算法的代表算法有 R-CNN (region-base convolutional neural networks)、Faster R-CNN 和 Cascade R-CNN 等,但上述算法最初并非针对遥感图像目标检测任务而设计。为此,一些研究人员通过改进现有的一些算法,使其能更好地应用于遥感图像目标检测任务。如:Li 等^[7]提出一种基于全局和局部注意力机制的特征融合模块,有效地平衡了局部细节和全局上下文的利用,提升了模型的检测精度;魏谢根等^[8]通过采用

多层次局部自注意力机制增强模型对浅层特征提取能力,提升了检测精度;Su 等^[9]设计了一种上下文感知模块,通过结合目标周围背景的上下文信息来减少分类混淆,有效地解决了模型在定位上表现良好但在分类上表现较差的问题;姚婷婷等^[10]通过构建感受野扩张模块,有效提升了模型对不同尺度目标的检测精度。除双阶段检测算法之外,单阶段检测算法通常具备更快的检测速度,所以常用于一些实时性要求较高的检测场景中。这类检测算法的代表算法有 YOLO (you only look once)、SSD (single shot multibox detector) 和 RetinaNet 等。针对单阶段检测算法,也有许多研究人员针对遥感图像目标检测任务提出一些改进。如:文斌等^[11]提出一种空间信息整合模块 SIM (spatial information integration module) 用于构建新型路径聚合网络,有效地解决了遥感场景下特征融合过程中信息冗余和忽略内层特征的问题;周沁坤等^[12]提出一种全新的特征增强模块用于缓解图像背景复杂和目标特征较弱等问题,有效增强了模型的特征表达能力;Li 等^[13]设计了一种面向遥感目标检测的自适应点表示方法,通过自适应点学习有效捕捉任意方向目标的几何结构信息,从而提升检测性能;赵文清等^[14]提出一种结合倒残差自注意力机制的目标检测方法,通过采用倒残差自注意力机制骨干网络充分提取目标特征,降低复杂背景的干扰;许思源等^[15]提出一种基于多尺度特征融合的遥感图像目标检测方法,通过使用 TIF (tripolar integrative fusion) 融合模块与 SlideLoss 损失函数,在多个遥感数据集上显著提升检测精度。

尽管现有的遥感图像目标检测算法在一定程度上提升了检测精度,但在处理复杂背景、图像模糊和小目标检测方面仍存在一定的局限性。为此,针对现有遥感图像目标检测所存在的问题,本文在 YOLOv8 模型的基础上提出了三点改进:

- 1) 设计层级分块注意力模块,替换 YOLOv8 算法模型主干网络上的卷积模块,增强特征提取能力并缓解由于跨步卷积所导致的细节信息丢失的问题。

- 2) 构建融合小波增强的对比驱动特征聚合模块,通过采用 Haar 小波变换提取特征图的高频信息与低频信息,构建局部注意力机制,引导模型聚焦潜在的目标区域,增强信息表达。

- 3) 提出浅层注意力增强检测头,通过引入更高分辨率的 160×160 特征图,并应用大型可分离卷积核注意力模块,帮助模型更好地捕捉小目标区域特征,加强模型的建模能力,提升模型在小

目标检测任务中的召回率。

1 YOLOv8 网络模型结构

YOLOv8 算法模型是当前目标检测领域中的

主流算法之一, 其检测精度和推理速度在许多目标检测任务中都取得了优异的成绩, 为此本文采用 YOLOv8 算法模型作为基线模型进行改进。

YOLOv8 网络结构如图 1 所示。

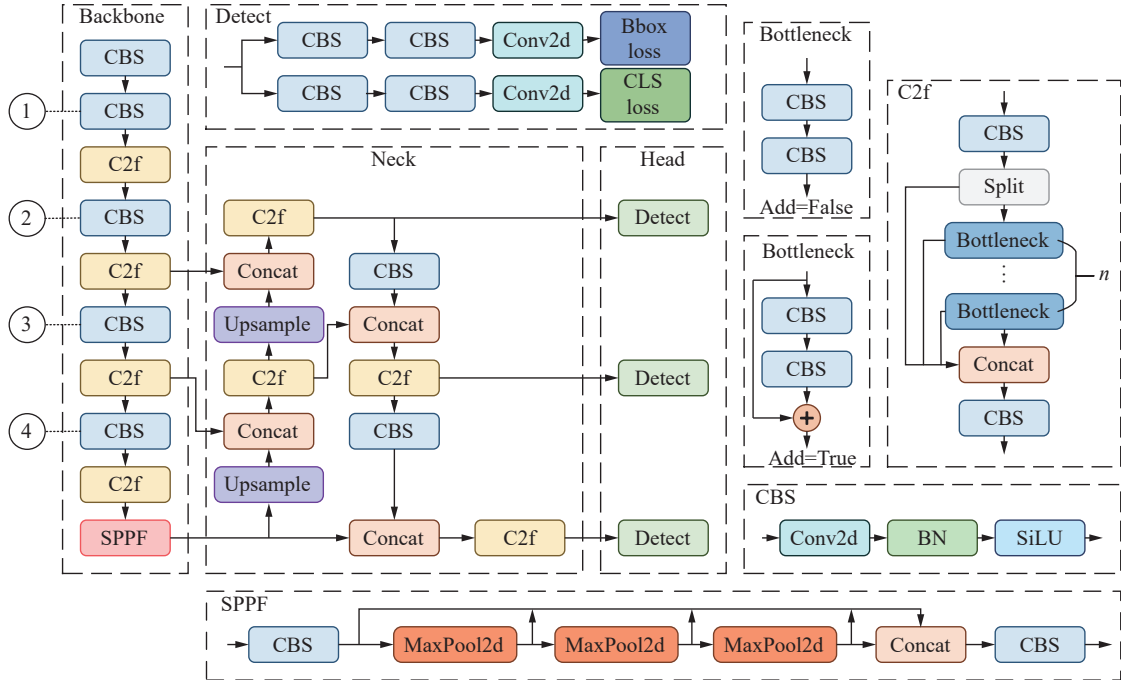


图 1 YOLOv8 网络结构

Fig. 1 YOLOv8 network architecture

YOLOv8 模型主要由 3 个部分组成, 分别为主干网络 (Backbone)、特征融合网络 (Neck) 和检测头 (Head)。其中, 主干网络通过使用 C2f 模块和卷积模块提取输入图像的语义特征信息, 在保持较低复杂度的同时, 增强了网络模型对图像局部结构与上下文信息的感知能力。特征融合网络通过采用 FPN+PAN 结构, 将特征提取阶段中提取的不同尺度的特征信息进行融合, FPN 通过自顶向下的路径传递高层语义信息, 而 PAN 通过自底向上路径强化底层细节特征, 从而实现语义信息和细节信息的充分交互, 增强模型对多尺度目标的检测能力。检测头将融合后的特征进行解码操作, 实现目标分类与边界框回归。

2 本文算法

2.1 HSAB 模块

在 YOLOv8 网络模型中, 主干网络采用了大量的标准卷积对输入图像及中间特征进行特征提取与下采样操作。然而在遥感图像目标检测中, 遥感图像通常存在目标呈现多尺度特征、背景复杂和图像模糊等特性, 这些特性进一步提升了遥感图像目标检测任务的难度。由于标准卷积对图

像的旋转、缩放等几何变换的鲁棒性较差, 所以其难以捕获图像中复杂的特征信息。此外, 使用标准卷积进行下采样操作时, 通常会将卷积的步长设置为 2, 但使用步长为 2 的卷积进行下采样操作时, 由于其是一种跨步卷积, 所以有可能会产生细粒度信息丢失的问题, 影响小目标物体和稀疏目标物体的检测精度。为解决上述问题, 本文设计了层级分块注意力模块 (hierarchical split attention block, HSAB) 用于替换主干网络上的部分标准卷积。

HSAB 模块首先通过分块下采样, 将输入特征 $X_{in} \in \mathbf{R}^{H \times W \times C}$ 分为 4 个大小相等的子特征 X_i , 通过这种方式进行下采样操作可以有效缓解使用步长为 2 的卷积所带来的细节特征信息丢失的问题。随后, 对每个子特征 X_i 进行特征提取和通道融合, 以增强模型的建模能力。然后, 对 X_i 按照通道拆分为 3 路特征, 得到特征 $X_i^{(j)}$ 。向特征 $X_i^{(1)}$ 中引入坐标注意力机制^[16](coordinate attention, CA), 将空间位置信息编码至特征图中, 引导网络聚焦关键目标区域。与此同时将特征 $X_i^{(2)}$ 与 $X_i^{(3)}$ 进行叠加, 强化重要特征, 并通过深度可分离卷积^[17](depthwise separable convolution, DSC) 进一步提取特征信息。随后将未经处理的特征 $X_i^{(2)}$ 与上述两

部分特征按照通道维数进行拼接,并通过卷积核大小为 1、步长为 1 的卷积进行跨通道融合。最

后将 4 部分信息按照通道维数拼接,并进行融合,得到该模块的输出。HSAB 模块如图 2 所示。

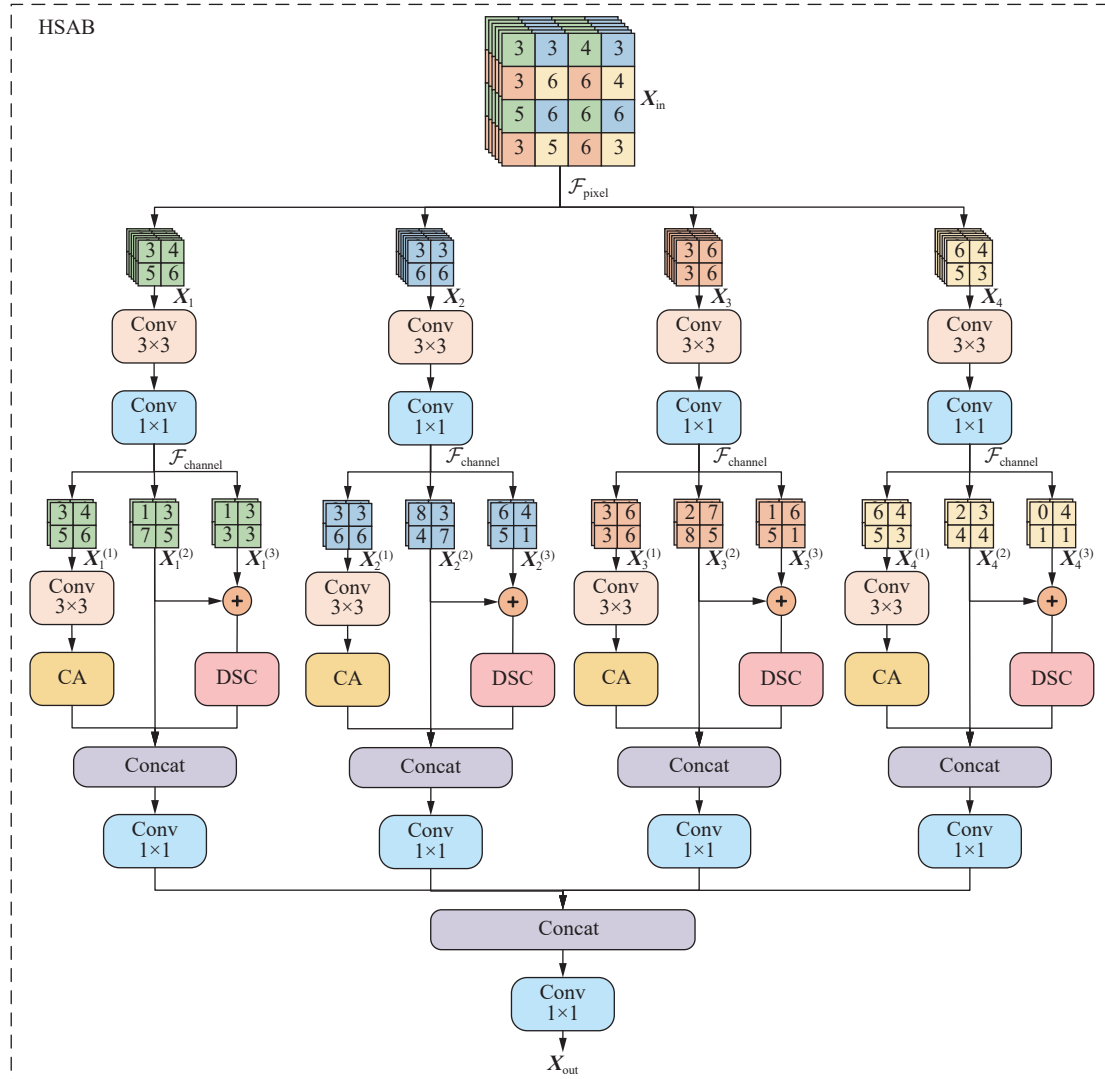


图 2 HSAB 模块

Fig. 2 HSAB module

HSAB 模块的公式表达分别为

$$\{X_i\} = \mathcal{F}_{\text{pixel}}(X_{\text{in}}), i \in \{1, 2, 3, 4\} \quad (1)$$

$$\{X_i^{(j)}\} = \mathcal{F}_{\text{channel}}(X_i), j \in \{1, 2, 3\} \quad (2)$$

$$C_i = \mathcal{F}_{\text{merge}}(\mathcal{F}_{\text{CA}}(X_i^{(1)}), X_i^{(2)}, \mathcal{F}_{\text{DSC}}(X_i^{(2)} + X_i^{(3)})) \quad (3)$$

$$X_{\text{out}} = \mathcal{F}_{\text{merge}}(C_1, C_2, C_3, C_4) \quad (4)$$

式中: $\mathcal{F}_{\text{pixel}}$ 代表分块下采样操作, X_{in} 代表从上层输入的特征, $\mathcal{F}_{\text{channel}}$ 代表通道切分操作, $\mathcal{F}_{\text{merge}}$ 、 \mathcal{F}_{CA} 和 \mathcal{F}_{DSC} 分别代表按通道维度拼接并进行 1×1 卷积融合运算、向特征图中融合坐标注意力操作和对特征图进行深度可分离卷积运算。

2.2 WE-CDFA 模块

在遥感图像目标检测任务中,遥感图像由于其成像高度较高、成像分辨率较低以及所要检测物体尺度差异大等因素,往往存在目标边界模糊、细节信息容易丢失的问题。此外,遥感图像

背景相对较复杂、干扰信息偏多,容易造成漏检误检等问题。为解决上述问题,本文通过引入 Haar 小波变换,并结合对比驱动特征聚合模块^[18] (contrast-driven feature aggregation, CDFA), 设计出融合小波增强的对比驱动特征聚合模块 (wavelet-enhanced contrast-driven feature aggregation, WE-CDFA)。

WE-CDFA 模块首先对输入特征 $F_{\text{in}} \in \mathbf{R}^{H \times W \times C}$ 使用两个卷积核为 3、步长为 1 的卷积提取特征图中的局部上下文信息,增强初始特征的代表能力。随后对输出特征 F' 通过线性层 $W_v \in \mathbf{R}^{C \times C}$, 映射为值向量 $V \in \mathbf{R}^{H \times W \times C}$, 并将值向量 V 在每个局部窗口内通过 Unflod 方法进行展开,以便为每个位置聚合邻域信息。令 $V_{\Delta i, j} \in \mathbf{R}^{C \times K^2}$ 代表以 (i, j) 为中心的局部窗口中的所有值,其定义为

$$V_{\Delta i, j} = \left\{ V_{i+p-\lfloor \frac{K}{2} \rfloor, j+q-\lfloor \frac{K}{2} \rfloor} \right\}, 0 \leq p, q < K \quad (5)$$

式中: p, q 代表局部窗口偏移量, K 代表窗口大小。

与此同时, 通过采用 Haar 小波变换, 提取特征图 F' 的水平高频分量 f_h 、垂直高频分量 f_v 、对角高频分量 f_d 以及低频分量 f_{low} , 并将所提取的 3 个高频分量进行叠加, 获得特征高频信息 f_{high} 。其中, 高频分量主要包含目标边缘和纹理细节等特征信息, 有利于加强模型对目标轮廓的感知能力。低频分量则保留了主要的背景信息, 有助于模型理解全局上下文和背景语义。随后, 对高频分量 f_{high} 和低频分量 f_{low} 分别通过两个不同的线性变换生成对应的注意力权重 A_{high} 与 A_{low} 。注意力权重的计算公式为

$$A_{high} = W_{high} \cdot f_{high}, \quad A_{low} = W_{low} \cdot f_{low} \quad (6)$$

式中 $W_{high} \in \mathbf{R}^{C \times K^4}$ 和 $W_{low} \in \mathbf{R}^{C \times K^4}$ 为线性变换权重矩阵。

接着, 位置 (i, j) 处的高频和低频注意力权重被重塑为 $A_{i,j}^{high} \in \mathbf{R}^{K^2 \times K^2}$ 与 $A_{i,j}^{low} \in \mathbf{R}^{K^2 \times K^2}$ 。并通过 Softmax 函数进行归一化。最后, 对展开后的值向量进行两步加权。WE-CDFA 模块的公式表达为

$$F_{out} = \mathcal{F}_{cbr}(\text{Softmax}(A_{i,j}^{low}) \otimes (\text{Softmax}(A_{i,j}^{high}) \otimes V_{\Delta i,j})) \quad (7)$$

其中 \mathcal{F}_{cbr} 代表两个标准卷积运算。

该模块将通过特征高频信息作为引导, 构建局部注意力机制, 引导模型聚焦潜在的目标区域, 增强信息表达。通过低频分量, 学习背景分布和上下文语义, 抑制冗余干扰信息。WE-CDFA 模块通过融合前景增强与背景建模后的特征, 实现了对遥感图像中细节目标的精准感知与定位, 可以有效降低复杂背景下的误检率与漏检率。WE-CDFA 模块的结构如图 3 所示。

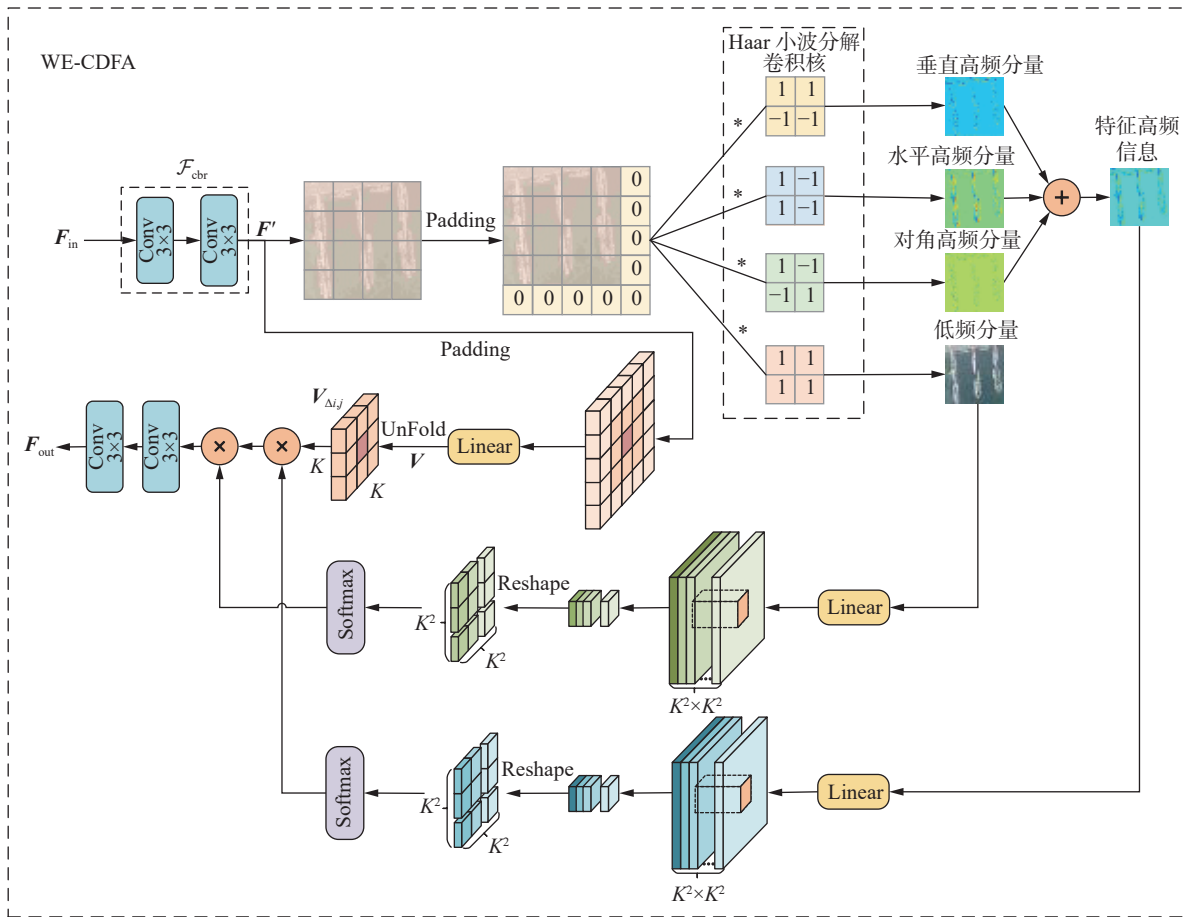


图 3 WE-CDFA 模块

Fig. 3 WE-CDFA module

2.3 SEAD 检测头

在 YOLOv8 算法中, 检测层所接收的特征图大小分别为 80×80 、 40×40 和 20×20 三种不同尺度的特征图, 其中 80×80 的特征图主要用于小目标物体的检测。为了进一步增强模型对小目标物体的检测精度, 本文提出一种浅层注意力增强检测头 (shallow enhancement attention detect, SEAD), 来进一步提升模型对小目标物体的检测精度。

SEAD 通过从 P1 层引出新的分支, 并通过大型可分离卷积核注意力模块^[19] (large selective kernel attention, LSKA) 和 C2f_Efficient 模块输出一个 160×160 的特征图, 用于小目标检测。C2f_Efficient 模块如图 4 所示。

C2f_Efficient 是通过将 C2f 模块中的 Bottleneck 模块替换为 FasterNet^[20] 中的 FasterNetBlock 得到的, 通过引入更加轻量化的基础单元, 来提升网络的推理速度。

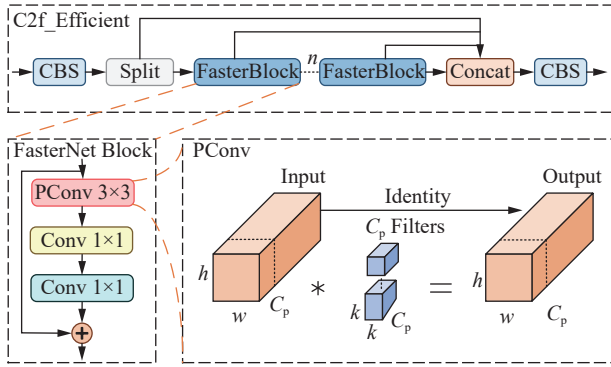


图 4 C2f_Efficient 模块
Fig. 4 C2f_Efficient module

LSKA 是在大核卷积注意力模块^[21] (large kernel attention, LKA) 基础上改进得到的。LKA 首先通过深度卷积和小感受野卷积提取特征图的局部特征信息, 随后通过使用扩张卷积和大感受野卷积核来捕捉远程依赖关系, 再采用卷积核大小为 1 的卷积生成空间注意力图, 并将生成的注意力图与原始输入进行相乘, 以此实现特征的自适应增强与细化。相比于 LKA, LSKA 通过将原本的 $K \times K$ 卷积核拆分为 $1 \times K$ 和 $K \times 1$ 两个卷积核串联的形式, 在保留大感受野建模能力的同时, 显著减少了卷积操作中的参数量与计算开销。LKA 与 LSKA 的结构如图 5 所示。

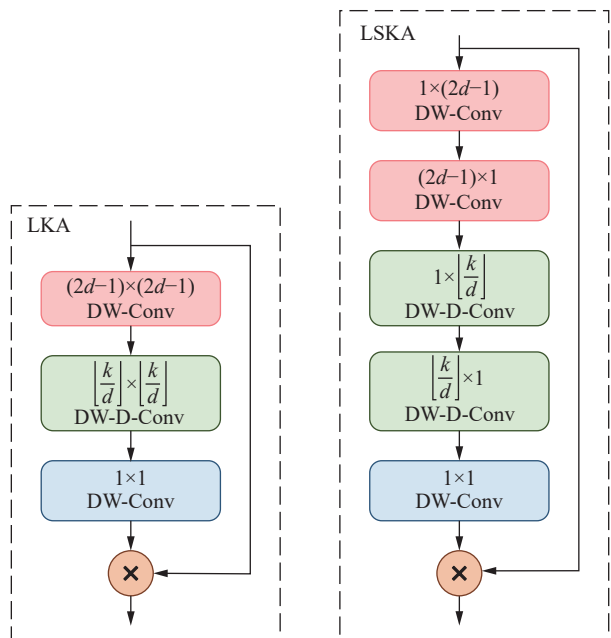


图 5 LKA 与 LSKA 模块
Fig. 5 LKA and LSKA module

LKSA 的输出表达式为

$$\bar{Z}^C = \sum_{h,w} W_{(2d-1) \times 1}^C * (W_{1 \times (2d-1)}^C * F^C) \quad (8)$$

$$Z^C = \sum_{h,w} W_{[k/d] \times 1}^C * \left(W_{1 \times [k/d]}^C * \bar{Z}^C \right) \quad (9)$$

式中: h 、 w 分别代表特征图的高度和宽度, W 代

表卷积核, 下标代表卷积核尺寸, d 代表空洞率。

改进后的模型结构图如图 6 所示。

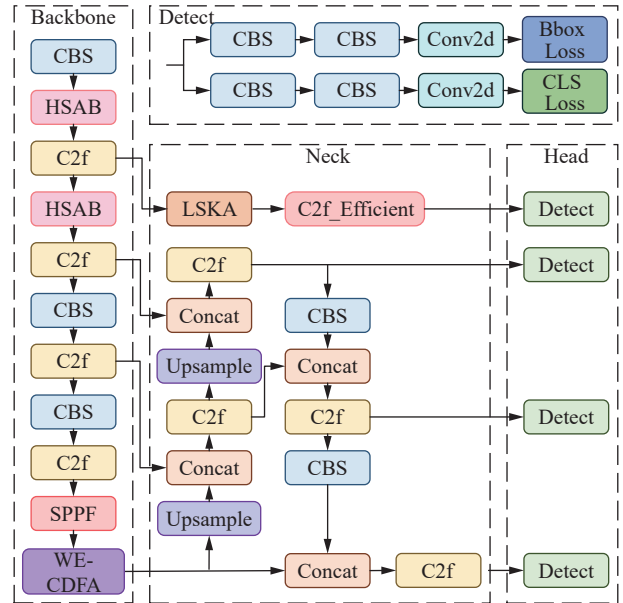


图 6 改进后的模型结构
Fig. 6 Improved model architecture

3 实验结果与分析

3.1 实验环境

为了防止由于硬件差异所导致的实验误差, 本文实验均在相同的实验环境, 相同的硬件配置下进行, 以确保实验结果的公平性与可比性, 最大程度降低因硬件差异给实验结果带来的干扰。具体配置如表 1 所示。

表 1 实验软硬件配置详细信息
Table 1 Detailed hardware configuration

名称	详细配置
操作系统	Ubuntu 20.04.6 LTS (GNU/Linux 5.15.0-136-generic x86_64)
处理器	12th Gen Intel(R) Core(TM) i5-12400F
显卡	NVIDIA GeForce RTX 3060 12 GB
内存	16 GB DDR4 3200 MHz × 2
CUDA	12.2
PyTorch	2.1.1

3.2 实验数据集与超参数设置

ShipRSImageNet 数据集共包含图像 3 435 张, 数据集来自多种传感器和卫星平台, 拍摄时间涵盖不同的方向与季节, 每张图片的分辨率为 930 像素 × 930 像素, 标注方式包括水平框与旋转框两种方式。在本文实验中采用 2 198 张图像作为训练集, 550 张图像作为测试集, 共计 2 748 张图像。目标分类等级采用 L0, 即仅对舰船和码头两个目标进行区分, 不再对舰船进行更细致的分类。

VisDrone2019 数据集共包含图像 8 629 张, 数据集收集自中国 14 个不同的城市, 包含 10 类目标物体, 涵盖多种地理环境和复杂场景, 是当前遥感小目标检测领域中比较有代表性和挑战性的数据集之一, 其标注方式为水平框标注。在本文实验中, 将采用 6 471 张图像作为训练集, 548 张图像作为验证集, 1 610 张图像作为测试集, 共计 8 629 张图像。

在对 ShipRSImageNet 数据集进行网络训练的过程中, 本文所采用的网络训练策略如下: 批量大小设置为 8, 采用随机梯度下降优化器 (stochastic gradient descent, SGD) 对网络参数进行迭代更新, 初始学习率设定为 0.01, 动量设置为 0.9, 训练轮次为 600 轮, CPU 线程数设置为 4, 开启图像增强算法, 输入图像尺寸固定为 640×640。为了更加准确地评估模型本身的性能, 在实验过程中不采用预训练模型权重对模型参数进行初始化。在对 VisDrone2019 数据集进行网络训练的过程中, 将上述训练方式的批量大小更改为 2, 训练轮次更改为 300, 其余超参数和训练策略不变。

3.3 评价指标

本实验采用平均精度均值 (mean average precision, mAP) 作为评价指标, mAP 值越高, 则表示模型在各类目标检测中的准确性和稳定性越强。所使用的 IoU 阈值均为 0.5。mAP 的计算公式为

$$m_{AP} = \frac{1}{c} \sum_{i=1}^n I_{AP}^i \quad (10)$$

式中: c 表示目标类别数, m_{AP} 是对每个类别的准确率 P (Precision) 和召回率 R (Recall) 绘制的曲线与横纵坐标轴所围成的面积的平均值。准确率和召回率计算公式分别为

$$P = \frac{P_T}{P_T + P_F} \times 100\% \quad (11)$$

$$R = \frac{P_T}{P_T + N_F} \times 100\% \quad (12)$$

式中: P_T 为正确检测的正样本的数量, P_F 为错误检测的负样本的数量, N_F 为错误检测的正样本的数量。

3.4 实验结果

3.4.1 HSAB 模块分析

为验证 HSAB 模块在数量与位置上对 YOLOv8s 模型性能的影响。本文通过在 YOLOv8s 网络模型的不同位置上使用 HSAB 模块, 并采用不同数量的模块来进行实验验证。替换位置编号如图 1 所示。本组实验共设计 8 组: 其中, 第 1 组为对照组; 第 2~5 组为分别替换主干网络上 1~4 号位置为 HSAB 模块; 第 6~8 组为分别替换主干网络上 HSAB 模块数量为 2、3 和 4 个。在本模块分析实验中, 批次大小设置为 4, 训练轮数为 300, 初始学习率为 0.01, 实验结果如表 2 所示。

表 2 HSAB 模块分析实验表
Table 2 HSAB module analysis experiment table

编号	替换位置	Ship50/%	Dock50/%	Precision/%	Recall/%	mAP50/%	mAP50-95/%	Params/ 10^6	FLOPs/ 10^9
G1	-	79.8	81.5	85.4	74.7	80.7	59.9	11.4	29.4
G2	1	83.2	82.1	83.6	76.1	82.7	61.2	11.5	45.1
G3	2	82.2	81.5	86.6	72.0	81.9	61.7	12.0	44.2
G4	3	79.3	81.0	83.0	73.6	80.2	59.8	14.0	43.8
G5	4	79.2	80.3	85.9	70.5	79.8	59.8	22.0	43.8
G6	1,2	81.2	86.2	89.2	73.3	83.7	62.0	12.2	59.8
G7	1,2,3	84.0	82.6	86.6	72.5	83.3	60.8	14.9	74.2
G8	1,2,3,4	81.6	83.7	81.4	76.6	82.6	61.5	25.6	88.6

由表 2 可知, 在不同位置引入不同数量的 HSAB 模块对 YOLOv8s 模型的检测性能产生了不同的影响。由第 2、3 组实验可以看出, 在 1 号或 2 号位置上添加 HSAB 模块后, 模型的检测性能相比未添加之前, mAP50 分别提升 2.0% 与 1.2%。而在 3 号或 4 号位置上引入 HSAB 模块后, 模型的部分指标有所下降。这一现象表明, 过深的位置引入 HSAB 模块可能会导致浅层的特征信息传递受限, 因而影响检测精度。

在多模块组合实验中, 由第 6 组实验可以看出, 当替换主干网络上 1 号和 2 号位置为 HSAB 模块后, 模型的检测性能得到了显著的提升, mAP50 相比于原始模型提升 3.0%。然而, 当进一步增加 3 个或 4 个 HSAB 模块时, 尽管部分指标有了小幅度提升, 但是模型整体性能相比于仅替换 1、2 号位置略有下降, 且过多数量的替换会导致模型的参数量与浮点运算量显著提升。为此, 综合模型精度与性能的考虑, 本文选取在 1 号和

2 号位置替换 HSAB 模块, 以达到检测精度与性能的均衡。

3.4.2 本文算法检测结果

为了相对全面地验证本文所提出的改进算法

在实际应用场景中的检测能力, 本文选取了两个在遥感图像目标检测中具有代表性的数据集来验证本文所提出模型的检测能力。具体实验结果如表 3 和表 4 所示。

表 3 模型在 ShipRSImageNet 上的实验结果
Table 3 Experimental results of the model on ShipRSImageNet

模型	Ship50/%	Dock50/%	Precision/%	Recall/%	mAP50/%	mAP50-95/%	Params/10 ⁶	FLOPs/10 ⁹
YOLOv8s	81.4	81.5	83.0	77.0	81.4	62.2	11.4	29.4
本文模型	87.6	86.9	88.6	78.9	87.3	66.5	23.2	85.4

表 4 改进前后模型在 VisDrone2019 上的实验结果
Table 4 Experimental results of the model before and after improvement on VisDrone2019

类别	YOLOv8s				本文算法			
	Precision/%	Recall/%	mAP50-%	mAP50-95-%	Precision/%	Recall/%	mAP50-%	mAP50-95-%
pedestrian	51.4	23.8	26.2	10.6	48.9	29.0	30.8 (+4.6)	12.9 (+2.3)
people	51.2	9.5	13.4	4.6	53.8	12.5	17.8 (+4.4)	6.6 (+2.0)
bicycle	28.3	10.7	8.4	3.1	36.3	11.7	12.6 (+4.2)	5.0 (+1.9)
car	65.4	71.7	71.0	45.5	67.4	76.4	75.7 (+4.7)	49.2 (+3.7)
van	39.4	41.2	36.6	24.3	45.3	42.2	40.4 (+3.8)	27.5 (+3.2)
truck	41.8	43.8	39.0	24.3	46.3	46.4	43.3 (+4.3)	28.1 (+3.8)
tricycle	23.7	26.2	16.6	8.3	30.4	28.7	20.4 (+3.8)	11.3 (+3.0)
awning-tricycle	40.1	20.7	18.5	10.4	47.5	22.2	23.3 (+4.8)	13.7 (+3.3)
bus	61.6	54.1	56.2	38.9	67.4	56.3	59.6 (+3.4)	42.2 (+3.3)
motor	40.4	32.5	28.0	11.2	46.3	32.7	30.8 (+1.2)	12.7 (+1.5)
所有	44.3	33.4	31.4	18.1	49.0	35.8	35.5 (+4.1)	20.9 (+2.8)

由表 3 可以看出, 本文改进算法相较于原始算法在 ShipRSImageNet 数据集上取得了较大幅度的提升, 其中 mAP50 提升 5.9%, mAP50-95 提升 4.3%, 对于舰船目标和码头目标 mAP50 精度分别提升 6.2% 与 5.4%。由此可以表明本文所提出算法针对现有模型细粒度信息易丢失, 以及针对复杂背景建模能力较差等问题有了很大的改善。

由表 4 可以看出, 本文所提出算法在 VisDrone-2019 数据集上提升比较明显, 其中 mAP50 提升 4.1%, mAP50-95 提升 2.8%。对所有类别目标检测精度均有提升, 由此可以看出本文所改进模型

具有很好的鲁棒性, 在多个数据集上表现优异。

3.4.3 消融实验

为了验证本文所提出改进的有效性, 本文在 YOLOv8s 模型的基础上设计了系统性的消融实验, 且实验环境与实验超参数完全一致, 确保实验结果公平合理。本消融实验共分为 8 组: 第 1 组为对照组, 即采用未改进的 YOLOv8s 网络模型进行实验; 第 2~4 组为分别添加一种改进后的模型; 第 5~7 组为将本文所提出改进两两进行结合之后的模型; 第 8 组为融合本文所提出所有改进后的模型。实验结果如表 5 所示。

表 5 模型在 ShipRSImageNet 上的消融实验结果
Table 5 Ablation experiment results of the model on ShipRSImageNet

HSAB	WE-CDFA	SEAD	Ship50/%	Ship50-95/%	Dock50/%	Dock50-95/%	Precision/%	Recall/%	mAP50/%	mAP50-95/%	Params/10 ⁶	FLOPs/10 ⁹
			81.4	63.0	81.5	61.4	83.0	77.0	81.4	62.2	11.4	29.4
√			84.7	65.2	84.0	62.4	86.1	78.6	84.3	63.8	12.2	59.8
	√		83.9	64.4	85.2	61.5	83.5	77.4	84.6	63.0	22.0	38.1
		√	84.0	64.7	83.6	62.0	82.5	78.0	83.8	63.4	11.7	48.6
√	√		84.6	65.0	84.5	65.1	81.1	79.8	84.5	65.1	22.9	68.6

续表 5

HSAB	WE-CDFA	SEAD	Ship50/ %	Ship50-95/ %	Dock50/ %	Dock50-95/ %	Precision/ %	Recall/ %	mAP50/ %	mAP50-95/ %	Params/ 10 ⁶	FLOPs/ 10 ⁹
√		√	88.3	67.7	84.9	63.5	87.2	79.5	86.6	65.6	12.6	79.0
	√	√	87.3	67.0	83.9	61.7	85.1	79.3	85.6	64.4	22.4	57.3
√	√	√	87.6	68.1	86.9	64.9	88.6	78.9	87.3	66.5	23.2	85.4

注: √表示应用此改进; 加粗项为在此改进下效果最优。

由表 5 可知, 在单独引入 HSAB 模块后, 模型精度即获得显著提升, 相较于基线模型 mAP50 评价指标提升 2.9%, mAP50-95 评价指标提升 1.6%, 这可以说明 HSAB 模块在增强模型特征表达、缓解细粒度信息丢失问题方面效果显著。在单独引入 WE-CDFA 模块后, 模型精度也获得了明显的提升, mAP50 评价指标提升 3.2%, 由此可以说明该模块可以有效缓解遥感图像中背景复杂、目标边界模糊等问题。在单独引入 SEAD 检测头后, 模型的 mAP50 精度提升约 2.4%, 由此可以看出, 该改进对模型在小目标物体的检测精度方面有所提升。

除此之外, 将多个改进模块进行组合之后, 会进一步提升模型的性能, 其中, 将 HSAB 模块与 WE-CDFA 模块进行组合后, 模型 mAP50-95 精度达到了 65.1%, 远大于单独采用 HSAB 或 WE-CDFA 模块的 63.8% 与 63.0%。将 3 个改进都应用于模型中后, 模型在 ShipRSImageNet 数据集上取得了最优的检测性能, mAP50 提升至 87.3%, 相较于基线模型性能提升高达 5.9%。mAP50-95 提升至 66.5%, 相比于基线模型性能提升 4.3%。这也进一步说明 HSAB 模块、WE-CDFA 模块和 SEAD 检测头具备良好的互补性。

为了进一步说明本文所改进模块的有效性, 本文采用热力图对引入 WE-CDFA 模块与 SEAD 模块前后的结果进行对比。对比图如图 7 所示。

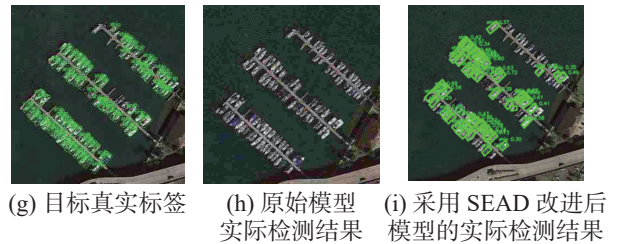
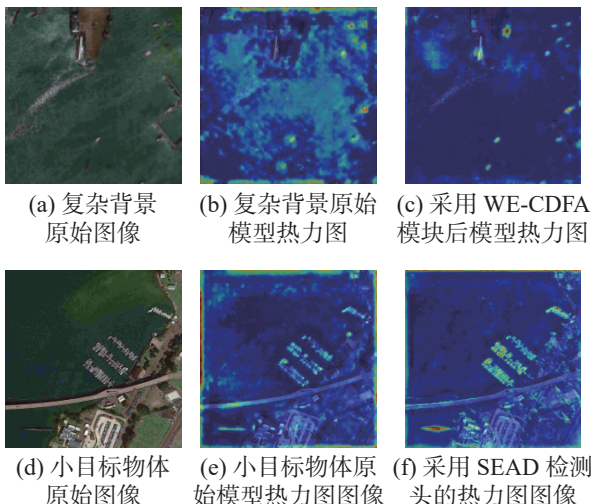


图 7 应用部分模块前后图像的热力图与部分实际检测结果

Fig. 7 Heatmaps before and after applying modules and corresponding detection results

图 7 中高亮部分代表模型对该区域的关注度较高, 暗部表示模型对该区域关注度较低。由图 7(b) 可知, 当模型没有采用 WE-CDFA 模块时, 模型的关注区域相对来说比较分散, 对图像背景区域也存在着较多的无效关注, 难以聚焦于真正的目标区域。由图 7(c) 可以看出, 采用 WE-CDFA 模块改进后的模型则更加关注于舰船目标, 对目标背景的干扰可以起到很好的抑制作用。由此可以看出 WE-CDFA 模块可以有效缓解遥感图像背景复杂所带来的问题, 有效提升了模型对前景目标的感知能力, 改善了模型在遥感图像中由于目标边界模糊、背景复杂所带来的检测性能下降问题。

由图 7(e) 与图 7(h) 可知, 原始模型对小目标物体的注意力并不突出, 缺乏对细小目标的有效响应, 并不能对遥感图像下的细小目标物体进行有效检测。由图 7(f) 与图 7(i) 可以看出, 采用 SEAD 改进后的模型可以更好地使模型关注小目标物体, 极大地缓解了模型对小目标物体的漏检问题。由此可见本文所提出改进的有效性。

3.4.4 不同模型对比实验

为了验证本文所提出的算法的有效性与先进性, 本文与最近几年的算法进行了对比, 分别在 ShipRSImageNet 数据集和 VisDrone2019 数据集上进行。实验结果如表 6 和表 7 所示。

从表 6 可以看出, 本文所提出的改进算法在 ShipRSImageNet 数据集上取得了 87.3% 的 mAP50 和 66.5% 的 mAP50-95, 均显著高于现有的主流算法, 具备良好的实际工程应用的潜力。

表 6 不同模型在 ShipRSImageNet 上的实验结果

Table 6 Performance comparison of various models on ShipRSImageNet

模型	mAP50/%	mAP50-95/%	Params/10 ⁶	FLOPs/10 ⁹
Faster R-CNN ^[22]	67.4	43.2	41.6	206.0
SSD ^[23]	60.5	33.8	24.4	30.7
FCOS ^[24]	38.3	20.0	32.1	198.0
YOLOX ^[25]	81.5	59.8	54.2	77.8
DAB-DETR ^[26]	78.3	57.7	43.7	102.0
DDQ ^[27]	78.8	59.9	48.3	274.0
CM-YOLO ^[28]	79.3	60.4	49.72	70.9
A2Net ^[29]	82.1	64.2	42.5	67.4
SVSDet ^[30]	84.5	63.6	—	—
YOLOv8n-obb	80.2	53.4	3.0	8.1
YOLOv8s-obb	81.4	62.2	11.4	29.4
本文模型	87.3	66.5	23.2	85.4

表 7 不同模型在 VisDrone2019 数据集上的实验结果

Table 7 Experimental results of different models on the VisDrone2019 dataset

模型	mAP50/%	mAP50-95/%	Params/10 ⁶	FLOPs/10 ⁹
YOLOX-Tiny ^[25]	27.8	14.8	5.0	7.5
RTMDet-Tiny ^[31]	31.2	18.4	4.8	8.0
RTDETR ^[32]	33.3	18.5	20.0	60.0
YOLOv8n	25.9	14.4	3.0	8.1
YOLOv8s	31.4	18.1	11.1	28.5
YOLOv8m	33.2	19.0	25.8	78.7
YOLOv10s	32.3	17.9	7.2	21.4
YOLOv10m	34.5	19.5	15.3	58.9
YOLOv11s	31.3	17.6	9.4	12.3
YOLOv11m	35.0	20.3	20.0	67.7
YOLOv12m	33.6	19.2	20.1	67.2
FBRT-YOLO-S ^[33]	32.3	18.3	2.9	22.9
FBRT-YOLO-M ^[33]	34.4	19.6	7.36	58.7
本文模型	35.5	20.9	22.9	85.4

从表 7 可以看出, 本文改进后的算法模型在 VisDrone2019 数据集上 mAP50 精度达到了 35.5%, 高于大多数主流算法, 相比于原始未改进模型 mAP50 精度提升 4.1%。进一步验证了本文算法在复杂场景下的鲁棒性和实用性。通过上面两个数据集的实验可以看出, 本文所提出的改进算法在不同类型的遥感数据集上都取得了不错的成绩, 具备较强的实际应用价值。

3.4.5 检测结果可视化

为了更加清晰和直观地展示本文所提出模型在遥感图像目标检测任务上的实际效果, 本文对比了改进前后模型在 ShipRSImageNet 和 VisDrone-2019 数据集上的检测效果。改进前后模型在 ShipRSImageNet 数据集上的检测效果如图 8 所示, 在 VisDrone2019 数据集上的检测效果如图 9 所示。

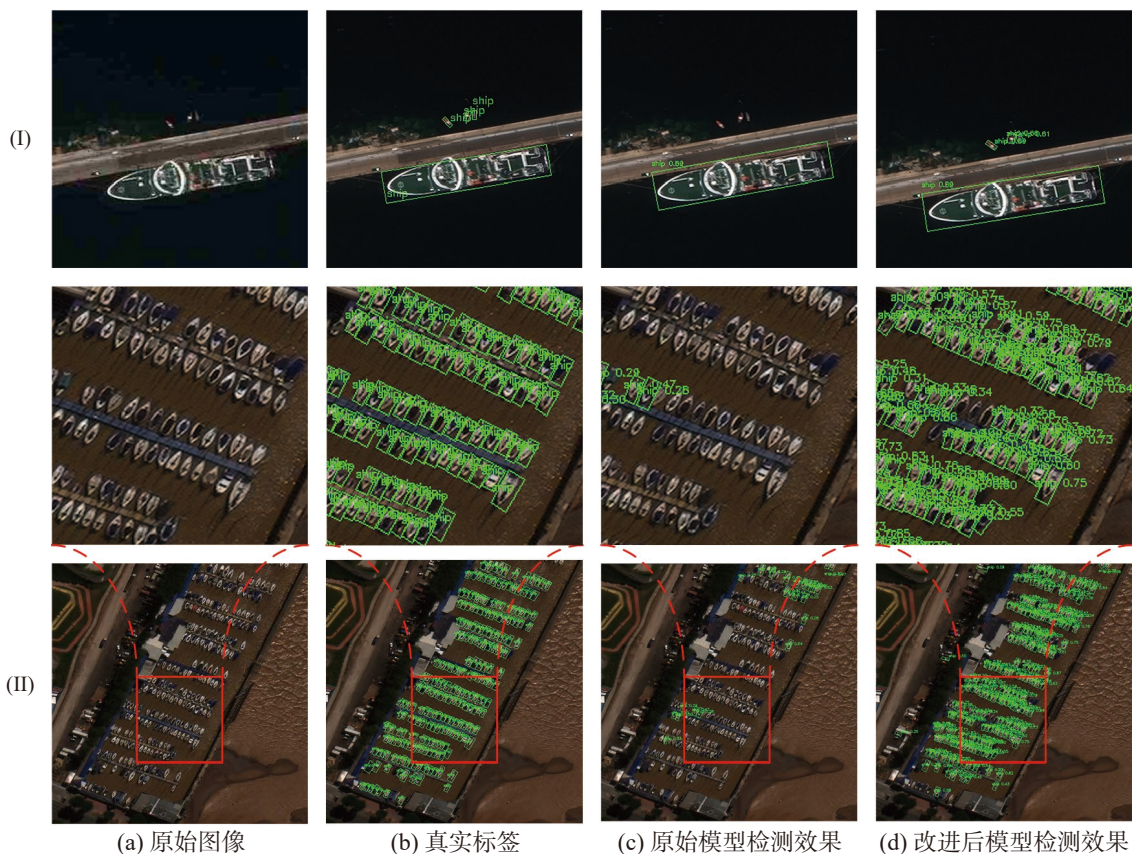


图 8 改进前后模型在 ShipRSImageNet 数据集上的检测效果

Fig. 8 Detection performance before and after improvement on ShipRSImageNet

由图 8 可以看出, 改进前模型在针对细小密集目标的检测效果较差, 存在着较多漏检问题, 而改进后模型在针对细小目标的检测效果得到了明显的提升, 漏检问题得到了极大的缓解。由此可以看出改进后模型的有效性。

由图 9(I) 可以看出, 改进后模型在对小目标

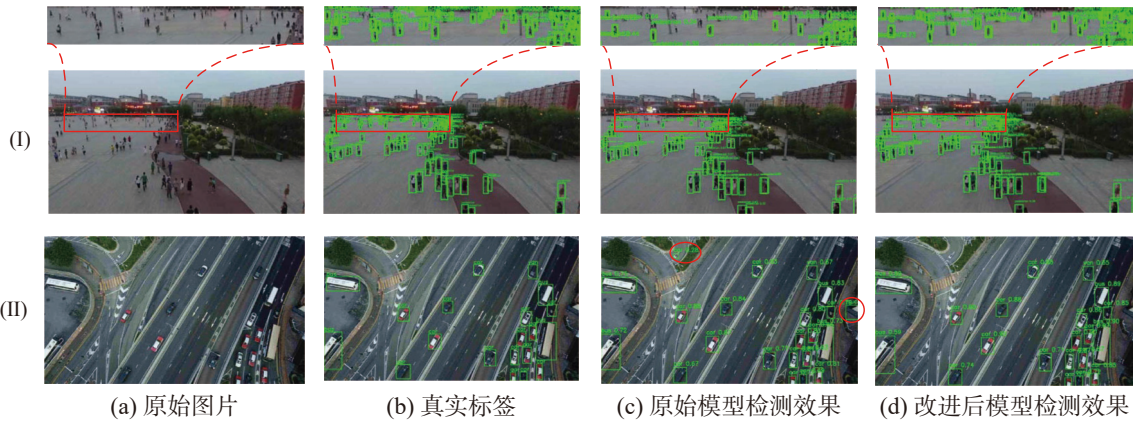


图 9 改进前后模型在 VisDrone2019 数据集上的检测效果

Fig. 9 Detection performance before and after improvement on VisDrone2019

由上述可视化检测效果可以说明, 本文所提出方法在遥感图像目标检测任务中的实际应用价值高于实际检测效果优异, 改进后模型不仅提升了检测准确率, 也缓解了复杂背景对模型所造成的干扰, 提升了模型在复杂环境下对小目标物体的感知能力。

4 结束语

本文针对遥感图像目标检测中存在的目标尺度变化较大、背景复杂和小目标物体检测困难等问题, 提出了一种基于分块注意力与小波增强特征聚合的遥感图像目标检测算法。通过采用 HSAB 模块替代主干网络中的卷积, 有效缓解了下采样过程中存在的细节信息丢失的问题, 并增强了模型对关键目标区域的感知能力; 设计了基于 Haar 小波变换与局部注意力机制的 WE-CDFA 模块, 实现了前景信息增强与背景建模的结合, 提高了复杂背景下模型的目标辨识能力; 设计 SEAD 浅层检测头, 增强模型对小目标物体的检测能力。通过大量的实验表明, 本文所提出改进算法在现有数据集上均取得了不错的成绩, 证明了所提出模型具备很好的鲁棒性。

尽管本文所提出方法在模型精度方面得到了显著提升, 但仍然存在很大的不足与待改进之处。由于改进后的模块参数量相较于原始模块较大, 使得模型参数和计算复杂度都有所提升, 不利于在边缘设备或资源受限平台上进行部署。未

物体检测性能方面有了明显提升。由图 9(II) 红圈处可以看出, 针对背景复杂的情况下所存在的误检问题, 改进后模型针对这一问题也得到了极大提升。进一步说明了本文所提出改进模型可以有效解决原始模型所存在的小目标检测效果较差、背景复杂情况下易出现漏检误检等问题。

来的研究将重点围绕模型轻量化以及模型推理高效化进行设计, 在提升检测性能的同时, 降低资源的开销, 使其更好地应用于广泛的计算设备之上。也期待更多的研究人员在遥感图像目标检测领域提出更好的思路 and 想法。

参考文献:

- [1] DALAL N, TRIGGS B. Histograms of oriented gradients for human detection[C]//2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05). Piscataway: IEEE, 2005: 886–893.
- [2] GAN Lu, LIU Peng, WANG Lizhe. Rotation sliding window of the hog feature in remote sensing images for ship detection[C]//2015 8th International Symposium on Computational Intelligence and Design. Piscataway: IEEE, 2016: 401–404.
- [3] BI Fukun, YANG Zhihua, LEI Mingyang, et al. Airport aircraft detection based on local context DPM in remote sensing images[C]//2019 IEEE International Geoscience and Remote Sensing Symposium. Piscataway: IEEE, 2019: 1362–1365.
- [4] DIVVALA S K, EFROS A A, HEBERT M. How important are “deformable parts” in the deformable parts model? [C]//Computer Vision—ECCV 2012. Workshops and Demonstrations. Berlin: Springer, 2012: 31–40.
- [5] WANG Xuan, WANG Aoran, YI Jinglei, et al. Small object detection based on deep learning for remote sensing: a comprehensive review[J]. *Remote sensing*, 2023, 15(13):

- 3265.
- [6] 姬晓飞, 秦宁丽, 刘洋. 多特征的光学遥感图像多目标识别算法[J]. 智能系统学报, 2016, 11(5): 655–662.
JI Xiaofei, QIN Ningli, LIU Yang. Research on multi-feature based multi-target recognition algorithm for optical remote sensing image[J]. CAAI transactions on intelligent systems, 2016, 11(5): 655–662.
- [7] LI Yuanyuan, ZHOU Zhengguo, QI Guanqiu, et al. Remote sensing micro-object detection under global and local attention mechanism[J]. *Remote sensing*, 2024, 16(4): 644.
- [8] 魏谢根, 曹林, 田澍, 等. 基于多层次自注意力增强的遥感目标检测[J]. *激光与光电子学进展*, 2023, 60(20): 2028004.
WEI Xiegen, CAO Lin, TIAN Shu, et al. Remote sensing target detection based on multilevel self-attention enhancement[J]. *Laser & optoelectronics progress*, 2023, 60(20): 2028004.
- [9] SU Haozheng, YOU Yanan, MENG Gang. Multi-scale context-aware R-cnn for few-shot object detection in remote sensing images[C]//2022 IEEE International Geoscience and Remote Sensing Symposium. Piscataway: IEEE, 2022: 1908–1911.
- [10] 姚婷婷, 肇恒鑫, 冯子豪, 等. 上下文感知多感受野融合网络的定向遥感目标检测[J]. *电子与信息学报*, 2025, 47(1): 233–243.
YAO Tingting, ZHAO Hengxin, FENG Zihao, et al. A context-aware multiple receptive field fusion network for oriented object detection in remote sensing images[J]. *Journal of electronics & information technology*, 2025, 47(1): 233–243.
- [11] 文斌, 张俊, 王浚银, 等. 基于 FNM-Net 的轻量级遥感目标检测算法[J]. *现代电子技术*, 2025, 48(13): 1–10.
WEN Bin, ZHANG Jun, WANG Junyin, et al. Lightweight remote sensing object detection algorithm based on FNM-Net[J]. *Modern electronics technique*, 2025, 48(13): 1–10.
- [12] 周沁坤, 周华平, 孙克雷, 等. ARST-YOLOv7: 用于航空遥感图像的小目标检测网络[J]. *计算机工程与应用*, 2025, 61(12): 232–242.
ZHOU Qinkun, ZHOU Huaping, SUN Kelei, et al. ARST-YOLOv7: small target detection network for aerial remote sensing images[J]. *Computer engineering and applications*, 2025, 61(12): 232–242.
- [13] LI Wentong, CHEN Yijie, HU Kaixuan, et al. Oriented RepPoints for aerial object detection[C]//2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2022: 1819–1828.
- [14] 赵文清, 赵振寰, 巩佳潇. 结合倒残差自注意力机制的遥感图像目标检测[J]. *智能系统学报*, 2025, 20(1): 64–72.
ZHAO Wenqing, ZHAO Zhenhuan, GONG Jiachao. Remote sensing image object detection based on inverted residual self-attention mechanism[J]. *CAAI transactions on intelligent systems*, 2025, 20(1): 64–72.
- [15] 许思源, 吴伟林. 多尺度特征融合的遥感图像目标检测算法研究[J]. *计算机工程与应用*, 2024, 60(23): 249–256.
XU Siyuan, WU Weilin. Research on object detection algorithm for remote sensing images based on multi-scale feature fusion[J]. *Computer engineering and applications*, 2024, 60(23): 249–256.
- [16] HOU Qibin, ZHOU Daquan, FENG Jiashi. Coordinate attention for efficient mobile network design[C]//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2021: 13708–13717.
- [17] HOWARD A G, ZHU Menglong, CHEN Bo, et al. MobileNets: efficient convolutional neural networks for mobile vision applications[EB/OL]. (2017–04–17)[2025–01–01]. <https://arxiv.org/abs/1704.04861>.
- [18] LEI Mengqi, WU Haochen, LYU Xinhua, et al. ConD-Seg: a general medical image segmentation framework via contrast-driven feature enhancement[J]. *Proceedings of the AAAI conference on artificial intelligence*, 2025, 39(5): 4571–4579.
- [19] LAU K W, PO Laiman, REHMAN Y A U. Large separable kernel attention: rethinking the large kernel attention design in CNN[J]. *Expert systems with applications*, 2024, 236: 121352.
- [20] CHEN Jierun, KAO S H, HE Hao, et al. Run, don't walk: chasing higher FLOPS for faster neural networks[C]//2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2023: 12021–12031.
- [21] GUO Menghao, LU Chengze, LIU Zhengning, et al. Visual attention network[J]. *Computational visual media*, 2023, 9(4): 733–752.
- [22] REN Shaoqing, HE Kaiming, 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.
- [23] LIU Wei, ANGUELOV D, ERHAN D, et al. SSD: single shot MultiBox detector[C]//Computer Vision–ECCV 2016. Cham: Springer, 2016: 21–37.
- [24] TIAN Zhi, SHEN Chunhua, CHEN Hao, et al. FCOS: fully convolutional one-stage object detection[C]//2019 IEEE/CVF International Conference on Computer Vision. Piscataway: IEEE, 2020: 9626–9635.
- [25] GE Zheng, LIU Songtao, WANG Feng, et al. YOLOX:

- exceeding YOLO series in 2021[EB/OL]. (2021-07-18) [2025-01-01]. <https://arxiv.org/abs/2107.08430>.
- [26] LIU Shilong, LI Feng, ZHANG Hao, et al. DAB-DETR: dynamic anchor boxes are better queries for DETR[EB/OL]. (2022-01-28)[2025-01-01]. <https://arxiv.org/abs/2201.12329>.
- [27] ZHANG Shilong, WANG Xinjiang, WANG Jiaqi, et al. Dense distinct query for end-to-end object detection[C]// 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2023: 7329-7338.
- [28] MIN Lingtong, DOU Feiyang, ZHANG Yani, et al. CM-YOLO: context modulated representation learning for ship detection[J]. *IEEE transactions on geoscience and remote sensing*, 2025, 63: 4202414.
- [29] ZHANG Yani, ER M J. Asymmetric aggregation network for accurate ship detection in optical imagery[J]. *IEEE transactions on geoscience and remote sensing*, 2024, 62: 4211914.
- [30] LIU Shanwei, BU Xi, XU Mingming, et al. SVSDet: a fine-grained recognition method for ship target using satellite video[J]. *IEEE journal of selected topics in applied earth observations and remote sensing*, 2024, 17: 4726-4742.
- [31] LYU Chengqi, ZHANG Wenwei, HUANG Haian, et al. RTMDet: an empirical study of designing real-time object detectors[EB/OL]. (2022-12-14)[2025-01-01]. <https://arxiv.org/abs/2212.07784>.
- [32] ZHAO Yian, LYU Wenyu, XU Shangliang, et al. DETRs beat YOLOs on real-time object detection[C]//2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2024: 16965-16974.
- [33] XIAO Yao, XU Tingfa, XIN Yu, et al. FBRT-YOLO: faster and better for real-time aerial image detection[J]. *Proceedings of the AAAI conference on artificial intelligence*, 2025, 39(8): 8673-8681.

作者简介:



牛为华, 副教授, 博士, 主要研究方向为数字图像处理、模式识别与计算机视觉。主持省部级教改项目 6 项, 主持科研项目 8 项, 参与科研项目 40 余项, 发表学术论文 20 余篇。E-mail: niuwh@ncepu.edu.cn。



郭迅, 硕士研究生, 主要研究方向为深度学习与遥感图像处理。E-mail: guo_x0315@163.com。