

# PSC-YOLO: 一种遥感图像小目标检测方法

孙永鑫, 王民水, 冷亮

吉林大学 地球探测科学与技术学院, 吉林 长春 130026

**摘要:** 针对无人机和卫星获取的高分辨率遥感图像在国土资源调查、城市规划、生态监测和灾害预警等应用中对目标检测精度要求不断提高的问题, 以及传统目标检测方法在处理目标尺度差异大、背景复杂、低对比度和样本稀缺等情况下表现不足, 特别是在多尺度目标和复杂背景场景中小目标检测能力有限的情况, 本文提出了一种改进的 YOLO 小目标遥感图像检测算法 PSC-YOLO。该算法基于 YOLO11n 框架, 并通过引入多种特征增强与优化机制提升模型对小目标及复杂背景的代表能力。首先, 在主干网络中引入 PMKCA 模块, 通过并行多核卷积与通道注意力机制增强多尺度特征提取能力, 提高模型的全局特征建模能力; 其次, 在 SPPF 特征融合层嵌入 SPPCA 模块, 通过多尺度池化与通道注意力机制提升模型对小目标的感知能力; 随后, 在 C3k2 结构中引入 C3k2\_HLCA 模块, 通过局部特征增强与通道注意力机制优化特征表达, 从而进一步提升小目标检测精度; 最后, 设计自适应置信度阈值调整函数, 对 NMS 过程进行优化, 以减少冗余检测框并提高检测性能。实验结果表明, 在 NWPU VHR-10 数据集上, PSC-YOLO 的检测性能优于 YOLOv5、YOLOv10、YOLO11n 和 YOLOv12。在 mAP50 和 mAP50-95 指标上, PSC-YOLO 分别比 YOLOv5 提升 0.8% 和 2.0%, 比 YOLOv10 提升 1.5% 和 2.0%, 比 YOLO11n 提升 0.8% 和 2.0%, 比 YOLOv12 提升 1.5% 和 2.0%。实验结果表明, 所提出的 PSC-YOLO 能够在复杂背景和多尺度目标场景中有效减少漏检, 提高小目标检测精度, 并具有更好的稳定性。该方法为遥感图像的自动化解译与智能化决策提供了更加精准的信息支持。

**关键词:** 小目标检测; 地物识别; 深度学习; YOLO; 遥感图像

## PSC-YOLO: a small target detection method for remote sensing images

Sun Yongxin, Wang Minshui, Leng Liang

College of Geo-Exploration Science and Technology, Jilin University, Changchun 130026, Jilin, China

**Abstract:** The increasing demand for target detection accuracy in applications such as land resource surveys, urban planning, ecological monitoring, and disaster early warning, using high-resolution remote sensing images from drones and satellites, has exposed limitations in traditional detection methods. These limitations are particularly

收稿日期/Received: 2025-08-11; 修订日期/Revised: 2026-03-17; 出版日期/Published: 2026-05-25

基金项目: 国家自然科学基金面上项目 (42171407); 吉林省教育厅科学研究项目 (JKH20241288KJ)

第一作者: 孙永鑫 (2001—), 男, 硕士研究生, 主要从事遥感数据处理与应用研究。E-mail: yongxin23@mails.jlu.edu.cn

\* 通信作者: 王民水 (1989—), 男, 硕士, 高级工程师, 主要从事遥感图像变化检测研究。E-mail: 1543383519@qq.com

© Editorial Office of World Geology. This is an open access article under the CC BY-NC-ND 4.0 license.

noticeable when there are significant scale differences, complex backgrounds, low contrast, and sparse samples. In particular, detecting small targets in multi-scale and complex background scenes has been challenging. To address these issues, this paper proposes an improved YOLO small target remote sensing image recognition algorithm called PSC-YOLO. It is based on the YOLO11n framework and incorporates multiple feature enhancement and optimization mechanisms. These additions improve the model's ability to handle small targets and complex backgrounds. First, the PMKCA module is introduced into the backbone network. It uses parallel multi-core convolution and channel attention mechanisms to enhance multi-scale feature extraction and improve global feature modeling. Second, the SPPCA module is embedded into the SPPF feature fusion layer. It employs multi-scale pooling and channel attention mechanisms to boost small target perception. Next, the C3k2\_HLCA module is integrated into the C3k2 structure. It enhances local feature representation and optimizes feature expression using channel attention mechanisms, which further improves small target detection accuracy. Finally, an adaptive confidence threshold adjustment function is designed to optimize the NMS process. This reduces redundant detection boxes and improves performance. Experimental results show that PSC-YOLO outperforms YOLOv5, YOLOv10, YOLO11n, and YOLOv12 on the NWPU VHR-10 dataset. For the mAP50 and mAP50-95 metrics, PSC-YOLO achieves a 0.8% and 2.0% improvement over YOLOv5, a 1.5% and 2.0% improvement over YOLOv10, a 0.8% and 2.0% improvement over YOLO11n, and a 1.5% and 2.0% improvement over YOLOv12. The results demonstrate that PSC-YOLO effectively reduces missed detections in complex background and multi-scale target scenarios. It also improves small target detection accuracy and provides better stability. This method offers more precise information support for the automated interpretation and intelligent decision-making of remote sensing images.

**Keywords:** small object detection; object recognition; deep learning; YOLO; remote sensing imagery

## 0 引言

遥感图像地物检测是国土资源调查、生态环境监测、城市规划和灾害预警等领域的重要技术支撑<sup>[1]</sup>。随着卫星与无人机载荷分辨率的不断提高,遥感影像在空间细节和地物类型上愈加丰富,同时呈现出目标尺度差异大、场景背景复杂和光照多变等特征<sup>[2-3]</sup>。受制于影像获取与精确标注的高成本,能够用于模型训练的样本数量普遍有限,使得现有检测算法在小目标场景下容易出现漏检、误检和过拟合等问题,严重制约了地物识别的精度与鲁棒性<sup>[4-5]</sup>。因此,如何在有限样本条件下实现多类别、小尺度目标的高效检测与识别,已成为当前遥感智能解译的关键难题之一<sup>[6]</sup>。

随着计算机技术的发展和深度学习网络的进步,目标检测技术呈现出跨越式进步<sup>[7]</sup>, Transformer 架构的兴起为全局特征建模提供了新的思路,其自注意力机制能够捕获长程依赖关系,检测效果显著提高。传统目标检测算法依赖手工特征与滑动窗口,在面对背景复杂、尺度跨度大的遥感影

像时难以满足高精度要求<sup>[8]</sup>。基于卷积神经网络的检测算法逐渐成为主流,主要分为一阶段和二阶段两类:一阶段目标检测方法,以端到端结构为特点,通过全局回归直接完成类别与位置预测,在保证精度的同时显著提升了检测速度,并在无人机航拍与高分辨率遥感影像的实时应用中得到广泛采用<sup>[9]</sup>;二阶段的目标检测是以区域提议为基础的兩阶段方法,通过候选区域生成与分类回归实现较高精度,但计算开销和推理速度相对较低<sup>[10]</sup>。YOLO系列作为典型的一阶段检测算法,通过端到端的高效结构在检测速度和精度之间取得了良好平衡,已成为遥感目标检测研究与应用中的重要基线方法。在遥感场景中仍然存在模型复杂度高、计算开销大的问题,因此研究者开始尝试结合卷积的局部建模优势,对模型进行轻量化改进<sup>[11]</sup>。

为了提高 YOLO 模型的目标检测效果,现有研究主要在多尺度特征增强、注意力机制引入以及网络轻量化设计等方面展开优化<sup>[12]</sup>。其中,多尺度特征增强通过构建特征金字塔、跨层融合等策略,有效强化了对不同尺寸目标的感知能力<sup>[13]</sup>;注意

力机制通过引导网络聚焦于关键区域与显著通道, 提升了在复杂背景和小目标场景下的识别效果<sup>[14]</sup>; 轻量化设计则以深度可分离卷积、结构重参数化等方法为代表, 在尽量保证检测精度的同时显著降低了模型参数量和计算开销, 从而更好适应无人机与卫星等资源受限平台上的遥感应用需求<sup>[15]</sup>。尽管这些方法在全局特征建模和小目标检测方面取得了积极进展, 但在小目标条件下同时兼顾检测精度与轻量化仍然存在较大挑战, 尤其是如何在有限训练数据下有效提升遥感图像中小目标的检测能力, 成为当前研究的重点与难点。

针对遥感图像中地物目标尺度差异大、背景复杂以及小目标易漏检的问题, 本文提出了一种改进的 PSC-YOLO 算法。该方法从增强多尺度特征表达和提升小目标感知能力两方面进行了改进: 在主干网络深层引入 PMKCA (parallel multi-kernel channel attention) 模块, 通过多尺度并行卷积与通道自适应加权, 显著增强了长程依赖建模和全局上下文特征提取能力, 且保持了较低的计算复杂度; 在多尺度特征融合阶段嵌入 SPPCA (parallel multi-kernel channel attention) 结构, 结合空间金字塔池化与通道注意力机制, 充分聚合了上下文信息。为了进一步提高小目标检测精度, 在主干网络的关键残差结构 C3k2 模块中加入 HLCA (hybrid local channel attention) 模块, 利用深度残差与局部特征增强, 提升了细粒度结构响应与小目标定位能力。此外, 针对不同图像和场景中的置信度分布差异, 笔者提出了一种自适应置信度阈值调整方法, 通过计算每批图像预测框的置信度分布并动态调整 NMS (Non-Maximum Suppression) 阈值, 有效减少冗余框并提升检测精度, 尤其在复杂背景和小目标场景中表现出显著优势。

## 1 PSC-YOLO 模型

为了解决多目标复杂场景下小目标易出现漏检与误判的问题, 在 YOLO11n 网络模型的基础上进行了结构与特征建模的优化设计: 在主干网络的高语义层引入 PMKCA 模块, 通过多核并行卷积与通道注意力的融合, 在保持轻量化的同时显著增强了全局特征建模与长程依赖表达能力; 在特征融合阶

段嵌入 SPPCA 结构, 结合空间金字塔池化与通道注意力, 有效捕获多尺度上下文信息并强化关键通道选择; 同时, 将骨干网络中的关键残差结构 C3k2 改进为 C3k2\_HLCA, 在残差连接中引入局部特征增强与通道注意力机制, 从而提升细粒度特征提取能力并增强小尺度目标的检测精度与鲁棒性。上述三种改进模块协同作用, 构成了 PSC-YOLO 网络整体架构, 其结构如图 1 所示。

### 1.1 多尺度卷积 PMKCA 模块

基础 YOLO11n 主干网络中主要依赖固定卷积核提取局部特征, 感受野有限, 对大范围上下文和跨尺度目标的长程依赖建模不足; 在小目标与背景差异微弱的遥感场景下, 弱语义的小目标更易被背景噪声淹没。PMKCA 模块在主干网络的高语义层并行采用多核深度可分离卷积, 以获取不同感受野的特征信息:

$$F_k = \varphi(\text{BN}(\text{DWCONV}_k(X))), k \in \{3, 5, 7\} \quad (1)$$

不同尺度特征在通道维度上拼接并经  $1 \times 1$  卷积压缩:

$$F = \varphi(\text{BN}(\text{CONV}_{1 \times 1}([F_3, F_5, F_7]))) \quad (2)$$

再利用全局平均池化和两层全连接生成通道权重:

$$s = \sigma(W_2 \varphi(W_1 \text{GAP}(F))) \quad (3)$$

最终输出由特征与权重逐通道相乘得到:

$$Y = X \odot s + X \quad (4)$$

式中:  $X$  为输入特征图, DWCONV 为深度可分离卷积,  $k$  是卷积核尺寸, BN 为批归一化,  $F_k$  为第  $k$  个卷积核提取的特征图,  $[ ]$  为特征拼接, CONV<sub>1×1</sub> 为一乘一卷积,  $F$  为融合后的多尺度特征图, GAP 为全局平均池化,  $W_1$  为第一层全连接层权重矩阵,  $W_2$  为第二层全连接层权重矩阵,  $s$  为通道注意力权重,  $\odot$  表示逐通道乘法,  $\varphi$  和  $\sigma$  分别为 ReLU 与 Sigmoid 激活函数,  $Y$  为 PMKCA 模块输出特征。

PMKCA 通过并行多尺度卷积与通道加权融合, 有效增强了多尺度上下文信息的捕获与关键通道的

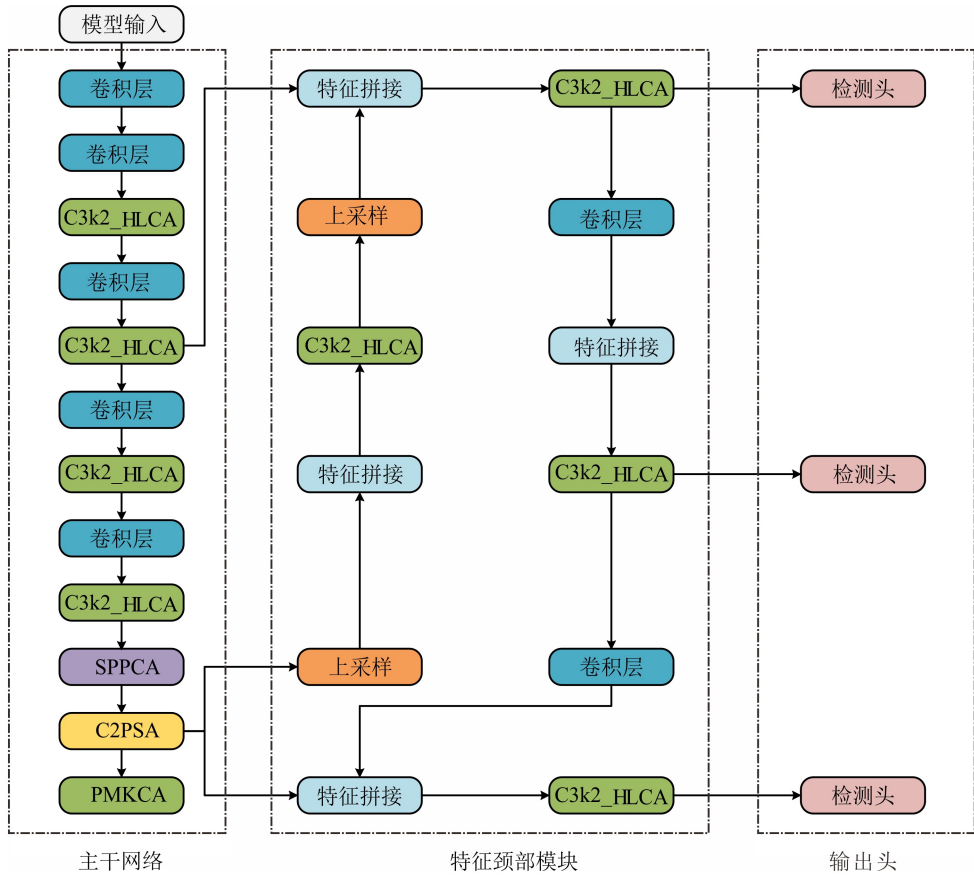


图1 PSC-YOLO 网络结构

Fig.1 PSC-YOLO network architecture

选择能力，为弱纹理或被遮挡的小目标提供了更充分的全局语义支持，从而提升其在复杂背景中的显著性与检测概率。其结构如图2所示。

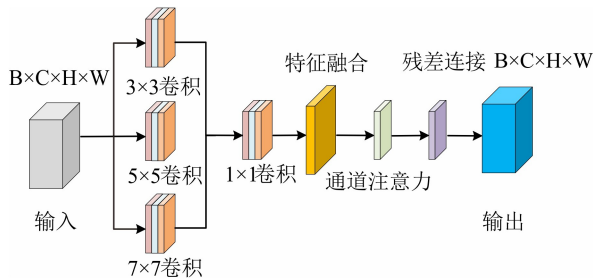


图2 PMKCA 模块结构图

Fig.2 Structure diagram of PMKCA module

## 1.2 空间金字塔池化 SPPCA 模块

YOLO11n 中的 SPPF 结构仅利用固定尺度的池化层扩展感受野，缺少通道自适应能力，难以充分整合不同尺度上下文并突出关键特征；在存在多尺

度目标和背景干扰的遥感图像中，小目标边缘和纹理易在融合过程中被过度平滑或弱化。在 SPPCA 模块中，笔者使用了不同尺寸的池化核（ $5 \times 5$ 、 $9 \times 9$ 、 $13 \times 13$ ），目的是捕捉不同范围的特征。小尺度的池化核（ $5 \times 5$ ）主要关注局部细节，而大尺度的池化核（ $13 \times 13$ ）可以捕捉更广泛的上下文信息。通过结合不同尺寸的池化核，模型能同时处理细节和大范围的特征，从而更好地检测遥感影像中的各种地物。其运算流程为：

$$P_k = \text{MaxPool}_k(X), k \in \{5, 9, 13\} \quad (5)$$

$$Z = [X, P_5, P_9, P_{13}] \quad (6)$$

$$F = \text{CONV}_{1 \times 1}(Z) \quad (7)$$

$$Y = \text{HLCA}(F) \quad (8)$$

式中： $X$  为输入特征图， $\text{MaxPool}_k$  表示池化核

大小为  $k$  的最大池化操作,  $P_k$  为池化后特征图, 表示特征拼接,  $\text{CONV}_{1 \times 1}$  用于特征压缩与融合, HLCA 为混合局部通道注意力模块多尺度池化能够有效扩大感受野并引入丰富的上下文信息, 而 HLCA 通道注意力在特征融合后可自适应增强与目标相关的关键信息。由此构建的 SPPCA 模块不仅能够在保持与 SPPF 相近计算开销和实时性的同时, 兼顾大范围语义背景信息的捕捉与细粒度目标特征的表达, 还在密集小目标或存在遮挡的复杂环境中显著提升了检测的召回率与定位精度。其具体结构如图 3 所示。

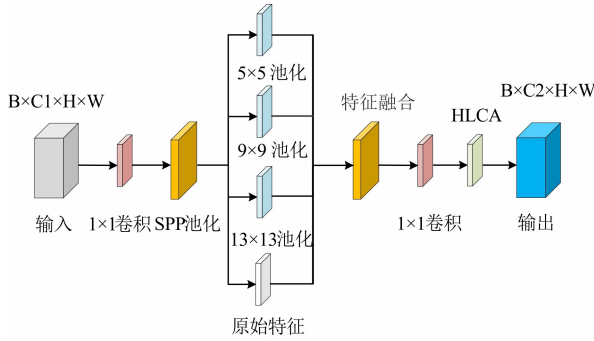


图 3 SPPCA 模块结构图

Fig. 3 Structure diagram of SPPCA module

### 1.3 多尺度特征增强 C3k2\_HLCA 模块

原始 C3k2 模块在小目标检测中对细节的保持能力有限。遥感图像中的小目标通常面积较小、边缘模糊且纹理对比度低, 经过多次卷积与下采样后, 细粒度特征容易丢失, 进而导致漏检和定位偏差。为缓解该问题, 本文设计 C3k2\_HLCA 模块, 在主分支中引入 Bottleneck\_HLCA 以强化局部结构响应, 其计算过程为:

$$B(X) = \text{MLCA}(\text{CONV}_{3 \times 3}(\text{CONV}_{1 \times 1}(X))) \quad (9)$$

主分支输出与旁分支经  $1 \times 1$  卷积后的特征拼接, 并通过  $1 \times 1$  卷积压缩:

$$Z = \text{CONV}_{1 \times 1}([B(X), \text{CONV}_{1 \times 1}(X)]) \quad (10)$$

最后经 MLCA 加权输出:

$$Y = \text{HLCA}(Z) \quad (11)$$

式中: CONV 为卷积计算,  $B(X)$  为输出的特

征图,  $Z$  为融合后的中间特征图, HLCA 为分层通道注意力,  $Y$  为最终输出特征图。通过残差连接与多层局部特征增强, C3k2\_HLCA 能够有效保留并强化小目标的边缘、纹理及高频结构特征。在保证网络轻量化与推理速度的同时, 该模块提升了小目标的特征表达与定位精度, 尤其在密集分布或背景复杂的场景中表现突出, 可显著降低漏检率并增强整体鲁棒性。其结构如图 4 所示。

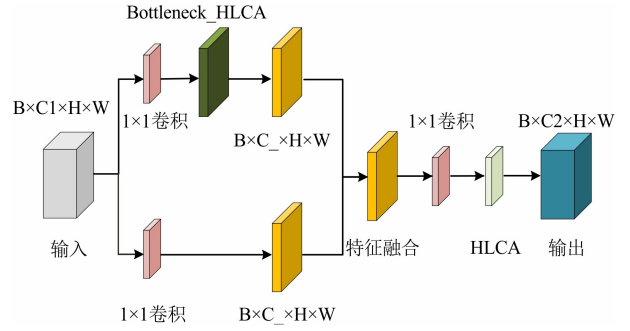


图 4 C3k2\_HLCA 模块结构图

Fig. 4 Structure diagram of C3k2\_HLCA module

### 1.4 自适应置信度阈值调整函数

在目标检测的非极大值抑制 (NMS) 过程中, 传统方法使用固定的置信度阈值来筛选候选框。但在目标分布复杂或置信度差异较大的情况下可能导致冗余框或漏检。为了解决这个问题, 笔者提出了一种自适应置信度阈值调整的方法, 通过动态计算每批图像中预测框的置信度分布, 并基于该分布动态设置置信度阈值, 从而优化候选框筛选过程。通过式 (12) 计算:

$$T_{\text{conf}} = \text{Quantile}(P(\text{score}), q) \quad (12)$$

式中:  $P(\text{score})$  表示当前批次所有预测框的置信度分布, Quantile 是用于计算置信度分布中某一百分位数的操作,  $q$  为该分位数。通过自适应函数, 阈值能够根据每次推理的得分分布自适应地调整, 确保只有高置信度的框被保留, 避免了固定阈值带来的局限性, 并在复杂场景中提升了目标检测的精度与召回率。

## 2 实验与结果分析

### 2.1 实验环境与参数设置

本实验在 Windows 11 + Anaconda 虚拟环境下完

成, 硬件平台为 Intel (R) Core (TM) Ultra 7 155H、32 GB 内存、NVIDIA GeForce RTX 4050 GPU (6 GB 显存)。软件环境为 Python 3.10.18、PyTorch2.6.0 + cu124、CUDA12.4、cuDNN9.1.0。模型训练采用 SGD 优化器, 初始学习率 0.01, 末端学习率衰减比例 0.01, 动量 0.937, 权重衰减 0.0005, 批次大小 8, 输入图像尺寸 640 × 640, 总训练轮数 300, 并启用 Mosaic 拼接、RandAugment 增强策略。

## 2.2 数据集

本文使用的实验数据集为 NWPU VHR-10 高分辨率遥感图像目标检测数据集。该数据集由西北工业大学于 2014 年发布, 影像数据来源于 Google Earth 和 Vaihingen 数据集, 并由遥感领域的专家进行了精确标注。所有图像的空间分辨率约为 0.5 m, 涵盖了丰富多样的地表场景和复杂背景, 具有目标尺度跨度大、类别多样、背景复杂等特点, 是检验小样本遥感目标检测方法的一个重要基准。

数据集共包含 800 张图像, 其中 650 张图像包含地物目标, 剩余 150 张图像为背景图像。本文实验仅使用包含目标的 650 张正样本, 这些样本涵盖了 10 类典型地物, 包括飞机、舰船、油罐、棒球场、网球场、篮球场、田径场、港口、桥梁和车辆。这些目标类别在实际遥感图像中具有很高的代表性, 涉及城市、工业、交通等多个领域, 且存在目标尺度差异、背景复杂、遮挡等问题, 符合小目标检测任务的研究需求。将 650 张样本按 8 : 2 的比例随机划分为训练集和验证集。为提升模型的鲁棒性和泛化能力, 训练前笔者对所有样本进行了尺寸归一化和标准化处理。采用了 Mosaic 拼接、随机翻转、颜色扰动和亮度调整等数据增强策略, 有效提升了模型在小目标场景下的表现, 帮助模型更好地适应不同场景变化, 提高了其对未见数据的适应性和检测精度。

## 2.3 评价指标

为全面衡量 PSC-YOLO 网络在遥感图像小目标检测任务中的性能, 本文选取以下指标进行评价:

### (1) 精确率 (precision, P)

精确率表示预测为正样本的目标中真实为正样本的比例, 用于衡量模型抑制误检的能力, 其计算公式为:

$$P = \frac{TP}{TP + FP} \quad (13)$$

式中:  $TP$  和  $FP$  分别表示真正例和假正例的数量。在小目标检测中, 背景复杂且与目标尺寸接近, 易产生伪检测。较高的精确率说明模型能够有效区分小目标与背景, 减少因背景噪声导致的误检, 提高检测结果的可靠性。

### (2) 召回率 (recall, R)

召回率表示真实正样本中被正确检测出的比例, 用于衡量模型降低漏检的能力, 其计算公式为:

$$R = \frac{TP}{TP + FN} \quad (14)$$

式中:  $FN$  为假负例的数量。小目标由于像素少、边缘模糊, 模型容易忽略而造成漏检。较高的召回率表明模型能够捕捉到更多小尺寸目标, 提高整体检出率, 是评价小目标检测性能的重要指标。

### (3) 平均精度均值 (mean average precision, mAP)

mAP 是综合评估检测精度的核心指标。单类别平均精度 (AP) 定义为精确率 - 召回率 (P-R) 曲线下的面积:

$$AP = \int_0^1 P(R) dR \quad (15)$$

mAP 为所有类别 AP 的算术平均。本文采用 mAP50 和 mAP50-95 两种形式, 其中 mAP50 在交并比 (IoU) 阈值为 0.5 时衡量常规检测表现, mAP50-95 则在 IoU 0.5 至 0.95 区间以 0.05 为步长取平均, 更严格地反映模型在不同匹配标准下的综合检测和定位能力。对于小目标检测任务, mAP50-95 对边界框的定位精度要求更高, 能够敏感地反映模型在小目标边界回归与精细定位方面的表现。

这些评价指标从误检控制、漏检控制及整体精度等方面全面刻画了模型在小样本遥感图像小目标检测中的性能, 为改进算法的有效性提供了有力支撑。

## 2.4 实验设计及结果分析

### 2.4.1 实验设计

为了验证所提出改进模块的有效性, 本文进行

了消融实验, 并与对照方法进行了对比实验。消融实验通过逐步引入 PMKCA、C3k2\_HLCA、SPPCA 和 NMS 模块, 分析了每个模块的独立贡献及模块间的协同增益。具体实验包括 YOLO11n 加入不同模块 (PMKCA、C3k2\_HLCA、SPPCA 和 NMS) 及其组合, 旨在评估各模块对模型性能的影响。对比实验比较了 YOLOv5、YOLOv10、YOLO11n、YOLOv12 和 PSC-YOLO 模型的性能。所有实验均在相同的数据集、训练环境和超参数设置下进行, 唯一变化是逐步引入模块或替换模型。通过这种设计, 评估了各模块的独立作用、模块间的协同效应及整体方法的有效性, 确保了实验结果的公平性和可比性。

#### 2.4.2 结果分析

表 1 展示了各消融实验的结果及基准模型性能。为了确保结果的可靠性, 表中所有数据均来自相同训练集和验证集下的实验, 训练环境、超参数保持一致, 唯一变化是逐步引入或去除网络模块。消融实验结果表明 YOLO11n 作为基准模型, 其 mAP50 为 0.917, mAP50-95 为 0.597。引入 PMKCA 模块后, 精度和 mAP50 分别提升了 0.8% 和 1.6%, 表明 PMKCA 模块通过增强特征提取能力, 特别是针对小目标的检测, 显著提高了模型的整体精度。这一提升表明, PMKCA 模块能够有效提升

模型的全局特征建模能力, 增强对小目标的检测能力。引入 C3k2\_HLCA 模块后, 召回率提升了 0.5%, mAP50 提升了 0.7%, mAP50-95 提升了 0.5%。这一结果表明, C3k2\_HLCA 模块通过加强局部特征的提取, 显著提升了小目标的召回率, 尤其是在复杂场景中的检测能力。这一模块对于提高小目标的识别和精确定位至关重要, 尤其在多目标和密集背景下效果更加明显。引入 SPPCA 模块后, 召回率提升了 2.7%, mAP50 提升了 1.5%, mAP50-95 提升了 1.5%。这一模块通过多尺度池化和通道注意力机制优化了模型在复杂场景中的表现, 尤其在小目标检测方面表现突出。SPPCA 能够有效增强模型在多尺度和变化背景下的适应能力, 提高了模型对小目标的感知能力, 显著减少了漏检现象。PSC-YOLO 在 YOLO11n + PMKCA + SPPCA + C3k2\_HLCA 的基础上, 精度提升了 0.8%, mAP50-95 提升了 2.0%。这一提升得益于 NMS 的引入, PSC-YOLO 有效减少了冗余检测框, 优化了目标框的选择, 显著提高了检测精度和稳定性。消融实验表明, 逐步引入不同模块 (PMKCA、C3k2\_HLCA、SPPCA) 有效提升了 PSC-YOLO 在小目标检测中的精度、召回率以及稳定性, 特别是在复杂背景下, 相比传统 YOLO 模型, 性能提升明显。

表 1 模型消融实验整体指标对比表

Table 1 Table of overall performance metrics of ablation experiments

模型	P	R	mAP50	mAP50-95
YOLO11n	0.935	0.872	0.917	0.597
YOLO11n + PMKCA	0.943	0.860	0.932	0.597
YOLO11n + C3k2_HLCA	0.921	0.868	0.923	0.600
YOLO11n + SPPCA	0.912	0.896	0.931	0.606
YOLO11n + PMKCA + SPPCA	0.893	0.889	0.918	0.604
YOLO11n + C3k2_HLCA + SPPCA	0.887	0.873	0.917	0.589
YOLO11n + PMKCA + C3k2_HLCA	0.918	0.884	0.929	0.601
YOLO11n + PMKCA + SPPCA + C3k2_HLCA	0.918	0.877	0.931	0.602
PSC-YOLO	0.924	0.875	0.931	0.609

表 2 展示了本文提出的 PSC-YOLO 模型与现有主流 YOLO 系列模型 (YOLOv5、YOLOv10、YOLO11n、YOLOv12) 的性能对比。为了确保对比

的公平性, 所有模型均在相同的数据集、训练环境和超参数设置下进行训练和测试, 训练轮数、学习率、批次大小等条件一致, 唯一变化是网络结构本

身。对比实验精度指标表明, PSC-YOLO 在融合了 PMKCA、SPPCA、C3k2\_HLCA 和自适应置信度阈值调整方法后, 表现出了显著的性能提升。相比于传统的 YOLO 模型, PSC-YOLO 在多个关键指标上都表现出了较强的优势, 特别是在 mAP50 和 mAP50-95 指标上。具体来说, 与基准模型 YOLO11n 相比, PSC-YOLO 在 mAP50 上提升了 0.8%, 在 mAP50-95 上提升了 2.0%。这一提升表明, PSC-YOLO 在处理复杂背景和多尺度目标时具有更强的鲁棒性, 能够有效提升小目标的检测精度。与 YOLOv5 相比, PSC-YOLO 的提升幅度相似, 在 mAP50 上提升了 0.8%, 在 mAP50-95 上提升了 2.0%。虽然 YOLOv5 的整体精度较高, 但 PSC-YOLO 通过引入多个创新模块, 特别是在小目标检测任务上, 展现了更强的优势, 尤其是在多尺度目标和复杂场景中的应用。与 YOLOv10 相比, PSC-YOLO 在 mAP50 和 mAP50-95 上分别提升了 1.5% 和 2.0%, 表现出明显的优势。YOLOv10 在准确率和召回率方面的表现相对较好, 但在复杂背景下的性能提升仍不如 PSC-YOLO。与 YOLOv12 相比, PSC-YOLO 在 mAP50 和 mAP50-95 上分别提升了 1.5% 和 2.0%。YOLOv12 虽然在一些场景下表现良好, 但在面对小目标和复杂场景时, PSC-YOLO 的优势更加明显, 进一步验证了本网络在小目标检测任务中的优越性。

表 2 模型对比实验整体指标结果表

Table 2 Table of overall performance metrics of model comparison experiments

模型	P	R	mAP50	mAP50-95
YOLOv5	0.941	0.896	0.932	0.597
YOLOv10	0.914	0.843	0.911	0.602
YOLO11n	0.935	0.872	0.917	0.597
YOLOv12	0.868	0.832	0.898	0.592
PSC-YOLO	0.924	0.875	0.931	0.609

为分析各模型在不同类别上的检测性能差异, 在相同数据集和训练条件下, 对 YOLOv5、YOLOv10、YOLO11n、YOLOv12 及 PSC-YOLO 进行了类别细节实验, 实验结果如表 3 所示, PSC-YOLO 在多个类别上展现了显著的性能提升。具体而言, 在桥梁类别中, PSC-YOLO 的 mAP50 为 0.775, 相比 YOLOv12 的 0.697 提升了 11.2%; 相比 YOLO11n 的 0.734 提升了 5.58%; 相比 YOLOv10 的 0.773 提升了 0.26%。在 mAP50-95 上, PSC-YOLO 为 0.317, 相比 YOLOv12 的 0.347 提升了 8.7%; 相比 YOLOv5 的 0.277 提升了 14.4%; 相比 YOLO11n 的 0.276 提升了 14.8%。图 5 展示了 PSC-YOLO 在桥梁类别中的检测性能。相较于其他 YOLO 模型, PSC-YOLO 在类别概率上表现突出, 达到了 0.89, 相比于 YOLOv5, PSC-YOLO 在 mAP50

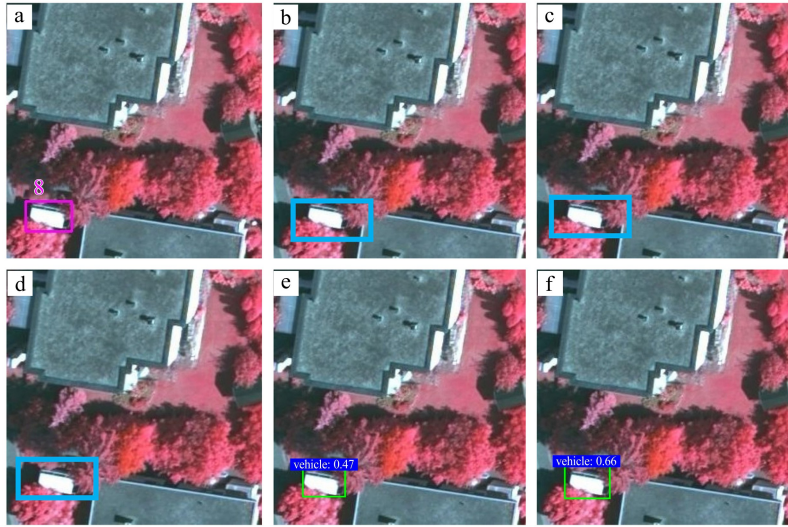
表 3 各模型 mAP50 和 mAP50-95 对比表

Table 3 Table of comparison of mAP50 and mAP50-95 among various models

类别	YOLOv5 (mAP50)	YOLOv10 (mAP50)	YOLO11n (mAP50)	YOLOv12 (mAP50)	PSC-YOLO (mAP50)	YOLOv5 (mAP50-95)	YOLOv10 (mAP50-95)	YOLO11n (mAP50-95)	YOLOv12 (mAP50-95)	PSC-YOLO (mAP50-95)
飞机	0.995	0.995	0.995	0.995	0.995	0.663	0.693	0.667	0.699	0.671
船舶	0.950	0.976	0.994	0.937	0.980	0.625	0.669	0.678	0.600	0.660
储油罐	0.965	0.957	0.923	0.958	0.954	0.521	0.596	0.550	0.604	0.568
棒球场	0.973	0.977	0.955	0.974	0.955	0.764	0.773	0.724	0.778	0.736
网球场	0.975	0.994	0.948	0.915	0.952	0.627	0.640	0.606	0.607	0.600
篮球场	0.743	0.767	0.797	0.770	0.822	0.511	0.544	0.543	0.520	0.520
田径场	0.995	0.906	0.995	0.973	0.995	0.876	0.777	0.864	0.842	0.855
港口	0.979	0.968	0.995	0.939	0.990	0.598	0.616	0.562	0.449	0.642
桥梁	0.870	0.773	0.734	0.697	0.775	0.277	0.493	0.276	0.347	0.317
车辆	0.876	0.856	0.836	0.819	0.890	0.508	0.539	0.496	0.470	0.494

上提升了 5.95%; 相比于 YOLOv10, 提升了 7.23%; 相比于 YOLO11n, 提升了 14.1%; 相比于 YOLOv12, 提升了 43.55%。这一系列提升充分显示了 PSC-YOLO 在复杂背景和小目标检测中的显著优势。尤其在目标识别和定位精度上, PSC-YOLO 通过引入改进的模块, 不仅提升了模型在复杂环境中的鲁棒性, 还在小目标的准确识别和定位方面展现出了优

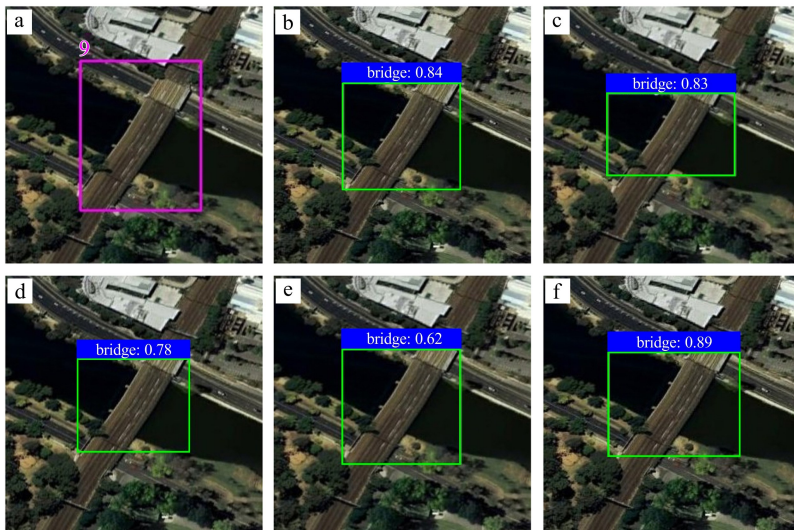
越的能力。在车辆类别中, PSC-YOLO 的 mAP50 为 0.890, 相比 YOLO11n 的 0.836 提升了 6.4%; 相比 YOLOv10 的 0.856 提升了 4.0%; 相比 YOLOv5 的 0.876 提升 1.6%。在 mAP50-95 上, PSC-YOLO 为 0.494, 相比 YOLOv12 的 0.470 提升了 5.1%, 图 6 展示了 PSC-YOLO 在车辆类别中的表现。相比 YOLOv5、YOLOv10 和 YOLO11n, 这



a. Lable; b. YOLOv5; c. YOLOv10; d. YOLO11n; e. YOLOv12; f. PSC-YOLO。

图 5 场景 1 目标检测结果

Fig. 5 Object detection results of situation1



a. Lable; b. YOLOv5; c. YOLOv10; d. YOLO11n; e. YOLOv12; f. PSC-YOLO。

图 6 场景 2 目标检测结果

Fig. 6 Object detection results of situation 2

些模型未能检测到目标, YOLOv12 的类别概率为 0.47, 而 PSC-YOLO 的类别概率达到了 0.66, 提升幅度高达 40.4%。这一显著提升表明, PSC-YOLO 能够有效检测到其他 YOLO 模型未能检测的小目标, 尤其是在复杂和密集目标的场景下, 展现了较强的鲁棒性。PSC-YOLO 通过自适应置信度阈值调整和模块的融合, 显著提高了对小目标的检测精度, 减少了漏检的发生, 尤其在目标密集的背景下, 其性能尤为突出。

### 3 结论

(1) PSC-YOLO 在小目标检测方面相比 YOLO11n 有明显的精度提升。该方法在 NWPU VHR-10 数据集上的 mAP50 和 mAP50-95 指标上, 均优于原始网络及单模块改进模型。通过引入 PMKCA、SPPCA、C3k2\_HLCA 以及自适应置信度阈值调整函数等创新模块, 显著增强了模型在复杂背景下小目标检测的能力。这些改进提升了模型的鲁棒性和稳定性, 为遥感图像中的高精度小目标检测提供了有效的解决方案。

(2) 通过实验验证, PSC-YOLO 在小目标检测中的显著提升得到了充分证明, 实验结果不仅体现在指标上, 还通过图像检测效果得到了体现。改进后的 mAP 得分以及可视化结果进一步验证了该模型在复杂背景中进行小目标检测的高精度和稳定性。实验表明, PSC-YOLO 能够有效解决遥感图像中小目标检测的问题, 尤其是在复杂环境下。

(3) 尽管 PSC-YOLO 在小目标检测上取得了显著提升, 但仍面临一些多尺度图像检测挑战。未来的研究可以进一步提升模型的泛化能力和鲁棒性, 探索更加高效的目标检测框架, 并结合更多上下文信息来优化目标检测。此外, 随着实时应用需求的增加, 未来的研究还应关注在保证高精度的同时实现模型的轻量化和高效推理。

### 参考文献:

[1] Li R, Zheng S, Duan C, et al. Land cover classification from remote sensing images based on multi-scale fully convolutional network [J]. *Geo-spatial Information Science*, 2022, 25 (2): 278-294.

[2] 韩冰冰, 陈圣波. 基于 GEE 时间序列遥感影像分类方法研究 [J]. *世界地质*, 2020, 39 (3): 706-713.

Han B B, Chen S B. Time series remote sensing image classification method based on GEE [J]. *Global Geology*, 2020, 39 (3): 706-713.

[3] 郑煜, 陈圣波, 陈彦冰, 等. 基于 Sentinel-1A 雷达数据和 Sentinel-2A 多光谱数据特征融合的地物分类 [J]. *世界地质*, 2021, 40 (2): 438-444.

Zheng Y, Chen S B, Chen Y B, et al. Ground object classification based on Sentinel-1A radar data and Sentinel-2A multispectral data [J]. *Global Geology*, 2021, 40 (2): 438-444.

[4] Zhang Z, Lu X, Cao G, et al. ViT-YOLO: transformer-based YOLO for object detection [C] // *Proceedings of the IEEE/CVF international conference on computer vision*. 2021: 2799-2808.

[5] Cao X, Su Y, Geng X, et al. YOLO-SF: YOLO for fire segmentation detection [J]. *Ieee Access*, 2023, 11: 111079-111092.

[6] Huang R, Pedoeem J, Chen C. YOLO-LITE: a real-time object detection algorithm optimized for non-GPU computers [C] // *2018 IEEE international conference on big data (big data)*. IEEE, 2018: 2503-2510.

[7] Dai Y, Liu W, Wang H, et al. Yolo-former: marrying yolo and transformer for foreign object detection [J]. *IEEE Transactions on Instrumentation and Measurement*, 2022, 71: 1-14.

[8] Liu W, Ren G, Yu R, et al. Image-adaptive YOLO for object detection in adverse weather conditions [C] // *Proceedings of the AAAI conference on artificial intelligence*. 2022, 36 (2): 1792-1800.

[9] 韩慧妍, 张秀权, 况立群, 等. 基于 YOLOv8L 遥感图像旋转目标检测 [J]. *激光与红外*, 2024, 54 (9): 1462-1468.

Han H Y, Zhang X Q, Kuang L Q, et al. Rotating object detection of remote sensing image based on YOLOv8L [J]. *Laser & Infrared*, 2024, 54 (9): 1462-1468.

[10] 李雪峰, 李宁, 吴迪, 等. 基于改进 Yolov8 的红外弱小目标识别算法 [J]. *激光与红外*, 2025, 55 (5): 789-797.

Li X F, Li N, Wu D, et al. Infrared dim small target detection method based on Yolov8 [J]. *Laser & Infrared*, 2025, 55 (5): 789-797.

[11] Cheng L, Li J, Duan P, et al. A small attentional YOLO model for landslide detection from satellite remote sensing images [J]. *Landslides*, 2021, 18 (8): 2751-2765.

[12] Fan K, Li Q, Li Q, et al. YOLO-remote: an object de-

- tection algorithm for remote sensing targets [J]. IEEE Access, 2024, 12: 155654–155665.
- [13] Liu F, Hu W, Hu H. YOLO-RMS: a lightweight and efficient detector for object detection in remote sensing [J]. IEEE Geoscience and Remote Sensing Letters, 2024, 21: 1–5.
- [14] Wang L, Cai J, Wang T, et al. detection of pine wilt disease using AAV remote sensing with an improved YOLO model [J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2024, 17: 19230–19242.
- [15] Ao D, Zhao W. RS-YOLO: a YOLO-Based method for small object detection in remote sensing [C] //2024 2nd International Conference on Machine Vision, Image Processing & Imaging Technology (MVIPIT). IEEE, 2024: 61–65.