

改进 YOLO11 的卫星遥感图像目标检测^①



潘童¹, 杜景林²

¹(南京信息工程大学 计算机学院, 南京 210044)

²(南京信息工程大学 人工智能学院, 南京 210044)

通信作者: 杜景林, E-mail: jldu@nuist.edu.cn

摘要: 现阶段卫星遥感图像尺寸较大, 检测目标大多较小且分布不均, 存在大量目标聚集在一起的现象, 且不同目标尺度差距较大, 背景较为复杂, 使得在土地利用和环境灾害检测方面面临相当大的挑战. 因此, 本文提出一种改进 YOLO11 的卫星遥感图像目标检测方法. 首先, 在 YOLO11 中的 C3k2 模块中引入注意力机制, 设计了 C3k2_DAB 模块, 在控制模型复杂度的同时提高模型在复杂背景影响下的检测性能. 其次, 在颈部网络后加入 PKI 模块, 促进了局部和全局上下文信息的自适应特征提取. 最后在检测端引入新的检测头 PConv 检测头, 在减少冗余计算和内存访问的前提下更快地提取空间特征. 实验结果表明, 改进的 YOLO11 网络模型在遥感图像目标检测任务中取得了优异性能, 相较于原 YOLO11 模型 mAP@0.5 提高了 2.4%, mAP@0.5:0.95 提高了 2.1%, 且优于其他主流目标检测模型, 为遥感目标检测算法的应用提供了新思路.

关键词: 卫星遥感; YOLO11; 目标检测; 卷积操作; 注意力机制

引用格式: 潘童, 杜景林. 改进 YOLO11 的卫星遥感图像目标检测. 计算机系统应用, 2026, 35(1): 209-218. <http://www.c-s-a.org.cn/1003-3254/10042.html>

Improved YOLO11 for Satellite Remote Sensing Image Object Detection

PAN Tong¹, DU Jing-Lin²

¹(School of Computer Science, Nanjing University of Information Science & Technology, Nanjing 210044, China)

²(School of Artificial Intelligence, Nanjing University of Information Science & Technology, Nanjing 210044, China)

Abstract: Currently, satellite remote sensing images feature a large size, and the targets to be detected are mostly small and unevenly distributed, with many targets gathering together. There are also significant scale differences among targets and the background is rather complex. All these factors pose great challenges to land utilization and environmental disaster detection. Therefore, this study proposes a satellite remote sensing image object detection method based on improved YOLO11. Firstly, this study introduces an attention mechanism into the C3k2 module of YOLO11 and designs the C3k2_DAB module. This enhances the detection performance under the influence of complex backgrounds while controlling model complexity. Secondly, a PKI module is added to the neck network to boost adaptive feature extraction of local and global context information. Finally, a new detection head PConv is introduced at the detection end to extract spatial features more swiftly under the prerequisite of reducing redundant computations and memory access. Experimental results demonstrate that the improved YOLO11 network model yields excellent performance in remote sensing image target detection. Compared to the original YOLO11 model, the proposed model's mAP@0.5 increases by 2.4% and mAP@0.5:0.95 improves by 2.1%. Additionally, this model outperforms other mainstream target detection models, thus providing new insights for the application of remote sensing target detection algorithms.

Key words: satellite remote sensing; YOLO11; object detection; convolutional operation; attention mechanism

① 收稿时间: 2025-06-05; 修改时间: 2025-06-30; 采用时间: 2025-08-01; csa 在线出版时间: 2025-11-17

CNKI 网络首发时间: 2025-11-18

我国土地利用类型主要包括耕地、林地、草地、水域、居民地和未利用土地等,在经济社会发展和城镇化进程加快的今天^[1],土地利用资源越来越紧缺,特别是城市的建设和耕地的保护之间,土地供需矛盾依然突出,而城镇和农村建设用地中存在大量低效和闲置土地,此土地上有大量难以识别的目标汇聚,这就需要卫星航拍来识别其中目标,通过遥感^[2]盘活利用此类土地.因此,在土地利用方面,卫星遥感目标检测的技术在该领域显得尤为重要^[3].

在遥感目标检测领域,传统的目标检测是基于手工特征的方法,如 Haar 特征^[4]、HOG 特征^[5]、SIFT 特征^[6]等,这些方法通过提取目标的边缘、纹理、形状等特征进行目标检测,此类方法是预先设计好的,无法根据具体数据或任务进行动态调整,从而造成效率低下,而遥感图像中的目标可能会因光照、角度、遮挡等因素发生变化,而手工特征的方法无法自动适应这些变化,导致检测性能较低,对目标变化不敏感等问题.

目前目标检测以深度学习中的两大类算法为主:双阶段目标检测算法和一阶段目标检测算法.双阶段目标检测以 Faster R-CNN^[7]为主的 R-CNN 系列算法为代表.此类算法将目标检测任务分解为两个阶段:区域提议和区域分类.这就有利于识别的精度更高,但是同时检测速度相对较慢,模型占有内存大,计算量也大,这在遥感检测方面并不是很契合.另一种是单阶段检测,以 SSD^[8]、RetinaNet^[9]和 YOLO^[10-12]系列为主.这类方法直接预测对象的边界框和类概率,无需单独的区域提议阶段.因此结构相对简单,检测速度相对较快,在遥感图像检测领域成为一个热点.但主流的 YOLO 系列对于遥感数据小目标类的检测也存在部分问题,因此研究人员开始在其基础上进行改进.绳鹏飞等人^[13]以 YOLOv5 为基础提出了 L-YOLOv5 网络模型,用残差空洞卷积对其主干网络进行模型压缩,减少模型的参数量与运算量,但是对于精度的提升却相对较少,存在漏检等精度与效率的平衡问题.崔家礼等人^[14]基于 YOLOv7 提出了一种改进的 RYOLOv7,通过引入选择性大核卷积增强网络对目标形状、类别、尺度等特征信息的感知能力,提高网络模型的精度,但与此同时,精度的提升伴随了计算量的大幅增加,且引入了许多背景噪声,所以成本也随之提高.鲁晓波等人^[15]提出基于改进 RRPN 模型的遥感图像目标检测方法,在特征提取网络中添加了 CBAM 注意力机制,提升了模型在

遥感图像目标特征提取方面的跨通道和空间处理能力,但是由于遥感图像的大像素问题,导致图像中出现多尺度且部分聚集的目标,因此对于小目标的检测仍然存在很多问题,且未曾调优,导致其模型泛化能力不足.吴温怡等人^[16]提出了基于多路稀疏上下文聚合与特征重筛选的方法,通过串联卷积和稀疏自注意力获取信息,通道拼合并残差聚合获取输出特征,利用注意力获取不同维度信息重要性指标,按指标针对不同维度通过特征重筛选提升提取效率.该方法有效提升了遥感目标检测精度,但是在实时检测场景中,遥感检测需要机载运行,由于机载的限制,对参数量与计算有一定的要求,因此针对此情况还存在缺陷.白晨帅等人^[17]提出超融合残差行进几何感知的算法,通过超融合残差行进模块优化网络结构,计算检测与真实结果的几何相似度来精准评估检测效果,最后考量契合度从而筛选最终结果,其提升了检测准确性与稳定性,但该方法对局部特征处理会产生一些偏差,还存在部分漏检问题,对局部目标的检测准确率有所下降,因而还有进一步提升空间.

综上所述,对于现代卫星遥感图像目标检测的领域技术仍然不是很成熟,在算法的计算量大小、计算复杂程度以及精度的方面需要平衡,且由于遥感图像的检测目标尺度变化大,整体图像像素较大的同时小目标又占据非常小的像素,再加上遥感图像中大多背景复杂,使得检测小目标异常困难.因此本文旨在在不引入过多背景噪声的前提下,提升多尺度遥感图像小目标的检测精度,同时适当平衡计算复杂性与精度,避免计算量的大幅度提升,从而提出一种基于改进 YOLO11 的卫星遥感图像目标检测方法,并在公开的 DOTA 数据集上验证模型改进后的性能.为了简化描述,后文所有遥感图像均指卫星遥感图像.本文主要工作如下.

1) 对原有 C3k2 模块进行了改进,加入了双注意力块 (DAB) 机制,从而提出了一种改进的 C3k2_DAB^[18] 模块,并嵌入到原主干网络中.针对计算量与精度的平衡问题,其使得模型在控制了参数量的同时,减少复杂背景对于精度的影响,增强模型对物体位置信息和通道语义信息的感知能力,提高目标检测的准确性和鲁棒性.

2) 由于遥感图像目标检测 (RSIs) 面临目标尺度的巨大变化,而通过大核卷积扩展主干的空间感受野的操作容易引起背景噪声的问题,本文加入了一种改进的 PKI 模块^[19],该模块由无扩展的多尺度卷积核来提

取不同尺度的目标特征并捕获局部上下文信息,同时通过上下文锚定注意力(CAA)模块来捕获远程上下文信息,多尺度卷积核和CAA模块并行组成PKI模块。

3) 针对遥感图像中部分聚集的小像素目标,本文在检测端针对原检测头进行了改进,引入了部分卷积PConv (partial convolution)^[20]的思想,通过对部分通道操作,增强局部特征,从而控制参数量的同时减少小目标检测准确率低、漏检、误检等问题频发。

1 YOLO11n 的目标检测算法

YOLO11是由Ultralytics公司开发的新一代目标检测算法,它包含n、s、m、l、x这5个版本,这5个版本的规模随顺序逐渐增大,精度与计算量也随之增加。为了适应遥感目标检测的需要,以及平衡效率与计算量之间的关系,本文选择了YOLO11n的版本为基准

模型。网络主要由4个部分组成:输入端、主干网络(backbone)、特征增强网络(neck)和检测头(head)。网络架构如图1所示。

1) 输入端:输入端设计灵活,能够适应不同的输入尺寸和数据格式,将其转换为特征图,这使得它能够广泛应用于各种目标检测任务。

2) Backbone^[21]:主干网络是YOLO11n核心部分,负责从输入图像中提取多尺度特征,由C3k2、卷积层、SPPF、C2PSA组成。C3k2模块是YOLO11n中的关键模块,替代了YOLOv8中的C2f模块。它通过分割特征图并应用一系列较小的内核卷积(3×3)来优化网络中的信息流。卷积层用于对图像进行下采样,提取基础特征。SPPF增强特征提取能力,适应多尺度目标检测。C2PSA通过空间注意力机制,进一步提升特征提取的精度。

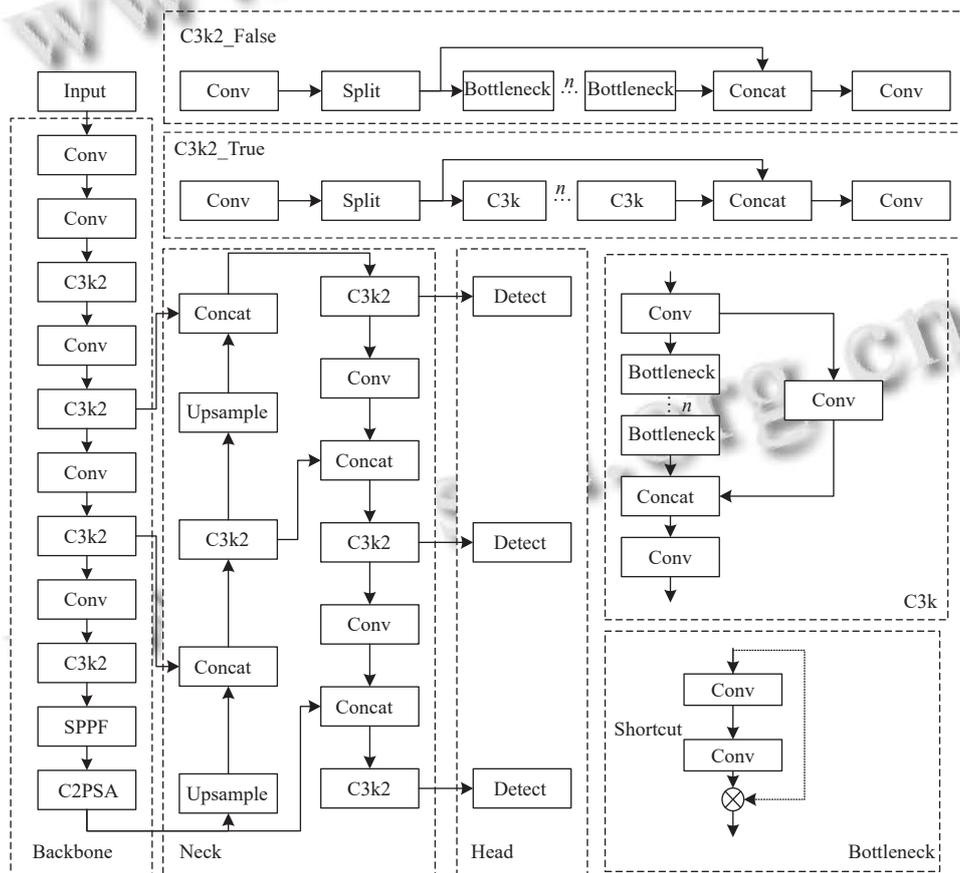


图1 YOLO11网络结构

3) Neck^[22]: Neck部分负责特征融合和增强,由FPN和PAN结构组成,通过自下而上的方式,通过上采样和下采样的方式,将主干网络传递来的浅层特征

和深层语义特征进行融合,增加了多尺度的信息检测,也保证了全局空间信息的完整性。

4) Head: Head部分负责生成最终预测的部分,主

要预测分类以及目标的坐标等信息,由3个检测头构成,分别是80×80(检测小目标)、40×40(检测中等大小的目标)、20×20(检测大目标)。

目前,YOLO11在遥感领域中广泛被应用,但是仍然存在许多的不足,首先是遥感硬件设备条件的限制,需平衡计算量与精度,其次对于小物体检测性能欠佳,对于较小的目标,YOLO11n的表现可能不如期望,由于网络结构的设计特点以及特征图分辨率的影响,小尺寸对象容易被忽略或误分类,因此,本文对此进行了改进。

2 改进 YOLO11 的遥感目标检测方法

本文基于YOLO11进行遥感目标检测改进,主要改进3个模块:首先,在YOLO11中的C3k2模块中引入注意力的思想,设计成为C3k2_DAB模块,其在控制模型复杂度的同时可以降低背景对于遥感图像的影响,更加精确地识别图像中的目标。其次,加入PKI模块,促进了局部和全局上下文信息的自适应特征提取。最后新的检测头PConv检测头,在减少冗余计算和内存访问的前提下更快地提取空间特征。其改进的网络结构图如图2所示。

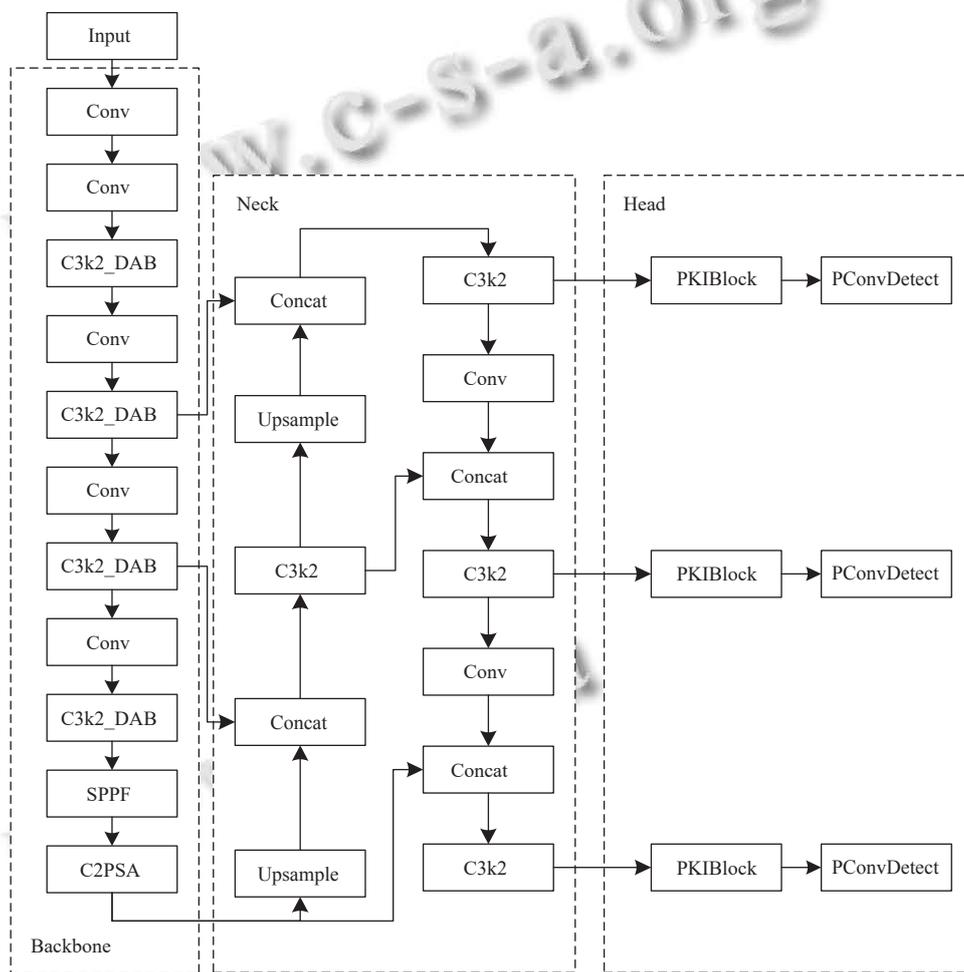


图2 改进后YOLO11网络结构

2.1 基于DAB注意力机制的C3k2_DAB模块

在原有的基准模型中,由于需要平衡计算量与检测精度的问题,再加上遥感图像目标检测一般都会受到复杂背景的影响,导致图像中的小目标并不清楚,从而难以识别。因此受到DAB注意力的启发,将其嵌入

C3k2模块中,这样做可以增加模型在复杂背景影响下对物体的检测性能,增强对物体位置信息和通道语义信息的感知能力,提升检测效果,同时也控制参数量的大幅度增长。

以往的注意力机制一般通过Q、K、V的点积操

作来捕获全局依赖关系,特别是在处理大尺寸特征图的时候,需要大量的点积操作来处理,而遥感图像恰好大部分的尺寸都比较大,因此计算量随之增加,同时一般注意力都聚焦于显著特征的提取,而忽略一些小目标等不显著的特征,因此综合来看许多注意力并不适用于此领域.而本文引入 DAB 注意力模块去替换原本 C3k2 中的 Bottleneck 模块,该模块将通道-空间注意力模块 (CSAM) 和并行注意力模块 (PAM) 结合串联起来,通过这两种注意力机制的协作,使网络能够更好地捕捉特征图中的小目标重要信息. DAB 的结构图如图 3 所示.



图 3 DAB 注意力机制结构

图 3 中 CSAM 块由批处理归一化、信道空间注意力 (CSA) 和多层感知器组成,具体结构图如图 4 所示. CSA 由通道和空间两个分支组成:一个通道分支使用全局平均池化 (global average pooling, GAP) 来获取每个通道的全局信息,然后利用两个线性层生成通道信息.另一条空间分支通过尺寸为 7×7 的大卷积核生成的广泛空间信息. CSA 公式如下:

$$X_{CSA} = Sigmoid(GELU(Linear(GAP(\hat{x})))) \otimes \hat{x} + Sigmoid(Conv_{7 \times 7}(\hat{x})) \otimes \hat{x} \quad (1)$$

PAM 块由批处理归一化、深度卷积、并行注意力和多层感知器组成, PAM 的结构图如图 5 所示.

而 PA 平行连接了 3 种不同的注意力机制:全局通道注意力 (GCA)、局部通道注意力 (LCA) 和空间注意力 (SA). GCA 使用 GAP 获取每个通道内部的全局压缩信息,然后通过两个 1×1 卷积层分别与 GELU 和 Sigmoid 函数生成特征图,提取跨通道的全局通道信息. LCA 通过将跨本地渠道的交互信息与本地渠道交互,本地渠道分支分别生成的每个渠道的本地信息相结合来保存渠道信息. SA 利用卷积产生的宽空间信息来生

成特征图.这 3 个注意力公式如下:

$$X_{GCA} = Sigmoid(Conv_{1 \times 1}(GELU(Conv_{1 \times 1}(GAP(\hat{x})))))) \quad (2)$$

$$X_{LCA} = (Sigmoid(1DConv(GAP(\hat{x}))) \times \hat{x} + \hat{x}) + Sigmoid(DWConv_{3 \times 3}(\hat{x})) \times \hat{x} \quad (3)$$

$$X_{SA} = Sigmoid(Conv_{7 \times 7}(\hat{x})) \times \hat{x} \quad (4)$$

最后将通道维度上的这 3 个特征映射通过 Concat 连接,从而保留图像细节,通过多层感知器从图像中捕获关键特征.

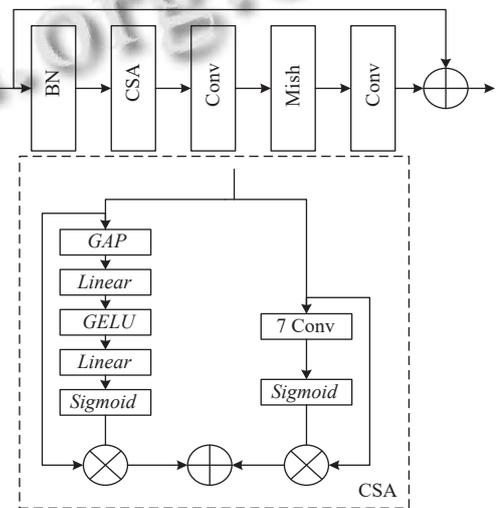


图 4 CSAM 的结构图

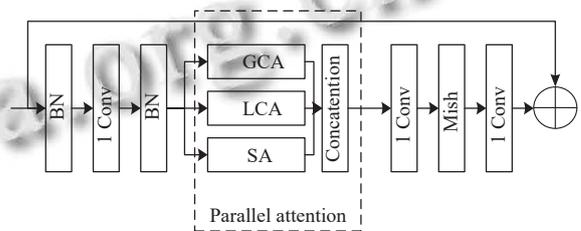


图 5 PAM 的结构图

2.2 PKI 模块

与一般的目标检测不同,遥感目标检测旨在定位和识别单幅图像内不同大小的目标,而遥感图像多为大像素、高分辨率的图像,这就造成图像中目标含量较多,同时同一类型的目标其大小形状也不同,例如车辆目标中既有小型轿车,也有大型卡车;建筑物目标中既有低矮的平房,也有摩天大楼,因此遥感图像中目标尺度变化较大成为目标检测的一个难题.为了解决这一挑战,传统的方法一般是用大核卷积去对特征图进行操作,这种做法有一定的效果,但是通常会引入相

当一部分的背景噪声,且计算成本高,容易忽视局部特征,因此其具有一定的局限性.为了进一步解决该问题,受到了 PKINet 的启发,本文将 PKI 模块引入 YOLO11 中,该模块采用无扩展的多尺度卷积核来提取不同尺度的目标特征并捕获局部上下文,但同时由于又要兼顾全局远程上下文,因此同时并行了上下文锚定注意力 (CAA) 模块来捕获远程上下文信息. PKI 模块就是由这两部分组成,其能够充分捕捉多尺度的信息,较为适用于遥感图像,减少了上下文特征的丢失,提高了遥感目标检测的性能. PKI 模块总体结构如图 6 所示.

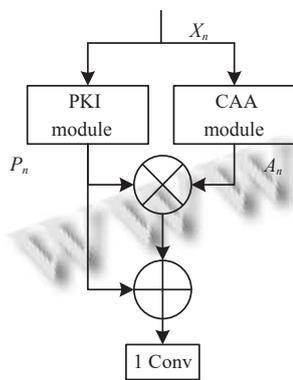


图 6 PKI 总体结构

PKI 模块是由 PKI module 和 CAA module 并行组成. 输入为 X_n , 其输入的形式是 $C \times H \times W$, C 是通道数, H 、 W 是特征图的高、宽, 分别经过两个并行的块 PKI 和 CAA, PKI 的输出是 P_n , CAA 的输出是 A_n . PKI 块如图 7 所示. 其先是一个小核卷积获取信息, 然后通过一组并行深度卷积来捕捉多尺度的上下文信息. 输入为 X_n , 先通过一个 3×3 的卷积输出 L_n , 然后再通过一组并行卷积得到输出 Z_n , 将 L_n 与 Z_n 相加后再用一个 1×1 的卷积得到输出 P_n .

另一个 CAA 块如图 8 所示. 其先通过一个平均池化层后经 1×1 的卷积获得局部特征, 由于条形卷积更适用于遥感图像中比如桥梁这种细长类的目标, 且其相比大核卷积更加轻量, 所以这里通过两个分别为 1×11 和 11×1 的条形卷积处理, 最后再经过一个 1×1 的卷积后激活输出结果 A_n . 将它与 PKI 块输出结果 P_n 点乘后再加上 P_n 残差连接, 最后通过 1×1 卷积输出结果.

根据上述分析, PKI 模块相较于传统的扩展感受野做法, 利用无扩展的初始式深度卷积来提取不同感受野之间的多尺度纹理特征, 在遥感目标检测领域更加适用, 既能够有效提取不同尺度的目标特征, 又能捕

获长距离上下文信息, 因而兼顾了局部和全局的上下文信息, 提升了多尺度目标的特征提取能力.

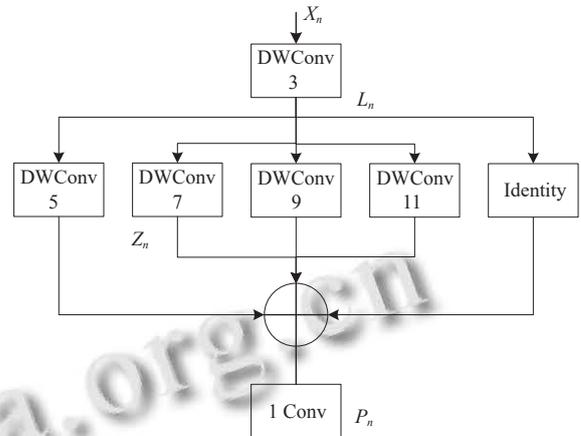


图 7 PKI module

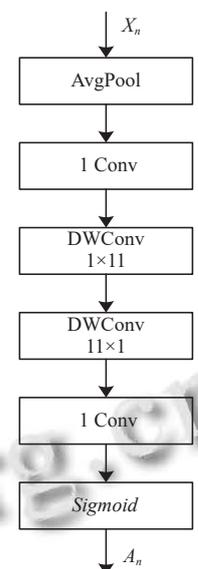


图 8 CAA module

2.3 检测头改进

YOLO11n 的检测头采用了解耦头设计, 通过定位检测头和分类检测头, 将目标位置和类别信息分别提取出来, 通过不同的网络分支分别学习, 最后再进行融合. 然而原本模型对于微小的目标的检测性能较差, 在网络的不断卷积之后, 会丢失一部分小目标的特征, 而小目标检测一直以来都是一个难点之一, 因此模型仍需考虑对小目标的检测性能, 尤其是对于高分辨率的遥感图像.

针对此问题, 在改进的遥感图像目标检测算法中需要对原本的检测头进行改进以便于更好地检测小目标, 并且需要考虑到控制计算量的问题. 本文引入了部分卷

积 (partial convolution, PConv) 的思想, 在 head 检测头中的定位和类别检测头之前在每一个通道上将输入特征图的通道分割为两部分, 对其中一部分通道进行卷积操作, 其余部分保持不变, 最后将卷积后的部分和未被卷积的部分重新组合. PConv 模块结构图如图 9 所示.

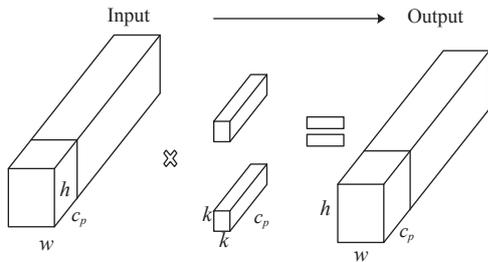


图 9 PConv 原理结构图

普通卷积的计算量和内存访问量如式 (5) 和式 (6) 所示:

$$FLOPs = h \times w \times k^2 \times c^2 \quad (5)$$

$$h \times w \times 2c + k^2 \times c^2 \approx h \times w \times 2c \quad (6)$$

而 PConv 的计算量和内存访问量如式 (7) 和式 (8) 所示:

$$FLOPs = h \times w \times k^2 \times c_p^2 \quad (7)$$

$$h \times w \times 2c_p + k^2 \times c_p^2 \approx h \times w \times 2c_p \quad (8)$$

其中, h 、 w 为输入特征图的高度和宽度, c 为输入的通道数, c_p 为部分输入的通道数, k 为卷积核的大小. 本文使用常用的 $c:c_p$ 为 1:4 的比例. 而 PConv 的 $FLOPs$ 仅为常规卷积的 1/16, 显著降低了所需的计算资源. 并且 PConv 的内存访问次数仅为常规卷积的 1/4. 这样既能通过减少浮点运算次数从而减少不必要的计算和内存访问, 提高模型的运行效率, 同时也能更加针对小目标的特征细节信息, 减少对于小目标特征信息的丢失, 提高小目标检测的效率和性能.

3 实验结果处理与分析

3.1 数据集介绍与预处理

本文选用由武汉大学的研究人员创建的 DOTA1.0 数据集, 该数据集是一个专门用于目标检测的遥感图像数据集, 主要用于检测遥感图像中的各种目标, 图像来自不同传感器和平台, 包括 Google Earth、JL-1 卫星等, 旨在解决遥感图像中目标检测的挑战性问题. 其收

录了 2806 张 800–4000 像素的图像, 总共包含 188282 个目标, 分为 15 个类别. 其中训练集 1411 张, 验证集 458 张, 测试集 937 张. 该数据集包括: 飞机、船只、储罐、棒球场、网球场、篮球场、地面田径场、港口、桥梁、小型车辆、大型车辆、直升机、环岛、足球场和游泳池目标的检测.

由于该数据集不符合 YOLO 标签形式, 先将其转化为 YOLO 检测模型适用的形式, 同时由于该数据集像素大, 分辨率高且不统一, 将数据集图像统一转成 1024 像素的图像, 并且数据集标签采用水平框标注^[23]. 数据集处理流程如图 10 所示.

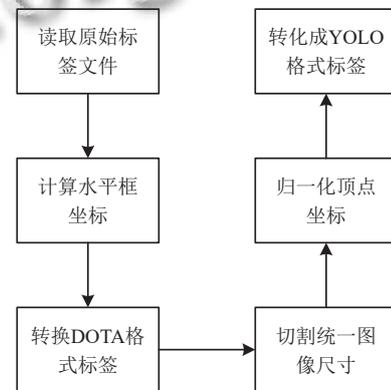


图 10 数据集处理流程

3.2 评价指标

为了更加直观地展示改进后的网络算法的有效性, 本文采用精度 (P)、召回率 (R)、平均精度 (AP) 以及平均精度均值 (mAP) 来评估模型的算法性能. 同时为了考虑计算量与检测复杂度, 采用计算量 ($GFLOPs$)、检测帧率 (FPS) 作为模型复杂度的评估指标. 其中以 $mAP@0.5$ 和 $mAP@0.5:0.95$ 作为主要评估指标, 其中 $mAP@0.5$ 是 IoU 阈值为 0.5 时的 mAP . 具体计算公式如式 (9):

$$\left\{ \begin{array}{l} FPS = \frac{1000}{Pre + Infer + Post} \\ P = \frac{TP}{TP + FP} \\ R = \frac{TP}{TP + FN} \\ F_{AP} = \int_0^1 p(R) dR \\ F_{mAP} = \frac{1}{n} \sum_{k=1}^n F_{AP_k} \end{array} \right. \quad (9)$$

其中, TP 表示正样本被预测为正样本的个数, FP 表示负样本被预测为正样本的个数, FN 表示正样本被预测为负样本的个数, n 为目标检测的总类别数, $p(R)$ 为以召回率为 x 轴和准确率为 y 轴绘制出的曲线, 又称 P-R 曲线, 其与坐标轴围成图形的面积大小即为平均准确率. FPS 为算法在 1 s 内能够检测的图片帧数 (包括前处理时间 Pre 、推理时间 $Infer$ 、后处理时间 $Post$), FPS 衡量算法检测速度, 其值越大模型检测速度越快, 实时性越好, 而 GFLOPs 表示每秒十亿次浮点运算, 表示算法参数等的总体计算量.

3.3 实验环境和参数配置

实验环境为 Windows 11 系统, GPU 为 NVIDIA GeForce RTX3090Ti 显卡. 平台结合 Anaconda 与 Py-Charm 作为代码运行环境. 环境配置包括 CUDA 12.2、Python 3.8、PyTorch 2.0.1 等. 参数配置方面, 初始学习率为 0.01, 动量为 0.937, 权重衰减系数为 0.0005, 批量大小为 8, 使用 SGD 优化器迭代训练 300 轮, 确保模型训练充分.

3.4 注意力消融对比实验

在 C3k2_DAB 模块中, 本文引入了双注意力机制 (DAB), 该注意力机制通过串联通道-空间注意力模块 (CSAM) 和并行注意力模块 (PAM) 组合而成, 为进一步验证 CSAM 与 PAM 组合方式以及各自注意力对模型性能以及计算量的影响, 包括 CSAM 与 PAM 的串联顺序对模型性能与计算量的影响, 在统一的实验环境下进行注意力消融实验, 同时将一个主流的注意力机制 LSKA 嵌入 C3k2 模块中与其进行对比. 表 1 展示了注意力的消融对比结果.

表 1 注意力消融对比实验结果

Attention	mAP@0.5 (%)	mAP@0.5:0.95 (%)	GFLOPs
Baseline	71.6	48.6	6.3
LSKA	71.8	47.9	7.2
CSAM	71.7	48.0	6.3
PAM	72.3	48.6	6.7
PAM→CSAM	72.5	48.9	7.1
CSAM→PAM	72.5	49.0	7.1

由表 1 可知, 在注意力消融实验中, PAM 注意力与 CSAM 注意力对于基线模型总体上来看有部分的性能提升, 且计算量的提升较小. 而串联 PAM 注意力与 CSAM 注意力的情况下, PAM 与 CSAM 注意力的串联顺序对模型性能与计算量的影响并不大, 相比基线模型 mAP@0.5 提高了 0.9%, mAP@0.5:0.95 提高了

0.4%, 计算量提升了 0.8×10^9 , 而相比主流的 LSKA 注意力机制, 其计算量相对较低的情况下检测性能更高, 由此可见 DAB 注意力机制在计算量没有明显增加的同时提升了模型的性能, 且此注意力设计最优.

3.5 消融实验

为探究本文改进的 YOLO11n 模型中各个模块对遥感目标检测性能的影响, 在统一的环境下进行消融实验. 表 2 展示了各个消融实验的结果, 其中“√”表示使用了该模块, ID 编号的具体含义如下.

表 2 消融实验结果

ID	C3k2_DAB	PKI	PConv	P (%)	R (%)	mAP@0.5 (%)	mAP@0.5:0.95 (%)	Parameter (MB)
①	—	—	—	76.8	67.8	71.6	48.6	9.86
②	√	—	—	76.9	68.7	72.5	49.0	10.44
③	—	√	—	79.1	67.7	72.8	49.4	14.77
④	—	—	√	77.6	69.3	73.1	49.6	10.38
⑤	√	√	—	79.2	68.6	73.2	50.1	15.35
⑥	√	—	√	78.3	68.2	72.6	49.1	10.96
⑦	—	√	√	77.8	68.5	73.8	50.6	15.28
⑧	√	√	√	79.3	69.4	74.0	50.7	15.86

① 指的是原 YOLO11n 模型的实验结果, 作为后续实验的对比基准, 为了对比出各个改进模块的性能提升效果.

② 为主干网络中的 C3k2 加入 DAB 注意力机制形成的 C3k2_DAB 模块, 旨在控制参数量的同时提升模型在复杂背景干扰下检测性能. 其参数量增加了 0.58 MB, 而 mAP@0.5 提高了 0.9%, mAP@0.5:0.95 提高了 0.4%.

③ 为在多尺度融合颈部网络后加入 PKI 模块, 促进局部和全局上下文信息的自适应特征提取. 因此, mAP@0.5 提高了 1.2%, mAP@0.5:0.95 提高了 0.8%.

④ 为改进检测头后的网络模型的实验结果, 旨在检测小目标的性能提升并且不大量提升参数量. 根据实验结果, mAP@0.5 提高了 1.5%, mAP@0.5:0.95 提高了 1%, 参数量增加了 0.52 MB.

⑤ 为同时加入 C3k2_DAB 模块和 PKI 模块的网络模型, 与实验②相比, 多引入 PKI 模块之后, mAP@0.5 提高了 0.7%, mAP@0.5:0.95 提高了 1.1%.

⑥ 为加入 C3k2_DAB 模块且改进检测头之后的模型, 与实验②相比, 多引入改进检测头之后, mAP@0.5 提高了 0.1%, mAP@0.5:0.95 提高了 0.1%.

⑦ 为加入 PKI 模块且改进检测头之后的模型, 与

实验③相比, $mAP@0.5$ 提高了 1%, $mAP@0.5:0.95$ 提高了 1.2%。

⑧ 为所有改进后的 YOLO11n 模型, 与基准模型相比, $mAP@0.5$ 提高了 2.4%, $mAP@0.5:0.95$ 提高了 2.1%。

综上, 消融实验验证了本文的各个改进点及其组合的有效性, 尤其在多尺度的小目标中检测突出, 证明了本文的方法在遥感目标检测领域的实用性。

3.6 对比实验

为了进一步体现出改进后的 YOLO11 算法的优势, 验证在遥感目标检测领域的效果, 除了 YOLO11n 以外, 选用了一些经典算法作为比较对象, 在 DOTA1.0 数据集上进行验证, 整理成表 3 以供分析。

表 3 不同算法在 DOTA1.0 数据集上的对比结果

模型	$mAP@0.5$ (%)	$mAP@0.5:0.95$ (%)	GFLOPs
Faster R-CNN	50.0	35.4	206
SSD	63.4	43.3	115
YOLOv5	69.8	46.8	7.1
YOLOv6	68.0	46.0	11.8
YOLOv8	68.2	46.2	8.9
YOLOv10	69.2	47.3	8.4
YOLO11	71.6	48.6	6.3
Ours	74.0	50.7	9.9

深入分析表 3 数据, 改进的方法在性能上相较于 YOLO11 基准模型提升明显, $mAP@0.5$ 提高了 2.4%, $mAP@0.5:0.95$ 提高了 2.1%, 且相比于其他主流模型, $mAP@0.5$ 和 $mAP@0.5:0.95$ 分别提升大约 5%–24% 和 3%–15%, 提升效果显著, 尽管 GFLOPs 相比原模型略微有点提升, 但在可接受范围之内, 且也低于其他大部分主流模型, 而在实际应用之中, 精度的提升更为重要。综合来说, 本文模型在精度指标上优于其他主流模型, 对遥感图像目标检测更为精确, 更适用于遥感目标检测领域。

3.7 可视化分析

为更加直观展示改进后的算法的检测性能, 本文在 DOTA1.0 数据集上选取密集、多尺度目标和复杂背景等情况下具有代表性的图像进行分析, 图 11 展示原模型与改进后的模型的对比结果。

根据图中所展示的效果来看, 相比于 YOLO11n 原模型, 对于第 1 列图片存在的大量密集的小目标, 检测结果更加细腻, 检测出了更多的密集目标, 且不同尺度的目标也被检测了出来。第 2 列图片中, 在较为复杂的

背景情况下, 也减少了漏检率, 漏检的小目标更容易被检测出来。在第 3 列图片中存在极其微小且肉眼难以可见的小目标, 对于该类型的目标, 改进后的模型能够更加精确地检测出来。

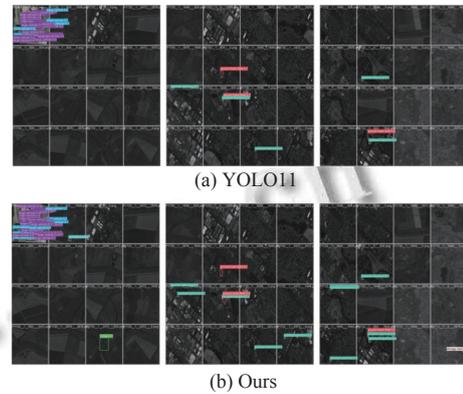


图 11 检测结果对比

由此可见, 改进后的算法能够明显提升检测性能, 减少漏检等问题, 特别是在密集、多尺度和复杂背景下的微小目标检测中表现突出, 因此本文模型能够更加适用于遥感图像目标检测领域。

4 结论与展望

针对遥感图像中目标尺度差异大、密度不均及控制参数量过多增长等问题, 提出了一种改进 YOLO11n 的遥感目标检测算法。通过在 YOLO11 中的 C3k2 模块中引入 DAB 注意力机制, 设计了 C3k2_DAB 模块, 其在控制参数的同时可以有助于提高模型在复杂背景影响下对目标的学习能力。其次, 在颈部网络后加入 PKI 模块, 促进了局部和全局上下文信息的自适应特征提取。最后在检测端引入新的检测头 PConv 检测头, 在减少冗余计算和内存访问的前提下更快速地提取空间特征。实验结果证明本文提出的算法在控制参数不剧烈增长的前提下, 在检测精度、检测性能上取得显著提升效果, 相较 YOLO11 原模型, $mAP@0.5$ 提高了 2.4%, $mAP@0.5:0.95$ 提高了 2.1%, 该方法有效解决了多尺度、密集的遥感小目标难以检测等问题, 也同时避免了参数量、计算量的大幅增长, 成功完成遥感目标检测任务。但本文提出的方法也存在一定局限性, 因此下一步的工作将针对此继续研究。

(1) 本文主要针对正常天气情况下的场景, 在受到低光照、云层遮挡等极端情况的影响下, 模型在检测

多尺度密集的遥感小目标方面的表现尚有提升空间,泛化能力有待加强.后续将会进一步引入针对云层遮挡、低光照等极端情况的目标检测技术,在对极端场景下的图像进行预处理以增强目标特征的同时对遥感图像目标检测方案进行优化,从而更好地满足实际应用中的多样化需求.

(2) 由于遥感目标检测通常是建立在资源受限的设备环境中执行,对模型计算量和参数量有一定的要求,需要更加轻量化的模型在保持高精度的同时完成检测任务.而高分辨率遥感图像包含大量的细节信息,但并非所有细节都对目标检测有用.因此后续将研究基于空间分辨率的通道剪枝策略,根据输入图像的目标大小和分布动态调整剪枝强度,在目标密集的区域保留更多通道,而在目标稀疏的区域更激进地剪枝,以实现更高的轻量化效果,从而进一步完善遥感目标检测任务.

参考文献

- 赵新天. 土地资源紧约束下房地产行业开发挑战与集约利用策略. *住宅与房地产*, 2025(6): 106–108.
- 赵昕, 李秀江, 张鑫. AI 遥感变化检测技术在土地卫片执法中的应用. *中国资源综合利用*, 2024, 42(11): 97–99.
- 王唯真. 数字遥感技术在土地利用动态监测中的应用. *中国高科技*, 2024(5): 90–92. [doi: 10.13535/j.cnki.10-1507/n.2024.05.20]
- 李景, 吴玉秀, 张捍东. 基于 Haar 特征的 Adaboost 算法的猪脸检测. *常州工学院学报*, 2023, 36(3): 14–19.
- 吴逸, 吴静静. 改进 HOG 特征的支持向量机工程车辆识别方法. *计算机与数字工程*, 2023, 51(10): 2405–2412. [doi: 10.3969/j.issn.1672-9722.2023.10.036]
- 李梓谦, 付志涛, 聂韩, 等. 基于图像块 SIFT 特征学习的多源遥感图像匹配方法. *遥感信息*, 2023, 38(1): 155–162. [doi: 10.20091/j.cnki.1000-3177.2023.01.020]
- Yang Q, Ma S, Guo DQ, *et al.* A small object detection method for oil leakage defects in substations based on improved Faster-RCNN. *Sensors*, 2023, 23(17): 7390. [doi: 10.3390/s23177390]
- Dong C, Luo XS. Research on a pedestrian detection algorithm based on improved SSD network. *Journal of Physics: Conference Series*, 2021, 1802(3): 032073. [doi: 10.1088/1742-6596/1802/3/032073]
- Lee SS, Lim LG, Palaiahnakote S, *et al.* Oil palm tree detection in UAV imagery using an enhanced RetinaNet. *Computers and Electronics in Agriculture*, 2024, 227(P1): 109530–109530.
- Nghiem VQ, Nguyen HH, Hoang MS. LEAF-YOLO: Lightweight edge-real-time small object detection on aerial imagery. *Intelligent Systems with Applications*, 2025, 25: 200484. [doi: 10.1016/j.iswa.2025.200484]
- 王兴涛, 单慧琳, 孙佳琪, 等. 基于改进 YOLOv3 的遥感目标检测算法. *兵器装备工程学报*, 2023, 44(11): 279–286.
- 张瑞芳, 刘占占, 程小辉, 等. 基于改进 YOLOv8 的无人机红外目标检测算法. *电子测量技术*, 2025, 48(7): 46–54.
- 绳鹏飞, 杨军. 采用轻量化主干网络的遥感影像目标检测. *测绘科学*, 2023, 48(5): 58–67. [doi: 10.16251/j.cnki.1009-2307.2023.05.008]
- 崔家礼, 刘远. 基于改进 YOLOv7 的遥感图像旋转目标检测. *微电子学与计算机*, 2025, 42(4): 48–57.
- 鲁晓波, 郭艳光, 辛春花. 改进 RRPN 模型的遥感图像目标检测. *现代电子技术*, 2025, 48(1): 8–16. [doi: 10.16652/j.issn.1004-373x.2025.01.002]
- 吴韞怡, 马明睿, 王靖宇, 等. 基于多路稀疏上下文聚合与特征重筛选的遥感图像目标检测方法. *第十三届中国指挥控制大会论文集(下册)*. 北京: 中国指挥与控制学会, 2025. 323–327. [doi: 10.26914/c.cnkihy.2025.007990]
- 白晨帅, 白晓凤, 邬开俊, 等. 超融合残差行进几何感知的遥感目标检测. *光学精密工程*, 2025, 33(8): 1289–1302.
- Kim N, Choi IS, Han SS, *et al.* DA-Net: Dual attention network for haze removal in remote sensing image. *IEEE Access*, 2024, 12: 136297–136312. [doi: 10.1109/ACCESS.2024.3459588]
- Cai XH, Lai QX, Wang YW, *et al.* Poly kernel inception network for remote sensing detection. *Proceedings of the 2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Seattle: IEEE, 2024. 27706–27716.
- Chen JR, Kao SH, He H, *et al.* Run, don't walk: Chasing higher FLOPS for faster neural networks. *Proceedings of the 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Vancouver: IEEE, 2023. 12021–12031.
- 朱飞雨, 刘容轲, 朱友文, 等. 面向轻量级无人机的 IB-YOLOv8 目标检测方法研究. *无线电工程*, 2025, 55(9): 1775–1784.
- Li YS, Wang JR, Zhang KH, *et al.* Lightweight object detection networks for UAV aerial images based on YOLO. *Chinese Journal of Electronics*, 2024, 33(4): 997–1009. [doi: 10.23919/cje.2022.00.300]
- 孙伟, 沈欣怡, 张小瑞, 等. 适应遥感船舶图像的轻量化旋转小目标检测网络. *电子测量与仪器学报*, 2025, 39(4): 122–131.

(校对责编: 张重毅)