

文章编号 1004-924X(2026)04-0624-16

融合背景信息和方向信息的遥感图像 旋转目标检测

胡贺南¹, 马奥阳¹, 李荣华^{2*}

- (1. 大连交通大学 机械工程学院, 辽宁 大连 116028;
2. 大连交通大学 电气工程学院, 辽宁 大连 116028)

摘要: 针对遥感图像目标检测中因目标尺度差异大、方向分布复杂导致的检测精度不足问题, 提出一种融合背景信息与方向信息的检测网络。首先, 针对目标尺度差异大的问题, 设计了感受野扩张模块(Receptive Field Extending, RFE)。与传统的固定感受野或复杂多分支的结构不同, 该模块通过大核分解、空洞卷积与并联分支结构设计, 在不显著增加计算量的前提下, 融合多尺度的背景信息, 解决了不同尺度目标对背景信息需求差异的问题。其次, 针对目标方向分布复杂的问题, 设计方向感知交互注意力机制模块(Orientation Aware Cross Attention, OACA)。与现有的注意力机制卷积核形状不同, 该模块通过水平与垂直方向可分离卷积提取方向纹理信息, 防止特征缺失与断裂; 同时设计交叉注意力机制, 抑制背景噪声并增强方向信息的交互性。实验结果表明, 提出的方法在 DOTA, HRSC2016 和 DIOR-R 数据集上检测精度分别达到 76.88%、98.43% 和 65.06%, 相比 Oriented R-CNN 方法分别提升了 1.01%、0.83% 和 0.76%, 进一步验证了背景信息和方向信息协同的有效性。

关键词: 遥感图像; 旋转目标检测; 感受野扩张; 方向感知交互注意力

中图分类号: TP751 文献标识码: A

doi: 10.37188/OPE.20263404.0624

CSTR: 32169.14.OPE.20263404.0624

Remote sensing image rotation object detection based on background information and direction information fusion

HU Henan¹, MA Aoyang¹, LI Ronghua^{2*}

- (1. Institute of Mechanical Engineering, Dalian Jiaotong University, Dalian 116028, China;
2. Institute of Electrical Engineering, Dalian Jiaotong University, Dalian 116028, China)

* Corresponding author, E-mail: lironghua705@163.com

Abstract: Aiming at the problem of insufficient detection accuracy caused by large differences in target scale and complex direction distribution in remote sensing image target detection, a detection network combining background information and direction information was proposed. Firstly, aimed at the problem of large differences in target scale, a Receptive Field Extending (RFE) module was designed. Different from the traditional fixed receptive field or complex multi-branch structure, this module integrated multi-scale background information through large kernel decomposition, dilated convolution, and parallel branch struc-

收稿日期: 2025-10-28; 修订日期: 2025-11-19.

基金项目: 国家自然科学基金(No. U24B20159); 辽宁省教育厅科学研究项目(No. LJ212410150036); 辽宁省交通科技资助项目(No. 202416); 大连市高层次人才创新支持计划(No. 2022RJ03)

ture design, and solved the problem of different background information requirements for different scale targets without significantly increasing the amount of calculation. Secondly, aimed at the problem of complex distribution of target directions, an Orientation Aware Cross Attention (OACA) module was designed. Different from the existing attention mechanism convolution kernel shape, the module extracted directional texture information through horizontal and vertical separable convolution to prevent feature loss and fracture; at the same time, a cross-attention mechanism was designed to suppress background noise and enhance the interactivity of directional information. The experimental results show that the detection accuracy of the proposed method on DOTA, HRSC2016 and DIOR-R datasets reaches 76.88%, 98.43% and 65.06%, respectively, which is 1.01%, 0.83% and 0.76% higher than that of the Oriented R-CNN method, respectively, which further verifies the effectiveness of background information and direction information collaboration.

Key words: remote sensing images; rotating object detection; receptive field extending; orientation aware cross attention

1 引言

遥感图像目标检测作为地理信息智能解译的核心任务,在灾害监测、城市规划及军事侦察等领域具有重要应用价值。与自然图像相比,遥感图像(Remote Sensing Images, RSIs)成像方式不同使得图像中的目标具有方向分布任意、空间分布密集、以及尺度差异巨大的特点。这些特点导致水平边界框(Horizontal Bounding Box, HBB)的检测模型难以紧密贴合任意方向的目标轮廓,最终会使检测过程中出现背景干扰加剧与定位精度下降的问题。

为解决上述问题,研究人员围绕定向边界框(Oriented Bounding Box, OBB)进行大量研究:在OBB检测器研发方面,Ding等人^[1]提出RoI-Transformer模块,通过旋转敏感特征与几何约束头解决了特征错位问题;Yang等人^[2]设计SmoothIoU-L1损失与密集锚框分配策略,有效解决了距离损失与IoU不匹配的问题;Huang等人^[3]将点预测与Transformer结合,实现了无需NMS的端到端旋转检测。在角度预测精度提升方面,Xu等人^[4]构建水平检测与角度推导的双阶段框架,避免复杂角度回归损失设计;Xu等人^[5]提出角度双重优化范式,解决了角度边界跳变问题;Zhang^[6]等人提出单位周期解析器(Unit-Cycle Resolver, UCR),解决角度预测的离散化误差。

虽然上述的方法推动了遥感检测技术的发

展,但大多数研究聚焦在优化边界框表示形式上,而在改进特征提取方面没有充分的探究。所以现有的方法在特征提取方面存在缺陷:首先,传统方法依赖堆叠卷积构建固定感受野,虽能覆盖中小尺度目标,却无法为大尺度目标提供充足全局信息,忽略了不同尺度的目标对背景信息需求的差异性;若单纯增大卷积核扩大感受野,又会导致小尺度目标特征被稀释,造成本义表征不完整。其次,标准的卷积核设计在面对方向分布复杂、高纵横比目标时,难以捕捉旋转目标的方向信息。同时,方向信息之间缺乏交互,使得不同方向的特征间难以建立有效的空间依赖关系,进一步削弱了模型对旋转目标的判别能力。

针对上述问题,本文设计了感受野扩张与方向感知交互网络(Receptive Field Extending Orientation-Aware Cross Net, RFONet),主要工作如下:

(1)基于ResNet骨干网络设计RFONet网络进行特征提取。利用ResNet的残差连接结构,有效缓解深度网络的梯度消失问题,从而能够构建深层的、稳定的特征提取器。同时,ResNet特征金字塔网络结构,为后续的感受野扩张模块提供了富含语义的深层特征,为方向感知交互注意力机制模块提供了富含细节纹理的浅层特征。

(2)感受野扩张(Receptive Field Expansion, RFE)模块,该模块通过大核分解、深度和空洞卷积组合以及并联分支结构,在不显著增加参数与

计算复杂度的前提下,融合多尺度特征图,扩大感受野。既增强了网络聚焦局部信息的能力又增强了捕获全局上下文信息的能力,解决了不同类型目标对上下文信息需求存在差异的问题。

(3)方向感知交互注意力机制(Orientation-Aware Cross Attention, OACA)模块,该模块通过使用水平与垂直卷积代替普通卷积核,显式地捕捉方向性信息,弥补了卷积过程中缺失的纹理;设计门控权重注意力机制,将提取的方向信息进行交互,增强长程关联性,抑制冗余噪声。

2 相关工作

2.1 遥感旋转目标检测框架

当前主流的遥感目标检测器大多基于Faster R-CNN^[7]这一经典的两阶段框架构建,其核心在于区域建议网络(Region Proposal Network, RPN)与区域检测头的协同。然而为适应遥感目标方向任意的挑战,遥感框架检测器经历了从HBB到OBB的转变。OBB检测器可大致分为两类:两阶段旋转检测器和单阶段旋转检测器。两阶段检测器,侧重于对RPN生成的初始候选框进行精细化旋转校正。例如,RoI-Transformer^[1]通过引入旋转RoI和学习仿射变换参数,实现对水平建议框的第一次方向修正;Oriented RCNN^[8]则通过设计新型的旋转框编码方案,解决了角度周期性导致的训练不稳定问题,提升了两阶段方法的精度与效率。而单阶段检测器,则是摒弃了区域建议步骤,致力于实现更快的检测速度。例如,S2A^[9]网络通过旋转特征对齐模块,直接从密集的锚点中预测旋转框,保证了特征与目标方向的一致性;R3Det^[10]则采用渐进式回归策略,以“先粗后细”的方式优化旋转框的定位精度。

尽管上述研究在检测头设计上有所突破,但在基础的特征提取方面的工作研究较少。因此,本文并没有把重心放在复杂的检测头部设计上,而是致力于优化主干网络的特征提取能力。通过设计主干网络,为上述先进的旋转检测框架^[1,8-9]提供更具鲁棒性的特征表达。

2.2 扩张感受野

扩张感受野是指通过特定的网络设计技术,扩大感受野范围,从而获取更丰富的上下文信

息。这对于解决遥感图像中目标尺度多样等问题具有重要意义。近年来,该领域的研究主要可以分为以下三种:首先,基于多分支结构和特征融合。通过并联多个具有不同感受野的分支,融合后并输出,以显式地捕获多尺度特征。例如,Bai等人^[11]设计的超融合残差行进模块与多路径特征融合模块、Li等人^[12]提出的多尺度上下文信息聚合模块、赵等人^[13]设计的多尺度深度可分离卷积模块,均通过构建复杂的支路来增强网络对多尺度目标的表征能力。其次,基于层级化构建感受野扩张模块,核心思想是将扩张策略嵌入到不同深度网络层。例如,Yao等人^[14]在特征金字塔的前四层网络中分别构建感受野扩张模块,梁等人^[15]在YOLOv8的特征金字塔和检测头中额外增加尺度扩张的特征图,均实现了在全尺度范围内对不同大小目标的精准感知。最后,基于动态选择机制、注意力融合。该方法致力于让网络根据输入内容自适应地调整感受野,或将其与特征增强模块结合。例如,Li等人^[16]利用空洞卷积核与空间选择机制,实现了感受野的动态调整;而Tie等人^[17]则将感受野扩增模块与双层路由注意力机制相结合,高效地捕获了全局上下文信息。

上述工作有效地验证了扩大感受野对遥感目标检测的积极影响。然而,多数方法侧重于通过复杂的多分支或特征融合来获得多尺度上下文,这不可避免地增加了计算复杂度。与之不同,本文设计的RFE模块是一种轻量化、结构化的设计。在增大感受野的同时,严格控制了参数与计算量的增长,从而实现了高效且丰富的背景信息提取工作。

2.3 注意力机制

近年来,随着深度学习技术的发展,注意力机制作为一种能够让模型聚焦于重要区域的技术,在计算机视觉领域取得了显著进展。早期研究主要聚焦于通道维度,如SE^[18]通过全局池化自适应校准通道重要性;随后扩展到空间维度,出现了增强空间信息聚合能力的GENet^[19];再到发展出通道与空间协同的注意力机制,CBAM^[20]和BAM^[21]便是通过双维度注意力实现更全面特征校准的代表。然而在面对遥感图像的应用中,研究人员根据RSIs的特性,开展了更具针对性

的创新:在特征感知层面, Lin等人^[22]通过双域注意力强化对飞机部件等关键特征的捕捉;在空间感知信息层面, Peng等人^[23]引入坐标注意力机制显式编码空间方位信息;在上下文关系建模层面, Song等人^[24]采用边界感知双流注意力实现前景和背景特征融合;在系统架构层面, Yang^[25]与 Yu^[26]等人将注意力机制继承到主干检测网络上, 分别实现了感受野的动态调整与多尺度特征的融合。

尽管现有注意力机制已取得成功, 但其在遥感场景下的应用仍存局限:一方面, 通用的注意力机制对方向性特征的提取不够直接, 另一方面, 它们在建立长程依赖时容易引入冗余噪声。本文设计的OACA模块通过引入水平与垂直卷积来显式地捕捉方向性信息, 弥补缺失的特征;通过门控权重对长程交互进行筛选, 在增强特征关联的同时, 有效抑制了背景噪声, 实现了更精准、更鲁棒的特征增强。

3 RFONet 网络结构

3.1 主体网络框架

图1为整体的网络框架图, 主要由四个阶段构成, 每个阶段包含 Patch embed 和 RFONet Block 两部分(N_l 代表模块中子块堆叠的个数, l 代表某个阶段)。其中 Patch embed 主要负责通道映射与阶梯式下采样, 为后续模块提供合适的特征输入; RFONet Block 负责增强目标特征和输出高质量特征, 是整个网络的核心构成。图1(b)为 RFONet Block 的结构, 主要由 RFO 子块和前馈网络(FFN)子块通过残差连接^[27]堆叠构成。图1(c)为 RFO 和 FFN 的主要结构, 其中 RFO 采用并联分支结构:一条分支为感受野扩张模块, 另一分支为方向感知交互注意力机制模块。FFN 由两个 1×1 卷积、一个激活函数和一个 Dropout 层组成。使用 1×1 卷积代替全连接层减少参数量同时保留了空间信息, 加入 Dropout 层防止过拟合。

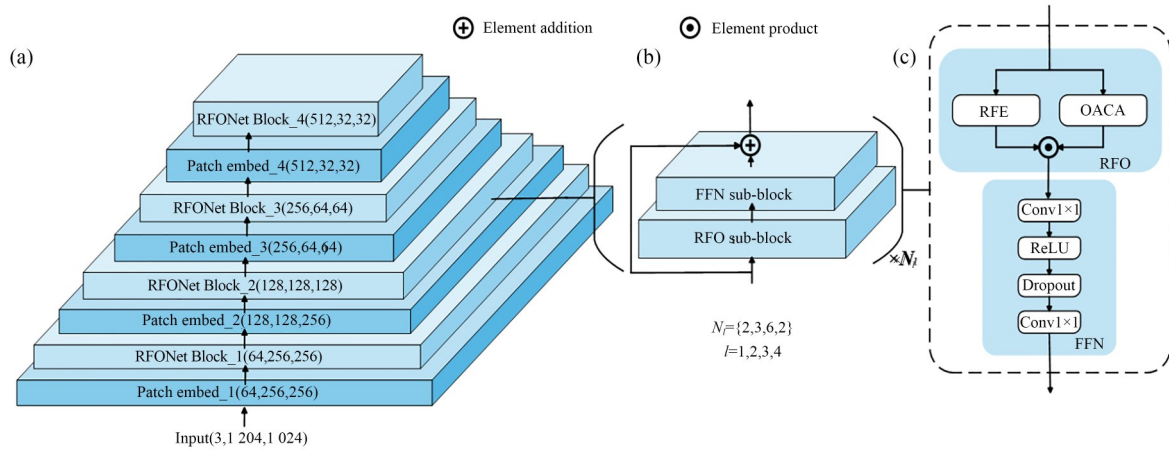


图1 RFONet的整体架构以及各个部分的组成结构

Fig. 1 Overall architecture of RFONet and the components of each part

输入、输出特征的通道数、维度以及 N_l 具体数值参考图1, 其中子块堆叠的数量对于目标检测起着关键的作用。由于前两个阶段图像经过下采样较少, 基础的边缘、纹理特征保留完整。少量子块堆叠能快速地构建低层特征、保存目标的信息避免过度抽象。随着第三阶段下采样到合适的尺度, 需要聚焦复杂的语义信息的提取。更多的子块的堆叠, 可以更快地提取高层信息,

建立目标间的关联。最后第四阶段, 由于特征图已经压缩到较低的分辨率, 需要巩固高层特征、优化输出。少量子块的堆叠降低计算量的同时, 对前期提取的复杂特征做精炼, 确保最终输出的特征图既有丰富语义, 又能适配检测头对特征维度、精度的要求, 更好地完成目标分类与定位。

3.2 感受野扩张模块

图2为 RFE 模块详细的结构图, 主要为了增

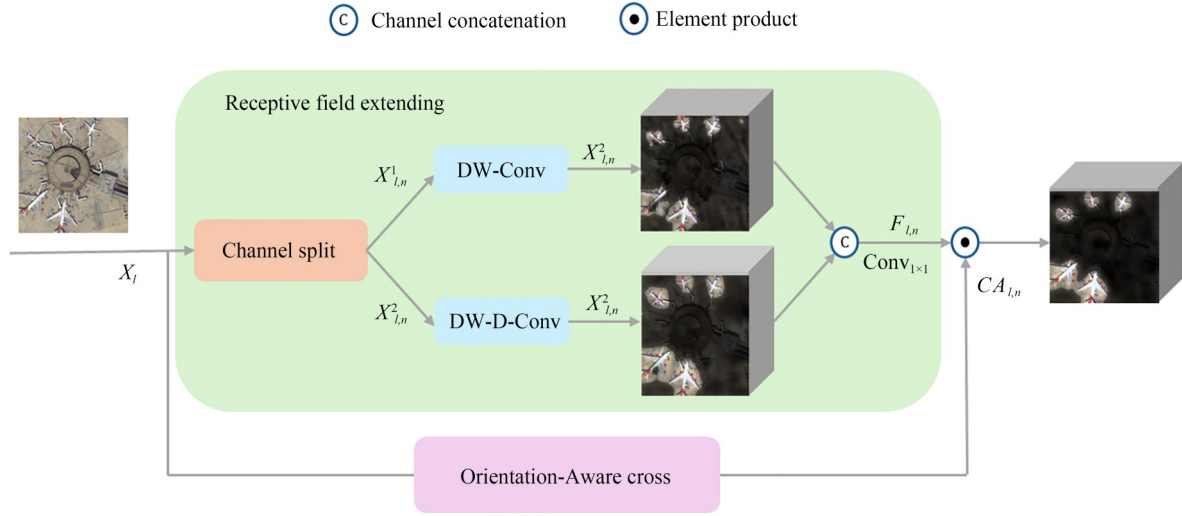


图 2 RFE 模块的详细结构

Fig. 2 Detailed structure of the RFE model

大感受野、提取多尺度的背景信息。首先为了便于提取不同尺度的背景信息,将输入特征图 $X_l \in \mathbb{R}^{C_l \times H_l \times W_l}$ 沿通道维度一分为二,生成两条路径:

$$X_{l,n}^1 = X_l[:\frac{1}{2}C_l, \dots], \quad (1)$$

$$X_{l,n}^2 = X_l[\frac{1}{2}C_l:, \dots], \quad (2)$$

其中: l 表示第 l 阶段, n 表示堆叠的第 n 个 RFO 子块。 $X_{l,n}^1 \in \mathbb{R}^{\frac{1}{2}C_l \times H_l \times W_l}$ 和 $X_{l,n}^2 \in \mathbb{R}^{\frac{1}{2}C_l \times H_l \times W_l}$ 是通道分割生成的特征,在特征提取过程中既能保留原本特征也不容易出现信息冗余。 C, H, W 分别代表通道数、特征高度和宽度。

同时为了降低并联卷积带来计算量,引进大核分解的策略。标准大核可以分成三个卷积核:深度卷积 (Depth-Wise Convolution, DW-Conv)、深度膨胀卷积 (Depth-Wise Dilation Convolution, DW-D-Conv) 和逐点卷积 (Point Convolution, P-Conv)。选取其中 DW-Conv 和 DW-D-Conv 作为两条路径的卷积分别提取不同尺度的感受野:

$$X_{l,n}^1 = DWConv_{k_1 \times k_2}(X_{l,n}^1), \quad (3)$$

$$X_{l,n}^2 = DWDCov_{k_2 \times k_2}(X_{l,n}^2), \quad (4)$$

其中 $X_{l,n}^1$ 的感受野 RF_1 , $X_{l,n}^2 \in \mathbb{R}^{\frac{1}{2}C_l \times H_l \times W_l}$ 的感受野 RF_2 、膨胀率 d 、深度卷积核 k_1 和深度膨胀卷积核 k_2 定义如式(5)~式(7)所示:

$$k_1 < k_2, d_1 = 1, d_1 < d_2 < RF_1, \quad (5)$$

$$RF_1 = k_1, RF_2 = d_2(k_2 - 1), \quad (6)$$

$$RF_{\text{last}} = \max(RF_1, RF_2). \quad (7)$$

为了避免空洞采样间隔超过已有感受野范围、特征提取出现间隙,所以 d 应该满足上述条件。同时卷积核之间采取并联,不存在感受野的递进,所以输出感受野 RF_{last} 由两者之中最大的决定。最后,将不同尺度的感受野进行融合得到最终特征图:

$$F_{l,n} = Conv_{1 \times 1}([X_{l,n}^1, X_{l,n}^2]), \quad (8)$$

其中: $[]$ 表示通道相加、 $Conv_{1 \times 1}$ 表示 1×1 卷积、融合的 $F_{l,n} \in \mathbb{R}^{C_l \times H_l \times W_l}$ 包含了精细的局部信息和丰富的背景信息。

3.3 方向感知交互注意力

为解决遥感目标方向分布任意的问题,设计 OACA 模块。OACA 模块如图 3 所示,首先输入特征经过平均池化后进行 1×1 卷积获取区域特征:

$$F_{l,n}^{\text{pool}} = Conv_{1 \times 1}(P_{\text{avg}}(X_l)), \quad (9)$$

其中, P_{avg} 表示平均池化,将得到的 $F_{l,n}^{\text{pool}}$ 经过大核深度带状卷积:

$$\begin{cases} F_{l,n}^h = DWConv_{1 \times k}(F_{l,n}^{\text{pool}}) \\ F_{l,n}^v = DWConv_{k \times 1}(F_{l,n}^{\text{pool}}) \end{cases}, \quad (10)$$

其中: $F_{l,n}^h$ 和 $F_{l,n}^v$ 分别表示经过水平卷积和垂直卷积得到特征图。引入条状卷积,主要的原因有两点。首先,大核深度带状卷积是轻量化级的。相比于 $k \times k$ 二维深度卷积的卷积核不断地变大使得参数量呈平方式的增长,一维的深度核通过串

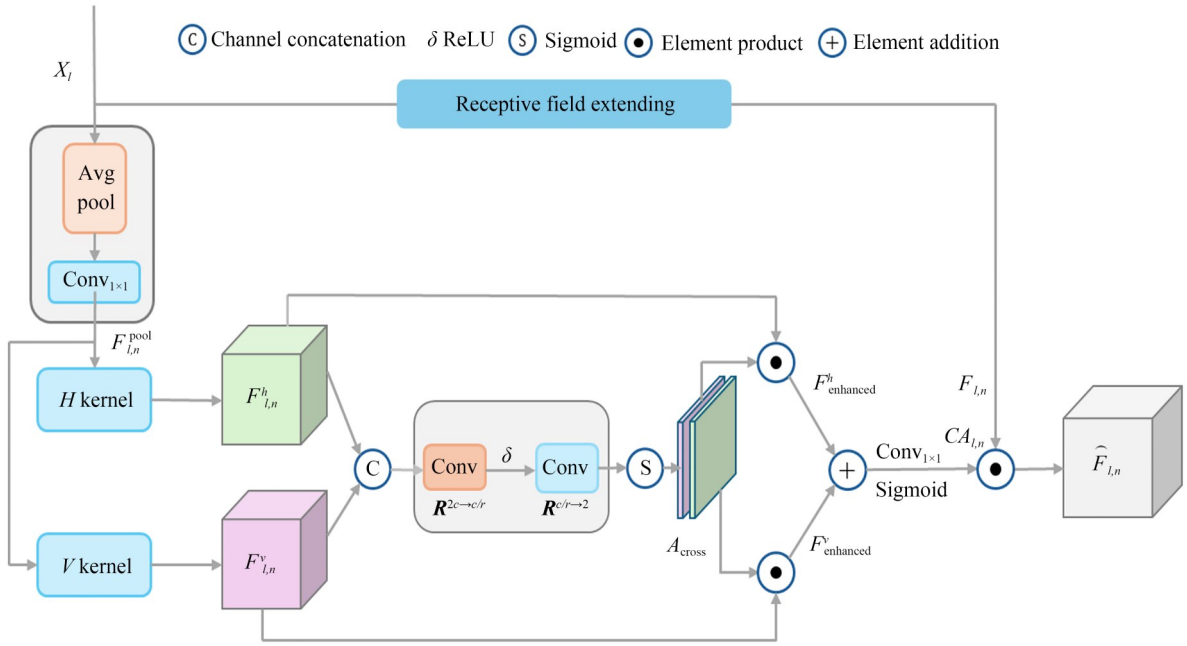


图3 OACA模块的详细结构

Fig. 3 Detailed structure of the OACA Model

联连接可以实现相似的效果,同时参数量减少。其次,水平和垂直条状卷积分别聚焦水平和垂直方向的结构信息,利用提取和识别图像中高纵横比物体(如船舶、桥梁等)的特征。为了提取更完整的精细特征,设置 $k=[21, 19, 17, 15]$ (后续展示带状卷积核大小的选取)。由于遥感目标的几何特征在不同方向呈现差异性,而条状卷积只能单纯捕捉单一的方向信息。为了实现不同方向的信息交互,设计了交叉感知系统。首先将得到方向特征图进行拼接,再经过注意力机制生成方向权重:

$$A_{cross} = S(W_2 \cdot \delta(W_1 \cdot [F_{l,n}^h, F_{l,n}^v])), \quad (11)$$

其中: $W_1 \in R^{2C \rightarrow (C/r)}$ 压缩通道至 C/r (r 为压缩比), δ 表示激活函数ReLU, $W_2 \in R^{(C/r) \rightarrow 2}$ 将通道数由 C/r 转换成2, S 表示Sigmoid函数。通过压缩比 r 控制参数量, r 取4时既不影响准确度又减少参数量。

将生成的方向权重 $A_{cross} \in R^{2 \times H_l \times W_l}$ 拆分为水平门控 A_h 和垂直门控 A_v 用作对原始方向加权:

$$\begin{cases} F_{l,n}^{h,enhanced} = F_{l,n}^h \odot A_v \\ F_{l,n}^{v,enhanced} = F_{l,n}^v \odot A_h \end{cases}, \quad (12)$$

其中: $F_{l,n}^{h,enhanced}$ 和 $F_{l,n}^{v,enhanced}$ 为增强的方向特征(方向

权重和方向特征交叉加权), $\odot$ 表示逐元素相乘。水平特征受垂直权重调控,垂直特征受水平权重调控,建立方向之间的交互。将增强的方向特征相加并通过 1×1 卷积和Sigmoid函数生成最终注意力图:

$$CA_{l,n} = \sigma(Con v_{1 \times 1}(F_{l,n}^{h,enhanced} + F_{l,n}^{v,enhanced})). \quad (13)$$

相比通道相加,采取直接相加实现轻量化同时,保留了水平和垂直增强特征在空间上互补的响应。最终将生成的方向感知注意力图 $CA_{l,n}$ 与RFE模块输出的特征 $F_{l,n}$ 相乘得到最终的输出特征:

$$\hat{F}_{l,n} = F_{l,n} \odot CA_{l,n}. \quad (14)$$

4 实验结果与分析

本文所用实验配置如下:CPU为13th Gen Intel(R) Core(TM) i7-13700K、内存32 GB、显卡型号为NVIDIA GeForce RTX 4090。基于Windows11操作系统,深度学习框架采用Pytorch1.8.1,Python 3.8,CUDA11.1。

4.1 数据集及预处理

对三种流行的遥感目标检测数据集进行实验:DOTA-v1.0^[28]是一个用于遥感检测的大规模数据集,大小从 800×800 到 $4\,000 \times 4\,000$,包

含 15 个类别、188 282 个实例:飞机(Plane, PL)、棒球场(Baseball diamond, BD)、桥梁(Bridge, BR)、操场(Ground track field, GTF)、小型车辆(Small vehicle, SV)、大型车辆(Large vehicle, LV)、船舶(Ship, SH)、网球场(Tennis court, TC)、篮球场(Basketball court, BC)、储罐(Storage tank, ST)、足球场(Soccer-ballfield, SBF)、环形交叉路口(Roundabout, RA)、港口(Harbor, HA)、游泳池(Swimming pool, SP)和直升机(Helicopter, HC)。如图 4(a), DOTA 具有多种尺度和方向的实例, 能更好验证本文的方法。HRSC2016^[29]是一个用于船舶探测的遥感数据集, 如图 4(b), 包含 1 061 张航拍图像, 大小从 300×300 和 $1\,500 \times 900$ 张不等。DIOR-R^[30]是专门为遥感图像中的旋转目标检测任务设计, 如图 4(c), 包含 20 个类别、190 288 个实例, 大小为 800×800 。

参考最近的先进的检测方法的实验, 将 train 和 val 用作训练集可以提高检测精度。所以在基

准数据集划分方面, 采用各数据集的训练验证集(trainval)进行模型训练。对于 DOTA-v1.0 高分辨率遥感数据集, 由于单张图片分辨率可达上万像素, 无法直接训练。所以采用滑动窗口策略将原始图像裁切为 $1\,024 \times 1\,024$ pixels 的样本块, $1\,024 \times 1\,024$ 是一个在计算效率和特征保留之间取得的平衡值。并设置 200 pixels 的重叠区域, 确保了大多数目标能完整地出现在至少一个样本块中。由于 HRSC2016 数据集分辨率较小、分布不规则, 则统一将长边调整为 800 pixels 来调整比例, 而不改变纵横比。优化器选用 AdamW^[32], 为了避免衰减系数过小引发过拟合, 衰减系数过大引发欠拟合, 所以权重衰减系数设为 0.05; 同时初始学习率设置为 0.000 1 来平衡衰减系数带来的约束, 避免收敛困难。值得注意的是, 所有模型的浮点运算量(FLOPs)均基于标准 1024×1024 输入分辨率进行测算, 评价指标采用的是平均精度均值(mean Average Precision, mAP)。



图 4 数据集样例图

Fig. 4 Sample diagram of datasets

4.2 对比实验

4.2.1 DOTA 数据集对比实验

将提出的 RFONet 在 DOTA 数据集上与其他主流方法进行了对比。实验对比如表 1 所示,相较于 EMO2-DETR,O2-DETR 和 ARS-DETR 基于 Transformer 检测器方法,以 O-RCNN 为框架的 RFONet 在 mAP 上分别提升了 5.97%,4.73% 和 2.72%;相较于 R3Det,SASM 和 RTMDet-R 单阶段检测器,RFONet 检测精度分别提高了 3.18%,1.96% 和 0.63%;相较于 SCRDet,G. V 和 ReDet 二阶段检测器,检测精度分别提高了 4.27%,1.86% 和 0.36%。同时与最近提出的基于滑动比 rs 彻底解决 OBB 边界框不连续的 COBB 检测方法相比,检测精度提升 0.36%。为了验证 RFONet 在旋转目标检测领域的泛化能力,将 RFONet 整合到 S2ANet,RoiTrans 和 O-RCNN 三个流行的旋转检测框架上。实验结果表明,RFONet 的 mAP 分别达到

75.47%,75.69% 和 76.88%,相较于其他三个框架分别提升了 1.35%,1.64% 和 1.01%。

如表 1 所示,相比于 SOTA 算法 AO2-DETR,RFONet 的检测精度略低 0.85%。主要原因在于两者所使用的架构不同:AO2-DETR 基于 Transformer 架构,采用端到端集合预测、深度可分离卷积替代自注意力机制优化编码器,并借助高召回率微调策略,使得算法在建模长程依赖与复杂上下文关系方面具有很大优势。但是 RFONet 为基于 CNN 轻量化网络,核心贡献在于设计了感受野扩张模块(RFE)与方向感知交互注意力模块(OACA),重点解决遥感图像目标检测中尺度差异大、方向分布复杂的问题。例如在船舶(SH)与大型车辆(LV)方向分布复杂的场景下,RFONet 检测精度取得最高,优于 AO2-DETR,这验证了 OACA 模块在方向感知上的有效性;在飞机(PL)、环形交叉口(RA)、小型车辆(SV)等需要丰富背景信息的场景下,RFONet 检

表 1 不同模型的精度对比(DOTA)

Tab. 1 Comparison of different models(DOTA)

Method	Back-bone	#P M	#F G	PL	BD	BR	GTF	SV	LV	SH
EMO ^[32]	R-50	40.8	340	87.99	79.46	45.74	66.64	78.90	73.90	73.30
CM ^[33]	R-50	41.1	198	89.02	80.56	49.41	61.98	77.99	74.19	83.74
O ² DETR ^[34]	R-50	74.3	304	86.01	75.92	46.02	66.65	79.70	79.93	89.17
SCRDet ^[2]	R-50	41.9	—	89.98	80.65	52.09	68.36	68.36	60.32	72.41
R3Det ^[10]	R-50	41.9	336	89.50	81.20	50.50	66.10	70.90	78.70	78.20
ARS ^[35]	R-50	41.6	—	86.97	75.56	48.32	69.20	77.92	77.94	87.69
SASM ^[36]	R-50	36.6	—	86.42	78.97	52.47	69.84	77.30	75.99	86.72
G. V. ^[37]	R-50	41.1	198	89.64	85.00	52.26	77.34	73.01	73.14	86.82
ReDet ^[38]	R-50	31.6	54	89.21	80.03	47.88	69.73	82.05	83.33	88.63
RTM-R ^[39]	R-50	4.9	20.5	88.79	82.64	53.97	74.00	78.13	84.06	88.04
COBB ^[40]	R-50	41.9	—	—	—	—	—	—	—	—
AO2 ^[41]	R-50	74.3	304	89.27	84.97	56.67	74.89	78.87	82.73	87.35
S2ANet ^[9]	R-50	38.5	198	89.11	82.84	48.37	71.11	78.11	78.39	87.25
	RFONet	26.9	168	89.62	83.60	50.63	71.63	78.46	80.43	89.24
RoiTrans ^[1]	R-50	55.1	200	89.01	77.48	51.64	72.07	74.43	77.55	87.76
	RFONet	46.9	168	89.46	80.57	49.4	72.57	77.64	81.37	87.39
ORCNN ^[8]	R-50	41.1	199	89.46	82.12	54.78	70.86	78.93	83.00	88.20
	RFONet	32.9	167	89.94	82.74	54.97	72.94	79.86	85.38	89.63

续表 1 不同模型的精度对比(DOTA)
Tab. 1 Comparison of different models(DOTA)

Method	Back-bone	TC	BC	ST	SBF	RA	HA	SP	HC	mAP
EMO ^[32]	R-50	90.40	80.55	85.89	55.19	63.62	51.83	70.15	60.04	70.91
CM ^[33]	R-50	89.44	78.01	83.52	47.64	65.93	63.68	67.07	61.59	71.59
O ² DETR ^[34]	R-50	90.44	81.19	76.00	56.91	62.45	64.22	65.80	58.96	72.15
SCRDet ^[2]	R-50	90.85	87.94	86.86	65.02	66.68	66.25	68.24	65.21	72.61
R3Det ^[10]	R-50	90.80	85.30	84.20	61.80	63.80	68.20	69.80	67.20	73.70
ARS ^[35]	R-50	90.50	77.31	82.86	60.28	64.58	74.88	71.76	66.62	74.16
SASM ^[36]	R-50	90.89	82.63	85.66	60.13	68.25	73.98	72.22	62.37	74.92
G. V. ^[37]	R-50	90.74	79.02	86.81	59.55	70.91	72.94	70.86	57.32	75.02
ReDet ^[38]	R-50	90.91	86.31	86.85	59.94	62.30	74.23	71.97	57.03	75.36
RTM-R ^[39]	R-50	90.89	87.78	85.75	61.76	60.39	75.96	68.07	63.59	76.25
COBB ^[40]	R-50	—	—	—	—	—	—	—	—	76.52
AO2 ^[41]	R-50	90.50	84.68	85.41	61.97	69.96	74.68	72.39	71.62	77.73
S2ANet ^[9]	R-50	90.83	84.90	85.64	60.36	62.60	65.26	69.13	57.94	74.12
	RFONet	89.50	86.39	85.82	63.4	64.74	67.83	70.36	60.34	75.47
RoiTrans ^[11]	R-50	90.81	79.71	85.27	58.36	64.11	76.50	71.99	54.06	74.05
	RFONet	90.51	82.90	85.62	64.98	70.17	71.70	72.34	58.72	75.69
ORCNN ^[8]	R-50	90.90	87.50	84.68	63.97	67.69	74.94	68.84	52.28	75.87
	RFONet	90.78	85.40	84.53	64.27	66.48	75.27	70.72	60.31	76.88

注释:R-50表示ResNet-50,红色和绿色加粗表示该类方法检测精度排名前两位(彩图见期刊电子版)

测精度也排在第二位,这得益于RFE模块融合多尺度上下文信息。除此之外,在模型效率方面,RFONet对比AO2-DETR具有很大优势。在参数量(#P)与计算量(#F)两个核心指标上,RFONet的参数量为32.9 M,计算量为167 GFLOPs;而AO2-DETR的对应值分别为74.3 M和304 GFLOPs。RFONet的模型复杂度和计算性能仅为AO2-DETR的54%,55%左右,实现了近乎减半的轻量化。相比之下,RFONet的效率优势来源于轻量化模块设计:RFE模块通过“大核分解”策略和OACA模块采用“条状卷积”与“压缩比”设计。综上所述,AO2-DETR在整体检测精度上的确存在先进性,但本文的贡献主要是通过轻量化的设计,降低参数量和计算量。同时研究背景信息与方向信息协同工作的有效性,解决遥感图像目标检测中目标尺度差异大、方向分布复杂的问题。

4.2.2 HRSC2016数据集对比实验

如表2所示,在HRSC2016数据集上的对比实验中,提出的RFONet模型表现出卓越的检测性能和较低的模型复杂度。在参数量方面,RFONet的模型规模为32.87 M,相比最近的方

表 2 不同模型的精度对比(HRSC2016)
Tab. 2 Comparison of different models(HRSC2016)

Method	#Params/M	mAP(07)	mAP(12)
Roi Trans ^[11]	55.1	89.85	97.37
G. V. ^[37]	41.1	88.20	—
R3Det ^[10]	41.9	89.26	96.01
S2ANet ^[9]	38.6	90.17	95.01
ReDet ^[38]	31.6	90.46	97.63
O-RCNN ^[8]	41.1	90.50	97.60
Yao ^[12]	—	90.50	98.26
RFONet	32.87	90.58	98.43

法显著减小,体现了模型设计的轻量化。在检测精度上,RFONet在mAP(07)和mAP(12)两种评价标准下均取得了最优结果,分别达到90.58%和98.43%,并且在绝大多数情况下显著优于其他方法。相比于利用全连接层对候选水平锚框进行旋转的算法Roi-Trans,mAP(07)提升了0.73%,mAP(12)提升了1.06%。相较于引入新型边框编码系统的算法O-RCNN和G.V来说,mAP(07)分别提升了0.08%和2.38%,mAP(12)提升了0.83%。相比于提出编码旋转等变特征和旋转不变特征的算法ReDet,mAP(07)提升了0.12%,mAP(12)提升了0.80%。相比于最新提出的算法Yao,mAP(07)提升了0.08%,mAP(12)提升了0.17%。同时对于单阶段检测网络R3Det和S2ANet,mAP(07)分别

提升了1.32%和0.41%,mAP(12)提升了2.42%和3.42%。

4.2.3 DIOR-R数据集对比实验

为验证本文所提RFONet模型的鲁棒性,在具有挑战性的DIOR-R数据集上进行了对比实验。如表3所示,RFONet检测精度达到了65.06%,相较于经典的单阶段基准方法RetinaNet-O(57.55%),mAP提升了7.51%;相较于两阶段基准FasterR-CNN-O(59.54%),mAP提升了5.52%。此外,RFONet的性能也优于主流的遥感图像检测器,相较于ROI-Trans(63.87%)、O-RCNN(64.30%),mAP分别提升了1.19%,0.76%。实验结果表明,无论是传统的基准模型还是先进的旋转检测框架,本文提出的RFONet检测性能更具优势。

表3 不同模型的精度对比(DIOR-R)

Tab. 3 A comparison of different models(DIOR-R)

Method	RetinaNet-O ^[42]	Faster RCNN-O ^[7]	ROI-Trans ^[1]	O-RCNN ^[8]	RFONet
mAP/%	57.55	59.54	63.87	64.3	65.06

注:-O表示检测方法采用旋转框

4.3 消融实验

4.3.1 大核分解和感受野大小

RFE模块中深度分离卷积和深度膨胀可分离卷积大小选择是影响获取感受野大小的关键因素。较小的感受野对先验信息涵盖较少,较大的感受野会引入过量的噪声。所以选取感受野为5,14,18和27进行实验。如图表4所示,过大过小的感受野都会影响模型的检测精度。当感受野达到14时,提出的RFONet可以既高效又精准地实现目标检测任务。同时针对模块中卷积核大小的选择也影响着特征提取效果, k_1 主要提取目标的纹理特征而 k_2 主要提取包含目标在内的背景信息。卷积核太小导致特征和背景信息提取不全;卷积核太大会导致提取特征的同时引入大量的无关信息;卷积核之间差距太大也会导致目标与背景信息之间出现割裂。所以当 $k_1=5, k_2=7, d_2=3$ 时,能在目标特征和背景信息之间实现良好的权衡,在FPS(每秒帧数)和mAP(平均精度)方面实现最佳性能。

表4 卷积核分解和感受野大小

Tab. 4 Convolutional Kernel Decomposition and the Receptive Field Size

(k_1, d_1)	(k_2, d_2)	RF_{last}	FPS	mAP/%
(3,1)	(5,2)	5	33.7	76.80
(5,1)	(7,3)	14	33.4	76.88
(5,1)	(9,3)	18	31.7	76.83
(7,1)	(9,4)	27	32.6	76.75

4.3.2 带状卷积核大小

带状卷积中的核大小对于提出的方向感知交互至关重要。较小的核在捕获高纵横比目标时会出现特征残缺问题,而较大的核在捕获的同时也会引入大量噪声。为此提出衰减策略,在浅层配备大核提取底层特征,在深层配备小核提取语义特征。正如表5所示,卷积核逆向尺寸序列(17→11)相较正向配置(11→17)带来0.64% mAP提升,验证了层级衰减策略的有效性。同时不断约束最大初始核尺寸,改变核尺寸缩减步长,当采用[21,19,17,15]配置时模型精度达到峰值。

表 5 带状卷积核大小选取

Tab. 5 Strip Convolution Kernel Size Selection

Kernel Size	#P/M	FLOPs/G	mAP/%
(11,13,15,17)	32.88	166.97	75.85
(17,15,13,11)	32.87	167.01	76.49
(19,17,15,13)	32.87	167.05	76.79
(21,19,17,15)	32.88	167.09	76.88
(23,21,19,17)	32.89	167.12	76.72
(21,17,13,9)	32.87	167.09	76.68

4.3.3 压缩比的大小

压缩比用于控制交叉注意力模块中通道压缩的比例,对模型的性能和计算效率有重要影响。为了确定最佳的压缩比,测试了不同的取值,包括 2,4 和 8。如表 6 所示,较小的压缩比虽然能够保留更多的通道信息,但计算量显著增加且性能提升有限。较大的压缩比虽然减少了计算量,但通道信息的丢失导致性能下降。最终,

表 6 压缩比大小

Tab. 6 reduction_ratio size

Reduction ratio	#P/M	FLOPs/G	mAP/%
2	34.38	170.74	76.90
4	32.89	167.08	76.88
8	30.64	160.26	75.93

当压缩比取 4 时,模型在性能和计算效率之间取得了最佳平衡。

4.3.4 RFE 与 OACA 模块协同作用分析

如表 7 所示,为探究 RFE 与 OACA 模块的协同作用,设计了五种不同组合方式的消融实验。实验结果表明:单模块性能方面,RFE 模块检测精度比 OACA 模块高了 0.48%。这说明 RFE 提供的多尺度信息具有通用性,更适配基础特征的建模。而 OACA 侧重于对高纵横比的特定目标精细化表征,适用场景相对局限;多模块协作方面,除“OACA+RFE”串联结构外,其余两个多模块方案检测精度分别达到了 76.72%,76.88%,均高于单模块最优的 RFE。

表 7 RFE 和 OACA 模块协同作用分析

Tab. 7 RFE and OACA synergy analysis module

Backbone	#P/M	FLOPs/G	mAP/%	CS	MI
RFE	30.38	158.38	76.52	—	—
OACA	31.81	163.06	76.04	—	—
OACA+RFE	34.62	172.64	76.45	0.177 2	0.174
RFE+OACA	34.62	172.64	76.72	0.003 1	0.189
(RFE, OACA)	32.87	167.25	76.88	0.002 4	0.232

注:OACA+RFE 表示:先 OACA 后 RFE 的串联方式,RFE+OACA 同理,(RFE, OACA)表示并联方式

为进一步分析不同连接方式对 RFE 与 OACA 模块协作的影响,本文计算了两个模块输出特征图间的余弦相似度(Cosine Similarity, CS)与互信息(Mutual Information, MI)。结果显示,“OACA+RFE”串联方式的 CS 值较高,MI 值较低,表明 OACA 模块与 RFE 模块输出的特征图相似,同时信息互补较少。这也是导致检测精度下降的原因:OACA 模块提取的方向信息抑制了背景信息,导致后续 RFE 模块输入受限、模块间功能发生冲突。而在“RFE+OACA”串联方式中,CS 值降低说明模块之间特征相互独

立,但其 MI 值也表明串联的连接方式下信息互补有限。相比之下,并联结构同时实现了最低的 CS 值(0.002 4)与最高的 MI 值(0.232),说明并联的连接方式既能保持 RFE 与 OACA 的功能独立性,又能最大程度促进二者间的信息互补。所以综上所述,RFE 和 OACA 模块之间采取并联的结构设计。

4.4 可视化

检测结果如图 5 所示(GT 代表真值,Ours 为本文提出的 RFONet,其中黄色、蓝色和绿色的虚线圆圈分别代表漏检、错检和缺陷检测,彩图见



图5 RFONet和主流算法的检测结果对比

Fig. 5 Comparison of detection results between RFONet and mainstream algorithms

期刊电子版),相比主流的检测方法出现了漏检、错检和残缺检测的问题,RFONet面对尺度跨度

大的飞机群、排列密集的小型汽车和船舶、纵横比大的港口仍保持较高的检测精度。图6展示了

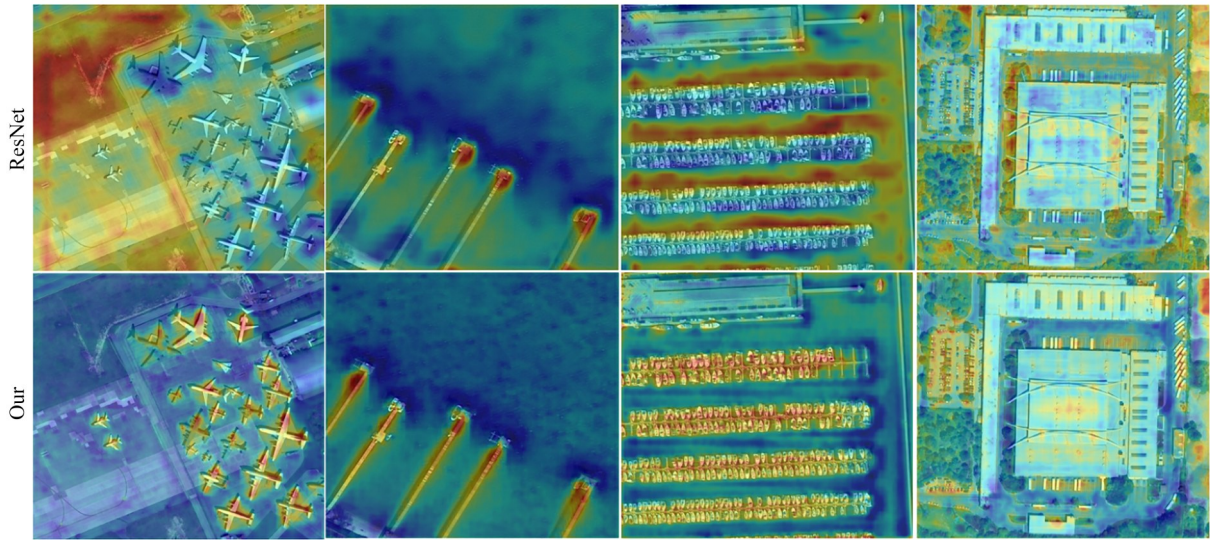


图6 RFONet和ResNet检测热力图对比

Fig. 6 Comparison of RFONet and ResNet detection heat maps

与ResNet对比的热力图,证明了RFONet面对遥感图像目标的特性仍具有精确检测。

4.5 错误案例分析

尽管RFONet在大部分场景下表现优异,但在极端密集、背景复杂、目标极小等场景下仍会出现漏检或错检。其根本原因在于:RFONet以CNN为基础骨干网络,架构存在局限性。在极

端密集场景中,目标之间的空间关联超出卷积的建模范围,导致相邻目标特征信息混淆了背景信息;在复杂背景场景中,CNN难以跨越局部像素级特征,提取的背景信息中存在大量噪声,削弱了模型对目标与背景的判别能力;在目标极小的场景中,卷积过程导致极小的目标的像素丢失,导致无法为模型提供足够的特征信息。

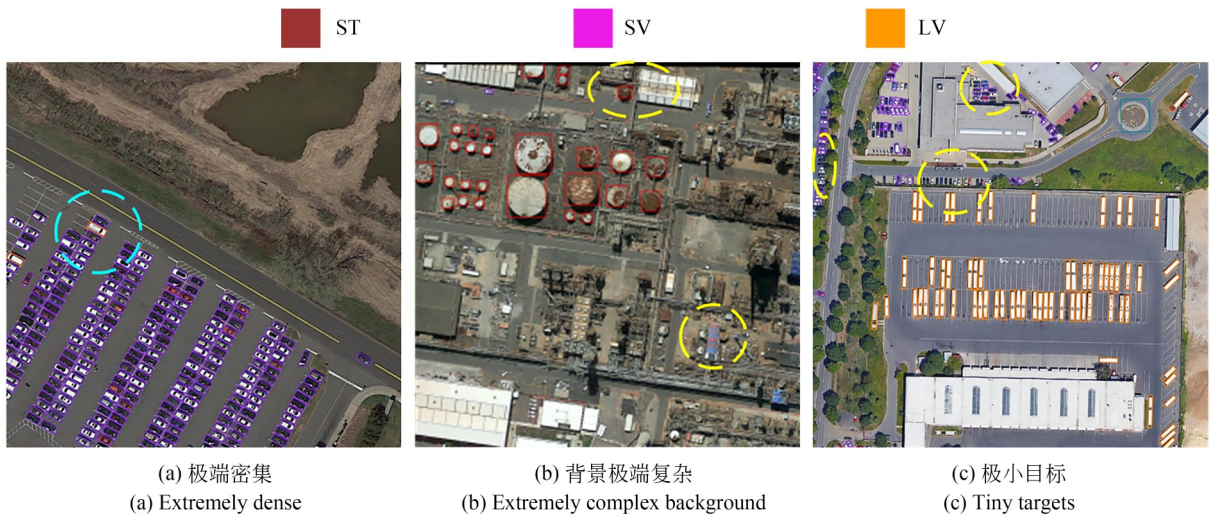


图7 特定场景下检测结果

Fig. 7 Detection results in specific scenarios

5 结 论

本文提出了一种融合背景信息与方向信息

的遥感目标检测网络,有效解决了因目标尺度差异大和方向分布复杂导致检测精度不足的问题。本文研究核心在于设计两个新颖的模块:

感受野扩张模块(RFE),通过轻量化的结构高效融合多尺度背景信息,克服了传统固定感受野的局限性;方向感知交互注意力模块(OACA),通过条状卷积与交叉注意力机制,精准捕捉方向信息并增强其交互,解决了通用注意力机制在方向感知上的缺陷。在DOTA, HRSC2016和DIOR-R数据集上的实验结果表明,提出的RFONet网络检测精度达到了76.88%,98.43%和65.06%,验证了背景信息

与方向信息协同工作的有效性。同时,RFONet在特定场景下会出现漏检或错检,未来工作从增强模型对目标和复杂背景的区分能力方面进行提升。

作者贡献声明:

胡贺南:方法的提出;

马奥阳:论文构思和撰写;

李荣华:论文审核与编辑写作。

参考文献:

- [1] DING J, XUE N, LONG Y, *et al.* Learning RoI transformer for oriented object detection in aerial images[C]. 2019 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. June 15-20, 2019, Long Beach, CA, USA. IEEE, 2020: 2844-2853.
- [2] YANG X, YANG J R, YAN J C, *et al.* SCRDet: towards more robust detection for small, cluttered and rotated objects[C]. 2019 *IEEE/CVF International Conference on Computer Vision (ICCV)*. October 27 - November 2, 2019, Seoul, Korea. IEEE, 2020: 8231-8240.
- [3] HUANG Y X, LIU H I, SHUAI H H, *et al.* DQ-DETR: DETR with dynamic query for tiny object detection [C]. *Computer Vision-ECCV 2024*. Cham: Springer, 2025: 290-305.
- [4] XU L, YU J W, MING D P. HODet: a new detector for arbitrary-oriented rectangular object in optical remote sensing imagery[J]. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2024, 17: 2918-2926.
- [5] XU H, LIU X Y, XU H N, *et al.* Rethinking boundary discontinuity problem for oriented object detection [C]. 2024 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. June 16-22, 2024, Seattle, WA, USA. IEEE, 2024: 17406-17415.
- [6] ZHANG X, YANG X, LI Y X, *et al.* RSAR: restricted state angle resolver and rotated sar benchmark[C]. 2025 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. June 10-17, 2025, Nashville, TN, USA. IEEE, 2025: 7416-7426.
- [7] REN S Q, HE K M, 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.
- [8] XIE X X, CHENG G, WANG J B, *et al.* Oriented R-CNN for object detection[C]. 2021 *IEEE/CVF International Conference on Computer Vision (ICCV)*. October 10-17, 2021, Montreal, QC, Canada. IEEE, 2022: 3500-3509.
- [9] HAN J M, DING J, LI J, *et al.* Align deep features for oriented object detection[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2022, 60: 5602511.
- [10] YANG X, YAN J C, FENG Z M, *et al.* R3Det: refined single-stage detector with feature refinement for rotating object [J]. *Proceedings of the AAAI Conference on Artificial Intelligence*, 2021, 35(4): 3163-3171.
- [11] 白晨帅, 白晓凤, 邹开俊, 等. 超融合残差行进几何感知的遥感目标检测[J]. *光学精密工程*, 2025, 33(8): 1289-1302.
BAI C S, BAI X F, WU K J, *et al.* Remote sensing object detection algorithm based on ultra fusion residual marching geometric perception [J]. *Opt. Precision Eng.*, 2025, 33(8): 1289-1302. (in Chinese)
- [12] 李智杰, 惠爱婷, 李昌华, 等. 面向遥感图像道路提取的多尺度上下文感知网络[J]. *光学精密工程*, 2025, 33(4): 610-623.
LI Z J, HUI A T, LI C H, *et al.* Multi-scale context-aware network for road extraction in remote sensing images [J]. *Opt. Precision Eng.*, 2025, 33(4): 610-623. (in Chinese)
- [13] 赵晓, 杜光磊. 基于条件先验增强和扩散模型的遥感图像超分辨重建算法[J]. *液晶与显示*,

- 2025, 40(7): 1067-1079.
- ZHAO X, DU G L. Super-resolution of remote sensing images based on conditional prior enhancement and diffusion models[J]. *Chinese Journal of Liquid Crystals and Displays*, 2025, 40(7): 1067-1079. (in Chinese)
- [14] 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 T T, ZHAO H X, FENG Z H, *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. (in Chinese)
- [15] 梁海翔, 唐艳慧, 王宇庆, 等. 高分辨率遥感图像的目标检测[J]. *液晶与显示*, 2024, 39(10): 1350-1360.
- LIANG H X, TANG Y H, WANG Y Q, *et al.* Object detection methods for high resolution remote sensing images[J]. *Chinese Journal of Liquid Crystals and Displays*, 2024, 39(10): 1350-1360. (in Chinese)
- [16] LI Y X, HOU Q B, ZHENG Z H, *et al.* Large selective kernel network for remote sensing object detection [C]. *2023 IEEE/CVF International Conference on Computer Vision (ICCV)*. October 1-6, 2023, Paris, France. IEEE, 2024: 16748-16759.
- [17] 帖军, 秦锦添, 郑禄, 等. 融合感受野扩增与特征增强的遥感小目标检测[J]. *激光杂志*, 2024, 45(12): 81-91.
- TIE J, QIN J T, ZHENG L, *et al.* Remote sensing small target detection integrating receptive field amplification and feature enhancement [J]. *Laser Journal*, 2024, 45(12): 81-91. (in Chinese)
- [18] HU J, SHEN L, SUN G. Squeeze-and-excitation Networks [C]. *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. June 18-23, 2018, Salt Lake City, UT, USA. IEEE, 2018: 7132-7141.
- [19] HU J, SHEN L, ALBANIE S, *et al.* Gather-Excite: Exploiting Feature Context in Convolutional Neural Networks [EB/OL]. 2018: *arXiv*: 1810.12348. <https://arxiv.org/abs/1810.12348>
- [20] WOO S, PARK J, LEE J Y, *et al.* CBAM: convolutional block attention module [C]. *Computer Vision-ECCV 2018*. Cham: Springer, 2018: 3-19.
- [21] PARK J, WOO S, LEE J Y, *et al.* BAM: bottleneck attention module [C]. *British Machine Vision Conference*. , 2018
- [22] 林珊玲, 张雪, 陈燕, 等. 融合全局信息与双域注意力机制的光学遥感飞机目标检测[J]. *光学精密工程*, 2024, 32(20): 3085-3098.
- LIN S L, ZHANG X, CHEN Y, *et al.* Integrating global information and dual-domain attention mechanism for optical remote sensing aircraft target detection [J]. *Opt. Precision Eng.*, 2024, 32(20): 3085-3098. (in Chinese)
- [23] PENG G L, YANG Z J, WANG S B, *et al.* AM-FLW-YOLO: a lightweight network for remote sensing image detection based on attention mechanism and multiscale feature fusion [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2023, 61: 4600916.
- [24] SONG J N, ZHOU M L, LUO J, *et al.* Boundary-aware feature fusion with dual-stream attention for remote sensing small object detection [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2025, 63: 5600213.
- [25] 杨志渊, 罗亮, 吴天阳, 等. 改进YOLOv8的轻量级光学遥感图像船舶目标检测算法[J]. *计算机工程与应用*, 2024, 60(16): 248-257.
- YANG Z Y, LUO L, WU T Y, *et al.* Improved lightweight ship target detection algorithm for optical remote sensing images with YOLOv8 [J]. *Computer Engineering and Applications*, 2024, 60(16): 248-257. (in Chinese)
- [26] 于银山, 丁明鉴, 唐旭, 等. 局部背景关联多注意力轻量化遥感飞机检测模型[J]. *计算机工程与应用*, 1-15.
- YU Y S, DING M J, TANG X, *et al.* Local context-aware multi-attention lightweight model for aircraft detection in remote sensing images [J]. *Computer Engineering and Applications*, 2025: 1-15. (in Chinese)
- [27] HE K M, ZHANG X Y, REN S Q, *et al.* Deep residual learning for image recognition [C]. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. June 27-30, 2016, Las Vegas, NV, USA. IEEE, 2016: 770-778.
- [28] XIA G S, BAI X, DING J, *et al.* DOTA: a large-scale dataset for object detection in aerial images [C]. *2018 IEEE/CVF Conference on Com-*

- puter Vision and Pattern Recognition. June 18-23, 2018, Salt Lake City, UT, USA. IEEE, 2018: 3974-3983.
- [29] LIU Z K, WANG H Z, WENG L B, *et al.* Ship rotated bounding box space for ship extraction from high-resolution optical satellite images with complex backgrounds [J]. *IEEE Geoscience and Remote Sensing Letters*, 2016, 13(8): 1074-1078.
- [30] CHENG G, WANG J B, LI K, *et al.* Anchor-free oriented proposal generator for object detection [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2022, 60: 5625411.
- [31] KINGMA D P, BA J. Adam: a method for stochastic optimization [EB/OL]. 2014: *arXiv*: 1412.6980. <https://arxiv.org/abs/1412.6980>
- [32] HU Z B, GAO K, ZHANG X D, *et al.* EMO₂DETR: efficient-matching oriented object detection with transformers [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2023, 61: 5616814.
- [33] LONG Y, XIA G S, LI S Y, *et al.* On creating benchmark dataset for aerial image interpretation: reviews, guidances, and million-AID [J]. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2021, 14: 4205-4230.
- [34] MA T L, MAO M Y, ZHENG H H, *et al.* Oriented Object Detection with Transformer [EB/OL]. 2021: *arXiv*: 2106.03146. <https://arxiv.org/abs/2106.03146>
- [35] ZENG Y, YANG X, LI Q Y, *et al.* ARS-DETR: Aspect ratio sensitive oriented object detection with transformer [J]. *arXiv preprint arXiv*: 2303.04989, 2023.
- [36] HOU L P, LU K, XUE J, *et al.* Shape-adaptive selection and measurement for oriented object detection [J]. *Proceedings of the AAAI Conference on Artificial Intelligence*, 2022, 36(1): 923-932.
- [37] XU Y C, FU M T, WANG Q M, *et al.* Gliding vertex on the horizontal bounding box for multi-oriented object detection [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2021, 43(4): 1452-1459.
- [38] HAN J M, DING J, XUE N, *et al.* ReDet: a rotation-equivariant detector for aerial object detection [C]. 2021 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. June 20-25, 2021. Nashville, TN, USA. IEEE, 2021: 2785-2794.
- [39] LYU C Q, ZHANG W W, HUANG H A, *et al.* RTMDet: An Empirical Study of Designing Real-Time Object Detectors [EB/OL]. 2022: *arXiv*: 2212.07784. <https://arxiv.org/abs/2212.07784>
- [40] XIAO Z K, YANG G Y, YANG X, *et al.* Theoretically achieving continuous representation of oriented bounding boxes [C]. 2024 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. June 16-22, 2024, Seattle, WA, USA. IEEE, 2024: 16912-16922.
- [41] DAI L H, LIU H, TANG H, *et al.* AO₂DETR: arbitrary-oriented object detection transformer [J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2023, 33(5): 2342-2356.
- [42] LIN T Y, GOYAL P, GIRSHICK R, *et al.* Focal loss for dense object detection [C]. 2017 *IEEE International Conference on Computer Vision (ICCV)*. October 22-29, 2017, Venice, Italy. IEEE, 2017: 2999-3007.

作者简介:



胡贺南(1996—),女,吉林长春人,博士,讲师,硕士生导师,2018年于长春理工大学获得学士学位,2023年于中国科学院长春光学精密机械与物理研究所获得博士学位,主要从事方向为数字图像处理,三维感知,计算机视觉,E-mail: huhenan@djtu.edu.cn

通讯作者:



李荣华(1983—),男,河北唐山人,博士,教授,博士生导师,2005年于大连交通大学获得学士学位,2011年于大连理工大学获得博士学位,主要从事方向为遥感图像处理、航天器姿态视觉测量,E-mail: lironghua705@163.com