



## 基于特征重用和膨胀卷积的遥感图像舰船检测

曲海成, 李瑞柯, 王蒙, 单以盟

引用本文:

曲海成, 李瑞柯, 王蒙, 单以盟. 基于特征重用和膨胀卷积的遥感图像舰船检测[J]. 智能系统学报, 2024, 19(5): 1298–1308.

QU Haicheng, LI Ruike, WANG Meng, et al. Ship detection in remote sensing images via feature reuse and dilated convolution[J]. *CAAI Transactions on Intelligent Systems*, 2024, 19(5): 1298–1308.

在线阅读 View online: <https://dx.doi.org/10.11992/tis.202304002>

## 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>

### 隔级融合特征金字塔与CornerNet相结合的小目标检测

Small target detection based on a combination of feature pyramid and CornerNet  
智能系统学报. 2021, 16(1): 108–116 <https://dx.doi.org/10.11992/tis.202004033>

### 基于注意力融合的图像描述生成方法

An image caption generation method based on attention fusion  
智能系统学报. 2020, 15(4): 740–749 <https://dx.doi.org/10.11992/tis.201910039>

### 基于注意力机制的显著性目标检测方法

Salient object detection method based on the attention mechanism  
智能系统学报. 2020, 15(5): 956–963 <https://dx.doi.org/10.11992/tis.201903001>

### 基于跳跃连接金字塔模型的小目标检测

Skip feature pyramid network with a global receptive field for small object detection  
智能系统学报. 2019, 14(6): 1144–1151 <https://dx.doi.org/10.11992/tis.201905041>

### 多层卷积特征的真实场景下行人检测研究

Research on pedestrian detection based on multi-layer convolution feature in real scene  
智能系统学报. 2019, 14(2): 306–315 <https://dx.doi.org/10.11992/tis.201710019>

DOI: 10.11992/tis.202304002

网络出版地址: <https://link.cnki.net/urlid/23.1538.TP.20240828.1059.026>

# 基于特征重用和膨胀卷积的遥感图像舰船检测

曲海成, 李瑞柯, 王蒙, 单以盟

(辽宁工程技术大学软件学院, 辽宁 葫芦岛 125105)

**摘要:** 在光学遥感图像中, 港口内的舰船目标通常处于密集的船只群中, 并受到周围环境的干扰和遮挡, 如集装箱、车辆等。为了进一步提高现有舰船目标检测算法的精度和泛化性能, 提出了一种基于特征重用和膨胀卷积的遥感图像舰船检测算法。首先构建了基于分组卷积和拆分注意力的残差块来提取特征, 同时嵌入可变形卷积提取更加符合舰船尺度变化的特征; 接着, 构造了多尺度感受野模块, 通过并行提取多尺度特征后再进行融合来减少信息损失; 最后, 在原有特征金字塔的基础上构建了一条自底向上的特征重用聚合路径以提高特征表示能力。在大型遥感数据集 DOTA 和舰船数据集 HRSC2016 上进行实验, 实验结果表明, 所提方法能够有效缓解舰船目标漏检和误检问题, 提高了遥感图像舰船目标检测的精度。

**关键词:** 遥感图像; 舰船检测; 特征重用; 膨胀卷积; 拆分注意力; 分组卷积; 特征金字塔; 可变形卷积

**中图分类号:** TP751; TN911.73 **文献标志码:** A **文章编号:** 1673-4785(2024)05-1298-11

中文引用格式: 曲海成, 李瑞柯, 王蒙, 等. 基于特征重用和膨胀卷积的遥感图像舰船检测 [J]. 智能系统学报, 2024, 19(5): 1298-1308.

英文引用格式: QU Haicheng, LI Ruike, WANG Meng, et al. Ship detection in remote sensing images via feature reuse and dilated convolution[J]. CAAI transactions on intelligent systems, 2024, 19(5): 1298-1308.

## Ship detection in remote sensing images via feature reuse and dilated convolution

QU Haicheng, LI Ruike, WANG Meng, SHAN Yimeng

(College of Software, Liaoning Technical University, Huludao 125105, China)

**Abstract:** In optical remote sensing images, ship targets in ports are often densely grouped and obstructed by the surrounding environment, such as containers and vehicles. To further improve the accuracy and generalization performance of existing ship target detection algorithms, this study proposes a remote sensing image ship detection algorithm based on feature reuse and dilated convolution. First, a residual block based on grouped convolution and split attention is constructed to extract features, with deformable convolution embedded to better handle ship scale variations. Afterward, a multiscale receptive field module is designed to reduce information loss by parallel extraction and fusion of multiscale features. Finally, a bottom-up feature reuse aggregation path is developed based on the original feature pyramid to enhance feature representation. Experiments were conducted on the large-scale remote sensing dataset, DOTA and the ship dataset, HRSC2016. The results show that the proposed method effectively alleviates the issues of missing and false detections of ship targets, leading to increased accuracy in ship target detection in remote sensing images.

**Keywords:** remote sensing image; ship detection; feature reuse; dilated convolution; split attention; group convolution; feature pyramid; deformable convolution

收稿日期: 2023-04-03. 网络出版日期: 2024-08-28.

基金项目: 国家自然科学基金面上项目 (42271409); 辽宁省高等学校基本科研项目 (LJKMZ20220699).

通信作者: 曲海成. E-mail: [quhaicheng@lntu.edu.cn](mailto:quhaicheng@lntu.edu.cn).

随着遥感技术的不断发展和进步, 遥感成像方式也呈现出多样化, 基于遥感图像的目标检测也因此成为计算机视觉领域中一个重要的研究分

支。舰船作为海上重要交通工具和经济活动的关键组成部分,对于国家安全和经济发展具有至关重要的作用<sup>[1]</sup>。光学遥感图像可以获取高分辨率的图像,能够提供更为详细和精准的地表信息,可以获取可见光、红外等多种波段信息,捕捉地表微小目标<sup>[2]</sup>,但是在海陆交接区域,受岸上车辆、集装箱等影响,进行舰船目标检测时会出现误检、漏检现象。另外近岸港口和海上舰船目标分布不平衡,港口通常会有大量密集停靠的船舶,而海上拍摄的图像可能只包含单个目标,这种分布不平衡的数据会影响机器学习模型的学习和识别能力。

传统的目标检测方法包括滑动窗口法<sup>[3]</sup>、模板匹配法<sup>[4]</sup>、背景建模法<sup>[5]</sup>、轮廓法<sup>[6]</sup>等,由于容易受目标尺寸、位置、形状和遮挡等因素的影响,这些方法的检测效果受到限制,对复杂的场景目标存在误检和漏检等问题,并且传统方法通常需要手动设计特征或分类器,难以适应不同场景目标,难以进行有效的迁移学习。2012年 AlexNet 网络<sup>[7]</sup>的提出给遥感图像舰船检测带来新的机遇和挑战,目前基于深度学习的目标检测方法分为基于锚框(anchor-base)和无锚框(anchor-free)检测方法,基于锚框的检测方法是指通过预定义的一组锚框,在图像上生成候选框,然后利用卷积神经网络对候选框进行分类和回归,具体细分为 Faster R-CNN(faster region-based convolutional neural network)<sup>[8]</sup>、Oriented R-CNN<sup>[9]</sup>、ReDet(rotation-equivariant detector)<sup>[10]</sup>、ROI Transformer(region of interest Transformer)<sup>[11]</sup>等两阶段检测方法和 YOLO(you only look once)<sup>[12]</sup>、RetinaNet<sup>[13]</sup>、S<sup>2</sup>A-Net(single-shot alignment net)<sup>[14]</sup>、STYOLO(ShuffleNet Tiny-YOLO)<sup>[15]</sup>等一阶段检测方法,基于锚框的检测方法可以在不同尺度上生成多个不同大小和长宽比的锚框,以适应不同尺寸和形状的目标,因此检测精度较高;无锚检测方法则直接在输入图像上对目标位置进行回归,不需要生成锚框,通常具有更简单的网络和更快的检测速度,例如 FCOS(fully convolutional one-stage object detector)<sup>[16]</sup>、Sasm\_rep-points(shape-adaptive selection and measurement rep-points)<sup>[17]</sup>、CornerNet<sup>[18]</sup>等。

随着深度学习的广泛应用,越来越多的研究者将其应用到遥感图像舰船检测领域, Yang 等<sup>[19]</sup>提出了自上而下的密集连接特征金字塔来融合多尺度特征并且设计特征对齐模块来提高舰船检测精度;成倩等<sup>[20]</sup>对聚合网络进行改进,同时使用 Swish 函数作为激活函数,提高网络对数据非线性特征的表征能力,提高模型对尺度变化大的数

据集的泛化能力; Ran 等<sup>[21]</sup>使用复平面代替传统平面角度解决了旋转检测框边界不连续问题;王昌安等<sup>[22]</sup>引入生成对抗网络增强细粒度的类别间判别能力,有效地提升了舰船目标的细粒度识别精度。尽管上述算法取得了不错的效果,提高了检测精度并解决了预测框与目标匹配不准确问题,但是在近岸舰船目标密集排列等复杂场景中,由于岸上车辆、集装箱等背景噪声的影响,模型不能够有效提取舰船目标特征,仍会存在目标的误检、漏检问题,检测算法精度仍然有待提高。针对以上问题,本文以两阶段目标检测算法为基础,提出了一种基于特征重用和膨胀卷积<sup>[23]</sup>的遥感图像舰船目标检测算法,本文的主要工作如下:1)使用多分支和通道注意力结合的拆分注意力机制构造新的残差块进行特征提取,同时嵌入可变形卷积网络(deformable convolution network, DCN)<sup>[24]</sup>提取更加符合目标形状的特征图。2)针对特征图通道数骤减造成的信息损失问题,构造了多尺度感受野(multi scale receptive field, MSRF)模块,使用不同膨胀率的膨胀卷积并行提取多感受野特征。3)针对原始特征金字塔特征提取不充分问题,在原始特征金字塔基础上构建了一条新的自底向上的特征重用聚合路径(feature reuse aggregation path, FRAP),提高特征提取能力。

## 1 理论基础

两阶段目标检测算法通过在特征图上均匀生成大量不同尺度和长宽比的候选框,然后对候选框进行分类和回归,从而得到最终的检测结果,因此两阶段检测算法精度较高。以 Oriented R-CNN 算法为例,整体结构由主干网络、特征金字塔网络(feature pyramid network, FPN)、方向区域生成网络(oriented region proposal network, Oriented RPN)和 Oriented R-CNN Head 4 部分组成。首先,输入图像经过主干网络进行特征提取,得到提取后的特征图,深层特征具有更强的语义信息,适用于目标分类,浅层特征具有更高的分辨率,适合候选框的回归,FPN 通过简单的自上而下路径将深层语义信息和浅层细节信息特征进行融合,以更好地适应大小不同的目标。

为了减少计算资源的消耗、提高检测效率, Oriented RPN 首先以  $[1:2, 1:1, 2:1]$  不同长宽比生成不同尺度的水平锚框,每组锚框用  $\mathbf{a} = (a_x, a_y, a_w, a_h)$  表示,其中  $(a_x, a_y)$  表示锚框的中心点坐标,  $a_w$  和  $a_h$  分别代表锚框的宽和高,使用  $1 \times 1$  卷积的回归分支生成锚框的 6 个偏移量  $\mathbf{J} = (\theta_x, \theta_y, \theta_w, \theta_h, \theta_a, \theta_\beta)$ ,接着对回归结果进行解码得到旋转框,公式表示为

$$\begin{cases} x = \vartheta_x \cdot a_w + a_x, y = \vartheta_y \cdot a_h + a_y \\ w = a_w \cdot e^{\vartheta_w}, h = a_h \cdot e^{\vartheta_h} \\ \Delta\alpha = \vartheta_\alpha \cdot w, \Delta\beta = \vartheta_\beta \cdot h \end{cases}$$

式中:  $(x,y)$ 表示预测框的中心点坐标,  $w$ 和 $h$ 分别代表预测框的宽和高,  $\Delta\alpha$ 和 $\Delta\beta$ 分别表示预测框顶点相对于外接矩形上侧和右侧中心点的偏差, 坐标偏差可视化如图 1 所示。  $(\Delta\alpha, 0)$ 表示旋转预测框顶点  $P_1$  相对于水平框上侧中心点的偏移量,  $(0, \Delta\beta)$ 表示旋转预测框顶点  $P_2$  相对于水平框右侧中心点的偏移量, 计算旋转检测框的 4 个顶点坐标公式为

$$\begin{cases} P_1 = (x, y - h/2) + (\Delta\alpha, 0) \\ P_2 = (x + w/2, y) + (0, \Delta\beta) \\ P_3 = (x, y + h/2) + (-\Delta\alpha, 0) \\ P_4 = (x - w/2, y) + (0, -\Delta\beta) \end{cases}$$

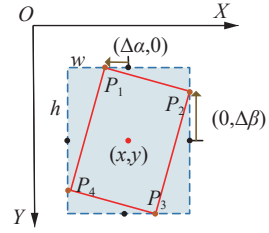


图 1 坐标偏差可视化

Fig. 1 Coordinate deviation visualization

RPN 输出特征图经过 Oriented R-CNN Head 的感兴趣区域对齐 (regions of interest align, ROI-Align) 模块, 将目标特征转化为固定大小区域, 然后连接全连接层, 最后由分类和回归分支对预测框进行分类和坐标回归调整, 得出最终的预测框。本文算法整体结构如图 2 所示。

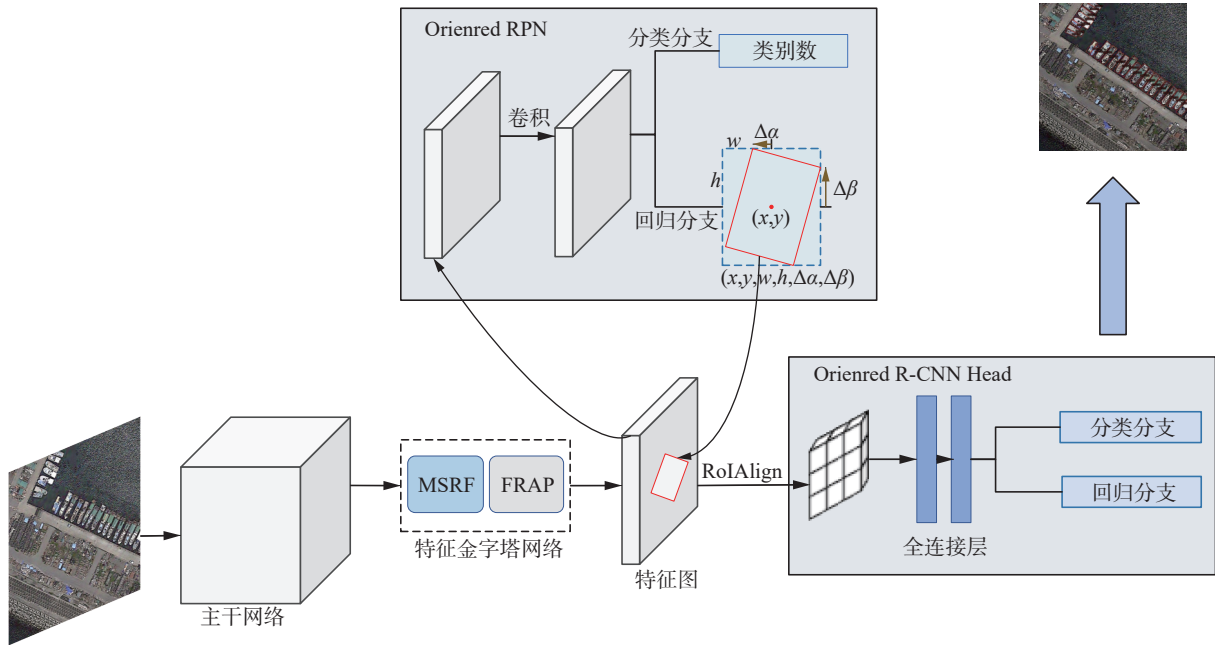


图 2 算法整体结构

Fig. 2 Overall structure of the algorithm

## 2 改进算法

### 2.1 主干网络 ResNeSt

ResNet50<sup>[25]</sup>已经广泛应用于各类计算机视觉任务, 并且表现出不错的性能, 但是在大规模、复杂图像数据场景中, 提取目标特征能力有限。ResNeSt50<sup>[26]</sup>在 ResNet50 基础上使用拆分注意力和分组卷积技术, 在保持计算效率的同时提高了模型精度和泛化能力。ResNet 残差块只使用简单的卷积和残差结构, 特征提取能力有限, 而 ResNeSt 残差块对特征图进行分组卷积, 并且使用拆分通道注意力, 在减少计算量的同时提高特征提取能力。

残差块结构如图 3 所示。首先将输入通道数

为  $c$ , 尺寸为  $h \times w$  的特征图平均划分为  $k$  个基组, 每个基组又均分为  $r$  个拆分单元, 一共有  $G = k \times r$  个拆分单元, 每个单元的通道数为  $c/(k \times r)$ , 特征图经过拆分单元卷积通道数变为  $c/k$ , 基组内经过拆分单元的特征图由拆分注意力进行合并, 最后将基组输出的特征图在通道维度进行拼接, 经过  $1 \times 1$  卷积融合特征后输出。在拆分单元中嵌入  $3 \times 3$  可变形卷积, 传统卷积每个卷积核的采样位置是固定的, 对于存在形变情况的目标, 规则卷积可能会引入噪声, 影响模型准确性, 而可变形卷积通过引入可变形卷积核, 更好地适应不同形状的目标, 提高模型的空间不变性。可变形卷积设计如图 4 所示。可变形卷积在每个位置生成 1 个偏移量, 用于控制卷积核在特征图上进行



可变形变换, 从而适应不同形状和尺度的目标物体。偏移量可以通过一个额外的可学习的偏移网

络进行计算, 图 4 中偏移特征通道维数为  $2N$ , “2” 分别对应  $X$  和  $Y$  2 个二维偏移,  $N$  是通道数。

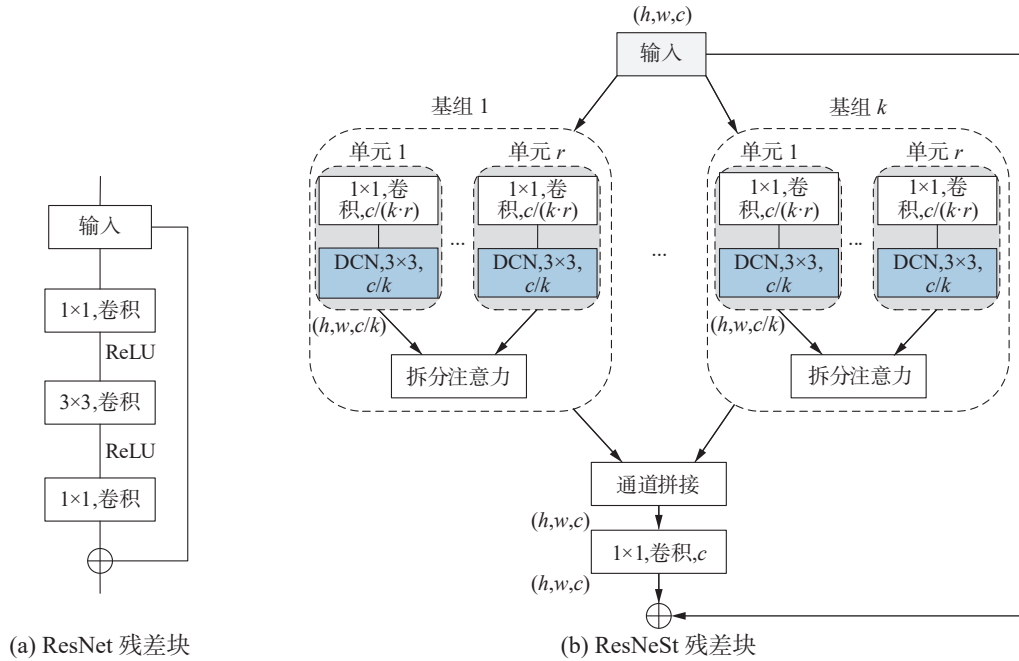


图 3 残差块对比

Fig. 3 Residual block comparison

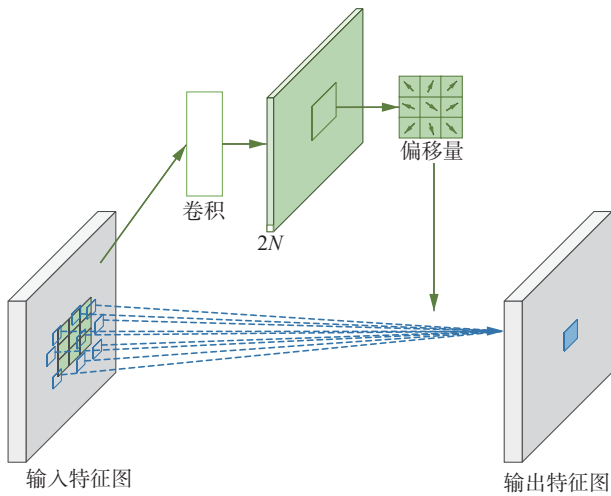


图 4 可变形卷积

Fig. 4 Deformable convolution

常规卷积分为 2 步: 1) 使用尺寸为  $c \times c$  的卷积核在特征图  $x$  上采样, 卷积核尺寸定义感受野大小和扩张; 2) 将采样范围内的像素与卷积核参数  $w$  进行加权求和。则卷积后特征图  $y$  上的任意像素位置值为

$$y(p_0) = \sum_{n=1}^{c \times c} w(p_n) \cdot x(p_0 + p_n)$$

式中  $p_n$  为卷积核中的位置向量。

可变形卷积在常规卷积计算公式上加了位置偏移向量  $\Delta p_n$ , 其中,  $\{\Delta p_n | n = 1, 2, \dots, N\}$ ,  $N = c \times c$ ,

可变形卷积计算公式为

$$y(p_0) = \sum_{n=1}^{c \times c} w(p_n) \cdot x(p_0 + p_n + \Delta p_n)$$

## 2.2 多尺度感受野模块

在深度神经网络中, 浅层特征具有更高的分辨率和丰富的细节信息, 但是缺乏语义信息; 深层特征具有更丰富的语义信息, 但是分辨率较低。在检测任务中, 通常使用特征金字塔结构来融合不同层次的特征信息, 以得到既具有丰富语义信息又保留高分辨率细节的特征表示。不同层次的特征通常具有不同的通道数, 为了将它们融合在一起, 会使用  $1 \times 1$  卷积将通道数进行统一, 但是这种通道数的统一可能会造成信息损失。针对此问题, 本文构造了多尺度感受野模块, 替换原有  $1 \times 1$  横向连接, 以减少信息损失。首先使用不同膨胀率的卷积并行提取多感受野特征, 进而提高特征多样性; 然后使用注意力机制强化舰船目标特征, 同时抑制噪声背景的干扰, 减少噪声和冗余。

本文 MSRF 结构设计如图 5 所示。遥感图像中的舰船目标存在尺度变化大等特点, MSRF 模块首先对  $C_i (i=2, 3, 4, 5)$  特征并行使用膨胀率为  $(5, 9, 13)$  的  $3 \times 3$  膨胀卷积提取多尺度、多感受野特征, 同时将  $C_i$  特征经过自适应平均池化和  $1 \times 1$  卷积来减少通道数, 然后与多感受野特征进行通道

维度上的拼接,接着使用卷积块注意力模块 (convolutional block attention module, CBAM)<sup>[27]</sup> 提高其

性能和泛化能力,最后将通过注意力机制的特征图和  $C_i$  特征图相加,减少信息损失。

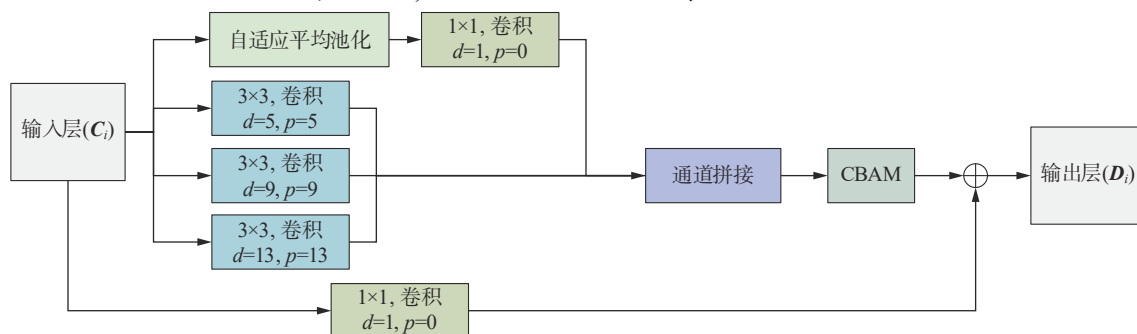


图 5 MSRF 结构

Fig. 5 MSRF structure

CBAM 结构设计如图 6 所示, CBAM 是由通道注意力和空间注意力结合来增强特征图的表示能力,通道注意力通过学习每个通道的重要性加权和,加强了特征图的通道关联性,以提

高特征的表达能力,空间注意力模块通过学习每个空间位置的重要性加权和,增强了特征图的空间位置关联性,以提高特征的局部感知能力。

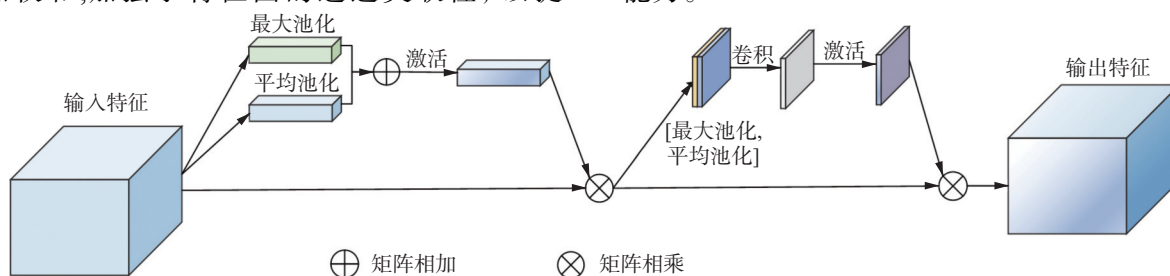


图 6 CBAM 模块

Fig. 6 CBAM module

### 2.3 特征重用聚合网络

主干网络会生成一组分辨率不同的特征图,为了使深层特征具有位置信息,浅层特征具有足够的语义信息,FPN 采用了自上而下的结构来优化网络,这样使得深层特征图的信息向低层特征进行传递,从而增强浅层特征的语义信息。然而,由于经过多次下采样,深层特征的语义信息

虽然很丰富,但是可能丢失小目标的语义信息,同时主干网络中生成的特征图拥有丰富的目标特征,但是未被重复利用。本文构建特征重用聚合路径,具体来说,在原有特征金字塔路径基础上,将原始特征进行重用,构建了自底向上的聚合路径,提高特征提取能力。FRAP 结构如图 7 所示。

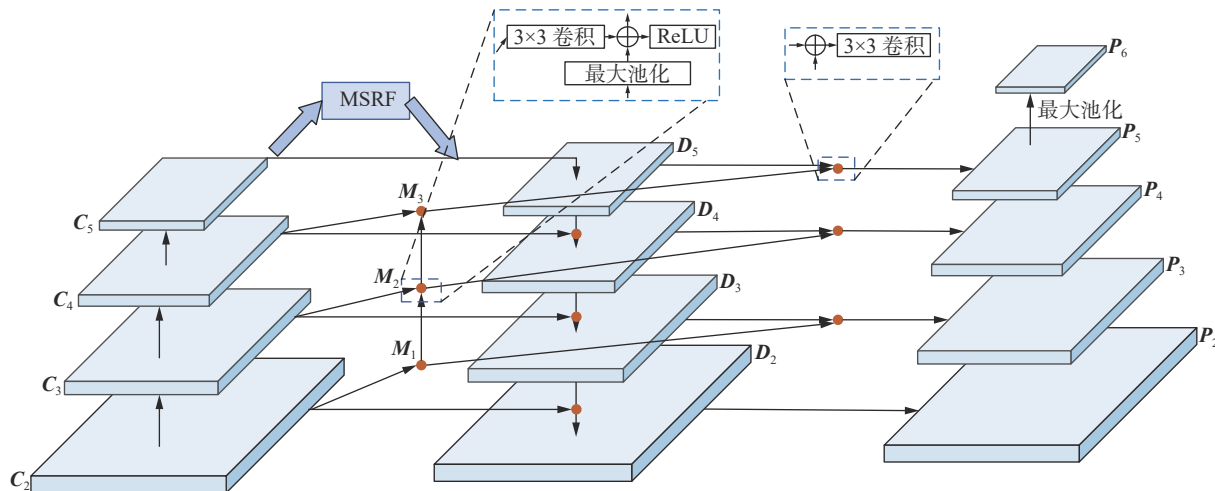


图 7 FRAP 结构

Fig. 7 Structure of FRAP

图像输入主干网络后得到特征图  $\{C_2, C_3, C_4, C_5\}$ , 接着, 这些特征图经过 MSRF 模块和自顶向下路径后得到特征图  $\{D_2, D_3, D_4, D_5\}$ 。其中,  $C_2$  经过  $3 \times 3$  卷积后, 特征图尺寸减半, 得到中间特征图  $M_1$ ,  $M_1$  与  $D_3$  层特征相加后经过  $3 \times 3$  卷积生成最终特征图  $P_3$ ; 同样将  $C_3$  经过  $3 \times 3$  卷积后与经过最大池化的  $M_1$  相加, 然后经过 ReLU 激活函数后再与  $D_4$  相加, 经过  $3 \times 3$  卷积后得到最终特征图  $P_4$ , 以此类推, 得到最终特征图  $P_5$ ,  $P_5$  层经过最大池化得到  $P_6$  层,  $P_2$  层特征图直接由  $D_2$  层得到。此聚合路径不仅继承了 FPN 的优点, 同时将主干特征图进行重用, 使得每层特征不仅直接融合了上一层特征信息, 还直接融合了下一层特征信息, 从而提高了网络的特征提取

能力。

### 3 实验结果与分析

#### 3.1 数据集

DOTA 数据集<sup>[28]</sup> 是一个大型遥感检测数据集, 收集了 Google Earth 和多个卫星平台的图像, 数据集包含了 16 个类别和 40 万个标注的实例对象, 包括了密集排列、任意方向的目标, 每个图片像素在  $800 \times 800 \sim 4\,000 \times 4\,000$ 。鉴于本文检测算法的特殊性, 只提取了舰船 1 个类别进行检测, 考虑到硬件的限制, 输入尺寸较大时难以训练, 将图片切割为 800 像素  $\times$  800 像素, 重叠像素为 200, 最终得到 3 868 张训练数据, 967 张测试数据。部分数据集实例如图 8(a) 所示。

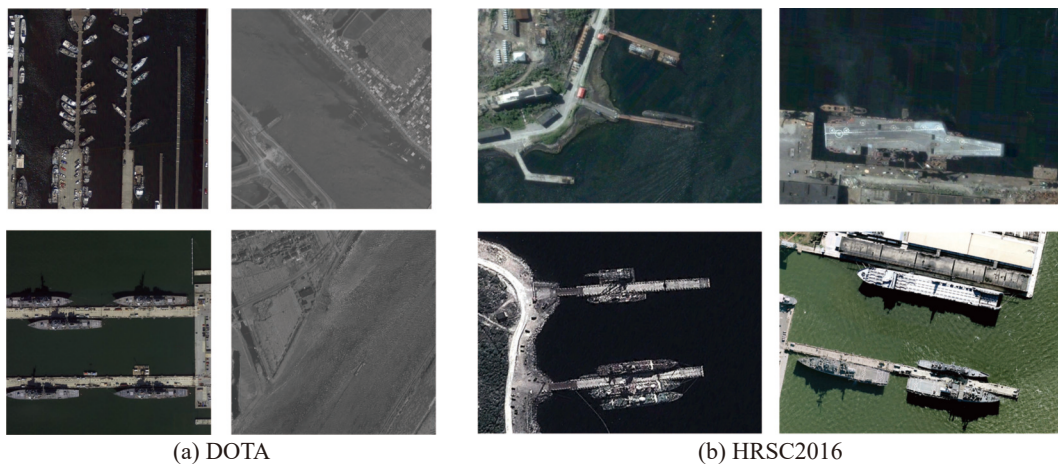


图 8 数据集可视化

Fig. 8 Dataset visualization

HRSC2016 数据集<sup>[29]</sup> 是由中国科学院自动化研究所发布的专门用于舰船检测遥感图像数据集, 该数据集涵盖了不同光照、天气条件下的舰船图像, 这些图像覆盖了不同的舰船尺度、类型和舰船姿态, 图像尺寸在  $300 \times 300 \sim 1\,500 \times 900$ , 分辨率介于  $0.4 \sim 2$  m。该数据集共有 1 061 张标注图像, 其中 617 张图像用于训练, 444 张图像用于测试, 实验中将图像统一调整为 800 像素  $\times$  512 像素。部分数据集实例如图 8(b) 所示。

#### 3.2 实验平台与参数设置

本次实验所使用的平台为: Dell Precision 7920 工作站, Inter(R) Xeon(R) Silver 4216 处理器, 128 GB 内存, GeForce RTX 2080Ti 显卡, 软件环境为: Windows10 操作系统, PyTorch 1.8.0、MMRotate 0.3.3、Python 3.8、CUDA 10.2 和 cuDNN 11。训练过程中使用 MMRotate 框架<sup>[30]</sup> 中的默认参数, 残差块基数  $k$  设置为 1, 拆分单元  $r$  设置为 2, 批次大小设置为 2, 采用 SGD 优化器进行优化, 动量设置为 0.9, 衰减权重为 0.000 1, 初始学习率为 0.005。

为了加速收敛, 所有实验使用在 ImageNet 数据集<sup>[31]</sup> 上的预训练权重。训练一共设置 12 个 epoch, 在第 8 个和第 11 个 epoch 时学习率乘以 0.1。训练过程中, 使用水平翻转、垂直翻转和对角翻转来增加数据多样性, 提高模型的泛化能力和鲁棒性。

从图 9 所示的训练损失函数曲线可以看出, 本文方法具有更低的损失, 说明模型可以更准确地识别出目标类别, 并更准确地预测出目标框的位置。

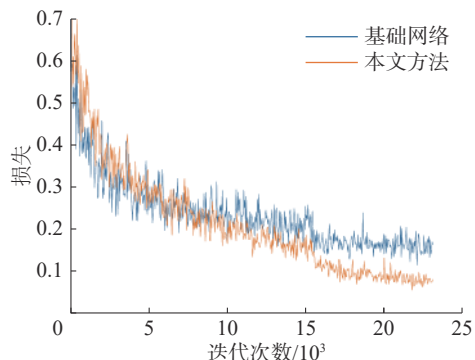


图 9 损失函数

Fig. 9 Loss function

### 3.3 评价指标

平均精度 (average precision,  $P^{AP}$ ) 已经被用于各类目标检测的评估, 其是由召回率 (recall,  $R$ ) 和精确率 (precision,  $P$ ) 计算得出的, 召回率是指在所有真实目标中检测正确的比例, 精确率是指检测出的目标中真实目标所占的比例, 召回率和精确率的计算公式为

$$P = \frac{N_{TP}}{N_{TP} + N_{FP}}$$

$$R = \frac{N_{TP}}{N_{TP} + N_{FN}}$$

式中:  $N_{TP}$  表示真正例 (true positive), 即被正确检测出来的目标数量,  $N_{FP}$  表示假正例 (false positive), 即被错误检测出来的目标数量,  $N_{FN}$  表示假负例 (false negative), 即未被检测出来的目标数量。

$P^{AP}$  是由 PR 曲线下方的面积计算得出, 计算公式为

$$P^{AP} = \int_0^1 P(R) dR$$

$P^{AP}$  通过不同的阈值细分为:  $P_{50}^{AP}$ 、 $P_{75}^{AP}$ 、 $P_{50:95}^{AP}$ 。

$P_{50}^{AP}$  和  $P_{75}^{AP}$  分别表示预测框和真实框的交并比 (intersection over union, IOU) 阈值取 0.5 和 0.75 时的精度,  $P_{50:95}^{AP}$  表示 IOU 阈值从 0.5 取到 0.95, 步长为 0.05 下的精度均值。IOU 阈值用来判断预测框和真实目标框的匹配程度, IOU 阈值越高, 预测框和真实目标框的重叠度就越高,  $P_{75}^{AP}$  指标需要检测方法在更高的 IOU 阈值下表现良好, 因此对于算法的准确性和稳定性提出更高的要求,  $P_{50:95}^{AP}$  是在不同阈值下对检测方法进行全面的评估, 从而更准确地反映方法的性能。

### 3.4 对比实验

为了验证本文所提方法的有效性, 与两阶段检测算法 Rotated Faster R-CNN<sup>[32]</sup>、ReDet<sup>[10]</sup>、ROI Transformer<sup>[11]</sup> 和一阶段检测算法 Sasm\_reppoints<sup>[17]</sup>、S<sup>2</sup>A-net<sup>[14]</sup> 等主流检测算法进行比较, 不同算法在 DOTA 数据集上的检测性能见表 1。在速度损耗有限的情况下, 本文方法保持较高的检测精度,  $P_{50}^{AP}$ 、 $P_{75}^{AP}$  和  $P_{50:95}^{AP}$  分别达到 91.95%、77.94% 和 62.08%。

表 1 不同检测方法在 DOTA 数据集上对比

Table 1 Comparison of different detection methods on the DOTA dataset

模型	输入大小	$P_{50}^{AP}/\%$	$P_{75}^{AP}/\%$	$P_{50:95}^{AP}/\%$	检测速度/(f/s)
Rotated Faster R-CNN	(800,800)	87.37	45.62	46.82	<b>18.3</b>
ReDet	(800,800)	85.14	47.08	46.69	8.0
ROI Transformer	(800,800)	89.80	53.90	51.07	15.1
Sasm_reppoints	(800,800)	87.42	45.01	46.62	17.2
S <sup>2</sup> A-net	(800,800)	88.47	37.66	44.70	17.4
Oriented R-CNN	(800,800)	89.83	54.88	51.26	16.4
本文方法	(800,800)	<b>91.95</b>	<b>77.94</b>	<b>62.08</b>	12.2

注: 加粗字体为本列最优结果。

为了进一步验证本文所提算法的泛化性和扩展性, 本文还将所提算法与其他几种算法在 HR-SC2016 专门舰船检测数据集上进行了对比, 实验

结果见表 2。从表 2 中可以看出, 本文所提算法依然保持较高的检测精度,  $P_{50}^{AP}$ 、 $P_{75}^{AP}$  和  $P_{50:95}^{AP}$  分别达到 93.13%、73.87% 和 61.27%。

表 2 不同检测方法在 HRSC2016 数据集上对比

Table 2 Comparison of different detection methods on the HRSC2016 dataset

模型	输入大小	$P_{50}^{AP}/\%$	$P_{75}^{AP}/\%$	$P_{50:95}^{AP}/\%$	检测速度/(f/s)
Rotated Faster R-CNN	(800,512)	71.00	14.20	26.94	24.3
ReDet	(800,512)	79.80	60.00	50.84	9.4
ROI Transformer	(800,512)	88.20	40.90	45.75	18.5
Sasm_reppoints	(800,512)	92.10	70.30	59.40	<b>27.6</b>
S <sup>2</sup> A-net	(800,512)	92.00	57.50	53.09	17.3
Oriented R-CNN	(800,512)	86.90	55.30	49.55	22.7
本文方法	(800,512)	<b>93.13</b>	<b>73.87</b>	<b>61.27</b>	16.5

注: 加粗字体为本列最优结果。

PR 曲线可以更好地体现召回率和准确率的关系, 曲线越接近 (1,1) 位置, 曲线下的面积越大,

说明模型检测性能越好。6 种检测算法在 2 个数据集上 IOU 阈值为 0.5 的 PR 曲线如图 10 所示,



在 DOTA 数据集上相比其他算法,本文算法在同一召回率下拥有较高准确率,在同一准确率下,召回率也是最高的,在 HRSC2016 数据集上,虽然

本文算法在准确率接近 1 时,召回率并未达到最高,但是整体曲线下的面积最大,模型整体性能较好。

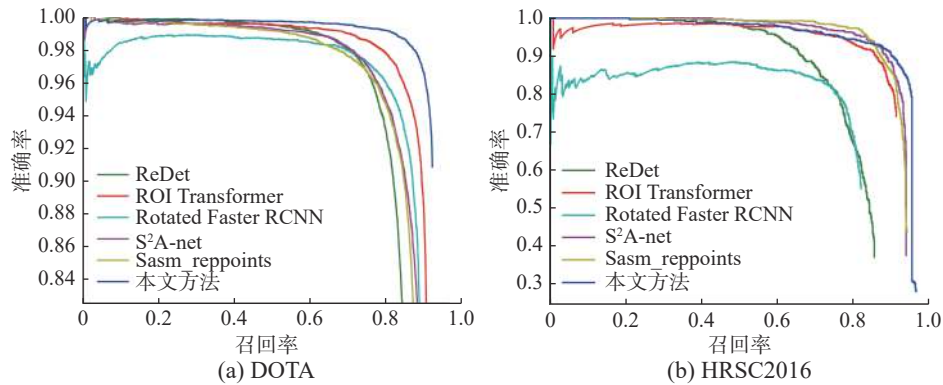


图 10 不同模型的 PR 曲线

Fig. 10 PR curves of different models

### 3.5 消融实验

为了证明本文所提模块的有效性,以 Oriented R-CNN 为基础网络,在 DOTA 数据集上进行消融实验,设置了 5 组实验,使用  $P_{50}^{AP}$ 、 $P_{75}^{AP}$ 、 $P_{50:95}^{AP}$ 、检测

速度、参数量和浮点运算次数 (floating point of operations, FLOPs) 作为评价指标,为公平地比较实验结果,所有参数均保持一致,消融实验结果见表 3。

表 3 消融实验结果

Table 3 Ablation experiment results

模型	输入大小	$P_{50}^{AP}/\%$	$P_{75}^{AP}/\%$	$P_{50:95}^{AP}/\%$	检测速度/(f/s)	参数量/ $10^6$	FLOPs
基础网络	(800,800)	89.83	54.88	51.26	<b>18.4</b>	41.13	121.58
+ ResNeSt50+DCN	(800,800)	91.47	74.85	60.06	15.4	43.01	128.06
+ MSRF	(800,800)	90.99	74.12	59.59	13.8	48.10	138.06
+ FRAP	(800,800)	90.36	72.96	58.61	16.3	45.13	132.66
本文方法	(800,800)	<b>91.95</b>	<b>77.94</b>	<b>62.08</b>	12.2	53.98	155.62

注:加粗字体为本列最优结果。

基础网络使用常规卷积提取舰船特征,容易引入背景噪声,无法提取有效的、符合舰船目标的特征,同时,基础网络没有将主干特征图进行重用,导致检测结果中的精度普遍较低,因此  $P_{75}^{AP}$  和  $P_{50:95}^{AP}$  较低,只有 54.88% 和 51.26%。+ResNeSt50+DCN 通过引入分组卷积和拆分注意力,同时使用可变形卷积提取更加符合舰船的特征,减少噪声和冗余,将  $P_{50}^{AP}$ 、 $P_{75}^{AP}$  和  $P_{50:95}^{AP}$  分别提升至 91.47%、74.85% 和 60.06%。MSRF 模块通过并行膨胀卷积提取多尺度特征,减少信息损失,将  $P_{50}^{AP}$ 、 $P_{75}^{AP}$  和  $P_{50:95}^{AP}$  分别提升至 90.99%、74.12% 和 59.59%; FRAP 通过构建自底向上的聚合路径将主干特征进行重用,提升了特征表达能力,将  $P_{50}^{AP}$ 、 $P_{75}^{AP}$  和  $P_{50:95}^{AP}$  分别提升至 90.36%、72.96% 和 58.61%,本文方法在基础网络上引入 ResNeSt50

和 DCN,同时整合 MSRF 和 FRAP,进一步提升了算法的检测性能,最终  $P_{50}^{AP}$ 、 $P_{75}^{AP}$  和  $P_{50:95}^{AP}$  分别达到了 91.95%、77.94% 和 62.08%。

由于在模型中引入了额外的卷积运算,因此模型的参数量和计算复杂度有所增加,但相较于模型精度的提升,额外付出的代价是可以接受的。

为了定性分析本文所提算法的特征提取能力,对特征金字塔输出的前 12 个通道的特征图进行可视化,可视化特征图如图 11 所示,图 11(a) 中基础网络输出的特征图目标特征比较模糊,轮廓不清晰,图 11(b) 中本文算法输出的特征图相比基础网络输出的特征图目标区域更加明显,可以更准确地识别目标所在位置和边界,同时目标与背景的差异更大,可以更好地区分目标和背景,从而更准确地识别目标的类别。

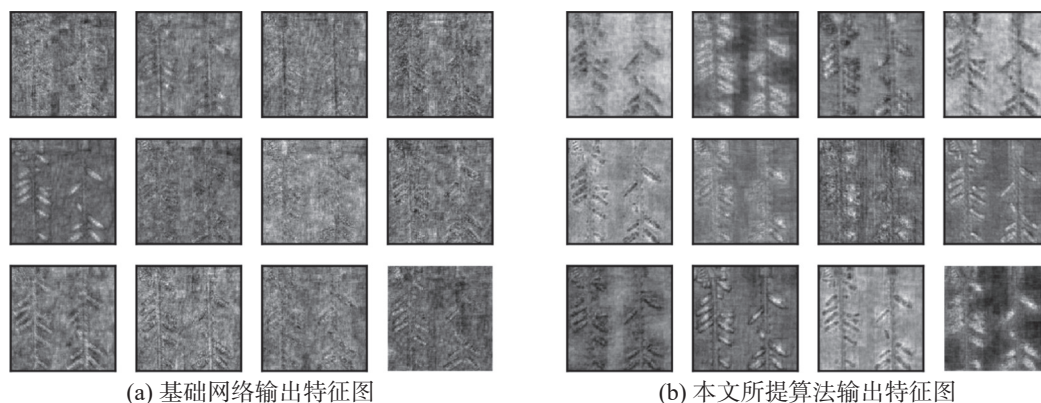


图 11 特征可视化

Fig. 11 Feature visualization

检测结果如图 12 所示, 从检测结果可以看出, 基础网络 Oriented R-CNN 算法在复杂背景下有误检、漏检, 黄色标记框内存在漏检舰船, 绿色标记框内将车辆误检为舰船, 而本文算法通过使

用嵌入可变形卷积的主干网络和构造多尺度感受野模块提取更有效的符合舰船目标的特征, 能够有效解决误检、漏检问题, 提高检测效果, 图 12 蓝色框内排除车辆误检, 并正确检测到舰船目标。

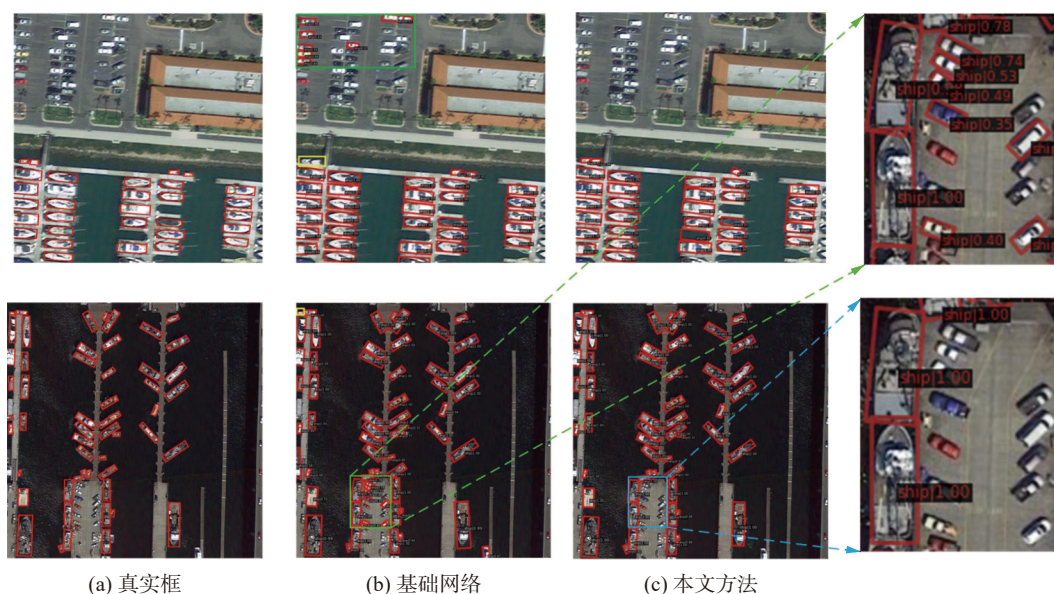


图 12 检测结果可视化

Fig. 12 Visualization of detection results

## 4 结束语

针对目标检测算法中对舰船小目标错检、漏检问题, 本文提出了一种基于特征重用和膨胀卷积的遥感图像舰船目标检测算法, 该算法使用具有更强大特征提取能力的主干网络, 并通过特征重用来提高特征提取能力, 另外, 还构造了多尺度感受野模块以减少特征在传递过程中因通道数骤减造成的信息损失。本文所提模块的协同合作使得目标检测精度明显提高, 最后在 DOTA 和 HRSC2016 数据集上  $P_{50}^{AP}$  分别达到 91.95% 和 93.13%, 虽然成功提高了模型的检测精度, 但也意味着需要更多的时间和计算资源来完成任务, 下一步将

进一步优化模型的结构和算法, 尝试在保持高精度的同时提高模型的处理速度。

## 参考文献:

- [1] 龚声蓉, 徐少杰, 周立凡, 等. 融入混合注意力的可变形空洞卷积近岸 SAR 小舰船检测 [J]. 中国图象图形学报, 2022, 27(12): 3663–3676.  
GONG Shengrong, XU Shaojie, ZHOU Lifan, et al. Deformable atrous convolution nearshore SAR small ship detection incorporating mixed attention[J]. Journal of image and graphics, 2022, 27(12): 3663–3676.
- [2] 黎纪元, 厉小润, 赵辽英. 基于边缘线分析与聚合通道特征的港口舰船检测 [J]. 光学学报, 2019, 39(8):

- 217–226.
- LI Jingyuan, LI Xiaorun, ZHAO Liaoying. Docked ship detection based on edge line analysis and aggregation channel features[J]. *Acta optica sinica*, 2019, 39(8): 217–226.
- [3] VIOLA P, JONES M. Rapid object detection using a boosted cascade of simple features[C]//Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Kauai: IEEE, 2001: 511–518.
- [4] 林鹏宇, 马晓珊, 彭晓东. 基于仿真模板和 SuperGlue 的舰船匹配检测 [J]. *计算机仿真*, 2023, 40(11): 11–15.
- LIN Pengyu, MA Xiaoshan, PENG Xiaodong. Ship target match method based on simulation template and SuperGlue feature matching[J]. *Computer simulation*, 2023, 40(11): 11–15.
- [5] STAUFFER C, GRIMSON W E L. Adaptive background mixture models for real-time tracking[C]//Proceedings of 1999 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Fort Collins: IEEE, 1999: 246–252.
- [6] 成艳, 孟祥勇, 于雪莲, 等. 海上舰船目标轮廓检测算法 [J]. *智能计算机与应用*, 2022, 12(8): 119–122.
- CHENG Yan, MENG Xiangyong, YU Xuelian, et al. Target contour detection algorithm for marine ships[J]. *Intelligent computer and applications*, 2022, 12(8): 119–122.
- [7] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. ImageNet classification with deep convolutional neural networks[J]. *Communications of the ACM*, 2017, 60(6): 84–90.
- [8] 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.
- [9] XIE Xingxing, CHENG Gong, WANG Jiabao, et al. Oriented R-CNN for object detection[C]//2021 IEEE/CVF International Conference on Computer Vision. Montreal: IEEE, 2021: 3500–3509.
- [10] HAN Jiaming, DING Jian, XUE Nan, et al. ReDet: a rotation-equivariant detector for aerial object detection[C]//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Nashville: IEEE, 2021: 2785–2794.
- [11] DING Jian, XUE Nan, LONG Yang, et al. Learning RoI transformer for oriented object detection in aerial images[C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach: IEEE, 2019: 2844–2853.
- [12] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: unified, real-time object detection[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas: IEEE, 2016: 779–788.
- [13] LIN T Y, GOYAL P, GIRSHICK R, et al. Focal loss for dense object detection[C]//2017 IEEE International Conference on Computer Vision. Venice: IEEE, 2017: 2999–3007.
- [14] HAN Jiaming, DING Jian, LI Jie, et al. Align deep features for oriented object detection[J]. *IEEE transactions on geoscience and remote sensing*, 2021, 60: 1–11.
- [15] 王慧赢, 王春平, 付强, 等. 面向嵌入式平台的轻量级光学遥感图像舰船检测 [J]. *光学学报*, 2023, 43(12): 121–134.
- WANG Huiying, WANG Chunping, FU Qiang, et al. Lightweight ship detection based on optical remote sensing images for embedded platform[J]. *Acta optica sinica*, 2023, 43(12): 121–134.
- [16] TIAN Zhi, SHEN Chunhua, CHEN Hao, et al. FCOS: fully convolutional one-stage object detection[C]//2019 IEEE/CVF International Conference on Computer Vision. Seoul: IEEE, 2019: 9626–9635.
- [17] YANG Ze, LIU Shaohui, HU Han, et al. RepPoints: point set representation for object detection[C]//2019 IEEE/CVF International Conference on Computer Vision. Seoul: IEEE, 2019: 9656–9665.
- [18] LAW H, DENG Jia. CornerNet: detecting objects as paired keypoints[C]//European Conference on Computer Vision. Cham: Springer, 2018: 765–781.
- [19] YANG Xue, SUN Hao, FU Kun, et al. Automatic ship detection in remote sensing images from google earth of complex scenes based on multiscale rotation dense feature pyramid networks[J]. *Remote sensing*, 2018, 10(1): 132.
- [20] 成倩, 李佳, 杜娟. 基于 YOLOv5 的光学遥感图像舰船目标检测算法 [J]. *系统工程与电子技术*, 2023, 45(5): 1270–1276.
- CHENG Qian, LI Jia, DU Juan. Ship target detection algorithm of optical remote sensing image based on YOLOv5 [J]. *Systems engineering and electronics*, 2023, 45(5): 1270–1276.
- [21] RAN Bohao, YOU Yanan, LI Zezhong, et al. Arbitrary-oriented ship detection method based on improved regression model for target direction detection network[C]//IGARSS 2020 - 2020 IEEE International Geoscience and Remote Sensing Symposium. Waikoloa: IEEE, 2020: 964–967.
- [22] 王昌安, 田金文. 生成对抗网络辅助学习的舰船目标精细识别 [J]. *智能系统学报*, 2020, 15(2): 296–301.
- WANG Chang'an, TIAN Jinwen. Fine-grained inshore ship recognition assisted by deep-learning generative adversarial networks[J]. *CAAI transactions on intelligent*



- systems, 2020, 15(2): 296–301.
- [23] WANG Panqu, CHEN Pengfei, YUAN Ye, et al. Understanding convolution for semantic segmentation[C]//2018 IEEE Winter Conference on Applications of Computer Vision. Lake Tahoe: IEEE, 2018: 1451–1460.
- [24] DAI Jifeng, QI Haozhi, XIONG Yuwen, et al. Deformable convolutional networks[C]//2017 IEEE International Conference on Computer Vision. Venice: IEEE, 2017: 764–773.
- [25] HE Kaiming, ZHANG Xiangyu, REN Shaoqing, et al. Deep residual learning for image recognition[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas: IEEE, 2016: 770–778.
- [26] ZHANG Hang, WU Chongruo, ZHANG Zhongyue, et al. ResNeSt: split-attention networks[C]//2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops. New Orleans: IEEE, 2022: 2735–2745.
- [27] WOO S, PARK J, LEE J Y, et al. CBAM: convolutional block attention module[C]//European Conference on Computer Vision. Cham: Springer, 2018: 3–19.
- [28] XIA Guisong, BAI Xiang, DING Jian, et al. DOTA: a large-scale dataset for object detection in aerial images[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City: IEEE, 2018: 3974–3983.
- [29] LIU Zikun, YUAN Liu, WENG Lubin, et al. A high resolution optical satellite image dataset for ship recognition and some new baselines[C]//Proceedings of the 6th International Conference on Pattern Recognition Applications and Methods. Porto: SCITEPRESS-Science and Technology Publications, 2017: 324–331.
- [30] ZHOU Yue, YANG Xue, ZHANG Gefan, et al. MMRotate: a rotated object detection benchmark using PyTorch[C]//Proceedings of the 30th ACM International Conference on Multimedia. Lisboa Portugal: ACM, 2022: 7331–7334.
- [31] DENG Jia, DONG Wei, SOCHER R, et al. ImageNet: a large-scale hierarchical image database[C]//2009 IEEE Conference on Computer Vision and Pattern Recognition. Miami: IEEE, 2009: 248–255.
- [32] YANG Sheng, PEI Ziqiang, ZHOU Feng, et al. Rotated faster R-CNN for oriented object detection in aerial images[C]//Proceedings of the 2020 3rd International Conference on Robot Systems and Applications. Chengdu: ACM, 2020: 35–39.

### 作者简介:



曲海成, 副教授, 博士, 辽宁工程技术大学软件学院/人工智能学院副院长, 主要研究方向为视觉信息计算。主持辽宁省自然科学基金项目 1 项、辽宁省教育厅面上项目 2 项, 发表学术论文 60 余篇。E-mail: [quhaicheng@lntu.edu.cn](mailto:quhaicheng@lntu.edu.cn)。



李瑞柯, 硕士研究生, 主要研究方向为遥感图像目标检测。E-mail: [lrk-19990101@163.com](mailto:lrk-19990101@163.com)。



王蒙, 硕士研究生, 主要研究方向为遥感图像目标检测。E-mail: [1377423034@qq.com](mailto:1377423034@qq.com)。