

## 基于语义特征的遥感舰船关重部位检测网络

张冬冬 王春平 付强

### Detection network of critical parts for remote sensing ship based on semantic features

ZHANG Dongdong, WANG Chunping, FU Qiang

引用本文:

张冬冬,王春平,付强. 基于语义特征的遥感舰船关重部位检测网络[J]. 应用光学, 2023, 44(3): 595–604. DOI: 10.5768/JAO202344.0303004

ZHANG Dongdong, WANG Chunping, FU Qiang. Detection network of critical parts for remote sensing ship based on semantic features[J]. Journal of Applied Optics, 2023, 44(3): 595–604. DOI: 10.5768/JAO202344.0303004

在线阅读 View online: <https://doi.org/10.5768/JAO202344.0303004>

---

## 您可能感兴趣的其他文章

### Articles you may be interested in

#### [基于深度学习的无人车夜视图像语义分割](#)

Semantic segmentation of night vision images for unmanned vehicles based on deep learning

应用光学. 2017, 38(3): 421–428 <https://doi.org/10.5768/JAO201738.0302007>

#### [结合深度学习的非下采样剪切波遥感图像融合](#)

Remote sensing image fusion based on deep learning non–subsampled shearlet

应用光学. 2018, 39(5): 655–666 <https://doi.org/10.5768/JAO201839.0502001>

#### [基于深度学习的铸件X射线图像分割研究](#)

Casting X-ray image segmentation based on deep learning

应用光学. 2021, 42(6): 1025–1033 <https://doi.org/10.5768/JAO202142.0602003>

#### [基于改进DeepLabv3+的无人车夜间红外图像语义分割](#)

Semantic segmentation of nocturnal infrared images of unmanned vehicles based on improved DeepLabv3+

应用光学. 2020, 41(1): 180–185 <https://doi.org/10.5768/JAO202041.0106002>

#### [复杂背景下红外图像弱小目标检测](#)

Dim–small targets detection of infrared images in complex backgrounds

应用光学. 2021, 42(4): 643–650 <https://doi.org/10.5768/JAO202142.0402002>

#### [基于密集连接结构的超分辨率精简网络](#)

Super–resolution simplification network based on densely connected structure

应用光学. 2019, 40(5): 805–811 <https://doi.org/10.5768/JAO201940.0502003>



关注微信公众号，获得更多资讯信息

文章编号:1002-2082(2023)03-0595-10

# 基于语义特征的遥感舰船关重部位检测网络

张冬冬, 王春平, 付 强

(陆军工程大学石家庄校区 电子与光学工程系, 河北 石家庄 050003)

**摘要:** 在近岸场景中, 受背景影响, 舰船关重部位误检概率高、检测精度低。针对以上问题, 提出了一种基于语义特征的舰船关重部位检测网络, 并命名为 CPDNet (critical part detection network)。通过优化网络结构及引入注意力机制, 提升网络的特征表达能力以及对关重部位的感知能力; 基于语义信息, 设计了语义掩膜模块, 以降低背景对检测精度的影响; 增加角度参数, 使网络适用于具有方向性的目标; 构建了舰船关重部位数据集, 命名为 CP-Ship, 以验证所提网络的有效性。在 CP-Ship 数据集上的实验结果表明: 所提网络的平均精度比 RetinaNet 提高了 11.35%, 与其他网络模型相比, 其检测精度和速度均表现优异。

**关键词:** 深度学习; 遥感图像; 关重部位检测; 语义特征

中图分类号: TN206; TP753

文献标志码: A

DOI: [10.5768/JAO202344.0303004](https://doi.org/10.5768/JAO202344.0303004)

## Detection network of critical parts for remote sensing ship based on semantic features

ZHANG Dongdong, WANG Chunping, FU Qiang

(Department of Electronic and Optical Engineering, Army Engineering University of PLA in Shijiazhuang Campus, Shijiazhuang 050003, China)

**Abstract:** In near-shore scenes, under the influence of background, the probability of false detection and low detection accuracy of ship critical parts are high. To address the above problems, a detection network of ship critical parts based on semantic features was proposed, which named critical part detection network (CPDNet). Firstly, by optimizing the network structure and introducing the attention mechanism, the feature expression ability and the perception ability of the ship critical parts were improved. Secondly, based on semantic information, a semantic mask module was designed to reduce the impact of background on detection accuracy. In addition, the angle parameter was added to make the network applicable to targets with orientation. Finally, a ship critical parts dataset, named CP-Ship, was constructed to verify the effectiveness of the proposed network. The experimental results on the CP-Ship dataset show that the average accuracy of the proposed network is 11.35% higher than that of RetinaNet. Compared with other network models, the proposed network performs well in both detection accuracy and speed.

**Key words:** deep learning; remote sensing image; critical part detection; semantic features

### 引言

舰船检测是计算机视觉领域的研究热点, 一直备受国内外研究人员的关注, 其在海上交通管理、安全威胁监测、海上救援等领域发挥着重要作用<sup>[1-2]</sup>。

在深度学习技术发展的驱动下, 舰船检测的精度

和速度得到了大幅提升<sup>[3]</sup>, 已基本能满足人们的日常需求。但随着信息技术的飞速发展, 一些特殊任务, 比如精确制导, 需要对舰船做进一步的精细化检测<sup>[4-6]</sup>。

精细化检测是指对目标相对重要的部位进行

收稿日期: 2022-06-06; 修回日期: 2022-08-09

基金项目: 军内某科研项目 (LJ20191A040155)

作者简介: 张冬冬 (1993—), 男, 硕士研究生, 主要从事光电侦察情报处理、目标识别技术研究。E-mail: [1042911849@qq.com](mailto:1042911849@qq.com)

通信作者: 付强 (1981—), 男, 讲师, 博士, 主要从事智能视觉与目标检测技术研究。E-mail: [1418748495@qq.com](mailto:1418748495@qq.com)

检测,对于舰船目标,指挥控制室、动力舱、导弹发射井和弹药舱等部位重要程度较高<sup>[4,7]</sup>。由于不同类型舰船的应用方向差异较大,导弹发射井和弹药舱仅存在于特殊用途的舰船上,并且由于光学遥感图像是俯瞰视角成像,动力舱的可视性差,因而上述几个部位不宜作为精细化检测的对象。相对而言,舰船的指挥控制室的可视性好、重要性强,且存在于多种舰船上。因此,本文选取指挥控制室作为舰船的关重部位并对其进行深入研究。

近年来,随着深度学习的发展成熟,基于卷积神经网络(convolutional neural network, CNN)的目标检测算法成为当前的主流方法。基于 CNN 的目标检测算法主要分为两类:一类是以 R-CNN<sup>[8]</sup>、Faster R-CNN<sup>[9]</sup>、Mask R-CNN<sup>[10]</sup>等为代表的双阶段网络;另一类是以 SSD<sup>[11]</sup>、YOLO<sup>[12]</sup>、RetinaNet<sup>[13]</sup>等为代表的单阶段检测网络。总体来说,双阶段网络检测精度高,但速度慢;单阶段网络检测速度快,但精度偏低。上述具有代表性的网络均是基于水平框的检测模型,并在多种类型的目标检测实验中取得了较优的结果,但对舰船这类具有方向性和大长宽比的目标检测效果不佳。针对这一问题,Liu 等通过引方向参数,将水平框转换为旋转框,使检测框更好地与舰船进行贴合,有效提高了舰船目标的检测精度<sup>[14]</sup>。Ding 等提出了 RoI Transformer<sup>[15]</sup>,通过空间变换将水平 RoI 转化为旋转 RoI,避免使用大量锚框,并能缓解错位问题。R3Det 通过由粗到细逐步精化的方式实现快速、准

确地检测目标<sup>[16]</sup>。在不断的优化下,基于 CNN 的舰船检测算法的精度得到了大幅提升,可以较好地实现舰船检测任务,但仍未对舰船进行精细化检测,无法应用于一些特殊任务中。

针对上述问题,本文以 RetinaNet 为基础,通过优化网络结构、引入注意力机制、构建语义掩膜网络(semantic mask network, SMN)以及增加方向属性,提出了一种用于复杂背景下舰船关重部位检测的网络(CPDNet, critical part detection network)。实验结果表明,与其他网络模型相比,所提网络可以实现对舰船关重部位的精确检测,并有较好的时效性。

## 1 网络模型设计

RetinaNet 是一种高效的单阶段检测网络,通过利用 Focal Loss 来缓解单阶段网络中存在的正负样本不平衡的问题。若将 RetinaNet 直接用于遥感舰船关重部位的检测,效果并不理想。主要原因有两点:1)关重部位特征不够显著,网络无法对其特征进行充分描述;2)受背景影响较大,漏检和误检概率增大。针对以上问题,本文以 RetinaNet 为基准,提出了 CPDNet,整体架构如图 1 所示。本文对 RetinaNet 的改进主要包括以下几个方面:1)优化特征金字塔网络(feature pyramid network, FPN)以更好的表征关重部位特征;2)构建 SMN 以降低背景影响;3)增加方向描述使定位更加精确。

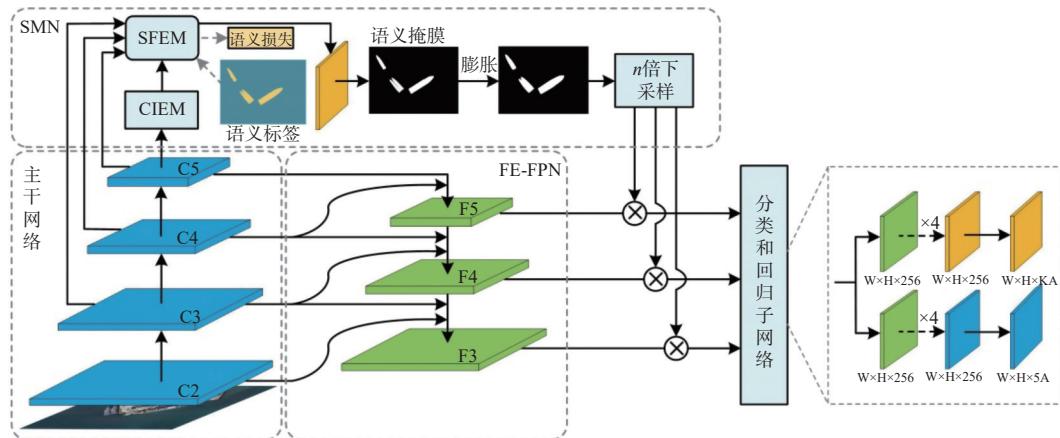


图 1 CPDNet 的整体架构

Fig. 1 Overall structure diagram of CPDNet

### 1.1 优化 FPN 结构

为提高对多尺度目标的检测能力,Retina-

Net 使用 FPN 进行特征融合,但这种融合方式对小尺度目标并不友好<sup>[17]</sup>。一方面,随着网络深度

的加深,低层细节特征不断地被整合成高层语义信息,导致高层特征图中缺少目标的细粒度信息。另一方面,在多次下采样的过程中,小尺度目标的像素逐步减少,受背景影响,特征显著性

难以被凸显,导致网络对小尺度目标的检测效果不佳。鉴于以上分析,我们从丰富特征信息和提高特征显著性两方面入手,设计了 FE-FPN,如图 2 所示。

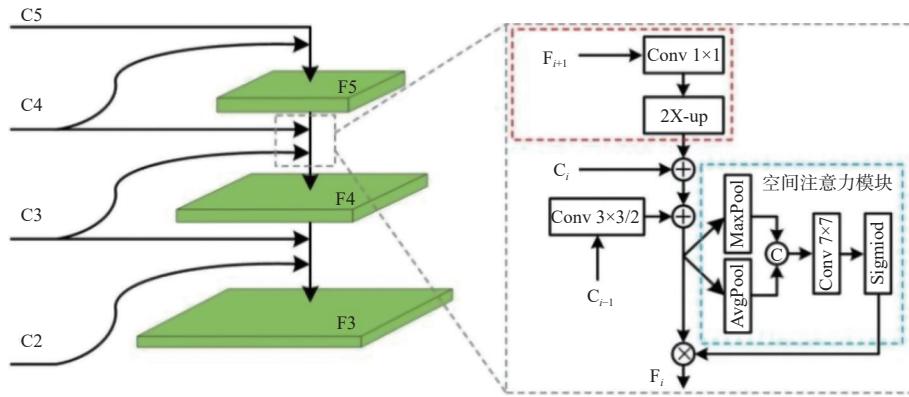


图 2 FE-FPN 结构图

Fig. 2 Structure diagram of FE-FPN

FE-FPN 在初始 FPN 的基础上增加了一条低层特征融合路径,通过融合低层特征来丰富高层特征图所含的细粒度特征。此外,为提高目标显著性,在金字塔自上而下的连接路径中增加空间注意力模块<sup>[18]</sup>。与 F3 和 F4 相比, F5 在构建过程中缺少了图 2 中虚框部分,其余构建过程均相同。特征图  $F_i \in R^{2W \times 2H \times 2C}$  的构建过程可表示为

$$\begin{aligned} F_i = & \text{Att}(P_i) \times P_i = \sigma \left\{ f^{7 \times 7} ([\text{AvgPool}(P_i) \oplus \text{MaxPool}(P_i)]) \right\} \times \\ & P_i \\ P_i = & g_{2X-up} \left( f^{1 \times 1} (F_{i+1}) \right) + C_i + g_{2X-down} (C_{i-1}) \end{aligned} \quad (1)$$

式中:  $F_{i+1} \in R^{W \times H \times 4C}$ 、 $C_i \in R^{2W \times 2H \times 2C}$ 、 $C_{i-1} \in R^{2W \times 2H \times C}$ 、 $\text{Att}(\cdot)$ , 表示空间注意力机制;  $\sigma(\cdot)$  表示 Sigmoid 函数;  $f^{7 \times 7}$  和  $f^{1 \times 1}$  分别表示  $7 \times 7$  卷积和  $1 \times 1$  卷积; AvgPool 和 MaxPool 分别表示平均池化和最大池化;  $g_{2X-up}$  和  $g_{2X-down}$  分别表示 2 倍上采样和 2 倍下采样;  $+$  和  $\oplus$  分别表示像素相加和通道叠加的融合方式。

## 1.2 SMN

光学遥感图像包含丰富的细节信息,能够直观呈现目标的纹理、颜色、形状等信息,同时也能清晰展现目标所处的场景。光学遥感图像舰船关重部位检测的一大难点是场景复杂度高,要想提高模型对关重部位检测的精度,需要尽可能降低背景对检测的影响。为降低无关区域对检测的干扰,本文提出了 SMN。该模块利用语义信息生成语义掩膜,语义掩膜经膨胀处理后作用于 FE-

FPN 生成的特征图,对不可能出现目标的区域进行遮蔽,从而降低背景对检测的影响。SMN 主要由上下文信息提取模块 (contextual information extraction module, CIEM) 和语义特征提取模块 (semantic feature extraction module, SFEM) 组成。

### 1.2.1 CIEM

上下文信息对于语义分割任务至关重要,它可以让网络更好地感知目标轮廓,为后续语义信息的提取奠定基础。如图 3 所示,CIEM 主要由 5 层组成,采用  $1 \times 1$  卷积增加通道相关性,  $3 \times 3$  卷积保留原始特征,2 个膨胀因子为 (2,5) 的  $3 \times 3$  卷积组成的膨胀序列和  $2 \times 2$  的最大池化层(后接上采样操作)提取多尺度上下文信息。将所有层以通道相加的方式进行融合,随后利用  $1 \times 1$  卷积来减少通道数量。

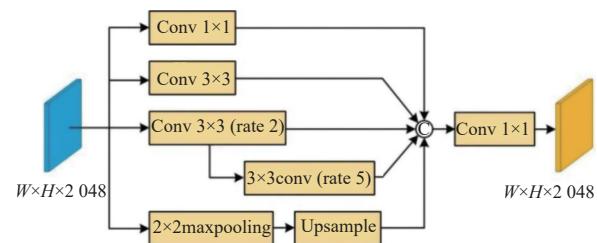


图 3 CIEM 结构图

Fig. 3 Structure diagram of CIEM

### 1.2.2 SFEM

在 SMN 中, SFEM 的构建是关键, 它对语义掩膜质量的好坏起着决定性作用。为获取精确的语义掩膜, 需要对特征图中的每个像素进行分类。采用跳跃连接<sup>[19]</sup>和反卷积操作<sup>[20]</sup>将带有上下文信息的高级特征图和带有位置信息的低级特征图结合起来, 通过多级特征图的叠加和上采样获得每个像素的分类结果。SFEM 的结构如图 4 所示, 各层的详细参数如表 1 所示。具体步骤可概括为: 首先将低层特征图与具有相同尺度的高层特征图按通道叠加的方式进行拼接; 然后将拼接的特征导入反卷积层, 降低特征图尺度并进一步提取图像特征; 最后使用较少的卷积核来调整特征图的深度。经过 Conv 5 后, 得到一个类别标签图和一个语义特征图。类别标签图用于计算语义分割损失, 语义特征图用于生成语义掩膜。

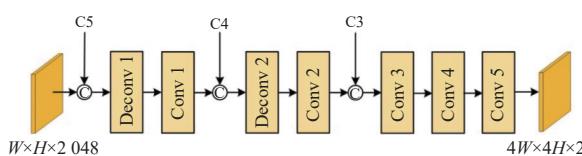


图 4 SFEM 结构图

Fig. 4 Structure diagram of SFEM

表 1 SFEM 各层的详细参数

Table 1 Detailed parameters of each SFEM layer

卷积层	通道数 (输入)	通道数 (输出)	大小 (输入)	大小 (输出)
Deconv 1	4 096	2 048	$W \times H$	$2W \times 2H$
Conv 1	2 048	1 024	$2W \times 2H$	$2W \times 2H$
Deconv 2	2 048	1 024	$2W \times 2H$	$4W \times 4H$
Conv 2	1 024	512	$4W \times 4H$	$4W \times 4H$
Conv 3	512	128	$4W \times 4H$	$4W \times 4H$
Conv 4	128	64	$4W \times 4H$	$4W \times 4H$
Conv 5	64	2	$4W \times 4H$	$4W \times 4H$

特征图 C5 经过 CIEM 和 SFEM 两个模块处理后得到了相应的语义特征图, 对语义特征图进行二值化处理得到语义掩膜图。在生成语义特征图的过程中, 语义分割结果不准确可能会导致舰船关重部位缺失。针对这一问题, 我们对语义掩膜图进行了膨胀处理, 通过扩张舰船区域来降低分割不准确带来的影响。在膨胀处理过程中, 若膨胀结构元素尺寸过小, 则无法弥补误分割的偏差; 若膨胀结构元素尺寸过大, 则会引入背景信息, 对检测精度造成影响。本文使用大小为  $11 \times 11$  的膨

胀结构元素, 相关实验在 2.2 节进行介绍。膨胀处理后的语义掩膜图经不同倍数的上采样操作后分别作用于 F5、F4 和 F3, 通过掩蔽非舰船区域来提高关重部位的检测精度。

### 1.3 角度参数

RetinaNet 使用水平框来定位目标, 而遥感图像中的舰船关重部位具有方向性, 水平框在标定时会包含大量的冗余信息, 影响检测器的精度。因此, 我们按照 R3Det 的构建方式对 RetinaNet 进行改进。在原有的四坐标基础上, 增加一个方向参数  $\theta$ , 即用 5 参数  $(x, y, w, h, \theta)$  来表示旋转框。其中  $(x, y)$  表示旋转框的中心坐标,  $\theta$  是旋转框与水平轴 ( $x$  轴) 夹角的锐角, 其范围为  $[-90^\circ, 0^\circ]$ , 用于定义  $\theta$  的边为宽  $w$ , 另一边为高  $h$ 。在回归子网络中需要增加对角度偏移量的预测, 旋转框偏移量计算公式如下:

$$\begin{aligned} t_x &= (x - x_a)/w_a, t_y = (y - y_a)/h_a \\ t_w &= \log(w/w_a), t_h = \log(h/h_a), t_\theta = \theta - \theta_a \\ t'_x &= (x' - x_a)/w_a, t'_y = (y' - y_a)/h_a \\ t'_w &= \log(w'/w_a), t'_h = \log(h'/h_a), t'_\theta = \theta' - \theta_a \end{aligned} \quad (2)$$

式中:  $(x_a, y_a, w_a, h_a, \theta_a)$  表示锚框的位置坐标;  $(x, y, w, h, \theta)$  和  $(x', y', w', h', \theta')$  分别表示真实框和预测框的位置坐标;  $(t_x, t_y, t_w, t_h, t_\theta)$  和  $(t'_x, t'_y, t'_w, t'_h, t'_\theta)$  分别表示锚框偏移量和预测框偏移量。

### 1.4 损失函数

所提网络的损失函数由语义分割损失  $L_{\text{seg}}$ 、分类损失  $L_{\text{cls}}$  和回归损失  $L_{\text{reg}}$  三部分组成, 计算公式如下所示。

$$Loss = \lambda L_{\text{seg}} + \alpha L_{\text{cls}} + \beta L_{\text{reg}} \quad (3)$$

式中:  $\lambda$ 、 $\alpha$  和  $\beta$  是权重参数, 用来平衡不同损失间的差距。

$L_{\text{reg}}$  由每个像素的预测标签和真实标签间的交叉熵损失计算得到, 计算方法如下:

$$L_{\text{seg}} = -\frac{1}{W \times H} \sum_i \sum_j M(i, j) \cdot \log(M'(i, j)) + (1 - M(i, j)) \cdot \log(1 - M'(i, j)) \quad (4)$$

式中:  $W \times H$  表示掩膜大小;  $M(\cdot)$  和  $M'(\cdot)$  分别表示标签的真实值和预测值。

$L_{\text{cls}}$  由下式计算得到:

$$L_{\text{cls}} = \frac{1}{N} \sum_{n=1}^N -\alpha_n (1 - p_n)^\gamma \log(p_n) \quad (5)$$

式中:  $N$  表示锚框数量;  $p_n$  表示样本被正确预测的概率;  $\alpha_n$  设为 0.25;  $\gamma$  设为 2。

$L_{\text{reg}}$  的计算方式与 R3Det 相同:

$$\begin{aligned} L_{\text{reg}} &= \frac{1}{N} \sum_{n=1}^N \text{obj}_n \frac{L(v'_n, v_n)}{|L(v'_n, v_n)|} |f(S \text{kewIoU})| \\ L(v', v) &= L_{\text{smooth-l1}}(v'_\theta, v_\theta) - \text{IoU}(v'_{\{x,y,w,h\}}, v_{\{x,y,w,h\}}) \\ L_{\text{smooth-l1}}(v'_\theta, v_\theta) &= \begin{cases} 0.5 \times (v'_\theta - v_\theta)^2, & |v'_\theta - v_\theta| < 1 \\ |v'_\theta - v_\theta| - 0.5, & \text{else} \end{cases} \end{aligned} \quad (6)$$

式中:  $\text{obj}_n$  取值为 1 或 0 ( $\text{obj}_n = 1$  表示前景,  $\text{obj}_n = 0$  表示背景);  $v'_n$  表示预测偏移量;  $v_n$  表示目标的真实向量;  $| \cdot |$  表示向量取模;  $f(\cdot)$  表示与相关的损失函数;  $\text{IoU}(\cdot)$  表示水平框的交并比。

## 2 实验

实验在配置为 E5-2630v2 CPU、NVIDIA GTX-1080Ti GPU(11G video memory)、64G RAM 的图像工作站上进行, 以 Ubuntu18.04 为平台, 在深度学习框架 Pytorch1.6 下编译完成。为从定量角度验证所提网络的有效性, 使用每秒处理帧数(frame per second, FPS)评价网络的检测速度, 利用准确率(Precision,  $P$ )、召回率(Recall,  $R$ )、平均精度(average precision,  $AP$ )评价网络的检测性能。各评价指标定义如下:

$$FPS = \frac{N_{\text{Figure}}}{\text{Time}} \quad (7)$$

$$P = \frac{TP}{TP + FP} \times 100\% \quad (8)$$

$$R = \frac{TP}{TP + FN} \times 100\% \quad (9)$$

$$AP = \int_0^1 P(R) dR \quad (10)$$

式中:  $N_{\text{Figure}}$  为所处理的图像数量; Time 是图像处理总时间;  $TP$  为正确分类的正样本数量;  $FP$  为错误分类为正样本的数量;  $FN$  为错误分类为负样本的数量。

### 2.1 数据集及训练细节

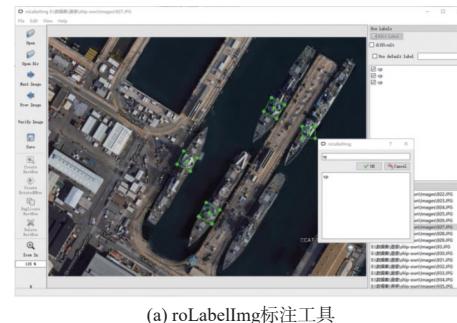
数据集: 当前用于舰船检测的光学遥感数据集, 比如 HRSC2016<sup>[21]</sup>、DIOR<sup>[22]</sup>、DOTA<sup>[23]</sup>等, 均以整体舰船作为目标进行标注, 多用于舰船位置的检测任务。将舰船部位作为研究对象时, 数据集中应包含明显的部位信息, 目前所存在的数据集无法满足这一需求。针对数据集欠缺的问题, 我们利用谷歌地图从世界上一些著名的港口(珍珠港、横须贺军港、朴次茅斯军港等)下载了 1 015 张图像, 构建了 CP-Ship 数据集, 数据集的详细信

息如表 2 所示。CP-Ship 中图像尺度像素为 800×600, 使用 roLabelImg 和 Labelme 分别对图像中的关重部位和舰船进行标注, 如图 5 所示。roLabelImg 的标注数据用于网络对舰船关重部位的训练和测试, 而 Labelme 的标注数据用于语义分割的训练。

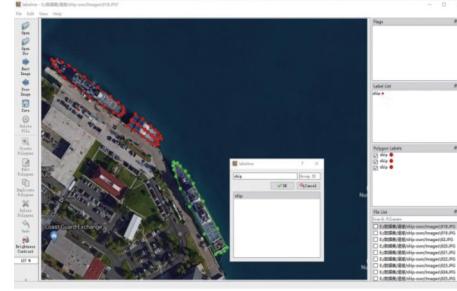
表 2 CP-Ship 数据集详细信息

Table 2 Details of CP-Ship dataset

数据集	CP-Ship	
	图像数量	关重部位个数
训练集	812	2 295
测试集	203	561
总计	1 015	2 856



(a) roLabelImg标注工具



(b) Labelme标注工具

图 5 标注示意图

Fig. 5 Schematic diagram of labeling

训练细节: 使用 ResNet-50<sup>[19]</sup> 作为主干网络, 利用在 ImageNet<sup>[24]</sup> 上训练的 ResNet-50 权重初始化主干网络。训练过程中, 网络在随机梯度下降法(stochastic gradient descent, SGD)下进行 36 轮迭代训练, 初始学习率为 0.0025, 批次大小为 2。在迭代 24 和 33 轮时, 学习率分别降低 10 倍。权重衰减和动量分别设置为 0.0001 和 0.9。

### 2.2 膨胀结构元素尺寸的选取

SMN 中使用膨胀操作的目的是降低误分割的影响, 合适的膨胀结构元素可以提高检测精度。为选取合适的膨胀结构元素, 改变模型中膨胀结构元素大小进行实验, 通过对比检测的平均精度

来确定膨胀结构元素大小, 相关实验结果如表 3 所示。

表 3 膨胀结构元素尺寸选取

Table 3 Size selection of expansion structure element

结构元素尺寸	AP/%	结构元素尺寸	AP/%
3×3	68.67	13×13	71.38
5×5	69.69	15×15	70.97
7×7	70.28	17×17	69.94
9×9	71.25	19×19	68.32
11×11	71.76	21×21	66.46

从表 3 可以看出, 随着结构元素尺寸的增大, 检测的精度呈现先增大后减小的态势。当结构元素大小为 11×11 时, 网络的检测精度最高。因此, 本文采用结构元素为 11×11 的膨胀操作。

### 2.3 消融实验

为评估 FE-FPN、SMN 以及角度参数对网络检测能力的贡献, 以 RetinaNet 为基准, 通过添加单个或多个模块进行实验, 实验结果如表 4 所示。

从表 4 中可以看出, 通过增加角度参数使 RetinaNet 的 AP 值提高了 2.23%, 证明了角度参数对于检测具有方向性目标的重要性。在增加角度的基

表 4 消融实验结果

Table 4 Results of ablation experiment

RetinaNet	角度参数	FE-FPN	SMN	AP/%
√				60.41
√	√			62.64
√	√	√		66.45
√	√	√	√	71.76

础上, 将 RetinaNet 中的 FPN 替换成 FE-FPN, AP 值提高了 3.81%, 说明在丰富特征的基础上增加注意力机制可以提高网络的检测能力。在上述基础上增加 SMN 大幅提高了检测精度, 说明在复杂背景下, 通过降低背景的影响可以提升网络的检测性能。

为更加直观说明 FE-FPN 和 SMN 两个模块的作用效果, 对 FPN 和 FE-FPN 的 F4 层特征图以及经过掩膜修饰用于检测的特征图进行可视化, 如图 6 所示。通过对图 6 中第 2 列和第 3 列可知, FE-FPN 在丰富特征的同时可以突出目标的显著性; 将第 4 列与前两列进行对比发现, 经过掩膜修饰后特征图所包含的背景信息大幅减少, 可以降低背景对检测的干扰, 目标的显著性也得到了凸显。

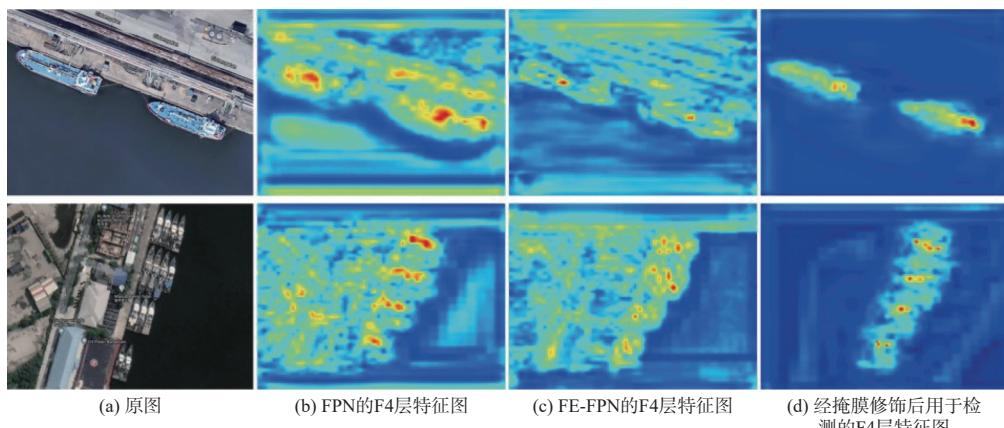


图 6 特征图可视化示例

Fig. 6 Visualization examples of feature images

### 2.4 对比实验

为验证所提网络的有效性, 将 CPDNet 与一些流行的目标检测算法进行对比分析。所用对比模型的相关代码来源于 OpenMMLab 平台。各网络模型在测试集上检测的定量结果和 PR 曲线如表 5 和图 7 所示。

从实验的定量结果中可以看出, 双阶段网络的检测性能要优于单阶段检测网络, 但在检测速度

上不占优势。单阶段网络和双阶段网络存在共点: 对于具有方向性的目标, 通过引入旋转框进行检测可有效提高检测精度, 但引入旋转框会增加网络的参数量, 使整个网络的检测速度大幅下降。RetinaNet 网络结构简单, 易受复杂背景的影响, 舰船关键部位目标的漏检和误检数量在所有对比算法中最多。与 RetinaNet 相比, 所提网络的 AP 值提高了 11.35%, 并且 FP 和 FN 各有不同程度

表 5 不同网络模型的定量结果

Table 5 Quantitative results of different network models

类别	方法	主干网络	锚框类型	TP	FP	FN	AP/%	FPS
双阶段网络	Faster R-CNN <sup>[9]</sup>	ResNet-50	水平框	411	152	150	65.78	19.2
	ReDet <sup>[25]</sup>	ResNet-50	旋转框	454	103	107	75.24	3.1
	Dynamic R-CNN <sup>[26]</sup>	ResNet-50	水平框	400	130	161	63.85	20.5
	Oriented R-CNN <sup>[27]</sup>	ResNet-50	旋转框	423	169	138	68.39	11.8
单阶段网络	YOLOv3 <sup>[28]</sup>	DarkNet-53	水平框	413	84	148	66.49	51.8
	R3Det <sup>[15]</sup>	ResNet-50	旋转框	393	146	168	64.28	11.3
	RetinaNet <sup>[13]</sup>	ResNet-50	水平框	387	229	174	60.41	23.2
	CFA <sup>[29]</sup>	ResNet-50	旋转框	414	225	306	63.61	19.3
	CPDNet	ResNet-50	旋转框	433	150	128	71.76	12.5

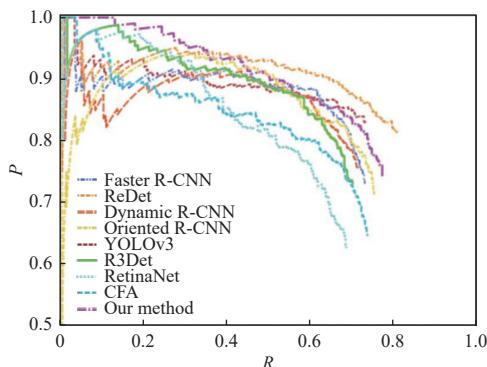


图 7 不同网络模型的 PR 曲线

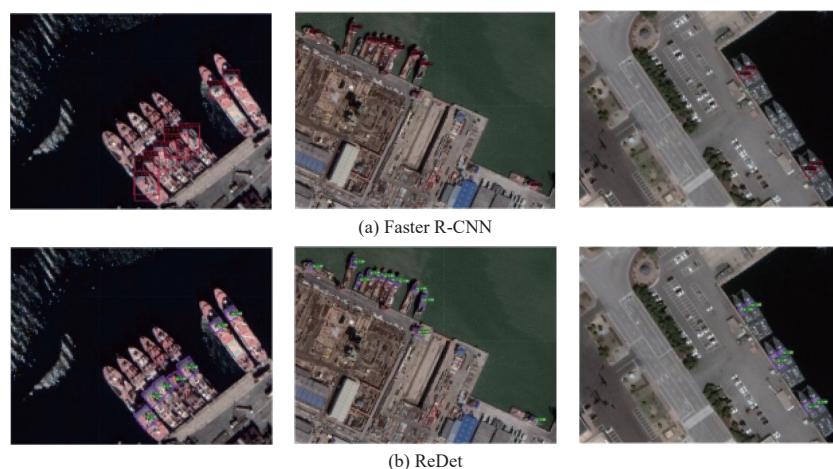
Fig. 7 PR curves of different network models

的降低, 虽然检测速度有所下降, 但在使用旋转框进行检测的网络中, 所提网络的检测速度占有一定的优势。从 PR 曲线图中可以看出, CPDNet 的 PR 曲线与坐标轴之间所围面积大于除 ReDet 外的所有模型, 并且曲线最为平稳, 表明所提网络的检测性能较好。

从各方法的检测速率来看, 水平框检测算法的

平均检测速率要快于旋转框检测算法, 究其原因是角度参数的引入会增加整个网络的计算量, 影响网络的推理速度。在旋转框检测算法中, CFA 以简单的网络结构和轻量的计算方式在检测速度上占绝对优势, ReDet 的双阶段网络结构和复杂的特征提取方式使其检测速度居于末尾。CPDNet 在 RetinaNet 基础上进行了多种改进, 致使其检测速度较 RetinaNet 略有不足, 但在水平框检测算法中处于前列。综上所述, 所提算法在同类型的算法中具有较高的优越性, 较好地兼顾了检测的速度和精度。

为更加直观地对以上 9 种方法进行比对, 图 8 给出了不同模型的可视化检测结果。从图 8 中可以看出, 与水平框相比, 旋转框与目标更加贴合, 可以更加直观反映出目标的位置; 对于并列目标, Faster R-CNN、Dynamic R-CNN、RetinaNet 因水平框之间存在交叠, 致使出现大量误检目标; ReDet、Oriented R-CNN、Oriented R-CNN 和 Oriented R-CNN 虽然借助旋转框改善了对并列目标的检测效果,



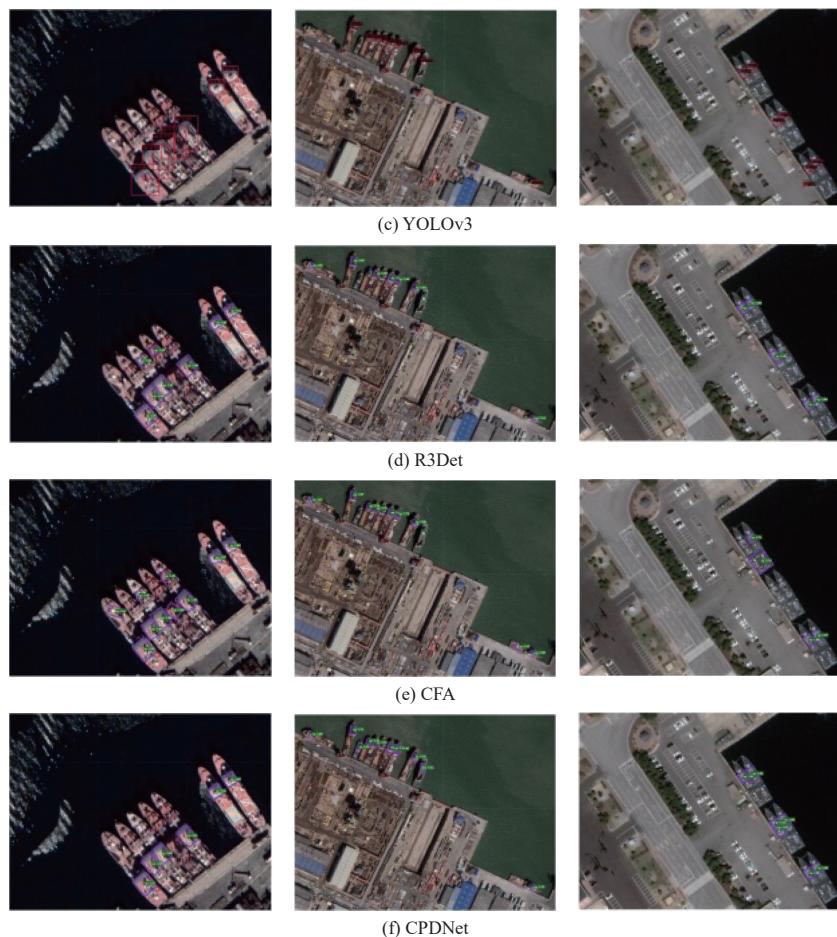


图 8 不同网络模型的检测结果可视化示例

Fig. 8 Visualization examples of detection results of different network models

但受复杂背景影响,存在大量的漏检目标。通过对比,所提方法的漏检和误检数量均最少,整体检测效果较好。总而言之,所提方法使用旋转框保证了对并列目标的检测效果,通过丰富特征及引入注意力机制,提升了应对各类背景的能力,使用语义掩膜规避背景对检测的影响,降低了误检概率,适用于复杂背景下舰船关重部位检测任务。

### 3 结论

本文对舰船的精细化检测作了进一步研究,在 RetinaNet 检测算法的基础上,提出了一种用于舰船关重部位检测的网络模型。该模型以 ResNet-50 为主干网络提取图像的深层次特征;通过优化 FPN 结构以及引入注意力机制的方式提高模型的特征表达能力,并提升模型对关重部位目标的敏感度;利用 SMN 网络模块构建出区域掩膜,降低背景对检测的影响,从而降低误检概率提高检测精度;检测器通过 5 参数的旋转框实现舰船关重部

位的定向检测。在 CP-Ship 数据集上的实验结果表明:所提网络的关重部位检测精度 AP 较 RetinaNet 提高了 11.35%;与其他旋转框检测模型相比,所提网络在速度和精度上均占有绝对的优势,验证了所提网络对关重部位检测的有效性和优越性。

### 参考文献:

- [1] HUANG Z. A sea-land segmentation algorithm based on graph theory[C]//2nd ISPRS International Conference on Computer Vision in Remote Sensing (CVRS 2015).Xiamen, China: ISPRS, 2016: 990111.
- [2] XU J, FU K, SUN X. An invariant generalized Hough transform based method of inshore ships detection [C]//2011 International Symposium on Image and Data Fusion. Tengchong, China: IEEE, 2011: 1-4.
- [3] 牛戈, 陈小前, 季明江, 等. 基于注意力机制特征重建网络的舰船目标检测[J]. 上海航天, 2021, 38(4): 128-136. NIU Ge, CHEN Xiaoqian, JI Mingjiang, et al. Ship detec-

- tion with feature reconstruction network based on attention mechanism[J]. *Aerospace Shanghai*, 2021, 38(4): 128-136.
- [4] 李晨瑄, 顾佼佼, 王磊, 等. 多尺度特征融合的Anchor-Free轻量化舰船要害检测算法[J]. *北京航空航天大学学报*, 2022, 48(10): 2006-2019.  
LI Chenxuan, GU Jiaojiao, WANG Lei, et al. Warship's vital parts detection algorithm based on lightweight Anchor-Free network with multi-scale feature fusion[J]. *Journal of Beijing University of Aeronautics and Astronautics*, 2022, 48(10): 2006-2019.
- [5] 李晨瑄, 钱坤, 胥辉旗. 基于深浅层特征融合的舰船要害关键点检测算法[J]. *系统工程与电子技术*, 2021, 43(11): 3239-3249.  
LI Chenxuan, QIAN Kun, XU Huiqi. Key-points detection algorithm based on fusion of deep and shallow features for warship's vital part[J]. *Systems Engineering and Electronics*, 2021, 43(11): 3239-3249.
- [6] 李晨瑄, 李添雨, 李梓正, 等. 智能化舰船要害检测、轨迹预测与位姿估计算法[J]. *北京航空航天大学学报*, 2023, 49(2): 444-456.  
LI Chenxuan, LI Tianyu, LI Zizheng, et al. Intelligent algorithm of warship's vital parts detection, trajectory prediction and pose estimation[J]. *Journal of Beijing University of Aeronautics and Astronautics*, 2023, 49(2): 444-456.
- [7] 赵微, 惠斌, 张玉晓. 红外舰船目标的要害点检测算法[J]. *红外与激光工程*, 2014, 43(1): 48-52.  
ZHAO Wei, HUI Bin, ZHANG Yuxiao. Detection algorithms of aim points of IR warship images[J]. *Infrared and Laser Engineering*, 2014, 43(1): 48-52.
- [8] GIRSHICK R. Fast R-CNN[C]//2015 IEEE International Conference on Computer Vision (ICCV). Santiago, Chile: IEEE, 2015: 1440-1448.
- [9] REN S, HE K, GIRSHICK R, et al. Faster r-cnn: towards real-time object detection with region proposal networks[J]. *Advances in Neural Information Processing Systems*, 2015, 28: 91-99.
- [10] HE K, GKIOXARI G, DOLL A R P, et al. Mask r-cnn[C]//Proceedings of the IEEE International Conference on Computer Vision. Venice, Italy: IEEE, 2017: 2961-2969.
- [11] LIU W, ANGUELOV D, ERHAN D, et al. SSD: single shot multibox detector[C]//ECCV 2016: Computer Vision – ECCV 2016. Amsterdam, The Netherlands: Springer C, 2016: 21-37.
- [12] REDMON J, FARHADI A. YOLO9000: better, faster, stronger[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition(CVPR) Honolulu.HI, USA: IEEE, 2017: 7263-7271.
- [13] LIN T Y, GOYAL P, GIRSHICK R, et al. Focal loss for dense object detection[C]//Proceeding of the IEEE International Conference on Computer Vision. Piscataway: IEEE, 2017: 2980-2988.
- [14] LIU Z, HU J, WENG L, et al. Rotated region based CNN for ship detection[C] //2017 IEEE International Conference on Image Processing (ICIP).Beijing, China: IEEE, 2017: 900-904.
- [15] 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), Long Beach, CA. USA: IEEE, 2019: 2844-2853.
- [16] YANG X, YAN J, FENG Z, et al. R3det: refined single-stage detector with feature refinement for rotating object[C]// AAAI conference on artificial intelligence Palo Alto California USA:AAAI,2021,35(4):3163-3171.
- [17] 韩子硕, 王春平, 付强, 等. 基于超密集特征金字塔网络的SAR图像舰船检测[J]. *系统工程与电子技术*. 2020, 42(10): 2214-2222.  
HAN Zishuo, WANG Chunping, FU Qiang, et al. Ship detection in SAR images based on super dense feature pyramid networks[J]. *Systems Engineering and Electronics*. 2020, 42(10): 2214-2222.
- [18] WOO S, PARK J, LEE J, et al. Cbam: convolutional block attention module[C]//ECCV 2018: Computer Vision – ECCV 2018.Munich, Germany: Springer C, 2018: 3-19.
- [19] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas, NV, USA: IEEE, 2016: 770-778.
- [20] ZEILER M D, TAYLOR G W, FERGUS R. Adaptive deconvolutional networks for mid and high level feature learning[C]//2011 International Conference on Computer Vision. Barcelona, Spain: IEEE, 2011: 2018-2025.
- [21] LIU Z, YUAN L, WENG L, et al. A high resolution optical satellite image dataset for ship recognition and some new baselines[C]//6th International Conference on Pattern Recognition Applications and Methods(ICPRAM

- 2017).Porto,Portugal:SCITE,2017:324-331.
- [22] LI K, WAN G, CHENG G, et al. Object detection in optical remote sensing images: a survey and a new benchmark[J]. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2020, 159: 296-307.
- [23] XIA G, BAI X, DING J, 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, UT, USA: IEEE, 2018: 3974-3983.
- [24] DENG J, DONG W, SOCHER R, et al. ImageNet: a large-scale hierarchical image database[C]//2009 IEEE Conference on Computer Vision and Pattern Recognition. Miami, FL, USA: IEEE, 2009: 248-255.
- [25] HAN J, 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).Nashville, TN, USA: IEEE, 2021: 2785-2794.
- [26] ZHANG H, CHANG H, MA B, et al. Dynamic R-CNN: towards high quality object detection via dynamic training[C]//ECCV 2020: Computer Vision – ECCV 2020.Glasgow, UK: Springer C, 2020: 260-275.
- [27] XIE X, CHENG G, WANG J, et al. Oriented R-CNN for object detection[C]//2021 IEEE/CVF International Conference on Computer Vision (ICCV).Montreal, QC, Canada: IEEE, 2021: 3500-3509.
- [28] WANG J T, XIAO W , NI T W. Efficient object detection method based on improved YOLOv3 network for remote sensing images[C]//2020 3rd International Conference on Artificial Intelligence and Big Data (ICAIBD). Chengdu, China: IEEE, 2020: 242-246.
- [29] GUO Z, LIU C, ZHANG X, et al. Beyond bounding-box: convex-hull feature adaptation for oriented and densely packed object detection[C]//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Nashville, TN, USA: IEEE, 2021: 8788-8797.