

自适应旋转区域生成网络的遥感图像舰船目标检测

徐志京, 丁莹*

上海海事大学信息工程学院, 上海 201306

摘要 针对遥感图像中舰船形状狭长、分布杂乱等特性导致检测难度增大的问题, 提出了一种基于快速区域卷积神经网络(Faster R-CNN)的舰船目标检测方法。采用双路网络提取舰船目标特征, 为了使特征图充分融合底层细节信息和高层语义信息, 用多尺度融合特征金字塔网络(MFPN)进行特征融合; 在候选框生成阶段, 提出了自适应旋转区域生成网络(AR-RPN), 集中在目标中心位置生成旋转锚框, 以高效获取优质的候选框。为了提升网络对舰船目标的检测率, 结合改进的损失函数对网络进行优化。在 HRSC2016 和 DOTA 舰船数据集上的测试结果表明, 本方法的平均精度分别为 89.10% 和 88.64%, 能很好地适应遥感图像中舰船的形状与分布特性。

关键词 图像处理; 舰船检测; 遥感图像; 多尺度特征融合; 自适应旋转区域生成网络

中图分类号 TP183, TP751.1

文献标志码 A

doi: 10.3788/LOP57.242805

Ship Object Detection of Remote Sensing Images Based on Adaptive Rotation Region Proposal Network

Xu Zhijing, Ding Ying*

College of Information Engineering, Shanghai Maritime University, Shanghai 201306, China

Abstract Aiming at the problem that increased difficulties in detection of ship detection in remote sensing images caused by the narrow and long shape, disorderly distribution and other characteristics, a ship target detection method based on faster region-convolution neural network (Faster R-CNN) is proposed in this paper. The method uses a two-way network to extract ship target features. In order to make the feature map fully integrate the low-level detail information and high-level semantic information, a multi-scale fusion feature pyramid network (MFPN) is used for feature fusion; in the candidate frame generation stage, an adaptive rotation region proposal network (AR-RPN) is proposed to generate a rotating anchor frame at the center of the target to efficiently obtain high-quality candidate frames. In order to improve the detection rate of the network to ship targets, the network is optimized with an improved loss function. The test results on the public ship data sets HRSC2016 and the DOTA show that the average accuracy of this method is 89.10% and 88.64%, respectively, which can well adapt to the shape and distribution characteristics of ships in remote sensing images.

Key words image processing; ship detection; remote sensing images; multi-scale feature fusion; adaptive rotation region proposal network

OCIS codes 100.2000; 110.2960; 150.4065

1 引言

遥感图像舰船检测是遥感信息处理的一个热点问题, 在海防建设、港口管理等方面发挥着重要作用。随着遥感技术的发展, 人们可以获得分辨率更高的光学遥感图像, 提供目标更多的细节信息, 对精确识别和定位舰船具有重要意义。同时, 光学图像

的检测结果也使信息更直观, 便于决策者进行判断^[1]。针对光学遥感图像舰船目标检测的研究已取得了丰硕的成果^[2-5], 传统目标检测方法都是利用人工构造的几何特征, 先提取图像中对应的特征, 然后用分类器训练完成分类, 最后进行边框修正实现检测^[6]。但这种基于机器学习的算法费时费力, 且鲁棒性差, 在实际光学遥感图像的复杂场景下表现较差。

收稿日期: 2020-06-15; 修回日期: 2020-06-23; 录用日期: 2020-06-24

基金项目: 国家自然科学基金(61673259)、航空科学基金(201955015001)

*E-mail: 1405530454@qq.com

随着深度学习的发展,目标检测网络也得到了飞速发展,目前主流的单阶段检测网络有单步多框检测(SSD)^[7]、YOLO(You only look once)^[8]等,双阶段网络有快速区域卷积神经网络(Faster R-CNN)^[9]、Mask R-CNN^[10]等。单阶段网络的目标检测速度优于双阶段网络,但双阶段网络的目标检测精度优于单阶段网络,两者在自然场景中均得到了广泛的应用。由于遥感图像都是俯拍图,物体的角度分布随机,且舰船的狭长形状导致其倾斜一定角度(如相对垂直方向倾斜 45°)时会引入大量的背景噪声,当舰船密集分布时也会造成漏检。

最初的斜框检测方法——旋转候选区域网络(RRPN)^[11]是基于旋转锚框的检测网络,RRPN原本用于多角度文本的检测,也适用于遥感舰船目标的检测。Liu等^[12]在2017年提出了旋转区域卷积神经网络(RR-CNN),将斜框检测应用到舰船目标检测领域。Yang等^[13-14]基于AO-RCNN(Arbitrary-oriented ships via multi task rotation RCNN)设计了自适应感兴趣区域对齐(Adaptive ROI Align)方法,可以减轻候选框中冗余噪声的影响。但大量的锚框增加了网络中的参数,降低了检测效率。Ding等^[15]设计了改进的ROI学习器和旋转位置敏感ROI对齐(RPS-ROI-Align)模块,通过提取旋转不变特征,解决了旋转锚框导致的计算复杂度提高问题,以促进后续的分类和回归。这些方法在DOTA数据集上的检测结果较好,但仍然存在较多的误检、漏检情况。

针对上述问题,本文基于Faster R-CNN框架提出了一种遥感图像舰船目标检测方法,通过双路组合的卷积神经网络(CNN)进行特征提取,提供不同的感受野以适应遥感图像的复杂背景;改进了多尺度融合特征金字塔网络(MFPN),充分融合了多

尺度特征信息,输出多级特征图以生成候选框;提出了自适应旋转区域生成网络(AR-RPN),可自适应地在目标中心位置生成锚框,并将90%的锚框集中在目标区域,以减少无效锚框、提升检测效率。通过位置语义信息指导锚框形状预测,有利于检测狭长的舰船目标;考虑到多种附加损失对网络进行优化,进一步提升了舰船目标检测的准确率。

2 模型设计

2.1 模型结构

本方法使用的模型结构以Faster R-CNN框架为基础,如图1所示。第一步将处理好的数据送入双路骨干网络进行特征提取,其中,A路为残差神经网络ResNet-101^[16]的基础网络,B路为基于可变形卷积的ResNext-101^[17],不同的特征提取网络可以提供不同的感受野,从而增强目标的检测性能。第二步将特征图分别送入特征金字塔网络(FPN)进行特征融合,考虑到舰船类内尺度差异较大,用改进的MFPN采用密集连接和跳跃连接的方式有效融合多级特征。第三步用设计的自适应旋转区域生成网络(AR-RPN)生成锚框(anchors),寻找潜在的舰船目标,输出最接近目标的前2000个锚框作为候选框(proposal)。第四步通过旋转感兴趣区域对齐(rotated ROI Align)网络微调候选框。第五步用旋转非极大值抑制(rotated NMS)筛选两路网络输出的候选框,输出最好的检测结果。

2.2 多尺度融合特征金字塔

随着CNN层数的加深,生成特征图的分辨率逐渐降低,语义特征逐渐增加,FPN结构是一种能有效融合多级信息的多尺度方法^[18]。由于舰船目标从航母到小艇尺度差异巨大,融合底层细节信息和高层语义信息对舰船检测效果的提升有重要意

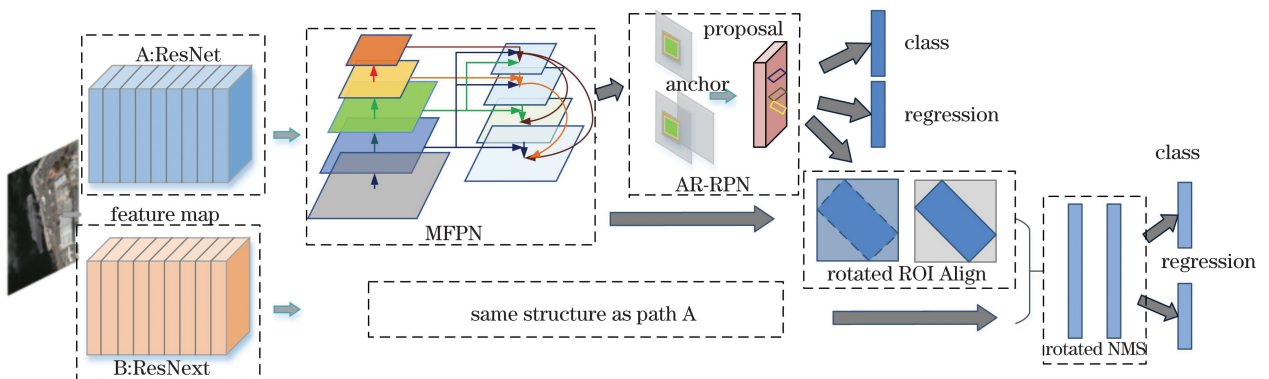


图1 本网络的模型结构

Fig. 1 Model structure of our network

义。实验设计的 MFPN 采用密集连接和跳跃连接方式融合多尺度舰船特征信息,如图 2 所示。虽然特征提取网络输出的第 4、第 5 层特征图 P4、P5 语义信息丰富,但缺失了舰船的位置、细节信息。因此,将 ResNet 第 2、第 3 层输出的 C2、C3 分别经过下采样后与 P4、P5 融合,在融合后的特征图上生成锚框。同理,低层特征图生成的 P2、P3 细节信息丰富,但语义信息较少,因此,将高层特征图经上采样后与 P2、P3 进行融合。

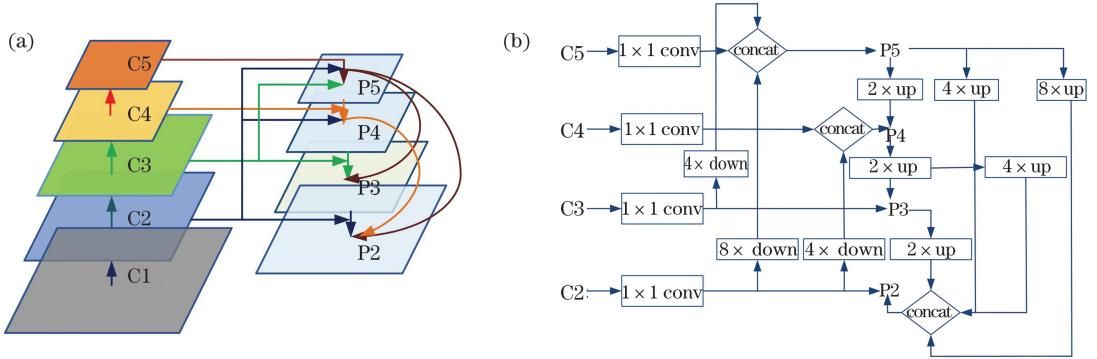


图 2 MFPN 的结构图。(a) 总体结构图;(b) 具体连接图

Fig. 2 Structure diagram of MFPN. (a) Overall structure diagram; (b) specific connection diagram

2.3 自适应旋转区域生成网络

传统 Faster R-CNN 中的区域生成网络(RPN)通过滑动窗口法生成锚框,然后从大量锚框中产生候选框,并对候选框进行分类,从而区分前景(目标区域)和背景(非目标区域)。RPN 在尺寸为 $W \times H$ 的特征图上,以 s 为步长均匀地生成锚框,但在大多图像中,目标不是均匀分布的,与检测目标无关的区域(如海洋、海岸)占大多数,在非目标区域设置的锚框会产生大量计算开销,同时干扰对目标的检测。只在目标中心位置生成锚框可以减少计算量与干扰项,根据目标中心位置的范围大小也可以相应地调整锚框尺寸、提高检测率,实验利用目标的位置信息指导旋转锚框的生成。

AR-RPN 分为位置预测、形状预测及特征自适应模块,其结构如图 3 所示。其中, F 为 MFPN 输出的特征图,实际舰船目标用数组 $(x_g, y_g, w_g, h_g, \theta_g)$ 标注。位置预测分支输出舰船目标中心点预测值 (x, y) ,形状预测分支在位置预测信息的指导下,输出目标形状预测值 (w, h, θ) ,其中, w 为宽度, h 为高度, θ 为角度,取值范围为 $[-30^\circ, 120^\circ]$,间隔为 30° 。在两个分支的输出中,通过选择最大可能位置和该位置最可能的形状生成一组锚框。特征自适应模块利用形状预测分支输出的目标偏移量(offset)调整原始

ResNet 的 5 层输出分别为 $\{C1, C2, C3, C4, C5\}$,经特征融合后 MFPN 的输出为 $\{P2, P3, P4, P5\}$ 。如图 2(b)所示, P2 由 C2 经 1×1 卷积降维、P5 经 8 倍上采样(upsampling)、P4 经 4 倍上采样、P3 经 2 倍上采样后组合得到。P3 由 C3 经 1×1 卷积降维、P5 经 4 倍上采样、P4 经 2 倍上采样后组合得到。P4 由 C4 经 1×1 卷积降维、P5 经 2 倍上采样、C2 经 4 倍下采样(downsampling)组合得到。P5 由 C5 经 1×1 卷积降维、C3 经 4 倍下采样、C2 经 8 倍下采样组合得到。

特征图,实现特征对齐,输出特征图 F_1 。

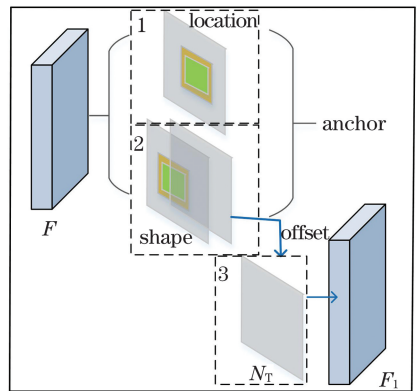


图 3 AR-RPN 的结构图

Fig. 3 Structure diagram of AR-RPN

2.3.1 位置预测

位置预测网络的目标是预测出可作为中心点生成锚框的区域,即区分目标中心区域和非目标中心区域。这类二分类问题一般都是先对特征图 F 进行 1×1 卷积运算,然后通过 Sigmoid 函数将卷积特征转化为该位置为目标中心的概率值。生成位置概率特征图后,设定一个位置阈值 ϕ_{loc} (实验设为 0.5)进行筛选。在高于阈值的位置进行形状预测,就能将锚框集中在舰船目标中心区域,过滤大部分无效的背景区域,从而降低计算量、提高准确率。具体的

预测步骤如下。

首先用数据集标注的边界框指导特征图上二进制标签的生成,将中心坐标为 (x, y) 、大小为 $(\sigma(\omega \cos \theta + h \sin \theta) \times (\omega \sin \theta + h \cos \theta))$ 的矩形区域定义为目标区域,其中, σ 为比例系数,取值范围为 $(0, 1)$ 。设 $\sigma \leq 0.2$ 时,该区域为目标绝对中心区域,标记为1; $\sigma \in (0.2, 0.5]$ 时,该区域为目标中心可能区域,不进行标记; $\sigma > 0.5$ 时,该区域为非目标中心区域,标记为0。将标记后的特征图送入位置预测网络进行训练,经过卷积层和 Sigmoid 函数转化输出概率大于0.5的中心点预测值。

2.3.2 形状预测

在确定目标中心区域后,形状预测分支需要确定每个位置目标的形状 (ω, h, θ) 。常规锚框生成方法通过锚框生成器在特征图上均匀生成多组锚框,输出与标注边界框交并比(IoU)最高的一组锚框。本方法的形状预测分支直接在目标中心位置生成锚框,生成的锚框与最近标注框的初始 IoU 较高,从而达到更好的检测效果。由于目标尺寸从1 pixel 到 1000 pixel 大小不一,直接预测的 (ω, h, θ) 不稳定,因此需要预测标准化坐标 $(\omega_s, h_s, \theta_s)$,可表示为

$$\begin{aligned} \omega &= \rho \cdot s \cdot \exp(\omega_s), h = \rho \cdot s \cdot \exp(h_s), \\ \theta &= \rho \cdot s \cdot \exp(\theta_s), \end{aligned} \quad (1)$$

式中, ρ 为一个经验比例因子(实验中取8)。这种非线性变换将目标尺寸从 $[0, 1000]$ 标准化到 $[-1, 1]$ 之间,从而使网络训练更稳定。具体的预测步骤:1)对输入特征图进行 $1 \times 1 \times 3$ 卷积得到形状特征图;2)根据舰船的比例特点设置锚框的长宽比例分别为 $\{1:7, 1:5, 1:3, 1:2, 1:1, 2:1, 3:1, 5:1, 7:1\}$,角度 θ 的取值范围为 $[-30^\circ, 120^\circ]$,间隔为 30° ;3)在目标中心位置生成不同形状的锚框,返回 IoU 最高的一组锚框作为候选框,经元素转换输出 (ω, h, θ) 。

2.3.3 特征自适应模块

传统 RPN 中,锚点在整个特征图上是均匀分布的,设置 RPN 在每个位置的形状和比例都相同,特征图可保持较好的一致性。而 AR-RPN 中,锚框形状会随锚框位置的变化而变化,直接在特征图上进行全卷积的效果较差。理想情况下的锚框大小应根据目标尺寸的大小而变化,而特征自适应模块 N_T 利用可变形卷积的思想,添加了一个 3×3 可变形卷积层,以实现特征对齐,根据锚框特征生成新的特征图 F_1 。

$$y(P_0) = \sum \omega(P_n) \cdot x(P_0 + P_n + \Delta P_n), \quad (2)$$

式中, $x(P_n)$ 为特征图 F 中第 n 个位置的特征, P_n 为位置坐标, P_0 为初始位置坐标, ΔP_n 为该位置对应的锚框形状偏移量, $\omega(P_n)$ 为 3×3 的卷积核, $y(P_0)$ 为可变形卷积后的特征。 N_T 利用(2)式对原始特征图 F 进行可变形卷积,得到新的特征图 F_1 ,之后可以在 F_1 上进行进一步的分类和边界框回归。可变形卷积的示意图如图4所示。

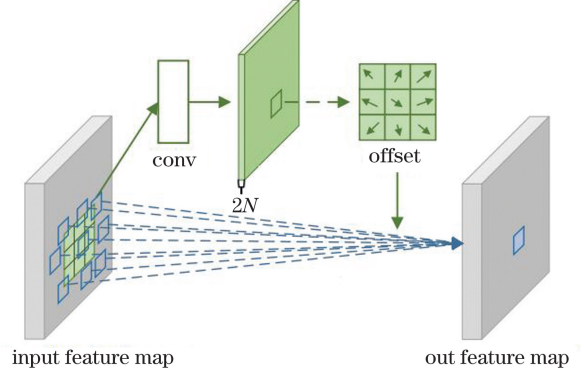


图4 可变形卷积示意图

Fig. 4 Schematic diagram of deformable convolution

3 损失函数

与传统 RPN 不同,AR-RPN 的损失函数在传统分类损失 L_{cls} 和回归损失 L_{reg} 的基础上,引入了锚框位置预测损失 L_{loc} 和锚框形状预测损失 L_{shape} 两种附加损失进行共同优化,总体损失函数可表示为

$$L = \lambda_1 L_{loc} + \lambda_2 L_{shape} + L_{cls} + L_{reg}, \quad (3)$$

式中,正则系数 $\lambda_1 = 1, \lambda_2 = 0.1$,其位置预测损失函数 L_{loc} 为 Sigmoid_focal_loss 损失,可表示为

$$\begin{cases} L_{loc} = \begin{cases} -\alpha [1 - \sigma(x)]^\gamma \log_e [\sigma(x)], & j+1 = y_{i,0} \\ -(1-\alpha) [\sigma(x)]^\gamma \log_e [1 - \sigma(x)], & \text{otherwise} \end{cases} \\ \sigma(x_{i,j}) = \frac{1}{1 + \exp(x_{i,j})} \end{cases} \quad (4)$$

式中, α 为平衡正样本和负样本的超参数,实验设置为0.25, γ 为平衡易分样本和难分样本的超参数,实验设置为2.0, $x_{i,j}$ 为样本的分类预测值, $y_{i,0}$ 为样本的分类目标值, i, j 为像素点在特征图中的位置。

锚框形状预测是通过计算锚框和标注框的 IoU 得到最优的 (ω, h, θ) ,传统的形状预测 IoU 损失不适用于 AR-RPN。原因是 (ω, h, θ) 为变量,用(5)式计算重叠面积时,目标重叠区域不是标准的长方形。如图5(a)中标注框 B_1 和候选框 B_2 的重叠区域 C_1 是一个不规则的多边形,需将其分解为多个三角形计算总面积,斜框的 IoU 可表示为

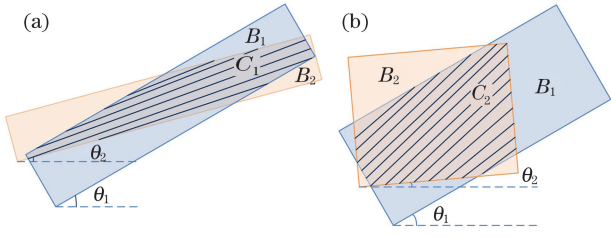


图 5 斜框的 IoU 示意图。(a)示意图 1;(b)示意图 2
Fig. 5 Schematic diagram of IoU in an oblique frame.

(a) Diagram1; (b) diagram2

$$X_{\text{IoU}} = \frac{X_{\text{area}}(B_1 \cap B_2)}{X_{\text{area}}(B_1 \cup B_2)} \quad (5)$$

$$\left\{ \begin{aligned} L_{\text{shape}} &= L_1 \left[1 - \min\left(\frac{w}{w_g}, \frac{w_g}{w}\right) \right] + L_1 \left[1 - \min\left(\frac{h}{h_g}, \frac{h_g}{h}\right) \right] + L_1 \left| \frac{1}{2\pi} [(\theta - \theta_g) \bmod 2\pi] \right| \\ L_1(x) &= \begin{cases} 0.5x^2, & \text{if } |x| < 1 \\ |x| - 0.5, & \text{otherwise} \end{cases} \end{aligned} \right. \quad (6)$$

式中, mod 为求余函数。

实验设定每 500 张图像进行一次评估, 设置位置阈值 ϕ_{loc} 为 0.5, 角度阈值 ϕ_{θ} 为 15° , 形状阈值 ϕ_{IoU} 为 0.6, 学习率为 0.02, 训练 24 个 epoch, 迭代 200000 次后, 最终损失函数的收敛结果如图 6 所示。

垂直框的 IoU 越大, 表示预测框与标注框的重合率越高, 因此, 可按照规定的阈值直接舍弃 IoU 较小的框。但斜框预测可能出现方向形状不一致但重叠面积较大的检测框, 如图 5(b) 所示, 此时 IoU 的筛选效果较差。先筛选出角度差 $\Delta\theta$ 小于阈值 ϕ_{θ} (实验取 15°) 的检测框, 进一步保留重叠面积 IoU 大于阈值 ϕ_{IoU} (实验取 0.6) 的检测框。

定义 AR-RPN 的 L_{shape} 为 Rotated_IoU_loss, 可表示为

4 分析与讨论

4.1 数据集预处理和评估指标

DOTA 数据集是目标检测中具有定向边界框标注的最大遥感图像数据集, 共 2806 张图像, 15 个类别, 包括大型车辆、油罐、足球场、舰船、飞机等, 完

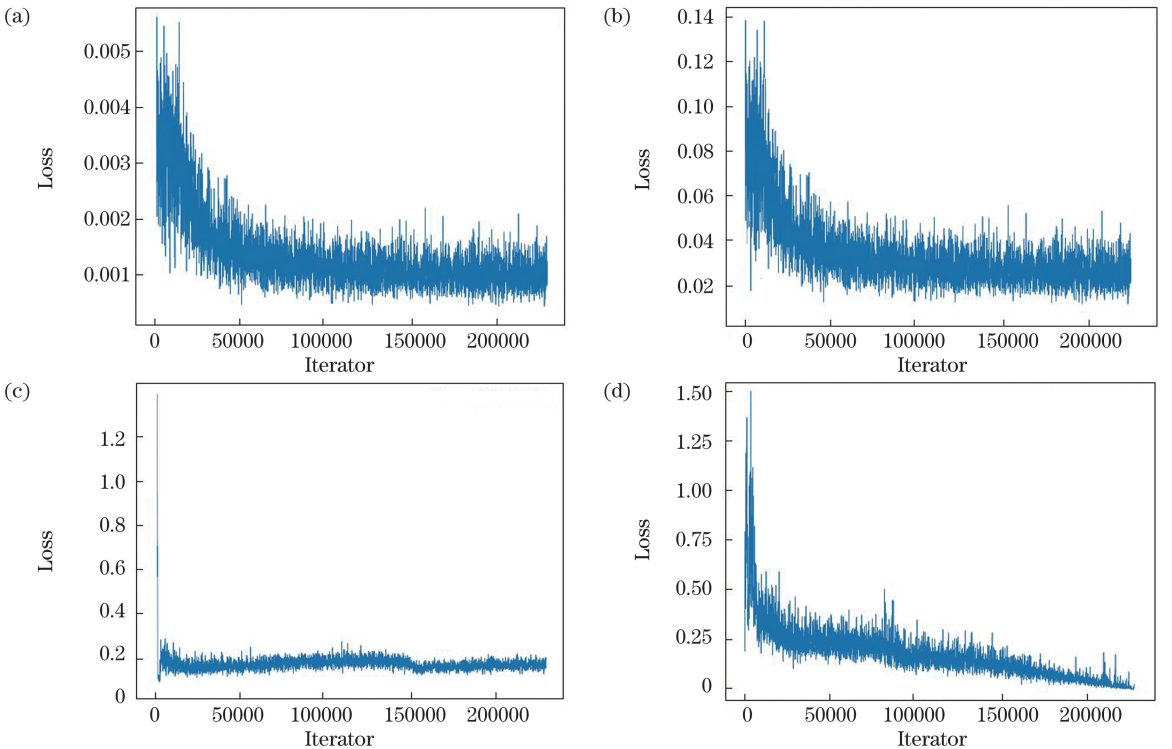


图 6 损失函数。(a)锚框位置损失;(b)锚框形状损失;(c)分类损失;(d)回归损失

Fig. 6 Loss function. (a) Anchor location loss; (b) anchor shape loss; (b) classification loss; (b) regression loss

整标注的 DOTA 图像包含 188282 个实例^[19]。实验主要针对舰船类的检测,在调整图像尺寸后,从原始图像中以 256 pixel 为步长裁剪一系列尺寸为 512 pixel×512 pixel 的图像。对于样本较少的舰船类别,从 4 个角度(0°,90°,180°,270°)进行旋转增强。

HRSC2016 数据集是具有挑战性的遥感舰船检测数据集,有 1061 张从 Google 地球收集的遥感图像。图像尺寸从 300 pixel×300 pixel 到 1500 pixel×900 pixel。训练集、验证集和测试集中分别包含 436、181、444 张图像,只采用水平移动进行数据增强,并将图像的尺寸调整为 512 pixel×800 pixel。

用平均精度(AP)定量评估舰船检测系统的性能^[20],可表示为

$$X_{AP} = \int_0^1 p(R) dR = \sum_{k=1}^N \max_{\tilde{k} \geq k} P(\tilde{k}) \Delta R(k), \quad (7)$$

式中, P 为准确率,即预测的正样本是实际正样本的概率, R 为召回率,即实际正样本中被预测为正样本的概率, N 为测试集中所有图像的数量, $P(\tilde{k})$ 为识别出 k 张图像时的最大 P , $\Delta R(k)$ 为识别图像数量从 $k-1$ 变化到 k 时(通过调整阈值) R 的变化情况。

4.2 实验结果分析

消融实验:基线网络(Baseline)基于 Faster R-CNN 结构,包括主干网络 ResNet101、FPN、RRPN、rotated ROI Align 和 rotated NMS。消融实验环境为 Ubuntu16.04, GPU 为 NVIDIA GEFORCE GTX 1080 Ti,内存为 11 GB,软件环境为 Python3.6.3、Pytorch 1.1。在 DOTA 数据集上进行消融实验,设置学习率为 0.02,每隔 500 张图像评估一次结果,训练 24 个 epoch 后的结果如表 1 所示。可以发现,对于 DOTA 数据集,利用改进 MFPN 可将网络的 AP 提高到 72.45%,MFPN 结合 AR-APN 后,AP 可达到 87.26%,最终得到本方法的 AP 为 88.64%,这表明本方法可以提升舰船目标的检测准确率。

表 1 消融实验的结果

Table 1 Result of ablation experiment

MFPN	—	✓	✓	✓
AR-RPN	—	—	✓	✓
ResNet+ResNext	—	—	—	✓
AP/%	68.13	72.45	87.26	88.64

图 7 为 Baseline 与本方法的实际检测结果,可以发现,RRPN 的检测结果中出现漏检、误检现象,且检测框准确性不高,而 AR-RPN 的检测结果更清楚。从图 7(b)可以看出,AR-RPN 能检测出倾斜角度较大且形状不一的舰船目标,图 7(a)中 RRPN 却漏检了这些目标,这表明 AR-RPN 更适应于遥感图像中舰船的形状分布与角度分布特性。图 7(c)中大小不一的舰船目标杂乱且密集分布时,RRPN 的漏检率较高,而 AR-RPN 和 MFPN 的组合能准确检测出密集排列的小目标,这表明 MFPN 可以提升网络的多尺度检测性能,尤其是对小目标检测效果的提升。

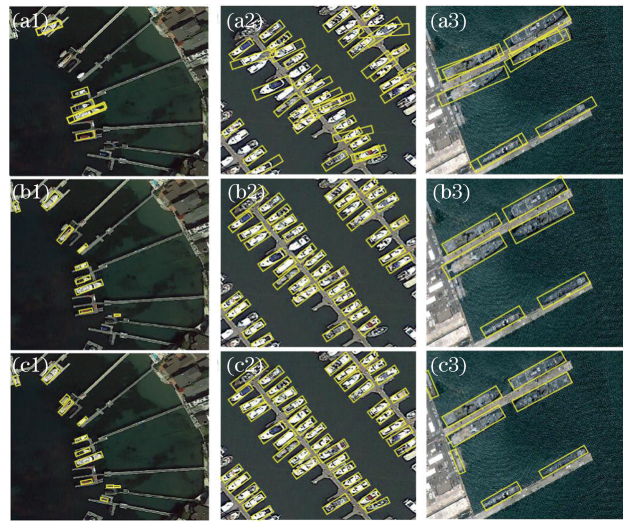


图 7 不同方法的检测结果。(a) RRPN; (b) AR-RPN; (c) MFPN

Fig. 7 Test results of different methods. (a) RRPN; (b) AR-RPN; (c) MFPN

4.3 对比实验

为了验证本方法的有效性,在 DOTA 数据集上将本方法与 Faster R-CNN,文献[20]中的方法 R^2 PN(Rotated region proposal and discrimination networks),AO-RCNN,CADNet(Context-aware detection network)^[21],ROI Trans 进行对比实验,结果如表 2 所示。可以看出,Faster R-CNN 的表现并不好,而 AO-RCNN 的检测效果有所提升, R^2 PN、CADNet 均额外引入了注意力机制或改进了损失函数,增加的模块会导致计算量增加,测试速度降低。针对舰船本身特性设计的 AR-RPN 更有利于遥感场景下的舰船检测,结合 MFPN 提升特征融合能力,最终达到了较好的检测效果。

表 2 不同方法在 DOTA 数据集上的 AP

Table 2 AP of different methods on the DOTA data set

Method	AP/%	Speed / (tasks · s ⁻¹)
Faster-RCNN ^[6]	38.30	10
AO-RCNN ^[14]	56.60	6.2
R ² PN ^[20]	68.13	5.7
CADNet ^[21]	76.70	5
RoI Trans ^[15]	83.59	5.9
Ours	88.64	6.4

在 HRSC2016 数据集上的实验结果如表 3 所示,可以发现,本方法的 AP 为 89.1%,相比 RC2、

RR-CNN、RRD^[22]、RoI Trans 这类基于斜框检测的方法,鲁棒性更强。

表 3 不同方法在 HRSC2016 数据集上的 AP

Table 3 AP of different methods on HRSC2016

Method	data set				unit: %
	RC2 ^[1]	RR-CNN ^[13]	RRD ^[22]	RoI Trans ^[15]	
AP	75.7	79.6	84.3	86.2	89.1

本方法在不同数据集上的检测结果如图 8 所示,可以发现,本方法在 DOTA 和 HRSC2016 数据集上均取得了较好的检测效果,可以检测密集排列的小目标,也可以检测角度各异的大目标。

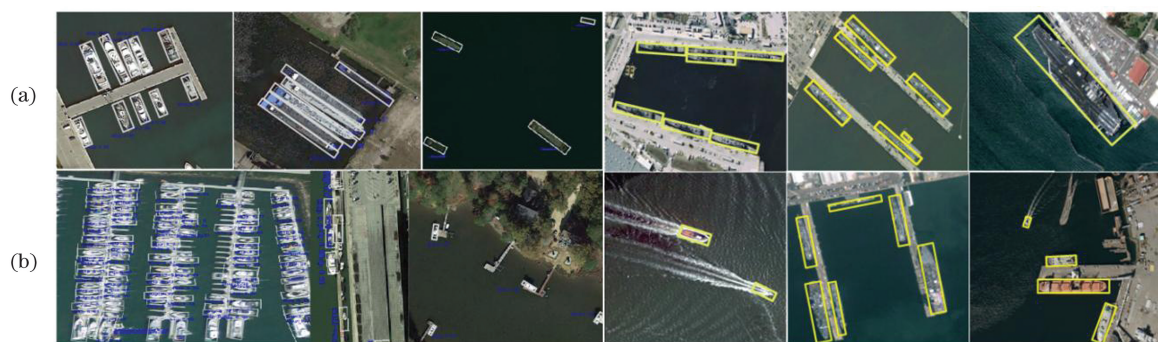


图 8 本方法在不同数据集上的检测结果。(a) HRSC2016 (b) DOTA

Fig. 8 Detection results of our method on different data sets. (a) HRSC2016 (b) DOTA

5 结 论

提出了一个以自适应旋转区域生成网络为核心部件的舰船目标检测框架,针对舰船本身的形状以及遥感俯拍图像舰船方向分布和尺度差异巨大的特点,设计了 MFPN,充分融合了底层细节和高层语义特征,利用底层感受野提升对小目标的检测性能;设计的 AR-RPN 可集中在目标中心位置生成锚框,减少冗余背景锚框的生成,在每个锚点处设计的长宽比例,更能适应舰船的极端长宽比。在传统分类损失和回归损失的基础上,引入了锚框位置预测损失和锚框形状预测损失两种附加损失,共同对网络进行优化。在不同数据集上将本方法与现有主流方法进行对比实验,结果表明,本方法可以有效提升遥感图像舰船目标检测的准确率,验证了本方法的有效性和泛化能力。

参 考 文 献

[1] 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.

[2] Ou P, Zhang Z, Lu K, et al. Object detection of remote sensing images based on convolutional neural networks [J]. Laser & Optoelectronics Progress, 2019, 56(5): 051002.

欧攀,张正,路奎,等.基于卷积神经网络的遥感图像目标检测[J].激光与光电子学进展,2019,56(5):051002.

[3] Du Z X, Yin J Y, Yang J. Remote sensing image detection based on dense connected networks [J]. Laser & Optoelectronics Progress, 2019, 56(22): 222803.

杜泽星,殷进勇,杨建.基于密集连接网络的遥感图像检测方法[J].激光与光电子学进展,2019,56(22):222803.

[4] Xu Z J, Huo Y H, Liu K, et al. Detection of ship targets in photoelectric images based on an improved recurrent attention convolutional neural network [J]. International Journal of Distributed Sensor Networks, 2020, 16(3): 155014772091295.

[5] Dai Y, Yi B S, Xiao J S, et al. Object detection of remote sensing image based on improved rotation

- region proposal network [J]. *Acta Optica Sinica*, 2020, 40(1): 0111020.
- 戴媛, 易本顺, 肖进胜, 等. 基于改进旋转区域生成网络的遥感图像目标检测 [J]. *光学学报*, 2020, 40(1): 0111020.
- [6] Nisa S U, Imran M. A critical review of object detection using convolution neural network[C]//2019 2nd International Conference on Communication, Computing and Digital systems (C-CODE), March 6-7, 2019, Islamabad, Pakistan. New York: IEEE, 2019: 154-159.
- [7] Liu W, Anguelov D, Erhan D, et al. SSD: single shot MultiBox detector[M]//Leibe B, Matas J, Sebe N, et al. *Computer Vision-ECCV 2016. Lecture Notes in Computer Science*. Cham: Springer, 2016, 9905: 21-37.
- [8] Redmon J, Divvala S, Girshick R, et al. You only look once: unified, real time object detection [C]// *Proceedings of the IEEE conference on computer vision and pattern recognition*, June 27-30, 2016, Las Vegas, NV. New York: IEEE, 2016: 779-788.
- [9] 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.
- [10] He K M, Gkioxari G, Dollár P, et al. Mask R-CNN [C] // *IEEE International Conference on Computer Vision*, October 22-29, 2017, Venice, Italy. New York: IEEE, 2017: 2980-2988.
- [11] Ma J Q, Shao W Y, Ye H, et al. Arbitrary-oriented scene text detection via rotation proposals [J]. *IEEE Transactions on Multimedia*, 2018, 20(11): 3111-3122.
- [12] Liu Z K, Hu J G, Weng L B, et al. Rotated region based CNN for ship detection [C] // *2017 IEEE International Conference on Image Processing (ICIP)*, September 17-20, 2017, Beijing, China. New York: IEEE, 2017: 900-904.
- [13] Yang X, Sun H, Fu K, et al. Automatic ship detection of remote sensing images from google earth in complex scenes based on multi-scale rotation dense feature pyramid networks [EB/OL]. [2020-05-21]. <https://arxiv.org/abs/1806.04331>.
- [14] Yang X, Sun H, Sun X, et al. Position detection and direction prediction for arbitrary-oriented ships via multitask rotation region convolutional neural network [J]. *IEEE Access*, 2018, 6: 50839-50849.
- [15] Ding J, Nan X, Yang L, et al. Learning RoI transformer for detecting oriented objects in aerial images [C] // *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, June 15-20, 2019, Long Beach, CA. New York: IEEE, 2019: 2844-2853.
- [16] 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)*, June 27-30, 2016, Las Vegas, Nevada. New York: IEEE, 2016: 770-778.
- [17] Dai J, Qi H, Xiong Y, et al. Deformable convolutional networks [C] // *2017 IEEE International Conference on Computer Vision (ICCV)*, October 22-29, 2017, Venice, Italy. New York: IEEE, 2017: 764-773.
- [18] Lin T Y, Dollár P, Girshick R, et al. Feature pyramid networks for object detection [C] // *2017 IEEE Conference on Computer Vision and Pattern Recognition*, July 21-26, 2017, Honolulu, HI, USA. New York: IEEE, 2017: 936-944.
- [19] Xia G S, Bai X, Ding J, et al. DOTA: a large-scale dataset for object detection in aerial images [EB/OL]. [2020-05-16]. <https://arxiv.org/abs/1711.10398>.
- [20] Zhang Z H, Guo W W, Zhu S N, et al. Toward arbitrary-oriented ship detection with rotated region proposal and discrimination networks [J]. *IEEE Geoscience and Remote Sensing Letters*, 2018, 15(11): 1745-1749.
- [21] Zhang G J, Lu S J, Zhang W. CAD-net: a context-aware detection network for objects in remote sensing imagery [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2019, 57(12): 10015-10024.
- [22] Liao M H, Zhu Z, Shi B G, et al. Rotation-sensitive regression for oriented scene text detection [EB/OL]. [2020-05-24]. <https://arxiv.org/abs/1803>.