

基于双重特征增强的遥感舰船小目标检测

徐志京, 柏雪*

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小目标占比大、多方向旋转等问题, 提出一种基于纹理和颜色双重特征增强的舰船小目标检测方法。首先, 设计生成对抗网络增强舰船小目标的纹理特征, 生成高分辨率的舰船图像。然后, 采用深度强化学习算法增强图像的颜色信息, 解决舰船目标与背景颜色低对比度的问题。接着, 设计自适应变换特征金字塔网络, 增强全局感受野, 有效解决深层网络中空间信息缺乏导致的小目标特征难以提取的问题。最后, 利用特征细化模块和圆形光滑标签完成对舰船目标边界框的特征点对齐和角度回归, 有效提高多方向旋转舰船目标的检测精度。此外, 在 HRSC2016 和 DOTA 两个公共数据集上进行了相关实验。可以发现, 所提方法在两个数据集上的平均精度均值可分别达到 72.87% 和 89.91%, 相比主流的目标检测算法得到了大幅提升。

关键词 遥感; 遥感图像; 舰船检测; 双特征增强; 小目标; 深度 Q 网络

中图分类号 TP391.4

文献标志码 A

DOI: 10.3788/AOS202242.1828002

Small Ship Target Detection Method for Remote Sensing Images Based on Dual Feature Enhancement

Xu Zhijing, Bai Xue*

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

Abstract In terms of the large proportion and multidirectional rotation of small ship targets in remote sensing images, a small ship target detection method based on texture and color enhancement is proposed. Firstly, a generative adversarial network is designed to enhance the texture features of small ship targets and generate high-resolution ship images. Secondly, the deep reinforcement learning algorithm is used to improve the image color, which solves the problem of the low contrast between the ship target and the background color. Thirdly, an adaptive transform feature pyramid network is designed to enhance the global receptive field and effectively deal with the hard extraction of small target features, which is caused by the lack of spatial information in the deep network. Finally, the feature refinement module and circular smooth label are utilized to align feature points and achieve angle regression in a ship target bounding box, which effectively increases the accuracy of detecting ship targets with multidirectional rotation. In addition, related tests are carried out on the public data sets of HRSC2016 and DOTA, and results show that the proposed method achieves an mean average precision of 72.87% and 89.91%, respectively, which is better than the existing mainstream small ship target detection methods.

Key words remote sensing; remote sensing images; ship detection; dual feature enhancement; small target; deep Q-network

1 引言

随着空间遥感检测技术的发展, 海洋监测中的舰船小目标检测已成为光学遥感领域中的研究热点^[1-3]。借助卫星遥感观测平台或飞机拍摄成像, 能够获取高分辨率和内容丰富的遥感图像, 这为海域内的舰船小

目标检测提供了可能。受限于遥感图像的图面距离, 地物的宏观信息在图像中表达时会出现部分区域的地表特征识别模糊、纹理细节信息缺失和目标易读性较差等问题。在遥感舰船应用场景中, 舰船外观特征所占空间范围很小(低至 10 pixel), 外表轮廓十分模糊, 同时舰船小目标的纹理和颜色信息往往隐藏于背景

收稿日期: 2022-01-07; 修回日期: 2022-02-17; 录用日期: 2022-02-28

基金项目: 国家重点研发计划(2019YFB1600605)、上海市扬帆计划(20YF1416700)

通信作者: *1185685025@qq.com

中,导致图像在噪声预处理过程中可能会丢失所需要的细节信息,这些问题给遥感舰船小目标检测带来了巨大的挑战。一般地,国际光学工程学会将物理尺寸小于 $256 \text{ pixel} \times 256 \text{ pixel}$ 的 12% 的目标定义为小目标,MS COCO^[4]中将尺寸小于 $32 \text{ pixel} \times 32 \text{ pixel}$ 的目标定义为小目标,本文将研究的 HRSC2016^[5]和 DOTA^[6]数据集中像素区域占比小于 0.03 的目标定义为小目标。

近年来,深度学习技术被广泛应用于小目标检测领域中^[7-9],取得了很好的效果。考虑到小目标包含的有效特征较少,一些研究人员试图通过增加图像分辨率的方法来提高目标的检测性能。Bai等^[10]设计了通用的多任务生成对抗网络框架,并与目标检测器相结合,以分离场景中的目标信息与背景信息。Rabbi等^[11]基于增强型超分辨率生成对抗网络(ESRGAN)^[12]将检测器损失反向传播到边缘增强型超分辨率产生式对抗网络(EESRGAN)中,提出了一种新的EESRGAN。Bashir和Wang^[13]结合循环生成对抗网络和残差特征聚合来改进分段路由(SR)框架,同时利用YOLO(You Only Look Once)检测器显著提高了检测性能。Xi等^[14]通过同时恢复高分辨率图像中高频分量和低频分量的缺失信息,提出了一种双流表示学习生成对抗网络,提高了低分辨率航空图像中的小目标检测精度。

同时,霍青兰^[15]指出由于环境因素的影响,获取的遥感图像容易在亮度或者色彩上出现分布不均匀的现象,这会直接影响生成的遥感图像的质量,因此经过人工调整颜色后遥感图像能展现出良好的视觉效果。Park等^[16]将高质量的图像作为参考,实现了一种与人类修图过程相似的颜色增强方法。Dharejo等^[17]利用暗通道增强算法和色彩校正直方图均衡处理技术改善了雾度和噪声对图像分辨率的影响。Xu等^[18]联合学习SR重建技术和感知图像增强(PIE)技术提高颜色和亮度对比度以补充图像的细节,进而提高图像分辨率。

当前的小目标增强算法大多是通过将生成对抗网络与目标检测算法相结合的方式来提升原始图像分辨

率的,并没有考虑到遥感海面舰船目标尺寸较小、多方向旋转和图像对比度低的特点。因此,本文设计了一种基于纹理和颜色双重增强的生成对抗网络(DEGAN)的遥感舰船小目标检测算法。在预处理阶段,设计纹理增强网络(TEN)提高遥感舰船图像的分辨率,以增强舰船小目标的边缘纹理信息。同时,设计颜色增强网络(CEN),以解决停靠在港口码头的内陆水域中的舰船与港口背景颜色相似的问题,进而改善图像亮度和对比度,提升舰船图像的视觉质量。在目标检测阶段,设计注意力聚合检测(A^2 Det)网络,通过自适应变换特征金字塔网络(ATFPN)充分融合语义信息和区域纹理特征,其中嵌入的坐标注意力(CA)模块^[19]能够自适应地调节感受野大小,更好地获得全局感受野。同时,在边界框粗调与细调阶段间加入特征细化(FR)模块以完成特征点的对齐,并利用圆形光滑标签(CSL)^[20]校正舰船目标的旋转角度,提高了对多方向旋转、密集排列的舰船目标的检测精度。

2 模型设计

本文设计的DEGAN小目标检测模型如图1所示。该模型由三个部分组成,即TEN、CEN和 A^2 Det网络。在将遥感舰船图像送入检测网络之前,对图像进行预处理。首先,将原始遥感图像送入TEN中获得纹理增强的SR图像,增强舰船小目标的边缘纹理特征。然后,将超分辨率图像送入CEN中,采用4层全连接层组成的颜色感知器进一步提高舰船目标与海面背景灰度、港口背景间的对比度,提高整幅图像的亮度。最后,设计 A^2 Det网络,在ATFPN中通过自上而下的路径拼接高层语义信息和低层空间特征,以丰富舰船目标的特征信息表达,并通过自下而上的路径增强低层定位信息,这有利于下一步的舰船小目标检测。每层的特征金字塔分别连接分类子网络和回归子网络,对输出的两个子网络组合进行特征重建,将新的特征图再次输入分类子网络和回归子网络中,即完成一次FR,设定两个阈值不同的级联细化模块,最终可输出舰船目标的检测结果。

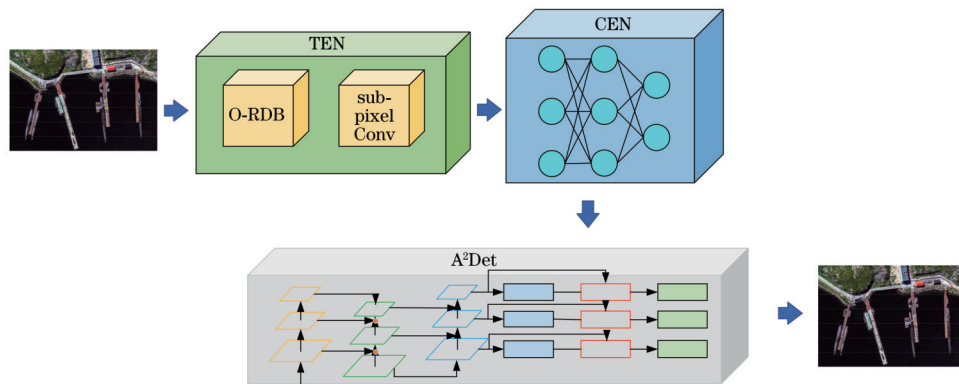


图1 DEGAN小目标检测模型

Fig. 1 DEGAN small object detection model

2.1 纹理增强网络

TEN 结构如图 2 所示,将级联的 Octave 卷积^[21]残差密集块(O-RDB)作为生成器网络的基本模块,这些密集连接的模块增加了网络的容量,保证了训练的稳定性。输入的遥感舰船图像先经过一个卷积块以提取

舰船图像的浅层特征。然后,利用多个密集连接的 O-RDB 残差嵌套结构和长跳跃连接来映射其深层特征。最后,利用两次子像素卷积(sub-pixel Conv)来发掘级联的 O-RDB 中更深层次的特征信息,进而生成具有丰富纹理信息的 4 倍放大的 SR 图像。

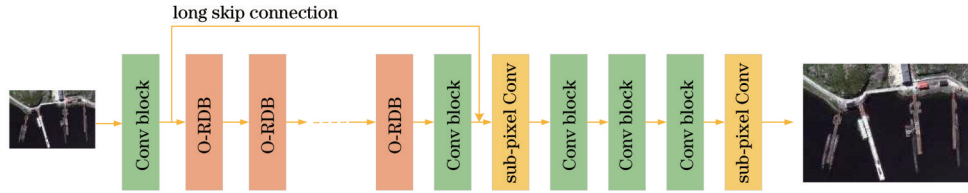


图 2 纹理增强模型结构图

Fig. 2 Structural diagram of texture enhancement model

深层特征映射部分的 O-RDB 结构由三个 Octave 卷积密集模块(ODB)和一个短跳跃连接构成,如图 3 所示,其中 β 为加权系数。一个 ODB 由 4 个密集级联的 Octave 卷积组合而成,Octave 卷积能够利用空间多尺度特征在 FirstOctConv 层中将图像分解成高频和低频两部分,降低低频特征图的分辨率能够使网络具有更大的感受野,同时能够有效减少存储量和计算量。接下来,利用线性整流函数(ReLU)学习密集块的参数。最后,在 LastOctConv 层中完成特征融合,进而提高图像 SR 的重构性能。

频两部分,降低低频特征图的分辨率能够使网络具有更大的感受野,同时能够有效减少存储量和计算量。接下来,利用线性整流函数(ReLU)学习密集块的参数。最后,在 LastOctConv 层中完成特征融合,进而提高图像 SR 的重构性能。

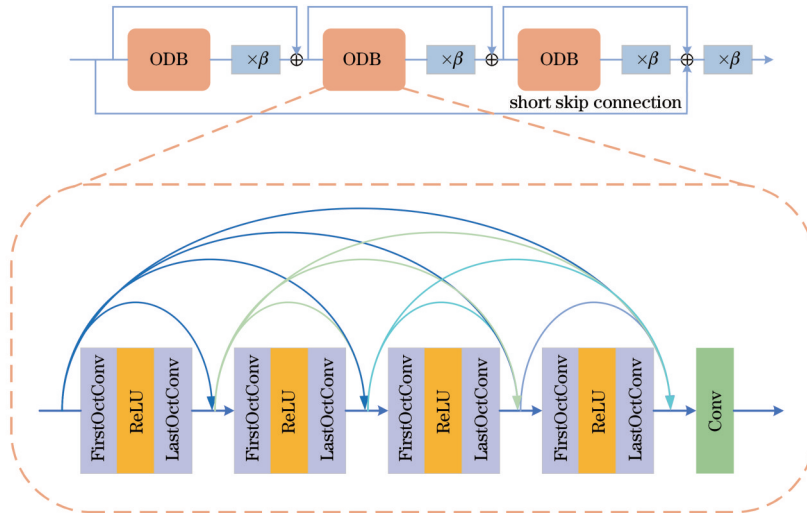


图 3 生成器中的 O-RDB 结构

Fig. 3 O-RDB structure in generator

2.2 颜色增强网络

CEN 结构如图 4 所示。通过 ResNet50 网络提取上一阶段的超分辨率图像的上下文特征 $F_{context}$,同时采用 CIELUV 颜色直方图提取超分辨率图像的颜色特征 F_{color} 。将获取的双重特征作为当前图像的状态 $S(F_{context}, F_{color})$,再在 4 层全连接层组成的代理网络中反复应用一个调整动作来增强输入图像 I 的颜色信息,并估计每个预定义调整动作 A 的操作值 $Q[S(t), A]$,最终选择具有最优序列的动作来增强图像的颜色。颜色增强的目标就是找到一个能让 $\|I(t_{final}) - I_{target}\|_2^2$ 最小化的最佳操作序列,其中 $I(t_{final})$ 为最终强化图像, t_{final} 为最终生成图像时间, I_{target} 为真实图像。同时,对完成操作后的图片进行一次前馈,计算整个网络的最大输出值,以更新目标 Q 值。此时,在

深度强化学习中深度 Q 网络(DQN)^[22]算法框架下,代理网络能够训练出最大化奖励 $\max\{Q[S(t), A]\}$ 。接着,利用反向传播更新代理网络权重,不断重复此过程,并在所有估计操作值均为负值时停止。

颜色增强的过程可以表述为

$$R(t) = \|I_{target} - I(t-1)\|_2^2 - \|I_{target} - I(t)\|_2^2, \quad (1)$$

前馈的过程可以描述为

$$Q(t) = R(t) + \max\{Q[S(t), A]\}, \quad (2)$$

式中: $R(t)$ 为依据欧氏距离的实时奖励; t 为当前时间; $\max(\cdot)$ 为取最大值函数。

动作颜色调整操作表如表 1 所示,将亮度、对比度和色彩饱和度等动作编码的颜色调整过程转化为马尔可夫决策过程(MDP)^[23],这是一种简单而直接的类似

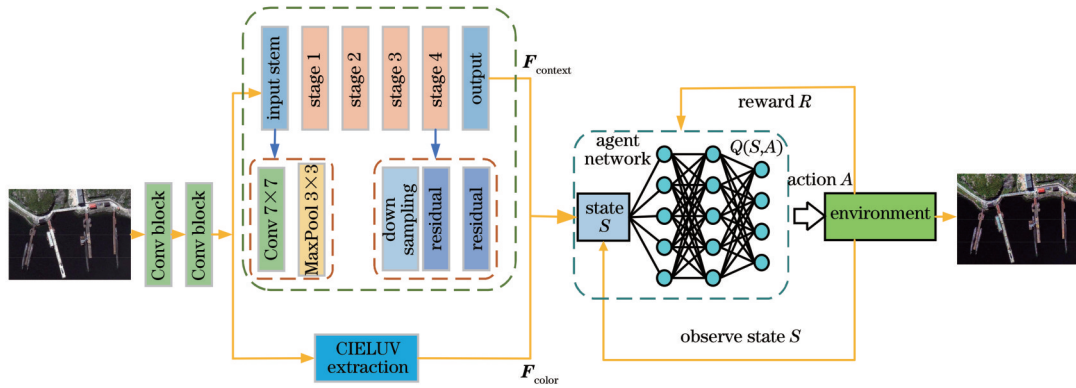


图 4 CEN 结构

Fig. 4 Structural diagram of CEN

SR 领域中失真-恢复的训练方案,可以避免寻找成对的遥感色彩增强数据集。同时,直接将 CEN 应用于

SR 图像中,可以降低使用成本。表 1 中括号中的内容为调整图片对应参数的倍数。

表 1 动作颜色调整操作

Table 1 Action color adjustment operation

Action code	Action description	Action code	Action description
1	Contrast(0.95×)	2	Contrast(1.05×)
3	Color saturation(0.95×)	4	Color saturation(1.05×)
5	Brightness(0.95×)	6	Brightness(1.05×)
7	Red+green(0.95×)	8	Red+green(1.05×)
9	Green+blue(0.95×)	10	Green+blue(1.05×)
11	Red+blue(0.95×)	12	Red+blue(1.05×)

在训练过程中,本文使用贪婪算法策略。贪婪算法以概率 σ 随机选择下一次动作,以概率 $1 - \sigma$ 的贪婪方式选择预期奖励最高的动作。在推理阶段中,使用 σ 为 0 的贪婪算法策略,即总是选择最高的预期收益,并重复这一过程,当所有预期回报均为负值时可以得到颜色增强后的图像。

2.3 注意力聚合检测网络

由于加深网络深度会造成小目标本身的细节信息缺失,浅层网络有利于定位舰船小目标,而深层网络更有利于获取大中型舰船目标的语义信息,因此设计了拓宽网络宽度的 A²Det 网络结构,如图 5 所示。整个网络结构可分为 ATFPN 模块和 FR 模块两个部分。先

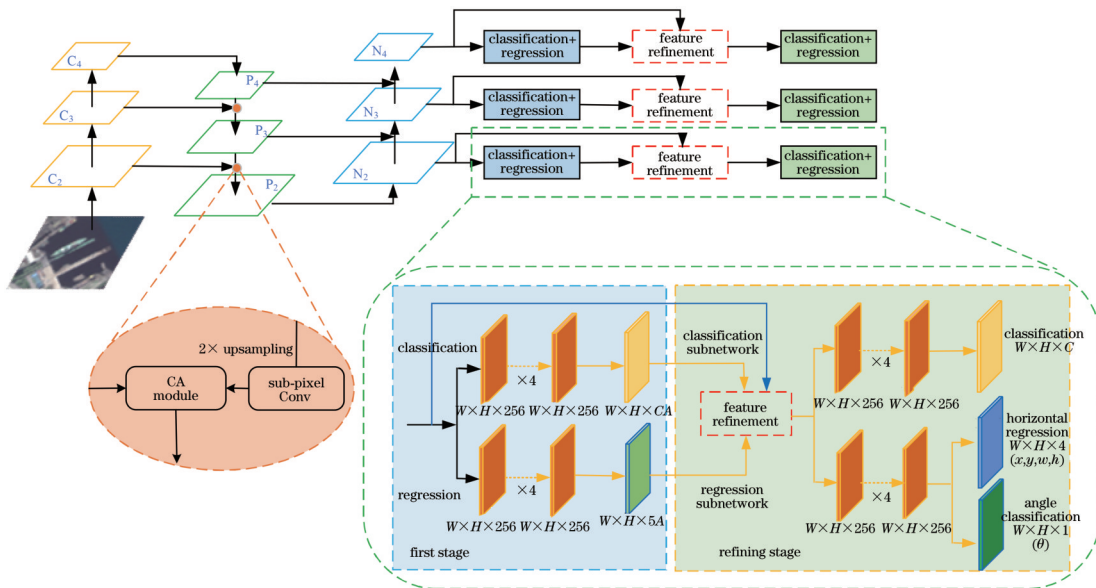


图 5 A²Det 网络结构

Fig. 5 A²Det network structure

将双重特征增强后的图像数据送入骨干网络中提取特征,并将得到的 C_2, C_3, C_4 特征图用来构建 ATFPN 模块。因为 ATFPN 的不同特征图具有不同的感受野,所以不宜采用共享权值的方法完成多尺度目标的位置信息编码。因此,在自上而下的 P_2, P_3, P_4 特征融合路径中加入子像素卷积完成 2 倍上采样,将上层特征映射图传递到当前层中,以融合不同层级的空间上下文信息,进而丰富了特征金字塔的多尺度特征表达。同时,嵌入 CA 模块,从位置通道和空间通道两方面对方向特征进行聚合,以学习舰船目标位置的敏感特征,使网络更容易注意到舰船小目标的位置。接下来,利用自下而上的 N_2, N_3, N_4 聚合路径融合低层和高层之间

的信息,以增强舰船小目标的定位信息流。每层的 ATFPN 分别连接分类和回归两个子网络,利用 FR 模块对输出的组合子网络进行特征重建,完成两次 FR 后输出最终的旋转舰船目标检测结果。图 5 中 W 为特征图的宽度, H 为特征图的高度, A 为每个特征点上的锚框数量, C 为类别数。

注意力聚合过程如图 6 所示。首先,利用子像素卷积分别扩展自顶向下的卷积层 $C_{l+1} (l=1, 2)$ 的分辨率和感受野。然后,利用 CA 模块捕获特征图位置的敏感信息。最后,将处理后的特征图与前一阶段横向连接的卷积层 $C_l (l=2, 3)$ 串联完成特征聚合,再使用 1×1 卷积层将特征通道数减半至 256。

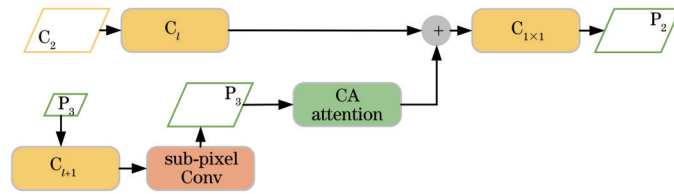


图 6 注意力聚合过程

Fig. 6 Attention aggregation process

考虑到旋转锚框轻微的角度变化会造成倾斜交并比分数骤降,从而可能丢失舰船目标。因此,需要在保证检测速度的前提下完成特征图的重建,如图 7 所示。首先,计算前一阶段分类和回归 2 个子网络得到的置信度最高的锚框。然后,通过特征插值对锚框中心位

置所确定的 5 个点(1 个中心点, 4 个角点)进行重新编码,并将计算结果映射回特征图对应的特征点上。最后,将计算结果叠加在原始特征图上以改变对应特征点的像素值,进而完成特征图的重建,在最终的分类型子网络和回归子网络中解决了特征错位的问题。

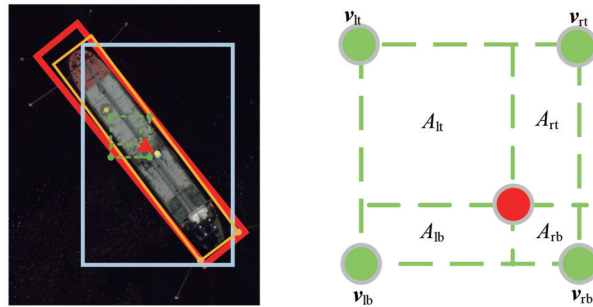


图 7 特征重建示意图

Fig. 7 Schematic diagram of feature reconstruction

在第一阶段生成水平检测框的条件下,按照如下的双线性插值方法设置所需的两个级联 FR 模块,即

$$v = v_{lt}A_{rb} + v_{rt}A_{lb} + v_{lb}A_{rt} + v_{rb}A_{lt}, \quad (3)$$

式中: $v_{lt}, v_{lb}, v_{rt}, v_{rb}$ 分别为左上、左下、右上、右下对应特征点的坐标值; $A_{lt}, A_{lb}, A_{rt}, A_{rb}$ 分别为左上区域、左下区域、右上区域、右下区域的面积。

2.4 损失函数

采用 5 个坐标参数 (x, y, w, h, θ) 表示任意方向的矩形包围盒,同时在回归子网中加入角度分类预测损失函数,这种做法解决了模型在边界处损失值突增的问题。在算法中,对旋转边界框的回归模型的定义为

$$\begin{cases} t_x = (x - x_a) / w_a \\ t_y = (y - y_a) / h_a \end{cases}, \quad (4)$$

$$\begin{cases} t_w = \ln(w / w_a) \\ t_h = \ln(h / h_a) \\ t_\theta = \theta - \theta_a \end{cases}, \quad (5)$$

$$\begin{cases} t'_x = (x' - x_a) / w_a \\ t'_y = (y' - y_a) / h_a \end{cases}, \quad (6)$$

$$\begin{cases} t'_w = \ln(w' / w_a) \\ t'_h = \ln(h' / h_a) \\ t'_\theta = \theta' - \theta_a \end{cases}, \quad (7)$$

式中: (x, y) 为真实框的中心坐标; w 为真实框的宽度; h 为真实框的高度; θ 为真实框的角度; (x', y') 为预测框的中心坐标; w' 为预测框的宽度; h' 为预测框的高度; θ' 为预测框的角度; (x_a, y_a) 为候选框的中心坐标;

w_a 为候选框的宽度; h_a 为候选框的高度; θ_a 为候选框的角度; t_x 为真实的水平平移变换; t_y 为真实的垂直平移变换; t_w 为真实的横向尺度缩放; t_h 为真实的纵向尺度缩放; t_θ 为真实的角度偏移; t'_x 为预测的水平平移变换; t'_y 为预测的垂直平移变换; t'_w 为预测的横向尺度缩放; t'_h 为预测的纵向尺度缩放; t'_θ 为预测的角度偏移。

由于遥感舰船图像包含大量密集排列、背景复杂的旋转船只, 细微的角度变化会使交并比分数迅速降低从而丢失舰船目标, 因此边界框的回归效果对旋转舰船的检测精度有很大的影响。利用 CSL 处理不同类别的角度损失值问题, 将角度的计算从回归问题转化为分类问题, 完成了对旋转舰船目标边界框的校正, 进而解决了边界位置损失值突然增加的问题。CSL 的计算公式为

$$L_{\text{cs}}(x) = \begin{cases} g(x), & \theta - r < x < \theta + r \\ 0, & \text{otherwise} \end{cases}, \quad (8)$$

式中: $g(x)$ 为窗口函数, 用于测量真实值和预测值之间的角度距离; r 为窗口半径。CSL 示意图如图 8 所示。

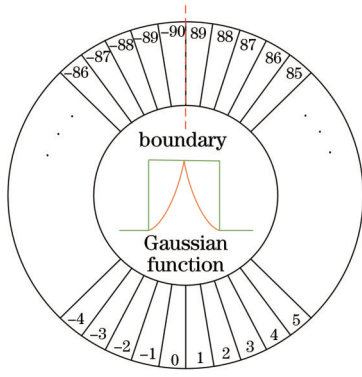


图 8 CSL 示意图

Fig. 8 Schematic diagram of CSL

多任务损失函数的定义公式为

$$L = \frac{1}{N} \left[\lambda_1 \sum_{n=1}^N O_n \sum_{j \in \{x, y, w, h\}} L_{\text{reg}}(\mathbf{v}'_{nj}, \mathbf{v}_{nj}) + \lambda_2 \sum_{n=1}^N L_{\text{cls}}(p_n, t_n) \right] + \lambda_3 \sum_{n=1}^N L_{\text{csl}}(Q'_n, Q_n), \quad (9)$$

式中: L_{reg} 为矩形框中心坐标、宽度和高度 4 个参数的回归损失函数; L_{cls} 为类别损失函数; L_{csl} 为角度损失函数; N 为所有预测框的数量; O_n 为一个二进制的值 ($O_n=1$ 代表前景, $O_n=0$ 代表背景); \mathbf{v}'_{nj} 为预测的偏移向量; \mathbf{v}_{nj} 为真实的目标向量; Q'_n 和 Q_n 分别为角度的标签和预测; t_n 为目标的标签; p_n 为用 Sigmoid 函数计算的各个类别的概率分布; $\lambda_1, \lambda_2, \lambda_3$ 为多任务损失超参数, 分别设置为 1.0、0.5、1.0, 用以控制三个任务损失之间的平衡。

L_{reg} 的表达式为

$$L_{\text{reg}}(\mathbf{v}'_{nj}, \mathbf{v}_{nj}) = \text{smoothL1}(|\Delta \mathbf{v}|) = \begin{cases} \frac{1}{2}(|\Delta \mathbf{v}|)^2, & |\Delta \mathbf{v}| \leq 1 \\ |\Delta \mathbf{v}| - 0.5, & |\Delta \mathbf{v}| > 1 \end{cases}, \quad (10)$$

式中: $\text{smoothL1}(\cdot)$ 为损失函数; $|\Delta \mathbf{v}|$ 为真实向量和预测向量差的模。

分类损失函数由类别损失函数 L_{cls} 和角度损失函数 L_{csl} 组合而成, L_{cls} 和 L_{csl} 的表达式为

$$L_{\text{cls}}(p_n, t_n) = F_L(p_n, t_n), \quad (11)$$

$$L_{\text{csl}}(Q'_n, Q_n) = F_L[L_{\text{cs}}(\theta'_n), L_{\text{cs}}(\theta_n)], \quad (12)$$

$$F_L = \begin{cases} -\alpha(1-p)^\gamma \ln p, & \gamma = 1 \\ -(1-\alpha)p^\gamma \ln(1-p), & \gamma = 0 \end{cases}, \quad (13)$$

式中: $L_{\text{cs}}(\theta'_n)$ 和 $L_{\text{cs}}(\theta_n)$ 分别为预测角度和真实角度对应的圆形平滑标签数值; α 为加权因子; $(1-p)^\gamma$ 为调节因子。

此外, 引入加权因子 $\alpha \in [0, 1]$ 和调节因子 $(1-p)^\gamma$ ($\gamma=0$) 来控制正样本和负样本的权重, 此处旨在说明训练更侧重于正样本。

3 分析与讨论

3.1 实验环境及过程

本文实验的环境配置为: 操作系统是 Ubuntu18.04, 中央处理器 (CPU) 是 I7-7700x, 图形处理器 (GPU) 是 NVIDIA GTX1080Ti, 内存是 32 GB。所有实验均在 Pytorch 框架下进行, 训练过程中均采用 ResNet50 作为主干网络。测试数据集选用两个在遥感舰船检测领域中广泛使用的 DOTA 和 HRSC2016 数据集, 并将 VOC2007 平均精度均值 (mAP) 作为实验的评估指标。

测试和训练过程中的相关参数设置为: 输入的图片大小为 800 pixel \times 800 pixel, 实验中的 O-RDB 包含 23 个 ODB 模块, 网络学习率为 0.0002, 批大小为 2, DQN 的推理最大次数为 12, 周期数为 200。将区域生成网络第一阶段交并比的正样本阈值设为 0.6, 负样本阈值设为 0.5, 保证了正负样本比例。在两次级联的 FR 阶段中, 将正负样本阈值设为 0.7 和 0.6。

3.2 DOTA 数据集实验结果分析

DOTA 数据集包含 2806 张航空图像, 图像分辨率高达 4000 pixel \times 4000 pixel, 总共包含 15 个目标类别, 188282 个实例。在实验过程中, 将数据集的舰船实例边界框所构成的任意四边形的左上角作为顶点, 按照顺时针方式 ($x_1, y_1, x_2, y_2, x_3, y_3, x_4, y_4$) 进行标注, 并将标注好的图像储存在相应的文件中。在标注图中可以看出, 舰船目标的旋转角度分布均衡, 但存在大量尺寸极小的舰船目标。将原始图像中的 1411 个作为训练集, 458 个作为验证集, 937 个作为测试集。因为

DOTA 数据集中原始图像尺寸过大,所以直接训练难度大。因此,利用一个步长为 512 的滑动窗口将图像裁剪到 $600 \text{ pixel} \times 600 \text{ pixel}$ 大小,裁剪后的数据集中总共有 27031 张图像,其中含有舰船的图像 4620 张,舰船目标总数达 76833 个。

分析表 2 结果可知,所提方法相比于 CSL^[20]、

RsDet^[24]、RiDet^[25]、R³Det^[26] 等单阶段检测算法和以 SCRDet^[27] 为代表的两阶段检测算法获得了更高的检测精度,mAP 达到了 72.87%。所提算法在 DOTA 数据集上的检测结果如图 9 所示。可以发现,所提算法可以准确检测具有方向性的密集排列的舰船小目标,并且几乎没有漏检、错检。

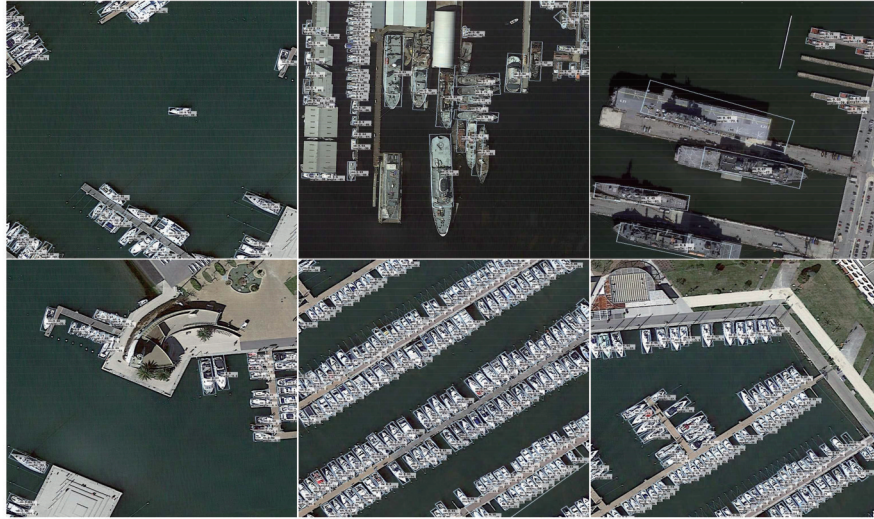


图 9 DOTA 数据集部分测试结果

Fig. 9 Partial test results of DOTA data set

表 2 DOTA 数据集对比实验结果

Table 2 Comparative experimental results in DOTA data set

Method	mAP / %
RetinaNet ^[28]	65.73
CSL ^[20]	67.38
RsDet ^[24]	66.05
RiDet ^[25]	66.06
SCRDet ^[27]	71.73
R ³ Det ^[26]	70.66
Proposed (DEGAN)	72.87

3.3 HRSC2016 数据集实验结果分析

HRSC2016 是目前唯一公开的专门用于光学遥感船舶实验的高分辨率船舶数据集,所有图像均来自 Google Earth 中 6 个著名的港口,总共包含 2976 张船舶目标数,船只实例基本分布在开放海域和近岸港口处。图像分辨率在 $0.4 \sim 2.0 \text{ m}$ 之间,尺寸范围在 $300 \text{ pixel} \times 300 \text{ pixel} \sim 1500 \text{ pixel} \times 900 \text{ pixel}$ 之间。数据集采用旋转包围盒标注的方式,总共包含 2976 张船舶目标数。训练集、验证集和测试集分别包含 436 个图像(包括 1207 个样本)、181 个图像(包括 541 个样本)和 444 个图像(包括 1228 个样本)。经计算,大约有 $2/3$ 的图像中舰船目标占图像面积的比例小于 0.03,这有利于舰船小目标的检测研究。

如表 3 所示,将传统的深度学习方法 RetinaNet^[28]、场景文本检测方法 R²CNN^[29]、面向任意目标的旋转检测器(如 RiDet^[25]、RsDet^[24]、R³Det^[26] 和

ReDet^[30]) 三类方法与所提算法的结果进行对比。可以看出,所提算法的 mAP 达到了 89.91%,相比 R³Det 检测器,检测精度提高了 3.39 个百分点。所提方法在 HRSC2016 数据集上的检测结果如图 10 所示,在大小舰船目标混合的遥感图像中,可以准确检测小型船只,这是因为在预处理阶段显著增强了舰船小目标的纹理特征和颜色特征,提高了图像的视觉特征。由于生成对抗网络的存在,故所提算法的检测速度远低于二阶段网络,但相比其他算法,其平均检测精度有较大提升。

表 3 HRSC2016 数据集对比实验结果

Table 3 Comparative experimental results in HRSC2016 data set

Method	mAP / %	Speed / (frame · s ⁻¹)
R ² CNN ^[30]	72.98	6.00
RetinaNet ^[28]	82.57	19.48
RiDet ^[25]	83.78	25.23
RsDet ^[24]	83.47	25.72
R ³ Det ^[26]	86.52	20.27
ReDet ^[30]	88.06	4.20
Proposed (DEGAN)	89.91	16.37

3.4 消融实验

本文在 HRSC2016 数据集上进行消融实验。实验结果如表 4 所示,可以看出:以 R³Det 为基准实验,模块 1 利用 TEN 对舰船小目标进行 SR 重建,检测精度提高了 0.93 个百分点;模块 2 在模块 1 的前提下结合

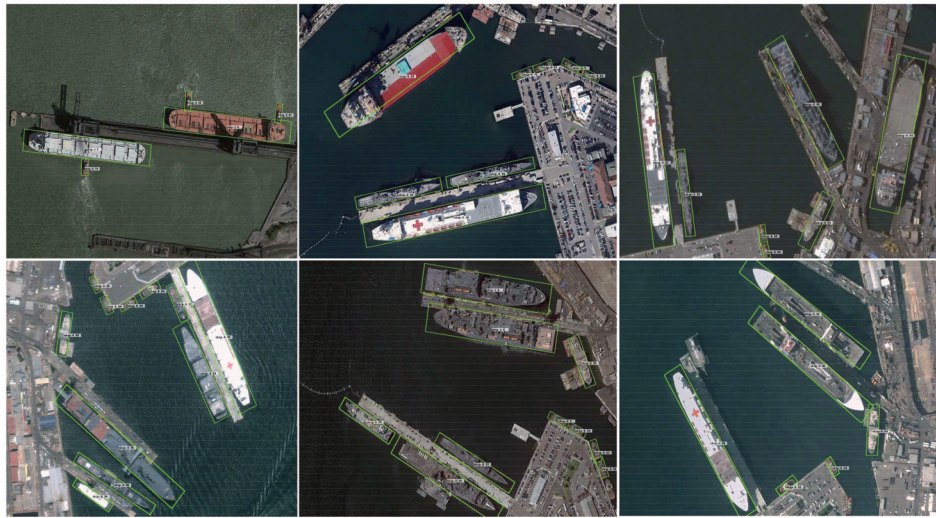


图 10 HRSC2016 数据集部分测试结果

Fig. 10 Partial test results of HRSC2016 data set

CEN, 检测精度提高了 1.40 个百分点; 模块 3 利用改进 ATFPN 可将网络的 mAP 提高 2.52 个百分点; 模块 4 利用 CSL 改进分类损失函数, 检测精度提高了 3.39 个百分点。实验结果表明, 所提算法针对遥感舰船小目标检测具有更强的鲁棒性。图 11 展示了所提方法 (每幅图的左侧) 和基准实验 (每幅图的右侧) 在 HRSC2016 数据集上的消融实验结果, 其中每一行代表使用不同模块增强后的检测效果对比。图 11(a1)、(a2) 说明通过增强船舶目标的纹理特征, 可解决检测过程中目标尺寸差异大所导致的漏检小目标的问题。图 11(b1)、(b2) 说明通过增强图像颜色信息, 视觉质量较差的遥感图像中的船舶检测精度也有明显提升。

图 11(c1)、(c2) 说明利用 CSL 能够解决角度未对齐的问题, 在提高包围盒精度的情况下, 降低了错检和漏检的概率。

表 4 消融实验结果

Table 4 Results of ablation experiment

Network	Module 1	Module 2	Module 3	Module 4
TEN	✓	✓	✓	✓
CEN		✓	✓	✓
ATFPN			✓	✓
CSL				✓
mAP / %	87.45	87.92	89.04	89.91

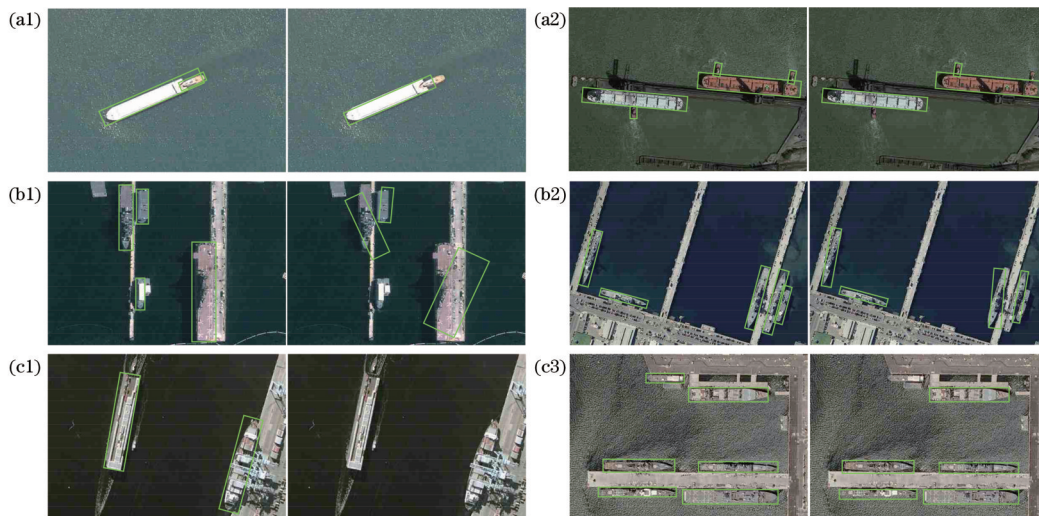


图 11 所提方法与 R³Det 基准实验的结果对比。(a1)(a2) 纹理增强; (b1)(b2) 颜色增强; (c1)(c2) 角度对齐

Fig. 11 Result comparison between proposed method and R³Det benchmark experiment. (a1)(a2) Texture enhancement; (b1)(b2) color enhancement; (c1)(c2) angle alignment

4 结 论

针对遥感图像中的舰船小目标进行研究, 提出一种基于双重特征增强的遥感舰船小目标的检测方法。

针对遥感图像中存在的颜色信息和纹理信息不足的特点, 在预处理阶段将生成对抗网络和强化学习结合, 增强了遥感图像的视觉效果。针对遥感舰船目标存在的目标尺度差异大的问题, 在多路径的特征金字塔中加

入注意力上采样机制,解决了特征细节信息丢失的问题,提高了舰船小目标的特征表现力。针对遥感船舶中大多数目标都具有方向性不定和长宽比过大的特点,在回归任务中加入级联的特征细化模块,重新设计了基于角度分类的损失函数,提高了角度回归的稳健性。为了验证所提方法的有效性,分别在 DOTA 和 HRSC2016 数据集上进行了对比实验。实验结果表明,所提方法在检测精度上优于其他现存的主流方法,能够为遥感舰船图像的小目标检测提供良好的解决方案。

参 考 文 献

- [1] Zhang D J, Zhan J, Tan L F, et al. Comparison of two deep learning methods for ship target recognition with optical remotely sensed data[J]. *Neural Computing and Applications*, 2021, 33(10): 4639-4649.
- [2] 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.
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.
- [3] 刘获, 张焱, 赵琰, 等. 基于特征重聚焦网络的多尺度近岸舰船检测[J]. *光学学报*, 2021, 41(22): 2215001.
Liu D, Zhang Y, Zhao Y, et al. Multi-scale inshore ship detection based on feature re-focusing network[J]. *Acta Optica Sinica*, 2021, 41(22): 2215001.
- [4] Lin T Y, Maire M, Belongie S, et al. Microsoft COCO: common objects in context[M]//Fleet D, Pajdla T, Schiele B, et al. *Computer vision-ECCV 2014. Lecture notes in computer science*. Cham: Springer, 2014, 8693: 740-755.
- [5] Liu Z K, Yuan L, Weng L B, 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*, February 24-26, 2017, Porto, Portugal. Setúbal: SCITEPRESS-Science and Technology Publications, 2017.
- [6] Xia G S, Bai X, Ding J, et al. DOTA: a large-scale dataset for object detection in aerial images[C]//*IEEE/CVF Conference on Computer Vision and Pattern Recognition*, June 18-23, 2018, Salt Lake City, UT, USA. New York: IEEE Press, 2018: 3974-3983.
- [7] 王勃凡, 赵海涛. 基于径向基激活函数的高光谱小目标检测[J]. *光学学报*, 2021, 41(23): 2311001.
Wang B F, Zhao H T. Small object detection in hyperspectral images based on radial basis activation function[J]. *Acta Optica Sinica*, 2021, 41(23): 2311001.
- [8] 刘峰, 郭猛, 王向军. 基于跨尺度融合的卷积神经网络小目标检测[J]. *激光与光电子学进展*, 2021, 58(6): 0610012.
Liu F, Guo M, Wang X J. Small target detection based on cross-scale fusion convolution neural network[J]. *Laser & Optoelectronics Progress*, 2021, 58(6): 0610012.
- [9] 奚琦, 张正道, 彭力. 基于改进 MDSSD 的小目标实时检测算法[J]. *激光与光电子学进展*, 2020, 57(20): 201009.
Xi Q, Zhang Z D, Peng L. Small target real-time detection algorithm based on improved MDSSD[J]. *Laser & Optoelectronics Progress*, 2020, 57(20): 201009.
- [10] Bai Y C, Zhang Y Q, Ding M L, et al. SOD-MTGAN: small object detection via multi-task generative adversarial network[M]//Ferrari V, Hebert M, Sminchisescu C, et al. *Computer vision-ECCV 2018. Lecture notes in computer science*. Cham: Springer, 2018, 11217: 206-221.
- [11] Rabbi J, Ray N, Schubert M, et al. Small-object detection in remote sensing images with end-to-end edge-enhanced GAN and object detector network[J]. *Remote Sensing*, 2020, 12(9): 1432.
- [12] Wang X T, Yu K, Wu S X, et al. ESRGAN: enhanced super-resolution generative adversarial networks[M]//Leal-Taixé L, Roth S. *Computer vision-ECCV 2018 workshops. Lecture notes in computer science*. Cham: Springer, 2019, 11133: 63-79.
- [13] Bashir S M A, Wang Y. Small object detection in remote sensing images with residual feature aggregation-based super-resolution and object detector network[J]. *Remote Sensing*, 2021, 13(9): 1854.
- [14] Xi Y, Jia W J, Zheng J B, et al. DRL-GAN: dual-stream representation learning GAN for low-resolution image classification in UAV applications[J]. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2020, 14: 1705-1716.
- [15] 霍青兰. 国产高分辨率光学遥感卫星影像地图制图技术研究[D]. 兰州: 兰州交通大学, 2020: 1-10.
Huo Q L. Research on mapping technology of domestic high-resolution optical remote sensing satellite image map [D]. Lanzhou: Lanzhou Jiatong University, 2020: 1-10.
- [16] Park J, Lee J Y, Yoo D, et al. Distort-and-recover: color enhancement using deep reinforcement learning [C]//*IEEE/CVF Conference on Computer Vision and Pattern Recognition*, June 18-23, 2018, Salt Lake City, UT, USA. New York: IEEE Press, 2018: 5928-5936.
- [17] Dharejo F A, Zhou Y C, Deeba F, et al. A remote-sensing image enhancement algorithm based on patch-wise dark channel prior and histogram equalisation with colour correction[J]. *IET Image Processing*, 2021, 15(1): 47-56.
- [18] Xu Y F, Zhang N, Li L, et al. Joint learning of super-resolution and perceptual image enhancement for single image[J]. *IEEE Access*, 2021, 9: 48446-48461.
- [19] Hou Q B, Zhou D Q, Feng J S. Coordinate attention for efficient mobile network design[C]//*IEEE/CVF Conference on Computer Vision and Pattern Recognition*, June 20-25, 2021, Nashville, TN, USA. New York: IEEE Press, 2021: 13708-13717.
- [20] Yang X, Yan J C. Arbitrary-oriented object detection with circular smooth label[M]//Vedaldi A, Bischof H,

- Brox T, et al. Computer vision-ECCV 2020. Lecture notes in computer science. Cham: Springer, 2020, 12353: 677-694.
- [21] Chen Y P, Fan H Q, Xu B, et al. Drop an octave: reducing spatial redundancy in convolutional neural networks with octave convolution[C]//IEEE/CVF International Conference on Computer Vision, October 27-November 2, 2019, Seoul, Korea (South). New York: IEEE Press, 2019: 3434-3443.
- [22] Mnih V, Kavukcuoglu K, Silver D, et al. Playing atari with deep reinforcement learning[EB/OL]. (2013-12-19) [2021-05-04]. <https://arxiv.org/abs/1312.5602>.
- [23] Ying M S, Feng Y, Ying S G. Optimal policies for quantum Markov decision processes[J]. International Journal of Automation and Computing, 2021, 18(3): 410-421.
- [24] Qian W, Yang X, Peng S L, et al. Learning modulated loss for rotated object detection[EB/OL]. (2019-11-19) [2021-05-08]. <https://arxiv.org/abs/1911.08299>.
- [25] Ming Q, Miao L J, Zhou Z Q, et al. Optimization for arbitrary-oriented object detection via representation invariance loss[J]. IEEE Geoscience and Remote Sensing Letters, 2022, 19: 21525304.
- [26] Yang X, Liu Q Q, Yan J C, et al. R3Det: refined single-stage detector with feature refinement for rotating object [EB/OL]. (2019-08-15)[2021-05-07]. <https://arxiv.org/abs/1908.05612>.
- [27] Yang X, Yang J R, Yan J C, et al. SCRDet: towards more robust detection for small, cluttered and rotated objects[C]//IEEE/CVF International Conference on Computer Vision, October 27-November 2, 2019, Seoul, Korea (South). New York: IEEE Press, 2019: 8231-8240.
- [28] Lin T Y, Goyal P, Girshick R, et al. Focal loss for dense object detection[C]//IEEE International Conference on Computer Vision, October 22-29, 2017, Venice, Italy. New York: IEEE Press, 2017: 2999-3007.
- [29] Jiang Y Y, Zhu X Y, Wang X B, et al. R2CNN: rotational region CNN for orientation robust scene text detection[EB/OL]. (2017-06-29) [2021-07-08]. <https://arxiv.org/abs/1706.09579>.
- [30] Han J M, Ding J, Xue N, et al. ReDet: a rotation-equivariant detector for aerial object detection[C]//IEEE/CVF Conference on Computer Vision and Pattern Recognition, June 20-25, 2021, Nashville, TN, USA. New York: IEEE Press, 2021: 2785-2794.