

# 基于双重标签分配的遥感有向目标检测方法

董燕<sup>1,2</sup> 魏铭宏<sup>1</sup> 高广帅<sup>1</sup> 刘洲峰<sup>1</sup> 李春雷<sup>1</sup>

1 中原工学院信息与通信工程学院 郑州 450007

2 电子科技大学自动化工程学院 成都 610000

(dy@zut.edu.cn)

**摘要** 由于遥感图像目标具有任意方向、大纵横比和密集排列等多样性分布特点,预设的锚框难以精准匹配所有真实目标,导致对大纵横比和密集排列的有向目标检测精度不高。为了解决上述问题,提出了一种基于双重标签分配的遥感有向目标检测方法。首先,提出双重标签分配策略为目标分配最大及次优交并比的候选框;其次,通过排斥损失(AP-Loss)和吸引损失(UP-Loss)约束相邻目标的候选框,以提高目标正确匹配概率;然后,为了提取适应于分类和回归分支的鲁棒特征,设计了一个特征增强模块(FEM),该模块基于偏振函数构造自适应特征,能够有效增强分类和回归任务所需的特征表达能力;最后,设计了一个定位指导分类(LGC)模块,该模块通过定位任务指导分类任务的采样位置,进行定位细化,以获取分类任务的关键特征,从而缓解分类与定位之间的不一致问题。在3个公开的遥感有向目标检测数据集DOTA,HRSC-2016和DIOR-R上进行了大量的实验,实验结果证明了所提方法的有效性,且优于现有主流方法。

**关键词:** 遥感图像;有向目标检测;双重标签分配;不一致问题;采样位置细化

**中图分类号** TP751

## Remote Sensing Oriented Object Detection Method Based on Dual-label Assignment

DONG Yan<sup>1,2</sup>, WEI Minghong<sup>1</sup>, GAO Guangshuai<sup>1</sup>, LIU Zhoufeng<sup>1</sup> and LI Chunlei<sup>1</sup>

1 School of Information and Communication Engineering, Zhongyuan University of Technology, Zhengzhou 450007, China

2 School of Automation Engineering, University of Electronic Science and Technology, Chengdu 610000, China

**Abstract** Due to the inherently diverse distribution characteristics of remote sensing image objects, such as their arbitrary orientation, large aspect ratio, and densely arranged, the utilization of preset anchor boxes makes it difficult to accurately match all real objects. This limitation results in low detection accuracy, particularly for oriented objects with large aspect ratios and densely arranged. To address this issue, an oriented object detection method in remote sensing images based on dual label assignment is proposed. Firstly, a dual-label assignment strategy is proposed to assign candidate boxes with maximum and suboptimal intersection union ratios to the real object. Then, the candidate boxes of adjacent objects are constrained by repulsion loss (AP Loss) and attraction loss (UP Loss) to improve the probability of correct object matching. In addition, to extract robust features suitable for classification and regression branches, a Feature Enhancement Module (FEM) is designed. This module constructs adaptive features based on polarization functions, which can effectively enhance the feature expression ability required for classification and regression tasks. Finally, a localization-guided classification (LGC) module is designed, which guides the sampling position of the classification task through localization tasks, performs localization refinement, and obtains key features of the classification task, thereby alleviating the inconsistency between classification and localization. A large number of experiments were conducted on three publicly available oriented object detection in remote sensing datasets, namely DOTA, HRSC-2016, and DIOR-R. The experimental results demonstrated the effectiveness of the proposed method and its accuracy (mAP) is better than existing mainstream methods.

**Keywords** Remote sensing images, Oriented object detection, Dual-label assignment, Inconsistency issues, Sampling position refinement

## 1 引言

有向目标检测通过旋转边界框来定位及识别感兴趣的目标。相对于水平边界框,旋转框更加精准并且保留目标的方向信息,已广泛应用于遥感目标检测、人脸识别、场景文本检

测和自然场景目标检测等领域<sup>[1-9]</sup>。然而,遥感图像中的目标通常具有密集排列、任意方向和大长宽比等多样性分布特性,有向目标检测在遥感领域仍然面临挑战。

传统有向目标检测方法<sup>[10]</sup>通过预设不同宽度、高度和角度的旋转锚框来解决锚框角度回归问题。然而,预设大量旋

基金项目:国家自然科学基金(62072489);中原科技创新领军人才项目(234200510009);河南省科技攻关项目(232102211002,232102211030)

This work was supported by the NSFC(62072489), Leading Talents of Science and Technology in the Central Plain of China(234200510009) and Henan Province Key Science and Technology Research Projects(232102211002,232102211030).

通信作者:李春雷(lichunlei1979@zut.edu.cn)

转锚框导致了正负样本的不平衡和大量冗余计算。为解决上述问题,保留水平锚框的方法<sup>[1,11]</sup>通过水平锚框到旋转框的转换实现回归。Ding 等<sup>[11]</sup>提出 RoI-transformer 来实现从水平锚框到旋转锚框的变换。 $S^2$  A-Net<sup>[1]</sup>提出一个特征对齐模块,仅预设单尺度水平锚框的同时生成高质量旋转锚框。放弃传统锚框的方法<sup>[12-13]</sup>也取得了良好的检测效果。Guo 等<sup>[12]</sup>提出一种凹包表示法来优化预测框回归。Xu 等<sup>[13]</sup>设计了四元组滑动顶点,用于实现精准定位。然而,在遥感场景中,由于目标的多样性分布特性,密集设置的锚框也难以实现高性能的检测。

为了更好地适应遥感图像的复杂场景,模型需要采用更精确的标签分配策略,以提高对多样性分布目标的匹配概率。在匹配大纵横比和角度变化的目标时,粗粒度标签分配策略(Maximum Intersection Over Union, MAXIOU)存在局限性,即无法保证样本分配阶段,所有目标均匹配到足量的正样本,从而影响了模型的检测性能。为了缓解样本不均衡和不充分问题,动态样本分配策略<sup>[14-15]</sup>采用动态的匹配度量进行样本分配任务。Zhang 等<sup>[14]</sup>指出是否基于锚框对模型检测性能的影响,并根据目标的统计特征构建了样本分配策略。Ming 等<sup>[15]</sup>根据捕获关键特征的能力自适应地选择高质量锚框。尽管基于交并比(IoU)的动态度量简单直观,但不能满足遥感图像目标多样性分布的需求。基于高斯分布的样本选择方法<sup>[16-18]</sup>采用高斯-瓦瑟斯坦距离(Gaussian Wasserstein Distance, GWD)来衡量样本与真实值之间的差异。Wang 等<sup>[17]</sup>反映了有向目标的形状和方向特征,有效地评估了样本与真实值之间的差距。Huang 等<sup>[18]</sup>提出了一种基于二维高斯分布表示正样本的通用方法。尽管上述策略有效缓解了样本不均衡问题,但是需要先验设定参数和复杂函数参与并且引入大量低质量样本。

现有的标签分配策略仅确保候选框接近其匹配的目标,却忽略了周围目标的影响。在边界框回归过程中,当预测框移动到其匹配目标周围时,并未对其施加额外的惩罚。如图 1 所示,当目标 A 匹配的候选框错误地匹配目标 B 时,导致检测不准确。此外,由于需要非极大值抑制(NMS)处理初始检测结果,易于产生匹配目标 A 的边界框被匹配目标 B 的边界框抑制,最终导致目标 A 漏检。即模型检测密集目标时,对 NMS 的阈值敏感;较高的阈值带来更多的假阳性样本,而较低的阈值导致更多的漏检现象。因此,在复杂场景中鲁棒地定位目标是遥感图像中有向目标检测的关键问题之一。基于上述问题,本文提出了一种双重标签分配策略。其结合最大交并比(MAXIOU)和次优交并比(Second Intersection Over Union, SECIOU)两种指标去匹配真实目标,以提

高目标被正确匹配的概率。MAXIOU 策略通过选择与候选框具有最大交并比的真实目标来进行标签分配。SECIOU 策略则补偿候选框匹配目标的周围目标,从而缓解漏检问题。

现有的目标检测方法<sup>[19-20]</sup>通常采用两个共享特征的分支分别实现分类任务和定位任务。然而,分类分支更关注旋转不变性特征,而定位分支更关注旋转可变性特征,导致遥感有向目标检测遭受不一致问题困扰。Cheng 等<sup>[21]</sup>提出大多数错误检测由不一致问题导致,即高分类分数对应低定位精度。为了缓解上述的不一致问题,Cheng 等<sup>[22]</sup>提出了一种解耦分类精化(DCR)方法,通过解耦分类和回归来选择样本,并训练另一个分类器来细化分类结果。Feng 等<sup>[23]</sup>提出了一种任务对齐的单阶段目标检测,来增强两个任务之间的交互。RefineDet++<sup>[24]</sup>通过自适应地调整空间特征来缓解空间特征和锚框之间的不匹配。

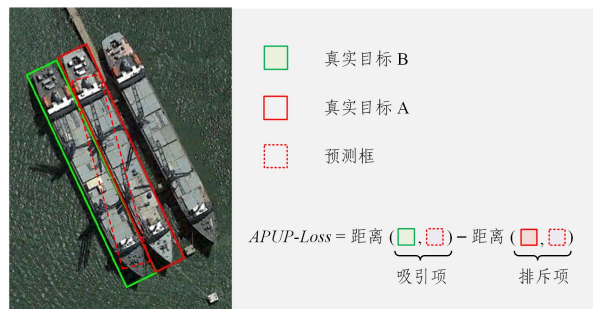


图 1 APUP-Loss 示意图

Fig. 1 Schematic diagram of APUP-Loss

通过对 CNN 提取的特征进行热力图可视化,发现目标检测所需的“关键特征”<sup>[25]</sup>在目标上并非均匀分布,而可能分布在局部区域上(见图 2(a)和图 2(b))。锚框需要捕获目标的关键特征以实现准确检测,即需要充分考虑分类和回归分支之间的特征倾向性。以分类任务为例,传统方法将图 2(c)和图 2(d)中的锚框判定为正样本,即锚框与真实框之间的交并比高于 0.5。但是图 2(c)中的锚框未能捕获到识别船舶所必需的关键特征,导致定位失准和分类错误,从而影响检测性能。只有捕获目标所需的关键特征的锚框(如图 2(d)中的船只),即正确的预测框应该准确地捕获区分不同类别的特征,以实现正确的分类。为了减轻上述问题的影响,本文设计了一个特征增强模块(Feature Enhancement Module, FEM),分别获得分类和回归两个分支的倾向性特征。并且设计一个定位指导分类(Localization-Guided Classification, LGC)模块来获取关键特征的采样位置。综上所述,本文的主要贡献可概括如下:

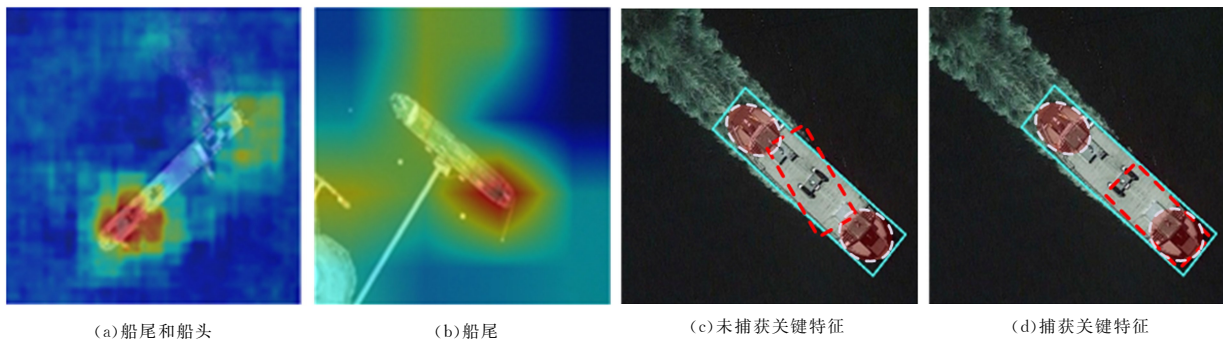


图 2 关键特征示意图

Fig. 2 Schematic diagram of critical feature

1)设计了一个特征增强模块(FEM),该模块旨在提取分类和回归两个分支的倾向性鲁棒特征,以提高模型的性能。

2)设计了一个定位指导分类(LGC)模块,利用定位任务的精确位置信息,指导分类任务中采样位置的局部细化,从而减轻分类和回归之间的不一致性。提出了一个双重标签分配策略,旨在优化多样性分布的遥感有向目标匹配正样本,以提高候选框的利用率,并减轻目标多样性分布带来的影响。本文设计的损失函数(APUP-Loss)通过排斥损失(Almighty-Push Loss, AP Loss)和吸引损失(Universal-Pull Loss, UP Loss)来约束候选框。

本文在DOTA<sup>[2]</sup>,HRSC-2016<sup>[26]</sup>和DIOR-R<sup>[27]</sup>数据集上分别取得了74.74%,89.90%和64.60%的mAP,均达到先进的准确性。本文第2章详细介绍了本文方法;第3章在

DOTA,HRSC-2016和DIOR-R3个公开的遥感数据集上进行了消融实验和对比实验;最后总结全文。

## 2 本文方法

本文方法的模型框架如图3所示。该模型由骨干网络、特征金字塔网络(FPN)、数据增强模块(FEM)、初始检测头和细化检测头组成。初始检测头生成高质量的旋转锚点,并为细化阶段提供对齐特征。细化阶段提取旋转不变的特征,通过分类和定位分支来得到最终结果。提出的双重标签分配策略只作用于初始阶段,并且与排斥损失(AP-Loss)和吸引损失(UP-Loss)共同作用,以约束目标匹配的候选框来获得高质量样本。LGC模块利用初始检测头的检测结果对分类分支的采样位置进行定位细化,通过分支来准确地提取关键特征。

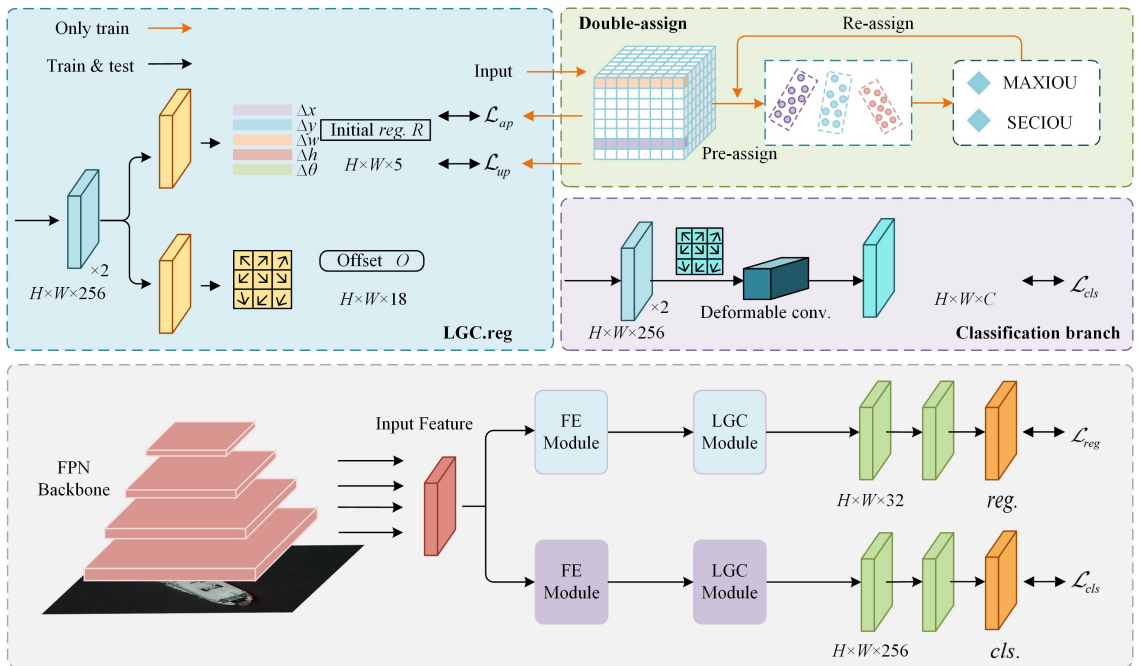


图3 模型框架图

Fig. 3 Model framework diagram

### 2.1 特征增强模块

传统的目标检测框架<sup>[19-20]</sup>中,分类分支和回归分支通常共享相同特征。然而,在有向目标检测中,回归分支需要对角度的变化敏感,以实现准确的角度预测,而分类分支应该对不同角度具有相同的响应。因此,本文设计了一个特征增强模块(FEM),生成分类和回归两分支的偏向性特征,以避免不同任务之间的特征干扰。模块总体结构如图4所示,该模块由注意力模块和特定任务的偏振函数组成。其包含通道方向的注意力(Channel Attention,CA)和空间方向的注意力(Spatial Attention,SA)机制,以此增强模型的特征表达能力。CA使用全局平均池化和自适应1D卷积运算来捕获通道间的依赖性,其表达式如下:

$$M_c = \sigma(C1D_k(F_{\text{gap}})) \times F \quad (1)$$

其中, $M_c$ 表示通道注意力图, $F$ 是FPN层输出的特征图, $F_{\text{gap}}$ 通过全局最大池化之后的特征图。C1D表示具有自适应内核大小 $k$ 的1D卷积。SA使用轻量级空间注意力模块沿着两个空间方向聚合特征,从而捕获输入图像中的像素之间的依赖关系。其中一个捕获通道之间的远程依赖性,而另一个

保留空间上精确的位置信息。其定义如下:

$$A_s(\omega, c) = \frac{1}{H} \sum_{i=0}^{H-1} F(i, \omega, c) \quad (2)$$

$$A_s(h, c) = \frac{1}{H} \sum_{j=0}^{W-1} F(j, h, c) \quad (3)$$

其中, $A_s(\omega, c)$ 和 $A_s(h, c)$ 表示FPN层的两个输出,而 $F(j, h, c)$ 和 $F(i, \omega, c)$ 对输入特征图 $F$ 的通道进行编码。然后,对所得输出进行级联操作,并将其输入到卷积和批量归一化(BN)层,以获得注意力图 $f$ 。其定义如下:

$$f = \delta(F_1[A_s(\omega, c)^h, A_s(h, c)^w]) \quad (4)$$

其中, $[\cdot, \cdot]$ 表示连接操作, $F_1$ 表示共享的 $1 \times 1$ 卷积后的特征图, $F_1$ 被送到非线性激活函数以获得注意力图 $f$ 。然后将该映射沿着空间维度分成两个单独的张量( $f^h \in \mathbb{R}^{C \times H \times W}$ 和 $f^w \in \mathbb{R}^{C \times H \times W}$ ),并输入到卷积层和非线性激活函数。最后,将获得的注意力权重乘以原始输入特征图以获得增强的注意力图。其定义如下:

$$M = M_c(F) \otimes M_s(F) \quad (5)$$

$$F' = M + \varphi(\sigma(M)) \odot F + F \quad (6)$$

其中,  $\otimes$  和  $\odot$  分别表示矢量和元素乘积,  $\sigma$  表示归一化函数,  $\varphi$  是两分支特定的激活函数。特征图  $F \in \mathbb{R}^{C \times H \times W}$  被输入到通道和空间注意力模块, 以产生通道注意力图  $M_c$  和空间注意力图  $M_s$ 。两个注意力特征图在水平方向上进行逐元素相乘操作, 以生成特征增强的注意力特征图  $M$ 。特征图  $M$  随后进行归一化处理, 并输入到任务特定的激活函数中, 以分别获得两分支的倾向性鲁棒特征。在分类任务中, 期望捕获更多的高响应区域特征, 并忽略一些用于定位的边缘特征。因此, 本文使用类 S 曲线函数来辅助过滤边缘信息, 并仅保留用于分类的高级语义特征。分类分支中使用的激发函数的表达式如下:

$$\varphi_{\text{cls}}(x) = \frac{1}{1 + e^{-\mu(x-0.5)}} \quad (7)$$

其中, 参数  $\mu$  用来调节特征被激活强度。由于 FPN 获得的特征图在训练初期具有不稳定性, 本文采用自适应参数阈值, 即参数  $\mu$  随着训练轮数增加而增加, 则用于分类任务的激发函数可以在训练初期保留额外的特征。随着训练稳定, 参数  $\mu$  的逐渐增加有助于过滤不重要的特征, 并生成适合特定分类

任务的鲁棒特征。分类任务激发函数如图 4 所示, 当参数阈值大于 0.5 时, 所有归一化特征都将被激活。因此, 分类分支可以更多关注对分类有利的区域, 减少过拟合并避免误判。参数  $\mu$  的表达式如下:

$$\begin{cases} process = \frac{s}{S} * l_s \\ \mu = -l * process - l_b \end{cases} \quad (8)$$

其中, 参数  $s$  是当前训练步骤,  $S$  是训练步骤的总数,  $l_s$  是控制曲线更新次数的超参数,  $l_b$  是控制曲线增长率的超参数。在回归分支, 本文期望更多关注目标的边缘信息。为此设计了一个回归任务抑制函数, 表达式如下:

$$\varphi_{\text{reg}}(x) = \begin{cases} x, & \text{if } x < 0.5 \\ 1-x, & \text{otherwise} \end{cases} \quad (9)$$

如图 2 所示, 与分类任务相比, 回归分支期望捕获更多的边缘信息, 并抑制不利于定位的高响应区域信息。在式(9)中, 回归函数抑制具有大于 0.5 的阈值的特征, 并且激活具有小于 0.5 的阈值的特征。最后, 将各分支的鲁棒特征送到多任务分支的分类和回归分支中, 以进一步提高模型的检测性能。

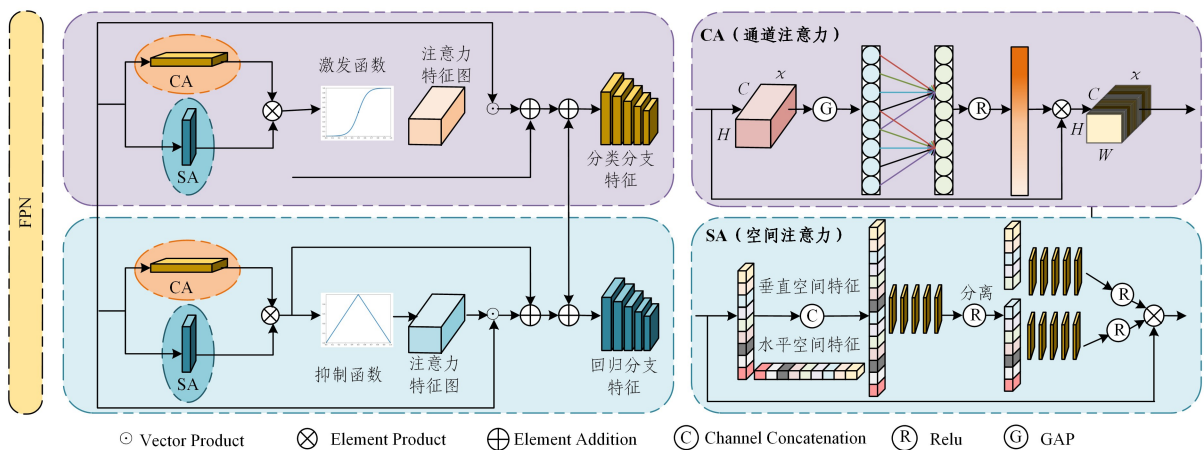


图 4 特征增强模块示意图

Fig. 4 Schematic diagram of feature enhancement module

## 2.2 LGC 模块

Cheng 等<sup>[28]</sup>通过回归分支的预测结果来指导分类分支, 以获得细化的判别性特征图。受此启发, 本文引入可变形卷积(DCN), 通过定位特征来细化分类分支提取关键特征的采样位置, 以缓解不一致问题。由于目标纵横比大和方向任意的特性, 现有的检测器能够准确定位目标, 但难以提取分类特征。只有捕获识别物体所需的关键特征的锚点(例如如图 2(d)中的船只), 才可以实现正确的分类。本文提出的定位指导分类(LGC)模块可以根据准确定位学习分类特征的采样位置, 并将采样位置馈送到可变形卷积以提取区分性特征。本文的参数和 LGC 模块的网络结构如图 3 所示。LGC 模块构造一组可学习的细化采样位置以提取关键特征, 如图 5(a)所示。LGC 模块构造一组可学习的细化采样位置以提取关键特征, 采样过程的表达式如下:

$$P = \{p_i\}_{i=1}^n, n=9 \quad (10)$$

其中,  $n$  是特征采样位置的总数,  $n$  默认设置为 9。通过初始采样位置和可学习的偏移量可以得到细化的特征采样位置。

$$P'_i = (x_i + w * \Delta x_i, y_i + h * \Delta y_i) \quad (11)$$

其中,  $(x_i, y_i)$  代表初始采样点的坐标,  $(\Delta x_i, \Delta y_i)$  代表可学习

的偏移量。其可以通过定位特征学习得到, 表达式如下:

$$\begin{cases} O = \delta(\text{conv}_o, \delta(\text{conv}_r(F_{\text{reg}}))) \\ R = \delta(\text{conv}_o, \delta(\text{conv}_r(F_{\text{reg}}))) \end{cases} \quad (12)$$

其中,  $F_{\text{reg}}$  表示位置特征,  $\delta$  表示激活函数,  $\text{conv}_o$  和  $\text{conv}_r$  是两个连续  $\text{conv}$  的两个分支, 其分别用于获得偏移量和定位特征。  $R \in \mathbb{R}^{H \times W \times 5}$  表示预测的边界框, 偏移量  $O \in \mathbb{R}^{H \times W \times 18}$  采用分类损失优化。偏移量  $O$  的通道数表示 9 个采样位置的横纵坐标偏移量。图 5(a)给出了细化的采样位置生成方法, 对于给定的旋转边界框, 取其中心点、4 个顶点和 4 个顶点的中点一共 9 个点作为初始采样位置。  $w$  和  $h$  分别是边界框的宽度和高度, 从而缩小偏移量的范围, 加快网络的收敛速度。图 5(b)给出了关键特征采样过程, 本文将学习到的偏移量与可变形卷积的偏移量相加来细化分类分支的采样点。最终得到的偏移量  $O$  参与可变形卷积计算, 其可以表示为:

$$F'_{\text{cls}} = \text{conv}(F_{\text{cls}} + \text{DCN}(F, O)) \quad (13)$$

其中, 特征图  $F$  是特征增强模块的输出,  $F_{\text{cls}}$  表示初始分类特征,  $F'_c \in \mathbb{R}^{H \times W \times C}$  是最终用于分类的细化特征。分类分支最终输出特征  $F_c$  是通过  $F_c$  与可变形卷积的输出之间的特征融合以及随后的分类卷积生成的。  $O$  是可变形卷积的偏移量输

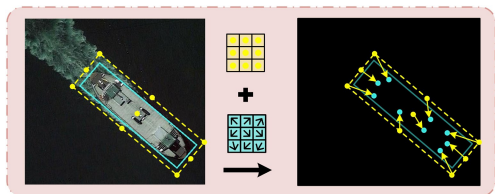
入。对于每个位置  $P_0$ , 基于可变形卷积的锚框采样位置可以被定义为:

$$Y(p_j) = \sum_{r \in \mathcal{R}; o \in \mathcal{O}} W(r) \cdot X(p_j + r + o) \quad (14)$$

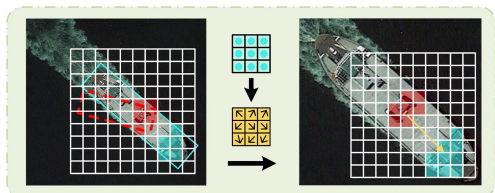
其中,  $\mathcal{R} = \{(r_x, r_y)\}$  是标准卷积的规则网格的集合。偏移量  $S$  和偏移量  $O$  的区别是:  $S$  是初始采样位置与细化采样位置的偏移。二者相加生成最终的偏移量  $O$ , 其被送到可变形卷积以实现关键分类特征提取。  $O$  是可变形卷积, 用于改变卷积的采样位置的偏移, 其可以通过式(5)获得:

$$\mathcal{O} = \{M \cdot p'_i - p_j - r\}_{r \in \mathcal{R}} \quad (15)$$

$M(\theta) = (\cos\theta, -\sin\theta, \sin\theta, \cos\theta)^T$  是空间变换矩阵, 其将细化特征采样位置  $P'_i$  的坐标转换为可变形卷积的采样位置。



(a) 可学习的采样点



(b) 采样点细化

图5 关键特征采样示意图

Fig. 5 Schematic diagram of critical feature sampling

## 2.3 分配策略和损失函数

本文采用长边定义(dle)通过5个参数任意方向的矩形, 分别代表矩形中心点坐标、宽度、长度以及方向角度。其中, 角度的范围取值为  $\theta \in [-\pi/4, 3\pi/4]$ 。所提方法的多任务损失的定义如下:

$$L = L_{\text{reg}}^i + L_{\text{cls}}^i + L_{\text{reg}}^r + L_{\text{cls}}^r \quad (16)$$

其中,  $L^i$  表示初始阶段的损失,  $L^r$  表示细化阶段的损失。分类损失  $L_{\text{cls}}^*$  均采用 Focal 损失<sup>[20]</sup>, 在本节中, 本文提出 APUP-loss 回归损失来解决旋转的候选框在回归过程中的漏检问题。回归损失由两个部分组成, 其定义为:

$$L_{\text{reg}} = \lambda_1 L_{\text{ap}} + \lambda_2 L_{\text{up}} \quad (17)$$

其中, 排斥项  $L_{\text{ap}}$  只作用在初始阶段, 其约束候选框远离其他周围真实目标, 吸引项  $L_{\text{up}}$  约束候选框接近其指定目标, 折衷系数  $\lambda_1$  和  $\lambda_2$  默认设置为 1,  $\mathcal{P}^+ = \{P\}$  是满足分配策略的候选框,  $\mathcal{G}^+ = \{G\}$  是真实框的集合。

### 2.3.1 吸引损失

假设所有真实目标来自同一类别, 采用长边定义法(dle)编码的两个集和  $P = (P_x, P_y, P_w, P_h, P_\theta)$  和  $G = (G_x, G_y, G_w, G_h, G_\theta)$  分别表示候选边界框和真实边界框。本文采用 SmoothL1 距离作为吸引项, 并将 SmoothL1 距离中的平滑参数设置为 2。假设给定其中一个候选框  $P \in \mathcal{P}^+$ , MAXIOU 分配策略将和候选框具有最大交并比的真实框命名为指定目标:  $G_{\text{ap}} = \arg\max_{G \in \mathcal{G}^+} IOU(G, P)$ 。吸引损失表示为:

$$L_{\text{up}} = \frac{1}{N_r} \sum_{n=1}^N t_n' \sum_{j \in \{x, y, w, h, \theta\}} L_{\text{reg}}(v_{nj}', v_{nj}) \quad (18)$$

其中,  $N_i$  表示初始阶段的正样本数量,  $t_n'$  是二进制数值(前景  $t_n' = 1$ , 背景  $t_n' = 0$ )。  $v_{*j}$  表示真实目标向量, 由  $G$  编码得到;  $v_{*j}'$  表示网络输出的预测偏移向量, 由匹配真实目标的候选框  $P$  回归编码得到。

### 2.3.2 排斥损失

给定一个候选框  $P \in \mathcal{P}^+$ , SECIUO 分配策略定义为其指定目标之外具有最大交并比的真实目标。其表达式如下:

$$G_{\text{ap}} = \arg\max_{G \in \mathcal{G}^+ \setminus \{G_{\text{ap}}\}} IOU(G, P) \quad (19)$$

区别于吸引损失, 本文设计排斥损失<sup>[29]</sup> ( $L_{\text{ap}}$ ) 将分配策略匹配的候选框  $P$  和真实目标  $G_{\text{ap}}$  之间的重叠施加排斥约束。  $L_{\text{ap}}$  的重叠(IoG)的定义如下:

$$IoG(B, G) \triangleq \frac{\text{area}(B \cap G)}{\text{area}(G)}$$

排斥损失定义如下:

$$L_{\text{ap}} = \frac{1}{N_r} \sum_{n=1}^N t_n' \text{Smooth}_{\ln}(IoG(B_n, G_n)) \quad (20)$$

$$\text{Smooth}_{\ln} = \begin{cases} -\ln(1-x), & x \leq \sigma \\ \frac{x-\sigma}{1-\sigma} - \ln(1-x), & x > \sigma \end{cases} \quad (21)$$

其中,  $N_r$  表示初始阶段的次级正样本数量, 是在  $(0, 1)$  中连续可微的光滑  $\ln$  函数,  $\sigma \in [0, 1)$  是调节排斥损失对离群点敏感性的光滑参数, 默认设定为 0.5。图6给出了具有不同  $\sigma$  的曲线。由式(21)所示, 候选框与非真实目标重叠越多, 排斥损失在回归过程中对边界框施加更大的惩罚。即排斥损失可以有效地阻止预测的边界框移位到不是其匹配目标的相邻目标。

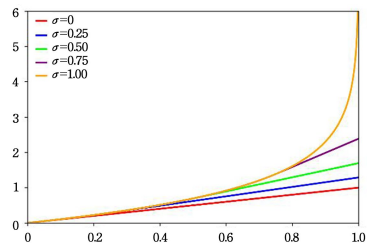


图6 ln函数示意图

Fig. 6 Schematic diagram of ln function

## 3 实验和结果分析

本文在3个公开的遥感图像数据上进行实验, 即 DOTA, HRSC-2016 和 DIOR-R 数据集。实验均在 NVIDIA Tesla V100 服务器上实现, 除非特殊声明, 实验默认由 ResNet101 骨干网络和 FPN 作为基本网络。网络模型均由 SGD 优化器使用 4 个 GPU 进行训练, 每个小批量总共 8 张图像。学习率、动量和权重衰减的超参数分别设置为 0.25, 0.9 和 0.0001。该模型分别在 DOTA, HRSC-2016 和 DIOR-R 数据集上分别训练了 36.48 和 36 个 epoch。在本小节将介绍数据集、消融实验和对比实验的结果分析。

### 3.1 数据集

DOTA 数据集包含 2806 幅卫星光学图像, 大小从  $800 \times 800$  到  $4000 \times 4000$  像素不等, 总共包含 188282 个实例。其被注释的 15 个常见目标类别具有各种各样的尺度、方向和形状。目标类别包括飞机(PL)、棒球场(BD)、桥梁(BR)、地面田径场(GTF)、小型车辆(SV)、大型车辆(LV)、船(SH)、网球

场(TC)、篮球场(BC)、储罐(ST)、足球场(SBF)、环形交叉路口(RA)、港口(HA)、游泳池(SP)和直升机(HC)。DOTA数据集使用原始图像的训练集和验证集进行训练。由于DOTA数据集中的图像尺寸太大,本文将图片按照大小为 $1024 \times 1024$ 像素、步长为200像素进行分割。

HRSC-2016数据集是一个高分辨率的船舶遥感数据集,由6个著名港口收集而来,其中包含1061幅图像,均使用旋转框标注。数据集被分为训练集、验证集和测试集,分别包含436,181和444张图像。本文使用PASCAL VOC指标和COCO测量结果来评估模型在该数据集上的表现。

DIOR-R数据集是一个公开的遥感有向目标检测数据集,它是DIOR数据集的扩展版本,并使用旋转框重新标注。该数据集包含了23463个图像和192518个目标实例。在数据集中选择并标注了20个常见类别,包括飞机(APL)、机场(APO)、棒球场(BF)、篮球场(BC)、桥(BR)、烟囱(CH)、大坝(DAM)、高速公路服务区(ESA)、高速公路收费站(ETS)、高尔夫球场(GF)、田径场(GTF)、港口(HA)、立交桥(OP)、船舶(SH)、体育场(STA)、储罐(STO)、网球场(TC)、火车站(TS)、车辆(VE)和风车(WM)。图像大小为 $800 \times 800$ 像素,空间分辨率为0.5~30m。训练集和验证集包含11725幅图像和68073个实例,测试集包含11738幅图像和124445个实例。

### 3.2 消融实验

为了验证所提方法的有效性,本文在DOTA, HRSC-2016和DIOR-R数据集上进行了一系列的对比实验。实验结果如表1、表2所列,其中包括了每个数据集上各类别的平均精度(AP)、整体平均精度(mAP)以及HRSC-2016数据集

上的AP50等性能指标。实验结果表明,本文方法使得每个数据集上的目标检测性能都得到了显著的提升。

1)FEM模块的影响:表1和表2分别列出了FEM模块在DOTA和HRSC-2016数据集上的结果。FEM模块实施后,DOTA数据集上的mAP性能提高了2.1%,HRSC-2016数据集上的mAP和AP50提高了7.64%和1.8%。结果证明分类和定位分支都得到了自适应的鲁棒特征,提高了检测器的检测性能。

2)LGC策略的效果:根据表1和表2的消融研究结果显示,LGC策略在DOTA和HRSC-2016数据集上取得了明显的提升。LGC策略实施后,DOTA数据集上的mAP提高了1.81%,而在HRSC-2016数据集上的mAP和AP50分别提高了9.71%和2.1%。这些结果表明,通过精确的定位来细化分类采样,从而提取更准确的分类特征,进而提高检测器的检测性能。

3)APUP-Loss的影响:根据表1可知,APUP-Loss实施后,在DOTA数据集上实现了74.46%的mAP,相对提高了2.28%的精度。如表2所列,在HRSC-2016数据集上实现了89.60%的AP50和52.33%的mAP,分别增加了1.17%和3.20%。表1中某些特定类别(如桥、小型车辆、大型车辆和港口)分别达到了51.66%,78.48%,78.16%和73.77%的mAP。上述类别的检测结果表明了双重分配标签策略对遥感中分布多样性目标的有效性。同时,图7给出了可视化结果。图中密集和长宽比大的目标检测性能得到明显提高,漏检问题也得到有效缓解。本文的猜测是,提出的分配策略在没有引入额外的候选框的前提下,提高了真实目标匹配到正样本的概率。

表1 在DOTA数据集上的消融实验

Table 1 Ablation experiments on DOTA dataset

Dataset	FEM	LGC	APUP-Loss	PL	BD	BR	GTF	SV	LV	SH	TC	BC	ST	SBF	RA	HA	SP	HC	mAP
DOTA	—	—	—	88.90	76.38	46.55	70.11	77.62	75.09	86.06	90.89	78.10	81.86	59.00	64.46	66.23	68.76	52.64	72.18
	✓	—	—	<b>89.53</b>	<b>81.00</b>	49.11	70.63	78.90	78.01	<b>87.36</b>	90.89	82.20	83.19	60.68	64.54	71.35	71.06	<b>55.73</b>	74.28
	—	✓	—	89.10	78.48	50.71	69.54	78.85	78.13	87.04	90.88	81.51	83.45	59.60	<b>67.49</b>	71.70	71.14	52.08	73.99
	—	—	✓	89.40	79.11	51.66	70.58	78.48	<b>78.16</b>	87.20	90.86	80.44	<b>83.94</b>	<b>61.61</b>	65.57	<b>73.77</b>	<b>71.23</b>	54.86	74.46
	✓	✓	✓	89.22	79.19	<b>52.17</b>	<b>72.44</b>	<b>79.08</b>	78.10	86.79	<b>90.90</b>	<b>84.53</b>	83.73	61.44	66.70	72.50	70.39	53.84	<b>74.74</b>

表2 HRSC数据集上的消融实验

Table 2 Ablation experiments on HRSC dataset

Dataset	FEM	LGC	APUP-Loss	mAP	AP50
HRSC	—	—	—	45.06	87.80
	✓	—	—	52.70	89.60
	—	✓	—	54.77	89.90
	—	—	✓	52.53	89.60
	✓	✓	✓	<b>56.03</b>	<b>89.90</b>

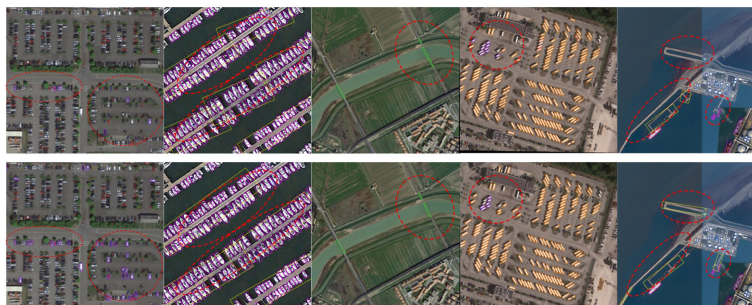


图7 多样性分布目标的检测结果对比

Fig. 7 Comparison of detection results for diversity distribution objects

### 3.3 对比实验

1)在DOTA数据集上的结果如下:本文在DOTA-v1.0 Task1上将所提方法与先进的方法进行了对比实验。为了公平起见,文中给出的结果都来自于官方网站的测试结果。如表3所列, $S^2A\text{-Net}^{[1]}$ 和 $R^3\text{Det}^{[5]}$ 都是先进的遥感有向目标检测器,同时具有最先进的性能。所提方法在二者

精度的基础上提高0.62%和0.95%的mAP。对于在mmrorate上的 $S^2A\text{-Net}^{[1]}$ 方法,本文只取得了72.18%的mAP,相比最终的结果74.74%相差了2.56%,证明了本文方法的有效性。本文在每个特征点处仅预设一个锚框(无冗余尺度和纵横比)来实现上述结果,最终的可视化结果如图8所示。

表3 与DOTA上先进的方法的比较

Table 3 Comparison with advanced methods on DOTA

Methods	backbone	PL	BD	BR	GTF	SV	LV	SH	TC	BC	ST	SBF	RA	HA	SP	HC	mAP
Two-stage																	
CADNet <sup>[30]</sup>	R-101	87.80	82.40	49.40	73.50	71.10	63.50	76.60	90.90	79.20	73.30	48.40	60.90	62.00	67.00	62.20	69.90
SCRDet <sup>[31]</sup>	R-101	<b>89.98</b>	80.65	52.09	68.36	68.36	60.32	72.41	90.85	<b>87.94</b>	<b>86.86</b>	<b>65.02</b>	66.68	66.25	68.24	65.21	72.61
RoI transformer <sup>[11]</sup>	R-101	88.64	78.52	43.44	<b>75.92</b>	68.81	73.68	83.59	90.74	77.27	81.46	58.39	53.54	62.83	58.93	47.67	69.56
One-stage																	
RetinaNet-O <sup>[20]</sup>	R-50	88.67	77.62	41.81	58.17	74.58	71.64	79.11	90.29	82.18	74.32	54.75	60.60	62.57	69.67	60.64	68.43
DRN <sup>[32]</sup>	H-104	88.91	80.22	43.52	63.35	73.48	70.69	84.94	90.14	83.85	84.11	50.12	58.41	67.62	68.60	52.50	70.70
$R^3\text{Det}^{[5]}$	R-101	88.76	<b>83.09</b>	50.91	67.27	76.23	<b>80.39</b>	86.72	90.78	84.68	83.24	61.98	61.35	66.91	70.63	53.94	73.79
PloU <sup>[33]</sup>	DLA-34	80.90	69.70	24.10	60.20	38.30	64.40	64.80	90.90	77.20	70.40	46.50	37.10	57.10	61.90	64.00	60.50
RSDet <sup>[34]</sup>	R-101	89.80	82.90	48.60	65.20	69.50	70.10	70.20	90.50	85.60	83.40	62.50	63.90	65.60	67.20	<b>68.00</b>	72.20
DAL <sup>[15]</sup>	R-50	88.68	76.55	45.08	66.80	67.00	76.76	79.74	90.84	79.54	78.45	57.71	62.27	69.05	<b>73.14</b>	60.11	71.44
$S^2A\text{-Net}^{[1]}$	R-50	89.11	82.84	48.37	71.11	78.11	78.39	<b>87.25</b>	90.83	84.90	85.64	60.36	62.60	65.26	69.13	57.94	74.12
Ours	R-50	89.11	79.23	50.10	71.00	78.94	76.28	87.04	90.90	81.66	83.28	59.63	66.66	67.43	70.63	52.57	73.63
Ours	R-101	89.22	79.19	<b>52.17</b>	72.44	<b>79.08</b>	78.10	86.79	<b>90.90</b>	84.53	83.73	61.44	<b>66.70</b>	<b>72.50</b>	70.39	53.84	<b>74.74</b>

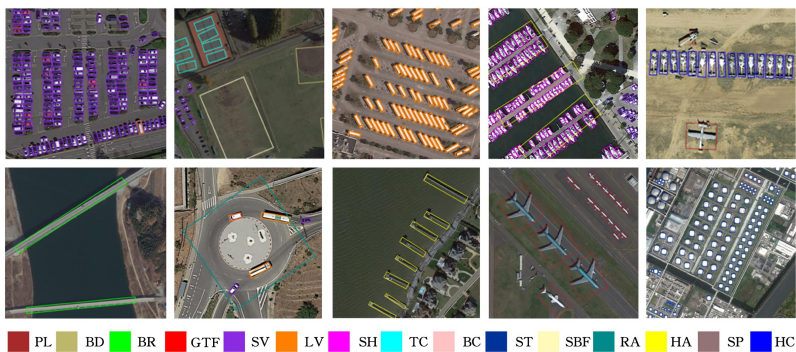


图8 DOTA数据集上的可视化

Fig. 8 Visualization on DOTA dataset

2)在HRSC-2016数据集上的结果如下:HRSC-2016数据集包含许多具有任意方向的细长船舶对象,这对检测提出了相当大的挑战。公平起见,本文在HRSC-2016数据集上的实验使用了PASCAL VOC 2007和COCO两个指标,即AP50和mAP。本文将所提方法与表4中的其他的先进方法进行了比较。实验结果表明,该方法的mAP达到56.03%,在AP50下的识别率为89.90%,优于其他方法的识别率。最终的可视化结果如图9所示。

表4 HRSC数据集上先进模型的对比

Table 4 Comparison of advanced models on HRSC dataset

Methods	Backbone	AP50
RepPoints <sup>[36]</sup>	R-101	85.16
RoI-Transformer <sup>[11]</sup>	R-101	86.20
GLIDING VECTORS <sup>[13]</sup>	R-101	88.20
$R^3\text{Det}^{[5]}$	R-101	89.26
DAL <sup>[15]</sup>	R-101	89.77
PloU <sup>[33]</sup>	DLA-34	89.20
Ours	R-101	<b>89.90</b>



图9 HRSC数据集上的可视化

Fig. 9 Visualization on HRSC dataset

3)在DIOR-R数据集上的结果如下:表5列出了本文方法和其他5种方法在DIOR-R数据集上的比较结果。实验结果表明,本文方法在挑战性较高的目标类别上表现出了优异的性能,例如机场、立交桥和桥。相比其他方法,该方法在这些类别上具有明显的优势,证明了其在复杂环境下进行目标检测任务的有效性和鲁棒性。例如,在机场类别中,Faster R-CNN-O<sup>[35]</sup>,RetinanNet-O<sup>[20]</sup>,Gliding Vertex<sup>[13]</sup>,RoI-transformer<sup>[11]</sup>和AOPG<sup>[27]</sup>的检测准确率分别只有26.80%,28.52%,28.87%,37.88%和37.79%。而本文方法可以获得45.50%的mAP,与其他5种算法相比有明显提高。上述结果证明了该方法的有效性,部分可视化结果如图10所示。

表 5 DIOR 数据集上与优秀模型的对比

Table 5 Comparison with excellent models on DOTA dataset

Methods	Backbone	APL	APO	BF	BC	BR	CH	DAM	ETS	ESA	GF	GTF	HA	OP	SH	STA	STO	TC	TS	VE	WM	mAP
Faster RCNN-O <sup>[35]</sup>	R-50	62.79	26.80	71.72	80.91	34.20	72.57	18.95	66.45	65.75	66.63	79.24	34.95	48.79	81.14	64.34	71.21	81.44	47.31	50.46	65.21	59.54
Retina Net-O <sup>[20]</sup>	R-50	61.49	28.52	73.57	81.17	23.98	72.54	19.94	72.39	58.20	69.25	79.54	32.14	44.87	77.71	67.57	61.09	81.46	47.33	38.01	60.24	57.55
GLIDING VECTORS <sup>[13]</sup>	R-50	65.35	28.87	<b>74.96</b>	81.33	33.88	74.31	19.58	70.72	64.70	72.30	78.68	37.22	49.64	80.22	69.26	61.13	81.49	44.76	47.71	65.04	60.06
ROI-TRANS. <sup>[11]</sup>	R-50	63.34	37.88	71.78	87.53	40.68	72.60	26.86	<b>78.71</b>	68.09	68.96	<b>82.74</b>	<b>47.71</b>	55.61	<b>81.21</b>	<b>78.23</b>	70.26	<b>81.61</b>	54.86	43.27	65.52	63.87
AOPG <sup>[127]</sup>	R-50	62.39	37.79	71.62	<b>87.63</b>	40.90	72.47	<b>31.08</b>	65.42	77.99	73.20	81.94	42.32	54.45	81.17	72.69	<b>71.31</b>	81.49	<b>60.04</b>	<b>52.38</b>	<b>69.99</b>	64.41
Ours	R-50	<b>70.30</b>	42.70	71.50	81.50	41.10	77.40	21.40	69.00	77.40	74.80	80.70	43.30	55.80	80.80	69.40	69.40	81.50	54.70	48.90	64.60	63.80
Ours	R-101	68.80	<b>45.50</b>	71.80	81.40	<b>41.90</b>	<b>74.90</b>	29.60	70.20	<b>79.90</b>	<b>78.10</b>	79.30	43.70	<b>56.00</b>	80.80	67.90	69.30	81.60	57.10	48.70	65.60	<b>64.60</b>

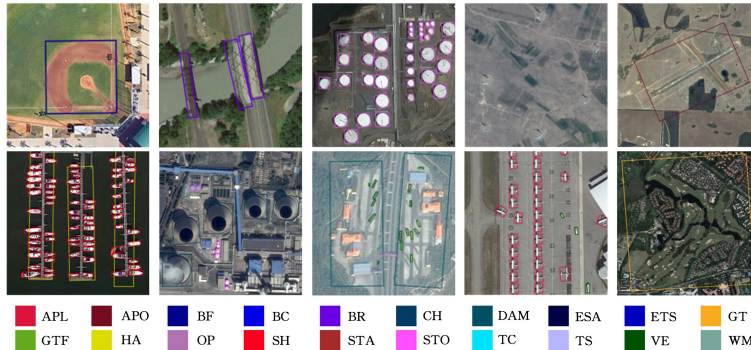


图 10 DIOR 数据集上的可视化

Fig. 10 Visualization on DIOR dataset

**结束语** 本文首先提出了一种双重标签分配策略,为多样性分布的遥感有向目标匹配正样本,设计的损失函数(AP-UP-Loss)通过排斥和吸引匹配的正样本来减缓多样性分布的影响,通过提高正确匹配概率,在 3 个数据集上的检测精度得到了显著的提升;其次,设计了一个特征增强模块(FEM),通过设计的偏振函数来获得分类和回归分支的鲁棒特征,解耦的鲁棒特征有利于准确地分类和回归;最后,设计了一个定位指导分类(LGC)模块。它通过分类任务的采样位置上的局部细化来提取关键特征,以减轻分类任务和回归任务之间的不一致问题。在 3 个公开的有向目标检测数据集上进行了大量实验,实验结果验证了所提方法对遥感图像中多样性分布目标的有效性。本文的标签分配策略忽视了目标的几何特征,后续工作将利用目标的空间属性来优化样本分配。

## 参考文献

- [1] HAN J, DING J, LI J, et al. Align deep features for oriented object detection[J]. IEEE Transactions on Geoscience and Remote Sensing, 2021, 60: 1-11.
- [2] XIA G S, BAI X, DING J, et al. DOTA: A large-scale dataset for object detection in aerial images[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018: 3974-3983.
- [3] RAN Y, ZHANG L. R-YOLOv5: Auto-cutting, R-otated Text Detection Model[J]. Computer Science, 2022, 49(S2): 637-642.
- [4] QIAN W, YANG X, PENG S, et al. RSDet++: Point-based modulated loss for more accurate rotated object detection[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2022, 32(11): 7869-7879.
- [5] YANG X, YAN J, FENG Z, et al. R3det: Refined single-stage detector with feature refinement for rotating object[C]// Proceedings of the AAAI Conference on Artificial Intelligence, 2021, 35(4): 3163-3171.
- [6] LI X, WANG J G. Optimization Algorithm for Detection and Localization of Pure Face in Color Images[J]. Computer Science, 2009, 36(7): 284-287.
- [7] WANG W, CHENG J W, WANG X. Remote Sensing Targets Detection Based on Adaptive Weighting Feature Dictionaries and Joint Sparse[J]. Computer Science, 2018, 45(10): 276-280.
- [8] ZHU Y, FANG G S, ZHENG B B, et al. Research on Detection Method of Refined Rotated Boxes in Remote Sensing[J]. Acta Automatica Sinica, 2023, 49(2): 415-424.
- [9] LU Q, YU Y Q, XU D M, et al. Improved YOLOv5 for Small Drones Target Detection Algorithm [J]. Computer Science, 2023, 50(S2): 212-219.
- [10] MA J, SHAO W, YE H, et al. Arbitrary-oriented scene text detection via rotation proposals[J]. IEEE Transactions on Multimedia, 2018, 20: 3111-3122.
- [11] 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), 2019: 2844-2853.
- [12] GUO Z, ZHANG X, LIU C, et al. Convex-hull feature adaptation for oriented and densely packed object detection[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2022, 32: 5252-5265.
- [13] XU Y, FU M, WANG Q, et al. Gliding vertex on the horizontal bounding box for multi-oriented object detection [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence,

- 2021,43:1452-1459.
- [14] ZHANG S,CHI C,YAO Y,et al. Bridging the gap between anchor-based and anchor-free detection via adaptive training sample selection[C]// 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition(CVPR). 2020:9756-9765.
- [15] MING Q,ZHOU Z,MIAO L,et al. Dynamic anchor learning for arbitrary-oriented object detection [C] // Proceedings of the AAAI Conference on Artificial Intelligence. 2021:2355-2363.
- [16] YANG X,YAN J,MING Q,et al. Rethinking rotated object detection with gaussian wasserstein distance loss[C]// International Conference on Machine Learning. PMLR, 2021: 11830-11841.
- [17] WANG J,XU C,YANG W,et al. A normalized gaussian wasserstein distance for tiny object detection [J]. arXiv, vol. abs/2110.13389,2021.
- [18] HUANG Z,LI W,XIA X G,et al. A general gaussian heatmap label assignment for arbitrary-oriented object detection [J]. IEEE Transactions on Image Processing,2022,31:1895-1910.
- [19] REN S,HE K,GIRSHICK R B,et al. Faster r-cnn:Towards real-time object detection with region proposal networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence,2015,39:1137-1149.
- [20] LIN T Y,GOYAL P,GIRSHICK R B,et al. Focal loss for dense object detection[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence,2020,42:318-327.
- [21] CHENG B,WEI Y,SHI H,et al. Revisiting rcnn:On awakening the classification power of faster rcnn[C]// Proceedings of the European Conference on Computer Vision (ECCV). 2018:453-468.
- [22] CHENG B,WEI Y,FERIS R,et al. Decoupled classification refinement: Hard false positive suppression for object detection [J]. arXiv:1810.04002,2018.
- [23] FENG C,ZHONG Y,GAO Y,et al. Tood:Task-aligned one-stage object detection[C]//2021 IEEE/CVF International Conference on Computer Vision (ICCV). 2021:3490-3499.
- [24] ZHANG S,WEN L,LEI Z,et al. Refinedet++:Single-shot refinement neural network for object detection[J]. IEEE Transactions on Circuits and Systems for Video Technology,2021,31:674-687.
- [25] MING Q,MIAO L,ZHOU Z,et al. CFC-Net: A Critical Feature Capturing Network for Arbitrary-Oriented Object Detection in Remote-Sensing Images[C]// IEEE Transactions on Geoscience and Remote Sensing,2022,60:1-14.
- [26] LIU Z,YUAN L,WENG L,et al. A high resolution optical satellite image dataset for ship recognition and some new baselines [C]//International Conference on Pattern Recognition Applications and Methods. 2017.
- [27] CHENG G,WANG J,LI K,et al. Anchor-free oriented proposal generator for object detection[J]. IEEE Transactions on Geoscience and Remote Sensing,2022,60:1-11.
- [28] CHENG G,YAO Y,LI S,et al. Dual-aligned oriented detector [J]. IEEE Transactions on Geoscience and Remote Sensing, 2022,60:1-11.
- [29] WANG X,XIAO T,JIANG Y,et al. Repulsion Loss: Detecting Pedestrians in a Crowd[C]// 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, UT, USA,2018:7774-7783.
- [30] ZHANG G,LU S,ZHANG W. CAD-Net: A context-aware detection network for objects in remote sensing imagery[J]. IEEE Trans. Geosci. Remote Sens. ,2019,57(12):10015-10024.
- [31] YANG X,YANG J,YAN J,et al. SCRDet: Towards more robust detection for small, cluttered and rotated objects[C]// Proceedings of the IEEE/CVF International Conference on Computer Vision. 2019:8232-8241.
- [32] PAN X,REN Y,SHENG K,et al. Dynamic refinement network for oriented and densely packed object detection [C] // 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition(CVPR). 2020:11204-11213.
- [33] CHEN Z,CHEN K,LIN W,et al. Piou loss: Towards accurate oriented object detection in complex environments[C]// Computer Vision — ECCV 2020: 16th European Conference, Glasgow, UK, August 23-28, 2020, Proceedings, Part V 16. Springer International Publishing, 2020:195-211.
- [34] QIAN W,YANG X,PENG S,et al. Learning modulated loss for rotated object detection[C]// Proceedings of the AAAI Conference on Artificial Intelligence. 2021:2458-2466.
- [35] XIE X,CHENG G,WANG J,et al. Oriented r-cnn for object detection[C]//2021 IEEE/CVF International Conference on Computer Vision (ICCV). 2021:3500-3509.
- [36] LI W,CHEN Y,HU K,et al. Oriented reppoints for aerial object detection[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2022:1829-1838.



**DONG Yan**, born in 1977, Ph.D, professor. Her main research interests include artificial intelligence, pattern recognition and surface defect detection based on machine vision.



**LI Chunlei**, born in 1979, Ph.D, professor. His main research interests include computer vision and pattern recognition.