



研究与开发

## 基于形状自适应标签分配的遥感有向目标检测网络

任王, 吴斌, 余长宏, 曾文捷

(浙江工商大学信息与电子工程学院, 浙江 杭州 310018)

**摘要:** 针对遥感图像中大纵横比目标因正样本不足而出现的 learning 不充分问题, 提出一种基于形状自适应标签分配的遥感有向目标检测网络 (shape-adaptive label assignment for oriented object detection network, SALANet)。首先, 引入纵横比敏感系数建立目标几何特征与正样本数量的动态映射关系, 缓解传统方法中固定分配规则引发的样本分布不平衡问题; 其次, 设计自适应标签分配策略, 通过对交并比 (intersection over union, IoU) 进行排名实现高质量正样本选择; 最后, 提出中心轴先验, 将圆形中心先验区扩展为目标中心轴的矩形区域, 增强大纵横比目标的几何特征表征能力。在 DOTAv1.0 和 HRSC2016 数据集上的对比实验表明, SALANet 分别取得 0.777 1 和 0.932 3 的平均精度均值 (mean average precision, mAP), 较基线方法 RoI Transformer 分别提升 8.15% 和 2.87%。

**关键词:** 遥感图像; 有向目标检测; 大纵横比目标; 形状自适应标签分配; 中心轴先验

**中图分类号:** TN957; TP751

**文献标志码:** A

**doi:** 10.11959/j.issn.1000-0801.2026035

## Shape-adaptive label assignment for oriented object detection network

Ren Wang, Wu Bin, Yu Changhong, Zeng Wenjie

School of Information and Electronic Engineering, Zhejiang Gongshang University, Hangzhou 310018, China

**Abstract:** To address the issue of insufficient learning caused by the lack of positive samples for large aspect ratio targets in remote sensing images, a shape-adaptive label assignment network for oriented object detection (SALANet) was proposed. Firstly, an aspect ratio sensitivity coefficient was introduced to establish a dynamic mapping relationship between target geometric features and the number of positive samples, alleviating the sample distribution imbalance caused by fixed allocation rules in traditional methods. Secondly, an adaptive label assignment strategy was designed to prioritize high-quality positive samples through intersection over union (IoU) ranking. Finally, a central axis prior was proposed, extending the circular central prior region to a rectangular region along the target's central axis, thereby enhancing the geometric feature representation capability for large aspect ratio targets. Comparative experiments on the DOTAv1.0 and HRSC2016 datasets demonstrated that SALANet achieved mAP scores of 0.777 1 and 0.932 3, respectively, representing improvements of 8.15% and 2.87% over the baseline method RoI Transformer.

**Key words:** remote sensing image, oriented object detection, large aspect ratio target, shape-adaptive label assignment, central axis prior

收稿日期: 2025-03-13; 修回日期: 2025-09-30

通信作者: 余长宏, yuch@mail.zjgsu.edu.cn

## 0 引言

近年来，随着对地观测技术的快速发展，航空遥感图像目标检测已成为智慧城市、环境监测等领域的关键技术支持。尽管基于深度学习的目标检测方法（如两阶段的Faster R-CNN<sup>[1]</sup>和单阶段的YOLO<sup>[2]</sup>系列）在自然场景中取得显著进展，但直接应用于遥感场景仍面临三重核心挑战。遥感场景下的挑战如图1所示。

- (1) 多尺度目标共存，尺寸差异可达2个数量级。
- (2) 大纵横比目标（如船舶、桥梁）的纵横比普遍超过5:1。
- (3) 任意方向分布导致传统水平框回归产生冗余背景区域。

因此，遥感图像的目标检测需要结合其多尺度、大纵横比及任意方向分布的数据特性，设计专门的算法和模型，以满足复杂场景下的检测需求<sup>[3]</sup>。

由于遥感影像目标的方向任意性与文本检测具有很大的相似性，许多遥感影像目标检测算法借鉴旋转区域提议网络（rotation region proposal network, RRPN）<sup>[4]</sup>等文本检测算法的思想，通过增加多角度预设描框进行预测，但该方法因增设大量描框会出现性能降低的问题。为此，文献[5]设计旋转感兴趣区域（rotated region of interest, RRoI）学习器模块，利用可变形卷积提取旋转不变特征，将旋转角度加入回归预测；文献[6]提出八参数法表示旋转边界框，通过偏移4个顶点解决角度周期性导致的损失函数梯度不连续问题；

文献[7]进一步将八参数表示法优化成六参数表示法，在保持旋转框精度的同时，将回归参数量降低25%，显著提升了模型收敛速度；文献[8]设计了具有自适应对准功能的即插即用动态旋转卷积来捕获航空目标的旋转角度。

针对大纵横比目标样本数量不足的问题，研究人员从标签分配策略入手展开研究<sup>[9]</sup>。基于锚框的检测方法通常依据锚框与标注框之间的交并比（intersection over union, IoU）进行标签分配，当IoU超过设定阈值的锚框时，被标记为正样本<sup>[10]</sup>，其余则被视为负样本或忽略。无锚框检测器常见的策略是将标注框内的特征点作为正样本<sup>[11-12]</sup>。然而，大纵横比目标对角度变化更加敏感，基于预定义规则的启发式标签分配方法容易导致大纵横比目标样本不足，造成样本分布不平衡<sup>[13]</sup>。不同纵横比目标对角度的敏感性如图2所示。为此，研究者们提出动态标签分配策略以优化训练过程。例如，文献[14]根据目标统计特征动态调整IoU阈值，从而自适应地筛选训练样本；文献[15]提出动态的软标签分配策略，自适应地为样本分配权重，更精准地引导网络的学习过程；文献[16]则引入匹配度替代IoU来衡量锚框的定位精度，并据此分配标签；文献[17]提出双重标签分配策略，为目标分配最大及次优IoU的候选框；文献[18]通过动态样本划分与联合优化，解决孪生跟踪器中分类与回归任务的不一致性问题；文献[19]使用高斯分布成本最优传输分配与解耦层注意力角头两大组件，提升检测的性

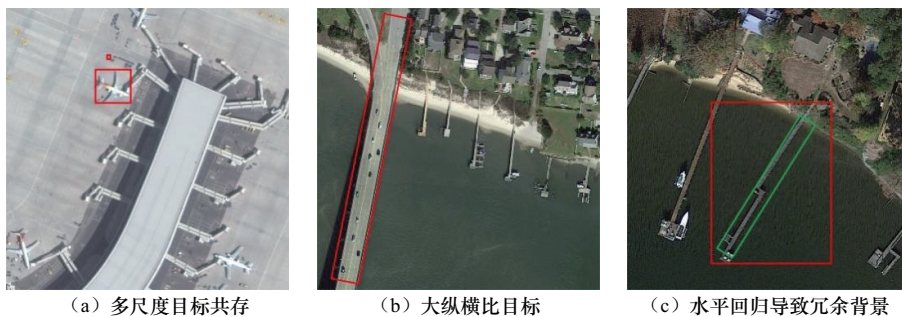


图1 遥感场景下的挑战



能。然而，现有动态标签分配策略普遍忽视目标的几何特征（如形状、方向），导致大纵横比目标的特征表征能力受限<sup>[18]</sup>。本文通过建立纵横比敏感的标签分配模型与中心轴约束机制，突破了传统方法的几何表征瓶颈。

因此，针对目标多尺度、大纵横比和方向任意所带来的检测不佳的问题，基于形状自适应标签分配的遥感有向目标检测网络（shape-adaptive label assignment for oriented object detection network, SALANet）从样本分配不均衡和几何表征不足两大方面进行改进，其主要贡献如下。

(1) 通过纵横比敏感 $k$ （aspect ratio sensitive  $k$ , ARSK）估计模块自适应调整不同纵横比目标所需的样本数量，从而缓解不同纵横比目标间样本数量不平衡的问题。

(2) 设计自适应标签分配策略（adaptive label assignment strategy, ALAS），通过对IoU进行排名实现高质量正样本选择。

(3) 采用中心轴先验（center axis prior, CAP）以适应不同形状和尺寸的目标，通过筛选高质量样本，在增强大纵横比目标的特征表征能力的同时，减轻计算压力。

## 1 基于形状自适应标签分配的遥感有向目标检测算法

### 1.1 模型整体框架

SALANet 的模型框架如图3所示。该模型主要由以下5个核心模块组成：骨干网络（backbone）、特征金字塔网络（feature pyramid network, FPN）、区域提议网络（region proposal network, RPN）、ALAS和检测头（head）。首先，骨干网络负责从输入图像中提取多层次的特征，FPN对这些特征进行融合，以增强多尺度特征的表达能力；RPN基于融合后的特征生成初始候选区域；ARSK模块根据目标的纵横比信息动态计算每个目标所需的样本数量，通过引入纵横比敏感系数，确保不同形状的目标都能获得适量的正样本，缓解正样本分布不均衡的问题；ALAS模块根据样本对目标的重要性动态调整样本分配策略，为每个候选区域分配类别标签，确保每个目标能够获取高质量样本，用于后续训练；通过一致性、可用性和

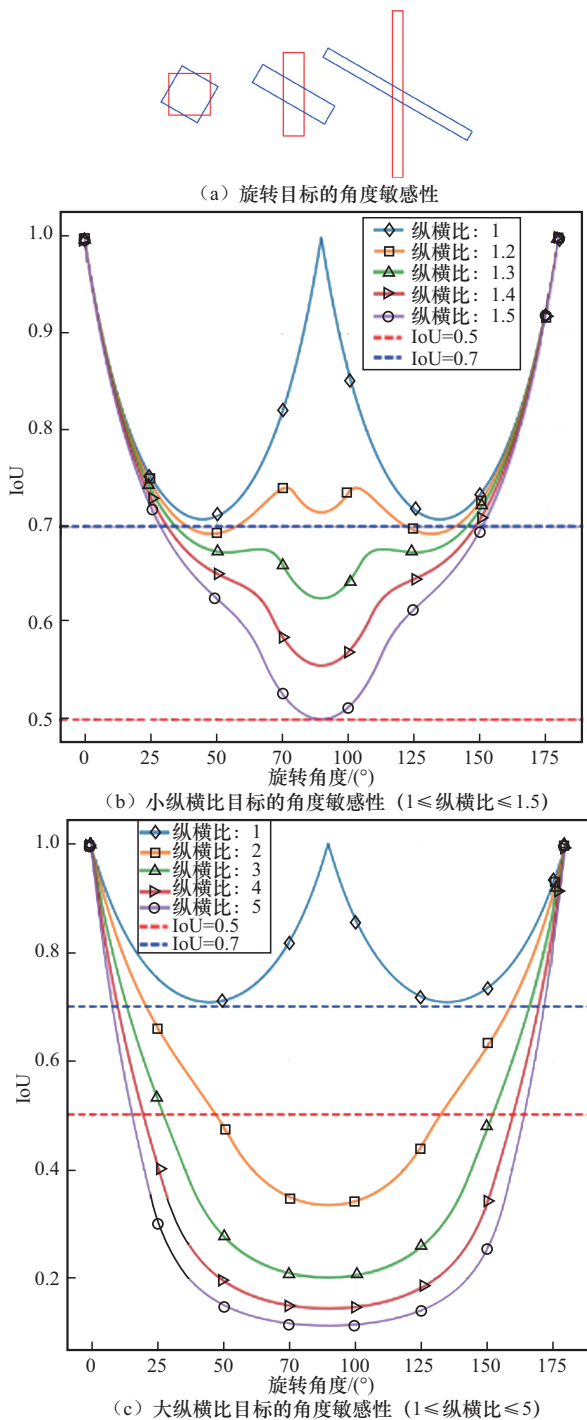


图2 不同纵横比目标对角度的敏感性

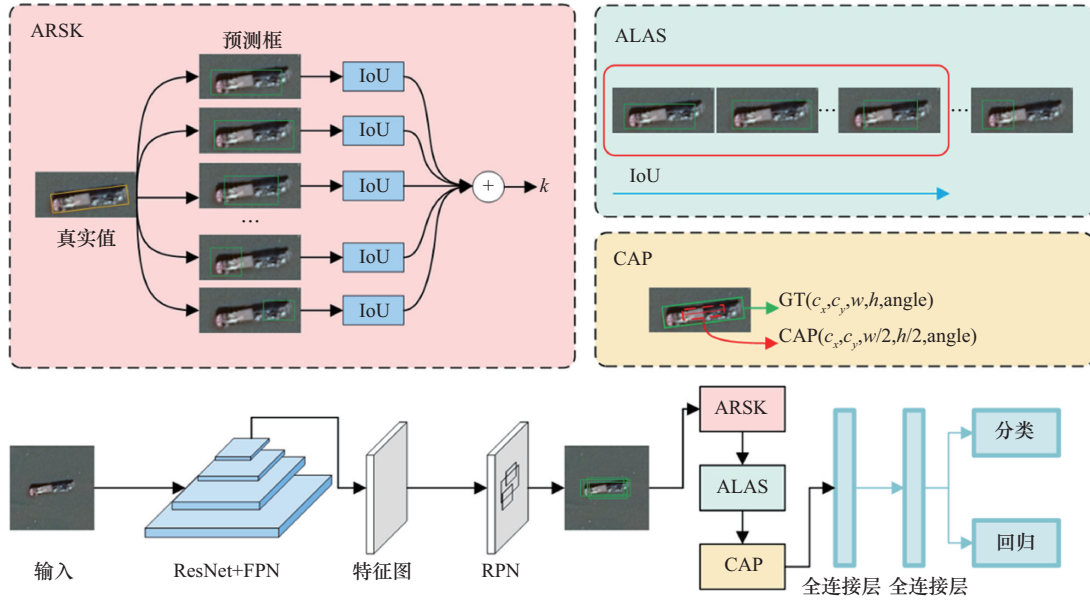


图3 SALANet的模型框架

分区容忍性 (consistency, availability, partition tolerance, CAP) 分析约束预测框和真实框的位置关系, 并将样本中心限制在中心轴区域附近, 筛选更能表征物体方向和形状信息的样本, 从而强化样本对目标几何信息的表征能力, 对于描框的分布描述其中心点位置、宽、高和旋转角度, 记为  $GT(c_x, c_y, w, h, angle)$ ; 最后, 检测头基于筛选后的样本进行分类预测和边界框回归。

### 1.2 纵横比敏感 $k$ 估计

SimOTA<sup>[19]</sup>首次提出动态  $k$  值估计方法, 通过自适应计算每个真实框所需的正样本数量来优化标签分配。SimOTA 先计算所有预测框和真实框之间的 IoU, 为每个真实框累积前  $q$  个最大 IoU 值, 并向下取整得到  $k$  作为目标所需正样本的数量。然而, 由于大纵横比目标的位置较难预测, 且其预测框和真实框之间的 IoU 普遍较小, 大纵横比目标的正样本数量明显少于小纵横比目标, 同时, 大纵横比目标在训练数据集中占比通常较低。这两个因素共同导致正样本数量分布失衡, 使模型对大纵横比目标的学习不够充分。

针对上述问题, 本文通过引入纵横比敏感系

数对动态  $k$  值估计算法进行改进, 以增加大纵横比目标的正样本数量, 如式 (1) 所示, 并通过调整累积 IoU 值的数量  $q^*$ , 为每个真实框累积前  $q^*$  个最大 IoU 值, 并四舍五入得到新的  $k$  值估计。该改进在增加大纵横比目标的正样本数量的同时, 对小纵横比目标的样本分布影响较小, 有效缓解了不同纵横比目标之间正样本数量分布不平衡的问题。ARSK 优化前后过程对比如图 4 所示, 图 4 标注了优化前后的 IoU 累积方式变化, 条纹填充表示优化后大纵横比目标增加的 IoU 累积数量。

$$q^* = \left(1 + \frac{s-1}{\alpha}\right) \times q \quad (1)$$

其中,  $q$  为一个初始参数, 用于确定累积最大 IoU 值的最小数量,  $\alpha$  为纵横比敏感系数, 用于调节  $q^*$  的值, 增加大纵横比目标的正样本数量,  $\alpha$  越小, 表明算法对目标纵横比变化的感知越明显,  $s$  为检测目标的纵横比, 定义为长边与短边的比, 如下所示:

$$s = \frac{\max(G_w, G_h)}{\min(G_w, G_h)} \quad (2)$$

其中,  $G_w$  表示目标旋转框的宽度,  $G_h$  表示目标旋转框的高度。

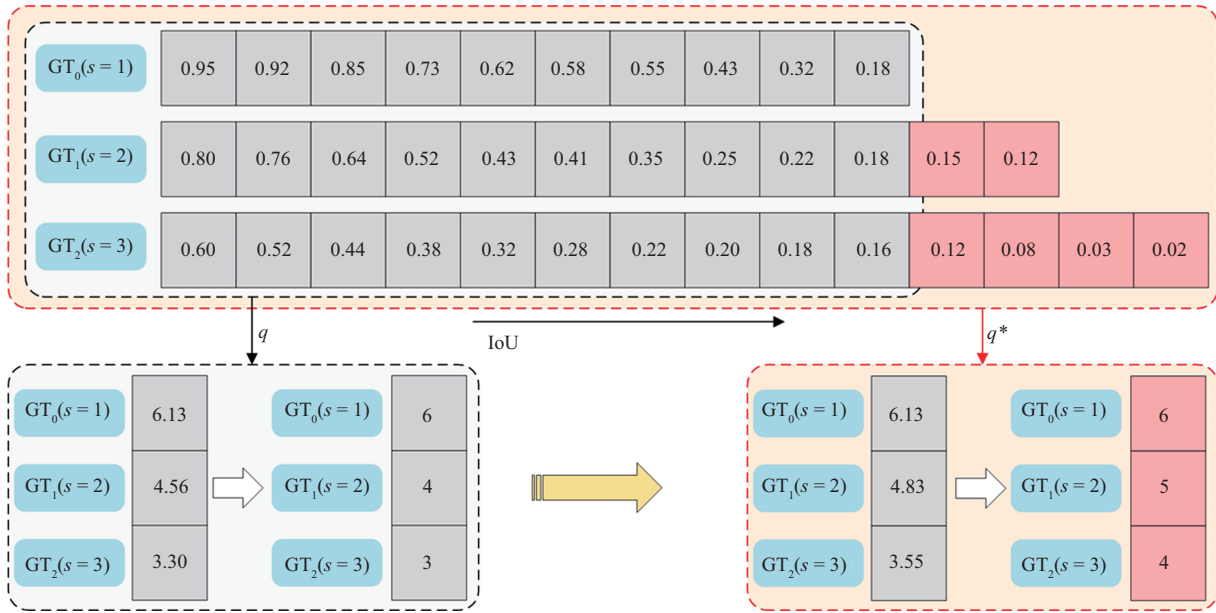


图4 ARSK 优化前后过程对比

### 1.3 自适应标签分配策略

由于不同纵横比物体的预测框与目标框之间的匹配程度存在差异，其在回归任务中的定位能力也有所不同<sup>[20]</sup>。为此，本文提出ALAS，其核心思想是为每个目标选择高质量的候选框作为正样本。假设输入图像有  $N$  个目标真实边界框，记为  $\{G_N\}$ ，其中第  $i$  个目标真实边界框为  $G_i = (c_x^{(i)}, c_y^{(i)}, w^{(i)}, h^{(i)}, \text{angle}^{(i)})$ ，通过RPN生成  $M$  个预测

边界框，记为  $\{P_M\}$ ，其中第  $j$  个预测边界框为  $P_j = (c_x^{(j)}, c_y^{(j)}, w^{(j)}, h^{(j)}, \text{angle}^{(j)})$ ， $K_N$  为 ARSK 模块计算的每个目标所需的样本数量。

ALAS 的标签分配过程如图 5 所示，其中 IoU 矩阵为每个目标和所有锚框的 IoU，排名矩阵为每个目标下样本根据 IoU 的降序排序值，条纹填充表示一个样本被分配多个目标的冲突情况，深色填充表示最后补充的样本，具体步骤详见算法 1。

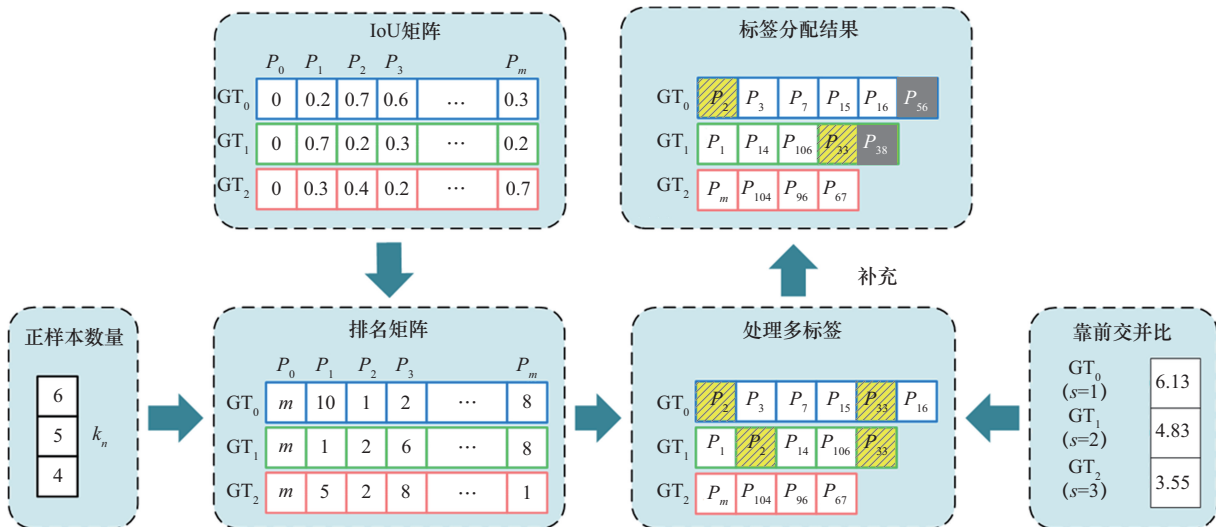


图5 ALAS 的标签分配过程

该策略计算每个目标  $G_i$  和每个预测边界框  $P_j$  的 IoU 矩阵, 如式 (3) 所示, 根据 IoU 值生成  $P_j$  对  $G_i$  的排名矩阵, 对于每个  $G_i$  挑选排名前  $K_i$  的预测框作为其正样本。然而, 传统的基于最大 IoU 的分配策略在处理预测框与多目标关联时 (如  $P_2$  和  $P_{33}$ ), 会导致类正方形目标优先获得正样本, 进一步加剧大纵横比目标样本数量不足的问题。

$$\text{IoU}_{ij} = \frac{\text{Area}(G_i \cap P_j)}{\text{Area}(G_i \cup P_j)}, \text{IoU} \in [0, 1] \quad (3)$$

为解决上述问题, ALAS 的创新在于, 优先将候选框分配给排名更靠前的目标 (如将  $P_2$  优先分配给  $GT_0$ ), 若多个目标的排名相同, 则根据预测框与目标的 Top IoU 值进行分配 (如将  $P_{33}$  优先分配给  $GT_1$ )。在实现过程中, Top IoU 值较大的目标表明该目标在当前图像中拥有较多高质量样本, 因此, 优先将 Top IoU 值较小的目标的标签分配给该候选框, 以平衡样本分布。最终按照排名矩阵的顺序补充新样本, 直至每个目标的样本数量与其  $K$  值相匹配, 从而更合理地分配正样本标签, 确保每个目标获得高质量的正样本用于训练。自适应标签分配过程如算法 1 所示。

#### 算法 1 自适应标签分配过程

**输入** 目标集合  $G_N$ , 预测框集合  $P_M$ , 目标的纵横比敏感  $k$  估计集合  $K_N$

**输出** 正样本集合  $\mathcal{P}$ , 负样本集合  $\mathcal{N}$

**for**  $G_i \in G_N, P_j \in P_M$  **do**:

$$\text{IoU}[i][j] = \frac{\text{Area}(G_i \cap P_j)}{\text{Area}(G_i \cup P_j)};$$

**end for**

**for**  $G_i \in G_N$  **do**:

Rank[ $i$ ] = sort(IoU $_i$ );

$\mathcal{P}_i = \text{top}(\text{Rank}[i], \text{round}(K_N[i]));$

**end for**

**for**  $P_j \in P_M$  **do**:

**if**  $P_j \rightarrow G_m, G_n$  **do**:

**if** Rank[ $m$ ][ $j$ ] == Rank[ $n$ ][ $j$ ] **do**:

$P_j \rightarrow \min(K_m, K_n);$

**else**:

$P_j \rightarrow \max(\text{Rank}[m][j], \text{Rank}[n][j]);$

**end if**

**end if**

**end for**

**for**  $G_i \in G_N$  **do**:

**if** sum( $\mathcal{P}_i$ ) <  $K_N[i]$  **do**:

从 Rank[ $i$ ] 中选择下一个  $P_{\text{next}}$  加入  $\mathcal{P}_i$

**end if**

**end for**

$\mathcal{N} = P_M - \mathcal{P}$

**return**  $\mathcal{P}, \mathcal{N}$

## 1.4 中心轴先验

为筛选高质量样本, 现有方法 (如 SimOTA<sup>[19]</sup>) 通常采用中心先验策略, 即优先选择位于目标中心附近圆形区域内的样本。该策略基于“靠近目标中心的描框能更好表征目标”的假设, 但对大纵横比目标的适用性有限, 这是因为中心先验仅提供目标的粗略中心位置信息, 无法有效反映其方向与形状特征。为此, 本文提出 CAP, 将传统的圆形先验区域扩展为矩形区域, 并根据目标的纵横比和尺度自适应调整其范围, 对于目标  $(c_x, c_y, w, h, \text{angle})$ , 其中心轴先验框定义为  $(c_x, c_y, \frac{w}{2}, \frac{h}{2}, \text{angle})$ , 通过限制样本中心点位于中心轴先验框内, CAP 策略能够更精确地捕捉目标的位置与方向信息, 从而筛选出更高质量的样本。中心先验与中心轴先验的对比如图 6 所示。

## 2 实验结果与分析

实验采用 NVIDIA A10 图形处理单元 (graph-

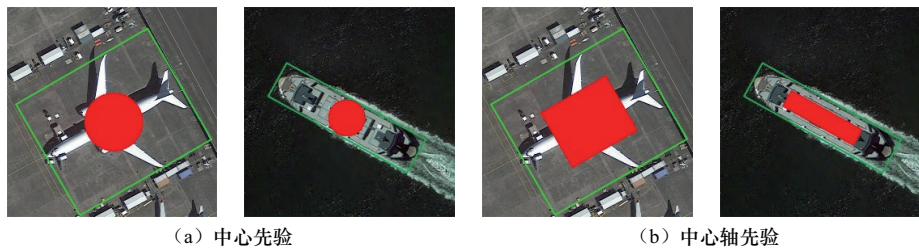


图6 中心先验与中心轴先验的对比

ics processing unit, GPU) 计算平台, 配置 Intel Xeon Platinum 8369B 处理器 (主频 2.90 GHz)。软件开发环境基于 Ubuntu 20.04 LTS 系统, 依托 PyTorch 1.11.0 框架实现算法, CUDA 11.3 提供 GPU 加速支持。此外, 本文分别在 DOTA 和 HRSC2016 数据集上进行实验。

## 2.1 实验参数

本文采用在 ImageNet 上预训练的 ResNet-50 作为骨干网络。训练过程中, 采用随机梯度下降 (stochastic gradient descent, SGD) 优化器, 在单块 A10 GPU 上以批量大小为 2 进行训练。

针对 DOTA 数据集的大尺寸图像, 本文采用单尺度与多尺度分割策略<sup>[21]</sup>。单尺度分割中, 通过滑动窗口的方式将图像切分成  $1024 \times 1024$  像素的切片, 滑动步长为 824 像素, 即相邻图像块的重叠区域为 200 像素, 缩放倍数为 1。多尺度分割中, 采用相同的方式将图像切分成  $1024 \times 1024$  像素大小的切片, 步长为 524 像素, 并在 0.5、1 和 2 这 3 个尺度上进行图像缩放。训练总轮数为 12, 权重衰减和动量分别设置为 0.000 1 和 0.9, 初始学习率为 0.01, 并在每次学习率衰减时降为原来的 1/10。训练集和验证集用于模型训练, 其余数据用于测试。在 DOTA 数据集上, 采用  $mAP_{50}$  (IoU 阈值为 0.5) 作为评价指标, 所有结果均通过 DOTA 官方评估服务器验证。

对于 HRSC2016 数据集, 保持图像原始纵横比不变, 将短边调整为 800 像素, 长边限制为不超过 1 333 像素。总训练轮数为 36, 初始学习率设置为 0.005, 并在第 24 轮和第 33 轮时将学习率

降为原来的 1/10。训练集和验证集用于模型训练, 其余数据用于测试。在 HRSC2016 数据集上, 采用平均精度均值 (mean average precision, mAP) 作为检测精度的评价指标。

## 2.2 对比实验

为全面评估 SALANet 的性能, 本文在 HRSC2016 数据集和 DOTA 数据集上进行了对比实验。不同方法在 HRSC2016 数据集和 DOTA 数据集上的检测结果分别见表 1、表 2。在 HRSC2016 数据集上, 将 SALANet 与 5 种主流定向目标检测方法进行对比; 在 DOTA 数据集上, 则与 8 种方法进行对比。其中, R-50 表示 50 层残差网络 (ResNet-50), R-101 表示 101 层残差网络 (ResNet-101), H-104 表示 104 层沙漏网络 (Hourglass-104), DLA-34 表示 34 层深层聚合网络 (deep layer aggregation), VIT 表示 vision Transformer 网络。表 1 和表 2 中的加粗数值表示最优精度。

表 1 不同方法在 HRSC2016 数据集上的检测结果

方法	骨干网络	mAP (07)	mAP (12)
RRPN <sup>[4]</sup>	R-101	0.790 8	0.856 4
S <sup>2</sup> ANet <sup>[15]</sup>	R-50	0.831 6	0.897 5
DAL <sup>[16]</sup>	R-50	0.832 2	0.899 0
RoI Transformer <sup>[5]</sup>	R-101	0.862 0	0.903 6
Gliding Vertes <sup>[6]</sup>	R-101	<b>0.882 0</b>	—
SALANet	R-101	0.875 0	<b>0.932 3</b>

HRSC2016 数据集<sup>[22]</sup>是由西北工业大学于 2016 年发布的公开的遥感图像舰船检测数据集。由表 1 可知, SALANet 在两个指标上分别达到

表2 不同方法在DOTA数据集上的检测结果

方法	骨干网络	不同类别的mAP										mAP
		棒球场 (BD)	桥梁 (BR)	小型车辆 (SV)	大型车辆 (LV)	船舰 (SH)	篮球场 (BC)	体育场 (ST)	港口 (HA)	游泳池 (SP)	直升机 (HC)	
IENet <sup>[11]</sup>	R-50	0.713 8	0.342 6	0.637 8	0.656 3	0.716 1	0.710 7	0.736 3	0.480 7	0.605 3	0.495 3	0.612 4
RoI Transformer <sup>*[15]</sup>	R-101	0.785 2	0.434 4	0.688 1	0.736 8	0.835 9	0.772 7	0.814 6	0.628 3	0.589 3	0.476 7	0.695 6
DAL <sup>*[16]</sup>	R-50	0.796 9	0.462 7	0.658 9	0.761 0	0.785 3	0.799 8	0.784 1	0.692 3	0.713 2	0.606 5	0.711 8
O <sup>2</sup> DETR <sup>*[17]</sup>	R-50	0.759 2	0.460 2	<b>0.797 0</b>	0.799 3	0.891 7	0.811 9	0.760 0	0.642 2	0.658 0	0.589 6	0.721 5
F-O <sup>2</sup> DETR <sup>[21]</sup>	R-50	0.819 1	0.512 0	0.776 4	0.804 7	0.878 4	0.845 6	0.816 8	0.675 0	0.6428	0.621 5	0.744 7
DAL <sup>[16]</sup>	R-101	0.791 9	0.521 7	0.790 8	0.781 0	0.867 9	0.845 3	0.837 3	0.725 0	0.703 9	0.538 4	0.747 4
AOPG <sup>[7]</sup>	R-101	0.827 4	0.518 7	0.776 5	0.824 2	0.880 8	0.862 6	0.851 3	0.740 5	0.677 6	0.587 7	0.753 9
无描框 Transformer <sup>[23]</sup>	VIT	0.793 4	0.569 1	0.785 3	0.842 9	<b>0.899 3</b>	0.865 6	0.836 0	0.658 9	0.696 3	0.645 9	0.758 3
SALANet	R-50	0.829 7	0.490 5	0.752 6	0.764 2	0.878 4	0.860 8	0.855 0	0.670 5	0.71 14	0.554 9	0.739 2
SALANet*	R-50	<b>0.833 5</b>	0.553 7	0.783 2	0.797 4	0.883 0	0.848 9	0.854 1	<b>0.760 8</b>	0.714 9	<b>0.688 0</b>	0.772 5
SALANet	R-101	0.819 5	0.517 3	0.744 4	0.789 8	0.861 1	0.863 4	0.850 1	0.752 8	0.660 6	0.533 8	0.744 5
SALANet*	R-101	0.831 2	<b>0.565 3</b>	0.783 5	<b>0.843 1</b>	0.887 2	<b>0.866 2</b>	<b>0.862 3</b>	0.756 5	<b>0.732 2</b>	0.650 3	<b>0.777 1</b>

注:\*表示采用多尺度分割作为数据增强手段。

0.875 0 和 0.9323, 相较于基线方法 (RoI Transformer) 分别提升约 1.5% 和 3.2%。这一性能提升主要归因于以下 3 方面的改进。

(1) 通过引入 ARSK 估计模块, 根据目标纵横比动态调整目标的样本分配数量, 增加大纵横比目标正样本数量, 有效缓解传统方法中正样本分布不均衡导致的大纵横比目标学习不充分的问题。

(2) 结合自适应标签分配策略, 优先选择与目标 IoU 较高的高质量样本, 确保模型充分学习大纵横比目标的方向和形状特征。

(3) 构建中心轴约束区域, 使正样本能更精准地捕捉到船舰等大纵横比目标两端的细节信息, 从而增强对目标几何特征的代表能力。

DOTA<sup>[24]</sup> 是一个大型公开的遥感图像基准数据集, 其 v1.0 版本包含 2 806 张图像, 分辨率范围从 800×800 像素到 4 000×4 000 像素。数据集按 3:1:2 的比例划分为训练集、验证集和测试集<sup>[25]</sup>。表 2 展示了 SALANet 在 DOTA 数据集上 10 种类别的检测结果。由表 2 可知, SALANet\* (ResNet-101)

的 mAP 达到 0.777 1, 优于 AOPG (0.753 9) 和无锚框 Transformer (0.758 3), 体现了其整体检测性能的优越性。在具体类别上, SALANet 在大型车辆 (LV)、港口 (HA) 和直升机 (HC) 等大纵横比类别上表现尤为突出, 分别达到 0.843 1、0.760 8 和 0.688 0, 显著优于其他方法。此外, SALANet 在不同骨干网络 (ResNet-50 和 ResNet-101) 上均表现出稳定的性能提升, 验证了其对不同网络的兼容性。尽管在船舰 (SH) 类别上指标略低于无锚框 Transformer, 但 SALANet 在多数类别上表现均衡且领先, 综合性能更具优势。SALANet 通过引入纵横比敏感系数增加大纵横比目标的正样本数量, 结合 ALAS 选取高质量样本, 并利用 CAP 增强几何特征提取能力, 在复杂场景下的旋转目标检测任务中展现出强大的竞争力。DOTA 数据集上使用 R-101 主干网络的 SALANet 检测结果示例如图 7 所示, 在可视化这些结果时, 置信度阈值设置为 0.3, 每个颜色代表一个分类。

桥梁检测结果如图 8 所示。图 8 (a) 中, 对于常规几何特征目标 (如纵横比接近 1 的桥梁结



图7 DOTA数据集上使用R-101主干网络的SALANet检测结果示例

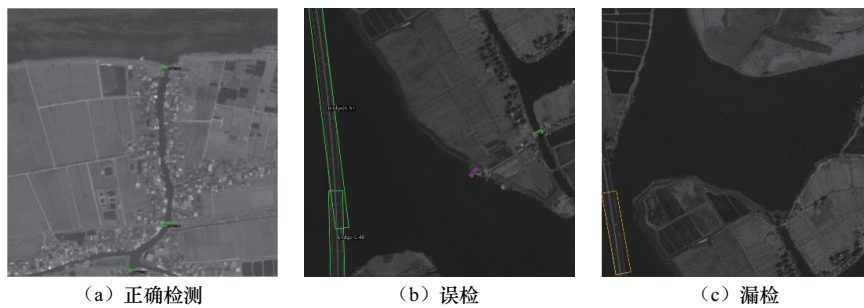


图8 桥梁检测结果

构), 检测系统能够实现精准定位与识别, 验证基础检测框架的有效性。然而, SALANet仍存在一定的局限性, 对于极端大纵横比目标, 模型可能出现目标分段误检现象, 表现为将单一桥梁结构错误分割为两段独立目标, 如图8(b)所示。此外, 当目标位于图像边缘并因截断使得特征不完整时, 模型在边界效应的影响下可能出现部分目标漏检, 如图8(c)所示。

在DOTA数据集上, 本文采用相同实验设置

对比了不同方法的精度与速度。不同方法下精度和速度对比如图9所示。所有方法均基于ResNet-50骨干网络, 并在单块RTX 2080Ti GPU上以批量大小为1、输入图像尺寸为1 024×1 024的条件下进行测试。不同方法下精度和速度对比结果见表3, SALANet在检测精度上显著优于对比方法, mAP达到0.739 2。然而, 由于增加了大纵横比目标的样本数量, 模型在推理阶段需要处理更多特征, 运行速度略有下降(仅13.9 FPS)。

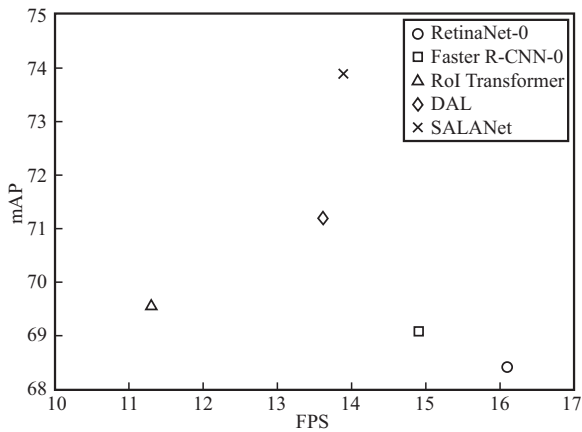


图9 不同方法下精度和速度对比

表3 不同方法下精度和速度对比结果

方法	FPS	mAP
RetinaNet-O <sup>[8]</sup>	16.1	0.684 3
Faster R-CNN-O <sup>[1]</sup>	14.9	0.690 5
RoI Transformer <sup>[5]</sup>	11.3	0.695 6
DAL <sup>[16]</sup>	13.6	0.711 8
SALANet	13.9	<b>0.739 2</b>

### 2.3 消融实验

在消融实验中，本文采用RoI Transformer作为基线模型，并采用ResNet-101作为骨干网络在DOTA数据集上进行实验，且未使用任何数据增强的方式。不同模块在DOTA1.0数据集上的消融实验结果见表4。其中，第1行为基线模型的检测结果，在此基础上，逐步引入各模块，以评估其对性能的提升效果。

表4 不同模块在DOTA1.0数据集上的消融实验结果

数据集	基线模型	ALAS	CAP	mAP
DOTA	√	—	—	0.695 6
	√	√	—	0.738 9 (+0.043 3)
	√	√	√	<b>0.744 5 (+0.048 9)</b>

#### 2.3.1 自适应标签分配策略的有效性

ALAS与传统的基于IoU阈值的标签分配方法不同，其核心在于，通过ARSK算法确定每个目标所需的正样本数量，并基于IoU排序机制优选高质量样本。该方法通过抑制小纵横比目标的

样本冗余，同时依据目标纵横比动态扩展大纵横比目标的候选样本池，有效缓解了样本分布不平衡问题。MaxIoU和ALAS对不同纵横比目标样本数量的影响如图10所示。由于ALAS有效缓解了小纵横比目标与大纵横比目标之间的样本数量差异，模型能够更专注于大纵横比目标的特征学习。实验结果表明，在不使用数据增强的情况下，ALAS模块在DOTA数据集上将SALANet的mAP提升至0.7389，验证了其在样本分布不平衡场景下的有效性。

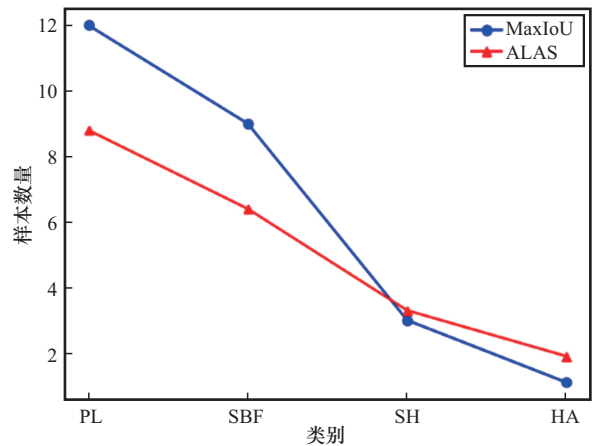


图10 MaxIoU和ALAS对不同纵横比目标样本数量的影响

#### 2.3.2 中心轴先验的有效性

中心轴先验的有效性主要源于其对大纵横比目标的先验区域的扩展设计。利用旋转目标中心轴的空间分布特性，该机制显著增强了模型对目标几何特征的建模能力。大纵横比目标的中心轴具有更强的方向一致性和空间连续性，使得中心轴先验能够精准定位目标的关键区域，从而优化特征表征。中心先验和中心轴先验对比如图11所示。与传统的中心先验相比，中心轴先验在特征提取过程中能够更有效地捕捉大纵横比目标两端的细节信息，同时强化对目标边缘特征的感知能力，显著提升了检测精度。实验结果表明，在DOTA数据集上，中心轴先验显著提升了目标的定位精度，将mAP提升至0.7445，验证了该方法的有效性。

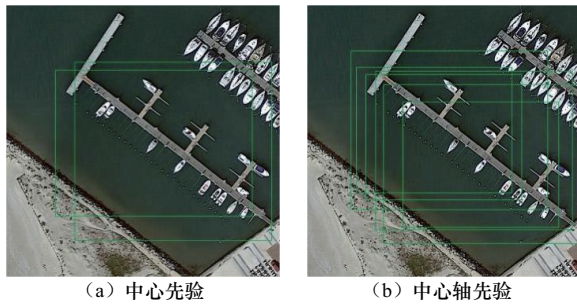


图11 中心先验和中心轴先验对比

### 2.3.3 纵横比敏感系数 $\alpha$ 的选择

针对大纵横比目标与预测框之间的IoU普遍较低的问题,本文在动态 $k$ 估计的基础上引入纵横比敏感系数 $\alpha$ ,以增加大纵横比目标的样本数量。为验证 $\alpha$ 对 $k$ 值的影响,在DOTA数据集上进行实验,记录了不同 $\alpha$ 值下多类代表性目标的实验结果。不同纵横比下 $\alpha$ 对 $k$ 值的影响结果见表5。表5中括号内标注为该类别在整个数据集中的平均纵横比, $k$ 值为同类别下所有正样本数量的均值。实验结果表明,随着系数 $\alpha$ 的减小,纵横比接近1的目标(如RA类别)的正样本数量基本保持稳定,而纵横比较大的目标则表现出显著的正样本数量增加趋势,验证了纵横比敏感系数对 $k$ 值的调节作用。进一步分析发现,随着 $\alpha$ 的减小,其对 $k$ 值的影响呈现非线性衰减趋势。尽管低 $\alpha$ 值增加了参与IoU统计的预测框数量,但由于大纵横比目标排名靠后的预测框IoU值较低,其对 $k$ 值的贡献逐渐减弱,体现了纵横比敏感系数在目标检测中的非线性调节特性。

表5 不同纵横比下 $\alpha$ 对 $k$ 值的影响结果

纵横比 敏感系数 $\alpha$	$k$ 值			
	RA (1.09)	SBF (1.65)	SH (3)	HA (4.5)
7	8.8	6.3	3.1	1.3
6	8.8	6.3	3.2	1.5
5	8.8	6.4	3.3	1.9
4	8.8	6.6	3.6	2.2
3	8.8	6.8	4.3	2.4
2	8.8	7.2	4.7	2.5
1	8.8	7.6	5.2	2.6

为探究纵横比敏感系数 $\alpha$ 对大纵横比目标识别精度的影响,本文在HRSC2016数据集上进行了实验分析, $\alpha$ 对精度的影响结果见表6。实验表明,随着 $\alpha$ 值的减小,大纵横比目标的正样本数量显著增加,同时检测精度逐步提升,并在 $\alpha=5$ 时性能达到峰值。然而,当 $\alpha$ 值进一步减小时,尽管正样本数量持续增加,但由于引入了IoU较低的低质量样本,识别精度出现下降。基于上述实验结果,本文将 $\alpha$ 值设置为5,以实现正样本数量与识别精度之间的最优平衡。

表6  $\alpha$ 对精度的影响结果

纵横比敏感系数 $\alpha$	mAP
7	0.893 4
6	0.921 5
5	<b>0.932 3</b>
4	0.915 4
3	0.881 6
2	0.862 3
1	0.811 0

## 3 结束语

本文针对遥感场景下大纵横比目标检测效果不佳的问题,提出一种基于形状自适应标签分配的有向遥感目标检测方法SALANet。该方法基于自适应的标签分配策略,将标签分配策略从使用固定的阈值转变为通过目标纵横比自适应调整,以缓解不同纵横比目标间样本数量不平衡的问题,并提出中心轴先验的区域,以更好地表征物体的形状和方向信息。在DOTA数据集上的测试表明,该方法取得了0.777 1的mAP,同时,在HRSC2016数据集上的测试结果进一步证明了SALANet的泛用性。

## 参考文献:

- [1] Ren S Q, He K M, Girshick R, et al. Faster R-CNN: towards real-time object detection with region proposal networks[C]//Proceedings of the IEEE Transactions on Pattern Analysis and Machine Intelligence. Piscataway: IEEE Press, 2016: 1137-1149.

- [2] Redmon J, Divvala S, Girshick R, et al. You only look once: unified, real-time object detection[C]//Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE Press, 2016: 779-788.
- [3] 姚婷婷, 肇恒鑫, 冯子豪, 等. 上下文感知多感受野融合网络的定向遥感目标检测[J]. 电子与信息学报, 2025, 47(1): 233-243.  
Yao T T, Zhao H X, Feng Z H, et al. A context-aware multiple receptive field fusion network for oriented object detection in remote sensing images[J]. Journal of Electronics & Information Technology, 2025, 47(1): 233-243.
- [4] 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.
- [5] Ding J, Xue N, Long Y, et al. Learning RoI transformer for oriented object detection in aerial images[C]//Proceedings of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE Press, 2020: 2844-2853.
- [6] Xu Y C, Fu M T, Wang Q M, 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(4): 1452-1459.
- [7] 胡齐. 基于特征回归和双中点向量表示的遥感目标检测算法研究[D]. 开封: 河南大学, 2024.  
Hu Q. Research on remote sensing object detection algorithm based on feature regression and double midpoint vector representation[D]. Kaifeng: Henan University, 2024.
- [8] Yu L B, Zhou Y, Li X L, et al. A rotating object detector with convolutional dynamic adaptive matching[J]. Applied Sciences, 2024, 14(2): 633.
- [9] 谢明杰. 基于自适应标签分配策略的遥感目标检测方法研究[D]. 哈尔滨: 哈尔滨工程大学, 2024.  
Xie M J. Research on remote sensing object detection method based on adaptive label assignment strategy[D]. Harbin: Harbin Engineering University, 2024.
- [10] 徐丹青, 吴一全. 光学遥感图像目标检测的深度学习算法研究进展[J]. 遥感学报, 2024, 28(12): 3045-3073.  
Xu D Q, Wu Y Q. Progress of research on deep learning algorithms for object detection in optical remote sensing images[J]. National Remote Sensing Bulletin, 2024, 28(12): 3045-3073.
- [11] Lin Y T, Feng P M, Guan J, et al. IENet: interacting embranchment one stage anchor free detector for orientation aerial object detection[PP]. arXiv (2019-12-02)[2021-03-29]. arXiv: arXiv. 1912.00969.
- [12] Xiao Z F, Qian L J, Shao W P, et al. Axis learning for orientated objects detection in aerial images[J]. Remote Sensing, 2020, 12(6): 908.
- [13] 李士博, 肖振久, 曲海成, 等. 面向SAR图像舰船检测的多粒度特征与形位相似度量方法[J]. 光电工程, 2025, 52(2): 44-58.  
Li S B, Xiao Z J, Qu H C, et al. Multi-granularity feature and shape-position similarity metric method for ship detection in SAR images[J]. Opto-Electronic Engineering, 2025, 52(2): 44-58.
- [14] Zhang S F, Chi C, Yao Y Q, et al. Bridging the gap between anchor-based and anchor-free detection via adaptive training sample selection[C]//Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE Press, 2020: 9756-9765.
- [15] 余颖. 复杂场景下遥感图像任意方向目标检测方法研究[D]. 西安: 西安电子科技大学, 2023.  
Yu Y. Research on arbitrary-oriented object detection methods in remote sensing images under complex scenes[D]. Xi'an: Xidian University, 2023.
- [16] Ming Q, Zhou Z Q, Miao L J, et al. Dynamic anchor learning for arbitrary-oriented object detection[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2021, 35(3): 2355-2363.
- [17] 董燕, 魏铭宏, 高广帅, 等. 基于双重标签分配的遥感有向目标检测方法[J]. 计算机科学, 2024, 51(S2): 496-504.  
Dong Y, Wei M H, Gao G S, et al. Remote sensing oriented object detection method based on dual-label assignment[J]. Computer Science, 2024, 51(S2): 496-504.
- [18] Zhang D W, Xiao X, Zheng Z L, et al. Probabilistic assignment with decoupled IoU prediction for visual tracking[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2024, 34(7): 5776-5789.
- [19] Huang Q Q, Yao R L, Lu X Q, et al. Oriented object detector with Gaussian distribution cost label assignment and task-decoupled head[J]. IEEE Transactions on Geoscience and Remote Sensing, 2024, 62: 5621916.
- [20] Ma Q H, Jin S H, Bian G, et al. S3DR-det: a rotating target detection model for high aspect ratio shipwreck targets in side-scan sonar images[J]. Remote Sensing, 2025, 17(2): 312.
- [21] Ge Z, Liu S T, Wang F, et al. YOLOX: exceeding YOLO series in 2021[PP]. V2 arXiv (2021-07-18) [2021-08-06]. arXiv: arXiv. 2107.08430.
- [22] 党思航, 李晓哲, 夏召强, 等. 采用自适应预筛选的遥感图像目标开集检测研究[J]. 电子与信息学报, 2024, 46(10): 3908-3917.  
Dang S H, Li X Z, Xia Z Q, et al. Research on open-set object detection in remote sensing images based on adaptive pre-screening[J]. Journal of Electronics & Information Technology, 2024, 46(10): 3908-3917.
- [23] 喻九阳, 胡天豪, 戴耀南, 等. 面向遥感目标检测的无锚框



Transformer算法[J]. 电子学报, 2023, 51(11): 3238-3247.

Yu J Y, Hu T H, Dai Y N, et al. Anchor-free Transformer algorithm for aerial remote sensing target detection[J]. Acta Electronica Sinica, 2023, 51(11): 3238-3247.

[24] Xia G S, Bai X, Ding J, et al. DOTA: a large-scale dataset for object detection in aerial images[C]//Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2018: 3974-3983.

[25] 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. Portugal: SCITEPRESS-Science and Technology Publications, 2017: 324-331.

[作者简介]



任王 (1981-), 男, 博士, 浙江工商大学信息与电子工程学院副教授, 主要研究方向为无线通信和信号处理。



吴斌 (2000-), 男, 浙江工商大学信息与电子工程学院硕士生, 主要研究方向为数字图像处理。



余长宏 (1978-), 男, 博士, 浙江工商大学信息与电子工程学院副教授, 主要研究方向为智能计算和大数据应用。



曾文捷 (2001-), 男, 浙江工商大学信息与电子工程学院硕士生, 主要研究方向为遥感图像目标检测。