中图法分类号:TP18;TP391.4　　**文献标识码**:A　　**文章编号**:1006-8961(2026)09-3427-14

论文引用格式:Wang Z Y, Ge Y, Wu M Q and Chen J L. 2026. Oriented object detection in remote sensing images with joint angle continuous representation and perceptual balance. Journal of Image and Graphics, 31(9):3427-3440(王朝阳, 葛芸, 吴梦奇, 陈金梁. 2026. 联合角度连续表示与感知平衡的遥感图像定向目标检测. 中国图象图形学报, 31(9):3427-3440)[DOI:10.11834/jig.250447]

联合角度连续表示与感知平衡的遥感图像定向目标检测

王朝阳,葛芸*,吴梦奇,陈金梁

1. 图像处理与模式识别江西省重点实验室,南昌　330063; 2. 南昌航空大学软件学院,南昌　330063

摘　要:**目的**　遥感图像定向目标检测在环境监测、城市规划等领域发挥着重要作用,基于二维高斯分布的定向边界框表示方法实现了对旋转交并比(skew intersection over union,SkewIoU)的近似计算,因此受到了广泛关注。但二维高斯分布表示方法中存在角度边界不连续问题,且对不同宽高比目标角度感知不平衡,导致定位不准确。因此,提出了一种联合角度连续表示与感知平衡的检测算法(joint angle continuous representation and perceptual balance,JCPB)。**方法**　首先,提出角度向量分解法(angular vector decomposition method,AVDM),通过在二维空间内将角度参数分解为向量编码,使得角度在边界和非边界位置处均能连续表示。并提出形状感知系数方法,根据目标宽高比动态调整系数值,在不增加编码长度的前提下调节角度损失函数的变化周期。其次,提出角度感知平衡策略(angle perceptual balance strategy,APBS),动态调节损失权重强化模型对小宽高比目标的方向判别能力。最后,基于二维高斯分布表示方法,通过联合角度连续表示和角度感知平衡策略,增强模型对目标方向的学习能力。**结果**　在DOTA(dataset for object detection in aerial images)、DIOR-R(dataset for object detection in aerial images)和HRSC2016(high-resolution ship collections)数据集上进行大量实验测试。结果表明,在DOTA、DIOR-R和HRSC2016数据集上,与基准方法相比,所提方法在平均精度均值(mean average precision,mAP)上平均提升分别为1.46%、1.31%和4.35%,在SkewIoU阈值为0.75时,平均精准率(average precision,AP)平均提升分别为3.31%、2.19%和6.07%,实验结果验证了所提算法的有效性。**结论**　本文算法通过将角度参数分解为连续的向量表示,并平衡模型对不同宽高比目标的感知能力,实现了遥感图像定向目标的高精度检测。

关键词:定向目标检测;高斯分布;边界不连续;角度向量分解;角度感知平衡

Oriented object detection in remote sensing images with joint angle continuous representation and perceptual balance

Wang Zhaoyang, Ge Yun*, Wu Mengqi, Chen Jinliang

1. Jiangxi Province Key Laboratory of Image Processing and Pattern Recognition, Nanchang　330063, China;

2. School of Software, Nanchang Hangkong University, Nanchang　330063, China

Abstract: **Objective** Remote sensing image object detection is fundamental to the automatic interpretation of aerial and satellite imagery. It aims to identify objects of predefined categories from given images and regress the localization informa-

收稿日期:2025-09-25;**修回日期**:2025-12-27;**预印本日期**:2026-01-03

*通信作者:葛芸　geyun@nchu.edu.cn

基金项目:国家自然科学基金项目(42261070,41801288);江西省自然科学基金项目(20252BAC240256)

Supported by: National Natural Science Foundation of China (42261070, 41801288); Natural Science Foundation of Jiangxi Province, China (20252BAC240256)

tion for each object. Remote sensing image object detection technology has broad applications in practical scenarios, such as intelligent surveillance, urban planning, port management, and military surveys. In contrast to natural images, where objects are typically horizontally aligned and backgrounds are relatively simple, remote sensing images are captured from a top-down perspective using spaceborne or airborne sensors, offering a much broader field of view. Consequently, objects in remote sensing images are often distributed in arbitrary orientations, and the background information is more complex and diverse. Although some horizontal object detectors perform well on natural images, horizontal bounding boxes (HBB) cannot describe the orientation of objects and often introduce unnecessary background redundancy. Moreover, in densely arranged scenes, the overlap between HBBs and neighboring boxes can be substantial, which may lead to missed or false detections. By contrast, oriented bounding boxes (OBB) can represent objects more accurately and are better suited for object detection in dense scenes. However, during regression, the skew intersection over union (SkewIoU) is difficult to compute and nondifferentiable, preventing the model from optimizing its predicted OBB parameters via backpropagation. To address this, current mainstream methods convert oriented bounding boxes into 2D Gaussian distribution and compute the distance between these distributions as the regression loss. This approach approximates the calculation of SkewIoU while remaining differentiable, and such methods are collectively referred to as Gaussian distance loss (GDL). However, given the periodic nature of angles, GDL-based methods exhibit discontinuities in angle regression. Furthermore, they suffer from imbalanced angle perception across targets with varying aspect ratios. The interplay between these issues considerably degrades the detection performance of the model. In this study, a detection algorithm named joint angle continuous representation and perceptual balance (JCPB) is proposed to mitigate the mentioned challenges. **Method** To address the angle boundary discontinuity problem in oriented object detection, this study proposes an angular vector decomposition method (AVDM), which decomposes angle parameters into an angle vector representation. To eliminate discontinuous regions inherent in traditional angular representations, this vector representation is further extended to encompass the entire unit circle plane, ensuring continuous variation of the angular vector during OBB regression. Aiming at the periodic discrepancy issue in angle loss calculation between rectangular and square targets, this study introduces an innovative shape aware coefficient method. Without increasing the encoding length of the angular vector, a shape-aware coefficient dynamically adjusts its value on the basis of the target's aspect ratio, thereby modulating the change period of the angular loss function. To address the problem of imbalanced angle perception, we proposed the angle perceptual balance strategy, which adaptively calculates loss weighting coefficients based on the aspect ratio of targets and adjusts their application timing according to the SkewIoU value. This enhances the model's angle perception capability for low-aspect-ratio targets during the later stages of training. Finally, on the basis of a 2D Gaussian distribution representation method, we further enhance the model's detection performance by joint angle continuous representation and perceptual balance. **Result** This study conducted comparative experiments using FCOS and RetinaNet detectors on the DOTA, DIOR-R, and HRSC2016 datasets to validate the effectiveness of the proposed methods. Furthermore, comparisons with current mainstream models on the DOTA dataset demonstrate the superiority of the proposed model. Evaluation metrics include average precision at 75% (AP_{75}) intersection over union threshold, mean average precision (mAP), and parameters. On the DOTA dataset, incorporating AVDM improves AP_{75} and mAP by 2.47% and 1.03% on average compared with the baseline model. Further incorporating APBS resulted in additional average improvements of 0.84% and 0.43% for AP_{75} and mAP, respectively. On the DIOR dataset, integrating AVDM achieved average gains of 1.37% and 0.95% for AP_{75} and mAP. Further incorporating APBS resulted in additional improvements of 0.82% and 0.36% for AP_{75} and mAP, respectively. On the HRSC dataset, the AVDM method achieved average gains of 5.63% and 3.86% for AP_{75} and mAP, respectively. After incorporating APBS, AP_{75} and mAP improved by 0.44% and 0.49% on average. Experimental results across multiple datasets demonstrate that APBS effectively enhances the model's angular learning capability for small aspect ratio targets, while AVDM achieves continuous angular representation, effectively resolving boundary discontinuity issues. Compared with other encoding methods, AVDM introduces only 0.04 M additional parameters, achieving continuous angular representation and periodic adaptation to angular losses for rectangular and square targets at minimal parametric cost. Experiments on the HRSC dataset show that detection accuracy for slender objects is highly susceptible to boundary discontinuity issues, while angle perceptual imbalance has a relatively weaker impact. Finally, comprehensive comparisons with current mainstream

remote sensing image orientation object detection methods on the DOTA dataset demonstrate that our approach achieves an mAP of 73.78%, outperforming other methods. Moreover, it achieves optimal accuracy on small vehicles, ships, tennis courts, storage tanks, and roundabouts categories. Additionally, when using Swin-Transformer as the backbone network, the model's performance is further improved, reaching an mAP of 74.71%. **Conclusion** Extensive experimental results demonstrate that the proposed AVDM achieves continuous angular representation, while APBS effectively enhances the model's perceptual capability for small aspect ratio targets. Building upon this foundation, JCPB remarkably improves the overall performance of the detection model in complex remote sensing scenarios by integrating continuous angular representation with perceptual balance.

Key words: oriented object detection; Gaussian distribution; boundary discontinuity; angle vector decomposition; angle perceptual balance

0 引　言

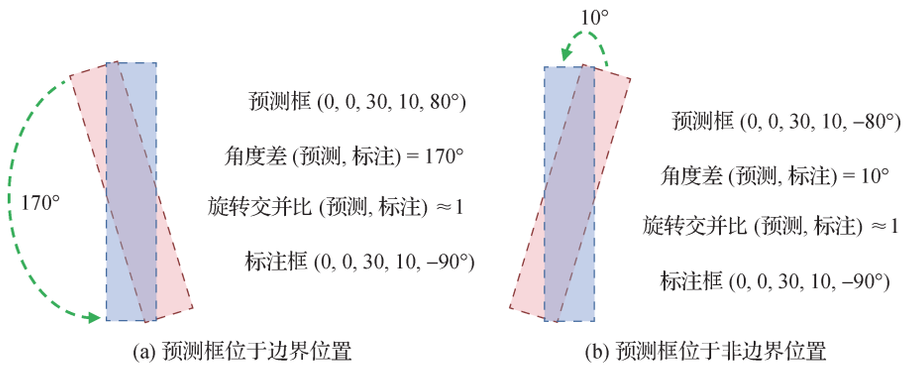
遥感图像目标检测任务旨在识别和定位遥感图像内感兴趣目标(徐丹青和吴一全,2024)。与传统自然图像不同,遥感图像中目标在复杂的背景下,通常呈现任意方向和密集分布的特点,且存在显著的宽高比差异。为提升复杂背景下目标的识别能力,Tu等人(2024)设计了一种双分支架构分别提取图像的高频局部特征与低频全局特征,实现了多源遥感数据在特征层面的有效融合。He等人(2024)通过设计多路径光谱空间扫描策略,有效建模三维高光谱数据的全局依赖关系,显著提升了分类精度与计算效率。

在目标检测方面,传统的水平边界框(horizontal bounding box,HBB)检测器,如FCOS(fully convolutional one-stage object detection)(Tian等,2019)、RetinaNet(Lin等,2017)、Faster R-CNN(faster region-based convolutional neural network)(Ren等,2017)和YOLO(you only look once)(Redmon等,2016)等,在目标分布密集的场景中容易发生检测框大面积重叠。尤其是当物体形状不规则或倾斜时,重叠框可能覆盖多个目标,导致后续非极大值抑制(non-maximum suppression,NMS)时难以区分正确检测框,造成漏检或误检。为解决上述问题,相关研究在HBB的基础上引入目标方向信息,提出基于定向边界框(oriented bounding boxes,OBB)的检测方法。Ma等人(2018)通过引入旋转锚框(rotation anchors)和RRoI(rotated region of interest)池化层检测定向目标。Ding等人(2019)提出的RoI Transformer实现了从水平RoI到RRoI的直接转换。张磊等人(2023)利用Faster R-CNN的检测分支对RRoI进行了精调。此外,佘浩东和赵良瑾(2024)通过中心点指向头部顶点的引导向量和目标宽度来表示定向边界框。

但是上述方法均存在角度边界不连续问题,如图1所示,图1(a)(b)中预测框与标注框之间的旋转交并比(skew intersection over union,SkewIoU)相同。在常用的长边定义法(Ma等,2018)下,图1(a)(b)中预测框的角度分别被定义为80°和-80°,标注框的角度均为-90°。在图1(a)中,预测框位于角度边界位置,其与标注框的角度差达170°。而在图1(b)中,预测框处于非边界位置,与标注框的角度差仅为10°。这种角度差的不一致现象源于角度的周期性,尽管两个角度在几何上接近,但跨越了角度定义的边界,角度差在数值上被放大。角度的周期性也导致正方形与长方形目标对应的角度损失具有不同的变化周期。上述问题共同导致SkewIoU与角度损失在优化方向上出现偏差,模型在训练过程中容易出现梯度信号混淆,进而影响网络的收敛性与检测精度。

在训练过程中,为了使角度损失与SkewIoU具有一致性,Yang和Yan(2020)提出CSL(circular smooth label),将角度估计从回归问题转化为分类问题,并通过环形的标签设置以增加相邻角度估计误差的容忍性。为了降低计算量和参数量,Yang等人(2021a)基于CSL角度分类方法提出了DCL(densely coded labels),通过密集编码标签的方式减少了编码长度。后续,Wang等人(2022)提出基于高斯衰减的动态加权损失函数,使模型更聚焦于精确角度的预测。Zeng等人(2024)提出的AR-CSL(aspect ratio aware circle smooth label)将宽高比引入角度编码标签,提高了大宽高比目标对角度损失变化的敏感性。

虽然角度分类方法避免了边界问题,但存在角度量化误差。同时,该方法未区分正方形和长方形



(a) 预测框位于边界位置　　(b) 预测框位于非边界位置

图1　角度边界不连续问题

Fig. 1　Angle boundary discontinuity problem ((a) predicted bounding box at the boundary position; (b) predicted bounding box at a non-boundary position)

在角度损失周期上的差异，并且单独预测角度缺少与中心点及宽高参数间的关联。因此，Yang等人(2021b)提出了一种创新的定位损失范式，通过将OBB建模为二维高斯概率分布，计算概率分布间的距离来实现OBB损失评估，该损失统称为高斯距离损失(Gaussian distance loss，GDL)。该系列方法包含基于Wasserstein距离的GWD(Gaussian Wasserstein distance)(Yang等，2021b)、采用KL散度的KLD(Kullback-Leibler divergence)(Yang等，2021c)以及利用卡尔曼滤波(Kalman filtering，KF)原理推导交并比(intersection over union，IoU)近似值的KFIoU(Yang等，2023)。与对边界框各参数独立优化的角度分类方法不同，GDL联合边界框的5个参数构建全局距离度量，这种联合优化的方式使得各个参数的优化方向与SkewIoU减小方向具有一致性，从而缓解了角度边界不连续问题。

然而，基于GDL的方法存在角度感知不平衡问题。如图2所示，对比两个具有相同中心、宽度和高度的边界框，在不同宽高比下GWD和KLD损失归一化后随角度差的变化曲线，可以看出随着边界框宽高比减小，两种GDL对角度偏差的敏感度呈现下降的趋势，从而削弱了网络对小宽高比目标角度的学习。为此，Li等人(2024)设计了一种基于GDL的梯度再平衡策略以缓解该问题。但是最近研究发现推理过程中也会存在边界问题，而GDL损失函数难以解决。因此，Yu和Da(2024)引入相移编码(phase-shifting coder，PSC)以实现角度的连续表示，并且采用双频编码(dual-frequency phase-shifting coder，PSCD)，解决长方形与正方形目标角度损失周期上的差异，但引入了额外的编码长度。

综上，本文提出联合角度连续表示与感知平衡的检测算法(joint angle continuous representation and perceptual balance，JCPB)。针对边界不连续问题，提出角度向量分解法(angular vector decomposition method，AVDM)，实现了角度的连续表示，并且避免了角度量化误差。针对长方形目标和类方形目标之

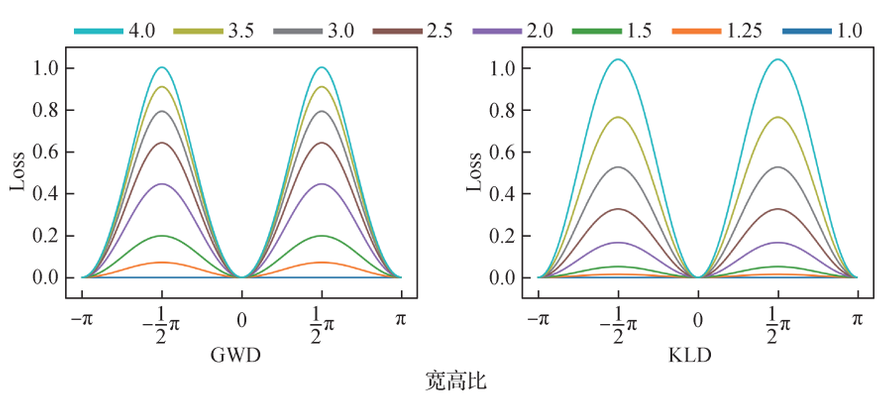


图2　不同宽高比目标在GDL上的对比

Fig. 2　Comparison of targets with different aspect ratios on GDL

间角度损失的周期差异，引入形状感知系数，在不增加编码长度的情况下处理类正方形目标。针对GDL对角度感知不平衡问题，提出角度感知平衡策略（angle perceptual balance strategy，APBS），基于目标宽高比自适应调节权重系数，增加模型对小宽高比目标角度信息的关注度。最后，在低参数的回归框架内，通过角度连续表示和角度感知平衡的协同优化，同时解决了角度边界不连续和感知不平衡两个问题，显著提升了模型对定向目标的检测精度。

1　研究方法

1.1　角度向量表示的定向目标检测框架

算法整体框架如图3所示，首先骨干网络对输入图像进行多尺度特征提取，将提取到的特征输入到颈部网络，对不同尺度的特征信息进行融合，再将特征送入检测头中进行目标边界框参数的预测，包括目标中心点及宽高(x, y, w, h)、目标角度向量表示$(\boldsymbol{m}_1, \boldsymbol{m}_2)$和目标类别得分（*classification*）。得到边界框的预测参数后，首先计算角度向量模长损失和分类损失，再将角度向量表示$(\boldsymbol{m}_1, \boldsymbol{m}_2)$解码得到预测角度$\theta$，用于计算角度损失。最后，将预测角度$\theta$与$(x, y, w, h)$预测参数联合得到预测框参数，再经过预测框参数解码得到预测框的表示，用于计算预测框损失。本文提出的JCPB算法，旨在从角度表示与损失优化两个方面协同提升模型对目标方向的学习能力，下文将详细介绍其核心组成部分。

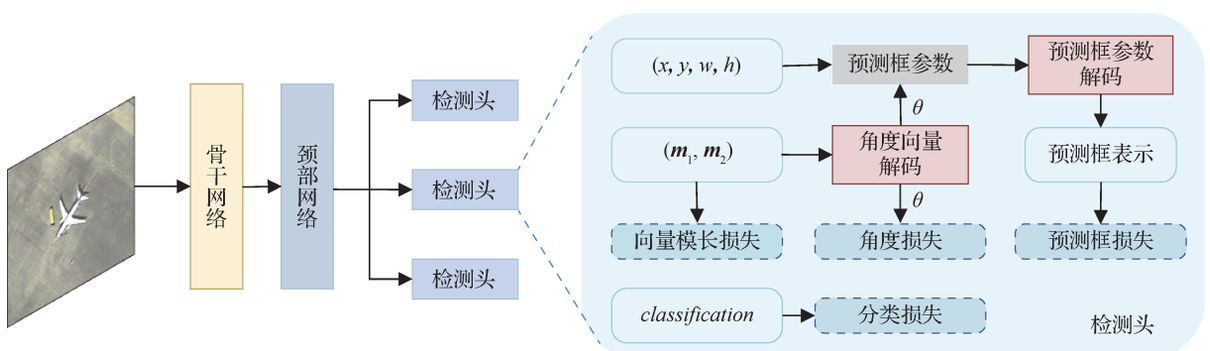


图3　角度向量表示的定向目标检测框架

Fig. 3　Oriented object detection framework with angular vector representation

1.2　角度向量分解方法

目标在旋转的过程中，中心点及宽高参数保持不变，由于定向边界框OBB角度的周期性导致了边界不连续问题。实现OBB角度的连续表示是解决边界不连续问题的关键。因此，本文提出角度向量分解法AVDM，以连续编码角度。

本文将预测OBB的角度定义为θ_{p}，长边定义法下OBB角度定义为θ_{box}，θ_{box}和θ_{p}之间的转换关系可以表示为

$$\theta_{\mathrm{box}}=\begin{cases}((\theta_{\mathrm{p}}+\pi)\bmod\pi)-\dfrac{\pi}{2} & h>w\\ ((\theta_{\mathrm{p}}+\dfrac{\pi}{2})\bmod\pi)-\dfrac{\pi}{2} & \text{其他}\end{cases}\tag{1}$$

式中，h表示短边，w表示长边。当$h>w$时需要交换长短边，以此统一OBB的表示，交换过程如图4所示。由图4可知，OBB旋转一周定义的角度θ_{box}不连续变化。

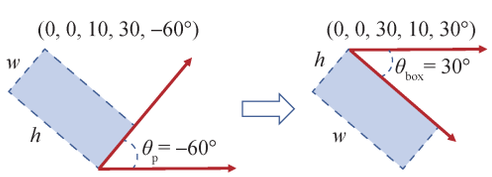


图4　长短边交换过程

Fig. 4　The process of swapping the long side and the short side

角度θ_{box}和向量编码数值的变化曲线如图5所示，θ_{box}的变化周期为π，并且旋转过程中θ_{box}在$\pm\pi/2$处不连续。为了得到连续角度表示，将角度向量的表示扩展至整个单位圆平面，使得其范围能够覆盖完整的2π周期，从而消除原有表示中的不连续区域（圆形坐标空间内阴影区域）。扩展后角度向量表示为$(\cos2\theta_{\mathrm{box}}, \sin2\theta_{\mathrm{box}})$，此时可以取单位圆上所有的向量来连续编码角度$\theta_{\mathrm{box}}$，并且OBB旋转过程中角度向量表示的数值始终保持连续变化。

θ_{box}与角度向量表示之间可以相互转化，它们之

间的关系为

$$\theta_{\text{box}} = \frac{1}{2}\arctan 2(\cos 2\theta_{\text{box}}, \sin 2\theta_{\text{box}}) \tag{2}$$

式中，arctan2 函数是一个象限函数，可以计算点在直角坐标轴上的弧度信息，其取值范围为$(-\pi, \pi]$，它的数学表达式为

$$\arctan 2(y,x) = \begin{cases} \arctan\left(\dfrac{y}{x}\right) & x > 0 \\ \arctan\left(\dfrac{y}{x}\right) + \pi & x < 0, y \geqslant 0 \\ \arctan\dfrac{y}{x} - \pi & x < 0, y < 0 \\ \dfrac{\pi}{2} & x = 0, y > 0 \\ -\dfrac{\pi}{2} & x = 0, y < 0 \end{cases} \tag{3}$$

在角度转化的过程中，编码函数可以将角度θ_{box}转化为向量编码，通过解码函数可以根据编码值还原出目标的角度θ_{box}，编码和解码函数可以表示为

$$\begin{cases} f_{\text{enc}}(\theta_{\text{box}}) = (\cos 2\theta_{\text{box}}, \sin 2\theta_{\text{box}}) = (\boldsymbol{m}_1, \boldsymbol{m}_2) \\ f_{\text{dec}}(\boldsymbol{m}_1, \boldsymbol{m}_2) = \dfrac{1}{2}\arctan 2(\boldsymbol{m}_2, \boldsymbol{m}_1) = \theta_{\text{box}} \end{cases} \tag{4}$$

式中，$f_{\text{enc}}(\cdot)$是将OBB的角度θ_{box}编码为角度向量表示$(\boldsymbol{m}_1, \boldsymbol{m}_2)$的函数，角度编码参数的数值为$(\cos 2\theta_{\text{box}}, \sin 2\theta_{\text{box}})$。$f_{\text{dec}}(\cdot)$是将角度向量表示$(\boldsymbol{m}_1, \boldsymbol{m}_2)$解码为$\theta_{\text{box}}$的函数。通过角度向量编码和解码的方式，解决了角度边界不连续问题。

1.3 损失函数

角度向量可以连续表示角度，但完整的检测框架还需分类损失与边界框回归损失。并且长方形与类正方形目标在角度损失上存在的周期性差异，也需在损失函数中考虑。为此，本节将详细介绍角度损失和边界框损失的具体计算方法。

首先，约束角度向量始终位于单位圆上，使得角度对应的向量编码具有唯一性，引入向量模长损失L_{cir}表示为

$$L_{\text{cir}} = \left|\boldsymbol{m}_1^2 + \boldsymbol{m}_2^2 - 1\right| \tag{5}$$

式中，$\boldsymbol{m}_1$和$\boldsymbol{m}_2$表示预测的角度向量表示参数。接下来，通过式(4)中解码函数$f_{\text{dec}}(\cdot)$，将$\boldsymbol{m}_1$和$\boldsymbol{m}_2$解码得到OBB的预测角度θ_{box}。

如图6(a)所示，长方形目标具有2种方向表示方式。在回归过程中，应选取与目标方向夹角最小的方向作为回归目标，即图中红色虚线箭头所指的方向。然而，由于存在边界不连续问题，与预测方向最接近的目标方向可能在不连续的区域。因此，进一步使用所提出的角度向量表示方法，消除了不连续区域，此时长方形目标的2个方向的向量表示保持一致。如图6(c)所示，在类正方形目标中也存在类似的问题，但类正方形目标具有4种方向表示方式。因此，长方形和类正方形目标的角度损失具有周期差异，角度损失周期应分别为π和$\pi/2$。

如图6(b)(d)所示，周期性的绝对值正弦函数与角度损失变化的周期一致，并且在边界位置处其数值可以反映OBB与目标真实的角度差。因此使用绝对值正弦函数作为角度损失，以此更新角度向量参数。角度损失函数L_{angle}表示为

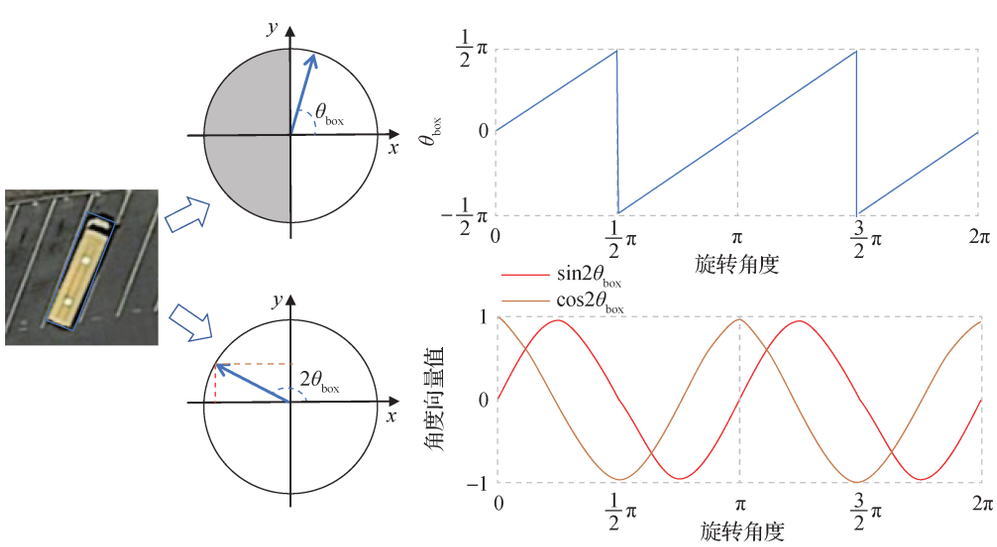


图5 角度非连续表示与连续表示值的比较

Fig. 5 Comparison between discontinuous and continuous angular representation values

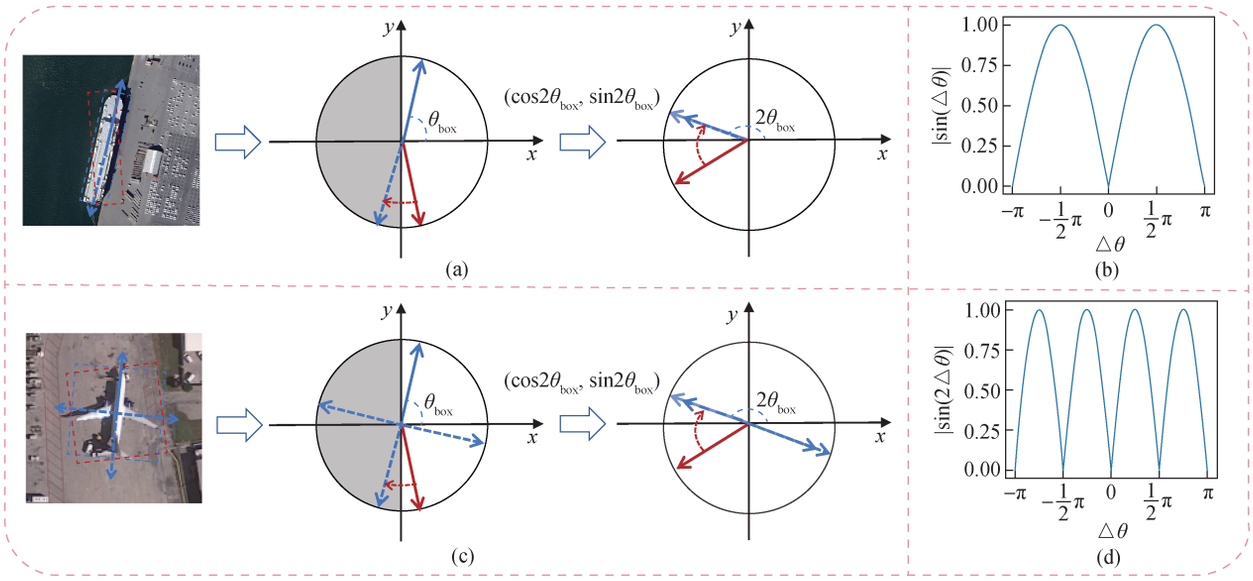


图 6　连续的角度向量表示

Fig. 6　Continuous angular vector representation

$$L_{angle} = \left|\sin\lambda(\theta_t - \theta_{box})\right| \tag{6}$$

式中，θ_t 为真实角度。λ 为形状感知系数，λ 值确定条件为

$$\lambda = \begin{cases} 1 & r_t \geqslant \varepsilon \\ 2 & 1 \leqslant r_t < \varepsilon \end{cases} \tag{7}$$

式中，r_t 为 GT(ground truth)框的宽高比，ε 为宽高比阈值。当 $r_t \geqslant \varepsilon$ 时目标为长方形，此时 λ 设置为 1；当 $r_t \geqslant 1$ 和 $< \varepsilon$ 时目标趋近于类正方形，此时 λ 设置为 2，通过动态 λ 值的设定，以应对长方形和类正方形角度损失不同的周期。

针对 GDL 中角度感知不平衡问题，本文提出角度感知平衡策略 APBS，通过将自适应权重系数与 GDL 损失函数直接相乘，得到新的损失，其中自适应权重系数设计为

$$s = 1 + \frac{IoU^{\beta}}{r_t - 1 + \alpha}, \quad r_t \geqslant 1 \tag{8}$$

式中，s 为自适应权重系数，超参数 β 用于控制权重曲线的上升速率，IoU 为每个正样本与标注框的交并比，超参数 α 用于确保数值计算的稳定性。自适应权重系数基于 IoU 进行动态计算，在训练初期 IoU 约等于 0，此时权重系数 s 约等于 1，对模型训练的影响较小。在训练后期模型对 OBB 预测较为准确，角度感知不平衡问题较为显著。此时，小宽高比目标将会获得一个较大的权重系数 s，将权重系数 s 与角度损失以及 GDL 相乘，从而平衡模型对不同宽高比目标角度的关注。

最后，将向量表示参数 $(\boldsymbol{m}_1, \boldsymbol{m}_2)$ 解码为 θ_{box} 后，联合中心点宽高参数 (x_p, y_p, w_p, h_p) 得到预测框的表示，计算边界框回归损失 L_{box}，损失表示为

$$L_{box} = L_{GDL}(B_p(x_p, y_p, w_p, h_p, \theta_{box}), B_t(x_t, y_t, w_t, h_t, \theta_t)) \tag{9}$$

式中，B_p 和 B_t 分别表示为预测 OBB 和 GT 的参数，L_{GDL} 表示基于 GDL 的边界框回归损失，包括 KLD、GWD 和 KFIoU。

联合本文提出的向量模长损失 L_{cir}、角度损失 L_{angle} 和自适应权重系数 s，总损失 L_{total} 可以表示为

$$L_{total} = w_{cls}L_{cls} + w_{box}L_{box} \times s + w_{angle}L_{angle} + w_{cir}L_{cir} \tag{10}$$

式中，L_{cls} 为分类损失，在这里本文使用 Focal Loss，w_{cls}、w_{box}、w_{angle} 和 w_{cir} 分别代表 L_{cls}、L_{box}、L_{angle} 和 L_{cir} 的权重。在边界框回归损失的基础上，联合角度损失和自适应权重系数 s，实现了对角度表示和边界框回归的同时优化。

2　实验结果

2.1　数据集

DOTA（dataset for object detection in aerial images）(Xia 等，2018)数据集由 2 806 幅大型航空图像组成，其中，1 411 幅用于训练，937 幅用于验证，458 幅用于测试。该数据集使用 15 个类别进行注释，总共有 188 282 个实例。类别定义为：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)、船舶(SH)、网球场(TC)、篮球场(BC)、储罐(ST)、足球场(SBF)、环岛(RA)、港口(HA)、游泳池(SP)和直升机(HC)。

DIOR-R (dataset for object detec tion in aerial images)(Cheng等,2022)是一个用于光学遥感图像中目标检测的大规模基准数据集,由23 463幅图像和192 518个带OBB注释的目标实例组成,其中,5 682幅用于训练,5 863幅用于验证,11 738幅用于测试。DIOR-R数据集涵盖了广泛的场景和20个常见的目标类别。类别定义为:飞机(PL)、机场(APO)、棒球场(BF)、篮球场(BC)、桥梁(BR)、烟囱(CH)、高速公路服务区(ESA)、高速公路收费站(ETS)、大坝(DAM)、高尔夫球场(GF)、地面田径场(GTF)、港口(HA)、立交桥(OP)、船舶(SH)、体育场(STA)、储罐(ST)、网球场(TC)、火车站(TS)、车辆(VE)和风车(WM)。

HRSC2016(high-resolution ship collections)(Liu等,2017)数据集包含海上和近海的船舶实例,具有任意方向。训练集、验证集和测试集分别包括436幅、181幅和444幅图像。

本文遵循MMRotate(Zhou等,2022)框架中标准数据集预处理流程,将DOTA数据集中高分辨率图像裁切为1 024×1 024像素的图像,设置重叠区域200像素,并将所有裁切后图像的检测结果合并以评估性能。DIOR-R数据集使用800×800像素图像大小进行训练和测试。HRSC2016数据集中图像被缩放到800×800像素进行训练和测试。

2.2 实验设置

实验使用单个显存为24 GB的NVIDIA RTX 3090 GPU,总批次大小为2进行训练和测试。训练使用AdamW(Loshchilov和Hutter,2019)优化器,初始学习率为5×10^{-5},一阶矩估计的指数衰减率为0.9,二阶矩估计的指数衰减率为0.99。在HRSC2016数据集上使用配置"6× RR",共训练72个epoch,其他数据集上使用"1×",共训练12个epoch,默认采用随机翻转。

2.3 消融实验

为验证不同超参数对本文方法的影响,进行了一系列消融实验。本节所有实验基于RetinaNet检测器,骨干网络采用ResNet-50(resdual network)(He等,2016),在联合优化方法中使用KLD(Yang等,2021c)损失进行训练,并在HRSC2016和DIOR-R数据集上对精度进行了评估。

1)向量模长损失权重w_{cir}和角度损失权重w_{angle}的消融。如表1所示,对式(10)中向量模长损失和角度损失进行不同权重的实验。实验结果表明,w_{cir}和w_{angle}最佳的设置为0.05和0.10,此时平均检测精度均值(mean average precision, mAP)为59.94%。此外,单独使用向量模长损失或角度损失的mAP值分别为57.81%和58.97%,低于最优的结果,说明基于GDL的联合优化方法中依然存在边界不连续问题。表中,AP_{50}表示IoU阈值为0.50时计算得到的平均精度;AP_{75}表示IoU阈值为0.75时计算得到的平均精度。

表1 HRSC2016数据集上参数w_{cir}和w_{angle}消融实验

Table 1 Ablation experiments of parameters w_{cir} and w_{angle} on the HRSC2016 dataset

/%

$(w_{\text{cir}}, w_{\text{angle}})$	AP_{50}	AP_{75}	mAP
(0.05, 0.05)	86.20	70.60	58.39
(0.05, 0.10)	**87.30**	70.80	**59.94**
(0.05, 0.15)	86.50	70.40	58.52
(0.05, 0.20)	86.10	70.50	57.60
(0.10, 0.10)	86.70	**71.60**	59.40
(0.15, 0.10)	86.20	69.90	57.29
(-, 0.10)	86.90	71.50	57.81
(0.05, -)	86.90	71.00	58.97

注:加粗字体表示各列最优结果,"-"表示参数设置为0。

2)宽高比阈值的消融。如表2所示,针对长方形目标与类正方形目标在边界回归问题中出现频率差异的现象,通过式(7)设置宽高比阈值ε进行优化调节。当宽高比阈值设为1.1时获得最佳检测性能,并且随着阈值的增大检测性能减弱,表明长方形目标与类正方形目标的角度损失具有周期差异。

3)自适应权重系数的消融。表3表示对式(9)中不同的自适应权重系数超参数β进行的实验。当β较小时,自适应权重系数过早地影响模型训练,导致模型检测性能下降。随着β的增大,检测性能逐渐提高;随着β的继续增大,IoU对平衡系数s的影响减弱,s将趋近于1,在训练过程中其自适应权重调节功能基本失效,此时模型检测性能趋近于不采用平衡策略时的精度。

4)形状感知系数的消融。如表4所示,对式(7)

表 2　DIOR-R 数据集上宽高比阈值 ε 消融实验

Table 2　Ablation experiments of aspect ratio thresholds ε on the DIOR-R dataset

ε	AP_{50}/%	AP_{75}/%	mAP/%
0.00	56.90	35.30	34.95
1.00	56.50	35.10	34.94
1.10	**57.00**	35.50	**35.34**
1.20	55.80	**35.70**	35.28
1.30	55.40	34.70	34.61

注：加粗字体表示各列最优结果。

表 3　DIOR-R 数据集上超参数 β 消融实验

Table 3　Ablation experiments of hyperparameter β on the DIOR-R dataset

β	AP_{50}/%	AP_{75}/%	mAP/%
0	**57.00**	35.50	35.34
2	54.90	35.20	34.71
3	56.10	36.20	35.59
4	56.00	36.00	35.73
5	**57.00**	**36.50**	**36.01**
6	55.70	35.70	35.40
7	56.10	35.80	35.36

注：加粗字体表示各列最优结果。

中形状感知系数 λ 进行了实验对比。其中形状感知系数表示根据目标宽高比动态调整 λ 值，λ 取 1 和 2 分别对应长方形和类正方形这两种典型情况，λ 为 1.5 代表一种介于二者之间的情况。实验结果表明，形状感知系数方法优于固定 λ 设置，验证了长方形与类正方形目标在角度损失周期上存在差异，而固定 λ 值无法适应不同宽高比目标的损失周期变化需求。

表 4　DIOR-R 数据集上形状感知系数消融实验

Table 4　Ablation experiments of shape awareness coefficient on the DIOR-R dataset

/%

方法	AP_{50}	AP_{75}	mAP
基准方法	56.47	34.40	34.72
λ = 1.0	56.90	35.20	35.17
λ = 1.5	56.50	35.40	34.87
λ = 2.0	56.90	**35.50**	35.00
形状感知系数	**57.00**	**35.50**	**35.34**

注：加粗字体表示各列最优结果。

图 7 展示了不同宽高比类别检测精度的对比，相比于固定 λ 方法，形状感知系数方法在小宽高比目标（飞机、棒球场）上检测精度最优，同时在大宽高比目标（篮球场、网球场）上保持了较高的检测精度，说明形状感知系数方法能够有效兼顾不同宽高比目标的角度损失变化周期。引入 APBS 后小宽高比目标（飞机、棒球场）的检测精度进一步提升，说明 APBS 通过宽高比自适应加权，增强了 GDL 对此类目标角度变化的感知能力。

2.4　联合优化方法实验

在联合优化方法中，AVDM 和 APBS 作为 GDL 方法上的附加模块。选择两个代表性检测器 RetinaNet、FCOS 以及 3 个代表性 GDL 损失 GWD（Yang 等，2021b）、KLD（Yang 等，2021c）和 KFIoU（Yang 等，2023）。通过 FCOS 和 RetinaNet 两个检测器在数据集上总提升精度的平均值，来验证 AVDM 和 APBS 是否可以提高 GDL 回归方法的性能。在 DOTA、DIOR-R 和 HRSC2016 数据集上的实验结果分别如表 5—表 7 所示，为确保公平比较，在表中每一行使

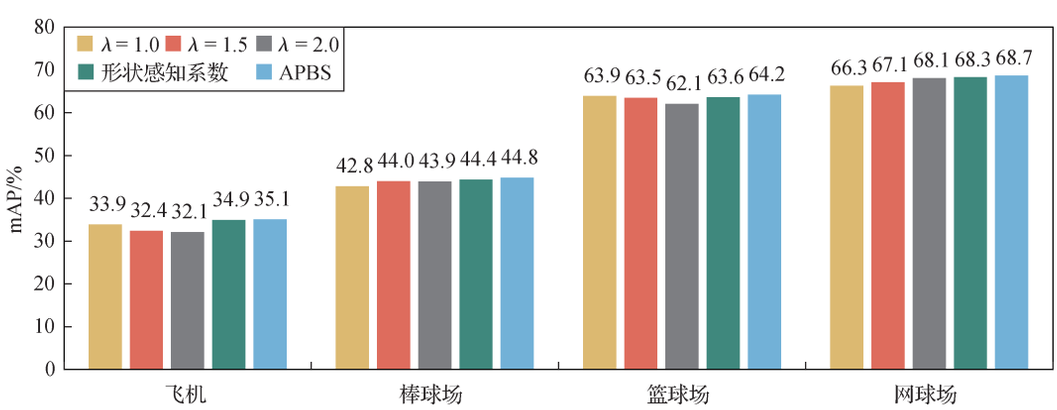


图 7　不同宽高比目标的检测精度比较

Fig. 7　Comparison of detection accuracy for targets with different aspect ratios

用相同的参数设置。

1)在DOTA数据集上的检测结果如表5所示，与基准方法相比，加入AVDM后AP_{50}、AP_{75}和mAP平均提升了0.52%、2.47%和1.03%，进一步引入APBS后AP_{50}平均降低了0.07%，AP_{75}和mAP再次平均提升了0.84%和0.43%。在GDL方法中加入AVDM带来了显著的性能提升，表明了处理边界不连续问题的必要性。相较于AVDM引入APBS后，精度指标AP_{75}和mAP平均提升0.84%和0.43%，并且与基准方法相比，两项指标平均提升3.31%和1.46，说明角度感知不平衡的问题依然存在，也表明了文中所提平衡策略的有效性。

2)在DIOR-R数据集上的检测结果如表6所示，AVDM有效提升了检测精度，3种精度指标AP_{50}、AP_{75}和mAP平均提升了0.82%、1.37%和0.95%，在此基础上引入APBS后，尽管精度指标AP_{50}轻微下降了0.05%，但更严格的定位指标AP_{75}和mAP分别提升了0.82%和0.36%，与基准方法相比，两项指标平均提升2.19%和1.31%。

3)HRSC2016数据集上的检测结果如表7所示，由于该数据集中的目标都为细长的船舶类，因此将形状感知系数λ始终设置为1。与基准方法相比，使用AVDM方法在精度指标AP_{50}、AP_{75}和mAP上平均提升了0.92%、5.63%和3.86%。在引入APBS后精度指标AP_{50}、AP_{75}和mAP再次平均提升了0.13%、0.44%和0.49%，与基准方法相比，AP_{75}和mAP平均提升6.07%和4.35。加入AVDM检测效果提升明显，表明细长目标的检测精度易受到边界不连续问题的影响，而角度感知不平衡问题对检测精度的影响相对较弱。

表5　不同GDL方法在DOTA数据集上应用AVDM和APBS的性能比较

Table 5　Performance comparison of different GDL methods with AVDM and APBS on DOTA dataset

/%

方法	损失函数	基准方法			AVDM(vs基准)			AVDM + APBS(vs AVDM)		
		AP_{50}	AP_{75}	mAP	AP_{50}	AP_{75}	mAP	AP_{50}	AP_{75}	mAP
RetinaNet	GWD	72.14	42.85	42.73	72.56(+0.42)	45.94(+3.09)	43.88(+1.15)	71.97(−0.59)	47.25(+1.31)	44.33(+0.45)
	KLD	72.68	42.61	42.41	72.80(+0.12)	44.65(+2.04)	43.37(+0.96)	72.13(−0.67)	45.61(+0.96)	43.45(+0.08)
	KFIoU	71.59	37.31	39.68	71.34(−0.25)	41.01(+3.70)	41.03(+1.35)	71.03(−0.31)	42.30(+1.29)	41.40(+0.37)
FCOS	GWD	71.99	44.42	43.32	72.51(+0.52)	45.21(+0.79)	43.91(+0.59)	73.05(+0.54)	44.55(−0.66)	44.11(+0.20)
	KLD	71.41	41.71	41.41	73.17(+1.76)	43.50(+1.79)	42.90(+1.49)	73.78(+0.61)	44.25(+0.75)	43.74(+0.84)
	KFIoU	72.23	36.08	40.06	72.77(+0.54)	39.48(+3.40)	40.68(+0.62)	72.77(+0.00)	40.90(+1.42)	41.33(+0.65)
平均值	GDL	72.01	40.83	41.60	72.53(+0.52)	43.30(+2.47)	42.63(+1.03)	72.46(−0.07)	44.14(+0.84)	43.06(+0.43)

注：括号内数值表示精度提升值(+)和降低值(−)。

表6　不同GDL方法在DIOR-R数据集上应用AVDM和APBS的性能比较

Table 6　Performance comparison of different GDL methods with AVDM and APBS on DIOR-R dataset

/%

方法	损失函数	基准方法			AVDM			AVDM + APBS		
		AP_{50}	AP_{75}	mAP	AP_{50}	AP_{75}	mAP	AP_{50}	AP_{75}	mAP
RetinaNet	GWD	56.48	35.49	35.61	56.20(−0.28)	36.40(+0.91)	36.05(+0.44)	56.10(−0.10)	37.10(+0.70)	36.50(+0.45)
	KLD	56.47	34.40	34.72	57.00(+0.53)	35.50(+1.10)	35.34(+0.62)	57.00(+0.00)	36.50(+1.00)	36.01(+0.67)
	KFIoU	58.74	31.13	32.86	59.60(+0.86)	33.50(+2.37)	34.09(+1.23)	59.60(+0.00)	33.90(+0.40)	34.25(+0.16)
FCOS	GWD	62.05	39.40	38.81	63.00(+0.95)	40.70(+1.30)	39.85(+1.04)	62.70(−0.30)	41.00(+0.30)	39.92(+0.07)
	KLD	61.42	38.15	37.67	63.40(+1.98)	39.10(+0.95)	38.62(+0.95)	63.20(−0.20)	40.00(+0.90)	38.87(+0.25)
	KFIoU	61.90	34.90	35.80	62.80(+0.90)	36.50(+1.60)	37.24(+1.44)	63.10(+0.30)	38.10(+1.60)	37.84(+0.60)
平均值	GDL	59.51	35.58	35.91	60.33(+0.82)	36.95(+1.37)	36.87(+0.95)	60.28(−0.05)	37.77(+0.82)	37.23(+0.36)

注：括号内数值表示精度提升值(+)和降低值(−)。

表 7　不同 GDL 方法在 HRSC2016 数据集上应用 AVDM 和 APBS 的性能比较

Table 7　Performance comparison of different GDL methods with AVDM and APBS on HRSC2016 dataset

/%

方法	损失函数	基准方法			AVDM			AVDM + APBS		
		AP_{50}	AP_{75}	mAP	AP_{50}	AP_{75}	mAP	AP_{50}	AP_{75}	mAP
RetinaNet	GWD	85.74	62.27	55.96	86.30(+0.56)	69.70(+7.43)	57.49(+1.53)	87.10(+0.80)	69.10(−0.60)	58.94(+1.45)
	KLD	86.39	62.90	56.14	87.30(+0.91)	70.80(+7.90)	59.94(+3.80)	87.40(+0.10)	71.60(+0.80)	59.59(−0.35)
	KFIoU	84.64	61.38	54.52	86.00(+1.36)	69.40(+8.02)	57.39(+2.87)	85.90(−0.10)	70.40(+1.00)	58.29(+0.90)
FCOS	GWD	88.68	74.63	61.41	89.90(+1.22)	78.30(+3.67)	67.71(+6.30)	89.90(+0.00)	78.50(+0.20)	67.59(−0.12)
	KLD	89.16	73.51	61.18	90.00(+0.84)	77.90(+4.39)	66.95(+5.77)	90.00(+0.00)	78.60(+0.70)	67.49(+0.54)
	KFIoU	89.20	75.10	62.33	89.80(+0.60)	77.50(+2.40)	65.21(+2.88)	89.80(+0.00)	78.00(+0.50)	65.74(+0.53)
平均值	GDL	87.30	68.30	58.59	88.22(+0.92)	73.93(+5.63)	62.45(+3.86)	88.35(+0.13)	74.37(+0.44)	62.94(+0.49)

注：括号内数值表示精度提升值(+)和降低值(−)。

4)如表 8 所示，基于 RetinaNet 检测器，在联合优化方法中使用 RotatedIoU(Zhou 等，2019)损失进行训练，在 DIOR-R 数据集上进行评估。与基准方法相比，使用 AVDM 方法后，AP_{50}、AP_{75}和 mAP 分别提升了 0.30%、0.80% 和 0.71%，说明基于 RotatedIoU 的损失方法同样受到边界不连续性问题的影响。在引入 APBS 后，模型性能提升有限，这主要是由于 RotatedIoU 损失直接以两个旋转边界框的交并比作为优化目标，而不是高斯分布距离，因此受角度感知不平衡问题的影响相对较小。

表 8　RotatedIoU 损失在 DIOR-R 数据集上的实验结果

Table 8　Experimental results of RotatedIoU loss on DIOR-R dataset

/%

方法	AP_{50}	AP_{75}	mAP
基准方法	56.30	35.70	35.75
AVDM	56.60(+0.30)	36.50(+0.80)	36.46(+0.71)
AVDM + APBS	56.40(−0.20)	36.40(−0.10)	36.48(+0.02)

注：括号内数值表示精度提升值(+)和降低值(−)。

2.5　与不同方法的性能比较

为进一步评估方法的有效性，基于 FCOS 检测器和 KLD 损失，将本文方法 JCPB 与当前主流方法进行了比较，包括 CSL(Yang 和 Yan，2020)、ABFL(angular boundary discontinuity free loss)(Zhao 和 Li，2024)、PSCD(Yu 和 Da，2023)、PSCD-KLD(Yu 和 Da，2024)、CTFC(composite trigonometric function coder)(Hu 等，2025)、Mask OBB(mask oriented bounding box)(Wang 等，2019)、CenterMap(center probability map)(Wang 等，2021)和 RsDet(Qian 等，2021)，在 DOTA 数据集上的评估结果如表 9 所示。以 ResNet50(R-50)为骨干网络，所提方法的 mAP 值为 73.78%，优于其他方法。并且，在小型车辆、船舶、网球场、储罐和环岛类别中精度最优。使用 Swin-Transformer(SwinT)为骨干网络，模型性能进一步提升，mAP 值为 74.71%。

将其中一些结果可视化，如图 8 所示。图中上侧为在 FCOS 中使用 KLD 的检测结果，下侧为使用本文方法 JCPB 的检测结果。从图中观察到，KLD 方法对小宽高比的目标角度感知不敏感，导致定位在方向上的偏差。本文方法对小宽高比目标的方向检测更加精确，如飞机和棒球场，并且保持了细长目标的检测精度，如船舶和港口。

在 RetinaNet 检测器中以 ResNet50 为骨干网络，输入图像大小为 1 024 × 1 024 像素的条件下，进一步对模型计算效率进行了分析。模型计算复杂度和参数量的评估结果如表 10 所示。其中，CSL(Yang 和 Yan，2020)为角度分类方法，ω 代表角度类别之间的间隔，当 $\omega = 4$ 时表示角度类别按每 4° 进行分布，PSCD(Yu 和 Da，2024)表示 PSC(Yu 和 Da，2024)的双频方式，采用 10 亿次浮点运算量(giga floating-point operations，GFLOPs)和参数量(parameters)作为指标来对比模型的时间和空间复杂度。结果表明，AVDM 仅增加了少量的计算量和参数量，且其增量低于其他方法。

表9　在DOTA数据集上与当前主流方法的精度对比

Table 9　Accuracy comparison with current state-of-the-art methods on DOTA dataset

方法	网络	AP_{50}/%							
		飞机	棒球场	桥梁	田径场	小型车辆	大型车辆	船舶	网球场
Mask OBB	R-50-FPN	–	–	–	–	–	–	–	–
PSCD-KLD	R-50-FPN	–	–	–	–	–	–	–	–
PSCD	R-50-FPN	88.04	73.95	48.83	63.44	80.01	80.75	87.58	90.88
RsDet	R-50-FPN	**89.30**	82.70	47.70	**63.90**	66.80	62.00	67.30	90.80
CSL	R-50-FPN	88.24	74.87	41.27	61.03	79.52	78.35	87.19	90.88
CenterMap	R-50-FPN	88.88	81.24	**53.15**	60.65	78.62	66.55	78.10	88.83
CTFC	R-50-FPN	88.52	78.24	50.68	60.97	80.32	80.40	87.57	**90.90**
ABFL	R-50-FPN	89.05	**83.09**	48.20	63.38	79.06	77.48	86.06	90.89
JCPB(本文)	R-50-FPN	88.77	78.37	49.65	61.24	80.24	79.79	**87.67**	**90.90**
JCPB(本文)	SwinT-FPN	88.65	81.90	50.48	62.15	**80.58**	**81.86**	**87.67**	90.88

方法	网络	AP_{50}/%							
		篮球场	储罐	足球场	环岛	港口	游泳池	直升机	mAP
Mask OBB	R-50-FPN	–	–	–	–	–	–	–	71.43
PSCD-KLD	R-50-FPN	–	–	–	–	–	–	–	72.32
PSCD	R-50-FPN	81.69	67.23	58.70	60.26	65.67	71.11	53.06	71.41
RsDet	R-50-FPN	85.30	82.40	**62.30**	62.40	65.70	68.60	**64.60**	70.80
CSL	R-50-FPN	81.50	84.53	54.70	62.65	62.84	68.45	46.50	70.83
CenterMap	R-50-FPN	77.80	83.61	49.36	66.19	**72.10**	72.36	58.70	71.74
CTFC	R-50-FPN	83.66	76.13	56.23	60.11	66.29	69.80	50.78	72.40
ABFL	R-50-FPN	82.24	84.60	56.66	66.13	64.75	69.82	54.33	73.01
JCPB(本文)	R-50-FPN	84.79	84.99	58.57	**67.97**	66.18	70.55	57.06	73.78
JCPB(本文)	SwinT-FPN	**86.80**	**86.65**	60.98	65.62	68.88	**73.64**	54.93	**74.71**

注：加粗字体表示各列最优结果。“–”表示检测结果原文未公布。



图8　在DOTA数据集上检测结果的比较

Fig. 8　Comparison of detection results on the DOTA dataset

表 10　计算复杂度和参数量的比较

Table 10　Comparison of GFLOPs and parameters

方法	GFLOPs	参数量/M
RetinaNet	215.92	36.42
+CSL(ω = 4)	236.29(+20.37)	37.35(+0.93)
+CSL(ω = 1)	297.41(+81.49)	40.15(+3.73)
+PSC	217.27(+1.35)	36.48(+0.06)
+PSCD	218.63(+2.71)	36.54(+0.12)
+AVDM	216.82(+0.90)	36.46(+0.04)

注：括号内数值表示相较基准方法的 GFLOPs 和参数量增加值。

3　结　论

本文提出一种遥感目标定向检测方法，针对角度边界不连续问题，提出将角度分解为连续向量表示，并引入形状感知系数，动态设置长方形与类正方形目标的角度损失周期。针对角度感知不平衡问题，提出角度感知平衡策略，增加模型对小宽高比目标角度信息的关注度。最后，通过联合角度连续表示和感知平衡，在保持低参数回归框架的前提下，同时解决角度边界不连续和感知不平衡两个问题。在 DOTA、DIOR 和 HRSC2016 数据集上的实验验证了所提方法的有效性。在 DOTA 数据集上的对比实验结果表明，本文方法检测精度优于其他先进方法。

本文方法在处理小宽高比目标时，其中自适应权重系数中的超参数β和形状感知系数中的宽高比阈值ε仍需手动调优，一定程度上限制了方法的泛化能力。为此，本文下一步工作将考虑设计形状自适应的高斯分布表示方法，通过度量该类分布之间的距离，指导模型平衡对不同宽高比目标的角度感知，避免依赖外部计算来分配不同目标的角度损失权重。同时，在该分布中嵌入连续角度表示机制，从而实现一个完全集成化且连续的定向目标检测框架。

参考文献(References)

Cheng G, Wang J B, Li K, Xie X X, Lang C B, Yao Y Q, et al. 2022. Anchor-free oriented proposal generator for object detection. IEEE Transactions on Geoscience and Remote Sensing, 60: #5625411 [DOI: 10.1109/TGRS.2022.3183022]

Ding J, Xue N, Long Y, Xia G S and Lu Q K. 2019. Learning RoI Transformer for oriented object detection in aerial images//Proceedings of 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach, USA: IEEE: 2844-2853 [DOI: 10.1109/CVPR.2019.00296]

He K M, Zhang X Y, Ren S Q and Sun J. 2016. Deep residual learning for image recognition//Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, USA: IEEE: 770-778 [DOI: 10.1109/CVPR.2016.90]

He Y, Tu B, Liu B, Li J and Plaza A. 2024. 3DSS-Mamba: 3D-spectral-spatial Mamba for hyperspectral image classification. IEEE Transactions on Geoscience and Remote Sensing, 62: #5534216 [DOI: 10.1109/TGRS.2024.3472091]

Hu J, Liang J W, Zhao M H, Du S L and Li P. 2025. Oriented object detection based on composite trigonometric function coder//Proceedings of 2025 IEEE International Conference on Image Processing. Anchorage, USA: IEEE: 277-282 [DOI: 10.1109/ICIP55913.2025.11084362]

Li Z H, Hou B, Wu Z T, Guo X P, Ren B, Ren Z L, et al. 2024. Rebalancing Gaussian location loss for high-precision detection on remote sensing images. IEEE Transactions on Geoscience and Remote Sensing, 62: #5645412 [DOI: 10.1109/TGRS.2024.3478364]

Lin T Y, Goyal P, Girshick R, He K M and Dollár P. 2017. Focal loss for dense object detection//Proceedings of 2017 IEEE International Conference on Computer Vision. Venice, Italy: IEEE: 2999-3007 [DOI: 10.1109/ICCV.2017.324]

Liu Z K, Yuan L, Weng L B and Yang Y P. 2017. A high resolution optical satellite image dataset for ship recognition and some new baselines//Proceedings of the 6th International Conference on Pattern Recognition Applications and Methods. Porto, Portugal: SciTePress, 1: 324-331 [DOI: 10.5220/0006120603240331]

Loshchilov I and Hutter F. 2019. Decoupled weight decay regularization//Proceedings of the 7th International Conference on Learning Representations. New Orleans, USA: OpenReview.net

Ma J Q, Shao W Y, Ye H, Wang L, Wang H, Zheng Y B, et al. 2018. Arbitrary-oriented scene text detection via rotation proposals. IEEE Transactions on Multimedia, 20(11): 3111-3122 [DOI: 10.1109/TMM.2018.2818020]

Qian W, Yang X, Peng S L, Yan J C and Guo Y. 2021. Learning modulated loss for rotated object detection//Proceedings of the 35th AAAI Conference on Artificial Intelligence. [s.l.]: AAAI: 2458-2466 [DOI: 10.1609/aaai.v35i3.16347]

Redmon J, Divvala S, Girshick R and Farhadi A. 2016. You only look once: unified, real-time object detection//Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, USA: IEEE: 779-788 [DOI: 10.1109/CVPR.2016.91]

Ren S Q, He K M, Girshick R and Sun J. 2017. Faster R-CNN: towards real-time object detection with region proposal networks. IEEE Transactions on Pattern Analysis and Machine Intelligence, 39(6): 1137-1149 [DOI: 10.1109/TPAMI.2016.2577031]

She H D and Zhao L J. 2024. Rotating target detection network that com-

bines key points and guide vectors. Journal of Image and Graphics, 29(2): 533-544 (余浩东, 赵良瑾. 2024. 结合关键点与引导向量的旋转目标检测网络. 中国图象图形学报, 29(2): 533-544) [DOI: 10.11834/jig.230207]

Tian Z, Shen C H, Chen H and He T. 2019. FCOS: fully convolutional one-stage object detection//Proceedings of 2019 IEEE/CVF International Conference on Computer Vision. Seoul, Korea (South): IEEE: 9626-9635 [DOI: 10.1109/ICCV.2019.00972]

Tu B, Ren Q, Li J, Cao Z L, Chen Y Y and Plaza A. 2024. NCGLF^2: network combining global and local features for fusion of multi-source remote sensing data. Information Fusion, 104: #102192 [DOI: 10.1016/j.inffus.2023.102192]

Wang J, Li F and Bi H X. 2022. Gaussian focal loss: learning distribution polarized angle prediction for rotated object detection in aerial images. IEEE Transactions on Geoscience and Remote Sensing, 60: #4707013 [DOI: 10.1109/TGRS.2022.3175520]

Wang J W, Ding J, Guo H W, Cheng W S, Pan T and Yang W. 2019. Mask OBB: a semantic attention-based mask oriented bounding box representation for multi-category object detection in aerial images. Remote Sensing, 11(24): #2930 [DOI: 10.3390/rs11242930]

Wang J W, Yang W, Li H C, Zhang H J and Xia G S. 2021. Learning center probability map for detecting objects in aerial images. IEEE Transactions on Geoscience and Remote Sensing, 59(5): 4307-4323 [DOI: 10.1109/TGRS.2020.3010051]

Xia G S, Bai X, Ding J, Zhu Z, Belongie S, Luo J B, et al. 2018. DOTA: a large-scale dataset for object detection in aerial images//Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, USA: IEEE: 3974-3983 [DOI: 10.1109/CVPR.2018.00418]

Xu D Q and Wu Y Q. 2024. Progress of research on deep learning algorithms for object detection in optical remote sensing images. National Remote Sensing Bulletin, 28(12): 3045-3073 (徐丹青, 吴一全. 2024. 光学遥感图像目标检测的深度学习算法研究进展. 遥感学报, 28(12): 3045-3073) [DOI: 10.11834/jrs.20243166]

Yang X, Hou L P, Zhou Y, Wang W T and Yan J C. 2021a. Dense label encoding for boundary discontinuity free rotation detection//Proceedings of 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Nashville, USA: IEEE: 15814-15824 [DOI: 10.1109/CVPR46437.2021.01556]

Yang X and Yan J C. 2020. Arbitrary-oriented object detection with circular smooth label//Proceedings of the 16th European Conference on Computer Vision. Glasgow, UK: Springer: 677-694 [DOI: 10.1007/978-3-030-58598-3_40]

Yang X, Yan J C, Ming Q, Wang W T, Zhang X P and Tian Q. 2021b. Rethinking rotated object detection with Gaussian Wasserstein distance loss//Proceedings of the 38th International Conference on Machine Learning. [s.l.]: PMLR: 11830-11841

Yang X, Yang X J, Yang J R, Ming Q, Wang W T, Tian Q, et al. 2021c. Learning high-precision bounding box for rotated object detection via Kullback-Leibler divergence//Proceedings of the 35th International Conference on Neural Information Processing Systems. [s.l.]: Curran Associates Inc., #1405

Yang X, Zhou Y, Zhang G F, Yang J R, Wang W T, Yan J C, et al. 2023. The KFIoU loss for rotated object detection//Proceedings of the 11th International Conference on Learning Representations. Kigali, Rwanda: OpenReview.net

Yu Y and Da F P. 2023. Phase-shifting coder: predicting accurate orientation in oriented object detection//Proceedings of 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Vancouver, Canada: IEEE: 13354-13363 [DOI: 10.1109/CVPR52729.2023.01283]

Yu Y and Da F P. 2024. On boundary discontinuity in angle regression based arbitrary oriented object detection. IEEE Transactions on Pattern Analysis and Machine Intelligence, 46(10): 6494-6508 [DOI: 10.1109/TPAMI.2024.3378777]

Zeng Y, Chen Y S, Yang X, Li Q Y and Yan J C. 2024. ARS-DETR: aspect ratio-sensitive detection transformer for aerial oriented object detection. IEEE Transactions on Geoscience and Remote Sensing, 62: #5610315 [DOI: 10.1109/TGRS.2024.3364713]

Zhang L, Chen W and Wang Y H. 2023. Key sub-region feature fusion network for fine-grained ship detection and recognition in remote sensing images. Journal of Image and Graphics, 28(9): 2940-2955 (张磊, 陈文, 王岳环. 2023. 用于遥感舰船细粒度检测与识别的关键子区域融合网络. 中国图象图形学报, 28(9): 2940-2955) [DOI: 10.11834/jig.220671]

Zhao Z F and Li S Y. 2024. ABFL: angular boundary discontinuity free loss for arbitrary oriented object detection in aerial images. IEEE Transactions on Geoscience and Remote Sensing, 62: #5611411 [DOI: 10.1109/TGRS.2024.3368630]

Zhou D F, Fang J, Song X B, Guan C Y, Yin J B and Dai Y C. 2019. IoU loss for 2D/3D object detection//Proceedings of 2019 International Conference on 3D Vision (3DV). Quebec City, Canada: IEEE: 85-94 [DOI: 10.1109/3DV.2019.00019]

Zhou Y, Yang X, Zhang G F, Wang J B, Liu Y Y, Hou L P, et al. 2022. MMRotate: a rotated object detection benchmark using PyTorch//Proceedings of the 30th ACM International Conference on Multimedia. Lisboa, Portugal: ACM: 7331-7334 [DOI: 10.1145/3503161.3548541]

作者简介

王朝阳,男,硕士研究生,主要研究方向为遥感目标检测。
E-mail: 2316085405023@stu.nchu.edu.cn
葛芸,通信作者,女,副教授,主要研究方向为人工智能、计算机视觉和遥感图像智能解译。E-mail: geyun@nchu.edu.cn
吴梦奇,女,硕士研究生,主要研究方向为遥感目标检测。
E-mail: 2316085405008@stu.nchu.edu.cn
陈金梁,男,硕士研究生,主要研究方向为遥感目标检测。
E-mail: 2420085405335@stu.nchu.edu.cn