

中图法分类号: TP18; TP391 文献标识码: A 文章编号: 1006-8961(2026)03-0912-15

论文引用格式: Yao R, Li Y M, Guo H F, Hu W T and Tian X R. 2026. Continuous geometric parameter representation for arbitrarily-oriented objects in remote sensing images. Journal of Image and Graphics, 31(3):0912-0926(姚睿, 李雨蔓, 郭浩帆, 胡文涛, 田祥瑞. 2026. 面向遥感图像任意方向目标的连续几何参数表示. 中国图象图形学报, 31(3):0912-0926)[DOI:10.11834/jig.250172]

面向遥感图像任意方向目标的连续几何参数表示

姚睿, 李雨蔓*, 郭浩帆, 胡文涛, 田祥瑞

南京航空航天大学自动化学院, 南京 210016

摘要: 目的 遥感图像中舰船、飞机等任意方向密集排布目标的检测是一项极具挑战的任务, 现有旋转目标检测器因定向边界框(oriented bounding box, OBB)表示不连续, 导致回归目标出现非连续跳跃, 从而限制了检测精度。为此, 提出一种定向边界框的连续几何表示(continuous geometric representation of oriented bounding boxes, CGR-OBB)。方法 CGR-OBB基于定向边界框的几何连续性约束, 构建包括位置参数、面积因子和比例参数的9维连续表示空间。具体而言, 基于严格的连续性指标, 首先通过目标角点相对于水平边界框(horizontal bounding box, HBB)边上中点的距离偏移量, 结合HBB的中心点、宽和高构成位置参数, 快速定位OBB; 其次, 利用OBB分割外接HBB的右上方面积计算面积因子, 以确定OBB的角点, 避免解码模糊; 最后, 引入OBB长宽比例参数约束其几何形态, 确保解码的OBB紧致包围目标。CGR-OBB进一步优化了舰船等高长宽比目标的参数定义和回归计算, 以提升检测稳定性。在实现层面, CGR-OBB被设计为编码—解码结构的插件模块, 可无缝集成至Faster R-CNN(region-based convolutional neural network)、RetinaNet等主流检测架构。**结果** 实验表明: 在HRSC2016(high-resolution ship collection 2016)数据集上, 使Rotated Faster R-CNN的mAP@75提升34.9%; 在DOTA(dataset for object detection in aerial images)数据集上, 使感兴趣区域(region of interest, RoI) Transformer和ReDet(rotation-equivariant detector)的mAP@75分别提升3.96%和4.31%, 展现了在高精度检测方面的显著优势, 与先进表示方法(DHRec(double horizontal rectangles)和COBB(continuous representation of oriented bounding boxes))相比在多个精度指标上改善效果更优。**结论** CGR-OBB能从表示层面直接解决OBB不连续问题, 有效减少了旋转目标检测器的漏检、错检情况, 具有高精度、易集成的特点, 展现较高的实用价值。

关键词: 任意方向目标检测; 连续表示; 定向边界框(OBB); 遥感图像; 深度学习

Continuous geometric parameter representation for arbitrarily-oriented objects in remote sensing images

Yao Rui, Li Yuman*, Guo Haofan, Hu Wentao, Tian Xiangrui

College of Automation Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China

Abstract: Objective Detecting arbitrarily oriented and densely arranged targets such as ships and aircraft in remote sensing images presents significant challenges. Traditional horizontal bounding boxes (HBBs) contain excessive background noise, substantially lowering the intersection over union (IoU) metric and compromising localization accuracy. Current rotated object detectors suffer from discontinuous representations of oriented bounding boxes (OBBs), causing abrupt

收稿日期: 2025-04-28; 修回日期: 2025-09-22; 预印本日期: 2025-09-29

* 通信作者: 李雨蔓 2303008@nuaa.edu.cn

基金项目: 航空科学基金项目(20240055052001)

Supported by: Aviation Science Foundation of China(20240055052001)

jumps in regression targets that limit detection precision. Current approaches to address OBB discontinuity typically follow three directions. First, improved loss functions can mitigate gradient conflicts in angle regression but ultimately fail to eliminate parameter jumps at periodic boundaries. Second, angle encoding methods effectively handle rotational discontinuity yet remain inadequate for aspect ratio variations. Third, and most fundamentally, novel OBB representations introduce continuous parameters through encode-decode modules, resolving the discontinuity issue at the representation level itself. However, current OBB representations have limitations, including incomplete continuity constraints, decoding ambiguity for special-shaped targets (square/extreme-aspect-ratio), and irregular parameter variations that restrict accuracy. To address these issues, this paper proposes a continuous geometric representation of OBBs (CGR-OBB) with rigorous continuity metrics to achieve high-precision detection. **Method** CGR-OBB constructs a nine-dimensional continuous representation space for OBBs on the basis of geometric continuity constraints, incorporating three key components: position parameters, area factors, and the aspect ratio parameter. The construction of these parameters strictly adheres to continuity metrics covering object rotation/aspect ratio continuity, loss rotation/aspect ratio continuity, and decoding completeness/robustness. First, the position parameters are formed by combining the distance offsets of target corners relative to the midpoints of HBB edges with the center point, width, and height of the HBB, enabling rapid OBB localization. Subsequently, the area factors are computed using the upper-right region areas obtained from the OBBs' partitioning of their circumscribed HBBs. These factors are then employed to determine the OBBs' vertex positions, effectively avoiding decoding ambiguity. Finally, the aspect ratio parameter of the OBB is applied to constrain its geometric shape, ensuring that the decoded OBB tightly encloses the target object. For high-aspect-ratio target detection, CGR-OBB enhances the detector's modeling capability by optimizing area factor definitions and performing logarithmic transformation on aspect ratio regression targets. CGR-OBB is implemented as a plug-and-play module with an encoder-decoder architecture. This modular component can be seamlessly integrated into mainstream detection frameworks such as Faster R-CNN and RetinaNet. During training, ground-truth bounding box annotations are transformed by the encoder module into a nine-dimensional continuous parameter representation. These parameters are then standardized according to regression parameter definitions and jointly computed with the detector's regression targets for loss calculation. During testing, the regression targets are decoded into bounding box parameters through the decoder module. The decoded bounding boxes are then quantitatively evaluated against ground-truth annotations, with simultaneous visualization of detection results on the input images.

Result Experimental results demonstrate CGR-OBB's effectiveness through comprehensive validation. First, continuity tests show that baseline detectors with CGR-OBB have reduced loss jumps and faster convergence, proving better training stability and efficiency. Second, unified benchmark tests on four detectors confirm accuracy improvements: on HRSC2016, mAP@75 increases by 34.9% for Rotated Faster R-CNN, 1.8% for Oriented R-CNN, 9.6% for RoI Transformer, and 1.0% for ReDet; on DOTA, improvements are 2.68% for Rotated Faster R-CNN, 0.66% for Oriented R-CNN, 3.96% for RoI Transformer, and 4.31% for ReDet. Third, compared with the latest other OBB representation methods (DHRec and COBB), CGR-OBB shows greater improvements across mAP@50, mAP@75, and mAP@50:95, averaging 2.08% over DHRec and 1.41% over COBB for RoI Transformer, and 1.69% over DHRec, and 1.79% over COBB for ReDet. Finally, in DOTA's mAP@50 tests against mmrotate's top detectors (including Gliding Vertex and R3Det + KFIOU), CGR-OBB achieves the best average accuracy while ranking first in eight out of 15 categories (containing extreme-aspect-ratio and square targets) and second in five others, demonstrating excellent multiscale target detection capability.

Conclusion The proposed CGR-OBB fundamentally addresses the OBB representation-level discontinuity, delivering significant improvements in detection accuracy. It constructs a continuous representation space composed of nine parameters, including position parameters, area factors, and aspect ratio parameters. These parameters vary continuously and predictably during regression, effectively eliminating sudden jumps in regression targets. The encode-decode structure ensures a strict reversible mapping between the OBB and its continuous representation, with the area factor resolving decoding ambiguities and the aspect ratio parameter enabling fine adjustment for high-precision detection. Its robust performance is particularly evident in challenging scenarios with dense target arrangements and complex backgrounds, where it effectively reduces missed detections and false alarms for rotated object detectors. Furthermore, the method combines high precision with easy integration capabilities, serving as a plug-and-play module that can be seamlessly incorporated into various main-

stream detection frameworks, demonstrating strong practical value for rotated object detection applications. Future work will focus on further enhancing performance under lower IoU thresholds to achieve comprehensive improvement across high and low precision metrics.

Key words: oriented object detection; continuous representation; oriented bounding box (OBB); remote sensing image; deep learning

0 引言

目标检测作为计算机视觉的基础任务(王晓燕等, 2025), 其定位精度直接影响图像解译的可靠性。在遥感图像中, 舰船、飞机等目标常以任意方向密集排布, 传统水平边界框(horizontal bounding box, HBB)包含大量背景噪声, 会显著降低交并比(intersection over union, IoU), 从而影响定位精度(石争浩等, 2023)。尽管主流定向边界框(oriented bounding box, OBB)表示方法(如旋转角参数化、顶点坐标编码)通过引入方向感知能力提升了检测适应性, 但其参数空间的不连续性仍构成根本性缺陷: 旋转角参数化在 0° 和 180° 临界状态会发生周期性跳变(Liu等, 2017a), 而顶点坐标编码因顶点顺序不确定性(如顺时针和逆时针互换)会导致回归目标的非连续跳跃(Ma等, 2018)。这种几何表示缺陷会引发损失函数振荡与梯度冲突, 破坏特征空间的一致性映射, 导致模型收敛不稳定及旋转鲁棒性下降。因此, 建立连续性的OBB表示对提升检测器对物体形态变化的泛化能力具有重要意义。

目前对OBB不连续问题的处理方法主要分为3类: 改进损失函数法、角度编码法和新OBB表示法。改进损失函数法通过设计平滑损失函数缓解角度回归在临界点的梯度冲突。例如, SCRDet(small, cluttered and rotated objects detector)(Yang等, 2019)在损失函数中引入平滑项, 以消除损失中的异常值; RSDet(rotation-sensitive detector)(Qian等, 2021)设计了可调制的损失函数以修正边界突变; GWD(Gaussian Wasserstein distance loss)(Yang等, 2021b)、KLD(Kullback-Leibler divergence loss)(Yang等, 2021c)与KFIoU(Kalman filter intersection over union)(Yang等, 2023)等方法将OBB转化为高斯分布后计算IoU损失, 但对方形物体的预测仍存在模糊性; ABFL(angular boundary discontinuity free loss)(Zhao和Li, 2024)通过将角度线性数据映射为圆形

数据计算损失, 进一步缓解了不连续性问题。然而, 这类方法难以完全消除周期边界处的参数跳变, 且复杂的损失函数设计可能增加优化难度。角度编码法通过离散化编码如CSL(circular smooth label)(Yang和Yan, 2020)和DCL(densely coded label)(Yang等, 2021a), 或连续编码如ACM(angle correct module)(Xu等, 2023)和PSC(phase-shifting coder)(Yu和Da, 2023)解决角度周期性导致的不连续问题。这类方法能够有效缓解目标旋转变化带来的方向角不连续, 但对长宽比变化引起的不连续性无用。新OBB表示法通过设计连续的回归参数并构造编码—解码模块, 从表示层面规避不连续问题。例如, RIE(rotated inscribed ellipse)(He等, 2021)和AEDet(arbitrary-oriented ellipse detector)(Zhou等, 2023)等方法利用向量分解表示OBB的边与角度, 实现了部分连续性, 但对方形物体的表示存在焦点消失问题; Gliding Vertex(Xu等, 2021)通过滑动OBB在最小外接HBB上的顶点实现长宽比变化的连续性, 但仍存在旋转不连续问题; DHRRec(double horizontal rectangles)(Nie和Huang, 2023)采用双水平矩形表示OBB实现了连续性编码, 但对高长宽比倾斜目标存在解码模糊问题, 参数较多且不能直接计算。COBB(continuous representation of oriented bounding boxes)(Xiao等, 2024)通过引入单一滑动比和IoU分数解码机制, 从理论上验证了方法的连续性。然而, 其IoU分数计算复杂, 且在回归过程中呈现不规则变化, 限制了方法的实用性。

因此, 为了从表示层面直接规避不连续问题, 应对现有方案的不足, 本文提出一种定向边界框的连续几何表示(continuous geometric representation of oriented bounding boxes, CGR-OBB)。首先, 定义9个满足连续性量化指标的参数共同表示OBB, 它们在回归过程中遵循定义规则变化, 可表示不同尺寸任意方向目标, 特别对高长宽比目标表示进行了优化。其次, 设计编码—解码模块实现OBB原表示与所提出表示之间的转换。编码模块基于目标形态

变化时OBB在最小外接HBB上的几何特性,将其映射为连续参数;解码模块通过解析几何约束实现参数的严格逆映射,确保OBB表示的唯一性,有效避免解码模糊问题。最后,设计回归目标与损失函数,并将编解码模块作为插件直接集成于现有检测器,在4个基线检测器和两个数据集上验证CGR-OBB能显著提升旋转目标检测的精度与鲁棒性。

1 CGR-OBB表示方法

1.1 连续性指标

OBB连续性量化指标已被严格定义(Xiao等, 2024)如下:

1) 目标旋转连续性

$$\forall x \in S, \lim_{\theta \rightarrow 0} \|f_E(x) - f_E(R(x, \theta))\| = 0 \quad (1)$$

2) 目标长宽比连续性

$$\forall x \in S, \lim_{r \rightarrow 1} \sum_{y \in A(x, r)} \|f_E(x) - f_E(y)\| = 0 \quad (2)$$

3) 损失旋转连续性

$$\forall x \in S, \lim_{\theta \rightarrow 0} \|L(f_E(x) - f_E(R(x, \theta)))\| = 0 \quad (3)$$

4) 损失长宽比连续性

$$\forall x \in S, \lim_{r \rightarrow 1} \sum_{y \in A(x, r)} \|L(f_E(x) - f_E(y))\| = 0 \quad (4)$$

5) 解码完整性

$$\forall x \in S, \exists t \in \mathbf{R}^n, IoU(x, f_D(t)) = 1 \quad (5)$$

6) 解码鲁棒性

$$\forall x \in S, \forall \varepsilon > 0, \exists \xi > 0, \forall \Delta d \in \mathbf{R}^n \wedge \|\Delta d\| < \xi, \\ 1 - IoU(x, f_D(f_E(x) + \Delta d)) < \varepsilon \quad (6)$$

式中, S 是OBB的集合, $\mathbf{R}^n$ 是表示参数的集合, f_E 是编码函数(将OBB映射为表示参数), f_D 是解码函数(将表示参数逆映射回OBB), L 是损失函数, $R(x, \theta)$ 表示做初始OBB x 旋转 θ 角度变换, $A(x, r)$ 表示做初始OBB x 改变其一边长至 r 倍变换。

1.2 连续性表示

CGR-OBB基于OBB与外接HBB之间的连续性几何变化构建表示参数,利用HBB长宽比的平滑性质(Xu等, 2021),确保了目标长宽比的连续性(即满足式(2)指标)。具体而言,CGR-OBB包含6个位置参数、2个面积因子和1个比例参数,如表1所示,本节将依次描述各参数定义,并阐述目标旋转连续性指标的实现机制。

表1 CGR-OBB的表示参数及其含义

Table 1 The representation parameters of CGR-OBB and their meanings

参数	变量
位置参数	$(x_c, y_c, w, h, \Delta\alpha , \Delta\beta)$
面积因子	(s_1, s_2)
比例参数	$ratio$

1.2.1 位置参数

位置参数由中心点 (x_c, y_c) , 外接HBB长 h 和宽 w , 以及距离偏移量 $(|\Delta\alpha|, |\Delta\beta|)$ 组成。如图1所示, $|\Delta\alpha|$ 和 $|\Delta\beta|$ 分别表示 $v_1(x_1, y_1)$ 角点相对于HBB上边中点的距离偏移量, 以及 $v_2(x_2, y_2)$ 角点相对于HBB右边中点的距离偏移量, 具体定义为

$$\begin{cases} |\Delta\alpha| = |x_1 - x_c| \\ |\Delta\beta| = |y_2 - y_c| \end{cases} \quad (7)$$

式中, $|\Delta\alpha|$ 和 $|\Delta\beta|$ 是无符号数。根据矩形框的对称性, $(w, h, |\Delta\alpha|, |\Delta\beta|)$ 随着目标旋转而连续改变, 满足旋转连续性要求(即满足式(1)指标)。

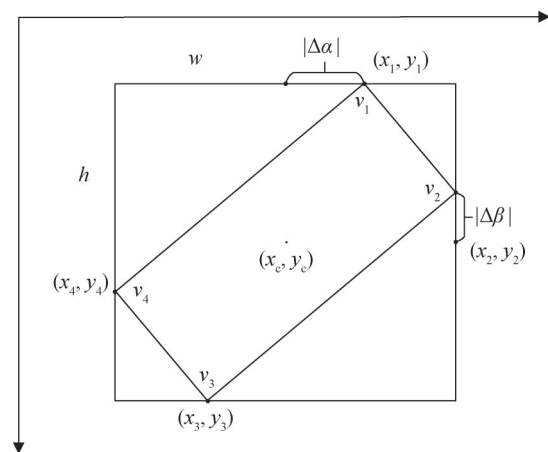


图1 CGR-OBB的位置参数表示图

Fig. 1 Illustration of the positional parameter representation of CGR-OBB

1.2.2 面积因子

一组位置参数最多可以表示4个形状不同的OBB, 如图2所示, 直接表示会造成解码混淆(Xiao等, 2024), 这也是大多数借助外水平框的OBB表示方法连续性设计需要解决的潜在问题。因此, 本节设计连续参数面积因子保证解码唯一性。

由图2可见, 外接HBB和OBB之间的面积差随

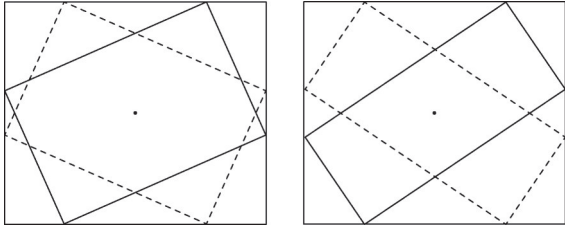


图2 给定位置参数后潜在的4个OBB表示图
Fig. 2 Illustration of potential OBB representations given the positional parameters

目标形态发生连续变化,若用OBB的4条边分割HBB,待确定的4个OBB分割出的4块面积(后文称分割面积)在同一方位各不相同。因此,对于任意给定的位置参数,只要计算出其中一个方位的分割面积,即可确定唯一的OBB表示;然而直接用分割面积判断会导致界定模糊。通过分割面积分类具体的OBB角点判断OBB的完整形状是最佳方案。因此,CGR-OBB将右上方的分割面积分为两部分,引入面积因子 s_1 和 s_2 ,其中 s_1 因子用于选择上角点, s_2 因子用于选择右上角点,以此推导唯一的OBB。

面积因子的定义规则借鉴DHRec中位置因子(Nie和Huang, 2023),但相较于位置因子具有以下优势:1)能够避免近水平目标和倾斜高长宽比目标的解码混淆;2)可以直接结合位置参数进行简单计算,降低计算复杂度;3)尤其擅长表示高长宽比目标。

如图3所示,首先连接中心点 (x_c, y_c) 与HBB右上顶点得到直线 CV_H , CV_H 交OBB于点A,再连接A点与HBB右上顶点得到线段 b 。线段 b 将分割面积分成两个部分(阴影填充区域)。以用于选择上顶点的 s_1 因子为例,如图3(a),引入两个三角形 T_{tr} 与 T_{tl} ,线段 b 作为三角形的底边,待选的上角点 $V_1(x_c + |\Delta\alpha|, y_c - 0.5h)$ 与 $V'_1(x_c - |\Delta\alpha|, y_c - 0.5h)$ 分别作为 T_{tr} 和 T_{tl} 的顶点,通过计算角点 V_1 和 V'_1 到直线 CV_H 的距离即可得到两个三角形的面积 S_{tr} 和 S_{tl} 。面积因子 s_1 定义为

$$s_1 = \frac{S_{tr} + \lambda \times S_{tl}}{S_{tr} + S_{tl} + 2\lambda \times S_{HBB}} \quad (8)$$

式中, S_{tr} 表示实际的上角点与底边 b 形成的面积,即图3(a)阴影区域面积, S_{tl} 表示外接HBB面积。 λ 是超参数,设置为接近0的值(如0.01),使用超参数是为了避免目标近水平时的困难检测(此时 S_{tr} 与 S_{tl} 约为0)。由式(8)可推断出,当上角点为 V_1 ,则 $s_1 <$

0.5;当上角点为 V'_1 ,则 $s_1 > 0.5$;当目标近水平,则 $s_1 = 0.5$ 。因此,面积因子 s_1 与阈值0.5的比较结果可以用于确定OBB的上角点。

同理,右上角点的确定过程如图3(b)所示,首先引入两个三角形 T_{tr} 与 T_{th} ,待选右角点分别表示为 $V'_2(x_c + 0.5w, y_c - |\Delta\beta|)$ 与 $V_2(x_c + 0.5w, y_c + |\Delta\beta|)$,同样以线段 b 为底,通过计算角点 V'_2 和 V_2 到直线 CV_H 的距离即可得到两个三角形的面积 S_{tr} 和 S_{th} 。面积因子 s_2 定义为

$$s_2 = \frac{S_{tr} + \lambda \times S_{th}}{S_{tr} + S_{th} + 2\lambda \times S_{HBB}} \quad (9)$$

式中, S_{tr} 表示实际的右角点与底边 b 形成的面积,即图3(b)阴影区域面积。由式(9)可推断出,当右角点为 V_2 ,则 $s_2 > 0.5$;当右角点为 V'_2 ,则 $s_2 < 0.5$;当目标近水平,则 $s_2 = 0.5$ 。因此,面积因子 s_2 与阈值0.5的比较结果可以用于确定OBB的右上角点。

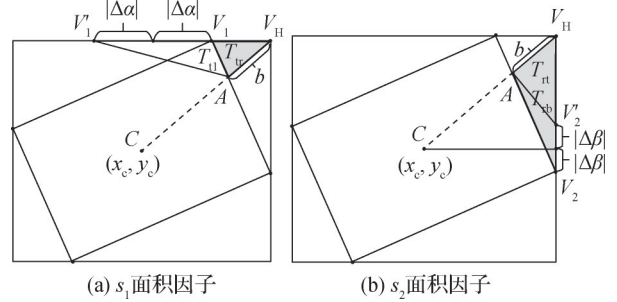


图3 面积因子定义说明图

Fig. 3 Illustration of area factor definition
(a) s_1 area factor; (b) s_2 area factor

当目标方向连续变化时,外接HBB和分割面积均为连续变量。基于连续函数的运算性质,由此导出的面积因子必然保持连续性(即满足式(1)指标)。值得注意的是,该面积因子表达不受目标形态限制,特别是在目标长宽比较大的情况下,候选角点间的面积比与阈值的比较结果将更为明晰,反而更有效地确定唯一的OBB表示。

1.2.3 比例参数

至此, $(x_c, y_c, w, h, |\Delta\alpha|, |\Delta\beta|, s_1, s_2)$ 8个参数解码出的OBB是平行四边形,直接用于预测目标会增加检测难度。针对这个问题,如图4(a)所示,CGR-OBB通过将较长对角线缩短到与较短对角线相同的长度,即将平行四边形转化为矩形,使其更紧密地包围目标,适应密集目标检测。

然而,从平行四边形到矩形的转换要求位置参

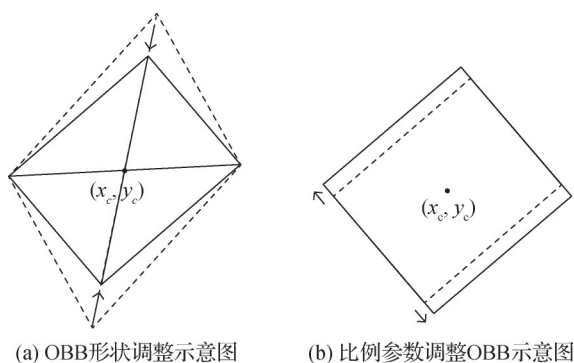


图4 OBB调整示意图

Fig. 4 Schematics of OBB adjustment ((a) schematic of OBB shape adjustment; (b) schematic of OBB scaling parameter adjustment)

数严格预测正确,否则将导致显著的回归参数误差,特别是当预测OBB与真实OBB尺寸相差较大时(Pu等,2024)。为了减轻对位置参数预测的压力,保证解码OBB准确定位,引入比例参数 $ratio$ 作为CGR-OBB的最后一个参数,其定义为

$$ratio = \begin{cases} \frac{h_o}{w_o} & h_o < w_o \\ \frac{w_o}{h_o} & w_o \leq h_o \end{cases} \quad (10)$$

式中, h_o 与 w_o 分别表示OBB的长度和宽度,当 $h_o < w_o$ 时,比例参数 $ratio$ 表示长度与宽度的比值,当 $w_o \leq h_o$ 时则表示宽度与长度的比值,方向改变不影响 $ratio$ 取值(即满足式(1)指标)。

1.3 在检测模型上的实现

CGR-OBB实例化为编码—解码模块后可作为插件直接应用到现有的基线模型。以Faster R-CNN (region-based convolutional neural network) (Ren等,2017)为例,如图5所示,首先在原模型的Box head回归分支的4个回归目标基础上添加其他5个回归目标;训练阶段,真值边界框标注通过编码模块与回归目标共同计算损失值;测试阶段,回归目标通过解码模块解算出边界框参数与真值边界框标注共同计算评估结果(VOC2012标准),同时在图像上可视化检测结果。

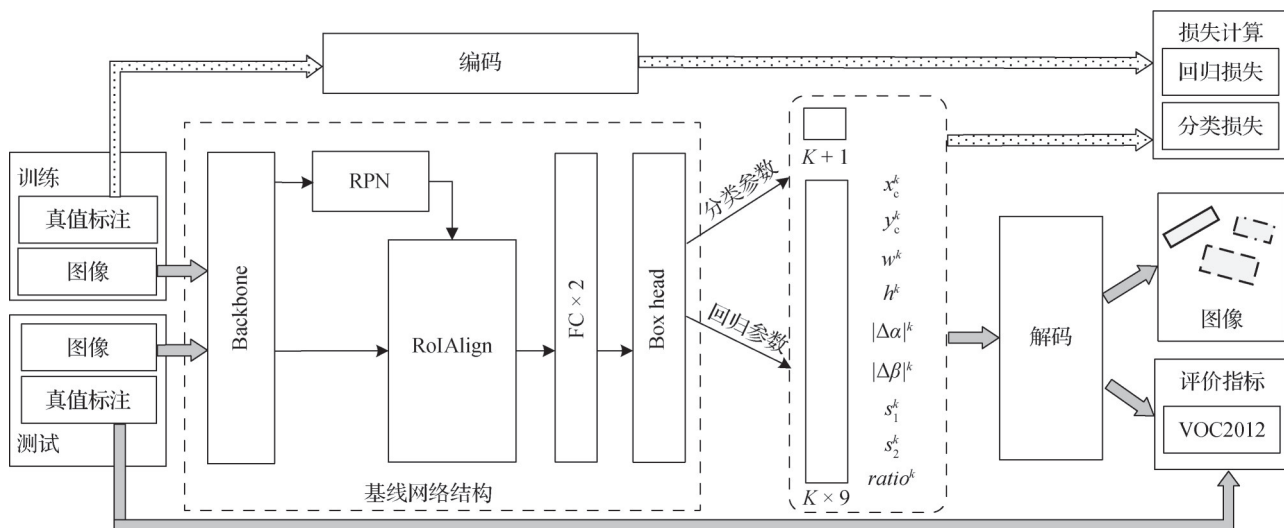


图5 基于Faster R-CNN的CGR-OBB实现

Fig. 5 The implementation of CGR-OBB based on Faster R-CNN

1.3.1 编码—解码模块

编码模块首先基于1.2节描述,利用HBB将真值边界框标注(4个角点坐标)编码成连续表示参数 $(x_c, y_c, w, h, |\Delta\alpha|, |\Delta\beta|, s_1, s_2, ratio)$ 。其次,对这些连续参数进行基于回归参数定义的标准转换,将其转化为回归形式以支持损失函数计算。

解码模块是CGR-OBB表示参数的严格逆映射的实现,即给定任意CGR-OBB表示,通过逆映射将

其还原为目标边界框参数(4个角点)。完整解码过程如下:首先,利用位置参数 $(x_c, y_c, w, h, |\Delta\alpha|, |\Delta\beta|)$ 得到待选的4个OBB,接着,使用面积因子 s_1 和 s_2 从其中选择唯一表示;随后,采用调整对角线的方式,将OBB由平行四边形转化成矩形;最后,利用比例参数 $ratio$ 精细调整矩形的短边以达到预测长宽比(如图4(b)所示),得到最终的唯一矩形OBB。整个解码过程中,表示参数有连续的微小变化,解码

的OBB也会有相应连续调整,因此CGR-OBB的解码是连续且完整的(即满足式(5)指标)。

为保证检测器的泛化能力,解码模块需要能抵抗输入的轻微改变。因此,对CGR-OBB从两个角度进行解码鲁棒性分析:1) s_1 和 s_2 数值轻微改变但与0.5阈值比较结果不变,即输入变化不影响OBB角点的选择。此时,因为其他的表示参数 $(x_c, y_c, w, h, |\Delta\alpha|, |\Delta\beta|, ratio)$ 是连续变化的,所以只会轻微改变OBB解码结果。2) 若 s_1 和 s_2 原值接近0.5阈值,轻微改变已经导致OBB角点的选择失误,因此时的目标近水平或OBB的角点靠近外接矩形HBB边上中点,则角点的改变对解码的OBB也只会产生轻微影响。总之,CGR-OBB解码表现出对抗扰的抵抗力,具备解码鲁棒性(即满足式(6)指标)。

1.3.2 回归目标与损失计算

大部分目标检测模型是计算真值标注与对应区域建议网络(regional proposal network, RPN)之间的偏差作为回归目标。水平框建议参数 (x_p, y_p, w_p, h_p) 回归与Faster R-CNN相同,具体为

$$\begin{cases} r_x = \frac{x_c - x_p}{w_p}, & r_y = \frac{y_c - y_p}{h_p} \\ r_w = \ln\left(\frac{w}{w_p}\right), & r_h = \ln\left(\frac{h}{h_p}\right) \end{cases} \quad (11)$$

式中, (r_x, r_y, r_w, r_h) 是计算得到的水平框偏差。

比例参数 $ratio$ 取值在 $(0, 1]$ 范围内,有两种回归方法,一种是直接回归,即 $r_{ratio_orig} = ratio$,另一种是利用对数函数进行域的转换,定义为

$$r_{ratio_ln} = 1 + \ln(ratio) \quad (12)$$

式中, r_{ratio_ln} 为使用对数转化回归方法的比例参数回归量,与 r_{ratio_orig} 相较鲁棒性更高,因为当目标为大长宽比形态时,若解码模块的输入出现轻微变化,直接回归方法将受到较大的影响,对数转化回归方法则比较稳健。

距离偏移量参数 $|\Delta\alpha|$ 和 $|\Delta\beta|$ 分别限制在 $[0, 0.5w]$ 与 $[0, 0.5h]$ 范围内,则这两个回归参数可以直接被标准化。具体为

$$r_{|\Delta\alpha|} = \frac{|\Delta\alpha|}{0.5w}, r_{|\Delta\beta|} = \frac{|\Delta\beta|}{0.5h} \quad (13)$$

式中, $r_{|\Delta\alpha|}$ 与 $r_{|\Delta\beta|}$ 是距离偏移量的回归目标,标准化后取值在 $[0, 1]$ 范围内。面积因子 s_1 和 s_2 本身取值在 $[0, 1]$ 范围内,可直接用于回归。定义损失函数为

$$\begin{aligned} L = & w_1 L_{cls} + w_2 L_{sig}(hbox_p, hbox_t) + \\ & w_3 L_{sig}((|\Delta\alpha|_p, |\Delta\beta|_p), (|\Delta\alpha|_t, |\Delta\beta|_t)) + \\ & w_4 L_{sig}(ratio_p, ratio_t) + w_5 L_{sig}(s_p, s_t) \end{aligned} \quad (14)$$

式中, L_{cls} 表示分类损失(与基线模型保持一致), L_{sig} 表示回归损失,采用Smooth L_1 损失(Ren等, 2017)。 $hbox_p$ 与 $hbox_t$ 分别表示水平边界框回归参数的预测值和真值标注。总损失等于分类损失与回归损失之和, w_1, w_2, w_3, w_4, w_5 分别表示各损失的权重,是预定义的超参数。参数回归量的连续性保证了损失量的连续性(即满足式(3)和式(4)指标)。

2 实验结果与分析

2.1 实验设置

2.1.1 数据集

1) DOTA (dataset for object detection in aerial images) (Xia等, 2018)数据集。是一个广泛使用的遥感图像目标检测基准,包含2 806幅高分辨率图像,图像尺寸范围从 $800 \times 800 \sim 4\,000 \times 4\,000$ 像素。数据集中共标注了188 282个目标实例,涵盖15个类别:飞机(PL)、棒球场(BD)、桥梁(BR)、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(SBT)、环形交叉路口(RA)、海港(HA)、游泳池(SP)和直升机(HC)。这些目标具有显著的尺度变化、方向多样性以及密集分布的特点,同时背景涵盖了城市、农田和港口等多种复杂场景。

2) HRSC2016 (high-resolution ship collection 2016) (Liu等, 2017b)数据集。是一个专注于船舶检测的高分辨率遥感图像数据集,图像分辨率从 $300 \times 300 \sim 1\,500 \times 900$ 像素不等。图像主要来源于港口和近海区域,目标具有明显的方向性和尺度变化,能够有效验证模型在复杂背景下的目标定位和方向预测能力。

2.1.2 实施细节

1) 数据集处理。对于DOTA数据集,根据标准协议,将所有图像裁剪为 $1\,024 \times 1\,024$ 像素图像块,相邻图像块之间重叠200像素;对于HRSC2016数据集,在不改变图像纵横比情况下将图像大小调整为 800×512 像素。对两个数据集数据增强包括随机翻转,采用水平翻转、垂直翻转和对角线翻转,翻转

概率均为 25%。实验使用训练集和验证集进行训练,其余的用于测试。

2)实验基准。本文使用基于 PyTorch 的开源旋转目标检测工具箱 mmrotate 作为统一的实验基准,它集成了多种先进的旋转目标检测模型,提供标准的评估工具,确保实验结果的可比性和可靠性。实验保留模型与 OBB 表示无关的超参数设置,训练时设置相同的随机种子。

3)其他设置。实验采用 Linux 操作系统、1 块 Tesla V100 GPU 训练,使用随机梯度下降优化,批处理大小为 2,每个 GPU (graphics processing unit) 工作 4 线程。在 DOTA 数据集上训练 12 个 epoch, HRSC2016 数据集上训练 36 个 epoch。

2.1.3 评价指标

使用平均精度均值 (mean average precision, mAP) 作为模型的性能评估指标。mAP 通过计算不同类别在不同 IoU 阈值下的平均精度 (average precision, AP),并对所有类别的 AP 取均值得到。例如, mAP@50 和 mAP@75 分别表示在 IoU 阈值为 0.5 与 0.75 时的平均精度均值, mAP@50:95 表示在 IoU 阈值为 0.5~0.95 (步长为 0.05) 范围内,计算每个阈值下的 AP 并取均值。mAP 是目标检测领域广泛使用的精度衡量标准,能够综合反映模型在检测准确性和召回率上的表现。

2.2 连续性验证与消融实验

2.2.1 连续性验证

本文从理论验证和应用验证两个层面开展连续性验证。理论验证旨在验证在目标长宽比与旋转角度连续变化时,表示参数的变化是否满足平滑性约束;应用验证旨在验证 CGR-OBB 编码—解码模块在训练过程中的应用连续性,通过分析损失函数曲线的收敛性与平滑性变化,评估其对训练稳定性的提升效果。

若 OBB 表示能够实现连续性表示,则在目标长宽比与旋转角度连续变化时能随之连续改变。其中,鉴于 $(x_c, y_c, w, h, ratio)$ 变化函数显见连续,这里重点验证距离偏移量参数 $|\Delta\alpha|$ 和 $|\Delta\beta|$ 与面积因子 s_1 和 s_2 随目标长宽比和旋转角度改变的情况。图 6 展示了距离偏移量的回归值与面积因子的回归值分别随角度与长宽比的变化。由图 6(a)(c) 可见,当角度在 $(-90^\circ, 90^\circ]$ 间以固定步长改变时, $|\Delta\alpha|$ 和 $|\Delta\beta|$ 与 s_1 和 s_2 随着角度的改变连续变化;由图 6(b)(d) 可见,

当长宽比在 $(0, 1]$ 间以固定步长改变时, $|\Delta\alpha|$ 和 $|\Delta\beta|$ 与 s_1 和 s_2 随着长宽比的改变连续变化。特别是在角度或长宽比变化的边界,这些表示参数也保持平滑变化。这一结果表明, CGR-OBB 的表示参数满足平滑性约束,体现了表示方法连续性。

若 OBB 表示能够实现连续性表示,则将其应用于检测器后,能够显著提升训练过程中的损失函数的平滑性。本节以 RoI Transformer (region of interest Transformer) (Ding 等, 2019) 作为基线检测器,对比分析了不同 OBB 表示方法对训练损失变化的影响,如图 7 所示。由图 7 可见,基线检测器与 CGR-OBB 结合后,损失曲线中的突变现象明显减少,且训练损失的收敛速度显著加快。这一结果表明, CGR-OBB 的连续性表示有效降低了损失函数的波动性,从而提升了模型训练的稳定性和效率。

2.2.2 消融实验

为了验证比例参数 *ratio* 的作用,在 RoI Transformer 上对比了 CGR-OBB 使用 *ratio* 参数前后的检测精度,如表 2 所示。结果表明,比例参数 *ratio* 通过约束 OBB 几何形态和对解码结果精细调整,对检测精度提升有明显贡献。

为了验证相对于直接回归,参数 *ratio* 用对数函数转化后回归更具鲁棒性,采用两种实验方法。

实验 1:通过下式量化了输入的轻微干扰对不同回归方式的影响。具体为

$$\Delta OBB = (1 - IoU[f_D(r), f_D(r + \varepsilon)])\varepsilon^{-1} \quad (15)$$

式中, r 是 r_{ratio_orig} 或 r_{ratio_ln} , $f_D(r)$ 是对近水平目标的解码结果, ε 是一个接近 0 的值 (设置为 0.01),用于模拟轻微干扰。结果表明, r_{ratio_ln} 面对输入轻微干扰更稳定,且不容易造成精度损失。轻微干扰对 r_{ratio_orig} 与 r_{ratio_ln} 影响的量化结果如图 8 所示。

实验 2:在 RoI Transformer 上对比了 CGR-OBB 的 *ratio* 参数使用不同回归方式的检测精度,如表 2 所示。

2.3 CGR-OBB 改善效果实验

为了验证所提出的 CGR-OBB 模块能够提升检测器对任意方向目标检测的检测精度,本文将其与 Rotated Faster R-CNN (Ren 等, 2017)、RoI Transformer、ReDet (rotation-equivariant detector) (Han 等, 2021) 和 Oriented R-CNN (Xie 等, 2021) 4 种先进的旋转目标检测器相结合,分别表示为 Rotated Faster R-

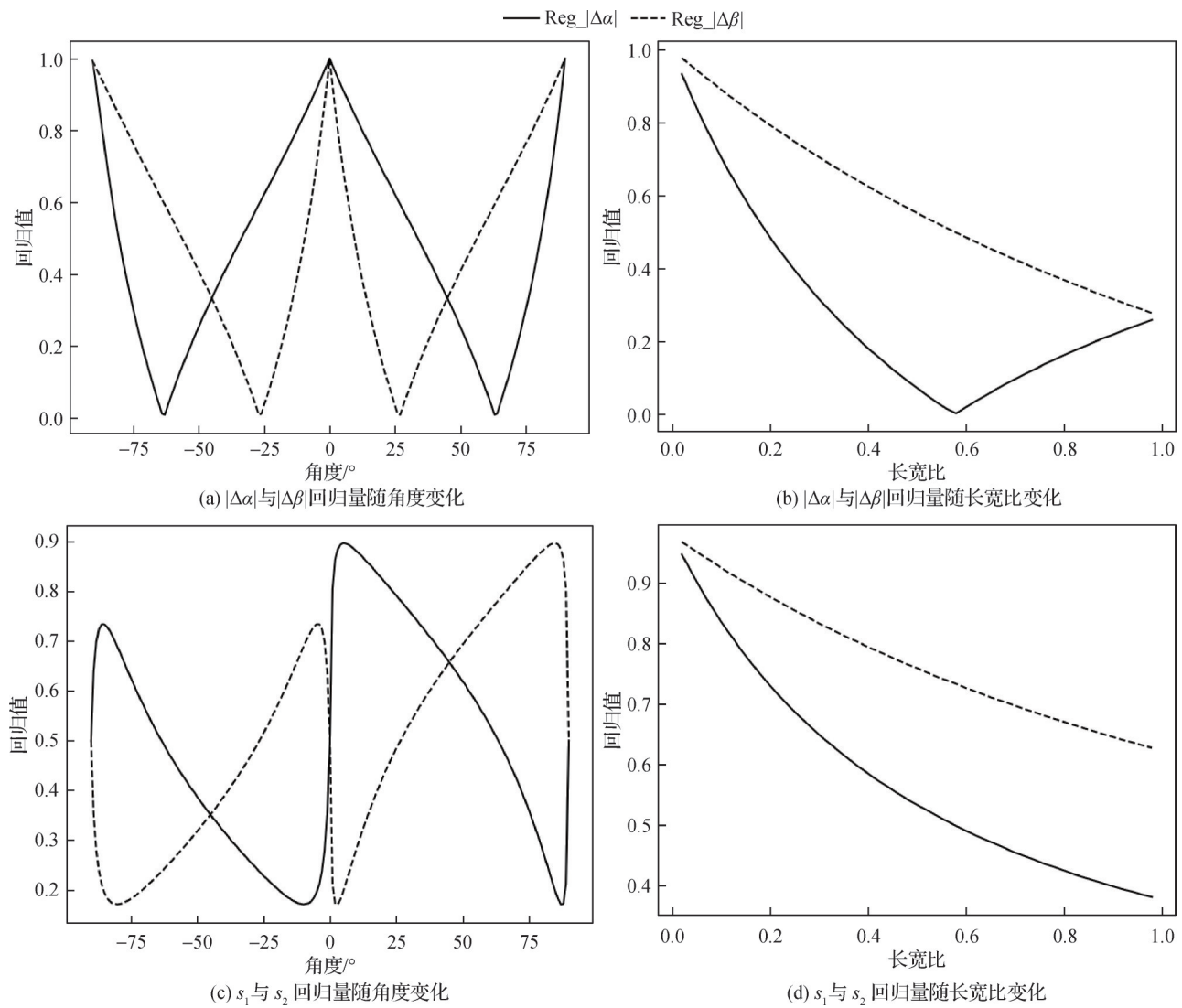


图6 距离偏移量的回归值与面积因子的回归值分别随角度与长宽比变化图

Fig. 6 Variation of regression values for distance offsets and area factors with angle and aspect ratio respectively
((a) distance offsets with angle; (b) distance offsets with aspect ratio; (c) area factors with angle;
(d) area factors with aspect ratio)

CNN+ *、Oriented-RCNN+ *、RoI Transformer+ * 和 ReDet+*。由于CGR-OBB基于水平建议框编码,因此分别应用于Rotated Faster R-CNN的box head、RoI Transformer 和 ReDet 的第一个 box head,以及 Oriented R-CNN 的 RPN head。本文对4种检测器在两个数据集上进行精度测试,实验定量结果如表3所示。

对于HRSC2016数据集,各检测器与CGR-OBB模块结合后在mAP@50和mAP@75指标上的精度均有不同程度的提升,且对于高精度指标mAP@75的提升效果更好。具体而言,Rotated Faster R-CNN分别提高了6.5%和34.9%,Oriented R-CNN分别提高

了0.6%和1.8%,RoI Transformer分别提高了2.4%和9.6%,ReDet分别提高了0.9%和1.0%。表明CGR-OBB对长宽比较大的船舶目标检测能力更强。

对于更具挑战性的DOTA数据集,CGR-OBB在mAP@50、mAP@75以及mAP@50:95这3个精度指标上的表现都优于基线检测器。具体而言,Rotated Faster R-CNN分别提高了0.51%、2.68%和1.35%,Oriented R-CNN分别提高了0.63%、0.66%和0.74%,RoI Transformer分别提高了1.42%、3.96%和2.27%,ReDet分别提高了0.35%、4.31%和2%。由此可见,CGR-OBB模块能够显著提升检测器的精度,尤其是在高精度检测任务中具有很大的优势。

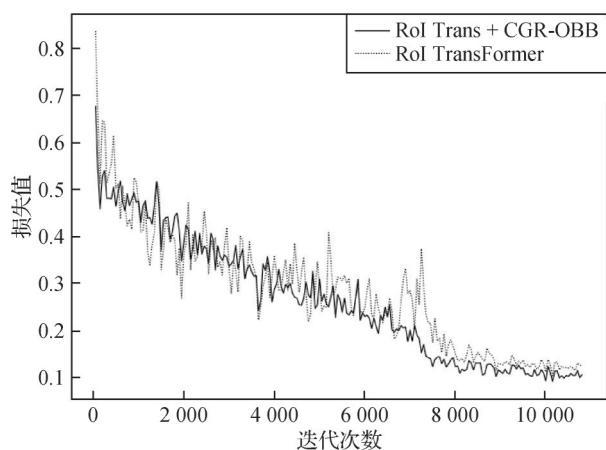


图7 CGR-OBB对训练损失收敛的平滑效果图

Fig. 7 Smoothing effect of CGR-OBB on training loss convergence

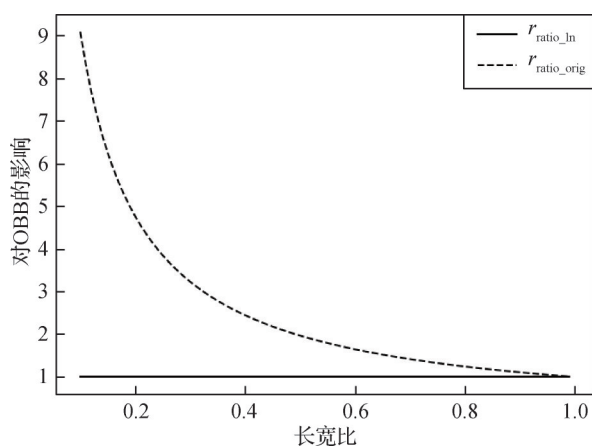
图8 轻微干扰对 r_{ratio_orig} 与 r_{ratio_ln} 影响的量化结果图Fig. 8 Quantitative results diagram of the impact of minor disturbances on r_{ratio_orig} and r_{ratio_ln}

表2 消融实验结果

Table 2 Ablation experiment results

表示方法	mAP@50	mAP@75	mAP@50:95
CGR-OBB*	75.84	49.36	46.32
CGR-OBB [^]	75.92	50.42	46.46
CGR-OBB	77.22	51.37	47.89

注: CGR-OBB*表示无ratio参数, CGR-OBB[^]表示ratio参数使用直接回归方式, CGR-OBB表示ratio参数使用对数转化回归方式。

为了直观地展现 CGR-OBB 模块对目标检测器的改善效果, 在 HRSC2016 和 DOTA 两个数据集中可视化了 CGR-OBB 对于各基线检测器的改善结果(仅显示置信度大于 0.3 的预测), 如图 9 和图 10

所示。

如图 9 所示, 通过 4 组对比图像(每列一组, 包含基线检测器及其与 CGR-OBB 结合后的检测结果)展示了 HRSC2016 数据集中舰船目标检测效果。由图 9(a)(e)可见, Rotated Faster R-CNN 对舰船目标的检测存在明显的方向预测偏差, CGR-OBB 的引入显著提升了方向检测准确度。由图 9(b)(d)可见, Oriented R-CNN 在复杂背景下以及 ReDet 对小目标都出现了不同程度的漏检(虚线框标出), 而图 9(f)(h)展现了 CGR-OBB 通过其连续表示优化了检测器的特征学习与回归过程, 有效减少了复杂背景和小目标的漏检情况。由图 9(c)(g)可见, RoI Transformer 在密集舰船检测中存在错检和漏检问题, 而 CGR-OBB 因为能正确表示目标方向特征且紧密包围目标, 可以更有效地检测这类场景目标。

图 10 展示了 DOTA 数据集中密集排布的大型车辆检测结果, 包含一幅原图和 4 列对照局部检测图(已从左至右分别在原图上框出相应位置)。结果显示, 各基线检测器都存在不同程度的检测偏差, 而 CGR-OBB 显著提高了基线检测器的准确性。

实验结果表明, 该方法能在多个场景任务中有效改善检测效果, 尤其在传统方法检测困难的任务(如密集排布、复杂背景)中表现出突出的能力。

2.4 与最先进方法的比较

基于 DOTA 数据集, 在 RoI Transformer 和 ReDet 检测器上对比了 CGR-OBB 与当前最先进的 DHRec 和 COBB 两种 OBB 表示方法的性能差异。

表 4 实验数据显示, CGR-OBB 在各项指标上均取得显著提升。对于 RoI Transformer, 其 mAP@50、mAP@75 和 mAP@50:95 分别较 DHRec 提高 0.99%、3.52% 和 1.73%, 较 COBB 提高 0.48%、2.09% 和 1.65%; 对于 ReDet, 其 mAP@75 和 mAP@50:95 分别较 DHRec 提升 4.23% 和 1.35%, 其 mAP@50、mAP@75 和 mAP@50:95 指标较 COBB 分别提升 0.21%、3.76% 和 1.41%。这些结果验证了 CGR-OBB 的双重优势: 一方面通过避免解码模糊和提升解码鲁棒性使 RoI Transformer 的 3 个精度指标均最优; 另一方面通过为各参数定义更为规则化的参数演化路径, 使其在高精度指标上, 对于具备对称性先验特征的等变检测器(如 ReDet)实现了更优提升效果。

同样, 基于 DOTA 数据集, 对比了 CGR-OBB 与 RoI Transformer 结合后与 mmrotate 框架下其他先进

表3 CGR-OBb模块在HRSC2016和DOTA数据集上的定量评估
Table 3 Quantitative evaluation of the CGR-OBb module on the HRSC2016 and DOTA datasets

检测器	骨干网络	HRSC2016		DOTA		
		mAP@50	mAP@75	mAP@50	mAP@75	mAP@50:95
Rotated Faster R-CNN	ResNet-50	85.60	35.20	73.41	39.62	40.76
Rotated Faster R-CNN+*		92.10	70.10	73.92	42.30	42.11
Oriented R-CNN	ResNet-50	96.50	87.10	75.54	47.47	45.22
Oriented R-CNN+*		97.10	88.90	76.17	48.13	45.96
RoI Transformer	ResNet-50	94.80	80.20	75.80	47.41	45.62
RoI Transformer+*		97.20	89.80	77.22	51.37	47.89
ReDet	ResNet-50	97.40	92.60	76.41	47.66	46.02
ReDet+*		98.30	93.60	76.76	51.97	48.02

注:加粗字体表示各组检测器中最优结果。

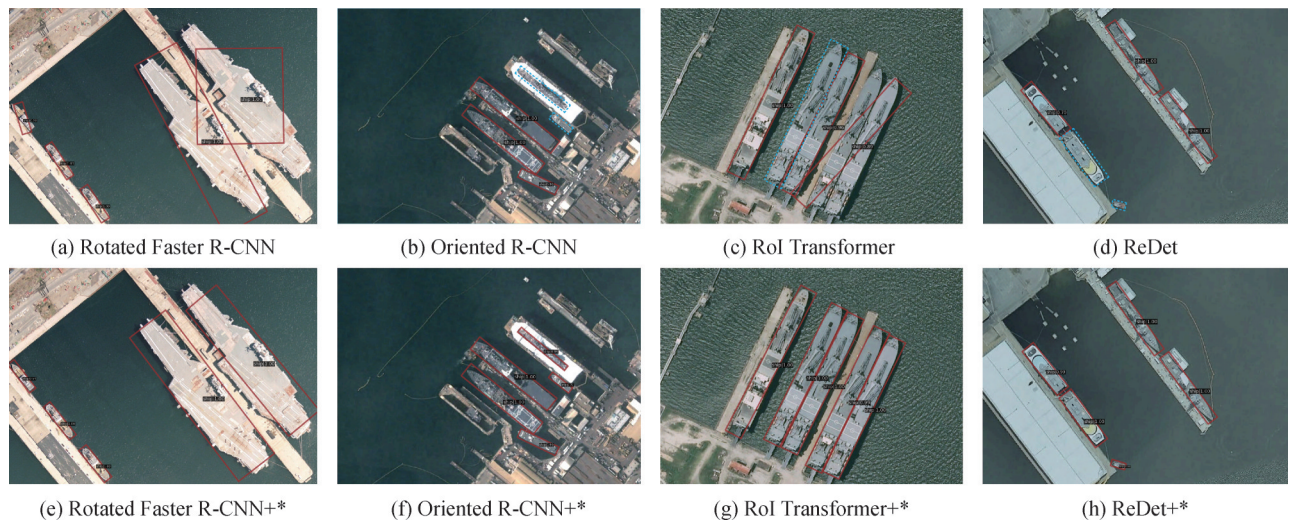


图9 基线检测器与CGR-OBb结合前后在HRSC2016数据集上可视化结果

Fig. 9 Visualization results on the HRSC2016 dataset before and after integrating the baseline detector with CGR-OBb
(a) Rotated Faster R-CNN; (b) Oriented R-CNN; (c) RoI Transformer; (d) ReDet; (e) Rotated Faster R-CNN+*;
(f) Oriented R-CNN+*; (g) RoI Transformer+*; (h) ReDet+*

旋转目标检测模型的mAP@50平均精度及15个类别的检测精度,详见表5。

由表5可见,CGR-OBb在平均精度上表现最优,且对不同尺寸目标具有稳健的检测性能(8类最优、5类次优),其中对于高长宽比目标如BR(桥梁)、LV(大型车辆)、SH(船舶)、HA(海港)检测精度很高。这一优势源于CGR-OBb的连续表示特性和针对性的参数优化设计:传统方法由于固有的角度不连续性,对高长宽比目标的微小角度偏差极为敏感,导致

IoU急剧下降;而CGR-OBb一方面通过参数连续化,从表示层面提升了检测器对目标特征的建模能力,另一方面其几何参数设计(如面积因子和比例参数回归目标定义)专门针对高长宽比目标特征进行了优化,因而对此类目标的检测性能优异。

各类别检测可视化结果如图11所示,结果表明,CGR-OBb表示方法不仅能有效检测多尺寸目标,而且在密集排布和高长宽比等更具挑战性的任务中表现出色,与现有最先进方法相比具有较强竞争力。

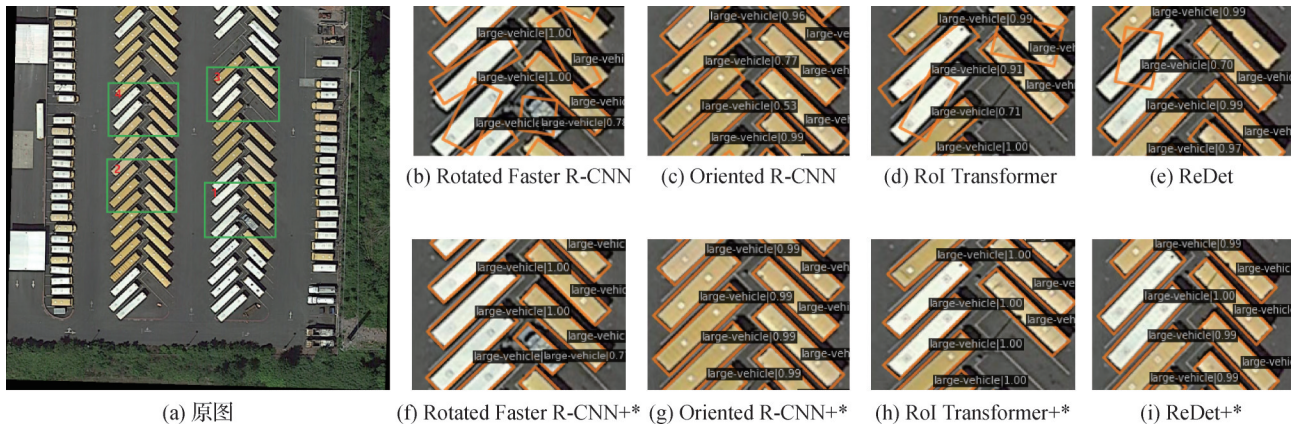


图 10 基线检测器与 CGR-OBb 结合前后在 DOTA 数据集上可视化结果

Fig. 10 Visualization results on the DOTA dataset before and after integrating the baseline detector with CGR-OBb
((a) original image; (b) Rotated Faster R-CNN; (c) Oriented R-CNN; (d) RoI Transformer; (e) ReDet; (f) Rotated Faster R-CNN+; (g) Oriented R-CNN+; (h) RoI Transformer+; (i) ReDet+)

表 4 DOTA 数据集上 CGR-OBb、DHRec 和 COBB 三种方法对基线检测器的精度改善情况对比
Table 4 Comparison of accuracy improvements by CGR-OBb, DHRec, and COBB methods over baseline detectors on the DOTA dataset

		/%		
基线检测器	表示方法	mAP@50	mAP@75	mAP@50:95
RoI Transformer	DHRec	+0.43	+0.44	+0.54
	COBB	+0.94	+1.87	+0.62
	CGR-OBb	+1.42	+3.96	+2.27
ReDet	DHRec	+0.84	+0.08	+0.65
	COBB	+0.14	+0.55	+0.59
	CGR-OBb	+0.35	+4.31	+2.00

注:加粗字体表示不同基线检测器的各精度指标的最优结果。

表 5 DOTA 数据集上 mmrotate 基准中旋转目标检测模型的 mAP@50 对比
Table 5 Comparison of mAP@50 for rotated object detection models on the mmrotate benchmark using DOTA dataset

检测器	PL	BD	BR	GTF	SV	LV	SH	TC	BC	ST	SBF	RA	HA	SP	HC	mAP@50
RetinaNet+GWD	88.93	77.04	45.91	69.30	72.53	64.06	76.35	90.88	79.21	80.46	57.68	64.37	63.60	64.75	48.27	69.56
FCOS+CSL	88.67	72.99	47.78	61.84	79.51	77.79	86.93	90.90	81.87	84.78	56.35	62.79	63.32	71.65	49.37	71.77
ATSS	88.50	77.74	49.60	69.87	76.87	72.52	82.49	90.80	80.31	82.97	62.35	64.67	64.82	66.82	53.98	72.29
R3Det + KFIoU	89.05	75.17	49.04	69.67	78.06	75.46	86.69	90.90	83.65	84.48	62.21	62.87	66.72	65.95	50.20	72.68
Gliding Vertex	88.93	81.09	50.97	67.67	77.74	75.91	86.54	90.89	79.77	85.31	55.48	64.62	66.43	68.92	52.23	72.83
Rotated Faster R-CNN	89.26	82.44	50.06	69.35	78.17	73.59	85.92	90.86	84.09	85.49	57.67	60.97	66.25	69.23	57.74	73.41
S2A-Net	89.25	81.20	51.56	71.40	78.62	77.38	86.77	90.90	86.28	84.64	61.22	65.65	66.07	67.58	50.19	73.91
Oriented R-CNN	89.34	81.63	53.05	71.55	78.94	82.06	87.94	90.88	85.60	85.11	60.20	<u>65.03</u>	73.60	68.46	59.69	75.54
RoI Transformer	89.12	<u>83.54</u>	<u>54.30</u>	<u>75.43</u>	79.08	77.47	87.87	90.90	86.17	84.90	<u>62.95</u>	61.44	74.31	70.55	58.94	75.80
ReDet	<u>89.32</u>	82.30	53.64	70.79	78.55	<u>82.27</u>	<u>88.08</u>	90.89	87.35	85.92	62.92	64.30	<u>75.15</u>	69.71	64.99	<u>76.41</u>
RoI Transformer+*	89.20	83.85	55.20	75.72	<u>79.35</u>	83.48	88.26	90.90	<u>87.19</u>	<u>85.86</u>	66.03	63.86	76.86	<u>71.14</u>	<u>61.44</u>	77.22

注:加粗、下划线字体表示各列最优、次优结果。

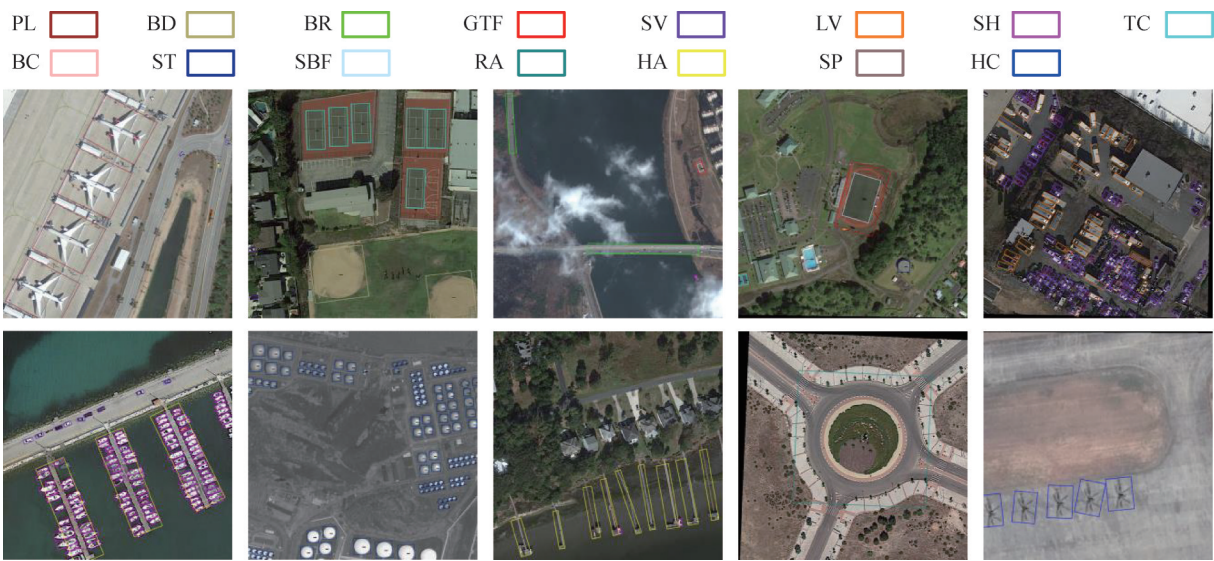


图 11 CGR-OBb 基于 RoI Transformer 在 DOTA 数据集上可视化结果
Fig. 11 Visualization results of CGR-OBb based on RoI Transformer on the DOTA dataset

3 结 论

针对旋转目标检测器因 OBB 表示不连续而限制检测精度的问题,所提出的 CGR-OBb 方法通过建立连续性表示,有效规避了由表示不连续性导致的检测误差。具体而言,CGR-OBb 被设计为编码模块和解码模块。编码模块基于目标形态变化时 OBB 在最小外接 HBB 上的几何特性,将目标四角点标注映射为 9 个连续表示参数,包括位置参数、面积因子和比例参数。这些参数在回归过程中不仅具备连续性,而且遵循定义规则变化。解码模块是表示参数的严格可逆映射,其中位置参数解码 OBB 角点坐标,面积因子解码唯一 OBB 表示,比例参数进行最后的精细调整以达成高精度检测,编码和解码过程满足严格的连续性指标要求。CGR-OBb 的连续性保证了模型收敛的稳定性,编码参数的规则变化使其能够有效应用于多种检测器(包括等变检测器),而解码的鲁棒性则进一步提升了检测器的泛化能力。实验结果表明,CGR-OBb 能够有效检测不同尺寸、任意方向密集排布的目标(尤其是高长宽比目标),在面对复杂背景等困难任务时也表现出色;且因为避免了解码模糊,在高精度指标上表现尤为突出,与现有最先进方法相比具有较强竞争力。然而,本文方法在 mAP@50 上的改善效果相对较弱,这是因为方法中引入的连续参数编码和解码过程更侧重

于精细的结构调整,对提升低 IoU 阈值性能的贡献较低,表明其在低精度检测任务中仍有优化空间。未来工作将兼顾高、低精度检测的优化,以更全面地提升性能。

参考文献(References)

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 (CVPR). Long Beach, USA: IEEE: 2844-2853 [DOI: 10.1109/CVPR.2019.00296]
Han J M, Ding J, Xue N and Xia G S. 2021. ReDet: a rotation-equivariant detector for aerial object detection//Proceedings of 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Nashville, USA: IEEE: 2786-2795 [DOI: 10.1109/CVPR46437.2021.00281]
He X, Ma S P, He L Y, Ru L and Wang C. 2021. Learning rotated inscribed ellipse for oriented object detection in remote sensing images. Remote Sensing, 13(18): #3622 [DOI: 10.3390/rs13183622]
Liu Z K, Hu J G, Weng L B and Yang Y P. 2017a. Rotated region based CNN for ship detection//Proceedings of 2017 IEEE International Conference on Image Processing (ICIP). Beijing, China: IEEE: 900-904 [DOI: 10.1109/ICIP.2017.8296411]
Liu Z K, Yuan L, Weng L B and Yang Y P. 2017b. 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 (ICPRAM). Porto, Portugal: SciTePress: 324-331 [DOI: 10.5220/0006120603240331]
Ma J Q, Shao W Y, Ye H, Wang L, Wang H, Zheng Y B, et al.

- Arbitrary-oriented scene text detection via rotation proposals. *IEEE Transactions on Multimedia*, 20(11): 3111-3122 [DOI: 10.1109/TMM.2018.2818020]
- Nie G T and Huang H. 2023. Multi-oriented object detection in aerial images with double horizontal rectangles. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 45(4): 4932-4944 [DOI: 10.1109/tpami.2022.3191753]
- Pu C D, Yu J, Su W and Liu T Y. 2024. Rotated R-CNN: a two-stage object detection method adapted to oriented bounding boxes//*Proceedings of 2024 IEEE International Conference on Image Processing (ICIP)*. Abu Dhabi, United Arab Emirates: IEEE: 256-262 [DOI: 10.1109/ICIP51287.2024.10647377]
- Qian W, Yang X, Peng S L, Yan J C and Guo Y. 2021. Learning modularized loss for rotated object detection//*Proceedings of the 35th AAAI Conference on Artificial Intelligence*. [s.l.]: AAAI Press: 2458-2466 [DOI: 10.1609/aaai.v35i3.16347]
- 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]
- Shi Z H, Wu C W, Li C J, You Z Z, Wang Q and Ma C C. 2023. Object detection techniques based on deep learning for aerial remote sensing images: a survey. *Journal of Image and Graphics*, 28(9): 2616-2643 (石争浩, 仵晨伟, 李建成, 尤珍臻, 王泉, 马城城. 2023. 航空遥感图像深度学习目标检测技术研究进展. *中国图象图形学报*, 28(9): 2616-2643) [DOI: 10.11834/jig.221085]
- Wang X Y, Liang W H, Li J, Mu J H and Wang X Y. 2025. Fusion of bounding box Gaussian modeling and feature aggregation distribution for fine-grained recognition of remote sensing aircraft images. *Journal of Image and Graphics*, 30(1): 282-296 (王晓燕, 梁文辉, 李杰, 牟建宏, 王禧钰. 2025. 融合边界框高斯建模与特征聚合分发的遥感飞机细粒度识别. *中国图象图形学报*, 30(1): 282-296) [DOI: 10.11834/jig.230862]
- 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]
- Xiao Z K, Yang G Y, Yang X, Mu T J, Yan J C and Hu S M. 2024. Theoretically achieving continuous representation of oriented bounding boxes//*Proceedings of 2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. Seattle, USA: IEEE: 16912-16922 [DOI: 10.1109/CVPR52733.2024.01600]
- Xie X X, Cheng G, Wang J B, Yao X W and Han J W. 2021. Oriented R-CNN for object detection//*Proceedings of 2021 IEEE/CVF International Conference on Computer Vision (ICCV)*. Montreal, Canada: IEEE: 3520-3529 [DOI: 10.1109/ICCV48922.2021.00350]
- Xu H, Liu X Y, Xu H N, Ma Y K, Zhu Z J, Yan C G, et al. 2023. Rethinking boundary discontinuity problem for oriented object detection//*Proceedings of 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. Seattle, USA: IEEE: 17406-17415 [DOI: 10.1109/CVPR52733.2024.01648]
- Xu Y C, Fu M T, Wang Q M, Wang Y K, Chen K, Xia G S, et al. 2021. Gliding vertex on the horizontal bounding box for multi-oriented object detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 43(4): 1452-1459 [DOI: 10.1109/tpami.2020.2974745]
- 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 (CVPR)*. 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 (ECCV)*. 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 (ICML)*. [s.l.]: PMLR: 11830-11841
- Yang X, Yang J R, Yan J C, Zhang Y, Zhang T F, Guo Z, et al. 2019. SCRDet: towards more robust detection for small, cluttered and rotated objects//*Proceedings of 2019 IEEE/CVF International Conference on Computer Vision (ICCV)*. Seoul, Korea (South): IEEE: 8231-8240 [DOI: 10.1109/ICCV.2019.00832]
- 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 (ICLR)*. 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 (CVPR)*. Vancouver, Canada: IEEE: 13354-13363 [DOI: 10.1109/cvpr52729.2023.01283]
- 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 K X, Zhang M, Zhao H H, Tang R, Lin S, Cheng X, et al. 2023. Arbitrary-oriented ellipse detector for ship detection in remote sensing images. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 16: 7151-7162 [DOI: 10.1109/JSTARS.2023.3267240]

作者简介

姚睿,女,副教授,主要研究方向为嵌入式系统及计算机测控技术、可重构硬件与系统容错设计、智能电路与自适应系统、智能算法与智能信息处理。

E-mail: yaorui@nuaa.edu.cn

李雨蔓,通信作者,女,硕士研究生,主要研究方向为计算机视觉与数字图像处理。E-mail: 2303008@nuaa.edu.cn

郭浩帆,男,硕士研究生,主要研究方向为图像及视频分析与处理。E-mail: guohaofan@nuaa.edu.cn

胡文涛,男,硕士研究生,主要研究方向为多维智能感知与演化修复技术。E-mail: huwentao@nuaa.edu.cn

田祥瑞,男,博士,讲师,主要研究方向为机器人智能感知和多无人系统协同技术。E-mail: xiangruitian@nuaa.edu.cn