

中图法分类号: TP75 文献标识码: A 文章编号: 1006-8961(XXXX)XX-0001-17

论文引用格式: Dong Yan, Jia Jijie, Gao Guangshuai, Gao Junyu, Li Xiangyun, Li Chunlei. Confidence-Weighted Guided Angle Correction for Remote Sensing Rotated Object Detection[J/OL]. Journal of Image and Graphics, XXXX: 1-17. DOI: 10.11834/jig.260167. (董燕, 贾季洁, 高广帅, 高君宇, 李湘云, 李春雷. 置信度加权引导角度修正的遥感旋转目标检测[J/OL]. 中国图象图形学报, XXXX: 1-17. DOI: 10.11834/jig.260167.) [DOI: 10.11834/jig.260167]

置信度加权引导角度修正的遥感旋转目标检测

董燕¹, 贾季洁¹, 高广帅¹, 高君宇², 李湘云¹, 李春雷¹

1. 中原工学院信息与通信工程学院, 郑州 450007; 2. 西北工业大学, 西安 710072

摘要: 目的 旋转目标检测是遥感图像处理领域的重要任务, 针对其存在的角度边界不连续和旋转框参数耦合导致的回归困难两大挑战, 本文提出置信度加权引导角度修正的旋转目标检测算法。方法 首先, 为解决角度边界不连续问题, 引入了角度调制模块(Angle Modulation Module, AMM), 通过复指数编解码机制将角度映射为连续可微信号, 从而使得边界回归平滑。其次, 设计三支解耦回归(Three-Branch Decoupled Regression, TBDR)模块, 将传统联合回归检测头解耦为中心坐标、尺度和角度三个独立分支, 实现各几何参数的独立学习, 解决参数特征混淆与梯度冲突。最后, 提出动态角度置信度加权(Dynamic Angle Confidence Weighting, DACW)机制, 通过轻量级置信度预测分支量化角度预测可靠性, 动态调控 AMM 模块的修正强度, 实现“高置信度预测少修正、低置信度预测多修正”的智能优化。结果 实验在 DOTA v1.0 和 HRSC2016 两个公开遥感数据集上与其他算法进行了比较, 结果验证了本文所提方法的有效性和优越性。所提方法在 DOTA v1.0 数据集上的 mAP 达到 76.52%, 相较基准算法提升了 8.09%, 在 HRSC2016 数据集上达 90.30%。消融实验进一步证实了所提模块的有效性。结论 本文提出了一种置信度加权引导角度修正的遥感旋转目标检测算法, 有效解决了角度边界不连续与旋转框参数耦合问题, 显著提升了复杂场景下任意方向目标的检测精度。

关键词: 遥感图像; 旋转目标检测; 角度边界不连续; 回归参数耦合; 置信度加权

Confidence-Weighted Guided Angle Correction for Remote Sensing Rotated Object Detection

Dong Yan¹, Jia Jijie¹, Gao Guangshuai¹, Gao Junyu², Li Xiangyun¹, Li Chunlei¹

1. School of Information and Communication Engineering, Zhongyuan University of Technology, Zhengzhou 450007, China; 2. Northwestern Polytechnical University, Xian 710072, China

Abstract: **Objective** Rotated Object Detection (ROD) is a critical and essential task in the field of remote sensing image processing, playing a vital role in numerous core application scenarios, including land surveying, national defense, agricultural monitoring, disaster emergency response, and urban planning. This task aims to accurately identify and locate direction-sensitive objects in remote sensing images, such as ships, aircraft, vehicles, buildings, and bridges. Unlike horizontal object detection in natural scenes, these remote sensing targets typically exhibit extreme scale variations, dense

收稿日期: 2026-04-01; 修回日期: 2026-05-14

基金项目: 国家自然科学基金(62301623, 62472463); 河南省重点研发计划项目(241111220700); 河南省高等学校重点科研项目(25A620001); 中原工学院自然科学基金重点项目(K2026ZD021); 河南省自然科学基金青年基金(B类)项目(262300421225)

Supported by: National Natural Science Foundation of China (62301623, 62472463); Henan Key Research and Development Project (241111220700); Key Scientific Research Project of Colleges and Universities in Henan Province (25A620001); Key Project of the Natural Science Foundation of Zhongyuan University of Technology (K2026ZD021); Youth Science Foundation Project (Category B) of the Natural Science Foundation of Henan Province (241111220700)

distribution, arbitrary orientations, and complex backgrounds. These characteristics render ROD a significantly more challenging research direction. Traditional Horizontal Bounding Box (HBB) detection methods fail to accurately fit rotating object contours, often introducing excessive background redundancy that leads to small object detection failures and positioning inaccuracies. In contrast, Oriented Bounding Box (OBB) detection techniques effectively address this issue by predicting five parameters: the object's center coordinates, width, height, and rotation angle. Consequently, OBB has become the mainstream research direction in remote sensing. However, existing ROD methods still face two core challenges constraining performance improvement: the angular boundary discontinuity problem and the parameter coupling issue in regression tasks. The first challenge, angular boundary discontinuity, fundamentally stems from the inherent contradiction between the periodic nature of angular parameters (0° and 180° being equivalent in practical applications) and the requirement for continuous differentiability in regression tasks. When predicted angles approach these boundaries, minor angular errors can trigger abrupt jumps in the loss function, leading to unstable training, slow convergence, and prediction biases. Although existing IoU-based joint optimization methods and various angle encoding schemes (e. g. , dual-angle approach, complex encoding) attempt to convert discrete angle prediction into a continuous regression task, they have not yet fundamentally realized stable and continuous angle estimation. The second challenge lies in the coupling of rotation box parameters. Mainstream rotation detectors (e. g. , Rotated RetinaNet, Rotated Faster R-CNN) typically employ a shared regression head to jointly predict center coordinates, scale, and rotation angle parameters. This design overlooks the fundamental differences in physical meaning and learning characteristics between these geometric parameters: center coordinates and scale represent translation and scaling transformations in Euclidean space, while the rotation angle represents a rotational transformation in periodic space. Feature confusion and gradient conflicts between these parameters compromise the learning process. In particular, unstable angle predictions propagate errors to other parameters through coordinate computations, significantly degrading the overall model performance. To address these issues, this paper proposes a high-precision ROD algorithm based on confidence-weighted guided angle correction, by extending the Rotated RetinaNet framework. **Method** First, an Angle Modulation Module (AMM) is introduced to resolve angular boundary discontinuity. Inspired by complex exponential encoding-decoding mechanisms, this module maps angular parameters to continuous quantizable complex-valued signals, then restores angle values via differentiable inverse transformations. When the angular frequency ω is set to 2, the encoding function ensures consistent encoding outputs for target rotation angles θ and $\theta + \pi$. This resolves prediction instability caused by segmented function fitting in traditional methods, guaranteeing continuity and differentiability in angle regression. Second, to address parameter coupling in rotation boxes, the Three-Branch Decoupled Regression (TBDR) module is designed. This module decouples the traditional joint regression head into three parallel, independent prediction branches, responsible for center coordinate, scale, and rotation angle prediction, respectively. These branches share underlying feature extraction to ensure parameter efficiency. Each branch independently learns the transformation rules for its corresponding geometric attribute: the center branch focuses on capturing positional offsets between predicted and ground truth bounding boxes; the scale branch handles width-height scaling relationships; and the angle branch concentrates on learning nonlinear periodic rotation patterns. This dual-layer architecture—combining shared feature extraction with dedicated prediction branches—effectively isolates feature interference and gradient conflicts between parameters, enabling specialized modeling for each geometric parameter. Finally, to maximize the correction capability of AMM while avoiding "over-correction" for high-confidence angle predictions, this paper innovatively introduces the Dynamic Angle Confidence Weighting (DACW) mechanism. This mechanism decouples a lightweight confidence prediction sub-branch from the angle branch, generating confidence scores within the $[0, 1]$ range via a Sigmoid activation function to quantify the reliability of each angle prediction. The confidence score serves as a dynamic weight to regulate the correction intensity of AMM outputs, while the hyperparameter λ fine-tunes the correction magnitude. Specifically, it reduces correction intensity for high-confidence predictions (high scores) to preserve original valid features, and enhances correction intensity for low-confidence predictions (low scores) to eliminate boundary discontinuities and prediction biases. The final angle result is obtained through a weighted fusion strategy combining the original angle prediction and the AMM-corrected angle. **Result** Experimental results on two publicly available remote sensing datasets (DOTA v1.0 and High-Resolution Ship Collection 2016, HRSC2016) validate the effectiveness and generalization capability

ity of the proposed method. Ablation experiments on DOTA v1.0 confirm that the optimal hyperparameter λ is 0.2. The combination of TBDR and DACW produces a synergistic effect, achieving a mean average precision (mAP) of 76.52%, representing an 1.80% improvement over the baseline model with only the Angle Modulation Module (AMM). Comparative experiments demonstrate that the proposed method achieves an mAP of 76.52% on the DOTA v1.0 test set (an 8.09% improvement over the baseline algorithm) and 90.30% on HRSC2016. This paper proposes a confidence-weighted guided angle correction algorithm for rotating target detection in remote sensing. By integrating three core modules—AMM, TBDR, and DACW—it systematically addresses two critical challenges in existing ROD: angular boundary discontinuity and rotation box parameter coupling. The algorithm achieves continuous angle regression through a complex exponential encoding-decoding mechanism, enables independent parameter optimization via a three-branch decoupling architecture, and implements adaptive angle correction using a dynamic confidence-weighted mechanism. The synergistic interaction of these three components effectively resolves angular discontinuity and rotation box coupling issues, significantly enhancing the detection accuracy and training stability of rotating objects. This approach provides a reliable technical basis for practical remote sensing applications in land surveying, national defense, and other fields.

Key words: Remote sensing images; rotated object detection; angular boundary discontinuity; Regression parameter coupling; confidence-weighted

论文引用格式: [DOI:10.11834/jig.260]

0 引言

遥感旋转目标检测作为计算机视觉与遥感图像处理交叉领域的核心任务,在国土测绘、国防安全、农业监测和灾害应急响应等关键领域发挥着不可替代的作用(Gui等,2024)。受俯拍视角影响,遥感图像中的船舶、飞机、建筑物、桥梁等目标普遍呈任意朝向分布,兼具尺度差异极端、分布密集、背景复杂的特性(高广帅等,2025;石争浩等,2023)。传统基于水平边界框(Horizontal Bounding Box, HBB)的检测方法因无法精准贴合旋转目标轮廓,易引入背景冗余信息,导致小目标漏检、边界框之间严重重叠和定位精度不足(袁翔等,2023;Wen等,2023)。为此,基于旋转边界框(Oriented Bounding Box, OBB)的旋转目标检测方法被提出,通过预测目标的中心坐标、尺度及旋转角度等几何参数,实现对任意朝向目标的精准定位,成为遥感领域的重要研究方向(Ge等,2024;余浩东等,2024)。当前方法主要聚焦于旋转敏感特征提取与边界框表示两个方面。

Ding等人(2019)提出RoI Transformer,实现了水平感兴趣区域(Horizontal Region of Interest, H-RoI)到旋转感兴趣区域(Rotated Region of Interest, R-RoI)的转换,并采用旋转RoI对齐(RRoI-Align)来提取特征。然而,该仅能在空间维度上对齐特征,无法实现方向维度对齐。为此,Han等人(2021)提出

ReDet,通过分组卷积生成方向通道特征,并结合RIRoI Align实现空间与方向双维度特征对齐,有效增强了方向敏感特征的鲁棒性。上述基于锚框的(Anchor-based)方法,尽管能够一定程度上提升旋转目标检测的精度,但是普遍存在预设锚框适配性差、易受角度周期性不连续问题的影响。为摆脱对锚框先验的依赖,研究者进一步提出了无锚框(Anchor-Free)方法,直接基于目标关键点或中心点预测OBB参数(Yang等,2019)。Zhou等人(2022)提出的Point RCNN和Yu等人(2024)提出的Point2RBox,均通过目标关键点直接生成旋转边界框,避免了锚框设计的繁琐过程。

旋转边界框表征作为旋转目标检测的基础,直接决定回归的精度和模型的优化效率(王晓燕等,2025)。目前应用最广泛的旋转边界框表示方法是五参数表示法,通过中心点坐标(x, y)、宽高(w, h)和旋转角度(θ)来定义OBB。其中最常见两种定义方式分别是OpenCV定义法和长边定义法两类,OpenCV定义法是将 x 轴逆时针旋转后与旋转边界框重合的第一条边视作参考边,角度范围为 $[-90^\circ, 0^\circ)$;长边定义法是将较长边视作参考边,角度范围为 $[0^\circ, 180^\circ)$,当较长边高于 x 轴时,角度 θ 为负,当其低于 x 轴时,角度 θ 为正。其次,还有多边形表示法,包括八参数表示法和点集表示法虽拟合精度更高,但存在参数冗余、回归成本高、框体易畸变等问题。本文采用五参数表示法中的OpenCV定义法开展研究。尽管现有表征方法持续优化,但仍然会存

在角度边界不连续和旋转参数耦合问题两个关键挑战。

1) **角度边界不连续问题**。由于角度参数具有周期性,角度边界处的微小偏差可能导致损失函数的剧烈跳变,导致基于平滑回归损失(如Smooth L1)的模型难以优化(Yang等,2019),现有基于IoU的联合优化方法(Yang等,2021b)与双角度预测(Ma等,2018)、复数编码(Ding等,2019)等角度编码方案,虽试图将离散的角度预测转化为连续空间中的回归问题,但未从根本上实现连续、稳定的角度预测估计。现有的解决方案主要集中在基于损失函数和基于角度编码设计两类方法。

(1) 基于损失函数的方法是通过设计更平滑的损失度量来缓解边界突变。具体而言,Yang等人(2019)提出IoU-Smooth L1损失抑制梯度波动,但在角度接近边界时仍存在轻微梯度波动问题。为此,Yang等人(2021a)提出GWD,将旋转边界框转换为二维高斯分布,并利用Wasserstein距离度量其差异,但因为高斯分布的协方差矩阵参数设定依赖目标形状先验,对遥感中不规则目标的形状表征能力不足。针对此问题,Yang等人(2021c)进一步利用高斯分布间的KL散度作为回归损失,强化了对目标朝向的判别能力,但因其存在固有的度量偏差问题,密集目标场景下易出现相邻目标回归混淆问题。为此,Yang等人(2022a)提出KFIoU,通过核函数映射与焦点机制优化复杂形状、密集场景的判别效果,仍未突破传统角度表示的局限。Jeffri等人(2024)提出ProbIoU将边界框建模为概率分布,以提升对噪声的鲁棒性,但依旧依赖传统角度参数形式,仍不能很好地解决角度边界不连续问题。这类方法仅从度量层面缓解角度回归波动,未从根本上突破传统角度表示的固有局限,难以同时兼顾复杂遥感场景下的回归稳定性与检测精度。

(2) 基于角度编码设计的方法是通过改变角度的表示形式,从而解决因角度跳变导致的回归不稳定性。早期经典方法中,Ding等人(2019)提出的复数编码借助复数空间连续性缓解周期问题,却未适配旋转框 180° 周期特性, θ 与 $\theta+\pi$ 的编码不一致,仍可能引发边界处的梯度跳变。近年来,Yang等人(2022b)提出圆形平滑标签(Circular Smooth Label, CSL)技术,将角度预测转为分类问题,但角度分类的离散性会限制精细角度的预测,无法满足高精度

检测需求。Yu等人(2023)提出相移编码器(Phase-shifting coder, PSC),将旋转角度编码为连续的正弦或余弦相位变化,利用相位连续性避免边界突变,但该方法未解决编码后角度与其他参数的耦合问题,且后续角度还原过程易引入累积误差。Xu等人(2024)提出角度修正模块,通过可微逆变换解码角度,将不连续的角度回归转化为连续的向量回归,但该模块未考虑角度预测的可靠性差异,采用固定强度的修正策略,易对高置信度预测造成过度修正。这类方法虽在一定程度上缓解了角度边界不连续问题,但仍存在参数耦合问题且缺乏对回归过程中角度预测可靠性的动态适配机制,无法满足高精度遥感旋转目标检测的需求。

区别于现有角度编码方法,本文所采用的基于复指数编解码的角度调制模块(Angle Modulation Module, AMM),其编码机制适配旋转框周期特性,且编解码全程连续可微,同时支持动态修正联动优化。

2) **旋转参数间的耦合干扰问题**(Han等,2021)。主流的旋转目标检测器(如Rotated RetinaNet(Lin等,2017),Rotated Faster R-CNN(Ren等,2016))采用共享检测头联合预测,忽略了不同几何参数在物理意义与学习特性上的本质差异:中心点坐标与尺度属于欧氏空间的平移与缩放变换,而角度属于周期空间的旋转变换(Zhou等,2023)。参数间的特征混淆与梯度冲突会相互干扰,尤其是不稳定的角度预测会通过坐标计算误差传导至其他参数,导致模型整体性能下降(Zhou等,2024)。即便部分方法尝试解耦,如Tian等(2019)通过回归目标四边距离分离位置与几何信息,Tian等(2023)改进CenterNet增设角度分支实现双分支解耦,但角度与尺度参数仍存在特征共享带来的耦合问题,无法实现各参数的完全独立优化,这一问题始终制约着检测精度的提升。

参数解耦是通过分离具有不同物理含义的回归任务,使各分支在专属特征空间中进行专门化学习,从而提升优化稳定性与预测精度。早期方法中, R^3 Det(Yang等,2021b)通过级联旋转检测头修正角度误差,SCRDet++(Yang等,2022c)引入特征增强模块与精细化损失优化,但二者均未实现角度与其他参数的完全解耦;CornerNet(Law等,2018)、FCOS(Tian等,2019)仅在形式上分离位置与几何信息,未

针对旋转角度设计专用解耦架构。

近年来,解耦回归方法聚焦分支结构优化与参数独立性强化。Tian 等人(2023)改进 CenterNet, 实现了角度与“位置/尺度”的双分支解耦,但仍存在中心坐标与尺度参数共享特征、角度分支未适配周期特性的问题。Zhang 等人(2023)改进了基于 FCOS 的遥感检测方法,因采用直接角度回归,仍受角度边界不连续问题影响,精度难以提升。总而言之,现有旋转目标检测方法在参数解耦方面虽取得了一定进展,但仍未实现旋转框各几何参数的完全独立优化,还是会受到参数耦合问题的制约。

为了提升角度预测结果的可靠性,研究者们开始在检测头中引入置信度预测分支,但现有方法存在明显局限。Chen 等人(2020)首次提出在点集回归框架中增加定位质量预测头,将 IoU 作为置信度用于损失加权与非极大值抑制(Non-Maximum Suppression, NMS),但未充分关联角度参数的周期性与边界特性;Dai 等人(2021)提出动态检测头,仅实现特征层面动态调节,未显式输出角度置信度,且未与角度修正模块联动;He 等人(2024)提出的方向不确定性估计模块,仅通过方差调整损失权重,未直接参与角度修正,且置信度分支独立设计,对边界区域、低质量特征的可靠性量化精度不足。总体而言,现有角度置信度加权方法仍存在诸多局限:1)置信度与角度特征关联薄弱,难以精准量化边界区域的预测可靠性;2)加权机制设计单一,仅用于损失加权或 NMS 筛选,未与角度修正、参数解耦形成协同联动,无法实现精准化的智能角度优化。

为解决上述问题,本文提出置信度加权引导角度修正的遥感旋转目标检测算法。首先,针对角度边界不连续的问题,引入角度调制模块(Angle Modulation Module, AMM),通过复指数编解码机制将角度参数映射为连续可微信号,保证角度回归过程的平滑稳定。其次,针对旋转参数间特征混淆与梯度冲突的问题,设计三支解耦回归(Three-Branch Decoupled Regression, TBDR)模块,将传统联合回归检测头解耦为中心点坐标、尺度和角度三个独立分支,让各几何参数能够独立学习对应特征,从结构上规避参数间的相互干扰。最后,针对角度修正强度难以自适应调控的问题,提出动态角度置信度加权(Dynamic Angle Confidence Weighting, DACW)机制,通过轻量级置信度预测分支量化每个

角度预测结果的可靠性,以此动态调控 AMM 模块的修正强度,实现“高置信度预测少修正、低置信度预测多修正”的智能优化。在两个大规模目标检测数据集 DOTA 和 HRSC2016 上进行大量的实验,证明了所提方法的有效性和优越性。

总之,本文贡献可以总结为以下三点:

1) 引入复指数编解码的角度调制模块(Angle Modulation Module, AMM),将角度值以连续光滑函数的形式进行编码和解码,缓解角度回归中的边界不连续问题。

2) 设计三支解耦回归(Three-Branch Decoupled Regression, TBDR)模块,针对旋转框的参数耦合问题,通过将传统的回归检测头解耦为三个独立分支,分别负责中心点、尺度和角度的预测,同时在底层共享特征提取,以实现参数高效且互不干扰的学习。

3) 提出动态角度置信度加权(Dynamic Angle Confidence Weighting, DACW)机制,自适应调整角度修正结果,实现“预测可靠则保持,预测不确定则加强修正”的智能优化。

1 提出方法

本文提出的置信度加权引导角度修正的遥感旋转目标检测算法,以 Rotated RetinaNet(Lin, 2017)为基准,引入角度调制模块(AMM),通过回归头参数解耦与置信度动态加权策略,提升角度预测精度,整体框架如图 1 所示。具体地,为了解决 OBB 中角度不连续问题,首先引入复指数编解码的角度调制模块,通过将旋转角度编码为复指数形式,来提高模型在方向性目标上的收敛稳定性。其次,考虑到不同回归参数具有不同的语义特征和回归特性,耦合回归会导致特征表示混淆;同时,又考虑到角度参数相比其他参数具有更强的敏感性和不确定性,统一的回归策略难以准确建模角度变化规律。因此,构建三支解耦回归网络来避免参数间的特征干扰和梯度冲突,此外,为进一步增强 AMM 的角度修正能力,同时避免对可靠角度预测的过度修正,设计动态角度置信度加权机制,实现“预测可靠则保持,预测不确定则加强修正”的智能优化策略。

1.1 基于复指数编-解码的角度调制模块(AMM)

在遥感旋转目标检测中,边界不连续问题普遍
© 中国图象图形学报版权所有

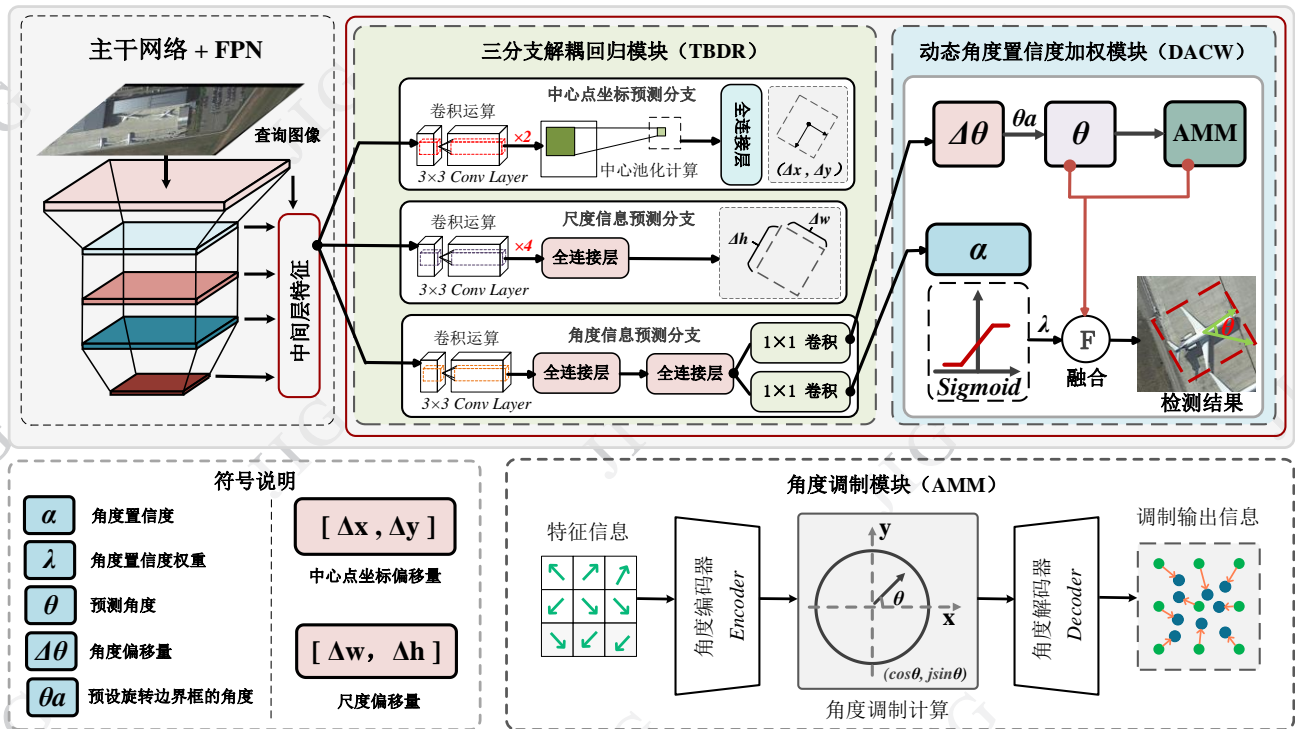


图1 模型框架图

Fig. 1 Model framework diagram

存在,这是由于边界处损失函数的剧烈变化所致。先前方法通常使用独立优化(Yu等,2023)和联合优化IoU类(Yang等,2023)的方法,通过损失平滑来进行处理。但是,理论研究发现,解决该问题的关键在于平滑函数的编码模式,而不是独立或者联合优化。现有的检测模型实质上是拟合边界框与目标的角度关系,但角度边界处的断点使得预测高度不稳定。

为解决以上问题,我们引入一种基于复指数编-解码的角度调制模块(Angle Modulation Module, AMM),通过引入基于复指数编-解码的角度修正方法,首先将角度编码为连续可微的复指数信号,然后通过连续可微的解码过程将其还原为角度值,从而实现角度边界校正和将角度与其他参数进行混合的目的。具体的编、解码的过程如公式(1)-(3)。

$$z = f(\theta) = e^{j\omega\theta} \quad (1)$$

$$\theta = f^{-1}(z) = -\frac{j}{\omega} \ln z \quad (2)$$

$$e^{j\theta} = \cos \theta + j \sin \theta \quad (3)$$

式中, z 是角度编码值, j 是虚数单位($j^2=-1$), ω 是角频率, $f(\cdot)$ 是角度的编码函数,当 $\omega=2$ 时,通过公式(1)的编码函数和公式(3)可使目标旋转角度 θ 与 $\theta+\pi$ 对应相同的编码结果,从根源上解决了传统方法中因拟合分段函数导致的预测不稳定问题,证明过程为:

1)当 $\omega=2$ 时,对于任意角度 θ , $\theta+\pi$ 的编码结果为 $f(\theta + \pi) = e^{j2(\theta + \pi)} = e^{j2\theta} \cdot e^{j2\pi}$,再根据公式(3)的欧拉公式可知,当旋转角度 $\theta=2\pi$ 时, $e^{j2\pi} = \cos 2\pi + j \sin 2\pi = 1 + j \cdot 0 = 1$,因此,编码函数 $f(\theta + \pi) = e^{j2\theta} = f(\theta)$,即 θ 与 $\theta+\pi$ 的编码结果完全一致;2)通过公式(2)将编码值解码为预测角度,以实现角度预测的连续建模。

由于Rotated RetinaNet默认使用的是OpenCV定义法,其核心规则为:将 x 轴逆时针旋转后与旋转边界框重合的第一条边视作参考边,角度范围是 $[-\pi/2, 0)$,这种表示法与AMM的编码方式不兼容,故需要将角度转换成长边定义法中的角度范围 $[0, \pi)$,而长边定义法的核心规则为:以旋转框的较长边为参考边,角度范围为 $[0, \pi)$,长边在 x 轴上方时角度为负,下方时为正。两种定义法的转换需结合宽高判断与角度偏移操作,完整的角度转换公式及推导如公式(4)与算法1所示。

$$\theta_{long} = \begin{cases} (\theta_{cv} + \pi/2) \bmod \pi, w \geq h \\ (\theta_{cv} + \pi) \bmod \pi, w < h \end{cases} \quad (4)$$

式中, θ_{cv} 是OpenCV定义法输出的角度, θ_{long} 是长边定义法输出的角度, w 和 h 分别是旋转框参数的宽和高。推导步骤如算法1所示,其中,定义长边是1,短

边是 s 。

算法 1: OpenCV 定义法转长边定义法

```

输入: OpenCV 格式旋转框参数  $(x, y, w, h, \theta_{cv})$ , 其中  $\theta_{cv} \in [-\pi/2, 0)$ 
输出: 长边定义格式旋转框参数  $(x, y, l, s, \theta_{long})$ , 其中  $\theta_{long} \in [0, \pi)$ 

1 //步骤 1: 判断选状况的长边与短边
2 若  $w \geq h$ 
3    $l=w; s=h$ ; //宽为长边, 高为短边
4   //步骤 2: 基于 OpenCV 角度计算初始转换角度
5    $\theta_{temp} = \theta_{cv} + \pi/2$ ;
6   //步骤 3: 结合长边调整角度
7    $\theta_{long} = \theta_{temp} \bmod \pi$ ;
8 否则
9    $l=h; s=w$ ; //高为长边, 宽为短边
10  //步骤 2: 基于 OpenCV 角度计算初始转换角度
11   $\theta_{temp} = \theta_{cv} + \pi$ ;
12  //步骤 3: 结合长边调整角度
13   $\theta_{long} = \theta_{temp} \bmod \pi$ ;
14  //步骤 4: 修正角度范围
15   $\theta_{long} = \theta_{long} \bmod \pi$ ;
16 返回  $(x, y, l, s, \theta_{long})$ ;

```

AMM 模块可以有效缓解边界不连续问题, 但是采用的优化范式依然是联合优化, 即各回归参数采用共享的特征, 容易出现耦合干扰和梯度冲突问题。

1.2 三支解耦回归头 (TBDR)

传统旋转目标检测的回归头普遍采用多参数联合优化的形式, 将边界框的 5 个参数 (x, y, w, h, θ) , 即位置参数 (中心点坐标 (x, y))、尺度参数 (宽、高 (w, h)) 与角度参数 (θ) , 通过单一分支同时预测。这种设计会导致两个问题: 1) **参数干扰**, 不同几何参数的物理意义与学习特性差异显著 (中心坐标和尺度属于欧氏空间的平移或缩放变换, 而角度属于周期空间的旋转变换), 共享特征空间会导致参数优化过程中的相互干扰; 2) **梯度冲突**, 不同参数对损失函数的敏感程度不同, 角度预测比较敏感, 其误差易通过坐标计算传导至中心与尺度估计, 反之中心和尺度的偏移也会干扰角度的梯度更新, 降低收敛效率。

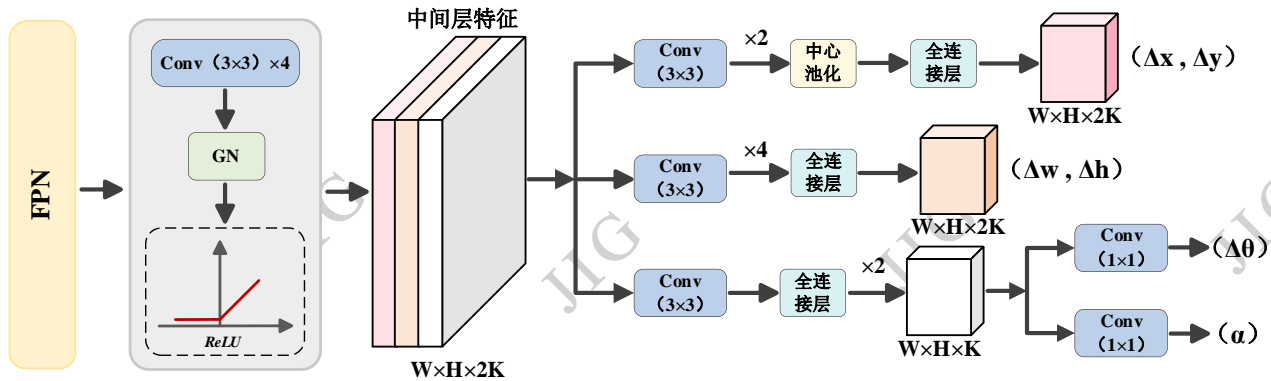


图 2 三支解耦回归示意图

Fig. 2 Schematic Diagram of Three-Branch Decoupled Regression

针对以上问题, 本文提出三支解耦回归 (Three-Branch Decoupled Regression, TBDR), 通过将回归任务解耦为三个并行分支, 分别负责中心坐标、尺度和旋转角度的预测, 同时共享中间层特征提取来保证参数效率, 如图 2 所示, 通过“共享特征提取+专属预测分支”的双层架构, 实现各几何参数的独立学习, 从而解决参数干扰和梯度冲突的影响。

具体地, 回归分支的首要任务是提取包含目标几何信息的基础特征。本文采用 4 层堆叠卷积对输入特征图进行编码, 提取适用于所有回归任务的通用特征, 减少冗余计算。具体过程如公式 (5) 所示。

$$F_{reg} = \text{ReLU}(\text{GN}(f_{conv}^4(\mathbf{x}))) \quad (5)$$

式中, $\mathbf{x} \in \mathbb{R}^{C \times H \times W}$ 是特征金字塔网络 (Feature Pyramid Network, FPN) 输出的特征图, $f_{conv}^4(\cdot)$ 代表 4 层堆叠 3×3 卷积操作, $\text{GN}(\cdot)$ 是分组归一化, $\text{ReLU}(\cdot)$ 是激活函数。此过程为后续三支提供统一、鲁棒的基础特征支撑, 避免特征冗余与信息丢失。

基于中间层特征 F_{reg} , 本文构建三个独立预测分支, 分别专注于特定几何参数的学习, 避免因参数特性差异导致的参数干扰和梯度冲突。1) 目标中心位置定位属于欧氏空间的平移变换, 学习目标为目标几何中心的位置偏移, 对特征的空间定位能力要求

高,且与绝对位置强相关,需独立捕获目标中心区域的局部特征,是检测任务的核心,不受尺度与角度变化干扰;2)尺度估计属于欧氏空间的缩放变换,学习目标为目标宽高的比例关系,与中心坐标的绝对位置无关,但宽与高需联合表征目标形状,二者学习特性高度相似,所以无需单独解耦,且与角度无关;3)角度预测具有周期性,其属于周期空间的旋转变换,学习目标为目标朝向,对特征的方向敏感性要求高,与欧氏空间参数无直接关联,需独立学习周期性变化规律需独立学习其非线性变化规律。

1)中心点坐标预测分支采用中心池化+基础卷积层+全连接层结构,如公式(6)所示,中心坐标回归依赖目标几何中心的局部特征,中心池化通过水平/垂直方向的特征最大化与融合,精准聚合中心区域上下文信息,有效抑制背景干扰,提升空间定位精度;2)尺度预测分支采用4层堆叠3×3卷积+全连接层结构,如公式(7)所示,尺度参数需捕获目标的全局形状特征,4层卷积可有效扩大感受野,适配遥感图像中目标的尺度变化,确保模型稳定学习并拟合目标的宽高比例关系;最终通过全连接层输出对数化的尺度回归结果;3)角度预测分支采用1层3×3卷积+全连接层结构,如公式(8)所示,角度回归依赖目标的方向敏感特征(如边缘朝向),浅层卷积可保留特征细节,避免深层卷积导致的方向特征模糊;同时,角度分支已结合AMM模块做复指数编码,浅层卷积结构足以捕获角度周期性特征,无需深层堆叠。

$$\text{center } r_{\text{pred}} = \text{CenterMaxPool} \left(f_{\text{conv}}^2 \left(\mathbf{F}_{\text{reg}} \right) \right) \in \mathbf{R}^{2 \times K \times H \times W} \quad (6)$$

$$\text{scale } e_{\text{pred}} = f_{\text{conv}}^4 \left(\mathbf{F}_{\text{reg}} \right) \in \mathbf{R}^{2 \times K \times H \times W} \quad (7)$$

$$\theta_{\text{pred}} = f_{\text{conv}} \left(\mathbf{F}_{\text{reg}} \right) \in \mathbf{R}^{1 \times K \times H \times W} \quad (8)$$

式中, $\text{center}_{\text{pred}}$ 、 $\text{scale}_{\text{pred}}$ 和 θ_{pred} 分别是由共享特征解耦出的中心点坐标预测分支、尺度预测分支和角度预测分支, $f_{\text{conv}}(\cdot)$ 代表3×3卷积操作, $f_{\text{conv}}^4(\cdot)$ 代表4层堆叠3×3卷积操作, $f_{\text{conv}}^2(\cdot)$ 代表2层堆叠3×3卷积操作, K 是每个位置的锚框数量,中心偏移分支和尺度缩放分支输出通道数都是2× K ,角度预测分支输出通道数是1× K ,每个锚框分别对应一组 $(\Delta x, \Delta y)$ 、 $(\Delta w, \Delta h)$ 和 $\Delta \theta$,分别专注于学习真实框中心与预测框中心的偏差、真实框尺度与预测框尺度的比例变换和真实框旋转角度与预测框初始角度的差异。

通过该模块,可以避免由多参数联合优化所导

致的参数干扰和梯度冲突问题,让不同的参数可以独立学习各自特征。在此基础上,我们将前文提到的AMM嵌入到角度预测分支中,进行角度修正,缓解边界不连续的问题。

1.3 动态角度置信度加权机制(DACW)

三支解耦回归(TBDR)模块实现了角度特征的独立学习,但角度预测仍面临两大挑战:一是角度周期性导致的边界不连续问题;二是不同样本的角度预测可靠性差异显著。传统方法中,角度修正通常采用固定权重,无法适应这种可靠性差异,我们通过在引入角度调制模块(AMM)一定程度上缓解了边界不连续问题,但是有可能对高置信度的角度预测造成“过度修正”,反而降低精度。为此,本文提出动态角度置信度加权(Dynamic Angle Confidence Weighting, DACW)机制,如图3所示,通过“置信度量化—权重调控修正强度—融合”的结构,同时引入超参数角度置信度权重控制置信度对角度修正的影响程度,实现角度修正强度与预测可靠性的精准适配。

为提升角度预测的精度,通过对角度预测分支进一步解耦生成的置信度预测分支进行激活,生成角度置信度,为每个角度预测结果分配一个置信度分数,作为修正强度的决策依据,具体过程由公式(9)给出。

$$\alpha_i = \sigma(\text{Conv}_{1 \times 1}(\theta_{\text{pred}})) \in [0,1] \quad (9)$$

式中, α_i 是角度置信度分数, $\sigma(\cdot)$ 代表 Sigmoid 激活函数。

角度置信度加权是通过结合角度置信度与AMM实现动态调整。AMM基于目标的角度特征生成角度修正量,而DACW则通过 α_i 并引入可调节权重参数 λ 来控制修正强度,即采用“可靠预测少修正,不可靠预测多修正”的策略,实现初始角度和经过AMM模块后的修正角度的融合,具体过程由公式(10)给出。

$$\theta_i^{\text{final}} = \theta_i + (\alpha_i \times \lambda) \odot (\theta_i^{\text{corr}} - \theta_i) \quad (10)$$

式中, \odot 表示向量乘法, θ_i 、 θ_i^{corr} 和 θ_i^{final} 分别表示原始角度预测值,角度修正量和最终角度值。当 $\alpha_i \rightarrow 1$ 时,表示角度预测可靠, $\theta_i^{\text{final}} \approx \theta_i$,少修正,避免AMM过度修正引入误差;当 $\alpha_i \rightarrow 0$ 时,表示预测不确定, $\theta_i^{\text{final}} \approx \theta_i^{\text{corr}}$,则多修正,通过AMM的连续编码特性修正边界突变与预测偏差。

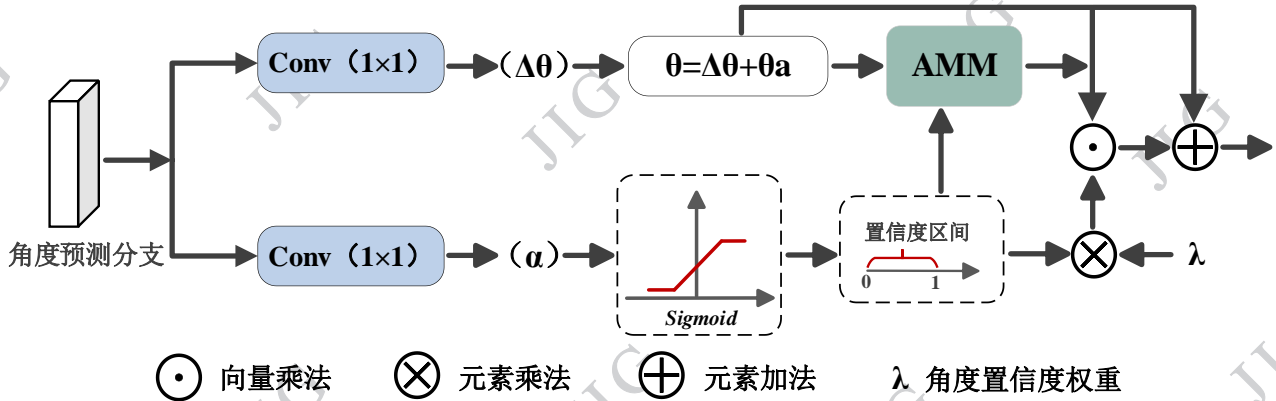


图3 动态角度置信度加权示意图

Fig. 3 Schematic Diagram of Dynamic Angle Confidence Weighting

回归子任务的损失为四个分支损失的加权和, 其中, 1) 中心损失采用 Smooth L1 损失, 缓解异常值对中心坐标回归的影响; 2) 尺度损失也采用 Smooth L1 损失, 增强对小尺度目标的敏感性; 3) 角度损失则采用改进的 Smooth L1 损失, 将角度约束在 $[-90^\circ, 90^\circ]$ 范围内; 4) 其置信度损失采用二元交叉熵 (BCE)。计算过程分别由公式(11)-(15)得到。

$$L_{reg} = \lambda_c L_{center} + \lambda_s L_{scale} + \lambda_\theta L_\theta + \lambda_{conf} L_{conf} \quad (11)$$

$$L_{center} = \frac{1}{N_{pos}} \sum_{i \in N_{pos}} \left\{ \begin{array}{l} SmoothL1(\Delta x_i, \Delta x_i') \\ + SmoothL1(\Delta y_i, \Delta y_i') \end{array} \right\} \quad (12)$$

$$L_{scale} = \frac{1}{N_{pos}} \sum_{i \in N_{pos}} \left\{ \begin{array}{l} SmoothL1(\log \frac{w_i}{w}, \log \frac{w_i'}{w}) \\ + SmoothL1(\log \frac{h_i}{h}, \log \frac{h_i'}{h}) \end{array} \right\} \quad (13)$$

$$L_\theta = \frac{1}{N_{pos}} \sum_{i \in N_{pos}} SmoothL1(\Delta \theta_i, \Delta \theta_i') \quad (14)$$

$$L_{conf} = \frac{1}{N_{pos} + N_{neg}} \sum_i \left\{ \begin{array}{l} \alpha_i' \log \alpha_i + \\ (1 - \alpha_i') \log (1 - \alpha_i) \end{array} \right\} \quad (15)$$

式中, $\lambda_c, \lambda_s, \lambda_\theta$ 和 λ_{conf} 分别是对应分支的平衡系数, λ_{conf} 在本文中设为 0.1, 其余三个系数均设为 1。 N_{pos} 和 N_{neg} 分别是正、负样本数量。 $(\Delta x_i, \Delta y_i)$ 是预测的中心偏移量, $(\Delta x_i', \Delta y_i')$ 是根据真实框与预设框计算出的理想中心偏移量; (w_i, h_i) 是预测框的尺度(宽和高), (w_i', h_i') 是真实框的尺度, (w, h) 是预设框的尺度; $\Delta \theta_i$ 是预测的角度偏移, $\Delta \theta_i'$ 是根据真实框与预设框计算出的理想角度偏移量。

2 实验与结果分析

本研究通过系统实验验证了所提方法在遥感旋

转目标检测任务中的有效性。在本小节将介绍数据集、实验设置、消融实验和对比实验的结果分析。

2.1 数据集和实验设置

DOTA (Xia 等, 2018) 是一个大规模的航拍图像数据集。我们将使用 DOTA v1.0 进行实验。DOTA v1.0 包含 2806 幅图像, 共计 188,282 个实例, 分辨率从 800×800 到 4000×4000 不等, 涵盖 15 个常见类别: 飞机 (Plane, PL)、棒球场 (Baseball-Diamond, BD)、桥梁 (Bridge, BR)、田径场 (Ground-Track-Field, GTF)、小型车辆 (Small-Vehicle, SV)、大型车辆 (Large-vehicle, LV)、船只 (Ship, SH)、网球场 (Tennis-Court, TC)、篮球场 (Basketball-Court, BC)、储油罐 (Storage-Tank, ST)、足球场 (Soccer-Ball-Field, SBF)、环形交叉路口 (Roundabout, RA)、港口 (Harbor, HA)、游泳池 (Swimming-Pool, SP) 和直升机 (Helicopter, HC)。由于 DOTA v1.0 数据集中的图像尺寸太大, 我们通过 DOTA_devkit 从原始图像中裁剪出一系列 1024×1024 像素、重叠 200 像素的图像块。

HRSC2016 (Liu 等, 2017) 是专为船舶检测开发的高分辨率遥感数据集, 包含 1061 幅图像, 共检测到 2976 艘船舶实例。训练集、验证集和测试集分别有 436 张、181 张和 444 张。由于图像尺寸存在差异 (从 300×300 到 1500×900 像素), 我们先将每幅图像的长边缩放到 800 像素, 短边按原始宽高比自适应缩放, 训练时将图像统一补边至 1024×1024 像素的固定输入尺寸, 全程保持原始宽高比。

本文中所有实验均在单个 GeForce RTX 4090 GPU 上进行的, 并在精心设计的训练设置下确保严

格评估。学习率、动量和权重衰减的超参数分别设置为 0.25、0.9 和 0.0001。对于 DOTA v1.0 数据集和 HRSC2016 数据集,模型在单一尺度下训练,输入尺寸固定为 1024×1024 像素,批次为 2,在 DOTA v1.0 数据集上的训练周期是 12epoch,而在 HRSC2016 数据集的训练周期是 72epoch。

评价指标采用平均精度均值(mean Average Precision, mAP)与 IoU 阈值为 0.5 时的平均精度,其中,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。

2.2 消融实验

本文是在 DOTA v1.0 数据集上做的消融实验,如表 1 所示,1)为了验证三支解耦模块的有效性,将回归部分进行两分支解耦(Dual-Branch Decoupled Regression, DBDR),即将回归分支解耦成中心、尺度联合子分支和角度子分支,该模块加入后,mAP 从 74.76% 提升至 75.28%(+0.58%),AP50 从 74.80% 提升至 75.30%(+0.50%),这一结果表明角度参数与中心和尺度参数分离的有效性:角度子分支摆脱了位置、尺度特征的干扰,能够专注学习角度的周期性变化规律,减少了跨参数的误差传导。

2)当 TBDR 加入后,模型的 mAP 从 74.76% 提升至 75.35%(+0.59%),AP50 从 74.80% 提升至 75.40%(+0.60%),较 DBDR 额外获得 0.07%(mAP)和 0.10%(AP50)的性能增益。这一结果证明即使中心坐标与尺度参数同属欧氏空间变换参数,二者的学习需求仍存在差异:中心定位对目标轮廓的几何中心敏感性更高,易受尺度变化的间接干扰;而尺度估计需聚焦宽、高比例关系,与中心位置的绝对坐标关联较弱。TBDR 设计的“共享特征提取+专属分支预测”的架构,将中心坐标、尺度、角度分配至独立分支进行专业化学习,有效避免了跨参数的特征干扰和梯度冲突。特别是角度参数作为旋转目标检测的敏感因子,其独立分支能够更专注于学习周期性变化规律,进一步减少了中心、尺度偏移对角度预测的误差传导,最终实现了优于 DBDR 的检测性能。

3)为了进一步验证三支解耦模块的有效性,将尺度参数进一步拆分为宽分支与高分支,实现四个参数的完全独立解耦。该模块加入后,mAP 为 75.11%,AP50 为 75.10%,虽然比未加入前精度有所提高,但是其精度不仅低于加入 TBDR 的精度,甚

至低于加入 DBDR 的精度。这一结果证明四分支架构存在一些问题:(1)宽与高同属尺度参数,二者的物理意义高度关联(共同表征目标的形状与大小),学习特性高度相似,单独解耦会导致特征冗余,在增加模型计算开销的同时,破坏了宽高比例的内在约束,易引发旋转框形态畸变(如宽高比例失衡);(2)宽、高分支的独立学习会分散模型对尺度特征的提取能力,导致尺度估计精度下降,进而间接影响中心坐标与角度参数的回归稳定性。因此,TBDR 架构是“解耦充分性”与“模型简洁性”的最优平衡,既通过三支设计解决了不同空间参数的耦合问题,又避免了过度解耦带来的特征冗余与约束失效。

表 1 验证解耦模块的消融实验
Table 1 Ablation Experiment for Verifying the /Decoupling Module

基线网络	DBDR	TBDR	FBDR	mAP/%	AP50/%
√	-	-	-	74.76	74.80
√	√	-	-	75.28	75.30
√	-	√	-	75.35	75.40
√	-	-	√	75.11	75.10

注:“√”为采用,“-”为未采用。加粗字体为每行最优值

如表 2 所示,对 DACW 中的权重参数的敏感度做消融实验,即确定 λ 的最优值,1)当 $\lambda=0.05$ 时,mAP 仅为 75.48%,此时权重过小,置信度对修正强度的调控能力不足;2)当 $\lambda=0.2$ 时,mAP 达到峰值 76.52%,此时置信度的修正增益与角度特征学习的平衡最优;3)当 $\lambda>0.2$ 时,mAP 逐渐下降至 75.47%,此时,权重过大导致置信度过度主导角度修正,掩盖了原始角度特征的有效信息。因此,我们在后续的实验均采用 $\lambda=0.2$ 。

如表 3 所示,为量化各核心模块的独立贡献,我们开展 1)单模块独立消融实验:(1)当仅引入 AMM,即采用耦合回归范式时,模型的 mAP 从 68.43% 提升至 74.76%(+6.33%),AP50 从 68.45 提升至 74.80%(+6.35%),这一结果说明 AMM 模块加入后,精度明显提升,可以有效缓解角度边界不连续问题;(2)单独引入 TBDR 模块时,模型 mAP 达到 74.92%(+6.49%),AP50 达到 74.90%(+6.45%),这一结果验证了三支解耦架构的有效性:将中心坐标、尺度修正模块,仍能通过参数解耦实现显著性能

表2 验证角度置信度权重敏感度的影响

Table 2 Verifying the Impact of Angle Confidence

/Weight Sensitivity		
λ	mAP/%	AP50/%
0.05	75.48	75.50
0.1	75.65	75.60
0.2	76.52	76.50
0.3	76.20	76.20
0.4	75.47	75.50

注:加粗字体为每行最优值

提升;(3)单独引入DACW模块时,模型mAP提升至74.96%(+6.53%),AP50提升至75.00%(+6.55%),这一结果表明,动态角度置信度加权机制通过量化角度预测可靠性并动态调整加权策略,能有效过滤低置信度预测的偏差,提升角度回归的精准度。

2)进一步开展模块间交叉消融实验,探究各模块的协同增益效应。(1)当模块AMM和TBDR组合实验时,模型mAP从只加入AMM时的74.76%提升至75.35%(+0.59%),AP50从只加入AMM时的74.8%提升至75.4%(+0.60%),这一结果表明,TBDR的解耦架构为AMM提供了角度专属特征,避免了耦合回归中位置、尺度特征对角度编码的干扰,让复指数编解码的调制效果更充分发挥;(2)当模块AMM和DACW组合实验时,模型mAP从只加入AMM时的74.76%提升至75.45%(+0.69%),AP50从只加入AMM时的74.8%提升至75.40%(+0.60%),这一结果表明,DACW能精准识别AMM调制后的角度预测可靠性,对高置信度修正结果少干预、对低置信度结果多优化,形成“编码修正+置信度筛选”的闭环,进一步降低角度回归误差;(3)当TBDR与DACW组合实验时,模型mAP达到76.07%,AP50达到76.10%,这一结果表明,TBDR为角度预测提供独立、鲁棒的特征基础,使DACW能更精准地量化角度预测可靠性,而DACW的动态加权策略又能优化TBDR分支的学习方向,减少低质量角度预测对模型优化的干扰,二者协同实现了角度回归精度的大幅提升。

(4)当AMM、TBDR和DACW三者组合实验时,模型性能达到最优,mAP提升至76.52%,AP50提升至76.50%,这一结果充分证明,TBDR的完全解耦架构能最大程度降低参数间的干扰,使角度分支输

表3 模块的独立、交叉消融实验

Table 3 Single Module Independent +Inter-Module Cross

/Ablation Experiment					
RetinaNet -R	AMM	TBDR	DACW	mAP/%	AP50/%
√	-	-	-	68.43	68.45
√	√	-	-	74.76	74.80
√	-	√	-	74.92	74.90
√	-	-	√	74.96	75.00
√	√	√	-	75.35	75.40
√	√	-	√	75.45	75.40
√	-	√	√	76.07	76.10
√	√	√	√	76.52	76.50

注:“√”为采用,“-”为未采用。加粗字体为每行最优值

出更精准、鲁棒的角度特征,为DACW的动态加权修正提供更优质的基础,而DACW则通过置信度引导进一步释放角度修正的潜力,三者协同互补,既解决了核心技术挑战,又最大化发挥了各模块的优势,最终实现了显著的性能突破。

2.3 与其他算法对比

我们在两个数据集上分别进行了实验对比,在DOTAv1.0数据集上做的实验,如表4所示,本文方法达到76.52%的mAP,优于所有其他对比方法,展现出较好的综合检测性能。从类别级表现来看,本文方法尤其在飞机(PL)、棒球场(BD)、田径场(GTF)、环形交叉路口(RA)、游泳池(SP)和直升机(HC)等类别上表现突出,AP值均处于所有对比算法的首位。

针对单个类别时,1)对于飞机(PL)这类几何结构清晰、多分布于空旷天空背景的目标,其明确的方向特征让AMM模块的复指数编解码充分发挥连续回归优势,TBDR的三支解耦也能精准捕获中心、尺度与角度参数;2)对于棒球场(BD)、田径场(GTF)、游泳池(SP)这三类均为规则多边形/圆形结构、边界特征鲜明的目标,其宽高比例稳定、无复杂轮廓导致的方向模糊,TBDR模块的尺度分支可高效学习形状特征,角度分支专注于周期性旋转变换建模,配合DACW机制基于高置信度预测精准调控修正强度,实现了角度回归精度的大幅提升;3)对于环形交叉路口(RA)这类环形布局显著、中心区域明确的目标,TBDR的中心分支可精准定位几何中心,

AMM 模块适配其平缓的角度变化,避免边界不连续问题;4)对于直升机(HC)这类方向分布多样、轮廓辨识度高的目标,TBDR的独立角度分支摆脱了位置与尺度特征的干扰,AMM的连续编码缓解了方向

多变带来的回归不稳定,DACW 机制则能动态修正低置信度预测的角度偏差,三大模块协同作用最大化发挥了技术优势,最终实现高 AP 表现。

表4 在DOTAv1.0数据集上与不同算法的比较

Table4 Comparison with Different Algorithms on the DOTAv1.0 Dataset

方法	PL	BD	BR	GTF	SV	LV	SH	TC	BC	ST	SBF	RA	HA	SP	HC	mAP/%
RetinaNet-R(Lin等,2017)	88.67	77.62	41.81	58.17	74.58	71.64	79.11	90.29	82.18	74.32	54.75	60.60	62.57	69.67	60.64	68.43
RADet(Li等,2020)	79.45	76.99	48.05	65.83	65.46	74.40	68.86	89.70	78.14	74.97	49.92	64.63	66.14	71.58	62.16	69.09
KLD(Yang等,2021c)	89.50	79.91	39.92	70.40	78.04	64.24	82.79	90.90	81.80	83.02	57.63	63.52	56.63	65.13	50.04	70.23
PSC(Yu等,2023)	89.32	82.29	37.92	71.52	78.40	66.33	78.01	90.89	84.21	80.63	60.22	64.33	59.69	68.37	53.85	71.09
DAL(Ming等,2021)	88.61	79.69	46.27	70.37	65.89	76.10	78.53	90.84	79.98	78.41	58.71	62.02	69.23	71.32	60.65	71.78
MA-RPDet(张家源等,2024)	89.54	75.59	42.90	73.93	83.77	75.12	88.05	90.81	82.60	86.68	56.28	60.59	65.84	66.84	46.72	72.35
SARD(Wang等,2019)	89.93	84.11	54.19	72.04	68.41	61.18	66.00	90.82	87.79	86.59	65.65	64.04	66.68	68.84	68.03	72.95
GWD(Yang等,2021a)	89.06	80.56	44.27	73.02	79.51	73.53	85.55	90.89	86.21	83.26	63.17	64.24	63.56	69.04	52.92	73.35
CFC-Net(Qi等,2021)	89.08	80.41	52.41	70.02	76.82	78.11	87.21	90.89	84.47	85.64	60.51	61.52	67.82	68.02	50.09	73.50
R ³ Det(Yang等,2021b)	88.80	83.10	50.90	67.30	76.20	80.40	86.70	90.80	84.70	83.20	62.00	61.40	66.90	70.60	53.90	73.80
S ² ANet(Han等,2021)	89.11	82.84	48.37	71.11	78.11	78.39	87.25	90.83	84.90	85.64	60.36	62.60	65.26	69.13	57.94	74.12
ARS-DETR(曾颖等,2025)	86.97	75.56	48.32	69.20	77.92	77.94	87.69	90.50	77.31	82.86	60.28	64.58	74.88	71.76	66.62	74.16
Gliding Vertex(Xu等,2020)	89.64	85.00	52.26	77.34	73.01	73.14	86.82	90.74	79.02	86.81	59.55	70.91	72.94	70.86	57.32	75.02
CSL(Yang等,2022a)	90.25	85.53	54.64	75.31	70.44	73.51	77.62	90.84	86.15	86.69	69.60	68.04	73.83	71.10	68.93	76.17
本文算法	90.60	89.10	52.30	83.90	74.20	62.50	76.00	90.80	79.00	68.70	70.70	85.60	60.10	81.30	83.00	76.52

注:加粗字体为每行最优值

虽然本文方法综合性能领先,但在大型车辆(LV)、储油罐(ST)、港口(HA)三个类别上的 AP 表现略低于部分对比算法,推测可能的原因有:1)对于大型车辆(LV)这类密集分布、与背景灰度值相似且重叠度高的目标,因为本文方法的核心设计聚焦于高方向敏感性场景的角度优化,未针对密集目标设计专用的 NMS 优化等后处理策略,因此在该极端场景下性能存在一定局限。经典算法 R³Det 通过级联旋转检测头逐步修正角度误差,有效提升了密集目标的检测精度,因此在该类别上表现更优,但其针对密集目标做了大量优化,模型复杂度、参数量与推理延迟显著更高,且在飞机(PL)、棒球场(BD)等强方向敏感目标上的检测精度远低于本文方法;2)对于储油罐(ST)这类无朝向意义的圆形目标且部分尺度极小的目标,因为本文方法的核心创新均围绕“角度参数的精准回归”展开,针对的是具有明确朝向的

目标,因此在这类无角度需求的目标上,模块优势无法充分发挥,属于设计侧重点的适配性问题,而非模块本身的缺陷。经典算法 Gliding Vertex 通过顶点偏移建模与自适应边界框优化,有效提升了圆形目标的检测精度,因此在该类别上表现更优,但其针对特定形状目标做了定制化设计,未针对角度边界不连续问题进行优化,且在飞机(PL)、直升机(HC)等强方向敏感目标上的检测精度显著低于本文方法;3)对于港口(HA)这类无明确几何轮廓、内部子目标密集且背景复杂的目标,因为本文方法是基于旋转框五参数表示法,核心优势是规则目标的精准回归,因此在该场景下的适配性有限,经典算法 ARS-DETR 基于 Transformer 架构实现了端到端检测,在该类别上表现更优,但该方法采用复杂的多尺度特征增强架构,模型参数量大、推理速度慢,且在旋转目标的角度回归精度、强方向敏感目标检测性能上远弱于

本文方法,综合检测性能不及本文方法,无法满足通用遥感旋转目标检测的需求。

综上,本文方法在具有明确几何轮廓、强方向特征、中等尺度的遥感旋转目标上,实现了全类别领先的检测性能,三大模块的协同设计充分发挥了技术优势;而在密集目标、无朝向圆形目标等特定场景下的性能局限,本质上是方法设计侧重点与目标特性的适配性差异,而非模块缺陷。对比算法在这些场景的优势均以牺牲通用性能、计算效率为代价,本文方法在综合检测精度、多场景泛化能力上实现了更优的平衡。

表5 在HRSC2016数据集上与不同算法的比较
Table 5 Comparison with Different Algorithms on the /HRSC2016 Dataset

方法	mAP/%
RetinaNet-R(Lin等,2017)	80.81
SARD(Wang等,2019)	85.40
KLD(Yang等,2021c)	87.45
Gliding Vertex(Xu等,2020)	88.20
DAL(Ming等,2021)	88.95
R ³ Det(Yang等,2021b)	89.26
CSL(Yang等,2022a)	89.62
CFC-Net(Qi等,2021)	89.70
GWD(Yang等,2021a)	89.85
PSC(Yu等,2023)	90.06
S ² ANet(Han等,2021)	90.17
本文算法	90.30

注:加粗字体为每行最优值

此外,为了验证所提算法的泛化性能,我们在表5中展示了所提方法在HRSC2016数据集上与现有算法的对比,也取得了最好的检测效果,说明我们所设计的模块,对于舰船这类纵横比较大的目标,也表现出不错的特征提取能力。

2.4 可视化结果对比与分析

为了更直观地展示本文方法的检测效果,图4展示了该方法在DOTAv1.0数据集上多个典型场景中的检测结果,并与基线RetinaNet-R(Lin等,2017)和CSL(Yang等,2022a)进行了对比,图中(a)是原始图像,(b)是真实框,(c)、(d)和(e)分别是RetinaNet-R、CSL和本文方法的检测结果。

1)如图4中第一行的飞机(PL)场景所示,(1)基线RetinaNet-R对强朝向敏感目标的检测框与真实目标的对齐度不足,且出现了背景误检,存在一定的定位误差;,(2)CSL方法虽提升了角度建模能力,但出现了背景误检问题,检测结果可靠性不足;,(3)而本文方法生成的检测框与真实目标朝向高度一致,无额外误检,展现出精准的方向感知与目标判别能力。

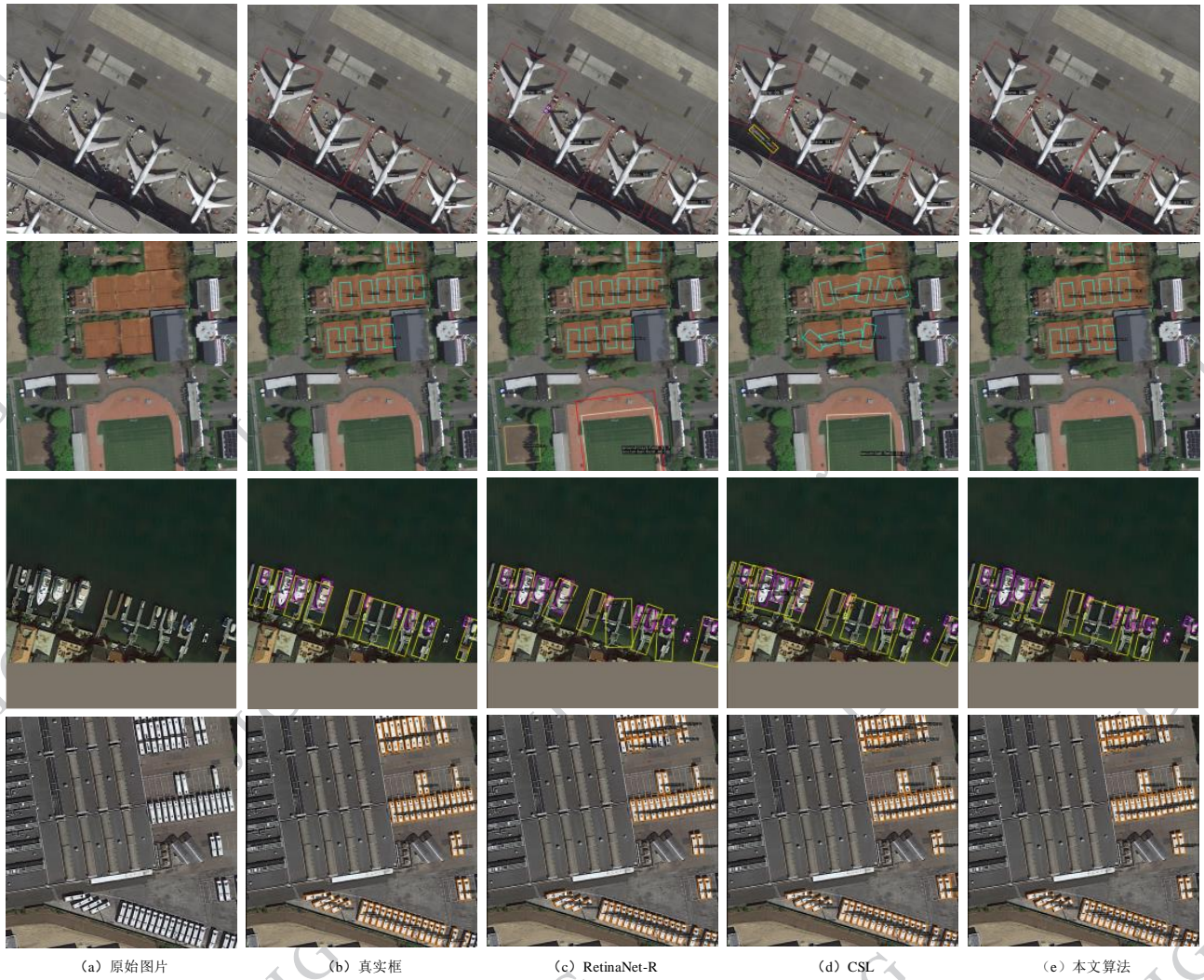
2)图中第二、三、四行中的网球场(TC)、篮球场(BC)、船只(SH)、小型车辆(SV)场景的检测结果进一步表明:(1)基线RetinaNet-R在复杂背景与密集目标场景下易产生误检与漏检,尤其对网球场等规则目标易出现类别误判;,(2)CSL方法对误检有一定抑制效果,但在密集目标场景下仍存在边界混淆、定位偏差问题;,(3)相比之下,本文方法在各场景中均有效降低了漏检误检率,检测框边界清晰、与真实标注高度贴合,充分体现了对背景干扰的强鲁棒性与定位精度优势。

图5则展示了本文方法在HRSC2016测试集上的可视化检测结果,进一步证明了我们所提方法在遥感场景中检测各类目标的有效性和泛化性。

尽管我们的方法取得了一定的检测性能,但也存在失败案例,如图6(a)所示,小尺寸小汽车(黄色圆圈位置)因目标尺寸过小,造成漏检;如图6(b)所示,飞机(橙色圆圈位置)因类别混淆,造成误检。这表明,本文方法对极小尺度目标的有效特征提取不充分,难以捕获足够的判别性信息,进而引发漏检;同时,不同类别目标在外观、轮廓上存在局部相似性,在复杂场景下易引发类别间特征混淆,导致分类分支判别置信度偏移,最终产生误检。后续可通过优化多尺度特征融合策略以增强小目标特征表征能力,并引入类别感知的特征约束机制,提升模型对相似目标的判别精度,从而缓解上述漏检与误检问题。

3 总结

本文提出了置信度加权引导角度修正的遥感旋转目标检测算法,首先,引入AMM,缓解了边界不连续问题;其次,设计了TBDR,有效避免了不同参数间的特征干扰与梯度冲突,为角度预测提供了更纯净的特征基础;再次,提出动态角度置信度加权机制



((a) original image; (b) ground truth; (c) RetinaNet-R; (d) CSL; (e) Our method)

图4 DOTA v1.0测试集中部分场景检测可视化结果对比图

Fig. 4 Detection Visualization Results Comparison of Partial Scenes in the DOTA v1.0 Test Set

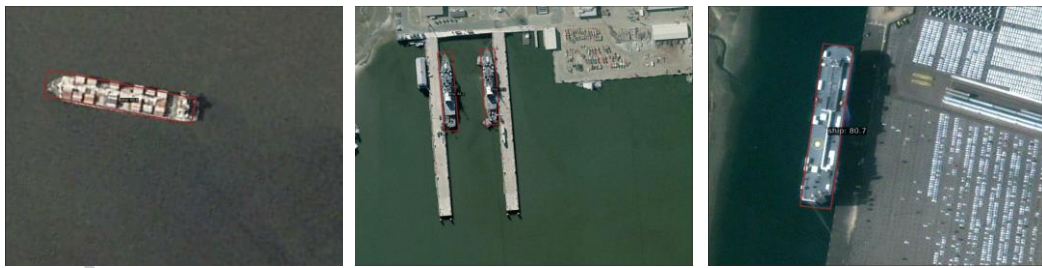


图5 HRSC2016测试集中的检测可视化效果

Fig. 5 Detection Visualization Effects in the HRSC2016 Test Set

(DACW),通过角度预测置信度自适应调控角度修正强度,强化了角度修正的针对性;最后,通过在DOTA v1.0、HRSC2016公开数据集上的实验,证明了所提模块的有效性,以及本文所提方法的优越性和泛化性。然而,本文方法在极小尺度目标与复杂

背景下的检测能力有限,这是因为方法的核心设计更侧重于强方向敏感目标的角度优化,未针对这类场景的特性进行设计。未来可以在现有工作的基础上,进一步融合多尺度特征感知与类别判别约束,优化动态置信度加权机制在复杂场景下的鲁棒性。



(a) (b)

图6 DOTAv1.0测试集中部分场景失败检测可视化结果

Fig. 6 Failed Detection Visualization Results of Partial Scenes in the DOTAv1.0 Test Set

参考文献 (References)

- Chen Y H, Zhang Z, Cao Y, Wang L W, Lin S and Hu H. 2020. Reppoints v2: Verification meets r-regression for object detection. *Advances in Neural Information Processing Systems*, 33: 5621-5631 [DOI:10.48550/arXiv.2007.08508]
- Dai X Y, Chen Y P, Siao B, Chen D D, Liu M C, Yuan L, et al. 2021. Dynamic head: Unifying object detection heads with attentions// *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Nashville, TN, USA. IEEE: 7369-7378 [DOI: 10.48550/arXiv.2106.08322]
- Ding J, Xue N, Long Y, et al. 2019. Learning roi transformer for oriented object detection in aerial images// *Proceedings of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Long Beach, CA, USA. IEEE: 2844-2853 [DOI: 10.1109/CVPR.2019.00296]
- Gao G S, Shang Y Q and Dong Y. 2025. Review of deep learning algorithms for small object detection in optical remote sensing images. *Journal of Image and Graphics*, 30(11):3479-3505 (高广帅, 尚云琦, 董燕. 2025. 光学遥感图像小目标检测深度学习算法综述. *中国图象图形学报*, 30(11): 3479-3505) [DOI: 10.11834/jig.240740]
- Ge Z, Yang Y, Li Q, Wang F and Luo X. 2024. An obb detection algorithm of maintenance components based on YOLOv5-OBBCR. *Information Technology and Control*, 53(1): 71-79 [DOI: 10.5755/j01.itc.53.1.33453]
- Gui S, Song S, Qin R and Tang Y. 2024. Remote sensing object detection in the deep learning era—a review. *Remote Sensing*, 16(2):327 [DOI:10.3390/rs16020327]
- Han J, Ding J, Li J and Xia G S. 2021. Align deep features for oriented object detection. *IEEE Transactions on Geoscience and Remote Sensing*, 60: 1-11 [DOI:10.48550/arXiv.2008.09397]
- Han J M, Ding J, Xue N, and Xia G S. 2021. Re-det: A rotation-equivariant detector for aerial object detection// *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Nashville, TN, USA. IEEE: 2786-2795 [DOI:10.48550/arXiv.2103.07733]
- He J Y and Law K L E. 2024. Deep learning models for rotated object detection in aerial images: survey and performance comparisons. *IEEE Access*, 12: 180436-180457 [DOI:10.1109/ACCESS.2024.3509745]
- Law H, Deng J. 2018. CornerNet: Detecting objects as paired keypoints// *Proceedings of the European Conference on Computer Vision (ECCV)* [EB/O-L]. [2018-09]. <https://arxiv.org/pdf/1808.01244v2.pdf>
- Li Y Y, Huang Q, Xuan P, Jiao L C and Shang R H. 2020. RADet: Refine feature pyramid network and multi-layer attention network for arbitrary-oriented object detection of remote sensing images. *Remote Sensing*, 12(3): 389 [DOI:10.3390/rs12030389]
- Lin T Y, Goyal P, Girshick R, et al. 2017. Focal loss for dense object detection// *Proceedings of the IEEE International Conference on Computer Vision*. Venice, Italy. IEEE: 2999-3007 [DOI: 10.1109/ICCV.2017.324]
- Ma J Q, Shao W Y, Ye H, Wang L, Wang H and Zheng Y B. 2018. Arbitrary-oriented scene text detection via rotation proposals. *IEEE Transactions on Multimedia*, 20(11): 3111-3122 [DOI:10.1109/TMM.2018.2818020]
- Ming Q, Miao L J, Zhou Z Q, and Dong Y P. 2021. CFC-Net: a critical feature capturing network for arbitrary-oriented object detection in remote-sensing images. *IEEE Transactions on Geoscience and Remote Sensing*, 60: 1-14 [DOI:10.1109/TGRS.2021.3095186]
- Ming Q, Zhou Z Q, Miao L J, Zhang H W and Li L H. 2021. Dynamic anchor learning for arbitrary-oriented object detection// *Proceedings of the AAAI Conference on Artificial Intelligence*. 35(3): 2355-2363 [DOI:10.1609/aaai.v35i3.16336]
- Murrugarra-Llerena J, Kirsten L N, Zeni L F, and J-ung C R. 2024. Probabilistic intersection-over-union for training and evaluation of oriented object detectors. *IEEE Transactions on Image Processing*, 33: 671-681 [DOI:10.1109/TIP.2023.3348697]
- Ren S Q, He K M, Girshick R and Sun J. 2016. 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, 李成建, 尹珍臻, 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]
- She H D and Zhao L J. 2024. Rotating target detection network that combines key points and guide vectors. *Journal of Image and Graphics*, 29(02):0533-0544 (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(02):0533-0544) [DOI:10.11834/jig.230207]

- Tian Z, Shen C H, Chen H and He T. 2019. Fcos: Fully convolutional one-stage object detection//P-roceedings of the IEEE/CVF International Conference on Computer Vision. Seoul, Korea (Sout-h). IEEE: 9627-9636 [DOI:10.1109/ICCV.2019.00972]
- Tian Z Z, Zhang H W, Wang K, Liu S Q, Zou Q J, Zhao Z, et al. 2023. Application of an impr-oved centernet in remote sensing images object detection. National Remote Sensing Bulletin. 27 (12):2706-2715 [DOI:10.11834/jrs.20231638]
- 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(01):0282-0296 (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(01): 0282-0296)[DOI:10.11834/jig.230862]
- Wang Y S, Zhang Y, Zhang Y, Zhao L J, Sun X and Guo Z. 2019. SARD: Towards scale-aware r-otated object detection in aerial imagery. IEEE Access. 7: 173855-173865 [DOI:10.1109/ACCE-SS.2019.2956569]
- Wen L, Cheng Y, Fang Y, and Li X Y. 2023. A co-mprehensive survey of oriented object detection in remote sensing images. Expert Systems with Applications, 224: 119960 [DOI:10.1016/j.eswa.2023.119960]
- Xu H, Liu X Y, Xu H N, Ma Y K, Zhu Z J, Yan C G, et al. 2024. Rethinking boundary discont-inuity problem for oriented object detection//Proc-eedings of the IEEE/CVF Conference on Comp-uter Vision and Pattern Recognition. Seattle, WA, USA. IEEE: 17406-17415 [DOI:10.48550/arXiv.2305.10061]
- Xu Y C, Fu M T, Wang Q M, Chen K and Xia G S. 2020. Gliding vertex on the horizontal bou-nding box for multi-oriented object detection. I-EEE transactions on pattern analysis and machine intelligence. 43 (4): 1452-1459 [DOI:10.1109/TPAMI.2020.2974745]
- Yang X, Yan J, Feng Z and He T. 2021b. R³det: R-efined single-stage detector with feature refinem-ent for rotating object//Proceedings of the AAAI conference on artificial intelligence. 35 (4): 3163-3171 [DOI:10.1609/aaai.v35i4.16426]
- Yang X, Yan J, Ming Q, Wang W, Zhang X and Tian Q. 2021a. Rethinking rotated object detection with gaussian wasserstein distance loss//Internat-ional Conference on Machine Learning. ICML: 11830-11841 [DOI:10.48550/arXiv.2106.06072]
- Yang X, Yan J C. 2022a. On the arbitrary-oriented object detection: classification based approaches r-evisited. International Journal of Computer Visio-n. 130 (5): 1340-1365 [DOI:10.48550/arXiv.2003.05597]
- Yang X, Yan J C, Liao W L, Yang X K, Tang J and He T. 2022b. Scrdet++ : Detecting small, clutt-ered and rotated objects via instance-level featu-re denoising and rotation loss smoothing. IEEE Transactions on Pattern Analysis and Machine I-ntelligence. 45 (2): 2384-2399 [DOI:10.1109/TPAMI.2022.3166956]
- Yang X, Yang J R, Yan J C, Zhang Y, Zhang T F and Guo Z. 2019. Scrdet: Towards more robustdetection for small, cluttered and rotated objects//Proceedings of the IEEE/CVF International Conference on Computer Vision. Seoul, Korea (So-uth). IEEE: 8232-8241 [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 b-ounding box for rotated object detection via ku-llback-leibler divergence. Advances in Neural In-formation Processing Systems. 34: 18381-18394 [DOI:10.48550/arXiv.2106.01883]
- Yang X, Zhou Y, Zhang G F, Yang J R, Wang W T, Yan J C, et al. 2023. The kfiou loss for rot-ated object detection. ArXiv Preprint. ArXiv:2201.12558 [DOI:10.48550/arXiv.2201.12558]
- Yang Z, Liu S H, Hu H, Wang L W and Lin S. 2019. Reppoints: Point set representation for objec-t detection//Proceedings of the IEEE/ CVF Intern-ational Conference on Computer Vision. Seoul, Korea (South). IEEE: 9657-9666 [DOI:10.1109/ICCV.2019.00975]
- Yuan X, Cheng G, Li G, Dai W, Yin W X, Feng Y C, et al. 2023. Progress in small object detect-ion for remote sensing images. Journal of Imag-e and Graphics, 28 (6): 1662-1684 (袁翔, 程堃, 李戈, 戴威, 尹文昕, 冯珉超, 等. 2023. 遥感影像小目标检测研究进展. 中国图象图形学报, 28(6): 1662-1684) [DOI:10.11834/jig.221202]
- Yu Y and Da F P. 2023. Phase-shifting coder: Predi-cting accurate orientation in oriented object dete-ction//Proceedings of the IEEE/ CVF Conference on Computer Vision and Pattern Recognition. V-ancouver, BC, Canada. IEEE: 13354-13363 [DOI:10.48550/arXiv.2211.06368]
- Yu Y, Yang X, Li Q Y, Da F P, Dai J F, Qiao Y, et al. 2024. Point2rbox: Combine knowledge fro-m synthetic visual patterns for end-to-end orient-ed object detection with single point superviseon//Proceedings of the IEEE/CVF Conference onComputer Vision and Pattern Recognition. 2024: 16783-16793 [DOI:10.48550/arXiv.2311.14758]
- Zeng Y. 2025. Research on remote sensing image ta-rget detection method based on Transformer. H-arbin: Harbin Institute of Technology. (曾颖.2025.基于Transformer的遥感图像目标检测方法研究[D].哈尔滨:哈尔滨工业大学)
- Zhang D D, Wang C P and Fu Q. 2023.OFCOS: A-n oriented anchor-free detector for ship detectionin remote sensing images. IEEE Geo-science andRemote Sensing Letters. 20: 1-5 [DOI:10.1109/LGRS.2023.3252572]
- Zhang J Y. 2024. Rotated object detection in remotesensing images based on keypoint sets. Guangd-ong:Guangdong University of Technology. (张家源.2024.基于关键点集的遥感图像旋转目标检测. 广东:广东工业大学)
- Zhou Q and Yu C H. 2022. Point RCNN: An angle-free framework for rotated object detection. Re-mote Sensing, 14 (11): 2605 [DOI:10.3390/rs14112605]

Zhou Q, Yu C H, Wang Z B, and Wang F. 2023. D2Q-DETR: Decoupling and dynamic queries for oriented object detection with transformers//I-EEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). Rhodes Is-land Greece. IEEE: 1-5 [DOI:10.17023/bhy5-4x03]

Zhou Z Z and Zhu Y Y. 2024. KLDet: Detecting tiny objects in remote sensing images via kullba-ck-leibler divergence. IEEE Transactions on Ge-oscience and Remote Sensing, 62: 1-16 [DOI: 10.1109/TGRS.2024.3382099]

Xia G S, Bai X, Ding J, Zhu Z, Belongie S, Luo JB, et al. 2018. DOTA: A large-scale dataset for object detection in aerial images// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Salt Lake City, UT, USA. I-EEE: 3974-3983 [DOI:10.48550/arXiv.1711.10398]

作者简介

董燕,女,教授,主要研究方向为人工智能、机器视觉中的模式识别。E-mail:dy@zut.edu.cn

贾季洁,女,硕士研究生,主要研究方向为遥感图像目标检测。E-mail:2024019009@zut.edu.cn。

高广帅,通信作者,男,讲师,主要方向为模式识别、图像处理以及遥感影像智能解译。E-mail:6911@zut.edu.cn。

高君宇,男,副教授,主要方向为计算机视觉与模式识别。E-mail:gjy3035@gmail.com。

李湘云,女,硕士研究生,主要研究方向为遥感图像目标检测。E-mail:2024019010@zut.edu.cn。

李春雷,男,教授,主要研究方向为计算机视觉与模式识别。E-mail:lichunlei1979@zut.edu.cn。