

文章编号:1671-1637(2026)03-0291-12

基于自适应微调的露天矿山低空无人机旋转目标检测

高铭,陈鑫,蒋烁,胡满江,秦洪懋*,边有钢

(湖南大学机械与运载工程学院,湖南长沙 410082)

摘要:为实现低空立体交通运输系统中露天矿山全场景实时视觉感知,提出基于自适应微调的无人机旋转目标检测方法(AFTDet)。针对矿用车辆从无人机视角观测时姿态变化显著的问题,设计自适应空间回归损失函数以优化角度学习并提高高纵横比目标的旋转边界框回归精度,提出微调非极大值抑制算法以利用重叠检测框的空间信息并通过定位参数差异的加权融合提升预测精度,构建包含乘用车、小型挖掘机、装载机和自卸卡车共4540个旋转标注样本的露天矿山旋转目标检测数据集(MineR),最终在公开遥感数据集DOTAv1.0和自建MineR数据集上对AFTDet进行验证。结果表明:在公开遥感数据集DOTAv1.0上,AFTDet取得78.61%的平均计算精度 AP_{50} 和55.45%的平均计算精度 AP_{75} ,较基准模型RTMDet-R-m分别提升0.47%和1.80%;在自建MineR数据集上取得76.25%的 AP_{50} 和44.38%的 AP_{75} ,较基准模型分别提升1.06%和3.62%;消融试验中自适应标签分配策略使 AP_{50} 提升0.99%、 AP_{75} 提升2.50%,微调非极大值抑制使 AP_{75} 进一步提升1.09%,检测速度达 $50.5 \text{ 帧} \cdot \text{s}^{-1}$,参数量维持 2.467×10^7 不变。自适应微调检测方法显著提升了旋转目标的姿态估计性能,尤其改善了大纵横比矿用车辆的检测召回率,在保持实时检测能力的同时为低空立体交通运输系统的无人机视觉感知提供了有效技术支持,促进了露天矿山智能监控与调度系统的发展。

关键词:低空立体交通运输系统;旋转目标检测;自适应微调;露天矿山;视觉感知;无人机检测

中图分类号:U495 **文献标志码:**A **DOI:**10.19818/j.cnki.1671-1637.2026.158

Rotated object detection using low-altitude UAVs for open-pit mines with adaptive fine-tuning

GAO Ming, CHEN Xin, JIANG Shuo, HU Man-jiang, QIN Hong-mao*, BIAN You-gang

(College of Mechanical and Vehicle Engineering, Hunan University, Changsha 410082, Hunan, China)

Abstract: To achieve full-scene real-time visual perception for open-pit mines in low-altitude three-dimensional transportation systems, an adaptive fine-tuning detection method (AFTDet) was proposed for unmanned aerial vehicle (UAV)-based rotated object detection. To address the significant pose variation of mining vehicles observed from UAV perspectives, an adaptive spatial regression loss function was designed to optimize angle learning and improve the rotated bounding box regression accuracy for high-

出版历程:2025-10-10 收稿,2025-12-22 修回,2026-01-23 录用

基金项目:国家自然科学基金项目(52472429)

作者简介:高铭(1991-),男,山东乳山人,副研究员,工学博士,博士后,E-mail:gaoming@hnu.edu.cn。

*通信作者:秦洪懋(1984-),男,江苏盐城人,副教授,博士生导师,工学博士,博士后,E-mail:qinhongmao@vip.sina.com。

引用格式:高铭,陈鑫,蒋烁,等.基于自适应微调的露天矿山低空无人机旋转目标检测[J].交通运输工程学报,2026,26(3):291-302.

Citation: GAO Ming, CHEN Xin, JIANG Shuo, et al. Rotated object detection using low-altitude UAVs for open-pit mines with adaptive fine-tuning[J]. Journal of Traffic and Transportation Engineering, 2026, 26(3): 291-302.

aspect-ratio targets. A fine-tuned non-maximum suppression algorithm was proposed to leverage spatial information from overlapping detection boxes and enhance prediction accuracy through weighted fusion of localization parameter differences. The open-pit mine rotated object detection dataset (MineR) was constructed, comprising 4 540 rotated annotated samples covering passenger vehicles, small excavators, loaders, and dump trucks. AFTDet was validated on both the public remote sensing dataset DOTAv1.0 and the self-built MineR dataset. The results demonstrate that AFTDet achieves 78.61% AP_{50} and 55.45% AP_{75} on DOTAv1.0, representing improvements of 0.47% and 1.80% respectively over the baseline model RTMDet-R-m. On the MineR dataset, it achieves 76.25% AP_{50} and 44.38% AP_{75} , with improvements of 1.06% and 3.62% over the baseline. Ablation studies indicate that the adaptive label assignment strategy improves AP_{50} by 0.99% and AP_{75} by 2.50%, while the fine-tuned non-maximum suppression further improves AP_{75} by 1.09%. The detection speed reaches 50.5 frames \cdot s $^{-1}$ with parameters maintained at 2.467×10^7 . The adaptive fine-tuning detection method significantly enhances pose estimation performance for rotated objects, particularly improving the detection recall of large-aspect-ratio mining vehicles, providing effective technical support for UAV visual perception in low-altitude three-dimensional transportation systems while maintaining real-time detection capabilities. In addition, it promotes the development of intelligent monitoring and scheduling systems for open-pit mines.

Keywords: low-altitude three-dimensional transportation system; rotated object detection; adaptive fine-tuning; open-pit mine; visual perception; UAV detection

Publication history: Received 2025-10-10; Received in revised form 2025-12-22; Accepted 2026-01-23

Funding: National Natural Science Foundation of China (52472429)

* **Corresponding author:** QIN Hong-mao, associate professor, PhD, E-mail: qinhongmao@vip.sina.com.

0 引 言

矿山智能化是中国矿山资源开采领域的发展方向。露天矿山作业范围广,道路环境复杂,同时运行矿用车辆众多,对于矿山无人驾驶的环境感知提出了新的挑战^[1]。目标检测是环境感知的典型任务,目前流行的目标检测算法可分为2类,一类是通过特征提取生成感兴趣区域,再对感兴趣区域进行分类和回归的两阶段方法,代表算法包括更快区域卷积神经网络(Faster Region-based Convolutional Neural Network, Faster R-CNN)^[2]、基于区域的全卷积网络(Region-based Fully Convolutional Network, R-FCN)^[3]等,另一类是只需一次特征提取的单阶段方法,代表算法为单次多框检测器(Single Shot MultiBox Detector, SSD)^[4]、视网膜网络(Retina Network, RetinaNet)^[5]和YOLO系列目标检测算法(You Only Look Once, YOLO)^[6]系列。目前露天矿山领域中,已有一些研究应用于矿用车辆的检测。然而这些研究偏向于部署在车载传感器和路侧传感器^[7]中,存在以下3方面的问题^[8]: ①基于车载和路侧的视觉传感器视场角度狭窄,难

以获取矿区全局信息,不利于矿区的整体调度;②矿区环境沙尘密布^[9],传感器感知性能易受影响,且路侧传感器长期处于扬尘环境,检测任务易受阻断;③矿山运输道路碎石繁多,导致矿用车辆行驶颠簸,车载环境感知数据跳变现象多发。

针对上述问题,本文基于无人机视角开展矿用车辆目标检测方法研究。无人机平台具有机动灵活、视角开阔的特点,在应用于露天矿区时有以下优势:①无人机远离扬尘环境,目标检测任务不会被长时间阻断;②无人机飞行平稳,视野覆盖范围宽广,能免除起伏地形与不平整地面对传感器的影响,从而实现稳定精确的数据采集与环境感知,获取矿区整体信息。将无人机系统引入矿山智能感知体系,不仅能够实现对作业车辆、障碍物及道路状况的实时识别,为无人驾驶系统提供可靠的环境感知信息^[10],也将拓展低空经济在矿业领域的应用场景,使矿山场景下的无人机任务成为“工业服务型低空经济”的重要组成部分。

旋转目标检测^[11]是一类典型的无人机环境感知任务,能在检测出目标位置的同时获取目标的方向信息。相较于水平目标检测,旋转目标检测模型在

检测头处增加了角度参数自由度,最终在图像上输出角度互不相同的矩形边界框^[12]。目前利用无人机拍摄数据集进行水平目标检测的研究逐渐成熟,相较于传统无方向的目标检测方式,利用无人机进行旋转目标检测可获取更多信息,是目标检测的发展趋势^[13-14]。在露天矿山领域,采用旋转目标检测具有以下3个方面的意义:①部分矿用车辆装备位置变化较少,但姿态变化频繁,采用旋转目标检测能有效捕捉其姿态变化情况;②对于纵横比较大,且在画面中姿态呈斜向的车辆,采用旋转目标检测能减少无用背景信息,降低分类任务的难度;③矿用运输车辆普遍以队列方式行进,通过旋转目标检测获取矿用运输车辆的位姿信息,能有效捕捉队列中个体车辆的异常行驶行为,减少安全隐患。

尽管近年来也出现了一些基于无人机影像的旋转目标检测研究,但相关方法大多基于DOTA等高空遥感数据集展开,其目标尺度、姿态变化和背景复杂度均与露天矿山的低空飞行场景存在显著差异。因此,现有无人机检测技术无法直接覆盖矿区车辆在低空俯视条件下呈现出的“大纵横比、姿态快速变化、角度多解”等特点。同时,现有旋转目标检测算法的技术路线也难以完全适应矿山无人机场景。大多数方法在标签匹配阶段仍主要依据交并比(Intersection over Union, IoU)进行空间回归代价计算,而IoU对角度变化不敏感,在大纵横比矿车目标上容易导致角度学习不足,进一步造成高姿态变化样本的召回率下降。

在水平目标检测中,假设目标外轮廓为平滑的曲线,那么定位最优解可以定义为:目标的外轮廓与边界框的所有边界线相切。由于水平目标检测可以通过计算轮廓的最小和最大坐标来确定最小边界框,因此定位最优解通常唯一。然而旋转目标检测加入了角度自由度,允许边界框在平面内自由旋转,以更紧密地贴合目标轮廓。即使要求所有边界线与目标外轮廓相切,边界框仍不唯一,即存在等效最优解。在此情况下,以非极大值抑制(Non-maximum Suppression, NMS)^[15]作为后处理的算法会片面地根据置信度筛选出唯一的边界框,忽略被筛掉的框的有效信息。

为此,本文提出一种基于旋转实时多任务检测器中型(Real-time Multi-task Detector-rotated-medium, RTMDet-R-m)模型^[16]的自适应微调检测器(Adaptive Fine-tuning Detector, AFTDet)。本文主要贡献如下。

(1)在露天矿山场景下采集了包含乘用车、小型挖掘机、装载机和自卸卡车共4540个样本的矿用车辆数据集MineR,并使用旋转矩形进行标注。此数据集在姿态变化和背景复杂度方面更符合露天矿山的低空飞行场景。

(2)本文提出自适应微调检测器AFTDet,使用自适应标签分配策略来优化样本的分配,通过改进空间回归代价函数,使模型对角度的学习更充分,提升大纵横比样本的召回率。

(3)针对边界框有效信息被忽略的问题,本文提出微调非极大值抑制(Fine-tuning Non-maximum Suppression, FTNMS)机制,综合考虑多个检测框的空间参数,以提高预测精度。

1 相关研究概述

1.1 矿区环境感知

目前,矿区环境感知传感器普遍部署于车体或路侧,传感器类别包括激光雷达和可见光相机^[17]。矿区的环境感知对象主要分为2个部分:①道路目标,包括车辆、行人和障碍;②可行驶区域,包括道路边界和坡度。

针对道路目标的感知,王仁炎等^[18]在YOLOv3算法上引入CIoU优化回归损失函数,提升了算法对矿车检测的效果。秦晓辉等^[19]将YOLOv5头部网络解耦,针对特征点,利用不平衡问题提出下采样优化,实现了对矿车和行人的检测。岳伟等^[20]对深度简单在线实时跟踪算法(Deep Simple Online and Realtime Tracking, DeepSORT)^[21]进行优化,通过降低置信度阈值,划分检测框送入不同的流程,实现对矿车和行人的跟踪。针对双目视觉成像特点,阮顺领等^[22]提出Fefol网络,结合焦点损失函数(Focal Loss, FL)、完全交并比损失(Complete Intersection over Union Loss, CIoU Loss)和距离交并比非极大值抑制(Distance Intersection over Union Non-maximum Suppression, DIoU-NMS),实现了可提供矿车避障决策信息的快速障碍检测算法。陈婷等^[23]将多尺度膨胀卷积融合模块和添加了挤压激励块的注意机制残差模块引入渐进式图像去雨网络(Progressive Image Deraining Network, PReNet),实现效果更优的雨天目标检测。

针对可行驶区域的感知,陈龙等^[24]利用激光雷达结合基于八叉树的路沿图构建方法,实现了对帧间路沿的平滑及全局路沿图构建,准确地实现对矿山道路边界的检测。孟德将等^[25]提出了栅格卡尔曼

道路坡度实时检测,采用二维栅格地图、感兴趣矩形区域迭代优化算法和卡尔曼滤波器处理惯性导航和激光点云数据,控制道路坡度检测的误差在较小范围内。

1.2 无人机旋转目标检测

边界不连续是旋转检测中独有的问题,由于旋转边界框角度存在周期性,导致损失不连续。因此,处在边界的角度略微变化会引起较大的损失浮动。Yang等^[26]针对旋转边界框回归的角度边界问题,在损失函数中加入交并比因子,限制损失浮动。现有的大多数角度预测方法都是基于回归方法,环形平滑标签技术(Circular Smooth Label, CSL)^[27]从角度分类的方式出发,通过设定有限的角度分类值将角度的回归问题转化为分类问题。Yu等^[28]提出一种角度相移编码器(Phase Shifting Coder, PSC),通过将角度编码成一个连续的周期相位,再对相位值进行回归,从而解决边界不连续问题。

正负样本的划分影响模型的收敛能力,更好的标签匹配策略往往能够使得网络更好学习到物体的特征,从而提高检测能力。Ming等^[29]发现在标签分配时使用IoU作为代价标准会导致回归不精确,因此提出一种动态锚框分配方法,通过设置衡量因子,筛选出高质量锚框。在EAST^[30]的启发下,Zhu等^[31]利用收缩法划分正负样本的定义空间,根据像素点所处的空间位置进行样本分配。Li等^[32]认为在无锚框的旋转目标检测中,点的质量影响模型的检测结果,进而提出了一种点集评估和分配方案,将点集的分类质量、定位质量、方向质量和点相关质量进行加权求和,从而评估点集的质量。

2 AFTDet算法模型

自适应微调检测器(AFTDet)以RTMDet-R-m为基础改进得到。旋转实时多任务检测器(Real-time Multi-task Detector-rotated, RTMDet-R)是一系列高性能、低延时的单阶段旋转目标检测算法。RTMDet-R在YOLOX^[33]的基础上将跨阶段部分网络(Cross Stage Partial Network, CSPNet)优化为CSPNeXt,并利用大核深度卷积^[34]减少参数数量和动态的软标签分配策略优化分类代价。AFTDet的网络结构与RTMDet-R相同,由CSPNeXt、CSPNeXt路径聚合特征金字塔网络(CSPNeXt Path Aggregation Feature Pyramid Network, CSPNeXtPAFPN)和可分离批归一化检测头(Seperable Batch Normalization

Head, SepBNHead)构成。如表1所示,根据参数数量的差异,RTMDet-R分为4种型号,综合权衡4种RTMDet-R在DOTAv1.0数据集^[35]上的检测精度和参数数量,本文选取检测速度与精度均衡的RTMDet-R-m算法模型作为AFTDet的基准模型。

表1 RTMDet-R各尺寸模型对比

Table 1 Comparison of RTMDet-R models of each size

算法模型	平均精度/%	参数量/ 10^6
RTMDet-R-tiny	75.36	4.88
RTMDet-R-s	76.93	8.86
RTMDet-R-m	78.24	24.67
RTMDet-R-l	78.85	52.27

AFTDet算法模型结构如图1所示,AFTDet在RTMDet-R-m的基础上,将训练过程的动态软标签分配替换为自适应标签分配策略,并在测试时采用微调非极大值抑制进行后处理。

3 自适应标签分配

RTMDet-R-m采用动态软标签分配策略,对中心先验代价 C_{center} 、空间回归代价 C_{reg} 和样本分类代价 C_{cls} 进行了Soft处理,以达到动态匹配效果,总代价 C 的计算式为

$$C = \lambda_1 C_{center} + \lambda_2 C_{reg} + \lambda_3 C_{cls} \quad (1)$$

式中: λ_1 、 λ_2 、 λ_3 分别为中心先验代价权重、空间回归代价权重和分类代价权重。

在计算分类代价时,将预测框与真实框的IoU得分分作为软标签。相比于二进制分类代价,采用软标签能克服较高分类分数的错误边界框得到较低代价的问题,分类代价公式为

$$C_{cls} = E_{CE}(P, Y_{soft}) \times (Y_{soft} - P)^2 \quad (2)$$

式中: E_{CE} 为交叉熵损失函数; P 为模型预测的类别概率; Y_{soft} 为软标签,即预测框与真实框的IoU值;叉乘符号代表矩阵元素依次相乘。

空间回归代价与回归损失一致,将IoU作为衡量标准,并使用对数来增大低质量与高质量匹配之间的差异,空间回归代价为

$$C_{reg} = -\ln(I) \quad (3)$$

式中: I 为交并比。

C_{center} 为中心先验代价,将位置先验引入代价矩阵的计算之中,既能够稳定收敛,也能使匹配的正样本区域更为灵活,中心先验代价公式为

$$C_{center} = \alpha^{|x_{pred} - x_{gt}| - \beta} \quad (4)$$

式中: α 为分类代价权重系数; β 为空间回归代价权

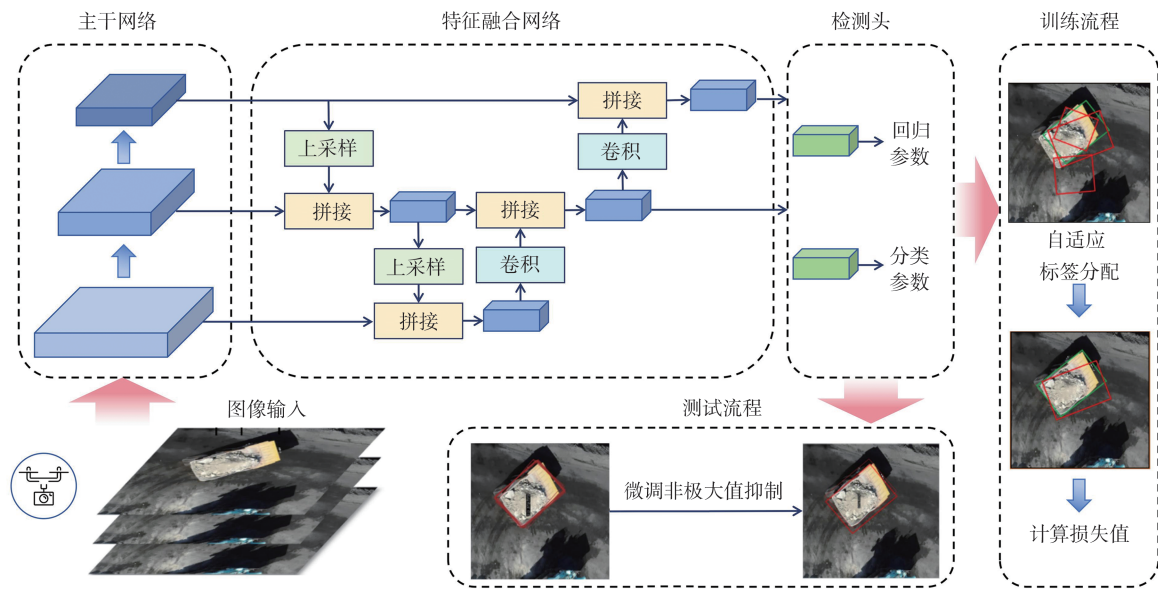


图1 AFTDet的模型结构

Fig. 1 Model structure of AFTDet

重系数; x_{pred} 为预测框中心点 x 坐标; x_{gt} 为真实框中心点 x 坐标。

由于RTMDet-R-m的回归损失标准是真实框和预测框的IoU值,所以采用 $-\ln(I)$ 作为样本分配的空间回归代价能使回归损失最小。然而,这种空间回归代价并没有考虑到不同纵横比样本的IoU值对角度的敏感性不同,仅依据IoU匹配正样本可能会将角度差异大的检测框误视为匹配目标,导致样本分配误筛选出次优的检测框。

标签分配流程示例如图2所示,绿色框为真实框,蓝色框和红色框为参与标签分配的边界框。

如表2所示,图2中红色框的 C_{reg} 小于蓝色框,因此红色框匹配优先级大于蓝色框。但是从图2中可以看出,在与真实框的匹配程度方面,蓝色框的角度和长宽比都优于红色框。上例所示情况会导致模型对角度参数的学习不充分,在不同角度下性能存在差异,从而出现部分角度目标漏检以至于召回率降低。

表2 图2中检测框的空间回归代价值

Table 2 Spatial regression value of the detection frame in Fig. 2

检测框	C_{reg}	$C_{ADPreg} (\beta=0.15)$
红色框	1.23	1.50
蓝色框	1.28	1.29

为解决此问题,本文提出一种自适应标签分配策略,使用新型自适应空间回归(Adaptive Regression, ADPreg)代价函数以克服原有的空间回归代价函数的角度学习不充分问题,ADPreg代价

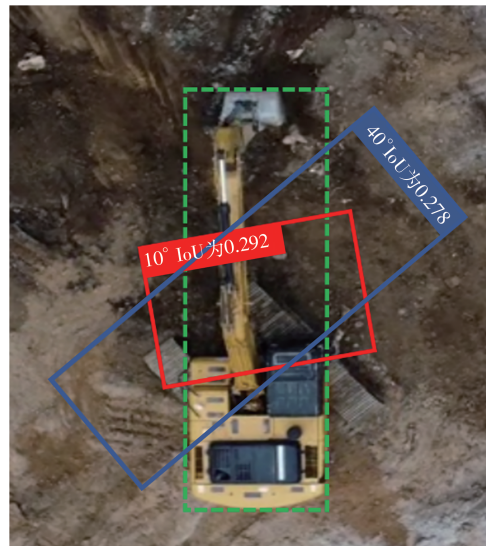


图2 标签分配流程示例

Fig. 2 Example of the label assignment process

C_{ADPreg} 计算式如式(5)所示

$$C_{ADPreg} = -\ln(I) \times [\beta \times f(\theta) + (1 - \beta)] \quad (5)$$

$$f(\theta) = e^{[1 - |\cos(\theta_{gt} - \theta_{pred})|]} \quad (6)$$

式中: θ_{gt} 和 θ_{pred} 分别为真实框和预测框的偏移角度。

式(6)用于获取预测框和真实框的角度差异绝对值信息,其中 $f(\theta)$ 与 $\theta_{gt} - \theta_{pred}$ 的关系曲线如图3所示。由于 $f(\theta)$ 取值范围较大,容易影响 C_{cls} 和 C_{center} 在总代价中的占比,导致训练初期模型过分关注空间回归代价,忽略分类代价和中心先验代价,因此通过超参数 β 调整 $f(\theta)$ 在 C_{ADPreg} 中的权值。 C_{ADPreg} 的效果如表2所示,蓝色框的 C_{ADPreg} 小于红色

框。使用 C_{ADPreG} 替换 C_{reg} , 使模型对角度的学习更充分。

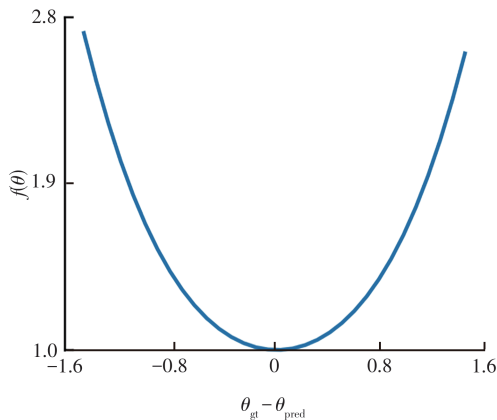


图3 $f(\theta)$ 与 $\theta_{gt} - \theta_{pred}$ 的关系曲线

Fig. 3 Relationship curve between $f(\theta)$ and $\theta_{gt} - \theta_{pred}$

4 微调非极大值抑制(FTNMS)

在目标检测研究中,典型的后处理算法包括非极大值抑制、软非极大值抑制(Soft Non-maximum Suppression, Soft-NMS)^[36]和加权非极大值抑制(Weighted Non-maximum Suppression, Weighted-NMS)^[37],它们被用于移除多余的检测边界框。NMS算法依据置信度和IoU,采用多轮迭代的方式过滤掉重叠的次优检测边界框,从而实现目标的精确定位。Soft-NMS在NMS的基础上,采用降低重叠框置信度的方式替换删除重叠框的方式,缓解重叠目标的漏检问题。然而,上述非极大值抑制算法仅依据置信度对边界框进行筛选,不能有效利用低置信度检测框提供的空间信息。针对上述问题,Weighted-NMS通过对多个边界框的空间参数加权计算,实现从低置信度边角框中获取有效信息。Weighted-NMS在水平目标检测中取得显著效果,然而在旋转目标检测中,对空间参数的加权会存在以下两方面的问题。

(1)受角度不连续性的影响,分布在临界位置两端的边界框容易合成明显错误的结果。如图4(a)所示,蓝色框和紫色框是参与Weighted-NMS的边界框,绿色框为真实框,红色框为合成框。2个边界框效果都接近真实框,但是由于角度的不连续性导致合成框拟合效果较差。

(2)对于类正方形的真实框,略微不准的长宽参数会导致合成后的角度参数偏差较大。如图4(b)所示,蓝色框和紫色框是参与Weighted-NMS的边界框,绿色框为真实框,红色框为合成框。2个边界框

效果都接近真实框,但是由于长宽的预测略微不准确导致合成框拟合效果较差。

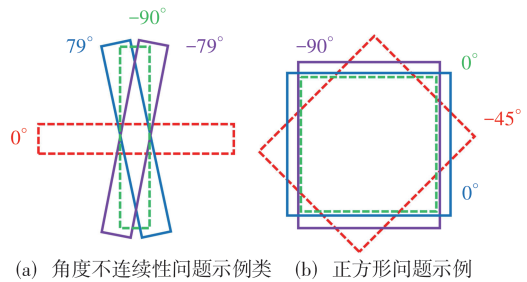


图4 角度不连续性问题示例类 (b) 正方形问题示例

Fig. 4 Angular discontinuity problems and square-like problems

在无人机图像上的旋转目标检测,充分利用重叠框的有效信息尤为重要。从无人机所采用的视角观察,目标在视觉上表现出的是一种具有角度偏斜的俯视效果,这是由无人机的飞行高度、相机的俯仰角度以及目标与无人机之间的相对位置3个因素共同作用的结果。上述情况会导致旋转目标检测算法在尝试捕捉目标的不同视角表现时,采用多个具有特定角度的旋转框来准确描述。如图5所示,黄色框和绿色框为参与NMS的2个边界框,两框的所有边界线都能与目标的外轮廓相切,黄色框对目标外轮廓的拟合显然更准确,而绿色框的角度信息与目标的行驶方向更接近,两框都存在各自的有效信息。



图5 保留框和高效框示例

Fig. 5 Examples of resbox and high-efficiency box

上述例子说明重叠的检测框存在有效信息。为了实现对有效信息的利用,本文提出FTNMS。相比于Weighted-NMS,FTNMS通过对角度微调的分类讨论和定义可靠度,避免角度不连续性的影响和类正方体影响。FTNMS公式为

$$t_m^* = t_m + \sum_{i=1}^{B_m} s_i \frac{\alpha_i}{k+1} (t_m - t_i) \quad (7)$$

$$\theta_m^* = \theta_m + \sum_{i=1}^{B_m} s_i \frac{\alpha_i}{k+1} q \quad (8)$$

$$pq = \begin{cases} \theta_i - \theta_m & |\theta_i - \theta_m| < \frac{\pi}{2} \\ \theta_i - \theta_m - \pi & \theta_i - \theta_m \geq \frac{\pi}{2} \\ \theta_i - \theta_m + \pi & \theta_i - \theta_m \leq -\frac{\pi}{2} \end{cases} \quad (9)$$

$$\alpha_i = 1 - |\sin(\theta_i - \theta_m)| \quad (10)$$

式中: t_i 为第*i*个有效框的中心点坐标; t_m 为原始保留框的中心点坐标; t_m^* 为微调后的主边界框中心点坐标; $B_m = \{t_i, \theta_i, S_i | 1 \leq i \leq k\}$ 为单个保留框对应的有效框集合; S_i 为第*i*个框的置信度; k 为有效框的数量; α_i 为第*i*个框的可靠度; q 为角度修正量; θ_i 为第*i*个有效框的角度; θ_m^* 和 θ_m 分别为微调后的目标方向和原始保留框的角度。

α_i 的设计基于以下考虑。

(1)单调递减性:随角度差增大, α_i 减小,符合“方向差异越大,可信度越低”的直觉。

(2)边界值合理:当 $\theta_i = \theta_m$ 时 $\alpha_i = 1$,完全可信;当 $|\theta_i - \theta_m| = \frac{\pi}{2}$ 时 $\alpha_i = 0$,完全不可信。

(3)周期性适应:正弦函数自然适应角度参数的周期性($0 \sim \pi$)。

(4)平滑调节:函数曲线平滑,避免突变,确保微调过程稳定

图 6 展示了 α_i 随角度差变化的曲线,可见其在 $[0, \frac{\pi}{2}]$ 区间内从 1 平滑递减至 0。仅当高效框与保留框方向基本一致(角度差小于 30°)时, α_i 较大(大于 0.5),微调效果显著;当角度差超过 60° 时, $\alpha_i < 0.15$,微调作用被有效抑制。

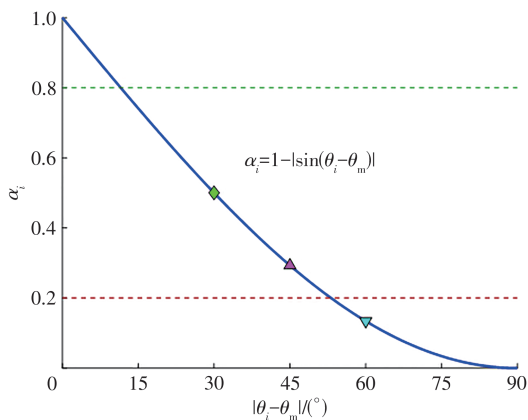


图 6 α_i 随 $|\theta_i - \theta_m|$ 变化的曲线

Fig. 6 Curve of α_i as a function of $|\theta_i - \theta_m|$

NMS 算法依据置信度和 IoU,采用多轮迭代的方式过滤掉重叠的边界框,留下的框称为保留框(Resbox)。而 FTNMS 则在 NMS 的基础上增加了对保留框进行微调的步骤。针对每一个保留框,FTNMS 首先通过设定 IoU 阈值从被过滤掉的框中筛选出 IoU 大于阈值的边界框(以下称为高效框),再计算高效框与保留框的定位参数差值,以置信度和可靠度作为权重,计算每个高效框对保留框的微调量,最后采用求和的方式得到高效框集合 B_m 对保留框的总微调量。式(8)在计算 θ_m^* 时根据角度差值大小分为 3 种情况,如式(9)所示,依靠分类讨论能排除角度不连续性的影响,有效控制微调方向。式(10)计算可靠度 α_i 用于根据角度差自适应调节权重,能降低方向与主框相差较大的有效框对主框的错误影响。

5 试验验证

5.1 矿车数据集

本文试验设计旨在系统验证 AFTDet 在目标场景下的有效性、模块贡献与实用性能。对比方法的选择基于经典性、代表性及与本文模型规模的可比性;消融试验则围绕 2 个核心创新模块(ADPreg 与 FTNMS)展开,从精度、速度、类别敏感性等多个维度评估其独立与协同作用。由于本文聚焦于露天低空这一特殊场景,与部分通用最新方法的直接对比存在领域差异与实现一致性等挑战,未来将在条件允许时补充更广泛的方法对比。

目前为止尚没有公开的俯视视角矿用车辆数据集。为解决露天无人机视角数据缺失难题,本研究基于网络搜索与实地采集融合方式构建露天旋转目标检测数据集 MineR。实地采集部分遵循以下规范,以确保数据集的准确性与多样性。

(1)采集设备与参数:使用大疆 Matrice 300 RTK 无人机进行数据采集,图像分辨率固定为 $5472 \text{ 像素} \times 3648 \text{ 像素}$ 。飞行高度保持在适合低空俯视的区间,以确保矿用车辆具有足够的像素细节并符合实际无人机巡检场景。

(2)时空跨度与环境多样性:采集工作覆盖了不同季节与一天中的若干典型时段,使数据能够包含矿区在自然光照与环境条件变化下的常见视觉特征。采集地点来自多个典型露天矿山,涵盖若干代表性作业区域,以保证场景分布具备一定的多样性。

(3)数据集构成:通过上述方式,MineR 数据集最终包含 386 张图像,按照 7:3 比例划分为训练集、

测试集。数据集包含4个类别,分别是乘用车辆、小型挖掘机、装载机、自卸卡车。该数据集共有4 540个样本,图7展示了4个类别的样本示例,每个类别的正样本统计如图8所示。使用与DOTA数据集相同的多尺度分割方式,获得29 434张图片切片用于训练和测试。

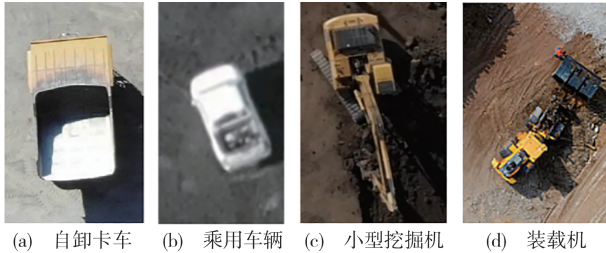


图7 MineR数据集的类别示例

Fig. 7 Category example of the MineR dataset

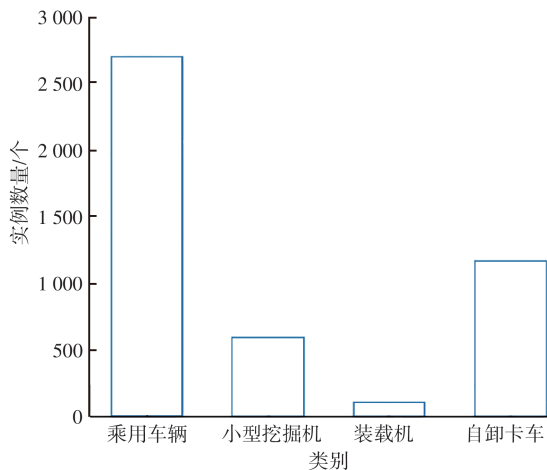


图8 各类别样本数量统计

Fig. 8 Statistics on the number of samples of each category

5.2 试验条件

为保证不同模型检测速度对比的公平性与可复现性,本文对所有实时性测试均在统一的硬件与软件环境下进行。试验平台采用单卡 NVIDIA RTX 4090 GPU(显存 24 GB),CPU 为 Intel Core i9-13900K,内存为 64 GB;软件环境为 Ubuntu 20.04 操作系统, CUDA 11.3, PyTorch 1.13.1, 使用 WarmUp 和余弦退火的学习率衰减,基准学习率为 0.000 25,优化器为 AdamW。试验基于 MMRotate 框架完成模型训练与推理测试。实时性测试采用 MMRotate 框架提供的官方基准测试工具,在不启用 TensorRT、FP16 或其他推理加速手段的条件下进行,测试模式为单卡、单尺度推理,批大小(Batch Size)设置为 1,输入图像尺寸与精度评估阶段保持一致。模型推理延迟通过多次前向推理取平均值的方式获得,并以毫秒(ms)为单位进行统计,从而反

映模型在标准离线推理场景下的实际检测速度。在上述统一测试条件下对不同算法的检测速度进行对比分析,确保本文给出的实时性结果具有良好的可比性和可复现性。试验中还使用了随机水平翻转、随机旋转的数据增强方法。

5.3 矿车数据集试验验证

本文开展消融试验以系统验证 ADPreg 与 FTNMS 两个核心创新模块的贡献。试验设计遵循渐进验证原则:首先评估 FTNMS 单独作用效果,再验证 ADPreg 单独作用效果,最后分析二者协同作用。此外,通过不同 IoU 阈值下的性能分析和各类别召回率影响分析,从多维度揭示模块改进机制,确保消融研究的系统性与说服力。

为了验证本文所提算法在矿用车辆无人机目标检测上的效果,在环境和参数设置保持不变的情况下,以原始的 RTMDet-R-m 网络为基准,在 MineR 数据集上对各个改进模块进行消融试验,结果见表 3。从表 3 可以看出,在 RTMDet-R-m 的基础上将回归标签分配改为 ADPreg 标签分配, AP_{50} 和 AP_{75} 的相对提升率分别为 0.99% 和 2.50%,并且不会增加模型的参数量。同时,把原来的非极大值抑制换成 FTNMS 后,可进一步使 AP_{75} 相对提升 1.09%, AFTDet 相比于 RTMDet-R-m AP_{50} 和 AP_{75} 的相对提升率分别为 1.06% 和 3.62%。

表3 MineR数据集上的消融试验结果

Table 3 Ablation study results on the MineR dataset

算法模型	FTNMS	ADPreg	$AP_{50}/\%$	$AP_{75}/\%$	参数量/ 10^7
RTMDet-R-m	×	×	75.45	42.83	2.467
	√	×	75.52	43.30	2.467
	×	√	76.20	43.90	2.467
AFTDet	√	√	76.25	44.38	2.467

注: AP_{50} 和 AP_{75} 分别为在 IoU 阈值为 0.50 和 0.75 时计算的平均精度,下同。

表 4 展示了本文所提算法对模型检测速度的影响,测试环境为 MMRotate 框架提供的基准测试工具。由于标签分配只在训练阶段进行,所以 ADPreg 不会对检测速度造成影响。FTNMS 在测试阶段进行,通过牺牲较小的检测速度带来可观的检测精度收益。

表4 算法对检测速度的影响

Table 4 Influence of the algorithm on the detection speed

算法模型	延迟/ms
RTMDet-R-m	19.2
AFTDet	19.8

本文所有速度指标均通过标准化测试流程获取。具体步骤为:①加载训练完成的模型权重;②使用MMRotate框架提供的标准测试图像(批大小为1,输入尺寸1 024像素×1 024像素)进行100次前向传播预热;③连续运行1 000次前向传播,记录总时间;④计算帧率时,用1 000除以总耗时秒数,得到的结果即为帧率值。该流程排除了输入/输出、数据预处理等额外开销,专注于模型推理计算时间。

表5为FTNMS和NMS在不同IoU阈值下的效果对比。从表5可以看出,FTNMS提升效果与评估时的IoU阈值呈正相关,证明被NMS筛掉的边界框存在被忽略的有效信息,而微调后的边界框能利用好这些信息,提升对真实框的拟合效果。

表5 FTNMS和NMS在不同IoU阈值下的效果对比

Table 5 Comparison of the effects of FTNMS and NMS at different IoU thresholds

算法模型	后处理 算法	平均计算精度/%			
		AP ₅₀	AP ₆₅	AP ₇₅	AP ₈₅
RTMDet-R-m	NMS	75.45	64.20	42.83	14.16
	FTNMS	75.52	64.54	43.30	14.71

表6展示了ADPreg对各类别召回率的影响,以纵横比较小的自卸卡车类别召回率微小下降为代价,实现了小型挖掘机和装载车2种大纵横比类别召回率的大幅提升,验证了ADPreg有效性。图9为改进前后算法在MineR数据集上的效果,图9(a)~(d)为RTMDet-R-m的检测结果,图9(e)~(h)为AFTDet的检测结果。检测结果a、b、c均存在样本漏检的情况,相比之下,检测结果e、f、g不存在样本漏检;检测结果d中挖掘机样本的边界框偏移目标,检测结果h中边界框对挖掘机样本的定位更准确。

表6 ADPreg对各类别召回率的影响

Table 6 Effect of ADPreg on the recall rate of each category

算法模型	ADPreg	召回率/%			
		乘用车 辆	小型挖 掘机	装载机	自卸 卡车
RTMDet-R-m	×	85.4	91.0	85.3	92.1
	√	87.1	92.8	90.7	91.7

5.4 公开数据集试验验证

本节选取在航空图像目标检测数据集(Dataset for Object Detection in Aerial Images, DOTA)上广泛报告结果且采用相似评估协议(单尺度训练与测试)的经典方法进行对比,包括KLD、Oriented-RepPoint、CSL等。这些方法已在旋转检测领域得到充分验证,能够为本文方法提供有意义的性能参照。随着领域发展,未来将考虑纳入更多最新方法

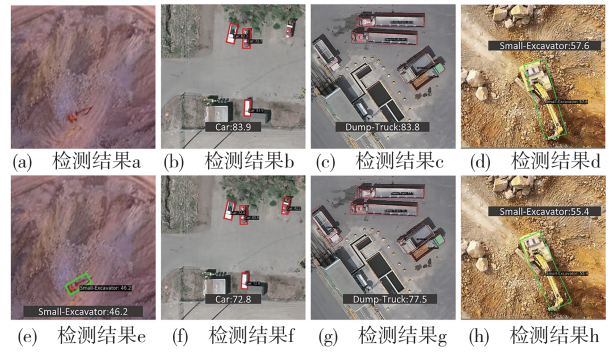


图9 MineR数据集的可视化效果对比

Fig. 9 Comparison of visualizations of the MineR dataset

进行扩展对比。

为了验证所提模型的泛化能力,本研究在公开的DOTAv1.0数据集上进行了试验。DOTA是目前最大且最常用的航空图像数据集之一,共有2 806张图片。DOTAv1.0包含15个类别和涵盖了各种尺寸、纵横比和方向的188 282个实例。

将模型在DOTAv1.0数据集上训练得到的权重文件用于测试,将测试结果传入官方评估服务器后获取AP₅₀和AP₇₅评价指标。表7展示了一些先进算法模型在DOTAv1.0数据集上的检测结果,这些模型均采用单尺度分割的数据集进行训练和测试。由表7可知,AFTDet的精度与更大尺寸的RTMDet-R-l接近,同时检测速度大幅度超过RTMDet-R-l,与其他模型相比,AFTDet在精度和速度方面均表现优异。

表7 先进算法在DOTAv1.0数据集上的对比试验结果

Table 7 Comparative experiment results of advanced algorithms on DOTA v1.0 dataset

算法模型	AP ₅₀ /%	帧率/(帧·s ⁻¹)	参数量/10 ⁷
KLD ^[38]	77.36	25.8	4.190
Oriented-RepPoint ^[39]	77.63	23.5	3.661
CFA ^[40]	76.67	24.1	3.661
CSL ^[41]	76.21	24.5	3.735
Gliding-Vertex ^[42]	75.02	17.4	6.013
PSC ^[29]	71.83	26.1	3.193
S2ANet ^[43]	76.11	17.4	3.624
RTMDet-R-m ^[16]	78.24	49.5	2.467
RTMDet-R-l ^[16]	78.85	29.1	5.227
AFTDet	78.61	45.6	2.467

表8展示了算法模型在DOTAv1.0数据集上的消融试验结果。单独加入ADPreg使AP₅₀和AP₇₅分别提高了0.46%和1.21%,单独加入FTNMS使AP₅₀和AP₇₅分别提高0.06%和0.72%,综合效果使AP₅₀和AP₇₅的相对提升率分别达到0.47%和1.80%。图10为改进前后算法在DOTAv1.0数据

集上的效果,图 10(a)~(d)为 RTMDet-R-m 的检测结果,图 10(e)~(h)为 AFTDet 的检测结果。检测结果 e 相比于检测结果 a,操场类别样本的的定位范围更准确,同时小型汽车漏检情况减少;检测结果 f 相比于检测结果 b,棒球类类别样本的角度预测更准确;检测结果 g、h 相比于检测结果 c、d,港口类别样本的定位更准确。

表 8 算法在 DOTA v1.0 数据集上的消融试验结果

Table 8 Ablation study results of the algorithm on the DOTA v1.0 dataset

算法模型	FTNMS	ADPreg	AP ₅₀ /%	AP ₇₅ /%
RTMDet-R-m	×	×	78.24	54.47
	√	×	78.29	54.86
	×	√	78.60	55.13
AFTDet	√	√	78.61	55.45

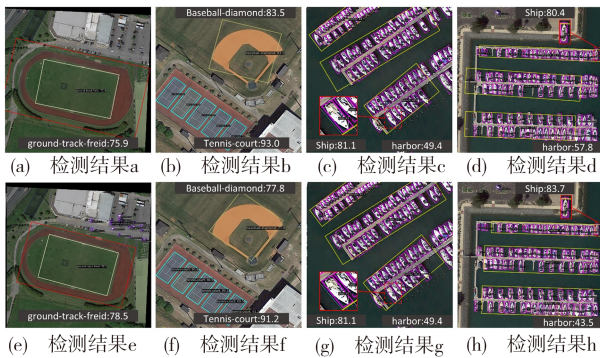


图 10 DOTA v1.0 数据集可视化效果对比

Fig. 10 Comparison of visualizations of the DOTA v1.0 dataset

需要说明的是,AFTDet在DOTA数据集上的性能提升(相较于基准模型,AP₅₀的相对提升率为0.47%)虽显有限,但应结合应用场景差异理解。DOTA作为高空遥感数据集,其目标特性与本文聚焦的矿山低空场景存在显著区别(如目标尺度更小、姿态相对固定)。AFTDet的设计更针对矿车辆的大纵横比与快速旋转特性,因此在MineR数据集上表现出更明显的优势(AP₅₀的相对提升率为1.06%)。未来工作将探索领域自适应方法,以增强算法在多样化场景中的泛化能力。

综合表7、8的分析,AFTDet在DOTA数据集上的性能表现应结合模型复杂度综合评估。与参数量相近的方法(如PSC、CSL)相比,AFTDet在精度上具有明显优势;与精度更高的方法(如RTMDet-R-l)相比,AFTDet在参数量与计算效率上更优。特别值得注意的是,AFTDet在不增加参数量的前提下实现了对基准模型RTMDet-R-m的性能提升,这验证了本文算法改进的有效性。在矿山实际应用中,

这种参数高效的改进方案更具实用价值。

6 结 语

(1)本文针对露天矿山车载与路侧传感器环境感知缺陷,提出基于自适应微调的无人机旋转目标检测方法AFTDet,通过自适应空间回归代价函数、微调非极大值抑制算法及MineR专用数据集三大创新,实现了矿用车辆全场景实时视觉感知。试验验证表明,该方法在DOTA v1.0和MineR数据集上的检测精度均优于RTMDet-R-m基准模型,且保持了良好的实时性,为低空立体交通运输系统的无人机视觉感知提供了有效技术支撑。

(2)研究成果可直接应用于露天矿山智能化监控与调度系统,精准捕捉矿用车辆的位置与姿态信息,减少队列行驶中的安全隐患,同时为城市复杂环境下“低慢小”目标精确感知奠定理论基础,拓展低空经济在矿业与交通领域的应用场景。

(3)本研究存在一定局限:一是MineR数据集样本数量与场景多样性有限,极端天气、复杂光照条件下的数据不足,可能影响模型泛化能力;二是FTNMS算法虽提升精度,但略微增加计算延迟,对高实时性场景适配性有待优化。

(4)未来可围绕以下问题开展研究:联合矿山企业扩充数据集,覆盖更多季节、天气及作业场景,提升模型鲁棒性;探索模型轻量化技术,平衡检测精度与计算效率,满足高实时性应用需求;融合旋转目标检测与语义分割、轨迹预测等任务,构建矿山多维感知系统,为矿用车辆自动驾驶提供更全面的环境感知支持。

参考文献:

Reference:

- [1] 李宏刚,王云鹏,廖亚萍,等. 无人驾驶矿用运输车辆感知及控制方法[J]. 北京航空航天大学学报, 2019, 45(11): 2335-2344.
LI Hong-gang, WANG Yun-peng, LIAO Ya-ping, et al. Perception and control method of driverless mining vehicle [J]. Journal of Beijing University of Aeronautics and Astronautics, 2019, 45(11): 2335-2344.
- [2] REN S Q, HE K M, GIRSHICK R, et al. Faster R-CNN: Towards real-time object detection with region proposal networks [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(6): 1137-1149.
- [3] DAI J F, LI Y, HE K M, et al. R-FCN: Object detection via region-based fully convolutional networks [C] // LEE D D, SUGIYAMA M, LUXBURG U V, et al. Advances in Neural

- Information Processing Systems 29. Red Hook: Curran Associates, 2016: 379-387.
- [4] LIU W, ANGUELOV D, ERHAN D, et al. SSD: Single shot MultiBox detector [C] // LEIBE B, MATAS J, SEBE N, et al. Computer Vision - ECCV 2016. Cham: Springer, 2016: 21-37.
- [5] LIN T Y, GOYAL P, GIRSHICK R, et al. Focal loss for dense object detection [C] // IEEE. 2017 IEEE International Conference on Computer Vision (ICCV). New York: IEEE, 2017: 2999-3007.
- [6] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: Unified, real-time object detection [C] // IEEE. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). New York: IEEE, 2016: 779-788.
- [7] HE Y, JIN L S, GUO B C, et al. Density-based road segmentation algorithm for point cloud collected by roadside LiDAR [J]. Automotive Innovation, 2023, 6(1): 116-130.
- [8] 张 涛,路向阳,李 雷,等. 露天矿山运输无人驾驶关键技术与标准 [J]. 控制与信息技术, 2019(2): 13-19.
ZHANG Tao, LU Xiang-yang, LI Lei, et al. Key technologies and standards of autonomous driving system applied in surface mines [J]. Control and Information Technology, 2019(2): 13-19.
- [9] FU Y, GAO M, XIE G T, et al. Density-aware U-Net for unstructured environment dust segmentation [J]. IEEE Sensors Journal, 2024, 24(6): 8210-8226.
- [10] 彭仲仁,刘晓锋,张立业,等. 无人飞机在交通信息采集中的研究进展和展望 [J]. 交通运输工程学报, 2012, 12(6): 119-126.
PENG Zhong-ren, LIU Xiao-feng, ZHANG Li-ye, et al. Research progress and prospect of UAV applications in transportation information collection [J]. Journal of Traffic and Transportation Engineering, 2012, 12(6): 119-126.
- [11] ZHOU Y, YANG X, ZHANG G F, et al. MMRotate: A rotated object detection benchmark using PyTorch [C] // ACM. Proceedings of the 30th ACM International Conference on Multimedia. New York: ACM, 2022: 7331-7334.
- [12] XIE X X, CHENG G, RAO C F, et al. Oriented object detection via contextual dependence mining and penalty-incentive allocation [J]. IEEE Transactions on Geoscience and Remote Sensing, 2024, 62: 5618010.
- [13] ZHAO J Q, DING Z Y, ZHOU Y, et al. OrientedFormer: An end-to-end transformer-based oriented object detector in remote sensing images [J]. IEEE Transactions on Geoscience and Remote Sensing, 2024, 62: 5640816.
- [14] ZHOU M H, LI T Y, QIAO C F, et al. DMM: Disparity-guided multispectral mamba for oriented object detection in remote sensing [J]. IEEE Transactions on Geoscience and Remote Sensing, 2025, 63: 5404913.
- [15] NEUBECK A, VAN GOOL L. Efficient non-maximum suppression [C] // IEEE. 18th International Conference on Pattern Recognition (ICPR'06). New York: IEEE, 2006: 850-855.
- [16] LYU C Q, ZHANG W W, HUANG H A, et al. RTMDet: An empirical study of designing real-time object detectors [EB/OL]. (2022-10-16). <https://arxiv.org/abs/2212.07784>.
- [17] LI G F, CHI X Y, QU X D. Depth estimation based on monocular camera sensors in autonomous vehicles: A self-supervised learning approach [J]. Automotive Innovation, 2023, 6(2): 268-280.
- [18] 王仁炎,陆占国,胡振涛,等. 基于改进YOLOv3的露天矿卡车目标检测方法 [J]. 矿业研究与开发, 2024, 44(2): 164-169.
WANG Ren-yan, LU Zhan-guo, HU Zhen-tao, et al. Target detection method of truck in open-pit mine based on improved YOLOv3 algorithm [J]. Mining Research and Development, 2024, 44(2): 164-169.
- [19] 秦晓辉,黄启东,常灯祥,等. 基于改进YOLOv5的露天矿山目标检测方法 [J]. 湖南大学学报(自然科学版), 2023, 50(2): 23-30.
QIN Xiao-hui, HUANG Qi-dong, CHANG Deng-xiang, et al. Object detection method in open-pit mine based on improved YOLOv5 [J]. Journal of Hunan University (Natural Sciences), 2023, 50(2): 23-30.
- [20] 岳 伟,林 军,康高强,等. 基于改进DeepSORT的路侧感知方法在露天矿山中的应用 [J]. 控制与信息技术, 2023(3): 89-94.
YUE Wei, LIN Jun, KANG Gao-qiang, et al. Application of roadside perception method based on improved DeepSORT in surface mine [J]. Control and Information Technology, 2023(3): 89-94.
- [21] WOJKE N, BEWLEY A, PAULUS D. Simple online and realtime tracking with a deep association metric [C] // IEEE. 2017 IEEE International Conference on Image Processing (ICIP). New York: IEEE, 2017: 3645-3649.
- [22] 阮顺颂,张回国,顾清华,等. 基于双目视觉的露天矿无人车前障碍检测研究 [J]. 煤炭学报, 2024, 49(增2): 1285-1294.
RUAN Shun-ling, ZHANG Hui-guo, GU Qing-hua, et al. Research on the detection of obstacles in front of unmanned vehicles in opencast mines based on binocular vision [J]. Journal of China Coal Society, 2024, 49(S2): 1285-1294.
- [23] 陈 婷,姚大春,高 涛,等. 基于PReNet和YOLOv4融合的雨天交通目标检测网络 [J]. 交通运输工程学报, 2022, 22(3): 225-237.
CHEN Ting, YAO Da-chun, GAO Tao, et al. A fused network based on PReNet and YOLOv4 for traffic object detection in rainy environment [J]. Journal of Traffic and Transportation Engineering, 2022, 22(3): 225-237.
- [24] 陈 龙,司译文,田 滨,等. 基于3D LiDAR的矿山无人驾驶车行驶边界检测 [J]. 煤炭学报, 2020, 45(6): 2140-2146.
CHEN Long, SI Yi-wen, TIAN Bin, et al. Boundary detection of mine drivable area based on 3D LiDAR [J]. Journal of China Coal Society, 2020, 45(6): 2140-2146.
- [25] 孟德将,田 滨,蔡 峰,等. 面向无人驾驶矿车的露天矿山道路坡度实时检测方法 [J]. 测绘学报, 2021, 50(11): 1628-

1638.
MENG De-jiang, TIAN Bin, CAI Feng, et al. Road slope real-time detection for unmanned truck in surface mine [J]. *Acta Geodaetica et Cartographica Sinica*, 2021, 50(11): 1628-1638.
- [26] YANG X, YANG J R, YAN J C, et al. SCRDet: Towards more robust detection for small, cluttered and rotated objects [C] // IEEE. 2019 IEEE/CVF International Conference on Computer Vision (ICCV). New York: IEEE, 2019: 8231-8240.
- [27] YANG X, YAN J C. On the arbitrary-oriented object detection: Classification based approaches revisited [J]. *International Journal of Computer Vision*, 2022, 130 (5) : 1340-1365.
- [28] YU Y, DA F P. Phase-shifting coder: Predicting accurate orientation in oriented object detection [C] // IEEE. 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). New York: IEEE, 2023: 13354-13363.
- [29] MING Q, ZHOU Z Q, MIAO L J, et al. Dynamic anchor learning for arbitrary-oriented object detection [J]. *Proceedings of the AAAI Conference on Artificial Intelligence*, 2021, 35(3): 2355-2363.
- [30] ZHOU X Y, YAO C, WEN H, et al. EAST: An efficient and accurate scene text detector [C] // IEEE. 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). New York: IEEE, 2017: 2642-2651.
- [31] ZHU Y X, DU J, WU X Q. Adaptive period embedding for representing oriented objects in aerial images [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2020, 58 (10): 7247-7257.
- [32] LI Z H, HOU B, WU Z T, et al. FCOSR: A simple anchor-free rotated detector for aerial object detection [J]. *Remote Sensing*, 2023, 15(23): 5499.
- [33] GE Z, LIU S T, WANG F, et al. YOLOX: Exceeding YOLO series in 2021 [EB/OL]. (2021-08-06). <https://arxiv.org/abs/2107.08430>.
- [34] DING X H, ZHANG X Y, HAN J G, et al. Scaling up your kernels to 31×31 : Revisiting large kernel design in CNNs [C] // IEEE. 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). New York: IEEE, 2022: 11953-11965.
- [35] XIA G S, BAI X, DING J, et al. DOTA: A large-scale dataset for object detection in aerial images [C] // IEEE. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. New York: IEEE, 2018: 3974-3983.
- [36] BODLA N, SINGH B, CHELLAPPA R, et al. Soft-NMS: Improving object detection with one line of code [C] // IEEE. 2017 IEEE International Conference on Computer Vision (ICCV). New York: IEEE, 2017: 5562-5570.
- [37] NING C C, ZHOU H J, SONG Y, et al. Inception single shot MultiBox detector for object detection [C] // IEEE. 2017 IEEE International Conference on Multimedia & Expo Workshops (ICMEW). New York: IEEE, 2017: 549-554.
- [38] YANG X, YANG X J, YANG J R, et al. Learning high-precision bounding box for rotated object detection via Kullback-Leibler divergence [C] // RANZATO M, BEYGEIR H, DAUPHIN Y, et al. *Advances in Neural Information Processing Systems 34*. Red Hook: Curran Associates, 2021: 18381-18394.
- [39] LI W T, CHEN Y J, HU K X, et al. Oriented RepPoints for aerial object detection [C] // IEEE. 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). New York: IEEE, 2022: 1819-1828.
- [40] GUO Z H, LIU C, ZHANG X S, et al. Beyond bounding-box: Convex-hull feature adaptation for oriented and densely packed object detection [C] // IEEE. 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). New York: IEEE, 2021: 8788-8797.
- [41] YANG X, YAN J C. Arbitrary-oriented object detection with circular smooth label [C] // VEDALDI A, BISCHOF H, BROX T, et al. *Computer Vision - ECCV 2020*. Cham: Springer International Publishing, 2020: 677-694.
- [42] XU Y C, FU M T, WANG Q M, et al. Gliding vertex on the horizontal bounding box for multi-oriented object detection [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2021, 43(4): 1452-1459.
- [43] HAN J M, DING J, LI J, et al. Align deep features for oriented object detection [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2022, 60: 5602511.