

基于改进YOLOv5算法的无人机遥感影像车辆检测研究

耿利川^{1,2,3}, 王帮进⁴, 崔云龙⁴, 缙艳方¹

¹许昌学院, 河南 许昌

²智慧中原地理信息技术河南省协同创新中心, 河南 郑州

³时空信息感知与融合技术自然资源部重点实验室, 河南 郑州

⁴浙江省测绘科学技术研究院, 浙江 杭州

收稿日期: 2022年7月31日; 录用日期: 2022年10月14日; 发布日期: 2022年10月28日

摘要

针对无人机遥感影像中的车辆检测问题, 本文对YOLOv5算法中损失函数仅利用边界框长宽比不能反映边界框宽高分别与其置信度的真实差异的缺点, 对损失函数进行了改进, 消除了因为不能反映宽高真实差异引起的收敛过慢问题。利用改进的YOLOv5算法进行了训练和检测, 验证了该算法对于无人机航拍影像中车辆检测的有效性。

关键词

目标检测, YOLOv5, 无人机遥感, 损失函数

Research on Vehicle Detection in UAV Remote Sensing Image Based on Improved YOLOv5

Lichuan Geng^{1,2,3}, Bangjin Wang⁴, Yunlong Cui⁴, Yanfang Gou¹

¹Xuchang University, Xuchang Henan

²Henan Collaborative Innovation Center of Geo-Information Technology for Smart Central Plains, Zhengzhou Henan

³Key Laboratory of Spatiotemporal Information Perception and Fusion Technology, Ministry of Land and Resources, Zhengzhou Henan

⁴Zhejiang Institute of Surveying and Mapping Science and Technology, Hangzhou Zhejiang

Received: Jul. 31st, 2022; accepted: Oct. 14th, 2022; published: Oct. 28th, 2022

文章引用: 耿利川, 王帮进, 崔云龙, 缙艳方. 基于改进YOLOv5算法的无人机遥感影像车辆检测研究[J]. 人工智能与机器人研究, 2022, 11(4): 325-332. DOI: 10.12677/airr.2022.114034

Abstract

Aiming at the problem of vehicle detection in UAV remote sensing images, this paper improves the loss function of YOLOv5 algorithm, which only uses the length-width ratio of the bounding box and cannot reflect the real difference between the width and height of the bounding box and its confidence. It eliminates the problem of slow convergence caused by the failure to reflect the real difference in width and height, and uses the improved YOLOv5 algorithm for training and detection. The effectiveness of the algorithm for vehicle detection in UAV aerial images is verified.

Keywords

Object Detection, YOLOv5, UAV Remote Sensing, Cost Function

Copyright © 2022 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

针对高分辨率无人机遥感影像中的交通车辆进行目标检测, 可用于分析交通拥堵、道路交通违法行为, 收集交通变化规律, 加入交通管制, 促进交通管理更加合理, 对缓解交通拥堵具有重要作用[1]。基于深度学习的车辆检测算法通过卷积神经网络自动提取从原始图像浅层位置信息到高层语义信息的多层次车辆特征, 利用基于神经网络的深度学习算法检测图像中的目标车辆已成为车辆检测领域的研究趋势[2][3]。由于无人机航拍影像的视角特殊性, 针对无人机遥感影像中的车辆检测目前主要面临的问题有: 复杂背景[4]、目标的尺度变化[5]、特殊视角[6][7][8]等问题。

目前基于深度学习的图像目标检测算法模型主要使用 Faster R-CNN [9], Mask R-CNN [10], Retina Net [11], Dynamic R-CNN [12], PAA [13] and ATSS [14], YOLO 系列[15][16][17][18]等算法框架。邢宇驰等[19]基于 YOLOv5 的遥感图像目标检测在公开的 10 类地理空间物体数据集进行了检测实验, 以评估所提出模型的目标检测性能。李珂泉等[20]按照是否存在显式的区域建议和是否定义先验锚框两种分类标准对已有的目标检测算法分类, 对代表性的目标检测算法在公开数据集中的表现进行了比较和分析。马跃等[21]在分段反卷积改进 SSD 的目标检测算法中提出了分段反卷积结构降低了噪声信息加入, 又采用了新的融合结构对高层特征图和低层特征图进行融合, 丰富了检测特征层的信息, 加入了不同长宽比例的候选框和多尺度特征图对小占比目标增强检测, 减小了小占比目标的误检和漏检。杨锦帆等[22]对比了各类常用单阶段车辆检测算法, 列举其改进措施以及在车辆检测方面存在的问题, 介绍了车辆检测相关数据集, 对现阶段车辆检测中亟待解决的问题与难点进行了分析, 提出了车辆检测未来的研究方向。冯加明等[23]提出通过删减部分残差以降低卷积层通道数并通过 K-均值聚类与手动调节相结合的方式提高检测的精度。汪昱东等[24]通过在检测网络中加入物浓度判断模块来提高网络的适应性和鲁棒性。YOLOv5 算法在 YOLOv4 的基础上提出, 其速度与精度都得到了极大的性能提升, 被广泛应用于遥感影像目标检测。本文针对 YOLOv5 算法中损失函数仅使用长宽比并不能反映边界框宽高分别与置信度的真实差异的缺点, 对损失函数进行了改进, 消除了损失函数因为不能反映宽高真实差异引起的收敛过慢问题, 提高了算法的检测精度。

2. 基本概念

2.1. YOLO 系列算法

YOLO 算法首先通过卷积神经网络将目标图像划分为 $S \times S$ 个方格的特征图像，计算机可以预测到方格中所落入的目标的边界框、置信度和所属类别的概率图，如果目标对象的中心落在正方形中，则会在正方形周围预测 N 个边界框，并计算每个边界框所包围的目标的置信度。最后通过非极大值抑制算法(NMS)消除单个目标冗余的边界框并得到检测结果。YOLOv2 是在 YOLOv1 的一些基础上做了优化，并提供了一个全新的联合训练算法。YOLOv2 算法将骨干网络由 224×224 换为 448×448 的 Darknet19，采用了全卷积网络结构 Conv + BatchNorm，并且在 YOLOv1 的基础网络上去掉了一个全卷积层，利用随机梯度的下降从而能够在 COCO 数据集上获得良好的检测结果。YOLOv3 算法使用了骨干网络 Darknet53、采用了多尺度预测、跨尺度特征融合等。YOLOv4 算法使用了 Mish 激活函数，引入了 CIoU，进一步提升了检测效果。YOLOv5 算法几乎与 YOLOv4 算法同时提出，在性能上略弱于 YOLOv4，但在灵活性和速度上比 YOLOv4 更好。YOLOv5 在 Backbone 部分没太大变化。在 Neck 部分的变化相对较大，首先是将 SPP 换成了 SPPF，在 PAN 结构中加入了 CSP，加强了网络特征融合的能力，在损失函数上 YOLOv5 采用了 CIoU。

2.2. IOU 介绍

IOU [25]的提出是为了衡量预测框与真实框的关系。IOU 即计算预测框与真实框的交并比，IOU 损失函数将四个点构成的边界框，作为整体进行回归。设置(xt, xb, xl, xr)进行计算，反应真实框和预测框的关系。 $IOU = 1 - IOU$ 。IOU 具有尺度不变性，但存在的缺点是当 IOU 等于 0 时，即预测框与真实框没有交集时，无法进行梯度回传，从而无法判断预测框与真实框的距离。另外 IOU 损失无法判断两个框的相交方式。GIoU [26]针对该问题做出了相应的改进，GIoU 对于任意的两个 A 、 B 框，首先确定一个能够包含它们的最小方框 C ，计算 $C - A \cup B$ ，然后计算该值与 C 面积的比值，再用 A 、 B 的 IOU 减去该比值得到 GIoU，GIoU 损失函数如公式(1)所示，

$$L_{\text{GIoU}} = 1 - \text{IOU} + \frac{C - A \cup B}{C} \quad (1)$$

虽然 GIoU 克服了 IOU 因为预测框和目标框没有交集时，梯度无法回传的问题，但当 A 包含 B 时， $|C - A \cup B|$ 将为零，反之亦然，在这种情况下，GIoU 损失减小为借据损失。因此，GIoU 损失的收敛率仍然很慢。针对该缺点，Zheng Z 等[27]提出了 DIOU，DIOU 的损失函数如公式(2)所示，

$$L_{\text{DIOU}} = 1 - \text{IOU} + \frac{\rho(\mathbf{b}, \mathbf{b}^{gt})}{C^2} \quad (2)$$

其中， \mathbf{b} 、 \mathbf{b}^{gt} 分别代表了预测框与真实框的中心点， ρ 代表两个中心点之间的欧式距离， C 代表能够同时包含预测框和真实框的最小闭包区域的对角线距离。DIOU 损失函数惩罚项可以直接最小化预测框和真实框中心点间的距离，因此比 GIoU 损失函数收敛快。CIoU 在 DIOU 基础上加了一个考虑预测框纵横比拟合真实框的纵横比的影响因子，进一步提升了算法的收敛速度，CIoU 的损失函数如公式(2)所示。

$$\text{CIoU} = \text{IOU} - \frac{\rho^2(\mathbf{b}, \mathbf{b}^{gt})}{c^2} - av \quad (3)$$

其中， $\rho^2(\mathbf{b}, \mathbf{b}^{gt})$ 表示真实框与预测框中心的欧式距离， c 表示真实框与预测框最小闭包区域对角线距离，

$$v = \frac{4}{\pi^2} \left(\arctan \frac{w^{gt}}{h^{gt}} - \arctan \frac{w}{h} \right)^2, \quad a = \frac{v}{1 - \text{IOU} + v}.$$

3. YOLOv5 算法的改进

CIOU 损失虽然考虑了边界框回归的重叠面积、中心点距离、纵横比。但是通过其公式中的 v 反映的是纵横比的差异，而不是宽高分别与其置信度的真实差异，所以会阻碍模型有效的优化相似性。针对这一问题，Zhang Y F 等[28]在 CIOU 的基础上将纵横比拆开，提出了 EIOU 损失函数，定义如公式(4)所示：

$$L_{\text{EIOU}} = L_{\text{IOU}} + L_{\text{dis}} + L_{\text{asp}} = 1 - \text{IOU} + \frac{\rho^2(\mathbf{b}, \mathbf{b}^{gt})}{c^2} + \frac{\rho^2(w, w^{gt})}{C_w^2} + \frac{\rho^2(h, h^{gt})}{C_h^2} \quad (4)$$

该损失函数包含三个部分：重叠损失，中心距离损失，宽高损失，前两部分延续 CIOU 中的方法，但是宽高损失直接使目标框与锚框的宽度和高度之差最小，使得收敛速度更快。其中 IOU 表示预测框和真实框相交区域面积和合并区域面积的比值， C_w 和 C_h 是包含预测框和真实框的最小闭包区域的宽度和高度， \mathbf{b} 与 \mathbf{b}^{gt} 分别表示预测框和真实框中心点的位置， $\rho^2(\mathbf{b}, \mathbf{b}^{gt})$ 表示预测框与真实框中心的欧式距离， w 与 w^{gt} 表示预测框与真实框的宽度， $\rho^2(w, w^{gt})$ 表示预测框与真实框宽度的欧式距离， h 与 h^{gt} 表示预测框与真实框的高度， $\rho^2(h, h^{gt})$ 表示预测框与真实框高度的欧式距离，将纵横比的损失项拆分成预测的宽高分别与最小外接框宽高的差值，加速了收敛提高了回归精度。本文对 YOLOv5 算法中的损失函数进行了改进，利用 EIOU 损失函数，替换原始 YOLOv5 算法中的 CIOU 损失函数，改进后的算法记为 YOLOv5-EIOU，改进前记为 YOLOv5-CIOU，并在无人机遥感影像数据集上进行了车辆检测实验验证。

4. 实验结果与分析

4.1. 实验环境配置

本文实验计算机配置 32 GB 的内存，NVIDIA GeForce GTX 2080Ti 显卡，CPU 主频 3.6 GHz。cuda 版本为 10.2，使用开发环境软件为 Pycharm，语言为 Python3.8，深度学习框架为 Facebook 开源的 pytorch。

4.2. 数据集准备与训练

本文实验使用数据集为 UAV-ROD，是国科大提供的面向无人机的汽车数据集，UAV-ROD 由 1577 张图像和 30,090 个汽车类别实例组成，训练集和测试集中的图像数分别设为 1150 和 427，实验使用其 coco 格式标注成果。实验对 YOLOv5 算法中的 s、m、l 三种模型分别进行了损失函数改进前后的实验验证，训练过程中 s、m 模型，设置 epoch 为 150，batch 大小设为 10，l 模型的 batch 大小设为 4。

4.3. 训练结果验证

模型训练完成后得出精确度(Precision)和召回率(Recall)两个数值指标，为综合评估目标检测的性能，通过平均精度均值(Mean Average Precision, mAP)来进行评价模型的优劣。以召回率 Recall 为横坐标轴，精确度 Precision 为纵坐标轴可以得到 PR 曲线，其中曲线下的面积是平均精确度 AP 值，目标检测模型中每种目标均可计算出一个 AP 值，对 AP 值求平均则可以得到模型的 mAP 值。mAP 是介于 0 到 1 之间的一个数值，该数值越接近于 1，就表示模型的性能越好。本文实验分别训练了 YOLOv5-CIOU 算法的 s、m、l 模型和 YOLOv5-EIOU 算法的 s、m、l 模型来检测目标，通过对模型的验证可以分别得到 3 类模型的 PR 曲线，计算将 IOU 设为 0.5 时，无人机遥感影像中汽车目标的 AP 值，由于本数据集只有一类目标，所以 AP 值即为 mAP 值，mAP 值越接近 1 表示所训练的模型的性能越好。实验同时给出了不同 IOU 阈值(从 0.5 到 0.95，步长 0.05)上的平均 mAP。结果如表 1 所示。

由表 1 可以看出经过将 YOLOv5 算法损失函数由 CIOU 改进为 EIOU 后 mAP@0.5 值都有不同程度的提升，其中 YOLOv5s 的 mAP@0.5 值由 0.971 提升到 0.973 提升了 0.2%，YOLOv5m 的 mAP@0.5 值

由 0.968 提升到 0.969 提升了 0.1%，YOLOv5l 的 $mAP@0.5$ 值由 0.972 提升到 0.973 提升了 0.3%，为 3 类模型中 mAP 值提升最多的算法。 $mAP@.5:.95$ 值在算法改进后同样有不同程度的提高，其中 YOLOv5m 模型在改进后提升了 1.5%，为三个模型中提升最多的模型。图 1 给出了各模型 PR 曲线的对比结果。由 PR 曲线下方面积对比也可看出改进后的算法 AP 值较原算法均有提升。

Table 1. Experience result comparison

表 1. 实验结果对比

模型	YO- LOv5s-CIOU	YO- LOv5s-EIOU	YO- LOv5m-CIOU	YO- LOv5m-EIOU	YO- LOv5l-CIOU	YO- LOv5l-EIOU
$mAP@0.5$ (%)	97.1	97.3	96.8	96.9	97.2	97.5
$mAP@.5:.95$ (%)	61.5	61.9	68.8	60.3	61.4	61.8

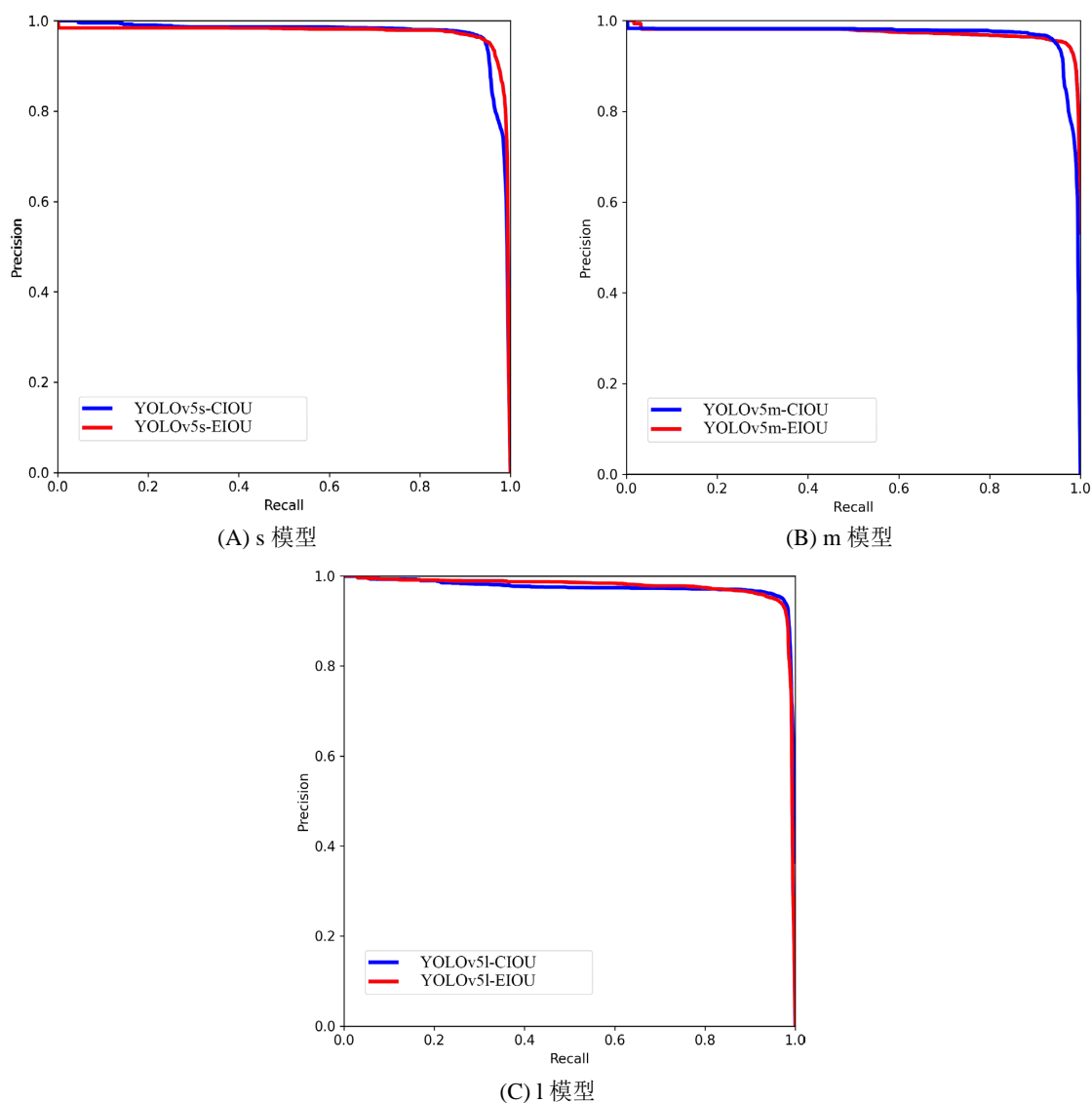


Figure 1. P-R Curve comparison of YOLOv5s, m, l-CIOU and YOLOv5s, m, l-EIOU model

图 1. YOLOv5s、m、l-CIOU 与 YOLOv5s、m、l-EIOU 三种模型 P-R 曲线对比

图 2 给出了三种模型部分检测结果对比,由对比结果可以看出 YOLOv5 算法各模型针对无人机遥感影像中的汽车车辆检测问题,均有很高的检测准确率。仅在图 2(C)中针对第 7 张影像的检测, YOLOv5I-CIOU 模型存在一个错检车辆,而 YOLOv5I-EIOU 模型没有错检该车辆。比较原始模型与损失函数优化后的 YOLOv5 模型之间的均值平均精度,发现优化后的 YOLOv5 算法表现出了较高的性能,究其根本,在于 EIOU 损失函数能够更好地反映目标框与锚框宽度和高度之差,加速了收敛提高了回归精度。



(A) s 模型车辆检测结果对比



(B) m 模型车辆检测结果对比



(C) l 模型车辆检测结果对比

Figure 2. Vehicle detection result comparison YOLOv5s, m, l-CIOU and YOLOv5s, m, l-EIOU model
图 2. YOLOv5s、m、l-CIOU 与 YOLOv5s、m、l-EIOU 车辆检测结果对比

5. 结论

针对无人机遥感影像中车辆目标的识别问题,本文对原始 YOLOv5 算法中的损失函数进行了改进,并分别训练了 YOLOv5 算法三种模型 YOLOv5s、YOLOv5m、YOLOv5l,对改进前后的算法利用无人机遥感影像数据集进行了训练和验证,验证了改进后的算法在车辆检测中的优越性。

基金项目

河南省科技攻关(212102310419), 智慧中原地理信息技术河南省协同创新中心、时空信息感知与融合技术自然资源部重点实验室联合基金(212109), 许昌学院科研反哺教学项目(2022FB004)。

参考文献

- [1] 王利鑫. 基于改进的 Faster R-CNN 的无人机影像车辆检测[D]: [硕士学位论文]. 湘潭: 湖南科技大学, 2019.
- [2] 董丽君, 曾志高, 易胜秋. 基于 YOLOv5 的遥感图像目标检测[J]. 湖南工业大学学报, 2022, 36(3): 44-50.
- [3] 谭路文, 哈斯巴干, 陈超民, 等. 基于深度学习的无人机遥感影像车辆检测[J]. 红外, 2022, 43(5): 41-48.
- [4] Wang, C., *et al.* (2019) Multiscale Visual Attention Networks for Object Detection in VHR Remote Sensing Images. *IEEE Geoscience and Remote Sensing Letters*, **16**, 310-314. <https://doi.org/10.1109/LGRS.2018.2872355>
- [5] Zhang, Y., *et al.* (2019) Hierarchical and Robust Convolutional Neural Network for Very High-Resolution Remote Sensing Object Detection. *IEEE Transactions on Geoscience and Remote Sensing*, **57**, 5535-5548. <https://doi.org/10.1109/TGRS.2019.2900302>
- [6] Cheng, G., *et al.* (2016) Learning Rotation-Invariant Convolutional Neural Networks for Object Detection in VHR Optical Remote Sensing Images. *IEEE Transactions on Geoscience and Remote Sensing*, **54**, 7405-7415. <https://doi.org/10.1109/TGRS.2016.2601622>
- [7] Li, K., *et al.* (2017) Rotation-Insensitive and Context-Augmented Object Detection in Remote Sensing Images. *IEEE Transactions on Geoscience and Remote Sensing*, **56**, 2337-2348. <https://doi.org/10.1109/TGRS.2017.2778300>
- [8] Wu, X., *et al.* (2019) ORSIm Detector: A Novel Object Detection Framework in Optical Remote Sensing Imagery Using Spatial-Frequency Channel Features. *IEEE Transactions on Geoscience and Remote Sensing*, **57**, 5146-5158. <https://doi.org/10.1109/TGRS.2019.2897139>
- [9] Ren, S., He, K., Girshick, R., *et al.* (2017) Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, **39**, 1137-1149. <https://doi.org/10.1109/TPAMI.2016.2577031>
- [10] He, K.M., Gkioxari, G., Dollar, P. and Girshick, R. (2017) Mask R-CNN. *IEEE International Conference on Computer Vision*, Venice, 22-29 October 2017, 2980-2988. <https://doi.org/10.1109/ICCV.2017.322>
- [11] Lin, T.Y., Goyal, P., Girshick, R., *et al.* (2017) Focal Loss for Dense Object Detection. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, **99**, 999-3007. <https://doi.org/10.1109/ICCV.2017.324>
- [12] Zhang, H.K., Chang, H., Ma, B.P., Wang, N.Y. and Chen, X.L. (2020) Dynamic R-CNN: Towards High Quality Object Detection via Dynamic Training. *European Conference on Computer Vision*, Glasgow, 23-28 August 2020, 260-275. https://doi.org/10.1007/978-3-030-58555-6_16
- [13] Kim, K.-J. and Lee, H.S. (2020) Probabilistic Anchor Assignment with IoU Prediction for Object Detection. 2020 *European Conference on Computer Vision, ECCV*, Glasgow, 23-28 August 2020, 1-17. https://doi.org/10.1007/978-3-030-58595-2_22
- [14] Zhang, S., Chi, C., Yao, *et al.* (2020) Bridging the Gap between Anchor-Based and Anchor-Free Detection via Adaptive Training Sample Selection. *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Seattle, 13-19 June 2020, 9759-9768. <https://doi.org/10.1109/CVPR42600.2020.00978>
- [15] Redmon, J., Divvala, S.K., Ross, B., *et al.* (2015) You Only Look Once: Unified, Real-Time Object Detection. 2016 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, 27-30 June 2016, 779-788. <https://doi.org/10.1109/CVPR.2016.91>
- [16] Redmon, J. and Farhadi, A. (2017) YOLO9000: Better, Faster, Stronger. *IEEE Conference on Computer Vision & Pattern Recognition*, Honolulu, 21-26 July 2017, 6517-6525. <https://doi.org/10.1109/CVPR.2017.690>
- [17] Redmon, J. and Farhadi, A. (2018) YOLOv3: An Incremental Improvement.
- [18] Bochkovskiy, A., Wang, C.Y. and Liao, H. (2020) YOLOv4: Optimal Speed and Accuracy of Object Detection.
- [19] 邢宇驰, 李大军, 叶茂. 基于 YOLOv5 的遥感图像目标检测[J]. 江西科学, 2021, 39(4): 725-732.
- [20] 李柯泉, 陈燕, 刘佳晨, 等. 基于深度学习的目标检测算法综述[J]. 计算机工程, 2022, 48(7): 1-12.
- [21] 马跃, 赵志浩, 尹震宇, 等. 分段反卷积改进 SSD 的目标检测算法[J]. 小型微型计算机系统, 2021, 42(7): 1415-1420.
- [22] 李科岑, 王晓强, 林浩, 等. 深度学习中的单阶段小目标检测方法综述[J]. 计算机科学与探索, 2022, 16(1): 41-58.

- [23] 冯加明, 储茂祥, 杨永辉, 等. 改进 YOLOv3 算法的车辆信息检测[J]. 重庆大学学报, 2021, 45(12): 71-79.
- [24] 汪昱东, 郭继昌, 王天保. 一种改进的雾天图像行人和车辆检测算法[J]. 西安电子科技大学学报, 2020, 47(4): 70-77.
- [25] Yu, J., Jiang, Y., Wang, Z., *et al.* (2016) UnitBox: An Advanced Object Detection Network. *Proceedings of the 24th ACM International Conference on Multimedia*, New York, October 2016, 516-520. <https://doi.org/10.1145/2964284.2967274>
- [26] Rezatofghi, H., Tsoi, N., Gwak, J.Y., *et al.* (2019) Generalized Intersection over Union: A Metric and a Loss for Bounding Box Regression. 2019 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Long Beach, 15-20 June 2019, 658-666. <https://doi.org/10.1109/CVPR.2019.00075>
- [27] Zheng, Z., Wang, P., Liu, W., *et al.* (2019) Distance-IoU Loss: Faster and Better Learning for Bounding Box Regression. *Proceedings of the AAAI Conference on Artificial Intelligence*, **34**, 12993-13000. <https://doi.org/10.1609/aaai.v34i07.6999>
- [28] Zhang, Y.F., Ren, W., Zhang, Z., *et al.* (2021) Focal and Efficient IOU Loss for Accurate Bounding Box Regression. *Neurocomputing*, **506**, 146-157. <https://doi.org/10.1016/j.neucom.2022.07.042>