

基于改进YOLOv8s的军用飞机目标检测算法

王广川, 赵寿为

上海工程技术大学数理与统计学院, 上海

收稿日期: 2024年4月30日; 录用日期: 2024年7月25日; 发布日期: 2024年8月1日

摘要

基于YOLOv8s目标检测算法, 提出了一种改进的军用飞机遥感图像目标检测算法。首先, 引入Mixup数据增强方法; 其次, 修改网络结构, 减少主干网络最后一个输出特征图的通道数为256; 然后, 在主干网络部分融合改进的SimAM注意力机制A; 最后, 使用改进的NWD损失作为位置损失函数。改进的算法在MAR20和NWPU VHR-10数据集上的mAP50分别比YOLOv8s提高了4.3%和2.2%, 验证了改进算法的有效性。

关键词

NWD损失, SimAM注意力机制, YOLOv8s, 遥感目标检测

Military Aircraft Target Detection Algorithm Based on Improved YOLOv8s

Guangchuan Wang, Shouwei Zhao

School of Mathematics, Physics and Statistics, Shanghai University of Engineering Science, Shanghai

Received: Apr. 30th, 2024; accepted: Jul. 25th, 2024; published: Aug. 1st, 2024

Abstract

Based on the YOLOv8s object detection algorithm, an improved algorithm for object detection in military aircraft remote sensing images is proposed. Firstly, the Mixup data augmentation method is introduced; secondly, the network structure is modified to reduce the number of channels in the

last output feature map of the backbone network to 256; then, an improved SimAM attention mechanism is integrated into the backbone network; finally, an improved NWD loss is used as the position loss function. The improved algorithm has increased the mAP50 on the MAR20 and NWPU VHR-10 datasets by 4.3% and 2.2% respectively compared to YOLOv8s, verifying the effectiveness of the improved algorithm.

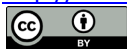
Keywords

NWD Loss, SimAM Attention Mechanism, YOLOv8s, Remote Sensing Object Detection

Copyright © 2024 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

随着遥感卫星技术的发展, 遥感图像的空间分辨率不断提升, 从海量遥感图像中自动提取有价值的目标信息逐渐成为研究热点。军用飞机作为有战略意义的目标, 对其精准定位和分类, 在军事领域具有重要作用[1]。由于遥感图像尺度变化大, 背景复杂等因素, 造成了遥感图像目标检测准确率低、漏检等问题[2]。为提高遥感图像目标检测的准确率, 不断有学者对此进行研究[3][4]。章程军等[5]在 YOLOv5s 的基础上通过引入 BiFPN [6]特征融合方法, 提高多尺度信息融合效率, 以提高遥感图像目标检测准确率。谢俊章等[7]基于 YOLOv4 目标检测算法, 设计轻量化的骨干网络, 以提高遥感图像目标检测准确率。本文基于 YOLOv8s 目标检测算法, 结合遥感图像的特点, 为 YOLOv8s 目标检测算法引入 Mixup [8]数据增强方法, 以提高数据集多样性; 其次, 使用改进的 NWD [9]损失函数替代原算法中的 CIoU 边框损失函数, 以更精确的计算边框损失; 接着引入改进的 SimAM [10]模块, 以提高主干网络的特征图质量; 最后, 考虑到遥感图像的目标较小, 减少了骨干网络第三个输出特征图的通道数, 以减少神经网络深层语义信息的参数量。基于以上改进, 构建了改进的 YOLOv8s 算法, 并在 MAR20 [11]军事飞机检测数据集和 NWPU VHR-10 数据集上做了实验, 实验结果验证了改进算法的有效性。

2. 改进的 YOLOv8s 算法

2.1. YOLOv8 算法原理

YOLOv8 目标检测算法是基于 YOLOv5 目标检测算法改进而来的单阶段目标检测算法, 它融合了众多最前沿的技术以提升检测效果。与 YOLOv5 相比, YOLOv8 的主要改动部分在于检测头部分和损失函数计算部分。YOLOv8 的检测头换成了目前的主流解耦头结构, 把分类问题和回归问题分离以提高检测准确率, 同时使用了无锚框的范式做边框回归。损失计算的时候, 则引入了 DFL [12]损失函数。YOLOv8 目标检测算法根据模型参数数量的不同, 有五个版本: YOLOv8n、YOLOv8s、YOLOv8m、YOLOv8l、YOLOv8x, 以适应不同的应用场景。

2.2. 改进的 YOLOv8s 算法

基于上述 YOLOv8 模型, 结合遥感图像背景复杂, 尺度变化较大的特点, 分别改进原算法的数据增强、主干网络、损失函数并引入注意力机制模块, 以提升算法的检测效果。改进的 YOLOv8s 算法结构如图 1 所示。

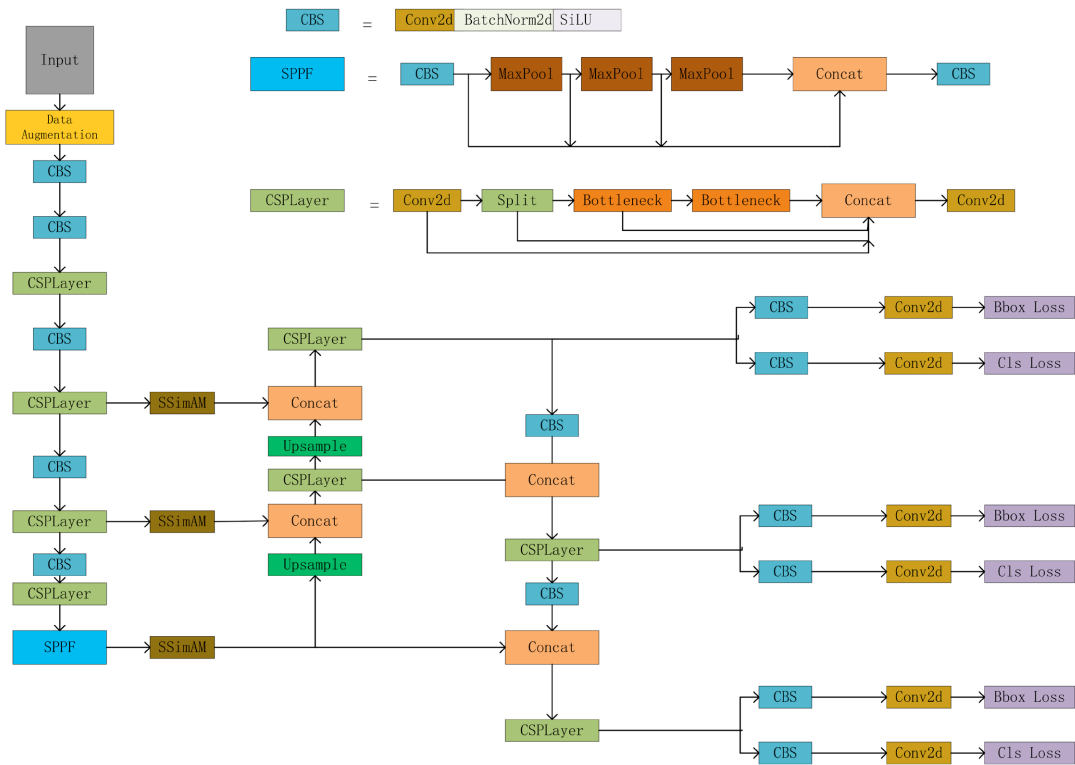


Figure 1. Improved YOLOv8 network structure
图 1. 改进的 YOLOv8 网络结构

2.2.1. Mixup 数据增强

训练过程中对图片进行数据增强是一种有效的提高模型泛化能力的做法。遥感图像中背景复杂，光照气候等条件变化大，需要基于原算法使用更多数据增强方法。Mixup 数据增强是一种与数据类型无关的数据增强操作，可以用在语音识别、计算机视觉、自然语言处理等领域，在 YOLO 系列目标检测算法中，YOLOv4 首先使用了 Mixup 数据增强操作，以提高模型泛化能力。由于不同参数量的模型应该使用不同的数据增强方式，原 YOLOv8s 模型未使用 Mixup 数据增强操作。然而在进行军事飞机遥感图像目标检测任务时，经过实验验证，引入 Mixup 数据增强方法可以有效提高 YOLOv8s 算法的准确率。Mixup 原理如公式(1)所示：

$$\begin{aligned} \tilde{x} &= \lambda x_i + (1-\lambda)x_j \\ \tilde{y} &= \lambda y_i + (1-\lambda)y_j \end{aligned} \tag{1}$$

x_i, x_j 是原始的输入图片， y_i, y_j 是独热编码之后的图片标签， $\lambda \in [0,1]$ 且服从 Beta 分布。Mixup 数据增强只增加很小的训练代价，并且可以提高模型的稳健性。

2.2.2. 改进的 NWD 距离损失

损失函数是目标检测算法的核心组件之一，它用来度量机器学习模型的预测值和真实值的差异。常见的衡量目标检测算法边框损失的函数经常是基于交并比计算的，基于交并比的度量微小物体的位置偏差非常敏感，遥感图像物体通常较小且密集，考虑使用 NWD (Normalized Wasserstein Distance, NWD) 损失缓解这种情况，同时结合 Alpha-IoU [13] 的思想，提出 α -NWD 损失函数，以进一步提高标检测准确度和收敛速度。 α -NWD 损失函数原理如下列公式所示：

$$W_2^2(\mu_a, \mu_b) = \left\| \left[\left[cx_a, cy_a, \frac{w_a}{2}, \frac{h_a}{2} \right]^T, \left[cx_b, cy_b, \frac{w_b}{2}, \frac{h_b}{2} \right]^T \right] \right\|_2^2 \quad (2)$$

$$NWD(\mu_a, \mu_b) = \exp\left(-\frac{1}{C}\sqrt{W_2^2(\mu_a, \mu_b)}\right) \quad (3)$$

$$\alpha - NWD = NWD^\alpha \quad (4)$$

$$Loss_{\alpha-NWD} = 1 - \alpha - NWD \quad (5)$$

上式中, cx_*, cy_*, w_*, h_* 分别是边界框的中心点横纵坐标和边界框的宽、高, C 是与数据集相关的常数, 一般设置为数据集目标物体的平均绝对大小, 实验中, α 的值设置为 64, α 是超参数, 实验中设置为 5.0。

2.2.3. 改进的 SimAM 注意力机制

神经网络中的注意力模块可以提高特征图的质量且不增加很高的计算成本。SimAM 是一种无参数的注意力机制模块, 它是受神经科学启发, 以最小化能量函数的方式直接学习特征图的空间信息和通道信息。SimAM 的原理如下列公式所示:

$$e_i^* = \frac{4(\tilde{\sigma}^2 + \lambda)}{t - \tilde{\mu}^2 + 2\tilde{\sigma}^2 + 2\lambda} \quad (6)$$

$$\tilde{X} = \text{sigmoid}\left(\frac{1}{e^i}\right) \cdot X \quad (7)$$

上式中, $\tilde{\mu}^2$ 是单个通道上所有神经元的均值, $\tilde{\sigma}^2$ 是单通道上所有神经元的方差。能量 e_i^* 的值越小, 代表神经元越重要。 \tilde{X} 代表表示输出特征图, X 表示输入特征图, $\text{sigmoid}(\cdot)$ 表示激活函数。

为了进一步提高 SimAM 注意力模块的性能, 结合 SA (Shuffle Attention, SA) [14] 的原理, 对 SimAM 的输入特征图进行分组并重新排列, 以此增加输入特征不同通道之间的信息交互, 构建 ShuffleSimAM 注意力模块。ShuffleSimAM 模块原理如图 2 所示:

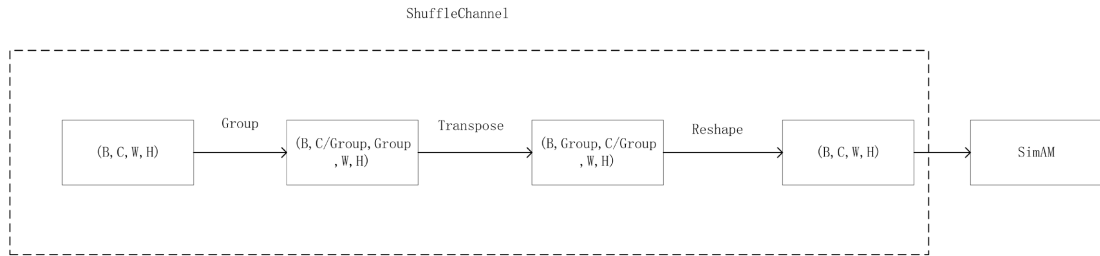


Figure 2. ShuffleSimAM module
图 2. ShuffleSimAM 模块

2.2.4. 修改主干网络结构

卷积网络的浅层特征图, 分辨率较高, 具有较低的语义信息, 深层特征图分辨率低, 具有较高的语义信息, 浅层的高分辨率特征图有利于目标对小目标的检测[15]。原 YOLOv8s 算法的骨干网络分别输出 $128 \times 80 \times 80$, $256 \times 40 \times 40$, $512 \times 20 \times 20$ 的特征图传入特征融合网络, 为了提高对小目标检测的准确率, 减少了骨干网络最后一个特征图输出层的特征通道数, 从 512 减少为 256, 发现这种方法可以有效提高 YOLOv8s 算法的准确率并可以减少参数量。

3. 实验结果分析

为了验证改进算法的性能, 使用 MAR20 军事飞机遥感图像目标检测数据集和 NWPU VHR-10 遥感图像数据集进行了实验并分析结果。

3.1. 实验环境和数据集

实验使用的操作系统是 Ubuntu 20.04.4, 显卡是 NVIDIA RTX3090, 总迭代次数为 200, 批量大小为 16, 学习率为 0.01。

MAR20 数据集是西北工业大学开源的军事飞机目标识别数据集, 包含 3842 张图像、20 种军用飞机型号以及 22,341 个目标实例。根据不同机场包含的各飞机型号和目标数量, 将 3842 张图像划分为训练集和测试集, 训练集包含 1331 张图像和 7870 个目标实例, 测试集包含 2511 张图像和 14,471 个目标实例。NWPU VHR-10 数据集是用于遥感图像目标检测的数据集, 包括 650 有标注的图像、10 个类别以及 3775 个标注实例, 实验中按照 4:1 划分训练集和测试集。

3.2. 评价指标

实验中使用三个指标来评价模型性能: 平均精度均值(mean Average Precision, mAP), 参数量和浮点计算量(Floating Point of Operations, FLOPs)。参数量越小, 目标检测算法对硬件的要求越低, FLOPs 用来衡量算法的计算复杂度, mAP 是指不同类型目标的平均精度(Average Precision, AP)。AP 是单类物体的 Precision-Recall 曲线下面积, Precision 和 Recall 的计算公式如下。

$$Precision = \frac{TP}{TP + FP} \quad (8)$$

$$Recall = \frac{TP}{TP + FN} \quad (9)$$

上式中, TP 是真正例的比例, FP 是假负例的比例。

3.3. 消融实验

Table 1. Experimental results of MAR20 dataset

表 1. MAR20 数据集实验结果

算法模块	mAP50:95 (%)	mAP50 (%)	mAP75 (%)	参数量(M)	FLOPs (G)
基线算法	66.1	87.6	82.9	11.143	14.292
Mixup	68.1	90.5	85.5	11.143	14.292
修改主干网络	66.6	88.4	83.7	6.732	12.448
NWD	66.1	88.4	82.6	11.143	14.292
α -NWD	66.6	89.1	83.4	11.143	14.292
SimAM	66.8	89.0	83.4	11.143	14.292
ShuffleSimAM	68.0	89.8	85.2	11.143	14.292
改进的算法	68.9	91.9	86.5	6.732	12.448

MAR20 数据集的实验结果如表 1 所示, 基线算法是指 YOLOv8s 目标检测算法, 改进的算法是指在此基础上使用了 Mixup 数据增强、修改骨干网络、ShuffleSimAM 注意力模块和 α -NWD 损失函数的检测

算法。由结果可知, 在 YOLOv8s 算法基础上加入本文介绍的模块后, mAP50 均有提升, 改进的算法 mAP50: 95 指标比 YOLOv8s 提高了 2.8%, mAP50 指标比 YOLOv8s 高了 4.3%, 且模型参数量减少了 39.5%, 浮点计算量减少了 12.9%, 实验结果验证了改进算法的有效性。

改进算法在 NWPU VHR-10 数据集的结果如表 2 所示。改进算法的 mAP50:95 指标提高了 0.6%, mAP50 指标提高了 2.2%, 验证了改进算法的泛化性。

Table 2. Experimental results of NWPU VHR-10 dataset

表 2. NWPU VHR-10 实验结果

算法模块	mAP50:95 (%)	mAP50 (%)	mAP75 (%)	参数量(M)	FLOPs (G)
基线模型	56.3	88.9	65.8	11.139	14.281
改进的算法	56.9	91.1	62.7	6.728	12.437

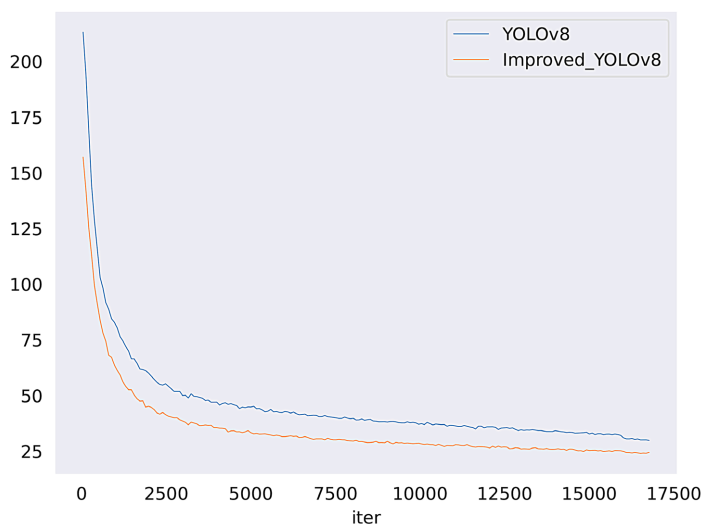


Figure 3. Loss changing on MAR20

图 3. MAR20 数据集损失变化

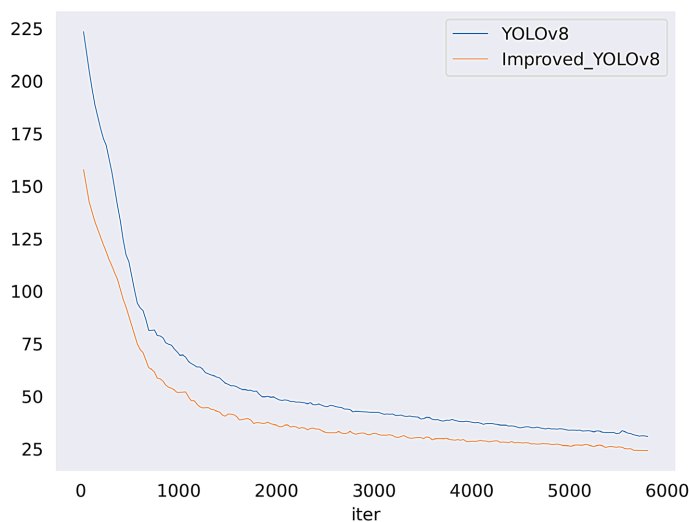


Figure 4. Loss changing on NWPU VHR-10

图 4. NWPU VHR-10 数据集损失变化

基线模型和改进模型在 MAR20 数据集和 NWPU VHR-10 数据集训练过程中损失的变化如图 3 和图 4 所示。可以看出改进算法的总损失小于 YOLOv8s 的总损失，可以实现更好的检测效果。

图 5 和图 6 是 YOLOv8s 的检测效果和改进算法的检测效果，图中飞机类别均为 A16，从图中可以看出，YOLOv8s 原算法在检测时有 1 个误检为 A5 类型，而改进算法的检测结果均正确。

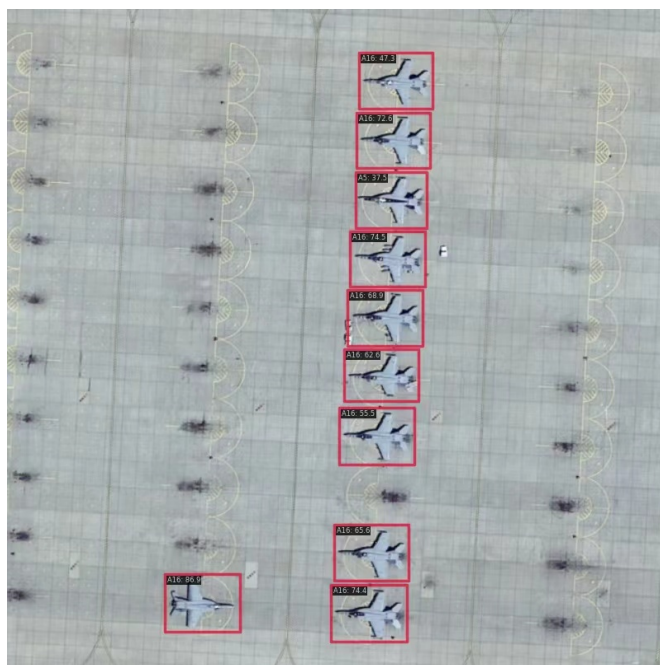


Figure 5. Detection result of the YOLOv8s algorithm

图 5. YOLOv8s 算法检测效果

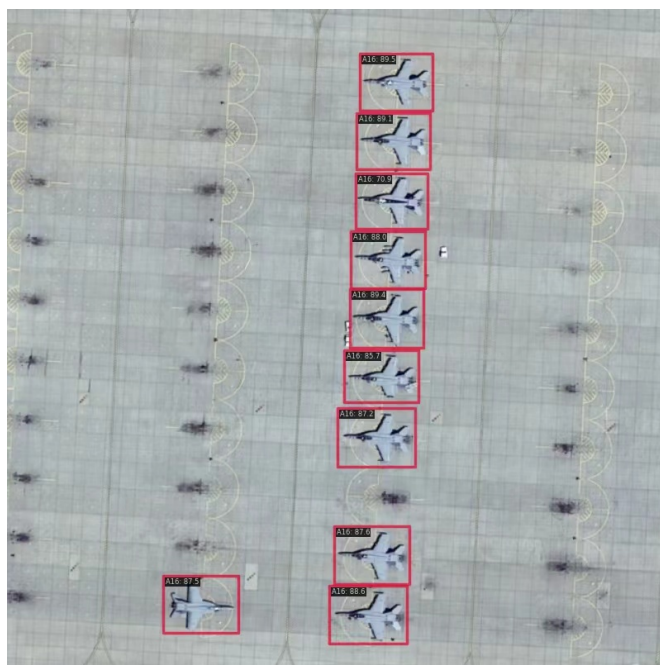


Figure 6. Detection result of the improved YOLOv8s algorithm

图 6. 改进的 YOLOv8s 算法检测效果

3.4. 对比实验

对改进的算法与其他常见目标检测算法在 MAR20 数据集上进行了对比实验, 包括 YOLOv5s, YOLOv6s [16], YOLOXs [17], 实验结果如表 3 所示。从表中结果可以看出, 改进算法的准确率比 YOLOv5s, YOLOv6s, YOLOX 更高且有更少的参数, 验证了本文改进的模型有效性。

Table 3. Comparison experiment results of MAR20 dataset

表 3. MAR20 数据集对比实验结果

算法	mAP50 (%)	mAP75 (%)	参数量(M)	FLOPs (G)
YOLOv5s	88.7	82.1	7.074	8.007
YOLOv6s	81.7	74.6	17.196	21.895
YOLOXs	90.2	79.0	8.945	13.339
YOLOv8s	87.6	82.9	11.143	14.292
改进算法	91.9	86.5	6.732	12.448

4. 结束语

基于 YOLOv8s 目标检测算法, 构建了改进的 YOLOv8s 算法: 引入 Mixup 数据增强, 提高训练集数据分布的多样性; 修改主干网络, 以减少深层语义信息的参数量; 为主干网络融合 ShuffleSimAM 注意力模块, 以提高输出的特征图质量; 在检测头部分, 使用 α -NWD 损失函数替换原本的 CIoU 损失函数, 以更好计算小物体的损失。在 MAR20 数据集和 NWPU VHR-10 数据集上的实验结果, 证明了改进算法的有效性。在后续研究中, 可以进一步探究不同注意力机制[18] [19]和损失函数的作用, 以构建更加高效的轻量化遥感图像目标检测算法。

参考文献

- [1] 王杰, 张上, 张岳, 等. 改进 YOLOv5 的军事飞机检测算法[J]. 无线电工程, 2024, 54(3): 589-596.
- [2] Xu, S., Chen, Z., Zhang, H., Xue, L. and Su, H. (2024) Improved Remote Sensing Image Target Detection Based on Yolov7. *Optoelectronics Letters*, **20**, 234-242. <https://doi.org/10.1007/s11801-024-3063-z>
- [3] Cui, C., Wang, R., Wang, Y., Zhou, F., Bian, X. and Chen, J. (2024) Research on Optical Remote Sensing Image Target Detection Technique Based on Dch-Yolov7 Algorithm. *IEEE Access*, **12**, 34741-34751. <https://doi.org/10.1109/access.2024.3368877>
- [4] Sunkur, R., Kantamaneni, K., Bokhoree, C., Rathnayake, U. and Fernando, M. (2024) Mangrove Mapping and Monitoring Using Remote Sensing Techniques Towards Climate Change Resilience. *Scientific Reports*, **14**, Article No. 6949. <https://doi.org/10.1038/s41598-024-57563-4>
- [5] 章程军, 胡晓兵, 魏上云, 等. 基于深度学习的遥感目标检测技术[J]. 计算机工程, 2024, 45(2): 594-600.
- [6] Tan, M., Pang, R. and Le, Q.V. (2020) Efficient Det: Scalable and Efficient Object Detection. 2020 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Seattle, 13-19 June 2020, 10778-10787. <https://doi.org/10.1109/cvpr42600.2020.01079>
- [7] 谢俊章, 彭辉, 唐健峰, 等. 改进 YOLOv4 的密集遥感目标检测[J]. 计算机工程与应用, 2021, 57(22): 247-256.
- [8] Zhang, H., Cisse, M., Dauphin, Y.N., et al. (2018) Mixup: Beyond Empirical Risk Minimization. *International Conference on Learning Representations*, Vancouver, 30 April-3 May 2018, 1-13.
- [9] Xu, C., Wang, J., Yang, W., Yu, H., Yu, L. and Xia, G. (2022) Detecting Tiny Objects in Aerial Images: A Normalized Wasserstein Distance and a New Benchmark. *ISPRS Journal of Photogrammetry and Remote Sensing*, **190**, 79-93. <https://doi.org/10.1016/j.isprsjprs.2022.06.002>
- [10] Yang, L., Zhang, R.-Y., Li, L. and Xie, X. (2021) SimAM: A Simple, Parameter-Free Attention Module for Convolutional Neural Networks. *Proceedings of the 38th International Conference on Machine Learning*, New York, 18-24

- July 2021, 11863-11874.
- [11] 禹文奇, 程焱, 王美君, 等. MAR20: 遥感图像军用飞机目标识别数据集[J]. 遥感学报, 2023, 27(12): 2688-2696.
 - [12] Li, X., Lv, C., Wang, W., *et al.* (2023) Generalized Focal Loss: Towards Efficient Representation Learning for Dense Object Detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **45**, 3139-3153. <https://doi.org/10.1109/TPAMI.2022.3180392>
 - [13] He, J., Erfani, S., Ma, J., *et al.* (2021) Alpha-IoU: A Family of Power Intersection over Union Losses for Bounding Box Regression. arXiv: 2110.13675. <https://doi.org/10.48550/arXiv.2110.13675>
 - [14] Zhang, Q.-L. and Yang, Y.-B. (2021) SA-Net: Shuffle Attention for Deep Convolutional Neural Networks. *ICASSP 2021-2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Toronto, 6-11 June 2021, 2235-2239. <https://doi.org/10.1109/ICASSP39728.2021.9414568>
 - [15] Lin, T., Dollar, P., Girshick, R., He, K., Hariharan, B. and Belongie, S. (2017) Feature Pyramid Networks for Object Detection. 2017 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, 21-26 July 2017, 936-944. <https://doi.org/10.1109/cvpr.2017.106>
 - [16] Li, C., Li, L., Jiang, H., *et al.* (2022) YOLOv6: A Single-Stage Object Detection Framework for Industrial Applications. arXiv:2209.02976. <https://doi.org/10.48550/arXiv.2209.02976>
 - [17] Ge, Z., Liu, S., Wang, F., *et al.* (2021) YOLOX: Exceeding YOLO Series in 2021. arXiv: 2107.08430. <https://doi.org/10.48550/arXiv.2107.08430>
 - [18] Guo, M., Lu, C., Liu, Z., Cheng, M. and Hu, S. (2023) Visual Attention Network. *Computational Visual Media*, **9**, 733-752. <https://doi.org/10.1007/s41095-023-0364-2>
 - [19] Lau, K.W., Po, L.-M. and Ur Rehman, Y.A. (2024) Large Separable Kernel Attention: Rethinking the Large Kernel Attention Design in CNN. *Expert Systems with Applications*, **236**, Article 121352. <https://doi.org/10.1016/j.eswa.2023.121352>