DOI:10.19651/j.cnki.emt.2314480

面向边缘计算平台及遥感影像的实时检测算法*

杨 洋[1] 宋品德[1] 杨思念[2] 曹立佳[2,3,4]
(1.四川轻化工大学自动化与信息工程学院 宜宾 644000; 2.四川轻化工大学计算机科学与工程学院 宜宾 644000;
3.人工智能四川省重点实验室 宜宾 644000; 4.企业信息化与物联网测控技术四川省高校重点实验室 宜宾 644000)

摘 要: 针对现有目标检测算法难以满足无人机遥感中实时检测的问题,提出了一种基于 ShuffleNetv2 及结构化剪枝的模型压缩方法。以 YOLOv5m 为基础,将 ShuffleNetv2 模型作为 YOLOv5m 的主干网络,减少模型的参数量及计算量,提升模型推理速度;其次,利用 ECA 注意力机制替换 ShuffleNetv2 中的 SE 模块,强化主干网络的特征提取能力;再者,以 FocalEIoU 作为 YOLOv5 算法的损失函数,提升模型的回归能力;最后,利用通道剪枝算法剔除 Neck 结构中冗余的参数,进一步压缩模型的参数及计算量,并通过模型微调的方式提升剪枝模型的精度。实验结果表明,在相同的测试环境下,与 YOLOv5m 相比,本文所提出模型的参数量及浮点运算量分别降低了 86.3%和 80.0%,mAP@0.5 和 mAP@0.5:0.95 达到了 92%及 50.4%,优于所对比的其他主流检测算法。此外,所提出的模型在 AGX 边缘计算平台上达到了 35 帧/s 的检测速度,满足实时检测的要求。

关键词: 遥感影像;剪枝;轻量化网络;FocalEIoU 损失函数;边缘计算平台
中图分类号: TP391 **文献标识码**: A **国家标准学科分类代码**: 520.20

Real-time detection algorithm for edge computing platforms and remote sensing imagery

Yang Yang[1] Song Pinde[1] Yang Sinian[2] Cao Lijia[2,3,4]
(1. School of Automation and Information Engineering, Sichuan University of Science & Engineering, Yibin 644000, China;
2. School of Computing Science and Engineering, Sichuan University of Science & Engineering, Yibin 644000, China;
3. Artificial Intelligence Key Laboratory of Sichuan Province, Yibin 644000, China; 4. Key Laboratory of Higher Education of Sichuan Province for Enterprise Informationalization and Internet of Things, Yibin 644000, China)

Abstract: To address the issue of existing object detection algorithms struggling to meet real-time detection requirements in UAV remote sensing, we propose a model compression method based on ShuffleNetv2 and structured pruning. Using YOLOv5m as the foundation, we incorporate the ShuffleNetv2 model as the backbone network of YOLOv5m, reducing the model's parameter count and computational complexity while improving inference speed. Furthermore, we employ the ECA attention mechanism to replace the SE module in ShuffleNetv2, enhancing the feature extraction capability of the backbone network. Additionally, we adopt FocalEIoU as the loss function for the YOLOv5 algorithm, improving the model's regression ability. Finally, we use channel pruning to eliminate redundant parameters in the Neck structure, further compressing the model's parameters and computational complexity, and enhancing the pruned model's accuracy through fine-tuning. Experimental results show that, under the same testing environment, compared to YOLOv5m, the proposed model reduces the parameter count and floating-point operations by 86.3% and 80.0%, respectively. The model achieves an mAP@0.5 of 92% and an mAP@0.5:0.95 of 50.4%, outperforming other mainstream detection algorithms. Moreover, the proposed model achieves a detection speed of 35 frames/s on the AGX edge computing platform, satisfying the requirements for real-time detection.

Keywords: remote sensing image; pruning; lightweight network; FocalEIoU loss; edge computing platform

0 引 言

遥感目标检测是计算机视觉领域的重点研究课题之一,其在军事战争、民用生活、灾后搜救等方面有广阔的应用场景,但遥感检测场景复杂,实际应用部署存在诸多局限,虽然目前主流的目标检测算法已经广泛应用于遥感领

收稿日期:2023-08-29
* 基金项目:国家自然科学基金(51905540)、中国高校产学研创新基金(2021ZYA11002)、研究生创新创业基金(Y2023271)项目资助

域,但仍存在一些问题,如模型参数量大、浮点运算量高以及小目标的检测能力差。因此,兼顾准确性与实时性的遥感目标检测算法具有重要的研究意义。

近年来,诸多学者在遥感目标检测领域开展了一系列的研究。范新南等[1]将 ResNet[2]与 Faster-RCNN[3]主干网络相结合,并在此基础上加入特征金字塔网络以解决遥感影像中语义和位置信息表达不全的问题,但由于 Faster-RCNN 是两阶段检测算法,因此,难以满足实时检测的要求。方青云等[4]针对在轨卫星实时检测的需求,提出了基于 MobileNet[5]与 YOLOv3[6]的轻量化遥感目标快速检测方法,在 Nvidia Titan xp 上达到了 101 fps 的检测速度,但该算法的检测精度较低,难以满足实际需求。韩要昌等[7]针对遥感目标检测模型压缩的问题,提出了基于 L1 正则化的模型剪枝方法,极大地压缩了模型的参数量,但其导致检测精度下降。宋中山等[8]为解决 YOLOv4 目标检测网络结构复杂、参数多、训练所需的配置高以及实时检测图片的传输帧数低等问题,使用 ShuffleNetv2 轻量级网络,提出一种基于 YOLOv4 改进的轻量化算法。王成龙等[9]针对遥感飞机目标检测场景中轻量级算法难以兼顾准确性与实时性的问题,提出了一种基于 YOLOv4[10]的结构化剪枝算法,并对剪枝后的模型进行微调,以恢复模型的检测精度,从而有效地保证了检测的实时性,但其检测精度还有待提高。鄢奉习等[11]针对遥感图像中存在背景复杂、密集小目标及目标遮挡等导致飞机检测精度不高的问题采用稠密连接和引入注意力机制,提出了基于 YOLOv5 的遥感图像飞机检测算法。钱承山等[12]针对遥感图像中目标排列紧密,背景复杂的问题,利用 Transformer 对 YOLOv5 进行改进。张洋等[13]为提高遥感目标检测的精度,利用融合多头注意力网络和特征金字塔网络,提出了基于 YOLOv5 的多尺度特征融合遥感目标检测方法。吴建成等[14]针对遥感图像背景复杂、小目标多和特征提取难等问题,设计了一种带通道注意力的主干网络模块和模糊池模块,并基于 YOLOv7 提出了 YOLO-Aff。郎磊等[15]为解决高分辨率遥感影像目标检测中复杂背景、密集物体及目标尺度差异大等问题,提出了一种基于 YOLOX-Tiny 的轻量级遥感目标检测网络,利用协通注意力模块和可变性卷积强化空间建模能力,并利用完全交并比(complete intersection over union,CIoU[16])损失函数改善遥感目标定位精度的问题,在 RTX2080 显卡上达到了 46 fps 的检测速度。闫钧华等[17]为解决遥感图像地面弱小目标检测中弱小目标信息量少、信息真假混杂的难题,对空间金字塔图进行跨层级通道特征融合,结合位置注意力机制,提出了基于 YOLOv5 的一种融合多层级特征的遥感图像地面弱小目标检测算法 CC-YOLO,在 RTX3070 显卡上达到了 58 fps 的检测速度。上述的算法虽然在一定程度上解决了遥感影像中检测精度的问题,但对于算力受限的嵌入式设备应用场景下,仍难以保证在较高的检测精度条件下满足实时的性能。此外,在嵌入式环境下,应进一步考虑内存访问效率的问题,以控制模型推理过程中的时间成本。

针对上述问题,本文提出了一种基于结构化剪枝的轻量级无人机遥感目标检测算法。在 YOLOv5m 模型架构的基础上,改进模型的骨干网络与损失函数,并通过通道剪枝的方式深度压缩模型参数及计算量。实验结果表明,本文所提出的算法在嵌入式设备中能较好地均衡检测精度与检测速度,且满足实时检测的要求。

1　YOLOv5 算法

YOLOv5 系列模型都包含相同的输入端、骨干网络、Neck 及预测头,仅仅通过宽度及深度参数控制模型的大小,包含以下不同参数量的 5 个版本,分别是 YOLOv5n、YOLOv5s、YOLOv5m、YOLOv5l 及 YOLOv5x,其区别在于网络之间的深度和宽度不同。

在数据增强方面,YOLOv5 使用了 Mosaic 及 Mixup,通过多图片组合的方式提升正样本数量,从而降低负样本对网络的干扰。此外,YOLOv5 利用 Kmeans 及遗传算法自适应 anchor 框,以增强网络对检测对象的拟合能力。

YOLOv5 的主干网络包含 Conv、C3 及 SPPF 模块。Conv 是 YOLOv5 的标准卷积单元,C3 是残差单元的堆叠,而 SPPF 则是 SPP 模块的改进。

在损失函数方面,YOLOv5 使用二值交叉熵作为类别和置信度的损失函数,如式 (1)所示。

$$loss(p,y)=-\frac{1}{N}\sum_{i}^{N}[y^i\times\log(p^i)+(1-y^i)\times\log(1-p^i)] \tag{1}$$

其中,p 代表预测类别,y 代表真实标签,i 代表第 i 个标签,N 代表标签数,y^i 代表第 i 个的真实类别,p^i 代表第 i 个标签的预测概率。

在边框回归损失中,YOLOv5 以 CIoU 作为其损失函数,如式 (2)所示。

$$\begin{aligned}CIoU_{\text{loss}}&=1-IoU+\frac{\rho^2(b^p,b^{gt})}{c^2}+\alpha\nu\\ \alpha&=\frac{\nu}{(1-IoU)+\nu}\\ \nu&=\frac{4}{\pi^2}\left(\arctan\frac{w^{gt}}{h^{gt}}-\arctan\frac{w}{h}\right)^2\end{aligned} \tag{2}$$

式中:α 是权重系数,υ 衡量预测框与真实框长宽比的相似性,c 为两框外接矩形的对角线距离,ρ 为两个框的欧式距离,b^{gt} 为真实标签,b^p 为预测框,w^{gt} 与 h^{gt} 为真实框的宽高,w 和 h 为预测框的宽高。

在本文中,为确保检测算法的精度与速度相均衡,以 YOLOv5m 为基准并对其进行改进,以提升该算法在遥感影像检测中的性能。

2　剪枝算法

CNN 在许多应用场景中的部署受到其高计算成本的

阻碍，Liu 等[18]提出了一种基于批归一化（batch normalization，BN）层通道的网络剪枝算法。首先，为每个 BN 通道添加缩放因子 γ；其次，引入 L1 正则化对 γ 进行稀疏化训练，如式（3）所示，使 γ 的分布趋近于 0；最后，对 γ 所对应的卷积核及 BN 层进行裁剪处理，并剪除下一层卷积核所对应的通道，其剪枝原理如图 1 所示。

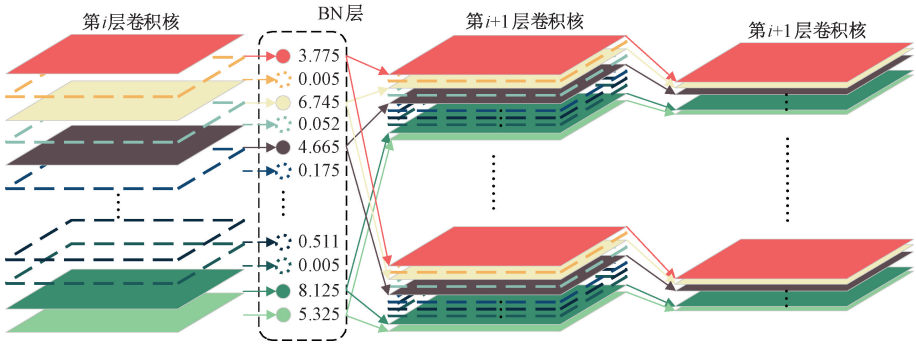


图 1　基于 L1 正则化稀疏训练的剪枝原理

$$L = L_{cls} + L_{obj} + L_{box} + L_1$$
$$L_1 = \lambda \sum_{\gamma \in \Gamma} g(\gamma) \tag{3}$$

式中：L_{cls}、L_{obj}、L_{box} 分别代表分类损失、置信度损失以及边框回归损失，L_1 为稀疏训练引入的正则项，λ 为惩罚项以均衡四者的损失，γ 为某 BN 层的缩放因子，Γ 为网络中需要稀疏训练的 γ 集合，$g(\gamma) = |\gamma|$。

图 1 中假设第 i 层的卷积核通道数为 1，当 BN 层的 γ 经过稀疏训练后，该值趋向于零分布，即上一层所对应趋向于零的 γ 的卷积核在网络中的收益较小，其原理如式(4)所示。

$$bn_{out} = \sum_{j=1}^{N} \gamma_j \frac{c_j - \mu_j}{\sqrt{\sigma_j^2 + \varepsilon}} + \beta_j \tag{4}$$

式中：μ_j 和 σ_j^2 为第 i 层第 j 个通道某批次的均值和方差，β_j 为第 i 层第 j 个通道的偏置，ε 为常量，防止分母为 0，c_j 为第 i 层第 j 个通道的卷积输入特征。由式(4)可知，当 γ_j 趋于 0 时，第 j 个 bn_{out} 的值也是趋于 0 的，即第 i 层的第 j 个卷积核收益趋近于 0，故裁剪该卷积核，从而实现网络的轻量化。完成模型的稀疏化及剪枝后，再利用 Adam 优化器训练 300epoch 以完成模型微调，微调后的网络性能优于原始网络。

3　YOLOv5m 网络结构改进设

为了实现遥感影像检测的精度和速度相均衡，本文提出了一种基于 YOLOv5m 的轻量化遥感影像检测模型，如图 2 所示。在该网络结构中，引入 ShuffleNetv2[19]骨干网络以减少模型的复杂度及计算量；其次，将高效的通道注意力机制（efficient channel attention，ECA[20]）与 ShuffleNetv2 相结合，其作用是减少模型参数量及增强网络的特征提取能力，通过突出遥感影像的关键信息，从而提高网络的检测能力。

3.1　主干网络改进

在轻量级网络的发展中，通常以浮点运算数（floating point operations，FLOPs）为依据来度量模型的计算复杂度，但在移动设备中，其运行速度不仅仅需要考虑 FLOPs，还需要考虑其他因素，如内存访问成本以及硬件平台对卷积核的优化。因此，Ma 等针对内存访问效率的问题提出了 ShuffleNetv2，该网络利用通道混淆降低模型的计算成本，并通过减少网络并行分支的方式，降低内存开销，其基本单元如图 3 所示。针对 YOLOv5 主干网络参数大，计算成本高的问题，本文将该模型与 YOLOv5 相结合，以提升网络整体的推理性能。

为进一步简化骨干网络的模型复杂度，本文将 ShuffleNetv2 基本卷积单元的挤压和激励（squeeze and excitation，SE）注意力机制替换为 ECA 注意力机制，在减少模型参数量的同时提升模型的特征提取能力。ECA 注意力机制如图 4 所示。

3.2　损失函数改进

YOLOv5 使用 CIoU 作为边框回归损失函数，该损失函数从 3 个角度评估预测框与真实框之间的差异，即重叠面积、中心点距离及长宽比，但由于反三角函数的存在，其计算量较大，需要更长的训练时间；其次，式（2）中 v 仅仅反映纵横比差异，没有独立考虑宽高各自的差异，从而阻碍模型减少预测值与真实框之间的差异。

为解决上述问题，本文将 FocalEIoU[21]损失函数与 YOLOv5 算法相结合，FocalEIoU 如式(5)所示。

$$L_{EIoU} = 1 - IoU + \frac{\rho(b, b^{gt})}{c^2} + \frac{\rho(w, w^{gt})}{c_w^2} + \frac{\rho(h, h^{gt})}{c_h^2}$$
$$L_{FocalEIoU} = IoU^{\gamma} L_{EIoU} \tag{5}$$

式中：c 为预测框与真实框外接矩形的对角线距离，c_w 为两框闭包矩形的宽，c_h 两框闭包矩形的高，ρ 为两框的欧式距离，b^{gt} 为真实标签，b 为预测框，w^{gt} 与 h^{gt} 为真实框的宽高，w 和 h 为预测框的宽高，γ 为超参数。

EIoU 将边长作为惩罚项，在一定程度上解决了边长被错误放大的问题，更有利于边框回归的收敛；其次，FocalLoss

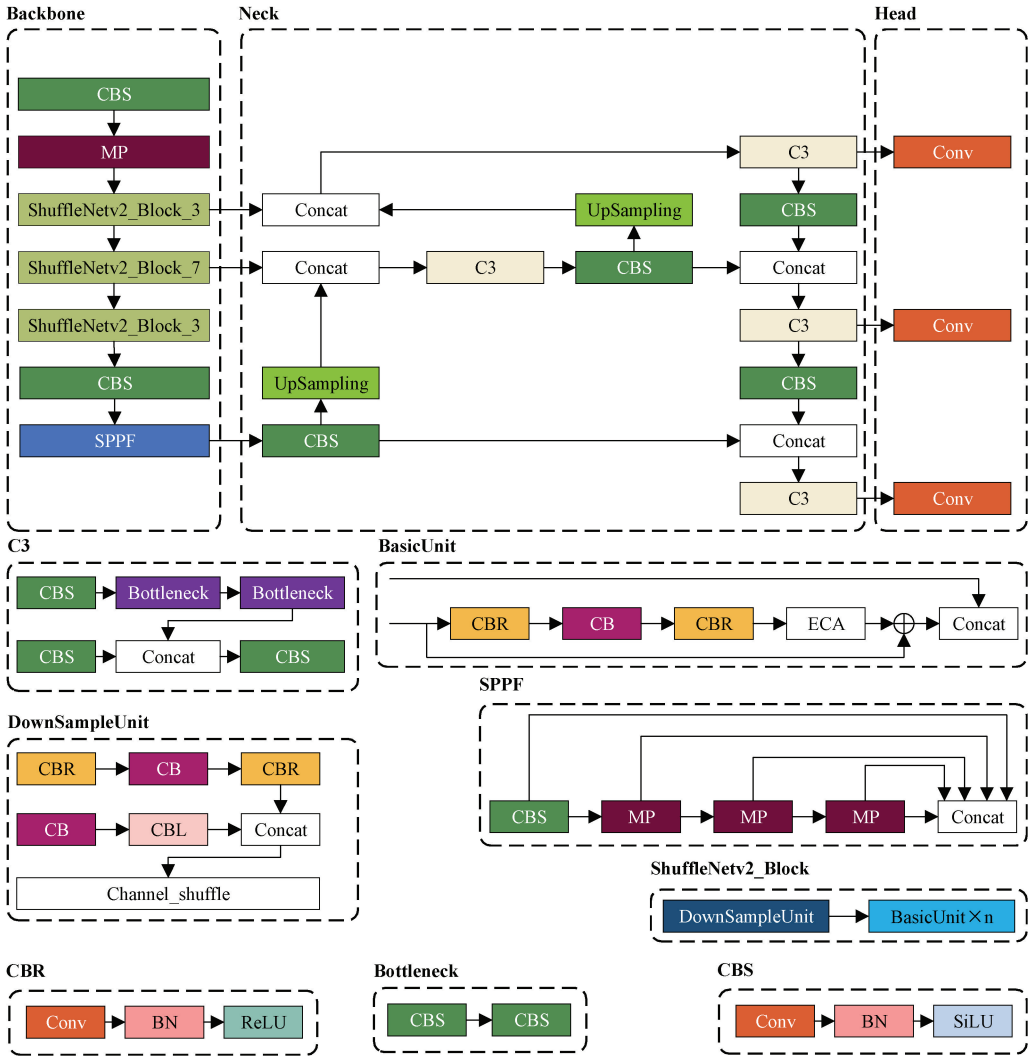


图 2　改进的 YOLOv5m 网络结构

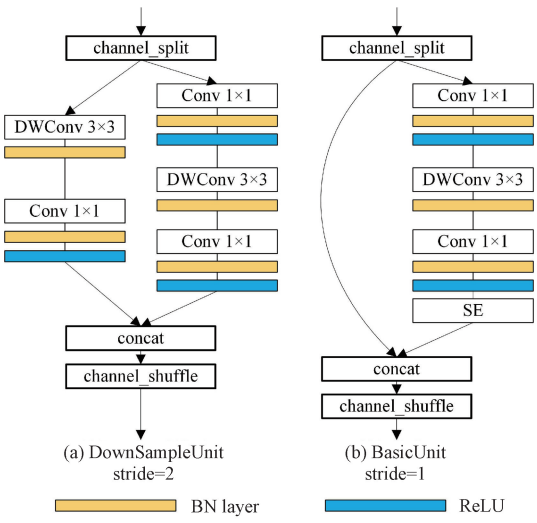


图 3　ShuffleNetv2 基本单元

思想的引入解决了正负样本不平衡的问题,从而减少低质量样本对梯度的影响。

3.3　模块剪枝

模型剪枝需要对指定 BN 层中的 γ 参数进行稀疏训练,由式 (3)可知 L1 惩罚项中需要设定一个 λ 参数以保证 γ 参数的稀疏化,因此,本文定义一个动态 λ 参数,如式(6)所示。

$$\lambda = \lambda_c(1-0.003x) \tag{6}$$

式中:λ_c 代表稀疏常数,x 代表当前训练轮次,λ 为实时稀疏常数。

本文仅对 ES2－YOLOv5m 模型里的 SPPF 及 C3 进行剪枝,故需要对上述模块 BN 层中的 γ 进行约束以达到稀疏化训练的目的。

完成模型稀疏化训练后,通过式 (3)中的 γ 参数作为评价 BN 层所属卷积核的重要性。本文定义 γ_h 为剪枝阈值,如式 (7)所示。

$$\gamma_h = \min(\max_{k=1}^{N}(\gamma_k)) \tag{7}$$

式中：γ_k 代表第 k 个稀疏训练 BN 层中的 γ 因子，$\max_{k=1}^{N}$ 代表对每一个受稀疏训练的 BN 层中的 γ 求最大值。

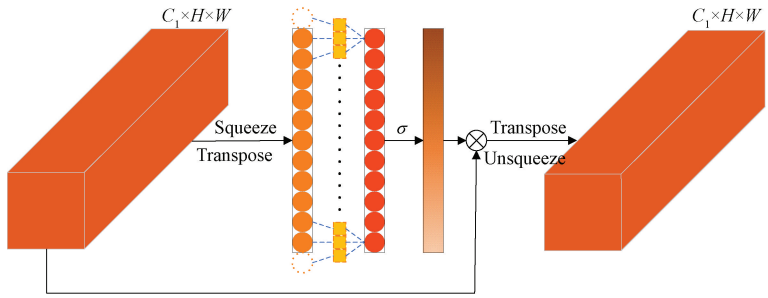


图 4　ECA 注意力机制

4　实验与分析

4.1　实验环境及训练细节

实验的测试环境为 Ubuntu 18.04 系统，YOLOv5 算法实现基于 PyTorch1.10.0 框架及 Python3.6.15。利用 mmdetection 工具箱验证主流算法的性能，其版本为 3.0.0rc2；服务器配置为双路 RTX8000 显卡、Intel Xeon Gold 6254 处理器；此外，边缘计算平台的性能验证基于 NVIDIA Jetson AGX Xavier。

Anchor 的生成基于 YOLOv5 的 Kmeans 聚类及遗传算法，本文所使用数据集的 anchor 如表 1 所示。

表 1　基于 NWPU VHR-10 数据集的 anchor 尺寸

特征尺寸	Anchor		
80×80	34,25	58,29	54,47
40×40	91,37	91,37	154,48
20×20	145,94	336,118	275,216

YOLOv5 训练中使用的超参数及配置如表 2 所示。

表 2　YOLOv5 训练超参数及其配置

参数及配置名	常规训练	稀疏训练	微调训练
优化器	SGD	Adam	Adam
初始学习率	0.01	0.01	32×10^{-4}
动量因子	0.937	0.937	0.843
权重衰减参数	5×10^{-4}	5×10^{-4}	36×10^{-5}
批次大小	32	32	32
迭代次数	300	100	300
预热批次	3	3	2
预热动量	0.8	0.8	0.5
预热偏置学习率	0.1	0.1	5×10^{-2}
输入图片尺寸	640×640	640×640	640×640

4.2　实验数据集

本文在 NWPU VHR-10 公开数据集上进行性能比对实验。该数据集为西北工业大学标注的航天遥感目标检测数据集，共包含 800 张图像，其中 650 张有 10 个类别的标签数据，其余为背景图像。此外，按照 8∶2 的比例随机将数据集划分为相互独立的训练集与验证集，即训练集 454 张图像，验证集 196 张图像。

4.3　评价指标

为评估改进 YOLOv5m 算法。本文使用准确度、回召率、mAP、参数量、浮点运算量及 FPS 来评估模型的性能。

准确率及回召率如式(8)与(9)所示。

$$Precision = \frac{TP}{TP + FP} \tag{8}$$

$$Recall = \frac{TP}{TP + FN} \tag{9}$$

其中，TP、FP 分别表示标签被正确或者错误地划分为正例的个数；FN 表示被错误地划分为负例的个数。

mAP 的定义如式 (10)所示。

$$AP = \int_0^1 P(R)$$
$$mAP = \frac{\sum_{i=1}^{N} AP_i}{N} \tag{10}$$

模型的参数量及浮点运算量以 thop 第三方库进行统计。FPS 的计算包含 3 个阶段，分别是图像预处理、模型推理、后处理，三者时间相加作为一帧图像预测的耗时成本。

4.4　主干网络改进实验

基于 YOLOv5m 将 ShuffleNetv2 作为该模型的主干网络，得到 YOLOv5m-SN2 模型。在此基础上基于 ECA 注意力机制将 YOLOv5m-SN2 重构为 YOLOv5m-SN2-ECA，即 ES2-YOLOv5m 模型，表 3 针对 NWPU VHR-10 数据集对上述模型在训练服务器上进行了参数量、浮点运算量、FPS 及 mAP 对比。

从表 3 可以看出替换主干网络后 YOLOv5m-SN2 的参数量降低了 42.43%，但引起了检测精度的降低，而引入 ECA 则提高了检测精度。

4.5　损失函数改进实验

本文将 FocalEIoU 边框回归损失函数引入 ES2-YOLOv5m 中，得到名为 ESF-YOLOv5m 的模型，并在此基础上对比了主流损失函数在遥感目标检测中的性能表

表 3　YOLOv5m 主干网络改进性能对比

方法	Params/M	FLOPs/G	mAP@0.5/%	mAP@0.5:0.95/%	帧率/fps
YOLOv5m	20.88	48.06	89.2	47.9	23.8
YOLOv5m-SN2	12.02	21.86	84.6	43.9	27.8
ES2-YOLOv5m	11.39	21.86	86.1	44.6	28.2

现,如图 5 所示。

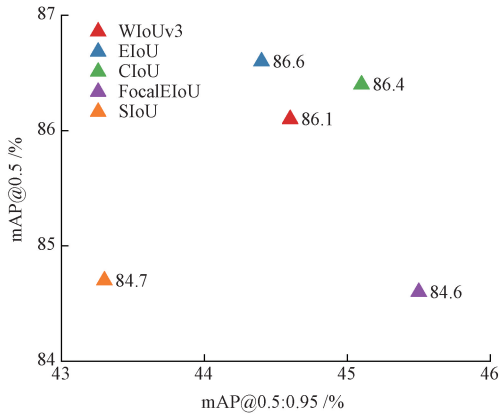


图 5　主流损失函数在遥感目标检测中的性能对比

从图 5 可以清晰地看出,EIoU 有最高的 mAP@0.5,但 FocalEIoU 在 mAP@0.5:0.95 中表现最为优异,说明在精确检测上 FocalEIoU 更具优势。

对比图 5 与表 3 中的两个 mAP 指标,本小节所引入的 FocalEIoU 在边框回归中提高了 ES2-YOLOv5m 的 mAP。由此证明:该损失函数更有利于遥感影像中的目标检测。

4.6　模型稀疏化及剪枝实验

设置不同的 λ_c 惩罚项对 ESF-YOLOv5m 进行稀疏化训练,训练过程中的 γ 分布如图 6 所示。

从图 6 可知,当稀疏常数为 0.1 时,γ 的值分布在 1.0 附近,但随着稀疏常数与 epoch 的增大,γ 逐渐趋于 0,如图 6(c)与(d)所示。

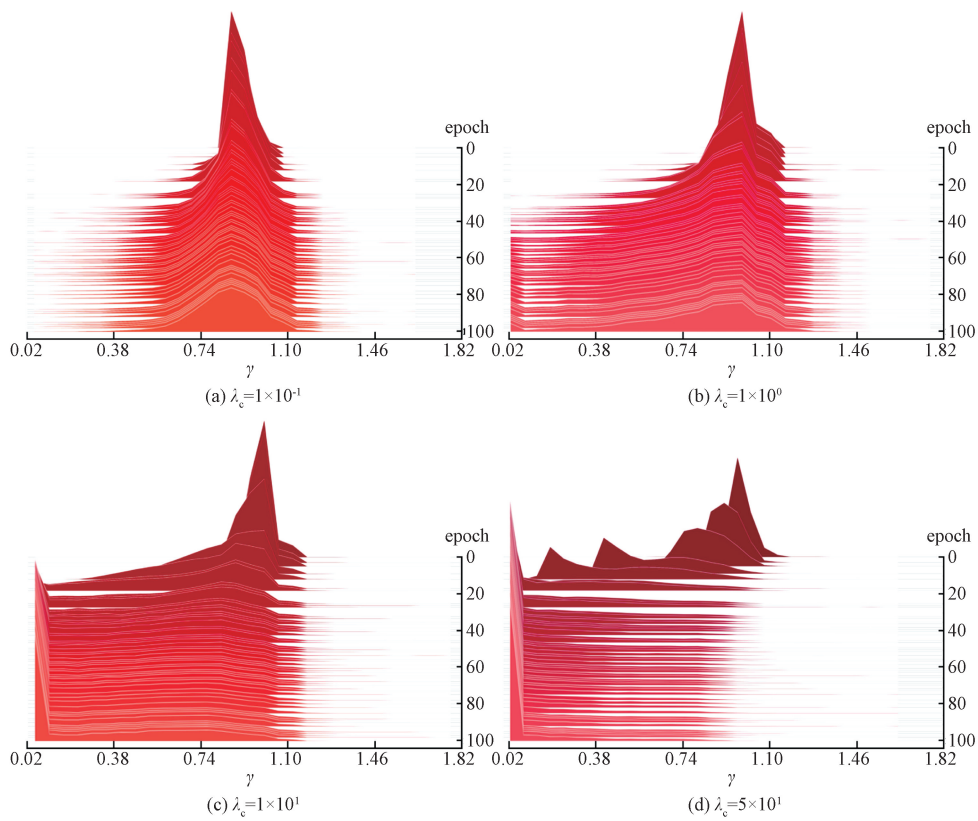


图 6　γ 随 λ_c 变化的分布情况

由图 6(d)可知,当稀疏常数为 50 时,多数 BN 层的参数分布在零附近,因此,该部分 BN 层所对应的参数对卷积核的增益较小,可做剪枝处理。为证明基于该稀疏常数下的稀疏训练的有效性,本文对图 6(d)所对应的模型进行裁剪和微调,其结果如表 4 所示。

由表 4 可知,随着裁剪比率的增加,ESF-YOLOv5m

表 4　不同剪枝率剪枝性能对比

方法	Rate/%	Params/M	FLOPs/G	mAP@0.5/%	mAP@0.5:0.95/%	帧率/fps
ESF-YOLOv5m	0.1	9.15	19.7	90.8	50.4	28.9
	0.2	4.77	18.0	91.3	49.9	29.6
	0.3	6.24	16.6	90.8	50.5	30.6
	0.4	5.39	15.4	91.8	51.0	31.3
	0.5	4.39	13.3	90.3	50.1	33.0
	0.6	3.54	11.4	91.0	50.2	34.5
	0.7	2.87	9.6	92.0	50.4	35.0

模型的参数量及浮点运算量均在降低，FPS 也有显著的提升。将使用裁剪比率 0.7 的 ESF-YOLOv5m 模型命名为 ESFP-YOLOv5m。ESFP-YOLOv5m 在 AGX 平台上达到了 30 fps 的检测速度。与表 1 中的 YOLOv5m 相比，参数量及浮点运算量分别减少了 86.3%和 80%，mAP@0.5 与 mAP@0.5:0.95 提升了 2.8%和 2.5%，在 AGX 边缘计算平台上，检测帧率较后者提升了 8 fps。实验结果证明，对大尺寸的模型做剪枝处理可有效地降低模型的参数量及浮点运算量，提升模型的推理性能。因此，该模型可有效地部署于边缘计算设备，从而完成遥感检测任务。

4.7　主流算法性能对比

为验证本文所提出算法的有效性，本文将改进的 ESFP-YOLOv5m 与主流的检测算法 SSD[22]、RetinaNet[23]、FCOS[24]、Faster-RCNN[3] 及 YOLO 系列[6,10,25-29]在参数量、浮点运算量、mAP@0.5、mAP@0.5:0.95 及 FPS 等指标下进行对比，其结果如表 5 所示。由于 Faster-RCNN 为两阶段检测算法，在 mmdetection 工具箱中无法完成参数及计算量统计；因 Jetpack 版本问题，未对 YOLOv8 做帧率测试。

表 5　改进算法与主流算法的性能对比

方法	Params/M	FLOPs/G	mAP@0.5/%	mAP@0.5:0.95/%	帧率/fps
SSDLite	3.2	0.7	68.6	32.1	29.3
YOLOv3-MobileNetv2	3.7	6.6	82.5	36.9	33.4
YOLOv4-Tiny	5.9	16.2	82.3	42.1	48.0
YOLOv5s	7.0	15.9	86.9	46.9	39.4
YOLOv7-Tiny	6.0	13.1	88.3	44.9	36.9
YOLOv8n	3.0	8.1	89.0	49.0	—
YOLOv8s	11.1	28.5	89.9	50.4	—
YOLOX-Tiny	5.0	7.6	81.7	42.0	30.1
YOLOXs	8.9	13.3	88.8	47.6	22.7
SSD300	25.0	34.8	77.7	36.4	14.5
SSD512	25.7	89.0	84.7	41.4	8.0
YOLOv3	61.6	77.6	89.7	44.0	9.7
YOLOv4	64.0	142.1	86.1	47.7	13.9
YOLOv7	36.5	103.3	90.3	48.8	24.3
RetinaNet	19.8	63.1	85.3	43.5	12.9
FCOS	31.9	78.8	88.5	48.3	9.5
Faster-RCNN	41.2	91.1	89.8	49.0	6.7
YOLOF	42.3	39.4	82.5	40.7	13.8
Cascade-RCNN	69.0	118.8	88.9	48.6	4.2
ESFP-YOLOv5m	2.87	9.6	92.0	50.4	35.0

由表 5 可知，在参数量及浮点运算量这两项指标中，除 SSDLite 与 YOLOX-Tiny 算法外，本文所提出的模型明显优于其他算法；其次，在检测精度方面，本文的算法达到了 92%的 mAP@0.5，其 mAP@0.5:0.95 与最新的单阶段检测算法 YOLOv8s 保持一致；在检测速度方面，YOLOv3-Tiny、YOLOv4-Tiny、YOLOv7-Tiny

及本文算法均达到了实时检测的要求,虽然上述轻量级算法在检测速度上优于本文算法,但在检测精度上,本文所提出的算法要显著地优于其他轻量级算法。因此,从参数量、浮点运算量、检测精度及 FPS 多个角度对比可以看出,本文所提出的算法更加适用于遥感场景的检测需求。

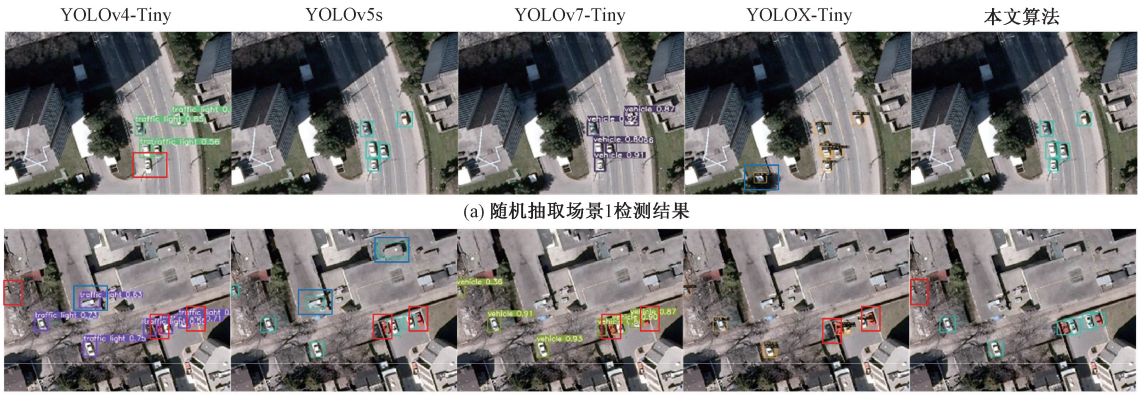
此外,为了进一步说明本文所提出算法的有效性,本文将改进的 ESFP-YOLOv5m 和引言中提及的遥感目标检测领域最新的成果进行对比,如表 6 所示。表中"—"表示原文献中未给出相关数据。mAP@0.5 仅与使用了 NWPU VHR-10 数据集相对比。TFLOPs 均以半精度衡量,其数值来自 NVIDIA 官方提供。

表 6　本文算法与目前遥感目标检测最新成果对比

方法	Params/M	mAP@0.5/%	帧率/fps	实验平台 GPU	算力/TFLOPs
文献[1]	—	93.4	10	NVIDIA TITAN V	24.4
文献[4]	—	82.22	33	NVIDIA Titan xp	24.2
文献[9]	4.4	—	61	NVIDIA GeForce RTX 2080super	22.4
文献[11]	7.19	92.15	28	NVIDIA Tesla T4	16.2
文献[12]	—	—	**98**	NVIDIA GeForce RTX 2080Ti	40.6
文献[13]	—	94.1	—	NVIDIA GeForce RTX3090	71
文献[14]	45.5	**96.0**	—	NVIDIA GeForce RTX3090	71
文献[15]	5.6	—	46	NVDIA GeForce RTX2080	20.2
文献[17]	—	—	58.8	NVIDIA GeForce RTX3090	71
ESFP-YOLOv5m	**2.87**	92.0	35.0	NVIDIA Jetson AGX Xavier	16

由表 6 可知,在检测精度方面,文献[14]精度达到了 96.0%,但其模型参数量却是本研究模型的 15 倍之多。相较之下,本研究的检测精度也达到了可观的 92.0%。在模型轻量化方面,本研究算法的参数量最小仅为 2.87,与文献[4]相比,本文在在显著压缩模型参数量的同时,仍然保持了较高的检测精度。在模型推理速度上,本文使用最大功耗仅 30 W 的边缘计算设备仍然达到了 35 帧,而其他算法虽然达到了较高的帧数,如文献[12]、文献[17]等,但其使用的实验平台均为功耗 200 W 以上的大型设备,说明其算法不适用于小型边缘计算设备。特别是与文献[11]相比,在实验平台算力相近的情况下,本文算法用其模型 0.4 倍的参数量,在检测精度上仅相差 0.15%,而且本文算法的 FPS 比其高 7 帧,有效证明了本文算法的有效性。

为了进一步验证所提出算法的有效性,在使用 NWPU VHR-10 数据集进行实验时,抽取了部分较新的主流轻量级算法进行可视化检测对比,如图 7 所示。同时,为了展示本文算法的优势和先进性,使用部分 google 地球的遥感地图数据,进行逐帧检测。在验证检测准确度的同时,与部分较新的主流轻量级算法对比模型推理速度,如图 8 所示。图 9 为图 8 中图片进行检测时的模型推理速度。



(a) 随机抽取场景1检测结果

(b) 随机抽取场景2检测结果

图 7　主流轻量级算法在 NWPU VHR-10 数据集检测结果对比

由图 7(a)和(b)可以看出,YOLOv4-Tiny 的检测效果较差,出现了大量的漏检目标,对于 YOLOv5s 及 YOLOX-Tiny 而言,不仅出现了漏检的目标,还有明显被误检的物体,而本文的算法在图 7(b)中仅有一个被漏检的边缘物体,无误检的情况,且优于 YOLOv7-Tiny 的检测效果。

由图 8(a)和(c)可以看出,YOLOv4-Tiny、YOLOv5s

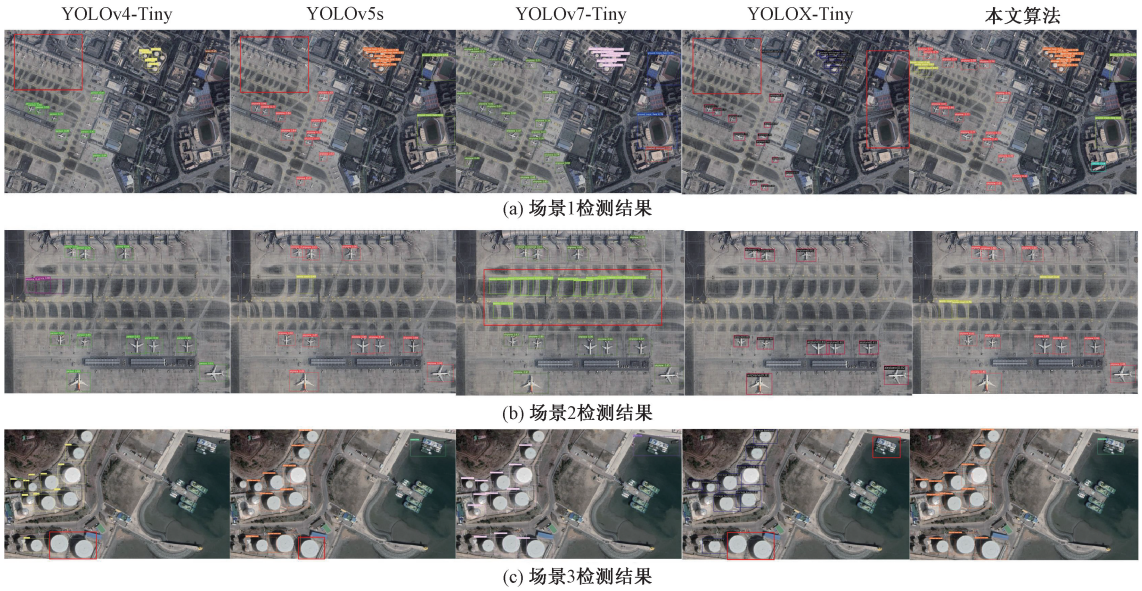


图 8　主流轻量级算法在 google 地球的遥感地图检测结果对比

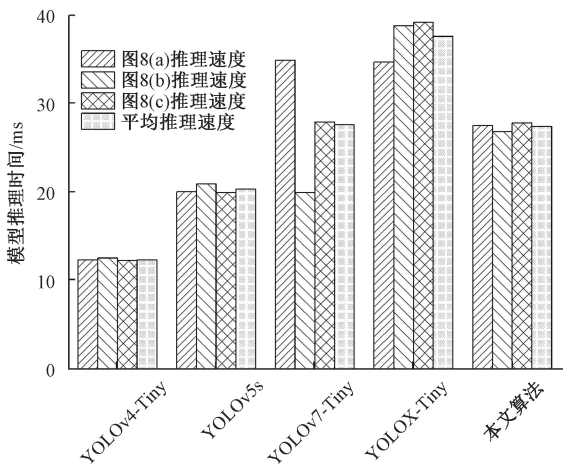


图 9　主流轻量级算法推理速度对比

和 YOLOX-Tiny 与本文算法对比均出现了较多漏检目标。由图 8(b)行可以看出 YOLOv7-Tiny 与本文算法对比出现了大量误检。由图 9 和 8 对比分析可知，YOLOX-Tiny 的检测精度尚可，但模型推理时间明显高于本文算法。YOLOv4-Tiny 和 YOLOv5s 虽然有较快的推理速度，但其检测精度较低。YOLOv7-Tiny 和本文算法的模型推理速度相近，但其出现了大量误检。因此，本文算法与目前主流算法相比，在压缩参数量及浮点运算量的同时仍能保持较优的检测性能及检测效果，可以满足边缘计算设备中的实时遥感影像目标检测需求。

5　结　　论

针对现有目标检测算法无法在边缘计算设备中兼顾检测速度和检测精度的问题，本文基于 YOLOv5 算法、ShuffleNetv2 模型、ECA 注意力机制、FocalEIoU 损失函数及剪枝方法，提出了 ESFP-YOLOv5m 模型。通过替换主干网络的方式，压缩 YOLOv5m 模型的参数量及浮点运算量，减少模型占用的内存资源，强化主干网络的特征提取能力。其次，引入 FocalEIoU 损失函数增强算法的边框回归能力。最后，利用通道剪枝的方法，极限压缩模型的参数量及浮点运算量，进一步提升模型的推理速度。

实验结果证明，本文所提出的方法能有效地压缩 YOLOv5m 模型的参数量及浮点运算量，较好地兼顾了检测精度与检测速度，并在 AGX 边缘计算平台中，满足了实时检测的需求。在后续的研究工作中，将尝试优化剪枝算法，进一步挖掘模型的性能。

参考文献

[1]　范新南，严炜，史朋飞，等. 多尺度深度特征融合网络的遥感图像目标检测[J]. 遥感学报，2022,26(11)：2292-2303.

[2]　HE K，ZHANG X，REN S，et al. Deep residual learning for image recognition[C]. Proceedings of the the 2016 IEEE Conference on Computer Vision and Pattern Recognition，Las Vegas，NV，USA，2016：770-778.

[3]　REN S，HE K，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.

[4]　方青云，王兆魁. 基于改进 YOLOv3 网络的遥感目标快速检测方法[J]. 上海航天，2019,36(5)：21-27.

[5]　HOWARD AG，ZHU M，CHEN B，et al. MobileNets：Efficient convolutional neural networks for mobile vision applications[J]. arXiv preprint，

2017, arXiv:1704-4861.
[6] REDMON J, FARHADI A. YOLOv3: An incremental improvement[J]. ArXiv Preprint, 2018, ArXiv:1804-2767.
[7] 韩要昌, 王洁, 鲁力, 等. 基于卷积核剪枝的遥感目标检测模型压缩方法[J]. 火力与指挥控制, 2021, 46(2): 23-29.
[8] 宋中山, 肖博文, 艾勇, 等. 基于改进 YOLOv4 的轻量化目标检测算法[J]. 2022,45(16): 142-152.
[9] 王成龙, 赵倩, 赵琰, 等. 基于结构化剪枝的遥感飞机检测算法[J]. 电光与控制, 2022,29(6): 37-41.
[10] BOCHKOVSKIY A, WANG C, LIAO H M. YOLOv4: Optimal speed and accuracy of object detection [J]. ArXiv Preprint, 2020, ArXiv: 2004-10934.
[11] 鄢奉习, 徐银霞, 蔡思远, 等. 基于改进 YOLOv5s 算法的遥感图像飞机检测[J]. 计算机工程与设计, 2023,44(9): 2794-2802.
[12] 钱承山, 沈有为, 孙宁, 等. 改进 YOLOv5s 的遥感图像检测研究[J]. 国外电子测量技术, 2022,41(11): 57-66.
[13] 张洋,夏英.多尺度特征融合的遥感图像目标检测方法[J/OL]. 计算机科学, 1-11[2024-02-28]. http://kns. cnki. net/kcms/detail/50. 1075. tp. 20230925. 1342. 094. html.
[14] 吴建成,郭荣佐,成嘉伟,等.注意力特征融合的快速遥感图像目标检测算法[J].计算机工程与应用,2024, 60(1):207-216.
[15] 郎磊, 刘宽, 王东. 基于 YOLOX-Tiny 的轻量级遥感图像目标检测模型[J]. 激光与光电子学进展, 2023, 60(2): 362-372.
[16] ZHENG Z, WANG P, LIU W, et al. Distance-IoU Loss: Faster and better learning for bounding box regression [J]. ArXiv Preprint, 2020, ArXiv: 1911-8287.
[17] 闫钧华, 张琨, 施天俊, 等. 融合多层级特征的遥感图像地面弱小目标检测[J]. 仪器仪表学报, 2022, 43(3): 221-229.
[18] LIU Z, LI J, SHEN Z, et al. Learning efficient convolutional networks through network slimming[C]. Proceedings of the the 2017 IEEE International Conference on Computer Vision, Venice, Italy, 2017: 2736-2744.
[19] MA N, ZHANG X, ZHENG H, et al. Shufflenet v2: Practical guidelines for efficient CNN architecture design[C]. Proceedings of the the 15th European Conference on Computer Vision, Munich, Germany, 2018: 116-131.
[20] WANG Q, WU B, ZHU P, et al. ECA-Net: Efficient channel attention for deep convolutional neural networks[C]. Proceedings of the the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 2020: 11531-11539.
[21] ZHANG Y, REN W, ZHANG Z, et al. Focal and efficient IOU loss for accurate bounding box regression [J]. Neurocomputing, 2022,506: 146-157.
[22] LIU W, ANGUELOV D, ERHAN D, et al. SSD: Single shot multibox detector[C]. Proceedings of the the 14th European Conference on Computer Vision, Amsterdam, Netherlands, 2016: 21-37.
[23] LIN T, GOYAL P, GIRSHICK R, et al. Focal loss for dense object detection[C]. Proceedings of the the 2017 IEEE International Conference on Computer Vision, Venice, Italy, 2017: 2980-2988.
[24] TIAN Z, SHEN C, CHEN H, et al. Fcos: Fully convolutional one-stage object detection [C]. Proceedings of the the 2019 IEEE/CVF International Conference on Computer Vision, Seoul, Korea (South), 2019: 9627-9636.
[25] REDMON J, FARHADI A. YOLO9000: Better, faster, stronger[C]. Proceedings of the the 2017 IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 2017: 6517-6525.
[26] CHEN Q, WANG Y, YANG T, et al. You only look one-level feature[C]. 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Nashville, TN, USA, 2021: 13034-13043.
[27] WANG C, BOCHKOVSKIY A, LIAO H M. YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors [J]. ArXiv Preprint, 2022,ArXiv: 2207-2696.
[28] GE Z, LIU S, WANG F, et al. YOLOX: Exceeding YOLO series in 2021[J]. ArXiv Preprint, 2021, ArXiv: 2107-8430.
[29] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: Unified, real-time object detection[C]. 2016 IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 2016: 779-788.

作者简介

杨洋,硕士研究生,主要研究方向为目标定位及跟踪。
E-mail: 322081104111@stu. suse. edu. cn

宋品德,硕士研究生,主要研究方向为轻量化模型设计。
E-mail: 321085404206@stu. suse. edu. cn

杨思念,硕士研究生,主要研究方向为目标识别。
E-mail: 1903437825@qq. com

曹立佳(通信作者),副教授,主要研究方向为目标识别、无人系统导航与控制。
E-mail: caolj@ suse. edu. cn