



无锚框的轻量级遥感图像目标检测算法

摘要

现有遥感图像目标检测算法存在参数量大、检测速度慢和难以部署于移动设备的问题,为此,本文提出了一种无锚框的轻量级遥感图像目标检测算法.首先设计了 DWS-Sandglass 轻量化模块以降低模型体积,并改进模型激活函数,以确保检测精度.然后引入无参数注意力模块 SimAM,使网络能够专注于更重要的特征信息.最后对无锚框算法的冗余通道进行剪枝操作以减少模型参数量,并通过微调回升精度.在 HRSC2016 数据集上的实验结果表明,与当前主流的无锚框检测算法相比,该算法在检测精度相当的情况下检测速度更快、模型体积更小,更适合在移动设备部署.

关键词

计算机应用;遥感目标检测;轻量级;模型剪枝

中图分类号 TP391.4

文献标志码 A

收稿日期 2023-08-29

资助项目 国家自然科学基金(61702347,62027801);河北省自然科学基金(F2022210007,F2017210161);河北省高等学校科学技术研究项目(ZD2022100,QN2017132);中央引导地方科技发展资金(226Z0501G)

作者简介

张云佐,男,博士,副教授,研究方向为计算机视觉、人工智能、大数据分析、复杂信息系统设计.zhangyunzuo888@sina.com

1 石家庄铁道大学 信息科学与技术学院,石家庄,050043

2 石家庄铁道大学 河北省电磁环境效应与信息处理重点实验室,石家庄,050043

3 石家庄铁道大学 管理学院,石家庄,050043

0 引言

航天遥感技术的飞速发展使高质量的遥感图像逐步增多.因此,遥感图像目标检测技术已成为计算机视觉领域的重要研究方向.传统目标检测方法通常可分为3步:1)通过区域选择器以遍历的方式选出候选区域;2)利用 HOG^[1]、Haar^[2] 等特征提取器进行特征提取;3)使用 AdaBoost^[3]、支持向量机^[4] 等分类器对提取到的特征进行分类.但该方法通过穷举候选框来得到感兴趣区域,不仅时间复杂度高,而且会产生大量窗口冗余.此外,手工设计的特征提取器泛化能力不足以应对遥感图像中的复杂场景和多类目标检测任务.

得益于硬件和算力的发展,基于深度学习的遥感图像目标检测算法逐渐代替传统方法成为了主流.与传统方法相比,基于深度学习的方法因其出色的特征表达和学习能力促进了遥感图像目标检测的发展.目前,基于深度学习的遥感图像目标检测已经在各个领域有广泛的应用^[5],为土地监测^[6]、资源调查^[7]、海洋测绘^[8] 等领域提供了新的解决方案.然而,传统的基于深度学习的遥感图像目标检测方法往往追求精度的提升,一般不考虑模型的复杂性和算法的计算速度.但在实际遥感应用场景中不仅要考虑检测的准确性,还要考虑工程应用中的运行效率和部署难度.随着卷积神经网络在可移动设备上的应用需求不断增加,轻量级网络结构的设计得到了极大的发展^[9].

SqueezeNet^[10] 的提出开辟了轻量化卷积神经网络的发展道路.SqueezeNet 通过使用 1×1 卷积减少参数量来缩小特征图,此外,它减少了 3×3 卷积的输入通道数量,以进一步降低参数量.同时,为了保证输出的准确性,在网络的后期使用池化操作,使卷积层具有更大的特征图.而更大的特征图可以容纳更多的信息,从而提高分类的准确性.尽管 SqueezeNet 在网络参数量上具有优势,但该网络的模型计算量相对较高,在速度方面和其他轻量级卷积神经网络模型相比较慢.

MobileNetV1^[11] 使用深度可分离卷积大大减少了参数量和计算量,并在此基础上使用两个超参数以进一步减少模型体积.在 MobileNetV1 整体网络中,大部分的参数和计算量都被 1×1 卷积所占据,所以,该方法的速度取决于 1×1 卷积.这也为后来的 MobileNetV2^[12] 提供了一个改进方向,即着重减少网络中 1×1 卷积的比例.

为了进一步降低模型的计算复杂度,许多学者都进行了研究.Mo-

bileNetv3^[13] 将 SENet^[14] 注意力机制引入 MobileNetv2 中, 并重新设计了尾部. 华为诺亚提出了 GhostNet^[15], 此网络通过恒等映射和线性变换的方式减少了模型的参数量. Li 等^[16] 提出 kernel-wise 的剪枝方法, 通过移除特定卷积核, 在对网络性能影响较小的前提下使模型尺寸得到相当大的压缩. 在卷积核剪枝的基础上, Liu 等^[17] 对剪枝方法进行了进一步细化, 提出了基于卷积核通道即所谓的 channel-wise 的通道剪枝算法, 通过联合 BN 层 γ 参数训练的方法, 进一步压缩模型体积, 减少网络参数量. 但是这些方法都需要启发式调整参数量.

基于以上分析, 本文提出了一种无锚框的轻量级遥感图像目标检测算法. 此算法基于无锚的 YOLOx 模型, 减少了需要启发式调整的设计参数量. 在此基础上设计了 DWS-Sandglass 轻量化模块以降低模型体积, 并通过改进模型激活函数以提高检测精度. 同时引入无参数注意力模块 SimAM, 使网络能够专注于更重要的特征信息. 最后, 对模型中的冗余通道进行剪枝操作以减少模型参数量.

1 本文所提方法

针对现有模型运行速度慢、体积大, 难以在移动端部署的问题, 本文提出一种基于无锚框的轻量级遥感图像目标检测算法, 网络结构如图 1 所示. 该算法以无锚框检测模型 YOLOx 为基础框架. 在 Backbone 中, 通过卷积和所提出的轻量化模块 DWS-

Sandglass 提取遥感图像目标的浅层纹理特征和深层语义特征, 并将 3 组不同尺度的特征图 (C3~C5) 输入到特征融合模块 FPN 中. SimAM 注意力机制的引入使网络能够在保持模型参数量和复杂度的情况下专注于主干网络提取的重要特征信息. 最后对模型中冗余通道进行剪枝操作进一步压缩模型体积, 并通过微调回升模型精度.

1.1 DWS-Sandglass 模块设计

逆残差模块中的逆残差学习和线性瓶颈层的设计使其取得了良好的效果, 但这样的设计存在问题, 逆残差模块中采用的先升维后降维的结构容易削弱梯度的跨层传播能力, 造成特征从高维空间压缩到低维空间时信息的丢失. 为此, MobileNeXt 中提出了一种新的沙漏瓶颈模块 Sandglass, 它运用了两次深度卷积这样的轻量级单元来编码更多的空间信息, 可以将更多的信息从底层传递到顶层. 另外, 还设计了一个更宽的网络结构缓解梯度混淆的问题. 具体实现方法是采用沙漏形的网络结构, 从通道维度上执行先降采样后升采样的操作, 沙漏瓶颈模块结构如图 2 所示.

Sandglass 模块的运算复杂度 S_c 和参数量 S_p 的计算公式分别如式(1) 和式(2) 所示:

$$S_c = H \times W \times (C_{in} \times 3 \times 3 + C_{in} \times \frac{C_{in}}{T} \times \frac{C_{in}}{T} \times C_{out} + C_{out} \times 3 \times 3), \quad (1)$$

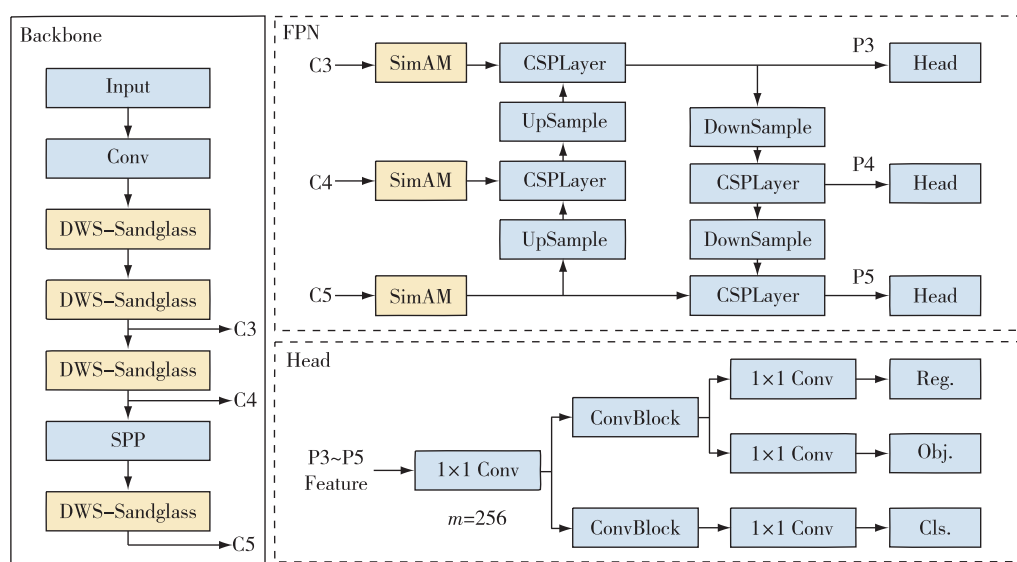


图 1 算法网络结构

Fig. 1 Algorithm network structure

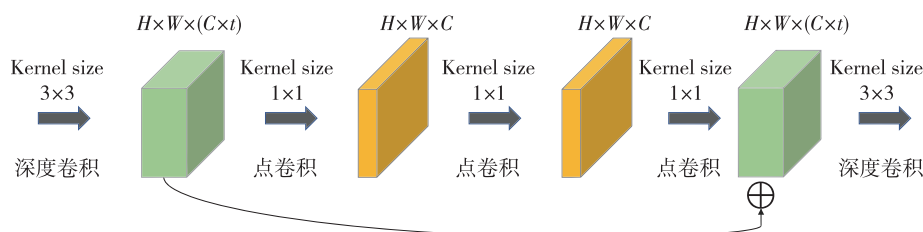


图2 沙漏瓶颈模块结构

Fig. 2 Sandglass bottleneck module structure

$$S_p = C_{in} \times 3 \times 3 + C_{in} \times \frac{C_{in}}{T} + \frac{C_{in}}{T} \times C_{out} + C_{out} \times 3 \times 3. \quad (2)$$

式(1)和式(2)中的 H 和 W 分别表示输入特征图和输出特征图的高和宽, T 表示 Sandglass 模块的缩放系数, C_{in} 和 C_{out} 分别表示输入特征图和输出特征图的通道数.

轻量级骨干网络 MobileNeXt 虽然可以减少网络中的计算量和参数数量,但其特征提取能力不足.本文在 Sandglass 的基础上提出 DWS-Sandglass 模块,如图 3 所示.

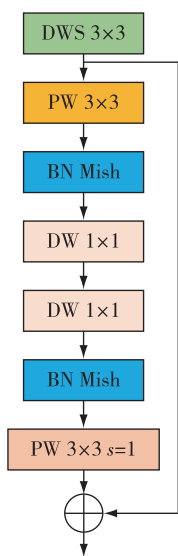


图3 DWS-Sandglass 模块

Fig. 3 DWS-Sandglass module

首先,利用第 1 个步长为 2 的 3×3 深度可分离卷积实现下采样操作,然后,为了减少特征提取过程中的信息丢失,将 2 个逐点卷积的步长设置为 1.为了提高网络的泛化性与准确性,此模块用 Mish 激活函数替换 ReLU 激活函数, Mish 激活函数对负值容许度高的特点容易使神经网络中包含更好的信息.另外,为了增强残差单元中梯度跨层传播的能力,在

DWS-Sandglass 模块中额外构建了跨层交互路径.

1.2 改进的激活函数

基线模型中使用的激活函数是 SiLU,它是 Sigmoid 函数的改进版本,具有无上界、下界、平滑和非单调的特点.非单调激活函数可以与其他激活函数区分开来.其作用可以表示为

$$y(x) = \frac{x}{1 + e^{-x}}. \quad (3)$$

ELiSH 激活函数采用了交叉算子的思想,能够更好地改善信息流,避免梯度消失的问题.在正数部分,它与 SiLU 激活函数具有相同的性质,而在负数部分,它可以看作是 Sigmoid 函数与 ELU 函数的乘积.其作用可以表示为

$$y(x) = \begin{cases} \frac{x}{1 + e^{-x}}, & x \geq 0; \\ \frac{e^x - 1}{1 + e^{-x}}, & x < 0. \end{cases} \quad (4)$$

ELiSH 分片混合激活函数结合了 SiLU、ELU 和 Sigmoid 函数,它可以同时继承 3 个激活函数的基因,对神经网络有更好的分类效果.因此,在本文算法中,使用 ELiSH 激活函数替换 SiLU 激活函数,以实现更好的检测效果.

1.3 无参注意力模块

近年来,注意力机制受到了广泛的关注,越来越多的研究者将注意力机制引入到网络中以提高检测结果的准确性,传统的注意力机制分为通道注意力机制和空间注意力机制.但是,通道注意力机制是针对不同的通道而区分的,空间注意力机制是针对不同的位置而区分的,而人脑中的两种注意往往是协同工作的.基于上述内容, Yang 等^[18]提出一种无参数的三维注意力机制 SimAM,并设计了一种能量函数来计算权重以评估每个神经元的重要性.此外,大多数算子是根据定义的能量函数的解来选择的,因此原始网络的结构不会被改变. SimAM 注意力机制如图 4

所示.

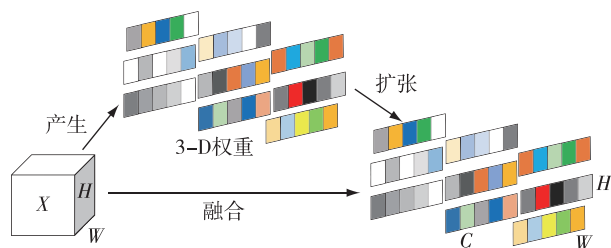


图4 SimAM 注意力机制

Fig. 4 SimAM attention mechanism

不同于常见的通道注意力和空间注意力机制, SimAM 的核心思想是为每个通道和空间位置上的每个像素生成对应的权重, 从而可以同时完成通道和空间之间关联信息的学习. 因此, 在算法中引入了三维注意力机制, 对网络模型的神经元进行建模, 并为最具区分力的神经元分配更高的权重, 以进一步提高细粒度图像分类的性能^[19]. SimAM 通过定义线性可分性的能量函数对每个网络中的每个神经元进行评估. 其中, 目标神经元的能量函数如下所示:

$$e_i^* = \frac{4(\sigma^2 + \lambda)}{(t - \mu)^2 + 2\sigma^2 + 2\lambda}. \quad (5)$$

式中, t 为目标神经元, λ 为超参数, μ 和 σ^2 分别是除 t 以外所有神经元的平均值和方差. 基于单个通道中所有像素遵循相同的分布的假设, 可以计算该通道内所有神经元的平均值和方差, 避免重复计算每个位置的平均值和方差, 降低计算成本. 平均值和方差计算公式分别如式(6)和式(7)所示:

$$\mu = \frac{1}{M} \sum_{i=1}^M x_i, \quad (6)$$

$$\sigma^2 = \frac{1}{M} \sum_{i=1}^M (x_i - \mu)^2. \quad (7)$$

式中, M 表示整体神经元数量, x_i 表示第 i 个神经元的参数.

e_i^* 能量越低, 神经元与整体区域神经元的区分度就越高, 神经元的重要程度也越高. 因此, 神经元的重要性可以通过 $1/e_i^*$ 得到. SimAM 模块最终优化为

$$\tilde{X} = \text{sigmoid}\left(\frac{1}{E}\right) \odot X. \quad (8)$$

式中, X 为输入特征, \odot 为哈达玛积运算, E 为每个通道上的能量函数, sigmoid 函数用于限制 E 中可能出现的过大值, 而输出结果 \tilde{X} 表示增强后的特征.

在图像特征识别中, 通道和空间位置特征非常

重要. 为了有效和全面地评估信道和空间位置特征的重要性, 本文将无参数注意力模块 SimAM 引入到 YOLOx 中. 在不添加模型参数的情况下, 它通过能量函数评估主干提取的特征. 换句话说, SimAM 注意力模块可以找到非常重要的神经元. 改进后的模型可以发现重要特征, 抑制无关特征的干扰, 提高模型的目标定位能力和网络的特征表达能力.

1.4 结构化剪枝

为获得能部署在移动设备上的小体积模型, 可以在训练过程中删减网络模型中的重复参数和连接, 这个过程为剪枝操作, 是压缩模型的主要方法之一. 剪枝可以分为结构化剪枝和非结构化剪枝两种. 与非结构化剪枝相比, 结构化剪枝最明显的特点是它产生的稀疏矩阵是有规则的, 剪枝后网络加速所需的运算量更少^[20]. 因此, 本文采用结构化剪枝策略裁剪网络中的不重要通道.

在卷积神经网络中, BN 层可以抑制内部协变量偏移, 基于 BN 层的缩放因子对网络进行剪枝的优点在于其可以在不引入其他卷积层和参数量的前提下提升剪枝效率. BN 层具有提升训练速度和网络泛化能力的作用, BN 层进行的变换为

$$\begin{cases} \hat{Z} = \frac{z_{in} - \mu_B}{\sqrt{\sigma_B^2 + \varepsilon}}, \\ z_{out} = \gamma \hat{Z} + \beta. \end{cases} \quad (9)$$

训练过程中可学习的缩放因子用 γ 来表示, 而平移参数用 β 表示. 选择 BN 层的缩放因子 γ 作为指标有两个原因: 一是结构化剪枝可以减少参数量和浮点运算量, 二是通道级剪枝方案具有一定的灵活性, 可以适用于大多数卷积神经网络. 式(9)中: B 批次所输入数据的标准差为 σ_B , 均值为 μ_B ; 为防止出现分母为 0 的情况, 在其中加入极小值参数 ε ; 而 BN 层的输入与输出分别用 z_{in} 和 z_{out} 表示. 一般来说, γ 较大的通道被认为对网络更重要, 应避免被裁剪. BN 层输出的激活值大多数不接近 0 并且呈正态分布, 这对稀疏化训练不利. 因此, 为约束通道重要性量化指标 γ 的值, 引入了 L1 正则约束, 改进后的损失函数为

$$L = L_{baseline} + \lambda \sum_{\gamma \in \Gamma} g(\gamma). \quad (10)$$

式中: 基础模型的损失函数为 $L_{baseline}$; λ 是用于平衡损失项的惩罚因子; L1 正则约束项为 $\sum_{\gamma \in \Gamma} g(\gamma)$, 其中 $g(\gamma) = |\gamma|$.

在训练过程中,首先采用稀疏训练对网络进行预训练,然后根据 BN 层缩放因子 γ 的绝对值开发了一种裁剪通道的算法.最后,设定一个较小的学习率来微调裁剪后的网络.经稀疏化训练后,BN 层的缩放因子 γ 通常接近零.因此可以通过适当的剪枝比例来删除不重要的通道.剪枝过程如图 5 所示.

2 实验结果与分析

2.1 实验环境及评价指标

实验环境:操作系统 CentOS7,显卡 NVIDIA Tesla V100S-PCIE-32 GB,处理器 Intel (R) Xeon (R) Gold 6226R CPU @ 2.90 GHz.

评价指标:本文中评价检测模型精度的指标为 AP,用以衡量遥感图像目标的检测效果;使用 Param 来测量模型体积的大小;采用 FPS 衡量检测器的检测速度.实验使用所提出方法的标准模型进行训练,此过程分为冻结训练和解冻训练.其中,冻结训练中学习率设为 0.001,迭代次数设置为 50.

在冻结训练过程中,基本模型主干部分的网络层被冻结.在解冻训练期间将学习率设置为 0.000 1,迭代次数设置为 150,批量大小设置为 8,在此过程中,取消了对模型参数的限制,能够使网络朝着更适合遥感图像目标检测应用场景的方向学习.在稀疏训练中,基准模型的权重参数被设置为初始权重,惩罚系数为 0.001、批量大小为 8、学习率为 0.000 1.根据比例系数的分布,使用 60%的剪枝率来避免严重的精度损失.为了提高实验数据的稳定性和可信度,平衡随机性带来的影响,本文在 5 次实验结果中取平均值.

2.2 实验数据集

本实验采用的 HRSC2016 数据集由西北工业大学于 2016 年发布,是光学遥感船只图像数据集,它包括近岸船只和海上船只,数据集中所有图像来自

6 个著名港口.该数据集中的图像尺寸范围在 $300 \times 300 \sim 1\,500 \times 900$ 之间,一共有 1 061 幅图像,其中,用于训练、验证和测试的图像分别为 436 幅、181 幅和 444 幅.使用训练集和验证集进行训练,使用测试集进行测试.

2.3 消融实验

为了促使模型关注重要特征,抑制无关特征的干扰,提高模型的目标定位能力和网络的特征表达能力,本文在特征融合部分引入了注意力机制指导模型训练.并且,验证了多种注意力模块对模型整体性能的影响.实验结果如表 1 所示.其中:CBAM 的添加使模型精度达到最好,但同时也带来了较多的额外参数量;SimAM 模块精度仅略微低于 CBAM,同时 SimAM 通过能量函数评估特征的方式也使得该模块不会带来过多参数量.因此,最终选取 SimAM 引导模型训练.

表 1 不同注意力模块性能对比

基准模型	SE	CBAM	ECA	SimAM	Param/MB	AP/%
✓					45.21	87.09
✓	✓				48.37	87.36
✓		✓			49.75	88.42
✓			✓		46.83	87.65
✓				✓	45.48	88.34

为了探究不同激活函数对模型精度的影响,本文选取 ReLU、SiLU 和 ELiSH 进行实验,结果如表 2 所示.实验结果表明,ELiSH 可以有效平衡线性和非线性特征的融合,提供更好的梯度传播和模型泛化能力,同时具备高效计算和稳定性.

为了验证各个模块对最终实验结果的影响,实验中以 Param、AP 和 FPS 作为评价标准,通过消融实验评估不同模块对整体模型性能的影响.消融实

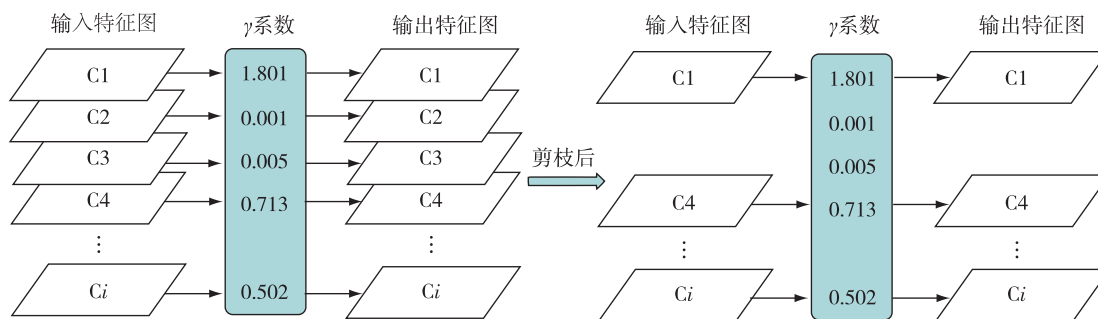


图 5 剪枝过程示意

Fig. 5 Schematic of pruning process

验结果如表 3 所示.

表 2 不同激活函数性能对比

Table 2 Performance comparison between different activation functions

基准模型	ReLU	SiLU	ELiSH	AP/%
✓	✓			86.13
✓		✓		87.09
✓			✓	87.67

表 3 消融实验

Table 3 Ablation experiment

基准模型	DWS-Sandglass	SimAM	ELiSH	Param/MB	AP/%
✓				45.21	87.09
✓	✓			43.37	86.56
✓	✓	✓		43.64	87.43
✓	✓	✓	✓	44.32	88.04

通过在基准模型中添加 DWS-Sandglass 模块,模型体积和检测精度都实现了小幅下降,检测速度得到了提升.接着加入 SimAM 注意力模块,检测精度得到提高,但模型的参数量增加了 0.27 MB.改进激活函数后,实现了 88.04% 的最佳检测精度.综上,相较于基准模型,改进后的算法具有更低的参数量和更高的检测精度.

最后,通过模型剪枝对最终模型进行压缩.在模型剪枝过程中,剪枝率是一个重要的超参数,它决定了保留在模型中的参数比例,较高的剪枝率可以显著减小模型大小和提高推理速度,但可能会导致性能下降.因此,为了探究剪枝率对模型性能的影响,本文测试了剪枝率为 50%、60%、70%、80% 情况下的模型参数和精度.修剪和微调的检测结果如表 4 所示.实验数据表明,随着修剪比例的增加,虽然参数数量会显著减少,但 mAP 也会降低.同时,在微调后,精度也会在一定程度上回升.综合比较和分析,本文最终将剪枝率设定为 60%.经过剪枝处理后,模型的体积显著减少,且由于微调策略,检测精度仅略微下降.

表 4 剪枝和微调对模型性能的影响

Table 4 Impact of pruning and fine-tuning on model performance

剪枝率/%	Param/MB	AP/%	
		微调前	微调后
50	36.21	82.77	87.43
60	27.59	78.21	86.94
70	14.37	66.75	80.31
80	9.80	62.12	74.36

2.4 对比实验

为了满足不同场景的需求,通过调整网络的深度和宽度,导出了两个不同规模的模型,即本文方法-L 和本文方法-S.为了对两种方法进行充分评估,验证其对遥感图像目标的检测能力,将两种方法与当前主流无锚框检测算法进行了对比实验分析,对比算法主要包括 CornerNet、CenterNet、VarifocalNet 等方法,实验结果如表 5 所示.结果表明,与其他遥感图像目标检测算法相比,本文方法在模型体积和检测速度方面均具有显著优势,满足了遥感图像目标准确度和实时检测的要求.

表 5 不同遥感目标检测算法在 HRSC2016 数据集上的性能比较

Table 5 Performance comparison between different remote sensing object detection algorithms on HRSC2016 dataset

算法	主干网络	AP/%	Param/MB	FPS/ (frame·s ⁻¹)
CornerNet ^[21]	Hourglass-104	60.71	150.95	1.9
CenterNet ^[22]	ResNet-18	75.25	16.38	46.6
YOLOX-L ^[23]	CSPDarkNet	87.09	45.21	27.8
VarifocalNet ^[24]	ResNet-50	85.04	32.67	20.8
BBAVectors ^[25]	ResNet-50	86.19	53.43	15.8
SASM ^[26]	ResNet-50	80.40	33.76	20.0
FCOS ^[27]	ResNet-50	78.06	41.08	25.7
本文方法-L	CSPDarkNet	86.94	27.59	35.4
本文方法-S	CSPDarkNet	82.74	15.24	48.5

为进一步说明本文所提算法的有效性,将对对比方法中 Param 和 FPS 的值进行了直观展示,如图 6 所示.可以直观看出,本文方法-S 实现了更低的模型参数量及更快的检测速度.其中:CornerNet 使用的主干网络为 Hourglass-104,其参数量过大,严重影响了网络的检测速度;CenterNet 虽然在 CornerNet 中添加了中心点预测,大大降低了错误检测的概率.然而,其主干网络使用 ResNet-18,特征提取能力不强,检测效果也不理想;VarifocalNet、BBAVectors、SASM 和 FCOS 4 种方法均使用 ResNet-50 作为主干网络,取得了不错的检测效果,但没有达到最优效果;YOLOx-L 虽然取得了最优的检测精度,但检测速度不高,模型体积也高达 45.21 MB;本文方法-L 的平均精度与 YOLOx-L 相当,其在检测速度上也具有明显的优势,模型体积压缩了 39%;本文方法-S 的检测精度虽然只有 82.74%,但基本可以实现遥感场景下目标的准确检测.且其模型体积达到最小体积,便于移动端部署,检测速度最佳,更适合实际应用.

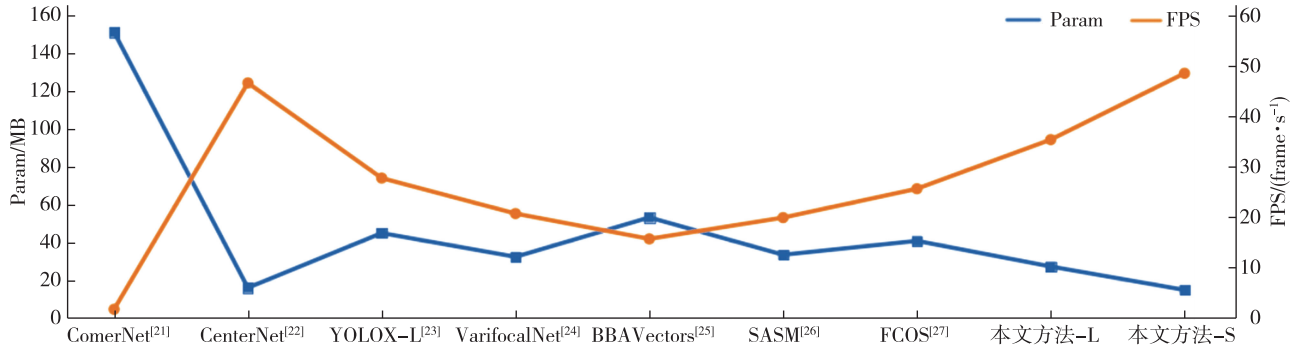


图6 对比算法直观展示结果

Fig. 6 Visual display of comparison results

在 HRSC2016 数据集上的部分检测可视化结果如图 7 所示,可以看出本文所提算法可以精准检测到遥感目标。

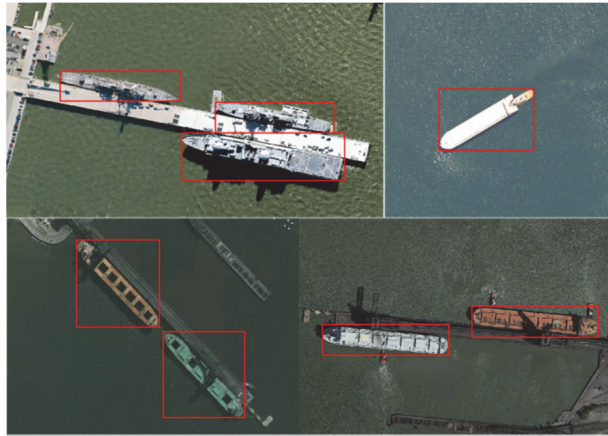


图7 HRSC2016 数据集检测结果

Fig. 7 Results of detection on HRSC2016 dataset

综上所述,本文所提算法在模型参数量和检测速度方面都实现了最优结果,虽然检测精度没有达到最佳值,但并不影响最终检测效果。

3 结论

为了保证模型的检测速度与检测精度,本文引入无锚框的检测算法,并在主干网络中引入轻量化模块 DWS-Sandglass 降低模型的体积,改进激活函数实现更好检测精度.在此基础上,为了获得体积更小的检测模型,使用剪枝策略裁剪模型中的冗余通道.并添加无参数注意力模块 SimAM,在不带来过多计算量和参数量的前提下评估主干提取的特征,提高网络的特征表达能力.最后,通过模型剪枝缩减模型体积,并通过微调回升精度.在 HRSC2016 数据集上的实验结果表明,本文所提算法在确保检测精度的

前提下具有检测速度快、模型体积小特点,更适合部署于移动设备。

参考文献

References

- [1] 王李祺,张成,侯宇超,等.基于深度学习特征融合的遥感图像场景分类应用[J].南京信息工程大学学报(自然科学版),2023,15(3):346-356
WANG Liqi, ZHANG Cheng, HOU Yuchao, et al. Remote sensing image scene classification based on deep learning feature fusion[J]. Journal of Nanjing University of Information Science & Technology (Natural Science Edition), 2023, 15(3): 346-356
- [2] Dalal N, Triggs B. Histograms of oriented gradients for human detection[C]//2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05). June 20-25, 2005, San Diego, CA, USA. IEEE, 2005: 886-893
- [3] Lienhart R, Maydt J. An extended set of Haar-like features for rapid object detection[C]//Proceedings of International Conference on Image Processing. September 22-25, 2002, Rochester, NY, USA. IEEE, 2002. DOI: 10.1109/icip.2002.1038171
- [4] Viola P, Jones M. Rapid object detection using a boosted cascade of simple features[C]//Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. December 8-14, 2001, Kauai, HI, USA. IEEE, 2003. DOI: 10.1109/CVPR.2001.990517
- [5] 张兵.光学遥感信息技术与应用研究综述[J].南京信息工程大学学报(自然科学版),2018,10(1):1-5
ZHANG Bing. A survey of developments on optical remote sensing information technology and applications[J]. Journal of Nanjing University of Information Science & Technology (Natural Science Edition), 2018, 10(1): 1-5
- [6] 黎江,许明慧,张羽.基于 Mask R-CNN 的遥感影像土地分割与轮廓提取[J].南京信息工程大学学报(自然科学版),2021,13(1):116-123
LI Jiang, XU Minghui, ZHANG Yu. Land segmentation and contour extraction of remote sensing image based on Mask R-CNN[J]. Journal of Nanjing University of Information Science & Technology (Natural Science

- Edition), 2021, 13(1): 116-123
- [7] 张雪红, 吴雨阳, 王永俊, 等. 基于热红外和夜间灯光遥感的 GDP 空间化研究: 以福建省为例 [J]. 南京信息工程大学学报(自然科学版), 2021, 13(6): 720-729
ZHANG Xuehong, WU Yuyang, WANG Yongjun, et al. GDP spatialization based on thermal infrared and night light remote sensing: a case study in Fujian province [J]. Journal of Nanjing University of Information Science & Technology (Natural Science Edition), 2021, 13(6): 720-729
- [8] 董超, 刘晶红, 徐芳, 等. 光学遥感图像舰船目标快速检测方法 [J]. 吉林大学学报(工学版), 2019, 49(4): 1369-1376
DONG Chao, LIU Jinghong, XU Fang, et al. Fast ship detection in optical remote sensing images [J]. Journal of Jilin University (Engineering and Technology Edition), 2019, 49(4): 1369-1376
- [9] Wu C P, Wen W, Afzal T, et al. A compact DNN: approaching GoogLeNet-level accuracy of classification and domain adaptation [C] // 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). July 21-26, 2017, Honolulu, HI, USA. IEEE, 2017: 761-770
- [10] Iandola F N, Han S, Moskewicz M W, et al. SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and < 0.5 MB model size [J]. arXiv e-print, 2016, arXiv:1602.07360
- [11] Howard A G, Zhu M L, Chen B, et al. MobileNets: efficient convolutional neural networks for mobile vision applications [J]. arXiv e-print, 2017, arXiv:1704.04861
- [12] Sandler M, Howard A, Zhu M L, et al. MobileNetV2: inverted residuals and linear bottlenecks [C] // 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). June 18-23, 2018, Salt Lake City, UT, USA. IEEE, 2018: 4510-4520
- [13] Howard A, Sandler M, Chen B, et al. Searching for MobileNetV3 [C] // 2019 IEEE/CVF International Conference on Computer Vision (ICCV). October 27 - November 2, 2019, Seoul, Korea (South). IEEE, 2020: 1314-1324
- [14] Hu J, Shen L, Sun G. Squeeze-and-excitation networks [C] // 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). June 18-23, 2018, Salt Lake City, UT, USA. IEEE, 2018: 7132-7141
- [15] Han K, Wang Y H, Tian Q, et al. GhostNet: more features from cheap operations [C] // 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). June 13 - 19, 2020, Seattle, WA, USA. IEEE, 2020: 1577-1586
- [16] Li H, Kadav A, Durdanovic I, et al. Pruning filters for efficient ConvNets [J]. arXiv e-print, 2016, arXiv: 1608.08710
- [17] Liu Z, Li J G, Shen Z Q, et al. Learning efficient convolutional networks through network slimming [C] // 2017 IEEE International Conference on Computer Vision (ICCV). October 22 - 29, 2017, Venice, Italy. IEEE, 2017: 2755-2763
- [18] Yang L X, Zhang R Y, Li L D, et al. SimAM: a simple, parameter-free attention module for convolutional neural networks [C] // 38th International Conference on Machine Learning, 2021: 11863-11874
- [19] 刘丽伟, 侯德彪, 侯阿临, 等. 基于 SimAM-YOLOv4 的自动驾驶目标检测算法 [J]. 长春工业大学学报, 2022, 43(3): 244-250
LIU Liwei, HOU Debiao, HOU Alin, et al. Automatic driving target detection algorithm based on SimAM-YOLOv4 [J]. Journal of Changchun University of Technology, 2022, 43(3): 244-250
- [20] Yu N G, Qiu S, Hu X L, et al. Accelerating convolutional neural networks by group-wise 2D-filter pruning [C] // 2017 International Joint Conference on Neural Networks (IJCNN). May 14 - 19, 2017, Anchorage, AK, USA. IEEE, 2017: 2502-2509
- [21] Law H, Deng J. CornerNet: detecting objects as paired keypoints [J]. International Journal of Computer Vision, 2020, 128(3): 642-656
- [22] Duan K W, Bai S, Xie L X, et al. CenterNet: keypoint triplets for object detection [C] // 2019 IEEE/CVF International Conference on Computer Vision (ICCV). October 27 - November 2, 2019, Seoul, Korea (South). IEEE, 2020: 6568-6577
- [23] Ge Z, Liu S T, Wang F, et al. YOLOX: exceeding YOLO series in 2021 [J]. arXiv e-print, 2021, arXiv: 2107.08430
- [24] Zhang H Y, Wang Y, Dayoub F, et al. VarifocalNet: an IoU-aware dense object detector [C] // 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). June 20-25, 2021, Nashville, TN, USA. IEEE, 2021: 8510-8519
- [25] Yi J R, Wu P X, Liu B, et al. Oriented object detection in aerial images with box boundary-aware vectors [C] // 2021 IEEE Winter Conference on Applications of Computer Vision (WACV). January 3 - 8, 2021, Waikoloa, HI, USA. IEEE, 2021: 2149-2158
- [26] Hou L P, Lu K, Xue J A, et al. Shape-adaptive selection and measurement for oriented object detection [J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2022, 36(1): 923-932
- [27] Tian Z, Shen C H, Chen H, et al. FCOS: fully convolutional one-stage object detection [C] // 2019 IEEE/CVF International Conference on Computer Vision (ICCV). October 27 - November 2, 2019, Seoul, Korea (South). IEEE, 2020: 9626-9635

Lightweight remote sensing image target detection without anchor frame

ZHANG Yunzuo^{1,2} WU Cunyu¹ GUO Wei¹ ZHAO Ning³

1 School of Information Science and Technology, Shijiazhuang Tiedao University, Shijiazhuang 050043, China

2 Hebei Key Laboratory of Electromagnetic Environmental Effects and Information Processing,
Shijiazhuang Tiedao University, Shijiazhuang 050043, China

3 School of Management, Shijiazhuang Tiedao University, Shijiazhuang 050043, China

Abstract The existing remote sensing image object detection algorithms have been frustrated by large parameter quantities, slow detection speed and inability to deploy on mobile devices. Here, we propose a lightweight remote sensing image object detection algorithm without anchor frames. First, a DWS-Sandglass lightweight module is designed to reduce the model volume, and the activation function of the model is improved to ensure detection accuracy. Then, a parameter free attention module SimAM is introduced to enable the network to focus on more important feature information. Finally, the redundant channels of the anchor frame free algorithm are pruned to reduce the number of model parameters, and the accuracy is improved by fine tuning. The experimental results on HRSC2016 dataset show that compared with current mainstream detection algorithms free of anchor frame, the proposed algorithm has faster detection speed and smaller model size, making it more suitable for deployment on mobile devices with comparable detection accuracy.

Key words computer applications; remote sensing target detection; lightweight; model pruning