

DOI:10.19651/j.cnki.emt.2106730

# 样本重采样小目标检测算法的改进<sup>\*</sup>

李睿康 焦新泉 陈建军

(中北大学 电子测试技术国家重点实验室 太原 030051)

**摘要:** 在小目标检测领域中,很多算法以增加模型的复杂度为代价提高了精度,但是带来了较大的计算负担和设备要求。针对模型中复杂度和检测精度之间的矛盾,提出了一种改进的图像金字塔上样本重采样策略算法。该算法只需计算少量的样本数据,并且引入少量参数的轻量化注意力机制模块。实验在 COCO 数据集上进行了训练和测试,重采样策略 mAP 值为 40.6%,引入注意力模块改进后为 42.1%,且引入注意力模块后权重文件大小只增加了 2%。实验结果表明,对样本重采样算法的改进能够在提升检测精度的同时,增加的计算负担较小,验证了轻量化注意力模块的有效性。

**关键词:** 目标检测;图像金字塔;注意力机制;深度学习

**中图分类号:** TP301.6 **文献标识码:** A **国家标准学科分类代码:** 520.20

## Improvement of small target detection algorithm based on sample resampling

Li Ruikang Jiao Xinquan Chen Jianjun

(State Key Laboratory of Electronic Measurement Technology, North University of China, Taiyuan 030051, China)

**Abstract:** In the field of small target detection, many algorithms improve the accuracy at the cost of increasing the complexity of the model, but it brings a large computational burden and equipment requirements. Aiming at the contradiction between complexity and detection accuracy in the model, an improved resampling strategy algorithm on image pyramid is proposed. The algorithm only needs to calculate a small amount of sample data and introduce a lightweight attention mechanism module with a few parameters. In the experiment, training and testing were carried out on the COCO dataset. The resampling strategy mAP value was 40.6%, and the improved value was 42.1% with the introduction of the attention module, and the weight file size only increased by 2% with the introduction of the attention module. The experimental results show that the improvement of the sample resampling algorithm can improve the detection accuracy while reducing the computational burden, which verifies the effectiveness of the lightweight attention module.

**Keywords:** object detection; image pyramid; attention mechanism; deep learning

### 0 引言

近年来,目标检测算法随着深度学习技术的火热发展,从基于手工特征的传统算法转向了基于深度神经网络的检测算法,取得了巨大的突破,出现了很多性能优异的目标检测算法,如 RCNN 系列<sup>[1-3]</sup>、YOLO 系列<sup>[4-5]</sup>、SSD<sup>[6-7]</sup>算法。这些算法在识别精度与速度上都有很好地表现,但是对于小目标检测识别精度并不高,还有很大地提升空间。原因是小目标本身的分辨率很小,其特征信息比中、大目标少很多,导致了特征表达能力较弱。在卷积

神经网络中,尺寸较大目标的语义信息出现在深层网络中,而小目标的语义信息出现在浅层,如果网络较深,经过卷积神经网络的多次下采样,小目标的特征信息在这一过程中逐步的减少、丢失,损失了很大一部分的细节信息。在 Faster RCNN 中小物体有时即使在锚框(anchors)中,也十分容易因为小面积导致交并比 IoU(intersection-over-union)过低,导致漏检。

而现实中的目标检测应用场景以小目标居多,例如民用领域的无人驾驶技术中,要对行人、交通标志、其他车辆实时检测,工地上实时监控施工人员是否佩戴安全帽;军事

收稿日期:2021-05-18

<sup>\*</sup> 基金项目:国家重点研发计划(2018YFF01010500)项目资助

领域中对卫星遥感图像<sup>[6]</sup>、雷达图像中的小飞机、小船只、小车辆检测以判断下一步的行动。

到目前为止,已经有针对小目标检测的一些解决方法。在数据预处理方面可以采用专门针对小目标的数据增强<sup>[9]</sup>,Mosaic数据增强,超分辨率数据增强<sup>[10]</sup>等数据增强方法,这些数据增强方法虽然可以增加数据量,提高模型的鲁棒性,但操作复杂,计算量大,实际情况中并不实用。在多尺度特征融合方面,有传统图像金字塔策略<sup>[11-12]</sup>、FPN网络<sup>[13]</sup>、和SNIP<sup>[14]</sup>与SNIPER<sup>[15]</sup>等方法,这些多尺度方法经过不断下采样,放大并提取了小目标的特征信息,但传统图像金字塔策略有计算量大、训练耗时长的问題,FPN网络、SNIP和SNIPER直接将不同阶段的两个差距较大语义信息的特征直接相加,减弱了多尺度特征的表达能。在注意力机制方面,该机制会分配更多的权重在局部关注的区域,增强了重点区域的信息表达,起到了抑制无用背景或者干扰信息的作用<sup>[16]</sup>。CBAM<sup>[17]</sup>分别从通道和空间两个维度上面计算特征图的注意力图,将注意力图与特征图逐元素相乘,进行特征的自适应学习。SENet<sup>[18]</sup>利用全局信息建立通道之间的交互联系,很大程度地减小了之前模

型的错误率,并且复杂度低,通过迭代学习的方式来获取每个通道的重要程度,依照重要程度增强那些有用的特征并抑制作用不大的特征。ECANet<sup>[19]</sup>在SENet的基础上做出了改进。捕获所有通道之间的依存关系效率不高且不必要,所以ECA模块不进行全通道的交互,而是只考虑了每个通道及其相邻的 $k$ 个通道来捕获跨通道交互以保证效率和有效性,减小了模型的复杂度和参数数量。

本文提出一种改进过的样本重采样策略,以Faster RCNN为基础网络,借鉴了SNIP和SNIPER的思想,利用真实标注框和区域候选网络RPN生成的候选框进行样本的重采样。除此之外,在卷积神经网络中加入了轻量化的通道注意力机制增强小目标特征信息的提取,引入较少参数的同时,能够明显的提高检测精度。在MSCOCO数据集上进行实验和分析。

## 1 模型实现

如图1所示为整体网络模型的结构,主要的改进是对以ResNet101为主干网络的Faster RCNN做了样本重采样处理和特征提取时加入了轻量化的注意力机制。

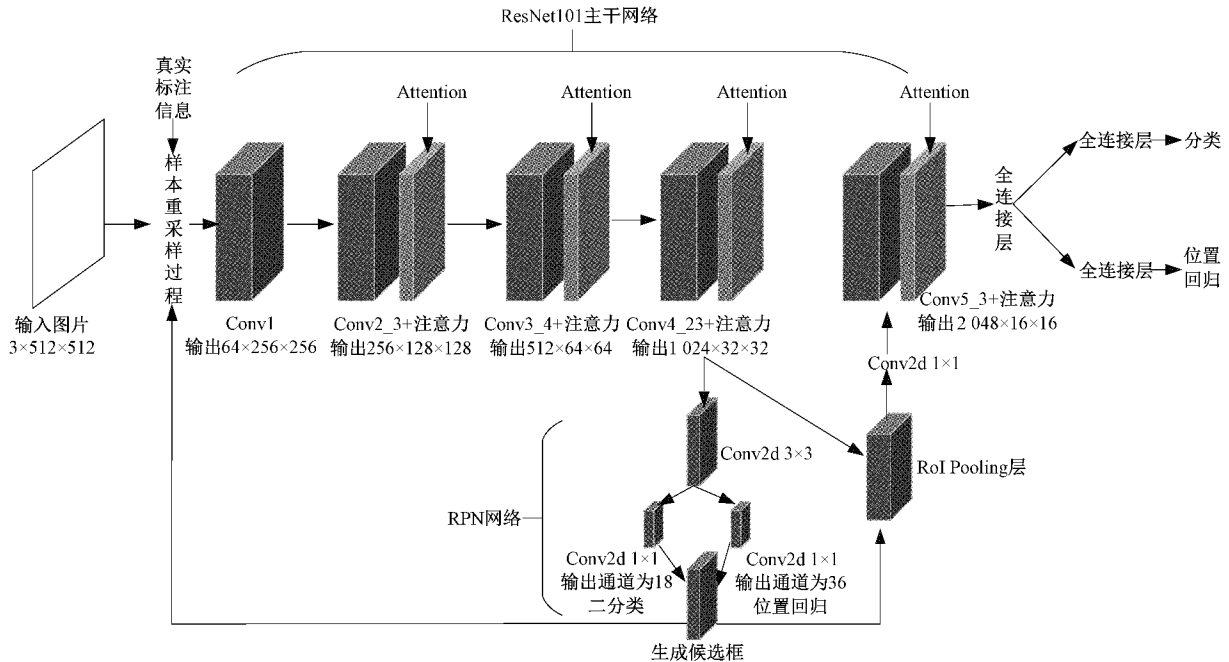


图1 整体网络模型结构

本文的模型采用了Faster RCNN作为基本框架,模型主要分为4个部分。

1)数据重采样加生成模块:生成出那些有助于训练的区域。

2)ResNet101网络模块<sup>[20]</sup>:负责从输入的重采样的数据中提取特征,输出为特征图。

3)注意力机制模块:在ResNet101的中间层中引入通道注意力机制模块,模块输入为经过卷积网络提取信息之

后的特征图,处理后得到了通道注意力特征图。

4)RPN网络和RoI Pooling层模块:对通道注意力特征图生成候选框并修正回归获得精确的位置,判断目标类别。

整体模型在Faster RCNN上进行了对样本的重采样和生成策略的优化,并对卷积神经网络提取到的特征进行注意力权重的分配,增强了重要通道特征的表现力,抑制了作用很小的通道特征。

通过样本重采样模块可以实现避免处理全部的像素,虽然上下文信息是起作用的,但是处理超过一定范围内的上下文信息是无效的,也是没有必要的。轻量化的通道注意力机制与 SNIPER 有着相同的思想,通道之间的交互也不是在全局都能起作用或者是十分低效的。通过两个轻量化的处理,使模型在图像金字塔和跨通道交互上的计算量大大减少。

### 1.1 样本重采样模块的实现

在图像金字塔上进行训练虽然可以提高模型的效果,但是由于在每个尺度上对每个像素都做了处理,极大地增加了计算量,而对图像上每一个像素点都做处理是不必要和十分低效的。实际上,训练只需要对目标本身和其附近的上下文区域,再加上一些容易误检或误报的区域进行即可。目标和附近的上下文区域可以定义为正样本碎片,用来训练目标的准确率,如果只用正区域进行训练,虽然也可以在当前的数据集上获得非常好的表现,但是这属于一种“温室训练”,包含的背景很少,一旦遇到复杂环境或干扰时,错误率增加很多,所以需要一些容易误检或误报的区域,定义为负样本碎片,用来增加一些带有挑战性的难以区分的背景或者干扰。这样,经过重采样的样本包括了目标区域和难度大的背景区域,排除了简单的、易于区分的背景区域,既保证了目标本身的训练,增加了抗干扰性,也同时只处理了图片的对训练有帮助的区域,大大减少了计算量。例如,样本重采样区域面积为  $512 \times 512$ ,图像金字塔上的分辨率分别为  $(480, 512)$ 、 $(800, 1\ 280)$ 、 $(1\ 400, 2\ 000)$ ,在低分辨率只取一个重采样碎片,高分辨率取 5 个碎片,重采样之后的样本像素点为 2 883 584 个,传统图像金字塔上样本的像素点为 4 005 760 个,重采样过后的计算量是传统图像金字塔策略的 71.29%,减少了近 30%。

碎片生成与正负区域的选择:与滑动窗口类似,在每个尺度上产生大小固定的碎片,以一个固定间隔的方式排列。

如图 2 所示,细线边框为真实标注的实例框,其他粗线边框为生成的碎片,正样本碎片集合包含了所有的标注框。



图 2 正样本碎片选择

与 SNIP 相同,在每个尺度上,只有大小合适的真实框才能参与训练,然后从所有生成的碎片中选取包含有效的真实框最多的作为正样本碎片,使每一个有效的真实框都被一个正样本碎片包含,所有的正样本碎片则覆盖全部有效的真实框。实现了在不同大小的尺度上,裁减出了大小相同的碎片,而这个大小相同的碎片中包含的同样一个目标的大小不相同,实现对目标多尺度的重采样。除了目标区域以外,模型还应该能够对背景正分类,降低误检率,所以需要生成一些负区域。借助 RPN 网络生成的候选框,在图像中的某部分如果没有任何候选框生成,则可以认为该部分很大可能是背景区域,训练时可以忽略这部分区域。之后再将被正样本碎片包含的候选框移除,再选择至少包含  $M$  个在有效范围内的候选框的碎片构成负区域样本碎片。这样训练时只需要处理正样本区域加上容易造成误检的难分样本区域,易于区分的背景区域直接忽略,不仅使模型更有针对性了,而且训练速度也大大提高了。前期可以预先训练 RPN 网络,筛选出正负样本区域。之后在进行多尺度归一化的端到端的训练。最终实现了在图像金字塔上进行重采样得到相同较低分辨率的采样区域,减少了训练的计算量,提高了训练的速度,对小目标检测有很大地提高,放大小目标特征的同时,也没有带来更大的复杂度和计算量。

### 1.2 通道注意力模块的实现

与 SNIPER 不使用图像的全部区域进行训练的思想类似,轻量化的通道注意力机制实现局部的跨通道交互,如图 3 所示为轻量化的通道注意力机制模型。具体实现可以先在空间上进行全局全局池化操作,再对通道维度进行一维卷积操作,一维卷积的卷积核的大小可以通过实验的方法对比选择一个最佳值,一维卷积需要进行填充,确保卷积前后通道维度不会发生改变。经过激活函数的处理,最后将通道注意力权重与原本的特征图进行逐元素相乘,得到最终的通道注意力特征图。之后的模块再对新的通道注意力特征图进行位置回归修正和分类判断。在 ResNet101 中有很多的中间层可以影响 ResNet 的性能,所以添加通道注意力机制时可以选择不同的一维卷积核  $k$  进行比较,参照了原文的设置,对较浅层次网络设置较小  $k$  值,较深层次网络设置较大  $k$  值。在 Faster RCNN 中 ResNet101 的 Conv2、Conv3、Conv4、Conv5 添加通道注意力机制,以增强通道之间的交互性,增强重要通道的特征表达,实现时可以将通道注意力模块加入每一个残差块中。最终实现了低模型复杂度、高效通道注意力的一个轻量化的通道注意力机制。

### 1.3 最终分类与位置回归

Faster RCNN 中的 RPN 网络会接收改进过的通道注意力特征图以生成候选框,RPN 网络分为两条路线,其中一条路线会对注意力特征图做  $1 \times 1$  卷积,调整输出通道维度为 18 以对应每个像素点 9 个锚框的正与负,利用

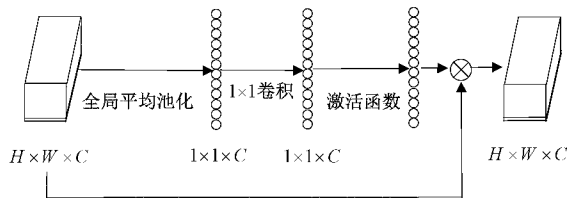


图 3 局部通道注意力模型

softmax 进行二分类判断锚框中是否包含检测物体。另一条路线计算候选框与真实框的映射回归由平移量 ( $dx, dy$ ) 和尺度因子 ( $dw, dh$ ) 体现, 通过平移、缩放方法实现。同样经过  $1 \times 1$  卷积调整输出通道维度为 36 以对应每个像素点 9 个锚框的 4 个回归点。最终 RPN 输出 softmax 分类特征矩阵和坐标回归特征矩阵。

如式(1)~(10)所示,  $x$  为预测框的坐标,  $x_a$  为锚框坐标;  $x^*$  为真实框坐标。通过真实框与锚框坐标计算得出  $d^*$  代表的是真实框与锚框的转化关系, RPN 网络通过学习这个转化关系, 优化和修正预测框和锚框的转化关系  $d$ , 再计算出预测框的真实坐标。

$$d_i = (dx, dy, dw, dh)_i \quad (1)$$

$$d_i^* = (dx^*, dy^*, dw^*, dh^*)_i \quad (2)$$

$$dx = (x - x_a) / w_a \quad (3)$$

$$dy = (y - y_a) / h_a \quad (4)$$

$$dw = \log(w / w_a) \quad (5)$$

$$dh = \log(h / h_a) \quad (6)$$

$$dx^* = (x^* - x_a) / w_a \quad (7)$$

$$dy^* = (y^* - y_a) / h_a \quad (8)$$

$$dw^* = \log(w^* / w_a) \quad (9)$$

$$dh^* = \log(h^* / h_a) \quad (10)$$

Faster RCNN 中的 RoI Pooling 层的输入是固定大小的注意力特征图和 RPN 网络输出的大小不同的候选框, 将候选框映射回注意力特征图, 将映射后的区域划分为与输出维度相同的块, 每个块做最大池化操作, 得到输出大小固定的特征图。分类部分对 RoI Pooling 层输出的候选区域的特征图使用全连接层和 softmax 层计算类别输出概率向量。

## 2 实 验

COCO 数据集是当前目标检测领域中最权威、最标准性的著名数据集。如图 4 所示 COCO 数据集上近 4 年的平均精确值 (average precision, AP) 表现, 其中小目标的 AP 值提升趋势与大、中目标的相同, 但是小目标的 AP 值仍然有较大程度的落后。为了评价模型的精度与速度, 实验在 COCO 数据集上进行训练和测试, 主要性能指标为 mAP 值。mAP 是从 0.5 以 0.05 的增量递增到 0.95 的 10 个 IoU 的平均精度之和取平均作为平均精度的平均值, 该指标最能体现网络的性能。APS 则是目标的面积小于  $32 \times 32$  的 mAP 值。

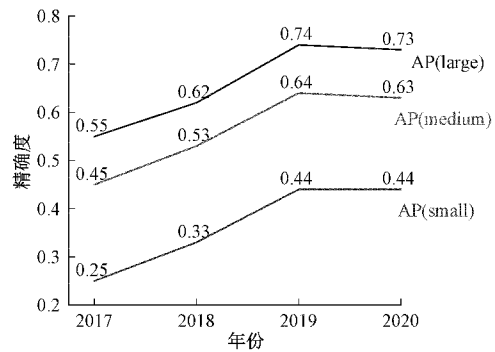


图 4 COCO 数据集近四年最高 AP 值

实验使用显卡 NVIDIA GeForce GTX 1080Ti, 显存为 11 G, 基于 Pytorch 框架实现神经网络。实验中一些重要的超参数及其对应的值如表 1 所示。

表 1 重要超参数的值

超参数	值
前 100 轮学习率	0.015
迭代 100 次之后学习率	0.002
batch-size	2
迭代轮数	300
多尺度尺寸及对应的有效真实框面积范围	(480, 512) (0, 80 <sup>2</sup> ) (800, 1 280) (32 <sup>2</sup> , 150 <sup>2</sup> ) (1 400, 2 000) (120 <sup>2</sup> , ∞)
碎片区域面积	512 × 512
正样本碎片的数量	5
负样本碎片的数量	2
RPN 网络预训练的迭代轮数	7

### 2.1 RPN 网络预训练

正样本碎片的生成只需要真实的标注信息, 负样本碎片生成需要 RPN 网络训练生成一系列的候选框, 负样本碎片会从这些候选框中排除真实实例区域之后再选择部分区域参与训练。如图 5 所示为模型训练流程, 训练过程中可以选择是否使用负样本碎片训练。在 COCO 数据集上训练时, 首先选择不使用负样本碎片, 在训练过程中将 RPN 网络生成的候选框位置信息保存, 在之后需要使用负样本碎片的训练再使用。实验中, 可以先训练较少的 7 轮迭代, 生成一个粗略训练、精度一般的 RPN 网络, 能生成准确率不高但是召回率还行的候选框, 这些候选框有两个作用, 其一是排除易分背景区域, 因为未被 RPN 网络选中的区域大概率是不存在目标的。另一个作用是排除正样本, 然后可以选择剩下的部分候选框作为难分样本区域参与训练。

将预训练的 RPN 网络权重保存下来, 在之后的样本重采样过程使用以获得负样本碎片参与训练。

为了验证使用难分背景区域是否对模型性能产生影

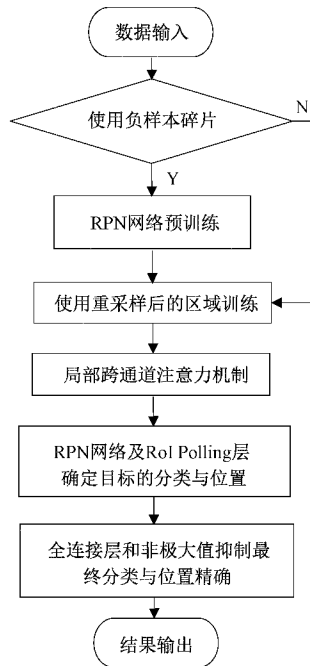


图 5 模型训练流程

响和预训练 RPN 网络在实验中对是否使用负样本碎片训练做对比,训练时不引入通道注意力机制。如表 2 所示,不使用负样本碎片时,召回率与使用负样本区域的大致相同,因为召回率计算只与正确识别到的正样本有关,与负样本关系不大。不使用负样本碎片进行训练时得到的 mAP 略低于使用负样本训练,由此验证了一些背景负样本区域参与训练可以提高模型的检测精度。

## 2.2 添加通道注意力机制

ResNet 中 Conv2、Conv3、Conv4、Conv5 层添加了通道注意力机制做出改进,其中计算通道注意力权重的一维卷积核为别设置为 3、3、3、5。为了保证卷积前后输入输出的形状一致,需要设置填充分别为 1、1、1、2。

如表 3 所示,使用添加了注意力机制的模型训练时, mAP 比原来高了 1.5%,其中小目标的  $AP_s$  高了 1.4%。训练过程中,模型的整体精度有所上升,但召回率实际比原来略有下降,原因可能是添加注意力机制后对提取的特征图做出改进后提前训练好 RPN 网络需要对这部分特征图重新学习和优化,对正样本区域与背景负区域的提取造成了影响。另外一个可能的原因是添加注意力机制之后,

表 2 使用负样本训练效果

方法	AP	AP <sup>50</sup>	AP <sup>75</sup>	%		
				AR max=1	AR max=10	AR max=100
使用负样本碎片	40.6	58.9	43.1	33.2	52.7	55.6
不使用负样本碎片	38.1	56.8	41.2	32.5	53.2	56.5

表 3 使用注意力机制效果

方法	AP	AP <sup>50</sup>	AP <sup>75</sup>	%		
				$\Delta P_s$	$\Delta P_M$	$\Delta P_L$
使用通道注意力	42.1	63.2	45.5	24.1	46.3	56.8
不使用通道注意力	40.6	58.9	43.1	22.7	44.2	56.3

整体模型的学习率与注意力机制的学习率不匹配,注意力机制在迭代过程中学习的比较慢,还处于欠拟合状态。

整体 AP 值的提高是在每个不同阈值下 IoU 的精度取平均而提高,由表 2 和 3 可知这些不同阈值下的 IoU 下的精度提高是不同的,表示当前算法的改进是对某一 IoU 值的效果最为突出。

## 2.3 推理与推理时间

如表 4 所示,3 种不同的尺度在推理时所需的时间,其中检测(480,512)尺寸的图像需要 0.34 s。

表 4 不同尺度推理时间

尺度	单次时间/s	单次处理数目/张
(1 400,2 000)	1.83	1
(800,1 280)	1.04	1
(480,512)	0.34	2

为了保证效率,在推理较高分辨率的图像时,选择每次处理一张图像,一方面是因为 GPU 显存有限,高分辨率占用较高的显存;另一方面是使用 CPU 来提取图片,为了提高 CPU 利用率、平衡提取效率和处理效率,单次处理的图像数日不宜过高。

如图 6 和 7 所示为是否使用通道注意力的对比结果,使用通道注意力模型的检测结果显示处在图像边缘位置上和被部分遮挡的鸟类都已成功识别,且被人物所遮挡的背包也识别出了。

## 2.4 模型的复杂度和性能

为了验证对样本重采样算法的改进能够改善模型复杂度与性能的矛盾,实验采用了 3 种不同的算法模型,分别是样本重采样算法、样本重采样加 SENet 注意力机制和样本重采样加局部通道注意力机制。

如表 5 所示,以模型权重文件的大小作为复杂度的参考,以模型的 mAP 值作为性能的参考,其中加入局部通道



图6 不使用通道注意力检测



图7 使用通道注意力检测

注意力模块后只额外增加了 7.39 MB 的权重文件大小, 占不加入局部通道注意力时权重文件大小的 2%, 但 mAP 值与 SENet 只差 0.1%, 较样本重采样算法提高了 1.5%, 可以得出经过轻量化注意力机制的模块提升了模型性能的同时并没有带来太多的复杂度。

表5 模型复杂度与性能

模型	模型权重文件大小/MB	mAP/%
样本重采样	361.06	40.6
+SENet	381.37	42.2
+局部通道注意力	368.45	42.1

### 3 结 论

为了增强小目标检测精度的同时, 减少图像金字塔上的计算量与计算负担, 在样本重采样策略的基础上添加了通道注意力机制, 平均检测精度提高了 1.5%, 在实际应用中证明了注意力机制的有效性。高效的局部通道注意力机制除了可以提高深度卷积神经网络的性能, 在未来还可以将该机制应用更多的网络之中, 一是可以结合重采样过程, 高效提取目标区域, 进一步减少计算量, 二是将该通道注意力机制与空间注意力机制结合。

### 参考文献

- [1] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]. IEEE Conference on Computer Vision and Pattern Recognition, Columbus, 2014:580-587.
- [2] GIRSHICK R. Fast R-CNN[C]. IEEE International Conference on Computer Vision (ICCV), Santiago, 2015:1440-1448.
- [3] REN SH 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.
- [4] REDMON J, FARHADI A. Yolov3: An incremental improvement[J]. ArXiv Preprint, 2018, ArXiv:1804.02767.
- [5] BOCHKOVSKIY A, WANG C Y, LIAO H Y. YOLOv4: Optimal speed and accuracy of object detection[J]. ArXiv Preprint, 2020, ArXiv:2004.10934.
- [6] LIU W, ANGELOV D, ERHAN D, et al. SSD: Single shot multibox detector [C]. European Conference on Computer Vision, 2016:21-37.
- [7] FU C, LIU W, RANGA A. DSSD: Deconvolutional single shot detector [J]. ArXiv Preprint, 2017, ArXiv:1701.06659.
- [8] 化嫣然, 张卓, 龙赛, 等. 基于改进 YOLO 算法的遥感图像目标检测[J]. 电子测量技术, 2020, 43(24):87-92.
- [9] KISANTAL M, WOJNA Z, MURAWSKI J, et al. Augmentation for small object detection [J]. ArXiv Preprint, 2019, ArXiv:1902.07296.
- [10] NOH J, BAE W, LEE W, et al. Better to follow, follow to be better: Towards precise supervision of feature super resolution for small object detection[C]. IEEE International Conference on Computer Vision (ICCV), Seoul, Korea (South), 2019: 9724-9733.
- [11] LIU Z, GAO G, SUN L, et al. IPGNet: Image pyramid guidance network for small object detection [C]. IEEE Conference on Computer Vision and Pattern Recognition Workshops, 2020: 4422-4430.
- [12] HE K, ZHANG X, REN S, et al. Spatial pyramid pooling in deep convolutional networks for visual recognition [C]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015: 1904-1916.
- [13] LIN T, DOLLÁR P, GIRSHICK R, et al. Feature pyramid networks for object detection [C]. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017: 936-944.
- [14] SINGH B, DAVIS L S. An analysis of scale

- invariance in object detection-SNIP [C]. IEEE Conference on Computer Vision and Pattern Recognition, 2018: 3578-3587.
- [15] SINGH B, NAJIBI M, DAVIS L S. SNIPER: Efficient multi-scale training[C]. Advances in Neural Information Processing Systems, 2018: 9310-9320.
- [16] YIN W, SCHUTZE H, XIANG B, et al. ABCNN: Attention based convolutional neural network for modeling sentence pairs [J]. Transactions of the Association for Computational Linguistics, 2016, 6(4): 259-272.
- [17] WOO S, PARK J C, LEE J Y, et al. CBAM: Convolutional block attention module [C]. European Conference on Computer Vision, 2018: 3-19.
- [18] HU J, SHEN L, SUN G. Squeeze-and-excitation networks [C]. IEEE Conference on Computer Vision and Pattern Recognition, 2018: 7132-7141.
- [19] WANG Q, WU B, ZHU P, et al. ECA-Net: Efficient channel attention for deep convolutional neural networks [C]. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2020: 11531-11539.
- [20] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition [C]. IEEE Conference on Computer Vision and Pattern Recognition, 2016: 770-778.

### 作者简介

李睿康, 硕士研究生, 主要研究方向为嵌入式开发、深度学习、计算机视觉。

E-mail: 605430725@qq.com

焦新泉(通信作者), 博士, 教授, 主要研究方向为微纳传感及测试技术研究。

E-mail: jiaoxinquan@nuc.edu.cn

陈建军, 博士, 讲师, 主要研究方向为动态测试技术、智能传感器、机器视觉、嵌入式系统开发。

E-mail: cjj@nuc.edu.cn