

控制与决策

Control and Decision

改进金字塔和跳跃连接的YOLOv5目标检测网络

刘建骐, 闫河, 王潇棠, 李梦雪

引用本文:

刘建骐, 闫河, 王潇棠, 李梦雪. 改进金字塔和跳跃连接的YOLOv5目标检测网络[J]. *控制与决策*, 2023, 38(6): 1730–1736.

在线阅读 View online: <https://doi.org/10.13195/j.kzyjc.2021.1414>

您可能感兴趣的其他文章

Articles you may be interested in

[YOLOv5预测边界框分簇自适应损失权重改进模型](#)

Enhanced self-adaptive loss weight YOLOv5 model based on predicted bounding boxes in clusters

控制与决策. 2023, 38(3): 645–653 <https://doi.org/10.13195/j.kzyjc.2021.1612>

[基于浅层定位的动态细化目标检测网络](#)

Dynamic refinement networks for object detection based on shallow localization

控制与决策. 2023, 38(1): 49–57 <https://doi.org/10.13195/j.kzyjc.2021.0700>

[基于注意力特征融合的无人机多目标跟踪算法](#)

UAV multi-target tracking algorithm based on attention feature fusion

控制与决策. 2023, 38(2): 345–353 <https://doi.org/10.13195/j.kzyjc.2021.1098>

[基于两阶段深度网络的输电线路异常目标检测方法](#)

Transmission line abnormal object detection method based on deep network of two-stage

控制与决策. 2022, 37(7): 1873–1882 <https://doi.org/10.13195/j.kzyjc.2020.1840>

[多目标小尺度车辆目标检测方法](#)

Multi-target and small-scale vehicle target detection method

控制与决策. 2021, 36(11): 2707–2712 <https://doi.org/10.13195/j.kzyjc.2020.0635>

改进金字塔和跳跃连接的 YOLOv5 目标检测网络

刘建骐, 闫河[†], 王潇棠, 李梦雪

(重庆理工大学 两江人工智能学院, 重庆 401135)

摘要: 针对 YOLOv5 网络模型的金字塔结构不能有效捕捉目标的跨尺度特征信息和存在梯度消失现象以及易导致目标检测精确率不高的问题, 利用 FPT (feature pyramid transformer) 的注意力机制有效提取目标的跨尺度特征信息, 把原 YOLOv5 网络模型中的 FPN (feature pyramid network) 和 PAN (path aggregation network) 结构替换为 FPT, 在 FPT 结构的两端加入跳跃连接 (skip connection) 并引入新的 Mish 激活函数, 从而提出一种改进金字塔和跳跃连接的 YOLOv5 目标检测网络模型 YOLO FS. 在 PASCAL VOC 和 MS COCO 数据集上的对比实验结果表明, 基于 YOLO FS 网络的目标检测在平均检测准确率、召回率和 F1 值上均有明显提升.

关键词: 目标检测; YOLOv5; FPT; 跳跃连接; Mish 激活函数

中图分类号: TP391

文献标志码: A

DOI: 10.13195/j.kzyjc.2021.1414

开放科学(资源服务)标识码(OSID):



引用格式: 刘建骐, 闫河, 王潇棠, 等. 改进金字塔和跳跃连接的 YOLOv5 目标检测网络[J]. 控制与决策, 2023, 38(6): 1730-1736.

YOLOv5 object detection network with improved pyramid and skip connection

LIU Jian-qi, YAN He[†], WANG Xiao-tang, LI Meng-xue

(Liang Jiang School of Artificial Intelligence, Chongqing University of Technology, Chongqing 401135, China)

Abstract: The pyramid structure of the YOLOv5 network model can not effectively capture the cross-scale feature information of the object and the gradient disappears which causes the problem that the accuracy of object detection is not high. The attention mechanism of FPT (feature pyramid transformer) can be used to extract the cross-scale feature information of the object, and then the FPN (feature pyramid network) and the PAN (path aggregation network) in the original YOLOv5 network model are replaced by the FPT. Skip connection is added at both ends of the FPT structure and an improved Mish activation function is introduced. Thus, an improved object detection network model, namely YOLO FS, is proposed. The experimental results on PASCAL VOC and MS COCO datasets show that the average detection accuracy, recall and F1-score of object detection based on the proposed network are significantly improved.

Keywords: object detection; YOLOv5; FPT; skip connection; mish activation function

0 引言

目前, 基于深度学习的目标检测方法在计算机视觉领域中十分活跃^[1]. 文献[2]提出了一种经典的端到端目标检测模型 YOLO, 它采用网格化的方式预测整体图像代替传统检测方法的候选框选取操作, 具有较快的检测速度, 能够满足实时性目标检测需求^[3]. YOLO 模型开拓了将目标检测视为回归任务的思想, 在目标检测领域具有里程碑意义, 但其存在目标检测精确率和召回率较低的问题^[4-5]. 针

对 YOLO 系列的检测精确度较低问题, 文献[6]提出的 YOLOv4 采用 CSP (cross-stage-partial) 结构提取特征, 引入 SSP (spatial pyramid pooling)、路径聚合网络 (PAN)^[7] 等结构, 通过将高层和低层特征信息融合再提取特征, 结合 Mosaic 数据增强方法, 提升了目标检测精度. 为了进一步提高目标检测精确度, 文献[8]提出的 YOLOv5 在 YOLOv4 网络的基础上引入自适应候选框 (anchor)^[9] 和自适应图片缩放, 增强数据并提高网络检测鲁棒性. YOLOv5 引入 Focus 结构并设

收稿日期: 2021-08-11; 录用日期: 2022-02-25.

基金项目: 国家重点研发计划“智能机器人”重点专项项目(2018YFB1308602); 国家自然科学基金面上项目(61173184); 重庆市自然科学基金项目(cstc2018jcyjA2328, cstc2018jcyjAX0694).

责任编辑: 张文安.

[†]通讯作者. E-mail: cqyanhe@163.com.

计两种CSP结构,加强了网络特征融合能力,保留更丰富的特征信息;YOLOv5采用GIOU_loss函数^[10],提高网络检测精确度和收敛稳定性;YOLOv5还通过引入网络深度、宽度两个缩放系数,将网络按照层数和通道数的不同划分为4种网络模型,分别为YOLOv5s、YOLOv5m、YOLOv5l、YOLOv5x. 4种网络模型体积依次增大,目标检测精确度也依次递增,这种结构令网络更加灵活. YOLOv5在其Neck部分使用了特征金字塔网络(FPN)结构^[11]和PAN结构,在多个尺度特征图之间传递语义特征和位置信息,目标检测精确率有一定提升,但是这两种结构提取跨尺度特征和同一尺度的相似特征能力不强,在预测目标时鲁棒性较差,检测精确度还有提升的空间. 本文通过引入特征金字塔transformer(FPT)结构^[12]替换原YOLOv5中的FPN结构和PAN结构,采用FPT结构中3种transformer结构,利用其注意力机制聚焦目标跨尺度特征以及提取同尺度相似特征,提升目标检测的精确度和鲁棒性;通过在FPT结构两端加入跳跃连接结构,解决网络结构加深时带来的梯度消失问题,并有效地传递显著性特征,提升目标检测精确度;同时引入Mish激活函数^[13]代替原YOLOv5的Backbone结构中的ReLU(rectified linear units)激活函数,提升网络的目标检测精确度;最后在PASCAL VOC数据

集^[14]和MS COCO数据集^[15]中进行训练、测试,并与YOLOv5网络和YOLOv4网络进行对比实验,结果表明YOLO FS网络在目标检测的精确度上具有明显的提升.

1 YOLOv5网络结构

1.1 YOLOv5

文献[8]提出的YOLOv5的特点是检测速度快,网络结构灵活,具有YOLOv5s、YOLOv5m、YOLOv5l、YOLOv5x四种网络结构. YOLOv5的4种网络结构只有网络深度和层通道数不同,网络中的其他结构相同,其中YOLOv5s网络结构在4种结构中最小,速度最快,因此本文选用YOLOv5s网络结构. YOLOv5网络的具体结构如图1所示. YOLOv5的网络结构分为输入端、Backbone、Neck、Prediction四个模块. Backbone模块中采用了Focus结构、CSP结构、空间金字塔池化(SPP)结构;Neck模块采用了FPN结构和PAN结构. YOLOv5采用了Mosaic数据增强方法,通过将图片随机缩放、随机裁剪、随机排布的方式,达到增强数据集维度的目的. Backbone模块中的Focus结构通过切片操作,增强了网络对图片的特征提取能力. YOLOv5网络重新设计了CSP结构,将其分为两种,一种主要应用于Backbone模块中,另一种结构应用于Neck模块.

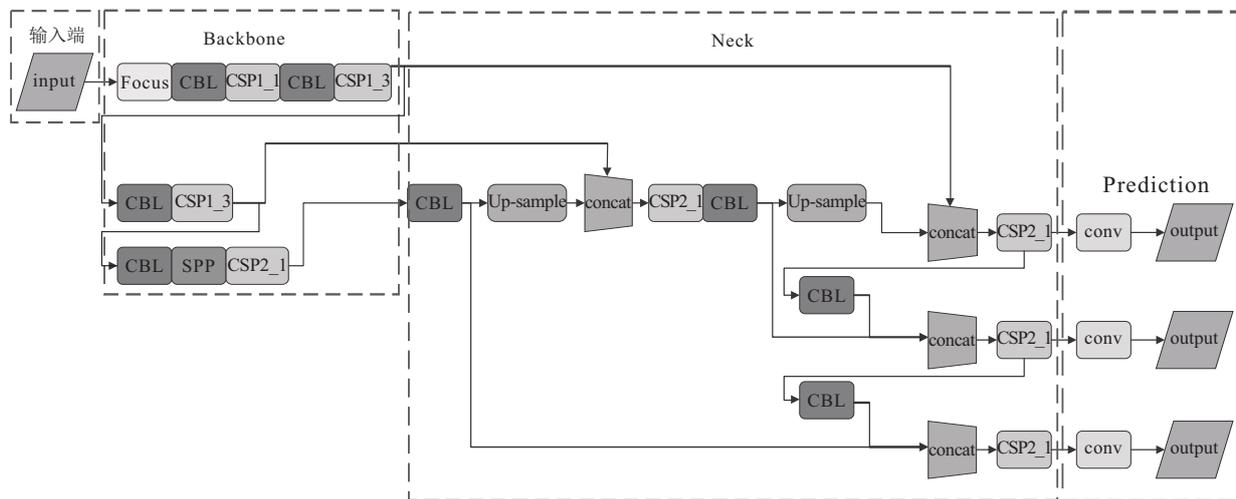


图1 YOLOv5网络模型结构

1.2 FPN与PAN结构

FPN结构通过下采样方法自顶向下传递高层的语义特征,增强了网络的语义信息. PAN结构是在FPN结构之后增加的一个自底向上的特征金字塔,将FPN结构输出的特征图自底向上传递特征信息,得到不同尺度的特征图. FPN结构包含3种连接:自

底向上连接、自顶向下连接和横向连接. 通过卷积神经网络有效地提取输入图像中各种维度的特征,尤其是可以提高小目标检测效果. 但是FPN结构只增强了网络获取语义信息的能力,而没有关注位置信息的传递. PAN结构通过自底向上的金字塔结构将网络底层位置信息向上传递,增强了网络获取位置信息的

能力,缩短了底层与高层之间的联系路径. YOLOv5网络的FPN结构和PAN结构将原本的PANet中add操作替换为concat函数,增强了网络特征融合能力. FPN与PAN结构如图2所示.

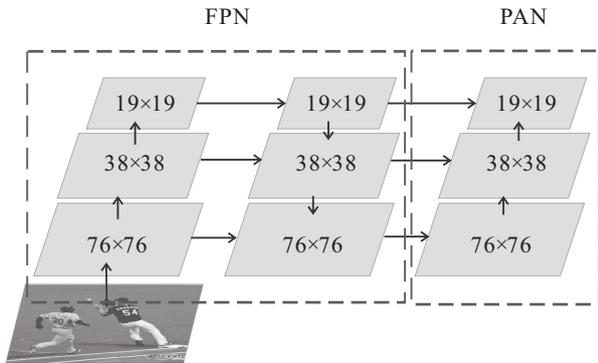


图2 FPN与PAN结构

2 YOLO FS网络结构

2.1 FPT网络结构

从FPN结构和PAN结构可以看出,结合不同尺度特征图之间的语义信息和位置信息,可以增强多尺度特征检测能力. 为了加强这种提取跨尺度特征能力以及增强提取同一尺度的相似特征能力,本文采用文献[12]提出的FPT网络结构. FPT基于transformer的思想,使用Non-local结构^[16]与自注意力机制,实现特征图跨空间和跨尺度的非局部交互. Non-local结构如图3所示,将特征图作为输入,经过3个 $1 \times 1 \times 1$ 的卷积核,将通道数缩小为原来的一半,再经过concat操作与一个 $1 \times 1 \times 1$ 的卷积核,所得特征图尺度与原来的一致,最后与输入的特征图进行add操作,得到完整的特征图输出.

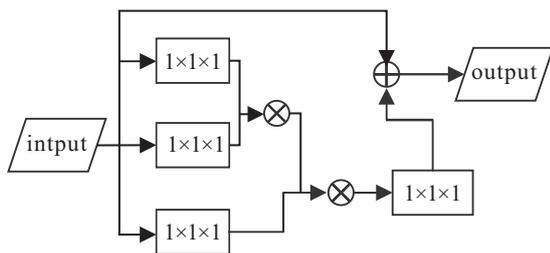


图3 Non-local结构

FPT网络整体结构包括5个模块,即输入层的特征金字塔结构模块、进行跨尺度特征交互的transformer模块、re-arrange模块、concat模块以及输出层. FPT网络结构输入为3种尺度特征图构成的特征图金字塔,输出为相同尺寸的特征图金字塔.

transformer模块包括了3种transformers结构,分别是ST(self transformer)结构、GT(grounding transformer)结构、RT(rendering transformer)结构^[12].

ST结构通过进行同等尺度特征图内的非局部交互,根据相似度函数和归一化函数计算出特征图内像素块间的权重,从而获取特征图上共现的目标的特征. GT结构采用自上而下的方式进行不同尺度特征图之间的交互,采用欧氏距离的负值计算相似度,将较小尺度的特征图的语义信息传递给较大尺度的特征图,是一种跨尺度局部交互方式,变换后的特征图与较小尺度的特征图尺度相同;RT结构则是采用自下而上的方式将较大尺度特征图中的位置信息传递给较小尺度特征图,通过全局平均池化计算出高层特征权重,然后利用 3×3 卷积和加和的方法传递特征信息.

在得到经过transformer模块转化的特征图后,通过re-arrange模块重新聚合相同尺度的特征图,再采用concat模块加入输入特征金字塔中对应尺度的特征图,经过输出层卷积变换,得到输出层特征图金字塔.

本文对FPT结构末端卷积层进行改进,将原FPT结构中输出端的卷积层中的 3×3 卷积层替换为 1×1 卷积层、 3×3 卷积层、 1×1 卷积层串联的结构,本文中称为Conv Block结构. Conv Block结构通过在 3×3 卷积层前后加入激励层,为卷积层添加了非线性激励,提升了网络的表达能力.

2.2 跳跃连接结构

ResNet网络模型^[17]中的跳跃连接结构(skip connection, S)已经证实其有效性,可以解决网络层数增加时由于梯度下降导致的精度降低问题^[18]. 如图4所示,本文在FPT结构的两端加入一个长跳跃连接结构,将通过FPT结构、Conv Block结构和上采样的多尺度特征与输入FPT结构的特征相结合. 跳跃连接结构通过传递来自相同尺度网络的底层特征,增强网络对于底层特征的聚焦,提高网络模型目标检测的精确率.

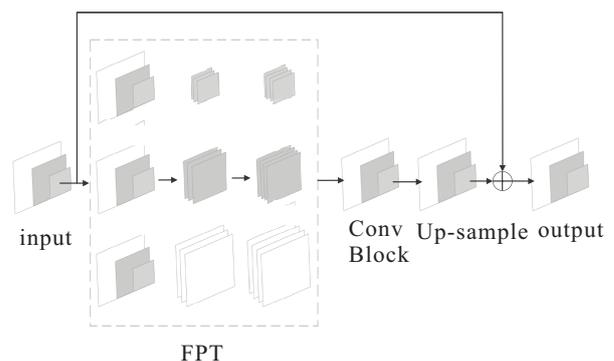


图4 跳跃连接结构

2.3 Mish激活函数

Mish 激活函数是光滑的非单调激活函数,其梯度变化的平滑性可以提高网络目标检测的精确率^[13],函数公式如下:

$$f(x) = x \cdot \tanh(\ln(1 + e^x)). \quad (1)$$

Mish 函数比起常用的 ReLu 函数虽然提高了目标检测精度,但是也增加了计算量,因此本文只将 Mish 函数用于网络的 Backbone 模块中.

2.4 YOLO FS 网络模型整体结构

YOLO FS 网络模型如图5所示,其中双实线表示的是跳跃连接结构. 本文将原 YOLOv5 网络模型中的 FPN 和 PAN 结构替换为 FPT,并在 FPT 结构的两端加入跳跃连接以及引入新的 Mish 激活函数. YOLO FS 网络通过输入标注好的数据集图片进行训练,检测时输入图像经过网络处理后输出3种尺度特征图像,再经过后处理得到标注好的结果图像. 实验结果显示, YOLO FS 网络模型相比于原 YOLOv5 网络模型有效地提高了目标检测精确率.

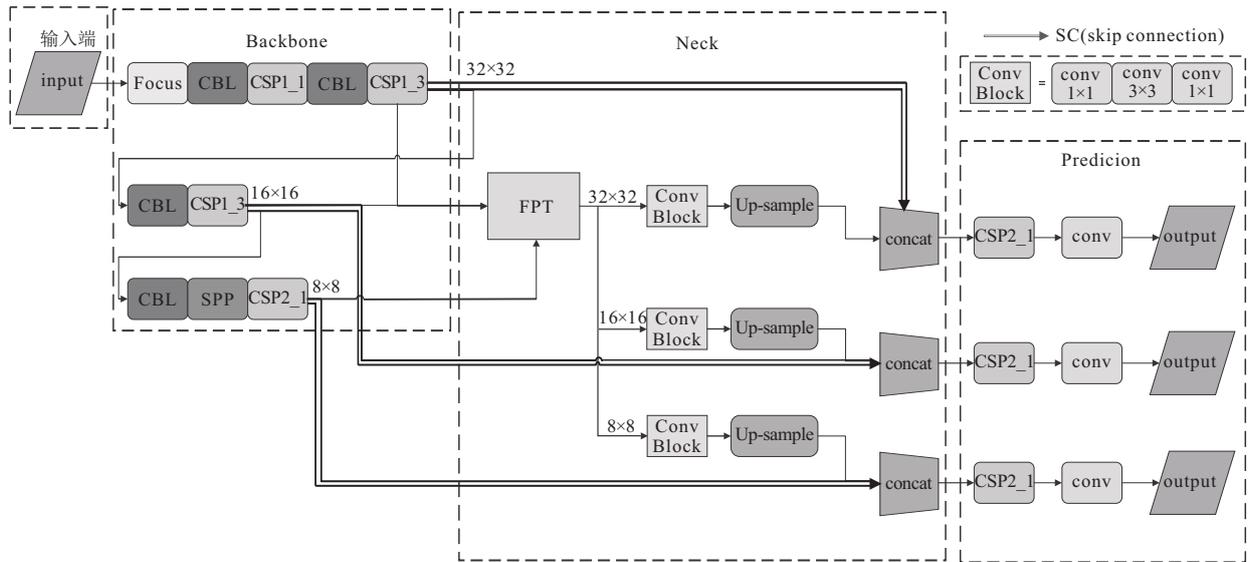


图5 YOLO FS网络模型结构

3 实验结果与分析

3.1 实验数据集

为了验证本文提出的网络模型在目标检测任务中的有效性,采用 PASCAL VOC 数据集和 MS COCO 数据集进行实验. PASCAL VOC 数据集是目标检测领域中最经典的数据集之一,包含 20 类可视目标,包含 51 700 张图像和 53 400 个注释目标,该数据集包括训练集、验证集和测试集^[14]. MS COCO 的全称是 Microsoft common objects in context,是微软团队提供的一个可以用来进行图像识别的数据集. MS COCO 数据集通过在 Flickr 上搜索 80 个对象类别和各种场景类型收集图像,其使用了亚马逊的 mechanical turk (AMT). MS COCO 数据集中的图像分为训练、验证和测试集,其中训练集包含 18 287 张图片,验证集包含 5 000 张图片,平均每张图片包含 3.5 个类别和 7.7 个实例目标,仅有不到 20% 的图片只包含一个类别,仅有 10% 的图片包含一个实例目标. MS COCO 数据集的数据量庞大,种类和实例数量较多,其中小目标数量占比较多^[15].

3.2 评价指标

为了评估本文提出的网络在解决目标检测问题的有效性,采用以下评价指标衡量模型性能:

1) PASCAL VOC 数据集的评价指标主要是平均精确率 (average precision)、平均召回率 (average recall)、F1 值 (F1-score) 和 FPS (frames per second). 评价指标计算公式如下:

$$\text{Precision} = \frac{TP}{TP + FP}, \quad (2)$$

$$\text{Recall} = \frac{TP}{TP + FN}, \quad (3)$$

$$AP = \int_0^1 \text{Precision} d \text{Recall}. \quad (4)$$

其中: Precision 代表精确率, Recall 代表召回率; TP 代表被正确分类的正样本, FP 代表被错误分类的负样本, FN 代表被错误分类的正样本.

$$F1 = \frac{2 \cdot \text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}}. \quad (5)$$

其中: Precision 代表精确率, Recall 代表召回率; F1 值

是模型精确率和召回率的一种平均值,同时兼顾模型的精确率和召回率,F1值越大,代表模型检测效果越好.网络精确率较高、召回率较低代表网络目标检测较为保守,精确率较低、召回率较高代表网络目标检测较为激进.

$$\text{FPS} = \frac{1}{t}, \quad (6)$$

其中 t 代表检测一张图片所需时间.FPS的值越大,代表网络检测时间越少,网络检测速度越快.

2) MS COCO数据集的评价指标主要是平均精确率均值、平均召回率、F1值和FPS.

3.3 实验结果与分析

YOLOv5网络模型具有YOLOv5s、YOLOv5m、YOLOv5l、YOLOv5x四种网络模型结构.4种网络模型只有网络的深度和层通道数不同,其他结构相同,故对其中体积最小的YOLOv5s网络进行对比实验.本文将YOLOFS(s版本)、YOLOv4网络和YOLOv5s网络在PASCAL VOC数据集上进行对比,将YOLOFS与YOLOv5s网络在MS COCO数据集上进行对比.实验结果如表1所示,其中*表示 AP_L 、 AR_L 的数值.

表1 各网络在PASCAL VOC数据集和MSCOCO数据集上性能的对比

数据集	网络	平均精确率/%	平均召回率/%	F1值	FPS
PASCAL VOC	YOLOv4 ^[6]	47.4	69.5	56.4	80
	YOLOv5s ^[8]	60.5	54.7	57.5	105
	YOLOv5s+Mish	63.5	53.0	57.8	105
	YOLO_FPT+Mish	63.2	57.6	60.2	53
	YOLO FS	66.6	56.3	61.1	53
MS COCO	YOLOv5s ^[8]	43.7*	69.7*	53.7	113
	YOLO FS	46.4*	71.9*	56.4	50

表1中所有对比实验模型均部署在Pytorch 1.8框架、Windows 10(64位)、NVIDIA Quadro RTX 5000(CUDA 10.2)同一软硬件环境下.表1中:YOLO_FPT表示使用FPT结构替代YOLOv5s网络中的FPN和PAN结构,YOLO FS表示使用FPT结构替代YOLOv5s网络中的FPN和PAN结构,在FPT结构的两端加入跳跃连接并引入新的Mish激活函数.对比其他方法,YOLO FS在YOLOv5网络的基础上使用FPT网络结构改进了FAN结构和PAN结构,同时结合跳跃连接模块提高了网络的整体性能.从表1的PASCAL VOC数据集实验结果可以看出:YOLOv5s在精确率上高于YOLOv4,在平均召回率上低于YOLOv4,在F1值上高于YOLOv4,说明YOLOv5s性能优于YOLOv4,但是在检测时较为保守;引入Mish激活函数之后的YOLOv5s在精确率上提升了3%,在召回率上降低了1.7%,在F1值上提升了0.3,说明Mish函数使YOLOv5s网络在检测效果上有所提升;在使用FPT结构替换FPN和PAN结构后,YOLO_FPT网络较加入Mish函数的YOLOv5s在精确率上降低了0.3%,在召回率上提升了4.6%,在F1值上提升了2.4,说明网络在检测时的性能和激进程度有所提升;在加入跳跃连接(skip connection, SC)之后,网络较替换FPT和Mish函数的YOLO_FPT网络在精确

率上提升了3.4%,在召回率上下降了1.3%,在F1值上提升了0.9;YOLO FS与YOLOv5s比较在精确率上提升了6.1%,在召回率上提升了1.6%,在F1值上提升了3.6.从表1的MS COCO数据集实验结果可以看出:YOLO FS网络相比于YOLOv5s在精确率上提升了2.7%,在召回率上提升了2.2%,在F1值上提升了2.7.对比实验结果验证了YOLO FS网络在目标检测任务中的有效性.除了客观指标以外,本文还展示了检测PASCAL VOC数据集和MS COCO数据集中的部分图像的主观效果对比.图6为YOLOv5s(最右一列)和YOLO FS网络(中间列)检测PASCAL VOC数据集以及MS COCO数据集中部分图片得到的结果,其中YOLO FS表示基于YOLOv5s框架进行改进的网络,最左一列为原图.如图6(a)~图6(d)所示,两种网络都能识别出图像中的物体,但当背景色与目标相近时,YOLOv5s网络会有误检、漏检问题,而YOLO FS网络对于此种情况具有较强的鲁棒性.从图6(e)可以看出,YOLO FS网络相较于YOLOv5s网络对于重叠图像的检测效果更好.从图6(f)可以看出,YOLO FS网络相较于YOLOv5s网络对于大量目标检测效果更好.从图6(g)可以看出,YOLO FS网络相较于YOLOv5s网络对于特征不明显图像检测效果更好.



图6 YOLO FS与YOLOv5s网络检测效果对比

4 结论

本文提出了一种改进金字塔和跳跃连接的YOLOv5目标检测网络YOLO FS,采用FPT结构替换原YOLOv5网络模型中的FPN结构和PAN结构,在FPT结构两端加入跳跃连接结构,引入Mish激活函数.其中FPT利用transformer模块的自注意力机制,并充分提取跨尺度特征信息以及同尺度相似特征信息,跳跃连接结构解决了梯度消失的问题并传递了显

著性特征,Mish函数提升了目标检测精确度.本文实验结果表明,融合以上多个结构能有效地提升网络目标检测精确率,且本文方法在PASCAL VOC数据集和MS COCO数据集的目标检测精确率优于现有YOLOv5网络.但本文提出的YOLO FS网络仍存在参数量较大以及检测速度较慢的问题,在之后的工作中会致力于降低网络的参数量,提高检测速度,以及进一步提升网络检测精度.

参考文献(References)

- [1] 李航. 基于深度学习目标检测的算法研究[D]. 北京: 中国科学院大学(中国科学院长春光学精密机械与物理研究所), 2020.
(Li H. Research on object detection algorithm based on deep learning[D]. Beijing: Institute of Physics, Chinese Academy of Sciences, 2020.)
- [2] 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, 2016: 779-788.
- [3] 安珊, 林树宽, 乔建忠, 等. 基于生成对抗网络学习被遮挡特征的目标检测方法[J]. 控制与决策, 2021, 36(5): 1199-1205.
(An S, Lin S K, Qiao J Z, et al. Object detection via learning occluded features based on generative adversarial networks[J]. Control and Decision, 2021, 36(5): 1199-1205.)
- [4] Redmon J, Farhadi A. YOLO9000: Better, faster, stronger[C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, 2017: 7263-7271.
- [5] 王殿伟, 杨旭, 韩鹏飞, 等. 复杂背景下全景视频运动小目标检测算法[J]. 控制与决策, 2021, 36(1): 249-256.
(Wang D W, Yang X, Han P F, et al. Panoramic video motion small target detection algorithm in complex background[J]. Control and Decision, 2021, 36(1): 249-256.)
- [6] Bochkovskiy A, Wang C Y, Liao H Y M. YOLOv4: Optimal speed and accuracy of object detection[J]. Computer Vision and Pattern Recognition, 2020, 17(9): 198-215.
- [7] Liu S, Qi L, Qin H F, et al. Path aggregation network for instance segmentation[C]. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, 2018: 8759-8768.
- [8] Jocher G. Yolov5[EB/OL]. (2020-08-10)[2020-08-10]. <https://github.com/ultralytics/yolo-v5>.
- [9] 邹逸群, 肖志红, 唐夏菲, 等. Anchor-free的尺度自适应行人检测算法[J]. 控制与决策, 2021, 36(2): 295-302.
(Zou Y Q, Xiao Z H, Tang X F, et al. Anchor-free scale adaptive pedestrian detection algorithm[J]. Control and Decision, 2021, 36(2): 295-302.)
- [10] Rezatofighi H, Tsoi N, Gwak J Y, et al. Generalized intersection over union: A metric and a loss for bounding box regression[C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2019: 658-666.
- [11] Lin T Y, Dollár P, Girshick R, et al. Feature pyramid networks for object detection[C]. 2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, 2017: 936-944.
- [12] Zhang D, Zhang H W, Tang J H, et al. Feature Pyramid Transformer[C]. Proceedings of the European Conf. on Computer Vision. Glasgow, 2020: 323-339.
- [13] Misra D. Mish: A self regularized non-monotonic neural activation function[C]. British Machine Vision Conference. England, 2020: 1-13.
- [14] Everingham M, van Gool L, Williams C K I, et al. The pascal visual object classes (VOC) challenge[J]. International Journal of Computer Vision, 2010, 88(2): 303-338.
- [15] Lin T Y, Maire M, Belongie S, et al. Microsoft coco: Common objects in context[C]. Proceedings of the European Conference on Computer Vision. Zurich, 2014: 740-755.
- [16] Wang X L, Girshick R, Gupta A, et al. Non-local neural networks[C]. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, 2018: 7794-7803.
- [17] He K M, Zhang X Y, Ren S Q, et al. Deep residual learning for image recognition[C]. 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, 2016: 770-778.
- [18] 王林, 李聪会. 基于多级注意力跳跃连接网络的行人属性识别[J]. 计算机工程, 2021, 47(2): 314-320.
(Wang L, Li C H. Pedestrian attribute recognition based on multi-level attention skip connection network[J]. Computer Engineering, 2021, 47(2): 314-320.)

作者简介

刘建骐(1998—), 男, 硕士生, 从事目标检测、视觉测量的研究, E-mail: 1304936474@qq.com;

闫河(1972—), 男, 教授, 博士, 从事小波分析、计算机视觉、模式识别等研究, E-mail: cqyanhe@163.com;

王潇棠(1997—), 女, 硕士生, 从事机器视觉、语义分割的研究, E-mail: 15730692969@163.com;

李梦雪(1997—), 女, 硕士生, 从事人脸识别、图像处理的研究, E-mail: limengxue@2020.cqut.edu.cn.