

基于深度学习的轻量化小目标检测算法研究

Lightweight Small Object Detection Algorithm

Based on Deep Learning

答辩人: 胡星晨导师: 范衠 教授

• 汕头大学 电子信息工程系

• 广东省数字信号与图像处理技术重点实验室

2021年5月4日

目录

- 1 研究背景与意义
- 2 本文的主要工作与创新
- 3 轻量化的多尺度特征融合网络研究
- 4 无锚框的轻量化小目标预测方法研究
- 5 综合实验
- 6 总结与展望

目标检测现状

- > 小目标检测效果差
- > 模型不够轻量化







图1目标检测在自动驾驶汽车、智能安防、机器人和无人机中的应用

目标检测现状

> 小目标检测效果差

表1 常用的目标检测算法在COCO数据集上的平均精度 (单位: %)

检测算法	骨干网络	分辨率	AP	AP ₅₀	APs	AP _M	AP _L
YOLOv2 ^[1]	DarkNet-19	288	21.6	44.0	5.0	22.4	35.5
YOLOv3 ^[2]	DarkNet-53	320	28.2	51.5	11.9	30.6	43.4
SSD ^[3]	VGG-16	300	25.1	43.1	6.6	25.9	41.4

- [1] Redmon J, Farhadi A. YOLO9000: Better, Faster, Stronger[C]. IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 2017: 6517-6525
- [2] Redmon J, Farhadi A. YOLOv3: An incremental improvement[J]. arXiv e-prints, 2018, arXiv:1804.02767
- [3] Liu W, Anguelov D, Erhan D, et al. SSD: Single shot multibox detector[C]. European Conference on Computer Vision, Amsterdam, Netherlands, 2016: 21-37

目标检测现状

▶ 模型不够轻量化

表2 常用的目标检测算法的参数量与浮点运算量

网络名称	骨干网络	输入分辨率	FLOPs	参数量
YOLOv3 ^[2]	Darknet-53	608*608	65.9B	62.5M
RetinaNet ^[4]	ResNet50-FPN	640*640	97B	34 M
Mask R-CNN ^[5]	ResNet101-FPN	800*800	149B	44.4M

[4] Lin T Y, Goyal P, Girshick R, et al. Focal loss for dense object detection[C]. IEEE International Conference on Computer Vision, Venice, Italy, 2017: 2999-3007

[5] He K, Gkioxari G, Dollar P, et al. Mask R-CNN[C]. IEEE International Conference on Computer Vision, Venice, Italy, 2017: 2961-2969

与前沿算法的对比

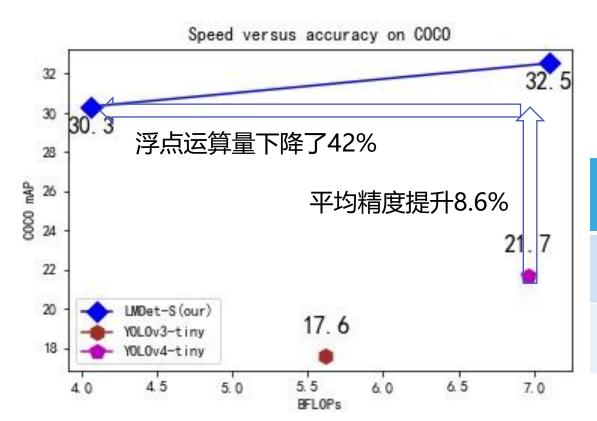


表3 在COCO数据集上本文算法与前沿轻量 化算法的性能比较

模型	mAP	APs	FLOPs	Params
YOLOv4- Tiny ^[6]	21.7	10.2	6.96B	6.06M
LMDet-S- 416	30.3	12.2	4.06B	4.0M

[6] Wang C-Y, Bochkovskiy A, Liao H-Y M. Scaled-YOLOv4: Scaling Cross Stage Partial Network[J]. arXiv e-prints, 2021, arXiv: 2011.08036v2

在移动平台的部署

表3 实机部署算法性能比较

模型	骨干网络	输入分辨率	FPS	模型尺寸
YOLOv4-Tiny ^[6]	Tiny-CSP- DarkNet53s	416*416	7.8	23.0MB
PP-YOLOv3 ^[7]	MobilenetV3	320*320	12	11MB
LMDet-S-416	EfficienNet-Lite1	416*416	13	3.9MB
LMDet-S-416	ShuffleNetV2 1x	416*416	30	2.0MB

[7] Deng K. PP-YOLO[DB/OL]. https://github.com/PaddlePaddle/PaddleDetection/blob/release /2.0-rc/configs/ppyolo/README_cn.md, 2020

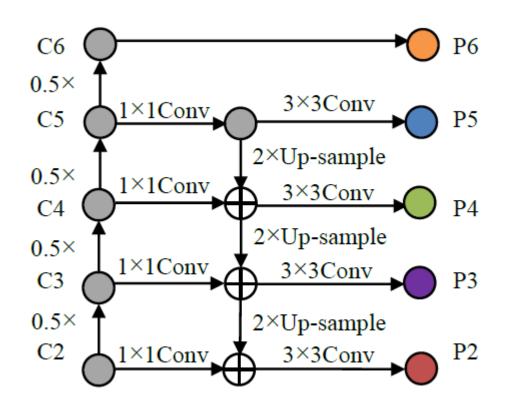
本文的主要创新

- ➤ 轻量化的多尺度特征融合网络—Lite-PAN
- ➤ 无锚框的轻量化检测头部网络—LAF-Head

3

轻量化的多尺度特 征融合网络研究

轻量化与路径增强的特征融合网络结构



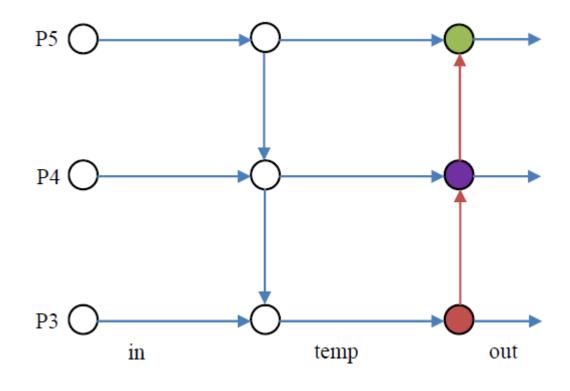
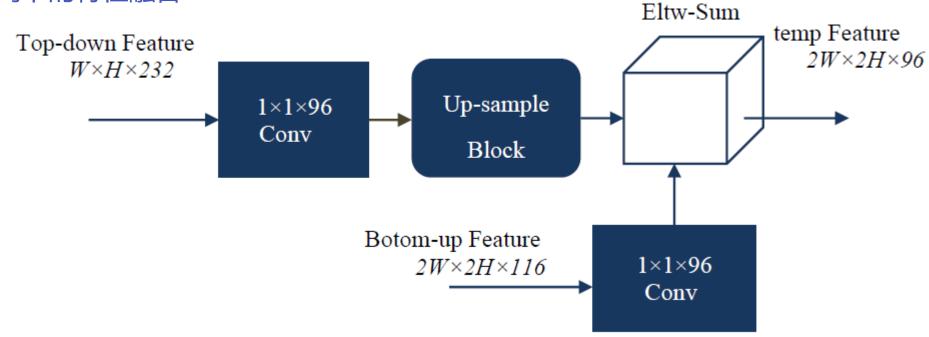


图2 FPN的网络结构

图3 Lite-PAN的网络结构

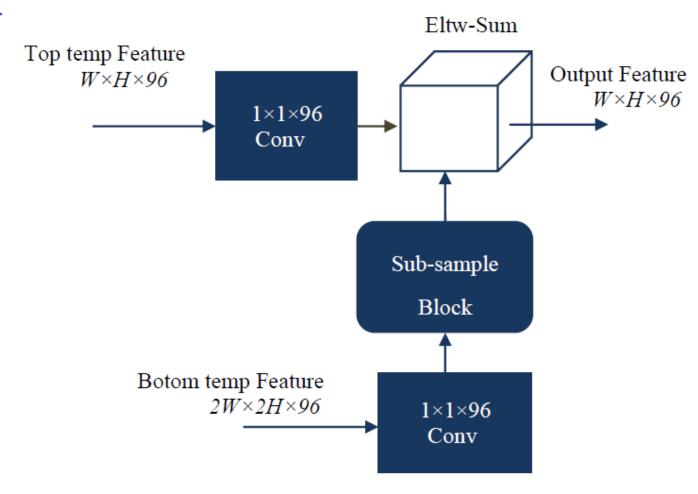
轻量化的特征融合路径

> 自顶向下的特征融合

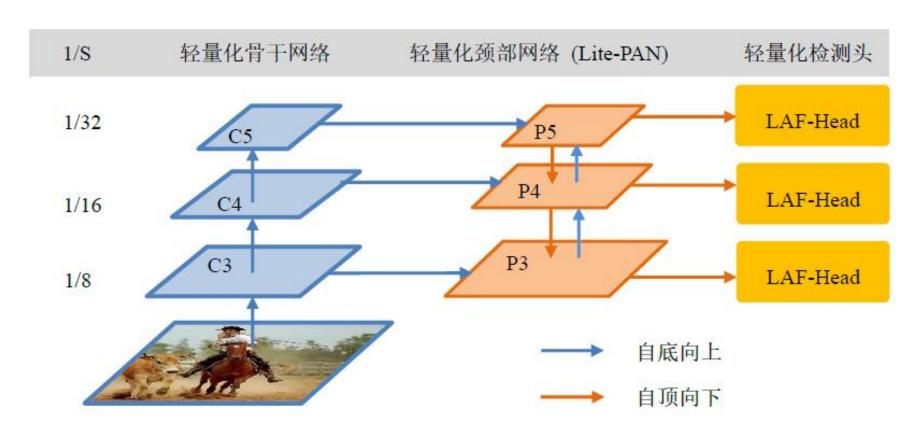


轻量化的特征融合路径

▶ 自底向上的特征融合



在本文轻量化框架中的应用



实验分析

> 消融实验

表4 有无Lite-PAN机制对算法性能的影响(在COCO数据集上的实验)

算法名称	骨干网络	分辨率	融合机制	mAP
Light-Head R-CNN ^[8]	ShuffleNetV2 1x	416×416	无	22.5%
LMDet	ShuffleNetV2 1x	416×416	Lite-PAN	23.5%

[8] Li Z, Peng C, YU G, et al. Light-head R-CNN: In defense of two-stage object detector[J]. arXiv e-prints, 2017, arXiv:1711.07264

实验分析

> 消融实验







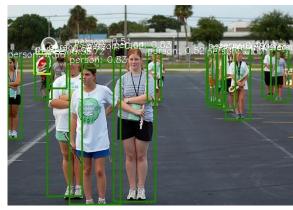






图4 有无Lite-PAN机制对算法性能的影响

方法背景

> 基于锚框的算法的缺点

鲁棒性差

泛化性能差

对超参数敏感

计算量大

> 无锚框的目标检测算法



基于中心的回归方法

> 中心点的选择

将图像划分成网格

以落在真值框内的网格的中心为正样本

对网格中心到真实框四边的距离进行回归

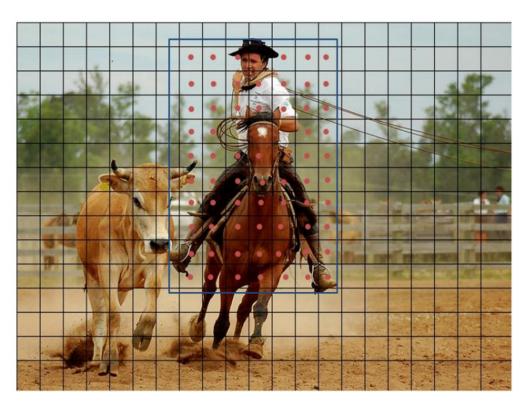
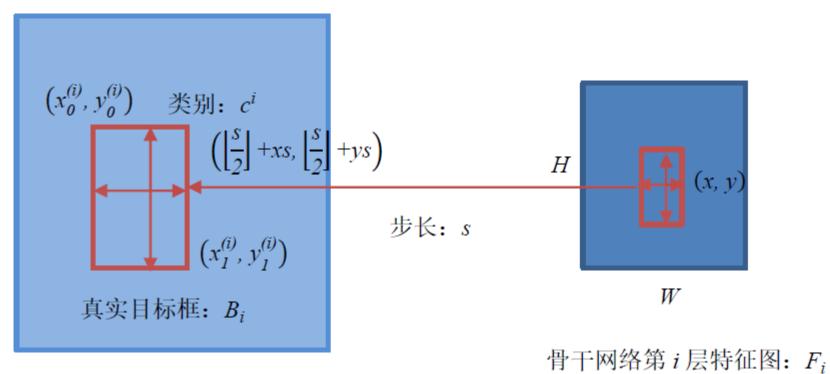


图5 中心点的选择,深蓝色边框为真实框, 红色点为落在真值框中的网格的中心点,也 即正样本点,网格为特征图映射回原图的形 象表示

基于中心的回归方法

> 从特征图到原图的映射

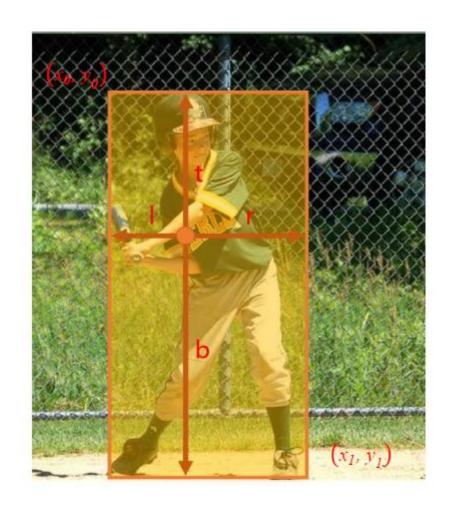


输入图像

基于中心的回归方法

▶ 回归目标

$$\begin{cases} l^* = x - x_0^{(i)} \\ t^* = y - y_0^{(i)} \\ r^* = x_1^{(i)} - x \\ b^* = y_1^{(i)} - y \end{cases}$$



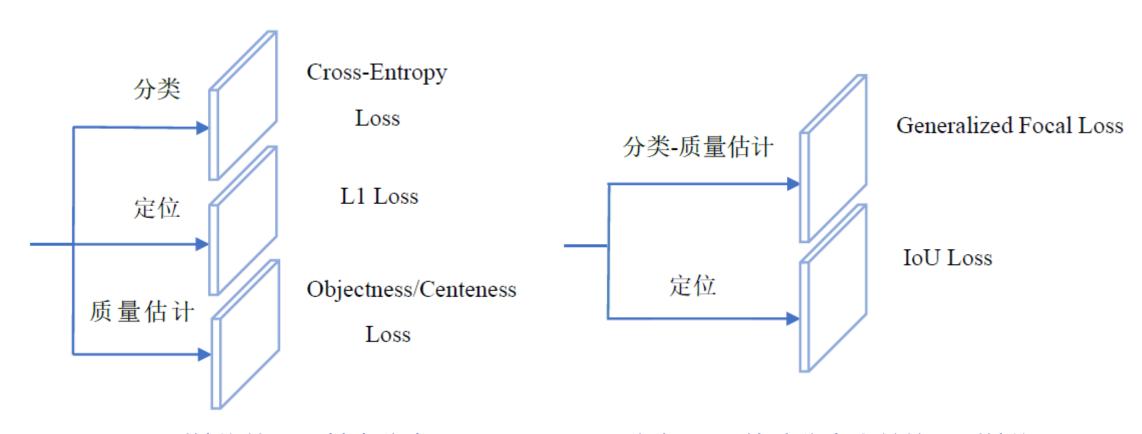
基于中心的回归方法

> 损失函数

上($\{c_{x,y}\}$, $\{t_{x,y}\}$)= $\frac{1}{N_{pos}}\sum_{x,y}L_{cls}(c_{x,y}, c_{x,y}^*)$ + $\frac{\lambda}{N_{pos}}\sum_{x,y}I_{\{c_{x,y}^*>0\}}L_{reg}(t_{x,y}, t_{x,y}^*)$

[9] Li X, Wang W, Wu L, et al. Generalized Focal Loss: Learning Qualified and Distributed Bounding Boxes for Dense Object Detection[J]. arXiv e-prints, 2020, arXiv:2006.04388v1

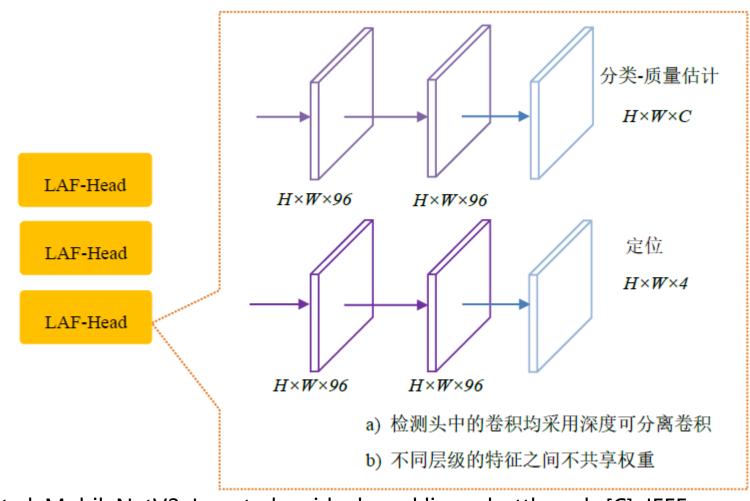
轻量化的预测分支



预测结构的三个基本分支

分类-质量估计分支合并的预测结构

轻量化的多层预测结构



[10] Sandler M, Howard A G, Zhu M, et al. MobileNetV2: Inverted residuals and linear bottlenecks[C]. IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 2018: 4510-4520

实验分析

> 与前沿算法的对比

表5 本文算法加入LAF-Head后与前沿算法的性能比较

模型	分辨率	mAP	AP ₅₀	APs	AP _M	AP _L
YOLOv4-Tiny ^[6]	416*416	21.7	42.1	10.2	26.3	30.9
LMDet-S-416	416*416	30.3	47.1	12.2	32.2	43.1

实验分析

> 与前沿算法的对比

表6 本文算法加入LAF-Head后与前沿算法的轻量化指标比较

模型	骨干网络	分辨率	mAP	FLOPs	参数量
YOLOv3-Tiny	DarkNet53	416*416	17.6%	5.6B	8.86M
YOLOv4-Tiny ^[6]	Tiny-CSP- DarkNet53s	416*416	21.7%	6.96B	6.06M
LMDet-S-416	EfficientNet-Lite1	416*416	30.3%	4.06B	4.0M
LMDet-S-512	EfficientNet-Lite1	512*512	32.5%	7.1B	4.7M

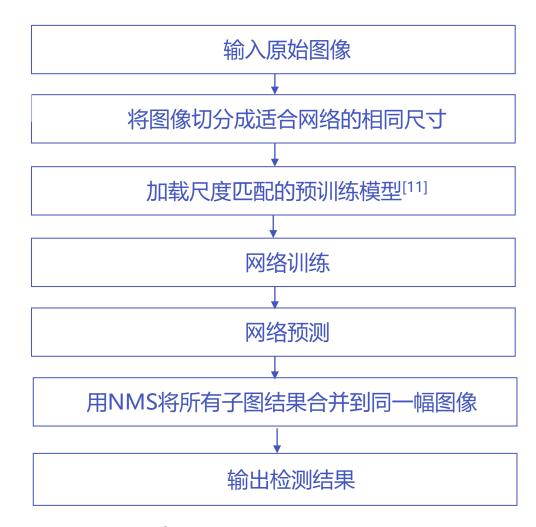
实验分析



上边是对比算法检测结果,下边是本文算法的检测结果

小目标数据集实验

> 实验流程



[11] Yu X, Gong Y, Jiang N, et al. Scale Match for Tiny Person Detection[C]. IEEE Winter Conference on Applications of Computer Vision. Snowmass Village. CO. USA. 2020. 1257-1265.

小目标数据集实验

> 实验结果

表7 LMDet-S在TinyPerson数据集上的检测精度 (单位: %)

模型	骨干网络	分辨率	AP_{50}^{small}	AP_{50}^{tiny}
LMDet-S	EfficientNet -Lite1	512*512	30.75	13.7













移动平台部署实验

> 实验流程



移动平台部署实验

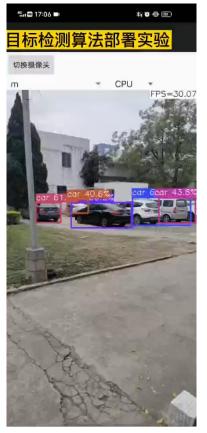
> 实验结果

表8 实机测试算法性能比较表

模型	骨干网络	输入分辨率	FPS	模型尺寸
YOLOv4-Tiny ^[6]	Tiny-CSP- DarkNet53s	416*416	7.8	23.0MB
PP-YOLOv3 ^[7]	MobilenetV3	320*320	12	11MB
LMDet-S-416	EfficienNet-Lite1	416*416	13	3.9MB
LMDet-S-416	ShuffleNetV2 1x	416*416	30	2.0MB

移动平台部署实验













YOLOv4(室外)本文算法(室内外)

YOLOv4(室内)

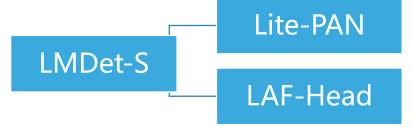
本文算法(室内)

本文算法(夜间)

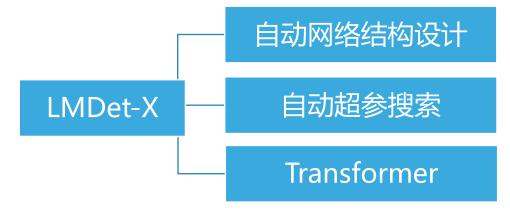
6 总结与展望

6 总结与展望

全文工作总结



未来研究展望



欢迎各位老师批评指正... THANK YOU!