

基于深度学习的无人机激光远程充电识别算法

李文峰, 杨雁南*

南京航空航天大学理学院, 江苏 南京 211100

摘要 为实现激光对飞行无人机远程充电过程中对充电目标的快速识别, 提出一种改进的 YOLOv3 算法, 采用轻量级网络模型作为特征提取网络, 实现了激光发射系统对充电无人机目标的精准快速识别。与原有的 YOLOv3 网络相比, 平均检测速度从 17 frame/s 提高到 33 frame/s, 并将网络模型权值大小从 236.0 MB 缩小到 29.7 MB, 大大减少了 YOLOv3 模型对硬件的依赖程度。研究结果表明, 改进后的算法具有较高的精确性和实时性, 为激光对无人机的实时远程充电提供了一种具有应用价值的技术手段。

关键词 机器视觉; 无线充电; YOLOv3 算法; 目标检测; 轻量级网络模型

中图分类号 TP391

文献标志码 A

DOI: 10.3788/LOP202259.1215012

Laser Remote Charging Recognition Algorithm for Unmanned Aerial Vehicle Based on Deep Learning

Li Wenfeng, Yang Yannan*

College of Science, Nanjing University of Aeronautics and Astronautics, Nanjing 211100, Jiangsu, China

Abstract In order to realize the fast recognition of charging target in the process of long-distance charging of flying unmanned aerial vehicle (UAV) by laser, an improved YOLOv3 algorithm is proposed. The lightweight network model is used as the feature extraction network to realize the accurate and fast recognition of charging UAV target by laser launching system. Compared with the original YOLOv3 network, the average detection speed is increased from 17 frame/s to 33 frame/s, and the weight of network model is reduced from 236.0 MB to 29.7 MB, which greatly reduces the dependence of YOLOv3 model on hardware. The research results show that the improved algorithm has high accuracy and real-time performance, which provides a valuable technical means for real-time remote charging of UAV by laser.

Key words machine vision; wireless charging; YOLOv3 algorithm; object detection; lightweight network model

1 引言

电动无人机具有轻便灵活、成本低、震动小、使用维护简单安全、可靠性高等优点, 在军事、民用等领域有广泛的应用。但由于该类无人机携带电池容量的限制, 工作时间短、续航能力差成为制约其发展的瓶颈问题。利用激光方向性好、能量密度

高, 以及专用激光电池允许辐照光强大、光电转换效率高的特点, 将激光束远程照射贴附在无人机上的激光电池进行实时充电, 是提高其续航时间的一种有效途径。但这种实时充电方式需要对飞行中的无人机进行精准、快速的跟踪瞄准, 保证激光束总是对准激光电池片, 才能实现高效电力补充^[1-3]。在对移动目标的充电过程中, 如何精准快速地识别

收稿日期: 2021-07-20; 修回日期: 2021-07-31; 录用日期: 2021-08-17

基金项目: 国家自然科学基金(51577091)

通信作者: *yangyn@nuaa.edu.cn

待充电目标,国内外学者进行了许多研究和探索。2010年 Becker等^[4]在充电目标上搭载全球定位系统(GPS)接收器件,采用高精度GPS定位技术获取目标位置并实时跟踪,给1 km外的机器人实施了无线充电;2019年,Zhang等^[5]采用四象限方法并结合卡尔曼滤波分析目标反射光斑的位置分布,解算出目标位置和实时偏移量实现对目标的跟踪。但这类方法需要在移动目标上安装合作器件如GPS接收器、反射棱镜等器件来间接反馈其位置,不仅增加了重量,而且增加了接受系统的复杂度,不太适用于对小型移动目标的识别。王超等^[6]采用双目相机获取移动小车的水平位置信息,通过算法得到水平偏转角来实现小车的跟踪。Bao^[7]和 Xu^[8]等通过双目相机来获取背景空间位置信息,结合卡尔曼滤波器修正,实现机器人对室内环境的识别和跟踪来达到自我定位。利用双目机器视觉的方法对目标进行识别跟踪的方法结构简单,无需在被识别目标上添加合作器件,并且成本低,适用于对小型移动目标的识别。因此,本文采用双目机器视觉来实现对无人机上光电池的识别和定位,并将整个识别装置与激光束发射系统置于二维转台上,实现激光束对无人机上光电池的实时跟踪充电。采用双目机器视觉方法实现激光束对飞行无人机的跟踪,其跟踪效果与目标检测精度和速度密切相关,目标检测算法的性能直接影响跟踪精度和速度,因此合适的目标检测算法十分重要。近年来,随着深度学习的快速发展,其已被广泛应用于目标检测^[9-10]、三维测距^[11]等方面,打破了传统计算机视觉算法的局限性,取得了良好成果。Yolo^[12]作为深度学习领域非常优秀的目标检测网络,其思想基于回归模型的one-stage检测算法,区别于two-stage的检测算法,更快的通过区域建议网络实现实时目标检测(Faster-RCNN)的算法^[13]和Yolo系列算法不需要事先进行目标候选框的预测,而是在已有的候选框上进行预测,稍微牺牲一点精度的同时大幅提升了检测速度。Yolov2^[14]和Yolov3^[15]在原有网络基础上进行了改进,性能得到很大提升。Yolov2主要借鉴了Faster-RCNN中Anchors的思想,通过K-means聚类对检测物体进行聚类分析,得到更好的候选框,并在卷积后采用归一化处理,防止梯度消失,从而加速了收敛速度。Yolov3则借鉴了ResNet的残差网络思想,在不损失梯度的情况下加深了网络深度^[16],采取了13 pixel×13 pixel、26 pixel×

26 pixel、52 pixel×52 pixel三个维度的多尺度检测,并且加入了类似特征金字塔网络(FPN)^[17]的结构,结合高层的语义信息和底层高分辨率信息,改善了单阶段多目标框检测(SSD)网络^[18]对小目标不敏感的现象。但这些网络仍然需要过多的计算能力,网络模型的复杂程度也导致了推理速度变慢,在实际的小型项目上不易部署。最新推出的Yolov3-tiny^[15]是Yolov3的简化版,它在Yolov3的基础上删减了大量特征层,只保留了13 pixel×13 pixel、26 pixel×26 pixel两个尺度,并且采用简单的套筒结构,未引入残差网络思想,虽提升了速度,但精度损失过大。在本文中,无人机飞行的速度约0.5 m/s,光电池尺寸为10 cm×10 cm,在飞行高度为3~4 m的距离上,使用双目摄像头对目标进行识别获取空间信息,并控制二维转台进行瞄准。要达到这一要求,对目标识别的误差必须控制在1 cm以内,识别速度要大于30 frame/s。在有限的硬件条件上,这要求系统在满足识别速度的同时,识别精度要达到毫米量级。由于无人机飞行空间背景环境复杂、光强影响大,直接识别目标难度较大且精度不高。为了精确识别目标位置,本文的做法是利用深度学习先框出预选框,排除大量背景区域后,在选定预选框的区域内,通过OpenCV开源库二值化处理,经逻辑判断,筛选出符合目标的像素点以达到精度要求。预选框越精确,背景排除率越高,对后续算法复杂度的要求就越低,整体速度和精度就越高。传统的Yolov3模型虽然可以满足预选框的精度要求,但是识别速度达不到要求;Yolov3-tiny模型可以加快预测速度,但精度对于光电池这类小目标效果比较差。因此,本文基于Yolov3模型提出一种既满足识别速度又满足识别精度的Yolov3-new模型。

2 Yolov3网络模型

Yolov3网络基础模型如图1所示。

darknet53为特征提取网络,负责提取图像上的信息,其中40549216个参数量为检测精度提供了保证。通过该网络提取特征后,输出13,26,52这3个维度分别负责预测大、中、小3类目标,之后进行上采样特征融合最终得到预测结果。可以看出,图像信息经特征提取网络得到3个尺度上的下采样特征图,在其最后一层,通道数达到了1024维。对于常规卷积(图2),输入 $n \times n \times h$ 的特征图经过 $m \times m$

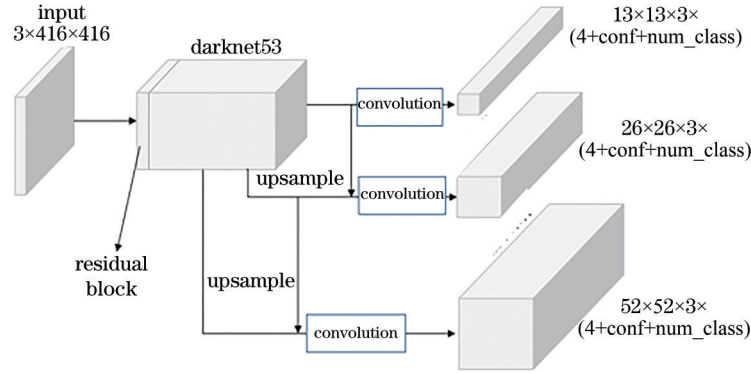


图 1 YOLOv3 网络结构图

Fig. 1 Structure diagram of YOLOv3 network

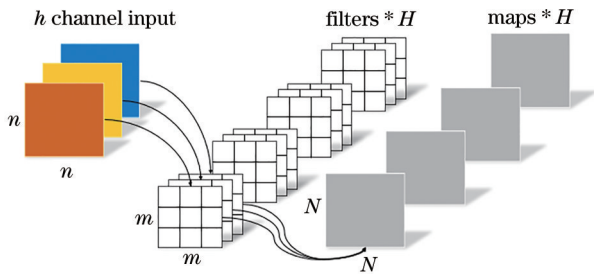


图 2 标准卷积工作方式

Fig. 2 Working mode of standard convolution

卷积核, 输出 $N \times N \times H$ 的特征图, 卷积层卷积一次的计算量由如下公式计算:

$$m \times m \times N \times N \times h \times H, \quad (1)$$

当网络的输入和输出通道为 1024, 输出大小为 13×13 时, 由式(1)计算得

$$R_{\text{result}} = 3 \times 3 \times 1024 \times 1024 \times 13 \times 13 = 1.59 \times 10^9,$$

即经过特征提取网络最后一层时, 卷积一次的计算量已达到 1.59×10^9 次, 这对硬件的计算能力要求很高, 因此并不合适部署在小型项目上。减少网络计算量是最有效的方法, YOLOv3-tiny 网络直接减少了网络层数, 虽提升了检测速度, 但对精度损失过大, 不能满足对移动目标识别误差小于 1 cm、识别速度大于 30 frame/s 的要求。因此引入新的轻量化卷积模块来代替 YOLOv3 中的卷积模块, 从而在确保识别精度的情况下提升网络速度。

3 新型网络模型 YOLOv3-new 的构建

为了同时满足对飞行无人机识别的精度和速度, 以下从两方面改进 YOLOv3 网络模型: 首先利用 ghostbottleneck 模块替换 darknet53 中的 residual block 基本卷积单元; 其次采用深度可分离卷积代替常规卷积, 从而达到网络的轻量化。

3.1 ghostmodule 模块

ghostnet^[19] 是一种新型的侧端神经网络框架, 其核心组件为 ghostbottleneck 模块。如图 3 所示, 该结构由两个 ghostmodule 模块构成: ghostmodule1 负责增加特征维度来更好地提取特征; ghostmodule2 负责减少维度。

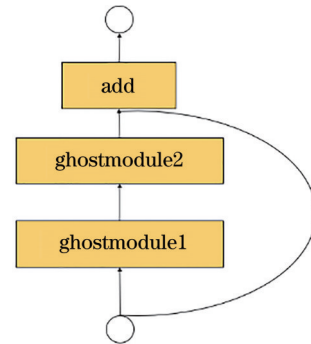


图 3 ghostbottleneck 模块结构图

Fig. 3 Structure diagram of ghostbottleneck module

ghostmodule 旨在通过“廉价”操作来获取更多的特征图。常规卷积神经网络过程中存在着大量冗余, 即相似特征图。而采用 ghostmodule 模块, 它们只需要采用“廉价”的线性变换即可得到。如图 4

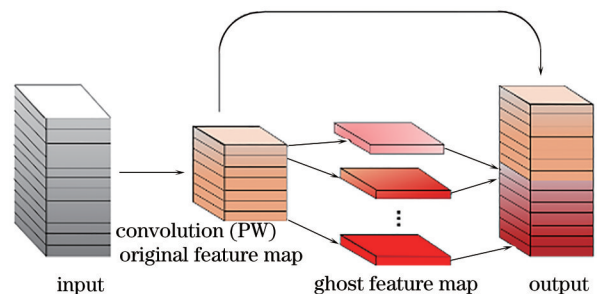


图 4 PW 卷积工作方式

Fig. 4 Working mode of PW convolution

所示,ghostmodule由卷积操作、线性变换、特征图拼接3部分组成。先对特征图进行逐点(PW)卷积操作得到固有特征图,再对固有特征图进行线性变换映射出更多的ghost特征图,最后将两部分特征图拼接起来得到最终的output。

由式(1)可知,采用PW卷积工作方式的计算量为

$$h \times 1 \times 1 \times H \times n \times n. \quad (2)$$

在得到固有特征图后,接下来通过逐通道(DW)卷积进行线性变换,如图5所示。此为分组卷积的极端形式,卷积核个数与上一层的通道数一样(通道数与卷积核数量一一对应),一个卷积核只负责一个通道,负责提取二维平面的特征信息,不考虑深度信息。由此得到ghost特征图。

此运算过程一个卷积核只包括一个维度,因此计算量为

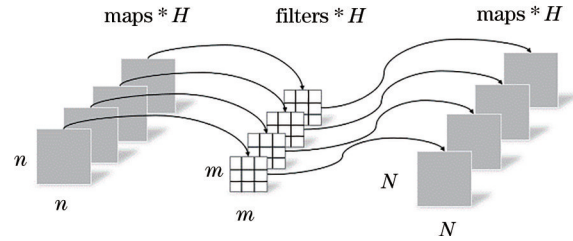


图5 DW卷积工作方式

Fig. 5 Working mode of DW convolution

$$m \times m \times H \times N \times N. \quad (3)$$

最终,将固有特征图与ghost特征图拼接起来得到output(见图4)。

这样做的优点是,先采用PW获取有价值的固有特征图时,只在深度上提取特征,参数量较少;再通过DW“线性变换”对固有特征图扩充。在相同输入的情况下,与传统卷积相比,采用ghostmodule计算压缩比为

$$\frac{(m \times m \times H \times N \times N + h \times 1 \times 1 \times H \times n \times n)}{m \times m \times h \times N \times N \times H} = \frac{n \times n}{m \times m \times N \times N} + \frac{1}{h}, \quad (4)$$

当h很大,n=N时,式(4)接近于1/m×m,采用3×3卷积核时,计算量降为原来的1/9,极大提高了计算效率。

借助于上述工作,采用ghostbottleneck来替换Yolov3中darknet53中的23个residual block基本卷积单元模块。如图6所示,residual block单元被替换成ghostbottleneck模块组成ghostnet。替换后的网络泛化能力得到提高,这是因为线性变换得到的冗余部分对网络贡献很小,不会对网络精度造成过多损失,且可有效防止过拟合现象,减少参数量,提升检测速度。

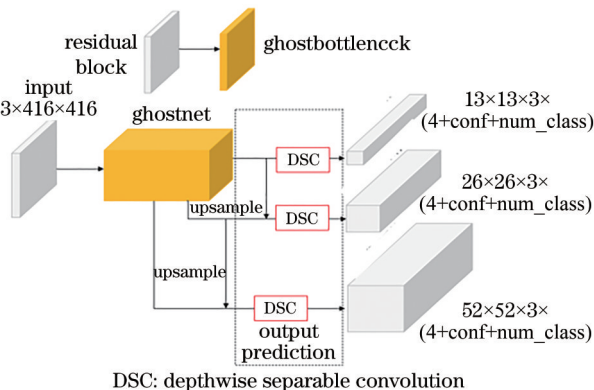


图6 Yolov3-new网络结构

Fig. 6 Structure of Yolov3-new network

3.2 深度可分离卷积(DSC)

改进特征提取网络后发现网络参数量只减少了一半左右,速度并未达到30 frame/s,无法满足识别无人机的要求。原因是图像经特征提取网络后通道数较大,不仅占用了过多内存,还影响了之后的卷积操作,由此可能影响了计算速度。为此,采用DSC代替预测输出模块的卷积操作,如图7所示。DSC也是将普通卷积拆分成PW和DW卷积,但它先通过DW获取二维平面特征,再通过PW提取深度方向特征,卷积一次计算压缩比为

$$\frac{(m \times m \times N \times N \times h + N \times N \times h \times H)}{m \times m \times h \times N \times N \times H} = \frac{1}{m \times m} + \frac{1}{H}, \quad (4)$$

相对普通卷积减少了计算量,加快了速度。

至此,完成了对Yolov3网络的改进,得到的结构如图7所示。该网络模型参数量大幅减少,改进前后网络参数量分别为61949149和7704233,在模型及网络层数相同的情况下,参数量锐减为原来的1/9。但仍可完整地保留原网络的模型结构,延续Yolov3的残差跳跃、特征融合等网络结构,使其在确保精度达到要求的情况下识别速度大于30 frame/s。

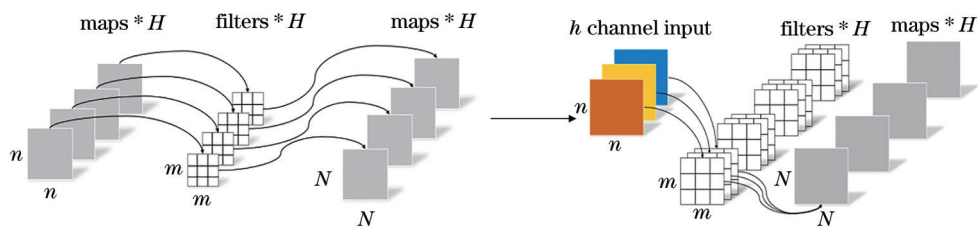


图 7 DSC 工作方式

Fig. 7 Working mode of DSC

4 分析与讨论

4.1 实验测试系统介绍

针对激光无人机远程充电的需求,将改进的 Yolov3 算法应用到无人机激光无线充电装置上,实际测试其工作性能。实验测试系统如图 8 所示。

该实验装置包括计算机、激光器、准直器、双目

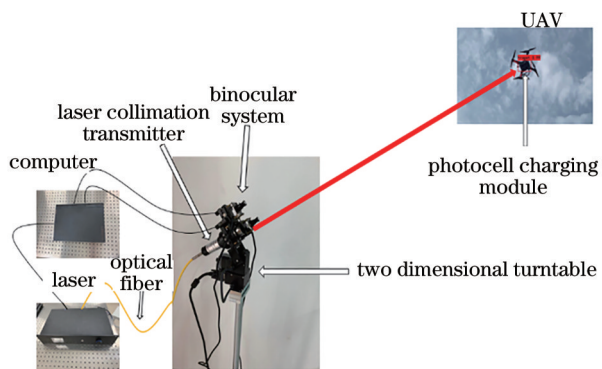


图 8 无人机充电实验测试系统图

Fig. 8 Test system diagram of UAV charging experiment

系统和待充电无人机等主要器件。整个系统的工作流程为:双目摄像头首先把采集到的图片信息输入到改进的 Yolov3-new 算法中,通过算法处理快速得到移动目标的方位信息;然后将此信息输入计算机,控制二维转台转动,让激光束对准充电目标;最后打开激光器给无人机充电。整个系统高效运行的关键在于对移动充电目标的快速高精度识别。

4.2 测试结果与分析

实验采用 voc2007^[20]数据集和本身实际项目的数据进行训练和预测,在 voc2007 数据集 1 万张图片下,利用 python 和 pytorch 框架进行训练,总体训练 500 次,前 150 次训练冻结部分网络参数加快收敛,70% 数据用来训练,30% 数据用来验证。在 NVIDIA GTX1650 上与 Yolov3、Yolov3-tiny 算法进行比较,评估指标包括检测精度(AP)、检测速度(FPS)和权重(weight)。

通过检测 voc2007 数据集中随机 1000 张图片,20 个类别的平均检测精度(mAP)值如表 1 所示。

表 1 3 种算法在 voc2007 数据集的检测性能对比

Table 1 Performance comparison of three algorithms for voc2007 data set detection

Performance	Class	Yolov3	Yolov3-tiny	Yolov3-new
AP / %	Aeroplane	82.52	80.37	77.75
	Bicycle	93.64	89.58	91.57
	Bird	87.67	74.93	83.80
	Boat	89.70	71.26	90.02
	Bottle	75.44	60.61	68.25
	Bus	93.05	88.07	96.00
	Car	87.25	84.90	88.29
	Cat	93.27	87.23	89.51
	Chair	84.63	71.48	76.48
	Cow	92.38	80.05	91.96
	Diningtable	88.56	78.01	79.53
	Dog	92.18	85.36	81.50
	Horse	94.06	86.50	93.57
	Motorbike	92.09	79.73	85.87
	Person	84.23	75.30	79.08
	Pottedplant	79.44	56.36	73.15
	Sheep	65.46	71.83	81.84
	Sofa	97.39	91.56	89.77
	Train	97.08	89.34	93.17
	Tvmonitor	86.10	80.66	91.89
mAP / %		87.81	75.16	85.15

训练 500 轮的损失值 (Loss) 如图 9 所示, 在训练 300 轮后, 损失值基本达到稳定状态, 训练完成。可以看出 YOLOv3 和 YOLOv3-new 网络训练的损失值下降得较快, YOLOv3-tiny 网络训练的损失值下降相对缓慢。主要原因是 YOLOv3-tiny 网络层数少和结构简单。

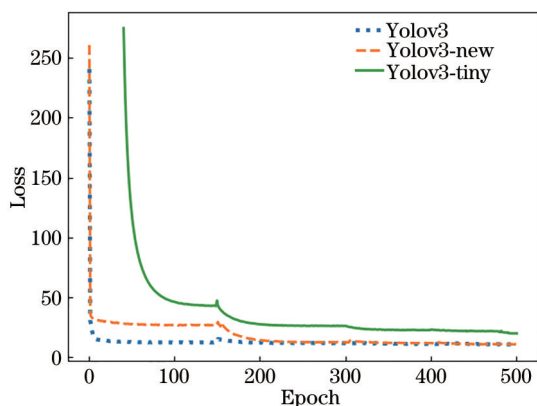


图 9 YOLOv3、YOLOv3-new、YOLOv3-tiny 训练损失值图
Fig. 9 Training loss chart of YOLOv3, YOLOv3-new, and YOLOv3-tiny

3 种算法检测性能对比如表 2 所示, 可以看出, 在训练轮数相同的情况下, 改进 YOLOv3-new 模型权值仅为 YOLOv3 的 1/8, 降至 29.7 MB, 但检测速度相对于 YOLOv3 有近 1 倍的提升, 精度和速度都满足激光为无人机实时充电的需求。

表 2 3 种算法检测性能对比
Table 2 Comparison of detection performance of three algorithms

Algorithm	Weight /MB	mAP /%	Speed / (frame·s ⁻¹)
YOLOv3	236.0	87.81	17
YOLOv3-tiny	33.8	75.16	47
YOLOv3-new	29.7	85.15	33

对于充电无人机的目标识别, 由于目标单一, 对网络性能要求不高, 但对速度要求较高。本文拍摄了 1500 张图片, 将图片进行反转、马赛克处理以增强数据集。同样进行 500 轮训练, 前 200 轮冻结部分参数, 后 300 轮正常训练, 结果如图 10 所示。由图可知, YOLOv3 和 YOLOv3-new 的损失值最终稳定在同一水平, 300 轮后损失值收敛, 训练完成。但 YOLOv3-tiny 网络对数据集不敏感, 训练 500 轮后损失值还维持在较高层面, 且预测结果差。

对飞行充电无人机测试集数据进行 AP 值测试, 结果如表 3 所示, 可以看出 YOLOv3-new 的检测

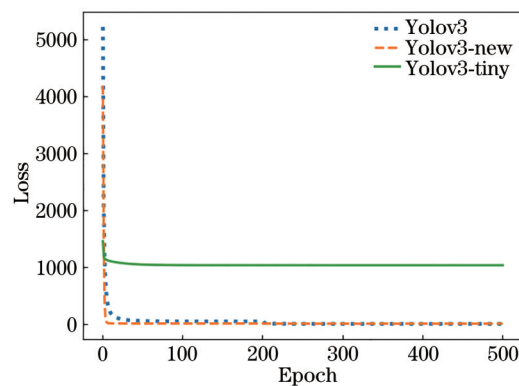


图 10 YOLOv3、YOLOv3-new、YOLOv3-tiny 无人机目标训练损失值图
Fig. 10 UAV target training loss chart of YOLOv3, YOLOv3-new, and YOLOv3-tiny

表 3 3 种算法在项目数据集整体检测性能对比
Table 3 Performance comparison of three algorithms in project data set detection

Algorithm	Weight /MB	AP /%	Speed / (frame·s ⁻¹)
YOLOv3	236.0	90.17	17
YOLOv3-new	29.7	89.33	33
YOLOv3-tiny	33.8	0	47

速度比 YOLOv3 提升近 1 倍, 权值锐减至 1/8, AP 值损失不到 1%。YOLOv3-tiny 的 AP 值为 0, 预测结果无法满足实验要求。因此, 综合考虑, YOLOv3-new 预测性能大幅提升。

将 3 种识别算法用于对无人机上光伏电池识别的实际测试效果如图 11 所示, 由图可知, YOLOv3-

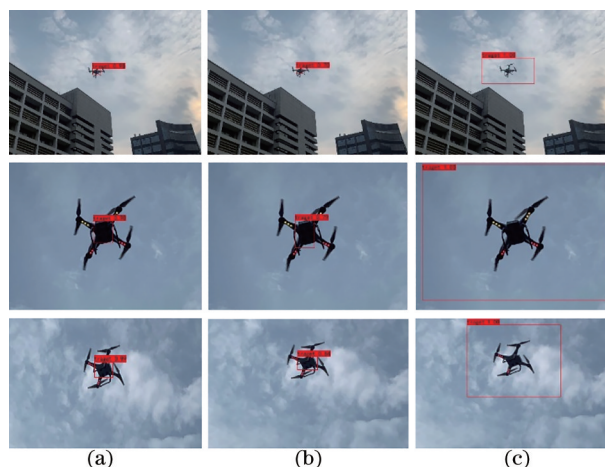


图 11 3 种算法在无人机目标数据集上的检测结果。
(a) YOLOv3; (b) YOLOv3-new; (c) YOLOv3-tiny
Fig. 11 Detection results of three algorithms on UAV target data set. (a) YOLOv3; (b) YOLOv3-new; (c) YOLOv3-tiny

new 在达到速度要求的情况下,效果与 Yolov3 差别细微,对识别结果影响很小,但比 Yolov3-tiny 精度高出很多。对比试验测试结果表明,采用 ghostmodule 和 DSC 替换网络卷积模块减少参数量的方法对该无人机充电目标的识别是可行的。

5 结 论

提出一种改进的 Yolov3-new 算法用于无人机实时远程充电识别的目标检测模型。对 Yolov3 特征提取网络 darknet53 中的 residual block 单元用 ghostmodule 轻量化网络代替,并在输出预测模块中的卷积处理上采用 DSC 进一步轻量化网络。与 Yolov3 网络相比,在检测精度几乎不变的情况下,网络参数量减少至 1/9,权值相当于 Yolov3 的 1/8,与 Yolov3-tiny 权值相当,但检测速度提升了近 1 倍。实验测试结果表明,在相同的训练轮数下,提出的新模型在目标检测速度上大大优于 Yolov3 模型,且 mAP 仅降低了 1% 左右,远优于 Yolov3-tiny 模型。该模型能在精度和速度之间达到很好的平衡,在较低硬件条件下能够满足对充电目标识别精度和速度的要求,因此该方法在激光为无人机实时远程充电技术中具有重要的应用价值。

参 考 文 献

- [1] Ouyang J, Che Y L, Xu J, et al. Throughput maximization for laser-powered UAV wireless communication systems[C]//2018 IEEE International Conference on Communications Workshops, May 20-24, 2018, Kansas City, MO, USA. New York: IEEE Press, 2018: 17896496.
- [2] Zhao M, Shi Q, Zhao M, Efficiency maximization for UAV-enabled mobile relaying systems with laser charging[J], IEEE Transactions on Wireless Communications, 2020, 19(5):3257
- [3] 何滔, 杨苏辉, 张海洋, 等. 高效激光无线能量传输及转换实验[J]. 中国激光, 2013, 40(3): 0317001.
He T, Yang S H, Zhang H Y, et al. Experiment of space laser energy transmission and conversion with high efficiency[J]. Chinese Journal of Lasers, 2013, 40(3): 0317001.
- [4] Becker D E, Chiang R, Keys C C, et al. Photovoltaic-concentrator based power beaming for space elevator application[C]. AIP Conference Proceedings, 2010, 1230(1): 271-281.
- [5] Zhang W G, Guo W, Zhang C W, et al. An improved method for spot position detection of a laser tracking and positioning system based on a four-quadrant detector[J]. Sensors, 2019, 19(21): 4722.
- [6] Wang C, Zhu L, Wang D, et al. Horizontal relative posture recognition based on binocular camera[C]//2018 37th Chinese Control Conference (CCC), July 25-27, 2018, Wuhan. New York: IEEE Press, 2018: 5566-5571.
- [7] Bao W, Zhang C W, Xiao B X, et al. Self-localization of mobile robot based on binocular camera and unscented Kalman filter[C]//2007 IEEE International Conference on Automation and Logistics, August 18-21, 2007, Jinan, China. New York: IEEE Press, 2007: 277-281.
- [8] Xu Y N, Dong Y, Li J Y, et al. Research on target tracking algorithm based on parallel binocular camera [C]//2019 IEEE 8th Joint International Information Technology and Artificial Intelligence Conference, May 24-26, 2019, Chongqing, China. New York: IEEE Press, 2019: 1483-1486.
- [9] 郭瑞鸿, 张莉, 杨莹, 等. 基于改进 SSD 的 X 光图像管制刀具检测与识别[J]. 激光与光电子学进展, 2021, 58(4): 040401.
Guo R H, Zhang L, Yang Y, et al. X-ray image controlled knife detection and recognition based on improved SSD[J]. Laser & Optoelectronics Progress, 2021, 58(4): 040401.
- [10] 苗壮, 张湧, 陈瑞敏, 等. 基于关键点的快速红外目标检测方法[J]. 光学学报, 2020, 40(23): 2312006.
Miao Z, Zhang Y, Chen R M, et al. Method for fast detection of infrared targets based on key points[J]. Acta Optica Sinica, 2020, 40(23): 2312006.
- [11] 伍俊龙, 郭正华, 陈先锋, 等. 基于深度学习的光场成像三维测量方法研究[J]. 中国激光, 2020, 47(12): 1204005.
Wu J L, Guo Z H, Chen X F, et al. Three-dimensional measurement method of light field imaging based on deep learning[J]. Chinese Journal of Lasers, 2020, 47(12): 1204005.
- [12] 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, June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE Press, 2016: 779-788.
- [13] Ren S 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.

- [14] Redmon J, Farhadi A. YOLO9000: better, faster, stronger[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition, July 21-26, 2017, Honolulu, HI, USA. New York: IEEE Press, 2017: 6517-6525.
- [15] Redmon J, Farhadi A. YOLOv3: an incremental improvement[EB/OL]. (2018-04-08) [2021-02-03]. <https://arxiv.org/abs/1804.02767>.
- [16] 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, June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE Press, 2016: 770-778.
- [17] 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, July 21-26, 2017, Honolulu, HI, USA. New York: IEEE Press, 2017: 936-944.
- [18] Liu W, Anguelov D, Erhan D, et al. SSD: single shot MultiBox detector[M]//Leibe B, Matas J, Sebe N, et al. Computer vision-ECCV 2016. Lecture notes in computer science. Cham: Springer, 2016, 9905: 21-37.
- [19] 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. New York: IEEE Press, 2020: 1577-1586.
- [20] 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.