

DOI:10.19651/j.cnki.emt.2209880

基于 YOLOv4 优化的轻量级无人机障碍物检测方法

白俊卿 张文静

(西安石油大学计算机学院 西安 710065)

摘要: 无人机平台内存小、计算资源有限。针对经典检测方法的网络结构复杂、检测速度慢等问题,提出一种基于轻量化的实时检测方法。首先,将轻量化模型 MobileNetv3 代替 CSPDarknet53 作为主干网络并融合有效通道注意力机制从而减小模型内存占用。其次,引入残差结构融合模块 RFM,增强网络的特征提取能力。为了进一步提高障碍物检测的泛化能力和算法的收敛速度,采用 Control Distance-IOU 损失函数替换 CIOU 损失函数进行网络训练。实验结果表明,在与原模型检测效果基本相同的情况下,改进后的模型内存占比减少了 80% 仅 39.5M, FPS 提高了 168% 达到 49.21 帧/s。

关键词: YOLOv4;轻量化 MobileNetv3 网络;ECA;障碍物检测;残差结构

中图分类号: TP391 **文献标识码:** A **国家标准学科分类代码:** 520.60

Lightweight UAV obstacle detection method optimized based on YOLOv4

Bai Junqing Zhang Wenjing

(School of Computer Science, Xi'an Shiyou University, Xi'an 710065, China)

Abstract: The drone platform has small memory and limited computing resources. Aiming at the problems of complex network structure and slow detection speed of classical detection methods, a real-time detection method based on lightweight is proposed. Firstly, the lightweight model MobileNetv3 replaces CSPDarknet53 as the backbone network and the effective channel attention mechanism is fused to reduce the memory occupation of the model. Secondly, the residual structure fusion module RFM is introduced to enhance the feature extraction capability of the network. To further improve the generalization ability of obstacle detection and the convergence speed of the algorithm, the Control Distance-IOU loss function is replaced by the CIOU loss function for network training. The experimental results show that memory occupation of the improved model is reduced by 80% to only 39.5M and the FPS is improved by 168% to 49.21 frames/s under the same basic detection effect as the original model.

Keywords: YOLOv4; lightweight MobileNetv3 network; ECA; obstacle detection; residual structure

0 引言

障碍物检测是无人机领域的研究热点,并且广泛应用于军用和民用领域。无人机在飞行中遇到障碍物会出现碰撞等现象,因此迫切需要通过无人机的探测来对无人机进行监控^[1]。目前无人机使用雷达、红外线、相机等传感器用于障碍物检测。相机因提供丰富视觉信息的同时还可得到低成本的高分辨率图像而被广泛应用^[2]。

传统机器视觉的方法会更多依赖于手工设计模板提取图像特征如 SIFT^[3]、SURF^[4],检测背景单一,与真实飞行环境相差较大。并且传统方法产生较多冗余信息,计算较复杂致使无法满足实时性的要求^[5]。深度学习能够利用数据集对模型进行训练,从而提升模型的鲁棒性,更易于适应

真实环境。

杨娟娟等^[6]虽然实现了对障碍物的检测,但因网络模型较大限制了检测速度。研究者提出的卷积神经网络参数量大、网络层数多例如 VGG-16^[7],虽然提高了检测精度,但是在内存受限的应用平台上进行部署比较困难。因此对模型进行轻量化和压缩操作,加快模型检测速度具有重要意义。

目前对模型进行压缩的方法主要有通道剪枝,知识蒸馏等,Chen 等^[8]在 Faster R-CNN^[9]模型框架中引入知识蒸馏,压缩模型,提高检测速度。田港^[10]采用 YOLOv4 的双尺度检测模型并对模型进行通道裁剪,通过剪枝去除对特征提取用处极小的特征层,提升检测速度。但是模型的压缩操作对于不同的数据集来说优化效果不稳定。针对模

收稿日期:2022-05-09

型的轻量化,许虞俊等^[11]使用轻量级网络 Ghost^[12],提高了参数的利用率。邵伟平等^[13]将 YOLOv3 的主干网络替换成 Mobilenet 主干网络,虽然引入轻量级网络减少了模型的参数量,但是平均检测精度下降的较为严重;

本文针对上述中提取能力特征能力有限,泛化能力较差的问题,提出了一种改进的轻量化 YOLOv4 无人机障碍物检测算法。主要改进如下:

- 1) 以 YOLOv4 模型为基础,使用改进后 MobileNetv3 替换 YOLOv4 的主干网络。并采用深度可分离卷积将原网络中 Head 模块的 3 × 3 标准卷积替换,极大的提高了检测速度。
- 2) 使用 ECA 注意力机制代替 MobileNetv3 网络中的 SE 模块,进一步减少模型的参数并保持检测性能。
- 3) 在 PANet 模块中引入融合残差模块,增强网络的特征提取能力,提高模型的检测精度。
- 4) 在损失函数方面,使用 CDIOULoss 替代原网络的

CIOULoss 来解决检测框定位不准确的问题,同时降低了计算复杂度,提升网络训练时的收敛速度。

1 本文整体网络结构

改进后的轻量化 YOLOv4 无人机障碍物检测网络(如图 1 所示)主要包含以下三部分:

第一部分是由 MobileNetv3 融合 ECA 注意力机制构成主干网络,完成对输入图片的初步特征提取。经过提取后获得 52×52 的大目标特征层、26×26 的中目标特征层、13×13 的小目标层。

第二部分是由 SPP 模块和 PANet-RFM 模块构成的颈部网络,将主干网络提取的三个特征层进行特征融合获得更具有泛化能力的特征层。

第三部分是由 YOLO-Head 构成的预测模块,将经过特征加强提取的特征层使用深度可分离卷积和 1 × 1 的卷积组合来获得最终的预测结果。

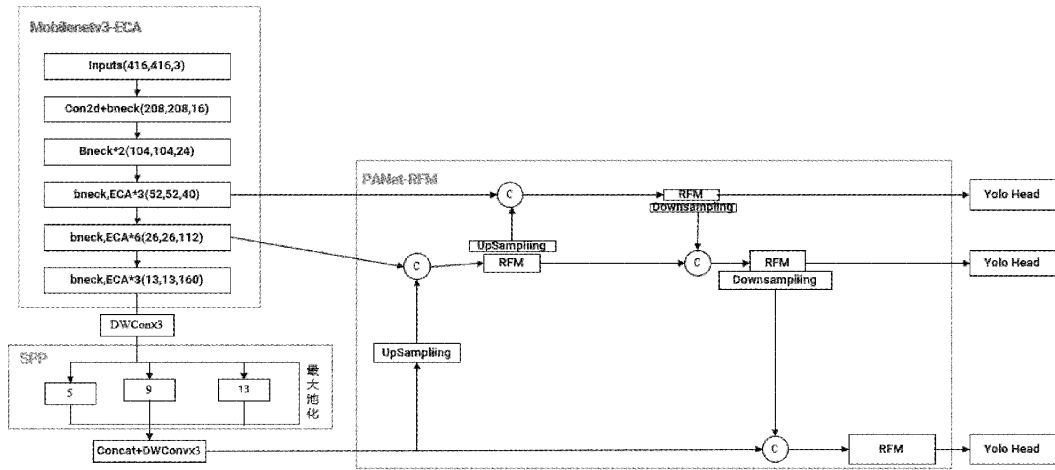


图 1 整体网络结构

1.1 MobileNetv3-ECA

YOLOv4 的主干网络为 CSPDarkNet53,目标检测分类精度较高。但是网络结构层数的加深带来了过多参数,计算量也随之加大,不利于实时性检测。针对上述问题本文引入 MobileNetv3 网络结构作为障碍物检测的主干网络,如图 2 所示。

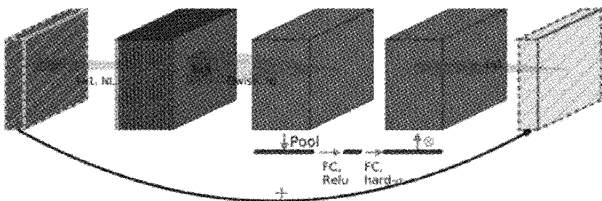


图 2 Mobilenetv3 的 bneck 核心模块

该网络中将深度可分离卷积代替标准卷积块,深度可分离卷积先分别对不同输入通道进行卷积,然后采用逐点卷积将输出的结果整合,极大地降低了模型的运算量。使

用深度可分离卷积与标准的卷积的运算量比例为:

$$\frac{D_K \cdot D_K \cdot M \cdot D_F + M \cdot N \cdot D_F \cdot D_F}{D_K \cdot D_K \cdot M \cdot N \cdot D_F \cdot D_F} = \frac{1}{N} + \frac{1}{D_K^2} \tag{1}$$

式中: D_K 为卷积核的尺寸; D_F 为输入大小; M 为输入通道数; N 为卷积核的个数。

MobileNetv3 在深度卷积之后引入 SENet^[14] 注意力机制(如图 3 所示的上半部分),获取每个输入特征通道的重要程度,根据不同的重要程度赋予通道不同的权值,从而使得卷积网络重点关注某些特征通道^[15]。

但是 MobileNetv3 中的 SE 模块使用降维致使部分特征信息丢失,降低通道的预测的准确度,同时需要捕获所有的通道之间的依赖,导致检测速度降低。因此考虑在 MobileNetv3 的 bneck 结构中引入有效注意力机制(efficient channel attention, ECA)来替换原网络中的 SE 模块,以达到通道间信息充分交流的效果。

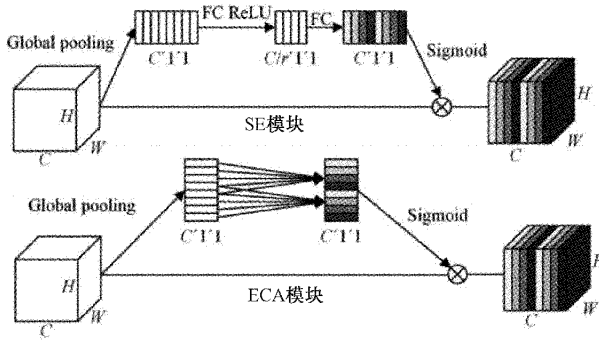


图3 SE模块与ECA模块

如图3所示的下半部分所示的ECA,输入的信息进行全局平均池化得到一维特征图,然后通过 $k \times 1$ 大小的一维卷积核得到本地的跨通道交互,其中 k 的大小是随着输入通道数的大小而改变的。根据权重占比将输入特征与通道权重结合从而获得具有通道注意力的特征。该方法中采用自适应核大小的一维卷积实现了不降维的局部跨信道交互策略,减少了特征信息的丢失,降低模型复杂度并保持了检测性能。

1.2 残差结构融合模块 RFM

YOLOv4的三个有效特征层使用了PANet结构,本文将继续沿用多层次特征融合方法,此外使用残差融合模块来替换五次的标准卷积(如图4所示)。首先根据通道维度分成两组:一组使用残差直接连接,一组使用 1×1 卷积和深度可分离卷积。接着使用channel shuffle将分组卷积之后的特征图进行重组,保证不同组之间信息可以互通,增强了特征的表达能力。最后通过卷积实现跨信道信息交互。

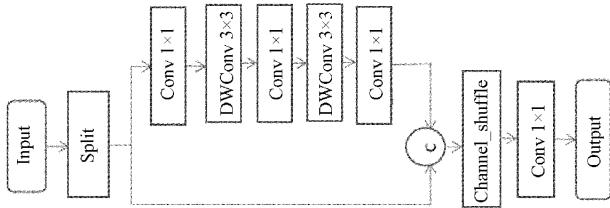


图4 RFM模块结构

1.3 Loss函数优化

在YOLOv4中使用CIoU来判断目标物体的检测框与真实框位置之间的定位置信度。CIoULoss^[26]丰富了IOULoss只考虑预测框与真实框交并比的情况。但CIoULoss计算过程中使用到反三角函数,增加了计算量导致训练时间加长并且检测精度并不高。因此本文中使用了CDIoULoss来替换CIoULoss,该方法直接检测真实框与检测框的相似性,降低模型计算量同时能够提升网络训练时的收敛速度,如图5所示。

$$diou = \frac{\|RP - GT\|_2}{4MBR'sdiagonal} = \frac{AE + BF + CG + DH}{4WY} \quad (2)$$

$$CDIOU = IOU + \lambda(1 - diou) \quad (3)$$

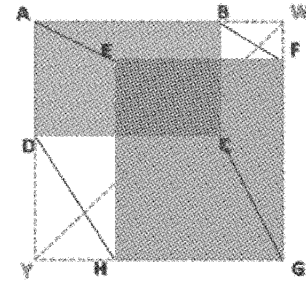


图5 CDIOU计算示意图

$$\iota_{CDIOU} = \iota_{IOU} + diou \quad (4)$$

2 模型的评判标准

本文方法的目的是在保证障碍物检测准确度的同时,减少模型内存的占比提高检测速度。因此在通过mAP(mean average precision)和FPS(frame per second)对网络性能进行测试,还应考虑到模型的内存占比。

$$P = \frac{TP}{TP + FP} \quad (5)$$

$$R = \frac{TP}{TP + FN} \quad (6)$$

$$AP = \int_0^1 P(R) dR \quad (7)$$

$$mAP = \frac{1}{n} \sum_{i=1}^n AP(i) \quad (8)$$

式(5)~(8)中 P 表示准确率, R 表示召回率, TP 指预测正确的正样本数, FP 指预测为正样本但实际结果是负样本的数量, FN 指预测为负样本但实际为正样本的数量, AP 指在不同召回率下准确率的均值。 FPS 表示每秒内可以处理的图片数,可以用来评价障碍物检测的速度。 W_i 表示该模型训练完成后的检测权重。

3 算法实验结果分析

3.1 实验环境

实验使用的CPU为AMD Ryzen 7 4800H with Radeon Graphics CPU, GPU为NVIDIA GeForce GTX1650, Windows操作系统,使用Pytorch框架,采用CUDA11.1与CUDNN并结合Python实现相关算法。

3.2 实验结果对比

考虑到无人机实际飞行的障碍物检测的实用性,需要采集现实环境中的图像作为数据集。本文障碍物数据集包括人,路标,建筑物,车辆等4个类别的13 000张图片。在训练中对采集的数据进行旋转,裁剪,添加噪声等处理,防止出现过拟合现象。输入图片的大小固定为 416×416 ,初始学习率设置为0.001,训练迭代次数设置为300,批处理的大小设置为8,动量因子为0.9。网络训练使用损失函数CDIoULoss来优化网络。网络的性能通过的平均精度、模型内存占比及FPS来衡量。障碍物检测的流程如图6所示。

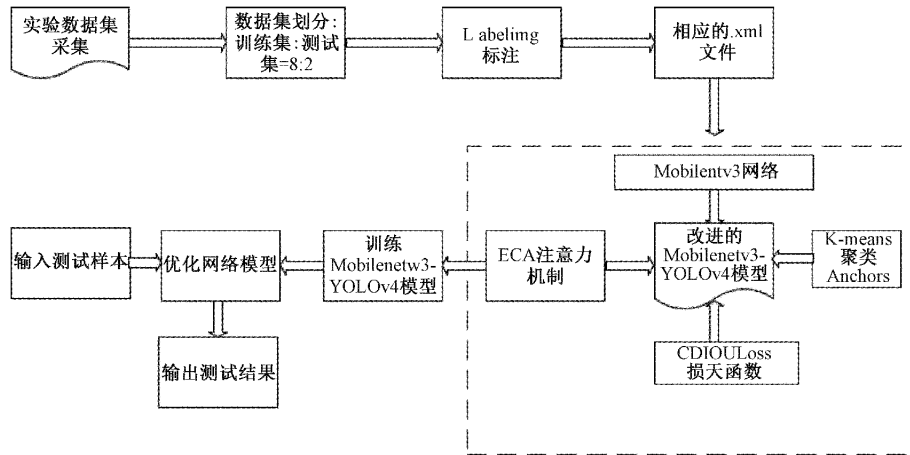


图 6 障碍物检测流程

本文模型在置信度为 0.5 的检测结果如图 7 所示,可以知道本文网络在大部分场景中检测精度较优,在不同的检测尺度检测中表现出较好的检测能力。黑色框标识数据集标注的真实框,其他颜色框标识预测框。通过比较框的位置发现使用 CDIOU 后的预测框位置较为准确。因此本文提出的模型利用网络中的有限参数可以提取到目标物体的特征信息。同时通过 CDIOULoss 能够更加准确有效的提取目标,提高预测框的精度,使网络性能得到进一步提升。

表 1 列出 YOLOv4 的消融实验结果,将主干网络替换为 MobileNetv3 后模型的内存占比降低了 188.6 MB,但是 mAP 降低了 3.5%。可以看出虽然 MobileNetv3 可以大幅度的简化网络结构,但也会伴随着 mAP 的大幅下降;将 MobileNetv3 的注意力机制替换为 ECA 后,模型的参数变化较小,mAP 提升了 1 个百分点,表明 ECA 采用的适当跨信道交互策略确实可以在降低模型复杂度的同时保持了检测性能;在网络中引入残差结构后检测精度提高了 1.31%,表明使用 RFM 后提高了特征提取能力。最终模

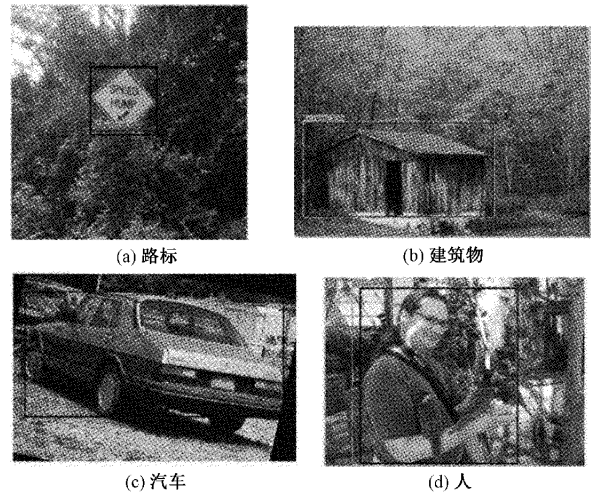


图 7 检测结果

型结构较原有 YOLOv4 在 mAP 仅降低 1%的前提下内存占比减少了 80%仅 39.50 MB,FPS 获得了 168%的提升,达到 49 帧/s。

表 1 消融实验结果

Algorithm	W_s /MB	mAP ₅₀ /%	Precision/%	FPS/(帧/s)
YOLOv4	238.35	81.79	94.15	18.23
YOLOv4+MobileNetv3	49.73	78.28	91.58	39.67
YOLOv4+MobileNetv3+ECA	43.28	79.33	89.67	40.24
YOLOv4+MobileNetv3+ECA+RFM	39.50	80.64	92.97	49.21

表 2 中对比各类别障碍物的平均检测精度,将主干网络替换为 MobileNetv3 时各类障碍物的检测精度大幅度降低,当引入注意力机制和残差融合模块各类障碍物的检测精度有一定的提升,其中 Street_sign,car, person 这三类小目标障碍物的平均检测精度提高十分明显。这表明本文提出的检测算法融合不同深度的特征以增强对浅层语义信息的表达,并利用有效通道注意力增强对重要通道的特

征信息,在某种程度上加强了对小目标的检测能力,从而提高对障碍物的检测精度。

由表 3 可知,本文设计的网络在内存模型占比、检测精度、在 GTX1650 检测速度以及预测准确度四方面均优于 SSD 的表现。相比于 Faster-RCNN 和 YOLOv4,虽然本文的网络在平均检测精度上有些许差距,但是在 GTX1650 上的检测速度是 YOLOv4 的 2.7 倍,Faster-RCNN

表2 数据集中各类障碍物的平均精度对比

网络模型	平均检测精度 mAP/%			
	路标	建筑物	汽车	人
YOLOv4	81.47	80.36	83.43	80.34
YOLOv4+MobileNetv3	78.57	78.34	80.37	77.52
YOLOv4+MobileNetv3+ECA	79.32	78.45	81.16	78.67
YOLOv4+MobileNetv3+ECA+RFM	80.21	78.46	82.54	79.21

表3 各算法对比结果

Algorithm	W_s / MB	mAP ₅₀ / %	Precision/ %	FPS/ (帧/ms)
Faster-RCNN	108.24	81.63	50.18	11.49
SSD	92.13	80.16	83.55	30.81
YOLOv4	238.35	81.79	94.15	18.23
Ours	39.50	80.64	92.97	49.21

的4.3倍,在模型内存占比上是Faster-RCNN的1/3,不足YOLOv4的1/6。根据检测精度、检测速度及模型内存占比,预测准确度四个方面综合考虑,本文提出的网络在较小的精度损失下减少网络模型参数,实现了检测精度与检测速度的平衡。

4 结 论

本文提出的轻量级目标检测算法,结合了轻量化结构MobileNetv3,融合了ECA注意力机制,同时在网络结构中引入残差结构、优化损失函数从而使得网络在减小模型内存占比和检测效果之间达到了较好的平衡。实验结果证明,改进后网络的内存模型占比不足YOLOv4的1/6,在保证精确度的情况下FPS达到49.21帧/s。同时将本文的网络与其他主干网络从内存模型占比、平均检测精度,以及检测速度进行比较实验,实验结果验证了改进算法在检测障碍物的有效性。虽然本文的改进网络的检测精度、速度在数据集上表现良好,但后续仍需继续关注在设备的实际测试部署工作。

参考文献

- [1] 邢姗姗,赵文龙.基于YOLO系列算法的复杂场景下无人机目标检测研究综述[J].计算机应用研究,2020,37(S2):28-30.
- [2] 彭继慎,孙礼鑫,王凯,等.基于模型压缩的ED-YOLO电力巡检无人机避障目标检测算法[J].仪器仪表学报,2021,42(10):161-170.
- [3] 杨岳航,陈武雄,朱明,等.基于机器视觉的无人机自主着陆技术[J].国外电子测量技术,2020,39(4):57-61.
- [4] 张绍荣,张闻宇,李云,等.基于FAST角点和FREAK描述符改进的无人机景象匹配算法[J].电子测量与仪

器学报,2020,34(4):102-110.

- [5] 解尧婷,张丕状.基于改进的YOLOv4输电线路小目标检测[J].国外电子测量技术,2021,40(2):47-51.
- [6] 杨娟娟,高晓阳,李红岭,等.基于机器视觉的无人机避障系统研究[J].中国农机化学报,2020,41(2):155-160.
- [7] 刘鸣璋,刘惠义.基于特征融合SSD的远距离车辆检测方法[J].国外电子测量技术,2020,39(2):28-32.
- [8] CHEN G, CHOI W, YU X, et al. Learning efficient object detection models with knowledge distillation[C]. Advances in Neural Information Processing Systems, 2017.
- [9] 张晓丽,董昱.面向铁路货车车号定位的Faster-RCNN卷积神经网络[J].电子测量与仪器学报,2020,34(10):65-73.
- [10] 田港.基于改进YOLOv4的无人机目标检测方法[J].空军工程大学学报,2021,22(4):9-14.
- [11] 许虞俊,李晨.基于YOLO优化的轻量级目标检测网络[J].计算机科学,48(11A):265-269.
- [12] 余洪山,郭丰,郭林峰,等.融合改进SuperPoint网络的鲁棒单目视觉惯性SLAM[J].仪器仪表学报,2021,42(1):116-126.
- [13] 邵伟平,王兴,曹昭睿,等.基于MobileNet与YOLOv3的轻量化卷积神经网络设计[J].计算机应用,2020,40(S01):8-13.
- [14] 蓝金辉,王迪,申小盼.卷积神经网络在视觉图像检测的研究进展[J].仪器仪表学报,2020(4):167-182.
- [15] 叶子勋,张红英.YOLOv4口罩检测算法的轻量化改进[J].计算机工程与应用,2021,57(17):157-168.
- [16] 韩玉洁,曹杰,刘琨,等.基于改进YOLO的无人机对地多目标检测[J].电子测量技术,2020,43(21):19-24.

作者简介

白俊卿,博士,副教授,硕士生导师,主要研究方向为计算机视觉,导航制导。

E-mail:1363684262@163.com

张文静,硕士研究生,主要研究方向为计算机视觉。

E-mail:3465752389@qq.com