

基于改进 YOLOv5s 的无人机火灾图像检测算法

苏小东¹, 胡建兴^{1,2}, 陈霖周廷¹, 高宏建¹

(1. 贵州理工学院 航空航天工程学院, 贵阳 550003; 2. 中航贵州飞机有限责任公司, 贵州 安顺 561000)

摘要: 为了解决现有火灾检测算法模型复杂, 实时性差, 难以部署在无人机平台的问题, 通过改进 YOLOv5s 算法对无人机火灾图像目标检测进行分析研究; 利用搭载高清摄像头的无人机设备获取的火灾图像、公开数据集、互联网航拍视频自主建立无人机火灾图像数据集; 采用轻量化模型 YOLOv5s 为基础模型, MobileNetV3 作为特征提取主干网络, 降低模型参数和计算量, 解决实时性差和模型部署的问题; 模型颈部引入注意力模块 CBAM, 综合了通道和空间信息, 加强网络对高层次语义信息的传递; 修改模型检测头部结构, 增强小目标检测能力; 通过消融试验对比分析各个模块对模型的影响, 与常见火灾模型进行对比分析, 分析本文算法的优劣; 算法在自建数据上的平均精度达到 76.9%, 模型大小为 6.7 M, 单帧 (640×640) 图像处理时间为 15.2 ms; 实验结果表明, 本文算法模型简单、实时性好, 为火灾检测算法部署在无人机平台奠定技术基础。

关键词: YOLOv5s; 航拍火灾图像; MobileNetV3; 小目标检测; 注意力机制

Fire Image Detection Algorithm for UAV Based on Improved YOLOv5s

SU Xiaodong¹, HU Jianxing^{1,2}, CHEN Linzhouting¹, GAO Hongjian¹

(1. Guizhou Institute of Technology School of aerospace engineering, Guiyang 550003, China;

2. China Aviation Industry Corp Guizhou Aircraft Co., Ltd., Anshun 561000, China)

Abstract: In order to solve the problems of complex model for existing fire detection algorithm, poor real-time performance, and difficult to deploy on the UAV platform, a UAV fire image target detection is analyzed and studied by improving YOLOv5s algorithm. The UAV equipment with the high-definition camera is used to obtain the fire image, public data set and Internet aerial video, and independently establish the UAV fire image data set; The lightweight model of the YOLOv5s is used as the basic model, and MobileNetV3 is used as the feature extraction backbone network to reduce the parameters and computation amount of the model, and solve the problems of poor real-time performance and model deployment; The attention module of convolution block attention module (CBAM) is introduced into the neck of the model, which integrates the channel and spatial information to strengthen the transmission of high-level semantic information; The head structure of the model is modified to enhance the ability of small target detection. Through the ablation test, the influence of each module on the model is compared and analyzed with common fire models, and the advantages and disadvantages of this algorithm are analyzed. The average accuracy of the algorithm on the self-built data reaches 76.9%, the model size is 6.7 M, and the image processing time for the single frame is 15 ms. The experimental results show that the algorithm model is simple and has good real-time performance, which lays a technical foundation for the deployment of fire detection algorithm in the UAV platform.

Keywords: YOLOv5s; aerial fire image; mobilenetv3; small target detection; attention mechanism

0 引言

火灾是世界性的自然灾害之一, 也是日常生活中主要灾害, 其发生发展不仅对自然环境有着严重影响, 还威胁着人们的生命与财产安全, 因此及时准确地发现火灾并预警具有十分重要的研究和现实意义。近年来由于自然和人为原因引发的火灾已造成极大的人员伤亡和生态破坏。2020 年四川凉山发生森林火灾, 过火面积超过 30 m², 火灾

造成 19 人死亡^[1]。在 2014 年, 拥有 300 多年历史的贵州报京侗寨失火, 烧毁房屋 100 余栋^[2]。

传统的火灾检测算法通过温度、烟雾传感器采集相关数据进行火灾预测, 但是检测准确性和实时性得不到保证。随着计算机技术和人工智能技术的不断发展, 深度学习技术已经在图像识别和检测方面展示了良好性能, 被广泛应用到火灾检测和预警。基于深度学习的目标检测算法, 如:

收稿日期: 2022-09-30; 修回日期: 2022-11-03。

基金项目: 国家自然科学基金地区基金项目(61763005); 贵州省科技计划项目(黔科合基础[2017]1069); 贵州省教育厅创新群体重大项目(黔教合 KY 字[2018]026); 贵州省普通高等学校工程研究中心(黔教合 KY 字[2018]007); 贵州省普通高等学校军民融合人才培养基地(黔科合基础[2020]011); 贵州省教育厅普通本科高校青年人才成长项目(黔教合 KY 字[2022]349); 贵州省基金基础研究计划项目(黔科合基础-ZK[2022]172)。

作者简介: 苏小东(1993-), 男, 陕西子长人, 硕士研究生, 讲师, 主要从事深度学习、目标检测方向的研究。

通讯作者: 胡建兴(1964-), 男, 山东人, 硕士, 研究员, 主要从事无人机控制方向的研究。

引用格式: 苏小东, 胡建兴, 陈霖周廷, 等. 基于改进 YOLOv5s 的无人机火灾图像检测算法[J]. 计算机测量与控制, 2023, 31(5): 41-47.

R-CNN^[3]、Fast R-CNN^[4]、Faster R-CNN^[5]、YOLO^[6-9]、SSD^[10]、RetinaNet^[11]、EfficientDet^[12]等。这些算法根据原理不同被分为：单阶段目标检测算法和双阶段目标检测算法。YOLO从2016年被提出就受到广泛关注，作为基于回归的单阶段目标检测算法，其在COCO数据集上均表现出良好性能，但对于无人机航拍图像，其在检测速度上仍然达不到实时要求，模型比较大，对设备硬件要求比较高，不能够很好的部署在无人机平台^[13-15]。

在火灾检测领域，晋耀^[16]采用Anchor-Free网络结构，在MobileNetV2网络后引入特征金字塔结构，模型AP50达到90%，单张火灾图片检测时间为40.1 ms。张竖鑫^[17]提出一种改进的Faster R-CNN检测火灾区域的目标检测框架，将主干网络替换为ResNet101，并加入注意力机制，模型mAP为0.851。任嘉锋^[18]利用FPN思想，将更低层次特征与高层次特征进行融合提高模型检测小目标能力，模型Map为0.87，相比YOLOV3原模型提高8个百分点。栗俊杰^[19]等针对无人机火灾检测问题，提出一种改进yolov2-tiny无人机火灾检测方法，并将模型部署至K210开发板，算法检出率为96.6%，检测速度为每秒14帧，为无人机火灾检测提供新思路，但其训练数据集为常规火灾数据集，飞机试验飞行高度为3米，与无人机实际工作环境有一定差异。杨文涛^[20]针对小目标检测问题，通过增加小目标检测层和特征融合的方法，提升模型对小目标检测准确度，为无人机图像目标检测提供借鉴。

针对以上问题，笔者利用自主航拍火灾视频、互联网公开无人机航拍视频和北亚利桑那大学^[21]公开FLAME航拍森林火灾图像自建无人机航拍火灾图像数据集，以YOLOv5s为基础模型，修改网络结构并引入注意力模块(CBAM, convolution block attention module)，提出一种改进YOLOv5s的无人机航拍火灾图像检测方法。并通过仿真实验验证算法可行性，为部署无人机平台奠定技术基础。

1 材料与方法

1.1 航拍火灾图像采集

数据集来源于无人机航拍火灾视频和北亚利桑那大学公开FLAME航拍森林火灾图像。无人机航拍火灾视频拍摄于贵州理工学院蔡家关校区足球场，采用大疆MATRICE 300 RTK无人机搭载高清摄像头与2020年10月20日对模拟火灾场景进行拍摄，通过调整无人机飞行高度获取不同高度的航拍火灾视频。如图1所示。

1.2 数据集构建

本文按照VOC格式自主构建数据集，由火灾视频、火灾干扰视频(类似火灾)和FLAME



图1 航拍火灾图像

数据集构成。视频利用DVDVideoSoft Free Studio软件对视频每10帧截取图像，共获得841张图像火灾图像和1161张干扰火灾图像，在FLAME中筛选1492张无人机火灾图像，自建3494张无人机火灾图像数据集。

利用标注软件labelImg对无人机火灾图像标注，文件包括图片的位置、名称、宽度、高度和维度，以及标注框对应的坐标信息。

1.3 研究方法

相比于其他目标检测算法，YOLO算法的特点是兼具精度和速度，这使得工程部署成为可能。YOLOv5是YOLO系列中比较容易部署的，其模型最小为1.9 M。本研究使用YOLOv5s为基础模型，算法结构如图3所示，其由主干网络(Backbone)、颈部(Neck)和检测层(Detect)3个部分组成。主干网络由卷积模块(Conv)、C3模块、SP-PPF模块。颈部为特征融合网络，主要包含卷积(Conv)、C3模块、上采(Upsample)、特征融合(Concat)等操作^[22]。检测层进行3个尺度的预测。

2 改进YOLOv5s算法

本研究引入CBAM注意力模块，修改主干网络和输出结构，在减少模型参数的同时进一步加强高层次特征和低层次特征融合，改进后的YOLOv5s结构如图2所示。

2.1 注意力机制CBAM

Sanghyun Woo(2018年)提出了一种轻量的注意力模块(CBAM, convolutional block attention module)，如图3所示。CBAM包含2个独立的子模块，通过通道注意力模

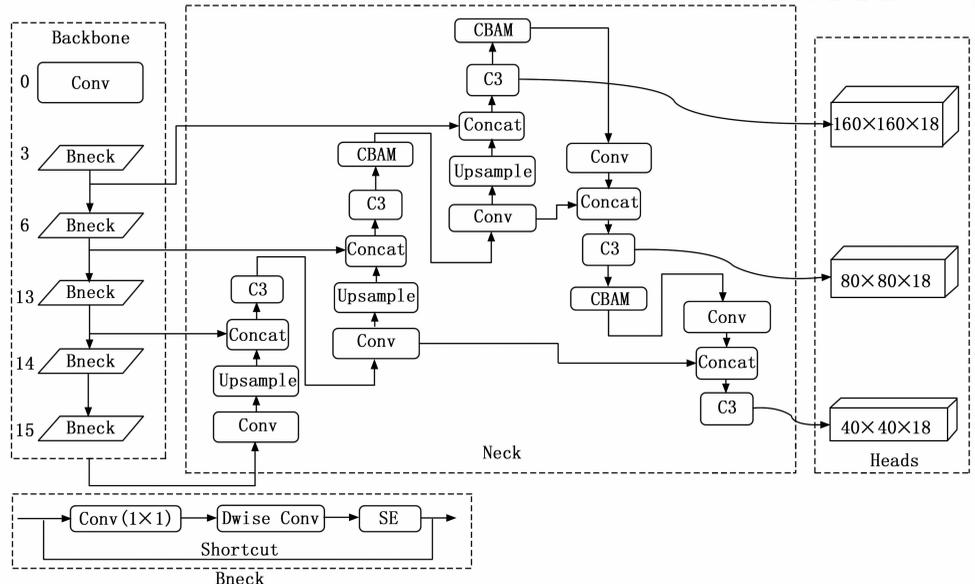


图2 改进YOLOv5s网络结构

块 (CAM, channel attention module) 和空间注意力模块 (SAM, spatial attention module) 获得通道与空间上的信息并融合^[23]。相比于 SENet, CBAM 不仅关注特征的位置信息, 还关注目标的语义特征, 分别在空间和通道两个维度。网络更加关注感兴趣的目标区域, 可以增强有效特征, 抑制无效特征或噪声, 尤其是对于小目标区域, 能获取更多的小目标的细节信息, 避免小目标特征不明显被当成背景信息^[24-25]。

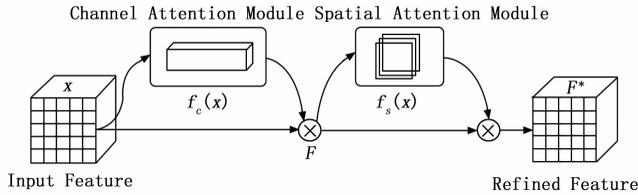


图 3 CBAM 模块

CBAM 中, 输入特征图 x , 通过通道注意力和空间注意力模块的串行计算, 得到输出特征 F^* , 即:

$$F = f_c(x) \otimes x \quad (1)$$

$$F^* = f_s(F) \otimes F \quad (2)$$

式中, $f_c(x)$, $f_s(F)$ 分别表示通道注意力和空间注意力, F 为通道注意力输出, F^* 为空间注意力输出, x 为输入特征图。

将输入特征经最大池化和平均池化后, 输入 MLP 多层感知器, 然后逐元素加和操作后经过激活函数生成通道注意力模块输出, 输出与输入特征图逐元素作乘法操作后作为空间注意力模块的输入。利用通道注意力模块输出的特征图作为空间注意力模块的输入, 通过通道注意模块的最大池化和平均池化得到两个 $H \times W \times 1$ 的特征图, 通过 Concat 方式拼接为 $H \times W \times 2$, 随后通过卷积操作 (7×7) 降维, 最后通过激活函数操作得到空间注意力模块输出, 输出与空间注意力模块输入逐元素相乘得到最终特征

$$f_c(x) = \sigma(\text{MLP}(\text{AvgPool}(x)) + \text{MLP}(\text{MaxPool}(x))) \quad (3)$$

$$f_s(x) = \sigma(\text{Conv}^{7 \times 7}(\text{Concat}(\text{AvgPool}(F); \text{MaxPool}(F)))) \quad (4)$$

式中, $\sigma(x)$ 为激活函数, MLP 为多层感知器, AvgPool 为平均池化操作, MaxPool 为最大池化操作, $\text{Conv}^{7 \times 7}$ 为卷积核大小为 7 的卷积运算, Concat 为拼接方式。

2.2 检测层改进

YOLOv5s 中有 3 个尺度的检测层, 分别为 20×20 、 40×40 、 80×80 (输入为 640×640), 可以检测大中小目标, 对图像具有很强的适应能力。本研究中自建数据集图像为航拍火灾图像, 目标多为小目标, 大尺度检测层的作用不明显, 为了在不降低检测性能的前提下尽可能减少参数, 删除大目标检测层 20×20 。在小目标检测中, 具有丰富位置信息的浅层特征比具有丰富语义信息的深层特征更重要^[26]。为了弥补高层次特征分辨率降低造成的空间信息丢失, 提高 YOLOv5s 模型对小目标的检测能力, 将第 3 层与第 27 层融合、第 6 与第 22 层的融合, 第 13 与第 18 层的

融合, 检测尺度变为 40×40 、 80×80 、 160×160 。

2.3 骨干网络修改

为了更好地将模型部署在无人机嵌入式平台, 使用 MobileNetV3^[27] 代替 YOLOv5s 骨干网络, 通过轻量型网络对模型进行优化, 减少参数数量和运算量, 提升检测实时性。

表 1 MobileNetV3-Large 网络结构

Input	Operator	exp size	out	SE	NL	S
$224^2 \times 3$	Conv2d	—	16	—	HS	2
$112^2 \times 16$	Bneck, 3×3	16	16	—	RE	1
$112^2 \times 16$	Bneck, 3×3	64	24	—	RE	2
$56^2 \times 24$	Bneck, 3×3	72	24	—	RE	1
$56^2 \times 24$	Bneck, 5×5	72	40	✓	RE	2
$28^2 \times 40$	Bneck, 5×5	120	40	✓	RE	1
$28^2 \times 40$	Bneck, 5×5	120	40	✓	RE	1
$28^2 \times 40$	Bneck, 3×3	240	80	—	HS	2
$14^2 \times 80$	Bneck, 3×3	200	80	—	HS	1
$14^2 \times 80$	Bneck, 3×3	184	80	—	HS	1
$14^2 \times 80$	Bneck, 3×3	184	80	—	HS	1
$14^2 \times 80$	Bneck, 3×3	480	112	✓	HS	1
$14^2 \times 112$	Bneck, 3×3	672	112	✓	HS	1
$14^2 \times 112$	Bneck, 5×5	672	160	✓	HS	2
$7^2 \times 160$	Bneck, 5×5	960	160	✓	HS	1
$7^2 \times 160$	Bneck, 5×5	960	160	✓	HS	1
$7^2 \times 160$	Conv2d, 1×1	—	960	—	HS	1
$7^2 \times 960$	Pool, 7×7	—	—	—	—	1
$1^2 \times 960$	Conv2d 1×1 , NBN	—	1 280	—	HS	1
$1^2 \times 1280$	Conv2d 1×1 , NBN	—	k	—	—	1

注: SE 表示该块中是否存在注意力 SE 模块。NL 表示所使用激活函数, HS 表示 h-swish, RE 表示 ReLU。NBN 表示没有批量标准化。s 表示步长。

MobileNetV3 主要针对嵌入式设备和移动端设备开发, 是一种轻量化网络结构, 它继承 V1^[28] 的深度可分离卷积, 降低网络的计算量, 引入 V2^[29] 的具有线性瓶颈的残差结构并引入轻量化 SE 注意力模块, 使网络关注更有用的信息来调整每个通道的权重, 同时该网络在结构中使用了 h-swish 激活函数, 代替 swish 函数, 减少运算量, 提高性能。其中深度可分离卷积是 MobileNet 系列的主要特点, 也是其发挥轻量级作用的主要因素, MobileNetV3 相比 MobileNetV2 可以在 COCO 数据集达到相同精度, 但速度提高了四分之一, 如图 4 所示。

与标准卷积相比, 深度可分离卷积将其拆分为逐通道卷积和点向卷积两个阶段。在逐通道卷积阶段, 输入特征图中的每个通道仅和与之对应的单通道卷积核进行卷积操作, 各输入通道之间相互独立, 取消了通道之间特征融合, 保持了卷积后输出特征图通道数量。点向卷积使用标准卷积方式, 采用 1×1 大小卷积核对所有通道的输出特征图进行整合处理, 进而改变输出特征图的通道数, 深度可分离卷积参数量

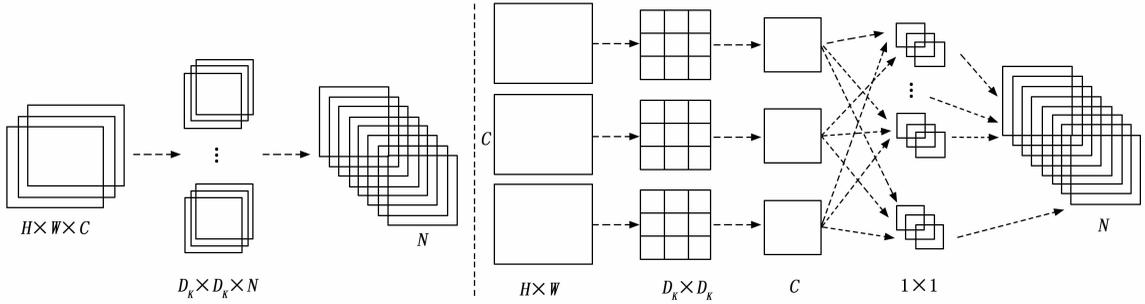


图 4 标准卷积与深度可分离卷积

与标准卷积参数数量的比值 Q 与计算量比值 R 为^[29]:

$$Q = R = \frac{1}{N} + \frac{1}{D_k^2} \quad (5)$$

式中, N 输出特征图数量, D_k 为输出特征图尺寸。

因此深度可分离卷积能降低卷积操作的参数量与计算量, 使网络更加轻量化, 提高模型检测效率。

3 试验环境与参数设置

3.1 训练环境

本实验选用开源的 PyTorch 框架, PyTorch 是一个开源的 Python 机器学习库, 该框架具有强大 GPU 加速的张量计算, 包含自动求导系统的深度神经网络, PyTorch 版本为 1.7.0, 编程语言选用 Python3.8.10, 硬件设备配置为 Interi7-9700, 内存为 8 G, GPU 为 GeForce RTX 2060 Super, 显存为 8 G, 操作环境为 Ubuntu20.04。

3.2 训练参数设置

模型训练参数设置 BitchSize 为 6, epoch 为 100。初始学习率 lr0 为 0.001, 动量因子为 0.937, 优化器选用 Adamw, 超参数 flipud=0.5, mixup=0.5, mosaic=1, copy_paste=1, 训练方式采用多尺度训练。

3.3 评价指标

目标检测的效果由预测框的分类精度和定位精度共同决定, 因此目标检测问题既是分类问题, 又是回归问题^[26]。目标检测模型的评价指标有精度 (Precision, P, %)、召回率 (Recall, R, %)、平均精度 (Average Precision, P_{AP} , %)、均值平均精度 (Mean Average Precision, mAP, %) 等。 P_{AP} 用来衡量数据集中某一类的平均分类精确率, 平均精确率均值 (Mean Average Precision, mAP), 用来衡量分类器对所有类别的分类精度, 也是目标检测算法最重要的指标。本文中只有一个类别, P_{AP} 与 mAP 相同。

为了综合评价模型, 本文使用平均精确度 P_{AP} 、模型参数量、模型大小和处理图像速度 FPS (帧/s)、浮点运算数 FLOPs 作为模型的主要评价指标。其中均值平均精度计算公式如下:

$$P = \frac{X_{TP}}{X_{TP} + X_{FP}} \times 100\% \quad (6)$$

$$R = \frac{X_{TP}}{X_{TP} + X_{FN}} \times 100\% \quad (7)$$

$$P_{AP} = \int_0^1 P(R) dR \times 100\% \quad (8)$$

式中, X_{TP} 为正样本被检测为正样本的数量; X_{FP} 为负样本被检测为正样本的数量; X_{FN} 为正样本被检测为负样本的数量。

4 试验与分析

为了验证小目标检测层、CBAM 注意力机制和轻量化骨架 MobileNet 在无人机火灾检测任务的有效性, 本文在自建数据集上进行了一系列消融试验, 以 YOLOv5s 为基线算法, mAP、P、模型大小、参数量、浮点运算次数为评价指标。

4.1 检测尺度改进试验

为了验证测试小目标检测尺度对模型的影响, 在基础模型的基础上增加一个 160×160 的小目标检测尺度, 构成 20×20 、 40×40 、 80×80 、 160×160 四个检测尺度的 YOLOv5s 模型。

表 2 中分析了增减不同检测层, 对模型性能的影响, 可以看出增加小目标检测层, 比基线模型平均精度上升 3 个百分点, 侧面说明小目标检测层可提升特征提取能力, 但网络层数增多, 参数量和模型大小分别增加 329 714 和 1.1 M; 增加小目标检测层, 并删除大目标检测层后, 平均精度提升 5 个百分点, 小目标检测能力得到明显提升, 同时参数下降 164 681 6, 模型大小减少 2.9 M, 主要得益于小目标检测层具有下采样的效果, 使网络输出特征图数量大幅减少, 参数数量下降; 删除 20×20 、 40×40 大目标检测层, 只保留 80×80 、 160×160 小目标检测层后, 模型精度比基础模型精度提升 3.6 个百分点, 参数量和模型大小分别减少 209 470 6 和 3.9 M。综合分析后, 检测头部网络选择 40×40 、 80×80 、 160×160 三个中小尺度来检测目标。

表 2 不同检测尺度性能对比

序号	检测尺度	AP/ %	P/ %	Params /M	ModelSize /M	GFL OPs
1	$20 \times 20, 40 \times 40$ 、 80×80	70.0	73.7	7012 822	14.4	15.8
2	$20 \times 20, 40 \times 40$ 、 $80 \times 80, 160 \times 160$	73	74.8	735 204 0	15.5	23.4
3	$40 \times 40, 80 \times 80$ 、 160×160	75.0	76.6	537 551 0	11.5	17.1
4	$80 \times 80, 160 \times 160$	73.6	70.5	492 762 0	10.5	15.7

注: 粗体为最优

4.2 消融试验

为进一步说明小目标检测层无人机图像检测的有效性和 CBAM 注意力机制和轻量化骨架 MobileNet 对模型性能的影响, 在自建数据集上进行消融试验。

4.2.1 CBAM 对模型的影响

由实验 1, 2; 3, 5; 4, 6; 7, 8 分别对比分析可知, 注意力机制 CBAM 不仅可加强对感兴趣区域的特征提取能力稳定提升模型平均精确度 P_{AP} , 还可节约参数和计算量。

4.2.2 改进检测层对模型的影响

由实验 1, 3; 2, 5; 4, 7; 6, 8 分别对比分析可知, 增加小目标检测层可提升模型对小目标的特征提取能力提升模型平均精确度 P_{AP} 并减少参数量。

4.2.3 MobileNetV3 对模型的影响

由实验 1, 4; 2, 6; 3, 7; 5, 8 分别对比分析可知, MobileNetV3 不仅具有强大的特征提取能力, 在不降低检测效果的前提下还因深度可分离卷积的作用可大幅度减少参数量和计算量。

由消融实验 (表 3) 可知, YOLOv5s 基础模型在自建数据集上 P_{AP} 为 70.0, 参数量为 702 232 6 个, 模型大小为 14.4 M, 计算量为 15.8 G, 将上述 3 种改进方法同时加入基线模型时, 模型 P_{AP} 提升 6.9 个百分点, 模型参数减少 404 710 0 个, 同比减少 57%, 模型大小减少 7.7 M, 同比减少 53.4%, 计算量减少 6.2 G, 同比减少 39.2%。验证了改进方案的可行性, 改进后的 YOLO5s 在不降低精度的前提下, 大幅减少模型参数和计算量, 可以有效指导无人机进行火灾检测。

表 3 自建数据集消融实验

实验序号	CBAM	检测层改进	MobileNetV3	平均精度@0.5 / %	精度 / %	参数量 / 个	模型大小 / M	FLOPs / G
1	N	N	N	70.0	73.7	701 282 2	14.4	15.8
2	Y	N	N	72.9	74.4	703 154 8	14.4	15.8
3	N	Y	N	75.0	76.6	537 551 0	11.5	17.1
4	N	N	Y	75.5	77.7	461 070 8	9.6	8.4
5	Y	Y	N	76.7	77.5	542 966 2	11.6	17.4
6	Y	N	Y	73.6	75.3	462 943 4	9.7	8.4
7	N	Y	Y	72.8	77.0	297 032 4	6.7	9.6
8	Y	Y	Y	76.9	81.3	297 522 6	6.7	9.6

注: Y 表示应用此模块, N 表示不应用此模块; 粗体为最优; M: 兆; G: 每秒浮点运算次数为 10^9 。

图 5 为训练过程中 P_{AP} 和损失的曲线图, P_{AP} 得到显著提升, 损失值进一步降低, 直观的说明改进方法对模型性能的有效提升。

图 6 为算法在测试集中火灾检测效果图, 图 (a, b, c) 为改进前的火灾检测效果, 图 (d, e, f) 为改进后的火灾检测效果。检测结果表明本文改进的 YOLOv5s 算法, 可快速准确检测出火灾位置, 降低了重复检测以及在干扰图像

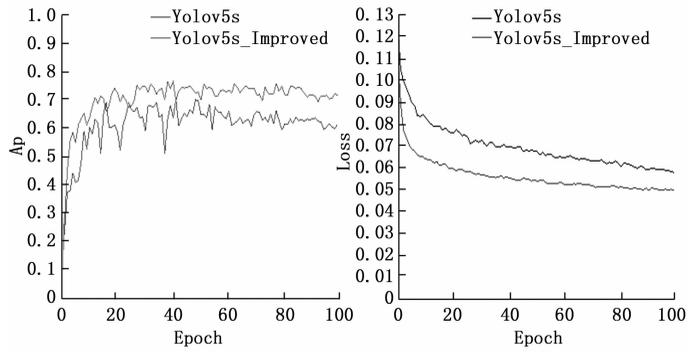


图 5 训练过程参数曲线

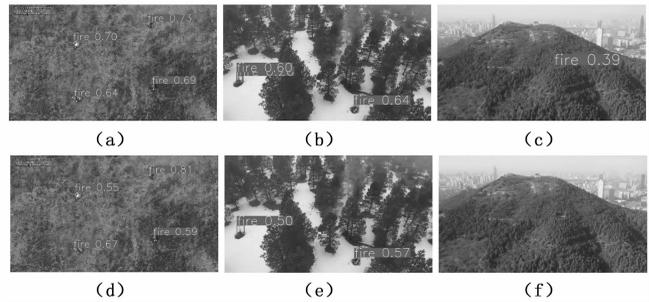


图 6 算法检测效果

中误检的情况, 提高了检测精度。

5 讨论

5.1 YOLO 系列检测算法在自建数据集上的性能对比

为证明改进后 YOLOv5s 算法对无人机火灾目标检测的有效性, 本文对 YOLO 系列算法在自建数据集上性能进行对比分析, 表 4 所示为各算法对自建数据集航拍火灾目标的 P 值、mAP 值、模型规模和浮点运算次数 FLOPs。从表 4 中可以看出, 改进 YOLOv5s 算法与其他先进算法相比取得了最优的综合性能, 检测性能与 YOLOv3 比较接近, 模型大小仅为 YOLOv3 的 5.4%, 浮点运算次数仅为 9.6 G, 计算量较小。由此可见本文提出改进 YOLOv5s 算法在处理无人机火灾图像目标检测任务时具有较大优势, 其检测精度达到 81.3%, 效果是十分可观的。

表 4 YOLO 系列检测算法在自建数据集上性能对比

方法	平均精度@0.5 / %	精度 / %	模型大小 / M	FLOPs / G
YOLOv3-tiny ^[8]	63.8	71.8	17.4	12.9
YOLOv3 ^[8]	74.5	77.1	123.5	154.9
YOLOv5l ^[30]	71.6	76.8	92.8	107.9
YOLOv5s ^[30]	70.0	73.7	14.4	15.8
本文方法	76.9	81.3	6.7	9.6

5.2 主流火灾检测算法性能对比

为证明改进后 YOLOv5s 算法相比其他主流火灾检测算法的优势, 将本文算法与当前主流火灾检测算法进行对比分析, 表 5 所示为各算法对常规火灾数据集上的 P 值、mAP 值、

表 5 主流火灾检测算法性能对比

方法	模型	主干网络	mAP@0.5(%)	P(%)	模型大小(M)	FPS(帧/s)	图像大小
文献 ^[31]	FasterRCNN	ResNet101	76.6	92.8	—	1.2	—
		ResNet101+多尺度融合	85.1	89.0	—	1.2	
		ResNet101+FPN	80.2	88.7	—	1.1	
		VGG16	61.4	70.1	—	1.9	
文献 ^[19]	YOLOv2-Tiny	Darknet19	—	94.6	—	14	210×297
文献 ^[32]	YOLOv3	Darknet53	81.6	—	—	36.2	640×640
	YOLOv5	CSPDarknet53	87.9	—	—	59.6	
	YOLOv5	CSPDarknet53+ResNet34	89.4	—	—	47	
文献 ^[33]	YOLOv3	Darknet53	—	97.0	—	26.0	416×416
文献 ^[34]	YOLOv4	CSP Darknet53	—	91.6	244	5.2	—
		CSP Darknet53+SE	—	92.5	247	5.1	
		CSP Darknet53-Tiny	—	89.9	22.6	21.5	
文献 ^[35]	SSD	VGG512	79.1	—	—	51.6	512×512
	FasterRCNN	ResNet50	81.4	—	—	11.5	
	CenterNet	ResNet50	82.5	—	31.9	42.0	
	CenterNet+ECA	ResNet50	86.3	—	31.9	43.3	
本文方法	YOLOv5s	MobileNetV3	76.9	81.3	6.7	66.7	640×640

注：表中文献 22 与本文使用航拍火灾数据集，其余均为普通火灾数据集。

模型规模和 FPS。可以看出当前主流火灾检测算法模型评价指标主要以精度为主，未涉及模型大小、参数、计算量等指标，但这些指标均影响着模型的部署难易程度和实时性，其在精度和均值平均精度都达到较高水平，但实时性比较差，尤其是采用两阶段目标检测算法（FasterRCNN），处理 1 帧火灾图像需要 1 秒左右时间，单阶段算法 YOLO、CenterNet 等最快可每秒处理 59.6 帧火灾图像。本文提出的改进 YOLOv5s 算法，在保证精度的前提下，优化网络结构，降低网络参数，模型大小为 6.7 M，每秒可处理 66.6 帧 640×640 火灾图像，容易部署，实时性比较好。

6 结束语

针对常见火灾检测算法应用场景受限，模型无法部署在无人机平台的问题，本文提出一种基于 YOLOv5s 轻量化航拍火灾检测算法。针对模型体量大，在无人机平台部署难度大的问题，以 MobileNetV3 为特征提取主干网络，减少模型参数和计算量。为了增强模型对感兴趣区域特征提取能力和抑制无效特征和背景信息，引入 CBAM 注意力机制，加强网络特征提取能力；针对无人机火灾图像主要为小目标的特点，修改网络头部检测层，提高模型小目标检测能力。通过消融实验验证各个模块对模型性能均有不同方面和不同程度的提升，改进后火灾检测平均精度达到 76.9%，模型大小仅为 6.7 M，满足部署无人机平台的要求，图像处理速度为 66.7 帧/s，即单帧图像处理时间为 15.2 ms，满足实时性要求。通过与主流目标检测算法和主流火灾检测算法对比分析可知，改进后的 YOLOv5s，不仅具有良好的特征提取能力，模型体量小，计算量小，实时性强，满足部署无人机进行火灾检测的要求，为无人机火灾检测提供技术基础，具有一定的科学和现实意义。

参考文献：

- [1] 凉山州西昌市“3·30”森林火灾事件调查结果公布 [J]. 中国消防, 2021 (1): 66-67.
- [2] 杨珂, 谭艳萍, 郭伟. 基于监测数据“云共享”的火灾预警系统设计——以贵州省木质住宅为例 [J]. 现代信息科技, 2021, 5 (12): 87-89.
- [3] KAIMING HE, GEORGIA GKIOXARI, PIOTR DOLLAR, et al. Mask R-CNN [C] //Proceedings of the IEEE International Conference on Computer Vision, 2017: 2961-2969.
- [4] ROSS GIRSHICK. Fast R-CNN [C] //Proceedings of the IEEE International Conference on Computer Vision, 2015: 1440-1448.
- [5] SHAOQING REN, KAIMING HE, ROSS GIRSHICK, et al. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks [J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2017, 39 (6): 1137-1149.
- [6] JOSEPH REDMON, SANTOSH DIVVALA, ROSS GIRSHICK, et al. You Only Look Once: Unified, Real-Time Object Detection [C] //Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016: 779-788.
- [7] JOSEPH, ALI FARHADI. YOLO9000: Better, Faster, Stronger [C] //Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017: 7263-7271.
- [8] JOSEPH REDMON, ALI FARHADI. YOLOv3: An Incremental Improvement [J]. ArXiv Preprint ArXiv: 1804.02767, 2018.
- [9] ALEXEY BOCHKOVSKIY, CHIENYAO WANG, HONGYUAN MARK LIAO. YOLOv4: Optimal Speed and Accuracy of Object Detection [J]. ArXiv Preprint ArXiv: 2004.10934, 2020.
- [10] WEI LIU, DRAGOMIR ANGUELOV, DUMITRU ERHAN, et al. SSD: Single Shot Multibox Detector [C] //European Con-

- ference on Computer Vision. Springer, Cham, 2016: 21-37.
- [11] TSUNGYI LIN, PRIYA GOYAL, ROSS GIRSHICK, et al. Focal Loss for Dense Object Detection [C] //Proceedings of the IEEE International Conference on Computer Vision, 2017: 2980-2988.
- [12] MINGXING TAN, ROUMING PANG, QUOC V. LE. EfficientDet: Scalable and Efficient Object Detection [C] //Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020: 10781-10790.
- [13] 王立舒, 秦铭霞, 雷洁雅, 等. 基于改进YOLOv4-Tiny的蓝莓成熟度识别方法 [J]. 农业工程学报, 2021, 37 (18): 170-178.
- [14] 鞠默然, 罗海波, 王仲博, 等. 改进的YOLO V3算法及其在小目标检测中的应用 [J]. 光学学报, 2019, 39 (7): 253-260.
- [15] 杨 铭, 文 斌. 一种改进的YOLOv3-Tiny目标检测算法 [J]. 成都信息工程大学学报, 2020, 35 (5): 531-536.
- [16] 晋 耀, 张 为. 采用Anchor-Free网络结构的实时火灾检测算法 [J]. 浙江大学学报 (工学版), 2020, 54 (12): 2430-2436.
- [17] 张坚鑫, 郭四稳, 张国兰, 等. 基于多尺度特征融合的火灾检测模型 [J]. 郑州大学学报 (工学版), 2021, 42 (5): 13-18.
- [18] 任嘉锋, 熊卫华, 吴之昊, 等. 基于改进YOLOv3的火灾检测与识别 [J]. 计算机系统应用, 2019, 28 (12): 171-176.
- [19] 栗俊杰, 毛鹏军, 谈文慧, 等. 基于YOLOv2-Tiny的无人机火灾检测与云台跟踪研究 [J]. 消防科学与技术, 2022, 41 (1): 108-112.
- [20] 杨文涛, 张维光. 基于改进YOLOv5m的弱小目标识别方法 [J]. 计算机测量与控制, 2022, 30 (12): 218-223, 231.
- [21] ALIREZA SHAMSOSHOARA, FATEMEH AFGHAH, A-BOLFAZL RAZI, et al. Aerial Imagery Fire Burn Detection Using Deep Learning: The FLAME Dataset [J]. Computer Networks, 2021, 193: 108001.
- [22] 赵 睿, 刘 辉, 刘沛霖, 等. 基于改进YOLOv5s的安全帽检测算法 [J/OL]. 北京航空航天大学学报: 1-16 [2022-11-01]. DOI:10.13700/j. bh. 1001-5965. 2021. 0595.
- [23] SANGHYUN WOO, JONGCHAN PARK, JOON-YOUNG LEE, et al. CBAM: Convolutional Block Attention Module [C] //Proceedings of the European Conference on Computer Vision (ECCV). 2018: 3-19.
- [24] 单鹏飞, 孙浩强, 来兴平, 等. 基于改进Faster R-CNN的综放煤矸混合放出状态识别方法 [J]. 煤炭学报, 2022, 47 (3): 1382-1394.
- [25] 姚齐水, 别帅帅, 余江鸿, 等. 一种结合改进Inception V2模块和CBAM的轴承故障诊断方法 [J]. 振动工程学报, 2022, 35 (4): 949-957.
- [26] 王建军, 魏 江, 梅少辉, 等. 面向遥感图像小目标检测的改进YOLOv3算法 [J]. 计算机工程与应用, 2021, 57 (20): 133-141.
- [27] ANDREW HOWARD, MARK SANDLER, GRACE CHU, et al. Searching for MobileNetV3 [C] //Proceedings of the IEEE/CVF International Conference on Computer Vision, 2019: 1314-1324.
- [28] ANDREW G. HOWARD, MENGLONG ZHU, BOCHEN, et al. MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications [J]. ArXiv Preprint ArXiv: 1704.04861, 2017.
- [29] MARK SANDLER, ANDREW HOWARD, MENGLONG ZHU, et al. MobileNetV2: Inverted Residuals and Linear Bottlenecks [C] //Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018: 4510-4520.
- [30] Ultralytics. YOLOv5 [EB/OL]. <https://github.com/ultralytics/YOLOv5>, 2020.06.
- [31] 张坚鑫, 郭四稳, 张国兰, 等. 基于多尺度特征融合的火灾检测模型 [J]. 郑州大学学报 (工学版), 2021, 42 (5): 13-18.
- [32] 张 全, 张 伟, 杨先凤, 等. 融合YOLOv5-ResNet级联网络的烟火检测方法 [J/OL]. 安全与环境学报: 1-10 [2022-11-01]. DOI:10.13637/j. issn. 1009-6094. 2021. 1645.
- [33] 张 为, 魏晶晶. 嵌入DenseNet结构和空洞卷积模块的改进YOLOv3火灾检测算法 [J]. 天津大学学报 (自然科学与工程技术版), 2020, 53 (9): 976-983.
- [34] 谢书翰, 张文柱, 程 鹏, 等. 嵌入通道注意力的YOLOv4火灾烟雾检测模型 [J]. 液晶与显示, 2021, 36 (10): 1445-1453.
- [35] 陆雅诺, 陈炳才. 融合注意力机制的无锚点森林火灾检测算法 [J]. 计算机与现代化, 2021 (11): 61-66.
- and optical sensor [C] //Acquisition, Tracking, and Pointing V. SPIE, 1991, 1482: 234-245.
- [15] 姚 郁, 郑天宇, 贺风华, 等. 飞行器末制导中的几个热点问题与挑战 [J]. 航空学报, 2015, 36 (8): 2696-2716.
- [16] 李 伟, 何鹏举, 高社生. 多传感器加权信息融合算法研究 [J]. 西北工业大学学报, 2010, 28 (5): 674-678.
- [17] 高 颖, 韩宏帅, 武梦洁, 等. 机动目标的IMM扩展卡尔曼滤波时间配准算法 [J]. 西北工业大学学报, 2016, 34 (4): 621-626.
- [18] 王树亮, 毕大平, 阮怀林, 等. 基于视觉注意机制的认知雷达数据关联算法 [J]. 航空学报, 2018, 39 (6): 152-160.

(上接第40页)

- [11] 丁光涛. 四元数的共轭数和乘法运算的有条件交换性 [J]. 安徽师范大学学报 (自然科学版), 2014, 37 (1): 1-5.
- [12] TAN Q, WANG P, HU J, et al. The application of multi-sensor target tracking and fusion technology to the comprehensive early warning information extraction of landslide multi-point monitoring data [J]. Measurement, 2020, 166: 108044.
- [13] 姚 郁, 郑天宇, 贺风华, 等. 飞行器末制导中的几个热点问题与挑战 [J]. 航空学报, 2015, 36 (8): 2696-2716.
- [14] BLAIR W D, RICE T R, ALOUANI A T, et al. Asynchronous data fusion for target tracking with a multitasking radar