

引用格式: 张旭彤, 胡鹏, 赵鑫, 等. 火灾预警中基于YOLO V5的火源智能检测定位方法[J]. 微电子学与计算机, 2023, 40(3): 67-74. [ZHANG X T, HU P, ZHAO X, et al. Intelligent detection and location method of fire source based on YOLO V5 in fire warning[J]. Microelectronics & Computer, 2023, 40(3): 67-74.] DOI: 10.19304/J.ISSN1000-7180.2022.0340

火灾预警中基于YOLO V5的火源智能检测定位方法

张旭彤, 胡鹏, 赵鑫, 丁云霞
(安徽理工大学人工智能学院, 安徽淮南232000)

摘要: 如何高效地检测出火灾初期的火源并对其进行准确定位,是有效遏制火情恶化和及时制定消防计划的重要前提。目前火源检测定位所面临的主要问题为火源检测与定位双任务相互分离,这严重制约了火灾预警的实时性。为了克服上述问题,提出以YOLO V5作为火源检测基础模型,同时利用CIOU(Complete intersection over union)损失函数对anchor(anchor-boxes)与GT(Ground Truth)进行精准框定以进一步提高模型标注精度,并将Leaky RELU激活函数替换为正则化和激活函数相结合的GELU(Gaussian Error Linear Unit)。另外,在准确检测出火源的同时,采用平行双目定位算法对火源进行空间定位,以实现火源检测与定位的智能一体化。实验结果表明,所提方法的火源检测mAP值比原始算法提高了9.8%,在保证检测火源精确性的同时能准确定位火源位置。

关键词: 目标检测; YOLO V5; 双目定位; 火源检测; 智慧消防

中图分类号: TP391.4

文献标识码: A

文章编号: 1000-7180(2023)03-0067-08

Intelligent detection and location method of fire source based on YOLO V5 in fire warning

ZHANG Xutong, HU Peng, ZHAO Xin, DING Yunxia

(College of artificial intelligence, Anhui University of Science and Technology, Huainan 232000, China)

Abstract: How to detect the fire source efficiently and accurately locate it is an important prerequisite for effectively controlling the deterioration of fire situation and making fire control plan in time. At present, the main problem faced by fire source detection and location is that the dual tasks of fire source detection and location are separated from each other, which seriously restricts the real-time performance of fire early warning. To overcome the above problems, YOLO V5 is proposed as the basic model of fire source detection, and CIOU (Complete intersection over union) loss function is used to accurately frame anchor (anchor-boxes) and GT (Ground Truth) to further improve the annotation accuracy of the model. The leaky RELU activation function is replaced by GELU (Gaussian Error Linear Unit), which combines regularization and activation function. In addition, while accurately detecting the fire source, the parallel binocular location algorithm is used to locate the fire source in space, to realize the intelligent integration of fire source detection and location. The experimental results show that the fire source detection map value of the proposed method is 9.8% higher than the original algorithm, which can accurately locate the fire source while ensuring the accuracy of fire source detection.

Key words: object detection; YOLO V5; Binocular positioning; Fire source detection; intelligent fire protection

收稿日期: 2022-06-01; 修回日期: 2022-07-24

基金项目: 安徽高校自然科学研究重点项目(KJ2020A0289); 安徽理工大学青年教师科学研究基金重点项目(QNZD2021-02); 安徽省大学生创新训练项目(S202210361271); 安徽理工大学科研基金(13210679); 淮南市科技规划项目(2021005)

1 引言

随着经济的飞速发展,城市规模不断扩大、建筑复杂性持续提高,使得火灾防控的难度和挑战性急剧攀升^[1]. 大多数火灾的产生都表现出鲜明的过程性,从产生火苗到出现火源直至恶化成火灾,一般需要一定时间的无干扰发展^[2]. 其中火源的出现既不像火苗那样难以发现,又不同于火灾难以控制. 大量的事实表明,如果能准确检测出火源并确定其位置,可有效遏制火情恶化,从而将火灾扑灭于产生之初. 因此,准确发现火源并对其进行定位对火灾的遏制起着决定性作用.

众所周知,火源检测在本质上属于目标检测的一种. 随着第四次工业革命智能化时代的发展,目标检测技术有了长足进展. 大量特征提取算法的提出开创了目标检测新技术,为目标检测技术奠定了扎实的理论基础与实践可能. 1999 年 David G.Lowe 等人发表了 SIFT 算法^[3],提出利用高斯函数在不同尺度空间查找关键点,通过尺度不变特征来匹配图像映射从而检测物体; Navneet Dalal 等人则在 2005 年提出了 HOG 算法^[4],其基本思想是利用梯度将图像局部外观和形状特征化,并用滑窗产生大小相同细胞单元,最后使用特征进行堆叠产生特征向量; 随后,2008 年 Felzenszwalb 等通过对 HOG 算法进行改进,提出了 DPM 算法^[5],并用其对特征图融合进行传统分类. 然而,上述算法普遍在查找关键点、生成描述子过程中计算复杂度较高,其低下的运算效率难以满足实时性的要求^[6]. 另外,上述算法所依赖的特征都来自人工设计,这类特征过度依赖先验知识且泛化能力十分有限^[7],这严重制约了检测算法在实际应用中的通用性.

近年来,随着深度学习的崛起,以数据驱动的自学习特征方法正在逐步取代人工特征. 大量基于自学习特征进行目标检测的方法已取得了众多引人注目的成果^[8-12]. 2015 年的 Faster-RCNN^[8-10] 在目标检测领域影响深远,该算法通过添加 RPN 网络,基于 anchor 机制来生成候选框,最终将特征提取、候选框选取、边框回归和分类整合到一个网络中,有效提升了检测精度和检测效率. 但从粗到细的检测使计算出现冗余,多次重复相同的特征提取导致实时检测效果差,而火灾的防控对实时性要求较高,延误警报往往会造成严重损失. 为了提高检测算法的实时性,端到端的一阶段检测方法^[11-12] 被相继提出. 其中 SSD^[13-14] 是这类算法典型代表,一阶段的 SSD 有效解决了因

检测速度带来的诟病,但随着卷积深度的不断加深,特征图中小目标特征极易丢失,致使 SSD 对小目标检测不够理想,这使得其对火灾早期小火源的检测容易出现漏检. 此外,火灾防控一般需借助于嵌入式设备进行部署,而上述 SSD、Faster-RCNN 由于网络结构拥有大量的参数,导致在算力有限的嵌入式部署时会带来困难^[15-17]. 上述弊端随着 YOLO V5^[18] 的出现有了明显改观,2020 年 Glenn Jocher 等人在 YOLO 前几代版本^[11-22] 的基础上发布一种新型轻量级检测模型 YOLO V5. 该算法利用残差网络融合高、低特征层之间的特征信息,有效解决了因卷积深度过深所造成的小目标特征丢失问题,从而显著提升了小目标的检测精度^[23]. 另外,在火源检测的基础上实现其准确定位是提高火灾防控成功率的决定性因素. 然而,双目视觉定位对 anchor 的精准度要求极高,原始 YOLO V5 的 GIOU 损失函数并不能精确的重合 anchor 与 GT. 因此,如何提高 anchor 和 GT 的重合度是提高检测精度的关键所在.

综上所述,本文针对火灾形成早期小火源难以及时发现并准确定位的问题,提出了一种基于 YOLO V5 和双目视觉相结合的火源智能检测定位一体化方法. 为了更精准的检测与定位到火源,通过改进新型轻量级模型 YOLO V5,将网络损失函数更换为 CIOU,对模型获得高重合度 anchor 与 GT 进行计算,同时特征提取网络更换为 GELU 激活函数,使得模型更精准的检测火源,利用 YOLO V5 实时性快的优点,实现了对火源的快速检测、精准定位.

2 火源检测网络结构

2.1 YOLO V5 目标检测算法

YOLO V5 作为轻量级新型网络,拥有实时性检测、轻量化部署等优点. 该网络结构主要由输入端、Backbone 网络、Neck 网络和 Prediction 四部分组成,如图 1 所示. 输入端输入 640*640 大小图像. Backbone 网络分为 Focus、CSP^[24] 结构. 利用 Focus 结构对图像进行切片操作,将 $n*n*n$ 的图像经卷积切片变为 $0.5n*0.5n*12$ 的特征图,通过 32 通道数 Conv 卷积层,输出 $0.5n*0.5n*32$ 大小特征映射. CBL 由 Conv 卷积、BN 标准化、Leaky RELU 激活函数构成,CSP 通过 CBL 输出处理对特征进行残差处理,避免因卷积层加深造成梯度消失,有效解决 SSD 检测算法对小目标的检测效果的弊端. 残差处理使得深度特征值与浅层特征值进行 Concat 特征拼接操作,最后利用融合特征进行 BN 标准化和 Leaky RELU 激活函数,通过

CBL 处理输出特征向量. Neck 网络使用 FPN 结构自顶向下 (top-down) 将深层语义特征向下传递, 增强语义信息, PAN 结构自底向上 (bottle-up) 对 FPN 进行补充, 将低层定位特征向上传递, FPN+PAN 结构对

深层和浅层网络特征进行融合, 提高目标提取的特征质量. Prediction 采用 GIOU 损失函数做 Bounding box 损失函数, 通过 anchor 和 GT 的重合与非重合面积来计算两框的相对位置, 从而对网络进行调整.

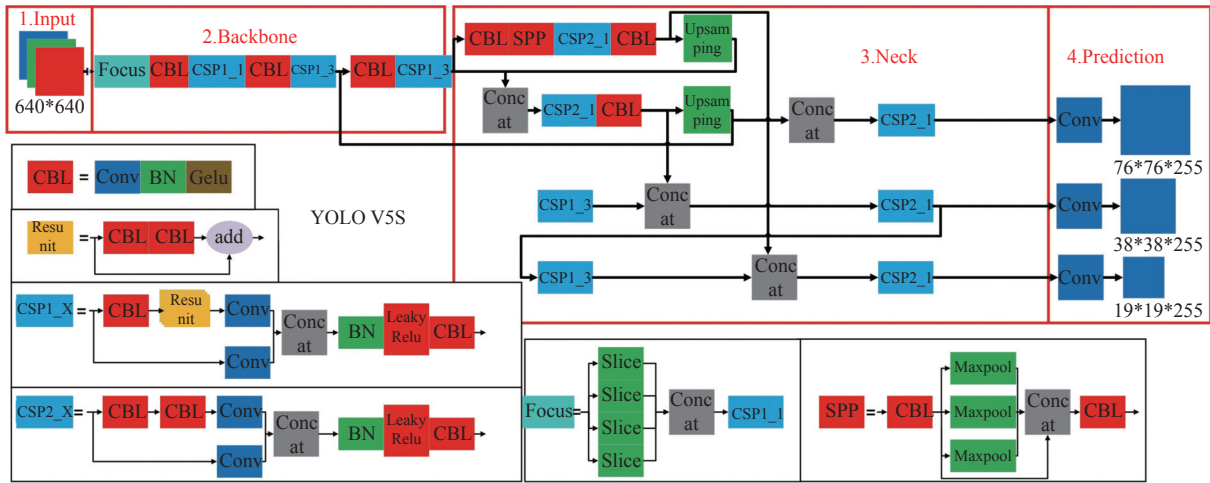


图 1 YOLO V5 网络结构

Fig. 1 YOLO V5 network structure

2.2 算法改进

YOLO V5 模型中, 输出端采用 G_{IOU} 损失函数作为边框预测计算损失方法, 对预测边框和实际标注边框进行比对并计算损失值. 该方法有效解决了 I_{OU} 损失函数 anchor 与 G_T 不相交导致损失函数不可导的问题, 公式如下:

$$I_{OU} = \frac{|A \cap B|}{|A \cup B|} \quad (1)$$

$$G_{IOU} = 1 - I_{OU} + \frac{|C - (A \cup B)|}{|C|} \quad (2)$$

式中, A, B 分别为 anchor 和 G_T 的大小, C 为 A, B 框所包起来的最小面积. 根据上式不难发现当 anchor 在 G_T 内时, G_{IOU} 便无法得到 G_T 与 anchor 的相对位置关系, 即退化为 I_{OU} 损失函数, 如图 2 所示, 蓝色框表示为 anchor, 绿色框表示为 G_T .



图 2 G_{IOU} 损失函数 anchor 在 G_T 内部图像

Fig. 2 The image of the anchor inside the G_T about G_{IOU} loss function

实际应用中, 火源定位对 anchor 要求极其高, 需要高重合度的 anchor 与 G_T 来获取火源的中心位置.

因此, G_{IOU} 算法依旧达不到精确获取火源实际的坐标. 为解决该问题, 本文采用 C_{IOU} 损失函数代替 G_{IOU} , 如式 (3) 所示. 此时, 当 anchor 被 G_T 包裹时, 如图 3 所示, 利用 anchor 和 G_T 中心点之间欧式距离, 并添加宽高比的尺度信息来计算损失值, 如式 (5) 所示. 以上改进不仅提升了目标的检测精度, 在双目火源检测中, 精准的 anchor 还可有效降低火源三维坐标的误差, 从而提高定位精度.

$$C_{IOU} = 1 - I_{OU} + \frac{\rho^2(b, b_g)}{c^2} + \alpha v \quad (3)$$

$$\alpha = \frac{v}{(1 - I_{OU}) + v} \quad (4)$$

$$v = \frac{4}{\pi^2} \left(\arctan \frac{w_{gt}}{h_{gt}} - \arctan \frac{w}{h} \right)^2 \quad (5)$$

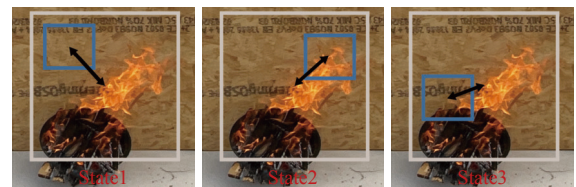


图 3 C_{IOU} 示意图

Fig. 3 C_{IOU} schematic diagram

式中, w_{gt}, h_{gt} 分别为 G_T 的宽和高, w, h 分别为 anchor 的宽和高, v 是衡量目标预测框长宽比一致性的参数, $\rho^2(b, b_g)$ 表示 anchor 与 G_T 之间的欧氏距离.

此外, 原网络的 Backbone 与 Neck 部分中采用

Leaky RELU 激活函数,因正、负函数不同导致无法为正、负输入值提供一致的预测关系,造成火源检测的精度有所损失. 为此将初始网络结构中 Leaky RELU 更改为 G_{ELU} 激活函数,如式(6)所示. 将正则化思想引入激活函数中,当 x 需分配更大值时,将 SOI 映射定义为随机地将单位映射应用于输入,引入不确定性,保持对输入值的依赖性. 通过大量实验证明,同一实验,实验效果 G_{ELU} 要比 ReLU、ELUs 等在 Classification Error(%) 检测中在同 Epoch 下相比错误率更小,改进后的网络如图 4 所示,新网络激活函数由红色框标出.

$$G_{ELU}(X) = x \times P(X \leq x) = x \times \phi(x), x \sim N(0, 1) \quad (6)$$

式中, x 是输入值, $P(X \leq x)$ 是 X 小于或等于 x 的概率, $\phi(x)$ 为正态分布概率密度. 可近似为公式(7).

$$G_{ELU}(X) = 0.5x(1 + \tanh[\sqrt{2/\pi}(x + 0.044715x^3)]) \quad (7)$$

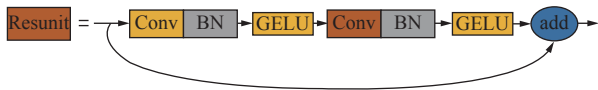


图 4 新网络结构

Fig. 4 New network structure

3 火源定位

火源定位方法主要可概括为单目和双目定位两类. 单目定位需根据检测视场,确定摄像机角度、水平高度等各种系统参数,若系统参数发生变化,则需重新进行参数标定^[25]. 火灾防控预警中,火源出现的位置不确定,往往需要大视场范围,而双目定位可以进行灵活的大视场检测,相比单目定位监控区域宽泛许多. 另外,利用 C_{IOU} 损失函数 G_T 与 anchor 的重合度更高使得双目定位效果更为显著,所以系统采用双目定位获得火源三维坐标.

双目定位通过视差原理来计算物体坐标,由于摄像头位置不同导致对同一物体进行检测会存在视差. 双目定位一般由摄像机放置不同位置可分为平行双目定位模型和双目汇聚定位模型^[26]. 本文利用平行双目定位模型计算火源三维坐标,调整两个摄像头使摄像头光轴平行,利用左右摄像头的视差原理分别在两个摄像头成像火源坐标,测量摄像头焦距与摄像头的间距,使用相似三角形法来得到物体坐标之间的关系,通过计算得出火源的世界坐标系的三维坐标. 如图 5 所示,令左摄像机为 C_1 ,右摄像机为 C_2 ,俩摄像头的

参数使光轴平行,测量摄像机焦距 f_1, f_2 令 $f_1 = f_2 = f$,测量两摄像机之间的距离间距 B ,设世界坐标系下火源的坐标为 $P(X, Y, Z)$,火源在左摄像机中映射图像坐标为 $P_1(u_1, u_2)$,右摄像机中映射图像坐标为 $P_2(u_1, u_2)$,

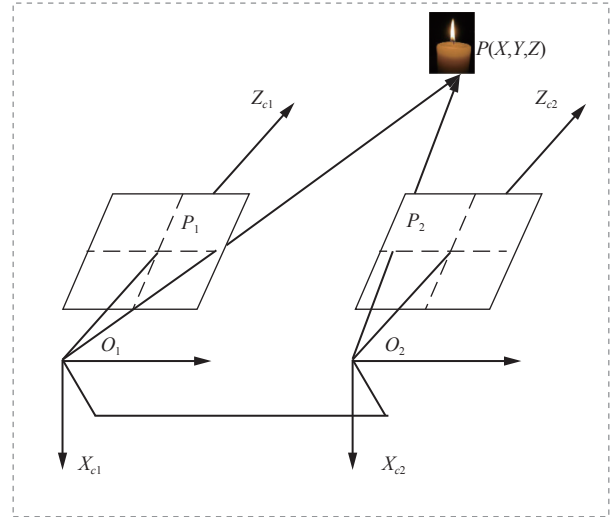


图 5 平行双目定位模型原理

Fig. 5 Principle of parallel binocular positioning model

P_1, P_2 之间距离为 a . 由 P 点投影在两幅图像上点之间的距离 P_1P_2 , 利用相似三角形得到与 X, Y 之间的关系,如式(8)所示:

$$\frac{a}{B} = \frac{Z-f}{Z}, \frac{2Y}{v_1+v_2} = \frac{Z}{f}, \frac{2X}{u_1+u_2} = \frac{Z}{f} \quad (8)$$

将世界坐标系与摄像机 C_1 重合,计算出 P_1, P_2 之间的距离:

$$a = B + u_1 - u_2 \quad (9)$$

视差 $d = u_1 - u_2$, 带入式(8)可得:

$$Z = B \frac{f}{d}, X = B \frac{u_1 + u_2}{2d}, Y = B \frac{v_1 + v_2}{2d} \quad (10)$$

则可得出 P 点三维坐标为 (X, Y, Z) .

4 模型部署

考虑火源检测实际运用的可行性,检测算法部署需小规模嵌入式装置,由于 YOLO V5 轻量级的优点,本文采用 Raspberry PI 作为嵌入式平台,如图 6 所示. 搭载 Intel AI-神经计算棒提升算力,增加检测处理速度;采用 CCD 视觉传感器,影像像素为 25 万、彩色分辨率为 330 线;PC 端训练火焰检测模型,Raspberry PI 使用 Opencvino 部署中间模型文件实现火源检测.

5 实验与分析

5.1 实验环境

环境配置如表 1 所示,实验训练参数:输入图像



图 6 模型部署单元

Fig. 6 Model deployment unit

大小为 640*640; 初始学习率设置为 0.01; 每个 batch 大小设置为 16; 最大迭代次数为 300; 动量因子为 0.937.

表 1 环境配置
Tab. 1 Environment configuration

参数	配置
操作系统	Window10
显存	8.0 G
内存	8.0 G
GPU	NVIDIA GoForce GTX 1 650
GPU环境	CUDA11.4
训练框架	Pytorch 1.7.0、python 1.8.0

5.2 网络训练与评价指标

实验采用自制数据集,对火源的各类情况进行数据收集,共包含 12 000 张静态图像,使用 Mosaic 数据增强的方式将图片随机缩放、裁剪、排布. 利用 labelimg 工具对火源图像进行标注. 按照 7 : 3 的比例将数据集分为训练集、测试集; 每张图片分辨率为 1 596*640. 选取部分标注过的数据集样本如图 7 所示,其中绿色框为标注过的框.

通过训练,比较改进前后损失值图像,结果如图 8 所示,在相同 Epoch 下,GELU 要比 Leaky RELU 的损失值更低,由此可得,对模型做出改进可有效加快损失值的下降速度,因此可适当减少学习率延缓模型训练中损失值的下降速度,用较小的学习率训练模型,可进一步提高目标的检测精度. 典型评价指标主要包含 mAP、AP_{0.5}、FPS 等,精度 AP 是指精确率 P 和召回率 R 所围成的曲线面积; mAP 为 IOU 阈值从 0.5 到 0.9,以 0.05 间隔,检测该阈值下 IOU 值,最后取十个值计算平均结果; AP_{0.5}是指阈值在 0.5 时的

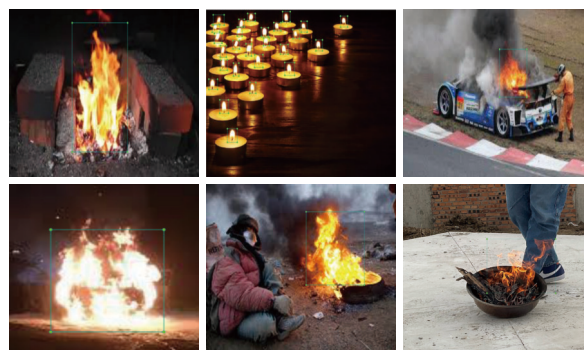


图 7 部分数据集

Fig. 7 Partial data set

IOU 值,FPS 是指每秒处理的图片数量. 本实验选取典型测试指标 mAP、AP_{0.5}、FPS 进行性能评估.

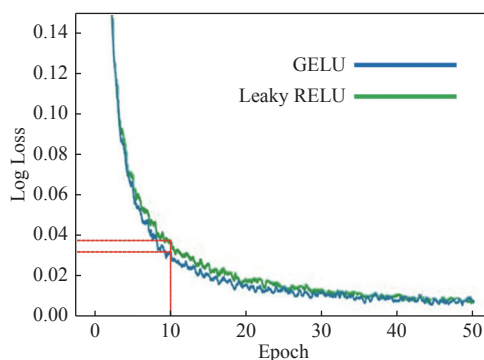


图 8 Leaky RELU 和 GELU 函数的损失值下降曲线

Fig. 8 Leaky RELU and GELU function of the loss value drop curve

5.3 结果与讨论

5.3.1 改进前后算法比较

图 9 给出了在添加 GELU 与 CIOU 后的 YOLO V5 与原始模型对同一张图像的检测结果,图中绿色框表示原 YOLO V5,蓝色框表示改进损失函数后的 YOLO V5,红色框表示对损失函数与激活函数同时做出改进的 YOLO V5,不难发现,在改进损失函数后模型的检测精度有所提升,但仍会出现部分无用背景信息,当损失函数与激活函数同时做出改进时,anchor 对火源的包裹更紧密,因此本次改进对火源的检测精度整体有显著提升.

利用 mAP、AP、FPS 值对不同模型性能进行比较,如表 2 所示. YOLO V5 在增加 GELU 激活函数后,模型 mAP_{0.5:0.9}和 mAP 分别提升了 1.9% 和 4.1%,精准率、召回率相比原网络均有提升. YOLO V5 在增加 CIOU 损失函数后,模型 mAP_{0.5:0.9}和 mAP 更是提升了 4% 和 6.8%,相比添加 GELU 激活函数性能提升更为明显. 将其共同融入网络,Our YOLO V5 模型

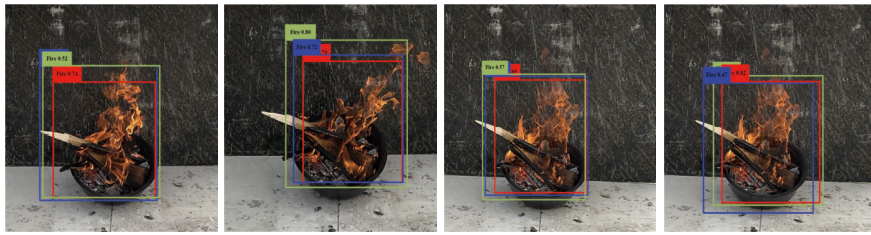


图 9 IOU 对比图

Fig. 9 IOU comparison diagram

mAP_{0.5,0.9} 可达 0.457, 较原 YOLO V5 提升了 9.8%, 改进后相比原 YOLO V5 提升了 14.2%, 整体检测精度有显著提升. 召回率相比原 YOLO V5 提升了 21.7%, 精准率提升了 7.2%, 检测性能也有大幅度的提升. 利用同一数据集对 YOLO、SSD 与 RCNN 检测模型进行训练, 改进的 YOLO V5 精准率、召回率与 mAP 均高于其他几种检测算法, 具有明显的优势.

5.3.2 双目定位与实际距离比较

Our YOLO V5 模型对火源的检测效果如图 10 所示, 改进后的检测模型减少了大量的无用背景信息, anchor 对火源的包裹更紧凑且检测结果的置信度更高.



图 10 检测效果图

Fig. 10 Check the effect drawing

对本方法进行实验分析, 结果如表 3 所示, z_1 、 z_2 分别表示距离摄像机三米与七米处的火源位置. 从表中可以看出, 三米与七米处的相对误差在 3.8%、7.1% 左右, 由于火源本身形状多变且不稳定, 所以该误差对定位精度带来的影响十分有限. 进一步观察表 3, 可以发现, 尽管本文所提方法可精确定位三米内的微型火源, 并有效定位七米内的初期火源, 但随着火源距离的增大, 误差也呈一定幅度上涨.

对本方法相对误差进行实验分析, 如图 11 所示, 图中实线表示相对误差拟合曲线, 虚线表示相对误差上下的实际浮动. 从图中可以看出, 将不同距离的实验数据进行对比, 得出 3 m 相比 7 m 处的误差幅度更加平缓. 在多数情况下, 在 3 m 处的微型火源获得了

表 2 不同算法检测效果对比

Tab. 2 Detection effect comparison of different algorithms

	mAP _{0.5,0.9}	mAP _{0.5}	P	R
Faster-RCNN	0.414	0.812	0.794	0.748
SSD	0.409	0.804	0.787	0.743
YOLO V3	0.391	0.756	0.778	0.731
YOLO V4	0.408	0.811	0.785	0.739
YOLO V5	0.416	0.813	0.796	0.750
YOLO V5+GELU	0.424	0.846	0.811	0.826
YOLO V5+CIUO	0.433	0.868	0.837	0.871
Our YOLO V5	0.457	0.929	0.854	0.913

相对稳定的定位实验效果, 而在 7 m 处的相对误差虽幅度相较偏大, 这与双目定位中定位误差随着目标距离的增加而增加的理论知识相吻合.

6 结束语

众所周知, 高效检测与精准定位火源是火灾预警的关键. 为此, 本文对 YOLO V5 激活函数与损失函数进行改进, 将原算法 GIOU 损失函数更改为 CIUO 损失函数, 有效的提高了模型 anchor 与 GT 的重合度; 使用 GELU 来替换原算法的 Leaky RELU 激活函数, 通过 GELU 函数的正则化思想进一步提高了火源的检测精度; 准确检测的同时利用平行双目定位模型对火源进行定位. 实验结果表明, 本文提出的算法相

表3 单目与双目误差对比

Tab. 3 Comparison of monocular and binocular errors

	x/m	y/m	z ₁ /m	z ₂ /m
实际距离	0.304	0.607	3.002	7.003
	0.308	0.599	3.121	7.542
	0.301	0.611	3.110	7.451
预测距离	0.306	0.601	3.115	7.500
	0.307	0.612	3.119	7.480
	0.309	0.614	3.122	7.530
相对误差	0.013	0.013	0.040	0.077
	0.007	0.007	0.036	0.064
	0.010	0.010	0.038	0.073
	0.010	0.008	0.039	0.068
	0.016	0.012	0.040	0.075

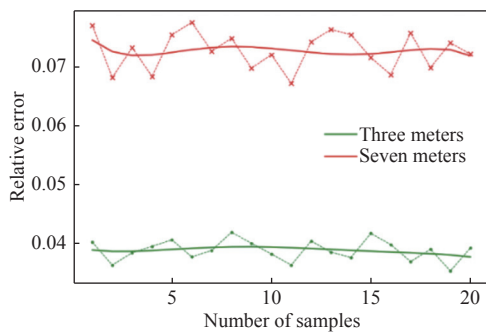


图11 误差比较

Fig. 11 Error comparison

比原算法 mAP 提升了 9.8 个百分点; 利用改进后的 YOLO V5 算法对火源进行双目定位, 实验证明可有效定位火源位置, 且随着火源的增大, 检测的距离也会随之增加. 如何进一步提升更远距离的精准定位是下一步主要的研究方向.

参考文献:

- [1] 王德帅. 高层建筑消防灭火救援的难点[J]. 居舍, 2021(7): 153-154.
WANG D S. Difficulties in fire fighting and rescue of high-rise buildings[J]. Residential Journal, 2021(7): 153-154.
- [2] 崔鹏程, 陈鑫, 张涛. 基于PyroSim的粮食储备库火灾仿真模拟分析[J]. 粮油食品科技, 2021, 29(2): 211-217. DOI: 10.16210/j.cnki.1007-7561.2021.02.029.
CUI P C, CHEN X, ZHANG T. Simulation analysis of the grain storage fire based on the PyroSim[J]. Science and Technology of Cereals, Oils and Foods, 2021, 29(2): 211-217. DOI: 10.16210/j.cnki.1007-7561.2021.02.029.
- [3] LOWE D G. Distinctive image features from scale-invariant keypoints[J]. International Journal of Computer Vision, 2004, 60(2): 91-110. DOI: 10.1023/B:VISI.0000029664.99615.94.
- [4] DALAL N, TRIGGS B. Histograms of oriented gradients for human detection[C]//Proceedings of 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. San Diego, CA, USA: IEEE, 2005: 886-893. DOI: 10.1109/CVPR.2005.177.
- [5] FELZENSZWALB P, MCALLESTER D, RAMANAN D. A discriminatively trained, multiscale, deformable part model[C]//Proceedings of 2008 IEEE Conference on Computer Vision and Pattern Recognition. Anchorage, AK, USA: IEEE, 2008: 1-8. DOI: 10.1109/CVPR.2008.4587597.
- [6] 刘鹏飞. 基于SIFT的图像配准方法研究[D]. 哈尔滨: 哈尔滨工业大学, 2019. DOI: 10.27061/d.cnki.gbgdu.2019.002598.
LIU P F. Research on image registration method based on SIFT[D]. Harbin: Harbin Institute of Technology, 2019. DOI: 10.27061/d.cnki.gbgdu.2019.002598.
- [7] 孙伟. 基于生成网络和分类器融合的小样本立体目标识别[D]. 西安: 西安电子科技大学, 2018.
SUN W. Stereo target recognition for small dataset based on generation network and classifiers fusion[D]. Xi'an: Xidian University, 2018.
- [8] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]//Proceedings of 2014 IEEE Conference on Computer Vision and Pattern Recognition. Columbus: IEEE, 2014. DOI: 10.1109/CVPR.2014.81.
- [9] GIRSHICK R. Fast RCNN[C]//Proceedings of 2015 IEEE International Conference on Computer Vision. Santiago: IEEE, 2015: 1440-1448. DOI: 10.1109/ICCV.2015.169.
- [10] REN S Q, HE K M, GIRSHICK R, et al. Faster RCNN: towards real-time object detection with region proposal networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(6): 1137-1149. DOI: 10.1109/TPAMI.2016.2577031.
- [11] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: unified, real-time object detection[C]//Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas: IEEE, 2016: 779-788. DOI: 10.1109/CVPR.2016.91.
- [12] TIAN Z, SHEN C H, CHEN H, et al. FCOS: a simple and strong anchor-free object detector[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2022, 44(4): 1922-1933. DOI: 10.1109/

- TPAMI.2020.3032166.
- [13] LIU W, ANGUELOV D, ERHAN D, et al. SSD: single shot MultiBox detector[C]//Proceedings of the 14th European Conference on Computer Vision. Amsterdam: Springer, 2016: 21-37. DOI: [10.1007/978-3-319-46448-0_2](https://doi.org/10.1007/978-3-319-46448-0_2).
- [14] 华夏, 王新晴, 王东, 等. 基于改进SSD的交通大场景多目标检测[J]. *光学学报*, 2018, 39(12): 1215003. DOI: [10.1587/transele.EI01.C.253](https://doi.org/10.1587/transele.EI01.C.253).
- HUA X, WANG X Q, WANG D, et al. Multi-objective detection of traffic scenes based on improved SSD[J]. *Acta Optica Sinica*, 2018, 39(12): 1215003. DOI: [10.1587/transele.EI01.C.253](https://doi.org/10.1587/transele.EI01.C.253).
- [15] 刘宽, 郎磊. 轻量化SSD目标检测方法研究[J]. *湖北民族大学学报(自然科学版)*, 2021, 39(4): 418-424. DOI: [10.13501/j.cnki.42-1908/n.2021.12.010](https://doi.org/10.13501/j.cnki.42-1908/n.2021.12.010).
- LIU K, LANG L. Research on lightweight SSD target detection method[J]. *Journal of Hubei Minzu University (Natural Science Edition)*, 2021, 39(4): 418-424. DOI: [10.13501/j.cnki.42-1908/n.2021.12.010](https://doi.org/10.13501/j.cnki.42-1908/n.2021.12.010).
- [16] 吴天成, 王晓荃, 蔡艺军, 等. 基于特征融合的轻量级SSD目标检测方法[J]. *液晶与显示*, 2021, 36(10): 1437-1444. DOI: [10.37188/CJLCD.2021-0007](https://doi.org/10.37188/CJLCD.2021-0007).
- WU T C, WANG X Q, CAI Y J, et al. Lightweight SSD object detection method based on feature fusion[J]. *Chinese Journal of Liquid Crystals and Displays*, 2021, 36(10): 1437-1444. DOI: [10.37188/CJLCD.2021-0007](https://doi.org/10.37188/CJLCD.2021-0007).
- [17] 黄真亮, 张继顺, 朱志武, 等. 目标检测算法在嵌入式平台中的应用[J]. *电子技术*, 2021, 50(8): 150-151.
- HUANG Z L, ZHANG J S, ZHU Z W, et al. Application of target detection algorithm in embedded platform[J]. *Electronic Technology*, 2021, 50(8): 150-151.
- [18] 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[2021-12-25]. <https://doi.org/10.13637/j.issn.1009-6094.2021.1645>.
- ZHANG Q, ZHANG W, YANG X F, et al. Smoke and flame detection method with YOLOv5-Resnet cascade network[J/OL]. *Journal of Safety and Environment*: 1-10[2021-12-25]. <https://doi.org/10.13637/j.issn.1009-6094.2021.1645>.
- [19] REDMON J, FARHADI A. YOLO9000: better, faster, stronger[C]//Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu: IEEE, 2017: 7263-7271. DOI: [10.1109/CVPR.2017.690](https://doi.org/10.1109/CVPR.2017.690).
- [20] REDMON J, FARHADI A. YOLOv3: an incremental improvement[J]. *Computer Vision and Pattern Recognition*, 2018, 276(7): 126-134.
- [21] ALEXEY B, WANG C Y, HONG Y. Yolov4: optimal speed and accuracy of object detection[J]. *Computer Vision and Pattern Recognition*, 2020, 10(9): 34-51.
- [22] DENG H F, CHENG J H, LIU T, et al. Research on iron surface crack detection algorithm based on improved YOLOv4 network[J]. *Journal of Physics:Conference Series*, 2020, 1631: 012081. DOI: [10.1088/1742-6596/1631/1/012081](https://doi.org/10.1088/1742-6596/1631/1/012081).
- [23] HE K M, ZHANG X Y, REN S Q, et al. Deep residual learning for image recognition[C]//Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas: IEEE, 2016: 770-778. DOI: [10.1109/CVPR.2016.90](https://doi.org/10.1109/CVPR.2016.90).
- [24] WANG C Y, LIAO H Y M, WU Y H, et al. CSPNet: a new backbone that can enhance learning capability of CNN[C]//Proceedings of 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops. Seattle: IEEE, 2020: 1571-1580. DOI: [10.1109/CVPRW50498.2020.00203](https://doi.org/10.1109/CVPRW50498.2020.00203).
- [25] 梁立超. 基于单目视觉的摄像机标定与立体匹配技术的研究[D]. 沈阳: 沈阳工业大学, 2018.
- LIANG L C. Research on camera calibration and stereo matching technology based on monocular vision[D]. Shenyang: Shenyang University of Technology, 2018.
- [26] 杨狄赛. 图像型消防水炮火点识别算法及执行机构控制[D]. 福州: 福建农林大学, 2016.
- YANG D S. Fire recognition algorithm and actuator control of fire protection cannon system based on image processing[D]. Fuzhou: Fujian Agriculture and Forestry University, 2016.

作者简介:

张旭彤 男,(2000-).研究方向为视觉检测、传感器信息融合.

胡鹏(通讯作者) 男,(1990-),博士,讲师.研究方向为计算机视觉、数字图像处理、多传感器信息融合等.

E-mail: aust_hp@163.com.

赵鑫 男,(1991-),博士,讲师.研究方向为机器视觉、激光技术、信号处理、光声成像技术.

丁云霞 女,(1993-),硕士,实验师.研究方向为计算机视觉、自然语言处理等.