

基于改进 YOLOv5s 的四旋翼 自主降落标识检测算法

李晓轩¹, 甄子洋¹, 刘彪², 梁永勋¹, 黄祎闻¹

(1. 南京航空航天大学 自动化学院, 南京 210016;

2. 沈阳飞机设计研究所 扬州协同创新研究院, 江苏 扬州 225006)

摘要: 无人机自主降落是无人机领域研究的热点之一, 导航信息在自主降落过程中又起到至关重要的作用, 而视觉导航相较于传统导航方式可以提供更多环境信息, 有利于提高无人机着陆安全性; 当无人机飞行高度越高, 机载相机捕获到的降落标识物就越小, 为了提升无人机识别标识物的能力, 基于 YOLOv5s 算法提出了一种改进的无人机实时小目标检测算法; 首先, 为了检测到更小尺度的目标在原算法基础上新增一个检测头; 然后采用 BiFPN 代替原先 PANet 结构, 提升不同尺度的检测效果; 最后将 EIoU Loss 替换 CIoU Loss 作为算法的损失函数, 在提高边界框回归速率的同时提高模型整体性能; 将改进算法应用于无人机自主降落场景下的二维码降落标识检测, 实验结果表明改进后的算法在小目标检测中相比于原始 YOLOv5s 算法的特征提取能力更强、检测精度更高, 证明了改进算法的优越性。

关键词: 无人机; 小目标检测; YOLOv5s; BiFPN; EIoU Loss; 嵌入式平台

Object Detection of Quadrotor UAV Auto-Landing Based on Improved YOLOv5s

LI Xiaoxuan¹, ZHEN Ziyang¹, LIU Biao², LIANG Yongxun¹, HUANG Yiwen¹

(1. College of Automation Engineering, NUAU, Nanjing 210016, China;

2. Aviation Industry, Yangzhou Research Institute, Yangzhou 225006, China)

Abstract: UAV auto-landing is one of the hotspots in UAV research field, its navigation information plays an important role in the process of autonomous landing. And compared with traditional navigation methods, visual navigation can provide more environmental information, which is prone to improve the landing safety of a UAV. when the UAV flies higher, the landing marker captured by an airborne camera is smaller. In order to improve the ability of the UAV to recognize small landing marker, an algorithm based on YOLOv5s is proposed. Firstly, a prediction head is added to detect smaller targets. Secondly, the bi-directional feature pyramid network (BiFPN) is replaced original path aggregation network (PANet) to improve the identification ability of the UAV. Finally, the loss function of the algorithm is taken by using EIoU Loss instead of CIoU Loss, which improves the regression rate of the boundary box and overall performance of the model. The improved algorithm is applied to the marker detection of UAV auto-landing scene. The results show that compared the original YOLOv5s algorithm, the proposed algorithm has stronger feature extraction ability and higher detection accuracy in small object detection, which proves the superiority of the improved algorithm.

Keywords: UAV; object detection; YOLOv5s; BiFPN; EIoU Loss; embedded platform

0 引言

随着无人机技术的飞速发展, 大量难题已被解决, 但仍然还存在一些需要解决的问题, 其中就包括无人机回收技术^[1-2]。在无人机导航领域中, 无人机自主降落系统已成为研究热点。传统的导航方式有 GPS 全球卫星定位导航、惯性导航、无线电导航等。GPS 采用的是多卫星单程测距

系统, 如果接收机测量到了其与空间中至少四颗导航卫星之间的距离, 那么就可以通过计算从而确定接收机的位置。但 GPS 受制于美军, 且信号接收容易受到无人机所处环境的影响。因此, 在无人机自主降落时仅使用 GPS 是非常受限的。惯性导航的应用是最广泛的, 它不依赖于外界信息, 仅通过其内部的加速度计以及陀螺仪就可以确定自身导航信息。但是随着时间增加会导致它的估计误差积累, 所以

收稿日期: 2022-10-18; 修回日期: 2022-11-07。

作者简介: 李晓轩(1997-), 男, 安徽滁州人, 硕士研究生, 主要从事计算机视觉与控制方向的研究。

甄子洋(1981-), 男, 浙江浦江人, 教授, 博士研究生导师, 主要从事舰载飞机着舰引导与控制、无人机集群智能协同控制方向的研究。

引用格式: 李晓轩, 甄子洋, 刘彪, 等. 基于改进 YOLOv5s 的四旋翼自主降落标识检测算法[J]. 计算机测量与控制, 2023, 31(6): 80-86.

它往往与其他导航方法结合使用。无线电导航利用了电磁波的传播特性, 缺点在于无线电信号容易被干扰且需要自身主动发射电磁波信号, 这就容易暴露自身位置。

由于传统导航方式有上述的各种缺点, 利用计算机视觉信息作为导航信息实现无人机自主降落已成为热点。计算机视觉是多学科交叉的产物, 是包含大量信息的被动信息源。无人机的计算机视觉导航通过在无人机上安装机载相机, 从而通过相机捕捉到实时图像, 并输入到机载计算机中, 通过各种算法计算出无人机的各项导航参数, 然后将这些导航信息提供给控制系统从而完成各项任务。相机是一种成本相对较低的传感器, 且可以提供丰富的环境信息。相对于其他传感器, 基于视觉的自主降落技术能够捕获大量的环境信息, 且不依赖外部的其他设备, 不易受干扰, 在电子对抗时有较大优势, 因此使用计算机视觉已经成为国内外的无人机自主降落技术以及着舰技术的主流^[3-4]。

确定了计算机视觉的导航方式后, 就要考虑选择合适的目标检测算法。传统的目标检测算法一般由人工提取目标特征, 由于检测过程比较复杂, 这就导致了目标检测的效率不高, 精度较低, 效果较差。近年来, 由于深度学习技术的研究取得了一定的突破, 基于深度学习的卷积神经网络 (CNN)^[5] 在目标检测中的应用也取得了长足的进展, 人们已经开始逐渐抛弃传统的人工提取图像中特征的方式, 而使用卷积神经网络来自动提取图像中的特征并进行分析。基于深度学习的目标检测算法大大提高了目标检测效果, 且泛化能力更强, 应用范围更广, 在实际情况下更容易应用, 已应用于自动驾驶^[6]、人脸识别^[7]、智慧医疗^[8-9]、工业检测^[10]、活动识别^[11] 等领域中。目前, 基于深度学习的检测算法最常见的两种方法分别是基于区域的卷积神经网络 (R-CNN) 系列为代表的基于候选框的二阶段目标检测算法和以 YOLO (You Only Look Once) 系列为代表的基于回归的单阶段目标检测算法。二阶段目标检测算法虽然在精度上有着较好的表现, 但检测速度较慢不适用于实时检测系统。而单阶段目标检测算法在有着较快检测速度的同时, 在检测精度上也表现不俗。

二阶段目标检测算法的代表是 R-CNN 系列算法。美国的工程师 R. Girshick 在 2013 年第一次将 CNN 应用到目标检测领域, 即 R-CNN^[12]。它使用 CNN 自动提取图片中的特征信息, 在训练网络时先采用大样本进行有监督的预训练, 再通过小样本进行微调, 以此来解决模型训练过度产生过拟合的问题。这种做法相对于传统方式极大地提升了检测性能, 但它需要对图像中的候选区域的卷积特征进行很多重复的计算。为了解决图片缩放后会产生形变, 提出了 SSP Net 网络^[13], 它的特点是无论输入什么尺寸的图片, 输出的大小都是固定的, 这就令 SSP Net 可以提取不同尺度的特征, 且仅需提取一次图片中候选区域的卷积特征, 大大节省了计算资源的浪费, 加快了检测速度。在此基础

上又提出了 Fast R-CNN^[14], 它在 R-CNN 中结合了 SSP Net 的思想, 不但提高检测的精度还提高了检测的速度。之后又提出了 Faster R-CNN^[15], 它引入了 RPN (region proposal networks, 区域检测网络), 进一步提升了检测的精度与速度。然而检测速度与单阶段目标检测算法相比仍有一定差距。

单阶段目标检测算法主要包括 YOLO 系列算法以及 SSD (single shot multibox detector) 算法^[16]。2015 年, 单阶段目标检测算法的开山之作 YOLO 系列算法被首次提出^[17]。YOLO 算法最后采用全连接层得到检测结果, 而 SSD 算法直接使用卷积提取不同尺度的特征图来得到检测结果。YOLOv2^[18] 在 YOLOv1 的基础上进行改进, 它在引入锚框机制的同时还引入了批标准化 (BN, batch normalization) 的概念。YOLOv3^[19] 又进行了进一步改进, 它的骨干网络 (Backbone) 使用 darknet53, 并将锚框的大小扩增到 3 种类型, 此外使用 sigmoid 函数作为逻辑分类器将输出限制在 0 和 1 之间, 这大大提高了 YOLOv3 的推理速度。在此基础上改进又提出了 YOLOv4 算法^[20], 它用 Mish 函数代替原先的 ReLU 函数作为激活函数, 因为 Mish 的函数图像比 ReLU 更加平滑, 在检测速度与精度之间做到了较好权衡。而 YOLOv5 将之前版本的优点融会贯通, 无论是在精度还是速度上, 都有了较大的进步^[21]。YOLOv5 算法包括 YOLOv5n, YOLOv5s, YOLOv5m, YOLOv5l, YOLOv5x 五个版本, 他们的参数量逐渐增加, 检测精度也随之升高, 但检测速度会逐渐减慢。YOLOv5s 在有着良好检测精度的同时速度也较快^[22], 适用于无人机自主降落实时系统中, 但在一些独特场景下对小目标的检测仍有一些不足。

本研究以二维码构成的降落标识物为检测目标, 如图 1 所示, 采集一千多张不同高度不同角度无人机拍摄的降落标识物图片, 构建降落标识物检测数据集。基于 YOLOv5s 检测网络, 为了增强对小目标检测的能力, 做出如下改进:

- 1) 增加小目标检测层, 即增加了一个预测头来检测更小尺度的目标;
- 2) 采用 BiFPN 加权双向特征金字塔网络代替原先 PA-Net 结构, 快速进行多尺度特征融合;
- 3) 将 EIou Loss 替换 CIou Loss 作为算法的损失函数, 在提高边界框回归速率的同时提高模型整体性能。

1 YOLOv5s 算法原理

YOLOv5 算法的速度快, 灵活性高, 该算法在 YOLOv4 的基础上做出了一些改进, 使得其速度与精度都有了一定提高。它由 4 个部分组成, 分别是输入端、Backbone、Neck 和 Head。YOLOv5s 的网络结构如图 2 所示。

输入端起到预处理输入数据集的作用, 包括自适应锚框计算、Mosaic 数据增强等步骤。在网络训练时, 模型首先会在初始锚点框的基础上输出得到的预测框, 然后计算预测框与真实框的差值, 接下来就可以反向更新网络中的

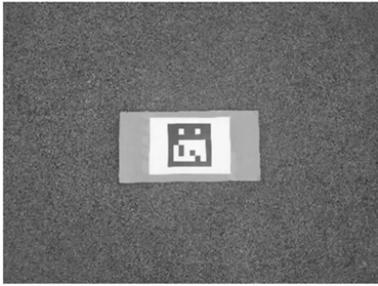


图 1 降落标识物

入特征相加；SPP 是空间金字塔池化层，它用于解决输入不是固定尺寸的情况，它能将任意大小的特征图转换成固定大小的特征向量，它在 3 种尺度上进行最大池化操作，然后将输出结果进行拼接。

Neck 是为了实现不同尺度特征信息的融合，它包括特征金字塔 (FPN)^[23] 结构和路径聚合网络 (PAN)^[24] 结构。由于高层的语义特征比低层的要强，为了增强整个金字塔的语义特征，FPN 自顶向下的将高层的强语义特征传递下来，但是它并没有传递定位信息。PAN 针对这一点不足，为了将低层较强的定位特征传递到上层，它在 FPN 的基础上又添加了一个自底向上的金字塔，弥补了 FPN 的不足。

Head 是 YOLOv5s 的检测部分，它继承了 YOLOv4 的 Head 进行多尺度预测，提高了对不同尺寸目标的检测性能，它可以输出大、中、小 3 种不同尺寸的特征图，分别用于检测小、中、大 3 种目标。

2 改进的 YOLOv5s 算法

本文以 YOLOv5s 网络为基础进行了如下 3 种改进：通过增加一个小目标检测头来提升对小尺度目标的检测能力；采用 BiFPN 加权双向特征金字塔网络代替原先 PANet 结构，快速进行多尺度特征融合；损失函数使用 EIoU Loss，提高边界框回归速率的同时提高模型整体性能。下面分别从以上 3 个方面进行详细介绍。

2.1 小目标检测头

原来的 YOLOv5s 骨干网络为了检测大中小 3 种尺度的目标使用了 3 种大小的特征图，这 3 种不同大小的特征图分别来自网络对原始输入图像进行 8 倍、16 倍和 32 倍下采样，如图 3 所示，得到这 3 种特征图后将其输入到特征融合网络中进行特征融合。虽然经过卷积次数越多的特征图的语义信息越丰富，但是代价是目标的位置信息也会慢慢丢失，这样就不利于检测小目标。相反的，虽然卷积次数少，特征图语义信息不足，但是目标的位置信息更加准确。因此为了提升小目标的检测效果，本文在 YOLOv5s 原有的网络上将 4 倍下采样得到的特征图也输入到特征网络中，从而得到新的大小的特征图，此特征图的感受野小，且拥有更加丰富的位置信息，能对更小尺度的目标有更好的检测效果，如图 4 所示。

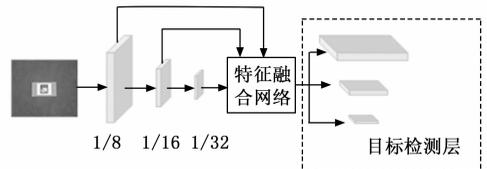


图 3 原始 YOLOv5s 目标检测层

2.2 BiFPN 结构

BiFPN 是加权双向特征金字塔，BiFPN 有两个主要思想：1) 高效的双向跨尺度连接；2) 加权特征融合。BiFPN 构造了自上而下、自下而上的双向通道，对主干网络中的

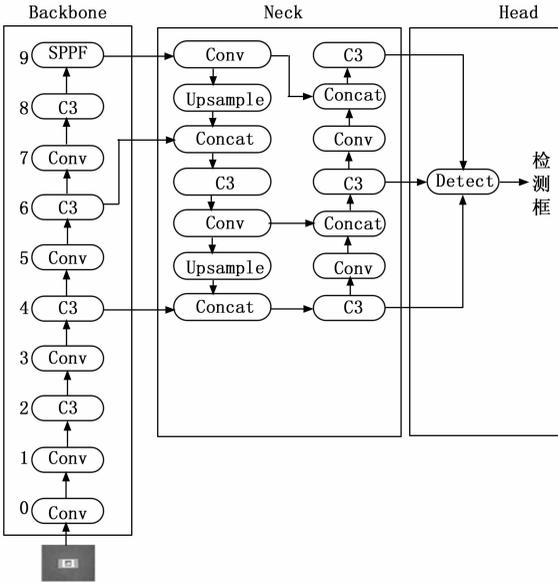


图 2 YOLOv5s 网络结构图

各项参数，因此设定初始锚点框也是比较关键的一环。在之前的 YOLO 系列检测算法中，训练不同的数据集时，为了获得初始锚点框，需要运行单独的程序来获得，而 YOLOv5s 的代码集成了此项功能，每次训练时，依赖数据集的名称，它可以进行自适应的计算，从而得到最佳的锚点框。Mosaic 数据增强借鉴了 CutMix 增强的方法，只不过本方法采用了 4 张图片，对其进行随机裁剪、缩放、旋转等操作，从而合成了 1 张图片，采用 Mosaic 数据增强的方法，不仅使图片能丰富检测目标的背景，而且能够提高小目标的检测效果。并且在 Batch Normalization 计算的时候一次性会处理四张图片。

Backbone 用于提取输入图像的信息特征，它使用了 CSPDarknet53 网络。Backbone 作为 YOLOv5s 的骨干结构，还包含 Focus、Conv、C3、SPP 等模块。其中 Focus 结构在 YOLOv1-YOLOv4 中没有引入，Focus 模块重要的是切片操作，它从纵向和横向两通道对输入进行切片后重新拼接；Conv 中主要的操作有卷积、正则化和激活层；C3 中包括 N 个残差网络 Bottleneck，输入的特征图首先经过 1×1 的卷积层，再经过 3×3 的卷积层，然后将得到的结果直接与输

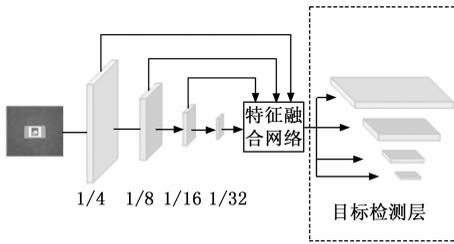


图 4 改进算法的目标检测层

不同尺度信息进行融合。在融合时, 通过上采样和下采样将不同尺度信息变为同一尺度, 并在同一特征的原始输入节点和输出节点之间添加横向连接, 这样不仅不增加成本还能融合更多特征。相比于 PANet, 它还移除了单输入边的结点, 因为单输入边的结点没有进行特征融合, 故具有的信息比较少, 对于最后的融合没有什么贡献度, 相反, 移除还能减少计算量。加权特征融合简单来说, 就是针对融合的各个尺度特征增加一个权重, 调节每个尺度的贡献度, 这其实是注意力机制与 FPN 的碰撞。此外, 还可以将 BiFPN 看成一个基本单元, 通过重复堆叠几次此单元的方式可以融合更多特征, 它的网络结构如图 5 所示。

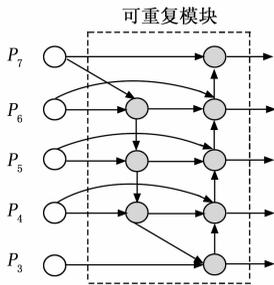


图 5 BiFPN 结构

2.3 EIoU Loss

本文采用 EIoU Loss 替代原始模型的 CIoU Loss 作为损失函数。EIoU Loss 的收敛速度更快, 评价也更加准确, 不但可以提高模型整体性能, 而且在一定程度上可以弥补多尺度特征融合模块增加导致的训练时间增加、收敛速度变慢的问题。EIoU Loss 是针对 CIoU Loss 进行改进的, CIoU Loss 的计算公式为:

$$L_{CIoU} = 1 - I + \frac{\rho^2(b, b^{gt})}{c^2} + \theta v \quad (1)$$

式中, I 为真实框与预测框的交并比, b^{gt} 与 b 分别代表真实框的中心点与预测框的中心点; ρ 代表真实框的中心点与预测框的中心点之间的欧式距离; c 代表可以同时包含真实框和预测框的最小外接四边形的对角线长度; θ 是一个权重系数, 其表达式为:

$$\theta = \frac{v}{(1 - I) + v} \quad (2)$$

式中, v 是两框长宽比的距离, 其表达式为:

$$v = \frac{4}{\pi^2} \left(\arctan \frac{w^{gt}}{h^{gt}} - \arctan \frac{w}{h} \right)^2 \quad (3)$$

式中, w^{gt} 、 h^{gt} 分别为真实框的长和宽; w 、 h 分别为预测框的长和宽。

从公式可以看出 CIoU Loss 的优点在于同时兼顾了真实框与预测框的交并比即重叠面积、中心点距离和纵横比, 但缺点是其公式中的 v 反映的是两框长宽比的差异, 没有考虑到宽度和高度分别与其置信度的真实差异, 所以有时会降低模型优化的速率。而 EIoU Loss 继承了 CIoU Loss 的优势, 沿用了交并比损失 L_{IoU} 、中心点距离损失 L_{dis} , 而宽高损失 L_{asp} 直接设置为真实框和预测框之间的宽度和高度最小差, 使收敛速度更快。计算公式如下:

$$L_{EIoU} = L_{IoU} + L_{dis} + L_{asp} = 1 - I + \frac{\rho^2(b, b^{gt})}{c^2} + \frac{\rho^2(w, w^{gt})}{c_w^2} + \frac{\rho^2(h, h^{gt})}{c_h^2} \quad (4)$$

式中, c_w 和 c_h 为可以包含真实框与预测框的最小外接矩形框的宽和高。

经过上述 3 种改进后, 算法的整体网络结构如图 6 所示。

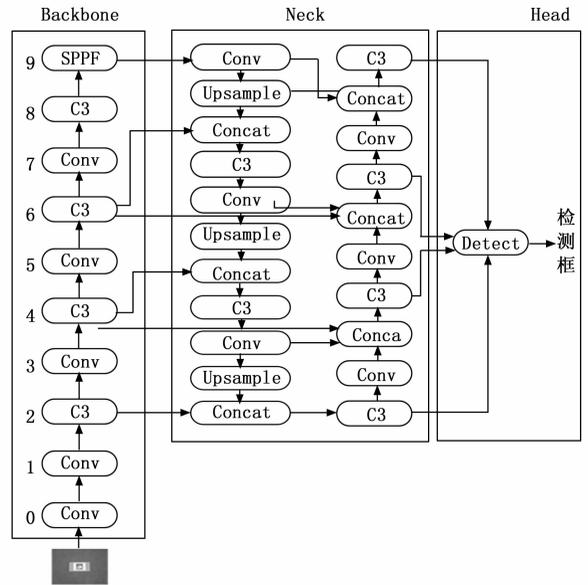


图 6 改进的 YOLOv5s 网络结构图

3 实验

3.1 数据集

本文主要针对无人机自主降落场景下的降落标识进行视频采集, 共计采集 1 350 张图片, 其中 80% 用于训练集, 15% 用于验证集, 5% 用于测试集, 其中部分数据如图 7 所示, 他们是在不同天气, 不同背景, 不同角度, 以及不同高度下拍摄的二维码降落标识的图片。采集完数据集后, 下一步就需要对数据集进行标注, 以便网络学习, 利用 Make Sense 网页对数据集进行在线标注, 标注成 YOLO 格式, 共有 1 个标记类别, 是 QRcode (降落标识)。完成标

注后，每一张图片都会生成一个 txt 文件，文件名与图片名相同。txt 文件中共有 n 行 5 列，每一行都代表一个检测目标的标记框，其中第一列代表该框的类别，第二列代表框中心的横坐标与图片宽度之比，第三列代表框中心的纵坐标与图片高度之比，第四列代表框的宽度与图片的宽度之比，最后一列代表框的高度与图片的高度之比。因为采集的降落标识数据集有限，改进后的算法与原算法的 map 值相差不大，在验证改进算法性能的优越性时，使用自制数据集并不太合适，所以采用 VOC2012 数据集中的部分数据作为数据集，将几种算法同时用在此数据集上，从而对比不同算法之间的性能。

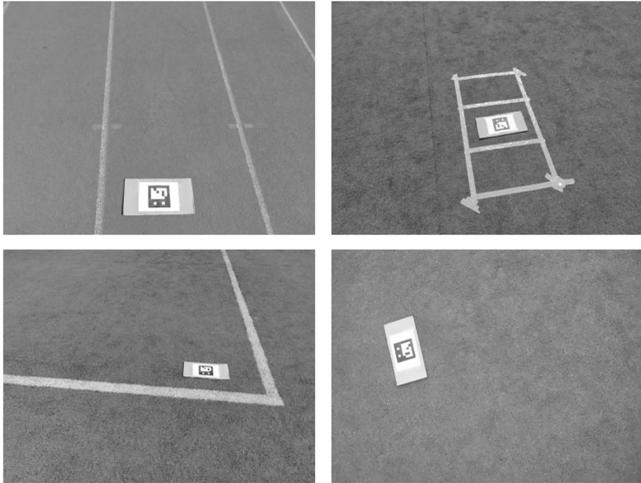


图 7 数据集中部分图片

3.2 实验环境与模型训练

实验环境如下：操作系统为 Ubuntu18.04，使用 GeForce GTX 1660 显卡进行运算，显存大小为 6 GB，CPU 配置为 Intel (R) Core (TM) i5-9400F @ 2.90 GHz，CUDA 的版本为 11.1.0，Python 的版本为 3.8.12，Pytorch 的版本为 1.8.0。本实验的一些参数设置如下：总迭代次数设置为 100 次，迭代批量大小设置为 16，优化器使用的是 SGD 算法。

3.3 评估指标

本文介绍 3 种常见的评估指标：平均精度 AP (average precision)、平均精度均值 mAP (mean AP) 和每秒检测图片的帧数 fps (frames per second)。其中，平均精度由精确率 (precision) 和查全率 (recall) 所决定。精确率是相对于预测结果而言的，它表示的是预测为正的样本中有多少是预测对的，所以用预测对的正样本数除以模型预测的正样本总数。查全率是相对于样本而言的，即样本中有多少正样本被预测正确，所以用预测对的正样本数除以实际样本总数。上述每种指标的计算公式如下：

$$AP = \int_0^1 PdR \quad (5)$$

$$mAP = \frac{\sum_{i=1}^N AP_i}{N} \quad (6)$$

$$precision = \frac{TP}{TP + FP} = \frac{TP}{all \ detections} \quad (7)$$

$$recall = \frac{TP}{TP + FN} = \frac{TP}{all \ ground \ truths} \quad (8)$$

式中，AP 是指精度-回归曲线 (Precision-Recall, P-R) 的面积。mAP 是通过对所有类别的 AP 进行平均得到的。N 表示检测到的类别总数。如果 mAP 值越大，那么算法的识别准确率就越高，检测效果就越好；TP 表示预测正确的正样本框的数量，FP 表示预测错误的正样本框的数量，FN 表示遗漏的样本框的数量。

3.4 改进后模型的性能分析

在 VOC2012 数据集上，改进后模型的 P-R 曲线如图 8 所示。改进后的模型性能与改进前的模型以及 YOLOv4 算法的对比如表 1 所示，改进后算法对所有类别的检测精确率、查全率、FPS、mAP@0.5 和 mAP@ [0.5; 0.95] 分别达到 77.8%，128，71.9% 和 48.2%。

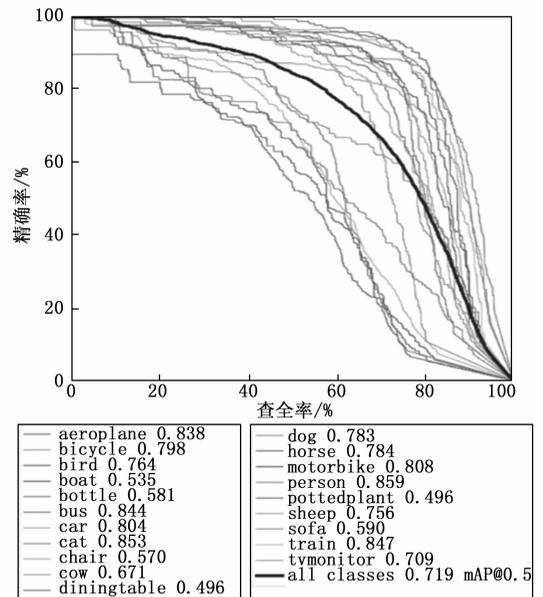


图 8 P-R 曲线

表 1 不同算法性能对比结果

模型	精确率 / %	查全率 / %	FPS	mAP@ 0.5 / %	mAP@[0.5; 0.95] / %
YOLOv4	77.4	66.6	41	72.5	48.9
YOLOv5s	76.5	65.0	142	71.5	47.3
改进 YOLOv5s	77.8	66.9	128	71.9	48.2

图 8 显示了 VOC 数据集中 20 个不同的类别以及所有类别的 P-R 曲线，图中以“aeroplane 0.838”为例，“aeroplane”表示类别名，“0.838”表示将 IoU 设为 0.5 时，该类别的 AP 值。由于 P-R 曲线只是宏观上展示模型性能，不同模型之间的性能差异主要通过 mAP 来衡量，且用 VOC 数据集仅是为了验证改进算法在性能上的优越性，并不用关注各个类别，只用关注所有类别的 P-R 曲线即可。

实验结果证明, 改进后的算法在 VOC2012 数据集上比原算法有更好的检测效果, 在精度与速度上的权衡是最好的, 虽然 FPS 略有降低但精度却得到了提升, 在无人机自主降落时, FPS 的略微降低的影响基本可以忽略不计, 但精度每提升一点, 无人机自主降落的安全性就会提升一点。由于 YOLOv4 的网络参数要多于 YOLOv5s, 所以 YOLOv4 算法在精度上会有一些优势, 但是 FPS 却较低, 即检测速度比较慢, 若将 YOLOv4 算法部署在无人机上的机载计算机系统中, 系统会缺乏实时性, 因此不适合应用于无人机上进行自主降落任务。

3.5 实飞实验

为了验证改进算法在无人机上的实际性能如何, 将改进算法应用到如图 9 所示的无人机上。采用的无人机的轴距为 410 mm, 使用碳纤维机身, 同时兼顾了机身的重量与续航。它还包含有 Pixhawk4 基础套件, 用来完成 PX4 最基本的飞行功能; 同时配备有单目视觉相机用于获取视觉信息, 作为改进算法的输入; 机载计算机采用 NVIDIA TX2, 它作为无人机自主降落任务的核心用来运行改进的目标检测算法来处理单目相机获取到的原始数据, 并通过相关算法计算出无人机与降落标识物之间的相对位姿关系, 最后通过 Mavros 将此信息发送至飞控之中, 从而引导无人机安全的自主降落到降落标识物上。



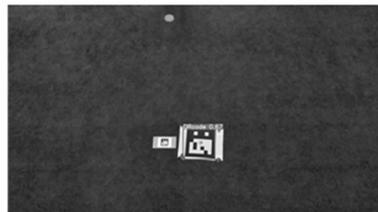
图 9 实验所用无人机

在无人机上的实际检测效果如图 10 所示。

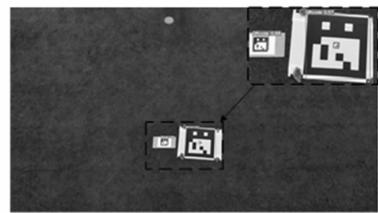
实验结果证明, 当无人机的高度越来越高时, 如果视野中同时存在一个大目标和小目标, 原 YOLOv5s 算法只能检测到大目标却不能检测到小目标, 而改进后的算法能同时检测到两种大小的目标, 证明了改进后的算法在二维码数据集中对小目标有更好的检测效果。同时改进后的算法在 TX2 上的 FPS 约为 32, 也满足嵌入式部署的实时性要求。

4 结束语

改进 YOLOv5s 算法首先通过增加小目标检测层来提升对小目标特征的检测能力; 然后采用 BiFPN 加权双向特征金字塔网络代替原先 PANet 结构, 快速进行多尺度特征融合; 最后损失函数使用 EIoU Loss, 使收敛更快。实验表明, 改进算法相较于原算法有更好的小目标检测能力, 能



(a) 原算法



(b) 改进算法

图 10 降落标识检测效果图

够获得较好的检测准确度和比较快的检测速度, 可以满足无人机自主降落的准确性和实时性的需求。由于在实际应用中场景是多种多样的, 本文选取的降落标识可能在某些背景中不是特别明显, 后续工作可以选取不同的降落标识物, 并采集更多的数据集, 通过实验结果来选择合适的降落标识物, 且可以在提升检测的实时性上作进一步研究。

参考文献:

- [1] 甄子洋. 舰载无人机自主着舰回收制导与控制研究进展 [J]. 自动化学报, 2019, 45 (4): 669-681.
- [2] WANG C, LI W, CHEN W, et al. Autonomous landing control of quadrotor UAVs [C] // 2015 IEEE International Conference on Cyber Technology in Automation, Control, and Intelligent Systems (CYBER). IEEE, 2015: 276-280.
- [3] 周城宇, 甄子洋, 黄一敏, 等. 基于视觉引导的无人直升机着舰技术研究 [J]. 导航定位与授时, 2018, 5 (4): 43-48.
- [4] 卫文乐, 芦利斌, 金国栋, 等. 基于单目 SLAM 的无人机视觉导航研究综述 [J]. 飞航导弹, 2020 (7): 17-22.
- [5] LE C Y, BOTTOU L, BENGIO Y, et al. Gradient-based learning applied to document recognition [J]. Proceedings of the IEEE, 1998, 86 (11): 2278-2324.
- [6] 侯志斌, 朱愿, 娄静涛. 面向自动驾驶的 3D 目标检测技术研究综述 [J]. 军事交通学报, 2022, 1 (8): 78-84.
- [7] 景晨凯, 宋涛, 庄雷, 等. 基于深度卷积神经网络的人脸识别技术综述 [J]. 计算机应用与软件, 2018, 35 (1): 223-231.
- [8] 糜泽花, 钱爱兵. 智慧医疗发展现状及趋势研究文献综述 [J]. 中国全科医学, 2019, 22 (3): 366-370.
- [9] LIU Y, MA Z, LIU X, et al. Privacy-preserving object detection for medical images with faster R-CNN [J]. IEEE Transactions on Information Forensics and Security, 2019 (99): 1.
- [10] 陈秋霞. 机器视觉技术在工业检测领域中的应用 [J]. 设备管理与维修, 2018 (16): 140-142.

- [11] RAGHUNANDAN A, RAGHAV P, ARADHYA H V R, et al. Object detection algorithms for video surveillance applications [C] //2018 International Conference on Communication and Signal Processing (ICCSP), IEEE, 2018: 0563 - 0568.
- [12] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation [C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. [S.l.]: IEEE, 2014: 580 - 587.
- [13] HE K, ZHANG X, REN S, et al. Spatial pyramid pooling in deep convolutional networks for visual recognition [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 37 (9): 1904 - 1916.
- [14] GIRSHICK R. Fast R-CNN [C] // Proceedings of the IEEE International Conference on Computer Vision. [S.l.]: IEEE, 2015: 1440 - 1448.
- [15] REN S, HE K, 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, 2016, 39 (6): 1137 - 1149.
- [16] LIU W, ANGUELOV D, ERHAN D, et al. SSD: Single shot multibox detector [C] //European Conference on Computer Vision. Cham: Springer, 2016: 21 - 37.
- [17] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: unified, real-time object detection [J]. CoRR, (上接第 79 页)
- [4] ULAPANE N, NGUYEN L. Review of pulsed-eddy-current signal feature-extraction methods for conductive ferromagnetic material-thickness quantification [J]. Electronics, 2019, 8: 470.
- [5] SZIELASKO K, MIRONENKO I, ALTPETER I, et al. Minimalistic devices and sensors for micromagnetic materials characterization [J]. IEEE Trans. Magn., 2013, 49 (1): 101 - 104.
- [6] LI K, QIU P C, WANG P, et al. Estimation method of yield strength of ferromagnetic materials based on pulsed eddy current testing [J]. J. Magn. Magn. Mater., 2021, 523: 167647.
- [7] LI K, LI L, WANG P, et al. A fast and non-destructive method to evaluate yield strength of coll-rolled steel via incremental permeability [J]. J. Magn. Magn. Mater., 2020, 498: 166087.
- [8] ZHANG Y, LIU W, LI K, et al. Application of a back-propagation neural network for mechanical properties prediction of ferromagnetic materials by magnetic Barkhausen noise technique [J]. Insight-Non-Destructive Testing and Condition Monitoring, 2019, 61 (2): 95 - 99.
- [9] ZHU B, XU Z, WANG K, et al. Nondestructive evaluation of hot stamping boron steel with martensite/bainite mixed microstructures based on magnetic Barkhausen noise detection [J]. J. Magn. Magn. Mater., 2020, 503: 166598.
- [10] 王礼立, 胡时胜, 杨黎明, 等. 材料动力学 [M]. 合肥: 中国
- 2015, abs/1506.02640.
- [18] REDMON J, FARHADI A. YOLO9000: Better, faster, stronger [C] //Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. [S.l.]: IEEE, 2017: 7263 - 7271.
- [19] REDMON J, FARHADI A. Yolov3: An incremental improvement [J]. arXiv preprint arXiv: 1804.02767, 2018.
- [20] BOCHKOVSKIY A, WANG C Y, LIAO H Y M. Yolov4: Optimal speed and accuracy of object detection [J]. ArXiv Preprint ArXiv: 2004.10934, 2020.
- [21] ZHENG Z H, WANG P, LIU W, et al. Distance-IoU Loss: Faster and Better Learning for Bounding Box Regression [J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2020, 34 (7): 12993 - 13000.
- [22] 李科岑, 王晓强, 林浩, 等. 深度学习中的单阶段小目标检测方法综述 [J]. 计算机科学与探索, 2022, 16 (1): 41 - 58.
- [23] LIN T Y, DOLLAR P, GIRSHICK R, et al. Feature pyramid networks for object detection [C] //Proceedings of the IEEE conference on computer vision and pattern recognition. 2017: 2117 - 2125.
- [24] WANG W, XIE E, SONG X, et al. Efficient and accurate arbitrary-shaped text detection with pixel aggregation network [C] //Proceedings of the IEEE/CVF International Conference on Computer Vision, 2019: 8440 - 8449.
- 科学技术大学出版社, 2017.
- [11] WILLIAM D. CALLISTER J. Fundamentals of materials Science and Engineering [J]. Wiley John + Sons, 2001, 34 (6): 824.
- [12] 钟群鹏, 赵子华. 断口学 [M]. 北京: 高等教育出版社, 2006.
- [13] 任威平. 电磁超声在钢板中的换能机理研究及应用 [D]. 北京: 北京科技大学, 2019.
- [14] ADLER E, PFEIFFER H. The influence of grain size and impurities on the magnetic properties of the soft magnetic alloy 47.5% NiFe [J]. IEEE Transactions on Magnetics, 1974, 10 (2): 172 - 174.
- [15] 王登泉, 杨明. 基于逆磁致伸缩效应的超声电机非接触扭矩测量 [J]. 传感器与微系统, 2011, 30 (9): 135 - 138.
- [16] 周子钦. 基于电磁超声的铁磁性材料屈服强度检测方法研究 [D]. 南京: 南京航空航天大学, 2020.
- [17] DOBMANN G. Non-destructive testing for ageing management of nuclear power components [J]. Nuclear Power-Control, Reliability and Human Factors, 2011, 34 (17): 311 - 338.
- [18] 黄松岭, 王坤, 赵伟, 等. 电磁超声导波理论与应用 [M]. 北京: 清华大学出版社, 2013.
- [19] 杜裕平, 冯雪, 盛宏威, 等. 基于多电磁无损检测的管线钢硬度检测模型 [J]. 理化检验 - 物理分册, 2022, 58 (12): 27 - 30, 72.
- [20] 李双池. 基于 RBF 神经网络的相贯节点网架结构损伤识别研究 [D]. 贵阳: 贵州大学, 2022.