

doi: 10.7690/bgzdh.2022.12.012

无人机空中回收视觉导航技术

闫留浩, 袁锁中

(南京航空航天大学自动化学院先进飞行器导航、控制与健康管理部重点实验室, 南京 211106)

摘要: 针对无人机空中回收过程中的导航问题, 提出一种利用深度学习进行目标检测并配合双目视觉进行位姿估计的技术。设计空中回收视觉导航系统, 通过改进原有目标检测算法 YOLOv3 框架提高回收过程中的检测精度和速度; 通过双目视觉系统对特征点进行 3 维位姿解算, 返回无人机和回收锥套中心相对位置信息。实验结果表明: 改进后的检测算法平均精度比 YOLOv3 提高了 3.2%, 检测速度提高到 73 FPS, 检测速度明显提升; 双目视觉算法的位姿解算精确度高, 两者同时满足导航系统精确性和实时性的要求。

关键词: 空中回收; YOLOv3; 双目视觉; 位姿解算

中图分类号: V279 **文献标志码:** A

Vision Navigation Technology for UAV Aerial Recovery

Yan Liuhaohao, Yuan Suozhong

(Key Laboratory of Navigation, Control and Health Management Technologies of Advanced Aircraft, Ministry of Industry and Information Technology, College of Automation, Nanjing University of Aeronautics and Astronautics, Nanjing 211106, China)

Abstract: Aiming at the navigation problem of unmanned aerial vehicle (UAV) in the process of aerial recovery, a technology of target detection based on deep learning and pose estimation based on binocular vision is proposed. A visual navigation system for aerial recovery is designed, which improves the detection accuracy and speed in the recovery process by improving the original target detection algorithm YOLOv3 framework. The 3D pose of the feature points is calculated by the binocular vision system, and the relative position information of the UAV and the recovery drogue center is returned. The experimental results show that the average accuracy of the improved algorithm is 3.2% higher than that of YOLOv3, and the detection speed is increased to 73 FPS, which shows that the detection speed is significantly improved. The pose calculation accuracy of the binocular vision algorithm is high, and both of them meet the requirements of accuracy and real-time of navigation system.

Keywords: aerial recovery; YOLOv3; binocular vision; pose calculation

0 引言

随着物联网、机器学习的快速发展, 小型无人机的空中回收被视为未来分布式空中作战概念中的重要突破口^[1-2], 可使“空中母舰”具备真实的重复作战能力, 也可改进当前以歼击机、轰炸机等大型多功能平台为核心的作战模式, 给敌人防御带来重大挑战, 将是大幅降低作战成本的重要途径。空中回收目前处于前期论证阶段^[3], 利用视觉导航回收有动力的无人机是技术难点。

传统视觉导航算法例如基于光学标志的导航系统和基于视觉传感器的导航系统^[4]均可以获得较好的精度, 但是很难突破系统计算速度、鲁棒性的瓶颈。近年来深度学习在目标检测方面展现了很高的鲁棒性和实时性。针对空中回收过程中视觉导航问题, 笔者提出利用深度学习进行目标检测并配合双

目视觉进行位姿结算的技术, 包括改进 YOLOv3 目标检测算法和基于双目视觉算法的 3 维位姿参数估计, 最后利用特征点 3 维坐标计算出回收锥套位置和姿态, 提升了视觉导航过程中目标检测的快速性和准确性以及相对位置信息的准确性。

1 空中回收导航系统设计

1.1 视觉导航方案设计

笔者受到飞机拖曳系统的启发, 设计末端安装有回收锥套的母机^[5], 如图 1 所示, 母机飞行轨迹处在半径为 R 的等角速度圆形轨道上, 速度保持恒定, 回收锥套具有足够的空气阻力, 则锥套的飞行轨迹为半径 r 的稳定圆形轨道。由于母机和回收锥套的角速率相同, 锥套的速度将明显小于母机的速度, 便于无人机以稍大于回收锥套的速度靠近。

收稿日期: 2022-08-24; 修回日期: 2022-09-28

基金项目: 国家自然科学基金(61273050); 中央高校基本科研业务费资助(XCA18155)

作者简介: 闫留浩(1996—), 男, 河南人, 硕士, 从事飞行制导与控制、图像处理与识别、空中回收研究。E-mail: 18115123932@163.com。

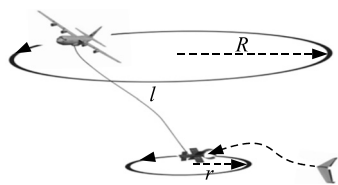


图 1 空中回收系统

在回收过程中，精准快速地识别到锥套是完成对接的前提，也是最关键的一步，笔者建立的视觉导航系统如图 2 所示。首先利用深度学习对回收锥套进行目标检测，当距离较近时切换到双目视觉系统，由机载双目摄像头来获取回收场景中锥套图像，视觉导航计算模块对所捕获的图像进行图像预处理、特征点提取和特征点匹配，最后返回无人机相对锥套的位置信息。

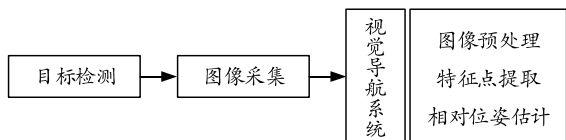


图 2 视觉导航系统总体结构

1.2 目标单步检测算法

传统的目标检测算法中大部分的流程是先进行区域选择，找到目标所在的候选区域，然后对候选区域进行特征提取，最后使用分类器分类。整个检测过程不仅会导致检测速度下降，另一方面候选框存在重叠的问题，增加算法处理的复杂性。为提升算法性能，笔者采用 YOLOv3 单步检测算法，其基本框架基于 Darknet-53。区别于两步检测器，如图 3 所示，单步检测器不再使用搜索遍历去寻找候选区域，基于直接回归就可确定多个位置所覆盖的目标进行类别预测，提高了检测模型快速性的同时避免了算法重复处理的复杂性^[6]。在检测回收锥套的应用场景中，需要增强算法对小目标的检测，所以考虑对 YOLOv3 网络进行重新聚类分析和多尺度训练在提高检测速度的同时提高对小目标的检测能力。单阶段检测器无需级联区域分类这一步骤，直接对特征图上每个位置所覆盖的目标进行类别预测。

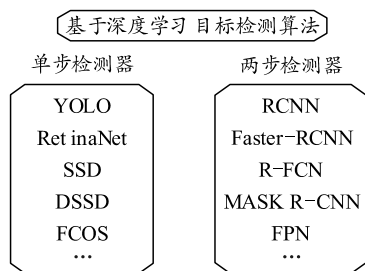


图 3 目标检测算法分类

2 改进 YOLO 算法检测模型

2.1 K 均值聚类优化锚框参数

单阶段检测算法原型 YOLOv3 虽然在检测速度上比 Fast-RCNN 快，但并不适用于空中小目标检测问题，鉴于本任务特点，对 YOLOv3 网络进行重新聚类分析，改进 YOLOv3 卷积神经网络特征结构，在提高检测速度的同时提高对小目标的检测能力。

YOLO 算法借鉴了 Faster-RCNN 算法在训练过程中使用锚框预测的做法，锚节点框的选取对于检测精度和速度有很大影响。但是由于 Faster-RCNN 需要手动设置锚节点框的尺寸大小和个数，而 YOLO 算法可对训练集使用聚类算法 K 均值聚类算法 (K-means) 对锚节点框进行自动迭代计算直到得到最优锚节点框^[7]，所以 K 均值聚类算法的实现使得网络更快地学习目标特征并提取。YOLOv3 对目标框使用 K 均值聚类得到 9 个基于 YOLOv3 的网络会在网络单元上预测出 3 个 anchor box^[8]，每个锚盒预测 3 个边界框，每个边界框预测出 4 个坐标值，如图 4 所示，偏移后的修正公式为：

$$\begin{aligned} b_x &= \sigma(t_x) + c_x; & b_y &= \sigma(t_y) + c_y; \\ b_w &= p_w e^{t_w}; & b_h &= p_h e^{t_h}. \end{aligned} \quad (1)$$

式中， p_w 和 p_h 分别代表锚节点框的宽、高。

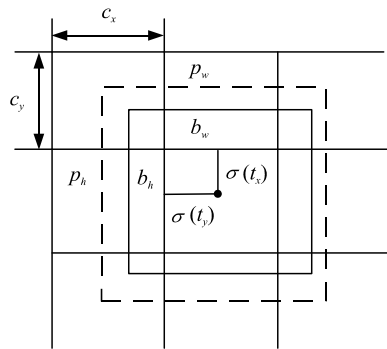


图 4 锚节点框修正

由于 YOLOv3 的锚节点框尺寸大小与个数是根据 COCO 数据集制作的，不适用空中小目标检测，所以首先对航空数据集进行重聚类分析。根据式(2)对锚节点框进行自动迭代计算直到得到最优锚节点框然后重新确定其尺寸大小与个数，使得其更容易学习到目标特征。

$$d(\text{box}, \text{centroid}) = 1 - \text{IOU}(\text{box}, \text{centroid}). \quad (2)$$

式中：box 为数据集的 anchor 框尺寸样本；centroid 为每一类簇的中心尺寸数值。YOLOv3 根据 COCO 数据集制作了 9 组锚节点框，尺寸分别是(10, 13)，(16, 30)，(33, 23)，(30, 61)，(62, 45)，(59, 119)，(116,

90), (156, 198), (373, 326)。利用 YOLOv3 对数据集锚节点框重新进行聚类分析后, 选择 $k=6$, 得到 6 组锚节点框尺寸(14, 10), (16, 15), (22, 20), (25, 16), (33, 19), (38, 33)。

2.2 改进 YOLOv3 的网络结构

检测模型的主干网络是 Darknet-53^[9], 为提升对空中小目标的检测能力, 笔者针对 YOLOv3 网络结构进行了改进。YOLOv3 网络模型的输入图像分辨率默认为 416×416, 最后经过上采样和一系列卷积操作后输出了 3 个特征层, 分别是最小尺度的 13×13×75、中等尺度的 26×26×75 和大尺度的 52×52×75。为增强网络模型对小目标的检测能力, 基于 YOLOv3 网络模型将输入图像分辨率设为 640×640, 包含了更多目标特征信息, 最后输出只保留 2 个尺寸较大的特征层, 使用得到的 2 个特征层对预测框进行回归分类, 增强模型对小目标特征敏感性的同时提升检测速度。改进后的网络结构如图 5 所示。

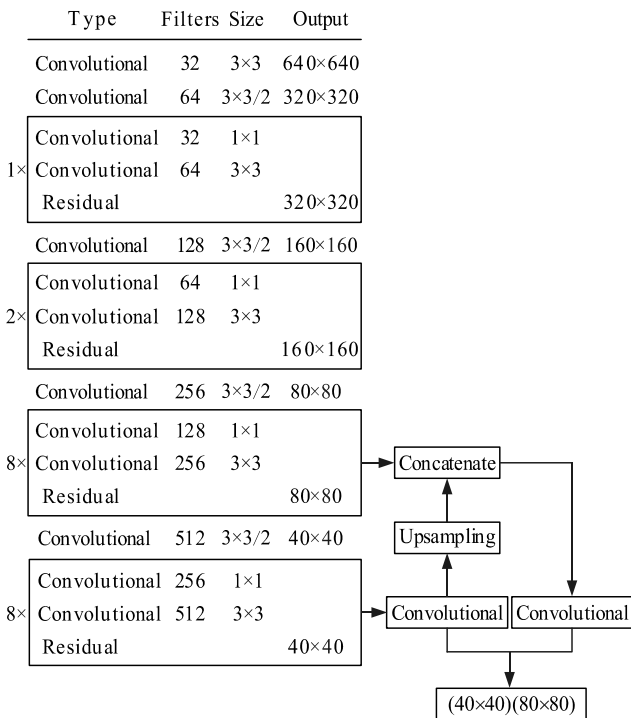


图 5 改进 YOLOv3 的网络结构

网络模型结构原始输入图像经过多次卷积和 4 次下采样后, 最小尺寸的特征图分辨率为 40×40, 为第一尺度特征图。然后对这个特征图进行卷积, 再与一个 80×80 的特征图进行融合, 形成一个新的特征图作为二阶特征。网络修改后, 使用 2 个更大尺寸的特征图进行目标检测, 使小目标在预测框中能够保留更多的信息, 增强小目标的检测能力。

3 基于双目视觉的导航方法

3.1 图像预处理与特征点提取

首先双目摄像头拍摄到图像以后, 5 个标记点采用 led 灯形式, 基于区域生长法查找图像连通区域质心作为特征点, 具体步骤如图 6 所示, 按扫描的顺序对所有标记出的连通区域用阿拉伯数字进行标识。最后将提取特征点以一定顺序的矩阵形式保存。由于左右特征点皆按相同的顺序保存, 该算法只需通过一次遍历就能找到标记区域并以每个标记点的质心坐标作为特征点, 为特征点匹配工作带来很大便利, 明显节省特征点提取的时间。

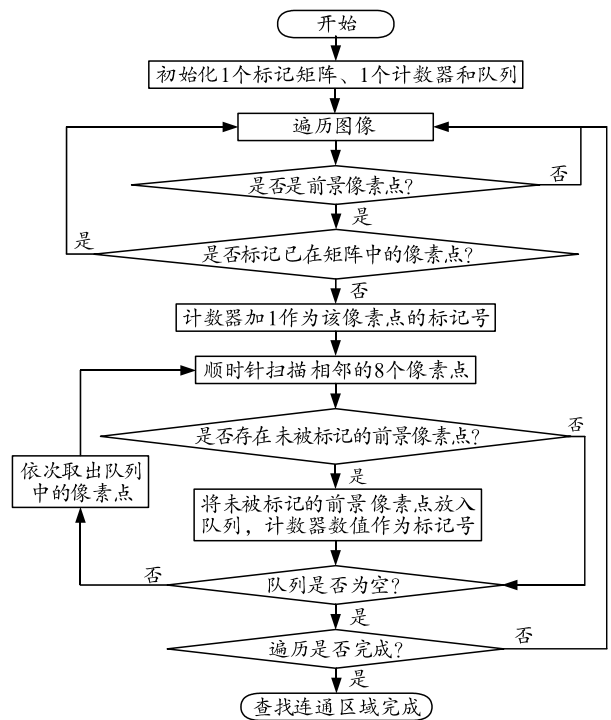


图 6 基于区域生长法查找连通区域流程

将白色光斑区域的中心坐标定义为近红外 led 特征点的坐标。采用灰度重心法^[10]对图像进行特征点提取, 通过得到的标记矩阵提取每个像素点的坐标值, 再利用以下几何关系求取区域质心坐标:

$$X_z = \sum p_i x_i / \sum p_i, Y_z = \sum p_i y_i / \sum p_i \quad (3)$$

式中: p_i 为第 i 个像素点的像素值; x_i, y_i 为第 i 个像素点坐标; X_z, Y_z 为连通区域质心的像素坐标。

3.2 特征点匹配与 3 维位姿解算

Surf 算法采用的是 haar 小波特征判断, 即通过一个 60°扇形区域选取圆形领域内 haar 小波特征点最多的方向作为特征点的主方向, 生成的特征点描述子比 Sift 算法少了一倍, 极大提高了效率。利用深度学习进行目标检测后得到的图像像素坐标系

下的位置信息后，对其进行 3 维位姿解算，使 2 维像素坐标转换成实际 3 维坐标。

设 $P_h=[X_h, Y_h, Z_h]^T$ 为空间一点 P 在无人机坐标系下的坐标，通过图像匹配可以确定点 P 在左右 2 个相机成像平面的投素坐标分别为 $[u_l, v_l]$, $[u_r, v_r]$ 。

$$w_l \begin{bmatrix} u_l \\ v_l \\ 1 \end{bmatrix} = M_l \begin{bmatrix} X_h \\ Y_h \\ Z_h \\ 1 \end{bmatrix} = \begin{bmatrix} n_{l11} & n_{l12} & n_{l13} & n_{l14} \\ n_{l21} & n_{l22} & n_{l23} & n_{l24} \\ n_{l31} & n_{l32} & n_{l33} & n_{l34} \end{bmatrix} \begin{bmatrix} X_h \\ Y_h \\ Z_h \\ 1 \end{bmatrix}; \quad (4)$$

$$w_r \begin{bmatrix} u_r \\ v_r \\ 1 \end{bmatrix} = M_r \begin{bmatrix} X_h \\ Y_h \\ Z_h \\ 1 \end{bmatrix} = \begin{bmatrix} n_{r11} & n_{r12} & n_{r13} & n_{r14} \\ n_{r21} & n_{r22} & n_{r23} & n_{r24} \\ n_{r31} & n_{r32} & n_{r33} & n_{r34} \end{bmatrix} \begin{bmatrix} X_h \\ Y_h \\ Z_h \\ 1 \end{bmatrix}。 \quad (5)$$

上式中首先分别消去 w_l 和 w_r ，通过计算得到关于 $P_h=[X_h, Y_h, Z_h]^T$ 的方程组：

$$\left. \begin{aligned} (u_l n_{l31} - n_{l11})X_h + (u_l n_{l32} - n_{l12})Y_h + \\ (u_l n_{l33} - n_{l13})Z_h = n_{l14} - u_l n_{l34} \\ (u_l n_{l31} - n_{l21})X_h + (u_l n_{l32} - n_{l22})Y_h + \\ (u_l n_{l33} - n_{l23})Z_h = n_{l24} - u_l n_{l34} \end{aligned} \right\}; \quad (6)$$

$$\left. \begin{aligned} (u_r n_{r31} - n_{r11})X_h + (u_r n_{r32} - n_{r12})Y_h + \\ (u_r n_{r33} - n_{r13})Z_h = n_{r14} - u_r n_{r34} \\ (u_r n_{r31} - n_{r21})X_h + (u_r n_{r32} - n_{r22})Y_h + \\ (u_r n_{r33} - n_{r23})Z_h = n_{r24} - u_r n_{r34} \end{aligned} \right\}。 \quad (7)$$

为求取空间点 P 在 3 维坐标系下的坐标，可以将上式进行联立求解得到点 P 3 维坐标，求解过程转换为矩阵形式：

$$A \begin{bmatrix} X_h \\ Y_h \\ 1 \end{bmatrix} = b。 \quad (8)$$

式中：

$$A = \begin{bmatrix} u_l n_{l31} - n_{l11} & u_l n_{l32} - n_{l12} & u_l n_{l33} - n_{l13} \\ v_l n_{l31} - n_{l21} & v_l n_{l32} - n_{l22} & v_l n_{l33} - n_{l23} \\ u_r n_{r31} - n_{r11} & u_r n_{r32} - n_{r12} & u_r n_{r33} - n_{r13} \\ v_r n_{r31} - n_{r21} & v_r n_{r32} - n_{r22} & v_r n_{r33} - n_{r23} \end{bmatrix}; \quad (9)$$

$$b = \begin{bmatrix} n_{l14} - u_l n_{l34} \\ n_{l24} - u_l n_{l34} \\ n_{r14} - u_r n_{r34} \\ n_{r24} - u_r n_{r34} \end{bmatrix}。 \quad (10)$$

$P_h=[X_h, Y_h, Z_h]^T$ 为空间一点 P 在无人机坐标系下的 3 维坐标，考虑到回收环境中噪声影响，利用最小二乘法简化式可得：

$$\begin{bmatrix} X_h \\ Y_h \\ Z_h \end{bmatrix} = (A^T * A)^{-1} * A^T * b。 \quad (11)$$

通过双目视觉算法计算得到图像像素坐标系下的像素位置信息和世界坐标系下的特征点 3 维坐标信息后，进行空间圆的拟合获取回收锥套和无人机的相对位置信息，即回收锥套平面底部的圆心坐标及法向量，拟合方法采用最小二乘法^[11]。锥套圆形端面可以当作是圆球和空间平面的集合，由圆球几何关系可知：

$$(x_h - u)^2 + (y_h - v)^2 + (z_h - w)^2 = R^2。 \quad (12)$$

分解上式圆球的 3 维方程式可得：

$$\begin{aligned} x_h^2 + y_h^2 + z_h^2 - 2ux_h - 2vy_h - \\ 2wz_h + u^2 + v^2 + w^2 = R^2。 \end{aligned} \quad (13)$$

由空间平面几何关系可得：

$$Ax_h + By_h + Cz_h - D = 0。 \quad (14)$$

矩阵形式可以表示为：

$$\begin{bmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ \vdots & \vdots & \vdots \\ x_n & y_n & z_n \end{bmatrix} \begin{bmatrix} A \\ B \\ C \end{bmatrix} - \begin{bmatrix} D \\ D \\ \vdots \\ D \end{bmatrix} = 0。 \quad (15)$$

4 仿真结果分析

4.1 目标检测算法结果分析

首先将标记好的航空目标数据集制作成 VOC 格式，训练集与测试集比例为 3:1，导入到改进后的 YOLO 算法模型中，在训练阶段，为了防止过拟合，采用图像批处理量大小为 64，网络参数动量为 0.9，初始学习率设置为 0.01，所有实验的最大迭代次数为 40 000，同时将训练参数变化以 log 文件形式保存。为进一步验证上述算法的有效性，将改进算法与未经过任何改进的 YOLOv3 算法进行对比实验，对比指标包括每秒帧率 (FPS) 和平均精度均值 (mAP)。实验结果如图 7 所示。



图 7 空中目标检测对比图

通过对航空目标数据集重聚类，获得适用于小

目标检测的锚框大小,同时对 Darknet-53 结构进行改进,有效提高了检测精度。如表 1 所示,改进后的模型平均精度比 YOLOv3 提高了 3.2%,检测速度明显提升,满足实时性(大于 50 FPS)的要求。

表 1 实验结果对比

算法模型	Map/%	速度/FPS	AP/%
YOLOv3	84.68	60	99.6
YOLOv3+	87.46	73	99.7

4.2 双目视觉算法结果分析

双目视觉系统基于 vs2015+opencv 环境搭建,由电脑控制左、右摄像头的采集、保存、图像处理等操作,双目相机获取的 8 组原始图像如图 8 所示。

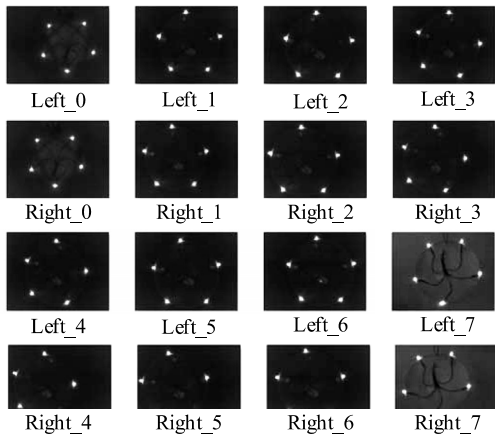


图 8 双目摄像头获取的原始图像

经过图像预处理后,采用八连通方式的像素扫描法按照由上到下、由左到右的方式对图像逐行逐列扫描,共需要 2 次扫描完成标记。之后利用灰度重心法对标记出的区域进行特征点提取,查找到连通区域后得到的标记结果如图 9 (a)所示,特征点提取结果如图 9 (b)所示,图像匹配结果如图 10 所示。

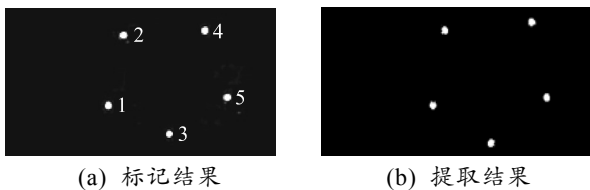


图 9 特征点标记结果与提取结果

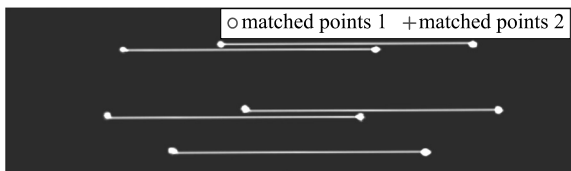


图 10 图像匹配结果

通过双目视觉算法及 3 维位姿估计算法得到一组特征点图像的 3 维坐标如表 2 所示。最后通过空间几何关系计算出锥套底面圆圆心的位姿信息才算

达到最终目的。记圆心空间坐标为 $O(X, Y, Z)$, 圆心姿态信息用切面的法向量 \vec{n} 表示,拟合方法采用最小二乘法可以求出球心处的法向量: $O(X, Y, Z)=(27.5\ 689, -3.2\ 964, 348.2\ 037)$, 其法向量 $\vec{n} = (-0.0\ 148, -0.0\ 581, 0.9\ 782)$, 锥套模型半径约为 107 mm。由此得到锥套的全部位姿信息,完成 3 维位姿解算。

表 2 1 组图像光点 3 维位置

光点	X 轴坐标	Y 轴坐标	Z 轴坐标
1	-18.072 5	93.569 6	353.313 2
2	95.893 1	79.119 6	354.105 7
3	127.220 9	-42.778 0	347.414 7
4	27.566 5	-110.314 6	341.831 8
5	-78.767 5	-16.614 0	345.810 2

5 结论

实验结果证明:改进后的检测算法平均精度比 YOLOv3 提高了 3.2%,检测速度提高到 73 FPS,提升了视觉导航过程中目标检测的快速性和准确性以及相对位置信息的准确性。

参考文献:

- [1] 郑晨明, 张民. 无人机协同目标跟踪分布式仿真[J]. 兵工自动化, 2020, 39(5): 45-48.
- [2] 宗群, 王丹丹, 邵士凯, 等. 多无人机协同编队飞行控制研究现状及发展[J]. 哈尔滨工业大学学报, 2017, 49(3): 1-14.
- [3] 黄雷. 美军小精灵无人机群项目发展现状综述[J]. 飞行导弹, 2018(7): 44-47.
- [4] 李飞, 谭乐祖, 唐亮. 基于跑道灯光的计算机视觉辅助无人机着陆[J]. 兵工自动化, 2012, 31(1): 11-13, 36.
- [5] 刘宏本, 姚开楦. 空中拖曳收放机构的设计[J]. 南京航空航天大学学报, 1960(1): 1-20.
- [6] 丁鹏程, 于进勇, 王超. 目标检测网络在无人机视觉定位的应用[J]. 兵工自动化, 2019, 38(7): 14-20, 29.
- [7] 张素洁, 赵怀慈. 最优聚类个数和初始聚类中心点选取算法研究[J]. 计算机应用研究, 2017, 34(6): 1617-1620.
- [8] 叶涛, 赵宗扬, 柴兴华, 等. 基于改进卷积神经网络的非合作无人机检测应用[J]. 科学技术与工程, 2021, 21(33): 14245-14250.
- [9] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. IEEE, 2016: 770-778.
- [10] 李柱. 基于双目视觉的无人机自主空中加油对接导航方法[D]. 厦门: 厦门大学, 2017.
- [11] 陈冠宇. 无人机空中加油对接引导技术[D]. 南京: 南京航空航天大学, 2019.