

基于深度学习的多旋翼无人机单目视觉 目标定位追踪方法

魏明鑫, 黄浩, 胡永明, 王德志, 李岳彬

(湖北大学 物理与电子科学学院, 武汉 430062)

摘要: 针对无人机对目标的识别定位与跟踪, 提出了一种基于深度学习的多旋翼无人机单目视觉目标识别跟踪方法, 解决了传统的基于双目摄像机成本过高以及在复杂环境下识别准确率较低的问题; 该方法基于深度学习卷积神经网络的目标检测算法, 使用该算法对目标进行模型训练, 将训练好的模型加载到搭载 ROS 的机载电脑; 机载电脑外接单目摄像机, 单目摄像头检测目标后, 自动检测出目标在图像中的位置, 通过采用一种基于坐标求差的优化算法进行目标位置准确获取, 然后将目标位置信息转化为控制无人机飞行的期望速度和高度发送给飞控板, 飞控板接收到机载电脑发送的跟踪指令, 实现对目标物体的跟踪; 试验结果验证了该方法可以很好地进行目标识别并实现目标追踪。

关键词: 计算机视觉; 深度学习; 无人机; 单目摄像机; 目标跟踪

Monocular Vision Target Tracking Method for Multi-rotor Unmanned Aerial Vehicle Based on Deep Learning

Wei Mingxin, Huang Hao, Hu Yongming, Wang Dezhi, Li Yuebin

(Hubei Key Lab of Ferro- & Piezoelectric Materials and Devices, Faculty of Physics and Electronic Science, Hubei University, Wuhan 430062, China)

Abstract: Aiming at the target recognition, location and tracking of unmanned aerial vehicles (UAV), a multi-rotor UAV monocular vision target recognition and tracking method based on deep learning is proposed, which solves the problems of high cost of traditional binocular camera and low recognition accuracy in complex environment. This method is based on the target detection algorithm of deep learning convolutional neural network, which is used to conduct target model training and load the trained model into the onboard computer equipped with ROS. Onboard computer external monocular camera, monocular camera detecting target, the automatically detect the target in the image position, by adopting a kind of optimization algorithm based on coordinate for poor get target location accurate, then the target position information into control of UAV flight speed and height expectation for flight control board, flight control board accepting follows commands sent to the airborne computer, realize the target object tracking. Experimental results show that this method can recognize and track targets well.

Keywords: computer vision; deep learning; unmanned aerial vehicles (UAV); monocular camera; target tracking

0 引言

无人机 (unmanned aerial vehicle, UAV) 主要分为旋翼无人机和固定翼无人机。多旋翼无人机是旋翼无人机的一种, 具有灵活性好、稳定性强、可垂直起降等特点, 比固定翼无人机机动性更强。因此, 多旋翼无人机的应用场景更加广泛。随着多旋翼无人机技术近几年的发展, 多旋翼无人机在民用和军用领域都有了很广的应用。民用上可以用在航拍、电力巡检、物流等领域; 军事上更是可以用在军事侦察、目标打击等方面。

收稿日期: 2019-08-19; 修回日期: 2019-08-30。

基金项目: 湖北省自然科学基金指导性计划项目 (2018CFC797)。

作者简介: 魏明鑫 (1996-), 男, 湖北武汉人, 硕士研究生, 主要从事计算机视觉, 机器人交互技术方向的研究。

胡永明 (1978-), 男, 江西资溪县人, 博士, 博士生导师, 教授, 主要从事光电传感器方向的研究。

传统视觉跟踪虽然有很强的自主性、宽测量范围和可以获取大量的环境信息等优点, 但它同时也存在许多缺点。

1) 强大的硬件系统是必需的: 为了获得精确的导航信息, 导航系统不可避免地配备了高分辨率相机和强大的处理器。从图像数据采集和处理, 这涉及到巨大的数据操作。许多图像算法都是非常复杂的, 这无疑给系统的集成设计、CPU、DSP、GPU 计算能力和内存大小带来了巨大的挑战。

2) 传统视觉导航跟踪的可靠性较差: 视觉导航有时很难在复杂的灯光和温度变化的环境下工作。环境适应性问题困扰视觉导航的难题。同时无法解决不同形状的物体在投影平面上的问题相同的图像视觉算法的高实时需求。

所以视觉导航不仅需要高性能的硬件来提高图像采集的速度和处理时间, 还需要在深度学习、神经网络和小波变换等软件算法上取得突破^[1-6]。

近几年, 由于深度学习的发展, 机器人无人驾驶等领

域又迎来了一次变革, 将朝着达到真正意义上的人工智能发展。目标跟踪无人机主要涉及视觉识别、目标检测与追踪等问题。文献 [7] 介绍了一种基于合作目标的无人机目标跟踪方法。能够实现准确跟踪, 此方法主要用在危险物排除。文献 [8] 介绍了基于视觉的无人机目标追踪方法, 该方法是采用数传将摄像机采集到的图像数据发送到远端进行处理后再发送指令给无人机。

本文介绍一种基于深度学习的无人机目标跟踪方法, 其自主性更强, 解决了传统的基于双目摄像机成本过高以及在复杂环境下识别准确率较低的问题。通过深度学习 SSD 算法训练自己的模型, 然后将训练好的模型移到嵌入式工控板并运行, 使其能够辨认出要识别并且跟踪的目标。在嵌入式工控板中装的是 Linux 操作系统, 并在 Linux 操作系统中安装了 ROS (robot operating system)。ROS 是一个用于编写机器人软件的灵活框架集成了大量的工具、库、协议, 提供类似操作系统所提供的功能, 可以极大简化繁杂多样的机器人平台下的复杂任务创建与稳定行为控制。因此嵌入式工控板可以通过 ROS 与飞控板建立通信。飞控板接收到目标物的位置并接受工控板发送的指令使无人机朝目标飞去, 并实现追踪。系统分为三层, 上层为视觉处理端, 底层为飞控端, 中间采用 ROS 作为通信机制。其系统代码结构图如图 1 所示。

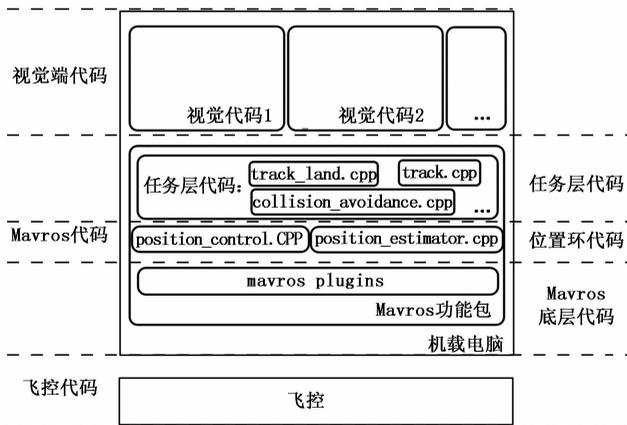


图 1 系统代码结构图

1 基于深度学习的目标检测

在计算机视觉领域, 基于深度学习的目标检测算法近几年发展迅速, 每年都有新的想法提出。R-CNN 提出于 2014 年, 是卷积神经网络在目标检测任务中的开山之作。R-CNN 因为要对每个候选区都要从头开始计算整个网络, 所以其运算效率低下。2015 年 RBG (Ross B. Girshick) 等结合了 SPP-Net 的共享卷积计算思想, 对 R-CNN 做出改进, 于是就有了 Fast R-CNN。RBG 团队在 2015 年, 与 Fast R-CNN 同年推出了 Faster R-CNN, Faster R-

CNN 的出现解决了 Fast R-CNN 的实时性问题, 在 Faster R-CNN 中提到了 RPN (Region Proposal Network) 网络, RPN 是一种全卷积网络, 它的前几层卷积层和 Faster R-CNN 的前五层是一样的, 所以 RPN 是在进一步的共享卷积层的计算, 以降低区域建议的时间消耗。较大的提高了目标检测精度和计算速度^[9-10]。

为了进一步提升目标检测的实时性, 基于单个网络的实时目标检测框架 YOLO (You Only Look Once), 框架和基于单个网络多层次预测框的目标检测算法 SSD (Single Shoot MultiBox Detector) 算法被提出。YOLO 虽然能够达到实时的效果, 但是其 mAP 与 SSD 的结果有很大的差距, 每个网格只能预测一个目标, 易导致漏检, 对于目标的尺度较敏锐, 对于尺度变化较大的目标泛化性较差^[11-14]。而无人机目标检测对准确性要求更高。综合对比下来, 本方案采用 SSD 目标检测算法。

SSD 的设计理念也是采用 CNN 网络来进行检测, 不同的是, 它采用了一个多尺度的特征图用来检测。多尺度检测顾名思义是采用尺寸不同的特征图, 分别采用一个大的特征图和一个比较小的特征图用来检测。用大的特征图检测小的目标, 小的特征图来检测大的目标。与 YOLO 采用的全连接层不同, SSD 最后直接采用卷积对不同的特征图进行检测提取。SSD 设置先验框, 这里是借鉴了 Faster R-CNN 中 anchor 的思想。图 2 展示了其基本架构。图 3 展示了大小两个特征图的效果。

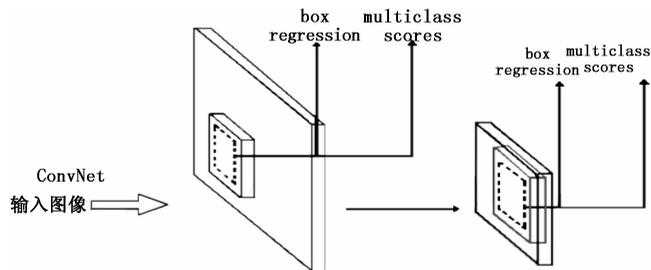


图 2 SSD 基本架构

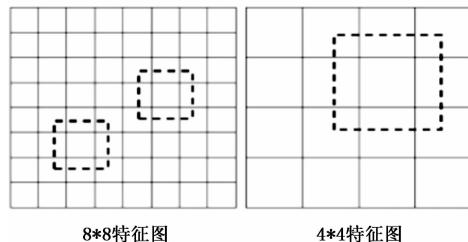


图 3 不同尺寸的特征图

SSD 以 VGG16 为基础模型, 在该模型上新增了卷积层来获取更多的特征图。从图 4 中, 可以看到该算法利用了多尺度的特征图来做检测。

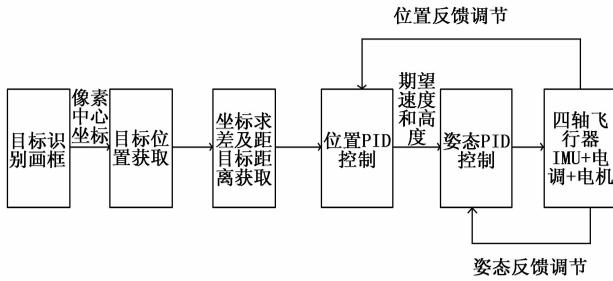


图 8 目标追踪控制流程图



图 9 自制模型训练流程图

3 试验结果及分析

3.1 硬件平台搭建

试验采用自行搭建的无人机飞行平台进行, 飞控板采用 Pixhawk 同时外接工业级 IMU, 使得无人机飞行时具有更好的稳定性, 嵌入式工控板采用 NVIDIA Jetson TX2 平台。

3.2 模型训练

训练过程主要分为两个步骤: 第一步是图片中的先验框与真实值进行匹配; 第二步是确定损失函数。

SSD 的先验框与真实值的匹配原则主要有两点: 第一, 对于图片中每个真实值, 找到与其重叠度最大的先验框, 然后该先验框与其进行匹配; 第二, 保证正负样本平衡, 保证正负样本的比例接近 1:3。

确定训练样本后, 然后确定损失函数。损失函数是由位置误差和置信度误差加权得到的, 其公式为:

$$L(x, c, l, g) = \frac{1}{N} (L_{conf}(x, c) + \alpha L_{loc}(x, l, g)) \quad (3)$$

式中, N 代表先验框的正样本的个数。这里 $x_{ij}^p \in \{1, 0\}$ 代表一个指示参数, 当 $x_{ij}^p = 1$ 时表示第 i 个先验框与第 j 个真实值进行匹配。对于位置误差, 其采用 Smooth L1 Loss, 定义如下:

$$L_{loc}(x, l, g) = \sum_{i \in Pos} * \sum_{m \in \{cx, cy, wx, h\}} x_{ij}^k smooth_{L1}(l_i^m - \hat{g}_j^m) \quad (4)$$

对于置信度误差, 其采用 SoftmaxLoss, 定义如下:

$$L_{conf}(x, c) = - \sum_{i \in Pos} x_{ij}^p \log(\hat{c}_i^p) \quad (5)$$

这里:

$$\hat{c}_i^p = \frac{\exp(c_i^p)}{\sum_p \exp(c_i^p)} \quad (6)$$

本次试验采用坦克战车作为目标物, 因此本文先训练模型识别坦克战车。这里训练模型使用 GPU 对训练过程进行加速。首先安装 Tensorflow Object Detection API 所需的开发环境, 并在官方的 Demo 上进行测试。然后运用数据进行训练和测试, 制作自己的模型。先从网上下载 1000 张坦克战车的图片。分为训练集与测试集。并设置标签 “tank”, 然后开始训练。图 9 是训练自己的模型的流程图。

模型训练结束后, 用训练好的模型检测效果。本文随机选取了几张坦克战车的照片, 以此来检验模型对于坦克战车识别的准确率。该模型在测试集的识别率为 96%。图 10 显示了在实际测试中, 目标均被准确识别出, 且识别为坦克的概率均在 95% 以上。



图 10 目标检测测试

3.3 目标识别与距离检测

如图 11 所示, 将训练好的模型移植到嵌入式工控板中, 调用单目摄像头并运行模型。然后进行目标识别。通过在 ROS 的可视化平台上看到工控板可以实时地检测到坦克战车, 同时将订阅到的距离信息也实时标示出来。



图 11 实时目标与距离检测试验结果

3.4 目标跟踪测试

我们接通飞控板与工控板，并建立通信机制。并且运行 MAVROS 包来监听一个飞控消息在 ROS 上订阅单目摄像头发布的目标位置信息，并通过 Mavlink 传输协议发送给飞控板。同时在工控板上发布追踪指令。飞控板接收到追踪指令后，开始朝电脑中的坦克目标飞去，并保持设定的距离。移动电脑，可以看到无人机朝着电脑的方向实现追踪。图 12 (a)、12 (b) 分别以第三视角与无人机第一视角拍摄了无人机对目标的追踪。

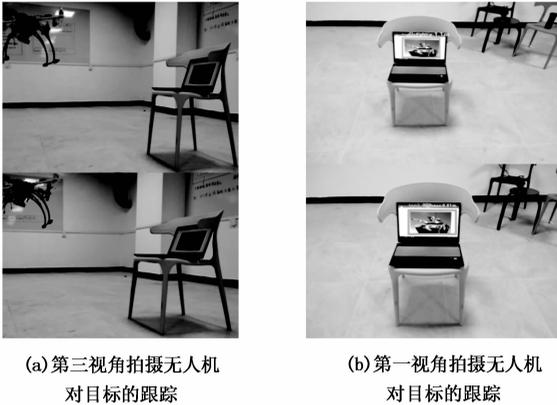


图 12 无人机对目标的追踪示意图

4 总结

本文介绍了利用单目摄像头，基于深度学习实现了对坦克战车的实时检测定位并追踪，此方法具有识别准确率高、跟踪效果好的特点，同时解决了传统的基于双目摄像机成本过高以及在复杂环境下识别准确率较低的问题。

通过试验验证了该方法能对目标进行高准确率识别，并在飞行试验平台上进行了实际飞行试验。由于这次选用的目标为坦克战车，实际的样本不好收集。所有的样本均来自网络，以及测试也采用在电脑上以图片的形式让无人机进行追踪。根据 SSD 算法，理论上是可以建立任意模型，也就是说，在模型训练比较好的情况下理论上无人机可以追踪任何物体。基于深度学习的多旋翼无人机目标跟踪通过实际的测试验证，试验结果证明了该方法的可行性。

参考文献:

- [1] 苏 贇, 王 挺, 姚 辰, 等. 基于合作目标的无人机目标跟踪方法 [J]. 机器人, 2018: 1-8.
- [2] 马 良, 穆朝絮, 杨万扣, 等. 四旋翼无人机目标跟踪系统设计 [J]. 控制工程, 2015, 22 (6): 1076-1081.
- [3] 张 谦, 邹依林. 基于智能视觉的无人机导航技术研究 [J]. 现代电子技术, 2018, 41 (16): 132-134, 139.

- [4] Zhang X M, Wang M G. Robust visual tracking based on adaptive convolutional features and offline Siamese tracker [J]. Sensors, 2018, 18 (7): 23-49.
- [5] Jackson Russell C, et al. Real-time visual tracking of dynamic surgical Suture threads [J]. IEEE transactions on automation science and engineering: a publication of the IEEE Robotics and Automation Society, 2018, 15 (3): 1078-1090.
- [6] 张 舸, 伊国兴, 高 翔. 基于视觉导航的旋翼无人机自主降落系统 [J]. 传感器与微系统, 2018, 37 (9): 77-80, 83.
- [7] 林 勇, 秦文静, 戚国庆. 基于机载视觉引导的无人机自主循迹研究 [J]. 电子设计工程, 2018, 26 (8): 21-24.
- [8] Qiang D. Visual UAV detection method with online feature classification [A]. Proceedings of 2017 IEEE 2nd Information Technology, Networking, Electronic and Automation Control Conference (ITNEC 2017) [C]. Beijing, 2017.
- [9] 刘丽军. 基于单目视觉的仓储物流机器人定位方法探讨 [J]. 电子制作, 2018 (12): 63-64.
- [10] 白志君. 四旋翼无人机室内自主导航系统的研究与实现 [D]. 厦门: 厦门大学, 2014.
- [11] Asa K, Funabora Y, Doki S, et al. Trajectory Planning of an UAV for automatic visual inspection system for infrastructure facilities considering measurement devices [A]. The Proceedings of JSME annual Conference on Robotics and Mechatronics (Robomec) [C]. 2016.
- [12] Xin L, Luo D L, Li H. A monocular visual measurement system for UAV probe-and-drogue autonomous aerial refueling [J]. International Journal of Intelligent Computing and Cybernetics, 2018, 11 (2): 166-180.
- [13] Liu W, et al. SSD: Single shot multibox detector [M]. Computer Vision - ECCV 2016. ECCV 2016. Lecture Notes in Computer Science, vol 9905. Springer, Cham. 2016.
- [14] Wang Y. Cognitive foundations of knowledge science and deep knowledge learning by cognitive robots [A]. 2017 IEEE 16th International Conference on Cognitive Informatics & Cognitive Computing [C]. 2017: 5-5.
- [15] Kriegman D, Triendl E, Binford T O. Stereo vision and navigation in buildings for mobile robots [J]. IEEE Transactions on Robotics & Automation, 1989, 5 (6): 792-803.
- [16] Zhang X M. Dynamic modelling of the hovering phase of a new V/STOL UAV and verification of the PID control strategy [A]. Proceedings of 2018 2nd International Conference on Advanced Technologies in Design, Mechanical and Aeronautical Engineering (ATDMAE 2018) [C]. Dalian University of Technology, Nazarbayev University, Hong Kong Society of Mechanical Engineers (HKSME), 2018: 6.
- [17] 赵搏欣. 无人机低成本微小型自主定位系统技术研究 [D]. 长沙: 国防科学技术大学, 2016.