

基于 YOLO 的四旋翼无人机人脸识别实验平台

陈鸿龙, 刘东永, 倪志琛, 孙 良, 刘 宝

(中国石油大学(华东) 控制科学与工程学院, 山东 青岛 266580)

摘 要: 该文简要介绍了 YOLO (you only look once) 人脸识别算法的基本工作原理, 设计了一套基于 YOLO 的四旋翼无人机人脸识别实验平台。该实验平台主要包括特洛 (Tello) 四旋翼无人机和地面站两个模块。地面站通过 Wi-Fi 模块与无人机通信来控制其飞行和图像采集, 接收来自无人机采集的图像信息, 并运行基于 YOLO 的人脸识别算法, 对图像信息进行人脸检测与识别。该人脸识别实验平台涵盖了无人机飞行控制、无线通信、图像处理以及深度学习算法等内容, 有助于学生深入学习和理解无人机控制和图像识别的原理及应用, 能够培养和提高学生针对复杂工程问题的创新和工程实践能力。

关键词: 深度学习; 无人机; 人脸识别; 无线通信

中图分类号: TP302.1 **文献标识码:** A **文章编号:** 1002-4956(2020)10-0107-05

Experimental platform for four-rotor unmanned aerial vehicle face recognition based on YOLO

CHEN Honglong, LIU Dongyong, NI Zhichen, SUN Liang, LIU Bao

(College of Control Science and Engineering, China University of Petroleum (East China), Qingdao 266580, China)

Abstract: The basic working principle of the YOLO (you only look once) face recognition is briefly introduced and an experimental platform for four-rotor unmanned aerial vehicle (UAV) face recognition based on YOLO is designed. The designed experimental platform consists of Tello UAV and ground control station. The ground control station communicates with the UAV using the Wi-Fi module. The UAV is controlled by the ground control station to fly and acquire the images. After receiving the images from the UAV, the ground control station can conduct the face detection and recognition by using the YOLO based face recognition algorithm. The experimental platform of face recognition covers UAV flight control, wireless communication, image processing and deep learning algorithm, which helps students to deeply learn and understand the principles and applications of UAV control and image recognition, and can cultivate and improve students' innovation and engineering practice ability for complex engineering problems.

Key words: deep learning; unmanned aerial vehicle; face recognition; wireless communication

无人机^[1-2]是指无需机载操作人员且具备动力装置的飞行器, 可以利用自身动力和控制系统进行自主飞行或由操作人员远程遥控飞行。无人机根据机身结构不同可分为固定翼无人机和多旋翼无人机。近年来随着控制理论与微电子技术的发展, 多旋翼无人机受

到了教学和科研人员的广泛关注。相比固定翼无人机, 多旋翼无人机具备机动性强、控制灵活等特点, 可在复杂环境中完成各种飞行任务。作为一种全新的技术, 无人机已经全面进入军用领域和民用商用领域, 为军事侦察^[3]、远程航拍探测^[4]、地形数据监测^[5]、物流运输^[6]和农业作业^[7]等应用提供了有效的解决方案。目前在无人机丰富的应用场景中, 利用四旋翼航拍无人机来获取图像信息是一种较为常见的应用。与手持式相机、监控摄像头等传统图像采集设备相比, 无人机可以对拟采集图像的目标进行主动的、多角度的拍摄, 能够极大地提高图像采集的质量和效率。

收稿日期: 2020-03-24

基金项目: 中国石油大学(华东) 校级教改项目 (JY-B201833); 山东省教学改革项目 (2015M016); 国家自然科学基金面上项目 (61772551)

作者简介: 陈鸿龙 (1984—), 男, 福建泉州, 博士, 副教授, 博士生导师, 中国石油大学(华东) 自动化系副主任, 研究方向为物联网。

E-mail: chenhl@upc.edu.cn

人脸识别^[8]是基于人的脸部特征信息进行身份识别的一种生物识别技术,人脸特征的唯一性和不易被复制的特性为身份识别奠定了基础。与指纹识别、虹膜识别等其他生物识别技术相比,人脸识别具备非强制性和非接触性,即对人脸进行采集时不需要强制目标完成特定的采集动作,同时在其采集过程中设备与目标也不需要直接接触。因此,利用人脸识别技术可以实现无感的身份确认,对于保护民众的个人隐私及人身安全都有重要的意义。同时,人脸识别系统集成人工智能、机器学习、数学建模、专家系统、视频图像处理等多种专业技术^[9],是生物特征识别的最新应用,其核心技术的实现展现了弱人工智能向强人工智能的转化。

本文通过无人机与地面站的无线通信,将四旋翼无人机与基于 YOLO 的人脸识别技术相结合,既可以充分发挥无人机图像采集灵活的优势,也可以充分利用深度学习人脸识别算法的鲁棒性,实现高效且精准的实时人脸识别效果。该实验平台有助于学生,尤其是自动化和测控等相关专业的学生,深入学习和理解无人机控制、无线通信、图像识别及深度学习等技术,有助于培养学生针对复杂工程问题的创新能力和工程实践能力^[10],提升他们今后的职业技能。

1 基于 YOLO 的人脸识别方法

利用无人机进行人脸图像采集可以极大地提升人脸图像采集的效率,但无人机多变的拍摄角度也给人脸识别算法的鲁棒性提出了更高的要求,同时为了实时的实时的人脸识别,算法应具备快速的图像处理能力,因此实验平台选用了基于深度学习的 YOLO 算法,并对其进行训练和优化使其具备人脸识别的能力。

YOLO 为 Redmond 等提出的一种基于深度学习的快速目标识别算法^[11],相比其他目标识别算法,YOLO 算法采用“一步走”策略,其将目标识别问题处理为回归问题,用一个卷积神经网络结构就可以直接预测出输入图像中目标的位置和类别概率。因此,该算法在具备较好的目标识别性能的同时还具备极快的检测速度,符合实验平台的需求。

YOLO 网络结构由前部 24 个负责特征提取的卷积层和后部 2 个负责回归预测的全连接层构成,简洁的网络结构使其具备极快的图像处理速度和端对端的优化能力。YOLO 网络的输入输出映射关系如图 1 所示,网络输入为分辨率 448×448 像素的 RGB 三通道图像,网络的输出结果为一个包含了图中所有目标位置信息与类别概率的张量。

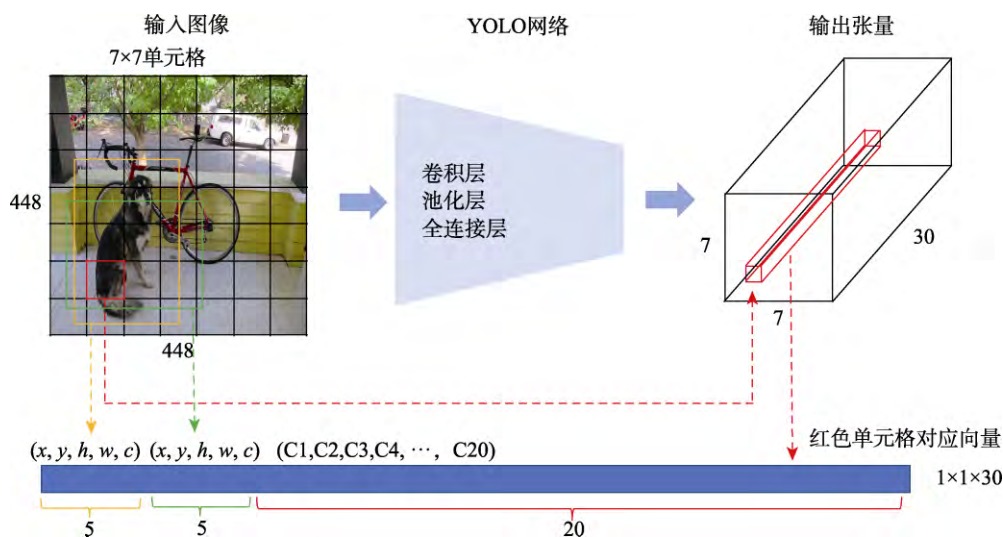


图 1 YOLO 算法输入输出映射关系图

448×448 像素的图像输入网络后,网络将图像划分为 $S \times S$ 个单元格,每个单元格负责预测落入该单元格的目标,每个单元格的预测信息由两部分构成:① B 个候选框信息,候选框信息包含候选框坐标 (x, y) 、宽高 (w, h) 及每个候选框的置信度 P 共 5 个值;② C 个概率值, C 为目标识别任务中需分类目标的总类别数,该概率值可显示单元格中目标的类别。因此 YOLO 网络输出的张量维度为 $S \times S \times (B \times 5 + C)$ 。其中: S

取值为 7,即输入图像被划分为 49 个单元格; B 取 2; C 取 20, C 的取值与训练算法所用的数据集有关。

因此图 1 中输入图像红色单元格所对应的预测信息为输出张量中红色的 1×30 的向量,该向量包含了红色单元格预测的两个候选框的信息及单元格内目标的概率。

该网络输出中,每个单元需要预测 B 个候选框的位置信息与每个边框的置信度 P , $P = \text{Pr}(\text{object}) \times \text{IoU}_{\text{pred}}^{\text{truth}}$ 。

其中: $\text{Pr}(\text{object})$ 为此预测的候选框的单元格内是否含有目标, 若有则为 1, 反之为 0; $\text{IoU}_{\text{pred}}^{\text{truth}}$ 为预测的边界框与输入图像中目标标注真实值 (ground truth) 的交并比。因此, P 代表预测该处是否有目标及候选框框选的准确度。

输出张量尾部的 C 类目标类别的条件概率定义为 $\text{Pr}(\text{Class}_i|\text{object})$, 该值表示单元格内存在物体且属于第 i 类的概率。

综上, 网络在测试时每个单元格预测最终输出的概率定义为

$$\frac{\text{Pr}(\text{Class}_i|\text{object}) \times \text{Pr}(\text{object}) \times \text{IoU}_{\text{pred}}^{\text{truth}}}{\text{Pr}(\text{Class}_i) \times \text{IoU}_{\text{pred}}^{\text{truth}}} \quad (1)$$

式 (1) 表示预测的边界框中目标属于某一类的概率及该边界框的准确度, 因此, 每个单元格最终输出 2 个候选框信息以及 2 列长度为 C 的概率值向量。最后, 将每个单元格输出的概率值进行非极大值抑制 (non maximum suppression, NMS), 即可得到最终预测的候选框及框内物体的类别概率。

2 基于深度学习的无人机人脸识别系统

本文设计的基于深度学习的无人机人脸识别实验平台的系统结构如图 2 所示, 主要包括四旋翼无人机和作为地面站的便携式计算机两部分。该平台中无人机与地面站通过 Wi-Fi 模块进行连接, 无人机由地面站远程控制, 按操作人员的指令进行飞行并对周围区域进行图像采集, 同时将采集的图像无线回传至地面站。地面站接收来自无人机的图像信息后对其进行实时计算与分析, 从而实现对无人机前方目标的人脸检测与识别。



图2 基于深度学习的无人机人脸识别实验平台系统结构图

本实验平台采用特洛 (Tello) 四旋翼无人机, 如图 3 所示。特洛无人机为大疆 (DJI) 公司与英特尔 (Intel) 公司共同技术支持、睿炽公司生产的微型无

人机, 采用四个空心杯电机为动力系统并采用 5 V 聚合物锂电池为动力来源。飞行过程中无人机内置的飞行控制模块将实时监测并计算无人机的当前姿态及坐标, 通过对四个空心杯电机转速的精准调节来实现稳定飞行^[12-13]。特洛无人机还同时具备下视定位模块及前向视觉模块, 飞行过程中无人机可以利用下视定位模块实现低空飞行的厘米级定位以保证飞行的稳定性, 前向视觉模块的高清摄像头可以实现高清图像采集以满足该实验平台人脸图像采集的要求。在通信方面, 特洛无人机内置了 Wi-Fi 通信模块, 利用该模块可以实时接收地面站的控制指令并且实现无人机与地面站之间的实时视频流传输。同时, 实验平台可通过该无人机开放的系统开发工具 (SDK) 来利用地面站程序对无人机进行指令控制与图像获取, 从而实现实验平台中无人机的飞行及图像采集任务。



图3 特洛四旋翼无人机

地面站主要负责运行该实验平台的软件程序以实现对无人机的飞行控制、回传图像的人脸识别以及识别结果的显示。在硬件上, 地面站需配备 Wi-Fi 模块以与无人机建立稳定的连接。由于本实验平台选用的人脸识别算法的实时运行需要较高计算能力, 因此该计算机内置高性能图像处理器 (GPU)。

在本实验平台中, 地面站装有 Ubuntu 操作系统并运行使用 Python2.7 语言编写的实验平台软件。该实验平台软件的工作流程如图 4 所示。在平台软件系统开始运行时, 地面站首先通过 Wi-Fi 模块与无人机进行连接, 并通过 UDP 协议实现地面站与无人机的通信, 然后调用无人机 SDK 使无人机进入指令控制模式, 并向无人机发送视频流回传请求。无人机在起飞后采集前方图像并将高清视频流实时回传至地面站。地面站接收到无人机回传的视频图像后便利用基于深度学习的 YOLO 算法对图像中的人脸进行快速检测与识别, 并显示识别结果。运行过程中地面站将持续根据键盘指令控制无人机飞行, 若输入降落指令则无人机降落且该平台软件结束运行。

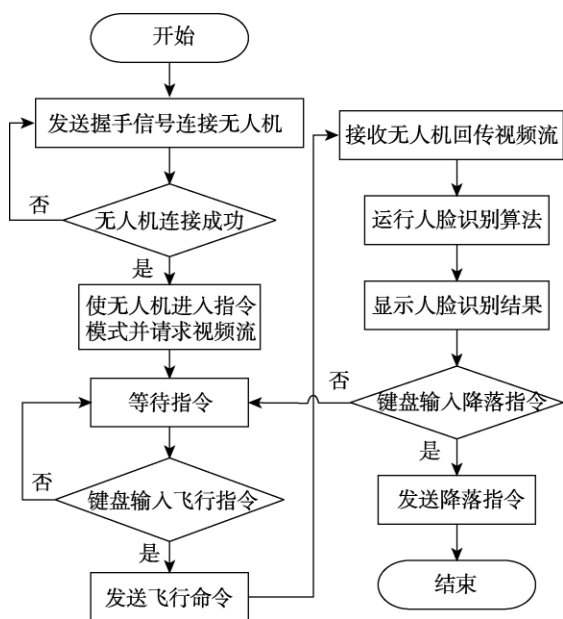


图4 实验平台地面站软件工作流程示意图

3 无人机人脸识别实验平台的测试

在该实验平台的测试中,首先需要建立人脸数据集并对平台软件中的YOLO人脸识别算法进行训练。数据集的样本采集示意图如图5所示。

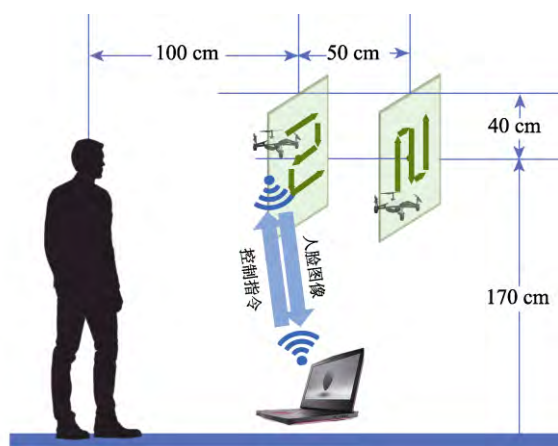


图5 人脸数据集样本采集示意图

本文在该实验平台中设计了一种利用无人机进行

人脸样本采集的程序以实现各种角度、距离条件的人脸图像采集。该程序由地面站运行以控制无人机完成图像采集,运行时地面站与无人机先进行Wi-Fi连接并使无人机进入指令控制模式,然后,地面站将控制无人机起飞并悬停于170 cm的高度,无人机位于采集目标正前方并距离目标100 cm。随后,特洛无人机将以当前所处位置为几何中心,在长宽均为80 cm的正方形竖直平面内进行水平与垂直的“弓”字形运动,在“弓”字形飞行期间无人机朝向始终保持为起飞状态的朝向,在飞行过程中,无人机将对前方人脸图像进行视频录制并将视频发送至地面站保存。在完成了当前距离的样本采集后,无人机将后退至距目标150 cm处以同样的方式进行人脸的图像采集,采集结束后无人机将降落并结束程序运行。在完成图像采集任务后,可以得到包含各个角度人脸图像且帧率为30帧每秒的高清视频,对视频文件进行抽帧操作即可获得包含丰富面部信息的静态图像。

在该实验平台的测试中,利用上述程序采集了来自6名志愿者共1996张图像样本,而后利用开源的LabelImg软件对每张图像中的人脸位置以及标签信息进行标注,最终获得包含1996个样本和标注信息的无人机人脸数据集。利用该数据集对YOLO算法进行50400次迭代训练,即可得到该网络具备人脸识别能力的最优权重文件。之后,对载入该权重文件的YOLO人脸识别算法及无人机人脸识别实验平台的人脸识别性能进行测试。

在对于YOLO人脸识别算法的离线测试中,从人脸数据集中随机选取了包含6名志愿者面部信息的300张图像样本对算法的性能进行测试,每张图像中至少包含1名志愿者的人脸图像。利用YOLO算法对测试集样本进行逐一人脸识别后,算法将输出每张图像中的人脸数量、人脸位置、人脸标签(志愿者信息)及每张人脸的置信度,对所有样本的识别结果进行统计即可证明该算法的人脸识别性能,测试集中各志愿者的人脸识别结果如图6所示。

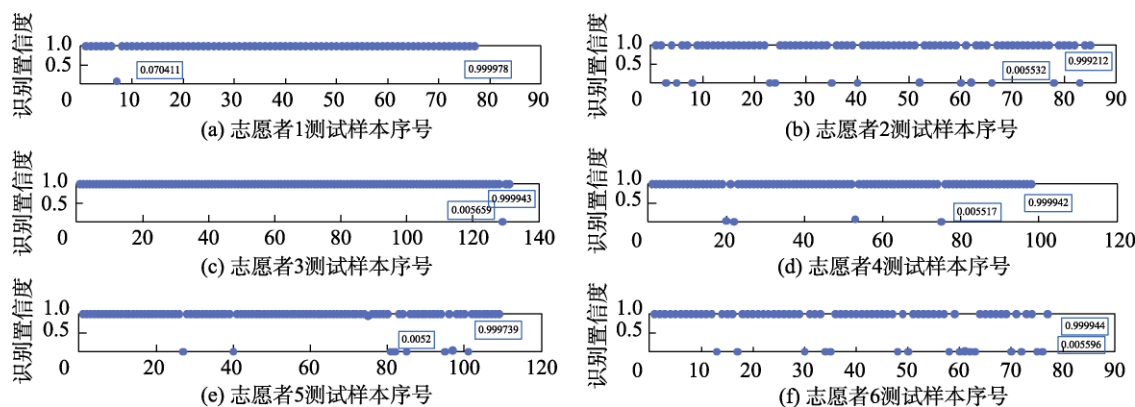


图6 测试集中各志愿者样本识别置信度统计图

经统计,在300张图像样本中共有585张来自志愿者的面部,由图6可以知,在经过基于YOLO的人脸识别算法处理后,样本中大部分面部都能被以接近100%的置信度进行有效识别,测试集的样本中共有43张人脸图像没有被有效地识别出来,因此YOLO人脸识别算法在该测试集中的人脸识别率为92.6%,具备较好的人脸识别能力。

本文还对实验平台进行了在线的人脸识别测试,在测试过程中地面站软件运行正常,无人机能够按照指令稳定飞行且进行图像采集,同时地面站中运行的基于YOLO的人脸识别算法能够实时地对无人机所采集图像进行人脸识别。

测试结果显示该实验平台在进行人脸识别时效果较好,无人机在面对一个或多个志愿者时,平台对数据集包含的6名志愿者的正脸识别置信度均在98%以上,测试过程及效果如图7所示。



图7 无人机人脸识别实验平台识别过程及测试结果示意图

4 结语

本文以专业课课程建设为出发点,设计了一套基于深度学习的四旋翼无人机人脸识别实验平台,地面站通过Wi-Fi通信模块控制无人机的飞行和图像采集,接收无人机的图像信息,并利用基于深度学习的人脸识别算法对图像信息进行人脸检测与识别。该实验平台可服务于本校自动化专业的“人工智能技术”

和“模式识别”两门课程,有助于学生深入学习和理解无人机飞行控制、无线通信、图像处理和深度学习算法等技术,培养针对复杂工程问题的创新能力和工程实践能力,助力学生今后职业技能的提升。

参考文献 (References)

- [1] 温凯,陈缪.全尺寸四旋翼无人机教学平台设计与课程开发[J].实验技术与管理,2018,35(9):100-203.
- [2] 罗偲,纪仁杰,韩彬,等.基于Odroid的四旋翼无人机教学平台设计[J].实验技术与管理,2016,33(11):165-167.
- [3] 李岳龙,李旭阳.无人机在军事领域的发展前景[J].科技创新,2018(12):59-60.
- [4] 刘梦亭,吴凡,陈志钦.基于ATmega328p与Android的航拍侦察六轴对转飞行器的设计[J].电子世界,2016(18):147-148.
- [5] 薛文斌,庞国楹,汪志远.应对突发事件的无人机监测任务分配和航行路径规划[J].军事交通学院学报,2017(4):90-94.
- [6] 翟少博,张雨佳,陈龙,等.无人机在电商物流领域应用可行性分析[J].全国流通经济,2018(9):19-20.
- [7] 陈鹏飞.无人机在农业中的应用现状与展望[J].浙江大学学报,2018(4):399-406.
- [8] 范哲意,曾亚军,蒋姣,等.视频人脸识别实验平台的设计与实现[J].实验技术与管理,2016,33(3):159-161.
- [9] 严严,章毓晋.基于视频的人脸识别研究进展[J].计算机学报,2009(5):878-886.
- [10] 刘杨,蒋文明,陈树军,等.基于响应面的电絮凝除乳化油研究[J].实验室研究与探索,2018,37(274):19-24.
- [11] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: Unified, real-time object detection[C]. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, Nevada, USA, 2016: 779-788.
- [12] 刘辉邦,褚金奎,支炜,等.基于STM32的无人机姿态测量系统设计[J].传感器与微系统,2013,32(8):108-110.
- [13] 张新英,余发军,刘聪.基于模糊PID控制的四旋翼无人机设计[J].实验室研究与探索,2017,36(4):56-59.

(上接第106页)

- [8] ERDEMIR A. Review of engineered tribological interfaces for improved boundary lubrication[J]. Tribology International, 2005, 38(3): 249-256.
- [9] 汪久根,张建忠.边界润滑膜的形成与破裂分析[J].润滑与密封,2005(6):4-8.
- [10] HAMROCK B J, SCHMID S R, JACOBSON B O. Fundamentals of fluid film lubrication[M]. Boca Raton, USA: CRC press, 2004.
- [11] 刘志全,葛培琪.弹流膜厚的测试方法及其特点[J].机械工

程师,1996(2):40-41.

- [12] TALLIAN T E, CHIU Y P, HUTTENLOCHER D F, et al. Lubricant films in rolling contact of rough surfaces[J]. ASLE Transactions, 1964, 7(2): 109-126.
- [13] FUREY M J. Metallic contact and friction between sliding surfaces[J]. ASLE Transactions, 1961, 4(1): 248-248.
- [14] 林福严,雷冀,王力伟.改造四球实验机进行摩擦因数测试[J].煤矿现代化,2009(6):58-59.
- [15] 王志芳,刘逸飞,傅行军,等.润滑油黏温特性实验研究[J].润滑与密封,2016,41(4):24-28.