

基于人像分割的智能搜救无人机系统设计

王 蓉, 吕祖盛, 孙 嘉, 江子妍, 肖 建
(南京邮电大学 电子与光学工程学院, 江苏 南京 210023)

摘 要: 为了在确保救援人员安全的同时, 迅速有效地开展野外搜救任务, 对搜救方式的选择变得非常重要。针对传统搜救系统消耗人力大、搜索效率低等弊端, 提出了一种无人机智能搜救系统。系统分为自动控制、图像拼接和人像检测三个子系统。控制系统保证无人机依照指定路线安全飞行, 实时地采集并回传图像; 图像拼接系统将回传的图像通过 ORB 特征提取算法合成大型高清航拍影像, 以便救援人员能在最短的时间内根据遇难者所在地地形和环境特点规划最优的救援路线。人像检测系统使用谷歌最新的语义分割模型 DeepLab V3+ 神经网络进行人像分割, 以实现人像检测的功能。该神经网络运行速度快、精度高、对截断人像识别效果较好, 适用于野外救援等紧急情况的实时应用场景。经过多次实验, 结果表明该无人机搜救系统测试稳定, 识别准确率高, 能在多种复杂的野外环境中发挥较好的救援作用。

关键词: 无人机救援; 人像分割; 图像拼接; 特征提取; 人体检测

中图分类号: TP391

文献标识码: A

文章编号: 1673-629X(2020)08-0147-05

doi:10.3969/j.issn.1673-629X.2020.08.025

Design of Intelligent Rescue UAV System Based on Portrait Segmentation

WANG Rong, LYU Zu-sheng, SUN Jia, JIANG Zi-qian, XIAO Jian

(School of Electronic and Optical Engineering, Nanjing University of Posts and Telecommunications,
Nanjing 210023, China)

Abstract: In order to quickly and efficiently carry out field search and rescue missions while ensuring the safety of rescue workers, the choice of search and rescue methods becomes extremely important. Aiming at the drawbacks of the traditional search and rescue system, such as high manpower consumption and low search efficiency, an intelligent search and rescue system for UAV is proposed. The system is divided into three subsystems: automatic control, image stitching and portrait detection. The control system ensures that the UAV can fly safely according to the planned route and collect and return images in real time. The image stitching system combines the returned images into large high-definition aerial images by ORB feature extraction algorithm, so that rescuers can plan the best rescue route according to the terrain and environmental characteristics of the location of the victims in the shortest time. The portrait detection system uses Google's latest semantic segmentation model DeepLab V3+ neural network for portrait segmentation to realize the function of portrait detection. The neural network has the advantages of fast running speed, high precision and excellent recognition effect for truncated portrait, which is suitable for the real-time application scenarios of field rescue. After many experiments, the results show that the UAV search and rescue system is stable in testing, high in recognition accuracy, and can play a better role in a variety of complex field environment.

Key words: UAV rescue; portrait segmentation; image stitching; feature extraction; human detection

0 引 言

近年来,许多“驴友”喜欢近距离探索自然。但是有些时候“驴友”会在野外步入险境,需要外界的救援。传统的搜救方式由于耗资大、范围不精确、巡查时间长,导致搜救效率低下。例如,2017年8月8日九寨沟地震中就暴露了现有搜救系统存在的巨大缺陷。由

于九寨沟地形复杂多样,传统的人力搜救不仅高成本低精度,还可能会因为灾区的天气和通信等条件过于恶劣而给搜救人员带来巨大的阻碍和危险。于是,方便灵活且精度较高的搜救方法成了野外救援的一个研究课题。传统的人力搜救暴露出的成本高、搜索速度慢等问题也促进了无人机智能搜救系统在野外救援方

收稿日期:2019-10-08

修回日期:2020-03-09

基金项目:国家自然科学基金(61874059);国家级大学生创新训练计划项目(SZDG2018011)

作者简介:王 蓉(1998-),女,研究方向为集成电路设计;肖 建,博士,副教授,硕导,CCF会员(A8837M),研究方向为嵌入式系统应用。

面的应用^[1]。无人机作为飞行机器人,具备小巧便捷、活动范围广、自由度高等特点,因此近年来被广泛应用于军事、农业等领域。文中设计了一款基于人像分割技术的搜救系统。无人机通过摄像头在空中实时采集图像并进行图像回传和 GPS 定位,便于搜救人员确定当前受困人员的位置、状态以及环境。以最快的速度规划最优路线并进行搜救。通过在公园、学校操场和山地等多种户外场景中的测试,证明该系统有效可行,能适用多种复杂的野外场景并满足实际使用需求。

1 搜救无人机整体方案设计

当搜救人员收到有人遇险的消息后,会给无人机下达救援指令。当无人机飞至事故区域上方时先进行自检,以确保能正常飞行,然后根据地面站传来的事故区域的两个对角坐标制定蛇形路径进行巡检。在巡检时,无人机回传当前视野图像并进行 GPS 定位。地面站将接收到的图像进行整合分析:采用人像分割技术进行人体目标检测^[2-3]来确定遇险人员,并通过 ORB 特征提取合成大型高清航拍影像,结合遇难者 GPS 位置坐标帮助搜救人员确定救援路线。

2 无人机硬件系统

使用 Pixhawk 作为飞行控制器,能更好地实现姿态调整和 GPS 巡航。使用 STM32F1 控制 Pixhawk 各通道油门,通过 MAVLink 协议进行飞行器自检和任务下达。无人机搭载树莓派作为图像采集和通信单元,用于图像获取、视频编码和远程图像传输。同时,树莓派也用于连接 STM32F1 和地面站,进行远程的数据传输和任务下达。本系统巡检设备安装在云台上,使用 8 mm 焦距的镜头,搭载大功率 WIFI 模块,保证无人机和地面站进行稳定的数据交互和图像传输。无人机飞行高度 17 m,速度 20 m/s,每秒采集 3 张图像,完成一平方公里巡检任务需要 8 分钟。

3 人体目标检测算法设计

3.1 神经网络的选择

在传统的行人检测等人像检测任务中,多用物体检测的神经网络框架进行检测,例如 Faster-RCNN^[4]、SSD^[5]、YOLO^[6]等,都能达到较好的识别精度和速度。但是在野外搜救中,由于环境较为复杂,有时候会出现人像被遮挡的情况,而且因为应用环境的特殊性,每一张图像都十分珍贵,不能容忍物体检测出现识别不到的情况,并且在识别过程中对预测区域与真实区域的交叠率要求也不太严格。因此,选用由谷歌开发的 DeepLab V3+^[7]图像分割神经网络来进行人像检测任务,使用 Xception^[8]神经网络进行特征提取。DeepLab

V3+结合了空间金字塔模块^[9-10]在输入 feature 上应用多采样率扩张卷积、多感受野卷积或池化,探索多尺度上下文信息和编解码器结构通过逐渐恢复空间信息来捕捉清晰的目标边界的优点。DeepLab V3+在 PASCAL VOC 2012 上验证了模型的有效性,在没有添加任何后端处理的情况下达到了 89% mIoU。

3.2 DeepLab V3+神经网络框架

使用的 DeepLab V3+神经网络框架如图 1 所示。编码器由 Xception 神经网络、ASPP 和一个 1×1 的卷积组成^[10]。DeepLab 使用了深度可分离卷积和带孔卷积以提升网络的性能。下面着重介绍 ASPP(带孔空间金字塔池化)的特点以及在文中的应用。

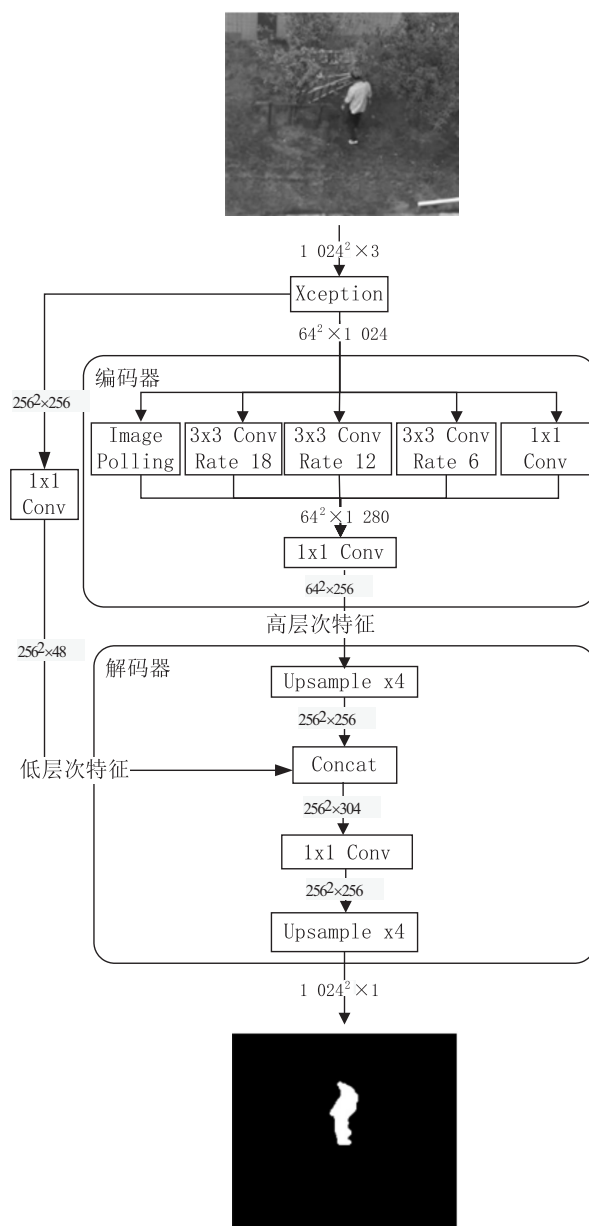


图 1 DeepLab V3+神经网络框架

ASPP(atrous spatial pyramid pooling,带孔空间金字塔池化),就是对同一特征图用不同大小的卷积核进行卷积操作,以实现对不同大小特征的感知。使用

带孔卷积来调整感受野的大小,同时使用深度可分离卷积来减少参数和提高运行效率。文中使用 $1\,024 \times 1\,024 \times 3$ 的图像作为 Xception 神经网络的输入,输出为 $1\,024 \times 1\,024 \times 1$ 的图像蒙版。ASPP 使用了 1×1 卷积, 3×3 rate=6、12、8 的带孔卷积和传统的 pooling 叠加。图像在经过神经网络运算后得到的蒙版,需要再经过膨胀运算以消除蒙版中的一些因识别不全而产生的空洞。在本次项目中测试使用的 GTX1060 GPU, DeepLab V3+神经网络框架总参数为 41.25 M,训练使用 10 万张航拍图像,测试 mIoU 可达 72%,运行速度可达 3.2 FPS。

图像分析系统运行流程如图 2 所示,主要分为人像分割模块和图像拼接模块。图像由无人机摄像头采集,使用 WIFI 模块实时地将图像信息与当前 GPS 坐标信息反馈给地面站,地面站同时将图像传入人像分割模块和图像拼接模块进行处理。

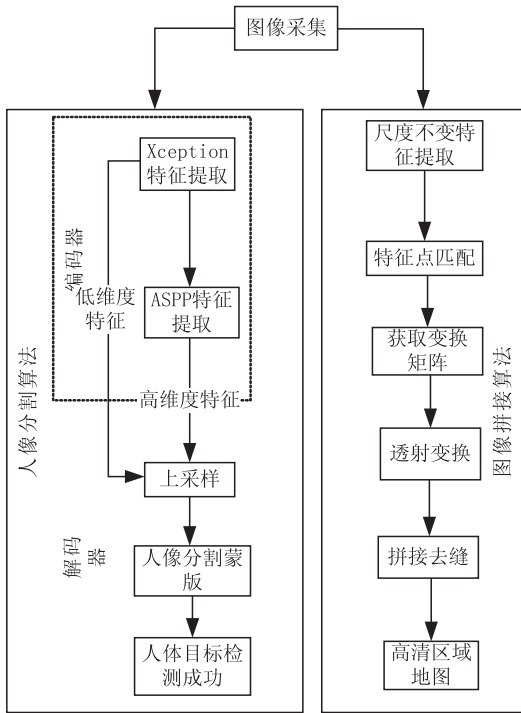


图 2 图像分析系统流程

其中,图像分割模块负责人像搜索的任务,将图像经过 Xception 特征提取、ASPP 多尺度特征提取以及上采样等一系列处理后得到人像分割蒙版。在本项目中,设定人像分割蒙版阈值为 0.5,若蒙版中存在置信度大于 0.5 的像素点,则认为是人像检测成功,随即在地面站上生成一条包含该图像、蒙版、无人机拍摄该图像时的 GPS 坐标和人像区域大小信息的记录,供搜救人员查看。

图像拼接模块负责将图像进行拼接,生成一张覆盖整个航拍区域的超高清全景影像。该模块对输入的图像进行尺度不变特征提取、特征点匹配、计算变换矩

阵和拼接去缝操作,将输入的图像依次进行拼接。当无人机完成一列巡检任务掉头进行第二列巡检时回传掉头信号,图像拼接模块完成当前列图像拼接任务,并保存图像。当无人机完成下一列巡检任务时,将列图像进行拼接,得到两列图像的拼接结果图,依次进行该过程直到无人机完成本次搜救任务。

4 图像拼接算法设计

4.1 图像特征提取

系统中使用 ORB (ORientedBrief) 算法进行特征提取^[11-13]。ORB 特征提取的优点在于效率高,约为 SIFT 运行速度的 100 倍, SURF 运行速度的 10 倍。提取特征的准确度相较于 FAST 算法也要好得多,综合性能较强。经多次测试,该算法对两张 500×500 的图像进行特征点提取仅耗时 200 ms。其不足之处在于提取的特征不具备尺度不变性。但是由于文中无人机始终与地面保持着一定高度飞行,所以几乎不存在尺度变化,因而该缺点影响可以忽略。

4.2 特征点匹配

对图像做特征点匹配,即使是 SIFT 算法效果也是非常糟糕的,为了减少特殊原因产生的误匹配关键点,可利用式(1)求得比例系数 R 。在本系统中,设定阈值 $T=0.5$,将图像 A 与图像 B 中求得 $R < T$ 的一对特征点作为优秀匹配特征点。

$$\frac{D_1}{D_2} = R \quad (1)$$

其中, D_1 为图像 A 中某一特征点与图像 B 所有特征点欧氏距离的最小值, D_2 为图像 A 中某一特征点与图像 B 所有特征点欧氏距离的次小值。

4.3 图像拼接与去缝

获得了匹配特征点后,用 RANSAC 方法^[14-15]计算多个二维点对之间的最优单映射变换矩阵,使用投射变换将图像 A 变换至图像 B 的坐标空间,并拷贝至图像 C 中。在拷贝的过程中由于光照等环境因素,直接拷贝拼接会使得接缝过于明显,文中使用加权相加的方式进行图像融合。图像融合公式为:

$$P_X = \begin{cases} P_{AX}, X \leq X_{begin} \\ P_{AX} \times \left(1 - \frac{X - X_{begin}}{X_{end} - X_{begin}}\right) + P_{BX} \times \frac{X - X_{begin}}{X_{end} - X_{begin}}, X_{begin} < X < X_{end} \\ P_{BX}, X \geq X_{end} \end{cases} \quad (2)$$

其中, P_X 为结果图像 C 在 X 位置处的像素值, P_{AX} 为图像 A 在 X 位置处的像素值, P_{BX} 为图像 B 在 X 位置处的像素值, X_{begin} 为拼接图像相交起点, X_{end} 为拼接图像相交终点。图 C 为图 A 和图 B 拼接后的结果(见

图3),实验证明 ORB 算法能够快速准确地匹配到有效特征点并进行图像拼接。

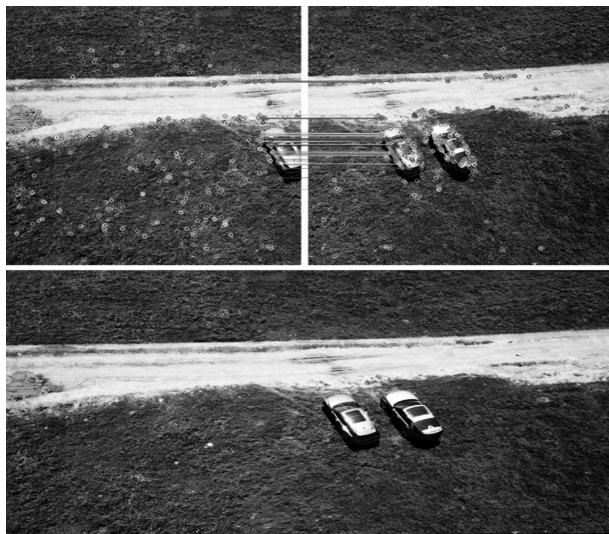
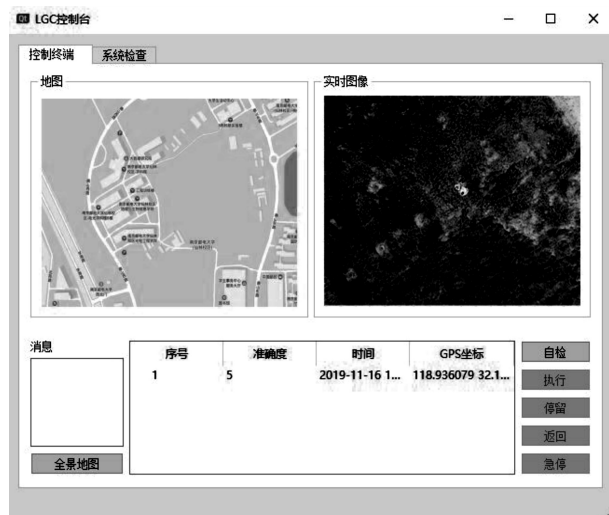


图3 图像拼接效果

(左上为图像 A,右上为图像 B,下方为图像 C)

5 上位机系统设计

文中开发了一款配套的用户界面(见图4),使用户能以更加便捷的方式操作本系统。系统由两大选项卡组成:控制台和系统检查。控制台负责用户在使用时的所有操作,系统检查负责用户在使用前进行无人机图像回传情况、人像分割运行情况等的检查。控制台界面由4大部分组成,分别是地图界面、无人机实时航拍图像界面、消息记录界面和控制按钮。地图界面可供用户定制巡检区域;实时航拍界面可供用户实时地观察到无人机的航拍影像;消息记录界面会在每次人像分割成功后生成一条记录,双击记录可打开人像分割查看界面,结束一次巡检即可点击全景地图按钮打开查看拼接界面;控制按钮进行无人机的自检、执行、停留、返回和急停。



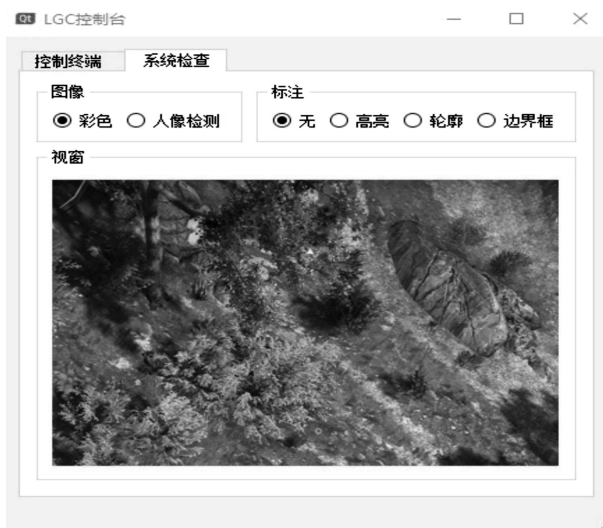
(a)控制台操作界面



(b)人像检测记录



(c)全景拼接



(d)系统检查

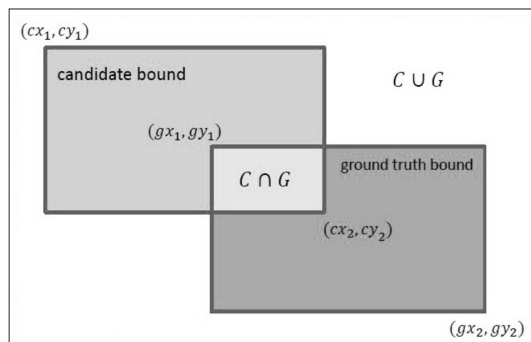
图4 上位机系统界面

6 实验与结果分析

交并比(intersection-over-union, IoU),是预测区域(candidate bound)与真实区域(ground truth bound)

的交叠率,即它们的交集与并集的比值。 IoU 在最理想情况下是完全重叠的,即比值为 1。 IoU 的平面示意图如图 5(a) 所示,由图 5(a) 可推导出 IoU 的计算公式为:

$$\text{IoU} = \frac{\text{are}(C) \cap \text{are}(G)}{\text{are}(C) \cup \text{are}(G)} \quad (3)$$

(a) IoU 平面示意图

mIoU 值是一个衡量图像分割精度的重要指标。 mIoU 可解释为平均交并比,即多张测试图像的 IoU 的均值。为了验证该无人机搜救系统在野外巡检时的人像检测精度,从 3 m ~ 23 m 拍摄高度的航拍图像中每隔 5 m 选取 250 张图像,共计 1 000 张图像进行 IoU 值和人像识别率测试(见表 1)。

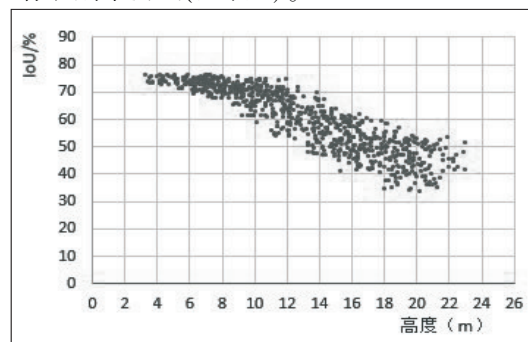
(b) IoU -航拍高度关系

图 5 测试结果分析

表 1 人像识别准确率测试结果

测试距离/m	数据集总数/张	识别正确数量/张	$\text{mIoU}/\%$	识别准确率/%	平均识别准确率/%
3 ~ 8	250	246	73.03	98.4	91.4
8 ~ 13	250	238	62.42	95.2	
13 ~ 18	250	223	49.49	89.2	
18 ~ 23	250	207	42.62	82.8	

准确率计算公式为:

$$\text{Acc} = \frac{R_s}{\text{Sum}} \quad (4)$$

其中,对于单个高度段的测量结果, Acc 表示识别准确率, R_s 表示识别成功的图片数量, Sum 表示测试集总数,实验中 Sum 默认取值为 250。整合四个高度段的测试结果并计算其平均准确率 mAcc ,计算公式为:

$$\text{mAcc} = \frac{\text{Acc}}{4} \quad (5)$$

测试结果显示,系统的视觉分析功能在无人机飞行高度为 18 m 以下时都有非常好的识别效果,识别准确率大多都能达到 90% 以上。但其中仍存在人像无法识别的情况,原因主要有两个方面:一方面,随着无人机飞行高度的不断增加,人像在图像中的占比越来越小,导致 IoU 在持续减小,具体变化情况如图 5(b) 所示;另一方面,在不同的遇险环境下,遇难者可能会被树木、草丛等物体遮挡,导致人像在图像中的占比大小不一致。这两种情况都会不同程度地对识别的准确率造成负面影响。实验表明,该系统测试稳定,识别准确率高,正常情况下能满足当前对野外遇险人员的搜救需求。

7 结束语

主要从目标检测、图像拼接、上位机系统等方面对智能搜救无人机进行了设计。设计的无人机搜救系统功能完善,操作简便,实时性好,能较大地减少野外救援的工作量。实验表明:该搜救系统稳定性好,能实现多方位巡检。目标检测及图像拼接的效率和精度也较高,可以满足多种复杂环境下的搜救需求。

参考文献:

- [1] 杨 铭,靳志勇,詹嘉文,等. 野外生命搜救无人机探测系统的设计[J]. 轻工科技,2018(4):73-75.
- [2] 侯 杰. 巡逻机器人中的行人检测技术研究[D]. 重庆:重庆邮电大学,2017.
- [3] 朱 玮. 基于视觉的四旋翼飞行器目标识别及跟踪[D]. 南京:南京航空航天大学,2014.
- [4] REN S, HE K, GIRSHICK R, et al. Faster R-CNN: towards real-time object detection with region proposal networks [C]//International conference on neural information processing systems. Cambridge: MIT Press, 2015:91-99.

(下转第 156 页)