文章编号:1001-5078(2022)03-0435-07

·图像与信号处理 ·

YOLOv5 与光流相结合的红外小目标检测算法

刘宝林,范有臣,秦明宇,谢鹏飞,郭惠超,张来线 (航天工程大学,北京101416)

摘 要:针对运动场景下红外小目标轨迹提取不准确的问题,提出采用 YOLOv5 与光流相结合的方法校正目标轨迹。首先 YOLOv5 网络,然后对比了 LK 和 HS 两种光流算法通过单样本 K-S 检验计算分布拟合度,构建 Q-Q 图得到平台真实运动量,最后结合 YOLOv5 校正目标轨迹。 实验结果表明,LK 算法更适合红外图像光流值的提取,YOLOv5 与光流相结合的方法在地/空 背景下红外图像数据集检测准确率达到 90 % 以上,损失率在 0.02 以下,对于区分真实和虚假 小目标有着重要意义。

关键词:YOLOv5;光流法;轨迹校正;红外小目标 中图分类号:TN215;TP391.41 文献标识码:A DOI:10.3969/j.issn.1001-5078.2022.03.020

Infrared small target detection algorithm combined with YOLOv5 and optical flow

LIU Bao-lin, FAN You-chen, QIN Ming-yu, XIE Peng-fei, GUO Hui-chao, ZHANG Lai-xian (Space Engineering University, Beijing 101416, China)

Abstract: Aiming at the problem of inaccurate extraction of the trajectory of small infrared targets in sports scenes, a method combining YOLOv5 and optical flow is proposed to correct the trajectory of the target. First, the training network of small infrared targets in YOLOv5 is introduced, and then the LK and HS optical flow algorithms are compared. A single-sample K-S test is used to calculate the distribution fit; the Q-Q diagram is constructed to obtain the true a-mount of background motion, and finally combined with YOLOv5 to correct the target trajectory. The experimental results show that the LK algorithm is more suitable for the extraction of infrared images optical flow value. The combination of YOLOv5 and optical flow has a detection accuracy of more than 90 % and a loss rate of less than 0.02 in infrared image datasets under ground/air background, which is of great significance for distinguishing between real and false small targets.

Keywords: YOLOv5; optical flow method; trajectory correction; infrared small target

1 引 言

高于热力学零度的物体都会辐射红外线,红外 辐射穿透云、雾、烟尘的能力强,红外探测成像是被 动接收红外辐射成像,具有良好的抗干扰能力和隐 蔽性,能够探测到几公里到几十公里远的目标。红 外成像系统视场大,远距离目标成像呈现点状,目标 成像大小介于3×3和9×9个像素^[1],所占的像素数少,成像面积占整幅图像比例很小,信杂比低^[2-3],检测难度大。由于散射的影响,目标边缘信息微弱,缺少固定的形状和纹理信息,不利于特征的提取。红外小目标检测算法分为基于单帧检测和基于多帧检测的算法^[4]。单步目标检测算法中较典

作者简介:刘宝林(1988 -),男,硕士,主要从激光图像处理方面的研究。E-mail:718413811@163.com 收稿日期:2021-06-07;修订日期:2021-07-25

型的是 YOLO 方法^[5-7],该方法由物体中心的格子 来完成目标预测,提取目标特征、候选框边界回归及 分类都在同一个网络中实现的,检测速度可达到 100 FPS 以上,解决了实时检测的任务要求。

复杂背景中包含多个与真实小目标相像的点, 目标的检测难度大,在训练数据集有限的条件下,利 用已有的小目标训练权重进行检测,容易漏掉真实 的小目标。为了将可能的小目标都检测出来,本文 在使用 YOLOv5 检测时,用低检测网络对复杂背景 图像进行检测,检测结果中除了真实小目标外,包括 多个虚假目标,不易提取目标的轨迹。文将 LK 和 HS 两种光流算法进行了对比,用单样本 K-S 检验得 到分布拟合度,构建 Q-Q 图来计算平台真实运动 量,结合 YOLOv5 校正检测目标校正目标轨迹,实现 了目标轨迹的提取。

2 红外小目标数据集的训练

实验室拍摄的旋翼无人机天空背景红外视频, 提取其中图片1651张,图像背景均匀,大小为567× 273,其中图像如图1所示,图1(a)中红外小目标的 灰度值较高,与背景的灰度值对比明显,图1(b)中 灰度值比图1(a)明显,不容易观察出来。



图1 单目标简单静止背景数据集图像

Fig. 1 Single target simple static background data set images

选择地/空背景下红外图像弱小飞机目标检测数 据图片^[8]文件夹 data5、data8 共 3396 张,图像大小为 256×256,其中典型图像如图 2 所示,图 2(a)中红外小 目标的灰度高于局部背景区域,图 2(b)中红外小目标 灰度值与局部背景区域灰度接近,不易与背景区分,同 样存在着与小目标灰度和形状接近的小区域。

实验是在安装 Window10 的服务器上完成,服务器

的 CPU:Intel(R) Xeon(R) Silver 4116 CPU @ 2.1 GHz 2.1GHz(2处理器),内存 32GB,显卡 NVIDIA GeForce RTX-3090 32 GB, cuda11.1.0_456.43_win10, cudnn-11.1-windows-x64-v8.0.4.30, Aanconda3 – 2020.02-Windows-x86_64 版本搭建 Pytorch1.7, python 3.7.6, pycharm-community – 2020.1.3×64。



图 2 单目标复杂背景数据集图像 Fig. 2 Single target complex background data set images

将单目标的简单和复杂背景图像使用 labelImg 软件标注,共得到标注图像 5047 张,对应标注 xml 文档 5047 个,区分训练图 4015 张,验证图 1032 张。 用 YOLOv5s、YOLOv5m、YOLOv51、YOLOv5x 四种模 型训练数据 50 次、100 次、200 次、300 次、400 次、 500 次和 600 次,如图 3 所示。



随着训练次数的增加,四种模型的训练准确率随着次数的增加,总体上是在逐渐提高,YOLOv5s、YOLOv5m、YOLOv51、YOLOv5x 训练 600 次的准确率分别为0.9608、0.9485、0.9605、0.977,在准确率上, 震荡程度最大的是 YOLOv5s,收敛最快的是 YOLOv5x。单一模型损失率的随训练次数增加损失 率总体呈现波动降低趋势,训练 600 次的最终损失 率在 0.01581、0.01478、0.0145、0.01408,数值相近, 全部低于 0.02,训练收敛效果较好。

3 运动背景下的单目标检测

3.1 光流算法

光流是图像中亮度保持不变,即运动前后对应 像素亮度值相同,而不随时间而变化可表示为:

I(x + dx, y + dy, t + dt) = I(x, y, t) (1) 其中,I(x, y, t)是图像中各点的像素值;(x, y)表示 空间中的位置;t是时间变量;dt是图像连续帧的间 隔时间;(dx, dv)表示在t时刻空间中的(x, y)点的 位置变化量,在相邻两帧图像中的点(x, y)有位移 (dx, dy),单位时间上的位置变化量就是速度矢量。

用泰勒级数实施一阶展开为:

$$I(x + u, y + u, t + dt) = I(x, y, t) + \frac{\partial I}{\partial x}dx + \frac{\partial I}{\partial y}dy + \frac{\partial I}{\partial t}dt + \varepsilon$$
(2)

$$I_{x} = \frac{\partial I}{\partial x}, I_{x} = \frac{\partial I}{\partial y}, I_{t} = \frac{\partial I}{\partial t}, x \text{ 轴方向速度矢量为u}$$
$$= \frac{dx}{dt}, y \text{ 轴方向速度矢量 } v = \frac{dy}{dt}, \text{ 简化后表示为}:$$

$$I_x u + I_y v + I_t = 0 aga{3}$$

对选择地/空背景下红外图像弱小飞机目标 检测数据中的连续两帧图像进行实验,图像大小 为256×256,如图4所示。图4中两幅图像的光 照条件没有变化,同一区域的灰度值也比较一致。 从两幅图像中地物对比可以看出,图4(b)帧明显 比图4(a)帧靠右,判断平台向右侧运动。



图 4 连续两帧图像 Fig. 4 Two consecutive images

选取 HS 和 LK 计算结果 50 × 150 的部分放大, 显示的结果如图 5 所示。

从选取部分放大后可以看出,HS 算法光流图中 光流值错误的多,且各个方向都有无法看出总体规 律。而 LK 算法光流值总体向右,与平台实际移动 方向较为一致。用 LK 法对这种场景计算运动量比 较合适。



Fig. 5 HS algorithm and LK algorithm results

3.2 光流值的校正

平台真实运动量需要通过取光流值的有效部分 得到,首先用 HS、LK 算法计算光流分量,然后通过 单样本 K-S 检验计算分布拟合度,构建 Q-Q 图得到 上下界数值,最后取上下界间的均值取整得到校 正值。

光流法对光线较为敏感,得到的光流场中存在 一些极端值,为了通过光流场解算出平台移动情况, 用不同光流算法得到的光流场对真实情况进行预 测,因此引入 *t* Location-Scale 分布函数对极端值进 行滤除。

1) t Location-Scale 分布基于 t 分布^[10],其概率

(5)

密度函数为:

$$f(x) = \frac{\Gamma(\frac{\nu+1}{2})}{\sqrt{\nu\pi}\Gamma(\frac{\nu}{2})} \left[1 + \frac{\left(\frac{x-\mu}{\sigma}\right)^2}{\nu}\right]^{-\frac{\nu+1}{2}}$$
(4)

其中, μ 为位置参数; σ 为尺度参数; ν 为自由度。 $\Gamma(x)$ 为伽马函数,相较于t分布,由于引入了位置 参数和尺度参数,t Location-Scale 分布具有更好的 自适应性。

2) 拟合优度检验,单样本 K-S 检验可用于检 验观测到的数据是否符合某已知的理论分布。 Matlab 中使用 kstest 函数实现 K-S 检验功能,返 回参数[*h*,*p*,ksstat,cv],其中,*h* 表示检验决策,*p* 值反映原假设成立的可能性,ksstat 为统计量的 值,cv 为接受域与拒绝域的边界值。使用 LK 算 法得出的光流场的竖直分量数据测试 *t* Location-Scale 分布的拟合程度,对比正态分布的拟合结果 数值,得到的拟合曲线如图 6 所示,参数计算结 果如表 1 所示。



图 6 t Location-Scale 拟合曲线与正态分布拟合曲线 Fig. 6 t Location-Scale fitting curve and normal distribution fitting curve

Tab. 1 K-S inspection result

	t Location-Scale	Normal		
h	1	1		
p	5.859156048772945e-49	3.469833027332950e-309		
ksstat	0. 086829643718679	0.219007809600211		
cv	0.015769137063221	0.015769137063221		

从表 1 中可以看出, t Location-Scale 分布的 ksstat 值远小于正态分布, 且 p 值大正态分布, 说明 tlocation-Scale 分布有着更好的拟合度。 3) Q-Q 图可用于评估两组数据是否来自同一分布,假设大部分数据数据均可用且符合 *t* Location-Scale 分布,用 Quantile-Quantile 图(Q-Q 图)对不可用部分数据进行筛选。

Q-Q 图的横坐标为理想分布的分位数,纵坐标 为样本分位数。因此要构建 Q-Q 图,首先要确定分 位数。

若有一容量为n的样本 $x_1, x_2, x_3, \dots, x_n, 其 p$ 分位数 x_n 的计算公式为:

$$x_p = \begin{cases} x_{([np]+1)} & np \subset Z \\ \frac{1}{2} [x_{(np)} + x_{(np+1)}] & np \not\subset Z \end{cases}$$

可使用 random 函数在理想分布模型中随机 抽取样本得到近似的理想分布样本分位数。结 合 qqplot 函数进行 Q-Q 图的构建,如图 7、图 8 所示。





由 Q-Q 图的特性可知,分布于直线附近的点与 理想 *t* Location-Scale 分布较为符合,可认为该点对 应数据为可用数据。得到 HS 和 LK 算法上下界,计 算上下界均值,得到实际速度如表 2 所示。

表 2 HS、LK 算法的水平和竖直分量上下界 Tab. 2 The upper and lower bounds of the horizontal and vertical components of the HS and LK algorithms

速度分量	下界	上界	真实速度
HS 法水平分量	- 1. 6717	4. 6349	1. 152270555326297
HS 法竖直分量	- 3. 5068	3.4087	-0. 126730718763922
LK 法水平分量	4. 4392	14. 111	9. 541400337785594
LK 法竖直分量	- 0. 37501	0. 44059	- 0. 029798676627955

LK 法得出的水平分量的有效取值范围为 [4.4392,14.111],LK 法得出的竖直分量的有效取 值范围为[-0.37501,0.44059],HS 法得出的水平 分量的有效取值范围为[-1.6717,4.6349],HS 法 得出的竖直分量的有效取值范围为[-3.5068, 3.4087]。对比 HS 和 LK 两种算法运动分量结果, 可以得出 LK 算法水平分量 9.54,竖直分量 -0.03, 水平分量四舍五入为10,竖直分量四舍五入为0,速 度分量结果与观察到的实际运动相符,选用 LK 算 法计算这种场景下的运动。

3.3 实验结果与分析

通常红外图像背景较为复杂,存在类似小目标 的虚假目标许多干扰点。使用训练权重对红外图像 中的小目标检测结果中包括真实的小目标和虚假目 标,光流法可计算得到帧与帧之间的横向和纵向运 动量,从而补偿平台运动的影响,将多帧图像中检测 点表示在一幅图像中,真实的目标轨迹是一条曲线, 而虚假目标点相对背景的位置没有发生变化,通过 光流补偿计算后的位置通常聚集在一小区域,能够 区分真实目标和虚假目标,从而检测出真实的目标。 检测流程如图9所示。





实验中,权重采用 YOLOv5s 训练 200 次的 best 权重,为了把所有可能的小目标都提取出来,需要设 置较低的置信度,将置信度设置为 0.001,检测地空 背景图像数据集 data18 中的 500 帧单个无人机目标 树林运动背景的红外图像,第 1、19、39、55 帧的检测 结果如图 10 所示。

其中,图10(a)图像中的目标是真实目标,从图 10(b)、(c)、(d)图像中可以看出,该目标从向树林 左下方运动,目标与背景位置有明显的相对运动,平 台随运动目标整体向左下移动,背景向左下运动,检 测到的假目标的坐标值发生变化。地面辐射强,图 像上的灰度值较高,形状与红外小目标类似,被检测 为小目标。图 10(b)中检测到1个目标,图 10(c) 检测到1个目标,这些检测到的固定目标的中心坐 标随着平台移动,发生变化了,但相对背景位置没有 发生变化。



图 10 数据集训练 200 次的运动背景检测结果 Fig. 10 Data set training 200 times of motion background detection results

选取其中第1到55的连续55帧图像,用LK方 法求相邻帧图像的横向和纵向光流变化量,从第2 帧到第*n*帧的光流量变化量取整后的累加值,就是 平台运动补偿量,如式(6)所示:

$$\begin{cases} U_n = u_1 + u_2 + \dots + u_n \\ V_n = v_1 + v_2 + \dots + v_n \end{cases}$$
(6)

其中,u,v分别是通过 LK 法计算相邻帧图像得到光 流的水平和垂直方向的取整结果; U_n , V_n 分别是第1 帧图像到第n帧图像间平台水平和垂直方向运动 量。从第1帧到第100帧图像预测点的横纵坐标 x_n , y_n 分别与平台运动补偿量求和,得到图像的中各 点的真实位置,如式(7)所示:

$$\begin{cases} X_n = x_n + U_n \\ Y = y + V \end{cases}$$
(7)

将校正前的坐标点和校正后的坐标点在第1帧 图像的 xy 坐标图上画出,结果如图 11 所示。

55 帧图像中12个目标的点数如表3所示。



Fig. 11 The points before and after complex background correction are expressed in xy coordinates

其中,图11(a)是光流法校正前的图像,目标 1、2、3、5、6、7、8、9、11、12对应的坐标比较集中,目 标1、3、4、6、7、9、12坐标点数是4个以内,是虚假的 目标,目标8、10、11的点数分别是7个、9个、7个, 目标4和目标10坐标变化较大,是位置发生变化的 点,点的轨迹呈与平台运动方向一致变化的曲线,可 判断为虚假的目标。图11(b)是经过光流法校正后 的图像,目标1的坐标轨迹是向左下方运动的曲线, 符合运动目标的运动方向,其他目标经校正后的坐 标都较集中,在第1帧图像中的检测目标附近,是背 景中的虚假目标。

表3 检测目标出现次数

Tab. 3 Number of occurrences of detection target

目标1	目标2	目标3	目标4	目标5	目标6
55	1	3	2	1	2
目标7	目标8	目标9	目标10	目标11	目标 12
2	7	4	9	7	3

4 结 论

文中采用基于 YOLOv5 的深度学习方法对红外 弱小目标提取,并对轨迹进行校正,较好的实现了目 标轨迹的复原。结合拍摄到的简单背景和公开的复 杂背景的数据集在网络中训练,而后对红外小目标 检测,检测准确率在90%以上,损失率在2%以下, 结果比较理想。给出了提出轨迹的流程,对比了两 种光流算法,得到适于推算场景变化量的算法,进行 分布拟合,构建 Q-Q 图得到平台真实运动量,通过 光流与 YOLOv5 相结合的方式校正了目标的运动轨 迹,克服了提出轨迹不正确的问题。

红外弱小目标检测的研究始终是一个热点问题。目标所处的背景千变万化,除了林地和天空背景外,还有云层、海洋、城市和乡村等多种多样的背景,如何从这些复杂的背景中检测到真实的目标有着很大的难度,需要不断加以学习研究。除了背景多变外,红外目标自身有着不同的形状,有的规则的方形,也有的形状是弧形等多种,实现正确的检测是今后研究工作中需要努力的。

参考文献:

 Wei Yuan, Yang Hua, Cheng Zhengdong, et al. Hysteresis threshold segmentation in infrared small target detection
 [J]. Laser & Infrared, 2020, 50(1):111 - 116. (in Chinese)

> 魏元,杨华,程正东,等. 红外小目标检测的滞后阈值 分割法[J]. 激光与红外,2020,50(1):111-116.

- [2] Duan Siwei, Wang Zhonghua, Ye Zheng. An infrared small object detection algorithm based on spatial weighted local contrast [J]. Laser & Infrared, 2020, 50(10):1200 1206. (in Chinese)
 段思韦,王忠华,叶铮. 空域加权局部对比度的红外小目标检测算法[J]. 激光与红外, 2020, 50(10): 1200 1206.
- [3] Qing Sun, Ling Li, Yun Hongxin. Research on the multiscale low rank method and its optimal parameters selection strategy for infrared small target detection [J]. Optik, 2019,192.

- [4] Li Junhong, Zhang Ping, Wang Xiaowei, et al. Infrared small-target detection algorithms: a survey [J]. Journal of Image and Graphics, 2020, 25(9):1739-1753. (in Chinese)
 李俊宏,张萍,王晓玮,等. 红外弱小目标检测算法综述[J].中国图象图形学报, 2020, 25(9):1739-1753.
- [5] Liu Zhijia, Wang Xuan, Zhao Jinbo, et al. An improved method of infrared image target detection based on YOLO algorithm[J]. Laser & Infrared, 2020, 50(12):1512 1520. (in Chinese)
 刘智嘉, 汪璇, 赵金博,等. 基于 YOLO 算法的红外图 像目标检测的改进方法[J]. 激光与红外, 2020, 50 (12):1512 1520.
- [6] Malta Ana, Mendes Mateus, Farinha Torres. Augmented reality maintenance assistant using YOLOv5 [J]. Applied Sciences, 2021, 11(11).
- [7] Wang Fei, Wang Qiu, Ren Jiayi, et al. Substation electrical equipment detection and three-dimensional modeling system[J]. Electrical Measurement and Instrumentation, 2021,58(3):160 167. (in Chinese)
 王菲,王球,任佳依,等. 变电站电气设备检测与三维 建模系统[J]. 电测与仪表,2021,58(3):160 167.
- [8] Hui Bingwei, Song Zhiyong, Fan Hongqi, et al. A dataset for infrared detection and tracking of dim-small aircraft targets under ground/air background[J]. China Scientific Data (Chinese and English online version), 2020,5(3): 291-302. (in Chinese) 回丙伟,宋志勇,范红旗,等.地/空背景下红外图像弱 小飞机目标检测跟踪数据集[J].中国科学数据(中英 文网络版),2020,5(3):291-302.
- [9] Zhang Yanyan. Research of moving target detection based on optical flow algorithm[D]. Xi'an:Xi'an Shiyou University,2018. (in Chinese)
 张艳艳. 基于光流算法的运动目标检测应用研究[D]. 西安:西安石油大学,2018.
- [10] Sadegh Etemad, Maryam Amirmazlaghani. A new multiplicative watermark detector in the contourlet domain using t location-scale distribution[J]. Pattern Recognition, 2018,77.