

文章编号: 2095-2163(2019)03-0171-05

中图分类号: TP391.4

文献标志码: A

复杂条件下小目标检测算法研究

彭小飞, 方志军

(上海工程技术大学 电子电气工程学院, 上海 201620)

摘要: 非海空背景小目标检测是图像处理最具挑战的任务之一。为了解决复杂条件下的小目标检测准确率不足的情况, 本文提出首先运用超分辨率模型对拍摄模糊图像进行重建, 将重建后的清晰图像进行小目标检测。另外, 对原始 FPN 模型进行改进, 利用浅层网络丰富的位置信息, 仅采用三层特征提取网络, 即可完成小目标全图搜索检测。实验表明, 本文方法在清晰图像直接进行重建准确率达到 81.82%, *map* 值为 0.895 1, 重建后的再进行小目标检测与清晰图像直接检测仅有一个未检测出。

关键词: 小目标检测; 超分辨率重建; 浅层; 全图搜索

Research on small target detection algorithm under complex conditions

PENG Xiaofei, FANG Zhijun

(School of Electronic and Electrical Engineering, Shanghai University of Engineering Science, Shanghai 201620, China)

[Abstract] Small target detection is one of the most challenging tasks in image processing. In order to solve the problem of insufficient accuracy of small target detection under complex conditions, this paper proposes to reconstruct the captured blurred image by using the super-resolution model, and to detect the reconstructed clear image for small target. In addition, the original FPN model is improved, and rich location information of the shallow target network is used, and only the three-layer feature extraction network is used to complete the small target full-image search detection. Experiments show that the accuracy of this method is 81.82% and the *map* value is 0.895 1. The reconstructed small target detection and clear image direct detection have only one undetected difference.

[Key words] small target detection; super-resolution reconstruction; shallow layer; full-image search

0 引言

随着视频监控的不断发展,使得硬件设备的性能在迅速提升的同时,监控涉及的领域也在拓展与增加。近年来,运用图像处理的方法对一些复杂的环境进行有目的监控已然成为现代社会保障公众人身安全的一项有益举措。例如,机场开阔区域对小目标的检测识别,远距离大范围监控等。

机场以及学校区域,安全管理至关重要。诸如众所皆知的是,近年来鸟类影响飞机起飞的事件时有发生。迄今为止,学界研究中的二维通用目标检测准确率以及速率均已达到商用的要求。2013年,RCNN^[1]将卷积神经网络运用到目标检测上,而后又相继涌现一系列基于卷积神经网络的目标检测框架。例如,基于区域候选提议的 SPP-Net^[2]、Fast-Net^[3]、Faster-Net^[4]等,此类检测主要是分为 2 个步骤,即:区域候选目标检测和细化打分分类,但在研发上却基本无法达到最基本的实时效果。后期主要基于特征提取网络的端对端目标检测方法,如

SSD^[5]、YOLOv1^[6]、YOLO9000^[7]等,此类方法主要就是将检测和分类融合起来,优点是速度较快,但是对小目标检测在效果上却仍然欠佳。

目前,国内外小目标检测研究主要停留在红外小目标检测^[8-10]、雷达空对地遥感小目标检测^[11-12]等。其研发原理是利用特殊的热传感器来分析景物不同物体温度,将图像转换为灰度图像的灰度值,物体在图像中的灰度值大小与物体温度有关。基于此,红外小目标检测即使用于诸如丛林等可见光很弱的恶劣条件下也能获得较强的辨别能力。但是现在研究指出,这种检测方法在通用的视频监控领域中对通用目标识别能力很差,只能大致分析出物体所在区域,并不能准确判断出属于哪类物体。

综上分析后可知,本文在 FPN^[13]的基础上提出利用浅层网络特征对复杂条件下的小目标进行检测。这里的复杂条件可描述为:场景内存在模糊图像,此时先要对模糊图像进行超分辨率重建;背景较复杂,非海空纯背景,有建筑物干扰;基于全图搜索

作者简介: 彭小飞(1994-),男,硕士研究生,主要研究方向:机器视觉。

收稿日期: 2019-01-16

小目标,目标很小,绝对像素大约为 $24 * 24$ 。

1 算法原理

1.1 超分辨率重建算法

本文采用的超分辨率重建算法是基于 SRN^[14] 模型,其研发设计主要源起自编码器-解码器思想。相应地,编码器是通过卷积神经网络对模糊图像特征进行提取并映射到一个矩阵空间,解码器是编码器反过程,就是模糊图像通过寻找映射空间中相似特征块进行重建。但在本次研究中,设计时在超分辨率模型中加入了多尺度训练模型,这样一来在使用不同尺度图像训练模型过程中则可以综合提取不同尺度模糊图像特征细节,由此将使最终超分辨率效果能有一定提升。并且,还可以通过在不同尺度训练网络的环节中做到权重共享,同时大大减少训练时间。另外,本文方法将循环神经网络(LSTM)应用到超分辨率重建训练过程,这种做法的好处就是在训练过程中即可以利用 LSTM 的记忆功能,从而不断优化参数,提升训练效率以及超分辨率重建效果。重建模型如图 1 所示。

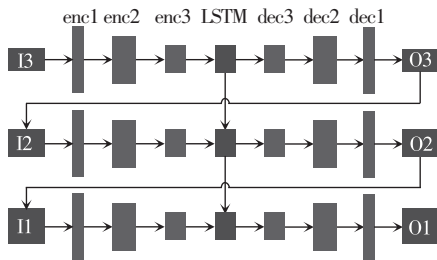


图 1 超分辨率重建模型

Fig. 1 Super-resolution reconstruction model

本文模型总共包含 3 个尺度,每个尺度以一张模糊图像和一张上采样的去模糊图像作为输入,并且用 ConvLSTM 循环神经网络来求得时间相关性和空间相关性。本文方法为了解决直接使用编码器带来的层数少导致感受野小、层数多导致空间信息不充分的问题,将残差块用于编码器网络,而且又采用跳跃连接就可以极大限度利用不同层的特征,且有利于梯度传播和模型加速收敛。

此模型可划分为 3 个部分,即:编码器模块、LSTM 模块、解码器模块。设计上,编码器由 3 层卷积块组成。具体来说,第一层卷积块包括 conv1_1、conv1_2、conv1_3、conv1_4,其中 conv1_1 输出 feature map 大小为 $32 * 32$,卷积核大小为 $5 * 5$,采用默认卷积步长为 1;conv1_2、conv1_3、conv1_4 采用的是 restnet 模块,输出 feature map 大小为 $32 *$

32 ,卷积核大小为 $5 * 5$,采用默认卷积步长为 1。第二层卷积块包括 conv2_1、conv2_2、conv2_3、conv2_4,其中 conv2_1 输出 feature map 大小为 $64 * 64$,卷积核大小为 $5 * 5$,步长为 2;conv2_2、conv2_3、conv2_4 同理采用的是 restnet 模块,输出 feature map 大小为 $64 * 64$,卷积核大小为 $5 * 5$;第三层卷积块包括 conv3_1、conv3_2、conv3_3、conv3_4,其中 conv3_1 输出 feature map 大小为 $128 * 128$,卷积核大小为 $5 * 5$,步长为 2;conv3_2、conv3_3、conv3_4 同理采用的是 restnet 模块,输出 feature map 大小为 $128 * 128$,卷积核大小为 $5 * 5$ 。

LSTM 模块采用的是 convLSTM, cell 尺寸为 $h/4 * w/4$,卷积核大小为 $3 * 3$,feature map 大小为 $128 * 128$ 。采用 convLSTM 模块,不仅具备 LSTM 能够得到时序关系,还能提取空间特征。

解码器模块与编码器模块对称,由一次卷积块和两层反卷积块组成,主要用于对编码模块数据进行解析。超分辨率重建设计流程如图 2 所示。

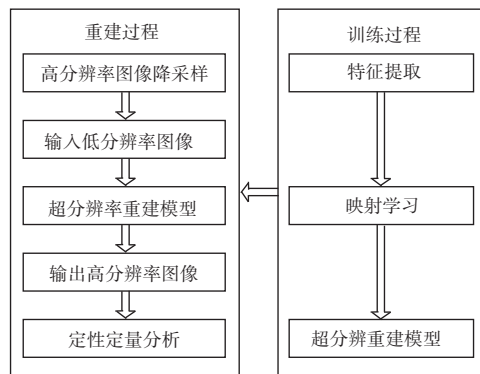


图 2 超分辨率重建流程

Fig. 2 Super-resolution reconstruction process

1.2 小目标检测算法

本文研究中,采用了改进的 FPN 特征金字塔网络对小目标进行检测。经过探索讨论后可知,浅层网络特征对于小目标检测更加有效,而且也将具有更为丰富的位置信息,如果将深层网络特征反卷积与浅层特征相融合却会对小目标检测起到负面作用。本文方法中,需要输入整张图像,这是为了缓解图像缩放导致的“绝对尺寸”缩小而随之出现无法检测的问题。在此基础上,通过实验验证后得知,对于本文单小目标进行检测,仅利用 3 层浅层网络对目标进行特征提取更有效,加入高层网络反传特征进行叠加,反而会影响最终的结果。究其原因则在于本文小目标绝对像素大小约为 $24 * 24$,经过第一个 Pooling 层之后,小目标绝对像素大小约为 $12 * 12$;经过第二个 Pooling 层之后,小目标绝对像素大

小约为 $6 * 6$ 。研究得到的小目标检测模型即如图 3 所示。

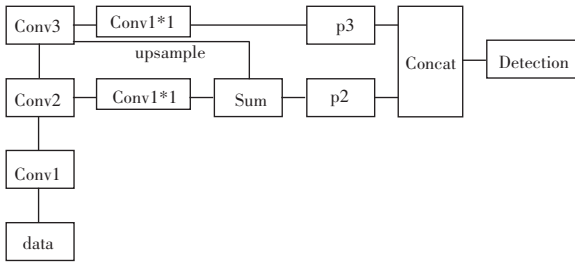


图 3 小目标检测模型

Fig. 3 Small target detection model

本文模型采用 Faster-RCNN 基本思想,对小目标进行检测。输入图像为整张图,如此则可以扩大相对像素大小,如 YOLO、SSD 输入首先要将图像变换到小于原始图像大小,这样就会进一步缩小原始目标图像,不利于小目标的特征提取。输入图像仅仅经过 3 层的 CNN 网络即可对小目标特征进行提取,其间采用了 ResNet 跳跃连接的思想。就设计整体而言,第一层包括:卷积层 Conv1,卷积核大小为 $5 * 5$, num_output 为 64, $stride$ 为 2,采用 BatchNorm 层和 Scale 层对卷积层输出进行归一化处理, ReLU 激活, pooling 层采用 Max pooling,核大小为 $3 * 3$,步长为 2。第二层包括:卷积层 res2a_branch1,卷积核大小为 $1 * 1$, num_output 为 256,采用 BatchNorm 层和 Scale 层对卷积层输出进行归一化处理, ReLU 激活;卷积层 res2a_branch2a,卷积核大小为 $1 * 1$, num_output 为 64,采用 BatchNorm 层和 Scale 层对卷积层输出进行归一化处理, ReLU 激活;卷积层 res2a_branch2b 连接 res2a_branch2a 的输出,卷积核大小为 $3 * 3$, num_output 为 64,采用 BatchNorm 层和 Scale 层对卷积层输出进行归一化处理;卷积层 res2a_branch2c 连接 res2a_branch2b 的输出,卷积核大小为 $1 * 1$, num_output 为 256,采用 BatchNorm 层和 Scale 层对卷积层输出进行归一化处理;res2a_branch2c 的输出和 res2a_branch1 的输出采用 Eltwise 连接输出得到 res2a,经过 ReLU 进行统一激活;res2a 输出作为 res2b 输入,具体参数设置如 res2b_branch1, res2b_branch2a, res2b_branch2b, res2b_branch2c 与 res2a 各模块类似;res2c 具体参数与上述类似;同理, res3a、res3b、res3c、res3d 参数设置与上述类似,不同的是, res3 层提取的特征经过上采样与 res2 进行特征融合,构成特征金字塔模型,最后利用 res2 和 res3 融合的特征对小目标检测。

本文采用 anchor 尺寸设置为 $[16, 16]$ 和 $[32,$

$32]$,分别对应 3 种 ratios $[0.5, 1, 2]$ 。原始相应层 anchor 尺寸设置为 $[64, 64]$ 和 $[128, 128]$,为了适应本文小目标检测任务,将对应 anchor 尺寸缩小,实验证明,此方法对于本文应用场景小目标检测具有很好的效果。

2 实验与分析

2.1 实验平台

实验基于 Ubuntu16.04,64 位操作系统,超分辨率模型训练平台为 Python2.7 和 Tensorflow,小目标检测模型训练平台为 Python2.7 和 caffe,硬件配置为 GTX1080Ti。

2.2 数据集制作

本文训练数据集为无人机小目标,背景为天空、教学楼以及不确定物体,为了使网络有更好的鲁棒性,训练数据采用隔帧获取,带小目标的清晰数据集为 863 张。模糊数据集采用方框滤波、均值滤波、高斯滤波三种线性滤波的方式和中值滤波、双边滤波两种非线性滤波的方式进行模糊处理,最终可得到 21 575 张模糊图像。小目标检测数据集在标注上采用了 label-image 开源标注工具,同时为了达到深度学习大数据训练要求,对标注图像进行扩充。主要扩充方式,包括将图像顺时针旋转 60° 、 90° 、 120° 、 150° 、 180° 、 210° 、 240° 、 270° 、 300° 、 330° ,并且将标注信息随着图像旋转,通过这种方法得到扩充图像为 8 300 张。另外,本文还将随机剪裁图像,此方法参考 SSD 数据扩充的方法,所剪裁的区块大小为原图大小的 0.9,经过 10 次随机剪裁,利用该方法得到的扩充数据集为 8 300 张,而且标注数据也将随着剪裁一起变化。

2.3 结果与分析

为了证明本文方法的有效性,将本超分辨率重建算法与经典超分辨率重建算法 SRCNN^[15]、FSRCNN^[16]、ESPCN^[17] 进行 3 倍、4 倍、5 倍放大情况下的研究对比,主要评价标准参考结构相似性 (Structural Similarity Index Measurement, SSIM)、峰值信噪比 (Peak Signal to Noise Ratio, PSNR) 两种指标。4 种不同对比实验结果详见表 1 和表 2。

分析可知,模糊图像对于小目标检测任务影响巨大,故而选取一个优质超分辨率重建模型对于小目标检测任务将尤为关键。由表 1 和表 2 可知,SRN 在同等条件要优于其它 3 种超分辨率算法,更符合此应用场景。

本文小目标检测对比实验分为 2 个部分,对此

阐释如下。

(1) 将本文模型和经典目标检测框架 FPN、YOLOv2 及 SSD 进行准确率和平均精度 map 值对比。

(2) 经过超分辨率重建和未经过超分辨率再进行小目标检测准确率对比。

表 1 数据集在 3 倍、4 倍、5 倍放大条件下的平均 SSIM

Tab. 1 Average SSIM For data sets at 3x, 4x, and 5x magnification

放大倍数	SRCNN	FSRCNN	ESPCN	SRN
3	0.512 0	0.509 0	0.737 5	0.971 8
4	0.473 0	0.475 0	0.746 8	0.956 6
5	0.439 0	0.458 0	0.753 6	0.972 1

表 2 数据集在 3 倍、4 倍、5 倍放大条件下的平均 PSNR

Tab. 2 Average PSNR For data sets at 3x, 4x, and 5x magnification

放大倍数	SRCNN	FSRCNN	ESPCN	SRN
3	37.549 9	38.335 9	40.702 7	52.010 9
4	36.470 9	36.470 9	39.336 9	46.522 7
5	34.788 8	34.856 7	37.838 0	45.567 3

在未参与训练的数据中随机抽取 11 张包含目标的数据集进行实验,实验结果见表 3 和表 4。

表 3 FPN、SSD、YOLOv2、FPN³ 准确率和 map 值对比实验

Tab. 3 FPN、SSD、YOLOv2、FPN³ accuracy and map value comparison experiment

算法	准确率/%	$map@0.5$	算法	准确率/%	$map@0.5$
FPN	54.55	0.673 5	YOLOv2	0	0
SSD	27.27	0.513 7	FPN ³	81.82	0.895 1

表 4 超分辨率重建小目标检测对比实验

Tab. 4 Super-resolution reconstruction small target detection contrast experiment

	模糊图像	超分辨率重建后图像	清晰图像
准确率	18.18	72.73	81.82

由表 3 可以得出,在此场景下,本文算法 FPN³ 具有更高的准确率,以及平均精度 map 值优于 YOLO、SSD、FPN 框架。

通过分析可得,FPN 原始模型运用的特征提取网络为 RestNet50,本文小目标绝对像素大小为 $24 * 24$ 左右,经过 3 层 Pooling 层操作之后,像素大小减小为 $3 * 3$,基本无目标特征,如果经过上采样和低层特征进行融合,反而会最终检测结果;相对于 YOLOv2 检测,SSD 准确率会高一些,这是因为 SSD 利用了多尺度特征图的思想对小目标进行预测,而 YOLOv2 采用全局对目标进行预测。但是 SSD、

YOLOv2 对于小目标检测效果却仍有待改进,追根溯源皆是因为两者均未使用低层高分辨率的位置信息,而只是在网络层最后一层做出预测。因此参考借鉴如上研究方案后,本文研究则采用 3 层网络对小目标进行预测,极大提高运算效率与准确率。

由表 4 分析后知道,图像模糊到一定程度时,基本无法进行小目标检测。经过超分辨率重建之后,准确率方面和清晰图像小目标检测准确率基本相当。故可以推论得出如下研究结论:经过对模糊图像超分辨率重建后进行小目标检测,具有一定研究和应用价值。本文的总体结果流程如图 4 所示。



(a) 模糊图像
(a) Blurred image



(b) 超分辨率重建后图像
(b) Super-resolution reconstructed image



(c) 小目标检测结果图
(c) Small target test result chart

图 4 整体结果流程图

Fig. 4 Overall result flow chart

3 结束语

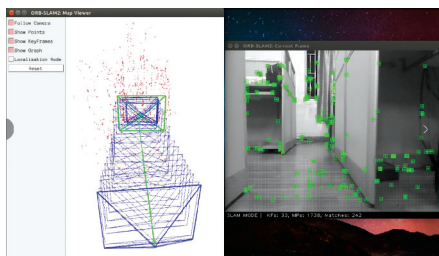
本文分析了时下目标检测以及小目标检测方法的不足,并基于此展开了在复杂条件下的小目标检测研究。首先运用超分辨率重建方法对模糊图像进

行重建,而后再将重建后清晰图像输入到小目标检测模型中进行小目标检测。不同于传统红外等灰度图像小目标检测的是,本文选择图像具有复杂背景,以及模糊图像干扰,且目标足够小,达到绝对像素大小为 24×24 。其清晰图像直接进行小目标检测准确率达到81.82%,*map*值达到0.895 1,进行超分辨率重建后、再进行小目标检测的准确率最终达到了72.73%。本文方法设计新颖,具有一定研究价值,但是目前仍处于2个独立阶段,后续还需将2个步骤予以系统整合。

参考文献

- [1] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation [C]// 2014 IEEE Conference on Computer Vision & Pattern Recognition. Columbus, OH, USA;IEEE, 2014:580-587.
- [2] HE Kaiming, ZHANG Xiangyu, REN Shaoqing, et al. Spatial pyramid pooling in deep convolutional networks for visual recognition[M]//FLEET D, et al. ECCV 2014, Part III, LNCS 8691.Switzerland;Springer International Publishing,2014:346-361.
- [3] GIRSHICK R. Fast R-CNN[J]. arXiv preprint arXiv:1504.08083, 2015.
- [4] REN Shaoqing, HE Kaiming, GIRSHICK R, et al. Faster R-CNN: Towards real-time object detection with region proposal networks[C]// IEEE Transactions on Pattern Analysis & Machine Intelligence,2017,39(6):1137-1149.
- [5] LIU Wei, ANGUELOV D, ERHAN D, et al. SSD: Single shot multibox detector[J]. arXiv preprint arXiv:1512.02325,2015.
- [6] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: Unified, real-time object detection [J]. arXiv preprint arXiv:1506.02640, 2015.
- [7] REDMON J, FARHADI A. YOLO9000: Better, faster, stronger [C]// 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu, Hawaii, USA;IEEE,2017:6517-6525.
- [8] 朱国强, 孟祥勇, 钱惟贤. 基于曲率的近地面红外小目标检测算法[J]. 光子学报, 2018, 47(10):1010001(1-12).
- [9] 胡洪涛, 敬忠良, 胡士强. 基于辅助粒子滤波的红外小目标检测前跟踪算法[J]. 控制与决策, 2005, 20(11):1208-1211.
- [10] 王军, 姜志, 孙慧婷, 等. 基于噪声方差估计的红外弱小目标检测与跟踪方法[J]. 光子学报, 2018,29(3):305-313.
- [11] 于晓涵, 陈小龙, 陈宝欣, 等. 快速高分辨稀疏 FRFT 雷达机动目标检测方法[J]. 光电工程, 2018,45(6):170702(1-7).
- [12] 李东, 赵婷, 宋伟, 等. 一种低信噪比下稳健的 ISAR 平动补偿方法[J]. 电子学报, 2018, 46(9):2049-2056.
- [13] LIN T Y, DOLLAR P, GIRSHICK R, et al. Feature pyramid networks for object detection [C]// 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu, Hawaii, USA;IEEE Computer Society, 2017:936-944.
- [14] TAO Xin, GAO Hongyun, WANG Yi, et al. Scale-recurrent network for deep image deblurring[J]. arXiv preprint arXiv:1802.01770, 2018.
- [15] DONG Chao, LOY C C, HE Kaiming, et al. Learning a deep convolutional network for image super-resolution [M]//FLEET D, et al. ECCV 2014, Part IV, LNCS 8692.Switzerland;Springer International Publishing,2014:184-199.
- [16] DONG Chao, LOY C C, TANG Xiaoou. Accelerating the super-resolution convolutional neural network [J]. arXiv preprint arXiv:1608.00369, 2016.
- [17] SHI Wenzhi, CABALLERO J, HUSZÁR F, et al. Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network [C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR).Las Vegas, NV, USA;IEEE, 2016:1874-1883.

(上接第170页)



(a) 室内运行地图 (b) 特征提取点
(a) Map of indoor operation (b) Feature extraction point

图4 建图效果

Fig. 4 Mapping effect

4 结束语

基于RGB-D相机的移动机器人以色彩图为图像信息进行ORB特征提取与匹配,并且提取速度快,匹配效果好,为下一步位姿估计提供了具有重要

价值的有效信息。经过重投影误差的计算,能够求出变化矩阵,也能计算得出机器人的位姿,最终以ORB-SLAM算法模拟核电设施室内环境实现了定位与地图构建。本文的研究可以解决在未知的室内环境下移动机器人定位问题,为智能化设备制造提供有益的技术支持与借鉴。

参考文献

- [1] 罗荣华,洪炳镛. 移动机器人同时定位与地图创建研究进展[J]. 机器人,2004,26(2):182-186.
- [2] 高翔,张涛. 视觉SLAM十四讲:从理论到实践[M]. 北京:电子工业出版社,2017.
- [3] MUR-ARTAL R, MONTIEL J M M, TARDOS J D. ORB-SLAM: A versatile and accurate monocular SLAM system[J]. IEEE Transactions on Robotics,2015,31(5):1147-1163.
- [4] RUBLEE E, RABAUD V, KONOLIGE K, et al. ORB: An efficient alternative to SIFT or SURF[C]//IEEE International Conference on Computer Vision, ICCV 2011. Barcelona, Spain;IEEE,2011:1-9.