



改进Center-Net网络的自主喷涂机器人室内窗户检测

洪恺临, 曹江涛, 姬晓飞

引用本文:

洪恺临, 曹江涛, 姬晓飞. 改进Center-Net网络的自主喷涂机器人室内窗户检测[J]. 智能系统学报, 2021, 16(3): 425–432.

HONG Kailin, CAO Jiangtao, JI Xiaofei. Indoor window detection of autonomous spraying robot based on improved CenterNet network[J]. *CAAI Transactions on Intelligent Systems*, 2021, 16(3): 425–432.

在线阅读 View online: <https://dx.doi.org/10.11992/tis.202005016>

您可能感兴趣的其他文章

基于注意力机制的显著性目标检测方法

Salient object detection method based on the attention mechanism

智能系统学报. 2020, 15(5): 956–963 <https://dx.doi.org/10.11992/tis.201903001>

注意力机制和Faster RCNN相结合的绝缘子识别

Insulator recognition based on attention mechanism and Faster RCNN

智能系统学报. 2020, 15(1): 92–98 <https://dx.doi.org/10.11992/tis.201907023>

基于改进的Faster R-CNN高压线缆目标检测方法

Object detection of high-voltage cable based on improved Faster R-CNN

智能系统学报. 2019, 14(4): 627–634 <https://dx.doi.org/10.11992/tis.201905026>

面向多机器人动态任务分配的事件驱动免疫网络算法

Event-driven immune network algorithm for multi-robot dynamic task allocation

智能系统学报. 2018, 13(6): 952–958 <https://dx.doi.org/10.11992/tis.201707022>

基于Kinect的改进移动机器人视觉SLAM

Improved V-SLAM for mobile robots based on Kinect

智能系统学报. 2018, 13(5): 734–740 <https://dx.doi.org/10.11992/tis.201705018>

基于图优化的移动机器人视觉SLAM

Visual-SLAM for mobile robot based on graph optimization

智能系统学报. 2018, 13(2): 290–295 <https://dx.doi.org/10.11992/tis.201612004>



关注微信公众号, 获取更多资讯信息

DOI: 10.11992/tis.202005016

改进 Center-Net 网络的自主喷涂机器人室内窗户检测

洪恺临¹, 曹江涛¹, 姬晓飞²

(1. 辽宁石油化工大学 信息与控制工程学院, 辽宁 抚顺 113001; 2. 沈阳航空航天大学 自动化学院, 辽宁 沈阳 110136)

摘要: 室内自主喷涂机器人可以实现室内墙面喷涂的自动化以此提升喷涂的效率, 减少人力物力的投入。而基于计算机视觉的室内窗户检测算法则是该机器人的关键技术。对于室内窗户检测, 由于环境光照、窗户形状和窗户透光属性的存在, 传统方法无法得到较好的效果。针对此问题, 设计一种基于深度学习的室内窗户检测算法。该算法主要对中心点网络 (CenterNet) 的特征提取网络进行修改, 减少部分卷积操作, 使用 Ghost 模块替换原始的卷积模块, 降低特征冗余, 并引入注意力机制, 让网络尽可能表达重要信息。实验结果表明, 改进的 CenterNet 在不损失网络精度的前提下, 大幅度提高了网络的运算速度, 使得该检测算法即使在机器人端的嵌入式系统上也可以达到实时检测的效果。

关键词: 喷涂机器人; 深度学习; 目标检测; 室内窗户检测; 中心点网络; Ghost 模块; 注意力机制; 嵌入式设备
中图分类号: TP391.1 **文献标志码:** A **文章编号:** 1673-4785(2021)03-0425-08

中文引用格式: 洪恺临, 曹江涛, 姬晓飞. 改进 Center-Net 网络的自主喷涂机器人室内窗户检测 [J]. 智能系统学报, 2021, 16(3): 425-432.

英文引用格式: HONG Kailin, CAO Jiangtao, JI Xiaofei. Indoor window detection of autonomous spraying robot based on improved CenterNet network[J]. CAAI transactions on intelligent systems, 2021, 16(3): 425-432.

Indoor window detection of autonomous spraying robot based on improved CenterNet network

HONG Kailin¹, CAO Jiangtao¹, JI Xiaofei²

(1. School of Information and Control Engineering, Liaoning Shihua University, Fushun 113001, China; 2. School of Automation, Shenyang Aerospace University, Shenyang 110136, China)

Abstract: An indoor autonomous spraying robot can realize the automation of indoor wall spraying to improve the efficiency of spraying and reduce the investment of manpower and material resources. The indoor window detection algorithm based on computer vision is the key technology of the robot. For indoor window detection, traditional methods cannot obtain good results owing to the actual scene's requirements for recognition speed and accuracy as well as the presence of lighting in the environment, shape of the window, and light transmission properties of the window. To solve this problem, an indoor window detection algorithm based on deep learning is designed. This algorithm mainly modifies the backbone feature extraction of the CenterNet network, reduces part of the convolution operation, replaces the original convolution module with ghost block, reduces the redundancy feature, and introduces an attention mechanism to keep the network under a limited number of parameters that express important information as much as possible. The experimental results show that the improved CenterNet algorithm greatly improves the operation speed of the network without losing the accuracy of the network so that the network can achieve a real-time detection effect even on the embedded system of the robot.

Keywords: spraying robot; deep learning; target detection; indoor window detection; Center-Net; Ghost block; attention mechanism; embedded device

随着城镇化水平的不断提高, 室内装修领域的自动化发展水平受到越来越多人的关注。目前

大部分的室内喷涂工作仍然是以人工喷涂为主, 喷涂的效率低且质量难以保证。现有的室内喷涂机器人虽然可以进行简单的墙面喷涂, 但是都缺少环境中不可喷涂区域的识别, 如果希望喷涂机器人真正地做到自主喷涂, 那么对于窗户的检测

收稿日期: 2020-05-12.

基金项目: 国家自然科学基金项目 (61673199); 辽宁省科技公益研究基金项目 (2016002006).

通信作者: 姬晓飞. E-mail: jixiaofei7804@126.com.

就是必不可少的。室内窗户的检测有如下几个难点: 1) 光照强度变化使得算法的准确性很容易受到影响; 2) 窗户透光属性的存在导致窗户外面的环境对窗户检测存在很大的干扰; 3) 喷涂过程中距离墙面很近时只能捕捉到窗户局部的影像。正是因为上述问题的存在使得室内窗户检测变得更加复杂。与此同时, 在实际的机器人开发过程中, 机器人的成本、功耗都要在考虑的范围内。因此应用于喷涂机器人上的窗户检测算法必须要保证所使用传感器和处理器的成本与功耗符合实际的开发需求。因此开发一种实时高效的室内窗户检测算法对机器人的落地具有十分重要的意义。

目前常用的窗户检测数据包括 3D 点云数据^[1]和图像数据^[2-4], 其中 3D 点云数据的获取需要搭载造价较为昂贵的双目相机或深度相机, 考虑到图像数据采集成本较低, 因此基于图像数据的研究更具有可行性。Ali 等^[2]使用多尺度小波对采集图像进行表示, 再将其输入到级联分类器中进行窗户检测; 孔倩倩等^[3]通过对窗户进行轮廓分析筛选出符合条件的轮廓, 并利用最小生成树对所有矩形进行合并以此确定窗户的位置; 缪君等^[4]结合窗户的直线特征、颜色特征和纹理特征使用分割技术提取完整的窗户。总体来说, 目前的方法都是基于手工设计的特征进行室内窗户检测, 该种方式受环境的影响较大, 无法在喷涂机器人上得到很好的应用。

而基于深度学习的方法因其适应性强, 易于转换等特点已被应用到各个领域。针对目标检测领域的深度学习方法无论是两阶段的 Fast R-CNN(fast region-based convolutional neural networks)^[5]、Faster R-CNN(faster region-based convolutional neural networks)^[6]、SPP-Net(spatial pyramid pooling networks)^[7], 还是单阶段的 YOLO(you only look once)^[8]、SSD(single shot multi-box detector)^[9]、Retina-Net^[10], 都先预设大量不同尺度和纵横比的锚框(anchor), 再通过非极大值抑制(non-

maximum suppression, NMS) 的方式进行筛选得到最终的检测框。但是对于室内窗户检测而言, 机器人所采集到的图片中窗户的数量比较少, 会导致大量的 anchor 被 NMS 去除掉, 使得网络效率不高。而无锚框(Anchor-free)^[11]类型的网络则通过在特征图中直接检测目标的关键点, 并回归出目标所在的位置, 极大地提升了网络的效率。Anchor-free 类型具有代表性的网络包括 Corner-Net^[12]、CornerNet-Lite^[13]、Center-Net^[14]、FCOS(fully convolutional one-stage object detection)^[15]等。其中, Center-Net 算法在识别速度和准确性上均具有非常好的效果, 但是由于其复杂特征提取网络的存在, 很难应用于机器人所搭载的嵌入式设备上。

本文设计一种基于深度学习的室内窗户检测算法, 引入 Anchor-free 类型的 Center-Net 网络作为基础网络, 并对其网络结构进行改进, 在不损失精度的情况下提升其检测速度, 减少网络的参数量, 使其适用于喷涂机器人窗户检测的特定应用。本文的创新点是简化 Center-Net 原始特征提取网络的结构, 在提升速度的同时也防止由于参数量过多而导致过拟合的问题; 将普通的卷积操作替换成更加高效的 Ghost 模块^[16], 以此进一步提升网络推理速度; 在编码结构的最后一层引入注意力机制, 通过添加压缩激励模块(squeeze and excitation block, SE-block)^[17]来保证网络在有限的参数量下尽可能表达重要信息。实验结果表明, 该网络即使在嵌入式设备上也可以满足其准确性和实时性的需求。

1 Center-Net 目标检测算法

Center-Net 是 Zhou 等于 2019 年 4 月提出的 Anchor-free 目标检测网络。算法框架如图 1 所示, 将待检测目标中心视为关键点, 然后再分别对目标的长宽以及中心点的偏移量进行回归以此完成目标检测任务。

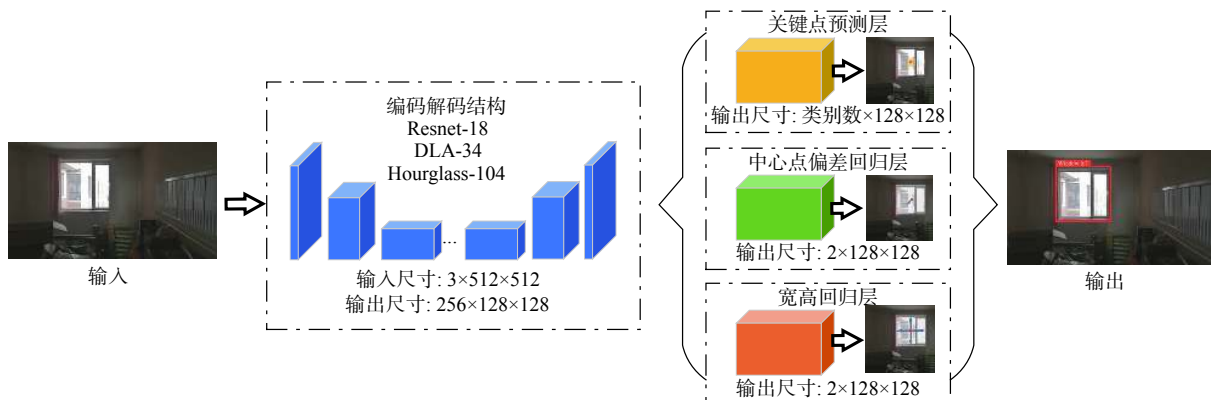


图1 Center-Net 算法框架

Fig. 1 Center-Net algorithm framework

1.1 Center-Net 特征提取网络

原始的 Center-Net 分别使用 Resnet-18(residual networks-18)^[18]、DLA-34(deep layer aggregation-34) 和 Hourglass-104^[19] 编码解码 (encoder-decoder) 结构作为特征提取网络, 实验结果表明, DLA-34 网络在速度和准确性上得到了很好的平衡。因此本文窗户检测算法也是在 DLA-34 网络基础上进行的改进。

DLA 是 Yu 等^[20] 于 2019 年提出的一种特征聚合的方式, 分为迭代深层聚合 (iterative deep aggregation, IDA) 和层次深度聚合 (hierarchical deep aggregation, HDA)。Center-Net 特征提取网络还对原始 DLA-34 进行了一些改进, 网络框架如图 2

所示, 图中绿色箭头对应 IDA, 红色方框对应 HDA。HDA 的作用是将数据的浅层与深层表达进行更好地融合, 这种融合方式比简单的残差网络 (ResNet) 和稠密网络 (DenseNet)^[21] 更好。而 IDA 的作用则是通过不断迭代将浅层特征重新定义从而完善最终输出的特征。网络的编码阶段就是将 IDA 与 HDA 相结合。解码部分则是对相邻的 HDA 结构的输出进行合并, 在合并的过程中将小分辨率的特征使用逆卷积 (deconv)^[22] 操作进行上采样并将得到的结果与相邻 HDA 的输出进行可形变卷积 (deformconv)^[23], 最后再相加, 以该种方式迭代生成树状结构, 最终得到 $256 \times 128 \times 128$ 的输出。

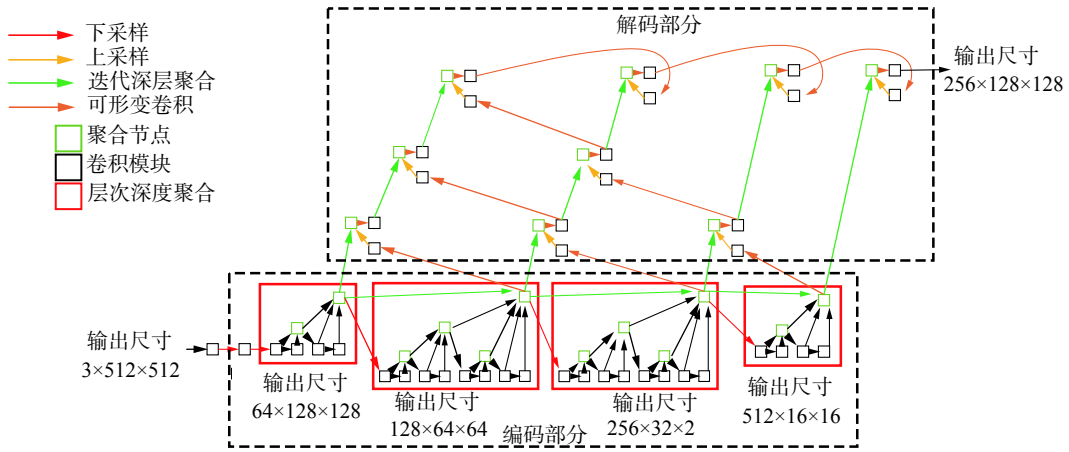


图 2 Center-Net 中的 DLA-34 算法框架

Fig. 2 DLA-34 algorithm framework in Center-Net

1.2 Center-Net 检测框回归网络

将特征提取网络的结果分别送入 3 个网络分支中, 分别是关键点预测层、中心点偏移量回归层以及检测框宽高回归层。

其中关键点预测层的输出为 $80 \times 128 \times 128$, 128 对应输出特征的长宽, 而 80 则对应 COCO 数据集的 80 类目标。首先将关键点通过高斯核分布到 128×128 分辨率的热图 (heatmap) 上, 再将热图与网络的输出逐点计算损失函数。该损失函数为了解决正负样本不平衡的问题, 使用了 Focal Loss 损失函数^[10] 的思想, 其公式为

$$L_k = \frac{-1}{N} \sum_{\text{yyc}} \left\{ \begin{array}{l} (1 - \hat{Y}_{\text{yyc}})^\alpha \lg(\hat{Y}_{\text{yyc}}), \quad \hat{Y}_{\text{yyc}} = 1 \\ (1 - Y_{\text{yyc}})^\beta (\hat{Y}_{\text{yyc}})^\alpha \lg(1 - \hat{Y}_{\text{yyc}}), \quad \text{其他} \end{array} \right.$$

式中: N 是图像中待检测目标的数量; α 和 β 是 Focal Loss 的超参数; Y_{yyc} 是真实框映射到 heatmap 之后的各点像素值; \hat{Y}_{yyc} 是网络输出结果的像素值。

目标框长宽和目标框的中心点坐标偏移量估计层的输出都是 $2 \times 128 \times 128$ 。中心点坐标偏移量估计层存在的意义是由于该网络中对图像进行

了 4 次下采样, 所以特征图重新映射到原始图像上的时候会带来精度损失, 因此对于每一个中心点, 都额外用一个偏移量去补偿它。这两个层的损失函数都使用 L_1 损失函数:

$$L_{\text{off}} = \frac{1}{N} \sum_p \left| \hat{o}_{\bar{p}} - \left(\frac{p}{R} - \bar{p} \right) \right|$$

$$L_{\text{size}} = \frac{1}{N} \sum_{k=1}^N |\hat{s}_{p_k} - s_k|$$

式中: R 对应下采样的次数; $\hat{o}_{\bar{p}}$ 是网络预测的偏移量; \bar{p} 是对实际框下采样之后进行取整。目标框长宽的回归也类似, 其中 \hat{s}_{p_k} 为回归后的尺寸, s_k 为原图尺寸。

在网络的测试阶段, 将测试图片送入到网络得到 3 个分支的输出之后, 由于该网络没有使用 anchor 所以在进行多余检测框剔除的时候不使用 NMS, 而是使用一个 3×3 的最大池化操作 (max-pool), 以此抑制关键点预测输出中极大值的 8 邻域, 起到去掉多余框的作用。最后将得到的输出进行解码得到检测框。

2 改进的 Center-Net 算法

Center-Net 网络虽然在 MSCOCO、PASCAL VOC 等数据集上表现良好,但是被应用于喷涂机器人上的目标检测算法在考虑其准确性的同时还需要考虑算法的实时性,所以网络的大小、参数量以及网络的运行时间在算法的评估层面都需要被考虑。本文在原有 Center-Net 的基础上进行优化改进,分别为:对特征提取网络结构进行简化;将网络中的普通卷积模块替换为高性能的

Ghost 模块;在编码阶段的最后一层添加 SE 模块。本文通过上述 3 种方式提升网络的速度与准确性。

2.1 网络结构的简化

在 COCO 数据集中待检测物体一共有 80 个类别,并且各类目标的形态尺寸变化较大,但是对于室内窗户检测这种单分类物体检测来说可以对网络进行一些简化,在提升速度的同时还可以防止因存在大量冗余的特征而导致网络过拟合。改进的特征提取网络如图 3 所示。

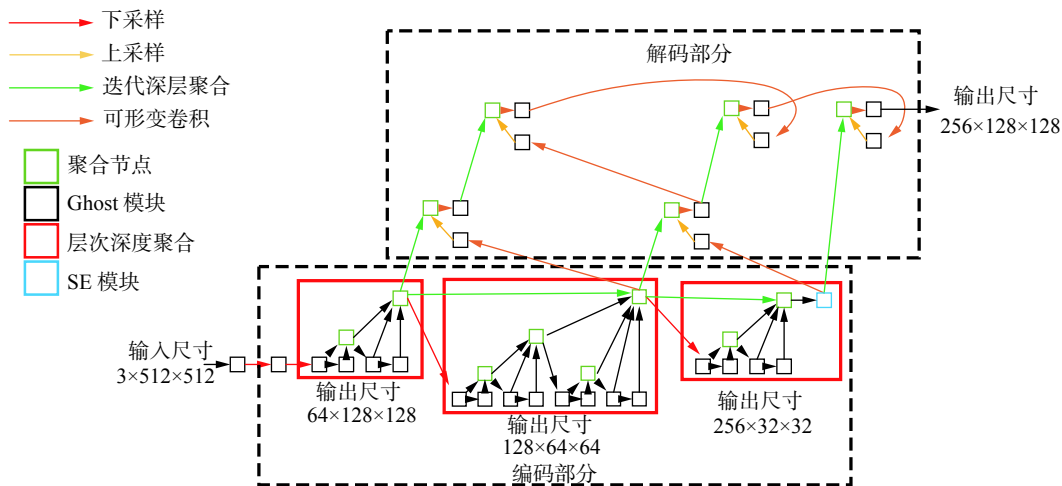


图 3 改进的 Center-Net 特征提取网络框架

Fig. 3 Improved Center-Net feature extraction network framework

对比图 2 与图 3,可以看出特征提取网络改进的部分在于编码阶段减少了一个 HDA 模块,这种操作使得网络下采样输出的分辨率从原来的 16×16 变到 32×32 ,该操作减少了编码解码所需要的时间。实验结果表明,网络的简化没有导致检测精度的过分下降。

2.2 高效卷积模块的使用

为了使网络的推理速度进一步提升,对普通的卷积模块也进行了优化。目前有很多高效的网络结构可以应用于移动设备,比如 MobileNet^[24]、ShuffleNet^[25]等。本文使用华为诺亚方舟实验室于 2020 年由韩凯等^[16]提出的 Ghost 结构来替换原始网络中所使用的普通卷积模块。Ghost 结构见图 4 左侧,其核心思想是先使用卷积核生成一部分特征,之后使用生成的特征通过一些计算代价小的线性变换来生成另外一部分特征,最终将这两部分特征叠加到一起得到最终的输出。

Ghost 结构这种高效的卷积操作可以使网络的运行时间进一步缩短。通过 Ghost 结构搭建的 Ghost 模块如图 4 右侧所示,网络输入首先经过一个 Ghost 结构,然后通过深度可分离卷积^[26]对特

征进行重新整合,再经过一个 Ghost 结构得到输出。其中, BN(batch normalization) 为批归一化操作, Relu (rectified linear unit) 为激活函数。在图 4 中,使用黑色小方框表示 Ghost 模块。

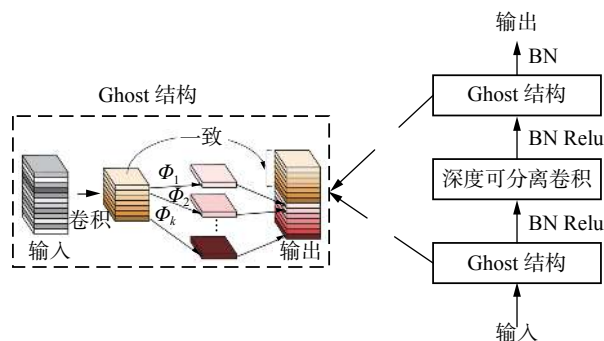


图 4 Ghost 模块结构

Fig. 4 Ghost block structure

2.3 注意力机制的引入

上述操作尽可能地减少了网络的参数量和运行时间,为了让网络在有限参数量下尽可能表达重要信息,在网络编码结构的最后一层引入了注意力机制,使用了压缩激励模块 (SE-block)^[17],如图 5 所示。

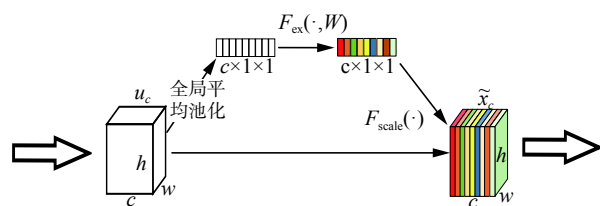


图5 压缩激励模块结构

Fig. 5 SE-block structure

该结构主要包括3个操作,首先压缩操作是对所有通道的特征图做一个全局平均池化(global average pooling, GAP),得到 $c \times 1 \times 1$ 个特征。其中每个特征都具有全局的感受野,表征特征通道上响应的全局分布。之后是激励操作,通过参数 W 为每个特征通道生成权重,公式为

$$s = F_{ex}(z, W) = \sigma(g(z, W)) = \sigma(W_2 \delta(W_1 z))$$

式中: δ 对应 Sigmoid 激活函数; σ 对应 Relu 激活函数; z 是上一层的输出。最后是权重调整操作,公式为

$$\tilde{x}_c = F_{scale}(u_c, s_c) = u_c \cdot s_c$$

式中: u_c 为上一个卷积层的输出; s_c 为经过压缩激励操作的输出。相乘的作用是对原始的特征进行重标定,这样有助于网络更关注具有代表性的特征。在图3中使用蓝色方框代表 SE 模块,从图3中可以看到,在网络完成3个HDA操作得到的聚合节点不直接送到解码网络中,而是经过 SE 模块之后再行解码。这种注意力机制的引入让网络更加集中于重要特征的学习,以此保证网络在数据量大幅度减小的情况下也能很好地完成检测任务。

3 实验测试与分析

3.1 数据集

本文所检测的目标还没有公开的数据集,所以文中使用的数据集是通过喷涂机器人所搭载的摄像头进行采集室内窗户影像,分辨率为 640×480 。使用 labeling 软件进行手工标注。一共 10000 张图片,其中 8000 张用作训练集,2000 张用作测试集。

3.2 图像预处理与增强

相较于大型开源数据集,本次实验所使用的数据集相对较小,所以需要通过图像预处理对图像进行数据增强,除了常规对图像进行简单的位移、缩放、颜色空间变换之外,本次实验还对室内喷涂场景进行了相应的分析,以此让数据集得到更为有效的增强。由于喷涂机器人在喷涂过程中具有距离墙面比较近的工作特性,更多时候需要识别的窗户是不完整的,所以本实验对于图像的预处理还加入了多图拼接,即每次随机选取4张

图片,然后对其进行随机拼接,得到训练图,如图6所示。这样做大幅度地增广了数据,还使得训练集中存在大量的局部窗户图。因此本实验将测试集的图片分成2个部分,1000张是采集的完整窗户,1000张是采集的局部窗户。以此验证改进网络的窗户检测效果,以及拼接方式增强的数据集是否更有利于网络学习局部窗户的特征。



图6 使用多图拼接得到的训练数据

Fig. 6 Training data obtained by multi-graph splicing

3.3 网络训练与测试

如表1所示,本实验在训练和测试阶段分别使用了2种不同的处理器。训练阶段使用图形工作站,加快网络的训练速度,节省实验时间。而在测试阶段,使用喷涂机器人上所搭载的移动端处理器对网络的准确性和实时性进行验证,以此检验算法是否符合要求。

表1 训练和测试过程实验平台对比

Table 1 Comparison of training and testing process experimental platform

实验平台	训练过程	测试过程
CPU	Intel(R) i7-7700	ARM Cortex-A57
GPU	NVIDIA GTX1080Ti	NVIDIA Pascal
系统	Ubuntu16.04LTS	Ubuntu16.04LTS
框架	Pytorch1.4.0	Pytorch1.4.0

训练阶段首先使用网络在 COCO 数据集上进行训练,以此得到较为健壮的权重参数,然后将网络在本次实验中的窗户数据集上进行迁移学习。将图片的分辨率转换为 512×512 ,然后送入网络训练200次,前100次冻结特征提取网络的权重,只训练3个网络分支,后100次将所有网络层解冻,并且使用不同的学习率进行训练,以此得到网络最终的权重值。

在测试阶段改变图像的分辨率,喷涂机器人上摄像头所采集的是 640×480 的分辨率,在训练的时候将图像统一映射到 512×512 的分辨率,如果测试阶段也沿用这种方式那么网络就会计算很多使用灰色填充的空缺位置,这无疑使得网络进行了很多没有价值的计算。所以在测试阶段,本文将图片映射成为 512×384 ,由于网络中没有全

连接层,所以该种分辨率网络仍然可以输出想要的结果。

3.4 实验结果及分析

本实验首先验证多图拼接训练窗口的可行性。使用2种不同的数据增强方式训练数据集,并且分别在完整窗口和局部窗口的测试集上进行验证,使用mAP(mean average precision)作为算法的评估指标,对比结果如表2所示。

表2 使用不同数据增强方式训练的结果对比

Table 2 Comparison of training results using different data enhancement methods

预处理方式	完整窗口mAP	局部窗口mAP
使用多图拼接	0.937	0.835
未使用多图拼接	0.855	0.623

通过表2,可以看出使用多图拼接的方式对训练集进行增强之后无论是完整窗口检测还是局部的窗口检测其mAP值都得到了大幅度的提升,其主要原因在于图像拼接可以使得训练集中的正样本数量得到提升,这样可以缓解其正负样本不平衡的问题,并且使用随机拼接的方式大大增强了训练数据集的多样性,很大程度上抑制了由于训练数据集相对较少而导致的过拟合问题。其中局部窗口检测的mAP值得到了更大的提升,原因是数据拼接使得网络在训练的过程中可以遇到大批量局部窗口的影像,并且随机裁剪的方式也使得局部窗口的多样性大幅度地提升。因此使用该种数据增强的方式可以让网络更好地学习完整和局部窗口的特征,以此更好地解决室内窗户识别的问题。

本实验还使用上述数据增强的方式分别对Center-Net、yolov3-tiny、yolo-nano以及本文改进的Center-Net进行了训练,得到的不同网络之间mAP值、运行时间和参数数量的对比,如表3所示。

表3 不同检测网络性能对比

Table 3 Performance comparison of different detection networks

网络类型	mAP	时间/s	参数量 $\times 10^6$
Yolov3_tiny	0.732	0.056	8 667
Yolo_nano	0.753	0.045	2 848
Center_Net	0.893	0.954	16 373
Center_Net改进	0.887	0.088	2 669

从测试结果可以看出,由于原始的Center-Net具有复杂的特征网络,其参数量巨大,而改进的Center-Net网络,通过对特征提取网络的简化,以及将剩余的普通卷积模块替换成为更加高效

的Ghost模块使得网络的参数量得到了大幅度的减少,甚至小于yolov3_tiny和yolo_nano这2个轻量级目标检测网络。与此同时,网络的运行时间也被大幅度缩短,可以满足在嵌入式设备上实时检测的需求。从表3中还可以看出,改进的Center-Net网络的精度没有因为网络结构的简化而大幅度降低,虽然其mAP值由原始网络的89.3%降低为88.7%,但是对比yolov3-tiny和yolo-nano网络mAP指标仍然有极大的提升。说明Ghost模块和SE模块可以让网络变得更加高效,并且Anchor-free这种思想的目标检测网络可以非常好地处理室内窗户目标检测问题,检测效果如图7所示。从图7中可以看出,无论是整体窗口还是局部窗口,都得到了准确的检测。



图7 改进的Center-Net检测效果

Fig. 7 Improved Center-Net detection renderings

4 结束语

针对室内窗户检测的问题,本文为自主喷涂机器人设计了一种基于深度学习的窗户检测算法。主要借鉴了Center-Net的思想,将其特征提取层进行简化,使用更加高效的卷积模块代替原始的卷积模块,并且引入注意力机制使得所提取的特征更有针对性。通过在实际场景中的验证,该网络可以很好地解决室内窗户检查的问题,具有广阔的应用前景。由于室内喷涂机器人在进行窗户检查的时候主要是对视频的分析,而本文算法并未考虑视频的时序信息,所以接下来的研究方向是根据视频中窗户的连续性等特点,结合时间信息进一步降低算法运算的时间成本,使得算法的性能再次得到提升。

参考文献:

- [1] 沈乐, 李桂清, 冼楚华, 等. 室内3D点云模型的门窗检测[J]. 计算机辅助设计与图形学学报, 2019, 31(9): 1494-1501.
SHEN Le, LI Guiqing, XIAN Chuhua, et al. Door and window detection in 3D point cloud of indoor scenes[J]. Journal of Computer-Aided Design and Graphics, 2019, 31(9): 1494-1501.

- al of computer-aided design & computer graphics, 2019, 31(9): 1494–1501.
- [2] ALI H, SEIFERT C, JINDAL N, et al. Window detection in facades[C]//14th International Conference on Image Analysis and Processing (ICIAP 2007). Modena, Italy, 2007: 837–842.
- [3] 孔倩倩, 赵辽英, 张莉. 基于图像轮廓分析的室内窗户检测[J]. 计算机与现代化, 2018(4): 56–61.
- KONG Qianqian, ZHAO Liaoying, ZHANG Li. Indoor window detection based on image contour analysis[J]. Computer and modernization, 2018(4): 56–61.
- [4] 缪君, 储珺, 张桂梅. 基于图像边缘与玻璃属性约束的窗户检测[J]. 图学学报, 2015, 36(5): 776–782.
- MIAO Jun, CHU Jun, ZHANG Guimei. Window detection based on constraints of image edges and glass attributes[J]. Journal of graphics, 2015, 36(5): 776–782.
- [5] GIRSHICK R. Fast R-CNN[C]//Proceedings of 2015 IEEE International Conference on Computer Vision. Santiago, Chile, 2015: 1440–1448.
- [6] REN Shaoqing, HE Kaiming, GIRSHICK R, et al. Faster R-CNN: towards real-time object detection with region proposal networks[J]. IEEE transactions on pattern analysis and machine intelligence, 2017, 39(6): 1137–1149.
- [7] HE Kaiming, ZHANG Xiangyu, REN Shaoqing, et al. Spatial pyramid pooling in deep convolutional networks for visual recognition[J]. IEEE transactions on pattern analysis and machine intelligence, 2015, 37(9): 1904–1916.
- [8] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: unified, real-time object detection[C]//Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, USA, 2016: 779–788.
- [9] LIU Wei, ANGUELOV D, ERHAN D, et al. SSD: single shot MultiBox detector[C]//14th European Conference on Computer Vision. Amsterdam, The Netherlands, 2016: 21–37.
- [10] LIN T Y, GOYAL P, GIRSHICK R, et al. Focal loss for dense object detection[C]//Proceedings of the IEEE International Conference on Computer Vision. Venice, Italy, 2017: 2999–3007.
- [11] ZHU Chenchen, HE Yihui, SAVVIDES M. Feature selective anchor-free module for single-shot object detection[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach, USA, 2019: 840–849.
- [12] LAW H, DENG Jia. CornerNet: detecting objects as paired keypoints[C]//Proceedings of the 15th European Conference on Computer Vision (ECCV). Munich, Germany, 2018: 765–781.
- [13] ZHANG Shifeng, CHI Cheng, YAO Yongqiang, et al. Bridging the gap between anchor-based and anchor-free detection via adaptive training sample selection[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle, USA, 2020: 9756–9765.
- [14] DUAN Kaiwen, BAI Song, XIE Lingxi, et al. CenterNet: keypoint triplets for object detection[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. Seoul, Korea (South), 2019: 6568–6577.
- [15] TIAN Zhi, SHEN Chunhua, CHEN Hao, et al. FCOS: fully convolutional one-stage object detection[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. Seoul, Korea (South), 2019: 9626–9635.
- [16] HAN Kai, WANG Yunhe, TIAN Qi, et al. GhostNet: more features from cheap operations[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle, USA, 2020: 1577–1586.
- [17] HU Jie, SHEN Li, SUN Gang. Squeeze-and-excitation networks[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, USA, 2018: 7132–7141.
- [18] HE Kaiming, ZHANG Xiangyu, REN Shaoqing, et al. Deep residual learning for image recognition[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, USA, 2016: 770–778.
- [19] NEWELL A, YANG Kaiyu, DENG Jia. Stacked hourglass networks for human pose estimation[C]//14th European Conference on Computer Vision. Amsterdam, The Netherlands, 2016: 483–499.
- [20] YU F, WANG Dequan, SHELHAMER E, et al. Deep layer aggregation[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, USA, 2018: 2403–2412.
- [21] HUANG Gao, LIU Shichen, VAN DER MAATEN L, et al. CondenseNet: an efficient DenseNet using learned group convolutions[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, USA, 2018: 2752–2761.
- [22] PAN Junting, SAYROL E, GIRO-I-NIETO X, et al. Shallow and deep convolutional networks for saliency prediction[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, USA,

2016: 598–606.

- [23] ZHU Xizhou, HU Han, LIN S, et al. Deformable ConvNets V2: more deformable, better results[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach, USA, 2019: 9300–9308.
- [24] SHENG Tao, FENG Chen, ZHUO Shaojie, et al. A quantization-friendly separable convolution for mobilenets [C]//2018 1st Workshop on Energy Efficient Machine Learning and Cognitive Computing for Embedded Applications (EMC2). Williamsburg, USA, 2018: 14–18.
- [25] ZHANG Xiangyu, ZHOU Xinyu, LIN Mengxiao, et al. Shufflenet: an extremely efficient convolutional neural network for mobile devices[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, USA, 2018: 6848–6856.
- [26] CHOLLET F. Xception: deep learning with depthwise separable convolutions[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, USA, 2017: 1800–1807.

作者简介:



洪恺临, 硕士研究生, 主要研究方向为计算机视觉、深度学习。



曹江涛, 教授, 博士, 主要研究方向为智能方法及其应用、视频分析与处理。主持国家自然科学基金项目1项、辽宁省自然科学基金项目1项。参与编著英文专著2部, 发表学术论文50余篇。



姬晓飞, 副教授, 博士, 主要研究方向为视频分析与处理、模式识别理论。主持国家自然科学基金项目1项、辽宁省自然科学基金项目1项。参与编著英文专著2部, 发表学术论文40余篇。

2021 中国粒计算与知识发现学术会议

2021 China Granular Computing and Knowledge Discovery Conference

由中国人工智能学会主办、中国人工智能学会粒计算与知识发现专委会协办、国际粗糙集学会支持, 华东交通大学承办的2021年中国粒计算与知识发现学术会议(第21届中国粗糙集与软计算学术会议、第15届中国粒计算学术会议、第9届三支决策学术会议)将于2021年8月20–22日在“英雄城——江西南昌”召开。现将会议有关征文事宜通知如下, 热忱欢迎相关研究人员踊跃投稿并参会。

征文范围 (包括但不限于):

1) 粗糙集与软计算; 2) 粒计算理论及其应用; 3) 三支决策模型与分析; 4) 知识发现与数据挖掘。

重要日期:

投稿截止日期: 2021年3月25日

录用通知日期: 2021年6月10日

终稿提交日期: 2021年6月25日

会议举办日期: 2021年8月20–22日

投稿要求及详情参见网址: <https://easychair.org/conferences/?conf=cgckd2021>

投稿与会务咨询:

钱老师(13775075661); 余老师(13755776891)

会务邮箱: cgckd2021@163.com

通信地址: 江西省南昌市双港东大街808号华东交通大学软件学院(330013)

会议网站: <http://cgckd2021.ecjtu.edu.cn>