

改进 yolov5 的口罩和安全帽佩戴人工智能检测识别算法

王 沔

(福建省建筑科学研究院有限责任公司 / 福建省绿色建筑技术重点实验室, 福建 厦门 361000)

摘要: 由于疫情的影响在特定施工环境下, 作业人员需同时佩戴口罩和安全帽。为了监控作业人员是否符合安全作业要求。本文将 yolo 目标检测算法应用于人脸口罩和安全帽佩戴检测, 通过改进 yolov5 算法, 使用实验室摄像头监控数据和爬取互联网建立了包含 21000 张图片的数据集。通过手工进行标注, 然后首先采用 K-means++ 算法进行 anchor 维度聚类, 确定 anchor 参数; 并将 ciou 和 diou_nms 应用于 yolov5 网络。经过神经网络训练后将生成的模型, 部署在 jet nano 上。结果表明应用该改进算法模型, 可以有效监控口罩和安全帽佩戴情况。

关键词: yolov5; 口罩; 安全帽

中图分类号: TU714; TP391.41 文献标志码: B

文章编号: 1673-0402(2020)11-0067-04

疫情期间, 在特定环境下进行建筑施工或建筑检测作业, 作业人员佩戴安全帽和口罩是标配。比如在密闭的门窗幕墙施工和检测现场, 如果没有佩戴安全帽和口罩, 将会造成严重的安全隐患和疫情的传播风险。使用监控摄像头, 通过人工智能算法实时监控现场作业人员的安全帽和口罩佩戴情况具有很大意义和价值。

近年来, 很多研究学者发布了大量基于人工智能的目标识别检测模型。比如 R-CNN、Fast R-CNN 和 Faster R-CNN 作为最为常见的 Two-Stage 算法模型, 最早是 Ross Girshick 发布的。这种模型首先从基于特征提取的基础上, 采用一个神经网络生成候选区域, 然后再对候选区域回归来确定目标的类别和位置, 检测速率可以达到 5f/s。由 Redmon 等提出的 YOLO 方法, Liu 等提出 SSD 方法, 是 One-Stage 算法, 检测速度达到了 45 f/s。

2019 年施辉采用 Image Pyramid 通过提取多种尺度的特征图, 结合 YOLO v3 进行安全帽佩戴的检测。2020 年王兵使用新的 Target 识别函数和 GIoU 计算方法来改进 YOLO v3 算

法, 训练检测模型。

本文使用门窗幕墙实验室监控录像数据, 以及网络收集的口罩佩戴数据集和安全帽佩戴数据, 制作了安全帽和口罩佩戴的检测数据集。基于 YOLO v5 模型为主体, 将多尺度的特征进行融合, 用于进行类别判断和位置识别; 然后用数据集对 Anchor 框进行统计, 增加 Anchor 框的个数, 通过预测偏移量, 基于 Anchor 框和 gird 的 sigmoid 偏移量; 最后, 采用多个分辨率的图片进行模型训练, 在预测的时候, 输入大尺寸的图片或大分辨率的流媒体数据, 保证多尺寸的训练效果。试验表明, 改进的 YOLO v5 算法, 不但有效提高了检测速率, 同时也能够保证检测精度。

1 yolov1-yolov4 算法介绍

传统目标检测方法: 滑动窗口分类法, 耗时, 操作复杂, 需要手动生成大量样本, 检测性能依靠的是分类器的性能和遍历的性能。因此 yolov1 实现了通过直接拟合坐标位置 x, y , 物体宽和高 w, h , 置信度的算法。存在很明显的缺

收稿日期: 2020-01-18

作者简介: 王 沔 (1986-), 男, 硕士, 工程师, 主要从事深度学习, 工程检测工作。

陷坐标位置 x, y 可以任意值, 也可以为没有意义的负值, 降低了训练速度和预测的准确率。

因此, yolov2 改为偏移量预测: 采用了 anchor 机制 (location prection) 局部预测, 也就是坐标位置 x, y 是通过 gird 预测, 然后宽和高 w, h 是依据 anchor 的宽和高预测。yolov2 的 backbone 改为了 darknet19, 训练不同 epoch 后的调整训练图, 也就是 Multi-Scale Training, 连接检测头部分取消了 FC 的连接方式, 改为卷积层。可是 yolov2 还是存在小目标预测不准确的问题。

随着残差网络 resnet 对 backbone 性能的提升, yolov3 又有了新的性能提升。首先是因为 resnet 出现了, 网络可以做的更深了, backbone 更改为 darknet53, 然后检测头也更改为多尺度, 统计数据得来的 9 个尺度, 32 倍下采样, 16 倍下采样, 8 倍下采样, 分别用 3 个检测头来预测。

到 yolov4 在 bakhone 更改为采用 CSPDarknet53, 同时采用了一些技巧进行提升性能, 首先数据增强采用了 cutmix and mosaic, class label smoothing, Mosaic 就是合并 4 张图, 起到了跨越上下进行标检测的作用, 同时相当于 minibatch 的数量扩大了 4 倍, 同样的 GPU 内存训练提升了性能。损失函数改为 mish。网络结构采用了 SPP、PAN、SAM 的网络, DropBleok CmBN。Yolo v1-v3 (location prection) 因为激活用的 sigmod 区间是 (0, 1) 开区间, yolov4 对此进行了改进。在 sigmod 乘上 1.01。

2 yolov5 网络分析

Yolov5 按照网络深度大小和特征图宽度大小分为 Yolov5s、Yolov5m、Yolov5l、Yolov5, 本文采用了 yolov5s 作为使用模型。Yolov5 的结构分为 input, backbone, Neck, 预测层。

(1) 在输入端使用了 Mosaic 的数据增强方式, 随机调用 4 张图片, 随机大小和分布, 进行堆叠, 丰富了数据, 增加了很多小目标, 提升小物体的识别能力。可以同时计算 4 张图片, 相当于增加了 Mini-batch 大小, 减少了 GPU 内存的消耗。Yolov5 首先也可以通过聚类设定 anchor 大小, 然后还可以在训练过程中, 在每次训练时, 计算不同训练集中的 anchor 值。然后在预测时使用了自适应图片大小的缩放模式, 通过减少黑边, 提高了预测速度。

(2) 在 Backbone 上的主要是采用了 Focus 结构, CSPnet 结构。

(3) 在 Neck 上采用了 FPN 结构和 PAN

结构。

(4) 在损失函数则使用了 GIOU_Loss。

3 改进的 yolov5 算法用于口罩和安全帽佩戴检测

3.1 使用 Ciou

Yolov5 中采用其中的 GIOU_Loss 做 Bounding box 的损失函数。GIOU_Loss 中, 增加了相交尺度的衡量方式, 解决了无法反应两个框距离的远近, 此时 loss 函数是不可求导的, GIOU_Loss 就无法计算两个框没有交集的情况。但是无法解决预测框在目标框内部且预测框大小一致的情况, 因为这时预测框和目标框的差集都是相同的, 因此, 这三种状态的 GIOU 值也都是相同的。因此更改为采用效果更好的 CIOU_Loss, 考虑了边界框中心点距离的信息, 同时考虑边界框宽高比的尺度信息。

3.2 使用 Diou_nms

非极大值抑制 nms, 用于推理过程中的预测框的选择, yolov5 算法中, 采用普通的 nms 的方式。本文将更改为 Diou_nms。使得 NMS 得到的结果更加合理和有效。

4 试验分析

4.1 试验数据集制作

本文的实验数据, 是针对安全帽和口罩同时佩戴检测的数据, 由于没有的公开数据集, 本文制作了一套安全帽和口罩佩戴的数据集, 制作数据集主要包含了四个步骤: 数据采集和整理、数据的预处理、数据的筛选和数据集标注。

4.2 数据收集

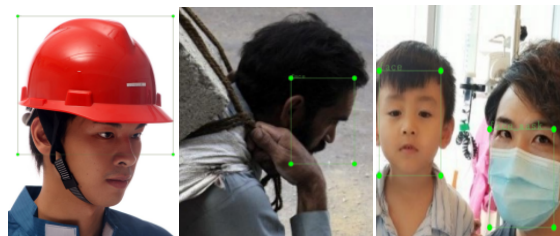


图1 图像标记示意图

本数据集的数据来源一个是建筑检测实验室的视频监控数据和工地施工现场视频监控数据, 以及网络上采集的图片数据。主要使用福建省建筑工程质量检测中心有限公司实验室门窗幕墙检测过程中, 工人安装门窗幕墙的视频监控数据。从监控中分帧得到的图片数

据,很多没有检测实体,将这些数据通过算法删除。然后使用labelimg工具对图像进行标记,标注类别为佩戴口罩,佩戴安全帽,人脸三类,如图1所示。

4.3 试验方案

选取20000张标注的图片作为训练样本,不同分辨率和大小的样本均分,以确保不同尺寸的图片都能得到充分的训练;另外选取1000张图片作为测试数据集。

4.4 试验平台

为了加快训练速度,快速达到训练收敛的目的。本次训练采用了网络上的算力平台。系统配置为 GeForce RTX 2080 Ti x4 GPU, Intel(R) Core(TM) i7-9800X CPU, 内存 16G, 操作系统 Ubuntu 19.10 5.3.0-42-generic GNU/Linux, CUDA 10.2。

5 试验结果分析

第3节提出的yolov5改进算法, Ciou 和 Diou_nms 对检测效果有很大的提升, 为了对比改进效果, 按照检测目标的大小分为大中小三类, 图2中列出了原始算法和 yolov5 改进算法对三类图片的检测效果。

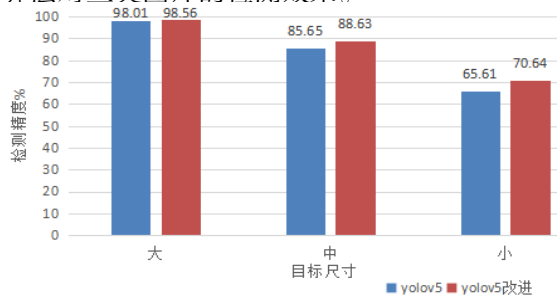


图2 原始算法和 yolov5 改进算法对比

由图2可以看出, 采用 Ciou 和 Diounms 算法的检测准确率均比原始算法检测精度更高。

6 试验测试

因为 Yolov5 的模型大小仅为 28m, 可以方便的部署在小型嵌入式设备中, 本文将其部署在 jet nano 上便携可移动。Nano 支持高分辨率摄像头, 可同时处理多个传感器, 支持常见的框架, 如 yolov5 模型使用的 pytorch 框架。本文最终设备如图3所示。包含了摄像头, 无

线网卡, 主板, 风扇, sd 卡。检测效果如图4所示。



图3 最终设备示意图



图4 检测效果示意图

因此, 改进 YOLO v5 算法能够满足高检测速率和高准确率。同时将模型部署在 nano 上是可行的。

7 结语

综上所述, 本文提出了识别安全帽和口罩佩戴的基于改进 YOLO v5 的检测算法。使用门窗幕墙检测实验室现场和施工现场监控视频作为主要数据来源进行检测试验, 通过 Ciou, Diou_nms 等方法改进 YOLO v5 网络, 并将其部署于 nano 设备上, 能够满足疫情期间, 特定环境安全帽和口罩佩戴监控需求。

参考文献:

- [1] 李星辰. 融合 YOLO 检测的多目标跟踪算法 [J]. 计算机工程与科学, 2020(4): 665-672.
- [2] 李章维. 基于视觉的目标检测方法综述 [J]. 计算机工程与应用, 2020(3): 1-9.
- [3] 刘磊. 一种 YOLO 识别与 Mean shift 跟踪的车流量统计方法 [J]. 制造业自动化, 2020(2): 16-20.
- [4] 王洋. 人工智能在门窗检测中图纸识别的应用 [J]. 四川水泥, 2019(9): 137-138.