

基于 yolov4 的车辆提醒系统

李 佳 周 丽

四川大学锦城学院 计算机与软件学院 四川 成都 611731

【摘要】随着人们生活水平的不断提高,交通越来越便利的同时,安全隐患也逐年上升,行人安全意识不足导致交通事故发生频率逐年增加,与此同时目标检测算法相对成熟,我们将yolo v4 算法进行了重新训练模型,采用VOC数据集来设计并训练了一个无红绿灯路口的车辆提醒系统,用于减少在光照不足,夜间照明条件不良的情况下提醒是否有车辆靠近,引导行人安全过马路,通过对模型结果分析,可以得到该模型可以在实际运用中具有很好的效果。

【关键词】 目标检测； yolov4 算法； 行人安全； 深度学习

1 引言

人们的生活水平的提高的同时, 交通出行工具的选择越来越多样, 道路以及道路上的车辆越来越多, 随之而来就是出现的安全隐患越来越大, 据我国统计年鉴显示, 2014—2018 年的交通事故发生量分别为 196812 187781 212846 203 049 和 244937, 交通事故发生量总体呈现出明显上升的趋势, 2018 年交通事故发生量达到了历年的最高值。城市公路在我国公路体系中占据重要地位, 也是交通事故的高发区域^[1], 所以面对这一系列惨剧的发生, 无时无刻不提醒我们提高行人过马路安全的重要性。

随着深度学习的近年来的火爆，目标检测算法近年来采用了很多改进，目前主流的算法包括两大类，一类是以 R-CNN^[2] 为代表的生成可能区域并进行 CNN 提取特征再将其放入分类器分类并修正位置，另一类是以 yolo 算法为代表的对预测物体进行回归，前者注重于准确率高后者注重于速度快，而本文中的 yolo 算法随着几代更新能够基本达到拥有快速的检测速度和高召回率。

2. 人脸检测 yolo 算法

2.1.yolo 算法概述

yolo 算法是 Redmon 等^[3]在 2015 年提出的第一个单目标检测方法算法, 通过滑动窗口得到不同大小的检测窗口, 然后采用非极大抑制算法^[4](Non-Maximum Suppression, NMS)的方法得到一个目标的检测框以及分类别, yolo 将整个图片分成 $S \times S$ 的格子, 然后判断检测目标中心点的位置, 将其得到的 boxes 值写入存放对应的格子里, 得到一个三维张量作为输出结果。yolo 算法由于检测速度快, 检测准确高的特点在工业实践中具有广泛的应用。

2.2.volov4 的改进

2.2.1 网络结构改进

yolov4^[5] 采用的是 CSPDarknet53 是在 Darknet53^[6] 的基础上, 依据 CSPNet 算法修改了 backbone 结构, 其基础实现思路如图 1 所示. 将基础层复制的特征映射图发送到下一层, 从而将基础层的特征映射图分离出来, 避免了因网络层次过深导致信号丢失的问题, 重用特征可以减少计算量, 这可以在提高模型的精度的同时减少计算量。

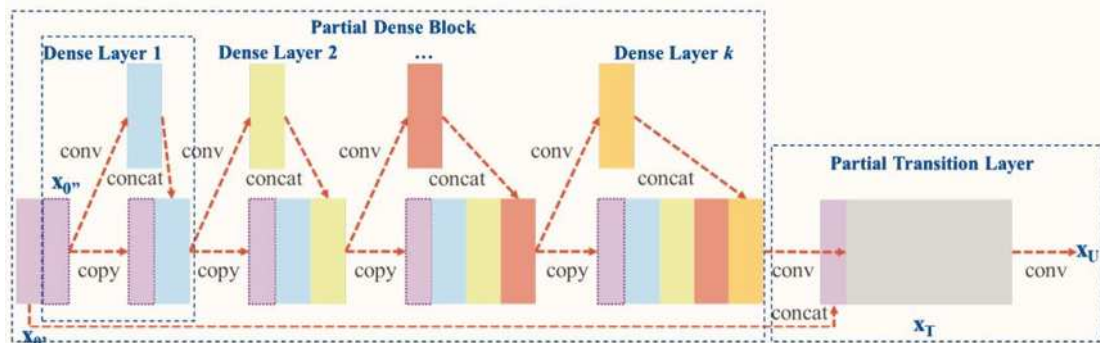


图 1 CSPDarknet53 结构图

2.2.2 特征提取改进

SPP 模块主要是通过加入最大池化层来增加特征的接受范围,只需要加入少量的计算代价就可以使性能大大提高。FPN 又名特征金字塔,原模型的 FPN 模块是通过自顶向下的上采样方式传递目标的强语义特征,但是

对目标的定位信息传递较少。针对原模型的定位问题,新模型在 FPN 模块的后面添加一个自底向上的金字塔结构的 PAN 模块,该模块通过改进 PANet[7] 结构得到,这不仅解决了检测中尺度问题还提升了网络的鲁棒性。经过 SPPNet 与 PANet 处理后,模型输出的特征维度大小分别为 (13, 13, 75)、(26, 26, 75)、(52, 52, 75),

特征中包含更多的细节和语义信息,能让特征处理层更好的利用这些特征进行分类及回归。^[8]

2.2.3 检测框改进

最开始的 IOU_Loss^[9] 计算检测框和目标框之间的重叠面积,GIOW_Loss 解决了 IOU 边界框不重合时出现的问题,而 DIOU_Loss 在 IOU 和 GIOU 的基础上,同时考虑了边界框中心点距离信息;CIOW_Loss 在 DIOU 的基础上加入了边界框宽高比的尺度信息考虑因子,得到的如下图 2 的公式。通过改进了非极大抑制算法,让其对密集物体也有着较高的检测。

$$CIOW_Loss = 1 - CIOW = 1 - \left(IOU - \frac{Distance^2}{Distance_C^2} - \frac{v^2}{(1 - IOU) + v} \right)$$

图 2 CIOW_Loss 公式

3 实验及结果分析

3.1 实验环境

本实验在百度开发的 AiStudio 平台,使用平台提供的 GPU 算力,相当于 TeslaV00, CPU: 4CoresRAM 32GB 进行的

3.1 系统实现流程

本系统如下图 3 所示,由两个部分组成,包括指示灯用告知行人附近是否有车,以及将检测算法嵌入进红外电子摄像头中。红外摄像头通过实时检测路上车辆信息并返回给给指示灯,行人在夜间照明条件不良车辆通过指示灯得到路口车辆状况时,若指示灯亮红灯证明有车靠近,若指示灯亮绿灯证明无车路口安全,从而保护行人安全和车辆行车安全。

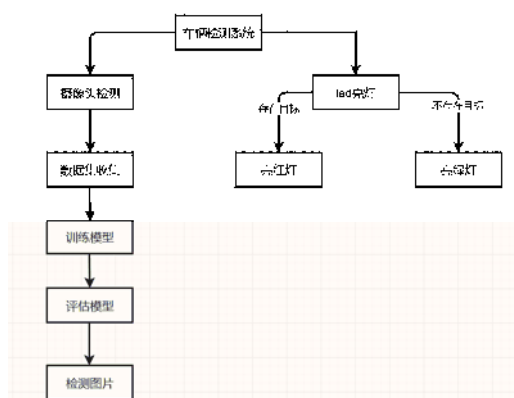


图 3 系统流程图

3.2 训练集准备

本实验采用百度 AiStudio 上 VOC 车辆数据集,总共包含 60000 张图片。大部分图片取自路面监控,大多数图片主要由如下图 4 所示含有多个类所组成。yolov4 中提供了不同大小的 anchor 的目标检测,本实验主要针对的是小目标检测,原模型采用了 20 个分类,而本实验只采用 car 类, bus 类, van 类, others 类这四个类,因此在进行数据清洗之后只会提取到训练图片中有 car 类, bus 类, van 类的数据的目标框的位置和类别用于训练,其他类均是 others 类。标签文件中存放了原模型中的 20 分类在照片出现的所有分类的类名和位置信

息由于本文所采用的数据集,对于大多是基于晴天的数据,对于晚上的数据收集难度较大,故应用对于夜晚识别准确度没有那么高。可能会出现一些漏检问题。不同的天气所带来的光照条件,都会给实时检测带来影响会导致检测准确度不同从而影响检测结果。因此本文基于上述问题在选择安装路口的时候选取具有良好的夜间照明系统条件下实施。



图 4 数据集

3.3 训练模型

总数据量有 60000 张图片,将其按 9:1 的比例划分为训练集和测试集,将其路径分别保存到训练文本和测试文本中,但是 yolov4 算法中会对数据进行一些数据增强和数据处理,例如从而使数据量加大,训练采用的学习率为 0.001,通过对 VOC 车辆数据中的图片中的汽车,公交,货车来获得图片中的每一张照片对应类的位置,然后通过文本里的图片路径在 yolov4 原模型的基础上重新训练自己的模型得到模型权重,通过模型权重来评估模型的好坏,最后通过模型来评估所要检测的图片。

3.4 套用模型

将路口监控视频每隔 7 帧保存一张图片,在主函数调用 video2frame(path0,path1,7) 函数,分别传入视频所在路径,图片保存路径与分帧间隔三个参数,将视频进行分帧,方便后续功能的处理^[10],将得到的图片用模型去检测,通过判断检测结果然后向 led 指示灯返回相应的结果,若存在则返回 1,对应的 led 灯亮红灯,不存在则返回 0,led 灯亮绿灯。这样就可以间接达到实时监测的目的。

3.5 模型效果展示及分析

为了验证模型的准确性和可行性,本文通过对模型的 MAP 函数如下图 5 所示来进行分析,由图可以看出当整个实验训练到第 30 轮时,MAP 的精度会达到 70% 左右,MAP 总体从第一轮开始呈现出上升的趋势,模型的 loss 值从第五轮之后保持在 0.94 左右,可以得出整个模型趋向于收敛,后续的收敛速度较慢,说明模型对车辆数据具有较好的学习能力。

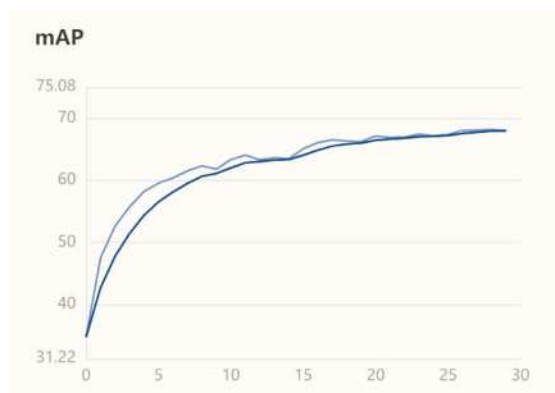


图 5MAP 图

通过对实际图片检测得到图 6 的检测结果,可以看出模型在能够检测出所有的车辆,说明模型在检测效果还是很好的。在实际测试效果的时候,此模型在检测准确率可以达到 $\geq 90\%$, 误报率: $\leq 5\%$, 这里的误报率是对车辆类型识别错误,但是在实际投入使用我们只需要检测出是否有车,所以不会对实际应用产生影响。



图 6. 实际检测图

虽然模型在检测上具有很好的效果,但是数据量收集不够,缺少多场景下的数据集,导致模型精度不是很高,同时现实条件下所考虑影响因子过多,例如行人和车辆的行驶路线,车辆的行驶速度与摄像头的捕获速度都会对此系统产生巨大的影响,这也是我们接下来要对这个系统进行改进的地方。

4 小结

本文通过在 yolov4 的基础上进行了一系列的调整,得到自己的模型, yolov4 的一些调整小技巧使数据集更加丰富,又因为引入了 CSPDarknet53 使得计算量没有较大的提高,训练速度相比于同类网络依旧很快,通过在 yolov4 的模型基础上采用 VOC 车辆数据集来训练对类的重新选取得到新的模型权重,对实验结果的分析我们可以进一步来得到模型效果,相比较于现在一些其他的目标检测模型, yolov4 在车辆识别精度和识别速度上要优于其他模型,随着人民生活水平的提高,人们对道路安全的重视将会越来越重视,所以此模型的精度和速度可以在实际运用中具有重大意义。

【参考文献】

[1] 苏芳,袁勤.城市公路交通事故关联规则分析[J].武汉理工大学学报(信息与管理工程

版),2020,42(04):313-318+331.

[2] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[J]. Computer Science, 2013: 580- 587.

[3] Redmon J, Divvala S, Girshick R, et al. You only look once: Unified, real-time object detection[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. New York: IEEE, 2016: 779-788.

[4] Neubeck A, Gool L. Efficient Non-Maximum Suppression[C]//International Conference on Pattern Recognition. IEEE Computer Society, 2006.

[5] Alexey Bochkovskiy, Chien-Yao Wang, Hong-Yuan Mark Liao. YOLOv4: Optimal Speed and Accuracy of Object Detection [C] //CVPR, Apr, 2020.

[6] Wang H B, Zhang Z D. Text detection algorithm based on improved YOLOv3[C]. Beijing: IEEE the Ninth International Conference on Electronics Information and Emergency Communication, 2019. [9] ZHENG Z, WANG P, LIU W, et al. Distance-IoU loss: faster and better learning for bounding box regression[C] // AAAI, 2020: 12993-13000.]

[7] LIU Shu, QI Lu, QIN Haifang, et al. Path aggregation network for instance segmentation[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018: 8759-8768.

[8] 高伟,周宸,郭谋发.基于改进 YOLOv4 及 SR-GAN 的绝缘子缺陷辨识研究 [J/OL]. 1-11 (2021-05-19 08:34).

[9] RAHMAN M A, WANG Y. Optimizing intersection-over-union in deep neural networks for image segmentation[J]. International Symposium on Visual Computing 2016.

[10] IAN Goodfellow, YOSHUA Bengio, AARON Courville. 深度学习 [M]. 北京: 人民邮电出版社, 2017.