

基于改进卷积神经网络的快速车辆检测*

朱锋彬, 应 娜

(杭州电子科技大学 通信工程学院 浙江 杭州 310018)

摘 要: 为了更加快速精准地识别和定位车辆,提出了一种基于卷积神经网络的改进算法。在建议网络(PN)以及特征提取上进行了优化,在提取基本特征之后嵌入一个轻量级的建议网络,产生候选区域,并在早期丢弃不可能的区域候选框减少背景数量;采用特征融合技术提取 Hyper 特征,细化了车辆的识别和定位,并对不同层提取的特征进行串联,有效地提高了车辆检测的质量和精度。在车辆检测和跟踪的大规模数据集和自己采集的车辆数据集上进一步验证,并与目前一些先进的算法进行比较,其中相比于 Faster RCNN 平均检测质量提高了 9.91%,充分表明了提出的算法的有效性。

关键词: 车辆检测; 卷积神经网络; Hyper 特征; 特征串联

中图分类号: TN391 **文献标识码:** A **文章编号:** 1000-9787(2018)12-0153-03

Fast vehicle detection based on improved CNN*

ZHU Feng-bin, YING Na

(College of Telecommunication Engineering, Hangzhou Dianzi University, Hangzhou 310018, China)

Abstract: In order to identify and locate vehicle more quickly and accurately, an improved algorithm based on convolutional neural networks(CNN) is proposed. The algorithm is optimized in the proposed network and feature extraction. Specifically, the framework is embedded with a light-weight proposal network to generate candidate region and early discard candidate frame of impossible regions to reduce background numbers; feature fusion technology is used to extract hyper feature, refine the identification and location of the vehicle, which improve the quality and precision of vehicle detection effectively. Further verify on large-scale datasets of vehicle detection and tracking and the self-collected vehicle datasets, and compared with current advanced algorithms, compared with Faster RCNN, average detecting quality is improved by 9.91%, which fully shows the effectiveness of the algorithm.

Keywords: vehicle detection; convolutional neural network(CNN); Hyper feature; feature concatenation

0 引 言

在计算机视觉领域,车辆检测作为一个实用而又具有挑战的课题,得到了很多研究者的关注。道路监管、辅助驾驶等人工智能都融入了车辆检测的应用。随着科学技术的不断发展,研究人员在不同阶段取得了不同的成就。从最初的区域选择^[1]到特征提取^[1],从机器学习卷积神经网络 R-CNN 到 Fast R-CNN, Faster R-CNN 模型,以及后来的 SSD(single shot multiBox detector),YOLO(you only look once)算法,车辆检测算法都在朝着更快更准的方向发展,在检测速度、定位精度上都得到了巨大提升。

在前人的基础上,为了达到更快速、准确地进行车辆检测目的,提出了一种改进算法,在一个深度卷积网络的基础上搭建了 2 个神经网络。算法首先改进 VGG16 网络中的

深度卷积神经网络(deep convolutional network, DCN),加入超特征技术,融合不同卷积层的特征,得到图片的超特征。然后在第一个建议网络(proposal network, PN)采用了一组卷积层来生成锚箱,并利用锚箱直接提取出候选框,同时利用非极大抑制算法(NMS)丢弃不太可能的候选框。并将剩余的候选框送入第二个微调网络(fine-tuning network, FTN)。算法融合 PN 网络、FTN 网络以及不同卷积层的特征,使得车辆的识别和定位都得到了显著的优化。

1 网络结构组成

如图 1 所示,提出的改进的卷积神经网络由: DCN, PN, FTN 3 个子网络组成。

1.1 DCN

首先利用 DCN 网络产生丰富的图像特征。采用 VGG-16

收稿日期: 2017-09-106

* 基金项目: 浙江省级公益性技术应用研究计划项目(2014C31056); 国家自然科学基金资助项目(61372157)

网络模型中的卷积层和最大池化层,其中权重由 ImageNet 数据集预训练完毕,以确保检测可以受益于不断变化的大数据集。用多层具有小卷积核滤波器的卷积层代替一层滤波器卷积核较大的卷积层。最终将输出的特征图送入 PN 网络和 FTN 网络进行进一步处理。另外,该网络还用到了各个卷积层之间的特征串联。

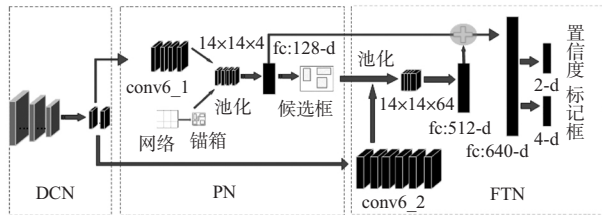


图1 改进的卷积层框架结构

1.2 PN

在图像检测过程中,于数量较大的背景物体中有效地筛选出候选物体汽车是建立一个高效的车辆检测系统必不可少的一步。可在特征提取顶端设计一个深度的分类网络实现。据此,设计了一个轻量级的特征提取网络——PN网络,利用该网络提取候选框,以及消除背景区域。PN网络是由一个ROI池化层,一个深度卷积网络和一个全连接层组成,具体步骤是:用一个带有4个滤波器的卷积层(conv6_1)连接到DCN的输出,得到一个 256×144 的特征映射。其中,conv6_1可使超特征更加精炼,同时减少占用内存,与以下conv6_2作用相同。将特征映射到 64×36 的网格。对于每个网格,生成一组固定尺度(32×32 , 64×64 , 128×128 , 256×256 , 512×512)和不同比例($1:2$, $2:1$, $1:1$)的锚箱。与Faster R-CNN^[2]不同的是,算法直接用锚箱提取出候选对象。ROI池化层从每个锚箱中取出一个 $14 \times 14 \times 4$ 特征向量;将经裁剪的特征向量送入一个128维的全连接层,得到候选框位置(x , y , w , h)和得分 s 。使用这个简单但有效的建议网络,将带来一个显著的速度提升。

1.3 FTN

FTN负责微调剩余的候选框。除了使用一个有64个滤波器的卷积层(conv6_2)产生 256×144 特征图,其他结构上类似之前的PN网络。FTN层主具体步骤是:ROI池化层为PN网络产生的车辆候选框提取一个 $14 \times 14 \times 64$ 的特征向量。连接一个640维的全连接层。本文涉及到两个阶段的特征串联:由现在的FTN网络形成512维特征;由PN网络形成128维特征(连接方式在图1中用符号表示)。然后再连接一个640维的全连接层。最后输出微调后的候选框位置和置信度——一个5维的向量:标出的物体框的位置、大小(x , y , w , h)以及区域得分 s 。

2 特征融合与特征连接

1) 依据文献[3]结论,本文采用融合多个不同层次卷

积层的特征图方法进行特征融合。具体操作如下:文献[4, 5]显示了将不同卷积层的特征图结合的效果。仿效这种方案,本文将DCN网络的多层特征进行融合从而产生超特征。图1显示了多特征图的连接。用卷积层计算整张输入图的特征图时,为了在同一分辨率下组合多层特征图,算法对不同层使用不同的采样策略。对低层进行最大池化下采样;高层特征使用反卷积操作进行上采样。将调整过的采样结果各自通过一层卷积层进行卷积操作,缩小维度并增强语义。采用局部响应归一化(local response normalization, LRN)^[6]对特征图进行归一化处理,并连接到一个单一的立方体输出,产生Hyper特征。Hyper特征有以下几个优势:多层抽象,分辨率适于检测,计算效率高。

2) 特征连接。将通过FTN网络的候选框产生的全连接层特征与此候选框在PN网络产生的全连接层特征进行串联。具体实现:在PN网络中记录保留下的候选框,在FTN网络中根据通过的候选框的序号找出PN网络中对应候选框产生的特征,进行连接。本文将FTN网络产生的128维特征图与PN网络产生的512维特征图串联得到640维的特征图。

3 训练

3.1 网络训练

初始化,网络中的DCN部分是由预训练好的ImageNet模型VGG-16的参数初始化,其他网络部分参数则是由零均值、标准差为0.01的高斯分布随机初始化。算法设置的初始化学习速率为0.001,经过 50×10^3 次迭代之后降为0.0001。该网络总共要经过 70×10^3 次的迭代训练。

算法利用随机梯度下降法对模型进行端到端的训练。将IOU(intersection over union)大于0.5的归于正样本候选框,小于0.3的归于负样本;利用阈值为0.7的非极大值抑制(non-maximum suppression, NMS)算法^[7]来消除多余的框。对于FTN网络,将 $\text{IOU} \geq 0.45$ 的归于正样本,将 $0.1 \leq \text{IOU} \leq 0.3$ 归于负样本。并在训练过程中加入难样本。按降序排序分类损失,并选择前70%个样本参与反向传播,并且忽略简单的例子。

3.2 多阶段损失函数

PN层和FTN层的全连接层分别产生 (x, y, w, h, s) 和 (x', y', w', h', s') 。类似于文献[8],算法用 $t = (t_x, t_y, t_w, t_h)$ 来参数化第一阶段PN产生的包围盒

$$\begin{cases} t_x = (x - \hat{x}) / \hat{w} & t_w = \log(w - \hat{w}) \\ t_y = (y - \hat{y}) / \hat{h} & t_h = \log(h - \hat{h}) \end{cases} \quad (1)$$

式中 \hat{x} , \hat{y} , \hat{w} , \hat{h} 为初始锚箱的位置。由 $t' = (t'_x, t'_y, t'_w, t'_h)$ 来参数化第二阶段FTN产生的得到改善了的候选框

$$\begin{cases} t'_x = (x' - x) / w & t'_w = \log(w' / w) \\ t'_y = (y' - y) / h & t'_h = \log(h' / h) \end{cases} \quad (2)$$

利用多级损失函数 L 共同训练 2 阶段的分类和回归

$$\begin{cases} L = \alpha L_{\text{pn}} + (1 - \alpha) L_{\text{ftn}} & L_{\text{pn}}(t, s) = L_{\text{cls}}(s) + \lambda L_{\text{loc}}(t, s) \\ L_{\text{ftn}}(t', s') = L_{\text{cls}}(s') + \lambda L_{\text{loc}}(t', s') \end{cases} \quad (3)$$

式中 L_{pn} 和 L_{ftn} 为 PN 网络和 FTN 网络各自单独的损失。 α 为平衡参数,平衡两个网络。类回归损失 $L_{\text{cls}}(s) = -\log s$ 为以类别得分 s 的对数形式表示。为了回归包围盒的位置,按照文献[9]的做法定义位置损失函数 L_{loc}

$$L_{\text{loc}}(t, s) = \sum_{i \in \{x, y, w, h\}} \sigma(t_i - s) \quad (4)$$

$$\sigma(x) = \begin{cases} 0.5x^2, & |x| < 1 \\ |x| - 0.5, & \text{其他} \end{cases} \quad (5)$$

认为 PN 网络和 FTN 网络重要性相同,则设置 $\sigma = 0.5$ 计算整个多阶段损失函数。

4 实验与结果分析

算法在最新的 DETRAC^[10] 车辆数据库上训练模型。数据库由 140 000 张捕捉帧和 1 200 000 标记了的车辆组成。包含了 84 000 张训练用的图片,而且进一步将这些图片划分成 56 000 张训练图片和 28 000 张验证图片。另外在晴天、阴天、雨天或者晚上等场景下采集视频,平均每个视频帧包含 8.6 辆车辆。

实验一 表 1 显示了用控制变量法关闭不同组件评估检测效果的对比实验。给出了整个网络的平均检测质量 (mean average precision, mAP) 以及不同情况下的 mAP。PN + FTN + Concat 表示不使用特征图融合 (融合第一、三、五层卷积层的特征图,产生超特征) 而是只用到了最后一层卷积层输出的特征图。PN + FTN + Hyper 表示关闭了两个阶段的特征串联。整个模型的结构是 PN + FTN + Hyper + Concat。算法还与 Faster RCNN 进行比较,Faster RCNN 可以看作是使用 PN 网络产生候选车辆,但直接回归车辆检测的结果,而且没有微调检测框。另外,Faster RCNN 也没有涉及使用特征融合和多阶段特征串联。

表 1 DETRAC 数据库上不同算法组合的 mAP 和运行时间

算法设定	整体	晴天	多云	雨天	夜晚	运行时间/ms
Faster RCNN	68.74	62.94	70.46	80.86	60.32	90
PN + FTN + Concat	73.84	69.65	74.23	81.97	67.59	79
PN + FTN + Hyper	82.84	86.09	83.80	84.95	70.01	110
PN + FTN + Hyper + Concat	84.43	87.48	85.88	85.65	70.86	120

从实验结果可以看出,提出的整体模型在性能上取得了显著的提升:1) 相比于 Faster RCNN,该模型取得了 15.96% mAP 上的性能提升。其中,特征融合是至关重要的,到了 14.1% 的增益。此外,多阶段特征连接也使得 mAP 提高了 1.59%;2) 虽然特征融合和特征连接带来了额外的运行时间,是整体的运行速度还是可以接受的。

实验二 将改进的算法和一些先进的算法在 DETRAC 数据集上进行了 mAP 和检测速度比较,实验结果如表 2 所

示。总体上,本文对比其他算法模型的模型表现最好。可以看出:改进的算法框架取得了显著的性能提升。相比于 RCNN 算法,该算法取得了 19.41% mAP 上的性能提升。而相比于 Faster RCNN,则是 9.91%。在检测速度上,该算法虽然略慢于 Faster RCNN,但在可接受范围内,且相比于其他算法取得了很大的提升。

表 2 不同算法在 DETRAC 测试集上的 mAP 和检测速度

模式	综合	简单	中等	困难	晴天	多云	雨天	夜晚	检测速度/(s/幅)
DPM	25.70	34.42	30.29	17.62	24.78	30.91	25.55	31.77	6.00
ACF	46.35	54.27	51.52	38.07	58.30	35.29	37.09	66.58	1.50
RCNN	48.95	59.31	54.06	39.47	59.73	39.32	39.06	67.52	10.00
Faster RCNN	58.45	82.75	63.05	44.25	62.34	66.29	45.16	69.85	0.09
本文	68.36	90.69	72.37	52.58	75.89	74.42	54.28	82.92	0.12

实验三 图 2 给出了 DETRAC 测试集上的检测结果,图 3 给出了本文自行采集的测试集上的检测结果。算法成功地检测到了不同外观下的大多数车辆,特别是交通堵塞或者车辆远离相机的情况。当然算法也有一些出错的情况。如车辆被分割成多个箱子,或未能识别出多个相邻的车辆。



图 2 DETRAC 测试集上的检测结果

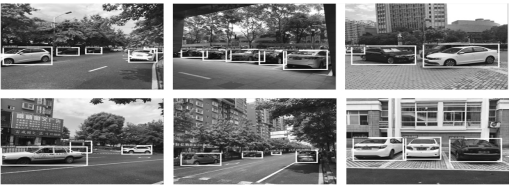


图 3 自行采集的测试集上的检测结果

5 结 论

提出了一种改进的物体检测框架,并将该框架应用于车辆检测。在这个框架中,多个神经网络的串联以及特征图的融合改善了候选框的质量。经实验研究发现,利用不同层的特征融合技术能更精准地实现物体的检测与定位。通过分析和优化检测过程每步,改进算法在 UA-DETRAC Benchmark 数据集上取得了较高的准确率。在自行采集的车辆数据集上进一步验证了该算法的有效性。

参考文献:

[1] 缪小冬,李舜韶,沈 恒,等.路面车辆的视觉检测方法[J].传感器与微系统,2010,29(2):5-8.
[2] Ren S, Girshick R, Girshick R, et al. Faster R-CNN: Towards real-time object detection with region proposal networks [J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2017, 39(6): 1137-1149.

(下转第 160 页)

b 为 0.5 m, 高 c 为 0.5 m, 送风量 L_0 为 $22 \text{ m}^3/\text{h}$, 换气次数 N 为 30 次/h。其中试验时室温温度为 26.3°C 左右, 实验设定温度为 23.0°C , 每隔 1 min 记录 1 次温度数据。实验结果如图 8。

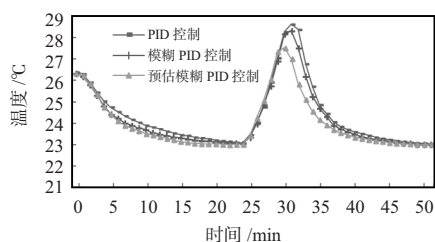


图8 3种控制方式实测温度曲线

从 0 min 开始计时, 预估模糊 PID 控制仅 17 min 后达到稳定值, PID 控制和模糊 PID 控制分别需要 23 min 和 21 min。当温度稳定后, 为模拟数据中心主机高负荷扰动, 从第 25 min 开始, 使用热风枪向模拟室内通入 45.0°C 热空气施加干扰, 持续时间 6 min。

在热空气干扰下, PID 控制温度最大值为 28.6°C , 超调量 24.5%, 停止干扰后 19 min 温度恢复稳定值; 模糊 PID 控制温度最大值为 28.3°C , 超调量 23.2%, 恢复时间为 17 min; 而预估模糊 PID 控制温度最大值仅为 27.5°C , 超调量 19.8%, 仅有 15 min 后就恢复温度稳定值。

与常规控制方式相比, 预估模糊 PID 控制调节时间分别减少 6 min 和 4 min, 高负荷干扰下温度波动降低 1.1°C 和 0.8°C , 恢复时间分别加快 21.0% 和 11.7%。实验结果表明, 该控制器响应能力更强, 能更好地改善系统动态性能, 在数据中心大负荷工作下比传统控制方式有更好表现。

4 结 论

针对数据中心机房高负荷运行的复杂工况下温度难以控制的问题, 设计了一种适合数据中心机房的 Smith 预估模糊 PID 控制器, 实验结果表明: 与常规控制方式相比, 预估模糊 PID 控制调节时间更少、响应速度更快、抗高负荷扰动能力更强, 可以一定程度上解决数据中心高负荷下温度控制问题。

(上接第 155 页)

- [3] Ghodrati A, Diba A, Pedersoli M, et al. DeepProposal: Hunting objects by cascading deep convolutional layers [C]// IEEE International Conference on Computer Vision, 2015: 2578 - 2586.
- [4] Ranjan R, Patel V M, Chellappa R. HyperFace: A deep multi-task learning framework for face detection, landmark localization, pose estimation, and gender recognition [J]. IEEE Transactions on Pattern Analysis & Machine Intelligence 2016, 17(7): 1.
- [5] Kong T, Yao A, Chen Y, et al. HyperNet: Towards accurate region proposal generation and joint object detection [C]// Conf on Computer Vision & Pattern Recognition 2016: 845 - 853.
- [6] Ioffe S, Szegedy C. Batch normalization: Accelerating deep network training by reducing internal covariate shift [C]// International Conference on Machine Learning 2015: 448 - 456.
- [7] Felzenszwalb P F, Girshick R B, Mcallester D, et al. Object de-

参考文献:

- [1] Lü H L, Duan P Y, Jia L. One novel fuzzy controller design for HVAC systems [C]// Chinese Control and Decision Conference, 2008: 2071 - 2076.
- [2] 游志宇, 陈维荣, 彭 赞, 等. 大功率空冷自增湿 PEMFC 温度控制方法 [J]. 西南交通大学学报 2015, 50(1): 44 - 50.
- [3] 李丽娜, 柳洪义, 罗 忠, 等. 模糊 PID 复合控制算法改进及应用 [J]. 东北大学学报: 自然科学版, 2009, 30(2): 274 - 278.
- [4] 朱万林, 陈懿华, 沙立民. 基于模糊控制技术的中央空调控制器的仿真研究 [J]. 工业仪表与自动化装置 2008(2): 57 - 59.
- [5] 段英宏. 空调房间温度预估模糊 PID 控制器的研究 [J]. 系统仿真学报 2008, 20(3): 620 - 622.
- [6] 叶大法, 杨国荣. 变风量空调系统设计 [M]. 北京: 中国建筑工业出版社 2007.
- [7] 王建明, 李训铭. 变风量系统空调房间建模与特性参数估算 [J]. 计算机仿真 2002, 19(4): 69 - 72.
- [8] 施俊良. 室温自动调节原理和应用 [M]. 北京: 中国建筑工业出版社 1983.
- [9] Zhang X, Wang R, Xing J. VAV air-conditioning system control of underground defensive project based on fuzzy predictive PID algorithm [C]// Chinese Control and Decision Conference 2016: 6286 - 6291.
- [10] 徐 凯, 李 琦. 变风量空调的自适应模糊 PID 复合控制 [J]. 计算机仿真 2011, 28(10): 151 - 155.
- [11] 修春波, 夏琳琳. 智能控制技术 [M]. 北京: 中国水利水电出版社 2013.
- [12] 陈小强, 马泳娟, 侯 涛, 等. 预测模糊 PID 在高速列车速度控制中的研究 [J]. 系统仿真学报 2014, 26(1): 191 - 196.
- [13] Choi O K, Kim J, Kim J, et al. BIBO stability analysis of TSK fuzzy PI/PD control systems [J]. Intelligent Automation & Soft Computing 2015, 21(4): 1 - 14.
- [14] 周向阳, 贾 媛. 航空遥感惯性稳定平台模糊/PID 复合控制 [J]. 仪器仪表学报 2016, 37(11): 2545 - 2554.

作者简介:

谢洁飞(1968 -), 男, 博士, 副教授, 研究生导师, 主要研究领域为制冷系统与空调, E-mail: nhxiejf@126.com。

tection with discriminatively trained part-based models [J]. IEEE Transactions on Pattern Analysis & Machine Intelligence 2010, 32(9): 1627 - 1645.

- [8] Girshick R, Donahue J, Darrell T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation [C]// IEEE Conference on Computer Vision & Pattern Recognition, 2014: 580 - 587.
- [9] Girshick R. Fast R-CNN [C]// IEEE International Conference on Computer Vision 2015: 1440 - 1448.
- [10] Wen L, Du D, Cai Z, et al. UA-DETRAC: A new benchmark and protocol for multi-object detection and tracking [J]. Computer Science 2016, 27(6): 18.

作者简介:

朱锋彬(1994 -), 男, 硕士, 主要研究方向为信号与信息处理。
应 娜(1994 -), 女, 博士, 副教授, 主要研究领域为信号与信息处理, 图像和语音处理。