

基于深度学习的目标检测综述

李 丹

(河北地质大学 信息工程学院, 河北 石家庄 050031)

【摘要】目标检测是计算机视觉中的一个重要问题,近年来传统检测方法已难以满足人们对目标检测效果的要求,随着深度学习在图像分类任务上取得巨大进展,基于深度学习的目标检测算法逐渐成为主流。对主流卷积神经网络的发展作一简述,并介绍几种基于深度学习的目标检测算法,同时对其存在的问题及未来可能的发展方向进行讨论。

【关键词】目标检测;计算机视觉;深度学习;卷积神经网络

【中图分类号】TP181 **【文献标识码】**A **【文章编号】**2096-1995(2019)13-0001-03

目标检测将目标定位和目标分类结合起来,随着计算机技术的发展,目标检测已在人脸识别、军事国防、人工智能等领域得到广泛应用。

现有的目标检测方法有传统检测方法和基于深度学习的检测方法。传统检测方法一般分三个阶段:首先在给定图像上采用不同大小的滑动窗口对整幅图像进行遍历选择候选区域;然后对这些区域提取特征;最后使用分类器进行分类。虽然这种方法取得了不错的结果,但是采用滑动窗口进行区域选择时没有针对性导致时间复杂度高且窗口冗余;另外手工设计的特征没有很好的鲁棒性。自从2012年AlexNet在ILSVRC比赛中获得冠军后,基于深度学习的目标检测逐渐崛起。

1 主流卷积神经网络的发展

1.1 LeNet-5

LeNet-5于1998年提出,主要用于手写数字识别。该网络共有7层网络,包括2个卷积层、2个池化层和3个全连接层,其准确率很高,但由于当时计算机的计算能力及数据量的不足,该模型未能受到重视,从而没有很好地发展起来。

1.2 AlexNet

直到2012年ILSVRC比赛中,AlexNet在算法准确率方面以超过使用传统方法的第二名11%的情况获得冠军后,才使卷积神经网络重回大众视野并得到重视。该网络是对LeNet-5的扩展,它将卷积神经网络的思想应用到了更深更宽的网络上,共有8层网络,包括5个卷积层和3个全连接层。

1.3 VGGNet

VGGNet则充分代表了从2012年到2014年在卷积神经网络结构上的一个进展。相比AlexNet来说该网络会更深一些,共有19层网络:16个卷积层和3个全连接层。该网络在2014年ILSVRC比赛中获得了第二名,在目标检测上获得了第一名。

1.4 GoogleNet

GoogleNet则是2014年ILSVRC比赛中获得第一名的网络。该网络通过增加网络的深度和广度来获取更好的结果,然而当网络加深加宽到一定程度时,继续加深加宽反倒不能再提高效果。因此,GoogleNet通过设计稀疏连接的Inception结构来解决这个问题。

1.5 ResNet

2015年ILSVRC比赛中ResNet网络获得冠军。同样该网络也是用于解决网络加深到一定程度就不能提高效果的问题,从而使得网络能够继续加深到更深的层次。

2 基于深度学习的目标检测算法

当前基于深度学习的目标检测算法主要有两类:一类是结

合region proposal和CNN的基于分类的R-CNN系列检测算法;一类是基于回归的检测算法。

2.1 基于分类的检测算法

2.1.1 R-CNN

2014年Girshick等人设计了R-CNN框架,使目标检测取得巨大突破。其流程如图1所示,首先使用Selective search算法在输入图像中提取大约2000个region proposal;其次为使CNN有标准的输入需要将region proposal的大小进行归一化;然后使用AlexNet提取特征;最后利用多个SVM进行分类,同时使用线性回归微调定位框。

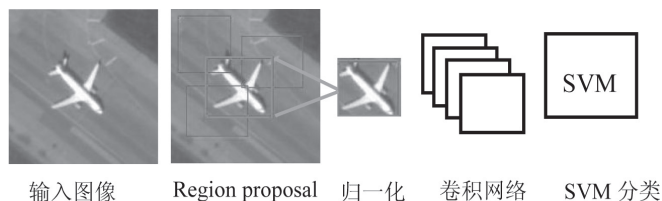


图1 R-CNN 结构示意图

R-CNN的应用使得在PASCAL VOC2007上的检测结果提升到了66%。然而R-CNN要对2000个左右的proposal分别进行特征提取,计算量大且处理速度慢。此外该算法将训练分为多个阶段导致步骤繁琐。

2.1.2 SPP-Net

为得到更快速更准确的算法,He等人提出了SPP-Net。如图2所示,针对R-CNN对每个proposal分别提取特征的缺点,SPP-Net对整张图像只进行一次卷积提取特征操作,后将region proposal在原图的位置映射到卷积层特征图上。此外为避免对region proposal进行归一化操作,SPP-Net在最后一个卷积层后加入了空间金字塔池化层(SPP),用于提取固定尺寸的特征向量。

SPP-Net使得检测速度得到了很大提升,然而同R-CNN一样,训练步骤依然繁琐。

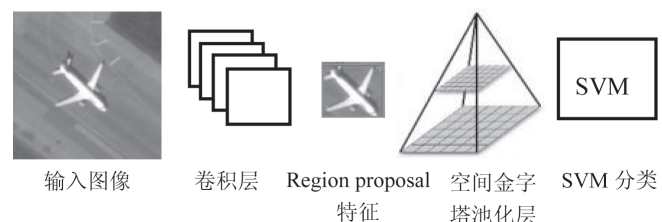


图2 SPP-Net 结构示意图

基金项目:河北省研究生创新资助项目(CXZZSS20181)。

作者简介:李丹(1993-),女,山西运城人,河北地质大学在读硕士研究生,主要从事机器学习与计算机视觉研究。

2.1.3 Fast R-CNN

2015年Ren等人提出Fast R-CNN。该算法在SPP-Net的基础上,使用一个简化的SPP层,即ROI Pooling层。同时使用SVD分解全连接层的输出,在得到softmax分类得分的同时进行边框回归。

Fast R-CNN使用softmax代替SVM,使用多任务损失函数将边框回归加入到网络中,使整个训练过程不再分多个阶段进行。使用VGG16代替AlexNet,使平均准确率达到70%的同时加快了训练速度。然而在选取region proposal时,Fast R-CNN依然使用计算量大的Selective Search方法。

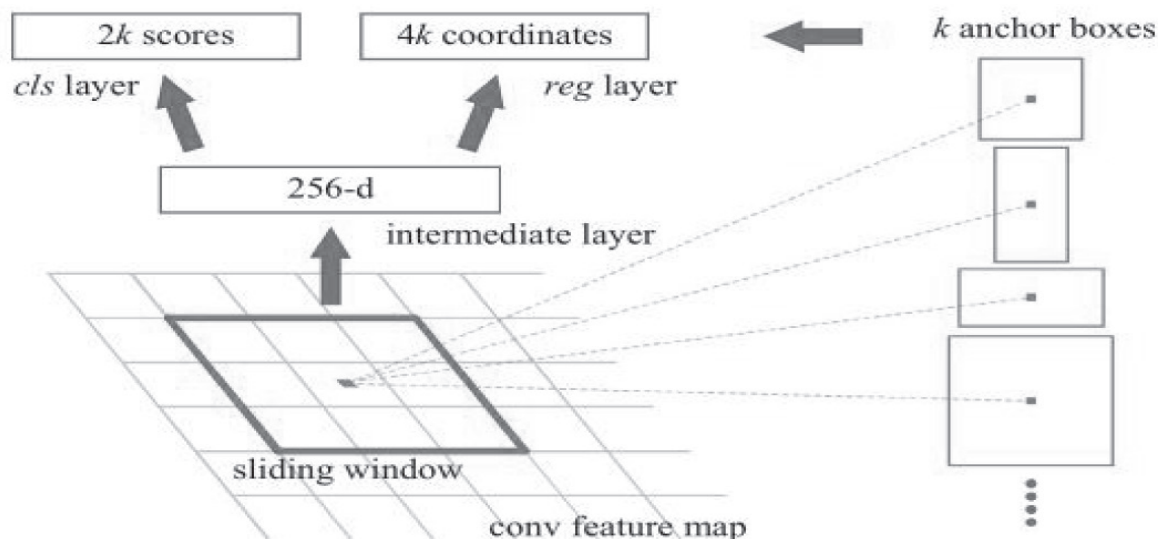


图3 RPN网络结构示意图

2.1.5 Faster R-CNN 改进算法

如图4所示,从R-CNN到Faster R-CNN,始终是将目标检测问题的关注点集中到分类上,采用region proposal与CNN feature结合的方法。整体看来这一系列算法在速度与精度上都在不断提高,此类算法仍然是当前的一大主流。

此后陆续出现了不少对Faster R-CNN的改进算法,如:拥有更好特征网络的HyperNet^[1]、MS-CNN^[2]和Light-Head R-CNN^[3];拥有更完善ROI pooling的R-FCN^[4]、Mask R-CNN^[5]和Cascade R-CNN^[6];拥有更精准RPN的FPN^[7]和CRAFT^[8];以及对样本后处理的OHEM^[9]、Soft-nms^[10]和A-Fast-RCNN^[11];还有更大mini-Batch的MegDet^[12]等改进算法。



图4 R-CNN、Fast R-CNN、Faster R-CNN对比

2.2 基于回归的检测算法

2.2.1 YOLO

Faster R-CNN系列算法在速度与精度上虽已取得了不错的提高,然而还尚未达到实时的效果。这类算法始终将目标检测分为选取region proposal及分类两步,而训练两个模型必然导致检测速度降低,因此2016年Redmon等人提出了YOLO算法,该算法不需要中间选取region proposal的操作,而是直接从给定图像的多个位置回归出目标框及目标类别。该算法在速度上

2.1.4 Faster R-CNN

为解决使用Selective Search计算量大的问题,Ren等人提出了Faster R-CNN。该算法引入RPN网络(Region Proposal Networks)来代替Selective Search,从而达到真正的端到端。

RPN本质还是滑动窗口,如图3所示为使用RPN产生region proposal的操作,使用预先设置的不同尺度不同长宽比的anchor映射到原图,得到region proposal。此后,使用与Fast R-CNN相同的结构进行分类及边框回归。RPN的使用,使Faster R-CNN能够在一个网络中完成选取region proposal、特征提取、分类及定位多个步骤。

取得了极大提高,精度却有所下降。因此Redmon等人对其进行了改进,同年提出YOLO9000^[13]算法,该算法在保持其速度的同时提升了精度。在2018年,Redmon等人又提出了速度与精度均有保障的YOLOv3^[14]算法。

2.2.2 SSD

2016年Liu等人提出了SSD^[15]算法。该算法结合了YOLO的回归思想以及Faster R-CNN的anchor机制,使得检测速度和精度均得到了保证。具体表现同YOLO一样通过回归获取目标的位置和类别。只不过YOLO是使用全图特征来预测某一位置,而SSD是使用某一位置周围的特征来进行预测。通过Faster R-CNN的anchor机制建立某一位置与其特征之间的对应关系。该算法既保持了YOLO速度快的优点,又像Faster R-CNN一样精准。

3 结语

目前已有众多基于深度学习的目标检测算法,且其相关研究仍在不断进行,几乎每年都有最新成果出现。即便如此,当前这些方法的检测效果仍处于一个较低的水平,从而导致不能广泛应用于实际通用的检测任务当中。此外,尽管当前基于深度学习的检测算法成为发展的主流,但传统的检测方法仍在使用,并未完全被抛弃,或许将传统方法有效应用于深度学习方法中会对检测效果的提高有所帮助。

【参考文献】

- [1]Kong T, Yao A, Chen Y, et al. HyperNet: Towards Accurate Region Proposal Generation and Joint Object Detection[J].2016.
- [2]Cai Z, Fan Q, Feris R S, et al. A Unified Multi-scale Deep Convolutional Neural Network for Fast Object Detection[J].2016.
- [3]Li Z, Peng C, Yu G, et al. Light-Head (下转P31)

压力,沿纵向施加到前端封头网格上。 P 可以取设计最大允许充装量或总重 R 与自重 T 的差值。对于最大允许充装量 $<(R-T)$ 的情况,建议按照 $(R-T)$ 进行货物惯性力或货物重力的加载,当最大允许充装量发生增大的变更时,由于所施加的载荷已经是最大的,故不需要重新进行分析。

竖直向下 $2R_g$ 工况,投影面积为内容器沿竖直向下的投影面积(等于沿筒体径向的投影面积),货物惯性力为 $2Pg$ 。等效均布压力沿竖直向下方向施加在内容器下半部网格上。

横向 $1R_g$ 工况,投影面积为内容器沿横向的投影面积(等于沿筒体径向的投影面积),力为 $1Pg$ 。等效均布压力沿横向施加在内容器侧半部网格上。

竖直向上 $1R_g$ 工况,投影面积为内容器沿竖直向上的投影面积(等于沿筒体径向的投影面积),力为 $1Pg$ 。等效均布压力沿竖直向上方向施加在内容器上半部网格上。

货物重力,可以参照上述货物惯性力的施加方法进行,即等效均布压力等于货物重力 $1Pg$ 除以内容器沿竖直向下的投影面积,然后沿竖直向下方向施加在内容器下半部网格上。另一方法是向内容器施加沿单元法线方向的梯度压力,即内容器壁任意点所受的压力(压强)等于货物密度与 g 及深度的乘积,方向朝向内容器外部。货物密度可以简化为等效密度,即货物重量 P 除以内容器的几何容积 V_G ,此时深度为任意点与内容器壁最高点的竖直距离。

需要注意的是纵向和横向这两个惯性力工况,重力作用与惯性力作用同时存在,因此需要施加箱体重力和货物重力。竖直向下和竖直向上这两个工况,不应考虑重力作用,即不施加箱体重力和货物重力。竖直向上工况,若施加重力作用,箱体惯性力被平衡掉;内容器的下半部和上半部同时承受数值相等,方向相反的力,对于框架、裙圈、外罐体和支撑结构而言,货物惯性力被平衡掉,合力为零。竖直向下工况,亦不需要再额外施加箱体重力和货物重力。若施加的话,惯性力和重力的合力为 $3R_g$,分析结果将因此变大。

纵向和横向惯性力工况是否施加货物重力,即 $1Pg$ 竖直向下的合力,对支撑结构及其附近区域、外罐体与框架连接位置的应力结果有较大的影响,而惯性力工况计算的关键正在于此,故不应忽略货物重力。这一点应与货物重力(静压头)对按规定设计计算容器壁厚的影响加以区别对待,计算容器壁厚时,货物重力按照液体静压力来考虑,且假设最大值均匀分布于整

个容器,即取最深处的液体静压力叠加到其它内压载荷上进行计算,由于所占比例小,对容器壁厚计算影响很小。当液柱静压力小于设计压力的5%时,可忽略不计^[5]。

注满水的重力,即水的密度 1000kg/m^3 与几何容积 V_G 及重力加速度 g 的乘积,施加时可以参照货物重力的两种施加方法进行,此处不再赘述。

堆码力,施加时根据设计堆码载荷值,考虑 $1.8g$ 加速度,偏心施加于顶角件顶面,横向偏心 25.4mm ,纵向偏心 38mm 。若设计堆码载荷值为 192000kg , $192000\text{kg}\cdot 1.8\cdot 9.81\text{m/s}^2=3390\text{kN}$,则每个顶角件承受的堆码力为 848kN ^[6]。由于框架结构总体上是 $1/4$ 对称的,故分析时可以只考虑一种整体偏心状态,横向向同一侧偏心,纵向向同一端偏心。堆码载荷是否偏心施加对该工况结果影响较大,故分析时应按规范和标准进行偏心加载。

4 结语

依据相关规范和标准,结合集装箱设计、分析和审图经验,对LNG罐箱有限元强度分析的方法和要求,作了详细深入的研究总结。所述的惯性力工况、堆码工况应考虑的载荷以及各载荷的施加方法,是符合相关规范和标准要求的,是保证合理、可靠地进行LNG罐箱有限元强度分析的基础。一些方法和要求也适用于其它类型罐箱的有限元强度分析,期望能对LNG罐箱及其它类型罐箱的设计、强度分析工作产生积极意义。

【参考文献】

- [1] 国家质量监督检验检疫总局.TSG R0005-2011 移动式压力容器安全技术监察规程[S].北京:新华出版社,2011:12-13.
- [2] 全国锅炉压力容器标准化技术委员会.NB/T 47059-2017 冷冻液化气体罐式集装箱[S].北京:新华出版社,2017:7-8.
- [3] 中国船级社.集装箱检验规范 2016[S].北京:人民交通出版社,2016:83-85.
- [4] IMO. IMDG CODE (amendment 38-16) International Maritime Dangerous Goods Code[S]. London, 2016: 409-410.
- [5] 全国锅炉压力容器标准化技术委员会.GB/T 150.1~150.4-2011 压力容器[S].北京:中国标准出版社,2012:10-12.
- [6] ISO. ISO 1496-3:1995(E) Series 1 freight containers-Specification and testing-Part 3: Tank containers for liquids, gases and pressurized dry bulk[S]. Switzerland, 1995:419-422.

(上接P2) R-CNN: In Defense of Two-Stage Object Detector[J].2017.

[4] Dai J, Li Y, He K, et al. R-FCN: Object Detection via Region-based Fully Convolutional Networks[J].2016.

[5] He K, Gkioxari G, Dollár P, et al. Mask R-CNN[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2017, PP(99):1-1.

[6] Cai Z, Vasconcelos N. Cascade R-CNN: Delving into High Quality Object Detection[J].2017.

[7] Lin T Y, Dollár P, Piotr, Girshick R, et al. Feature Pyramid Networks for Object Detection[J].2016.

[8] Yang B, Yan J, Zhen L, et al. CRAFT Objects from Images[C]// Computer Vision & Pattern Recognition.2016.

[9] Shrivastava A, Gupta A, Girshick R. Training Region-Based Object Detectors with Online Hard Example

Mining[J].2016:761-769

[10] Bodla N, Singh B, Chellappa R, et al. Improving Object Detection With One Line of Code[J].2017.

[11] Wang X, Shrivastava A, Gupta A. A-Fast-RCNN: Hard Positive Generation via Adversary for Object Detection[J].2017:3039-3048.

[12] Peng C, Xiao T, Li Z, et al. MegDet: A Large Mini-Batch Object Detector[J].2017.

[13] Redmon J, Farhadi A. YOLO9000: Better, faster, stronger[C]// Proceedings of CVPR 2016, 2016.

[14] Redmon J, Farhadi A. YOLOv3: An Incremental Improvement[J].2018.

[15] Liu W, Anguelov D, Erhan D, et al. SSD: Single Shot MultiBox Detector[C]// European Conference on Computer Vision.2016.