

基于深度学习的油库安全监测方法研究

林海, 胡旭晓, 吴跃成, 汪燕超

(浙江理工大学机械与自动控制学院, 浙江 杭州 310018)

✉2049535531@qq.com; huxuxiao@zju.edu.cn; wuyuechen@126.com; 2597123426@qq.com



摘要: 存储石油化工产品时有极高的安全监测要求, 传统油库监测系统采用人工观察视频监控的方式检测产品的存储是否存在安全问题, 该方法在火灾应急行动中存在时效性低、响应慢及漏检等问题。针对上述问题, 文章提出了一种基于深度学习的油库安全监测方法, 在原始YOLOv5(单阶段目标检测算法)的主干网络添加坐标注意力机制(Coordinate Attention, CA), 该机制可以得到方向感知和位置感知的注意力图, 将这些图应用于输入特征图, 可以更加丰富地表示感兴趣对象; 使用回归损失函数EIOU_Loss代替原有的GIOU_Loss进行预测框的定位回归损失计算, 该方法能提高Bounding Box(边界框)的回归精度; 使用DIOU-NMS替换NMS(非极大抑制), 可以提高对遮挡目标的辨识度。该方法在自制的测试数据集中的mAP达到了91.2%, 说明该方法在油库安全监测领域具有较高的应用价值。

关键词: 油库; 目标检测; YOLOv5; 安全监测

中图分类号: TP311 **文献标识码:** A

Research on Safety Monitoring Method of Oil Depot based on Deep Learning

LIN Hai, HU Xuxiao, WU Yuecheng, WANG Yanchao

(School of Mechanical Engineering and Automation, Zhejiang Sci-Tech University, Hangzhou 310018, China)

✉2049535531@qq.com; huxuxiao@zju.edu.cn; wuyuechen@126.com; 2597123426@qq.com

Abstract: Petrochemical products storage requires extremely high level of safety monitoring. Traditional oil depot monitoring system uses manual observation video monitoring to detect whether there are safety problems. This method has problems such as low timeliness, slow response and missed detection. In response to these problems, this paper proposes an oil depot safety monitoring method based on deep learning. The Coordinate Attention (CA) mechanism is added to the backbone network of original YOLOv5 (a single-stage target detection algorithm), so as to obtain the attention maps of direction awareness and position awareness. These maps are applied to input feature maps so that the objects of interest are represented more abundantly. Then, the original GIOU_Loss is replaced by EIOU_Loss function to calculate the localization regression loss of the prediction box, which can improve the regression accuracy of Bounding Box. Finally, DIOU-NMS is used to replace NMS (Non-Maximum Suppression) to improve the recognition of occluded targets. The mAP of this method in the self-made test dataset reaches 91.2%, indicating that the proposed method has high application value in the safety monitoring of oil depots.

Keywords: oil depot; target detection; YOLOv5; safety monitoring

1 引言(Introduction)

石油化工行业是我国最大的基础性制造行业, 石油化工产品储存、运输和生产各个环节均有极高的安监、环保、应急管理要求^[1]。油库储备的原油具有易燃、易爆、易挥发、易产生静电等危险特性, 安全风险非常高, 因此加强对油库的安全监测, 及时制止工作人员错误行为和发现安全隐患显得尤为重要。随着大数据和人工智能技术的发展, 在视频监控中加入对视频的智能监测, 实现对视频源实时分析, 可以

避免遗漏信息, 使各种安全风险在发生之初就被发现, 并提醒工作人员及时解决, 从而避免安全事故发生, 不但降低了工作人员的工作强度, 还有效地提高了油库安全监测的效果和质量。

近年来, 深度学习飞速发展, 计算机视觉领域中很多优秀的算法都是基于深度学习的。目标检测领域主流的方法目前分为两类: 一类是基于候选框的两阶段算法, 如RCNN^[2]、Fast-RCNN^[3]、Faster-RCNN^[4]、Mask-RCNN^[5]等。另

一类是基于边框回归的一阶段算法,如SSD^[6]和YOLO^[7]系列算法。两阶段目标检测算法精度高,但是检测速度相对较慢。与两阶段目标检测算法相比,一阶段目标检测算法速度远快于两阶段目标检测算法,经过许多研究者的不断努力,目前一阶段目标检测算法基本能达到两阶段目标检测算法的精度,所以实际应用中主要采用一阶段目标检测算法,例如YOLO系列目标检测算法。

本文选用目前表现优秀的YOLOv5目标检测算法为基础,通过加入CA^[8]提取更加丰富的特征,同时修改边框回归损失函数以提升边框回归精度,最后在输出结果中用DIOU-NMS替换传统的NMS,提高对遮挡目标的辨识度。目标检测对象主要为油库工作人员是否佩戴安全头盔,以及烟雾和火焰。实验结果表明,改进后的YOLOv5算法比原始的YOLOv5算法的 mAP 高2.9%。

2 YOLOv5模型结构(YOLOv5 model structure)

YOLOv5是Ultralytics团队于2020年提出的,虽然YOLOv5并没有对YOLO算法系列进行全新的模型架构重建,但是YOLOv5在目标检测领域中的应用表现优异。YOLOv5在网络体积上分为YOLOv5s、YOLOv5m、YOLOv5l、YOLOv5x四个不同的版本。YOLOv5各个型号的网络模型结构大致相同,通过不同的宽度系数和深度系数控制模型体积的大小。本文以最小的网络模型YOLOv5s为例,介绍YOLOv5的模型结构,其结构主要由输入端、主干网络、Neck网络和输出端四个部分组成,详细网络结构图如图1所示。

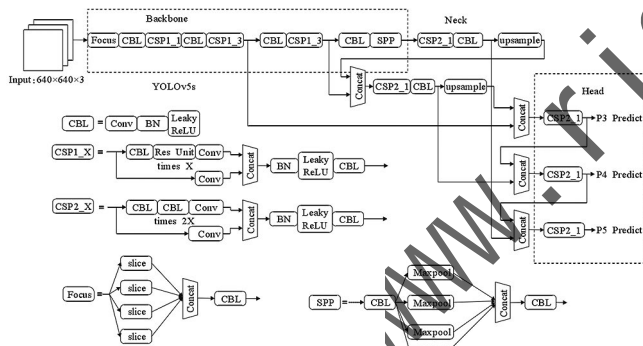


图1 YOLOv5网络结构图

Fig.1 YOLOv5 network structure diagram

YOLOv5的输入端和YOLOv4一样采用了Mosaic数据增强,Mosaic数据增强就是随机选取4张图片采用随机缩放、随机裁剪、随机排布的方式将其拼接为一张图片用于训练。这种方法不但丰富了数据集,而且能随机缩放,增加了很多小目标,在一定程度上提高了网络对小目标的检测,大大减少了GPU内存的消耗,即使使用单张GPU训练,也能达到很好的效果。同时,输入端还使用自适应锚框主动学习并调整锚框大小和自适应图片缩放来减少信息冗余并提升推理速度。

相比YOLOv4,YOLOv5的主干网络在前面添加了一个Focus结构,主要用于切片操作。以YOLOv5s为例,原始 $608 \times 608 \times 3$ 的图片经过Focus结构切片后,先得到 $304 \times 304 \times 12$ 的特征图,然后经过一次输出通道为32的卷积操作,最终得到 $304 \times 304 \times 32$ 的特征图。与YOLOv4相比,本

研究在YOLOv5中设计了两种CSP结构,CSP1_X结构用于主干网络,CSP2_X结构用于Neck网络中。

YOLOv5的Neck网络采用PAN结构和FPN结构,PAN层自底向上传达强定位特征,FPN层自顶向下传达强语义特征,把这两种结构结合在一起,不同的主干层对不同的检测层进行参数聚合,进一步提高特征提取能力。

YOLOv5的输出端采用GIOU_Loss作为Bounding box的回归损失函数,GIOU_Loss缓解了IOU_Loss边界框不重合情况下的损失计算问题,最后使用非极大抑制NMS,进而选取最优的目标框。

3 网络结构改进(Network structure improvement)

3.1 添加坐标注意力模块

大部分注意力机制用于深度神经网络都可以带来不错的性能提升,例如使用Self-Attention,但是这种注意力带来的计算复杂度对轻量级网络而言是无法承受的。所以,在轻量级网络上主要使用SE(Squeeze-and-Excitation)和CBAM(Convolutional Block Attention Module),但SE只考虑内部通道信息,忽略了位置信息,而计算机视觉任务中目标的空间结构是十分重要的。CBAM提出通过在通道上进行全局池化来捕获位置信息,但是这种方式只能捕获局部的信息,而无法获得大范围相互依赖的信息。

坐标注意力是新加坡国立大学的HOU等^[8]提出了一种为轻量级网络设计的新的注意力机制。CA(Coordinate Attention)是一种新的高效注意力机制,其原理是将位置信息嵌入到通道注意力中,使得轻量级网络能够在更大的区域进行注意力,同时减轻了大量的计算,降低了计算的复杂度。为了避免2D全局池化造成的位置信息丢失,研究人员将通道注意力分解为两个并行的1D特征编码过程,有效地将空间坐标信息整合到生成的注意力图中。CA模块结构如图2所示。

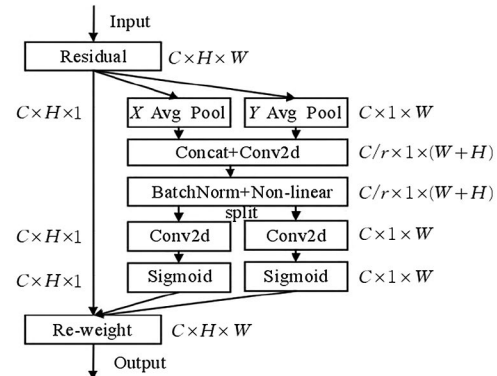


图2 CA模块结构示意图

Fig.2 CA module structure diagram

CA模块工作流程如下,首先假定输入的特征图大小为 $C \times H \times W$,分别用 $H \times 1$ 和 $W \times 1$ 大小的池化核沿着水平方向与垂直方向对每个通道进行编码,得到水平方向和垂直方向的特征图 z^h 和 z^w , z^h 的大小为 $C \times H \times 1$, z^w 的大小为 $C \times 1 \times W$ 。其次将上述两个带有方向信息的特征图进行拼接且利用权值共享的 1×1 卷积和进行卷积操作,经过一个非线性激活函数输出特征图 $f \in R^{(C/r) \times 1 \times (H+W)}$, r 为卷积过程中下采样的比例。最后将 f 在通道进行拆分,分别经过 1×1 卷积核进行通

道调整, 得到 $f^h \in R^{C \times H \times l}$ 和 $f^w \in R^{C \times l \times W}$, 再经过Sigmoid激活函数得到两个不同空间方向的注意力权重 g^h 和 g^w ^[9]。将输入特征图与上述两个注意力权重相乘, 从而得到表现能力更强的特征图。

CA是一个简单的即插即用的轻量级网络注意力机制, HOU等^[8]在EfficientNet和MobileNet加入了CA模块, 在图像分类任务上均取得了不错的效果。同时, HOU等^[8]也在目标检测和语义分割任务上进行实验, 发现模型性能均得到不小的提升。所以, 本研究同样将CA模块添加到YOLOv5的主干网络中, 它既能捕获通道之间的依赖, 也能很好地捕获位置信息的空间长距离依赖^[10]。

3.2 使用EIOU Loss代替GIOU Loss

YOLOv5的边框回归损失函数使用的GIOU Loss, 其计算式如式(1)所示:

$$GIOU_{Loss} = 1 - IOU + \frac{|C - A \cup B|}{C} \quad (1)$$

式(1)中, IOU表示交并比, 即预测框和真实框相交区域面积与合并区域的面积比。IOU Loss为1减去IOU后的值, 当预测框和真实框不相交时, IOU的值为0, 这就导致在一定范围内损失函数没有梯度, 无法进行学习和训练。针对这个问题, 有研究者提出了GIOU Loss, GIOU Loss的原理是假定有两个框A和B, 然后找到一个最小的封闭形状C, 再计算C中没有覆盖A和B的面积与C总面积的比值, 最后用A和B的IOU减去这个值。GIOU Loss不仅关注重合区域, 还关注其他非重合区域, 能更好地反映二者的重合度。

边框回归的过程中需要考虑三个主要问题: 重叠面积、中心点距离和长宽比。无论是IOU Loss还是GIOU Loss都只考虑了重叠面积, 因此有学者提出DIOU Loss, 计算式如式(2)所示:

$$DIOU_{Loss} = 1 - IOU + \frac{\rho^2(b, b^{gt})}{c^2} \quad (2)$$

式(2)中, ρO 表示预测框和真实框两个中心点之间的欧氏距离, c 表示能够同时包含预测框和真实框的最小区域的对角线距离。

同样, DIOU Loss只考虑了重叠面积和中心点距离, 没有考虑到长宽比, CIOU Loss在DIOU Loss的基础上进行了一定的改进, 计算式如式(3)所示:

$$CIOU_{Loss} = 1 - IOU + \frac{\rho^2(b, b^{gt})}{c^2} + \alpha v \quad (3)$$

式(3)中, v 表示预测框与真实框长宽比的相似性:

$$v = \frac{4}{\pi^2} \left(\arctan \frac{w^{gt}}{h^{gt}} - \arctan \frac{w}{h} \right)^2 \quad (4)$$

α 为一个权重参数:

$$\alpha = \frac{v}{(1 - IOU) + v} \quad (5)$$

CIOU Loss虽然考虑了边框回归的重叠面积、中心点距离及纵横比, 但是公式中反映的是纵横比的差异, 而不是宽高与其置信度真实差异, 有时会阻碍模型有效的优化相似性。所以, 有学者在CIOU Loss的基础上将纵横比拆开, 提出EIOU Loss, 计算式如式(6)所示^[11]:

$$EIOU_{Loss} = 1 - IOU + \frac{\rho^2(b, b^{gt})}{c^2} + \frac{\rho^2(w, w^{gt})}{C_w^2} + \frac{\rho^2(h, h^{gt})}{C_h^2} \quad (6)$$

式(6)中, 前三项与DIOU Loss一样, 分别表示重叠损失和中心点距离损失, 后两项表示将纵横比拆分为预测的宽高分别与最小外接框宽高的差值, 这样使得收敛速度更快, C_w 和 C_h 表示覆盖两个边框的最小外接框的宽度和高度。

基于上述分析, 本文选择EIOU Loss作为边框回归的损失函数, 不但能加速收敛, 而且能提高边框回归精度。

3.3 使用DIOU-NMS替换NMS

目标检测最后的预测阶段会输出很多候选的anchor box, 其中同一个目标往往会有许多的接近重叠的预测框, 这时候就需要使用NMS挑选一个最好的边框保留。在原始的NMS中, 得分最高的框与周围其他框分别作IOU计算, 如果IOU大于一个给定的阈值, 则认为是同一个物体, 将删除这个边框。如果两个物体十分接近且给定的IOU阈值较高时, 经过NMS后就只剩下一个物体, 这样就造成漏检的情况。为了减少这种漏检现象, 本文用DIOU-NMS代替原始的NMS, DIOU-NMS不但考虑两个框体的IOU, 还考虑两框中心点的距离, 即使两个框的IOU较大时, 如果这两个框的中心距离也较大, 则可认为这是两个物体的候选框而不会被抑制掉, 这样就在一定程度上减少了漏检现象和提高被遮挡物体的辨识度。

4 实验及结果分析(Experiment and result analysis)

4.1 实验数据集和实验环境

由于油库视频监控中很少能采集到烟雾、火焰和工人信息, 所以从互联网上爬取需要的图片。爬取的图片可能未包含研究人员所需要的物体或者掺杂了许多的噪音, 所以需对图片进行初步筛选, 最终本文选取了4,000张图片作为标注数据集, 对标注数据集进行预处理, 如重命名和将图片统一转换成jpg格式, 使用labelImg对图片进行手工标注, 标注对象为烟雾(Smoke)、火焰(Fire)、佩戴安全帽(Hat)和未佩戴安全帽(Person)四类, 标注的数据集格式为PASCAL VOC格式。将标注好的数据集按8:2的比例随机划分为训练集和测试集, 最终有3,200张图片用于训练, 800张图片用于测试。

此次训练的模型是在YOLOv5s的基础上进行改进的, 在主干网络中添加CA模块, 使用EIOU Loss代替原来的GIOU Loss, 使用DIOU-NMS代替传统的NMS。本次实验在Ubuntu 18.04操作系统、NVIDIA RTX2080Ti显卡、CUDA版本为10.1、深度学习框架为PyTorch 1.9.0的环境下进行, 选取的batch_size为32, 训练轮数epochs为280。

4.2 实验结果

将800张测试集图放在YOLOv5测试文件夹中, 测试集中烟雾目标包含256个, 火焰目标包含186个, 佩戴安全帽目标包含891个, 未佩戴安全帽目标包含431个。分别用原始的YOLOv5s训练出的模型和改进后的YOLOv5训练出的模型对测试集图片进行检测, 最终危险目标识别结果如表1所示, 模型性能对比结果如表2所示。

表1 危险目标识别结果

Tab.1 Dangerous target identification results

模型	烟雾AP/%	火焰AP/%	佩戴安全帽AP/%	未佩戴安全帽AP/%
YOLOv5s	85.5	86.7	90.4	90.6
改进后模型	86.3	88.4	94.6	95.5

(下转第14页)