



一种轻量化油田危险区域入侵检测算法

田枫, 白欣宇, 刘芳, 姜文文

引用本文:

田枫,白欣宇,刘芳,姜文文. 一种轻量化油田危险区域入侵检测算法[J]. 智能系统学报, 2022, 17(3): 634–642.

TIAN Feng,BAI Xinyu,LIU Fang,JIANG Wenwen. A lightweight intrusion detection algorithm for hazardous areas in oilfields[J]. *CAAI Transactions on Intelligent Systems*, 2022, 17(3): 634–642.

在线阅读 View online: <https://dx.doi.org/10.11992/tis.202107033>

您可能感兴趣的其他文章

基于注意力机制的显著性目标检测方法

Salient object detection method based on the attention mechanism

智能系统学报. 2020, 15(5): 956–963 <https://dx.doi.org/10.11992/tis.201903001>

注意力机制和Faster RCNN相结合的绝缘子识别

Insulator recognition based on attention mechanism and Faster RCNN

智能系统学报. 2020, 15(1): 92–98 <https://dx.doi.org/10.11992/tis.201907023>

基于注意力融合的图像描述生成方法

An image caption generation method based on attention fusion

智能系统学报. 2020, 15(4): 740–749 <https://dx.doi.org/10.11992/tis.201910039>

层次化双注意力神经网络模型的情感分析研究

Hierarchical double-attention neural networks for sentiment classification

智能系统学报. 2020, 15(3): 460–467 <https://dx.doi.org/10.11992/tis.201812017>

基于跳跃连接金字塔模型的小目标检测

Skip feature pyramid network with a global receptive field for small object detection

智能系统学报. 2019, 14(6): 1144–1151 <https://dx.doi.org/10.11992/tis.201905041>

基于双向消息链路卷积网络的显著性物体检测

Salient object detection based on bidirectional message link convolution neural network

智能系统学报. 2019, 14(6): 1152–1162 <https://dx.doi.org/10.11992/tis.201812003>



微信公众平台



期刊网址

DOI: 10.11992/tis.202107033

网络出版地址: <https://kns.cnki.net/kcms/detail/23.1538.TP.20220329.1834.002.html>

一种轻量化油田危险区域入侵检测算法

田枫, 白欣宇, 刘芳, 姜文文

(东北石油大学 计算机与信息技术学院, 黑龙江 大庆 163318)

摘要: 油田危险区域入侵是油田安防领域的核心问题, 以目标检测的方式捕获实时发生的危险是区域入侵任务的重点。为了提高模型的实时性, 本文提出结合跨阶段线性瓶颈模块和通道注意力机制的轻量化 YOLO 检测算法。首先以轻量化卷积模块与跨阶段局部残差模块级联的跨阶段线性瓶颈模块搭建特征提取网络, 大大减少了模型的参数量。在特征金字塔的特征融合模块前使用改进的通道注意力机制, 增强特征的表达能力与特征的全局的关联性。在特征推理模块, 使用中心归一化非极大值抑制方法进行输出优化, 避免了对邻近目标的错误抑制。本算法在 VOC2007 数据集实验, 精确率可达 74.9%, 优于大多轻量化检测算法, 已在冀东油田部署应用, 有效保证了油田作业人员生命财产安全。

关键词: 油田危险区域入侵; 目标检测; 深度可分离卷积; 轻量化; 通道注意力; 深度学习; 特征融合; 特征提取
中图分类号: TP37 **文献标志码:** A **文章编号:** 1673-4785(2022)03-0634-09

中文引用格式: 田枫, 白欣宇, 刘芳, 等. 一种轻量化油田危险区域入侵检测算法 [J]. 智能系统学报, 2022, 17(3): 634-642.

英文引用格式: TIAN Feng, BAI Xinyu, LIU Fang, et al. A lightweight intrusion detection algorithm for hazardous areas in oilfields[J]. CAAI transactions on intelligent systems, 2022, 17(3): 634-642.

A lightweight intrusion detection algorithm for hazardous areas in oilfields

TIAN Feng, BAI Xinyu, LIU Fang, JIANG Wenwen

(School of Computer and Information Technology, Northeast Petroleum University, Daqing 163318, China)

Abstract: Hazardous areas in oilfields are a core problem in oilfield security, and capturing hazards occurring in real-time with target detection is the focus of intrusion detection models. To improve the real-time performance of such models, this paper proposes a lightweight YOLO detection algorithm combining the cross-stage linear bottleneck module and channel attention mechanism. First, a feature extraction network is built with the cross-stage linear bottleneck module cascaded with the lightweight convolutional and cross-stage local residual modules; this greatly reduces the parameters of the model. Then, an improved channel attention mechanism is used in front of the feature fusion module of the feature pyramid to enhance feature expressiveness and global relevance. In the feature inference module, the output is optimized using a central, normalized, nonmaximal suppression method to avoid false suppression of neighboring targets. This algorithm is tested in the VOC2007 dataset, and the accuracy rate can reach 74.9%, which is better than most lightweight detection algorithms. The algorithm has been deployed and applied in the Jidong Oilfield, where it effectively ensures the safety of life and property of oilfield operators.

Keywords: intrusion into hazardous areas in oilfields; target detection; depth separable convolution; lightweight; channel attention; deep learning; feature fusion; feature extraction

随着深度学习等大量智能算法的出现以及计算机性能的不断提升, 使用智能算法对监控视频进行快速分析成为可能。在我国各大油田由建设

数字化油田转变建设智慧油田的背景下, 油田危险区域入侵智能综合识别技术彻底取代了基于人力监控油田的高成本低准确率的方式, 为油田安全提供了有力保障, 但油田作业区设备算力有限, 摄像头数量达上百路, 如何在保证模型精度的基础上, 减少模型的参数量, 尽可能地提高模型的运算速度是目前亟待解决的问题。

目前基于深度学习的目标检测算法根据预测

收稿日期: 2021-07-19. 网络出版日期: 2022-03-30.

基金项目: 国家自然科学基金项目 (61502094); 黑龙江省自然科学基金项目 (LH2021F004); 黑龙江省优秀青年科学基金项目 (YQ2020D001); 东北石油大学优秀中青年科研创新团队项目 (KYCXTD201903).

通信作者: 刘芳. E-mail: lfliufang1983@126.com.

的流程可以分为两类^[1]:第1类为基于回归的深度卷积的目标检测算法,代表性的算法有 YOLO (you only look once) 系列的 YOLOv3^[2],其使用深层的特征提取网络 Darknet-53^[2]以及3个尺度的特征图进行边界框的预测,同时增加了 anchor 的数量,YOLOv4^[3]提出使用 CSPDarknet53^[3]来替换主干网络并采用多种训练策略,YOLOv5^[4]设计了两种 CSP 结构来适应不同的任务,并加入 Focus 做切片操作提高速度。YOLO 系列算法在不同比例的目标尺度上泛化性能不好,需要多次下采样来获取标准特征。SSD (single shot multi-box detector) 目标检测算法是基于前馈卷积网络,其特征提取基础网络中每个特征层用于检测的卷积核大小是不同的,卷积后可以得到多个尺度检测的预测值,以实现图像的多尺度检测,它使用非极大抑制 (non maximal suppression, NMS) 算法对预测结果进行处理后获得最终检测结果。Fu 等^[5]在 SSD 算法的基础上提出了反卷积单阶段多框探测器 DSSD 算法,在分类回归之前引入了残差模块,加深了网络层数。该算法有效地联系上下文,将各类语义信息进行融合,提高了检测精度。Jeong 等^[6]研究提出了一种 RSSD 融合算法,该算法通过特征图池化加反卷积的操作方式进一步融合不同层的网络特征,池化与特征图加反卷积^[7]的步骤同步进行,有效地解决了原始 SSD 特征图中存在重复框的问题,同时提升了小目标物体检测的成功率。Li 等^[8]提出 FSSD,该算法将网络中部分特征调整为同一尺寸再进行连接,得到一个像素层,并以此层为基础层来生成特征金字塔。第2类是基于候选区域的深层卷积神经网络目标检测算法,代表算法有 R-CNN^[9]、R-FCN^[10]、Fast R-CNN^[11]和 Faster R-CNN^[12],将目标检测分为两步,先通过区域建议算法生成可能包含目标的候选区域,再通过 CNN 对候选区域进行分类和位置回归,得到最终的检测框。这类算法需要逐个处理产生的候选框^[13],检测速度受到限制。

经筛选,YOLOv5 是目前目标检测中平衡速度与精度最好的算法,本文针对 YOLOv5 参数量多,计算量大的问题进行改进,提出结合深度可分离卷积的跨阶段卷积模块与改进的 SE^[14] (squeeze-and-excitation) 通道注意力模块的算法模型,并使用 DIOU-NMS^[15] (distance intersection over union-non maximum suppression) 中心归一化非极大值抑制算法进行特征推理;提出 OilPerson 数据集 (油田现场工人数据集),最后使用 TensorRT 加速与射线法共同完成区域入侵的判定。本文算法在油田作业现场应用良好,有效保障油田施工人员

的生命财产安全。

1 相关理论

1.1 YOLOv5 模型

YOLOv5 的网络结构如图 1 所示, YOLOv5 在 Backbone(特征提取网络部分) 中使用 Focus 结构减少计算量。使用 CSPNet^[3] 的 CSP 模块, 增加 CNN 的学习能力, 降低计算瓶颈并节约内存成本。使用 SPPNet^[16] 的特征金字塔模块, 将卷积神经网络的输入转化为任意尺寸。

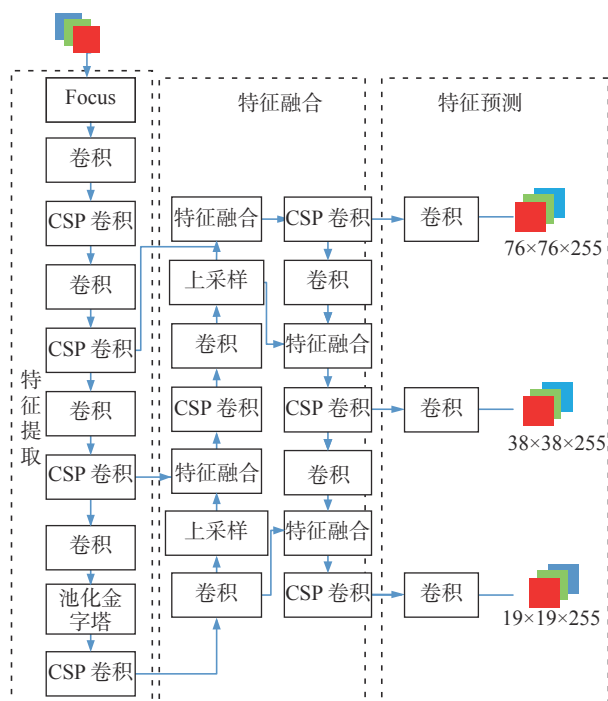


图1 YOLOv5 检测网络结构

Fig. 1 YOLOv5 detection network structure

在 Neck(特征融合部分), 采用 FPN 和 PAN^[17] 结构, 使用自上而下的路径和横向连接以及自底向上的路径增强, 对语义信息与定位信息进行特征融合, 用以提高多尺度检测的准确性。

在 Prediction(特征预测部分), 在特征图上应用锚定框生成带有类概率、对象得分和包围框的 3 种不同大小的特征图向量。使用 NMS^[2] 检测算法对同一目标产生多次检测的结果进行推理, 保证每个目标只检测一次, 找到检测效果最好的框。

1.2 基于深度可分离卷积的线性瓶颈模块

普通卷积模块在进行卷积操作时, 同时在空间和通道两个维度进行, 而深度可分离卷积^[18] 在空间和通道两个维度分开进行卷积操作, 形成了空间维度的逐层卷积和通道维度的逐点卷积两个部分。具体来说, 它将标准卷积分成了 2 步, 即 3×3 的逐层卷积和 1×1 逐点卷积, 如图 2 所示。

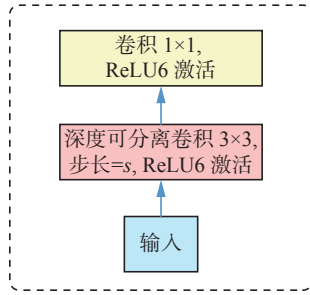


图 2 深度可分离卷积

Fig. 2 Depth separable convolution

这种分解方式明显地减少了计算量。MobileNetv2^[19] 则在此基础上, 为了改进 3×3 的逐层卷积在低维空间提取特征效果较差的缺点, 在逐层卷积前多加入了一个逐点卷积, 如图 3 所示, 进而完成对卷积神经网络中的特征图升维。

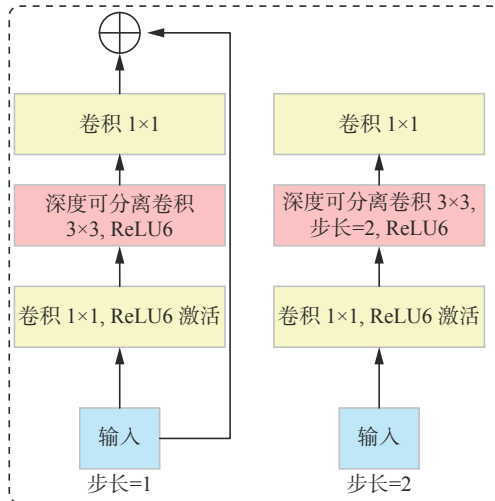


图 3 Inverted residuals 模块

Fig. 3 Inverted residuals bottleneck

在逐点卷积的末尾去掉了激活函数, 此结构即线性瓶颈模块。MobileNetv2 由多个瓶颈模块堆叠成, 其速度与准确率均得到了提升。

2 本文算法

2.1 特征提取网络

本文提出的网络结构如图 4 所示, 相对于原 YOLOv5 模型, 在 Backbone 特征提取部分由多个深度可分离卷积与线性瓶颈模块堆叠而成, 使用 Inver 在图中进行表示。将线性瓶颈模块与跨阶段局部残差网络结合形成跨阶段线性残差模块, 大幅减少了模型的运算量, 在每个特征金字塔采样前使用改进的 SE 模块提高模型定位精度, 使用 SE-h-sigmoid 来表示, 最后在 Prediction 特征推理部分使用 DIOU-NMS^[15] 替换 NMS, 使得其避免重叠度较高的同类目标出现漏检的问题。此模型在 YOLOv5 的基础上降低了模型参数量。

2.1.1 本文卷积模块

D_{Conv} 逐层卷积与 P_{Conv} 逐点卷积的计算成本如式 (1) 所示:

$$\begin{aligned} D_{Conv} &= D_K D_K C_M D_F D_F \\ P_{Conv} &= C_M C_N D_F D_F \end{aligned} \quad (1)$$

式中: C_M 代表输入通道的数量; C_N 代表输出通道的数量; D_K 代表卷积核大小; D_F 为特征映射图。深度可分离卷积与标准卷积的计算成本的比例如式 (2) 所示:

$$\frac{D_K D_K C_M D_F D_F + C_M C_N D_F D_F}{D_K D_K C_M C_N D_F D_F} = \frac{1}{C_N} + \frac{1}{D_K^2} \quad (2)$$

使用 3×3 的深度可分离卷积时, 在少量降低精度的前提下, 计算量相对于标准卷积减少 8 倍以上。

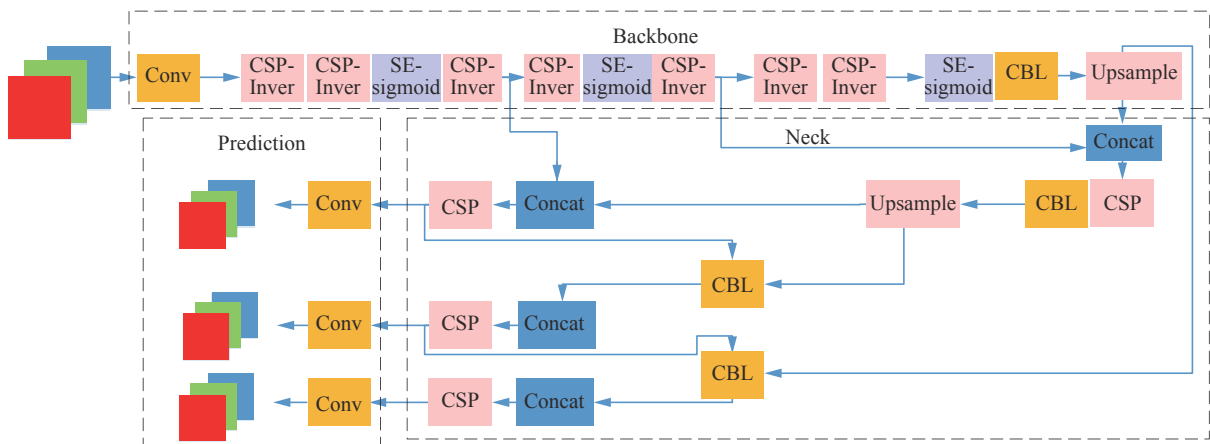


图 4 网络结构

Fig. 4 Network structure

Inverted Resblock 结构在轻量化网络方面有良好的性能表现, 但 Sandler 等^[19] 解释线性瓶颈模块存在梯度混淆的情况, 同时 1×1 的卷积减少了

空间信息, 本文根据线性瓶颈模块中存在的问题将 CSPNet 的跨阶段局部模块与 MobileNetv2 的线性瓶颈模块进行融合, 形成跨阶段线性瓶颈

卷积模块。输入的特征映射分为两个分支, 分别进行特征提取, 其中一支后接 Inver 卷积模块, 进行进一步特征提取, 然后将两分支通过级联操作进行合并。

如图 5 所示, 本文使用的 CSP-Inver 卷积块由两条分支组成, 上面的分支首先经过 3×3 卷积, 然后经过 Inver 卷积, 与下面的 3×3 卷积的分支进行 concat 操作。因两个分支的梯度信息不同, 不会出现重复的梯度信息情况, 利于学习到更优的梯度信息。此外, 级联操作是对卷积结果的拼接, 不含矩阵计算, 因此不会增加过多的计算量。

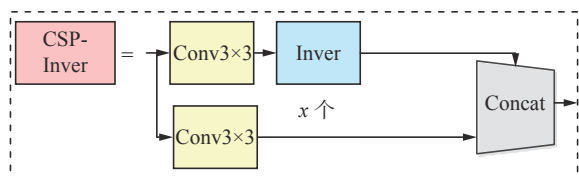


图 5 CSP-Inver: 跨阶段线性瓶颈模块

Fig. 5 CSP-Inver: Cross phase linear bottleneck module

从计算量的角度分析, 若线性瓶颈卷积模块中共 x 个卷积模块, 则卷积层的输入输出的内存流量可记为 $xc_{out} + \frac{x^2 + x}{2}$, 再引入 CSP 后可记为 $(xc_{out} + (x^2 + x))/2$, 通过式 (3) 可计算出使用 CSP 后减少的计算量比例。融合后线性瓶颈卷积模块的输入输出的内存流量减少近一半。

$$\text{ratio} = \frac{(xc_{out} + (x^2 + x))/2}{xc_{out} + \frac{1}{2}(x^2 + x)} = 1 - \frac{1}{2 + \frac{x^2 + x}{xc_{out}}} \quad (3)$$

因此将重新构成的跨阶段线性瓶颈模块与通道注意力模块融入 YOLOv5 网络结构中, 因使用 CSP-Inver 卷积块的 Inver 部分先进行 1×1 逐点卷积, 然后进行 3×3 的逐层卷积与 1×1 的逐点卷积取代普通卷积, 因此大大减少了网络模型的复杂度。

跨阶段线性瓶颈模块大幅降低了计算量, 减少了模型参数量, 但也存在着局限性, 当检测小目标物体或目标物体被遮挡时检测精度并不高, 出现此问题的原因分析如下:

对于一帧输入图像而言, 其多尺度特征图在 YOLOv5 网络模型中的数学定义如式 (4)、(5):

$$F_n = s_n(F_{n-1}) = s_n(s_{n-1}(\cdots(s_1(t)))) \quad (4)$$

$$D = f(r_n(F_n), r_{n-1}(F_{n-1}), \cdots, r_{n-k}(F_{n-k})), \quad n > k > 0 \quad (5)$$

式 (4) 中: 函数 s_n 代表第 n 层特征图与第 $n-1$ 层特征图之间的非线性映射函数, 其中函数主要操作有 3 个, 分别为池化、卷积和非线性激活函数操作; F_n 表示第 n 层的特征图; $s_1(t)$ 中的 t 表示输入图像, 1 表示第 1 层特征; $s_n(t)$ 表示输入图像和第 n 层的特征图之间的非线性函数。再用非线性函

数对特征图进行卷积操作, 并获得最终的检测结果对应的非线性函数, 在式 (5) 中用 $f(\cdot)$ 表示; 采用非线性映射的方法对第 n 层特征图进行处理, 进而可以获得位于某一范围内的检测结果, 这个结果用 $r_n(\cdot)$ 表示。

由式 (4)、(5) 得, 因不同的特征层对应多种不同尺度大小的目标结果, 为保证所获得的检测结果更加准确, 通常需要应用各种非线性函数, 需要保证网络结构中的每一层特征图包含更多的有效信息。因此以跨阶段非线性纺锤形模块构建特征提取网络, 更加需要全局信息进行特征筛选, 对重要特征强调并对非重要特征进行抑制, 所以本文在多尺度特征金字塔的每个尺度中, 需要添加改进的通道注意力机制, 增强整个网络的表示能力, 进而对小目标物体或被其他物体遮挡的目标物体进行较为准确的检测。

2.1.2 改进的 SE 通道注意力机制

特征金字塔网络结构虽然可以通过提取不同尺寸的特征信息来增加感受野, 但是金字塔网络在将具有很强代表性和区分性的高层语义信息通过上采样操作传递到其他层时, 高层语义信息将会被逐渐淡化^[20], 但高层语义信息因其区分性与强代表性能够更好地识别物体类别。为解决这个问题, 通过在特征金字塔中加入通道注意力层, 如图 6 所示, 该模块建立在自底向上路径的末端, 通过通道注意力层, 对重要的语义信息给予更多的关注, 这些包含有效语义信息的多尺度特征经过通道注意力层的处理后, 在自顶向下的过程中连接在一起, 补充自上而下的高级语义信息。

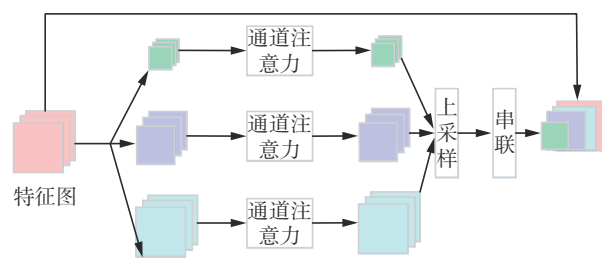


图 6 通道注意力特征金字塔

Fig. 6 Pyramid of channel attention characteristics

在每个特征金字塔的不同尺度模块采样前加入 SENet 网络模块作为注意力机制, 通过增强建模通道之间的相互依赖关系, 自适应地调整通道的特征响应, 但通道注意力机制结合特征金字塔带来了部分计算参数与模型参数, 降低了模型的运算速度, 增加了模型的运行内存。因此对 SE 通道注意力机制进行改进, 少量增加计算量的同时对重要特征进行强调。

改进前的 SE 通道注意力机制主要步骤: 对输

入的特征图进行全局平均池化, 得到长度等于通道数 M 的实数列 Z_{gap} 。在压缩率 $r=16$ 的情况下对特征图 $X = [x_1 \ x_2 \ \cdots \ x_n]$ 进行全局平均池化, 得到 F_{gap} , 计算过程如式 (6)、(7) 所示:

$$Z_{\text{gap}} = F_{\text{gap}}(x_c) = \frac{1}{HW} \sum_{m=1}^H \sum_{n=1}^W x_c(m, n) \quad (6)$$

式中: F_1 为进行降维的全连接层, $F_1 \in R^{M/r \times M}$; F_2 为升维的全连接层, $F_2 \in R^{M/r \times M}$; C 为特征图的索引。

原始的 SE 卷积注意力模块的最后一层使用 sigmoid 函数激活, 在反向传播更新梯度时, 求导做除法运算更加消耗资源。因为 sigmoid 激活函数其指数运算具有以上缺点, 如式 (7) 所示:

$$\text{sigmoid}[x] = \frac{1}{1 + e^{-x}} \quad (7)$$

所以使用由 Relu^[20] 激活函数表示的分段线性函数 hsigmoid^[21] 降低部分计算成本, 改进 SE 通道注意力机制, R 代表 Relu6^[20] 激活函数, 如式 (8) 所示:

$$\text{hsigmoid}[x] = \frac{R(x+3)}{6} \quad (8)$$

生成每个通道对应的权重信息 $S = [s_1 \ s_2 \ \cdots \ s_c]$, h 代表 hsigmoid 函数, 如式 (9) 所示:

$$S = h(F_2(R(F_1 \times Z_{\text{gap}}))) \quad (9)$$

根据权值对输入的特征图加权更新, 得到更新后的通道特征 $Y = [y_1 \ y_2 \ \cdots \ y_c]$, 如式 (10) 所示:

$$Y_c = s_c x_c \quad (10)$$

更换 hsigmoid 激活函数的 SE 通道注意力, 如图 7 所示。

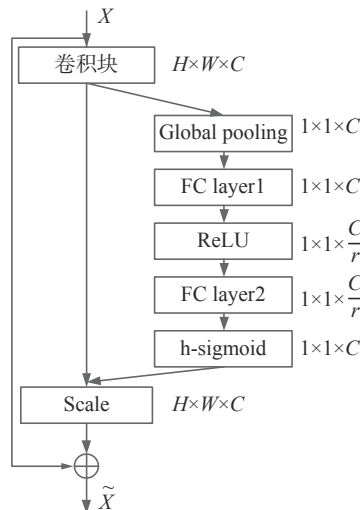


图 7 更换激活函数后的通道注意力

Fig. 7 Channel attention after changing the active function

因 YOLOv5 是由自顶向下的特征提取网络和自底向上的多尺度目标检测网络两部分组成, 前者负责提取的是输入图像的低层细节特征, 后者从前者的低层特征中提取高级的语义特征, 然后将不同尺度的高级特征与相对应尺度的低层特征通过

特征金字塔进行结合, 实现了多尺度的目标检测。

因此在特征提取网络中的特征金字塔的采样模块前使用改进的 SE 注意力模块, 能够更好地利用全局信息进行特征筛选, 对重要特征强调并对非重要特征进行抑制, 增强了整个网络的表示能力。

2.1.3 DIOU-NMS

本文使用 DIOU-NMS 替换 YOLOv5 的 NMS, 提升检出率。YOLOv5 使用的 NMS 其 IOU 指标常用于抑制冗余框, 遮挡时重叠区域经常导致 NMS 产生错误抑制。IOU 如式 (11) 所示:

$$U = \frac{|B \cap B^{\text{gt}}|}{|B \cup B^{\text{gt}}|} \quad (11)$$

式中: $B^{\text{gt}} = (x^{\text{gt}}, y^{\text{gt}}, w^{\text{gt}}, h^{\text{gt}})$ 代表真实的边界框; $B = (x, y, w, h)$ 为预测框。

DIOU 在 IOU 的基础上增加了一个惩罚项, 该惩罚项用于最小化两个 box 的中心点距离, 如式 (12) 所示:

$$R_{\text{DIOU}} = \frac{\rho^2(b, b^{\text{gt}})}{c^2} \quad (12)$$

式中: ρ 是欧几里得距离; c 是覆盖两个框的对角线长度; b 、 b^{gt} 分别是 B 、 B^{gt} 的中心点。

对于预测分数较高的 box, DIOU-NMS 除考虑重叠区域外, 还将两个检测输出的 box 的中心点距离作为考虑因素, DIOU-NMS 如式 (13) 所示:

$$s_i = \begin{cases} s_i, & \text{IOU} - R_{\text{DIOU}}(M, B_i) < \varepsilon \\ 0, & \text{IOU} - R_{\text{DIOU}}(M, B_i) \geq \varepsilon \end{cases} \quad (13)$$

式中: s_i 代表分类得分; ε 为 NMS 阈值; M 为得分最高的 box。因为考虑中心距离, DIOU-NMS 不抑制两个中心点较远、得分较高的 box, 并将其识别为两个目标对象。通过这个方式可提升检出率。

2.2 区域入侵

考虑不同视角下的摄像头位置含有景深信息, 而非仅有俯视的二维表面, 在倾角较小或者近似平行的摄像头位置中, 仅根据检测目标与区域判断是否发生重合来判定是否发生区域入侵并不准确。

人物是否走入危险区域内, 通过人物落脚点的位置进行判断。因此考虑人物的落脚点是否在危险区域内, 如图 8 所示, 先通过目标检测获取目标位置, 然后将人物踏入危险区域的区域入侵问题抽象成为落脚点与危险区域多边形是否相交的问题, 即判断落脚点是否在不规则多边形内。

计算油田工人落脚点 P_x , 如式 (14) 所示:

$$P_x = b\left(x_1, \frac{h}{2} + y_1\right) \quad (14)$$

式中: b 为获取到的检测框坐标函数, x_1 、 y_1 为中心点的坐标, h 为输出的定位框的高。

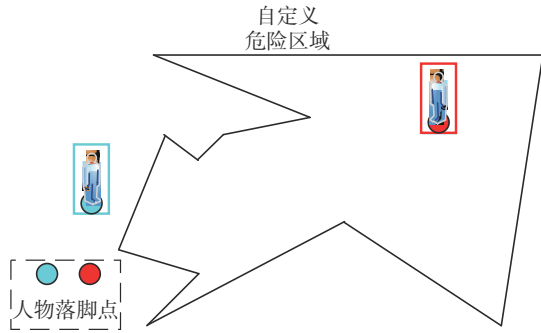


图8 射线法
Fig.8 Ray method

对于平面内任意闭合曲线,曲线都把平面分割成了内、外两部分。在平面内对于任意一条直线,在穿越多边形边界时,仅有两种情况:进入或穿出多边形。因此本文设计射线法进行区域入侵的判断。

如图9所示,因在不同视角下的摄像头位置含有景深信息,例如:90°、60°、30°摄像头位置悬挂下,同一危险区域映射到摄像头图像的位置不同,若仅凭借目标检测框与危险区域的面积计算交并比这一方式判断目标发生危险区域入侵不可取,本文以射线法判别目标是否处于危险区域中,即判别目标的落脚点是否位于不规则的危险区域对边形中,进行区域入侵判断,这样可以有效解决不同视角下的摄像头是否发生危险区域入侵问题。

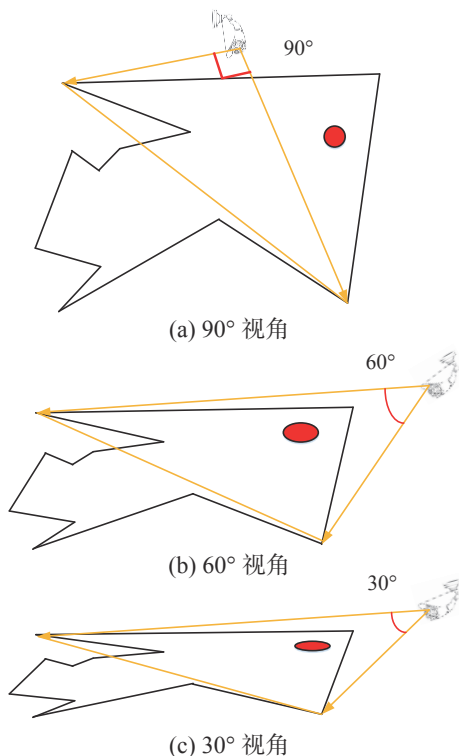


图9 不同角度的危险区域
Fig.9 Dangerous areas from different angles

根据式(15)判断是否发生多边形区域入侵:

$$R = s((q(P_x, x)) \% 2) \quad (15)$$

式中: q 为点 P_x 向任意方向做射线; s 表示求和; x 为自定义危险区域; R 表示发生区域入侵; $\%2$ 表示与2相除后取余数。当目标在画面中出现时,如图10所示,从目标的落脚点对任意方向做射线,当射线的交点个数为奇数时发生区域入侵。

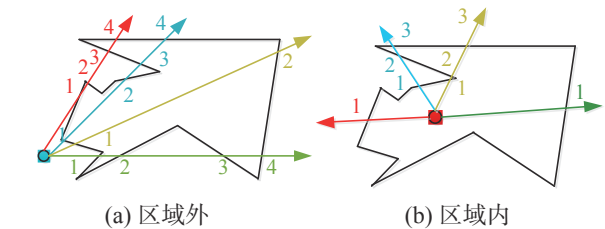


图10 判断点是否在多边形内
Fig.10 Determine whether the point is in the polygon

3 实验结果与分析

实验硬件环境:硬件平台为联想工作站, i7-6700 3.4 GHz CPU、NVIDIA CTX 3080 GPU, 操作系统为 Ubuntu 18.04。软件环境:程序编写语言为 Python3.7。

3.1 VOC数据集实验结果对比与分析

使用 Mosaic^[22](马赛克)数据增强策略随机改变训练样本,随机读取4张训练图像,进行反转和旋转操作,组合成为一张训练图片,提高模型的泛化能力。batch-size(每批训练的图片量)设置为32, momentum(动量值)设置为0.9,学习率(learning-rate)初始值为0.0001,通过余弦退火策略^[23]调整学习率,权重衰减参数为0.005, epoch设置为300。

本文选用 PASCAL VOC2007^[24]和 PASCAL VOC 2012数据集联合训练模型,图片数量为16551张,在测试过程中选择的是 PASCAL VOC 2007测试集,图片数量为4952张。本文选用的评价指标为:检测精度(mean average precision, mAP)、运行速度(frames per second, f/s)、模型大小(model size)。

表1实验结果显示,网络结构复杂的模型检测准确率高,但检测速度较慢。轻量级检测网络优点是检测速度较快,缺点是检测的准确率低。在模型体积类似的 YOLOv3-tiny 和 YOLOv4-tiny 中, mAP 分别提高了17.8%和9.6%,其主要原因是本文的级联基于深度可分离卷积的线性瓶颈模块与改进的 SE 通道注意力模块在特征金字塔上的使用,弥补了特征提取能力不足的问题,增强了模型对特征的利用率,提高了准确率。使用 DIIOU-NMS 替换 YOLOv5 的 NMS 增强了模型的推理能力,避免了在目标距离较近情况

下的错误抑制,提升了准确率。对训练集使用 Mosaic 处理丰富了物体的背景,提高了模型的泛化能力。

表 1 不同网络在 VOC 数据集的评价值

Table 1 Evaluation value of different models in VOC dataset

算法	基础网络	检测精度/%	模型大小/MB	运行速度/f·s ⁻¹
Faster RCNN	VGG16	73.2	528	28
SSD-MobileNet	MobileNet	77.4	101	51
YOLOv3	Darknet53	78.3	248	29
YOLOv3-tiny	—	57.1	36	133
YOLOv4-tiny	—	65.3	24	99
YOLOv5l	CSPNet	79.2	94.6	43
本文	—	74.9	7.4	142

在 VOC 数据集上的消融实验如表 2 所示,用√表示使用当前模块,在仅使用 CSP-Inver 卷积代替 CSP 卷积减少了模型的参数量,但模型精度有所下降。在基准网络中加入 FPN-SE 模块(即在特征金字塔进行特征融合前的卷积模块使用通道注意力机制),mAP 提高了 2.3%,说明本文使用的 FPN-SE 模块能增强特征提取能力,但也带来了 2.7 MB 的模型参数量;使用 FPN-SE-h-sigmoid 模块替换 FPN-SE 模块,在保证检测精度的同时,减少了 0.5 MB 的模型参数量。加入 DIOU-NMS 进行特征推理后,在不增加网络参数的前提下,mAP 提高了 0.5%。实验证明:CSP-Inver 卷积大大减少了模型的参数量,FPN-SE-h-sigmoid 与 DIOU-NMS 模块在提升模型精度的同时少量增加模型参数,保证了模型的运行速度。

表 2 本文消融实验

Table 2 Ablation experiments

CSP-Inver	FPN-SE	FPN-SE-h-sigmoid	DIOU-NMS	模型大小/MB	检测精度/%
√	—	—	—	5.2	72.2
√	√	—	—	7.9	74.5
√	—	√	—	7.4	74.4
√	—	√	√	7.4	74.9

3.2 OilPerson 数据集实验结果对比分析

油田场景中,背景复杂,因油田工人着装为红色所以易与红色集装箱、抽油机等设备混淆,且油田作业中,工人作业姿态复杂,例如在拧天然气或石油阀门时侧拧、下蹲等作业姿态。因此对油田中典型的复杂场景(输油泵房、输气泵房、石油阀门、H²S 钻井口、天然气阀门、抽油机、油水

分类器、钻井平台、电力高压间、露台/密闭泥浆池、原油仓库)进行视频采样(约 30 000 张)。针对工人在远景摄像头中目标较小、数量密集、作业姿势复杂等问题,人工制作油田作业现场工人数据集 20 000 张,覆盖油田大部分复杂场景。

本文划分训练集、测试集、验证集的标准为:覆盖本油田下全部复杂场景、多种摄像头倾角、多种人员作业姿态。Oilperson 数据集按照 60%、20%、20% 的比例进行划分。

Oilperson 实验结果如表 3 所示,本文算法模型较其他模型,在大幅减少模型参数的情况下保证了检测精度,证明了其对油田场景的有效性。实验结果显示,在 Oilperson 油田工人数据集中,因标注的数据集覆盖所有场景、多种作业人员姿态以及不同倾角和不同尺度的目标,所以各个检测算法在本数据集上的检测结果均有提升,验证了本文 Oilperson 对油田工人检测的有效性。本文检测算法 mAP 为 89.9%,模型大小为 7.65 MB,运行速度为 140 f/s,与表 3 中其他轻量化模型算法相比,速度、模型参数量达到最优。

表 3 不同网络在 Oilperson 数据集上的评价

Table 3 Evaluation value of different networks on Oilperson dataset

算法	检测精度/%	模型大小/MB	运行速度/f·s ⁻¹
YOLOv3-tiny	75.6	40	132
YOLOv4-tiny	84.2	26	99
SSD-MobileNet	86.5	101	51
YOLOv5l	91.4	94.6	39
本文	89.9	7.65	140

图 11 的检测结果显示,各个场景的作业人员都有较高的检测准确率。针对远景小目标,对比其他轻量化算法具有较好的检出率,验证了本文数据集的有效性以及本文算法的鲁棒性。

3.3 算法部署与应用

将生成的模型经过 TensorRT 进行 int8 类型的量化。经过对 500 张图片进行测试,模型平均处理一张图片的时间为 1.34 ms。

TensorRT^[24]加速后的模型与射线法进行结合,共同进行区域入侵的判定,不同场景的入侵结果如图 12 所示。

在河北唐山冀东油田进行危险区域入侵算法的部署与测试:在不同摄像头倾角、不同的目标大小、不同的危险场景下,结果图 12 所示。其中,图(a)摄像机为俯角,划定油罐车上方为危险

区域,当工人进入此区域时本算法判定为发生区域入侵;图(b)摄像机为俯角,阀门的正前方区域设置为危险区域,当工人停留时间达到5 s时,算法判定为发生区域入侵(工人可能在正拧阀门,而正拧阀门是一种错误的操作行为,当阀门压力

过大时,易导致阀门喷出致人受伤死亡);图(c)摄像机为平角,加热炉下方危险区域禁止进入,发生入侵;图(d)摄像机为俯角,在近海区域发生翻墙行为,油田工人抄近路易造成危险,本算法判定发生区域入侵并显示工人的行踪轨迹。



图 11 油田不同场景的检测效果

Fig. 11 Detection effect of different scenes in Oilfield

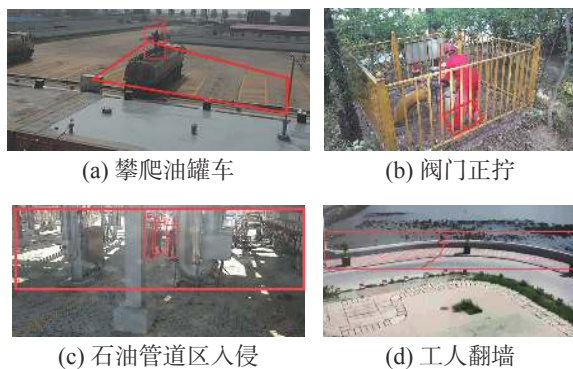


图 12 区域入侵效果

Fig. 12 Regional intrusion effect

4 结束语

本文使用基于深度可分离卷积的线性瓶颈模块与 CSP 跨阶段局部特征模块级联形成 CSP-Inver(跨阶段线性瓶颈模块),大幅减少了模型参数,提高了运算速度。每个特征金字塔的特征融合层添加改进的通道注意力模块,提高了模型的特征提取能力。特征推理部分使用 DIOU-NMS 减少了误检次数。经过 TensorRT 加速后,本文模型与射线法结合,在多种角度、不同景深信息的摄像

头下进行部署应用。部署结果显示,本文算法在冀东油田作业现场应用良好,有效地保障了油田施工人员的生命财产安全。

参考文献:

- [1] 赵永强, 饶元, 董世鹏, 等. 深度学习目标检测方法综述[J]. 中国图象图形学报, 2020, 25(4): 629–654.
ZHAO Yongqiang, RAO Yuan, DONG Shipeng, et al. Survey on deep learning object detection[J]. *Journal of image and graphics*, 2020, 25(4): 629–654.
- [2] REDMON J, FARHADI A. YOLOv3: an incremental improvement[EB/OL]. (2018–04–08)[2021–07–19]<https://arxiv.org/abs/1804.02767>.
- [3] YU Zhenwei, SHEN Yonggang, SHEN Chenkai. A real-time detection approach for bridge cracks based on YOLOv4-FPM[J]. *Automation in construction*, 2021, 122: 103514.
- [4] CAO Weixing. A wheat spike detection method in UAV images based on improved YOLOv5[J]. *Remote sensing*, 2021, 13(16): 3095.
- [5] FU Chengyang, LIU Wei, RANGA A, et al. DSSD: deconvolutional single shot detector[EB/OL]. (2017–01–23) [2021–07–19]<https://arxiv.org/abs/1701.06659>.
- [6] JEONG J, PARK H, KWAK N. Enhancement of SSD

- by concatenating feature maps for object detection[EB/OL]. (2017-05-26)[2021-07-19]<https://arxiv.org/abs/1705.09587>.
- [7] 郭川磊, 何嘉. 基于转置卷积操作改进的单阶段多边框目标检测方法[J]. 计算机应用, 2018, 38(10): 2833-2838.
- GUO Chuanlei, HE Jia. Improved single shot multibox detector based on the transposed convolution[J]. Journal of computer applications, 2018, 38(10): 2833-2838.
- [8] LI Zuoxin, ZHOU Fuqiang. FSSD: feature fusion single shot Multibox detector[EB/OL]. (2017-12-04)[2021-07-19]<https://arxiv.org/abs/1712.00960>.
- [9] HE Kaiming, GKIOXARI G, DOLLÁR P, et al. Mask R-CNN[J]. *IEEE transactions on pattern analysis and machine intelligence*, 2020, 42(2): 386-397.
- [10] ZHANG Hanwang, KYAW Z, YU Jinyang, et al. PPR-FCN: weakly supervised visual relation detection via parallel pairwise R-FCN[C]//Proceedings of 2017 IEEE International Conference on Computer Vision. Venice, Italy, 2017: 4233-4241.
- [11] GIRSHICK R. Fast R-CNN[C]//Proceedings of 2015 IEEE International Conference on Computer Vision. Santiago, Chile, 2015: 1440-1448.
- [12] REN Shaoqing, HE Kaiming, GIRSHICK R, et al. Faster R-CNN: towards real-time object detection with region proposal networks[J]. *IEEE transactions on pattern analysis and machine intelligence*, 2017, 39(6): 1137-1149.
- [13] WANG Qiong, ZHANG Lu, LI Yan, et al. Overview of deep-learning based methods for salient object detection in videos[J]. *Pattern recognition*, 2020, 104: 107340.
- [14] HU Jie, SHEN Li, ALBANIE S, et al. Squeeze-and-excitation networks[J]. *IEEE transactions on pattern analysis and machine intelligence*, 2020, 42(8): 2011-2023.
- [15] XIE Huosheng, WU Zesen. A robust fabric defect detection method based on improved RefineDet[J]. *Sensors*, 2020, 20(15): 4260.
- [16] HE Kaiming, ZHANG X, REN S, et al. Spatial pyramid pooling in deep convolutional networks for visual recognition[J]. *IEEE transactions on pattern analysis and machine intelligence*, 2014, 37(9): 1904-1916.
- [17] LIU Shu, QI Lu, QIN Haifang, et al. Path aggregation network for instance segmentation[C]//Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City: IEEE, 2018: 8759-8768.
- [18] HOWARD A G, ZHU Menglong, CHEN Bo, et al. MobileNets: efficient convolutional neural networks for mobile vision applications[EB/OL]. (2017-04-17)[2021-07-19]<https://arxiv.org/abs/1704.04861>.
- [19] SANDLER M, HOWARD A, ZHU Menglong, et al. MobileNetV2: inverted residuals and linear bottlenecks[C]//Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City: IEEE, 2018: 4510-4520.
- [20] 郑浦, 白宏阳, 李伟, 等. 复杂背景下的小目标检测算法[J]. 浙江大学学报(工学版), 2020, 54(9): 1777-1784.
- ZHENG Pu, BAI Hongyang, LI Wei, et al. Small target detection algorithm in complex background[J]. Journal of Zhejiang University (engineering science edition), 2020, 54(9): 1777-1784.
- [21] DUAN Chunmei, ZHANG Taoquan. Two-stream convolutional neural network based on gradient image for aluminum profile surface defects classification and recognition[J]. *IEEE access*, 2020, 8: 172152-172165.
- [22] ZHAO Boya, WU Yuanfeng, GUAN Xinran, et al. An Improved aggregated-mosaic method for the sparse object detection of remote sensing imagery[J]. *Remote sensing*, 2021, 13(13): 2602.
- [23] 赵凯琳, 靳小龙, 王元卓. 小样本学习研究综述[J]. 软件学报, 2021, 32(2): 349-369.
- ZHAO Kailin, JIN Xiaolong, WANG Yuanzhuo. Survey on few-shot learning[J]. *Journal of software*, 2021, 32(2): 349-369.
- [24] EVERINGHAM M, GOOL L V, WILLIAMS C, et al. The pascal visual object classes (VOC) challenge[J]. *International journal of computer vision*, 2010, 88(2): 303-338.

作者简介:



田枫, 教授, 博士生导师, 东北石油大学计算机与信息技术学院院长, 主要研究方向为计算机视觉、智能数据分析处理。主持和参与国家自然科学基金项目、国家科技重大专项项目 8 项。获授权专利 16 项, 发表学术论文 31 篇。



白欣宇, 硕士研究生, 主要研究方向为计算机视觉。



刘芳, 副教授, 博士研究生, 主要研究方向为智慧教育、多媒体与现代教育技术、计算机视觉、智能数据分析处理。获黑龙江省科技进步二等奖 1 项、大庆市科技进步二等奖 1 项, 主持和参与国家自然科学基金项目、黑龙江省自然科学基金项目 6 项。发表学术论文 22 篇。