

基于 YOLO v3 的机场场面飞机检测方法

郭进祥^{1,2}, 刘立波^{1*}, 徐峰¹, 郑斌¹

¹宁夏大学信息工程学院, 宁夏 银川 750021;

²中国民用航空西北地区空中交通管理局宁夏分局, 宁夏 银川 750009

摘要 小目标、飞机相互遮挡等难以检测的问题,对飞机检测的准确性及实时性提出很大的挑战。将实时性较高的 YOLO v3 算法应用到机场场面飞机检测领域,并提出两点改进:将骨干网络中的卷积层替换为空洞卷积,保持较高分辨率及较大感受野,提高模型对小目标检测的准确率;通过线性衰减置信得分的方式,对非极大值抑制(NMS)算法进行优化,以提升模型对被遮挡飞机的检测能力。结果表明,改进后的 YOLO v3 能够较好地检测小目标和遮挡飞机,且在保证实时性的前提下,将检测准确率从 72.3% 提高到 83.7%。

关键词 图像处理; 场面飞机检测; YOLO v3; 空洞卷积; 非极大抑制算法

中图分类号 TP391.7

文献标识码 A

doi: 10.3788/LOP56.191003

Airport Scene Aircraft Detection Method Based on YOLO v3

Guo Jinxiang^{1,2}, Liu Libo^{1*}, Xu Feng¹, Zheng Bin¹

¹ School of Information Engineering, Ningxia University, Yinchuan, Ningxia 750021, China;

² Ningxia Branch, Northwest Regional Air Traffic Management Branch of Civil Aviation Administration of China, Yinchuan, Ningxia 750009, China

Abstract The difficulty to detect small targets or occlusion aircrafts poses a great challenge to the accuracy and real-time of aircraft detection. In this paper, YOLO v3 algorithm with high real-time performance is applied to the field of aircraft detection in airport scene, and two improvements are proposed: replacing the convolution layer in backbone network with void convolution, maintaining high resolution and large field of receptivity and improving the accuracy of small target detection; optimizing the NMS algorithm by linear attenuation confidence score to improve the detection accuracy of occlusion aircrafts. The results show that the improved YOLO v3 can well detect small targets and occlusion aircraft, and the detection accuracy is improved from 72.3% to 83.7% as the real-time performance is ensured.

Key words image processing; scene aircraft detection; YOLO v3; dilated convolution; non-max suppression algorithm

OCIS codes 100.2000; 100.3008; 100.4999; 100.4996

1 引言

场面飞机检测是智慧机场、远程塔台的重要组成部分,它能自动识别机场图像中的飞机,辅助塔台管制员保障飞行安全,在场面飞机冲突预警、防跑道侵入、远程指挥等方面具有重要的应用价值^[1]。

场面飞机图像通常是由航站楼、塔台的摄像机

采集。相比于轮廓完整、姿态单一的遥感飞机图像,场面飞机图像存在飞机相互遮挡,及拍摄距离远造成飞机较小等难以检测的问题,这使得基于滑动窗口的传统方法受到限制^[2-4]。近些年,随着深度学习技术的快速发展,基于深度学习的方法广泛应用于目标检测领域。目前基于深度学习的方法主要分为两类:一是基于候选区域的方法,例如 Fast-

收稿日期: 2019-03-22; 修回日期: 2019-04-07; 录用日期: 2019-04-16

基金项目: 国家自然科学基金(61862050)、西部一流大学科研创新项目(ZKZD2017005)、宁夏自然科学基金(2018AAC03219)

* E-mail: 10801317@qq.com

RCNN^[5]、Faster-RCNN^[6]；二是基于回归的方法，例如 You Only Look Once (YOLO)^[7]、YOLO9000^[8]、SSD^[9]、YOLO v3^[10-11]。其中，基于候选区域的方法具有很高的检测准确率，但是检测速度较慢，不适用于场面飞机检测，而基于回归的检测方法检测速度快，满足场面飞机检测实时性要求，且 YOLO v3 融合锚 (Anchor)^[12]、特征金字塔 (FPN)^[13] 等策略，大大提高了检测准确率。

YOLO v3 是目前最优秀的检测算法之一，国内外诸多研究者将其应用于飞机检测领域，并提出了不同的改进方案。李耀龙等^[14]将 Ting-YOLO^[15]应用到机场飞机目标识别中，并提出一种简化的 Ting-YOLO 算法，提高了检测速度。Kharchenko 等^[16]将 YOLO v3 应用于无人机拍摄的地面小物体检测，该方法具有较高的检测能力、定位精度和实时处理速度。戴伟聪等^[17]将 YOLO v3 应用于遥感图像中的飞机检测，并将密集相连模块用于骨干网络，实现了多尺度遥感图像中的飞机检测。上述文献将 YOLO v3 应用于飞机检测，并通过减少骨干网络层数、改变卷积层等方式对算法进行改进，但是在小目标及飞机相互遮挡的检测方面，效果并不理想。

针对上述问题，本文将 YOLO v3 算法应用于场面飞机检测，并提出两点改进：1) 在骨干网络中，使用空洞卷积层^[18]替换卷积层，保持较高分辨率和感受野，提高小目标检测的能力；2) 在检测网络中，根据多个网格预测的候选框交并比 (IOU)^[19]的大小，线性衰减置信得分，保留被遮挡的飞机，解决飞机相互遮挡检测问题。本文使用自行组建并标注的场面飞机数据集 Aeroplane，在此数据集上，对多种算法进行对比测试，实验结果表明，本文方法在小目

标及相互遮挡的飞机检测方面取得良好的效果。

2 YOLOv3 算法

2016 年, Redmon 等^[7]在 YOLO v2^[8]算法的基础上, 融合 ResNet^[20-21]、FPN、二元交叉熵损失等方法, 提出了 YOLO v3 算法。YOLO v3 的网络结构分为骨干网络 (Darknet-53) 和检测网络。Darknet-53 由 52 个卷积层和 1 个全连接层组成, 构成了一个 53 层的卷积神经网络 (CNN), 并输出 13×13 、 26×26 及 52×52 三种尺度的特征, 送入检测网络^[16]。检测网络对三种尺度的特征回归, 预测出多个预测框, 并使用非极大抑制 (NMS) 算法, 去除 IOU 较大与置信得分较低的预测框, 保留置信得分较高的预测框为目标检测框^[19]。

YOLO v3 的检测流程, 如图 1 所示。首先, 输入的图像整体缩放为 416×416 , 并将图像划分成大小为 $S \times S$ 的网格, 如果某一目标的中心落入某一网格, 则该网格负责预测该目标^[22]。每个网格负责预测出 b 个预测框, 每个预测框对应 $(C+5)$ 个值, C 为检测物体类别数, 5 表示预测框的坐标 (x, y) 、尺寸 (宽 w 、高 h) 以及存在物体的置信得分 $P(O_{\text{object}}) = 1$ 这 5 个参数, 预测框置信得分 S_{confi} 为

$$S_{\text{confi}} = P(C_i O_{\text{object}}) \times P(O_{\text{object}}) \times I(\text{truth}, \text{pred}), \quad (1)$$

式中: 如果预测框中存在检测物体, 则 $P(O_{\text{object}}) = 1$, 否则 $P(O_{\text{object}}) = 0$; $P(C_i O_{\text{object}})$ 为网格预测 i 类的置信得分; C_i 为第 i 类物体, i 为类别编号; O_{object} 为存在物体; $I(\text{truth}, \text{pred})$ 为预测目标框与真实目标框的交并比。

最后, 使用 NMS 算法筛选出置信得分 S_{confi} 较

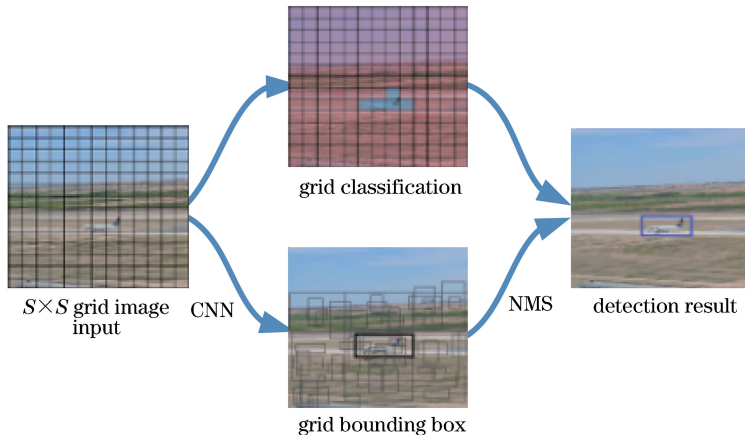


图 1 YOLO v3 的检测流程

Fig. 1 Detection flow of YOLO v3

高的预测框为检测框。NMS 可用公式表示为

$$S_{\text{confi}} = \begin{cases} S_{\text{confi}}, & I(M, b_i) < N_t \\ 0, & I(M, b_i) \geq N_t \end{cases}, \quad (2)$$

式中, M 为置信得分较大的候选框, b_i 为被比较的物体预测框, $I(M, b_i)$ 为 M 与 b_i 的交并比, N_t 为抑制阈值。

YOLO v3 中, Darknet-53 使用步长为 2, 大小为 3×3 的卷积核进行降采样, 感受野区域广, 牺牲了空间分辨率, 致使深层网络的小目标信息损失, 缺乏细节信息。虽然 YOLO v3 采用 FPN, 将高层特征与浅层特征信息融合, 提高了小目标的检测效果, 但是步长较大的高层导致小目标的语义特征丢失严重, 该特征融合方法对精确定位飞机和检测较小的场面飞机有不利的影响。NMS 算法通过判断 IOU 值是否大于阈值的方式, 删除多余的预测框, 保留物体的检测框。该方式较好地筛选了预测框, 但势必会删除遮挡飞机的预测框, 影响相互遮挡飞机的检测^[19]。

3 改进的 YOLO v3 模型

针对场面飞机存在的小目标与相互遮挡难以检测的问题, 对 YOLO v3 的骨干网络 Darknet-53 和 NMS 算法提出两点改进: 一是采用高分辨率、大感受野的骨干网络; 二是采用线性衰减的 NMS 算法。

3.1 高分辨率、大感受野的骨干网络

对于小目标的检测, 一方面需要高分辨率的特征图, 才能检测较小区域的目标信息; 另一方面需要较开阔的感受野或者更全局的信息, 才能准确判断目标的位置和语义特征。为了改善场面图像中小目

标的检测, 结合空洞卷积和 FPN 结构, 提出一种高分辨率、大感受野的骨干网络用于特征提取。

空洞卷积是通过改变卷积核的内部间隔来扩大卷积核的感受野。图 2 为 3 种不同间隔的空洞卷积核, 其中 r_{rate} 表示卷积核内部空洞的间隔。图 2(a) 为 $3 \times 3, r_{\text{rate}}=1$ 的空洞卷积, 卷积核的感受野范围是 3×3 。图 2(b) 为 $3 \times 3, r_{\text{rate}}=2$ 的空洞卷积, 卷积核的感受野增大到 7×7 。图 2(c) 为 $3 \times 3, r_{\text{rate}}=3$ 的空洞卷积, 卷积核的感受野增大到 15×15 , 保证卷积网络能够提取到较大视野内的特征信息^[23]。

改进后的 YOLO v3 的骨干网络及 FPN 结构, 如图 3 所示。相比于 YOLO v3 骨干网络, 改进后的 YOLO v3 没有对输入图像做过多的下采样, 最后三个阶段的分辨率保持为 26×26 , 增加最终输出特征图的分辨率, 减少深层网络的小尺寸目标语义特征丢失, 有利于识别较小的目标。该结构减少下采样的数量后进行特征提取, 势必将带来深层网络的感受野减小的问题^[23]。故改进后的 YOLO v3 沿用 FPN 结构, 引入空洞卷积瓶颈层, 如图 4(a) 所示, 带 1×1 卷积结构的空洞卷积瓶颈层, 如图 4(b) 所示。这两种复杂度较低的空洞卷积残差结构中, Conv 为卷积层, Add 为相加运算, ReLU 为激活函数。因使用大小为 $3 \times 3, r_{\text{rate}}=2$ 的空洞卷积, 从而整体上增大了骨干网络的感受野和特征表达能力^[24]。同时, 空洞卷积残差结构仍具有残差单元的网络参数较少、计算复杂度较低的优点, 图 4(b) 结构使用 1×1 的卷积实现跨通道的特征融合, 较好地整合了信息。

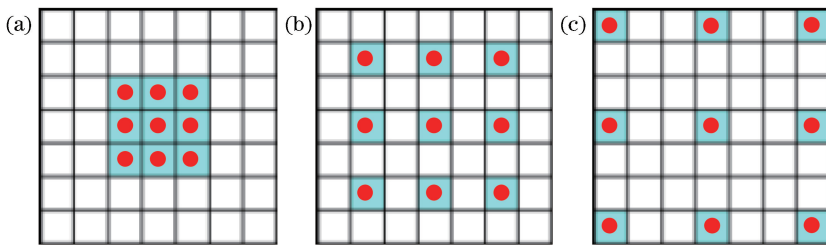


图 2 空洞卷积。(a) $r_{\text{rate}}=1$; (b) $r_{\text{rate}}=2$; (c) $r_{\text{rate}}=3$

Fig. 2 Dilated convolutions. (a) $r_{\text{rate}}=1$; (b) $r_{\text{rate}}=2$; (c) $r_{\text{rate}}=3$

特征的分辨率将直接影响小目标的检测和整体性能指标, 13×13 的小目标语义特征丢失严重, 52×52 的高分辨率特征会造成较大的计算量和内存存储量, 所以考虑到不降低骨干网络的整体性能, 在网络改进的过程中选择计算量较为适中的分辨率 26×26 。最终, 改进的 YOLO v3 能够在深层网络中保持较高分辨率和较大的特征图的感受野, 增强

了 YOLO v3 算法对小目标的感受野和检测能力。

3.2 线性衰减的 NMS 算法

如图 5 所示, 停靠在机场内的 A、B 两架飞机, A 飞机预测框的置信得分为 0.8, B 飞机预测框的置信得分为 0.6, A 飞机严重遮挡 B 飞机, 且 A、B 两架飞机预测框的 IOU 大于 0.5。使用 NMS 算法处理多余预测框, 当 A、B 飞机预测框的 IOU 高于 0.5

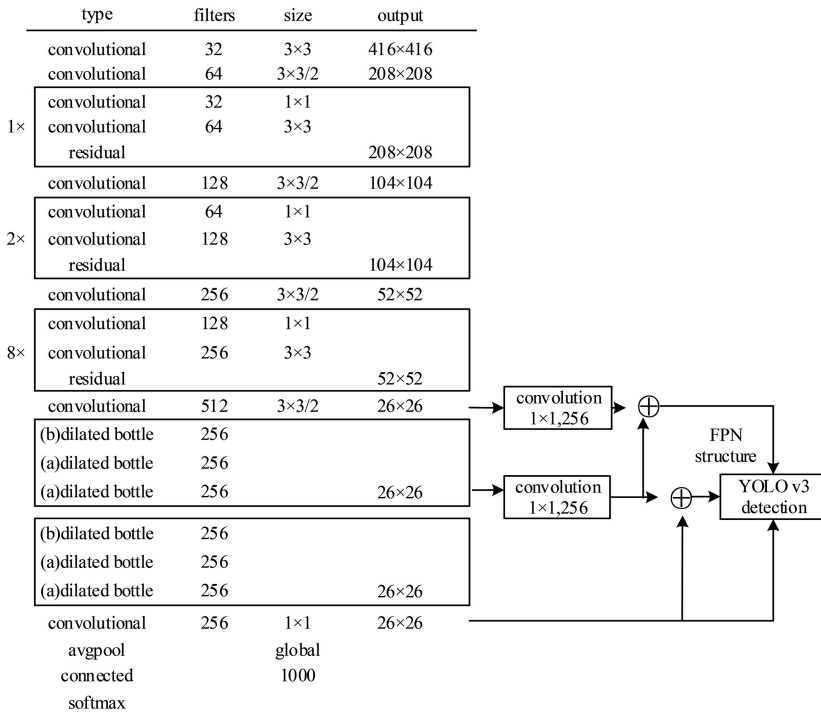


图 3 改进后 YOLO v3 的骨干网络及 FPN 结构

Fig. 3 Backbone network and FPN architecture of the improved YOLO v3

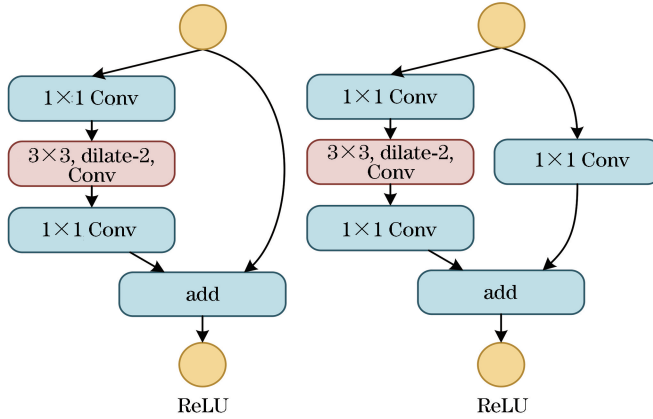


图 4 空洞卷积残差结构。(a)空洞卷积瓶颈层;(b)带 1×1 卷积结构的空洞卷积瓶颈层

Fig. 4 Structure of dilated convolution residuals. (a) Dilated convolution bottleneck; (b) dilated convolution bottleneck with 1×1 Conv projection

(YOLO v3 设定阈值为 0.5)时,保留置信得分较高的 A 飞机,将 B 飞机的预测框的置信得分置为 0,从而造成无法检测出 B 飞机的结果^[25]。本文使用线性衰减的 NMS 算法,解决了 YOLO v3 中 NMS 造成遮挡飞机无法准确检测的问题。当交并比 $I(M, b_i)$ 高于抑制阈值 N_t 时,对(2)式中置信得分 S_{confi} 进行线性平滑处理,优化后的 NMS 算法表示为

$$S_{confi}^* = \begin{cases} S_{confi}, & I(M, b_i) < N_t \\ S_{confi} [1 - I(M, b_i)], & I(M, b_i) \geq N_t \end{cases} \quad (3)$$

式中, S_{confi}^* 为线性平滑处理后的置信得分。

优化后 NMS 算法处理流程图,如图 6 所示。具体步骤为:1)根据置信得分的大小,将回归产生的 N 个预测框依次排序;2)选择置信分数最大的预测框,并计算与其他预测框的 IOU 值;3)比较计算所得的 IOU 值与阈值大小,如果 IOU 值大于阈值,将使用(3)式重新计算置信分数,否则保持原值不变;4)将置信得分最大的预测框存入 L 中,执行 $N-1$,剩余的预测框重复步骤 1)~4),直到 $N=1$,即处理完所有预测框;5)删除 L 中置信分数小于阈值的预



图5 两架相互遮挡的飞机

Fig. 5 Two planes of occlusion

测框,显示置信分数大于阈值的预测框,即该预测框为飞机的最终检测结果。

改进后的骨干网络具有较高的分辨率,能够检测小目标所在区域的特征信息,且该网络具有更开阔的感受野,有助于从全局信息中定位场面图像中的小目标。改进后的 NMS 算法不再是简单的阈值比较法,而是一种通过采用 IOU 值线性衰减置信得分的方法,能避免误删遮挡目标的预测框,提升对遮挡目标的检测能力。

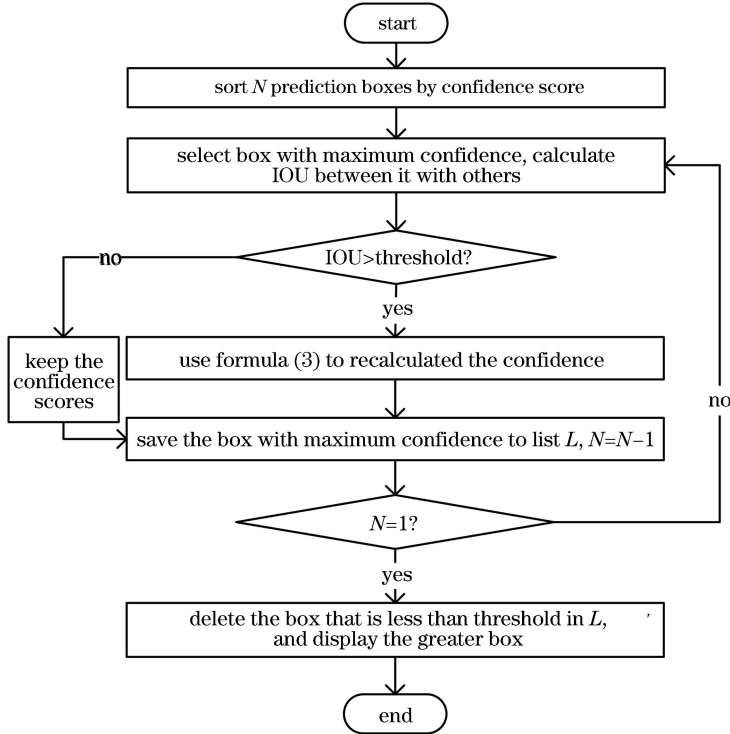


图6 优化后的 NMS 处理流程图

Fig. 6 Flow chart of the optimized NMS processing

4 实验与结果分析

实验使用 Darknet、TensorFlow^[26] 深度学习框架,在 2 块 NVIDIA 1070 Ti 显卡上完成训练及测试。

4.1 场面飞机数据集

目前,在机场及场面飞机检测的应用研究方面,缺乏公开的适合深度学习训练的数据集。本文构建场面飞机数据集 Aeroplane,数据主要来源于宁夏银川市河东国际机场监控视频系统、PASCAL VOC2007 与 VOC2012^[27] 的飞机图像。

Aeroplane 数据集包含了各种型号的单架客机、相互遮挡客机、小目标、其他目标(机场内的鸟、引导车、摆渡车、行人)等图像,共计 10000 张,如表 1 所示。其中 8000 张为训练集,2000 张为测试









集。此外,使用图像工具对所有图像裁剪及统一命名,且横、纵尺寸均保持在 375~500 pixel 范围内,横、纵偏差不超过 100 pixel。最后,采用 PASCAL VOC 格式的标注工具,人工对图像进行标注,并分为飞机和其他物体两种类别。

4.2 网络训练

基于数据集 Aeroplane 进行训练,类别为 2,输出特征图的数量为 21,Anchor 为 10×13 、 16×30 、 33×23 、 30×61 、 62×45 、 59×119 、 116×90 、 156×198 、 373×326 ^[26]。训练的批次为 64,动量为 0.9,衰减为 0.0005,最大迭代次数为 60000。开始训练时采用 10^{-3} 的学习率,以达到稳定整个网路的效果,迭代 20000 次后调整为 10^{-2} ,迭代 30000 次后调整为 10^{-3} ,迭代 30000 次后调整为 10^{-4} 。

表1 场面飞机数据集

Table 1 Airport scene aircraft data sets

| Category | Single aircraft | Overlapped aircraft | Small target | Others |
|----------------|---|---|--|---|
| Original image |  |  |  |  |
| Labeled image |  |  |  |  |
| Count | 4500 | 2000 | 2000 | 1500 |

在训练过程中,通过绘制损失(Loss)曲线能够直观地观察到训练的动态过程。图7为在训练过程中对应的Loss曲线,横坐标表示训练迭代次数,纵坐标表示训练过程中的Loss值。可知在开始训练时,Loss值下降较为迅速;当迭代训练20000次后,Loss曲线平稳下降;当迭代至60000次时,Loss收敛至0.2,结束训练。

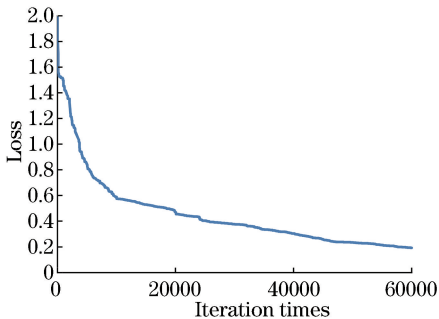


图7 Loss曲线

Fig. 7 Loss curve

4.3 结果及分析

为验证本方法能够较好地检测小目标和相互遮挡的飞机,进而整体地提升场面飞机的检测效果,相关人员将在Aeroplane的测试集上进行3个对比实验。

1) 小目标检测

从测试集中随机抽取尺寸为10 pixel × 20 pixel、20 pixel × 40 pixel、30 pixel × 60 pixel、40 pixel × 80 pixel、50 pixel × 100 pixel、60 pixel × 120 pixel的场面飞机图像,使用梯度直方图和支持向量机(HOG + SVM)^[4]、Faster-RCNN^[6]、SSD^[9]、YOLO v3^[10-11]和本文方法进行对比实验。

如图8所示,使用不同检测方法对不同尺寸的小目标飞机进行检测。横坐标为飞机尺寸(Size),纵坐

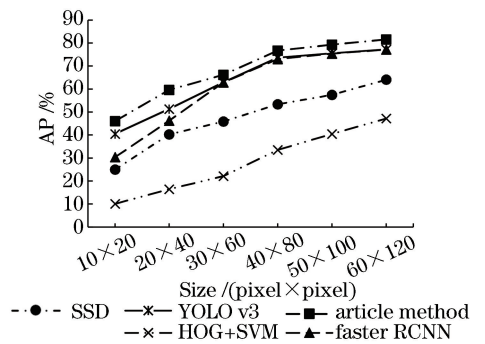


图8 不同方法检测多尺度小目标的结果

Fig. 8 Detecting results of multi-scale small targets by different methods

标为平均准确率(AP),随着飞机尺寸的增加,AP数值逐渐升高。当飞机尺寸为60 pixel × 120 pixel时,HOG + SVM、Faster-RCNN、YOLO v3、SSD和本文方法的AP分别达到了47.3%、77.2%、77.2%、64.1%、81.6%。在飞机尺寸为40 pixel × 80 pixel以下,本文方法的AP明显高于其他方法,由此可见,本方法在较小目标的检测方面具有优势。

2) 飞机相互遮挡检测

从测试集中随机抽取200张飞机遮挡图像,按照横向、纵向遮挡比例进行分组实验。不同遮挡比例的检测性能对比,如表2所示。由表2结果可知,在横向遮挡中,遮挡比例在20%以下,两种检测方法的检测准确率相同;当遮挡比例为20%~40%时,本文方法比YOLO v3检测准确率高30%;遮挡比例更高时,本文方法更能体现出在遮挡飞机检测方面的优势。在纵向遮挡中,两种方法的检测均较低,主要原因为纵向飞机尺寸较小,很难依据IOU值区别出两架飞机。本文方法采取高分辨率的高层特征信息,根据线性衰减NMS的方式,保留被遮挡飞机预测区域。在纵向遮挡中,两种检测方法具

表2 不同遮挡比例的检测性能对比

Table 2 Detection performance comparison of different overlapped proportions

| Method | Overlapped direction | Overlapped proportion / % | AP / % |
|----------------|----------------------|---------------------------|--------|
| YOLOv3 | Horizontal | 0-20 | 90 |
| | | 20-40 | 60 |
| | | 40-60 | 40 |
| | | 70-90 | 10 |
| | | Vertical | 0-90 |
| Article method | Horizontal | 0-20 | 90 |
| | | 20-40 | 90 |
| | | 40-60 | 60 |
| | | 70-90 | 20 |
| | | Vertical | 20-80 |



图9 不同遮挡程度的飞机检测对比实验。(a)(b)遮挡比例接近于20%;(c)(d)遮挡比例接近于60%;(e)(f)有明显颜色特征、遮挡比例接近于60%;(g)(h)遮挡比例接近于80%

Fig. 9 Contrast experiments of aircraft detection with different occlusion proportions. (a)(b) Occlusion is close to 20%; (c)(d) occlusion is close to 60%; (e)(f) obvious color characteristics, occlusion is close to 60%; (g)(h) occlusion is close to 80%

3) 不同检测方法的性能对比

表3为本文方法及各类检测方法在场面飞机数据集上测试性能对比, P 表示飞机检测的准确率, v_{FPS} 表示检测的速度,本文方法的 P 和 v_{FPS} 分别为83.7%、26 frame \cdot s $^{-1}$,比传统的HOG+SVM检测方法提升了34.1%和12 frame \cdot s $^{-1}$,得益于深度学习的检测方法具有提取深层特征的能力,同时YOLO、SSD系列算法将目标检测看成回归问题,提升了检测速度;本文方法与Faster-RCNN相比,检测精度也提高了4.1%,主要原因是本文方法沿用了Anchor机制、FPN结构,提升网络对目标的检测准确率;本文方法的检测准确率较SSD、YOLO v3分别提高了13.2%和11.4%,表明本文方法的骨干网络能保持较高的分辨率、开阔的感受野,有利于提高场面飞机检测的准确率。本文方法的检测速度较YOLO v3降低了8 frame \cdot s $^{-1}$,主要是由于本文方法采用了空洞卷积(比普通卷积耗时),且较高的特

有相同的检测准确率。本文方法在一定程度上弥补了YOLO v3对场面飞机间遮挡造成漏检的不足。

图9为本文方法与YOLO v3对比例遮挡飞机的检测效果对比图。图9(a)(b)中飞机遮挡比例较小,两种算法均能检测到三架飞机。图9(c)(d)中遮挡比例接近于60%,YOLO v3只能检测到靠前的飞机,无法检测被遮挡的飞机。图9(e)(f)中两架飞机具有明显的颜色特征的遮挡飞机图像,YOLO v3检测效果与图9(c)相同,而本文方法却能够检测到被遮挡的飞机。图9(g)(h)存在较为严重的遮挡,遮挡比例接近于80%,YOLO v3无法检测被遮挡的飞机,本文方法对被遮挡的飞机有良好的特征提取效果,且改进后的NMS算法对预测框的置信得分衰减为非零,可保证对相互遮挡飞机的检测。

征图分辨率造成检测网络的计算量增加,从而降低了检测速度。虽然检测速度有所降低,但检测准确率得到提升,且满足不低于25 frame \cdot s $^{-1}$ 的实时性要求。

表3 多种检测方法性能对比

Table 3 Performance comparison of various detection methods

| Method of detection | P / % | AP / % | v_{FPS} / (frame \cdot s $^{-1}$) |
|---------------------|---------|--------|--|
| HOG+SVM | 49.6 | 43.6 | 14 |
| Faster RCNN | 79.6 | 71.8 | 12 |
| SSD | 70.5 | 63.1 | 28 |
| YOLO v3 | 72.3 | 68.4 | 34 |
| Article method | 83.7 | 74.2 | 26 |

5 结 论

将YOLO v3应用到机场场面飞机检测领域,可以实现场面飞机的实时有效检测。针对小目标及遮挡飞机难以检测的问题,在YOLO v3模型的基础

基础上,对骨干网络和 NMS 算法进行改进,在 Aeroplane 数据集上训练及进行对比测试。实验表明本文方法较大地改善了小目标、飞机遮挡的检测效果,将场面飞机检测准确率从 72.3% 提高到 83.7%,速度达到 $26 \text{ frame} \cdot \text{s}^{-1}$,满足实时性检测要求。但本文算法在检测严重遮挡的飞机图像方面仍有不足,这也是后续的研究方向。

参 考 文 献

- [1] Yang H Y, Yang B, Wu X P, *et al.* Research and prospect of intellectualized air traffic management technology [J]. *Advanced Engineering Sciences*, 2018, 50(4): 12-21.
杨红雨, 杨波, 武喜萍, 等. 智能化空管技术与展望[J]. *工程科学与技术*, 2018, 50(4): 12-21.
- [2] Li X B, Wang S J, Jiang B T, *et al.* Airplane detection using convolutional neural networks in a coarse-to-fine manner [C] // 2017 IEEE 2nd Information Technology, Networking, Electronic and Automation Control Conference (ITNEC), December 15-17, 2017, Chengdu, China. New York: IEEE, 2017: 235-239.
- [3] Huang J, Jiang Z G, Zhang H P, *et al.* Ship object detection in remote sensing images using convolutional neural networks[J]. *Journal of Beijing University of Aeronautics and Astronautics*, 2017, 43(9): 1841-1848.
黄洁, 姜志国, 张浩鹏, 等. 基于卷积神经网络的遥感图像舰船目标检测[J]. *北京航空航天大学学报*, 2017, 43(9): 1841-1848.
- [4] Yao Y, Jiang Z G, Zhang H P. Aircraft detection in remote sensing image based on hierarchical classifiers [J]. *Spacecraft Recovery & Remote Sensing*, 2014, 35(5): 88-94.
姚远, 姜志国, 张浩鹏. 基于层次化分类器的遥感图像飞机目标检测[J]. *航天返回与遥感*, 2014, 35(5): 88-94.
- [5] Girshick R. Fast R-CNN [C] // 2015 IEEE International Conference on Computer Vision (ICCV), December 7-13, 2015, Santiago, Chile. New York: IEEE, 2015: 1440-1448.
- [6] Ren S Q, He K M, 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.
- [7] Redmon J, Divvala S, Girshick R, *et al.* You only look once: unified, real-time object detection [C] // 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE, 2016: 779-788.
- [8] Redmon J, Farhadi A. YOLO9000: better, faster, stronger [C] // 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), July 21-26, 2017, Honolulu, HI. New York: IEEE, 2017: 7263-7271.
- [9] Liu W, Anguelov D, Erhan D, *et al.* SSD: single shot multibox detector[M] // Leibe B, Matas J, Sebe N, *et al.* *Computer vision-ECCV 2016. Lecture notes in computer science*. Cham: Springer, 2016, 9905: 21-37.
- [10] Tumas P, Serackis A. Automated image annotation based on YOLOv3 [C] // 2018 IEEE 6th Workshop on Advances in Information, Electronic and Electrical Engineering (AIEEE), November 8-10, 2018, Vilnius, Lithuania. New York: IEEE, 2018: 18356488.
- [11] Qu H Q, Yuan T Y, Sheng Z Y, *et al.* A pedestrian detection method based on YOLOv3 model and image enhanced by retinex [C] // 2018 11th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI), October 13-15, 2018, Beijing, China. New York: IEEE, 2018: 18439663.
- [12] Dai C K, Li Y. Aeroplane detection in static aerodrome based on faster RCNN and multi-part model[J]. *Journal of Computer Applications*, 2017, 37(s2): 85-88.
戴陈卡, 李毅. 基于 Faster RCNN 以及多部件结合的机场场面静态飞机检测[J]. *计算机应用*, 2017, 37(s2): 85-88.
- [13] Lin T Y, Dollár P, Girshick R, *et al.* Feature pyramid networks for object detection [C] // 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), July 21-26, 2017, Honolulu, HI. New York: IEEE, 2017: 936-944.
- [14] Li Y L, Zhang Y K, Luo Z B. Target recognition technology of airport and airport aircraft based on deep convolution neural network [J]. *Journal of Chongqing Institute of Technology(Natural Science)*, 2018, 32(3): 210-216.
李耀龙, 张永科, 罗镇宝. 应用深度卷积神经网络的机场及机场内飞机目标识别技术[J]. *重庆理工大学学报(自然科学)*, 2018, 32(3): 210-216.

- [15] Ma J, Chen L, Gao Z Y. Hardware implementation and optimization of tiny-YOLO network[M] // Zhai G, Zhou J, Yang X. Digital TV and wireless multimedia communication. IFTC 2017. Communications in computer and information science. Singapore: Springer, 2018, 815: 224-234.
- [16] Kharchenko V, Chyrka I. Detection of airplanes on the ground using YOLO neural network[C] // 2018 IEEE 17th International Conference on Mathematical Methods in Electromagnetic Theory (MMET), July 2-5, 2018, Kiev, Ukraine. New York: IEEE, 2018: 294-297.
- [17] Dai W C, Jin L X, Li G N, *et al.* Real-time airplane detection algorithm in remote-sensing images based on improved YOLOv3 [J]. Opto-Electronic Engineering, 2018, 45(12): 81-89.
戴伟聪, 金龙旭, 李国宁, 等. 遥感图像中飞机的改进 YOLOv3 实时检测算法[J]. 光电工程, 2018, 45(12): 81-89.
- [18] Yu F, Koltun V, Funkhouser T. Dilated residual networks[C] // 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), July 21-26, 2017, Honolulu, HI. New York: IEEE, 2017: 472-480.
- [19] Bodla N, Singh B, Chellappa R, *et al.* Soft-NMS: improving object detection with one line of code[C] // 2017 IEEE International Conference on Computer Vision (ICCV), October 22-29, 2017, Venice, Italy. New York: IEEE, 2017: 5561-5569.
- [20] Chen Z X, Xie Z F, Zhang W B, *et al.* ResNet and model fusion for automatic spoofing detection[C] // Interspeech 2017, August 20-24, 2017, Stockholm, Sweden. [S.l.: s.n.], 2017: 102-106.
- [21] He K M, Zhang X Y, Ren S Q, *et al.* Deep residual learning for image recognition [C] // 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE, 2016: 770-778.
- [22] Wei Y M, Quan J C, Houyu Q Y. Aerial image location of unmanned aerial vehicle based on YOLOv2[J]. Laser & Optoelectronics Progress, 2017, 54(11): 111002.
魏湧明, 全吉成, 侯宇青阳. 基于 YOLOv2 的无人机航拍图像定位研究[J]. 激光与光电子学进展, 2017, 54(11): 111002.
- [23] Li Z M, Peng C, Yu G, *et al.* DetNet: a backbone network for object detection[J/OL]. (2018-04-19) [2019-02-19]. <https://arxiv.org/abs/1804.06215>.
- [24] Wu C Y, Yi B S, Zhang Y G, *et al.* Retinal vessel image segmentation based on improved convolutional neural network [J]. Acta Optica Sinica, 2018, 38(11): 1111004.
吴晨玥, 易本顺, 章云港, 等. 基于改进卷积神经网络的视网膜血管图像分割[J]. 光学学报, 2018, 38(11): 1111004.
- [25] Hua X, Wang X Q, Wang D, *et al.* Multi-objective detection of traffic scenes based on improved SSD[J]. Acta Optica Sinica, 2018, 38(12): 1215003.
华夏, 王新晴, 王东, 等. 基于改进 SSD 的交通大场景多目标检测 [J]. 光学学报, 2018, 38(12): 1215003.
- [26] Feng X Y, Mei W, Hu D S. Aerial target detection based on improved faster R-CNN [J]. Acta Optica Sinica, 2018, 38(6): 0615004.
冯小雨, 梅卫, 胡大帅. 基于改进 Faster R-CNN 的空中目标检测 [J]. 光学学报, 2018, 38(6): 0615004.
- [27] Ma Y J, Li X Y, Song X F. Traffic sign recognition based on improved deep convolution neural network [J]. Laser & Optoelectronics Progress, 2018, 55(12): 121009.
马永杰, 李雪燕, 宋晓凤. 基于改进深度卷积神经网络的交通标志识别[J]. 激光与光电子学进展, 2018, 55(12): 121009.