

基于双重分类深度学习的低空目标自动检测方法

钟立军, 林 彬, 王 杰, 甘叔玮, 张小虎
(中山大学 航空航天学院, 广东 广州 510275)

摘 要: 由于图像数量多, 因此准确、高效的目标检测是提升靶场光测图像处理自动化程度的关键步骤。针对低空目标图像及目标类型多、目标特性变化等情况导致传统目标检测算法适应性差的问题, 提出了一种基于双重分类深度学习的低空目标自动检测方法。该方法基于深度学习目标检测框架 YOLO V3, 根据低空目标的亮度和形状的双重属性特征, 将网络输出层中的单属性分类改进为双属性分类; 基于目标区域生长实现样本自动标注, 利用序列图像目标约束增加检测结果确认环节。靶场低空场景下的实际图像训练及检测结果表明: 该方法初步检测成功率高于 90%, 后处理之后取得了 99% 的检测成功率和 62% 的平均定位精度。

关键词: 双属性分类; 目标检测; 深度学习; 自动标注; 后处理

中图分类号: V 217

文献标志码: A

DOI: 10.19328/j.cnki.2096-8655.2022.02.014

An Automatic Detection Method for Low-Altitude Targets Based on Dual-attribute Classification Deep Learning

ZHONG Lijun, LIN Bin, WANG Jie, GAN Shuwei, ZHANG Xiaohu

(School of Aeronautics and Astronautics, Sun Yat-sen University, Guangzhou 510275, Guangdong, China)

Abstract: Due to the large number of images, accurate and efficient target detection is the key step to enhance the automation of the shooting-range photometric image processing. Aiming at the problem of poor adaptability of traditional target detection algorithms due to multiple low-altitude target images and target types and changes in target characteristics, this paper proposes an automatic low-altitude target detection method based on dual-attribute classification deep learning. The method is based on YOLO V3, a deep learning target detection framework, and improves the single-attribute classification in the output layer of the network to dual-attribute classification based on the dual-attribute features of luminance and shape of the low-altitude target; achieves automatic sample annotation based on target region growth, and confirms the detection results using sequential image target constraints. The actual image training and detection results in the low-altitude scenario of the range show that the initial detection success rate of the method is higher than 90%, and 99% detection success rate and 62% average localization accuracy are achieved after post-processing.

Key words: dual-attribute classification; target detection; deep learning; automatic labeling; post processing

0 引言

利用光电经纬仪进行跟踪式测量是靶场进行低空目标测量的主要手段。在对目标进行跟踪测量的过程中会产生大量图像, 目标类型一般为飞机和导弹, 其中, 飞机分为载机和靶机, 导弹分为试验弹和靶弹。一个任务会有多台设备参与整个任务

的跟踪, 以任务时间 5 min, 帧率为 100 帧/s、6 台设备、每台设备 3 个探头为例计算, 图像总量约为 $3 \times 6 \times 5 \times 60 \times 100 = 540\,000$ 帧, 而每个设备包含目标的有效跟踪段落只是其中一部分, 人工检查包含目标的段落会消耗大量时间。因此, 快速、准确地检测图像中的目标, 对靶场快速甚至实时进行数据

收稿日期: 2021-12-17; 修回日期: 2022-01-22

作者简介: 钟立军(1986—), 男, 博士, 主要研究方向为图像处理与测量、人工智能。

处理以及其他图像分析有着重要意义。

目前,目标检测方法主要有基于变化检测、基于人工设计特征检测和深度学习检测。

基于变化检测的思路不太适用于靶场目标检测,其原因主要有以下 2 点:1) 靶场低空图像主要为天空背景,纹理弱,甚至主要以噪声为主,若基于经纬仪角度进行运动补偿,很容易出现虚假目标;2) 为确保跟踪稳定,经纬仪会尽可能保证目标在图像中心,此时目标在图像上无明显运动规律,在目标保持匀速的情况下甚至会短时间内保持不动,且背景会变化,此时基于运动的检测方法将会失效。

基于人工设计特征和分类器的方法很难适用于靶场低空目标检测的原因有以下 3 点:1) 图像类型、目标表现形式、试验场景多,且由于跟踪距离远,存在大量的弱小目标;2) 基于人工设计特征的检测方法在定位目标和给出目标框的大小时依靠滑动窗口的方式进行,算法耗时大,很难实现快速检测;3) 基于人工设计特征和分类器的方法很难设计出能鲁棒的区分靶场各类目标的特征,也同样很难训练出能区分各类特征的分类器。因此,基于人工设计特征和分类器的方法在靶场应用具有很大的局限性。

靶场低空目标图像由于场景和目标的特殊性,其图像特点与公开数据集中的图像区别很大。公开数据集中的图像目标为生活中常见的人和物,成像清晰,纹理丰富;而靶场图像目标主要为武器,成像模糊,纹理较弱,一般为单通道灰度或红外图像,大多体现为弱目标或小目标。因此,目前的深度学习成果无法直接应用于靶场目标检测,需要针对靶场图像的特点进行分析与研究。本文结合多年靶场图像处理经验,研究如何将深度学习应用于靶场图像处理,使深度学习在靶场发挥更大的作用。由前述分析可知,对于靶场实际需求,检测速度是其最重要的需求,因此采用 YOLO V3 算法作为检测算法的基础框架,结合低空目标图像的特点进行改进,尤其是红外目标纹理和轮廓不清晰的情况,给出了一种改进的双重分类检测方法并结合场景特性的后处理方法。

1 相关工作

目标检测是计算机视觉领域的热点问题之一,相应的研究非常多,从早期的人工特征及变化检测到基于深度学习的方法,各类方法层出不穷,性能

也有了大步提升,尤其是目前基于深度学习的方法已经在检测性能上远超传统方法。

自 2012 年 KRIZHEVSKY 等^[1]在 ImageNet 大规模目标识别比赛(ImageNet Large Scale Visual Recognition Challenge, ILSVRC)中使用 AlexNet 卷积神经网络(Convolutional Neural Network, CNN)模型以绝对优势获得冠军后,基于深度学习的目标识别算法才引起了众多研究者的热切关注。深度学习方法是多层级的特征学习方法,能够将从原始数据开始的特征转化为更高阶、更抽象化的层级特征。CNN 可以有效地捕获隐藏的数据内部结构,得到大量更具代表性的特征信息,从而对样本进行更高精度的分类和预测。尤其是 ImageNet、VOC (Visual Object Class)、MS COCO (Microsoft Common Objects in Context) 等包含大量图像及标注信息数据集的发布,给深度学习的发展提供了必不可少的数据支撑。基于多层卷积网络的深度学习在图像处理领域本来就具有得天独厚的优势,涌现出众多具有代表性的目标检测算法,主要分为两阶段检测方法和单阶段检测方法。

1.1 两阶段的深度学习目标检测算法

2014 年 GIRSHICK 等^[2]提出区域卷积神经网络 (Regions with Convolutional Neural Network, R-CNN) 模型,成为目标检测深度学习算法的开端。与传统的目标检测方法相比,R-CNN 算法先进行选择性搜索以确定候选区域再进行识别和定位,增强了检测的针对性,提高了检测精度。该方法在 PASCAL VOC 2007 (Pattern Analysis, Statistical Modeling and Computational Learning Visual Object Class 2007) 上的检测率从 35.1% 提升到 53.7%。随后 HE 等^[3]在 R-CNN 模型的基础上提出了空间金字塔池化网络 (Spatial Pyramid Pooling Convolutional Neural Network, SPP NET),该模型在全连接层前面添加一个可伸缩的池化层-空间金字塔池化层 SPP,避免了因尺寸归一化而引起的信息失真问题。2014 年 GIRSHICK^[4]结合 SPP NET 的优点提出了具有并行结构的 Fast R-CNN 模型。该模型的主要特点是先对整幅图像进行卷积操作,再从特征映射中选择候选区域,这样极大地提高了训练速度;引入相当于单层 SPP NET 的候选区域池化层感兴趣区 (Region Of Interest, ROI),共享候选区域在卷积网络的前向传播过程提取每个候选区域的固定维

度特征表示;引入多任务损失函数,将支持向量机(Support Vector Machine, SVM)分类器边框回归合并到一个端到端的网络中。Fast R-CNN 模型融合了 R-CNN 和 SPP NET 的优点,大幅提升了训练与检测速度。2015 年 REN 等^[5]提出了 Faster R-CNN 算法,该算法的创新点主要体现在候选区域提取网络设计、候选区域提取网络训练和共享卷积特征 3 个方面,并提出了一种全新的提取候选区域网络(Region Proposal Network, RPN)。该网络被置于全幅图像 CNN 的后面,通过 Anchor 机制,确定滑动窗口内是否存在所要检测的目标,并进行窗口边界回归,提取候选区域。较之前的 Fast R-CNN 算法, Faster R-CNN 算法的检测速度提升了 10 倍。

2017 年在 Faster R-CNN 算法基础上, HE 等^[6]提出了 Mask R-CNN 算法。该算法可以实现目标识别、定位和分割 3 项任务。与 Faster R-CNN 算法相比, Mask R-CNN 算法在生成候选区域后添加了用来输出分割掩码的掩码层(Mask Branch),该层并行于分类层和边界回归层,其实质上是一组全卷积层 FCN,用来预测每个像素所属目标类别实现分割。随着 R-CNN、Fast R-CNN、Faster R-CNN、Mask R-CNN 等算法的不断改进,基于候选区域的两阶段深度学习算法的检测率不断提高,检测速度也不断加快。

1.2 单阶段的深度学习目标检测算法

端对端单阶段的目标检测方法极大地加快了检测速度,弥补了 Faster R-CNN 实时性较差的缺点,具有代表性的此类方法有 YOLO 系列、单炮多盒探测器(Single Shot MultiBox Detector, SSD)。

2016 年 REDMON 等^[7]提出了 YOLO 算法,其网络结构是在 GoogleNet 模型之上建立的 YOLO 算法,与 Faster R-CNN 算法相比,采用这种统一模型实现了端对端的训练和预测,其检测速度更快,背景误判率低,泛化能力和鲁棒性较好。但由于每个单元格仅对同一组类别进行边界框预测,这使得 YOLO 算法的定位准确率受到影响。

随后 YOLO 9000 算法对 YOLO 算法进行了改进,在每一个卷积层后添加批量归一化层,使用高分辨率分类器,采用新的特征提取网络模型 Darknet-19,采用 K-means 聚类方法对边界框做了聚类分析,将检测数据集和分类数据集联合训练^[8]。

2018 年原作者提出了 YOLO V3 算法,与

YOLO 9000 算法相比,该算法采用残差网络模型 Darknet-53 进行特征提取,并且利用特征金字塔网络实现多尺度检测。实验表明:YOLO V3 算法在检测速度上具有明显优势^[9]。

2020 年 YOLO V4 被提出,相比 YOLO V3 而言,性能有进一步提升^[12]。

2016 年鉴于 YOLO 算法定位准确度差,尤其是对小目标的检测遗漏问题, LIU 等^[10]提出了 SSD 算法。该算法去除了 YOLO 算法的全连接层,并借鉴 Faster R-CNN 算法的 Anchor 机制提出了默认框方法,对于底层的卷积层输出使用较小尺度的限定框,实现对小目标的检测。对于高层的卷积层输出,使用较大尺度的限定框实现对大目标的检测。

2017 年 FU 等^[11]对 SSD 算法进一步优化,提出了反卷积单炮多盒探测器(Deconvolutional Single Shot MultiBox Detector, DSSD)算法。不同于 SSD 算法直接在卷积层输出上做预测, DSSD 算法在更深层的反卷积模块上做预测,由于这种深层特征具有更大的感受野,因此这种深层特征和浅层特征相融合的方法为小目标提供了上、下文信息,进一步提升了对小目标的检测精度。

由于 YOLO 系列具有速度优势,因此在实际工程中有广泛应用。文献[13-17]结合不同的场景及目标对 YOLO 系列进行改进。

2 本文方法框架及流程

2.1 总体框架

本文方法主要流程如图 1 所示,主要过程包括 5 个步骤:目标自动标注、图像预处理、利用改进网络进行训练、检测、结合场景先验知识的后处理。



图 1 算法流程

Fig.1 Flow chart of the algorithm

2.2 靶场图像的多样性及特点

2.2.1 成像设备及图像类型

靶场的成像设备主要为光电经纬仪,同时搭载多个传感器,多目标进行跟踪测量。传感器类型主要包括可见光及红外探头,其中,可见光包含捕获电视和测量电视,捕获电视焦距较短,视场较大,用于捕获目标;测量电视焦距相对较长,观测距离较远,视场角小,一般仅为 $30'$ 左右,用于对目标进行精确跟踪及测量;红外探测器主要包含长、中波 2 种,其跟踪距离比可见光传感器远,且能全天候工作,但分辨率及帧频相对可见光较小。

2.2.2 目标类型及场景

针对的场景类型为武器试验及鉴定,目标主要为飞机和导弹,不包括火箭。按在任务中的作用,飞机分为靶机和载机,导弹分为试验弹和靶弹。

靶场主要场景包括空空、空地、低空、海面发射、水下发射等场景,各类场景特点不一。

2.2.3 图像及目标特点

光电经纬仪布置在地面上,往上观测空中目标,背景主要为天空,红外图像背景比较平整,呈现圆形斑。其目标显著性主要取决于目标与背景的温度差以及成像距离,对于主动段的导弹和带曳光管的飞机而言,成像较明显,此时较明显部分为尾焰及曳光管,表现为亮目标,弹体和飞机可能无法识别,需要针对具体的成像场景进行分析。对于被动段导弹及不带曳光管的飞机,由于相对背景温差较小,其红外图像表现为较弱的目标,而对于可见光图像其目标显著性主要取决于其颜色对比度与成像距离,同时对云等的遮挡比较敏感,作用距离相对于红外较短。

本文针对光电经纬仪拍摄的图像进行研究,即主要针对可见光图像与红外图像进行研究。考虑到飞机和导弹都会有暗、亮目标 2 种情况,比如导弹在被动段时其光学特性不明显,相对有尾焰的情况表现较暗,在有尾焰的情况下相对较亮;对于飞机,在其带曳光管时表现较亮,不带时表现较暗。同时考虑到飞机和导弹的特征差异,因此存在暗导弹、亮导弹、暗飞机和亮飞机共 4 类典型目标。

2.3 基于历史判读结果与区域增长的标注过程

由于目标图像较多,若完全依赖于人工标注,工作量巨大,且容易出错,因此利用目标序列图像

的特点,采用靶场的判读结果进行标注。利用寻找直方图波谷的方法获得分割自适应阈值,利用目标位置初值,通过区域增长获取目标区域。具体增长方法步骤如下:

步骤 1 统计得到全图灰度均值 I_A 及直方图;

步骤 2 统计得到以判读点为中心的 21×21 区域内的最大值 $I_{\max}$;

步骤 3 在直方图中由 I_A 到 $I_{\max}$ 进行遍历,寻找第一个像素点数 < 5 的灰度作为分隔阈值;

步骤 4 若不存在满足条件的点则取 $I_A + 0.05(I_{\max} - I_A)$ 。

标注结果如图 2 所示。除类似如图 3 所示的由于背景干扰导致框过大之外,其他样本均能较好地实现目标框自动标注,且由该方法得到的框相比人工标注带来的随机性而言更具有一致性,即目标框满足一定的标准。在自动标注的基础上对一些明显出错的框进行人工修正。

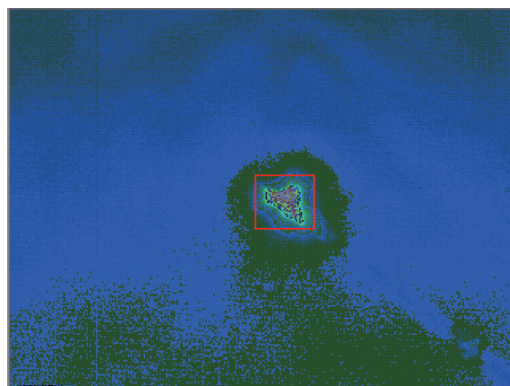


图 2 标注结果

Fig.2 Schematic diagram of labeling results

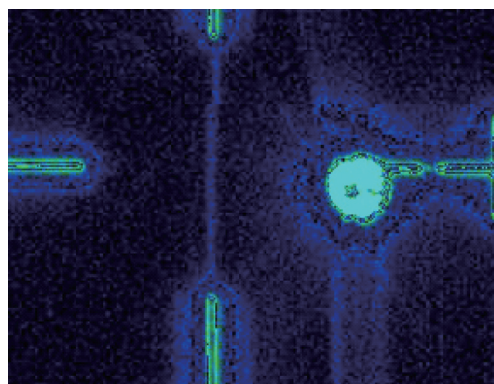


图 3 自动标注时受背景影响较大的目标

Fig.3 Schematic diagram for a target greatly affected by the background under automatic labeling

2.4 针对靶场图像的预处理

在将图像输入到网络之前,需要对其进行预处理,其主要包含 3 个步骤:去噪、增强、样本增强。

1) 去噪。采用高斯滤波对图像进行去噪。

2) 增强。由之前的分析可知,图像背景主要为天空,背景灰度较均匀,对于暗目标,背景灰度与天空灰度区别不大,需要进行一定程度的拉伸,提升可辨识度。同时可见光图像一般为 8 位或 10 位灰度图像,目标的显著度取决于目标的相对亮度,红外图像目标的显著度取决于目标的相对温度。因此可以利用该共同点对可见光与红外图像进行统一预处理。采用分段线性拉伸的方法对图像进行处理,以增强目标对比度。经大量靶场低空目标图像处理验证,该方法能较好地弱化背景,突出目标。

3) 丰富样本。深度学习中常用丰富样本的方法有随机切割、图像旋转、增加噪声、改变图像对比度和饱和度等。由于存在目标位于边缘的情况,因此在采

用随机切割时考虑目标是否位于边缘。对于目标位于图像边缘时,需保证目标位于切割后的图像内。

2.5 基于双重分类的网络结构改进

由前面对目标特点的分析可知,暗导弹、亮导弹、暗飞机和亮飞机以及其他目标最初按照 YOLO V3 的处理方法,则直接将其编号为 $1, 2, \dots, n$,但在实验中发现该分类方法检测结果并不理想。其主要原因是类型之间有共同点,1、2 和 3、4 分别为同一类目标,而 1、3 和 2、4 之间又分别具有同样的目标特点,目标同样较亮或较暗,导致网络分类困难,且解决 4 分类问题本身就比较解决 2 分类问题困难。

基于上述原因,将网络改成双重分类网络,即对于目标类型分成单独的 2 分类:一类针对目标亮度,即亮目标和暗目标;另一类针对目标固有类型,如导弹、飞机、火箭等。改进前的网络分类与改进后的网络分类如图 4 所示。

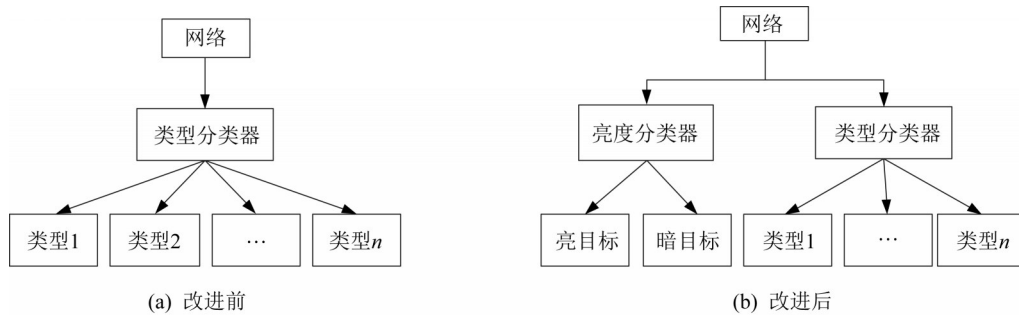


图 4 改进前与改进后的分类器

Fig.4 Classifier before and after improvement

因此对于每个框,输出向量变成 $(4 + 1 + C_{\text{classes}} + 2)$ 个,增加的 2 表示目标为亮目标或暗目标。总输出向量长度为 $w \times h \times 3 \times (4 + 1 + C_{\text{classes}} + 2)$,在本文试验中 class 为 2。

2.6 检测后处理

对于靶场序列图像进行目标检测,由于检测中不可避免地存在错检和漏检的情况,因此需要考虑序列的约束,去掉其中的错误检测,为后续的自动化处理提供准确的检测结果。以下介绍 2 种典型任务情况下的序列约束:

1) 对于空空或空地任务,即导弹由空中发射,目标位于空中或地面的情况,其任务过程为载机空中发射导弹到导弹击靶。因此,对于载机和导弹的

观测设备而言,在载机飞行过程中只有载机一个目标,从发射开始到载机出视场的过程中有 2 个目标,在导弹飞行过程中直到遇靶均只有一个目标,在相遇段过程中有 2 个目标,分别为导弹和靶。对于观测靶的设备而言,其在相遇段之前只有一个目标,在相遇段有 2 个目标。

2) 对于地空或地地任务,即导弹由地面发射,目标位于空中或地面的情况,其任务过程为地面发射导弹到导弹击靶。对于导弹的观测设备而言,在导弹遇靶前均只有一个目标,在相遇段过程中有 2 个目标,分别为导弹和靶;对于观测靶的设备而言,其在相遇段之前只有一个目标,在相遇段有 2 个目标。

对于同一个目标,其由像机光心到目标的角度满足一定的规律,不会存在大的跳变,具体规律可参考文献[18]。

3 试验过程及结果

采用基于判读结果与区域增长的标注方法对靶场实际序列图像进行标注,并对网络进行训练,同样用靶场实际图像进行验证,数据总量为 45 209 张,来自于 20 组序列图像,样本中总共标注目标 45 462 个,其中,导弹目标 9 185 个,飞机目标 36 277 个。

3.1 样本情况及训练过程

3.1.1 目标类型

在本文试验中,考虑到在后处理中需要利用目标的一致性约束,本文按实际情况将目标分为飞机与导弹 2 类。对于同一个探头而言,一般情况是跟踪同一个目标,部分探头会有相遇段出现,即出现 2 个目标的情况。

3.1.2 图像类型

图像类型主要包括 14 位红外图像、8 位可见光图像、10 位可见光图像及 24 位彩色图像。

3.1.3 目标框聚类

本文采用 K -means 聚类方法对数据集中的所有目标框进行聚类,聚类数设为 9,框之间的距离采用宽和高的平方和。由于针对的任务场景为目标跟踪过程,其目标相对距离为由远及近,再由近及远的过程,目标表现为由小到大,再由大到小的趋势。因此,初值采用均分的方式设置,即采用最大面积和最小面积之间的等分间隔作为每个聚类的大小初值,聚类获得的 9 组框宽高依次为 18.30、13.80、18.40、23.40、32.70、23.90、20.20、42.30、47.30、21.40、27.50、67.60、45.50、42.00、28.70、94.00、68.37、61.00。

3.1.4 训练过程

将样本标注好之后就输入网络进行训练,训练时先将图像进行预处理,然后将其统一为三通道图像,输入网络进行训练,在检测时同样进行相应的预处理。初始学习率为 0.001,对所有图像训练 50 个批次之后调整为 0.000 1。训练过程中需要调高目标比例权重。

3.1.5 提高目标比例权重

对于生活中的现实图像,图像中目标较多,因此 YOLO V3 中将是否存在目标参数 objectness 的比例设为 1:1,即跟非目标区域相等。在最初对靶场图像进行训练时,本文同样将该参数的权重设为 1,结果见表 1。表中:IOU (Intersection Over Union)

为训练过程中的预测区域重合度,应接近于 1;Class 为目标类型准确度,应接近于 1;Obj 为目标所在框有目标的概率,应接近于 1;表中的 No obj 为无目标区域有目标的概率,应接近于 0。由表中结果看出,在 No obj 趋向于 0 的同时,Obj 同样趋向于 0,在检测时,YOLO V3 网络用 $O_{obj} \times C_{classes}$ 表示该区域目标的概率,因此在网络检测时无法正确识别目标。其主要原因是对于靶场图像,大部分图像上只有一个目标,极少数图像上有 2 个目标,即网络得到的特征图上的大部分均没有目标,因此在残差整体调整时,其贡献较小,比较容易被其他无目标区域的调整淹没,收敛较慢,甚至不收敛,因此需要提高目标区域的权重。因此将目标区域的权重调整为 20。调整后的训练结果见表 1,此时网络残差均值为 0.3。

表 1 目标区域权重调整前后训练结果对比

Tab.1 Comparison of the training results before and after the target area weight adjustment

状态	Avg IOU	分类器 1	分类器 2	Obj	No obj
调整前	0.480	0.430	0.230	0.010	0.010
调整后	0.830	0.990	0.990	0.990	0.001

3.2 实验结果对比

3.2.1 试验设计

为分别验证改进后 YOLO V3 框架的有效性,进行 2 组试验:一组是用所有序列图的 1/10 进行训练,即训练样本为 4 520 张,另外 9/10 的图像用于检测;另一组是用其中 10 000 张图像用于检测,待检测样本来自于连续的序列图像,检测中的目标样本接近 1:1,用剩下图像的 1/10 用于训练。

目标检测问题同时包含分类与目标位置,即如何判断目标检测位置是否正确,考虑到本文图像中小目标较多,本文取 IOU 为 0.25 作为判断检测是否正确的阈值。

3.2.2 检测精度及可靠性

试验 1 检测阈值为 0.50,即目标可信度大于 0.50 时才认为是可靠目标。在验证过程中发现,由于目标较小,若采用 0.50 的 IOU 判别阈值容易导致一些检测到的目标视为无效,将 IOU 判别阈值采用 0.25,得到精确度与召回率曲线如图 5 所示。

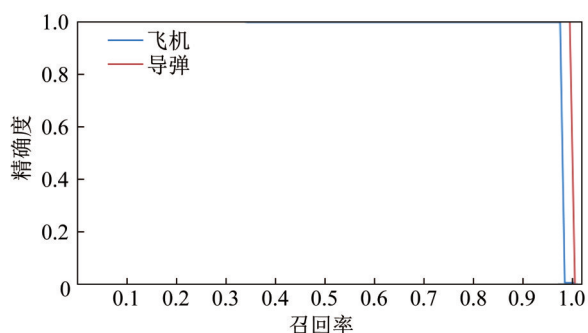


图5 试验1的精确度-召回率曲线

Fig.5 Precision-recall curve of Experiment 1

由图5可得,该方法针对靶场图像的检测率非常高,在召回率0.9时导弹和飞机的检测精度均在99%以上。其主要原因是序列图像中目标具有一定的相似性。其中, T_p (True Positive)=45 157, F_p (False Positive)=13 264, F_N (False Negative)=305,IOU均值为56.31%, $F_{1-Score}=0.87$ 。经过对检测结果进行分析,其中漏检的目标主要为 x 和 y 均低于图像区域3%的小目标,因此需要进一步提升小目标检测成功率的方法。

由 T_p 、 F_p 和 F_N 的值可知,仍然存在不少错误的检测,对于实现自动检测及后续处理而言,还需要利用序列图像之间的约束将其中的错误检测排除。在经过后处理之后,检测成功率均接近100%, $T_p=45\ 158$, $F_p=56$, $F_N=304$,IOU均值为72.77%, $F_{1-Score}=0.99$ 。经后处理确认后,去除了大量的错误检测。

在试验2中,检测阈值采用0.50,即目标可信度大于0.50时才认为是可靠目标,IOU判别阈值采用0.25,得到精确度与召回率曲线如图6所示。

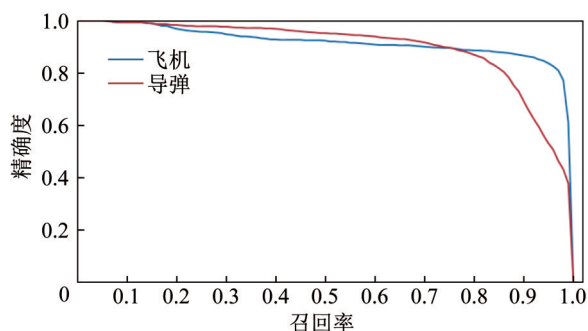


图6 试验2的精确度-召回率曲线

Fig.6 Precision-recall curve of Experiment 2

相比试验1的结果,试验2的导弹和飞机的检测率均有降低。检测的IOU均值为42.00%,相比之前也有一定程度的降低。此时 $T_p=9\ 042$, $F_p=4\ 442$, $F_N=658$, $F_{1-Score}=0.78$,平均精度为90.27%。

经过检测确认后, $T_p=9\ 015$, $F_p=5$, $F_N=685$, $F_{1-Score}=0.96$,检测的IOU均值为62.60%,检测确认后的PR曲线如图7所示。由上述数据与图可看出,检测处理后的效果有明显提升。飞机和导弹的检测效果如图8~图9所示。

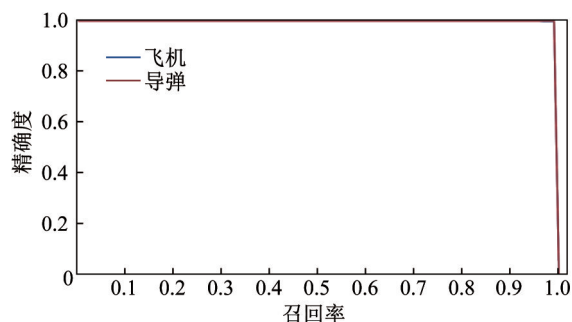


图7 试验2后处理之后的精确度-召回率曲线

Fig.7 Precision-recall curve of Experiment 2 after post-process

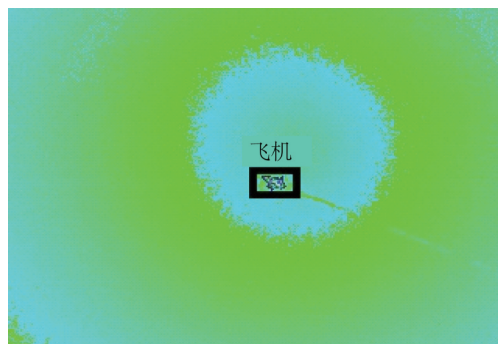


图8 飞机检测结果

Fig.8 Detection result of a plane

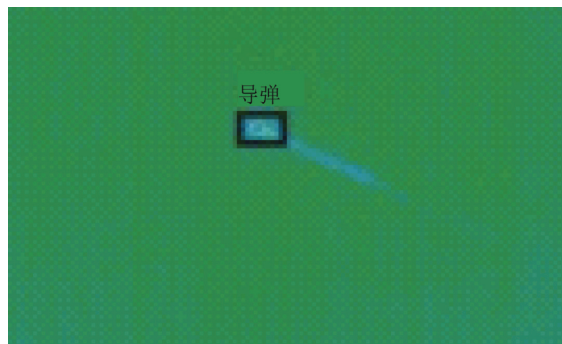


图9 导弹检测结果

Fig.9 Detection result of a missile

3.2.3 检测速度

该算法的主体框架为 YOLO V3 实现,在 YOLO V3 的全连接层上增加一个分类属性,以及在全连接层之前的滤波层增加了 2 个卷积核,其他结构完全一致,因此其运算量接近于 YOLO V3。在 GeForce RTX2080Ti 上运行时,平均每帧耗时 12.4 ms,帧频达到 80.6 帧/s,可以实现靶场低空目标快速检测需求。

3.3 试验结果分析

由试验结果得出,结合双重分类的 YOLO V3 网络结合靶场低空序列图像的特点,对靶场目标具有非常高的检测精度和速度,能同时适用靶场多种图像类型,可以有效提升靶场目标检测的自动化程度。

4 结束语

本文给出一种适用于靶场低空场景的红外与可见光目标检测的方法,靶场多种场景下的实际图像训练及检测结果表明:该方法能有效应用于靶场,提升靶场低空光测图像处理的自动化程度。该方法进一步的工作包括以下 2 个方面:1) 目前使用的样本主要针对低空场景,只包含导弹和飞机两类目标,需要用更多的目标类型进行验证;2) 目前结果弱、小目标的检测结果相对较差,后续需要考虑如何提升对于靶场低空弱、小目标的检测成功率,并针对更多场景目标的图像进行进一步的研究与测试。

参考文献

- [1] KRIZHEVSKY A, SUTSKEVER I, HINTON G. ImageNet classification with deep convolutional neural networks [J]. *Advances in Neural Information Processing Systems*, 2012, 25(2): 84-90.
- [2] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]// *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2014: 580-587.
- [3] HE K M, 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*, 2015, 37(9): 1904-1916.
- [4] GIRSHICK R. Fast R-CNN[C]// *Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition*. 2015: 1440-1448.
- [5] REN S Q, HE K M, ROSS G, 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.
- [6] HE K M, GKIOXARI G, DOLLÁR P, et al. Mask R-CNN [J]. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 2017, 5: 386-397.
- [7] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: unified, real-time object detection [C]// *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2016: 779-788.
- [8] REDMON J, FARHADI A. YOLO 9000: better, faster, stronger [C]// *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2017: 7263-7271.
- [9] REDMON J, FARHADI A. YOLO V3: an incremental improvement [EB/OL]. (2018-04-01) [2021-12-01]. <https://arxiv.org/pdf/1804.02767.pdf>.
- [10] LIU W, ANGUELOV D, ERHAN D, et al. SSD: single shot multibox detector [C]// *European Conference on Computer Vision*. Berlin: Springer, 2016: 21-37.
- [11] FU C, LIU W, RANGA A, et al. DSSD: deconvolutional single shot detector [EB/OL]. (2017-01-23) [2021-12-01]. <https://arxiv.org/pdf/1701.06659.pdf>.
- [12] ALEXEY B, WANG C Y, LIAO H Y M, et al. YOLO V4: optimal speed and accuracy of object detection [EB/OL]. (2020-04-23) [2021-12-01]. <https://arxiv.org/pdf/2004.10934.pdf>.
- [13] 牛戈, 陈小前, 季明江, 等. 基于注意力机制特征重建网络的舰船目标检测 [J]. *上海航天*, 2021, 38(4): 128-136.
- [14] 鞠默然, 罗海波, 王仲博, 等. 改进的 YOLO V3 算法及其在小目标检测中的应用 [J]. *光学学报*, 2019, 39(7): 245-252.
- [15] 马立, 巩笑天, 欧阳航空. Tiny YOLO V3 目标检测改进 [J]. *光学精密工程*, 2020, 28(4): 988-995.
- [16] 聂鑫, 刘文, 吴巍. 复杂场景下基于增强 YOLO V3 的船舶目标检测 [J]. *计算机应用*, 2020, 40(9): 2561-2570.
- [17] 王战涛, 张策, 王晓田. 基于 YOLO V3 的改进目标检测识别算法 [J]. *上海航天*, 2021, 38(6): 60-70.
- [18] 钟立军, 于起峰, 周颀鑫, 等. 基于位置预测的靶场图像实时判读方法 [J]. *国防科技大学学报*, 2020, 42(2): 85-91.