

基于深度学习的无人机识别方法现状与挑战

李其骏¹, 杨明^{1,2,3}, 任昊天¹, 常皓亮¹, 张小强^{1,2}, 朱新宇^{1,2}

(1. 中国民用航空飞行学院 航空电子电气学院, 广汉 618307)

(2. 四川省通用航空器维修工程技术研究中心, 广汉 618307)

(3. 电子科技大学 光电科学与工程学院, 成都 610054)

摘要: 无人机在军事/民用/商业领域的广泛应用促使对其识别和分类的需求。随着人工智能的不断发展, 深度学习作为一种机器学习技术, 在目标检测领域展现出良好的性能, 也可应用于无人机识别领域。本文首先介绍了无人机识别的背景和意义, 回顾了深度学习的发展历程, 并分别介绍了目标检测中两种重要的算法结构: 两阶段目标检测算法和单阶段目标检测算法; 然后, 对目标检测常用算法以及算法中的骨干网络等进行了阐述, 归纳了近年来无人机识别改进算法的改进策略, 总结了改进效果及其缺点和局限性; 最后, 针对目前无人机识别的研究现状, 提出展望和挑战, 有望在建立无人机数据集, 提高无人机检测的准确性、实时性等方面取得更大突破, 推动无人机技术在各个领域的应用。

关键词: 目标检测; 无人机; 深度学习; 计算机视觉; 神经网络

中图分类号: V279

文献标识码: A

DOI: 10.16615/j.cnki.1674-8190.2025.02.01

Status and challenges of UAV recognition methods based on deep learning

LI Qiqin¹, YANG Ming^{1,2,3}, REN Haotian¹, CHANG Haoliang¹, ZHANG Xiaoqiang^{1,2}, ZHU Xinyu^{1,2}

(1. Institute of Electronic and Electrical Engineering, Civil Aviation Flight University of China, Guanghan 618307, China)

(2. Sichuan Engineering Technology Research Center of General Aircraft Maintenance, Guanghan 618307, China)

(3. School of Optoelectronic Science and Engineering, University of Electronic Science and Technology of China, Chengdu 610054, China)

Abstract: The wide range of military, civil, and commercial applications of UAVs has prompted the need for their recognition and classification. With the development of artificial intelligence, deep learning, as a machine learning technique, has shown good performance in the field of object detection, and is also applied to the field of UAV recognition. This paper firstly introduces the background and significance of UAV recognition, reviews the development history of deep learning, and introduces two important algorithm structures in object detection: two-stage detector and single-stage detector. Secondly, it describes the common algorithms for object detection and the backbone network in the algorithms, and then summarises the improvement strategies of improved algorithms for UAV recognition in recent years, and summarises the improvement effect and its shortcomings and limitations. Finally, the outlook and challenges are discussed with respect to the current research status of UAV recognition, which is expected to make greater breakthroughs in establishing UAV datasets, improving the accuracy and real-time performance of UAV detection, and promoting the application of UAV technology in various fields.

Key words: object detection; unmanned aerial vehicle; deep learning; computer vision; neural networks

收稿日期: 2023-11-22; 修回日期: 2024-05-20

基金项目: 中央高校基本科研业务费专项资金资助(ZJ2023-012, PHD2023-007); 四川省通用航空器维修工程技术研究中心资助课题(GAMRC2021ZD01)

通信作者: 杨明(1986—), 男, 博士, 讲师。E-mail: yangming932@163.com

引用格式: 李其骏, 杨明, 任昊天, 等. 基于深度学习的无人机识别方法现状与挑战[J]. 航空工程进展, 2025, 16(2): 1-11.

LI Qiqin, YANG Ming, REN Haotian, et al. Status and challenges of UAV recognition methods based on deep learning[J]. Advances in Aeronautical Science and Engineering, 2025, 16(2): 1-11. (in Chinese)

0 引言

在无人机技术迅速发展的背景下,无人机作为一种具有众多优点的飞行器,广泛应用于侦察、监视、搜索救援、航拍摄影等领域。相较于传统飞行器,无人机可以在无人操控的情况下自主飞行,其飞行动作也更加灵活机动,且成本更低,资源利用效率高,能适应不同的环境和任务需求^[1]。

然而,无人机的应用也带来了一系列问题和潜在危害。近年来,关于无人机的社会危害事件频发,包括威胁国防安全、扰乱社会秩序、破坏交通秩序、危害航空安全以及侵害个人权益等^[2]。例如,无人机干扰了美国加利福尼亚州的森林大火扑灭行动,延误了灭火工作;在英国伦敦希思罗机场,一架无人机的飞行引发了航空安全警报,导致航班延误和取消;在中国也有类似的无人机危害事件,例如成都双流机场曾出现过几起无人飞行器危害事件,航空飞行安全受到了严重威胁。

为了加强对无人机的管控,无人机识别成为当今社会面临的重要挑战之一,需要加大研究和探索的力度。准确识别无人机是实施有效管控的基础和前提,并且可以有效预防非法活动和安全威胁。

从无人机的应用场景和特点来看,无人机识别的难点有:

1) 无人机的尺寸小:无人机存在不同的型号,大部分正在使用的无人机尺寸较小,在图像中的可见性和识别度都会受到限制。

2) 复杂的背景:一般有无人机出现的场景,都是在低空,其背景中有天空、建筑物、树木等与无人机外观相似或容易导致误判、误识别的物体,会对目标识别产生一定的干扰。如果无人机处于复杂的环境中,他们会与背景混合并融入其中,更加难以识别。

目前,现代无人机主要采用雷达检测^[3]、射频检测^[4]、光电检测^[5]和声波检测^[6]等方法来进行

识别。

深度学习作为现代非常实用的有效学习方法,也在无人机目标识别中展现出了巨大的潜力和应用前景。深度学习模型具有较强的特征提取和表示能力,可以从大量的无人机图像数据中学习复杂的视觉特征。利用深度卷积神经网络等模型,可以自主学习图片中的纹理、形状、色彩等特征,从而有效地区分不同类别的无人机目标^[7-8]。相较于传统方法,深度学习模型能够更快速地实现目标检测、分类和定位等任务,提高了识别效率和准确性^[9]。

因此,深度学习在无人机目标识别中的应用前景广阔。通过深度学习模型的引入和改进,可以提高识别无人机的准确性和实时性,为无人机在军事、安防、航空等领域的应用带来更大的潜力和机会。

本文讨论目标检测的体系结构及其发展历程,在回顾数据集及评价指标后,对近年来无人机识别领域中的算法改进作出分析和整理;总结针对无人机目标识别的解决方案,并综合当前研究现状,提出未来的挑战与发展方向。

1 目标检测的体系结构

目标检测旨在识别图像或视频中的对象并确定其方位。在早期检测任务中,一般使用传统目标检测方法,这种检测方法中,特征设计往往依赖人工设计,存在移植性差、时间复杂度高等缺点,具有一定的局限性^[10]。随着深度学习的发展,不少目标检测方法都开始采用深度神经网络,其在检测识别方面,满足大数据时代的要求,具备强大的信息处理能力,并且出现了很多优秀的方法。而目前目标检测算法的结构主要可以分为两种:两阶段(two-stage)目标检测和单阶段(one-stage)目标检测,其发展历程如图1所示。

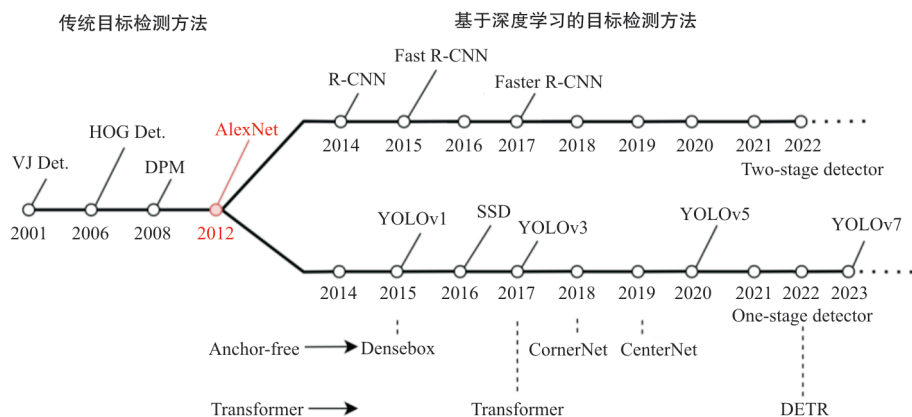


图1 目标检测的发展历程

Fig. 1 Evolution of object detection

1.1 两阶段目标检测

常见的两阶段目标检测算法有 R-CNN、Fast R-CNN、Faster R-CNN 等,两阶段目标检测主要将检测任务分为两个部分进行,第一个任务首先是寻找检测对象的位置以得到建议框,在有建议框之后,识别目标的位置得以确定,其最终识别的准确度和召回率都会有所提高;第二个任务是根据第一个任务中所得的建议框进行分类,进一步确定位置以保证精确性。使用两阶段检测算法可以获得更好的检测效果,主要表现在鲁棒性和准确性较高,针对不同的场景,可以通过调整参数来适应,但是这类算法通常因其复杂度较高而导致速度太慢。

2014 年, Girshick 等^[11]提出最初的区域卷积神经网络 R-CNN 算法,它使用选择性搜索方法提取候选目标区域,并对每个区域进行特征提取,然后通过卷积神经网络(CNN)进行分类和边界框回归,其基本结构如图 2 所示。该算法是目标检测算法中重要的里程碑,而其后以该算法为基础的改进算法,其精确度和实时性都在不断得到提升。

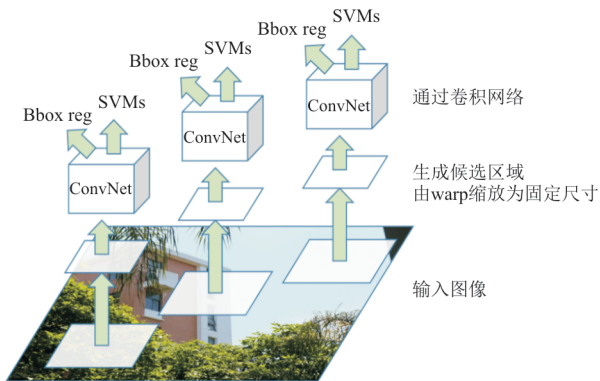


图 2 R-CNN 基本结构^[11]

Fig. 2 Basic structure of R-CNN^[11]

2015 年, Girshick^[12]在 R-CNN 算法的基础上进行了改进,提出了 Fast R-CNN 算法。它在卷积计算部分使用了 VGG16 网络代替了原有的 AlexNet 网络,并且设计了 ROI(Region of Interest) Pooling 层,解决了权值更新的问题。相比 R-CNN, Faster R-CNN 减少了对储存空间的占用,其精确性得到了有效的改善。

2017 年, Ren 等^[13]提出 Faster R-CNN 算法。经过 R-CNN 和 Fast R-CNN 的积淀, Faster R-CNN 沿用了 ROI Pooling 层,并将特征抽取 (fea-

ture extraction)、proposal 提取、bounding box regression、classification 都整合在一个网络中^[14],其最大的改进是引用了 RPN(Region Proposal Network)网络生成区域建议,基本结构如图 3 所示,检测速度获得显著提升。随着 Faster R-CNN 的问世,研究人员开始将这种算法投入到小目标检测的研究当中,邓珊珊等^[15]在针对小目标的识别中,对 Faster R-CNN 算法做出改进,通过使用对比度增强图像和高频增强图像作为训练样本,同时,将特征提取网络改为多层融合底层细节特征和高层语义特征,避免了在卷积过程中小目标特征消失的问题。这些针对小目标识别的算法改进,都对识别小型无人机有一定的启发作用。

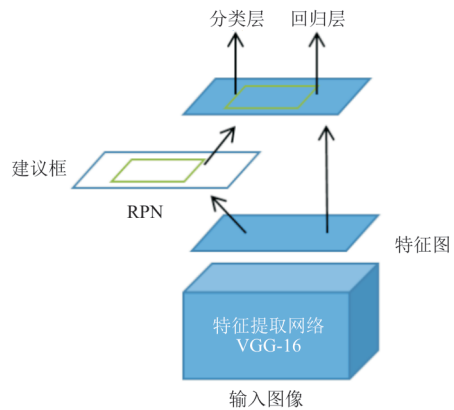


图 3 Faster R-CNN 基本结构^[13]

Fig. 3 Basic structure of Faster R-CNN^[13]

1.2 单阶段目标检测

常见的单阶段目标检测算法有 YOLO 系列和 SSD 算法,这类算法只需要一次检测就可以获得检测结果,因为其不同于两阶段目标检测算法,不需要得到建议框阶段,物体的类别概率和位置的坐标值可以直接产生。相比于两阶段目标检测,单阶段目标检测具有简单、速度快等优点,但是检测精度稍差。

2015 年, Redmon 等^[16]提出了 YOLO 模型,这是一种用于实时目标检测的深度学习模型,它的核心思想是将目标检测任务视为一个回归问题,在单个网络中同时预测目标的边界框和类别。与传统的目标检测方法相比, YOLO 模型将预测任务中的图片划分为几个正方形的网格,其中正方形的网格中又包含数个单元,在每一个网格单元中,又包含着边界框。YOLO 模型就是在这每个

单元中预测目标所属的类别。

2016年, Liu W 等^[17]提出 SSD 算法, SSD 算法的核心思想是利用卷积神经网络(CNN)同时进行目标的位置定位和分类。与两阶段目标检测算法(如 R-CNN 系列)相比, SSD 不需要使用复杂的区域建议网络(RPN)来生成候选框,从而在很大程度上简化了检测过程。它的检测速度快,可以多尺度检测,准确性较高,但是 SSD 算法也存在一定的缺陷,对于尺寸较小的目标可能无法有效检测,而对 SSD 模型进行算法改进后,可以有效识别小目标^[18]。

2017年, Redmon 等^[19]对 YOLO 模型进行了改进, 移除了全连接层, 并使用了与 Faster R-CNN 同样的做法: 使用卷积核 anchor boxer 预测边界框, 精度得到了提升。

2018年, Redmon 等^[20]又提出了 YOLOv3 算法, 使用 K-means 聚类 anchor 框, 采用多尺度训练, 并加入了 DarkNet53 残差网络模型、特征金字塔结构等改进, 使得 YOLOv3 算法拥有更高的精确度和更快的速度。

YOLOv4 算法^[21]在原有骨干网络 Darknet-53 的基础上增加了 CSP(Cross Stage Partial)模块, 在维持轻量的条件下, 既增强了 CNN 的学习能力, 又保证了精度。Shi Q B 等^[22]尝试在检测无人机目标中使用 YOLOv4 算法, 使得检测速度提升 4 帧/秒, m_{AP} 值达到了 95.64%。虽然目前使用 YOLOv4 模型检测无人机的研究较少, 但 YOLOv4 的检测精度高于传统的 YOLOv3 和 SSD 网络, 可以用于检测无人机。

2020年, Jocher 等^[23]提出 YOLOv5 模型。YOLOv5 采用了轻量级的主干网络, 如 CSPNet 和 EfficientNet 等, 以进一步提高计算效率。此外, YOLOv5 引入单阶段的训练过程, 通过不同大小的图像进行训练, 以实现多尺度的目标检测。但是 YOLOv5 对小目标检测有一定的限制。

2023年, YOLOv7 模型^[24]问世, 在整个 YOLO 系列中拥有更高的检测精度和更快的检测速度, 模型多、适应性强, 可以在不同的识别领域中得到应用。

从 YOLO 到 YOLOv7, 每个版本都引入了新的技术和方法, 推动着目标检测领域的发展。未来的 YOLO 版本有望进一步提升性能, 并应用于更多领域和应用场景。近几代 YOLO 模型在小目

标检测领域的研究和应用还不够成熟, 存在光照差, 遮挡严重情况下小目标检测准确性较差的问题等。目前 YOLOv3 在检测无人机领域中得到了广泛应用, 不少研究人员通过改进 YOLOv3 算法, 获得了较好的无人机检测效果。而具有更优模型的 YOLOv5 和 YOLOv7 也有望在识别无人机领域得到优化和应用。

1.3 Anchor-free 及 Transformer 方法

在目标检测领域, 单阶段目标检测算法和两阶段目标检测算法一直是研究的热点, 它们在检测准确性和速度之间提供了不同的权衡。最近, 一些新技术的出现引起了广泛的关注, 其中包括 Anchor-free 方法和 Transformer 方法。

在以往的目标检测任务中, 研究者们将其归结为一种从上到下的问题, 将这种方法称为 Anchor-based。作为目标检测方法中一种主流的方法, 它的核心思想是在图像中预定义一系列锚框(anchors), 这些锚框通常具有多个尺寸和宽高比。目标检测模型通过在这些锚框上进行分类和回归, 来检测图像中的目标。

具体而言, 锚框通常位于图像的不同位置, 并且与每个锚框相关联的是多个类别的概率分布及边界框的回归参数。在训练过程中, 模型通过与真实边界框进行比较, 来调整锚框的位置和形状。例如 Faster R-CNN 利用的就是典型的 Anchor-based 方法, 第一阶段生成候选框(Region Proposal), 第二阶段对这些候选框进行分类和精细的边界框回归。

但是这种方法可能存在某些缺陷, 比如在大部分没有物体的背景区域, 均匀采样 anchor 会导致大部分样本是无效的负样本。大小和长宽比都固定的 anchor 也使得检测具有局限性, 对于不同的检测任务需要设计不同的 anchor。

而 Anchor-free 方法摆脱了传统目标检测中使用锚框(anchor)的限制, 它通过直接从图像中学习目标的位置和形状, 消除对锚框的依赖。这种方法的优势在于其更好地适应不同尺寸和形状的目标, 同时简化了模型的设计。

虽然 Anchor-free 在 2015 年被首次提出^[25], 但当时并未引起重视, 直到 2018 年, 这种方法才被广泛应用于目标检测中, 研究人员提出了 CornerNet^[26]、ExtremeNet^[27]、CenterNet^[28] 等网络, 有效

地降低了超参数数量,模型得到了极大优化,目标检测的难度得到了进一步降低。

Transformer方法^[29]于2017年被提出,该方法最初是为自然语言处理任务设计的,但后来被成功地引入到计算机视觉领域,包括目标检测。与传统的卷积神经网络(CNN)不同,Transformer通过自注意力机制允许模型在处理图像时更好地捕捉全局上下文信息。

在目标检测中,Transformer方法通常使用注意力机制来动态地关注图像中的不同区域,从而有效地捕捉目标的相关信息。

DETR(Detection Transformer)是一个典型的例子,2020年,Carion等^[30]提出了一种将目标检测视为直接集预测问题的新方法,它通过将目标检测问题转化为集合的形式,通过Transformer模型直接输出目标的位置和类别,避免了传统两阶段方法的繁琐流程。

通过引入Transformer方法,目标检测模型能够更好地处理目标的上下文关系,提高了模型对目标复杂关系的建模能力,直到现在,这种方法都在优化和更新,为目标检测任务带来了新的思路和突破。

2 数据集与评价指标

2.1 数据集

数据集是用于训练、验证和测试机器学习或深度学习模型的样本集合。构建一个好的数据集对于训练准确和泛化能力强的模型至关重要。常见的数据集有MNIST^[31]、ImageNet^[32]、Tiny Images^[33]等,这些数据集内容丰富,涵盖了物体的不同尺寸、不同种类、环境光照变化、多种角度、复杂背景等情况,具有代表性、多样性和完整性。

然而目前对于无人机等低空小目标的研究非常有限,很少有统一的无人机数据集^[34],为了解决这个问题,大部分研究人员选择通过收集和拍摄无人机图像来作为自己的数据集。例如张杨等^[35]拍摄了目标与探测器距离约为15、70、100 m的可见光图像约1 500张;马旗等^[36]在不同时间、不同背景对3种民用无人机拍摄多角度、多姿态图像共3 524张;王靖宇等^[37]建立了15 000张低空无人机目标图像数据库,包括天空、草地、阴天、树林等不同背景的多姿态无人机图像等。

2.2 评价指标

在目标检测算法中,评价指标反映了目标检测算法对实际应用效果体现,主要有精确率(Precision)、召回率(Recall)、平均精度(Average Precision)和平均精度均值(Mean Average Precision)、每秒帧率(Frame Per Second)等,其定义及公式如下:

1) 精确率(P):预测真正样本与所有预测正样本的比值。

$$P = \frac{T_P}{T_P + F_P} \quad (1)$$

式中: T_P 为真正例; F_P 为假正例。

2) 召回率(R):预测真正样本与实际正样本的比值。

$$R = \frac{T_P}{T_P + F_N} \quad (2)$$

式中: F_N 为假反例。

3) 平均精度(A_P):正确检测个体的个体数占总个体的个数的百分率。

$$A_P = \int P(R) dR \quad (3)$$

4) 平均精度均值(m_{AP}):所有类别检测的平均精度, m_{AP} 值越大,检测精度越高。

$$m_{AP} = \frac{1}{M} \sum_{i=1}^M AP_i \quad (4)$$

5) 每秒帧数(F_{PS}):每秒检测的帧数, F_{PS} 值越大,检测性能越好,检测速度越高。

在无人机识别领域中,精确性和实时性是识别的关键,故研究人员通常采用 m_{AP} 和 F_{PS} 对改进的算法模型进行评估。

3 无人机识别的研究进展

随着深度学习的迅猛发展,在识别领域出现了许多卓越的方法,但是当这些方法被应用于无人机识别领域时,效果并不尽如人意,因为无人机的特殊性质,如其小尺寸、低速飞行以及多样化的外观和形状,使得传统的深度学习方法在无人机识别任务中面临一些困难。

Faster R-CNN和YOLO系列算法在检测小目标方面表现出色,SSD算法经过改进也能应用于小目标检测领域,这些算法同样适合用于识别无人机。为了提升无人机识别的效果,研究人员使用不同的改进方法对这三类算法模型进行了积

极探索和改进。

3.1 替换骨干网络

骨干网络是深度学习中用于提取高级特征表示的核心网络结构,它在目标检测中是不可或缺的基础模块之一。骨干网络主要将输入图像进行特征提取,转换为对应特征图的输出。骨干网络可分为复杂骨干网络(CBN)和轻量骨干网络(LBN),而复杂骨干网络拥有更强的特征表达能力、更高的检测精度,还拥有处理复杂背景、噪声等情况的优势,在识别无人机领域有着更广泛的应用。常见的骨干网络有VGGNet、ResNet(Residual Network)、HRNet、DenseNet等,不同的骨干网络对不同的目标也有着不同的效果,因此选对骨干网络是提升目标检测精度的关键。

ResNet是一种在深度学习中常用的网络结构,残差网络的提出旨在解决深层神经网络的训练困难和梯度消失问题^[38]。在传统的神经网络中,信息通过层与层之间的连接流动,每一层都需要逐步学习输入和输出之间的映射关系。然而,随着网络层数的增加,梯度逐渐衰减,导致训练过程变得困难,网络的性能也不一定提升,反而有可能因为深度增加而降低。

而残差网络通过引入残差连接来解决这个问题。残差连接是指将输入特征直接添加到网络的输出中,使得网络可以学习输入和输出之间的残差,如图4所示。这样,网络可以更容易地学习残差部分,而不需要从头开始学习输入和输出之间的映射。通过残差连接,网络可以有效地传递梯度并保持梯度的有效流动,从而促进深层网络的训练。

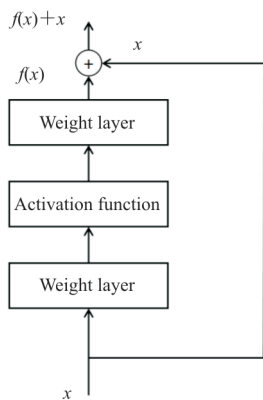


图4 残差网络学习模块^[38]

Fig. 4 Residual network learning module^[38]

ResNet在小目标识别领域十分重要,不少研究人员在优化网络结构时都选择了引入深度残差网络。

张杨等^[35]在研究无人机识别时,为了提高Faster R-CNN的检测效果采取了一系列措施,首先使用ResNet-18代替原有卷积层更为复杂的ResNet-50作为算法的特征提取网络,算法参数量得到有效减少,再通过改进FPN中特征的融合方式,增强目标与背景之间的特征差异,最后 m_{AP} 值提高了0.7%,算法速度提高至34.5帧/秒,为原算法的一倍。

Cheng J H等^[39]为了解决无人机与空中背景产生混淆问题,根据无人机的体积、形状等特征,采用区域目标识别算法,吸收了ResNet-50中群卷积的特征,对Fast R-CNN算法进行改进,得到优化后的算法GC-faster-RCNN,在1080p图像质量下,该算法的准确度达到了94.8%,识别速度为8帧/秒,可以对图片中1~5架小型无人机进行定位。尽管地面的建筑物和车辆干扰了实验的结果,并且和YOLO系列算法相比,检测速度仍有待提高,但对于低空无人机仍然是一种有效的定位检测方法。

康兢^[40]针对单目标物的检测,以Faster R-CNN算法为基础,将AlexNet替换为VGG网络结构,并对损失函数进行改进,加入了假阳性样本的训练损失,改进后 m_{AP} 值达到85.4%,但所用的数据集都是从网络上获得的无人机图像,数量和种类都偏少,在实际应用中如果只使用这样的数据集,无法保证检测的多样性和准确性。

李彦锴等^[41]通过优化YOLOv3,使用EfficientNet骨干网络代替YOLOv3算法中的DarkNet-53网络,提升了对空中红外目标的检测精度,同时简化了模型;为了进一步优化模型和提升精度,他们选择了CIoU作为网络的损失函数。结果表明,在使用EfficientNet-B4作为网络的主干特征提取网络时, m_{AP} 值相比YOLOv3提高了0.38%,而加入CIoU函数后, m_{AP} 值相比YOLOv3提高了1.17%,但是这两种方法在检测速度上略低于原YOLOv3算法,不过仍然可以满足红外场景下检测目标的实时性要求,且模型大小降低了50.03%。

刘鹏飞等^[42]针对无人机在高层卷积特征下的表达较差问题,引入VGG16特征提取网络的低层特征图Conv3_3,无人机的识别精度得到了一定的

改善,相比于原始的SSD算法,检测准确度提升7.32%。

3.2 图像融合

图像融合是将来自不同传感器或模态的图像信息进行整合,以提供更全面、准确的信息。红外与可见光图像融合是一种常见的融合方式,通过整合可见光图像和红外图像的信息,可以在目标检测等应用中取得更好的效果。

可见光图像对于人眼是可见的,包含了丰富的颜色和形状信息。而红外图像记录了场景中物体的热辐射信息,对于在可见光中难以察觉的目标(如在低光或无光条件下)具有很好的灵敏度。

红外与可见光图像融合的目标是将两种模态的信息结合,以充分利用它们的互补性,提供更全面的场景描述。

马旗等^[36]提出了采用双通道技术的无人机识别方法,即将红外图像和可见光图像一同传入深度残差网络,该方法能够提取特征并融合特征图,以增强特征的表达能力,并在此基础上采用文献[43]首创的多尺度预测网络(如图5所示)对无人机目标进行类别判断和位置回归。结果显示,采用平均融合的双通道YOLOv3-D相较于单一数据源的YOLOv3, m_{AP} 提升了约6.1%,同时,检测速度也得到了显著提升。

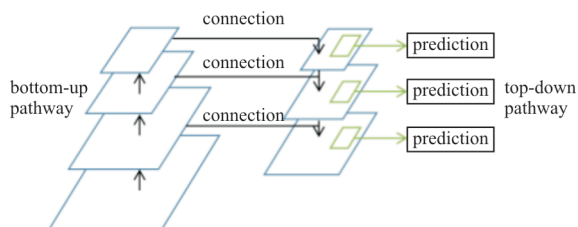


图5 多尺度预测网络的特征融合^[43]

Fig. 5 Feature fusion for multi-scale prediction networks^[43]

黄颖杰^[44]在研究无人机识别时,基于YOLOv3算法,选择了像素级融合,将红外和可见光图像融合,通过对融合前后的图像进行无人机识别测试,发现图像在经过融合后,平均精度均值提升了0.94%,证明了图像融合对无人机的识别具有正向作用。

3.3 多尺度特征融合

拍摄空中物体时,由于无人机飞行的高度不同,目标的尺度会根据距离的不同而出现“近大远

小”的多尺度特征,要解决无人机图像中尺度变化大一类的问题,可以通过多图像分辨率的方法来采集多尺度特征,并将多尺度进行融合^[45]。

多尺度特征融合是目标检测领域中常用的一种技术,用于提取和整合不同尺度下的图像特征,以改善对复杂场景或对象的识别性能。

多尺度特征融合的主要目的是克服单一尺度下的限制,使得系统能够更好地适应不同大小、形状和姿态的目标。而对于识别复杂背景下的小目标无人机,研究人员给出了相关方案。

马旗等^[46]以YOLOv3网络为基础,通过多尺度预测以及ResNet优化网络结构,提出O-YOLOv3网络,模型的精确性和泛化性得到了有效的提高,但因为网络参数的增加,检测速度有所下降,且数据集的类别有限,飞行器图像均是在光照良好的条件下获取,在其他环境下无人机的检测是否能高效地进行,还有待进一步验证。

梁日杰^[47]认为YOLOv3算法虽然对一般大小的目标能进行准确识别,但在对小目标例如无人机检测时,准确性明显有所降低,于是将传统的特征金字塔结构改进为双向特征金字塔,经过改进,检测精度提升了3.8%。但是如果在神经网络中算法的体积较大,就会存在资源大量消耗的问题,文献[41]指出,随着神经网络深度的加深,消耗的资源增多,其实时跟踪性甚至不如结构更为简单的滤波方法,如果只是靠采集视觉图像来检测无人机,会存在一些难以避免的问题。为了提升目标检测和识别的可靠性,还需要进行更深入的研究。

苏康^[48]基于SSD算法将推导的广义交并比损失公式应用到总损失函数中,替换原来的Smooth L1 loss,提高模型的检测性能;随后引入两个特征融合模块,通过提高位置损失占比来提高模型的检测性能;最后基于ShuffleNetv2算法搭建新的主干网络,降低了模型参数量和对内存的使用量。

3.4 引入注意力机制

注意力机制是一种模仿人类视觉注意力机制的方法,用于在深度学习模型中增强对输入数据的关注程度,使模型能够更有针对性地处理信息。注意力机制在图像处理、自然语言处理和其他领域中得到了广泛的应用。

人类的感知和认知过程通常会集中在特定区域,而忽略或减弱其他区域的信息。这种集中注

意力的机制使得人们能够更有效地处理信息,尤其是在复杂环境中。在深度学习模型中,注意力机制通过对输入的不同部分分配不同的权重来模拟注意力的分配。

常见的注意力机制有CA(Channel Attention)注意力机制和CBAM(Convolutional Block Attention Module)。CA注意力机制主要考虑增强网络对不同通道的关注程度。通过学习每个通道的权重,使网络更有针对性、选择性地加强或减弱输入的通道,以提高对特定特征的敏感性。CBAM是在CA注意力机制的基础上进一步引入了空间注意力,除了对通道进行关注外,CBAM还关注了每个空间位置的重要性,以捕捉更全面的特征信息。

杨辉跃等^[49]在解决无人机有效实施探测问题时,为了应对低空无人机灵活、微小等特性,除了

使用轻量级网络简化网络结构,使用双向加权特征金字塔替换PANet结构外,还特别在骨干网络中引入CA注意力机制,以提高无人机识别的精度。其提出的GCB-YOLOv5s算法的检测速度从48.38 帧/秒提升至53.97 帧/秒,平均精度达到97.5%。

张丰节^[50]为解决在复杂环境下,小目标无人机的特征提取不足问题,在CBAM的基础上提出了一种通道——空间注意力机制,以改变通道和空间的连接结构,除此之外,在特征融合时还采用了跳跃式连接方法,在预测网络中使用高斯加权的Soft-NMS替换原来的NMS非极大值抑制,改进后 m_{AP} 提高了5.8%。

上述使用不同改进策略对无人机识别的算法改进如表1所示。

表1 无人机识别的算法改进分析
Table 1 Algorithm improvement analysis for UAV recognition

算法来源	改进方法类别	基础算法结构	主要改进方法	改进效果	缺点以及局限性
文献[35]	替换骨干网络	Faster R-CNN	骨干网络中用ResNet-18代替ResNet-50	m_{AP} 提升至92.8%; F_{PS} 提升至34.5	距离超过100 m后识别效果较差
文献[39]	替换骨干网络	Faster R-CNN	骨干网络使用ResNet+Y+Gse	m_{AP} 有所提升,达到94.8%	检测速度过慢,只有8帧/秒,地面车辆和建筑物可能会对检测产生干扰
文献[40]	替换骨干网络	Faster R-CNN	将改进后的VGG网络结构替换掉AlexNet网络结构	m_{AP} 达到85.4%	数据集在数量和种类上不够全面,目标检测算法应用环境过于理想
文献[41]	替换骨干网络	YOLOv3	使用EfficientNet作为主干特征提取网络	模型尺寸减少了50.03%,精准度提升了1.17%	检测速度略低于原算法
文献[42]	替换骨干网络	SSD	引入VGG16特征提取网络的低层特征图Conv3_3	检测 A_P 提高了7.32%	可识别无人机图像光照条件单一,无人机类别较少
文献[36]	图像融合	YOLOv3	红外与可见光图像融合	m_{AP} 提升约6.1%,检测速度提升	复杂背景下配准困难,不能对此类数据进行验证
文献[44]	图像融合	YOLOv3	红外与可见光图像融合,将Canny和Orb算法融合	m_{AP} 为93.39%,比原来提高了0.94%	背景较为简单
文献[46]	多尺度特征融合	YOLOv3	使用搭建了深度残差网络和多尺度融合优化后的O-YOLOv3	m_{AP} 提高了8.29%	检测速度稍低于原始方法,夜间和昏暗光线下的识别还需要进一步验证
文献[47]	多尺度特征融合	YOLOv3	采用双向特征金字塔结构	检测精度提高了3.8%	目标跟踪算法体积过大,会消耗大量计算资源
文献[48]	多尺度特征融合	SSD	替换损失函数,引入特征融合模块,基于ShuffleNetv2算法重新搭建主干网络,提出SSD_GIS算法	降低模型参数量和内存使用量,速度提高71.4%	检测速度和复杂背景下的检测性能有待提高
文献[49]	引入注意力机制	YOLOv5	在骨干网络中引入CA注意力机制	F_{PS} 从48.38帧/秒提升至53.97帧/秒,平均精度达到97.5%	应用效果较好,没有讨论缺点
文献[50]	引入注意力机制	YOLOv5	在CBAM基础上提出了通道——空间注意力机制,以改变通道和空间的连接结构	m_{AP} 提高了5.8%	应用效果较好,没有讨论缺点

4 总结与展望

综合当前的研究现状,除了通过替换骨干网络^[35,39-42]、使用图像融合技术^[36,44]等方法提升无人机的检测效果和检测速度外,对于无人机尺寸小的难题,研究人员特别提出了以下解决方案:

1) 多尺度特征融合,引入金字塔结构可以通过在网络中构建不同分辨率的特征图来处理不同尺度的目标^[47,49];

2) 调整非极大值抑制(NMS)参数,使其更适应小目标检测^[50]。

对于在复杂背景下识别无人机的难题,研究人员特别提出了以下解决方案:

1) 引入注意力机制^[49-50],帮助模型更有效地关注目标区域,并抑制来自复杂背景的干扰;

2) 修改损失函数^[48],引入调制因子,将背景部分视为简单样本,减小背景损失在置信损失中的占比,从而提高在复杂背景下的精度。

综合当前的研究现状,目前利用深度学习识别无人机还面临着以下挑战:

1) 从数据集出发:目前还没有统一的无人机数据集,且研究人员自行创建的训练数据集数据较少,多为拍摄条件较好的白天,所训练出的模型不能够很好地反映在如黑夜、恶劣天气等复杂环境下识别无人机的精度^[42]。在实际应用中,需要更多的数据集来增加不同环境下识别不同种类无人机的准确性。

2) 从算法的模型和体积出发:为了使目标识别技术更易于应用,算法的模型也需要进行优化。在改进的算法中,尽管识别的精确度得到了提升,但是由于网络的深度增加,算法的体积增大,在计算上消耗了很多资源,提升了少量精度而大幅增加复杂度,在结果上可能会造成“弊大于利”的问题^[47]。处理大规模的数据集需要庞大的存储和计算资源,同时也要面临着复杂性和计算复杂度变高的挑战。希望在未来能够看到更加轻量和高效率的模型,为目标识别的应用提供更多的机会。

3) 从识别速度出发:精准度一直是算法模型发展的重要目标,研究人员一直期望能够构建更加准确的识别模型,对各种不同的数据进行精细化分析。然而在许多现实场景里,识别的速度同样很重要,例如在智能安防系统里及时发现未知无人机入侵,就需要一定的识别速度。但是在优

化算法提升识别精度的过程中,检测速度往往受到一定的损失^[48]。未来,根据特定的场景需求,在满足精确度的条件下,尽可能提升识别的速度是一项重要的挑战。

随着技术的不断进步和深度学习算法的改进,相信无人机检测方法会取得更大的突破,并在准确性、快速性和鲁棒性等方面得到更广泛的应用。

参考文献

- [1] 金伟, 尚勇. 中国无人机安全监管[J]. 科技导报, 2019, 37(14): 66-77.
JIN Wei, SHANG Yong. The safety supervision of unmanned aircraft systems in China[J]. Science & Technology Review, 2019, 37(14): 66-77. (in Chinese)
- [2] 刘静. 民用无人机滥用风险监管: 美国实践及其启示[J]. 时代法学, 2022, 20(5): 103-111.
LIU Jing. Risk regulation of civil UAV abuse: American practice and its enlightenment[J]. Presentday Law Science, 2022, 20(5): 103-111. (in Chinese)
- [3] 许道明, 张宏伟. 雷达低慢小目标检测技术综述[J]. 现代防御技术, 2018, 46(1): 148-155.
XU Daoming, ZHANG Hongwei. Overview of radar LSS target detection technology[J]. Modern Defence Technology, 2018, 46(1): 148-155. (in Chinese)
- [4] BISIO I, GARIBOTTO C, LAVAGETTO F, et al. Unauthorized amateur UAV detection based on WiFi statistical fingerprint analysis[J]. IEEE Communications Magazine, 2018, 56(4): 106-111.
- [5] MARVASTI F S, MOSAVI M R, NASIRI M. Flying small target detection in IR images based on adaptive toggle operator[J]. IET Computer Vision, 2018, 12(4): 527-534.
- [6] 朱少雄. 声纹识别系统与模式匹配算法研究[D]. 大连: 大连理工大学, 2006.
ZHU Shaoxiong. Research on voiceprint recognition system and pattern matching algorithm[D]. Dalian: Dalian University of Technology, 2006. (in Chinese)
- [7] 于越. 基于深度学习的无人机目标检测算法研究[D]. 西安: 中国科学院西安光学精密机械研究所, 2020.
YU Yue. Research on UAV target detection algorithm based on deep learning[D]. Xi'an: Xi'an Institute of Optics and Precision Mechanics, Chinese Academy of Sciences, 2020. (in Chinese)
- [8] 魏瑞歌. 基于深度学习的无人机小目标检测算法研究[D]. 成都: 电子科技大学, 2022.
WEI Ruige. Research on UAV small target detection algorithm based on deep learning[D]. Chengdu: University of Electronic Science and Technology of China, 2022. (in Chinese)
- [9] 刘伟. 基于视觉的无人机识别与跟踪技术研究[D]. 成都:

- 电子科技大学, 2022.
- LIU Wei. Vision-based UAV detection and tracking technology research[D]. Chengdu: University of Electronic Science and Technology of China, 2022. (in Chinese)
- [10] 司可意. 多源图像中低空小型无人机目标检测技术研究[D]. 沈阳: 沈阳航空航天大学, 2022.
- SI Keyi. Research on object detection technology of low altitude small UAV in multi-source images[D]. Shenyang: Shenyang Aerospace University, 2022. (in Chinese)
- [11] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]// 2014 IEEE Conference on Computer Vision and Pattern Recognition. Columbus, OH, USA: IEEE, 2014: 580-587.
- [12] GIRSHICK R. Fast R-CNN[C]// 2015 IEEE International Conference on Computer Vision (ICCV). Santiago, Chile: IEEE, 2015: 1440-1448.
- [13] 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.
- [14] 兰建雄, 苏义坤, 谭勇, 等. 基于无人机的仁新高速养护智能巡检系统设计[J]. 黑龙江交通科技, 2021, 44(9): 214-215.
- LAN Jianxiong, SU Yikun, TAN Yong, et al. Design of intelligent inspection system for maintenance of Renxin expressway based on UAV[J]. Communications Science and Technology Heilongjiang, 2021, 44(9): 214-215. (in Chinese)
- [15] 邓姗姗, 黄慧, 马燕. 基于改进 Faster R-CNN 的小目标检测算法[J]. 计算机工程与科学, 2023, 45(5): 869-877.
- DENG Shanshan, HUANG Hui, MA Yan. A small object detection algorithm based on improved Faster R-CNN[J]. Computer Engineering & Science, 2023, 45(5): 869-877. (in Chinese)
- [16] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: unified, real-time object detection[C]// 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas, NV, USA: IEEE, 2015: 779-788.
- [17] LIU W, ANGUELOV D, ERHAN D, et al. SSD: single shot MultiBox detector[M]. Cham: Springer International Publishing, 2016: 21-37.
- [18] 甄然, 苏康, 周金星, 等. 改进 SSD 的无人机目标检测研究[J]. 现代电子技术, 2022, 45(22): 33-37.
- ZHEN Ran, SU Kang, ZHOU Jinxing, et al. Research on UAV target detection based on improved SSD[J]. Modern Electronics Technique, 2022, 45(22): 33-37. (in Chinese)
- [19] REDMON J, FARHADI A. YOLO9000: Better, faster, stronger[C]// 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu, HI, USA: IEEE, 2017: 6517-6525.
- [20] REDMON J, FARHADI A. YOLOv3: an incremental improvement [EB/OL]. (2018-04-08) [2023-11-22]. <https://arxiv.org/abs/1804.02767>.
- [21] BOCHKOVSKIY A, WANG C Y, LIAO H Y M. YOLOv4: optimal speed and accuracy of object detection[EB/OL]. (2020-04-23) [2023-11-22]. <https://arxiv.org/abs/2004.10934>.
- [22] SHI Q B, LI J. Objects detection of UAV for anti-UAV based on YOLOv4[C]// 2020 IEEE 2nd International Conference on Civil Aviation Safety and Information Technology (ICCSIT). Weihai: IEEE, 2020: 1048-1052.
- [23] JOCHER G, CHAURASIA A, STOKEN A, et al. Ultralytics/YOLOv5: v7.0-YOLOv5 sota realtime instance segmentation [EB/OL]. (2020-11-22) [2023-11-22]. <https://ui.adsabs.harvard.edu/abs/2022zndo.7347926J/abstract>.
- [24] WANG C Y, BOCHKOVSKIY A, LIAO H M. YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors[C]// 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Vancouver, BC, Canada: IEEE, 2023: 7464-7475.
- [25] HUANG L, YANG Y, DENG Y, et al. DenseBox: unifying landmark localization with end to end object detection [EB/OL]. (2015-09-16) [2023-11-22]. <https://arxiv.org/abs/1509.04874>.
- [26] LAW H, DENG J. CornerNet: detecting objects as paired keypoints [M]. Cham: Springer International Publishing, 2018: 765-781.
- [27] ZHOU X Y, ZHUO J C, KRÄHENBÜHL P. Bottom-up object detection by grouping extreme and center points[C]// 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Long Beach, CA, USA: IEEE, 2019: 850-859.
- [28] DUAN K W, BAIS, XIE L X, et al. CenterNet: keypoint triplets for object detection[C]// 2019 IEEE/CVF International Conference on Computer Vision (ICCV). Seoul, Korea (South): IEEE, 2019: 6568-6577.
- [29] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need [EB/OL]. (2017-06-12) [2023-11-22]. <http://export.arxiv.org/abs/1706.03762>.
- [30] CARION N, MASSA F, SYNNAEVE G, et al. End-to-end object detection with transformers[M]. Cham: Springer International Publishing, 2020: 213-229.
- [31] LECUN Y, BOTTOU L, BENGIO Y, et al. Gradient-based learning applied to document recognition[J]. Proceedings of the IEEE, 1998, 86(11): 2278-2324.
- [32] DENG J, DONG W, SOCHER R, et al. ImageNet: a large-scale hierarchical image database [C] // 2009 IEEE Conference on Computer Vision and Pattern Recognition. Miami, FL, USA: IEEE, 2009: 248-255.
- [33] TORRALBA A, FERGUS R, FREEMAN W T. 80 million tiny images: a large data set for nonparametric object and scene recognition [J]. IEEE Transactions on Pattern

- Analysis and Machine Intelligence, 2008, 30(11): 1958-1970.
- [34] ASHRAF M W, SULTANI W, SHAH M. Dogfight: detecting drones from drones videos[C]// 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Nashville, TN, USA: IEEE, 2021: 7063-7072.
- [35] 张杨, 程正东, 朱斌. 基于改进 Faster R-CNN 的无人机小目标检测算法[J]. 量子电子学报, 2022, 39(3): 354-363. ZHANG Yang, CHENG Zhengdong, ZHU Bin. Small target detection algorithm of drones based on improved Faster R-CNN [J]. Chinese Journal of Quantum Electronics, 2022, 39(3): 354-363. (in Chinese)
- [36] 马旗, 朱斌, 程正东, 等. 基于双通道的快速低空无人机检测识别方法[J]. 光学学报, 2019, 39(12): 1210002. MA Qi, ZHU Bin, CHENG Zhengdong, et al. Detection and recognition method of fast low-altitude unmanned aerial vehicle based on dual channel [J]. Acta Optica Sinica, 2019, 39(12): 1210002. (in Chinese)
- [37] 王靖宇, 王霁禹, 张科, 等. 基于深度神经网络的低空弱小无人机目标检测研究[J]. 西北工业大学学报, 2018, 36(2): 258-263. WANG Jingyu, WANG Xianyu, ZHANG Ke, et al. Small UAV target detection model based on deep neural network [J]. Journal of Northwestern Polytechnical University, 2018, 36(2): 258-263. (in Chinese)
- [38] 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). Las Vegas, NV, USA: IEEE, 2016: 770-778.
- [39] CHENG J H, LIU Y Y, LI G N, et al. An efficient detection approach for unmanned aerial vehicle (UAV) small targets based on group convolution [J]. Applied Sciences, 2022, 12(11): 5402.
- [40] 康兢. 基于卷积神经网络的无人机目标检测算法研究[D]. 石家庄: 河北科技大学, 2021. KANG Jing. Research on UAV object detection algorithm based on convolutional neural network [D]. Shijiazhuang: Hebei University of Science and Technology, 2021. (in Chinese)
- [41] 李彦锴, 许媛媛, 刘子琪, 等. 基于改进 YOLO v3 算法的空中红外目标检测[J]. 红外技术, 2023, 45(4): 386-393. LI Yankai, XU Yuanyuan, LIU Ziqi, et al. Aerial infrared target detection based on improved YOLO v3 algorithm [J]. Infrared Technology, 2023, 45(4): 386-393. (in Chinese)
- [42] 刘朋飞, 周海, 冯水春, 等. 基于改进 SSD 的多尺度低空无人机检测[J]. 计算机工程与设计, 2021, 42(11): 3277-3285. LIU Pengfei, ZHOU Hai, FENG Shuichun, et al. Multi-scale low altitude UAV detection model based on improved SSD [J]. Computer Engineering and Design, 2021, 42(11): 3277-3285. (in Chinese)
- [43] 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). Honolulu, HI, USA: IEEE, 2017: 936-944.
- [44] 黄颖杰. 基于红外图像与可见光图像融合的无人机探测研究[D]. 广州: 广东工业大学, 2022. HUANG Yingjie. Research on UAV detection technology based on infrared and visible image fusion [D]. Guangzhou: Guangdong University of Technology, 2022. (in Chinese)
- [45] 祝宁华, 郑江滨, 张阳. 无人机航拍野生动物智能检测与统计方法综述[J]. 航空工程进展, 2023, 14(1): 13-26. ZHU Ninghua, ZHENG Jiangbin, ZHANG Yang. Review of intelligent detection and statistical methods of wild animals in UAV aerial photography [J]. Advances in Aeronautical Science and Engineering, 2023, 14(1): 13-26. (in Chinese)
- [46] 马旗, 朱斌, 张宏伟, 等. 基于优化 YOLOv3 的低空无人机检测识别方法[J]. 激光与光电子学进展, 2019, 56(20): 279-286. MA Qi, ZHU Bin, ZHANG Hongwei, et al. Low-altitude UAV detection and recognition method based on optimized YOLOv3 [J]. Laser & Optoelectronics Progress, 2019, 56(20): 279-286. (in Chinese)
- [47] 梁日杰. 小型四轴无人机目标识别与跟踪技术研究[D]. 长春: 长春理工大学, 2022. LIANG Rijie. Research on object recognition and tracking technology of small four axis UAV [D]. Changchun: Changchun University of Science and Technology, 2022. (in Chinese)
- [48] 苏康. 基于改进 SSD 的无人机目标检测研究[D]. 石家庄: 河北科技大学, 2022. SU Kang. Research on UAV target detection based on improved SSD [D]. Shijiazhuang: Hebei University of Science and Technology, 2022. (in Chinese)
- [49] 杨辉跃, 容易圣, 简钰洪, 等. 低空无人机实时探测的 GCB-YOLOv5s 算法[J]. 兵器装备工程学报, 2023, 44(7): 1-8. YANG Huiyue, RONG Yisheng, JIAN Yuhong, et al. GCB-YOLOv5s algorithm for real-time detection of a low altitude UAV [J]. Journal of Ordnance Equipment Engineering, 2023, 44(7): 1-8. (in Chinese)
- [50] 张丰节, 郁书好, 王冠凌, 等. 基于 YOLOv5 无人机小目标检测技术研究[J]. 牡丹江师范学院学报(自然科学版), 2023(4): 23-29. ZHANG Fengjie, YU Shuhao, WANG Guanling, et al. Research on small target detection technology of UAV based on YOLOv5 [J]. Journal of Mudanjiang Normal University (Natural Sciences Edition), 2023(4): 23-29. (in Chinese)

(编辑:马文静)