

# 基于深度学习的无人机视觉目标检测与跟踪

王博文 乔艳球\*

三亚学院 海南 三亚 572000

**摘要:** 深度学习驱动的无人机视觉目标检测与跟踪技术被深入探讨,这一技术对于无人机应用至关重要,显著拓宽了其应用领域并提升工作效率。无人机视觉目标检测涉及图像捕获、预处理、特征提取和分类器判断等环节。而深度学习在此方面的应用展现了卓越的优势,尤其是在网络架构设计、数据集准备和处理上。为了确保研究的有效性,实验设计和数据集的精心准备同样不可或缺。

**关键词:** 无人机;视觉;目标检测;跟踪

## 1 视觉目标检测与跟踪在无人机应用中的重要性

视觉目标检测与跟踪在无人机应用中占据着举足轻重的地位,这不仅是因为它为无人机赋予了“看”和“追踪”的能力,更是因为它极大地扩展了无人机的应用场景和实用价值。首先,视觉目标检测使得无人机能够在复杂的环境中识别出特定的物体或目标。在搜救任务中,无人机可以通过视觉检测迅速定位到失踪人员;在环境监测领域,它可以识别出非法排放源;在农业领域,它可以帮助农民精准定位病虫害区域。而视觉目标跟踪则更进一步,它不仅能让无人机看到目标,还能持续锁定并跟随目标。在军事侦察中,这意味着无人机能够紧跟敌方车辆或人员,实时传递情报;在物流配送中,无人机能够跟随货物车辆,确保货物安全送达;在野生动物保护领域,它可以帮助研究人员跟踪动物迁徙的轨迹<sup>[1]</sup>。随着深度学习技术的不断发展,视觉目标检测与跟踪的准确性和实时性也在不断提高,这使得无人机在实际应用中更加可靠和高效。

## 2 无人机视觉目标检测基本原理

无人机视觉目标检测的基本原理主要依赖于计算机视觉和深度学习技术。第一,无人机通过其搭载的摄像头捕获环境的实时图像或视频数据。这些数据是目标检测的原始输入。第二,图像经过预处理步骤,包括缩放、裁剪、色彩空间转换等,以提高后续处理的效率和准确性。第三,预处理后的图像被送入一个预先训练好的深度学习模型,通常是卷积神经网络(CNN)。这个网络通过学习大量标注过的图像数据,自动提取图像中的有用特征,如边缘、纹理、形状等。这些特征对于区分不同的物体和目标至关重要。第四,在特征提取的基础上,网络中的分类器会对这些特征进行进一步分析,并判断图像中是否存在特定的目标物体。分类器会根据学习到的知识,为每个潜在的目标生成一个置信度分

数,该分数表示该位置存在目标的可能性。第五,通过设定合适的阈值,可以筛选出置信度高的目标,并进行位置标注和可视化展示。这样,无人机就能够实时地检测出图像中的目标物体,并进行后续的处理和分析。无人机视觉目标检测的基本原理就是这样,它结合了计算机视觉和深度学习的先进技术,为无人机提供了强大的目标识别和定位能力。

## 3 基于深度学习的目标检测方法设计

### 3.1 深度学习在无人机视觉目标检测中的优势

深度学习在无人机视觉目标检测中展现出显著的优势,这些优势主要体现在以下几个方面:(1)高准确性:深度学习模型,特别是卷积神经网络(CNN),通过在大规模数据集上进行训练,能够自动学习和提取图像中的复杂特征,从而在目标检测任务中达到极高的准确性。这种准确性使得无人机能够在各种复杂环境中准确识别目标。(2)自动特征学习:深度学习模型能够自动从原始图像数据中学习和提取特征,无需手动设计特征工程。这不仅减轻工程师的负担,还提高了目标检测的效率和准确性。(3)处理大规模数据:随着无人机应用范围的扩大,视觉数据量也呈现爆发式增长。深度学习模型能够有效处理这些大规模数据,从而确保无人机能够在海量数据中进行目标检测。(4)适应性强:深度学习模型具有较强的通用性和泛化能力,可以适应不同的环境和任务。这使得无人机能够在各种复杂场景下进行有效的目标检测<sup>[2]</sup>。(5)实时性:随着计算机算力的发展,深度学习的目标检测算法进一步提升了运算速度,满足了无人机视觉目标检测的实时性要求。在无人机进行实时监测和追踪时,这一优势尤为关键。

### 3.2 目标检测深度学习网络架构设计

基于深度学习的目标检测方法设计在近年来取得了显著的进步,其核心在于构建高效且精准的目标检测网

络架构。这类网络架构通常采用卷积神经网络（CNN）作为基础，通过多层卷积操作来提取图像中的特征信息。在设计目标检测网络架构时，数据是至关重要的因素。首先，需要收集并标注大量的目标检测数据集，这些数据集应包含各种复杂场景下的目标，并标注出目标的位置和类别。这些数据将用于训练深度学习模型，使模型能够学习到目标的有效特征表示。在网络架构的设计方面，可以采用两阶段或单阶段的目标检测算法。两阶段算法如Faster R-CNN，首先生成候选区域，再对候选区域进行分类和边界框回归；而单阶段算法如YOLO和SSD，则直接通过单次前向传播预测目标的位置和类别。这两种算法各有优缺点，可以根据实际需求进行选择。为了提高目标检测的准确性和效率，还可以在网络架构中引入注意力机制、多尺度特征融合等技巧。注意力机制可以帮助模型关注图像中的关键区域，提高目标检测的准确性；而多尺度特征融合则可以利用不同尺度的特征信息，提高模型对不同大小目标的检测能力。

### 3.3 数据集获取和预处理

#### 3.3.1 数据集获取

（1）常用数据集：在目标检测领域，常用数据集如VOC和COCO。VOC数据集包含约10,000张带有边界框的图片，涵盖车辆、家具、动物、人等四大类，共计20个小类。而COCO数据集是一个大型、丰富的物体检测、分割数据集，包含超过33万张图片和超过150万个目标标签。（2）自定义数据集：针对特定任务或场景，可以构建自定义数据集。这通常涉及收集并标注与目标相关的图像数据，确保数据的多样性和代表性。

#### 3.3.2 数据预处理

（1）数据清洗：在获取到原始数据集后，首先需要进行数据清洗，去除重复、模糊、低质量或标注错误的图像。这可以通过自动化工具或人工审查来完成。（2）数据增强：为了增加模型的泛化能力和鲁棒性，可以对数据集进行数据增强。这包括但不限于随机旋转、翻转、缩放、裁剪、添加噪声等操作。这些变换可以增加模型的训练样本数量，同时模拟真实世界中可能出现的各种变化。（3）归一化：将数据集中的图像归一化到相同的尺度，以便后续处理和分析。归一化可以减少计算量，提高模型的训练速度和准确性<sup>[3]</sup>。（4）标签处理：将图像的标注信息（如边界框坐标、目标类别等）转换为适用于深度学习模型的格式。这通常涉及坐标转换、类别编码等操作。在基于深度学习的目标检测方法设计中，数据集获取和预处理是不可或缺的一步。通过选择合适的数据集、进行数据清洗和增强、以及标签处理，

可以为模型的训练和测试提供高质量的数据支持，从而提高模型的准确性和泛化能力。

## 4 无人机视觉目标跟踪技术研究

### 4.1 无人机目标跟踪的重要性

无人机目标跟踪的重要性不仅体现在提升无人机智能化和自主化水平上，还在实际应用中发挥着至关重要的作用。首先，无人机目标跟踪技术对于实现无人机的智能化和自主化至关重要。通过视觉感知和图像处理技术，无人机能够实时获取环境信息，识别并跟踪目标，自主进行飞行控制和决策。这不仅减少人为操作的依赖，还提高无人机的反应速度和任务执行效率。其次，无人机目标跟踪技术在实际应用中具有广泛的应用价值。在军事领域，无人机可以实时跟踪敌方目标，为导弹打击提供精确情报支持，提高作战效果。在搜索与救援任务中，无人机能够迅速定位被困人员或重要物资，为救援行动提供关键信息，增加救援成功率。在公共安全和监控领域，无人机可用于跟踪和识别可疑目标，及时发现潜在的安全隐患，提高公共安全水平。随着技术的不断进步，无人机目标跟踪技术还将有更广泛的应用前景。例如，在智能交通系统中，无人机可用于实时追踪交通流量，优化交通管理；在环境保护领域，无人机可用于跟踪和监测野生动物的活动轨迹，为生态保护提供数据支持。

### 4.2 目标跟踪算法综述

目标跟踪算法大致可分为生成式模型和判别式模型两类。生成式模型主要关注目标本身的特征表示，通过构建目标模板并在后续帧中搜索最相似的区域来实现跟踪。而判别式模型则将目标跟踪看作是一个二分类问题，通过训练一个分类器来区分目标和背景，从而实现目标跟踪。在无人机视觉目标跟踪中，常用的目标跟踪算法包括KCF（Kernelized Correlation Filters）算法、TLD（Tracking-Learning-Detection）算法和DeepSORT（Deep Simple Online and Realtime Tracking）算法等。其中，KCF算法以其高效性和鲁棒性广泛应用于无人机目标跟踪中；TLD算法结合了目标跟踪和学习的思想，能够在目标发生形变、遮挡等情况下实现较好的跟踪效果；而DeepSORT算法则利用深度学习技术，通过特征提取和数据关联，实现对多目标的长时间、准确跟踪。无人机视觉目标跟踪技术的研究需要大量的标注数据集进行算法的训练和评估。目前，已有多个公开可用的无人机视觉目标跟踪数据集，如VisDrone、UAVDT等。这些数据集包含了丰富的无人机拍摄场景和多样化的目标类型，为无人机视觉目标跟踪算法的研究提供了有力支持。

### 4.3 基于深度学习的目标跟踪方法设计

在无人机视觉目标跟踪技术的研究中,基于深度学习的目标跟踪方法已经展现出强大的潜力和优势。设计基于深度学习的无人机视觉目标跟踪方法,通常涉及到深度神经网络的构建和优化。一种常见的做法是利用卷积神经网络(CNN)来提取图像中的目标特征,并结合循环神经网络(RNN)或长短时记忆网络(LSTM)来捕获目标在连续帧间的运动信息。为了训练这样的模型,需要大量的标注数据,包括无人机拍摄的视频序列和对应的目标位置信息。这些数据可以通过专业的标注工具进行人工标注,或者利用自动化标注方法进行预处理。一些公开的无人机视觉目标跟踪数据集,如VisDrone、UAV123等,为这类方法的研究提供宝贵的数据资源。基于深度学习的目标跟踪方法不仅可以在复杂环境中准确跟踪目标,还可以通过在线学习的方式不断适应目标外观和背景的变化<sup>[4]</sup>。此外,这类方法还具有较好的实时性和鲁棒性,能够满足无人机在实际应用中对于目标跟踪的需求。

### 4.4 目标跟踪实验设计和数据集准备

在无人机视觉目标跟踪技术的研究中,目标跟踪实验设计和数据集准备是至关重要的一环。一个合理的实验设计能够确保研究结果的有效性和可靠性,而精心准备的数据集则是模型训练和评估的基础。要明确实验的目标和假设,例如验证某种目标跟踪算法在无人机场景下的性能。选择合适的无人机平台和视觉传感器,根据实验目标设定相应的飞行参数和拍摄条件。设计多样化的实验场景,包括不同的光照、天气和背景条件,以模拟真实世界中的复杂环境。在实验过程中,要确保数据的准确性和一致性,通过多次重复实验以获得可靠的结果。针对无人机视觉目标跟踪的特点选择合适的数

集。现有的无人机视觉数据集如VisDrone、UAVDT等,提供了大量包含标注目标信息的无人机拍摄视频,可以作为实验的基础数据。在准备数据集时,还需要对原始数据进行预处理,包括去噪、增强和归一化等操作,以提高数据的质量和模型的训练效果。为了更好地评估目标跟踪算法的性能,还需要准备合适的评估指标和测试数据。常用的评估指标包括精度、成功率和速度等,测试数据则应包括不同场景和条件下的视频序列,以全面评估算法的性能。

### 结束语

随着深度学习技术的不断进步,无人机视觉目标检测与跟踪领域取得显著的研究成果。这些技术为无人机在各种复杂环境中准确识别与跟踪目标提供强有力的支持,推动无人机在军事侦察、智能监控、物流配送等多个领域的应用。未来,期待更多创新的研究能够推动无人机视觉目标检测与跟踪技术的进一步发展,为无人机的智能化和自主化贡献更多力量。

### 参考文献

- [1]鞠默然,罗海波,王仲博,等.改进的YOLOV3算法及其在小目标检测中的应用[J].光学学报,2019,39(7):0715004.
- [2]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):95-104.
- [3]刘大伟,韩玲,韩晓勇.基于深度学习的高分辨率遥感影像分类研究[J].光学学报,2016,36(4):298-30.
- [4]蒲良,张学军.基于深度学习的无人机视觉目标检测与跟踪[J].北京航空航天大学学报,2022,48(5):872-880. DOI:10.13700/j.bh.1001-5965.2020.0664.