**开题检查记录表**

|  |  |  |  |
| --- | --- | --- | --- |
| 学生姓名 | 屈丹卉 | 学 号 | 20201105660 |
| 专 业 | 计算机科学与技术 | 开题日期 | 2023.10.26 |
| 指导教师 | 戚桂美 | 职 称 | 讲师 |
| 设计（论文）题目 | 基于YOLOv8的无人机图像目标检测算法研究 | | |
| 指导教师评语：  针对无人机图像中目标尺度大，现有检测算法精度低的问题，将最先进One-stage目标检测算法Yolov8应用于无人机图像，并进行算法的改进，提高检测精度。该研究可有效提升无人机图像在农业、畜牧业等领域的应用，提高工作效率，同意开题。  签字：戚桂美 | | | |
| 检查记录（答辩记录）： | | | |
| 是否通过开题： □通过   □不通过 | | | |
| 答辩组长签字： | | 答辩组成员： | |
| 答辩秘书签字： | |



**内蒙古师范大学计算机科学技术学院**

**毕业设计（论文）开题报告**

**题 目：基于YOLOv8的无人机图像目标检测算法研究**

**专 业 计算机科学与技术**

**姓 名 屈丹卉**

**学 号 20201105660**

**指导教师 戚桂美**

**日 期 2023.10.26**

**计算机科学技术学院制**

**说 明**

**一、开题报告主要内容**

1．课题来源及研究的目的和意义

1.1 课题来源：

无人机(Unmanned Aerial Vehicle, UAV)最早出现在20世纪的军事战场上，是通过无线遥感设备或车载计算机完全或间歇操作的非载人飞机。在军用领域，无人机能够在军事侦察、军事打击等任务中有效辅助军事人员进行战场态势感知[1]。 随着无人机小型化、智能化的发展，近年无人机在智慧城市、地理勘测、科技农业等多个民用领域也有着广泛的应用。在智慧城市领域，无人机可以监控城市交通动态，辅助道路摄像头及时将车流动态回传给交通管制，提升交通通行能力[2]。在地理勘测领域，无人机可以辅助专业人员提前进行地质分析和路线勘测。在科技农业领域，无人机可以用于农作物监控和害虫检测。总而言之，随着无人机技术的发展，无人机航拍图象的视觉分析具有广大的应用前景与研究意义。 而目标检测是计算机视觉的一项重要任务，其旨在从图像或视频中精准定位每个目标的位置并判断目标的所属类别。自2014年Girshick等[3]将卷积神经网络（Convolutional Neural Network，CNN）成功应用于目标检测领域，标志着基于深度学习的目标检测技术的发展到达了新的阶段。在此之后，卷积神经网络已成为多目标检测和目标跟踪的主流框架，研究者们在此基础上展开一系列的视觉研究任务。

1.2 研究的目的：

提高无人机图像目标检测的准确性和实时性：通过研究基于YOLOv8的无人机图像目标检测算法，可以提高无人机在复杂环境中对目标的识别和定位能力，从而提高无人机在实际应用中的效率和安全性；

降低无人机图像目标检测的计算复杂度：传统的目标检测算法通常具有较高的计算复杂度，限制了无人机在实时性要求较高的场景中的应用。基于YOLOv8的无人机图像目标检测算法可以有效降低计算复杂度，提高无人机的实时性能。

扩展无人机图像目标检测的应用范围：通过研究基于YOLOv8的无人机图像目标检测算法，可以使其适应更多类型的目标和场景，从而拓展无人机在农业、环保、安防等领域的应用。

优化无人机图像目标检测的模型结构：通过对YOLOv8模型进行改进和优化，可以提高无人机图像目标检测算法的性能，为后续相关研究提供参考。

1.3 研究的意义：

基于YOLOv8的无人机图像目标检测算法的研究可以为无人机技术提供更先进的技术支持，推动无人机技术在各个领域的应用和发展；而通过提高无人机图像目标检测的准确性和实时性，可以有效降低无人机在执行任务过程中发生意外的风险，还可以拓展无人机在各领域的应用，使无人机在农业、环保、安防等领域的应用更加广泛，为这些领域的发展提供有力支持；通过对YOLOv8模型的改进和优化，可以为后续相关研究提供有价值的参考，为相关研究提供参考，推动目标检测技术的进步。

1. 国内外在该方向的研究现状及分析

目标检测主要有两个发展阶段：基于传统方法的目标检测算法；基于深度学习 的目标检测算法。

2.1 基于传统方法的目标检测算法

早期传统的目标检测算法依赖于人工设定特征，所以图像的表达十分有限。其中，选择区域、提取特征、分类三个步骤是传统目标检测算法的要点问题。滑动窗口是最直接的区域选择方式，通过对图像上下、左右进行遍历来寻找目标位置，虽然操作简单，但是对于特征的提取分类计算过于复杂，效率较低。这些传统的识别方法无法解决复杂场景的检测问题。各种传统的目标检测方法都是采用了人为设计的特征，其性能完全取决于人工经验，很难充分发挥庞大数据的优点[4]，此外，传统的目标检测算法针对的目标类别有限，只在特定某几类目标中效果较好，普适性较差。并且在现实生活中，也存在应用场景受限，泛化性较差等问题。这也是为什么现在基于深度学习的算法受到广泛应用的原因。

2.2 基于深度学习的目标检测算法

深度学习发展迅速，基于深度学习的目标检测方法得到了广泛研究与应用，这种方法克服了传统人工特征提取方法在背景复杂和目标遮挡特点下的缺陷。目标检测轻量化模型在保证检测精度的情况下，有效提升了目标检测模型的检测效率。2012 年，Krizewski等人提出了AlexNet深度卷积神经网络，取得了有史以来最高的的目标分类准确度[5]。此后，许多学者开始研究基于卷积神经网络的深度学习算法，并将其应用到不同的目标检测方向。基于深度学习的目标检测算法分为一阶段和两阶段两大类[6]，一阶段检测算法基于回归方法，不需要事先选择候选域，使用卷积神经网络直接提取特征，对域进行分类、定位，具有较快的检测速度；二阶段检测算法先进行候选域的选择再进行特征的提取和目标种类的识别，速度较慢。SSD和YOLO系列是一阶段目标检测算法的代表之作。2015年，YOLO首次提出将目标检测问题转化为回归问题[7]，实现了45帧每秒的检测速度和实时图像处理，但目标定位不准确，检测质量较差。随后的SSD算法融合了二阶段算法Faster R-CNN锚框和一阶段算法YOLO回归的优势[8]，实现了与 YOLO相似的检测速度和比Faster R-CNN更好的检测精度。2017年，YOLOv2 算法也加入了锚框的思想[9]，并构建了骨干网络 Darknet19，检测精度与SSD相当，却达到了更快的速度。2018年，YOLOv3算法借鉴ResNet，提出了 Darknet-53骨干网络[10]，融合三种不同尺度的特征信息，检测精度明显高于YOLOv2，且速度没有降低。2020年提出的 YOLOv4是基于YOLOv3的更高性能网络模型[11]，其结合加权残差连接、跨小批量归一化、自对抗训练、Mish 激活函数、CIOU损失等机制，有效提高了检测精度和速度，实现更高效、鲁棒的目标检测网络。同年6月， Ultralytics 提出了一个具有较快的检测速度以及良好的检测精度的版本YOLOv5，得到了很多学者的广泛应用，其模型体积极小且精度损失不多，具有更高的工程实现意义。上述基于深度学习的目标检测算法不论在检测速度还是精度方面都胜过传统算法，对于不同背景的适应性和泛化能力更好，因此成为了目前目标检测领域的研究热点。

1. 主要研究内容

3.1 数据集的下载及处理

本次实验所使用的数据集是VisDrone2019数据集，VisDrone2019数据集是由天津大学机器学习和数据挖掘实验室AISKYEYE团队收集的一个大规模基准数据集。这个数据集包括了288个视频片段，这些片段由各种无人机摄像头捕获，总共包含了261908帧和10209幅静态图像。它的覆盖面非常广泛，包括了位置（来自中国相隔数千公里的14个不同城市）、环境（城市和农村）、物体（行人、车辆、自行车等）和密度（稀疏和拥挤）等多个方面。因此，VisDrone2019数据集可以广泛应用于各种重要的计算机视觉任务，如目标检测和追踪等。

3.2 深度学习——卷积神经网络（CNN）

深度学习是一类[模式分析](https://baike.baidu.com/item/%E6%A8%A1%E5%BC%8F%E5%88%86%E6%9E%90/12598452?fromModule=lemma_inlink" \t "https://baike.baidu.com/item/%E6%B7%B1%E5%BA%A6%E5%AD%A6%E4%B9%A0/_blank)方法的统称，就具体研究内容而言，主要涉及三类方法：基于卷积运算的神经网络系统，即[卷积神经网络](https://baike.baidu.com/item/%E5%8D%B7%E7%A7%AF%E7%A5%9E%E7%BB%8F%E7%BD%91%E7%BB%9C/17541100?fromModule=lemma_inlink" \t "https://baike.baidu.com/item/%E6%B7%B1%E5%BA%A6%E5%AD%A6%E4%B9%A0/_blank)（CNN）。基于多层神经元的自编码[神经网络](https://baike.baidu.com/item/%E7%A5%9E%E7%BB%8F%E7%BD%91%E7%BB%9C/16600562?fromModule=lemma_inlink" \t "https://baike.baidu.com/item/%E6%B7%B1%E5%BA%A6%E5%AD%A6%E4%B9%A0/_blank)，包括自编码（Auto encoder）以及近年来受到广泛关注的[稀疏编码](https://baike.baidu.com/item/%E7%A8%80%E7%96%8F%E7%BC%96%E7%A0%81/10289670?fromModule=lemma_inlink" \t "https://baike.baidu.com/item/%E6%B7%B1%E5%BA%A6%E5%AD%A6%E4%B9%A0/_blank)两类（Sparse Coding）。以多层自编码神经网络的方式进行预训练，进而结合鉴别信息进一步优化神经网络权值的[深度置信网络](https://baike.baidu.com/item/%E6%B7%B1%E5%BA%A6%E7%BD%AE%E4%BF%A1%E7%BD%91%E7%BB%9C/52718542?fromModule=lemma_inlink" \t "https://baike.baidu.com/item/%E6%B7%B1%E5%BA%A6%E5%AD%A6%E4%B9%A0/_blank)（DBN）。

卷积神经网络(CNN)是一种具有前反馈的网络。与传统的全连接网络结构不同，每一层都进行特征的提取，减少了参数量，因此网络的计算量减少。网络的复杂度降低，防止模型训练过程中出现过拟合的现象，也解决了传统的目标检测特征提取比较复杂的问题。卷积神经网络的特征是三维的：长、宽和高，这些特征在图像的处理中占据很大的优势，尤其对图像的纹理、颜色、形状和空间的提取表现更好的鲁棒性[12]。 卷积神经网络是由原本的神经网络上发展而来，传统的神经网络由输入层、隐藏层和全连接层三部分组成。在输入层中具有大量的神经元(Neuron)，每一个神经元都会接受许多的输入的非线性数据。在输入层和输出层两者之间有许多的神经网络，相互链接从而构成了隐藏层。节点的神经元数目越多表示神经网络的非线性就越显著，隐藏层主要用于存储处理特征和存储信息。输出层中神经元经过传输分析后，形成输出结果。

卷积神经网络长期以来是图像识别领域的核心算法之一，并在学习数据充足时有稳定的表现。对于一般的大规模图像分类问题，卷积神经网络可用于构建阶层分类器，也可以在精细分类识别中用于提取图像的判别特征以供其它分类器进行学习 。对于后者，特征提取可以人为地将图像的不同部分分别输入卷积神经网络，也

可以由卷积神经网络通过非监督学习自行提取。在深度学习方面，YOLOv8算法的核心是卷积神经网络（CNN），通过训练大量的标注数据来学习目标的特征表示，从而实现目标检测任务。

3.3 机器视觉和目标检测

机器视觉是一项涉及光学、机械、计算机、模式识别、图像处理、人工智能和信号处理等多个领域的综合性研究领域，其主要目的是让机器模拟并超越人类的视觉功能，以实现各种测量和判断。在机器视觉的应用中，目标检测是一项至关重要的任务，早期的目标识别方法包括Blob分析法（BlobAnalysis）、模板匹配法以及深度学习法等。其中，Blob分析法主要针对图像中的相同像素的连通域进行分析，其过程就是将图像进行二值化，分割得到前景和背景，然后进行连通区域检测，从而得到Blob块的过程。然而随着深度学习技术的发展，深度学习模型已经逐渐取代了传统的机器视觉方法，成为了目标检测领域的主流算法。相较于目标定位，目标检测任务更具挑战性，因为其图像中出现的目标种类和数目都不定。目前计算机视觉的三大核心任务是：图像分类、目标检测和图像分割，其中目标检测不仅要识别出图像中的物体，还需要给出这些物体的具体位置信息。例如，在一张图片中，目标检测不仅要识别出图片中的阿猫阿狗，还得给出阿猫阿狗的具体位置。

3.4 YOLOv8

YOLOv8算法是一种基于深度学习的目标检测算法，是ultralytics公司在2023年1月10号开源的YOLOv5的下一个重大更新版本，它建立在以前YOLO版本的成功基础上，并引入了新的功能和改进，以进一步提升性能和灵活性。YOLOv8是目前YOLO系列算法中最新推出的检测算法，可以完成检测、分类、分割任务。从整体设计到模型结构，YOLOv8算法的核心特性和改动可以归结为如下：首先，它采用了一个新的骨干网络；其次，它引入了一个全新的Ancher-Free检测头；最后，它还采用了一个全新的损失函数。这些创新使得YOLOv8可以在从CPU到GPU的各种硬件平台上运行。在训练策略方面，YOLOv8进行了详细的数据增强处理，包括Mosaic、Mixup、CutMix等，以提高模型的泛化能力。而在推理过程中，YOLOv8则采用了多尺度测试和自适应阈值的方法来提高检测精度。

1. 研究方案

数据集准备：收集无人机拍摄的图像数据，并进行标注。VisDrone2019数据集是一个用于视觉无人机目标跟踪的大规模数据集，首先需要从网站或GitHub仓库下载VisDrone2019数据集。数据集通常包括训练、验证和测试子集。将下载的数据集文件解压缩到一个目录中。确保有足够的磁盘空间来存储这些文件。在开始训练之前，需要对数据进行预处理。包括图像缩放、裁剪、旋转等操作。此外，还需要为每个视频创建一个标签文件，其中包含物体类别、边界框坐标等信息。根据需求，将数据集划分为训练集和验证集。通常可以使用80%的数据作为训练集，20%的数据作为验证集。为了方便地加载数据，需要创建一个数据加载器。可以通过使用深度学习框架（如PyTorch或TensorFlow）提供的数据加载器函数来实现。根据模型和硬件配置，设置合适的训练参数，如学习率、批次大小、迭代次数等。

模型训练：使用YOLOv8算法对标注好的数据集进行训练，得到目标检测模型。使用创建的数据加载器和训练参数，训练模型。在训练过程中，可以监控模型的性能指标，如准确率、召回率等，并根据需要调整训练参数。

模型优化：在训练完成后，使用验证集评估模型的性能。如果性能不佳，可以尝试调整模型结构、训练参数或数据预处理方法。对训练得到的模型进行优化，包括调整超参数、增加正则化等方法，提高模型的准确率和鲁棒性。在训练过程中，可以定期保存模型权重。在完成训练后，将最佳模型权重加载到新的实例中进行推理或进一步优化。

模型测试：使用测试集对优化后的模型接着进行测试，评估其性能指标，如准确率、召回率、F1值等。

5．进度安排，预期达到的目标

|  |  |  |
| --- | --- | --- |
| 时间 | 进度安排 | 预期目标 |
| 2023.10.10 | 确定论文选题 | 确定论文选题 |
| 2023.10.11-2023.10.15 | 查阅参考文献，进行背景调查，目的分析等 | 目的背景完成 |
| 2023.10.16-2023.10.29 | 纂写开题报告，修改开题报告，完成开题答辩 | 开题答辩 |
| 2023.10.30-2023.12.25 | 整理和总结论文资料，确定论文大纲 | 论文大纲 |
| 2023.12.26-2024.01.25 | 围绕论文论点查阅相关数据和写作资料 | 资料查找 |
| 2024.01.26-2023.02.20 | 撰写论文初稿，与指导教师沟通并及时进行修改 | 论文初稿 |
| 2024.02.21-2023.04.20 | 完成论文中期检查，并针对检查中存在问题进行修改 | 论文中期检查 |
| 2023.04.21-2023.06.20 | 完善论文撰写，提交论文终稿 | 论文终稿 |

1. 课题已具备和所需的条件、经费

外部条件：网络在线资源、图书馆资料

硬件条件：笔记本电脑

软件条件：PyCharm、Window 10

1. 研究过程中可能遇到的困难和问题，解决的措施

1）问题： 数据集标注问题，如果训练数据集中存在错误标注或者标签文件内容有误，可能会导致训练效果不佳。

解决方法：仔细检查和清理数据集，确保所有的图像和标签文件都符合要求。

2）问题： 模型训练不稳定，训练过程中可能会出现一些报错信息，例如"no labels found in xxx.cache"等。

解决方法：首先需要确认图像和标签文件夹的内容是否正确，然后检查训练过程中使用到的文件，如dataset.py等是否存在问题。

3）问题： 计算资源消耗大，由于YOLOv8模型较为复杂，因此在训练和测试过程中可能需要大量的计算资源。

解决方法：可以考虑使用一些优化技术，如模型压缩、量化等方法来减少模型的计算需求。

8．主要参考文献

[1] 惠国保. 军事小目标智能感知认知技术研究[C] // 第九届中国指挥控制大会, 2021: 101-106.

[2] 葛园园, 许有疆, 赵帅, 等. 自动驾驶场景下小且密集的交通标志检测[J]. 智能系 统学报, 2018, 13(03): 366-372.

[3] 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

[4]李柯泉 , 陈 燕 , 刘佳晨 , 等 . 基于深度学习的目标检测算法综述 [J]. 计算机工 程,2022,48(07):1-12.

[5]Krizhevsky A, Sutskever I, Hinton G. ImageNet Classification with Deep Convolutional Neural Networks[J]. Advances in neural information processing systems, 2017,60(06):84-90.

[6]裴伟,许晏铭,朱永英,等.改进的 SSD 航拍目标检测方法[J].软件学报, 2019, 030(03):738- 758.

[7]Redmon J, Divvala S, Girshick R, et al. You Only Look Once: Unified, Real-Time Object Detection[C]// Computer Vision & Pattern Recognition. IEEE, 2016:779-788.

[8]Liu Wei, Anguelov D, Erhan D, et al. SSD: Single Shot MultiBox Detector[C]// European Conference on Computer Vision, 2016:21-37.

[9]Redmon J, Farhadi A. YOLO9000: Better, Faster, Stronger[C]// IEEE conference on computer vision and pattern recognition, 2017: 6517-6525.

[10]Redmon J, Farhadi A. YOLOv3: An Incremental Improvement[J]. arXiv preprint arXiv, 2018, 1804(04):1-6.

[11]Bochkovskiy A, Wang Chienyao, Liao Hongyuan. YOLOv4: Optimal Speed and Accuracy of Object Detection[J]. arXiv preprint arXiv, 2020,2004(04):1-17.

[12]胡广. 基于深度学习语义分割的断层识别应用研究[D].中国地质大学,2021。