

基于 Soft-NMS 改进 YOLOv8 算法的奶牛识别

刘潇阳 韩泽宇 闫艺菲 杨济嘉 张佳欣

中国农业大学 山东烟台 100083

摘要: 目前, 人工只能和物联网技术正广泛应用在农业领域, 中国智慧农业正稳步发展。计算机视觉技术的引入, 使农场等动物识别正逐渐摆脱传统监控系统对传感器的依赖。而密集场景下的动物识别具有空间小、数量多、个体体积大、遮挡和粘连问题严重等特点, 使得在动物群体中准确地识别出每一只禽畜成为一个难点。本文使用 YOLOv8 模型实现奶牛识别, 并使用 Soft-NMS 算法替代传统非极大抑制算法。通过对比, 解决了大体积动物存在严重遮挡情况下的个体识别, 为下一阶段的动物行为识别作好铺垫。

关键词: 目标检测; 非极大抑制算法; Yolo。

1. 绪论

1.1 背景和意义

随着农业物联网技术的普及, 加速了我国农业和畜牧业和人工智能技术的结合。使用计算机系统协助农民和牧场主对动植物完成实时检测和异常预警可以极大程度解放劳动力, 并可以更快更好地完成传统监管任务。尤其对畜牧业中如牛、羊等动物的行为和身体状况进行实时监控和预警可以实现更快更优质的肉类等其他相关产品生产。

长期以来, 我国畜牧业中对动物的监管依赖于传感器, 常见方法是将集合温度、三轴加速度和角速度等传感器佩戴在动物颈部或腿部。通过接收传感器数据来判断当前动物身体状态和行为。尤其是为了获取更细致的数据, 部分研究者将传感器注射如动物体内, 这种方法一定程度上实现了对每一只动物的管控。但是, 动物佩戴或注射传感器本身是对动物行为的一种限制, 尤其注射传感器方式经常引起动物局部发炎但无法处理。因此, 需要寻找更优的方法完成这类任务。

近年来计算机视觉技术的发展突飞猛进, 短短十年中, 研究者们已经从 ImageNet 上只有百分之六十准确率, 发展到如今在目标检测、行为识别、物体分割、图片生成等任务上均实现极高准确率。得益于此, 我国智慧农业正向着人工智能和物联网技术融合的方向发展。在农牧场中安装集成了环境参数传感器, 并搭载了深度学习模型和处理器的高精度摄像头可以完美替代传感器方法和解决传感器方法的一系列问题。本文基于对计算机视觉模型 YOLOv8 的

改进, 使之更适合于小空间, 大数量, 大体积, 高遮挡性的奶牛识别任务。

1.2 文献综述

深度学习方法的目标检测可以在分类模型训练的前几层自动完成特征提取任务。目前常见的目标检测任务模型分为一阶段 (以 YOLO、SSD 为主) 和两阶段 (以 Fast-RCNN 为主) 两种方法。郭阳阳将空间注意力机制和 YOLOv4 模型融合, 提出了 YOLOv4-SAM 深度学习网络, 在提取奶牛多尺度特征基础上突出生物特征等相关特征, 提高了生物视觉的表征能力, 实现了对奶牛个体的高精度检测^[1]。冯策在研究猪的进食检测中, 提出了通过猪头部和食槽的交互来定义进食行为, 并采用 VGG16 建立猪头部识别模型, 实现了结构化场景下猪头部的精准识别^[2]。WANG R 等人使用 GhostConv 替换 YOLOv5 的 Conv 并且将轻量化注意机制融入 YOLOv5 模型, 使用 YOLOv5s 预训练, 实现了对奶牛骑乘行为的精准识别^[3]。

2. 相关技术综述

2.1 基于机器视觉的牲畜识别框架

基于机器视觉的牲畜识别技术在各种研究中所使用数据采集设备、算法、针对目标等略有不同, 但其整体的研究框架基本类似。。

首先是数据采集和数据标注。通过采集不同场景下各种牲畜的视频或图片数据, 使用 Labelimg 或 cvat 等工具对其进行 yolo 格式的标注。再划分训练集和验证集, 形成一个可训练的数据集。

然后是对模型的导入和修改。随着目标检测任务愈发困难，传统的卷积神经网络等模型无法实现较好的检测效果。因此多数研究者们正在使用 YOLO 系列模型完成目标检测任务。针对研究者个人任务的特殊性，对原生 YOLO 模型进行修改，使之更适配特定任务。

最后，模型训练和测试完成后，将模型装载到摄像机上，完成后续任务。

2.2 相关技术概述

2.2.1 Yolo 系列模型及 YOLOv8

YOLO (You Only Look Once) 系列算法是为处理实时目标检测任务而生,至 2023 年 4 月,Ultralytics 发布 YOLOv8 模型,代表着 YOLO 系列模型的一大突破。YOLOv8 模型在各种数据集、应用领域上表现良好,只需稍加调整,即可应对特定任务需求。如图 1 所示,为 YOLOv8 结构图。可以看出,在进行不同任务时,YOLOv8 将此前的耦合头改为解耦头。针对不同任务,在最后一步前,路过不同的卷积模块。

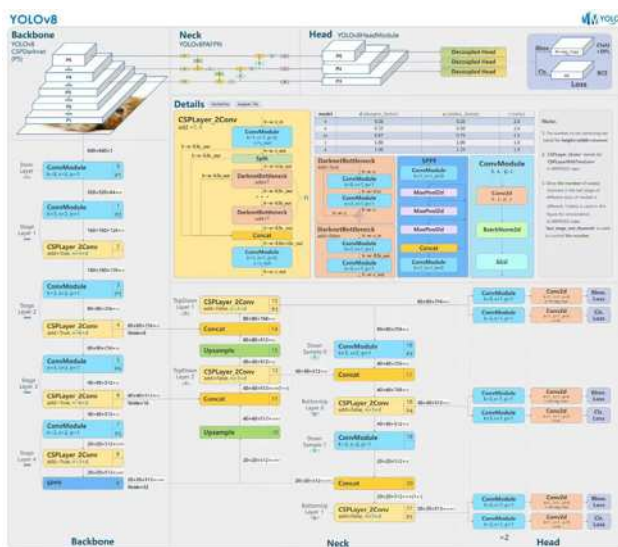


图 1 YOLOv8 模型结构图

2.2.2 IOU

IoU(Intersection over Union),即交并比。对于两个重叠的框选区域而言, IoU 等于两个区域交叉重叠部分除以两个区域的并集得到的结果,计算方法如图二所示。一般来说,这个 score > 0.5 就可以被认为是一个不错的结果了。

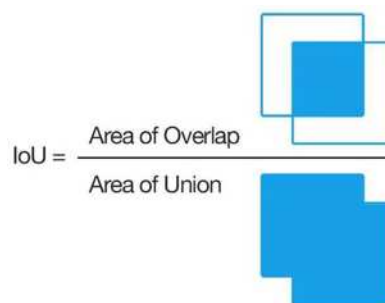


图 2 IOU (交并比) 图示

2.2.3 NMS 算法

在目标检测任务中,物体遮挡和粘连情况下很容易造成漏检。如图三,定位某个物体时算法找出很多方框,此时需要保留一个最好的框。非极大值抑制 (NMS) 即是在多个对目标的标注框中找到最优的一个保留,并计算交并比。

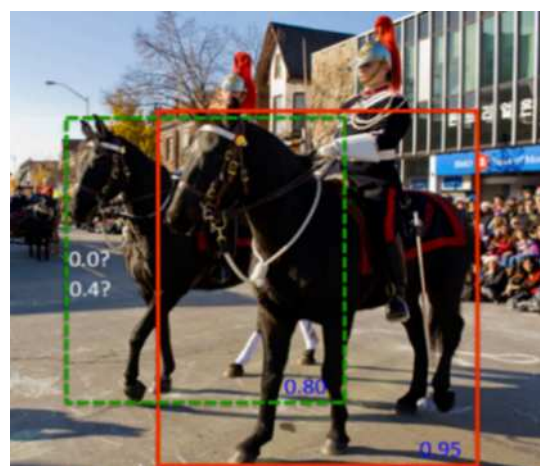


图 3 传统 NMS 算法举例

3. 数据集简介

本文使用的奶牛数据集是各种场景下收集得到的。该数据集仅手动标注了奶牛一个类别,标注文件保存了对应图片中的类别数量(标签为 0,类别为奶牛)、这个类别中每个个体所在标注框的左上和右下对角线点的坐标,通过这两个点的坐标即可确定一个矩形框。本文实验的奶牛检测识别数据集包含训练集 3250 张图片,验证集 860 张图片,选取部分数据部分样本数据集和标注情况如下图 3 所示,本文对比实验也将基于示例进行展开。YOLOv8 可以接受任何尺寸的图片数据并在模型内将接收到的图片数据调整为 640x640 的大小,并保持原有的宽高比例,而这一工作,在使用 YOLOv5 时,我们需要手动调整图片再送入模型。此外,

为了增强模型的泛化能力和鲁棒性,本文使用了旋转、缩放、和颜色变换等数据增强技术,以扩充数据集并减少过拟合风险。



图 4 数据集部分图片展示

4. Soft-NMS 算法及其对 YOLOv8 的改进

4.1 Soft-NMS 算法

继续看图三,传统的 NMS 算法会将绿色框当成红色框目标的一个较差的目标框进行删除操作。这导致了严重的漏检。

为了改进贪心的 NMS 算法,本文引入了 Soft-NMS 算法。在该算法中,基于重叠部分的大小为相邻检测框设置一个衰减函数而非彻底将其分数置为零。其思路是:不删除所有 IOU 大于阈值的框,而是降低其置信度。本文将此算法引入进行实验。

4.2 实验方法

本文实验环境是 Pytorch 2.0.0, Python 3.8(ubuntu20.04), Cuda 11.8, GPU 使用 RTX 4090(24GB)。

本文实现方法主要修改了 Ultralytics/Utils/ops.py 根据 Improving Object Detection With One Line of Code 作者的开源代码,将 Soft-NMS 代码作为一个函数加入到 ops.py 文件中添加 soft_nms() 函数后。在 ops.py 文件中找到的 non_max_suppression() 函数中的下面代码。

```
i = torchvision.ops.nms(bboxes, scores, iou_thres) # NMS
```

修改为刚刚添加进来函数。修改后的代码如下:

```
i = soft_nms(bboxes, scores, iou_thres) # NMS
```

此时我们已实现在 YOLOv8 模型中使用 Soft-NMS 算法代替原本的 NMS 算法。最后,导入 Ultralytics 官方给出的 yolov8n.pt 预训练文件即可。

4.3 实验结果和分析

实验结果分析中,本文使用精度和召回率等指标来评估模型的性能,通过损失曲线和 PR 曲线来分析训练过程。本文使用 YOLOv8 模型对数据集总计训练了 100 个 epochs。在训练过程中,记录了模型在训练集和验证集上的损失曲线。如下图 5,模型的训练损失和验证损失均在逐步降

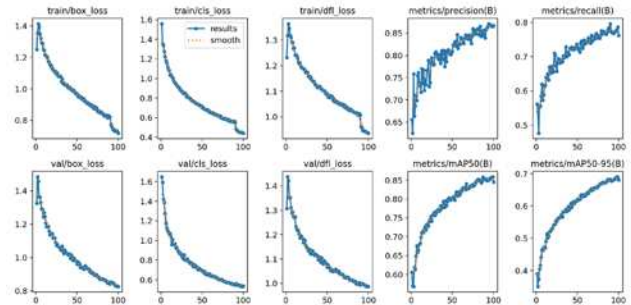


图 5 训练过程中训练集和验证集上的损失和精度

直观而言,本文使用训练好的模型在上图三的验证集上进行预测。如下图 6,可以清楚看出,在奶牛密集存在遮挡的情况下,仍对遮挡目标进行了良好标注。整体看来, YOLOv8 模型在数据集上表现良好,具有较高的检测精度和鲁棒

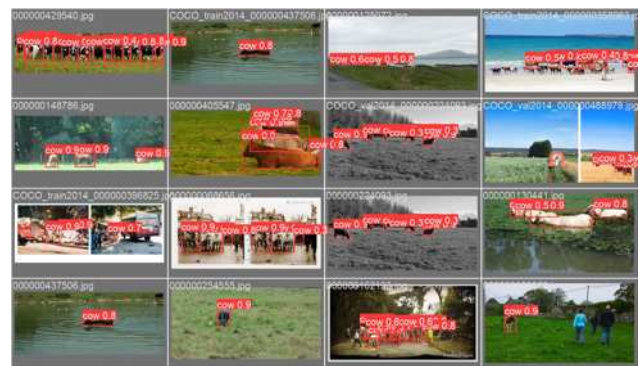


图 6 在图三打标签的验证集上的预测结果

5. 结论和展望

综上所述,将 Soft-NMS 算法引入 YOLOv8 模型后,在遮挡严重和牛类密集的图片的识别中有很大改进。而在部分图片中,对小目标和粘连目标的检测较差,且模型整体单张图片检测耗时为 0.17-0.22 秒,对目前大多数帧率摄像机而言,其检测速度较慢。这与 YOLOv8 模型较大有关。所以后续工作可以从加入轻量化注意力机制和将卷积模块替换为轻量化的卷积模块展开。

参考文献

- [1] 基于机器视觉的奶牛身体区域检测与典型行为分类方法 _ 郭阳阳 .pdf[J].
- [2] 基于多目标跟踪的群养猪进食行为识别方法研究 _ 冯策 .pdf[J].
- [3] WANG R, GAO R, LI Q, et al. A lightweight cow mounting behavior recognition system based on improved YOLOv5s[J].
- [4] BODLA N, SINGH B, CHELLAPPA R, et al. Soft-NMS — Improving Object Detection with One Line of Code[C/OL]//2017 IEEE International Conference on Computer Vision (ICCV), Venice. 2017. <http://dx.doi.org/10.1109/iccv.2017.593>. DOI:10.1109/iccv.2017.593.