

基于感受野动态注意力的苹果检测研究

万福涛^{1*} (通讯作者) 王洋 戚俊杰

西京学院计算机学院 陕西 西安 710018

摘要: 传统的物体检测方法在农业环境中, 尤其是在苹果检测应用中, 常因复杂背景、苹果间的遮挡以及苹果本身大小变化大等问题而面临识别精度不足和误检率高的困难。为了解决这些困难, 本文引入了一种基于感受野的动态注意力机制, RFACnv (Receptive-Field Attention Convolution), 并将其应用于 YOLOv8s 模型, 用于提高农业环境中的苹果检测精度。通过在自制的苹果数据集上测试, 集成了 RFACnv 的 YOLOv8 模型在多个性能指标上显著优于基线模型, 展示了其在农业中苹果检测的应用潜力。

关键词: 物体检测; 动态注意力机制; RFACnv; YOLOv8

1、引言

随着深度学习技术在图像处理领域的迅猛发展, 特别是在农业领域, 对象检测技术已经取得了巨大的进步 [1-2]。近年来, 随着空间注意力机制^[3]在深度学习领域引入, 它通过赋予神经网络关注图像中重要区域的能力, 显著提高了网络在各种任务上的识别性能。然而, 大多数卷积网络依赖于参数共享机制, 这意味着同一个卷积核被用于处理整个图像的所有位置。虽然这种设计大大减少了模型的参数数量, 提高了计算效率, 但它也限制了网络在处理具有丰富内部变异的图像时的灵活性和效果^[4]。

为了解决上述问题, 我们提出将 RFACnv 融入 YOLOv8s 模型中, 以期在苹果检测任务上实现更高的准确率和更好的模型泛化能力。本文的主要贡献包括: (1) 提出一种结合 RFACnv 和 YOLOv8s 的新框架, 用于改进苹果检测的精度; (2) 在自建的苹果数据集上进行实验, 验证所提方法的有效性; (3) 通过这项研究, 我们希望为农业图像处理领域提供一个高效、精确的检测工具, 以支持农业自动化和智能化的发展。

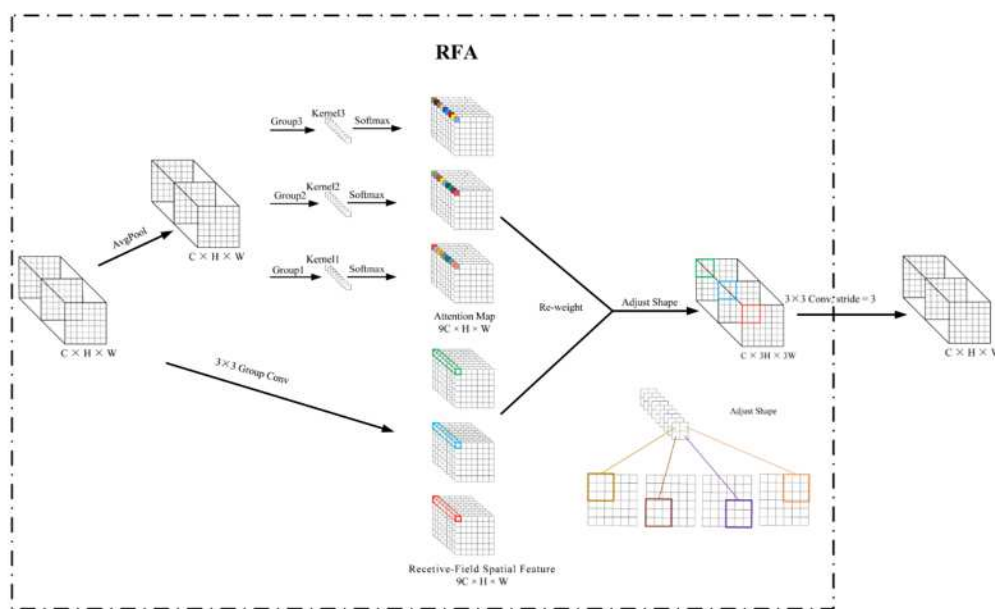


图 1 RFACnv 结构图

2 基于 YOLOv8 的改进算法

2.1 空间注意力机制模块

空间注意力机制的核心思想是使计算机视觉模型能够自动识别并集中处理图像中最重要的部分，从而忽略那些不太相关或干扰性的信息。通过利用空间注意力机制能显著增强模型在特征提取阶段的效率和效果。CBAM (Convolutional Block Attention Module) 通过一个模块化的设计，顺序集成通道注意力和空间注意力两个子模块，分别从通道和空间两个维度优化特征的表征力^[5]。SENet (Squeeze-and-Excitation Networks) 通过对卷积网络中的通道关系进行动态重新校准，来增强网络的表现^[6]。其核心是一个“squeeze-and-excitation”操作，这使得模型能够自动强化有用的特征并抑制无关的特征，从而在不同的视觉任务中实现更精确的特征表达。

2.2 RFACConv 模块

RFACConv 通过引入基于感受野的注意力机制来优化传统卷积操作。其核心思想是为每个感受野赋予一个独立的注意力权重，这些权重根据感受野中的内容动态调整，如图 1 所示。具体来说，RFACConv 包括以下几个关键步骤：(1)

输入特征图被分解为多个感受野块，每个块对应于原始特征图中的一个局部区域。(2) 对每个感受野块，一个小型的注意力网络被用来计算一个权重值，这个权重值反映了该块的重要性。

3 实验结果与分析

3.1 图像采集与预处理

本文中的数据集为本人自主拍摄而来，由 4200 张高分辨率图像组成。每张图像大约包含 8 个苹果，由两种不同类型的高清摄像机拍摄。图像有两种分辨率：4032*3024 和 1920*1080，并以 JPG 格式存储。图像大小和质量的多样性有助于形成适合各种分析目的的综合数据集。实施了严格的质量控制措施。质量不佳的图像和不包含苹果的图像被手动丢弃，以保持数据集的完整性和相关性。

3.2 实验结果及分析

3.2.1 消融实验

为了更好的验证改进后模型的有效性，在各实验参数相同的前提下，本文进行了注意力机制消融实验和对比实验。

表 1 注意力机制消融实验和对比实验

模型	mAP@0.5%	mAP@0.5-0.95%	P	R
YOLOv8s(baseline)	85.1	75.6	85.4	82.9
YOLOv8+CBAM	86.5	78.6	86.3	85.5
YOLOv8+SE	86.9	79.1	87.1	84.6
YOLOv8+RFACConv	88.2	82.4	88.5	86.8
SSD	55.6	46.8	48.5	49.6
Faster R-CNN	62.3	54.2	51.8	50.2
YOLOv3	65.6	60.2	59.4	56.7
YOLOv5s	82.1	70.5	76.9	74.8
YOLOv7	82.2	72.3	79.6	75.4
YOLOv8+RFACConv	88.2	82.4	88.5	86.8

如表所示，YOLOv8+CBAM 在 mAP@0.5 上比基线模型提高了 1.4%，在 mAP@0.5-0.95 上提高了 3%。YOLOv8+SE 在 mAP@0.5 提高了 1.8%，mAP@0.5-0.95 提高了 3.5%，YOLOv8+RFACConv 是使用了新提出的 RFACConv 注意力机制的 YOLOv8，展示了所有模型中最高的性能提升，mAP@0.5 和 mAP@0.5-0.95 分别比基线模型提高了 3.1% 和 6.8%。SSD 和 Faster R-CNN 是较早期的模型，在所有指标上性能较低。YOLOv3 是 YOLO 系列的早期版本，其性能虽然优于 SSD 和 Faster R-CNN，但与后续的 YOLO 版本相比还有提升空间。YOLOv5s 和 YOLOv7 的性能有了明显的提升，尤其

在 mAP@0.5 上分别达到了 82.1% 和 82.2%，显示了 YOLO 模型的迅速发展和性能提升。YOLOv8+RFACConv 在所有指标上均有最高分。mAP@0.5 达到了 88.2%，mAP@0.5-0.95 达到了 82.4%，这意味着该模型不仅能检测出大部分的对象，而且对对象边界的识别也非常准确。精确率和召回率分别达到了 88.5% 和 86.8%，意味着它在减少误检（高精确率）和减少漏检（高召回率）方面都取得了最优结果。

4 结束语

在本研究中，我们提出了一种结合了感受野动态注意力机制 (RFACConv) 的改进 YOLOv8s 模型，旨在提高农业

图像处理中的苹果检测精度。我们在自建的苹果数据集上进行了广泛的实验验证。这一结果不仅验证了 RFACConv 在处理复杂背景和遮挡条件下的优越性，也展示了其在实际农业生产环境中的潜在应用价值。未来，我们将继续优化 RFACConv 的结构和参数，探索其在其他农业监测任务中的应用，以支持农业自动化和智能化的持续进步。

参考文献

[1] 邓玲黎, 沈侃. 基于机器视觉的农业机械自动导航与作业优化 [J]. 农业技术与装备, 2023, (12): 43-45+48.

[2] Salazar-Gomez A, Darbyshire M, Gao J, et al. Towards practical object detection for weed spraying in precision agriculture[J]. arXiv preprint arXiv: 2109.11048, 2021.

[3] 程德强, 王培杰, 董彦强, 等. 基于多尺度空间注意力引导的图像超分辨率重建网络 [J/OL]. 北京航空航天大学学报, 1-12[2024-04-21]. <https://doi.org/10.13700/j.bh.1001-5965.2023.0547>.

[4] Cheng W, Ma T, Wang X, et al. Anomaly detection for internet of things time series data using generative adversarial networks with attention mechanism in smart agriculture[J]. Frontiers in Plant Science, 2022, 13: 890563.

[5] Liu B, Zhao X, Hu H, et al. Detection of Esophageal Cancer Lesions Based on CBAM Faster R-CNN[J]. Journal of Theory and Practice of Engineering Science, 2023, 3(12): 36-42.

[6] Huang Y, Shi P, He H, et al. Senet: spatial information enhancement for semantic segmentation neural networks[J]. The Visual Computer, 2023: 1-14.

作者简介:

万福涛 (1995.11-) 男, 汉族, 安徽六安人, 西京学院计算机学院硕士研究生在读。

主要研究方向: 图像处理与目标检测