

基于 FFR-YOLO 的森林火灾多目标实时检测方法研究

罗佳慧

大连民族大学

DOI:10.12238/acair.v3i2.13541

[摘要] 本研究提出FFR-YOLO模型以解决森林火灾检测在复杂地形中的准确性和实时性问题。通过结构化无锚框模块和动态非极大值抑制算法,显著降低了计算量,使非极大值抑制处理时延缩短。采用非中心化局部感受野和步进式下采样策略,增强了多尺度火焰特征捕捉能力。改进的AF模块(含编码器和Fusion结构)提升了跨尺度特征融合效果。实验表明,FFR-YOLO的mAP@0.5达到97.4%,较YOLOv8提升5.7%,同时保持131 FPS的实时性能。该研究为复杂环境下的森林火灾实时检测提供了高效解决方案,兼具生态保护和社会经济价值。

[关键词] 深度学习; 森林火灾检测; YOLOv8; 注意力机制

中图分类号: S762.1 **文献标识码:** A

Research on real-time multi-target detection of forest fire based on FFR-YOLO

Jiahui Luo

Dalian National University

[Abstract] This study proposes the FFR-YOLO model to address the issues of accuracy and real-time performance in forest fire detection in complex terrains. By using a structured anchor-free module and a dynamic non-maximum suppression algorithm, the computational load is significantly reduced, and the processing delay of non-maximum suppression is shortened. The non-centralized local receptive field and stepwise downsampling strategy enhance the ability to capture multi-scale flame features. The improved AF module (including the encoder and Fusion structure) improves the cross-scale feature fusion effect. Experiments show that the mAP@0.5 of FFR-YOLO reaches 97.4%, which is 5.7% higher than that of YOLOv8, while maintaining a real-time performance of 131 FPS. This research provides an efficient solution for real-time forest fire detection in complex environments, with both ecological protection and social economic value.

[Key words] deep learning; forest fire detection; YOLOv8; attention mechanism

引言

森林是陆地生态系统的核心组成部分^[1], 为人类生存发展提供基础资源保障, 在全球碳平衡、气候稳定^[2]、水源调控等关键生态过程中具有决定性的影响。然而, 近年来由于气候变化和人类活动加剧, 森林资源正面临前所未有的威胁, 其中森林火灾是最具破坏性的因素之一^[3]。

然而目前存在的双阶段^[4]和单阶段^[5]算法, 双阶段目标检测算法例如R-CNN^[6]、Faster R-CNN^[7]、R-FCN^[8]等, 其特点是运行速度较慢, 并且要占用大量的磁盘空间, 难以满足实时检测的场景要求, 而单阶段目标检测算法, 如OverFeat^[9]和SSD^[10]、由于将区域建议和目标检测融合为一个阶段, 它比双阶段目标检测算法在速度上更有优势, 应用场景更加广泛。但是, 单阶段的目标检测算法在检测精度上仍有提升空间。为提高模型对于多目标火灾的检测能力的同时提高检测速度, 本文改进YOLOv8模型

的网络结构, 以求在速度和精度上达到更佳的平衡状态, 完成森林火灾多目标实时检测方法的研究。

1 研究方法

1.1 改进YOLOv8检测模型

针对森林火灾的多目标实时检测问题, 提出了S-HG模块, 解决YOLO模型中存在的非最大值抑制(NMS)速度延迟及优化困难的问题。在主干网络中整合高效的HG模块, 并采用Shuffle卷积替代传统卷积操作, 实现了实时检测的目标。同时, 设计了一层Transformer Encoder, 仅处理来自主干网络的特征, 这一部分被称为AF模块。之后在检测头部分, 更进一步提出了非中心化的局部注意力机制HelloNet, 以及步进自注意力下采样操作, 这些创新有效减少了计算复杂度和内存需求, 从而提升了模型的效率和准确性。这些改进共同推动了森林火灾检测技术的发展。

1.2 S-HG模块

本研究提出S-HG模块优化了YOLO模型的非极大值抑制延迟问题。该模块通过整合高效HG模块和采用Shuffle卷积替代传统卷积,显著提升了处理速度和检测效率,同时降低了计算量和内存占用。

具体实现上,采用CSP策略和通道洗牌方法重组HG模块。首先对 1×1 卷积进行通道洗牌,然后将改组后的特征图与原始特征图连接,最后通过第二个 1×1 卷积生成输出。这种设计通过通道随机排序增强了特征融合能力,如图1所示。

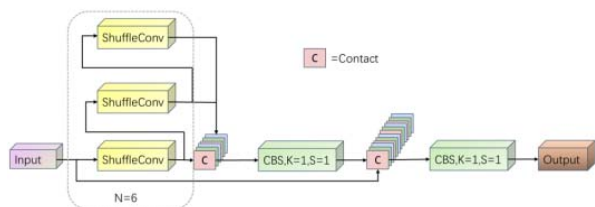


图1 S-HG Block架构

1.3 HelloNet检测头

检测头部分提出非中心化的局部注意力和步进自注意力下采样操作,减少计算复杂度和内存需求,提高模型的效率和准确性。

本研究提出helloing操作来解决模块内信息损失问题。通过在模块外用原始图像信息padding一圈,有效扩大了感受野,使每个模块能获取更丰富的上下文信息。相比传统方法仅关注模块内部信息,该操作显著提升了特征提取能力。

在下采样方面,创新性地采用分模块采样结合注意力的机制,替代传统的平均/最大池化。该方法对每个模块采样后进行注意力计算,在保持精度的同时减少了4倍FLOPs。这种设计通过对模块间信息的高效处理,实现了计算量的大幅降低,为模型轻量化提供了新思路。

HelloNet引入了一个新颖的局部自注意力机制,该机制通过helloing技术高效映射到现有硬件上。与全局自注意力相比,局部自注意力通过限制每个注意力操作的范围来减少计算复杂度和内存需求,使得模型能在不牺牲性能的情况下,有效扩展到更大的数据集和更复杂的任务上。

1.4 AF模块

特征提取模块是指从低级特征中提取出高级特征,这些特征包含了图像中物体的丰富语义信息。但从直观上讲,在连接多尺度特征上进行特征交互是多余的。因此,对于该部分重新设计了一系列具有不同编码器的变种AF,以及帮助完成跨尺度融合的Fusion模块共同构成AF模块,如图2所示。

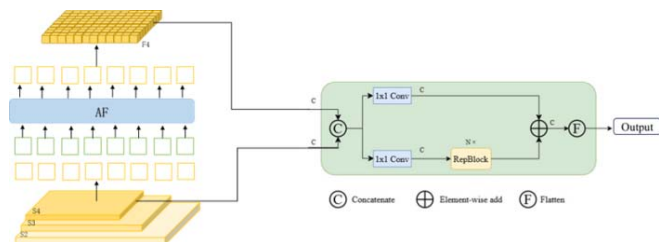


图2 AF模块

对于混合编码器的部分,它将二维的S4特征拉成向量,然后交给模块处理,其数学过程就是多头自注意力与FFN,随后,再将输出调整回二维,以便去完成后续的跨尺度特征融合。

对于Fusion模块,将其YOLO的角度看这个结构更便于理解,这个Fusion模块就是一个FPN/PAN结构。该部分是由2个 1×1 卷积和N个RepBlock构成的,可通过调整Fusion中RepBlock的数量和Encoder的编码维度控制AF中的深度和宽度。

值得注意的是,AF模块只对S4特征图处理。可以用如下公式表示。

$$Q = K = V = \text{Flatten}(S4)$$

$$F4 = \text{Reshape}(\text{Attn}(Q, K, V))$$

$$\text{Output} = \text{Fussion}(\{S4, F4\})$$

AF通过引入多尺度特征并提出了可变形注意力机制来减少计算量却存在的输入序列的大幅增加而导致编码器成为计算瓶颈的问题。

2 实验结果与分析

2.1 实验设置

为了确保高效的模型训练过程,需要处理大量的图像数据,通过从互联网和公开数据集上采集相关火灾图像,严格标注构建了数据集,提升模型对复杂火灾场景中多目标的识别精度。最终构建的数据集规模达到38267张图像,满足了深度学习算法的训练需求。此次实验采用GPU进行加速,GPU型号为NVIDIA RTX A4000,使用Python3.9作为编译语言。

2.2 评价指标

本实验选用的性能指标有精度(Precision)、召回率(Recal)、平均精度(mAP),计算公式如下所示:

$$\text{Precision} = \frac{TP}{TP + FP}$$

公式中TP表示实际上为正例并且被分类器划分为正例的个数。FP表示实际为负例但是被分类器划分为正例的实例数。TN实际为负例并且被分类器划分为负例的实例数。FN实际为正例但被分类器划分为负例的实例数。

召回率的具体公式如下所示:

$$\text{Recall} = \frac{TP}{TP + FN}$$

召回率衡量的是实际正例中被模型正确识别出来的比例。高召回率意味着模型能够更准确地找出图像中实际存在的火焰和烟雾,减少漏检的情况。

mAP是目标检测任务中常用的一个评估指标。在面对复杂场景和多种类别检测任务时具有重要意义。具体公式如下所示:

$$mAP = \frac{\sum_{n=1}^N \int_0^1 Precision_n d(Recall_n)}{N}$$

它衡量的是模型在所有类别上的平均精度,能够综合反映模型在不同类别上的性能。

2.3对比实验

本研究通过对比实验验证了改进YOLOv8模型在火焰和烟雾检测任务中的优越性能。如表1所示,与YOLOv4至YOLOv11等经典算法相比,该模型在速度-精度权衡方面表现最佳,不仅显著优于原始YOLOv8,也超越了YOLOv5、YOLOv7等同类模型。

具体而言,改进后的模型在检测精度上提升了5.7个百分点,召回率提高了8.8个百分点,有效减少了误报和漏报。同时,模型帧率达到131.2FPS,完全满足实时检测需求。

这些结果表明,本文提出的优化策略显著提升了火灾检测的准确性和稳定性,为实际应用提供了可靠的技术支持,具有重要的研究价值和应用前景。

表1 对比实验结果

Dataset	Model	Precision(%)	Recall(%)	mAP@0.5(%)
森林火灾数据集	YOLOv4	73.2	60.0	66.9
	Faster R-CNN	\	\	75.7
	YOLOv5	87.4	85.4	95.4
	YOLOv7	91.5	83.6	91.1
	YOLOv8	92.2	86.5	91.7
	YOLOv11	90.5	81.5	91.4
	Ours	97.8	95.3	97.4

2.4消融实验

表2的消融实验验证了所提改进策略的有效性。实验在相同条件下对比了S-HG模块、AF模块和Hello检测头的作用。单独使用S-HG模块时,mAP@0.5略有下降,这是由于引入了梯度流信息的注入,促进了网络的学习和特征提取,因此该下降并不会太多影响,所以结合S-HG和AF模块,并利用实时检测的原理后,mAP@0.5提升了0.8%,同时FPS增加。替换为Hello检测头后,mAP@0.5上升2.8%,因其优化了下采样效率。最终,调整损失函数进一步提升了FPS。综合所有改进,模型mAP@0.5提高5.7%,FPS达131.2,显著提升了检测速度和精度,减少了漏检,验证了方法的合理性和实用性。

表2 消融实验结果

NO.	Model	mAP@0.5(%)	mAP@0.5:0.95(%)	FPS(fps)
1	YOLOv8	91.7	63.4	106.7
2	YOLOv8+S-HG	83.4	57.9	99.9
3	YOLOv8+S-HG+AF	92.5	67.5	115.3
4	YOLOv8+S-HG+AF+Hello	97.4	80.4	131.2

综上,本文所提出的模型更加适应森林火灾检测任务,能够在兼顾速度和精度的最佳平衡状态下完成森林火灾的多目标实时检测任务。

3 结论

本文在森林火灾多目标实时检测领域展开了研究,本文使用YOLOv8作为基础网络,将骨干网络进行修改,使用S-HG模块得

到一个更有利于GPU推理的backbone,精度、速度上超越其它CNN模型。再引入AF模块提高特征融合能力。在进一步提升精度的同时,也提高了实时性。改进了Head部分,使用Hello技术,在其引入局部自注意力机制限制每帧减少计算的复杂度和内存的需求,使原模型在不牺牲性能的前提下,提升扩展到更大的数据集和更复杂的任务上面。最后实验结果表明,优化后的模型相较于原始模型来说,在森林火灾检测任务上表现更加出色。

参考文献

- [1]Dixon R K, Solomon A M, Brown S, Houghton R A, Trexler M C, Wisniewski J. Carbon pools and flux of global forest ecosystems. Science, 1994, 263(5144): 185-190.
- [2]刘世荣,代力民,温远光.面向生态系统服务的森林生态系统经营:现状、挑战与展望[J].生态学报,2015,35(1):1-9.
- [3]Filkov A I, Ngo T, Matthews S, et al. Impact of Australia's Black Summer fires on fauna and ecosystems[J]. Nature Climate Change, 2020, 10(3): 171-176.
- [4]GIRSHICK R. Fast R-CNN[C]//IEEE International Conference on Computer Vision (ICCV). 2015: 1440-1448.
- [5]REDMON J, DIVVALA S, GIRSHICK R, et al. You Only Look Once: Unified, Real-Time Object Detection[C]//IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2016: 779-788.
- [6]HE K, ZHANG X, REN S, et al. Spatial Pyramid Pooling in Deep Convolutional Networks for Visual Recognition[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 37(9): 1904-1916.
- [7]REN S, HE K, 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.
- [8]DAI J, LI Y, HE K, et al. R-FCN: Object Detection via Region-based Fully Convolutional Networks[J]. Advances in Neural Information Processing Systems, 2016, 29: 379-387.
- [9]ERMANET P, EIGEN D, ZHANG X, et al. OverFeat: Integrated Recognition, Localization and Detection using Convolutional Networks[C]//International Conference on Learning Representations (ICLR). 2014.
- [10]LIU W, ANGUELOV D, ERHAN D, et al. SSD: Single Shot Multi-Box Detector[C]//European Conference on Computer Vision (ECCV). 2016: 21-37.

作者简介:

罗佳慧(2000—),女,汉族,河北省秦皇岛市人,硕士研究生在读,研究方向:人工智能。