

基于改进 YOLOv8 的森林火灾检测算法研究

陈义啸, 沈景凤, 仲梁维

(上海理工大学机械工程学院, 上海 200093)

✉ cyx1873891950@163.com; sh_jf@163.com; zlv@usst.edu.cn



摘要:随着全球天气持续变暖,高温、干旱、大风等极端天气呈现多发并发态势,导致全球范围内森林火灾频繁爆发。为了提升森林火灾检测精确率和实时性,提出了一种基于改进 YOLOv8 的森林火灾检测算法模型。该模型在 Neck 端上采样阶段中的 Upsample 结构后面以及每个 C2F 模块后面添加 CBAM (Convolutional Block Attention Module) 注意力机制,通过对特征通道和空间的学习,提高模型的特征提取能力,解决火焰和烟雾图像分辨率低和受不同场景因素影响导致的识别率低的问题。将该算法模型应用于火灾数据集进行训练、验证、测试发现,与原算法模型相比,经改进的算法模型的准确率和召回率分别提高了 6.5% 和 6.8%,其中 $mAP@0.5$ 提高了 4.8%。实验结果表明,改进后的算法模型能够实现对森林火灾的实时监测与精确识别。

关键词:YOLOv8; CBAM 注意力机制; 森林火灾检测

中图分类号:TP391 **文献标志码:**A

Research on Forest Fire Detection Algorithm Based on Improved YOLOv8

CHEN Yixiao, SHEN Jingfeng, ZHONG Liangwei

(School of Mechanical Engineering, University of Shanghai for Science and Technology, Shanghai 200093, China)

✉ cyx1873891950@163.com; sh_jf@163.com; zlv@usst.edu.cn

Abstract: With the continuous warming of global weather, extreme weather such as high temperatures, droughts, and strong winds are becoming more frequent and concurrent, leading to frequent outbreaks of forest fires around the world. In order to improve the accuracy and real-time performance of forest fire detection, this paper proposes a forest fire detection algorithm model based on improved YOLOv8. This model adds CBAM (Convolutional Block Attention Module) attention mechanism after the Upsample structure in the Neck upsampling stage and after each C2F module, which improves the feature extraction ability of the model through learning feature channels and spaces and solves the problem of low resolution of flame and smoke images and low recognition rate caused by different scene factors. Applying the algorithm model to the fire dataset for training, validation, and testing, it has been found that compared with the original algorithm model, the accuracy and recall rates of the improved algorithm model increase by 6.5% and 6.8%, respectively and $mAP@0.5$ increases by 4.8%. The experimental results show that the improved algorithm model can achieve real-time monitoring and accurate identification of forest fires.

Key words: YOLOv8; CBAM attention mechanism; forest fire detection

0 引言 (Introduction)

随着全球天气持续变暖,高温、干旱、大风等极端天气呈现

多发并发态势,导致全球范围内森林火灾频繁爆发。据统计,2022 年前 10 个月,我国共发生森林火灾 648 起^[1]。为有效监

测和防范森林火灾,近年来基于深度学习的目标检测技术逐渐成为研究热点。现阶段,传统的森林火灾检测算法在精确性和实时性方面存在诸多挑战,尤其是受限于火焰和烟雾图像分辨率较低和不同场景因素的影响。

当前,YOLOv8 在目标检测领域的表现出色,但在复杂场景下,特别是在森林火灾检测中,其性能仍有不足,包括分辨率低、易受环境影响等。针对以上问题,本文提出一种新的算法模型 YOLOv8,并在改进算法的基础上加入了 CBAM 注意力机制^[2]。通过实验数据对比,验证了与原始算法相比,改进之后的算法具有更高的检测精度。

1 相关研究 (Related research)

近年来,研究者将深度学习和机器学习应用于图像领域,取得了较好的成果,搭建了一些预测模型。黄翰鹏等^[3]提出一种新型 FNN-TSM 火灾预警算法,将模糊神经网络模型和温度时序模型结合,有效解决了传统火灾检测器误报率高的问题。卢鹏等^[4]提出一种基于 SSD_MobileNet 的复杂环境火焰区域标记方法,适用于实时火焰检测领域。张彬彬等^[5]提出将改进的 CIOU 和 Focal loss 作为损失函数,能满足火焰检测的实时性和高效性要求,但是训练和检测速度有所下降。王冠博等^[6]通过 K-Means 聚类 and 引入注意力机制,提出改进型 YOLOv4 火焰图像实时检测模型,解决火焰受周围因素干扰的问题,使检测效果得到了很大的提升。目前,以上算法大都针对城市和周边居民区的火灾检测,有些算法只能对火焰进行检测,但忽略了对烟雾的检测^[7],而在森林火灾发生的初期,对烟雾的检测恰恰是不能缺失的,并且自然环境下的火焰和烟雾的形态与城市环境下的火焰和烟雾的形态差别较大。在不同场景下,面对不同天气变化和周围环境变化,精确且实时地检测烟雾和火焰是检测算法研究人员面临的一个难题。因此,必须设计一个合适的预测模型,既能够同时检测火焰和烟雾,又要保证其检测精度。本文采用 YOLO 系列中最新算法模型 YOLOv8,在 Neck 层中改进算法和加入注意力机制,对常见的森林火灾发生场景进行识别,验证其识别效果。

2 YOLOv8 原理及模型架构 (YOLOv8 principles and model architecture)

目前,YOLO 系列中最新的算法模型是 YOLOv8,其在公开的数据集上的检测准确率和速度均超过其他 YOLO 系列的算法模型。YOLOv8 算法由 Glenn-Jocher 提出,与 YOLOv3 算法、YOLOv5 算法一脉相承^[8]。YOLOv8 网络结构图如图 1 所示,主要组成部分为 Input、Backbone、Neck 和 Head。

2.1 Input

在输入端部分,主要包括 Mosaic 图像增强、自适应锚框计算和自适应图片缩放。

采用 Mosaic 数据增强方法的思想是随机使用 4 张不同图像,将其随机拼接成一张大的图像,可以增加训练集的多样性

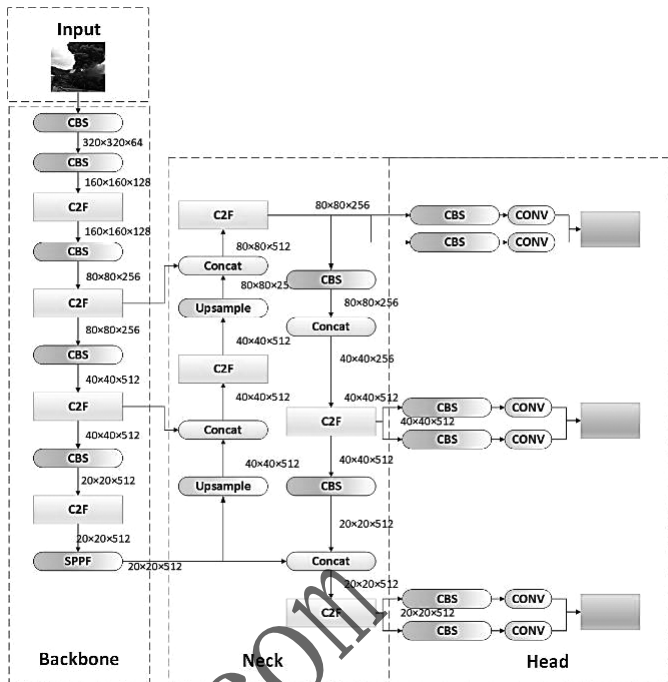


图 1 YOLOv8 网络结构图

Fig. 1 YOLOv8 network structure diagram

和难度,有助于提高目标检测模型的泛化能力。

在网络训练前,自适应锚框通过学习的方式自动计算出最适合输入图像的锚框参数,不需要手动设置。这种方法可以提高目标检测的精度和鲁棒性。

在 YOLOv4 之后加入了图像混叠增强技术,但是在训练的过程中全程启用图像混叠增强技术会影响训练效果。YOLOv8 会在最后训练时期关闭图像混叠增强技术,从而得到更好的训练效果。

2.2 Backbone

Backbone 由 CBS、C2F、SPPF 等模块组成,主要是对目标特征进行提取。其中,C2F 模块参考了 YOLOv7^[9] 中 ELAN 的设计思想以及 YOLOv5 中的 C3 模块进行设计,将原先的 C3 模块全部换成 C2F 模块。同时,在该模块中加入更多的分支。图 2 为 YOLOv8 的 C2F 模块结构图,图 3 为 YOLOv5 的 C3 模块结构图。SPPF 仍然使用 YOLOv5 中的模块。Bottleneck 是一种特殊的残差结构。

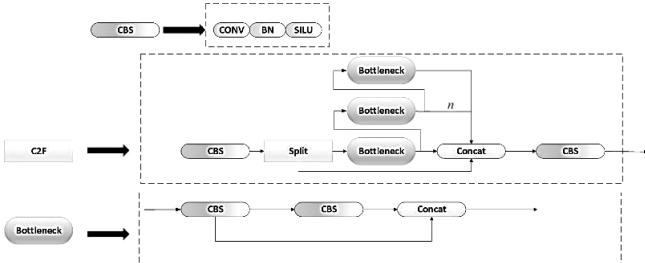


图 2 YOLOv8 的 C2F 模块结构图

Fig. 2 C2F module structure diagram of YOLOv8

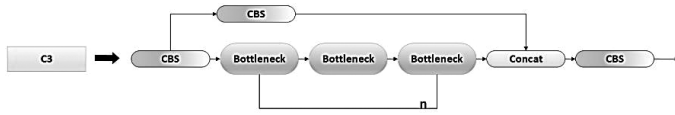


图3 YOLOv5 的 C3 模块结构图

Fig. 3 C3 module structure diagram of YOLOv5

2.3 Neck

Neck 层仍然采用了 YOLOv5 中的特征金字塔网络 (Feature Pyramid Network, FPN) [10] 和路径聚合网络 (Path Aggregation Network, PAN) [11], 对不同层的特征信息进行融合。FPN 和 PAN 相互配合使用, 可以提高模型的检测性能。FPN 结构可以增强模型的感知能力和尺度不变性, 而 PAN 结构可以增强多尺度特征的融合能力。在结构上仅需要将 C3 模块替换为 C2F 模块, 其余结构不变。

2.4 Head

本文设计的算法模型有一部分改进是针对 Head 层, 与 YOLOv5 相比, 有两大改进之处。首先使用了目前主流的解耦头 (Decoupled-Head) 结构, 将分类和检测头分离。其次将 Anchor-Based 换成 Anchor-Free。该模块负责检测大尺度目标、中尺度目标和小尺度目标, 用于将特征图转换为目标框的预测结果, 能够准确地预测图像中目标的位置、类别和边界框信息。

目前, YOLOv8 官方代码中给出了 YOLOv8n、YOLOv8s、YOLOv8m、YOLOv8l 和 YOLOv8x 共 5 个不同网络宽度和深度的版本。考虑到模型的检测速度和精度, 本文采用 YOLOv8n 为基线模型对目标进行检测, 并在此基础上进行了改进, 检测速度快、稳定性高, 具有深度和特征图宽度最小的优点, 可以很好地用于火焰和烟雾目标的实时监测。

3 改进 YOLOv8 识别模型 (Improvement of YOLOv8 recognition model)

3.1 CBAM 注意力机制

CBAM (Convolutional Block Attention Module) 是一种组合模型, 它将通道注意力和空间注意力结合, 可以提高模型的表现力。CBAM 模块包含通道注意力和空间注意力两个子模块, 可以对输入特征进行通道级和空间级的注意力调控。通道注意力模块计算每个通道的重要性, 以便更好地区分不同通道之间的特征; 空间注意力模块计算每个像素在空间上的重要性, 帮助模型更好地捕捉图像中的空间结构。

3.2 CBAM 注意力机制原理

CBAM 是用于增强卷积神经网络 (CNN) 性能的注意力机制模块, 能够应用于很多卷积神经网络中, 如 GoogLeNet [12]、VGGNet [13] 和 ResNet [14] 等。CBAM 的主要目标是通过在 CNN 中引入通道注意力模块和空间注意力模块提高模型的感

知能力, 从而在不增加网络复杂性的情况下提高其性能。

通道注意力模块是一种用于确定每个通道对最终特征的重要性的方法, 通道注意力机制模型结构如图 4 所示, 它通过对卷积后的特征图进行全局平均池化, 得到每个通道的全局特征描述。通过两个全连接层将全局特征描述映射为通道权重, 表示每个通道对最终特征的重要性, 然后将通道权重与原始特征图相乘, 得到经过通道注意力调整的特征图。这样可以更好地区分不同通道之间的特征, 提高特征的表达能力。

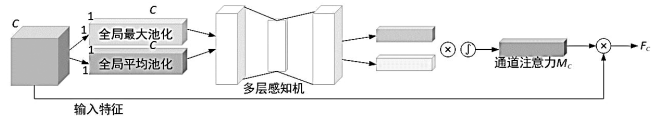


图4 通道注意力机制模型结构

Fig. 4 Structure of the channel attention mechanism model

$$M_c(F) = \sigma(MLP(AvgPool(F)) + MLP(MaxPool(F))) \\ = \sigma(W_1(W_0(F_{avg}))) + W_1(W_0(F_{max}))) \quad (1)$$

公式(1)中的 σ 表示 Sigmoid 函数, $W_0 \in R^{C \times C/r}$, $W_1 \in R^{C \times C/r}$ 。MLP 的权重由 W_0 和 W_1 共享。 W_0 前有 ReLU 激活函数。

空间注意力模块则用于计算每个像素在空间上的重要性, 以更好地捕捉图像中的空间结构。空间注意力机制模型结构如图 5 所示, 首先通过对通道注意力模块输出的特征图进行最大池化和平均池化操作, 得到每个通道的最大值和平均值。将这些值拼接并输入一个全连接层中, 生成表示每个像素空间权重的向量。将空间权重向量与经过通道注意力调整的特征图相乘, 得到最终经过空间注意力调整的特征图。这样可以更好地区分不同像素空间的重要性, 提高特征的表达能力。

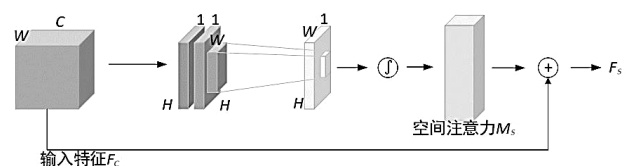


图5 空间注意力机制模型结构

Fig. 5 Structure of the spatial attention mechanism model

$$M_s(F) = \sigma(f^{7 \times 7}([AvgPool(F); MaxPool(F)])) \\ = \sigma(f^{7 \times 7}([F_{avg}^s; F_{max}^s])) \quad (2)$$

公式(2)中的 σ 表示 Sigmoid 函数, $f^{7 \times 7}$ 表示 7×7 大小的卷积核。

CBAM 注意力机制能够自适应地调整不同的通道和像素在特征映射中的权重, 它可以针对不同的任务和数据集自动选择最优的特征组合, 从而提高模型的性能。

3.3 改进模型结构

在初始 YOLOv8 网络结构的基础上, 加入效果更好的混

4.3 实验平台配置

本次实验中使用的 YOLOv8 模型是在 Windows 10 的操作系统下配置,CPU 型号为 Intel(R)Core(TM)i7-8750 H CPU @2.2 GHz 2.21 GHz,运行内存为 16 GB,GPU 型号为 NVIDIA GeForce GTX 1050 Ti,深度学习框架为 Pytorch2.0.1+cu118,运算架构为 Cuda11.8,开发环境是 Python 3.8。

4.4 训练过程及评价指标

为了验证图像的检测性能,给出模型的评价指标。

(1)准确率 P (Precision)。准确率可以衡量模型的精确性,可以定义为模型预测的正类别中有多少是真正的正类别,其公式如下:

$$P=\frac{TP}{TP+FP}$$
 (5)

其中: TP 表示模型正确地将正例判定为正例的数量, FP 表示模型错误地将负例判定为正例的数量。

(2)召回率 R (Recall)。召回率可以衡量模型的全面性,可以定义为所有真正的正类别中有多少被模型成功检测到,其公式如下:

$$R=\frac{TP}{TP+FN}$$
 (6)

其中: TP 表示模型正确地将正例判定为正例的数量, FN 表示模型错误地将正例判定为负例的数量。

(3)平均准确率 AP (Average Precision)。平均准确率可以衡量模型的检测性能,定义为对各个样本的检测精度,其公式如下:

$$AP=\int_0^1 P(R) dR$$
 (7)

(4)平均准确率均值 mAP (mean Average Precision)。平均准确率均值^[14]是一种综合评价指标,它综合考虑了模型在不同类别上的精准率,是所有类别 AP 的均值,其公式如下:

$$mAP=\frac{1}{C}\sum_{i=1}^C AP_i$$
 (8)

其中: AP_i 表示单个类别的平均精度, C 表示所有类别的数量。

5 实验结果展示 (Presentation of experimental results)

为了验证改进 YOLOv8 算法模型的有效性,将其与原始 YOLOv8 分别应用于数据集中进行训练实验,实验检测结果如表 1 所示。从表 1 中的数据可见,改进后的 YOLOv8 模型相比原始的 YOLOv8 模型, P 、 R 和 $mAP@0.5$ 指标分别提高了 6.5%、6.8%和 4.8%,表明采用了 CBAM 注意力机制模块后,较好地满足了对火焰和烟雾的检测要求。

表1 实验检测结果

Tab.1 Test results of the experiment

算法	准确率/%	召回率/%	$mAP@0.5$ /%
原始 YOLOv8	68.4	59.9	65.5
改进 YOLOv8	74.9	66.7	70.3

图 9(a)和图 9(c)为原始的 YOLOv8 的检测结果,图 9(b)和图 9(d)为添加了 CBAM 注意力机制模块的火灾检测结果。对比 4 张图片发现,两种网络模型均能够对火灾进行识别,但是未添加注意力机制的网络对于检测框的置信度低于添加注意力机制的改进网络,所以经改进的 YOLOv8 能更精确地对火焰和烟雾进行识别。



(a)原始 YOLOv8 fire 模型



(b)改进 YOLOv8 fire 模型



(c)原始 YOLOv8 smoke 模型



(d)改进 YOLOv8 smoke 模型

图9 实验结果

Fig. 9 Experimental results

6 结论(Conclusion)

为提升森林火灾检测精确率和实时性差等,本文提出一种基于改进 YOLOv8 的森林火灾检测算法模型,引入 CBAM 注意力机制到 Neck 层中 C2F 模块后面和 Upsample 结构后面,能够提高模型特征提取的准确率,使网络模型更好地学习森林火灾场景,最终提高其对火灾和烟雾的识别灵敏度和辨别能力。从最终的实验结果可知,相比于原模型,改进后的 YOLOv8 模型的 $mAP@0.5$ 提高了 4.8%,具有较高的检测精度。目前,该模型还处于优化阶段,对于一些特殊情况还要进一步研究、调整,之后的研究中将加入更多不同的数据集对不同环境下的火灾进行检测,并进一步研究算法对目标检测的识别性能。

参考文献(References)

- [1] 杜秋洋,杨传杰,刘鹏科,等. 2022 年全球野火发生情况及应对策略分析[J]. 森林防火,2023,41(2):49-53.
- [2] WOO S, PARK J, LEE J Y, et al. CBAM: convolutional block attention module[C]//FERRARI V, HEBERT M, SMINCHISCU C, et al. Computer Vision—ECCV 2018: 15th European Conference. Cham: Springer, 2018: 3-19.
- [3] 黄翰鹏,李柏林,欧阳,等. 融合模糊神经网络与时序模型的火灾预警算法[J]. 计算机工程与设计,2020,41(6): 1639-1644.
- [4] 卢鹏,赵亚琴,陈越,等. 复杂背景环境下基于 SSD MobileNet 深度学习模型的火焰图像识别研究[J]. 火灾科学,2020,29(3):142-149.
- [5] 张彬彬,帕孜来·马合木提. 基于 YOLOv3 改进的火焰目标检测算法[J]. 激光与光电子学进展,2021,58(24): 289-296.
- [6] 王冠博,杨俊东,李波,等. 改进 YOLOv4 的火焰图像实时检测[J]. 计算机工程与设计,2022,43(5):1358-1365.
- [7] 皮骏,刘宇恒,李久昊. 基于 YOLOv5s 的轻量化森林火灾检测算法研究[J]. 图学学报,2023,44(1):26-32.
- [8] 韩强. 面向小目标检测的改进 YOLOv8 算法研究[D]. 长春:吉林大学,2023.
- [9] WANG C Y, BOCHKOVSKIY A, LIAO H Y M. YOLOv7: trainable bag-of-freebies sets new state-of-the-art for real-time object detectors[C]//IEEE. Proceedings of the IEEE: 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition(CVPR). Piscataway: IEEE, 2023: 7464-7475.
- [10] LIN T Y, DOLLÁR P, GIRSHICK R, et al. Feature pyramid networks for object detection[C]//IEEE. Proceedings of the IEEE: 2017 IEEE Conference on Computer Vision and Pattern Recognition(CVPR). Piscataway: IEEE, 2017: 936-944.
- [11] LIU S, QI L, QIN H F, et al. Path aggregation network for instance segmentation[C]//IEEE. Proceedings of the IEEE: 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition(CVPR). Piscataway: IEEE, 2018: 8759-8768.
- [12] SZEGEDY C, LIU W, JIA Y Q, et al. Going deeper with convolutions[C]//IEEE. Proceedings of the IEEE: 2015 IEEE Conference on Computer Vision and Pattern Recognition(CVPR). Piscataway: IEEE, 2015: 1-9.
- [13] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition[J]. Computer science, 2014: 1-14.
- [14] HE K M, ZHANG X Y, REN S Q, et al. Deep residual learning for image recognition[C]//IEEE. Proceedings of the IEEE: 2016 IEEE Conference on Computer Vision and Pattern Recognition(CVPR). Piscataway: IEEE, 2016: 770-778.

作者简介:

陈义啸(1999-),男,硕士生。研究领域:计算机辅助智能设计制造。

沈景凤(1968-),女,博士,副教授。研究领域:现代设计理论及方法,优化设计。

仲梁维(1962-),男,硕士,教授。研究领域:计算机辅助智能设计制造。

关于《软件工程》期刊增页的通知

近年来,在主管、主办单位的正确领导下,广大作者、读者的大力支持下,《软件工程》的办刊质量稳步提升,影响力不断扩大,来稿数量与日俱增。但由于版面限制,导致大量优秀稿件积压,影响了论文的时效性与先进性。为了优化资源配置,提升期刊出版服务能力。经辽宁省新闻出版局审核批准,本刊从 2024 年第 1 期起由原来的每期 64 页增加至 80 页。每本定价仍为 10 元。特此通知。欢迎广大作者、读者踊跃投稿与订阅。

《软件工程》编辑部