

基于改进YOLOv8的森林火灾检测方法研究

雷建云^{ab}, 田祚汉^{ab}, 夏梦^{ac*}, 雷瑞璠^{ac}

(中南民族大学 a. 计算机学院; b. 湖北省制造企业智能管理工程技术研究中心; c. 农业区块链与智能管理湖北省工程研究中心, 武汉 430074)

摘要 针对森林火灾检测对实时性要求较高的问题, 提出了一种基于改进YOLOv8的森林火灾检测方法. 在YOLOv8的基础上, 采用轻量化特征提取网络EfficientNet优化YOLOv8原主干网络CSPDarknet53, 以减少计算量并提高模型的收敛速度, 进而提高火灾检测速度. 此外, 融入了SENet注意力机制模块, 以增强网络对火灾检测的准确性. 使用 α -IoU损失函数代替YOLOv8原始损失函数中的CIoU损失函数来计算定位损失, 该函数能够自适应地调整IoU的阈值, 更好地处理不同大小和形状的目标, 同时提高模型对小目标的检测性能. 结果表明: 该方法的平均准确率(mA@0.5P)达到了87.2%, 帧率(FPS)提升了17帧, 显著提高了火灾检测的实时性.

关键词 深度学习; YOLOv8模型; 森林火灾检测; 实时性

中图分类号 TP391.4 文献标志码 A 文章编号 1672-4321(2026)01-0097-09

doi: 10.20056/j.cnki.ZNMDZK.20250823

Research on forest fire detection method based on improved YOLOv8

LEI Jianyun^{ab}, TIAN Zuohan^{ab}, XIA Meng^{ac*}, LEI Ruifa^{ac}

(South-Central Minzu University, a. School of Computer Science; b. Hubei Provincial Engineering Research Center for Intelligent Management of Manufacturing Enterprises; c. Hubei Provincial Engineering Research Center of Agricultural Blockchain and Intelligent Management, Wuhan 430074, China)

Abstract An improved forest fire detection method based on YOLOv8 is proposed to solve the challenge of achieving high real-time performance in forest fire detection. Building upon YOLOv8, the lightweight feature extraction network EfficientNet is utilized to optimize the original YOLOv8 backbone network CSPDarknet53 to diminish computational requirements and accelerate model convergence, thereby accelerating the fire detection speed. Additionally, the SENet attention mechanism module is integrated to bolster the model's accuracy in fire detection. The α -IoU loss function is implemented to supplant the CIoU loss function from YOLOv8's original loss function for calculating positioning loss. This function can adaptively fine-tune the IoU threshold to more effectively handle targets of varying sizes and shapes, while also enhancing the model's capability to detect small targets. The outcomes demonstrate that the proposed method achieves an average accuracy of 87.2% at mA@0.5P, with a 17-frame increase in the frame rate per second (FPS), significantly enhancing the real-time capabilities of fire detection.

Keywords deep learning; YOLOv8 model; forest fire detection; real-time

随着科技的迅猛发展, 人工智能在各个领域中发挥着愈发重要的作用, 特别是在森林火灾检测领域. 森林火灾作为一种破坏性极强的自然灾害, 不仅会造成严重的经济损失, 还会对生态环境和人类

的生命安全构成威胁. 因此, 如何实现火灾的早期检测和及时预警, 成为当前研究的重点方向之一^[1-2]. 在传统的火灾检测方法中, 基于图像处理的技术占据主导地位. 例如, 许多早期研究依赖于火焰的颜

收稿日期 2024-10-13 * 通信作者 夏梦, 研究方向: 深度学习与图形识别, E-mail: xiameng@mail.scuec.edu.cn

作者简介 雷建云(1972-), 男, 教授, 博士, 研究方向: 大数据与网络空间安全, E-mail: lejianyun@mail.scuec.edu.cn

基金项目 湖北省技术创新计划重点研发资助专项(2023BAB087); 中央引导地方科技发展资金资助项目(ZYYD2024QY08); 武汉市重点研发计划资助项目(2023010402010614); 武汉东湖新技术开发区“揭榜挂帅”资助项目(2023KJB204)

色、形状以及烟雾的特征来识别火灾.这些方法的主要优势在于实现简单、计算代价较低,但在实际应用中常受到环境光线、天气状况等外部因素的干扰,导致误检和漏检问题^[3].随着深度学习的兴起,学者们逐渐将目光转向更为智能化和自动化的检测技术.深度学习中的卷积神经网络(CNN)为火灾检测提供了新的思路,能够通过训练大量的火灾图像数据来自动提取火焰和烟雾的特征,进而提升检测的准确性^[4-5].刘磊等利用CNN构建了一个火焰检测系统,该系统通过学习火焰在不同条件下的形态变化,实现了火灾的较为精准的认识^[6].尽管如此,由于CNN的计算复杂度较高,实时性仍然是一个重要的挑战.为了解决火灾检测中的实时性问题,YOLO(You Only Look Once)模型被广泛应用于该领域.YOLO模型作为一种统一的目标检测框架,YOLO通过将目标检测问题简化为一个回归问题,从而大幅提高了检测速度^[7-8].在YOLO的多个版本中,YOLOv3在目标检测的准确性和速度之间取得了良好的平衡,因此在许多实时检测任务中得到了应用.然而,YOLOv3在检测较小的火灾目标时表现不佳,这限制了其在森林火灾检测中的应用.为了解决上述问题,研究人员不断对YOLO模型进行改进.邵林等提出了一种基于YOLOv3的改进方法,使用多尺度特征融合技术增强了对小目标的检测能力,并在森林火灾检测中取得了较好的效果^[9].与此同时,YOLOv4和YOLOv5的发布进一步提升了目标检测的性能,尤其是在火灾检测的复杂场景中表现更为稳定^[10].与YOLOv4相比,YOLOv8的检测精度进一步提升,尤其是在实时目标检测任务中表现突出.这一进步为火灾检测提供了新的研究思路和应用场景^[11-12].然而,尽管YOLOv8在许多方面具有优势,如何在保持高准确度的同时进一步提升检测速度,仍然是一个值得深入探讨的问题^[13].综上所述,本文基于YOLOv8提出了一种改进的实时目标检测模型,旨在解决火灾检测中实时性与精度之间的平衡问题,通过优化模型结构并结合轻量化技术,改进后的模型不仅稳定了检测速度,还能够更加准确地识别早期火源,为火灾的早期预警提供了更为有效的技术支持.

文章的主要贡献如下.

(1)采用轻量化特征提取网络EfficientNet优化YOLOv8原主干网络CSPDarknet53.EfficientNet利用深度可分离卷积和通道注意力机制等轻量化技术,使得在推理阶段的计算速度更快.它具有更小的模型

尺寸,这使得在森林火灾检测场景中,使用EfficientNet可以减少模型的存储空间和内存占用,更适合在资源受限的环境中进行部署.虽然EfficientNet是一种轻量级网络,但它在图像分类和目标检测任务中仍然能够保持较高的性能.

(2)融入SENet注意力机制模块以提升网络对火灾检测的准确性.SENet注意力机制模块通过学习和利用不同通道之间的关系来提高模型性能.该模块通过全局平均池化操作将每个通道的特征图压缩为一个特征向量,这一操作有助于减少计算量和参数量,并引入通道间的关系.接着,通过一个全连接层来学习通道间的依赖关系.这个全连接层对压缩后的特征向量进行非线性变换,得到每个通道的权重.将得到的通道权重与原始的特征图相乘,实现对原始特征图的重标定,使模型对重要特征给予更高的关注,从而提高模型的性能.SENet注意力机制模块有助于YOLOv8模型更好地关注火灾相关特征,提高模型对火灾检测的准确性.

(3)使用 α -IoU损失函数代替YOLOv8原始损失函数中的CIoU损失函数来计算定位损失. α -IoU损失函数引入了自适应调整IoU值的机制,它根据目标的尺寸和形状动态调整IoU的阈值,以更好地匹配目标特征.这种方法的优势在于,对于森林中小目标,如初期火苗或者烟雾,可以使用较低的IoU阈值,提高小目标的检测率,而对于大目标,则可以使用较高的IoU阈值,确保大目标的定位精度.这种自适应调整IoU阈值的机制使模型具备更好的灵活性和鲁棒性,能够适应不同目标特征.

1 改进的YOLOv8模型

1.1 YOLOv8结构

YOLOv8是YOLO系列中最新的目标检测模型,它在先前版本的基础上进行了多项改进,旨在提高性能和灵活性.相较于之前的版本,YOLOv8的设计思想受到了YOLOv7 ELAN的启发,同时也借鉴了YOLOv5的一些特点^[14].首先,YOLOv8的骨干网络进行了优化,采用了梯度流更丰富的C2f结构,替代了YOLOv5中的C3结构.通过这一改变,YOLOv8能够更有效地捕捉图像中的特征,并提高模型的感知能力.此外,为了适应不同尺度的目标,YOLOv8对多尺度模型进行了通道数的调整,以进一步增强模型的表达能力^[15-16].

与其他传统的卷积神经网络模型相比,CNN在图像分类、目标检测等视觉任务上表现优秀.CNN的结构包括输入层、卷积层、激活函数、池化层、全连接层和输出层.CNN的卷积层和池化层能有效提取图像的多尺度特征,这对于目标检测等任务至关重要.由于CNN在图像处理领域的成熟性和卓越的特征提取能力,YOLOv8选择了CNN结构作为其主要组件^[17].YOLOv8的设计保留了CNN的优点,并在其基础上进行了多项创新和优化^[18].在Head部分,YOLOv8进行了两个重要改进:从原先的耦合头变为解耦头,这意味着YOLOv8将目标的类别和位置信息分别进行预测,使得模型更加灵活和可解释.其次,YOLOv8从YOLOv5的Anchor-Based方式转变为Anchor-Free方式,这一改变消除了先验框(Anchor)的需求,使模型能更好地适应不同尺度和形状的目标.这种Anchor-Free的设计减少了预测中的偏差,提高了目标检测的

准确性.YOLOv8还引入了一种新的正负样本匹配方式,称为Task-Aligned Assigner,与以往的IoU匹配或单边比例分配方式相比,Task-Aligned Assigner能够更精确地匹配正负样本,提升了模型的训练效果和稳定性,这种匹配方式能更好地平衡正负样本的数量,并使模型更好地学习目标特征^[19].Transform是一种新型的深度学习网络结构,它在目标检测方面同样表现出色,最早由谷歌Brain团队于2017年提出.Transform具有多个优点,例如其注意力机制可以自适应地关注不同尺度的目标,这有助于提高对不同大小目标的检测效果^[20-21].然而,Transform模型的参数量大,计算开销较高,这在实时检测等场景下可能存在效率瓶颈,相比之下,YOLOv8具有更高的计算效率和目标检测能力,YOLO系列模型已在工业界广泛应用,并拥有较为成熟的训练方案,因此针对本次森林火灾监测的任务,选择使用YOLOv8模型.

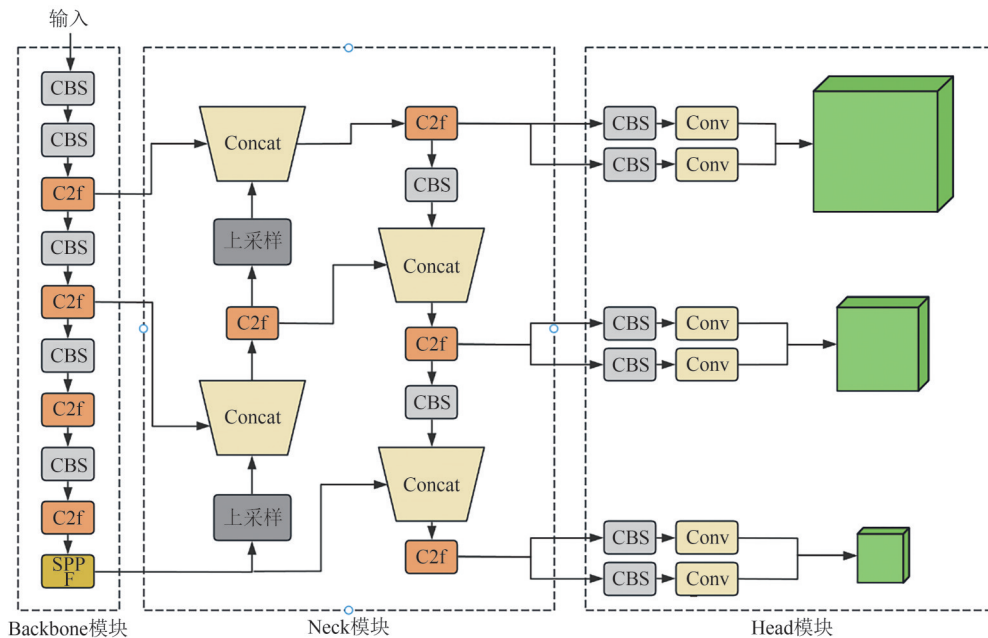


图1 YOLOv8整体网络结构

Fig. 1 YOLOv8 Overall network structure

1.2 轻量化的YOLOv8

使用轻量化特征提取网络EfficientNet优化YOLOv8的原主干网络CSPDarknet53,形成了新的EfficientNet-CSPDarknet主干网络,EfficientNetCSP

替代了CSPDarknet53中的C2f模块.

在保持网络深度的同时,引入EfficientNet的轻量级特性,提升了YOLOv8的检测实时性.EfficientNet是一种轻量化的特征提取网络,能够在保持高精度

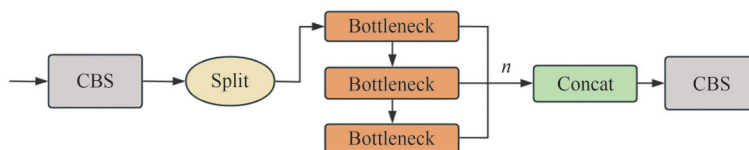


图2 C2f整体网络结构图

Fig. 2 C2f Overall network structure

的同时,减少模型的计算量和推理时间.尽管YOLOv8能够通过深层的卷积层和池化层捕捉图像中的边缘、纹理和形状等复杂特征,但在森林火灾检测中,对这些特征的要求并不高.因此,为了减少计算量并缩短推理时间,适宜引入更轻量化的模型结构,

以便在火灾初期更快速地进行检测和应对,特别是在森林火灾的初期阶段,这是进行控制和扑救的最佳时机.通过研究YOLOv8在火灾检测的实时性,能够在火灾刚开始时立即发出警报,使相关人员能够迅速做出反应,减少火灾的蔓延速度和范围.

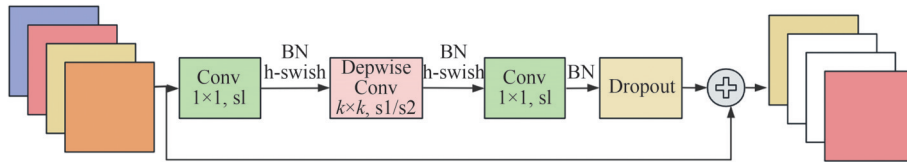


图3 EfficientNet网络结构图

Fig. 3 EfficientNet network structure diagram

EfficientNet网络包含一个 1×1 的普通卷积,其主要作用是升维,并包含批量归一化(BN)和h-swish激活函数;一个 $k \times k$ 的深度可分离卷积(Depwise Conv),其中 k 的取值有 3×3 和 5×5 两种情况;一个 1×1 的普通卷积,其作用是降维,包含BN和线性激活(线性激活函数定义为 $y=x$);以及一个Droupout层.EfficientNet原本使用非线性函数swish,这能够显著提升神经网络的精度,swish函数定义如下:

$$\text{swish } x = x \cdot \sigma(x), \quad (1)$$

式中, $\sigma(x) = \frac{1}{1 + e^{-x}}$,即常见的Sigmoid函数.虽然swish这种非线性激活函数提高了精度,但它在嵌入式环境中的计算成本较高,因为在移动设备上计算Sigmoid函数或对其求导代价较大.本文引入了公式(2)中的函数:

$$h - \text{Sigmoid}(x) = \frac{\text{ReLU6}(x + 3)}{6}, \quad (2)$$

该函数解决了Sigmoid函数带来的计算复杂性和求导复杂性.文中采用的激活函数如公式(3)所示:

$$h - \text{swish}[x] = x \frac{\text{ReLU6}(x + 3)}{6}. \quad (3)$$

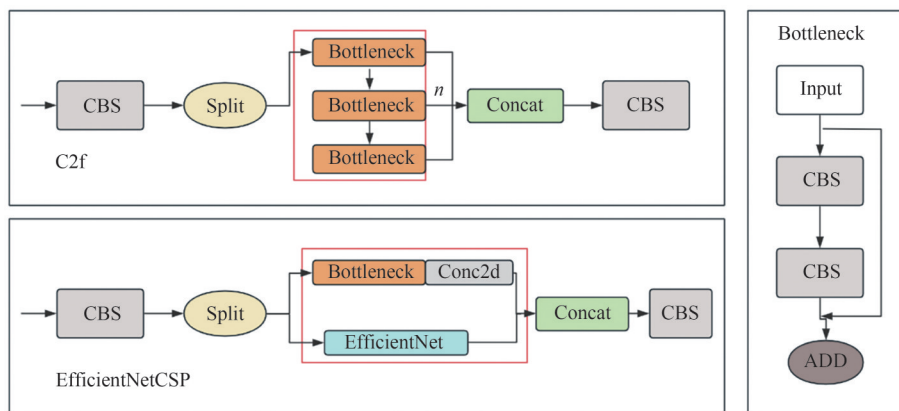


图4 EfficientNetCSP网络结构图

Fig. 4 EfficientNetCSP network structure diagram

1.3 SENet注意力机制模块

在EfficientNet中引入SENet注意力机制模块,形成了Efficient-SENet,这有助于EfficientNet在每个阶段自适应地调整特征通道的权重,增强模型对火苗重要特征的代表能力,提高森林火灾检测的准确性.同时,它还可以帮助模型更好地捕捉目标的语义信息,并降低对噪声和不相关特征的敏感性.

通过EfficientNet的主干网络提取语义信息和边缘等低级图像特征.对于每个特征图,使用全局

平均池化操作将其压缩为一个特征向量.这个特征向量捕捉了整个特征图的全局信息.对于每个特征向量,引入两个全连接层,其中一个用于降维,另一个用于升维.这些层可以学习到特征通道的相关性和重要性.通过使用Sigmoid函数将Excitation操作的输出限制在0到1之间,得到一个注意力权重向量.这个向量将根据特征通道的重要性,对特征图的每个通道进行缩放.将注意力权重向量与原始特征图相乘,得到加权后的特征图.这样,模型可以更加关

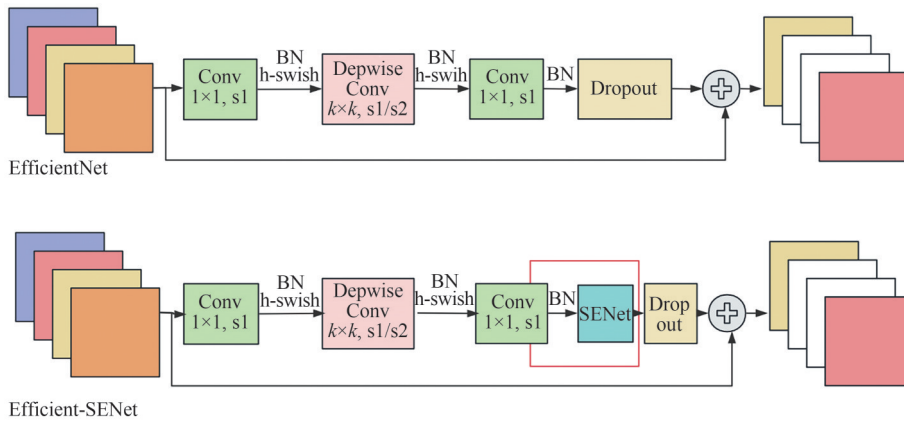


图5 Efficient-SENet网络结构图

Fig. 5 Efficient-SENet network structure diagram

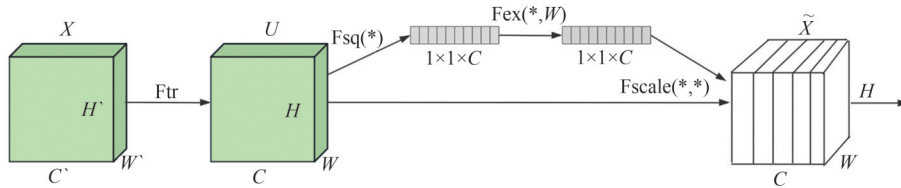


图6 SENet注意力机制结构图

Fig. 6 Structure diagram of SENet attention mechanism

注那些被认为是重要的特征通道,提高对目标的感知能力. X 为输入, U 为通过函数变换 F_{tr} 得到的特征图,即通过卷积得到的结果,在GoogleNet中就是Inception,在残差网络中就是Residual block. U 后面的上面那根线表示SE block的操作,最后再和原输出 U 相结合.输入 X 的通道数是 C' ,输出特征图 U 的通道数是 C ,要找到通道之间的关系,需要用到所有通道的信息,把每个通道都压缩成一个点,直接求平均,相当于就是这个通道的特征值.这个求平均的过程就是所谓的Squeeze. z_c 的表达如公式(4)所示:

$$z_c = F_{sq}(u_c) = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W u_c(i, j), \quad (4)$$

式中: u_c 表示特征图 U 中的某个通道, z_c 表示经过压缩之后的向量 Z ,从公式中可以看出,每个尺寸为 $H \times W$ 的特征图通道,经过全局平均之后,就只剩下一个像素点.因此,一个 $H \times W \times C$ 的特征图经过squeeze之后,就变成了一个 $1 \times 1 \times C$ 的向量,记作 Z .特征图经过压缩后,第二步就是综合这些通道的特征值.这个综合方法需要满足两个条件:其一要可以表达非线性关系,其二要可以提取处非互斥的关系.因此,提出了一个计算方法,如公式(5)所示:

$$s = F_{ex}(z, W) = \sigma(g(z, W)) = \sigma(W_2 \delta(W_1 z)), \quad (5)$$

式中: δ 为ReLU, $W_1, W_2 \in R^{C \times C}$ 为权重矩阵, $W_2, W_2 \in R^{C \times r}$ 为权重矩阵.从上面可以看出,这个Excitation就是

两个全联接层:第一个连接层把 $1 \times 1 \times C$ 的向量 Z 变成一个 $1 \times 1 \times \frac{C}{r}$ 的向量.然后后面跟一个ReLU激活层.第二个连接层把 $1 \times 1 \times \frac{C}{r}$ 的向量重新恢复为一个 $1 \times 1 \times C$ 的向量.然后后面跟一个Sigmoid的激活层.

1.4 损失函数优化

用 α -IoU损失函数替代原来的CIoU损失函数^[22],CIoU损失函数主要关注预测框与真实框之间的重叠程度、宽高比和中心距离等因素,然而,对于远处的目标,由于其尺度较小、特征不明显,CIoU损失函数在检测这些目标时显得不够敏感.这是因为CIoU损失函数在计算损失时,主要考虑的是预测框和真实框的重叠程度,而对于小目标,这种重叠程度不够显著.在森林火灾检测中,火源的形状可能会随着火势的变化发生剧烈变化.然而,CIoU损失函数在处理形状变化方面存在不足,它主要关注的是预测框和真实框的重叠程度及中心点距离,对形状变化并没有进行有效的建模.CIoU损失函数的计算公式如下(6):

$$CIoU Loss = 1 - IoU + (\alpha - IoU) \times \frac{v^2}{c^2}, \quad (6)$$

式中:IoU是传统的IoU计算结果,表示预测框和真实框的交集与并集的比值; α 是一个可调参数,用于平衡IoU项和长宽比一致性项的权重; v 是目标框的

长宽比的对数差值,用于衡量两个矩形框相对比例的一致性; c 是一个用于归一化的参数,通常取为两个矩形框闭合区域的对角线距离.

基于上述原因,改用 α -IoU 损失函数替代原来的 CIoU 损失函数. α -IoU 损失函数通过引入幂变换,能够更精确地进行边界框回归.在火灾检测中,该模型能够更准确地定位火源,从而提高检测的准确性,同时 α -IoU 损失函数对远处火源的检测更加敏感.火灾是一个动态变化的场景,火源的位置和形状可能会随着时间发生快速变化, α -IoU 损失函数通过引入幂变换,使得损失函数能更好地适应这种动态场景.结合随机裁剪技术.在训练过程中,通过随机裁剪图像并更新目标框坐标,确保模型能够处理不同尺度和位置的火源. α -IoU 损失函数的计算公式如下(7):

$$\alpha - \text{IoU Loss} = 1 - \text{IoU}^\alpha + \text{Regularization Term}, \quad (7)$$

式中:IoU 是传统的 IoU 计算结果,表示预测框和真实框的交集与并集的比值; α 是一个可调参数,用于控制幂变换的程度,当 $\alpha > 1$ 时,可以增加高 IoU 目标的损失和梯度,从而提高边界框回归的精度;Regularization Term 是一个正则化项,用于防止过拟合.

2 实验结论与分析

2.1 实验环境与数据

在本文的实验中,使用了配备 NVIDIA Tesla V100 SXM2 32GB 显卡的硬件环境进行训练. NVIDIA Tesla V100 是一款高性能计算显卡,采用了 NVIDIA 的 Volta 架构,拥有 32 GB 的显存,提供了强大的计算和内存容量,非常适合进行深度学习任务的训练和推理.在训练过程中,使用 CUDA (Compute Unified Device Architecture) 和 CUDNN (CUDA Deep Neural Network library) 等工具和库,以最大化利用 Tesla V100 的性能优势.实验环境配置为操作系统 Ubuntu 20,采用 Python 3.8 作为编程语言,并使用深度学习开源框架 Pytorch 进行编码实现,具体实验环境如下表所示.

在整个实验中,使用了 FLAME 公开数据集,该数据集总共包含 5500 张图片,包括红外相机拍摄的照片记录和热图.数据集中的每一张图片都进行了帧级标注,标注内容包括火灾的具体区域、火焰的边界、温度分布等.FLAME (Fire Luminosity Airborne-based Machine learning Evaluation) 是一个基于航拍

表1 实验环境

Tab. 1 Experimental environment

| 配置 | 参数 |
|---------|-----------------------------|
| GPU | NVIDIA Tesla V100 SXM2 32GB |
| 编程语言 | Python 3.8 |
| 深度学习框架 | Pytorch 1.6.0 |
| GPU 加速库 | CUDA 10.2 |
| 显卡驱动版本 | NVIDA 450.80 |

图像的火灾数据集,由北亚利桑那大学等单位的学者创建,该数据集是通过无人机在亚利桑那州松树林的规定燃烧堆积杂物期间收集的,旨在促进森林火灾的监控和报警,可以帮助研究人员轻松建立模型.

为了进行训练、验证和测试,将数据集按照 8:1:1 的比例进行划分.训练集包含 4400 张图片,这个子集用于模型的训练阶段,通过对数据集进行迭代学习,模型可以学会识别和区分火灾和烟雾的特征.验证集包含 550 张图片,这个子集用于在训练过程中评估模型的性能和进行超参数的调整.通过对验证集的准确度和损失进行监控,可以及时发现模型的过拟合或欠拟合问题,并进行相应的调整和改进.测试集也包含 550 张图片,这个子集用于最终评估模型的性能和泛化能力.在模型训练完成后,使用测试集进行评估可以提供对模型在未见过的数据上的表现的客观度量.在实验中,在满足计算资源限制的前提下选择了 640×640 的图像尺寸.通过经验和实验,确定了 600 次的训练轮数,以在保证模型充分训练的同时避免过拟合.

2.2 实验指标

实验采用了平均精度均值 (mAP) 和帧率 (FPS) 作为评估指标.准确率、精度、召回率、平均精度均值 (mAP) 和帧率 (FPS) 的计算公式如下(8)-(13).平均精度 (AP): AP 是汇总了在每个阈值下达到的精度的加权平均值,其中从上一个阈值到当前阈值的召回率增加作为权重.AP 的计算公式为:

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FN} + \text{FP}}, \quad (8)$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}, \quad (9)$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}, \quad (10)$$

$$\text{AP} = \sum_n (R_n - R_{n-1}), \quad (11)$$

式中, P_n 和 R_n 分别是第 n 个阈值处的精度和召回率.

平均精度均值 (mAP),是所有类别的平均分类性能的平均值,计算公式为:

$$mAP = \frac{1}{n} \sum_{i=1}^n (AP)_i, \quad (12)$$

式中: $(AP)_i$ 表示第 i 类的 AP 值, 当预测框与真实框的重合度为 0.5 时, 表示为 $mAP@0.5$. 本实验中的 $mAP@0.5$ 涉及两个二分类问题的叠加, 因此在推理过程中将 n 值取为 2 进行计算.

帧率(FPS): FPS 通常是连续图像(帧)被捕获或显示的频率, FPS 的一种常见计算方法是使用固定间隔法, 其公式为:

$$FPS = \frac{\text{frameNum}}{\text{elapsedTime}}, \quad (13)$$

式中: frameNum 是帧数, elapsedTime 是经过的时间.

2.3 消融实验

为了验证基于轻量化特征提取网络、SENet 注意力机制和 α -IoU 损失函数的改进对 YOLOv8 模型的有效性, 进行了一系列消融实验. 实验结果如下表所示, 其中粗体表示最优值.

表 2 不同改进策略对 YOLOv8n 的影响

Tab. 2 Effects of different improvement strategies on YOLOv8n

| 模型 | ③ | ② | ① | mAP@0.5/% | Precision/% | Size/MB | FPS |
|---------|---|---|---|-------------|-------------|------------|------------|
| YOLOv8n | — | — | — | 86.8 | 91.2 | 8.5 | 278 |
| ① | √ | — | — | 84.3 | 80.3 | 6.3 | 302 |
| ② | √ | √ | — | 86.4 | 87.3 | 7.1 | 293 |
| ③ | √ | √ | √ | 87.2 | 89.2 | 7.1 | 295 |

注: “—”表示在模型中未使用该方法; “√”表示在模型中使用了该方法; ①代表 EfficientNet-CSPDarknet; ②代表 SENet 注意力机制; ③代表 α -IoU 损失函数.

原始 YOLOv8 模型含有大量的参数和计算量, 不适合部署在内存要求严格的嵌入式设备上. 改进点 1 引入了轻量化特征提取网络, 采用 EfficientNet 模块来减少模型的内存占用率, 这导致网络的特征提取能力稍有降低, 从而使得 $mAP@0.5$ 的值有所降低. 但是, 改进点 1 的模型大小显著减小, 减少了 1.4 MB, FPS 提升了 24, 证明了 EfficientNet 在网络压缩方面的有效性.

改进点 2 在改进点 1 的基础上引入了 SENet 注意力机制, 增强了网络对空间和通道的关注. 与改进点 1 相比, 改进点 2 的 $mAP@0.5$ 提高了 2.1%, 模型大小增加了 0.8 MB, FPS 下降了 9. 这说明 SENet 注意力机制对模型的性能的提升起到了积极作用.

改进点 3 在改进点 2 的基础上应用了 α -IoU 损失函数和数据增强. 相对于改进点 2, 改进点 3 的 $mAP@0.5$ 进一步提高了 0.8%, 这主要归功于数据增强算法扩增了训练样本, 提升了模型的鲁棒性. 同时, 改进点 3 的模型大小和浮点运算量与改进点 2

相比几乎没有变化.

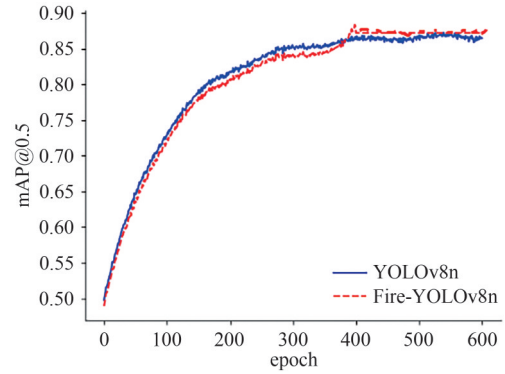


图 7 模型改进前和改进后的 $mAP@0.5$ 对比情况

Fig. 7 Comparison of $mAP@0.5$ before and after model improvement

综上所述, 基于轻量化特征提取网络、SENet 注意力机制和 α -IoU 损失函数的改进, 使得 YOLOv8 模型在保持较高 $mAP@0.5$ 的同时, 显著减小了模型大小, 加快了检测速度. 这使得改进后的 YOLOv8 模型更适合在算力有限、功耗较低的嵌入式设备上运行.

2.4 实验结果

基于轻量化特征提取网络、SENet 注意力机制和 α -IoU 损失函数的创新, 在 YOLOv8 模型上进行了森林火灾检测实验. 使用了包含森林火灾图像及其相应标注框的数据集进行训练和测试. 实验结果如下.

表 3 实验结果

Tab. 3 Experimental result

| 模型 | mAP@0.5/% | Precision/% | Size/MB | FPS |
|-------------|-----------|-------------|---------|------|
| SSD | 77.2 | 79.3 | 108.1 | 10.8 |
| YOLOv5n | 81.3 | 86.6 | 18.2 | 62.5 |
| YOLOv7-tiny | 83.6 | 88.3 | 12.3 | 83 |
| YOLOv8 原模型 | 86.8 | 91.2 | 8.5 | 278 |
| SRN-YOLO | 88.1 | 91.8 | 9.2 | 256 |
| Ours | 87.2 | 89.2 | 7.1 | 295 |

原始的 YOLOv8 模型在森林火灾检测任务上达到了 86.8% 的 $mAP@0.5$. 然而, 其模型大小为 8.5 MB, FPS 为 278, 这不太适合在资源受限的环境中部署. 参考论文中改进的 SRN-YOLO 模型在本文相同的数据集上, 虽然在原始 YOLOv8 的基础上精度有所上升, 但检测速度显著下降. 应用了创新点后, 改进版 YOLOv8 模型在相同数据集上取得了 87.2% 的 $mAP@0.5$, 相较于原始模型略有提升. 同时, 模型大小减少至 1.4 MB, FPS 增加了 17. 这表明改进使得模型更轻量化、高效, 更适合在计算资源有限的嵌入式设备上运行森林火灾检测.

综上所述, 基于轻量化特征提取网络、SENet 注意



图8 森林火灾检测图

Fig. 8 Forest fire detection map

力机制和 α -IoU损失函数的创新,对YOLOv8模型在森林火灾检测任务上进行了有效改进.改进后的模型在保持高mAP@0.5的同时,显著减少了模型大小,大幅提升了检测速度,提高了模型的适用性和效率.

3 结语

本文针对森林火灾检测任务,基于YOLOv8模型进行了三个方面的改进:轻量化特征提取网络优化主干网络、引入SENet注意力机制和采用 α -IoU损失函数.通过消融实验和性能评估,验证了这些创新点的有效性.引入了轻量化特征提取网络,采用EfficientNet优化了特征提取模块.实验结果表明:该轻量化模块成功减小了模型的大小,降低了1.4 MB,检测速度FPS提升了24.这证明了EfficientNet在网络压缩方面的有效性,使得模型更适合在资源受限的嵌入式设备上部署.随后,引入了SENet注意力机制,增强了网络对空间和通道的关注度.实验结果显示:SENet注意力机制显著提升了模型的识别准确率,尽管模型大小略有增加,表明SENet注意力机制对于提升模型性能具有积极作用.采用了 α -IoU损失函数和数据增强,进一步提升了模型性能.实验结果表明:在不增加模型计算量的前提下, α -IoU损失函数和数据增强使mAP@0.5相较于基准模型提高了0.8%.这主要归功于数据增强算法扩充了训练样本,提升了模型的鲁棒性.

综上所述,基于轻量化特征提取网络、SENet注意力机制和 α -IoU损失函数的创新成功改进了YOLOv8模型在森林火灾检测任务上的性能.改进

后的模型在保持较高的mAP@0.5的同时,显著减少了模型大小、提升了检测速度,提高了模型的适用性和效率.这些创新点为嵌入式设备上的森林火灾检测应用提供了有力支持.未来的研究可以进一步优化这些创新点,以提升模型性能,并探索其他创新点来进一步改进火灾检测系统的性能和准确性.

参考文献

- [1] REDMON J. YOLOv3: An incremental improvement [J]. arXiv Preprint arxiv:1804.02767, 2018.
- [2] 李萌娜. 基于深度学习的无人机森林火灾图像识别与分割研究[D]. 西安: 西安理工大学, 2022.
- [3] BOCHKOVSKIY A, WANG C Y, LIAO H Y M. YOLOv4: Optimal speed and accuracy of object detection [J]. arXiv Preprint arxiv: 2004.10934, 2020.
- [4] WANG C Y, BOCHKOVSKIY A, LIAO H Y M. Scaled-YOLOv4: Scaling cross stage partial network [C]//IEEE/CVF Conference on Computer Vision and Pattern Recognition. Virtual: IEEE, 2021: 13029-13038.
- [5] LIU W, ANGUELOV D, ERHAN D, et al. SSD: Single shot MultiBox detector [M]//Lecture Notes in Computer Science. Cham: Springer International Publishing, 2016: 21-37.
- [6] LIU L, OUYANG W, WANG X, et al. Deep learning for generic object detection: A survey [J]. International Journal of Computer Vision, 2020, 128(2): 261-318.
- [7] LI X, CHEN C, HU X, YANG J, ZHANG X. Fire detection in videos using YOLO-based deep learning [J]. Applied Sciences, 2020, 10(5): 1599.
- [8] GARG N, YADAV N. Fire detection in videos using YOLO object detection and optical flow analysis [J]. IEEE

- Access, 2021, 9: 2560-2571.
- [9] SHAO L, SONG Z, WANG L. Fire detection based on YOLOv3 and deep residual attention network [J]. IEEE Access, 2020, 8: 165190-165201.
- [10] LI W, ZUO W. Fast fire detection in video based on improved YOLOv4 and deep learning [J]. IEEE Access, 2020, 8: 203973-203982.
- [11] CAO J, LI X, WANG X, YANG J, ZHANG X. An efficient fire detection method based on YOLOv8 and channel attention mechanism [J]. IEEE Access, 2023, 9: 20788-20796.
- [12] ZHANG X, CAO J, YANG J, ZHOU X. Fire detection in videos using improved YOLOv8 and spatial-temporal features [J]. Sensors, 2021, 21(4): 1246.
- [13] 苏盈盈, 何亚平, 喻骏, 等. 基于改进YOLOv4的轻量化目标检测方法[J]. 中南民族大学学报(自然科学版), 2024, 43(3): 370-377.
- [14] GENG Y, LI J, ZHANG T, ZHAO Y. Fire detection in videos using YOLOv8 and improved deep learning [J]. IEEE Access, 2022, 9: 120033-120042.
- [15] 王晨灿, 李明. 基于YOLOv8的火灾烟雾检测算法研究[J]. 北京联合大学学报, 2023, 37(5): 69-77.
- [16] HUANG Z, LI X, XU Y, ZHANG X, CHEN C. Fire detection based on deep learning with spatial pyramid pooling [J]. IEEE Access, 2018, 6: 12482-12492.
- [17] HUANG Z, LI X, XU Y, ZHANG X. Fire detection in video surveillance using a three-dimensional convolutional neural network [J]. Sensors, 2018, 18(3): 787-789.
- [18] CHEN P, HUANG C, LIN Y. Fire detection using multi-task learning and deep convolutional neural network [J]. Sensors, 2019, 19(19): 4259.
- [19] SHAFIEE M J, YAGHMAEE F. Fire detection using deep learning: A review [J]. Fire Safety Journal, 2019, 107: 27-51.
- [20] GARG N, YADAV N. Early fire detection using YOLO-based object detection and deep belief network [J]. IEEE Access, 2020, 8: 147353-147363.
- [21] RANA S, KUMAR A. Real-time fire detection and monitoring system using deep learning [C]// 2020 4th International Conference on Computing Methodologies and Communication (ICCMC). Erode: IEEE, 2020: 207-212.
- [22] LIU X, ZHANG Y, WANG H, et al. α -IoU: An adaptive IoU loss function for enhanced object detection accuracy [C]// IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). New Orleans: IEEE, 2022: 5123-5131.

(责编&校对 刘钊)