

DOI:10.19651/j.cnki.emt.2517944

融合多尺度注意力的轻量化初期林火检测算法^{*}徐瑞杰¹ 谢辉¹ 姜吴瑾² 李洪兵² 肖扬¹

(1.重庆三峡学院电子与信息工程学院 重庆 404100;2.重庆三峡学院计算机科学与工程学院 重庆 404100)

摘要: 针对森林火灾初期检测领域中环境背景复杂、小目标火焰和烟雾纹理特征不明显、模型参数量大、实际部署计算资源受限的问题,提出基于改进 YOLOv5s 的轻量化森林火灾初期检测算法 YOLO-VRG。首先,使用极简主义网络 VanillaNet 作为特征提取网络,以显著降低模型复杂度,实现高效特征提取;其次,设计了空间特征和特征通道重建注意力卷积 RVBC3EMA 模块,以减少空间和通道维度上的特征冗余,提高特征表达能力;最后,充分利用分组混洗卷积以进一步降低模型参数量和计算量。实验结果表明:YOLO-VRG 模型相比于原始模型,mAP@0.5 达到了 87.6%,提升了 3.2%的同时,模型参数量减少了 74.1%,GFLOPs 降低了 71.9%,仅有 2.1 M 的参数量和 4.5 的 GFLOPs,该算法在提高检测精度的同时大幅度减少了模型参数量,能更好的适用于计算资源受限的场景。

关键词: 火灾初期检测;YOLOv5s;轻量化;EMA 注意力;目标检测

中图分类号: TP391.4;TN911.73 **文献标识码:** A **国家标准学科分类代码:** 510.4050

Lightweight early-stage forest fire detection algorithm integrating multi-scale attention

Xu Ruijie¹ Xie Hui¹ Jiang Wujin² Li Hongbing² Xiao Yang¹

(1. School of Electronic and Information Engineering, Chongqing Three Gorges University, Chongqing 404100, China;

2. School of Computer Science and Engineering, Chongqing Three Gorges University, Chongqing 404100, China)

Abstract: To address challenges in early forest fire detection—including complex environmental backgrounds, indistinct texture features of small flame/smoke targets, and high computational demands in resource-constrained deployments—we propose YOLO-VRG, a lightweight detection algorithm based on improved YOLOv5s. First, we employ VanillaNet as the feature extraction backbone to significantly reduce model complexity while maintaining efficient feature capture. Second, we design the RVBC3EMA module with spatial-channel reconstruction attention to minimize feature redundancy and enhance discriminative representation. Third, we implement grouped shuffle convolution to further optimize parameter efficiency. Experimental results demonstrate that YOLO-VRG achieves 87.6% mAP@0.5 (3.2% improvement over baseline) with only 2.1 M parameters (74.1% reduction) and 4.5 GFLOPs (71.9% reduction). This balanced architecture enables superior detection accuracy and hardware efficiency for edge deployment scenarios.

Keywords: early fire detection; YOLOv5s; lightweight; EMA attention; object detection

0 引言

森林作为地球陆地生态系统的核心组成部分,在维持全球碳循环平衡和 2060 前中国实现碳中和目标中具有不可替代的生态价值^[1]。然而,森林火灾作为最具破坏性的自然灾害之一,其突发性和蔓延性对生态系统和人类安全构成重大威胁,这使得早期精准检测技术成为防灾减灾的

关键环节^[2]。

森林火灾早期的监测技术经历了从传统方法到智能化手段的重要演变。传统的监测手段主要依赖人工巡检、瞭望塔观测和传感器系统,这些方法存在响应滞后、成本高昂、覆盖范围受限等显著缺陷^[3]。随着技术进步,基于无人机和计算机视觉的智能监测方案逐渐成为主流^[4]。计算机视觉和无人机技术相结合的森林火灾智能监测,不仅能够

收稿日期:2025-01-16

^{*} 基金项目:重庆市自然科学基金(2022NSCQ-MSX4084)、重庆市教委科学技术研究项目(KJZD-M202201204, KJZD-M202301203)、重庆市教委科学技术研究项目(KJQN202401237)资助

实现全天候、大范围的火情监控,而且可以为消防人员提供更全面的现场态势信息,有效降低了人力物力消耗。与传统计算机视觉方法相比,基于深度学习的视觉检测方法具有自动特征提取能力,能够捕捉火灾更深层次的特征,有效避免了人工特征选择的复杂性和局限性。因此,基于深度学习的目标检测网络被越来越多的应用于森林火灾检测任务当中^[5-6]。在深度学习目标检测领域,检测算法主要分为两类^[7]:一类是基于区域提取的两阶段检测策略;一类是基于目标回归的单级检测策略。

两阶段检测算法以 R-CNN 系列为代表,其检测过程包括区域建议生成和目标分类定位两个阶段。戴颖成等^[8]利用红外影像的成像特点结合 Faster R-CNN 网络对火灾特征进行放大,提高检测精度,但模型参数量较大,检测速度慢。Gagliardi 等^[9]使用基于快速区域的卷积神经网络(region-based convolutional neural network, R-CNN)来基于可疑火灾区域和非火灾区域的空间特征来检测,通过减少错误检测成功地提高火灾检测精度,然而检测速度相对较慢。两阶段检测算法虽具有较高的检测精度,但模型参数量多,计算资源消耗大,难以满足实时性要求^[10]。

单阶段检测算法以 YOLO 系列为代表,通过直接回归方式预测目标位置和类别,省略了区域建议步骤,显著提升了检测速度,但在检测精度上略有损失^[11]。在森林火灾检测中,考虑到实时检测的重要性,单阶段检测算法因其较快的推理速度和较好的精度平衡而受到广泛关注。吴福明等^[12]提出一种自适应样本均衡林火数据增强方法,该方法在 SSD、YOLOv3、YOLOv4 林火检测中引入自适应参数解决了样本分布不均问题,但检测精度相对较低,无法实现小目标,遮挡目标的精准检测。赵松怀等^[13]通过引入 GhostConv 对 YOLOv5 进行轻量化改进,设计了 C3Ghost 模块以提高烟火检测性能,但检测精度较低。Lin 等^[14]将自关注机制的 Transformer 编码器与 CNN 相结合作为 YOLOv5 特征提取网络,增强对森林火灾目标全局信息的提取,但模型计算量也随之增加。曹云刚等^[15]在 YOLOv7 上引入动态非单调聚焦机制的边界框定位损失函数,提高了对小目标火点的检测能力,但模型参数量较大,不适用于计算资源受限设备的部署。马跃等^[16]在 YOLOv8 特征融合部分采用跨尺度特征融合模块 CCFM,减少了模型参数的同时增强了对火焰尺度变化的适应性,但研究只针对火焰的检测,缺少对烟雾的检测。Lei 等^[17]通过采用 DWConv 和 GhostConv 构建了 DSDetect 和 GhostC2f 模块替换 YOLOv8 的 C2f 模块和检测头,提高了模型检测精度,但检测背景单一,缺少火焰和烟雾的协同检测。He 等^[18]提出了 FireNet 检测网络,集成了简化的视觉变换器(revisiting mobile cnn from vit perspective, RepViT)来增强全局特征学习,同时减少计算开销,但缺少对火灾初期小目标火焰和烟雾的检测。董润华等^[19]提出一个基于 RT-DETR 改进的火灾检测算法 FDH-DETR,通过 HiLo-AIFI

高低频尺度内特征交互模块,增强对高低频特征的提取能力,但算法计算量和参数量大,推理成本较高。

综上所述,深度学习方法在森林火灾检测领域展现出良好的应用前景,但在森林火灾初期复杂环境下,小目标火焰和遮挡火焰极易出现漏检、检测精度低的问题;烟雾和火焰协同检测研究不足;并且还要考虑模型参数规模,在边缘计算等资源受限环境下的实际部署问题都导致森林火灾早期预警效果不佳。针对以上问题,YOLOv5s 检测网络较好的平衡了检测精度和模型体积,本文以 YOLOv5s 网络为基础,提出了 YOLO-VRG 轻量化森林火灾初期检测算法。使用轻量化主干网络 VanillaNet 进行森林火灾特征提取,以减少模型参数量;通过 RepViT Block 可视化模块和跨空间学习的高效多尺度注意力(efficient multi-scale attention, EMA),重新设计 C3 结构为 RVBC3EMA,优化特征提取和融合,专注提高准确性;同时使用分组混洗卷积(grouped spatial convolution, GSConv)替换特征融合部分用于下采样的两个标准卷积,分组混洗卷积的特殊结构设计更能在资源受限的环境下提高模型的计算效率和推理速度。

1 YOLOv5s 模型结构

YOLOv5s 作为 YOLO 系列的重要迭代版本,是一个高效的单阶段目标检测框架。该网络采用了模块化设计思想,主要由输入层(Input)、主干网络(Backbone)、特征融合网络(Neck)和检测头(Head)4 个核心组件构成。其主干网络采用改进的 CSPDarknet 作为特征提取器;在特征融合层面,采用双向特征融合结构 PANet;检测头部分采用非解耦结构,通过 3 个不共享权重的卷积层对特征图进行变换,输出不同尺度的预测结果。YOLOv5s 的网络结构如图 1 所示。本文所提轻量化 YOLO-VRG 结构如图 2 所示。

2 YOLO-VRG 林火初期检测模型

2.1 VanillaNet 轻量化主干网络

为实现轻量级网络设计,借鉴 AlexNet^[20]和 VGG^[21]等传统深度网络的结构,采用一种极简主义的架构思路 VanillaNet^[22]。将该网络架构整合到 YOLOv5s 中作为主干网络,通过纯粹的卷积运算和池化操作,摒弃传统网络中常见的复杂连接机制,大幅降低网络的运算开销和存储需求,以优化模型的整体结构。其网络结构如图 3 所示。

从架构设计来看, VanillaNet 主要由 3 个核心组件构成:主干(Stem)、主体(Main Body)和全连接层(fully connected layer)。在主干部分,采用 $4 \times 4 \times 3 \times C$ 的卷积层进行特征变换,将输入的 3 通道图像通过下采样的方式映射至含有 C 个通道特征图。主体部分由多个阶段组成,每个阶段使用步长为 2 的最大池化层进行下采样操作,同时将特征通道数扩充为上一层的 2 倍。值得注意的是,第 5 层卷积层不增加通道数,并使用平均池化层。最后一层

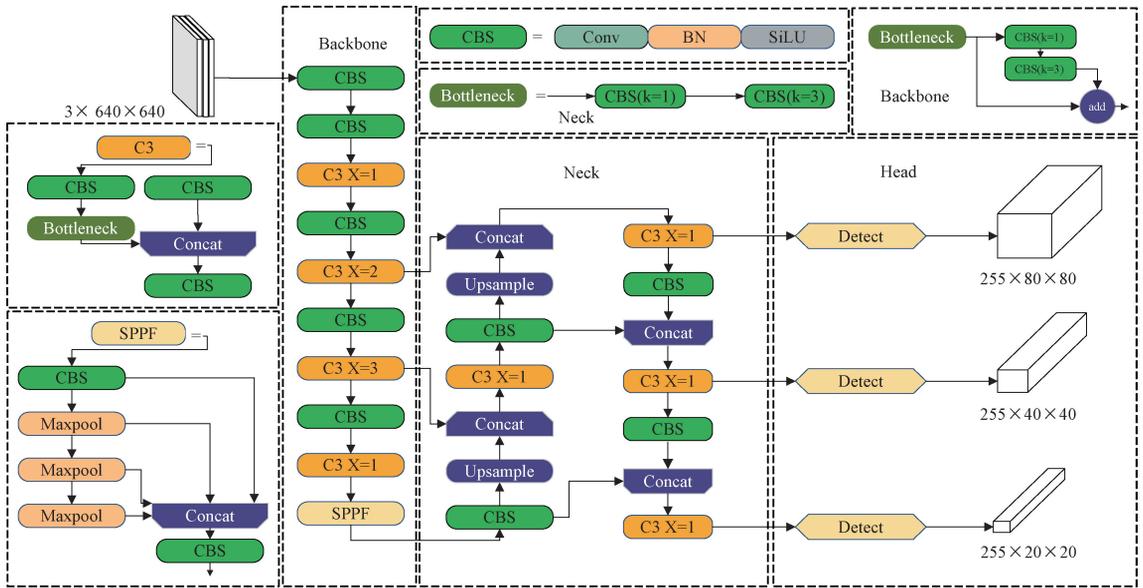


图 1 YOLOv5s 网络结构

Fig. 1 YOLOv5s network structure

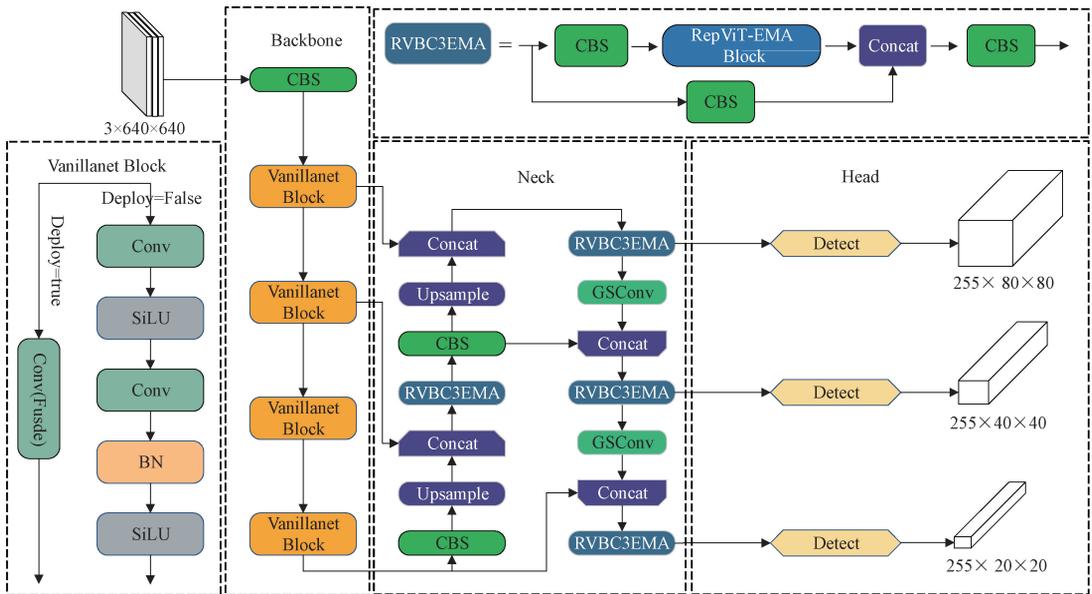


图 2 YOLO-VRG 网络结构

Fig. 2 YOLO-VRG network structure

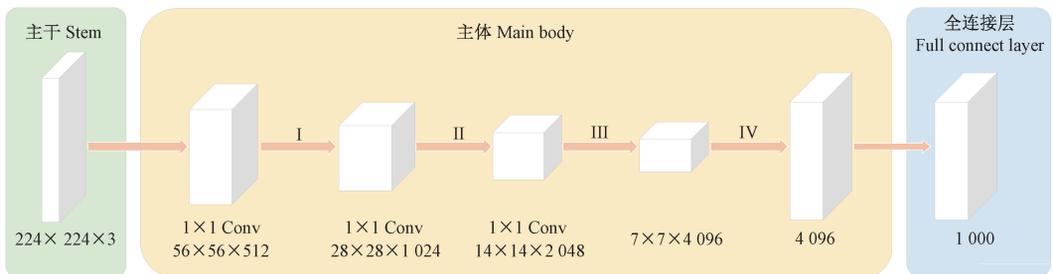


图 3 VanillaNet 网络结构

Fig. 3 VanillaNet network structure

为全连接层,用于输出对图像进行分类的结果。为了最小化计算开销并保持特征信息的完整性,网络采用 1×1 卷积核进行特征处理,并在每个卷积层后引入批量归一化(batch normalization, BN)层和激活函数,不仅简化了网络训练过程,还提高了模型的收敛效率。

为解决网络深度减少导致的非线性拟合能力下降问题,本文使用一种深度训练策略:在训练初期使用两个带 $SiLU$ 激活函数的堆叠卷积层,随训练进行激活函数逐渐简化为恒等映射,最终将两个卷积层合并为单层。这种策略有效地平衡了网络简化与性能之间的权衡, $ReLU$ 激活函数与 $SiLU$ 激活函数的表达式为:

$$ReLU = \begin{cases} x, & x \geq 0 \\ 0, & x < 0 \end{cases} = \max(0, x) \quad (1)$$

$$SiLU = \frac{x}{1 + e^{-x}} \quad (2)$$

通过式(1)和(2)看出, $ReLU$ 函数在输入为负值时输出恒为零,这一特性导致网络中的部分神经元出现“死区”现象,无法有效参与参数更新过程,从而影响模型的特征提取能力。与之相比, $SiLU$ 函数在零值附近表现出更为平滑的过渡特性,并通过引入 $Sigmoid$ 函数将输出值规范化到 $[0, 1]$ 区间内。这种特性使得 $SiLU$ 函数能够更好地保持输入信息的完整性,更适合本文中对信息保留要求较高的应用场景。

2.2 RVBC3EMA 模块

为解决现有 YOLOv5s 网络中, C3 结构存在特征冗余、全局特征捕获能力不足、尤其是在处理小目标和复杂背景时表现欠佳的问题,本文使用高效多尺度注意力 EMA^[23] 替换 RepViT Block^[24] 中的 SE 模块,设计了一个增强的 Transformer^[25] 可视化模块 RepViT-EMA Block。再将 RepViT-EMA Block 替换原 C3 模块中的 Bottleneck 结构构成新型 RVBC3EMA 模块,以提升特征提取效率和检测性能。C3 模块结构如图 4 所示, RVBC3EMA 模块结构如图 5 所示。

RepViT Block 通过重新排列 MobileNetV3^[26] 模块的 3×3 深度可分离卷积(depthwise separable convolution, DWConv)^[27] 并整合成统一分支。MobileNetV3 和 RepViT Block 结构如图 6 所示。

RepViT-EMA block 采用了多层级的特征提取与融合策略。首先,通过 3×3 和 1×1 DWConv 结合残差连接的复合结构,实现了输入特征的多尺度融合,该设计既保证了空间信息的有效捕获,又通过残差路径维持了特征的完整性。其次,对 DW 层引入结构重参数化技术,在训练阶段构建多分支拓扑结构以增强模型的特征学习能力,而在推理阶段将多分支结构合并为单一结构,有效降低了在推理过程中与残差连接相关的计算开销,提升了推理效率。

随后,模块引入 EMA 多尺度注意力机制,通过全局信息编码来自适应调节并行分支的通道权重,并引入跨维度

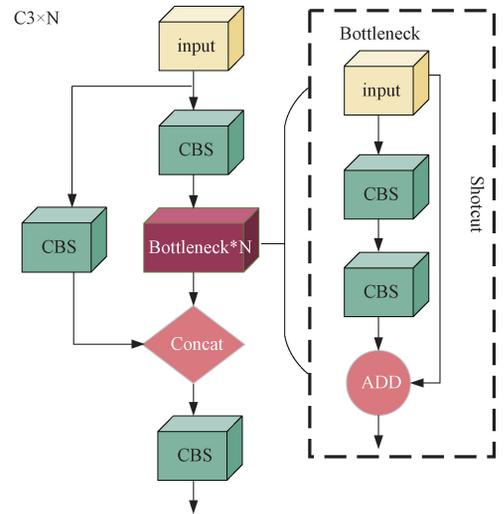


图 4 C3 模块结构

Fig. 4 C3 module structure

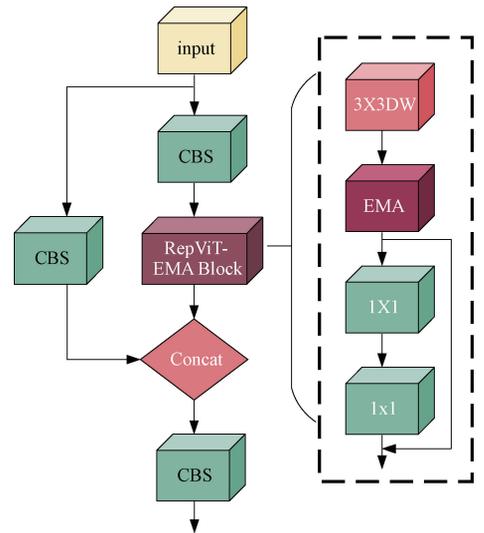


图 5 RVBC3EMA 模块结构

Fig. 5 RVBC3EMA module structure

交互机制实现像素级特征关联的建模。最终,采用 1×1 扩展卷积实现了空间特征混合器(Token Mixer)与特征通道混合器(Channel Mixer)的解耦,这种设计既符合 ViT 的 MetaFormer 架构思想,又保持了卷积运算的高效性。基于该设计, RepViT-EMA block 在提升特征提取效率和多尺度特征融合方面展现出显著优势,有效提高了复杂场景下的目标检测性能。RepViT-EMA block 其结构如图 7 所示。

2.3 跨空间学习的高效多尺度注意力

高效多尺度注意力 EMA 将任意输入的特征图 $X \in R^{c \times h \times w}$, 式中: c 表示输入特征映射的通道数量, h 和 w 分别表示输入特征映射的高度和宽度,划分为 G 个跨通道维度的子特征,来学习不同的特征信息。每组表示为:

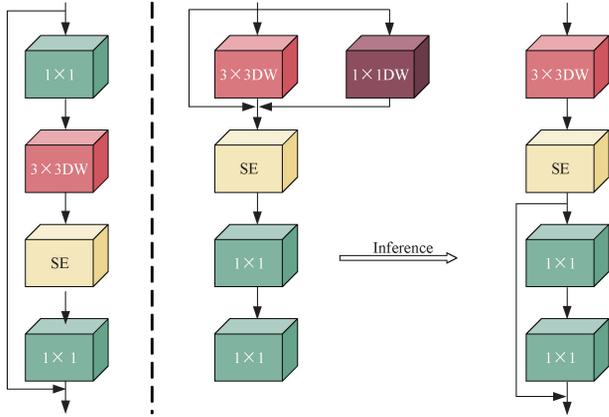


图 6 MobileNetV3 和 RepViT Block 结构
Fig. 6 MobileNetV3 and RepViT Block structure

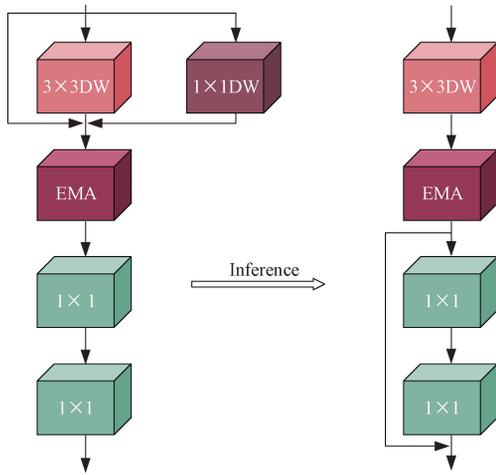


图 7 RepViT-EMA block 结构
Fig. 7 RepViT-EMA block structure

$$X = [X_0, X_i, \dots, X_{G-1}], X_i \in R^{c // c \times h \times w}, G \ll C \quad (3)$$

EMA 模块采用三路并行分支架构提取特征图注意力权重,包含两个 1×1 分支和一个 3×3 分支。在 1×1 分支中,分别沿垂直和水平方向执行一维全局平均池化操作进行通道编码,将编码特征拼接后共享同一 1×1 卷积,输出经分解为两个向量并通过 *Sigmoid* 函数拟合二项分布,实现跨通道交互特征的聚合,而不进行降维。 3×3 分支则通过卷积操作捕获局部跨通道交互以扩展特征空间。在跨空间信息处理阶段,模块对 1×1 分支输出应用二维全局平均池化进行编码,池化操作计算公式为:

$$Z_c = \frac{1}{H \times W} \sum_j \sum_i x_c(i, j) \quad (4)$$

通过非线性变换层对特征进行映射处理,并利用点乘操作将其与三通道卷积分支的输出相结合,构建第 1 层空间关系表征;另一方面,对三通道卷积输出施加全局均值池化操作后,与单通道卷积分支进行点乘运算,生成保持空间精确定位的第 2 层注意力表征。最终将这两个空间维度的权重信息融合,与原始输入特征进行自适应加权,

从而在像素级建立关联并提升特征的全局语义表达能力。其结构如图 8 所示。

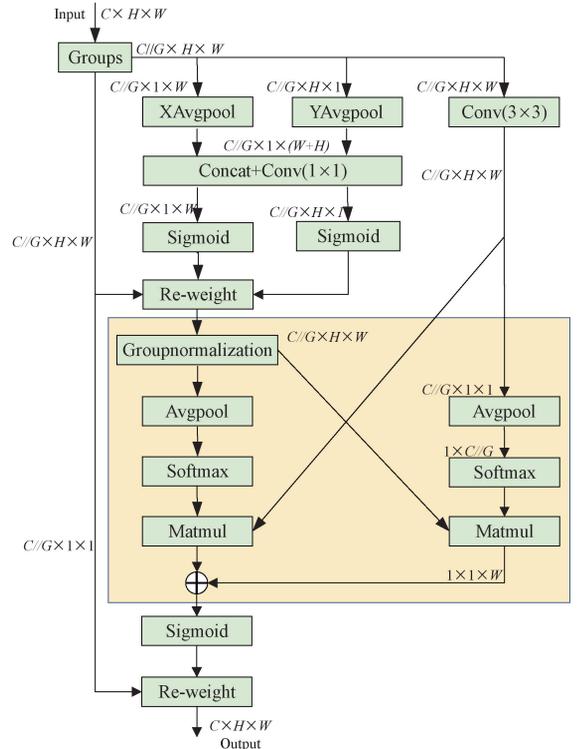


图 8 EMA 模块结构
Fig. 8 EMA module structure

2.4 分组混洗卷积

使用 $GSConv^{[28]}$ 替换特征融合部分用于下采样的两个标准卷积,以平衡特征提取能力和计算效率。通过分析 YOLOv5s 模型的特征分布特性发现,当特征图传递至网络颈部区域时,其通道维度达到峰值而空间维度(高度和宽度)降至最小,此时特征信息的冗余度较低。因此本文选择性地在该区域引入 $GSConv$ 模块替代传统卷积。

$GSConv$ 模块首先将输入的 C_1 通道特征图进行二分处理,形成两个独立的特征子集;其中一个子集经过标准卷积操作输出 $C_2/2$ 通道特征图,另一个子集通过 $DWConv$ 进行处理,该操作对每个输入通道独立进行卷积运算,有效降低了参数量和计算开销。随后,将两路处理得到的特征图进行 Channel-Wise 的拼接,获得 C_2 通道的特征表示。最终通过 Shuffle 操作重新排列通道顺序,促进跨通道信息的有效融合与交互,输出通道数为 C_2 的特征图。其结构如图 9 所示。

标准卷积和 $GSConv$ 模块的计算时间复杂度表示为:

$$T_{SConv} = (W \times H \times K_1 \times K_2 \times C_1 \times C_2) \quad (5)$$

$$T_{GSConv} = [W \times H \times K_1 \times K_2 \times C_2/2 \times (C_1 + 1)] \quad (6)$$

式中: W 和 H 分别表示输出特征图的宽度和高度维度, $K_1 \times K_2$ 代表卷积核的尺寸大小, C_1 和 C_2 则分别对应输

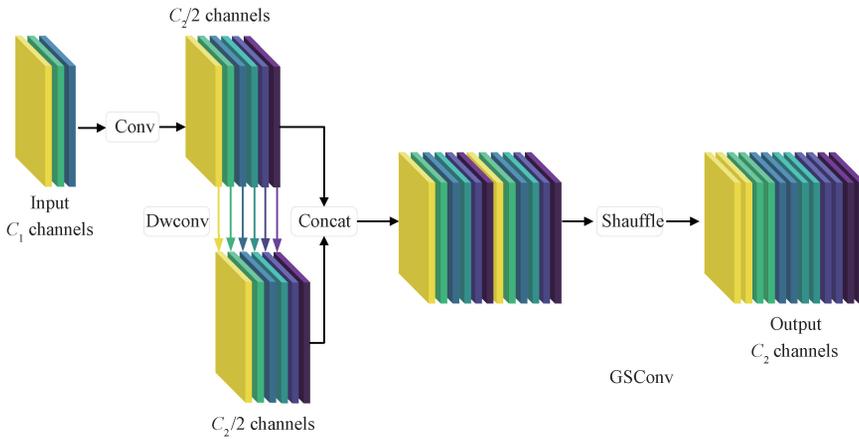


图 9 GSConv 模块结构

Fig. 9 GSConv module structure

入特征图和输出特征图的通道数。

3 实验与结果分析

3.1 实验环境及森林火灾初期数据集

由于目前尚没有用于森林火灾初期检测的公开数据集,而数据集的质量对模型训练结果有着至关重要的作用,因此本实验通过编写爬虫脚本,从 Google 图像搜索中获取了部分森林火灾初期图像。此外,还从公开的火灾数据集中汇总了部分图像,例如 FLAME^[29] 数据集、中国科学技术大学火灾科学国家实验室的烟雾数据集、KMU Fire & Smoke Database^[30] 等共收集了 2 652 张原始图像。为提升模型的泛化能力,采用了 Mosaic 数据增强方法对数据进行了扩充,通过精细化筛查流程,最终筛选出 3 942 幅具备无人机俯瞰视角的早期森林火情图像样本。为了构建模型训练体系,采用科学的数据集划分策略,将图像数据按照 70%、20% 和 10% 的比例分别用于模型训练、验证和性能测试,制作成一个多尺度、多背景、可用于实验的森林火灾初期数据集。

在本研究中,对采集的森林烟火图像样本进行了规范化的边界框标定工作。通过交互式的可视化界面,完成了目标区域的精确定位和类别标记。部分图像样本如图 10 所示,展示了数据集中典型的烟火目标。本文实现模型训练的实验环境及相关参数设置如表 1 所示。

3.2 模型评价指标

实验使用参数量 (Params)、浮点运算次数 (giga floating point operations, GFLOPs) 来综合衡量模型轻量化水平,使用帧率 (frames per second, FPS) 衡量模型的检测速度,选取 IoU 阈值为 0.5 的平均精度均值 mAP@0.5 (mean average precision, mAP), IoU 在 0.5 到 0.95 (其中步长为 0.05) 加权平均后的平均精度即 mAP@0.5 : 0.95 来评估模型的检测准确度。使用 P、R 曲线图和 mAP 图来评估模型的整体性能。其计算公式如下:

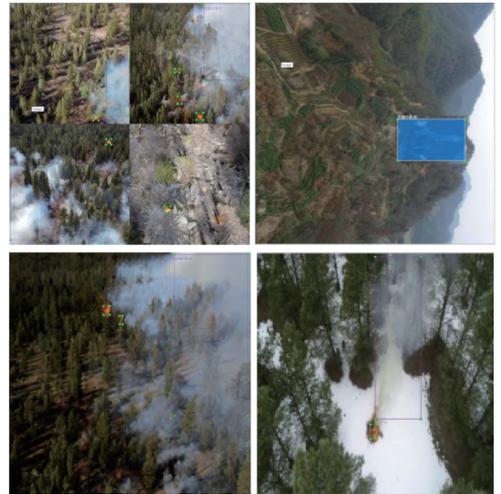


图 10 森林火灾初期数据集中的部分图像样本

Fig. 10 Partial image samples from the initial forest fire datas

表 1 实验环境及参数配置

Table 1 Experimental environment and parameter configuration

名称	参数
操作系统	windows 10
CPU	AMD Ryzen 9 3950X
GPU	NVIDIA GeForce RTX 3070
显存	12 G
编程语言	python3.8
深度学习框架	Pytorch 2.2.2+CUDA 12.1
优化函数	SGD
训练轮数	300
初始学习率	0.01
批量	16
输入图像尺寸	640×640

$$P = \frac{TP}{TP + FN} \quad (7)$$

$$R = \frac{TP}{TP + FN} \quad (8)$$

$$F_1 = 2 \times \frac{P \times R}{P + R} \quad (9)$$

$$A_{AP} = \int_0^1 P dR \quad (10)$$

$$m_{mAP} = \frac{1}{N} \sum_{i=1}^N A_{AP_i} \quad (11)$$

式中: TP (true positives) 表示正确识别森林火灾目标的数

量; FP (false positives) 表示错误识别森林火灾目标的数量; FN (false negative) 表示未识别出森林火灾目标的数量。如果林火目标的真实框和预测框的 $IoU \geq 0.5$ 则检测框被标记为 TP , 反之记为 FP 。 A_{AP} (average precision, AP) 是 P-R 曲线与坐标轴围成的面积。

3.3 消融实验

为验证本实验所提改进算法的性能, 在森林火灾初期数据集设计了 8 组消融实验。通过这些实验的比较, 系统的分析各模型不同结构的表现差异。消融实验检测结果如表 2 所示, 检测结果对比图如图 11 所示。

表 2 森林火灾初期数据集上消融对比实验

Table 2 Comparison experiment of ablation on forest fire initial dataset

模型	VanillaNet	RVBC3EMA	GSConv	mAP@0.5/%	mAP@0.5:0.95/%	参数量/M	GFLOPs/G	帧率
YOLOv5s				84.4	51.4	8.1	16.0	131.7
YOLOv5s-V	✓			85.1	51.7	4.2	8.2	158.1
YOLOv5s-R		✓		86.3	53.5	5.4	11.1	114.7
YOLOv5s-G			✓	83.9	51.1	6.3	14.7	138.8
YOLOv5s-VR	✓	✓		86.8	53.1	2.5	6.1	129.4
YOLOv5s-VG	✓		✓	85.8	52.3	2.8	6.3	149.4
YOLOv5s-RG		✓	✓	86.1	52.6	4.9	10.4	119.4
YOLO-VRG	✓	✓	✓	87.6	53.9	2.1	4.5	133.4



(a) 原始图片 (a) Original image (b) YOLOv5s (c) YOLOv5s-V (d) YOLOv5s-R (e) YOLOv5s-G (f) YOLO-VRG

图 11 消融实验检测结果对比图

Fig. 11 Comparison of ablation experiment detection results

针对森林火灾初期检测, 使用轻量化主干网络 VanillaNet 替换原主干网络后 $mAP@0.5$ 从 0.844 提升至 0.851, 提升了 0.7%, 模型参数量减少了 48.1%, GFLOPs 下降了 7.8 G, FPS 提升了 26.4 帧/秒, 表明 VanillaNet 主干网络能够极大地减少特征提取过程中的重复操作,

节约了计算成本, 帧速率得到了可观的提升; 在颈部采用 RVBC3EMA 模块进行改进时, $mAP@0.5$ 从 0.844 提升至 0.863, 提升了 1.9%, 模型参数减少了 33.3%, GFLOPs 下降了 4.9 G, 表明 RVBC3EMA 模块能够在减少模型参数的同时有效提升了检测精度, 在轻量化和检测精度间达

到了很好的平衡,能大幅度减少特征图像中的冗余;在使用了新的卷积网络 GSCnv 后,模型参数减少了 22.2%,GFLOPs 下降了 1.3 G,FPS 提升了 7.1 fps,表明 GSCnv 卷积在提高模型的检测帧率的同时实现了模型参数的有效降低。综合实验结果,同时采用 VanillaNet 主干网络、RVBC3EMA 模块和 GSCnv 卷积的 YOLO-VRG 的 mAP@0.5 提升了 3.2%,模型参数仅为 2.1 M,减少了 74.1%,GFLOPs 下降了 11.5 G,FPS 提升了 1.7 fps;同时通过图 11 检测结果可以看出,第 1 张图片右下角火灾初期烟雾,第 2 张图片左上角被遮挡的小目标火焰和第 3 张图

片右下角小目标火焰均有更好的检测效果。YOLO-VRG 在有效降低模型参数量的同时提高了检测精度,实现了模型的轻量化,可以有效解决模型在资源受限设备部署难题并降低维护成本。

3.4 不同目标检测算法对比试验

为更全面地评价本文算法 YOLO-VRG 在林火初期检测的优势,将本文算法与一些先进的目标检测算法进行了对比实验,包括 Faster R-CNN、SSD、YOLOv7、YOLOv8s、YOLOv11s、文献[13]、文献[31]、文献[32]。实验结果如表 3 所示。

表 3 森林火灾初期数据集上对比实验

Table 3 Comparative experiments on the initial forest fire dataset

模型	mAP@0.5/%	mAP@0.5:0.95/%	参数量/M	GFLOPs/G	帧率
YOLOv5s	84.4	51.4	8.1	16.0	131.7
Faster R-CNN	80.3	44.4	136.7	401.7	28.4
SSD	75.7	43.0	26.8	62.2	69.5
YOLOv7	83.6	47.8	36.5	103.2	83.3
YOLOv8s	83.7	48.3	11.8	28.6	159.4
YOLOv11s	84.9	53.2	9.4	21.3	164.2
文献[13]	84.7	48.4	5.0	4.8	109.3
文献[31]	85.1	48.6	3.8	8.6	96.7
文献[32]	85.5	50.8	7.0	16.3	105.1
YOLO-VRG	87.6	53.9	2.1	4.5	133.4

与目标检测算法 Faster R-CNN、SSD 相比,本文算法 YOLO-VRG 在检测精度 mAP@0.5 分别提升了 7.3%和 11.9%,mAP@0.5:0.95 分别提升了 9.5%和 10.9%,平均精度均值、模型参数量、浮点运算次数、检测速度方面均表现更好。与 YOLOv5s、YOLOv7、YOLOv8s 算法相比,YOLO-VRG 在保证推理速度的前提下,mAP@0.5 分别提升了 3.2%、4.0%、3.9%;mAP@0.5:0.95/% 分别提升了 2.5%、6.1%、5.6%;检测精度提升的同时,模型参数分别下降了 74.1%、94.2%、82.2%;GFLOPs 分别下降了 11.5、98.7、24.1 G。与最新推出的 YOLOv11s 和改进的先进算法文献[13]、文献[31]、文献[32]相比,YOLO-VRG 的 mAP@0.5 提升了 2.7%、2.9%、2.5%、2.1%;mAP@0.5:0.95/% 提升了 0.7%、5.5%、5.3%、3.1%;模型参数下降了 77.7%、58.0%、44.7%、70.0%;GFLOPs 下降了 16.8、0.3、4.1、11.8 G。由此可得,在以上算法中,本文算法 YOLO-VRG 的 mAP@0.5 最高,达到了 87.6%,模型参数量为 2.1 M,远远小于其他算法模型,且计算量仅为 4.5 GFLOPs,因此该算法具备高准确率和轻量化的特点。同时,算法的检测速度达到了 133.4 fps,能够满足森林火灾初期检测对实时性的需求。通过一系列主流算法的对比实验,进一步验证了本文提出的 YOLO-VRG 算法在提升检测精度的同时能够大幅度减少参数量

和计算复杂度,有效降低了使用成本、提高了检测效率,在实现模型轻量化设计与部署方面更具优势。

3.5 算法检测效果展示

将 YOLO-VRG 与 YOLOv11s、文献[13]、文献[31]、文献[32]在测试集的检测效果进行对比,验证改进算法的有效性。原始 YOLOv5s 算法和改进 YOLO-VRG 算法 P、R 曲线和 mAP 曲线如图 12 所示,不同算法检测结果对比图如图 13 所示。

由图 12 可以看出,改进后的 YOLO-VRG 算法的 P、R 曲线所包围的面积更大,mAP 值更高,表明算法的检测精度更高,性能更好,检测效果有了显著提升。在图 13 第 1 张照片中,YOLOv11s、文献[13]、文献[31]和文献[32]均漏检了照片右下角的第 2 处初期小型烟雾,只有 YOLO-VRG 能够检测出。在图 13 第 2 张照片中存在初期小目标和被遮挡火焰,YOLOv11s、文献[13]、文献[31]和文献[32]都有不同程度的漏检,只有 YOLO-VRG 检测出所有小目标火焰并检测出了左上角被遮挡的小目标火焰。在图 13 第 3 张照片中,也仅有文献[31]、文献[32]和 YOLO-VRG 能够检测出右下角小目标火焰,YOLOv11s 和文献[13]均存在漏检,文献[13]还把右下角的湖泊误检成烟雾;YOLO-VRG 检测出的所有目标置信度得分最高,检测结果更具有可靠性。通过检测结果对比,可以直观地

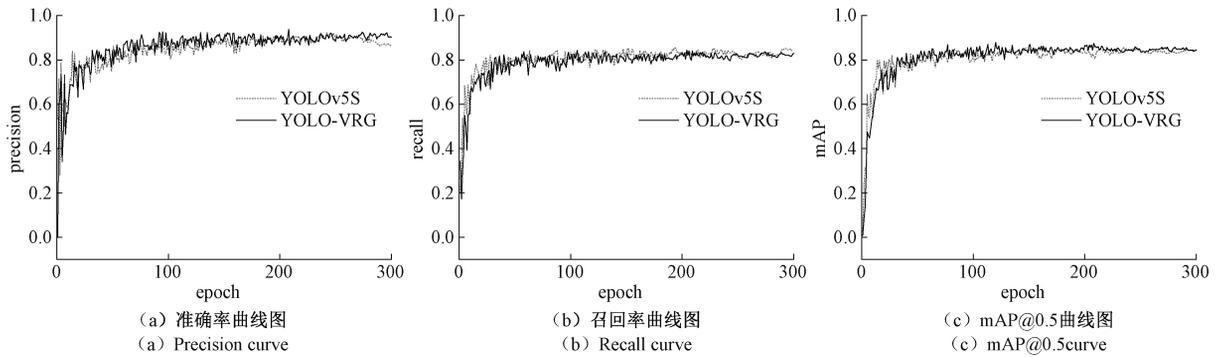


图 12 YOLOv5s 和 YOLO-VRG 算法检测结果 P,R 曲线图和 mAP 曲线图

Fig. 12 The P,R curve and mAP curve of YOLOv5s and YOLO-VRG algorithm detection results

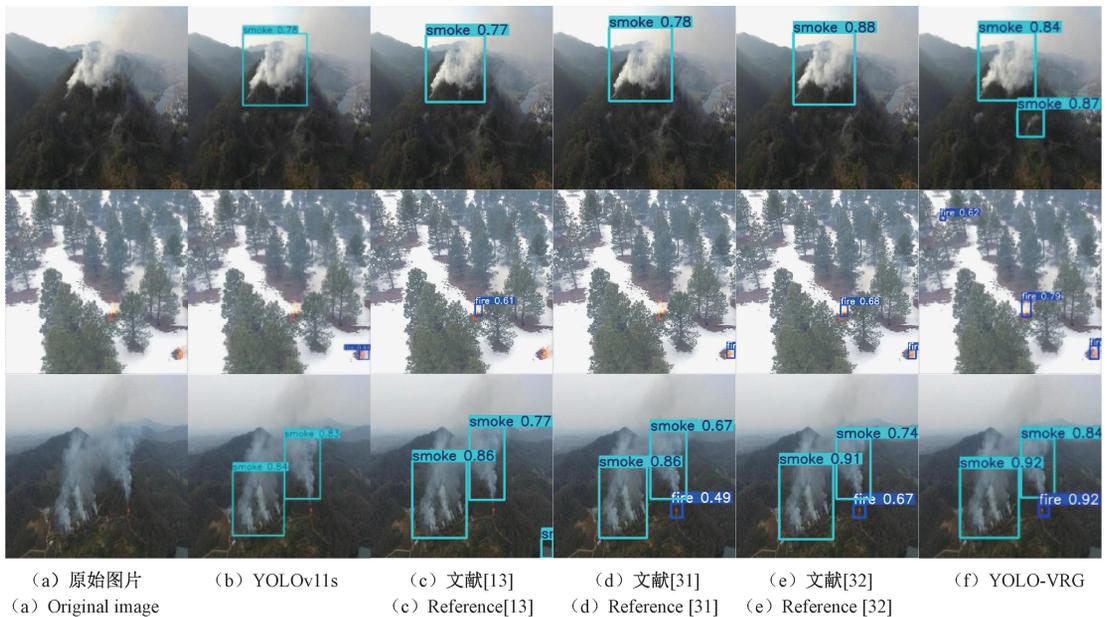


图 13 算法检测效果对比图

Fig. 13 Comparison of algorithm detection results

看出本文提出的 YOLO-VRG 算法表现出更好的检测性能,特别是在处理被遮挡目标和小目标方面。该算法不仅提升了检测准确性,降低了漏检率和误检率,还增强了在复杂场景中的泛化性,能对森林火灾初期小目标火焰和烟雾进行有效检测。

4 结 论

在森林火灾初期检测领域中,由于负责检测的终端设备计算资源有限,导致基于深度学习的森林火灾检测算法部署困难。本文基于 YOLOv5s 目标检测算法进行了一系列创新增强,提出了轻量级森林火灾初期检测算法 YOLO-VRG。首先,在 YOLOv5s 原始模型的基础上使用 VanillaNet 网络替换原主干网络进行特征提取,通过减少特征提取层数、简化操作以及避免复杂的连接方式(如自注意力和残差连接),显著降低了复杂度,优化了资源利

用,实现了高效的特征提取。其次,设计了轻量级模块 RVBC3EMA,将 RepViT Block 与 EMA 注意力机制融入 C3 模块,在减少计算开销的同时增强特征多样性和表达力,提升模型在实际应用中的检测精度和鲁棒性。最后,在模型颈部采用 GSConv 模块替代普通卷积,降低模型参数数量和复杂度,增强网络灵活性。

实验结果表明,相比于 YOLOv5s 基准模型,YOLO-VRG 模型在参数量减少了 74.1%,GFLOPs 降低了 71.9%的同时 mAP@0.5 提升了 3.2%,模型仅有 2.1 M 的参数量和 4.5 的 GFLOPs。通过降低参数量和计算复杂度,实现了模型的轻量化设计,使得模型更能适用于计算资源受限的环境。同时,在复杂林地环境中,模型展现出优异的火灾识别能力,尤其是对于初期小火焰和部分遮挡情况下的火焰目标,均能实现准确识别,有效降低了漏检率。未来,将继续针对 YOLO-VRG 的实际部署进行研

究,提升模型在更复杂场景下的整体检测性能。

参考文献

- [1] 贾敏,叶舒,李明煜,等. 2020—2060年中国森林碳汇及碳移除潜力分析[J]. 环境工程学报, 2024, 18(10): 2677-2685.
JIA M, YE SH, LI M Y, et al. Analysis of forest carbon sinks and carbon removal potential in China from 2020 to 2060 [J]. Journal of Environmental Engineering, 2024, 18(10): 2677-2685.
- [2] WANG S, LI H L, NIU SH K. Empirical research on climate warming risks for forest fires: A case study of grade I forest fire danger zone, Sichuan Province, China[J]. Sustainability, 2021, 13(14): 7773.
- [3] ABDUSALOMOV A B, ISLAM B M D S, NASIMOV R, et al. An improved forest fire detection method based on the detectron2 model and a deep learning approach[J]. Sensors, 2023, 23(3): 1512.
- [4] LI CH D, LI G Y, SONG Y CH, et al. Fast forest fire detection and segmentation application for uav-assisted mobile edge computing system [J]. IEEE Internet of Things Journal, 2023, 11 (16): 26690-26699.
- [5] 宋晔,吴一全. 基于无人机航拍的风力发电机叶片表面缺陷检测综述[J]. 仪器仪表学报, 2024, 45(10): 1-25.
SONG Y, WU Y Q. Review of wind turbine blade surface defect detection based on UAV aerial photography [J]. Chinese Journal of Scientific Instrument, 2024, 45(10): 1-25.
- [6] XIAO H G, LIU SH H, ZUO K, et al. Multiple adverse weather image restoration: A review [J]. Neurocomputing, 2024, 68: 129044.
- [7] 朱弥雪,刘志强,张旭,等. 林火视频烟雾检测算法综述[J]. 计算机工程与应用, 2022, 58(14): 16-26.
ZHU M X, LIU ZH Q, ZHANG X, et al. Review of research on video-based smoke detection algorithms[J]. Computer Engineering and Applications, 2022, 58(14): 16-26.
- [8] 戴颖成,陈知明,刘峰,等. 基于无人机红外影像的森林火灾燃烧点检测方法[J]. 中南林业科技大学学报, 2022, 42(9): 102-114.
DAI Y CH, CHEN ZH M, LIU F, et al. Forest fire detection method based on uav infrared images and deep learning[J]. Journal of Central South University of Forestry & Technology, 2022, 42(9): 102-114.
- [9] GAGLIARDI A, VILLELLA M, PICCIOLINI L, et al. Analysis and design of a YOLO like DNN for smoke/fire detection for low-cost embedded systems [J]. Lecture Notes in Electrical Engineering, 2021, 738: 12-22.
- [10] 孙备,孙晓永,钱翰翔,等. 动态大视角场景融合帧间信息与模板匹配的低慢小无人机目标检测[J]. 仪器仪表学报, 2024, 45(7): 64-74.
SUN B, SUN X Y, QIAN H X, et al. Low speed small unmanned aerial vehicle target detection based on dynamic large angle scene fusion of inter frame information and template matching [J]. Chinese Journal of Scientific Instrument, 2024, 45(7): 64-74.
- [11] ZOU ZH X, CHEN K Y, SHI ZH W, et al. Object detection in 20 years: A survey [J]. IEEE, 2023, 111(3): 257-276.
- [12] 吴福明,宋智豪,王超,等. 基于自适应样本均衡与信息融合的林火检测数据增强方法[J]. 林业科学, 2023, 59(6): 88-101.
WU F M, SONG ZH H, WANG CH, et al. Adaptive sample equalization and information fusion augmentation method for forest fire data [J]. Scientia Silvae Sinicae, 2023, 59(6): 88-101.
- [13] 赵松怀,周敏,申飞,等. 基于改进YOLOv5s的烟火轻型检测算法[J]. 电子测量技术, 2024, 47(17): 140-146.
ZHAO S H, ZHOU M, SHEN F, et al. A lightweight detection algorithm for fireworks based on improved YOLOv5s [J]. Electronic Measurement Technology, 2024, 47(17): 140-146.
- [14] LIN J, LIN H F, WANG F. A semi-supervised method for real-time forest fire detection algorithm based on adaptively spatial feature fusion [J]. Forests, 2023, 14(2): 361.
- [15] 曹云刚,曾雅慧,程海波,等. 基于全维动态卷积与聚焦IoU的多视角森林火点检测方法[J/OL]. 西南交通大学学报, 1-9 [2024-12-11]. <http://kns.cnki.net/kcms/detail/51.1277.U.20240719.0940.002.html>.
CAO Y G, ZENG Y H, CHENG H B, et al. A method for forest fire detection based on omnidimensional dynamic convolution and an IoU with dynamic focusing mechanism for multi-view images [J/OL]. Journal of Southwest Jiaotong University, 1-9 [2024-12-11]. <http://kns.cnki.net/kcms/detail/51.1277.U.20240719.0940.002.html>.
- [16] 马跃,黄周睿,周雯,等. 基于感受野注意力的轻量化林火检测算法[J/OL]. 计算机工程, 1-14 [2024-12-11]. <https://doi.org/10.19678/j.issn.1000-3428.0069454>.
MA Y, HUANG ZH R, ZHOU W, et al. Lightweight forest fire detection algorithm based on receptive field attention [J/OL]. Computer Engineering, 1-14 [2024-12-11]. <https://doi.org/10.19678/j.issn.1000-3428.0069454>.

- [17] LEI L, DUAN R F, YANG F, et al. Low complexity forest fire detection based on improved YOLOv8 network[J]. *Forests*, 2024, 15(9): 1652.
- [18] HE Y H, SAHMA A, HE X, et al. FireNet: A lightweight and efficient multi-scenario fire object detector[J]. *Remote Sensing*, 2024, 16(21): 4112.
- [19] 董润华, 常青, 孔鹏伟, 等. FDH-DETR 工况场景工人行为及火灾检测算法[J]. *电子测量技术*, 2025, 48(3): 145-153.
- DONG R H, CHANG Q, KONG P W, et al. FDH-DETR worker behavior and fire detection algorithm in working condition online first [J]. *Electronic Measurement Technology*, 2025, 48(3): 145-153.
- [20] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. ImageNet classification with deep convolutional neural networks[J]. *Communications of the ACM*, 2017, 60(6): 84-90.
- [21] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition [J]. *ArXiv preprint arXiv:1409.1556*, 2014.
- [22] CHEN H T, WANG Y H, GUO J Y, et al. VanillaNet: The power of minimalism in deep learning[C]. *37th International Conference on Neural Information Processing Systems*, 2023: 7050-7064.
- [23] OUYANG D L, HE S, ZHANG G ZH, et al. Efficient multi-scale attention module with cross-spatial learning [C]. *ICASSP 2023-2023 IEEE International Conference on Acoustics, Speech and Signal Processing*, 2023: 1-5.
- [24] WANG A, CHEN H, LIN Z J, et al. Repvit: Revisiting mobile cnn from vit perspective[C]. *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2024: 15909-15920.
- [25] DOSOVITSKIY A, BEYER L, KOLESNIKOV A, et al. An image is worth 16×16 words: Transformers for image recognition at scale [J]. *ArXiv preprint arXiv:2010.11929*, 2020.
- [26] HOWARD A, SANDLER M, CHU G, et al. Searching for mobilenetv3 [C]. *IEEE/CVF International Conference on Computer Vision*, 2019: 1314-1324.
- [27] ZHANG T X, XU W J, LUO B, et al. Depth-wise convolutions in vision transformers for efficient training on small datasets [J]. *Neurocomputing*, 2024: 128998.
- [28] 杨宇龙, 张银胜, 段修贤, 等. 基于轻量级卷积网络的风力发电机表面缺陷检测研究[J]. *电子测量与仪器学报*, 2024, 38(8): 36-45.
- YANG Y L, ZHANG Y SH, DUAN X X, et al. Research on surface defect detection of wind turbine based on lightweight convolutional network [J]. *Journal of Electronic Measurement and Instrumentation*, 2024, 38(8): 36-45.
- [29] SHAMSOSHOARA A, AFGHAH F, RAZI A, et al. Aerial imagery pile burn detection using deep learning: The FLAME dataset [J]. *Computer Networks*, 2021, 193: 108001.
- [30] KHAN S, MUHAMMAD K, MUMTAZ S, et al. Energy-efficient deep CNN for smoke detection in foggy IoT environment[J]. *IEEE Internet of Things Journal*, 2019, 6(6): 9237-9245.
- [31] 张立国, 张琦, 金梅, 等. 一种改进 YOLOv5s 的森林火灾烟雾检测算法 [J]. *计量学报*, 2024, 45(9): 1314-1323.
- ZHANG L G, ZHANG Q, JIN M, et al. An improved YOLOv5s forest fire smoke detection algorithm[J]. *Acta Metrologica Sinica*, 2024, 45(9): 1314-1323.
- [32] 程飘, 王秋华, 崔荣凯, 等. 基于航拍可见光图像的早期林火探测优化 [J]. *中南林业科技大学学报*, 2025, 45(6): 78-88.
- CHENG P, WANG Q H, CUI R K, et al. Early forest fire detection optimization based on aerial visible light images[J]. *Journal of Central South University of Forestry & Technology*, 2025, 45(6): 78-88.

作者简介

徐瑞杰, 硕士研究生, 主要研究方向为机器视觉、目标检测、深度学习。

E-mail: xrj13996124558@163.com

谢辉(通信作者), 教授, 硕士生导师, 主要研究方向为人工智能、图像处理。

E-mail: 20030041@sanxiau.edu.cn