

基于频空卷积改进YOLOv11的火灾检测算法

殷 波

贵州交通职业大学智能制造学院, 贵州 贵阳

收稿日期: 2026年7月2日; 录用日期: 2026年7月30日; 发布日期: 2026年8月10日

摘 要

本文针对火灾检测任务提出了一种基于频空卷积(Frequency-Spatial Convolution, FSConv)改进的火灾检测算法——FSConv-YOLOv11。通过引入FSConv模块重构了YOLOv11的主干与颈部特征提取网络, 实现了检测精度与计算效率的双重优化。具体而言, FSConv模块利用哈尔离散小波变换(Haar DWT)将特征图无损解耦为高频细节与低频背景, 利用高频特征引导空间注意力机制, 实现了频域与空域特征的深度融合; 同时, 该模块采用无参数的小波变换与深度可分离卷积, 大幅削减了传统标准卷积带来的计算冗余。实验结果表明, 改进后的FSConv-YOLOv11算法检测精度(mAP@0.5)达到了0.632, 较YOLOv11提升了1.2%。另外, 本文所提算法的复杂度大大降低, 具体包括: 模型参数量减少了15.4%, 梯度计算量降低了15.4%, GFLOPs降低了9.2%, 本文所提算法为复杂场景下的火灾智能预警及边缘端部署提供了高效可靠的技术支撑。

关键词

频空卷积, YOLOv11, 火灾检测, 复杂度

Fire Detection Algorithm Based on Improved YOLOv11 with Frequency-Spatial Convolution

Bo Yin

School of Intelligent Manufacturing, Guizhou Communications Polytechnic University, Guiyang Guizhou

Received: July 2, 2026; accepted: July 30, 2026; published: August 10, 2026

Abstract

Aiming at the fire detection task, this paper proposes an improved fire detection algorithm named FSConv-YOLOv11 based on Frequency-Spatial Convolution (FSConv). The FSConv module is embedded

to restructure the backbone and neck feature extraction networks of YOLOv11, realizing dual optimization of detection accuracy and computational efficiency. Specifically, the FSConv module adopts the discrete Haar Wavelet Transform (Haar DWT) to losslessly decouple feature maps into high-frequency detail features and low-frequency background features, and high-frequency features are used to guide the spatial attention mechanism to achieve deep fusion of frequency-domain and spatial-domain features. Meanwhile, the module combines parameter-free wavelet transform with depthwise separable convolution to greatly eliminate computational redundancy caused by traditional standard convolution. Experimental results verify that the improved FSConv-YOLOv11 obtains an mAP@0.5 of 0.632, which is 1.2% higher than the original YOLOv11. Besides, the proposed algorithm achieves obvious reduction in model complexity: the number of model parameters is reduced by 15.4%, the amount of gradient computation drops by 15.4%, and GFLOPs is decreased by 9.2%. The proposed algorithm provides efficient and reliable technical support for intelligent fire early warning and edge-end deployment in complex scenarios.

Keywords

Space-Frequency Convolution, YOLOv11, Fire Detection, Complexity

Copyright © 2026 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

随着现代社会的快速发展与城市化进程的加速,火灾作为一种频发且极具破坏性的灾害,对人民生命财产安全和自然生态环境构成了巨大威胁。因此,建立高效、准确的火灾检测与预警系统是公共安全领域的重要研究课题。传统的火灾检测主要依赖于物理传感器,如感烟报警器、感温探测器和红外传感器等。然而,这些传统设备在实际应用中存在诸多局限性:例如,在火灾初期烟雾浓度不足、或在森林、大型厂房等开阔空间中,传感器往往难以及时触发响应;此外,环境中的水汽、粉尘以及气温的剧烈变化也极易导致传感器产生误报或漏报[1]。

近年来,计算机视觉与深度学习技术的蓬勃发展为火灾检测提供了全新的解决思路。基于深度卷积神经网络(CNN)的目标检测算法,能够直接从监控视频或无人机航拍图像中自动提取特征,实现对火焰和烟雾的快速定位与分类[2]。在众多目标检测算法中,YOLO (You Only Look Once)系列算法凭借其端到端的单次预测架构、高精度的识别能力以及极快的推理速度,成为了实时目标检测领域的标杆[3]。文献[4]提出了一种基于改进 YOLOv5s 的小目标烟雾火焰检测算法,有效提升了早期火灾的识别率。文献[5]针对隧道等低能见度环境,引入注意力机制改进 YOLOv8,增强了模型对复杂背景下火灾特征的提取能力。文献[6]则针对大型机械设备舱内的火灾检测,提出了轻量化的 YOLO-MSA 模型,在保障精度的前提下降低了计算开销。

本文实验所采用的 YOLOv11 基线模型来源于 Ultralytics 官方开源实现,需要说明的是,当前“YOLOv11”更多是工程实现与开源版本中的命名方式,严格意义上其权威来源主要体现在 Ultralytics 官方代码仓库及其模型文档,而非已广泛确立的正式期刊论文[7]。因此,本文将作为实验基线时,统一以 Ultralytics 官方开源版本为准,并在参考文献中给出对应的代码仓库与官方文档链接,以保证实验的可复现性。YOLOv11 通过引入全新的 C3k2 模块、C2PSA 空间注意力机制以及优化的多尺度特征融合策略,在检测精度与推理速度的平衡上达到了新的高度[8]。然而,在实际的火灾检测应用中,直接部署 YOLOv11 仍面临两大挑战:一是火灾场景(尤其是森林野火或复杂城市背景)存在大量的环境噪声,初起阶段的小面积火源极易与背景中的高亮物体混淆,导致检测精度受限;二是 YOLOv11 的核心模块为了

追求极致的特征提取能力，堆叠了大量的标准卷积与残差结构，导致模型参数量庞大、计算复杂度较高，这给算力受限的无人机挂载设备或边缘计算节点带来了极大的部署压力[9]。

针对上述痛点，本文以降低模型复杂度、提升抗环境干扰能力为切入点，提出了一种基于频空卷积(Frequency-Spatial Convolution, FSConv)改进的轻量化火灾检测算法——FSConv-YOLOv11。该算法的核心思想是将信号处理领域的频域分析引入卷积神经网络。由于火焰在视觉上呈现出极高的对比度和锐利的边缘轮廓，属于典型的高频信号，而大部分自然背景则表现为平滑的低频信号[10]。本文设计将原生YOLOv11中的复杂特征提取模块替换为FSConv模块，利用离散小波变换(DWT)从物理频域层面对火焰目标与背景噪声进行有效剥离，并利用高频特征引导空间注意力。该方法不仅从根本上减少了冗余的浮点运算与参数规模，更大幅度提升了模型对复杂背景下火源特征的感知敏感度。本文详细阐述了网络结构的改进策略，并通过丰富的仿真实验验证了该算法在火灾检测任务中的高效性与可靠性。

2. 本文总体方案

2.1. YOLOv11 网络结构分析

YOLOv11 是目标检测领域的有效突破，其网络架构设计在保持实时性的前提下显著增强了特征表征能力[11]。YOLOv11 的整体网络架构如图 1 所示，主要由骨干网络(Backbone)、颈部网络(Neck)和头部网络(Head)三大核心组件构成。在主干网络方面，YOLOv11 放弃了 YOLOv8 中广泛使用的 C2f 模块，全面引入了全新的 C3k2 模块作为特征提取的主力。C3k2 模块采用了双并行的特征流设计：一部分特征直接传递以保留浅层的细粒度空间信息，另一部分特征则通过深度的 Bottleneck (瓶颈)层或 C3k 子模块进行高阶语义特征提取，最后进行通道拼接与融合。此外，主干网络末端集成了 SPPF (空间金字塔池化)和 C2PSA (部分空间注意力)模块，以增强模型对全局上下文信息的捕获能力。颈部网络采用 PANet (路径聚合网络)架构，通过自顶向下和自底向上的双向特征传递，实现多尺度特征的高效融合。头部网络则延续了解耦头(Decoupled Head)设计，将分类与边界框回归任务分开处理，并引入了更加灵活的动态分配策略。

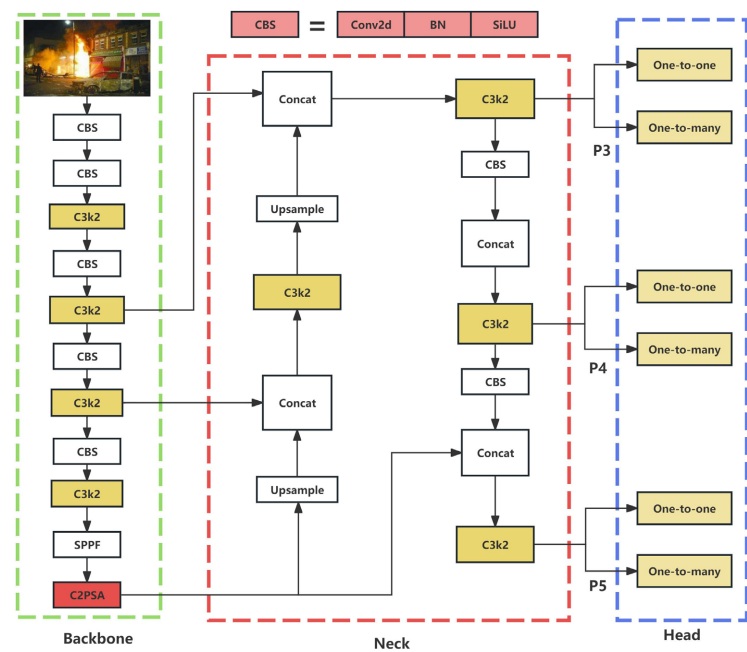


Figure 1. Network structure of YOLOv11
图 1. YOLOv11 的网络结构

尽管 YOLOv11 的 C3k2 模块在特征提取上表现卓越，但其内部包含的多个 Bottleneck 结构依赖大量的 3×3 标准卷积进行密集的矩阵乘加运算。在火灾检测这类特定任务中，背景信息中往往存在大量冗余特征，C3k2 模块的全通道密集计算不仅造成了算力浪费，而且在提取特征时并未对火灾的“高频视觉特性”进行针对性处理，容易导致微小火源特征在深层网络中被背景特征同化。因此，有必要对现有网络结构进行轻量化与特定领域的针对性重构。

2.2. 频空卷积模块设计

为了打破传统空间域卷积在提取极小目标时的局限性，并大幅度降低模型的参数规模，本文引入了频空卷积(FSConv)模块。该模块的设计灵感来源于信号处理理论，即图像可以被视为不同频率信号的叠加。火灾火焰由于其剧烈的色彩变化和锐利的轮廓，属于高频信息，而天空、建筑外墙等大面积平缓区域属于低频信息。FSConv 通过引入二维哈尔离散小波变换(2D Haar DWT)，实现了在 CNN 框架内对频域特征与空域特征的双线并行处理。FSConv 模块的具体内部结构如图 2 所示，其核心处理流程分为以下几个阶段：

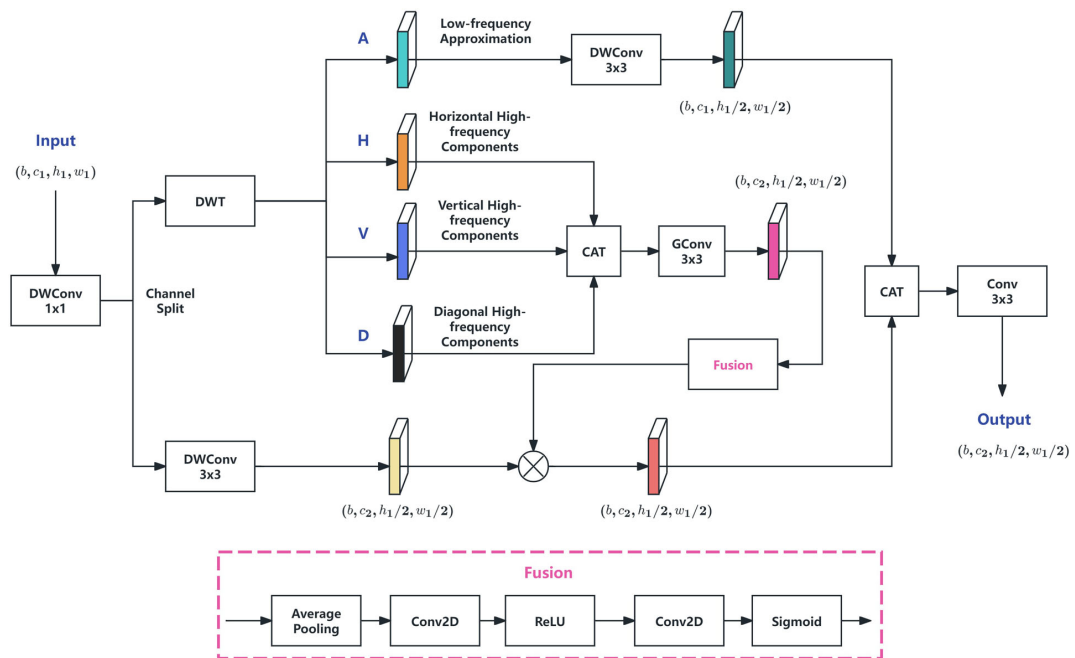


Figure 2. Network structure of FSConv

图 2. FSConv 网络结构

(1) 通道扩充与频空分流

当输入特征图进入 FSConv 模块时，首先经过一个深度可分离卷积(DWConv 1×1)。DWConv 通过将深度卷积和逐点卷积分离，在参数量极低的情况下将输入特征的通道数翻倍。随后，将翻倍后的通道对半切分，一半输入至频域分支，另一半输入至空域分支。这种设计保证了模型在提取跨域特征时拥有充足的维度容量，避免了信息瓶颈。

(2) 频域分支(Frequency Branch)的高低频解耦

在频域分支中，模块对输入特征应用离散小波变换(Discrete Wavelet Transform, DWT)。DWT 是一种无参数的数学变换，它利用四个简单的卷积核将特征图无损分解为四个子带：

低频近似分量(图 2 里的 A 分支)：包含了图像的低频全局结构信息(如火灾现场的大面积背景)。该分

量被单独送入 DWConv 进行特征提取。

高频细节分量(图 2 里的 H、V、D 分枝): 分别包含了水平、垂直和对角线方向的剧烈像素变化(即火焰的边缘、尖角等局部细节)。这三个分量被拼接后, 通过分组卷积(Group Conv, GConv)提取出融合的高频特征表示。

通过 DWT 的物理性解耦, 网络能够独立地对火灾的边缘特征与背景噪声进行过滤与学习, 从根本上缓解了复杂背景对火源的干扰。三个高频分量通过通道拼接(Concat, CAT)之后进行分组卷积(GConv 3×3), 卷积后的特征信息进入 Fusion 模块进行信息融合(融合过程如图 2 里的小图所示), 融化后的进行最后与空域分支进行乘积运算, 实现了频域与空域的信息融合。

(3) 空域分支(Spatial Branch)与高频注意力引导

在空域分支, 特征图首先通过一个深度可分离卷积(DWConv 3×3)提取空间局部特征, 之后与来自频域的特征信息进行乘积运算, 实现频域与空域信息的融合。FSConv 模块设计了一种极具创新的跨域注意力机制: 模块将前述频域分支中提取到的高频细节特征进行全局平均池化, 随后经过 Sigmoid 激活函数生成通道维度的注意力权重, 并将该权重直接乘回给空域分支的特征图。

这一设计的物理意义在于: 利用从频域中剥离出的极度纯粹的“火灾高频边缘”信号作为“向导”, 动态指导空间域卷积的激活重点。换言之, 网络被强制要求在空间域中更加关注那些高频特征响应强烈的通道, 从而实现了精准的火源定位。

(4) 频空特征融合

最后, 模块将经过高频注意力强化的空域局部特征(包含精确的火源位置)与频域分支提取的低频背景特征(包含全局环境上下文)进行通道拼接(Concat, CAT), 然后经过普通的 Conv 3×3 卷积进行跨通道信息整合与降维输出。将输入信号 Input (b, c_1, h_1, w_1) 和输出信号 Output $(b, c_2, h_1/2, w_1/2)$ 进行对比, 其中 b, c, h, w 分别表示批量大小、通道数、特征图的高度和宽度, 假设输入和输出信号的通道数相等, 即 $c_1 = c_2$, 那么输出信号特征图信息的高度和宽度分别下降了一半, 总数据量相当于输入信号的 1/4, 大大降低了后续处理的数据量。

2.3. 改进算法

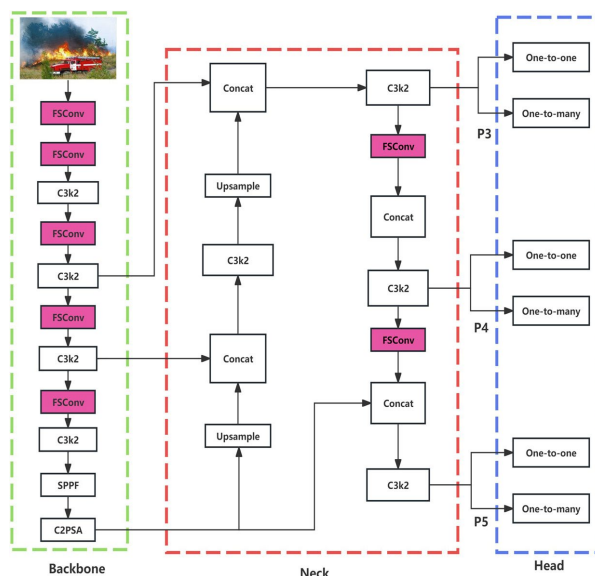


Figure 3. Network structure of improved algorithm

图 3. 改进算法的网络结构

基于上述理论分析,本文对 YOLOv11 的网络架构进行了深度重构。考虑到网络浅层主要提取边缘、角点等基础几何特征,若过早引入频域分析容易引入大量的背景高频噪声(如树叶纹理、建筑纹理等);而网络中深层的特征图已经具备了较高阶的语义抽象能力,此时引入频域分析能更精准地剥离真正的火源与宏观背景。因此,本文将 YOLOv11 骨干网络和颈部网络的 C3k2 模块,替换为 FSConv 模块。这一重构策略将原本基于单纯空间域连续密集卷积堆叠,转化为“空间提取-频域解耦-跨域融合”的新型范式。重构后的 FSConv-YOLOv11 网络,不仅大幅剔除了 C3k2 模块带来的冗余计算和过拟合风险,更赋予了模型在嘈杂环境中对微弱高频火灾信号的敏锐捕捉能力,改进算法的网络结构如图 3 所示。

3. 仿真实验与结果分析

3.1. 实验数据集

在深度学习模型的训练与评估中,数据集的质量与多样性直接决定了模型的泛化能力。为此,本文结合开源数据平台(如 AI Studio)以及自主收集的真实火灾图像,构建了一个多场景、高挑战性的综合火灾检测数据集。该数据集如图 4 所示,共包含 2059 张高质量火灾标注图像,广泛覆盖了真实世界中可能发生火灾的各类复杂场景,具体包括:森林与草原野火(背景多为绿色植被,存在大量高频纹理干扰)、城市建筑物火灾(背景包含复杂的几何线条与光影变化)、公路与车辆火灾(存在尾灯、路灯等高亮光源干扰)、以及室内工厂起火(光线昏暗且存在遮挡)。为了确保模型训练的有效性与测试结果的客观性,本文将数据集以 8:1:1 的比例严格划分为训练集、验证集和测试集,其中,训练集包含 1647 张图像,用于模型参数的反向传播优化,验证集和测试集各包含 206 张图像,分别用于训练过程中的超参数调优与最终模型性能的独立评估。



Figure 4. Data set
图 4. 数据集

3.2. 实验结果分析

本节通过仿真实验验证本文所提算法的性能,具体的实验参数如表 1 所示。

Table 1. Experimental parameters
表 1. 实验参数

参数名	参数值
训练集图片	1647
验证集图片	206
测试集图片	206
epoch	300
学习率	0.01
batchsize	8
优化器	SGD

图 5 为本文所提算法的整体训练过程。图中前三列子图反映各类损失指标在训练过程中的变化规律，横轴为训练迭代轮次 epoch，纵轴对应 loss 数值；后两列子图则展示模型精度指标的变化趋势，坐标轴设置与前者保持一致，横轴仍为迭代轮次，纵轴代表准确率。从训练过程可见，随着迭代轮次不断递增，分类损失 `clc_loss`、边框回归损失 `box_loss` 与分布特征损失 `dfl_loss` 均呈逐步下降并收敛的趋势；与此同时，精确率 `precision`、召回率 `recall` 以及全局平均精度 `mAP` 等性能指标持续上升，后期逐渐收敛并趋于平稳。

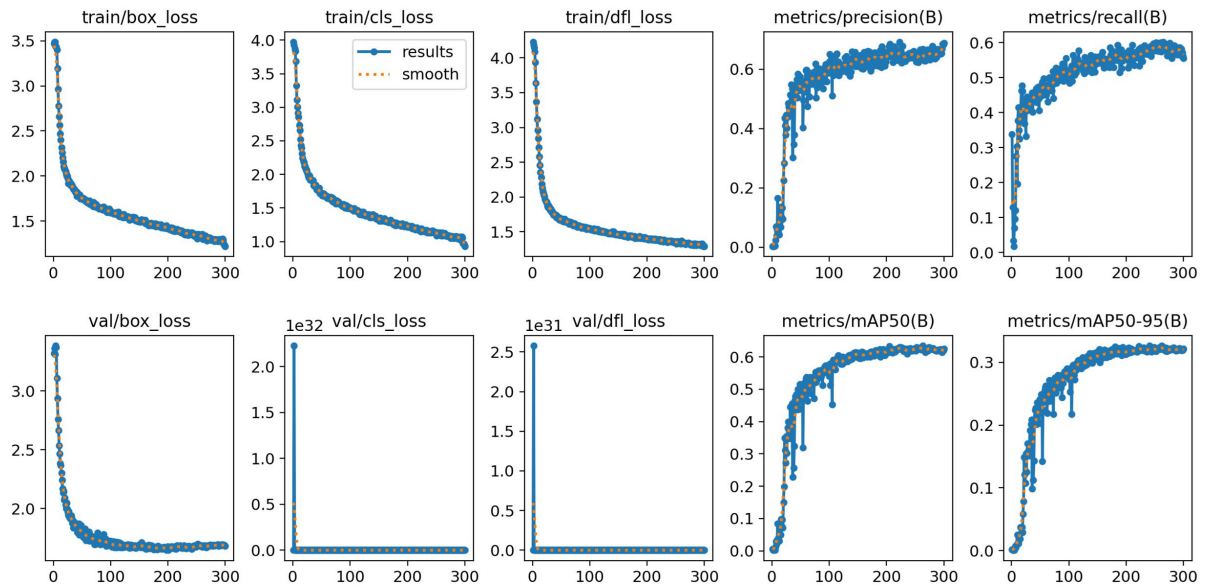


Figure 5. Training process
图 5. 训练过程

为了科学评估本文提出的 FSConv-YOLOv11 算法的综合性能，我们选取了全类平均精度($mAP@0.5$)作为衡量检测精度的核心指标，同时，参数量(Parameters)、梯度计算量(Gradients)和每秒十亿次浮点运算数(GFLOPs)作为衡量模型复杂度与轻量化程度的指标。我们将 YOLOv11 模型与本文改进的 FSConv-YOLOv11 模型在相同的数据集和实验环境下进行了严格对比，实验结果如图 6 和表 2 所示。图 6 显示的检测精度的对比，与 YOLOv11 相比，本文所提算法的 $mAP@0.5$ 指标从 0.620 提升至 0.632，提升了 1.2%，

这表明本文所提算法能够保持与 YOLOv11 相当的检测精度，甚至还实现了稍许的提高。

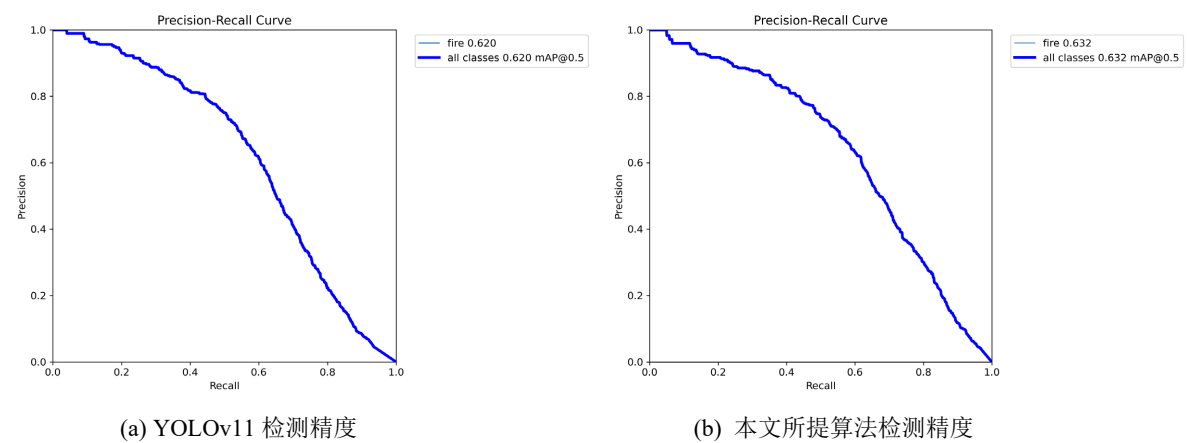


Figure 6. Comparison of the algorithm detection accuracy (mAP50)
图 6. 算法检测精度对比(mAP50)

表 2 显示的是复杂度指标的对比情况，我们加入了 YOLOv8 这一经典算法进行对比，从表 2 中可以清晰地看到，YOLOv11 算法相较 YOLOv8 算法已经全方面地大幅降低了复杂度，另外比较本文算法和 YOLOv11 算法，本文所提算法 FSCnv-YOLOv11 在提升精度的同时，进一步地降低了复杂度，本文所提算法的参数量(Parameters)由 2.61 M 下降至 2.21 M，降幅高达 15.4%；梯度计算量(Gradients)同样下降了 15.4%；而直接反映模型运算负担的 GFLOPs 指标则从 6.5 降低至 5.9，降幅达到 9.2%。这一成果主要归功于 FSCnv 模块成功替代了内部包含繁重标准卷积的 C3k2 模块，以极低计算代价的深度可分离卷积和无参数的小波变换算子完成了高效的特征提取。

Table 2. Comparison of the algorithm complexity
表 2. 算法复杂度对比

模型	Parameters	Gradients	GFLOPs
YOLOv8n	3,157,200	3,157,184	8.7
YOLOv11n	2,616,248	2,616,232	6.5
本文所提算法	2,213,268★	2,213,252★	5.9★

★表示这一类数据对比，结果最好的那一个。

为进一步验证所提算法的性能，我们使用测试数据集在多种场景下进行了火灾检测实验，结果如图 7 所示。从图 7 中可以看到，本文所提算法能够在不同环境中有效识别火灾目标，即使同一图像中存在多个火灾对象也能保持精准检测，另外对于图像中较小尺寸的火灾目标，本文所提算法也具备出色的识别能力。





Figure 7. Testing situation of the proposed algorithm

图 7. 本文所提算法的测试情况

4. 结论

本文针对传统视觉算法及重型深度学习网络在复杂场景火灾检测中存在的抗干扰能力差、模型部署困难等问题,提出了一种基于 FSConv 改进的轻量化火灾检测算法——FSConv-YOLOv11。该算法打破了传统 YOLO 模型仅在空间域提取特征的局限,巧妙地融合了信号处理领域的小波变换技术,将火灾特征在频域进行高低频解耦,利用高频细节引导空间注意力机制,极大增强了模型对微弱火源轮廓的感知能力。实验结果表明:本文所提算法在保持甚至超越原生 YOLOv11 检测精度的同时,显著减少了网络参数数量和浮点运算数。本文所提算法在检测精确性与计算轻量化之间取得了优异的平衡,有效克服了传统目标检测模型在复杂自然环境或工业场景下的瓶颈。该算法不仅具备强大的鲁棒性与泛化能力,更因其极低的硬件算力门槛,展现出了在智能无人机巡护、林区边缘计算监控网络及家庭智能安防等实际工程部署中的巨大应用潜力,为智慧城市与公共安全体系的建设贡献了坚实的技术力量。未来的工作将尝试引入轻量化的时空序列融合机制,利用火焰动态闪烁的时序特征进一步提升对极早期微小火源的辨识能力。

基金项目

贵州交通职业大学高层次人才科研启动基金项目:复杂交通环境下的多目标检测研究(KYQD2022004)。

参考文献

- [1] 吴方卫, 石翟, 吴毅, 等. 基于多传感器融合的火灾检测算法综述[J]. 数字技术与应用, 2024, 42(12): 204-207.
- [2] Redmon, J., Divvala, S., Girshick, R. and Farhadi, A. (2016) You Only Look Once: Unified, Real-Time Object Detection. 2016 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, 27-30 June 2016, 779-788. <https://doi.org/10.1109/cvpr.2016.91>
- [3] 殷波. 基于改进 YOLOv8 的轻量化火灾检测算法[J]. 计算机科学与应用, 2024, 14(9): 47-55.
- [4] 王一旭, 肖小玲, 王鹏飞, 等. 改进 YOLOv5s 的小目标烟雾火焰检测算法[J]. 计算机工程与应用, 2023, 59(1): 72-81.
- [5] Bu, L., Li, W., Zhang, H., Wang, H., Tian, Q. and Zhou, Y. (2025) FIRE-YOLOv8s: A Lightweight and Efficient Algorithm for Tunnel Fire Detection. *Fire*, 8, Article 125. <https://doi.org/10.3390/fire8040125>
- [6] 张建华, 郭晓, 黄刚. 基于改进轻量级 YOLOv8 的电铲机房火灾检测算法研究[J]. 武汉理工大学学报, 2025, 47(1): 68-74, 97.
- [7] Ultralytics (2025) Ultralytics YOLO11 Official Repository. GitHub. <https://github.com/ultralytics/ultralytics>
- [8] 孔垂乐, 孟宇宇, 霍久远, 等. 改进 YOLOv11 的无人机海上小目标检测算法[J]. 计算机工程与应用, 2026, 62(1): 151-161.
- [9] 高俊义, 张伟, 李泽林. YOLO-BFEPS: 一种高效注意力增强跨尺度 YOLOv10 火灾检测模型[J]. 计算机科学, 2025, 52(z1): 412-420.
- [10] Xu, W., Ding, Z., Wang, Z., Cui, Z., Hu, Y. and Jiang, F. (2025) Think Locally and Act Globally: A Frequency-Spatial

-
- Fusion Network for Infrared Small Target Detection. *IEEE Transactions on Geoscience and Remote Sensing*, **63**, 1-17. <https://doi.org/10.1109/tgrs.2025.3612417>
- [11] Nyi Myo, N., Boonkong, A., Khampitak, K. and Hormdee, D. (2025) A Two-Point Association Tracking System Incorporated with YOLOv11 for Real-Time Visual Tracking of Laparoscopic Surgical Instruments. *IEEE Access*, **13**, 12225-12238. <https://doi.org/10.1109/access.2025.3529710>