

基于Faster R-CNN的火灾检测方法研究

刘倚帆, 韩响策, 陆思彤, 凌天靖, 程广涛*

天津商业大学信息工程学院, 天津

收稿日期: 2026年3月21日; 录用日期: 2026年4月14日; 发布日期: 2026年4月23日

摘要

火灾早期精准检测是降低生命财产损失、实现防灾减灾目标的核心技术, 而烟雾作为火灾初期最显著的视觉表征, 其浓度扩散、形态演变等动态特征可有效反映火势蔓延趋势。本研究以Faster R-CNN目标检测框架为基础开展系统性优化。首先, 将原框架中的ResNet-50主干网络替换为ResNeXt-50网络, 提升模型的特征表达能力; 其次, 将原框架中的FPN替换成PAFPN, 显著提升了模型对小目标烟雾的检测能力, 同时保持高效推理速度。实验结果表明, 改进后模型的整体检测准确率较原模型提高2.6个百分点, 同时处理时间维持在0.22 ms/帧。特别值得注意的是, 改进模型对不同尺度烟雾目标的检测性能有所提升, 其中小目标检测准确率提升1.2个百分点, 大目标提升3.8个百分点。本研究为火灾烟雾的实时监测提供了高效可靠的解决方案。

关键词

深度学习, 火灾检测, 烟雾识别, Faster R-CNN, ResNeXt

Research on Fire Detection Method Based on Faster R-CNN

Yifan Liu, Xiangce Han, Sitong Lu, Tianjing Ling, Guangtao Cheng*

School of Information Engineering, Tianjin University of Commerce, Tianjin

Received: March 21, 2026; accepted: April 14, 2026; published: April 23, 2026

Abstract

Early and precise detection of fire is a core technology for reducing life and property losses and achieving the goal of disaster prevention and mitigation. As the most prominent visual manifestation of fire in its early stage, smoke, with its dynamic characteristics such as concentration diffusion

*通讯作者。

文章引用: 刘倚帆, 韩响策, 陆思彤, 凌天靖, 程广涛. 基于 Faster R-CNN 的火灾检测方法研究[J]. 软件工程与应用, 2026, 15(2): 241-253. DOI: 10.12677/sea.2026.152023

and shape evolution, can effectively reflect the trend of fire spread. This study conducts systematic optimization based on the Faster R-CNN object detection framework. Firstly, the ResNet-50 backbone network in the original framework is replaced with the ResNeXt-50 network to enhance the model's feature expression ability. Secondly, the FPN in the original framework is replaced with PAFPN, significantly improving the model's detection ability for small smoke targets while maintaining efficient inference speed. Experimental results show that the overall detection accuracy of the improved model is 2.6 percentage points higher than that of the original model, and the processing time remains at 0.22 ms per frame. Notably, the detection performance of the improved model for smoke targets of different scales has been enhanced, with the detection accuracy for small targets increasing by 1.2 percentage points and that for large targets by 3.8 percentage points. This study provides an efficient and reliable solution for real-time monitoring of fire smoke.

Keywords

Deep Learning, Fire Detection, Smoke Recognition, Faster R-CNN, ResNeXt

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

当前社会发展阶段, 火灾对人类生存与生态可持续性已构成显著威胁。据国家消防救援部门统计, 仅 2024 年全国记录火灾事件即达 90.8 万起, 造成直接经济损失高达 77.4 亿元。火灾不仅损毁建筑设施、导致人员伤亡, 更对森林地带及草原生态系统的稳定性产生深远影响。由于火灾在发生初期往往难以察觉, 一旦蔓延至高温有毒烟气快速扩散阶段, 扑救与逃生难度将呈几何级数增加, 因此, 实现火情的早期发现与预警是防灾减灾的关键。

烟雾作为火灾的重要表征, 在早期火灾检测阶段十分重要。烟雾在燃烧过程中, 烟雾的颜色、形态和浓度会呈现动态变化, 其物理、化学特性则蕴含着丰富的火情信息: 在火灾初期的阶段, 烟雾是色泽浅淡、低密度状态, 这时的烟雾呈现出丝状或片状升腾的形态; 随着火势的发展, 烟雾会偏向灰黑色, 烟雾的浓度也会显著提升, 此时烟雾呈现出的形态是翻滚涌动的湍流, 通过算法捕捉烟雾的这些特性, 对于火灾检测十分重要。

目标检测算法的发展为火灾烟雾检测技术的智能化奠定了坚实基础。自 Girshick 等提出 RCNN 经典模型[1]以来, Fast RCNN [2]与 Faster R-CNN [3]相继问世。其中, Ren 等提出的 Faster R-CNN 通过引入区域建议网络(RPN), 成功实现了检测速度与精度的良好平衡, 而 He 等优化的 Mask R-CNN 进一步在实例分割中表现卓越[4]。基于此类深度学习框架, 祝玉华等系统综述了其在火灾检测中的应用进展[5]。围绕火灾烟雾检测的特定需求, 国内外学者开展了大量优化工作: 在数据层面, Zhang 等通过合成烟雾图像有效缓解了数据短缺及真实场景的分布差异[6]; 在网络结构与优化策略上, 学者们通过构建基于 Resnet50 或 ResNet-101 的深层特征提取网络、引入 RoI Align 及 soft-NMS 算法等方式提升了模型识别率[7], 并相继探索了分裂机制[8]、鲸鱼优化算法调节损失权重[9]、嵌入注意力机制的轻量化 YOLOv4 [10]以及运动与多特征融合策略[11], 全方位提升了火灾目标的检测效能。

尽管现有研究取得了显著进展, 但在火灾初期的实际复杂场景下, 早期烟雾检测仍面临严峻挑战。火灾初期的烟雾往往呈现小目标特性, 涵盖像素占比低且语义信息匮乏, 致使模型特征提取困难、检

测效能受限。针对这一难点, 诸多研究提出了多尺度特征融合的改进方案: 如引入特征金字塔(FPN)与可变形卷积[12]、结合关联网络融合形状与位置特征[13]、多层频域特征融合与锚点优化[14], 以及优化RPN锚点尺寸融合上下文信息[15]。此外, 研究者还设计了多尺度空间增强的预测金字塔及正负样本平衡策略[16], 或采用深度可分离分组卷积来弥补小目标特征提取能力的不足[17]。然而, 当前模型在面对复杂环境时, 尤其是对边缘模糊、形态多变的早期薄烟敏感度依然有限。因此, 如何在兼顾计算效率的前提下, 进一步增强对多尺度早期烟雾的特征感知与长程依赖建模能力, 是当前火灾检测领域亟待突破的技术瓶颈。

针对火灾早期烟雾检测中特征提取能力不足与多尺度目标识别困难等痛点, 本研究以 Faster R-CNN [3]框架为基础, 提出了一种优化的火灾烟雾检测模型。首先, 考虑到传统卷积神经网络难以有效应对多尺度目标的挑战, 本文采用基于分组卷积策略的 ResNeXt-50 结构[18]替换原始的 ResNet-50 主干网络, 在保持计算复杂度可控的同时, 显著提升了模型对复杂场景下烟雾多尺度特征的表达能力。其次, 针对传统单向特征金字塔网络(FPN) [19]在多尺度融合时不充分且小目标信息丢失严重的问题, 本文引入路径聚合特征金字塔网络 PAFPN [20]对特征融合模块进行改进。该结构通过自顶向下与自底向上的双向特征传播及横向连接, 构建了高保真的多层次融合机制, 使基于 ResNet-50 模型的小目标检测精度由 72.0%大幅升至 82.8%。最后, 在涵盖 4539 张多尺度样本的自建烟雾数据集上进行的系统对比与消融实验表明, 改进后的模型平均精度达到 69.9%, 较基准 ResNet-50 模型提升了 2.6 个百分点。该方法在维持 0.220 ms/帧实时处理速率的同时, 有效提升了烟雾的识别性能, 实现了检测准确性与实时性的良好平衡。

2. Faster R-CNN 基础模型

本文所改进的火灾烟雾检测算法以 Faster R-CNN 为基础架构。该模型作为经典的高精度双阶段目标检测网络, 通过整合区域提议网络(RPN)与 Fast R-CNN, 实现了特征提取、候选区域生成及分类回归的端到端计算。Faster R-CNN 的整体结构如图 1 所示。

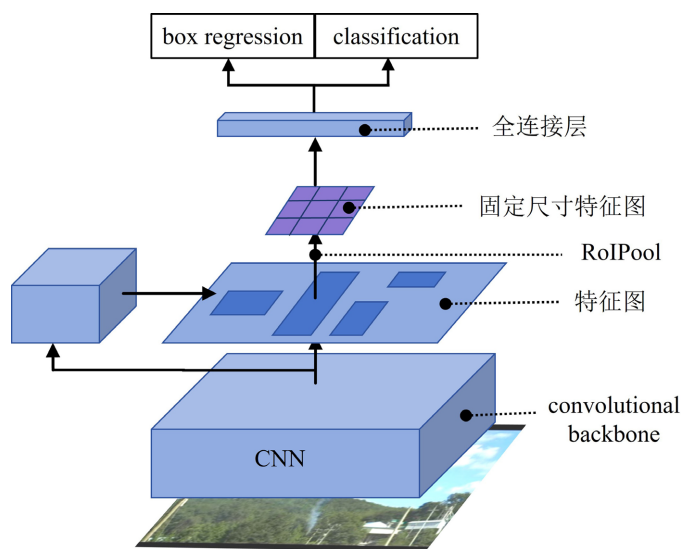


Figure 1. Faster R-CNN architecture
图 1. Faster RCNN 结构图

在其整体工作流程中, 输入图像首先由卷积主干网络提取多尺度特征, 随后共享特征图被送入 RPN 模块。RPN 采用双分支结构, 通过在共享特征图的每个位置密集生成不同尺寸与比例的预定义锚框

(Anchors), 并结合非极大值抑制(NMS)有效处理重叠的候选框、剔除冗余, 从而高效输出高质量的目标候选区域。为了进一步应对多尺度目标的检测难题, 原始框架引入了单向的特征金字塔网络(FPN)。

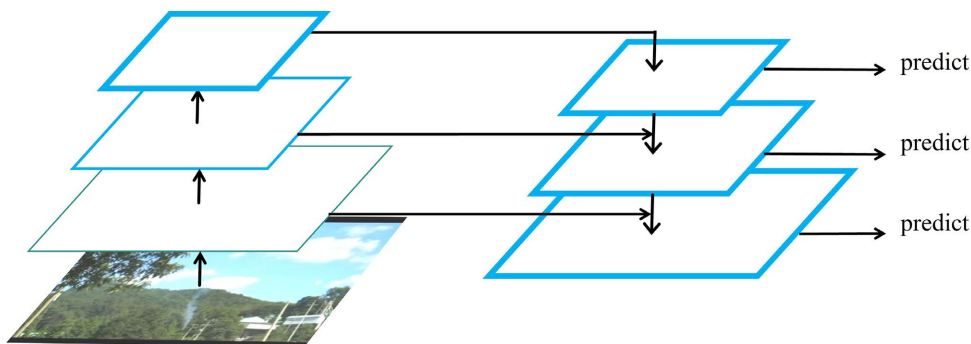


Figure 2. FPN architecture
图 2. FPN 网络结构

如图 2 所示, FPN 网络通过自下而上的特征提取与自上而下的语义传递, 并辅以横向连接, 将深层的丰富语义信息与浅层的高分辨率局部细节相融合, 构建起多层次的特征表达体系。

然而, 面对火灾早期烟雾这一特定检测任务, 传统的 Faster R-CNN 架构暴露出明显的局限性: 一方面, 火灾初期的烟雾形态多变且边缘模糊, 传统主干网络(如单一路径的 ResNet-50)特征表达能力不足, 难以充分应对复杂环境下的多尺度特征提取挑战; 另一方面, 早期烟雾在图像中通常表现为像素占比极低的小目标, 传统 FPN 仅依赖单向的自上而下融合, 导致底层富含细节的空间信息在向高层传递时损耗严重, 多尺度特征融合极不充分, 极易造成小目标烟雾的漏检与信息丢失。因此, 为提升火灾早期预警的准确性, 必须对该基础模型的主干网络与特征融合机制进行深度重构与优化。

3. 基于改进 Faster R-CNN 的火灾检测方法

本研究提出的火灾检测网络模型基于改进的 Faster R-CNN 架构, 通过多维度优化实现了对烟雾目标的精准识别。本研究所提出的火灾烟雾检测网络模型如图 3 所示。

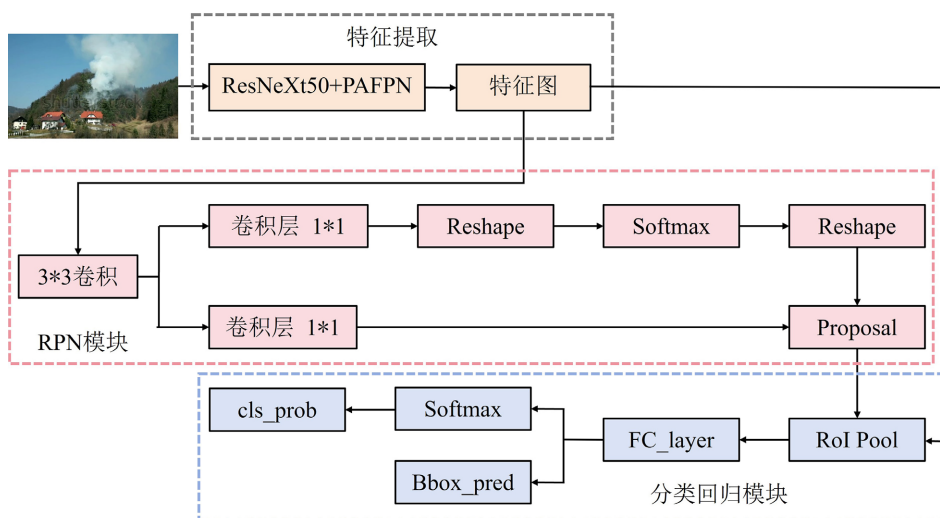


Figure 3. Fire smoke detection model
图 3. 火灾烟雾检测模型

3.1. 特征提取

本文对输入的烟雾图像采用 ResNeXt50 + PAFPN 组合进行特征提取。二者结合，从输入图像中生成包含丰富语义和多尺度信息的特征图。下面将对两个改进模块进行详细介绍。

3.1.1. ResNeXt

ResNeXt 是由 Facebook AI Research (FAIR) 提出的一种高效卷积神经网络架构，其核心思想是通过分组卷积和基数的引入，在保持模型复杂度的同时，显著提升特征的表达能力[18]。该网络在 ImageNet 分类任务中表现出色，同时也被广泛应用于目标检测、语义分割等计算机视觉任务，包括火灾烟雾检测。

ResNeXt 的核心创新在于其“分裂 - 变换 - 合并”策略，即在残差块内部使用多组并行的卷积路径，每组执行不同的特征变换，最终合并结果。相较于传统 ResNet 的单一变换路径，ResNeXt 通过增加基数 (Cardinality, 即分组数) 来提升模型的表达能力，而非单纯增加深度或宽度。数学上，ResNeXt 的单个残差块可表示为如式(1)所示。其中： x 为输入特征， C 为基数，即分组数， T_i 表示第 i 组的变换函数。

$$y = x + \sum_{i=1}^C T_i(x) \quad (1)$$

ResNet 和 ResNeXt 的模块结构对比如图 4 所示。在 ResNet 中，输入的具有 256 个通道的特征经过 1×1 卷积压缩 4 倍到 64 个通道，之后 3×3 的卷积核用于处理特征，经 1×1 卷积扩大通道数与原特征残差连接后输出。ResNeX 也是相同的处理策略，但输入的具有 256 个通道的特征被分为 32 个组，每组被压缩 64 倍到 4 个通道后进行处理。32 个组相加后与原特征残差连接后输出。

在火灾烟雾检测任务中，由于烟雾的形态多变，ResNeXt 的分组卷积机制能够并行学习不同抽象层次的特征，提高对不同烟雾模式的适应性。相较于标准 ResNet-50，ResNeXt-50 的参数量仅增加约 10%，但特征多样性显著提升，由实验表明在烟雾检测任务中 mAP 提高约 1.6%。ResNeXt 凭借其分组卷积机制，在火灾烟雾检测任务中展现出更强的特征学习能力，特别是在复杂背景和多尺度目标检测场景下。

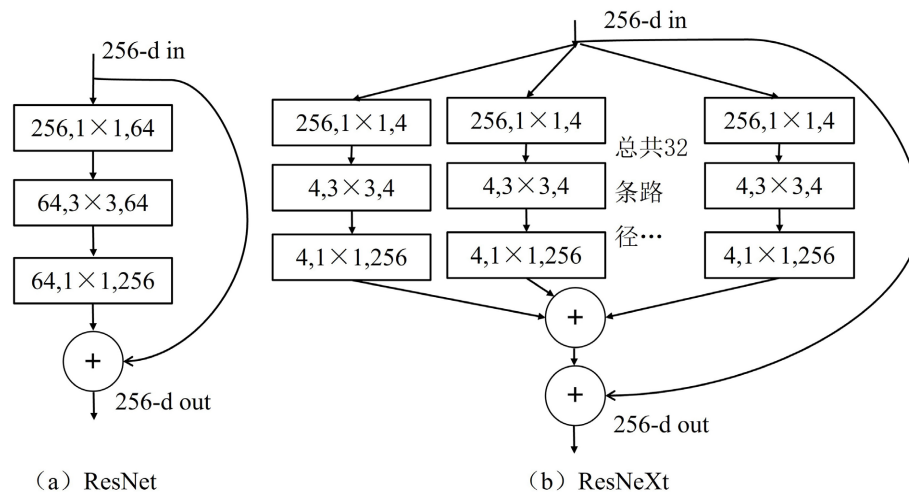


Figure 4. Comparison of module structures between ResNet and ResNeXt
图 4. ResNet 和 ResNeXt 的模块结构对比图

3.1.2. PAFPN

PAFPN 是一种高效的多尺度特征融合网络[20]，本文通过改进传统 FPN [19] 结构，显著提升了目标检测任务中对不同尺度目标的识别能力。其核心创新在于构建双向(自顶向下 + 自底向上)的特征传播路

径,并引入横向连接增强特征复用。传统FPN仅通过自上而下和横向连接进行特征融合,对于底层特征的利用不够充分,在检测小目标时存在局限性。PAFPN引入额外的路径聚合结构,旨在缩短底层特征与高层特征的传递路径,减少信息损耗,使底层细节信息能更高效地融入高层语义特征中,从而提升网络对小目标的检测能力。

PAFPN框架如图5所示。其基础金字塔结构沿用FPN[19]的自下而上和自上而下路径,再在FPN基础上,添加路径聚合路径,即从最底层特征图开始,经过一系列 3×3 卷积和上采样操作,将底层特征逐层向上传递,并与自上而下路径生成的特征图进行逐元素相加,使得底层的位置细节信息能够快速传递到高层,加强了特征的空间信息表达。在实际火灾烟雾检测任务中,PAFPN相较于FPN,能够使Faster R-CNN[3]在多尺度目标检测任务中获得更高的平均精度(AP),尤其在小目标检测指标上有显著提升。

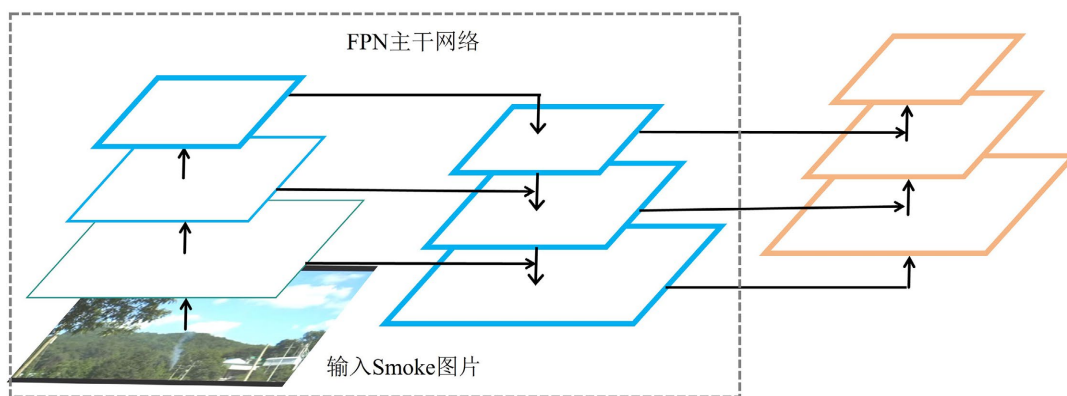


Figure 5. Schematic diagram of PAFP framework
图5. PAFP框架示意图

3.2. RPN

以ResNeXt50+PAFPN提取的特征图为输入,先会经过一个 3×3 卷积,对输入特征图进行初步处理,之后分两路分别经过 1×1 卷积,一路经过Reshape、Softmax等操作,用于生成目标的置信度得分,判断特征图中区域是否为目标;另一路经过后续操作生成目标的候选区域Proposal,即初步确定可能存在目标的区域位置。RPN模块的作用是快速生成一系列目标候选区域,减少后续分类回归模块的处理范围,提升检测效率。

3.3. 分类回归模块

将RPN模块生成的候选区域通过RoI Pool(感兴趣区域池化)操作,将不同大小的候选区域特征图转换为固定大小,便于后续处理。接着通过全连接层FC_layer进一步提取特征,再分两路分别经过Softmax和其他处理。如图3所示,cls_prob表输出类别概率,用于判断目标属于不同类别的可能性,实现目标分类;Bbox_pred表输出边界框预测值,对目标的位置和大小进行精确回归调整,使检测框能更准确地框定目标。

4. 实验与结果分析

4.1. 实验准备

4.1.1. 实验环境配置

本研究采用AutoDL云平台作为实验环境,其高性能计算资源为火灾烟雾检测模型的开发提供了可靠保障。硬件方面,配置了NVIDIA vGPU(32 GB显存)和12 GB RAM内存,这种组合设计有效解决了

深度学习中的两大关键问题：一方面，大显存支持批量处理更多图像数据，显著提升训练效率；另一方面，充足的内存避免了数据加载时的性能瓶颈。具体的软硬件配置信息如表 1 所示。

Table 1. Software and hardware environment configuration

表 1. 软硬件环境配置

实验环境	信息
GPU	vGPU-32 GB (32 GB)
CPU	16 vCPU AMD EPYC 9654 96-Core Processor
RAM	12 GB
操作系统	Windows11, 64 位操作系统, 基于 x64 的处理器
Python	3.8
PyTorch	1.10.0
CUDA	11.3

4.1.2. Smoke 自建数据集

在数据构建方面，本研究构建的数据集共包含 4539 张 smoke 图片，采用行业广泛应用的 COCO 格式进行存储和管理。该数据集在采集过程中，充分考虑了实际火灾场景的复杂性，全面覆盖了不同的应用场景。在数据集的构成比例上，小目标烟雾图像占比约 25%，中目标约 42%，大目标约 33%。Smoke 数据集部分图像如图 6 所示。在本研究中，训练、验证和测试的划分比例约 8:1:1，其中训练集包含 3631 张图片，验证集包含 453 张图片，测试集包含 455 张图片。

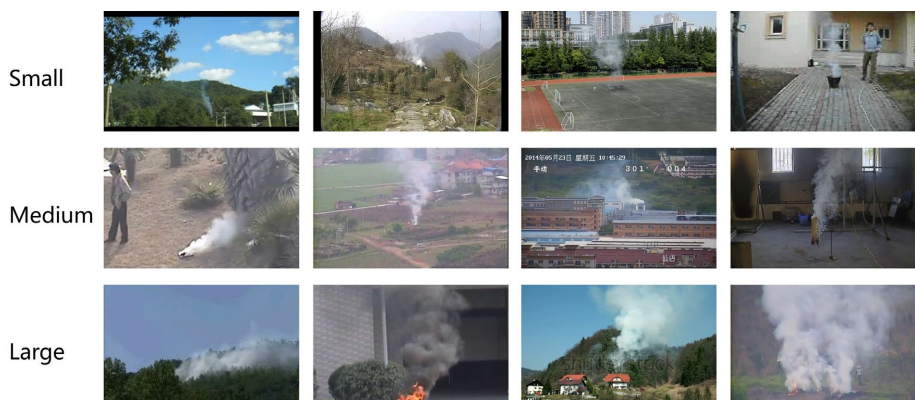


Figure 6. Display of selected images from the smoke dataset categorized by small, medium, and large scales

图 6. Smoke 数据集部分图像按小中大展示

4.1.3. 评价指标

为验证火灾烟雾检测模型的有效性，本文主要采用平均精度(*mAP*)为评价指标，对实验结果进行评估。

(1) 交并比(*IoU*): 预测框与真实框的重叠程度，计算公式如式(2)所示，其中 B_{pred} 为预测框， B_{gt} 为真实框。通常以 $IoU \geq 0.5$ 作为正样本阈值。

$$IoU = \frac{B_{pred} \cap B_{gt}}{B_{pred} \cup B_{gt}} \quad (2)$$

(2) 准确率(Precision, 指正确预测为正例的实例个数占有所有被预测为正例的实例个数的比重)与召回

率(Recall, 指将正确预测实例个数占所有正确标注实例个数的比重)的表达式分别如式(3)、式(4)所示。

$$P = \frac{TP}{TP + FP} \tag{3}$$

$$R = \frac{TP}{TP + FN} \tag{4}$$

其中 TP 为真阳性: 常指将正样本预测成为正样本的个数, FP 为假阳性: 指将负样本预测成为正样本的个数, FN 为假阴性: 将负样本预测成为负样本的个数。

(3) PR 曲线: P 表示为准确率, R 表示为 Recall (召回率)。在目标检测中, 准确率 P 和召回率 R 的数值越大, 检测的性能越好。

(4) 平均精度(Average Precision, AP): 表预测类别的平均精度。计算方法是坐标轴与 PR 曲线所围成的面积, 也可通过其他方法来计算, 例如图片有 n 张, 其中类别 c 的平均精度 AP_c 。可以表示为式(6)所示。

$$P_{c,i} = \frac{TP}{TP + FP} \tag{5}$$

$$AP_c = \frac{1}{n} \sum_{i=1}^n P_{c,i} \tag{6}$$

(5) 平均精度均值(mean Average Precision, mAP): 若检测类别数为 m , 则 COCO 标准 mAP 表达式如式(7)所示。 mAP 表示每个类别的平均精度的算数平均值。 mAP_{50} 指在宽松阈值($IoU \geq 0.5$)下, 预测框的平均精度, 同样 mAP_{75} 反映严格定位要求($IoU \geq 0.75$)时的平均精度。 mAP 综合反映模型定位与分类能力, mAP 值越大, 表明算法检测性能越好, 其是火灾检测任务的首选指标。

$$mAP = \frac{1}{m} \sum_{i=1}^m AP_c \tag{7}$$

4.2. 实验结果比较与可视化

4.2.1. 实验结果比较

在火灾烟雾检测研究领域, 为了能够科学、严谨地验证本文所提出研究方法的优越性, 构建可靠且具有说服力的实验论证体系至关重要。基于此, 本研究在实验流程的设计上, 采用了逐步递进、层层深入的研究策略。首先, 成功运行基线模型 Faster R-CNN [3], 以此作为后续实验对比的基准。Faster R-CNN 作为经典的目标检测框架, 在诸多视觉任务中展现出良好性能, 将其作为基线模型, 能够为评估改进方法的有效性提供稳定且具有代表性的参照标准。

Table 2. Comparison of network structures between ResNet-50 and ResNeXt-50

表 2. ResNet-50 与 ResNeXt-50 的网络结构对比

Stage	Output	ResNet-50	ResNeXt-50 (32 × 4d)
Conv1	112 × 112	7 × 7, 64, stride2 3 × 3 max pool, stride2	7 × 7, 64, stride2 3 × 3 max pool, stride2
Conv2	56 × 56	$\begin{bmatrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 128 \\ 3 \times 3, 128, C = 32 \\ 1 \times 1, 256 \end{bmatrix} \times 3$
Conv3	28 × 28	$\begin{bmatrix} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 512 \end{bmatrix} \times 4$	$\begin{bmatrix} 1 \times 1, 256 \\ 3 \times 3, 256, C = 32 \\ 1 \times 1, 512 \end{bmatrix} \times 4$

续表

Conv4	14×14	$\begin{bmatrix} 1 \times 1, 256 \\ 3 \times 3, 256 \\ 1 \times 1, 1024 \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1, 512 \\ 3 \times 3, 512, C = 32 \\ 1 \times 1, 1024 \end{bmatrix} \times 6$
Conv5	7×7	$\begin{bmatrix} 1 \times 1, 512 \\ 3 \times 3, 512 \\ 1 \times 1, 2048 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 1024 \\ 3 \times 3, 1024, C = 32 \\ 1 \times 1, 2048 \end{bmatrix} \times 3$
	1×1	global average pool 1000-d fc, softmax	global average pool 1000-d fc, softmax
#params		25.5×10^6	25.0×10^6
FLOPs		4.1×10^9	4.2×10^9

在完成基线模型的搭建与验证后，本研究精心设计并开展了一系列消融实验。消融实验旨在通过逐步改变模型的关键组成部分或添加特定技术，系统地检验这些改动对模型性能产生的影响，明确各部分在整体模型中所发挥的作用，从而精准定位能够提升模型性能的核心要素。具体而言，本研究尝试了以下多种改动：

(1) 将 FPN 替换为 PAFPN

特征金字塔网络是目标检测中常用的多尺度特征融合模块，但在处理烟雾这类具有复杂形态和尺度变化的目标时，其融合效果存在一定局限性。路径聚合特征金字塔网络 PAFPN [20] 通过引入双向的特征传播路径，能够更有效地融合不同尺度的特征信息，增强模型对小目标烟雾的检测能力。将原来的 FPN 替换为 PAFPN [20]，旨在探究这种改进后的多尺度融合策略是否能够显著提升模型在火灾烟雾检测任务中的性能表现。

(2) 在 ResNet-50 基础上加入 SE 注意力机制

组合构成 SE-ResNet50，ResNet-50 作为经典的卷积神经网络主干架构，在图像特征提取方面具有良好的性能。而挤压激励注意力机制能够通过学习自动获取每个特征通道的重要程度，有针对性地增强关键特征通道的响应，抑制不重要的通道。将 SE 注意力机制融入 ResNet-50，构建 SE-ResNet50 模型，期望通过增强网络对烟雾特征的选择性关注，提升模型对烟雾目标的识别精度，同时观察这种改动对模型整体性能的影响，包括是否会引入额外的计算负担或导致过拟合等问题。

(3) 将主干网络替换为 ResNeXt-50

ResNeXt-50 [18] 基于分组卷积的设计理念，通过增加网络的基数而非单纯增加网络深度，在不显著增加计算量的前提下，有效提升网络的特征表达能力和模型性能。ResNeXt-50 的网络结构对比 ResNet-50 如表 2 所示，ResNeXt-50 ($32 \times 4d$) 与 ResNet-50 在整体架构上保持高度一致，均包含 5 个 stage 的卷积层设计，输出特征图尺寸从 112×112 逐步降采样至 7×7 。二者的核心差异体现在 conv2 至 conv5 阶段的 bottleneck 结构中：ResNeXt-50 将传统的 3×3 卷积替换为分组卷积 ($group = 32$)，同时在中间层将通道数从 64 提升至 128 (conv2 阶段)，这种设计通过增加并行路径数量来增强特征多样性。然而尽管 ResNeXt-50 增加了特征处理路径，但其参数量 (25.0 M) 与 ResNet-50 [4] (25.5 M) 基本相当，且计算量 (FLOPs) 仅从 4.1 G 略微增加至 4.2 G。实验证明，相较于 ResNet-50，ResNeXt-50 能够更好地捕捉烟雾的多尺度特征和复杂形态变化。将主干网络替换为 ResNeXt-50，旨在验证这种更先进的主干架构在火灾烟雾检测任务中，是否能够为模型带来更强大的特征提取能力，进而提升检测精度和效率。

(4) 在主干网络替换为 ResNeXt-50 基础上，再将 FPN 替换为 PAFPN

该改进结合了上述两种关键改进，即采用性能更优的 ResNeXt-50 作为主干网络以增强特征提取能

力,同时利用 PAFPN [20]的优势提升多尺度特征融合效果。通过这种组合,期望能够充分发挥两者的协同效应,进一步优化模型对火灾烟雾目标的检测性能,尤其是在小目标检测和复杂场景下的检测能力,验证多方面改进对模型性能提升的综合影响。最后为了确保实验结果的公平性和准确性,避免因实验环境和数据差异导致的结果偏差,上述所有实验均严格遵循统一的实验标准,均使用同一数据集,即包含 4539 张 smoke 图片的 COCO 格式数据集,该数据集覆盖了不同场景、光照条件及烟雾浓度梯度,且具有明确的小目标、中目标和大目标烟雾图像比例分布,确保了实验数据的一致性和可比性,从而为准确评估各改动对模型性能的影响提供了坚实的基础。

将本文提出的 ResNeXt-50 + PAFPN 网络模型与上述其他结构的网络模型进行比较,模型实验数据如表 3 所示,其中 Time 的单位为 ms/帧。

Table 3. Model experimental data

表 3. 模型实验数据

网络模型	<i>mAP</i>	<i>mAP</i> ₅₀	<i>mAP</i> ₇₅	<i>mAP</i> _s	<i>mAP</i> _m	<i>mAP</i> _l	Time
ResNet50	67.0%	96.8%	78.4%	72.0%	65.3%	67.7%	0.216
ResNet50-PAFPN	68.6%	95.7%	80.9%	82.8%	64.4%	70.4%	0.220
SE-ResNet50	62.9%	96.1%	73.5%	75.9%	59.7%	63.7%	0.216
ResNeXt50	68.9%	97.0%	81.3%	80.8%	66.5%	69.9%	0.218
ResNeXt50-PAFPN	69.6%	96.9%	80.9%	73.2%	64.8%	71.5%	0.221

通过严谨的实验设计与数据分析,本研究所得实验结果明确揭示了模型架构改进在火灾烟雾检测领域所发挥的重要作用,即模型架构的改进对火灾烟雾检测性能具有显著影响。在网络架构的对比研究中,以 ResNeXt50 [18]作为主干网络展现出了相较于原始 ResNet50 更为卓越的特征提取能力。具体而言,ResNeXt50 在平均精度均值(*mAP*)这一重要指标上,较 ResNet50 提升了 1.9 个百分点,从 67.0% 提升至 68.9%。这一数据有力地验证了分组卷积在捕捉烟雾多尺度特征方面的有效性,分组卷积通过独特的结构设计,能够更细致、全面地对不同尺度的烟雾特征进行提取和分析,从而提升了模型对烟雾的识别能力。

在多尺度特征融合模块的研究中,路径聚合特征金字塔网络 PAFPN [19]的引入显著提升了小目标检测性能。当将 PAFPN [19]与 ResNet50 相结合,构建以 ResNet50-PAFPN 为骨干网络的烟雾检测模型时,该模型在小目标检测性能(*mAP*_s)方面取得了令人瞩目的提升,其小目标检测性能(*mAP*_s)从 72.0% 提升至 82.8%,增幅 10.8 个百分点,同时整体 *mAP* 从 67.0% 提升至 68.6%。然而,值得深入探究的是,这一显著优势在 ResNeXt50-PAFPN 组合中并未完全保持,在 ResNeXt50-PAFPN 组合中,小目标检测性能(*mAP*_s)为 73.2%,较 ResNeXt50 [18]单独使用时 80.8% 有所下降。这一现象表明,多尺度融合策略必须与主干网络自身的特性相匹配。从特征复杂性与定位精度的权衡来看,ResNeXt50 通过分组卷积和增加基数显著增强了对复杂烟雾形态的语义表征能力,但对于像素占比极低的小目标而言,过强的语义抽象可能导致空间位置信息被“过度平滑”,相比之下,结构更简单的 ResNet50 保留了更多的原始像素特征和几何边缘信息,当与具有双向传播机制的 PAFPN 结合时,ResNet50 提供的底层高分辨率特征能更直接地映射到检测框定位中,从而在极小尺度目标的定位精度上占据优势。此外,从任务重心与泛化能力的表现分析,虽然 ResNet50-PAFPN 在小目标上表现惊人,但其在中目标(64.4%)和大目标(70.4%)上的表现均逊色于最终模型 ResNeXt50-PAFPN,这表明 ResNet50-PAFPN 可能产生了针对小尺度样本的某种程度的特征偏置,而 ResNeXt50-PAFPN 虽然牺牲了部分极小目标的绝对精度,却换取了在全尺度范围内更

稳健的检测效能，其整体 mAP 达到 69.6%，并实现了大、中目标检测的均衡表现。

在注意力机制的研究层面，挤压激励模块的应用带来了复杂的结果。虽然 SE 模块能够使基础检测 (mAP_{50}) 达到 96.1% 的较高水平，但却导致整体 mAP 下降了 4.1 个百分点。经过深入分析，这一现象可能与烟雾检测任务对空间信息的特殊需求密切相关。烟雾在实际场景中的形态和分布具有多样性，对空间位置信息的准确捕捉至关重要，而 SE 模块在强调通道注意力的同时，可能在一定程度上削弱了对空间信息的处理能力，从而影响了整体检测性能。

综合各项指标和实验结果，ResNeXt50-PAFPN 最终脱颖而出，以 69.6% 的 mAP 成为最优方案，较基准模型 ResNet50 提升 2.6 个百分点。其推理速度为 0.221 ms/帧，并实现了大目标 ($mAP_l = 71.5%$) 和中目标 ($mAP_m = 64.8%$) 检测的均衡表现。值得注意的是，ResNeXt50-PAFPN 的小目标检测性能 ($mAP_s = 73.2%$) 较 ResNet50 (72.0%) 仅提升 1.2 个百分点，而 ResNet50-PAFPN 在该指标上表现更优，达到 82.8%，说明不同架构的优势需根据具体任务需求权衡。这一成果为火灾预警系统提供了更为可靠、高效的技术支持，能够更早、更准确地检测到火灾烟雾，为人员疏散和火灾扑救争取宝贵时间。

进一步对模型优化过程进行深入分析。在网络结构设计中，网络深度与特征多样性的平衡至关重要。ResNeXt50 [18] 摒弃了单纯增加网络深度来提升性能的传统方式，而是通过增加基数 ($\text{cardinality} = 32$) 这一创新途径，在不显著增加网络复杂度和计算量的前提下，有效提升了网络的性能。这种方式能够在保证网络对特征进行充分提取和表达的同时，避免因深度增加带来的梯度消失、过拟合等问题，实现了网络深度与特征多样性的良好平衡。多尺度特征融合虽然在显著改善小目标检测性能方面发挥了重要作用，但这一模块的应用需要配合适当的主干网络和训练策略。不同的主干网络具有不同的特征提取能力和结构特点，只有选择与之相适配的主干网络，并采用合理的训练策略，才能充分发挥 PAFPN [20] 在多尺度特征融合方面的优势，实现对小目标的精准检测。

4.2.2. 实验结果可视化

ResNet50 和 ResNeXt50-PAFPN 实验检测结果对比图如图 7 和图 8 所示。通过对比可清晰观察到 ResNet50 与 ResNeXt50-PAFPN 两种模型在烟雾检测任务中的表现差异。在图 7 中，ResNet50 对烟雾目标检测的置信度为 98.2%，而 ResNeXt50-PAFPN 达到了 100.0%。图 8 中，ResNet50 检测置信度为 93.2%，ResNeXt50-PAFPN 则为 98.4%。由此可见，ResNeXt50-PAFPN 在置信度上普遍高于 ResNet50，说明其对烟雾目标的判定更为精准、可靠，误判可能性更低。这得益于 ResNeXt50 通过增加基数提升特征表达能力，以及 PAFPN 有效融合多尺度特征，增强了模型对烟雾特征的捕捉与识别。

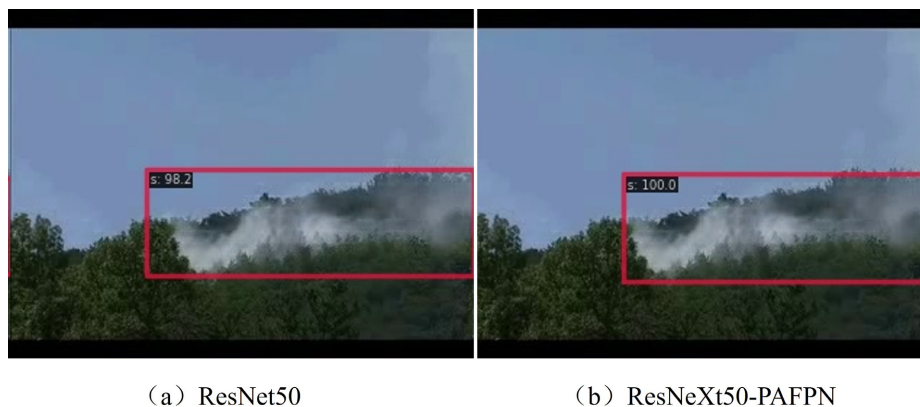


Figure 7. Comparison of experimental detection results
图 7. 实验检测结果对比

从检测框的划定来看, ResNeXt50-PAFPN 的检测框在定位烟雾范围时更为贴合实际烟雾区域。如在图 8 中, ResNet50 的检测框出现了一定程度的偏移, 未能精准覆盖烟雾区域, 而 ResNeXt50-PAFPN 的检测框能更准确地框定烟雾, 这体现了 ResNeXt50-PAFPN 在目标定位上的优势。PAFPN 通过优化特征融合路径, 让模型能更好地利用底层细节信息和高层语义信息, 从而提升了对烟雾目标位置的定位精度。



(a) ResNet50

(b) ResNeXt50-PAFPN

Figure 8. Comparison of experimental detection results

图 8. 实验检测结果对比

综上, 相较于 ResNet50, ResNeXt50-PAFPN 在烟雾检测任务中展现出更高的置信度和更精准的定位效果, 证明了 ResNeXt50 网络结构与 PAFPN 特征融合模块相结合在火灾烟雾检测中的有效性和优越性。本研究的这些发现为后续火灾检测研究提供了重要的参考依据。在未来的研究中, 需充分考虑烟雾的形态多变性和尺度多样性特点, 有针对性地设计专用的特征提取与融合策略, 进一步提升火灾烟雾检测模型的性能, 为火灾预防和安全保障提供更有力的技术支撑。

5. 结束语

针对火灾早期烟雾检测任务中存在的小目标敏感度不足及复杂背景特征提取困难等问题, 本文提出了一种基于改进 Faster R-CNN 的高效火灾烟雾检测模型。本研究首先引入 ResNeXt-50 主干网络, 通过分组卷积策略在不显著增加计算量的前提下, 有效增强了模型对烟雾多尺度特征的表达能力; 其次, 采用 PAFPN 替换传统的单向 FPN 结构, 构建了双向特征传播与多层次融合机制, 显著改善了底层细节向高层传递的效率, 极大缓解了小目标烟雾特征丢失的问题。在自建的多尺度烟雾数据集上的实验结果表明, 优化后的 ResNeXt50-PAFPN 模型平均精度均值(mAP)达到了 69.6%, 较基准 ResNet50 模型提升了 2.6 个百分点, 且维持了 0.221 ms/帧的实时推理速度。本研究成功实现了烟雾检测在准确性与实时性上的良好平衡, 为城市消防与森林防护的早期智能预警提供了可靠的技术支撑。未来, 在此框架基础上进一步融合视频时序动态特征或探索更为先进的网络架构, 将有望持续突破复杂火灾场景下的检测瓶颈。

基金项目

2025 年天津商业大学大学生创新创业训练计划项目(202510069029)。

参考文献

- [1] Girshick, R., Donahue, J., Darrell, T. and Malik, J. (2014) Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation. 2014 *IEEE Conference on Computer Vision and Pattern Recognition*, Columbus, 23-28 June 2014, 580-587. <https://doi.org/10.1109/cvpr.2014.81>
- [2] Girshick, R. (2015) Fast R-CNN. 2015 *IEEE International Conference on Computer Vision (ICCV)*, Santiago, 7-13 December 2015, 1440-1448. <https://doi.org/10.1109/iccv.2015.169>

- [3] Ren, S., He, K., Girshick, R., *et al.* (2015) Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. *NIPS'15: Proceedings of the 29th International Conference on Neural Information Processing Systems*, Volume 1, 91-99.
- [4] He, K., Gkioxari, G., Dollar, P. and Girshick, R. (2020) Mask R-CNN. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **42**, 386-397. <https://doi.org/10.1109/tpami.2018.2844175>
- [5] 祝玉华, 司艺艺, 李智慧. 基于深度学习的烟雾与火灾检测算法综述[J]. 计算机工程与应用, 2022, 58(23): 1-11.
- [6] Zhang, Q., Lin, G., Zhang, Y., Xu, G. and Wang, J. (2018) Wildland Forest Fire Smoke Detection Based on Faster R-CNN Using Synthetic Smoke Images. *Procedia Engineering*, **211**, 441-446. <https://doi.org/10.1016/j.proeng.2017.12.034>
- [7] 殷小芳, 辛月兰, 兰天. 改进 Faster R-CNN 的目标检测算法[J]. 计算机与数字工程, 2021, 49(4): 744-749.
- [8] 王宪保, 朱啸咏, 姚明海. 基于改进 Faster RCNN 的目标检测方法[J]. 高技术通讯, 2021, 31(5): 489-499.
- [9] 刘琳. 基于改进的 Faster R-CNN 目标检测算法研究[D]: [硕士学位论文]. 长春: 吉林大学, 2021.
- [10] 秦常帅. 基于深度学习的森林火灾烟雾检测算法研究[D]: [硕士学位论文]. 上海: 东华大学, 2022.
- [11] 赵坤. 基于深度学习的火灾烟雾检测算法研究及应用[D]: [硕士学位论文]. 西安: 西安理工大学, 2020.
- [12] 彭豪, 李晓明. 基于改进 Faster R-CNN 的小目标检测模型[J]. 电子测量技术, 2021, 44(24): 122-127.
- [13] 汪常建, 丁勇, 卢盼成. 融合改进 FPN 与关联网络的 Faster R-CNN 目标检测[J]. 计算机工程, 2022, 48(2): 173-179.
- [14] 邓姗姗, 黄慧, 马燕. 基于改进 Faster R-CNN 的小目标检测算法[J]. 计算机工程与科学, 2023, 45(5): 869-877.
- [15] 任钰. 基于 Faster R-CNN 的小目标检测研究与应用[D]: [硕士学位论文]. 安庆: 安庆师范大学, 2022.
- [16] Ren, Y., Zhu, C. and Xiao, S. (2018) Small Object Detection in Optical Remote Sensing Images via Modified Faster R-CNN. *Applied Sciences*, **8**, Article No. 813. <https://doi.org/10.3390/app8050813>
- [17] 王忠源. 基于改进 Faster RCNN 的小目标检测技术研究[D]: [硕士学位论文]. 长春: 长春理工大学, 2021.
- [18] Xie, S., Girshick, R., Dollar, P., Tu, Z. and He, K. (2017) Aggregated Residual Transformations for Deep Neural Networks. 2017 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, 21-26 July 2017, 1492-1500. <https://doi.org/10.1109/cvpr.2017.634>
- [19] Lin, T., Dollar, P., Girshick, R., He, K., Hariharan, B. and Belongie, S. (2017) Feature Pyramid Networks for Object Detection. 2017 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, 21-26 July 2017, 2117-2125. <https://doi.org/10.1109/cvpr.2017.106>
- [20] Liu, S., Qi, L., Qin, H., Shi, J. and Jia, J. (2018) Path Aggregation Network for Instance Segmentation. 2018 *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Salt Lake City, 18-22 June 2018, 8759-8768. <https://doi.org/10.1109/cvpr.2018.00913>