Computer Science and Application 计算机科学与应用, 2024, 14(9), 47-55
Published Online September 2024 in Hans. https://www.hanspub.org/journal/csa
https://doi.org/10.12677/csa.2024.149186

基于改进YOLOv8的轻量化火灾检测算法

殷　波

贵州交通职业大学机械电子工程系，贵州　贵阳

收稿日期：2024年7月30日；录用日期：2024年8月29日；发布日期：2024年9月5日

摘　要

本文提出了一种基于改进YOLOv8的轻量化火灾检测算法，旨在优化传统目标检测模型在火灾场景中的实时性和准确性，以满足复杂环境下的高效监测需求。针对YOLOv8算法在资源受限设备上的部署挑战，主要进行了以下两方面的创新工作：(1) 设计了FasterC2f模块，作为YOLOv8算法中C2f模块的有效替代。该模块通过创新的降维策略，对输入特征图进行高效处理，显著降低了模型的参数总量与计算复杂度。(2) 提出了一种轻量化的检测头架构——Lightweight Shared Convolutional Detection (LSCD)，进一步提升模型的检测效率和精度。LSCD通过引入共享卷积机制，有效减少了检测头的参数量，同时增强了特征图之间的全局信息融合能力，确保了即使在复杂多变的火灾场景中也能维持较高的检测精度。实验结果表明，本文提出的算法在火灾检测任务上取得了优异的表现，不仅在模型体积和计算量上实现了显著轻量化，而且保持了与原版YOLOv8相当的甚至更高的检测精度。

关键词

火灾检测，YOLOv8，轻量化，共享卷积

Lightweight Fire Detection Algorithm Based on an Improved YOLOv8

Bo Yin

Department of Mechanical and Electronic Engineering, Guizhou Communications Polytechnic University, Guiyang Guizhou

Received: Jul. 30th, 2024; accepted: Aug. 29th, 2024; published: Sep. 5th, 2024

Abstract

This paper presents a lightweight fire detection algorithm based on improved YOLOv8, aiming to optimize the real-time and accuracy of traditional target detection models in fire scenarios to meet the requirements of efficient monitoring performance in complex environments. In view of the

文章引用: 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.
DOI: 10.12677/csa.2024.149186

deployment challenges of the YOLOv8 algorithm on resource-constrained devices, the following two innovative works are mainly carried out: (1) the FasterC2f module is designed as an effective alternative to the C2f module in the YOLOv8 algorithm. Through the innovative dimension reduction strategy, this module efficiently processes the input feature graph, which significantly reduces the total amount of parameters and the computational complexity of the model. (2) A lightweight detection head architecture called Lightweight Shared Convolutional Detection (LSCD) is proposed to further improve the detection efficiency and accuracy of the model. By introducing the shared convolution mechanism, LSCD effectively reduces the number of detection heads and enhances the global information fusion ability between feature graphs, ensuring that high detection accuracy can be maintained even in complex and variable fire scenarios. The experimental results show that the proposed algorithm achieves excellent performance in the fire detection task, not only achieving significant lightweight in model volume and computational quantity, but also maintaining even higher detection accuracy comparable to the original YOLOv8.

Keywords

Fire Detection, YOLOv8, Lightweight, Shared Convolution

Copyright © 2024 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

火灾，作为一种常见的灾害形式，其突发性和破坏性给人们的生命财产安全带来了极大的威胁。随着现代科技的不断进步，火灾检测技术的发展与应用也日益受到重视。传统的火灾烟雾目标检测往往基于传感器检测与数字图像处理技术实现，由于传感器检测灵敏度受环境温度、湿度等影响较大，且其检测基于阈值设定完成，往往难以达到早期火灾烟雾检测效果[1] [2]。基于数字图像处理的传统计算机视觉识别检测方法则由于特征选取的不确定性、预处理阶段计算的冗杂性，在实际场景的应用中存在精确度不够、适应力不强、检测效率不足的问题[3] [4]。近年来，随着以深度卷积神经网络为核心的深度学习领域的迅速发展，其为实现早期火灾烟雾检测提供了许多新的解决方案[5]-[7]。

YOLO 算法，即 You Only Look Once，是一种将目标检测任务转化为单次回归问题的算法[8]。其核心原理在于，通过一个深度卷积神经网络，模型能够直接在图像的多个位置上预测目标的边界框和所属类别。这一流程避免了传统方法中复杂的区域提议和后续分类步骤，从而实现了高效的实时检测。YOLO 算法采用了端到端的训练方式能够快速处理图像，并在保持较高准确性的同时实现实时目标检测。此外，YOLO 算法还能有效地处理背景信息，减少误检率，提高检测的准确性。虽然 YOLO 系列算法在实时目标检测领域应用广泛，但是由于网络模型较大、参数量较多、运算推理速度较慢的原因，使得其在一些实时性要求高的场景中应用受到限制，在实时火灾检测场景中，这个问题显得尤为突出。因此，如何在保证检测准确率的同时，实现模型的轻量化，成为了当前火灾检测领域亟待解决的关键问题。

由于 YOLO 系列算法的快速发展，YOLOv8 算法在目标检测领域已经展现出强大的性能。论文[9]提出了一种基于改进的 YOLOv8s 模型的番茄自动检测方法，通过深度可分离卷积(DSConv)技术替换普通卷积，减少了特征图的计算量，降低了计算复杂性。论文[10]提出了一种水下目标检测方法，采用轻量级的 FasterNet 模块取代原有的 YOLOv8s 骨干网，减少计算量，提高网络性能。另外将 C2f 的最后一个标

准卷积、瓶颈模块分别替换为 GSConv 和部分卷积，进一步降低了计算量，提高了运算速度。论文[11]出了一个轻量级的模型，YOLOv8-EMSC，采用 C2f_EMSC 取代了 C2f 模块，显著降低了增强的 YOLOv8-EMSC 模型中的模型参数，从而提高了模型推理速度。因此针对火灾目标的实时检测问题，本文提出一种基于 YOLOv8 的轻量化改进算法，在保留 YOLOv8 算法的高精确度的同时，降低算法的复杂度，提高其实时性。

2. 本文总体方案

2.1. YOLOv8 网络结构

YOLOv8 算法是目标检测领域中的一种深度学习算法，它在继承 YOLO 系列优秀性能的基础上，通过改进网络结构和引入新的技术手段，进一步提升了目标检测的精度和速度。该算法的核心思想是通过卷积神经网络进行特征提取和目标检测，它采用了一种独特的双路径预测和紧密连接的卷积网络结构，使得算法能够更好地处理不同大小的目标。此外，YOLOv8 还引入了级联和金字塔的思想，通过多尺度的特征图进行物体检测，能够检测不同大小的物体，并具有较好的尺度适应性。

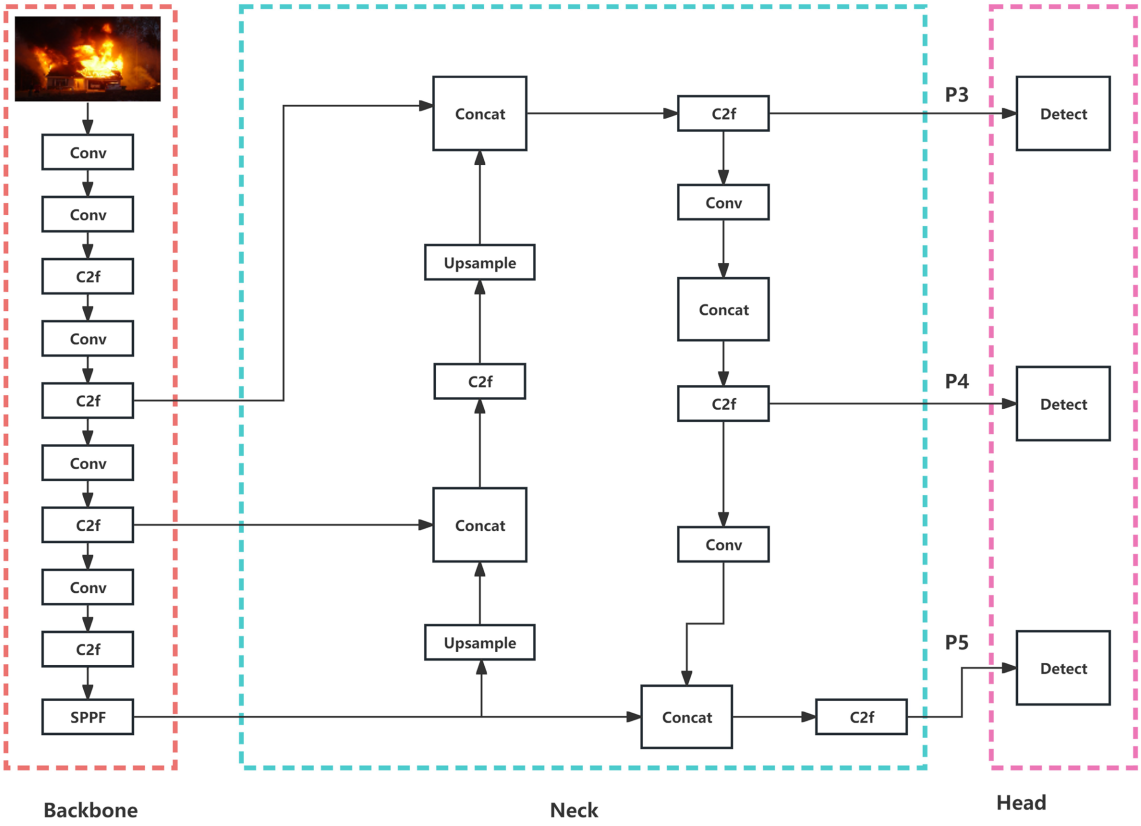

Figure 1. Network structure of YOLOv8
图 1. YOLOv8 网络结构

YOLOv8 算法网络模型如图 1 所示，主要包括三个部分：骨干网络(Backbone)、颈部网络(Neck)和头部网络(Head)。在图 1 所示的网络结构中，主要模块 Conv、C2f、SPPF、Concat、Upsample 和 Detect 模块的操作流程和 YOLOv8 一样，在此本文不做过多说明。YOLOv8 的骨干网络有 Conv、C2f 和 SPPF (spatial pyramid pooling fusion)结构，其中，C2f 模块是对残差特征进行学习的主要模块，通过更多的分支

跨层连接，使模型具有更为丰富的梯度流，形成具有更强特征表示能力的神经网络模块。颈部网络采用路径聚合网络(path aggregation network, PAN)结构，可加强网络对不同缩放尺度对象的特征融合能力。头部网络将分类和检测过程进行解耦，主要包括损失计算和目标检测框筛选，解耦头的设计让每个尺度都有独立的检测器，这种设计有助于更好地处理不同尺度的目标，提高检测的准确性[12]。

2.2. FasterC2f 轻量化模块

FasterNet 网络是一种高效的神经网络架构，其设计目标是在不牺牲准确性的前提下提高计算速度，特别是在视觉任务中表现突出[13]。该网络引入了一种名为部分卷积(PConv)的新型卷积方法，通过仅处理输入通道的一部分来减少计算量和内存访问，从而实现了高效的特征提取。这种设计使得 FasterNet 在多种设备上都能实现比其他网络更快的运行速度，同时保持较高的准确度。

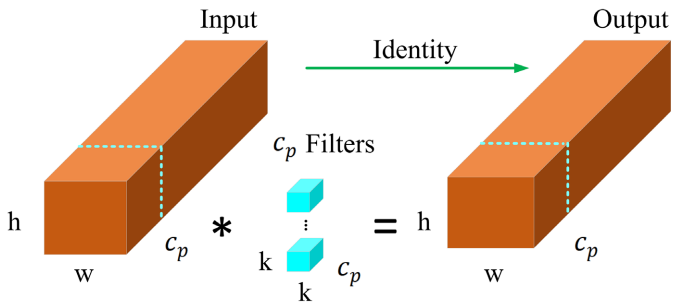

Figure 2. PConv modular structure
图 2. PConv 模块结构

部分卷积(PConv)的操作流程如图 2 所示，PConv 模块仅选择部分输入通道执行普通的 Conv 操作以提取空间特征，而剩余通道则维持原样不进行处理。针对连续或规则的内存访问模式，该模块通常选择首个或末尾连续的c_p通道作为特征图的代表进行运算。在一般情况下，我们可以假定输入与输出特征图的通道数量保持一致。因此，对于 PConv 模块，其计算复杂度(FLOPs)可以表示为：$h*w*k^2*c_p^2$，其中 h 和 w 分别表示特征图的高和宽，k 代表卷积核的大小，而 c_p 则是参与常规卷积计算的通道数量。与传统卷积相比，这里的 c_p 相当于普通卷积中数量的 1/4，从而使得 PConv 的 FLOPs 仅为传统卷积的 1/16，大大降低了计算量。至于内存访问情况，PConv 模块在运算过程中的内存访问量大致为：$h*w*2*c_p$。与常规卷积相比，PConv 的内存访问量仅仅是常规卷积的 1/4，从而无需频繁地进行大规模的内存访问操作，提高了推理速度。

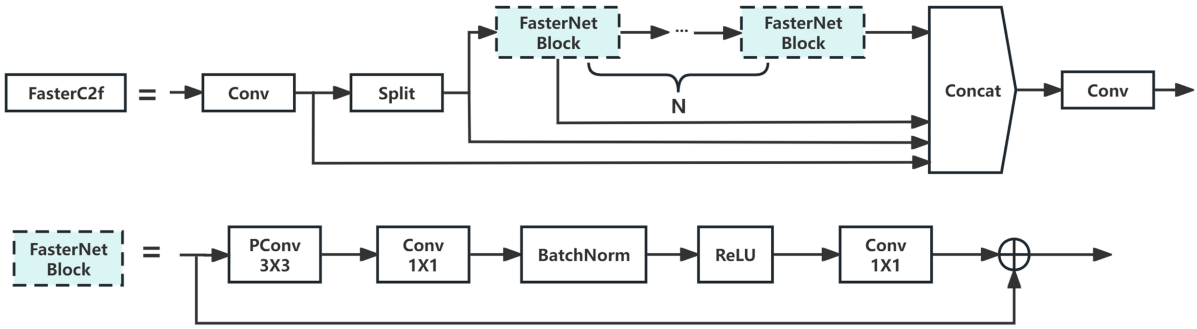

Figure 3. FasterC2f modular structure
图 3. FasterC2f 模块结构

本文利用 FasterNet 网络的部分卷积机制来对 YOLOv8 算法进行改进，来降低冗余计算，实现网络轻量化。由于部分卷积机制主要是用在 FasterNet 网络的 FasterNet Block 模块实现，因为改进措施是将 FasterNet Block 模块引入到 YOLOv8 的 C2f 模块中，用 FasterNet Block 模块代替 C2f 模块中的 Bottleneck 模块，从而构建成新的轻量化模块，命名为 FasterC2f 模块，其网络结构如图 3 所示。最后，用构建的 FasterC2f 模块来替代 YOLOv8 网络结构中的所有 C2f 模块，实现对 YOLOv8 算法的轻量化改进。

2.3. LSCD 轻量化检测头

在深度学习网络中，为了加速模型的收敛速度、提高泛化能力、防止梯度消失或爆炸，一般需要对数据进行归一化处理，常用的归一化处理技术是批归一化 BatchNorm。BatchNorm 具有以下优点：(1) 通过增大学习率，加快模型收敛至最佳参数，减少训练时间；(2) 减少对初始权重的敏感性，帮助模型在不同初始值下保持相似性能；(3) 减少了对如 Dropout 等正则化技术的依赖，降低过拟合风险；(4) 调整激活值分布，防止梯度消失或爆炸，使训练更稳定。虽然 BatchNorm 具有很多优点，但是对 batch size 的大小比较敏感，由于每次计算均值和方差是在一个 batch 上进行的，所以如果 batch size 太小，则计算的均值和方差可能不足以代表整个数据分布，可能导致模型性能下降。本文选择使用 GroupNorm 代替 BatchNorm 主要是因为 GroupNorm 具有更好的适应性、能够保持通道之间的相对关系、提高模型的鲁棒性、在一些场景下性能更佳以及减少对小批量大小的依赖等优点[14]。另外，在文献[15]中已经证明在处理一些目标检测任务中，GroupNorm 可以提升检测头定位和分类的性能。

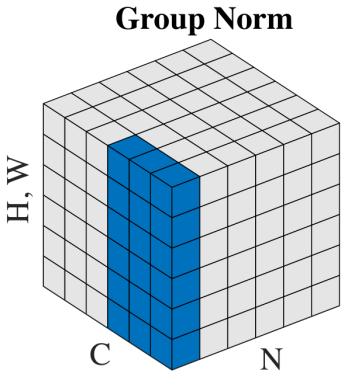

Figure 4. GroupNorm method
图 4. GroupNorm 操作方法

GroupNorm 的操作方法如图 4 所示，其中 N 表示数据的批次轴，C 作为通道轴，H，W 作为空间轴。GroupNorm 对输入大小为[N, C, H, W]的图像，首先将通道设定成若干组(如图 2 蓝色部分)，在每个组内计算方差和平均值，然后根据计算结果对这组内的所有数据进行归一化。GroupNorm 的计算区域与输入的通道数 C 有关，而不是批次大小 N。因此，它不依赖批量大小，在目标检测、视频分类等大型计算机视觉应用中，当计算机内存受到限制或者必须设置较小的样本数量的情况下，群归一化会非常有用。

共享卷积是卷积神经网络(CNN)中的一个核心概念。在 CNN 中，卷积层由一系列的卷积核(或称为过滤器、滤波器)组成，这些卷积核与输入图像的不同位置进行卷积运算，生成对应的特征图。而共享卷积的原理在于，对于同一个卷积核，它在不同位置的卷积运算使用的是相同的权重参数。这意味着，无论卷积核在图像的哪个位置进行卷积，其权重参数都是相同的，从而实现了权重的共享。共享卷积的核心优势在于其权重共享的机制，这一机制显著减少了网络需要训练的参数数量，从而提高了计算效率，并降低了过拟合的风险。通过保持卷积核对整个图像的权重一致性，共享卷积能够更有效地提取图像中的

局部特征，同时保持空间结构信息。此外，其移不变性使得模型对输入数据的平移变换具有鲁棒性，进一步增强了模型的泛化能力。最后，共享卷积支持多通道输入处理，能够充分利用不同通道间的信息，增强模型的表达能力。这些优点共同使得共享卷积在图像处理等领域展现出卓越的性能。

本文设计了一种基于群归一化 GroupNorm 和共享卷积的轻量化检测头 Lightweight Shared Convolutional Detection-LSCD，其网络结构如图 5 所示。LSCD 检测头充分利用了 GroupNorm 和共享卷积的优势，能够在保持特征信息有效融合的情况下，尽量减少计算量和运算复杂度。

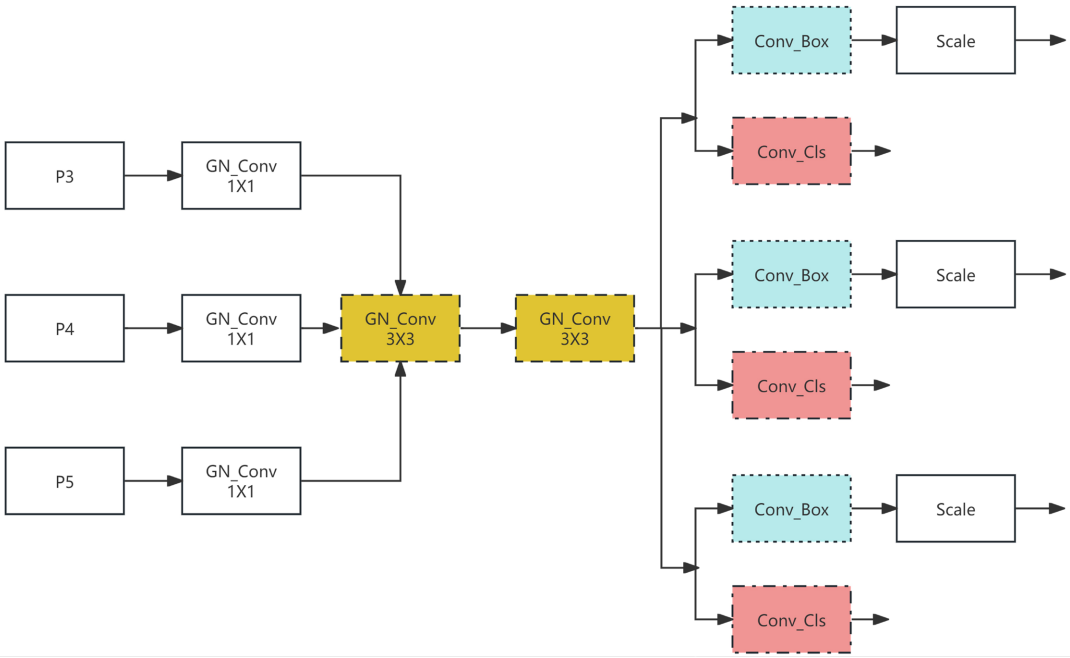

Figure 5. LSCD detection head structure
图 5. LSCD 检测头结构

在图 5 所示的 LSCD 检测头网络结构中，GN_Conv 1 × 1 模块表示的是群归一化 GroupNorm + 卷积 Conv 操作，其中 1 × 1 表示卷积核大小为 1 × 1。两个黄色群归一化卷积模块 GN_Conv 3 × 3 共享权重，三个蓝色预测框卷积模块 Conv_Box 共享权重，三个红色分类卷积模块 Conv_Cls 共享权重。在每个 Conv_Box 模块后面连接着 Scale 模块，Scale 模块主要是一个缩放系数，用来匹配对不同尺度目标的检测。

3. 仿真实验与结果分析

3.1. 实验数据集

在本次实验过程中，我们利用 AI Studio 平台与中国视觉资源网站的协作，精心构建了一个多元化的火灾场景数据集。此数据集广泛涵盖了从森林与草地野火、家庭外部环境中的建筑火灾，到实验室内发生的火灾等多种情景。整个数据集共计收录了 2059 幅图像，这些图像分别捕捉了森林区域、建筑外部、交通地带以及室内环境下的火灾现场。为了进一步促进数据的有效利用与模型训练，我们将该数据集细致地划分为三个独立但相互关联的子集：训练集、验证集与测试集，且严格遵循 8:1:1 的比例分配原则。具体而言，训练集占据了数据集的主要部分，包含 1647 幅图像，用于模型的训练与学习；验证集与测试集则各含 206 幅图像，分别用于模型训练过程中的效果验证与最终的性能评估。图 6 中展示的是该数据集中部分代表性图像，直观呈现了数据集的多样性和丰富性。

Figure 6. Dataset images
图 6. 数据集图片

3.2. 仿真实验

为评价本文所提算法的有效性，以全类平均精度(mean Average Precision, mAP)、每秒钟检测帧数(Frames Per Second, FPS)、每秒 10 亿次浮点运算数(Giga Floating-point Operations Per Second, GFLOPs)、模型层数(Layers)和参数量(Parameters/M)作为评估指标。其中，mAP 作为评价模型检测精度提升的效果，其值越大代表模型检测精度越高；FPS 代表模型每秒钟能够检测图像的张数，数值越大则反映模型实时性越好；GFLOPs、Layers 和 Parameters 反映模型轻量化程度，这些值越小则代表模型轻量化程度越高，对硬件性能要求就越低。

本文采用消融实验对不同模型进行性能对比：以 YOLOv8n 模型作为基础网络，在 YOLOv8n 中将 C2f 模块替换为 FasterC2f 模块得到的模型为 FasterC2f-YOLOv8n，以自研 LSCD 检测头替换 YOLOv8 原检测头的模型为 LSCD-YOLOv8n，将前面两种策略同时应用得到的模型为 FasterC2f-LSCD-YOLOv8n。本文的实验是在下述条件下进行的：Python-3.8.18，torch-1.13.1+cu116，CUDA:0 (NVIDIA GeForce RTX 3080, 10240MiB)。消融实验结果如表 1 所示，所示结果是网络模型进行融合后的数值。

Table 1. Performance comparison of different models
表 1. 不同模型性能对比

模型	FasterC2f	LSCD	mAP50	FPS	GFLOPs	Layers	Parameters/M
YOLOv8n			65.1%	96.7 ★	8.7	168	3.15
FasterC2f-YOLOv8n	√		66.9%	95.6	7.0	198	2.45
LSCD-YOLOv8n		√	68.4%	94.0	6.6	157 ★	2.37
FasterC2f-LSCD-YOLOv8n	√	√	68.5% ★	93.7	5.6 ★	175	2.00 ★

首先，比较 YOLOv8n 和 FasterC2f-YOLOv8n 模型，使用 FasterC2f 模块代替 C2f 模块后得到如下提升：检测精度 mAP50 提升 1.8%，运算量 GFLOPs 降低 19.54%，参数量 Parameters/M 降低 22.22%。从对比结果中可以看出，FasterC2f 模块相较于 C2f 模块能够提升一定的检测精度，最重要的作用的能够大幅

降低模型参数和运算量。

其次，比较 YOLOv8n 和 LSCD-YOLOv8n 模型，使用 LSCD 检测头后得到如下提升：检测精度 mAP50 提升 3.3%，运算量 GFLOPs 降低 24.14%，模型层数 Layers 降低 6.55%，参数量 Parameters/M 降低 24.76%。从对比结果中可以看出，LSCD 检测头相较于 YOLOv8 原始的检测头能够有效地提供检测精度，并且能够更大幅度地降低模型参数和运算量。

最后，综合比较四种模型，可以看出同时使用 FasterC2f 模块和 LSCD 检测头的 FasterC2f-LSCD-YOLOv8n 模型在多个方面都占据最优(表 1 中，★表示在某方面性能对比中最优的模型)。FasterC2f-LSCD-YOLOv8n 模型在检测精度 mAP50、模型参数 Parameters/M 和运算量 GFLOPs 方面都是最优。在 FPS 方面，原始的 YOLOv8 占据最优，说明本文所做的改进并不能提高模型的推理速度，但可以接受的是各模型推理速度相差不多，即最大差值为 3.10% (96.7→93.7)。

Figure 7. Testing situation of the proposed algorithm
图 7. 本文所提算法的测试情况

在测试集上，验证本文所提算法的火灾检测能力。图 7 显示的是本文所提算法在测试集上的火灾检测情况。从图 7 中可以看到，对于不同场景下(公路、森林、家里、工厂)的火灾，本文所提算法都能准确地检测出来，对于同一场景下的多个火灾目标(如图 7 第二行第三列的子图)都能准确识别，对于一些小目标(如图 7 第三行第二列的子图)也能够准确地识别。综合来看，本文所提算法虽然大幅度地降低了原始 YOLOv8 算法的复杂度，实现了轻量化的设计，但依然保留了较高的检测精度，针对不同场景下的多个不同尺寸的火灾目标都能实现精准检测。

4. 结论

本文通过对 YOLOv8 算法进行针对性优化，提出了 LSCD-FasterC2f-YOLOv8 轻量化火灾检测算法，有效解决了传统目标检测模型在火灾监测场景中面临的实时性与准确性平衡难题，特别是在资源受限设备上的部署挑战。通过引入 FasterC2f 模块，实现了对特征图的高效处理，大幅降低了模型参数与计算复

杂度；同时，设计的 LSCD 检测头架构，利用共享卷积机制显著减少了检测头参数，并增强了特征融合能力，确保了复杂火灾场景下的高精度检测。实验结果表明，LSCD-FasterC2f-YOLOv8 算法在保持甚至超越原版 YOLOv8 检测精度的同时，实现了模型体积与计算量的显著轻量化，为火灾预警系统的智能化、实时化升级提供了强有力的技术支持。该算法不仅提升了火灾检测的实时响应速度，还增强了系统的鲁棒性与泛化能力，对于提升城市安防、工业监控等领域的火灾防控水平具有重要意义。未来，随着技术的不断迭代与优化，LSCD-FasterC2f-YOLOv8 算法有望进一步拓展其应用场景，为构建更加安全、智能的社会环境贡献力量。

基金项目

省级科研平台项目资助(黔科合平台人才-CXTD[2021]008)。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)。贵州交通职业大学 2024 年度校级科研项目：基于深度学习的交通多目标实时检测算法研究(2024YB02ZK)。

参考文献

[1] 张思祥, 甘凯, 周围. 提高火灾烟雾传感器检测精度的方法[J]. 传感器与微系统, 2021, 40(1): 151-153.

[2] 程文彬, 杜磊, 刘依依. 无线多传感器数据融合的家庭火灾检测系统设计[J]. 电信科学, 2017, 33(9): 174-181.

[3] 袁斌. 基于图像处理技术的火灾识别方法的应用与研究[J]. 现代电子技术, 2018, 41(13): 43-46.

[4] 李先锋, 徐森, 花义明. 基于 OpenCV 计算机视觉的海事视频船舶火灾烟雾检测技术[J]. 舰船科学技术, 2021, 43(22): 202-204.

[5] 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.

[6] 蔡静, 张瓒, 冉光金, 等. 基于改进 YOLOv5s 的火灾烟雾检测算法研究[J]. 智能计算机与应用, 2023, 13(5): 75-81.

[7] Bo, Y. (2023) A Fire Detection Algorithm Based on Improved YOLOv5. *Fourth International Conference on Mechanical Engineering, Intelligent Manufacturing, and Automation Technology*, Guilin, 20-22 October 2023, 130820R. https://doi.org/10.1117/12.3026293

[8] 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

[9] Yang, G., Wang, J., Nie, Z., Yang, H. and Yu, S. (2023) A Lightweight YOLOv8 Tomato Detection Algorithm Combining Feature Enhancement and Attention. *Agronomy*, **13**, Article 1824. https://doi.org/10.3390/agronomy13071824

[10] Guo, A., Sun, K. and Zhang, Z. (2024) A Lightweight YOLOv8 Integrating Faster Net for Real-Time Underwater Object Detection. *Journal of Real-Time Image Processing*, **21**, Article No. 49. https://doi.org/10.1007/s11554-024-01431-x

[11] Deng, L., Tan, Y., Zhou, J., Wu, S.-Q. and Liu, Q.-Y. (2024) YOLOv8-EMSC: A Lightweight Fire Recognition Algorithm for Large Spaces. *Journal of Safety Science and Resilience*, In Press. https://doi.org/10.1016/j.jnlssr.2024.06.003

[12] 马超伟, 张浩, 马新明, 等. 基于改进 YOLOv8 的轻量化小麦病害检测方法[J]. 农业工程学报, 2024, 40(5): 187-195.

[13] Chen, J., Kao, S., He, H., Zhuo, W., Wen, S., Lee, C., *et al*. (2023) Run, Don't Walk: Chasing Higher FLOPS for Faster Neural Networks. 2023 *IEEE/CVF Conference on Computer Vision and Pattern Recognition* (*CVPR*), Vancouver, 17-24 June 2023, 12021-12031. https://doi.org/10.1109/cvpr52729.2023.01157

[14] Wu, Y. and He, K. (2019) Group Normalization. *International Journal of Computer Vision*, **128**, 742-755. https://doi.org/10.1007/s11263-019-01198-w

[15] Tian, Z., Shen, C., Chen, H. and He, T. (2020) FCOS: A Simple and Strong Anchor-Free Object Detector. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **44**, 1922-1933. https://doi.org/10.1109/tpami.2020.3032166