

## 渐进自适应特征融合的轻量化火焰检测算法研究

李啸澜, 马勇

引用本文

李啸澜, 马勇. 渐进自适应特征融合的轻量化火焰检测算法研究[J]. 计算机科学, 2025, 52(4): 64-73.

LI Xiaolan, MA Yong. Study on Lightweight Flame Detection Algorithm with Progressive Adaptive Feature Fusion [J]. Computer Science, 2025, 52(4): 64-73.

---

### 相似文章推荐 (请使用火狐或 IE 浏览器查看文章)

Similar articles recommended (Please use Firefox or IE to view the article)

[基于多尺度融合注意力的多视角文档图像篡改检测与定位](#)

Multi-view and Multi-scale Fusion Attention Network for Document Image Forgery Localization  
计算机科学, 2025, 52(4): 327-335. <https://doi.org/10.11896/jsjcx.240100142>

[大选择性核双边网络的长尾分布医学图像分类方法](#)

Long-tail Distributed Medical Image Classification Based on Large Selective Nuclear Bilateral-branch Networks  
计算机科学, 2025, 52(4): 231-239. <https://doi.org/10.11896/jsjcx.240700039>

[基于YOLO-Unet组合网络的牛只个体识别方法研究](#)

Research on Individual Identification of Cattle Based on YOLO-Unet Combined Network  
计算机科学, 2025, 52(4): 194-201. <https://doi.org/10.11896/jsjcx.240100144>

[基于虹膜纹理感知的自适应关联学习方法](#)

Adaptive Contextual Learning Method Based on Iris Texture Perception  
计算机科学, 2025, 52(4): 185-193. <https://doi.org/10.11896/jsjcx.250100022>

[基于边缘约束和改进Swin Unetr的复杂器官分割方法](#)

Complex Organ Segmentation Based on Edge Constraints and Enhanced Swin Unetr  
计算机科学, 2025, 52(4): 177-184. <https://doi.org/10.11896/jsjcx.240600007>

# 渐进自适应特征融合的轻量化火焰检测算法研究

李啸澜 马勇

南京理工大学计算机科学与工程学院 南京 210094

(xiaolan.li.work@qq.com)

**摘要** 针对视觉安防系统在边缘计算平台部署火焰检测模型时面临的精度与实时性难以平衡的问题,提出一种渐进自适应特征融合的轻量化火焰检测算法。首先,设计轻量级稀疏卷积算子降低模型计算复杂度与内存访问开销。其次,针对分组卷积的通道间信息交互缺陷,基于残差思想构建长距离上下文特征增强的轻量级特征提取组件。为解决深度骨干网络中特征丢失及背景干扰问题,创新性地提出基于高频增强的轻量级特征强化机制,优化空间域和通道域参数,缓解背景干扰问题。在此基础上,建立特征增强-渐进自适应特征融合框架,促进不同尺度特征图充分融合,提高特征图利用率,增强对多尺度目标的识别效果。实验结果表明,所提方法在实时推理速度最高达到 27.1 FPS 的同时,参数量降低至  $2.1 \times 10^6$ ,较基准模型减少 69.5%,并达到 83.4% 的 mAP@0.5 检测精度,显著优于现有主流方法。

**关键词**:深度学习;计算机视觉;目标检测;轻量化神经网络;特征提取网络;特征融合网络;特征增强

**中图分类号** TP391

## Study on Lightweight Flame Detection Algorithm with Progressive Adaptive Feature Fusion

LI Xiaolan and MA Yong

School of Computer Science and Engineering, Nanjing University of Science and Technology, Nanjing 210094, China

**Abstract** In response to the challenge of balancing accuracy and real-time performance when deploying flame detection models on edge computing platforms for visual security systems, a lightweight flame detection algorithm featuring progressive adaptive feature fusion is proposed. Firstly, a lightweight sparse convolution operator is designed to reduce the model's computational complexity and memory access cost. Subsequently, to address the shortcomings of inter-channel information exchange in grouped convolutions, a lightweight feature extraction component is constructed based on the residual concept, enhancing long-distance contextual features. To tackle the issues of feature loss and background interference in deep backbone networks, an innovative lightweight feature enhancement mechanism based on high-frequency augmentation is proposed, optimizing the parameters in both spatial and channel domains to mitigate background disturbances. On this basis, a feature enhancement-progressive adaptive feature fusion framework is established to facilitate the thorough integration of feature maps at different scales, thereby improving the utilization of feature maps and enhancing the recognition effectiveness of multi-scale targets. Experimental results demonstrate that this method achieves a real-time inference speed of up to 27.1 FPS, reduces the parameter count to 2.1 M, which is a 69.5% reduction compared to the baseline model, and attains a detection accuracy of 83.4% mAP@0.5, significantly outperforming existing mainstream methods.

**Keywords** Deep learning, Computer vision, Object detection, Lightweight neural network, Feature extraction network, Feature fusion network, Feature enhancement

## 1 引言

火灾是一种常见而危险的灾害,其蔓延迅速,突发性强,对人类的生命和财产安全造成了严重威胁。在许多场合,如工厂、仓库、实验室和住宅区等,火焰的实时检测和及时报警可以有效避免财产损失和人员伤亡。因此,成本合理的区域性实时火焰检测方案意义重大。现有的火灾报警系统大多基于各类传感器<sup>[1]</sup>,价格低廉,易于部署,但仅适用于室内环境,无法满足室外开放环境的需求,且往往存在一定的误报

现象<sup>[2]</sup>。近年来,随着计算机视觉技术的进步,基于图像的火焰检测技术得到了不断的发展,该类方法旨在采用图像或视频作为输入,基于火焰图像级特征,对火焰进行检测。相比于基于传感器的检测方法,基于视觉的检测方法受环境影响较小,识别速度较快,适用场景更为广泛。

基于机器学习的火焰检测算法主要依赖于人工构建特征模板来提取感兴趣区域,采用机器学习分类方法,对火焰进行检测。现有算法主要基于单帧图像或多帧连续图像。基于单帧图像的算法<sup>[3]</sup>针对火焰颜色和形状结构特征设计匹配

到稿日期:2024-10-18 返修日期:2025-02-14

基金项目:国家自然科学基金(61773210)

This work was supported by the National Natural Science Foundation of China(61773210).

通信作者:马勇(mayong@njust.edu.cn)

规则,提取感兴趣区域(Region of Interest, ROI),并使用支持向量机<sup>[4]</sup>(Support Vector Machine, SVM)完成分类。此类算法速度快,但存在一定的误检和漏检情况。基于多帧连续图像的算法<sup>[5-7]</sup>在设计单帧图像的火焰特征的基础上,针对帧间火焰运动规律设计动态特征,提取更为精准的ROI,然后使用SVM或极限学习机(Extreme Learning Machines, ELM)完成分类,提高了算法的检测性能,但计算复杂,对部署平台算力资源要求高。基于机器学习的算法难以在大规模部署的低算力平台上兼顾高精度和运行实时性需求,因此使用得较少。

基于深度学习的火焰检测算法基于特征提取网络和特征融合网络,能够自适应搜索网络参数,无需手动设计,且算法鲁棒性强,能够应对复杂的应用场景,在精度和速度方面也表现出一定的优势。基于深度学习的目标检测算法分为一阶段算法和两阶段算法:一阶段算法通过直接回归得到检测框;两阶段算法先划分出感兴趣区域,然后在区域内生成检测框。一阶段算法的检测速度快,可满足实际应用需求。近年来,考虑成本限制,面对低算力平台的模型规模化部署问题,轻量化神经网络逐渐成为研究热点。研究人员在降低模型计算量方面进行了大量研究。MobileNetV1<sup>[8]</sup>, MobileNetV2<sup>[9]</sup>, MobileNetV3<sup>[10]</sup>和ShuffleNet<sup>[11]</sup>基于深度可分离卷积技术(Deepwise Convolution, DWConv)设计轻量化的特征提取网络,有效降低了模型的浮点运算数(Floating Point of Operations, FLOPs)。但深度可分离卷积将空间域和通道域特征提取过程分离,难以捕捉跨通道域的复杂特征,逐像素卷积导致通道域参数冗余,增加了推理时的额外访存次数。FasterNet<sup>[12]</sup>和GhostNet<sup>[13]</sup>指出常规卷积存在特征图冗余和计算量浪费问题,并基于分组卷积技术减少访存次数,增强模型捕捉复杂模式的能力,但仍然无法解决通道间信息交互受限的问题,并引发深层主干网络中浅层特征一定程度的丢失。以上工作局限在降低模型的计算量,忽视了轻量级卷积表征能力受限的问题,无法在维持空间域和通道域特征提取连贯性的同时缓解特征丢失问题,难以适应高度非线性和复杂特征整合的轻量化火焰目标检测任务。

火焰颜色为红、黄色且明亮,在检测场景中,极易和日光、灯光等混淆,存在严重的背景干扰。此外,其边界形状复杂,颜色过渡不显著,前向传播时边界特征极易丢失。同时,考虑到轻量化特征提取网络特征提取能力减弱的问题,在对特征信息进行处理前应对特征图的浅层尺度信息进行特征强化<sup>[14]</sup>。注意力机制参考哺乳动物视觉成像机制,基于通道域或空间域对像素进行微调,以引导模型关注信息更丰富的区域。A2Net<sup>[15]</sup>和SANNet<sup>[16]</sup>使用卷积对空间域像素进行信息聚集,并计算像素级的权重,在通道内完成权重扩散与

分配,引导模型关注特征图内更有效的区域;SENet<sup>[17]</sup>和ECANet<sup>[18]</sup>以通道为单位输入卷积网络,计算通道域的权重,筛选出高信息密度的特征图并进行通道域强化,引导模型关注更有价值的特征图。以上工作局限于通道域或空间域单个维度,限制了模型同时学习跨通道和跨空间的三维信息的能力。SimAM<sup>[19]</sup>和CBAM<sup>[20]</sup>结合空间域和通道域,解决单一维度的权重调整能力不足的问题,引导模型从更高维度关注特征图中的有效信息,形成3D维度的广域参数微调。但這些方法依赖于复杂的神经网络结构,计算量大,访存次数多,难以兼顾轻量化火焰检测网络的实时性和高精度要求。

火焰蔓延迅速,形状特征各异,尺度变化剧烈,随着网络层次的加深,语义信息不断丰富,特征信息不断抽象,常规的无特征融合网络无法应对多尺度检测的挑战。特征融合技术将不同尺度的特征进行融合,构建特征金字塔,增强网络的多尺度检测能力。FPN<sup>[21]</sup>和PAN<sup>[22]</sup>结构基于插值和池化技术,将主干网络提取的不同层次的特征图进行融合,相邻特征图互补语义信息和特征信息,提高了多尺度目标的检测能力。但当网络层次逐渐深入时,简单的上采样和线性组合会引入噪声和偏差,无权的多级特征图导致了空间信息浪费,单向的特征融合则引发了特征不平衡现象。为了缓解以上问题,BiFPN<sup>[23]</sup>, NAS-FPN<sup>[24]</sup>和ASFF<sup>[25]</sup>增加了融合边,并引入加权融合,考虑特征图不同程度的贡献,在FPN结构中使不同层次的特征尽量平衡。然而,以上方法计算量大,无法解决不同层次特征图间尺度性和语义性差异带来的信息矛盾问题,导致了浅层特征的退化,并引入了噪声,不平衡的特征融合也引发了浅层信息的丢失。

针对以上问题,本文构建了一种渐进自适应特征融合的轻量化火焰检测算法。首先,构建基于通道级稀疏连接的轻量化卷积算子,通过分组卷积解决现有算子计算量大和访存次数多的问题。为缓解该算子通道间信息交互的不足,构建了基于残差思想的双分支轻量级特征提取组件,并构造了轻量级骨干网络。针对轻量级骨干网络深层特征丢失、特征直接融合时产生噪声和不同层级特征的信息矛盾问题,提出一种基于高频增强的特征强化-渐进自适应特征融合框架。特征增强模块实现轻量化的广域特征联合学习,对骨干网络特征图进行强化;轻量化的渐进自适应特征融合模块则完成对不同尺度的特征图的自适应降噪融合。

## 2 算法设计

本文基于轻量级骨干网络和特征增强-渐进自适应特征融合框架构建的轻量级火焰检测算法的整体架构如图1所示。

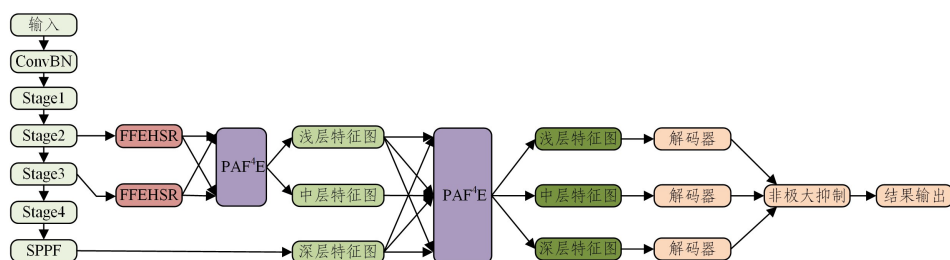


图1 算法的整体架构图

Fig. 1 Overall framework of the proposed algorithm

## 2.1 基于通道级稀疏连接的轻量化卷积算子

面对低成本、低算力平台的火焰检测模型规模化部署问题,轻量化卷积算子的设计面临巨大挑战。模型推理速度不仅受 FLOPs 影响,也与该模型在部署设备上的访存数(Memory Access Cost, MAC)有关。定义 Latency 为算子理论推理速度,可得:

$$Latency = \frac{FLOPs}{\min\left(\frac{FLOPs}{MAC} \times Bandwidth, FLOPs\right)} \quad (1)$$

其中, FLOPs 是算子计算量, MAC 是算子访存次数, Bandwidth 是设备带宽, FLOPs 是设备理论算力。

传统深度可分离卷积算子以增加 MAC 为代价降低了 FLOPs, 导致访存密集操作中推理时间延长, 无法充分发挥其轻量化优势。针对该问题, 本文构建了一种基于通道级稀疏连接的轻量化卷积算子(Channel Sparse Convolution, CSCConv), CSCConv 包含两个分支; 并将特征图划分为激活组和复用组, 激活组参与卷积分支操作, 提取特征, 复用组参与恒等映射分支操作, 完成特征复用。

设卷积算子的输入特征图  $I \in \mathbb{R}^{c_1 \times h_1 \times w_1}$ , 滤波器  $K \in \mathbb{R}^{k \times k}$ , 输出特征图  $O \in \mathbb{R}^{c_2 \times h_2 \times w_2}$ , 激活组比例为  $\alpha$ , 可知:

$$\begin{cases} FLOPs = \alpha(h_1 - k + 1)(w_1 - k + 1)k^2 c_1 c_2 \\ MAC = \alpha(h_1 w_1 c_1 + h_2 w_2 c_2) + \alpha^2 k^2 c_1 c_2 \end{cases} \quad (2)$$

易得:

$$c_1 c_2 = \frac{FLOPs}{\alpha(h_1 - k + 1)(w_1 - k + 1)k^2} \quad (3)$$

因此有:

$$\begin{aligned} MAC &= \alpha(h_1 w_1 c_1 + h_2 w_2 c_2) + \alpha^2 k^2 c_1 c_2 \\ &\geq 2\alpha \sqrt{h_1 w_1 h_2 w_2 c_1 c_2} + \alpha^2 k^2 c_1 c_2 \\ &= 2\alpha \sqrt{h_1 w_1 h_2 w_2 \frac{FLOPs}{\alpha(h_1 - k + 1)(w_1 - k + 1)k^2}} + \\ &\quad \alpha^2 k^2 \frac{FLOPs}{\alpha(h_1 - k + 1)(w_1 - k + 1)k^2} \end{aligned} \quad (4)$$

当且仅当  $c_1 = c_2$  时, MAC 极小化。因此, CSCConv 的输入通道数  $c_1$  和输出通道数  $c_2$  设计为相等时, MAC 最小。

故 CSCConv 的运算过程为:

$$\begin{cases} Y = conv(X_{activation}) \oplus X_{identity} \\ X_{activation} = X[:, :, I] \\ X_{identity} = X[:, :, J] \\ I = \{1, 2, \dots, \lfloor ac_1 \rfloor\} \\ J = \{\lfloor ac_1 \rfloor + 1, \lfloor ac_1 \rfloor + 2, \dots, c_1\} \end{cases} \quad (5)$$

其中,  $X$  是输入特征图,  $Y$  是输出特征图,  $conv$  代表卷积,  $\oplus$  代表 concat 操作。

图 2 和图 3 分别展示了常规 DWConv 和 CSCConv 的运算过程。

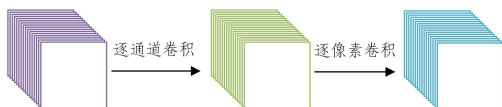


图 2 DWConv 运算过程图

Fig. 2 DWConv process diagram

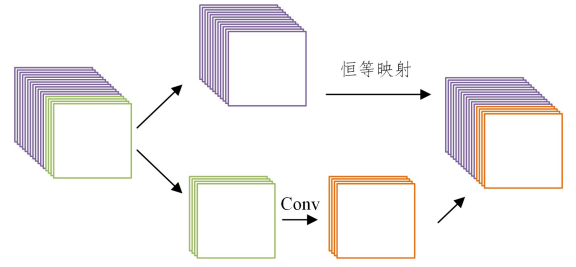


图 3 CSCConv 运算过程图

Fig. 3 CSCConv process diagram

对于常规卷积算子 Conv, 由计算可得:

$$\begin{cases} FLOPs = (h_1 - k + 1)(w_1 - k + 1)k^2 c_1 c_2 \\ MAC = h_1 w_1 c_1 + h_2 w_2 c_2 + k^2 c_1 c_2 \end{cases} \quad (6)$$

对于 DWConv 而言, 由于缺乏通道间信息的交互, 需要使用  $c'$  ( $c' \geq 6c$ ) 个滤波器, 以维持与 Conv 相当的精度。由计算可得:

$$\begin{cases} FLOPs = (h_1 - k + 1)(w_1 - k + 1)k^2 c' \\ MAC = c'(h_1 w_1 + h_2 w_2) + k^2 c' \end{cases} \quad (7)$$

对于 CSCConv, 则有:

$$\begin{cases} FLOPs = \alpha^2 (h_1 - k + 1)(w_1 - k + 1)k^2 c_1 c_2 \\ MAC = \alpha(h_1 w_1 c_1 + h_2 w_2 c_2) + \alpha^2 k^2 c_1 c_2 \end{cases} \quad (8)$$

当  $c_1 = c_2 = c$  时, 由计算可知, DWConv 和 Conv 的访存次数比值约为  $c'/c$  ( $c' \geq 6$ ), 对于 IO 能力极为有限的低算力平台而言, 其会对推理延迟产生严重的影响。CSCConv 与 Conv 的访存次数比值约为  $\alpha$  ( $\alpha < 1$ ), 与 Conv 的 FLOPs 比值约为  $\alpha^2$ 。因此, CSCConv 解决了 DWConv 访存密集的问题, 同时相比于 Conv 有效降低了计算量。

## 2.2 轻量化特征提取组件

基于 CSCConv, 本文设计了基于残差连接的轻量化特征提取组件 BasicBlock。为确保骨干网络随着层次加深, 形成的映射关系优于浅层骨干网络, 并缓解梯度消失导致的高层网络参数固化问题, BasicBlock 基于残差连接思想, 采用双分支设计(特征提取分支和上下文特征增强分支)。BasicBlock 的运算过程如图 4 所示。

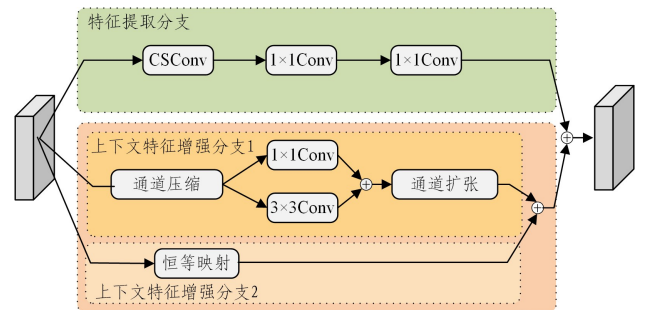


图 4 BasicBlock 结构图

Fig. 4 BasicBlock structure diagram

特征提取分支采用倒残差结构, 先使用  $3 \times 3$  的 CSCConv 提取火焰特征。由于 CSCConv 的通道间信息交互受限, 因此采用  $1 \times 1$  卷积进行通道扩张, 增强特征提取能力, 促进通道间信息的交互。最后, 使用  $1 \times 1$  卷积对通道数进行压缩, 得

到特征提取分支的输出。

上下文特征增强分支采用残差结构。火焰边界形状复杂,高频信号占比多,浅层特征信息在网络深层容易丢失。因此,利用残差思想将特征图的浅层特征作为先验信息,可以有效缓解深度骨干网络在特征提取中信息丢失的现象。分支1使用 $1 \times 1$ 卷积对输入进行通道压缩,先采用 $3 \times 3$ 卷积分支强化语义特征,增强目标框定位能力,然后并联 $1 \times 1$ 卷积强化细节特征,细化目标框边界。两种卷积的输出逐像素相加,最后使用 $1 \times 1$ 卷积恢复特征图维度。分支2完成恒等映射,优化梯度流动。将分支1和分支2结果逐像素相加,得到上下文特征增强分支的输出。

为了强化骨干网络中全局信息的流动,维持长距离上下文信息和局部信息的有效平衡,增大组件感受野,特征增强分支的桥接位置选择在组件的输入点和输出点。

BasicBlock 的运算如下:

$$\begin{cases} \mathbf{Y}_{\text{extraction}} = \text{conv1}(\text{conv1}(\text{CSConv}(\mathbf{X}_{\text{in}}))) \\ \mathbf{Y}_{\text{enhancement}} = \text{bottle\_out}(\mathbf{Y}_1 + \mathbf{Y}_2) + \mathbf{X}_{\text{in}} \\ \mathbf{Y}_1 = \text{conv1}(\text{bottle\_in}(\mathbf{X}_{\text{in}})) \\ \mathbf{Y}_2 = \text{conv2}(\text{bottle\_in}(\mathbf{X}_{\text{in}})) \\ \mathbf{Y}_{\text{out}} = \mathbf{Y}_{\text{extraction}} + \mathbf{Y}_{\text{enhancement}} \end{cases} \quad (9)$$

其中, $\mathbf{X}_{\text{in}}$ 是组件输入, $\text{conv1}$ 是 $1 \times 1$ 卷积, $\mathbf{Y}_{\text{extraction}}$ 是特征提取分支输出, $\mathbf{Y}_1$ 和 $\mathbf{Y}_2$ 分别是上下文特征增强分支2中 $1 \times 1$ 卷积和 $3 \times 3$ 卷积的输出, $\text{bottle\_in}$ 是通道压缩操作, $\text{bottle\_out}$ 是通道扩张操作, $\mathbf{Y}_{\text{out}}$ 是组件输出。

### 2.3 轻量化骨干网络

受 CSPDarkNet 的启发,本文基于 BasicBlock 设计了32倍下采样的骨干网络。首先使用 $3 \times 3$ 卷积对RGB三通道图像进行维度膨胀,然后使用4个由 ConvBN 和 BasicBlock 组成的阶段(stage)进行下采样和特征提取,最后采用空间金字塔池化对骨干网络进行尺度不变性增强。骨干网络结构如表1所列,设输入图像为 $640 \times 640 \times 3$ 的RGB图像。

表1 骨干网络结构

Table 1 Backbone network structure

Stage	Input	OutPut	Depth
ConvBN	$640 \times 640 \times 3$	$320 \times 320 \times 32$	—
Stage1	$320 \times 320 \times 32$	$160 \times 160 \times 64$	2
Stage2	$160 \times 160 \times 64$	$80 \times 80 \times 128$	4
Stage3	$80 \times 80 \times 128$	$40 \times 40 \times 256$	8
Stage4	$40 \times 40 \times 256$	$20 \times 20 \times 512$	2
SPPF	$20 \times 20 \times 512$	$20 \times 20 \times 512$	—

表1中,Stage代表网络阶段,Input代表输入特征图尺寸,Output代表输出特征图尺寸,Depth代表Stage1—Stage4中BasicBlock的重复次数。为适应火焰特征从浅层细节特征到深层语义特征的过渡,在Stage1—Stage3中,Stage的Depth被递进地设置为2,4,8;为了在最深层进行有效的信息筛选和压缩,避免过拟合的产生,Stage4的Depth被设置为2。

骨干网络的结构如图5所示。

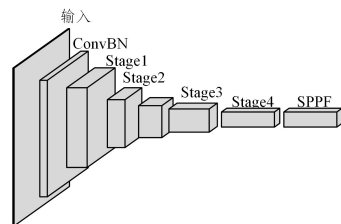


图5 骨干网络结构图

Fig. 5 Diagram of backbone network structure

### 2.4 基于高频增强的特征强化模块

在火焰检测任务中,火焰容易与日光、灯光等无关信息产生混淆,在轻量级骨干网络的前向传播中,边界特征极易丢失。考虑到火焰的整体形状、颜色渐变和光度变化通常属于低频成分,而边界处的细微纹理、光度波动和颜色突变属于高频成分,为了提高模型对高效信息的关注率,同时考虑到现有特征强化方法计算复杂和参数量大的问题,本文在火焰检测模型中构建了一种基于高频信号增强的火焰特征强化模块(Flame Feature Enhancement with High-Frequency Signal Reinforcement, FFEH-SR),用于对特征进行激活,提高骨干网络的信息利用率,实现对网络参数的微调优化,同时增强通道间信息的交互。

化学热力学表明,燃烧反应中,局部的热能和光辐射会发生显著改变。受该理论启发,为特征图定义像素级的能量函数,用以量化像素显著性并区分像素级差异。设计二值化标签 $y_i = 1, y_j = -1$ ,定义线性决策边界 $y = wx + b$ ,其中 $w$ 和 $b$ 是决策边界的待优化对象。定义能量函数为:

$$F = (y_i - y)^2 + \frac{1}{M-1} \sum_{i=1}^{M-1} (y_j - y_i)^2 \quad (10)$$

其中, $x$ 和 $x_i$ 分别代表被分析像素值和当前通道内的其他像素值,且它们属于输入特征图的同一通道; $M$ 代表当前通道内的总像素数,且有:

$$\begin{cases} y = wx + b \\ y_i = wx_i + b \end{cases} \quad (11)$$

式(10)的能量函数值越小,表明线性决策边界对于高频和低频信号的区分效果越好。将 $y$ 和 $y_i$ 替换,然后引入二值化标签 $y_i = 1$ 和 $y_j = -1$ ,并添加正则项约束,可得:

$$F = \lambda w^2 + (1 - (wx + b))^2 + \frac{1}{M-1} \sum_{i=1}^{M-1} (-1 - (wx_i + b))^2 \quad (12)$$

对 $w$ 和 $b$ 分别求偏导,令偏导为0,经整理后联立可得:

$$\begin{bmatrix} \lambda + x^2 + \frac{\sum_{i=1}^{M-1} x_i^2}{M-1} & x + \frac{\sum_{i=1}^{M-1} x_i}{M-1} \\ x + \frac{\sum_{i=1}^{M-1} x_i}{M-1} & 2 \end{bmatrix} \begin{bmatrix} w \\ b \end{bmatrix} = \begin{bmatrix} x - \frac{\sum_{i=1}^{M-1} x_i}{M-1} \\ -1 \end{bmatrix} \quad (13)$$

使用Cramer法则求解该方程,整理结果,可得:

$$\begin{cases} w = \frac{(3x - \frac{\sum_{i=1}^{M-1} x_i}{M-1})(M-1)^2}{(M-1)^2(2\lambda + x^2) + 2 \sum_{i=1}^{M-1} (x_i^2 - xx_i) - (\frac{\sum_{i=1}^{M-1} x_i}{M-1})^2} \\ b = \frac{-(M-1)^2(\lambda + 2x^2) - \sum_{i=1}^{M-1} x_i^2 + (\frac{\sum_{i=1}^{M-1} x_i}{M-1})^2}{(M-1)^2(2\lambda + x^2) - 2 \sum_{i=1}^{M-1} (xx_i + x_i^2) - (\frac{\sum_{i=1}^{M-1} x_i}{M-1})^2} \end{cases} \quad (14)$$

将结果代入能量函数,经过整理,并将  $w$  和  $b$  代入  $F$ , 易得:

$$\begin{cases} w = \frac{2(x-\mu)}{(x-\mu)^2 + 2\sigma^2 + 2\lambda} \\ b = -\frac{w(x+\mu)}{2} \\ F_{\min} = \frac{4(\sigma^2 + \lambda)}{(x-\mu)^2 + 2(\sigma^2 + \lambda)} \end{cases} \quad (15)$$

其中,  $\mu$  和  $\sigma$  分别是目标像素所在通道内其他像素的均值和方差, 满足:

$$\begin{cases} \mu = \frac{1}{M-1} \sum_{i=1}^{M-1} x_i \\ \sigma^2 = \frac{1}{M-1} \sum_{i=1}^{M-1} (x_i - \mu)^2 \end{cases} \quad (16)$$

$\lambda$  是为了防止过拟合引入的正则项, 取典型值 0.0001。

计算时, 假设单通道内像素服从同一分布<sup>[26]</sup>。

能量函数值越小, 该像素蕴含的有效信息越多。被分析神经元的权重因子  $\beta$  可以表示为:

$$\beta = \frac{1}{F_{\min}} = \frac{(x-\mu)^2}{4(\sigma^2 + \lambda)} + 0.5 \quad (17)$$

为了对权重因子的数值范围进行约束, 避免产生过大的数值, 使用 sigmoid 激活函数对权重因子处理, 并将其作为增益或减益与被分析神经元相乘。

$$\tilde{\mathbf{X}} = \mathbf{X} \times \text{sigmoid}(\beta) \quad (18)$$

FFEHSR 特征强化方法的计算过程如图 6 所示。就骨干网络结构而言, Stage1 层次最浅, 细节特征丰富, 无需进行特征强化; Stage4 层次最深, 语义信息丰富, 难以进行特征强化。因此, 在网络中通道数与空间深度适中的 Stage2 和

Stage3 之后引出连接, 接入 FFEHSR 模块。

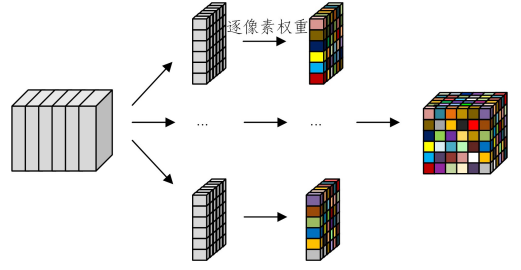


图 6 FFEHSR 机制示意图

Fig. 6 FFEHSR mechanism diagram

## 2.5 特征增强-渐进自适应特征融合框架

火焰是一种尺度变化剧烈的目标, 因此单一尺度的特征信息无法对其进行有效描述, 必须采用特征融合策略。传统的 FPN 和 PAN 结构仅使用简单的上下采样操作处理特征图尺度问题, 非相邻层存在较大的信息维度差距, 直接叠加会导致信息丢失, 引入噪声和偏差。此外, 传统特征融合框架生成多级特征图, 但只在特定尺度的特征图上筛选出前景目标, 导致特征图空间信息利用不充分。为应对上述问题, 同时得益于特征强化模块提供的高质量特征图, 本文构建了一种特征增强-渐进自适应特征融合框架 (Progressive Adaptive Feature Fusion Framework with Feature Enhancement, PAF<sup>4</sup>E)。PAF<sup>4</sup>E 采用渐进式的自适应特征融合设计, 缓解了不同层级特征的信息矛盾问题, 减少了融合过程中产生的噪声; 设计全连接式融合方案, 减少了浅层特征信息的丢失。其结构如图 7 所示。

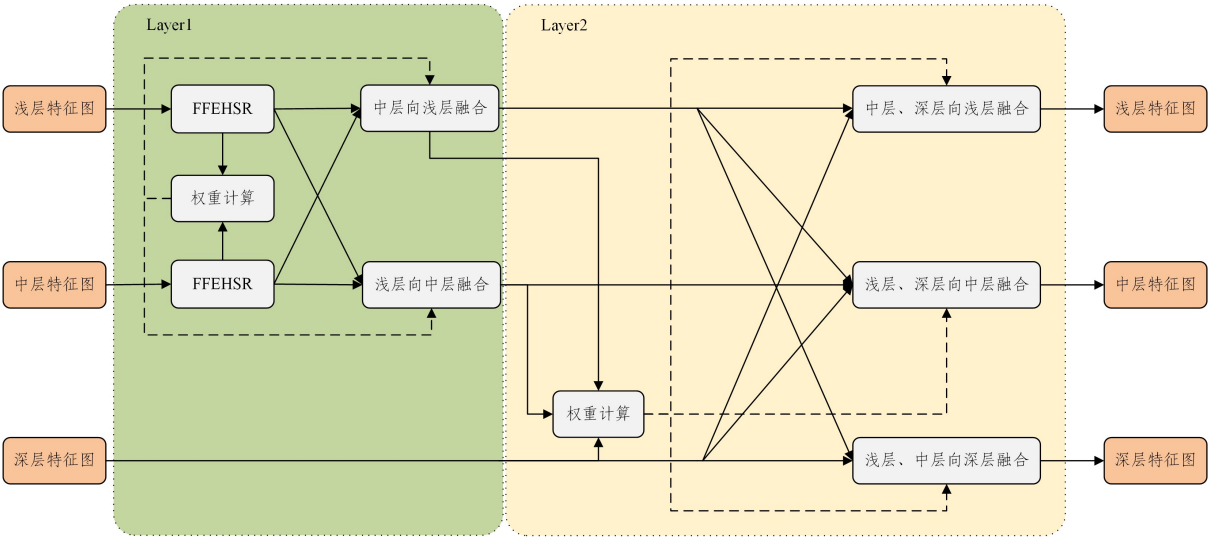


图 7 PAF<sup>4</sup>E 结构图

Fig. 7 PAF<sup>4</sup>E structure diagram

PAF<sup>4</sup>E 分为两个 Layer, 每个 Layer 分为权重计算阶段和特征融合阶段。

在 Layer1, 接收特征增强模块输出的两个层次的特征图, 在空间域将其转换为相同尺寸, 使用  $1 \times 1$  卷积计算权重并归一化; 然后基于权重, 分别完成浅层特征图向中层特征图的预融合以及中层特征图向浅层特征图的预融合; 最后使用  $3 \times 3$  卷积抑制特征退化。其运算过程如下:

$$\begin{cases} \mathbf{Y}_1 = \text{Enhance}(\mathbf{X}_{\text{stage2}}) \\ \mathbf{Y}_2 = \text{Enhance}(\mathbf{X}_{\text{stage3}}) \\ \mathbf{Y}_1' = \text{resize}(\mathbf{Y}_1) \\ \mathbf{Y}_2' = \text{resize}(\mathbf{Y}_2) \\ \mathbf{W} = \text{softmax}(\text{concat}(\text{conv1}(\mathbf{Y}_1'), \text{conv1}(\mathbf{Y}_2'))) \\ \mathbf{Y}_{\text{pre}}^i = \mathbf{Y}_i \otimes \mathbf{W}_i, i \in \{1, 2\} \\ \mathbf{Y}_{\text{out}}^i = \text{conv3}(\mathbf{Y}_{\text{pre}}^i + \text{resize}(\mathbf{Y}_{\text{pre}}^{3-i})) \end{cases} \quad (19)$$

其中,  $\mathbf{X}_{\text{stage2}}$  和  $\mathbf{X}_{\text{stage3}}$  分别代表骨干网络 Stage2 和 Stage3 输出

的特征图; Enhance 代表特征增强操作;  $Y_1$  和  $Y_2$  分别代表  $X_{stage2}$  和  $X_{stage3}$  经 FFEHSR 增强后的结果; resize 代表上采样或下采样操作,用于统一特征图大小,resize 的实际操作取决于所处理的特征图层次;  $W$  代表合并且归一化后的权重;  $W_i$  代表  $W$  中对应第  $i$  层次的权重张量,  $i=1$  时代表浅层次,  $i=2$  时代表中层次;  $\otimes$  代表逐像素乘;  $+$  代表逐像素加;  $Y_{pre}^i$  代表预融合后第  $i$  层次的输出;  $Y_{out}^i$  代表抑制特征退化后第  $i$  层次的输出,即 PAF<sup>E</sup> 的输出。

在 Layer2,接收 Layer1 输出的两个层次的特征图和骨干网络输出的特征图,在空间域 resize 为相同尺寸,并依次进行权重计算、预融合并抑制特征退化,将输出结果送入检测头。

### 3 实验和结果分析

#### 3.1 数据集构建

在训练目标检测模型时,数据集是至关重要的因素,其质量与规模对目标检测模型的性能有直接的影响。本文研究的目标检测算法基于全监督式的深度学习方法,需要大量的标注数据以训练模型。在火焰检测领域,尚无高质量的公开数据集,因此通过 Kaggle 平台、MIVIA 数据集等互联网资源收集火焰图片,筛选出 3000 张图片,使用 Labelimg 工具进行标注,并按照约 8:1:1 的比例划分的训练集、验证集和测试集,其分别有 2443,246,311 张图片。

基于标注好的数据集,采用 K-means 聚类算法对锚框大小进行聚类,在深层次、中层次和浅层次 3 个尺度分别聚类出 3 个锚框,聚类结果为:  $([30, 35], [61, 56], [87, 107])$ ,  $([184, 122], [124, 209], [215, 195])$ ,  $([374, 169], [256, 329], [437, 360])$ 。

#### 3.2 模型训练

在高性能服务器上进行识别模型训练,实验环境如表 2 所列。

表 2 服务器环境

名称	信息
CPU	Intel Xeon E5-2630
GPU	NVIDIA RTX2080Ti
内存	110 GB
Python	3.8.16
Pytorch	1.8.0

模型训练 200 个周期,批次大小设置为 16,输入图片大小统一缩放为  $640 \times 640$ ,使用随机梯度下降法作为优化器,学习率设置为 0.01,学习率动量设置为 0.937,借助高性能服务器,在数据集上完成识别模型训练。

部分典型检测结果如图 8 所示。



图 8 目标检测效果图

Fig. 8 Target detection effect

在面对尺度较小的目标时,本模型也具有较好的检测效果。所提模型对于小目标的部分检测结果如图 9 所示。



图 9 小目标检测效果图

Fig. 9 Small target detection effect

在实际场景中,背景中可能含有灯光和日光等易混淆的目标,检测模型可能会受到背景干扰,本文也挑选了符合以上特征的图片来进行测试,检测效果如图 10 所示。



图 10 易混淆背景下的检测效果图

Fig. 10 Detection effect under confusion background

#### 3.3 评价指标

在目标检测领域中,常用的评价指标有精确率(Precision)、召回率(Recall)和平均精度均值(mAP)等。精确率,即查准率,表示预测为正确且和事实相符的数量与所有预测为正确的样本数量的比值:

$$Precision = \frac{TP}{TP + FP} \quad (20)$$

召回率,即查全率,表示预测为正确且和事实相符的数量与所有该种待检测目标数量的比值:

$$Recall = \frac{TP}{TP + FN} \quad (21)$$

其中,  $TP$  指正样本被成功识别的样例数量;  $FP$  指负样本被错误识别的数量;  $FN$  指正样本被错误识别的样例数量。

尽管精确率和召回率都可以衡量模型的性能,但二者耦合,无法全面衡量模型的性能,因此引入平均精度均值。以 Recall 为横坐标, Precision 为纵坐标,绘制出曲线并对横坐标进行积分,得到 AP 值。将所有类的 AP 值取平均,即得到平均精度均值 mAP。本文以 mAP@0.5 作为性能指标。

#### 3.4 消融实验结果与分析

为了验证所提方法中各个模块的有效性,针对骨干网络和特征融合框架进行消融实验。采用 CSPDarknet53+PANet 为基准模型,分别去掉本文方法中的轻量级骨干网络和 PAF<sup>E</sup>; 同时,为验证长距离上下文特征增强分支的有效性,在 BasicBlock 中分别采用长距离上下文特征增强分支、短距离上下文特征增强分支和单一恒等映射特征增强分支策略,将其分别命名为 LongBasicBlock, ShortBasicBlock 和 IdentityBasicBlock; 为证明 FFEHSR 模块的有效性,在 PAF<sup>E</sup> 中去掉

FFEHSR 模块。消融实验结果如表 3 所列。

表 3 消融实验  
Table 3 Ablation experiment

模型	mAP@0.5	参数量	计算量/FLOPs
baseline	80.9	7 012 822	16.0
baseline+IdentityBasicBlock	80.4(-0.5)	3 847 782	8.0
baseline+ShortBasicBlock	80.5(-0.4)	4 150 430	8.8
baseline+LongBasicBlock	82.2(+1.3)	3 984 222	8.4
baseline+PAF <sup>1</sup> E(no-FFEHSR)	81.9(+1.0)	5 106 435	13.2
baseline+PAF <sup>1</sup> E	82.5(+1.6)	5 106 435	13.2
baseline+LongBasicBlock+PAF <sup>1</sup> E	83.4(+2.5)	2 141 323	5.8

分析表 3 可知,相比于基线模型,在引入轻量化骨干网络改进后,参数量和计算量均大幅下降;但由于通道间信息交互的不足和浅层细节信息的退化,mAP@0.5 有 0.5% 的下降。对 BasicBlock 引入短距离上下文特征增强分支后,mAP@0.5 依然存在 0.4% 的下降;引入长距离上下文特征增强分支后,mAP@0.5 提升 1.3%。这充分说明轻量级骨干网络在层次加深时存在浅层细节信息的退化,且短距离分支和 BasicBlock 中串联的两层 1×1 卷积功能基本相同,无法有效增强通道间信息的交互。而长距离上下文特征增强分支通过将浅层特征作为先验信息,缓解了通道稀疏卷积的通道间信息交互不足的问题和深度骨干网络中信息丢失的现象。对 baseline 引入 PAF<sup>1</sup>E 改进,模型参数量和计算量有了一定的下降,mAP@0.5 提升了 1.6%,表明了渐进式自适应特征融合框架在火焰检测任务中的高效性,并且其能够有效缓解特征融合时产生的噪声与不同层级特征间的信息矛盾。在去掉 FFEHSR 后,mAP@0.5 仅提升 1.0%,但模型参数量和计算量不变,这充分说明 FFEHSR 具有不引入额外参数的优秀特性和其在特征强化工作中的有效性。

### 3.5 对比实验结果与分析

为了验证所提算法的先进性,选取先进检测方法构建火焰检测网络来进行对比实验,选取的算法包括 ResNet<sup>[27]</sup> 系列、VGG<sup>[28]</sup> 系列、YOLO<sup>[29]</sup> 系列、ShuffleNet2<sup>[30]</sup>、MobileNetv3、GhostNet 和 MobileNeXt<sup>[31]</sup>。对比实验结果如表 4 所列。

表 4 对比实验  
Table 4 Comparative experiment

模型	mAP@0.5	参数量	计算量/FLOPs
ResNet50	82.5	26 967 094	73.9
ResNet101	83.2	45 933 110	134.5
VGG-13	71.9	5 149 910	51.8
VGG-16	76.4	6 477 654	68.8
VGG-19	78.2	7 199 062	80.8
YOLOv3	81.9	9 301 222	23.7
YOLOv4	81.6	10 062 822	21.6
YOLOv5	80.9	7 012 822	16.0
YOLOv6	82.5	4 629 007	11.3
ShuffleNetv2	76.4	2 865 718	5.8
MobileNetv3	81.0	4 900 516	9.0
GhostNet	83.2	3 675 726	8.0
MobileNeXt	81.7	4 420 630	8.5
本文算法	<b>83.4</b>	<b>2 141 323</b>	<b>5.8</b>

对比实验结果表明,相较于 ResNet50, ResNet101,

YOLOv3, YOLOv4 和 YOLOv5 等算法,本文算法在 mAP@0.5 分别提高 0.9%, 0.2%, 1.5%, 1.8% 和 2.5% 的情况下,参数量分别降低 92.1%, 95.3%, 77.0%, 78.7% 和 69.5%,计算量分别降低 92.2%, 95.7%, 75.5%, 73.1% 和 63.8%。这充分说明了本文算法在提升 mAP@0.5 指标的同时,可以大幅降低计算量和参数量。相较于 YOLOv6, ShuffleNetv2, MobileNetv3, GhostNet 和 MobileNeXt 等轻量化算法,所提算法在参数量分别降低 53.7%, 25.3%, 56.3%, 41.8% 和 51.6%,计算量分别降低 48.7%, 0.35.6%, 27.5% 和 31.8% 的情况下,mAP@0.5 指标分别提高 0.9%, 7.0%, 2.4%, 0.2% 和 1.7%。这充分说明本文算法在进一步降低参数量和计算量的同时,可以有效提高 mAP@0.5 指标。

图 11 中,横坐标代表参数量,纵坐标代表 mAP@0.5,圆点面积代表计算量,靠近左上且面积小的圆点代表对应算法参数量和计算量更低,且 mAP@0.5 更高。

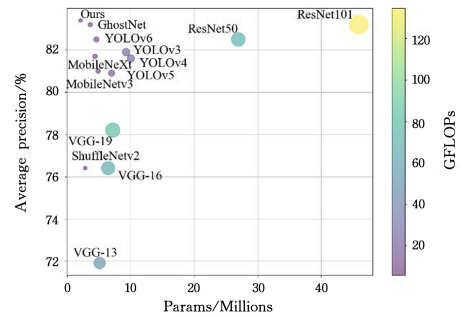


图 11 对比实验结果

Fig. 11 Comparative experiment results

为进一步证明所提算法在火焰检测领域的先进性,选取了文献[32-34]中的火焰检测算法进行对比实验,结果如表 5 所列。

表 5 相关火焰检测算法对比实验

Table 5 Comparative experiment on related flame detection algorithms

模型名称	mAP@0.5	参数量	计算量/FLOPs
文献[32]算法	82.9	5 331 649	10.6
文献[33]算法	83.1	2 176 492	12.5
文献[34]算法	82.6	3 357 092	11.7
本文算法	83.4	2 141 323	5.8

对比实验结果表明,相较于文献[32]、文献[33]和文献[34]提出的算法,本文算法在 mAP@0.5 分别提高 0.5%, 0.3% 和 0.8% 的情况下,参数量分别降低 59.8%, 1.6% 和 36.2%,计算量分别降低 45.3%, 53.6% 和 50.4%。这充分说明了本文算法在提升 mAP@0.5 指标的同时,能够一定程度地降低计算量和参数量。

### 3.6 嵌入式平台部署实验

为验证所提出算法的轻量化特性,共选用 4 种典型的嵌入式平台计算资源,包含 Intel N5105 CPU 运算模块、Intel N5105 GPU 运算模块、NVIDIA MAXWELL 架构 GPU 和 4 核 ARM A57 CPU,进行嵌入式边缘计算设备的部署实验。与 3.4 节中算法保持一致,进行模型推理时间的消融实验,实验结果如表 6 所列。

表6 模型推理时间消融实验

Table 6 Ablation experiment on model inference time

模型名称	(ms)			
	N5105CPU 推理时间	N5105GPU 推理时间	NVIDIA Maxwell™ GPU 推理时间	4核 ARM A57 推理时间
baseline	355.5	52.9	130.4	1273.3
baseline+IdentityBasicBlock	214.6	33.0	79.6	717.2
baseline+ShortBasicBlock	231.9	37.3	94.4	824.8
baseline+LongBasicBlock	229.8	37.4	92.3	866.8
baseline+PAF <sup>4</sup> E(no-FFEHSR)	302.2	50.7	109.1	1184.0
baseline+PAF <sup>4</sup> E	332.1	52.1	113.8	1233.0
baseline+LongBasicBlock+PAF <sup>4</sup> E	<b>234.9</b>	<b>36.9</b>	<b>90.5</b>	<b>785.5</b>

分析表6可知,相比于baseline,引入轻量化骨干网络改进后,在4类嵌入式平台上的推理时间分别减少39.6%,37.6%,39.0%和43.7%,充分说明了本文中轻量化骨干网络的优秀轻量化特性。在引入长距离上下文特征增强分支后,由于网络结构的复杂化、参数数量的增加和计算量的增加,推理时间增加,相较于基线模型,推理速度分别降低35.4%,29.3%,29.2%和31.9%,但其依然具有良好的轻量化特性。对baseline引入PAF<sup>4</sup>E改进后,由于模型参数数量和计算量

下降,推理时间相较于baseline分别降低6.6%,1.5%,12.7%和3.2%,表明渐进式特征融合框架存在轻量化特性。将全部改进结合,推理时间相较于baseline分别降低33.9%,30.2%,30.6%和38.3%,充分说明了各个模块的优秀轻量化特性和在不同嵌入式平台上的泛化性。

为了验证所提算法在嵌入式平台上推理速度的先进性,采用与3.4节中相同的算法,在嵌入式平台上进行推理时间的对比实验,实验结果如表7所列。

表7 模型推理时间对比实验

Table 7 Comparative experiment on model inference time

模型名称	(ms)			
	N5105CPU 推理时间	N5105GPU 推理时间	NVIDIA Maxwell™ GPU 推理时间	4核 ARM A57 推理时间
ResNet50	1380.0	237.3	463.0	4134.2
ResNet101	2441.0	375.7	743.4	6761.0
VGG13	1055.0	139.1	252.3	3539.0
VGG16	1325.0	178.0	302.5	3996.0
VGG19	1521.0	205.1	333.8	4678.0
YOLOv3	533.4	66.5	158.8	1725.6
YOLOv4	601.5	73.3	213.0	2233.6
YOLOv5	355.5	52.9	130.4	1273.3
GhostNet	244.0	53.1	131.9	5242.2
MobileNetv3	287.2	49.9	162.1	5542.5
ShuffleNetv2	144.0	29.3	76.7	2421.5
MobileNeXt	270.4	41.1	133.8	4836.6
本文算法	<b>234.9</b>	<b>36.9</b>	<b>90.5</b>	<b>785.5</b>

对比实验结果表明,在Intel N5105 CPU运算模块、Intel N5105 GPU运算模块、NVIDIA MAXWELL架构GPU和4核ARM A57 CPU这4类计算资源上,相较于经典的ResNet50算法,本文算法推理时间分别大幅减少83.0%,84.5%,80.5%和80.1%;相较于先进的轻量化MobileNeXt算法,本文算法推理时间分别减少13.1%,10.2%,32.3%和83.8%。本文提出的算法在4核ARM A57 CPU平台上的推理时间最短,在其余3类计算资源上的推理时间仅次于Shuf-

fleNetv2,但ShuffleNetv2的mAP@0.5指标仅为76.4,而本文算法高达83.4,二者存在高达7.0%的差距,因此推理时间表现略差于ShuffleNetv2的现象合理。基于嵌入式设备的推理时间对比实验结果充分表明了本文提出的算法在不同嵌入式平台上的优秀轻量化特性。

为进一步证明所提算法的轻量化优势,同样选取文献[32]、文献[33]和文献[34]中的火焰检测算法进行模型推理时间的对比实验。实验结果如表8所列。

表8 相关火焰检测算法模型推理时间的对比实验

Table 8 Comparative experiment on model inference time with related flame detection algorithms

模型名称	(ms)			
	N5105CPU 推理时间	N5105GPU 推理时间	NVIDIA Maxwell™ GPU 推理时间	4核 ARM A57 推理时间
文献[32]	332.3	49.4	126.7	1017.4
文献[33]	602.4	84.1	219.7	2274.2
文献[34]	300.6	43.5	107.1	941.2
本文算法	234.9	36.9	90.5	785.5

对比实验结果表明,在Intel N5105 CPU运算模块、Intel

N5105 GPU运算模块、NVIDIA MAXWELL架构GPU和

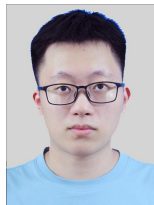
4核 ARM A57 CPU4类计算资源上,相较于文献[32]算法,本文算法的推理时间分别减少 29.3%,25.3%,28.6%和 22.8%;相较于文献[33]算法,本文算法推理时间分别减少 61.0%,56.1%,58.8%和 65.5%;相较于文献[34]算法,本文算法推理时间分别减少 21.9%,15.2%,15.5%和 16.5%。实验结果充分说明:相较于部分其他的火焰检测相关工作,本文算法在推理时间方面也具有一定的先进性。

**结束语** 本文提出了一种渐进自适应特征融合的轻量化火焰检测算法,针对现有算法的模型计算量较大、访存次数多的问题,构建了基于通道级稀疏连接的轻量化卷积算子和基于残差思想的双分支轻量级特征提取组件,在充分降低计算量和访存量增强的情况下增强了通道间信息的交互,缓解了深层骨干网络的信息丢失问题。针对特征直接融合易产生噪声和轻量级骨干网络深层特征丢失的问题,提出一种基于高频增强的特征强化方法,对骨干网络特征图进行强化,并构建了一种渐进自适应特征融合框架,对不同尺度的特征图之间建立联系并完成渐进自适应融合,增强模型对多尺度目标的检测能力。训练前采用 K-means 算法聚类先验锚框,以提升框选质量。实验结果表明,相比于基线模型、现有先进目标检测模型和部分其他火焰检测方法,本文方法能够有效降低参数量和计算量,取得更高的 mAP@0.5 指标。在 4 类嵌入式计算资源上的部署实验表明,本文算法综合优势最大。但由于数据集的限制,小目标和不同视角数据集数量较少,后续应通过增补数据集,增强模型的鲁棒性与泛化性。

## 参 考 文 献

- [1] CHEN J, HE Y, WANG J. Multi-feature fusion based fast video flame detection[J]. *Building and Environment*, 2010, 45(5): 1113-1122.
- [2] DEVE K B, HANCKE G P, SILVA B J. Design of a smart fire detection system[C]// *IECON 2016 - 42nd Annual Conference of the IEEE Industrial Electronics Society*. IEEE, 2016: 6205-6210.
- [3] WU Q Y, YAN Y Y, DU J, et al. Flame Detection Algorithm Based on Multi-feature Fusion[J]. *Journal of Intelligent Systems*, 2015, 10(2): 240-247.
- [4] CORTES C, VAPNIK V. Support-vector networks[J]. *Machine learning*, 1995, 20: 273-297.
- [5] JIA Y, W H Q, HU Y, et al. Image-based Flame Recognition Based on Improved Hierarchical Clustering and SVM[J]. *Computer Engineering and Applications*, 2014(5): 165-168.
- [6] EMMY PREMA C, VINSLEY S S, SURESH S. Efficient flame detection based on static and dynamic texture analysis in forest fire detection[J]. *Fire technology*, 2018, 54: 255-288.
- [7] YAN Y Y, CHEN C X, LIU Y A, et al. Flame Detection of Dynamic Texture Features in Dimension-Weighted Mode[J]. *Chinese Journal of Intelligent Systems*, 2017, 12(4): 548-555.
- [8] HOWARD A G, ZHU M, CHEN B, et al. Mobilenets: Efficient convolutional neural networks for mobile vision applications[J]. *arXiv: 1704. 04861*, 2017.
- [9] SANDLER M, HOWARD A, ZHU M, et al. Mobilenetv2: Inverted residuals and linear bottlenecks[C]// *2018 IEEE Conference on Computer Vision and Pattern Recognition*. IEEE, 2018: 4510-4520.
- [10] HOWARD A, SANDLER M, CHU G, et al. Searching for mobilenetv3[C]// *2019 IEEE/CVF International Conference on Computer Vision*. IEEE, 2019: 1314-1324.
- [11] ZHANG X, ZHOU X, LIN M, et al. Shufflenet: An extremely efficient convolutional neural network for mobile devices[C]// *2018 IEEE Conference on Computer Vision and Pattern*. IEEE, 2018: 6848-6856.
- [12] CHEN J, KAO S, HE H, et al. Run, Don't Walk: Chasing Higher FLOPS for Faster Neural Networks[C]// *2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. IEEE, 2023: 12021-12031.
- [13] HAN K, WANG Y, TIAN Q, et al. Ghostnet: More features from cheap operations[C]// *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. IEEE, 2020: 1580-1589.
- [14] TAN M, CHEN B, PANG R, et al. Mnasnet: Platform-aware neural architecture search for mobile[C]// *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. IEEE, 2019: 2820-2828.
- [15] CHEN Y, KALANTIDIS Y, LI J, et al. A<sup>2</sup>-nets: Double attention networks[C]// *Proceedings of the 32nd International Conference on Neural Information Processing Systems*. 2018: 350-359.
- [16] ZHAO H, JIA J, KOLTUN V. Exploring self-attention for image recognition[C]// *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2020: 10076-10085.
- [17] HU J, SHEN L, SUN G. Squeeze-and-excitation networks[C]// *2018 IEEE Conference on Computer Vision and Pattern Recognition*. IEEE, 2018: 7132-7141.
- [18] WANG Q, WU B, ZHU P, et al. ECA-Net: Efficient channel attention for deep convolutional neural networks[C]// *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2020: 11534-11542.
- [19] YANG L, ZHANG R Y, LI L, et al. Simam: A simple, parameter-free attention module for convolutional neural networks[C]// *2021 International Conference on Machine Learning*. PMLR, 2021: 11863-11874.
- [20] WOO S, PARK J, LEE J Y, et al. Cbam: Convolutional block attention module[C]// *2018 European Conference on Computer Vision*. Springer, 2018: 3-19.
- [21] LIN T Y, DOLLAR P, GIRSHICK R, et al. Feature pyramid networks for object detection[C]// *2017 IEEE Conference on Computer Vision and Pattern Recognition*. IEEE, 2017: 2117-2125.
- [22] LIU S, QI L, QIN H, et al. Path aggregation network for instance segmentation[C]// *2018 IEEE Conference on Computer Vision and Pattern Recognition*. IEEE, 2018: 8759-8768.
- [23] TAN M, PANG R, LE Q V. Efficientdet: Scalable and efficient object detection[C]// *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. IEEE, 2020: 10781-10790.
- [24] GHIASI G, LIN T Y, LE Q V. Nas-fpn: Learning scalable fea-

- ture pyramid architecture for object detection[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2019;7036-7045.
- [25] LIU S, HUANG D, WANG Y. Learning spatial fusion for single-shot object detection[J]. arXiv;1911.09516, 2019.
- [26] HARIHARAN B, MALIK J, RAMANAN D. Discriminative decorrelation for clustering and classification[C]//European Conference on Computer Vision. Berlin: Springer, 2012: 459-472.
- [27] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2016;770-778.
- [28] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition[J]. arXiv;1409.1556, 2014.
- [29] LI C, LI L, JIANG H, et al. YOLOv6: A single-stage object detection framework for industrial applications[J]. arXiv;2209.02976, 2022.
- [30] MA N, ZHANG X, ZHENG H T, et al. Shufflenet v2: Practical guidelines for efficient cnn architecture design[C]//2018 European Conference on Computer Vision. Munich: Springer, 2018: 116-131.
- [31] DAQUAN Z, HOU Q, CHEN Y, et al. Rethinking Bottleneck Structure for Efficient Mobile Network Design[J]. arXiv;2007.02269, 2020.
- [32] HE Y, SAHMA A, HE X, et al. FireNet: A Lightweight and Efficient Multi-Scenario Fire Object Detector[J]. Remote Sensing, 2024, 16(21):4112.
- [33] SHEN P, SUN N, HU K, et al. FireViT: An Adaptive Lightweight Backbone Network for Fire Detection[J]. Forests, 2023, 14(11):2158.
- [34] LIANG Y, CHEN T, ZHANG W. Multi-Scale Fire Detection Algorithm with Adaptive Attention[J]. Transactions of Beijing institute of Technology, 2024, 44(1):91-101.



**LI Xiaolan**, born in 2001, postgraduate, is a student member of CCF (No. U9103G). His main research interests include target detection and edge computing.



**MA Yong**, born in 1983, Ph.D, associate professor. His main research interests include intelligent robotics and intelligent hardware design.

(责任编辑:柯颖)