

光电工程

Opto-Electronic Engineering

中文核心期刊 中国科技核心期刊
Scopus CSCD

基于级联稀疏查询机制的轻量化火灾检测算法

张小雪, 王雨, 吴思远, 孙帮勇

引用本文:

张小雪, 王雨, 吴思远, 等. 基于级联稀疏查询机制的轻量化火灾检测算法[J]. 光电工程, 2023, 50(10): 230216.

Zhang X X, Wang Y, Wu S Y, et al. An improved lightweight fire detection algorithm based on cascade sparse query[J]. *Opto-Electron Eng*, 2023, 50(10): 230216.

<https://doi.org/10.12086/oe.2023.230216>

收稿日期: 2023-09-01; 修改日期: 2023-11-15; 录用日期: 2023-11-15

相关论文

基于改进YOLOv5s的无人机图像实时目标检测

陈旭, 彭冬亮, 谷雨

光电工程 2022, 49(3): 210372 doi: 10.12086/oe.2022.210372

一种用于驾驶场景下手机检测的端到端的神经网络

戴腾, 张珂, 尹东

光电工程 2021, 48(4): 200325 doi: 10.12086/oe.2021.200325

更多相关论文见光电期刊集群网站 



光电工程
Opto-Electronic Engineering

<http://cn.ojournal.org/oe>



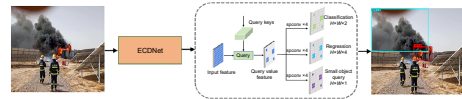
 OE_Journal



Website

DOI: 10.12086/oe.2023.230216

基于级联稀疏查询机制的轻量化火灾检测算法

张小雪¹, 王雨¹, 吴思远², 孙帮勇^{1,2*}¹西安理工大学印刷包装与数字媒体学院, 陕西 西安 710054;²中国科学院西安光学精密机械研究所光谱成像技术重点实验室, 陕西 西安 710119

摘要: 针对现有火灾检测算法仍存在的模型复杂、检测速度慢、误检率高等问题, 提出一种基于级联稀疏查询机制的轻量化火灾检测网络 LFNNet。首先, 建立了轻量化的图像特征提取模块 ECDNet, 其通过在 YOLOv5s 主干网络中嵌入轻量化注意力模块 ECA (efficient channel attention), 用于解决火灾检测中火焰与烟雾的多尺度难点; 其次, 利用深层特征提取模块 FPN+PAN, 对不同层级的特征图进行深度处理和多尺度融合; 最后, 利用嵌入轻量化的级联稀疏查询模块 CSQ (cascade sparse query) 提升对早期火灾中的小火焰与薄烟雾的检测准确率。实验表明, 本文方法在 mAP 和 Precision 等客观指标上的综合表现达到最优, 同时在实现较高检测精度时的参数量也较低, 能够满足实际场景的火灾检测要求。

关键词: 目标检测; 火灾检测; 轻量化; 级联稀疏查询机制; Slimming**中图分类号:** TN911.73**文献标志码:** A

张小雪, 王雨, 吴思远, 等. 基于级联稀疏查询机制的轻量化火灾检测算法 [J]. 光电工程, 2023, 50(10): 230216

Zhang X X, Wang Y, Wu S Y, et al. An improved lightweight fire detection algorithm based on cascade sparse query[J]. *Opto-Electron Eng*, 2023, 50(10): 230216

An improved lightweight fire detection algorithm based on cascade sparse query

Zhang Xiaoxue¹, Wang Yu¹, Wu Siyuan², Sun Bangyong^{1,2*}¹College of Printing, Packaging and Digital Media, Xi'an University of Technology, Xi'an, Shaanxi 710054, China;²Key Laboratory of Spectral Imaging Technology, Xi'an Institute of Optics and Precision Mechanics, Chinese Academy of Sciences, Xi'an, Shaanxi 710119, China

Abstract: To address the challenges of complex models, slow detection speed, and high false detection rate during fire detection, a lightweight fire detection algorithm is proposed based on cascading sparse query mechanism, called LFNNet. In the study, firstly, a lightweight feature extraction module ECDNet is established to extract more fine-grained features in different levels of feature layers by embedding the lightweight attention module ECA (efficient channel attention) in YOLOv5s backbone network, which is used to solve the multi-scale of flame and smoke in fire detection. Secondly, deep feature extraction module FPN+PAN is adopted to improve multi-scale fusion of feature

收稿日期: 2023-09-01; 修回日期: 2023-11-15; 录用日期: 2023-11-15

基金项目: 国家自然科学基金资助项目 (62076199); 陕西省重点研发计划项目 (2022ZDLGY01-03); 瞬态光学与光子技术国家重点实验室开放基金项目 (SKLST202214); 陕西省教育厅重点科学研究计划项目 (23JY063); 西安市科技计划项目 (22GXFW0088).

*通信作者: 孙帮勇, sunbangyong@xaut.edu.cn.

版权所有©2023 中国科学院光电技术研究所

maps at different levels. Finally, the Cascade Sparse Query embedded lightweight cascade sparse query module is applied to improve the detection accuracy of small flames and thin smoke in early fires. Experimental results show that the comprehensive performance of the proposed method in objective indicators such as mAP and Precision is the best on SF-dataset, D-fire and FIRESENSE. Furthermore, the proposed model achieves lower parameters and higher detection accuracy, which can meet the fire detection requirements of challenge scenes.

Keywords: object detection; fire detection; lightweight; cascade sparse query; Slimming

1 引言

火灾是人类日常生活中的主要灾害之一, 及时、准确地发现火灾可以为消防灭火、人员救助赢得宝贵时间, 有效保障人们生命财产安全。随着成像传感器和智能计算的发展, 火灾检测已逐渐发展为宽波段探测和实时信息处理的自动检测过程。

分析火灾检测技术的发展过程可知, 火灾检测主要分为两类: 基于传感器原理的火灾检测和基于计算机视觉的火灾检测。基于传感器原理的火灾检测主要利用烟雾传感器、颜色传感器或者热传感器进行火焰或者烟雾的识别。应用该类传感器检测火灾时, 只有火焰或者烟雾弥散至传感器一定距离范围内且达到浓度阈值才能触发报警, 因此无法对早期火灾进行准确预警, 不适合用于较大场景或者相对开放的空间。同时, 基于传感器的火灾检测方法误检率较高, 在实际应用中存在一定缺陷。

近年来, 计算机视觉技术快速发展, 在火灾检测领域围绕计算机视觉技术可分为基于机器学习和基于深度学习两类检测方法。基于机器学习的火灾检测, 肖潇等^[1]主要利用模式识别原理进行火焰特征与烟雾特征提取, 进而实现火灾检测。火灾检测的关键是对火灾特征的提取和识别, 现有机器学习的检测方法大多是通过研究火焰和烟雾的静态特征和动态特征完成检测。其中, 静态特征主要包括颜色特征、形状特征、纹理特征和边缘特征等; 动态特征主要包括运动特征、形状的变化等。例如, Demirel 等人^[2]提出一种对火焰像素在 YCrCb 颜色空间进行分类的算法, 通过对火焰和非火焰像素进行训练, 构建了火焰像素分类器。但是该方法误报率较高, 无法运用到实际场景。Qiu 等人^[3]提出一种边缘检测算法用于分割火焰前景, 但主要适用于背景简单的火灾场景, 在复杂背景下难以正确分割出火焰。邵婧等人^[4]提出基于动态纹理特征分析的火灾检测算法, 对较大空间的火灾检测率较高, 但难以区分实际火灾区域和颜色与火焰近似的非火灾

区域。为了降低误报率, Surit 等人^[5]通过提取火焰的静态与动态特征, 对接近火焰和烟雾特征的区域进行分析并提出一种误检率较低的火灾检测模型。总体上, 大部分基于机器学习的火灾检测方法是针对特定的单一场景提出, 所建立的检测模型泛化能力差, 容易受火灾区域周围环境的影响, 特别是光照、天气等因素, 难以适应复杂实际场景的火灾检测需求。

基于深度学习的火灾检测方法在最近几年取得了显著性进展, 具有准确性高、成本低、速度快等优势, 已被证明可应用于多种现实场景的火灾检测任务。比如, Zhang 等人^[6]提出了一种联合的深度卷积神经网络火灾分类器, 首先在全局图像上对火灾特征进行提取, 然后在感兴趣火灾区域使用细粒度面片分类器检测火灾面片的精确位置。Dunnings 等人^[7]在训练时间不变的情况下, 使用降低复杂度的深卷积神经网络 Inceptionv1 进行实时火灾区域检测。Sharma 等人^[8]在数量较少的火灾数据集上使用 VGG16 和 ResNet50 预训练的卷积神经网络进行火灾数据集的训练, 并在全连接层上进行微调以增强这些网络。后来, 为了进一步提高火灾检测性能, Akhlofi 等人^[9]提出一种 Deep-Fire 卷积神经网络算法, 在复杂的火灾图像上取得了较高检测精度。任嘉锋等人^[10]基于改进的 YOLOv3 网络提出火灾检测方法, 其针对 YOLOv3 小目标识别性能不足的问题进行改进, 将火灾检测与识别任务转换为多分类识别和坐标回归问题。而缪伟志等人^[11]通过改进 YOLOv4 网络并将其部署在嵌入式设备上对森林火灾检测, 获得了较低的误检率, 不过实时性较差。刘凯等人^[12]提出支持向量机算法, 建立了实时监测系统对森林火灾进行全天候自动检测, 但网络的训练过程较复杂。为了减少模型复杂度, 提高检测速度, 皮骏等人^[13]基于 YOLOv5s 的轻量化森林火灾检测方法, 进行了分类实验, 采用 Shufflenetv2 替换主干网络, 有效降低了参数量, 不过检测精度仍有待提高。以上基于深度学习的火灾检测网络, 在与

训练集类似的场景下都获得了相对满意的检测结果,但仍存在一些不足:1)模型的泛化能力仍有待提高,难以满足复杂场景中多尺度火灾目标的检测要求;2)对小目标的火灾检测精度低,尤其是远距离的小型火灾目标经常被漏检;3)模型的速度与精度难以达到平衡,在现实场景应用中仍存在缺陷。

针对以上问题,本文基于YOLOv5s网络基础框架,提出了一种基于级联稀疏查询机制的轻量化火灾检测网络LFNet。首先,对传统YOLOv5s的主干网络CSPDarkNet进行改进,在网络中嵌入了一种轻量化的注意力模块ECA^[14](efficient channel attention),建立了新的主干网络模块ECDNet(efficient channel attention darkNet)。ECDNet能够在不同级别的特征层中提取细粒度更高的特征信息,解决了多尺度火焰特征导致的检测精度低等问题;同时设计了一种新的轻量化级联查询模块CSQH(cascade sparse query head, CSQH)用于预测火焰概率和位置,能够在不增加网络参数量的情况下提高较小火焰目标的检测性能。

2 网络架构与基本原理

针对火灾检测任务,本文设计了一种轻量化的检

测网络LFNet,结构上主要包含:主干网络ECDNet、深度特征提取模块FPN+PAN以及火灾预测模块CSQH。如图1所示,火灾检测过程中,首先,输入彩色图像在主干网络ECDNet中进行特征提取,获得不同级别和不同细粒度的烟雾特征;然后,主干网络提取的特征被送入深度特征提取模块FPN+PAN进行深度特征提取和融合,从而对烟雾特征进行深度处理和提取;最后,在设计的轻量化预测模块CSQH中引入级联查询机制,计算出不同尺度火焰目标的类别概率和位置。

2.1 改进的主干网络模块ECDNet

现有火灾检测系统多基于YOLOv5主干网络进行特征提取,然而通过大量实验发现,YOLOv5主干网络中的残差结构易造成部分细粒度特征丢失,使得一些类火和类烟目标被检测为火焰,导致火灾检测系统的误检率偏高。为此,本文对传统的YOLOv5主干网络中的残差结构进行改进。如图1所示,主干网络ECDNet主要包括Focus模块、CBL层、ECSP1_3模块、ECSP1_9模块、SPP(spatial pyramid pooling)模块以及ECSP2_3模块。首先输入640×640×3的彩色图像到主干网络ECDNet中,Focus模块对进入主

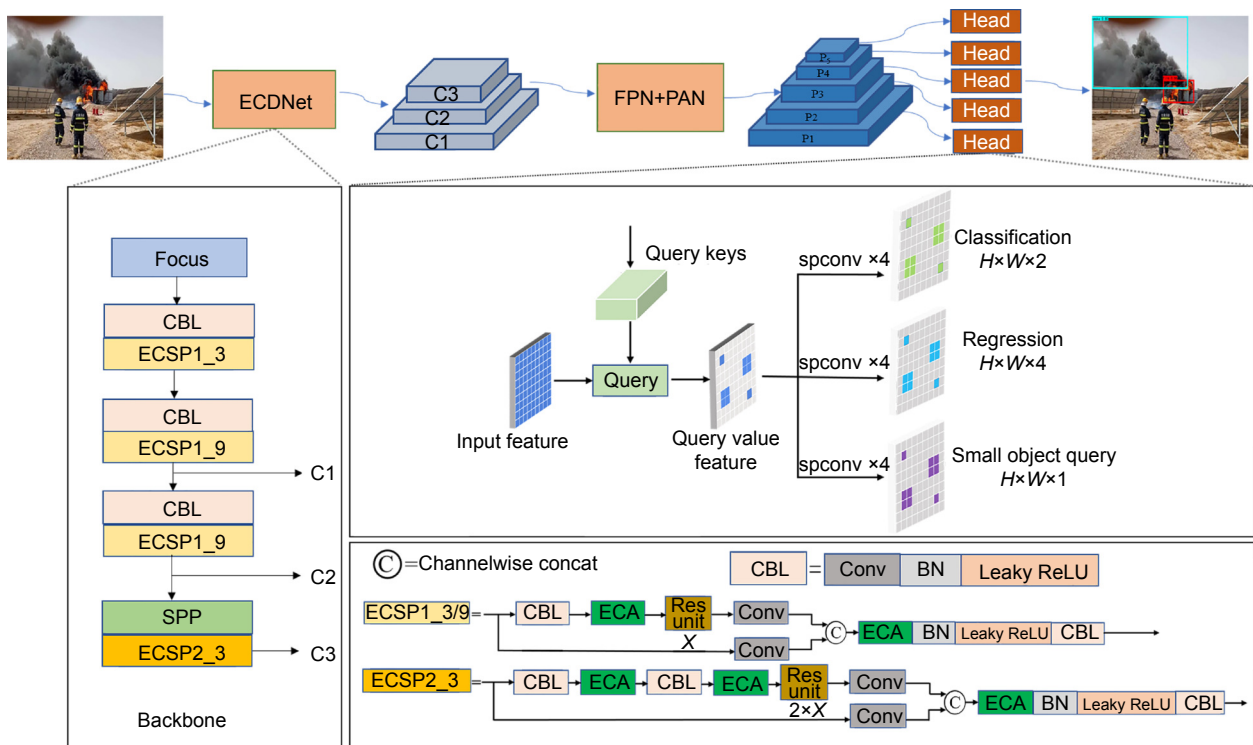


图 1 LFNet 网络架构图

Fig. 1 Network architecture of LFNet

干网络的图片进行切片处理得到 $320 \times 320 \times 64$ 的特征图, 切片操作目的解决图像下采样中的信息丢失问题。接着通过 CBL 层和 ECSP 模块进行特征提取, 其中残差结构 ECSP1_3、ECSP1_9 以及 ECSP2_3 中嵌入了轻量化注意力模块 ECA, 有助于提高不同细粒度特征的提取精度, SPP 模块有效地避免了对特征信息的重复提取。为了得到不同级别和不同细粒度的特征图, 文中采用步长为 2 的 3×3 卷积核提取特征, 卷积核个数分别为 256、512、1024, 选择 C1: $80 \times 80 \times 256$ 、C2: $40 \times 40 \times 512$ 、C3: $20 \times 20 \times 1024$ 三种不同尺寸的特征图送入深度特征提取模块进行特征融合。

ECSP 模块是一种嵌入了 ECA 注意力的轻量化残差模块, 能够在增加少量参数的情况下, 有效避免网络残差结构中细粒度特征的丢失, 从而实现网络特征提取能力的提升。ECA 注意力模块本质是一种未降维的跨通道交互模块, 其中, 不降维有利于细粒度特征的提取, 跨通道交互能够降低模型复杂度并提高检测精度。ECA 模块的结构如图 2 所示, H 、 W 代表高和宽, C 表示通道维数。若定义 K 代表跨通道交互的相邻通道数, 则其与所采用 1×1 卷积的卷积核数相同, 也与通道维数 C 成正比。在 ECA 模块中, 利用非线性函数, 当确定通道维数 C 后, K 值就可以自适应地通过式 (1) 计算:

$$K = \Psi(C) = \left\lfloor \frac{\log_2(C)}{m} + \frac{o}{m_{\text{odd}}} \right\rfloor, \quad (1)$$

其中: $\lfloor \cdot \rfloor_{\text{odd}}$ 表示取最邻近奇数操作, m 和 o 是常数, 本文分别取为 2 和 1。另外, 在特征提取部分的通道数 C 值设定为 32, 64, 128, 256, 则通过公式计算可得, 本文方法的 K 值分别为 3、3、4、4。利用式 (1) 自适应确定 K 值, 能够显著降低特征提取过程的运算量。

ECA 注意力模块通过对特征图进行通道维度的注意力加权, 可以捕捉到空间上不同位置之间的关联

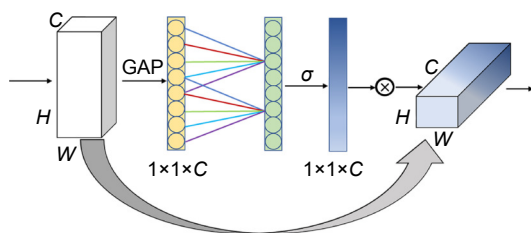


图 2 ECA 注意力模块原理示意图
Fig. 2 Efficient channel attention module

性。如图 3 所示, 嵌入 ECA 注意力模块后可帮助 ECDNet 更好地捕捉火灾特征。此外 ECA 注意力模块在增强特征表达的同时, 可以减少计算资源的消耗, 从而提高模型检测性能和实时性。



图 3 (a) 原图; (b) 注意力热力图
Fig. 3 (a) Original; (b) Attention mechanism heat map

2.2 深层特征提取模块 FPN+PAN

深层特征提取模块主要对主干网络输出的不同等级的特征图进行多尺度融合, 并将图像特征传递给预测模块。如图 1 所示, 深层特征提取采用了 FPN+PAN 结合的双金字塔结构。FPN (Feature pyramid network, 特征金字塔网络) 自顶向下, 将高层的语义特征往下传递, 得到预测的特征图^[15]。而 PAN (path aggregation network, 路径聚合网络) 能够针对 FPN 传递的语义特征信息进一步增强和融合^[16]。深层特征提取模块通过在特征金字塔网络 FPN 的后面增加一个 PAN 自底向上的路径, 用底层精确的定位信息来增强整个特征层次结构, 从而缩短底层特征与高层特征之间的信息路径。因此, 主干网络提取的浅层特征, 通过 FPN+PAN 模块进一步提取后, 得到特征图更精确的语义信息和定位信息, 这将有利于提高预测模块对于小目标火灾的检测精度。

2.3 预测模块 CSQH

为了提高小目标烟雾的检测精度, 本文在预测部分 Head 中嵌入了用于提高小目标检测精度的轻量化级联稀疏查询机制^[17], 建立了新的预测模块 CSQH。该模块不仅能够提高小火焰目标的检测性能, 而且所采用的级联稀疏卷积降低了模型参数量, 能够提高模型速度。

CSQH 预测模块的查询操作包括查询键、查询层、查询值。具体实现过程如图 1 所示, 图像被送入主干和深层特征提取模块, 经过 FPN+PAN 加工生成了一系列不同分辨率的特征图 P_l (l 为特征图的级别)。 P_l 同时也作为预测模块 CSQH 中的查询层 (通过实验, 本文选择从 P_3 层开始查询), 每一层从上一层接收一

组小目标的查询键, 即小目标的粗略位置, 并应用查询操作生成稀疏特征图。然后稀疏检测头和稀疏查询头利用查询键预测对应目标的查询值, 即上一层对应目标的检测框位置和类别概率, 从而逐渐得到图像中不同尺度的火灾类别概率和具体位置。

如图 4 所示, 预测模块包含两个级联的查询操作, 即: Large->Medium 和 Medium->Small, 其中 Large, Medium, Small 指的是深层特征提取模块所对应的特征图 P_1 中目标的不同尺度的级别。以 Large->Medium 为例。首先, 模块在 Large 级别层次的图像对小目标进行标记(将规模小于预设阈值 s 的对象定义为小目标, 阈值 s 为特征层 P_1 上的最小锚框尺度); 其次, 在预测推理的过程中, 模块选择预测分数小于阈值 s 的位置作为查询, 并将这个位置映射到 Medium 的特征图中, 具体的映射原理如式 (2) 所示:

$$\{k_{l-1}^o\} = \{(2x_l^o + i, 2y_l^o + j), \forall i, j \in \{0, 1\}\}, \quad (2)$$

其中, $\{k_{l-1}^o\}$ 为 Large 层中对应的低特征级别层 P_{L-1} 上被判别为小目标的关键位置集合。最后 Medium 上对应的轻量化预测头(本文实验选用 P_3 、 P_4 、 P_5 对应的三个 Head) 通过级联稀疏查询, 将只处理在 Large 级别层中无法被检测到的小目标并对其进行查询预测, 整个计算过程通过稀疏卷积减少预测模块的计算开销。同样, Medium->Small 也通过相同的原理进行预测。

火灾检测中目标类别概率通过 CSQH 中的 Classification 进行预测, 如图 5 所示, 原始图像通过提取特征得到一系列特征图, Classification 层分类时

采用全连接层进行目标类别概率计算, 经过 *softmax* 函数, 将输出形式转化为概率形式。*softmax* 函数定义如:

$$S(x_i) = \frac{e^{x_i}}{\sum_{j=1}^n e^{x_j}}, \quad (3)$$

其中, x_i 表示向量 x 中第 i 个元素。

目标的位置信息通过 Regression 进行预测, 目标的具体位置通过坐标偏移量来表示, 原理如:

$$\begin{cases} b_x = 2\sigma(t_x) - 0.5 + c_x \\ b_y = 2\sigma(t_y) - 0.5 + c_y \\ b_w = p_w(2\sigma(t_w))^2 \\ b_h = p_h(2\sigma(t_h))^2 \end{cases}, \quad (4)$$

其中: b_x 、 b_y 、 b_w 、 b_h 分别指目标锚框的中心坐标与宽高尺寸, p_w 和 p_h 代表预设锚框的宽与高, σ 指 sigmoid 激活函数; t_x 、 t_y 、 t_w 、 t_h 分别是目标框中心相对于预设锚框中心的偏移量; c_x 和 c_y 代表目标所在的网格距图像左上角距离。最后坐标偏移量与预设锚框进行转换, 输出目标框位置。图 5 列出了一种火灾检测中的代表图像, 检测结果输出以包围框形式输出。

2.4 k-Means++ 聚类生成锚框

YOLOv5 中锚框的生成采用 k-means 聚类方法, 在传统的 k-means 算法中, 初始聚类中心的选择具有随机性, 会导致聚类结果不稳定和收敛速度慢的问题。k-means++ 方法在 k-means 聚类基础上, 改善了聚类中心的初始化问题。此外, 不同的数据集上预设的锚框尺寸各不相同, 因此为了生成适用于不同尺度火灾

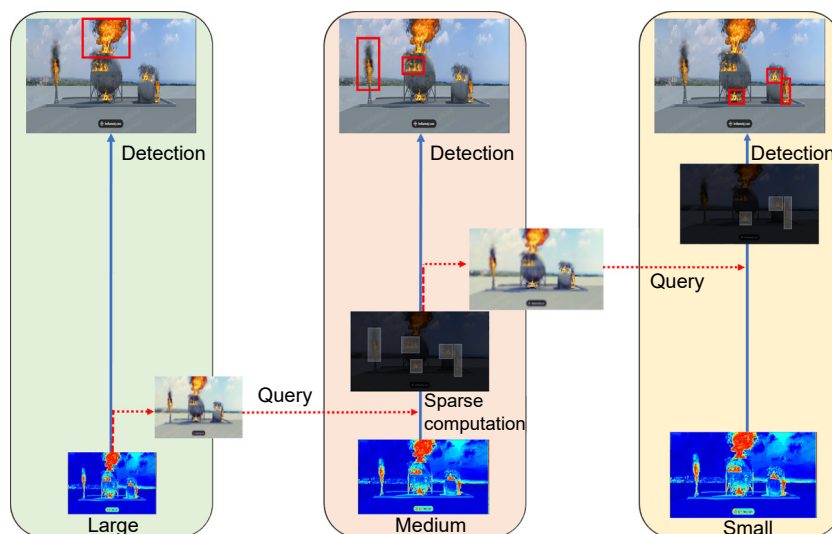


图 4 级联查询机制模块原理示意图

Fig. 4 Cascade sparse query module

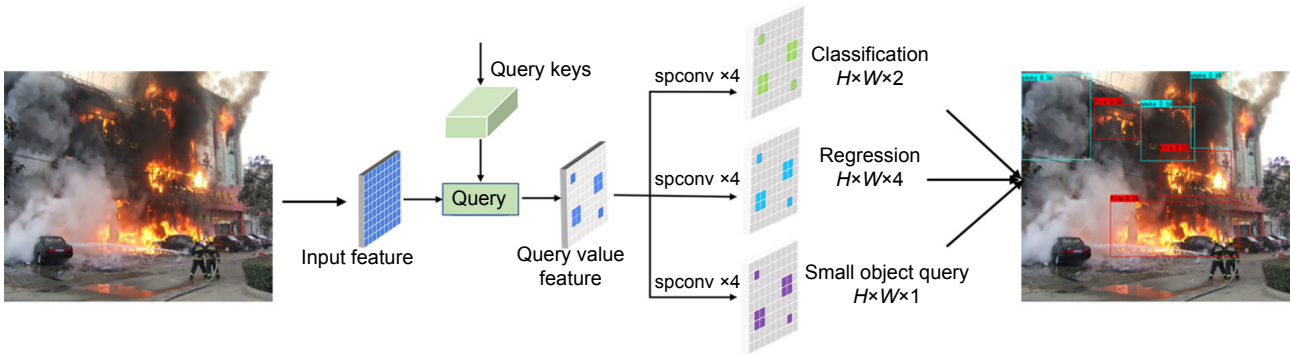


图 5 头部预测原理示意图
Fig. 5 Cascade sparse query head module

的锚框, 利用 k-means++ 聚类方法进行锚框计算。具体聚类过程如下:

a) 从数据集 X 中随机选择一个真实框的宽高作为初始聚类中心 c_1 。

b) 计算每个样本与聚类中心的最小交并比距离 $D(x)$, 依据概率计算式 (5) 选择下一个聚类中心, 式 (5) 为:

$$P(x) = \frac{D(x)^2}{\sum_{x \in X} D(x)^2} \quad (5)$$

c) 重复步骤 b), 直到选出 K 个聚类中心。

d) 根据选定的聚类中心, 计算 x_i 到 K 个聚类中心的交并比距离, 将每个样本分配到最近的聚类中心。

e) 计算每个聚类的新中心, 若聚类中心发生变化, 更新聚类中心。原理为:

$$\mu_i = \frac{1}{c_i} \sum_{i \in c_i} x_i, \quad (6)$$

其中 μ_i 为聚类质心, 即聚类中所有样本的平均值。

f) 重复步骤 d) 和步骤 e), 直到所有锚框的宽高不再变化或变化幅度很小, 即生成最优结果。

依据锚框数量与平均 IOU 曲线, 图 6 所示。实

验最终选择 9 个适用于火灾数据集的锚框尺寸, 分别为 SF-dataset 数据集: (26, 33)、(47, 68)、(93, 69)、(91, 128)、(64, 241)、(185, 126)、(159, 259)、(320, 247)、(453, 459); D-fire 数据集: (10, 13)、(18, 25)、(23, 54)、(42, 36)、(46, 94)、(81, 59)、(88, 145)、(158, 83)、(221, 205); FIRESENSE 数据集: (15, 20)、(31, 33)、(36, 64)、(67, 76)、(42, 158)、(141, 87)、(90, 156)、(158, 183)、(305, 267)。三个数据集的聚类结果如图 7 所示。

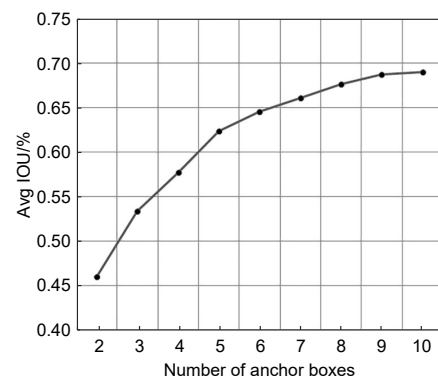


图 6 SF-dataset 数据集上聚类实验结果
Fig. 6 Clustering experiment results on SF-dataset

因此使用 k-means++ 聚类算法, 可保证每个锚框的差异性尽可能大, 以满足火灾检测的聚类要求, 从

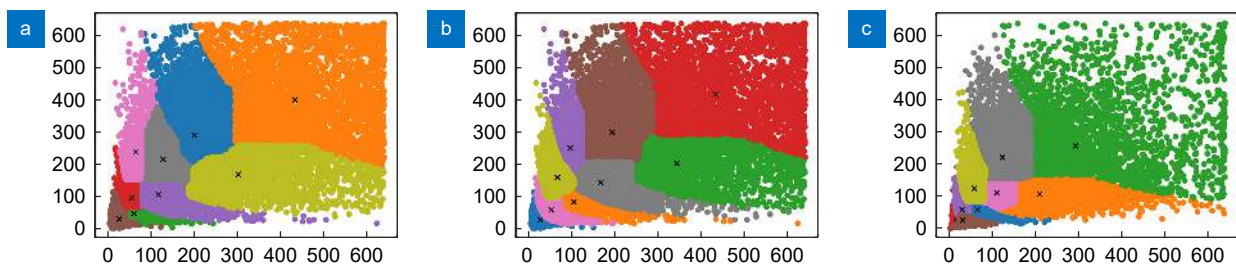


图 7 聚类结果。(a) SF-dataset; (b) D-fire; (c) FIRESENSE
Fig. 7 Clustering results. (a) SF-dataset; (b) D-fire; (c) FIRESENSE

而提升模型训练时边界框回归的速度与精度。

2.5 损失函数

有效的损失函数可提高模型的收敛速度与检测精度。实验设计的火灾检测损失函数由置信度损失, 类别损失, 定位损失三部分加权组合而成:

$$Loss = aLoss_{con} + bLoss_{cla} + cLoss_{reg}, \quad (7)$$

其中: $Loss_{con}$ 代表置信度损失函数, 为 BCE 二值交叉熵损失函数, $Loss_{cla}$ 代表类别损失函数, 为 Quality Focal 损失函数, $Loss_{reg}$ 代表边界框回归损失函数, 为 CIOU 损失函数。 a, b, c 代表三个可学习的损失函数权。 BCE 损失函数用于预测矩形框的可信程度, 取值范围 0~1, 值越大说明矩形框中越可能存在目标, 其原理如:

$$L_{bce} = \frac{1}{n} \sum_x [y \ln \varphi + (1-y) \ln (1-\varphi)], \quad (8)$$

其中: x 表示样本, y 表示标签, φ 表示预测的输出, n 表示样本的总数量。

Quality Focal 损失函数用于解决类别中正负样本不平衡分布的问题, 其计算原理如:

$$L_{qfl} = -\alpha y(1-y)^\gamma \log y' - (1-\alpha)(1-y)^\gamma \log(1-y') \\ = \begin{cases} -\alpha(1-y)^\gamma \log y', y = 1 \\ -(1-\alpha)y^\gamma \log(1-y'), y = 0 \end{cases}, \quad (9)$$

其中: y 为标签, y' 为预测结果, 指数因子 γ 减少了易分样本的关注度, 特征贡献率因子 α 用于调节难分和易分 2 种样本的特征权重贡献率。

CIOU 损失函数用于预测边界框的回归位置, 计算原理如式 (10):

$$CIOU(B, B_{gt}) = -\eta v - \frac{\rho^2(B, B_{gt})}{d^2} + IOU(B, B_{gt}), \quad (10)$$

$$v = \frac{4}{\pi} \left(\arctan \frac{w^{gt}}{h^{gt}} - \arctan \frac{w}{h} \right)^2, \quad (11)$$

$$\eta = \frac{v}{1 - IOU(B, B_{gt}) + v}, \quad (12)$$

其中: B 和 B_{gt} 为预测框和真实框, ρ 为预测框与真实框中心点的距离, d 为包围预测框和真实框的最小包围框的对角距离。 w 和 h 为预测矩形框的宽和高, w^{gt} 和 h^{gt} 分别为真实矩形框的宽与高。 v 为归一化了预测框和真实框的比例差值, η 为权衡长宽比例的平衡因子。实验通过采用类别损失与边界框回归损失相

结合的 Quality Focal-CIOU 损失函数, 从而实现较快的训练收敛和预测框定位。

2.6 模型压缩 Slimming

减少模型的参数量和计算量可以通过多种方法实现, 例如优化主干网络, 模型压缩等。但是考虑到模型实时性, 仅仅关注参数量与计算量是不够的, 还要关注模型在优化之后的推理速度, 有些优化方案虽然减少了参数量与计算量, 但模型的推理速度也在相应增长, 这将导致模型的实时性变差。因此实验在对比了多种优化方案后选择使用 Slimming 剪枝算法^[18]对模型进行压缩。

Slimming 剪枝算法利用 BN (batch normalization) 层的 γ 参数来衡量当前通道的重要程度, 并使用 L1 损失来稀疏化 γ 值。BN 层作为一种能加速网络收敛及提高网络性能的标准方法, 被广泛应用于深度学习中, 其基本公式如:

$$\hat{z} = \frac{z_{in} - \mu_B}{\sqrt{\sigma_B^2 + \epsilon}}; z_{out} = \gamma \hat{z} + \beta \quad (13)$$

其中: z_{in} 和 z_{out} 分别是 BN 层的输入和输出, μ_B 表示本批次该层输入的均值, σ_B 表示本批次该层输入的方差。BN 层会对卷积后的输出特征乘上 γ 因子, 即缩放因子来调整其分布, 如果 γ 因子很小, 那么这个通道的输出值也会很小, 进而就代表着该通道对于后续网络的贡献就很小, 在剪枝时就会删减这些不重要的通道, 保留较为重要的通道, 从而实现模型的压缩。Slimming 剪枝算法的定义公式如:

$$L = \sum_{(x,y)} l(f(x, W), y) + \lambda \sum_{\gamma \in \Gamma} g(\gamma), \quad (14)$$

其中: (x, y) 是训练的输入和目标, 为网络中的可训练参数, 前一项代表卷积神经网络中的训练损失函数, $g(\cdot)$ 是缩放因子上的惩罚项, λ 是两项的平衡因子。

在显示研究中, L1 为正则 ($g(s)=|s|$), 被广泛地用于稀疏化。该剪枝算法包含三个步骤: 稀疏正则化, 剪枝, 微调, 其中, 稀疏正则化就是在模型训练时利用 L1 损失来标记通道的重要程度, 剪枝则是根据通道的重要程度进行修剪, 而修剪不重要的通道可能会造成模型短暂的性能降低, 因此在剪枝后还需要进行模型的微调。本文在进行模型稀疏正则化时仅利用原始训练集对模型进行稀疏训练, 而在微调阶段时才利用数据增强后的训练集进行稀疏训练, 实现高效的轻量化算法。

3 实验与分析

3.1 实验设置

本文在检测网络训练过程中, 主要实验参数设置如下: Batch size 为 64, 学习率为 0.001, 循环迭代轮次为 300, 采用 Pytorch 框架, 实验 GPU 为 NVIDIA GTX 3090。

3.2 数据集

实验中选用了三个火灾数据集, 分别为 D-fire 数据集、FIRESENSE 数据集以及 SF-dataset 数据集。D-fire 为公开的数据集, 由 21527 张不同尺度的火灾烟雾图片组成, 包括森林中远距离的火灾烟雾, 郊外不同尺度的火灾烟雾以及草坪远距离的小型火灾目标, 数据集链接地址为: <https://github.com/gaiasd/DFireDatase>; FIRESENSE 为公开的数据集, 由 49 个视频组成, 包括日常生活类火物体背景下的火灾烟雾以及浓烟背景下的火灾烟雾, 数据集链接地址为: <https://www.kaggle.com/datasets/chrisfilo/firesense>。考虑到 D-fire 数据集和 FIRESENSE 数据集中场景比较单一, 特别是缺少室内外建筑的多样化火灾场景, 本文建立了 SF-dataset 数据集。SF-dataset 数据集包含多种挑战性火灾场景, 由 18218 张火灾烟雾图片组成, 包括室外火灾烟雾以及室内火灾烟雾等。本文将三种数据集分别按 8:1:1 比例划分为训练集、验证集、测试集, 具体分配如表 1 所示。训练过程中, 首先对火灾数据集进行数据预处理, 通过读取数据集的 XML 文本来获得训练图像的存储路径以及图像中目标的标记像素坐标和目标类别。此外, 为了使网络能够更好地学习到火灾的目标特征, 需要对数据集进行数据增强。本文对训练图像采用了翻转、旋转角度、随机颜色、对比度增强、亮度增强等数据增强手段。

3.3 评价指标

本文使用训练的 LNet 网络对测试图像进行了一系列实验, 以验证所提出方法的有效性。实验采用模

型精度 (Precision)、召回率 (Recall)、准确率 (Accuracy)、类别精度 (AP)、类别平均精度 (mAP)、模型参数 (Parameter)、检测速度 (FPS) 等指标进行评价, 各项指标定义如:

$$Precision = \frac{TP}{TP + FP}, \quad (15)$$

$$Recall = \frac{TP}{TP + FN}, \quad (16)$$

$$Accuracy = \frac{TP + FN}{TP + FP + FN + TN}, \quad (17)$$

$$AP = \int_0^1 Precision(t) dt, \quad (18)$$

$$mAP = \frac{\sum_{i=1}^C AP_i}{C}, \quad (19)$$

其中: TP (true positives) 为正样本被检测正确的数量; FP (false positives) 为负样本被检测为正样本的数量; FN (false negatives) 为背景被错误检测为正样本的数量; C 表示类别数。

3.4 对比实验

为了验证提出的 LNet 网络的有效性, 实验使用检测精度 (Precision)、召回率 (Recall) 等指标在三个火灾数据集上与经典的检测算法进行了对比实验。对比方法包括 3 类: 3 种传统的火灾检测方法^[2,19-20]; 3 种现有的基于深度学习的火灾检测方法 Fire-YOLO^[21]、EFDNet^[22]、Pruned+KD^[23] 以及 3 种主流的通用型目标检测方法 YOLOX-Tiny^[24]、PicoDet^[25]、YOLOv7^[26] 进行比较。此外, 为了验证网络模型的轻量化, 利用模型参数 (Parameter)、检测速度 (Speed) 等指标在 SF-dataset 数据集上与主流的轻量化检测算法进行了对比实验。对比方法包括两类: 5 种现有的基于深度学习的轻量化火灾检测方法: M-YOLO^[27]、Fire-YOLO^[21]、EFDNet^[22]、Y-Edge^[25]、Pruned+KD^[23] 以及 5 种主流的轻量化目标检测方法 PPYOLO-Tiny^[28]、YOLOX-Tiny^[24]、PicoDet^[25]、PP-YOLOE^[29]、YOLOv7^[26]。

表 2 为不同对比方法在不同火灾数据集上以 mAP、Precision、Recall 以及 Accuracy 为客观指标的

表 1 三个数据集的训练集, 验证集, 测试集的详细统计数据
Table 1 Numbers of training set, validation set and testing set for the three datasets

数据集	训练集			验证集			测试集			总数
	Fire	Smoke	None	Fire	Smoke	None	Fire	Smoke	None	
SF-dataset	4859	4859	4859	607	607	607	607	607	607	18219
D-Fire	4658	4693	7870	582	587	984	582	587	984	21527
FIRESENSE(video)	9	11	21	1	1	2	1	1	2	49

表 2 不同数据集上不同方法的精度对比实验结果
Table 2 Experimental precision results of different comparative methods on different datasets

数据集	方法	指标			
		Recall/%	Precision/%	Accuracy/%	mAP/%
SF-dataset	Celik 等 ^[19]	38.45	69.54	72.12	41.58
	Demirel 等 ^[2]	42.28	73.65	78.54	47.15
	Zhang 等 ^[20]	40.12	72.77	73.15	45.4
	Fire-YOLO ^[21]	49.93	87.05	90.09	69.38
	EFDNet ^[22]	44.27	80.45	82.28	59.89
	Pruned+KD ^[23]	47.25	83.16	85.64	63.1
	YOLOX-Tiny ^[24]	49.95	86.89	89.24	69.08
	PicoDet ^[25]	50.1	87.1	90.13	69.42
	YOLOv7 ^[26]	54.12	88.64	94.88	71.69
	LFNet	54.98	89.12	98.5	71.76
D-fire	Celik 等 ^[19]	35.90	65.78	68.42	39.65
	Demirel 等 ^[2]	40.68	73.12	78.27	45.85
	Zhang 等 ^[20]	38.67	70.45	70.86	43.94
	Fire-YOLO ^[21]	51.18	84.12	88.21	68.88
	EFDNet ^[22]	43.68	76.57	77.94	58.77
	Pruned+KD ^[23]	46.02	79.85	82.4	62.71
	YOLOX-Tiny ^[24]	51.06	83.94	86.14	68.14
	PicoDet ^[25]	51.27	84.32	88.26	68.95
	YOLOv7 ^[26]	53.12	86.44	93.85	70.65
	LFNet	53.35	87.68	97.92	71.15
FIRESENSE	Celik 等 ^[19]	33.7	56.07	60.21	36.48
	Demirel 等 ^[2]	38.55	64.28	69.44	42.66
	Zhang 等 ^[20]	36.96	61.58	62.38	40.36
	Fire-YOLO ^[21]	52.47	79.88	85.19	68.12
	EFDNet ^[22]	42.74	68.25	70.22	56.97
	Pruned+KD ^[23]	46.80	72.94	75.1	61.35
	YOLOX-Tiny ^[24]	53.94	80.44	84.12	68.02
	PicoDet ^[25]	52.96	80.92	85.91	68.89
	YOLOv7 ^[26]	54.53	89.76	96.28	70.15
	LFNet	53.19	92.44	98.42	70.61

注: 加粗字体表示最优结果

实验结果。通过实验结果, 可以发现, 对于相同的方法, 无论在 mAP、Recall、Precision 以及 Accuracy 指标上, SF-dataset 数据集训练的模型的性能均优于 D-fire 和 FIRESENSE 数据集, 这主要是由于涵盖了更多具有挑战性的检测场景, 有助于网络训练获得更好

的学习能力。在 SF-dataset 数据集上, LFNet 网络实现了最佳的 mAP、Recall、Precision 以及 Accuracy, 分别为 71.76%、54.98%、89.12% 以及 98.5%。而在三个数据集上, 相比于其他的方法, 传统方法的各种性能指标均处于最低, 这主要是由于传统火灾检测方

法侧重于颜色, 运动以及纹理等手工特征来检测火灾, 不能完全提取火灾的特征, 从而导致检测精度较低。相比之下, 本文方法可以有效地提取火灾的特征, 实现最佳的检测性能。

表 3 为不同方法在 SF-dataset 数据集上以 mAP、Speed、Flops、Parameter 以及 Infer time 为指标的实验结果。如表 3 所示, 本文方法实现了 71.76% 的 mAP, 在所有方法中实现了最佳的检测性能。在 Flops、Speed 以及 Infer time 方面, PicoDet 达到了最优的效果, 但其 mAP 为 69.42%, 比本文方法低 2.34%, 比次优方法 YOLOv7 低 2.27%。通过实验结果可以看出, 尽管本文方法在检测速度以及模型参数方面没有实现最优, 但比大多数方法表现更好, 实现了实时的火灾检测。因此综合检测性能的角度考虑, 本文方法在各项检测精度中最优, 同时模型参数大小合适, 处理速度较快, 是一种轻量化, 低延时和高精度的火灾检测算法。此外, 为了进一步验证本文方法的有效性, 本文将 LNet 方法与几个具有代表性的方法进行了可视化的对比检测实验。图 8 为 LNet、EFDNet、Y-Edge、M-YOLO、Fire-YOLO、YOLOX-Tiny、PicoDet、PPYOLOE 以及 YOLOv7 方法分别在三种火灾数据集上不同场景下可视化的火灾检测效果。其中图 8 主要包含建筑物火灾和浓烟同时存在的检测场景以及空旷的室外不同程度的火焰同时存在的检测场景; 图 9 主要包含森林中远距离不同尺度的火灾同时存在的检测场景以及草坪上远距离火焰和近距离的小火焰同时存在的检测场景; 图 10 主要包含日

常生活中火灾与类火物体同时存在的检测场景以及雾霾与火灾同时存在的检测场景。从可视化的检测结果来看, 本文方法无论是对于远距离的小火灾还是不同尺度火灾变化的场景, 都能实现一个最佳的检测效果。

3.5 消融实验

本文使用 YOLOv5s 作为 BaseNet, 并采用五种策略在 SF-dataset 上测试最有效的火灾检测精度。为了验证这些策略的性能, 本文方法使用不同的设置进行了消融实验: 1) 改进 YOLOv5s 的主干网络模块为 ECDNet; 2) 使用 FPN+PAN 作为深层特征提取模块; 3) 使用 CSQH 预测模块替换 YOLOv5s 的预测头部分; 4) 使用 Quality-CIOU 作为网络的损失函数; 5) 利用 Slimming 剪枝算法对网络进行压缩。

实验结果如表 4 所示, 尽管 BaseNet 的 mAP 达到 67.89%, 但其性能仍不能满足火灾检测的要求, 例如对于一些有类火物体存在的场景以及远距离不同尺度的火灾场景, 而且其模型参数大小达到了 27.6 M。与 BaseNet 相比, 改进的主干网络 ECDNet 的 mAP 达到了 68.89%, 在不增加模型参数的情况下, 获得了 1% 的提高。在网络 ECDNet 的基础上添加 FPN+PAN, 模型参数大小仅增加 0.1 M, 但模型 mAP 从 68.89% 增加到 70.15%, Recall 从 50.96% 增加到 51.53%。由此可见, 使用 FPN+PAN 的特征融合方式有效提高了模型的检测性能。而当将网络的 BaseNet 的预测头部分替换成 CSQH 时, 模型的检测速度达到了 91 f/s, mAP 提高了 0.46%, 特别是对于远距离小

表 3 SF-dataset 上不同方法的速度对比实验结果

Table 3 Experimental speed results of different comparative methods on SF-dataset

方法	指标				
	Flops/G	Parameter/M	Speed/(f/s)	Infer time/ms	mAP/%
M-YOLO ^[27]	7.54	23.8	18	50.6	66.6
Fire-YOLO ^[21]	45.12	62	28	32.17	69.38
EFDNet ^[22]	1.99	3.6	63	7.21	59.89
Y-Edge ^[25]	30.47	53.8	36	27.97	65.95
Prund+KD ^[23]	16.8	26.3	45	17.25	63.1
PPYOLO-Tiny ^[26]	4.96	17.84	42	23.8	68.36
YOLOX-Tiny ^[24]	5.42	19.19	40	25.21	69.08
PicoDet ^[25]	0.73	4	105	6.65	69.42
PPYOLOE ^[29]	29.42	52.20	98	25.64	70.85
YOLOv7 ^[26]	18.42	36.9	88	19.85	71.69
LNet	3.85	12.6	98	10.24	71.76

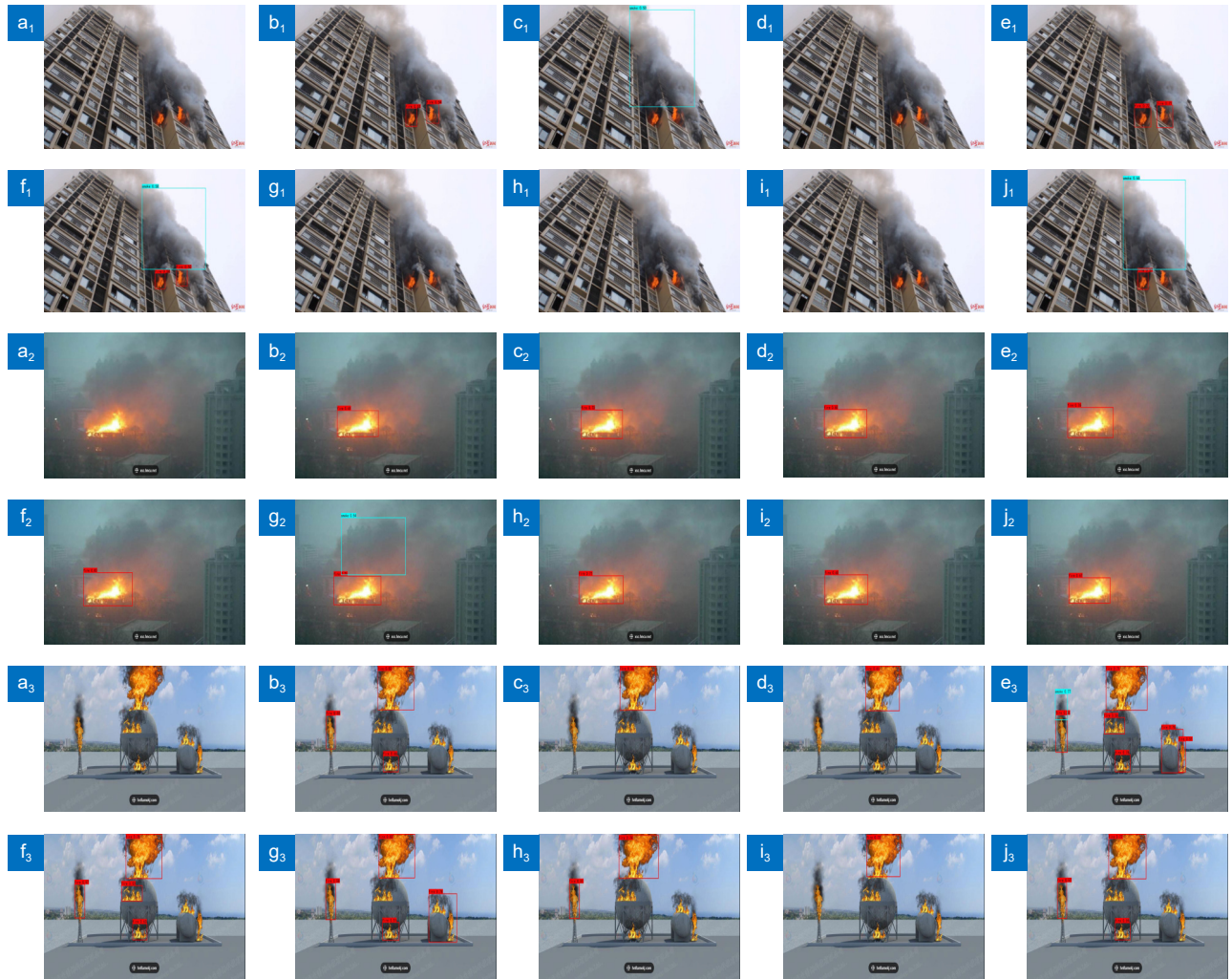


图 8 SF-dataset 数据集比较实验检测结果。(a) Images; (b) Ours; (c) EFDNet; (d) Y-Edge; (e) M-YOLO; (f) Fire-YOLO; (g) YOLOX-Tiny; (h) PicoDet; (i) PP-YOLOE; (j) YOLOv7
 Fig. 8 Comparison experiment detection results for the SF-dataset. (a) Images; (b) Ours; (c) EFDNet; (d) Y-Edge; (e) M-YOLO; (f) Fire-YOLO; (g) YOLOX-Tiny; (h) PicoDet; (i) PP-YOLOE; (j) YOLOv7

尺度火灾的检测性能, 有了明显的改进。最后网络在分别使用了 Quality Focal-CIOU 损失函数和 Slimming 剪枝算法后, 模型的检测精度、模型参数大小以及检测速度均有了一定的提升, 模型的 mAP 最终实现了 71.76%, 模型参数大小仅为 12.6M, 速度达到了 98 f/s。因此综合实验结果的各项指标, 可以看出每种策略对于 LFNet 的有效性和优越性。

3.6 参数实验

为了验证不同大小的训练数据集样本比例、Batch size、不同模型输入尺寸以及迭代次数 epoch 对于所提出的 LFNet 网络火灾检测精度的影响, 实验在 SF-dataset 数据集上进行了多个参数测试, 实验结果如图 11 所示。

图 11(a) 显示了不同大小训练数据集样本比例对于模型精度的影响, 在给定占比 50% 的训练数据集样本的条件下, 本文方法的 mAP 指标已达到 69.53%; 在一定的范围内, 当不断提高训练样本占比时, 在 80% 的训练数据集样本占比条件下, 本文方法的 mAP 达到了 71.76%。

图 11(b) 给出了不同 Batch size 大小对于模型精度的影响, 选定的 Batch size 分别为 8,16,32,64,128。从图中可以看出, 不同 Batch size 条件下的 mAP 指标存在一定差异, 总体波动在 2% 之内。

图 11(c) 给出了不同的模型输入尺寸对于模型精度的影响。从图示结果可以看出, 不同的模型输入尺寸与模型精度成正比, 模型输入尺寸越大, 模型精度



图 9 D-fire 数据集比较实验检测结果。(a) Images; (b) Ours; (c) EFDNet; (d) Y-Edge; (e) M-YOLO; (f) Fire-YOLO; (g) YOLOX-Tiny; (h) PicoDet; (i) PP-YOLOE; (j) YOLOv7;

Fig. 9 Comparison experiment detection results for the D-fire dataset. (a) Images; (b) Ours; (c) EFDNet; (d) Y-Edge; (e) M-YOLO; (f) Fire-YOLO; (g) YOLOX-Tiny; (h) PicoDet; (i) PP-YOLOE; (j) YOLOv7;

也随之提高,但同时也伴随着模型参数大小的增加,这主要是由于随着模型输入尺寸的增加,网络能够捕捉更多的空间信息。因此,综合模型的检测精度与模型参数大小,本文最终选择了 640×640 作为模型输入尺寸,在参数量足够小的同时提高模型精度,实现了速度与精度的相对平衡。

图 11(d) 给出了不同迭代次数 (epoch) 对于模型精度的影响,图示可以看出不同 epoch 对于 mAP 指标的影响有一定的影响,波动在 2% 之内。综合考虑模型速度与精度,本次迭代次数选用 300。

4 结论

本文提出了一种基于级联稀疏查询机制的轻量化火灾检测算法,有效缓解了当前火灾检测算法面临的多尺度火灾的检测精度低,对小目标的漏检率较高以及速度与精度难以达到平衡等主要问题。本文采用轻量化的主干网络 ECDNet 在不同级别的特征层中提取细粒度更高的特征信息,解决了多尺度火焰特征导致的检测精度低等问题;通过引入轻量化的 CSQH 模块提高对小目标的检测性能;利用 Slimming 剪枝算法



图 10 FIRESENSE 数据集比较实验检测结果。(a) Images; (b) Ours; (c) EFDNet; (d) Y-Edge; (e) M-YOLO; (f) Fire-YOLO; (g) YOLOX-Tiny; (h) PicoDet; (i) PP-YOLOE; (j) YOLOv7

Fig. 10 Comparison experiment detection results for the FIRESENSE dataset. (a) Images; (b) Ours; (c) EFDNet; (d) Y-Edge; (e) M-YOLO; (f) Fire-YOLO; (g) YOLOX-Tiny; (h) PicoDet; (i) PP-YOLOE; (j) YOLOv7

表 4 消融实验的实验结果

Table 4 Experimental results of Ablation experiments

Model	方法					指标			
	ECDNet	FPN+P	CSQH	QFocal-CIOU	Slimming	mAP(%)	Recall/%	Parameter/M	Speed/(f/s)
YOLOv5s						67.89	49.38	27.6	84
	√					68.89	50.96	27.6	84
	√	√				70.15	51.53	27.7	80
	√	√	√			70.61	52.12	27.7	91
	√	√	√	√		71.15	53.35	27.7	91
	√	√	√	√	√	71.76	52.98	12.6	98

注: 加粗字体表示最优结果

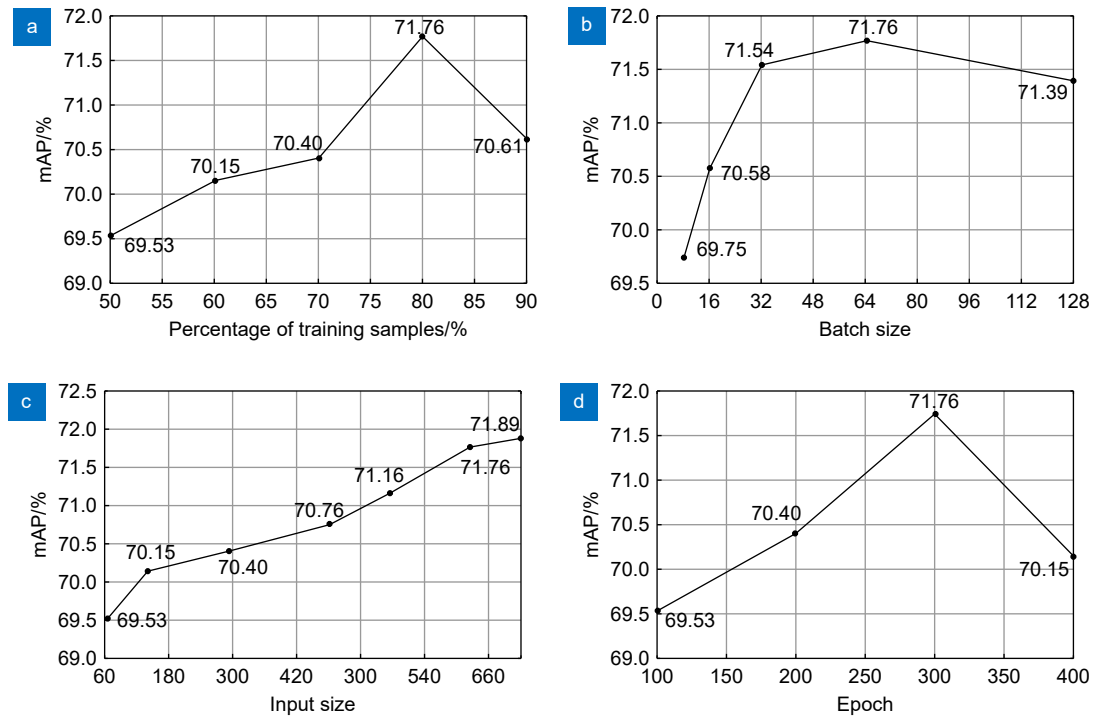


图 11 在 SF-dataset 数据集上训练数据集样本比例、Batch size、模型输入尺寸以及迭代次数 epoch 实验结果

(a) 训练数据集样本比例; (b) Batch size; (c) 模型输入尺寸; (d) 迭代次数 epoch

Fig. 11 Parameters experiment of percentage of training samples, batch size and patch size on the Santa Barbara dataset

(a) Percentage of training dataset samples; (b) Batch size; (c) Model input size; (d) Epoch

来进一步减少模型参数量, 使模型获得更好的性能提升。实验通过自建数据集以及公开的火灾数据集训练并搭建火灾检测模型。实验结果表明, 在模型参数大小合适, 速度较快的情况下, 本文方法无论是在自建的火灾数据集还是公开的火灾数据集均达到了一个最优的检测效果, 这将有利于基于计算机视觉的火灾检测方法在工业的应用。

参考文献

- [1] Xiao X, Kong F Z, Liu J H. Monitoring video fire detection algorithm based on dynamic characteristics and static characteristics[J]. *Comput Sci*, 2019, 46(S1): 284–286,299. 肖潇, 孔凡芝, 刘金华. 基于动静态特征的监控视频火灾检测算法[J]. *计算机科学*, 2019, 46(S1): 284–286,299.
- [2] Çelik T, Demirel H. Fire detection in video sequences using a generic color model[J]. *Fire Saf J*, 2009, 44(2): 147–158.
- [3] Qiu T, Yan Y, Lu G. An autoadaptive edge-detection algorithm for flame and fire image processing[J]. *IEEE Trans Instrum Meas*, 2012, 61(5): 1486–1493.
- [4] Shao J, Wang G X, Guo W. Fire detection based on video dynamic texture[J]. *J Image Graph*, 2013, 18(6): 647–653. 邵婧, 王冠香, 郭蔚. 基于视频动态纹理的火灾检测[J]. *中国图象图形学报*, 2013, 18(6): 647–653.
- [5] Surit S, Chatwiriya W. Forest fire smoke detection in video based on digital image processing approach with static and dynamic characteristic analysis[C]//*Proceedings of the 2011 First ACIS/JNU International Conference on Computers, Networks, Systems and Industrial Engineering*, 2011: 35–39. <https://doi.org/10.1109/CNSI.2011.47>.
- [6] Zhang Q J, Xu J L, Xu L, et al. Deep convolutional neural networks for forest fire detection[C]//*Proceedings of 2016 International Forum on Management, Education and Information Technology Application*, 2016: 568–575. <https://doi.org/10.2991/ifmeita-16.2016.105>.
- [7] Dunning A J, Breckon T P. Experimentally defined convolutional neural network architecture variants for non-temporal real-time fire detection[C]//*Proceedings of 2018 25th IEEE International Conference on Image Processing*, 2018: 1558–1562. <https://doi.org/10.1109/ICIP.2018.8451657>.
- [8] Sharma J, Granmo O C, Goodwin M, et al. Deep convolutional neural networks for fire detection in images[C]//*Proceedings of the 18th International Conference on Engineering Applications of Neural Networks*, 2017: 183–193. https://doi.org/10.1007/978-3-319-65172-9_16.
- [9] Akhloufi M A, Tokime R B, Ellassady H. Wildland fires detection and segmentation using deep learning[J]. *Proc SPIE*, 2018, 10649: 106490B.
- [10] Ren J F, Xiong W H, Wu Z H, et al. Fire detection and identification based on improved YOLOv3[J]. *Comput Syst Appl*, 2019, 28(12): 171–176. 任嘉锋, 熊卫华, 吴之昊, 等. 基于改进YOLOv3的火灾检测与识别[J]. *计算机系统应用*, 2019, 28(12): 171–176.
- [11] Miao W Z, Lu Z N, Wang J L, et al. Fire detection research based on vision[J]. *For Eng*, 2022, 38(1): 86–92,100. 缪伟志, 陆兆纳, 王俊龙, 等. 基于视觉的火灾检测研究[J]. *森林工程*, 2022, 38(1): 86–92,100.

- [12] Liu K, Wei Y X, Xu J G, et al. Design of forest fire identification algorithm based on computer vision[J]. *For Eng*, 2018, **34**(4): 89–95.
刘凯, 魏艳秀, 许京港, 等. 基于计算机视觉的森林火灾识别算法设计[J]. *森林工程*, 2018, **34**(4): 89–95.
- [13] Pi J, Liu Y H, Li J H. Research on lightweight forest fire detection algorithm based on YOLOv5s[J]. *J Graph*, 2023, **44**(1): 26–32.
皮骏, 刘宇恒, 李久昊. 基于YOLOv5s的轻量化森林火灾检测算法研究[J]. *图学学报*, 2023, **44**(1): 26–32.
- [14] Wang Q L, Wu B G, Zhu P F, et al. ECA-Net: efficient channel attention for deep convolutional neural networks[C]// *Proceedings of 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2020: 11534–11542. <https://doi.org/10.1109/CVPR42600.2020.01155>.
- [15] Lin T Y, Dollár P, Girshick R, et al. Feature pyramid networks for object detection[C]// *Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition*, 2017: 2117–2125. <https://doi.org/10.1109/CVPR.2017.106>.
- [16] Wang W H, Xie E Z, Song X G, et al. Efficient and accurate arbitrary-shaped text detection with pixel aggregation network[C]// *Proceedings of 2019 IEEE/CVF International Conference on Computer Vision*, 2019: 8440–8449. <https://doi.org/10.1109/ICCV.2019.00853>.
- [17] Yang C H Y, Huang Z H, Wang N Y. QueryDet: cascaded sparse query for accelerating high-resolution small object detection[C]// *Proceedings of 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2022: 13668–13677. <https://doi.org/10.1109/CVPR52688.2022.01330>.
- [18] Liu Z, Li J G, Shen Z Q, et al. Learning efficient convolutional networks through network slimming[C]// *Proceedings of 2017 IEEE International Conference on Computer Vision*, 2017: 2736–2744. <https://doi.org/10.1109/ICCV.2017.298>.
- [19] Celik T, Ozkaramanli H, Demirel H. Fire pixel classification using fuzzy logic and statistical color model[C]// *Proceedings of 2007 IEEE International Conference on Acoustics, Speech and Signal Processing*, 2007: I-1205–I-1208. <https://doi.org/10.1109/ICASSP.2007.366130>.
- [20] Zhang D Y, Han S Z, Zhao J H, et al. Image based forest fire detection using dynamic characteristics with artificial neural networks[C]// *Proceedings of 2009 International Joint Conference on Artificial Intelligence*, 2009: 290–293. <https://doi.org/10.1109/IJCAI.2009.79>.
- [21] Zhao L, Zhi L Q, Zhao C, et al. Fire-YOLO: a small target object detection method for fire inspection[J]. *Sustainability*, 2022, **14**(9): 4930.
- [22] Li S B, Yang Q D, Liu P. An efficient fire detection method based on multiscale feature extraction, implicit deep supervision and channel attention mechanism[J]. *IEEE Trans Image Proc*, 2020, **29**: 8467–8475.
- [23] Wang S Y, Zhao J, Ta N, et al. A real-time deep learning forest fire monitoring algorithm based on an improved pruned + KD model[J]. *J Real-Time Image Proc*, 2021, **18**(6): 2319–2329.
- [24] Ge Z, Liu S T, Wang F, et al. YOLOX: exceeding YOLO series in 2021[Z]. arXiv: 2107.08430, 2021. <https://arxiv.org/abs/2107.08430>.
- [25] Yu G H, Chang Q Y, Lv W Y, et al. PP-PicoDet: a better real-time object detector on mobile devices[Z]. arXiv: 2111.00902, 2021. <https://arxiv.org/abs/2111.00902>.
- [26] Wang C Y, Bochkovskiy A, Liao H Y M. YOLOv7: trainable bag-of-freebies sets new state-of-the-art for real-time object detectors[C]// *Proceedings of 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2023: 7464–7475. <https://doi.org/10.1109/CVPR52729.2023.00721>.
- [27] Wang S Y, Chen T, Lv X Y, et al. Forest fire detection based on lightweight Yolo[C]// *Proceedings of 2021 33rd Chinese Control and Decision Conference*, 2021: 1560–1565. <https://doi.org/10.1109/CCDC52312.2021.9601362>.
- [28] Long X, Deng K P, Wang G Z, et al. PP-YOLO: an effective and efficient implementation of object detector[Z]. arXiv: 2007.12099, 2020. <https://arxiv.org/abs/2007.12099>.
- [29] Xu S L, Wang X X, Lv W Y, et al. PP-YOLOE: an evolved version of YOLO[Z]. arXiv: 2203.16250, 2022. <https://arxiv.org/abs/2203.16250>.

作者简介



张小雪(2000-), 女, 硕士研究生, 主要研究方向为计算机视觉和深度学习。

E-mail: 18792399592@163.com



【通信作者】孙帮勇(1980-), 男, 教授, 博士生导师, 主要研究方向为多光谱成像技术和图像处理。

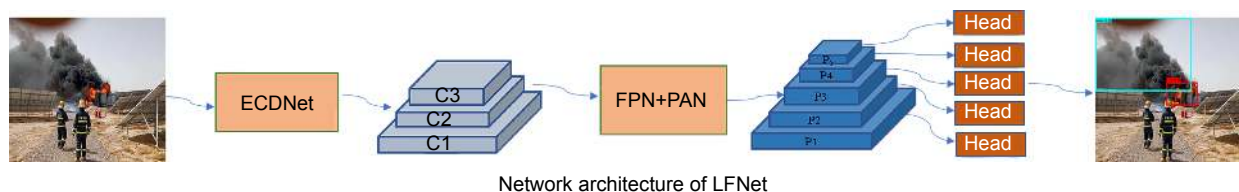
E-mail: sunbangyong@xaut.edu.cn



扫描二维码, 获取PDF全文

An improved lightweight fire detection algorithm based on cascade sparse query

Zhang Xiaoxue¹, Wang Yu¹, Wu Siyuan², Sun Bangyong^{1,2*}



Overview: The Visual Fire Detection task is designed to detect flames and smoke using visual algorithms from images and videos to achieve fire alarms. In recent years, fire detection algorithms based on convolutional neural networks have greatly improved the detection accuracy of flames and smoke. However, the following questions still exist in the current methods: 1) The generalization ability of the model still needs to be improved; 2) Low fire detection accuracy for small objects; 3) The tradeoff between the detection accuracy and speed fails to achieve. In order to overcome the above problems, a lightweight fire detection algorithm is proposed based on cascading sparse query mechanism, called LFNet. In this study, firstly, a lightweight feature extraction module ECDNet is established to extract more fine-grained features in different levels of feature layers by embedding the lightweight attention module ECA (Efficient Channel Attention) in YOLOv5s backbone network, which is used to solve the multi-scale of flame and smoke in fire detection. Secondly, deep feature extraction module FPN+PAN is adopted to improve multi-scale fusion of feature maps at different levels. Finally, the Cascade Sparse Query embedded lightweight cascade sparse query module is applied to enhance the detection accuracy of small flames and thin smoke in early fires. Furthermore, to further decrease the parameters and calculation of the model, the Slimming pruning algorithm is adopted to change the size of the model. The experimental results on the three fire datasets of SF-dataset, D-fire and FIRESENSE show that the comprehensive performance of the proposed method on objective indicators such as mAP and Recall is best. On the SF-dataset dataset, the LFNet achieves the best mAP and Recall, which are 71.76% and 52.98%, respectively. On the D-fire dataset, the mAP of our method reaches 71.76%, which is far superior to other fire detection methods. On the FIRESENSE dataset, our method achieves 70.61% mAP. Our method effectively alleviates the main problems of current fire detection algorithms, such as low detection accuracy, high missed detection rate for small objects, and difficulty in balancing speed and accuracy. The network trains and builds a fire detection model on self-built datasets and other fire datasets. The experimental results show that on the condition that the model size is suitable and the speed is relatively fast, our method achieves an optimal detection effect on both the self-built fire dataset and the public fire datasets, and will potentially promote the application of deep learning-based fire detection methods in industries.

Zhang X X, Wang Y, Wu S Y, et al. An improved lightweight fire detection algorithm based on cascade sparse query[J]. *Opto-Electron Eng*, 2023, 50(10): 230216; DOI: [10.12086/oe.2023.230216](https://doi.org/10.12086/oe.2023.230216)

Foundation item: Project supported by National Natural Science Foundation of China (62076199), Key Research and Development Program of Shaanxi (2022ZDLGY01-03), Open Fund of the State Key Laboratory of Transient Optics and Photonic Technology (SKLST202214), Key Scientific Research Program of Shaanxi Provincial Department of Education (23JY063), and Scientific Research Program of Xi'an (22GXFW0088)

¹College of Printing, Packaging and Digital Media, Xi'an University of Technology, Xi'an, Shaanxi 710054, China; ²Key Laboratory of Spectral Imaging Technology, Xi'an Institute of Optics and Precision Mechanics, Chinese Academy of Sciences, Xi'an, Shaanxi 710119, China

* E-mail: sunbangyong@xaut.edu.cn