

DOI:10.19651/j.cnki.emt.2313981

复杂环境下的 YOLOv5s 烟火检测方法^{*}

王铮帅 邱联奎 李迎港

(河南科技大学信息工程学院 洛阳 471023)

摘要: 针对复杂环境中烟火检测困难,检测精度低的问题,提出一种改进 YOLOv5s 的烟火检测方法。首先,针对 Neck 层烟火特征融合不准确、效果差的问题,提出一种通道注意力机制—Scoring module,对每个通道的特征打分,选择分数高的特征进行特征融合,过滤分数低的特征,避免引入过多冗余特征,在增加少量的可训练参数情况下提高特征融合能力和检测精度;然后,为了提高 Head 层的预测能力,使用 α -EIOU 替换 GIOU 损失函数,提高预测框的定位和检测性能;最后,为了改善数据集数据量少、数据形式单一的问题,使用改进的 Mosaic 数据增强法扩充样本数据,提高模型泛化能力。实验结果表明,改进后的方法比原 YOLOv5s 平均精度均值高 4.7%,检测速度为 212 fps,同时在与其它改进型 YOLOv5s 的对比实验中表现较好。在环境复杂的图像中取得了较好的检测效果,可以满足复杂环境下的烟火检测任务。

关键词: 烟火检测;YOLOv5s;特征选择模块; α -EIOU;数据增强

中图分类号: TP391 **文献标识码:** A **国家标准学科分类代码:** 520.6040

Smoke and flame detection method with YOLOv5s in complex environment

Wang Zhengshuai Qiu Liankui Li Yinggang

(College of Information Engineering, Henan University of Science and Technology, Luoyang 471023, China)

Abstract: Aiming at the problems of complex smoke occurrence scene and low accuracy in smoke, an improved YOLOv5s smoke and flame detection method is proposed. Firstly, in order to solve the problem that the Neck feature fusion of smoke and flame is not accurate and the effect is poor, a new channel attention mechanism, Scoring module, is proposed to score features of each channel. Features with high scores are selected for feature fusion and features with low scores are filtered to avoid introducing too many redundant features. On the premise of not increasing too much computation burden, the module can enhance feature fusion ability and detection accuracy. Then, in order to improve the prediction ability of the Head layer, α -EIOU is used to replace GIOU as the prediction box regression loss to improve the prediction accuracy of the prediction box. Finally, the improved Mosaic data enhancement method is used to solve the problem of small data set and single data form, expand the sample data, and improve the generalization ability of the model. As a result, the mean average precision of the modified YOLOv5s model is improved by 4.7%, while the detection speed reaches 212 frames per second. Meanwhile, it performs well in the comparison experiment with other improved YOLOv5s. It achieves good detection effect in the image with complex environment, and can meet the task of smoke and flame detection in complex environment.

Keywords: smoke and flame detection;YOLOv5s;feature select module; α -EIOU;data enhancement

0 引言

火灾由于其不可预见性,时刻威胁着人民的生命财产安全。据中华人民共和国应急管理部统计:2020 年全国范围共上报火灾 25.2 万起,死亡 1 183 人,受伤 775 人,直接财产损失 40.09 亿元;2021 年全国范围共上报火灾 74.8 万

起,死亡 1 987 人,受伤 2 225 人,直接财产损失 67.5 亿元,因此,烟火检测有着重大的研究意义。早期的烟火检测方法依靠各类传感器对现场的信息进行采集^[1],这类方法实时性较差,且检测空间有限^[2]。传统的图像处理方法通过人工选取烟火的部分特征,如颜色、光流、运动方向、空间域、频域等特征,Yuan 等^[3]将局部二值模式、核主成分分

收稿日期:2023-06-28

^{*} 基金项目:国家自然科学基金(U1504617)项目资助

析、高斯回归模型相结合,提高了传统烟雾检测的精度。然而,传统图像处理方法泛化能力不足,检测速度慢,难以满足越来越复杂的火灾场景需要。

近年来,基于深度学习的烟火检测方法开始成为主流研究方向。文献[4]率先提出基于神经网络的烟火检测方法,并取得了显著的效果。其中基于目标检测模型的烟火检测方法效果最好,研究最为广泛,典型代表网络有 YOLO、SSD、Faster R-CNN 等。文献[5]使用改进的 YOLOv3 网络实现火灾检测,文献[6]基于 SSD 网络,采用递进池化和特征融合预测烟雾位置。文献[7]使用 Faster R-CNN 进行烟雾检测,取得了较高的检测精度,但是速度较慢。

目标检测方法的改进主要集中在检测精度和检测速度两个方面,针对不同场景,王一旭等^[8]通过引入无参数的注意力机制(a simple, parameter-free attention module, SimAM)增强 YOLOv5 对烟火小目标特征提取能力,并使用遗传算法优化超参数,但是增加了计算量,导致检测速度下降。宋华伟等^[9]针对烟火图像的特征规律,在 YOLOv5s 的 Neck 层使用双向交叉尺度融合模块,同时在 head 层加入协调注意力机制增强推理能力,提高网络对该类型图像的检测精度和速度,但是样本集数量少,不充分,实际检测环境中表现一般。文献[10]针对野外火灾的特点,使用 K 均值聚类算法(K-means clustering algorithm, K-means++)提取 YOLOv5 的锚框,并使用剪枝优化网络检测速度,但是检测精度并不高。文献[11]为满足工业生产需要,提出一

种 YOLO 网络和残差神经网络(deep residual network, ResNet)相结合的分阶段烟火检测方法,在检测精度上的表现良好,但双网增加了计算成本,导致检测速度不高。文献[12]将 YOLOv5s 与对抗神经网络(super-resolution generative adversarial network, SRGAN)模型相结合,重点检测灯光昏暗场景下的小火焰目标,但对于发展中及扩散的火灾场景置信度偏低。

目前针对复杂场景下的烟火检测研究较少,且检测精度普遍较低,为了提高在多场景、复杂环境中烟雾检测精度及检测速度,本文提出一种改进 YOLOv5s 的烟火检测方法。对收集到的公开烟火数据集进行数据增强,提高模型泛化能力,然后在瓶颈层的路径聚合网络(path aggregation network, PAN)中嵌入一种评分注意力机制—Scoring module,提高特征融合能力,增强对不同场景下烟火特征的识别能力,最后,使用 α -IOU (alpha-intersection over union, α -IOU) 替换头部网络的损失函数(intersection over union, IOU),目的是提升预测框的性能,增强对复杂场景中烟火锚框的准确定位。

1 YOLOv5s 模型

YOLO 系列^[13-15]是目标检测领域中广泛使用的网络之一,YOLOv5 是较新的版本,包括 s、m、l、x、xl 等多种版本。本文选取综合性能较好,并且更加稳定的 YOLOv5s 作为烟火检测模型,模型结构如图 1 所示。

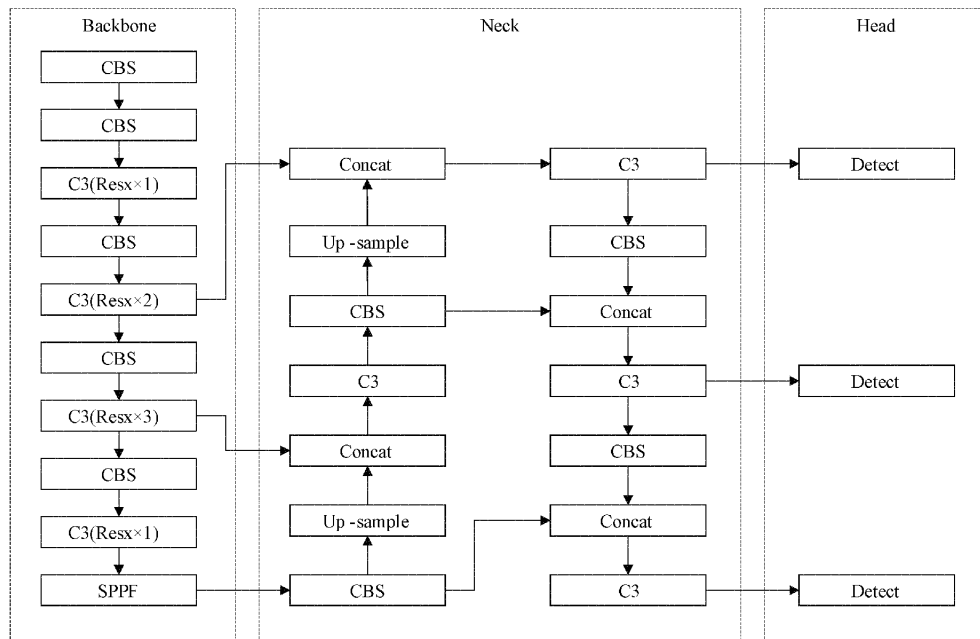


图 1 YOLOv5s 结构图

YOLOv5s 网络主要包含三部分:主干网络(Backbone)、颈部(Neck)和头部(Head)^[16]。主干网络包括 CBS(Conv、batchnormalization、silu)、C3(cross stage partial bottle)、SPPF(spatial pyramid pooling fast)3 个模

块,CBS 模块采用批量归一化降低网络参数量,提高检测速度;C3 模块使用了跨阶段局部网络理论,改善主网络推理过程的计算瓶颈,减少内存占用;SPPF 模块在 SPP 的基础上将并行连接卷积核调整为串行,旨在降低计算量,提

升检测效率。颈部使用路径聚合网络融合特征,包含自顶向下的PAN和自底向上的特征金字塔网络(feature pyramid network, FPN),其中自顶向下路径用来传递图像语义信息,自底向上路径用来传递图像位置信息。头部也称为输出阶段,根据颈部输出的特征图生成多个锚框,根据非极大值抑制(non-maximum sup-pression, NMS)原理^[17]和边界框损失函数(complete intersection over union, CIoU)去除多余的锚框,输出预测结果。YOLOv5s的部分模块结构如图2所示。

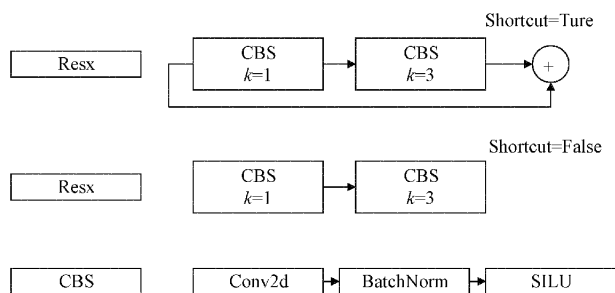


图2 YOLOv5s部分模块结构图

2 改进YOLOv5s的烟火检测方法

2.1 评分模块

传统特征融合算法生成的特征图往往维数很大,因为它只是不同特征图的拼接操作。更高的特征维度并不意味着更好的性能,甚至可能降低检测器的性能。此外,并非所有特征都对输出具有积极影响。这些冗余特征也被称为神经网络中的“扰动变量”,影响检测器的性能^[18]。文献[19]提出了一种通道注意力机制(squeeze-and-excitation networks, SE),对每个通道进行“Squeeze”和“Excitation”,目的是尽可能地减少扰动变量,同时加强最具代表性的特征,本文采用通道评分的思想,将通道注意力和跳跃连接^[20]结构相结合,提出一种评分模块 Scoring module, Scoring module 模块与 SE 模块的机制类似,分别对输入特征图中每个通道进行评分,过滤掉低分值的特征通道,仅保留满足 Top-K 分数的通道,避免特征融合阶段引入过多的不相关特征。评分模块的整体架构如图3所示。

评分模块可以分解为两个分支:特征压缩分支(feature compression branch, FCB)和特征选择分支(feature selection branch, FSB),如图4所示。

该模块包括上下两个分支,上层分支负责提取输入特征并生成特征得分图。下层分支使用特征得分图来选择输入特征中得分较高的特征,最后,将两个分支的输出合并并输出。此外,使用跳跃连接确保输入特征图可以直接转移到输出特征图,从而简化评分模块的学习复杂度。

FCB的主要任务是从输入中获得最具代表性的特征,使用 1×1 卷积核计算输入特征图,生成 Top-K 得分图和

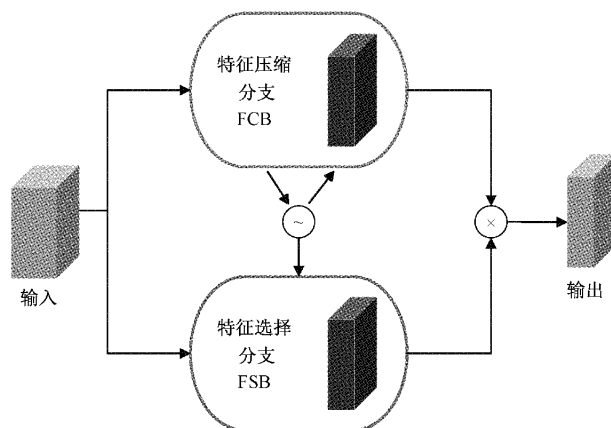
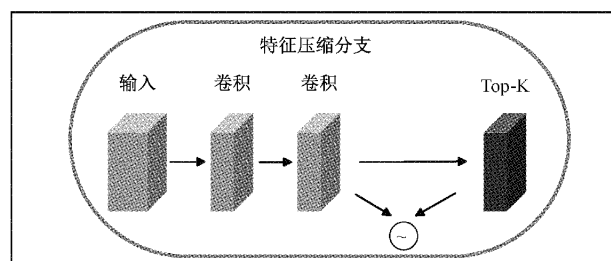
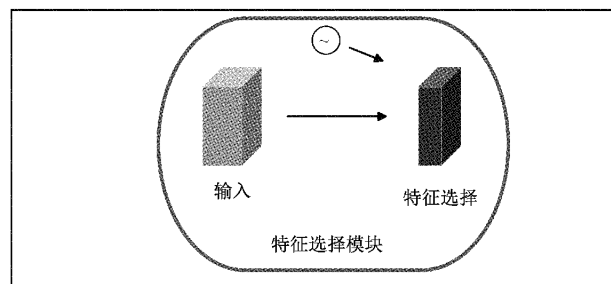


图3 评分模块结构图



(a) 特征压缩分支结构图



(b) 特征选择分支结构图

图4 两个分支的主要结构图

相应的分数信息。FSB 只有一个非常简单的特征选择操作,并且不对输入进行任何卷积计算,因此没有额外的参数被添加到网络中。FSB 使用 FCB 生成的 Top-K 图来收集相应的输入特征以重构特征图。

评分模块的计算过程如下:

第1步,在 FCB 分支中建立变换 F_{fs} , 将输入 $X \in \mathbf{R}^{H \times W \times C}$ (H, W, C 分别为张量的高、宽和通道数)映射到特征图 $M \in \mathbf{R}^{H \times W \times K}$, 其中 $V = \{v_1, v_2, \dots, v_c\}$ 是 F_{fs} 中的一组卷积核集合。FCB 的卷积过程如式(1)所示。

$$X^{conv} = V_2 \otimes (V_1 \otimes X) \quad (1)$$

其中, $V_1 = \{v_1^1, v_1^2, \dots, v_1^K\}$ 和 $V_2 = \{v_2^1, v_2^2, \dots, v_2^C\}$ 表示在 FCB 内构建的第一和第二卷积层, \otimes 代表 2D 卷积运算。

第2步,计算 FSB 生成的 Top-K 图。

$$S_{h,w} = \text{TopK}(\text{sig}(X^{conv})) \quad (2)$$

其中, $TopK$ 是通道特征选择算子, 可以过滤掉通道中分数为非 $Top-K$ 的特征, 仅留下分数为 $Top-K$ 的特征, $TopK$ 使用 FCB 来学习高分图, 并将其应用于 FCB 模块 M^{FCB} 和 FSB 模块 M^{FSB} 。 S 是布尔矩阵, 根据 $TopK$ 算子生成分数为 $Top-K$ 的特征。下标 h, w 表示特征像素的位置。

第3步, 计算 FCB 的输出特征图。

$$M = M^{FCB} \cdot M^{FSB} \quad (3)$$

第4步, 计算 FSB 的输出特征图。

$$M_{h,w}^{FCB} = S_{h,w} \cdot X_{h,w}^{conv} \quad (4)$$

第5步, 计算特征融合输出。

$$M_{h,w}^{FSB} = S_{h,w} \cdot X_{h,w} \quad (5)$$

式(6)、(7)详细介绍了 $TopK$ 运算符。

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} = TopK([\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3]^T), \varphi_1 < \varphi_2 < \varphi_3 \quad (6)$$

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \times [\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3]^T = \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_3 \end{bmatrix} \quad (7)$$

在特征 $X = [\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3]$, φ_i 用于表示对应 i 特征的得分。通过所生成的矩阵和特征 X 的叉乘, 可以过滤掉最低得分 \mathbf{x}_2 。

下面计算引入评分模块后, 原网络增加的计算成本。增加的额外参数源于添加评分机制的两个 Conv2d 层, 对于 640×640 的输入图像, 下行信道速率设置为 0.5, 评分模块引入对网络增加的参数总数由式(8)给出。

$$\sum_{n=1}^N 2r \times C_{in}^2 \quad (8)$$

N 表示网络中应用的评分模块的数量, 而 r 表示下行信道速率, 并且在 C 中表示输入信道的维度。当在 Conv2d 层中使用偏置项时, 所引入的参数和计算成本为 $(1+r)C_{in}$, 通常可忽略。经计算, 在 Neck 层加入评分模块后可训练参数增加了 1.169%。因此, Scoring module 是一种轻量级的神经网络嵌入模型, 对模型可训练参数的影响很小。

2.2 α -EIOU 损失函数

IOU 损失表示目标位置预测值和真实值之间的差值, 原始 YOLOv5 采用 GIOU (generalized-IOU, GIOU) 作为边界框回归损失函数, GIOU 的计算公式如下:

$$IOU = \frac{A \cap B}{A \cup B} \quad (9)$$

$$GIOU = IOU - \frac{C - (A \cup B)}{C} \quad (10)$$

其中, C 为预测框与真实框的最小外接矩形面积, A 为预测框面积, B 为真实框面积。当预测框位于真实框内部, 即 $A \cup B = C$ 时, 由式(10)可知, GIOU 的输出变为 IOU, 这种现象又称为“退化”, 会造成预测框无法准确回归。

为了解决这一问题, CIOU (complete intersection over

union, CIOU) 通过反映纵横比的差异提升预测框的精度, 但由于不能反映实际宽高与其置信度间的差异, 因此又有学者提出了 EIOU (efficient intersection over union, EIOU) 损失函数^[21], EIOU 明确地衡量了重叠区域、中心点距离和宽高真实差; α -IOU 根据自适应加权 IOU 目标的损失和梯度, 通过权重参数 α 对这些损失进行幂变换, 使回归过程专注于高质量的锚框, 以提高模型定位和检测的性能。本文将 EIOU 和 α -IOU^[22] 相结合, 提出一种 α -EIOU 函数修正预测框的位置, α -EIOU 的计算公式如下:

$$L_{EIOU} = 1 - IOU + \frac{\rho^2(b, b^{gt})}{c^2} + \frac{\rho^2(w, w^{gt})}{c_w^2} + \frac{\rho^2(h, h^{gt})}{c_h^2} \quad (11)$$

$$L_{\alpha EIOU} = 1 - IOU^\alpha + \frac{\rho^{2\alpha}(b, b^{gt})}{c^{2\alpha}} + \frac{\rho^{2\alpha}(w, w^{gt})}{c_w^{2\alpha}} + \frac{\rho^{2\alpha}(h, h^{gt})}{c_h^{2\alpha}} \quad (12)$$

其中, C_w 和 C_h 是真实框和预测框的最小外接矩形的宽度和高度, b 和 b^{gt} 分别表示预测框和真实框的中心点, $\rho(b, b^{gt})$ 计算两个中心点间的欧氏距离, $\rho(w, w^{gt})$ 和 $\rho(h, h^{gt})$ 分别表示预测框与最小外接框宽高的差值, $\alpha > 1$ 用于对锚框设置绝对对梯度权重, IOU 越大, 边界框损失也越大, 从而加速对高质量锚框的学习, 提高对梯度自适应加权的边界框回归精度。根据文献^[22], 本文 α 取值为 3 时, 梯度下降较快, 训练效果较好。

3 实验与分析

3.1 数据集增强与标注

考虑到烟火种类丰富, 场景难以统一, 目标尺度可大可小, 小目标的烟火检测往往更加困难, Mosaic 数据增强法^[23]可以将 4 幅图像拼接为一副新图, 增加对小目标的检测性能。

Mosaic 增强法分为 4 个步骤:

- 1) 从数据集任意 batch 中随机选取 4 幅图像。
- 2) 对 4 幅图像随机进行旋转、缩放、色差调节等操作。
- 3) 将 4 幅图像从左上开始按顺时针拼接为新图, 然后利用矩阵的方式截取 4 幅图像的固定区域。
- 4) 将截取区域拼接为一张新的图像样本, 并变换对应的标签。

但是也存在相应的问题。首先, 它先进行数据增强(旋转、缩放等操作), 之后对图像进行拼接操作, 导致模型训练时间长, 且裁剪操作为固定区域, 存在裁剪区域目标不存在目标或目标过小等问题, 容易造成无效操作。

针对这些问题, 本文对 Mosaic 数据增强法进行改进, 首先, 针对裁剪操作, 使用真实框中心点的均值 (ground truth, GT) 和标签^[24]防止裁剪区域不存在目标或目标过小等情况, 其次, 先使用 Mosaic 数据增强, 后对图像进行翻转

等操作,目的是节省数据增强过程的时间,提高效率。修改后的流程图如图5所示。

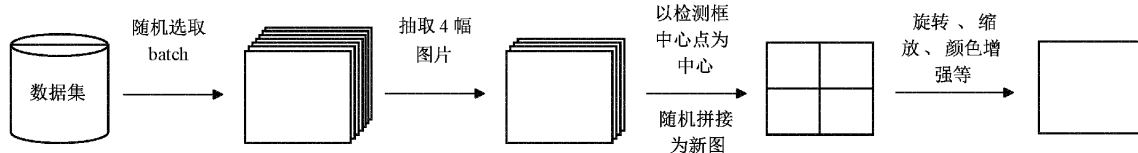


图5 改进的Mosaic数据增强流程图

评价指标包括:准确度(precision, P)、召回率(recall, R)、平均精度均值(mean average precision, mAP)。

$$P = \frac{TP}{TP + FP} \quad (13)$$

$$R = \frac{TP}{TP + FN} \quad (14)$$

$$AP = \int_0^1 P(R) dR \quad (15)$$

$$mAP = \frac{1}{m} \sum_{i=1}^m AP_i \quad (16)$$

其中, TP 表示实际正类且预测正类; FP 表示实际负类但预测正类; FN 表示实际正类预测负类, AP 为平均精度,代表某一类的均值, mAP 为全部类别的均值。

表1是Mosaic增强法改进前后对比情况,选取训练集进行验证。

表1 Mosaic增强法改进前后对比情况 %

数据增强方法	准确度	召回率	平均精度均值
Mosaic	82.3	74.1	80.2
改进 Mosaic	83.1	74.3	81.0

从实验结果看,改进后的Mosaic在 P 和 R 上分别提升了0.8%和0.2%, mAP 提高了0.8%,表明该方法可以提高模型的性能。

烟火检测目前尚未有公开的权威数据集,本文选取的数据来源包括:江西财经大学公开数据集^[1]、Bilkent大学的火焰视频集、韩国启明大学火灾视频库、中国科学技术大学火灾科学国家实验室的数据集、BoWFire数据集^[25]以及使用互联网爬虫工具获取的火焰和烟雾图片,从中筛选出共计9613张图片,包含城市、野外、晴天、雾天、傍晚、夜间、户内、户外等多种复杂场景以及不同颜色和形状的火焰、烟雾。按照8:1:1比例划分为训练集、验证集和测试集。使用LabelImg图像标注工具标注对建立的数据集标注,标注格式为YOLO的txt文件,标注方式如图6所示。

为保证实验结果的准确性与严谨性,所有数据皆为3次实验后统计取平均值。

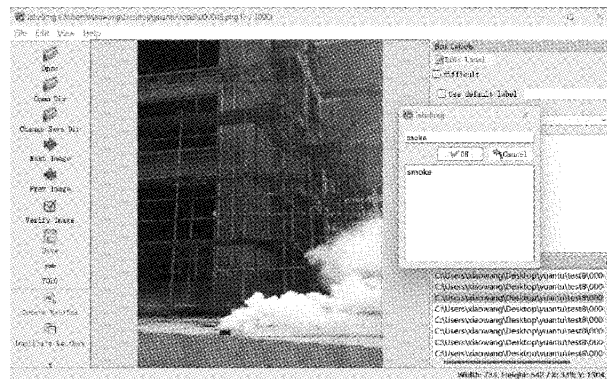


图6 标注示例图

3.2 实验环境

实验环境及配置如表2所示。

表2 实验环境及配置

实验环境	参数配置
硬件	Intel(R)Core(TM)i7-2600CPU@3.40 GHz; 内存 512 G;RTX3050Ti
软件	Ubuntu 18.04.1,CUDA 11.0,Pytorch1.9.0

使用改进后的Mosaic数据增强法对数据集进行训练,输入图像调整为 640×640 ,不满足时等比缩放。初始学习率0.01,SGD动量值0.937,权重衰减0.0004,批处理数16,训练次数200次。

3.3 实验

本文选取的评价指标有:准确度 P 、召回率 R 、平均精度均值 mAP 、帧率(frame per second, FPS)。

1) 消融实验

为验证本文方法的性能,对改进的YOLOv5s进行消融实验,选用增强后的数据集进行验证,实验结果如表3所示。

表3 消融实验

改进	评分模块	α -EIOU $\alpha=3$	准确度/%	召回率/%	平均精度均值/%	帧率
YOLOv5s	—	—	83.1	74.3	81.0	230
1	✓	—	86.4	75.9	84.3	221
2	—	✓	85.7	75.0	83.3	208
3	✓	✓	87.5	77.9	85.7	212

消融实验中,仅使用 Scoring module 情况下 mAP 提升 3.3%, FPS 下降 0.04%;仅使用 α -EIOU 情况下 mAP 提升 2.3%, FPS 下降 0.1%;改进后的 YOLOv5s 相比原模型 mAP 提高了 4.7%, FPS 小幅下降,在检测精度提升的同时仍然保持了实时性的优势。

2) 对比实验

为验证本文烟火检测方法的优劣,设计了一组对比实验,选择相同的数据集和实验环境,将本文方法与 YOLO 系列模型分别进行实验,得到表 4 实验结果。

可以看出, YOLOv3-Tiny 的检测精度最低,但是检测速度高于 YOLOv4, YOLOv4-Tiny 相比原 YOLOv4 在检测速度上提升明显,但检测精度下降了 5.4%, YOLOv5s 作为 YOLOv5 模型参数量最小的版本,对比 YOLOv4-Tiny 检测精度高 7.5%, FPS 提升 9 fps,而本文改进后的 YOLOv5s 相比原网络检测精度提高 5.4%, FPS 只是略

表 4 YOLO 系列实验结果对比

模型	准确度/ %	召回率/ %	平均精度 均值/%	帧率
YOLOv3-Tiny	71.6	55.7	64.6	82
YOLOv4	80.1	69.5	73.8	58
YOLOv4-Tiny	75.6	58.2	68.4	221
YOLOv5s	83.1	74.3	81.0	230
本文方法	88.5	82.4	87.0	212

微下降,仍然满足检测实时性的要求,在检测速度与检测精度之间取得了更优的平衡。

为进一步验证本文烟火检测方法的性能,选择目前烟火检测效果较好的其他改进型 YOLOv5s 模型以及其他主流目标检测算法进行对比实验,选择相同的数据集和实验环境,对比结果如表 5 所示。

表 5 不同方法实验结果对比

	模型	准确度/%	召回率/%	平均精度均值/%	帧率
改进型 YOLOv5s	文献[8]	83.9	75.5	82.0	98
	文献[9]	85.0	76.5	83.8	223
	文献[10]	89.2	80.6	87.1	64
	文献[11]	79.1	68.4	76.8	154
其他目标检测方法	VGG16+ResNet50 ^[26]	78.9	72.1	74.3	53
	SSD+MobileNetV2 ^[27]	80.7	66.8	70.3	125
	本文方法	88.5	82.4	87.0	212

表 5 中, VGG 和 SSD 目标检测方法在烟火检测上的 mAP 值分别为 74.3%、70.3%,相比之下, YOLO 系列算法在烟火检测领域效果更好,普遍为 80%以上(除文献[11]),并且 SSD 网络的 P 值高, R 值低, VGG 网络则刚好相反。检测速度上,由于 SSD 网络引入了 MobileNetV2, FPS 提升明显,但 mAP 值并不高, VGG 网络的 FPS 较低,但是 mAP 较高,也说明轻量级网络以检测精度的下降来换取更高的检测速度。

改进型的 YOLOv5s 中文献[9]的 FPS 最高,达到 223 fps,本文方法为 212 fps,其余方法均低于本文方法;文献 10 的 mAP 值为 87.1%,本文方法 mAP 值为 87%,其余方法均低于本文方法。分析原因是本文提出的评分模块虽然增加的参数量只有 1.169%,但在 Head 层引入的 α 权重参数降低了网络约 10%的 FPS ,所以两者叠加后,后者的负效应对网络的推理速度起主导作用,造成检测速度下降了 18 fps,但两者叠加后网络 mAP 值提高 4.7%。

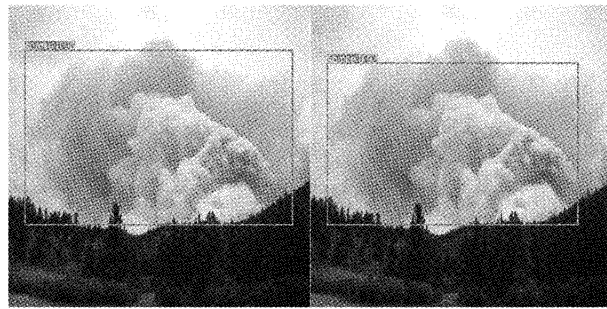
检测精度的提升通常意味着检测速度的下降,在对比实验中,本文方法对二者的平衡最优,实现检测速度和精度均较高。文献[9]是在 Neck 层基础上引入跨尺度连接进行特征融合,没有增加新的参数,因此检测速度没有降低,但检测精度提升较小;文献[10]在对 YOLOv5s 改进后

串联了 ResNet34 网络进行烟火细分类,最后结合图 7 多帧检测结果判断烟火目标,在检测精度上的表现最好,但也因为级联两个网络带来大量的运算参数导致网络的推理速度下降较多。

3) 检测结果

为了直观展现改进 YOLOv5s 算法的检测效果,选取了不同环境下的检测效果图,包括烟雾的干扰物、复杂环境、小目标、火焰的干扰物、夜间、遮挡这几类典型环境,最终检测效果如图 7 所示,其中左侧为 YOLOv5s 的检测效果图,右侧为改进 YOLOv5s 后的检测效果图。

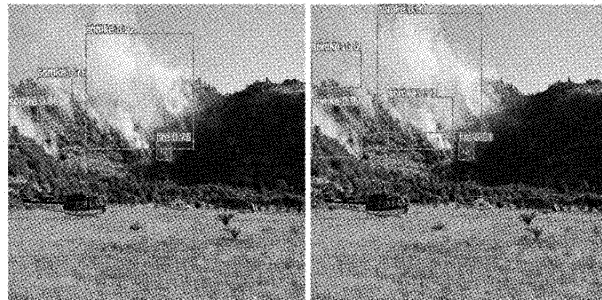
图 7(a)中存在强烈的日光,并且烟雾颜色呈现白色,容易和白色的天空背景混淆,本文方法依然取得了较高的置信度;图 7(b)图中工厂的黑色烟雾与一般火灾烟雾颜色差异较大,容易被漏检,白色的水柱和烟雾轮廓也极其相似,很容易造成误检,且伴随着火焰,本文方法依然对图像中烟雾和火焰区域给出了较高的置信度;图 7(c)对远距离情况下的烟火小目标进行检测,两种方法整体精度较高;图 7(d)图野草的颜色与火焰相似,又存在多个着火点,且图像中存在稀疏烟雾区域,本文方法对火焰较小的区域灵敏度更高,烟雾检测框也更加精确;图 7(e)图是夜间的火灾图,其中明亮的火焰区域都被准确找出;图 7(f)图中存



(a) 烟雾的干扰物



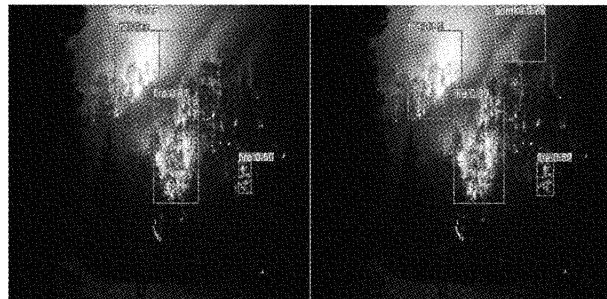
(b) 复杂环境



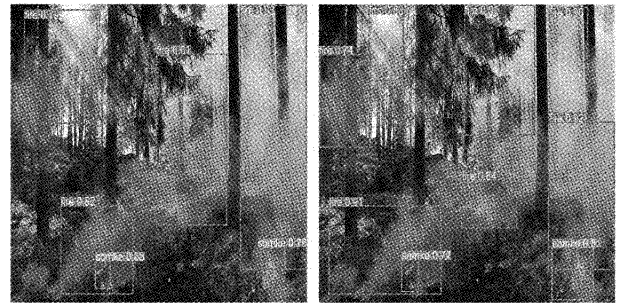
(c) 小目标



(d) 火焰的干扰物



(e) 夜间



(f) 遮挡

图7 YOLOv5s模型改进前后的效果图对比

在许多被遮挡的烟雾区域,且烟雾非常分散,但最终检测结果中可以看出YOLOv5s的检测框对层叠烟区分类较差,置信度不高,而本文方法的检测框能够准确捕捉到主要的烟区,且置信度更好。综上,本文算法的检测框在准确度上有明显的改善,在烟、火分类上更加精确,在复杂环境下有着较好的烟火检测精度。

4 结 论

针对复杂环境下烟火检测困难,检测精度低的问题,提出了一种改进YOLOv5s的烟火检测方法,相比原网络,主要改进有三方面,首先对Mosaic数据增强法进行改进,增强模型的泛化能力和提高训练效率,然后在Neck层的PAN结构中加入评分模块,对每个通道的特征进行评分,过滤低分数的特征,高分数的特征被保留进行特征融合,最后在Head层,将GIOU损失函数替换为 α -EIOU,进一步提高预测框的性能。对比实验的结果表明,本文提出的复杂环境下的烟火检测算法检测精度高,同时也保持了较高的检测速度,可以满足烟雾实时检测的要求。未来会扩大烟火数据集,寻求分辨率更高,环境更加复杂的视频和图片,以达到更好的检测效果。

参考文献

- [1] 史劲亭,袁非牛,夏雪. 视频烟雾检测研究进展[J]. 中国图象图形学报, 2018, 23(3): 303-322.
- [2] GAUR A, SINGH A, KUMAR A, et al. Video flame and smoke based fire detection algorithms: A literature review[J]. Fire technology, 2020, 56(5): 1943-1980.
- [3] YUAN F N, XIA X, SHI J T, et al. Non-linear dimensionality reduction and Gaussian process based classification method for smoke detection[J]. IEEE Access, 2017, 5: 6833-6841.
- [4] FRIZZI S, KAABI R, BOUCHOUICHA M, et al. Convolutional neural network for video fire and smoke detection[C]. IECON 2016-42nd Annual Conference of the IEEE Industrial Electronics Society, IEEE, 2016: 877-882.

- [5] 罗小权,潘善亮.改进YOLOV3的火灾检测方法[J].计算机工程与应用,2020,56(17):187-196.
- [6] 刘丽娟,陈松楠.一种基于改进SSD的烟雾实时检测模型[J].信阳师范学院学报(自然科学版),2020,33(2):305-311.
- [7] ZHANG Q X, LIN G H, ZHANG Y M, et al. wildland forest fire smoke detection based on faster R-CNN using synthetic smoke images [J]. Procedia Engineering, 2018, 211(1):441-446.
- [8] 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.
- [9] 宋华伟,屈晓娟,杨欣,等.基于改进YOLOv5的火焰烟雾检测[J].计算机工程,2023,49(6):250-256.
- [10] MUKHIDDINOV M, ABDUSALOMOV A B, CHO J. A wildfire smoke detection system using unmanned aerial vehicle images based on the optimized YOLOv5 [J]. Sensors, 2022, 22(23), DOI:10.3390/S22239384.
- [11] 张全,张伟,杨先凤,等.融合YOLOv5-ResNet级联网络的烟火检测方法[J].安全与环境学报,2023,23(2):397-405.
- [12] 常丽,张雪,蒋辉,等.融合YOLOv5s与SRGAN的实时隧道火灾检测[J].电子测量与仪器学报,2022,36(8):223-230.
- [13] REDMON J, FARHADI A. YOLOv3: An incremental improvement[J]. ArXiv Preprint, 2018, ArXiv:1804.02767.
- [14] 李澎林,章军伟,李伟.基于光流改进与YOLOv3的烟雾检测方法[J].浙江工业大学学报,2021,49(1):9-15.
- [15] 谢书翰,张文柱,程鹏,等.嵌入通道注意力的YOLOv4火灾烟雾检测模型[J].液晶与显示,2021,36(10):1445-1453.
- [16] 邵延华,张铎,楚红雨,等.基于深度学习的YOLO目标检测综述[J].电子与信息学报,2022,44(10):3697-3708.
- [17] 黄磊,杨媛,杨成煜,等.FS-YOLOv5:轻量化红外目标检测方法[J].计算机工程与应用,2023,59(9):215-224.
- [18] HAN K, WANG Y, TIAN Q, et al. GhostNet: More features from cheap operations[J]. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2020, 1577-1586, DOI:10.1109/CVPR42600.2020.00165.
- [19] NIALL M, JESUS M, PAUL M. 3-D human pose estimation using iterative conditional squeeze and excitation networks [J]. IEEE transactions on cybernetics, 2022, 52(1):687-699.
- [20] DROZDZAL M, VORONTSOV E, CHARTRAND G, et al. The importance of skip connections in biomedical image segmentation [J]. ArXiv Preprint, 2016, ArXiv:1608.04117.
- [21] ZHANG Y F, REN W Q, ZHANG Z, et al. Focal and efficient IOU loss for accurate bounding box regression [J]. Neurocomputing, 2022, 506:146-157.
- [22] HE J B, ERFANI S, MA X J, et al. Alpha-IOU: A family of power intersection over union losses for bounding box regression [J]. ArXiv Preprint, 2021, ArXiv:2110.13675.
- [23] 张欣怡,张飞,郝斌,等.基于改进YOLOv5的口罩佩戴检测算法[J].计算机工程,2023,49(8):265-274.
- [24] GUO R P, LIN C Q, LI C G, et al. Deep group-shuffling dual random walks with label smoothing for person reidentification [J]. IEEE Access, 2020, 8: 40018-40028, DOI:10.1109/access.2020.2976849.
- [25] 夏雪,袁非牛,章琳,等.从传统到深度:视觉烟雾识别、检测与分割[J].中国图象图形学报,2019,24(10):1627-1647.
- [26] 王洋,程江华,刘通,等.一种多网络模型融合的烟雾检测方法[J].计算机工程与科学,2019,41(10):1771-1776.
- [27] SANDLER M, HOWARD A, ZHU M, et al. Mobilenetv2: Inverted residuals and linear bottleNecks[C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018: 4510-4520.

作者简介

王铮帅(通信作者),硕士研究生,主要从事图像处理、烟雾检测方面的研究。

E-mail:wang16696370466@163.com

邱联奎,博士,副教授,主要从事图像处理、机器学习深度学习方面的研究。

E-mail:qiulk@haust.edu.cn

李迎港,硕士研究生,主要从事图像处理方面的研究。

E-mail:634974137@qq.com