

文章编号: 2095-2163(2023)05-0075-07

中图分类号: TP391.41

文献标志码: A

基于改进 YOLOv5s 的火灾烟雾检测算法研究

蔡静, 张讚, 冉光金, 李震, 李良荣

(贵州大学 大数据与信息工程学院, 贵阳 550025)

摘要: 为了解决火灾烟雾检测算法中存在的错检、漏检以及实时性差等问题, 提出了一种基于 YOLOv5s 的火灾烟雾检测模型。首先, 使用 Ghost Convolution 模块代替原 YOLOv5s 网络结构中的常规卷积模块, 在保持相同性能的基础上, 降低检测模型的计算成本、减少模型参数; 其次, 在原 YOLOv5s 模型骨干网络中加入 Vision Transformer 结构, 减少对卷积神经网络的依赖性, 同时提高获取全局和局部特征的能力; 最后, 引入 Coordinate Attention 注意力机制, 有效地提取特征信息, 进一步提高检测的准确率。实验结果表明, 所提出的火灾烟雾检测模型参数减少 17%, 准确率提高 0.73%, 检测速度提升 22.5%, 可以满足实际场景下的火灾烟雾检测。

关键词: 火灾烟雾检测; YOLOv5s; Vision Transformer; Coordinate Attention

The fire smoke detection algorithm based on improved YOLOv5s research

CAI Jing, ZHANG Zan, RAN Guangjin, LI Zhen, LI Liangrong

(College of Big Data and Information Engineering, Guizhou University, Guiyang 550025, China)

[Abstract] In order to solve the problems of false detection, missed detection and poor real-time performance in fire smoke detection algorithm, a fire smoke detection model based on yolov5s is proposed. Firstly, ghost revolution module is used to replace the conventional convolution module in the original yolov5s network structure to reduce the calculation cost and model parameters of the detection model on the basis of maintaining the same performance; Secondly, the vision transformer structure is added to the backbone network of the original yolov5s model to reduce the dependence on convolutional neural network and improve the ability to obtain global and local features; Finally, the coordinated attention mechanism is introduced to effectively extract feature information and further improve the accuracy of detection. The experimental results show that the parameters of the proposed fire smoke detection model are reduced by 18%, the accuracy is improved by 3.12%, and the detection speed is improved by 20%, which can meet the fire smoke detection in the actual scene.

[Key words] fire smoke detection; YOLOv5s; Vision Transformer; Coordinate Attention

0 引言

火灾是造成环境问题、人员伤亡和经济损失的重大灾害之一。在国外许多地方, 火灾的面积和强度都以惊人的速度增加。这不仅仅是火灾数量在增加, 火灾的性质也在发生变化, 尤其是在巴西和澳大利亚。例如, “地球之肺” 亚马逊雨林发生的大火, 已烧毁了总计 4920 平方公里的雨林, 面积超过 50 万个足球场, 给人类和自然环境带来了难以估量的破坏。因此, 为了有效控制和减轻火灾, 对火灾的萌

发进行预警尤为重要。对于火灾的预警, 与火灾火焰相比, 烟雾出现时间更早, 传播速度更快, 体积更大, 更易于识别。

传统的烟雾检测主要是基于对纹理、温度、颜色、运动特征和空气透明度的分析^[1], 最常见的烟雾探测器是基于红外/紫外摄像头。基于此, Deldjoo Y 等人^[2]使用稳健的烟雾特征评估方法区分烟雾和非烟雾运动对象, 利用模糊推理系统, 以模糊化的方式将烟雾特征结合起来进行最终评估, 并决定何时发出火灾警报。J Qian 等人^[3]利用延时参数改进高斯混合

基金项目: 国家自然科学基金(61361012); 贵州省科技计划项目(黔科合平台人才[2017]5788)。

作者简介: 蔡静(1997-), 女, 硕士研究生, 主要研究方向: 电路与系统、目标检测; 张讚(1997-), 男, 硕士研究生, 主要研究方向: 高性能计算; 冉光金(1997-), 男, 硕士研究生, 主要研究方向: 神经网络压缩、电路与系统; 李震(1987-), 男, 博士, 高级实验师, 主要研究方向: 电路与系统、密码学; 李良荣(1963-), 男, 学士, 教授, 主要研究方向: 电路与系统。

通讯作者: 李良荣 Email: lrli@gzu.edu.cn

收稿日期: 2022-06-15

模型提取候选烟雾区域,通过烟雾的面积变化率和运动方式等运动特征,从候选区域中选择烟雾区域。Yuan F 等人^[4]利用 Gabor 网络对烟雾图像进行识别以及对其纹理信息进行分类。殷梦霞等人^[5]利用烟雾图像块的 HSV (Hue、Saturation、Value) 颜色特征、能量特征等物理特征,提出了基于多特征融合的自适应烟雾检测算法。刘长春等人^[6]根据可见光视频图像处理原理,以及烟雾块的纹理特征、HSV 颜色空间等物理特征,提出了一种基于局部区域图像动态特征的火视频烟雾检测方法。王伟刚等人^[7]提出一种 TDFP (Triple Multi Feature Local Binary Patterns and Derivative Gabor Feature Fusion) 的烟雾检测算法,采用 T-MFLBP (Triple Multi-Feature Local Binary Patterns) 算法获得烟雾的纹理特征,利用高斯核函数进一步优化图像边缘灰度信息,最后对融合后的特征进行训练,对烟雾区域进行识别。

但是,当背景复杂以及影响因数较多时,传统的检测算法会受到限制。由于烟雾为非刚性物体,其形状、颜色和纹理等物理特征容易随时间变化,导致无法提取烟雾的最本质特征^[8],且传统的烟雾检测传感器大多用于室内,难以在森林和草原等户外场景中发挥有效作用,有时也会有较高的错检率以及误警率^[9]。

得益于各种人工智能领域的进步,图像处理和计算机视觉等基于视觉的研究领域已经取得了一定的成果,基于计算机视觉的火灾烟雾检测模型也通过这些技术得到了改进。与传统的检测模型相比,基于计算机视觉的火灾检测模型在成本、准确性、鲁棒性和可靠性等方面都具有许多优势。Zhang^[10]和张倩^[11]等人利用 Faster R-CNN 对烟雾图像进行识别检测。Lee Y 等人^[12]利用三帧差分算法和均方误差获得输入帧图像,通过 Faster R-CNN 提取火焰和烟雾候选区域,最后利用局部 HSV 和 RGB 颜色直方图,确定最终的火灾和烟雾区域。谢书翰等人^[13]通过改进 YOLOv4 模型对火灾烟雾图片进行了检测识别。Saponara S 等人^[14]提出了一个非常轻量级的神经网络 FireNet,开发了一个物联网 (IoT) 火灾探测单元取代当前基于物理传感器的火灾探测器,在 Raspberry Pi 3B 等嵌入式设备上训练;联合物联网功能允许探测单元在发生火灾紧急情况时,向用户提供实时视觉反馈和火灾警报。刘丽娟等人^[15]通过改进 SSD 算法对火灾烟雾图像进行识别。叶寒雨等人^[16]将光流估计与 YOLOv4 算法结合,提出了 SmokeNet 算法对烟雾进行检测。

基于计算机视觉的检测算法虽然在精度和速度上都提高,但是当背景复杂、检测目标距离较远时,仍然存在漏检以及错检等问题。针对以上问题,本文提出一种基于改进的 YOLOv5s 的火灾烟雾检测算法,在 YOLOv5s 的骨干网上引入 Vision Transformer 模块,增强其特征提取能力;使用 Ghost Block 卷积代替常规卷积,减少模型参数和计算量。实验结果表明,该算法的检测性能均提高且模型参数显著减少。

1 相关工作

根据型号的大小, YOLOv5 分为 YOLOv5s、YOLOv5m、YOLOv5l 和 YOLOv5x 4 个版本。模型越大,单个图像的精度越高,检测时间越长。由于 4 种版本只是在其模型宽度和深度的不同,且对于火灾烟雾检测而言,不仅对检测精度有极高的要求,对检测速度也要求实现实时检测,所以选用 YOLOv5s 作为本文的基本检测模型。YOLOv5s 在数据输入端使用了 Mosaic 数据增强^[17]、自适应锚计算和自适应图像缩放等技术,来提高对小目标的检测。主干使用了跨阶段局部网络 (CSP)^[18] 结构,提取输入图像特征,头部主要解决从主干提取特征映射的定位问题,并执行类概率预测。Neck 结构由特征金字塔网络 (FPN)^[19] 和路径聚合网络 (PAN)^[20] 组成, Neck 结构是连接主干和头部的部分,主要是对特征图进行细化和重构。此外 YOLOv5s 的 Neck 结构中,还借鉴了 CSPNet 中设计的 CSP2_X 结构,加强网络特征融合的能力^[21]。YOLOv5s 的网络结构如图 1 所示, YOLOv5s 结构中的组件结构如图 2 所示。

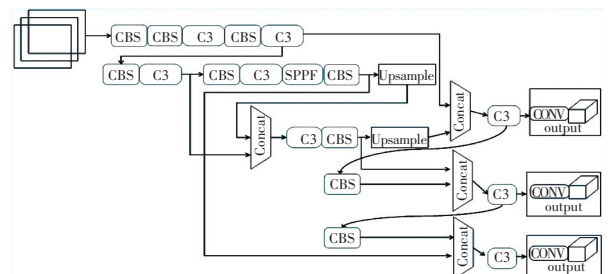


图 1 YOLOv5s 结构

Fig. 1 YOLOv5s structure

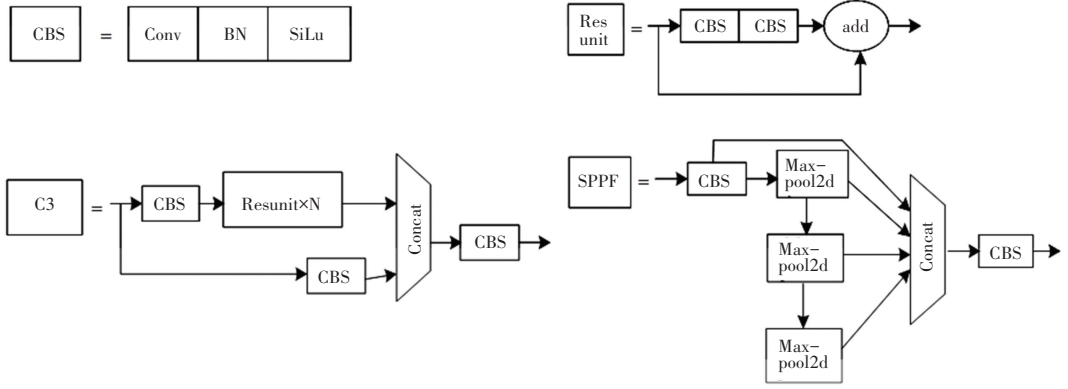


图 2 YOLOv5s 组件图

Fig. 2 YOLOv5s component diagram

2 算法的改进

2.1 Ghost Convolution 模块

在深层网络中常规的特征提取方式会堆叠大量的卷积层,产生丰富的特征图,但是同时需要保存大量参数,占用其有限内存,不利于在嵌入式设备上部署网络模型。本文使用一种 Ghost Convolution 模块^[22]替换原网络中的常规卷积模块,可以通过较少的参数来提取特征。与常规卷积相比,在不改变输出特征映射大小的情况下,该 Ghost Convolution 模块所需的参数和计算量都有所降低,其结构如图 3 所示。

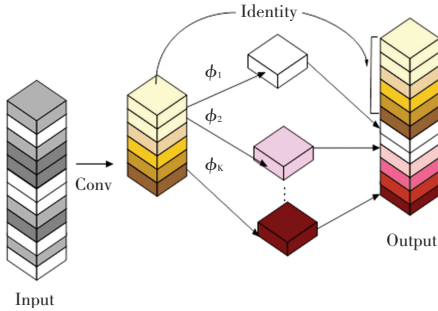


图 3 Ghost Convolution 模块

Fig. 3 Ghost Convolution module

Ghost Convolution 模块通过常规卷积生成几个固有的特征映射,然后使用线性运算来扩展特征和增加通道。给定输入数据 $X \in R^{c \times h \times w}$,其中 c 为输入通道数, h, w 是 X 的高度和宽度,任意卷积层生成特征图的操作可表示为

$$Y = X * f + b \quad (1)$$

式中: $*$ 表示卷积运算, b 表示偏差, $Y \in R^{h' \times w' \times n}$ 是具有 n 个通道的输出特征图, $f \in R^{c \times k \times k \times n}$ 是这一层中的卷积滤波器。此外, h' 和 w' 表示输出特征图 Y 的高度和宽度, $k \times k$ 表示 f 的核大小。

使用常规卷积生成 m 个固有特征图的操作可表示为

$$Y' = X * f' \quad (2)$$

其中, $f' \in R^{c \times k \times k \times m}$ ($m \leq n$) 表示所用滤波器,滤波器的大小、填充、步幅等超参数与式(1)中的常规卷积相同。

对 y 中每个固有的特征图进行线性运算,生成 s 个 Ghost 特征,其计算公式可表示为

$$y_{ij} = \Phi_{i,j}(y'_i), \forall i = 1, \dots, m, j = 1, \dots, s, \quad (3)$$

式中: y'_{ij} 是 Y' 中的第 i 个特征图, $\Phi_{i,j}$ 是用于生成第 j 个 Ghost 特征图 y_{ij} 的第 j 个线性运算, $\Phi_{i,s}$ 用于保存内在特征映射的恒等映射。

由公式(3)得到 $n = m \cdot s$ 个特征图 $Y = y_{11}, y_{12}, \dots, y_{ms}$ 作为 Ghost 模块的输出数据。

2.2 网络改进

2.2.1 Vision Transformer 模块

在机器视觉领域中,为了改善卷积神经网络(Convolutional Neural Network, CNN)^[23]在低层特征依赖关系范围较大时存在的局限性。Dosovitskiy A 等人^[24]在 Transformer 的基础上提出了 Vision Transformer。与 CNN 网络不同, Vision Transformer 在获取全局信息能力上很强大,将 Vision Transformer 与 CNN 模型结合,通过自注意力机制改善底层特征提取。Vision Transformer 网络模块结构如图 4 所示。

为了将图片转换为标准的 Transformer 编码器,用于处理序列数据,将输入二维图像 $x \in R^{H \times W \times C}$ 分割成固定大小的图片 $x_p \in R^{N \times (P^2 \cdot C)}$ 。其中 (H, W) 是原始图像的分辨率, C 表示通道数, (P, P) 是每个图片块(patch)的分辨率, $N = HW/P$ 表示获得的图

片块总数,其也用作变换器的有效输入序列长度;然后线性嵌入每个面片,添加位置嵌入,并将生成的矢量序列提供给标准的 Transformer 编码器;再将 Transformer 的第一个输出送入 MLP Head 得到预测结果。此外,在输入的序列数据之前添加了一个分类标志位(class),可以更好地表示全局信息。

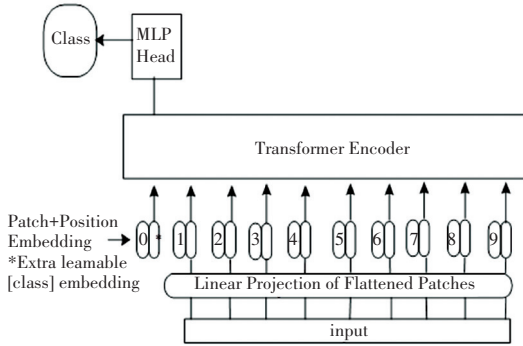


图4 Vision Transformer 模块

Fig. 4 Vision Transformer module

2.2.2 嵌入 Coordinate Attention 注意力

在某些情况下,摄像头捕捉到的物体图像较小,而 Yolov5s 模型对小目标的检测效果欠佳,此时模型预测易受颜色、亮度等因素的影响。因此,本文引入 Coordinate Attention 注意力机制^[25],可以有效地提高骨干网络的特征提取能力,进一步提升准确率。

Coordinate Attention 模块可以视为一个计算单元,有效提升网络的表达能力,并且可以充分利用捕获的位置信息,从而准确地定位感兴趣的区域。Coordinate Attention 注意力的实现过程如图 5 所示。

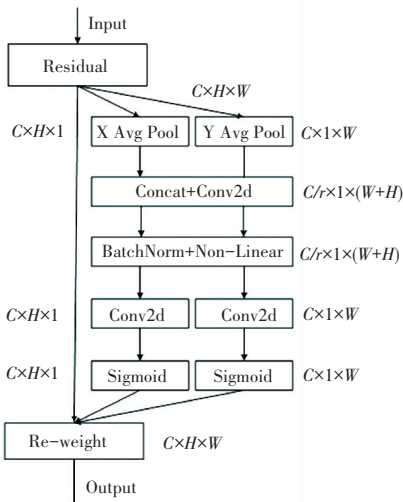


图5 CA 模块结构图

Fig. 5 CA module structure diagram

Coordinate Attention 注意力为了获取输入图像

全局上的注意力,并对每个通道信息进行编码。首先将输入特征图分为宽度和高度两个方向进行全局平均池化,获得在宽度和高度两个方向的特征图。其中,分别逐通道使用两个大小为 $(H, 1)$ 和 $(W, 1)$ 池化核,在高度为 h 的第 c 个通道的输出可表示为

$$z_c^h(h) = \frac{1}{W} \sum_{0 \leq i < W} x_c(h, i) \quad (4)$$

其中, $z_c^h(h)$ 表示在高度为 h 的第 c 个通道的输出特征图。

同样,在宽度为 w 的第 c 个通道的输出特征图 $z_c^w(w)$ 可以表示为

$$z_c^w(w) = \frac{1}{H} \sum_{0 \leq j < H} x_c(j, w) \quad (5)$$

通过式(4)、式(5)获得全局感受野并编码精确的位置信息,两种变换分别沿空间两个方向聚合特征,生成一对特征映射张量,这有助于网络更准确地定位感兴趣的对象。

随后将两个方向的特征图进行融合,使用大小为 1×1 的卷积模块,把维度降低为原来的 C/r ,再通过批量归一化处理送入非线性激活函数中获得输出特征图,如式(6):

$$f = \delta(F_1([z^h, z^w])) \quad (6)$$

其中, $[\cdot, \cdot]$ 表示沿空间维度的融合操作, F_1 为 1×1 卷积, $f \in R^{C/r \times (H+W)}$ 表示在水平方向和垂直方向上编码空间信息的中间特征图, δ 是非线性激活函数, r 表示下采样比例。然后沿着空间维度将 f 分解成两个独立的张量 $f^h \in R^{C/r \times H}$ 和 $f^w \in R^{C/r \times W}$,再使用两个 1×1 卷积 F_h 和 F_w 将特征图变换为与输入 X 相同数量的通道。得到结果如式(7)、式(8):

$$g^h = \delta(F_h(f^h)) \quad (7)$$

$$g^w = \delta(F_w(f^w)) \quad (8)$$

式中: δ 为 Sigmoid 函数, g^h, g^w 分别用作两个空间上注意力的权重。最后在原始特征图上通过乘法加权计算,将得到在宽度和高度方向上带有注意力权重的特征图,如式(9):

$$y_c(i, j) = x_c(i, j) \times g_c^h(i) \times g_c^w(j) \quad (9)$$

2.3 改进算法结构

本文将改进模块加入到 YOLOv5s 模型中,优化结构如图 6 所示。

图 7 中,GBS 表示使用 Ghost Convolution 模块替换原网络中的常规卷积模块的结构图,C3_TR 表示使用 Transformer 替换原 CSP 结构中的 Resunit 模块的结构图。

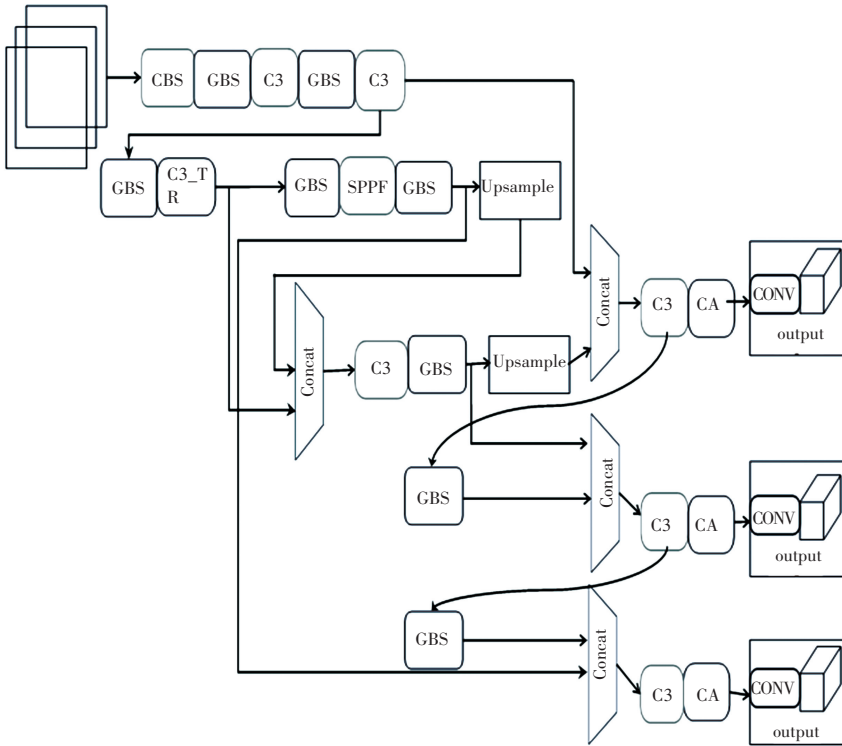


图 6 改进后的结构

Fig. 6 Improved structure diagram

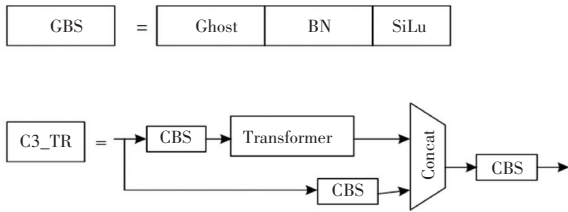


图 7 改进结构图的组件图

Fig. 7 Component diagram of improved structure diagram

中部分火灾烟雾如图 8 所示。



图 8 部分数据集

Fig. 8 Partial Dataset

3 实验对比

3.1 实验环境

实验的操作系统为 windows7 系统,训练框架为 PyTorch1.7.1。CPU 为 8 GB RAM 的 Intel(R) Core (TM) i7 - 6700k @ 4.00 GHz, GPU 为 NVIDIA GeForce GTX 970。

3.2 火灾烟雾数据集

由于公开的火灾烟雾数据集较少,通过互联网采集数据,最终经过清洗获得 20 114 张图片,随机选取 18 287 张作为训练集,1 827 张作为测试集。数据集图像包含建筑、草原、森林、车辆起火 4 种场景,并且涵盖白天、黑夜以及其它背景干扰因素。数据集

3.3 评价指标

实验使用召回率 (Recall)、精准率 (Precision)、平均精度 AP (Average Precision)、平均精度均值 mAP (mean Average Precision) 来评价检测模型准确性。评价指标计算公式如下:

$$Precision = \frac{TP}{TP + FP} \tag{10}$$

$$Recall = \frac{TP}{TP + FN} \tag{11}$$

$$AP = \int_0^1 p(r) dr \quad (12)$$

$$mAP = \frac{1}{m} \sum_{i=1}^m AP_i \quad (13)$$

其中, m 表示样本类别数; $p(r)$ 表示 *Precision* 以 *recall* 为参数的一个函数; *TP* (True Positives) 是指被正确识别的正样本; *TN* (True Negatives) 为被正确识别的负样本; *FP* (False Positives) 表示负样本被错误识别为正样本; *FN* (False Negatives) 表示正样本被错误识别为负样本。

3.4 实验结果与分析

本文主要从实际检测结果图进行对比分析,检测对比数据见表1,实际检测结果如图9、图10所示。

表1 改进前后算法测试结果对比表

Tab. 1 Comparison of test results before and after the improvement

模型	<i>Precision</i>	<i>Recall</i>	<i>mAP</i>	速度/ (帧·s ⁻¹)	模型参数/ M
YOLOv5s	80.41	62.71	71.82	41	7.06
本文	83.52	63.34	72.55	50	5.87



图9 YOLOv5s 检测结果

Fig. 9 YOLOv5s test result chart



图10 改进 YOLOv5s 检测结果

Fig. 10 Image of improved YOLOv5s test results

表1中, mAP 为 $mAP@0.5$ 值。通过测试结果可以得出,改进后 mAP 提高了 0.73%,检测速度提高了 22.5%,模型参数比原网络参数减少了 17%。

通过图9、图10对比结果可知,改进后的 YOLOv5s 模型在不同场景中对烟雾、火焰目标算法都可以有效地解决原 YOLOv5s 存在的漏检、错检。

在相同的运算环境下,将深度学习算法 YOLOv4-Tiny、传统单阶段算法 SSD、原网络 YOLOv5s 同本文算法进行比较分析。算法的对比结果见表2。

表2 检测算法测试结果对比表

Tab. 2 Comparison table of test results of detection algorithm

模型	<i>Precision</i>	<i>Recall</i>	<i>mAP</i>	速度/ (帧·s ⁻¹)
SSD	75.34	51.43	61.32	23
Mobilenetv2-SSD	82.82	58.43	71.34	38
YOLOv4-tiny	85.20	47.87	67.10	62
YOLOv5s	80.41	62.71	71.82	41
本文	83.52	63.34	72.55	50

由表2可见,在检测精度和检测速度上, YOLOv4-tiny 检测算法相对本文改进的模型有一定的优势,而检测速度的快慢与算法的模型参数量等因素成反比,即模型参数量越小,检测速度则越快。YOLOv4-tiny 是 YOLOv4 的轻量化网络架构,极大地减少了模型参数量以及降低了模型的计算量,提高了检测速度,但是该算法的总体性能略低。

通过对比实验可以看出,本文算法总体上优于其它几种算法,该检测算法在实际工程应用中更具优势。

4 结束语

本文提出了一种基于 YOLOv5s 的火灾烟雾检测模型,在 YOLOv5s 模型的基础上进行优化,并引入注意力机制、Vision Transformer 模块提高检测精度,解决目标漏检和错检等问题;引入 Ghost Convolution 模块,减少模型参数,提高火灾烟雾检测算法的性能。与同一环境下的其他检测模型相比,改进后的模型在识别精度和速度上都有良好优势。

但是,本文的召回率还有很大的进步空间,在之后的工作中,在保证准确率的前提下,提升召回率;其次,还应扩充火灾烟雾数据集,丰富火灾烟雾检测场景,增强火灾烟雾图片质量。

参考文献

[1] CHEN T H, KAO C L, CHANG S M. An intelligent real-time

- fire-detection method based on video processing[C]// IEEE 37th Annual 2003 International Carnahan Conference on Security Technology, 2003. Proceedings. IEEE, 2003: 104-111.
- [2] DELDJOO Y, NAZARY F, FOTOUHI A M. A novel fuzzy-based smoke detection system using dynamic and static smoke features [C]//2015 23rd Iranian Conference on Electrical Conference on Electrical Engineering. IEEE, 2015:729-733.
- [3] QIAN J, FU J, QIAN J, et al. Automatic early forest fire detection based on gaussian mixture model [C]//2018 IEEE 18th international conference on communication technology (ICCT). IEEE, 2018: 1192-1196.
- [4] 殷梦霞,王理,孙连营. 基于多特征融合的自适应烟雾检测算法[J]. 建筑科学,2019,35(9):26-31.
- [5] LU X, ZHANG X, LI F, et al. Detection of Fire Smoke Plumes Based on Aerosol Scattering Using VIIRS Data over Global Fire-Prone Regions[J]. Remote Sensing, 2021, 13(2):196.
- [6] 刘长春,刘鹏举,季烨云. 基于视频区域动态特征的林火烟雾检测技术研究[J]. 北京林业大学学报,2021,43(1):10-19.
- [7] 王伟刚,王炳蔚,张云伟. TDFP:一种强鲁棒性的烟雾图像检测算法[J]. 激光与光电子学进展,2021,58(4):273-280.
- [8] 朱弥雪,刘志强,张旭,等. 林火视频烟雾检测算法综述[J/OL]. 计算机工程与应用: 1-13[2022-05-20]. <http://kns.cnki.net/kcms/detail/11.2127>.
- [9] TANG Y, GE L. Development of scattering-style infrared smog photoelectric detector [J]. Journal of Sichuan University (Engineering Science Edition), 2002, 34(4): 117-120.
- [10] 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, 211:441-446.
- [11] 张倩,周平平,王公堂,等. 基于合成图像的 Faster R-CNN 森林火灾烟雾检测[J]. 山东师范大学学报(自然科学版), 2019, 34(2): 180-185.
- [12] LEE Y, SHIM J. Deep learning and color histogram based fire and smoke detection research [J]. International journal of advanced smart convergence, 2019, 8(2): 116-125.
- [13] 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.
- [14] SAPONARA S, ELHANASHI A, GAGLIARDI A. Real-time video fire/smoke detection based on CNN in antifire surveillance systems[J]. Journal of Real-Time Image Processing, 2021, 18(3): 889-900.
- [15] 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.
- [16] 叶寒雨,李传昌,刘森,等. 基于稠密光流和目标检测的烟雾检测算法[J/OL]. 电光与控制:1-7[2022-05-20]. <http://kns.cnki.net/kcms/detail/41.1227.tn.20220314.1205.006.html>.
- [17] BOCHKOVSKIY A, WANG C Y, LIAO H Y M. Yolov4: Optimal speed and accuracy of object detection[J]. arXiv preprint arXiv: 2004.10934, 2020.
- [18] WANG C Y, LIAO H Y M, WU Y H, et al. CSPNet: A new backbone that can enhance learning capability of CNN [C]// Proceedings of the IEEE/ CVF conference on computer vision and pattern recognition workshops. 2020: 390-391.
- [19] LI X, LAI T, WANG S, et al. Weighted feature pyramid networks for object detection [C]// 2019 IEEE Intl Conf on Parallel & Distributed Processing with Applications, Big Data & Cloud Computing, Sustainable Computing & Communications, Social Computing & Networking (ISPA/ BDCloud/ SocialCom/ SustainCom). IEEE, 2019: 1500-1504.
- [20] LIU S, QI L, QIN H, et al. Path aggregation network for instance segmentation [C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2018: 8759-8768.
- [21] 王立辉. 基于卷积神经网络的行人检测与跟踪算法研究[D]. 湖北:武汉科技大学,2021.
- [22] KETKAR N. Convolutional neural networks[M]. Deep Learning with Python. Apress, Berkeley, CA, 2017: 63-78.
- [23] 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 (CVPR), 2020: 1577-1586.
- [24] DOSOVITSKIY A, BEYER L, KOLESNIKOV A, et al. An image is worth 16x16 words: Transformers for image recognition at scale[J]. arXiv preprint arXiv: 2010.11929, 2020.
- [25] HOU Q, ZHOU D, FENG J. Coordinate attention for efficient mobile network design [C]// Proceedings of the IEEE/ CVF Conference on Computer Vision and Pattern Recognition. 2021: 13713-13722.

(上接第74页)

- [9] 牛群峰,袁强,靳毅,等. 基于改进VGG-16卷积神经网络的烟丝类型识别[J]. 国外电子测量技术,2022,41(9):149-154.
- [10] 陈嘉,季雪,阙云,等. 改进型VGG算法对小样本路面破损的分类识别[J/OL]. 湖南大学学报(自然科学版):1-11[2023-03-20]. <http://kns.cnki.net/kcms/detail/43.1061.N.20221201.1259.001.html>.
- [11] 季菁菁,奚峥皓,李忠峰. 基于深度迁移学习的煤岩显微组分识别算法[J]. 智能计算机与应用,2023,13(2):92-97,102.
- [12] 王桂平,陈旺桥,杨建喜,等. 基于迁移学习的桥梁表观病害检测技术研究[J]. 铁道科学与工程学报,2022,19(6):1638-1646.
- [13] 周义飏. 基于VGG的人脸表情识别与分类[J]. 智能计算机与应用,2021,11(9):35-41.
- [14] 张士豹,王文韬. 基于改进VGG模型的人脸表情识别研究[J]. 现代信息科技,2021,5(23):100-103.
- [15] 王江晴,冀星,莫海芳,等. 基于轻量化VGG的植物病虫害识别[J]. 中国农机化学报,2022,43(4):25-31.
- [16] 周安众,罗可. 一种卷积神经网络的稀疏性Dropout正则化方法[J]. 小型微型计算机系统,2018,39(8):1674-1679.
- [17] GOLLE P. Machine learning attacks against the Asirra CAPTCHA [C]//Proceedings of the 15th ACM conference on Computer and communications security. 2008: 535-542.