

文章编号: 2095-2163(2022)03-0158-04

中图分类号: TP399

文献标志码: A

基于YOLOv5网络模型的火焰检测

涂沛驰¹, 傅钰雯², 熊宇璇³, 杨健晟¹

(1 贵州大学 电气工程学院, 贵阳 550025; 2 贵州交通职业技术学院 物流系, 贵阳 550025;

3 贵州大学 外国语学院, 贵阳 550025)

摘要: 煤炭资源在开采的过程中会伴随着产生一种名为煤层气的产物,煤层气又称为瓦斯,若将瓦斯直接排放至大气中,则会造成严重的温室效应,研究发现瓦斯可以通过燃烧用来发电,但瓦斯在发电过程中遇到明火,则会发生爆炸,给工作人员及企业会造成不可估量的损失,因此检测瓦斯发电站内的火焰情况,成为了解决瓦斯发电站爆炸事故的重要目标。基于火焰识别问题,采用传统目标检测算法难以满足精度要求,后续出现了基于深度学习的两步目标检测算法,虽在识别精度上能有效满足要求,但检测的实时性上存在不足。本文综合考虑目标检测的检测精度以及检测实时性,采用了最新的YOLOv5模型进行火焰的实时检测。

关键词: 火焰检测; YOLOv5; 目标检测; 深度学习

Flame detection based on YOLOv5 network model

TU Peichi¹, FU Yuwen², XIONG Yuxuan³, YANG Jiansheng¹

(1 College of Electrical Engineering, Guizhou University, Guiyang 550025, China;

2 College of Logistics, Guizhou Communication Vocational College, Guiyang 550025, China;

3 College of Foreign Languages, Guizhou University, Guiyang 550025, China)

[Abstract] In the process of exploitation of coal resources can produce a product called coalbed methane, with gas, coalbed methane is also known as if gas direct emissions to the atmosphere, will cause serious greenhouse effect, the study found that the gas can be used to generate electricity by burning, but the gas power generation in flame, will be an explosion, to the staff and enterprise can create an immeasurable loss, Therefore, the detection of the flame in the gas power station has become an important target to solve the gas power station explosion accident. Based on the flame identification problem, it is difficult to meet the accuracy requirements of the traditional target detection algorithm. The subsequent two-step target detection algorithm based on deep learning appears. Although it can effectively meet the requirements in recognition accuracy, it has shortcomings in real-time detection. In this paper, the latest YOLOv5 model is used to detect flame in real time.

[Key words] flame detection; YOLOv5; object detection; deep learning

0 引言

近年来,随着现代工业的不断发展,发电站的发电量与日俱增。其中,燃煤发电作为主要的发电方式之一,有着十分重要的地位。在煤矿的开采过程中,会伴随着采出一种伴生资源煤层气,煤层气俗称瓦斯。在煤矿的开采过程中,如果将瓦斯直接排放至大气中,将会产生十分严重的温室效应。研究发现瓦斯热值高于通用煤1~4倍,与天然气相当,且燃烧后较为纯净,不会产生工业废气。因此,瓦斯可以作为一种良好的工业化工发电的燃料,但当瓦斯

中空气浓度达到5%~16%时,遇到火焰会发生爆炸,这也是瓦斯发电站爆炸事故的根源。在瓦斯发电站中,火焰可能导致严重安全事故的发生,目前大多数工厂仍采用人工巡检的方式对瓦斯发电站进行检查,巡检效率会随着时间的延长而降低;由于火势蔓延十分的迅速,人工巡检很难早期发现并及时扑灭火焰。因此,急需一种能实时定位、识别火焰的方法来监测瓦斯发电站内可能出现的火焰^[1-2]。

与人工巡检监测火焰的方式相比,利用计算机视觉检测火焰具有更高的效率。张汝峰等人^[3]以及苏展等人^[4]将火焰图片经过图像预处理操作后,

基金项目: 贵州省科学技术基金(黔科合基础[2018]1030); 贵州省教育厅创新群体(黔教合KY字[2021]012)。

作者简介: 涂沛驰(1995-),男,硕士研究生,主要研究方向:机器视觉、人工智能;傅钰雯(1988-),女,硕士,讲师,主要研究方向:大数据分析;熊宇璇(1999-),女,硕士研究生,主要研究方向:非文学翻译;杨健晟(1987-),男,博士,副教授、硕士生导师,主要研究方向:机器视觉检测技术、燃烧诊断技术。

通讯作者: 杨健晟 Email:jsyang3@gzu.edu.cn

收稿日期: 2021-11-01

再通过采用 RGB、GBR 与 HSI 两两结合的颜色模型,根据火焰的颜色、亮度等进行检测;何爱龙等人^[5]采用视频检测及图像处理的方法,通过将火焰图片进行预处理后提取火焰区域,并选取其中较为明显的火焰特征构建火焰特征工程,进行火焰的识别;侯易呈等人^[6]通过采用 Faster R-CNN (Faster Region-Convolutional Neural Networks) 网络与深度残差网络相结合的方法提取火焰特征,采用激励网络减少相关度较低的冗余特征,最后使用多尺度特征融合结构,丰富深层特征,以此来提高火焰识别的精度。但以上方法,对火焰的检测主要集中在对火焰特征的提取,进而以良好的准确率识别检测火焰,但在现实生活中,火势蔓延十分迅速,故火焰检测的实时性也是十分重要的一个环节。

随着深度学习的发展,基于深度学习的目标检测算法发展迅速,其主要分为单步目标检测算法以及双步目标检测算法。双步目标检测算法分为图像的识别和定位两个步骤,首先使用区域候选算法从待检测图像中提取候选区域;通过卷积神经网络对图像进行特征提取及分类识别。双步目标检测算法具有较高的准确率,但是计算量大,运算速度慢,针对于实时性要求高的目标检测不太适用。单步目标检测算法是将提取候选区域以及分类识别两个部分融合到一个网络里,直接由输入图像得到图像中存在的物体类别及位置。单步目标检测算法,在精度不会丢失太多的同时兼具了运算速度快,实时性好,对硬件要求低等优点,被广泛应用于目标实时检测领域中。

1 YOLOv5 模型的组成及原理

YOLO (You Only Look Once) 算法是一种单步端到端的目标检测算法,将提取候选区域及识别分类融合在一起,具有检测速度快、模型文件小等优点。通过更新迭代,现如今更新到了 YOLOv5 模型。YOLOv5 模型共有 YOLOv5s、YOLOv5m、YOLOv5l 和 YOLOv5x 4 个版本,其中 YOLOv5s 网络是 YOLOv5 系列中深度最小、模型文件最小,是其他版本网络的基础。综合检测的实时性与模型准确性,本文采用 YOLOv5s 网络作为火焰检测的网络。

YOLOv5 的网络架构主要可分为 4 部分,依次为输入端、骨干网络、颈部网络、输出端^[7]。

1.1 输入端

YOLOv5 网络的输入端采用 Mosaic 数据增强、自适应锚框计算、自适应图片缩放 3 种算法。

Mosaic 数据增强方法在输入的数据中随机选取 4 张图片进行随机的缩放、裁剪及排列。可将数据扩容,使得网络的鲁棒性得以提升;其次,通过图片的拼接可以一次计算 4 张图片的数据,从而提升模型的训练速度,降低模型的内存要求。Mosaic 数据增强操作原理如图 1 所示。

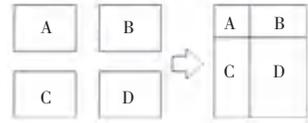


图 1 Mosaic 数据增强操作原理

Fig. 1 Mosaic data enhancement operation principle

自适应锚框计算方法是在建立的数据集中设定指定长宽的锚点框,以输入图像的每一个像素作为中心点,以此生成多个边界框。当网络进入到训练阶段时,会在初始锚点框上输出与其对应的预测框,计算其与真实框的差距,进行反向更新,进而迭代更新网络参数。

1.2 骨干网络

YOLOv5 网络的骨干网络采用了 Focus 结构及 CSP (Cross Stage Partial) 结构。Focus 结构对输入图片切片操作,每一张图片采集某个像素点的值后,不采集相邻像素点的值,而是中间隔开一个像素点,采集下一个像素点,从而扩充输入通道数;最后,将获得的拼接图片通过卷积操作,得到没有特征信息丢失的二倍下采样特征图。Focus 切片操作原理如图 2 所示。

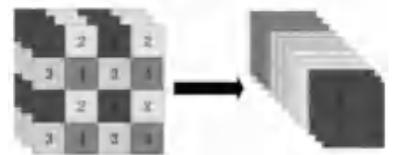


图 2 Focus 切片操作原理

Fig. 2 Focus slicing operation principle

以 YOLOv5s 为例,输入图像尺寸为 $640 \times 640 \times 3$,将图像输出至 Focus 结构中,通过切片操作,得到 $320 \times 320 \times 12$ 的特征图,最后经过卷积核为 32 的卷积层,得到 $320 \times 320 \times 32$ 的特征图。

CSP 结构在 YOLOv5s 网络中有两处应用,Csp_1 结构应用于骨干网络中,Csp_2 结构则是应用于颈部网络中。Csp_1 结构中包含了 CBL (Conv + BN + LeakyRelu) 模块以及多个残差组件 (Res unit)。CBL 模块与残差组件的组成,如图 3 所示。

残差组件是借鉴了 Resnet 网络中的残差结构,应用于较深的网络中,可以解决网络层数过深而导致的梯度消失的问题。Csp_2 结构中,使用普通的

CBL 模块代替了残差组件。Csp 结构用两条支路,通过支路实现特征融合,提取更为丰富的特征信息。

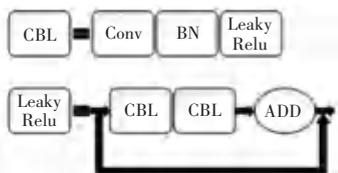


图3 CBL 模块及残差组件的组成

Fig. 3 Composition of CBL module and residual components

SPP (Spatial Pyramid Pooling) 是为了解决卷积神经网络输入图像数据大小固定的问题而提出的一种结构,卷积神经网络是由卷积层和全连接层构成,卷积层对输入图像数据大小没有规定,但第一个全连接层对输入图像的数据大小是有要求的。因此卷积神经网络基本上对输入数据的大小有所规定,针对许多数据高宽比不固定的情况,如果直接对图片进行切割,那么会存在丢失特征信息的可能性。因此,为了解决上述问题,提出了 SPP 结构,SPP 结构工作原理如图 4 所示。

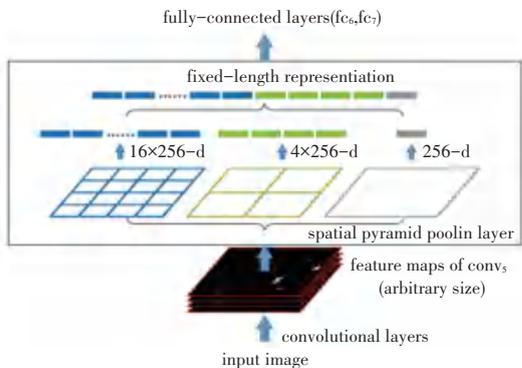


图4 SPP 结构工作原理

Fig. 4 SPP structure working principle

YOLOv5 中引进了 SPP 结构,用来解决图像由于裁剪、缩放等操作而引起的特征丢失等问题。

1.3 颈部网络

YOLOv5 网络的颈部网络采用的是 FPN (Feature Pyramid Networks)+PAN (Path Aggregation Network) 结构。FPN 是自顶向下的,将高层特征通过上采样和低层特征做融合得到进行预测的特征图。FPN 是自顶向下的,将高层的强语义特征传递下来,对整个金字塔进行增强,不过只增强了语义信息,对定位信息没有传递,而在 FPN 的后面添加一个自底向上的金字塔,是对 FPN 的补充,将低层的强定位特征传递上去。

FPN 就是把深层的语义特征传到浅层,从而增强多个尺度上的语义表达。而 PAN 则是把浅层的

定位信息传导到深层,增强多个尺度上的定位能力。FPN+PAN 结构如图 5 所示。

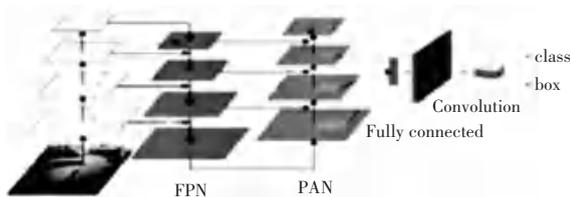


图5 FPN+PAN 结构示意图

Fig. 5 Schematic diagram of FPN+PAN structure

1.4 输出端

YOLOv5 的输出端中包含了预测框、损失函数、非极大值抑制 3 部分。其中,YOLOv5 的损失函数包括 GIoU (Generalized Intersection over Union) 损失函数、目标置信度损失函数、分类损失函数。GIoU 计算公式如式(1):

$$GIoU = IoU - \left| \frac{A_c - U}{A_c} \right| \quad (1)$$

其中, A_c 为预测框与真实框的最小外接矩形, U 为 A_c 中不属于预测框与真实框的部分。当真实框与预测框越接近时, U 越小,从而使得 $GIoU$ 越小。

YOLOv5 网络处于预测阶段时,会产生许多围绕同一目标的预测框,通过使用非极大值抑制使得预测框合并。将 GIoU 损失函数与非极大值抑制结合可以更好的识别图像中有重叠的目标。

2 数据采集及数据集制作

2.1 数据来源

本文的火焰图片数据来自公开数据集以及手动拍摄两部分,手动拍摄的数据为丁烷点火器生成的火焰视频,将视频截取成火焰图片。此次数据共采集 600 张火焰图片,通过数据增强的方法将原有数据进行旋转及亮度变换等操作,扩容至 3 000 张图片,火焰数据图片如图 6 所示。



图6 部分火焰图片数据

Fig. 6 Part of the flame picture data

2.2 数据集建立

将得到的火焰图片数据使用 Labelimg 软件进行标注。在标注过程中,尽可能的选取完整的火焰区域,防丢失图像特征。本次实验对象为火焰一个类,通过手动标注后自动生成 txt 文件,标签文件中包含了标签类别、预测框坐标等信息。标注结果如图 7 所示。



图 7 火焰数据标注图像

Fig. 7 Flame data labeled image

火焰样本数据标注完成后,通过脚本文件自动划分成训练集和测试集。以便后续输入模型训练。

3 实验结果及分析

3.1 实验环境

此次实验采用 Windows10 专业版操作系统,Pytorch1.7 版本框架,集成开发环境为 Pycharm。平台硬件参数如下:CPU 为 i9,内存大小为 32 GByte, GPU 为 Nvidia QUADRO RTX4000。

3.2 实验结果

本文实验采用 YOLOv5 目标检测模型,总体训练轮数设置为 200 轮,分组大小设置为 20,学习率设置为 0.001。网络训练过程中,训练前期模型的损失函数迭代收敛速度较快,随着训练次数的增多,损失函数收敛速度逐渐趋于缓和,最终实现收敛,并结束模型的训练。训练完成后将保存一个最新训练的权值文件与一个最优权值文件。网络训练结果如图 8 所示,可以看出, mAP 值为 0.96。

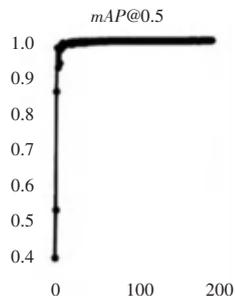


图 8 模型训练结果

Fig. 8 Model training results

平均精度与召回率是目标检测中常用的评价指标,其计算公式如式(2)~(3)。

$$P = \frac{TP}{TP + FP} \quad (2)$$

$$R = \frac{TP}{TP + FN} \quad (3)$$

其中, TP 表示将正类别预测为正确类别的个数; FP 表示将负类别预测为正确类别的个数; FN 表示将正类别预测为负类别的个数。

将训练好的权重文件导入模型之后,输入火焰图片进行识别,识别所耗时间为 0.245 s,识别效果如图 9 所示。



图 9 火焰识别效果图

Fig. 9 Flame recognition effect diagram

从图 9 可以看出,针对不同情况下的火焰,训练后的模型对其识别后置信度为 91%~93%,因此,经过训练的 YOLOv5 模型对火焰的识别具有较高的准确性,同时训练完成的检测模型兼具良好的实时性。

4 结束语

针对火焰检测存在实时性不足的问题,本文提出了基于深度学习的火焰检测方法,使用 YOLOv5s 目标检测算法模型,通过 3 000 张火焰图片数据的训练后,使用训练结果最优的权重参数文件。实验后期采用火焰图片数据对模型进行检测,具有较高的准确性和良好的实时性,可以满足实时性要求较高的火焰检测场景。

参考文献

- [1] 孙恒,孙守军,耿金亮,等. 新型含氧煤层气直接液化分离工艺模拟和优化[J]. 低温与超导,2021,49(10):40-45.
- [2] 王雅丽,李治刚,郭红光. 超临界 CO₂ 压裂煤岩储层增产煤层气应用发展前景[J]. 中国矿业,2021,30(10):160-167.
- [3] 张汝峰,高峰,项璟,等. 基于多空间颜色特征的火灾检测[J]. 电子测试,2021(11):47-48.
- [4] 苏展,张防震,王瑞,等. 基于 RGB+HSI 颜色模型的早期火焰检测算法研究[J]. 河南教育学院学报(自然科学版),2021,30(3):28-32.
- [5] 何爱龙,陈美娟. 基于多特征融合的视频火灾检测方法[J]. 软件导刊,2020,19(7):198-203.
- [6] 侯易呈,王慧琴,王可. 改进的多尺度火焰检测方法[J]. 液晶与显示,2021,36(5):751-759.
- [7] 马琳琳,马建新,韩佳芳,等. 基于 YOLOv5s 目标检测算法的研究[J]. 电脑知识与技术,2021,17(23):100-103.