

# 融合运动特征的高效视频火焰检测算法

孙维亚<sup>1</sup>, 陈恺鑫<sup>2</sup>, 吴 铭<sup>2</sup>, 王 丹<sup>3</sup>, 杜立轩<sup>2</sup>, 马占宇<sup>2</sup>

(1. 南水北调中线建管局, 北京 100053; 2. 北京邮电大学人工智能学院, 北京 100876; 3. 南水北调中线信息科技有限公司, 北京 100053)

**摘 要:** 提出一种轻量高效的视频火焰检测算法。该算法以基于深度学习的卷积神经网络目标检测算法为主体, 提取监控视频中的图像帧, 识别并定位火焰区域。加入运动目标检测模块作为后处理机制, 依据连续视频帧中火焰的运动特性, 采用基于混合高斯模型的运动目标检测算法对火焰目标检测结果进行合理化判断, 减少类似火焰的静止物体或光线造成的误报, 效率高且资源消耗少。此外, 收集并标注了一套火焰检测数据集 (Fire detection dataset, FDD), 包含多种场景下多类型燃烧物产生的火焰图片 2 487 张以及 15 段不同场景下的火灾视频数据。在 FDD 的视频检测实验中本文算法准确率达到 98.94%, 证明了本文算法的有效性。

**关键词:** 火焰检测; 目标检测; 运动目标检测; 视频监控; 深度学习

**中图分类号:** TP391      **文献标志码:** A

## An Efficient Video Flame Detection Algorithm Integrating Motion Features

SUN Weiya<sup>1</sup>, CHEN Kaixin<sup>2</sup>, WU Ming<sup>2</sup>, WANG Dan<sup>3</sup>, DU Lixuan<sup>2</sup>, MA Zhanyu<sup>2</sup>

(1. Construction and Administration Bureau of South-to-North Water Diversion Middle Route Project, Beijing 100053, China; 2. School of Artificial Intelligence, Beijing University of Posts and Telecommunications, Beijing 100876, China; 3. South-to-North Water Diversion Middle Route Information Technology Co., Ltd., Beijing 100053, China)

**Abstract:** This paper proposes a lightweight and efficient flame detection algorithm of videos. The flame detection algorithm is based on the convolutional neural network of deep learning. Considering the motion characteristics of the flame in the continuous video frame, this paper evaluates the flame detection results by the motion object detection and removes the false positive results caused by stationary objects or lights. The motion object detection algorithm based on the Gaussian mixture model is highly efficient. In addition, we collect and label a set of fire detection dataset (FDD), including 2 487 flame images and 15 fire videos under various scenarios with different flammable materials. In conclusion, the proposed algorithm obtains 98.94% accuracy on FDD test videos.

**Key words:** flame detection; object detection; motion object detection; video monitor; deep learning

## 引 言

现实场景中的火焰检测在科学研究领域和实际应用场景中都具有重大意义。本文主要目的是设

计出高效可应用的检测算法,因此本文以南水北调中线干线工程智能视频监控数据为研究背景,聚焦到实际场景。南水北调中线干线工程是一项跨流域、长距离的特大型、综合性调水工程,各站配备了包括启闭机、柴油发电机、网络交换机、大型计算机房及高压电源等设备设施。对于这些设备而言,火灾是最常见的严重危害之一,它会造成不可逆转的严重危害,是中线工程中重点防范的灾害。

现有的火灾报警大多依赖于烟雾报警器和红外传感器,这些物理传感器的覆盖范围十分有限,且易受到外界干扰,诸如大型室内仓库或室外场景均无法有效报警,对应用场景要求较高<sup>[1]</sup>。采用基于紫外线或红外线的光谱传感器会受到大型电子设备和其他光传感器的影响,可靠性不高,不能满足安全管控需求<sup>[2]</sup>。基于视频监控的火焰检测系统可以对摄像头覆盖的区域进行实时火焰检测,覆盖范围更广,检测效率更快,而且基本不会受到温度等自然因素的影响,在实际安全监控中具有重要意义。

基于视频的火焰检测可以分为两类。第一类是传统的机器学习方法,利用人为设定的特征对火焰进行检测<sup>[3-7]</sup>。这些传统检测方法统计火焰的色彩、运动情况以及边缘及纹理等特征,并针对这些特征设计检测方式。但是大多数人为特征没有平移、旋转和缩放不变性,当背景发生变化时,算法鲁棒性不足,因此效果都不甚理想。第二类是基于深度学习的方法。随着深度学习技术的出现及普及,一些研究者开始选择用深度神经网络提取有用的特征来进行火焰检测,将火灾检测视为目标检测任务,利用成熟的神经网络进行检测,算法准确率更高,鲁棒性更好<sup>[8-11]</sup>。这些方法分别通过反向传播神经网络、前馈神经网络和卷积神经网络解决了基于人为特征火焰检测算法准确率低和鲁棒性差的问题。

本文将上述两类方法结合,提出了一种结合动静特征的室内火焰检测系统。该系统的核心模块共分为2部分:(1)针对火焰的静态图像特征,在图像中定位火焰,为了到达较好的检测效果,该部分利用神经网络;(2)运动检测模块,加入火焰运动特征,消除虚警。由于本文研究拥有视频资源,可获取连续相邻的视频帧,结合火焰闪烁摇晃的运动特性,对所有的疑似火焰报警进行筛选,进一步判断是否为火焰,这样可以极大地避免与火焰相似的静态物体所触发的误报,提高算法的准确率。

火焰检测安全监控系统在实际应用中往往需要能够满足多路运行的条件,因此除了准确率,硬件资源也是重要决定因素。神经网络模型在图形处理器(Graphics processing unit, GPU)下运行可以发挥更好的效果,但普通GPU缓存有限,无法支持多路运行。相比之下,中央处理器(Central processing unit, CPU)具有多核独立运行的特点,可以满足系统多路运行的要求,也能够实现对现有CPU服务器的利旧。本文算法针对CPU平台进行一定优化,使得该系统可以在同时多路运行的情况下保证较好的检测准确率和实时性,具有较好的实际应用效果。除去检测部分,本文还完成了获取实时视频帧的程序,与检测模块结合完成整套火焰检测系统。此外,为了进行有效的实验,本文创建了一份标注完善的火焰数据集。

## 1 本文算法

本文提出的基于视频的实时火焰检测系统主要分为2个主要部分:第1部分负责获取摄像机中的视频帧;第2部分负责进行火焰检测。检测算法首先利用卷积神经网络做目标检测,识别并定位火焰,再通过运动检测对候选区域进行筛选,最终获得检测结果。本文还针对CPU进行了优化,使其更具实际意义。

### 1.1 系统整体结构与目标检测网络

为了实现实时火焰检测的功能,本文系统分为2条支路,整体结构如图1所示。图1中,视频读取模块与检测模块在2个线程中独立运行,这样设计的原因是如果顺序运行2个模块会出现检测延迟问题。视频提取模块为了保证视频连续性,内置缓冲空间,从与摄像头建立连接开始便自动缓存视频图像。如果应用在同一个线程中,检测模块会占用一定的时间,而这段时间缓冲区内一直在缓存图像,读取时

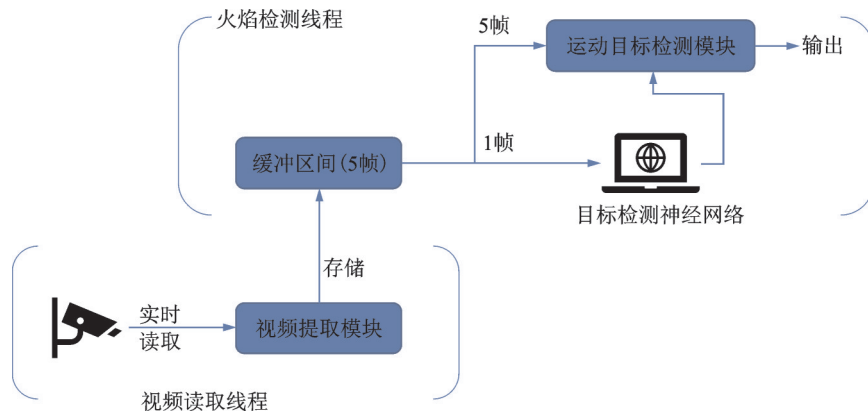


图1 系统的整体结构

Fig.1 Overall structure of the system

会从缓存区域内第1张开始读取,这样下一次检测读取到的图片依旧是第1次存入的图片,与当前实际时间就会有一定延迟,无法达到实时效果。本文采用双线程设计模式,创建独立的缓冲区域,获取视频帧的线程负责将视频提取模块缓存区内的图像读取出来,并覆盖式地存入独立的缓冲区内,使得本文创建的缓冲区内永远都是实时的图像。检测线程在独立缓冲区内读取图片,确保检测的实时性。

对于火焰检测这种安全检测,检测的效率与警报的准确性都十分重要,因此模型的实时性和准确率成为选择模型的重要指标。目前已经成熟的目标检测模型可按照运行流程分为单阶段网络和双阶段网络,双阶段目标检测网络,诸如RCNN<sup>[12]</sup>、Fast RCNN<sup>[13]</sup>和Faster RCNN<sup>[14]</sup>等,检测时间较长但准确率较高;单阶段网络,诸如Yolo<sup>[15]</sup>、Yolov3<sup>[16]</sup>、SDD<sup>[17]</sup>和RetinaNet<sup>[18]</sup>等,速度快但准确率不理想。由于本系统是应用在视频监控中,可以在一定时间内检测多次,准确率的要求可以适当降低,因此为了满足实时性要求,本文选择在单阶段网络中挑选模型作为基线模型。深度检测模型可表示为

$$[x, y, w, h] = \text{NMS}(\text{CNN}(I)) \quad (1)$$

式中: $x$ 、 $y$ 分别代表目标框左上角点纵横坐标; $w$ 、 $h$ 表示目标框宽度与目标框高度;NMS为非极大值抑制,过滤同一目标多个检测框,保留得分最高者;CNN为卷积神经网络; $I$ 为输入图片。

本文对现有较为成熟的网络进行了对比实验。根据实验结果,本文选择Yolov3作为目标检测基线模型。这是由于Yolov3整体网络结构采用了类似特征金字塔网络<sup>[19]</sup>的上采样和多尺度特征融合,具有更丰富的语义信息,同时图像细节丢失的深层特征与包含更多局部细节特征的浅层特征图进行融合,获得既有语义信息又具有局部信息特征图,可以实现对不同尺度目标检测的效果。特殊的网络结构使其可以对小目标物体也具有较好的识别能力,可以在火苗较小时及时发现并及时报警。虽然准确率相比于双阶段网络较差,但由于本文系统应用于视频检测中,单阶段网络所具有的速度优势可以帮助其在相同时间内检测多帧视频图像,从而弥补准确率的不足。

## 1.2 运动目标检测

本文火焰检测算法基于实时视频进行检测,由于火焰形状不确定,在实际应用场景中,室内的阳光、红色的灯等都对算法有一定的欺骗性。此外,由于无法达到百分百的准确率,经过一天不间断的检测也会出现一部分误报,如图2所示。为了尽量消除这些误报,本文尝试在算法中增加一些后处理操作。由于火焰是运动的,而本文研究又拥有视频资源,可以对相邻几帧进行操作,因此选择利用运动检测,对目标区域进行检测判断是否发生运动,可以过滤掉一些静态的误报。

运动目标检测涉及多步图像处理操作,具体流程如下:首先去除图像背景,将1组相邻视频图像作为输入;然后采用基于高斯混合模型<sup>[20]</sup>的运动目标检测算法提取运动区域,剔除背景并最终输出最后1帧图像中的运动区域,输出格式为灰度图。高斯混合模型是由K个单高斯模型组合而成,原始单高斯模型表达式为

$$P(x|\theta) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \quad (2)$$

式中: $\mu$ 为数据期望; $\sigma$ 为数据标准差。在火焰检测任务中图像数据为多维数据,因此高斯模型转变为

$$P(x|\theta) = \frac{1}{(2\pi)^{\frac{D}{2}} |\Sigma|^{\frac{1}{2}}} \exp\left(-\frac{(x-\mu)^T \Sigma^{-1} (x-\mu)}{2}\right) \quad (3)$$

式中: $\Sigma$ 为数据协方差; $D$ 为数据维度。本文可以进一步得到高斯混合模型的数学表达式为

$$P(x|\theta) = \sum_{k=1}^K \alpha_k \Phi(x|\theta_k) \quad (4)$$

式中: $\alpha_k$ 为数据属于第k个子模型的概率,满足 $\alpha_k \geq 0, \sum_{k=1}^K \alpha_k = 1$ ;  $\Phi(x|\theta_k)$ 为第k个子模型的单高斯分布密度函数, $\theta_k = (\mu_k, \sigma_k^2)$ ,展开形式与上文介绍的单高斯模型相同。混合模型中包含的参数为每个子模型的期望、方差和混合模型中发生的概率。从理论上来说,一个混合模型可以选择任何模式的子概率模型,但高斯分布具备良好的数学特性,因此本文采用高斯混合模型来处理数据分布。

由于本文要获取火焰的运动区域,而火焰的颜色基本都趋于明亮,因此将阈值设定为(50~100),这样过滤掉一些非火焰的运动区域。然后对图像进行腐蚀和膨胀,腐蚀是为了让图像边缘更加平滑,剔除微小的噪声,膨胀是为了使区域更加连贯,对处理后的结果查找图像轮廓,获取一系列候选运动区域。最后对图像轮廓进行尺寸筛选,由于火焰尺寸不会太小,因此剔除像素尺寸较小的候选框,最后得到的结果便是视频中的运动区域。运动检测模块处理流程可视化效果如图3所示。

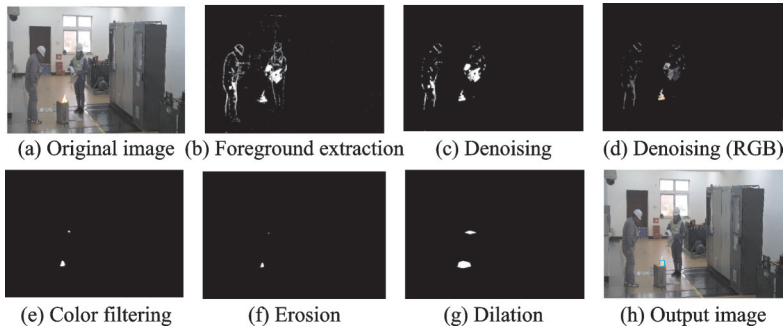


图3 运动目标检测流程图  
Fig.3 Flow chart of motion detection



图2 错误检测结果样例  
Fig.2 Some false detection results

运动检测方式是将运动区域与检测结果做比较,如果检测的火焰位置与运动区域重合,则认为该部分为火焰,否则认为是静态的误检。判断重合即计算火焰的边界框(Bounding box, B-box)与运动区域的交并比(Intersection-over-union, IOU),表达式为

$$\text{IOU} = \frac{S_i}{S_1 + S_2 - S_i} \quad (5)$$

式中: $S_i$ 表示交集面积; $S_1$ 和 $S_2$ 表示2个被计算框面积。由于运动检测主要作用是过滤误检,为了防止过滤掉正确的检测结果,不要求2个候选框处于包含状态,而是当交并比达到一定阈值时便认为二者包含区域相同,属于重叠状态,即可以认为是火焰。

### 1.3 CPU的适应性

本文系统运行效果主要依靠CPU,因此模型的选择和优化都需满足CPU的性能。由于采用Pytorch加载模型的方式在CPU上会占用大量的硬件资源,导致运行速率受到严重影响,而利用OpenCV中的DNN模块加载模型可以避免这一问题,所以本文将目标检测网络按照OpenCV的加载方式进行了调整,更改了参数模型以及配置文件,使得本文系统可以在CPU上顺利运行。

## 2 实验和结果

### 2.1 评价指标

在常规的目标检测对比实验中,通常选择准确率与误报率的综合评价结果平均准确率(Average precision, AP)作为评价指标,但对于安全检测系统,准确率、召回率和误报率都十分重要。因此,本文实验采用主要评价指标包括准确率、召回率和误报率。

#### (1)准确率

准确率 $P$ 表示被判别为正确目标数量在所有正确目标中所占的比例。通常来说,准确率越高,模型的检测效果越好,计算公式为

$$P = \text{TP}/(\text{TP} + \text{FP}) \quad (6)$$

式中:TP为有火焰且判断为火焰的数量;FP为无火焰但判断为有火焰的误报数量。

#### (2)召回率

召回率 $R$ 是算法对正例敏感性的评价指标,度量有多少个正例被成功检测。召回率可以帮助度量模型发生漏报的情况,计算公式为

$$R = \text{TP}/(\text{TP} + \text{FN}) \quad (7)$$

式中:FN表示未被检测到的正确示例,即有火焰的漏报。

#### (3)误报率

误报率ER是描述检测错误的比例,为被判断错误的目标数与所有样本数的比值,计算公式为

$$\text{ER} = (\text{FP} + \text{FN})/(\text{TP} + \text{FN} + \text{FP} + \text{TN}) \quad (8)$$

式中:TN表示将负类预测为负类数。

### 2.2 实验数据

为了完成实验,本文创建了一份包含图片与视频的火焰数据集FDD,其中数据来源分为2种。第1种来源于网络视频,本文在网上选取了一些不同场景的室内着火视频,截取出实验所需的图片构成数据集图像部分,原视频也保存为数据集中的视频部分。由于同一视频的背景一般保持不变,因此只是挑选火焰形态有明显变化的图像作为数据。第2种是在实际场景中多种不同房间内点火,按照实验要求创造一些新的室内着火数据,同样分别按照视频与图像进行存储。由于图像来源不一,共有3种尺寸,分别为960像素×540像素、1280像素×720像素和1920像素×1080像素。经过筛选和标注后,数据集共包含2487张图像,其中2048张正样本、439张负样本。数据集中的正样本包括房屋着火、家用燃气、电器着火和木炭燃烧等多种情况,能够使模型更好地学习到火焰的本质特征,而不会局限于某

一种燃料燃烧时的火焰。负样本也均为类似火焰的灯光或阳光投影,可以有效降低该类物体的干扰,降低模型的误报率,正负样本比例约为8:2。数据集按照8:2的比例分为训练集和验证集,即1 990张训练集、497张验证集。数据集部分图片样本如图4所示。

本文实验包括2类:第1类是通过图像数据对目标检测进行实验;第2类是通过视频数据评价运动检测算法的性能。网络模型测试实验选择的测试集采用火焰数据集FDD,共包括249张图像,其中205张包含火焰的正样本,44张不包含明火的负样本,图像尺寸统一为1 280像素 $\times$ 720像素。对于第2类实验,本文从15条火焰视频中选择了3条较为适合的测试视频,视频时长均为5 min,视频帧率为25帧/s,每条视频共计7 500张视频图像,视频中均包括有明火的正样例和无火焰的负样例。



图4 数据集FDD中部分图像样本

Fig.4 Some sample images of FDD

### 2.3 实验细节

本文实验将提出算法在GPU平台上训练,并最终通过优化部署到CPU平台。为了提高实验速度,对网络模型的效果测试实验和运动检测模块的对比实验都在GeForce RTX 2080Ti上进行,显卡缓存为16 GB。最终运行在英特尔至强E5-2630V4处理器上进行,该处理器包含20个CPU内核,并且实现了英特尔超线程技术,可看作40线程处理器。

为了丰富训练数据的多样性、提升训练效果,采用多种数据增强方式。针对火焰特性,本文选择更改颜色空间、随机左右翻转、上下翻转和仿射变换。颜色变换主要是在HSV(Hue, saturation, value)空间上对S和V元素进行调整,即对火焰的亮度做数据扩增。由于火焰形状的不固定性,四面翻转都可以起到较好的效果。仿射变换包括旋转、平移、错切和尺度变化。仿射变换特点为直线经过仿射变换仍然是直线。这一变化在目标检测中经常被使用,它的多种变换可以更改火焰的形态、大小和位置,有助于算法学习到更多种类的火焰。同时由于在基础网络部分常常会生成比原图小数10倍的特征图,采用多尺度训练方式会导致小物体的特征描述不容易被检测网络捕捉。通过输入更大、更多尺寸的图像进行训练,能够在一定程度上提高检测模型对物体大小的鲁棒性。在训练前首先会将输入图像尺寸统一为416像素 $\times$ 416像素,在训练时会按照设定好的多尺度范围改变图像大小。实际设定的范围是288~608像素,每间隔32像素设立为一个尺度标准。

运动目标检测模块包含众多参数,下文会着重介绍对检测效果影响较大的参数。腐蚀和膨胀的迭代次数分别为2和5,结构元素的大小分别为(3,3)和(8,3)。筛选图像轮廓的尺寸大小为400,判断运动区域与火焰候选框重合的比例为0.7。

## 2.4 目标检测网络对比实验

本文首先对不同目标检测网络模型的检测效果进行对比。选择目前较为成熟的几个单阶段目标检测网络作为实验对象,实验数据为2.2节中介绍的图像数据集,统计结果如表1所示,检测效果如图5所示。表1中运行时间表示单张图像所需的检测时间。从表1可以看到Yolov3网络从检测效果到检测速度都比另外2个模型更加优秀,共报警190张,其中正确186张、误报4张和漏报15张,同时误报基本都为红色的类似火焰灯光,运动检测可以有效避免这些误报。

表1 不同网络检测效果

Table 1 Performance comparison of different detection networks

方法	准确率/%	召回率/%	误报率/%	运行时长/ms
SSD <sup>[17]</sup>	91.16	80.49	16.06	61
RetinaNet50 <sup>[18]</sup>	96.17	85.85	11.65	73
Yolov3 <sup>[16]</sup>	97.89	90.73	7.63	25

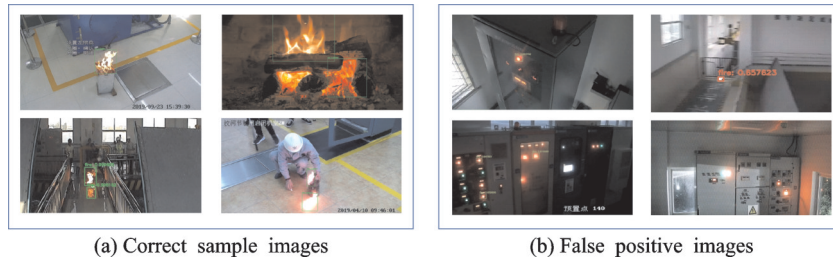


图5 目标检测网络检测正确样本图与误报样本图

Fig.5 Correct sample images and false positive images

## 2.5 运动目标检测模块对比实验

### 2.5.1 颜色过滤区间对比实验

本文还对运动目标检测进行了不同参数的对比实验,以达到最好的检测效果。实验中选择的第1个参数是二值化阈值,这一参数主要是根据火焰的颜色特性从运动区域中剔除非火焰部分,实际效果如图6所示。

不同颜色区间对火焰的敏感效果不同,针对火焰黄红颜色的特性将二值化阈值从低到高逐步进行实验。第1次实验数据为视频数据,选择视频的前50帧,对比不同颜色区间对火焰的运动检测效果,结果如表2所示。

本文研究主要作用是选出所有疑似火焰的运动区域,和神经网络检测结果作对比,从而剔除误报,



图6 二值化处理剔除非火焰效果图

Fig.6 Effect picture of eliminating non flame by binarization

表2 第1次实验不同颜色过滤区域检测结果

Table 2 Detection results of different color filter intervals in the first experiment

颜色区域	50~75	75~100	100~150	150~200	200~255
火焰数目	37	33	21	16	8
误检数目	2	3	6	12	14
漏检数目	4	8	20	25	33

因此应尽量保证每一张火焰都被检测到,这样才不会影响正常的报警。因此运动检测可以具有一定的误检,但应尽量保证所有的火焰必须都在检测框内。因而第2次实验放宽了检测区间以保证检测效果,结果如表3所示。

表3 第2次实验不同颜色过滤区域检测结果

Table 3 Detection results of different color filter intervals in the second experiment

颜色区域	50~100	75~150	100~200	150~255
火焰数目	40	33	22	16
误检数目	3	6	13	14
漏检数目	1	8	19	25

根据实验结果,本文最终选择50~100作为火焰的二值化颜色区间。在该情况下视频中的火焰基本可以完全检测到,而且不会有过多的其他运动区域影响。

### 2.5.2 前景提取时视频帧数对比实验

第2个参数是获取前景图像的视频帧数。在做背景提取时,需要输入一组相邻图像,图像的数目会影响提取效果与提取时间,本文对不同的图像数量进行了测试,测试条件与二值化阈值实验相同。通过实验发现:输入图像数量小于5张时,程序无法准确找到背景,提取出来的运动区域不准确;而大于5张时可以找到清晰的运动区域,且无明显区别。由于图像越多处理时间越长,因此最终选择5帧作为输入数量,有利于进一步挑选运动火焰区域。

## 2.6 加运动目标检测模块前后对比实验

本文研究在目标检测模型上加入了运动目标检测模块以提高模型的准确率,降低误报率。通过对比实验证明这一模块的实际效果。本实验采用2.2节中介绍的视频数据作为测试数据,对比对象分别为目标检测网络与加了运动检测模块后的目标检测网络,统计二者的最终检测结果。实验时将视频按帧统计,包含明火的记为正样本,不含明火的记为负样本。模型成功检测并准确回归火焰边界框记为TP,回归边界框与真实边界框IOU小于0.5时记为FP,含有明火但并未被检测到记为FN,不含明火也未被检测到记为TN。考虑到检测任务存在有火焰被检测到但回归区域错误的情况,算法召回率计算公式调整为

$$R = TP/T \quad (9)$$

式中: $T$ 为所有正样本数目。本文对3条视频进行逐帧测试,运动检测模块以重叠的方式对当前帧的相邻帧进行检测,实验设备选择的是Nvidia 2080Ti GPU和英特尔至强E5-2630V4 CPU。依照式(6,8,9)计算统计结果,并记录每帧视频平均检测时常,结果如表4所示。

该对比实验可以清晰表明运动检测模块的有效性。从表4中可以看出算法的准确率有明显提升,误报率也有明显下降,这是因为运动检测模块借助运动特性过滤背景中静态目标,减少了误报数目。

表4 运动目标检测模块对比实验

Table 4 Comparison experiments of motion object detection module

检测模型	平均准确率/%	平均召回率/%	平均误报率/%	GPU运行时长/ms	CPU运行时长/s
CNN	94.28	96.51	4.67	26	2.16
CNN + 运动检测	98.94	96.51	2.26	30	2.16



同时算法的召回率没有发生变化,这表明该模块的过滤策略并不会对正确的报警造成影响,正确的火焰结果不会被过滤。实验结果说明,运动检测模块可以在不影响正确报警的前提下有效剔除部分静态误报,增强算法效果。

此外,本文还分析了算法的误报样本。未加运动检测的模型检测到的误报多为红色灯,而这些都可以有效地被运动检测模块剔除。加运动检测后的误报主要是由于摄像头发生移动,导致画面模糊,形状类似火焰,误报场景如图7所示。



图7 误报场景图片示例

Fig.7 One of false positives

### 3 结束语

针对南水北调中线室内火焰检测需求,本文提出了一种基于视频监控的实时室内火焰检测算法。通过分析火焰的运动特性,该算法将传统视频检测算法与深度学习算法相结合,在传统目标检测的基础上利用运动检测有效消除了静态背景产生的误报,提升了火焰目标的准确率。为了保证视频检测的实时性,本算法采用双线程处理模式,并通过建立缓存区域,实现了视频读取和视频处理的独立运行。同时,本文基于CPU平台进行了算法和流程优化,通过调整模型加载和运行方式,满足了实际应用场景中的多路并发需求,充分利用了硬件资源。本文提出的算法已成功应用于中线视频智能分析系统中,经过多次现场模拟测试证实,本文算法能够快速检测室内火焰,检测效果良好。将基于视频监控的检测手段与传统消防检测手段相结合,可以有效缩短事件响应时间,提高应急处理能力,对中线运维管理水平的提升具有十分重要的现实意义。如何解决误报样本将是下一步研究方向。

### 参考文献:

- [1] 高爱莲,刘辉,林宏,等. 基于视频的火焰检测技术研究[J]. 云南民族大学学报(自然科学版), 2008, 17(1): 94-96.  
GAO Ailian, LIU Hui, LIN Hong, et al. Flame-detection technique based on video processing[J]. Journal of Yunnan Nationalities University, 2008, 17(1): 94-96.
- [2] 王华伟. 基于计算机视觉的视频火焰检测技术[J]. 消防界(电子版), 2017, 1: 76.
- [3] ZHANG Z, SHEN T, ZOU J. An improved probabilistic approach for fire detection in videos[J]. Fire Technology, 2014, 50(3): 745-752.
- [4] HAN X F, JIN J S, WANG M J, et al. Video fire detection based on Gaussian mixture model and multi-color features[J]. Signal Image & Video Processing, 2017, 11(5): 1-7.
- [5] RONG J, ZHOU D, YAO W, et al. Fire flame detection based on GICA and target tracking[J]. Optics & Laser Technology, 2013, 47(4): 283-291.
- [6] WANG D C, CUI X, PARK E J, et al. Adaptive flame detection using randomness testing and robust features[J]. Fire Safe, 2013, 55(328): 116-125.
- [7] LIU Z G, YANG Y, JI X H. Flame detection algorithm based on a saliency detection technique and the uniform local binary pattern in the YCbCr color space[J]. Signal, Image and Video Processing, 2016, 10(2): 277-284.
- [8] GUO W Q, YAN F, HAN N. Wildfire smoke detection using dynamic features and neural network[J]. International Journal of Digital Content Technology and Its Applications, 2012, 6(11): 61-68.
- [9] GUO Q, DAI J, WANG J. Study on fire detection model based on fuzzy neural network[C]//Proceedings of International Workshop on Intelligent Systems and Applications. Wuhan, China: IEEE, 2010: 1-4.
- [10] FRIZZI S, KAABI R, BOUCHOUICHA M, et al. Convolutional neural network for video fire and smoke detection[C]//Proceedings of Conference of the IEEE Industrial Electronics Society. Florence, Italy: IEEE, 2016: 877-882.
- [11] ZHANG Q, XU J, 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. France: Atlantis Press, 2016:

568-575.

- [12] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]//Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition. [S.l.]: IEEE, 2014.
- [13] GIRSHICK R. Fast R-CNN[C]//Proceedings of the IEEE International Conference on Computer Vision. [S.l.]: IEEE, 2015.
- [14] REN S, HE K, GIRSHICK R, et al. Faster R-CNN: Towards real-time object detection with region proposal networks[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2017, 39(6): 1137-1149.
- [15] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: Unified, real-time object detection[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, NV, USA: IEEE, 2016: 779-788.
- [16] REDMON J, FARHADI A. Yolov3: An incremental improvement[EB/OL].(2018-04-08)[2020-12-20]. <https://arxiv.org/abs/1804.02767>.
- [17] LIU W, ANGUELOV D, ERHAN D, et al. SSD: Single shot multibox detector[C]//Proceedings of European Conference on Computer Vision. Cham: Springer, 2016: 21-37.
- [18] LIN T Y, GOYAL P, GIRSHICK R, et al. Focal loss for dense object detection[C]//Proceedings of the IEEE International Conference on Computer Vision. Venice, Italy: IEEE, 2017: 2980-2988.
- [19] LIN T Y, DOLLÁR P, GIRSHICK R, et al. Feature pyramid networks for object detection[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, HI, USA: IEEE, 2017: 2117-2125.
- [20] 马义德, 朱望飞, 安世霞, 等. 改进的基于高斯混合模型的运动目标检测方法[J]. 计算机应用, 2007(10): 2544-2546.  
MA Yide, ZHU Wangfei, AN Shixia, et al. Improved moving objects detection method based on Gaussian mixture model[J]. Computer Applications, 2007(10): 2544-2546.

## 作者简介:



孙维亚(1980-),男,高级工程师,研究方向:水利信息化,E-mail:sunweiya@nsbd.cn。



陈恺鑫(1997-),男,硕士研究生,研究方向:计算机视觉、模式识别。



吴铭(1977-),通信作者,女,博士,副教授,研究方向:计算机视觉、模式识别,E-mail:wuming@bupt.edu.cn。



王丹(1989-),女,工程师,研究方向:视频智能分析。



杜立轩(1999-),男,本科,研究方向:计算机视觉、模式识别。



马占宇(1982-),男,博士,教授,博士生导师,研究方向:计算机视觉、模式识别与机器学习。

(编辑:张黄群)