



改进 Yolo-v3 的视频图像火焰实时检测算法

赵媛媛¹ 朱 军¹ 谢亚坤¹ 李维炼¹ 郭煜坤¹

¹ 西南交通大学地球科学与环境工程学院,四川 成都,611756

摘要:为解决现有视频图像火焰检测方法精度低、速度慢的问题,提出了改进 Yolo-v3 的视频火焰实时检测算法。首先,在特征提取阶段,通过进一步融合多尺度特征提高网络对图像浅层信息的学习能力,以实现小火焰区域的精准识别;其次,在目标检测阶段,利用改进的 K-means 聚类算法优化多尺度先验框以适应火焰不同尺寸;最后,在改进 Yolo-v3 的视频火焰检测之后,利用火焰特有的闪烁特征对检测结果中的误检帧进行排除,进一步提高检测精度。从精度和速度两个方面对视频火焰进行检测,并与近年来先进的火焰检测算法对比,结果表明,该方法准确率均值可达到 98.5%,误检率低至 2.3%,平均检测速率为 52 帧/s,在精度和速度方面皆有更好的表现。

关键词:火焰检测;视频图像;Yolo-v3;闪烁特征;多尺度检测

中图分类号:P237;TP391

文献标志码:A

火灾本身具有不确定性、复杂性、快速性等特征,一旦发生容易造成严重的经济损失和人员伤亡。准确的火焰预警可以极大地减少损失。随着智能监控处理能力的提升,基于视频的火焰自动检测方法不断涌出,其中,利用颜色进行火焰检测是最早的方法^[1-2]。但该方法需要大范围火焰区域。为了提高火焰识别的准确率,研究人员在颜色特征的基础上加入了火焰的运动特征^[3],提高了火焰识别的准确率,但误检率仍旧较高。

近年来,各种深度学习网络,如卷积神经网络^[4]、递归神经网络^[5]、深度信念网络^[6]等完成了如图像处理、分类和检索^[5-8]、目标检测^[9]、定位^[10]、分割^[11]、自然语言处理^[12]、语音识别^[13]等方面的应用。其中,卷积神经网络在图像分类方面取得了较好的效果,因此,许多卷积神经网络算法被应用于视频火焰检测中^[14]。在火焰检测研究中,文献[15]设计了一个9层卷积神经网络;文献[16]设计了一种两级结构卷积神经网络;文献[17]设计了一个12层卷积神经网络,虽然大幅度地提高了火焰检测精度,降低了误检率,但是检测速率仍然很慢,无法实现对视频中火焰的实时检测;文献[18]提出了 Yolo 算法,其检测速度可

达 45 帧/s,真正实现了实时检测;并在此基础上,又提出了 Yolo-v2 算法^[19];此后,Redmon 等^[20]在 Yolo-v2 算法的基础上,融合 ResNet 网络、二元交叉熵损失等的思想,提出了 Yolo-v3 算法,该算法在大幅度提高检测速率的同时也提高了检测精度。但该算法对小目标检测效果仍不乐观,不适合对小火焰区域进行识别。

为解决上述问题,本文提出了改进 Yolo-v3 的视频图像火焰识别算法,在保证视频图像火焰检测速率的同时,大幅度提高视频图像火焰识别的精度,特别是实现对小火焰区域的高效检测。此外,本文结合火焰闪烁特征,进一步排除视频火焰误检情况,提高检测精度。

1 改进算法

1.1 Yolo-v3 算法简介

Yolo-v3 通过引入了残差模块,降低深层网络训练复杂度,提高目标检测精度和速率。其中,Darknet-53(如图1所示)作为 Yolo-v3 的骨干网络,使用步长为 2 的卷积进行 32、16、8 倍下采样操作,并进行目标检测,拼接特征信息,融合 13×13 、 26×26 、 52×52 这 3 种不同尺度的特征,连接深、浅层特征,达到同时学习多尺度特征的

收稿日期:2020-08-05

项目资助:国家自然科学基金(41871289);四川省自然资源厅科研项目(KJ-2020-4);四川省青年科技创新研究团队项目(2020JDTD0003)。

第一作者:赵媛媛,硕士生,主要从事虚拟地理环境与灾害场景建模研究。3011441848@qq.com

通讯作者:朱军,博士,教授。vgezj@163.com

效果。尽管 Yolo-v3 目标检测能力较强,但对小目标识别能力仍然有限,而在火焰检测时,需要对小区域火焰进行实时精确的检测,防止火势蔓延。针对 Yolo-v3 在视频图像火焰检测上存在的问题,本文提出了改进 Yolo-v3 的视频图像火焰实时检测算法,并通过实验验证了本文方法在火焰检测中的可行性。

1.2 小火焰区域精准识别

Darknet-53 网络提取特征信息时,浅层特征

网格划分较小,主要提供位置信息;深层特征网格划分较大,主要提供语义信息。为了能够利用浅层位置信息,提高目标检测精度,Yolo-v3 在进行多尺度检测时使用了 3 个尺度(13×13 、 26×26 、 52×52)融合的方法。分辨率最小的特征尺度为 52×52 ,相对 416×416 像素的输入图像,网格划分还是不够精细,经过层层卷积计算后,部分信息会丢失,造成浅层特征信息浪费,导致其对小火焰区域的检测精度不高。

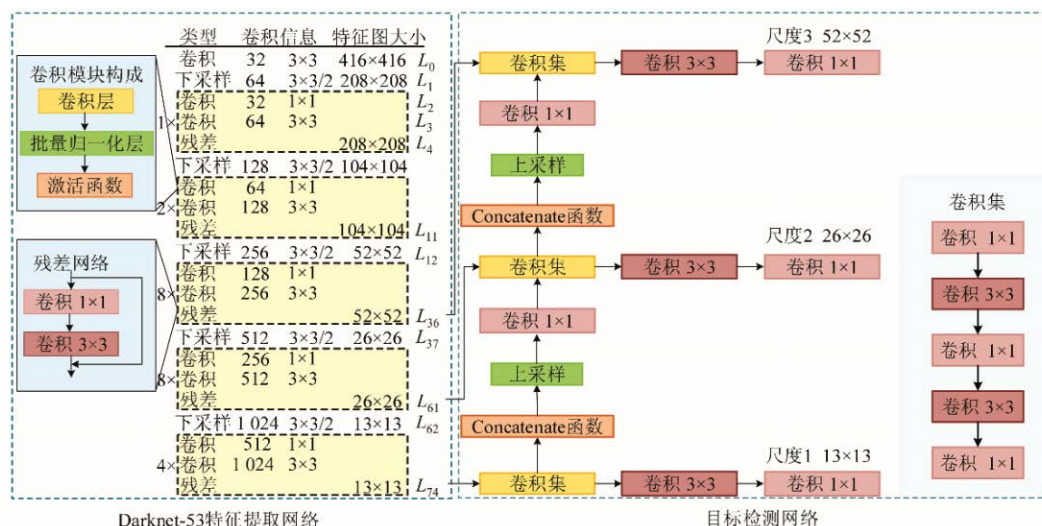


图1 Yolo-v3网络结构

Fig.1 Yolo-v3 Network Structure

为提高网络融合多尺度特征能力,准确识别小火焰区域,本文改进了 Yolo-v3 多尺度检测网络,添加分辨率更小的特征尺度以充分学习浅层特征。由于新增特征尺度会增加网络复杂程度,

拉低检测速率,综合考虑火焰检测精度和速率要求,增添 104×104 尺度,改进为 4 尺度检测。本文输入图像大小为 416×416 像素,改进的网络模型如图 2 所示。

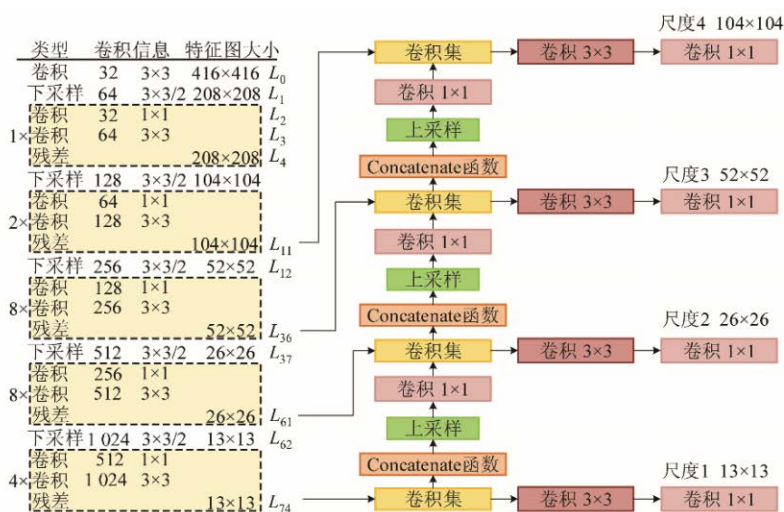


图2 改进后的多尺度检测网络

Fig.2 Improved Multi-scale Detection Network

首先,将 L_{74} (74 层) 作为第 1 尺度特征 13×13 ,进行卷积和上采样操作,与 L_{61} 融合得到 $26 \times$

26 的第 2 尺度特征;其次,将第 2 尺度特征与 L_{37} 结果融合作为 52×52 的第 3 尺度特征;最后,将

第3尺度特征进行上采样与 L_{11} 结果进行融合,得到 104×104 的第4尺度特征。

改进后的多尺度特征融合方法可以更多地获取图像浅层特征,在满足火焰检测速度的同时,进一步提高对小火焰区域的检测效果。

1.3 先验框设置优化

Yolo-v3算法将输入图像分成若干网格,如果检测目标实际边框的中心处于某个网格中,则这个网格负责预测这个目标,如图3所示。

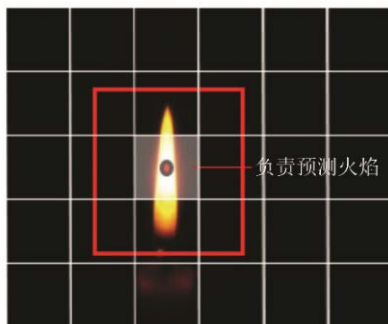


图3 目标预测示意图

Fig.3 Diagram of Target Prediction

视频中火焰区域大小具有不定性,为了能够实现高精度的火焰检测,需要对Yolo-v3的先验框尺寸进行优化,使之更适合火焰检测,本文利用K-means聚类算法分析统计本文数据集的实际边框,聚类个数 K 即为先验框的个数,聚类中心框的宽、高即为先验框宽、高与图像大小的比值。

由于先验框大小不一,采用K-means欧氏测距法进行距离计算,产生的误差与先验框的大小成正相关性,即先验框越大,产生的误差越大。为解决上述问题,同时,使先验框与检测目标实际边框的交并比更大,本文用改进的距离公式^[20]进行计算,利用 I 来计算先验框之间的距离,其值越大,先验框之间的距离越小。先验框之间的距离为:

$$d(b, c) = 1 - I(b, c) \quad (1)$$

式中, b 为任意一个先验框; c 为对应先验框的中心; I 为先验框之间重叠面积与总面积的比值。利用优化后的K-means算法,得出先验框个数 K 与 I 的关系,如图4所示。

图4中,横、纵坐标分别为先验框个数 K 和 I 值,随着 K 值的增加, I 值会随之增大,当 $K=12$ 时, I 值趋于平缓,综合考虑网络火焰检测的速率和精度,设置先验框个数为12,并随机产生12个初始聚类中心。按照式(1)行距离迭代计算,直到聚类中心距离变化小于阈值(本文中为0.005)停止计算,得出最终聚类中心坐标,如表1所示。

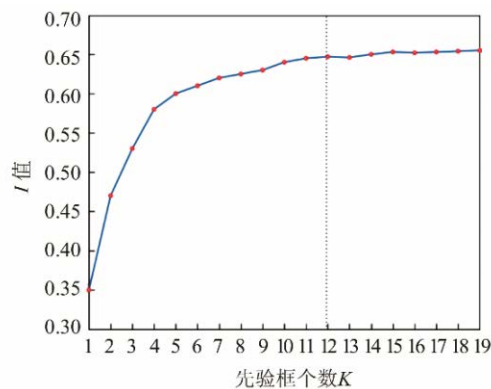


图4 先验框个数 K 与 I 关系图

Fig.4 Relationship Between K and I

通过迭代运算得到的聚类中心坐标是相对于图像大小的比例,要与输入图像的尺寸 416×416 像素相乘,得到先验框尺寸,并按尺寸大小进行排列,结果如表1所示。

表1 聚类中心坐标和先验框尺寸

Tab. 1 Coordinates of Cluster Center and Size of the Anchors

聚类中心坐标		先验框尺寸	
聚类中心	坐标	先验框	尺寸/像素
1	(0.15, 0.23)	1	17×25
2	(0.04, 0.06)	2	29×37
3	(0.07, 0.19)	3	29×79
4	(0.29, 0.71)	4	46×150
5	(0.11, 0.36)	5	58×54
6	(0.25, 0.35)	6	62×96
7	(0.17, 0.51)	7	71×212
8	(0.28, 0.23)	8	104×146
9	(0.43, 0.42)	9	117×96
10	(0.07, 0.09)	10	121×295
11	(0.14, 0.13)	11	179×175
12	(0.5, 0.75)	12	208×312

改进后的目标检测网络有4种尺寸特征图,本文沿用Yolo-v3算法,每种尺寸特征图匹配3个先验框。 13×13 的特征图感受野最大,适合检测大区域火焰,相应匹配 121×295 、 179×175 、 208×312 这3种大尺寸先验框; 26×26 的特征图适合检测较大区域火焰,匹配 71×212 、 104×146 、 117×96 这3种较大尺寸的先验框; 52×52 的特征图适合检测中等区域火焰,匹配 46×150 、 58×54 、 62×96 这3种中等尺寸的先验框;而新增的 104×104 的特征图感受野最小,适合检测小区域火焰,所以匹配 17×25 、 29×37 、 29×79 这3种小尺寸的先验框。

1.4 火焰闪烁特征提取

火焰特有的闪烁频率范围为 7~10 Hz^[21],这是和其他类火干扰物最显著的区别。本文利用视频相邻多帧图像像素点亮度值变化来计算火焰闪烁特征,对火焰检测结果进行闪烁特征分析,能够有效地排除检测结果中误判的情况,提高火焰检测精度。

本文利用相邻帧之间亮度值累积差分法计算闪烁频率,建立闪烁计数矩阵 $M(x, y, t)$,亮度值矩阵 $I(x, y, t)$ 。计算公式为:

$$\begin{cases} M(x, y, t) = \begin{cases} M(x, y, t-1) + 1, \Delta I > T \\ M(x, y, t-1), \text{其他} \end{cases} \\ \Delta I = |I(x, y, t) - I(x, y, t-1)| \end{cases} \quad (2)$$

式中, $I(x, y, t)$ 为像素点 (x, y) 在 t 时刻的亮度值; $M(x, y, t)$ 为像素点 (x, y) 在 t 时刻的闪烁计数; T 为相邻两帧亮度差阈值, ΔI 超过阈值 T , 则相应的闪烁计数加 1, 否则不变。

t 时刻, 通过该像素点一段时间 T 内的闪烁次数判定其是否符合闪烁特征, T 通常取值为 1 s, 即帧速率, 1 s 内, 若闪烁计数 M 值变化超过阈值 T_M (取值为 10)^[21], 即可判定该区域是闪烁的:

$$M(x, y, t) - M(x, y, t - T) > T_M \quad (3)$$

最后, 计算一段时间内图像的闪烁频率 F 为:

$$F = \frac{M(x, y, t_n) - M(x, y, t_1)}{t_n - t_1} \quad (4)$$

1.5 视频图像火焰实时检测算法

本文算法流程如图 5 所示。

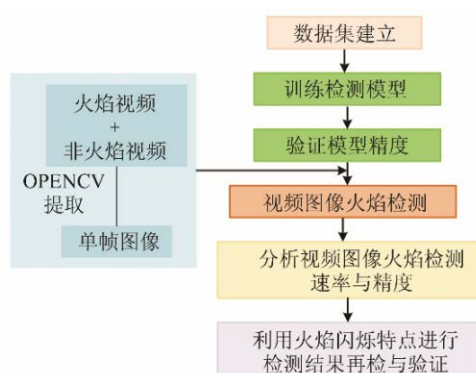


图 5 视频图像火焰检测流程

Fig.5 Flowchart of Video Image Flame Detection

首先, 通过网络获取图像和视频, 对视频进行单帧提取, 构建数据集, 并对数据集中图像进行 Resize、归一化处理。因 Darknet-53 中存在 5 次步幅为 2×2 的下采样过程, 会将特征图缩小 32 倍, 故要将图像大小 Resize 为 32 的倍数。本文将数据集图像 Resize 为 416×416 像素的尺寸。图

像归一化处理主要是在不改变图像信息的前提下, 把图像从 0~255 像素变成 0~1 像素的范围, 以加快训练网络的收敛性。其次, 利用改进的 Yolo-v3 网络以网格法分割图像并分别输出特征图, 学习图像中火焰深层语义信息和浅层位置信息, 并利用 4 个尺度特征 (如图 6 所示) 融合实现精确实时的视频图像火焰检测。最后, 利用火焰的闪烁特征, 排除火焰视频图像检测结果中误检的情况, 以提高火焰检测精度。

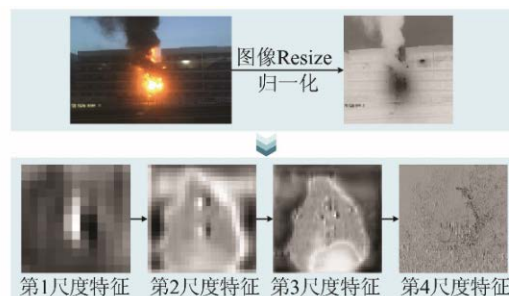


图 6 火焰特征图

Fig.6 Flame Feature Image

2 实验结果与分析

2.1 实验环境

本文所用实验数据集图像大小为 416×416 像素。改进 Yolo-v3 模型的架构、训练及测试都是在 TensorFlow 的深度学习框架 Keras 上进行的, 在带有 Win10 系统 NVIDIA GeForce GTX 1080 with Max-Q Design 8 GB 的 GPU 上实现。

2.2 实验数据集

本文实验数据集如表 2 所示。

表 2 数据集描述/张
Tab. 3 Data Description/piece

数据集名称	火焰图像	非火图像	总数
训练集	3 370	1 650	5 020
验证集	600	300	900
测试集	14 667	6 418	21 085
总数	18 637	8 368	27 005

1) 由从公开的火焰数据集 ImageNet、BoW-Fire、文献[22]中获得的 4 000 张图像, 以及通过互联网下载的 1 020 张图像, 构成本次实验的训练集, 如图 7 所示。训练集数据分为火焰图像和非火图像。

2) 由从互联网上下载的 600 张火焰数据和 300 张非火数据构成的验证集。

3) 由从 Bilkent 大学公开火焰视频库^[23]以及



图7 训练集部分数据

Fig.7 Part of the Data in Training Set

文献[24]提供的视频中选取火焰视频和非火焰视频,进行视频单帧提取获得21 085张图像,构成本次实验的测试集。测试集中视频1~6为火焰视频,视频7~10为非火焰视频。

2.3 改进 Yolo-v3 的视频图像火焰识别与分析

为全面评估基于改进 Yolo-v3 的火焰识别效果,本文利用验证集图像数据进行火焰识别测试,以验证本文模型的可靠性,测试结果如图8所示。



图8 模型验证结果

Fig.8 Model Verification Results

定义查准率 A 为正确判定火焰或非火焰的概率,误检率 P_f 为在无火的情况下误判为有火的概率,漏检率 N_f 为有火的情况下漏检的概率,计算公式为:

$$\begin{cases} A = \frac{T_p + T_n}{N_{pos} + N_{neg}} \times 100\% \\ P_f = \frac{F_p}{N_{neg}} \times 100\% \\ N_f = \frac{F_n}{N_{pos}} \times 100\% \end{cases} \quad (5)$$

式中, T_p 、 T_n 分别为火焰图像、非火焰图像被正确检测的图像数量; F_p 为非火焰图像被检测为火焰的图像数量; F_n 为火焰图像被漏检的图像数量; N_{pos} 、 N_{neg} 分别为样本中的火焰图像、非火焰图像的数量。

基于改进的 Yolo-v3 算法,火焰识别的查准率为 98.1%,误检率为 3.5%,漏检率为 1.1%。说明本文算法对火焰检测精度较高,可以将本文方法运用到视频火焰检测中。基于改进 Yolo-v3 的视频图像火焰识别测试结果如表3所示。

利用改进的模型进行视频火焰检测,查准

率可达 97.6%,误检率为 5.3%,漏检率为 1.1%。改进 Yolo-v3 的视频图像火焰检测结果如图9所示。其中,标注框左上角的数字代表模型判定该区域为火焰的概率。检测中出现的误检情况主要是闪烁灯光及阳光等类火对象的干扰所致,为排除这些情况造成的干扰,本文引入了火焰特有的闪烁特征,排除非火对象,降低视频火焰识别的误检率,优化检测精度。



图9 改进 Yolo-v3 的视频图像火焰检测结果

Fig.9 Flame Recognition Results of Video Images Based on Improved Yolo-v3

2.4 视频火焰检测结果优化与分析

根据火焰特有的闪烁特征,对检测结果中判定为火焰的帧进行再次检测,计算这些帧所在的时间间隔内火焰频率。本文以视频5、视频7、视频9和视频10为例,提取检测为火焰的帧所在的10 s间隔的视频,进行闪烁频率变化分析,结果如图10所示。

视频5闪烁频率在 8~10 Hz 内波动,在火焰的正常闪烁频率 7~10 Hz 范围内,可以断定为火焰;视频7中,仅 4~6 s 的帧符合闪烁频率,且低于 7 Hz,可判断为非火对象;视频9中闪烁频率波动范围最大值小于 7 Hz,由此可以判断为此视频中检测到的目标为非火对象;视频10中闪烁频率在 5~9 Hz 波动,根据火焰闪烁特征,可排除闪烁

表 3 火焰视频和非火焰视频检测结果

Tab. 3 Detection Results on Flame Videos and Non-flame Videos

火焰视频检测结果						非火焰视频检测结果			
视频编号	总帧数	正检帧数	查准率/%	漏检帧数	漏检率/%	视频编号	总帧数	误检帧数	误检率/%
1	650	643	98.9	7	1.1	7	2 975	67	2.3
2	1 500	1 476	98.4	24	1.6	8	2 225	0	0.0
3	1 300	1 279	98.4	21	1.9	9	550	39	7.1
4	5 875	5 823	99.1	52	0.9	10	668	234	15.4
5	3 095	3 057	98.8	38	1.2				
6	2 247	2 229	99.2	18	0.8				
均值	2 445	2 418	98.9	27	1.1	均值	1 604	85	5.3

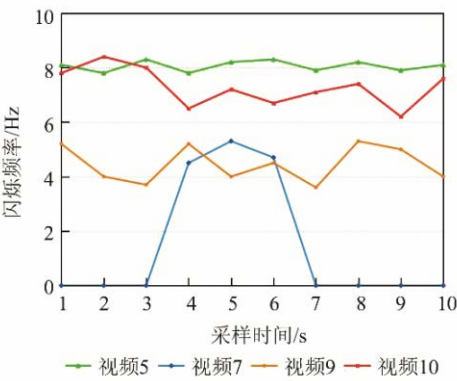


图 10 闪烁频率检测

Fig.10 Flicker Frequency Detection

频率在 5~7 Hz 的误检帧。

基于以上方法,排除类火对象对检测结果造成的干扰,可以有效地降低误检率,提高视频图像火焰检测精度,视频图像火焰检测优化结果如表 4 所示。

表 4 视频检测优化结果

Tab. 4 Video Detection Optimization Results

视频编号	总帧数	误检帧数	误检率/%
7	2 975	0	0.0
8	2 225	0	0.0
9	550	14	0.0
10	668	135	9.3
均值	1 604	37	2.3

通过对检测对象的频率分析,有效地排除类火物体对检测结果造成的干扰,误检率由 5.3% 下降到 2.3%,查准率从 97.6% 提高到 98.5%,漏检率低至 1.1%,大幅度地提高了视频图像火焰检测精度。

2.5 视频图像火焰检测速率分析

改进 Yolo-v3 的视频图像火焰检测速度统计如表 5 所示。

表 5 视频检测速度统计

Tab. 5 Video Detection Speed Statistics

视频编号	总帧数	总耗时/s	速率/(帧·s ⁻¹)
1	650	12.75	51
2	1 500	28.30	53
3	1 300	25.00	52
4	2 247	43.21	52
5	3 095	60.69	51
6	5 875	117.50	50
7	2 975	56.13	53
8	2 225	41.98	53
9	550	10.19	54
10	668	12.85	52
均值	2 109	40.86	52

本文方法的视频图像火焰检测平均速率可达 52 帧/s,完全可以满足视频图像火焰实时检测的要求。

2.6 对比实验

为了进一步证明本文算法在火焰检测上的精确性与实时性,将本文算法与文献[15]的 9 层卷积神经网络、文献[16]的两级结构卷积神经网络、文献[17]的 12 层卷积神经网络算法以及 Yolo 系列的 Yolo、Yolo-v2、Yolo-v3 算法基于本实验所用的 10 个火焰视频数据集进行对比分析,7 种算法测试结果如表 6 所示。

表 6 7 种算法结果对比

Tab. 6 Results Comparison of Seven Algorithms

算法	查准率/%	误检率/%	速率/(帧·s ⁻¹)
文献[15]算法	81.6	17.8	29
文献[16]算法	88.7	18.7	27
文献[17]算法	90.3	19.9	21
Yolo 算法	82.7	15.2	43
Yolo-v2 算法	85.4	9.4	55
Yolo-v3 算法	96.1	6.3	62
本文算法	98.5	2.3	52

由表6可知,本文算法相对于9层卷积神经网络、两级结构卷积神经网络以及12层卷积神经网络3种算法,在查准率提高的同时,检测速率也有了大幅度的提升;与Yolo系列算法相比,本文算法相比于Yolo、Yolo-v2、Yolo-v3算法,查准率分别提高了15.8%、13.1%、2.4%,误检率分别降低了12.9%、7.1%、4.0%,在检测速率上,本文算法虽然比Yolo-v3和Yolo-v2算法稍逊,但是也达到了52帧/s的检测速率,完全可以达到火焰实时检测要求。相对于其他6种算法,本文算法对于小火焰区域的检测能力以及类火物体的排除能力有了明显提升,如图11所示。

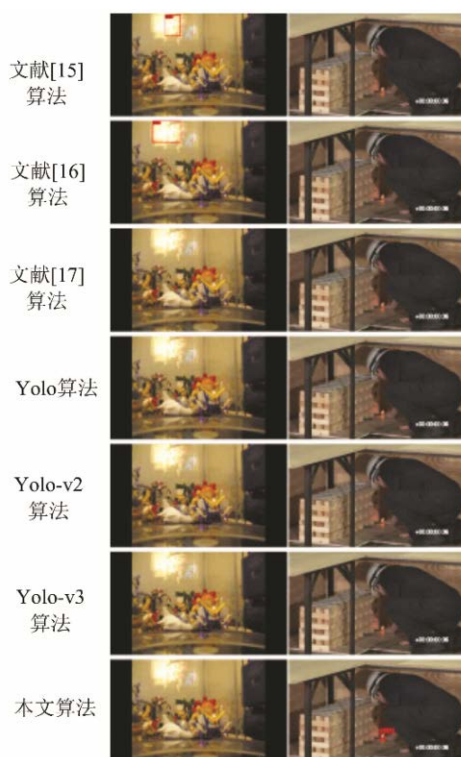


图11 7种算法检测结果对比

Fig.11 Detection Results Comparison of Seven Algorithms

从图11可以看出,本文改进的Yolo-v3算法能够准确地识别出小火焰,并能够准确排除类火物体的干扰,文献[15]的9层卷积神经网络、文献[16]的两级结构卷积神经网络均不能实现,而Yolo系列算法和文献[17]的12层卷积神经网络算法虽然能够排除类火物体的干扰,但是并不能实现对小火焰的识别。

本文算法以52帧/s的检测速率,在满足火焰检测的实时性要求下,获得了98.5%的查准率和2.3%的误检率,实验证明了本文算法在火焰实时高效检测中的可行性。

3 结 语

随着监控系统智能化发展,利用监控系统实现火焰实时高效检测对于控制火势蔓延带来的损失具有积极意义。本文优化了原Yolo-v3算法中的多尺度检测网络 and 先验框尺寸,并将优化后的算法应用到视频图像火焰识别领域,加强了网络对小火焰区域的识别能力,在提高了视频图像火焰识别精度、降低误检率的同时,也实现了实时检测。此外,本文还结合火焰的闪烁特征,在很大程度上排除了误判情况,降低了误检率,提高了检测精度。通过多组实验证明了本文算法的有效性,与现有技术相比,该算法能够更加有效地应用于视频火焰检测。

参 考 文 献

- [1] Çelik T, Demirel H, Ozkaramanli H. Automatic Fire Detection in Video Sequences [C]. The 14th European Signal Processing Conference, Florence, Italy, 2006
- [2] Çelik T, Demirel T. Fire Detection in Video Sequences Using a Generic Color Model [J]. *Fire Safety Journal*, 2009, 44(2): 147-158
- [3] Pasquale F, Alessia S, Mario V. Real-Time Fire Detection for Video-Surveillance Applications Using a Combination of Experts Based on Color, Shape, and Motion [J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2015, 25(9): 1 545 - 1 556
- [4] Ju H, Li W, Tighe S, et al. CrackU-Net: A Novel Deep Convolutional Neural Network for Pixelwise Pavement Crack Detection [J]. *Structural Control and Health Monitoring*, 2020, 27(8), DOI: 10. 1002/stc. 2551
- [5] Sun W, Paiva A, Xu P, et al. Fault Detection and Identification Using Bayesian Recurrent Neural Networks [J]. *Computers and Chemical Engineering*, 2020, 141, DOI: 10. 1016/j. compchemeng. 2020. 106991
- [6] Rei S, Yuhei H, Ayako O. Non-destructive Detection of Tea Leaf Chlorophyll Content Using Hyperspectral Reflectance and Machine Learning Algorithms [J]. *Plants*, 2020, 9(3), DOI: 10. 3390/plants9030368
- [7] Ji Shunping, Tian Siqu, Zhang Chi. Using All-Hole Convolutional Neural Network for Urban Land Cover Classification and Change Detection [J]. *Geomatics and Information Science of Wuhan University*, 2020, 45(2): 233-241(季顺平, 田思琦, 张驰. 利

- 用全空洞卷积神经网络进行城市土地覆盖分类与变化检测[J]. 武汉大学学报·信息科学版, 2020, 45(2): 233-241
- [8] Chirra V, Uyyala S, Kolli V. Deep CNN: A Machine Learning Approach for Driver Drowsiness Detection Based on Eye State[J]. *International Information and Engineering Technology Association*, 2019, 33(6): 461-466
- [9] Hu Tao, Zhu Xinyan, Guo Wei, et al. A Moving Object Detection Method Combining Color and Depth Data[J]. *Geomatics and Information Science of Wuhan University*, 2019, 44(2): 276-282(胡涛, 朱欣焰, 吕维, 等. 融合颜色和深度信息的运动目标提取方法[J]. 武汉大学学报·信息科学版, 2019, 44(2): 276-282)
- [10] Kantorov V, Oquab M, Cho M, et al. Context LocNet: Context - Aware Deep Network Models for Weakly Supervised Localization[C]. *European Conference on Computer Vision*, Amsterdam, Netherlands, 2016
- [11] Xie Y, Zhu J, Cao Y, et al. Refined Extraction of Building Outlines from High - Resolution Remote Sensing Imagery Based on a Multifeature Convolutional Neural Network and Morphological Filtering[J]. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2020, 13: 1 842-1 855
- [12] Mikolov T, Chen K, Corrado G, et al. Efficient Estimation of Word Representations in Vector Space[EB/OL]. (2013-09-07) [2019-09-25]. <https://arxiv.org/abs/1301.3781>
- [13] Hinton E, Osindero S, Teh Y. A Fast Learning Algorithm for Deep Belief Nets[J]. *Neural Computation*, 2006, 18(7): 1 527-1 554
- [14] Xie Y, Zhu J, Cao Y, et al. Efficient Video Fire Detection Exploiting Motion-Flicker-Based Dynamic Features and Deep Static Features[J]. *IEEE Access*, 2020, 8: 81 904 - 81 917
- [15] Frizzi S, Kaabi R, Bouchouicha M, et al. Convolutional Neural Network for Video Fire and Smoke Detection[C]. *Conference of the IEEE Industrial Electronics Society*, Florence, Italy, 2016
- [16] Zhang Q, Xu J, Xu L, et al. Deep Convolutional Neural Networks for Forest Fire Detection[C]. *International Forum on Management, Education and Information Technology Application*, Paris, France, 2016
- [17] Fu Tianju, Zheng Chang'e, Tian Ye, et al. Forest Fire Recognition Based on Deep Convolutional Neural Network Under Complex Background[J]. *Computer and Modernization*, 2016 (3): 52-57(傅天驹, 郑嫦娥, 田野, 等. 复杂背景下基于深度卷积神经网络的森林火灾识别[J]. 计算机与现代化, 2016 (3): 52-57)
- [18] Redmon J, Divvala S, Girshick R, et al. You Only Look Once: Unified, Real-Time Object Detection[C]. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Lasvegas, USA, 2016
- [19] Redmon J, Farhadi A. YOLO9000: Better, Faster, Stronger[C]. *IEEE Conference on Computer Vision and Pattern Recognition*, Honolulu, USA, 2017
- [20] Redmon J, Farhadi A. YOLOv3: An Incremental Improvement[EB/OL]. (2018-04-08) [2019-09-25]. <https://arxiv.org/abs/1804.02767>
- [21] Chen Chuixiong, Yan Yunyang, Liu Yian, et al. Fire Detection Based on Parallel Computing of Motion and Flicker Frequency Feature[J]. *Journal of Data Acquisition and Processing*, 2017, 32(2): 424-430(陈垂雄, 严云洋, 刘以安, 等. 运动区域提取和闪频分析并行的火焰检测算法[J]. 数据采集与处理, 2017, 32(2): 424-430)
- [22] Toulouse T, Rossi L, Campana A, et al. Computer Vision for Wildfire Research: An Evolving Image Dataset for Processing and Analysis[J]. *Fire Safety Journal*, 2017, 92: 188-194
- [23] Cetin A. Fire Detection Samples[EB/OL]. (2013-09-03) [2019-09-25]. <http://Signal.ee.bilkent.edu.tr/VisiFire>
- [24] Steffens, Botelho S, Rodrigues R. A Texture Driven Approach for Visible Spectrum Fire Detection on Mobile Robots[C]. *The XIII Latin American Robotics Symposium and IV Brazilian Robotics Symposium (LARS/SBR)*, Recife, Brazil, 2016

A Real-Time Video Flame Detection Algorithm Based on Improved Yolo-v3

ZHAO Yuanyuan¹ ZHU Jun¹ XIE Yakun¹ LI Weilian¹ GUO Yukun¹

¹ Faculty of Geosciences and Environment Engineering, Southwest Jiaotong University, Chengdu 611756, China

Abstract: **Objectives:** In order to solve the problems of the low accuracy and the slow speed of the existing video image flame detection methods, we propose a real-time video flame detection algorithm based on improved Yolo-v3 to achieve real-time and efficient detection of flames in the video. **Methods:** Firstly, in the feature extraction stage, the multi-scale detection network is improved. We add a new-scale feature and then improved the network's ability to learn the shallow information of the images by further integrating multi-scale features. Using this method, the accurate identification of small flame is achieved. Secondly, in the target detection stage, we use the improved *K*-means clustering algorithm to optimize the multi-scale prior frames, and make them adapt to the changing posture and shape of the flame. Finally, after detecting video flames based on improved Yolo-v3, we use the unique flicker characteristics of the flame to check the video again, and eliminate the false detection frame in the detection result. And in this method the detection accuracy is further improved. **Results:** In order to prove the effectiveness of our method, the video flames are detected from both accuracy and speed, and the results are compared with the advanced flame detection methods in recent years. The results show that the average accuracy rate of our method can reach 98.5%, the false detection rate is as low as 2.3%, and the average detection rate is 52 frames/s, so our method has better performance in terms of accuracy and speed. **Conclusions:** The effectiveness of this method is proved through multiple sets of experiments. Comparing with the existing flame detection methods, our method can be more effectively applied to video flame detection.

Key words: flame detection; video image; Yolo-v3; flicker characteristics; multi-scale detection

First author: ZHAO Yuanyuan, postgraduate, specializes in virtual geographic environment and disaster scenario modeling. E-mail: 3011441848@qq.com

Corresponding author: ZHU Jun, PhD, professor. E-mail: vgezj@163.com

Foundation support: The National Natural Science Foundation of China(41871289); the Scientific Research Project of Sichuan Provincial Department of Natural Resources(KJ-2020-4); the Sichuan Youth Science and Technology Innovation Team (2020JDTD0003).

引文格式: ZHAO Yuanyuan, ZHU Jun, XIE Yakun, et al. A Real-Time Video Flame Detection Algorithm Based on Improved Yolo-v3[J]. Geomatics and Information Science of Wuhan University, 2021, 46(3):326-334. DOI:10.13203/j.whugis20190440(赵媛媛, 朱军, 谢亚坤, 等. 改进 Yolo-v3 的视频图像火焰实时检测算法[J]. 武汉大学学报·信息科学版, 2021, 46(3):326-334. DOI:10.13203/j.whugis20190440)