

基于目标跟踪与深度学习的视频火焰识别方法

耿梦雅, 张国平, 徐洪波, 张莹莹

(华中师范大学 物理科学与技术学院, 湖北 武汉 430079)

摘要: 近年来火灾事故频发, 对生态环境, 社会经济都造成了严重影响, 视频监控系统在火灾预防和环境监控中都有非常重要的作用; 针对传统的视频火焰检测方法需要手工提取火焰特征且识别率低、误检率高的缺点, 提出了一种基于特征检测, 多目标跟踪和深度学习的火焰检测算法; 通过高斯混合模型运动检测方法对视频中的动态目标进行提取, 再经过 HSI 与 RGB 结合的颜色模型进行筛选, 得到疑似火焰目标, 对提取的目标进行多目标跟踪算法跟踪, 最终对稳定存在的目标通过 CaffeNet 模型进行判别, 得到火焰判别结果; 实验证明, 本算法实现了对视频火焰的准确检测, 能对火焰进行有效识别, 对火焰视频数据集上的平均识别精度达到 98.79%, 并能适应实时检测火灾的需求。

关键词: 火焰检测; 图像处理; 卷积神经网络; 多目标跟踪

Video Flame Detection Method Based on Target Tracking and Deep Learning

Geng Mengya, Zhang Guoping, Xu Hongbo, Zhang Yingying

(College of Physical Science and Technology, Central China Normal University, Wuhan 430079, China)

Abstract: In recent years, frequent fire accidents have seriously affected the ecological environment and social economy, video surveillance system plays a very important role in fire prevention and environmental monitoring. To overcome the shortcomings of traditional video flame detection methods, such as manual extraction of flame features, low recognition rate and high false detection rate, a flame detection algorithm based on feature detection, multi-target tracking and deep learning is proposed. The algorithm extracts the dynamic object in video by the Gaussian mixture model motion detection method, then filters the suspected flame object through the color model combined with HSI and RGB, and tracks the extracted object with multi-target tracking algorithm. Finally, the stable target is judged by CaffeNet model and the result of flame discrimination is obtained. Experiments show that the algorithm can detect the video flame accurately, identify, and the average recognition accuracy of the flame video data set is 98.79%, which can meet the demand of real-time flame detection.

Keywords: flame detection; image processing; convolution neural network; multi-target tracking

0 引言

由于火灾的突发性、频发性和危害性, 人们一直在寻找更加及时有效的火灾检测方法。传统的火灾探测系统存在探测范围有限, 误报率高, 安装密度大, 可拓展性不强, 智能化程度较低等缺点。近年来, 随着视频监控的普及、视频图像处理技术和计算机视觉技术的发展, 视频火焰识别一直是火灾检测领域的热点。

传统的视频火焰检测算法大多关注检测火焰的存在性, 而不能对火焰的燃烧状态进行持续有效的跟踪, 使用常见的目标跟踪算法, 如 kalman 滤波器算法对火焰目标进行跟踪, 其对于火焰这种原地闪动且存在被遮挡可能的目标跟踪效果欠佳^[1]。大多数早期的火焰检测算法都是通过先人工提取火焰的特征, 再将多种特征进行融合识别火焰。人

工提取的火焰特征主要分为静态与动态两类, 早期的火焰识别算法通过提取典型的静态特征如圆形度序列, 面积序列, 面积序列频谱, 强度序列等进行融合判别火焰^[2], 动态特征方面, 火焰的尖角运动方向^[3]、闪烁特征与边缘一致性特征^[4]等能有效区分疑似火焰的干扰物。将火焰的尖角数目与火焰高度的频谱变化等静态特征与动态特征相结合进行火焰识别, 也成为一个研究热点^[5]。但是, 人工提取特征有容错性差, 鲁棒性低的缺点。

火焰识别是视频火焰检测最关键的一步, 其识别效果决定了算法整体的实用性与可靠性, 早期的火焰识别通过建立概率模型实现^[2-5], 这种方法逐渐被机器学习算法取代, 如支持向量机^[6-7]、贝叶斯分类器^[8-9]等, 但该类方法依赖于特征的选取。由于火灾发生的背景较复杂, 容易出现干扰物, 为使算法适用于复杂的环境情况, Oleksii Maksymiv^[10]通过 AdaBoost 和 LBP 进行前期的检测, 将检测为疑似发生火灾的区域送至卷积神经网络 (CNN) 中处理, 进行最终判断, 但烟雾与火灾的分类结果影响了其识别精度。

针对上述算法的优缺点, 同时考虑到火焰检测的实时性与准确性, 本文提出了一种基于多目标跟踪与深度学习

收稿日期: 2019-01-04; 修回日期: 2019-02-11。

基金项目: 湖北省自然科学基金(2017CFB591)。

作者简介: 耿梦雅(1996-), 女, 安徽阜阳人, 硕士, 主要从事计算机视觉与火灾检测方向的研究。

通讯作者: 张国平(1969-)男, 湖北汉川人, 博士, 教授, 主要从事光电子技术、信号与信息处理方向的研究。

的火焰识别算法, 首先通过运动检测与颜色检测提取出具有火焰颜色的运动目标, 然后对疑似火焰目标进行跟踪, 进而对每个目标通过训练好的卷积神经网络模型进行预测, 得到最后的火焰识别结果。

1 疑似火焰目标提取与跟踪

对于待检测的视频序列, 对单帧图像进行检测运算量大, 不利于火焰特征的提取, 因此提取图像中疑似火焰区域将是识别前的必要步骤。传统的火焰识别算法是直接对提取到的疑似火焰目标进行识别, 判断其是否是火焰, 但这种方式不能对火焰的状态进行记录, 无法排除瞬时出现的干扰物, 导致算法误检率升高, 鲁棒性较差, 因此本算法在提取火焰目标后对目标进行持续跟踪, 从而为火焰识别阶段提供稳定存在的待识别目标。

1.1 火焰目标区域提取

火焰目标提取采用先运动检测再颜色检测的方式提取疑似火焰区域, 首先, 运动检测算法与颜色检测算法相比, 运算速度较快, 因此先运动检测, 筛选出图像中运动的物体能有效提高算法整体运算速度。其次, 根据火焰的持续不断运动的特征, 先进行运动检测可以筛选出视频中的运动物体, 进而排除具有类似火焰颜色但静止的干扰物。

运动检测利用了火焰持续不断进行随机运动的特征, 对比多种运动检测方法后使用自适应的高斯混合模型的背景建模法提取运动物体。该算法提取的前景较为连续且完整, 同时运算速度较快, 可以满足实时检测运动物体的要求。

接下来将运动前景进行分割, 对该分割区域进行颜色检测, 得到具有火焰颜色的运动目标。颜色检测使用 Chen^[11] 提出的火焰颜色模型进行识别。

颜色模型公式:

$$\begin{cases} R > R_{avg} \\ R \geq G \geq B \\ S > 0.4 \\ S \geq ((255 - R) * S_{th}) / R_{th} \end{cases} \quad (1)$$

式 (1) 即本文使用的火焰的颜色检测模型, 其采用基于 RGB 与 HSV 颜色空间相结合的火焰颜色判据, R、G、B 代表图像的 R、G、B 通道值, S 代表图像的 HSV 颜色空间的 Saturation 值, R_{avg} 代表图像中 R 通道的均值, 在本算法中, $S_{th} = 0.4, R_{th}$ 为 115。

图 1 (a)、(b)、(c) 为对一段视频进行运动检测与颜色检测后的结果。图 1 (c) 中白色部分即为二值化后的火焰区域。由检测结果可以看出, 运动检测可以检测出视频中火焰、烟雾等运动物体, 再对运动区域进行颜色检测, 可以提取出较完整的疑似火焰区域。

1.2 火焰目标跟踪

针对火焰运动速度较慢的特点, 结合多目标跟踪算法的处理速度需要, 设计了一种基于帧间距离匹配的多目标跟踪算法, 处理时先将每一帧提取的目标的外接矩形进行

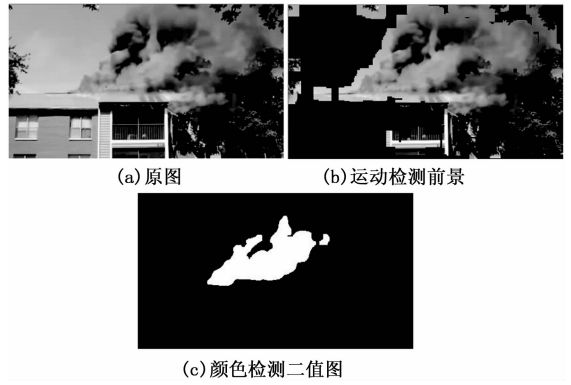


图 1 运动检测与颜色检测结果

整体融合处理。融合时与目标匹配时所依据的距离匹配关系相同, 这样可以确保接下来要融合的目标的唯一性。两目标按照距离相交或相邻的位置示意图见图 2, 图 2 中 a 图表示的情况即两目标相交, b 图表示的情况即两目标相邻, (X_{a1}, Y_{a1}) 、 (X_{a2}, Y_{a2}) 为矩形 a 的左上角坐标与右下角坐标, (X_{b1}, Y_{b1}) 、 (X_{b2}, Y_{b2}) 为矩形 b 的左上角坐标与右下角坐标, W_a 、 W_b 即矩形 a、b 的宽, H_a 、 H_b 即矩形 a、b 的高。

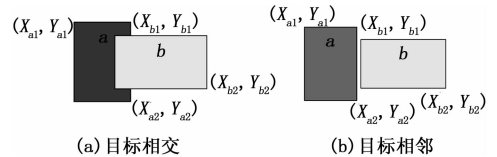


图 2 目标相交相邻位置示意图

具体的距离匹配公式见式 (2),

$$\begin{cases} \left| \frac{X_{a1} + X_{a2}}{2} - \frac{X_{b1} + X_{b2}}{2} \right| - \frac{W_a + W_b}{2} < \frac{\max(W_a, W_b)}{6} \\ \left| \frac{Y_{a1} + Y_{a2}}{2} - \frac{Y_{b1} + Y_{b2}}{2} \right| - \frac{H_a + H_b}{2} < \frac{\max(H_a, H_b)}{6} \end{cases} \quad (2)$$

若前后帧有两个目标相邻或相近则视该目标状态为存在, 即若相邻帧两个目标的外接矩形距离关系满足式 (2) 则判定其为同一目标。若前一帧的目标在当前帧无匹配结果, 则视为目标消失。类似的, 若当前帧的目标在前一帧无匹配结果, 则视为目标新增。

当出现目标离散或目标融合情况时, 可使用最大目标信息继承方式处理相应目标, 即当多个目标融合为一个目标时, 当前目标继承历史目标中面积最大目标的信息, 当历史目标离散为多个目标时, 当前多个目标中面积最大者继承历史目标的信息。为了处理短时间的目标遮挡问题, 在目标消失后算法将依据该目标的存在时间长短, 对目标信息进行一定时间保存, 当目标遮挡消失, 仍能对该目标进行持续不断的跟踪, 若该消失目标长时间没有再次出现, 再删除目标信息。这样的处理既适用于火焰的目标跟踪也保证了跟踪的稳定性。

多目标跟踪算法的跟踪效果如图 3 所示, 方框代表疑似火焰目标, 其附近的标号代表该目标的序号, 由跟踪效

果可以看出, 本算法能够处理短时间目标被遮挡以及目标离散或融合等复杂运动情况, 较好地实现了针对火焰疑似目标的实时跟踪功能。火焰作为一种运动区域相对固定的物体, 在检测火焰时进行目标跟踪可以在前期排除一些诸如汽车、自行车等快速移动的物体, 同时可以提高系统的鲁棒性, 为接下来的目标识别做准备。

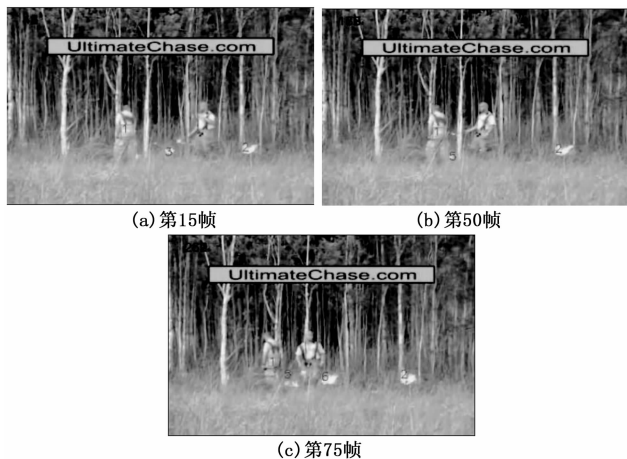


图3 多目标跟踪结果

2 基于深度学习的火焰识别

深度学习类似人脑学习的机制, 从底层特征到高层特征逐层进行抽象提取, 进而得到分类识别所需要的关键特征, 提升分类的准确性。本文使用卷积神经网络, 其可以压缩输入端维度, 在多个空间位置共享参数, 进行公共特征的提取, 尤其适合处理图片这种二维格状数据, 其适用性强, 泛化能力好, 具有全局优化能力, 且训练参数少。深度神经网络是由机器学习中的神经网络算法演变而来, 且与传统的机器学习方法相比, 深度神经网络能一次完成特征提取与分类识别, 从而避免了手工提取特征的步骤, 实现火焰识别。

Lecun 在 1998 年提出的深度卷积神经网络, 是解决模式识别问题的有效工具^[12]。原始图像输入到有多个神经元的卷积层, 其中的卷积核函数会提取输入图像的边缘, 轮廓, 尖角等局部特征。卷积层后面是池化层, 主要用于对卷积层进行特征图的下采样运算, 降低分辨率, 从而降低模型的复杂度。下采样后的特征也具有一定的尺度不变能力。卷积层和池化层交替出现在网络中。通常, 一层卷积层和一层池化层完成一次特征提取过程。在卷积神经网络中, 最后几层通常是全连接层, 这使得最终生成的特征图连接在一起。输出节点的数量可以根据分类要求进行调整, 最终得到分类结果。综上, 卷积神经网络的主要优点是不依赖于先验知识, 可以有效地解决手工特征的设计难度问题。

本文算法选用的特征提取网络为 CaffeNet, 其网络结构如图 4 所示。该网络由五层卷积层、三层全连接层组成, 其图像的输入尺寸为 227×227 , 通道数为 3, 即图像的 RGB 三通道, 网络层数为 8 层, 网络参数大小为 60 MB, 神经元个

数 650 000, 分类数目为 1 000, 并且已经在 ImageNet 数据集上做了预训练。CaffeNet 使用 RELU 函数进行激活, 将负激活部分全部清零, 以改善使用 sigmoid 激活函数带来的收敛波动以及梯度消失缺陷, 在保留神经网络非线性能力的同时, 加快训练速度。使用数据集扩展和 dropout 的方法防止过拟合。CaffeNet 的降采样技术将输入为 227×227 的图像降采样为 6×6 , 最终得到抽象后的特征。

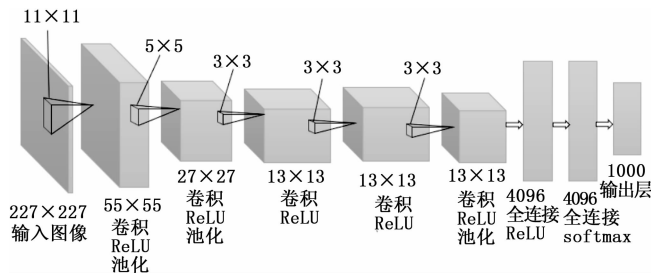


图4 CaffeNet结构示意图

微调是在数据量不足时的解决方案, 它使用复杂网络在小数据集上训练模型, 结合网络结构, 恢复已经训练好的模型权重进行训练的方式。在大规模的网络结构中, 如果训练数据集的样本量和特征维数较小, 系统可能出现局部收敛, 并导致过拟合, 对测试数据进行分类结果较差, 分类精度不高。微调可以有效地缓解这一问题。训练时, 直接在模型中使用其他层的权重, 通过改变最后一层并利用现有的数据对训练好的模型进行微调操作。该方法不仅大大提高了训练速度, 还能减少过拟合问题。在本文中, 我们对 CaffeNet 的 Caffe 模型进行微调, 以适应火焰的分类任务。

首先, 网络中的 `crop_size` 为 227×227 。在 Caffe 的训练过程中, 训练和验证同时进行。`test_iter` 值设置为 260 代表每 1000 次迭代计算一次误差, `batch_size` 设置为 16, 即每次迭代的图像数是 `batch_size` \times `crop_size` $= 16 \times 1000 = 16\ 000$ 。每一个迭代过程, 所有的训练数据集都将通过网络训练一次。在 `solover.prototxt` 中, 迭代的最大次数为 `max_iteration` $= 50\ 000$, 因此 `epoch` $= 50\ 000 / 1\ 000 = 50$ 。当初始学习率 `base_lr` $= 0.01$ 时, 训练中 `loss` 值不会下降, 甚至在一定的值停止, 通过调整 `base_lr` 参数来解决这个问题。在本实验中, 将 `base_lr` 设置为 0.001, `lr_policy` 设置为 `step`, `Gamma` $= \gamma = 0.1$, `stepsize` $= 20\ 000$ 。根据 Step 学习策略, 学习速率的衰减时间与 `step` 有关, 衰减程度与 “ γ ” 有关。例如, 当设置 `step` $= 500$, `base_lr` $= 0.000\ 1$, `Gamma` $= 0.1$ 时, 达到第一次 500 迭代次数时, 学习率将开始衰减。衰减后的学习率为 `a` \times `a` $\times \gamma = 0.000\ 1 \times 0.1 = 10^{-5}$ 。即 `step` 表示学习速率的衰减步长, “ γ ” 表示衰减因子。优化后的网络训练参数见表 1。

3 实验结果

本算法的整体流程图如图 5 所示。其中程序初始化主要实现阈值初始化、火焰识别模型加载等步骤, 接着对输入的视频进行运动检测提取出运动前景, 并对运动前景进行

表 1 CaffeNet 优化前后网络参数

参数名	原始参数	优化参数
test_iter	1000	260
base_lr	0.01	0.001
step_size	100000	20000
max_iter	450000	50000
solver_mode	GPU	CPU
batch_size	256	16

颜色检测提取出面积较大的显著目标，然后对疑似火焰目标进行多跟踪跟踪并对目标信息进行存储与更新，并对持续存在的目标进行火焰识别，最后根据当前目标被识别为火焰的频次得到识别结果，实现火灾检测。

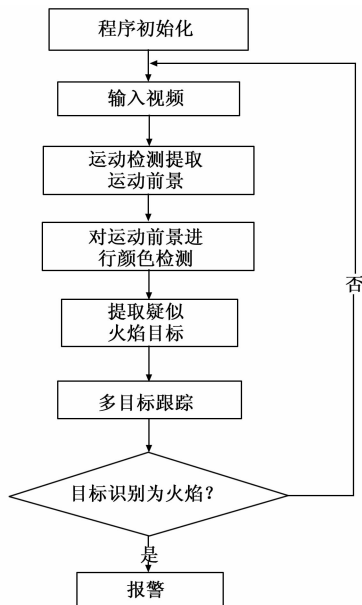


图 5 算法整体流程图

本实验软件环境为 Windows10 系统下 Caffe 开源框架，硬件配置为：英特尔 Core i7-8550U 八核处理器，8 G 内存，128 G 固态硬盘等。由于火焰检测领域还没有公开、统一的标准数据集，因此本实验使用的火焰样本视频部分来自 Bilkent 大学火灾火焰视频库 (<http://Signal.ee.bilkent.edu.tr/VisiFire>)，部分来自网络及作者拍摄。

3.1 数据集准备

火焰识别模型的训练标签有两类：火焰与非火焰，为实现卷积神经网络的训练，从火焰视频数据集与火焰图片数据集中整理了 2400 张火焰图片作为火焰训练集，1020 张火焰图片作为火焰测试集，2400 张疑似火焰图片作为疑似火焰训练集，1020 张疑似火焰图片作为疑似火焰测试集，两类样本数量达到 1:1 平衡，且训练集与测试集数量基本满足 2.4:1，符合训练集与测试集的合理划分比例。为满足 CaffeNet 的输入要求，图像大小统一为 227pixel×227pixel，将数据集转化为 LMDB 格式存储，并输入 CaffeNet 处理。

3.2 CaffeNet 微调结果

实验过程中，网络训练采用微调策略，每迭代 1000 次

进行一次预测，迭代训练 50000 次后网络已经具备了识别火焰与疑似火焰目标的能力，并可通过输出对应各类别的概率实现对目标是否是火焰的量化评价。网络训练过程的损失函数如图 6 所示，从图 6 中可以看出，经过 10000 次的迭代训练后迅速收敛，并保持平稳。网络的训练误差在 20000 次迭代训练后逐步下降，并在 50000 次附近达到平稳，实验结果表明网络得到了有效学习。如图 7 所示，网络的测试精度也逐步上升，最终网络模型的精度达到 97.79%。

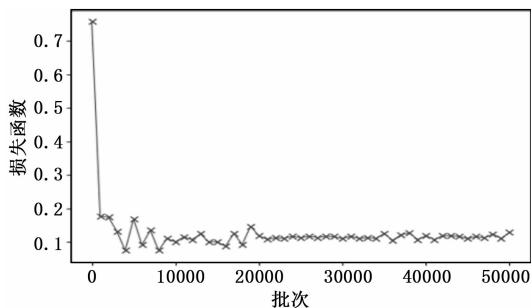


图 6 CaffeNet 训练阶段误差

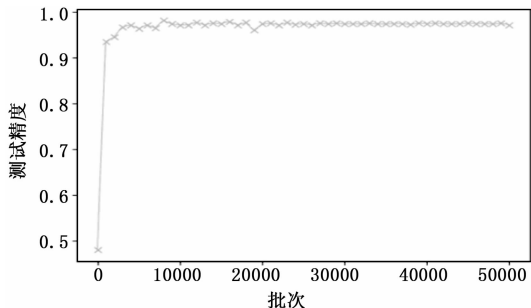


图 7 CaffeNet 测试阶段精度

3.3 算法识别结果

本文为测试算法检测火焰的准确性，使用了多个火焰与非火焰视频进行了测试，火焰视频包含了多种场景下，各个阶段的火焰燃烧情况，非火焰视频包含了一些生活中较为常见的与火焰相似的物体，具有较强的干扰性。

算法实验结果如图 8 所示，对于识别为火焰的目标，用红色框高亮显示，并在旁边标注为火焰；对于识别为非火焰的目标，用绿色框高亮显示，并在旁边标注为非火焰。火焰误报率即为将非火焰目标识别为火焰目标的概率，本文算法的平均检测率达到了 98.79%，且对各种场景的火焰都具有较好的检测效果。图 8 是不同场景的测试结果图，结果表明本算法能成功排除相似颜色的树干，移动灯光，带有火焰颜色的烟雾等干扰，稳定地检测出不同场景、不同光照下的火焰。

为了验证算法的识别准确率，使用上述视频样本对算法进行测试，检测结果如表 2 所示，同时，将本算法的识别准确率与三种典型火焰检测算法进行对比，文献 [11] 采用的是 RGB-HSI 混合颜色模型，通过运动检测与颜色检测结合面积变化特征识别火焰，该方法对稳定燃烧的火



图8 不同场景的测试结果

焰识别效果不好且容易误判。文献 [6] 使用水平、垂直和对角方位的亮度图结合 SVM 分类器建立了一种基于像素的非线性火焰分类方法, 实验结果表明该方法对于火焰检测具有较强的鲁棒性, 但存在对比度较高时容易误判的问题。文献 [13] 提出的基于 RGB 颜色空间统计模型的火焰识别算法, 提取火焰的分层与闪烁特征, 利用神经网络进行火焰的识别, 该算法鲁棒性较好, 但对大空间内火焰面积占比较小的视频检测效果较差。由表 2 和表 3 可见, 与类似算法相比本算法具有较高的可靠性与准确性, 抗干扰能力强, 误报率低, 检测效果好。在算法运行效率上, 火焰检测识别算法中前期的特征提取与跟踪部分计算量较小, 但火焰判别部分对算法运行效率影响较大, 在测试实验中, 本文算法平均帧处理时间为 147 ms。利用每三帧识别一次的方式处理, 在加快算法运行速率的同时也能保证算法的识别效果。在应用中火焰检测的平均处理效率达到了 20 帧/s, 能很好地满足工业需求。

表2 火焰视频检测结果

视频 编号	总帧 数/帧	含火帧 数/帧	识别率/%			
			文献[11]	文献[6]	文献[13]	本文算法
视频 1	439	439	95.7	96.0	96.1	97.7
视频 2	260	260	95.2	96.5	85.9	99.6
视频 3	246	246	90.7	96.8	99.6	100

表3 非火焰视频检测结果

视频 编号	总帧 数/帧	含火帧 数/帧	误报率/%			
			文献[11]	文献[6]	文献[13]	本文算法
视频 4	155	0	67.2	7.7	4.6	0.3
视频 5	160	0	71.4	45.3	5.7	0.1
视频 6	333	0	57.3	12.5	2.4	0.0

4 总结

本文提出了一种基于多目标跟踪与深度学习的火焰检测算法, 使用多目标跟踪的方法使得算法能对火焰目标进行持续稳定的跟踪, 增强系统的鲁棒性, 使用深度学习的

方法进行火焰目标识别, 使得本算法与传统机器学习方法进行火焰检测相比, 拥有更好的准确率与更低的误报率, 同时可以避免人工提取火焰特征的缺点。测试结果表明, 本算法具有较为理想的火焰检测效果, 同时具备较好的实时性与抗干扰能力, 具有较好的应用前景。但本算法仅能检测到火焰的发生, 对于阴燃及火焰被完全遮挡的情况还没有考虑, 因此后面将增加烟雾检测模块, 实现对火焰更为及时有效的预警。

参考文献:

- [1] 张学敏, 蔡晓东, 梁玉敏, 等. 基于目标跟踪和多特征融合的火焰检测算法 [J]. 电视技术, 2013, 37 (15): 205-210.
- [2] Li H, Liu Q, Wang S. A novel fire recognition algorithm based on flame's Multi-features Fusion [A]. International Conference on Computer Communication and Informatics [C]. IEEE, 2012: 1-6.
- [3] Dattathreya, Kim H, Park J, et al. Fire flame detection based on color model and motion estimation [A]. IEEE International Conference on Consumer Electronics-Asia [C]. IEEE, 2017: 1-2.
- [4] Liu Q, Wang S, Zhang X H, et al. Flame Recognition Algorithm Research under Complex Background [A]. IEEE, International Conference on Computer and Information Technology [C]. IEEE, 2010: 503-510.
- [5] 张进华, 庄健, 杜海峰, 等. 一种基于视频多特征融合的火焰识别算法 [J]. 西安交通大学学报, 2006, 40 (7): 811-814.
- [6] Ko B C, Cheong K H, Nam J Y. Fire detection based on vision sensor and support vector machines [J]. Fire Safety Journal, 2009, 44 (3): 322-329.
- [7] Yang N, Wang H, Zhang Q, et al. Adaptive flame image detection algorithm [A]. International Congress on Image and Signal Processing [C]. IEEE, 2010: 1787-1791.
- [8] Ko B C, Cheong K H, Nam J Y. Early fire detection algorithm based on irregular patterns of flames and hierarchical Bayesian Networks [J]. Fire Safety Journal, 2010, 45 (4): 262-270.
- [9] Koeirich Borges P V, Mayer J, Izquierdo E. Efficient visual fire detection applied for video retrieval [C]. Signal Processing Conference, 2008, European. IEEE, 2008: 1-5.
- [10] Maksymiv O, Rak T, Peleshko D. Real-time fire detection method combining AdaBoost, LBP and convolutional neural network in video sequence [A]. Experience of Designing and Application of Cad Systems in Microelectronics [C]. IEEE, 2017: 351-353.
- [11] Chen T H, Wu P H, Chiou Y C. An early fire-detection method based on image processing [A]. International Conference on Image Processing [C]. IEEE, 2005 (3): 1707-1710.
- [12] Le Y. Cun: Generalization and Network Design Strategies, Connectionism in Steels [C]. Elsevier Science Publishers Bv, North-holland. 1989.
- [13] 陈嘉卿, 张冰, 宋英磊. 基于 RGB 统计颜色模型的火焰识别 [J]. 江苏科技大学学报 (自然科学版), 2017, 31 (2): 178-184.