

# 基于 Faster R-CNN 模型的火焰检测

严云洋<sup>1,2</sup>, 朱晓妤<sup>1,2</sup>, 刘以安<sup>2</sup>, 高尚兵<sup>1</sup>

(1. 淮阴工学院计算机与软件工程学院, 江苏 淮安 223003)

(2. 江南大学物联网工程学院, 江苏 无锡 214122)

[摘要] 常规的火焰检测一般是提取火焰的静态或动态特征, 然后进行火焰的判别。但是传统特征无法全面描述火焰特性, 会导致识别的准确率降低。本文提出一种基于 Faster R-CNN 模型的火焰检测算法。首先利用候选区域生成网络(Region Proposal Network, RPN)提取火焰候选区域, 然后对候选区域进行卷积及池化操作, 提取火焰特征, 最后利用联合训练的快速区域卷积神经网络(Fast R-CNN)进行火焰识别。实验结果表明该方法能够自动提取火焰特征, 有效提高复杂背景下的火焰识别的准确率, 具有良好的泛化能力和鲁棒性。

[关键词] Faster R-CNN, 候选区域生成网络, 快速区域卷积神经网络, 火焰检测

[中图分类号] TP391.41 [文献标志码] A [文章编号] 1001-4616(2018)03-0001-05

## Flame Detection Based on Faster R-CNN Model

Yan Yunyang<sup>1,2</sup>, Zhu Xiaoyu<sup>1,2</sup>, Liu Yi'an<sup>2</sup>, Gao Shangbing<sup>1</sup>

(1. Faculty of Computer & Software Engineering, Huaiyin Institute of Technology, Huai'an 223003, China)

(2. School of Internet of Things Engineering, Jiangnan University, Wuxi 214122, China)

**Abstract:** Usually the static or dynamic characteristics of the flame is extracted for flame detection. But the traditional characteristics can not fully describe the characteristics of flame, which leads to the reduction of recognition accuracy. To solve this problem, a flame detection based on Faster R-CNN model is proposed in this paper. First, the candidate region of the flame is extracted by RPN. Then the convolution and pool operation of candidate regions are performed to extract the flame characteristics. Finally, Fast R-CNN is used to identify the flame. The experimental results show that the method can automatically extract the flame characteristics, effectively improve the accuracy of flame recognition in the complex background, and have good generalization ability and robustness.

**Key words:** Faster R-CNN, RPN, Fast R-CNN, flame detection

火焰是火灾发生时的重要视觉标志之一, 因此火焰的研究对于火灾的准确监控具有重要作用。传统的火灾探测器一般是基于感温、感烟和感光等探测设备进行火灾的检测, 容易受到自然环境的限制<sup>[1]</sup>。与基于传感器的火焰检测系统相比, 基于图像处理火焰检测系统不仅可以对火灾进行快速及时的响应, 同时可以提供清楚的火灾现场情况、保存事故发生的信息以便及时施救和事后处理, 正成为火焰探测技术的研究热点。

熊国良等<sup>[2]</sup>根据火焰燃烧时的特征, 选取火焰的圆形度、尖角个数、面积的变化和质心点的运动等特征作为区分火焰的主要标志。吴茜茜等<sup>[3]</sup>根据 HSI 和 RGB 颜色模型之间的关系, 提出了一种新的颜色判断规则, 更加完整地提取了火焰候选区域, 然后对该区域提取闪频特征, 最后使用逻辑回归进行火焰检测。胡勤等<sup>[4]</sup>结合火焰动态纹理特征, 在火灾发生的初期和中后期使用滑动窗口机制对火焰进行检测。陈磊等<sup>[5]</sup>提出对视频图像中的连续帧进行分块差分, 根据像素间的关联情况获取稳定的运动目标, 从而减少噪声对运动区域的影响。Stadler 等<sup>[6]</sup>使用加权帧间差分法提取火焰候选区域, 通过高闪烁频率区分火焰和非火焰, 将阈值滤波和高通滤波器应用于强度变化从而提高识别率。但是火焰的颜色、纹理、形状等特征, 只反映了火焰的部分特性, 应用这些特征的火焰检测率不高, 因此需要更强的抽象特征表示。相比

收稿日期: 2018-04-16.

基金项目: 国家自然科学基金(61402192)、江苏省“六大人才高峰”项目(2013DZXX-023)、江苏省“青蓝工程”、淮安市“533 英才工程”。

通讯联系人: 严云洋, 博士, 教授, 研究方向: 数字图像处理、模式识别。E-mail: areyyyke@163.com

传统人工提取特征,基于深度卷积神经网络对火焰目标具有更强的表达能力.卷积神经网络(Convolutional Neural Networks,CNN)<sup>[7]</sup>、基于区域的卷积神经网络(Region-based CNN,R-CNN)<sup>[8]</sup>和基于快速区域的卷积神经网络(Fast Region-based CNN,Fast R-CNN)<sup>[9]</sup>等深度神经网络在图像检测和图像识别上取得了较好的效果.但是传统的区域生成策略非常耗时,而在此基础上改进的 Faster R-CNN 模型采用共享的卷积网络,大幅度地提升了目标检测的速度.

本文端到端地实现了自动提取火焰特征并进行检测的过程,通过联合训练方法实现定位与识别的有机统一.利用 RPN 在表征原始图像的特征图上生成火焰区域候选框,最后通过联合训练的 Fast R-CNN 实现火焰的检测.实验结果表明,该方法在复杂的场景中具有较好的火焰识别率和鲁棒性.

1 基于 Faster R-CNN 模型的火焰检测

1.1 Faster R-CNN 整体结构

Faster R-CNN 主要由候选区域生成网络(RPN)<sup>[10]</sup>和快速区域卷积神经网络(Fast R-CNN)两个模块组成.基于 Faster R-CNN 模型的火焰检测主要有 4 个步骤:火焰候选区域的生成、火焰特征的提取、火焰分类和边框回归.图 1 为 Faster R-CNN 模型结构图.首先,对于任意大小的输入图片,将其尺寸归一化为 224×224.然后,通过 CNN 中的卷积层和池化层产生特征映射图(Feature Map),其中 CNN 模型包含 13 个卷积层(卷积层+激励层)、5 个最大池化层和 2 个全连接层.在所有 Feature Map 上通过候选区域生成网络(Region Proposal Network,RPN)提取火焰候选区域(Region of Interest,RoI)和区域得分,对区域得分进行非极大值抑制操作,将得分最高的区域作为最终的火焰候选区域.任意一个火焰候选区域经过 RoI 池化层映射到 Feature Map 的相应位置上,并对该区域下采样一个固定大小的特征图,然后对每个输入的特征图通过全连接层提取一个固定维度的特征向量.最后将这个特征向量输入到两个同级的输出层:一个是 Softmax 分类层,判断目标是否是火焰;另一个是边界回归层,主要对 RoI 边框位置和大小进行微调.

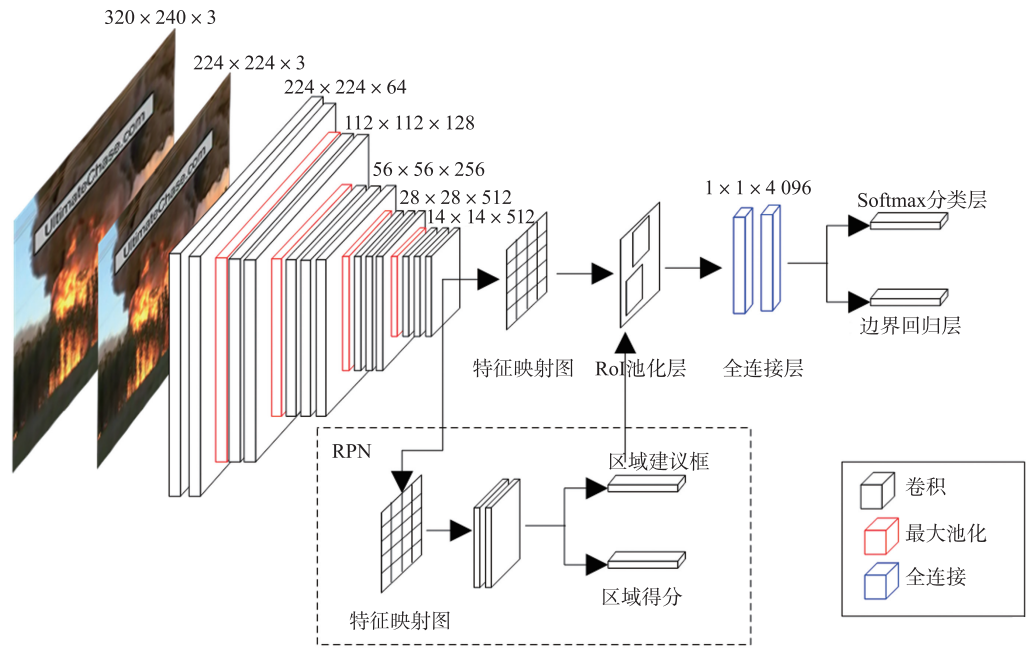


图 1 Faster R-CNN 模型结构图  
Fig. 1 Structure diagram of Faster R-CNN model

1.2 候选区域生成网络

RPN 用来生成高质量的火焰候选区域框,它是一个全卷积神经网络,通过 RPN 提取火焰候选区域,与火焰检测网络共享卷积特征.本文采用 3 种不同大小的 Anchor(128×128,256×256,512×512),用 3 种长宽比(1:1,1:2,2:1)进行缩放滑动窗口,因此共用 9 种类型的 Anchor 对火焰候选区域进行预测.将每个滑动窗口映射成一个低维特征向量,将这些特征作为两个全连接层(分类层和边界回归层)的输入,分类层给出火焰候选区域中的物体是火焰和非火焰的概率,边界回归层给出火焰候选区域的坐标.由于提取的

火焰候选区域有很多交叉重叠,因此根据非极大值抑制原则选取前 300 个候选区域框作为最终的火焰候选区域. 图 2 为 RPN 模型的结构图.

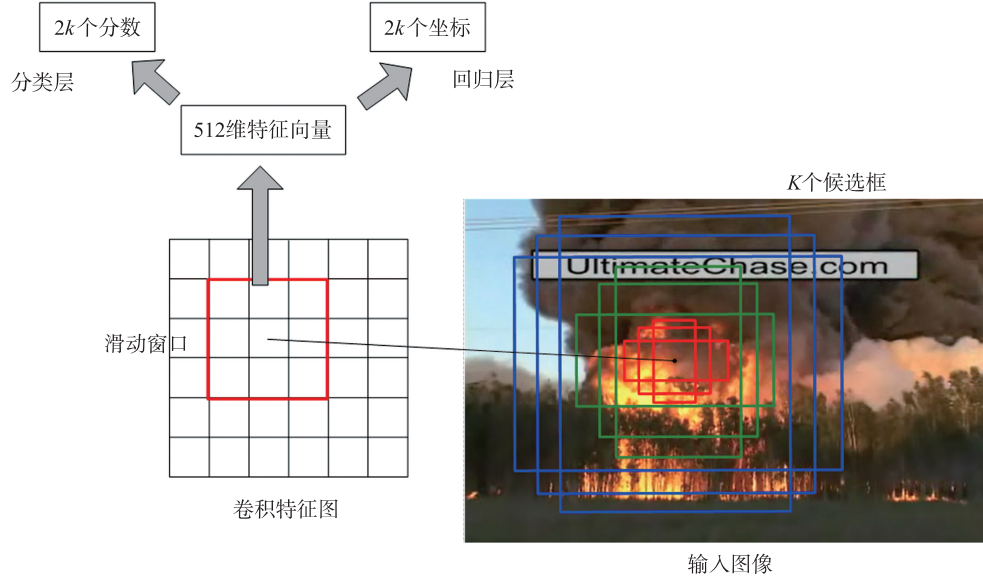


图 2 RPN 结构图

Fig. 2 Structure diagram of RPN

### 1.3 模型训练

Faster R-CNN 模型采用交替训练的方式,使得 RPN 和 Fast R-CNN 共享卷积层参数. 具体的训练过程如下:

Step 1: 训练 RPN. 利用 ImageNet 预训练的模型初始化网络参数,通过端到端微调生成候选区域;

Step 2: 训练 Fast R-CNN 模型. 使用 Step1 中 RPN 生成的候选区域训练 Fast R-CNN 模型,同样采用 ImageNet 预训练的模型初始化检测网络参数;

Step 3: 初始化新的 RPN. 利用 Step 2 中得到的检测网络模型初始化一个新的 RPN,同时共享卷积层的参数,微调 RPN 独有部分的参数.

Step 4: 微调 Faster R-CNN 的全连接层. 保持共享的卷积层参数不变,利用 RPN 的结果对全连接层进行微调,构成一个统一的网络.

在训练的过程中,如果一个候选框和任意一个火焰的真实包围框的边框重叠度 (Intersection over Union, IoU) 最大或者大于 0.7,该候选框被标记为正样本,如果一个候选框和某一个火焰真实包围框的 IoU 值小于 0.3,该候选框被标记为负样本. 对于每一个区域候选框其损失函数为:

$$L(p_i, t_i) = \frac{1}{N_{cls}} \sum_i L_{cls}(p_i, p_i^*) + \lambda \frac{1}{N_{reg}} \sum_i p_i^* L_{reg}(t_i, t_i^*). \quad (1)$$

式中,  $i$  为候选区域的编号,  $p_i$  是第  $i$  个候选区域是火焰的概率.  $p_i^* = 1$  表示候选区域为正样本,  $p_i^* = 0$  表示候选区域为负样本.  $t_i = (t_x, t_y, t_w, t_h)$  表示边框的坐标值,  $\lambda$  为平衡权重,  $\lambda = 1$ . 其中  $L_{cls}$  为分类损失函数,  $L_{reg}$  为回归损失函数:

$$L_{cls}(p_i, p_i^*) = -\ln[p_i^* p_i + (1 - p_i^*)(1 - p_i)]. \quad (2)$$

$$L_{reg}(t_i, t_i^*) = \sum_{i \in \{x, y, w, h\}} \text{smooth}_{L1}(t_i, t_i^*). \quad (3)$$

$$\text{smooth}_{L1}(x) = \begin{cases} 0.5x^2, & |x| \leq 1, \\ |x| - 0.5, & \text{其他}. \end{cases} \quad (4)$$

## 2 实验及结果分析

### 2.1 实验数据

在训练过程中使用的数据集来源于互联网,并将视频按帧转化成图片,然后根据 PASCAL VOC 的格

式手工标注图片中的目标. 实验中的训练集共有 20 个场景,5 840 张图片,将火焰区域标注为 fire,将疑似火焰区域(如车灯、指示灯等)标注为 likefire,未标注的区域为背景,火焰目标个数和疑似火焰目标个数分别为 4 000 张.

2.2 结果与分析

本文实验所用机器的基本配置为 GPU: GTX 1070Ti, CPU: AMD Ryzen 71700, 内存 8G, 操作系统: Ubuntu 16.04, 使用深度学习中的 Caffe<sup>[11]</sup> 框架. 实验视频集均来自于 Bilkent 大学火焰视频库, 视频帧的频率为 25 帧/s, 大小为 320×240. 包括 6 段火焰视频和 2 段非火焰视频(视频样例及检测结果如图 3 所示, 视频的描述如表 1 所示).



图 3 视频样本示例及检测结果  
Fig. 3 Video examples and detection results

表 1 数据库视频集内容描述

Table 1 Description of the video

视频	视频内容描述
1	庭院火焰, 香炉内的火焰随风摇曳, 后面的墙壁倒映着火焰的光影.
2	森林火焰, 草丛中燃烧着火焰, 人在火焰附近走动, 带动火焰摆动.
3	森林火焰, 树木剧烈燃烧, 火焰燃烧面积较大, 且燃烧过程中伴随着浓烟.
4	庭院火焰, 即将燃尽的火焰, 面积较小, 地面颜色与火焰颜色相似.
5	森林火焰, 整片森林剧烈燃烧, 并且伴随着大量的黑色的浓烟.
6	公路火焰, 草地和火焰的颜色相似, 远处的公路光照较强, 颜色较亮.
7	公路场景, 深夜的马路上车辆从远处驶来, 明亮的车灯照亮四周, 使得周围的颜色与火焰相似.
8	公路场景, 路边停靠的汽车打开车灯, 照亮了整个路面. 右侧行人前行, 从身后经过一辆开着车灯的大卡车.

实验结果及效果对比如表 2 和表 3 所示. 其中 TPR(True Positive Rate)代表火焰检测率, (False Negative Rate)代表火焰漏检率, (True Negative Rate)代表非火焰正确率, (False Positive Rate)代表非火焰误检率.

表 2 火焰视频检测结果

Table 2 Detection results on fire videos

视频信息			TPR/%				FNR/%			
视频	总帧数	火焰帧数	文献[ 12]	文献[ 13]	文献[ 14]	本文算法	文献[ 12]	文献[ 13]	文献[ 14]	本文算法
1	428	428	92.5	90.7	93.0	97.5	7.5	9.3	7.0	2.5
2	249	249	90.4	77.9	89.2	95.8	9.6	22.1	10.8	4.2
3	235	235	95.7	94.7	86.8	96.7	4.3	5.3	13.2	3.3
4	697	697	87.9	90.5	90.4	95.8	12.1	9.5	9.6	4.2
5	234	234	91.0	95.7	88.9	99.0	9.0	4.3	11.1	1.0
6	1 190	1 190	88.6	90.6	98.8	94.9	11.4	9.4	1.2	5.1

表 3 非火焰视频检测结果

Table 3 Detection results on non-fire videos

视频信息			TPR/%				FNR/%			
视频	总帧数	火焰帧数	文献[ 12]	文献[ 13]	文献[ 14]	本文算法	文献[ 12]	文献[ 13]	文献[ 14]	本文算法
7	144	0	95.1	93.3	82.6	98.3	4.9	6.7	17.4	1.7
8	149	0	94.0	83.6	83.9	97.4	6.0	16.4	16.1	2.6



文献[12]先使用 HSV 颜色模型提取疑似火焰区域,排除与火焰颜色相近的干扰物.然后利用高斯混合模型(Gauss Mixture Model,GMM)检测运动目标.该算法对简单场景下的火焰具有较好的检测效果,但是对于复杂的室外或森林火焰,由于背景变化较快且复杂,因此检测率比较低只能达到 88.6%.文献[13]提取出面积、边界粗糙度和颜色等火焰特征,然后将统计特征输入到 BP 神经网络中进行判别.但是由于特征提取不够全面,所以当出现与火焰纹理相似的运动区域时,误检率较高,如视频 8 中的车灯导致周围的环境亮度较高,纹理与火焰相似,导致误检率达到 16.4%.相比文献[12-13]中传统的火焰识别算法,文献[14]使用深度卷积神经网络对火焰进行检测,提高了准确率.但是深度卷积神经网络需要固定的输入尺寸,因此对输入数据进行归一化操作时,导致了图像的失真,造成准确率的下降.同时使用基于 CPU 的搜索选择方法进行候选区域的提取,造成过多冗余区域框的产生,效率降低.本文使用 RPN 提取高质量的火焰候选区域,便于后续火焰目标的提取和检测.然后使用卷积神经网络提取抽象特征,能够更加全面地描述火焰,因此本文算法对不同场景都具有较好的检测能力,平均火焰检测率达到了 96.93%.

### 3 结论

本文通过深度卷积神经网络 Faster R-CNN 模型进行火焰检测,首先使用卷积神经网络对图片进行卷积操作得到特征映射图,然后在特征映射图上使用 RPN 得到火焰候选框的信息,最后对候选框提取的特征使用联合训练的 Fast R-CNN 进行识别,实现了特征提取和火焰检测的端到端的过程.实验结果表明,该方法在不同的场景中都具有较高的检测率,同时提高了复杂场景下的火焰识别率.

#### [参考文献]

- [1] 相徐斌. 基于视频的烟雾检测算法研究[D]. 杭州:浙江大学,2017.
- [2] 熊国良,苏兆熙,刘举平,等. 火焰特性识别的 Matlab 实现方法[J]. 计算机工程与科学,2013,35(7):131-136.
- [3] 严云洋,吴茜茜,杜静,等. 基于色彩和闪频特征的视频火焰检测[J]. 计算机科学与探索,2014,8(10):1271-1279.
- [4] 胡勤,陈琛,刘敏. 一种基于动态纹理的烟雾和火焰检测方法[J]. 消防科学与技术,2014,33(6):667-669.
- [5] 陈磊,黄继凤. 基于视频的火焰检测方法[J]. 计算机工程与设计,2014,35(9):3143-3147.
- [6] STADLER A, WINDISCH T, DIEPOLD K. Comparison of intensity flickering features for video based flame detection algorithms[J]. Fire safety journal,2014,66:1-7.
- [7] 王华利,邹俊忠,张见,等. 基于深度卷积神经网络的快速图像分类算法[J]. 计算机工程与应用,2017,53(13):181-188.
- [8] GIRSHICK R, DONAHUE J, DARRELL T, et al. Region-based convolutional networks for accurate object detection and segmentation[J]. IEEE Trans Pattern Anal Mach Intell,2015,38(1):142-158.
- [9] 曹诗雨,刘跃虎,李辛昭. 基于 Fast R-CNN 的车辆目标检测[J]. 中国图象图形学报,2017,22(5):671-677.
- [10] 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.
- [11] JIA Y, SHELHAMER E, DONAHUE S, et al. Caffe:convolutional architecture for fast feature embedding[C]//Proceedings of the 22nd ACM International Conference on Multimedia. Orlando, Florida:ACM,2014.
- [12] CHMELAR P, BENKRID A. Efficiency of HSV over RGB Gaussian Mixture Model for fire detection[C]//Radioelektronika. Slovakia:IEEE,2014.
- [13] RONG J Z, ZHOU D C, YAO W, et al. Fire flame detection based on GICA and target tracking[J]. Optics & laser technology, 2013,47:283-291.
- [14] FRIZZI S, KAABI R, BOUCHOUICHA M, et al. Convolutional neural network for video fire and smoke detection[C]//Industrial Electronics Society, IECON 2016-42nd Annual Conference of the IEEE. USA:IEEE,2016.

[责任编辑:黄 敏]