

# 基于改进 YOLOv3 的火灾检测与识别<sup>①</sup>



任嘉锋<sup>1</sup>, 熊卫华<sup>1</sup>, 吴之昊<sup>1</sup>, 姜明<sup>2</sup>

<sup>1</sup>浙江理工大学 机械与自动控制学院, 杭州 310018)

<sup>2</sup>杭州电子科技大学 计算机学院, 杭州 310018)

通讯作者: 熊卫华, E-mail: xiongwh@zstu.edu.cn

**摘要:** 现阶段火灾频发, 需要自动进行火灾的检测与识别, 虽然存在温度、烟雾传感器等火灾检测手段, 但是检测实时性得不到保证. 为了解决这一问题, 提出了基于改进 YOLOv3 的火灾检测与识别的方法. 首先构建一个多场景大规模火灾目标检测数据库, 对火焰和烟雾区域进行类别和位置的标注, 并针对 YOLOv3 小目标识别性能不足的问题进行了改进. 结合深度网络的特征提取能力, 将火灾检测与识别形式化为多分类识别和坐标回归问题, 得到了不同场景下火焰和烟雾两种特征的检测识别模型. 实验表明, 本文提出的改进 YOLOv3 算法对不同拍摄角度、不同光照条件下的火焰和烟雾检测都能得到理想的效果, 同时在检测速度上也满足了实时检测的需求.

**关键词:** 深度学习; 机器视觉; YOLOv3; 火灾检测

引用格式: 任嘉锋, 熊卫华, 吴之昊, 姜明. 基于改进 YOLOv3 的火灾检测与识别. 计算机系统应用, 2019, 28(12): 171-176. <http://www.c-s-a.org.cn/1003-3254/7184.html>

## Fire Detection and Identification Based on Improved YOLOv3

REN Jia-Feng<sup>1</sup>, XIONG Wei-Hua<sup>1</sup>, WU Zhi-Hao<sup>1</sup>, JIANG Ming<sup>2</sup>

<sup>1</sup>(Faculty of Mechanical Engineering and Automation, Zhejiang Sci-Tech University, Hangzhou 310018, China)

<sup>2</sup>(School of Computer Science and Technology, Hangzhou Dianzi University, Hangzhou 310018, China)

**Abstract:** At the present stage, fires occur frequently, and fire detection and identification are required automatically. Although there are fire detection methods such as temperature and smoke sensors, the real-time detection is not guaranteed. To solve this problem, a method based on improved YOLOv3 fire detection and identification is proposed. Firstly, a multi-scenario large-scale fire target detection database was constructed to mark the categories and locations of the flame and smoke areas, and the problem of insufficient performance of YOLOv3 small target recognition was improved. Combined with the feature extraction ability of deep network, the fire detection and recognition were formalized into multi-class recognition and coordinate regression problems. The detection and recognition models of flame and smoke were obtained under different scenarios. Experiments show that the improved YOLOv3 algorithm proposed in this study can achieve ideal results for flame and smoke detection under different shooting angles and different illumination conditions, and also meets the real-time detection requirements in terms of detection speed.

**Key words:** deep learning; machine vision; YOLOv3; fire detection

## 1 引言

在各类灾害中, 火灾是对公共安全和社会发展最

常见和最普遍的威胁之一. 随着社会的不断发展, 在社会财富不断增加的同时, 火灾的危险性也在增加<sup>[1]</sup>.

① 基金项目: 国家自然科学基金 (61803339, 61503341); 浙江省自然科学基金 (LQ18F030011); 浙江省重点研发计划项目 (2019C03096)

Foundation item: National Natural Science Foundation of China (61803339, 61503341); Natural Science Foundation of Zhejiang Province (LQ18F030011); Key Research and Development Program of Zhejiang Province (2019C03096)

收稿时间: 2019-05-13; 修改时间: 2019-05-31; 采用时间: 2019-06-04; csa 在线出版时间: 2019-12-05

传统的火灾检测算法通过温度、烟雾传感器采集相关数据进行火灾概率预测,但是检测实时性得不到保证.而图像识别由于其具有直观,响应速度快等特点,越来越多地被应用于火灾监控和报警.

近年来,深度学习成为图像识别的主要手段.现有的基于深度学习的火灾检测,主要通过卷积神经网络对火灾图片进行特征提取,然后使用分类器进行分类.文献[2]提出了一种通过 AlexNet 实现早期火灾识别的方法,通过 AlexNet 进行图像特征提取,并通过全连接层和 Softmax 进行分类,但 AlexNet 参数量较大,在样本数据不足的情况下泛化性能较差.文献[3]提出了一种基于 MPCANet 的火灾图像识别模型,提高了火灾图像识别的泛化能力,但本质上还是通过特征提取跟分类器分类来解决静态火灾图像的分类问题.文献[4]提出了一种双流序列回归深度网络的视频火灾检测方法,通过融合静态特征和动态特征提升了火灾识别精度,但其工作重点依旧是火灾图像的识别分类,而未给出火灾目标检测的解决方案.

基于深度学习的火灾检测,在静态火灾图像识别领域表现出了强大的鉴别力,而在动态火灾目标检测方面的工作相对较少.现有火灾检测方法将火灾检测看成一个二分类问题,实现火灾和非火灾的分类.而本文根据不同的火灾特征,如火焰和烟雾,将火灾检测形式化为多分类和坐标回归问题.

本文构建了一个多场景大规模火灾目标检测数据库,并对火焰和烟雾区域进行类别和位置信息的标注.另外,本文提出了一种基于改进 YOLOv3 的火灾检测与识别方法,针对 YOLOv3 小目标检测能力不足的问题进行了改进,并在本文构建的火灾目标检测数据库上进行离线训练,从而获得相应的火灾检测模型.本文主要对火灾的火焰和烟雾特征进行识别和检测,验证了深度学习在火灾检测上的可行性.

## 2 算法理论

### 2.1 YOLOv3

目标检测是计算机视觉的实际应用中一个重要的领域,目标检测算法不仅需要对目标进行分类,还需要返回目标的坐标以实现定位<sup>[5-7]</sup>.YOLOv3 算法是目标检测领域的 SOTA 算法.与 Faster R-CNN 相比,YOLOv3 可以实现端到端的训练,减少了模型训练的复杂度,并且图像的处理速度大幅度提高<sup>[8]</sup>.而与 SSD<sup>[9]</sup>算法相比,YOLOv3 拥有更高的检测精度以及更快的处理速度.在 Pascal Titan X 的硬件条件下,YOLOv3 每秒可以处理 30 帧 COCO 数据集的图像,能满足火灾检测实时处理的需求<sup>[10]</sup>.

YOLOv3 算法的整体结构如图 1 所示,可将其分为 3 个部分,包括主干网络、多尺度特征提取器和输出层.

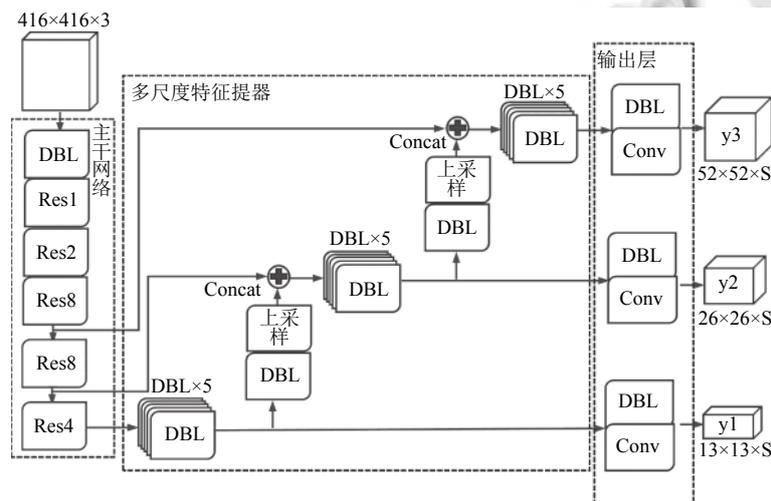


图 1 YOLOv3 网络结构

## 2.2 算法优化

### 2.2.1 改进的 YOLOv3 检测模型

虽然 YOLOv3 引入了特征金字塔 (FPN) 的思想,

并配以不同大小的锚点来改善小目标检测效果,但在实际测试过程中发现,YOLOv3 对于部分小目标的检测效果并不好,存在误检与漏检的情况.

YOLOv3 网络分别利用 8 倍、16 倍以及 32 倍降采样输出 3 种不同尺寸的特征图, 并且主要在 8 倍降采样的输出特征图上进行小目标检测. 这意味着 YOLOv3 在对小于  $8 \times 8$  的特征图进行检测时会出现困难<sup>[11]</sup>. 虽然 YOLOv3 对 16 倍以及 32 倍降采样的输出特征图进行上采样, 并与 8 倍降采样特征图进行融合以增强 8 倍降采样特征图, 但这一操作并不能提供更多小目标的特征信息.

如图 2, 本文对主干网络中输出 8 倍下采样特征图

进行 2 倍上采样, 使尺寸与主干网络的 4 倍下采样特征图匹配. 然后, 将经 2 倍上采样处理后的 8 倍降采样特征图与 4 倍降采样特征图进行 Concat 融合, 在得到  $104 \times 104$  大小的特征图上进行分类和坐标回归. 通过使用更大特征图, 可获得更丰富的上下文信息, 从而改善模型的小目标检测性能. 同时, 本文将最终得到的 8 倍和 16 倍下采样特征图分别 Maxpool 至与 32 倍下采样特征图同尺寸后进行 Concat 融合, 实现特征复用, 以增加 32 倍降采样特征图中的细节特征.

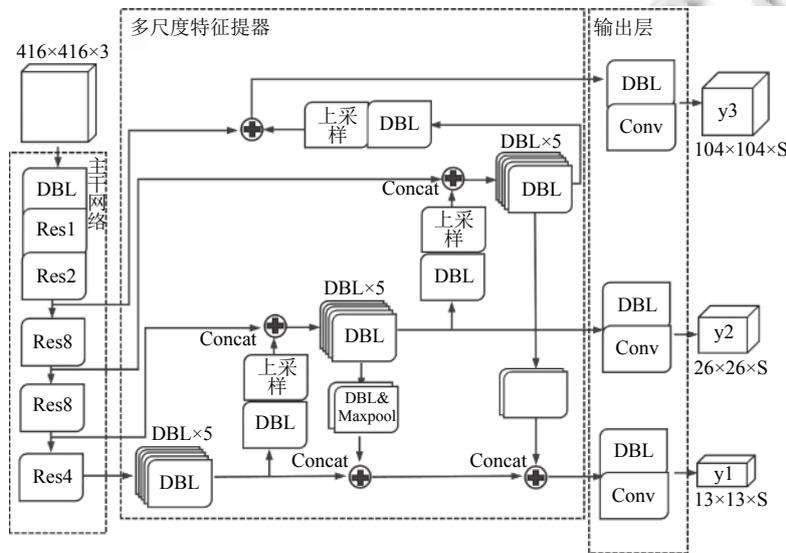


图 2 改进后的 YOLOv3 网络结构

### 2.2.2 K-means 聚类

YOLOv3 引入了 Faster R-CNN 中所使用的锚点 (anchor boxes) 的思想. 在 YOLOv3 输出的 3 个尺度的特征图中, 每组特征图都对应了通过 K-means 聚类计算得到的 3 个不同尺度的锚点. YOLOv3 的输出层通过预测不同锚点对应的偏移量来实现目标边界框的回归.

K-means 聚类是最常用的无监督学习算法之一, 其中的 K 代表类簇个数, means 代表类簇内数据对象的均值<sup>[12]</sup>. K-means 聚类算法是一种以距离作为数据对象间相似性度量的标准, 从而实现数据划分的聚类算法. 当数据对象之间的距离越小, 则代表数据之间的具有越高的相似度, 越有可能被划分在同一类簇中.

数据间距离的计算方式有很多种, 典型的 K-means 聚类通常采用欧氏距离来计算数据对象间的距离. 算法的详细流程如下:

#### 算法 1. K-means 聚类算法

1. 从数据中随机选取 K 个样本作为初始的 K 个质心;
2. 计算剩余所有样本与各个质心之间的欧式距离, 将各个样本划分至距离其最近的质心所在的类簇;
3. 重新计算每一个类簇的质心;
4. 如果所有的 K 个质心都没有发生变化则输出簇划分结果, 否则返回至第 2 步.

而从数据集的标注框聚类出先验的锚点, YOLOv3 并不使用常用的欧式距离来获得锚点, 而是采用使用 IOU 来计算标注框之间的距离, 如式 (1) 所示:

$$d(box, centroid) = 1 - IOU(box, centroid) \quad (1)$$

即 IOU 越大, 代表标注框之间的距离越小.

通过 K-means 算法在本文构建的火灾目标检测数据集中聚类得到 9 个锚点, 其中尺度 1 对应的锚点是 (171, 133), (143, 228), (329, 264), 尺度 2 对应的锚点是

(42, 80), (106, 73), (75, 126), 尺度 3 对应的锚点是 (19, 22), (30, 42), (59, 48).

YOLOv3 在 *K-means* 算法得到的锚点的基础上进行偏移量的回归, 从而得到目标的实际检测框. 针对 YOLOv3 小目标检测效果不佳的情况, 本文对不同的大小的质心计算距离时, 分别增加了对应的权重. 从而在 *K-means* 迭代的过程中, 约束其偏向于较小的锚点. 通过使用更小的锚点来改善 YOLOv3 的小目标检测能力. 具体如式 (2)、式 (3) 所示.

$$d(box, centroid)_{centroid \leq (0.3, 0.3)} = \lambda_1 [1 - IOU(box, centroid)] \quad (2)$$

$$d(box, centroid)_{centroid > (0.3, 0.3)} = \lambda_2 [1 - IOU(box, centroid)] \quad (3)$$

最终, 本文通过改进后的 *K-means* 聚类算法在构建的火灾目标检测数据库上得到 9 个锚点. 通过对比测试最终选定一组锚 (5, 6), (25, 10), (20, 34), (35, 79), (52, 43), (59, 115), (115, 90), (156, 197), (372, 326) 并分别分配给 3 种尺度的特征图. 最终聚类得到的 Bounding-box 类簇分布如图 3 所示.

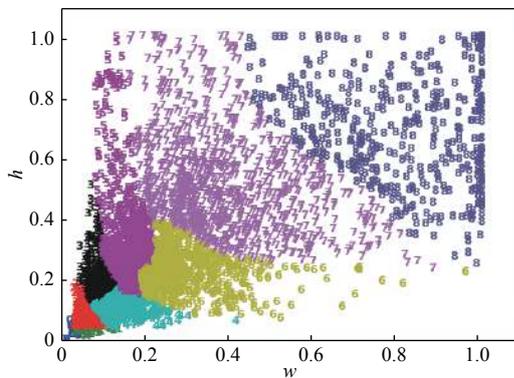


图 3 Bounding-box 类簇分布

### 2.3 迁移学习

迁移学习 (transfer learning) 是一种机器学习方法, 迁移学习可以把源域的知识迁移到目标域, 使得算法在目标域能够取得更好的学习效果<sup>[13]</sup>.

由于目前公开的火灾检测数据集较少, 而人工标注则需要耗费大量的人力物力, 因此通过使用迁移学习在有限的数据集上进一步提高模型的泛化能力十分必要. 本文将 YOLOv3 在 COCO 数据集上的预训练权重作为源域, 将构建的火灾检测数据库作为目标域. 使用预训练权重可以使本文的改进 YOLOv3 模型学习到

其他数据集中的“知识”, 从而增强算法的泛化性能. 其中, 使用迁移学习的步骤如下: 首先, 将预训练权重载入改进 YOLOv3 模型, 并冻结模型除了最后三层外的所有参数, 并以较低的学习率 ( $lr \leq 0.001$ ) 在本文的火灾目标检测数据库上进行训练. 较低的学习率可以最小化地调整先前学习的参数, 以防止迁移学习失效. 在训练过程中逐渐开放改进的 YOLOv3 其他层的参数, 以便对整个网络的参数实现微调.

## 3 实验

### 3.1 数据集制作

通过获取网络上的火灾图像和截取公开的火灾视频数据集实现训练样本的采集, 一共采集 1 万多张图片, 最后挑选出其中的 3835 张图片作为实际训练测试数据集, 如图 4 所示.



图 4 训练样本

利用公开的 labeling 标注系统对整个数据集中每张图像的火焰和烟雾区域进行人工标注, 标注系统界面如图 5 所示, 可以对火焰和烟雾区域进行类别选择和定位. 并将所有的样本图像集按照 PASAL VOC2007 样本集的格式处理并存储, 然后利用随机函数将数据集随机的分成训练集和测试集, 其中训练集和测试集的图片数量比例为 4:1.

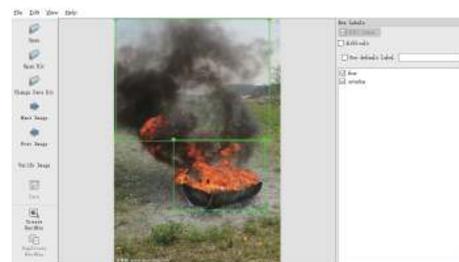


图 5 数据集标注

### 3.2 训练

本文实验过程均在以下的工作条件下完成训练与测试: IntelCore i5 8500 CPU, 配备 8 GB RAM 与 Win10 操作系统, NVIDIA GTX1070 配备 8 GB 显存, 以及 Keras 深度学习框架. 并配置 YOLOv3 的运行环境, 然后将标定完成后的样本数据集使用 YOLOv3 进行训练. 训练 Loss 曲线如图 6 所示, 可以看出, 网络迭代到 Epoch30 之后, 损失值趋于平缓.

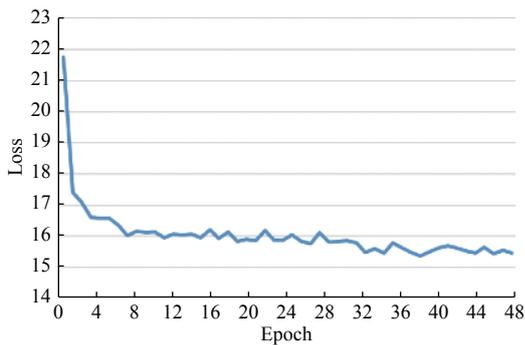


图 6 训练 Loss 曲线

### 3.3 评价指标

在本文的实验中, 均使用 mAP (mean Average Precision) 来评判模型识别的效果. mAP 是目标检测中衡量检测精度的常用指标, 指的是多个类别的平均准确率的平均值. AP 指的是单个类别的平均准确率, 衡量的是模型在某一类别上识别效果的好坏. 本文使用的 AP 默认为 AP<sub>50</sub>, 即预测框与 ground truth 的 IOU 在大于 50% 前提下的单类别的平均准确率. 而 mAP 即为所有 AP 的平均值. mAP 的值越大表示该模型的整体识别准确率越高<sup>[14]</sup>.

### 3.4 测试结果

本文提出的改进 YOLOv3 与 YOLOv3 的对比测试结果如表 1 所示.

表 1 对比测试结果 (%)

模型	YOLOv3	增强 K-means 的 YOLOv3	本文
mAP	79.34	80.24	87.07
火焰 AP	77.01	78.34	90.75
烟雾 AP	81.67	82.14	83.38

可以看出本文的改进 YOLOv3 在火焰和烟雾检测的平均精度上均有所提升, 尤其在火焰的检测精度上有大幅提升. 在各种场景的火灾图像中, 小尺寸的火焰目标较多, 而烟雾对象由于其自身的扩散特性, 在图像

中一般占据较大的尺寸. 本文选用了更大尺寸的特征图进行小目标检测, 并使用偏好小尺寸的 anchor 聚类生成策略来改善模型的小目标检测性能, 因此在火焰目标的检测性能上有较好的提升, 而对于本身尺寸较大的烟雾对象则提升有限, 总体上有效提高了模型的火灾检测精度.

最终训练得到的火灾检测模型在测试样本上的检测结果如图 7 所示, 可以看出检测模型在图像中存在小目标的情况下也能实现准确检测与识别, 整体检测效果较为理想.

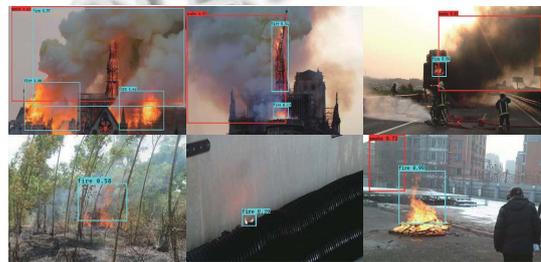


图 7 测试结果

## 4 结论与展望

针对目前火灾频发, 火灾检测实时性难以保证的情况, 本文提出了一种基于改进 YOLOv3 的火灾检测与识别方法. 实验结果表明, 本文使用的方法在火灾检测和识别上能够取得较好的检测效果, 尤其在小目标检测上有显著提升. 但是, 由于火焰和烟雾本身具有形状不固定的特点, 通过固定锚点实现坐标回归的方法难以获得较高精度的目标检测框. 因此, 在以后的研究中, 将重点解决火灾目标检测中锚点方法的改进, 以获得更高的检测精度.

### 参考文献

- 段锁林, 毛丹. 关于火灾火焰图像检测算法研究. 计算机仿真, 2016, 33(2): 393-398. [doi: 10.3969/j.issn.1006-9348.2016.02.082]
- Muhammad K, Ahmad J, Baik SW. Early fire detection using convolutional neural networks during surveillance for effective disaster management. Neurocomputing, 2018, 288: 30-42. [doi: 10.1016/j.neucom.2017.04.083]
- 张秀玲, 侯代标, 张逞逞, 等. 深度学习的 MPCANet 火灾图像识别模型设计. 红外与激光工程, 2018, 47(2): 0203006.
- 孔亚奇, 郎丛妍, 冯松鹤, 等. 双流序列回归深度网络的视

- 频火灾检测方法. 中国科技论文, 2017, 12(14): 1590–1595, 1663. [doi: 10.3969/j.issn.2095-2783.2017.14.005]
- 5 华夏, 王新晴, 王东, 等. 基于改进 SSD 的交通大场景多目标检测. 光学学报, 2018, 38(12): 1215003.
- 6 王文秀, 傅雨田, 董峰, 等. 基于深度卷积神经网络的红外船只目标检测方法. 光学学报, 2018, 38(7): 0712006.
- 7 罗海波, 徐凌云, 惠斌, 等. 基于深度学习的目标跟踪方法研究现状与展望. 红外与激光工程, 2017, 46(5): 0502002.
- 8 高露, 马元婧. 基于 Faster R-CNN 的设备故障检测与识别. 计算机系统应用, 2019, 28(4): 170–175. [doi: 10.15888/j.cnki.csa.006877]
- 9 Liu W, Anguelov D, Erhan D, *et al.* Ssd: Single shot multibox detector. Proceedings of the 14th European Conference on Computer Vision. Amsterdam, The Netherlands. 2016. 21–37.
- 10 Redmon J, Farhadi A. YOLOv3: An incremental improvement. arXiv: 1804.02767, 2018.
- 11 鞠默然, 罗海波, 王仲博, 等. 改进的 YOLO V3 算法及其在小目标检测中的应用. 光学学报, 2019, 39(7): 0715004.
- 12 蒋华, 季丰, 王慧娇, 等. 改进 Kmeans 算法的海洋数据异常检测. 计算机工程与设计, 2018, 39(10): 3132–3136.
- 13 袁功霖, 侯静, 尹奎英. 基于迁移学习与图像增强的夜间航拍车辆识别方法. 计算机辅助设计与图形学学报, 2019, 31(3): 467–473.
- 14 冯小雨, 梅卫, 胡大帅. 基于改进 Faster R-CNN 的空中目标检测. 光学学报, 2018, 38(6): 0615004.