

基于YOLO模型的火焰图像检测

太原科技大学 李钰冰 冯泽文 王 强 崔志浩 许 奇

DOI:10.19353/j.cnki.dzsj.2020.12.013

由于YOLOv1、YOLOv2在小目标的检测上表现不佳, YOLOv3模型中引入特征金字塔网络(FPN)以及残差网络(RS-net)后, 检测性能得到很大的提升, 本文采用YOLOv3模型实现对不同环境下的火焰图像进行检测。首先建立了火焰图像数据集, 并对其进行标注, 其次用标注好的图片在linux操作平台下对YOLOv3模型进行训练, 得到合适的加权系数, 使模型损失函数值最小, 最后将训练好的模型对实际环境下火焰进行检测。实验结果显示, 不同场景下火焰检测的准确率和实时性都比较好。

火灾对于社会经济以及人的生命安全隐患巨大, 一点小小的火星都可能导致严重的火灾, 因此我们需要密切监控火灾可能发生的区域, 并且配有烟雾报警器等辅助设备提高准确率, 把火灾发生的概率降到最小, 因此对于火灾的实时监控以及有效预测得到了越来越多的研究, 这些年在这一领域也取得了许多的进展。

熊国良(熊国良, 苏兆熙, 刘举平, 等. 火焰特性识别的Matlab实现方法)等根据火焰燃烧时的特征, 选取火焰的圆弧度、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(产云洋, 吴茜茜, 杜静, 等. 基于色彩和闪频特征的视频火焰检测)等根据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(胡勤, 陈琛, 刘敏. 一种基于动态纹理的烟雾和火焰检测方法)等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。Stadler(STADLER A, WINMSCH T, DIEPOLD K. Comparison of intensity flickering features for video based flame detection algorithm)等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。Yamada(Yamada M, Shiratori S S. Smoke sensor using mass controlled layer-by-layer self-assembly of polyelectrolytes films)等提出基于逐层自组装电解质薄膜的烟雾传感器来进行火灾提前预警。Cheon(Cheon J, Lee J, Lee I, et al. A Single-Chip CMOS Smoke and Temperature Sensor for an Intelligent Fire Detector)等将温度传感器与烟雾传感器结合进行火灾提前预警。Hart利用RGB中R分量的饱和度, 设置阈值来判断疑似火焰区域。静态特征能够较好地捕捉到火焰的边缘、颜色空间等信息。

本文主要论述了YOLO算法的原理与流程, 并在此基础上搭建合适的神经网络模型, 用标注好的火焰图像数据集训练和优化YOLO模型, 并对不同类型的火焰图像进行检测, 得出实验结果并对其进行分析。

1 建立数据集

要采用深度学习的方法对火焰图像进行研究, 就需要建立较为全面的、体系完善的火焰数据集。本文通过人工拍摄、互联网下

载和视频截帧的方式对火焰原始图片进行采集。初步采集了约有12000张火焰图片, 从中选取了火焰特征明显、类型比较全面的约1700张火焰图片做为本次实验的数据集。由于火焰种类众多, 将数据集分为四类子集, 分别是室内火焰、室外火焰、森林火焰、其他类别。通过labelimg对图片完成标记分类, 如果火焰形状较为规则可以直接用一个矩形框完成标注, 但是对于多尖火焰, 为了减少不必要的计算量可以使用多个矩形框分别标注。

2 基于YOLOv3模型的火焰图像检测算法

YOLO算法的核心是不再需要区域候选网络, 跳过分类器完成物体检测, 直接使用神经网络通过全局信息预测包络框和类别, 以达到端到端的目标检测。YOLO检测容易区分主体与背景, 易于对端到端进行优化。

2.1 算法流程

相比较于YOLOv1和YOLOv2, 改进后的YOLOv3网络模型对于图片对象的检测和识别更加快速和精确。检测过程如下:

(1) 首先要对输入的图像进行统一的缩放处理, 使其大小全部变为 416×416 , 即32的整数倍, 这是基于后续处理的需要。

(2) 接下来将经过缩放处理后的输入图像分割为个网格, 通过新的Darknet-53网络架构提取特征后, 提取出3个尺度各不相同的特征图, 长度规律为13, 26, 52。

(3) 上述步骤中, Darknet-53已经将提取出的特征图缩小到了相对于输入图像的 $1/32$, 每一种尺度的特征图利用3个对待测物体做出预测, 详细来说, 对数据集使用k均值聚类算法帮助确定了anchor box的边框尺寸。

(4) 由于YOLOv3网络架构中出现了多种尺度的特征图, 进而产生了更多大小不同的预测anchor box, 故要想更精确的预测目标, 需要完成跨尺度的预测, 即将深、浅两个层次的特征完成融合。

(5) 对于最终的检测和分类, YOLOv3网络架构使用了二值交叉熵以及逻辑回归的方法, 以完成端到端的训练。YOLOv3的结构如图1所示。

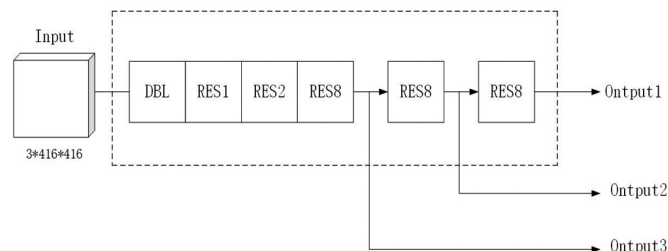


图1 YOLOv3网络架构图

2.2 算法原理

YOLOv1网络对于首先将输入图像分割, 随后对每个网格设置B个包络框, 其中每一个包络框全部都要预测5个参数值: $x, y,$

w, h, c 。待检测物体中心位置相对于网格位置的偏移程度用 x 和 y 表示。 w 和 h 则为待测物体相对于图像的宽、高比例。 c 表示置信度，它的值反映了每一个网格在有无包含待检测物体的条件下，其包络框位置准确性的信息，计算 c 如公式(1)所示：

$$confidence = P_r(Object) * IOU_{pred}^{truth} \quad (1)$$

若有目标落入其中某一个网格中, $P_r(Object) = 1$, 否则为0。随后要预测边界框与真实物体框之间的重叠度 (IOU)。接下来对每一个单元格做出类别信息C的预测, 这是由网格本身决定的, 预测边界框筛选的衡量标准如公式(2)所示:

$$P_r(Class_i/Object) * P_r(Object) * IOU_{pred}^{truth} = P_r(Class_i) * IOU_{pred}^{truth} \quad (2)$$

式(2)的结果就是每个网格的分类置信度得分, 具体来说, 这个等式包含了待预测的类别信息以及包络框准确度信息。完成所有网格的得分计算之后, 需要设置一个阈值用于筛选。得分低的框将被剔除, 剩余保留框经过NMS处理以得出最终实验结果。

YOLOv1的网络结构具体的包括了24个卷积层以及2个全连接层，在完成B个包围框以及C个类别的检测之后，最终得出了 $S \times S \times (5 \times B + C)$ 维的张量。总体来说，YOLOv1网络检测的效果一般，它的网络结构如图2所示。

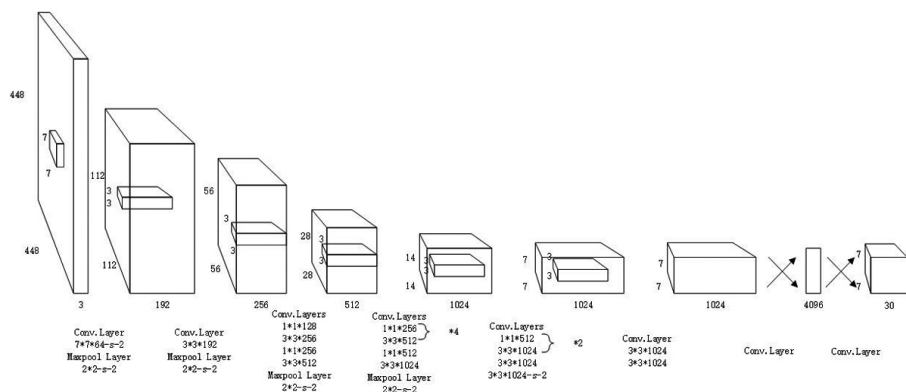


图2 YOLOv1网络结构图

YOLOv2的改进结构是Darknet-19（杨俊.基于卷积神经网络的目标检测研究）网络框架，它拥有19个数量的Convolutional层，Bathch Normalization（王宇，杨艺，王宝山，王田，卜旭辉，王传云.深度残差神经网络高分辨率遥感图像建筑物分割）层被放到了每一个Convolutional层的位置之后，这就加快了收敛速度。同时5个Maxpool层也可以降低网络的过拟合。

YOLOv3使用新的网络架构darknet-53(徐国标, 侯明利, 熊辉.基于YOLOv3改进算法的远程塔台运动目标检测)来完成改进,达到提取特征的目的。如图3所示结构图,新的网络架构中池化层以及全连接层被取消,而且卷积核步长增加,并采用了残差的设计思想。新的网络几乎比ResNet-101(J.Sang, P.Guo, Z.Xiang, 等.Vehicle detection based on faster-RCNN)的性能增加了50%,与ResNet-152(Alistarh, Dan, Grubic, Demjan, 等.QSGD:Communication-Efficient SGD via Gradient Quantization and Encoding)的效果也相比也毫不逊色。

2.3 实验结果与分析

首先安装linux系统配置训练环境，在该系统下安装Open cv和CUDA，安装完毕后测试安装是否正确并及时修改。打开命令终端，下载

yolo源代码文件后重新编译文件。在一边训练一边测试的情况下,最终选择迭代340000次之后的权重文件,进而得到了YOLOv3火焰检测模型。



图3 室内火焰检测图



图4 室外火焰检测图



图5 森林火焰检测图



图6 其他类别火焰检测图

由图3-图6可知，基于YOLO算法的火焰图片检测模型在训练样本充足，迭代次数足够的基础上，能够得到非常不错效果。室内和森林火焰的识别效果较好，室外其他类别火焰识别的效果一般造成的原因可能是参与训练的图片标记细节不够明显，室外的背景过于复杂，

室内和森林的背景单一；受光照的作用，或是火焰的颜色不同也会对检测结果造成影响。

3 结论

本文详细论述了YOLO算法的原理和网络结构模型训练，在linux系统下用火焰数据集训练YOLOv3模型。采集火焰图片，从中筛选出火焰特征明显和不同背景下的火焰图片作为数据集，对图片标记后与标签整合到一起。在训练的过程中，将测试集的图片分批送入网络，经卷积、池化等等前向传播后再反向传播更新权重系数，多次迭代得到最终的神经网络模型。需要注意的是，迭代次数不应过多，迭代过了会使结果过于拟合，能识别的图片反而无法正常识别。

山西省大学生创新创业训练项目“多模特征融合的火焰烟雾视频监控系统设计”（项目编号：2019358）。

作者简介:

李钰冰(1998—)，男，汉族，山西晋中人，大学本科，现就读于太原科技大学电子信息工程学院。

冯泽文(1998—)，男，汉族，山西长治人，大学本科，现就读于太原科技大学电子信息工程学院。

王强(1997—)，男，汉族，天津武清人，大学本科，现就读于太原科技大学电子信息工程学院。

崔志浩（1997—），男，汉族，山西运城人，大学本科，现就读于太原科技大学华科学院。

许奇(1998—)，男，汉族，山西大同人，大学本科，现就读于太原科技大学电子信息工程学院。