

# 基于动态形状特征提取及增强的改进 YOLOv3 火焰检测算法

丁浩, 王慧琴\*, 王可

西安建筑科技大学信息与控制工程学院, 陕西 西安 710055

**摘要** 针对现有多目标检测网络对动态火焰特征提取及增强能力不足, 检测效果不佳的问题, 提出基于动态形状特征提取及增强的改进 YOLOv3 火焰检测算法。采用小尺寸结构的 ResNet50\_vd 作为 YOLOv3 的主干网络, 减少特征信息冗余; 在主干网络 stage 4 和 stage 5 中加入可变形卷积模块, 控制采样网格随火焰目标形状的动态变化; 引入交并比 (IoU) Aware 模块, 增加置信度得分与 IoU 定位精度的相关性, 提高网络的火焰特征提取能力; 同时在 YOLOv3 Head 中加入 Drop Block, 引入 IoU 预测分量优化损失函数, 提高模型学习过程中的特征增强能力。通过消融实验验证各改进部分对模型的提升效果, 实验结果表明, 改进模型对火焰的检测精度达 94.11%, 推理速度达 73.52 frame/s, 能够有效满足对动态形状火焰的检测要求。

**关键词** 火焰检测; 动态形状; ResNet50\_vd; 卷积神经网络

中图分类号 TP391.41

文献标志码 A

DOI: 10.3788/LOP202259.2410003

## Improved YOLOv3 Flame Detection Algorithm Based on Dynamic Shape Feature Extraction and Enhancement

Ding Hao, Wang Huiqin\*, Wang Ke

College of Information and Control Engineering, Xi'an University of Architecture and Technology,  
Xi'an 710055, Shaanxi, China

**Abstract** To address the issue that the existing multitarget detection network cannot extract and enhance dynamic flame features, resulting in poor detection results, this paper presents an improved YOLOv3 flame detection algorithm based on dynamic shape feature extraction and enhancement. ResNet50\_vd with a small size structure is used as the backbone network of YOLOv3 to reduce the redundancy of feature information. To control the dynamic change of the sampling grid with the shape of the flame target, deformable convolutional neural network modules are added to the backbone network stage 4 and stage 5. The IoU Aware module is introduced to increase the correlation between the confidence score and the positioning accuracy of the IoU, and to enhance the flame feature extraction ability of the network. Simultaneously, the Drop Block module is added to the YOLOv3 Head, and the IoU prediction component is introduced to optimize the loss function, which improves the feature enhancement ability during the model learning process. The ablation experiments were performed to verify the effect of each improvement on the proposed model. The results show that the improved model for flame detection has a detection accuracy of 94.11% and an inference speed of 73.52 frame/s, which can effectively meet the detection requirements of dynamic shape flames.

**Key words** flame detection; dynamic shape; ResNet50\_vd; convolutional neural network

## 1 引言

火灾一直是威胁人们生命和财产安全的常见隐患之一。对于生产车间、储存仓库等大空间场所, 火灾发

生所造成的后果极其严重。快速且准确地检测和预警火灾, 有助于及时发现火情, 减少火灾对生命和财产所造成的损失。

传统火灾报警传感器一般是基于感知环境中烟尘

收稿日期: 2021-09-09; 修回日期: 2021-10-18; 录用日期: 2021-10-27

基金项目: 陕西省自然科学基金基础研究计划(2021JM-377)、陕西省科技厅科技合作项目(2020KW-012)、陕西省教育厅智库项目(18JT006)、西安市科技局高校人才服务企业项目(GXYD10.1)

通信作者: \*hqwang@xauat.edu.cn

浓度、温度等物理量变化的,当变化量达到设定阈值时,报警器进行报警,具有一定的时间滞后性。传统基于机器学习的火焰检测主要利用人工方法提取火焰特征,利用机器学习分类器进行火焰检测。荣建忠等<sup>[1]</sup>引入并改进了统计地形特征的纹理描述方法,该方法基于神经网络算法,融合多个火焰特征快速识别火焰。李巨虎等<sup>[2]</sup>提出了分块的局部二值模式(LBP)直方图特征结合局部相位量化(LPQ)直方图特征的火焰识别算法。Borges等<sup>[3]</sup>首先依据颜色判据对图像进行预判,然后计算统计疑似区域的火焰面积、边界粗糙度等特征,最终利用贝叶斯分类器实现决策。传统火焰检测算法受火焰颜色、纹理等人为规定特征的影响,对火焰动态特征的提取和增强能力不足,在环境背景复杂、多变时,易出现漏检和误检的情况。

基于深度学习的火焰图像检测方法近年来发展迅速,一般由多层卷积神经网络自动提取图像中火焰的深度特征,并利用分类器或多尺度检测模块进行分类和定位。赵飞扬等<sup>[4]</sup>提出了一种基于改进YOLOv3算法的火焰检测算法,在特征融合过程中提出同尺度特征密集相连结构和空间金字塔卷积运算,降低浅层位置特征信息融合不足导致的漏检率。熊卫华等<sup>[5]</sup>提出了一种基于混合卷积神经网络(CNN)的火灾识别方法,通过其中一路特征提取器提取图像中的深层语义信息,另一路特征提取器提取图像的浅层上下文信息,实现对多尺度特征的融合。严云洋等<sup>[6]</sup>利用候选区域提取网络提取火焰特征,结合后续的池化操作和分类、回归分支进行火焰检测,构建新的火焰检测网络。罗小权等<sup>[7]</sup>改进特征提取网络和多尺度检测模块,再通过聚类算法生成全新的火焰目标锚框,有效提升火焰目标检测精度。喻丽春等<sup>[8]</sup>提出一种基于Mask R-CNN改进的火焰目标识别算法,该算法通过改进特征金字塔结构有效提取火焰特征,再利用新的损失函数优化网络训练过程,使得火焰检测更加准确。

上述基于深度学习的火焰检测算法没有充分考虑火焰动态形状特征与传统目标特征的区别,对火焰动态形状特征的提取和增强能力存在一定不足,火焰检测精度仍需提高。针对此问题,本文提出一种基于火焰动态形状提取及增强的火焰检测算法。采用小尺寸的ResNet50\_vd网络替换DarkNet53,再将可变形卷积网络(DCN)加入ResNet50\_vd的stage 4和stage 5中,构建ResNet50\_vd-DCN作为模型的主干网络,以更好地学习火焰的动态形状。在网络Head中加入Drop Block模块,提高特征增强能力;在置信度得分中加入IoU Aware,优化非极大值抑制(NMS)中置信度得分排序的可靠性,提高模型定位精度;在损失函数中加入交并比(IoU)预测分量,优化模型参数。

## 2 YOLOv3算法基本原理

YOLOv3算法是经典的一阶段端到端目标检测

算法<sup>[9-11]</sup>,通过输入固定尺寸的图片,将整张图片划分为多个相同尺寸的网格,利用“框海”战术,再通过回归得到锚框的中心偏移量、宽高拉伸量,从而得到预测框的位置和其所属的类别,实现端到端目标检测。二阶段算法检测速度较为缓慢,不符合检测火焰及时性的基本要求,而YOLOv3算法的响应速度较快,满足火焰快速、及时的检测要求。YOLOv3在应对主干网络深度过大从而产生梯度弥散问题时,吸收残差网络的思想,在DarkNet53中加入残差模块,增强了网络的传播性能,更好地表征了图像的信息。YOLOv3在应对目标尺度多变从而导致识别困难问题时,首先采用多尺度检测方法,通过32倍、16倍、8倍降采样,输出三个不同尺度的特征图,它们分别对应检测的大、中、小物体;其次借鉴特征金字塔网络的思想,对三种降采样的特征图进行特征融合,增强语义信息,从而提升了网络对不同尺寸目标的检测精度。

YOLOv3的损失函数由位置损失、置信度损失、类别损失三部分构成,表达式为

$$L_{\text{total}} = L_{\text{loc}} + L_{\text{conf}} + L_{\text{class}}, \quad (1)$$

其中,位置损失采用均方差损失函数,表达式为

$$L_{\text{loc}} = \lambda_{\text{coord}} \sum_{i=0}^{S^2} \sum_{j=0}^B I_{ij}^{\text{obj}} \left[ (x_i - \hat{x}_i^j)^2 + (y_i - \hat{y}_i^j)^2 \right] + \lambda_{\text{coord}} \sum_{i=0}^{S^2} \sum_{j=0}^B I_{ij}^{\text{obj}} \left[ (\sqrt{w_i^j} - \sqrt{\hat{w}_i^j})^2 + (\sqrt{h_i^j} - \sqrt{\hat{h}_i^j})^2 \right], \quad (2)$$

置信度损失和分类损失均采用二分类交叉熵损失函数,表达式分别为

$$L_{\text{conf}} = - \sum_{i=0}^{S^2} \sum_{j=0}^B I_{ij}^{\text{obj}} \left[ \hat{C}_i^j \log C_i^j + (1 - \hat{C}_i^j) \log (1 - C_i^j) \right] - \lambda_{\text{noobj}} \sum_{i=0}^{S^2} \sum_{j=0}^B I_{ij}^{\text{noobj}} \left[ \hat{C}_i^j \log C_i^j + (1 - \hat{C}_i^j) \log (1 - C_i^j) \right], \quad (3)$$

$$L_{\text{class}} = - \sum_{i=0}^{S^2} I_{ij}^{\text{obj}} \sum_{c \in c_{\text{classes}}} \left[ \hat{P}_i^j \log P_i^j + (1 - \hat{P}_i^j) \log (1 - P_i^j) \right], \quad (4)$$

式中: $\lambda_{\text{coord}}$ 和 $\lambda_{\text{noobj}}$ 表示权重系数; $x_i$ 和 $y_i$ 、 $w_i^j$ 和 $h_i^j$ 、 $C_i^j$ 和 $P_i^j$ 分别表示目标坐标预测值、置信度预测值、类别预测值; $\hat{x}_i^j$ 和 $\hat{y}_i^j$ 、 $\hat{w}_i^j$ 和 $\hat{h}_i^j$ 、 $\hat{C}_i^j$ 和 $\hat{P}_i^j$ 分别表示目标坐标真实值、置信度真实值、类别真实值; $I_{ij}^{\text{obj}}$ 表示第*i*个网格的第*j*个锚框是否负责该目标的预测,若负责, $I_{ij}^{\text{obj}}$ 为1,否则 $I_{ij}^{\text{obj}}$ 为0; $I_{ij}^{\text{noobj}}$ 表示第*i*个网格的第*j*个锚框是否不负责该目标的预测,若不负责, $I_{ij}^{\text{noobj}}$ 为1,否则 $I_{ij}^{\text{noobj}}$ 为0。

火焰检测任务与多目标检测任务有着明显的不同,将用于多目标检测任务的YOLOv3网络直接应用于火焰检测任务时会产生以下问题:没有重点考虑火焰目标与传统多目标特征的区别,即火焰动态形状的充分提取和增强问题,容易导致检测效果不佳;多目标模型参数数量过大,容易产生数据冗余,且过大的模型参数需要更多的储存空间,不利于网络模型的现场部署。

### 3 基于动态形状特征提取及增强的改进 YOLOv3 火焰检测算法

在文献[11]的基础上设计检测动态形状特征火焰目标的特定网络模型,具体改进如下。

1) 采用小尺寸结构的 ResNet50\_vd 作为主干网络,避免信息冗余;在 ResNet50\_vd 残差网络的 stage 4 和 stage 5 中加入可变形卷积,代替原有的 3×3 卷积,使采样网格的位置随火焰动态形状变化而变化,增强主干网络对火焰特征的提取能力。提出 ResNet50\_vd-DCN 模型,将其作为火焰检测算法的主干网络。

2) 训练时,在 YOLO Head 中的 1×1 卷积前加入 Drop Block 模块,提高模型对残缺火焰特征的增强能力,优化火焰目标的检测效果。

3) 将 IoU Aware 加入模型,即在火焰置信度得分

中加入 IoU 预测分量,考虑定位精度,减少准确的检测框置信度得分不高,反而被 NMS 算法抑制的情况,加强模型对火焰目标特征的代表能力,提高特征提取能力。

4) 在  $L_{total}$  中加入  $L_{IoU}$ ,优化 IoU 预测,提高模型检测精度,表达式为

$$L_{total} = L_{loc} + L_{conf} + L_{class} + \lambda_{IoU} L_{IoU}, \quad (5)$$

式中: $\lambda_{IoU}$  为权重参数。

所提基于动态形状特征提取及增强的改进 YOLOv3 火焰检测算法的网络结构如图 1 所示,其中 Conv\_block\_DCIN 和 Identity\_block\_DCIN 模块中省去了 BN 和 ReLU 层。基本参数中加了 IoU 预测值,由 5 个变为 6 个,网络输入图像的尺寸大小为 3×608×608,输出图像  $y_1$ 、 $y_2$ 、 $y_3$  的尺寸大小分别为 19×19、38×38、76×76,分别对应大、中、小三种尺寸的火焰,从而提高算法对火焰目标的检测精度。

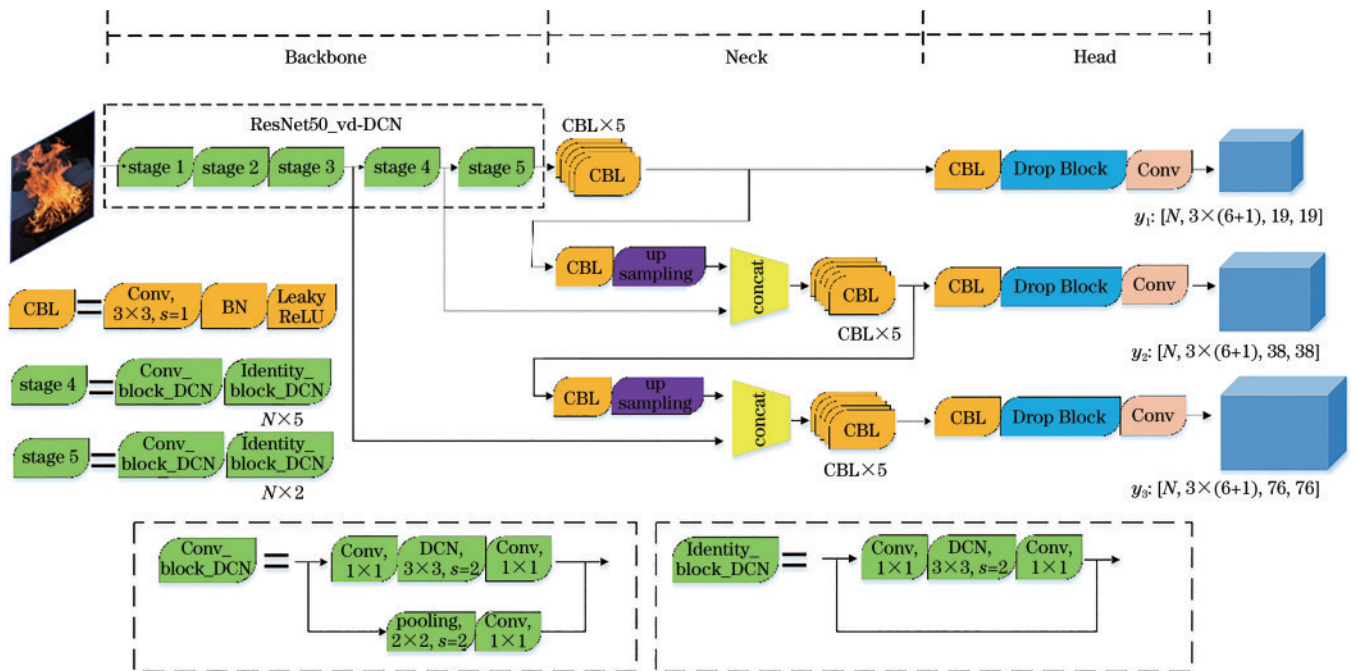


图 1 火焰检测算法的网络结构

Fig. 1 Network structure of flame detection algorithm

#### 3.1 ResNet50\_vd-DCN 主干网络设计

将 ResNet50\_vd 和 DCN 相结合,构建 ResNet50\_vd-DCN 作为改进模型的主干网络,将 ResNet50\_vd 中 stage 4、stage 5 结构内 3×3 的标准卷积替换成 3×3 的可变形卷积,如图 1 中 Conv\_block\_DCIN 和 Identity\_block\_DCIN 所示。保持 ResNet50\_vd 原有的小尺寸结构,同时可变形卷积能够增强网络对火焰动态特征的提取能力,提高模型识别精度。

##### 3.1.1 小尺寸主干网络引入

多目标检测模型的主干网络的层数较深,以便于学习不同目标的深层特征,增强模型的代表能力,但对于单目标火焰检测,模型参数过多,容易产生信息冗余,所以选取小尺寸结构的 ResNet50\_vd<sup>[12]</sup>作为主干

网络。ResNet50\_vd 包括 5 个 stage 结构:stage 1 由 1 个卷积层和 1 个最大池化层组成;stage 2 包括 1 个 Conv\_block 和 2 个 Identity\_block;stage 3 包括 1 个 Conv\_block 和 3 个 Identity\_block;stage 4 包括 1 个 Conv\_block 和 5 个 Identity\_block;stage 5 包括 1 个 Conv\_block 和 2 个 Identity\_block。ResNet50\_vd 在 ResNet50 基础上对 Conv\_block 进行了改进,如图 2 所示,左侧分支卷积结构不变,将步长为 2 的 1×1 卷积转换为 3×3 卷积,右侧分支加入步长为 2 的平均池化层进行降采样,整体改进减少了卷积操作过程中图像信息的丢失,保证感受野的有效感知范围。

##### 3.1.2 可变形卷积替换

基于 CNN 的深度学习算法在刚性物体的目标识



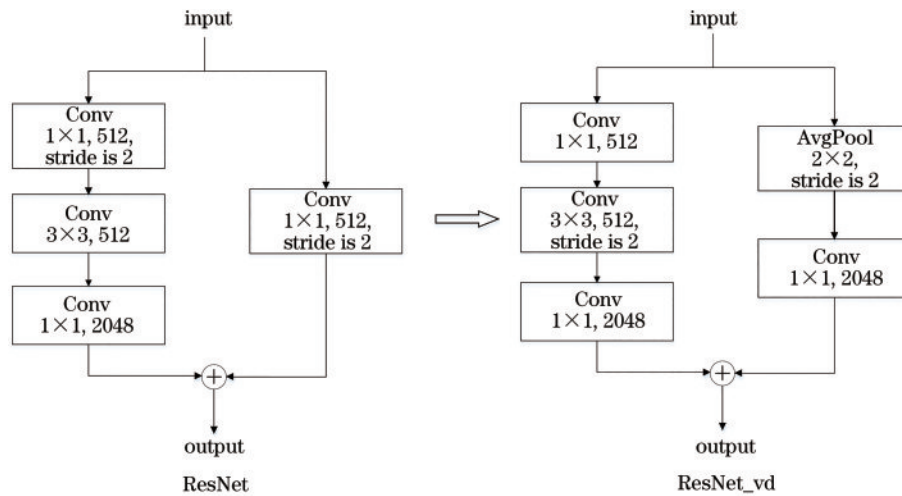


图2 Conv\_block改进前后结构对比

Fig. 2 Conv\_block structure comparison before and after improvement

别任务中已取得了优异的成绩,但由于CNN固定的几何结构(一个卷积单元在固定位置对输入特征图进行采样,池化层以固定比例下采样),缺少处理几何变换的内部机制,所以CNN在处理存在未知几何变换的任务时,这些变换没有被适当建模。同时,对于过于复杂的变换,手工设计算法往往过于困难或不可完成,这对火焰的动态形状特征检测显然存在一定的弊端,在算法预测时,无法有效地对火焰动态形状的各类形变进行建模。

文献[13-14]用DCN代替标准卷积网络。可变形卷积将一个二维偏移量加到标准卷积中的常规网格采样位置,采样网格可以进行位置变换,卷积核的感受野不再固定,可进行变换,但这变换可能超出目标物体的区域,导致提取的特征受到不相关背景的影响。因此,引入一种扩展变形建模范围的调节机制,增强可变形卷积的有效建模能力。文献[15-16]在主干网络中加入可变形卷积,通过可变形卷积中采样网格的位置能够对图像内容特征进行动态调整,精确地检测目标位置,文献实验结果中检测精度皆有显著提高,验证了可变形卷积对动态特征提取的有效性。3×3标准卷积与可变形卷积对比如图3所示。图3(a)为标准卷积(深色圆点),图3(b)为可变形卷积,箭头为偏移量,浅

色圆点为变形采样位置,图3(c)和图3(d)为可变形卷积特例,改变宽高比例和进行旋转变换。将标准3×3卷积替换成可变形3×3卷积,利用可变形卷积采样网格位置随火焰形状变化而变化的特性,增强网络提取火焰动态形状特征的能力。

卷积核为标准卷积时,输出特征图的每个位置为

$$y(p_0) = \sum_{p_n \in R} w(p_n) \cdot x(p_0 + p_n), \quad (6)$$

式中: $w(p_n)$ 为网络学习的权重,即卷积核; $x(p_0 + p_n)$ 为 $p_0 + p_n$ 的像素值; $R$ 为卷积核; $p_n$ 为 $R$ 中除去 $p_0$ 的采样位置。

卷积核为可变形卷积时,输出特征图的每个位置为

$$y(p_0) = \sum_{p_n \in R} w(p_n) \cdot x(p_0 + p_n + \Delta p_n) \cdot \Delta m_n, \quad (7)$$

式中: $\Delta m_n$ 为调节机制标量,且 $\Delta m_n \in [0, 1]$ ;  $\Delta p_n$ 为增加的二维偏移量,且 $\{\Delta p_n | n = 1, \dots, N\}, N = |R|, \Delta p_n$ 一般为小数。计算 $x(p_0 + p_n + \Delta p_n)$ 的双线性插值公式为

$$x(p) = \sum_q G(q, p) \cdot x(q), \quad (8)$$

式中: $q$ 为特征图中所有采样位置; $G(q, p)$ 为格林函数的二阶离散拉普拉斯算子。 $G(q, p)$ 可转换为两个一

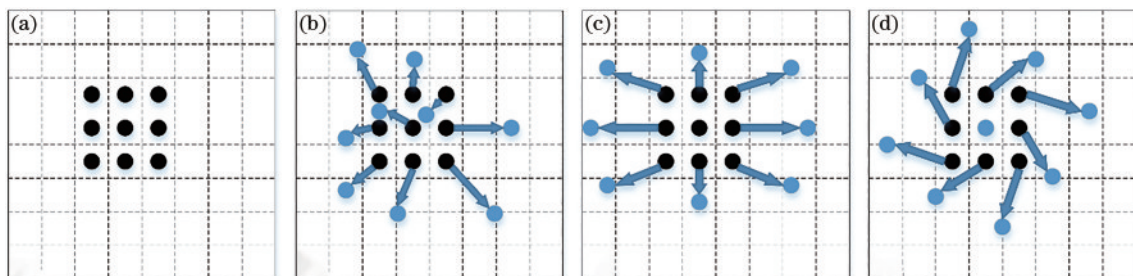


图3 3×3标准卷积与可变形卷积对比。(a)标准卷积;(b)~(d)可变形卷积

Fig. 3 Comparison between 3×3 standard convolution and deformable convolution. (a) Standard convolution; (b)~(d) deformable convolution

阶微分算子,表达式为

$$G(q, p) = g(q_x, p_x) \cdot g(q_y, p_y), \quad (9)$$

式中:  $g(a, b) = \max(0, 1 - |a - b|)$ 。

标准卷积固定感受野和可变形卷积自适应感受野图解如图 4 所示。底层表示在火焰特征图上进行  $3 \times 3$  卷积采样处理,高亮的蓝色部分对应上面的高亮部分;

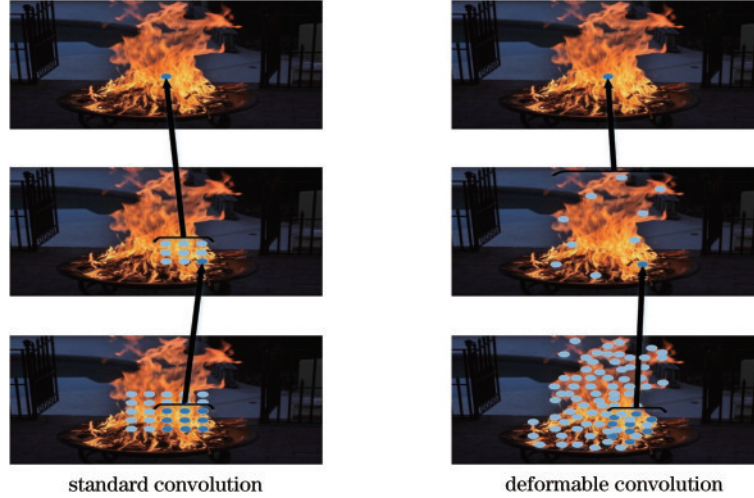


图 4 标准卷积固定感受野和可变形卷积自适应感受野图解

Fig. 4 Diagram of standard convolution fixed receptive field and deformable convolution adaptive receptive field

### 3.2 网络 Head 改进

为了更好地提高模型检测效果,改进了 YOLOv3 Head。在激活函数之后加入了 Drop Block 模块对火焰特征进行增强,同时加入 IoU Aware 模块,优化基于置信度得分的预测框排序操作,提高网络对火焰特征的提取和表征能力。

#### 3.2.1 增强火焰特征

由于火焰动态特征的随机和不固定性,为了应用于多变的现实场景,在训练过程中需要对其特征进行增强,提高模型检测精度,在网络 Head 中加入了 Drop Block 模块,对残缺火焰特征进行增强,如图 1 中改进 Head 结构所示。

Drop Out 通常是用于全连接层的正则化技术,但对卷积层的效果有限。文献[17]提出了一种模块化的具有丢弃形式的 Drop Block 正则化技术,该技术将输出特征图中相邻区域的单元一起丢弃,有效避免 Drop Out 点状丢弃时相邻单元依然保留部分该位置的语义信息的情况。文献[18]在网络中加入 Drop Block 模块,通过 Drop Block 模块丢弃特征图中块状的连续区域,使网络学习到鲁棒性更好的特征,增强了网络的特征提取能力,该实验结果中平均精度得到了有效的提高,验证了 Drop Block 模块对增强网络特征提取的能力。这种特征图中丢弃连续区域而不是随机丢弃单元的形式能够有效模拟现实中火焰被树木、车辆、桌椅等物体遮挡的情况,增强没有被遮挡的火焰特征,优化火焰残缺时的检查效果,增强模型的鲁棒性。

中层表示标准卷积和可变形卷积采样位置的变化;顶层表示经过卷积变换后两个卷积的最终位置。由图 4 可知,相比于标准卷积固定采样模式只能学习火焰目标的权重,可变形卷积既可以学习火焰目标的权重,也能学习火焰目标的动态形状,提高了模型的检测精度。

Drop Block 含有两个主要参数:  $block\_size(s_{block})$  和  $\gamma$ 。  $block\_size$  是丢弃方块的边长大小,当  $s_{block} = 5$  时,丢弃方块的形状大小为  $5 \times 5$ ;  $\gamma$  表示要删除的激活单元个数,控制要进行 Drop Block 操作的特征图数量。  $\gamma$  的计算公式为

$$\gamma = \frac{1 - p_{keep}}{s_{block}^2} \times \frac{s_{feat}^2}{(s_{feat} - s_{block} + 1)^2}, \quad (10)$$

式中:  $p_{keep}$  为在 Drop Out 中每个激活单元被保留的概率;  $s_{feat}^2$  为特征图大小;  $(s_{feat} - s_{block} + 1)^2$  为经过 Drop Block 操作后有效区域的大小。训练过程中,以固定  $p_{keep}$  的 Drop Block 模块优化网络的效果欠佳,所以设置  $p_{keep}$  为  $[0.75, 0.95]$ 。

#### 3.2.2 置信度得分优化

在 YOLOv3 网络结构中,分类概率乘以物体得分即为最后的检测置信度,表达式为

$$s_{score} = p_{objectness} \times p_{classification}, \quad (11)$$

式中:  $s_{score}$  为检测置信度;  $p_{objectness}$  表示是否有物体;  $p_{classification}$  表示物体属于哪一类。但这其中没有考虑定位精度,会出现准确的检测框置信度得分不高,反而被 NMS 的情况,不利于特征提取和增强,同时降低了网络对火焰目标特征的表征能力。

文献[19]提出 IoU Aware 优化算法,在置信度得分中增加 IoU 预测分量衡量定位的精度。在训练过程中,所提方法在置信度得分的逻辑回归中加入 IoU 预测,增加一个并行通道预测 IoU。每个网格单元预测 3 个预测框,每个预测框的基本参数个数由 5 增加到 6,

加上类别数,得到每个预测框输出的参数量,如图 1 中改进 Head 结构所示。预测时,预测 IoU 乘以分类概率,得到最终的检测置信度。该置信度用于 NMS 计算,使得分高但定位不佳的预测框排名下降,得分高且定位精准的预测框排名上升,有效避免了准确的火焰目标检测框得分不高,被 NMS 的情况,提高网络对火焰特征的特征和提取能力,优化模型参数。

IoU 预测损失函数的计算公式为

$$L_{IoU} = \frac{1}{N_{P_{pos}}} \sum_{t \in P_{pos}} BCE(R_{IoU_t}, \hat{R}_{IoU_t}), \quad (12)$$

式中: $R_{IoU_t}$ 表示第  $t$  个预测框的 IoU; $\hat{R}_{IoU_t}$ 表示第  $t$  个预测框与真实框所计算出的 IoU; $BCE(\cdot)$ 表示二分类交叉熵函数; $P_{pos}$ 表示正例的 IoU; $N_{P_{pos}}$ 表示正例 IoU 的数量。

IoU 梯度的计算公式为

$$\frac{\partial BCE(R_{IoU_t}, \hat{R}_{IoU_t})}{\partial \hat{R}_{IoU_t}} = \log \frac{1 - R_{IoU_t}}{R_{IoU_t}}. \quad (13)$$

在预测时,检测置信度得分的计算公式为

$$S_{det} = c_t R_{IoU_t}^{1-\alpha}, \quad (14)$$

式中: $c_t$ 表示分类得分; $\alpha \in [0, 1]$ 用于控制分类分数和预测  $R_{IoU_t}$ 对检测置信度得分的贡献。

## 4 实验及结果分析

所提网络及其他用于对比实验的网络均在如下的工作条件下完成:操作系统为 Windows10 64 bit、内存为 32 GB、GPU 设备为 Tesla V100、深度学习框架为 Paddle、Python 版本为 3.6。

### 4.1 火焰数据集

火焰数据集部分为网上公开的火焰图像和从火焰视频中截取的图像,同时加入实验室自身拍摄的部分火焰图像扩充火焰数据集,增加数据的多样性。实验数据为 10023 张火焰图像数据,从中随机挑选出 8023 张火焰图像作为训练集,1000 张作为验证集,1000 张作为测试集。数据集包括多种不同场景下的火焰图像,如建筑、森林、车辆、公路、庭院、夜晚等。

### 4.2 训练与实验结果分析

训练时,初始学习率设为常见的 0.0001,采用分段衰减策略进行优化。训练批次大小根据 GPU 性能设为 16,模型的训练损失函数收敛情况如图 5 所示。

由图 5 可知:开始时,损失下降迅速;50 至 200 个 epoch 之间,损失下降缓慢;200 个 epoch 之后,损失趋于稳定,最终为 5 左右。

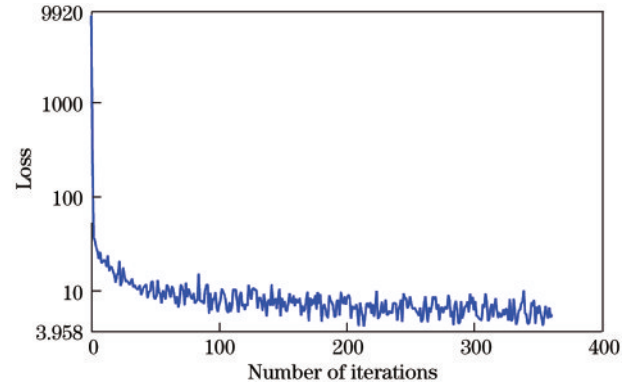


图 5 网络迭代损失变化

Fig. 5 Variation of network iteration loss

所提改进 YOLOv3 火焰检测网络在原始 YOLOv3 网络基础上进行了多点改进:将 DCN 加入到 ResNet50\_vd 的 stage 4 和 stage 5 模块中,采用改进后的 ResNet50\_vd-DCN 作为模型的主干网络;引入 Drop Block 模块;引入 IoU Aware 模块。为了验证 DCN 放置位置的有效性和各模块的作用,进行了消融实验,结果如表 1 所示。

由表 1 可知:将主干网络换成 ResNet50\_vd 后,模型体积减小 1/4,推理速度上升;DCN 加入 stage 4 和 stage 5 时,模型精度提升 4.75 个百分点,效果明显优于模型 B 和模型 C,证明加入 DCN 后,通过采样网格可偏移的特性,卷积核感受野范围灵活可变,所提网络更好地捕获火焰的动态形状特征,增强了对火焰特征提取能力;在 D 模型的基础上,加入 Drop Block 模块,精度上升 0.21 个百分点,推理速度不变,证明加入 Drop Block 模块后,通过丢弃特征图中连续的区域,网络更好地学习火焰其他位置的特征,增强了对不完整火焰特征的提取能力;在 E 模型的基础上,加入 IoU Aware 模块,精度上升 0.09 个百分点,但由于增加了一个通道预测 IoU 值,推理速度降低 3.86 frame/s。从消融实验的结果来看,加入各项改进后,模型的精度和推理速度都提升明显,且模型大小减小 1/4,降低了硬件存储的要求,便于模型布置于实际生活。

表 1 消融实验

Table 1 Ablation experiment

Model	YOLOv3 and its improvement	AP / %	Speed / (frame · s <sup>-1</sup> )	Model volume / MB
A	DarkNet53	89.06	70.41	246
B	ResNet50_vd-DCN (stage 5)	90.30	83.19	181
C	ResNet50_vd-DCN (stage 4)	90.10	81.52	181
D	ResNet50_vd-DCN (stage 4, stage 5)	93.81	77.38	182
E	D+Drop Block	94.02	77.38	182
F	E+IoU Aware(ours)	94.11	73.52	184



为了进一步说明所提火焰检测算法的有效性,设计了对比实验。将所提算法与一阶段和二阶段经典目标检测算法进行对比,包括 Faster R-CNN<sup>[20]</sup>、Cascade R-CNN<sup>[21]</sup>、YOLOv3、YOLOv3\_IoULoss<sup>[22]</sup>、YOLOv4<sup>[23]</sup>、RetinaNet<sup>[24]</sup>,结果如表 2 所示。

由表 2 可知:所提算法的检测精度和推理速度皆高于 Faster R-CNN、Cascade R-CNN、YOLOv3、YOLOv3\_IoULoss、YOLOv4、RetinaNet\_R50 算法,验

证了所提火焰检测算法的有效性;同时,模型大小也只有 184 MB,由于火焰检测算法一般布置于性能较低的现场设备,存储空间受限,小模型更具有优势。不同算法的检测结果如图 6 所示,选取表现较好的二阶段算法 Cascade R-CNN 和一阶段算法 YOLOv4 作为比较对象,前两张图片为正常火焰目标,第三张图片为残缺火焰目标,第四张图片为小火焰目标。

表 2 不同目标检测算法的对比

Table 2 Comparison of different target detection algorithms

Model	Algorithm	AP / %	Speed / (frame · s <sup>-1</sup> )	Model volume / MB
1	Faster R-CNN	90.43	5.28	301
2	Cascade R-CNN	90.87	12.96	355
3	YOLOv3	89.06	70.41	246
4	YOLOv3_IoULoss	89.12	70.55	247
5	YOLOv4	90.84	60.41	266
6	RetinaNet_R50	90.10	61.92	229
7	Propose algorithm	94.11	73.52	184

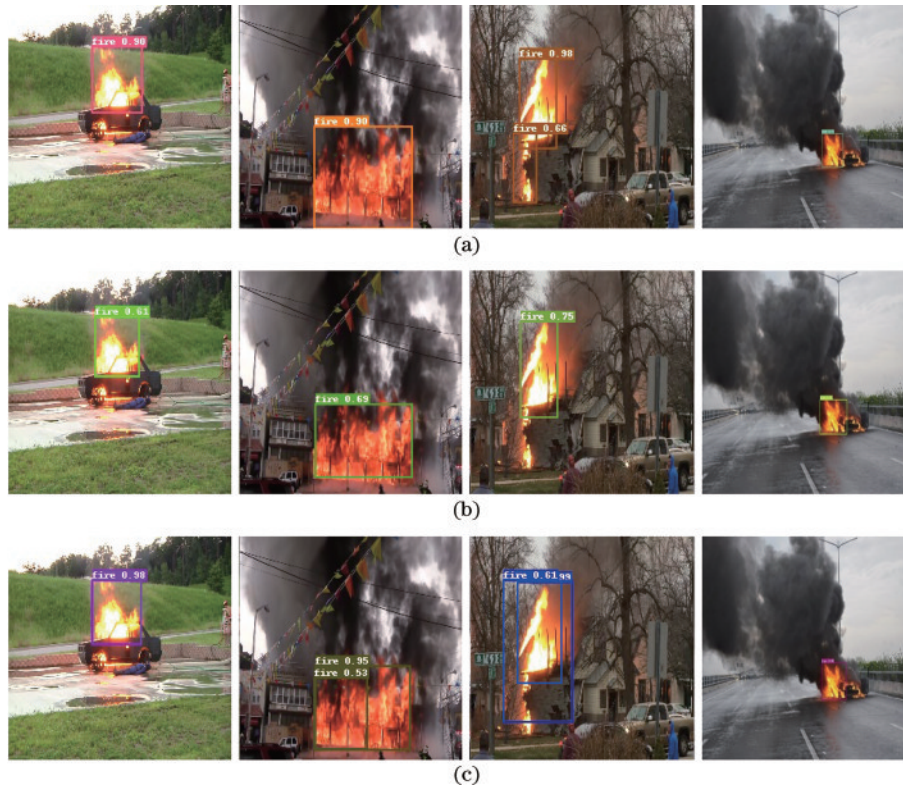


图 6 不同算法的检测效果对比。(a)所提算法;(b) YOLOv4;(c) Cascade R-CNN

Fig. 6 Comparison of detection effects of different algorithms. (a) Proposed algorithm; (b) YOLOv4; (c) Cascade R-CNN

由图 6 可知,在正常火焰目标情况下:YOLOv4 算法的预测框不够精准,预测框偏小,在第二张图片上表现尤为明显,且置信度得分也较低,只有 0.61 和 0.69; Cascade R-CNN 算法对第一张图片的预测效果良好,但第二张图片预测框重叠;所提算法表现优秀,预测框精准,置信度得分达 0.90,预测效果最优。在残缺火焰目标情况下:YOLOv4 算法对残缺火焰小目标存在

漏检; Cascade R-CNN 算法中预测框存在重叠情况;所提算法对残缺火焰目标检测精准,预测效果最优。在小火焰目标情况下:YOLOv4 算法的预测框置信度得分为 0.71; Cascade R-CNN 算法的预测框置信度得分为 0.98;所提算法预测框置信度得分为 0.73。可以看出 Cascade R-CNN 算法预测效果最佳,所提算法与 YOLOv4 算法预测效果较为优秀,但所提模型尺寸最

小,预测速度最快。综上可知,所提检测算法能够有效地提取和增强火焰的动态形状特征,检测效果优秀。

## 5 结 论

火焰检测是机器视觉中的常见检测任务之一,针对现有多目标检测网络对动态形状火焰特征提取及增强能力不足,检测效果不佳的问题,设计了基于动态形状特征提取及增强的改进 YOLOv3 火焰检测算法。将主干网络替换为 ResNet50\_vd,加入 DCN 模块和 IoU Aware 模块,提高了模型对火焰动态形状特征的提取能力;加入 Drop Block 模块,并改进损失函数,提高了模型对火焰动态形状特征的增强能力。实验结果证明,所提模型检测效果优异,有效地提取和增强了火焰的动态形状特征。

### 参 考 文 献

- [1] 荣建忠,姚卫,高伟,等.基于多特征融合技术的火焰视频探测方法[J].燃烧科学与技术,2013,19(3):227-233.  
Rong J Z, Yao W, Gao W, et al. Fire video detection method based on multi-features fusion[J]. Journal of Combustion Science and Technology, 2013, 19(3): 227-233.
- [2] 李巨虎,范睿先,陈志泊.基于颜色和纹理特征的森林火灾图像识别[J].华南理工大学学报(自然科学版),2020,48(1):70-83.  
Li J H, Fan R X, Chen Z B. Forest fire recognition based on color and texture features[J]. Journal of South China University of Technology (Natural Science Edition), 2020, 48(1): 70-83.
- [3] Borges P V K, Izquierdo E. A probabilistic approach for vision-based fire detection in videos[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2010, 20(5): 721-731.
- [4] 赵飞扬,罗兵,林国军,等.基于改进 YOLOv3 的火焰检测[J].中国科技论文,2020,15(7):820-826.  
Zhao F Y, Luo B, Lin G J, et al. Flame detection based on improved YOLOv3[J]. China Sciencepaper, 2020, 15(7): 820-826.
- [5] 熊卫华,任嘉锋,吴之昊,等.基于混合卷积神经网络的火灾识别研究[J].计算机技术与发展,2020,30(7):81-86.  
Xiong W H, Ren J F, Wu Z H, et al. Research on fire identification based on hybrid convolutional neural network[J]. Computer Technology and Development, 2020, 30(7): 81-86.
- [6] 严云洋,朱晓妤,刘以安,等.基于 Faster R-CNN 模型的火焰检测[J].南京师大学报(自然科学版),2018,41(3):1-5.  
Yan Y Y, Zhu X Y, Liu Y A, et al. Flame detection based on Faster R-CNN model [J]. Journal of Nanjing Normal University (Natural Science Edition), 2018, 41(3): 1-5.
- [7] 罗小权,潘善亮.改进 YOLOv3 的火灾检测方法[J].计算机工程与应用,2020,56(17):187-196.  
Luo X Q, Pan S L. Improve the fire detection method of YOLOv3[J]. Computer Engineering and Application, 2020, 56(17): 187-196.
- [8] 喻丽春,刘金.基于改进 Mask R-CNN 的火焰图像识别算法[J].计算机工程与应用,2020,56(21):194-198.  
Yu L C, Liu J. Flame image recognition algorithm based on improved Mask R-CNN[J]. Computer Engineering and Application, 2020, 56(21): 194-198.
- [9] 孙迎春,潘树国,赵涛,等.基于优化 YOLOv3 算法的交通灯检测[J].光学学报,2020,40(12):1215001.  
Sun Y C, Pan S G, Zhao T, et al. Traffic light detection based on optimized YOLOv3 algorithm[J]. Acta Optica Sinica, 2020, 40(12): 1215001.
- [10] Redmon J, Farhadi A. YOLO9000: better, faster, stronger[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition, July 21-26, 2017, Honolulu, HI, USA. New York: IEEE Press, 2017: 6517-6525.
- [11] Redmon J, Ali F. Yolov3: an incremental improvement [EB/OL]. (2018-04-08)[2021-04-05]. <https://arxiv.org/abs/1804.02767>.
- [12] He T, Zhang Z, Zhang H, et al. Bag of tricks for image classification with convolutional neural networks[C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 15-20, 2019, Long Beach, CA, USA. New York: IEEE Press, 2019: 558-567.
- [13] Dai J F, Qi H Z, Xiong Y W, et al. Deformable convolutional networks[C]//2017 IEEE International Conference on Computer Vision, October 22-29, 2017, Venice, Italy. New York: IEEE Press, 2017: 764-773.
- [14] Zhu X Z, Hu H, Lin S, et al. Deformable ConvNets V2: more deformable, better results[C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 15-20, 2019, Long Beach, CA, USA. New York: IEEE Press, 2019: 9300-9308.
- [15] 文奴,郭仁忠,贺彪.基于 DCN-Mobile-YOLO 模型的多车道车辆计数[J].深圳大学学报(理工版),2021,38(6):628-635.  
Wen N, Guo R Z, He B. Multi-lane vehicle counting based on DCN-Mobile-YOLO model[J]. Journal of Shenzhen University (Science and Engineering), 2021, 38(6): 628-635.
- [16] 李国进,荣誉.基于 DCN+SERes-YOLOv3 的人脸佩戴口罩检测算法[J].计算机与现代化,2021(9):12-20,30.  
Li G J, Rong Y. Face mask detection algorithm based on DCN-SERes-YOLOv3[J]. Computer and Modernization, 2021(9): 12-20, 30.
- [17] Ghiasi G, Lin T Y, Le Q V. DropBlock: a regularization method for convolutional networks[EB/OL]. (2018-10-30)[2021-04-05]. <https://arxiv.org/abs/1810.12890>.
- [18] 朱浩,宁芊,雷印杰,等.基于注意力机制-Inception-CNN 模型的滚动轴承故障分类[J].振动与冲击,2020,39(19):84-93.  
Zhu H, Ning Q, Lei Y J, et al. Fault classification of rolling bearing based on attention mechanism-Inception-CNN Model[J]. Journal of Vibration and Shock, 2020, 39(19): 84-93.



- [19] Wu S K, Li X P, Wang X G. IoU-aware single-stage object detector for accurate localization[J]. *Image and Vision Computing*, 2020, 97: 103911.
- [20] Ren S Q, He K M, Girshick R, et al. Faster R-CNN: towards real-time object detection with region proposal networks[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, 39(6): 1137-1149.
- [21] Cai Z W, Vasconcelos N. Cascade R-CNN: delving into high quality object detection[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, June 18-23, 2018, Salt Lake City, UT, USA. New York: IEEE Press, 2018: 6154-6162.
- [22] 陈俊. 基于YOLOv3算法的目标检测研究与实现[D]. 成都: 电子科技大学, 2020.  
Chen J. Research and implementation of target detection based on YOLOv3 algorithm[D]. Chengdu: University of Electronic Science and Technology of China, 2020.
- [23] Bochkovskiy A, Wang C Y, Liao H Y M. YOLOv4: optimal speed and accuracy of object detection[EB/OL]. (2020-04-23) [2021-04-05]. <https://arxiv.org/abs/2004.10934>.
- [24] Lin T Y, Goyal P, Girshick R, et al. Focal loss for dense object detection[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2020, 42(2): 318-327.