



激光与光电子学进展
Laser & Optoelectronics Progress
ISSN 1006-4125,CN 31-1690/TN

《激光与光电子学进展》网络首发论文

题目： 基于 YOLOv3 改进的火焰目标检测算法
作者： 张彬彬，帕孜来·马合木提
收稿日期： 2021-01-20
网络首发日期： 2021-03-12
引用格式： 张彬彬, 帕孜来·马合木提. 基于 YOLOv3 改进的火焰目标检测算法[J/OL]. 激光与光电子学进展.
<https://kns.cnki.net/kcms/detail/31.1690.TN.20210311.1628.047.html>



网络首发：在编辑部工作流程中，稿件从录用到出版要经历录用定稿、排版定稿、整期汇编定稿等阶段。录用定稿指内容已经确定，且通过同行评议、主编终审同意刊用的稿件。排版定稿指录用定稿按照期刊特定版式（包括网络呈现版式）排版后的稿件，可暂不确定出版年、卷、期和页码。整期汇编定稿指出版年、卷、期、页码均已确定的印刷或数字出版的整期汇编稿件。录用定稿网络首发稿件内容必须符合《出版管理条例》和《期刊出版管理规定》的有关规定；学术研究成果具有创新性、科学性和先进性，符合编辑部对刊文的录用要求，不存在学术不端行为及其他侵权行为；稿件内容应基本符合国家有关书刊编辑、出版的技术标准，正确使用和统一规范语言文字、符号、数字、外文字母、法定计量单位及地图标注等。为确保录用定稿网络首发的严肃性，录用定稿一经发布，不得修改论文题目、作者、机构名称和学术内容，只可基于编辑规范进行少量文字的修改。

出版确认：纸质期刊编辑部通过与《中国学术期刊（光盘版）》电子杂志社有限公司签约，在《中国学术期刊（网络版）》出版传播平台上创办与纸质期刊内容一致的网络版，以单篇或整期出版形式，在印刷出版之前刊发论文的录用定稿、排版定稿、整期汇编定稿。因为《中国学术期刊（网络版）》是国家新闻出版广电总局批准的网络连续型出版物（ISSN 2096-4188, CN 11-6037/Z），所以签约期刊的网络版上网络首发论文视为正式出版。

基于 YOLOv3 改进的火焰目标检测算法

张彬彬¹, 帕孜来·马合木提^{1*}

¹新疆大学电气工程学院, 新疆 乌鲁木齐 830047

摘要 针对火灾小目标、多目标、边缘模糊等情况而导致其检测精度不佳的问题, 提出了一种基于 YOLOv3 改进的火焰目标检测算法。通过改进特征金字塔 FPN 使得局部信息得到二次利用, 设计大尺度全卷积模块获取不同尺度的全局空间信息, 采用改进的通道空间注意力机制提高有效信息并抑制无用信息, 利用 CIoU 和 Focal Loss 作为损失函数提高难识别目标的精度和缓解数据集不平衡的问题。通过自制火焰数据进行实验, 并与其他先进算法对比, 改进后的算法平均精度均值 MAP 高达 89.82%, 较原算法提高了 4.32%, 检测速度达 20.2FPS, 具有更高的检测精度和较快的检测速度, 满足了火灾检测的实时性和高效性要求。

关键词 目标检测; 特征金字塔; 全卷积模块; 注意力机制; 损失函数

中图分类号 TP391.41 **文献标志码** A

Flame target detection algorithm based on YOLOv3 improvement

Zhang Binbin¹, Pazilai · Mahemuti^{1*}

¹School of Electrical Engineering, Xinjiang University, Urumqi, Xinjiang 830047, China

Abstract Aiming at the problem of poor detection accuracy caused by fire small target, multi-target and fuzzy edge, an flame target detection algorithm based on YOLOv3 improvement is proposed. By improving the feature pyramid FPN, local information can be used twice,a large-scale full convolution module is designed to obtain global spatial information of different scales, and an improved passageway space attention mechanism is used to improve effective information and suppress useless information,CIoU and Focal Loss are used as the loss function improves the accuracy of difficult to identify targets and alleviates the problem of data set imbalance. Experiments are carried out with self-made flame data and compared with other advanced algorithms, the average accuracy MAP of the improved algorithm is as high as 89.82%, which is 4.32% higher than that of the original algorithm, and the detection speed reaches 20.2 FPS, compared with other advanced algorithms, the improved algorithm has higher detection accuracy and faster detection speed, It meets the real-time and high-efficiency requirements of fire target detection.

Key words target detection; feature pyramid; full convolution module; attention mechanism; loss function

OCIS codes 110.2970; 100.3008; 110.2960; 150.1135

1 引言

火灾是人类面临的主要自然灾害之一, 全球火灾发生频率居高不下。2018 年美国发生大约 1,318,500 起火灾, 直接导致了 720 平民死亡和 256 亿美元损失。2019 年四川凉山发生的森林火灾造成了 31 人死亡和财产损失, 其灭火工作难度之大更甚于城市火灾。因此,

基金项目: 新疆维吾尔自治区自然科学基金(2016D01C038)

E-mail: zhangken178@126.com; *E-mail: 294625876@qq.com

对于火灾早期的有效预防、预警和应对方法变得尤为重要。

早期火灾预警大多采用传统的火灾探测器，即火灾产生的烟雾、温度或其他因素组合触发探测器启动预警和灭火系统^[1-4]。这种探测方法受多种因素影响，其响应时间慢，有漏报误报等问题存在，且响应延迟可能会导致火势迅速蔓延，无法控制。机器学习的发展带来了监控视频或图像数据探测火灾的新技术^[5]，但此方法的效果取决于选取的人工特征的好坏^[6]。由于火灾类型难以界定，且边缘模糊加上环境复杂，难以提取到理想特征，因而此方法存在检测精度低、延迟和计算量大等问题。深度学习方法可以解决传统机器学习在特征提取上所面临的问题，在火灾检测领域的发展空间极大^[7]，它可以让大量标注好的图片输入到深层非线性网络自行学习特征，因而对图片的变形具有高度的不变性。目前，主流检测算法包括以 RCNN^[8]为代表的两阶段算法和以 YOLO^[9]系列为代表的单阶段算法。Jiuling 等^[10]设计了以 Faster-RCNN 模型与软件平台结合的火灾检测方法，取得了较好精度，但不能达到实时检测的效果。Fu 等^[11]提出以 AlexNet 模型为基础，以多个 SVM 作为分类器的森林火灾检测网络，并用迁移学习训练了整个网络模型，取得了较高检测精度，但由于 SVM 二分类的特性和数据集背景单一，存在泛化能力差等问题。Fang 等人^[12]基于 Tinier-YOLO 设计出了微型检测网络，在检测速度方面取得较大优势。Frizz 等^[13]提出一种卷积神经网络来识别视频中的火灾，说明了火灾和烟雾的运动信息可以被编码，大大降低了时间成本，但仅在目标分类领域表现出了良好的性能。

针对以上火灾检测算法的不足，本文以自制火焰数据集作为研究对象，提出了一种改进的 YOLOv3 检测算法。首先，重新设计特征金字塔（Feature Pyramid Networks, FPN）结构，将局部信息重新有效利用到整个网络，实现局部信息与全局信息的二次融合。其次，在深层特征提取路径上通过注意力机制模块，实现在通道和空间上的信息交互，使得模型能筛选出不重要的信息特征，以提高网络有效信息的提取能力和抗干扰能力。设计了大尺度全卷积模块，使用了对称的卷积组合实现大卷积功能，减少了计算量并丰富了特征图的全局信息。在此基础上，使用 CIoU 损失函数替换了坐标误差损失函数，考虑到了目标与预测框的中心距离、重叠率和长宽比信息，提高了难识别图片的精度。采用 Focal Loss 替换类别损失函数，缓解二分类数据集不平衡的问题。实验结果表明，改进后的算法在小目标、多目标和复杂背景图片上具有更高精度并保持较快的检测速度。

2 YOLOv3 算法概述

2.1 YOL0v3 结构

YOL0v3 采用 Darknet53 为主干特征提取网络^[14], Darknet53 包含 53 个卷积层, 结合 Resnet^[15] 网络思想加入了残差结构, 该结构能使网络在进行深层卷积提取特征的同时又避免了梯度消失和梯度爆炸的问题。残差模块 ResX 由残差单元 Resunit 堆叠而成, 如图 1 所示。其中, DBL 为深度卷积, 包含了卷积、标准化(Batch Normalization, BN) 和 Leaky ReLU 激活函数操作, 每次卷积均用了 L2 正则化处理, 减轻了网络训练中过拟合情况。Resunit 的过程就是将输入的特征图进行一次大小为 3×3 、步长为 2 的卷积, 保存该卷积的 Layer, 再进行一次 1×1 的卷积压缩 2 倍通道和 3×3 的卷积扩张 2 倍通道, 把最后得到的特征和保存的 Layer 进行堆叠。最终残差模块由 DBL 单元和若干个 Resunit 组合而成, 图 1 中 X 代表该模块用了 X 个残差单元进行堆叠。在 Darknet53 中一共用了 5 个残差模块, 所用的残差单元数量分别为 1、2、8、8、4。

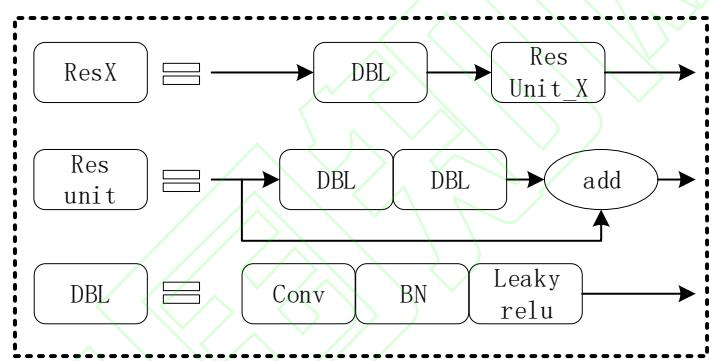


图 1 残差模块

Fig.1 Residual module

YOL0v3 的整体结构如图 2 所示。输入的图片经过 Reshape, 尺寸统一到 416×416 , 经过 5 个步长为 2 的残差模块下采样, 分别得到 5 个尺度的特征图。YOL0v3 网络将最后三个残差模块得到的特征图输入到 FPN 网络中, 其中 13×13 的特征经过 5 次卷积后送到检测网络中, 同时将同一特征图进行卷积压缩通道数和上采样扩充尺寸, 与 26×26 的特征图进行 Concat 融合, 融合后的特征再次进行 5 次卷积送到检测网络。同上述方法, 将同一特征经过卷积和上采样后与尺寸为 52×52 特征图进行 Concat 融合, 经 5 次卷积后送到检测网络。最终检测网络得到 $(75, 13, 13)$, $(75, 26, 26)$, $(75, 52, 52)$ 共 3 个不同尺度的特征进行多尺度检测, 分别对应检测大中小物体。

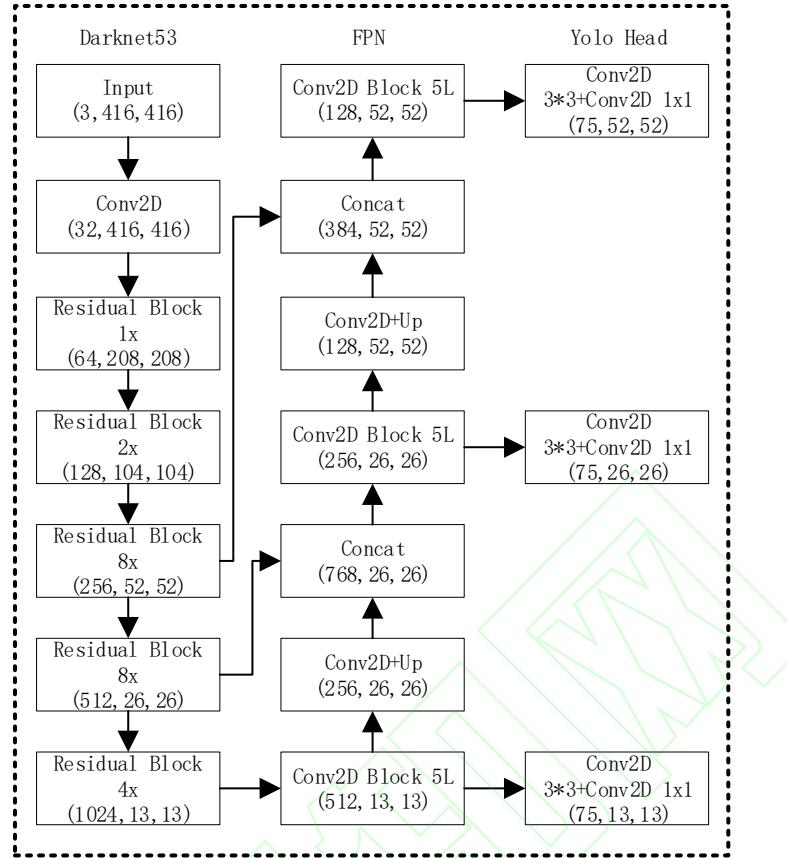


图 2 YOLOv3 算法结构示意图

Fig.2 Schematic diagram of YOLOv3 algorithm structure

采用 YOLOv3 算法检测物体前，先用 K-means 聚类算法根据数据集标注框算出 9 组锚框 (anchor)，平均分配到 3 个不同的尺度上。每个尺度中的每个锚点分别对应 3 组锚框，总计 10647 个锚框，由此，可以得到 3 个特征层先验框的位置^[16]。

2.2 目标边界框预测

YOLOv1 算法的预测框坐标 (x, y, w, h) 是直接预测出来的，当坐标有微小的变化时，其映射到实际图片中会造成很大变化，从而预测失败。YOLOv3 针对这一问题对预测框公式进行了改进：

$$b_x = \sigma(t_x) + c_x \quad (1)$$

$$b_y = \sigma(t_y) + c_y \quad (2)$$

$$b_w = p_w e^{t_w} \quad (3)$$

$$b_h = p_h e^{t_h} \quad (4)$$

其中， t_x, t_y, t_w, t_h 是网络输出的坐标， p_w, p_h 为 anchor 的宽高， c_x, c_y 为相

对左上角的位置，每个小格的长度为 1，如图 3 所示。

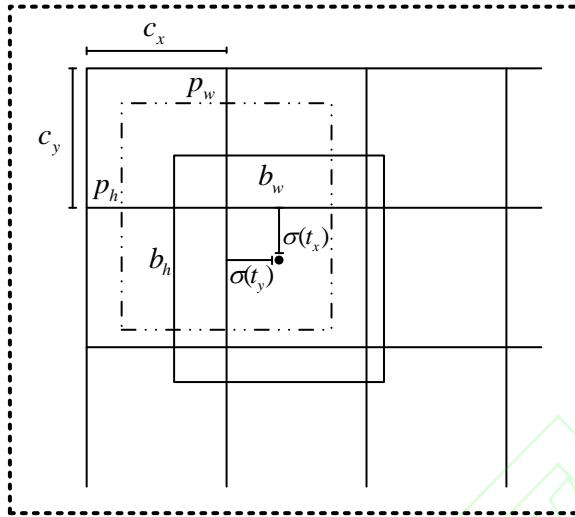


图 3 YOLOv3 边框预测示意图
Fig.3 YOLOv3 border prediction diagram

t_x 与 t_y 经过与 sigmoid 输出为 $0 \sim 1$ 之间的偏移量，然后与 c_x 、 c_y 相加得到预测框的中心点。 p_w 和 p_h 分别与 t_w 和 t_h 作用后得到预测框的宽和高，得到一系列预测框的坐标。最终，通过置信度阈值和非极大值抑制 NMS (Non-Maximum Suppression) 处理后得到最终预测框结果。相对于 YOLOv1，YOLOv3 的边框回归计算方式可以有效防止偏移过大带来预测结果失败等问题。

3 改进的 YOLOv3 算法

3.1 大尺度全卷积模块

现代目标检测网络大多采用了将特征图 Feature Map 直接送到特征融合网络，对 Feature Map 进行多尺度融合。计算量虽然减少了，但不同层次的 Feature Map 会丢失部分全局信息。针对以上两个问题，本文设计出大尺度全卷积模块 GCM (Global Convolutional module)，在尽可能不增加计算量的基础上，对不同层级的 Feature Map 提取全局信息，其计算方式如图 4。

卷积操作中的卷积核大小会直接影响到全局信息的获取。将卷积核大小增加到 Feature Map 的空间尺寸，其在原图上产生的有效域能尽可能的包含目标，从而可以有效的获取全局信息，但卷积核的增大会带来巨大的计算量。本文通过 $K \times 1$ 和 $1 \times K$ 的对称独立卷积组合，在增加少量计算量的基础上实现了 $K \times K$ 的大卷积操作，通过取消非线性操作采用全卷积，减少了部分特征丢失，得到的特征与原来的特征进行叠加，丰富了其上下文的全局信息，同时也

起到防止退化的作用。

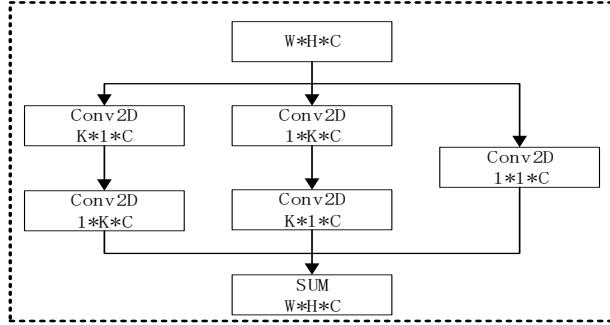


图 4 大尺度全卷积模块

Fig.4 Large-scale full convolution module

3.2 通道空间注意力机制

本文在通道注意力(Efficient Channel Attention, ECA)^[17]基础上，添加空间注意力，设计出通道空间注意力机制(Channel Spatial Attention, CSA)，如图 5 所示。该模块通道注意力部分：Feature Map 通过全局平均池化，使用大小为 K 的快速 1D 卷积，实现每个通道和 K 个邻近通道来获取局部跨通道交互信息，其中 K 的大小为局部跨信道交互的覆盖率，即此通道附近有多少通道参与这个信道的注意力部分，再经过 Sigmoid 函数生成 $1 \times 1 \times C$ 尺寸的通道注意力分支，与原 Feature Map 进行点乘。空间注意力部分：通过 1×1 卷积和 Sigmoid 函数，生成了空间维度的权重与原 Feature Map 进行点乘。最后将两者的结果进行 Add 叠加，从而实现通道和空间特征的权重分配。

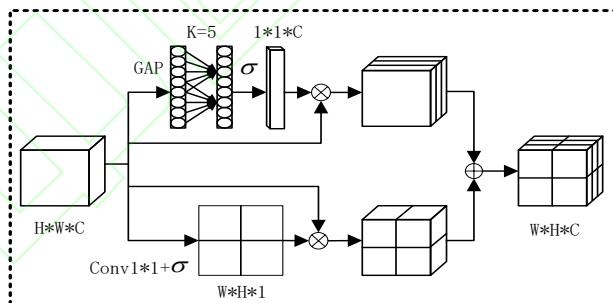


图 5 通道空间注意力机制

Fig.5 Channel Spatial Attention (CSA)

3.3 改进的 FPN 层

在图 2 中，YOLOv3 算法的 FPN 层没有充分利用局部信息。在小目标检测中，局部信息有利于小目标的检测，而全局信息侧重于大目标检测。主干特征提取网络，其浅层提取局部信息，深层能提取全局信息。因此本文重新设计了 FPN 层，改进的 YOLOv3 算法如图 6 所示。将原 YOLOv3 在自下而上的融合策略基础上增加自上而下的融合路径，并将第二个残差模块

提取的局部信息重新加以利用。

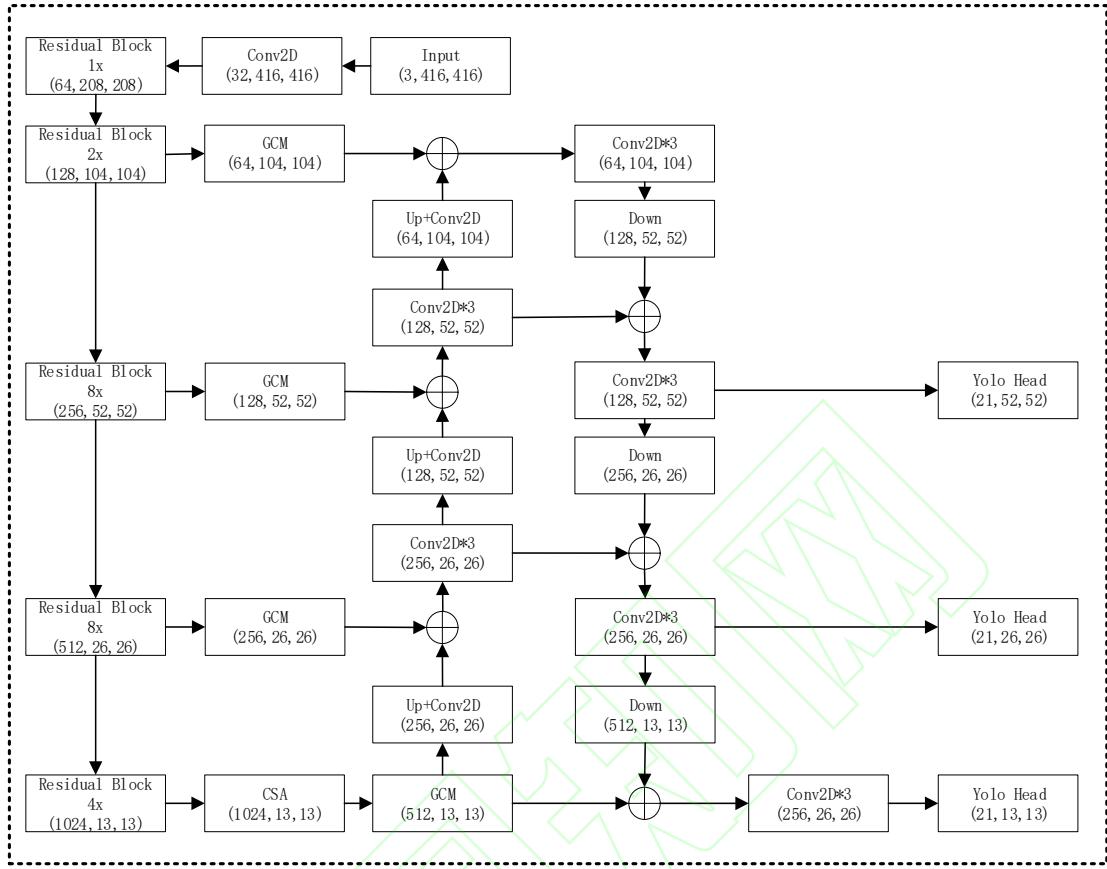


图 6 本文算法结构
Fig.6 The algorithm structure of this paper

本文算法结构最深层的 Feature Map 利用率最高，在此路径上先通过注意力机制模块 CSA 提取有效信息，抑制无用信息。在 4 个不同层级的 Feature Map 路径上，分别通过 GCM 模块丰富其全局信息。从图 6 中可以看出，本文引入了第二个 Residual Block 提取的局部信息，并将 YOLOv3 原 FPN 层自下而上的特征融合策略增加了自上而下的特征融合。整个算法流程如下：首先，最底层(1024, 13, 13)的 Feature Map 经过 2 倍上采样压缩尺寸和 1×1 的卷积扩充通道，并保留原特征 Feature Map1，与(256, 26, 26)的 Feature Map 进行 Concat 后，进行 3 次卷积保留其特征 Feature Map2，经过 2 倍上采样和 1×1 的卷积后与(128, 52, 52)的 Feature Map 进行 Concat，经过 3 次卷积后保留特征为 Feature Map3，再经过 2 倍上采样和 1×1 卷积与(64, 104, 104)的 Feature Map 进行 Concat，经过 3 次卷积和下采样后与 Feature Map3 进行 Concat，接着 3 次卷积后保留特征为 Feature Map4，再经过下采样与 Feature Map2 进行 Concat，同样经过 3 次卷积后保留特征为 Feature Map5，接着下采样与 Feature Map1 进行 Concat，经过 3 次卷积后保留为 Feature Map6，最终将 Feature Map4、Feature Map5、Feature Map6 送入 Yolo Head 得到三个不同尺度的特征预测层。其输出尺

度分别为(21, 13, 13)、(21, 26, 26)和(21, 52, 52)，通道数为21是因为本文采用的数据集为2分类，通道数为，其中3为对应的三个anchor的通道，5为先验框坐标和分类的通道数，2为类别的通道数。

3.4 损失函数

YOLOv3的损失函数由三部分构成：由中心点和宽高坐标部分带来的误差、置信度误差和分类误差。其中坐标误差采用真实框与预测框的宽高差方计算，分类误差采用交叉熵损失函数。IoU是目标检测中的常见指标，计算公式如下：

$$IoU = \frac{|M \cap N|}{|M \cup N|} \quad (5)$$

M 代表预测框坐标， N 代表真实框坐标。IoU作为损失函数，可以直接反映出真实框与预测框的距离，但是在两个框没有相交时，产生的Loss值为0，所以无法进行反向梯度更新。本文采用CIoU作为坐标误差损失函数，CIoU考虑到了真实框与预测框中心点的距离、重叠部分和长宽比的信息，在训练过程中缓解了发散等问题。计算公式如下：

$$Loss_{CIoU} = 1 - IoU + \frac{\rho^2(b, b^{gt})}{c^2} + \alpha v \quad (6)$$

$$v = \frac{4}{\pi^2} (\arctan \frac{\omega^{gt}}{h^{gt}} - \arctan \frac{\omega}{h})^2 \quad (7)$$

$$\alpha = \frac{v}{1 - IoU + v} \quad (8)$$

其中， b, b^{gt} 表示预测框与真实框的中心点， ρ^2 为欧式距离， c 为同时包含预测框与真实框的最小闭包区域的对角线距离， v 是用来衡量长宽比的相似性， α 为权重。

由于本文自制数据集二分类存在不平衡的问题，所以分类误差采用二分类的Focal Loss^[18]，相比于交叉熵Loss，Focal Loss加入了平衡因子，用来调节正负样本损失，加入的gamma因子减少易分类样本的损失。计算公式如下：

$$Loss_{cls} = \begin{cases} -\alpha(1 - \hat{y})^\gamma \log(\hat{y}) & y = 1 \\ -(1 - \alpha)\hat{y}^\gamma \log(1 - \hat{y}) & y = 0 \end{cases} \quad (9)$$

其中， $y \in \{0, 1\}$ 为真实的标签， \hat{y} 为预测值，平衡因子 $\alpha = 0.25$ ， $\gamma = 2$ 。Focal loss的引入可以有效解决one-stage目标检测中的分类样本不平衡问题

4 实验结果与分析

4.1 数据集和实验平台

目前，网络上并没有公开的较为全面的火灾图像数据集。为此，本文利用人工拍摄和网络资源等方法制作了火焰数据集，用于模拟火灾数据，以便网络学习到火焰特征。为了提高检测模型的实用性，加入了人造光源作为干扰项，这是因为卷积神经网络在提取的多层次特征中，可以学到多层次抽象的数据信息，其中浅层提取主要是物体颜色和边缘信息，深层提取主要是物体的轮廓信息，火焰图片经过网络层层特征提取，最后其轮廓信息容易和光源等发光物产生混淆。最终得到训练集 3816 张，测试集 562 张，火焰与人造光源占比为 3 比 1 的二分类数据集。实验设备为 I5-9400F，GPU 为 NVIDIAGT2060S，使用 Pytorch 框架，采用 Adam 优化器，迭代轮次(epoch) 为 100 轮，初始学习率为 0.001，动量因子为 0.9。

4.2 评价指标

使用目标检测常见评价指标，其中包括精确率(Precision)、召回率(Recall)、平均精度(Average-Precision, AP)、平均精度均值(Mean Average Precision, MAP)、F1 Score 和检测时间 FPS。其中精确率和召回率计算公式如下：

$$\text{Precision} = \frac{tp}{tp + fp} \quad (10)$$

$$\text{Recall} = \frac{tp}{tp + fn} \quad (11)$$

$$F1 = 2 \times \frac{\text{precision} \times \text{recall}}{\text{precision} + \text{recall}} \quad (12)$$

其中， tp 为正样本被正确识别为正样本的数量， fp 为负样本被错误识别为正样本的数量， fn 为正样本被错误识别为负样本的数量。

平均准确率(AP)是不同置信度下 Precision 和 Recall 的曲线面积。AP 值兼顾了精确率和召回率，采用 AP 值来衡量模型检测每个类别的性能，对于整个模型性能的评价采用各类 AP 值的均值 MAP 衡量。F1 值为作为 Precision 和 Recall 的调和平均数，同样是多分类问题的参考指标。检测时间用每秒生成图片的张数 FPS 来衡量。

4.3 实验结果分析

本文提出的算法是基于 YOLOv3 改进，为了对比两种算法的性能差异大小，分别打印出训练过程中指标变化用以对比，实验结果如图 7 所示。

从图 7 中可以看出，本文算法在 MAP 值和 F1 较 YOLOv3 算法均取得较好的性能。在 MAP 指标对比中，YOLOv3 的 MAP 指标最终稳定在 85.5%，而本文算法 MAP 值高达 89.82%，相比于 YOLOv3 提高约 4 个百分点。在 F1 值对比中，YOLOv3 的 F1 值在 78.34%，本文算法的 F1 值高达 86.59%。替换后的 LOSS 最终收敛低于 YOLOv3 的 LOSS 值。总体而言，本文改进后的算法在检测精度上优于 YOLOv3 检测算法。

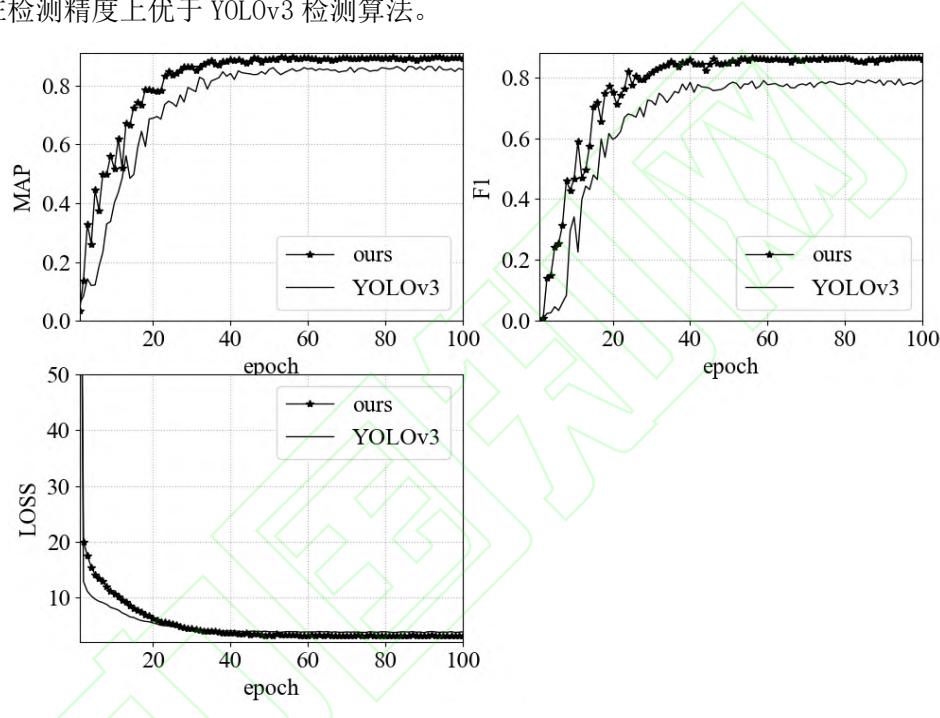
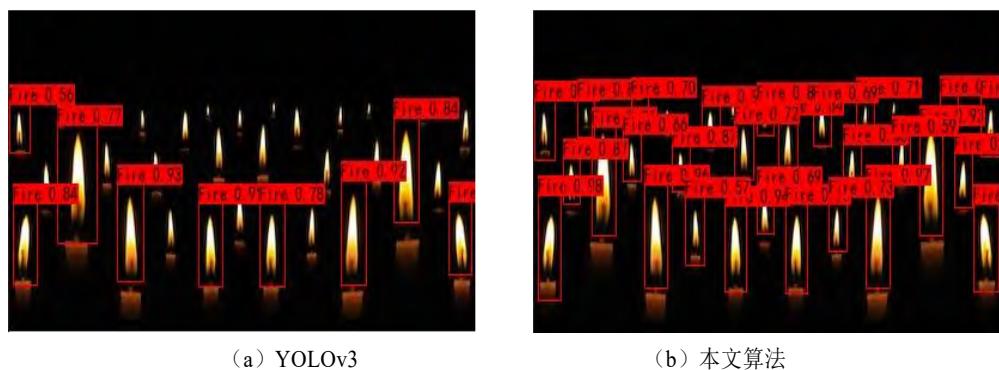


图 7 本文算法与 YOLOv3 对比
Fig.7 This algorithm is compared with YOLOv3

本文选取了小目标、多目标和复杂背景实际效果图 8 所示。从图 8 检测效果对比可以看出，对于火焰在多目标、小目标和复杂场景上，较 YOLOv3 均取得良好效果，尤其在多目标和小目标的实际检测效果上，本文算法取得了较大优势。



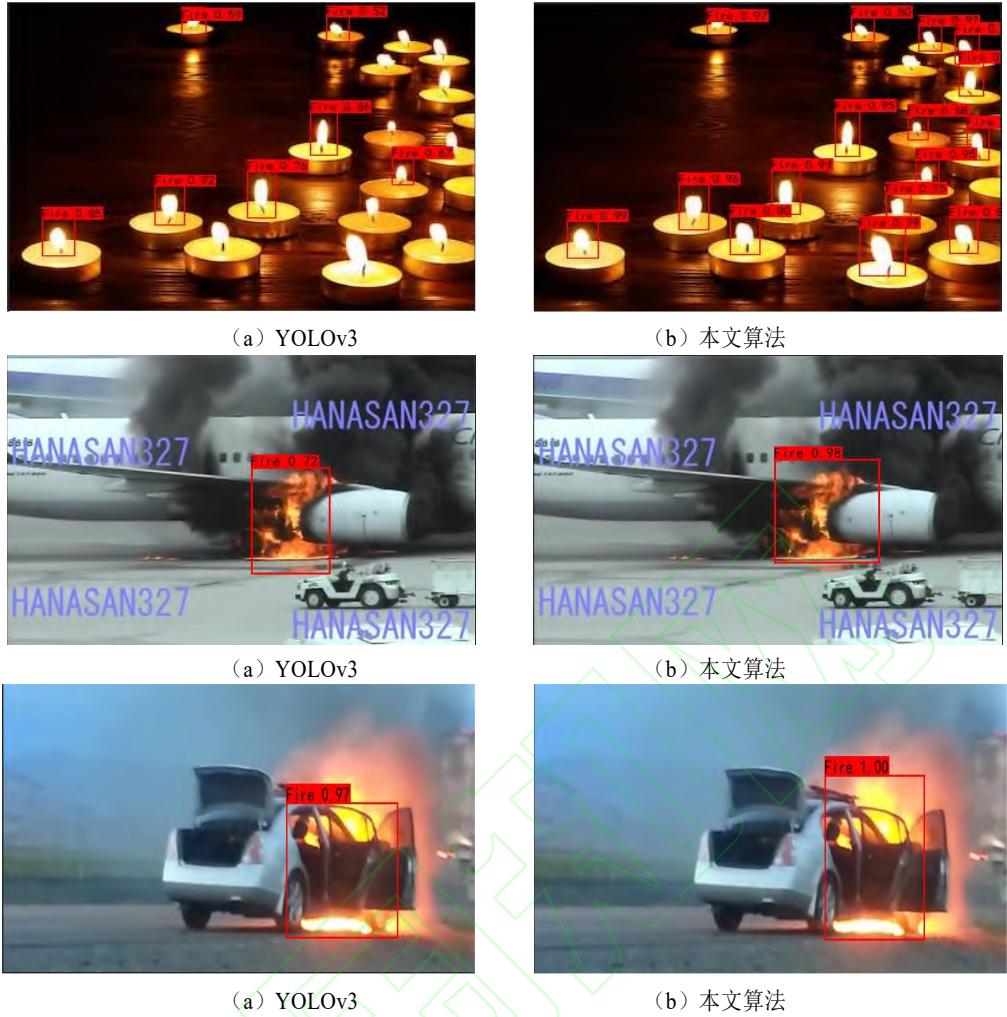


图 8 实际检测效果对比

Fig.8 Comparison of actual detection effect

4.4 本文算法与其他算法比较

为了进一步验证本文算法的性能，将其他先进算法与本文算法进行对比，比较的结果如表 1 所示。

表 14 种算法性能比较

Table 1 Performance comparison of four algorithms

| Network | Category | Precision/% | Recall/% | F1/% | MAP/% | FPS |
|----------------------|----------|-------------|----------|-------|-------|------|
| YOLOv3 | Fire | 91.72 | 69.76 | 79.00 | 85.35 | 26.9 |
| | Light | 82.17 | 75.71 | 79.00 | | |
| Retinanet et al [18] | Fire | 90.71 | 78.43 | 84.00 | 86.22 | 18.5 |
| | Light | 83.21 | 81.43 | 82.00 | | |
| Centernet et al [19] | Fire | 93.57 | 77.95 | 85.00 | 88.19 | 15.3 |
| | Light | 88.46 | 82.14 | 85.00 | | |
| Ours | Fire | 95.07 | 78.90 | 86.00 | 89.82 | 20.2 |
| | Light | 89.39 | 84.29 | 87.00 | | |

从表 1 中可以看出，本算法对于各个分类的精确率、召回率和 F1 值均优于其他各类算

法。由于实验平台的限制, Retinanet 采用 Resnet50 作为特征提取网络, 其 FPN 网络加 Focal Loss 的引入使得 Retinanet 的检测精度略高于 YOLOv3 算法, 精度低于 Centernet-HG。Centernet-HG 的特征提取网络是 Hourglass, Hourglass Network 有 19000 万的参数量, 有优秀的特征提取能力, 同时该网络利用高分辨率特征图取得较好的检测效果, 但巨大的网络参数在检测速度上不理想。本文提出的算法在精度和速度上有较大优势, 与 Centernet-HG 算法相比, 其轻量化的结构更加适合实时检测。

4.5 消融实验结果

本算法在改变网络结构的同时添加了大尺度全卷积模块和注意力机制模块, 更改了原有 Loss 的计算方式。为进一步分析改进算法相较于 YOLOv3 的影响, 将本文算法拆分为 5 组实验, 其实验结果如表 2 所示。

表 2 消融实验结果

Table 2 Ablation experiment results

| NET | FPN | CSA | GCM | LOSS | MAP/% |
|-----|-----|-----|-----|------|-------|
| 1 | × | × | × | × | 85.35 |
| 2 | √ | × | × | × | 87.80 |
| 3 | √ | √ | × | × | 88.52 |
| 4 | √ | √ | √ | × | 88.97 |
| 5 | √ | √ | √ | √ | 89.82 |

从表 2 可以看出, 在第 2 组实验中, 网络结构 FPN 层, 引入第二个 Residual Block 提取的局部信息, 并加入自上而下的融合路径, 提高了局部信息的利用率, 输出的有效特征图更加丰富, 其 MAP 值提升了 2 到 3 个百分点。在第 3 组实验中 CSA 注意力机制的加入, 使网络更加专注有用信息而抑制无用信息, 相对于第 2 组实验结果 MAP 值提升 0.7 个百分点。第 4 组实验在 4 个 Residual Block 提取特征图的路径上分别加入 4 个 GCM 模块, 在提取的特征图上丰富了全局信息, 由于数据集中大目标占比较小, 其 MAP 值较第 3 组提高 0.4 个百分点。第 5 组实验更换了损失函数的计算方式, 引入了 CIoU 和 Focal Loss, 缓解了数据集二分类不平衡的问题, 较第 4 组实验结果提升了 0.85 个百分点。综上, 本文针对 YOLOv3 算法的改进策略均能够提升对本文数据集的检测效果。

5 结 论

本文通过引入主干特征提取网络第 2 个 Residual Block 提取局部信息, 并增加自上而下的特征融合结构, 增加了局部信息利用率; 在 4 个 Residual Block 特征路径上加入大尺

度全卷积模块丰富了全局信息；引入了注意力机制模块使网络专注于有用信息；将损失函数更改为 CIoU 和 Focal Loss，增强了难识别等目标的精度以及缓解二分类数据集不平衡的问题。这些基于 YOLOv3 算法的改进，使目标检测具有较高精度和检测速度，下一步将考虑自行设计网络结构，以构造性能更优越的轻量级网络，从而提升模型的检测精度和实时性。

参考文献

- [1] Wu S, Guo C, Yang J. Using PCA and one-stage detectors for real-time forest fire detection[J]. The Journal of Engineering, 2020(13): 383-387.
- [2] Vikram R, Sinha D, De D, et al. EEFFL: energy efficient data forwarding for forest fire detection using localization technique in wireless sensor network[J]. Wireless Networks, 2020(2): 1-29.
- [3] Jingmin D, Haiy Y, Fang S, et al. An early fire gas sensor based on $2.33\mu\text{m}$ DFB laser[J]. Infrared Physics & Technology, 2018, 92: 84-89.
- [4] Toptas B, Hanbay D. A new artificial bee colony algorithm-based color space for fire/flame detection[J]. Soft Computing, 2019(2): 1-12.
- [5] Tian L, Wang J, Zhou H, et al. Automatic detection of forest fire disturbance based on dynamic modelling from MODIS time-series observations[J]. International Journal of Remote Sensing, 2018, 39(11-12): 3801-3815.
- [6] Fu T J. Forest fire image recognition algorithm and implementation based on deep learning[D]. Beijing: Beijing Forestry University, 2016.
傅天驹. 基于深度学习的林火图像识别算法及实现[D]. 北京：北京林业大学, 2016.
- [7] Gaur A, Singh A, Kumar A, et al. Video flame and smoke based fire detection algorithms: A Literature Review[J]. Fire Technology, 2020(9): 1-38.
- [8] Shoulin Y, Hang L, Lin T. Airport detection based on improved faster RCNN in large scale remote sensing images[J]. Sensing and Imaging, 2020, 21(1): 1747-1755.
- [9] Redmon J, Farhadi A. YOLOV3: an incremental improvement[C]//Proceedings of 2018 IEEE Conference on Computer Vision and Pattern Recognition. Washington D.C., USA: IEEE Press, 2018: 1-6.
- [10] Jiuling Z, Yang J, Ding Z, et al. Study on the situational awareness system of mine fire rescue using faster ross girshick-convolutional neural network[J]. IEEE Intelligent Systems, 2019, 35(1): 54-61.
- [11] Fu Y J, Zhang H L. Forest fire detection method based on transfer learning of convolutional neural network[J]. Laser & Optoelectronics Progress, 2020, 57(4): 041010.
富雅捷, 张宏立. 基于迁移学习的卷积神经网络森林火灾检测方法[J]. 激光与光电子学进展, 2020, 57(4): 041010.
- [12] Fang W, Wang L, Ren P, et al. YOLO: A real-time object detection method for constrained environments[J]. IEEE Access, 2019, PP(99): 1-1.
- [13] Frizz S, Kaabi R, Bouchouicha M, et al. Convolutional neural network for video fire and smoke detection[C]. Conference of the IEEE Industrial Electronics Society. IEEE, 2016: 877-882.

- [14] Ju M R, Luo H B, Wang Z B, et al. Improved YOLO V3 algorithm and its application in small target detection[J]. Acta Optica Sinica, 2019, 39(7): 0715004.
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[J].光学学报,2019,39(7): 0715004.
- [15] He K, Zhang X, Ren S, et al. Deep residual learning for image recognition[C]/IEEE Conference on Computer Vision & Pattern Recognition. IEEE Computer Society,2016.
- [16] Zhao Q, Li B Q, Li T W. Target detection algorithm based on improved YOLO v3[J]. Laser & Optoelectronics Progress, 2020, 57(12): 121502.
赵琼,李宝清,李唐薇.基于改进 YOLO v3 的目标检测算法[J].激光与光电子学进展,2020,57(12):121502.
- [17] Wang Q, Wu B, Zhu P, et al. ECA-Net: efficient channel attention for deep convolutional neural networks[C]/2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2020.
- [18] Lin T Y, Goyal P, Girshick R, et al. Focal Loss for Dense Object Detection[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2020, 42(2): 318-327.
- [19] Zhou X, Wang D, Krahenbuhl P, et al. Object as points[J]. arXiv: Computer Vision and Pattern Recognition, 2019.

网络首发：

标题：基于YOLOv3改进的火焰目标检测算法

作者：张彬彬,帕孜来·马合木提

收稿日期：2021-01-20

录用日期：2021-03-04

DOI：10.3788/llop58.2411003

引用格式：

张彬彬,帕孜来·马合木提. 基于YOLOv3改进的火焰目标检测算法[J]. 激光与光电子学进展, 2021, 58(24): 2411003.

网络首发文章内容与正式出版的有细微差别, 请以正式出版文件为准!

您感兴趣的其他相关论文：

红外单帧图像弱小目标检测技术综述

王好贤 董衡 周志权

哈尔滨工业大学(威海)信息与电气工程学院, 山东 威海 264209

激光与光电子学进展, 2019, 56(8): 080001

基于深度学习航拍图像检测的梯度聚类算法

解博 朱斌 张宏伟 马旗 张扬

国防科技大学脉冲功率激光技术国家重点实验室, 安徽 合肥 230037

激光与光电子学进展, 2019, 56(6): 061007

基于卷积神经网络的遥感图像目标检测

欧攀 张正 路奎 刘泽阳

北京航空航天大学仪器科学与光电工程学院, 北京 100191

激光与光电子学进展, 2019, 56(5): 051002

基于改进特征金字塔的Mask R-CNN目标检测方法

任之俊 蔺素珍 李大威 王丽芳 左健宏

中北大学大数据学院, 山西 太原 030051

激光与光电子学进展, 2019, 56(4): 041502

基于马尔可夫随机场的植被环境中的障碍物识别

程子阳 任国全 张银

陆军工程大学车辆与电气工程系, 河北 石家庄 050003

激光与光电子学进展, 2019, 56(3): 031010