

文章编号: 2095-2163(2023)10-0137-06

中图分类号: TH391.41

文献标志码: A

融合 BiFPN 与 YOLO v5 网络的工厂火灾检测

尚明鹏^{1,2}, 周敏^{1,2}, 陈燕军^{1,2}, 李鑫炎^{1,2}

(1 武汉科技大学 冶金装备及其控制教育部重点实验室, 武汉 430081;

2 武汉科技大学 机械传动与制造工程湖北省重点实验室, 武汉 430081)

摘要: 针对工厂火灾检测中存在环境复杂、目标密集、火焰初期目标较小、样本像素较低以及火焰边界特征不明显等问题, 提出一种基于改进 YOLO v5 的工厂火灾检测算法 Bi-YOLO v5。该方法通过融合加权双向特征金字塔(BiFPN)来增强特征信息;通过锚框参数优化增强了网络的泛化能力;通过损失函数改进提升了网络的收敛速度。实验结果表明,相比 YOLO v5 算法 Bi-YOLO v5 模型的准确率和平均精度分别提高了 2.2% 和 1.7%, 并且每帧推理时间降低到了 27 ms, 达到了在复杂的工厂环境下对火灾检测的要求。

关键词: 工厂火灾检测; YOLO v5 算法; 加权双向特征金字塔; 锚框参数; 损失函数

Factory fire detection with BiFPN and YOLO v5

SHANG Mingpeng^{1,2}, ZHOU Min^{1,2}, CHEN Yanjun^{1,2}, LI Xinyan^{1,2}

(1 Key Laboratory of Ministry of Education for Metallurgical Equipment and Control, Wuhan University of Science and Technology, Wuhan 430081, China; 2 Key Laboratory of Mechanical and Manufacturing Engineering Transmission Hubei Province, Wuhan University of Science and Technology, Wuhan 430081, China)

[Abstract] In factory fire detection, there are problems such as complex environment, dense targets, small targets at the initial stage of the flame, low sample pixels, and inconspicuous flame boundary characteristics. Aiming at above problems, a factory fire detection algorithm Bi-YOLO v5 based on improved YOLO v5 is proposed. This proposed method enhances feature information by fusing weighted bidirectional feature pyramids (BiFPN) and improves the generalization ability of the network through the optimization of anchor frame parameters. The loss function is improved to boost the convergence of the network. Experimental results show that the recall rate and average accuracy of the Bi-YOLO v5 model are increased by 2.2% and 1.7% compared with YOLO v5, and the inference time per frame reaches 27ms, which meets the requirements for fire detection in complex factory environments.

[Key words] factory fire detection; YOLO v5 algorithm; weighted two-way feature pyramid; anchor box parameters; loss function

0 引言

制造业拥有十分繁杂的生产工艺,而机械加工设备一旦出现电器过载或线路绝缘层老化以及短路等情况,就会发生工厂火灾^[1],对工业企业带来巨大的危害。及时发现火情和控制火势,能够尽可能的减少火灾所造成的损失,因此工厂火灾的早期检测和预警有着重要意义。

随着近年来图像处理技术的发展,由于基于机

器视觉的检测方法其检测速率快且受环境因素干扰程度小等优点,成为众学者研究的焦点^[2]。胡勒等^[3]依据阴燃和火羽流阶段的热物理现象,提出了一种基于动态纹理的火焰检测算法;邓理文等^[4]通过选取温度、烟雾浓度和 CO 浓度 3 种参数,基于模糊神经网络对火灾进行探测。但是,火灾探测传感器不仅会受到空气湿度、温度以及气流等环境因素的影响,而且工厂环境复杂不利于传感器放置,所以传感器往往无法及时检测到火灾。此外,一些学者

基金项目: 国家自然科学基金(51975431)。

作者简介: 尚明鹏(1997-),男,硕士研究生,主要研究方向:机器视觉、制造业信息化;周敏(1967-),女,博士,教授,博士生导师,主要研究方向:制造业信息化、智能制造、工业工程与管理;陈燕军(1989-),男,博士研究生,主要研究方向:智能制造理论及应用;李鑫炎(1998-),男,硕士研究生,主要研究方向:机器视觉、机械设计制造。

通讯作者: 周敏 Email:zhoumin@wust.edu.cn

收稿日期: 2022-10-31

尝试通过人工选取特征来对火焰进行描述并结合浅层机器学习模型进行训练和分类。如:TRUONG T^[5] 基于多阶段模式提取火焰特征;DIMITROPOULOS K^[6] 基于动态纹理分析来提取特征;KONG S^[7] 和 HAN X^[8] 通过回归和混合背景高斯模型来进行机器学习模型的构建等等。但这些检测方法的效果取决于人工特征选取,模型识别率较低。随后,孟令昀^[9] 对火焰检测算法进行了特征优化,但该算法的泛化性还有待改进;富雅捷等^[10] 采用迁移学习结合卷积神经网络的方法来训练火灾数据,该算法虽然误报率低但检测速度有待提高。

在目标检测领域中,基于区域的双阶段算法(如 Faster R-CNN^[11] 系列算法),虽然精度较高,但其大量预选框的需求导致了检测速度极慢。基于回归的单阶段算法(如 SDD^[12] 算法、YOLO^[13] 系列算法),推理速度更快,但其高级语义特征较多,且网络后部分特征图尺寸较小,所以在小目标检测上存

在不足。而工厂火灾检测中存在环境复杂、目标密集、火焰初期目标较小、样本像素较低以及火焰边界特征不明显等问题,因此以上算法对于工厂火灾的检测效果并不理想。

针对上述问题,本文基于目前较为优秀的 YOLO v5 模型算法,提出了一种融合了 BiFPN^[14] 和 YOLO v5 网络(以下简称 Bi-YOLO v5)的工厂火灾检测算法。

1 基于 YOLO v5 的工厂火灾检测模型构建

YOLO v5 是 Ultralytics LLC 公司基于 Pytorch 框架提出的轻量级目标检测算法,在单阶段检测算法中比较成熟,相比其它目标检测模型,体积更加轻便且兼具准确度和检测速率,适合嵌入在工厂火灾探测设备中。YOLO v5 将网络结构分成 Input、Backbone、Neck 和 Prediction 4 个部分,其网络结构如图 1 所示。

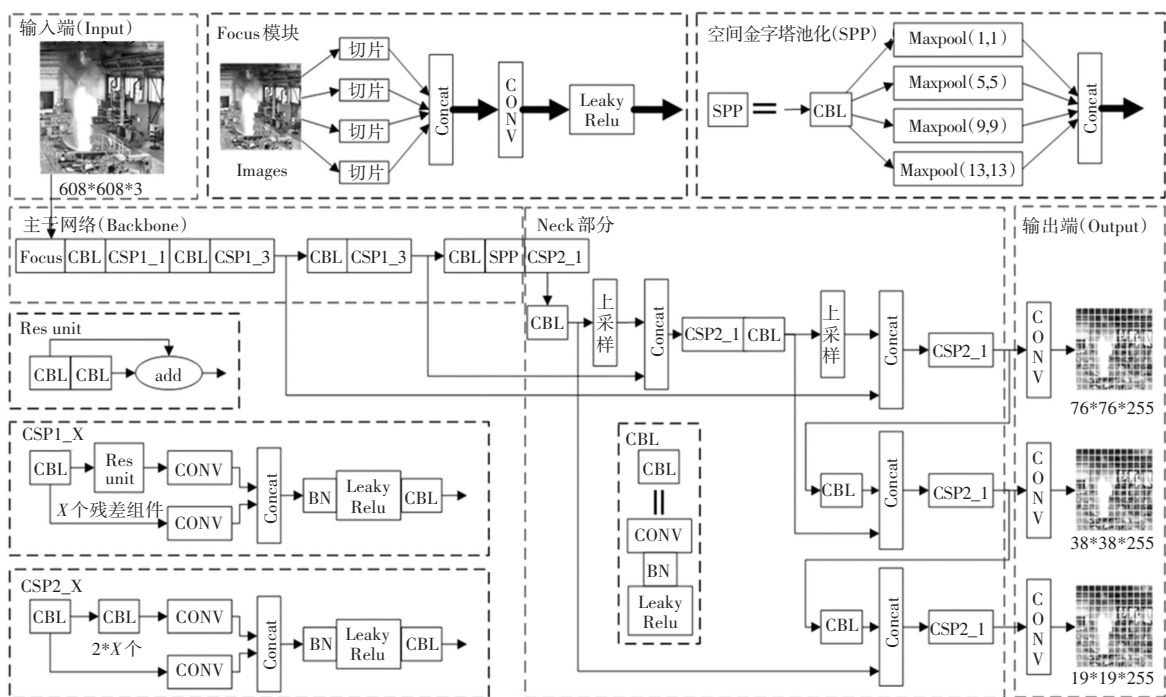


图1 YOLO v5 网络结构

Fig. 1 The structure of YOLO v5 network

YOLO v5 目标检测模型的输入端(Input)主要包含自适应锚框值选取、自适应缩放图像和马赛克(Mosaic)数据增强技术 3 个模块。模型主干网络(Backbone)是在分层图像细粒度上聚合生成图像特征的卷积神经网络,主要包含了 Focus、跨阶段局部融合网络(Cross Stage Partial Networks, CSPNet)和

空间金字塔池化(Spatial Pyramid Pooling, SPP)模块。其中,第一层 Focus 从输入端图像中相邻的 4 个位置进行堆叠,在 C 通道空间中聚焦 W、H 维度信息,能够减少计算量。CSPNet 利用密集跳层连接来进行局部跨层融合,获取多层次的特征图。最后一层 SPP 使用 5、9、13 的 Maxpool 最大池化层,在低

速度差的情况下进行 Concat 融合来提高感受野,有助于特征层对齐。Neck 部分采用特征金字塔结构 (Feature Pyramid Networks, FPN)^[15] 和路径聚合网络结构 (Path Aggregation Networks, PAN), 以加强网络结构对多种缩放尺度的图像特征融合。FPN 通过自上向下、自下向上的高低层特征融合,提高了对小目标的检测准确度;而 PAN 结合不同层的 CSPNet 模块对输出的深浅层特征进行聚合,减少底层定位信息向顶层传递过程中的损耗。输出端 (Output) 使用 GIOU_Loss^[16] 函数作为损失函数,相比于之前 YOLO 版本的 IOU_Loss 函数,采用非极大值抑制 (NMS) 对多目标的边界框进行筛选,增加了相交尺度衡量,提升了遮挡目标和多目标的检测能力。

2 模型改进与优化

2.1 特征金字塔网络融合

由于网络层次连续加深,在特征语义由低维向高维转换过程中,每经过一层都会丢失部分特征信息。为了丰富特征语义信息,实现多层级的特征融合,构造特征金字塔是目前主流的融合思想。

FPN 结构如图 2(a) 所示,该结构搭建了一条自顶向下的通路,把更抽象的高语义信息特征图进行上采样,融合后的语义信息通过在两层相同尺寸的特征空间中横向连接传播。但是,单向的信息流通在一定程度上限制了底层位置信息的传递。针对于此,PAN 结构 (图 2(b)) 在自顶向下通路的基础上添加了一条自底向上的通路,使预测层中也可以对底层的位置信息进行定位,这样预测层就兼具了底层位置信息和顶层语义信息,提升了目标检测的准确度。

加权双向特征金字塔网络结构 (Bidirectional Feature Pyramid Network, BiFPN) 是由谷歌大脑团队在 PAN 的基础上提出的双向融合方式 (图 2(c))。BiFPN 基于路径增强的双向融合思想,实现了高效的双向跨尺度连接,搭建了自顶而下和自底向上的特征融合通道。由于在多尺度特征图的融合过程中,输出特征的重要程度受到输入特征在不同分辨率下的影响,所以在每个节点设置权重来平衡多尺度的特征信息,能够进一步提高检测精度。

基于该结构的优点,本文将原 YOLO v5 特征金字塔网络中的 PAN 结构改为 BiFPN 结构,以增强特征融合,提高检测精度。融合了 BiFPN 的 YOLO v5 网络的特征金字塔 (如图 3 所示),从 Backbone 的

Conv 4、Conv 5 和 Conv 6 层中提取 3 种不同尺度的特征进行通道压缩,然后进行跨尺度连接和加权特征融合,并在最后设置特征分辨率为 19×19 、 38×38 和 76×76 的预测分支,减少了特征融合时小目标信息的损失。

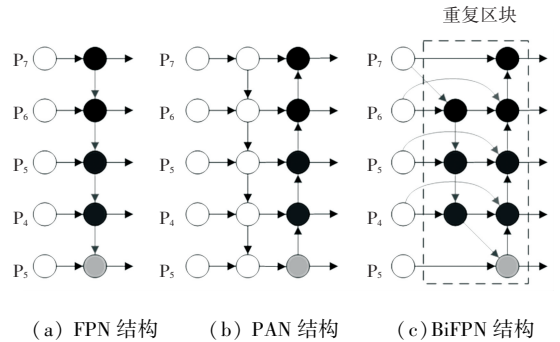


图 2 FPN、PAN 和 BiFPN 结构

Fig. 2 The structure of FPN, PAN and BiFPN

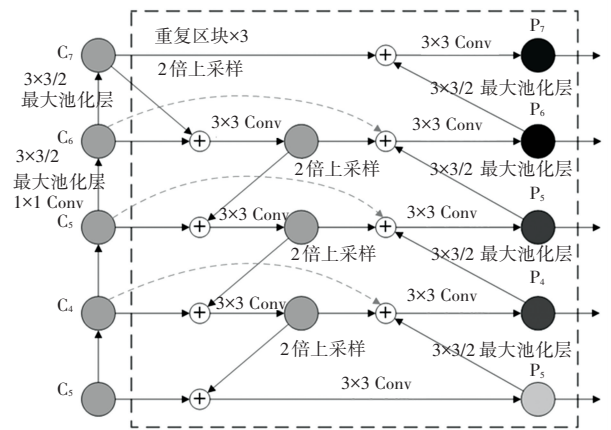


图 3 Bi-YOLO v5 的特征金字塔结构

Fig. 3 The structure of feature pyramid in Bi-YOLO v5

2.2 锚框参数优化

原 YOLO v5 中的自适应计算锚框值的思想,使用 K 均值 (K-Means) 聚类算法,依照标注好的目标框来自动计算并选取适合的锚定框,该过程在模型训练数据集时自动完成。

由于公开数据集检测的目标尺寸与自定义数据集有所不同,YOLO v5 根据 MSCOCO (Microsoft Common Objects in Context) 数据集重新自动学习并设置了 9 种锚定框,从大到小依次为 (373, 326)、(156, 198)、(116, 90)、(59, 119)、(62, 45)、(30, 61)、(33, 23)、(16, 30) 和 (10, 13)。由于工厂环境复杂且初期火焰目标较小,所以本文添加了用于检测工厂火灾的 3 种小尺寸的锚定框,分别为 (16, 14)、(9, 15) 和 (7, 8)。依据检测层不同的尺度对锚

定框进行归类,统计出的锚定框分配情况见表1,用于检测目标较小的初期火焰。

表1 锚定框分配表

Tab. 1 Anchor box allocation

特征图	19×19	38×38	76×76	
感受野	大	中	小	
	(116,90)	(30,61)	(10,13)	(7,8)
锚定框	(156,198)	(62,45)	(16,30)	(9,15)
	(373,326)	(59,119)	(33,23)	(16,14)

2.3 边框损失函数改进

由于火势的大小受到燃烧环境和燃烧时长的影响,造成了火焰边界框尺寸在不同尺度图像中存在差异,所以对边框损失函数进行改进和优化。虽然 GIOU 函数对 IOU 的梯度问题进行了改进,但仍存在收敛效率慢和回归不稳定的弊端。因此,本文采用 CIOU 函数来为 bounding box 计算边框回归损失,如式(1)~式(3)所示:

$$L_{CIOU} = 1 - IOU + \frac{\rho^2(b, b^{gt})}{c^2} + av \quad (1)$$

$$a = \frac{v}{(1 - IOU) + v} \quad (2)$$

$$v = \frac{4}{\pi^2} \left(\arctan \frac{w^{gt}}{h^{gt}} - \arctan \frac{w}{h} \right)^2 \quad (3)$$

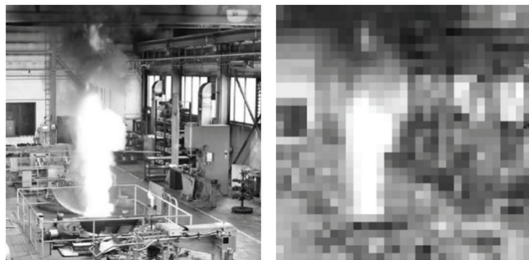
式中: ρ 代表计算预测框与实际框两个中心点间的欧式距离, b, b^{gt} 分别代表预测框和实际框的中心点, c 是最小包围预测框和实际框的对角线距离, α 是权重参数, v 是用来计算预测框和实际框之间宽高比的相似性, w, w^{gt} 分别是预测框和实际框的宽度, h, h^{gt} 分别是预测框和实际框的高度。

CIOU 纳入了边界回归的中心点距离、重叠面积和纵横比,多维度的参数更好的反映了预测框和实际框的差异,提高了收敛速度和预测精度。

3 实验与结果分析

3.1 数据集

本文通过 Python 编写脚本从网络上获取火灾图片,并使用缩放、翻转、剪切、加噪、平移和镜像等方法对数据集进行数据扩充,经扩充后的样本数量为 8 435 张图片。对自建数据集通过增加图片饱和度和调整色彩空间和 Mosaic 数据增强技术,来获得更好的训练效果。图 4 为对火灾数据集样本进行 Mosaic 数据增强的实例。处理后的数据集保证了本文提出的 Bi-YOLO v5 模型能在复杂的工厂环境中良好泛化。



(a) Mosaic 数据增强前 (b) Mosaic 数据增强后

图4 Mosaic 数据增强实例

Fig. 4 The instance of Mosaic data enhancement

3.2 评价指标

为了验证 Bi-YOLO v5 模型的性能,本文用真阳性 (True Positive, TP)、假阳性 (False Positive, FP) 和假阴性 (False Negative, FN) 来计算召回率 ($Recall$)、准确率 ($Precision$)、平均精度 (Average Precision, AP) 和平均精度均值 (Mean Average Precision, mAP), 以此作为评价指标。

$Recall$ 表示真阳性和真阳性与假阴性之和的比值,如式(4)所示:

$$Recall = \frac{TP}{TP + FN} \quad (4)$$

$Precision$ 表示真阳性占所有阳性的比值,如式(5)所示:

$$Precision = \frac{TP}{TP + FP} \quad (5)$$

AP 用来衡量模型在每个类别上的学习程度,其数学意义是以 $Recall$ 为横轴和以 $Precision$ 为纵轴围成的曲线面积,如式(6)所示:

$$AP = \int_0^1 P(R) \quad (6)$$

mAP 表示模型里所有类别的 AP 均值,如式(7)所示:

$$mAP = \frac{1}{n} \sum_{i=1}^n AP_i \quad (7)$$

其中, n 代表模型里所有类别的个数。

3.3 模型训练

本文将火灾数据集按照 8 : 2 的比例随机划分成训练集和测试集。其中训练集和测试集的数量分别为 6 748 和 1 687。采用 608×608 作为输入图像尺度,并使用随机梯度下降 (Stochastic Gradient Descent, SGD) 函数来为模型的网络参数进行优化。训练参数设置见表 2,本次实验共训练 4 个模型,分别是 Faster R-CNN、SDD、YOLO v5 和本文提出的 Bi-YOLO v5。

表 2 参数设置表
Tab. 2 Parameter settings

参数名称	参数值
批量大小 (Batch_size)	32
迭代轮数 (Epochs)	100
动量 (Momentum)	0.9
权重衰减 (Weight_decay)	0.000 5
学习率 (Learning_rate)	0.01

3.4 实验结果分析

根据日志文件绘制出 Bi-YOLO v5 和 YOLO v5 模型在训练过程中的 Loss 值对比曲线, 如图 5 所示。

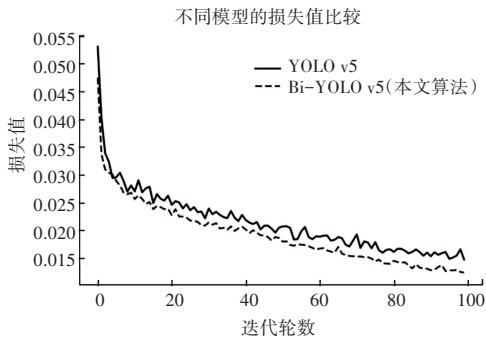


图 5 Loss 值对比曲线图

Fig. 5 Comparison curve of Loss

由图可以看出, Bi-YOLO v5 的 Loss 值曲线始终在 YOLO v5 下方, 且迭代 30 轮后 Loss 值变化更为平缓, 说明 Bi-YOLO v5 模型比 YOLO v5 模型收敛更快, 损失值更小。

此外, 通过对比召回率、准确率、平均精度和每帧推理时间, 验证了几种现有算法和 Bi-YOLO v5 的性能, 结果见表 3。

表 3 与现有算法的对比

Tab. 3 Comparison with existing algorithms

算法	召回率/ %	准确率/ %	平均 精度/%	每帧推理 时间/ms
Faster R-CNN	89.2	90.7	92.1	124
SDD	76.8	78.0	86.4	89
YOLO v5	94.3	96.2	97.5	28
Bi-YOLO v5 (本文算法)	96.1	98.4	99.2	27

由表 3 可知, 在相同数据集中, Bi-YOLO v5 的平均精度为 99.2%, 相比于 Faster R-CNN、SDD 和 YOLO v5 模型, 平均精度分别提高了 7.1%、12.8% 和 1.7%。并且改进后算法的每帧推理时间降低到了 27 ms。综合来看, 本文提出的 Bi-YOLO v5 更好的满

足了火灾检测对于检测精度和实时性上的需求。

3.5 实际场景检测

为了展示 Bi-YOLO v5 对工厂火灾的检测效果, 在网络中随机选取一些背景复杂且像素较低的工厂监控视频, 以及暗黑环境下的火灾图片进行了检测, 检测结果如图 6 所示。

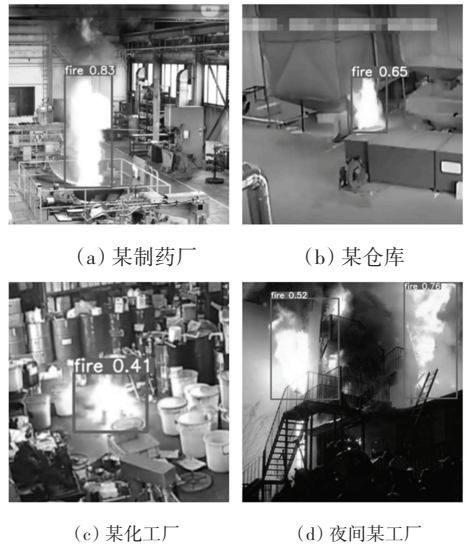


图 6 实际火灾现场检测结果图例

Fig. 6 Fire detection in practical scenes

由检测效果可见, 无论是白天还是夜晚, Bi-YOLO v5 都能在环境复杂的工厂中通过像素不高的监控系统准确的识别火焰, 体现了模型优越的性能以及良好的泛化能力。

4 结束语

针对工厂环境复杂、目标密集、火焰初期目标较小、样本像素较低以及火焰边界特征不明显等问题, 本文提出了一种 Bi-YOLO v5 算法模型。该算法通过融合 BiFPN 和 YOLO v5 的特征网络、锚框参数优化和损失函数改进来提高模型对多尺度火焰的检测效率, 进而提高了火灾检测在工厂环境中的泛化能力。实验结果表明, Bi-YOLO v5 模型的准确率和平均精度比 YOLO v5 提高了 2.2% 和 1.7%, 并且每帧推理时间降低到了 27 ms。与现有算法相比, Bi-YOLO v5 的准确率最高且时间最快。综上, 本文算法为进一步部署工厂火灾监控系统提供了模型支持, 对工厂火灾检测系统的推广有着实际意义。

参考文献

[1] 马永昌. 机械制造业消防安全管理问题与对策探索[J]. 科技创新导报, 2020, 17(5): 73-74.