DOI:10.16136/j.joel.2023.08.0406

基于改进 YOLOV4 的铁轨裂缝目标检测算法

苗新法*,李晓琴,刘宝莲,侯 越

(兰州交通大学 电子与信息工程学院,甘肃 兰州 730070)

摘要:针对铁轨表面裂缝的小目标特征及传统检测方法精度低,速度慢等问题,提出一种基于改进YOLOV4的目标检测算法。首先,使用改进的RFB(receptive field block)模块替换空间金字塔池化(spatial pyramid pooling,SPP)结构,以获取特征图更大的有效感受野区域,提升算法的检测精度;其次,采用深度可分离卷积结构替代网络模型中的普通卷积结构,使网络轻量化、提升检测速度;同时,利用 K-means++算法重新获取锚框,再对得到的锚框进行线性尺度变化,解决原锚框不适合小目标检测的问题。结果表明改进的YOLOV4算法,平均精度均值(mean average precision, *mAP*)达到 84.8%,相对于原 YOLOV4 算法提高了 3.4%;检测速度(frames per second,*FPS*)为 62.39 帧/s,提高了 4.07 帧/s。

关键词:目标检测;裂缝检测;RFB模块;深度可分离卷积;K-means++ 中图分类号:TP391 文献标识码:A 文章编号:1005-0086(2023)08-0816-07

Research on rail crack detection algorithm based on improved YOLOV4

MIAO Xinfa*, LI Xiaoqin, LIU Baolian, HOU Yue

(College of Electronic and Information Engineering , Lanzhou Jiaotong University , Lanzhou , Gansu 730070, China)

Abstract: Aiming at the small target characteristic of rail surface crack and the low precision and slow speed of traditional detection methods, we propose an object detection method based on improved YOLOV4 network for cracks on the surface of rails in this paper. Firstly, in order to obtain the larger effective receptive field area of the feature map and improve the detection accuracy, we use the improved receptive field block (RFB) module to replace the spatial pyramid pooling (SPP) structure; Secondly, we use the deep separable convolution structure to replace the common convolution structure in the network model, so that the network is lightweight and the detection speed is improved; At the same time, we use K-means + algorithm to reacquire the anchor frame, and then change the linear scale of the anchor frame to solve the problem that the original anchor frames are not suitable for small target detection. The results show that the mean average precision (mAP) of the improved YOLOV4 is 84.8%, which is 3.4% higher than that of the original YOLOV4 algorithm; The detection speed (FPS) is 62.39 frame/s, which increases by 4.07 frame/s,

Key words:object detection; crack detection; receptive field block (RFB) module; depth-separable convolution; K-means++

0 引 言

轨道检测是维护铁路安全运营的重要手段, 截至 2021 年 12 月,中国铁路的总运营里程超十 五万公里。铁轨长时间承受着高速重压,再加上 外部自然环境的影响,极易产生裂纹等结构损伤, 若不及时对裂纹检测维修,将给铁路的运行安全 带来极大的威胁。因此,对铁轨裂缝进行及时准 确的检测是十分有必要的。

当前,传统的测定方法有人工目检法、电涡流法、超声波法和磁粉法等,但这些方法所测定的精确度较低,可靠性差且操作复杂,无法满足对铁轨

 ^{*} E-mail:1459084294@qq.com
 收稿日期:2022-05-30 修订日期:2022-09-10
 基金项目:国家自然科学基金(62063014)资助项目

裂缝实时检测的要求^[1]。卷积神经网络概念的提出,给目标检测打开了新研究方向的大门,大量研究者开始将算法与其结合运用于各种场景的目标检测与识别中。采用深度学习的卷积神经网络检测方式具有高非接触性、检测速率快、精度高、可移植性强等优势。目前,常用的检测算法主要有R-CNN(region conventional neural network)、Faster-RCNN^[2]系列的双阶段算法和 SSD(single shot multibox detector)^[3]、YOLO(you only look once)^[4,5]系列的单阶段检测方法。前者是通过区域生成网络将待检图片转换成含有检测对象特征的候选框,然后再对候选区域进行特征提取进而实现识别分类。后者不要求采集候选区的特征,而是直接提取被检对象的特征信息,此类算法虽牺牲了少量精度却大大提高了检测的速度。

近年来,视觉领域的目标检测掀起研究的热 潮,其中就有大量学者将其应用于铁轨表面的缺 陷研究。文献[6]针对铁轨固件缺陷,将传统的图 像处理方法与深度学习网络 Faster-RCNN 结合, 实现轨道扣检缺陷的智能识别检测。文献[7]提 出一种基于 DenseNet201 的深度迁移学习扣件检 测方法,通过重组模型结果达到最佳的注意力分 布,利用迁移学习使模型训练达到较好的鲁棒性。 文献[8]通过融合两种轻量级的网络 SqueezeNet 和 MobileNetV2 来检测钢轨表面的缺陷,在低对 比度下取得了较好的结果。文献「9]提出了基于 YOLOV3 的改进模型用于铁路轨道的缺陷检测, 可以有效地监测铁路轨道故障的状态。以上方法 利用不同的卷积神经网络针对铁轨表面缺陷进行 检测,并侧重不同方面进行检测算法的改进,使得 检测精度有了不错的提升,模型的运行速度也得 到了有效的保证,但目前针对铁轨裂纹的目标检 测研究的较少,且在铁轨裂纹图像中目标像素占 比较小,目标特征不明显,一些细小裂纹与背景特 征高度相似,使得网络提取出的有用特征信息少, 导致检测的准确度低,甚至出现小目标检测丢失 等问题。

对此,本文提出一种改进 YOLOV4 的铁轨裂 缝目标检测算法。YOLOV4 以其较快的检测速 度和精确的检测能力征服了世人,成为在目标检 测精度要求较高场所中的优先考虑者。在 YOLOV4 算法的基础上,本文在颈部网络使用改 进的 RFB(receptive field block)结构替代空间金字 塔池化(spatial pyramid pooling, SPP)结构,以获得 更大的有效感受野,使网络更好地保留特征图中 深层次小尺寸的裂缝目标;其次,将网络中的普通 卷积结构替换为深度可分离卷积,减少网络参数 量以此提升系统的处理速度。同时使用 K-means ++算法结合 IOU(intersection over union)损失函 数重聚锚框,再根据数据集的特点将得到的锚框 进行尺度伸缩变换,以此来克服所标记的目标框 尺寸较集中的问题,提高整体的检测精度,从而有 效地检测铁轨表面的裂缝。

1 原 YOLOV4 算法

YOLOV4 算法是在 YOLOV3 的基础上改进的 单阶段检测算法,它在同一个网络中完成物体的分 类并对目标进行定位^[10]。YOLOV4 网络结构主要 由特征提取网络、SPP结构、路径聚合网络(path aggregation network, PANet) 和头部检测网络构成。 CSPDarknet53 是其特征提取网络,它是由一系列的 多次残差网络结构构成,这些残差组件增强了卷积 网络的特征提取能力,残差块的堆叠方式解决了网 络深度与梯度之间难以平衡的问题[11],同时主干网 络采用 mish 激活函数。颈部的 SPP 结构则通过最 大池化的方式将特征图融合,从而使得网络的感受 野得到了极大的拓展。PANet 是在特征金字塔网络 (feature pyramid network, FPN)中加入下采样的路 径增强结构来加强网络的特征提取能力。PANet 从 不同的检测层进行特征数据的融合,增强了目标特 征的提取。最后再对输出的特征层进行目标分类并 定位位置坐标,其整体结构图如图1所示。

2 改进的 YOLOV4 算法

2.1 SPP 结构替换为改进的 RFB 模块

SPP模块对特征图通过 4 个不同尺寸的最大池 化处理,突出不同尺度下的数据特征,增加了网络的 感受野,不仅能保持图像的尺度不变,而且也减少了 网络过拟合的现象^[12]。但是由于铁轨表面有很多细 小的裂缝,且它们占整幅图的像素比重小,在不断下 采样和特征提取的过程中使得能够提取的裂缝特征 信息较少,致使小目标漏检。而 RFB模块将输入特 征进行多分支空洞卷积,并加入了输入与输出直连 的一个大的残差边,使得模型易优化且增强了深层 次的有效感受野区域,对复杂特征的小目标有较好 的学习能力。本文将 SPP 结构替换为改进的 RFB 模块,并将其添加在尺度为 26 和 52 分支的颈部网络 输入端。

图 2 所示为 RFB 模块,它由多个分支结构构成, 每一分支上由不同尺度的空洞卷积和普通卷积共同 参与特征运算。输出的特征图相同时,使用两种小 尺寸的卷积核进行级联计算时比使用一种大尺寸卷 积核所需的参数量更少,而采用多分支卷积使模块 有较好的非线性结构;多尺度空洞卷积是为了在保 持参数量不变的情况下增大有效感受野,优化图像 相邻间的冗余信息,从而获得更高分辨率的特征^[13]。 原 RFB模块共有 3 个分支与一个直连残差边,各分 支尺度分别为 1、3、5。特征图输入后先通过 1×1 的 卷积层来调控通道数,然后在各分支上进行不同尺 度的普通卷积与不同膨胀率的空洞卷积计算,最后 将 3 个分支结果进行融合后与直连残差边拼接输 出。本文改进的 RFB 模块是在原有模块的基础上增 加了一个尺度为 7 的分支。同时,为了加深非线性 层,使用1×n加n×1的卷积层替代原来n×n的卷



Fig. 1 The network structure of the YOLOV4 algorithm











2.2 卷积方式的改进

YOLOV4 模型的网络结构较深,在检测的过程 中通过不断地下采样来获取特征信息,虽然网络的 加深有利于提取更多的细节信息,但同时增加了复 杂度与计算量,使得检测的过程耗时长,计算机承担 较大的运算负荷。为了轻量化网络模型且保持较高 分辨率,使用深度可分离卷积结构来代替普通的卷 积结构^[15]。深度可分离卷积是将整个卷积过程分步 进行,先进行每个通道的卷积再进行每个点的卷积。

图 4 所示为普通卷积的过程示意图,若输入层 图片的大小为 H×W、M 通道的图片,经过 N 个大 小为 K×K 的标准卷积核运算后,卷积层的参数量 C1 为:



Fig. 4 Ordinary convolution

(1)

 $C1 = M \times N \times K \times K_{\circ}$

图 5 为深度可分离卷积的示意图,首先将大小 为 H×W×M 的输入图像进行单通道卷积,通过 M 个大小为K×K×1 的卷积核,从而得到 M 个特征 图。该过程中参数计算量 Cd 为:

$$Cd = M \times K \times K_{\circ} \tag{2}$$

然后,再使用 N 个大小为 1×1×M 的卷积核调 控深度,该过程中卷积运算的参数量 Cp 为:

$$Cp = 1 \times 1 \times M \times N_{\circ}$$
 (3)
综上可得深度可分离卷积的参数量 C2 为:

$$C2 = Cd + Cp = M \times K \times K +$$

$$1 \times 1 \times M \times N_{\circ}$$
 (4)

两种卷积方式的参数量之比表示为:

$$\frac{C2}{C1} = \frac{1}{N} + \frac{1}{K^2}.$$
 (5)

由上述推理可以看出卷积方式的改进减少了网 络的参数量,从根本上加快了算法的检测速度。



2.3 改进 K-means++聚类锚框

在 YOLOV4 算法中,真实框的聚类采取的是 Kmeans 算法,但 K-means 的初始聚类点选取随机,会 造成聚类偏差,使最终得到的聚类中心与实际位置 相差太远,易形成局部最优,使得锚框与真实框不匹 配,影响检测的准确率,故利用 K-means++聚类算 法结合 IOU 损失函数来选择锚框,K-means++算 法的随机性小,使距离较远的点较易被选中,克服了 K-means 对初始点选取的敏感性^[16]。

通过聚类方法产生的锚框跟数据集本身有着很 大的关系,本文数据类型单一,且标记框的大小差距 不大,尺寸比较集中,所以聚类得到的锚框也很集 中,这就体现不出模型的多尺度训练优势。所以在 通过聚类产生锚框后,再进行线性尺度缩放,将锚框 尺寸伸缩变换,使其能够更贴近真实框的位置。其 计算式如下:

$$X'_{\min} = \alpha X_{\min}, \qquad (6)$$

$$X'_{\max} = \beta X_{\max}, \qquad (7)$$

$$X'_{i} = [(X_{I} - X_{\max})(X_{I} - X_{\min})](X'X_{\max} - X'_{\min}) + X'_{\min}, \qquad (8)$$

$$Y'_{i} = X'_{\min} \times \frac{Y_{i}}{X_{i}}, \qquad (9)$$

式中, X'_{min} 表示缩放后的最小锚框, X'_{max} 表示缩放 后的最大锚框, α 和β分别为缩放倍数, X'_i 和 Y'_i 代表 缩放后锚框的 x 轴和 y 轴坐标。图 6 展示了改进方 法后的聚类中心图, 改进的 K-means + +聚类算法 的输入为数据集中所有标定的锚框数据, 图中的横 坐标和纵坐标分别代表锚框的宽和高, 由图可以看 出中心点在整个数据集上分布均匀。



Fig. 6 Improved clustering center graph for K-means++

3 实验结果分析

3.1 实验环境及参数设定

本文使用的铁轨裂缝数据集包含 2526 张图片。 训练集与验证集按照 8:2 比例进行划分。实验是 在云服务器 AutoDL上进行,镜像环境为 python 3.8、pytorch 1.10、cuda 11.3,实验配置为 GPU: RTX3090,显存 24 GB。对比实验配置环境相同。 本文设定样本的整个训练过程为 300 个 epoch,前 50 个 epoch采用冻结训练方式,后 250 个 epoch进行解 冻训练,网络训练时参数优化方法选择随机梯度下 降法,本文模型训练损失曲线如图7所示,由图7可



以看出迭代次数为50次左右时,训练过程达到了 收敛。

3.2 评价准则

表1为二分类的混淆矩阵用于评价分类 质量^[17]。

表1 混淆矩阵

Tab. 1 Confounding matrix

	Condition positive	Condition negative
	condition positive	condition negative
Test outcome positive	True positive(TP)	False negative(FN)
Test outcome negative	False positive(FP)	True negative(TN)

1) 准确率:正确检测的样本占检测总数的比例, 其式如下:

$$P = \frac{TP}{TP + FP} \,^{\circ} \tag{10}$$

2) 召回率:预测为正的样本占实际正样本的比例,计算式如下:

$$R = \frac{TP}{TP + FN} \,^{\circ} \tag{11}$$

3) AP 指数: AP 是精准度, 用来反映模型全局的 性能, 在准确率的基础上考虑了模型对位置的回归 质量, AP 值越大表示模型精度越高, 性能越好, mAP 为平均精度均值(mean average precision, mAP), n 为检测目标类别数, 计算式如下:

$$AP = \int_{0}^{1} P(R) dR, \qquad (12)$$

$$mAP = \frac{\sum AP}{n}_{\circ} \tag{13}$$

 4) FPS(frames per second):每秒处理的帧数, 用来反映模型的检测速度。

$$FPS = \frac{frameNum}{elapsedTime}$$
(14)

3.3 主观评价

为了直观看到本文算法的可行性,将本文算法 与常见的目标检测算法进行对比实验,图 8 为检测 结果。

由图 8 可以看出,YOLOV3 算法对于特征清晰 的裂缝目标有较好的检测效果,但检测的精度偏低 YOLOV4-tiny 对细微小裂纹会有较严重的漏检现 象。YOLOV4 算法的细小裂缝的检测力增强了,但 还是存在细小裂缝的漏检。改进后的 YOLOV4 算 法能更好地识别小目标,对于特征不明显或细小的 裂纹能够准确检测,检测精度也有了进一步的提升。



图 8 铁轨裂缝实验结果对比

Fig. 8 Comparison of the experimental results of the rail cracks

3.4 客观评价

本文采用定量指标的客观数据进一步验证所改 进算法的性能。对图 8 所示的结果进行平均精度均 值 mAP 和检测速度FPS 指标对比分析,数值越高说 明算法的性能越好,表 2 为各种算法的指标数据。

从表 2 中可以看出, YOLOV3 和 YOLOV4 算

法的 mAP 高,但 FPS 慢,是因为这两种算法的参数 多,网络深度较深;YOLOV4-tiny 的 FPS 显著地提 高了,达到了 76.22 帧/s,但 mAP 却较 YOLOV3 有 所降低。而本文改进的 YOLOV4 算法对目标检测 的整体效果都有了一定的提升,虽然 FPS 稍逊于 YOLOV4-tiny,但较原 YOLOV4 算法, mAP 提升了 3.4%, FPS 提升了 4.07 帧/s。

表 2 不同检测算法的结果对比

 Tab. 2 Comparison of results from different

 detection algorithms

uccetion argorithmis			
Model	mAP / %	FPS /(frame/s)	
YOLOV3	76.8	40.41	
YOLOV4-tiny	73.9	76.22	
YOLOV4	81.4	58.32	
Ours	84.8	62.39	

同时,图9更清晰地展现了不同目标检测算法



图 9 不同算法的准确率对比曲线

Fig. 9 Accuracy comparison curves for different algorithms

的准确率曲线变化。由图可以看出,改进后的 YOLOV4曲线在其他算法线的上方,表明改进的 YOLOV4网络拥有较高的准确率,能够较好地满足 铁轨裂缝小目标的识别需要。

3.5 消融实验

为了进一步验证本文改进的各个模块的检测能力,将算法分为4组进行消融实验。第一组为原 YOLOV4模型,"√"与"×"分别表示添加了此模块 和未添加此模块,消融实验对比结果如表3所示。

在表 3 中可以看出, 网络替换为深度可分离卷 积后, 检测的精度没有发生太大的提升, 但网络检测 的速度有较为明显的提高, FPS 提高了 2.61 帧/s, 是有效的网络轻量化方法; 由 2、3 组对比可知, 加入 改进的 RFB 模块后, mAP 和 FPS 分别为 83.5%和 57.31, 检测的精度提升了 2.7 个百分点, 证明添加 改进的 RFB 模块增大了有效感受野, 获得了深层次 的复杂特征, 使得网络的精度有所提升。多加入的 模块不可避免地会引起参数量的增加, 所以对比第 2 组检测速度有稍许的降低; 在第 3 组中加入改进的 K-means++算法后, 网络模型的 mAP 和 FPS 均有 所提升, mAP 值提升了 1.3%, FPS 提升了 3.01 帧/s, 且与第 4 组对比可知, 检测速度的提升并不是 依赖深度可分离卷积参数量减少而产生的。改进的

Depth-separable mAP/%Group Improved RFB Improved K-means++ FPS /(frame/s) convolution 1 \times Х Х 81.4 56.23 2 81.8 \checkmark \times \times 58.84 3 $\sqrt{}$ \times 83.5 57.31 \sim $\sqrt{}$ $\sqrt{}$ 57.68 4 \times 84.3

. /

表 3 消融实验对比 Tab. 3 Comparison of the ablation experiments

聚类算法使得锚框更加符合铁轨裂缝数据的目标, 使网络的定位更加准确,识别度更高,识别速度 更快。

~/

4 结 论

5

本文提出了一种改进的 YOLOV4 的铁轨裂缝 检测算法,使用改进的 RFB 模块提高了网络的有效 感受野,K-means++聚类锚框的伸缩变换使小目标 检测率得到有效提高,最终 mAP 达到 84.8%。深度 可分离卷积的引入大大地降低了网络模型的参数 量,FPS 提高了 2.61 帧/s。在铁轨裂缝数据集上进 行训练验证,并与原 YOLOV4 算法进行对比,结果显示改进后的算法检测速度为 62.39 帧/s, FPS 提高了 4.07 帧/s。本文提出的算法对铁轨裂缝检测的精确度较高,检测时间较短,可以有效地实现铁轨裂缝的精确、快速检测。后续工作将关注数据集的前期处理,对铁轨裂缝小目标数据集进行数据增强并构建大规模数据集,从数据集的处理来进一步提高检测的准确率。

62.39

84.8

参考文献:

based on convolutional neural network [D]. Mianyang: Southwest University of Science and Technology, 2018; 7-9.

刘雄祥.基于卷积神经网络的铁轨表面缺陷识别研究 [D].绵阳:西南科技大学,2018:7-9.

- [2] 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.
- [3] LIU W, ANGUELOV D, ERHAN D, et al. SSD: single shot multibox detector [C]//European Conference on Computer Vision, October 11-14, 2016, Amsterdam, The Netherlands. Cham: Springer, 2016: 21-37.
- [4] REDMON J,FARHADI A. YOLO9000: Better, faster, stronger[C]//IEEE Conference on Computer Vision and Pattern Recognition(CVPR), July 21-26, 2017, Honolulu, HI, USA. NewYork: IEEE, 2017:6517-6525.
- [5] REDMON J, FARHADI A. YOLOv3: an incremental improvement[EB/OL]. (2018-04-08) [2022-05-10]. https: //arxiv.org/abs/1804.02767.
- [6] WEI X K, YANG Z M, LIU Y X, et al. Railway track fastener defect detection based on image processing and deep learning techniques: A comparative study[J]. Engineering Applications of Artificial Intelligence, 2019, 80: 66-81.
- [7] WU S Y, LIU L Y, ZHANG H, et al. Research on defect morphology detection algorithm of railway fastener based on transfer learning[J/OL]. Journal of Railway Science and Engineering. (2022-04-11) [2022-05-16]. https:// doi.org/10.19713/j.cnki.43-1423/u.T20211508. 吴送英,刘林芽,张洪,等. 基于迁移学习的铁路扣件缺 陷形态检测算法研究[J/OL]. 铁道科学与工程报. (2022-04-11)[2022-05-16]. https://doi.org/10.19713/ j.cnki.43-1423/u.T20211508.
- [8] AYDIN I, AKIN E, KARAKOSE M. Defect classification based on deep features for railway tracks in sustainable transportation [J]. Applied Soft Computing, 2021, 111: 107706.
- [9] WAN Z J. CHEN S Q. Railway tracks defects detection based on deep convolution neural networks[M]//Artificial Intelligence in China. Singapore: Springer, 2021:119-129.
- [10] BOCHKOVSKIY A, WANG C Y, LIAO H Y M. YOLOv4:optimal speed and accuracy of object detection [J]. Visio and Pattern Recognition, 2020, 17(9):198-215.

[11] WANG Z Y, ZHANG G, YANG Q, et al. Research on surface defect recognition of copper strip based on YOLOV4
[J]. Journal of Optoelectronics • Laser, 2022, 33 (2): 163-170.
王紫玉,张果,杨奇,等.基于 YOLOv4 的铜带表缺陷识

别研究[J].光电子・激光,2022,33(2):163-170.

- [12] LIU S, HUANG D, WANG Y. Receptive field block net for accurate and fast object detection [C]//European Conference on Computer Vision, September 8-14, 2018, Munich, Germany. Singapore: Springer 2018:404-419.
- [13] CHEN M T, YU S. Research on traffic sign recognition based on improved YOLOV4 model[J]. Microelectronics & Computer, 2021, 39(1):17-25.
 陈梦涛,余粟.基于改进 YOLOV4 模型的交通标志识别 研究[J]. 微电子学与计算机, 2021, 39(1):17-25.
- [14] LUO H, JIA C, LI J. Road surface disease detection algorithm based on improved YOLOV4[J]. Laser & Optoelectronics Progress, 2021, 58(14):1410025.
 罗晖, 贾晨, 李健. 基于改进 YOLOV4 的公路路面病害 检测算法[J]. 激光与光电子学进展, 2021, 58(14): 1410025.
- [15] WEI X Z, ZHAO Q, ZHOU H. Capacitor appearance defect detection based on improved YOLOV3[J]. Journal of Optoelectronics • Laser, 2021, 32(12):1278-1284.
 魏相站,赵麒,周骅.基于改进 YOLOV3 的电容器外观 缺陷检测[J].光电子 • 激光, 2021, 32(12):1278-1284.
- [16] ZHANG B P, KANG Q Z, LI J M, et al. Lightweight improved YOLOV4 target detection algorithm[J]. Computer Engineering, 2021, 48(8): 206-214.
 张宝朋,康谦泽,李佳萌,等轻量化改进的 YOLOV4 目标检测算法[J]. 计算机工程, 2021, 48(8): 206-214.
- [17] SUN W K. Research on crack detection in high resolution image based on deep learning[D]. Jinan: Shandong University, 2021; 29-31.

孙伟凯.基于深度学习的高分辨率图像裂缝检测研究 [D].济南:山东大学,2021:29-31.

作者简介:

苗新法 (1979-),男,硕士,副教授,硕士生导师,主要从事图像处 理、目标识别、电子技术方面的研究.