

基于 YOLOv5 的行李运输系统轮对裂纹检测算法

方浩楠[†], 李登鹏

(中国民航大学 电子信息与自动化学院, 天津 300300)

摘 要:为了检测机场行李运输系统中轮对存在的裂纹,提出了一种基于 YOLOv5 的机场行李运输系统轮对裂纹检测方法。该方法针对裂纹细小、密集等特点采取了以下措施:在 YOLOv5 网络的 Head 部分使用 SIoU 替换了原有的 Ciou;在 Backbone 部分加入了 SE 注意力机制;在 Neck 部分引入了 Swin-Transformer 模块;在整个 YOLOv5 网络中使用 SPD-Conv 替代了传统的 Conv 卷积模块;在图像的预处理方面,使用了图像分割与子图反向拼接技术。通过这些改进,有效地改善了 YOLOv5 对于细小、密集裂纹的特征提取能力,相较于传统的 YOLOv5 算法,裂纹检测的能力得到了有效提升。

关键词:裂纹检测;YOLOv5;SPD-Conv(space-to-depth-Conv);SIoU;注意力机制;图像预处理

中图分类号:TP391.4

文献标识码:A

Wheelset Crack Detection Algorithm of Baggage Transportation System Based on YOLOv5

FANG Haonan[†], LI Dengpeng

(School of Electronic Information and Automation, Civil Aviation University of China, Tianjin 300300, China)

Abstract: In order to detect the cracks existing in the wheelset of airport baggage transportation system, a new wheelset crack detection method based on YOLOv5 was proposed. In view of the characteristics of small and dense cracks, the following measures were taken. The original Ciou is replaced by SIoU in the Head part of YOLOv5 network. Then the SE attention mechanism is added to the Backbone part and the Swin-Transformer module is introduced in the Neck part. In addition, The traditional Conv convolution module is replaced by space-to-depth-Conv in the entire YOLOv5 network. In terms of image preprocessing, image segmentation and subgraph reverse stitching technology are used. Through these improvements, the feature extraction ability of YOLOv5 for fine and dense cracks is effectively improved, and compared with the traditional YOLOv5 algorithm, the ability of crack detection has been effectively improved.

Key words: crack detection; YOLOv5; SPD-Conv(space-to-depth-Conv); SIoU; attention mechanism; image preprocessing

调研显示,我国机场早期建设的行李运输线,如广州白云机场 1 号航站楼内的输送线,其行李分拣输送系统均为滑轮结构,当运送行李的传送带承载行李高速(≥ 1.7 m/s)运动时,关键轮位就会由于同钢轨摩擦发生劣化,产生裂纹。关键轮位上的裂纹会导致整个传送系统运动失衡,从而使整条

输送线停机,给机场造成重大损失。该风险通常被机场判定为行李延误系统性风险,为重点关注防范的重大故障类别。基于此,提出的基于 YOLOv5 的裂纹识别算法具有较重大的实际意义。

目标识别方向上的主流算法大致分为两类,其一是以 R-CNN^[1] (region with CNN features)、

收稿日期:2023-01-18

基金项目:天津市 2022 年大学生创新创业训练计划资助项目(中国民航大学大学生创新创业训练计划资助项目)(202210059165)

作者简介:方浩楠(2002—),男,安徽马鞍山人,本科,研究方向:自动化技术及应用。

[†]通信联系人, E-mail: 1824134874@qq.com

Fast R-CNN^[2] (fast region-based convolutional network)为代表的二阶段(Two-stage)目标识别,另一类是以 YOLO^[3] (you only look once)为代表的一阶段(One-stage)目标检测。一阶段检测不需要生成候选区域(region proposals),而是直接生成检测类别的概率与位置坐标,往往有更快的检测速度,其中 YOLOv5 凭借其又快又准的检测效果,应用于许多场所,考虑到实时检测对于速度的要求,因此选择了以 YOLOv5 为基础的检测方式来实现对机场轮对裂纹的检测。

1 YOLOv5 算法介绍

YOLOv5^[4] 是一种较好的一阶段目标检测算法,其具有一个典型的深度学习网络结构,具体包括主干(Backbone)、颈部(Neck)和输出(Head)三部分。此外,YOLOv5 还使用了许多先进的训练

策略来提高其精度,如数据增强(data augmentation)、多尺度训练(multi-scale training)等。优秀的网络结构和训练策略保证了 YOLOv5 的准确性和实时性,使得 YOLOv5 算法可以较好地完成常见的检测任务。

2 裂纹检测算法优化

本文对传统 YOLOv5 算法的优化改良大体可分为:图像分割预处理与子图反向拼接,以及主干网络的改良。其中主干网络的改良包括:在网络的 Backbone 部分加入了通道注意力机制,使用 C3SE 模块代替了原来的 C3 模块;在网络的 Neck 部分以 C3STR 的形式增加了 Swin-Transformer 模块,使用 SPD-Conv 代替了原模型的传统卷积模块 Conv;在 Head 部分中将原来的 CIoU_Loss 改进为 SIoU_Loss。改进后的网络结构如图 1 所示。

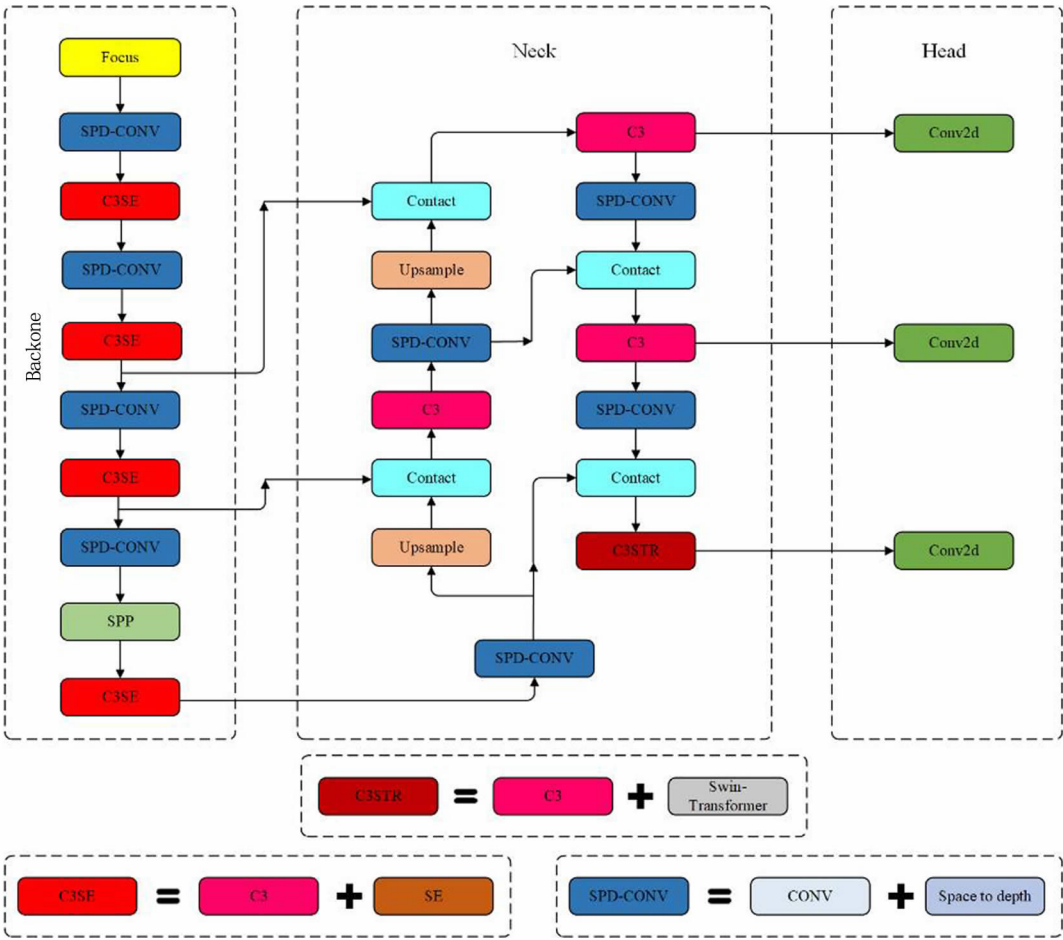


图 1 改进后的网络结构

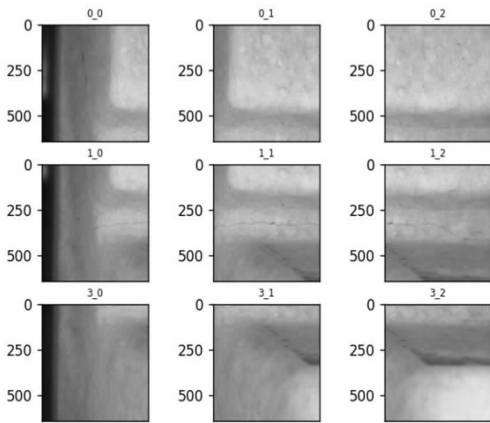
2.1 图像分割预处理与子图反向拼接

图像预处理采用了一种图像分割的处理方法,通过分割训练集以及训练集标签的形式,将一张大像素图片分割为多张等大子图。虽然在一定程度上增加了训练与检测的计算量,但大大降低了对于硬件设备的要求,同时提高了网络对于一些细小裂纹的识别能力。

在图像分割的设计中,采用了滑动窗口切图^[5]的方式。利用大小确定的滑窗对高分辨率的图片进行滑动切分,为了避免切口处存在裂纹,在设置滑窗大小的同时还规定了相邻子图之间的重叠率。分割预处理后的结果对比图如图 2 所示。



(a) 切割前原图(像素为 1400×1200)



(b) 切割后子图(每张像素为 640×640)

图 2 分割预处理前后对比图

检测过程中采用相同的处理思路,先对原图进行预处理分割,然后对分割后的子图进行裂纹识别,检测完成后再利用得到的裂纹坐标与子图重新反向拼接,完成一次裂纹的检测。

2.2 SPD-Conv 卷积模块

卷积神经网络(CNN)^[6]在计算机视觉领域取

得了巨大成就。但由于 CNN 体系结构中跨卷积和池化层的使用,使传统的神经网络对于细粒度信息的获取能力较差,无法很好地作用于小目标检测任务,因此采用了一种 SPD-Conv^[7]的卷积方式代替传统的卷积模块。SPD-Conv 具体由一个空间到深度(SPD)层和一个无卷积步长(Conv)层组成,它在设计上完全消除了步长和池化操作,使得应用了该模块后的网络对于细小的数据集往往有更好的检测效果。

2.3 改进的通道域注意力机制(SEnet)模块

SE^[8]注意力机制是一种通道注意力增强机制,其主要由压缩(squeeze),激活(excitation)以及重加权(reweight)三部分组成。通过对原特征图进行全局平均池化、ReLU 函数激活、逐通道重加权的方式提升模型对不同通道特征的敏感性,进而改善其捕捉密集目标的能力。网络中 C3SE 的结构即在原网络的 C3 模块中加入 SE 注意力机制。

2.4 Swin-Transformer 注意力机制的增加

Swin-Transformer^[9]是一种注意力机制,可以帮助模型在进行预测时将注意力集中在图像的最重要特征上。该机制采用逐层次化设计,共包含四个 stage,每个 stage 都会缩小输入特征图的分辨率,并逐层扩大感受野。对比于传统的 Transformer, Swin-Transformer 将注意力的计算限制在每个窗口内,进而减少了计算量。通过在原网络 Neck 部分的 C3 模块中增加 Swin-Transformer 机制(即 C3STR),提高 Neck 部分对于裂纹特征的提取能力,从而直接影响 Head 部分的检测效果。

2.5 SIoU 损失函数

SIoU^[10](scale-aware intersection over union)是一种用于提高目标检测精度的新度量,具体由角度损失函数、距离损失函数、形状损失函数和 IoU 损失函数组成。由于该损失函数重新定义了惩罚度量,考虑了期望与回归之间的向量夹角,使之可以有效提高模型对小目标的检测性能,有助于稳定训练过程,对于裂纹等小目标具有较好的检测效果。

3 实验结果及分析

3.1 实验配置

所有模型的训练和测试均在 Linux 系统下完成,选用 PyTorch1.7.1 框架。具体的实验环境配置如表 1 所示。

表 1 环境配置

配置名称	配置信息
操作系统	Linux
语言	python3.8
CPU	Xeon E5-2686 v4
GPU	NVIDIA GeForce RTX 3080Ti
GPU 加速	CUDA 11.0 cuDNN 8.0

3.2 模型评价指标

为了对改进后的模型进行性能评估对比,我们以准确率(Precision,P)、召回率(Recall,R)和平均精度均值(mAP)作为模型检测效果的评估指标。通常来说,P和R是相互制约的关系,而mAP值往往能体现模型的综合检测能力,因此在模型的检测效果评价中以mAP为主要参考依据。模型的复杂程度则与计算量(GFLOPs)和参数量(parameter,p)相关,通常GFLOPs与parameters越大,模型越复杂。

3.3 训练结果和分析

实验训练操作中首先对训练的图片进行分割处理以及Mosaic数据增强操作,然后再按照训练集:测试集:验证集=6:2:2的比例对分割后的数据集进行划分,最后以输入图像分辨率640×640,训练总轮数为400轮的参数进行训练。

3.3.1 使用 SIoU 后的实验分析

为了验证 SIoU 对于模型综合性能的影响,实验一对比了 YOLOv5s 模型在不同目标框损失函数下的表现,对比结果如表 2 所示。通过表 2 可知,在使用 SIoU_Loss 替换 CIoU_Loss 后,模型的召回率、准确率与综合检测能力均得到了提高。

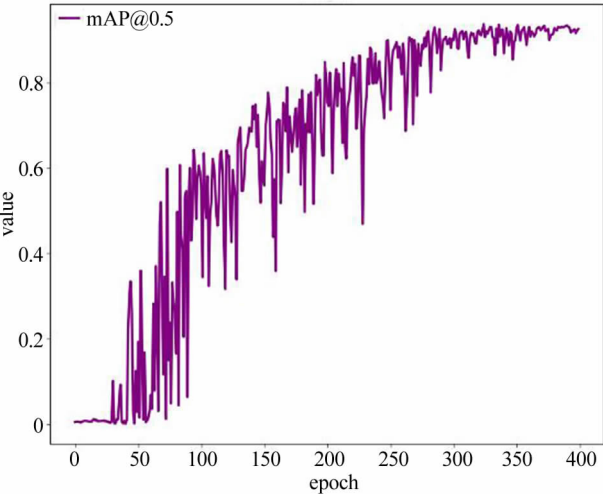
表 2 不同损失函数下的实验对比

模型	mAP@0.5	mAP@0.5 : 0.95	准确率 P	召回率 R
YOLOv5s_CIoU	91.07%	60.87%	88.18%	86.17%
YOLOv5s_SIoU	91.92%	61.19%	87.02%	87.23%

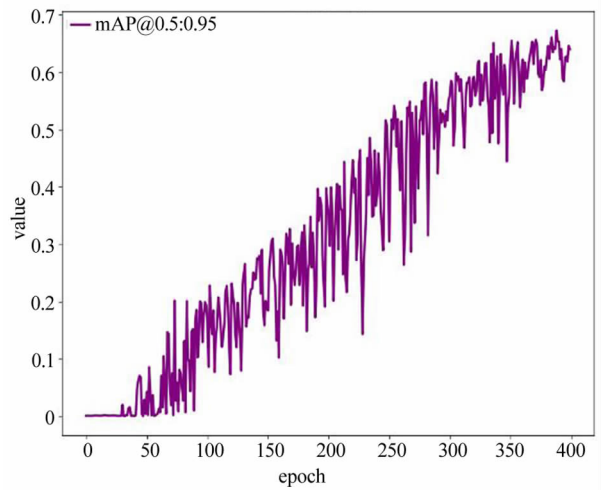
3.3.2 修改网络结构后的实验分析

在 YOLOv5s 算法的基础上对网络结构进行修改,修改后的网络结构如图 1(记作模型 A)。为了探究模型 A 的检测能力,在实验一的基础上进行实验二,具体操作为在 SIoU_Loss 的基础上运行修改后的模型 A,并在完成训练后对模型的训练结果进行可视化处理,其中 $mAP^{IoU=0.5}$ 以及

$mAP^{IoU=0.5 : 0.95}$ 的具体结果如图 3 所示。



(a) mAP@0.5 结果可视化



(b) mAP@0.5 : 0.95 结果可视化

图 3 mAP 结果可视化

从图 3 可知,随着训练轮数的增加,mAP@0.5与mAP@0.5 : 0.95 大体上均呈上升趋势,在到达 300 轮后模型基本收敛,此时的 mAP@0.5 稳定在 92.5%左右,训练过程中模型的最高 mAP@0.5值为 93.51%;mAP@0.5 : 0.95 最终保持在 65%附近。相比于传统的 YOLOv5s 网络,mAP 值的增大说明了修改后的模型 A 对密集、细小的裂纹数据集有着更好的识别能力。

模型 A 与 YOLOv5s、YOLOv5m 的结果对比如表 3 所示。通过表 3 可以发现模型 A 的参数量与计算量相较于 YOLOv5s 有所上升,但复杂度却远小于 YOLOv5m,且模型 A 的检测效果无论是对比轻量的 YOLOv5s 还是复杂的 YOLOv5m,都有着一定的提升。

表 3 实验二对比结果

模型	mAP@0.5	mAP@0.5 : 0.95	准确率 P	召回率 R	参数量 p	计算量 GFLOPs
YOLOv5s	91.07%	60.87%	88.18%	86.17%	7,063,542	16.5
YOLOv5m	91.59%	63.12%	88.81%	85.52%	21,056,406	50.6
模型 A	93.51%	65.10%	86.57%	91.25%	10,671,478	35.0

3.4 分割拼接处理的实验结果与分析

为了验证分割拼接处理对于检测效果的影响,设计了一组对照试验,在相同的检测权重和图片集

条件下,一组按照先分割,再检测,最后反向拼接还原的方式进行检测操作;另一组则是直接进行检测操作。两组实验的检测结果对比如图 4。

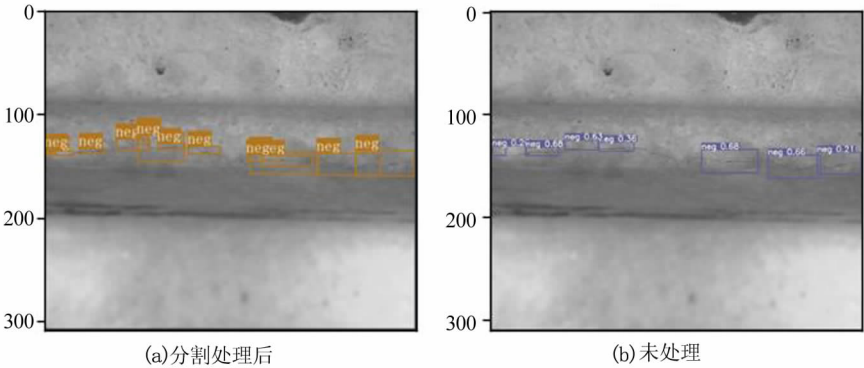


图 4 分割处理后与未处理检测结果对比

通过图 4 发现,经过分割拼接方法处理后的实验组可以检测出一些难以发现的细小的、不明显的裂纹,但由于存在拼接过程,不可避免地造成了检测框重叠的现象。

4 结 论

为了实现机场行李运输系统轮对裂纹的实时检测,探索了一种基于 YOLOv5 的裂纹检测算法,改进的主要方向为:修改了原网络的损失函数,在网络中增加了 SE 注意力机制和 Swin-Transformer 模块;使用 SPD-Conv 替代了传统的 Conv 模块;增加了图片分割与子图反向拼接的预处理。通过实验可知,修改后的模型与 YOLOv5s 相比,mAP@0.5 提升了 2.44%,召回率 P 提升了 5.08%。在检测实验中,利用图像分割与拼接处理,即使在相同的数据集与检测权重下,也能更好地识别出细小裂纹。

综上所述,修改后的模型基本能满足机场对于行李运输系统轮对裂纹的实时检测要求。

参考文献

[1] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic

segmentation[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition,2014: 580—587.
[2] GIRSHICK R. Fast r-cnn[C]//Proceedings of the IEEE International Conference on Computer Vision, 2015: 1440—1448.
[3] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: unified, real-time object detection[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016: 779—788.
[4] 邵延华,张铎,楚红雨,等. 基于深度学习的 YOLO 目标检测综述[J]. 电子与信息学报,2022,44(10):3697—3708.
[5] CHEN Xinlei, GIRSHICK R, HE K, et al. Tensormask: a foundation for dense object segmentation[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision, 2019: 2061—2069.
[6] KRIZHEVSKY A, SUTSKEVER I, HINTON G. Imagenet classification with deep convolutional neural networks[J]. Communications of the ACM, 2017, 60(6): 84—90.
[7] SUNKARA R, LUO T. No more strided convolutions or pooling: a new CNN building block for low-resolution images and small objects[C]//Joint European Conference on Machine Learning and Knowledge Discovery in Databases. Cham: Springer Nature Switzerland, 2022: 443—459.
[8] HU Jie, LI Shen, SUN Guang. Squeeze-and-excitation networks[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition,2018: 7132—7141.
[9] LIU Ze, LIN Yutong, CAO Yue, et al. Swin transformer: hierarchical vision transformer using shifted windows[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision,2021: 10012—10022.
[10] GEVORGYAN Z. SiOU loss: more powerful learning for bounding box regression[J]. arXiv preprint arXiv: 2205.12740, 2022.