

改进 YOLOv5s 的钢材表面缺陷检测

谷长江, 高法钦

(浙江理工大学计算机科学与技术学院, 浙江 杭州 310018)

✉ siesuta@outlook.com; gaofaqin@zstu.edu.cn



摘要: 缺陷检测是生产中重要的环节, 基于钢板表面缺陷特征不明显和难以提取导致的检测精度不足问题, 文章在 YOLOv5s 检测网络的基础上进行改进, 首先基于 DO-Conv 过参数化模块改进网络特征提取模块, 然后使用 ULSAM 注意力机制改进网络的颈部(Neck), 提出改进的 YOLOv5s 缺陷检测网络。基于 NEU-DET 数据集的实验结果表明, 改进的 YOLOv5s 缺陷检测网络检测平均准确率达 76.6%, 较 YOLOv5s 和 YOLOv4 分别提升了 7.8% 和 6.3%, 有效提高了钢材表面缺陷检测精度。

关键词: YOLOv5s; 缺陷检测; 注意力机制; 过参数化

中图分类号: TP399 **文献标志码:** A

Steel Surface Defect Detection Based on Improved YOLOv5s

GU Changjiang, GAO Faqin

(Department of Computer Science and Technology, Zhejiang Sci-tech University, Hangzhou 310018, China)

✉ siesuta@outlook.com; gaofaqin@zstu.edu.cn

Abstract: As defect detection is a crucial part of production, this paper proposes an improved YOLOv5s detection network to solve the problem of insufficient detection accuracy caused by unclear and difficult extraction of surface defect features on steel plates. Firstly, the network feature extraction module is improved based on the DO-Conv over-parameterized module, and then the neck of the network is improved by using the ULSAM attention mechanism. Finally, an improved YOLOv5s defect detection network is proposed. The experimental results based on the NEU-DET dataset show that the improved YOLOv5s defect detection network has mean average precision of 76.6%, which is 7.8% and 6.3% higher than YOLOv5s and YOLOv4, respectively. The improved network effectively improves the accuracy of steel surface defect detection.

Key words: YOLOv5s; defect detection; attention mechanisms; over-parameterization

0 引言(Introduction)

钢材作为国民生产中重要的材料, 在其生产、加工及运输过程中不可避免地有缺陷产生, 其中尤为严重的就是生产过程中产生的缺陷, 这类缺陷的影响较大, 常常造成残次品或无法满足实际使用需要等。传统的目标检测主要是采用人工方式进行检测, 这种方法受主观因素影响较大, 因此研究基于深度学习的钢材表面缺陷检测算法具有重大意义。

随着深度学习在图像领域的不断应用, 缺陷检测也逐渐成为其中一个应用分支^[1]。传统的机器图像如无损检测、微波检测方法的应用存在对材料有限制等问题而无法广泛应用^[2-3]。基于图像深度学习的缺陷检测技术没有目标材料限制要求且不需要直接接触被检测物体而逐渐被应用于更多领域的材料缺陷检测任务中。

研究者在钢材表面缺陷检测领域已进行了大量研究。

2017年,邢健夫^[4]使用 AlexNet 构建了一套对钢材缺陷进行分类的网络,并使用数据增强技术对数据集进行扩充,最终提高了缺陷分类的准确性。2020年,徐镒等^[5]在 YOLOv3 网络的基础上进行改进,采用 MobileNet 模块提升了带钢表面缺陷的检测速度。

然而,对于钢材表面缺陷检测任务,上述研究仍存在很多问题,例如对小目标不敏感、缺陷检测的精度依旧有待提高等。针对小目标检测问题,本文通过使用 ULSAM 注意力机制,加强了上下文特征信息的融合和小目标的特征提取能力。针对缺陷检测精度不高的现象,本文提出基于 DO-Conv 的特征提取模块,利用过参数化带来的强大特征提取能力提升模型的性能^[6]。并通过实验证明,本文所提方法有效提升了钢材表面缺陷检测精度。

1 YOLOv5s 算法及其改进 (YOLOv5s algorithm and its improvements)

1.1 YOLOv5s 网络结构

YOLOv5 是继 YOLOv4 之后提出的 YOLO 系列的又一名新成员,它的特征提取网络是基于跨阶段的局部网络架构的暗网(CSPDarknet)^[7]。CSPDarknet 拥有较高的计算效率,同时 YOLOv5 系列为了面向不同的检测任务,提出了如 YOLOv5-t、YOLOv5-s、YOLOv5-m、YOLOv5-l 等多个不同版本,这使得 YOLOv5 系列比 YOLOv4 系列拥有更广的适用领域^[8]。其中应用最多的就是 YOLOv5s 网络,它在网络复杂性与网络精度方面取得了良好的协调。因此,本文针对 YOLOv5s 网络进行研究。

图 1 是 YOLOv5s 网络的结构图。YOLOv5s 网络分为 Backbone 特征提取网络、neck 上下文特征融合网络及 Head 预测网络三个部分。其中,SPPF 结构使用集连方法构建模块,这在很大程度上扩大了模块的感受野。

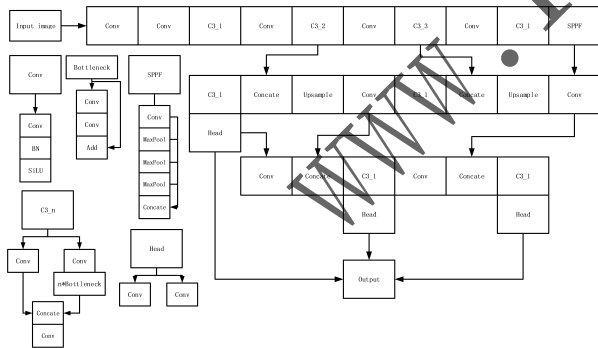


图 1 YOLOv5s 结构图

Fig. 1 YOLOv5s structure diagram

1.2 YOLOv5s 网络检测原理

YOLOv5s 是一种基于深度学习的目标检测模型,它使用卷积神经网络实现对图像中物体的检测。YOLOv5s 采用了单阶段目标检测的方法,该方法可以在单次推理中完成目标检测、定位和分类任务。

在 YOLOv5s 中,网络使用多个卷积层提取图像特征。Neck 部分完整对 Backbone 部分提取到的特征进行进一步融合,由此得到更丰富的目标特征,以便实现对目标的预测。Neck 网络部分采用的是特征金字塔+路径聚合(FPN+PAN)

结构。LIN 等^[9]提出著名的特征金字塔网络 (Feature Pyramid Networks, FPN),目的是对网络内部不同深度的特征信息进行融合。Mosaic 数据增强技术是在 YOLOv4 网络的基础上提出的,经实验证明, Mosaic 能有效缓解特征不明显的现象,因此 YOLOv5s 中也使用了这种方法^[10]。

首先,网络使用一组卷积层提取图像特征。其次,网络使用类似于 FPN 的结构生成多个分辨率的特征图。这些特征图可以用来预测目标的位置和类别。最后, YOLOv5s 使用了一种称为非极大值抑制 (Non-Maximum Suppression) 的方法移除重叠的检测框,并输出最终的检测结果。

YOLOv5s 的损失函数如公式 (1):

$$L_{\text{loss}} = \alpha \cdot L_{\text{obj}} + \beta \cdot L_{\text{bbox}} + \gamma \cdot L_{\text{conf}} \quad (1)$$

其中, α, β, γ 是协调三个损失的超参数, L_{conf} 表示该检测点是否存在目标的置信度损失, L_{obj} 表示目标分类的损失, L_{bbox} 表示目标真实框与预测框之间偏离导致的损失。

1.3 YOLOv5s 的改进

1.3.1 基于 DO-Conv 改进 YOLOv5s

DO-Conv 是一种过参数化 (Over-Parameter) 方法^[11],它的思想是通过将一次卷积分离成两个可学习的训练步骤,最终在推理时将这两个步骤结合成一个卷积操作,它包含一个深度卷积操作和一个普通的矩阵乘法即线形层。其中,深度卷积公式如下:

$$O_{DC_{in}} = \sum_i^{M \times N} W_i DC_{in} \quad (2)$$

通过一次深度卷积可以让原有尺度为 $C_{in} \times (M \times N)$ 的输入特征转换为尺度为 $C_{in} \times D$ 的输入特征,对于 $h \times w$ 大小的特征图而言, D 指的是 $h \times w$ 个空间点位, $M \times N$ 指的是滑动窗口时卷积核 W 的个数, C_{in} 表示输入特征的尺度为 C 。

DO-Conv 的具体流程图如图 2 所示,其中 W 表示卷积权重, I 表示输入特征, P 表示卷积参数, O 表示输出特征, $\circ$ 为公式 (2) 表示的运算操作, $*$ 表示普通卷积操作。图 2 是 DO-Conv 推理时的流程,可以看出,输入特征相当于进行了两次卷积,根据 DO-Conv 研究推理得出,图 2 中的计算流程等价于图 3 所示的流程。

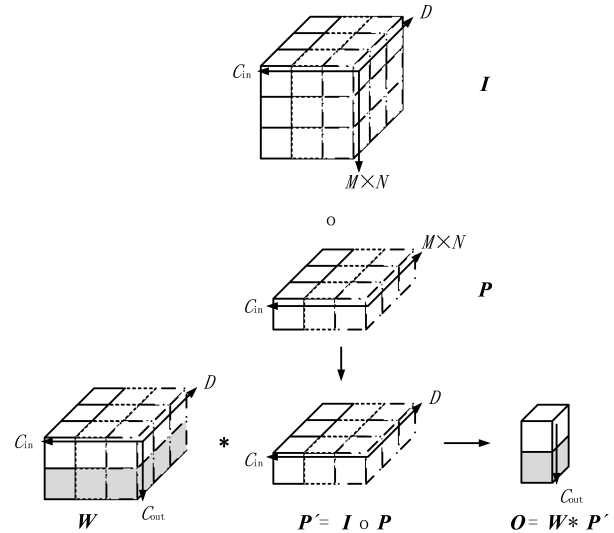


图 2 DO-Conv 训练过程

Fig. 2 DO-Conv training process

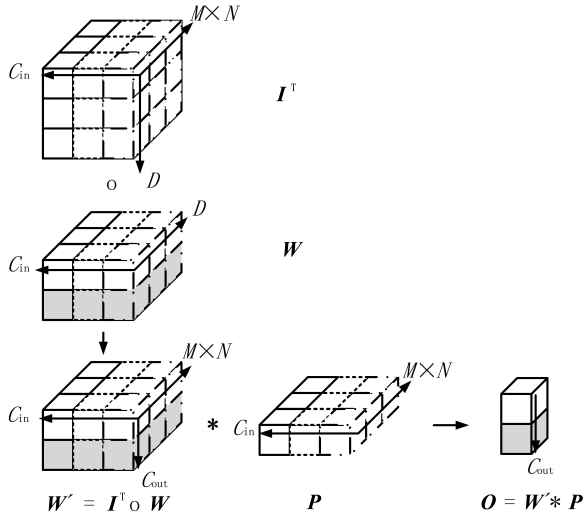


图3 DO-Conv 推理过程

Fig. 3 DO-Conv inference process

从图3中可以看出,DO-Conv 采用了类似 Inception 的思想,但不同的是 DO-Conv 采用深度方向的堆叠方法。

1.3.2 ULSAM

ULSAM(超轻量级子空间注意力模块),是一种用于紧凑型 CNN 的新型注意力模块。ULSAM 可以为每个特征图子空间学习不同的注意力图,减少紧凑型 CNN 的特征图中的空间和通道冗余。同时,为每个子空间学习不同的注意力图,可以实现多尺度的特征表示,特别是对于细粒度的图像分类任务。ULSAM 是一个能够高效地学习特征图子空间的模块,它是基于跨通道相互依赖性的注意力模块构建的。

ULSAM 计算流程图如图4所示,DW 表示分组卷积,PW 表示点积, 1×1 表示卷积核大小,Softmax 是激活函数。

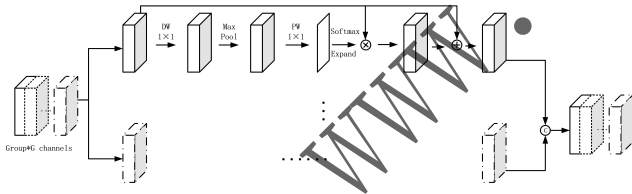


图4 ULSAM 计算流程图

Fig. 4 ULSAM calculation flowchart

为了增强网络对特征的识别能力,本文在将特征送入预测网络之前,使其通过一个 ULSAM 模块,以更好地提取特征。

改进后的网络结构图如图5所示。

2 实验与分析(Experiment and analysis)

2.1 实验环境

为了验证本研究改动对小目标缺陷件检测的有效性和先进性,在 CPU 为 RX3600, GPU 为 Ge Force RTX3060,开发语言为 Python,深度学习框架为 pytorch 的环境下进行实验。

2.2 数据集简介

数据集来源于东北大学的 NEU-DET 钢材表面缺陷数据集^[12]。图像类别包括裂纹(CR, Cracking)、夹杂(IN, Inclusion)、

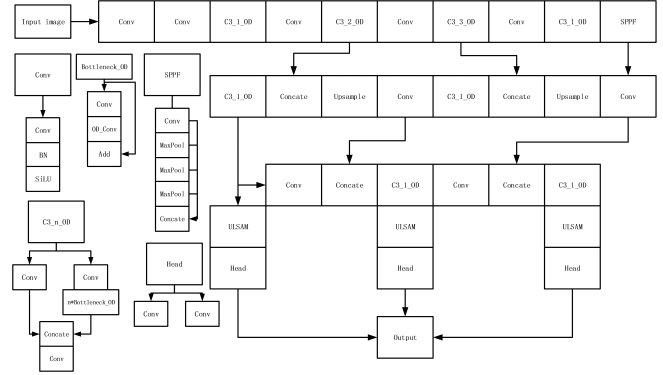


图5 改进 YOLOv5s 结构图

Fig. 5 Improved structure diagram of YOLOv5s

斑块(PA, Patches)、麻点(PS, Pitted_Surface)、压入氧化铁皮(RS, Rolled-in_Scale)及划痕(SC, Scratches)。每种类型的缺陷图片包含300幅尺寸为 200×200 的图像,共计1800张。使用xml格式定位缺陷位置并对其进行分类。数据集按7:2:1的比例随机分为训练集、验证集和测试集。

2.3 模型训练

本文提出的改进 YOLOv5s 算法的初始设置情况如下:初始学习率为0.01,训练300个epoch,若训练结束验证集损失未稳定,则相应增加epoch值,如果训练长时间不收敛,超过50个epoch,则提前结束。模型训练目标的类别置信度阈值是0.5,DIoU 阈值是0.45。当两个预测框之间的DIoU值超过了设置阈值时,则删除该预测框。将输入图像进行分类处理后,输出结果作为最终结果。

评价指标主要分为模型指标和精度指标两大类。模型指标为模型的参数量大小Params,模型推理时的浮点数GFLOPs。精度指标采用精确率(Precision, P)、召回率(Recall, R)和平均准确率(Mean Average Precision, MAP)三项作为网络模型的性能指标。P、R、MAP 和 AP 公式如下:

$$P = \frac{TP}{TP + FP} \quad (3)$$

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

$$MAP = \frac{\sum_{i=0}^n AP(i)}{n} \quad (5)$$

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

其中,TP 为正确预测的正例;FP 为错误预测的负例;FN 为错误预测的正例;n 为类别数;AP 为每种检测类型的准确性。

2.4 结果与分析

为了验证本文方法的优越性,首先将其与主流算法进行了性能比较,然后通过消融实验验证了改进 YOLOv5s 算法的有效性。

2.4.1 算法的对比实验

改进 YOLOv5s 与 YOLOv3、YOLOv4 和 YOLOv5s 三种常用算法的评价指标见表1^[13-16]。下述算法的GFLOPs 是在输入图像大小为 640×640 时获得的。

表1 算法结果对比表

Tab.1 Comparison table of algorithm results

算法模型	P/%	R/%	MAP/%	Params/M	GFLOPs/G
YOLOv3	56.6	62.4	59.9	61.55	155.2
YOLOv4	60.4	72.4	70.3	63.96	141.52
YOLOv5s	67.5	65.8	68.8	7.036	15.9
改进 YOLOv5s	76.2	70.3	76.6	7.041	15.9

上述改进模型的 *Params* 和 *GFLOPs* 是在推理时获得的。所有模型均是在初始学习率为 0.01、*IoU* 阈值为 0.5 的条件下计算,训练结束时模型损失均已收敛。由表 1 中的结果可知:本文提出的改进 YOLOv5s 算法以 76.6% 的平均准确率、76.2% 的精确率、70.3% 的召回率排第一,其综合性能优于其他几种算法。

2.4.2 消融实验

本文设立 3 组不同条件的实验,未改动的 YOLOv5s、仅添加 DO-Conv 的 YOLOv5s,以及添加了 ULSAM 注意力机制与 DO-Conv 的 YOLOv5s(本文改进算法),通过比较 3 组实验的 MAP 验证本文所提算法的有效性。这 3 组实验的 MAP 见表 2。

表2 消融实验结果表

Tab.2 Ablation experiment results table

算法	P/%	R/%	MAP/%
YOLOv5s	67.5	65.8	68.8
YOLOv5s+DO-Conv	69.6	68.0	71.0
YOLOv5s+DO-Conv+ULSAM	76.2	70.3	76.6

由表 2 可知,改动后算法的检测精度得到了大幅提升,通过使用注意力机制融合特征以及通过 DO-Conv 增强特征提取能力,可以有效提升模型的精度。

3 结论(Conclusion)

本文提出的基于 YOLOv5s 的钢板缺陷检测方法通过 DO-Conv 增强特征提取能力,引入 ULSAM 模块提升了算法对 NEU-DET 钢板缺陷识别的精度。针对钢材表面缺陷检测问题,本文对 YOLOv5s 算法进行改进:引入 DO-Conv 模块,进一步增强模型对特征的提取能力;引入 ULSAM 注意力机制,增强特征融合能力,更好地提取目标特征。实验结果验证了本文提出的新算法的有效性,研究成果为钢材表面缺陷检测提供了一种算法基础。接下来,将在提升模型检测速度及检测精度方面做进一步研究。

参考文献(References)

[1] 袁洪强,杜国锋,余泽禹,等. 基于轻量神经网络的钢材表面缺陷快速识别[J]. 科学技术与工程,2021,21(34): 14651-14656.

[2] 李翰超,古阳,于亚婷. 基于涡流无损检测技术的金属表面缺陷三维模拟分析[J]. 电子科技大学学报,2020,49(6): 942-948.

[3] GAO B,ZHANG H,WOO W L,et al. Smooth nonnegative matrix factorization for defect detection using micro-

wave nondestructive testing and evaluation[J]. IEEE Transactions on Instrumentation and Measurement,2014, 63(4):923-934.

[4] 邢健夫. 基于卷积神经网络的热轧带钢表面缺陷识别及系统开发[D]. 沈阳:东北大学,2017.

[5] 徐镔,朱洪锦,范洪辉,等. 改进的 YOLOv3 网络在钢板表面缺陷检测研究[J]. 计算机工程与应用,2020,56(16): 265-272.

[6] YE H,ZHU X D,LIU C Y,et al. Furniture image classification based on depthwise group over-parameterized convolution[J]. Electronics,2022,11(23):3889.

[7] WANG X,GAO J S,HOU B J,et al. A lightweight modified YOLOX network using coordinate attention mechanism for PCB surface defect detection[J]. IEEE Sensors Journal,2022,22(21):20910-20920.

[8] 杨畅畅,李慧玲,潘广通. 基于改进 YOLOv4 算法的铝型材表面缺陷检测[J]. 组合机床与自动化加工技术,2022(11): 66-69.

[9] LIN T Y,DOLLÁR P,GIRSHICK R,et al. Feature pyramid networks for object detection[C]//IEEE. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. New York:IEEE,2017:936-944.

[10] 顾学坤,楚建安. 基于 YOLOv5 改进算法的印花图案疵点检测[J]. 电子测量技术,2022,45(4):59-65.

[11] PANDEY E,KUMAR S. Exploring the generalization capacity of over-parameterized networks[J]. International Journal of Intelligent Systems and Applications in Engineering,2023,11(1s):97-112.

[12] HE Y,SONG K C,MENG Q G,et al. An end-to-end steel surface defect detection approach via fusing multiple hierarchical features[J]. IEEE Transactions on Instrumentation and Measurement,2020,69(4):1493-1504.

[13] 程婧怡,段先华,朱伟. 改进 YOLOv3 的金属表面缺陷检测研究[J]. 计算机工程与应用,2021,57(19):252-258.

[14] 方叶祥,甘平,陈俐. 金属表面缺陷检测的改进 YOLOv3 算法研究[J]. 机械科学与技术,2020,39(9):1390-1394.

[15] 庾彩红,黄海松,曾锋,等. 基于 YOLO_v4 的空心杯电枢表面缺陷实时检测[J]. 组合机床与自动化加工技术, 2021(6):59-62,66.

[16] CHEN S H,LAI Y W,KUO C L,et al. A surface defect detection system for golden diamond pineapple based on CycleGAN and YOLOv4[J]. Journal of King Saud University-Computer and Information Sciences,2022,34(10):8041-8053.

作者简介:

谷长江(1997-),男,硕士生。研究领域:目标检测,缺陷检测,图像处理。本文通信作者。

高法钦(1974-),男,博士,副教授。研究领域:机器学习与智能感知,机器视觉定位与产品外观检测重构研究。